Molecular and Machine Learning Based Characterization of Human Skeletal Muscle to Decipher Complex Biological Processes Governing Muscle Wasting in Surgical Patients with Cancer

by

Bhumi Jagdish Bhatt

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

Department of Laboratory Medicine and Pathology
University of Alberta

Abstract

Cancer cachexia is a multifactorial syndrome characterized by progressive loss of weight (WL), muscle, and fat tissues. Skeletal muscle wasting, in particular, is strongly associated with morbidity and mortality. Understanding the pathophysiological mechanisms underlying muscle wasting in humans is urgently needed. While clinical variables, e.g., Body Mass Index (BMI), WL, and Skeletal Muscle Index (SMI), are helpful in prognostication, these variables provide no insight into the underlying molecular mechanisms of cachexia. Molecular mechanisms are described in animal models of cachexia with unknown translational relevance; clinical studies of human skeletal muscle biopsies are sparse and did not yield consistent findings. Patient classification heterogeneity, limited sample size, aggregate cancer types, and sex, and focus only on coding transcripts also limited the value of molecular findings in humans.

An emerging consensus is that non-coding RNAs are key players in gene regulation and may play a role in muscle homeostasis. Therefore, comprehensive human studies are needed to address these gaps. My thesis addresses these gaps, and all subsequent work describes the findings in a sex-specific context.

I hypothesize that expression profiling of coding and non-coding RNAs from skeletal muscle of patients with cancer, when subjected to unsupervised machine learning clustering approaches, would facilitate the identification of the underlying molecular architecture (subtypes) of skeletal muscle.

Rectus abdominis skeletal muscle biopsies (n=84; males, n=48; and females, n=36) were obtained from a surgical cohort of patients with cancer. Clinical data included age, sex, cancer diagnosis, WL, BMI, plasma C-reactive protein, and SMI determined by computed tomography.

mRNAs, long non-coding RNAs/ lncRNAs, and small RNAs (miRNA, piRNA, snoRNA, tRNA) were profiled using high-throughput Next Generation Sequencing. Transcriptome profiles were subjected to unsupervised clustering using the integrative Non-negative Matrix Factorization (intNMF) algorithm to address the clinical classification heterogeneity. K=2 clusters (or subtypes) were identified in patients of the male and female sex. Differential Expression (DE) analysis of the subtypes identified dysregulated RNAs. However, the intricate interplay through which RNAs co-regulate each other via post-transcriptional competing endogenous RNA (ceRNA) mechanism is not studied in clinical cachexia. ceRNAs communicate through common miRNA binding sites or response elements (MREs). I identified lncRNAs and mRNAs acting as ceRNAs. Amongst the top six identified hub lncRNAs, three were common.

Although 2 muscle subtypes were determined, there is no standard method to discern the intrinsic characteristics of subtypes. I adopted two independent benchmark paradigms and referred to them as (i) clinical and (ii) molecular and functional benchmarking.

I performed clinical benchmarking based on WL grading (severity Grade 0-4) and age- and sex-specific SMI z-score. These analyses revealed that Subtype 1 in both male and female patients was proportionally associated with high-grade WL and low SMI z-score and regarded as a cachexia group; Subtype 2 had minimal WL and high SMI.

There are no available datasets in the literature to guide functional benchmarking. I employed two experimental model systems to approach molecular benchmarking. mRNA profiles were generated from (i) human Rhabdomyosarcoma cells, a model of myoblast proliferation and differentiation. and (ii) from gastrocnemius muscle of rats bearing a colon adenocarcinoma. Comparison of pathways from DE mRNAs in these models with that in human Subtype 1 vs. Subtype 2 showed

overlap. These findings led me to propose a schematic model explaining the disruption of homeostatic response from component pathways necessary for skeletal muscle integrity and function. These include Extracellular Matrix regulation, nNOS signaling, Matrix Metalloproteases, calcium signaling, and transcriptional regulatory network in embryonic stem cells.

Overall, my work has identified two de novo clinically applicable molecular subtypes. This is the first large-scale investigational study that provides a transcriptional landscape of human skeletal muscle from patients with cancer along with multilayered RNA crosstalk. The study results in a paradigm shift from applying heterogeneous patient classification criteria to the pragmatic utility of unsupervised machine learning algorithms for pursuing molecular studies in the cachexia domain. Subtype 1 is demonstrated to be a group affected by cachexia by WL and muscle depletion.

Data curated for the entire transcriptional atlas of mRNAs, lncRNAs, and small ncRNAs will be accessible publicly to enable cachexia researchers to examine their genes or pathways of interest.

Preface

This doctoral thesis is an original work by Bhumi Bhatt. Conjoint Health Research Ethics Board approved the research work conducted as part of this thesis at the University of Calgary (Ethics ID: E17213) for the accrual of skeletal muscle biopsies. Whole transcriptome profiling of skeletal muscle biopsy from patients with cancer and access to clinical information was approved by the Health Research Ethics Board of Alberta (HREBA) – Cancer Committee under the ethics protocol ETH 17-0517 (valid until 31st March and renewed yearly) ETH-21709 with Dr. Sambasivarao Damaraju as principal investigator and Dr. Vickie Baracos as co-investigator. The animal work presented in Chapter 5 was approved by the University of Alberta Institutional Animal Care Committee (protocol ID RES006429), with Dr. Vera Mazurak as the principal investigator.

Bhumi Bhatt conceptualized the study designs of Chapters 2, 3 with Dr. Baracos and Dr. Damaraju, and Chapter 4 with Dr. Damaraju. Dr. Oliver Bathe provided human skeletal muscle biopsy specimens. Bhumi Bhatt and Dr. Baracos collected and assembled clinical and demographic data. Bhumi Bhatt performed Next Generation Sequencing and machine learning data analyses. Bhumi Bhatt performed statistical analyses under the guidance of Dr. Sunita Ghosh.

Bhumi Bhatt conceptualized the study design of Chapter 5 with Dr. Baracos and Dr. Damaraju. The animal model work represents a collaboration between Dr. Damaraju and Dr. Mazurak's labs at the University of Alberta. Dr. Mazurak developed the tumor-bearing rat model and extracted gastrocnemius muscle for mRNA profiling experiments. Dr. Damaraju's lab generated the next-generation sequencing profiles for rat muscle. Bhumi Bhatt performed bioinformatic analyses for animal and cell-line models.

Dedication

God,

my ever-loving parents,

and

my gurus, Dr. Sambasivarao Damaraju and Dr. Vickie Baracos

A special thanks to my parents for everything.

Words cannot describe how blessed, grateful and fortunate I am to be your daughter.

Acknowledgments

First and foremost, I would like to express my deepest gratitude to my supervisor Dr. Sambasivarao Damaraju. Words would never be enough to describe the mentorship, guidance, motivation, relentless work ethic, valuable life lessons, and immense knowledge you have shared with me over the years. Thank you for teaching me the art, science, and philosophy of scientific research. You are indeed the *father of my scientific journey*. Thank you for your omics expertise and for encouraging me to undertake statistical, programming, and machine-learning courses that have been instrumental in building my career profile. Your ever-enthusiastic and anytime approach to students is exemplary. Thank you for developing inquisitive out-of-the-box thinking, enriching my translational abilities, and making me a confident researcher. Thank you very much for sharing this truly memorable and quintessential journey with me and for my being my *guru in all endeavors*.

I am deeply indebted and would like to express my gratitude to my co-supervisor, Dr. Vickie Baracos. Thank you, mama Vickie for being my cachexia guru. I am incredibly grateful and fortunate to have worked with you. Every conversation with you has helped me learn something new about cachexia and life philosophies. Your three words – upwards and onwards- inspire me to be confident and perform my tasks efficiently. Thank you for teaching valuable life lessons and for your scientific insights and constructive feedback.

I thank my committee members, Dr. Sunita Ghosh and Dr. Vera Mazurak. Thank you, Dr. Ghosh, for being my statistics guru, providing excellent and valuable statistical inputs for this project, and for your unconditional support, care, and encouragement. Thank you, Dr. Mazurak, for the collaboration and for sharing your muscle biology expertise. I would like to thank Dr. Oliver Bathe for his Surgical Oncology expertise and for providing valuable biopsy specimens for this project.

I sincerely thank Dr. Simon Wing and Dr. Paul Stothard for serving on my examining committee and Dr. Jelena Holovati for chairing my thesis defense examination.

I thank Dr. Monika Keelan for her help and support throughout my graduate program. Thank you for all the great conversations and for being approachable.

I am thankful to Ms. Cherly Titus, Mr. Ronald Marinas, and Ms. Femka Williams for their tremendous administrative support. I would like to thank Ms. Jennifer Dufour and Ms. Abha Dunichand-Hoedl for their technical assistance and for being friendly colleagues.

I want to thank all my friends and colleagues on the CCI's fourth floor of experimental oncology. I had the pleasure of working with you all. Thank you, Dr. Mahalakshmi Kumaran, for being a supportive and friendly lab colleague. Thank you, Dr. Micheal Weinfeld, for the collaborative opportunity, and Ms. Yan Yang, for the incredible and unconditional friendship and support. Thank you, Mr. Mesfin Fanta and Mr.Wisdom Kate, for being great friends. Without your support and friendship, this journey would not be as fun and memorable. Thank you, Kaushiki, Sargun, Faisal, Amir, April, and Ferri. I would like to thank my cachexia research family. Thank you, Dr. Ashok Narasimhan, Dr. Lisa Martin, Dr. Ana Anoveros, Dr. Cathy Kubarak, Fateema, Peter, Monir, and Magaly.

I would like to thank Alberta Cancer Research Biobank for housing the muscle biopsies used in this study. I am grateful to all study participants for providing the clinical specimens. I would like to thank the Canadian Institutes of Health Research for funding this research (to SD and VB). I would like to express my gratitude to the Ph.D. program at the FoMD. The graduate program has enlightened me in several aspects of my scientific career and personal growth and development. I am thankful to Grey Nuns Community Hospital's volunteer services. Thank you,

Teresa and Shelly; I feel grateful for the volunteering opportunity. It has made me connect and add value to my research.

I am forever grateful and blessed to my family for all their unconditional love and support. Thank you to my brother Kalpesh Bhatt. Thank you for inspiring me to pursue research. Thank you to the people I love the most– my parents - Ila Bhatt and Jagdish Bhatt. Thank you, mom and dad, for all the love, care, sacrifices, values, and motivation; everything is dedicated to you both. Words will never be enough to express my gratitude and love. You are indeed my role model and an inspiration in whatever I pursue. You have taught me the value of education, to follow my goals and ambitions and achieve them. Thank you for teaching me all the valuable life lessons through various phases of life. In all, you two are my biggest blessings.

Thank you, God, for being at the forefront of everything I do and helping and guiding me at every step of life.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AGO	Argonaute
BMI	Body Mass Index
CBP	Complex Biological Processes
CC	Cachexia Cachexia
CCI	Charlson Comorbidity Index
CDS	Coding Domain Sequence
ceRNA	Competing Endogenous RNA
CPI	Cluster Prediction Index
CRP	C-Reactive Protein
CSA	Cross-Sectional Area
CT	Computed Tomography
DE	Differential Expression
DGC	Dystrophin Glycoprotein Complex
DGCR8	DiGeorge Syndrome Critical Region 8
ECM	Extracellular Matrix
FC	Fold change
FDA	Food and Drug Administration
FMI	Fat Mass Index
HU	Hounsfield Unit
IFN	Interferon
IL	Interleukin
IMAT	Intramuscular Adipose Tissue

intNMF	Integrative Non-negative Matrix Factorization
IPA	Ingenuity Pathway Analysis
KL	Kullback-Leilber
lncRNA	Long non-coding RNA
miRNA	microRNA
ML	Machine Learning
MMP	Matrix Metalloprotease
MRE	miRNA Response Element
MRF	Myogenic regulatory factor
mRNA	Messenger RNA
MYH	Myosin heavy chain
MYOD1	Myogenic differentiation 1
ncRNA	Non-coding RNA
NEB	New England BioLabs
NGS	Next Generation Sequencing
NMF	Non-negative Matrix Factorization
NMJ	Neuromuscular junction
NOS	Nitric oxide synthase
OS	Overall Survival
PAX	Paired Box
PDAC	Pancreatic ductal adenocarcinoma
piRNA	Piwi-interacting RNA
PIWI	P-element induced wimpy testis in drosophila

qRT-PCR	quantitative reverse transcription polymerase chain reaction
RBP	RNA binding protein
RD	Rhabdomyosarcoma
RI	Rand Index
RIN	RNA integrity number
RISC	RNA-induced Silencing Complex
RNA	Ribo-Nucleic Acid
SAT	Subcutaneous Adipose Tissue
SC	Satellite cell
SMI	Skeletal Muscle Index
snoRNA	Small nucleolar RNA
SW	Silhouette Width
TAT	Total Adipose Tissue
TIMP	Tissue inhibitors of MMP
TNF	Tumor necrosis factor
TPA	12-O-tetradecanolyphorbol-13-acetate
tRNA	Transfer RNA
UPP	Ubiquitin proteasome pathway
UTR	Untranslated Region
VAT	Visceral Adipose Tissue
VST	Variance Stabilized Transformation
WL	Weight Loss

Chapter 1 Introduction

1.1 Cancer-associated muscle wasting: Criteria for the classification of human skeletal muscle and their utility for molecular association studies

Cancer Cachexia (CC), as defined in 2011 by a DELPHI consensus, is "*a multifactorial syndrome characterized by an ongoing loss of skeletal muscle mass (with or without loss of fat mass) that cannot be fully reversed by conventional nutritional support and leads to progressive functional impairment*" [1]. Patients are classified as cachexic if there is weight loss greater than 5% or ongoing weight loss of greater than 2% in individuals already showing depletion based on Body Mass Index (BMI) < 20 kg/m^2 or skeletal muscle mass[1] (sarcopenia, defined using literature reported cut-offs). Cachexia occurs in about 50-80% of patients with advanced cancer, contributing to mortality in about 20% of patients [2-4]. There are currently no Food and Drug Administration (FDA) approved agents for the treatment of CC, and it is thus an unmet medical need and represents a significant clinical burden. Since the provisional diagnostic criteria for CC were proposed in 2011, classification based on weight loss cut-offs was refined using a large contemporary population cohort applying the criteria of percent Weight Loss (% WL) and BMI grading system[5] (Grades 0,1,2,3, and 4 representing graded severity). The grading system has been used in independent cohorts of patients with cancer for their association with clinical outcomes [6,7].

Cancer-associated muscle wasting is the cardinal feature of cachexia and is distinct from age-related sarcopenia, which relies primarily on muscle function [8,9]. Sarcopenia is characterized by reduced muscle strength, physical inactivity, and physiological changes associated with aging or occurs secondary to the disease. Conversely, CC is characterized by weight, muscle, and fat loss caused by manifestations of host-tumor interactions.

Skeletal muscle is the most extensive tissue in the body, accounting for 40-60% of body mass, and is associated with regular bodily functions such as movement, maintenance of posture, breathing, and thermogenesis. Our group pioneered the opportunistic use of Computed Tomography (CT) images obtained during routine patient care to quantify body composition in patients with cancer [10]. Image analysis at the third lumbar vertebrae (L3) is considered the de facto standard landmark correlating with whole-body muscle and fat components [11]. CT-derived muscle mass assessment of Skeletal Muscle Index (SMI) provides a measure of muscularity defined as Muscle Cross-Sectional Area in cm^2/height in m^2. Low muscle mass in patients with cancer is associated with adverse clinical outcomes such as length of hospital stays, chemotherapy toxicity, quality of life, and mortality [12-19].

Despite the advances in the clinical setting for numerous criteria to define cachexia, there is no consensus regarding patient stratification for application in molecular or bio-profiling studies. The definition and diagnostic criteria for cancer-associated muscle wasting are thus still in progress.

I refer to bio-profiling studies as those encompassing omic approaches (transcriptomics, epigenomics, proteomics, metabolomics) to identify potential molecular markers for cancer-associated muscle wasting. Clinically relevant cut-offs for low muscularity vary widely across molecular studies. These cut-offs were adopted from Prado et al. [15]. (2008) based on sex-specific muscle mass cut-offs associated with mortality in a population cohort of obese individuals with locally advanced/metastatic solid tumors, Martin et al. [16] (2013) sex-specific SMI cut-offs for varying BMI categories, and Kazemi et al. [20] (2015) sex- and age-stratified cut-offs. A systematic review and meta-analysis of the published cut-offs for low muscle mass and clinical outcomes are described [21,22].

Non-invasive CT-based prognostication is well-established and does not involve complex bio-profiling techniques, often not within reach of a clinical setting. While the SMI cut-offs or WL-BMI grading schema were informative for prognostication, applying the same criteria for association studies for molecular bio-profiling skeletal muscle is not ideal. This is because the endpoint in profiling studies using human skeletal muscle is to understand the underlying pathophysiology of muscle wasting, not the prognostic gene signatures. On the other hand, disease diagnostic markers are distinct from prognostic markers [23]. CC diagnostic markers from gene signatures are beneficial for identifying patients at risk of developing CC. Since CC is a multifactorial syndrome, developing diagnostic markers that show high specificity and sensitivity involves comprehensive bio-profiling of multiple tissues, including tumors, skeletal muscle, adipose tissue, and physiological fluids such as blood and urine.

Thus, there is no consensus on the clinical criteria for application to molecular studies. There is an urgent need for molecular studies encompassing unbiased omic approaches for patient stratification for identifying Complex Biological Processes (CBPs) driving the wasting phenomena. The vast majority of understanding of molecular mechanisms was obtained from experimental and model systems described below.

1.2 Bio-profiling attempts conducted to date to decipher underlying molecular mechanisms of cancer-associated muscle wasting (clinical studies and experimental models)

The pathophysiology of cancer-associated muscle wasting is multifaceted, and the exact mechanisms are not well-defined. Identifying biomarkers to predict the development of cancer-associated muscle wasting and response to therapeutic interventions is the primary outcome in experimental and clinical studies. Several factors impeding the validity and reproducibility across

studies are as follows: (i) limited sample size, (ii) heterogeneity in the patient stratification strategies employed in these studies, (iii) sex-specific differences are an emerging theme and were unaccounted for in studies using human muscle, (iv) use of multiple cancer types, (v) focus on a limited repertoire of candidate gene(s) or protein(s) from presently known pathways, and (vi) primarily studying protein-coding genes.

1.2.1 Sample size limitations

Samples for biological profiling of muscles are not readily available owing to the inherently invasive nature of skeletal muscle biopsy. This limits the available sample size and to conducting sex-specific stratified molecular association analyses. One means of increasing the sensitivity of analyses with relatively small samples is to capture the extremes of phenotype. For example, defining obesity-related biology by comparing individuals in the average body weight range (18.5-24.9 kg/m^2) to morbidly obese (>40 kg/m^2) [24]. For cachexia, the corollary would be a comparison of weight–stable individuals with those experiencing >10% weight loss (or weight loss-BMI ≥Grade 3).

In a recent study, Anoveros-Barerra et al. conducted a literature review of n=59 studies reporting human skeletal muscle biopsies from patients with cancer [25]. These studies had a limited sample size ranging from n=3 to 41, and 76% of studies included ≤ 30 participants with cancer. Moreover, 98% of the studies did not account for sex differences, and the results presented were aggregated.

Although the review addresses the critical limitations across human muscle bio-profiling studies, an in-depth understanding of gene expression profiles, their integrated higher-level regulatory mechanisms, and transcriptional changes in human skeletal muscle is yet to be determined.

1.2.2 Heterogeneity of cancer diagnosis and comparisons with healthy controls

The Prevalence of CC varies based on the cancer type, with the highest incidence in pancreas cancer (80%), about 50% of patients with lung, prostate, or colon cancer, and up to 40% in breast cancers and leukemias[26,27]. It is not presently known whether different primary cancers invoke different mechanisms at the tissue level in muscle, as muscle biopsy research has used aggregate cancer types for their analyses. Therefore, the bio-profiling studies may limit understanding of the molecular mechanisms when combined cancer types are used and preclude stratified analysis based on cancer types due to the sample size. In an attempt to identify cancer-specific mediators and secreted proteins with different prevalences of cachexia, Friere et al. [28] (2020) performed the characterization of expression profiling of secreted genes in 12 tumor types from publicly available TCGA (https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga) and GTEx databases (https://gtexportal.org/home/). The authors identified 25 candidate genes from a previous study [29]and termed Cachexia Inducing Factors (CIFs) that showed the highest number of up-regulated genes in pancreas cancer (cachexia is most prevalent in this cancer type as described above) and were associated with WL, cachexia prevalence, and neoplastic cellularity. A global perspective of the transcriptome profiling and regulatory mechanisms in skeletal muscle from patients with cancer is unclear.

Comparisons with healthy controls: Some gene expression studies compared cancer cachexic muscle and healthy controls without accounting for all potentially confounding variables [30-33]. It was recently shown that different skeletal muscles have intrinsic transcriptional diversity [34]; hence, comparisons across diverse muscle types or disease states may not be a feasible approach. Due to the limitations mentioned above, an unambiguous and coherent approach, independent of clinical

labels, is needed for molecular profiling studies to understand the molecular mechanisms contributing to disease etiology.

1.2.3 Candidate genes/pathways

Many experimental model systems have contributed to the underlying mechanisms of cancer-associated muscle wasting.[35-38]

Skeletal muscle mass is maintained by the balance between protein synthesis and degradation, promoting muscle hypertrophy (increase in the myofibrillar muscle mass) and atrophy (reduction in the myofibrillar muscle mass), respectively. Candidate or representative genes, proteins, and pathways implicated in the anabolic and catabolic processes are explicitly studied in cachexia literature. Such hypothesis testing approaches, however, are confined to (i) patient classification (case vs. control) based on investigator-defined cut-offs and criteria for %WL and SMI and (ii) quantifying select genes or proteins in a highly selected pathway of interest. Representative genes within the pathways regulating muscle mass were quantified using the Real-Time Polymerase Chain Reaction (qRT-PCR) technique or northern blot[31,33]. However, the results across the studies were inconsistent, and their role remains debated about their association with classical cachexia variables or underlying pathophysiological mechanisms[39,40]. For instance, the gene expression profiles within pathways associated with muscle protein degradation via Ubiquitin-Proteasome Pathway (UPP)[30,31,41] and autophagy and those mediating the processes of muscle atrophy or regeneration were inconsistent and variously showed no association or increased expression in muscle from the cachexic group[32,42,43]. The focus on the Ubiquitin-Proteasome Pathway came from experimental models that identified and suggested a group of genes termed atrogenes, including E3 Ubiquitin Ligases muscle atrophy F-box/atrogin1 (FBXO32) and Muscle RING finger 1 (MuRF1) contributing to cancer-associated muscle wasting. Hypothesis-testing studies from the

experimental model systems have contributed to a vast understanding of upstream regulators and mediators of cachexia (examples include pro-inflammatory cytokines, Transforming Growth Factor TGF-β superfamily members such as Activin A, Myostatin, Growth Differentiation Factor/GDF11, GDF-15, Parathyroid Hormone release Peptide/PTHrP, and several miRNAs).

Systemic inflammation is considered the hallmark of cachexia. Plasma-derived acute phase protein (C-Reactive Protein, CRP) is a marker of systemic inflammation. However, the clinical application of CRP as a marker of cachexia is not well-defined [44]. Inflammatory mediators, including TNF-α, IL6, IL1β, and Interferon γ (INF-γ), elicit catabolic responses, promoting skeletal muscle wasting. Hypothesis-testing studies alluded to individual pathways contributing to muscle hypertrophy and atrophy, including the PI3K/Akt/mTOR pathway, GDF1 mediated Bcl-/caspase 3 pathway, and IL6/STAT3 signaling, amongst several others [45].

Despite the experimental advances in understanding the catabolic processes driving muscle atrophy, the role of muscle regeneration is also debated. Skeletal muscle exhibits plasticity and the ability to regenerate following injury. This is driven by adult stem cells, termed satellite cells [46], located in close proximity to myofibers[47]. Upon injury or damage to muscle, satellite cells undergo a myogenic cascade of events, including myoblast proliferation, differentiation, and, eventually, fusion to form mature myofibers. It has been suggested that impaired skeletal muscle regeneration may contribute to muscle wasting in cancer patients,[48,49]. However, the role of dysregulation of skeletal muscle regeneration in cachexia is not well-defined.

Studying one pathway at a time does not provide mechanistic insights from a holistic perspective, given the complex and multifaceted nature of the syndrome.

The findings thus elucidate that complex, combinatorial, and multifactorial regulatory mechanisms drive the wasting phenomenon in patients with cancer. A more comprehensive approach to understanding global transcriptional changes is critical.

1.2.4 Sexual dimorphism

It is well-established from the clinical and body composition studies that the prevalence of cachexia is higher in males [50,51]. Sex-specific physiological differences in skeletal muscle are well described[51]. Men have more muscle mass, lower fat mass, a higher proportion of fast-twitch (type II, glycolytic) muscle fibers, and a higher rate of muscle loss compared to women [52,53]. Men oxidize less fat and more carbohydrate than women during women[54]. Women are more resistant to fatigue than men. These morphological, physiological, and metabolic differences accompany the molecular differences at the transcriptional and post-transcriptional levels. Overall, human skeletal muscle is sexually dimorphic concerning gene expression profiles, fiber type specifications, fatigue resistance, muscle strength, mitochondrial function, muscle protein turnover, and intrinsic signaling pathways within the muscle microenvironment [51,52,54-58].

Skeletal muscle exhibits diversity in the transcriptional changes reflected at the gene expression levels in a sex-specific manner [55]. Considerable efforts in gaining mechanistic insights from molecular profiling studies focus on males or remain unaccounted for sex as a potential confounding variable. Recent experimental hypothesis-testing studies have started acknowledging the importance of sex differences for understanding the molecular mechanisms of cancer-associated muscle wasting. Zhong et al. (2022) recently demonstrated the role of activin as a mediator of Pancreatic Ductal Adenocarcinoma (PDAC) associated cachexia in KPC mouse model [59]. They showed the importance of sex as a significant contributor to the underlying molecular mechanisms. Lindholm et al. (2014) performed RNA sequencing of *vastus lateralis* muscle biopsy

in 48 muscle biopsies (2 biopsies from each leg; and 2 biopsies within each leg ~ 3cm apart from n=6: 3 males and 3 females and a single biopsy from each leg of n=12 subjects: 6 males and 6 females)[58]. Although skeletal muscle tissue showed spatial tissue homogeneity, transcriptome profiles exhibited sex differences. A recent study used RNA-sequencing data from 700 individuals from 24 tissues from the Genotype-Tissue Expression project. Tissue-specific gene co-expression maps were generated, and it was found that 29.5% of the gene co-expression is influenced by sex: skeletal muscle of males showed predominantly stronger gene co-expression. Anoveros et al. reported sexual dimorphism at L3-total muscle cross-sectional area and rectus abdominis muscle, as well as in the mean fiber CSA and gene expression changes from literature-reported pathways in cachexia such as muscle growth, apoptosis, and inflammation[25]. Of the total n=59 studies the authors reviewed, n=50 studies that included both male and female patients were disproportionate. About 98% of studies on human muscle reported aggregate results. Despite the overarching sex-related differences in human skeletal muscle, no studies have construed comprehensive mechanisms underlying cancer-associated muscle wasting, broadscale RNA crosstalk, and gene regulation contributing to disease etiology. Although sex-specific differences are identified in skeletal muscle in aging and exercise settings [55,58], gene expression showed distinct profiles from those with cancer-associated wasting and needed to be thoroughly investigated.

1.2.5 Whole transcriptome profiling of human skeletal muscle and the role of non-coding RNAs

In contrast to the human muscle studies undertaken using candidate molecular approaches described in the previous section, literature on whole transcriptome profiling of human skeletal muscle is limited. This section introduces a description and justification of whole transcriptome profiling studies. An extensive repertoire of the human genome (~80-90%) previously considered

is transcribed into nonprotein-coding RNAs (ncRNAs) [60,61]. ncRNAs regulate protein-coding genes at epigenetic, transcriptional, and post-transcriptional levels and mediate changes in gene expression levels. ncRNAs have gained prominence as their expression dysregulation has severe implications in several diseases, including various cancer types, and are essential players in modulating complex cellular and molecular processes. There are two distinct classes of ncRNAs based on their length: (i) long ncRNAs (lncRNAs, > 200 nucleotides in length) and (ii) small ncRNAs (< 200 nucleotides in length) that include microRNAs, piwi-interacting RNAs (piRNAs), small nucleolar RNAs (snoRNAs), and transfer RNAs (tRNAs). The details for each are provided below:

1.2.5.1 *MicroRNAs (miRNAs)*

miRNAs (~ 22 nucleotides in length) represent the most studied class of small regulatory RNAs[62]. miRNAs are transcribed from the introns of the protein-coding genes, lncRNAs, or miRNA clusters. Recent evidence also suggests that they could be embedded within other small non-coding RNAs such as small nucleolar RNAs [63,64] (snoRNAs), and transfer RNAs [65] (tRNAs) as well as long non-coding RNAs [66,67] (lncRNAs). miRNAs (primary miRNAs/pri-miRNAs) are transcribed canonically from their genes and cleaved into 60-70nt precursor miRNAs/pre-miRNAs by DiGeorge Syndrome Critical Region 8 (DGCR8) and by Drosha, ribonuclease III that are part of microprocessor complex. Exportin 5 and Ran GTPase export pre-miRNAs into the cytoplasm from the nucleus, which is further processed and cleaved by Dicer, an RNAase III endonuclease. The guide strand or the functional strand is then incorporated into Argonaute protein as part of an RNA-induced silencing complex, which can target mRNAs and repress their expression by mediating mRNA degradation or translational inhibition [68].

Currently, there are about 2600 miRNAs reported in miRbase v.22 [69] (https://www.mirbase.org/). miRNAs function by base-pairing with the 3' untranslated region (3'UTR) of the target mRNAs and mediate mRNA degradation or translational repression depending on the sequence homology of the seed sequence. Perfect base pair complementarity (6-8 nucleotide seed sequence of miRNA) could lead to mRNA degradation, whereas imperfect complementarity results in translational repression. Recently, miRNAs are also known to interact with the 5'UTR, CDS, the coding sequence, and gene promoter regions, thereby regulating transcription and translation processes. miRNAs are the critical regulators of skeletal muscle function, and their role has been depicted in the processes of skeletal muscle atrophy and regeneration [70,71]. However, only two studies profiled using human skeletal muscle, one in lung [72], and the other in the pancreas and colorectal cancer[73], respectively, have elucidated the role of miRNAs in the pathophysiology of cancer cachexia. Narasimhan et al. profiled and identified 777 tissue-specific miRNAs in the *rectus abdominis* muscle (n=22 cachexic cases and n=20 non-cachexic cases) stratified based on the international consensus diagnostic criteria of CC, of which 82 miRNAs were expressed post-filtering of read counts of greater than 5 read counts in 75% of the samples[73]. van de Worp et al. (2020) profiled 754 miRNAs in *vastus lateralis* muscle using n=8 cachexic lung cancer patients and n=8 age-matched healthy controls [72]. While these studies provided the premise to identify the potential role of miRNAs underlying the pathophysiology of muscle wasting in cancer patients, they did not account for sex-specific differences due to the limited sample size. Also, they were limited in the availability of matched mRNA expression datasets for deciphering potential gene targets. miRNAs exhibit a pleiotropic effect on gene regulation, i.e., a single miRNA can regulate multiple gene targets, and several miRNAs can regulate a single gene target. A matched dataset to profile and identify miRNA-mRNA interactions from individual patient muscle biopsy is not yet generated;

hence, understanding miRNA-mediated gene regulation and its role in human muscle need further interpretations. Expression profiles generated from the same technology platform (for instance, NGS for both small RNA and RNA sequencing) open the avenues for performing correlation analysis and understanding whole transcriptome level gene regulation.

1.2.5.2 *Piwi-interacting RNAs (piRNAs)*

piRNAs are small single-stranded ncRNAs (26-31 nucleotides in length) [74], transcribed from introns of protein-coding genes, piRNA clusters, or snoRNAs [64,75] bind to Piwi proteins and mediate gene repression via epigenetic silencing and gene degradation [76]. piRNAs were previously thought to be expressed in the mammalian germline acting on transposon silencing and thus protecting the integrity of the genome from transposons, autonomous pieces of DNA that replicate and insert into the genome and have detrimental effects. However, recent evidence reveals their tissue-specific expression in various human tissues along with Piwi (P-element induced wimpy testis) proteins (PIWIL1/HIWI, PIWIL2/HILI, PIWIL3, and PIWIL4/HIWI2)[77]. It has shown to be novel diagnostic and prognostic in several cancer types [78,79]. Unlike miRNAs, piRNAs lack sequence conservation and are produced in a dicer-independent manner. Another distinct feature of piRNAs from miRNAs is that piRNA precursors do not exhibit hairpin structures and are not double-stranded. The exact mechanism of action of piRNA is still emerging. However, they are thought to exhibit similar mechanisms as miRNAs[80] (i.e., piRNA/piwi complex represses the translation via binding to the protein-coding gene). piRNAs are essential for stem cell pluripotency and are developmentally regulated, and their expression regulation has been identified in somatic cells. A recent study (La Greca et al., 2020) found 447 piRNAs (241 detected in pluripotency, 218 in the mesoderm, and 171 in cardiac cells) to have a role in fine-tuning the gene expression involved in the differentiation of pluripotent cells to cardiomyocytes [81]. Muscle satellite cells may

also harbor piRNA sequences and play a role in myofiber formation; however, this in the context of cachexia remains unknown. Coenen-Stass et al. (2018) identified differentially abundant small RNAs in Duchenne muscular dystrophy (DMD) muscle and serum samples in a mouse model of DMD [82]. No studies to date have identified and profiled piRNAs in the human skeletal muscle of patients with cancer. piRNA expression profiles and their role in undifferentiated and differentiated myoblast leading to mature myofiber formation remain to be determined.

1.2.5.3 *Small nucleolar RNAs (snoRNAs)*

snoRNAs (60-200 nucleotides in length) function canonically in the nucleolus. They guide the chemical modification of ribosomal RNAs (rRNAs) and small nuclear RNAs (snRNAs) by binding to the ribonucleoprotein complexes (snoRNPs) [83]. snoRNAs as classified into C/D (SNORD) or H/ACA box snoRNAs (SNORA), depending on the conserved sequence motifs. Recent studies have also depicted the non-canonical role of snoRNAs acting as the regulators of gene expression, wherein they harbor miRNA or piRNA sequences [64,75,84,85]. snoRNAs regulate their targets through complementary base-pairing with the target sequence. McCann et al. (2020) profiled snoRNAs in mouse embryonic stem cells and mouse myoblasts using small RNA sequencing and found that a subset of H/ACA snoRNAs are regulated during differentiation regulating the production of the ribosomes [86]. snoRNA expression profiles in the human skeletal muscle of patients with cancer and their role in muscle wasting remain obscure.

1.2.5.4 *Long non-coding RNAs (lncRNAs)*

lncRNAs are > 200 base pairs to several kilobases in length. lncRNA structure, interactions with DNA, RNA, and proteins, and cellular localization are the key aspects that dictate their functions. lncRNAs are the master regulators and act at epigenetic, transcriptional, post-transcriptional, and post-translational levels [87,88]. The same lncRNA could exhibit several roles depending on their

cellular localization and tissue-specific expression and are developmentally regulated. The human genome consists of about 16,000 lncRNA genes, and the functional role of lncRNAs is emerging. lncRNAs exhibit high tissue specificity and serve as excellent biomarkers and prognosticators. Recent studies have deciphered the functional role of lncRNAs in skeletal muscle development. However, these studies are limited to experimental model systems. One of the most studied mechanisms of action of lncRNAs is that they compete for the miRNA binding sites (miRNA Response Elements, MRE) with the mRNAs, regulating their expression levels [89-91] and thereby contributing to a particular phenotype. The ceRNA role of lncRNA lnc-MD1[92], a cytoplasmic lncRNA expressed during myoblast differentiation, was first studied in mouse and human myoblasts. lnc-MD1 acts as a ceRNA for miR-133 and miR-135 to regulate the expression of a mastermind-like transcriptional coactivator 1 (MAML1) and Myocyte Enhancer Factor 2C (MEF2C) transcription factors and thereby governs the timing of muscle differentiation. lncRNA lnc-mg, myogenesis-associated lncRNA overexpression in mice is known to promote hypertrophy and acts functionally as ceRNA for microRNA-1325b to control protein abundance of insulin-like growth factor 2 [93]. Li et al. (2021) recently demonstrated in a mouse C2C12 model that lncRNA Mir22hg harbors miR-22-3p and inhibits its target gene Histone Deacetylase 4 (HDAC4) and thereby increases the MEFC2, thus promoting myoblast differentiation[94]. Despite the advances in the functional characterization of lncRNAs in experimental model systems to understand skeletal muscle development and atrophy mechanisms, there is a paucity of literature regarding their expression and functional relevance to human skeletal muscle.

Studies utilizing mRNA profiling techniques in the skeletal muscle of patients with cancer are scarce. Researchers thus far have utilized qRT-PCR and microarray techniques to profile and identify mRNA dysregulation in cancer cachexia. qRT-PCR is an excellent technique for gene

expression and renders high specificity; however, it can be used to quantify representative candidate marker genes with apriori-defined sets of genes. Microarray quantifies thousands of genes from multiple samples, but it suffers from cross-hybridization of probes across the array, and there is reliance upon pre-existing knowledge about genome sequences. In contrast, Next Generation Sequencing (NGS) has revolutionized the understanding of health and disease by its ability to perform massively parallel sequencing of millions of sequence reads simultaneously per run. Depending on the sequencing depth, it also enables higher detection sensitivity. Moreover, the detection of non-coding transcripts, novel splice variants, and gene fusions can be uncovered using NGS technologies.

A recent study (Narasimhan et al., 2021) is the first study to generate mRNA expression profiles for muscle using ion proton sequencing using *rectus abdominis* muscle from patients with pancreas cancer (n=24, n=12 each of males and females) compared with patients diagnosed with benign lesions (n=11, n=5 and 6 respectively for males and females)[95]. The patient classification was based on consensus classification criteria for CC. The analyses schema did not encompass sex-related expression differences due to the limited sample size.

Overall, the expression landscape and higher-order interactions of RNAs in the human skeletal muscle of patients with cancer remain undetermined. No study in cachexia research has illuminated the unsupervised "omic" approach to decipher the intrinsic characteristics of skeletal muscle from surgical patients with cancer, independent of predefined phenotypes of interest.

1.3 Unsupervised clustering and Machine Learning methods for identification of potential molecular subtypes from the skeletal muscle of patients with cancer

Supervised learning is a machine learning approach that uses labeled information for classification or predicting outcomes of interest. In contrast, in unsupervised learning, the algorithms are applied to cluster unlabelled data and identify hidden patterns using clustering or dimensionality reduction methods. High throughput omics data suffer from a high dimensionality problem. There is a low sample size (typically less than 100) and massive expression data (expression of thousands of genes) acquired from whole transcriptome profiles generated from Next-generation sequencing or array technologies. One of the major challenges is to discover coherent gene expression patterns by a set of genes across a group of experimental conditions. These patterns may offer insights into biological processes associated with different physiological states. Unsupervised machine learning approaches such as Non-negative Matrix Factorization (NMF) [96,97] (also referred to as blind source separation or pattern recognition) are performed to reduce the dimensionality of the data yielding a representation of conditions as a linear combination of a reduced set of k-factors. These techniques aid in discovering patterns (clusters or subtypes) from high-throughput data, and their characterizations help identify Complex Biological Processes (CBP). The unsupervised methods have several advantages over the supervised techniques as they are independent of *a priori* assigned phenotype used for molecular association studies. The raw sequencing reads are aligned and quantified to the reference genome and transcriptome (i.e., annotated transcriptome databases for mRNAs, lncRNA, miRNAs, piRNAs, snoRNAs, and tRNAs), followed by normalization and log-transformation. These are then subjected to the Matrix Factorization algorithm. The Gene x Sample matrix (genes represented in the rows and samples in columns, known as Gene by Sample

matrix) is used as the input data. Matrix Factorization could be used to cluster genes or samples depending on the study objectives.

Given the normalized and transformed positive data matrix (A consisting of N rows and M columns), the NMF algorithm[96] iteratively computes an approximation A ~ W x H, where W and H are non-negative matrices with respective sizes N x k, and k x M. k here represents the number of clusters or subtypes. For the given rank *k*, NMF groups the samples into clusters. For every run, the algorithm may or may not converge to the same solution for rank k, depending on the random initial conditions. Therefore, several iterations or runs must be performed to reach consensus matrices. This means that if clustering into k classes is robust, the sample assignments to the respective clusters will remain invariant from every run. The model selection is based on consensus clustering to evaluate the robustness of the factorization (or decomposition). Quality matrices such as Silhouette Width (SW) are used to assess the goodness of the clustering technique. The value of silhouette width ranges from -1 to 1, wherein 0 means that the clusters are indifferent; 1 means that the clusters are well apart from each other and are clearly distinguishable, and -1 implies that the clusters are wrongly assigned. A silhouette plot can measure how close each sample in one cluster is compared to samples in the neighboring clusters. While NMF can identify clusters or subtypes for a single RNA class, integrative NMF (intNMF) can incorporate multi-omic or multi-RNA classes to identify molecular subtypes [98]. Thus far, studies addressing the cancer-associated muscle wasting condition considered mRNA or miRNA profiling from skeletal muscle to understand muscle biology based on predefined phenotypes. However, such an approach is still incomplete in understanding the RNA-cross talk and higher-order gene regulatory networks when other molecular entities (tRNAs, snoRNAs, piRNAs, and lncRNAs) are now known to contribute to the overall gene regulatory mechanisms. Readouts from high-throughput omics data are

complex and require integrative analytic approaches. Comprehensive characterization of molecular subtypes is now studied in several cancer types using these integrative approaches, such as Pancreatic Ductal Adenocarcinoma (PDAC), muscle-invasive bladder cancer, colon and rectal cancer, particularly in the consortia studies such as The Cancer Genome Atlas Research Network (TCGA) and others [99-101]. Moffitt et al. (2015) applied NMF to PDAC gene expression microarray data, including primary tumor, metastatic and normal samples, and identified distinct tumor and stroma-specific subtypes of PDAC [102].

Given the complexity of the muscle-wasting condition in patients with cancer, it is crucial to investigate if the global transcriptomic data, when subjected to unsupervised methods, help decipher molecular subtypes of human skeletal muscle. Such a systematic approach has never been undertaken in cancer cachexia research for molecular characterization and functional relevance.

1.4 Gene Expression Omnibus (GEO) datasets on mRNA and miRNA profiling studies in the skeletal muscle of patients with cancer

The literature-reported transcriptome profiling studies using mRNA or miRNAs in the human skeletal muscle of patients with cancer are described in table 1.1.

Table 1.1 Description of human skeletal muscle studies utilizing transcriptome methods to understand disease etiology

GEO dataset (Year)	Study objective (s)	Sample size and patient classification	Skeletal muscle anatomical location	Technology
GSE133979 (2021) Zimmers lab	Identified skeletal muscle and adipose tissue gene (mRNA) expression profiles in Pancreatic Ductal Adenocarcinoma (PDAC) cachexia patients	N=11 benign, n=24 PDAC patients. Classification based on WL-BMI grading system	Rectus abdominis	Ion Ampliseq
GSE130563 (2019) Judge lab	Skeletal muscle fibrosis in pancreatic cancer patients with respect to survival	N=16 non-cancer control, n=15 cachectic (>5%WL) PDAC patients and N=5 non-cachectic PDAC patients	Rectus abdominis	Microarray
GSE85017 (2018) Damaraju and Baracos lab	Differentially expressed alternatively spliced genes (mRNAs) in human skeletal muscle of patients with cancer and cachexia	N=21 cachexic, N=19 non-cachexic patients; classification based on International consensus	Rectus abdominis	Human Transcriptome Array (HTA 2.0)
GSE41726 (2018) Damaraju and Baracos lab	Sarcopenia and myosteatosis are accompanied by distinct profiles in patients with pancreatic and periampullary adenocarcinomas	N=29, data mined from a subset of samples utilized from the previously generated dataset in 2013, n=8 sarcopenia, n=10 myosteatosis, and n=15=neither	Rectus abdominis	Agilent Whole Human Genome Array

GSE75473 (2017) Damaraju and Baracos lab	miRNA profiling in human skeletal muscle of patients with cancer and cachexia	N=22 cachexic, N=20 non-cachexic patients; classification based on international consensus criteria	Rectus abdominis	MiSeq Next Generation Sequencing
GSE41726 (2013) Damaraju and Baracos lab	Sexual dimorphism in skeletal muscle gene (mRNA) expression was used to assess the effect of sample size on the differential expression. Rank order and prediction tasks	N=134, 69 men and 65 women	Rectus abdominis	Agilent Whole Human Genome Array
GSE34111 (2012) Fearon lab	Suppression of skeletal muscle turnover gene (mRNAs) in weight-losing Upper GI cancer	N=6 healthy controls, N=6 Upper Gastro-Intestinal Cancer/UGIC pre- and post-surgical resection; mean % WL 7%	Quadriceps	Affymetrix U133+2 platform
GSE18832 (2010) Fearon and Timmons lab	mRNA expression profiling in upper GI cancer patients. Identified CaMKIIβ and TIE1 genes associated with WL. FOXO proteins and ubiquitin E3 ligases commonly found in experimental studies were not related to WL in clinical study. Performed targeted and global gene profiling	N=21 (18 Upper Gastro-Intestinal Cancer/UGIC, 2 non-cancer controls) ; %WL > 5%	Rectus abdominis	Affymetrix HGU-133 + 2 array

Table 1.1: N=8 studies applied transcriptome profiling techniques from the human skeletal muscle of patients with cancer to elucidate the molecular mechanisms underlying the pathophysiology of cachexia. The GEO series number of these studies are listed as follows: GSE133979[95], GSE85017[103], GSE75473[73], GSE41726[104,105], GSE34111[40], GSE18832[39], GSE130563[106].

1.5 The rationale for conducting the study

Although mRNA expression profiles have advanced the understanding of cancer-associated muscle-wasting, there is a lack of consensus on comprehensive integrative molecular mechanisms. The expression landscape of human skeletal muscle from cancer patients using single-matched biopsy has never been generated. Attempts in this study are to generate comprehensive transcriptome profiles and systematically integrate RNAs to reveal the complex underlying biology, at least in a single biopsy specimen, as a representative specimen along the disease trajectory. Integrative analytical and experimental approaches could facilitate an understanding of molecular mechanisms underlying cancer-associated muscle wasting pathophysiology.

My thesis provides the cornerstone for future molecular profiling and association studies utilizing human skeletal muscle biopsy to understand the functional aspects of muscle wasting in patients with cancer.

1.6 Hypothesis and Objectives

Hypothesis: I hypothesize that the whole-genome transcriptome profiling and sex-specific gene expression profiles generated from human skeletal muscle of patients with cancer, when subjected to unsupervised machine learning clustering approaches, would facilitate the identification of the molecular subtypes and sex-specific pathophysiology. The dysregulated transcriptome and RNA networks from the identified subtypes may explain the Complex Biological Processes (CBP) and potential functional consequences.

Study objectives are organized as part of individual chapters as follows:

a) To accrue *rectus abdominis* muscle biopsies from patients with cancer intraoperatively, perform data acquisition from clinical and body composition analyses, and their phenotypic characterization using diagnostic criteria of %WL, and SMI **(Chapter 2)**
b) To generate transcriptome profiles from each patient biopsy and to perform unsupervised clustering to identify molecular subtypes of muscle. Two sub-objectives of this chapter include: (i) to identify Differentially Expressed (DE) genes from subtypes to explain complex regulatory pathways within the skeletal muscle and (ii) to perform clinical benchmarking analyses of the subtypes **(Chapter 3)**
c) To investigate multilayered RNA crosstalk through Competing Endogenous RNA (ceRNA) network analysis (lncRNA-miRNA-mRNA triplets) and delineate pathways regulating the Complex Biological Processes (CBPs) in human skeletal muscle **(Chapter 4)**
d) To identify converging pathways from transcriptional profiles in patients and in model systems **(Chapter 5),** to gain insights into the biological processes governing skeletal muscle homeostasis (functional benchmarking).

Overall, my study encompasses profiling and generating the entire transcriptional landscape of human skeletal muscle *(rectus abdominis)* in patients with cancer and their analysis to identify molecular subtypes. N=84 muscle biopsy specimens were procured and subjected to whole transcriptome profiling (Next-generation RNA sequencing and small RNA sequencing). I utilized in silico computational tools and experimental techniques to address the sex-specific molecular and functional characterization of human skeletal muscle subtypes.

1.7 Scope of this work

In Differential Expression profiling experiments, the intent is to get a high-level view of the dysregulated pathways, and in my data analysis, individual gene-based interpretations were avoided. Applying machine learning models to identify muscle subtypes, their transcriptome profiles, and integrative analyses of RNA networks offer insights into these subtypes. Further clinical relevance and functional insights could aid in characterizing the identified subtypes. My interest is to look for common or distinct pathways in males and females and determine if the sex-specific gene expression profiles and pathways are comparable with experimental model systems. I address this using the first experimental model of a cell-line model system in which proliferative myoblasts are chemically induced to differentiate into mature myofibers. The second experimental system of gastrocnemius skeletal muscle transcriptome from tumor-induced rodents is also explored for their pathway analysis overlap with human muscle subtypes. The scope of this work is limited to performing molecular characterization of human skeletal muscle from cancer patients. While other omic techniques, such as those using proteomics and metabolomics, are emerging in cachexia literature, my research addresses muscle biology at the post-transcriptional regulatory mechanisms.

1.8 References

1. Fearon K, Strasser F, Anker SD, Bosaeus I, Bruera E, Fainsinger RL, et al. Definition and classification of cancer cachexia: an international consensus. *Lancet Oncol.* 2011;12(5):489-495.
2. Fearon KC, Glass DJ, Guttridge DC. Cancer cachexia: mediators, signaling, and metabolic pathways. *Cell Metab.* 2012;16(2):153-166.
3. von Haehling S, Anker MS, Anker SD. Prevalence and clinical impact of cachexia in chronic illness in Europe, USA, and Japan: facts and numbers update 2016. *J Cachexia Sarcopenia Muscle.* 2016;7(5):507-509.
4. von Haehling S, Anker SD. Prevalence, incidence and clinical impact of cachexia: facts and numbers-update 2014. *J Cachexia Sarcopenia Muscle.* 2014;5(4):261-263.
5. Martin L, Senesse P, Gioulbasanis I, Antoun S, Bozzetti F, Deans C, et al. Diagnostic criteria for the classification of cancer-associated weight loss. *J Clin Oncol.* 2015;33(1):90-99.
6. Daly L, Dolan R, Power D, Ni Bhuachalla E, Sim W, Fallon M, et al. The relationship between the BMI-adjusted weight loss grading system and quality of life in patients with incurable cancer. *J Cachexia Sarcopenia Muscle.* 2020;11(1):160-168.
7. Vagnildhaug OM, Blum D, Wilcock A, Fayers P, Strasser F, Baracos VE, et al. The applicability of a weight loss grading system in cancer cachexia: a longitudinal analysis. *J Cachexia Sarcopenia Muscle.* 2017;8(5):789-797.
8. Dunne RF, Loh KP, Williams GR, Jatoi A, Mustian KM, Mohile SG. Cachexia and Sarcopenia in Older Adults with Cancer: A Comprehensive Review. *Cancers (Basel).* 2019;11(12).

9. Moreira-Pais A, Ferreira R, Oliveira PA, Duarte JA. Sarcopenia versus cancer cachexia: the muscle wasting continuum in healthy and diseased aging. *Biogerontology.* 2021;22(5):459-477.

10. Mourtzakis M, Prado CM, Lieffers JR, Reiman T, McCargar LJ, Baracos VE. A practical and precise approach to quantification of body composition in cancer patients using computed tomography images acquired during routine care. *Appl Physiol Nutr Metab.* 2008;33(5):997-1006.

11. Shen W, Punyanitya M, Wang Z, Gallagher D, St-Onge MP, Albu J, et al. Total body skeletal muscle and adipose tissue volumes: estimation from a single abdominal cross-sectional image. *J Appl Physiol (1985).* 2004;97(6):2333-2338.

12. Vaughan VC, Martin P, Lewandowski PA. Cancer cachexia: impact, mechanisms and emerging treatments. *J Cachexia Sarcopenia Muscle.* 2013;4(2):95-109.

13. Fearon K, Arends J, Baracos V. Understanding the mechanisms and treatment options in cancer cachexia. *Nat Rev Clin Oncol.* 2013;10(2):90-99.

14. Prado CM, Baracos VE, McCargar LJ, Mourtzakis M, Mulder KE, Reiman T, et al. Body composition as an independent determinant of 5-fluorouracil-based chemotherapy toxicity. *Clin Cancer Res.* 2007;13(11):3264-3268.

15. Prado CM, Lieffers JR, McCargar LJ, Reiman T, Sawyer MB, Martin L, et al. Prevalence and clinical implications of sarcopenic obesity in patients with solid tumours of the respiratory and gastrointestinal tracts: a population-based study. *Lancet Oncol.* 2008;9(7):629-635.

16. Martin L, Birdsell L, Macdonald N, Reiman T, Clandinin MT, McCargar LJ, et al. Cancer cachexia in the age of obesity: skeletal muscle depletion is a powerful prognostic factor, independent of body mass index. *J Clin Oncol.* 2013;31(12):1539-1547.

17. Barret M, Antoun S, Dalban C, Malka D, Mansourbakht T, Zaanan A, et al. Sarcopenia is linked to treatment toxicity in patients with metastatic colorectal cancer. *Nutr Cancer.* 2014;66(4):583-589.

18. Prado CM, Baracos VE, McCargar LJ, Reiman T, Mourtzakis M, Tonkin K, et al. Sarcopenia as a determinant of chemotherapy toxicity and time to tumor progression in metastatic breast cancer patients receiving capecitabine treatment. *Clin Cancer Res.* 2009;15(8):2920-2926.

19. Lieffers JR, Bathe OF, Fassbender K, Winget M, Baracos VE. Sarcopenia is associated with postoperative infection and delayed recovery from colorectal cancer resection surgery. *Br J Cancer.* 2012;107(6):931-936.

20. Kazemi-Bajestani SM, Mazurak VC, Baracos V. Computed tomography-defined muscle and fat wasting are associated with cancer clinical outcomes. *Semin Cell Dev Biol.* 2016;54:2-10.

21. Rinninella E, Cintoni M, Raoul P, Pozzo C, Strippoli A, Bria E, et al. Muscle mass, assessed at diagnosis by L3-CT scan as a prognostic marker of clinical outcomes in patients with gastric cancer: A systematic review and meta-analysis. *Clin Nutr.* 2020;39(7):2045-2054.

22. Xia L, Zhao R, Wan Q, Wu Y, Zhou Y, Wang Y, et al. Sarcopenia and adverse health-related outcomes: An umbrella review of meta-analyses of observational studies. *Cancer Med.* 2020;9(21):7964-7978.

23. Oldenhuis CN, Oosting SF, Gietema JA, de Vries EG. Prognostic versus predictive value of biomarkers in oncology. *Eur J Cancer.* 2008;44(7):946-953.

24. Lu Z, Meng L, Sun Z, Shi X, Shao W, Zheng Y, et al. Differentially Expressed Genes and Enriched Signaling Pathways in the Adipose Tissue of Obese People. *Front Genet.* 2021;12:620740.

25. Anoveros-Barrera A, Bhullar AS, Stretch C, Esfandiari N, Dunichand-Hoedl AR, Martins KJB, et al. Clinical and biological characterization of skeletal muscle tissue biopsies of surgical cancer patients. *J Cachexia Sarcopenia Muscle.* 2019;10(6):1356-1377.

26. Argiles JM, Busquets S, Stemmler B, Lopez-Soriano FJ. Cancer cachexia: understanding the molecular basis. *Nat Rev Cancer.* 2014;14(11):754-762.

27. Baracos VE, Martin L, Korc M, Guttridge DC, Fearon KCH. Cancer-associated cachexia. *Nat Rev Dis Primers.* 2018;4:17105.

28. Freire PP, Fernandez GJ, de Moraes D, Cury SS, Dal Pai-Silva M, Dos Reis PP, et al. The expression landscape of cachexia-inducing factors in human cancers. *J Cachexia Sarcopenia Muscle.* 2020;11(4):947-961.

29. Talbert EE, Lewis HL, Farren MR, Ramsey ML, Chakedis JM, Rajasekera P, et al. Circulating monocyte chemoattractant protein-1 (MCP-1) is associated with cachexia in treatment-naive pancreatic cancer patients. *J Cachexia Sarcopenia Muscle.* 2018;9(2):358-368.

30. Williams A, Sun X, Fischer JE, Hasselgren PO. The expression of genes in the ubiquitin-proteasome proteolytic pathway is increased in skeletal muscle from patients with cancer. *Surgery.* 1999;126(4):744-749; discussion 749-750.

31. Khal J, Hine AV, Fearon KC, Dejong CH, Tisdale MJ. Increased expression of proteasome subunits in skeletal muscle of cancer patients with weight loss. *Int J Biochem Cell Biol.* 2005;37(10):2196-2206.

32. Pessina P, Conti V, Pacelli F, Rosa F, Doglietto GB, Brunelli S, et al. Skeletal muscle of gastric cancer patients expresses genes involved in muscle regeneration. *Oncol Rep.* 2010;24(3):741-745.

33. Rhoads MG, Kandarian SC, Pacelli F, Doglietto GB, Bossola M. Expression of NF-kappaB and IkappaB proteins in skeletal muscle of gastric cancer patients. *Eur J Cancer.* 2010;46(1):191-197.

34. Terry EE, Zhang X, Hoffmann C, Hughes LD, Lewis SA, Li J, et al. Transcriptional profiling reveals extraordinary diversity among skeletal muscle tissues. *Elife.* 2018;7.

35. Olson B, Norgard MA, Levasseur PR, Zhu X, Marks DL. Physiologic and molecular characterization of a novel murine model of metastatic head and neck cancer cachexia. *J Cachexia Sarcopenia Muscle.* 2021;12(5):1312-1332.

36. Huot JR, Novinger LJ, Pin F, Narasimhan A, Zimmers TA, O'Connell TM, et al. Formation of colorectal liver metastases induces musculoskeletal and metabolic abnormalities consistent with exacerbated cachexia. *JCI Insight.* 2020;5(9).

37. Talbert EE, Cuitino MC, Ladner KJ, Rajasekerea PV, Siebert M, Shakya R, et al. Modeling Human Cancer-induced Cachexia. *Cell Rep.* 2019;28(6):1612-1622 e1614.

38. Suzuki T, Von Haehling S, Springer J. Promising models for cancer-induced cachexia drug discovery. *Expert Opin Drug Discov.* 2020;15(5):627-637.

39. Stephens NA, Gallagher IJ, Rooyackers O, Skipworth RJ, Tan BH, Marstrand T, et al. Using transcriptomics to identify and validate novel biomarkers of human skeletal muscle cancer cachexia. *Genome Med.* 2010;2(1):1.

40. Gallagher IJ, Stephens NA, MacDonald AJ, Skipworth RJ, Husi H, Greig CA, et al. Suppression of skeletal muscle turnover in cancer cachexia: evidence from the transcriptome in sequential human muscle biopsies. *Clin Cancer Res.* 2012;18(10):2817-2827.

41. Bossola M, Muscaritoli M, Costelli P, Grieco G, Bonelli G, Pacelli F, et al. Increased muscle proteasome activity correlates with disease severity in gastric cancer patients. *Ann Surg.* 2003;237(3):384-389.

42. Schersten T, Lundholm K. Lysosomal enzyme activity in muscle tissue from patients with malignant tumor. *Cancer.* 1972;30(5):1246-1251.

43. D'Orlando C, Marzetti E, Francois S, Lorenzi M, Conti V, di Stasio E, et al. Gastric cancer does not affect the expression of atrophy-related genes in human skeletal muscle. *Muscle Nerve.* 2014;49(4):528-533.

44. Tavares P, Goncalves DM, Santos LL, Ferreira R. Revisiting the clinical usefulness of C-reactive protein in the set of cancer cachexia. *Porto Biomed J.* 2021;6(1):e123.

45. Mueller TC, Bachmann J, Prokopchuk O, Friess H, Martignoni ME. Molecular pathways leading to loss of skeletal muscle mass in cancer cachexia--can findings from animal models be translated to humans? *BMC Cancer.* 2016;16:75.

46. Mauro A. Satellite cell of skeletal muscle fibers. *J Biophys Biochem Cytol.* 1961;9(2):493-495.

47. Le Grand F, Rudnicki MA. Skeletal muscle satellite cells and adult myogenesis. *Curr Opin Cell Biol.* 2007;19(6):628-633.

48. Arneson PC, Doles JD. Impaired Muscle Regeneration in Cancer-Associated Cachexia. *Trends Cancer.* 2019;5(10):579-582.

49. Bossola M, Marzetti E, Rosa F, Pacelli F. Skeletal muscle regeneration in cancer cachexia. *Clin Exp Pharmacol Physiol.* 2016;43(5):522-527.

50. Zhong X, Zimmers TA. Sex Differences in Cancer Cachexia. *Curr Osteoporos Rep.* 2020;18(6):646-654.

51. Montalvo RN, Counts BR, Carson JA. Understanding sex differences in the regulation of cancer-induced muscle wasting. *Curr Opin Support Palliat Care.* 2018;12(4):394-403.

52. Smith GI, Mittendorfer B. Sexual dimorphism in skeletal muscle protein turnover. *J Appl Physiol (1985).* 2016;120(6):674-682.

53. Bredella MA. Sex Differences in Body Composition. *Adv Exp Med Biol.* 2017;1043:9-27.

54. Landen S, Hiam D, Voisin S, Jacques M, Lamon S, Eynon N. Physiological and molecular sex differences in human skeletal muscle in response to exercise training. *J Physiol.* 2021.

55. Welle S, Tawil R, Thornton CA. Sex-related differences in gene expression in human skeletal muscle. *PLoS One.* 2008;3(1):e1385.

56. Maher AC, Fu MH, Isfort RJ, Varbanov AR, Qu XA, Tarnopolsky MA. Sex differences in global mRNA content of human skeletal muscle. *PLoS One.* 2009;4(7):e6335.

57. Norman K, Stobaus N, Reiss J, Schulzke J, Valentini L, Pirlich M. Effect of sexual dimorphism on muscle strength in cachexia. *J Cachexia Sarcopenia Muscle.* 2012;3(2):111-116.

58. Lindholm ME, Huss M, Solnestam BW, Kjellqvist S, Lundeberg J, Sundberg CJ. The human skeletal muscle transcriptome: sex differences, alternative splicing, and tissue homogeneity assessed with RNA sequencing. *FASEB J.* 2014;28(10):4571-4581.

59. Zhong X, Narasimhan A, Silverman LM, Young AR, Shahda S, Liu S, et al. Sex specificity of pancreatic cancer cachexia phenotypes, mechanisms, and treatment in mice and humans: role of Activin. *J Cachexia Sarcopenia Muscle.* 2022;13(4):2146-2161.

60. Mattick JS, Makunin IV. Non-coding RNA. *Hum Mol Genet.* 2006;15 Spec No 1:R17-29.

61. Carninci P, Kasukawa T, Katayama S, Gough J, Frith MC, Maeda N, et al. The transcriptional landscape of the mammalian genome. *Science.* 2005;309(5740):1559-1563.

62. Bartel DP. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell.* 2004;116(2):281-297.

63. Scott MS, Ono M. From snoRNA to miRNA: Dual function regulatory non-coding RNAs. *Biochimie.* 2011;93(11):1987-1992.

64. Krishnan P, Ghosh S, Wang B, Heyns M, Graham K, Mackey JR, et al. Profiling of Small Nucleolar RNAs by Next Generation Sequencing: Potential New Players for Breast Cancer Prognosis. *PLoS One.* 2016;11(9):e0162622.

65. Krishnan P, Ghosh S, Wang B, Heyns M, Li D, Mackey JR, et al. Genome-wide profiling of transfer RNAs and their role as novel prognostic markers for breast cancer. *Sci Rep.* 2016;6:32843.

66. He D, Wu D, Muller S, Wang L, Saha P, Ahanger SH, et al. miRNA-independent function of long noncoding pri-miRNA loci. *Proc Natl Acad Sci U S A.* 2021;118(13).

67. Ulitsky I. Interactions between short and long noncoding RNAs. *FEBS Lett.* 2018;592(17):2874-2883.

68. Ha M, Kim VN. Regulation of microRNA biogenesis. *Nat Rev Mol Cell Biol.* 2014;15(8):509-524.

69. Kozomara A, Birgaoanu M, Griffiths-Jones S. miRBase: from microRNA sequences to function. *Nucleic Acids Res.* 2019;47(D1):D155-D162.

70. Ge Y, Chen J. MicroRNAs in skeletal myogenesis. *Cell Cycle.* 2011;10(3):441-448.

71. Marceca GP, Nigita G, Calore F, Croce CM. MicroRNAs in Skeletal Muscle and Hints on Their Potential Role in Muscle Wasting During Cancer Cachexia. *Front Oncol.* 2020;10:607196.

72. van de Worp W, Schols A, Dingemans AC, Op den Kamp CMH, Degens J, Kelders M, et al. Identification of microRNAs in skeletal muscle associated with lung cancer cachexia. *J Cachexia Sarcopenia Muscle.* 2020;11(2):452-463.

73. Narasimhan A, Ghosh S, Stretch C, Greiner R, Bathe OF, Baracos V, et al. Small RNAome profiling from human skeletal muscle: novel miRNAs and their targets associated with cancer cachexia. *J Cachexia Sarcopenia Muscle.* 2017;8(3):405-416.

74. Aravin A, Gaidatzis D, Pfeffer S, Lagos-Quintana M, Landgraf P, Iovino N, et al. A novel class of small RNAs bind to MILI protein in mouse testes. *Nature.* 2006;442(7099):203-207.

75. Zhong F, Zhou N, Wu K, Guo Y, Tan W, Zhang H, et al. A SnoRNA-derived piRNA interacts with human interleukin-4 pre-mRNA and induces its decay in nuclear exosomes. *Nucleic Acids Res.* 2015;43(21):10474-10491.

76. Ozata DM, Gainetdinov I, Zoch A, O'Carroll D, Zamore PD. PIWI-interacting RNAs: small RNAs with big functions. *Nat Rev Genet.* 2019;20(2):89-108.

77. Han YN, Li Y, Xia SQ, Zhang YY, Zheng JH, Li W. PIWI Proteins and PIWI-Interacting RNA: Emerging Roles in Cancer. *Cell Physiol Biochem.* 2017;44(1):1-20.

78. Li J, Wang N, Zhang F, Jin S, Dong Y, Dong X, et al. PIWI-interacting RNAs are aberrantly expressed and may serve as novel biomarkers for diagnosis of lung adenocarcinoma. *Thorac Cancer.* 2021;12(18):2468-2477.

79. Muller S, Raulefs S, Bruns P, Afonso-Grunz F, Plotner A, Thermann R, et al. Next-generation sequencing reveals novel differentially regulated mRNAs, lncRNAs, miRNAs, sdRNAs and a piRNA in pancreatic cancer. *Mol Cancer.* 2015;14:94.

80. Krishnan P, Ghosh S, Graham K, Mackey JR, Kovalchuk O, Damaraju S. Piwi-interacting RNAs and PIWI genes as novel prognostic markers for breast cancer. *Oncotarget.* 2016;7(25):37944-37956.

81. La Greca A, Scarafia MA, Hernandez Canas MC, Perez N, Castaneda S, Colli C, et al. PIWI-interacting RNAs are differentially expressed during cardiac differentiation of human pluripotent stem cells. *PLoS One.* 2020;15(5):e0232715.

82. Coenen-Stass AML, Sork H, Gatto S, Godfrey C, Bhomra A, Krjutskov K, et al. Comprehensive RNA-Sequencing Analysis in Serum and Muscle Reveals Novel Small RNA Signatures with Biomarker Potential for DMD. *Mol Ther Nucleic Acids.* 2018;13:1-15.

83. Kiss T. Small nucleolar RNAs: an abundant group of noncoding RNAs with diverse cellular functions. *Cell.* 2002;109(2):145-148.

84. Krishnan P, Ghosh S, Wang B, Li D, Narasimhan A, Berendt R, et al. Next generation sequencing profiling identifies miR-574-3p and miR-660-5p as potential novel prognostic markers for breast cancer. *BMC Genomics.* 2015;16:735.

85. Wajahat M, Bracken CP, Orang A. Emerging Functions for snoRNAs and snoRNA-Derived Fragments. *Int J Mol Sci.* 2021;22(19).

86. McCann KL, Kavari SL, Burkholder AB, Phillips BT, Hall TMT. H/ACA snoRNA levels are regulated during stem cell differentiation. *Nucleic Acids Res.* 2020;48(15):8686-8703.

87. Kung JT, Colognori D, Lee JT. Long noncoding RNAs: past, present, and future. *Genetics.* 2013;193(3):651-669.

88. Kopp F, Mendell JT. Functional Classification and Experimental Dissection of Long Noncoding RNAs. *Cell.* 2018;172(3):393-407.

89. Kartha RV, Subramanian S. Competing endogenous RNAs (ceRNAs): new entrants to the intricacies of gene regulation. *Front Genet.* 2014;5:8.

90. Salmena L, Poliseno L, Tay Y, Kats L, Pandolfi PP. A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language? *Cell.* 2011;146(3):353-358.

91. Bosia C, Pagnani A, Zecchina R. Modelling Competing Endogenous RNA Networks. *PLoS One.* 2013;8(6):e66609.

92. Cesana M, Cacchiarelli D, Legnini I, Santini T, Sthandier O, Chinappi M, et al. A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA. *Cell.* 2011;147(2):358-369.

93. Zhu M, Liu J, Xiao J, Yang L, Cai M, Shen H, et al. Lnc-mg is a long non-coding RNA that promotes myogenesis. *Nat Commun.* 2017;8:14718.

94. Ye J, She X, Liu Z, He Z, Gao X, Lu L, et al. Eukaryotic Initiation Factor 4A-3: A Review of Its Physiological Role and Involvement in Oncogenesis. *Front Oncol.* 2021;11:712045.

95. Narasimhan A, Zhong X, Au EP, Ceppa EP, Nakeeb A, House MG, et al. Profiling of Adipose and Skeletal Muscle in Human Pancreatic Cancer Cachexia Reveals Distinct Gene Profiles with Convergent Pathways. *Cancers (Basel).* 2021;13(8).

96. Brunet JP, Tamayo P, Golub TR, Mesirov JP. Metagenes and molecular pattern discovery using matrix factorization. *Proc Natl Acad Sci U S A.* 2004;101(12):4164-4169.

97. Stein-O'Brien GL, Arora R, Culhane AC, Favorov AV, Garmire LX, Greene CS, et al. Enter the Matrix: Factorization Uncovers Knowledge from Omics. *Trends Genet.* 2018;34(10):790-805.

98. Chalise P, Fridley BL. Integrative clustering of multi-level 'omic data based on non-negative matrix factorization algorithm. *PLoS One.* 2017;12(5):e0176278.

99. Cancer Genome Atlas Research Network. Electronic address aadhe, Cancer Genome Atlas Research N. Integrated Genomic Characterization of Pancreatic Ductal Adenocarcinoma. *Cancer Cell.* 2017;32(2):185-203 e113.

100. Robertson AG, Kim J, Al-Ahmadie H, Bellmunt J, Guo G, Cherniack AD, et al. Comprehensive Molecular Characterization of Muscle-Invasive Bladder Cancer. *Cell.* 2018;174(4):1033.

101. Cancer Genome Atlas N. Comprehensive molecular characterization of human colon and rectal cancer. *Nature.* 2012;487(7407):330-337.

102. Moffitt RA, Marayati R, Flate EL, Volmar KE, Loeza SG, Hoadley KA, et al. Virtual microdissection identifies distinct tumor- and stroma-specific subtypes of pancreatic ductal adenocarcinoma. *Nat Genet.* 2015;47(10):1168-1178.

103. Narasimhan A, Greiner R, Bathe OF, Baracos V, Damaraju S. Differentially expressed alternatively spliced genes in skeletal muscle from cancer patients with cachexia. *J Cachexia Sarcopenia Muscle.* 2018;9(1):60-70.

104. Stretch C, Khan S, Asgarian N, Eisner R, Vaisipour S, Damaraju S, et al. Effects of sample size on differential gene expression, rank order and prediction accuracy of a gene signature. *PLoS One.* 2013;8(6):e65380.

105. Stretch C, Aubin JM, Mickiewicz B, Leugner D, Al-Manasra T, Tobola E, et al. Sarcopenia and myosteatosis are accompanied by distinct biological profiles in patients with pancreatic and periampullary adenocarcinomas. *PLoS One.* 2018;13(5):e0196235.

106. Judge SM, Nosacka RL, Delitto D, Gerber MH, Cameron ME, Trevino JG, et al. Skeletal Muscle Fibrosis in Pancreatic Cancer Patients with Respect to Survival. *JNCI Cancer Spectr.* 2018;2(3):pky043.

Chapter 2 Challenges in patient stratification using clinical labels for cachexia research and molecular association studies

2.1 Introduction

The defining characteristic of cancer cachexia is unintentional weight loss, primarily due to skeletal muscle depletion [1]. Traditionally, the determination of the presence of cachexia in patients with cancer relied on patient-reported weight changes or those obtained from medical charts. WL is known to mask the hidden muscle loss or dynamic changes within the muscle during patients' disease trajectory [2-4]. Imaging modalities such as the opportunistic use of Computed Tomography (CT) scan and the quantitative image analysis at a specific landmark (Lumbar vertebrae, L3) representative of whole body muscle and fat components have revolutionized cachexia research in the last decade [5-7]. The ability of image-based technologies to differentiate body composition components and their high specificity to discriminate organs and tissues have provided value-added information for cachexia in patients with cancer. Several researchers have utilized retrospective datasets from medical records and image repositories to assess low muscle mass, radiodensity, and components of adipose tissue, including Subcutaneous Adipose Tissue (SAT), Intermuscular Adipose Tissue (IMAT), and Visceral Adipose Tissue (VAT). The total muscle and fat cross-sectional area segmented at L3 normalized by stature provide indices known as the Skeletal Muscle Index (SMI, cm^2/m^2) and Fat Mass Index (FMI, cm^2/m^2), respectively, as a measure of muscle and fat components. Clinical association studies using CT-based image analysis and their association with surgical and oncological outcomes have substantially advanced the field of cachexia [2,3,5,8-11].

On the molecular forefront, despite decades of research in understanding mechanisms underlying the pathophysiology of cachexia, there are no druggable targets or FDA-approved agents to treat

or reverse cachexia. An overall consensus is that more human studies are required to understand the molecular mechanisms governing cachexia. Skeletal muscle biopsy acquired by the dedicated surgical oncologist during the surgical resection of the tumor is an excellent tool to explore biochemical, histological, and molecular data. The accrual of muscle biopsy specimens (which otherwise would be an invasive procedure) is considered a minimally invasive procedure since the sample collection is performed during surgery.

To date, molecular studies exploring gene expression and regulation using human skeletal muscle biopsies from patients with cancer have not yielded consensus results with animal models or across human studies. The research in this area is sparse since the muscle biopsy accrual, otherwise, is an arduous process and requires extensive collaboration between surgical oncologists and researchers. Ana Anoveros and colleagues performed a state-of-the-science literature review on human skeletal muscle biopsy studies (n=59). They reported inconsistencies across studies regarding sample size limitations, aggregate sex measures, and differing classification criteria, contributing to low quality and risk of sampling bias [12]. Comparisons at the molecular level changes were not the focus of the review and hence were not reported.

Decades of research have focused on quantifying gene expression using northern blot or qRT-PCR. These studies were mainly hypothesis-testing and presented with no consensus regarding the comprehensive changes at the molecular level. The hypothesis-generating whole genome studies are limited. N=8 mRNA [13-18]/miRNA [19,20] profiling studies from the human skeletal muscle of patients with cancer are reported. A robust dataset with complete clinical annotation and an entire transcriptional landscape from the muscle biopsy of patients with cancer has never been generated.

A compelling question that remains unanswered is the classification or phenotyping of patients for molecular profiling studies. The investigators in human cachexia research have applied varying

classification strategies to classify patients as cachexic or non-cachexic based on the conventionally defined variables of interest, as mentioned in the previous sections. These include varying degrees of percent Weight Loss (%WL, which also includes investigator-defined cut-off values for %WL to classify patients) and low SMI determined using literature-derived cut-offs or combinations of these as described by the 2011 international diagnostic criteria for cancer cachexia. Since muscle mass loss could be hidden within the weight-stable patients and the heterogeneity related to classification strategies has never been looked at in the molecular association study.

The traditional molecular association studies using muscle biopsies relied on phenotypic binary (clinical labels). However, there is a lack of consensus on the application of clinical labels for molecular association studies, which in part, could be one of the reasons for the incongruency and validity of findings. Another significant aspect impeding the replicability of findings is that the studies have not accounted for sex-specific differences, which is an important confounding variable. Sex-specific differences in the biological mechanisms driving muscle wasting phenomenon are not understood in human skeletal muscle studies from patients with cancer.

Thus far, no study has identified clinical heterogeneity within individual male and female patients for the clinical labels. Whether the application of these for patient classification, particularly for molecular association studies, is unknown.

Hypothesis: I hypothesize that applying conventional clinical labels (WL, SMI, and BMI) for cachexia could present heterogeneity and represent inter-individual variations (concerning the clinical labels) and limit their applicability for molecular association studies.

Objectives: (i) to accrue clinical data and perform body composition analyses for patients with cancer undergoing surgery for tumor resection (ii) to perform bivariate analyses of clinical labels and identify the presence of patient-level heterogeneity (iii) to provide recommendations for future molecular profiling studies

This study focusses on the clinical annotation (or phenotyping, used interchangeably) of samples for which the muscle biopsy specimens were accrued.

2.2 Methods*

2.2.1 Patient recruitment, demographic, and clinical data acquisition

Medical records were accessed to obtain patient clinical history (demographic, clinical, and survival data)†. The data from medical records included: age (>18 years), detailed weight history over time, height, Body Mass Index [BMI, weight (kg) /height (m^2)], sex, cancer type, date of surgery, and time to death or censoring where applicable. Inclusion criteria: Pancreatic and colorectal patients with cancer (n=84) from the biobank specimens were selected to reduce cancer site heterogeneity and relatively high prevalence of WL and muscle loss. Exclusion criteria included missing weight histories, imaging history, and insufficient availability of samples for bio-profiling.

The percent weight loss (%WL) was calculated from medical records as follows:

(current weight – previous weight / previous weight) x 100

Patients were assigned composite grades based on the Weight Loss-Body Mass Index (WL-BMI) grading system ranging from 0-4 [21].

* Study was conceptualized by Bhumi Bhatt, Dr. Vickie Baracos and Dr. Sambasivarao Damaraju
† Medical records were obtained by Bhumi Bhatt and Dr. Vickie Baracos.

2.2.2 Body composition analysis

Cross-sectional CT imaging was used to assess the body composition in patients with cancer[2,3,21,22]. Pre-operative CT images were analyzed[‡] for quantifying muscle and fat tissue cross-sectional area (cm^2) using Slice-O-Matic software Tomovision, Montreal, Quebec, Canada. Single and consecutive CTs obtained pre-operatively at Lumbar vertebrae 3 (L3) were analyzed where applicable to quantify and assess the dynamic changes in muscle and fat components. Muscle radiation attenuation was measured and reported in Hounsfield Units (HU) (range: -29 to +150 HU). The specific Hounsfield Unit (HU) range for fat components is as follows: -190 to -30 for Subcutaneous Adipose Tissue (SAT), -190 to -30 for Intramuscular Adipose Tissue (IMAT), and -150 to -50 for Visceral Adipose Tissue (VAT). Total Adipose Tissue (TAT) was calculated as the sum of SAT, VAT, and IMAT and reported in cm^2. Muscle and fat cross-sectional area was normalized for stature and reported as Skeletal Muscle Index (SMI, cm^2/m^2) and Fat Mass Index (FMI, cm^2/m^2), respectively. Estimates of muscle and fat stores (in kg) were calculated using the following regression equation[5]:

Whole-body muscle mass (kg) = 0.30 x [skeletal muscle at L3 using CT (cm^2)] + 6.06

Whole-body fat mass (kg) = 0.042 x [fat tissue at L3 using CT (cm^2)] + 11.2

Net %WL was inferred from the cumulative net muscle and fat loss quantified using consecutive CTs wherever patient-reported weight histories were missing. WL, SMI, TAT, and FMI were interchangeably referred to as body composition variables or clinical labels. Sex-specific median bivariate cut-offs were generated for WL, SMI, and FMI.

[‡] Bhumi Bhatt obtained certification to perform CT image analysis at the Cross Cancer Institute (2019).

2.2.3 Statistical analysis

Statistical analyses were performed using Rstudio and SPSS v27. Differences between groups were analyzed using an independent t-test for continuous variables and Pearson's χ^2 test for categorical variables. Non-normally distributed variables were compared using the Mann-Whitney U test. Correlations between continuous variables were assessed using Pearson correlation coefficients using Hmisc and corrplot R packages. Survival analysis was performed using the Kaplan-Meier method, and the differences were compared using log-rank tests. A nominal p-value of 0.05 was considered statistically significant unless stated otherwise.

2.3 Results

2.3.1 Study cohort and patient characteristics

Each *Rectus abdominis* muscle biopsy (n=84) was procured and served to generate small RNA and RNA sequencing and downstream data pre-processing. Clinical data, including weight history, cancer diagnosis, and survival data, were obtained, and body composition analysis was used to assess skeletal muscle and fat components.

This study focused on studying the sex-related differences in the skeletal muscle of patients with cancer, and the analysis presented will be sex-specific unless stated otherwise. The cohort was 57% male (see clinical characteristics in **Table 2.1)**. No statistically significant differences were observed in clinical variables such as age, BMI, cancer type, and % weight changes between males and females. Significant differences in body composition variables such as SMI and FMI were observed. SMI was higher in males, and FMI was higher in females.

Table 2.1 Clinical, body composition and survival analysis for male and female patients with cancer

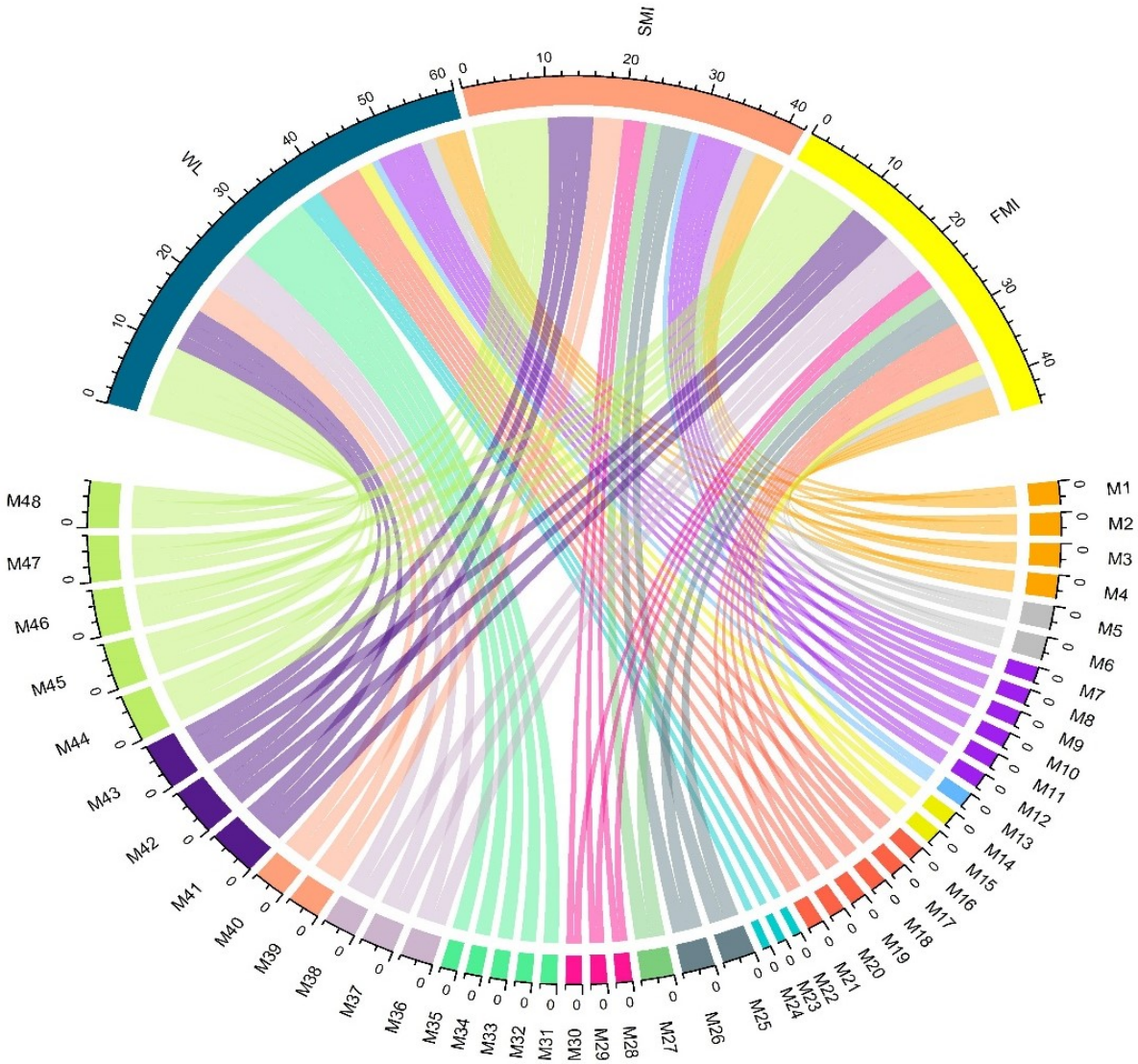
Demographic or clinical variable	Male (n=48)	Female (n=36)	P-value
Age, years	64 ± 9	63 ± 10	0.6[a]
BMI, kg/m^2	26.3 ± 4.3	26.6 ± 7.0	0.8[a]
Cancer type			0.2[b]
Pancreas, n	19	19	
Colorectal, n	29	17	
Body composition variable			
% Weight loss	-3.44 ± 8.8	-5.50 ± 9.4	0.3[a]
Skeletal Muscle Index, cm^2/m^2	48.2 ± 7.8	39.1 ± 8.1	<0.0001[a]
Skeletal Muscle Radiodensity, HU	36.0 ± 9.2	28.5 ± 6.8	<0.0001[a]
Total Adipose Tissue, cm^2	311.3 ± 135.0	358.3 ± 188.9	0.3[c]
Fat Mass Index, cm^2/m^2	103.0 ± 44.7	141.5 ± 80.6	0.02[c]
Survival analysis			
Median OS survival (days)	1293 ± 343	764 ± 59	0.08[d]

Data presented as mean ± standard deviation.
[a] Independent t-test for normally distributed variables
[b] Pearson chi-square test for cancer site
[c] Mann-Whitney U test for non-normally distributed variables.
[d] Log-rank test.

2.3.2 Challenges using clinical labels for the identification of molecular markers

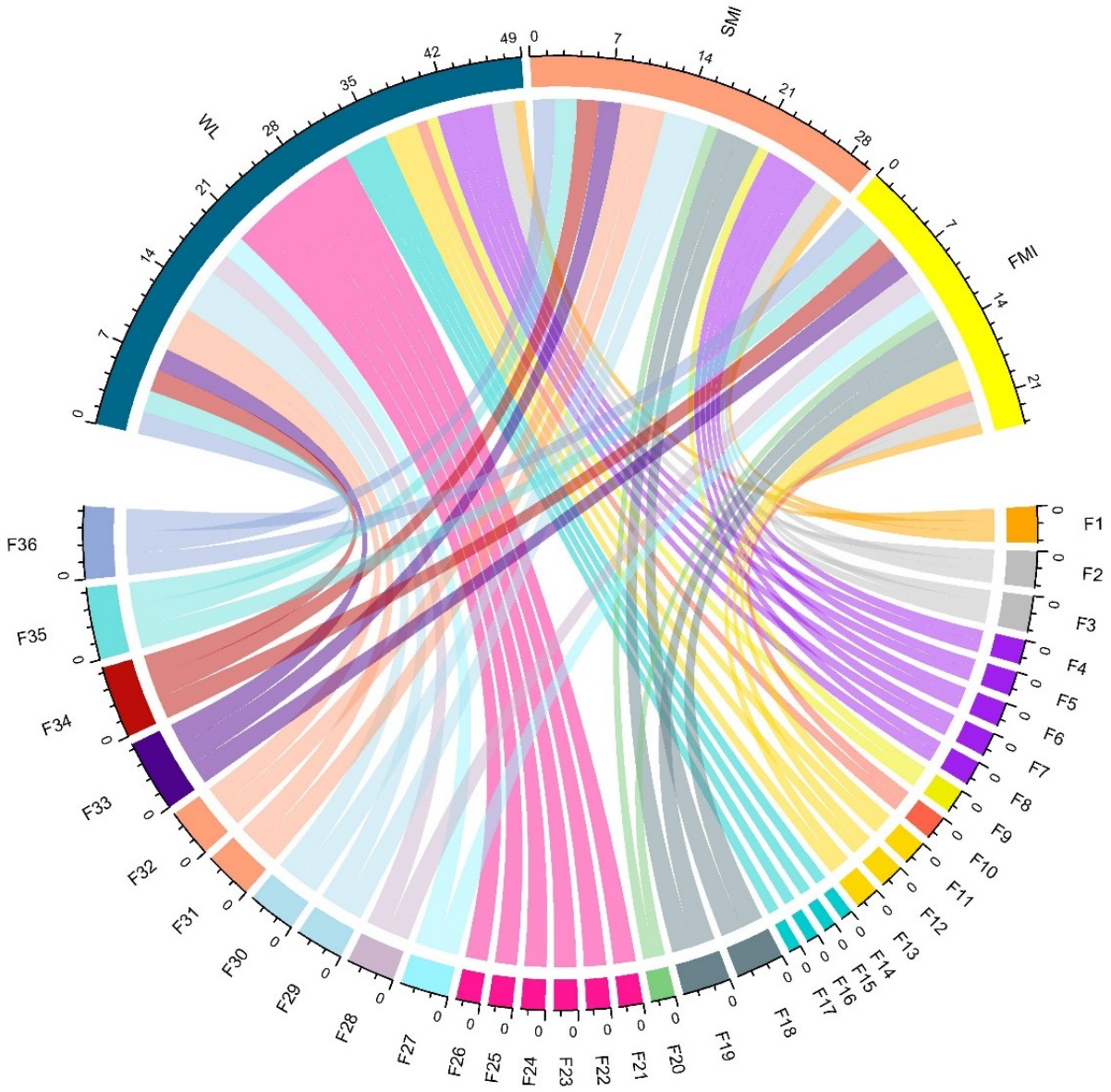
Patient-level heterogeneity from clinical labels limits their applicability to molecular characterization of skeletal muscle from patients with cancer. Inter-individual variations in the clinical labels were observed. Sex-specific median bivariate cut-offs were generated for clinical labels. However, the analyses resulted in disparities in conventionally defined parameters for

patient stratification, as represented in **Figure 2.1 and Figure 2.2**. For example, if one patient was classified as a case defined by low skeletal muscle mass, the same patient was classified as a control if the criterion was determined by weight stability. This disparity in classifying patients as cachexic or non-cachexic based on specific parameters of interest when used for molecular association studies (for example, gene expression analyses) could result in some overlap between the gene signatures. Hence, these analyses suggest using clinical labels cautiously, particularly when considering molecular profiling studies.



Sex	Category	%WL	SMI	FMI
Male	High	≥ -2	≥ 47.1	≥ 97.1
	Low	< -2	< 47.1	< 97.1

Figure 2.1 Heterogeneity of phenotypes in patients with cancer depicted as a chord diagram
Sex-specific median bivariate cut-offs for this cohort were generated, as shown in the table depicted below the chord diagram. The right end of the sample spectrum (from M1) and the left end of the sample spectrum (M48) represent the highest and the lowest median measures, respectively, of %weight, SMI, and FMI groups (represented in various colored arcs shown in scale as degrees).



Sex	Category	%WL	SMI	FMI
Females	High	≥ -3	≥ 38.0	≥ 123.3
	Low	< -3	< 38.0	< 123.3

Figure 2.2 Heterogeneity of phenotypes in patients with cancer depicted as a chord diagram
Sex-specific median bivariate cut-offs for this cohort were generated, as shown in the table depicted below the chord diagram. The right end of the sample spectrum (from F1) and the left end of the sample spectrum (F36) represent the highest and the lowest median measures, respectively, of %weight, SMI, and FMI groups (represented in various colored arcs shown in scale as degrees).

According to the above criteria, it is evident that the patient stratification presents a wide phenotypic variation within the individual patient, precluding the approach to classical association studies based on a single defined phenotype. For instance, an individual patient exhibiting different combinations of the phenotypic spectrum based on WL, SMI, or FMI (proxy to BMI) poses challenges to identifying individual RNA classes in the association analysis. This precludes integrative analysis and interpretation for a holistic understanding of muscle biology. Such complexities in the heterogeneity of skeletal muscle from cancer patients necessitate an unbiased approach independent of clinical labels to stratify patients based on transcriptome profiles. For example, the patient classified as a case based on percent %WL could represent a control if the criterion considered Skeletal Muscle Index (SMI) or Fat Mass Index (and vice versa) in the association study design. As such, the cut-offs adopted in literature for clinical labels are highly variable, lack clear consensus, and are often based on investigator-defined criteria for patient stratification.

2.4 Discussion

Classification criteria for molecular profiling studies are heterogenous and have impeded the progress in collating, replicating and validating the findings across the molecular studies. There is no concordance in implementing a specific classification criterion based on cachexia parameters such as % weight changes or SMI.

This investigation examined the interindividual variations and disparities at the individual patient level for phenotypes of weight loss, SMI, or FMI. My study suggests re-evaluating the criteria established from clinical cachexia parameters and implementing an unbiased and systematic methodology to decode the mechanisms underlying the pathophysiology of cancer-associated muscle wasting, particularly for molecular association studies.

Currently, there is no preferred animal model to study human cachexia [22-24], and the findings across human studies are also not encouraging, owing to the discrepancies in the array of cancer diagnoses in sample collection, low sample size, aggregate results for males and females, and classification schemas. Johns et al. (2014) analyzed muscle fiber size, biochemical composition, and markers for select pathways, including autophagy, SMAD signaling, and inflammation, using four classification schemas based on > 5% WL, >10% WL, lean muscle mass (LM), and LM+2%WL. They found that muscle fiber size, biochemical composition, and markers for representative pathways varied based on the classification strategy [25]. Ramge et al. (2016) used rectus abdominis muscle biopsies from n=32 patients with cancer. They analyzed protein content which, when correlated with weight loss or CT-derived measure of muscle mass, SMI or radiodensity, SMD, the protein content varied widely [26]. The criterion for %WL has also been debated as it masks the skeletal muscle depletion occurring within these cohorts of patients and should be used with caution when the study design involves skeletal muscle biopsy from patients with cancer [27].

