



Faculty of health- and Social Sciences
Department of Sports Science

Candidate nr: 301

Master's thesis
The missing Incs

The missing Incs

IDR3005

2020

Acknowledgement

First I would like to give a huge thank you, to my supervisor at Department for Sport Science Stian Ellefsen. You were always available for questions and good advices. Especially the advices you gave regarding data wrangling in R studio. You had an idea of how it should be done, and look like in the end. Luckily, I had Daniel Hammarstrøm to turn to when the data wrangling became challenging. Thank you, Daniel!

Writing this paper has been quite the journey. I started out as a novice in the lab, not knowing anything. After some time and a little help, I started to feel confident on qPCR analysis. The collection of data from qPCR analysis started in 2018, and I had all the time in the world. After months of data wrangling in R, waiting for the new lab and for the results from the RNA-seq analysis, I was suddenly in a hurry. It did not help that the SARS-cov-2 virus emerged and shut down everything. Luckily I made it to the end after months and several hundred hours of data-wrangling in R.

Thanks to the rest of the staff at the Department of Sport Science. You have all helped me on the journey to complete this master's thesis.

Ragnvald B. Steile

Jessheim, 15.09.2020

Abstract

Introduction: Resistance training is volume dependent. The muscular adaptations to resistance training vary between moderate and low training volume. Little is known about how these muscular adaptations happen. LncRNAs have emerged as an interesting regulator of different signaling pathways connected to cell proliferation and growth. The aim of this study was to explore and identify differentially expressed lncRNAs in m. vastus lateralis, and thus, explore volume- and time-effects on muscular adaptations.

Method: Forty-one female and male participants were recruited to the study, of which 25 had biopsies, from all three timepoints, with sufficient RNA-quality. Strength tests and muscle biopsies were taken before, in the middle and after a 12week contralateral, within subject, resistance training intervention. Biopsies from the 25 participants were sent to RNA sequencing. RNA-seq data was analyzed with Mixed-effects negative binomial count models, and differential expression and log₂fold-change was calculated on all three timepoints.

Results: Analysis of RNA-seq data identified 1400 lncRNAs, of which ~12% percent were differentially expressed (DE). Between timepoint w2pre and w12, 169 lncRNAs were differentially expressed. Most of the lncRNAs identified were upregulated, and 17 lncRNAs were DE at all three timepoints. No significant difference was found between low and moderate volume.

Conclusion

As many as 17 DE lncRNAs were found on all three timepoints, suggesting that they are important in muscle adaptations to resistance training. Resistance training with low and moderate volume resulted in similar changes in lncRNA expression, reiterating on the fact that the different volume conditions do not lead to substantial differences in cellular phenotypes measured per unit muscle tissue (though higher volume is associated with larger increases in muscle mass). More research is needed to expand the entrezgene id database and allocate gene annotations.

Keywords

RNA-seq, qPCR, resistance training, long non-coding RNA, skeletal muscle.

The Missing lncs

Table of Contents

Acknowledgement.....	1
Abstract.....	2
Abbreviations.....	4
I. Theory.....	6
I.I Training volume and muscle growth.....	7
I.II Cellular pathways.....	8
I.III Long non-coding RNA.....	9
I.III.I Cell growth and proliferation.....	10
I.III.II Methods used to study lncRNAs in human muscle cells.....	11
1. Introduction.....	13
2. Methods.....	14
2.1. Ethical approval.....	14
2.2. Intervention and participants overview.....	15
2.3. Training protocol and testing of muscle strength.....	15
2.4. Muscle biopsies.....	16
2.5. Total RNA extraction.....	16
2.6. RNA sequencing data.....	17
2.7. Quantitative real-time reverse transcription polymerase chain reaction (qPCR).....	18
2.8. Statistics.....	19
3. Results.....	19
3.1. LncRNAs.....	20
4. Discussion.....	25
5. Conclusion.....	28
References.....	29
Appendix 1, Khan et al. (2020).....	33
Appendix 2, All lncRNA identified in RNA-seq muscle biopsies data.....	67

Abbreviations

Differentially expressed (DE)

Eukaryotic translation initiation factor 4E-binding protein 1 (4E-BP1)

Extracellular matrix (ECM)

Enhancer of zeste 2 polycomb repressive complex 2 subunit(EZH2)

Focal adhesion kinase (FAK)

Fold change (FC)

Growth arrest specific 5 (GAS5)

Long non coding RNA (lncRNA)

Mammalian target of rapamycin (mTOR)

Mammalian target of rapamycin complex 1 (mTORC1)

Mammalian target of rapamycin complex 2 (mTORC2)

Mitogen-activated protein kinase (MAPK)

Mitogen-activated protein kinase kinase (MAP2K)

Mitogen-activated protein kinase kinase kinase (MAP3K)

Muscle protein synthesis (MPS)

Over-representation analysis (ORA),

Phosphatidic acid (PA)

Protein Kinase B (Akt)

Quantitative /real time polymerase chain reaction (qPCR)

Ribosomal protein S6K (p70S6K)

RNA component of mitochondrial RNA processing endoribonuclease (RMRP)

40S ribosomal protein S6 kinase (RSK)

I. Theory

Humans are made for movement. As such, it is important to sustain proper skeletal muscle functions throughout the life course. This can effectively be achieved through exercise training directed at improving muscle functions, with resistance training standing out as the preferred training modality, improving exercise performance and functionality, as well as promoting systemic health (Kraemer, Ratamess, & French, 2002). Resistance training exposes skeletal muscle to mechanical and metabolic stress, thus triggering cellular signaling cascades and changes in gene expression that eventually leads to muscle growth and increased muscle strength and endurance (Hughes, Ellefsen, & Baar, 2018). Despite this simplified view, there are vast numbers of different resistance training programs and methods, varying in training volume, repetitions or load (Egan & Zierath, 2013). Training volume is interesting. During the past few years, studies have shown that higher training volume increases the changes in muscle growth and strength. Contralateral resistance training protocols have shown high correlations between volume and muscular adaptations (Hammarström et al., 2020). Although relatively much is known about the gross adaptations to different resistance training protocols, and thus how to maximize strength and muscle gain (Hughes et al., 2018), little is known about the microbiology controlling muscle adaptations. Whereas we know fairly well which main signaling pathways that are involved, such as Akt and MAPK (Bodine et al., 2001) our knowledge about the detailed changes in muscle biology are far from complete. For example, we hardly know anything about the role of long non-coding RNAs (lncRNAs), which are protruding as important regulators of cellular growth and differentiation in other experimental human cell models.

lncRNAs thus remain a bit of a mystery. For a long time they were solely regarded as transcriptional noise, but in recent years they have gained reputation as important contributors to and regulators of cellular functions (Kung, Colognori, & Lee, 2013). They represent a diverse class of long RNAs that are not typically translated into protein (Chen et al., 2018). Instead they affect cellular functionality by interacting with other types of RNA (e.g. mRNA and microRNA) or by altering micropeptide functions (Cesana et al., 2011; Douglas et al., 2015; G. Hu et al., 2018; G.-Q. Wang et al., 2016), and are often seen to affect cellular growth, so also in muscle fibre. As such, it remains plausible that lncRNAs are involved in adaptations to resistance

training in humans. However, so far the effects of such training on their expression and their implications for muscle adaptations and functions remain elusive (Hughes et al., 2018).

I.I Training volume and muscle growth

Different resistance training protocols yields different training effects. Some are more effective than others, and especially training volume correlates with muscle mass (Folland & Williams, 2007; Hammarström et al., 2020). Muscle mass correlates with muscle strength, and this correlation is primarily visible after long periods of resistance training (Folland & Williams, 2007). Muscle strength is also correlated with pennation angle and other muscle biological features like ECM. Adaptation to resistance training is an individual response and are not yet fully understood. Usually the load vary between 1RM and 10RM, the repetitions between 4-12 and the sets between 1-6 (Fry, 2004). Prediction models on how the human body responds and adapts to various stimuli, are made to understand muscle adaptations. The results from these models advocates that there are a strong correlation between training volume and muscle growth (Hester, Iliescu, Summers, & Coleman, 2011). For a long time this was debated, but recent meta-studies have shown that the former is true (Ralston, Kilgore, Wyatt, & Baker, 2017).

Resistance training induces mechanical, and metabolic, stimuli to skeletal muscles (Folland & Williams, 2007). and the muscle adapts, among others, by adding sarcomeres in parallel in muscle fibers (Folland & Williams, 2007). Resistance training results in neural adaptations, improved strength, alter muscle phenotype and increased cross sectional area (CSA) of the muscle fiber. Other adaptations to resistance training are increase in noncontractile tissue, e.g collagen, and change in muscle fibers angle of pennation. Adaptations to resistance training can be identified after 8-12 weeks of repeated training (Folland & Williams, 2007). Over time, the strength gain will be more due to muscle growth than neural adaptations. There will be an increase in muscle net protein synthesis (MPS). MPS is increased due to lower protein degradation and higher, maybe more efficient, protein synthesis (Damas et al., 2016). The central neural component is important for muscle adaptations due to resistance training, especially with unilateral training. With the latter, the CSA in the untrained leg does not change, but one can observe an increase in strength (Munn, Herbert, & Gandevia, 2004).