Cancer-associated skeletal muscle wasting is an intricate and multifaceted syndrome. Identification of molecular markers that could potentially help in identifying cachexia in patients with cancer is imperative. A pre-requisite to it is the application of unsupervised methods to ascertain patterns based on the expression profiling of patients with cancer. Such broad-scale strategies were investigated in diverse cancer settings. However, these need to be determined in human cachexia research to facilitate an understanding of muscle wasting in patients with cancer. The robust classification system to stratify patients or alternative pattern recognition techniques could potentially enhance replicability and validity in identifying biomarkers for putative preclinical and clinical trials. The advantage of applying unsupervised learning techniques is that they are independent of predefined phenotypic stratifications and can discover inherent patterns

within the data. The expression analysis of coding and non-coding RNAs for each clinical label is infeasible since permutations and combinations of association analyses would limit the integrative analyses and the interpretations. For instance, if three clinical labels were individually assessed for the sex-specific expression profiles for coding and non-coding RNAs (total RNA type, n=6), then the resulting association studies that would need to be performed would be n=36 (n=3x6=18 studies, individually for males and females). This concept would potentially affect the aggregation and interpretations based on individual RNA types and phenotypes analyzed. Integrative analyses are therefore necessary to understand RNA crosstalk.

2.5 Conclusion

My study is the first to identify inter- and intra-patient-level heterogeneity for the clinical labels used for molecular studies and suggests their use with caution. The clinical labels showed little to no overlap across patient groups in this cohort of patients undergoing tumor resection and using biopsy collection for expression profiling. Therefore these clinical labels in a muscle biopsy study may not fully reflect underlying changes in the muscle of patients with cancer. There is an urgent need to address this challenge with an unsupervised machine learning approach to stratify patients for molecular association studies.

2.6 References

1. Fearon K, Strasser F, Anker SD, Bosaeus I, Bruera E, Fainsinger RL, et al. Definition and classification of cancer cachexia: an international consensus. *Lancet Oncol.* 2011;12(5):489-495.
2. Prado CM, Lieffers JR, McCargar LJ, Reiman T, Sawyer MB, Martin L, et al. Prevalence and clinical implications of sarcopenic obesity in patients with solid tumours of the respiratory and gastrointestinal tracts: a population-based study. *Lancet Oncol.* 2008;9(7):629-635.
3. Martin L, Birdsell L, Macdonald N, Reiman T, Clandinin MT, McCargar LJ, et al. Cancer cachexia in the age of obesity: skeletal muscle depletion is a powerful prognostic factor, independent of body mass index. *J Clin Oncol.* 2013;31(12):1539-1547.
4. Gallagher D, Ruts E, Visser M, Heshka S, Baumgartner RN, Wang J, et al. Weight stability masks sarcopenia in elderly men and women. *Am J Physiol Endocrinol Metab.* 2000;279(2):E366-375.
5. Mourtzakis M, Prado CM, Lieffers JR, Reiman T, McCargar LJ, Baracos VE. A practical and precise approach to quantification of body composition in cancer patients using computed tomography images acquired during routine care. *Appl Physiol Nutr Metab.* 2008;33(5):997-1006.
6. Shen W, Punyanitya M, Wang Z, Gallagher D, St-Onge MP, Albu J, et al. Total body skeletal muscle and adipose tissue volumes: estimation from a single abdominal cross-sectional image. *J Appl Physiol (1985).* 2004;97(6):2333-2338.

7. Shen W, Punyanitya M, Wang Z, Gallagher D, St-Onge MP, Albu J, et al. Visceral adipose tissue: relations between single-slice areas and total volume. *Am J Clin Nutr.* 2004;80(2):271-278.

8. Lieffers JR, Bathe OF, Fassbender K, Winget M, Baracos VE. Sarcopenia is associated with postoperative infection and delayed recovery from colorectal cancer resection surgery. *Br J Cancer.* 2012;107(6):931-936.

9. Prado CM, Sawyer MB, Ghosh S, Lieffers JR, Esfandiari N, Antoun S, et al. Central tenet of cancer cachexia therapy: do patients with advanced cancer have exploitable anabolic potential? *Am J Clin Nutr.* 2013;98(4):1012-1019.

10. Kazemi-Bajestani SM, Mazurak VC, Baracos V. Computed tomography-defined muscle and fat wasting are associated with cancer clinical outcomes. *Semin Cell Dev Biol.* 2016;54:2-10.

11. Baracos VE, Reiman T, Mourtzakis M, Gioulbasanis I, Antoun S. Body composition in patients with non-small cell lung cancer: a contemporary view of cancer cachexia with the use of computed tomography image analysis. *Am J Clin Nutr.* 2010;91(4):1133S-1137S.

12. Anoveros-Barrera A, Bhullar AS, Stretch C, Esfandiari N, Dunichand-Hoedl AR, Martins KJB, et al. Clinical and biological characterization of skeletal muscle tissue biopsies of surgical cancer patients. *J Cachexia Sarcopenia Muscle.* 2019;10(6):1356-1377.

13. Narasimhan A, Greiner R, Bathe OF, Baracos V, Damaraju S. Differentially expressed alternatively spliced genes in skeletal muscle from cancer patients with cachexia. *J Cachexia Sarcopenia Muscle.* 2018;9(1):60-70.

14. Narasimhan A, Zhong X, Au EP, Ceppa EP, Nakeeb A, House MG, et al. Profiling of Adipose and Skeletal Muscle in Human Pancreatic Cancer Cachexia Reveals Distinct Gene Profiles with Convergent Pathways. *Cancers (Basel).* 2021;13(8).

15. Judge SM, Nosacka RL, Delitto D, Gerber MH, Cameron ME, Trevino JG, et al. Skeletal Muscle Fibrosis in Pancreatic Cancer Patients with Respect to Survival. *JNCI Cancer Spectr.* 2018;2(3):pky043.

16. Stretch C, Khan S, Asgarian N, Eisner R, Vaisipour S, Damaraju S, et al. Effects of sample size on differential gene expression, rank order and prediction accuracy of a gene signature. *PLoS One.* 2013;8(6):e65380.

17. Stephens NA, Gallagher IJ, Rooyackers O, Skipworth RJ, Tan BH, Marstrand T, et al. Using transcriptomics to identify and validate novel biomarkers of human skeletal muscle cancer cachexia. *Genome Med.* 2010;2(1):1.

18. Gallagher IJ, Stephens NA, MacDonald AJ, Skipworth RJ, Husi H, Greig CA, et al. Suppression of skeletal muscle turnover in cancer cachexia: evidence from the transcriptome in sequential human muscle biopsies. *Clin Cancer Res.* 2012;18(10):2817-2827.

19. Narasimhan A, Ghosh S, Stretch C, Greiner R, Bathe OF, Baracos V, et al. Small RNAome profiling from human skeletal muscle: novel miRNAs and their targets associated with cancer cachexia. *J Cachexia Sarcopenia Muscle.* 2017;8(3):405-416.

20. van de Worp W, Schols A, Dingemans AC, Op den Kamp CMH, Degens J, Kelders M, et al. Identification of microRNAs in skeletal muscle associated with lung cancer cachexia. *J Cachexia Sarcopenia Muscle.* 2020;11(2):452-463.

21. Martin L, Senesse P, Gioulbasanis I, Antoun S, Bozzetti F, Deans C, et al. Diagnostic criteria for the classification of cancer-associated weight loss. *J Clin Oncol.* 2015;33(1):90-99.

22. Tomasin R, Martin A, Cominetti MR. Metastasis and cachexia: alongside in clinics, but not so in animal models. *J Cachexia Sarcopenia Muscle.* 2019;10(6):1183-1194.

23. Baracos VE. Bridging the gap: are animal models consistent with clinical cancer cachexia? *Nat Rev Clin Oncol.* 2018;15(4):197-198.

24. Mueller TC, Bachmann J, Prokopchuk O, Friess H, Martignoni ME. Molecular pathways leading to loss of skeletal muscle mass in cancer cachexia--can findings from animal models be translated to humans? *BMC Cancer.* 2016;16:75.

25. Johns N, Hatakeyama S, Stephens NA, Degen M, Degen S, Frieauff W, et al. Clinical classification of cancer cachexia: phenotypic correlates in human skeletal muscle. *PLoS One.* 2014;9(1):e83618.

26. Ramage MI, Johns N, Deans CDA, Ross JA, Preston T, Skipworth RJE, et al. The relationship between muscle protein content and CT-derived muscle radio-density in patients with upper GI cancer. *Clin Nutr.* 2018;37(2):752-754.

27. Roeland EJ, Ma JD, Nelson SH, Seibert T, Heavey S, Revta C, et al. Weight loss versus muscle loss: re-evaluating inclusion criteria for future cancer cachexia interventional trials. *Support Care Cancer.* 2017;25(2):365-369.

Chapter 3 Whole transcriptome profiling of human skeletal muscle and unsupervised clustering to identify molecular subtypes in cancer-associated muscle wasting

3.1 Introduction

Muscle wasting, with or without the loss of fat mass, is one of the hallmarks of cachexia [1]. We have little if any, insight into the inherent biological characteristic of human muscle biology in patients with cancer. Understanding molecular mechanisms underlying the pathophysiology of cancer-associated muscle wasting is derived from experimental model systems, including in vivo and in vitro systems. These, however, have limited applicability to the human condition due to varied reasons such as (i) heterogeneity in cancer types and models studied, (ii) the rapid development of cachexia in animal models (typically timeframe of tumor development is only a few weeks), the tumors encompassing about 10% of the entire body mass, (iii) the commonly used animal models showed little overlap with the molecular signatures from human muscle biopsies [2,3]. A consensus from animal models of cachexia is that no experimental model exists to study cancer cachexia encompassing the effects of metastases [4], chemotherapy [5], and recapitulating human clinical cachexia condition. Anoveros and colleagues conducted a systematic review of n=59 muscle biopsy studies. They reported fundamental limitations across studies concerning heterogenous patient populations due to constraints in accessibility to muscle biopsies, sample size, aggregate sex in the documented association analysis, and varying classification criteria, amongst others [6]. Gene expression studies using human skeletal muscle biopsy from patients with cancer have to date, quantified candidate gene expression profiles to elucidate molecular mechanisms from a focussed lens. These studies have delineated the role of several pro-inflammatory mediators and pathways contributing to protein imbalance, muscle wasting,

regeneration, and autophagy [7-16]. Previous studies were limited to mRNA species, and the data generated was from microarray platforms. Further, the classification criteria varied and were subject to all crucial limitations explained in Chapter 2.

Advances in research paradigms in clinical cachexia setting to study human muscle biology are needed. The cancer surgery setting serves as the best resource to obtain the relevant tissue of interest. The tissue collection is considered minimally invasive, given that the tissue is attained simultaneously with the surgical tumor resection. Collaboration between cancer surgeons and researchers can enable molecular profiling in cachexia-susceptible cancers. Morphological and histological[17-19], fiber type composition[20], biochemical[20], protein[21], and mRNA expression profiling analysis[16,19,22-26] were conducted using the acquired muscle biopsies from the susceptible cachexia patient populations. The human genome comprises ~2-3% protein-coding genes, whereas an extensive repertoire of the transcribed genome is non-coding RNAs regulating gene expression at epigenetic, transcriptional, and post-transcriptional levels[27,28].

Studying whole transcriptional machinery is essential to understand muscle biology in patients with cancer since the signaling pathways target the transcriptional machinery. Advances in omics and computational biology provide tools to integrate different layers of information to understand gene expression regulation. mRNA profiling studies are sparse and heterogenous in patient populations. Studies in cancer n=6 only reviewed protein-coding (i.e., mRNA);[16,19,22-25] whereas n=2 studied miRNAs from skeletal muscle biopsies of patients with cancer[29,30]. Whole transcriptome profiling studies are needed to gain mechanistic insights and an understanding of pathophysiology from a vantage point of view.

Non-coding RNAs are essential regulators of gene expression and their dysregulation results in several diseases. miRNAs and lncRNAs are the most studied non-coding RNAs in the experimental model systems studying skeletal muscle biology [31]. However, their application in clinical samples from patients with cancer is limited to only the studies on miRNAs mentioned above. Integrated networks and the crosstalk mechanism between these RNAs contribute to gene regulation and may help delineate the processes involved in cancer-associated muscle wasting. These need to be studied in integrated analyses to understand the transcriptional changes in the skeletal muscle of patients with cancer. Therefore it is necessary to perform unbiased analyses for a molecular understanding of human skeletal muscle at the whole transcriptome level.

Cluster analyses, an unsupervised machine learning method, is a task of grouping a set of data objects similar to one another in one group or cluster than those in other groups or clusters. It can be used to classify patients based on their expression profiles, help delineate patterns from complex high-dimensional datasets and aid in identifying Complex Biological Processes (CBPs). Non-negative Matrix Factorization (NMF) is an unsupervised learning method that was first applied in facial recognition, signal processing, and text mining [32,33]. NMF is a matrix decomposition approach that decomposes a non-negative matrix into lower-rank non-negative matrices. Brunet et al. were among the first to expand NMF application to gene expression studies. They applied the NMF algorithm to identify metagenes (a set of genes behaving in a functionally correlated manner within the genome) using four independent cancer datasets [34]. The challenges inherent to high-dimensional datasets are that the underlying biological processes are hidden (when no a priori hypothesis is presented), and the unsupervised clustering algorithms allow data dimensionality reduction and deconvolution of complex biological data into potentially interpretable data entities. I used integrative NMF (intNMF) to identify the clusters from NGS-generated whole

transcriptome data for human skeletal muscle mass for males and females, respectively. The identified clusters in skeletal muscle going forward are referred to as molecular subtypes, or subtypes throughout the thesis, analogous to the molecular classifications of tumors from transcriptome datasets reported in the literature.

Several implementations of the NMF algorithm have been proposed and have revolutionalized genomics [35-39]. Details on the methodology are described in chapter 1. An unbiased approach to identifying molecular subtypes of human skeletal muscle from patients with cancer has never been undertaken. The inherent variability of the data is unknown; therefore, grouping data according to transcriptional patterns is imperative to gain insights into muscle biology, which thus far has never been attempted in cachexia research.

Investigations employing the human skeletal muscle to understand pathophysiological mechanisms of cancer-associated muscle wasting have not executed sex-related expression analyses. Human skeletal muscle is sexually dimorphic concerning patterns of gene expression, fiber type specifications, fatigue resistance, muscle strength, mitochondrial function, muscle protein turnover, and intrinsic signaling within the muscle microenvironment [40-46]. Cachexia research relating to human muscle biopsies from patients with cancer has presented aggregate results or has not accounted for sex as a potential confounding variable [6]. There have been no comprehensive studies to date that have addressed sex-specific differences using whole transcriptome datasets.

I hypothesize that the whole transcriptome dataset generated from the human skeletal muscle of patients with cancer, when subjected to unsupervised clustering, would identify molecular subtypes. The dysregulated gene expression profiles from the subtypes could explain sexually

dimorphic pathways in the muscle of patients with cancer. The objectives of this study are to (i) accrue human skeletal muscle biopsies from patients with cancer, (ii) perform whole transcriptome Next-Generation RNA sequencing and small RNA sequencing on the obtained muscle biopsies, (iii) utilize machine learning-based unsupervised clustering methods to identify sex-specific molecular subtypes of human skeletal muscle (iv) to implement clinical benchmark (clinical relevance) analyses of subtypes (iv) perform Differential Expression (DE) of RNAs within these subtypes (v) identify distinct and common canonical pathways in males and females.

3.2 Methods[§]

The overall study design is shown in **Figure 3.1,** represented in two panels. The top panel A depicts the acquisition of skeletal muscle biopsy and patient inclusion criteria. The bottom panel B illustrates the flow diagram (left to the right arrow) of the body composition assessment using CT image analysis, clinical data acquisition, and bio profiling methodology steps. The schematic flow diagram for methodological data pre-processing and clustering are outlined in Figures 3.2 and 3.3, respectively.

[§] Study was conceptualized by Bhumi Bhatt, Dr. Vickie Baracos and Dr. Sambasivarao Damaraju

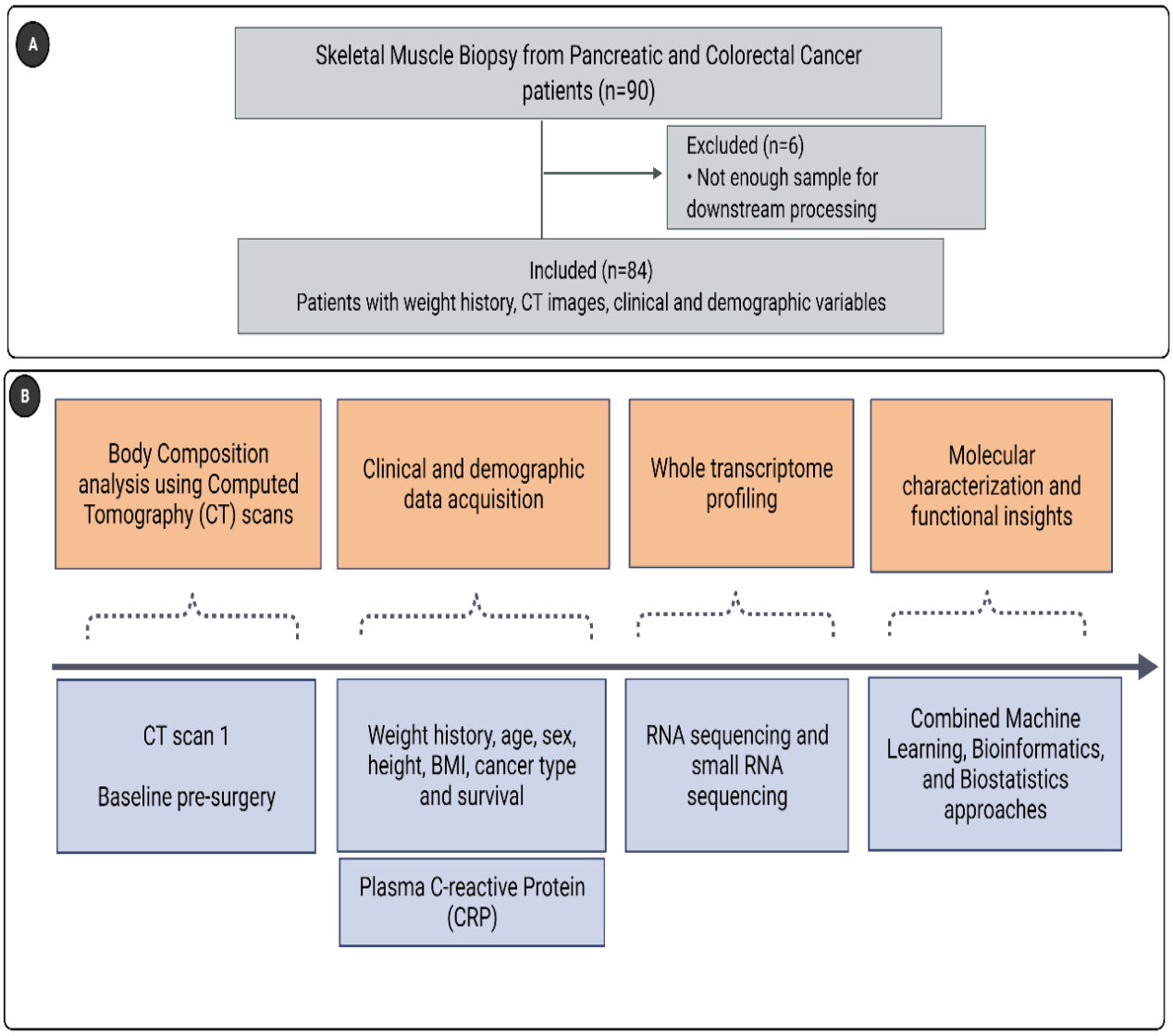


Figure 3.1 Schematic of the study design
Panel A depicts the exclusion and inclusion criteria of the study cohort for the bio profiling of RNAome; Panel B depicts the component data acquisition and downstream analysis modules (left to the right arrow).

3.2.1 Accrual of human skeletal muscle biopsies

Rectus abdominis muscle biopsies[**] of pancreatic and colorectal patients with cancer with liver metastasis were procured from the regional Hepatopancreaticobiliary/Gastrointestinal Tumor Bank (University of Calgary, Alberta, Canada). Patients had undergone laparotomy at the Foothills

[**] Skeletal muscle biopsies provided by Dr. Oliver Bathe.

Hospital, Calgary, from 2006 to 2015. Standard procedures were followed for tissue procurement and storage. The operating surgeon (Dr. Oliver Bathe, Surgical Oncologist) took the biopsy specimen within 30 mins of the start of the surgery using sharp dissection, immediately flash-frozen in liquid nitrogen to minimize ischemic shock post-devitalization, and stored it at -80°C until further use. The study participants provided written informed consent and were approved by the Conjoint Health Research Ethics Board at the University of Calgary (Ethics ID: E17213). Health Research Ethics Board of Alberta (HREBA) – Cancer Committee approved the current study protocol for whole transcriptome profiling and access to the patient's clinical information (protocol number ETH21709).

3.2.2 Next-Generation Sequencing (NGS) profiling of human skeletal muscle from patients with cancer

RNA extraction was performed using the Trizol method and Qiagen RNAeasy midi kit (Mississauga, ON). The optical density (OD) 260/280 ratio was measured using Nanodrop, and RNA integrity number (RIN) was assessed using Agilent Bio-analyzer 2100 for all the samples.

Services from the Genome Quebec facility (Montreal, Canada) were utilized for library preparation and whole transcriptome sequencing of RNA from human skeletal muscle. Briefly, pre-processing of isolated total RNA and downstream processing of the samples were as per the manufacturer's instructions summarized below:

a rRNA depleted sequencing (RNA sequencing)

Total RNA was quantified, and its integrity (RNA integrity Number, RIN) was assessed using 5K / RNA / Charge Variant Assay LabChip and RNA Assay Reagent Kit (Perkin Elmer). All samples had RIN values above 5 (90% of samples with RIN >6), and one sample had at RIN 4.3. The RNA quality was based on RIN number for paired-end sequencing using rRNA-depleted libraries for

RNASeq, and small RNA sequencing was shown to have no dependence on RIN number (except for poly-A enriched RNA library preparations). Quality indices were consistently obtained and surpassed the recommended thresholds for assessing library and sequence quality scores. rRNA was depleted from 250 ng of total RNA using QIAseq FastSelect (Human 96rxns). New England BioLabs (NEB) provided the following reagents and kits, including adapters and primers for the cDNA synthesis: NEBNext RNA First-Strand Synthesis and NEBNext Ultra Directional RNA Second Strand Synthesis Modules were used. The remaining steps of the library preparation were performed using NEBNext Ultra II DNA Library Prep Kit for Illumina. Libraries were quantified using the Quant-iT™ PicoGreen® dsDNA Assay Kit (Life Technologies) and the Kapa Illumina GA with Revised Primers-SYBR Fast Universal kit (Kapa Biosystems). The average size fragment was determined using a LabChip GX (PerkinElmer) instrument.

The libraries were normalized, pooled, and then denatured in 0.05N NaOH and neutralized using HT1 buffer. The pooled libraries were loaded at 225pM on an Illumina NovaSeq S4 lane using Xp protocol per the manufacturer's recommendations. The run was performed for 2x100 cycles (paired-end mode). A phiX library was used as a control and mixed with libraries at 1% level. All samples passed a Phred Quality Score of >30, as per the manufacturer's recommendations. Base-calling was performed with RTA v3.4.4. Program bcl2fastq2 v2.20 was then used to demultiplex samples and generate fastq reads.

b Small RNA sequencing (Small RNA seq)

Total RNA was quantified, and its integrity (RIN) was assessed as described above for the RNASeq protocol. Libraries were generated from 300 to 800ng of total RNA using the NEBNext Multiplex Small RNA Library Prep Kit for Illumina, as per the manufacturer's recommendations. cDNA construct purification has been performed using SparQ beads (Qiagen). Final libraries were

quantified using the Quant-iT™ PicoGreen® dsDNA Assay Kit (Life Technologies) and the Kapa Illumina GA with Revised Primers-SYBR Fast Universal kit (Kapa Biosystems). The average size fragment was determined using a LabChip GX (PerkinElmer) instrument.

The libraries were normalized and pooled and then denatured in 0.05N NaOH and neutralized using HT1 buffer. The pool was loaded at 225pM on an Illumina NovaSeq SP lane using Xp protocol as per the manufacturer's recommendations. Downstream steps in the sequencing protocol are the same as adopted for RNASeq, including base calling, demultiplexing, and generating fastq reads.

3.2.3 Analysis of raw sequence files and Differential Expression (DE) of RNAs††

The data analyses of raw fastq sequence files were performed using Partek Flow software v10.0.21.0929 (Copyright ©; 2018 Partek Inc., St. Louis, MO, USA) unless specified otherwise. The raw fastq files were subjected to Cutadapt [47] for the 3' adapter trimming. The trimmed reads were then aligned to Human Genome (reference index hg38) using STAR aligner [48] (v2.7.3a) and Bowtie [49] (v.2.2.5) for RNA seq and small RNA seq data, respectively. The generated .bam files were quantified to transcriptome using RNAs obtained from different annotation databases: Ensembl transcripts v102 was used to quantify lncRNA and mRNA, miRNA using miRbase v22[50], snoRNA using snoDB [51] v1.2.0, piRNA using piRNAdb v1.8.0 [52], and tRNA using GtRNAdb[53]. The features (the terminology is used interchangeably referring to RNAs) were filtered for ten read counts in 90% of samples following quantification of the reads. Differential Expression (DE) analysis of the RNAs was performed using the DESeq2 [54] R package. RNAs were considered to be DE at a Fold-change cut-off of 1.5 and a nominal $P < 0.05$ (unadjusted), and a False Discovery Rate (FDR, adjusted) $P<0.05$.

†† Bioinformatic analyses were performed by Bhumi Bhatt.

3.2.4 Non-negative Matrix Factorization (NMF) clustering for single RNA species[‡‡]

Non-negative Matrix Factorization (NMF) is an unsupervised machine learning method conventionally used in molecular tumor profiling studies to identify clusters (groups) from multidimensional data and to uncover Complex Biological Processes (CBPs) [34,55]. The [gene x sample] expression matrix from the whole transcriptome datasets was generated and used as input for the algorithm. Data pre-processing steps are shown, and the overarching flow diagram is described in **Figure 3.2.** Briefly, the raw read counts were filtered, and feature selection was applied to filter out features that did not vary across the samples but selected only those that showed the most variance. The analysis was performed using the R package NMF [56]. Variance Stabilized Transformed (vst) counts [gene x sample] matrix was subjected to the downstream analysis to identify molecular subtypes using the methods of Kullback-Leibler (KL) divergence and Euclidean. The algorithm was run for 30 iterations and fixed initializing seed parameters (with the predefined lower and upper limit of k clusters as 2:4) to estimate the optimal number of clusters. After ascertaining the number of clusters, the algorithm was run for 200 iterations to arrive at a consensus number of clusters using the same parameters above. The quality of the clusters and the optimal number of clusters were assessed using cophenetic coefficient (Ω), consensus maps (or connectivity matrix plots), Silhouette Width (SW), and the considerations of study sample size within the clusters. Rand Index (RI) scores, a metric of the similarity between the clustering results from two different RNA classes, were used to determine the similarity between the cluster annotations within the individual RNA class. I computed the RI using the R package Aricode. Rand Index [57] is calculated as follows:

[‡‡] Machine learning analyses were performed by Bhumi Bhatt.

$$\boldsymbol{R} = \frac{\boldsymbol{a+b}}{\boldsymbol{a+b+c+d}} = \frac{\boldsymbol{a+b}}{\binom{\boldsymbol{n}}{\boldsymbol{2}}}$$

a = number of pairs of clusters identified from the NMF clustering algorithm and the pairwise comparisons of sample identities being the same in a given set of RNA classes

b = number of pairs of clusters identified from the NMF clustering algorithm and the pairwise comparisons of sample identities being different in a given set of RNA classes

c = number of pairs of clusters identified from the NMF clustering algorithm and the pairwise comparisons of sample identities being the same in RNA class 1 and different in RNA class 2

d = number of pairs of clusters identified from NMF clustering algorithm and the pairwise comparisons of sample identities being different in RNA class 1 and same in RNA class

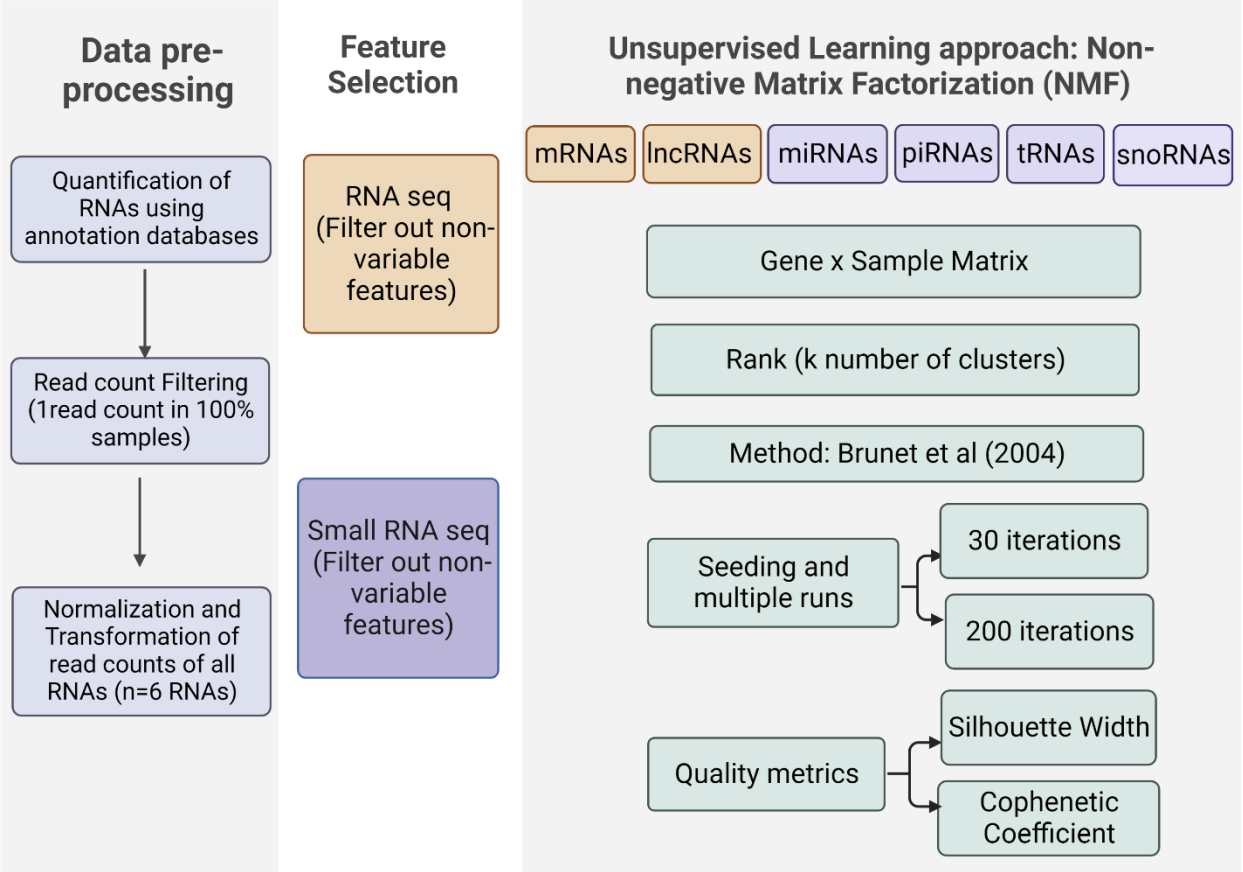


Figure 3.2 Flow-diagram for identifying molecular subtypes from single RNA species using Non-negative Matrix Factorization (NMF)
Individual RNAs (mRNA, lncRNA, miRNA, piRNA, tRNA, and snoRNA) were subjected to data pre-processing steps, feature selection, and clustering using the NMF algorithm to identify consensus clusters for each RNA class (each RNA class boxed in the figure). Data pre-processing steps from the raw sequence data to normalized counts are shown on the **left panel**, feature selection is shown in the **middle panel**, and workflow for NMF for individual RNA classes is shown in the **right panel.**

3.2.5 Integrative Non-negative Matrix Factorization (intNMF) clustering to identify molecular subtypes of human skeletal muscle

The flow diagram for identifying consensus molecular subtypes using the intNMF algorithm [58] is shown in **Figure 3.3**. Using one RNA class to identify the molecular subtype of muscle does not provide an intuitive understanding and holistic picture of the underlying pathophysiology. Integrative unsupervised approaches are therefore necessary. Sex-specific integrative analyses were performed for all RNA classes. The filtered, normalized, and transformed [sample x gene]

matrix was subjected to intNMF analysis using the R package intNMF. An integrative analysis was implemented, combining all the RNA classes to determine the molecular subtypes of human skeletal muscle. The intNMF algorithm was run at 100 iterations using a random initializing seed parameter of an optimal number of clusters, 2-4 (k=2:4), with 5-fold cross-validation and Euclidean as a distance metric. The optimal number of clusters was selected based on the considerations of study sample size within the clusters, Cluster Prediction Index (CPI), and SW as quality indices.

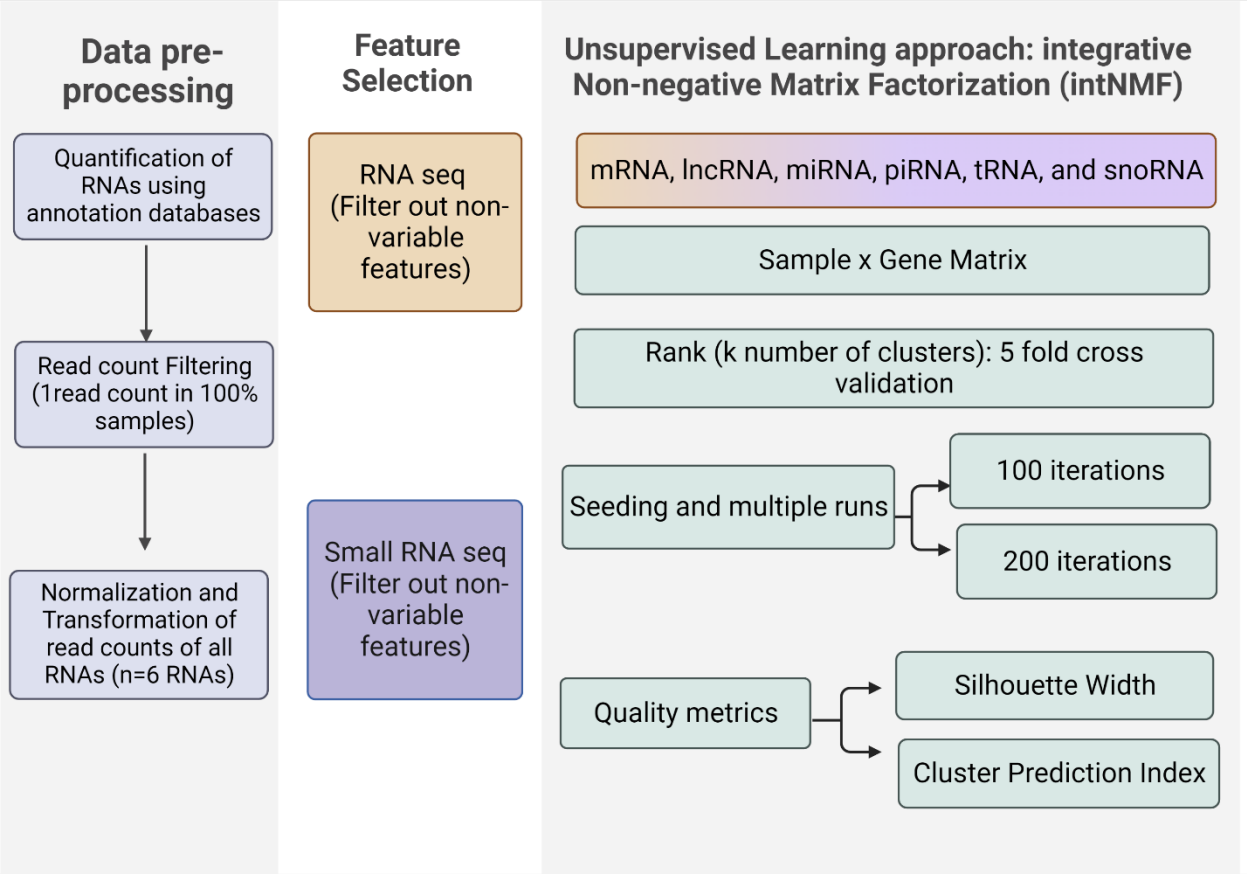


Figure 3.3 Flow diagram for identifying molecular subtypes using integrative Non-negative Matrix Factorization (intNMF)
The whole transcriptome normalized and transformed read count data (mRNA, lncRNA, miRNA, piRNA, tRNA, and snoRNA, all RNAs considered together and marked within a single box in the figure) were subjected to data pre-processing steps, feature selection, and intNMF algorithm to identify consensus and coherent molecular subtypes of human skeletal muscle.

3.2.6 Statistical analysis

Statistical analyses were performed using Rstudio and SPSS v27. Differences between groups were analyzed using an independent t-test for continuous variables and Pearson's χ^2 test for categorical variables. Non-normally distributed variables were compared using the Mann-Whitney U test. Correlations between continuous variables were assessed using Pearson correlation coefficients using Hmisc and corrplot R packages. Survival analysis was performed using the Kaplan-Meier method, and the differences were compared using log-rank tests. Statistical significance was determined at $P < 0.05$ unless otherwise specified within the text.

3.2.7 Functional insights into pathway-level regulation using Ingenuity Pathway Analysis (IPA)

Differentially Expressed (DE) genes were subjected to pathway analysis and functional annotation using Ingenuity Pathway Analysis [59](IPA,

QIAGEN Inc., https://www.qiagenbioinformatics.com/products/ingenuitypathway-analysis). P-value in IPA is calculated using right-tailed Fisher's exact test, and it tests the likelihood that the association or overlap between a set of molecules from the experimental data set and the associated pathway or process predicted from the IPA knowledge base is due to random chance alone. The smaller the p-value, the less likely that the association is random. -log(P-value) threshold of 1.3 (equivalent to a nominal p-value of 0.05) was used to define statistically significant pathways.

3.3 Results

3.3.1 Next-Generation Sequencing profiling of human skeletal muscle

The overall study design is shown in **Figure 3.1**. *Rectus abdominis* muscle biopsies (n=84) and each biopsy specimen served to generate small RNA and RNA sequencing and downstream data

pre-processing. Each patient's muscle biopsy specimen was used in this study to generate global transcriptome expression profiles. The RNA species studied were mRNA (protein-coding), small RNAs (<200 nt, miRNAs, snoRNAs, piRNAs, tRNAs), and long non-coding RNAs (>200 nt, lncRNAs). The expression landscape (normalized and transformed read counts) for all RNA species for males and females is summarized. No significant differences in the relative abundance of RNA species between males and females were observed, as shown in **Figure 3.4,** ruling out bias at the expression level (read counts) per se between the sexes. Amongst the RNA species, mRNAs were highly abundant, followed by:

snoRNAs > lncRNAs > tRNAs > miRNAs > piRNAs.

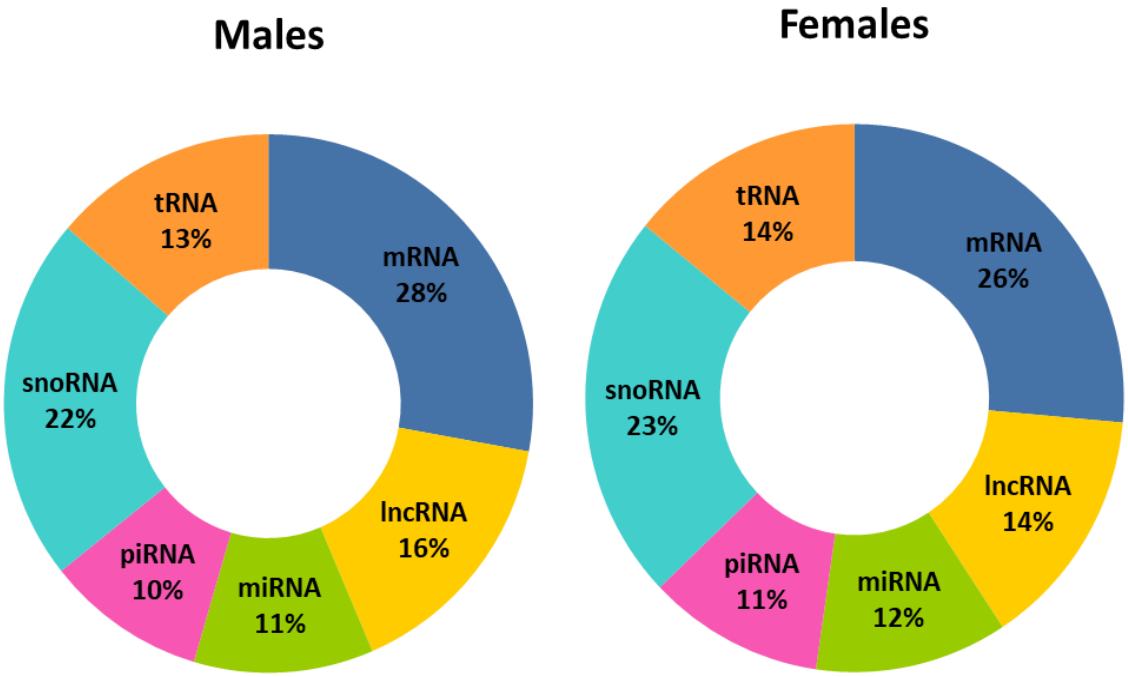


Figure 3.4 Relative abundance of RNA classes in human skeletal muscle from patients with cancer
A total of six RNAs were profiled from n=84 muscle biopsy specimens from patients with cancer, and their expression profile distribution (normalized read counts) is shown individually for males (left panel) and females (right panel), respectively. The relative distribution of RNAs within skeletal muscle expression profiles of males and females was not statistically significant (P-value for the test of proportions >0.8 for all RNAs between males and females).

3.3.2 Individual RNAs in clustering analyses using Non-negative Matrix Factorization (NMF)

RNAs identified from the human skeletal muscle of patients with cancer are described: (*Males*: 18586 mRNAs, 16204 lncRNAs, 479 mature miRNAs, 868 piRNAs, 608 snoRNAs, tRNAs 408; *Females*: 18548 mRNAs, 16183 lncRNAs, 565 mature miRNAs, 838 piRNAs, 604 snoRNAs, tRNAs 410). Individual RNA species were subjected to an unsupervised NMF algorithm to determine human skeletal muscle molecular subtypes. Read count filtering of 1 read count in 100% of samples was applied, and the features that did not vary across the patient samples were removed. The optimal number of clusters k=2 was selected for all the RNA classes based on the quality metrics (Silhouette width and cophenetic coefficient) and sample size considerations.

From the obtained two-cluster model, I compared the sample similarity between clusters using Rand Index (RI) derived from individual RNA classes **(Figure 3.5**). Up to 50% of the sample labels were identical in all NMF clusters, and the remaining sample labels were disseminated widely based on the class of RNA used for clustering. These results highlight that individual RNA expression profiles have the intrinsic ability to partition the data into unique and distinct clusters, and the lack of complete overlap of sample labels supports the earlier observations that clinical cachexia labels alone are insufficient to guide the clustering of molecular data.

Figure 3.5 Sample similarity matrix for NMF generated clusters computed using Rand Index (RI) RIs identify the sample similarity between the RNA classes and are shown individually for males (top panel) and females (bottom panel). RIs range from 0-1, with 1 being a 100% overlap of samples in the class comparisons. Pairwise RNA matrices were delineated separately for the sexes.

Note that in traditional association studies, samples were binarized as cases or controls based on the investigator-defined phenotypic criteria. The above sample similarity matrix analysis shows that some RNAs split the samples differently; males show an extensive sample overlap within the clusters driven by miRNA with piRNAs. In females, sample overlap was seen with miRNAs, piRNAs, and snoRNAs, whereas the sample overlap was 49-61% for other RNA classes and their pair-wise comparison.

3.3.3 Integrative NMF (intNMF) clustering of human skeletal muscle transcriptome

High dimensional data from profiling studies may capture a holistic view of the complex biological patterns in the health and diseased states. The integrated approach helps identify disease subtypes, potential biomarkers, and underlying pathophysiological insights. The whole transcriptome profiles with annotated RNA species are as follows: *Males:* 18586 mRNAs, 16204 lncRNAs, 479 mature miRNAs, 868 piRNAs, 608 snoRNAs, and 408 tRNAs. *Females:* 18548 mRNAs, 16183 lncRNAs, 565 mature miRNAs, 838 piRNAs, 604 snoRNAs, and 410 tRNAs; The intNMF clustering algorithm [58] was applied using a combination of all input RNA classes to classify patients into disjoint and coherent clusters based on global transcriptome expression profiles. A predefined cluster range K=2:5 was initially investigated with 100 iterations for each of the sexes and using 5-fold cross-validation. Silhouette width (SW) and Cluster Prediction Index (CPI) were quality metrics used to assess cluster solutions. k=2 cluster solution was consistent in all iterations when a high stringency of SW or CPI was considered. k=2 cluster solution was also consistent when sample considerations were adhered to maintain statistical rigor. In my analysis, the sample size for males (n=48) and females (n=36) align well with previous studies that reported a minimum sample size of n=36 as a prerequisite to confidently calling the optimum number of clusters. Given these considerations, males showed n=27 and n=21 samples in Cluster 1: Cluster 2, and females

showed n=18 samples in each cluster, respectively. Hereafter I use the terminology of clusters 1 and 2 synonymously with skeletal muscle molecular subtype 1 and subtype 2, respectively. The intNMF method encompassing the entire transcribed genome offers an unbiased stratification of patients, independent of clinical labels, with cancer from their muscle transcriptome, providing a premise for the association of the subtypes with differential expression of genes and an in-depth analysis of the subtypes for tissue-level regulation of genes.

3.3.4 Clinical benchmark analyses of the intNMF subtypes

I performed two independent analyses to understand the clinical relevance of the intNMF subtypes. I refer to them as clinical benchmark analyses or clinical benchmarking of the intNMF subtypes throughout the thesis document. I conducted the first association analysis of the subtypes using the WL-BMI grading system [60] and the second association analysis using the age- and sex-adjusted distribution of CT-defined body composition at the L3 level in the single largest population of patients with cancer [61]. The analysis methods are described in the subsequent sections.

3.3.5 Association of clusters with the (a) WL-BMI grades and (b) ZSMI

One of the first steps to understanding the inherent characteristics of clusters was to identify their association with the WL-BMI grading schema proposed by Martin et al. (2015) [60] and the ZSMI distribution based on literature-reported ZSMI cutoffs.[61]. Of note, in the low-grade cluster (cluster 2) in both male (n=4) and female patients (n=4), total n=8 patients showed SMI values at or above those reported for healthy 30-year old kidney donors in male and female subjects, respectively [62]. There were no such observations for subtype 1. An interesting observation from this analysis was that although 10% of patients in cluster 2 (i.e., the less severe group related to cachexia severity) had high WLBMI grades (**Figure 3.6**), they had higher muscle mass (i.e., from ZSMI values). Similar trends also followed for the same cluster 2 wherein, although patients belonged to low

muscle mass distribution, they were in the stable weight category. This analysis output reinstates using unsupervised methods to stratify patients for molecular association studies in lieu of heterogenous clinical labels.

Despite the similarities or differences in the proportional association of subtypes with WL-BMI grades and z-SMI distribution, the overall take-home message is that subtype 1 is a relatively high-risk group affected by cachexia compared to subtype 2. The given comparisons could be used as a working hypothesis for lack of a better metric to annotate the subtypes. This is corroborated by an independent experimental approach to compare the subtypes with model systems referred to as functional benchmarking, described in chapter 5. It is not to say that one approach is better and adequate. The overall consensus is reached from all of the above analytical methods. In this context, one should recall that the use of ML is to decipher subtypes in an unbiased manner, independent of clinical labels. Therefore, the benchmark analyses described in this thesis attempt to understand the subtype characteristics.

The association analysis results of samples within the clusters to the WL-BMI grades and ZSMI is represented in **Figure 3.6** for males and **Figure 3.7** for females. The aggregate analysis performed by combining all samples (n=84) is represented in **Figure 3.8**, respectively. The association statistics of the clusters for males, females, and aggregate samples (combined for males + females) are presented in **Table 3.1**. Statistical significance was ascertained using Fisher's exact test. * represents statistical significance at P-value < 0.05 and ** represents P-value < 0.001. Significant differences were observed between the assignment of two clusters based on the grading system and ZSMI distribution, with the high-grade cluster showing a higher proportional association with high WL and low muscle mass in males and females, as well as for the combined sample analyses.

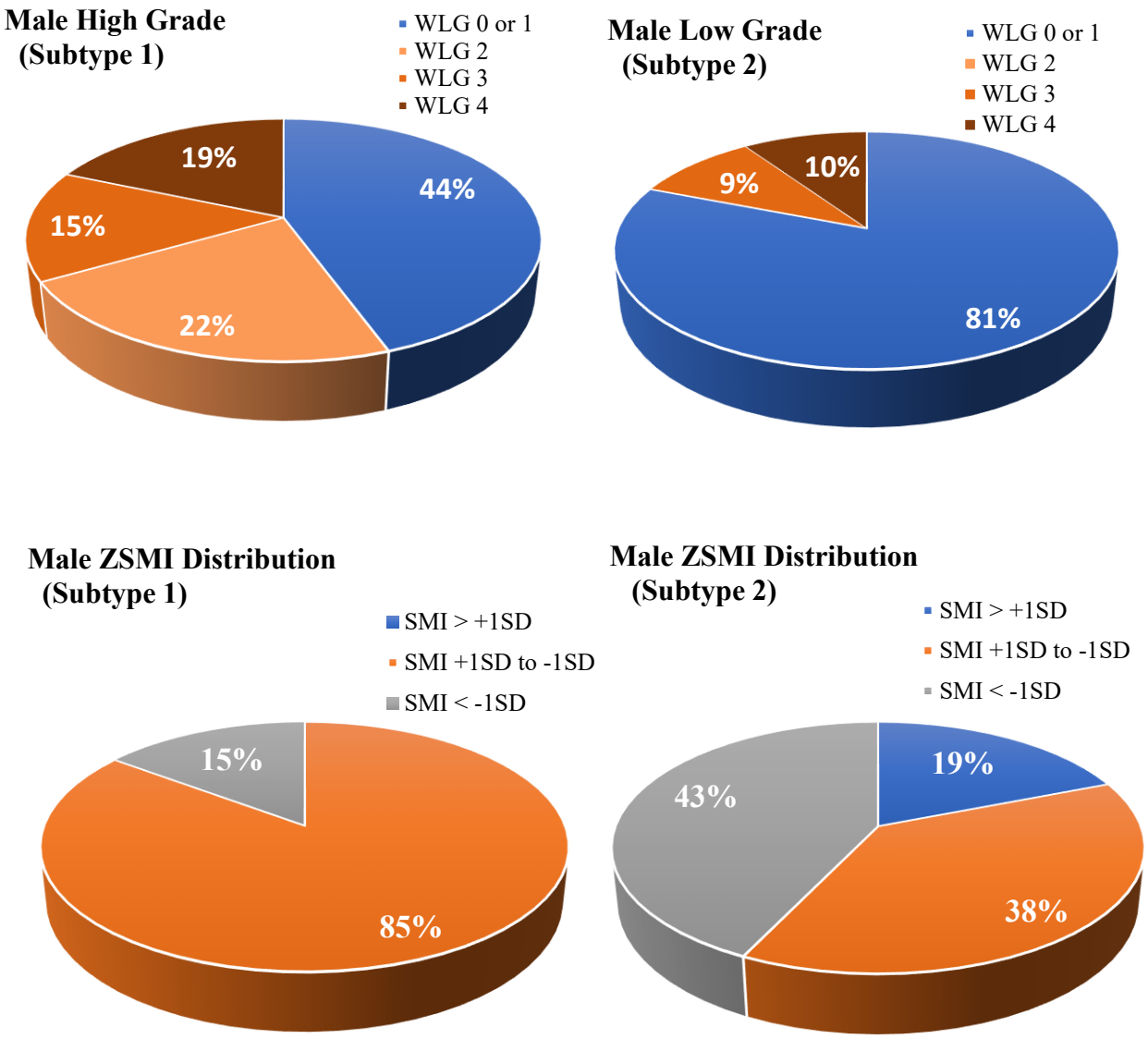


Figure 3.6 Clinical benchmark analyses for males.
The pie charts depict the distribution of samples within the clusters based on the WL-BMI grades and z-SMI distribution for the subtypes in males. High and Low-Grade labels correspond to the WL-BMI classification schema for the associated skeletal muscle subtypes.

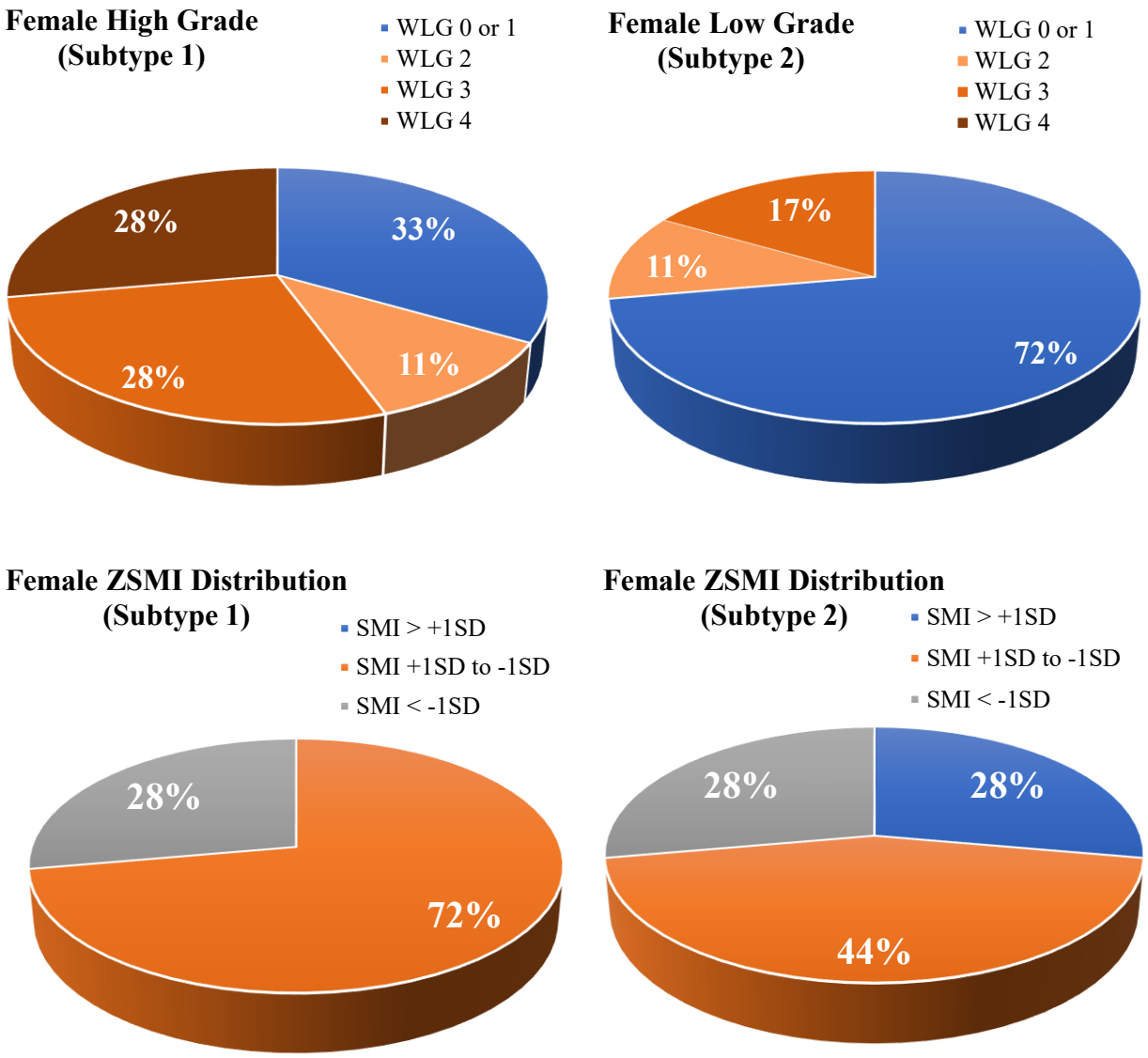


Figure 3.7 Clinical benchmark analyses for females
The pie charts depict the distribution of samples within the clusters based on the WL-BMI grades and z-SMI distribution for the subtypes in females. High and Low-Grade labels correspond to the WL-BMI classification schema for the associated skeletal muscle subtypes.

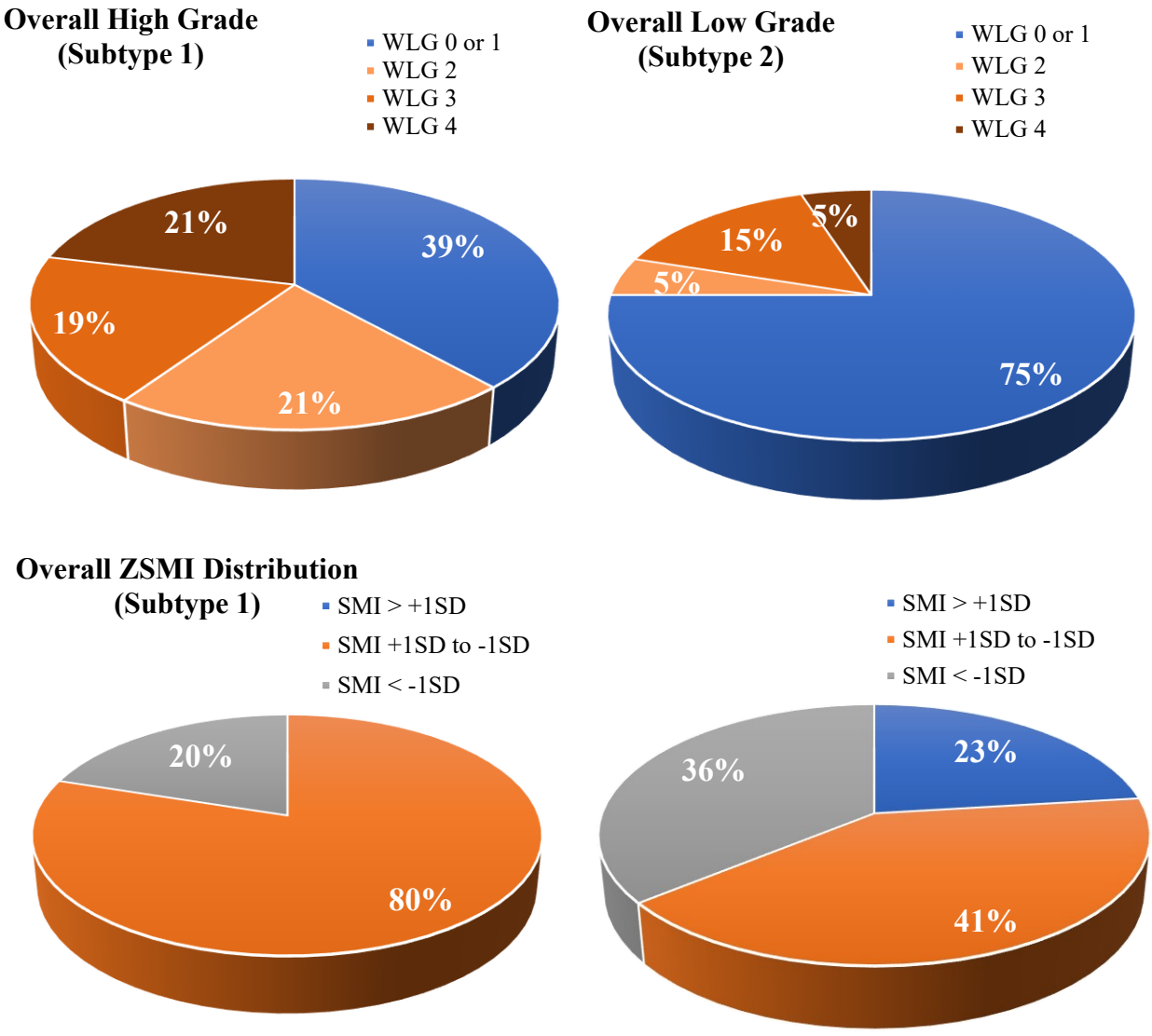


Figure 3.8 Clinical benchmark analyses for aggregate samples
The pie charts depict the distribution of samples within the clusters based on the WL-BMI grades and z-SMI distribution for the subtypes in overall samples. High and Low-Grade labels correspond to the WL-BMI classification schema for the associated skeletal muscle subtypes.

Table 3.1 Association statistics of the sex-specific clusters with WL-BMI grades and Z-SMI distribution

	Males (n=48)			Females (n=36)			Males + Females (n=84)		
Clinical Labels	**Subtype 1 (n=27)**	**Subtype 2 (n=21)**	**P value**	**Subtype 1 (n=18)**	**Subtype 2 (n=18)**	**P value**	**Subtype 1 (n=45)**	**Subtype 2 (n=39)**	**P value^{+}**
WL-BMI grade									
Grade 0 or 1	12 (44%)	17 (81%)	0.04*	6 (33%)	13 (72%)	0.03*	18 (38%)	30 (75%)	0.003**
Grade 2	6 (22%)	0 (0%)		2 (11%)	2 (11%)		10 (21%)	2 (5%)	
Grade 3	4 (15%)	2 (10%)		5 (28%)	3 (17%)		9 (19%)	6 (15%)	
Grade 4	5 (19%)	2 (10%)		5 (28%)	0 (0%)		10 (21%)	2 (5%)	
Z-SMI distribution									
> +1 SD	0 (0%)	4 (19%)	0.001**	0 (0%)	5 (28%)	0.04*	0 (0%)	9 (23%)	0.0001***
+1SD to -1SD	23 (85%)	8 (38%)		13 (72%)	8 (44%)		36 (80%)	16 (41%)	
< -1SD	4 (15%)	9 (43%)		5 (28%)	5 (28%)		9 (20%)	14 (36%)	

Subtype 1 = WL/low muscle mass
Subtype 2 = No WL/high muscle mass
$^{+}$Fisher's exact test
* represents statistical significance at P-value < 0.05
** represents P-value < 0.001
*** represents P-value < 0.0001.

3.3.6 Differential Expression (DE) of coding and non-coding genes in the muscle subtypes from intNMF classification

Molecular subtypes of human skeletal muscle derived from unsupervised ML algorithms do not directly allude to the underlying complex biological processes. Therefore, I subjected the global gene profiling data to differential expression analysis to discern the dysregulated genes in muscle subtype 1 relative to subtype 2 based on intNMF sample compositions for these subtypes in sexes **(Appendix Tables A.1-A.11).** All RNA classes were analyzed individually. The following DE RNAs were identified in males: 3519 mRNAs, 4231 lncRNAs, 59 snoRNAs, 72 miRNAs, and 136 piRNAs. The DE RNAs identified in females are 2847 mRNAs, 2638 lncRNAs, 57 snoRNAs (57), 39 miRNAs, and 26 piRNAs. tRNAs were not found to be DE in either sex except for one DE tRNA in males **(Appendix Table A.6)**. Sex-related differences in DE profiles showed an overlap between the sexes as follows: mRNAs (n= 2376, 59%) and lncRNAs (n= 2385, 53%). Interestingly, a limited overlap of small ncRNAs was observed between the sexes (n=17, 34.7% for miRNAs, n=7, 6.4% for piRNAs, and n=7, 6.4% for snoRNAs), indicating the small ncRNAs are highly sexually dimorphic compared to mRNAs and lncRNAs.

Overall, there is a trend towards higher DE coding and non-coding transcripts in males. There were no DE tRNAs in females, and one DE tRNA was observed in males between the subtypes**.** It is noted that tRNAs governing protein synthesis machinery were not differentially expressed between the subtypes identified in males and females. Thus far, mechanisms have highlighted the importance of protein degradation pathways relative to impaired protein synthesis. tRNAs have shown high DE in various cancers compared to healthy non-cancerous counterparts as a necessary condition under which the growth of tumor mass, increased protein synthesis, dysregulated apoptotic machinery, and uncontrolled cell division was explained [63]. Human skeletal muscle is

predominantly a differentiated tissue, and it is not surprising that the tRNA expression is not dysregulated between the muscle subtypes and appears to reflect a state of skeletal muscle homeostasis in patients with cancer. In contrast, other coding and non-coding RNAs are highly differentially expressed. There are 497 tRNAs in the human genome [53], and my profiling study identified 409 and 411 in male and female muscle biopsies, respectively. RNA modifications affecting sequencing were not accounted for, and previous studies in other models have not supported this premise for tRNA sequencing-based DE profiles [63].

3.3.7 Comprehensive heatmap representation of the DE profiles in muscle subtypes

The unsupervised clustering method, intNMF, identified two molecular subtypes of muscle for males and females. A Heatmap of row-scaled normalized and log-transformed expression of 200 DE mRNAs is illustrated, wherein the top 100 up-and down-regulated genes were considered for the schematic representation (Panel A in Figures 3.9 and 3.10 for males and females, respectively). The clinical labels for the corresponding samples are shown below the heatmap for individual subtypes and age and cancer diagnosis (Panels B and C in **Figures 3.9 and 3.10** for males and females, respectively). Overall the individual parameters do not define the subtypes based on the binary classification of clinical labels and show significant non-overlap with the machine learning-identified subtypes of skeletal muscle. Therefore the premise presented herein offers a novel modality to reveal complex biological processes underlying muscle wasting in patients with cancer

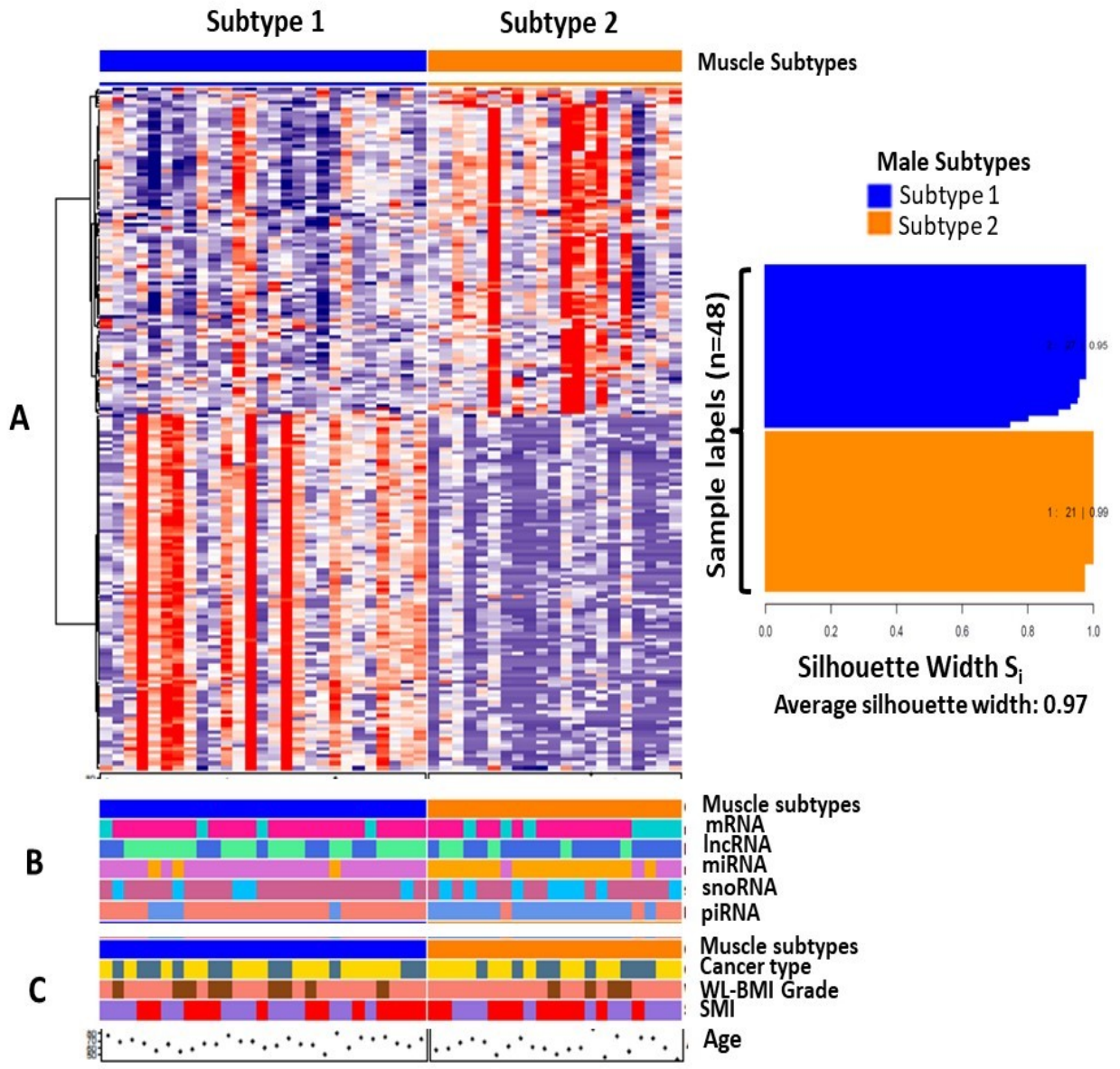


Figure 3.9 Heatmap representation of differentially expressed genes from human skeletal muscle molecular subtypes in males

The Heatmap schematic depicted represents the top 200 Differentially Expressed (DE) mRNAs with identified muscle subtypes 1 and 2. **Panel A** represents row-scaled normalized and log-transformed expression for mRNAs between the two-consensus clusters computed using the intNMF algorithm. Columns in the heatmap annotation represent samples. **Panel B** shows sample label distribution from individual RNAs using NMF clusters for comparison. **Panel C** is the annotation for clinical and body composition variables. SMI, cancer type, and WL-BMI grades are depicted. **The right panel (next to the heatmap)** is a schematic representation of the Silhouette Width (SW). SW is an index representing the validity of clusters within a range 0-1, with 1 meaning that the samples within the clusters are closest compared to their neighboring clusters. 0 indicates samples within the cluster are primarily distinct and do not belong to the same cluster.

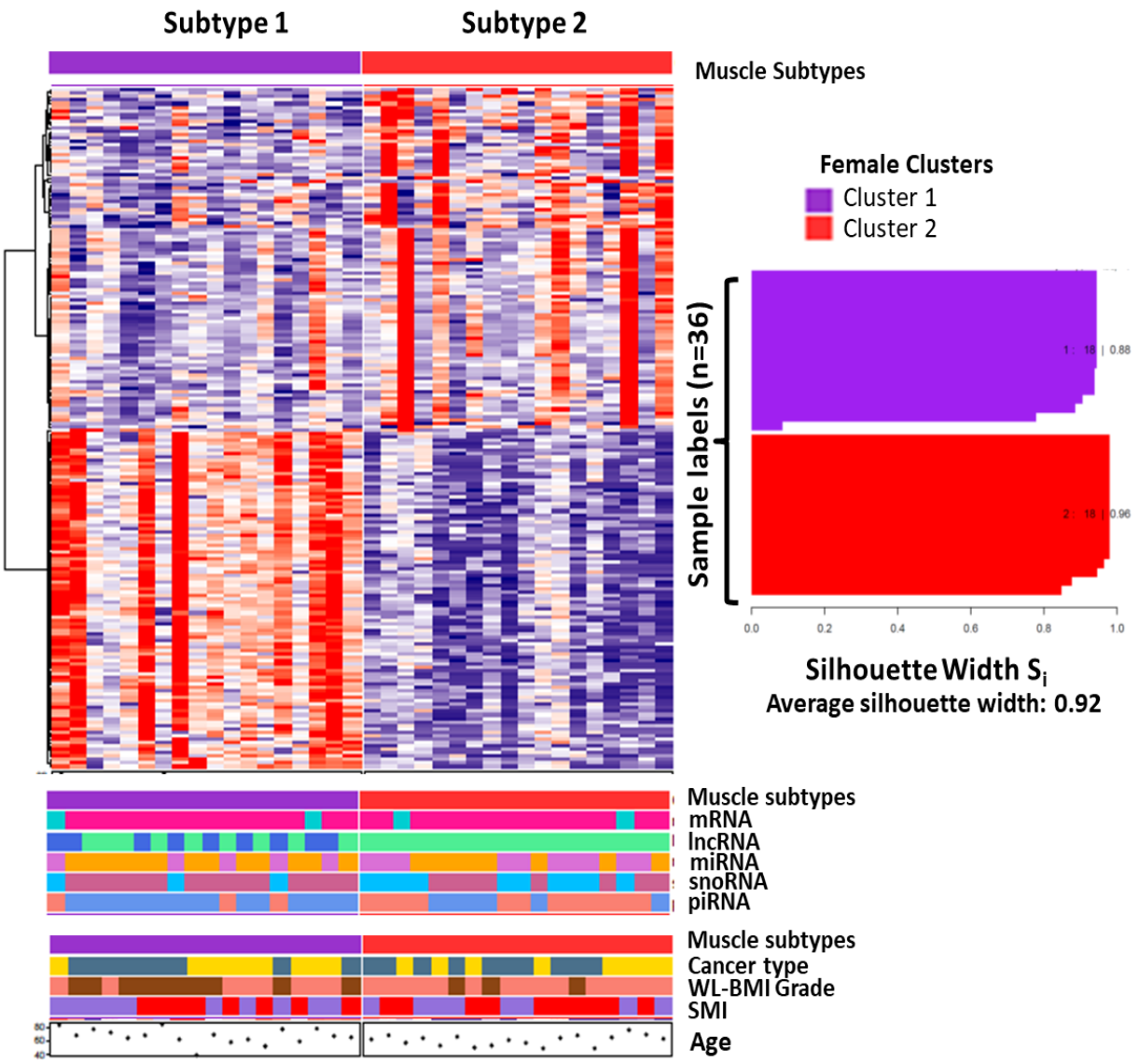


Figure 3.10 Heatmap representation of differentially expressed genes from human skeletal muscle molecular subtypes in females

The Heatmap schematic depicted represents the top 200 Differentially Expressed (DE) mRNAs with identified muscle subtypes 1 and 2. **Panel A** represents row-scaled normalized and log-transformed expression for mRNAs between the two-consensus clusters computed using the intNMF algorithm. Columns in the heatmap annotation represent samples. **Panel B** shows sample label distribution from individual RNAs using NMF clusters for comparison. **Panel C** is the annotation for clinical and body composition variables. SMI, cancer type, and WL-BMI grades are depicted. **The right panel (next to the heatmap)** is a schematic representation of the Silhouette Width (SW). SW is an index representing the validity of clusters within a range 0-1, with 1 meaning that the samples within the clusters are closest compared to their neighboring clusters. 0 indicates samples within the cluster are primarily distinct and do not belong to the same cluster.

3.3.8 Ingenuity Pathway Analysis (IPA) for the DE mRNAs between skeletal muscle molecular subtypes

There is a lack of a dedicated knowledge base for the functional annotation of human skeletal muscle tissue-specific gene expression profiles. IPA data are not enriched for human skeletal muscle transcriptome, and the functional annotations are based on published findings and manual curation of the data fields, primarily from cancer tissues and cell lines. Nonetheless, this study demonstrates the utility of IPA as a knowledge base in interpreting skeletal muscle transcriptome profiles. The compelling premise of this study rests on the fact that biology is redundant in that several genes contribute to the regulation of the same pathway in different cells and tissues and exhibit pleiotropy: the same gene might contribute to multiple pathways. Hence sexual dimorphism in the regulatory mechanisms of skeletal muscle in cachexia is interrogated at the gene and pathway levels using IPA knowledge.

There were two prime objectives for performing pathway analyses: one was to identify sexually dimorphic patterns of pathways, and the second was the functional annotation of subtypes at the DE gene level. I subjected the DE mRNAs within the muscle subtypes in males and females to IPA core analysis. The Venn diagram **(Figure 3.11)** indicates males' and females' overlapping and distinct canonical pathways. There was a significant overlap of canonical pathways between males and females at 66.4% (n=71 pathways, P-value <0.05 or -log(P-value) ≥ 1.3), with 19.6% (n=21 pathways) and 14% (n=15 pathways) unique pathways enriched in males and females, respectively. Detailed pathway analysis results are reported in the bubble plots of the distinctive pathways in males **(Figure 3.12)** and females **(Figure 3.13)**. Although a substantial overlap of the pathways was observed between males and females (n=71), the results are reported independently as bubble plots for males **(Figure 3.14)** and females **(Figure 3.15).** This is due to the differing

count of molecules or genes within the pathways, and their differing statistical significance, -log(P-value) and gene expression differences (fold-change) within each pathway for the individual sexes. The considerable overlap at the level of pathways supports the view that redundancy and pleiotropy are characteristics of the genes conferring pathways and phenotypes. This aspect has been discussed in detail in the ensuing Chapter 5, wherein I provided a detailed list of the molecules regulating the pathways as supporting evidence for the same gene contributing to different pathways.

There is no apriori information on the human skeletal muscle whole transcriptome sequencing dataset and no standard way to annotate the clusters identified in this study. Clinical benchmarking using the WL-BMI grading system and z-SMI distribution indicated subtype 2 to be a relatively less severe group in terms of WL-BMI grades and z-SMI distribution compared to subtype 1. In silico pathway analysis was performed to elucidate the common and distinct pathways enriched in males and females. Male pathway analysis results revealed pathways relatively enriched in inflammatory and immune cell regulation, and female expression analyses revealed pathways relatively more enriched in metabolic regulation. For example, acute phase response signaling, IL17 signaling, and the role of cytokines in mediating communication between immune cells, amongst others, were found to be exclusively regulated in males. In contrast, the urea cycle, triacylglycerol degradation, ethanol degradation, phospholipases, and eicosanoid signaling were found to be regulated solely in females. Overall, the overlapping pathways involved neuroendocrinal, metabolic, and immune cell regulatory pathways were enriched in males and females. Thus, from the clinical and in silico bioinformatic analyses, it is evident that subtype 1 is a relatively more severe muscle wasting group than subtype 2**.**

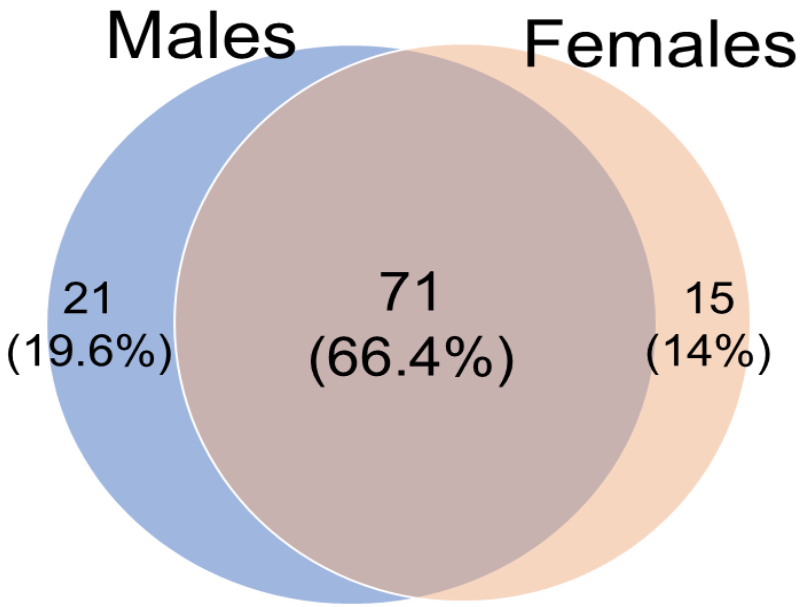


Figure 3.11 Venn diagram of canonical pathways between males and females
Sex-specific differentially expressed genes between the subtypes subjected to Ingenuity Pathway Analysis. The Venn diagram indicates the overlapping (brown intersection) and unique pathways between males (blue) and females (peach). The pathways significantly enriched in males and females at -log(P-value) ≥ 1.3 overlapped to examine the sexually dimorphic expression patterns of the canonical pathways.

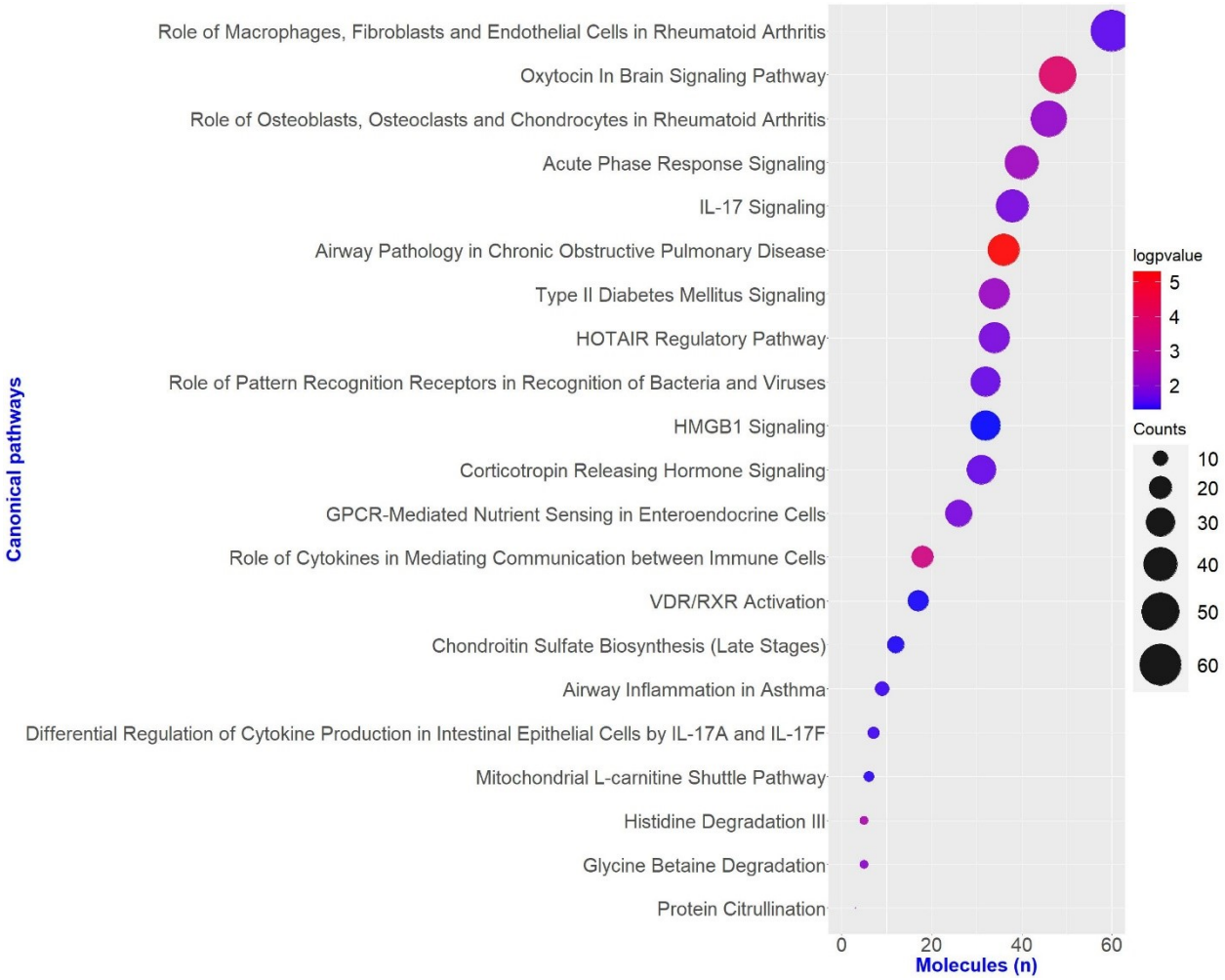


Figure 3.12 Bubble plot of unique pathways in males from skeletal muscle gene expression profiles N=21 statistically significant canonical pathways were exclusively enriched in males (Figure 3.11, left side of the Venn diagram). The bubble plot illustrates these 21 canonical pathways in males and is arranged based on the increasing order of the count of molecules or genes within a given pathway. The X-axis represents the number of molecules or genes within the individual pathways. Y-axis represents the terminology of the canonical pathways identified from IPA. The bubble size corresponds to the count of molecules within the pathways, as indicated in the legend on the bottom right side of the figure, i.e., small to large bubbles representing a low to a high number of molecules within the given pathway, respectively. The color gradient within the bubbles corresponds to the statistical significance, as indicated in the figure's legend on the top right side. Blue to red is indicative of low to high statistical significance, respectively.

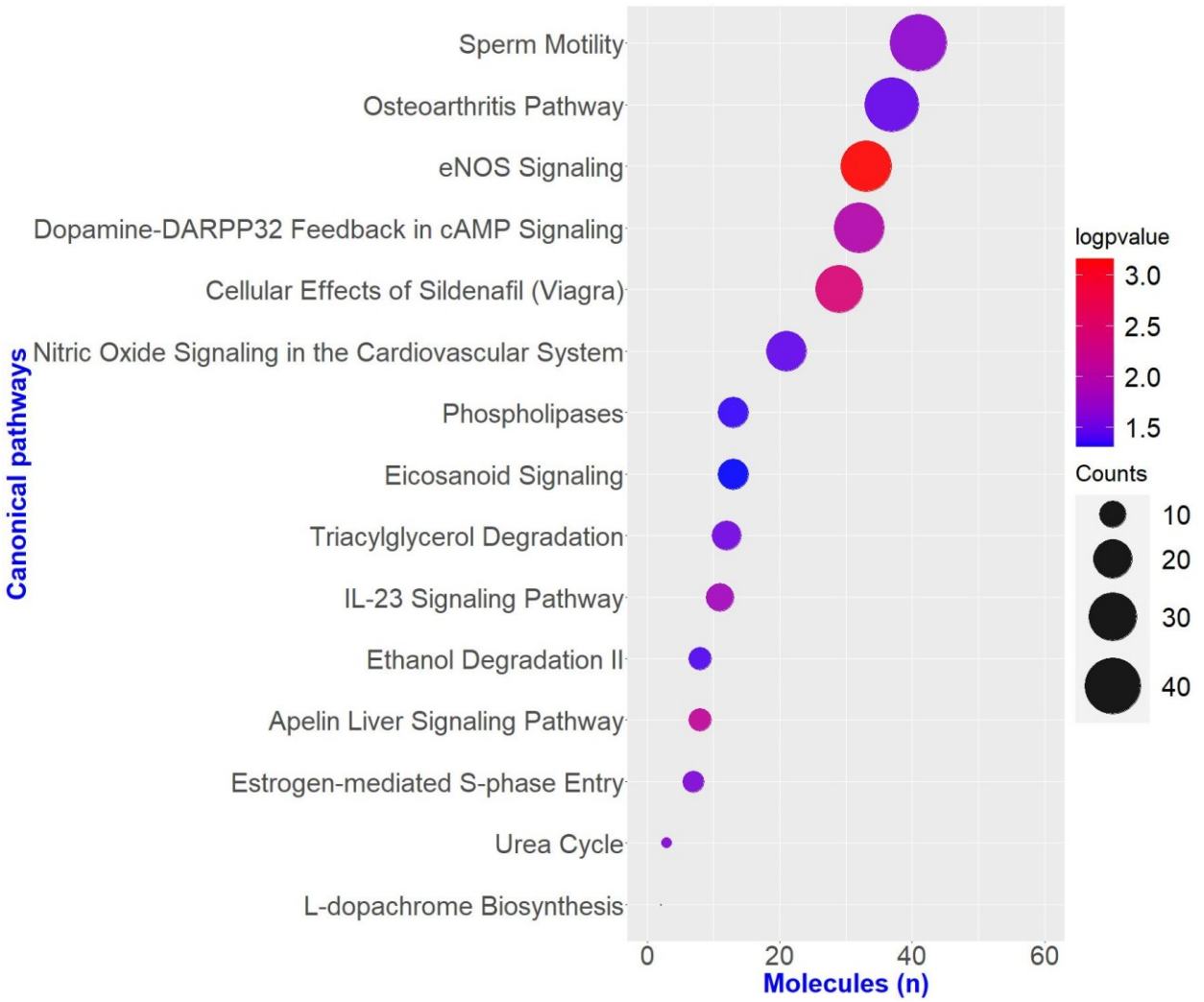


Figure 3.13 Bubble plots of unique pathways in females from skeletal muscle gene expression profiles

N=15 statistically significant canonical pathways exclusively enriched in females (Figure 3.11, right side of the Venn diagram). The bubble plot illustrates these 15 canonical pathways in females and is arranged based on the increasing order of the count of molecules or genes within a given pathway. The X-axis represents the number of molecules or genes within the individual pathways. Y-axis represents the terminology of the canonical pathways identified from IPA. The bubble size corresponds to the count of molecules within the pathways, as indicated in the legend on the bottom right side of the figure, i.e., small to large bubbles representing a low to a high number of molecules within the given pathway, respectively. The color gradient within the bubbles corresponds to the statistical significance, as indicated in the figure's legend on the top right side. Blue to red is indicative of low to high statistical significance, respectively.

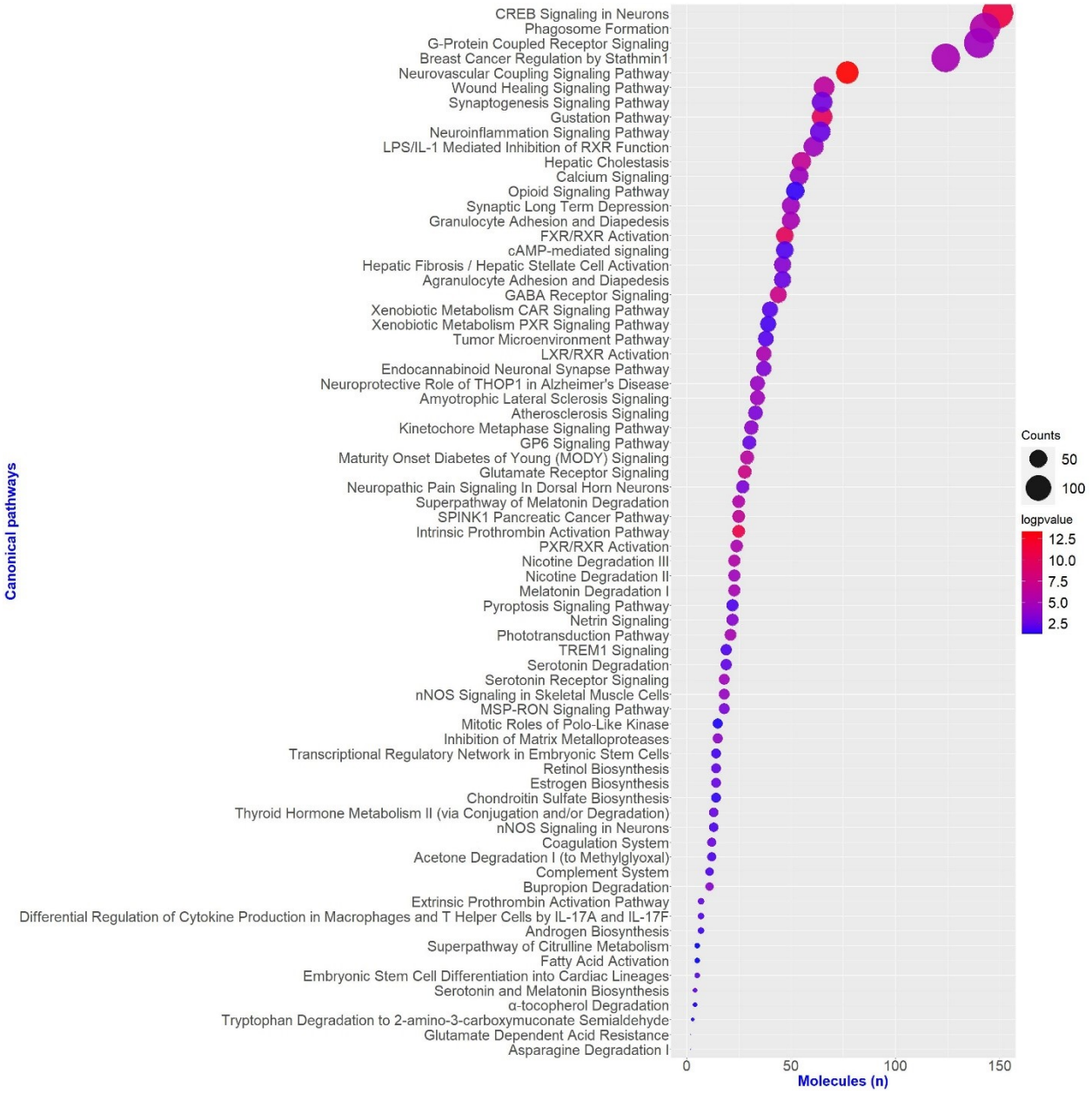


Figure 3.14 Bubble plot of the overlapping pathways from the male pathway analysis dataset
N=71 statistically significant canonical pathways overlapped between males and females (Figure 3.11, overlapping Venn diagram intersection). Despite the considerable overlap, the bubble plots are represented separately for males and females due to the differing count of molecules within the pathways and their statistical significance. The bubble plot in this figure illustrates 71 canonical pathways from the male dataset. It is arranged based on the increasing order of the count of molecules or genes within a given pathway. The X-axis represents the number of molecules or genes within the individual pathways. Y-axis represents the terminology of the canonical pathways identified from IPA. The bubble size corresponds to the count of molecules within the pathways, i.e., small to large bubbles representing a low to a high number of molecules within the given pathway, respectively. The color gradient within the bubbles corresponds to the statistical significance, as indicated in the figure's legend on the top right side. Blue to red is indicative of low to high statistical significance, respectively.