Results from animal studies indicate that lncRNAs may regulate satellite cell biology (Li, Chen, Sun, & Wang, 2018). Satellite cells are important for muscle regeneration. If an injury occurs, the satellite cells will be activated and become myoblasts. Pax7 is downregulated and myogenic regulatory factors (MRFs) are activated to start cell differentiation, thereby making new muscle fibers and replenish the damaged muscle cells (Kuang, Kuroda, Le Grand, & Rudnicki, 2007).

I.II Cellular pathways

Muscle growth is facilitated by organelles in the muscle cell, and involves translation capacity changes and gene expression patterns. Satellite cells facilitate muscle growth by providing myocores, and different signalling pathways control cellular growth and differentiation. Usually changes in protein and RNA expression, including rRNA, mRNA and other RNA species, are a result of the latter (Hughes et al., 2018). Changes in gene expression are crucial in muscular adaptations to resistance training, and the latter are controlled by cellular pathways. Many different signaling pathways have been identified and explored. Some of them are connected to muscle adaptation after resistance training. One of the most important to resistance training adaptations is mammalian target of rapamycin (mTOR) (Hoppeler, 2016). mTOR is a part of the phosphatidylinositol 3-kinase-related kinase family and are important in two distinct multi-protein complexes, mTOR C1 and C2, that regulate muscle growth (Hoppeler, 2016). The former is recognized as the one most important to muscle adaptation. The complete function of mTORC2 is still eluded, but research advocates that it may be associated with regulating ribosomal activity, and cell survival (Chaillou, Kirby, & McCarthy, 2014). Initiation of protein translation is activated by phosphatidic acid (PA), the latter activates mTORC1 and thereafter ribosomal protein S6K (p70S6K) (Bodine et al., 2001). mTORC1 is targeting different signaling pathways and proteins, and is probably very important for muscle protein synthesis (MPS) (Mirzoev & Shenkman, 2018). Mechanical stress, such as resistance training, activates mitogen-activated protein kinase (MAPK), which can phosphorylate C-myc. C-myc regulates transcription factors and thereby may regulate transcription of proteins (Hoppeler, 2016). Insulin-like growth factor 1 (IGF-1) is also activated, by mechanical stress, and binds to receptors in the cell membrane, and this initiates stimulation of Phosphoinositide 3-kinase PI3K- and protein kinase B (Akt)-activity (Bodine et al., 2001; McCarthy & Esser, 2010). Akt phosphorylates downstream effectors and that

activates G-protein Rheb (Ras-homolog enriched in brain) and that again activates mTORC1 (Bodine et al., 2001) Ras is activated when extracellular mitogen binds to the membrane receptor (McCarthy & Esser, 2010). Followed by activation of Mitogen-activated protein kinase kinase kinase (MAP3K), Mitogen-activated protein kinase kinase (MAP2K) and Mitogen-activated protein kinase (MAP), thus activate Myc, or other transcription factors (McCarthy & Esser, 2010).

Little is known of proteins that regulates satellite cells. McCroskery, Thomas, Maxwell, Sharma, and Kambadur (2003) advocates that satellite cells are regulated by myostatin. Upregulated myostatin levels increase p21, a cyclin-dependent kinase inhibitor, and inhibits differentiation. Satellite cells, or muscle stem cells, are small cells that can evolve to skeletal muscle cells. When activated they can proliferate and transform to myoblasts. The latter can induce muscle fiber hypertrophy or make new muscle cells (Morgan & Partridge, 2003). They are situated between the sarcolemma and membrane of the muscle fiber. Activated upon mechanical strain. Exercise triggers a cascade of different signaling molecules, e.g. growth factors and cytokines. HGF activates satellite cells, fibroblast growth factor (FGF) and insulin-like growth factor-I (IGF-1) increases proliferation. Translation of ribosome and the making of proteins is the key to muscle growth. Translation depends on two variables, translation capacity and translation efficiency. Translation capacity is the number of ribosome available, tRNA and translation factors. And translation efficiency is the efficiency of the protein synthesis (Chaillou et al., 2014). Increase in the latter is likely one of the main variables behind elevated MPS as a response to resistance training (O'Neil, Duffy, Frey, & Hornberger, 2009). Mdm2 -p53 stress response pathway regulates cellular homeostasis. If activated it results in apoptosis, cell cycle arrest, DNA repair or replicative senescence. It is important to regulate cell growth (Bartlett, Close, Drust, & Morton, 2014).

I.III Long non-coding RNA

The lncRNAs are made up of over 200 nucleotides (Ponting, Oliver, & Reik, 2009). Little is known of their function and the evidence for lncRNAs function is scarce. But due to more advanced research methods, and higher interest the recent years, many lncRNAs has been

identified and annotated. It may be plausible to claim that lncRNAs could be very responsive to resistance training, since studies have shown that they control different groups of genes and proteins responsible for cell- differentiation, proliferation and cell growth (Chen et al., 2018). As late as in the early 1990s, Brannan, Dees, Ingram, and Tilghman (1990) were the first to discover that the lncRNA H19 was involved in epigenetic regulation. Later, many more lncRNAs have been identified and described. Their function in epigenetic regulation are still eluded, but some of them are well known. Some of the lncRNAs can alter the coding gene by pairing with mRNA (G.-Q. Wang et al., 2016), and other can interact with microRNAs and make them miss their target mRNA (Cesana et al., 2011). Yet another group of lncRNAs can encode micropetides, that are shorter than 100 amino acids, and by doing that alter the micropeptide induced functions (Douglas et al., 2015). Most of the lncRNAs directly linked to myogenesis acts as transcriptional or epigenetic regulators (Li et al., 2018). LncRNAs can be allocated to five categories based on where they are situated in the genome: 1) sense, 2) antisense, 3)bidirectional, 4)intronic and 5)intergenic (Ponting et al., 2009). In the first category they overlap one or more exons of another exon on the same strand. The second is the same as the first, except the lncRNA is on the opposite strand. In the third category, the lncRNA is in close genomic proximity to a coding transcript on the opposite strand. The fourth is when it is derived from an intron on a second transcript, and the fifth is when it is in the genomic interval between two different genes.

I.III.I Cell growth and proliferation

Some of the most known lncRNAs induces cell growth and cell proliferation. Knockout gene studies have been conducted to explore lncRNAs role in tumor growth. Another important feature for the lncRNAs is that they play a role in epigenetic and transcriptional regulation of chromatins. The lncRNAs interacts with chromatins, and may inhibit other transcriptional regulators activities (Han et al., 2014). In mice, lncRNA SYISL interacts with polycomb repressive complex 2 and regulates myogenesis (Jin et al., 2018). SYISLs human homolog is lncRNA AK021986, but no ensemble ID has been allocated and the latter is therefore removed from the qPCR analysis. The lncRNA H19 induces muscle differentiation in mice, but its function in human muscle cells are not fully understood (Kallen et al., 2013). The lncRNA Growth arrest specific 5

(GAS5), suppress MYC translation (Pickard, Mourtada-Maarabouni, & Williams, 2013). Parrot/LINP1 is a positive regulator of c-Myc and ribosomal biogenesis (Zhang et al., 2016). PVT1 is activated in the early phase of muscle atrophy. PVT1 alter mitochondrial respiration, myofiber size, apoptosis and mito/autophagy (Tseng et al., 2014). RNA component of mitochondrial RNA processing endoribonuclease (RMRP) alters the transport through the mitochondrial membrane (X. Wang et al., 2018). Linc-MD1 has many functions related to muscle adaptations. It regulates myogenic differentiation, myogenesis and hypertrophy. Linc-MD1 downregulates myogenic markers when depleted. It is reported to influence the mRNA levels of miRNA-targeted muscle differentiation genes, by and blocking the target mRNA for miR-133 and miR-135 (Cesana et al., 2011).

I.III.II Methods used to study lncRNAs in human muscle cells

As mentioned, little is known of lncRNAs function and expression in human biology. And especially in muscle biology. How their expression changes dependent on training stimuli, are largely unknown. The lncRNAs may alter signaling pathways and cellular responses. By altering signaling pathways, they may alter muscle growth and differentiation. Discovery and characterization of lncRNAs has sped up due to the recent years leap in high throughput gene sequencing technology (Jason, Spacek, & Michael, 2015). qPCR analysis is considered the gold standard when analyzing gene expression, but it is time-consuming to explore big data frames (Adamski, Gumann, & Baird, 2014). The RNA sequencing method yields massive data, and mining for lncRNAs are possible with the right tools. The polyA-primed sequencing method used in this paper needs lncRNAs with poly A tails. A poly A tail consists of many adenosine monophosphates and helps preventing degradation of mRNA. LncRNAs normally have low expression and are highly tissue specific. They usually have a poly A⁺ or poly A⁻ tail at the 3' end of the transcript (Cabili et al., 2011). X. Sun et al. (2016) defined 7692 lncRNAs in bovine skeletal muscle using Ribo-Zero RNA-seq. This sequencing technology can identify both poly A⁺ and poly A⁻ transcripts. GO analysis are important when exploring RNA-seq data. It describes the genes cellular location, molecular function and biological functions (Yamaguchi et