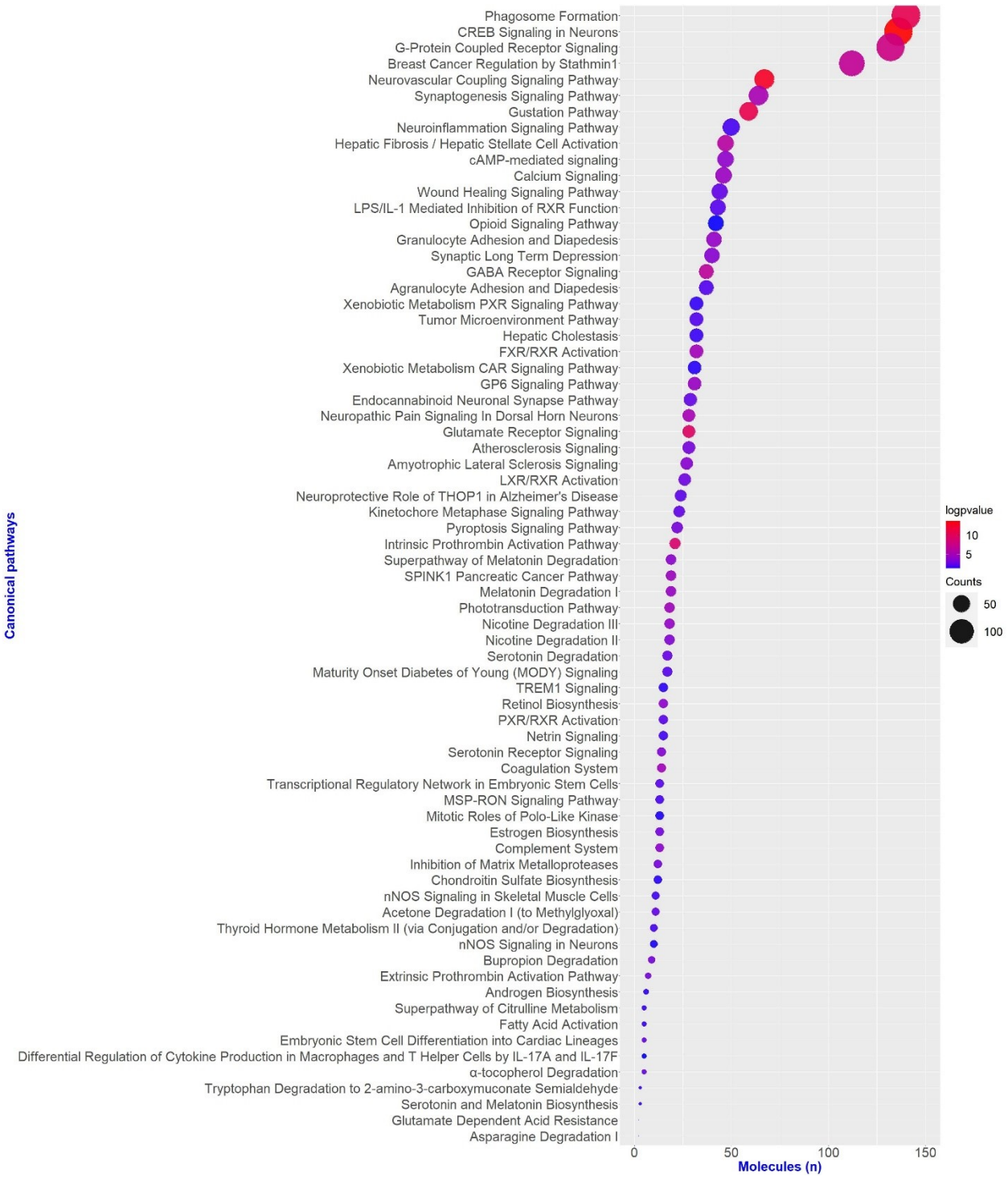


Figure 3.15 Bubble plot of the overlapping pathways from the female pathway analysis dataset
N=71 statistically significant canonical pathways overlapped between males and females (Figure 3.11, overlapping Venn diagram intersection). See Figure 3.14 for caption details.

3.4 Discussion

Cancer-associated muscle wasting emanates from synergistic interactions of myriad mediators, perturbations of signaling events, and biological pathways contributing to the wasting phenomenon. This investigation provides insights into the unsupervised clustering-based identification of sex-specific molecular clusters or subtypes of human skeletal muscle from patients with cancer. Unsupervised clustering facilitated the identification of two sex-specific coherent molecular subtypes of muscle using the entire transcribed genome within the skeletal muscle of patients with cancer. The subtypes were consequently subjected to molecular characterization (DE of genes within subtypes and sexes) to gain functional insights. Clinical and in silico methods were used to ascertain the intrinsic characteristics of the subtypes.

To date, molecular profiling studies using human skeletal muscle biopsies quantified candidate genes using qRT-PCR or microarray technology to profile mRNAs. The human muscle studies in cachexia literature have not characterized sex-specific transcriptional changes. In a recent study, the authors utilized an Ampliseq transcriptome gene expression kit to perform mRNA profiling in rectus abdominis muscle and subcutaneous adipose tissue specimens of Pancreatic Ductal Adenocarcinoma (PDAC) vs. healthy controls [24] and the study was not optimized to identify differences within the expression profiles in patients with cancer due to small sample size. While the targeted sequencing generated mRNA expression profiles, it does not represent the entire transcribed genome, and their regulation and sex-related differences remain unaccounted [24]. Human skeletal muscle gene expression data vary according to sex [42,43,45]. Therefore, in this study, sex-specific molecular clusters were generated from the human skeletal muscle of patients with cancer. The single RNA clustering was inadequate to classify patients owing to clustering based on the RNA types and the individual association statistics. Performing association analyses for

individual RNAs is challenging and enigmatic, given that there are n=6 RNAs that would result in multiple association analyses if individual phenotypic variables were to be used. Therefore, integrated transcriptome analyses were performed using intNMF and revealed two cohesive sex-specific subtypes of skeletal muscle in patients with cancer, with silhouette width corresponding to 0.97 and 0.92 (1.0 being ideal) for males and females, respectively.

Unsupervised clustering-based methods have advantages over conventionally used criteria for patient classification, particularly in cachexia research due to the fact that (i) the diagnostic criteria are not consistent across molecular studies, (ii) it is imperative to study muscle transcriptome independently of any clinical or body composition variables to understand the state of muscle biopsied at the time of surgery irrespective of weight loss measured months before sample collection (iii) supervised methods use case vs. control study designs with aggregate cancer types, sex, and investigator defined weight loss or muscle loss criteria and then perform correlative analyses to represent the association between a representative set of genes of interest. Body mass is an aggregate measure encompassing body fat and lean mass (that includes muscle, organs, and bone). Therefore, classifying patients based on differing criteria of %WL on the samples obtained from human skeletal muscle biopsies of patients with cancer at a static time point is elusive. Weight stability masks the underlying skeletal muscle loss [64,65]. While the supervised association approach defines a set of correlated genes, it fails to unveil the underlying changes at the whole transcriptional level. Importantly it also invokes challenges in the replicability and validity of findings. Vidman et al. (2019) evaluated the clustering performance based on sample size, distribution of subtypes, and sample heterogeneity from the publicly available breast, brain, kidney, and stomach cancer gene expression datasets. Sample size beyond a critical threshold (n=36) had a limited effect on the performance of cluster analyses, whereas homogeneity with

respect to sex was important for performance [66]. Therefore, this study suggested optimum sample size and identified sex-specific clusters and not aggregate measures to characterize muscle from patients with cancer.

This is the first report to indicate the prominence of sexual dimorphism in the non-coding part of the genome. The percent overlap of expression of RNAs in the descending order is as follows mRNAs (59% overlap between males and females), lncRNAs (53% overlap between males and females), miRNAs (34.7% overlap between males and females), piRNAs (6.4% overlap between males and females) and snoRNAs (6.4% overlap between males and females). These results illustrate the importance of sexually dimorphic patterns of expression of ncRNAs. Some of the sex-related differences likely reflect the differences in sample size, with males at n=48 vs. females at n=36. If such differences could be ascribed to sample size, independent analysis at higher sample sizes is needed to replicate and validate the findings. As such, this could be noted as a potential limitation to the study findings.

At the pathway analyses level, although 66.4% overlapping pathways were found between males and females, these were performed considering the differential gene expression (mRNA) levels. Crosstalk mechanisms of the ncRNA-mediated regulation (provided in the ensuing chapter) may provoke higher-order integrative regulatory mechanisms of the regulated gene targets.

This study reports a sex-specific transcriptional landscape of human skeletal muscle from whole transcriptome profiling and unsupervised clustering. There is no standard process to determine the constituent peculiarity of these subtypes, and there is also not a previously generated sequencing dataset available to compare the subtypes. I, therefore, approached this challenge with two precise benchmarking methods. The intrinsic characteristics of the subtypes in males and females were

articulated using clinical benchmarks and expression analyses contributing to pathway-level regulation.

Clinical benchmarking of the subtypes was decoded using the WL-BMI grading system [60] and z-score SMI distribution (ZSMI). Statistically significant differences were observed between the severity of the grading system and the two subtypes of muscle in males and females, wherein subtype 1 was indicative of higher WL-BMI grades and low muscle mass or, in other words, showed relatively higher severity of WL-BMI grades and low muscularity than subtype 2 in males and females, respectively. Similar trends followed for the aggregate statistics (i.e., combining male and female samples), wherein subtype 1 was a higher grade severity group than subtype 2. Thus, from the clinical viewpoint, subtype 1 was a more severe group compared to subtype 2.

The muscle transcriptome expression profiling and pathways elicited the perturbations in metabolic, immune cell regulation, with a central emphasis on the perturbations involved in skeletal muscle microenvironment. Cancer-associated muscle wasting is not a single mediator or single pathway-derived mechanism. Studies conducting blockage of single mediators or pathways leading to muscle wasting have failed due to the complex heterogeneous nature of the syndrome. Cachexia is characterized by the complex interplay between the myriad set of mediators acting in concert. The acute inflammatory response is important for skeletal muscle regeneration. However, in cachexia, there is a chronic sustained inflammatory response resulting in muscle wasting by activating procatabolic pathways and inhibiting anabolic pathways. Pro-inflammatory cytokines released from host-tumor interactions such as TNF, IL1B, and IL6 are well-characterized in rodent model systems and are known to be associated with skeletal muscle, and fat wasting, as well as are essential contributors to hypothalamic inflammation. Direct effects of these catabolic mediators facilitate the activation of downstream signaling pathways such as NFkB, JAK/STAT

signaling, reduced PI3K//Akt signaling, and activation of myostatin and activin, amongst others[67]. These culminate in the proteolysis of myofibrillar proteins, thereby causing muscle wasting. In a normal physiological state, muscle mass is maintained by a balance between anabolic and catabolic signals leading to muscle protein synthesis and degradation . In cachexia, however, there is an imbalance of protein homeostasis, with an increase in protein degradation defined by Ubiquitin Proteasome Pathway (UPP), Autophagy, and calcium-dependent calpains and caspases. Experimental model systems suggest the involvement of these pathways in the pathophysiology of cachexia, whereas results from clinical studies using human muscle biopsies and mRNA profiling suggest otherwise. I did not capture proteolysis pathways in my datasets. These findings corroborate the findings from clinical studies conducted using human muscle [21,22]. It is not surprising that the ubiquitin-proteasome or autophagy pathways were not identified in this study due to the rapid development of cachexia in model systems compared to human clinical cachexia.

Skeletal muscle serves as a precursor of amino acids to the liver for synthesizing acute-phase proteins. The liver sustains the cachexic state by increased energy expenditure and hypermetabolism through increased gluconeogenesis, secretion of acute phase reactant proteins, and reduced low-density lipoproteins, all these resulting in muscle proteolysis and fat breakdown in adipose tissue. A recent study demonstrated a marked increased production of acute phase reactants by glucocorticoids and pro-inflammatory cytokines in C26 cancer cachexia model, and in patients with cancer [64]. The acute phase signaling pathway was highly enriched in males, specifically in subtype 1. These findings suggest the proteolysis to be an upstream mechanism and hence could be one of the plausible reasons for not being represented in a cachexic state in human muscle data from patients with cancer, unlike murine models that are characterized by the rapid development of cachexia and acute muscle wasting rather than the chronic wasting observed in the

human condition [65]. The expression levels of the pro-inflammatory markers were upregulated in subtype 1 vs. 2, indicative of higher expression of these mediators in subtype 1.

While this is the first time the unsupervised method has been applied to human muscle from patients with cancer, these were previously employed in consortia studies in several tumor types. Previous gene expression profiling analyses have identified subtypes of pancreatic [68], colorectal[69], breast cancer [70,71], muscle-invasive bladder cancer [72], and B-cell lymphoma [73], amongst several other cancer types, using integrative clustering methods. Utilizing the multi-omic approaches combining mRNA, ncRNA, and methylation datasets provide novel insights into molecular subtyping in addition to previously known histological subtypes. Skeletal muscle is composed of myofibers, and changes in muscle function are reflected upon by the changes in the myofiber composition. Data-driven strategies were undertaken for fiber type classification to identify myofiber clusters in older human muscles wherein unsupervised clustering revealed the existence of six distinct myofiber clusters from *vastus lateralis* muscle of n=56 healthy elderly human subjects [74]. Several in vitro studies have identified abundant lncRNAs and precursor miRNAs in muscle cell differentiation processes using single nucleus RNA-seq [75]. This emphasizes the importance of studying the non-coding aspect of gene regulation in human skeletal muscle. Micheli et al. (2020) profiled the transcriptome of individual muscle stem cells and muscle-resident cells from adult human muscle samples using single-cell RNA sequencing.[76] Role of muscle fiber types in the degradation of muscle mass remains debated. I did not have fiber-type data for my study cohort. Skeletal muscle heterogeneity is reflected by fiber type composition differing in myosin heavy chain isoforms, metabolism, and neural response. Type 1, or slow-twitch fibers express MYH7 that are highly expressed in females, and type II, or fast-twitch fiber type, express MYH2 (type 2A), MYH1 (type 2x), and MYH4 (type 2B). There were no significant

expression fold differences between MYH expression in males. Females have higher type I fibers expressing MYH7. In this study, MYH7 was downregulated by -1.5 fold in subtype 1 vs. 2. Fiber typing annotation was not an objective of this study and is out of scope; however, downregulated expression of MYH7 in subtype 1 solidifies the annotation obtained from clinical benchmarking and literature reported marker genes and pathway analyses.

Transcriptome profiles from cancer-associated muscle wasting and those obtained from a healthy condition are distinct. Global mRNA profiles from the human skeletal muscle of healthy donors in the setting of aging [68,69] and exercise [70-73] have been described using data generated from the *vastus lateralis* muscle. However, studies in human skeletal muscle from patients with cancer revealed distinct gene expression profiles and suggested discrete mechanism that prevails within these cohorts of patients [22,24]. The strength of this study is that a systematic, comprehensive transcriptional landscape of human skeletal muscle was generated from the biopsies. The biopsies were collected from the same anatomical location (rectus abdominis muscle) by the designated surgical oncologist, and the subsequent sequencing experiments, data generation, and downstream analyses were executed in a single batch from each patient's muscle biopsy specimen. The study cohort included all patients diagnosed with pancreatic and colorectal cancer to restrict the heterogeneity related to cancer type.

I used matrix factorization to cluster samples. However, the applicability of matrix factorization is colossal to finding complex structures or relationships between genes, gene set discovery, and time course analysis, among others [36]. Brunet et al. were the first researchers to utilize NMF in the gene expression domain using three cancer datasets of leukemia, medulloblastoma, and central nervous system tumors and identified biologically significant phenotypes [34]. Attempts in cachexia literature to adopt machine learning models to detect changes in body composition using imaging techniques

have gained prominence in recent years. While these are the preliminary findings, the basis of these has been the input variables from CT assessment or weight loss [74-76]. The central question that remains unaddressed is the use of varying cut-offs, the applicability of appropriate machine learning models to assess body composition, and the underlying biology insights. This investigation advances the field forward using integrative techniques adopted to identify molecular subtypes from the transcriptional profiles and to understand biological insights into the muscle of patients with cancer.

The novel aspect of this study is the sex-specific molecular characterization of human skeletal muscle from patients with cancer using whole transcriptome profiles by NGS. This study demonstrates the importance of unsupervised methods and the utilization of transcriptional expression profiles to allow reproducible data across labs rather than relying on arbitrary cut-offs for a phenotype binary in the association studies. My study identified two sex-specific molecular subtypes using the intNMF algorithm. Future studies with higher sample sizes may potentially identify additional clusters or subtypes representative of the disease trajectory of cachexia severity.

Given the multifactorial and complex nature of cachexia syndrome, no single pathway could be accredited to the causality of the syndrome. Identifying molecular subtypes and functional insights into mRNA-based regulation is the first step toward discovering complex biological regulation and the underlying pathophysiological mechanisms. Recent reports have illuminated the crosstalk mechanisms and the integrative networks of RNA crosstalk (mediated by ncRNAs) that regulate mRNAs and thereby confer the phenotypic changes and pathophysiology. These higher levels of RNA crosstalk eliciting gene-level regulations are independently explored in the ensuing chapter.

3.5 Conclusion

Construing insights into the biological mechanisms in patients with cancer from the transcriptional regulation standpoint is the investigational theme of this chapter. I used unsupervised clustering algorithms from transcriptional expression profiles from diverse RNA classes generated from rectus abdominis skeletal muscle biopsies of patients with cancer. I determined two sex-specific molecular subtypes independent of clinical or body composition variables. Sexually dimorphic expression profiles were prominent in small non-coding RNAs (miRNA, snoRNA, piRNA). lncRNAs and small non-coding RNA classes studied using NGS are reported for the first time in cachexia literature using human skeletal muscle biopsies. mRNA profiling by NGS is also first of its kind. Another unique feature of this study is that a single biopsy from each individual patient served to profile all RNA classes (in independent small RNA seq and RNA seq experiments), eliminating batch effects in sequencing studies.

Clinical and in silico benchmarking explained the relative intrinsic characteristics of muscle subtypes, with subtype 1 being a group associated with cachexia severity reflected by higher grades of WL-BMI. In silico IPA pathway analysis results indicated a multitude of signaling mediators underlying the molecular landscape of cachexia. The complex multilayered RNA cross-talk constituting the muscle subtype characteristics is explored independently and is summarized in the ensuing chapter.

3.6 References

1. Fearon K, Strasser F, Anker SD, Bosaeus I, Bruera E, Fainsinger RL, et al. Definition and classification of cancer cachexia: an international consensus. *Lancet Oncol.* 2011;12(5):489-495.
2. Talbert EE, Cuitino MC, Ladner KJ, Rajasekerea PV, Siebert M, Shakya R, et al. Modeling Human Cancer-induced Cachexia. *Cell Rep.* 2019;28(6):1612-1622 e1614.
3. Baracos VE. Bridging the gap: are animal models consistent with clinical cancer cachexia? *Nat Rev Clin Oncol.* 2018;15(4):197-198.
4. Tomasin R, Martin A, Cominetti MR. Metastasis and cachexia: alongside in clinics, but not so in animal models. *J Cachexia Sarcopenia Muscle.* 2019;10(6):1183-1194.
5. Suzuki T, Von Haehling S, Springer J. Promising models for cancer-induced cachexia drug discovery. *Expert Opin Drug Discov.* 2020;15(5):627-637.
6. Anoveros-Barrera A, Bhullar AS, Stretch C, Esfandiari N, Dunichand-Hoedl AR, Martins KJB, et al. Clinical and biological characterization of skeletal muscle tissue biopsies of surgical cancer patients. *J Cachexia Sarcopenia Muscle.* 2019;10(6):1356-1377.
7. Schersten T, Lundholm K. Lysosomal enzyme activity in muscle tissue from patients with malignant tumor. *Cancer.* 1972;30(5):1246-1251.
8. Lundholm K, Bylund AC, Holm J, Schersten T. Skeletal muscle metabolism in patients with malignant tumor. *Eur J Cancer (1965).* 1976;12(6):465-473.
9. Lundholm K, Edstrom S, Ekman L, Karlberg I, Bylund AC, Schersten T. A comparative study of the influence of malignant tumor on host metabolism in mice and man: evaluation of an experimental model. *Cancer.* 1978;42(2):453-461.

10. Williams A, Sun X, Fischer JE, Hasselgren PO. The expression of genes in the ubiquitin-proteasome proteolytic pathway is increased in skeletal muscle from patients with cancer. *Surgery.* 1999;126(4):744-749; discussion 749-750.

11. Khal J, Hine AV, Fearon KC, Dejong CH, Tisdale MJ. Increased expression of proteasome subunits in skeletal muscle of cancer patients with weight loss. *Int J Biochem Cell Biol.* 2005;37(10):2196-2206.

12. Pessina P, Conti V, Pacelli F, Rosa F, Doglietto GB, Brunelli S, et al. Skeletal muscle of gastric cancer patients expresses genes involved in muscle regeneration. *Oncol Rep.* 2010;24(3):741-745.

13. Rhoads MG, Kandarian SC, Pacelli F, Doglietto GB, Bossola M. Expression of NF-kappaB and IkappaB proteins in skeletal muscle of gastric cancer patients. *Eur J Cancer.* 2010;46(1):191-197.

14. D'Orlando C, Marzetti E, Francois S, Lorenzi M, Conti V, di Stasio E, et al. Gastric cancer does not affect the expression of atrophy-related genes in human skeletal muscle. *Muscle Nerve.* 2014;49(4):528-533.

15. Stephens NA, Skipworth RJ, Gallagher IJ, Greig CA, Guttridge DC, Ross JA, et al. Evaluating potential biomarkers of cachexia and survival in skeletal muscle of upper gastrointestinal cancer patients. *J Cachexia Sarcopenia Muscle.* 2015;6(1):53-61.

16. Stephens NA, Gallagher IJ, Rooyackers O, Skipworth RJ, Tan BH, Marstrand T, et al. Using transcriptomics to identify and validate novel biomarkers of human skeletal muscle cancer cachexia. *Genome Med.* 2010;2(1):1.

17. Boehm I, Miller J, Wishart TM, Wigmore SJ, Skipworth RJ, Jones RA, et al. Neuromuscular junctions are stable in patients with cancer cachexia. *J Clin Invest.* 2020;130(3):1461-1465.

18. Anoveros-Barrera A, Bhullar AS, Stretch C, Dunichand-Hoedl AR, Martins KJB, Rieger A, et al. Immunohistochemical phenotyping of T cells, granulocytes, and phagocytes in the muscle of cancer patients: association with radiologically defined muscle mass and gene expression. *Skelet Muscle.* 2019;9(1):24.

19. Judge SM, Nosacka RL, Delitto D, Gerber MH, Cameron ME, Trevino JG, et al. Skeletal Muscle Fibrosis in Pancreatic Cancer Patients with Respect to Survival. *JNCI Cancer Spectr.* 2018;2(3):pky043.

20. Johns N, Hatakeyama S, Stephens NA, Degen M, Degen S, Frieauff W, et al. Clinical classification of cancer cachexia: phenotypic correlates in human skeletal muscle. *PLoS One.* 2014;9(1):e83618.

21. Ebhardt HA, Degen S, Tadini V, Schilb A, Johns N, Greig CA, et al. Comprehensive proteome analysis of human skeletal muscle in cachexia and sarcopenia: a pilot study. *J Cachexia Sarcopenia Muscle.* 2017;8(4):567-582.

22. Gallagher IJ, Stephens NA, MacDonald AJ, Skipworth RJ, Husi H, Greig CA, et al. Suppression of skeletal muscle turnover in cancer cachexia: evidence from the transcriptome in sequential human muscle biopsies. *Clin Cancer Res.* 2012;18(10):2817-2827.

23. Narasimhan A, Greiner R, Bathe OF, Baracos V, Damaraju S. Differentially expressed alternatively spliced genes in skeletal muscle from cancer patients with cachexia. *J Cachexia Sarcopenia Muscle.* 2018;9(1):60-70.

24. Narasimhan A, Zhong X, Au EP, Ceppa EP, Nakeeb A, House MG, et al. Profiling of Adipose and Skeletal Muscle in Human Pancreatic Cancer Cachexia Reveals Distinct Gene Profiles with Convergent Pathways. *Cancers (Basel).* 2021;13(8).

25. Stretch C, Khan S, Asgarian N, Eisner R, Vaisipour S, Damaraju S, et al. Effects of sample size on differential gene expression, rank order and prediction accuracy of a gene signature. *PLoS One.* 2013;8(6):e65380.

26. Stretch C, Aubin JM, Mickiewicz B, Leugner D, Al-Manasra T, Tobola E, et al. Sarcopenia and myosteatosis are accompanied by distinct biological profiles in patients with pancreatic and periampullary adenocarcinomas. *PLoS One.* 2018;13(5):e0196235.

27. Carninci P, Kasukawa T, Katayama S, Gough J, Frith MC, Maeda N, et al. The transcriptional landscape of the mammalian genome. *Science.* 2005;309(5740):1559-1563.

28. Mattick JS, Makunin IV. Non-coding RNA. *Hum Mol Genet.* 2006;15 Spec No 1:R17-29.

29. Narasimhan A, Ghosh S, Stretch C, Greiner R, Bathe OF, Baracos V, et al. Small RNAome profiling from human skeletal muscle: novel miRNAs and their targets associated with cancer cachexia. *J Cachexia Sarcopenia Muscle.* 2017;8(3):405-416.

30. van de Worp W, Schols A, Dingemans AC, Op den Kamp CMH, Degens J, Kelders M, et al. Identification of microRNAs in skeletal muscle associated with lung cancer cachexia. *J Cachexia Sarcopenia Muscle.* 2020;11(2):452-463.

31. Liu Q, Deng J, Qiu Y, Gao J, Li J, Guan L, et al. Non-coding RNA basis of muscle atrophy. *Mol Ther Nucleic Acids.* 2021;26:1066-1078.

32. Lee DD, Seung HS. Learning the parts of objects by non-negative matrix factorization. *Nature.* 1999;401(6755):788-791.

33. Hutchins LN, Murphy SM, Singh P, Graber JH. Position-dependent motif characterization using non-negative matrix factorization. *Bioinformatics.* 2008;24(23):2684-2690.

34. Brunet JP, Tamayo P, Golub TR, Mesirov JP. Metagenes and molecular pattern discovery using matrix factorization. *Proc Natl Acad Sci U S A.* 2004;101(12):4164-4169.

35. Wang G, Kossenkov AV, Ochs MF. LS-NMF: a modified non-negative matrix factorization algorithm utilizing uncertainty estimates. *BMC Bioinformatics.* 2006;7:175.

36. Carmona-Saez P, Pascual-Marqui RD, Tirado F, Carazo JM, Pascual-Montano A. Biclustering of gene expression data by Non-smooth Non-negative Matrix Factorization. *BMC Bioinformatics.* 2006;7:78.

37. Kim H, Park H. Sparse non-negative matrix factorizations via alternating non-negativity-constrained least squares for microarray data analysis. *Bioinformatics.* 2007;23(12):1495-1502.

38. Moffitt RA, Marayati R, Flate EL, Volmar KE, Loeza SG, Hoadley KA, et al. Virtual microdissection identifies distinct tumor- and stroma-specific subtypes of pancreatic ductal adenocarcinoma. *Nat Genet.* 2015;47(10):1168-1178.

39. Robertson AG, Kim J, Al-Ahmadie H, Bellmunt J, Guo G, Cherniack AD, et al. Comprehensive Molecular Characterization of Muscle-Invasive Bladder Cancer. *Cell.* 2018;174(4):1033.

40. Montalvo RN, Counts BR, Carson JA. Understanding sex differences in the regulation of cancer-induced muscle wasting. *Curr Opin Support Palliat Care.* 2018;12(4):394-403.

41. Landen S, Hiam D, Voisin S, Jacques M, Lamon S, Eynon N. Physiological and molecular sex differences in human skeletal muscle in response to exercise training. *J Physiol.* 2021.

42. Welle S, Tawil R, Thornton CA. Sex-related differences in gene expression in human skeletal muscle. *PLoS One.* 2008;3(1):e1385.

43. Maher AC, Fu MH, Isfort RJ, Varbanov AR, Qu XA, Tarnopolsky MA. Sex differences in global mRNA content of human skeletal muscle. *PLoS One.* 2009;4(7):e6335.

44. Norman K, Stobaus N, Reiss J, Schulzke J, Valentini L, Pirlich M. Effect of sexual dimorphism on muscle strength in cachexia. *J Cachexia Sarcopenia Muscle.* 2012;3(2):111-116.

45. Lindholm ME, Huss M, Solnestam BW, Kjellqvist S, Lundeberg J, Sundberg CJ. The human skeletal muscle transcriptome: sex differences, alternative splicing, and tissue homogeneity assessed with RNA sequencing. *FASEB J.* 2014;28(10):4571-4581.

46. Smith GI, Mittendorfer B. Sexual dimorphism in skeletal muscle protein turnover. *J Appl Physiol (1985).* 2016;120(6):674-682.

47. Kechin A, Boyarskikh U, Kel A, Filipenko M. cutPrimers: A New Tool for Accurate Cutting of Primers from Reads of Targeted Next Generation Sequencing. *J Comput Biol.* 2017;24(11):1138-1143.

48. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics.* 2013;29(1):15-21.

49. Langmead B. Aligning short sequencing reads with Bowtie. *Curr Protoc Bioinformatics.* 2010;Chapter 11:Unit 11 17.

50. Kozomara A, Birgaoanu M, Griffiths-Jones S. miRBase: from microRNA sequences to function. *Nucleic Acids Res.* 2019;47(D1):D155-D162.

51. Bouchard-Bourelle P, Desjardins-Henri C, Mathurin-St-Pierre D, Deschamps-Francoeur G, Fafard-Couture E, Garant JM, et al. snoDB: an interactive database of human snoRNA sequences, abundance and interactions. *Nucleic Acids Res.* 2020;48(D1):D220-D225.

52. Piuco R, Galante PAF. piRNAdb: A piwi-interacting RNA database. *bioRxiv.* 2021:2021.2009.2021.461238.

53. Chan PP, Lowe TM. GtRNAdb 2.0: an expanded database of transfer RNA genes identified in complete and draft genomes. *Nucleic Acids Res.* 2016;44(D1):D184-189.

54. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014;15(12):550.

55. Stein-O'Brien GL, Arora R, Culhane AC, Favorov AV, Garmire LX, Greene CS, et al. Enter the Matrix: Factorization Uncovers Knowledge from Omics. *Trends Genet.* 2018;34(10):790-805.

56. Gaujoux R, Seoighe C. A flexible R package for nonnegative matrix factorization. *BMC Bioinformatics.* 2010;11:367.

57. Rand WM. Objective Criteria for the Evaluation of Clustering Methods. *Journal of the American Statistical Association.* 1971;66(336):846-850.

58. Chalise P, Fridley BL. Integrative clustering of multi-level 'omic data based on non-negative matrix factorization algorithm. *PLoS One.* 2017;12(5):e0176278.

59. Kramer A, Green J, Pollard J, Jr., Tugendreich S. Causal analysis approaches in Ingenuity Pathway Analysis. *Bioinformatics.* 2014;30(4):523-530.

60. Martin L, Senesse P, Gioulbasanis I, Antoun S, Bozzetti F, Deans C, et al. Diagnostic criteria for the classification of cancer-associated weight loss. *J Clin Oncol.* 2015;33(1):90-99.

61. Kazemi-Bajestani SM, Mazurak VC, Baracos V. Computed tomography-defined muscle and fat wasting are associated with cancer clinical outcomes. *Semin Cell Dev Biol.* 2016;54:2-10.

62. Derstine BA, Holcombe SA, Ross BE, Wang NC, Su GL, Wang SC. Skeletal muscle cutoff values for sarcopenia diagnosis using T10 to L5 measurements in a healthy US population. *Sci Rep.* 2018;8(1):11369.

63. Krishnan P, Ghosh S, Wang B, Heyns M, Li D, Mackey JR, et al. Genome-wide profiling of transfer RNAs and their role as novel prognostic markers for breast cancer. *Sci Rep.* 2016;6:32843.

64. Martin L, Birdsell L, Macdonald N, Reiman T, Clandinin MT, McCargar LJ, et al. Cancer cachexia in the age of obesity: skeletal muscle depletion is a powerful prognostic factor, independent of body mass index. *J Clin Oncol.* 2013;31(12):1539-1547.

65. Prado CM, Lieffers JR, McCargar LJ, Reiman T, Sawyer MB, Martin L, et al. Prevalence and clinical implications of sarcopenic obesity in patients with solid tumours of the respiratory and gastrointestinal tracts: a population-based study. *Lancet Oncol.* 2008;9(7):629-635.

66. Vidman L, Kallberg D, Ryden P. Cluster analysis on high dimensional RNA-seq data with applications to cancer research - An evaluation study. *PLoS One.* 2019;14(12):e0219102.

67. Webster JM, Kempen L, Hardy RS, Langen RCJ. Inflammation and Skeletal Muscle Wasting During Cachexia. *Front Physiol.* 2020;11:597675.

68. Mahmassani ZS, Reidy PT, McKenzie AI, Stubben C, Howard MT, Drummond MJ. Age-dependent skeletal muscle transcriptome response to bed rest-induced atrophy. *J Appl Physiol (1985).* 2019;126(4):894-902.

69. Tumasian RA, 3rd, Harish A, Kundu G, Yang JH, Ubaida-Mohien C, Gonzalez-Freire M, et al. Skeletal muscle transcriptome in healthy aging. *Nat Commun.* 2021;12(1):2014.

70. Rundqvist HC, Montelius A, Osterlund T, Norman B, Esbjornsson M, Jansson E. Acute sprint exercise transcriptome in human skeletal muscle. *PLoS One.* 2019;14(10):e0223024.

71. Makhnovskii PA, Bokov RO, Kolpakov FA, Popov DV. Transcriptomic Signatures and Upstream Regulation in Human Skeletal Muscle Adapted to Disuse and Aerobic Exercise. *Int J Mol Sci.* 2021;22(3).

72. Pourteymour S, Eckardt K, Holen T, Langleite T, Lee S, Jensen J, et al. Global mRNA sequencing of human skeletal muscle: Search for novel exercise-regulated myokines. *Mol Metab.* 2017;6(4):352-365.

73. Dickinson JM, D'Lugos AC, Naymik MA, Siniard AL, Wolfe AJ, Curtis DR, et al. Transcriptome response of human skeletal muscle to divergent exercise stimuli. *J Appl Physiol (1985).* 2018;124(6):1529-1540.

74. Yoon HG, Oh D, Noh JM, Cho WK, Sun JM, Kim HK, et al. Machine learning model for predicting excessive muscle loss during neoadjuvant chemoradiotherapy in oesophageal cancer. *J Cachexia Sarcopenia Muscle.* 2021;12(5):1144-1152.

75. Kim J, Han SH, Kim HI. Detection of sarcopenic obesity and prediction of long-term survival in patients with gastric cancer using preoperative computed tomography and machine learning. *J Surg Oncol.* 2021;124(8):1347-1355.

76. Anyene I, Caan B, Williams GR, Popuri K, Lenchik L, Giri S, et al. Body composition from single versus multi-slice abdominal computed tomography: Concordance and associations with colorectal cancer survival. *J Cachexia Sarcopenia Muscle.* 2022;13(6):2974-2984.

Chapter 4 Competing Endogenous RNA (ceRNA) integrative networks for mechanistic insights in human skeletal muscle from patients with cancer

4.1 Introduction

This chapter focuses on post-transcriptional gene regulation, mediated by multilayered RNA crosstalk networks, including protein-coding mRNA and non-protein-coding RNAs (ncRNAs) [1]. About 80-90% of the human genome is transcribed into non-coding RNAs (ncRNAs), suggesting that ncRNAs represent most of the human transcriptome [2]. Two classes of ncRNAs are identified. Long non-coding RNAs (lncRNAs, > 200 nucleotides to several kilobases in length), and several species of small non-coding RNAs, the most studied being miRNAs[3] (20-22 nucleotides in length). miRNAs are transcribed from the intron of the protein-coding genes, lncRNAs, or independently from miRNA clusters. Recent evidence also suggests that they could be embedded within other small non-coding RNAs such as snoRNAs [4,5] and tRNAs [6]. miRNAs are known as the master regulators of gene expression. miRNAs are transcribed as precursor molecules and their subsequent processing by endoribonucleases, such as Drosha and Dicer, in the nucleus and cytoplasm, respectively, to form the mature miRNAs. miRNAs bind to the Argonaute family of proteins (AGO) to form the RNA-induced silencing complex (RISC), a process known as RISC loading. When this complex is loaded onto the miRNA response element (MRE) on the target gene, it results in gene silencing by translational repression or degradation. Gene regulation mediated by miRNAs is through the base-pair complementarity or sequence homology of the seed sequence (2-8 nucleotide long miRNA region that determines target specificity) between miRNA and target gene [7-9]. Perfect base-pair complementarity leads to mRNA degradation, and imperfect base pairing leads to translational repression. Canonical sites of gene regulation are 3'UTRs.

However, miRNAs also bind to 5'UTR and coding domain sequence (CDS) of the target genes [10-12]. miRNAs exhibit the property of redundancy and pleiotrophy, i.e., a single miRNA can regulate multiple mRNA targets, and several miRNAs can regulate a single mRNA target [7,13]. More than 60% of human protein-coding genes have at least one conserved binding site for the miRNAs and several non-conserved sites [14]. This illustrates the complex regulation and control mediated by miRNAs. A repertoire of miRNAs is known to be expressed in the skeletal muscle of experimental model systems and human biopsy specimens and are termed myomiRs [15]. Dysregulated expression of miRNAs has been reported in the experimental model systems [16,17] contributing to the pathophysiology of cancer-associated muscle wasting and in the muscle biopsies obtained from *rectus abdominis* and *vastus lateralis* muscle of patients with cancer cachexia from independent research groups, including our lab group [18-20].

lncRNAs, on the other hand, are a diverse class of molecules regulating gene expression via epigenetic, transcriptional, and post-transcriptional levels [21-24]. Their interacting partners could be DNA, RNA, or proteins [25]. Most lncRNAs share similar biogenesis with mRNAs; lncRNAs are also transcribed by RNA polymerase II (Pol II), are often capped by 7-methyl guanosine at their 5' ends, and are polyadenylated at the 3' ends, and spliced similar to mRNAs [26]. lncRNA loci are encoded from intronic regions and intergenic regions, from both strands of DNA in a sense as well as antisense orientation. Even though the lncRNAs are lowly expressed than the protein-coding counterparts, their expression patterns are tissue-specific and developmentally regulated. The mechanism of action of lncRNAs is influenced by their subcellular localization in cytoplasmic and nuclear compartments.

Recent reports and experimental evidence pinpoints the importance of the post-transcriptional regulatory mechanisms wherein crosstalk between RNA species act by sponging or sequestering

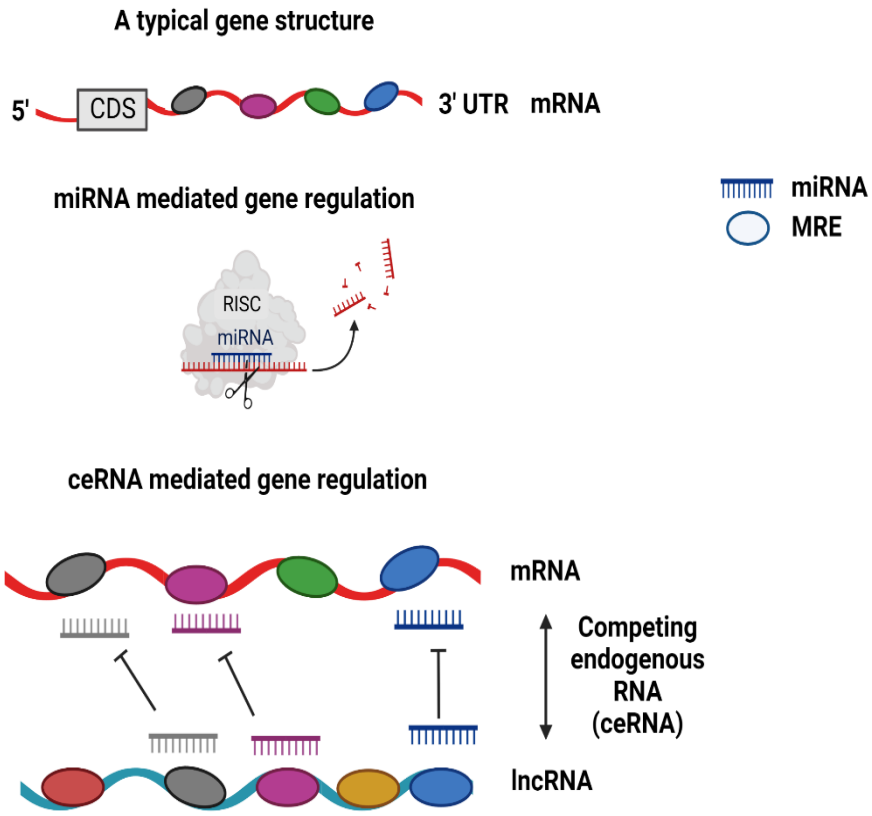
miRNAs [27], thereby releasing the inhibitory action of miRNAs on the target mRNAs. The RNA molecules that compete for the miRNA binding sites are referred to as competing endogenous RNA (ceRNA) [28-30]. The effect of miRNA competition on regulating protein-coding and non-coding RNAs represents a multilayered complexity[31,32] of ceRNA crosstalk that has large-scale implications and may, therefore, contribute to the pathophysiological basis of skeletal muscle wasting. A schematic representation of ceRNA is depicted in **Figure 4.1**. Several lncRNAs acting as ceRNA have been characterized in the skeletal muscle of experimental model systems in the lncRNA-miRNA-mRNA crosstalk networks, regulating muscle differentiation and regeneration process. Myogenesis, or the process of mature myofiber formation, is a highly regulated process where the progenitor cells proliferate to myoblast that undergoes differentiation and fuse to form multinucleated myofibers [33]. The functional role of lncRNAs in experimental model systems (i.e., animal models and C2C12 myoblast proliferation and differentiation models) to understand skeletal muscle biology is well-defined [34-44].

The expression profiling and crosstalk mechanisms contributing to higher-order integrative gene networks [45] are important to understand overall regulation. Such an approach to offer mechanistic insights into the pathophysiology of muscle wasting in patients with cancer have not yet been elucidated in human skeletal muscle.

Hypothesis: I hypothesize that the competing endogenous RNAs (ceRNAs) crosstalk mechanisms characterized by the triplet (lncRNA-miRNA-mRNA) could provide insights into lncRNA-mediated gene and pathway regulation within the skeletal muscle of patients with cancer.

***Objectives*:** (i) to perform ceRNA analysis and the subsequent network analysis for identifying lncRNA-miRNA-mRNA triplet interacting partners in males and females and (ii) to perform IPA pathway analysis of the ceRNA-mediated gene targets.

In this study, I utilized the expression profiles of mRNA, lncRNA, and miRNA from the skeletal muscle of patients with cancer, described in this thesis in the previous chapter (chapter 3). I subjected the Differentially Expressed (DE) lncRNAs, miRNAs, and mRNAs from the two molecular subtypes in males and females (FDR < 0.05 and fold-change >1.5) to the competing endogenous RNAs, network, and pathway analysis.



MRE: miRNA Response Element; **miRNA**: microRNA; **mRNA**: messenger RNA; **lncRNA**: long non-coding RNA; **RISC**: RNA induced silencing complex; **ceRNA**: Competing endogenous RNA; CDS: Coding sequence

Figure 4.1 Schematic representation of ceRNA crosstalk mechanism

A typical gene structure with 5'UTR, 3'UTR, and coding domain sequence is shown in the **top** panel consisting of MREs, wherein the multicolored ovals represent MREs, and mRNA is represented with a single strand. The **middle** panel represents the canonical role of miRNA-mediated gene regulation. , wherein miRNA bound to the RNA-induced silencing complex (RISC) targets the 3'UTR and leads to translational repression or mRNA degradation. miRNAs are represented as single-stranded short-chain molecules. The **bottom** panel represents the ceRNA-mediated gene regulation wherein ceRNA, which in the present case scenario is represented by lncRNA, competes with the mRNA for MREs and hence inhibits the binding and effect of miRNAs to their corresponding target genes. Created with BioRender.com.

4.2 Methods[§§]

4.2.1 Bioinformatic target prediction of miRNA binding sites in lncRNA and mRNA[***]

The expression profiles of miRNAs, lncRNAs, and mRNAs were generated from *rectus abdominis* muscle biopsy specimens (refer to chapter 3). In silico target prediction of miRNA binding sites were performed individually for lncRNA (full-length non-coding RNAs) and mRNAs (3'UTR regions). DE miRNAs, mRNAs, and lncRNAs within the sex-specific subtypes were subjected to in silico target prediction. Target prediction tools, such as miRanda [46], were used to determine miRNA binding sites in DE lncRNAs and mRNAs within the datasets. Functionally validated target predictions for mRNAs and lncRNAs were also accessed from publicly available databases such as Tarbase v8.0 [47], lncBase v3.0 [48,] and mirtarbase [49], which utilize different prediction algorithms from expression datasets from tissues and cell lines. To identify skeletal muscle-specific genes (mRNAs) as targets, I overlapped mRNAs from male and female datasets in this study cohort with the gene targets from the publicly available databases.

4.2.2 Competing Endogenous RNA (ceRNA) analysis

R package GDCRNAtools [50] was used to perform ceRNA analysis. Competing lncRNA-mRNA pairs were determined using three criteria as follows: (i) Hypergeometric test was performed to determine if DE lncRNA-mRNA pairs share a significant number of miRNA binding sites as follows:

$$P = 1 - \sum_{k=0}^{m} \frac{\binom{K}{k}\binom{N - K}{n - k}}{\binom{N}{n}}$$

[§§] Study was conceptualized by Bhumi Bhatt and Dr. Sambasivarao Damaraju.
[***] Bioinformatic analyses were performed by Bhumi Bhatt.

m = number of shared miRNAs between lncRNA-mRNA competing pairs

N = total number of miRNAs

n = number of miRNAs targeting the lncRNA

K = number of miRNAs targeting mRNA

(ii) For lncRNA-mRNA to be competing pairs, they must be co-expressed. Therefore, I determined this using the Pearson correlation test.

(iii) To ascertain the regulatory role of miRNA on lncRNA and mRNA, regulation similarity score (regSim) and sensitivity correlation were calculated. The regulation similarity score is used to check the similarity of expression correlation between lncRNA-miRNA and mRNA-miRNA pairs. Sensitivity correlation is used to measure if the correlation between mRNA and lncRNA competing pairs is mediated by miRNA.

The filtering criteria of Hypergeometric P-value < 0.01, correlation P-value < 0.01, and regulation similarity score > 0 (wherein 0 signifies no interactions) was applied to define competing endogenous RNAs (lncRNA-miRNA-mRNA).

4.2.3 Network Analysis

The network visualization of the triplet lncRNA-miRNA-mRNA as ceRNAs was performed to identify hub lncRNAs for individual male and female datasets using Cytoscape v3.9.0 [51]. Network analysis was performed, and the top six lncRNAs with the highest node degrees and betweenness centrality were selected as the Hub lncRNAs. The regulation network of the top six hub lncRNAs and their related mRNAs and miRNAs are depicted in the figures.

4.2.4 Functional annotation and pathway analysis

The lncRNA-mediated target genes from ceRNA analysis in males and females were subjected to pathways analysis and functional annotation using Ingenuity Pathway Analysis [52] (IPA, QIAGEN Inc., https://www.qiagenbioinformatics.com/products/ingenuitypathway-analysis). P-value in IPA is calculated using Fisher's exact test and determines the probability that the association between the genes identified within a given pathway is explained by chance alone. -log(P-value) threshold of 1.3 (equivalent to a nominal p-value of 0.05) was used to define statistically significant pathways.

4.3 Results

4.3.1 Crosstalk of RNAs using Competing Endogenous RNA (ceRNA) analysis and mechanistic insights

Recent studies have accentuated the crosstalk between various classes of RNAs, such as mRNAs and lncRNAs (sponge RNAs). These RNAs harboring miRNA binding sites regulate the gene expression levels by competing for miRNA Recognition Elements (MREs). These are known as "competing endogenous RNAs" (ceRNAs) [28,53-55]. The expression and ceRNA regulatory network and their perturbations potentially contribute to a disease phenotype. However, the presence and contribution of such a complex regulatory network in human skeletal muscle remain undetermined. Each muscle biopsy specimen was subjected to next-generation rRNA depleted sequencing and small RNA sequencing to generate mRNA, lncRNA expression profiles, and miRNA expression profiles, respectively. I used the previously identified DE profiles from the molecular subtype analysis in males and females (refer to Chapter 3), to construct the competing RNA triplets. The higher-order integrative analyses could potentially aid in the identification of mechanisms of gene regulation in patients undergoing muscle loss in cancer.

I used several target prediction databases to identify miRNA binding sites for lncRNAs and mRNAs to identify the ceRNAs. The stringency cut-offs applied for ceRNA analysis was as follows: Hypergeometric P-value < 0.01, Pearson's correlation P-value < 0.01, correlation R-value > 0.7, and Regulation Similarity Score (RegSim score) > 0. A total of non-redundant lncRNA-miRNA-mRNA triplets identified in males were: 961 lncRNAs, 71 miRNAs, 2081 mRNAs. A total of non-redundant lncRNA-miRNA-mRNA triplets identified in females were: 394 lncRNAs, 39 miRNAs, and 722 mRNAs from the ceRNA analysis.

Network analysis was performed and visualized using Cytoscape v 3.9.0. We identified the top six lncRNAs based on the highest node degree and betweenness Centrality (bC) as the hub lncRNAs for males and females, respectively **(Figure 4.2 and Figures 4.3-4.5 and 4.6-4.8).** The top six hub lncRNAs among the vast repertoire of the hub lncRNAs in ceRNA analysis and their interacting partners are summarized here for illustrative purposes: 6 hub lncRNAs, 60 miRNAs, and 202 mRNAs in males **(Appendix Table A.12)**, and 6 hub lncRNAs, 31 miRNAs and 114 mRNAs in females **(Appendix Table A.13)** with the stringency cut-offs as described above.

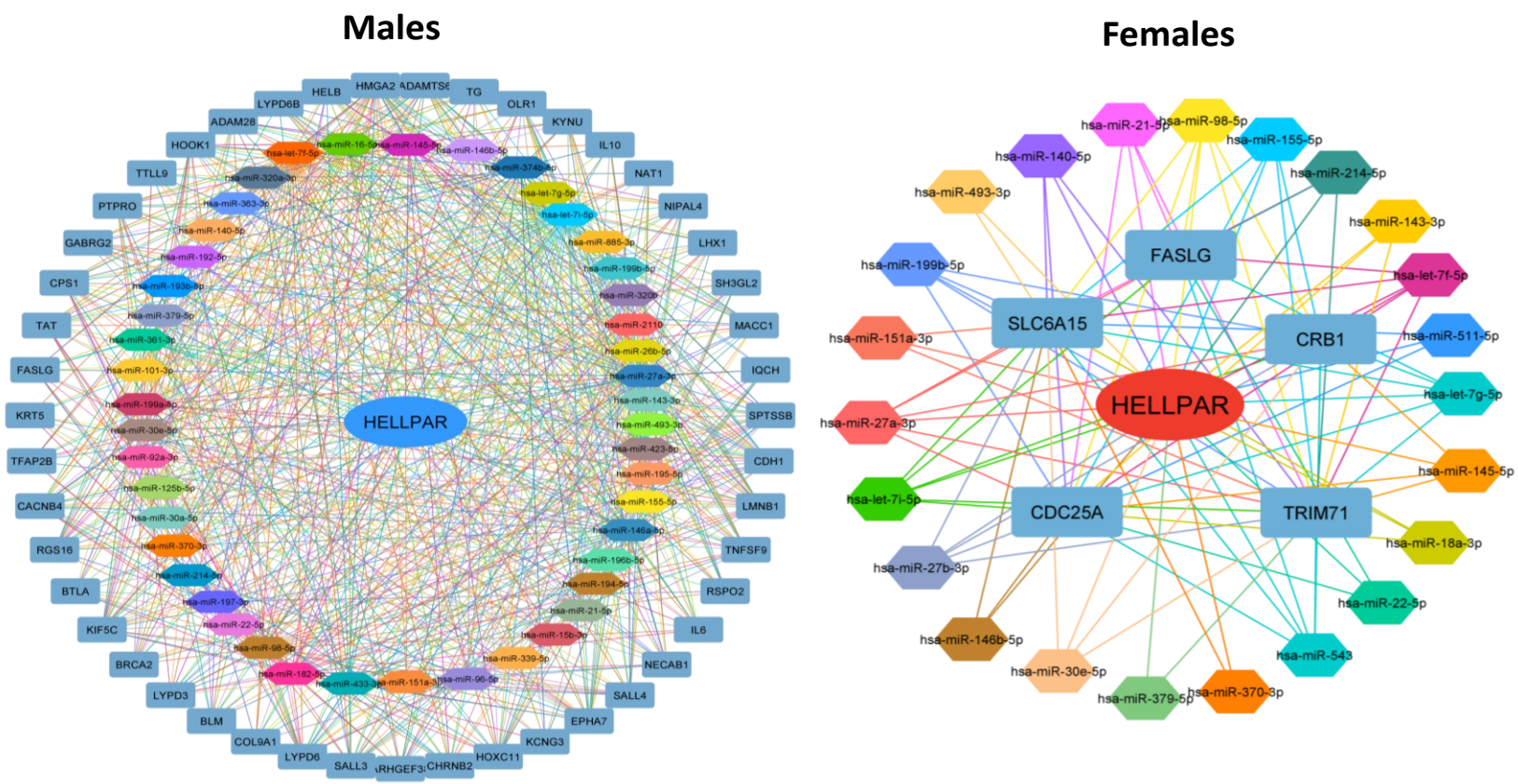


Figure 4.2 HELLPAR lncRNA and its interacting mRNA and miRNA partners in males and females
Top hub lncRNAs representing the highest node degree (i.e., the highest number of interacting miRNA and mRNA partners) from the Network analysis are summarized using Cytoscape. One of the representative hub lncRNA, namely HELLPAR lncRNA (205kb in length), is common among males (shown in blue on the left side) and females (shown in red on the right side). Sexually dimorphic patterns of expression were observed at the ceRNA level, and the figure depicts that HELLPAR interacts and regulates 98% and 62% of the unique mRNAs and miRNAs, respectively, in individual sexes. Note that the number of mRNAs and miRNAs interacting with HELLPAR are significantly different between sexes.

4.3.2 Hub lncRNA networks in males and females

Networks were generated and visualized using Cytoscape. lncRNA (oval) is illustrated at the center of each network, surrounded by the interacting partners, mRNAs (rectangle), and miRNAs (hexagon). Note that lncRNAs HELLPAR (shown in figure 4.1), CASC19, and LINC00943 were common in both sexes, but the interacting partners show a variable degree of similarities. Amongst the common lncRNAs, HELLPAR is represented in **Figure 4.2**. The other two common lncRNAs are not represented format. Unique lncRNAs in males (DLEU1, LINC00511, and XACT) and in females (AC016717.2, DLX6-AS1, and FIRRE) are represented in **Figures 4.3-4.5 and Figures 4.6-4.8**, respectively. The detailed description of the six hub lncRNAs and interacting partners, their fold-change in expression, and expression correlations are summarized in **Appendix Table A.12 (males) and Appendix Table A.13 (females),** respectively.

The top hub lncRNAs and their MREs for males are HELLPAR (205 kb and 46 MREs), CASC19 (154 kb and 22 MREs), LINC00943 (25 kb and 22 MREs), LINC00511 (350 kb and 25 MREs), DLEU1 (from autosomes, 824 kb, and 23 MREs), and XACT (X-chromosome, 442 kb, and 39 MREs). The top hub lncRNAs in males and females were higher in Subtype 1 high grade weigh loss vs. Subtype 2 low grade weight loss, and the same pattern of expression followed for their co-expressed mRNAs. The top hub lncRNAs for females are HELLPAR (205 kb and 17 MREs), CASC19 (154 kb and 14 MREs), LINC00943 (25 kb and 12 MREs), AC016717.2 (5 kb and 14 MREs), and DLX6-AS1 (from autosomes, 58 kb, and 12 MREs) and FIRRE (X-chromosome, 142 kb and 11 MREs). I observed sexual dimorphism at the ceRNA level as well. Although three hub lncRNAs were found to overlap between males and females, their interacting miRNAs and mRNAs showed differences. I also observed sex-specific differences in gene co-expression patterns, with males exhibiting higher gene co-regulation than females. For example, hub lncRNA

HELLPAR despite their average co-expression correlation r value: 0.85 in males and 0.8 in females, is comparable, whereas mRNA level expression fold-changes (upregulated) showed differences. The average fold-change was 3.5 and 2.9 in males and females, respectively. Their corresponding number of mRNAs expressed and regulated through lncRNA showed differences; in males (n=43 mRNAs, fold-change range: 1.6, 10.0) and females (n=5 mRNAs, fold-change range: 1.8, 3.3), concerning the identified subtypes 1 vs. 2. This demonstrates that miRNAs alone are not contributing to the overall regulation of mRNAs, but the dynamic interactions between the RNAs bring about the gene regulatory events.

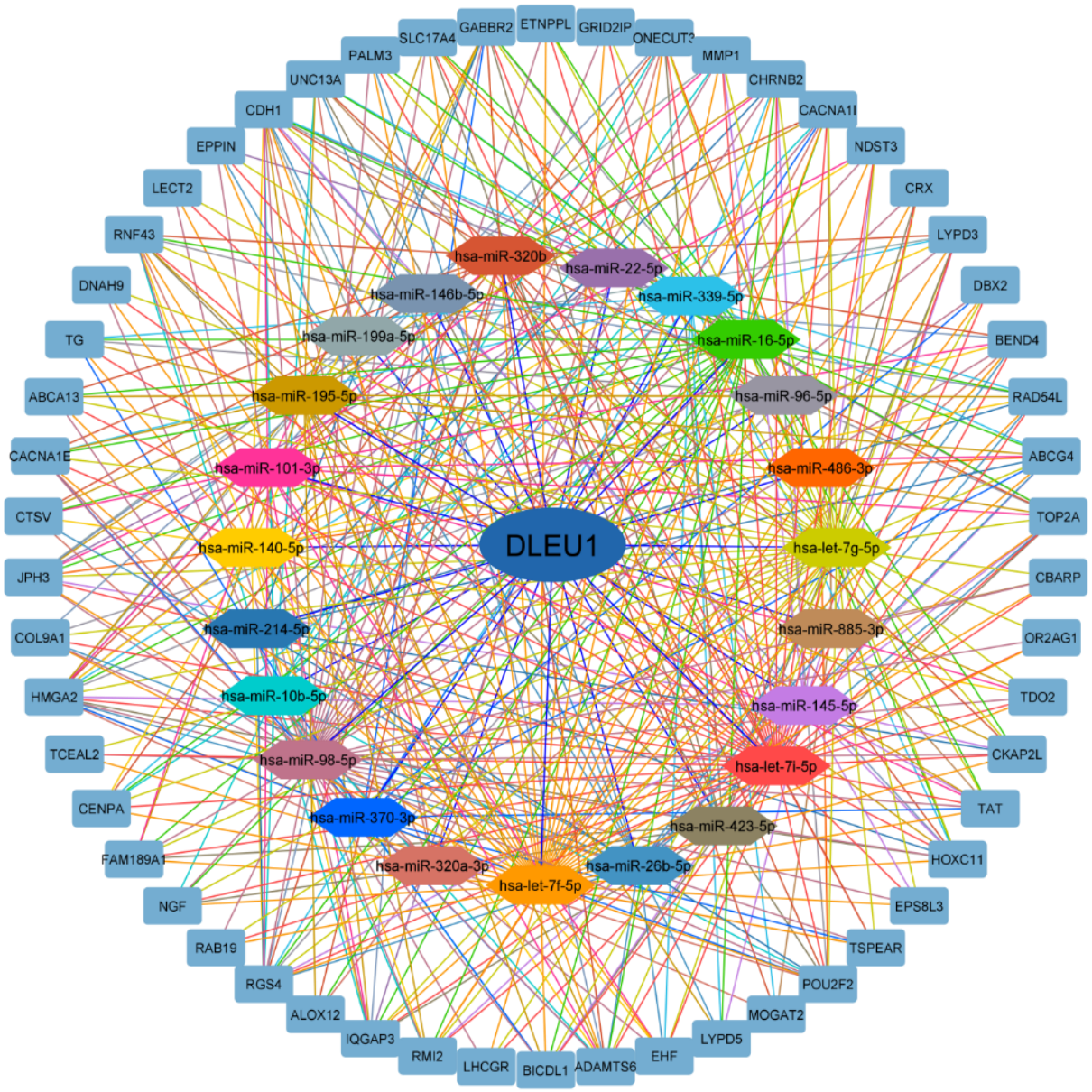


Figure 4.3 DLEU1 lncRNA and its interacting mRNA and miRNA partners in males (unique)

The figure depicts the network diagram generated from Cytoscape. DLEU1 (lncRNA deleted in lymphocytic leukemia), lncRNA (center, oval), mRNA (hexagon), and mRNA (rectangle) are described. Edges are denoted with the lines (interactions), and nodes are denoted in different shapes and colors for different RNA types and annotations.

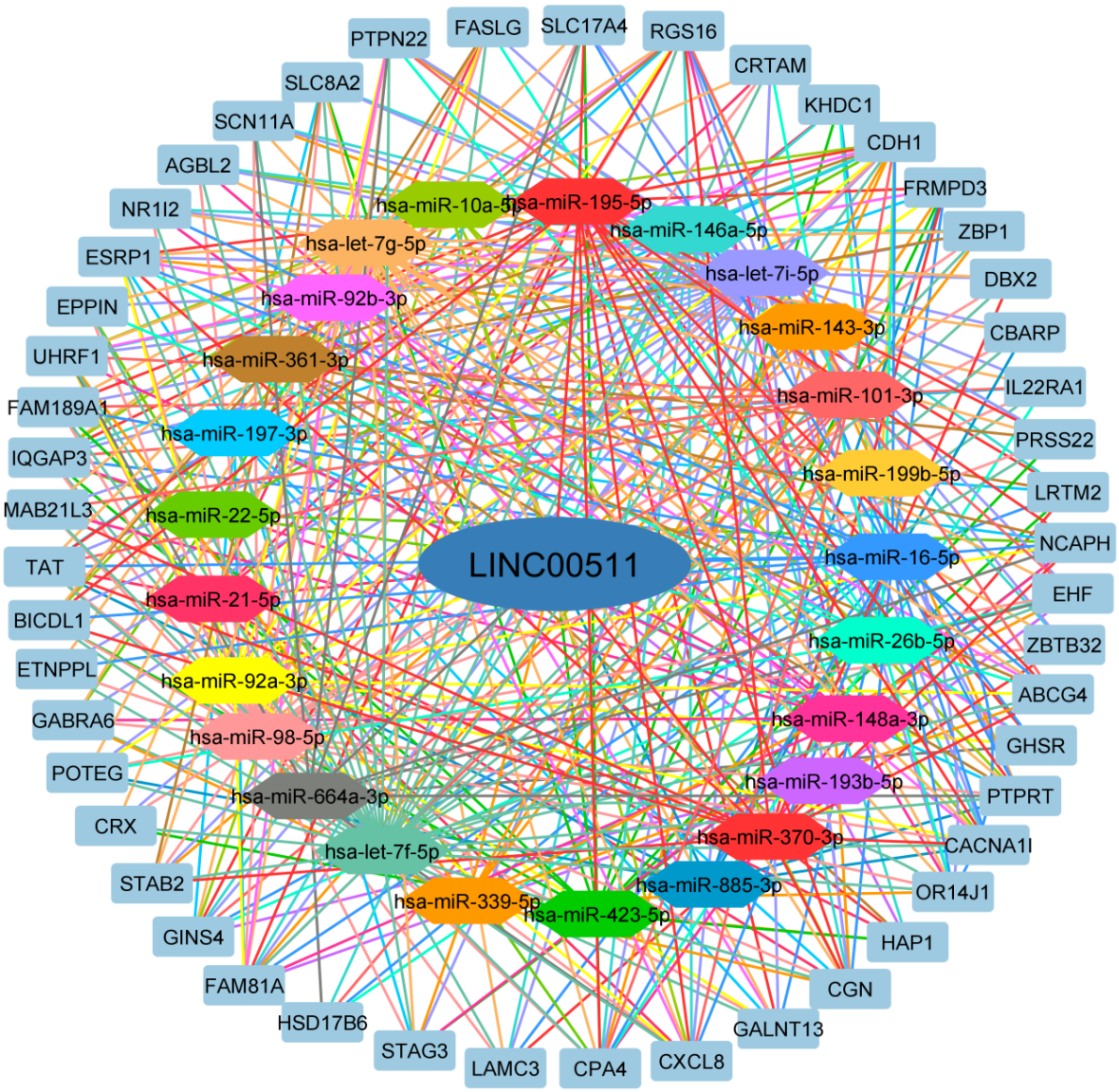


Figure 4.4 LINC00511 lncRNA and its interacting mRNA and miRNA partners in males (unique)
The figure depicts the network diagram generated from Cytoscape. LINC00511 (long intergenic non-coding RNA 00511), lncRNA (center, oval), mRNA (hexagon), and mRNA (rectangle) are described. Edges are denoted with the lines (interactions), and nodes are denoted in different shapes and colors for different RNA types and annotations.

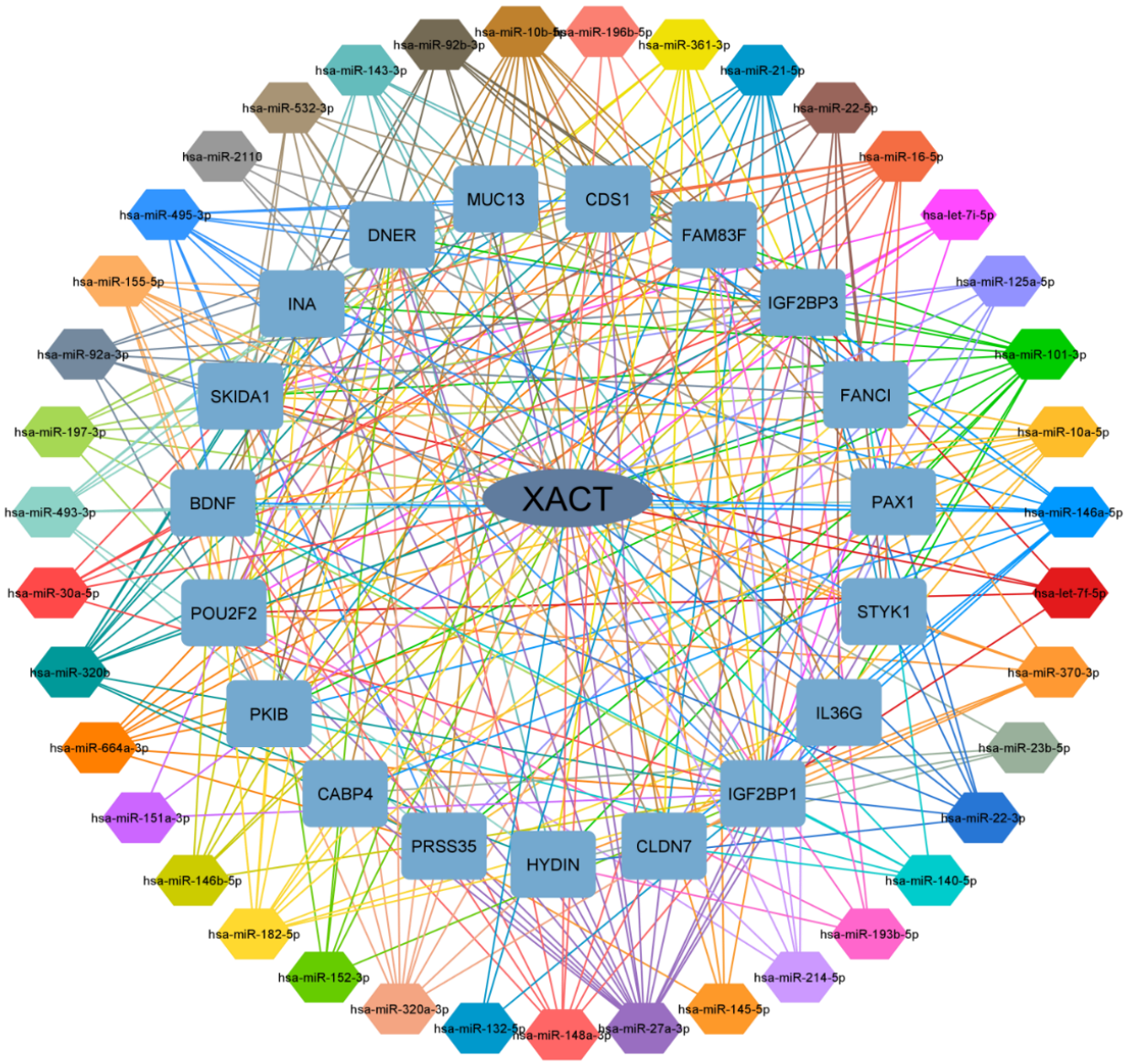


Figure 4.5 XACT lncRNA and its interacting mRNA and miRNA partners in males (unique)
The figure depicts the network diagram generated from Cytoscape. XACT (X active specific transcript) lncRNA (center, oval), mRNA (hexagon), and mRNA (rectangle) are described. Edges are denoted with the lines (interactions), and nodes are denoted in different shapes and colors for different RNA types and annotations.

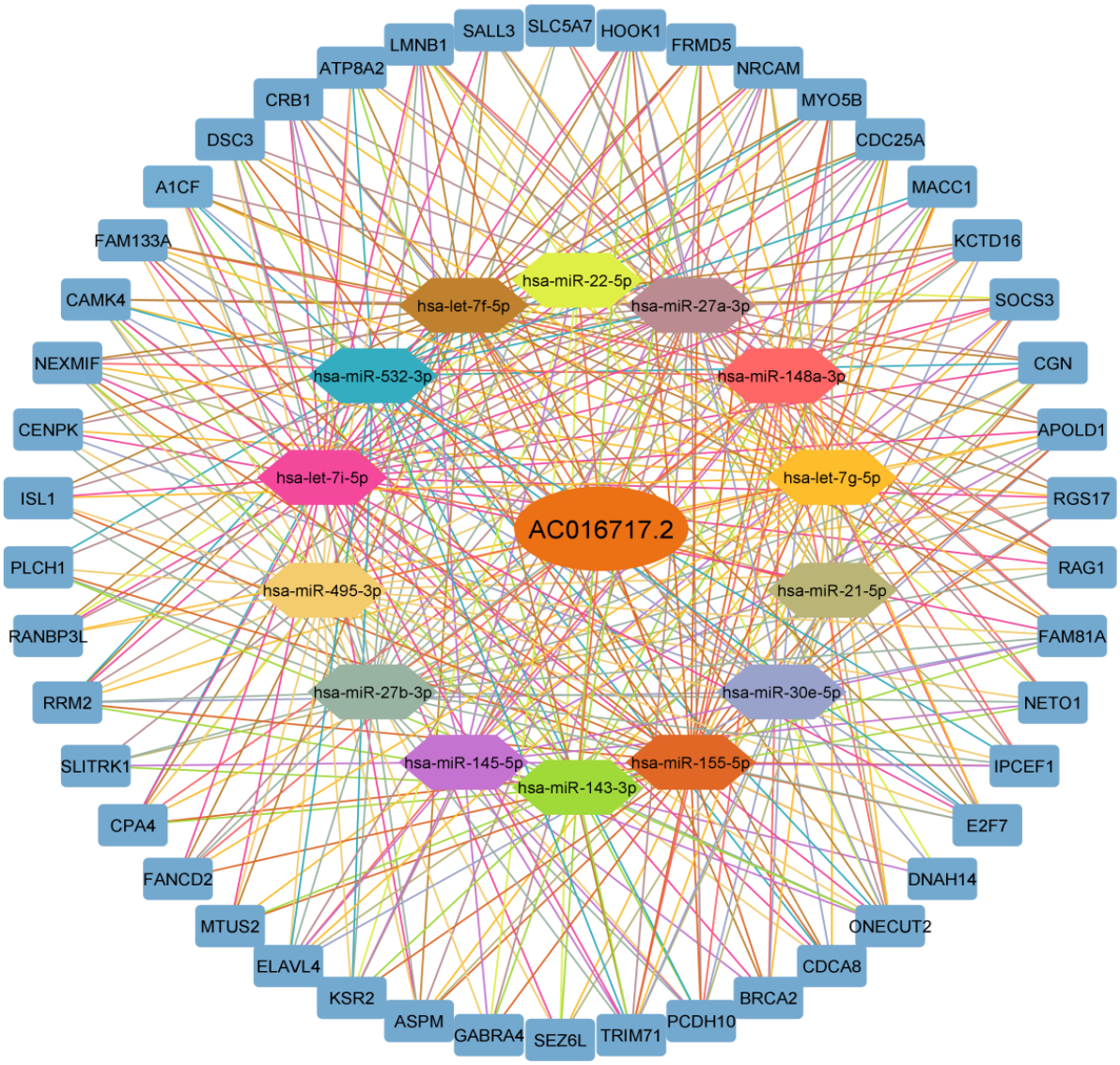


Figure 4.6 AC016717.2 lncRNA and its interacting mRNA and miRNA partners in females (unique) The figure depicts the network diagram generated from Cytoscape. AC016717.2 lncRNA (center, oval), mRNA (hexagon), and mRNA (rectangle) are described. Edges are denoted with the lines (interactions), and nodes are denoted in different shapes and colors for different RNA types and annotations.

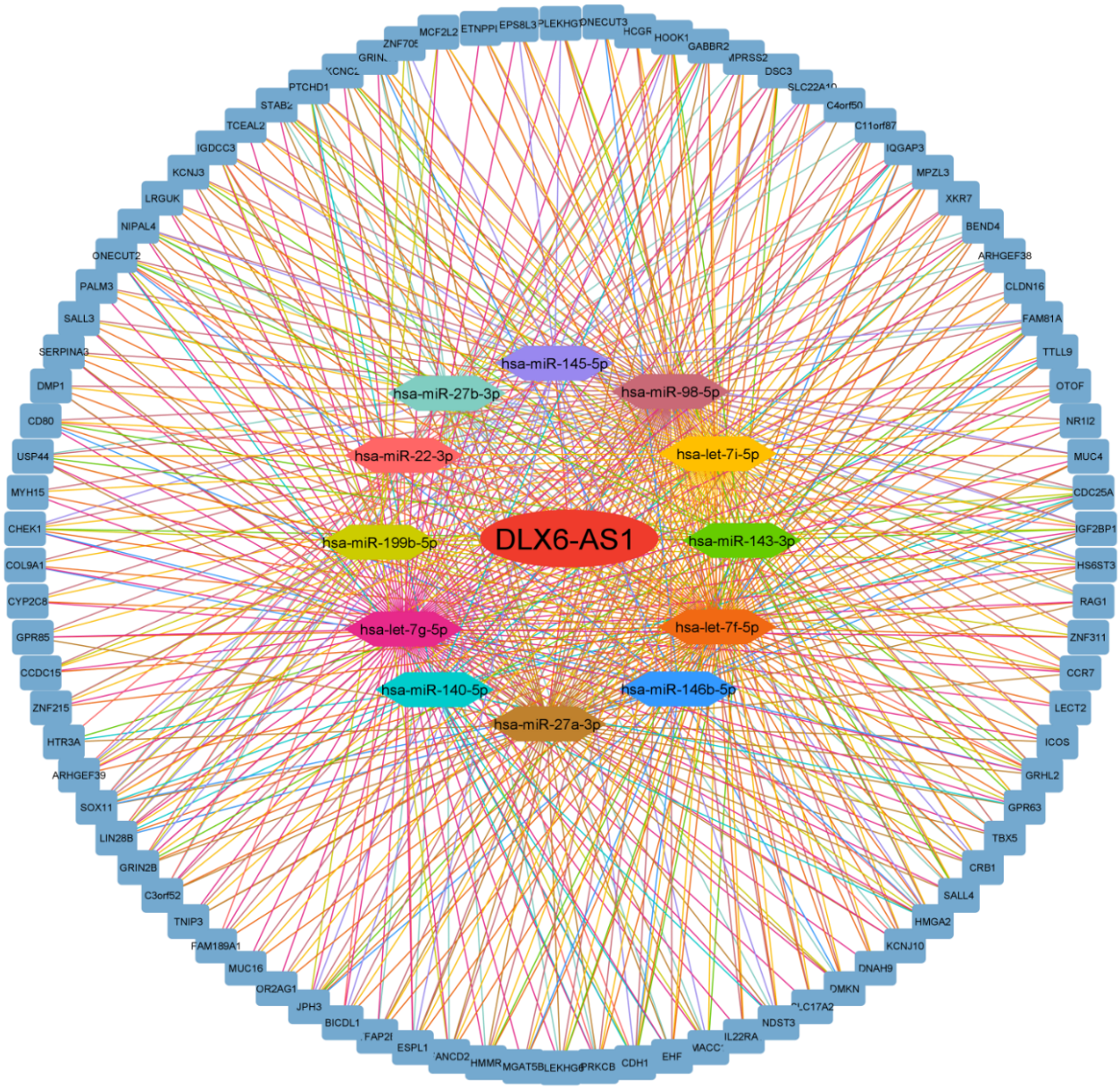


Figure 4.7 DLX6-AS1 lncRNA and its interacting mRNA and miRNA partners in females (unique) The figure depicts the network diagram generated from Cytoscape. DLX6-AS1 (Distal-less homeobox 6 antisense 1) lncRNA (center, oval), mRNA (hexagon), and mRNA (rectangle) are described. Edges are denoted with the lines (interactions), and nodes are denoted in different shapes and colors for different RNA types and annotations.

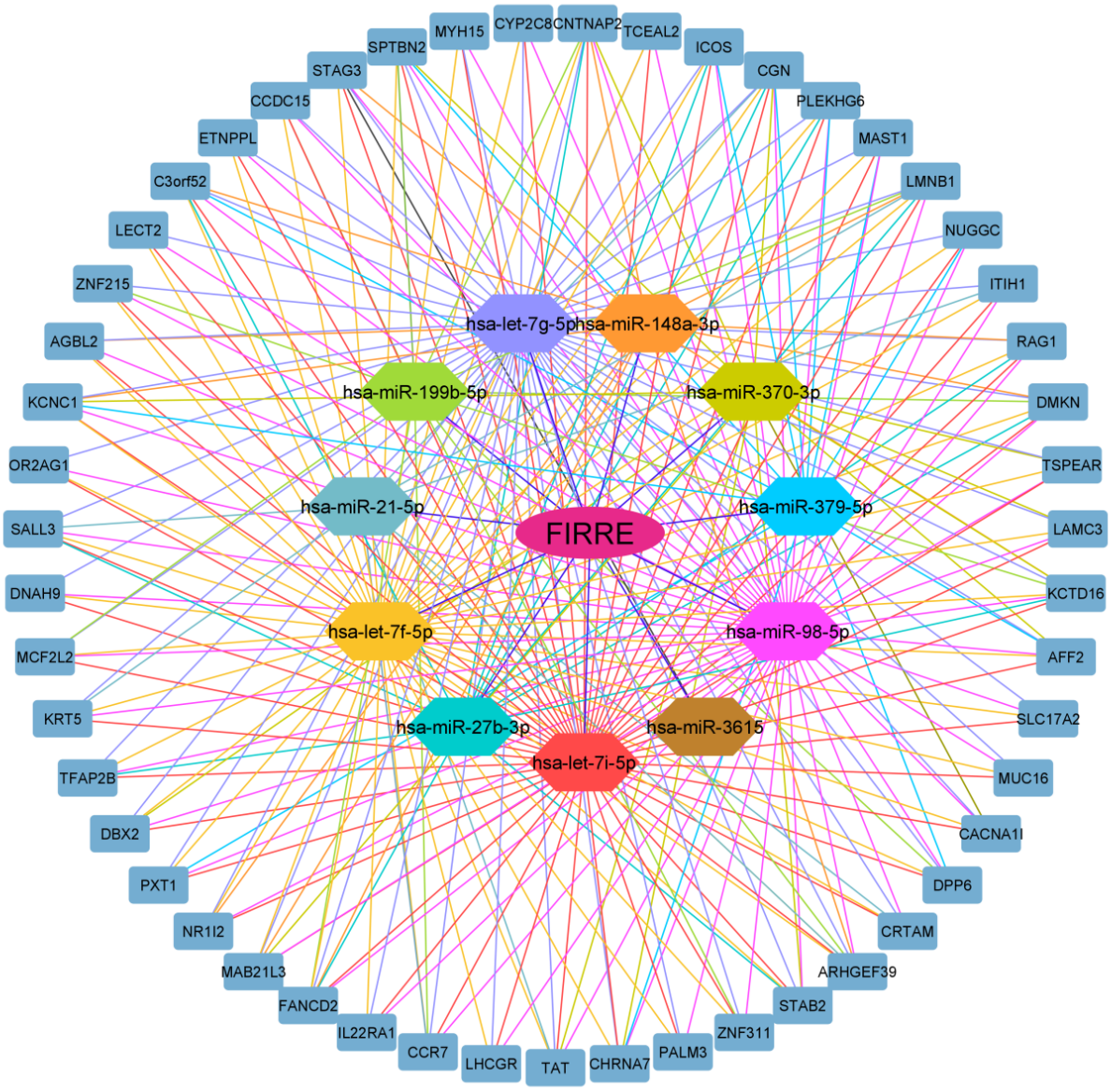


Figure 4.8 FIRRE lncRNA and its interacting mRNA and miRNA partners in females (unique)
The figure depicts the network diagram generated from Cytoscape. FIRRE (Functional intergenic repeating RNA element), lncRNA (center, oval), mRNA (hexagon), and mRNA (rectangle) are described. Edges are denoted with the lines (interactions), and nodes are denoted in different shapes and colors for different RNA types and annotations.

4.3.3 Functional annotation of mRNAs acting as competing partners with lncRNAs

There is a lack of a dedicated knowledge base for the functional annotation of human skeletal muscle tissue-specific gene expression profiles. IPA canonical pathways and biological functions. The pathway and function labels are primarily developed on non-muscle cell types. This practically means that IPA might identify a statistically significant inflammatory signature as being associated with "arthritis" or a metabolic signature associated with "hepatic cholestasis" in skeletal muscle by identifying DE elements characteristic of inflammation and metabolism, respectively, and are not enriched for skeletal muscle transcriptome. Nonetheless, this study demonstrates the utility of IPA as a knowledge base in interpreting skeletal muscle transcriptome profiles. The compelling premise of this study rests on the fact that biology is redundant in that several genes contribute to the regulation of the same pathway in different cells and tissues and exhibit pleiotropy: the same gene might contribute to multiple pathways. Hence sexual dimorphism in the regulatory mechanisms of skeletal muscle in cachexia is interrogated at the gene and pathway levels using IPA knowledge.

As described in the previous sections, the molecular subtypes of human skeletal muscle from patients were identified, and their Differentially Expressed RNA profiles were performed (refer to Chapter 3). The clinical benchmarking revealed that subtype 1 was associated with poor WL-BMI grades and low skeletal muscle mass scored as z-SMI distribution.

In this study, the premise is to understand post-transcriptional gene regulation via ceRNA analysis. mRNAs constitute <2% of the human genome, and due to the higher order regulation of mRNAs by lncRNAs in concert with miRNA, integrative ceRNA analyses were performed. The hub lncRNAs (those with the highest miRNA and mRNA interacting partners) are important regulators of gene expression since they exhibit the highest node degree, i.e., the number of interacting

miRNA and mRNA partners. Hence, ceRNA-regulated mRNAs that passed the stringency threshold were subjected to pathway analysis to identify significantly enriched pathways independently for males and females as well as the ones that regulate common molecular mechanisms.

The Venn diagram (**Figure 4.9**) indicates males' and females' overlapping and distinct pathways. Significant canonical pathways that overlapped between males and females were at 44.4% (n=44 pathways, P-value < 0.05 or -log(P-value ≥ 1.3), with 45.5% (n=45 pathways) and 10.1% (n=10 pathways) unique pathways enriched in males and females, respectively. Detailed pathway analysis results are reported in the bubble plots of the distinctive pathways in males **(Figure 4.10)** and females **(Figure 4.11)**. Although a substantial overlap of the pathways was observed between males and females (n=44), the results are reported independently as bubble plots for males **(Figure 4.12)** and females **(Figure 4.13)**, respectively. This is due to the differing count of molecules or genes within the pathways, their differing statistical significance, -log(P-value), and gene expressions within each pathway for the individual sexes.

The proinflammatory gene signatures attributed to the cachexia in rodent model systems showed upregulation in subtype 1 vs. 2 in males and females. In this chapter, the top hub lncRNAs are known to be involved in tumorigenesis in various cancer types (see ensuing discussion section for details) and were upregulated in subtype 1 (the subtype that was associated with poor WL-BMI grades and low muscle mass based on z-SMI distribution, refer to Chapter 3 for details).

The pathways unique to males (n=44, Figure 4.10 and **Appendix Table A.15**) were enriched in immune cell regulation, components important for skeletal muscle regeneration, such as extracellular matrix remodeling pathways, matrix metalloproteases, nNOS and eNOS signaling, amongst others. The pathways unique to females (n=10, Figure 4.11 and **Appendix Table A.16**)

were enriched in metabolic pathways such as the urea cycle, serotonin degradation, and components important for skeletal muscle homeostasis. A detailed report on the overlapping pathways between males and females is provided in **Appendix Table A.14**.

Overall, the inferences obtained from pathway analysis in males and females (both the common and unique pathways) suggest that the components of skeletal muscle homeostasis appear essential, and their associated perturbations could help understand the pathophysiology of cachexia. If validated in the experimental model systems, these pathways could shed light on the holistic view of homeostatic regulation rather than studying a single pathway at a time, as has been the norm thus far.

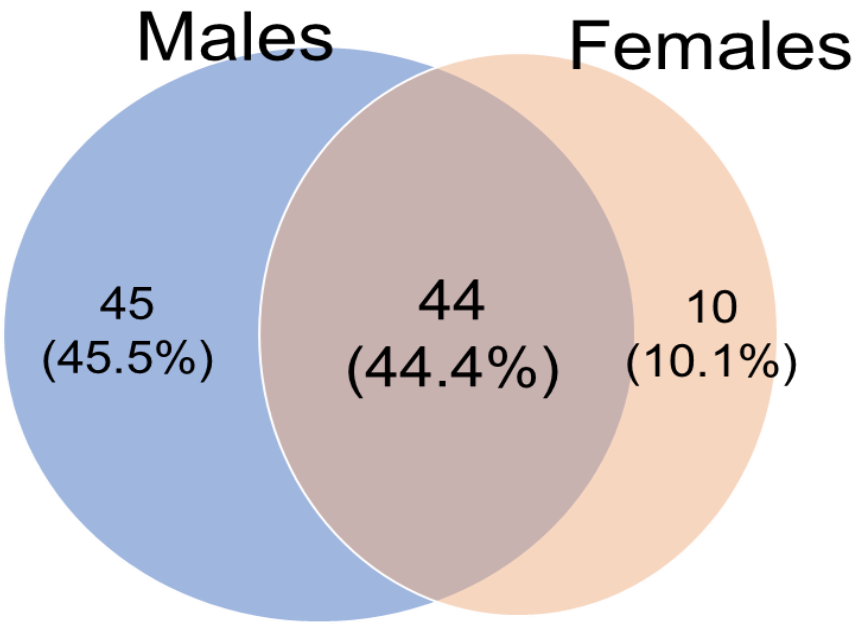


Figure 4.9 Venn diagram of canonical pathways between males and females
Sex-specific lncRNA-mediated competing mRNA targets subjected to the Ingenuity Pathway Analysis. The Venn diagram indicates the overlapping (brown intersection) and unique pathways between males (blue) and females (peach). The pathways significantly enriched in males and females at -log(P-value) ≥ 1.3 overlapped to examine the sexually dimorphic expression patterns of the canonical pathways.

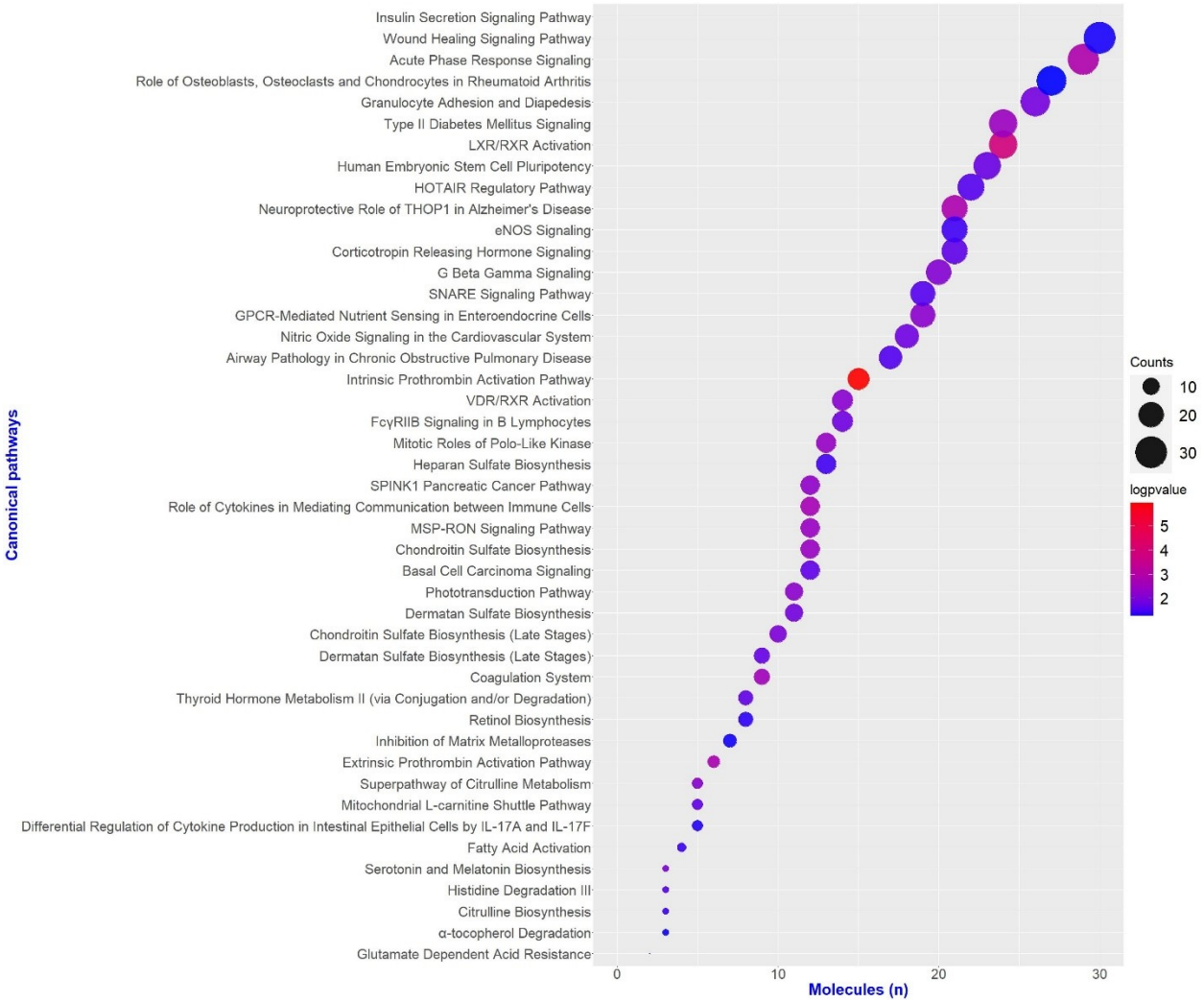


Figure 4.10 Bubble plot of unique pathways in males from ceRNA-mediated gene targets
N=45 statistically significant canonical pathways were exclusively enriched in males (Figure 4.9, left side of the Venn diagram). The bubble plot illustrates these 45 canonical pathways in males and is arranged based on the increasing order of the count of molecules or genes within a given pathway. The X-axis represents the number of molecules or genes within the individual pathways. Y-axis represents the terminology of the canonical pathways identified from IPA. The bubble size corresponds to the count of molecules within the pathways, as indicated in the legend on the bottom right side of the figure, i.e., small to large bubbles representing a low to a high number of molecules within the given pathway, respectively. The color gradient within the bubbles corresponds to the statistical significance, as indicated in the figure's legend on the top right side. Blue to red is indicative of low to high statistical significance, respectively.

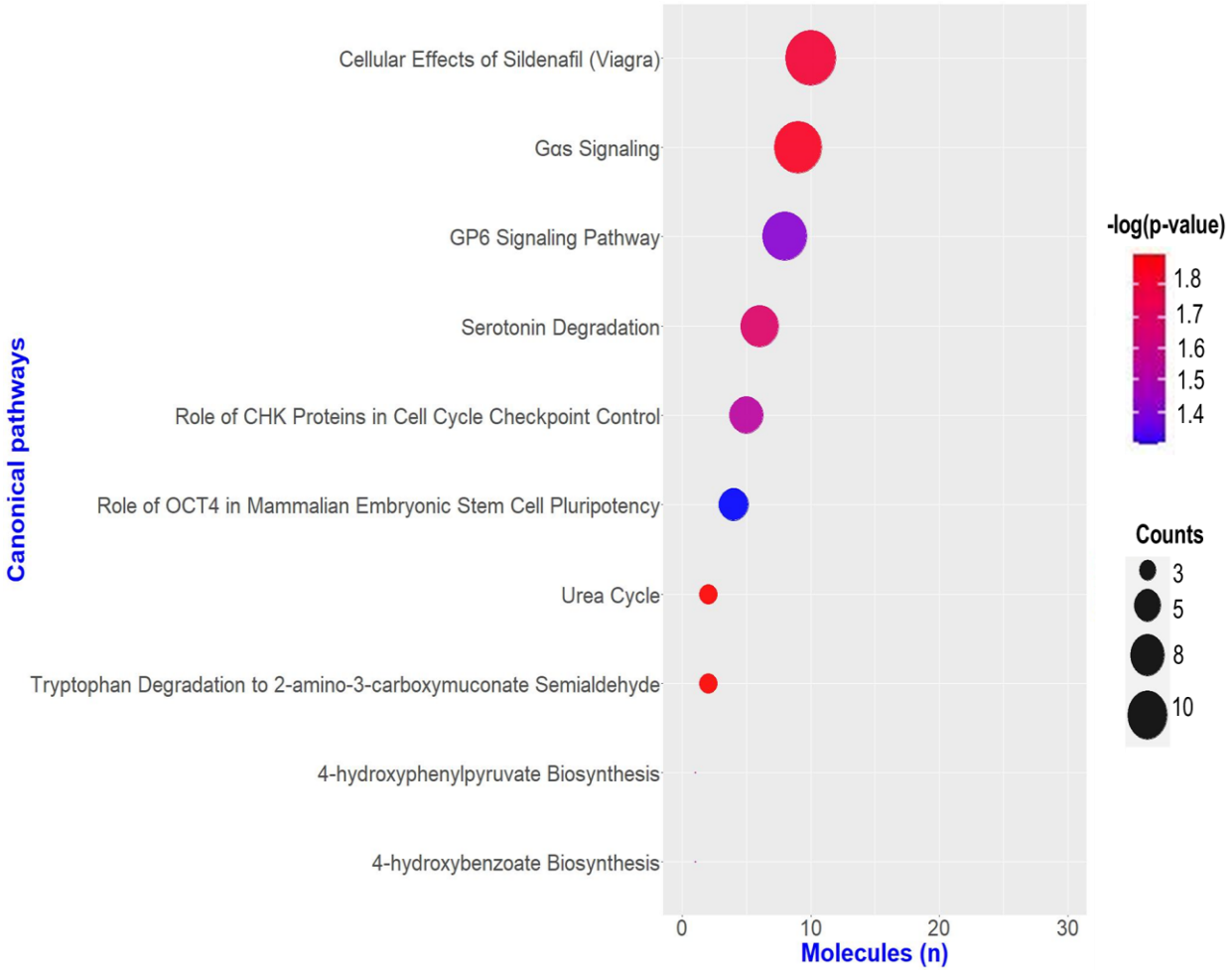


Figure 4.11 Bubble plot of unique pathways in females from ceRNA-mediated gene targets
N=10 statistically significant canonical pathways exclusively enriched in females (Figure 4.9, right side of the Venn diagram). The bubble plot illustrates these 10 canonical pathways in females and is arranged based on the increasing order of the count of molecules or genes within a given pathway. The X-axis represents the number of molecules or genes within the individual pathways. Y-axis represents the terminology of the canonical pathways identified from IPA. The bubble size corresponds to the count of molecules within the pathways, as indicated in the legend on the bottom right side of the figure, i.e., small to large bubbles representing a low to a high number of molecules within the given pathway, respectively. The color gradient within the bubbles corresponds to the statistical significance, as indicated in the figure's legend on the top right side. Blue to red is indicative of low to high statistical significance, respectively.