al., 2008). Few annotated lncRNAs complicates GO analysis. An alternative approach could be to explore shared DE lncRNAs across different timepoints or top 5 genes based on log₂FC/p-value.

The lack of knowledge on how lncRNAs adapts to mechanical stimuli, resistance training, in human muscle, advocates further research on the topic. The goal of this study was to (i) identify lncRNAs that are expressed in m. vastus lateralis, to (ii) explore and identify lncRNAs responding to low resistance training volume in contrast to moderate volume, (iii) validate expression patterns of lncRNAs identified in RNA-seq data using gene-specific qPCR, and (iiii) to explore lncRNAs that are differentially expressed at all three timepoints. Within subject RNA-seq data and qPCR data were used to compare the benefits of moderate and low resistance training volume.

1. Introduction

Humans are made for movement. As such, it is important to sustain proper skeletal muscle functions throughout the life course. This can effectively be achieved through exercise training directed at improving muscle functions, with resistance training standing out as the preferred training modality, improving exercise performance and functionality, as well as promoting systemic health (Kraemer et al., 2002). Resistance training exposes skeletal muscle to mechanical and metabolic stress, thus triggering cellular signaling cascades and changes in gene expression that eventually leads to muscle growth and increased muscle strength and endurance (Hughes et al., 2018). Despite this simplified view, there are vast numbers of different resistance training programs and methods, varying in training volume, repetitions or load (Egan & Zierath, 2013). Training volume is interesting. During the past few years, studies have shown that higher training volume increases the changes in muscle growth and strength. Contralateral resistance training protocols have shown high correlations between volume and muscular adaptations (Hammarström et al., 2020). Although relatively much is known about the gross adaptations to different resistance training protocols, and thus how to maximize strength and muscle gain (Hughes et al., 2018), little is known about the microbiology controlling muscle adaptations. Whereas we know fairly well which main signaling pathways that are involved, such as Akt and MAPK (Bodine et al., 2001), our knowledge about the detailed changes in muscle biology are far from complete. For example, we hardly know anything about the role of long non-coding RNAs (lncRNAs), which are protruding as important regulators of cellular growth and differentiation in other experimental human cell models.

lncRNAs thus remain a bit of a mystery. For a long time they were solely regarded as transcriptional noise, but in recent years they have gained reputation as important contributors to and regulators of cellular functions (Kung et al., 2013). They represent a diverse class of long RNAs that are not typically translated into protein (Chen et al., 2018). Instead they affect cellular functionality by interacting with other types of RNA (e.g. mRNA and microRNA) or by altering micropeptide functions (Cesana et al., 2011; Douglas et al., 2015; G. Hu et al., 2018; G.-Q. Wang et al., 2016), and are often seen to affect cellular growth, so also in muscle fibre. As such, it remains plausible that lncRNAs are involved in adaptations to resistance training in humans. However, so far the effects of such training on their expression and their implications for muscle

adaptations and functions remain elusive (Hughes et al., 2018). Improved RNA-sequencing methods have made it possible to statistically explore massive data frames with millions of gene counts. qPCR is considered gold standard for gene expression analysis (Adamski et al., 2014), but the method is cumbersome and time-consuming. Thus, studies on the correlation between qPCR and RNA-seq are important.

The lack of knowledge on how lncRNAs adapts to mechanical stimuli, resistance training, in human muscle, advocates further research on the topic. The goal of this study was to (i) identify lncRNAs that are expressed in m. vastus lateralis, to (ii) explore and identify lncRNAs responding to low resistance training volume in contrast to moderate volume, (iii) validate expression patterns of lncRNAs identified in RNA-seq data using gene-specific qPCR, and (iiii) to explore lncRNAs that are differentially expressed at all three timepoints. Within subject RNA-seq data and qPCR data were used to compare the benefits of moderate and low resistance training volume.

2. Methods

This study is based on the 1/3 set study completed by Hammarström et al. (2020). A contralateral leg resistance training protocol was used. The aim was to evaluate how single and multiple set resistance training affected muscle hypertrophy, strength gain, fibre-type total RNA, mRNA and ribosomal RNA. The effects on mTORC1 related protein phosphorylation were also explored (Hammarström et al., 2020).

2.1. Ethical approval

Information about potential discomforts and risks associated with the study were given to all the participants and they gave their informed consent before study enrolment. All procedures were performed in accordance to the Declaration of Helsinki. The study design was pre-registered (ClinicalTrials.gov Identifier: NCT02179307) and approved by the local ethics committee at Lillehammer University College, Department of Sport Science (no. 2013-11-22:2).

2.2. Intervention and participants overview

Forty-one female and male participants were recruited to the study. Twenty-five of them had biopsies from all the different timepoints, with sufficient RNA quality, and they were sent to RNA sequencing. The eligibility criteria were age between 18 and 40 years and non-smoking. The exclusion criteria were impaired muscle strength due to ongoing or previous injury, prescribed medicine that could alter exercise adaptations, more than one strength exercise bout weekly during the last 12 months or local anesthetic intolerance. 7 participants were excluded during data analysis due to different reasons. Details can be found in Hammarström et al. (2020).

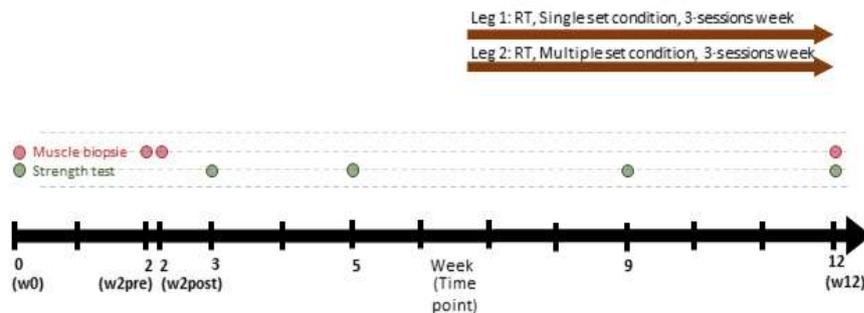


Figure 1. Study overview. The long arrow represents the 12 weeks of resistance training intervention. Muscle biopsies (red dots) were taken from *m. vastus lateralis* before the intervention, pre and post resistance training at week 2 and at 12 weeks. Strength tests (green dots) were conducted at week 0, 3, 5, 9 and 12. A contralateral, within subject, training program was performed and the resistance training was made up by three weekly sessions. The participant performed unilaterally leg exercises that were randomly assigned to one leg with single set resistance training and the other three set (multiple set) resistance training.

2.3. Training protocol and testing of muscle strength

The training bouts always started with 5 min of ergometer cycling with Borgs RPE 12-14 as a warm-up. Afterwards 4 bodyweight exercises (sit-ups, push-ups, back-extensions and squats) with 10 repetitions each. Followed by 10 reps at 50% of 1 repetition maximum (1RM) for each strength exercise. Thereafter unilateral leg press, leg curl and knee extension either as one set or three sets for the latter. The leg exercises were performed unilaterally to differentiate between single and multiple set. Thus, a contralateral protocol was induced to explore within subject

volume differences. After the lower leg exercises, they performed two sets of pull-down, seated rowing or shoulder-press and bench press. The intensity was progressed from 10RM (2 weeks), 8RM (3 weeks) to 7RM (7weeks). The rest period between the latter sets was 90-180 seconds. Strength tests were performed at week 0, 3, 5, 9 and 12. A dynamometer (Cybex 6000, Cybex International, Medway, MA, USA) was used to assess isometric and isokinetic unilateral knee-extension strength. Knee extension and unilateral leg press, tested at 1RM, determined Maximal strength (Hammarström et al., 2020). For more details see Hammarström et al. (2020).

2.4. Muscle biopsies

The muscle biopsies were taken, within 10 minutes, bilaterally from m. vastus lateralis using a spring-loaded biopsy instrument (Bard Magnum, Bard, Rud, Norway) with a 12-gauge needle (Universal-plus, Medax, San Possidonio, Italy). Local anaesthetics (Xylocaine, 10 mg ml⁻¹ with adrenaline 5µgml⁻¹, AstraZeneca AS,Oslo,Norway) was used during the latter protocol. The resting samples were taken after a standardized meal, at the same timepoint in the morning. Biopsies were taken minimum 48 hours after ended resistance training bout. Patella and spina iliaca anterior superior (SIAS) were used as landmarks, and the first biopsy was taken from 1/3 of the latter distance. Consecutive biopsies were taken 2 cm proximal to the latter sample. Ice cold saline solution (0,9%) was used when dissecting the muscle samples free from connective tissue and blood. The muscle tissue (~60 mg) that were to be used in RNA- and protein-analysis were quickly frozen in isopentane and stored at -80 degrees Celsius