Figure 4.12 Bubble plot of overlapping pathways in males from ceRNA-mediated gene targets
N=44 statistically significant canonical pathways overlapped between males and females (Figure 4.9, brown intersection). Despite the considerable overlap, the bubble plots are represented separately for males and females due to the differing count of molecules within the pathways and their statistical significance.

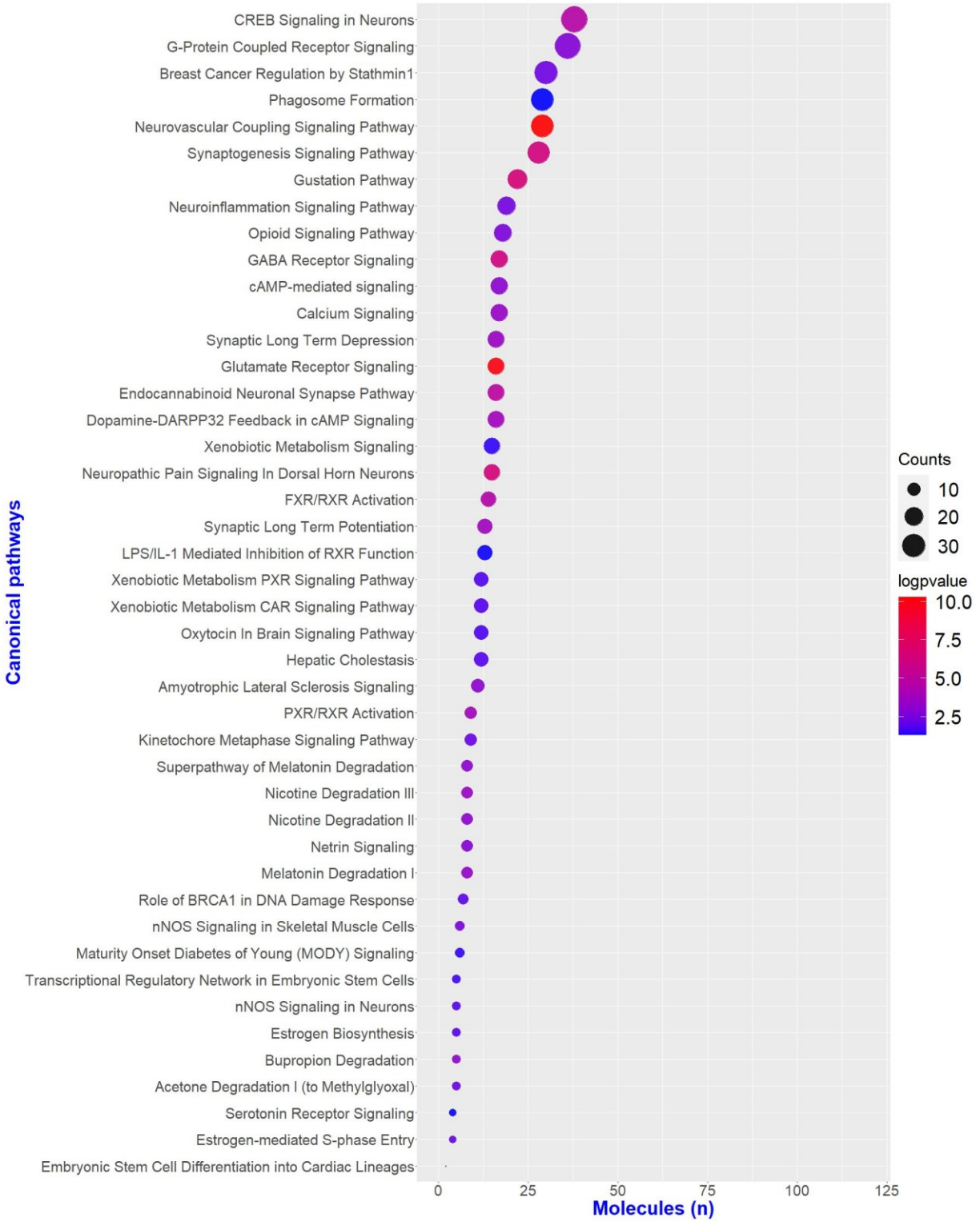


Figure 4.13 Bubble plot of overlapping pathways in females from ceRNA-mediated gene targets
N=44 statistically significant canonical pathways overlapped between males and females (Figure 4.9, overlapping Venn diagram intersection). Despite the considerable overlap, the bubble plots are represented separately for males and females due to the differing count of molecules within the pathways and their statistical significance.

4.4 Discussion

This investigation demonstrates the mechanistic insights and RNA crosstalk via the competing endogenous networks of lncRNA-miRNA-mRNA. The differential expression profiles from the identified subtypes (refer to chapter 3) in the skeletal muscle of patients with cancer were investigated to identify hub lncRNA and their interacting miRNA and mRNA partners. Mechanistic insights also helped decipher the intrinsic characteristics of molecular subtypes of skeletal muscle in conjunction with the previously annotated subtypes based on the clinical (WL-BMI grading system and ZSMI distribution) [56,57] and in silico benchmarks (at the level of differential gene expression patterns and regulation). (Refer to Chapter 3). This study solidifies the previously annotated subtypes, and the findings are encouraging, given the higher-order gene regulation contemplated by the RNA crosstalk mechanism. While the described lncRNA-miRNA-mRNA interactions are at the post-transcriptional level (in the cytosolic cellular compartment), the role of regulatory lncRNAs at transcriptional and epigenetic levels needs to be ascertained in the future.

Muscle differentiation has been one of the most studied models of ceRNAs as it can be recapitulated in vitro, and the transcription factors and signaling events in muscle growth and differentiation are highly conserved across the species [58,59]. Lnc-MD1 was the first identified ceRNA in mouse and human myoblasts and was found to act as ceRNA for miRNAs miR-133 and miR-135 to regulate the expression of MAML1 and MEF2C transcription factors that activate muscle specific gene expression [41]. Since then, many functionally important lncRNAs have been identified that play vital roles in myogenesis and skeletal muscle growth. This study identified sex-specific ceRNAs and their regulatory networks. Sexually dimorphic patterns of expression were observed at the ceRNA level. For instance, HELLPAR lncRNA was found to be one of the top

hub lncRNAs in both sexes. However, the interacting miRNA and mRNA partners were significantly different in males and females. This is the first study to identify sexually dimorphic ceRNA regulatory networks and their functional annotation in the skeletal muscle of patients with cancer.

The top hub lncRNAs identified in this study have also been investigated for their roles in tumorigenesis. Cancer susceptibility 19 (CASC19) acts as ceRNA via CASC16/miRNA-130b-3p/ZBR2 and regulates proliferative, migratory, and invasive capacities of tumor cells [60]. CASC19 was upregulated by 6 and 5.4-fold in subtype 1 vs. subtype 2 in the skeletal muscle of male and female participants. lncRNA deleted in lymphocytic leukemia (DLEU1), another tumorigenic lncRNA, was upregulated in males by 1.6 fold and is known to accelerate colorectal cancer carcinogenesis through the miR-320b/PRPS1 axis. Knockdown of DLEU1 repressed cell proliferation, migration, and invasion[61]. Distal-less homeobox 6 antisense 1 (DLX6-AS1, also called Evf2) was one of the top hub lncRNA in females and is known to be aberrantly expressed in several cancer types, and is involved in signaling pathways such as Notch signaling, Wnt/β-catenin/PI3k/Akt/mTOR, and STAT3 signaling[62]. It serves an oncogenic role and is upregulated in colorectal cancer and pancreas cancer. It is also involved in other cancer types, such as lung, renal, hepatocellular, endometrial, and gastric cancers. Hub lncRNAs involved in the tumorigenic process were found to be upregulated in subtype 1 compared to subtype 2 in males in the current study.

Given the complexity and multifactorial nature of the syndrome, it is evident that no single therapeutic modality or candidate biomarker approach is ideal. Several pathways involved in the skeletal muscle microenvironment, such as extracellular matrix regulation, nNOS signaling, and calcium signaling, amongst others, were found in the integrative analyses. Further, signaling

pathways from inflammatory, metabolic, and endocrinal pathways were also found to be statistically significant. Tumor-host interactions release proinflammatory cytokines, enhancing the series of signaling events leading to muscle wasting. Acute phase reactants/complement and coagulation proteins were increased in cancer cachexia [63]. Massart et al. (2020) characterized the changes in protein expression using proteomic analysis during cancer cachexia and identified increased muscular production of acute phase reactants in the C26 cancer cachexia model compared with vehicle control. The top canonical pathway involved in Soluble Fraction and Myofibrillar fractions of the C26 cancer cachexia model were acute phase reaction and complement and coagulation cascades. They showed that proinflammatory cytokines and glucocorticoids contributed to producing acute-phase reactants in muscle cells[63]. We found immune- and proinflammatory-related pathways significantly enriched in males. For instance, acute phase response signaling and coagulation cascade were activated in males in subtype 1. IL6 is a well-established marker contributing to cancer-associated muscle wasting [64]. One of the diverse effects of IL6 is that it promotes a coagulation cascade. Fibrinogen, an acute phase reactant that is a central player in coagulation, is elevated in the C26 model of cachexia. Reddel et al. (2017) found a mechanistic link between cachexia and thrombosis. They discovered that thrombin generation and coagulation parameters were partially but not completely corrected by blocking tumor-derived IL6 in C26 mice[65]. The expression levels of IL6 mRNA were higher by 2.17 and 7.28 in males and females, respectively, in subtype 1 vs. subtype 2. One of the earliest metabolic abnormalities associated with cancer was glucose intolerance[66]. When insulin binds to its receptor phosphatidyl-3-kinase (PI3K) and Akt, the activation of Akt leads to the suppression of Foxo and caspase 3, leading to decreased expression of atrogenes, atrogin-1, and MuRF-1. When the activity of PI3k is reduced, as it occurs in cancer cachexia and insulin resistance, inhibitory action is

released, and hence proteolysis is increased. Through this anabolic pathway, insulin controls muscle degradation and is one of the mechanisms through which insulin resistance can lead to increased protein degradation and skeletal muscle wasting. Unlike the studies pursued using experimental model systems of cachexia wherein ubiquitin-proteasome pathway and autophagy pathways were predominant in cancer-associated muscle wasting, my dataset did not identify the proteolysis pathway. These findings align with studies conducted in human skeletal muscle wherein molecular mechanisms in animal models were not recapitulated in the transcriptional changes in human muscle[67-69]. Overall, an emerging theme was the involvement of pathways in the skeletal muscle microenvironment that seemed perturbed and were activated in subtype 1. In addition, pathways (not individual genes due to the redundancy of their functions) contributing to insulin resistance were enriched in subtype 1. These pathways enrichment in subtype 1 reveal a relatively poorer group than subtype 2 regarding cachexia severity than subtype 2.

Cachexia-induced experimental models are highly controlled systems. They do not fully recapitulate the complexities, comorbidities, and wasting phenomenon experienced in the muscle wasting of cancer patients in a clinical setting [69]. Nevertheless, most of the understanding of individual pathways and mediators was derived from these model systems and is thus helpful to collate the findings from this study undertaken using whole transcriptome profiling of human skeletal muscle. This is the first study to provide the entire transcribed genome within the skeletal muscle of patients with cancer, independent of any assigned weight loss or CT-based muscle parameter-generated classification. In contrast to lncRNA-mediated pathway regulation in males, unique pathways in females were pertaining to Gαs Signaling, urea cycle, GP6 Signaling Pathway, Tryptophan, and serotonin degradation.

The overall perturbations and overlap of canonical pathways between males and females was 44.4% (n=44 common pathways between males and females). Of note, despite the pathway overlap between the sexes, the gene overlap within the common pathways was 33.3% (n=146 common genes of the 293 total non-redundant genes), with a higher number of genes uniquely present in male pathways at 62.9% (n=276) and 3.9% (n=17) in females pathways.

A comparison of the pathways regulated by DE mRNAs (Chapter 3) between the subtypes (1 vs. 2) in males with the ceRNA analysis (mRNAs, Chapter 4) revealed an overlap of 66.1% (n=72 common canonical pathways). The number of unique pathways, 18.3% (n=20 canonical pathways) and 15.6% (n=17 canonical pathways), were found in DE mRNAs vs. ceRNAs (mRNA analysis), respectively. A similar comparison between subtype 1 vs. 2 in females at the DE mRNAs (Chapter 3) and ceRNA analysis (mRNAs, Chapter 4) revealed an overlap of 44.4% (n=40 common canonical pathways). The number of unique pathways was 4.4% (n=4 canonical pathways), and 51.1% (n=46 canonical pathways) were found in DE mRNAs vs. ceRNAs (mRNA analysis), respectively.

As alluded to in Chapter 3, some sex-related differences at the pathway level and the networks likely reflect the differences in sample size with males at n=48 vs. females at n=36 and the number of lncRNAs, miRNAs, and mRNAs identified. If such differences could be ascribed to sample size, independent analysis at a higher sample size is needed to replicate and validate the findings. As such, this could be noted as a potential limitation to the study findings.

Overall, the common pathways in males and females revealed pathways involved in immune cell regulation, metabolism-related pathways, and pathways important for maintaining skeletal muscle homeostasis. Although IPA is a pathway and functional annotation database, it should be noted that it is not skeletal muscle-specific. Hence, the genes regulating several pathways involved in

other organ systems are not discussed. The expression levels of the proinflammatory cytokines such as IL1β, TNFα, and IL6 were elevated in the skeletal muscle subtype 1 vs. subtype 2, indicative of the higher expression of these mediators in subtype 1. Overall, the perturbations of these mediators and canonical pathways involved in the above processes were enriched and predicted to be activated. Hence, the annotation of subtypes based on the mechanistic insights denote it to be associated with muscle wasting in subtype 1 in patients with cancer, i.e., subtype 1 was concluded to be a relatively more severe group in terms of clinical benchmarks and expression profiling, and pathway level in silico benchmarking. Hence, this study provides another level of validation to the molecular subtype annotation and strengthens the findings from previous analyses reported in chapter 3.

The application of competing endogenous RNA crosstalk mechanism is not limited to lncRNAs; pseudogenes and circular RNAs can also act as competing partners for mRNAs [70,71]. The expression levels of individual RNAs within the ceRNA network can influence their cross-regulation. Since these mechanistic insights are gained from the ceRNA networks, future studies should perform knockdown or knockout (loss-of-function) using RNA interference (RNAi), antisense oligonucleotides (ASOs) or CRISPR/cas9 genome editing, and overexpression (gain-of-function)[72] experiments with the hub lncRNA that consists of the highest interacting partners. In this study, skeletal muscle transcriptome profiles were generated from biopsy specimens of patients with cancer to understand their expression dysregulation and complex integrative higher-order gene regulatory mechanisms. The expression profiles of lncRNAs and miRNAs are not limited to tissue specimens. They are also known to be expressed in the plasma and serum and are known to be transported via exosomes and hence were identified as circulating biomarkers [73,74]. Future studies should identify their expression profiles using non-invasive sampling from patients

with cancer to detect their expression regulation and to compare their expression regulation from the skeletal muscle transcriptome generated in the present study. lncRNAs and miRNAs are known regulators of skeletal muscle development, as elucidated from the studies conducted using experimental model systems in vivo and in vitro systems. Hence, exploring the expression dysregulation in a cell-line model system representative of the myogenic cascade of events and replicating the findings from the human muscle is another approach to studying the lncRNAs that could potentially serve as important markers regulating the process of cancer-associated muscle wasting.

This study aimed to determine the ceRNA mechanism via lncRNA-miRNA-mRNA ceRNA networks. piRNAs exhibit similar functions to the miRNAs [12]. However, their expression profiling and gene targets have never been attempted in human skeletal muscle from patients with cancer. I have annotated piRNAs and their expression profiles in subtypes. Future work should analyze lncRNA-piRNA-mRNA ceRNA networks and pinpoint any overlapping or distinct molecular signatures and pathways regulated by miRNA and piRNAs, respectively.

4.5 Conclusion

This is the first study to identify ceRNA-mediated lncRNA-miRNA-mRNA higher-order or multilayered crosstalk as a potential post-transcriptional regulatory mechanism contributing to the overall homeostasis of the skeletal muscle of patients with cancer. Since ceRNA profiles were generated comparing subtype 1 vs. subtype 2, the ensuing post-transcriptional regulatory mechanisms underlying the complex interplay of RNA crosstalk are elucidated in the cachexia risk group. The ceRNA analysis revealed the pathways and mediators emerging from the skeletal muscle milieu representing integrated responses from the muscle microenvironment and metabolic and immune response signals. Molecular subtype annotation from ceRNA-mediated gene

regulation supports the evidence from clinical benchmarking and mRNA expression analyses (chapter 3), confirming subtype 1 as a relatively severe group affected by cachexia based on the above systematic analytical approaches.

4.6 References

1. Cabili MN, Trapnell C, Goff L, Koziol M, Tazon-Vega B, Regev A, et al. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev.* 2011;25(18):1915-1927.
2. Carninci P, Kasukawa T, Katayama S, Gough J, Frith MC, Maeda N, et al. The transcriptional landscape of the mammalian genome. *Science.* 2005;309(5740):1559-1563.
3. Lee RC, Ambros V. An extensive class of small RNAs in Caenorhabditis elegans. *Science.* 2001;294(5543):862-864.
4. Scott MS, Ono M. From snoRNA to miRNA: Dual function regulatory non-coding RNAs. *Biochimie.* 2011;93(11):1987-1992.
5. Krishnan P, Ghosh S, Wang B, Heyns M, Graham K, Mackey JR, et al. Profiling of Small Nucleolar RNAs by Next Generation Sequencing: Potential New Players for Breast Cancer Prognosis. *PLoS One.* 2016;11(9):e0162622.
6. Krishnan P, Ghosh S, Wang B, Heyns M, Li D, Mackey JR, et al. Genome-wide profiling of transfer RNAs and their role as novel prognostic markers for breast cancer. *Sci Rep.* 2016;6:32843.
7. Lewis BP, Shih IH, Jones-Rhoades MW, Bartel DP, Burge CB. Prediction of mammalian microRNA targets. *Cell.* 2003;115(7):787-798.
8. Friedman RC, Farh KK, Burge CB, Bartel DP. Most mammalian mRNAs are conserved targets of microRNAs. *Genome Res.* 2009;19(1):92-105.
9. Bartel DP. MicroRNAs: target recognition and regulatory functions. *Cell.* 2009;136(2):215-233.

10. Lee I, Ajay SS, Yook JI, Kim HS, Hong SH, Kim NH, et al. New class of microRNA targets containing simultaneous 5'-UTR and 3'-UTR interaction sites. *Genome Res.* 2009;19(7):1175-1183.

11. Brummer A, Hausser J. MicroRNA binding sites in the coding region of mRNAs: extending the repertoire of post-transcriptional gene regulation. *Bioessays.* 2014;36(6):617-626.

12. Krishnan P, Ghosh S, Graham K, Mackey JR, Kovalchuk O, Damaraju S. Piwi-interacting RNAs and PIWI genes as novel prognostic markers for breast cancer. *Oncotarget.* 2016;7(25):37944-37956.

13. Yao S. MicroRNA biogenesis and their functions in regulating stem cell potency and differentiation. *Biol Proced Online.* 2016;18:8.

14. Ha M, Kim VN. Regulation of microRNA biogenesis. *Nat Rev Mol Cell Biol.* 2014;15(8):509-524.

15. Siracusa J, Koulmann N, Banzet S. Circulating myomiRs: a new class of biomarkers to monitor skeletal muscle in physiology and medicine. *J Cachexia Sarcopenia Muscle.* 2018;9(1):20-27.

16. Fernandez GJ, Ferreira JH, Vechetti IJ, Jr., de Moraes LN, Cury SS, Freire PP, et al. MicroRNA-mRNA Co-sequencing Identifies Transcriptional and Post-transcriptional Regulatory Networks Underlying Muscle Wasting in Cancer Cachexia. *Front Genet.* 2020;11:541.

17. Xie K, Xiong H, Xiao W, Xiong Z, Hu W, Ye J, et al. Downregulation of miR-29c promotes muscle wasting by modulating the activity of leukemia inhibitory factor in lung cancer cachexia. *Cancer Cell Int.* 2021;21(1):627.

18. van de Worp W, Schols A, Dingemans AC, Op den Kamp CMH, Degens J, Kelders M, et al. Identification of microRNAs in skeletal muscle associated with lung cancer cachexia. *J Cachexia Sarcopenia Muscle.* 2020;11(2):452-463.

19. Narasimhan A, Ghosh S, Stretch C, Greiner R, Bathe OF, Baracos V, et al. Small RNAome profiling from human skeletal muscle: novel miRNAs and their targets associated with cancer cachexia. *J Cachexia Sarcopenia Muscle.* 2017;8(3):405-416.

20. Li X, Du L, Liu Q, Lu Z. MicroRNAs: Novel players in the diagnosis and treatment of cancer cachexia (Review). *Exp Ther Med.* 2022;24(1):446.

21. Rinn JL, Chang HY. Genome regulation by long noncoding RNAs. *Annu Rev Biochem.* 2012;81:145-166.

22. Kung JT, Colognori D, Lee JT. Long noncoding RNAs: past, present, and future. *Genetics.* 2013;193(3):651-669.

23. Yao RW, Wang Y, Chen LL. Cellular functions of long noncoding RNAs. *Nat Cell Biol.* 2019;21(5):542-551.

24. Iyer MK, Niknafs YS, Malik R, Singhal U, Sahu A, Hosono Y, et al. The landscape of long noncoding RNAs in the human transcriptome. *Nat Genet.* 2015;47(3):199-208.

25. Statello L, Guo CJ, Chen LL, Huarte M. Gene regulation by long non-coding RNAs and its biological functions. *Nat Rev Mol Cell Biol.* 2021;22(2):96-118.

26. Guttman M, Amit I, Garber M, French C, Lin MF, Feldser D, et al. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature.* 2009;458(7235):223-227.

27. Ebert MS, Neilson JR, Sharp PA. MicroRNA sponges: competitive inhibitors of small RNAs in mammalian cells. *Nat Methods.* 2007;4(9):721-726.

28. Salmena L, Poliseno L, Tay Y, Kats L, Pandolfi PP. A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language? *Cell.* 2011;146(3):353-358.

29. Smillie CL, Sirey T, Ponting CP. Complexities of post-transcriptional regulation and the modeling of ceRNA crosstalk. *Crit Rev Biochem Mol Biol.* 2018;53(3):231-245.

30. Kartha RV, Subramanian S. Competing endogenous RNAs (ceRNAs): new entrants to the intricacies of gene regulation. *Front Genet.* 2014;5:8.

31. Tay Y, Rinn J, Pandolfi PP. The multilayered complexity of ceRNA crosstalk and competition. *Nature.* 2014;505(7483):344-352.

32. Figliuzzi M, Marinari E, De Martino A. MicroRNAs as a selective channel of communication between competing RNAs: a steady-state theory. *Biophys J.* 2013;104(5):1203-1213.

33. Karalaki M, Fili S, Philippou A, Koutsilieris M. Muscle regeneration: cellular and molecular events. *In Vivo.* 2009;23(5):779-796.

34. Zhang ZK, Li J, Guan D, Liang C, Zhuo Z, Liu J, et al. A newly identified lncRNA MAR1 acts as a miR-487b sponge to promote skeletal muscle differentiation and regeneration. *J Cachexia Sarcopenia Muscle.* 2018;9(3):613-626.

35. Li Z, Cai B, Abdalla BA, Zhu X, Zheng M, Han P, et al. LncIRS1 controls muscle atrophy via sponging miR-15 family to activate IGF1-PI3K/AKT pathway. *J Cachexia Sarcopenia Muscle.* 2019;10(2):391-410.

36. Dey BK, Pfeifer K, Dutta A. The H19 long noncoding RNA gives rise to microRNAs miR-675-3p and miR-675-5p to promote skeletal muscle differentiation and regeneration. *Genes Dev.* 2014;28(5):491-501.

37. Li R, Li B, Cao Y, Li W, Dai W, Zhang L, et al. Long non-coding RNA Mir22hg-derived miR-22-3p promotes skeletal muscle differentiation and regeneration by inhibiting HDAC4. *Mol Ther Nucleic Acids.* 2021;24:200-211.

38. Cai R, Zhang Q, Wang Y, Yong W, Zhao R, Pang W. Lnc-ORA interacts with microRNA-532-3p and IGF2BP2 to inhibit skeletal muscle myogenesis. *J Biol Chem.* 2021;296:100376.

39. Fu X, Li S, Jia M, Xu B, Yang L, Ma R, et al. Myogenesis controlled by a long non-coding RNA 1700113A16RIK and post-transcriptional regulation. *Cell Regen.* 2022;11(1):13.

40. Zhu M, Liu J, Xiao J, Yang L, Cai M, Shen H, et al. Lnc-mg is a long non-coding RNA that promotes myogenesis. *Nat Commun.* 2017;8:14718.

41. Cesana M, Cacchiarelli D, Legnini I, Santini T, Sthandier O, Chinappi M, et al. A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA. *Cell.* 2011;147(2):358-369.

42. Zhang ZK, Li J, Guan D, Liang C, Zhuo Z, Liu J, et al. Long Noncoding RNA lncMUMA Reverses Established Skeletal Muscle Atrophy following Mechanical Unloading. *Mol Ther.* 2018;26(11):2669-2680.

43. Lu L, Sun K, Chen X, Zhao Y, Wang L, Zhou L, et al. Genome-wide survey by ChIP-seq reveals YY1 regulation of lincRNAs in skeletal myogenesis. *EMBO J.* 2013;32(19):2575-2588.

44. Jin JJ, Lv W, Xia P, Xu ZY, Zheng AD, Wang XJ, et al. Long noncoding RNA SYISL regulates myogenesis by interacting with polycomb repressive complex 2. *Proc Natl Acad Sci U S A.* 2018;115(42):E9802-E9811.

45. Merico D, Gfeller D, Bader GD. How to visually interpret biological data using networks. *Nat Biotechnol.* 2009;27(10):921-924.

46. Enright AJ, John B, Gaul U, Tuschl T, Sander C, Marks DS. MicroRNA targets in Drosophila. *Genome Biol.* 2003;5(1):R1.

47. Karagkouni D, Paraskevopoulou MD, Chatzopoulos S, Vlachos IS, Tastsoglou S, Kanellos I, et al. DIANA-TarBase v8: a decade-long collection of experimentally supported miRNA-gene interactions. *Nucleic Acids Res.* 2018;46(D1):D239-D245.

48. Karagkouni D, Paraskevopoulou MD, Tastsoglou S, Skoufos G, Karavangeli A, Pierros V, et al. DIANA-LncBase v3: indexing experimentally supported miRNA targets on non-coding transcripts. *Nucleic Acids Res.* 2020;48(D1):D101-D110.

49. Huang HY, Lin YC, Li J, Huang KY, Shrestha S, Hong HC, et al. miRTarBase 2020: updates to the experimentally validated microRNA-target interaction database. *Nucleic Acids Res.* 2020;48(D1):D148-D154.

50. Li R, Qu H, Wang S, Wei J, Zhang L, Ma R, et al. GDCRNATools: an R/Bioconductor package for integrative analysis of lncRNA, miRNA and mRNA data in GDC. *Bioinformatics.* 2018;34(14):2515-2517.

51. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 2003;13(11):2498-2504.

52. Kramer A, Green J, Pollard J, Jr., Tugendreich S. Causal analysis approaches in Ingenuity Pathway Analysis. *Bioinformatics.* 2014;30(4):523-530.

53. Jacobsen A, Wen J, Marks DS, Krogh A. Signatures of RNA binding proteins globally coupled to effective microRNA target sites. *Genome Res.* 2010;20(8):1010-1019.

54. Ala U, Karreth FA, Bosia C, Pagnani A, Taulli R, Leopold V, et al. Integrated transcriptional and competitive endogenous RNA networks are cross-regulated in permissive molecular environments. *Proc Natl Acad Sci U S A.* 2013;110(18):7154-7159.

55. Bosia C, Pagnani A, Zecchina R. Modelling Competing Endogenous RNA Networks. *PLoS One.* 2013;8(6):e66609.

56. Martin L, Senesse P, Gioulbasanis I, Antoun S, Bozzetti F, Deans C, et al. Diagnostic criteria for the classification of cancer-associated weight loss. *J Clin Oncol.* 2015;33(1):90-99.

57. Kazemi-Bajestani SM, Mazurak VC, Baracos V. Computed tomography-defined muscle and fat wasting are associated with cancer clinical outcomes. *Semin Cell Dev Biol.* 2016;54:2-10.

58. Ballarino M, Morlando M, Fatica A, Bozzoni I. Non-coding RNAs in muscle differentiation and musculoskeletal disease. *J Clin Invest.* 2016;126(6):2021-2030.

59. Buckingham M, Vincent SD. Distinct and dynamic myogenic populations in the vertebrate embryo. *Curr Opin Genet Dev.* 2009;19(5):444-453.

60. Qu CX, Shi XC, Zai LQ, Bi H, Yang Q. LncRNA CASC19 promotes the proliferation, migration and invasion of non-small cell lung carcinoma via regulating miRNA-130b-3p. *Eur Rev Med Pharmacol Sci.* 2019;23(3 Suppl):247-255.

61. Xu D, Yang F, Fan Y, Jing W, Wen J, Miao W, et al. LncRNA DLEU1 Contributes to the Growth and Invasion of Colorectal Cancer via Targeting miR-320b/PRPS1. *Front Oncol.* 2021;11:640276.

62. Hu C, Liu K, Wang B, Xu W, Lin Y, Yuan C. DLX6-AS1: An Indispensable Cancer-related Long Non-coding RNA. *Curr Pharm Des.* 2021;27(9):1211-1218.

63. Massart IS, Paulissen G, Loumaye A, Lause P, Potgens SA, Thibaut MM, et al. Marked Increased Production of Acute Phase Reactants by Skeletal Muscle during Cancer Cachexia. *Cancers (Basel).* 2020;12(11).

64. Carson JA, Baltgalvis KA. Interleukin 6 as a key regulator of muscle mass during cachexia. *Exerc Sport Sci Rev.* 2010;38(4):168-176.

65. Reddel CJ, Allen JD, Ehteda A, Taylor R, Chen VM, Curnow JL, et al. Increased thrombin generation in a mouse model of cancer cachexia is partially interleukin-6 dependent. *J Thromb Haemost.* 2017;15(3):477-486.

66. Dev R, Bruera E, Dalal S. Insulin resistance and body composition in cancer patients. *Ann Oncol.* 2018;29(suppl_2):ii18-ii26.

67. Stephens NA, Gallagher IJ, Rooyackers O, Skipworth RJ, Tan BH, Marstrand T, et al. Using transcriptomics to identify and validate novel biomarkers of human skeletal muscle cancer cachexia. *Genome Med.* 2010;2(1):1.

68. Gallagher IJ, Stephens NA, MacDonald AJ, Skipworth RJ, Husi H, Greig CA, et al. Suppression of skeletal muscle turnover in cancer cachexia: evidence from the transcriptome in sequential human muscle biopsies. *Clin Cancer Res.* 2012;18(10):2817-2827.

69. Mueller TC, Bachmann J, Prokopchuk O, Friess H, Martignoni ME. Molecular pathways leading to loss of skeletal muscle mass in cancer cachexia--can findings from animal models be translated to humans? *BMC Cancer.* 2016;16:75.

70. Thomson DW, Dinger ME. Endogenous microRNA sponges: evidence and controversy. *Nat Rev Genet.* 2016;17(5):272-283.

71. Jayarathna DK, Renteria ME, Sauret E, Batra J, Gandhi NS. Identifying Complex lncRNA/Pseudogene-miRNA-mRNA Crosstalk in Hormone-Dependent Cancers. *Biology (Basel).* 2021;10(10).

72. Konermann S, Brigham MD, Trevino AE, Joung J, Abudayyeh OO, Barcena C, et al. Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature.* 2015;517(7536):583-588.

73. Arita T, Ichikawa D, Konishi H, Komatsu S, Shiozaki A, Shoda K, et al. Circulating long non-coding RNAs in plasma of patients with gastric cancer. *Anticancer Res.* 2013;33(8):3185-3193.

74. Tong YS, Wang XW, Zhou XL, Liu ZH, Yang TX, Shi WH, et al. Identification of the long non-coding RNA POU3F3 in plasma as a novel biomarker for diagnosis of esophageal squamous cell carcinoma. *Mol Cancer.* 2015;14:3.

Chapter 5 Molecular and functional benchmarking of the identified skeletal muscle subtypes using experimental model systems

5.1 Introduction

The current investigation focuses on determining the intrinsic molecular and putative functional characteristics of the skeletal muscle subtypes in patients with cancer. There is no prior human skeletal muscle sequencing dataset to understand the characteristics of the identified subtypes. Therefore, I adopted two independent, pragmatic approaches to ascertain the intrinsic characterization of the subtypes.

I referred to the method to understand the clinical relevance of subtypes as clinical benchmarking in Chapter 3. I have described the proportional association of the muscle subtypes with the WL-BMI grading system [1] and age- and sex- adjusted ZSMI distribution [2] as a metric to understand the clinical relevance (Chapter 3).

The complementary approach for developing a molecular metric, referred to as molecular and functional benchmarking (hereafter, the terms molecular and functional benchmarking are used interchangeably), is conceived here as an extension for the molecular characterizations of the subtypes beyond the gene-level differential expression analysis. In this preliminary analysis, the pathway level analysis is the focus rather than gene level expression differences described in chapter 3 since gene expressions are tissue-specific, exhibit pleiotropy, and show redundancy of functions regulating a diverse set of pathways.

Skeletal muscle wasting is the cardinal feature of cancer cachexia [3]. Despite several attempts to understand the underlying mechanisms of muscle wasting as a consequence of experimental

systems or human muscle biopsies, there is a lack of congruence on the mechanisms contributing to the pathophysiology of cachexia. Several decades of investigations were based on increased protein degradation, decreased protein synthesis, or a combination of both, leading to cancer-associated muscle wasting [4-11]. However, the findings remained inconsistent across the studies[10-12], as already reported in the preceding chapters.

The importance of muscle environment for maintaining skeletal muscle integrity and function in cancer-associated muscle wasting and their role in the regenerative process has not been explicitly identified in human skeletal muscle studies. Adult skeletal muscle (comprises ~ 40-60% body mass) can regenerate post-injury; therefore, skeletal muscle regeneration is a hallmark of adult muscle tissue. Skeletal muscle regeneration is a highly orchestrated process. It is mediated by the interactions of satellite cells (adult stem cells) and their microenvironment. Satellite cells are characterized by their special anatomical localization between the basal lamina and sarcolemma, present in a particular microenvironment called "niche." The satellite niche comprises cellular and acellular components such as Extracellular Matrix (ECM) Proteins, fibroadipogenic progenitors, matrix metalloproteases (MMPs), growth factors interstitial cells, endothelial cells, chemokines, and motor neurons. Satellite cells are in a quiescent or dormant state characterized by Pax7 gene expression . However, in response to muscle injury, satellite cells are activated, proliferate, and differentiate into myoblasts. Myoblasts then fuse to form multinucleated myotubes that result in the formation of mature myofibers. The relative expression of myogenic regulatory factors (MYOD1, MYOG, and MYF5) is vital for the activation of satellite cells. A subset of the activated satellite cells can self-renew and revert to a quiescent state to maintain the satellite cell pool. The regenerative process is highly coordinated and requires participation from diverse mediators in a stage-specific manner, such as the factors released from ECM, growth factors, encoded proteins

from damage-associated molecular patterns (DAMPs), and infiltrating inflammatory immune cells.

Impaired skeletal muscle regeneration could eventually contribute to muscle wasting in patients with cancer [13,14]. A detailed review of skeletal muscle regeneration in cancer cachexia from the in vitro studies of human primary myoblasts and muscle progenitor cells as well as animal studies is provided by Bossola et al [15]. Overall, different models of animal studies suggested the involvement of various steps implicating the process of skeletal muscle regeneration. Persistent high expression of Pax7 in the muscle of C26-bearing animals is noted compared to controls, whereas myogenin levels were reduced, suggestive of an impaired muscle regenerative mechanism [13,16]. Pessina et al. (2010) showed that the rectus abdominis muscle of patients with gastric cancer showed higher expression of genes involved in muscle regeneration compared to non-cancer controls [17]. They studied candidate genes such as PAX7, MYOD1 and necdin (NDN) using qRT-PCR analysis. The higher expression of these genes in the muscle of patients with cancer suggested that regeneration is not inhibited but rather activated to counteract tumor-induced muscle wasting. These studies highlighted the potential for an injured muscle (in patients with cancer in cachectic conditions) to adapt to the changes in response to tumor-induced cachectic factors to preserve muscle homeostasis. This premise needs to be tested in human skeletal muscle in patients with cancer. There is a paucity of comprehensive and systematic human skeletal muscle study analyses addressing these gaps [18,19].

***Study rationale*:** Tissue homeostasis is a dynamic process and is highly regulated through tissue-specific proliferation and terminal differentiation processes. In patients with cancer, the specific stages of muscle loss/regeneration are difficult to ascertain due to interindividual variability in the disease and muscle loss trajectory. However, in model systems, it was documented that the satellite

cells retain the regenerative ability [20,21]. Independent investigations also supported this premise in human muscle stem cells [21]. Therefore, the ongoing interest is to develop an appropriate model system (s) to assess human skeletal muscle atrophy or regenerative ability within the subtypes of muscle groups identified in patients with cancer.

Proinflammatory cytokines promote activation and proliferation of satellite cells but sustained high systemic levels of these cytokines in cancer impede the differentiation potential by several mechanisms. In these aspects, animal models of cancer-induced muscle loss differ from those observed in patients with cancer. A further lack of expression of genes to promote differentiation was also reported in animal models, further emphasizing a need to interrogate the molecular mechanisms to delineate muscle atrophy and regeneration in human skeletal muscle. Hence, molecular level and functional benchmarking may serve as a composite metric between the subtypes of patients exhibiting differences in severity of muscle loss as judged from pathway analysis. There is no literature to date addressing these objectives.

To address these gaps in the literature, two experimental model systems were developed to investigate if such model systems facilitate the molecular and functional benchmarking of muscle subtypes. The first system describes proliferating Rhabdomyosarcoma (RD) cells with the potential to differentiate and offers a model to emulate states of regenerative processes. The second system describes a rodent model to emulate the state of muscle atrophy in the presence of the tumor. I reasoned that the subtype 1 exhibiting higher severity of muscle loss as judged from the proportional association with poor WL-BMI grades and low muscle mass (z-SMI distribution of subtypes), would display characteristics of both muscle injury/repair as well as muscle atrophy features. The choice of these model systems is based on the observation that no prior reference

datasets are available to perform molecular and functional characterizations of the muscle subtypes identified in this study in patients with cancer.

Hypothesis: Adopting model systems that characterize the muscle regenerative process and atrophy, generating NGS data from these model systems and their pathway level regulations when compared with pathways identified in human skeletal muscle subtypes, offers insights into muscle homeostasis in patients with cancer.

Study-specific objectives are to (i) perform transcriptome analyses of RD cells in undifferentiated and differentiated states mimicking the process of skeletal muscle regeneration, (ii) perform transcriptome analyses of skeletal muscle of rodent model-bearing tumors mimicking the process of tumor-induced muscle atrophy, and (iii) identify overlapping and convergent pathways from model systems with that of human skeletal muscle tissue for molecular and functional benchmarking of the muscle subtypes to reveal pathways pertinent to muscle homeostasis.

5.2 Methods[†††]

5.2.1 Experimental design and analysis

The experimental design and analysis plan is presented in **Figure 5.1**. Two independent experimental model systems were generated and subjected to mRNA expression profiling analyses. The first human cell-line model system consisted of Rhabdomyosarcoma cells (RD cells) derived from a female donor in myoblast undifferentiated (n=2, proliferating) and differentiating (n=2) conditions were compared. The second model system consisted of a rodent model (female Fischer rats) characteristic of muscle wasting (rats bearing the ward colon tumor, n=5 and control animals, n=5).

The rationale for the choice of the above model systems was based on (i) limited resources to generate NGS profiling in multiple cell lines or muscle from animal models; (ii) the availability of NGS data profiled for RD cells and rat bearing the tumor and data generated from the skeletal muscle (gastrocnemius muscle); (iii) all sample processing and data generation was from the single lab limiting or eliminating the batch effects (reagents, similar NGS platforms, and sample handling levels); (iv) RD cells were extensively used as a model system for myoblast differentiation and yet, a comprehensive approach to generate NGS data was lacking in the literature. I also acknowledge that other cell lines or model systems could potentially serve to extend this premise and will be deferred for future investigations.

[†††] Study was conceptualized by Bhumi Bhatt, Dr. Vickie Baracos and Dr. Sambasivarao Damaraju.

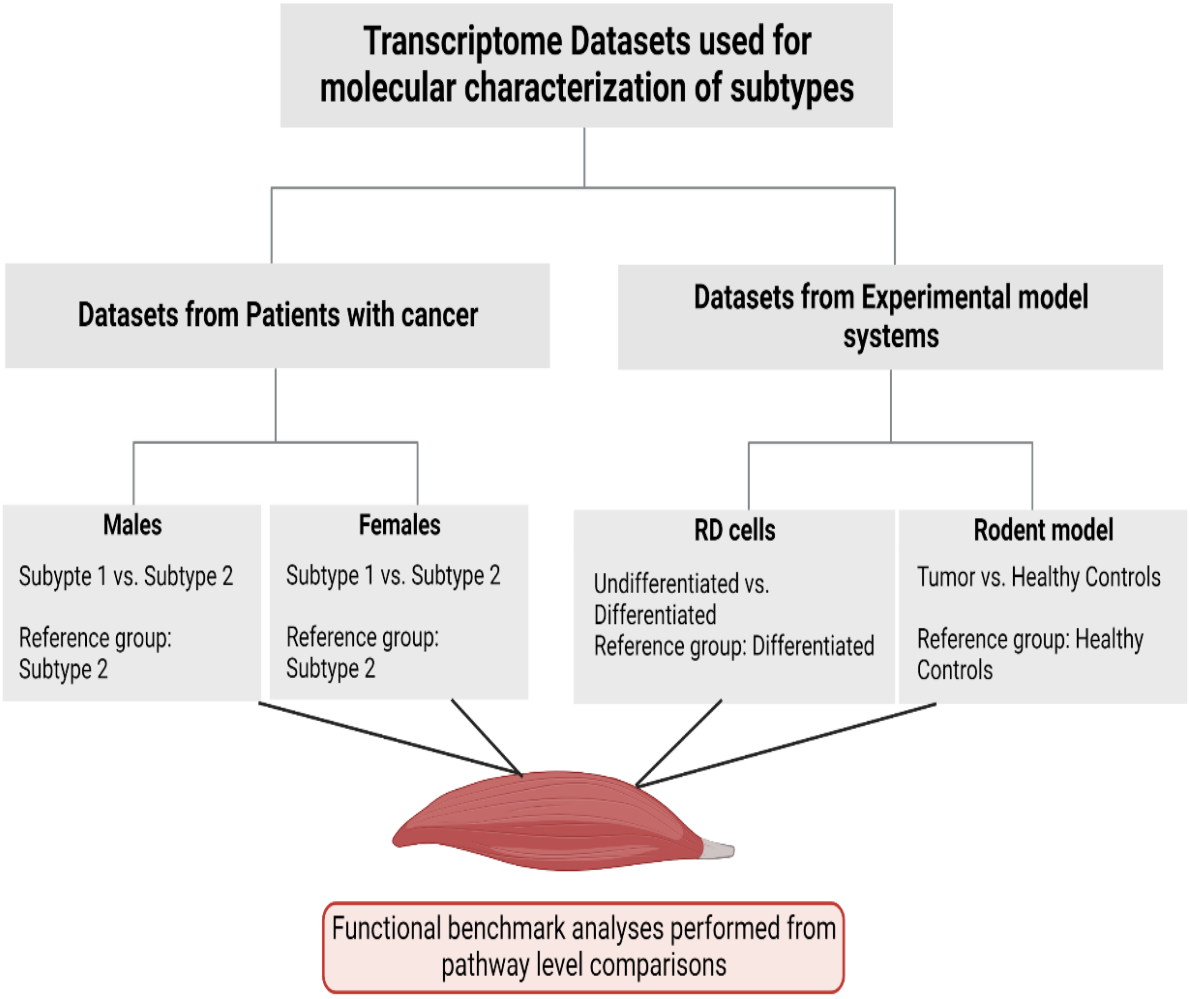


Figure 5.1 Experimental design and analysis plan to perform molecular and functional benchmark analyses of the subtypes

The schematic presents the analysis plan and the individual datasets with their respective comparisons and comparator reference groups. Pathway-level comparisons were used in the above systems to perform functional benchmarking of the molecular subtypes. Created with BioRender.com

5.2.2 Cell Culture and qRT-PCR validation of myogenic markers

RD cells were purchased from ATCC by Prof. Toshifumi Yokota, University of Alberta, Edmonton, Alberta, Canada, and the same were provided as a gift for the experiments described herein. RD cells were cultured in Dulbecco's Modified Eagle Medium (DMEM) containing 10% Fetal Bovine Serum (FBS) at 37° C in a humidified atmosphere containing 95% air and 5% CO_2. For myogenic differentiation, proliferating (undifferentiated) myoblasts were made to differentiate using a chemical trigger (12-O-tetradecanolyphorbol-13-acetate, TPA, P1585, Sigma-Aldrich). TPA was used at a final concentration of 100nM. The cells were allowed to differentiate under these conditions for 6-10 days in T175 flasks with 35 ml culture medium. Cell cultures were set up in duplicates, and cells were harvested under proliferating and differentiating conditions. Cell suspensions were centrifuged to collect the cell pellets. These were resuspended in Trizol reagent for the extraction of total cellular RNA, as described in Chapter 3 in the RNA isolation procedures from muscle tissues. The total RNA was reverse transcribed for RT-PCR analysis using miRScript Plant RT kit routinely used for RT-PCR analysis to quantify miRNAs and mRNAs. QuantiTech Primer assays and QuantiTect SYBR Green Master mix within the miRScript SYBR Green PCR kit were used for the relative quantification of canonical myogenic gene markers. GAPDH was used as an internal control. Expression fold changes in undifferentiated (proliferating) and differentiating (differentiated) states of RD cells were calculated using the $2^{-\Delta\Delta Ct}$ method [22].

5.2.3 Next-Generation RNA sequencing (NGS) of RD Cells

RNA extraction was performed using the Trizol method and Qiagen RNAeasy midi kit (Mississauga, ON). The optical density (OD) 260/280 ratio was measured using Nanodrop, and RNA integrity number (RIN) was assessed using Agilent Bio-analyzer 2100 for all the samples.

Services from the Genome Quebec facility (Montreal, Canada) were utilized for library preparation and whole transcriptome sequencing of RNA from human skeletal muscle. Briefly, pre-processing of isolated total RNA and downstream processing of the samples were as per the manufacturer's instructions summarized for rRNA depleted sequencing (RNA sequencing).

Total RNA was quantified, and its integrity (RNA integrity Number, RIN) was assessed using 5K / RNA / Charge Variant Assay LabChip and RNA Assay Reagent Kit (Perkin Elmer). Purified RNA (RNA Integrity Values range 9.5 to 9.7) was used to construct RNA-seq library preparations from biological replicates (independently cultured and isolated cell pellets). Quality indices were consistently obtained and surpassed the recommended thresholds for assessing library and sequence quality scores (see details below). rRNA was depleted from 250 ng of total RNA using QIAseq FastSelect (Human 96rxns). New England BioLabs (NEB) provided the following reagents and kits, including adapters and primers for the cDNA synthesis: NEBNext RNA First-Strand Synthesis and NEBNext Ultra Directional RNA Second Strand Synthesis Modules were used. The remaining library preparation steps were performed using NEBNext Ultra II DNA Library Prep Kit for Illumina. Libraries were quantified using the Quant-iT™ PicoGreen® dsDNA Assay Kit (Life Technologies) and the Kapa Illumina GA with Revised Primers-SYBR Fast Universal kit (Kapa Biosystems). The average size fragment was determined using a LabChip GX (PerkinElmer) instrument.

The libraries were normalized, pooled, and then denatured in 0.05N NaOH and neutralized using HT1 buffer. The pooled libraries were loaded at 225pM on an Illumina NovaSeq S4 lane using Xp protocol per the manufacturer's recommendations. The run was performed for 2x100 cycles (paired-end mode). A phiX library was used as a control and mixed with libraries at 1% level. All samples passed a Phred Quality Score of >35, as per the manufacturer's recommendations. Base-

calling was performed with RTA v3.4.4. Program bcl2fastq2 v2.20 was then used to de-multiplex samples and generate fastq reads.

5.2.4 Analysis of raw sequence files and Differential Expression (DE) of RNAs

The data analyses of raw fastq sequence files were performed using Partek Flow software v10.0.21.0929 (Copyright ©; 2018 Partek Inc., St. Louis, MO, USA) unless specified otherwise. The raw fastq files were subjected to Cutadapt [23] for the 3' adapter trimming. The trimmed reads were aligned to Human Genome (reference index hg38) using STAR aligner [24] (v2.7.3a). The generated .bam files were quantified to transcriptome using RNAs obtained from different annotation databases: Ensembl transcripts v102 was used to quantify mRNA. The features (the terminology is used interchangeably referring to RNAs) were filtered for ten read counts in 90% of samples following quantification of the reads. Differential Expression (DE) analysis of the RNAs was performed using the DESeq2 [25] R package. RNAs were considered DE at a Fold-change cut-off of 1.5 and $P < 0.05$.

5.2.5 Next Generation RNA sequencing and differential expression analysis of gastrocnemius in a rodent model of muscle wasting

The development of the preclinical animal model for cancer-associated muscle wasting has been described [26]. Briefly, the Ward colorectal carcinoma (0.05g) was subcutaneously implanted into the flank of female Fischer 344 rats aged 11-12 weeks. The tumor was subcutaneously introduced to enable assessment of the rate of tumor growth. Following the tumor implantation, animals were sacrificed on day 14, when 22.2% mean muscle cross-sectional area reduction was recorded relative to healthy control animals. RNA was isolated from Gastrocnemius muscle tissue using a MagMax-95 total RNA isolation kit (Ambion, Austin, Texas, USA) following the manufacturer's instructions. RNA-seq libraries were prepared using TruSeq stranded total RNA with Ribo-Zero

™ Human/Mouse/Rat, TruSeq stranded total RNA according to manufacturer's instructions. Total RNA (1µg per sample and a RIN value of > 8.0) was used as an input material, depleted of rRNA, and the remaining RNA was purified, fragmented, and used for cDNA synthesis. The samples were sequenced on Illumina NextSeq 500 using high throughput 2x150 nt runs (paired-end read) with the density of 35 samples per flow cell to generate 10-13 million reads per sample. Sequencing services were from PlantBiosis Ltd, Lethbridge, Alberta, Canada. Base-calling and de-multiplexing were performed using 87 Illumina CASAVA1.9 with default settings. Adapter trimming was done using Trim Galore v.0.4.1. Quality control of the sequenced reads was performed using FastQC v0.11.4. Trimmed sequences were aligned to the rat reference genome using Tophat 2.0.10 with Bowtie2. Rat Genome (Rnor 6, Ensembl) was downloaded from the iGENOME website and served as a rat reference genome. Aligned sequences were saved as .sam files, which were then converted to .bam files and used for further data processing. Data analysis of .bam files was performed using Partek Flow software. mRNAs were annotated using Ensembl Rattus norvegius Rnor-6.0.92. Differential Expression analyses were performed using DESeq2 with the Fold-change cut-off of 1.5 and P <0.05.

5.2.6 Ingenuity Pathway analysis

Differentially Expressed (DE) genes from the skeletal muscle subtypes and the two experimental model systems (RD cell and rodent models) were subjected to pathway analysis and functional annotation using Ingenuity Pathway Analysis [27] (IPA, QIAGEN Inc., https://www.qiagenbioinformatics.com/products/ingenuitypathway-analysis). In characterizing the muscle subtypes, and muscle loss in patients with cancer, relative pathway enrichment analysis were used to create a schematic of networks that are integrative and interactive to potentially infer the complexity of the muscle loss in the cancer trajectory. II used functional and/or molecular

benchmarking as terms synonymous with integrating information from an IPA platform regarding molecular functions, cellular functions, biological processes (genes contributing to pathways), interactions, and gene regulations as described (https://digitalinsights.qiagen.com/technical-support/manuals/). P-value in IPA is calculated using Fisher's exact test. It determines if the probability that the association between the genes identified within a given pathway is explained by chance alone. -log(P-value) threshold of 1.3 (equivalent to a nominal p-value of 0.05) was used to define statistically significant pathways.

5.3 Results

5.3.1 Cross-platform validation of gene markers regulating the processes of skeletal muscle regeneration

The relative quantification of genes in undifferentiated and differentiated states of the skeletal muscle regenerative process is depicted in **Figure 5.2** (top panel). PAX7, a proliferation marker, is shown in green bars (up-regulated in the proliferative state and down-regulated in the differentiated state). The MYH7 marker of differentiation is shown in pink bars (up-regulated in the differentiated state, down-regulated in the proliferative state). The NGS results were aligned with qRT-PCR platform results and are in agreement. Therefore, the relative expression of additional markers of differentiation in RD cells such as MYOD1, MYOG, and DES was presented from NGS findings.

Overall, these results conform with the expected expressions of the candidate gene markers in proliferating and differentiating states of RD cells. The same total RNA from this batch was subsequently processed for RNASeq analysis. Cross-platform concordances between qRT-PCR and NGS further strengthen the experimental premise described.

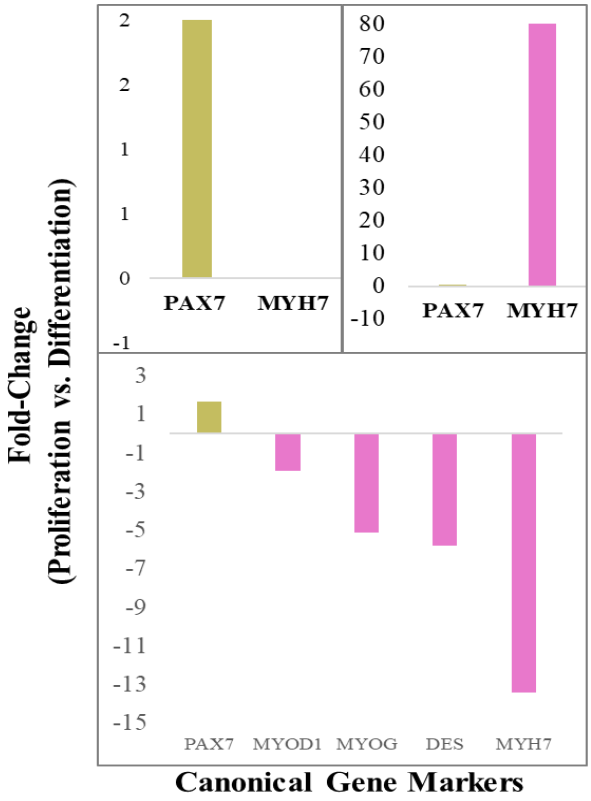


Figure 5.2 Expression fold differences in proliferating and differentiating states of RD cells
Expression fold change differences in proliferating and differentiating states of RD cells mimic the myogenic cascade of events. ***Top panel:*** qRT-PCR validation of select canonical gene markers: PAX7, for myoblast proliferating, and MYH7, for differentiating states. ***Bottom panel:*** Expression fold-change of select gene markers from Next-Generation Sequencing profiling study in proliferating vs. differentiating states. The X-axis represents canonical gene markers. Y-axis in both panels represents relative expression changes in proliferating vs. differentiating states of mature myofiber formation. Proliferating and differentiating gene markers are indicated in green and pink colored bars, respectively.
PAX7: Paired box 7; MYH7: Myosin Heavy Chain; MYOD1: Myogenic Differentiation 1; MYOG: Myogenin and DES: Desmin.

5.3.2 NGS profiling of RD cells and rat gastrocnemius muscle

The expression profiling results from the experimental model systems are as follows: n=4595 DE mRNAs were identified in the RD cells (undifferentiated vs. differentiated states), and n=1283 DE mRNAs were identified in the rodent gastrocnemius muscle (in tumor-bearing vs. healthy controls). These DE genes were subjected to pathway enrichment analysis using IPA.

5.3.3 Functional benchmark analyses of the identified molecular subtypes from patients with cancer using experimental model systems

I interrogated the transcriptome data from two independent experimental model systems to extract functional understanding and interpretation of the identified molecular subtypes. Given the tissue/cell level heterogeneity of expression profiles and gene pleiotrophy [28], the DE genes from human rectus abdominis muscle, gastrocnemius muscle, and RD cells were subjected to IPA pathway analysis. The subsequent findings and data representation are from pathway-level analysis results.

Individual analysis of the genes within each model system and their contribution to muscle homeostasis is beyond the scope of the current objectives.

Detailed pathway analysis results in the tabular format with the -log(P-value) and the overlapping or convergent and distinct pathways within the datasets are presented in **Table 5.1** and **Appendix Tables A.17-05A-20.**

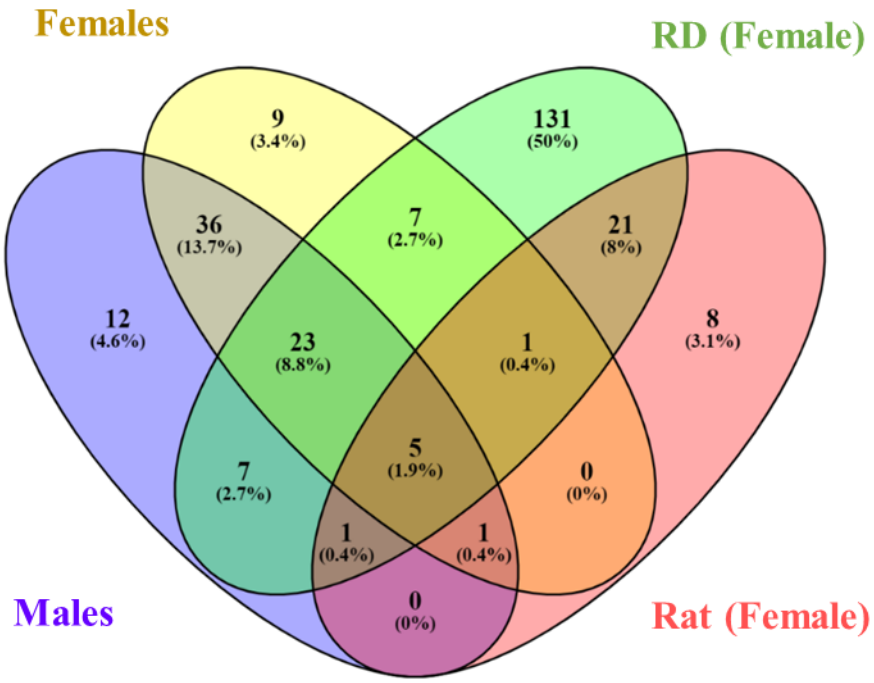
The total number of pathways determined in individual datasets were as follows: n=85 pathways in males, n=82 pathways in females were enriched in subtype 1 relative to subtype 2. In RD, n=196 pathways were enriched in proliferating vs. differentiating states of cells. In rat gastrocnemius muscle in tumor-bearing vs. healthy rats, n=37 enriched pathways were identified.

Overall, five convergent pathways were identified from patients with cancer (male and female) and the experimental model systems (**Figure 5.3**). These include Calcium signaling, nNOS signaling in Skeletal Muscle Cells, GP6 (Glycoprotein 6, a part of the Extracellular matrix) Signaling Pathway, Tumor Microenvironment Pathway, and Transcriptional Regulatory Network in Embryonic Stem Cells. Detailed descriptions of the five convergent pathways described above,

genes regulating those pathways, and the statistical significance -log(P-value) are described in **Table 5.1.**

Despite the pathway overlap, the individual genes within the convergent pathways showed limited overlap, shown in the Venn diagram in appendix **Figure A.1**. The gene comparisons revealed that they were particularly distinct, which is expected due to the tissue-specific and species-specific transcriptional profile differences. Therefore, it is imperative to perform pathway-level comparisons, not gene expression comparisons owing to the limited species/tissue-specific gene expression patterns.

Besides the convergent pathways, the individual pathway analysis within males and females also showed a certain overlap with immune cell regulatory, extracellular matrix regulation and nNOS signaling, cellular metabolism, transcriptional regulatory network, and neuroendocrinal regulation (**Appendix Table A.19, and Table A.20**).



Datasets that overlap	Pathway overlap from datasets	Number of pathways (n)
All four	Male, Female, RD and Rat	5
Any three	Male, Female and RD	23
	Male, Female and Rat	1
	Male, RD and Rat	1
	Female, RD and Rat	1
Any two	Male and Female	36
	Male and RD	7
	Female and RD	7
	RD and Rat	21
Unique pathways	Female	9
	Male	9
	RD	131
	Rat	8

Figure 5.3 Venn diagram of the pathway overlap between male, female, RD, and rat pathway analysis datasets

The figure represents the Venn diagram and table explaining the overlapping and distinct pathways from four dataset comparisons. IPA pathway analysis was performed for sex-specific DE mRNAs from the subtypes identified in patients with cancer, RD cell, and rat datasets. Overall, five pathways were found to be convergent (**Table 5.1**). The other overlapping pathways between any three or two datasets and the distinct pathway results are presented in **Table 05A.17- 05A.20**, respectively.

Table 5.1 Convergent pathways in patients with cancer and those overlapping with experimental model systems

Calcium Signaling (one of the convergent pathways)		
Males	-log(P-value)	4.5
	Genes	AKAP5,ATP2B3,CACNA1A,CACNA1B,CACNA1D,CACNA1E,CACNA1F,CACNA1G,CACNA1I,CACNA2D4,CACNB4,CACNG2,CACNG3,CACNG5,CACNG7,CACNG8,CAMK1G,CAMK4,CATSPER1,CATSPER4,CHRNA2,CHRNA3,CHRNA4,CHRNA5,CHRNA6,CHRNA7,CHRNA9,CHRNB2,CHRNB3,CHRNB4,GRIA1,GRIA2,GRIA4,GRIK1,GRIN1,GRIN2A,GRIN2B,GRIN2D,GRIN3A,HTR3A,MYH13,MYH8,MYO1A,MYO1H,PPP3R2,RYR2,SLC8A2,TNNI3,TRPC3,TRPC4,TRPC5,TRPC7,TRPM8,TRPV6
Females	-log(P-value)	4.4
	Genes	AKAP5,ATP2B3,CACNA1A,CACNA1B,CACNA1D,CACNA1F,CACNA1I,CACNG2,CACNG5,CACNG8,CAMK1D,CAMK4,CATSPER4,CHRNA2,CHRNA3,CHRNA4,CHRNA6,CHRNA7,CHRNA9,CHRNB2,CHRNB3,CHRNB4,GRIA2,GRIA4,GRIK1,GRIN1,GRIN2A,GRIN2B,GRIN3A,HTR3A,MYH10,MYH13,MYH7,MYL2,MYL6B,MYO1A,MYO1H,PPP3R2,PRKAR1B,RYR2,TRPC3,TRPC4,TRPC5,TRPC7,TRPM8,TRPV6
RD	-log(P-value)	15.0
	Genes	ACTA1,ACTC1,AKAP5,ATP2A3,ATP2B1,ATP2B4,CACNA1A,CACNA1B,CACNA1C,CACNA1D,CACNA1E,CACNA1S,CACNA2D3,CACNB1,CACNG1,CACNG4,CACNG6,CACNG8,CALM1,CAMK1D,CAMK2A,CAMK2B,CASQ1,CASQ2,CATSPER2,CHRNA1,CHRNA10,CHRNA4,CHRNA5,CHRNA9,CHRNB1,CHRNB2,CHRND,CHRNG,CREB3,CREB5,GRIK1,GRIN2B,GRIN2C,GRIN2D,GRIN3A,GRINA,HDAC11,HDAC2,HDAC5,HDAC8,HDAC9,ITPR2,MEF2A,MEF2C,MEF2D,MYH10,MYH2,MYH3,MYH6,MYH7,MYH7B,MYH8,MYH9,MYL1,MYL4,MYL5,MYL6,MYL6B,MYL9,MYO18A,MYO18B,NFATC2,PRKACA,PRKACB,PRKAG2,PRKAR1A,PRKAR1B,PRKAR2B,RCAN2,RCAN3,RYR3,SLC8A2,SLC8A3,TNNC1,TNNC2,TNNI1,TNNI2,TNNT1,TNNT2,TNNT3,TP63,TPM1,TPM2,TRDN,TRPM8,TRPV6
Rat	-log(P-value)	2.1
	Genes	ACTA1,ATP2A2,ATP2B1,CACNA1D,CACNA1S,CACNA2D3,CACNB4,CACNG1,CASQ2,GRIA4,HDAC9,MYH10,MYH6,MYH7B,RYR3,TNNC1,TNNT1,TPM3,TRPC1

nNOS Signaling (one of the convergent pathways)	
Males -log(P-value)	4.2
Genes	CACNA1A,CACNA1B,CACNA1D,CACNA1E,CACNA1F,CACNA1G,CACNA1I,CACNA2D4,CACNB4,CACNG2,CACNG3,CACNG5,CACNG7,CACNG8,CAMK4,CATSPER1,CATSPER4,RYR2
Females -log(P-value)	1.7
Genes	CACNA1A,CACNA1B,CACNA1D,CACNA1F,CACNA1I,CACNG2,CACNG5,CACNG8,CAMK4,CATSPER4,RYR2
RD -log(P-value)	4.1
Genes	CACNA1A,CACNA1B,CACNA1C,CACNA1D,CACNA1E,CACNA1S,CACNA2D3,CACNB1,CACNG1,CACNG4,CACNG6,CACNG8,CALM1,CAPN3,CATSPER2,CHRNA1,ITPR2,NOS1,RYR3,SNTA1,SNTB1
Rat -log(P-value)	2.1
Genes in Rat dataset	CACNA1D,CACNA1S,CACNA2D3,CACNB4,CACNG1,RYR3,SNTB1

GP6 Signaling Pathway (one of the convergent pathways)	
Males -log(P-value)	2.5
Genes	CAMK4,COL17A1,COL19A1,COL1A1,COL1A2,COL20A1,COL26A1,COL2A1,COL3A1,COL5A1,COL5A2,COL6A1,COL6A2,COL6A3,COL6A5,COL9A1,COL9A2,COL9A3,FGA,FGB,FGG,GP6,GRAP2,ITGA2B,LAMA1,LAMC3,NOX1,PIK3C2G,PRKCG,PRKCZ
Females -log(P-value)	4.3
Genes	CAMK4,COL10A1,COL17A1,COL1A1,COL1A2,COL20A1,COL23A1,COL24A1,COL25A1,COL2A1,COL3A1,COL5A1,COL6A1,COL6A2,COL6A3,COL6A5,COL9A1,COL9A2,COL9A3,FGA,FGB,FGG,GRAP2,LAMA1,LAMC2,LAMC3,NOX1,PIK3C2G,PIK3R2,PRKCB,PRKCZ
RD -log(P-value)	3.6
Genes	CALM1,COL15A1,COL16A1,COL1A2,COL20A1,COL22A1,COL25A1,COL26A1,COL28A1,COL2A1,COL3A1,COL4A3,COL5A1,COL5A2,COL6A1,COL6A2,COL6A3,COL7A1,FCER1G,GRAP2,ITGB3,ITK,LAMA3,LAMA5,LAMB2,LAMB3,LAMC2,PIK3C2B,PIK3CD,PIK3CG,PIK3R5,PLCG2,PRKCE,PRKCQ,PRKCZ,PRKD1,PRKD3,RASGRP2,SCHIP1,SYK,VAV1
Rat -log(P-value)	2.1
Genes in Rat dataset	APBB1IP,COL22A1,COL27A1,COL3A1,COL4A1,COL4A3,COL5A1,COL5A3,ITGB3,LAMB1,LAMC1,PIK3CB,VAV2

Tumor Microenvironment Pathway (one of the convergent pathways)	
Males -log(P-value)	2.1
Genes	ARG1,COL1A1,COL1A2,COL3A1,CTLA4,CXCL8,FASLG,FGF14,FGF20,FGF23,FGF5,FN1,FOXG1,ICAM1,IDO2,IL10,IL13,IL1B,IL6,MMP1,MMP10,MMP12,MMP13,MMP14,MMP2,MMP20,MMP21,MMP24,MMP26,MMP7,MMP8,NOS2,OSM,PIK3C2G,PROK1,SLC1A4,SLC2A2,TNF
Females -log(P-value)	2.0
Genes	ARG1,COL1A1,COL1A2,COL3A1,CTLA4,FASLG,FGF14,FGF17,FGF23,FN1,IDO2,IGF2,IL1B,IL6,MMP10,MMP12,MMP13,MMP19,MMP20,MMP21,MMP25,MMP3,MMP8,MYC,NOS2,PDGFD,PIK3C2G,PIK3R2,SLC2A2,SLC2A3,TNF,VEGFB
RD -log(P-value)	6.7
Genes	BAD,BCL2,CCL2,CCND1,CD44,COL1A2,COL3A1,CSF1,CSPG4,CXCL12,CXCL8,CXCR4,EGF,FGF1,FGF13,FGF7,FN1,FOXG1,FOXO6,HLA-A,HLA-B,HLA-C,HLA-E,ICAM1,IL1B,IL6,ITGA5,ITGB3,LEPR,MMP1,MMP11,MMP14,MMP15,MMP17,MMP19,MMP2,MMP23B,MMP24,MMP9,MRAS,MYC,NFKB2,NRAS,PDCD1LG2,PDGFB,PDGFC,PIK3C2B,PIK3CD,PIK3CG,PIK3R5,PLAU,RASD1,RASD2,RELB,RRAS,SLC1A4,SLC2A1,TGFB1,TGFB2,TNC,VEGFA,VEGFB,VEGFC
Rat -log(P-value)	1.9
Genes	BCL2,COL3A1,CXCR4,FGF1,FGF7,FGF9,FOXO4,HLAA,ITGB3,LEP,MMP14,MMP19,MYC,PIK3CB,RRAS,TGFB1

Transcriptional Regulatory Network in Embryonic Stem Cells (one of the convergent pathways)	
Males -log(P-value)	1.8
Genes	EOMES,GATA4,GJD2,HNF4A,ISL1,LHX5,NANOG,ONECUT1,OTX1,PAX6,POU5F1,RFX4,SOX2,ZIC3
Females -log(P-value)	2.1
Genes	EOMES,GATA4,HESX1,HNF4A,ISL1,NANOG,ONECUT1,OTX1,PAX6,POU5F1,RFX4,SOX2,ZIC3
RD -log(P-value)	3.3
Genes	FOXC1,GATA6,GJD2,H3A/3B,H4C1,H4C11,H4C12,H4C13,H4C2,H4C3,H4C4,H4C8,H4C9,HNF4A,MEIS1,OTX1,PAX6,RIF1,SET,SKIL,TCF7L1
Rat -log(P-value)	3.0
Genes	CDYL,H4C1,H4C14,H4C4,H4C6,H4C8,RIF1,SKIL,SMARCAD1

The male dataset is indicated in blue; the female dataset in peach color; RD in yellow; and the rat dataset in green.

5.4 Discussion

In this study, I have attempted to elucidate molecular and functional benchmark analyses of subtypes from patients with cancer. Two independent experimental model systems were used to accomplish the study endpoints and to perform molecular characterization of muscle from patients with cancer. NGS-generated whole transcriptome datasets facilitated comparisons between experimental model systems and exemplified the pathway-level regulation across independent datasets. This is the first study to demonstrate the characterization of molecular subtypes of muscle from patients with cancer from comparisons using the experimental model system to gain insights into components of muscle and regenerative processes. The pathways were compared between human skeletal muscle datasets and those identified from the experimental model systems (RD cells and tumor-implanted rodent model). Perturbations of the multiple components in the local skeletal muscle microenvironment with sustained tumor-elicited inflammatory responses were observed. From this analysis, it appears that skeletal muscle homeostasis is maintained by coordinated and integrated actions of various components and is vital for maintaining skeletal muscle mass, strength, fatigue, resistance, and metabolic properties.

Skeletal muscle regeneration requires crosstalk between satellite cells and their microenvironment, also called niche [29,30]. Skeletal muscle is a highly organized tissue consisting of bundles of myofibers (muscle fibers). Each myofiber composed of several myofibrils is known as a muscle cell with the basic structural and the contractile unit called a sarcomere. Bundles of myofibers form the fascicles, and bundles of fascicles form the muscle tissue. An extracellular matrix and cytoskeletal networks cover each layer, from muscle cells to muscle tissue [29,31]. Skeletal muscle is an important metabolic regulator and a highly vascularized and innervated tissue. The functional components of skeletal muscle include Neuromuscular junctions, Excitation-contraction coupling,

contractile unit (sarcomere), the homeostatic balance between inflammation, oxidative stress, extracellular matrix components, cytoskeleton supporting the muscle structure, circulation, and mitochondrial energy metabolism. [29] The activity and regulation of these interconnected structures are crucial for maintaining muscle mass, structure, and function. Any derangements in these coordinated processes due to external physical insults or chronic disease may lead to the loss of muscle fibers, affecting overall muscle health and function.

The key processes affecting skeletal muscle integrity and function, as depicted in **Figure 5.4**, are as follows: (1) the inflammatory perturbations leading to a sustained proinflammatory response. In an ideal state, for a muscle to regenerate, early response is mediated by the complement system followed by mast cell and neutrophil activation. This is followed by the activation of M1 macrophage (proinflammatory), and the last phase is characterized by the activation of M2 macrophage (anti-inflammatory). (2) Satellite cells (muscle stem cells) reside in a specialized microenvironment referred to as a "niche"[29,32]. The niche is important for maintaining satellite cell quiescence, provides a signaling mechanism for activation, proliferation, and differentiation of satellite cells in response to tissue injury, and maintains the structural integrity of the muscle. Satellite cells are located between the basal lamina and sarcolemma of the myofiber (**Figure 5.4**). The ECM surrounding the muscle fiber comprises collagens, laminins, fibronectin, and proteoglycans [33]. ECM consists of two layers, an outer layer, known as the reticular lamina, and an inner basal lamina. Satellite cells are connected to the basal lamina via integrins and laminins, which are connected to collagen IV of the basal lamina. Satellite cell also connects to the dystrophin complex on the sarcolemma via laminin. (3) Basal lamina maintains the quiescence state of satellite cells. In order for satellite cells to migrate to the injured site, it requires the degradation of ECM components which is mediated by Matrix MetalloProteases (MMPs). The

fragments and growth factors released during the ECM degradation are important for satellite cell activation, proliferation, and differentiation. A tight regulation between MMPs and Tissue inhibitors of MMPS (TIMPs) is required for ECM regulation [34]. Imbalance in this pathway results in ECM accumulation leading to fibrosis [35,36]. (4) The two important signaling mediators of the ECM are integrins and Dystrophin Glycoprotein Complex (DGC), consisting of dystroglycan, sarcoglycans, dystrobrevins, and syntrophins. DGC is the structural link between the cytoskeleton and ECM, and interactions with nNOS mediate the signaling via DGC [31,37]. (5) Nitric oxide synthase (NOS) is attached to syntrophin and binds to dystrophin. NOS catalyzes the formation of Nitric Oxide and L-citrulline from L-arginine. Skeletal muscle fibers are a vital source of nNOS [38,39]. Three forms of nNOS are neuronal (nNOS, activated by interaction with calcium and calmodulin), endothelial (eNOS, activated by interaction with calcium and calmodulin), and inducible or macrophage or calcium-independent (iNOS). nNOS is differentially distributed near mitochondria, presynaptic and postsynaptic motor nerve terminals, sarcoplasmic reticulum, and sarcolemma. NO is a multifunctional molecule that affects muscle contraction, the mitochondrial respiratory chain, glucose metabolism, and the neuromuscular junction. It serves as an activity sensor and responds to the effects of prolonged disuse as well as exercise-mediated changes. (6) Mitochondria are essential for ATP production and are one of the primary sources of Reactive Oxygen Species/Reactive Nitrogen Species. Dysregulation in mitochondrial energetics results in elevated levels of ROS that are detrimental to skeletal muscle homeostasis. (7) Dysregulation in nNOS signaling results in the delocalization of NOS from sarcolemma, and cytoplasmic accumulation leads to Ryanodine receptor (RYR, the calcium channels that control intracellular calcium levels by releasing calcium from the sarcoplasmic reticulum)nitrosylation and thereby leads to Ca^{2+} leakage from the sarcoplasmic reticulum. Ca^{2+} leak can further activate nNOS in a

feed-forward mechanism [40]. This affects the overall dynamics of skeletal muscle contractility and causes muscle damage and fatigue. (8) Dysregulated calcium signaling affects NMJ signaling. Besides acetylcholine, glutamate is also shown to stimulate nNOS acting through N-methyl-D-aspartate receptors and is colocalized with nNOS at the sarcolemma [37]. (9) Transcriptional regulation plays an essential role in the maintenance of skeletal muscle homeostasis. This vicious cycle or series of imbalanced component mechanisms with the sustained inflammatory response in the local milieu may contribute to the pathophysiology of cachexia and therefore needs to be carefully investigated.

The components contributing to skeletal muscle homeostasis and their disturbances resulting in the imbalance were found to be altered in subtype 1 relative to subtype 2, and that subtype 1 is likely a severe group affected by cachexia. These findings that subtype 1 is a skeletal muscle wasting group align with the results from clinical benchmark analyses (chapter 3), wherein subtype 1 was associated with high-grade WL-BMI and with low muscle mass (evident from z-SMI distribution). Overall, subtype 1 is inferred as a group affected by cachexia from clinical benchmark analyses (Chapter 3) and the functional benchmark analyses performed in this study.

Despite the large number of pathways identified in each dataset, the limited overlap across all datasets. This was expected due to the heterogeneity in the model systems. Heterogeneity in the transcriptome and metabolic profiles of the model systems exists due to species- and sex-specific differences within study datasets, differential skeletal muscle anatomical localization (rectus abdominis skeletal muscle obtained from patients with cancer, and gastrocnemius obtained from rats bearing tumors and those from the healthy control rats), different comparator groups and reference groups. Human muscle subtypes were identified from patients with cancer, whereas the

rat model consisted of tumor-bearing and healthy reference controls. The rat model system is a highly controlled system wherein a tumor is implanted based on the specified experimental timeline [26], and therefore transcriptional changes are captured accordingly. On the other hand, the human clinical cachexia condition is the chronic wasting process, and the specific stages of muscle wasting (i.e., when the muscle wasting exactly occurred) or the damage caused to muscle due to the presence of a tumor cannot be captured. Abdelmoez and colleagues performed comparative profiling of skeletal muscle models (Rat L6, mouse C2C12, and primary human skeletal muscle cells) and identified transcriptome and metabolic differences based on the choice of the independent model systems [41]. While the current study focused on the convergent pathways, the importance of the pathways overlapping with any two comparators is not to be undermined (Subtypes in individual sex vs. rat model or subtypes vs. RD model, or RD vs. rat model). Further investigations are warranted on the relative importance of these pathways in the above comparisons. In the absence of an ideal model of cancer-associated muscle wasting, I utilized the model systems developed in-house mimicking muscle cell regenerative process (RD) and a model mimicking cancer-induced muscle atrophy (Rat) and was able to successfully explore the potential molecular characteristics of skeletal muscle subtypes from patients with cancer.

The overarching theme emerging from this study underpins the implications and dysregulation of several coordinated events that regulate skeletal muscle structure, function, and regenerative process. Numerous investigations have been conducted on experimental systems mainly focussed on the catabolic pathways with a particular emphasis on protein degradation pathways (ubiquitin-proteasome pathway, autophagy, calpain proteases, and caspases). Protein homeostasis is essential for skeletal muscle mass but is not the only contributory factor. Emphasis was placed on enhanced protein degradation and decreased protein synthesis in the experimental systems. Human skeletal

muscle studies, on the contrary, presented conflicting results [10,11]. Observations from model systems were not recapitulated in the human muscle studies [42,43]. Given these limitations, the current study design and conceptualization justify using the two experimental model systems to understand and conduct benchmark analysis of the human skeletal muscle subtypes. While common pathways across all datasets highlight shared processes, this is not to undermine the importance of other pathways in individual datasets or pair-wise comparisons between any two datasets. Regardless of the number of pathways in individual datasets, the overarching theme appears to be the key findings summarized in **Figure 5.4**.

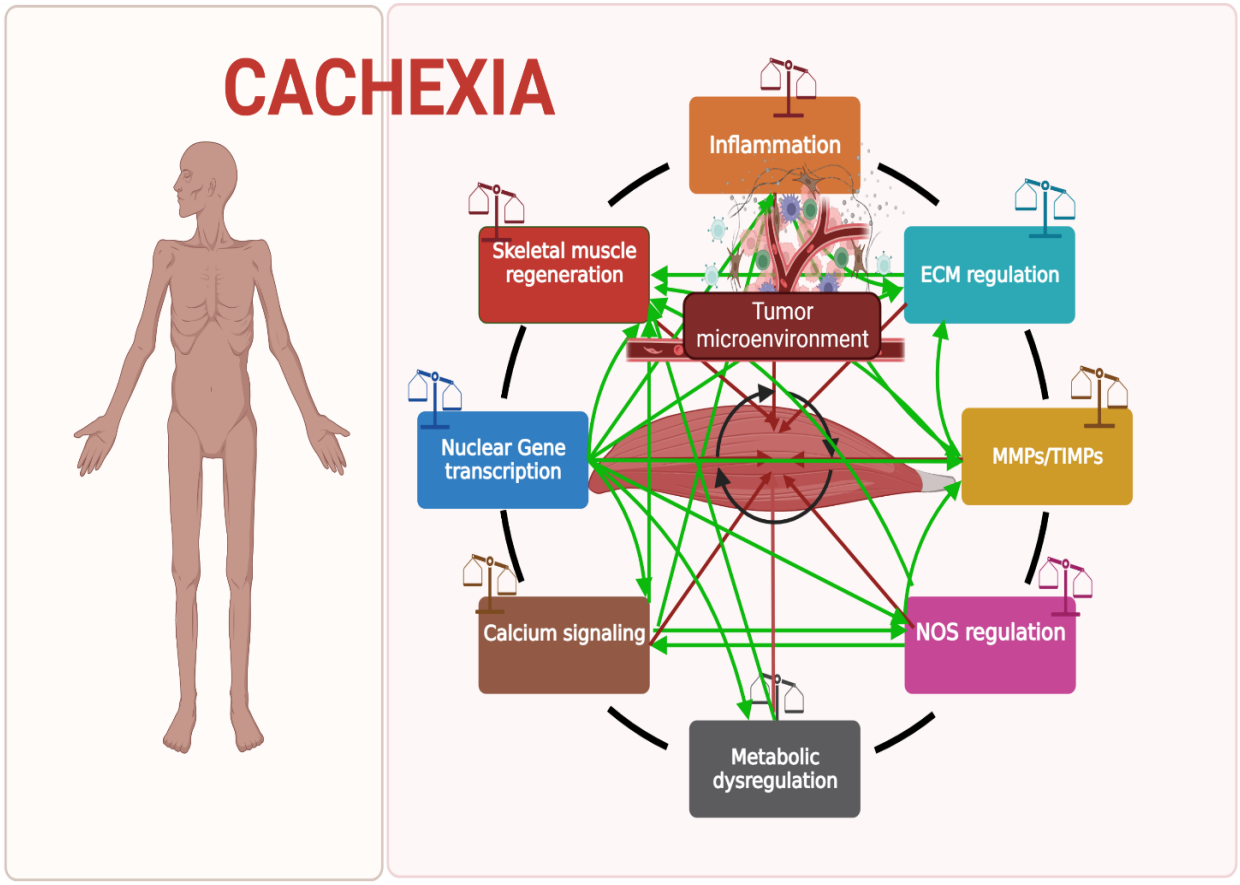


Figure 5.4 Holistic view of pathway implications from functional benchmark analysis
Overarching mechanistic insights from molecular characterization of subtypes identified in patients with cancer. The figure depicts the holistic view, including the coordinated mechanisms involving individual components and their crosstalk occurring within the milieu of skeletal muscle along with the tumor-elicited sustained inflammatory responses. Created with Biorender.com

This is the first study using the human skeletal muscle that has demonstrated the involvement of several fundamental factors integral to skeletal muscle integrity, the most important being the crosstalk between muscle and its microenvironment. A series of events affecting the muscle regenerative process, such as sustained chronic inflammatory response, perturbations in Extracellular Matrix regulation, Calcium signaling, nNOS signaling, transcriptional dysregulation, and metabolic disturbances, were some of the key mechanisms observed. The process of regeneration begins with the necrosis of damaged fibers and the infiltration of immune cells. The initial processes involve the activation of the coagulation cascade followed by immune cell response to the injury [44]. The complement system serves as the initial trigger in response to muscle injury. The activation by the complement system is followed by the activation and infiltration of mast cells and neutrophils (early phase). This is followed by the activation of T cells, and M1 macrophages (pro-inflammatory) stimulate satellite cell proliferation and prepress differentiation; in the middle phase, these processes secrete TNFα, IL1β, IFNγ, and IL6. The last phase is characterized by M2 macrophages along with regulatory T cells. M2 macrophages secrete IL4, IL10, and IL13 and are characterized by differentiation and maturation of myofibers. Acute inflammatory response plays an important role in initiating the tissue remodeling process. However, sustained chronic inflammatory responses are deleterious for muscle health and function. A sustained inflammatory process disrupts ECM remodeling (degradation and reformation). The interplay and balance between hydrolytic activity and its inhibition are important for ECM homeostasis. Coordination between MMPs/TIMPs is required for skeletal muscle regeneration, as disruption of this balance has deleterious effects on skeletal muscle structure and function. MMP2 and MMP14 deficiency were found to contribute to defective maturation of skeletal muscle and promote cell death [45,46]. MMP14 regulates the migration of myoblasts and

therefore plays an important role in skeletal muscle regeneration. MMP2 and MMP14 were found to be DE in male subtypes but not female ones. MMP2 and MMP14 expression were downregulated in male subtype 1. 18 collagen genes were found to be perturbed in subtypes in both the sexes and 12 and 9 MMP genes in males and females, respectively. The beneficial or disruptive effects of NOS are based on the spatial localization of NOS synthases. nNOS is important for contractile function and maintaining skeletal muscle integrity by regulating muscle size, strength, and fatigue resistance. NOS expression increases after endurance exercise, whereas inactivity or prolonged muscle disuse leads to decreased expression. In an experimental model of tail-suspension-induced muscle atrophy in mouse models, nNOS from the DGC complex induced muscle atrophy via upregulation of FoxO [47]. nNOS dissociated from α-1syntrophin within the sarcolemma and dislocated to the cytoplasm. The delocalization of NOS led to the generation of NO that regulates Foxo transcription factors, thereby promoting proteolysis by the ubiquitin-proteasome system. Alterations in the nNOS pathway impair muscle fiber growth by causing a decrease in the number of myonuclei per fiber and defects in the fusion of myogenic precursor cells during regeneration, mitochondrial dysfunction (alterations in shape, morphology, size, and energy metabolism), Unfolded Protein Response, and autophagy and were observed *in vivo* mouse model and isolated satellite cells in vitro models [48].

Judge et al. performed morphological and microarray analysis on the rectus abdominis muscle biopsies from PDAC patients and weight-stable control subjects and found increased collagen content in the cachexic PDAC patients representative of a fibrotic phenotype (cachexia defined using 5% WL criteria) [49]. My study findings using muscle from the same anatomical location align with their findings, suggesting the importance of studying components described as a schematic model. Reports from another human skeletal muscle study affirm these findings. Stephens et al.

evaluated biomarkers for cancer cachexia from rectus abdominis muscle of gastrointestinal cancer patients and concluded that structural elements of skeletal muscle were potential cancer cachexia biomarkers [50]. Amongst all the atrophy and structural genes quantified, the study suggested β-dystroglycan as a potential biomarker of cancer cachexia [50]. Overall, the inferences obtained from the literature-reported findings and those observed in the muscle of patients with cancer in my study suggest that these mechanisms need to be studied in detail, and the therapeutic developments or drug repurposing should target the components described in **Figure 5.4**.

It is important to note that the findings reported in this study are focused on the mRNA expression profiles generated and comparative pathway analyses at the mRNA level. Future implementations should incorporate multilayered RNA crosstalk from post-transcriptional ceRNA (lncRNA-miRNA-mRNA) mechanisms in the experimental model systems. These would facilitate the comparisons of hub lncRNAs and their regulatory networks identified from human skeletal muscle subtypes (described in Chapter 4). There is evidence of lncRNAs involved in regulating ECM in skeletal muscle [51]. Therefore, if the ceRNA mediated-gene regulation is operated by the same hub lncRNAs, then further knockdown or overexpression experiments could be performed. It would help confirm the mRNA expression changes and thereby governing pathways. These need to be determined.

5.5 Conclusions

This is the first investigation towards generating a transcriptional landscape of skeletal muscle from patients with cancer and comparisons to cellular model systems representative of muscle regenerative process and a skeletal muscle model of tumor-induced muscle atrophy. Invoking model systems for molecular and functional benchmarking of the muscle subtypes is also the first

in the literature to align the human skeletal muscle transcriptome for putative functional insights. Implementation of the comparative analyses in this study illuminated the importance of the component mechanisms concerning skeletal muscle homeostasis. I provide a schematic model from the pathway analyses comparisons across different datasets wherein disturbances occurring within the skeletal muscle and its microenvironment and the tumor-elicited effects aggravating those disturbances are captured. The resulting disruptions from all data sets' component mechanisms from converging pathways are depicted and summarized from the in silico IPA pathway analysis. These analyses provide insights into the skeletal muscle integrity, function, and overall homeostasis in patients with cancer. Distinct component mechanisms within the skeletal muscle milieu epitomize the series of events (cycle) in a feed-forward manner, including sustained inflammatory response, disturbances in ECM regulation, perturbations in calcium signaling, and transcriptional networks. In this study, molecular and functional benchmarking methods for identifying intrinsic characteristics of skeletal muscle from patients with cancer corroborate with the findings obtained from the clinical benchmarking methods, i.e., the clinical relevance (WL-BMI graded severity and z-SMI distribution) of the subtypes. Future studies should target the discrete component mechanisms (**Figure 5.4**) acting in a coordinated fashion in concert rather than in isolation in future studies.

5.6 References

1. Martin L, Senesse P, Gioulbasanis I, Antoun S, Bozzetti F, Deans C, et al. Diagnostic criteria for the classification of cancer-associated weight loss. *J Clin Oncol.* 2015;33(1):90-99.
2. Kazemi-Bajestani SM, Mazurak VC, Baracos V. Computed tomography-defined muscle and fat wasting are associated with cancer clinical outcomes. *Semin Cell Dev Biol.* 2016;54:2-10.
3. Fearon K, Strasser F, Anker SD, Bosaeus I, Bruera E, Fainsinger RL, et al. Definition and classification of cancer cachexia: an international consensus. *Lancet Oncol.* 2011;12(5):489-495.
4. Sartori R, Romanello V, Sandri M. Mechanisms of muscle atrophy and hypertrophy: implications in health and disease. *Nat Commun.* 2021;12(1):330.
5. Bossola M, Muscaritoli M, Costelli P, Grieco G, Bonelli G, Pacelli F, et al. Increased muscle proteasome activity correlates with disease severity in gastric cancer patients. *Ann Surg.* 2003;237(3):384-389.
6. Schersten T, Lundholm K. Lysosomal enzyme activity in muscle tissue from patients with malignant tumor. *Cancer.* 1972;30(5):1246-1251.
7. Williams A, Sun X, Fischer JE, Hasselgren PO. The expression of genes in the ubiquitin-proteasome proteolytic pathway is increased in skeletal muscle from patients with cancer. *Surgery.* 1999;126(4):744-749; discussion 749-750.
8. Khal J, Wyke SM, Russell ST, Hine AV, Tisdale MJ. Expression of the ubiquitin-proteasome pathway and muscle loss in experimental cancer cachexia. *Br J Cancer.* 2005;93(7):774-780.

9. Khal J, Hine AV, Fearon KC, Dejong CH, Tisdale MJ. Increased expression of proteasome subunits in skeletal muscle of cancer patients with weight loss. *Int J Biochem Cell Biol.* 2005;37(10):2196-2206.

10. Stephens NA, Gallagher IJ, Rooyackers O, Skipworth RJ, Tan BH, Marstrand T, et al. Using transcriptomics to identify and validate novel biomarkers of human skeletal muscle cancer cachexia. *Genome Med.* 2010;2(1):1.

11. Gallagher IJ, Stephens NA, MacDonald AJ, Skipworth RJ, Husi H, Greig CA, et al. Suppression of skeletal muscle turnover in cancer cachexia: evidence from the transcriptome in sequential human muscle biopsies. *Clin Cancer Res.* 2012;18(10):2817-2827.

12. D'Orlando C, Marzetti E, Francois S, Lorenzi M, Conti V, di Stasio E, et al. Gastric cancer does not affect the expression of atrophy-related genes in human skeletal muscle. *Muscle Nerve.* 2014;49(4):528-533.

13. He WA, Berardi E, Cardillo VM, Acharyya S, Aulino P, Thomas-Ahner J, et al. NF-kappaB-mediated Pax7 dysregulation in the muscle microenvironment promotes cancer cachexia. *J Clin Invest.* 2013;123(11):4821-4835.

14. Arneson PC, Doles JD. Impaired Muscle Regeneration in Cancer-Associated Cachexia. *Trends Cancer.* 2019;5(10):579-582.

15. Bossola M, Marzetti E, Rosa F, Pacelli F. Skeletal muscle regeneration in cancer cachexia. *Clin Exp Pharmacol Physiol.* 2016;43(5):522-527.

16. Penna F, Costamagna D, Fanzani A, Bonelli G, Baccino FM, Costelli P. Muscle wasting and impaired myogenesis in tumor bearing mice are prevented by ERK inhibition. *PLoS One.* 2010;5(10):e13604.

17. Pessina P, Conti V, Pacelli F, Rosa F, Doglietto GB, Brunelli S, et al. Skeletal muscle of gastric cancer patients expresses genes involved in muscle regeneration. *Oncol Rep.* 2010;24(3):741-745.

18. Johns N, Stephens NA, Fearon KC. Muscle wasting in cancer. *Int J Biochem Cell Biol.* 2013;45(10):2215-2229.

19. Talbert EE, Guttridge DC. Impaired regeneration: A role for the muscle microenvironment in cancer cachexia. *Semin Cell Dev Biol.* 2016;54:82-91.

20. Inaba S, Hinohara A, Tachibana M, Tsujikawa K, Fukada SI. Muscle regeneration is disrupted by cancer cachexia without loss of muscle stem cell potential. *PLoS One.* 2018;13(10):e0205467.

21. Novak JS, Mazala DAG, Nearing M, Hindupur R, Uapinyoying P, Habib NF, et al. Human muscle stem cells are refractory to aging. *Aging Cell.* 2021;20(7):e13411.

22. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods.* 2001;25(4):402-408.

23. Kechin A, Boyarskikh U, Kel A, Filipenko M. cutPrimers: A New Tool for Accurate Cutting of Primers from Reads of Targeted Next Generation Sequencing. *J Comput Biol.* 2017;24(11):1138-1143.

24. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics.* 2013;29(1):15-21.

25. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014;15(12):550.

26. Almasud AA, Giles KH, Miklavcic JJ, Martins KJB, Baracos VE, Putman CT, et al. Fish oil mitigates myosteatosis and improves chemotherapy efficacy in a preclinical model of colon cancer. *PLoS One.* 2017;12(8):e0183576.

27. Kramer A, Green J, Pollard J, Jr., Tugendreich S. Causal analysis approaches in Ingenuity Pathway Analysis. *Bioinformatics.* 2014;30(4):523-530.

28. Narasimhan A, Zhong X, Au EP, Ceppa EP, Nakeeb A, House MG, et al. Profiling of Adipose and Skeletal Muscle in Human Pancreatic Cancer Cachexia Reveals Distinct Gene Profiles with Convergent Pathways. *Cancers (Basel).* 2021;13(8).

29. Mukund K, Subramaniam S. Skeletal muscle: A review of molecular structure and function, in health and disease. *Wiley Interdiscip Rev Syst Biol Med.* 2020;12(1):e1462.

30. Loreti M, Sacco A. The jam session between muscle stem cells and the extracellular matrix in the tissue microenvironment. *NPJ Regen Med.* 2022;7(1):16.

31. Thomas K, Engler AJ, Meyer GA. Extracellular matrix regulation in the muscle satellite cell niche. *Connect Tissue Res.* 2015;56(1):1-8.

32. Dumont NA, Bentzinger CF, Sincennes MC, Rudnicki MA. Satellite Cells and Skeletal Muscle Regeneration. *Compr Physiol.* 2015;5(3):1027-1059.

33. Zhang W, Liu Y, Zhang H. Extracellular matrix: an important regulator of cell functions and skeletal muscle development. *Cell Biosci.* 2021;11(1):65.

34. Csapo R, Gumpenberger M, Wessner B. Skeletal Muscle Extracellular Matrix - What Do We Know About Its Composition, Regulation, and Physiological Roles? A Narrative Review. *Front Physiol.* 2020;11:253.

35. Devine RD, Bicer S, Reiser PJ, Velten M, Wold LE. Metalloproteinase expression is altered in cardiac and skeletal muscle in cancer cachexia. *Am J Physiol Heart Circ Physiol.* 2015;309(4):H685-691.

36. Alameddine HS, Morgan JE. Matrix Metalloproteinases and Tissue Inhibitor of Metalloproteinases in Inflammation and Fibrosis of Skeletal Muscles. *J Neuromuscul Dis.* 2016;3(4):455-473.

37. Grozdanovic Z, Baumgarten HG. Nitric oxide synthase in skeletal muscle fibers: a signaling component of the dystrophin-glycoprotein complex. *Histol Histopathol.* 1999;14(1):243-256.

38. Stamler JS, Meissner G. Physiology of nitric oxide in skeletal muscle. *Physiol Rev.* 2001;81(1):209-237.

39. Kobayashi J, Uchida H, Kofuji A, Ito J, Shimizu M, Kim H, et al. Molecular regulation of skeletal muscle mass and the contribution of nitric oxide: A review. *FASEB Bioadv.* 2019;1(6):364-374.

40. Percival JM. nNOS regulation of skeletal muscle fatigue and exercise performance. *Biophys Rev.* 2011;3(4):209-217.

41. Abdelmoez AM, Sardon Puig L, Smith JAB, Gabriel BM, Savikj M, Dollet L, et al. Comparative profiling of skeletal muscle models reveals heterogeneity of transcriptome and metabolism. *Am J Physiol Cell Physiol.* 2020;318(3):C615-C626.

42. Mueller TC, Bachmann J, Prokopchuk O, Friess H, Martignoni ME. Molecular pathways leading to loss of skeletal muscle mass in cancer cachexia--can findings from animal models be translated to humans? *BMC Cancer.* 2016;16:75.

43. Baracos VE. Bridging the gap: are animal models consistent with clinical cancer cachexia? *Nat Rev Clin Oncol.* 2018;15(4):197-198.

44. Yang W, Hu P. Skeletal muscle regeneration is modulated by inflammation. *J Orthop Translat.* 2018;13:25-32.

45. Oh J, Takahashi R, Adachi E, Kondo S, Kuratomi S, Noma A, et al. Mutations in two matrix metalloproteinase genes, MMP-2 and MT1-MMP, are synthetic lethal in mice. *Oncogene.* 2004;23(29):5041-5048.

46. Snyman C, Niesler CU. MMP-14 in skeletal muscle repair. *J Muscle Res Cell Motil.* 2015;36(3):215-225.

47. Suzuki N, Motohashi N, Uezumi A, Fukada S, Yoshimura T, Itoyama Y, et al. NO production results in suspension-induced muscle atrophy through dislocation of neuronal NOS. *J Clin Invest.* 2007;117(9):2468-2476.

48. De Palma C, Morisi F, Pambianco S, Assi E, Touvier T, Russo S, et al. Deficient nitric oxide signalling impairs skeletal muscle growth and performance: involvement of mitochondrial dysregulation. *Skelet Muscle.* 2014;4(1):22.

49. Judge SM, Nosacka RL, Delitto D, Gerber MH, Cameron ME, Trevino JG, et al. Skeletal Muscle Fibrosis in Pancreatic Cancer Patients with Respect to Survival. *JNCI Cancer Spectr.* 2018;2(3):pky043.

50. Stephens NA, Skipworth RJ, Gallagher IJ, Greig CA, Guttridge DC, Ross JA, et al. Evaluating potential biomarkers of cachexia and survival in skeletal muscle of upper gastrointestinal cancer patients. *J Cachexia Sarcopenia Muscle.* 2015;6(1):53-61.

51. Sun Y, Sun X, Liu S, Liu L, Chen J. The overlap between regeneration and fibrosis in injured skeletal muscle is regulated by phosphatidylinositol 3-kinase/Akt signaling pathway - A bioinformatic analysis based on lncRNA microarray. *Gene.* 2018;672:79-87.

Chapter 6 Discussion

This thesis is the first and the most extensive investigation in cachexia literature. In this thesis, I have performed sex-specific molecular characterization of human skeletal muscle from patients with cancer. I performed data integration using an unsupervised clustering algorithm from the entire transcribed genome of the human skeletal muscle to identify molecular subtypes. The subtype identification was independent of clinical cachexia parameters, also referred to as clinical labels throughout the thesis. Further clinical relevance and functional insights were gained using independent analytical approaches. The key findings of this study are described below:

6.1 Sample size considerations and acquisition of skeletal muscle biopsy

I utilized N=84 (N=48 males and N=36 females) human *rectus abdominis* skeletal muscle biopsy specimens obtained from patients undergoing surgical tumor resection in a regional cancer center in Alberta, Canada. Biopsy specimen from each patient was subjected to high-throughput Next Generation rRNA and small RNA sequencing (mRNAs, lncRNAs, and small non-coding RNAs) to generate transcriptional expression profiles. To date, molecular profiling studies using human skeletal muscle biopsies have used qRT-PCR or northern blot analysis to measure expression levels of candidate gene markers or various microarray technology platforms to profile mRNAs. High-throughput NGS sequencing has several advantages over qRT-PCR and microarray-based techniques described in chapter 1.

None of the human muscle studies from patients with cancer have characterized sex-specific transcriptional changes. A recent study used targeted sequencing (representation of over 20k RefSeq genes) using an Ampliseq transcriptome gene expression kit. Sequencing was performed on rectus abdominis muscle and subcutaneous biopsies of Pancreatic Ductal Adenocarcinoma (PDAC) vs. healthy controls to identify a distinct and overlapping set of muscle and fat-related

genes and pathways [1]. Although the targeted sequencing was performed on protein-coding genes and is therefore not a complete representation of the entire transcriptome, the sex-related differences were not elucidated, possibly due to the limited sample size. My study overcomes this limitation. I tried to minimize the effects of the confounding variables in the statistical models considered and the ensuing analysis plan.

6.2 Molecular profiling studies and challenges in the patient classification

Despite the higher prevalence of cachexia in patients with cancer, clinically meaningful definitions and diagnostic criteria, molecular staging, and understanding of pathophysiologic mechanisms are still rudimentary. Data emerging from decades of work using experimental model systems are beyond the realms of understanding cachexia as a holistic problem emulating the complex human cachexia condition. Nevertheless, a vast interpretation of representative pathways contributing to the pathophysiology of cancer-associated muscle wasting was acquired from varied animal models of cachexia due to the supporting experimental evidence. Some reasons for the lack of translatability include the rapid growth of tumors in these models [2,3], comorbidities in clinical settings that are not reflected in the model systems, and the lack of appropriate models that represent the human cachexia condition, including the state of metastasis and cancer-specific treatment effects. One of the caveats of the animal models for cachexia is that the tumor proliferates and reaches 10% of the body mass within a limited period of a few days or weeks.

In contrast, in humans, the tumor burden is < 1% of the body mass [2,4]. Human skeletal muscle biopsy is an excellent tool to infer the muscle-wasting phenomenon in patients with cancer [5]. However, the accrual of muscle biopsy specimens is an invasive procedure, usually obtained during surgical resection of the tumor, and requires extensive collaboration between cachexia researchers and surgical oncologists. Therefore, studies employing muscle biopsy are limited.

Until now, studies that have used human skeletal muscle biopsy to gain insights into molecular mechanisms of muscle wasting have applied case vs. control strategies. The case herein is defined as a patient with cancer experiencing a phenotype of % WL or low skeletal muscle mass defined using low Skeletal Muscle Index (SMI) or radiodensity as a metric of muscle quality. (parameters obtained from image analysis). The weight changes often introduce recall bias, incomplete records from the medical charts, and are not insightful of underlying dynamic changes of muscle mass; [6,7] hence, using weight changes as a sole diagnostic criterion to stratify patients is challenging [8].

Lack of assessment of weight changes was also reported in a recent study from our group, wherein of the n=30 patients, n= 3 did not have records of weight changes, and for n=11, data for weight changes were taken several months before surgery [9]. It is evident from the literature that when weight loss is recognized in patients with cancer, cachexia is often in a state of irreversibility, and the outcome measures in patients deteriorate [10]. This has also resulted in the failures of clinical trials because the recruitment of patients was at a refractory stage wherein weight loss is severe, and the catabolic events are exalted. One of the other muscle measures obtained from CT image analysis, which I have not considered in my study, is Skeletal Muscle Radiodensity (SMR, measured in HU). Although SMR is not a canonical parameter used in classifying cases as cachexic, it is associated with the amount of lipid content in muscle or fatty infiltration in muscle [11], known as myosteatosis. It is associated with poor clinical outcomes [12-14]. However, a recent study found that intravenously administered contrast material during scans and different beam energies (in kilovolts) serve as confounding variables. Muscle density increased or decreased based on high contrast or low beam energy. These variables are often not reported in clinical studies or profiling studies [15]. Hence, these variables should be used cautiously, especially when designing molecular studies, due to the labor-intensive, resource, and time-consuming

experimental design and processing. No consensus exists on the classification strategies, and variable criteria employed in molecular or profiling studies have generated conflicting results.

For the initial analyses, I performed phenotyping of patients based on cachexia parameters of % WL, SMI, and FMI (proxy to BMI). I generated bivariate sex- and cohort-specific median cutoffs to reduce the bias of using literature-derived cutoffs of SMI or arbitrary investigator-defined WL cutoffs. I found that patients in my cohort exhibited interindividual variations and distributed to a wide range of the phenotypic spectrum (**Figures 2.1 and 2.2**). For instance, I found that a weight-losing patient was on the high SMI spectrum and vice versa. The inference was that the clinical and body composition parameters did not yield consensus; hence, I needed alternative unbiased methods that were independent of these parameters to classify patients with cancer into subtypes to understand the pathobiology of the skeletal muscle.

To summarize the findings from the clinical phenotyping analyses, there is a lack of a clear understanding of the definition and diagnostic criteria that can be applied to classify patients as cachexic both in a clinical setting and when pursuing molecular profiling studies. The use of heterogeneous patient classification strategies is one of the key impediments to the replicability of molecular findings across the studies.

6.3 Unsupervised machine learning approaches to identify sex-specific molecular subtypes of human skeletal muscle from patients with cancer

Unsupervised clustering methods are well-established and were applied in consortia studies for various cancer types. I used Non-negative Matrix Factorization (NMF), a dimensionality reduction method that models data by additive combinations of non-negative basis vectors, i.e., metagenes, to cluster individual samples relying on individual RNA expression profiles. Brunet et al. were the

first researchers to utilize NMF in the gene expression domain using three cancer datasets of leukemia, medulloblastoma, and central nervous system tumors and identified biologically significant phenotypes [16]. Since then, the applicability of NMF has increased tremendously. NMF utilizes non-negative gene by sample matrix (gene expression matrix) and reduces the dimensionality to a few metagenes reflective of the original matrix. In this study, I applied the NMF algorithm to cluster samples. However, the applicability of NMF is vast and can be used to cluster genes and gene set discovery, among others. In the NMF analysis employing single RNA species, the resulting association statistics with clinical and body composition variables of the NMF clusters showed striking patterns reflecting the sample heterogeneity across clusters and their differential associations with clinical labels. The biological mechanisms are attributed to the complex interplay between various RNA species and are not dependent on a single RNA species. Therefore, I investigated the interactions amongst different RNAs to sample clustering to mirror the complex biological patterns inherent to the subtype, a proxy to a phenotype. I used the intNMF algorithm to identify molecular subtypes of skeletal muscle from patients with cancer. Unlike other models that require specific distributional assumptions to be met, the intNMF algorithm does not assume any distributional form of the data. Two cluster model was selected, considering the cluster quality indices (cophenetic coefficient, cluster prediction index, and silhouette width) and sample size utilized to confidently ascertain the subtypes and facilitate downstream analysis and data interpretations. This is the first study wherein whole transcriptome profiles were used to generate subtypes of muscle, unlike any other previously conceived study based on supervised classification using case vs. control strategies dependent on phenotypic characterizations of the cohorts.

Several attempts were made to apply machine learning models to detect changes in body composition using imaging techniques, including muscle loss [17], weight loss [18], and application of

auto segmentation of CT image data (instead of manual segmentation) and analysis of muscle and fat components of body composition [14,19]. While these are preliminary findings, the central question that remains to be determined is the use of varying cutoffs, the applicability of appropriate machine learning models to assess changes in body composition, and the underlying biological insights. My thesis addresses these questions and advances the field using integrative approaches to gain biological insights using human skeletal muscle.

6.4 Sex-specific expression analyses and pathway regulation

Sex-related differences are prominent in skeletal muscle. Males have higher muscle mass, a higher composition of fast-twitch Type II fibers (i.e., glycolytic phenotype), and are less fatigue resistant [20]. Sexual dimorphism is also extensively studied in the gene expression profiles from human skeletal muscle biopsies in healthy aging [21,22] and exercise settings [23-26]. My study identified sex-specific Differentially Expressed (DE) profiles of the protein-coding mRNAs and the long and small ncRNAs. Notably, the DE profile overlap between males and females for mRNAs and lncRNAs was about 50%. In contrast, I observed a higher diversity of expression differences in the small ncRNAs, with miRNAs exhibiting an overlap of 34.7% genes between males and females. Other small RNAs demonstrated < 10% gene overlap between the sexes. Therefore, studying whole transcriptome profiles rather than mRNAs in isolation is imperative. Protein-coding genes are under tight regulation by lncRNAs or small non-coding RNAs. While mechanisms of miRNA regulation of mRNAs are well studied, the regulatory potential of snoRNAs, piRNAs, and tRNAs is currently being explored for their non-canonical functions. Recognizing the regulatory roles of all small RNAs and lncRNAs, I included these RNAs along with protein-coding RNAs in the intNMF clustering models to identify the skeletal muscle subtypes in patients with cancer. This is the first report to show sexually dimorphic expression

patterns of these RNA species in the human skeletal muscle of patients with cancer, with the highest sample size conducted to date in the literature.

Transcriptome profiling has revealed that in comparisons of muscle biopsy specimens from different anatomical locations of skeletal muscle, those obtained from the same site showed homogeneity [27], whereas intrinsic transcriptional diversity is prevalent among other skeletal muscle tissues [28]. About 55.2% of transcripts were differentially expressed among 11 mouse skeletal muscles, and about 13% were differentially expressed between at least two skeletal muscles [28], highlighting the importance of using a single muscle type and limiting the number of cancer types in the study to limit heterogeneity to allow interpretability and replication of findings in independent studies.

In this study, rectus abdominis skeletal muscle was obtained from each patient to perform sequencing of small ncRNAs, lncRNAs, and mRNAs. Several measures were taken to reduce sampling bias and to ensure the maintenance of the quality of samples and consequent data analyses. Patient biopsy specimens were obtained from a single center by a designated surgical oncologist. Sample processing for NGS was carried out in the same batch eliminating the introduction of batch effects.

6.5 Intrinsic characterization of molecular subtypes of skeletal muscle from patients with cancer

Since my study is the only investigation performed in cachexia literature that identified subtypes of skeletal muscle from patients with cancer, there was no standard method I could undertake to understand which of the two subtypes is affected by cachexia. Therefore, I performed two independent benchmark analyses described as follows:

Clinical benchmarking analyses: I first explored the skeletal muscle subtypes for their clinical relevance. This was accomplished using the association statistics using the distribution of samples from the two subtypes in males and females to (i) WL-BMI grading schema [10](grades 0-4 represent increased graded severity) and (ii) z-score SMI distribution generated from literature-reported sex- and age-adjusted SMI values [29]. Reports from the individual analyses construed subtype one as a relatively high-grade cachexia group or the group affected by cachexia (with high WL-BMI and low SMI). Further, a slight variation was observed in subtype 2 (low-grade cachexia group/high SMI), with 10% of the patients within subtype 2 falling under the WL-BMI grade 4.

Molecular and functional benchmarking analyses: Next, I interrogated the subtypes transcriptome profiles and their pathway analysis results from IPA to the in-house generated experimental model systems (RD cells in undifferentiated and differentiated skeletal muscle from tumor-implanted rodents). I acknowledge the limitations of using these experimental model systems (RDs being a female cell lineage and rodents being all females). However, the premise was to test the pathways identified from muscle subtypes in a model that emulates myoblast proliferating and differentiating states (i.e., RD dataset) and an animal model of cancer-induced muscle wasting (i.e., rat transcriptome dataset). It is not to conclude the use of any particular systems but to understand the intrinsic characteristics of skeletal muscle subtypes from patients with cancer. These datasets were generated in-house. I performed all the data preprocessing and analysis regimen, which is also one of the strengths of this study, that all sequencing datasets are generated from the same lab.

The overarching theme emerging from this study underpins the implications and dysregulation of several coordinated events that regulate skeletal muscle structure, function, and regenerative process. The disturbances in the component mechanisms may contribute to the pathophysiology

of cancer-associated muscle wasting. I present a schematic model of perturbations in the skeletal muscle microenvironment and tumor-elicited inflammatory responses in **Figure 6.1.**

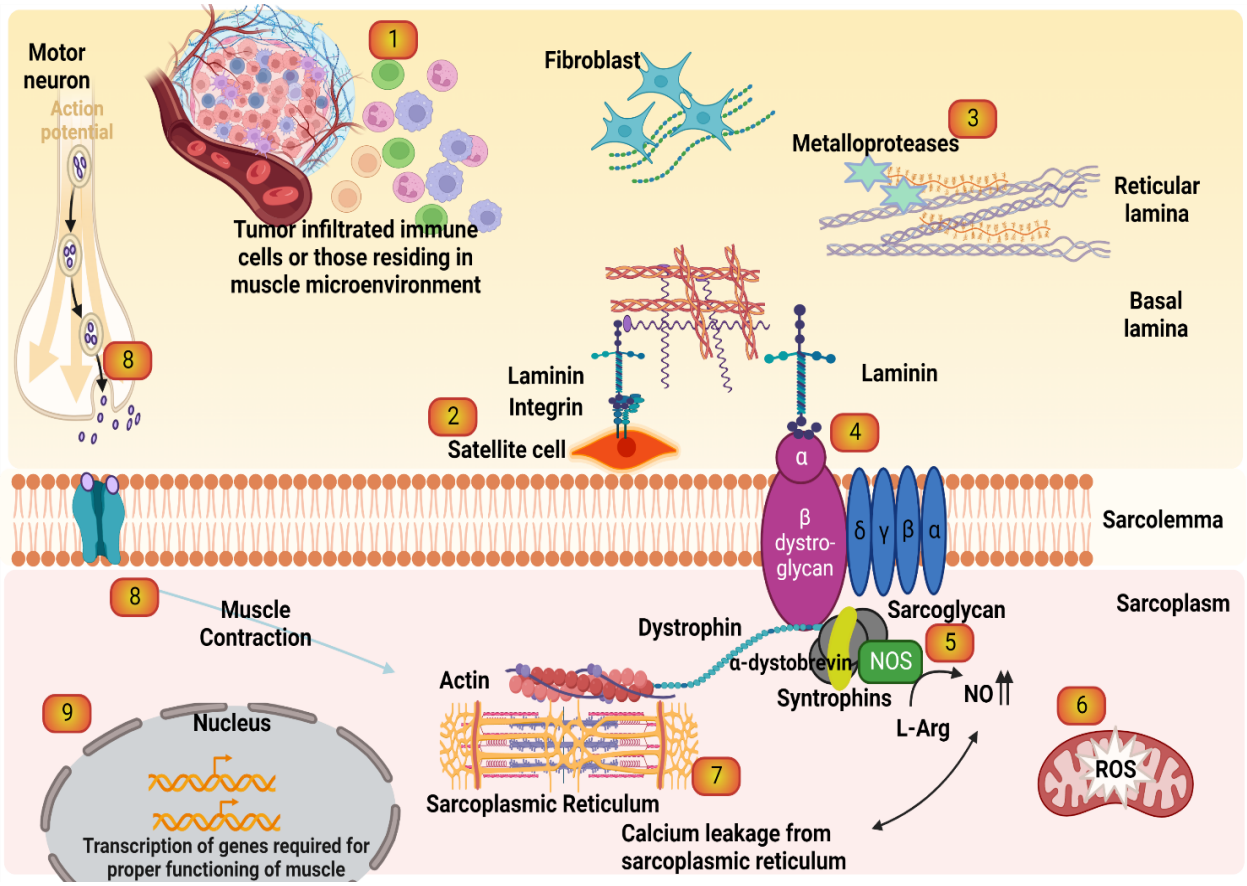


Figure 6.1 A proposed schematic of various cellular and functional components contributing to skeletal muscle homeostasis

The disrupted homeostatic response from multi-component mechanisms regulating skeletal muscle structure and function is shown. Skeletal muscle wasting in patients with cancer appears to be highly orchestrated. Individual component systems were shown in the numbered format in the above schematic, and their detailed description is shown below. (1) Sustained inflammatory response (2) Satellite cell and their microenvironment (3) ECM regulation by MMPs (4) DGC complex (5) nNOS signaling (6) Mitochondrial dysfunction (7) Calcium signaling (8) Synaptic junction (9) Transcriptional regulation. Created with Biorender.com

Pathways illustrated were found to be perturbed in subtype 1 (IPA predictions). The component systems concerned were described in detail in Chapter 5, section 5.4, and Figure 5.4.

A proposed schematic of various cellular and functional components contributing to skeletal muscle homeostasis

I also present a model of the series of events or cycle of a feed-forward process from individual components that could contribute to the pathophysiology of cachexia is depicted in **Figure 5.4**.

This is the first study using the human skeletal muscle that has demonstrated the involvement of several fundamental factors integral to skeletal muscle integrity, the most important being the crosstalk between muscle and its microenvironment. A series of events affecting the muscle regenerative process, such as sustained chronic inflammatory response, perturbations in Extracellular Matrix regulation, Calcium signaling, nNOS signaling, transcriptional dysregulation, and metabolic disturbances, were some of the key mechanisms observed.

This is the first human skeletal muscle study from patients with cancer to investigate the pathways from subtypes and report the importance of the skeletal muscle microenvironment and the interactions of various components in a sex-specific manner. This further explains the complexity of cachexia syndrome and the rationale for not looking from a focused lens but rather understanding cachexia as a multifaceted syndrome involving multiple mediators and signaling pathways. Overall, results from both the benchmark analyses align; therefore, subtype 1 is assigned as a group affected by cachexia. This study provides a roadmap for future implementations wherein a higher sample size could segregate the subtypes into multiple other subtypes that could indicate the graded severity of cachexia.

Systemic inflammation is one of the hallmarks of cancer-associated muscle wasting. Proinflammatory factors are released by complex host-tumor interactions and act centrally to directly mediate muscle loss by activating pathways leading to muscle protein degradation or peripherally through interactions with other organ systems having an indirect effects on skeletal muscle wasting. Direct activation of proinflammatory cytokines leads to the activation of pro-catabolic pathways leading to muscle wasting; these include NF-kB signaling, JAK/STAT signaling, MAPK, and SMAD signaling pathways [30]. All these eventually culminated in increased proteolysis or muscle protein breakdown and decreased protein synthesis, as have been identified

in experimental model systems. However, activation of proteolytic systems or decreased protein synthesis were not recapitulated in human studies [31,32], including this study. Proinflammatory cytokines such as IL6, IL1B, and TNF-α stimulate the release of acute phase protein (APP) from the liver. These APPs include C-Reactive Protein (CRP), coagulation, complement proteins, and transport proteins, as well as proteinase inhibitors such as α1 antitrypsin. Massart et al. (2020) suggested skeletal muscle as being another source of APPs using C26 cancer cachexia model, and that activation of APPs were mediated by glucocorticoids and proinflammatory cytokines [33].

Elevated levels of plasma or serum CRP is recognized as a marker of systemic inflammation and cachexia in patients with cancer [34]. However, the clinical utility of CRP is still debated due to the varying cutoffs; the most commonly used is greater than 10mg/L [35]. I was fortunate to have plasma CRP data from a subset of patients for whom I had the muscle biopsy specimen and transcriptome datasets to perform correlation analyses. However, when plasma CRP was correlated with tissue muscle expression profiles of proinflammatory cytokines and candidate genes selected from cachexia literature related to skeletal muscle regeneration, it showed no correlation or poor correlation in male and female patients, suggesting the criteria to re-evaluate the significance of a single marker to address the multifactorial problem of cachexia (data not shown). Despite the activation of downstream catabolic events by these cytokines, antibody targetting individual cytokines is ineffective in reversing all the features of cancer-associated muscle wasting. This is because the activation of one cytokine entity triggers the downstream effector molecules and complementary cytokines that bring about changes in signaling pathways.

In addition to the primary role of inflammatory cascades, skeletal muscle serves as an important metabolic center of regulation. Skeletal muscle is the largest reservoir of amino acids and is important for glucose homeostasis. The amino acids released from the breakdown of proteins in

muscle are utilized by the liver for gluconeogenesis. Insulin resistance has been shown to be another feature of cachexia described in experimental models. All these metabolic perturbations were observed in my dataset, along with the upregulation of proinflammatory markers such as IL6, IL1B, and TNF in males and females in subtype 1.

In a recent study by Abdelmoez et al. (2020) it was identified that the in vitro model systems used, such as rat L6 cells [36] derived from the thigh muscle, mouse C2C12 derived from thigh muscle [37], and primary human skeletal muscle cells (HSMCs) derived from human biopsies [38], exhibit heterogeneity in their transcriptome profiles and metabolism [39]. Hence, caution should be applied in using these model systems. Researchers should utilize these models to address their specific research question to gain biological insights considering their strengths and limitations. Also, these studies have been performed using mRNA expression from representative pathways (select candidate gene expression profiles). As my study explained, higher-order regulatory mechanisms have never been performed. The top six hub lncRNAs identified with high stringency cutoffs as described in this thesis, and with a greater proportion of interacting miRNAs and mRNAs demonstrated, could be subject to further interrogation using in vitro and in vivo studies since they regulate a large number of targets. This could potentially be one of the ways to study the integrated gene and pathway networks instead of targeting individual pathways. These could possibly be druggable targets, provided their validity in cell line models, or validation in non-invasive serum or plasma samples, provided these samples show circulating RNAs of sufficient selectivity and specificity for the diagnostic, prognostic, or predictive value to guide treatment decisions.

6.6 Limitations of the study

Since this is the first study to perform molecular subtyping of human skeletal muscle from patients with cancer, there was no previously generated reference dataset to predict the inherent type or

ability to characterize the subtypes functionally or to correlate the findings from this study. Nevertheless, I successfully characterized the subtypes in males and females to overcome such limitations.

Inherent characterization of skeletal muscle subtypes encompassed several validation methods, including aligning samples within the subtypes with those classified based on the WL-BMI grading system and z-SMI distribution. Subsequently, I performed functional benchmarking to understand the intrinsic characteristics of skeletal muscle subtypes. The perturbations in pathways related to the proinflammatory signaling cascade, nNOS signaling, tumor microenvironment, inhibition of matrix metalloproteases, calcium signaling, extracellular matrix regulation, activation of Acute Phase Reactant (APR) proteins, and metabolic aberrations were enriched in subtype 1 relative to subtype 2. While the molecular subtype identification and benchmarking methods (both clinical and functional benchmarking) are novel and presented for the first time in literature, this is not to imply that one method is adequate. I undertook two independent paradigms to reach a consensus on the subtype characteristics.

The sample size in this cohort was n=84; although this is the single most extensive sex-specific study of skeletal muscle transcriptome in cachexia, increasing the sample size could possibly result in identifying subclusters within these clusters and could be representative of trajectory-specific markers. A higher sample size will also allow analysis and interpretations from a single cancer diagnosis should there be any dependence on cancer type when the rectus abdominis is the only muscle analyzed. However, this needs to be determined.

Published studies in human skeletal muscle transcriptome were limited to mRNAs. Such datasets in the public domain could not be used to validate the classification of subtypes in my study since

these lack profiling of lncRNAs and small non-coding RNAs from the same patient's rectus abdominis muscle. Where miRNA data is available [40], lack of access to RNA from published cohorts or availability of RNA isolated from rectus abdominis muscle from our previous studies (Damaraju and Baracos labs) hindered profiling of all RNA classes.

Data on muscle fiber typing and its role in the selective degradation of muscle mass remain controversial [41]. I did not have immunohistological or fiber-type data for my study cohort. Muscle heterogeneity is reflected by fiber type composition differing in myosin heavy chain isoforms, metabolism, and neural response. Slow type I fibers express MYH7, and fast type 2 fibers express MYH2 (type 2A), MYH1 (type 2x), and MYH4 (type 2B). While no significant differences were observed in MYH expression for males, women have more type I fiber expressing MYH7; in this cohort, MYH7 expression was down-regulated by -1.52 fold in subtype 1 vs. 2. I cannot comment on the fiber type data as it was not one of the study objectives and hence, it is out of the scope of the current study. Nevertheless, the lower expression of myosin heavy chain in subtype 1 further strengthens the annotation of the subtypes.

For this project, the skeletal muscle biopsy sampling procurement was at a fixed time point during the surgery. Hence dynamic changes in the muscle of patients with cancer are unattainable. The subtype annotation limitations are also from single biopsy profiling, whereas WL and SMI are from varying time points. Collection of serial biopsies pre- and post-surgery along the disease trajectory would be useful information to annotate subtype-related transcriptional changes. However, this is not a feasible or ethical approach considering the invasiveness of the surgical procedure. Nevertheless, the unsupervised clustering and identification of subtypes are encouraging, given the above sampling limitations.

6.7 Strengths of this study

- This is the first ever largest dataset generated of human skeletal muscle from patients with cancer that has complete (i) clinical and body composition annotation for each patient included in the study (ii) Muscle transcriptional landscape was generated for each patient using Next Generation rRNA depleted RNA sequencing (to profile mRNAs and lncRNAs) and Small RNA sequencing (to profile miRNAs, piRNAs, snoRNAs, and tRNAs). Using a high-throughput NGS platform has significant advantages over qRT-PCR and microarray-based technology.
- My entire dataset from the initial data acquisition of muscle biopsy (from a designated surgical oncologist) and subsequent NGS profiling experiments was handled in a single batch to overcome reagent and processing artifacts. Quality checks for all samples enabled the downstream data processing to ensure confidence in the data analysis and interpretations. Such a rigorous approach may facilitate the reproducibility of this study's findings in independent cohorts.
- Patient stratification using cohort-specific median cutoffs and their distributions led me to infer that the application of classically used cachexia variables or clinical labels was not an appropriate method for the profiling study conducted in this thesis. This is important because muscle biopsy was obtained at a single time at the time of surgery, and it becomes imperative to study the transcriptome independent of clinical labels. It opens new avenues to study cachexia and its underlying pathophysiological mechanisms from a pragmatic view.
- Overcoming the challenge of classification bias was a significant breakthrough for my downstream analyses. The unsupervised clustering method identified sex-specific differential gene expression profiles and molecular subtypes of skeletal muscle in patients with cancer.

- I adopted two de novo avenues to establish subtype characterization: (i) clinical and (ii) molecular and functional benchmarking. The validation from both approaches led to determining subtype one as a group affected by cachexia. Overall, the molecular and functional benchmarking suggested disturbances within the skeletal muscle milieu contributing to the pathophysiology of cachexia.

6.8 References

1. Narasimhan A, Zhong X, Au EP, Ceppa EP, Nakeeb A, House MG, et al. Profiling of Adipose and Skeletal Muscle in Human Pancreatic Cancer Cachexia Reveals Distinct Gene Profiles with Convergent Pathways. *Cancers (Basel).* 2021;13(8).
2. Fearon KC, Glass DJ, Guttridge DC. Cancer cachexia: mediators, signaling, and metabolic pathways. *Cell Metab.* 2012;16(2):153-166.
3. Baracos VE. Bridging the gap: are animal models consistent with clinical cancer cachexia? *Nat Rev Clin Oncol.* 2018;15(4):197-198.
4. Mueller TC, Bachmann J, Prokopchuk O, Friess H, Martignoni ME. Molecular pathways leading to loss of skeletal muscle mass in cancer cachexia--can findings from animal models be translated to humans? *BMC Cancer.* 2016;16:75.
5. Anoveros-Barrera A, Bhullar AS, Stretch C, Esfandiari N, Dunichand-Hoedl AR, Martins KJB, et al. Clinical and biological characterization of skeletal muscle tissue biopsies of surgical cancer patients. *J Cachexia Sarcopenia Muscle.* 2019;10(6):1356-1377.
6. Prado CM, Lieffers JR, McCargar LJ, Reiman T, Sawyer MB, Martin L, et al. Prevalence and clinical implications of sarcopenic obesity in patients with solid tumours of the respiratory and gastrointestinal tracts: a population-based study. *Lancet Oncol.* 2008;9(7):629-635.

7. Martin L, Birdsell L, Macdonald N, Reiman T, Clandinin MT, McCargar LJ, et al. Cancer cachexia in the age of obesity: skeletal muscle depletion is a powerful prognostic factor, independent of body mass index. *J Clin Oncol.* 2013;31(12):1539-1547.

8. Roeland EJ, Ma JD, Nelson SH, Seibert T, Heavey S, Revta C, et al. Weight loss versus muscle loss: re-evaluating inclusion criteria for future cancer cachexia interventional trials. *Support Care Cancer.* 2017;25(2):365-369.

9. Anoveros-Barrera A, Bhullar AS, Stretch C, Dunichand-Hoedl AR, Martins KJB, Rieger A, et al. Immunohistochemical phenotyping of T cells, granulocytes, and phagocytes in the muscle of cancer patients: association with radiologically defined muscle mass and gene expression. *Skelet Muscle.* 2019;9(1):24.

10. Martin L, Senesse P, Gioulbasanis I, Antoun S, Bozzetti F, Deans C, et al. Diagnostic criteria for the classification of cancer-associated weight loss. *J Clin Oncol.* 2015;33(1):90-99.

11. Goodpaster BH, Kelley DE, Thaete FL, He J, Ross R. Skeletal muscle attenuation determined by computed tomography is associated with skeletal muscle lipid content. *J Appl Physiol (1985).* 2000;89(1):104-110.

12. Kroenke CH, Prado CM, Meyerhardt JA, Weltzien EK, Xiao J, Cespedes Feliciano EM, et al. Muscle radiodensity and mortality in patients with colorectal cancer. *Cancer.* 2018;124(14):3008-3015.

13. Carvalho ALM, Gonzalez MC, Sousa IM, das Virgens IPA, Medeiros GOC, Oliveira MN, et al. Low skeletal muscle radiodensity is the best predictor for short-term major surgical complications in gastrointestinal surgical cancer: A cohort study. *PLoS One.* 2021;16(2):e0247322.

14. Kim J, Han SH, Kim HI. Detection of sarcopenic obesity and prediction of long-term survival in patients with gastric cancer using preoperative computed tomography and machine learning. *J Surg Oncol.* 2021;124(8):1347-1355.

15. Lortie J, Gage G, Rush B, Heymsfield SB, Szczykutowicz TP, Kuchnia AJ. The effect of computed tomography parameters on sarcopenia and myosteatosis assessment: a scoping review. *J Cachexia Sarcopenia Muscle.* 2022;13(6):2807-2819.

16. Brunet JP, Tamayo P, Golub TR, Mesirov JP. Metagenes and molecular pattern discovery using matrix factorization. *Proc Natl Acad Sci U S A.* 2004;101(12):4164-4169.

17. Yoon HG, Oh D, Noh JM, Cho WK, Sun JM, Kim HK, et al. Machine learning model for predicting excessive muscle loss during neoadjuvant chemoradiotherapy in oesophageal cancer. *J Cachexia Sarcopenia Muscle.* 2021;12(5):1144-1152.

18. Yin L, Cui J, Lin X, Li N, Fan Y, Zhang L, et al. Identifying cancer cachexia in patients without weight loss information: machine learning approaches to address a real-world challenge. *Am J Clin Nutr.* 2022.

19. Anyene I, Caan B, Williams GR, Popuri K, Lenchik L, Giri S, et al. Body composition from single versus multi-slice abdominal computed tomography: Concordance and associations with colorectal cancer survival. *J Cachexia Sarcopenia Muscle.* 2022;13(6):2974-2984.

20. Kilgour RD, Vigano A, Trutschnigg B, Hornby L, Lucar E, Bacon SL, et al. Cancer-related fatigue: the impact of skeletal muscle mass and strength in patients with advanced cancer. *J Cachexia Sarcopenia Muscle.* 2010;1(2):177-185.

21. Tumasian RA, 3rd, Harish A, Kundu G, Yang JH, Ubaida-Mohien C, Gonzalez-Freire M, et al. Skeletal muscle transcriptome in healthy aging. *Nat Commun.* 2021;12(1):2014.

22. Mahmassani ZS, Reidy PT, McKenzie AI, Stubben C, Howard MT, Drummond MJ. Age-dependent skeletal muscle transcriptome response to bed rest-induced atrophy. *J Appl Physiol (1985).* 2019;126(4):894-902.

23. Makhnovskii PA, Bokov RO, Kolpakov FA, Popov DV. Transcriptomic Signatures and Upstream Regulation in Human Skeletal Muscle Adapted to Disuse and Aerobic Exercise. *Int J Mol Sci.* 2021;22(3).

24. Rundqvist HC, Montelius A, Osterlund T, Norman B, Esbjornsson M, Jansson E. Acute sprint exercise transcriptome in human skeletal muscle. *PLoS One.* 2019;14(10):e0223024.

25. Dickinson JM, D'Lugos AC, Naymik MA, Siniard AL, Wolfe AJ, Curtis DR, et al. Transcriptome response of human skeletal muscle to divergent exercise stimuli. *J Appl Physiol (1985).* 2018;124(6):1529-1540.

26. Pourteymour S, Eckardt K, Holen T, Langleite T, Lee S, Jensen J, et al. Global mRNA sequencing of human skeletal muscle: Search for novel exercise-regulated myokines. *Mol Metab.* 2017;6(4):352-365.

27. Lindholm ME, Huss M, Solnestam BW, Kjellqvist S, Lundeberg J, Sundberg CJ. The human skeletal muscle transcriptome: sex differences, alternative splicing, and tissue homogeneity assessed with RNA sequencing. *FASEB J.* 2014;28(10):4571-4581.

28. Terry EE, Zhang X, Hoffmann C, Hughes LD, Lewis SA, Li J, et al. Transcriptional profiling reveals extraordinary diversity among skeletal muscle tissues. *Elife.* 2018;7.

29. Kazemi-Bajestani SM, Mazurak VC, Baracos V. Computed tomography-defined muscle and fat wasting are associated with cancer clinical outcomes. *Semin Cell Dev Biol.* 2016;54:2-10.

30. Webster JM, Kempen L, Hardy RS, Langen RCJ. Inflammation and Skeletal Muscle Wasting During Cachexia. *Front Physiol.* 2020;11:597675.

31. Stephens NA, Gallagher IJ, Rooyackers O, Skipworth RJ, Tan BH, Marstrand T, et al. Using transcriptomics to identify and validate novel biomarkers of human skeletal muscle cancer cachexia. *Genome Med.* 2010;2(1):1.

32. Gallagher IJ, Stephens NA, MacDonald AJ, Skipworth RJ, Husi H, Greig CA, et al. Suppression of skeletal muscle turnover in cancer cachexia: evidence from the transcriptome in sequential human muscle biopsies. *Clin Cancer Res.* 2012;18(10):2817-2827.

33. Massart IS, Paulissen G, Loumaye A, Lause P, Potgens SA, Thibaut MM, et al. Marked Increased Production of Acute Phase Reactants by Skeletal Muscle during Cancer Cachexia. *Cancers (Basel).* 2020;12(11).

34. Mallard J, Gagez AL, Baudinet C, Herbinet A, Maury J, Bernard PL, et al. C-Reactive Protein Level: A Key Predictive Marker of Cachexia in Lymphoma and Myeloma Patients. *J Hematol.* 2019;8(2):55-59.

35. Tavares P, Goncalves DM, Santos LL, Ferreira R. Revisiting the clinical usefulness of C-reactive protein in the set of cancer cachexia. *Porto Biomed J.* 2021;6(1):e123.

36. Richler C, Yaffe D. The in vitro cultivation and differentiation capacities of myogenic cell lines. *Dev Biol.* 1970;23(1):1-22.

37. Yaffe D, Saxel O. Serial passaging and differentiation of myogenic cells isolated from dystrophic mouse muscle. *Nature.* 1977;270(5639):725-727.

38. Blau HM, Webster C. Isolation and characterization of human muscle cells. *Proc Natl Acad Sci U S A.* 1981;78(9):5623-5627.

39. Abdelmoez AM, Sardon Puig L, Smith JAB, Gabriel BM, Savikj M, Dollet L, et al. Comparative profiling of skeletal muscle models reveals heterogeneity of transcriptome and metabolism. *Am J Physiol Cell Physiol.* 2020;318(3):C615-C626.

40. Narasimhan A, Ghosh S, Stretch C, Greiner R, Bathe OF, Baracos V, et al. Small RNAome profiling from human skeletal muscle: novel miRNAs and their targets associated with cancer cachexia. *J Cachexia Sarcopenia Muscle.* 2017;8(3):405-416.

41. Cosper PF, Leinwand LA. Myosin heavy chain is not selectively decreased in murine cancer cachexia. *Int J Cancer.* 2012;130(11):2722-2727.

Chapter 7 Future directions

My thesis focussed on identifying skeletal muscle subtypes from patients with cancer using transcriptome profiles and unsupervised clustering algorithms. The intrinsic characteristics attained from clinical (graded WL-BMI severity and low z-SMI scores) and functional (or molecular) benchmarking analyses demonstrated subtype 1 as a group affected by cachexia. This is the first pivotal transcriptome investigational study, and the consecutive text describes future considerations.

7.1 Validation of the present study in an independent patient cohort

This study focussed on the accrual of *rectus abdominis* skeletal muscle biopsies from a cohort of patients with pancreas and colorectal cancer types due to the higher prevalence of cachexia in these cancer types [1]. Restricted choice of the cancer types was to prevent the heterogeneity related to them since different cancer types may exert distinct or overlapping muscle wasting phenomena in terms of their transcriptional landscape, which needs to be ascertained. Studying cancer-type-specific muscle wasting in individual cancer types is also imperative. Hence, the study findings must be replicated and validated in an independent patient cohort. Since this is the first whole transcriptome study of human skeletal muscle within patients with cancer, it could serve as a reference dataset for future validation studies.

7.2 Sample size considerations: Identification of cancer-specific and sex-specific molecular subtypes with higher sample size

While this study focussed on identifying sex-specific molecular subtypes, it would be equally essential to analyze cancer-type- and sex-specific molecular subtypes. In this study, the molecular subtypes or clusters were restricted to the k=2 cluster model, considering the sample size

limitations. However, this is not to say that there are no subclusters within a given cluster. Increasing the sample size would help elucidate the subclusters that could potentially corroborate the stage-specific trajectory of cachexia, i.e., determine subtypes in early vs. late phases based on the associated graded cachexia severity. It needs to be determined using an independent patient cohort and would require collaborative efforts between researchers and surgical oncologists to perform an extensive cancer cachexia-omic study. I did not have any previously generated RNA sequencing data from the skeletal muscle of patients with cancer to compare the identified clusters or to perform functional benchmarking of the clusters. Hence the dataset in my study could serve as a reference cohort in future cachexia research.

7.3 Functional characterization

LncRNAs are diverse molecules and exhibit functions based on their cellular localization. lncRNAs interact with DNA, RNA, and protein to exert their gene regulatory functions. I generated the ceRNA networks in this study, and hub lncRNAs were determined based on the highest interacting miRNA and mRNA partners. As the transcriptome profiles derived from tissue data do not entirely recapitulate the expression profiles from the cell-line data, the next step would be to confirm the expression of ceRNAs, and their interacting miRNA and mRNAs. One of the ways to validate the ceRNA mechanism would be performing overexpression or knockdown experiments of lncRNAs and confirming the similar abundance of corresponding mRNA targets. Further confirmation of the ceRNA mechanism could be obtained by generating Dicer knockout cell lines, i.e., that are essentially deficient in miRNA biogenesis. The Dicer knockout cells should not have any effects on the other miRNAs. To initiate these, the specific targets should be the hub lncRNAs that demonstrate the highest miRNA binding sites and interacting target genes. In addition to the ceRNA-mediated regulation, the cis- and trans-regulatory genes could be predicted

using bedtools suite [2] that could have functions in skeletal muscle wasting or muscle pathology. Further experimental models could be generated for candidate lncRNAs via knockdown or overexpression followed by sequencing to confirm the gene targets. lncRNAs also have epigenetic roles wherein their association with chromatin-modifying complexes, for example, polycomb repressive complex 2 (PRC2), leads to changes in gene expression [3-5]. Identifying such mechanisms in human skeletal muscle could shed light on lncRNA-mediated gene expression at the epigenetic levels.

7.4 Expand the ceRNA premise to piRNA-mediated crosstalk with lncRNAs and mRNAs to find unique and overlapping gene targets and pathways

I performed molecular subtype identification using unsupervised clustering in the human skeletal muscle of patients with cancer. Non-negative Matrix Factorization[6,7] (NMF) was used to identify clusters based on a single RNA class. The final ascertainment of molecular subtypes was based on integrative NMF[8] (intNMF) clusters or molecular subtypes. To gain mechanistic insights using the post-transcriptional ceRNA mechanism[9], I utilized lncRNA-miRNA-mRNA triplet RNA crosstalk and network-based analysis. piRNAs have a similar post-transcriptional mechanism of action as miRNAs, i.e., they bind to the 3'UTR of the mRNAs and lead to translational repression. Previously thought to be expressed in the germline, they are now known to be expressed in somatic tissues, and their dysregulation contributes to pathological states. I identified the Differentially Expressed (DE) piRNAs in subtype 1 vs. 2 in males and females. Future studies could explore lncRNA-piRNA-mRNA crosstalk and compare the gene targets to that of lncRNA-miRNA-mRNA crosstalk to determine piRNA- and miRNA-specific gene targets and pathway level regulation. This would emphasize the importance of the distinct or shared mechanisms of action between miRNA and piRNA-mediated gene regulation.

7.5 lncRNA and snoRNA embedded small RNAs in human skeletal muscle

Recent studies have identified that lncRNA and snoRNA can serve as precursors to miRNA and piRNAs, although the exact mechanism is still unknown[10-13]. It would be interesting to determine if any of the profiled lncRNAs and snoRNAs in the skeletal muscle dataset from my study harbor any miRNAs and piRNAs. The embedded miRNA and piRNAs could be checked for expression in the in-house skeletal muscle dataset generated. Further, their corresponding gene targets regulate common and distinct pathways. It is a novel phenomenon and remains to be determined in human skeletal muscle from patients with cancer.

7.6 Multi-omic molecular characterization of human skeletal muscle and their integrative gene regulatory mechanisms in cancer-associated muscle wasting

Future advancements in the field would be significant to identify molecular subtypes using multi-omic platforms to unravel the mechanisms. However, acquiring a skeletal muscle biopsy is an invasive and arduous procedure and requires extensive collaborative efforts to accomplish the task. Few projects worldwide perform data integration at the systems biology levels, particularly in the cancer context requiring large-scale consortia efforts from different expertise and sampling collection. Pursuing such a comprehensive study in cachexia research would require extensive funding, sampling and clinical data collection, and collaborative efforts between clinicians and researchers. It could potentially open up avenues to form consortia in the research area focussing on cachexia in patients with cancer. The data integration methods could also help identify the subtype of patients undergoing muscle-wasting to target therapeutic modalities and identify druggable targets for those patients. This is the first report to suggest nuance approaches as the one undertaken in the current study.

7.7 References

Baracos VE, Martin L, Korc M, Guttridge DC, Fearon KCH. Cancer-associated cachexia. *Nat Rev Dis Primers.* 2018;4:17105.

2. Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics.* 2010;26(6):841-842.

3. Khalil AM, Guttman M, Huarte M, Garber M, Raj A, Rivea Morales D, et al. Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. *Proc Natl Acad Sci U S A.* 2009;106(28):11667-11672.

4. Gupta RA, Shah N, Wang KC, Kim J, Horlings HM, Wong DJ, et al. Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature.* 2010;464(7291):1071-1076.

5. Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, Brugmann SA, et al. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell.* 2007;129(7):1311-1323.

6. Brunet JP, Tamayo P, Golub TR, Mesirov JP. Metagenes and molecular pattern discovery using matrix factorization. *Proc Natl Acad Sci U S A.* 2004;101(12):4164-4169.

7. Gaujoux R, Seoighe C. A flexible R package for nonnegative matrix factorization. *BMC Bioinformatics.* 2010;11:367.

8. Chalise P, Fridley BL. Integrative clustering of multi-level 'omic data based on non-negative matrix factorization algorithm. *PLoS One.* 2017;12(5):e0176278.

9. Salmena L, Poliseno L, Tay Y, Kats L, Pandolfi PP. A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language? *Cell.* 2011;146(3):353-358.

10. He D, Wu D, Muller S, Wang L, Saha P, Ahanger SH, et al. miRNA-independent function of long noncoding pri-miRNA loci. *Proc Natl Acad Sci U S A.* 2021;118(13).

11. Ulitsky I. Interactions between short and long noncoding RNAs. *FEBS Lett.* 2018;592(17):2874-2883.

12. Zhong F, Zhou N, Wu K, Guo Y, Tan W, Zhang H, et al. A SnoRNA-derived piRNA interacts with human interleukin-4 pre-mRNA and induces its decay in nuclear exosomes. *Nucleic Acids Res.* 2015;43(21):10474-10491.

13. Krishnan P, Ghosh S, Graham K, Mackey JR, Kovalchuk O, Damaraju S. Piwi-interacting RNAs and PIWI genes as novel prognostic markers for breast cancer. *Oncotarget.* 2016;7(25):37944-37956.

Chapter 8 Conclusions

Molecular association studies applying binary phenotype classes have limitations when multiple RNA classes are to be studied. Investigator-defined cut-offs primarily drive phenotype-specific patient stratification in the cachexia context for WL, BMI, SMI, or combinations of these variables. Literature-reported findings from molecular association studies attempted in isolation for mRNA and miRNA using human skeletal muscle biopsies have limited value in understanding the complex interplay of RNA molecules in a cellular milieu. My thesis surmounts these limitations. Overall, two de novo molecular subtypes using the intNMF algorithm were identified from the skeletal muscle of patients with cancer. This is the first large-scale investigation providing a transcriptional atlas of human skeletal muscle from patients with cancer. Matched expression profiles using high-throughput NGS sequencing of small RNA dataset and RNA sequencing dataset is one of a kind and the first in cachexia research. Data curated for the entire transcriptional atlas of mRNAs, lncRNAs, and small ncRNAs will be accessible publicly. This would help enable cachexia researchers to study their genes or pathway of interest unique to a particular sex or overlapping targets. It is a prerequisite to deposit sequencing data and cite the Gene Expression Omnibus (GEO; https://www.ncbi.nlm.nih.gov/geo/) Accession Number before communication. GEO releases the data upon publication of the original findings.

The study results in a paradigm shift from applying clinical labels to the pragmatic application of classifying patients with cachexia and pursuing molecular studies using unbiased pattern recognition ML algorithms. Further, my study provides insights into multilayered RNA crosstalk via competing endogenous RNA (ceRNA) analysis of lncRNA-miRNA-mRNA triplets. The hub lncRNAs demonstrated the sexually dimorphic pattern of post-transcriptional gene regulation.

Two independent paradigm avenues revealed subtype 1 as a group affected by cachexia. *Clinical benchmarking* showed an association of subtype 1 with increased severity of WL-BMI grades and low muscle mass; subtype 2 had minimal WL and high SMI. Further clinical relevance of subtypes was obtained using survival analysis. Subtype 1 showed poor overall survival (Kaplan-Meier method), and differences in survival were statistically significant at a log-rank P-value of 0.05 (data not shown). *Molecular and functional benchmarking* using experimental model systems led me to generate a pathway map of component pathways important for skeletal muscle integrity and function. Drugs developed thus far using targeted mechanisms at a single gene product or a single pathway was inadequate as judged from the clinical benefit to patients to alleviate cachexia condition. Stratifying patients for clinical trials should consider the complex interplay of clinical variables and an unbiased approach to identifying at-risk patients for cachexia. Though not ready for clinical implementation, the study premise described in this thesis suffices to bring to the fore independent methods to understand the complex muscle homeostasis in patients with cancer at a molecular level. My study could be a reference for future studies to validate the findings.

Bibliography

Abdelmoez AM, Sardon Puig L, Smith JAB, Gabriel BM, Savikj M, Dollet L, et al. Comparative profiling of skeletal muscle models reveals heterogeneity of transcriptome and metabolism. *Am J Physiol Cell Physiol.* 2020;318(3):C615-C626.

Ala U, Karreth FA, Bosia C, Pagnani A, Taulli R, Leopold V, et al. Integrated transcriptional and competitive endogenous RNA networks are cross-regulated in permissive molecular environments. *Proc Natl Acad Sci U S A.* 2013;110(18):7154-7159.

Alameddine HS, Morgan JE. Matrix Metalloproteinases and Tissue Inhibitor of Metalloproteinases in Inflammation and Fibrosis of Skeletal Muscles. *J Neuromuscul Dis.* 2016;3(4):455-473.

Almasud AA, Giles KH, Miklavcic JJ, Martins KJB, Baracos VE, Putman CT, et al. Fish oil mitigates myosteatosis and improves chemotherapy efficacy in a preclinical model of colon cancer. *PLoS One.* 2017;12(8):e0183576.

Anoveros-Barrera A, Bhullar AS, Stretch C, Dunichand-Hoedl AR, Martins KJB, Rieger A, et al. Immunohistochemical phenotyping of T cells, granulocytes, and phagocytes in the muscle of cancer patients: association with radiologically defined muscle mass and gene expression. *Skelet Muscle.* 2019;9(1):24.

Anoveros-Barrera A, Bhullar AS, Stretch C, Esfandiari N, Dunichand-Hoedl AR, Martins KJB, et al. Clinical and biological characterization of skeletal muscle tissue biopsies of surgical cancer patients. *J Cachexia Sarcopenia Muscle.* 2019;10(6):1356-1377.

Anyene I, Caan B, Williams GR, Popuri K, Lenchik L, Giri S, et al. Body composition from single versus multi-slice abdominal computed tomography: Concordance and

associations with colorectal cancer survival. *J Cachexia Sarcopenia Muscle.* 2022;13(6):2974-2984.

Aravin A, Gaidatzis D, Pfeffer S, Lagos-Quintana M, Landgraf P, Iovino N, et al. A novel class of small RNAs bind to MILI protein in mouse testes. *Nature.* 2006;442(7099):203-207.

Argiles JM, Busquets S, Stemmler B, Lopez-Soriano FJ. Cancer cachexia: understanding the molecular basis. *Nat Rev Cancer.* 2014;14(11):754-762.

Arita T, Ichikawa D, Konishi H, Komatsu S, Shiozaki A, Shoda K, et al. Circulating long non-coding RNAs in plasma of patients with gastric cancer. *Anticancer Res.* 2013;33(8):3185-3193.

Arneson PC, Doles JD. Impaired Muscle Regeneration in Cancer-Associated Cachexia. *Trends Cancer.* 2019;5(10):579-582.

Ballarino M, Morlando M, Fatica A, Bozzoni I. Non-coding RNAs in muscle differentiation and musculoskeletal disease. *J Clin Invest.* 2016;126(6):2021-2030.

Baracos VE. Bridging the gap: are animal models consistent with clinical cancer cachexia? *Nat Rev Clin Oncol.* 2018;15(4):197-198.

Baracos VE, Martin L, Korc M, Guttridge DC, Fearon KCH. Cancer-associated cachexia. *Nat Rev Dis Primers.* 2018;4:17105.

Baracos VE, Reiman T, Mourtzakis M, Gioulbasanis I, Antoun S. Body composition in patients with non-small cell lung cancer: a contemporary view of cancer cachexia with the use of computed tomography image analysis. *Am J Clin Nutr.* 2010;91(4):1133S-1137S.

Barret M, Antoun S, Dalban C, Malka D, Mansourbakht T, Zaanan A, et al. Sarcopenia is linked to treatment toxicity in patients with metastatic colorectal cancer. *Nutr Cancer.* 2014;66(4):583-589.

Bartel DP. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell.* 2004;116(2):281-297.

Bartel DP. MicroRNAs: target recognition and regulatory functions. *Cell.* 2009;136(2):215-233.

Blau HM, Webster C. Isolation and characterization of human muscle cells. *Proc Natl Acad Sci U S A.* 1981;78(9):5623-5627.

Boehm I, Miller J, Wishart TM, Wigmore SJ, Skipworth RJ, Jones RA, et al. Neuromuscular junctions are stable in patients with cancer cachexia. *J Clin Invest.* 2020;130(3):1461-1465.

Bosia C, Pagnani A, Zecchina R. Modelling Competing Endogenous RNA Networks. *PLoS One.* 2013;8(6):e66609.

Bossola M, Marzetti E, Rosa F, Pacelli F. Skeletal muscle regeneration in cancer cachexia. *Clin Exp Pharmacol Physiol.* 2016;43(5):522-527.

Bossola M, Muscaritoli M, Costelli P, Grieco G, Bonelli G, Pacelli F, et al. Increased muscle proteasome activity correlates with disease severity in gastric cancer patients. *Ann Surg.* 2003;237(3):384-389.

Bouchard-Bourelle P, Desjardins-Henri C, Mathurin-St-Pierre D, Deschamps-Francoeur G, Fafard-Couture E, Garant JM, et al. snoDB: an interactive database of human snoRNA sequences, abundance and interactions. *Nucleic Acids Res.* 2020;48(D1):D220-D225.

Bredella MA. Sex Differences in Body Composition. *Adv Exp Med Biol.* 2017;1043:9-27.

Brummer A, Hausser J. MicroRNA binding sites in the coding region of mRNAs: extending the repertoire of post-transcriptional gene regulation. *Bioessays.* 2014;36(6):617-626.

Brunet JP, Tamayo P, Golub TR, Mesirov JP. Metagenes and molecular pattern discovery using matrix factorization. *Proc Natl Acad Sci U S A.* 2004;101(12):4164-4169.

Buckingham M, Vincent SD. Distinct and dynamic myogenic populations in the vertebrate embryo. *Curr Opin Genet Dev.* 2009;19(5):444-453.

Cabili MN, Trapnell C, Goff L, Koziol M, Tazon-Vega B, Regev A, et al. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev.* 2011;25(18):1915-1927.

Cai R, Zhang Q, Wang Y, Yong W, Zhao R, Pang W. Lnc-ORA interacts with microRNA-532-3p and IGF2BP2 to inhibit skeletal muscle myogenesis. *J Biol Chem.* 2021;296:100376.

Cancer Genome Atlas N. Comprehensive molecular characterization of human colon and rectal cancer. *Nature.* 2012;487(7407):330-337.

Cancer Genome Atlas Research Network. Electronic address aadhe, Cancer Genome Atlas Research N. Integrated Genomic Characterization of Pancreatic Ductal Adenocarcinoma. *Cancer Cell.* 2017;32(2):185-203 e113.

Carmona-Saez P, Pascual-Marqui RD, Tirado F, Carazo JM, Pascual-Montano A. Biclustering of gene expression data by Non-smooth Non-negative Matrix Factorization. *BMC Bioinformatics.* 2006;7:78.

Carninci P, Kasukawa T, Katayama S, Gough J, Frith MC, Maeda N, et al. The transcriptional landscape of the mammalian genome. *Science.* 2005;309(5740):1559-1563.

Carson JA, Baltgalvis KA. Interleukin 6 as a key regulator of muscle mass during cachexia. *Exerc Sport Sci Rev.* 2010;38(4):168-176.

Carvalho ALM, Gonzalez MC, Sousa IM, das Virgens IPA, Medeiros GOC, Oliveira MN, et al. Low skeletal muscle radiodensity is the best predictor for short-term major surgical complications in gastrointestinal surgical cancer: A cohort study. *PLoS One.* 2021;16(2):e0247322.

Cesana M, Cacchiarelli D, Legnini I, Santini T, Sthandier O, Chinappi M, et al. A long noncoding RNA controls muscle differentiation by functioning as a competing endogenous RNA. *Cell.* 2011;147(2):358-369.

Chalise P, Fridley BL. Integrative clustering of multi-level 'omic data based on non-negative matrix factorization algorithm. *PLoS One.* 2017;12(5):e0176278.

Chan PP, Lowe TM. GtRNAdb 2.0: an expanded database of transfer RNA genes identified in complete and draft genomes. *Nucleic Acids Res.* 2016;44(D1):D184-189.

Coenen-Stass AML, Sork H, Gatto S, Godfrey C, Bhomra A, Krjutskov K, et al. Comprehensive RNA-Sequencing Analysis in Serum and Muscle Reveals Novel Small RNA Signatures with Biomarker Potential for DMD. *Mol Ther Nucleic Acids.* 2018;13:1-15.

Cosper PF, Leinwand LA. Myosin heavy chain is not selectively decreased in murine cancer cachexia. *Int J Cancer.* 2012;130(11):2722-2727.

Csapo R, Gumpenberger M, Wessner B. Skeletal Muscle Extracellular Matrix - What Do We Know About Its Composition, Regulation, and Physiological Roles? A Narrative Review. *Front Physiol.* 2020;11:253.

D'Orlando C, Marzetti E, Francois S, Lorenzi M, Conti V, di Stasio E, et al. Gastric cancer does not affect the expression of atrophy-related genes in human skeletal muscle. *Muscle Nerve.* 2014;49(4):528-533.

Daly L, Dolan R, Power D, Ni Bhuachalla E, Sim W, Fallon M, et al. The relationship between the BMI-adjusted weight loss grading system and quality of life in patients with incurable cancer. *J Cachexia Sarcopenia Muscle.* 2020;11(1):160-168.

De Palma C, Morisi F, Pambianco S, Assi E, Touvier T, Russo S, et al. Deficient nitric oxide signalling impairs skeletal muscle growth and performance: involvement of mitochondrial dysregulation. *Skelet Muscle.* 2014;4(1):22.

Derstine BA, Holcombe SA, Ross BE, Wang NC, Su GL, Wang SC. Skeletal muscle cutoff values for sarcopenia diagnosis using T10 to L5 measurements in a healthy US population. *Sci Rep.* 2018;8(1):11369.

Dev R, Bruera E, Dalal S. Insulin resistance and body composition in cancer patients. *Ann Oncol.* 2018;29(suppl_2):ii18-ii26.

Devine RD, Bicer S, Reiser PJ, Velten M, Wold LE. Metalloproteinase expression is altered in cardiac and skeletal muscle in cancer cachexia. *Am J Physiol Heart Circ Physiol.* 2015;309(4):H685-691.

Dey BK, Pfeifer K, Dutta A. The H19 long noncoding RNA gives rise to microRNAs miR-675-3p and miR-675-5p to promote skeletal muscle differentiation and regeneration. *Genes Dev.* 2014;28(5):491-501.

Dickinson JM, D'Lugos AC, Naymik MA, Siniard AL, Wolfe AJ, Curtis DR, et al. Transcriptome response of human skeletal muscle to divergent exercise stimuli. *J Appl Physiol (1985).* 2018;124(6):1529-1540.

Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics.* 2013;29(1):15-21.

Dumont NA, Bentzinger CF, Sincennes MC, Rudnicki MA. Satellite Cells and Skeletal Muscle Regeneration. *Compr Physiol.* 2015;5(3):1027-1059.

Dunne RF, Loh KP, Williams GR, Jatoi A, Mustian KM, Mohile SG. Cachexia and Sarcopenia in Older Adults with Cancer: A Comprehensive Review. *Cancers (Basel).* 2019;11(12).

Ebert MS, Neilson JR, Sharp PA. MicroRNA sponges: competitive inhibitors of small RNAs in mammalian cells. *Nat Methods.* 2007;4(9):721-726.

Ebhardt HA, Degen S, Tadini V, Schilb A, Johns N, Greig CA, et al. Comprehensive proteome analysis of human skeletal muscle in cachexia and sarcopenia: a pilot study. *J Cachexia Sarcopenia Muscle.* 2017;8(4):567-582.

Enright AJ, John B, Gaul U, Tuschl T, Sander C, Marks DS. MicroRNA targets in Drosophila. *Genome Biol.* 2003;5(1):R1.

Fearon K, Arends J, Baracos V. Understanding the mechanisms and treatment options in cancer cachexia. *Nat Rev Clin Oncol.* 2013;10(2):90-99.

Fearon K, Strasser F, Anker SD, Bosaeus I, Bruera E, Fainsinger RL, et al. Definition and classification of cancer cachexia: an international consensus. *Lancet Oncol.* 2011;12(5):489-495.

Fearon KC, Glass DJ, Guttridge DC. Cancer cachexia: mediators, signaling, and metabolic pathways. *Cell Metab.* 2012;16(2):153-166.

Fernandez GJ, Ferreira JH, Vechetti IJ, Jr., de Moraes LN, Cury SS, Freire PP, et al. MicroRNA-mRNA Co-sequencing Identifies Transcriptional and Post-transcriptional Regulatory Networks Underlying Muscle Wasting in Cancer Cachexia. *Front Genet.* 2020;11:541.

Figliuzzi M, Marinari E, De Martino A. MicroRNAs as a selective channel of communication between competing RNAs: a steady-state theory. *Biophys J.* 2013;104(5):1203-1213.

Freire PP, Fernandez GJ, de Moraes D, Cury SS, Dal Pai-Silva M, Dos Reis PP, et al. The expression landscape of cachexia-inducing factors in human cancers. *J Cachexia Sarcopenia Muscle.* 2020;11(4):947-961.

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

Fu X, Li S, Jia M, Xu B, Yang L, Ma R, et al. Myogenesis controlled by a long non-coding RNA 1700113A16RIK and post-transcriptional regulation. *Cell Regen.* 2022;11(1):13.

Gallagher D, Ruts E, Visser M, Heshka S, Baumgartner RN, Wang J, et al. Weight stability masks sarcopenia in elderly men and women. *Am J Physiol Endocrinol Metab.* 2000;279(2):E366-375.

Gallagher IJ, Stephens NA, MacDonald AJ, Skipworth RJ, Husi H, Greig CA, et al. Suppression of skeletal muscle turnover in cancer cachexia: evidence from the transcriptome in sequential human muscle biopsies. *Clin Cancer Res.* 2012;18(10):2817-2827.

Gaujoux R, Seoighe C. A flexible R package for nonnegative matrix factorization. *BMC Bioinformatics.* 2010;11:367.

Ge Y, Chen J. MicroRNAs in skeletal myogenesis. *Cell Cycle.* 2011;10(3):441-448.

Goodpaster BH, Kelley DE, Thaete FL, He J, Ross R. Skeletal muscle attenuation determined by computed tomography is associated with skeletal muscle lipid content. *J Appl Physiol (1985).* 2000;89(1):104-110.

Grozdanovic Z, Baumgarten HG. Nitric oxide synthase in skeletal muscle fibers: a signaling component of the dystrophin-glycoprotein complex. *Histol Histopathol.* 1999;14(1):243-256.

Gupta RA, Shah N, Wang KC, Kim J, Horlings HM, Wong DJ, et al. Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature.* 2010;464(7291):1071-1076.

Guttman M, Amit I, Garber M, French C, Lin MF, Feldser D, et al. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature.* 2009;458(7235):223-227.

Ha M, Kim VN. Regulation of microRNA biogenesis. *Nat Rev Mol Cell Biol.* 2014;15(8):509-524.

Han YN, Li Y, Xia SQ, Zhang YY, Zheng JH, Li W. PIWI Proteins and PIWI-Interacting RNA: Emerging Roles in Cancer. *Cell Physiol Biochem.* 2017;44(1):1-20.

He D, Wu D, Muller S, Wang L, Saha P, Ahanger SH, et al. miRNA-independent function of long noncoding pri-miRNA loci. *Proc Natl Acad Sci U S A.* 2021;118(13).

He WA, Berardi E, Cardillo VM, Acharyya S, Aulino P, Thomas-Ahner J, et al. NF-kappaB-mediated Pax7 dysregulation in the muscle microenvironment promotes cancer cachexia. *J Clin Invest.* 2013;123(11):4821-4835.

Hu C, Liu K, Wang B, Xu W, Lin Y, Yuan C. DLX6-AS1: An Indispensable Cancer-related Long Non-coding RNA. *Curr Pharm Des.* 2021;27(9):1211-1218.

Huang HY, Lin YC, Li J, Huang KY, Shrestha S, Hong HC, et al. miRTarBase 2020: updates to the experimentally validated microRNA-target interaction database. *Nucleic Acids Res.* 2020;48(D1):D148-D154.

Huot JR, Novinger LJ, Pin F, Narasimhan A, Zimmers TA, O'Connell TM, et al. Formation of colorectal liver metastases induces musculoskeletal and metabolic abnormalities consistent with exacerbated cachexia. *JCI Insight.* 2020;5(9).

Hutchins LN, Murphy SM, Singh P, Graber JH. Position-dependent motif characterization using non-negative matrix factorization. *Bioinformatics.* 2008;24(23):2684-2690.

Inaba S, Hinohara A, Tachibana M, Tsujikawa K, Fukada SI. Muscle regeneration is disrupted by cancer cachexia without loss of muscle stem cell potential. *PLoS One.* 2018;13(10):e0205467.

Iyer MK, Niknafs YS, Malik R, Singhal U, Sahu A, Hosono Y, et al. The landscape of long noncoding RNAs in the human transcriptome. *Nat Genet.* 2015;47(3):199-208.

Jacobsen A, Wen J, Marks DS, Krogh A. Signatures of RNA binding proteins globally coupled to effective microRNA target sites. *Genome Res.* 2010;20(8):1010-1019.

Jayarathna DK, Renteria ME, Sauret E, Batra J, Gandhi NS. Identifying Complex lncRNA/Pseudogene-miRNA-mRNA Crosstalk in Hormone-Dependent Cancers. *Biology (Basel).* 2021;10(10).

Jin JJ, Lv W, Xia P, Xu ZY, Zheng AD, Wang XJ, et al. Long noncoding RNA SYISL regulates myogenesis by interacting with polycomb repressive complex 2. *Proc Natl Acad Sci U S A.* 2018;115(42):E9802-E9811.

Johns N, Hatakeyama S, Stephens NA, Degen M, Degen S, Frieauff W, et al. Clinical classification of cancer cachexia: phenotypic correlates in human skeletal muscle. *PLoS One.* 2014;9(1):e83618.

Johns N, Stephens NA, Fearon KC. Muscle wasting in cancer. *Int J Biochem Cell Biol.* 2013;45(10):2215-2229.

Judge SM, Nosacka RL, Delitto D, Gerber MH, Cameron ME, Trevino JG, et al. Skeletal Muscle Fibrosis in Pancreatic Cancer Patients with Respect to Survival. *JNCI Cancer Spectr.* 2018;2(3):pky043.

Karagkouni D, Paraskevopoulou MD, Chatzopoulos S, Vlachos IS, Tastsoglou S, Kanellos I, et al. DIANA-TarBase v8: a decade-long collection of experimentally supported miRNA-gene interactions. *Nucleic Acids Res.* 2018;46(D1):D239-D245.

Karagkouni D, Paraskevopoulou MD, Tastsoglou S, Skoufos G, Karavangeli A, Pierros V, et al. DIANA-LncBase v3: indexing experimentally supported miRNA targets on non-coding transcripts. *Nucleic Acids Res.* 2020;48(D1):D101-D110.

Karalaki M, Fili S, Philippou A, Koutsilieris M. Muscle regeneration: cellular and molecular events. *In Vivo.* 2009;23(5):779-796.

Kartha RV, Subramanian S. Competing endogenous RNAs (ceRNAs): new entrants to the intricacies of gene regulation. *Front Genet.* 2014;5:8.

Kazemi-Bajestani SM, Mazurak VC, Baracos V. Computed tomography-defined muscle and fat wasting are associated with cancer clinical outcomes. *Semin Cell Dev Biol.* 2016;54:2-10.

Kechin A, Boyarskikh U, Kel A, Filipenko M. cutPrimers: A New Tool for Accurate Cutting of Primers from Reads of Targeted Next Generation Sequencing. *J Comput Biol.* 2017;24(11):1138-1143.

Khal J, Hine AV, Fearon KC, Dejong CH, Tisdale MJ. Increased expression of proteasome subunits in skeletal muscle of cancer patients with weight loss. *Int J Biochem Cell Biol.* 2005;37(10):2196-2206.

Khal J, Wyke SM, Russell ST, Hine AV, Tisdale MJ. Expression of the ubiquitin-proteasome pathway and muscle loss in experimental cancer cachexia. *Br J Cancer.* 2005;93(7):774-780.

Khalil AM, Guttman M, Huarte M, Garber M, Raj A, Rivea Morales D, et al. Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. *Proc Natl Acad Sci U S A.* 2009;106(28):11667-11672.

Kilgour RD, Vigano A, Trutschnigg B, Hornby L, Lucar E, Bacon SL, et al. Cancer-related fatigue: the impact of skeletal muscle mass and strength in patients with advanced cancer. *J Cachexia Sarcopenia Muscle.* 2010;1(2):177-185.

Kim H, Park H. Sparse non-negative matrix factorizations via alternating non-negativity-constrained least squares for microarray data analysis. *Bioinformatics.* 2007;23(12):1495-1502.

Kim J, Han SH, Kim HI. Detection of sarcopenic obesity and prediction of long-term survival in patients with gastric cancer using preoperative computed tomography and machine learning. *J Surg Oncol.* 2021;124(8):1347-1355.

Kiss T. Small nucleolar RNAs: an abundant group of noncoding RNAs with diverse cellular functions. *Cell.* 2002;109(2):145-148.

Kobayashi J, Uchida H, Kofuji A, Ito J, Shimizu M, Kim H, et al. Molecular regulation of skeletal muscle mass and the contribution of nitric oxide: A review. *FASEB Bioadv.* 2019;1(6):364-374.

Konermann S, Brigham MD, Trevino AE, Joung J, Abudayyeh OO, Barcena C, et al. Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature.* 2015;517(7536):583-588.

Kopp F, Mendell JT. Functional Classification and Experimental Dissection of Long Noncoding RNAs. *Cell.* 2018;172(3):393-407.

Kozomara A, Birgaoanu M, Griffiths-Jones S. miRBase: from microRNA sequences to function. *Nucleic Acids Res.* 2019;47(D1):D155-D162.

Kramer A, Green J, Pollard J, Jr., Tugendreich S. Causal analysis approaches in Ingenuity Pathway Analysis. *Bioinformatics.* 2014;30(4):523-530.

Krishnan P, Ghosh S, Graham K, Mackey JR, Kovalchuk O, Damaraju S. Piwi-interacting RNAs and PIWI genes as novel prognostic markers for breast cancer. *Oncotarget.* 2016;7(25):37944-37956.

Krishnan P, Ghosh S, Wang B, Heyns M, Graham K, Mackey JR, et al. Profiling of Small Nucleolar RNAs by Next Generation Sequencing: Potential New Players for Breast Cancer Prognosis. *PLoS One.* 2016;11(9):e0162622.

Krishnan P, Ghosh S, Wang B, Heyns M, Li D, Mackey JR, et al. Genome-wide profiling of transfer RNAs and their role as novel prognostic markers for breast cancer. *Sci Rep.* 2016;6:32843.

Krishnan P, Ghosh S, Wang B, Li D, Narasimhan A, Berendt R, et al. Next generation sequencing profiling identifies miR-574-3p and miR-660-5p as potential novel prognostic markers for breast cancer. *BMC Genomics.* 2015;16:735.

Kroenke CH, Prado CM, Meyerhardt JA, Weltzien EK, Xiao J, Cespedes Feliciano EM, et al. Muscle radiodensity and mortality in patients with colorectal cancer. *Cancer.* 2018;124(14):3008-3015.

Kung JT, Colognori D, Lee JT. Long noncoding RNAs: past, present, and future. *Genetics.* 2013;193(3):651-669.

La Greca A, Scarafia MA, Hernandez Canas MC, Perez N, Castaneda S, Colli C, et al. PIWI-interacting RNAs are differentially expressed during cardiac differentiation of human pluripotent stem cells. *PLoS One.* 2020;15(5):e0232715.

Landen S, Hiam D, Voisin S, Jacques M, Lamon S, Eynon N. Physiological and molecular sex differences in human skeletal muscle in response to exercise training. *J Physiol.* 2021.

Langmead B. Aligning short sequencing reads with Bowtie. *Curr Protoc Bioinformatics.* 2010;Chapter 11:Unit 11 17.

Le Grand F, Rudnicki MA. Skeletal muscle satellite cells and adult myogenesis. *Curr Opin Cell Biol.* 2007;19(6):628-633.

Lee DD, Seung HS. Learning the parts of objects by non-negative matrix factorization. *Nature.* 1999;401(6755):788-791.

Lee I, Ajay SS, Yook JI, Kim HS, Hong SH, Kim NH, et al. New class of microRNA targets containing simultaneous 5'-UTR and 3'-UTR interaction sites. *Genome Res.* 2009;19(7):1175-1183.

Lee RC, Ambros V. An extensive class of small RNAs in Caenorhabditis elegans. *Science.* 2001;294(5543):862-864.

Lewis BP, Shih IH, Jones-Rhoades MW, Bartel DP, Burge CB. Prediction of mammalian microRNA targets. *Cell.* 2003;115(7):787-798.

Li J, Wang N, Zhang F, Jin S, Dong Y, Dong X, et al. PIWI-interacting RNAs are aberrantly expressed and may serve as novel biomarkers for diagnosis of lung adenocarcinoma. *Thorac Cancer.* 2021;12(18):2468-2477.

Li R, Li B, Cao Y, Li W, Dai W, Zhang L, et al. Long non-coding RNA Mir22hg-derived miR-22-3p promotes skeletal muscle differentiation and regeneration by inhibiting HDAC4. *Mol Ther Nucleic Acids.* 2021;24:200-211.

Li R, Qu H, Wang S, Wei J, Zhang L, Ma R, et al. GDCRNATools: an R/Bioconductor package for integrative analysis of lncRNA, miRNA and mRNA data in GDC. *Bioinformatics.* 2018;34(14):2515-2517.

Li X, Du L, Liu Q, Lu Z. MicroRNAs: Novel players in the diagnosis and treatment of cancer cachexia (Review). *Exp Ther Med.* 2022;24(1):446.

Li Z, Cai B, Abdalla BA, Zhu X, Zheng M, Han P, et al. LncIRS1 controls muscle atrophy via sponging miR-15 family to activate IGF1-PI3K/AKT pathway. *J Cachexia Sarcopenia Muscle.* 2019;10(2):391-410.

Lieffers JR, Bathe OF, Fassbender K, Winget M, Baracos VE. Sarcopenia is associated with postoperative infection and delayed recovery from colorectal cancer resection surgery. *Br J Cancer.* 2012;107(6):931-936.

Lindholm ME, Huss M, Solnestam BW, Kjellqvist S, Lundeberg J, Sundberg CJ. The human skeletal muscle transcriptome: sex differences, alternative splicing, and tissue homogeneity assessed with RNA sequencing. *FASEB J.* 2014;28(10):4571-4581.

Liu Q, Deng J, Qiu Y, Gao J, Li J, Guan L, et al. Non-coding RNA basis of muscle atrophy. *Mol Ther Nucleic Acids.* 2021;26:1066-1078.

Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods.* 2001;25(4):402-408.

Loreti M, Sacco A. The jam session between muscle stem cells and the extracellular matrix in the tissue microenvironment. *NPJ Regen Med.* 2022;7(1):16.

Lortie J, Gage G, Rush B, Heymsfield SB, Szczykutowicz TP, Kuchnia AJ. The effect of computed tomography parameters on sarcopenia and myosteatosis assessment: a scoping review. *J Cachexia Sarcopenia Muscle.* 2022;13(6):2807-2819.

Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014;15(12):550.

Lu L, Sun K, Chen X, Zhao Y, Wang L, Zhou L, et al. Genome-wide survey by ChIP-seq reveals YY1 regulation of lincRNAs in skeletal myogenesis. *EMBO J.* 2013;32(19):2575-2588.

Lu Z, Meng L, Sun Z, Shi X, Shao W, Zheng Y, et al. Differentially Expressed Genes and Enriched Signaling Pathways in the Adipose Tissue of Obese People. *Front Genet.* 2021;12:620740.

Lundholm K, Bylund AC, Holm J, Schersten T. Skeletal muscle metabolism in patients with malignant tumor. *Eur J Cancer (1965).* 1976;12(6):465-473.

Lundholm K, Edstrom S, Ekman L, Karlberg I, Bylund AC, Schersten T. A comparative study of the influence of malignant tumor on host metabolism in mice and man: evaluation of an experimental model. *Cancer.* 1978;42(2):453-461.

Maher AC, Fu MH, Isfort RJ, Varbanov AR, Qu XA, Tarnopolsky MA. Sex differences in global mRNA content of human skeletal muscle. *PLoS One.* 2009;4(7):e6335.

Mahmassani ZS, Reidy PT, McKenzie AI, Stubben C, Howard MT, Drummond MJ. Age-dependent skeletal muscle transcriptome response to bed rest-induced atrophy. *J Appl Physiol (1985).* 2019;126(4):894-902.

Makhnovskii PA, Bokov RO, Kolpakov FA, Popov DV. Transcriptomic Signatures and Upstream Regulation in Human Skeletal Muscle Adapted to Disuse and Aerobic Exercise. *Int J Mol Sci.* 2021;22(3).

Mallard J, Gagez AL, Baudinet C, Herbinet A, Maury J, Bernard PL, et al. C-Reactive Protein Level: A Key Predictive Marker of Cachexia in Lymphoma and Myeloma Patients. *J Hematol.* 2019;8(2):55-59.

Marceca GP, Nigita G, Calore F, Croce CM. MicroRNAs in Skeletal Muscle and Hints on Their Potential Role in Muscle Wasting During Cancer Cachexia. *Front Oncol.* 2020;10:607196.

Martin L, Birdsell L, Macdonald N, Reiman T, Clandinin MT, McCargar LJ, et al. Cancer cachexia in the age of obesity: skeletal muscle depletion is a powerful prognostic factor, independent of body mass index. *J Clin Oncol.* 2013;31(12):1539-1547.

Martin L, Senesse P, Gioulbasanis I, Antoun S, Bozzetti F, Deans C, et al. Diagnostic criteria for the classification of cancer-associated weight loss. *J Clin Oncol.* 2015;33(1):90-99.

Massart IS, Paulissen G, Loumaye A, Lause P, Potgens SA, Thibaut MM, et al. Marked Increased Production of Acute Phase Reactants by Skeletal Muscle during Cancer Cachexia. *Cancers (Basel).* 2020;12(11).

Mattick JS, Makunin IV. Non-coding RNA. *Hum Mol Genet.* 2006;15 Spec No 1:R17-29.

Mauro A. Satellite cell of skeletal muscle fibers. *J Biophys Biochem Cytol.* 1961;9(2):493-495.

McCann KL, Kavari SL, Burkholder AB, Phillips BT, Hall TMT. H/ACA snoRNA levels are regulated during stem cell differentiation. *Nucleic Acids Res.* 2020;48(15):8686-8703.

Merico D, Gfeller D, Bader GD. How to visually interpret biological data using networks. *Nat Biotechnol.* 2009;27(10):921-924.

Moffitt RA, Marayati R, Flate EL, Volmar KE, Loeza SG, Hoadley KA, et al. Virtual microdissection identifies distinct tumor- and stroma-specific subtypes of pancreatic ductal adenocarcinoma. *Nat Genet.* 2015;47(10):1168-1178.

Montalvo RN, Counts BR, Carson JA. Understanding sex differences in the regulation of cancer-induced muscle wasting. *Curr Opin Support Palliat Care.* 2018;12(4):394-403.

Moreira-Pais A, Ferreira R, Oliveira PA, Duarte JA. Sarcopenia versus cancer cachexia: the muscle wasting continuum in healthy and diseased aging. *Biogerontology.* 2021;22(5):459-477.

Mourtzakis M, Prado CM, Lieffers JR, Reiman T, McCargar LJ, Baracos VE. A practical and precise approach to quantification of body composition in cancer patients using computed tomography images acquired during routine care. *Appl Physiol Nutr Metab.* 2008;33(5):997-1006.

Mueller TC, Bachmann J, Prokopchuk O, Friess H, Martignoni ME. Molecular pathways leading to loss of skeletal muscle mass in cancer cachexia--can findings from animal models be translated to humans? *BMC Cancer.* 2016;16:75.

Mukund K, Subramaniam S. Skeletal muscle: A review of molecular structure and function, in health and disease. *Wiley Interdiscip Rev Syst Biol Med.* 2020;12(1):e1462.

Muller S, Raulefs S, Bruns P, Afonso-Grunz F, Plotner A, Thermann R, et al. Next-generation sequencing reveals novel differentially regulated mRNAs, lncRNAs, miRNAs, sdRNAs and a piRNA in pancreatic cancer. *Mol Cancer.* 2015;14:94.

Narasimhan A, Ghosh S, Stretch C, Greiner R, Bathe OF, Baracos V, et al. Small RNAome profiling from human skeletal muscle: novel miRNAs and their targets associated with cancer cachexia. *J Cachexia Sarcopenia Muscle.* 2017;8(3):405-416.

Narasimhan A, Greiner R, Bathe OF, Baracos V, Damaraju S. Differentially expressed alternatively spliced genes in skeletal muscle from cancer patients with cachexia. *J Cachexia Sarcopenia Muscle.* 2018;9(1):60-70.

Narasimhan A, Zhong X, Au EP, Ceppa EP, Nakeeb A, House MG, et al. Profiling of Adipose and Skeletal Muscle in Human Pancreatic Cancer Cachexia Reveals Distinct Gene Profiles with Convergent Pathways. *Cancers (Basel).* 2021;13(8).

Norman K, Stobaus N, Reiss J, Schulzke J, Valentini L, Pirlich M. Effect of sexual dimorphism on muscle strength in cachexia. *J Cachexia Sarcopenia Muscle.* 2012;3(2):111-116.

Novak JS, Mazala DAG, Nearing M, Hindupur R, Uapinyoying P, Habib NF, et al. Human muscle stem cells are refractory to aging. *Aging Cell.* 2021;20(7):e13411.

Oh J, Takahashi R, Adachi E, Kondo S, Kuratomi S, Noma A, et al. Mutations in two matrix metalloproteinase genes, MMP-2 and MT1-MMP, are synthetic lethal in mice. *Oncogene.* 2004;23(29):5041-5048.

Oldenhuis CN, Oosting SF, Gietema JA, de Vries EG. Prognostic versus predictive value of biomarkers in oncology. *Eur J Cancer.* 2008;44(7):946-953.

Olson B, Norgard MA, Levasseur PR, Zhu X, Marks DL. Physiologic and molecular characterization of a novel murine model of metastatic head and neck cancer cachexia. *J Cachexia Sarcopenia Muscle.* 2021;12(5):1312-1332.

Ozata DM, Gainetdinov I, Zoch A, O'Carroll D, Zamore PD. PIWI-interacting RNAs: small RNAs with big functions. *Nat Rev Genet.* 2019;20(2):89-108.

Penna F, Costamagna D, Fanzani A, Bonelli G, Baccino FM, Costelli P. Muscle wasting and impaired myogenesis in tumor bearing mice are prevented by ERK inhibition. *PLoS One.* 2010;5(10):e13604.

Percival JM. nNOS regulation of skeletal muscle fatigue and exercise performance. *Biophys Rev.* 2011;3(4):209-217.

Pessina P, Conti V, Pacelli F, Rosa F, Doglietto GB, Brunelli S, et al. Skeletal muscle of gastric cancer patients expresses genes involved in muscle regeneration. *Oncol Rep.* 2010;24(3):741-745.

Piuco R, Galante PAF. piRNAdb: A piwi-interacting RNA database. *bioRxiv.* 2021:2021.2009.2021.461238.

Pourteymour S, Eckardt K, Holen T, Langleite T, Lee S, Jensen J, et al. Global mRNA sequencing of human skeletal muscle: Search for novel exercise-regulated myokines. *Mol Metab.* 2017;6(4):352-365.

Prado CM, Baracos VE, McCargar LJ, Mourtzakis M, Mulder KE, Reiman T, et al. Body composition as an independent determinant of 5-fluorouracil-based chemotherapy toxicity. *Clin Cancer Res.* 2007;13(11):3264-3268.

Prado CM, Baracos VE, McCargar LJ, Reiman T, Mourtzakis M, Tonkin K, et al. Sarcopenia as a determinant of chemotherapy toxicity and time to tumor progression in metastatic breast cancer patients receiving capecitabine treatment. *Clin Cancer Res.* 2009;15(8):2920-2926.

Prado CM, Lieffers JR, McCargar LJ, Reiman T, Sawyer MB, Martin L, et al. Prevalence and clinical implications of sarcopenic obesity in patients with solid tumours of the respiratory and gastrointestinal tracts: a population-based study. *Lancet Oncol.* 2008;9(7):629-635.

Prado CM, Sawyer MB, Ghosh S, Lieffers JR, Esfandiari N, Antoun S, et al. Central tenet of cancer cachexia therapy: do patients with advanced cancer have exploitable anabolic potential? *Am J Clin Nutr.* 2013;98(4):1012-1019.

Qu CX, Shi XC, Zai LQ, Bi H, Yang Q. LncRNA CASC19 promotes the proliferation, migration and invasion of non-small cell lung carcinoma via regulating miRNA-130b-3p. *Eur Rev Med Pharmacol Sci.* 2019;23(3 Suppl):247-255.

Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics.* 2010;26(6):841-842.

Ramage MI, Johns N, Deans CDA, Ross JA, Preston T, Skipworth RJE, et al. The relationship between muscle protein content and CT-derived muscle radio-density in patients with upper GI cancer. *Clin Nutr.* 2018;37(2):752-754.

Rand WM. Objective Criteria for the Evaluation of Clustering Methods. *Journal of the American Statistical Association.* 1971;66(336):846-850.

Reddel CJ, Allen JD, Ehteda A, Taylor R, Chen VM, Curnow JL, et al. Increased thrombin generation in a mouse model of cancer cachexia is partially interleukin-6 dependent. *J Thromb Haemost.* 2017;15(3):477-486.

Rhoads MG, Kandarian SC, Pacelli F, Doglietto GB, Bossola M. Expression of NF-kappaB and IkappaB proteins in skeletal muscle of gastric cancer patients. *Eur J Cancer.* 2010;46(1):191-197.

Richler C, Yaffe D. The in vitro cultivation and differentiation capacities of myogenic cell lines. *Dev Biol.* 1970;23(1):1-22.

Rinn JL, Chang HY. Genome regulation by long noncoding RNAs. *Annu Rev Biochem.* 2012;81:145-166.

Rinn JL, Kertesz M, Wang JK, Squazzo SL, Xu X, Brugmann SA, et al. Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell.* 2007;129(7):1311-1323.

Rinninella E, Cintoni M, Raoul P, Pozzo C, Strippoli A, Bria E, et al. Muscle mass, assessed at diagnosis by L3-CT scan as a prognostic marker of clinical outcomes in patients with gastric cancer: A systematic review and meta-analysis. *Clin Nutr.* 2020;39(7):2045-2054.

Robertson AG, Kim J, Al-Ahmadie H, Bellmunt J, Guo G, Cherniack AD, et al. Comprehensive Molecular Characterization of Muscle-Invasive Bladder Cancer. *Cell.* 2018;174(4):1033.

Roeland EJ, Ma JD, Nelson SH, Seibert T, Heavey S, Revta C, et al. Weight loss versus muscle loss: re-evaluating inclusion criteria for future cancer cachexia interventional trials. *Support Care Cancer.* 2017;25(2):365-369.

Rundqvist HC, Montelius A, Osterlund T, Norman B, Esbjornsson M, Jansson E. Acute sprint exercise transcriptome in human skeletal muscle. *PLoS One.* 2019;14(10):e0223024.

Salmena L, Poliseno L, Tay Y, Kats L, Pandolfi PP. A ceRNA hypothesis: the Rosetta Stone of a hidden RNA language? *Cell.* 2011;146(3):353-358.

Sartori R, Romanello V, Sandri M. Mechanisms of muscle atrophy and hypertrophy: implications in health and disease. *Nat Commun.* 2021;12(1):330.

Schersten T, Lundholm K. Lysosomal enzyme activity in muscle tissue from patients with malignant tumor. *Cancer.* 1972;30(5):1246-1251.

Scott MS, Ono M. From snoRNA to miRNA: Dual function regulatory non-coding RNAs. *Biochimie.* 2011;93(11):1987-1992.

Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 2003;13(11):2498-2504.

Shen W, Punyanitya M, Wang Z, Gallagher D, St-Onge MP, Albu J, et al. Visceral adipose tissue: relations between single-slice areas and total volume. *Am J Clin Nutr.* 2004;80(2):271-278.