2.5. Total RNA extraction

RNA extraction was done in accordance with the protocol found in Hammarström et al. (2020) article. 1 ml of TRIzol reagent (Invitrogen, Life technologies AS, Oslo, Norway) was used to homogenize about 25 mg of wet muscle tissue. RNase-free zirconium oxide beads was then added to the solution and run in a Bullet blender (Bullet Blender,NextAdvanced,Averill Park,NY, USA). 400 µl of the phase was allocated and isopropanol was used to precipitate a RNA pellet.

Before the RNA pellet was eluted in TE buffer, 70% EtOH was used in three washing steps. A spectrophotometer determined quality and amount of RNA.

2.6. RNA sequencing data

RNA-sequencing was performed in accordance to the unpublished paper conducted by Khan, Hammarstrøm, Rønnestad, Ellefsen & Ahmad (2020) at Norwegian Sequencing center (Appendix1).

LncRNAs were identified with BiomaRT R package and Ensemble ID. With this method is it not possible to find lncRNAs without Ensemble id. The latter applied to AKO21986, one of the lncRNAs analyzed with qPCR, and it was removed from further analysis. Data mining showed that Parrot was annotated with the Hgnc-symbol LINP1.

Only some of the lncRNAs identified in the gene sequence data have entrezgene id id number. Entrezgene id id is needed to perform a gene ontology analysis (GO). A GO analysis usually gives information about genes cellular location, molecular function and biological functions (Yamaguchi et al., 2008). GO makes it possible to sort genes based on the three latter variables and thus make a picture of the gene expression. National Center for Biotechnology Information (NCBI) hosts the gene specific database Entrez Gene. The database generates unique and stable gene identifier integers (Maglott, Ostell, Pruitt, & Tatusova, 2011). This gene ID are then used to integrate different information about the specific gene, such as nomenclature, sequence, pathways and protein interaction (Maglott et al., 2011). The information in the database is based on results from NCBI's other databases. Research on lncRNAs and their functions are therefore important and needed to evolve the database and increase the number of annotated lncRNAs. Thus, little information about lncRNAs in the NCBI database gives low quality gene analysis, which in return gives little information to the database.

2.7. Quantitative real-time reverse transcription polymerase chain reaction (qPCR)

qPCR was performed on selected lncRNAs (Table 1). cDNA synthesis was done, in accordance to Hammarström et al. (2020) prior to qPCR analysis. Oligo-dT random hexamer primers (Thermo Scientific) and Super Script IV Reverse Transcriptase (Invitrogen) were used to reverse transcribe 500 nanograms of RNA. A tissue offset normalisation factor was created based on the amount of tissue used in cDNA synthesis, and qPCR results (rested state samples from w2pre and w12) normalized, in accordance to Hammarström et al. (2020). The acute results (from w2post) were normalized to lib-size (Khan et al., 2020).

The qPCR was done with a qPCR machine (Applied Biosystems 7500 fast Real-Time PCR Systems, Life Technologies AS). Used 384 well plates filled with total 10 µl solution. The latter consisting of 2 µl cDNA, specific primers (Forward and Reverse, total 1µl), H2O and a prepared master mix (2X SYBR Select Master Mix, Applied Biosystems, Life Technologies AS). The qPCR protocol was 40 cycles (3 s 95°C denaturing and 30 s 60°C annealing).

Primers were designed for all selected long non coding RNAs (lncRNAs) with Primer3Plus (Untergasser et al., 2012) and ordered from Thermo Scientific. Primertests were performed and the primers with the best melt-curves, with no biproduct or primerdimers, was selected (single product amplification) (Table 1).

Table 1. Primers used for qPCR analysis of m.vastus lateralis biopsies. Primertests were conducted, and the primers with single product amplification, visually controlling the melt curves, were chosen.