Shen W, Punyanitya M, Wang Z, Gallagher D, St-Onge MP, Albu J, et al. Total body skeletal muscle and adipose tissue volumes: estimation from a single abdominal cross-sectional image. *J Appl Physiol (1985).* 2004;97(6):2333-2338.

Siracusa J, Koulmann N, Banzet S. Circulating myomiRs: a new class of biomarkers to monitor skeletal muscle in physiology and medicine. *J Cachexia Sarcopenia Muscle.* 2018;9(1):20-27.

Smillie CL, Sirey T, Ponting CP. Complexities of post-transcriptional regulation and the modeling of ceRNA crosstalk. *Crit Rev Biochem Mol Biol.* 2018;53(3):231-245.

Smith GI, Mittendorfer B. Sexual dimorphism in skeletal muscle protein turnover. *J Appl Physiol (1985).* 2016;120(6):674-682.

Snyman C, Niesler CU. MMP-14 in skeletal muscle repair. *J Muscle Res Cell Motil.* 2015;36(3):215-225.

Stamler JS, Meissner G. Physiology of nitric oxide in skeletal muscle. *Physiol Rev.* 2001;81(1):209-237.

Statello L, Guo CJ, Chen LL, Huarte M. Gene regulation by long non-coding RNAs and its biological functions. *Nat Rev Mol Cell Biol.* 2021;22(2):96-118.

Stein-O'Brien GL, Arora R, Culhane AC, Favorov AV, Garmire LX, Greene CS, et al. Enter the Matrix: Factorization Uncovers Knowledge from Omics. *Trends Genet.* 2018;34(10):790-805.

Stephens NA, Gallagher IJ, Rooyackers O, Skipworth RJ, Tan BH, Marstrand T, et al. Using transcriptomics to identify and validate novel biomarkers of human skeletal muscle cancer cachexia. *Genome Med.* 2010;2(1):1.

Stephens NA, Skipworth RJ, Gallagher IJ, Greig CA, Guttridge DC, Ross JA, et al. Evaluating potential biomarkers of cachexia and survival in skeletal muscle of upper gastrointestinal cancer patients. *J Cachexia Sarcopenia Muscle.* 2015;6(1):53-61.

Stretch C, Aubin JM, Mickiewicz B, Leugner D, Al-Manasra T, Tobola E, et al. Sarcopenia and myosteatosis are accompanied by distinct biological profiles in patients with pancreatic and periampullary adenocarcinomas. *PLoS One.* 2018;13(5):e0196235.

Stretch C, Khan S, Asgarian N, Eisner R, Vaisipour S, Damaraju S, et al. Effects of sample size on differential gene expression, rank order and prediction accuracy of a gene signature. *PLoS One.* 2013;8(6):e65380.

Sun Y, Sun X, Liu S, Liu L, Chen J. The overlap between regeneration and fibrosis in injured skeletal muscle is regulated by phosphatidylinositol 3-kinase/Akt signaling pathway - A bioinformatic analysis based on lncRNA microarray. *Gene.* 2018;672:79-87.

Suzuki N, Motohashi N, Uezumi A, Fukada S, Yoshimura T, Itoyama Y, et al. NO production results in suspension-induced muscle atrophy through dislocation of neuronal NOS. *J Clin Invest.* 2007;117(9):2468-2476.

Suzuki T, Von Haehling S, Springer J. Promising models for cancer-induced cachexia drug discovery. *Expert Opin Drug Discov.* 2020;15(5):627-637.

Talbert EE, Cuitino MC, Ladner KJ, Rajasekerea PV, Siebert M, Shakya R, et al. Modeling Human Cancer-induced Cachexia. *Cell Rep.* 2019;28(6):1612-1622 e1614.

Talbert EE, Guttridge DC. Impaired regeneration: A role for the muscle microenvironment in cancer cachexia. *Semin Cell Dev Biol.* 2016;54:82-91.

Talbert EE, Lewis HL, Farren MR, Ramsey ML, Chakedis JM, Rajasekera P, et al. Circulating monocyte chemoattractant protein-1 (MCP-1) is associated with cachexia in treatment-naive pancreatic cancer patients. *J Cachexia Sarcopenia Muscle.* 2018;9(2):358-368.

Tavares P, Goncalves DM, Santos LL, Ferreira R. Revisiting the clinical usefulness of C-reactive protein in the set of cancer cachexia. *Porto Biomed J.* 2021;6(1):e123.

Tay Y, Rinn J, Pandolfi PP. The multilayered complexity of ceRNA crosstalk and competition. *Nature.* 2014;505(7483):344-352.

Terry EE, Zhang X, Hoffmann C, Hughes LD, Lewis SA, Li J, et al. Transcriptional profiling reveals extraordinary diversity among skeletal muscle tissues. *Elife.* 2018;7.

Thomas K, Engler AJ, Meyer GA. Extracellular matrix regulation in the muscle satellite cell niche. *Connect Tissue Res.* 2015;56(1):1-8.

Thomson DW, Dinger ME. Endogenous microRNA sponges: evidence and controversy. *Nat Rev Genet.* 2016;17(5):272-283.

Tomasin R, Martin A, Cominetti MR. Metastasis and cachexia: alongside in clinics, but not so in animal models. *J Cachexia Sarcopenia Muscle.* 2019;10(6):1183-1194.

Tong YS, Wang XW, Zhou XL, Liu ZH, Yang TX, Shi WH, et al. Identification of the long non-coding RNA POU3F3 in plasma as a novel biomarker for diagnosis of esophageal squamous cell carcinoma. *Mol Cancer.* 2015;14:3.

Tumasian RA, 3rd, Harish A, Kundu G, Yang JH, Ubaida-Mohien C, Gonzalez-Freire M, et al. Skeletal muscle transcriptome in healthy aging. *Nat Commun.* 2021;12(1):2014.

Ulitsky I. Interactions between short and long noncoding RNAs. *FEBS Lett.* 2018;592(17):2874-2883.

Vagnildhaug OM, Blum D, Wilcock A, Fayers P, Strasser F, Baracos VE, et al. The applicability of a weight loss grading system in cancer cachexia: a longitudinal analysis. *J Cachexia Sarcopenia Muscle.* 2017;8(5):789-797.

van de Worp W, Schols A, Dingemans AC, Op den Kamp CMH, Degens J, Kelders M, et al. Identification of microRNAs in skeletal muscle associated with lung cancer cachexia. *J Cachexia Sarcopenia Muscle.* 2020;11(2):452-463.

Vaughan VC, Martin P, Lewandowski PA. Cancer cachexia: impact, mechanisms and emerging treatments. *J Cachexia Sarcopenia Muscle.* 2013;4(2):95-109.

Vidman L, Kallberg D, Ryden P. Cluster analysis on high dimensional RNA-seq data with applications to cancer research - An evaluation study. *PLoS One.* 2019;14(12):e0219102.

von Haehling S, Anker MS, Anker SD. Prevalence and clinical impact of cachexia in chronic illness in Europe, USA, and Japan: facts and numbers update 2016. *J Cachexia Sarcopenia Muscle.* 2016;7(5):507-509.

von Haehling S, Anker SD. Prevalence, incidence and clinical impact of cachexia: facts and numbers-update 2014. *J Cachexia Sarcopenia Muscle.* 2014;5(4):261-263.

Wajahat M, Bracken CP, Orang A. Emerging Functions for snoRNAs and snoRNA-Derived Fragments. *Int J Mol Sci.* 2021;22(19).

Wang G, Kossenkov AV, Ochs MF. LS-NMF: a modified non-negative matrix factorization algorithm utilizing uncertainty estimates. *BMC Bioinformatics.* 2006;7:175.

Webster JM, Kempen L, Hardy RS, Langen RCJ. Inflammation and Skeletal Muscle Wasting During Cachexia. *Front Physiol.* 2020;11:597675.

Welle S, Tawil R, Thornton CA. Sex-related differences in gene expression in human skeletal muscle. *PLoS One.* 2008;3(1):e1385.

Williams A, Sun X, Fischer JE, Hasselgren PO. The expression of genes in the ubiquitin-proteasome proteolytic pathway is increased in skeletal muscle from patients with cancer. *Surgery.* 1999;126(4):744-749; discussion 749-750.

Xia L, Zhao R, Wan Q, Wu Y, Zhou Y, Wang Y, et al. Sarcopenia and adverse health-related outcomes: An umbrella review of meta-analyses of observational studies. *Cancer Med.* 2020;9(21):7964-7978.

Xie K, Xiong H, Xiao W, Xiong Z, Hu W, Ye J, et al. Downregulation of miR-29c promotes muscle wasting by modulating the activity of leukemia inhibitory factor in lung cancer cachexia. *Cancer Cell Int.* 2021;21(1):627.

Xu D, Yang F, Fan Y, Jing W, Wen J, Miao W, et al. LncRNA DLEU1 Contributes to the Growth and Invasion of Colorectal Cancer via Targeting miR-320b/PRPS1. *Front Oncol.* 2021;11:640276.

Yaffe D, Saxel O. Serial passaging and differentiation of myogenic cells isolated from dystrophic mouse muscle. *Nature.* 1977;270(5639):725-727.

Yang W, Hu P. Skeletal muscle regeneration is modulated by inflammation. *J Orthop Translat.* 2018;13:25-32.

Yao RW, Wang Y, Chen LL. Cellular functions of long noncoding RNAs. *Nat Cell Biol.* 2019;21(5):542-551.

Yao S. MicroRNA biogenesis and their functions in regulating stem cell potency and differentiation. *Biol Proced Online.* 2016;18:8.

Ye J, She X, Liu Z, He Z, Gao X, Lu L, et al. Eukaryotic Initiation Factor 4A-3: A Review of Its Physiological Role and Involvement in Oncogenesis. *Front Oncol.* 2021;11:712045.

Yin L, Cui J, Lin X, Li N, Fan Y, Zhang L, et al. Identifying cancer cachexia in patients without weight loss information: machine learning approaches to address a real-world challenge. *Am J Clin Nutr.* 2022.

Yoon HG, Oh D, Noh JM, Cho WK, Sun JM, Kim HK, et al. Machine learning model for predicting excessive muscle loss during neoadjuvant chemoradiotherapy in oesophageal cancer. *J Cachexia Sarcopenia Muscle.* 2021;12(5):1144-1152.

Zhang W, Liu Y, Zhang H. Extracellular matrix: an important regulator of cell functions and skeletal muscle development. *Cell Biosci.* 2021;11(1):65.

Zhang ZK, Li J, Guan D, Liang C, Zhuo Z, Liu J, et al. A newly identified lncRNA MAR1 acts as a miR-487b sponge to promote skeletal muscle differentiation and regeneration. *J Cachexia Sarcopenia Muscle.* 2018;9(3):613-626.

Zhang ZK, Li J, Guan D, Liang C, Zhuo Z, Liu J, et al. Long Noncoding RNA lncMUMA Reverses Established Skeletal Muscle Atrophy following Mechanical Unloading. *Mol Ther.* 2018;26(11):2669-2680.

Zhong F, Zhou N, Wu K, Guo Y, Tan W, Zhang H, et al. A SnoRNA-derived piRNA interacts with human interleukin-4 pre-mRNA and induces its decay in nuclear exosomes. *Nucleic Acids Res.* 2015;43(21):10474-10491.

Zhong X, Narasimhan A, Silverman LM, Young AR, Shahda S, Liu S, et al. Sex specificity of pancreatic cancer cachexia phenotypes, mechanisms, and treatment in mice and humans: role of Activin. *J Cachexia Sarcopenia Muscle.* 2022;13(4):2146-2161.

Zhong X, Zimmers TA. Sex Differences in Cancer Cachexia. *Curr Osteoporos Rep.* 2020;18(6):646-654.

Zhu M, Liu J, Xiao J, Yang L, Cai M, Shen H, et al. Lnc-mg is a long non-coding RNA that promotes myogenesis. *Nat Commun.* 2017;8:14718.

Appendix

Table A.1 Differentially Expressed (DE) mRNAs between subtypes in males

Top 200 (up- and down-regulated) DE mRNAs (Subtype 1 vs. Subtype 2) in males Genome assembly GRCh 38 Human Genome (hg) build 38							
mRNA	**Chr**	**Start**	**Stop**	**P-value**	**FDR**	**FC**	**log2 FC**
FSHB	11	30231014	30235262	3.12E-09	3.95E-07	14.8	3.9
OR14J1	6	29301701	29313018	8.55E-11	1.01E-07	14.3	3.8
DUOXA2	15	45114326	45118422	5.89E-08	1.97E-06	13.9	3.8
LRRC19	9	26993136	27005673	4.88E-09	5.03E-07	13.6	3.8
IL36A	2	113005459	113008045	9.78E-09	7.55E-07	13.3	3.7
ANKRD40CL	17	50761029	50767558	2.15E-09	3.28E-07	13.2	3.7
OR8D4	11	123902167	123909230	3.27E-11	6.67E-08	13.1	3.7
TMEM213	7	138797952	138838102	2.43E-10	1.38E-07	13.0	3.7
ZSWIM2	2	186827475	186849209	5.47E-10	1.94E-07	13.0	3.7
SPATS1	6	44342650	44380180	8.69E-09	6.99E-07	12.5	3.6
PGK2	6	49785660	49787286	8.09E-07	1.09E-05	12.1	3.6
SMIM31	4	164754064	164803796	1.76E-09	2.92E-07	12.0	3.6
TTLL8	22	50015123	50056936	9.62E-08	2.64E-06	11.2	3.5
BPIFB6	20	33031648	33044048	2.37E-08	1.19E-06	11.0	3.5
GSTA3	6	52896639	52909699	7.30E-08	2.23E-06	10.9	3.4
OR2J1	6	29099657	29102702	1.52E-06	1.68E-05	10.8	3.4
APOBEC1	12	7649400	7665909	9.03E-07	1.17E-05	10.7	3.4
OR4Q3	14	19743571	19749470	1.11E-08	7.97E-07	10.6	3.4
OR4K2	14	19875142	19883933	4.88E-10	1.85E-07	10.4	3.4
OR2T2	1	248445512	248455726	4.45E-06	3.69E-05	10.4	3.4
VSX2	14	74239449	74262739	1.59E-06	1.73E-05	10.3	3.4
NEUROD6	7	31337465	31340727	7.91E-07	1.08E-05	10.3	3.4
OR10H5	19	15787661	15801026	5.09E-10	1.85E-07	10.2	3.4
SUN5	20	32983773	33004434	6.71E-08	2.11E-06	10.2	3.3
IL36G	2	112973203	112985659	5.08E-08	1.81E-06	10.1	3.3
C2orf80	2	208165343	208190031	2.17E-08	1.14E-06	10.1	3.3
C17orf78	17	37375985	37392709	4.36E-09	4.65E-07	10.1	3.3
OR7G2	19	9100407	9107476	1.40E-07	3.40E-06	10.1	3.3
OR4D9	11	59511368	59520704	1.49E-10	1.21E-07	10.0	3.3
INSL6	9	5123880	5185648	1.31E-09	2.69E-07	10.0	3.3
LHX1	17	36936785	36944613	1.00E-07	2.70E-06	10.0	3.3
FRG2C	3	75664330	75667221	2.62E-09	3.66E-07	9.9	3.3
RFPL2	22	32190435	32205002	3.48E-08	1.47E-06	9.9	3.3
PIWIL1	12	130337887	130372638	2.22E-10	1.38E-07	9.8	3.3

MEIOB	16	1833986	1884295	7.61E-09	6.43E-07	9.8	3.3
MGAT4D	4	140442262	140498378	1.18E-09	2.54E-07	9.7	3.3
CACNG2	22	36560857	36703753	5.12E-10	1.85E-07	9.7	3.3
SPIC	12	101475336	101486998	1.32E-07	3.26E-06	9.7	3.3
PRAMEF2	1	12857086	12861910	5.04E-10	1.85E-07	9.7	3.3
MOG	6	29657002	29672373	6.92E-08	2.16E-06	9.6	3.3
EXD1	15	41182725	41230744	3.34E-10	1.58E-07	9.6	3.3
PADI6	1	17372196	17401700	5.28E-09	5.24E-07	9.6	3.3
SLC13A2	17	28473293	28497782	2.44E-07	4.87E-06	9.6	3.3
NAT16	7	101170496	101180294	2.78E-07	5.32E-06	9.6	3.3
CEACAM20	19	44501677	44529789	9.39E-10	2.40E-07	9.5	3.3
TRIM77	11	89710299	89717873	3.81E-09	4.38E-07	9.5	3.2
OR5K1	3	98463201	98472925	9.63E-09	7.54E-07	9.5	3.2
BEST2	19	12751702	12758459	4.00E-07	6.76E-06	9.5	3.2
KCNJ13	2	232765802	232776566	1.97E-10	1.35E-07	9.4	3.2
OR7A17	19	14878203	14886133	4.17E-10	1.73E-07	9.4	3.2
OTOP1	4	4188726	4226930	4.07E-08	1.61E-06	9.3	3.2
NPHS2	1	179550539	179575953	1.69E-06	1.81E-05	9.3	3.2
UGT3A2	5	36035021	36071359	8.69E-08	2.47E-06	9.3	3.2
SCGB3A2	5	147870682	147882192	3.72E-10	1.58E-07	9.3	3.2
LIPK	10	88724544	88752757	5.94E-08	1.98E-06	9.3	3.2
SLC34A1	5	177379235	177398849	7.84E-08	2.31E-06	9.3	3.2
OR4K1	14	19930917	19936758	3.13E-09	3.95E-07	9.3	3.2
TEX101	19	43401496	43418598	9.77E-07	1.24E-05	9.2	3.2
OR4N2	14	19719015	19830254	7.85E-10	2.39E-07	9.2	3.2
ASIC5	4	155829729	155866278	3.23E-08	1.43E-06	9.2	3.2
PLCZ1	12	18683169	18738101	1.94E-10	1.35E-07	9.2	3.2
OR4M1	14	19773504	19783697	1.37E-07	3.34E-06	9.2	3.2
AC013470.2	7	12469621	12542223	3.67E-12	3.43E-08	9.1	3.2
RNF212B	14	23185316	23273478	6.06E-07	8.91E-06	9.1	3.2
NANOGNB	12	7765216	7774122	1.35E-06	1.55E-05	9.1	3.2
DAOA	13	105465867	105491035	4.47E-08	1.70E-06	9.1	3.2
NLRP9	19	55708438	55738403	8.72E-08	2.48E-06	9.0	3.2
TMEM225	11	123882920	123885671	1.29E-06	1.50E-05	9.0	3.2
CD200R1L	3	112815709	112846857	3.33E-08	1.44E-06	8.9	3.2
MS4A12	11	60492778	60507431	2.53E-10	1.38E-07	8.9	3.2
BTNL3	5	180988846	181006728	1.54E-09	2.90E-07	8.9	3.2
NLRP11	19	55785397	55836801	1.54E-08	9.44E-07	8.9	3.1
OR5H1	3	98130721	98138549	6.86E-09	6.13E-07	8.8	3.1
BRINP3	1	190097658	190478405	3.55E-10	1.58E-07	8.8	3.1
PRR27	4	70133616	70176800	1.29E-09	2.69E-07	8.8	3.1

PRR23C	3	139042102	139044893	5.16E-06	4.14E-05	8.8	3.1
OR6N2	1	158774222	158781205	5.88E-08	1.97E-06	8.7	3.1
SERPINB13	18	63586989	63604640	5.31E-09	5.24E-07	8.7	3.1
TMIGD1	17	30316333	30334060	3.54E-06	3.11E-05	8.7	3.1
ELSPBP1	19	47994632	48025155	3.47E-07	6.10E-06	8.7	3.1
OR51A7	11	4903783	4909463	3.97E-07	6.73E-06	8.7	3.1
TSPAN16	19	11296139	11326997	8.41E-08	2.42E-06	8.7	3.1
SPANXB1	X	141002594	141003707	1.82E-05	1.12E-04	8.7	3.1
CCDC63	12	110846769	110907536	1.03E-08	7.63E-07	8.7	3.1
HMGB4	1	33860475	33864792	1.04E-05	7.17E-05	8.7	3.1
SELENOV	19	39515113	39520687	7.63E-07	1.05E-05	8.6	3.1
DNAJC22	12	49346888	49357547	3.79E-08	1.54E-06	8.6	3.1
TRIM51GP	11	48975324	48983827	8.45E-09	6.95E-07	8.6	3.1
NLRP7	19	54923509	54966313	2.49E-09	3.60E-07	8.6	3.1
ACSBG2	19	6135247	6193095	3.83E-08	1.55E-06	8.6	3.1
HTR3A	11	113974881	113990314	1.62E-06	1.76E-05	8.5	3.1
ARGFX	3	121567949	121590623	2.32E-08	1.19E-06	8.5	3.1
SULT2A1	19	47870467	47886316	1.10E-06	1.33E-05	8.5	3.1
SLC17A1	6	25782915	25832053	9.46E-11	1.03E-07	8.5	3.1
SLC9A4	2	102473226	102533973	9.84E-10	2.43E-07	8.5	3.1
ADAM2	8	39743735	39838290	4.54E-09	4.78E-07	8.5	3.1
SLC17A6	11	22338381	22379504	2.04E-09	3.16E-07	8.5	3.1
TSGA13	7	130668643	130687433	1.85E-06	1.93E-05	8.5	3.1
ANKRD20A3P	9	66106815	66179475	7.12E-08	2.18E-06	8.5	3.1
OR2V1	5	181123122	181131170	9.86E-07	1.25E-05	8.5	3.1
C1QC	1	22643633	22648111	1.65E-03	5.51E-03	-1.5	-0.6
DLK1	14	100725705	100738225	2.40E-03	7.70E-03	-1.5	-0.6
OAF	11	120211032	120230335	2.26E-04	9.39E-04	-1.5	-0.6
XG	X	2752040	2816501	6.00E-04	2.22E-03	-1.5	-0.6
DCLK1	13	35768652	36131383	4.30E-03	1.30E-02	-1.5	-0.6
PDGFRL	8	17576433	17644072	1.05E-07	2.76E-06	-1.5	-0.6
ECM2	9	92493554	92536656	1.52E-03	5.11E-03	-1.5	-0.6
EZR	6	158765741	158819369	1.67E-05	1.04E-04	-1.5	-0.6
SCN4B	11	118133377	118152889	3.17E-04	1.26E-03	-1.5	-0.6
SCARA3	8	27633868	27676777	2.24E-03	7.24E-03	-1.5	-0.6
LRRC56	11	537527	554913	1.56E-02	4.07E-02	-1.5	-0.6
PTPRF	1	43525187	43623667	1.97E-02	4.97E-02	-1.5	-0.6
GPC3	X	133535745	133985595	3.46E-03	1.07E-02	-1.5	-0.6
CYBB	X	37780059	37813462	6.55E-04	2.40E-03	-1.5	-0.6
THBS2	6	169215780	169254051	7.26E-04	2.64E-03	-1.5	-0.6
FBLN1	22	45502238	45601136	4.78E-05	2.49E-04	-1.5	-0.6

SLC1A4	2	64988477	65023866	1.85E-04	7.92E-04	-1.6	-0.6
FBN1	15	48408313	48645722	6.09E-05	3.05E-04	-1.6	-0.6
IGSF10	3	151425384	151458710	1.30E-02	3.45E-02	-1.6	-0.6
ABI3BP	3	100749156	100993516	4.90E-05	2.54E-04	-1.6	-0.6
COL6A2	21	46098112	46132849	1.77E-05	1.09E-04	-1.6	-0.6
OIP5	15	41309273	41332592	4.06E-04	1.57E-03	-1.6	-0.6
ARRDC2	19	18001132	18014103	4.85E-03	1.44E-02	-1.6	-0.7
FSTL1	3	120392293	120450994	9.26E-05	4.33E-04	-1.6	-0.7
COL6A1	21	45981769	46005051	6.00E-06	4.64E-05	-1.6	-0.7
MRC2	17	62627670	62693598	1.77E-04	7.63E-04	-1.6	-0.7
PCOLCE	7	100602363	100608176	8.13E-04	2.92E-03	-1.6	-0.7
OLFML3	1	113979391	114035573	7.11E-05	3.48E-04	-1.6	-0.7
PLXNB2	22	50274979	50307647	2.53E-05	1.46E-04	-1.6	-0.7
OMD	9	92414245	92424462	2.36E-03	7.58E-03	-1.6	-0.7
FN1	2	215360440	215436074	4.07E-04	1.57E-03	-1.6	-0.7
SMOC2	6	168441151	168673446	4.75E-05	2.48E-04	-1.6	-0.7
MFAP2	1	16974502	16980633	8.24E-03	2.31E-02	-1.6	-0.7
MGP	12	14880864	14885858	6.87E-04	2.51E-03	-1.6	-0.7
S100A4	1	153543613	153550137	2.48E-04	1.02E-03	-1.6	-0.7
PI16	6	36948263	36964838	1.94E-04	8.21E-04	-1.6	-0.7
MMP14	14	22836560	22849028	2.51E-04	1.03E-03	-1.6	-0.7
CD248	11	66314494	66317045	3.86E-04	1.50E-03	-1.6	-0.7
SPARC	5	151661096	151686976	3.18E-04	1.26E-03	-1.6	-0.7
HTRA1	10	122458551	122514908	3.01E-06	2.75E-05	-1.6	-0.7
FST	5	53480626	53487135	1.08E-02	2.94E-02	-1.6	-0.7
CHRDL1	X	110673856	110795820	2.51E-04	1.03E-03	-1.6	-0.7
SESN3	11	95165513	95232542	7.70E-04	2.78E-03	-1.6	-0.7
HS6ST2	X	132626016	132961396	4.14E-03	1.25E-02	-1.6	-0.7
SMAD9	13	36844831	36920766	6.76E-03	1.94E-02	-1.6	-0.7
SCPEP1	17	56978129	57006769	2.09E-05	1.25E-04	-1.7	-0.7
CAPN6	X	111245099	111270484	2.46E-03	7.87E-03	-1.7	-0.7
COL6A3	2	237324003	237414208	4.59E-06	3.78E-05	-1.7	-0.7
SSC5D	19	55488404	55519100	3.85E-05	2.08E-04	-1.7	-0.7
LAPTM5	1	30732469	30757775	2.14E-03	6.94E-03	-1.7	-0.8
FSCN1	7	5592816	5606656	2.92E-06	2.70E-05	-1.7	-0.8
FMOD	1	203340628	203351759	1.45E-02	3.81E-02	-1.7	-0.8
NGFR	17	49495293	49515009	1.05E-02	2.87E-02	-1.7	-0.8
COL5A2	2	189031898	189225313	5.34E-05	2.73E-04	-1.7	-0.8
MMP2	16	55389700	55506692	2.79E-06	2.61E-05	-1.7	-0.8
EDA2R	X	66595637	66639299	6.55E-03	1.88E-02	-1.7	-0.8
GXYLT2	3	72888046	72998139	1.32E-04	5.88E-04	-1.7	-0.8

BGN	X	153494980	153509547	6.94E-03	1.98E-02	-1.8	-0.8
COL14A1	8	120059780	120373574	1.10E-03	3.83E-03	-1.8	-0.8
HES6	2	238238267	238240663	2.83E-03	8.90E-03	-1.8	-0.8
PDZD4	X	153802166	153830566	1.55E-04	6.77E-04	-1.8	-0.8
MMRN1	4	89879532	89954630	1.29E-03	4.41E-03	-1.8	-0.8
HPDL	1	45326895	45328680	1.40E-02	3.69E-02	-1.8	-0.9
GPNMB	7	23235967	23275109	1.03E-05	7.12E-05	-1.8	-0.9
COL19A1	6	69866556	70212469	5.98E-03	1.74E-02	-1.9	-0.9
FRZB	2	182833275	182866638	2.80E-03	8.84E-03	-1.9	-0.9
FAM43B	1	20552573	20555021	5.30E-05	2.71E-04	-1.9	-0.9
THY1	11	119415476	119424986	5.34E-03	1.57E-02	-1.9	-0.9
LUM	12	91102629	91111495	4.20E-04	1.62E-03	-1.9	-0.9
POSTN	13	37562583	37598845	4.98E-03	1.48E-02	-1.9	-0.9
LOX	5	122063195	122078414	3.82E-04	1.49E-03	-1.9	-0.9
COL5A1	9	134641803	134844844	7.04E-05	3.45E-04	-1.9	-0.9
CEBPA	19	33299934	33302535	8.65E-05	4.10E-04	-1.9	-0.9
CCL21	9	34709005	34710137	4.75E-03	1.42E-02	-1.9	-0.9
FNDC1	6	159169400	159272109	2.24E-03	7.22E-03	-1.9	-1.0
THBS4	5	79991311	80083288	4.92E-03	1.46E-02	-2.0	-1.0
OLFML2B	1	161983192	162023855	1.23E-04	5.53E-04	-2.0	-1.0
ADAMTS2	5	179110853	179345462	3.07E-04	1.22E-03	-2.0	-1.0
TMEM119	12	108589851	108598321	2.90E-04	1.16E-03	-2.0	-1.0
CCDC80	3	112596797	112649531	3.18E-04	1.26E-03	-2.0	-1.0
H1-5	6	27866792	27867589	2.34E-03	7.51E-03	-2.0	-1.0
ELN	7	74027789	74069908	1.66E-04	7.19E-04	-2.1	-1.1
KAZALD1	10	101061989	101068132	7.90E-05	3.80E-04	-2.1	-1.1
ASPN	9	92456205	92482507	5.93E-04	2.20E-03	-2.1	-1.1
SCD	10	100347233	100364827	1.70E-02	4.37E-02	-2.1	-1.1
APOE	19	44905791	44909394	4.95E-05	2.56E-04	-2.2	-1.1
MYH8	17	10390322	10421951	9.78E-03	2.69E-02	-2.2	-1.1
LRRC15	3	194355249	194369744	5.38E-03	1.58E-02	-2.3	-1.2
CHAD	17	50464496	50468907	8.49E-05	4.03E-04	-2.4	-1.2
MXRA5	X	3308565	3346653	3.86E-04	1.50E-03	-2.5	-1.3
UTS2R	17	82371400	82375587	2.26E-03	7.30E-03	-2.6	-1.4
COL1A2	7	94394895	94431228	1.58E-05	9.95E-05	-2.6	-1.4
SCT	11	626309	627182	1.65E-04	7.15E-04	-2.8	-1.5
SFRP2	4	153780591	153789084	5.10E-04	1.93E-03	-2.8	-1.5
MYBPH	1	203167811	203175827	4.53E-04	1.73E-03	-3.0	-1.6
COL3A1	2	188974373	189012747	3.03E-05	1.69E-04	-3.0	-1.6
NPR3	5	32689070	32791725	8.48E-05	4.03E-04	-3.2	-1.7
SFRP4	7	37905932	38025696	4.61E-07	7.44E-06	-3.4	-1.8

SAMD11	1	923928	944582	6.37E-06	4.86E-05	-4.2	-2.1
COL1A1	17	50184101	50201633	7.19E-06	5.34E-05	-4.3	-2.1

N=48 male samples were subjected to the DE analysis. The table represents top 100 up-regulated and 100 down-regulated mRNAs between subtype 1vs. subtype 2 in males. Total n= 3519 mRNAs were found to be DE between the subtypes.

Table A.2 Differentially Expressed (DE) lncRNAs between subtypes in males

Top 200 (up- and down-regulated) DE lncRNAs (Subtype 1 vs. Subtype 2) in males Genome assembly GRCh 38 Human Genome (hg) build 38							
lncRNA	**Chr**	**Start**	**Stop**	**P-value**	**FDR**	**FC**	**log2FC**
AC087639.2	15	68880663	68928942	6.80E-10	2.17E-07	18.9	4.2
LINC02681	10	101252821	101263551	9.45E-09	7.52E-07	16.4	4.0
LINC01549	21	17438821	17450105	6.41E-09	5.89E-07	15.0	3.9
LINC02558	22	31970823	32037996	2.64E-09	3.66E-07	14.0	3.8
LINC02267	4	96310701	96818865	2.80E-10	1.46E-07	14.0	3.8
AC008060.1	7	155382076	155396357	5.51E-08	1.89E-06	13.8	3.8
AC100834.2	15	34657736	34670513	2.21E-08	1.15E-06	13.7	3.8
LINC02378	12	17509353	17590050	3.64E-10	1.58E-07	13.5	3.8
AL139090.1	6	88378280	88385448	5.32E-08	1.84E-06	12.9	3.7
AF279873.4	8	33973701	34009596	1.12E-07	2.90E-06	12.6	3.7
AC114760.2	2	196151263	196154882	1.64E-07	3.73E-06	12.6	3.7
AC107373.1	8	24992002	25009778	5.46E-08	1.88E-06	12.3	3.6
EFCAB6-AS1	22	43516107	43560064	1.27E-10	1.11E-07	12.1	3.6
AC134508.1	3	72321051	72324028	1.07E-08	7.76E-07	12.0	3.6
LINC02796	1	72765031	72791283	8.39E-12	3.95E-08	12.0	3.6
AC103957.1	8	8414204	8424649	4.44E-08	1.70E-06	12.0	3.6
AP003774.1	11	64420311	64432671	6.74E-08	2.12E-06	12.0	3.6
AC019155.2	7	125229579	125264292	1.03E-09	2.43E-07	11.9	3.6
AC023136.1	4	162022733	162047568	3.56E-08	1.49E-06	11.9	3.6
AL139393.2	6	160926269	160943111	1.36E-08	8.97E-07	11.5	3.5
SNCA-AS1	4	89836408	89841979	1.08E-08	7.84E-07	11.4	3.5
TLX1NB	10	101089321	101131127	7.67E-08	2.29E-06	11.3	3.5
AP004247.2	11	58044110	58060139	2.17E-08	1.14E-06	11.3	3.5
LINC02058	5	92907180	92939592	1.62E-07	3.69E-06	11.3	3.5
FAM230C	13	18195297	18232025	4.01E-14	9.84E-10	11.2	3.5
AC009019.1	16	22083256	22092486	1.26E-09	2.66E-07	11.1	3.5
AC018692.1	21	9913117	9914847	1.50E-08	9.30E-07	11.0	3.5
LINC02087	17	15806241	15847997	1.02E-08	7.63E-07	11.0	3.5
LINC00485	12	102809280	102824400	5.86E-09	5.59E-07	10.9	3.5
AC093765.2	4	116754453	116921793	3.64E-08	1.51E-06	10.9	3.4
AL732437.2	10	5563295	5566882	1.10E-07	2.85E-06	10.9	3.4
AC011287.2	7	13339201	13365050	7.95E-08	2.34E-06	10.9	3.4
AC016383.2	18	52283909	52309225	6.54E-07	9.39E-06	10.8	3.4
AC091885.2	5	24554018	24613223	3.46E-08	1.47E-06	10.8	3.4
AL360169.3	6	158296671	158303373	4.12E-08	1.63E-06	10.7	3.4
LINC00992	5	117415509	117546299	1.30E-08	8.75E-07	10.7	3.4

AC138123.2	12	93090480	93108630	1.51E-09	2.90E-07	10.6	3.4
DSCAM-AS1	21	40383083	40385359	3.89E-08	1.57E-06	10.6	3.4
AL355347.1	6	61870097	61893883	1.02E-10	1.03E-07	10.6	3.4
AC021134.2	4	162806370	162844691	3.68E-07	6.36E-06	10.5	3.4
LINC01514	10	101176306	101194148	1.71E-07	3.82E-06	10.5	3.4
AP001255.1	21	23106928	23119904	1.63E-06	1.77E-05	10.5	3.4
AC104257.1	8	131308545	131317633	1.49E-08	9.30E-07	10.5	3.4
AC003092.1	7	94022833	94066662	5.18E-07	8.03E-06	10.4	3.4
LINC02510	4	137193756	137198368	4.95E-06	4.01E-05	10.4	3.4
AL441943.2	10	2150480	2169461	7.12E-07	9.98E-06	10.4	3.4
AC125613.1	3	181534177	181700612	2.90E-10	1.48E-07	10.4	3.4
LINC02797	1	72793104	72854476	3.71E-08	1.52E-06	10.4	3.4
AL359987.1	6	92631066	92633269	6.75E-06	5.07E-05	10.4	3.4
LINC01525	1	117272182	117321337	9.51E-08	2.62E-06	10.3	3.4
AC018767.3	16	8309962	8357861	2.50E-05	1.44E-04	10.3	3.4
LINC01518	10	42644445	42691724	6.42E-09	5.89E-07	10.3	3.4
LINC00547	13	37534940	37551537	9.96E-08	2.70E-06	10.2	3.4
AL591543.1	9	38650215	38662718	4.15E-09	4.52E-07	10.2	3.4
LINC02221	5	9854371	9899971	4.98E-08	1.80E-06	10.2	3.3
LINC01818	2	150169474	150356461	3.60E-07	6.26E-06	10.1	3.3
AC090337.2	18	69240448	69250782	1.96E-10	1.35E-07	10.1	3.3
AC104623.1	2	16523176	16555565	1.13E-07	2.91E-06	10.1	3.3
AC113348.3	5	178690804	178693682	1.96E-06	2.02E-05	10.1	3.3
LINC02134	16	47965620	47971694	2.36E-06	2.30E-05	10.0	3.3
AC079154.1	2	124011984	124025174	2.10E-10	1.36E-07	10.0	3.3
AC093426.1	1	239247808	239250819	3.95E-08	1.58E-06	10.0	3.3
LINC01966	2	144877734	144882034	1.93E-06	1.99E-05	10.0	3.3
AC109441.1	5	144439204	144468472	1.48E-08	9.28E-07	10.0	3.3
MARCHF10-DT	17	62808500	62836372	1.29E-07	3.19E-06	9.9	3.3
LINC01441	20	55420336	55427198	2.74E-07	5.26E-06	9.9	3.3
TTC21B-AS1	2	165933857	165949892	2.99E-08	1.35E-06	9.9	3.3
AL353730.1	9	25579657	25584273	6.89E-06	5.16E-05	9.8	3.3
LINC01179	4	165684639	165762779	1.08E-08	7.84E-07	9.8	3.3
AL450307.1	10	131776386	131795278	5.31E-07	8.16E-06	9.8	3.3
AL445985.2	13	23954493	23970189	2.44E-06	2.36E-05	9.8	3.3
AC097654.1	4	32005076	32022484	8.11E-10	2.39E-07	9.8	3.3
FABP6-AS1	5	160195744	160204827	1.19E-07	3.02E-06	9.8	3.3
AP001120.1	18	11666456	11670166	3.50E-07	6.13E-06	9.8	3.3
AC016245.1	17	80801640	80805633	8.67E-08	2.47E-06	9.8	3.3
LINC02487	6	167679626	167696291	2.54E-08	1.24E-06	9.8	3.3
AL353784.1	10	54486230	54656052	3.84E-09	4.38E-07	9.7	3.3

AL137018.1	9	74630794	74641759	2.99E-09	3.90E-07	9.7	3.3
LINC02529	6	159383422	159401148	1.19E-07	3.02E-06	9.7	3.3
AC116611.1	4	32583855	32745526	1.51E-09	2.90E-07	9.7	3.3
AC093515.1	16	7878799	8112757	8.36E-08	2.41E-06	9.7	3.3
AC025180.1	5	52930606	52990279	4.17E-08	1.64E-06	9.7	3.3
LINC00381	13	74419158	74444736	7.08E-09	6.13E-07	9.7	3.3
LMNB1-DT	5	126751963	126776487	2.73E-06	2.56E-05	9.7	3.3
AC105114.2	18	25196663	25218242	1.27E-09	2.67E-07	9.6	3.3
AC019330.1	2	198493242	198772357	2.03E-10	1.35E-07	9.6	3.3
AL356295.1	13	53207831	53801490	4.54E-08	1.71E-06	9.6	3.3
AC068189.1	8	95071732	95087925	4.30E-05	2.28E-04	9.6	3.3
AL139280.3	10	3270090	3277650	1.48E-08	9.28E-07	9.6	3.3
LINC01964	2	85061213	85067348	3.77E-08	1.54E-06	9.6	3.3
AC022201.1	2	70402934	70422679	9.04E-07	1.17E-05	9.6	3.3
FAM245A	10	87396561	87435036	3.15E-07	5.75E-06	9.6	3.3
AC008966.3	5	53024924	53035885	1.04E-07	2.76E-06	9.5	3.3
AC026462.4	16	52622115	52677748	1.23E-10	1.11E-07	9.5	3.2
LINC00838	10	33759713	33772703	5.12E-10	1.85E-07	9.5	3.2
LINC01120	2	131402778	131410051	3.33E-09	4.03E-07	9.5	3.2
AC011284.1	7	3642815	3643785	1.72E-08	1.00E-06	9.5	3.2
LINC00911	14	85393828	85420551	4.85E-09	5.03E-07	9.5	3.2
LINC01548	21	33164809	33170650	8.01E-08	2.35E-06	9.4	3.2
AP000439.2	11	69425690	69429622	4.58E-05	2.40E-04	9.4	3.2
AC092115.3	16	69727013	69742564	1.79E-02	4.58E-02	1.6	0.7
LINC01348	1	235065479	235074221	1.40E-02	3.70E-02	1.6	0.7
AC073130.1	7	116275606	116286735	1.87E-02	4.74E-02	1.6	0.7
AC023481.1	3	8359344	8366428	4.05E-06	3.41E-05	1.6	0.7
AC044893.1	8	49086301	49229175	5.22E-03	1.54E-02	1.6	0.7
LINC01801	19	34788527	34832870	6.89E-03	1.97E-02	1.6	0.7
AL161756.1	14	64329431	64338640	1.86E-02	4.72E-02	1.6	0.7
AC010332.3	19	52453580	52458837	1.30E-02	3.46E-02	1.6	0.7
KCNMB2-AS1	3	178526505	178937353	3.16E-03	9.85E-03	1.6	0.7
AC090260.1	15	78141243	78143174	1.79E-02	4.57E-02	1.6	0.7
PRC1-AS1	15	90966340	90988626	6.69E-07	9.53E-06	1.6	0.7
AC015712.2	15	100892343	100919392	3.50E-03	1.08E-02	1.6	0.6
AC000403.1	13	76887551	76891136	5.24E-03	1.55E-02	1.6	0.6
ZNF503-AS1	10	75243568	75373501	1.68E-02	4.32E-02	1.6	0.6
AC131934.1	11	665910	678392	7.44E-03	2.11E-02	1.6	0.6
AL110114.1	20	1023874	1118468	4.20E-03	1.27E-02	1.6	0.6
AC092802.1	1	95161676	95233983	1.71E-03	5.67E-03	1.6	0.6
CCDC28A-AS1	6	138725211	138773653	5.64E-03	1.65E-02	1.6	0.6

AP001122.1	11	128614340	128686923	5.44E-04	2.04E-03	1.6	0.6
AL034550.1	20	32449755	32453608	3.27E-03	1.02E-02	1.6	0.6
AC012443.2	2	152635268	152641276	8.98E-04	3.19E-03	1.6	0.6
HOXA-AS2	7	27107777	27134303	2.01E-04	8.49E-04	1.6	0.6
AC004263.2	12	120218070	120222669	1.87E-02	4.75E-02	1.6	0.6
AC008679.1	5	129150677	129173258	1.00E-02	2.75E-02	1.6	0.6
AL359764.3	1	241453751	241484996	2.33E-04	9.63E-04	1.6	0.6
AF131215.4	8	11123381	11126065	1.03E-03	3.62E-03	1.6	0.6
HCG25	6	33249534	33255168	1.89E-03	6.20E-03	1.6	0.6
SEC24B-AS1	4	109347475	109433818	1.09E-03	3.82E-03	1.6	0.6
LINC00484	9	91118592	91182763	7.14E-04	2.60E-03	1.6	0.6
AL445250.1	6	57961438	58438365	7.19E-03	2.05E-02	1.6	0.6
AC008496.3	5	79840797	79845240	1.35E-02	3.57E-02	1.6	0.6
AC026469.1	16	85142696	85146002	1.81E-02	4.62E-02	1.6	0.6
EPS15-AS1	1	51518309	51561630	6.90E-03	1.97E-02	1.6	0.6
AP000845.1	18	268148	270279	4.78E-03	1.43E-02	1.6	0.6
AC109587.1	3	69013941	69056623	1.99E-04	8.42E-04	1.6	0.6
AC090518.1	15	56542952	56629593	1.01E-03	3.55E-03	1.6	0.6
APCDD1L-DT	20	58515379	58619889	4.27E-03	1.29E-02	1.5	0.6
AC017050.1	2	33599442	33657461	4.97E-04	1.88E-03	1.5	0.6
SPART-AS1	13	36346431	36369602	3.47E-03	1.07E-02	1.5	0.6
AC139887.5	4	843963	846295	1.92E-02	4.87E-02	1.5	0.6
AL137026.1	10	44282489	44293999	5.86E-03	1.71E-02	1.5	0.6
TMSB15B-AS1	X	103845151	103919549	6.67E-03	1.91E-02	1.5	0.6
EAF1-AS1	3	15436171	15455941	7.81E-03	2.20E-02	1.5	0.6
CPEB2-DT	4	14909961	15002046	9.81E-05	4.55E-04	1.5	0.6
AL117336.1	10	35210416	35210751	4.23E-03	1.28E-02	1.5	0.6
ACOXL-AS1	2	111098345	111116408	1.18E-02	3.17E-02	1.5	0.6
AL133482.1	10	113482069	113492055	1.30E-02	3.46E-02	1.5	0.6
AC002467.1	7	107739999	107744582	2.63E-03	8.34E-03	1.5	0.6
LINC01460	2	27705786	27715733	7.19E-03	2.05E-02	1.5	0.6
AC073530.1	12	66950754	67096411	2.54E-03	8.10E-03	1.5	0.6
WDR11-AS1	10	120759898	120851458	5.29E-05	2.71E-04	1.5	0.6
AC004241.1	12	47706083	47742295	9.46E-04	3.35E-03	1.5	0.6
KCTD21-AS1	11	78139771	78175324	1.74E-03	5.77E-03	1.5	0.6
AC003086.1	7	91880791	91886491	9.44E-03	2.61E-02	1.5	0.6
AC040160.1	16	67261108	67263785	1.60E-06	1.74E-05	1.5	0.6
LINC02553	11	97222644	97259988	1.20E-02	3.23E-02	1.5	0.6
AC004943.2	16	72665123	72822782	4.73E-05	2.47E-04	1.5	0.6
AL133257.1	6	123589711	123685324	2.44E-03	7.82E-03	1.5	0.6
AL078604.2	6	159586955	159589170	1.93E-04	8.19E-04	1.5	0.6

AL022069.3	6	166342632	166351467	8.28E-03	2.32E-02	1.5	0.6
AL359710.1	9	99396987	99819922	1.07E-03	3.73E-03	1.5	0.6
LINC02646	10	130213488	130483194	6.26E-03	1.81E-02	1.5	0.6
LINC02615	4	128292751	128519399	5.09E-04	1.92E-03	1.5	0.6
AL008628.1	6	170414139	170419326	5.24E-04	1.97E-03	1.5	0.6
AL596442.4	6	170135374	170142927	1.39E-02	3.68E-02	1.5	0.6
AP001029.2	18	12438890	12448206	1.89E-04	8.04E-04	1.5	0.6
AFDN-DT	6	167822103	167826813	1.86E-02	4.72E-02	1.5	0.6
LIX1-AS1	5	97089075	97437218	7.06E-03	2.01E-02	1.5	0.6
AC017083.1	2	68252870	68253849	6.23E-04	2.30E-03	1.5	0.6
AC130814.1	2	18939697	18946651	1.48E-02	3.88E-02	1.5	0.6
AL390198.1	20	25697003	25752777	6.52E-03	1.88E-02	1.5	0.6
ALDH1L1-AS2	3	126180012	126210170	5.35E-03	1.58E-02	1.5	0.6
AL133215.1	10	100980507	100985615	2.52E-03	8.05E-03	1.5	0.6
AC025165.1	12	57612118	57619639	8.07E-03	2.27E-02	1.5	0.6
LINC00924	15	95326528	95507866	1.04E-02	2.83E-02	1.5	0.6
AC109927.1	4	138819954	139012647	8.94E-03	2.48E-02	1.5	0.6
AC105285.1	4	173094868	173169653	4.02E-03	1.22E-02	1.5	0.6
AC008771.1	5	80411231	80488096	1.18E-03	4.08E-03	1.5	0.6
ERVK9-11	19	38935297	38938633	8.15E-03	2.29E-02	1.5	0.6
AC098484.4	1	42658687	42682161	5.45E-03	1.60E-02	1.5	0.6
AL359715.1	6	80441295	80466675	7.74E-04	2.79E-03	1.5	0.6
AC026401.2	16	15683290	15684571	7.11E-03	2.03E-02	1.5	0.6
AC073655.2	12	94277758	94282845	1.72E-03	5.72E-03	1.5	0.6
FLG-AS1	1	152168125	152445457	1.16E-02	3.14E-02	1.5	0.6
AC024060.2	3	3152942	3153436	2.33E-03	7.50E-03	1.5	0.6
MIR223HG	X	66015461	66020423	3.94E-03	1.20E-02	-1.5	-0.6
AC004584.3	17	56982749	56985105	1.31E-02	3.49E-02	-1.5	-0.6
WAKMAR2	6	137823673	137868234	9.91E-03	2.72E-02	-1.6	-0.7
AL359397.2	14	50326526	50327910	1.65E-03	5.51E-03	-1.6	-0.7
AC069224.1	3	160753428	160755143	2.54E-03	8.08E-03	-1.6	-0.7
AC020909.2	19	50486810	50487639	1.82E-03	5.99E-03	-1.6	-0.7
MDC1-AS1	6	30703067	30713185	8.73E-03	2.43E-02	-1.7	-0.7
GAS6-AS1	13	113815630	113845745	2.67E-03	8.47E-03	-1.7	-0.8
AC021016.2	2	218398743	218399220	9.92E-05	4.59E-04	-1.7	-0.8
AC254633.1	1	12618900	12619245	2.73E-03	8.62E-03	-1.7	-0.8
AL158195.1	13	50372788	50387159	7.16E-03	2.04E-02	-1.8	-0.8
AC011239.1	2	23507043	23524345	3.19E-05	1.77E-04	-2.6	-1.4
MIR503HG	X	134543119	134546643	1.06E-03	3.72E-03	-2.7	-1.4
LINC02593	1	916865	921017	1.88E-03	6.19E-03	-2.8	-1.5

N=48 male samples were subjected to the DE analysis. The table represents top 100 up-regulated and 100 down-regulated lncRNAs between subtype 1vs. subtype 2 in males. Total n= 4232 lncRNAs were found to be DE between the subtypes.

Table A.3 Differentially Expressed miRNAs between subtypes in males

DE miRNAs (Subtype 1 vs. Subtype 2) in Males Genome assembly GRCh 38 Human Genome (hg) build 38							
miRNA	**Chr**	**Start**	**Stop**	**P-value**	**FDR**	**FC**	**log2 FC**
hsa-miR-296-5p	20	58817661	58817682	4.42E-11	3.50E-09	3.1	1.6
hsa-miR-4485-3p	11	10508277	10508297	7.19E-04	2.64E-03	2.6	1.4
hsa-miR-2110	10	114174151	114174173	2.14E-09	9.42E-08	2.6	1.4
hsa-miR-885-3p	3	10394499	10394521	2.83E-04	1.15E-03	2.4	1.3
hsa-miR-501-3p	X	50009772	50009794	1.83E-06	1.83E-05	2.2	1.1
hsa-miR-320b	1	116671787	116671809	4.33E-06	3.91E-05	2.1	1.0
hsa-miR-433-3p	14	100881949	100881971	6.65E-04	2.48E-03	2.0	1.0
hsa-miR-139-3p	11	72615066	72615089	1.35E-04	6.21E-04	2.0	1.0
hsa-miR-339-5p	7	1022990	1023013	5.42E-05	3.31E-04	1.9	1.0
hsa-miR-532-3p	X	50003204	50003226	5.72E-07	6.99E-06	1.9	0.9
hsa-miR-193a-5p	17	31560016	31560038	5.80E-06	4.56E-05	1.9	0.9
hsa-miR-486-3p	8	41660443	41660464	2.62E-04	1.11E-03	1.8	0.9
hsa-miR-320a-3p	8	22244975	22244997	4.84E-06	4.10E-05	1.8	0.9
hsa-miR-92a-3p	13	91351361	91351383	1.32E-04	6.16E-04	1.8	0.8
hsa-miR-92b-3p	1	155195237	155195259	3.98E-04	1.56E-03	1.8	0.8
hsa-miR-193b-5p	16	14303980	14304002	9.02E-05	4.73E-04	1.8	0.8
hsa-miR-23b-5p	9	95085227	95085249	3.38E-04	1.35E-03	1.7	0.8
hsa-miR-664a-3p	1	220200549	220200572	7.98E-05	4.36E-04	1.7	0.8
hsa-miR-145-5p	5	149430661	149430684	9.57E-05	4.80E-04	1.7	0.7
hsa-miR-423-5p	17	30117095	30117118	8.12E-05	4.36E-04	1.7	0.7
hsa-miR-361-3p	X	85903641	85903664	1.97E-09	9.42E-08	1.6	0.7
hsa-miR-125a-5p	19	51693268	51693292	2.53E-03	7.53E-03	1.6	0.7
hsa-miR-486-5p	8	41660484	41660506	6.18E-03	1.67E-02	1.5	0.6
hsa-miR-197-3p	1	109598940	109598962	1.55E-02	3.63E-02	1.5	0.6
hsa-miR-125b-5p	11	122099809	122099831	1.98E-04	8.79E-04	1.5	0.6
hsa-miR-199a-5p	1	172144592	172144615	2.26E-03	7.00E-03	-1.5	-0.6
hsa-miR-22-5p	17	1713952	1713974	2.00E-04	8.79E-04	-1.6	-0.6
hsa-miR-148a-5p	7	25949960	25949982	1.80E-02	3.96E-02	-1.6	-0.6
hsa-miR-152-3p	17	48037174	48037195	4.62E-03	1.29E-02	-1.6	-0.7
hsa-miR-10a-5p	17	48579904	48579927	9.59E-05	4.80E-04	-1.6	-0.7
hsa-miR-22-3p	17	1713914	1713936	9.58E-06	6.80E-05	-1.6	-0.7
hsa-miR-363-3p	X	134169382	134169404	8.49E-03	2.25E-02	-1.6	-0.7
hsa-miR-145-3p	5	149430699	149430721	7.85E-04	2.78E-03	-1.7	-0.7
hsa-miR-27a-3p	19	13836447	13836468	5.27E-08	8.28E-07	-1.7	-0.8
hsa-miR-143-3p	5	149428978	149428999	8.95E-04	3.08E-03	-1.7	-0.8
hsa-miR-182-5p	7	129770447	129770471	2.37E-03	7.23E-03	-1.8	-0.8

hsa-miR-30a-5p	6	71403595	71403617	6.55E-05	3.74E-04	-1.8	-0.8
hsa-miR-151a-3p	8	140732587	140732608	2.87E-08	5.75E-07	-1.8	-0.8
hsa-miR-30e-5p	1	40754371	40754393	5.11E-05	3.21E-04	-1.8	-0.8
hsa-miR-196b-5p	7	27169528	27169550	8.90E-04	3.08E-03	-1.8	-0.9
hsa-miR-194-5p	1	220118206	220118228	1.55E-03	5.00E-03	-1.8	-0.9
hsa-miR-99a-5p	21	16539101	16539123	4.19E-08	7.53E-07	-1.9	-0.9
hsa-miR-335-3p	7	130496162	130496184	1.60E-02	3.64E-02	-1.9	-0.9
hsa-miR-26b-5p	2	218402657	218402678	4.44E-06	3.91E-05	-1.9	-0.9
hsa-miR-132-5p	17	2049965	2049987	3.46E-03	9.88E-03	-1.9	-0.9
hsa-miR-155-5p	21	25573983	25574007	1.19E-04	5.71E-04	-2.0	-1.0
hsa-miR-192-5p	11	64891203	64891224	1.82E-03	5.71E-03	-2.0	-1.0
hsa-miR-379-5p	14	101022071	101022092	6.04E-03	1.66E-02	-2.0	-1.0
hsa-miR-10b-5p	2	176150329	176150352	1.05E-06	1.15E-05	-2.0	-1.0
hsa-miR-16-5p	3	160404754	160404776	6.12E-06	4.65E-05	-2.0	-1.0
hsa-let-7g-5p	3	52268336	52268358	1.69E-07	2.32E-06	-2.0	-1.0
hsa-miR-146a-5p	5	160485372	160485394	6.62E-05	3.74E-04	-2.0	-1.0
hsa-miR-499a-5p	20	34990408	34990429	2.49E-04	1.08E-03	-2.0	-1.0
hsa-miR-96-5p	7	129774739	129774762	1.30E-03	4.33E-03	-2.1	-1.1
hsa-let-7i-5p	12	62603691	62603713	2.51E-11	3.50E-09	-2.3	-1.2
hsa-miR-493-3p	14	100869116	100869138	7.24E-06	5.31E-05	-2.3	-1.2
hsa-let-7f-5p	9	94176353	94176375	4.45E-08	7.53E-07	-2.3	-1.2
hsa-miR-140-5p	16	69933103	69933125	1.01E-05	6.95E-05	-2.3	-1.2
hsa-miR-148a-3p	7	25949922	25949944	1.80E-06	1.83E-05	-2.3	-1.2
hsa-miR-98-5p	X	53556299	53556321	8.13E-09	2.56E-07	-2.4	-1.2
hsa-miR-370-3p	14	100911186	100911208	4.78E-11	3.50E-09	-2.5	-1.3
hsa-miR-146b-5p	10	102436520	102436543	1.17E-05	7.81E-05	-2.5	-1.3
hsa-miR-214-5p	1	172138857	172138879	9.70E-09	2.67E-07	-2.5	-1.3
hsa-miR-199b-5p	9	128244783	128244806	1.19E-08	2.90E-07	-2.5	-1.3
hsa-miR-15b-3p	3	160404645	160404667	6.62E-04	2.48E-03	-2.6	-1.4
hsa-miR-126-3p	9	136670653	136670675	1.45E-07	2.12E-06	-2.6	-1.4
hsa-miR-195-5p	17	7017667	7017688	5.51E-06	4.49E-05	-2.6	-1.4
hsa-miR-374b-5p	X	74218587	74218609	2.81E-07	3.63E-06	-2.6	-1.4
hsa-miR-21-5p	17	59841273	59841295	2.68E-08	5.75E-07	-2.7	-1.4
hsa-miR-451a	17	28861403	28861425	1.12E-04	5.49E-04	-2.7	-1.4
hsa-miR-495-3p	14	101033804	101033826	6.34E-07	7.34E-06	-2.7	-1.4
hsa-miR-101-3p	1	65058442	65058463	4.11E-09	1.51E-07	-3.3	-1.7

N=48 male samples were subjected to the DE analysis. The table represents n=72 DE miRNAs between subtype 1 vs. subtype 2 in males.

Table A.4 Differentially Expressed piRNAs between subtypes in males

DE piRNAs (Subtype 1 vs. Subtype 2) in Males Genome assembly GRCh 38 Human Genome (hg) build 38							
piRNA	**Chr**	**Start**	**Stop**	**P-value**	**FDR**	**FC**	**log2 FC**
hsa-piR-33031_752301	12	34205795	34205821	8.06E-20	1.10E-17	5.9	2.6
hsa-piR-23127_440092	6	4428054	4428084	2.59E-20	8.61E-18	5.9	2.5
hsa-piR-33031_752321	12	38161464	38161490	9.95E-20	1.10E-17	5.5	2.5
hsa-piR-33031_752312	1	228614748	228614774	1.95E-19	1.62E-17	5.4	2.4
hsa-piR-33031_752307	1	228634869	228634895	9.55E-18	3.17E-16	5.0	2.3
hsa-piR-33031_752305	1	228623665	228623691	1.44E-18	9.59E-17	4.8	2.3
hsa-piR-33031_752319	1	228641566	228641592	1.06E-16	2.70E-15	4.7	2.2
hsa-piR-33031_752309	1	228612507	228612533	4.77E-18	1.98E-16	4.7	2.2
hsa-piR-33031_752304	1	228621445	228621471	5.36E-17	1.48E-15	4.6	2.2
hsa-piR-33031_752318	1	228643807	228643833	2.98E-18	1.65E-16	4.6	2.2
hsa-piR-33031_752315	1	228637094	228637120	7.70E-18	2.84E-16	4.5	2.2
hsa-piR-33031_752313	1	228625907	228625933	4.11E-18	1.95E-16	4.4	2.1
hsa-piR-33031_752302	X	28982850	28982876	1.17E-17	3.53E-16	4.4	2.1
hsa-piR-32913_735926	X	69672573	69672600	6.33E-16	1.31E-14	4.1	2.0
hsa-piR-33031_752316	1	228632629	228632655	1.99E-15	3.30E-14	3.9	2.0
hsa-piR-33031_752310	1	228630388	228630414	1.71E-16	3.78E-15	3.9	2.0
hsa-piR-32913_735925	12	38161465	38161492	1.32E-11	1.62E-10	3.9	1.9
hsa-piR-33031_752317	1	228639335	228639361	1.25E-16	2.98E-15	3.7	1.9
hsa-piR-33031_752308	1	228619230	228619256	7.92E-16	1.55E-14	3.7	1.9
hsa-piR-32913_735927	19	21113127	21113154	1.37E-15	2.40E-14	3.6	1.9
hsa-piR-33031_752303	1	228558284	228558310	1.12E-15	2.07E-14	3.6	1.8
hsa-piR-33031_752314	1	228646038	228646064	7.25E-15	1.09E-13	3.6	1.8
hsa-piR-33031_752311	1	228610266	228610292	2.51E-15	3.97E-14	3.6	1.8
hsa-piR-32913_735913	1	228619231	228619258	1.00E-08	7.22E-08	3.6	1.8
hsa-piR-33031_752320	1	228628146	228628172	1.35E-14	1.95E-13	3.5	1.8
hsa-piR-32913_735921	1	228639336	228639363	3.66E-09	3.04E-08	3.4	1.8
hsa-piR-32913_735914	1	228630389	228630416	4.96E-08	3.11E-07	3.3	1.7
hsa-piR-33031_752306	1	228616989	228617015	2.11E-14	2.80E-13	3.3	1.7
hsa-piR-32913_735924	1	228628147	228628174	4.00E-10	4.28E-09	3.2	1.7
hsa-piR-32913_735923	1	228641567	228641594	8.39E-08	5.16E-07	3.2	1.7
hsa-piR-32913_735917	1	228625908	228625935	4.33E-09	3.34E-08	3.2	1.7
hsa-piR-32913_735910	1	228623666	228623693	1.08E-07	6.30E-07	3.2	1.7
hsa-piR-32913_735911	1	228616990	228617017	2.75E-07	1.52E-06	3.0	1.6
hsa-piR-32913_735907	7	140386780	140386807	1.53E-14	2.12E-13	2.9	1.6
hsa-piR-32913_735915	1	228610267	228610294	4.53E-08	2.95E-07	2.9	1.5

hsa-piR-32913_735908	1	228558285	228558312	2.06E-06	9.77E-06	2.8	1.5
hsa-piR-32913_735922	1	228643808	228643835	5.20E-07	2.66E-06	2.7	1.5
hsa-piR-32913_735920	1	228632630	228632657	1.19E-08	8.07E-08	2.7	1.5
hsa-piR-32913_735905	1	228612508	228612535	1.28E-07	7.31E-07	2.7	1.4
hsa-piR-32913_735912	1	228634870	228634897	2.82E-05	1.08E-04	2.6	1.4
hsa-piR-32913_735906	X	28982851	28982878	2.10E-06	9.83E-06	2.6	1.4
hsa-piR-32913_735919	1	228637095	228637122	7.31E-07	3.68E-06	2.5	1.3
hsa-piR-32913_735909	1	228621446	228621473	1.94E-06	9.35E-06	2.5	1.3
hsa-piR-32913_735918	1	228646039	228646066	1.22E-06	6.02E-06	2.3	1.2
hsa-piR-32919_735943	3	186784832	186784861	3.71E-04	1.18E-03	2.3	1.2
hsa-piR-33112_752509	4	7582605	7582639	4.44E-04	1.39E-03	2.3	1.2
hsa-piR-32913_735916	1	228614749	228614776	1.51E-05	6.21E-05	2.2	1.1
hsa-piR-32976_744093	M	2552	2578	2.45E-03	6.95E-03	2.1	1.1
hsa-piR-32941_743234	1	227561275	227561303	1.47E-09	1.39E-08	2.1	1.1
hsa-piR-33171_752575	19	12703635	12703669	8.93E-04	2.72E-03	2.1	1.1
hsa-piR-28172_625444	19	12706490	12706518	1.16E-03	3.44E-03	2.1	1.0
hsa-piR-33055_752427	6	85677587	85677622	2.84E-05	1.08E-04	1.9	0.9
hsa-piR-33090_752480	1	173864368	173864403	5.50E-03	1.39E-02	1.8	0.9
hsa-piR-32845_734843	8	123044780	123044810	8.81E-03	2.13E-02	1.7	0.8
hsa-piR-33161_752563	1	173864413	173864447	1.65E-02	3.83E-02	1.5	0.6
hsa-piR-32952_743942	21	8400810	8400838	3.23E-04	1.06E-03	-1.5	-0.6
hsa-piR-28085_601020	11	77886523	77886552	5.69E-03	1.43E-02	-1.5	-0.6
hsa-piR-32914_735935	21	8438155	8438180	1.53E-02	3.59E-02	-1.5	-0.6
hsa-piR-28255_625563	21	8218531	8218556	3.30E-03	8.87E-03	-1.6	-0.6
hsa-piR-32914_735932	11	45826690	45826715	2.08E-02	4.70E-02	-1.6	-0.6
hsa-piR-28085_601015	7	69062492	69062521	5.04E-03	1.30E-02	-1.6	-0.6
hsa-piR-1014_303	21	8398881	8398911	2.48E-04	8.48E-04	-1.6	-0.6
hsa-piR-1014_305	21	8443110	8443140	1.28E-03	3.77E-03	-1.6	-0.7
hsa-piR-1014_304	21	8260076	8260106	2.26E-04	7.81E-04	-1.6	-0.7
hsa-piR-32835_734800	1	198859106	198859131	6.09E-03	1.52E-02	-1.6	-0.7
hsa-piR-1014_302	21	8215847	8215877	1.58E-05	6.41E-05	-1.6	-0.7
hsa-piR-28247_625544	12	124921797	124921827	7.00E-03	1.73E-02	-1.7	-0.7
hsa-piR-28255_625564	21	8445801	8445826	9.99E-04	3.02E-03	-1.7	-0.7
hsa-piR-1911_208229	7	48632773	48632803	7.43E-04	2.31E-03	-1.7	-0.8
hsa-piR-3440_757356	21	8218493	8218518	1.09E-02	2.63E-02	-1.7	-0.8
hsa-piR-3440_757357	21	8445763	8445788	1.75E-02	4.01E-02	-1.7	-0.8
hsa-piR-14992_62869	21	8398762	8398791	2.61E-03	7.27E-03	-1.7	-0.8
hsa-piR-1911_208228	21	8401537	8401567	2.61E-04	8.84E-04	-1.7	-0.8
hsa-piR-32837_734802	12	20551529	20551555	2.53E-05	9.90E-05	-1.7	-0.8
hsa-piR-33041_752376	7	148987140	148987165	1.13E-02	2.70E-02	-1.8	-0.8
hsa-piR-28085_601016	21	8217623	8217652	3.67E-04	1.18E-03	-1.8	-0.8

hsa-piR-32837_734804	21	8438258	8438284	3.40E-06	1.50E-05	-1.8	-0.8
hsa-piR-32896_735112	16	47504906	47504931	4.97E-03	1.29E-02	-1.8	-0.8
hsa-piR-1911_208227	21	8445769	8445799	4.11E-04	1.30E-03	-1.8	-0.8
hsa-piR-3440_757359	X	89480866	89480891	4.35E-03	1.14E-02	-1.8	-0.9
hsa-piR-14992_62868	21	8215728	8215757	2.86E-03	7.84E-03	-1.8	-0.9
hsa-piR-1911_208226	21	8218499	8218529	1.50E-04	5.26E-04	-1.8	-0.9
hsa-piR-33041_752353	1	10999866	10999891	7.55E-03	1.86E-02	-1.9	-0.9
hsa-piR-32837_734805	21	8394048	8394074	3.33E-07	1.78E-06	-1.9	-0.9
hsa-piR-3440_757355	7	48632767	48632792	7.98E-03	1.95E-02	-1.9	-0.9
hsa-piR-28255_625566	17	77162271	77162296	1.09E-04	3.90E-04	-1.9	-0.9
hsa-piR-27080_595436	14	89875055	89875082	4.05E-06	1.77E-05	-1.9	-0.9
hsa-piR-28085_601017	21	8444893	8444922	1.92E-05	7.60E-05	-1.9	-0.9
hsa-piR-28398_626068	21	8218422	8218453	1.49E-05	6.18E-05	-1.9	-0.9
hsa-piR-14992_62867	21	8259957	8259986	8.94E-05	3.26E-04	-1.9	-0.9
hsa-piR-32837_734807	16	34161416	34161442	4.63E-07	2.40E-06	-1.9	-1.0
hsa-piR-32837_734806	21	8211013	8211039	4.72E-08	3.01E-07	-2.0	-1.0
hsa-piR-28398_626069	21	8445692	8445723	2.32E-06	1.06E-05	-2.0	-1.0
hsa-piR-28085_601018	21	8400661	8400690	1.38E-05	5.82E-05	-2.0	-1.0
hsa-piR-28398_626070	21	8401460	8401491	2.83E-07	1.54E-06	-2.0	-1.0
hsa-piR-28247_625547	6	28658279	28658309	2.54E-03	7.15E-03	-2.0	-1.0
hsa-piR-28255_625562	21	8401569	8401594	3.16E-05	1.19E-04	-2.0	-1.0
hsa-piR-14992_62870	21	8442991	8443020	6.76E-05	2.52E-04	-2.0	-1.0
hsa-piR-601_774896	21	8393728	8393757	3.11E-06	1.39E-05	-2.1	-1.1
hsa-piR-28085_601019	1	237603102	237603131	1.23E-06	6.02E-06	-2.1	-1.1
hsa-piR-28255_625565	12	127166401	127166426	4.12E-06	1.77E-05	-2.1	-1.1
hsa-piR-33117_752514	10	68755205	68755240	2.30E-03	6.59E-03	-2.1	-1.1
hsa-piR-1588_175572	2	206161922	206161951	8.89E-04	2.72E-03	-2.1	-1.1
hsa-piR-33080_752465	20	17962709	17962744	2.88E-04	9.57E-04	-2.1	-1.1
hsa-piR-28734_627512	1	30968167	30968196	1.46E-03	4.26E-03	-2.1	-1.1
hsa-piR-32837_734803	21	8255221	8255247	8.71E-09	6.43E-08	-2.1	-1.1
hsa-piR-28764_627566	5	181188459	181188489	5.33E-03	1.36E-02	-2.1	-1.1
hsa-piR-33160_752562	20	2656938	2656970	9.96E-05	3.59E-04	-2.2	-1.1
hsa-piR-28247_625550	6	28673878	28673908	2.72E-04	9.11E-04	-2.2	-1.2
hsa-piR-601_774894	21	8254901	8254930	9.53E-08	5.65E-07	-2.2	-1.2
hsa-piR-27080_595438	21	8394147	8394174	4.63E-10	4.65E-09	-2.3	-1.2
hsa-piR-601_774897	21	8210693	8210722	2.33E-07	1.31E-06	-2.3	-1.2
hsa-piR-601_774893	22	11630607	11630636	3.95E-07	2.08E-06	-2.3	-1.2
hsa-piR-427_764105	21	8255327	8255355	1.09E-08	7.53E-08	-2.3	-1.2
hsa-piR-427_764102	20	29298647	29298675	2.25E-09	1.97E-08	-2.3	-1.2
hsa-piR-27080_595440	21	8255321	8255348	1.08E-10	1.23E-09	-2.4	-1.3
hsa-piR-27080_595437	21	8438357	8438384	1.69E-09	1.55E-08	-2.4	-1.3

hsa-piR-427_764104	21	8211118	8211146	9.32E-08	5.63E-07	-2.4	-1.3
hsa-piR-601_774895	21	8437938	8437967	4.11E-09	3.25E-08	-2.4	-1.3
hsa-piR-27080_595439	21	8211112	8211139	3.86E-09	3.12E-08	-2.4	-1.3
hsa-piR-33036_752333	1	40754372	40754397	1.35E-04	4.75E-04	-2.5	-1.3
hsa-piR-25987_572571	7	129770440	129770471	3.69E-04	1.18E-03	-2.5	-1.3
hsa-piR-427_764103	21	8394153	8394181	2.04E-09	1.83E-08	-2.5	-1.3
hsa-piR-1056_12506	17	77162277	77162303	1.74E-05	6.97E-05	-2.5	-1.3
hsa-piR-427_764101	21	8438363	8438391	1.89E-10	2.09E-09	-2.5	-1.3
hsa-piR-29006_628299	7	100094034	100094064	7.24E-05	2.67E-04	-2.5	-1.3
hsa-piR-32899_735175	1	237603164	237603190	3.94E-08	2.62E-07	-2.6	-1.4
hsa-piR-32899_735179	21	8400723	8400749	5.59E-10	5.46E-09	-2.9	-1.5
hsa-piR-816_788048	5	138561047	138561074	2.26E-06	1.04E-05	-2.9	-1.5
hsa-piR-32899_735176	X	109054275	109054301	3.44E-09	2.93E-08	-3.2	-1.7
hsa-piR-32899_735174	21	8217685	8217711	1.03E-08	7.30E-08	-3.2	-1.7
hsa-piR-32899_735180	11	77886585	77886611	2.48E-11	2.94E-10	-3.3	-1.7
hsa-piR-32899_735177	8	69690206	69690232	4.44E-13	5.67E-12	-3.5	-1.8
hsa-piR-32899_735178	21	8444955	8444981	4.27E-10	4.43E-09	-3.5	-1.8
hsa-piR-32927_743195	17	76561160	76561188	7.15E-09	5.39E-08	-4.8	-2.3
hsa-piR-33048_752419	7	48632774	48632801	8.73E-06	3.72E-05	-9.6	-3.3

N=48 male samples were subjected to the DE analysis. The table represents n=136 DE piRNAs between subtype 1 vs. subtype 2 in males.

Table A.5 Differentially Expressed snoRNAs between subtypes in males

DE snoRNAs (Subtype 1 vs. Subtype 2) in Males Genome assembly GRCh 38 Human Genome (hg) build 38							
snoRNA	**Chr**	**Start**	**Stop**	**P-value**	**FDR**	**FC**	**log2 FC**
SNORA74B_URS0000 00D517	5	173020727	173020930	8.18E-14	3.72E-11	3.5	1.8
SNORD3A_URS00005 A4CE1	17	19188017	19188715	2.09E-08	1.90E-06	3.2	1.7
SNORA14A_URS0000 6C8B95	7	75943783	75943917	6.10E-05	9.57E-04	2.9	1.6
SNORA13_URS00001 BD443	5	112161486	112161618	2.15E-06	8.91E-05	2.6	1.4
SCARNA4_URS00007 1234F	1	155925959	155926086	5.49E-03	2.58E-02	2.6	1.4
SNORD145_URS0000 ABD825	1	44670958	44671057	1.00E-03	7.72E-03	2.4	1.3
SNORA75_URS000046 40CB	2	231455801	231455937	1.43E-03	9.72E-03	2.2	1.2
SNORA10_URS000003 4D55	16	1962335	1962467	6.65E-04	5.82E-03	2.2	1.1
SNORA66_URS000065 6AFD	1	92840720	92840852	1.88E-04	2.25E-03	2.1	1.1
SNORA31B_URS0000 721537	13	45336315	45336448	1.98E-08	1.90E-06	2.1	1.0
SCARNA4_URS00005 C0742	1	155925959	155926087	2.28E-03	1.31E-02	2.0	1.0
SNORD3C_URS00006 6CC58	17	19190029	19190246	1.54E-05	3.33E-04	2.0	1.0
SNORD3B-2_URS0000A9F9BD	17	19063347	19064137	9.71E-03	3.98E-02	2.0	1.0
SNORA11_URS000032 FEC4	X	54814371	54814501	1.70E-04	2.19E-03	2.0	1.0
SNORA51_URS00004F 2BDE	20	2655068	2655199	1.75E-04	2.19E-03	2.0	1.0
SNORA6_URS0000714 944	3	39408390	39408540	2.83E-06	9.91E-05	2.0	1.0
SNORA18_URS000020 2A89	11	93733467	93733598	1.73E-07	1.31E-05	2.0	1.0
SNORA65_URS000081 59A6	9	127448502	127448638	1.97E-03	1.17E-02	2.0	1.0
SNORA27_URS00005F 1DD0	13	27255402	27255527	2.76E-04	3.13E-03	1.9	0.9
SNORA46_URS000001 8549	16	58548500	58548634	1.74E-03	1.11E-02	1.8	0.9

SNORA2A_URS00002E5E66	12	48656649	48656783	2.50E-07	1.63E-05	1.8	0.9
SNORA109_URS00008E39C3	X	55183481	55183608	1.22E-02	4.81E-02	1.8	0.8
SNORA17A_URS000008230A	9	136726748	136726880	1.92E-05	3.79E-04	1.8	0.8
SNORA52_URS00003E7565	11	811682	811815	5.90E-05	9.57E-04	1.8	0.8
SNODB2078_URS0000EFB92F	1	88794475	88794587	1.20E-02	4.77E-02	1.8	0.8
SNORD22_URS000066FABB	11	62852911	62853036	4.56E-04	4.41E-03	1.7	0.8
SNORD124_URS000062A15C	17	40027543	40027646	4.11E-04	4.07E-03	1.7	0.8
SNORD3D_URS0000686F5C	17	19112420	19112637	9.96E-06	2.39E-04	1.7	0.8
SNORA44_URS000002FDBC	1	28580382	28580513	9.68E-06	2.39E-04	1.7	0.8
SNORA12_URS000002084A	10	100237157	100237303	5.49E-03	2.58E-02	1.6	0.7
SNORA71A_URS000002B3FE4	20	38427307	38427444	4.36E-03	2.20E-02	1.6	0.7
SNORA62_URS00001696A6	3	39411055	39411208	1.48E-03	9.74E-03	1.6	0.7
SNORA57_URS00000CC074	11	62665423	62665571	1.10E-03	8.30E-03	1.6	0.7
SNORA50_URS0000A9F116	17	64146338	64146472	1.94E-03	1.17E-02	1.6	0.7
SCARNA10_URS0000569A4A	12	6510223	6510552	8.32E-05	1.26E-03	1.6	0.7
SNORD115-44_URS00006346FD	15	25250860	25250941	7.97E-03	3.42E-02	1.6	0.7
SNORA64_URS0000210FD6	16	1962974	1963107	2.83E-05	5.15E-04	1.6	0.7
SNORA31_URS000019135F	13	45337481	45337610	6.49E-04	5.79E-03	1.6	0.7
SCARNA13_URS000026BDF0	14	95533356	95533630	7.01E-06	1.99E-04	1.6	0.7
SNORD3B-1_URS0000720EFA	17	19061913	19062130	6.45E-04	5.79E-03	1.6	0.7
SNORD3B-2_URS0000720EFA	17	19063920	19064137	3.16E-04	3.50E-03	1.6	0.7
SNORD73A_URS000051B5A1	4	151103828	151103892	1.78E-04	2.19E-03	1.6	0.6

SNORA20_URS00005682D3	6	159780251	159780382	1.07E-04	1.52E-03	1.6	0.6
SNORA63_URS00004B4AC7	3	186787300	186787434	3.47E-03	1.86E-02	1.6	0.6
snosnR60_Z15_URS000072E51D	1	173864828	173864909	1.24E-03	8.57E-03	1.5	0.6
SNORA3B_URS0000451E99	11	8685440	8685570	1.26E-02	4.91E-02	1.5	0.6
SNORD80_URS0000832FA3	1	173864830	173864907	1.46E-03	9.74E-03	1.5	0.6
SNORA23_URS0000064325	11	9428767	9428955	9.87E-04	7.72E-03	1.5	0.6
SNORA73B_URS0000ABD8E7	1	28508559	28508763	4.00E-03	2.07E-02	1.5	0.6
SNORD82_URS0000626A32	2	231460372	231460441	9.30E-07	4.23E-05	-1.5	-0.6
SNORD51_URS0000A77ED7	2	206161879	206161958	1.78E-05	3.67E-04	-1.6	-0.6
SNORD51_URS000070BB97	2	206161879	206161957	3.31E-06	1.07E-04	-1.6	-0.7
SNORD144_URS0000A76DCB	4	82898141	82898226	4.33E-07	2.46E-05	-1.8	-0.9
SNORD24_URS000052FCE0	9	133349397	133349471	3.71E-11	8.44E-09	-2.0	-1.0
snoU2-30_URS00006CFD5E	11	93721798	93721867	4.78E-05	8.06E-04	-2.1	-1.0
SNORD36B_URS0000687979	9	133350096	133350169	2.39E-06	9.07E-05	-2.9	-1.5
SNORD103B_URS000075CFA2	1	30949119	30949206	4.62E-03	2.29E-02	-2.9	-1.5
SCARNA15_URS00005172B7	15	82755946	82756072	5.53E-09	8.38E-07	-2.9	-1.6
SNORD103A_URS000075CFA2	1	30935690	30935777	7.59E-04	6.40E-03	-3.1	-1.7

N=48 male samples were subjected to the DE analysis. The table represents n=59 DE snoRNAs between subtype 1 vs. subtype 2 in males.

Table A.6 Differentially Expressed tRNAs between subtypes in males

DE tRNAs (Subtype 1 vs. Subtype 2) in Males Genome assembly GRCh 38 Human Genome (hg) build 38							
tRNA	**Chr**	**Start**	**Stop**	**P-value**	**FDR**	**FC**	**log2 FC**
tRNA-Val-CAC-2-1	6	27280270	27280343	4.95E-04	3.35E-02	-2.0	-1.0

N=48 male samples were subjected to the DE analysis. The table represents one DE tRNA between subtype 1 vs. subtype 2 in males.

Table A.7 Differentially Expressed mRNAs between subtypes in females

Top 200 (up- and down-regulated) DE mRNAs (Subtype 1 vs. Subtype 2) in females Genome assembly GRCh 38 Human Genome (hg) build 38							
mRNA	**Chr**	**Start**	**Stop**	**P-value**	**FDR**	**FC**	**Log2 fc**
CEACAM5	19	41708585	41730434	4.69E-10	1.87E-08	15.4	3.9
TRIM31	6	30102897	30113091	6.03E-13	2.23E-10	14.0	3.8
ONECUT3	19	1753506	1780989	1.19E-09	3.87E-08	13.3	3.7
NLRP4	19	55836540	55881856	4.37E-12	7.83E-10	12.1	3.6
SPATA3	2	230990324	231025056	1.18E-10	6.96E-09	11.1	3.5
BTBD18	11	57743514	57751782	1.66E-09	4.95E-08	11.1	3.5
CCDC63	12	110846769	110907536	1.87E-11	2.12E-09	10.8	3.4
AC097636.1	3	51798803	51833390	2.41E-10	1.16E-08	10.6	3.4
CKMT1B	15	43593054	43604902	2.50E-10	1.20E-08	10.4	3.4
UGT2B11	4	69199951	69214749	2.42E-11	2.44E-09	10.1	3.3
UNC93A	6	167271169	167316015	1.23E-09	3.95E-08	9.9	3.3
NXPH1	7	8433609	8752962	7.11E-09	1.54E-07	9.8	3.3
WNT8A	5	138083990	138092366	8.36E-10	2.96E-08	9.7	3.3
OR5K1	3	98463201	98472925	9.74E-11	6.17E-09	9.7	3.3
C10orf53	10	49679651	49710262	5.35E-11	4.11E-09	9.5	3.3
DAO	12	108858932	108901044	2.82E-10	1.30E-08	9.5	3.3
CLEC4M	19	7763210	7769606	3.46E-08	5.47E-07	9.4	3.2
MAT1A	10	80271820	80289659	1.65E-08	3.03E-07	9.3	3.2
KRT32	17	41459513	41467387	8.51E-08	1.14E-06	9.2	3.2
EXD1	15	41182725	41230744	1.80E-10	9.42E-09	9.2	3.2
TENT5D	X	80335504	80445312	1.06E-09	3.55E-08	8.8	3.1
LIPM	10	88802730	88820547	1.50E-08	2.80E-07	8.8	3.1
MAJIN	11	64937517	64972109	1.16E-10	6.92E-09	8.7	3.1
ANKRD66	6	46746917	46759507	2.31E-10	1.13E-08	8.7	3.1
EDAR	2	108894471	108989373	1.54E-09	4.65E-08	8.7	3.1
IGFL4	19	46039748	46077119	1.92E-12	4.50E-10	8.6	3.1
PAX4	7	127610292	127618143	1.85E-07	2.17E-06	8.6	3.1
TEDDM1	1	182398117	182400668	4.32E-09	1.04E-07	8.5	3.1
CLUL1	18	596988	650335	5.48E-10	2.11E-08	8.5	3.1
PLA2G2F	1	20139323	20150382	3.72E-08	5.77E-07	8.4	3.1
CEACAM6	19	41750977	41772212	5.97E-08	8.48E-07	8.4	3.1
NLRP9	19	55708438	55738403	3.32E-11	3.03E-09	8.3	3.1
AFM	4	73481745	73504002	1.67E-11	2.00E-09	8.3	3.1
CYLC2	9	102995311	103018489	3.76E-14	4.92E-11	8.3	3.0
SULT2B1	19	48552172	48599426	1.43E-07	1.75E-06	8.3	3.0
STMND1	6	17102050	17131379	3.79E-10	1.62E-08	8.2	3.0

AGR2	7	16791811	16833434	5.17E-11	4.00E-09	8.2	3.0
SLC7A3	X	70925579	70931126	9.39E-12	1.31E-09	8.2	3.0
PRAMEF1	1	12791397	12796629	9.32E-08	1.22E-06	8.2	3.0
CRX	19	47819779	47843331	8.84E-10	3.09E-08	8.2	3.0
HTR1A	5	63957874	63962508	3.58E-10	1.55E-08	8.2	3.0
SLC34A1	5	177379235	177398849	4.45E-10	1.80E-08	8.1	3.0
OR4Q3	14	19743571	19749470	1.05E-08	2.12E-07	8.1	3.0
OR51D1	11	4637477	4643061	7.67E-07	7.40E-06	8.1	3.0
OR2M3	1	248197265	248212926	4.91E-11	3.93E-09	8.0	3.0
GOLGA6L2	15	23439038	23447244	1.61E-12	4.21E-10	8.0	3.0
OR1I1	19	15082211	15092971	1.34E-09	4.23E-08	8.0	3.0
CATSPERD	19	5720637	5778735	2.35E-12	5.13E-10	8.0	3.0
DSC1	18	31129236	31162857	7.48E-11	5.22E-09	8.0	3.0
LIN28B	6	104936616	105083333	1.88E-11	2.12E-09	8.0	3.0
PDC	1	186443566	186461115	2.77E-10	1.28E-08	8.0	3.0
REG4	1	119794017	119811581	3.00E-09	7.91E-08	8.0	3.0
PP2D1	3	19979961	20012268	3.20E-10	1.43E-08	7.9	3.0
EHF	11	34621093	34663289	2.01E-14	3.39E-11	7.8	3.0
OR1A1	17	3207539	3218897	1.59E-15	8.83E-12	7.8	3.0
GOLGA6L6	15	20531856	20541801	2.10E-07	2.42E-06	7.8	3.0
SLC5A8	12	101155493	101210239	7.18E-09	1.55E-07	7.7	3.0
OR10A2	11	6863057	6874718	3.57E-09	9.01E-08	7.7	2.9
GRXCR2	5	145858521	145937127	5.14E-11	4.00E-09	7.7	2.9
TMPRSS7	3	112034843	112081270	3.74E-10	1.61E-08	7.7	2.9
GLRA1	5	151822513	151924852	7.49E-07	7.24E-06	7.6	2.9
SNTN	3	63652675	63679021	1.89E-11	2.12E-09	7.6	2.9
OR4K2	14	19875142	19883933	3.47E-11	3.12E-09	7.6	2.9
SLC30A8	8	116950273	117176715	2.63E-13	1.39E-10	7.6	2.9
PRAMEF19	1	13369067	13371901	1.49E-04	8.24E-04	7.6	2.9
UGT2B28	4	69280475	69295051	2.39E-05	1.58E-04	7.6	2.9
ADGRD2	9	124451425	124478590	5.90E-08	8.42E-07	7.6	2.9
DNAAF6	X	107206611	107244248	2.68E-09	7.26E-08	7.6	2.9
RS1	X	18639688	18672109	3.88E-10	1.64E-08	7.5	2.9
B4GALNT2	17	49132460	49176841	4.71E-12	8.12E-10	7.5	2.9
KASH5	19	49388219	49417991	2.69E-12	5.59E-10	7.5	2.9
ZP4	1	237877864	237890923	7.01E-07	6.84E-06	7.5	2.9
SLC1A6	19	14950034	15022991	1.05E-09	3.53E-08	7.5	2.9
CYP4F8	19	15615218	15630640	1.96E-11	2.17E-09	7.5	2.9
ANKS4B	16	21233699	21253851	2.21E-09	6.21E-08	7.5	2.9
CNGB3	8	86553977	86743676	5.47E-12	9.08E-10	7.5	2.9
FAM71C	12	99647753	99650115	7.83E-08	1.06E-06	7.5	2.9

UGT2B15	4	68646597	68670653	1.17E-09	3.84E-08	7.4	2.9
OR51A4	11	4942831	4947606	5.00E-11	3.97E-09	7.4	2.9
MUC15	11	26559032	26572264	1.43E-09	4.43E-08	7.4	2.9
MAGEA11	X	149688228	149717269	1.36E-08	2.57E-07	7.3	2.9
C6orf118	6	165279664	165309606	2.08E-09	5.96E-08	7.3	2.9
POTEB3	15	21405401	21440500	1.14E-08	2.26E-07	7.3	2.9
IRGM	5	150846523	150900737	5.01E-07	5.12E-06	7.3	2.9
OR14J1	6	29301701	29313018	6.94E-09	1.51E-07	7.3	2.9
IL6	7	22725884	22732003	5.15E-05	3.16E-04	7.3	2.9
CNGB1	16	57882340	57971129	1.75E-08	3.17E-07	7.3	2.9
IDO2	8	39934614	40016392	5.14E-11	4.00E-09	7.3	2.9
GPRC6A	6	116792085	116829084	1.25E-09	4.00E-08	7.2	2.9
FRMPD2	10	48153088	48274697	3.65E-11	3.16E-09	7.2	2.8
OR51G2	11	4912588	4919351	1.43E-08	2.69E-07	7.2	2.8
CD5L	1	157830911	157898257	4.58E-09	1.09E-07	7.2	2.8
CCKBR	11	6259806	6272128	2.35E-09	6.56E-08	7.2	2.8
OR5H14	3	98147479	98156615	2.93E-10	1.33E-08	7.2	2.8
CTXN2	15	48178122	48203759	3.10E-09	8.08E-08	7.2	2.8
TMPRSS4	11	118077012	118121891	3.56E-13	1.64E-10	7.2	2.8
TMEM132D	12	129071725	129904026	5.90E-12	9.57E-10	7.2	2.8
TMEM207	3	190428655	190449877	1.94E-10	9.93E-09	7.1	2.8
GUCY2D	17	8002615	8020343	4.25E-07	4.44E-06	7.1	2.8
BRINP2	1	177170958	177282423	2.17E-11	2.28E-09	7.1	2.8
CD248	11	66314494	66317045	9.98E-03	3.34E-02	-1.7	-0.7
RFLNA	12	123973241	124316025	1.24E-03	5.55E-03	-1.7	-0.7
TECRL	4	64275257	64409469	2.73E-03	1.10E-02	-1.7	-0.7
STMN3	20	63639705	63657683	5.32E-03	1.97E-02	-1.7	-0.7
TPPP3	16	67389809	67393519	8.99E-03	3.06E-02	-1.7	-0.7
LGALS3BP	17	78971238	78979948	3.14E-04	1.61E-03	-1.7	-0.7
COL6A1	21	45981769	46005051	2.54E-03	1.04E-02	-1.7	-0.7
C1QTNF4	11	47589667	47594412	9.18E-03	3.12E-02	-1.7	-0.8
ANKRD65	1	1418420	1421770	2.33E-04	1.24E-03	-1.7	-0.8
IFI6	1	27666064	27672199	2.57E-05	1.68E-04	-1.7	-0.8
MFAP4	17	19383442	19387191	1.38E-03	6.07E-03	-1.7	-0.8
PRKAR1B	7	549197	727651	1.71E-04	9.32E-04	-1.7	-0.8
CBR1	21	36069941	36073167	3.94E-04	1.99E-03	-1.7	-0.8
ABI3BP	3	100749156	100993516	3.77E-03	1.46E-02	-1.7	-0.8
C1QB	1	22652762	22661638	1.73E-03	7.41E-03	-1.7	-0.8
NT5E	6	85449584	85495792	4.32E-03	1.65E-02	-1.7	-0.8
OMD	9	92414245	92424462	1.47E-02	4.61E-02	-1.7	-0.8
SCARA3	8	27633868	27676777	6.49E-03	2.33E-02	-1.7	-0.8

B3GNT9	16	67148104	67150999	8.78E-03	3.01E-02	-1.7	-0.8
SEMA3C	7	80742538	80922360	5.27E-08	7.68E-07	-1.7	-0.8
APLN	X	129645259	129654957	9.61E-03	3.24E-02	-1.7	-0.8
OSGIN2	8	89901849	89927889	1.23E-05	8.60E-05	-1.7	-0.8
FBN1	15	48408313	48645722	1.14E-03	5.15E-03	-1.7	-0.8
DPT	1	168695468	168729207	1.48E-03	6.44E-03	-1.7	-0.8
HSD11B2	16	67430652	67437554	1.51E-02	4.73E-02	-1.7	-0.8
OXCT1	5	41730065	41870426	6.30E-04	3.03E-03	-1.7	-0.8
COL5A1	9	134641803	134844844	5.74E-03	2.10E-02	-1.7	-0.8
NES	1	156668763	156677408	7.28E-04	3.44E-03	-1.8	-0.8
ACOT11	1	54542257	54639193	5.09E-03	1.90E-02	-1.8	-0.8
ECM1	1	150508062	150513790	2.83E-03	1.14E-02	-1.8	-0.8
FSCN1	7	5592816	5606656	1.83E-03	7.80E-03	-1.8	-0.8
F2R	5	76716126	76735771	8.41E-04	3.92E-03	-1.8	-0.8
ASB18	2	236194872	236264410	2.78E-03	1.12E-02	-1.8	-0.8
POMC	2	25160853	25168904	3.87E-03	1.50E-02	-1.8	-0.8
SLC25A34	1	15736258	15741393	4.89E-04	2.42E-03	-1.8	-0.8
OLFML3	1	113979391	114035573	2.29E-03	9.48E-03	-1.8	-0.8
PRICKLE1	12	42456757	42590356	8.00E-04	3.75E-03	-1.8	-0.8
CILP	15	65194760	65211474	7.52E-04	3.55E-03	-1.8	-0.8
MAP1A	15	43510958	43531621	1.20E-03	5.36E-03	-1.8	-0.8
SOX12	20	325595	330225	1.06E-04	6.02E-04	-1.8	-0.9
ASPN	9	92456205	92482507	7.76E-03	2.71E-02	-1.8	-0.9
COL6A2	21	46098112	46132849	5.85E-04	2.84E-03	-1.8	-0.9
IGSF10	3	151425384	151458710	8.23E-03	2.85E-02	-1.8	-0.9
GPNMB	7	23235967	23275109	1.37E-03	6.02E-03	-1.9	-0.9
SMOC2	6	168441151	168673446	2.31E-04	1.23E-03	-1.9	-0.9
FSTL1	3	120392293	120450994	5.48E-04	2.68E-03	-1.9	-0.9
OSR1	2	19351485	19358624	3.02E-03	1.20E-02	-1.9	-0.9
PI16	6	36948263	36964838	1.49E-03	6.50E-03	-1.9	-0.9
QPRT	16	29663279	29698700	1.48E-03	6.44E-03	-1.9	-0.9
KCTD15	19	33795933	33815764	6.04E-04	2.92E-03	-1.9	-0.9
C14orf132	14	96039362	96093972	2.05E-03	8.61E-03	-1.9	-0.9
CYS1	2	10056780	10080945	8.30E-03	2.87E-02	-1.9	-0.9
SLC45A3	1	205657851	205680510	6.62E-03	2.37E-02	-1.9	-0.9
LOXL1	15	73925989	73952138	7.81E-04	3.67E-03	-1.9	-0.9
FAP	2	162170684	162245152	7.34E-04	3.47E-03	-1.9	-0.9
FBLN1	22	45502238	45601136	1.17E-03	5.26E-03	-1.9	-0.9
SPOCK1	5	136975298	137598380	2.43E-03	9.96E-03	-1.9	-0.9
IGF2	11	2129112	2141239	1.19E-03	5.33E-03	-1.9	-0.9
C1QTNF3	5	34017858	34043214	7.05E-04	3.35E-03	-1.9	-0.9

ACOX2	3	58505136	58537284	1.17E-06	1.06E-05	-2.0	-1.0
MRC2	17	62627670	62693598	1.46E-03	6.39E-03	-2.0	-1.0
CLIC6	21	34668994	34718228	4.38E-04	2.19E-03	-2.0	-1.0
KDELR3	22	38468078	38483448	1.02E-03	4.64E-03	-2.0	-1.0
KAZALD1	10	101061989	101068132	9.93E-04	4.55E-03	-2.0	-1.0
MGP	12	14880864	14885858	1.33E-03	5.88E-03	-2.0	-1.0
CCL21	9	34709005	34710137	6.17E-03	2.23E-02	-2.0	-1.0
MYL6B	12	56152256	56159648	1.20E-02	3.90E-02	-2.0	-1.0
COL6A3	2	237324003	237414208	3.77E-04	1.90E-03	-2.0	-1.0
UCHL1	4	41256413	41268456	4.72E-04	2.34E-03	-2.1	-1.0
FN1	2	215360440	215436074	1.72E-03	7.36E-03	-2.1	-1.1
FAM180A	7	135728348	135748814	4.20E-03	1.61E-02	-2.1	-1.1
OLFML2B	1	161983192	162023855	5.90E-04	2.86E-03	-2.1	-1.1
EXTL1	1	26019884	26036465	2.09E-04	1.12E-03	-2.1	-1.1
SLC1A7	1	53087179	53142639	8.69E-04	4.03E-03	-2.1	-1.1
SMIM32	5	136191468	136193163	6.32E-03	2.28E-02	-2.2	-1.1
BICD1	12	32106835	32383634	5.76E-05	3.49E-04	-2.2	-1.1
PCOLCE	7	100602363	100608176	1.16E-03	5.22E-03	-2.3	-1.2
COL1A2	7	94394895	94431228	3.33E-03	1.31E-02	-2.3	-1.2
THY1	11	119415476	119424986	3.39E-03	1.33E-02	-2.4	-1.3
FNDC1	6	159169400	159272109	4.61E-03	1.74E-02	-2.4	-1.3
CKB	14	103519667	103522834	2.47E-04	1.30E-03	-2.5	-1.3
TYRP1	9	12685439	12710286	2.37E-03	9.75E-03	-2.6	-1.4
ALX4	11	44260440	44310140	1.42E-03	6.23E-03	-2.6	-1.4
THBS4	5	79991311	80083288	1.00E-03	4.59E-03	-2.6	-1.4
LRRC15	3	194355249	194369744	5.60E-03	2.05E-02	-2.6	-1.4
FNDC10	1	1598012	1600136	4.04E-03	1.55E-02	-2.8	-1.5
MXRA5	X	3308565	3346653	1.19E-03	5.34E-03	-2.8	-1.5
SLC16A7	12	59596029	59789856	2.95E-06	2.38E-05	-2.8	-1.5
BGN	X	153494980	153509547	6.10E-04	2.95E-03	-2.9	-1.5
COL3A1	2	188974373	189012747	5.92E-04	2.87E-03	-3.0	-1.6
ARFGEF3	6	138161939	138344664	3.44E-04	1.76E-03	-3.0	-1.6
CCDC80	3	112596797	112649531	2.88E-04	1.49E-03	-3.1	-1.7
SFRP2	4	153780591	153789084	6.39E-04	3.07E-03	-3.2	-1.7
LUM	12	91102629	91111495	5.82E-05	3.51E-04	-3.2	-1.7
SBK2	19	55528611	55537134	2.94E-03	1.18E-02	-3.4	-1.8
HMGCS2	1	119748002	119768906	1.71E-03	7.33E-03	-3.6	-1.9
METTL21C	13	102685744	102694505	5.19E-03	1.93E-02	-4.0	-2.0
DIRAS1	19	2714567	2721373	2.45E-03	1.00E-02	-4.2	-2.1
COL1A1	17	50184101	50201633	9.51E-06	6.82E-05	-5.4	-2.4
FAM240C	2	241893988	241902552	2.11E-04	1.13E-03	-7.4	-2.9

N=36 female samples were subjected to the DE analysis. The table represents top 100 up-regulated and 100 down-regulated mRNAs between subtype 1vs. subtype 2 in females. Total n= 2847 mRNAs were found to be DE between the subtypes.

Table A.8 Differentially Expressed lncRNAs between subtypes in females

Top 200 (up- and down-regulated) DE lncRNAs (Subtype 1 vs. Subtype 2) in females Genome assembly GRCh 38 Human Genome (hg) build 38							
lncRNA	**Chr**	**Start**	**Stop**	**P-value**	**FDR**	**FC**	**Log2 FC**
CACNA2D1-AS1	7	82009177	82029956	1.28E-13	8.89E-11	21.5	4.4
PRICKLE2-DT	3	64445231	64456071	1.37E-10	7.73E-09	16.6	4.0
LINC00410	13	90890954	90926598	8.55E-12	1.22E-09	14.2	3.8
LINC02253	15	97215812	97432095	1.51E-10	8.18E-09	14.0	3.8
AL034427.1	20	9835556	9873901	5.28E-10	2.06E-08	13.0	3.7
LINC02885	22	34756676	35002863	5.81E-13	2.19E-10	12.6	3.7
SHANK2-AS1	11	70626441	70635659	1.81E-12	4.37E-10	12.3	3.6
AC090993.1	8	34228439	34346732	3.34E-09	8.55E-08	12.2	3.6
AC092053.6	3	39213069	39216586	1.74E-08	3.16E-07	12.1	3.6
AC068535.1	2	104406471	104416721	4.94E-14	5.78E-11	12.1	3.6
LINC02369	12	128086621	128118657	7.29E-11	5.10E-09	11.9	3.6
AL117353.1	1	69706950	69710275	1.54E-09	4.65E-08	11.6	3.5
AC005863.1	17	14767583	14780687	2.79E-08	4.59E-07	11.2	3.5
AL355997.1	6	51599723	51623429	5.96E-11	4.43E-09	10.6	3.4
LINC02096	17	14834557	14957217	1.12E-10	6.81E-09	10.5	3.4
ELF3-AS1	1	201995696	202010464	2.05E-12	4.65E-10	10.4	3.4
CASC17	17	71097774	71202204	3.73E-14	4.92E-11	10.1	3.3
AC010230.1	5	114475339	114668411	1.40E-08	2.64E-07	10.1	3.3
AC093791.1	4	28343862	28402937	6.21E-15	1.97E-11	10.0	3.3
AC096644.3	1	220401122	220404034	6.26E-09	1.41E-07	10.0	3.3
AC009505.1	2	105874509	105962387	6.85E-10	2.54E-08	9.9	3.3
AL133259.1	6	14152196	14211187	1.27E-09	4.05E-08	9.9	3.3
AL157359.2	21	18917934	18935860	1.08E-08	2.17E-07	9.8	3.3
AC090337.2	18	69240448	69250782	5.65E-12	9.23E-10	9.8	3.3
CSMD2-AS1	1	33868953	33893727	2.45E-09	6.77E-08	9.6	3.3
AC095350.1	12	130651342	130669234	1.61E-10	8.63E-09	9.6	3.3
SPATA3-AS1	2	230984368	230996033	1.74E-08	3.17E-07	9.6	3.3
AP001150.1	11	120249759	120265933	1.17E-11	1.55E-09	9.4	3.2
AL109653.3	X	145699056	145820633	2.14E-10	1.06E-08	9.2	3.2
AC096719.1	4	21949015	22330331	7.08E-14	7.08E-11	9.1	3.2
LINC01646	1	4571481	4594017	3.82E-10	1.63E-08	9.1	3.2
AC022133.2	5	52853183	52872926	9.25E-11	5.91E-09	9.1	3.2
LINC02401	12	103547751	103578948	4.05E-13	1.77E-10	9.0	3.2
AL158038.1	13	65941073	65973700	1.48E-09	4.56E-08	9.0	3.2
AL136524.1	13	104125930	104134258	6.21E-07	6.16E-06	9.0	3.2
AC125437.3	18	79185801	79191560	1.50E-14	3.35E-11	9.0	3.2

AC007920.1	3	187197090	187207630	3.52E-10	1.53E-08	8.9	3.2
LINC02824	12	126688341	126720334	7.19E-10	2.63E-08	8.9	3.2
AC026746.1	5	61658474	61698433	4.91E-12	8.40E-10	8.9	3.2
ARNTL2-AS1	12	27389789	27446626	2.07E-08	3.58E-07	8.9	3.2
AC109441.1	5	144439204	144468472	3.08E-09	8.04E-08	8.9	3.1
L3MBTL4-AS1	18	6256747	6260935	1.53E-12	4.06E-10	8.8	3.1
AC114550.1	8	8561386	8574039	2.66E-10	1.24E-08	8.8	3.1
AC007402.2	2	51441959	51455914	1.36E-10	7.70E-09	8.7	3.1
LINC01541	18	71519962	71578957	7.57E-11	5.24E-09	8.7	3.1
LINC02003	17	72342692	72355137	4.53E-08	6.76E-07	8.7	3.1
DAOA-AS1	13	105459055	105505682	2.57E-12	5.50E-10	8.7	3.1
AC076968.2	12	53739611	53898977	8.66E-14	7.70E-11	8.7	3.1
LINC00485	12	102809280	102824400	6.99E-12	1.07E-09	8.6	3.1
MKNK1-AS1	1	46538611	46570256	2.13E-07	2.44E-06	8.6	3.1
AC244100.2	17	35983656	35990271	1.64E-10	8.76E-09	8.5	3.1
AC124067.3	8	37559992	37567478	1.53E-09	4.63E-08	8.5	3.1
AC093838.1	2	131555007	131568600	8.79E-09	1.82E-07	8.5	3.1
LINC02338	13	58165827	58209484	1.03E-14	2.87E-11	8.5	3.1
LCAL1	6	79307669	79313385	1.15E-10	6.88E-09	8.4	3.1
AC022433.1	5	51438792	51665049	4.04E-10	1.67E-08	8.4	3.1
LINC01551	14	28771676	28823818	1.94E-13	1.20E-10	8.4	3.1
LARGE-AS1	22	33725007	33750844	2.29E-14	3.39E-11	8.4	3.1
AC020743.2	7	50202001	50262995	1.72E-08	3.14E-07	8.4	3.1
AC127540.2	17	16525832	16529291	8.13E-09	1.71E-07	8.4	3.1
AC093515.1	16	7878799	8112757	2.14E-11	2.28E-09	8.4	3.1
AL117372.1	20	60755001	60767280	3.71E-10	1.60E-08	8.4	3.1
LINC00547	13	37534940	37551537	7.14E-10	2.62E-08	8.3	3.1
AC010976.1	2	127455394	127514624	2.65E-10	1.24E-08	8.3	3.1
CARS1-AS1	11	3029009	3041261	1.99E-07	2.31E-06	8.3	3.0
LINC01297	14	19344578	19384588	1.33E-08	2.53E-07	8.3	3.0
AC016687.3	4	33775498	34039915	2.90E-10	1.33E-08	8.2	3.0
LINC02671	10	64901136	65017640	6.94E-10	2.56E-08	8.2	3.0
AC113385.1	5	100399047	100526913	2.54E-13	1.38E-10	8.2	3.0
AL033381.1	6	1079929	1104947	2.76E-07	3.05E-06	8.2	3.0
LINC01791	19	31167457	31225144	1.33E-10	7.58E-09	8.2	3.0
AC144833.1	15	27362310	27366399	5.47E-10	2.11E-08	8.1	3.0
AP001360.1	11	120008044	120013621	5.44E-08	7.89E-07	8.1	3.0
MIR137HG	1	97933474	98049864	2.27E-13	1.26E-10	8.1	3.0
Z82186.1	22	47345569	47373542	2.20E-06	1.84E-05	8.1	3.0
AC096585.1	4	96841995	96857901	5.12E-09	1.20E-07	8.0	3.0
AC023824.3	16	35207884	35284147	2.23E-08	3.81E-07	8.0	3.0

LINC02548	11	13784017	13848003	4.63E-13	1.98E-10	8.0	3.0
AC060765.1	8	82514568	82677154	5.97E-08	8.49E-07	8.0	3.0
OSBPL10-AS1	3	31704058	31721653	3.60E-11	3.15E-09	8.0	3.0
AC108073.3	4	187942164	187963346	1.92E-10	9.90E-09	8.0	3.0
LINC02380	4	57424495	57478342	6.16E-12	9.93E-10	7.9	3.0
LINC00474	9	115888169	115925208	4.05E-08	6.19E-07	7.9	3.0
AC015813.8	17	58050789	58060740	6.60E-11	4.76E-09	7.9	3.0
AL157886.1	9	84316514	84657078	7.89E-11	5.37E-09	7.9	3.0
AL355481.1	13	83877888	83895212	1.71E-11	2.03E-09	7.9	3.0
DAB1-AS1	1	57860532	57880906	1.91E-12	4.50E-10	7.9	3.0
LINC02476	7	119495024	119907398	1.92E-10	9.88E-09	7.9	3.0
AC008894.2	19	16123661	16139893	1.78E-07	2.11E-06	7.8	3.0
AC021351.1	15	35939167	36252258	1.61E-08	2.96E-07	7.8	3.0
AC007106.1	4	27964517	27985064	1.45E-11	1.80E-09	7.8	3.0
AL355337.1	6	142671980	142681268	3.67E-09	9.24E-08	7.8	3.0
LINC02074	17	74043452	74154213	1.62E-08	2.98E-07	7.8	3.0
LINC02627	10	105808225	105820350	7.54E-09	1.61E-07	7.8	3.0
BX255923.2	9	41100793	41119910	1.36E-09	4.27E-08	7.8	3.0
AC126323.2	15	62196066	62222912	1.83E-09	5.35E-08	7.8	3.0
AL158058.2	14	29415856	29500461	2.13E-11	2.28E-09	7.8	3.0
AL772363.1	9	137867925	137892571	1.14E-08	2.26E-07	7.7	3.0
LINC02201	5	122628952	122730686	3.53E-13	1.64E-10	7.7	2.9
LINC02488	5	87662040	87705338	1.06E-09	3.55E-08	7.7	2.9
AC125437.1	18	79117207	79117921	9.61E-05	5.53E-04	1.6	0.7
AC105285.1	4	173094868	173169653	1.13E-04	6.41E-04	1.6	0.7
LINC02009	3	46416524	46423592	8.76E-03	3.00E-02	1.6	0.7
BAALC-AS1	8	103153394	103298773	6.15E-04	2.97E-03	1.6	0.7
LINC00623	1	120913184	121009292	1.20E-02	3.91E-02	1.6	0.7
AL121938.1	6	125268087	125274829	4.02E-03	1.55E-02	1.6	0.7
Z93943.1	X	102945971	102952040	1.99E-04	1.07E-03	1.6	0.7
AL121992.3	1	15565611	15565957	1.17E-02	3.81E-02	1.6	0.7
AC005562.1	17	30576464	30672790	1.08E-02	3.55E-02	1.6	0.7
AC018816.1	3	4814294	4887294	1.98E-03	8.34E-03	1.6	0.7
AL023806.1	6	145735570	145737219	1.09E-02	3.59E-02	1.6	0.7
AC078881.1	4	177442519	177681382	5.53E-05	3.37E-04	1.6	0.7
LINC00652	20	18786065	18794580	2.35E-03	9.68E-03	1.6	0.7
AC007785.3	19	46027648	46031668	6.39E-03	2.30E-02	1.6	0.7
LIVAR	18	70335929	70337311	7.57E-03	2.65E-02	1.6	0.7
AC107071.1	4	108669949	108678875	1.41E-03	6.20E-03	1.6	0.7
AC007877.1	2	40534806	40545782	9.04E-03	3.08E-02	1.6	0.7
AP002784.1	11	94185439	94282204	8.35E-03	2.89E-02	1.6	0.7

CAVIN2-AS1	2	191846539	192044526	1.09E-03	4.93E-03	1.6	0.7
TUB-AS1	11	8060038	8069374	1.33E-02	4.26E-02	1.6	0.7
AC022154.1	19	48619272	48624133	1.59E-02	4.92E-02	1.6	0.7
AL365361.1	1	110653560	110657041	1.51E-02	4.71E-02	1.6	0.7
AC004160.1	7	11180902	11520176	9.70E-03	3.27E-02	1.6	0.7
AL359764.3	1	241453751	241484996	1.79E-03	7.64E-03	1.6	0.7
AL157911.1	14	59969116	60091784	3.72E-03	1.45E-02	1.6	0.7
AC012464.3	12	93945041	93947814	8.15E-03	2.83E-02	1.6	0.7
LINC02680	10	27564430	27594589	2.67E-03	1.08E-02	1.6	0.7
AC004672.2	12	1385833	1386988	9.17E-03	3.12E-02	1.6	0.7
AL157827.2	9	92600094	92612369	1.22E-02	3.97E-02	1.6	0.7
LINC00989	4	79491802	79623480	2.39E-03	9.81E-03	1.6	0.7
KRTAP5-AS1	11	1571353	1599185	7.85E-03	2.74E-02	1.6	0.7
AL445250.1	6	57961438	58438365	3.04E-03	1.21E-02	1.6	0.7
LINC01933	5	151949571	152270450	9.80E-03	3.30E-02	1.6	0.7
PGR-AS1	11	101129077	101209592	4.15E-03	1.59E-02	1.6	0.7
TMSB15B-AS1	X	103845151	103919549	5.58E-04	2.72E-03	1.6	0.7
LINC01619	12	91984976	92142915	1.45E-03	6.33E-03	1.6	0.7
AC016831.1	7	130876809	130913311	1.30E-02	4.17E-02	1.6	0.7
AL512353.1	1	42832522	42846423	5.89E-03	2.15E-02	1.6	0.7
AP002907.1	8	102256392	102257822	5.82E-03	2.13E-02	1.6	0.7
AC002480.2	7	22589698	22602537	1.50E-02	4.69E-02	1.6	0.7
FSIP2-AS1	2	185652374	185800152	1.82E-03	7.76E-03	1.6	0.6
AC234782.2	X	103707034	103919048	4.00E-04	2.01E-03	1.6	0.6
LINC00640	14	51333068	51365558	5.36E-03	1.98E-02	1.6	0.6
AC015802.1	17	76569792	76571241	1.28E-02	4.11E-02	1.6	0.6
CPEB2-DT	4	14909961	15002046	9.08E-06	6.55E-05	1.6	0.6
AC138627.1	16	73943078	74296763	5.11E-05	3.13E-04	1.6	0.6
ECI2-DT	6	4136072	4157386	1.92E-03	8.14E-03	1.6	0.6
LINC01819	2	43027823	43040663	3.40E-03	1.34E-02	1.6	0.6
USP2-AS1	11	119364359	119526665	2.91E-03	1.17E-02	1.6	0.6
AP001122.1	11	128614340	128686923	5.72E-05	3.47E-04	1.6	0.6
AF165147.1	21	28328738	28674849	5.30E-03	1.97E-02	1.5	0.6
MMADHC-DT	2	149587196	150047448	1.07E-03	4.87E-03	1.5	0.6
AC002451.1	7	95545191	95615133	1.42E-04	7.88E-04	1.5	0.6
AC073530.1	12	66950754	67096411	2.39E-03	9.83E-03	1.5	0.6
AL096870.10	14	24271726	24279090	2.78E-03	1.12E-02	1.5	0.6
LINC01266	3	536062	846562	1.95E-03	8.23E-03	1.5	0.6
AC100827.4	15	72615810	72618251	1.34E-03	5.92E-03	1.5	0.6
AC004834.1	7	99252452	99325827	7.14E-03	2.53E-02	1.5	0.6
AL662884.5	6	32108406	32112850	5.01E-04	2.47E-03	1.5	0.6

AC024060.2	3	3152942	3153436	6.38E-03	2.30E-02	1.5	0.6
AL359220.1	14	61294909	61322839	1.19E-03	5.32E-03	1.5	0.6
GEMIN7-AS1	19	45076510	45092636	1.13E-02	3.69E-02	1.5	0.6
LINC02656	10	6350316	6352763	3.19E-04	1.64E-03	1.5	0.6
AC004076.2	19	57449689	57453012	5.16E-05	3.17E-04	1.5	0.6
AL157400.3	10	89667181	89701275	9.84E-03	3.31E-02	1.5	0.6
AL160272.1	9	117759455	117879989	1.35E-02	4.31E-02	1.5	0.6
AC009228.1	2	24210650	24222202	7.31E-03	2.58E-02	1.5	0.6
AC022211.1	17	75145261	75146547	9.95E-03	3.34E-02	1.5	0.6
AC015871.6	15	79894502	79896696	1.16E-03	5.22E-03	1.5	0.6
LINC00471	2	231508422	231514367	1.24E-02	4.01E-02	1.5	0.6
AC109587.1	3	69013941	69056623	1.67E-03	7.19E-03	1.5	0.6
LINC01352	1	220829255	220832430	1.25E-02	4.02E-02	1.5	0.6
AL352979.4	14	52951432	52952865	5.64E-03	2.07E-02	1.5	0.6
AC097717.1	2	199867396	199911160	4.31E-03	1.64E-02	1.5	0.6
LINC01273	20	50171809	50176677	4.25E-03	1.62E-02	1.5	0.6
AL022098.1	6	11291299	11296225	1.58E-02	4.90E-02	1.5	0.6
AC011603.2	12	49127782	49188485	5.62E-04	2.74E-03	1.5	0.6
AC090518.1	15	56542952	56629593	6.12E-04	2.96E-03	1.5	0.6
LINC00240	6	26956932	27059750	1.57E-03	6.79E-03	1.5	0.6
MCM3AP-AS1	21	46229196	46259391	1.38E-05	9.55E-05	1.5	0.6
LINC02246	21	14819699	14918553	5.48E-04	2.68E-03	1.5	0.6
EIF2AK3-DT	2	88627539	88631822	2.61E-03	1.06E-02	1.5	0.6
INTS6-AS1	13	51452364	51554679	3.35E-05	2.12E-04	1.5	0.6
AC002467.1	7	107739999	107744582	4.13E-03	1.59E-02	1.5	0.6
AL049828.1	14	39432599	40390548	1.71E-03	7.36E-03	1.5	0.6
AC002398.2	19	35754566	35755491	4.70E-04	2.33E-03	-1.5	-0.6
AC026748.7	5	1345197	1350787	5.32E-04	2.60E-03	-1.6	-0.7
AL356218.2	9	33401478	33410554	3.44E-03	1.35E-02	-1.6	-0.7
AL450326.1	10	43420738	43422101	2.35E-05	1.55E-04	-1.7	-0.7
AL590004.3	6	3904920	3911980	7.06E-03	2.50E-02	-1.7	-0.7
AC012555.1	12	103841451	103844665	5.47E-03	2.02E-02	-1.7	-0.8
AL513217.1	1	201893842	201899979	5.75E-04	2.79E-03	-1.7	-0.8
AC113189.3	17	7439159	7443328	7.47E-03	2.63E-02	-1.7	-0.8
AL161669.3	14	103208100	103208877	1.33E-03	5.90E-03	-1.8	-0.8
AC048344.4	12	32109076	32109603	2.04E-04	1.10E-03	-1.9	-0.9
AC026356.1	12	32339368	32340725	2.87E-04	1.49E-03	-1.9	-0.9
LINC01405	12	110934590	110959094	1.14E-04	6.46E-04	-2.0	-1.0
AC026356.2	12	32352349	32354145	4.59E-05	2.84E-04	-2.1	-1.1
LINC02188	16	86710122	86742084	1.23E-03	5.49E-03	-2.5	-1.3
AC073349.4	7	64801958	64804118	7.70E-06	5.64E-05	-2.5	-1.3

N=36 female samples were subjected to the DE analysis. The table represents top 100 up-regulated and 100 down-regulated lncRNAs between subtype 1vs. subtype 2 in males. Total n= 2637 lncRNAs were found to be DE between the subtypes.

Table A.9 Differentially Expressed miRNAs between subtypes in females

DE miRNAs (Subtype 1 vs. Subtype 2) in Females Genome assembly GRCh 38 Human Genome (hg) build 38							
miRNA	**Chr**	**Start**	**Stop**	**P-value**	**FDR step up**	**FC**	**Log2 FC**
hsa-miR-3615	17	74748663	74748684	6.09E-03	3.88E-02	2.2	1.2
hsa-miR-18a-3p	13	91350797	91350820	1.19E-03	1.60E-02	2.1	1.0
hsa-miR-1306-5p	22	20086072	20086094	5.12E-03	3.63E-02	2.1	1.0
hsa-miR-145-5p	5	149430661	149430684	9.43E-04	1.34E-02	1.6	0.7
hsa-miR-532-3p	X	50003204	50003226	4.76E-04	8.45E-03	1.6	0.7
hsa-miR-652-3p	X	110055389	110055410	4.32E-04	8.45E-03	1.5	0.6
hsa-miR-22-5p	17	1713952	1713974	5.34E-03	3.68E-02	-1.5	-0.6
hsa-miR-22-3p	17	1713914	1713936	2.82E-03	2.52E-02	-1.5	-0.6
hsa-miR-27b-3p	9	95085505	95085526	4.78E-04	8.45E-03	-1.5	-0.6
hsa-miR-143-3p	5	149428978	149428999	4.04E-03	3.04E-02	-1.6	-0.6
hsa-miR-99a-5p	21	16539101	16539123	7.38E-03	4.45E-02	-1.6	-0.7
hsa-miR-30e-5p	1	40754371	40754393	6.12E-03	3.88E-02	-1.6	-0.7
hsa-miR-151a-3p	8	140732587	140732608	4.04E-03	3.04E-02	-1.7	-0.7
hsa-let-7g-5p	3	52268336	52268358	2.41E-03	2.48E-02	-1.7	-0.8
hsa-miR-27a-3p	19	13836447	13836468	3.06E-05	1.47E-03	-1.7	-0.8
hsa-miR-145-3p	5	149430699	149430721	5.97E-05	2.40E-03	-1.8	-0.8
hsa-miR-409-3p	14	101065346	101065368	3.89E-03	3.04E-02	-1.9	-0.9
hsa-let-7f-5p	9	94176353	94176375	2.91E-03	2.52E-02	-1.9	-0.9
hsa-miR-340-3p	5	180015319	180015341	2.57E-03	2.48E-02	-1.9	-0.9
hsa-let-7i-5p	12	62603691	62603713	5.26E-04	8.45E-03	-1.9	-1.0
hsa-miR-155-5p	21	25573983	25574007	1.45E-03	1.84E-02	-1.9	-1.0
hsa-miR-493-5p	14	100869075	100869097	2.95E-03	2.52E-02	-2.0	-1.0
hsa-miR-98-5p	X	53556299	53556321	1.75E-03	1.97E-02	-2.1	-1.0
hsa-miR-487b-3p	14	101046505	101046527	5.56E-03	3.72E-02	-2.1	-1.1
hsa-miR-126-3p	9	136670653	136670675	1.80E-03	1.97E-02	-2.1	-1.1
hsa-miR-199b-5p	9	128244783	128244806	1.59E-03	1.92E-02	-2.1	-1.1
hsa-miR-140-5p	16	69933103	69933125	4.20E-03	3.07E-02	-2.2	-1.1
hsa-miR-511-5p	10	17845122	17845143	6.60E-03	4.08E-02	-2.2	-1.2
hsa-miR-214-5p	1	172138857	172138879	3.03E-03	2.52E-02	-2.3	-1.2
hsa-miR-148a-3p	7	25949922	25949944	3.76E-04	8.45E-03	-2.5	-1.3
hsa-miR-543	14	101032033	101032055	5.19E-04	8.45E-03	-2.6	-1.4
hsa-miR-21-5p	17	59841273	59841295	1.97E-04	5.27E-03	-2.8	-1.5
hsa-miR-499a-5p	20	34990408	34990429	1.50E-05	9.03E-04	-2.8	-1.5
hsa-miR-654-3p	14	101040269	101040291	3.77E-06	4.55E-04	-2.8	-1.5
hsa-miR-379-5p	14	101022071	101022092	2.51E-03	2.48E-02	-3.1	-1.6
hsa-miR-370-3p	14	100911186	100911208	1.26E-05	9.03E-04	-3.1	-1.6

hsa-miR-146b-5p	10	102436520	102436543	1.42E-04	4.40E-03	-3.9	-2.0
hsa-miR-495-3p	14	101033804	101033826	1.46E-04	4.40E-03	-4.6	-2.2
hsa-miR-493-3p	14	100869116	100869138	2.29E-07	5.52E-05	-5.5	-2.5

N=36 female samples were subjected to the DE analysis. The table represents n=39 DE miRNAs between subtype 1 vs. subtype 2 in females.

Table A.10 Differentially Expressed piRNAs between subtypes in females

DE piRNAs (Subtype 1 vs. Subtype 2) Genome assembly GRCh 38 Human Genome (hg) build 38							
piRNA	**Chr**	**Start**	**Stop**	**P-value**	**FDR**	**FC**	**Log2 FC**
hsa-piR-32158_692422	20	18329074	18329099	2.70E-04	1.56E-02	2.5	1.3
hsa-piR-32913_735924	1	228628147	228628174	1.02E-03	2.35E-02	2.2	1.2
hsa-piR-32913_735914	1	228630389	228630416	1.36E-03	2.93E-02	1.9	1.0
hsa-piR-32952_743942	21	8400810	8400838	3.05E-03	4.49E-02	-1.6	-0.7
hsa-piR-28085_601016	21	8217623	8217652	3.13E-03	4.49E-02	-1.8	-0.9
hsa-piR-32835_734800	1	198859106	198859131	8.65E-04	2.30E-02	-2.0	-1.0
hsa-piR-28004_600873	6	27902950	27902980	9.89E-04	2.35E-02	-2.0	-1.0
hsa-piR-28004_600872	16	70789507	70789537	3.37E-03	4.49E-02	-2.0	-1.0
hsa-piR-28004_600876	1	16861921	16861951	2.27E-03	4.14E-02	-2.1	-1.1
hsa-piR-1044_11926	16	70778253	70778283	6.99E-04	2.13E-02	-2.1	-1.1
hsa-piR-1044_11927	16	70788694	70788724	1.73E-03	3.32E-02	-2.1	-1.1
hsa-piR-1044_11924	1	161523889	161523919	1.55E-04	1.19E-02	-2.2	-1.1
hsa-piR-427_764104	21	8211118	8211146	3.11E-03	4.49E-02	-2.2	-1.1
hsa-piR-1044_11928	2	156401189	156401219	1.73E-04	1.19E-02	-2.2	-1.1
hsa-piR-1044_11923	5	86372858	86372888	5.89E-04	2.13E-02	-2.2	-1.2
hsa-piR-28278_625759	6	28838485	28838516	3.12E-03	4.49E-02	-2.3	-1.2
hsa-piR-28004_600875	1	16545981	16546011	3.28E-04	1.62E-02	-2.3	-1.2
hsa-piR-32946_743914	1	161469799	161469829	2.92E-03	4.49E-02	-2.3	-1.2
hsa-piR-1044_11925	16	70779081	70779111	6.28E-04	2.13E-02	-2.3	-1.2
hsa-piR-33041_752380	7	39679583	39679608	3.30E-03	4.49E-02	-2.3	-1.2
hsa-piR-32946_743913	1	161455008	161455038	1.02E-04	1.17E-02	-2.7	-1.4
hsa-piR-28278_625758	6	28863726	28863757	6.75E-04	2.13E-02	-2.8	-1.5
hsa-piR-33041_752353	1	10999866	10999891	9.72E-05	1.17E-02	-2.9	-1.6
hsa-piR-32927_743195	17	76561160	76561188	1.55E-03	3.16E-02	-3.0	-1.6
hsa-piR-33036_752333	1	40754372	40754397	7.88E-05	1.17E-02	-3.1	-1.6
hsa-piR-32915_735937	5	136080573	136080599	7.39E-04	2.13E-02	-3.6	-1.8

N=36 female samples were subjected to the DE analysis. The table represents n=26 DE piRNAs between subtype 1 vs. subtype 2 in females.

Table A.11 Differentially Expressed snoRNAs between subtypes in females

DE snoRNAs (Subtype 1 vs. Subtype 2) Genome assembly GRCh 38 Human Genome (hg) build 38							
snoRNA	**Chr**	**Start**	**Stop**	**P-value**	**FDR**	**FC**	**Log2 FC**
SNORD115-41_URS000071C0E6	15	25245479	25245560	1.08E-03	1.15E-02	2.0	1.0
SNORD115-29_URS00000F2AEF	15	25223247	25223328	1.56E-04	3.57E-03	1.9	0.9
SNORD115-38_URS000065B7B7	15	25239839	25239920	6.52E-04	8.34E-03	1.9	0.9
SNORD115-35_URS0000682126	15	25234248	25234329	3.34E-04	6.02E-03	1.8	0.9
SNORD3A_URS00005A4CE1	17	19188017	19188715	6.26E-03	4.80E-02	1.7	0.8
SNORD115-16_URS00005CE711	15	25199449	25199530	1.66E-03	1.66E-02	1.7	0.8
SNORD115-30_URS00003737BE	15	25225204	25225285	7.38E-04	8.66E-03	1.7	0.8
SNORD115-39_URS000068E625	15	25241747	25241828	2.30E-03	2.04E-02	1.6	0.7
SNORD144_URS0000A76DCB	4	82898141	82898226	1.89E-03	1.84E-02	-1.6	-0.6
SNORD45B_URS0000725609	1	75789478	75789549	8.42E-04	9.40E-03	-1.6	-0.7
SNORD45B_URS00002F8E76	1	75789478	75789548	7.26E-04	8.66E-03	-1.6	-0.7
SNORD24_URS000052FCE0	9	133349397	133349471	1.31E-03	1.34E-02	-1.7	-0.7
snoU2-30_URS00006CFD5E	11	93721798	93721867	6.64E-03	4.80E-02	-1.8	-0.9
SNORD114-22_URS00006CE133	14	100982927	100982998	4.85E-03	3.99E-02	-1.8	-0.9
SNORD114-11_URS000075D5CC	14	100968112	100968185	2.00E-03	1.88E-02	-1.8	-0.9
SNORD114-22_URS000075A044	14	100982927	100982997	5.06E-03	4.09E-02	-1.8	-0.9
SNORD114-25_URS000075DB96	14	100986058	100986128	5.69E-03	4.52E-02	-1.8	-0.9
SNORD114-1_URS000067B8ED	14	100949834	100949905	6.06E-04	8.22E-03	-1.9	-0.9
SNORD114-14_URS000069C4EF	14	100972104	100972178	8.26E-04	9.40E-03	-1.9	-1.0
SNORD114-1_URS000075D742	14	100949834	100949904	5.51E-04	7.84E-03	-2.0	-1.0
SNORD114-14_URS000075DABA	14	100972104	100972177	6.13E-04	8.22E-03	-2.0	-1.0

SNORD114-23_URS00006EEEE6	14	100983877	100983948	1.00E-04	2.95E-03	-2.0	-1.0
SNORD114-13_URS000075AB92	14	100969880	100969952	3.64E-03	3.05E-02	-2.1	-1.1
SNORD112_URS00007014CB	14	100897921	100897997	7.28E-04	8.66E-03	-2.1	-1.1
SNORD113-6_URS00006CC403	14	100939557	100939631	3.57E-04	6.17E-03	-2.2	-1.1
SNORD114-23_URS000075B92D	14	100983877	100983947	3.79E-05	2.21E-03	-2.2	-1.1
SNORD114-21_URS00006B8E06	14	100981976	100982047	6.58E-04	8.34E-03	-2.2	-1.1
SNORD113-6_URS000075E6A0	14	100939557	100939630	2.76E-04	5.40E-03	-2.2	-1.1
SNORD114-9_URS00006DFE33	14	100966030	100966101	9.12E-05	2.85E-03	-2.3	-1.2
SNORD113-7_URS00006B57E0	14	100941127	100941203	2.22E-03	2.00E-02	-2.3	-1.2
SNORD114-21_URS000075C687	14	100981976	100982046	1.28E-04	3.34E-03	-2.3	-1.2
SNORD113-7_URS000075ED12	14	100941127	100941202	2.15E-03	1.97E-02	-2.3	-1.2
SNORD114-3_URS00007234FC	14	100953350	100953424	3.31E-05	2.21E-03	-2.4	-1.2
SNORD114-11_URS00006926AF	14	100968112	100968186	5.97E-05	2.27E-03	-2.4	-1.2
SNORD114-3_URS000075C3B9	14	100953350	100953423	2.77E-05	2.21E-03	-2.4	-1.3
SNORD114-28_URS00006E38EC	14	100989131	100989202	6.30E-05	2.27E-03	-2.4	-1.3
SNORD114-28_URS000075A673	14	100989131	100989201	7.03E-05	2.36E-03	-2.4	-1.3
SNORD113-9_URS0000674865	14	100945650	100945721	3.81E-04	6.17E-03	-2.4	-1.3
SNORD113-9_URS000075A3C2	14	100945650	100945720	3.76E-04	6.17E-03	-2.4	-1.3
SNORD114-9_URS000075F03E	14	100966030	100966100	3.70E-05	2.21E-03	-2.4	-1.3
SNORD51_URS000004E6C0	2	206161882	206161951	3.02E-03	2.58E-02	-2.5	-1.3
SNORD114-5_URS000075B3D8	14	100955371	100955439	8.72E-04	9.52E-03	-2.5	-1.3
SNORD113-3_URS00006762EA	14	100929920	100929991	2.14E-04	4.36E-03	-2.5	-1.3
SNORD113-3_URS000075A0ED	14	100929920	100929990	1.13E-04	3.13E-03	-2.6	-1.4
SNORD113-8_URS0000729F65	14	100943452	100943525	4.25E-05	2.21E-03	-2.6	-1.4
SNORD114-12_URS0000676D30	14	100968949	100969023	6.30E-05	2.27E-03	-2.7	-1.4

SNORD113-8_URS000075D37A	14	100943452	100943524	2.73E-05	2.21E-03	-2.7	-1.4
SNORD113-4_URS000075EC65	14	100936492	100936565	1.43E-04	3.54E-03	-2.7	-1.4
SNORD113-4_URS000064C51A	14	100936492	100936566	4.40E-04	6.45E-03	-2.8	-1.5
SNORD114-12_URS000075AF2D	14	100968949	100969022	4.83E-05	2.27E-03	-2.8	-1.5
SCARNA15_URS00005172B7	15	82755946	82756072	2.05E-04	4.36E-03	-2.8	-1.5
SNORD114-26_URS00006AAF5B	14	100987047	100987118	9.19E-06	1.44E-03	-2.9	-1.5
SNORD114-15_URS000075AFFC	14	100972671	100972741	1.60E-04	3.57E-03	-2.9	-1.5
SNORD114-26_URS000075CBCD	14	100987047	100987117	8.91E-06	1.44E-03	-2.9	-1.5
SNORD114-15_URS000063F2D8	14	100972671	100972742	3.24E-04	6.02E-03	-2.9	-1.5
SNORD103A_URS000075CFA2	1	30935690	30935777	4.15E-04	6.29E-03	-5.4	-2.4
SNORD103B_URS000075CFA2	1	30949119	30949206	4.10E-04	6.29E-03	-5.9	-2.6

N=36 female samples were subjected to the DE analysis. The table represents n=57 DE snoRNAs between subtype 1 vs. subtype 2 in females.

Table A.12 ceRNA analysis results for n=6 hub lncRNAs in males

CASC19 hub lncRNA in Males								
mRNA	**miR binding sites**	**Hyper-geo metric P-value**	**Interacting miRNAs**	**lncRNA -mRNA Corr R-value**	**lncRNA -mRNA Corr P-value**	**Regulation Similarity Score (RegSim)**	**FC lncRNA (Subtype 1 vs. 2)**	**FC mRNA (Subtype 1 vs. 2)**
IL12A	5	8.9E-03	hsa-miR-101-3p,hsa-miR-21-5p,hsa-miR-26b-5p,hsa-miR-27a-3p,hsa-miR-10a-5p	0.8	5.1E-12	1.0	6.0	3.0
ZIC5	16	2.6E-03	hsa-miR-101-3p,hsa-miR-21-5p,hsa-miR-98-5p,hsa-miR-148a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-96-5p,hsa-let-7g-5p,hsa-miR-26b-5p,hsa-miR-27a-3p,hsa-miR-10a-5p,hsa-miR-486-5p,hsa-miR-125a-5p,hsa-miR-423-5p,hsa-miR-92b-3p,hsa-miR-92a-3p	0.9	2.6E-16	0.9	6.0	5.6
NR1I2	6	8.4E-03	hsa-miR-98-5p,hsa-miR-148a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-27a-3p	0.9	3.3E-16	0.7	6.0	4.6
OR2AG1	4	7.1E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p	0.9	2.6E-18	1.0	6.0	5.7
SYT16	8	5.5E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-96-5p,hsa-let-7g-5p,hsa-miR-22-5p,hsa-miR-486-5p,hsa-miR-125a-5p	0.9	1.7E-21	0.9	6.0	5.5
RGS4	16	8.8E-04	hsa-miR-101-3p,hsa-miR-21-5p,hsa-miR-98-5p,hsa-miR-140-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-96-5p,hsa-miR-499a-5p,hsa-let-7g-5p,hsa-miR-26b-5p,hsa-miR-27a-3p,hsa-miR-10a-5p,hsa-miR-125b-5p,hsa-miR-486-5p,hsa-miR-125a-5p,hsa-miR-320b	0.9	1.0E-14	0.7	6.0	2.1
MOXD1	9	8.2E-03	hsa-miR-21-5p,hsa-miR-98-5p,hsa-miR-140-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-96-5p,hsa-let-7g-5p,hsa-miR-26b-5p,hsa-miR-27a-3p	0.8	4.1E-14	0.5	6.0	2.2

ICOS	7	7.0E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-miR-320a-3p,hsa-miR-320b	0.8	2.4E-12	1.0	6.0	2.6
CYP2C8	6	8.4E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-92b-3p,hsa-miR-92a-3p	0.8	1.2E-10	1.0	6.0	2.9
RTP1	6	2.6E-03	hsa-miR-27a-3p,hsa-miR-125b-5p,hsa-miR-125a-5p,hsa-miR-423-5p,hsa-miR-320a-3p,hsa-miR-320b	0.8	4.4E-13	0.9	6.0	3.1
FGF5	13	4.9E-03	hsa-miR-21-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-miR-125b-5p,hsa-miR-486-5p,hsa-miR-125a-5p,hsa-miR-423-5p,hsa-miR-92b-3p,hsa-miR-92a-3p,hsa-miR-320b	0.8	3.5E-11	0.9	6.0	2.5
GPR85	7	7.4E-04	hsa-miR-101-3p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-96-5p,hsa-let-7g-5p,hsa-miR-27a-3p	0.9	7.5E-17	0.9	6.0	3.2
BEND4	11	9.4E-04	hsa-miR-101-3p,hsa-miR-98-5p,hsa-miR-148a-3p,hsa-miR-140-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-26b-5p,hsa-miR-27a-3p,hsa-miR-320a-3p,hsa-miR-320b	0.9	9.7E-15	0.9	6.0	4.1
IL22RA1	7	7.4E-04	hsa-miR-101-3p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-26b-5p,hsa-miR-125a-5p	0.7	4.5E-10	0.7	6.0	1.9
C3orf52	8	5.5E-03	hsa-miR-101-3p,hsa-miR-98-5p,hsa-miR-148a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-26b-5p,hsa-miR-27a-3p	0.8	1.2E-13	0.9	6.0	1.9
SALL1	12	2.5E-03	hsa-miR-101-3p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-499a-5p,hsa-let-7g-5p,hsa-miR-26b-5p,hsa-miR-27a-3p,hsa-miR-423-5p,hsa-miR-92b-3p,hsa-miR-92a-3p,hsa-miR-320b	0.9	2.8E-15	1.0	6.0	5.0
TFAP2B	7	7.0E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-miR-320a-3p,hsa-miR-320b	0.9	2.1E-19	0.9	6.0	4.9
MAST1	5	8.9E-03	hsa-miR-101-3p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p	0.8	9.1E-13	1.0	6.0	6.1

TDO2	6	8.4E-03	hsa-miR-101-3p,hsa-miR-98-5p,hsa-miR-140-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p	0.9	5.5E-17	1.0	6.0	5.2
IL23A	4	7.1E-03	hsa-miR-101-3p,hsa-miR-499a-5p,hsa-miR-27a-3p,hsa-miR-10a-5p	0.8	1.1E-11	0.7	6.0	4.5
CRTAM	6	4.8E-04	hsa-miR-21-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-26b-5p	0.8	8.0E-14	0.7	6.0	3.5
TOP2A	11	7.0E-03	hsa-miR-101-3p,hsa-miR-21-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-96-5p,hsa-let-7g-5p,hsa-miR-26b-5p,hsa-miR-27a-3p,hsa-miR-22-5p,hsa-miR-423-5p,hsa-miR-320b	0.8	1.1E-13	0.8	6.0	2.0
SKIDA1	15	9.3E-04	hsa-miR-101-3p,hsa-miR-21-5p,hsa-miR-98-5p,hsa-miR-148a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-miR-10a-5p,hsa-miR-125b-5p,hsa-miR-125a-5p,hsa-miR-92b-3p,hsa-miR-92a-3p,hsa-miR-320a-3p,hsa-miR-320b	0.9	3.5E-15	0.9	6.0	2.0
CSPG5	5	8.9E-03	hsa-miR-26b-5p,hsa-miR-27a-3p,hsa-miR-486-5p,hsa-miR-125a-5p,hsa-miR-92a-3p	0.7	2.5E-09	0.6	6.0	1.9
PTPRO	11	3.8E-03	hsa-miR-101-3p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-26b-5p,hsa-miR-27a-3p,hsa-miR-125b-5p,hsa-miR-423-5p,hsa-miR-92b-3p,hsa-miR-92a-3p	0.8	1.0E-13	1.0	6.0	2.2
DNAH9	4	7.1E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p	0.9	4.1E-22	1.0	6.0	4.5
FAM133A	11	7.0E-03	hsa-miR-101-3p,hsa-miR-21-5p,hsa-miR-98-5p,hsa-miR-148a-3p,hsa-miR-140-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-96-5p,hsa-miR-499a-5p,hsa-let-7g-5p,hsa-miR-320b	0.8	8.1E-13	0.8	6.0	6.5
RAB19	4	7.1E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p	0.8	2.6E-11	1.0	6.0	5.3
EHF	8	5.5E-03	hsa-miR-101-3p,hsa-miR-98-5p,hsa-miR-140-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-26b-5p,hsa-miR-27a-3p	0.9	6.8E-22	1.0	6.0	4.8
RAB3C	12	1.3E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-96-5p,hsa-miR-499a-5p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-miR-22-5p,hsa-miR-92b-	0.8	2.2E-12	0.8	6.0	1.8

			3p,hsa-miR-92a-3p,hsa-miR-320a-3p,hsa-miR-320b					
FASLG	10	5.5E-04	hsa-miR-21-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-26b-5p,hsa-miR-27a-3p,hsa-miR-22-5p,hsa-miR-92b-3p,hsa-miR-92a-3p	0.8	5.2E-12	0.8	6.0	3.1
PAX9	12	7.7E-03	hsa-miR-101-3p,hsa-miR-21-5p,hsa-miR-140-5p,hsa-miR-499a-5p,hsa-miR-27a-3p,hsa-miR-125b-5p,hsa-miR-486-5p,hsa-miR-125a-5p,hsa-miR-423-5p,hsa-miR-92b-3p,hsa-miR-92a-3p,hsa-miR-320b	0.8	1.1E-11	0.9	6.0	1.9
ONECUT2	21	3.6E-03	hsa-miR-101-3p,hsa-miR-21-5p,hsa-miR-98-5p,hsa-miR-148a-3p,hsa-miR-140-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-96-5p,hsa-miR-499a-5p,hsa-let-7g-5p,hsa-miR-26b-5p,hsa-miR-27a-3p,hsa-miR-22-5p,hsa-miR-125b-5p,hsa-miR-486-5p,hsa-miR-125a-5p,hsa-miR-423-5p,hsa-miR-92b-3p,hsa-miR-92a-3p,hsa-miR-320a-3p,hsa-miR-320b	0.9	3.5E-15	0.9	6.0	3.6
CACNA1I	11	3.8E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-10a-5p,hsa-miR-125b-5p,hsa-miR-125a-5p,hsa-miR-92b-3p,hsa-miR-92a-3p,hsa-miR-320a-3p,hsa-miR-320b	0.8	5.3E-12	0.8	6.0	3.2
SLC6A14	14	4.9E-03	hsa-miR-21-5p,hsa-miR-98-5p,hsa-miR-140-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-26b-5p,hsa-miR-27a-3p,hsa-miR-125b-5p,hsa-miR-486-5p,hsa-miR-125a-5p,hsa-miR-92b-3p,hsa-miR-92a-3p,hsa-miR-320b	0.9	7.4E-15	0.9	6.0	7.2
ARHGEF38	9	1.8E-03	hsa-miR-101-3p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-96-5p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-miR-320a-3p,hsa-miR-320b	0.9	2.7E-20	1.0	6.0	6.3
DMP1	5	8.9E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-27a-3p	0.9	5.1E-15	0.8	6.0	4.5
HMGA2	18	4.2E-03	hsa-miR-101-3p,hsa-miR-98-5p,hsa-miR-148a-3p,hsa-miR-140-5p,hsa-let-7f-5p,hsa-let-7i-	0.9	2.7E-22	0.9	6.0	5.7

			5p,hsa-miR-96-5p,hsa-let-7g-5p,hsa-miR-26b-5p,hsa-miR-27a-3p,hsa-miR-22-5p,hsa-miR-125b-5p,hsa-miR-125a-5p,hsa-miR-423-5p,hsa-miR-92b-3p,hsa-miR-92a-3p,hsa-miR-320a-3p,hsa-miR-320b					
SLC8A2	7	7.0E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-125b-5p,hsa-miR-125a-5p,hsa-miR-423-5p	0.8	6.4E-12	1.0	6.0	2.9
TCEAL2	4	7.1E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p	0.9	3.5E-15	0.7	6.0	2.8
ONECUT3	9	8.2E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-96-5p,hsa-let-7g-5p,hsa-miR-10a-5p,hsa-miR-423-5p,hsa-miR-320a-3p,hsa-miR-320b	0.9	4.3E-15	0.7	6.0	5.7
KIF5C	17	2.1E-03	hsa-miR-101-3p,hsa-miR-21-5p,hsa-miR-98-5p,hsa-miR-148a-3p,hsa-miR-140-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-96-5p,hsa-let-7g-5p,hsa-miR-26b-5p,hsa-miR-125b-5p,hsa-miR-486-5p,hsa-miR-125a-5p,hsa-miR-423-5p,hsa-miR-92b-3p,hsa-miR-92a-3p,hsa-miR-320b	0.9	7.7E-20	0.9	6.0	3.1
CXCL5	11	7.0E-03	hsa-miR-21-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-499a-5p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-miR-92b-3p,hsa-miR-92a-3p,hsa-miR-320a-3p,hsa-miR-320b	0.8	7.6E-13	0.8	6.0	2.8
E2F7	18	5.9E-04	hsa-miR-101-3p,hsa-miR-21-5p,hsa-miR-98-5p,hsa-miR-148a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-26b-5p,hsa-miR-27a-3p,hsa-miR-10a-5p,hsa-miR-125b-5p,hsa-miR-486-5p,hsa-miR-125a-5p,hsa-miR-423-5p,hsa-miR-92b-3p,hsa-miR-92a-3p,hsa-miR-320a-3p,hsa-miR-320b	0.9	1.0E-20	0.9	6.0	4.7
LHCGR	4	7.1E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p	0.8	4.0E-11	0.9	6.0	1.9
SHANK2	17	5.5E-03	hsa-miR-101-3p,hsa-miR-98-5p,hsa-miR-148a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-96-5p,hsa-miR-499a-5p,hsa-let-7g-5p,hsa-miR-26b-5p,hsa-miR-27a-3p,hsa-miR-10a-5p,hsa-	0.9	8.0E-19	0.8	6.0	2.9

			miR-22-5p,hsa-miR-125b-5p,hsa-miR-125a-5p,hsa-miR-423-5p,hsa-miR-92b-3p,hsa-miR-92a-3p					
GRID2	11	2.0E-03	hsa-miR-101-3p,hsa-miR-148a-3p,hsa-let-7f-5p,hsa-miR-96-5p,hsa-miR-27a-3p,hsa-miR-486-5p,hsa-miR-125a-5p,hsa-miR-92b-3p,hsa-miR-92a-3p,hsa-miR-320a-3p,hsa-miR-320b	0.9	1.0E-18	0.8	6.0	4.1
CELSR3	14	2.9E-03	hsa-miR-101-3p,hsa-miR-98-5p,hsa-miR-140-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-miR-486-5p,hsa-miR-125a-5p,hsa-miR-423-5p,hsa-miR-92b-3p,hsa-miR-92a-3p,hsa-miR-320a-3p,hsa-miR-320b	0.8	3.7E-12	0.9	6.0	1.8
SALL3	11	2.0E-03	hsa-miR-21-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-26b-5p,hsa-miR-27a-3p,hsa-miR-92b-3p,hsa-miR-92a-3p,hsa-miR-320a-3p,hsa-miR-320b	0.8	9.3E-14	1.0	6.0	4.4
NRCAM	12	4.5E-03	hsa-miR-101-3p,hsa-miR-21-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-96-5p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-miR-423-5p,hsa-miR-92b-3p,hsa-miR-92a-3p,hsa-miR-320b	0.8	5.1E-11	0.7	6.0	1.8
C3orf52	8	5.5E-03	hsa-miR-101-3p,hsa-miR-98-5p,hsa-miR-148a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-26b-5p,hsa-miR-27a-3p	0.8	1.2E-13	0.9	6.0	1.9
DLX1	9	8.2E-03	hsa-miR-101-3p,hsa-miR-148a-3p,hsa-miR-96-5p,hsa-miR-27a-3p,hsa-miR-10a-5p,hsa-miR-125b-5p,hsa-miR-125a-5p,hsa-miR-320a-3p,hsa-miR-320b	0.8	1.5E-12	0.8	6.0	3.8
NRIP3	6	8.4E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-499a-5p,hsa-let-7g-5p,hsa-miR-27a-3p	0.8	1.1E-13	1.0	6.0	2.4
SPAG1	11	2.0E-03	hsa-miR-101-3p,hsa-miR-98-5p,hsa-miR-140-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-96-5p,hsa-let-7g-5p,hsa-miR-26b-5p,hsa-miR-27a-3p,hsa-miR-10a-5p,hsa-miR-22-5p	0.8	5.4E-13	0.6	6.0	1.7
TSPEAR	6	8.4E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-320a-3p,hsa-miR-320b	0.8	9.4E-14	1.0	6.0	4.1

CRX	5	1.9E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-423-5p	0.9	6.2E-18	1.0	6.0	5.8
GABRB3	7	7.0E-03	hsa-miR-26b-5p,hsa-miR-27a-3p,hsa-miR-125b-5p,hsa-miR-486-5p,hsa-miR-125a-5p,hsa-miR-92b-3p,hsa-miR-92a-3p	0.8	3.4E-12	0.9	6.0	2.3
SALL3	11	2.0E-03	hsa-miR-21-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-26b-5p,hsa-miR-27a-3p,hsa-miR-92b-3p,hsa-miR-92a-3p,hsa-miR-320a-3p,hsa-miR-320b	0.8	9.3E-14	1.0	6.0	4.4
SAMD12	20	2.4E-04	hsa-miR-101-3p,hsa-miR-98-5p,hsa-miR-148a-3p,hsa-miR-140-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-96-5p,hsa-miR-499a-5p,hsa-let-7g-5p,hsa-miR-26b-5p,hsa-miR-27a-3p,hsa-miR-10a-5p,hsa-miR-22-5p,hsa-miR-125b-5p,hsa-miR-486-5p,hsa-miR-125a-5p,hsa-miR-423-5p,hsa-miR-92b-3p,hsa-miR-92a-3p,hsa-miR-320b	0.8	3.3E-14	0.8	6.0	1.8
SALL3	11	2.0E-03	hsa-miR-21-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-26b-5p,hsa-miR-27a-3p,hsa-miR-92b-3p,hsa-miR-92a-3p,hsa-miR-320a-3p,hsa-miR-320b	0.8	9.3E-14	1.0	6.0	4.4
NDST3	7	2.6E-03	hsa-miR-101-3p,hsa-miR-98-5p,hsa-miR-140-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-27a-3p	0.9	2.5E-17	0.9	6.0	5.7
ADAM28	9	4.1E-03	hsa-miR-101-3p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-125b-5p,hsa-miR-125a-5p,hsa-miR-320a-3p,hsa-miR-320b	0.8	7.2E-14	0.9	6.0	1.8
MAB21L3	6	8.4E-03	hsa-miR-101-3p,hsa-miR-98-5p,hsa-miR-148a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p	0.9	3.5E-18	0.9	6.0	4.3
ESPL1	9	8.2E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-miR-10a-5p,hsa-miR-125b-5p,hsa-miR-125a-5p,hsa-miR-423-5p	0.8	1.9E-14	0.8	6.0	3.2
IL13	6	8.4E-03	hsa-miR-101-3p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-125a-5p	0.7	5.6E-10	1.0	6.0	4.6

IQGAP3	12	6.2E-04	hsa-miR-101-3p,hsa-miR-98-5p,hsa-miR-148a-3p,hsa-miR-140-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-10a-5p,hsa-miR-22-5p,hsa-miR-125b-5p,hsa-miR-125a-5p,hsa-miR-423-5p	0.9	6.9E-16	0.7	6.0	3.0
EPPIN	6	2.6E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-26b-5p,hsa-miR-22-5p	0.9	5.4E-20	0.8	6.0	6.1
TDO2	6	8.4E-03	hsa-miR-101-3p,hsa-miR-98-5p,hsa-miR-140-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p	0.9	5.5E-17	1.0	6.0	5.2
DIO1	7	1.2E-04	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-26b-5p,hsa-miR-486-5p,hsa-miR-125a-5p	0.9	2.1E-15	1.0	6.0	6.9
TBX5	7	7.4E-04	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-96-5p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-miR-10a-5p	0.8	5.4E-14	0.6	6.0	2.4
RRM2	13	4.9E-03	hsa-miR-101-3p,hsa-miR-98-5p,hsa-miR-140-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-96-5p,hsa-miR-499a-5p,hsa-let-7g-5p,hsa-miR-26b-5p,hsa-miR-27a-3p,hsa-miR-125b-5p,hsa-miR-125a-5p,hsa-miR-423-5p	0.9	2.5E-18	0.9	6.0	3.7
AGBL2	6	8.4E-03	hsa-miR-98-5p,hsa-miR-148a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-10a-5p	0.8	1.5E-13	0.8	6.0	2.3
LIN28B	18	1.7E-03	hsa-miR-101-3p,hsa-miR-21-5p,hsa-miR-98-5p,hsa-miR-148a-3p,hsa-miR-140-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-499a-5p,hsa-let-7g-5p,hsa-miR-26b-5p,hsa-miR-27a-3p,hsa-miR-10a-5p,hsa-miR-125b-5p,hsa-miR-486-5p,hsa-miR-125a-5p,hsa-miR-92b-3p,hsa-miR-92a-3p,hsa-miR-320b	0.9	1.1E-21	0.9	6.0	7.1
RASGRF1	5	8.9E-03	hsa-miR-101-3p,hsa-miR-26b-5p,hsa-miR-27a-3p,hsa-miR-125b-5p,hsa-miR-125a-5p	0.9	2.5E-17	1.0	6.0	4.5
LIN28A	11	3.8E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-26b-5p,hsa-miR-27a-3p,hsa-miR-125b-5p,hsa-miR-125a-5p,hsa-miR-92b-3p,hsa-miR-92a-3p,hsa-miR-320b	0.9	3.5E-19	1.0	6.0	5.6

PXT1	5	8.9E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-423-5p	0.9	5.0E-15	1.0	6.0	3.8
LRGUK	5	1.9E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-27a-3p	0.8	3.4E-14	1.0	6.0	1.9
KIF5C	17	2.1E-03	hsa-miR-101-3p,hsa-miR-21-5p,hsa-miR-98-5p,hsa-miR-148a-3p,hsa-miR-140-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-96-5p,hsa-let-7g-5p,hsa-miR-26b-5p,hsa-miR-125b-5p,hsa-miR-486-5p,hsa-miR-125a-5p,hsa-miR-423-5p,hsa-miR-92b-3p,hsa-miR-92a-3p,hsa-miR-320b	0.9	7.7E-20	0.9	6.0	3.1
SAXO1	6	8.4E-03	hsa-miR-101-3p,hsa-let-7f-5p,hsa-miR-125b-5p,hsa-miR-125a-5p,hsa-miR-320a-3p,hsa-miR-320b	0.8	1.9E-14	1.0	6.0	2.7
MUC16	5	8.9E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-125b-5p	0.9	4.5E-18	1.0	6.0	3.9
IMPG2	7	7.0E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-499a-5p,hsa-let-7g-5p,hsa-miR-26b-5p,hsa-miR-27a-3p	0.8	9.8E-13	1.0	6.0	1.7
LECT2	4	7.1E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p	0.9	1.2E-21	1.0	6.0	6.6
SAMD3	5	8.9E-03	hsa-miR-21-5p,hsa-miR-140-5p,hsa-miR-27a-3p,hsa-miR-320a-3p,hsa-miR-320b	0.8	5.4E-12	0.7	6.0	2.0
AGBL2	6	8.4E-03	hsa-miR-98-5p,hsa-miR-148a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-10a-5p	0.8	1.5E-13	0.8	6.0	2.3

DLEU1 hub lncRNA in Males								
mRNA	**miR binding sites**	**Hyper-geo metric P-value**	**Interacting miRNAs**	**lncRNA-mRNA Corr R-value**	**lncRNA-mRNA Corr P-value**	**Regulation Similarity Score (RegSim)**	**FC lncRNA (Subtype 1 vs. 2)**	**FC mRNA (Subtype 1 vs. 2)**
TAT	10	4.6E-03	hsa-miR-101-3p,hsa-miR-195-5p,hsa-miR-370-3p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-145-5p,hsa-miR-885-3p	0.7	2.9E-09	0.5	1.6	6.3
ETNPPL	5	2.4E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p	0.7	1.0E-08	0.6	1.6	2.1
NGF	7	9.6E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-10b-5p,hsa-miR-423-5p	0.7	3.6E-09	0.9	1.6	2.0
RAD54L	9	6.1E-03	hsa-miR-195-5p,hsa-miR-214-5p,hsa-miR-146b-5p,hsa-miR-140-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-16-5p,hsa-miR-26b-5p,hsa-miR-339-5p	0.7	1.5E-08	0.4	1.6	3.8
CTSV	7	3.7E-03	hsa-miR-101-3p,hsa-miR-195-5p,hsa-miR-140-5p,hsa-miR-16-5p,hsa-miR-26b-5p,hsa-miR-320a-3p,hsa-miR-320b	0.7	3.2E-09	0.3	1.6	4.7
ADAMTS6	13	8.8E-03	hsa-miR-101-3p,hsa-miR-195-5p,hsa-miR-146b-5p,hsa-miR-370-3p,hsa-miR-98-5p,hsa-miR-140-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-10b-5p,hsa-miR-26b-5p,hsa-miR-320b	0.7	3.9E-09	0.8	1.6	1.6
IQGAP3	12	1.1E-03	hsa-miR-101-3p,hsa-miR-195-5p,hsa-miR-146b-5p,hsa-miR-98-5p,hsa-miR-140-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-10b-5p,hsa-miR-22-5p,hsa-miR-423-5p	0.7	1.3E-09	0.8	1.6	3.0

TDO2	7	1.0E-03	hsa-miR-101-3p,hsa-miR-195-5p,hsa-miR-98-5p,hsa-miR-140-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p	0.7	1.3E-09	0.5	1.6	5.2
BICDL1	10	9.0E-04	hsa-miR-195-5p,hsa-miR-370-3p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-10b-5p,hsa-miR-423-5p,hsa-miR-145-5p	0.8	2.3E-10	1.0	1.6	3.7
EPPIN	6	3.5E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-26b-5p,hsa-miR-22-5p	0.7	4.6E-09	0.4	1.6	6.1
HMGA2	19	2.3E-03	hsa-miR-101-3p,hsa-miR-195-5p,hsa-miR-146b-5p,hsa-miR-370-3p,hsa-miR-98-5p,hsa-miR-140-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-96-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-26b-5p,hsa-miR-22-5p,hsa-miR-199a-5p,hsa-miR-423-5p,hsa-miR-145-5p,hsa-miR-320a-3p,hsa-miR-339-5p,hsa-miR-320b	0.7	6.0E-09	0.6	1.6	5.7
LYPD5	5	2.4E-03	hsa-miR-195-5p,hsa-miR-16-5p,hsa-miR-26b-5p,hsa-miR-145-5p,hsa-miR-339-5p	0.7	1.7E-09	0.5	1.6	2.3
RGS4	15	9.5E-03	hsa-miR-101-3p,hsa-miR-214-5p,hsa-miR-146b-5p,hsa-miR-98-5p,hsa-miR-140-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-96-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-10b-5p,hsa-miR-26b-5p,hsa-miR-145-5p,hsa-miR-339-5p,hsa-miR-320b	0.7	3.1E-09	0.8	1.6	2.1
ABCA13	7	3.7E-03	hsa-miR-195-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-320b	0.7	2.4E-09	0.5	1.6	5.1
CKAP2L	9	3.4E-04	hsa-miR-195-5p,hsa-miR-214-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-145-5p,hsa-miR-320b	0.7	1.1E-09	0.5	1.6	3.8

TDO2	7	1.0E-03	hsa-miR-101-3p,hsa-miR-195-5p,hsa-miR-98-5p,hsa-miR-140-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p	0.7	1.3E-09	0.5	1.6	5.2
BEND4	10	8.9E-03	hsa-miR-101-3p,hsa-miR-146b-5p,hsa-miR-98-5p,hsa-miR-140-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-26b-5p,hsa-miR-320a-3p,hsa-miR-320b	0.7	7.1E-09	0.6	1.6	4.1
CACNA1E	10	8.9E-03	hsa-miR-195-5p,hsa-miR-370-3p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-320a-3p,hsa-miR-339-5p,hsa-miR-320b	0.7	9.8E-09	0.7	1.6	3.7

HELLPAR hub lncRNA in Males								
mRNA	**miR binding sites**	**Hyper-geometric P-value**	**Interacting miRNAs**	**lncRNA-mRNA Corr R-value**	**lncRNA-mRNA Corr P-value**	**Regulation Similarity Score (RegSim)**	**FC lncRNA (Subtype 1 vs. 2)**	**FC mRNA (Subtype 1 vs. 2)**
ADAMTS6	23	1.9E-04	hsa-miR-101-3p,hsa-miR-374b-5p,hsa-miR-195-5p,hsa-miR-146b-5p,hsa-miR-370-3p,hsa-miR-98-5p,hsa-miR-140-5p,hsa-let-7f-5p,hsa-miR-493-3p,hsa-let-7i-5p,hsa-miR-146a-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-155-5p,hsa-miR-26b-5p,hsa-miR-30e-5p,hsa-miR-30a-5p,hsa-miR-27a-3p,hsa-miR-363-3p,hsa-miR-193b-5p,hsa-miR-92a-3p,hsa-miR-433-3p,hsa-miR-320b	0.8	2.8E-13	0.8	3.5	1.6
CACNB4	22	5.2E-03	hsa-miR-101-3p,hsa-miR-374b-5p,hsa-miR-195-5p,hsa-miR-146b-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-96-5p,hsa-miR-146a-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-155-5p,hsa-miR-26b-5p,hsa-miR-196b-5p,hsa-miR-30e-5p,hsa-miR-151a-3p,hsa-miR-30a-5p,hsa-miR-182-5p,hsa-miR-	0.9	4.6E-17	0.8	3.5	2.0

			143-3p,hsa-miR-27a-3p,hsa-miR-361-3p,hsa-miR-433-3p					
GABRG2	13	9.7E-03	hsa-miR-21-5p,hsa-miR-374b-5p,hsa-miR-195-5p,hsa-miR-146b-5p,hsa-let-7f-5p,hsa-miR-96-5p,hsa-miR-146a-5p,hsa-miR-16-5p,hsa-miR-26b-5p,hsa-miR-27a-3p,hsa-miR-423-5p,hsa-miR-320b,hsa-miR-2110	0.9	2.0E-20	0.8	3.5	6.4
LYPD6	24	5.6E-03	hsa-miR-101-3p,hsa-miR-21-5p,hsa-miR-15b-3p,hsa-miR-199b-5p,hsa-miR-98-5p,hsa-miR-140-5p,hsa-let-7f-5p,hsa-miR-493-3p,hsa-let-7i-5p,hsa-miR-96-5p,hsa-let-7g-5p,hsa-miR-379-5p,hsa-miR-192-5p,hsa-miR-155-5p,hsa-miR-151a-3p,hsa-miR-182-5p,hsa-miR-143-3p,hsa-miR-27a-3p,hsa-miR-199a-5p,hsa-miR-125b-5p,hsa-miR-197-3p,hsa-miR-361-3p,hsa-miR-433-3p,hsa-miR-320b	0.9	1.0E-15	0.9	3.5	3.0
TTLL9	10	7.6E-03	hsa-miR-140-5p,hsa-let-7f-5p,hsa-miR-493-3p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-196b-5p,hsa-miR-27a-3p,hsa-miR-199a-5p,hsa-miR-339-5p,hsa-miR-433-3p	0.9	1.4E-15	1.0	3.5	3.2
ARHGEF38	13	1.4E-03	hsa-miR-101-3p,hsa-miR-374b-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-96-5p,hsa-let-7g-5p,hsa-miR-196b-5p,hsa-miR-182-5p,hsa-miR-27a-3p,hsa-miR-320a-3p,hsa-miR-320b,hsa-miR-2110	0.9	1.4E-19	0.9	3.5	6.3
HOOK1	23	9.0E-03	hsa-miR-374b-5p,hsa-miR-195-5p,hsa-miR-199b-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-155-5p,hsa-miR-26b-5p,hsa-miR-196b-5p,hsa-miR-30e-5p,hsa-miR-30a-5p,hsa-miR-143-3p,hsa-miR-27a-3p,hsa-miR-363-3p,hsa-miR-22-5p,hsa-miR-199a-5p,hsa-miR-125b-5p,hsa-miR-361-3p,hsa-miR-145-5p,hsa-miR-92a-3p,hsa-miR-339-5p	0.9	3.6E-15	0.9	3.5	2.0

OLR1	17	5.5E-03	hsa-miR-21-5p,hsa-miR-195-5p,hsa-miR-370-3p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-miR-493-3p,hsa-let-7i-5p,hsa-miR-96-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-379-5p,hsa-miR-155-5p,hsa-miR-182-5p,hsa-miR-143-3p,hsa-miR-27a-3p,hsa-miR-423-5p,hsa-miR-320b	0.8	9.3E-11	0.9	3.5	3.5
RSPO2	11	4.4E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-miR-493-3p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-155-5p,hsa-miR-196b-5p,hsa-miR-27a-3p,hsa-miR-197-3p,hsa-miR-320a-3p,hsa-miR-320b	0.9	3.0E-19	0.9	3.5	4.4
KIF5C	29	3.2E-03	hsa-miR-101-3p,hsa-miR-21-5p,hsa-miR-374b-5p,hsa-miR-195-5p,hsa-miR-146b-5p,hsa-miR-370-3p,hsa-miR-98-5p,hsa-miR-140-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-96-5p,hsa-miR-146a-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-379-5p,hsa-miR-192-5p,hsa-miR-155-5p,hsa-miR-26b-5p,hsa-miR-30e-5p,hsa-miR-30a-5p,hsa-miR-182-5p,hsa-miR-143-3p,hsa-miR-363-3p,hsa-miR-125b-5p,hsa-miR-361-3p,hsa-miR-423-5p,hsa-miR-92a-3p,hsa-miR-433-3p,hsa-miR-320b	0.9	8.7E-19	0.9	3.5	3.1
KIF5C	29	3.2E-03	hsa-miR-101-3p,hsa-miR-21-5p,hsa-miR-374b-5p,hsa-miR-195-5p,hsa-miR-146b-5p,hsa-miR-370-3p,hsa-miR-98-5p,hsa-miR-140-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-96-5p,hsa-miR-146a-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-379-5p,hsa-miR-192-5p,hsa-miR-155-5p,hsa-miR-26b-5p,hsa-miR-30e-5p,hsa-miR-30a-5p,hsa-miR-182-5p,hsa-miR-143-3p,hsa-miR-363-3p,hsa-miR-125b-5p,hsa-miR-361-3p,hsa-miR-423-5p,hsa-miR-92a-3p,hsa-miR-433-3p,hsa-miR-320b	0.9	8.7E-19	0.9	3.5	3.1
BRCA2	13	9.7E-03	hsa-miR-195-5p,hsa-miR-146b-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-146a-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-	0.8	1.7E-14	0.8	3.5	2.0

			miR-192-5p,hsa-miR-155-5p,hsa-miR-30e-5p,hsa-miR-30a-5p,hsa-miR-145-5p					
SALL4	18	3.3E-03	hsa-miR-101-3p,hsa-miR-195-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-96-5p,hsa-miR-146a-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-155-5p,hsa-miR-30e-5p,hsa-miR-30a-5p,hsa-miR-182-5p,hsa-miR-143-3p,hsa-miR-27a-3p,hsa-miR-22-5p,hsa-miR-433-3p,hsa-miR-320b	0.9	6.5E-16	0.9	3.5	4.7
BLM	14	5.8E-03	hsa-miR-195-5p,hsa-miR-146b-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-146a-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-192-5p,hsa-miR-26b-5p,hsa-miR-30e-5p,hsa-miR-30a-5p,hsa-miR-27a-3p,hsa-miR-197-3p,hsa-miR-361-3p	0.9	6.0E-17	0.9	3.5	3.2
IQCH	11	4.4E-03	hsa-miR-21-5p,hsa-miR-199b-5p,hsa-miR-146b-5p,hsa-miR-96-5p,hsa-miR-146a-5p,hsa-miR-194-5p,hsa-miR-196b-5p,hsa-miR-182-5p,hsa-miR-27a-3p,hsa-miR-22-5p,hsa-miR-199a-5p	0.9	1.4E-16	0.9	3.5	2.7
MACC1	16	8.9E-03	hsa-miR-374b-5p,hsa-miR-195-5p,hsa-miR-199b-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-155-5p,hsa-miR-26b-5p,hsa-miR-143-3p,hsa-miR-27a-3p,hsa-miR-199a-5p,hsa-miR-197-3p,hsa-miR-145-5p,hsa-miR-193b-5p,hsa-miR-433-3p,hsa-miR-2110	0.9	2.6E-21	0.9	3.5	5.1
FASLG	13	9.7E-03	hsa-miR-21-5p,hsa-miR-15b-3p,hsa-miR-214-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-26b-5p,hsa-miR-27a-3p,hsa-miR-363-3p,hsa-miR-22-5p,hsa-miR-361-3p,hsa-miR-92a-3p	0.9	1.3E-15	0.7	3.5	3.1
KRT5	10	7.6E-03	hsa-miR-21-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-146a-5p,hsa-let-7g-5p,hsa-miR-155-5p,hsa-miR-151a-3p,hsa-miR-182-5p,hsa-miR-125b-5p	0.9	1.0E-16	0.8	3.5	4.4
KYNU	13	9.7E-03	hsa-miR-21-5p,hsa-miR-374b-5p,hsa-miR-195-5p,hsa-miR-146b-5p,hsa-miR-370-	0.8	1.8E-12	0.8	3.5	1.9

			3p,hsa-miR-146a-5p,hsa-miR-16-5p,hsa-miR-155-5p,hsa-miR-26b-5p,hsa-miR-30a-5p,hsa-miR-27a-3p,hsa-miR-145-5p,hsa-miR-320b					
TG	10	7.6E-03	hsa-miR-195-5p,hsa-miR-370-3p,hsa-let-7f-5p,hsa-miR-493-3p,hsa-let-7i-5p,hsa-miR-96-5p,hsa-miR-16-5p,hsa-miR-182-5p,hsa-miR-361-3p,hsa-miR-339-5p	0.9	1.8E-19	0.9	3.5	3.2
TAT	16	2.4E-04	hsa-miR-101-3p,hsa-miR-21-5p,hsa-miR-374b-5p,hsa-miR-195-5p,hsa-miR-370-3p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-miR-493-3p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-182-5p,hsa-miR-197-3p,hsa-miR-145-5p,hsa-miR-885-3p,hsa-miR-2110	0.9	3.0E-18	0.9	3.5	6.3
NIPAL4	16	2.0E-03	hsa-miR-101-3p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-miR-493-3p,hsa-let-7i-5p,hsa-miR-96-5p,hsa-miR-146a-5p,hsa-let-7g-5p,hsa-miR-30e-5p,hsa-miR-30a-5p,hsa-miR-182-5p,hsa-miR-143-3p,hsa-miR-27a-3p,hsa-miR-125b-5p,hsa-miR-145-5p,hsa-miR-339-5p	0.8	3.8E-13	1.0	3.5	2.9
SH3GL2	13	9.7E-03	hsa-miR-101-3p,hsa-miR-21-5p,hsa-miR-195-5p,hsa-miR-140-5p,hsa-miR-146a-5p,hsa-miR-16-5p,hsa-miR-155-5p,hsa-miR-26b-5p,hsa-miR-30a-5p,hsa-miR-182-5p,hsa-miR-143-3p,hsa-miR-27a-3p,hsa-miR-320b	0.8	7.6E-14	0.9	3.5	3.7
LMNB1	26	5.6E-03	hsa-miR-101-3p,hsa-miR-21-5p,hsa-miR-195-5p,hsa-miR-199b-5p,hsa-miR-98-5p,hsa-miR-140-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-192-5p,hsa-miR-155-5p,hsa-miR-26b-5p,hsa-miR-30a-5p,hsa-miR-182-5p,hsa-miR-27a-3p,hsa-miR-363-3p,hsa-miR-22-5p,hsa-miR-199a-5p,hsa-miR-125b-5p,hsa-miR-361-3p,hsa-miR-145-5p,hsa-miR-92a-3p,hsa-miR-433-3p,hsa-miR-320b,hsa-miR-2110	0.7	9.2E-09	0.9	3.5	1.6
SALL3	17	1.2E-03	hsa-miR-21-5p,hsa-miR-195-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-146a-	0.8	1.4E-13	0.8	3.5	4.4

			5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-26b-5p,hsa-miR-194-5p,hsa-miR-196b-5p,hsa-miR-27a-3p,hsa-miR-363-3p,hsa-miR-92a-3p,hsa-miR-320a-3p,hsa-miR-433-3p,hsa-miR-320b					
CPS1	22	5.2E-03	hsa-miR-101-3p,hsa-miR-21-5p,hsa-miR-374b-5p,hsa-miR-15b-3p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-miR-493-3p,hsa-let-7i-5p,hsa-miR-146a-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-155-5p,hsa-miR-26b-5p,hsa-miR-30e-5p,hsa-miR-30a-5p,hsa-miR-363-3p,hsa-miR-22-5p,hsa-miR-197-3p,hsa-miR-92a-3p,hsa-miR-339-5p,hsa-miR-433-3p,hsa-miR-320b	0.8	2.0E-13	0.9	3.5	2.1
COL9A1	10	7.6E-03	hsa-miR-146b-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-miR-493-3p,hsa-let-7i-5p,hsa-miR-96-5p,hsa-miR-146a-5p,hsa-let-7g-5p,hsa-miR-26b-5p,hsa-miR-197-3p	0.9	1.9E-20	0.9	3.5	6.8
HELB	14	8.0E-04	hsa-miR-101-3p,hsa-miR-374b-5p,hsa-miR-195-5p,hsa-miR-493-3p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-26b-5p,hsa-miR-30e-5p,hsa-miR-143-3p,hsa-miR-22-5p,hsa-miR-197-3p,hsa-miR-145-5p,hsa-miR-433-3p,hsa-miR-320b	0.8	4.0E-10	0.9	3.5	1.9
SPTSSB	21	2.6E-03	hsa-miR-101-3p,hsa-miR-21-5p,hsa-miR-374b-5p,hsa-miR-370-3p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-96-5p,hsa-miR-146a-5p,hsa-let-7g-5p,hsa-miR-379-5p,hsa-miR-26b-5p,hsa-miR-30e-5p,hsa-miR-30a-5p,hsa-miR-182-5p,hsa-miR-27a-3p,hsa-miR-22-5p,hsa-miR-361-3p,hsa-miR-423-5p,hsa-miR-145-5p,hsa-miR-433-3p	0.9	5.5E-20	1.0	3.5	4.7
CHRNB2	13	9.7E-03	hsa-miR-101-3p,hsa-miR-195-5p,hsa-miR-214-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-	0.9	2.8E-15	0.9	3.5	4.3

			361-3p,hsa-miR-145-5p,hsa-miR-339-5p,hsa-miR-433-3p,hsa-miR-2110					
LYPD6B	11	4.4E-03	hsa-miR-374b-5p,hsa-miR-195-5p,hsa-miR-146a-5p,hsa-miR-16-5p,hsa-miR-155-5p,hsa-miR-143-3p,hsa-miR-27a-3p,hsa-miR-361-3p,hsa-miR-423-5p,hsa-miR-433-3p,hsa-miR-320b	0.9	2.3E-18	0.9	3.5	5.3
BTLA	14	5.8E-03	hsa-miR-101-3p,hsa-miR-21-5p,hsa-miR-195-5p,hsa-miR-16-5p,hsa-miR-30e-5p,hsa-miR-151a-3p,hsa-miR-30a-5p,hsa-miR-182-5p,hsa-miR-143-3p,hsa-miR-27a-3p,hsa-miR-363-3p,hsa-miR-92a-3p,hsa-miR-339-5p,hsa-miR-433-3p	0.8	2.0E-14	0.9	3.5	3.1
CDH1	27	8.9E-03	hsa-miR-101-3p,hsa-miR-195-5p,hsa-miR-199b-5p,hsa-miR-146b-5p,hsa-miR-98-5p,hsa-miR-140-5p,hsa-let-7f-5p,hsa-miR-493-3p,hsa-let-7i-5p,hsa-miR-96-5p,hsa-miR-146a-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-379-5p,hsa-miR-26b-5p,hsa-miR-30a-5p,hsa-miR-143-3p,hsa-miR-363-3p,hsa-miR-22-5p,hsa-miR-199a-5p,hsa-miR-197-3p,hsa-miR-361-3p,hsa-miR-145-5p,hsa-miR-92a-3p,hsa-miR-320a-3p,hsa-miR-339-5p,hsa-miR-320b	0.8	2.3E-14	0.8	3.5	2.5
ADAM28	13	9.7E-03	hsa-miR-101-3p,hsa-miR-374b-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-146a-5p,hsa-let-7g-5p,hsa-miR-30e-5p,hsa-miR-30a-5p,hsa-miR-125b-5p,hsa-miR-145-5p,hsa-miR-320a-3p,hsa-miR-320b	0.9	1.3E-15	0.9	3.5	1.8
LHX1	12	2.5E-03	hsa-miR-101-3p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-96-5p,hsa-miR-146a-5p,hsa-let-7g-5p,hsa-miR-379-5p,hsa-miR-26b-5p,hsa-miR-30e-5p,hsa-miR-30a-5p,hsa-miR-182-5p	0.8	1.6E-13	0.9	3.5	10.0
NAT1	10	7.6E-03	hsa-miR-101-3p,hsa-miR-199b-5p,hsa-miR-155-5p,hsa-miR-26b-5p,hsa-miR-30e-5p,hsa-	0.8	6.3E-12	1.0	3.5	1.8

			miR-30a-5p,hsa-miR-27a-3p,hsa-miR-199a-5p,hsa-miR-320a-3p,hsa-miR-320b					
IL10	10	7.6E-03	hsa-miR-21-5p,hsa-miR-374b-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-miR-493-3p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-155-5p,hsa-miR-27a-3p,hsa-miR-2110	0.8	2.2E-11	1.0	3.5	2.3
NECAB1	18	3.3E-03	hsa-miR-101-3p,hsa-miR-374b-5p,hsa-miR-195-5p,hsa-miR-146b-5p,hsa-miR-96-5p,hsa-miR-146a-5p,hsa-miR-16-5p,hsa-miR-379-5p,hsa-miR-192-5p,hsa-miR-155-5p,hsa-miR-26b-5p,hsa-miR-30e-5p,hsa-miR-30a-5p,hsa-miR-182-5p,hsa-miR-27a-3p,hsa-miR-145-5p,hsa-miR-193b-5p,hsa-miR-433-3p	0.8	2.7E-14	0.8	3.5	3.4
PTPRO	17	5.5E-03	hsa-miR-101-3p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-155-5p,hsa-miR-26b-5p,hsa-miR-182-5p,hsa-miR-27a-3p,hsa-miR-363-3p,hsa-miR-125b-5p,hsa-miR-361-3p,hsa-miR-423-5p,hsa-miR-145-5p,hsa-miR-92a-3p,hsa-miR-433-3p	0.8	2.4E-13	1.0	3.5	2.2
KCNG3	10	7.6E-03	hsa-miR-101-3p,hsa-miR-21-5p,hsa-miR-374b-5p,hsa-miR-195-5p,hsa-miR-146b-5p,hsa-miR-96-5p,hsa-miR-146a-5p,hsa-miR-16-5p,hsa-miR-182-5p,hsa-miR-339-5p	0.8	2.1E-12	0.7	3.5	3.4
SALL3	17	1.2E-03	hsa-miR-21-5p,hsa-miR-195-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-146a-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-26b-5p,hsa-miR-194-5p,hsa-miR-196b-5p,hsa-miR-27a-3p,hsa-miR-363-3p,hsa-miR-92a-3p,hsa-miR-320a-3p,hsa-miR-433-3p,hsa-miR-320b	0.8	1.4E-13	0.8	3.5	4.4
TFAP2B	10	7.6E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-146a-5p,hsa-let-7g-5p,hsa-miR-182-5p,hsa-miR-27a-3p,hsa-miR-320a-3p,hsa-miR-320b,hsa-miR-2110	0.9	4.8E-18	0.9	3.5	4.9

LHX1	12	2.5E-03	hsa-miR-101-3p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-96-5p,hsa-miR-146a-5p,hsa-let-7g-5p,hsa-miR-379-5p,hsa-miR-26b-5p,hsa-miR-30e-5p,hsa-miR-30a-5p,hsa-miR-182-5p	0.8	1.6E-13	0.9	3.5	10.0
HOXC11	20	1.1E-03	hsa-miR-195-5p,hsa-miR-214-5p,hsa-miR-146b-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-miR-493-3p,hsa-let-7i-5p,hsa-miR-146a-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-379-5p,hsa-miR-151a-3p,hsa-miR-182-5p,hsa-miR-143-3p,hsa-miR-27a-3p,hsa-miR-361-3p,hsa-miR-423-5p,hsa-miR-145-5p,hsa-miR-339-5p,hsa-miR-320b	0.8	1.2E-12	0.7	3.5	5.6
EPHA7	24	1.8E-03	hsa-miR-101-3p,hsa-miR-374b-5p,hsa-miR-195-5p,hsa-miR-199b-5p,hsa-miR-214-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-miR-493-3p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-155-5p,hsa-miR-26b-5p,hsa-miR-196b-5p,hsa-miR-30e-5p,hsa-miR-30a-5p,hsa-miR-182-5p,hsa-miR-27a-3p,hsa-miR-199a-5p,hsa-miR-125b-5p,hsa-miR-92a-3p,hsa-miR-339-5p,hsa-miR-433-3p,hsa-miR-320b	0.9	8.7E-15	0.9	3.5	2.3
SALL3	17	1.2E-03	hsa-miR-21-5p,hsa-miR-195-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-146a-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-26b-5p,hsa-miR-194-5p,hsa-miR-196b-5p,hsa-miR-27a-3p,hsa-miR-363-3p,hsa-miR-92a-3p,hsa-miR-320a-3p,hsa-miR-433-3p,hsa-miR-320b	0.8	1.4E-13	0.8	3.5	4.4
HMGA2	32	4.2E-03	hsa-miR-101-3p,hsa-miR-195-5p,hsa-miR-199b-5p,hsa-miR-146b-5p,hsa-miR-370-3p,hsa-miR-98-5p,hsa-miR-140-5p,hsa-let-7f-5p,hsa-miR-493-3p,hsa-let-7i-5p,hsa-miR-96-5p,hsa-miR-146a-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-379-5p,hsa-miR-26b-5p,hsa-	0.9	1.2E-20	0.9	3.5	5.7

			miR-196b-5p,hsa-miR-151a-3p,hsa-miR-182-5p,hsa-miR-143-3p,hsa-miR-27a-3p,hsa-miR-363-3p,hsa-miR-22-5p,hsa-miR-199a-5p,hsa-miR-125b-5p,hsa-miR-423-5p,hsa-miR-145-5p,hsa-miR-92a-3p,hsa-miR-320a-3p,hsa-miR-339-5p,hsa-miR-433-3p,hsa-miR-320b					

LINC00511 hub lncRNA in Males								
mRNA	**miR binding sites**	**Hyper-geo metric P-value**	**Interacting miRNAs**	**lncRNA-mRNA Corr R-value**	**lncRNA-mRNA Corr P-value**	**Regulation Similarity Score (RegSim)**	**FC lncRNA (Subtype 1 vs. 2)**	**FC mRNA (Subtype 1 vs. 2)**
IQGAP3	12	5.3E-03	hsa-miR-101-3p,hsa-miR-195-5p,hsa-miR-98-5p,hsa-miR-148a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-146a-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-10a-5p,hsa-miR-22-5p,hsa-miR-423-5p	0.9	3.8E-20	0.7	4.5	3.0
NR1I2	7	2.7E-03	hsa-miR-98-5p,hsa-miR-148a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-146a-5p,hsa-let-7g-5p,hsa-miR-16-5p	0.9	1.9E-19	1.0	4.5	4.6
SCN11A	7	8.9E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-146a-5p,hsa-let-7g-5p,hsa-miR-143-3p,hsa-miR-664a-3p	0.9	6.7E-19	1.0	4.5	2.9
FAM81A	14	6.2E-03	hsa-miR-101-3p,hsa-miR-195-5p,hsa-miR-98-5p,hsa-miR-148a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-146a-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-143-3p,hsa-miR-197-3p,hsa-miR-193b-5p,hsa-miR-92b-3p,hsa-miR-92a-3p	0.9	2.1E-17	0.8	4.5	2.2
GINS4	10	7.5E-03	hsa-miR-101-3p,hsa-miR-195-5p,hsa-let-7i-5p,hsa-miR-146a-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-26b-5p,hsa-miR-10a-5p,hsa-miR-197-3p,hsa-miR-193b-5p	0.9	8.1E-16	1.0	4.5	2.6

FRMPD3	9	8.4E-03	hsa-miR-195-5p,hsa-miR-370-3p,hsa-miR-16-5p,hsa-miR-197-3p,hsa-miR-423-5p,hsa-miR-92b-3p,hsa-miR-92a-3p,hsa-miR-339-5p,hsa-miR-885-3p	0.9	2.3E-16	0.7	4.5	2.5
ESRP1	12	2.6E-03	hsa-miR-195-5p,hsa-miR-370-3p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-26b-5p,hsa-miR-143-3p,hsa-miR-197-3p,hsa-miR-92b-3p,hsa-miR-92a-3p	0.9	6.7E-19	1.0	4.5	6.7
EPPIN	6	7.6E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-26b-5p,hsa-miR-22-5p	0.9	4.9E-23	0.9	4.5	6.1
OR14J1	5	4.7E-03	hsa-miR-148a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p	0.9	5.3E-16	0.6	4.5	14.3
ETNPPL	5	4.7E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p	0.8	1.9E-12	0.8	4.5	2.1
PTPRT	12	5.3E-03	hsa-miR-195-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-143-3p,hsa-miR-10a-5p,hsa-miR-197-3p,hsa-miR-361-3p,hsa-miR-664a-3p,hsa-miR-885-3p	0.9	8.2E-24	0.9	4.5	3.7
FASLG	10	3.3E-03	hsa-miR-21-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-26b-5p,hsa-miR-22-5p,hsa-miR-361-3p,hsa-miR-92b-3p,hsa-miR-92a-3p	0.8	3.0E-13	0.8	4.5	3.1
SLC17A4	8	3.3E-03	hsa-miR-195-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-423-5p,hsa-miR-664a-3p	0.9	5.0E-22	0.9	4.5	6.7
OR14J1	5	4.7E-03	hsa-miR-148a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p	0.9	5.3E-16	0.6	4.5	14.3
BICDL1	10	3.3E-03	hsa-miR-195-5p,hsa-miR-370-3p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-10a-5p,hsa-miR-361-3p,hsa-miR-423-5p	0.9	7.6E-20	0.7	4.5	3.7
NCAPH	10	3.3E-03	hsa-miR-195-5p,hsa-miR-370-3p,hsa-miR-148a-3p,hsa-miR-16-5p,hsa-miR-22-5p,hsa-	0.9	2.2E-16	0.9	4.5	2.9

			miR-197-3p,hsa-miR-361-3p,hsa-miR-423-5p,hsa-miR-664a-3p,hsa-miR-92a-3p					
CDH1	18	4.9E-03	hsa-miR-101-3p,hsa-miR-195-5p,hsa-miR-199b-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-146a-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-26b-5p,hsa-miR-143-3p,hsa-miR-10a-5p,hsa-miR-22-5p,hsa-miR-197-3p,hsa-miR-361-3p,hsa-miR-92b-3p,hsa-miR-92a-3p,hsa-miR-339-5p	0.9	8.7E-17	0.7	4.5	2.5
PALM3	7	4.5E-04	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-26b-5p,hsa-miR-339-5p	0.9	3.2E-19	0.9	4.5	5.7
PTPN22	11	1.2E-03	hsa-miR-101-3p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-146a-5p,hsa-let-7g-5p,hsa-miR-26b-5p,hsa-miR-143-3p,hsa-miR-664a-3p,hsa-miR-92b-3p,hsa-miR-92a-3p	0.7	6.6E-10	0.7	4.5	1.6
OR14J1	5	4.7E-03	hsa-miR-148a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p	0.9	5.3E-16	0.6	4.5	14.3
CGN	15	8.2E-03	hsa-miR-21-5p,hsa-miR-195-5p,hsa-miR-370-3p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-146a-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-143-3p,hsa-miR-10a-5p,hsa-miR-197-3p,hsa-miR-423-5p,hsa-miR-193b-5p,hsa-miR-339-5p	0.9	4.5E-24	0.8	4.5	4.3
AGBL2	7	2.7E-03	hsa-miR-98-5p,hsa-miR-148a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-146a-5p,hsa-let-7g-5p,hsa-miR-10a-5p	0.8	1.4E-12	0.8	4.5	2.3
STAG3	8	8.8E-04	hsa-miR-98-5p,hsa-miR-148a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-197-3p,hsa-miR-423-5p,hsa-miR-339-5p	0.9	1.4E-16	1.0	4.5	2.6
SLC8A2	8	3.3E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-146a-5p,hsa-let-7g-5p,hsa-miR-197-3p,hsa-miR-361-3p,hsa-miR-423-5p	0.8	5.7E-14	1.0	4.5	2.9
HAP1	7	8.9E-03	hsa-miR-21-5p,hsa-miR-370-3p,hsa-miR-146a-5p,hsa-miR-361-3p,hsa-miR-423-5p,hsa-miR-193b-5p,hsa-miR-339-5p	0.8	1.8E-11	0.6	4.5	2.3

CPA4	13	4.3E-03	hsa-miR-101-3p,hsa-miR-195-5p,hsa-miR-98-5p,hsa-miR-148a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-146a-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-26b-5p,hsa-miR-143-3p,hsa-miR-22-5p,hsa-miR-423-5p	0.9	1.3E-16	1.0	4.5	3.6
PALM3	7	4.5E-04	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-26b-5p,hsa-miR-339-5p	0.9	3.2E-19	0.9	4.5	5.7
DBX2	5	4.7E-03	hsa-miR-370-3p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p	0.8	1.9E-11	1.0	4.5	2.5
STAB2	7	2.7E-03	hsa-miR-370-3p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-143-3p,hsa-miR-361-3p	0.9	4.4E-16	0.7	4.5	2.2
FAM189A1	8	8.8E-04	hsa-miR-101-3p,hsa-miR-195-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-423-5p	0.8	1.3E-13	0.9	4.5	3.0
FAM189A1	8	8.8E-04	hsa-miR-101-3p,hsa-miR-195-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-423-5p	0.8	1.3E-13	0.9	4.5	3.0
OR14J1	5	4.7E-03	hsa-miR-148a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p	0.9	5.3E-16	0.6	4.5	14.3
MAB21L3	7	2.7E-03	hsa-miR-101-3p,hsa-miR-370-3p,hsa-miR-98-5p,hsa-miR-148a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p	0.9	4.3E-17	0.8	4.5	4.3
AGBL2	7	2.7E-03	hsa-miR-98-5p,hsa-miR-148a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-146a-5p,hsa-let-7g-5p,hsa-miR-10a-5p	0.8	1.4E-12	0.8	4.5	2.3
GALNT13	9	3.5E-03	hsa-miR-195-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-197-3p,hsa-miR-92b-3p,hsa-miR-92a-3p	0.9	1.0E-14	1.0	4.5	2.9
TAT	11	3.0E-03	hsa-miR-101-3p,hsa-miR-21-5p,hsa-miR-195-5p,hsa-miR-370-3p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-197-3p,hsa-miR-885-3p	0.9	1.1E-23	0.9	4.5	6.3

LRTM2	7	2.7E-03	hsa-miR-195-5p,hsa-miR-370-3p,hsa-miR-146a-5p,hsa-miR-16-5p,hsa-miR-197-3p,hsa-miR-361-3p,hsa-miR-423-5p	0.8	7.6E-11	1.0	4.5	5.3
OR14J1	5	4.7E-03	hsa-miR-148a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p	0.9	5.3E-16	0.6	4.5	14.3
CRTAM	6	1.5E-03	hsa-miR-21-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-26b-5p	0.8	1.5E-14	0.7	4.5	3.5
LRTM2	7	2.7E-03	hsa-miR-195-5p,hsa-miR-370-3p,hsa-miR-146a-5p,hsa-miR-16-5p,hsa-miR-197-3p,hsa-miR-361-3p,hsa-miR-423-5p	0.8	7.6E-11	1.0	4.5	5.3
CRX	5	4.7E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-423-5p	0.9	3.1E-17	1.0	4.5	5.8
ZBTB32	6	7.6E-03	hsa-miR-370-3p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-146a-5p,hsa-let-7g-5p	0.8	1.8E-13	0.8	4.5	4.9
CACNA1I	12	5.3E-03	hsa-miR-195-5p,hsa-miR-370-3p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-146a-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-10a-5p,hsa-miR-92b-3p,hsa-miR-92a-3p,hsa-miR-885-3p	0.9	6.6E-16	0.9	4.5	3.2
GABRA6	7	8.9E-03	hsa-miR-98-5p,hsa-miR-148a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-22-5p,hsa-miR-339-5p	0.9	5.9E-18	1.0	4.5	7.5
GHSR	6	1.5E-03	hsa-miR-195-5p,hsa-miR-16-5p,hsa-miR-26b-5p,hsa-miR-664a-3p,hsa-miR-92b-3p,hsa-miR-92a-3p	0.9	1.8E-16	0.9	4.5	7.3
ZBP1	7	4.5E-04	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-146a-5p,hsa-let-7g-5p,hsa-miR-361-3p,hsa-miR-339-5p	0.8	1.3E-13	0.8	4.5	2.3
LAMC3	7	8.9E-03	hsa-miR-370-3p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-146a-5p,hsa-let-7g-5p,hsa-miR-16-5p	0.9	3.4E-17	1.0	4.5	3.5
HSD17B6	5	4.7E-03	hsa-miR-101-3p,hsa-miR-146a-5p,hsa-miR-16-5p,hsa-miR-26b-5p,hsa-miR-664a-3p	0.9	7.2E-15	0.8	4.5	4.5

KHDC1	5	4.7E-03	hsa-miR-21-5p,hsa-miR-146a-5p,hsa-miR-16-5p,hsa-miR-26b-5p,hsa-miR-423-5p	0.8	1.0E-12	0.8	4.5	2.6
PRSS22	8	3.3E-03	hsa-miR-195-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-146a-5p,hsa-let-7g-5p,hsa-miR-361-3p,hsa-miR-423-5p	0.8	1.4E-12	0.9	4.5	5.3
OR14J1	5	4.7E-03	hsa-miR-148a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p	0.9	5.3E-16	0.6	4.5	14.3
PRSS22	8	3.3E-03	hsa-miR-195-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-146a-5p,hsa-let-7g-5p,hsa-miR-361-3p,hsa-miR-423-5p	0.8	1.4E-12	0.9	4.5	5.3
POTEG	7	8.9E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-26b-5p,hsa-miR-92a-3p,hsa-miR-885-3p	0.8	1.2E-12	1.0	4.5	7.1
EHF	9	3.5E-03	hsa-miR-101-3p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-26b-5p,hsa-miR-664a-3p,hsa-miR-193b-5p,hsa-miR-339-5p	0.9	2.7E-21	0.9	4.5	4.8
IL22RA1	7	2.7E-03	hsa-miR-101-3p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-26b-5p,hsa-miR-197-3p	0.8	3.4E-11	0.8	4.5	1.9
ABCG4	11	1.2E-03	hsa-miR-101-3p,hsa-miR-195-5p,hsa-miR-370-3p,hsa-let-7f-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-26b-5p,hsa-miR-361-3p,hsa-miR-92b-3p,hsa-miR-92a-3p,hsa-miR-339-5p	0.9	3.1E-15	0.6	4.5	3.0
OR14J1	5	4.7E-03	hsa-miR-148a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p	0.9	5.3E-16	0.6	4.5	14.3

LINC00943 hub lncRNA in Males								
mRNA	miR binding sites	Hyper-geo metric P-value	Interacting miRNAs	lncRNA -mRNA Corr R-value	lncRNA -mRNA Corr P-value	Regulation Similarity Score (RegSim)	FC lncRNA (Subtype 1 vs. 2)	FC mRNA (Subtype 1 vs. 2)
TAT	11	4.0E-04	hsa-miR-101-3p,hsa-miR-21-5p,hsa-miR-374b-5p,hsa-miR-195-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-197-3p,hsa-miR-145-5p	0.9	1.3E-21	0.9	6.2	6.3
DNAH9	4	7.1E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p	0.9	1.0E-23	1.0	6.2	4.5
TCEAL 2	4	7.1E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p	0.9	1.3E-15	0.7	6.2	2.8
MUC4	6	8.4E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-197-3p,hsa-miR-145-5p	0.9	1.7E-17	1.0	6.2	3.8
MUC4	6	8.4E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-197-3p,hsa-miR-145-5p	0.9	1.7E-17	1.0	6.2	3.8
MUC4	6	8.4E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-197-3p,hsa-miR-145-5p	0.9	1.7E-17	1.0	6.2	3.8
SLC6A 14	16	1.3E-04	hsa-miR-21-5p,hsa-miR-374b-5p,hsa-miR-195-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-26b-5p,hsa-miR-125b-5p,hsa-miR-197-3p,hsa-miR-486-5p,hsa-miR-125a-5p,hsa-miR-92b-3p,hsa-miR-92a-3p,hsa-miR-433-3p	0.8	7.1E-14	0.9	6.2	7.2
RRM2	13	4.9E-03	hsa-miR-101-3p,hsa-miR-195-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-499a-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-26b-5p,hsa-miR-125b-5p,hsa-miR-197-3p,hsa-miR-125a-5p,hsa-miR-423-5p	0.9	1.7E-16	0.9	6.2	3.7
NGF	7	7.0E-03	hsa-miR-374b-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-423-5p	0.8	5.0E-11	0.6	6.2	2.0
MUC4	6	8.4E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-197-3p,hsa-miR-145-5p	0.9	1.7E-17	1.0	6.2	3.8

KIF5C	17	2.1E-03	hsa-miR-101-3p,hsa-miR-21-5p,hsa-miR-374b-5p,hsa-miR-195-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-26b-5p,hsa-miR-125b-5p,hsa-miR-486-5p,hsa-miR-125a-5p,hsa-miR-423-5p,hsa-miR-92b-3p,hsa-miR-92a-3p,hsa-miR-433-3p	0.9	1.1E-20	0.8	6.2	3.1
LRFN5	5	8.9E-03	hsa-miR-195-5p,hsa-miR-16-5p,hsa-miR-486-5p,hsa-miR-125a-5p,hsa-miR-145-5p	0.7	3.1E-09	0.5	6.2	1.6
PTPN22	9	8.2E-03	hsa-miR-101-3p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-26b-5p,hsa-miR-145-5p,hsa-miR-92b-3p,hsa-miR-92a-3p	0.7	5.3E-10	0.7	6.2	1.6
MUC4	6	8.4E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-197-3p,hsa-miR-145-5p	0.9	1.7E-17	1.0	6.2	3.8
IL12A	5	8.9E-03	hsa-miR-101-3p,hsa-miR-21-5p,hsa-miR-26b-5p,hsa-miR-10a-5p,hsa-miR-433-3p	0.8	1.2E-10	0.8	6.2	3.0
OMG	7	7.0E-03	hsa-miR-195-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-499a-5p,hsa-let-7g-5p,hsa-miR-16-5p	0.8	1.7E-10	0.9	6.2	3.2
MUC4	6	8.4E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-197-3p,hsa-miR-145-5p	0.9	1.7E-17	1.0	6.2	3.8
SLC22A10	7	1.2E-04	hsa-miR-374b-5p,hsa-miR-15b-3p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-145-5p	0.8	1.9E-14	0.9	6.2	6.6
SLC8A2	8	7.8E-04	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-125b-5p,hsa-miR-197-3p,hsa-miR-125a-5p,hsa-miR-423-5p	0.8	3.2E-13	1.0	6.2	2.9
BICDL1	9	4.1E-03	hsa-miR-195-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-10a-5p,hsa-miR-423-5p,hsa-miR-145-5p	0.9	1.7E-19	0.8	6.2	3.7
ITIH1	7	2.6E-03	hsa-miR-21-5p,hsa-miR-195-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p	0.9	3.8E-15	0.9	6.2	4.9
HOOK1	14	4.9E-03	hsa-miR-374b-5p,hsa-miR-195-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-	0.8	8.3E-14	0.9	6.2	2.0

			miR-16-5p,hsa-miR-26b-5p,hsa-miR-10a-5p,hsa-miR-125b-5p,hsa-miR-125a-5p,hsa-miR-145-5p,hsa-miR-92b-3p,hsa-miR-92a-3p					
RGS4	16	8.8E-04	hsa-miR-101-3p,hsa-miR-21-5p,hsa-miR-374b-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-499a-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-26b-5p,hsa-miR-10a-5p,hsa-miR-125b-5p,hsa-miR-486-5p,hsa-miR-125a-5p,hsa-miR-145-5p,hsa-miR-433-3p	0.8	2.6E-14	0.9	6.2	2.1
POTEG	7	2.6E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-26b-5p,hsa-miR-92a-3p,hsa-miR-433-3p	0.8	5.2E-12	1.0	6.2	7.1
MUC4	6	8.4E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-197-3p,hsa-miR-145-5p	0.9	1.7E-17	1.0	6.2	3.8
XRCC2	10	5.7E-03	hsa-miR-374b-5p,hsa-miR-195-5p,hsa-miR-499a-5p,hsa-miR-16-5p,hsa-miR-26b-5p,hsa-miR-10a-5p,hsa-miR-197-3p,hsa-miR-486-5p,hsa-miR-125a-5p,hsa-miR-145-5p	0.7	4.1E-09	0.7	6.2	1.5
ABCG4	9	8.2E-03	hsa-miR-101-3p,hsa-miR-195-5p,hsa-let-7f-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-26b-5p,hsa-miR-145-5p,hsa-miR-92b-3p,hsa-miR-92a-3p	0.8	3.8E-14	0.7	6.2	3.0
ADAM 28	9	4.1E-03	hsa-miR-101-3p,hsa-miR-374b-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-125b-5p,hsa-miR-125a-5p,hsa-miR-145-5p	0.8	1.2E-14	0.9	6.2	1.8
PALM3	6	2.6E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-26b-5p	0.9	9.4E-16	0.8	6.2	5.7
MUC16	5	8.9E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-125b-5p	0.9	1.9E-17	1.0	6.2	3.9
IQGAP 3	11	3.8E-03	hsa-miR-101-3p,hsa-miR-195-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-10a-5p,hsa-miR-125b-5p,hsa-miR-125a-5p,hsa-miR-423-5p	0.9	8.8E-18	0.8	6.2	3.0
LIPH	11	9.4E-04	hsa-miR-374b-5p,hsa-miR-195-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-	0.9	7.7E-16	0.8	6.2	3.2

			miR-16-5p,hsa-miR-486-5p,hsa-miR-125a-5p,hsa-miR-92b-3p,hsa-miR-92a-3p					
FASLG	9	4.1E-03	hsa-miR-21-5p,hsa-miR-15b-3p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-26b-5p,hsa-miR-92b-3p,hsa-miR-92a-3p	0.8	1.6E-11	0.8	6.2	3.1
LECT2	4	7.1E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p	0.9	1.9E-18	1.0	6.2	6.6
MUC4	6	8.4E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-197-3p,hsa-miR-145-5p	0.9	1.7E-17	1.0	6.2	3.8
PLEKHG7	6	2.6E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-145-5p,hsa-miR-433-3p	0.9	2.0E-19	1.0	6.2	4.7
PAQR5	12	1.3E-03	hsa-miR-101-3p,hsa-miR-21-5p,hsa-miR-195-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-499a-5p,hsa-let-7g-5p,hsa-miR-26b-5p,hsa-miR-486-5p,hsa-miR-125a-5p,hsa-miR-423-5p	0.7	1.5E-09	0.8	6.2	1.5
IL13	6	8.4E-03	hsa-miR-101-3p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-125a-5p	0.7	4.4E-09	1.0	6.2	4.6
PALM3	6	2.6E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-26b-5p	0.9	9.4E-16	0.8	6.2	5.7
ESRP1	11	2.0E-03	hsa-miR-374b-5p,hsa-miR-195-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-26b-5p,hsa-miR-197-3p,hsa-miR-92b-3p,hsa-miR-92a-3p	0.9	6.4E-20	0.9	6.2	6.7
RNF43	11	7.0E-03	hsa-miR-195-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-10a-5p,hsa-miR-125b-5p,hsa-miR-486-5p,hsa-miR-125a-5p,hsa-miR-423-5p	0.7	4.7E-10	0.6	6.2	2.3
TNIP3	7	2.6E-03	hsa-miR-101-3p,hsa-miR-15b-3p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-433-3p	0.9	4.8E-19	0.9	6.2	5.9
LHCGR	4	7.1E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p	0.7	3.1E-09	1.0	6.2	1.9
FAM189A1	8	1.9E-04	hsa-miR-101-3p,hsa-miR-195-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-423-5p	0.8	1.8E-13	0.9	6.2	3.0

GALNT13	9	6.9E-04	hsa-miR-195-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-197-3p,hsa-miR-92b-3p,hsa-miR-92a-3p	0.9	9.3E-16	0.9	6.2	2.9
NTF3	5	1.9E-03	hsa-miR-101-3p,hsa-miR-21-5p,hsa-miR-374b-5p,hsa-miR-195-5p,hsa-miR-16-5p	0.7	4.4E-09	0.8	6.2	1.7
SYT14	5	8.9E-03	hsa-miR-21-5p,hsa-miR-374b-5p,hsa-miR-195-5p,hsa-miR-16-5p,hsa-miR-125a-5p	0.8	2.4E-13	0.6	6.2	3.4
NR1H4	4	7.1E-03	hsa-miR-195-5p,hsa-miR-16-5p,hsa-miR-197-3p,hsa-miR-92a-3p	0.9	3.6E-18	0.7	6.2	6.1
WNK3	13	2.8E-03	hsa-miR-101-3p,hsa-miR-21-5p,hsa-miR-195-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-499a-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-26b-5p,hsa-miR-10a-5p,hsa-miR-125a-5p,hsa-miR-433-3p	0.8	2.7E-11	0.9	6.2	1.7
CHEK1	15	5.1E-04	hsa-miR-101-3p,hsa-miR-374b-5p,hsa-miR-195-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-125b-5p,hsa-miR-197-3p,hsa-miR-125a-5p,hsa-miR-145-5p,hsa-miR-92b-3p,hsa-miR-92a-3p,hsa-miR-433-3p	0.8	1.9E-11	0.9	6.2	2.0
RAB19	4	7.1E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p	0.8	9.2E-11	1.0	6.2	5.3
CRTAM	6	4.8E-04	hsa-miR-21-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-26b-5p	0.8	1.9E-13	0.7	6.2	3.5
KIF5C	17	2.1E-03	hsa-miR-101-3p,hsa-miR-21-5p,hsa-miR-374b-5p,hsa-miR-195-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-26b-5p,hsa-miR-125b-5p,hsa-miR-486-5p,hsa-miR-125a-5p,hsa-miR-423-5p,hsa-miR-92b-3p,hsa-miR-92a-3p,hsa-miR-433-3p	0.9	1.1E-20	0.8	6.2	3.1
GHSR	5	8.9E-03	hsa-miR-195-5p,hsa-miR-16-5p,hsa-miR-26b-5p,hsa-miR-92b-3p,hsa-miR-92a-3p	0.8	4.3E-14	0.8	6.2	7.3
SALL1	12	2.5E-03	hsa-miR-101-3p,hsa-miR-195-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-499a-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-26b-	0.9	2.7E-15	0.8	6.2	5.0

			5p,hsa-miR-423-5p,hsa-miR-92b-3p,hsa-miR-92a-3p					
CHRNB2	9	4.1E-03	hsa-miR-101-3p,hsa-miR-195-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-145-5p,hsa-miR-433-3p	0.9	2.9E-15	0.9	6.2	4.3
PXT1	5	8.9E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-423-5p	0.8	9.7E-13	1.0	6.2	3.8
SYT16	8	5.5E-03	hsa-miR-374b-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-486-5p,hsa-miR-125a-5p,hsa-miR-433-3p	0.9	2.2E-19	0.9	6.2	5.5
MAST1	5	8.9E-03	hsa-miR-101-3p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p	0.9	5.7E-16	1.0	6.2	6.1
CPS1	14	1.7E-03	hsa-miR-101-3p,hsa-miR-21-5p,hsa-miR-374b-5p,hsa-miR-15b-3p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-26b-5p,hsa-miR-197-3p,hsa-miR-92b-3p,hsa-miR-92a-3p,hsa-miR-433-3p	0.8	2.2E-12	0.9	6.2	2.1
OR2AG1	4	7.1E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p	0.9	1.2E-15	1.0	6.2	5.7
TDO2	6	8.4E-03	hsa-miR-101-3p,hsa-miR-195-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p	0.9	5.9E-15	0.8	6.2	5.2
CACNA1I	11	3.8E-03	hsa-miR-195-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-10a-5p,hsa-miR-125b-5p,hsa-miR-125a-5p,hsa-miR-92b-3p,hsa-miR-92a-3p	0.8	4.0E-14	1.0	6.2	3.2
TOP2A	11	7.0E-03	hsa-miR-101-3p,hsa-miR-21-5p,hsa-miR-195-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-26b-5p,hsa-miR-197-3p,hsa-miR-423-5p,hsa-miR-145-5p	0.8	1.1E-12	0.8	6.2	2.0
EPS8L3	5	1.9E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-145-5p	0.9	1.9E-17	0.9	6.2	7.0
HELB	9	4.1E-03	hsa-miR-101-3p,hsa-miR-374b-5p,hsa-miR-195-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-26b-5p,hsa-miR-197-3p,hsa-miR-145-5p,hsa-miR-433-3p	0.7	1.1E-08	0.9	6.2	1.9

HTR4	6	8.4E-03	hsa-miR-195-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p	0.9	5.5E-17	0.7	6.2	2.7
NRIP3	6	8.4E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-499a-5p,hsa-let-7g-5p,hsa-miR-16-5p	0.8	6.6E-13	0.8	6.2	2.4
CDC25A	16	2.6E-03	hsa-miR-21-5p,hsa-miR-374b-5p,hsa-miR-195-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-26b-5p,hsa-miR-125b-5p,hsa-miR-197-3p,hsa-miR-125a-5p,hsa-miR-423-5p,hsa-miR-145-5p,hsa-miR-92a-3p,hsa-miR-433-3p	0.8	5.2E-14	0.9	6.2	2.1
XDH	4	7.1E-03	hsa-miR-195-5p,hsa-miR-16-5p,hsa-miR-26b-5p,hsa-miR-197-3p	0.9	1.1E-19	0.5	6.2	2.9
ASPM	11	3.8E-03	hsa-miR-21-5p,hsa-miR-195-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-499a-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-26b-5p,hsa-miR-92b-3p,hsa-miR-92a-3p,hsa-miR-433-3p	0.8	5.0E-13	0.8	6.2	2.1
DIAPH3	12	4.5E-03	hsa-miR-374b-5p,hsa-miR-195-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-26b-5p,hsa-miR-197-3p,hsa-miR-145-5p,hsa-miR-92b-3p,hsa-miR-92a-3p	0.9	3.6E-15	0.9	6.2	4.0
SHANK2	19	1.3E-04	hsa-miR-101-3p,hsa-miR-374b-5p,hsa-miR-195-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-499a-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-26b-5p,hsa-miR-10a-5p,hsa-miR-125b-5p,hsa-miR-197-3p,hsa-miR-125a-5p,hsa-miR-423-5p,hsa-miR-145-5p,hsa-miR-92b-3p,hsa-miR-92a-3p,hsa-miR-433-3p	0.9	6.8E-21	0.9	6.2	2.9
GRID2IP	4	7.1E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p	0.7	3.4E-09	1.0	6.2	3.2
CYP2C8	6	8.4E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-92b-3p,hsa-miR-92a-3p	0.7	5.4E-09	1.0	6.2	2.9
DMP1	5	8.9E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-197-3p	0.9	2.0E-15	0.8	6.2	4.5
GINS4	9	8.2E-03	hsa-miR-101-3p,hsa-miR-195-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-26b-	0.8	6.5E-14	0.9	6.2	2.6

			5p,hsa-miR-10a-5p,hsa-miR-197-3p,hsa-miR-433-3p					
KCNK1	11	3.8E-03	hsa-miR-101-3p,hsa-miR-195-5p,hsa-miR-499a-5p,hsa-miR-16-5p,hsa-miR-26b-5p,hsa-miR-10a-5p,hsa-miR-125b-5p,hsa-miR-125a-5p,hsa-miR-145-5p,hsa-miR-92b-3p,hsa-miR-92a-3p	0.9	5.8E-18	0.9	6.2	2.7
FGF5	14	9.1E-04	hsa-miR-21-5p,hsa-miR-374b-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-125b-5p,hsa-miR-486-5p,hsa-miR-125a-5p,hsa-miR-423-5p,hsa-miR-145-5p,hsa-miR-92b-3p,hsa-miR-92a-3p,hsa-miR-433-3p	0.8	1.1E-10	1.0	6.2	2.5
RASGRF1	5	8.9E-03	hsa-miR-101-3p,hsa-miR-374b-5p,hsa-miR-26b-5p,hsa-miR-125b-5p,hsa-miR-125a-5p	0.9	5.5E-21	1.0	6.2	4.5
ARHGEF39	13	2.8E-03	hsa-miR-21-5p,hsa-miR-195-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-26b-5p,hsa-miR-125b-5p,hsa-miR-125a-5p,hsa-miR-423-5p,hsa-miR-145-5p,hsa-miR-433-3p	0.9	2.2E-16	0.9	6.2	3.1
EPGN	6	8.4E-03	hsa-miR-374b-5p,hsa-miR-195-5p,hsa-miR-16-5p,hsa-miR-197-3p,hsa-miR-486-5p,hsa-miR-125a-5p	0.8	1.5E-11	0.5	6.2	3.8
TEX15	6	8.4E-03	hsa-miR-21-5p,hsa-miR-374b-5p,hsa-miR-195-5p,hsa-miR-499a-5p,hsa-miR-16-5p,hsa-miR-145-5p	0.9	3.5E-19	0.8	6.2	4.5
DIO1	7	1.2E-04	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-26b-5p,hsa-miR-486-5p,hsa-miR-125a-5p	0.8	6.7E-14	0.9	6.2	6.9
POLQ	12	1.3E-03	hsa-miR-101-3p,hsa-miR-374b-5p,hsa-miR-195-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-197-3p,hsa-miR-92b-3p,hsa-miR-92a-3p,hsa-miR-433-3p	0.8	1.0E-13	0.9	6.2	2.6
SALL3	11	2.0E-03	hsa-miR-21-5p,hsa-miR-195-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-	0.8	2.5E-14	0.9	6.2	4.4

			miR-16-5p,hsa-miR-26b-5p,hsa-miR-92b-3p,hsa-miR-92a-3p,hsa-miR-433-3p					
TDO2	6	8.4E-03	hsa-miR-101-3p,hsa-miR-195-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p	0.9	5.9E-15	0.8	6.2	5.2
SCN8A	15	1.7E-03	hsa-miR-101-3p,hsa-miR-21-5p,hsa-miR-374b-5p,hsa-miR-195-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-197-3p,hsa-miR-486-5p,hsa-miR-125a-5p,hsa-miR-145-5p,hsa-miR-92b-3p,hsa-miR-92a-3p	0.9	7.6E-17	0.8	6.2	2.9
SLC17A4	7	7.0E-03	hsa-miR-195-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-423-5p	0.9	1.2E-20	0.9	6.2	6.7
ZWINT	13	2.8E-03	hsa-miR-374b-5p,hsa-miR-195-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-26b-5p,hsa-miR-10a-5p,hsa-miR-197-3p,hsa-miR-486-5p,hsa-miR-125a-5p,hsa-miR-92a-3p	0.8	1.3E-12	0.8	6.2	2.3
CRB1	9	8.2E-03	hsa-miR-374b-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-486-5p,hsa-miR-125a-5p,hsa-miR-145-5p	0.8	8.5E-11	0.9	6.2	1.9
NWD1	7	7.4E-04	hsa-miR-374b-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-26b-5p,hsa-miR-197-3p	0.8	6.7E-13	1.0	6.2	2.6
SALL3	11	2.0E-03	hsa-miR-21-5p,hsa-miR-195-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-26b-5p,hsa-miR-92b-3p,hsa-miR-92a-3p,hsa-miR-433-3p	0.8	2.5E-14	0.9	6.2	4.4
SALL3	11	2.0E-03	hsa-miR-21-5p,hsa-miR-195-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-26b-5p,hsa-miR-92b-3p,hsa-miR-92a-3p,hsa-miR-433-3p	0.8	2.5E-14	0.9	6.2	4.4
PAX9	12	7.7E-03	hsa-miR-101-3p,hsa-miR-21-5p,hsa-miR-195-5p,hsa-miR-499a-5p,hsa-miR-16-5p,hsa-miR-125b-5p,hsa-miR-486-5p,hsa-miR-125a-5p,hsa-	0.8	5.8E-11	0.9	6.2	1.9

			miR-423-5p,hsa-miR-92b-3p,hsa-miR-92a-3p,hsa-miR-433-3p					
FAM189A1	8	1.9E-04	hsa-miR-101-3p,hsa-miR-195-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-423-5p	0.8	1.8E-13	0.9	6.2	3.0
SAMD12	19	1.9E-03	hsa-miR-101-3p,hsa-miR-195-5p,hsa-miR-15b-3p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-499a-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-26b-5p,hsa-miR-10a-5p,hsa-miR-125b-5p,hsa-miR-197-3p,hsa-miR-486-5p,hsa-miR-125a-5p,hsa-miR-423-5p,hsa-miR-145-5p,hsa-miR-92b-3p,hsa-miR-92a-3p	0.8	9.2E-14	0.9	6.2	1.8
CLGN	9	1.8E-03	hsa-miR-101-3p,hsa-miR-374b-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-92b-3p,hsa-miR-92a-3p	0.8	1.6E-12	1.0	6.2	2.8
OPRM1	10	5.7E-03	hsa-miR-374b-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-26b-5p,hsa-miR-10a-5p,hsa-miR-92b-3p,hsa-miR-92a-3p	0.9	7.1E-23	1.0	6.2	5.0
CRX	5	1.9E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-423-5p	0.9	1.2E-17	1.0	6.2	5.8
CXCL5	11	7.0E-03	hsa-miR-21-5p,hsa-miR-374b-5p,hsa-miR-195-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-499a-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-92b-3p,hsa-miR-92a-3p	0.8	4.0E-12	0.7	6.2	2.8
EPHA7	14	4.9E-03	hsa-miR-101-3p,hsa-miR-374b-5p,hsa-miR-195-5p,hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-499a-5p,hsa-let-7g-5p,hsa-miR-16-5p,hsa-miR-26b-5p,hsa-miR-125b-5p,hsa-miR-125a-5p,hsa-miR-92a-3p,hsa-miR-433-3p	0.8	1.4E-11	0.9	6.2	2.3
FGFR2	10	5.7E-03	hsa-miR-374b-5p,hsa-miR-195-5p,hsa-miR-16-5p,hsa-miR-125b-5p,hsa-miR-125a-5p,hsa-miR-423-5p,hsa-miR-145-5p,hsa-miR-92b-3p,hsa-miR-92a-3p,hsa-miR-433-3p	0.9	4.9E-16	0.9	6.2	2.9
ETNPPL	5	1.9E-03	hsa-miR-98-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-let-7g-5p,hsa-miR-16-5p	0.8	8.0E-12	0.8	6.2	2.1

XACT hub lncRNA in Males								
mRNA	**miR binding sites**	**Hyper-geo metric P-value**	**Interacting miRNAs**	**lncRNA -mRNA Corr R-value**	**lncRNA -mRNA Corr P-value**	**Regulation Similarity Score (RegSim)**	**FC lncRNA (Subtype 1 vs. 2)**	**FC mRNA (Subtype 1 vs. 2)**
FANCI	14	7.1E-03	hsa-miR-21-5p,hsa-miR-370-3p,hsa-miR-140-5p,hsa-miR-16-5p,hsa-miR-10b-5p,hsa-miR-182-5p,hsa-miR-22-3p,hsa-miR-10a-5p,hsa-miR-22-5p,hsa-miR-197-3p,hsa-miR-92b-3p,hsa-miR-92a-3p,hsa-miR-532-3p,hsa-miR-320b	0.9	2.1E-15	0.8	5.0	2.1
CLDN7	12	7.9E-03	hsa-miR-21-5p,hsa-miR-214-5p,hsa-miR-370-3p,hsa-miR-148a-3p,hsa-miR-140-5p,hsa-miR-146a-5p,hsa-miR-10b-5p,hsa-miR-27a-3p,hsa-miR-10a-5p,hsa-miR-125a-5p,hsa-miR-361-3p,hsa-miR-532-3p	0.7	2.2E-09	0.5	5.0	1.8
POU2F2	17	8.4E-03	hsa-miR-214-5p,hsa-miR-370-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-10b-5p,hsa-miR-155-5p,hsa-miR-196b-5p,hsa-miR-182-5p,hsa-miR-143-3p,hsa-miR-27a-3p,hsa-miR-10a-5p,hsa-miR-197-3p,hsa-miR-361-3p,hsa-miR-193b-5p,hsa-miR-320a-3p,hsa-miR-532-3p,hsa-miR-320b	0.7	1.8E-09	0.5	5.0	1.5
HYDIN	12	7.9E-03	hsa-miR-101-3p,hsa-miR-370-3p,hsa-miR-493-3p,hsa-miR-182-5p,hsa-miR-143-3p,hsa-miR-27a-3p,hsa-miR-22-3p,hsa-miR-361-3p,hsa-miR-23b-5p,hsa-miR-193b-5p,hsa-miR-320a-3p,hsa-miR-320b	0.9	3.9E-19	0.9	5.0	5.3
CDS1	19	6.3E-03	hsa-miR-101-3p,hsa-miR-495-3p,hsa-miR-21-5p,hsa-miR-146b-5p,hsa-miR-148a-3p,hsa-miR-146a-5p,hsa-miR-16-5p,hsa-miR-10b-5p,hsa-miR-143-3p,hsa-miR-27a-3p,hsa-miR-152-3p,hsa-miR-22-5p,hsa-miR-197-3p,hsa-miR-145-5p,hsa-miR-664a-3p,hsa-miR-92b-	0.8	2.0E-11	0.9	5.0	2.1

			3p,hsa-miR-92a-3p,hsa-miR-320a-3p,hsa-miR-320b					
PRSS35	10	7.5E-03	hsa-miR-101-3p,hsa-miR-495-3p,hsa-miR-493-3p,hsa-miR-10b-5p,hsa-miR-27a-3p,hsa-miR-10a-5p,hsa-miR-125a-5p,hsa-miR-664a-3p,hsa-miR-532-3p,hsa-miR-320b	0.8	7.7E-12	0.6	5.0	2.9
IGF2BP1	34	2.3E-03	hsa-miR-495-3p,hsa-miR-21-5p,hsa-miR-214-5p,hsa-miR-146b-5p,hsa-miR-370-3p,hsa-miR-148a-3p,hsa-miR-140-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-146a-5p,hsa-miR-16-5p,hsa-miR-10b-5p,hsa-miR-155-5p,hsa-miR-132-5p,hsa-miR-196b-5p,hsa-miR-151a-3p,hsa-miR-30a-5p,hsa-miR-182-5p,hsa-miR-143-3p,hsa-miR-27a-3p,hsa-miR-22-3p,hsa-miR-10a-5p,hsa-miR-152-3p,hsa-miR-22-5p,hsa-miR-197-3p,hsa-miR-125a-5p,hsa-miR-361-3p,hsa-miR-145-5p,hsa-miR-664a-3p,hsa-miR-23b-5p,hsa-miR-320a-3p,hsa-miR-532-3p,hsa-miR-320b,hsa-miR-2110	0.9	1.1E-15	0.9	5.0	5.9
SKIDA1	22	8.8E-04	hsa-miR-101-3p,hsa-miR-495-3p,hsa-miR-21-5p,hsa-miR-370-3p,hsa-miR-148a-3p,hsa-let-7f-5p,hsa-miR-493-3p,hsa-let-7i-5p,hsa-miR-16-5p,hsa-miR-155-5p,hsa-miR-30a-5p,hsa-miR-143-3p,hsa-miR-27a-3p,hsa-miR-10a-5p,hsa-miR-152-3p,hsa-miR-125a-5p,hsa-miR-361-3p,hsa-miR-92b-3p,hsa-miR-92a-3p,hsa-miR-320a-3p,hsa-miR-532-3p,hsa-miR-320b	0.8	2.1E-14	1.0	5.0	2.0
FAM83F	12	1.9E-03	hsa-miR-21-5p,hsa-miR-214-5p,hsa-miR-146a-5p,hsa-miR-132-5p,hsa-miR-30a-5p,hsa-miR-143-3p,hsa-miR-361-3p,hsa-miR-145-5p,hsa-miR-664a-3p,hsa-miR-193b-5p,hsa-miR-92b-3p,hsa-miR-532-3p	0.8	2.7E-14	0.6	5.0	2.8
PKIB	14	2.1E-03	hsa-miR-101-3p,hsa-miR-21-5p,hsa-miR-146b-5p,hsa-miR-148a-3p,hsa-miR-140-5p,hsa-miR-146a-5p,hsa-miR-16-5p,hsa-miR-10b-5p,hsa-miR-155-5p,hsa-miR-182-5p,hsa-miR-27a-	0.8	1.8E-11	0.9	5.0	1.9

			3p,hsa-miR-664a-3p,hsa-miR-92b-3p,hsa-miR-92a-3p					
IGF2BP3	27	4.0E-04	hsa-miR-101-3p,hsa-miR-495-3p,hsa-miR-21-5p,hsa-miR-214-5p,hsa-miR-146b-5p,hsa-miR-148a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-146a-5p,hsa-miR-16-5p,hsa-miR-10b-5p,hsa-miR-155-5p,hsa-miR-196b-5p,hsa-miR-30a-5p,hsa-miR-182-5p,hsa-miR-27a-3p,hsa-miR-22-3p,hsa-miR-197-3p,hsa-miR-125a-5p,hsa-miR-361-3p,hsa-miR-664a-3p,hsa-miR-193b-5p,hsa-miR-92b-3p,hsa-miR-92a-3p,hsa-miR-320a-3p,hsa-miR-320b,hsa-miR-2110	0.9	2.3E-19	1.0	5.0	4.8
IL36G	8	5.1E-03	hsa-miR-101-3p,hsa-miR-146a-5p,hsa-miR-10b-5p,hsa-miR-155-5p,hsa-miR-182-5p,hsa-miR-143-3p,hsa-miR-27a-3p,hsa-miR-22-5p	0.8	2.5E-13	0.8	5.0	10.2
CLDN7	12	7.9E-03	hsa-miR-21-5p,hsa-miR-214-5p,hsa-miR-370-3p,hsa-miR-148a-3p,hsa-miR-140-5p,hsa-miR-146a-5p,hsa-miR-10b-5p,hsa-miR-27a-3p,hsa-miR-10a-5p,hsa-miR-125a-5p,hsa-miR-361-3p,hsa-miR-532-3p	0.7	2.2E-09	0.5	5.0	1.8
STYK1	8	5.1E-03	hsa-miR-101-3p,hsa-miR-495-3p,hsa-miR-21-5p,hsa-miR-16-5p,hsa-miR-155-5p,hsa-miR-27a-3p,hsa-miR-22-3p,hsa-miR-22-5p	0.7	1.9E-09	1.0	5.0	2.6
HYDIN	12	7.9E-03	hsa-miR-101-3p,hsa-miR-370-3p,hsa-miR-493-3p,hsa-miR-182-5p,hsa-miR-143-3p,hsa-miR-27a-3p,hsa-miR-22-3p,hsa-miR-361-3p,hsa-miR-23b-5p,hsa-miR-193b-5p,hsa-miR-320a-3p,hsa-miR-320b	0.9	3.9E-19	0.9	5.0	5.3
INA	12	7.9E-03	hsa-miR-101-3p,hsa-miR-495-3p,hsa-miR-146b-5p,hsa-miR-493-3p,hsa-miR-146a-5p,hsa-miR-16-5p,hsa-miR-155-5p,hsa-miR-182-5p,hsa-miR-27a-3p,hsa-miR-197-3p,hsa-miR-320a-3p,hsa-miR-320b	0.8	2.5E-13	0.9	5.0	4.6
CABP4	8	5.1E-03	hsa-miR-146a-5p,hsa-miR-10b-5p,hsa-miR-182-5p,hsa-miR-27a-3p,hsa-miR-10a-5p,hsa-miR-152-3p,hsa-miR-145-5p,hsa-miR-23b-5p	0.7	1.3E-09	0.5	5.0	1.8

MUC13	10	7.5E-03	hsa-miR-495-3p,hsa-miR-493-3p,hsa-miR-143-3p,hsa-miR-27a-3p,hsa-miR-197-3p,hsa-miR-361-3p,hsa-miR-92b-3p,hsa-miR-92a-3p,hsa-miR-320a-3p,hsa-miR-320b	0.9	4.7E-16	1.0	5.0	4.4

ceRNA results for n=6 hub lncRNAs are presented for males.

Table A.13 ceRNA analysis results for n=6 hub lncRNAs in females

AC016717.2 hub lncRNA in Females								
mRNA	**miR binding sites**	**Hyper-geo metric P-value**	**Interacting miRNAs**	**lncRNA -mRNA Corr R-value**	**lncRNA -mRNA Corr P-value**	**Regulation Similarity Score (RegSim)**	**FC lncRNA (Subtype 1 vs. 2)**	**FC mRNA (Subtype 1 vs. 2)**
CENPK	7	5.2E-03	hsa-miR-27b-3p,hsa-miR-30e-5p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-495-3p	0.7	1E-07	1.0	2.7	2.7
MACC1	8	1.5E-03	hsa-miR-145-5p,hsa-miR-532-3p,hsa-miR-27b-3p,hsa-miR-143-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7i-5p,hsa-miR-155-5p	0.7	4E-07	1.0	2.7	5.0
CRB1	8	4.5E-03	hsa-miR-145-5p,hsa-miR-27b-3p,hsa-miR-30e-5p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-155-5p	0.7	5E-07	1.0	2.7	2.4
FANCD 2	7	5.2E-03	hsa-miR-27b-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-155-5p,hsa-miR-148a-3p	0.7	8E-07	1.0	2.7	1.9
IPCEF1	8	1.5E-03	hsa-miR-145-5p,hsa-miR-532-3p,hsa-miR-27b-3p,hsa-miR-27a-3p,hsa-miR-155-5p,hsa-miR-148a-3p,hsa-miR-21-5p,hsa-miR-495-3p	0.7	5E-07	0.7	2.7	2.4

CASC19 hub lncRNA in Females								
mRNA	**miR binding sites**	**Hyper-geo metric P-value**	**Interacting miRNAs**	**lncRNA -mRNA Corr R-value**	**lncRNA -mRNA Corr P-value**	**Regulation Similarity Score (RegSim)**	**FC lncRNA (Subtype 1 vs. 2)**	**FC mRNA (Subtype 1 vs. 2)**
RPH3A	7	1.4E-03	hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-493-5p,hsa-miR-98-5p,hsa-miR-148a-3p	0.8	2E-10	1.0	5.4	3.4
DMP1	5	3.5E-03	hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.8	5E-10	1.0	5.4	3.8
FAM133A	8	4.5E-03	hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-140-5p,hsa-miR-148a-3p,hsa-miR-21-5p,hsa-miR-499a-5p	0.8	9E-11	1.0	5.4	5.2
GRIN3A	7	5.2E-03	hsa-miR-27b-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-493-5p,hsa-miR-98-5p	0.9	5E-12	0.9	5.4	2.9
ITIH1	5	3.5E-03	hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-21-5p	0.8	1E-08	1.0	5.4	6.2
WNK3	11	1.9E-04	hsa-miR-27b-3p,hsa-miR-151a-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-493-5p,hsa-miR-98-5p,hsa-miR-543,hsa-miR-21-5p,hsa-miR-499a-5p	0.7	1E-07	0.7	5.4	2.0
KCNC2	8	4.5E-03	hsa-miR-27b-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-543,hsa-miR-499a-5p	0.7	1E-07	1.0	5.4	2.9
C11orf87	7	1.4E-03	hsa-miR-27b-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-493-5p,hsa-miR-98-5p	0.8	2E-09	1.0	5.4	3.4
GABRA6	8	3.6E-04	hsa-miR-22-5p,hsa-miR-151a-3p,hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-493-5p,hsa-miR-98-5p,hsa-miR-148a-3p	0.9	4E-14	1.0	5.4	4.6
EHF	9	3.7E-04	hsa-miR-27b-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-493-5p,hsa-miR-98-5p,hsa-miR-140-5p,hsa-miR-543	0.9	2E-11	1.0	5.4	7.8

FAXC	9	3.4E-03	hsa-miR-27b-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-493-5p,hsa-miR-98-5p,hsa-miR-543,hsa-miR-499a-5p	0.8	4E-11	0.9	5.4	3.4
MEI4	5	3.5E-03	hsa-miR-151a-3p,hsa-miR-493-5p,hsa-miR-543,hsa-miR-21-5p,hsa-miR-499a-5p	0.8	5E-11	1.0	5.4	4.9
SALL3	7	2.2E-04	hsa-miR-27b-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-21-5p	0.8	2E-09	1.0	5.4	3.9
ARHGEF38	7	1.4E-03	hsa-miR-27b-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-493-5p,hsa-miR-98-5p	0.9	2E-17	1.0	5.4	5.6
ONECUT2	13	9.6E-03	hsa-miR-22-5p,hsa-miR-27b-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-493-5p,hsa-miR-98-5p,hsa-miR-140-5p,hsa-miR-148a-3p,hsa-miR-543,hsa-miR-21-5p,hsa-miR-499a-5p	0.8	3E-11	0.9	5.4	4.5
TFAP2B	6	9.2E-04	hsa-miR-27b-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.9	4E-13	1.0	5.4	6.5
PARPBP	8	4.5E-03	hsa-miR-27b-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-148a-3p,hsa-miR-499a-5p	0.7	3E-07	0.9	5.4	1.6
GRM5	9	1.3E-03	hsa-miR-22-5p,hsa-miR-27b-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-493-5p,hsa-miR-98-5p,hsa-miR-543	0.9	2E-12	1.0	5.4	3.6
RSPO2	7	5.2E-03	hsa-miR-27b-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-493-5p,hsa-miR-98-5p	0.7	5E-07	0.9	5.4	3.0
BEND4	9	3.4E-03	hsa-miR-27b-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-140-5p,hsa-miR-148a-3p,hsa-miR-543	0.9	3E-12	1.0	5.4	3.8
FANCD2	7	5.2E-03	hsa-miR-27b-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-148a-3p	0.8	5E-10	1.0	5.4	1.9

NEXMIF	10	5.2E-03	hsa-miR-27b-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-493-5p,hsa-miR-98-5p,hsa-miR-148a-3p,hsa-miR-543,hsa-miR-21-5p	0.8	8E-11	1.0	5.4	2.3
TDO2	6	5.1E-03	hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-493-5p,hsa-miR-98-5p,hsa-miR-140-5p	0.7	7E-07	0.8	5.4	3.1
KCNJ10	8	4.5E-03	hsa-miR-27b-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-493-5p,hsa-miR-98-5p,hsa-miR-21-5p	0.8	7E-11	1.0	5.4	4.7
AP3B2	6	5.1E-03	hsa-miR-27b-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.8	1E-08	0.5	5.4	2.8
NR1I2	6	9.2E-04	hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-148a-3p	0.8	1E-10	0.8	5.4	4.6
NDST3	8	1.5E-03	hsa-miR-27b-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-493-5p,hsa-miR-98-5p,hsa-miR-140-5p	0.9	7E-12	1.0	5.4	4.6
XKR7	5	3.5E-03	hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.8	2E-09	1.0	5.4	4.0
GALNT13	6	5.1E-03	hsa-miR-27b-3p,hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-543	0.8	4E-09	1.0	5.4	2.2
TBX5	6	5.1E-03	hsa-miR-27b-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.8	1E-08	1.0	5.4	2.8
NTRK3	9	3.7E-04	hsa-miR-22-5p,hsa-miR-151a-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-148a-3p,hsa-miR-543	0.7	6E-07	0.9	5.4	1.5
SPAG1	9	1.3E-03	hsa-miR-22-5p,hsa-miR-27b-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-140-5p,hsa-miR-543	0.8	6E-08	1.0	5.4	1.7
MAB21L3	6	5.1E-03	hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-148a-3p,hsa-miR-543	0.8	5E-08	0.7	5.4	3.9

DLX6-AS1 hub lncRNA in Females								
mRNA	**miR binding sites**	**Hyper-geometric P-value**	**Interacting miRNAs**	**lncRNA-mRNA Corr R-value**	**lncRNA-mRNA Corr P-value**	**Regulation Similarity Score (RegSim)**	**FC lncRNA (Subtype 1 vs. 2)**	**FC mRNA (Subtype 1 vs. 2)**
CHEK1	8	2.4E-03	hsa-miR-145-5p,hsa-miR-143-3p,hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-199b-5p,hsa-miR-146b-5p	0.7	5E-07	0.8	5.2	1.7
SALL4	7	9.2E-03	hsa-miR-27b-3p,hsa-miR-143-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.7	8E-08	1.0	5.2	3.1
CDH1	10	1.2E-03	hsa-miR-145-5p,hsa-miR-22-3p,hsa-miR-143-3p,hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-199b-5p,hsa-miR-140-5p,hsa-miR-146b-5p	0.7	1E-07	0.8	5.2	3.7
EPS8L3	5	1.4E-03	hsa-miR-145-5p,hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.9	1E-11	1.0	5.2	6.5
FANCD2	7	1.4E-03	hsa-miR-27b-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-199b-5p	0.8	2E-10	1.0	5.2	1.9
SALL3	6	1.7E-03	hsa-miR-27b-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.8	6E-10	1.0	5.2	3.9
SLC22A10	5	1.4E-03	hsa-miR-145-5p,hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.9	5E-15	1.0	5.2	5.8
BEND4	8	5.4E-03	hsa-miR-27b-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-140-5p,hsa-miR-146b-5p	0.9	1E-14	1.0	5.2	3.8
LECT2	4	6.0E-03	hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.9	7E-13	1.0	5.2	6.6
HOOK1	10	5.4E-04	hsa-miR-145-5p,hsa-miR-22-3p,hsa-miR-27b-3p,hsa-miR-143-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-199b-5p	0.9	3E-11	1.0	5.2	2.9
ZNF311	5	6.8E-03	hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-199b-5p	0.8	1E-08	1.0	5.2	2.6

COL9A1	5	6.8E-03	hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-146b-5p	0.9	6E-18	1.0	5.2	4.3
TNIP3	5	6.8E-03	hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.8	8E-10	1.0	5.2	4.1
ONECUT3	5	1.4E-03	hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-146b-5p	0.9	1E-11	0.8	5.2	13.3
DMP1	5	1.4E-03	hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.9	7E-12	1.0	5.2	3.8
OTOF	5	1.4E-03	hsa-miR-143-3p,hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.8	5E-09	1.0	5.2	2.3
ESPL1	8	2.8E-04	hsa-miR-145-5p,hsa-miR-22-3p,hsa-miR-27b-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.7	1E-07	1.0	5.2	2.8
IL22RA1	4	6.0E-03	hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.8	3E-08	1.0	5.2	2.9
EHF	7	9.2E-03	hsa-miR-27b-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-140-5p	0.9	2E-13	1.0	5.2	7.8
DSC3	8	9.1E-04	hsa-miR-22-3p,hsa-miR-27b-3p,hsa-miR-143-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.8	5E-10	1.0	5.2	3.3
LIN28B	11	5.9E-03	hsa-miR-145-5p,hsa-miR-22-3p,hsa-miR-27b-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-199b-5p,hsa-miR-140-5p,hsa-miR-146b-5p	0.9	7E-16	1.0	5.2	8.0
IGF2BP1	12	7.8E-03	hsa-miR-145-5p,hsa-miR-22-3p,hsa-miR-27b-3p,hsa-miR-143-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-199b-5p,hsa-miR-140-5p,hsa-miR-146b-5p	0.9	3E-15	0.9	5.2	5.0
TBX5	6	1.7E-03	hsa-miR-27b-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.8	1E-09	1.0	5.2	2.8
CYP2C8	4	6.0E-03	hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.7	7E-08	1.0	5.2	2.3

PTCHD1	8	2.8E-04	hsa-miR-27b-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-199b-5p,hsa-miR-140-5p	0.8	1E-09	1.0	5.2	1.7
STAB2	7	3.6E-04	hsa-miR-27b-3p,hsa-miR-143-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.8	5E-08	1.0	5.2	2.4
TTLL9	5	6.8E-03	hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-140-5p	0.9	2E-11	0.8	5.2	3.1
GRIN2B	10	5.4E-04	hsa-miR-145-5p,hsa-miR-22-3p,hsa-miR-27b-3p,hsa-miR-143-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-199b-5p	0.9	4E-18	1.0	5.2	3.5
HTR3A	4	6.0E-03	hsa-miR-22-3p,hsa-miR-143-3p,hsa-miR-27a-3p,hsa-miR-140-5p	0.8	1E-08	0.7	5.2	6.5
CLDN16	5	6.8E-03	hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.9	2E-11	1.0	5.2	4.2
C11orf87	6	5.6E-03	hsa-miR-27b-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.8	9E-10	1.0	5.2	3.4
TFAP2B	6	2.8E-04	hsa-miR-27b-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.9	8E-16	1.0	5.2	6.5
ONECUT2	12	2.5E-03	hsa-miR-145-5p,hsa-miR-22-3p,hsa-miR-27b-3p,hsa-miR-143-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-199b-5p,hsa-miR-140-5p,hsa-miR-146b-5p	0.9	2E-12	0.9	5.2	4.5
GPR85	5	6.8E-03	hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.8	5E-10	1.0	5.2	2.1
RAG1	6	5.6E-03	hsa-miR-27b-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.7	8E-08	1.0	5.2	1.7
IGDCC3	8	2.4E-03	hsa-miR-145-5p,hsa-miR-22-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-199b-5p	0.8	4E-08	0.6	5.2	2.1
TMPRSS2	6	5.6E-03	hsa-miR-27b-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.9	9E-14	1.0	5.2	4.9
MUC4	6	1.7E-03	hsa-miR-145-5p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.8	3E-09	1.0	5.2	4.1

CDC25A	11	4.0E-04	hsa-miR-145-5p,hsa-miR-27b-3p,hsa-miR-143-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-199b-5p,hsa-miR-140-5p,hsa-miR-146b-5p	0.9	6E-12	0.7	5.2	1.8
OR2AG1	4	6.0E-03	hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.9	3E-13	1.0	5.2	4.7
TCEAL2	4	6.0E-03	hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.8	5E-11	1.0	5.2	3.7
KCNJ10	7	9.2E-03	hsa-miR-22-3p,hsa-miR-27b-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.9	2E-11	1.0	5.2	4.7
HS6ST3	8	5.4E-03	hsa-miR-145-5p,hsa-miR-22-3p,hsa-miR-27b-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-199b-5p	0.8	7E-11	0.8	5.2	2.8
HMMR	7	3.9E-03	hsa-miR-27b-3p,hsa-miR-143-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.8	7E-10	1.0	5.2	1.7
HMGA2	12	2.5E-03	hsa-miR-145-5p,hsa-miR-22-3p,hsa-miR-27b-3p,hsa-miR-143-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-199b-5p,hsa-miR-140-5p,hsa-miR-146b-5p	0.9	7E-18	1.0	5.2	4.4
ZNF215	5	6.8E-03	hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-199b-5p	0.8	5E-09	1.0	5.2	2.0
DMKN	8	2.8E-04	hsa-miR-27b-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-199b-5p,hsa-miR-146b-5p	0.8	2E-08	0.9	5.2	2.5
ICOS	6	1.7E-03	hsa-miR-27b-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.7	4E-07	1.0	5.2	2.1
SOX11	11	5.9E-03	hsa-miR-145-5p,hsa-miR-22-3p,hsa-miR-27b-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-199b-5p,hsa-miR-140-5p,hsa-miR-146b-5p	0.7	6E-07	0.7	5.2	2.0
MYH15	4	6.0E-03	hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.8	2E-09	1.0	5.2	2.6

DNAH9	4	6.0E-03	hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.9	2E-17	1.0	5.2	4.5
FAM81A	9	1.3E-03	hsa-miR-145-5p,hsa-miR-27b-3p,hsa-miR-143-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-146b-5p	0.9	8E-12	0.9	5.2	2.0
MACC1	7	3.9E-03	hsa-miR-145-5p,hsa-miR-27b-3p,hsa-miR-143-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7i-5p,hsa-miR-199b-5p	0.8	6E-09	1.0	5.2	5.0
USP44	9	1.3E-03	hsa-miR-22-3p,hsa-miR-27b-3p,hsa-miR-143-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-199b-5p	0.8	3E-08	1.0	5.2	2.0
PALM3	4	6.0E-03	hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.9	2E-11	1.0	5.2	5.9
XKR7	5	1.4E-03	hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.9	2E-13	1.0	5.2	4.0
IQGAP3	7	3.9E-03	hsa-miR-22-3p,hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-140-5p,hsa-miR-146b-5p	0.7	7E-07	0.9	5.2	2.4
NR1I2	5	6.8E-03	hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.8	5E-11	1.0	5.2	4.6
NIPAL4	8	2.4E-03	hsa-miR-145-5p,hsa-miR-27b-3p,hsa-miR-143-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.8	3E-08	1.0	5.2	3.9
PLEKHG7	6	2.8E-04	hsa-miR-145-5p,hsa-miR-143-3p,hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.9	4E-14	1.0	5.2	6.0
LRGUK	5	1.4E-03	hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.7	6E-07	1.0	5.2	1.9
CRB1	8	9.1E-04	hsa-miR-145-5p,hsa-miR-27b-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-199b-5p	0.9	2E-13	1.0	5.2	2.4
NDST3	8	2.8E-04	hsa-miR-145-5p,hsa-miR-27b-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-140-5p	0.9	2E-12	1.0	5.2	4.6

FAM189A1	5	1.4E-03	hsa-miR-22-3p,hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.7	4E-07	0.9	5.2	2.2
C4orf50	4	6.0E-03	hsa-miR-145-5p,hsa-miR-22-3p,hsa-miR-27b-3p,hsa-miR-27a-3p	0.9	2E-15	1.0	5.2	5.7
SLC17A2	4	6.0E-03	hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.9	2E-16	1.0	5.2	6.1
CCDC15	5	1.4E-03	hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-199b-5p	0.7	3E-07	1.0	5.2	1.6
ARHGEF38	6	5.6E-03	hsa-miR-27b-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.9	4E-17	1.0	5.2	5.6
ARHGEF39	7	3.9E-03	hsa-miR-145-5p,hsa-miR-22-3p,hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-199b-5p	0.7	4E-07	0.7	5.2	1.8
GRHL2	10	8.9E-03	hsa-miR-22-3p,hsa-miR-27b-3p,hsa-miR-143-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-199b-5p,hsa-miR-146b-5p	0.9	4E-17	0.9	5.2	6.9
BICDL1	6	5.6E-03	hsa-miR-145-5p,hsa-miR-22-3p,hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.8	4E-08	1.0	5.2	2.3
ETNPPL	4	6.0E-03	hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.8	2E-08	0.9	5.2	2.3
KCNC2	7	9.2E-03	hsa-miR-27b-3p,hsa-miR-143-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.8	4E-08	1.0	5.2	2.9
MGAT5B	5	6.8E-03	hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-146b-5p	0.8	2E-09	0.8	5.2	2.4
GRIN3A	7	1.4E-03	hsa-miR-27b-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-146b-5p	0.9	2E-13	0.9	5.2	2.9
OR2AG1	4	6.0E-03	hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.9	3E-13	1.0	5.2	4.7

FIRRE hub lncRNA in Females								
mRNA	**miR binding sites**	**Hyper-geometric P-value**	**Interacting miRNAs**	**lncRNA-mRNA Corr R-value**	**lncRNA-mRNA Corr P-value**	**Regulation Similarity Score (RegSim)**	**FC lncRNA (Subtype 1 vs. 2)**	**FC mRNA (Subtype 1 vs. 2)**
LECT2	4	4.0E-03	hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.7	1E-06	1.0	2.0	6.6
FANCD2	7	6.1E-04	hsa-miR-27b-3p,hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-199b-5p,hsa-miR-148a-3p	0.8	2E-08	0.9	2.0	1.9
STAG3	6	1.4E-04	hsa-miR-3615,hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-148a-3p	0.7	8E-08	1.0	2.0	3.1
SLC17A2	4	4.0E-03	hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.7	4E-07	1.0	2.0	6.1
KRT5	5	4.1E-03	hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-21-5p	0.7	3E-07	1.0	2.0	5.2
CACNA1I	5	4.1E-03	hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-370-3p	0.7	4E-07	0.9	2.0	2.3
CRTAM	5	8.0E-04	hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-21-5p	0.7	8E-08	0.6	2.0	2.1
TFAP2B	5	4.1E-03	hsa-miR-27b-3p,hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.7	6E-07	1.0	2.0	6.5
SALL3	6	8.6E-04	hsa-miR-27b-3p,hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-21-5p	0.7	8E-07	1.0	2.0	3.9
MAST1	5	8.0E-04	hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-379-5p	0.8	4E-09	1.0	2.0	3.4
DNAH9	4	4.0E-03	hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.8	5E-08	1.0	2.0	4.5
PALM3	4	4.0E-03	hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.7	7E-07	1.0	2.0	5.9
IL22RA1	4	4.0E-03	hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.7	8E-07	1.0	2.0	2.9

CNTNAP2	8	9.1E-04	hsa-miR-27b-3p,hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-199b-5p,hsa-miR-148a-3p,hsa-miR-370-3p	0.8	6E-08	0.7	2.0	5.0
CGN	8	4.5E-03	hsa-miR-27b-3p,hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-21-5p,hsa-miR-379-5p,hsa-miR-370-3p	0.7	4E-07	0.7	2.0	3.6

LINC00943 hub lncRNA in Females								
mRNA	**miR binding sites**	**Hyper-geo metric P-value**	**Interacting miRNAs**	**lncRNA-mRNA Corr R-value**	**lncRNA-mRNA Corr P-value**	**Regulation Similarity Score (RegSim)**	**FC lncRNA (Subtype 1 vs. 2)**	**FC mRNA (Subtype 1 vs. 2)**
TRIM71	17	8.6E-03	hsa-miR-145-5p,hsa-miR-27b-3p,hsa-miR-143-3p,hsa-miR-30e-5p,hsa-miR-151a-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-155-5p,hsa-miR-98-5p,hsa-miR-199b-5p,hsa-miR-140-5p,hsa-miR-214-5p,hsa-miR-543,hsa-miR-21-5p,hsa-miR-379-5p	0.8	1E-08	0.9	2.9	5.7
CDC25A	17	5.6E-04	hsa-miR-18a-3p,hsa-miR-145-5p,hsa-miR-22-5p,hsa-miR-27b-3p,hsa-miR-143-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-155-5p,hsa-miR-98-5p,hsa-miR-199b-5p,hsa-miR-140-5p,hsa-miR-511-5p,hsa-miR-543,hsa-miR-21-5p,hsa-miR-146b-5p	0.8	1E-09	0.6	2.9	1.8
FASLG	9	6.2E-03	hsa-miR-22-5p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-214-5p,hsa-miR-543,hsa-miR-21-5p	0.7	9E-08	0.9	2.9	3.2
SLC6A15	18	4.0E-03	hsa-miR-18a-3p,hsa-miR-27b-3p,hsa-miR-30e-5p,hsa-miR-151a-3p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-155-5p,hsa-miR-98-5p,hsa-miR-199b-5p,hsa-miR-140-5p,hsa-miR-511-5p,hsa-miR-214-5p,hsa-	0.8	6E-09	1.0	2.9	3.4

			miR-543,hsa-miR-370-3p,hsa-miR-146b-5p,hsa-miR-493-3p					
CRB1	11	1.5E-03	hsa-miR-145-5p,hsa-miR-27b-3p,hsa-miR-30e-5p,hsa-let-7g-5p,hsa-miR-27a-3p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-155-5p,hsa-miR-98-5p,hsa-miR-199b-5p,hsa-miR-543	0.8	2E-10	1.0	2.9	2.4

HELLPAR hub lncRNA in Females								
mRNA	**miR binding sites**	**Hyper-geo metric P-value**	**Interacting miRNAs**	**lncRNA -mRNA Corr R-value**	**lncRNA -mRNA Corr P-value**	**Regulation Similarity Score (RegSim)**	**FC lncRNA (Subtype 1 vs. 2)**	**FC mRNA (Subtype 1 vs. 2)**
TAT	6	5.6E-03	hsa-miR-145-5p,hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-21-5p	0.8	6E-09	1.0	6.1	5.2
ETNPPL	4	6.0E-03	hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.7	6E-07	1.0	6.1	2.3
WNK3	9	2.8E-03	hsa-miR-30e-5p,hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-493-5p,hsa-miR-98-5p,hsa-miR-543,hsa-miR-21-5p,hsa-miR-499a-5p	0.8	3E-08	0.7	6.1	2.0
FAM189A1	5	1.4E-03	hsa-miR-22-3p,hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.7	3E-07	1.0	6.1	2.2
HTR4	5	1.4E-03	hsa-miR-30e-5p,hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.8	6E-09	1.0	6.1	2.7
LHCGR	4	6.0E-03	hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.7	6E-07	0.8	6.1	1.9
PALM3	4	6.0E-03	hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.9	2E-12	1.0	6.1	5.9
NTRK3	7	9.2E-03	hsa-miR-22-3p,hsa-miR-99a-5p,hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-543	0.8	5E-08	1.0	6.1	1.5

KRT5	5	6.8E-03	hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-21-5p	0.8	1E-10	0.9	6.1	5.2
BICDL1	6	5.6E-03	hsa-miR-145-5p,hsa-miR-22-3p,hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.8	2E-08	1.0	6.1	2.3
SPTSSB	10	5.4E-04	hsa-miR-145-5p,hsa-miR-22-3p,hsa-miR-30e-5p,hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-543,hsa-miR-21-5p,hsa-miR-499a-5p	0.7	2E-07	1.0	6.1	2.8
GRM5	8	2.4E-03	hsa-miR-22-3p,hsa-miR-30e-5p,hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-493-5p,hsa-miR-98-5p,hsa-miR-543	0.9	3E-11	1.0	6.1	3.6
TCEAL2	4	6.0E-03	hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.8	5E-09	1.0	6.1	3.7
LAMC3	5	6.8E-03	hsa-miR-30e-5p,hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.8	6E-09	1.0	6.1	3.1
SLITRK5	7	3.6E-04	hsa-miR-99a-5p,hsa-miR-30e-5p,hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-493-5p,hsa-miR-98-5p	0.7	4E-07	1.0	6.1	1.7
DNAH9	4	6.0E-03	hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.9	8E-15	1.0	6.1	4.5
MYH15	4	6.0E-03	hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.8	1E-08	1.0	6.1	2.6
ZNF695	9	1.3E-03	hsa-miR-145-5p,hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-493-5p,hsa-miR-98-5p,hsa-miR-543,hsa-miR-21-5p,hsa-miR-499a-5p	0.7	7E-08	0.9	6.1	5.6
KCNJ10	7	9.2E-03	hsa-miR-22-3p,hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-493-5p,hsa-miR-98-5p,hsa-miR-21-5p	0.9	1E-12	1.0	6.1	4.7
FAM133A	7	9.2E-03	hsa-miR-30e-5p,hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-21-5p,hsa-miR-499a-5p	0.8	9E-10	1.0	6.1	5.2

EPS8L3	5	1.4E-03	hsa-miR-145-5p,hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.8	1E-10	1.0	6.1	6.5
OR2AG1	4	6.0E-03	hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.9	2E-12	1.0	6.1	4.7
RFX6	6	5.6E-03	hsa-miR-30e-5p,hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-543	0.9	9E-12	1.0	6.1	5.6
SLC17A2	4	6.0E-03	hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.9	1E-14	1.0	6.1	6.1
CRTAM	5	1.4E-03	hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-21-5p	0.7	2E-07	0.6	6.1	2.1
GALNT13	6	1.7E-03	hsa-miR-30e-5p,hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-543	0.8	3E-09	1.0	6.1	2.2
LECT2	4	6.0E-03	hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.9	4E-12	1.0	6.1	6.6
OMG	6	5.6E-03	hsa-miR-30e-5p,hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-499a-5p	0.8	2E-08	1.0	6.1	2.3
SLC22A10	5	1.4E-03	hsa-miR-145-5p,hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.9	7E-13	1.0	6.1	5.8
PLEKHG7	5	6.8E-03	hsa-miR-145-5p,hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.9	2E-12	1.0	6.1	6.0
ITIH1	5	1.4E-03	hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-21-5p	0.8	2E-09	1.0	6.1	6.2
CRB1	7	9.2E-03	hsa-miR-145-5p,hsa-miR-30e-5p,hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p,hsa-miR-543	0.8	1E-10	1.0	6.1	2.4
CYP4F2	5	1.4E-03	hsa-miR-99a-5p,hsa-let-7g-5p,hsa-let-7f-5p,hsa-let-7i-5p,hsa-miR-98-5p	0.8	4E-11	1.0	6.1	5.5

ceRNA results for n=6 hub lncRNAs are presented for females.

Table A.14 Common pathways in males and females from ceRNA-mediated gene regulation

N=44 common pathways between Male and Female ceRNA-mediated gene regulation		
Pathway	**Male -log(P-value)**	**Female -log(P-value)**
Acetone Degradation I (to Methylglyoxal)	2.8	2.2
Amyotrophic Lateral Sclerosis Signaling	6.4	3.1
Breast Cancer Regulation by Stathmin1	5.6	2.3
Bupropion Degradation	3.8	3.1
Calcium Signaling	5.8	3.4
cAMP-mediated signaling	2.2	3.1
CREB Signaling in Neurons	12.1	4.7
Dopamine-DARPP32 Feedback in cAMP Signaling	2.7	3.8
Embryonic Stem Cell Differentiation into Cardiac Lineages	2.1	1.5
Endocannabinoid Neuronal Synapse Pathway	4.6	4.9
Estrogen Biosynthesis	2.9	2.0
Estrogen-mediated S-phase Entry	2.3	2.1
FXR/RXR Activation	10.1	4.6
GABA Receptor Signaling	8.8	6.2
Glutamate Receptor Signaling	11.0	10.1
G-Protein Coupled Receptor Signaling	5.5	2.8
Gustation Pathway	9.5	6.6
Hepatic Cholestasis	4.0	1.9
Kinetochore Metaphase Signaling Pathway	5.7	2.2
LPS/IL-1 Mediated Inhibition of RXR Function	5.5	1.3
Maturity Onset Diabetes of Young (MODY) Signaling	5.1	1.5
Melatonin Degradation I	4.3	3.3
Netrin Signaling	4.0	2.8
Neuroinflammation Signaling Pathway	3.4	2.4
Neuropathic Pain Signaling In Dorsal Horn Neurons	5.4	6.4
Neurovascular Coupling Signaling Pathway	14.9	10.3
Nicotine Degradation II	3.4	3.1
Nicotine Degradation III	4.1	3.5
nNOS Signaling in Neurons	2.7	1.9
nNOS Signaling in Skeletal Muscle Cells	4.4	2.5
Opioid Signaling Pathway	2.7	2.6
Oxytocin In Brain Signaling Pathway	2.8	1.8
Phagosome Formation	4.7	1.3
PXR/RXR Activation	6.5	3.9
Role of BRCA1 in DNA Damage Response	1.7	2.0
Serotonin Receptor Signaling	2.4	1.3
Superpathway of Melatonin Degradation	3.8	3.0
Synaptic Long Term Depression	5.8	3.5
Synaptic Long Term Potentiation	2.1	3.8

Synaptogenesis Signaling Pathway	6.3	6.3
Transcriptional Regulatory Network in Embryonic Stem Cells	2.7	1.6
Xenobiotic Metabolism CAR Signaling Pathway	3.7	1.9
Xenobiotic Metabolism PXR Signaling Pathway	3.2	1.8
Xenobiotic Metabolism Signaling	2.2	1.5

Table A.15 Unique pathways in males from ceRNA-mediated gene regulation

N=45 unique pathways in Males from ceRNA-mediated gene regulation	
Pathway	**-log(P-value)**
Acute Phase Response Signaling	2.9
Airway Pathology in Chronic Obstructive Pulmonary Disease	1.6
Basal Cell Carcinoma Signaling	1.7
Chondroitin Sulfate Biosynthesis	2.5
Chondroitin Sulfate Biosynthesis (Late Stages)	2.0
Citrulline Biosynthesis	1.4
Coagulation System	2.6
Corticotropin Releasing Hormone Signaling	1.7
Dermatan Sulfate Biosynthesis	1.9
Dermatan Sulfate Biosynthesis (Late Stages)	1.7
Differential Regulation of Cytokine Production in Intestinal Epithelial Cells by IL-17A and IL-17F	1.4
eNOS Signaling	1.5
Extrinsic Prothrombin Activation Pathway	2.8
Fatty Acid Activation	1.5
FcγRIIB Signaling in B Lymphocytes	1.9
G Beta Gamma Signaling	2.1
Glutamate Dependent Acid Resistance	2.1
GPCR-Mediated Nutrient Sensing in Enteroendocrine Cells	2.2
Granulocyte Adhesion and Diapedesis	1.9
Heparan Sulfate Biosynthesis	1.5
Histidine Degradation III	1.6
HOTAIR Regulatory Pathway	1.6
Human Embryonic Stem Cell Pluripotency	1.8
Inhibition of Matrix Metalloproteases	1.3
Insulin Secretion Signaling Pathway	1.9
Intrinsic Prothrombin Activation Pathway	6.0
LXR/RXR Activation	3.9
Mitochondrial L-carnitine Shuttle Pathway	1.7
Mitotic Roles of Polo-Like Kinase	2.4
MSP-RON Signaling Pathway	2.5
Neuroprotective Role of THOP1 in Alzheimer's Disease	2.8
Nitric Oxide Signaling in the Cardiovascular System	1.9

Phototransduction Pathway	2.2
Retinol Biosynthesis	1.4
Role of Cytokines in Mediating Communication between Immune Cells	2.7
Role of Osteoblasts, Osteoclasts and Chondrocytes in Rheumatoid Arthritis	1.3
Serotonin and Melatonin Biosynthesis	2.0
SNARE Signaling Pathway	1.6
SPINK1 Pancreatic Cancer Pathway	2.3
Superpathway of Citrulline Metabolism	2.2
Thyroid Hormone Metabolism II (via Conjugation and/or Degradation)	1.7
Type II Diabetes Mellitus Signaling	2.5
VDR/RXR Activation	2.2
Wound Healing Signaling Pathway	1.3
α-tocopherol Degradation	1.3

Table A.16 Unique pathways in females from ceRNA-mediated gene regulation

N=45 unique pathways in Females from ceRNA-mediated gene regulation	
Pathway	**-log(P-value)**
4-hydroxybenzoate Biosynthesis	1.5
4-hydroxyphenylpyruvate Biosynthesis	1.5
Cellular Effects of Sildenafil (Viagra)	1.8
GP6 Signaling Pathway	1.4
Gαs Signaling	1.9
Role of CHK Proteins in Cell Cycle Checkpoint Control	1.5
Role of OCT4 in Mammalian Embryonic Stem Cell Pluripotency	1.3
Serotonin Degradation	1.7
Tryptophan Degradation to 2-amino-3-carboxymuconate Semialdehyde	1.9
Urea Cycle	1.9

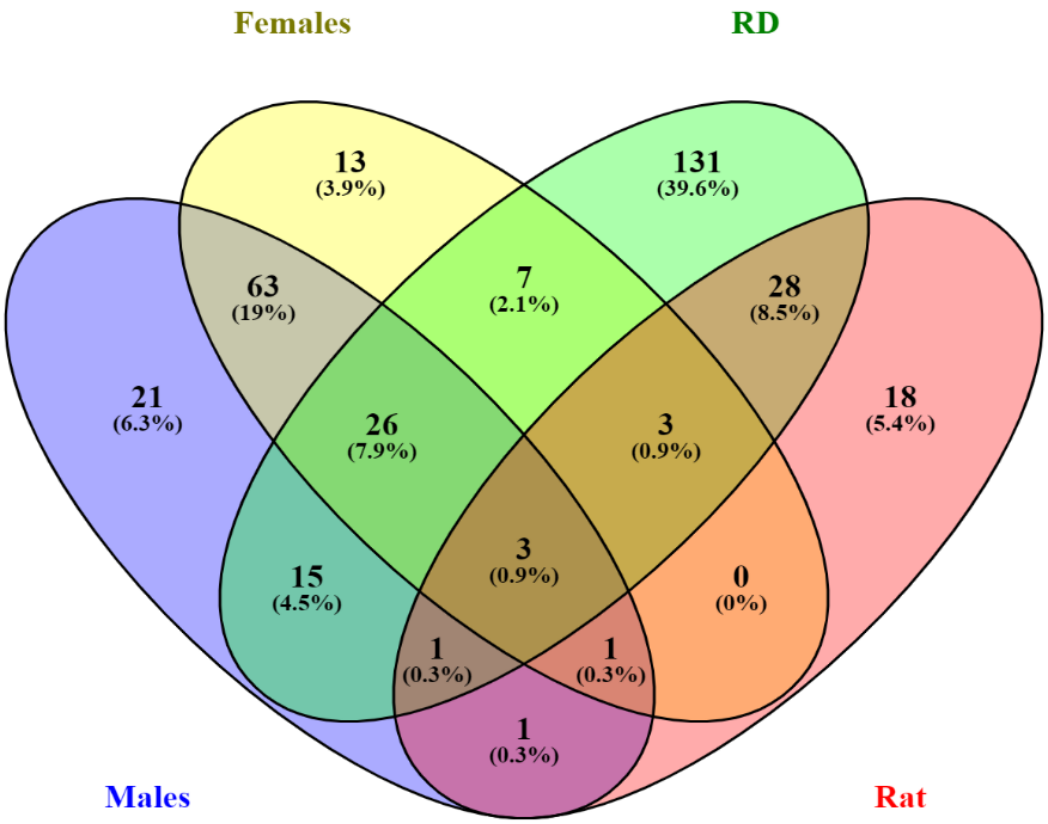


Figure A.1 Gene overlap within the five convergent pathways identified from patient dataset and experimental systems

The individual and non redundant genes were overlapped between all datasets from the five convergent IPA identified pathways. The Venn diagram shows the overlapping and unique genes within all datasets. Of the five convergent pathways, the overlapping and distinct genes regulating the pathways are shown in the figure.

Table A.17 Overlapping pathways between any three datasets

N=23 common pathways between Male, Female, and RD pathway analysis			
Pathway	**Male -log(P-value)**	**Female -log(P-value)**	**RD -log(P-value)**
Neuroinflammation Signaling Pathway	2.6	1.8	3.3
Neurovascular Coupling Signaling Pathway	13.4	12.7	7.0
CREB Signaling in Neurons	11.1	13.7	4.0
Gustation Pathway	10.0	10.8	9.3
Glutamate Receptor Signaling	7.6	9.6	1.5
GABA Receptor Signaling	7.5	6.5	4.8
Synaptic Long Term Depression	4.7	3.6	3.2
Endocannabinoid Neuronal Synapse Pathway	3.3	2.4	4.8
Neuropathic Pain Signaling In Dorsal Horn Neurons	3.1	5.1	5.9
Synaptogenesis Signaling Pathway	2.8	5.2	5.0
nNOS Signaling in Neurons	1.9	1.4	2.4
Opioid Signaling Pathway	1.6	1.3	5.1
G-Protein Coupled Receptor Signaling	4.6	7.8	3.0
Kinetochore Metaphase Signaling Pathway	3.9	2.4	9.0
Mitotic Roles of Polo-Like Kinase	1.4	1.4	5.0
Netrin Signaling	3.5	1.8	6.2
Inhibition of Matrix Metalloproteases	3.8	2.9	4.6
cAMP-mediated signaling	2.0	3.9	1.5
Coagulation System	2.6	4.8	1.3
Wound Healing Signaling Pathway	6.3	2.4	3.4
Phagosome Formation	5.6	10.5	3.3
Agranulocyte Adhesion and Diapedesis	2.6	2.0	3.1
Maturity Onset Diabetes of Young (MODY) Signaling	6.2	2.1	2.4

Table A.18 Overlapping pathways between any three datasets

N=1 common pathways between Male, Female, and Rat			
Pathway	**Male -log(P-value)**	**Female -log(P-value)**	**Rat -log(P-value)**
Complement system	1.9	3.8	1.5

N=1 common pathways between Male, RD, and Rat			
Pathway	**Male -log(P-value)**	**RD -log(P-value)**	**Rat -log(P-value)**
Role of Pattern Recognition Receptors in Recognition of Bacteria and Viruses	1.7	1.6	1.5

N=1 common pathways between Female, RD, and Rat			
Pathway	**Female -log(P-value)**	**RD -log(P-value)**	**Rat -log(P-value)**
Cellular Effects of Sildenafil (Viagra)	2.4	11.4	1.9

Table A.18 Overlapping pathways between any two datasets

N=36 common pathways between Male and Female datasets		
Pathway	**Male -log(P-value)**	**Female -log(P-value)**
Intrinsic Prothrombin Activation Pathway	10.8	9.0
Granulocyte Adhesion and Diapedesis	5.1	4.2
MSP-RON Signaling Pathway	3.1	1.8
Extrinsic Prothrombin Activation Pathway	2.4	2.9
Differential Regulation of Cytokine Production in Macrophages and T Helper Cells by IL-17A and IL-17F	2.0	1.3
Pyroptosis Signaling Pathway	2.0	3.1
TREM1 Signaling	2.0	1.5
IL-23 Signaling Pathway	1.3	1.8
FXR/RXR Activation	9.9	4.8
PXR/RXR Activation	5.3	2.2
LXR/RXR Activation	5.3	2.7
LPS/IL-1 Mediated Inhibition of RXR Function	4.6	2.1
Xenobiotic Metabolism CAR Signaling Pathway	2.1	1.5
Xenobiotic Metabolism PXR Signaling Pathway	1.8	1.6
Nicotine Degradation III	5.7	4.2
Superpathway of Melatonin Degradation	5.6	3.8
Melatonin Degradation I	5.1	4.3
Serotonin Receptor Signaling	4.7	3.4
Nicotine Degradation II	4.6	3.4
Bupropion Degradation	3.5	2.9
Serotonin Degradation	2.4	2.6
Retinol Biosynthesis	2.3	3.8
Serotonin and Melatonin Biosynthesis	2.3	1.6
Acetone Degradation I (to Methylglyoxal)	1.9	2.3
Histidine Degradation III	2.6	1.3
Thyroid Hormone Metabolism II (via Conjugation and/or Degradation)	2.6	1.8
Estrogen Biosynthesis	2.5	2.8
Asparagine Degradation I	1.7	1.9
Androgen Biosynthesis	1.6	1.5
Tryptophan Degradation to 2-amino-3-carboxymuconate Semialdehyde	1.4	1.6
α-tocopherol Degradation	1.4	2.5
Fatty Acid Activation	1.3	1.6
Superpathway of Citrulline Metabolism	1.3	1.6
Chondroitin Sulfate Biosynthesis	1.5	1.5
Glutamate Dependent Acid Resistance	1.7	1.9
Embryonic Stem Cell Differentiation into Cardiac Lineages	2.1	2.5

Table A.19 Overlapping pathways between any two datasets

N=7 common pathways between Male and RD datasets		
Pathway	**Male -log(P-value)**	**RD -log(P-value)**
HOTAIR Regulatory Pathway	1.8	2.8
Role of Macrophages, Fibroblasts and Endothelial Cells in Rheumatoid Arthritis	1.6	3.6
FcγRIIB Signaling in B Lymphocytes	1.3	2.8
Type II Diabetes Mellitus Signaling	2.3	2.4
Corticotropin Releasing Hormone Signaling	1.6	4.1
GPCR-Mediated Nutrient Sensing in Enteroendocrine Cells	1.8	4.1
VDR/RXR Activation	1.3	1.5

N=7 common pathways between Female and RD datasets		
Pathway	**Female -log(P-value)**	**RD -log(P-value)**
eNOS Signaling	3.2	4.3
Nitric Oxide Signaling in the Cardiovascular System	1.5	4.7
Phospholipases	1.4	2.3
Eicosanoid Signaling	1.3	2.2
Apelin Liver Signaling Pathway	2.1	1.4
Dopamine-DARPP32 Feedback in cAMP Signaling	2.0	5.1
Estrogen-mediated S-phase Entry	1.6	2.3

Table A.19 Overlapping pathways between any two datasets

N=21 common pathways between RD and Rat datasets		
Pathway	**RD -log(P-value)**	**Rat -log(P-value)**
Actin Cytoskeleton Signaling	7.0	2.8
ATM Signaling	2.5	1.8
Cardiac Hypertrophy Signaling	3.7	1.4
Cardiac Hypertrophy Signaling (Enhanced)	10.5	2.2
Caveolar-mediated Endocytosis Signaling	7.3	1.6
Cell Cycle: G1/S Checkpoint Regulation	1.3	1.4
DNA Methylation and Transcriptional Repression Signaling	2.0	1.6
Ferroptosis Signaling Pathway	2.7	1.8
NAD Signaling Pathway	3.3	2.3
NER (Nucleotide Excision Repair, Enhanced Pathway)	4.4	2.0
Neuregulin Signaling	4.9	1.3
Regulation of Actin-based Motility by Rho	3.9	1.7
Regulation of Cellular Mechanics by Calpain Protease	6.6	1.6
Regulation of eIF4 and p70S6K Signaling	1.3	1.3
RHOGDI Signaling	6.5	1.3
Semaphorin Neuronal Repulsive Signaling Pathway	9.5	1.9
Senescence Pathway	3.5	1.3
Signaling by Rho Family GTPases	4.7	1.5
STAT3 Pathway	3.9	2.3
VEGF Signaling	1.3	1.3
White Adipose Tissue Browning Pathway	3.5	2.3

Table A.19 Unique pathways in individual datasets

N=9 exclusive pathways in the Female dataset	
Pathway	**-log(P-value)**
Ethanol Degradation II	1.4
L-dopachrome Biosynthesis	1.9
NAD biosynthesis II (from tryptophan)	1.3
Neuroprotective Role of THOP1 in Alzheimer's Disease	2.3
Osteoarthritis Pathway	1.5
Phototransduction Pathway	4.6
SPINK1 Pancreatic Cancer Pathway	4.5
Triacylglycerol Degradation	1.5
Urea Cycle	1.6

N=9 exclusive pathways in the Male dataset	
Pathway	**-log(P-value)**
Acute Phase Response Signaling	2.3
Chondroitin Sulfate Biosynthesis (Late Stages)	1.3
Differential Regulation of Cytokine Production in Intestinal Epithelial Cells by IL-17A and IL-17F	1.4
Glycine Betaine Degradation	2.1
HMGB1 Signaling	1.3
IL-17 Signaling	1.8
L-serine Degradation	1.3
Mitochondrial L-carnitine Shuttle Pathway	1.4
Oxytocin In Brain Signaling Pathway	3.9

Table A.20 Unique pathways in individual datasets

N=131 exclusive pathways in the RD dataset	
Pathway	**-log(P-value)**
3-phosphoinositide Biosynthesis	1.7
Acetate Conversion to Acetyl-CoA	1.6
Actin Nucleation by ARP-WASP Complex	2.8
Adrenomedullin signaling pathway	3.0
Aldosterone Signaling in Epithelial Cells	1.5
AMPK Signaling	3.7
Androgen Signaling	3.2
Angiopoietin Signaling	1.3
Apelin Cardiac Fibroblast Signaling Pathway	1.7
Apelin Cardiomyocyte Signaling Pathway	3.7
Apelin Endothelial Signaling Pathway	3.5
Apelin Pancreas Signaling Pathway	2.1
Apoptosis Signaling	1.8
Aryl Hydrocarbon Receptor Signaling	1.7
Autophagy	1.9
Axonal Guidance Signaling	12.7
BER (Base Excision Repair) Pathway	1.9
BMP signaling pathway	2.3
Cardiac β-adrenergic Signaling	4.3
CDK5 Signaling	3.7
CDP-diacylglycerol Biosynthesis I	1.5
Cell Cycle Control of Chromosomal Replication	6.8
Cell Cycle: G2/M DNA Damage Checkpoint Regulation	1.7
Ceramide Signaling	1.3
Circadian Rhythm Signaling	3.8
Creatine-phosphate Biosynthesis	2.3
CXCR4 Signaling	1.8
Cyclins and Cell Cycle Regulation	2.2
Death Receptor Signaling	2.0
D-myo-inositol (1,4,5,6)-Tetrakisphosphate Biosynthesis	1.4
D-myo-inositol-5-phosphate Metabolism	1.5
DNA Double-Strand Break Repair by Homologous Recombination	2.8
Endocannabinoid Cancer Inhibition Pathway	2.9
Endocannabinoid Developing Neuron Pathway	1.6
Endothelin-1 Signaling	3.9
Ephrin Receptor Signaling	5.7
ERK/MAPK Signaling	4.4
Estrogen Receptor Signaling	6.0
Factors Promoting Cardiogenesis in Vertebrates	4.9
FAK Signaling	3.8
Fc Epsilon RI Signaling	1.3

G Beta Gamma Signaling	5.6
GADD45 Signaling	2.8
Gap Junction Signaling	2.6
GDNF Family Ligand-Receptor Interactions	1.4
GDP-glucose Biosynthesis	1.4
Glycine Degradation (Creatine Biosynthesis)	1.4
Glycogen Biosynthesis II (from UDP-D-Glucose)	1.6
GNRH Signaling	3.8
GPCR-Mediated Integration of Enteroendocrine Signaling Exemplified by an L Cell	1.4
Granzyme A Signaling	2.3
Guanosine Nucleotides Degradation III	1.6
Gα12/13 Signaling	1.6
Gαi Signaling	1.7
Hepatic Fibrosis Signaling Pathway	11.9
HGF Signaling	3.5
HIF1α Signaling	3.0
Human Embryonic Stem Cell Pluripotency	2.7
IGF-1 Signaling	1.8
IL-1 Signaling	2.9
IL-15 Production	3.3
IL-6 Signaling	1.7
IL-8 Signaling	3.8
ILK Signaling	5.6
Inhibition of Angiogenesis by TSP1	2.2
Insulin Secretion Signaling Pathway	2.3
Integrin Signaling	2.7
Interferon Signaling	4.0
Iron homeostasis signaling pathway	1.4
JAK/STAT Signaling	1.3
Leptin Signaling in Obesity	2.6
Leukocyte Extravasation Signaling	4.1
Macropinocytosis Signaling	4.6
Melanocyte Development and Pigmentation Signaling	2.0
Melatonin Signaling	2.2
Mismatch Repair in Eukaryotes	3.0
Molecular Mechanisms of Cancer	8.5
MYC Mediated Apoptosis Signaling	1.4
NAD Salvage Pathway II	1.7
NF-κB Activation by Viruses	2.7
Notch Signaling	2.1
Oncostatin M Signaling	1.6
Oxytocin Signaling Pathway	7.1
P2Y Purigenic Receptor Signaling Pathway	4.3
p38 MAPK Signaling	1.5

p53 Signaling	1.6
PAK Signaling	6.9
Paxillin Signaling	4.7
Pentose Phosphate Pathway (Oxidative Branch)	1.3
PFKFB4 Signaling Pathway	2.5
Phagosome Maturation	3.4
Phosphatidylglycerol Biosynthesis II (Non-plastidic)	1.3
PI3K/AKT Signaling	5.5
PPARα/RXRα Activation	2.7
Prolactin Signaling	1.3
Protein Kinase A Signaling	6.8
PTEN Signaling	7.0
Pulmonary Healing Signaling Pathway	4.9
Purine Nucleotides De Novo Biosynthesis II	2.1
Pyrimidine Deoxyribonucleotides De Novo Biosynthesis I	1.3
Pyrimidine Ribonucleotides Interconversion	1.3
RAC Signaling	4.4
RAR Activation	2.3
Reelin Signaling in Neurons	2.0
Relaxin Signaling	3.3
Renin-Angiotensin Signaling	3.4
RHOA Signaling	1.6
Role of CHK Proteins in Cell Cycle Checkpoint Control	1.8
Role of Hypercytokinemia/hyperchemokinemia in the Pathogenesis of Influenza	1.6
Role of JAK1 and JAK3 in γc Cytokine Signaling	1.8
Role of MAPK Signaling in Inhibiting the Pathogenesis of Influenza	1.3
Role of MAPK Signaling in Promoting the Pathogenesis of Influenza	1.5
Role of NANOG in Mammalian Embryonic Stem Cell Pluripotency	1.8
Role of NFAT in Cardiac Hypertrophy	7.7
Role of Tissue Factor in Cancer	2.6
Sonic Hedgehog Signaling	1.9
Superpathway of Inositol Phosphate Compounds	2.1
Telomere Extension by Telomerase	1.3
TGF-β Signaling	1.3
Thrombin Signaling	2.5
Tight Junction Signaling	2.5
tRNA Splicing	1.7
Urate Biosynthesis/Inosine 5'-phosphate Degradation	1.5
UVA-Induced MAPK Signaling	1.6
UVB-Induced MAPK Signaling	1.6
VEGF Family Ligand-Receptor Interactions	3.3
Virus Entry via Endocytic Pathways	4.7
α-Adrenergic Signaling	3.6
β-alanine Degradation I	1.4

γ-glutamyl Cycle	1.6
γ-linolenate Biosynthesis II (Animals)	1.3

Table A.20 Unique pathways in individual datasets

N=8 exclusive pathways in the Rat dataset	
Pathway	**-log(P-value)**
Adipogenesis pathway	4.7
EIF2 Signaling	3.7
Dilated Cardiomyopathy Signaling Pathway	2.8
Ephrin B Signaling	1.7
Glycine Cleavage Complex	1.5
Stearate Biosynthesis I (Animals)	1.4
Ceramide Degradation	1.4
Agrin Interactions at Neuromuscular Junction	1.3