NAME	GENE	TRANSCRIPT	FORWARD_PRIMER	REVERSED_PRIMER
GAS5 FIR1	ENSG00000234741	ENST00000650796.1	TGAAGAAATGCAGGCAGACC	CACTCTAGCTTGGGTGAGGC
LINCAKO17368 F2R2	ENSG00000268518	ENST00000595005.1	CCATCTGTCCGGAACCTCTGG	AGGCAAGTTGCTTCCTGTCT
LINCMD1 F3R3	ENSG00000225613	ENST00000418518.2	AGGTAGTGTGTCCCCAGCAC	CCTGTCTGGAAAGCCTTCAT
LNC1405 F3R3	ENSG00000185847	ENST00000657482.1	AACGGCTGGTCTTGAAGTCC	ATTGTGTCTTGGCTGTGCAC
LNC310CON1 F2R2	ENSG00000249515	ENST00000510302.1	GGATGACAGTGTCCAGGTCCC	ATAATGGTGGGGTGGCTGTG
LNC310CON2 F5R5	ENSG00000249515	ENST00000510302.1	TGAACAAATGAGACAAGGCTGC	GACAAGAGTCGGGCCTGAG
LINPI FIR1	ENSG00000223784	ENST00000650334.1	ACAGCCCTTAGGCTTGGACT	TCCCCATACCCTCTCCTACC
PVT1 FIR1	ENSG00000249859	ENST00000660438.1	CTGCATGGAGCTTCGTCAAG	CGTGTGTCATTCCAGTGCATG
RMRP FIR1	ENSG00000269900	ENST00000602361.1	CTCTGTTCCTCCCCTTTCCG	TCTTGGCGGACTTTGGAGTG

2.8. Statistics

All data-analysis and wrangling was done in RStudio (RStudio Team, 2016). LncRNAs were identified using biomaRT package in R. Mixed-effects negative binomial count models were fitted and saved in: ./R/dge_list_models.R. Results saved in RDS files for easy loading. The fixed effects are reduced to only contain gene-specific time + time:sets according to Hammarström et al. (2020). The RNAseq data was normalized to tissue weight, according to (Khan et al.2020). Statistical significance was set to $\alpha = 0.01$ and significant fold-change (FC) were below -0.5 and above 0.5. Fold change analysis was done at all timepoints and between low and moderate training volume. With the α set to 0.01 and the fold change significance level set to $< -0.5 : >0.5$, no significant difference was found between low and moderate volume. The same applied to the data when adjusting α level to 0.05 and 0.1

Raw data was exported from the qPCR machine and uploaded to RStudio and analyzed with the qpcR-package (Ritz & Spiess, 2008) written for R (Team, 2013). Threshold cycles (Ct) were estimated within the latter. Gene expression data were log-transformed prior to statistical analysis. The qPCR data was normalized to tissue weight based on amount of tissue used in cDNA synthesis (Hammarström et al., 2020). A Correlation test between qpcr data and RNAseq was performed using Pearson test. A GO analysis ,enrichGO from clusterprofiler (Yu, Wang, Han, & He, 2012) was performed, but no relevant genes could be sorted due to not annotated any GO groups. A literature survey was performed on top 5 lncRNAs for three timepoints, based on log2FC and adjusted p-value. The latter method was also applied to the 17 differentially expressed (DE) lncRNAs that were present at all three timepoints.

All datafiles, scripts, figures and code can be found at github.com “Innlevering_masteroppgave”

https://github.com/ragnvalds/Innlevering_masteroppgave

3. Results

Analysis was based on based on the muscle samples from Hammarström et al. (2020). Twelve weeks of moderate-volume resistance training led to larger increased in muscle strength

compared to low-volume training (3.4–7.7% difference, all $P < 0.05$) and CSA (5.2 (3.8)% versus 3.7 (3.7)%, $P < 0.001$) (Hammarström et al., 2020).

3.1. LncRNAs

In the entire RNA-seq data set, 15025 genes were identified as being expressed in *m. vastus lateralis*, 1400 of which were identified as lncRNAs (appendix 2). Analyses of pooled data (both legs combined) yielded differential expression (DE) of 169 transcripts at w2 (Figure 2A; 164 of which increased), 64 transcripts at w12 (Figure 2B; all of which increased) and 102 transcripts at w2post (Figure 2C; 40 of which increased), with as many as 17 transcripts being shared DE-genes between the three timepoints (Figure 2G). In analyses of the effects of resistance training volume on differential expressed lncRNAs, no differential responses for any of the lncRNAs were found (Figure 2D-F). RNA-seq-based estimates of the effects of resistance training on the expression of 6 selected lncRNAs (LincMD1, GAS5, LINP1 and PVT1, LINC01405 and RMRP) were correlated with qPCR-based estimates ($r=0.75$, $p=0.088$, 95% CI, -0.1650172 to 0.9702552; Figure 2H). Suggesting that RNA-seq and qPCR provided similar estimates of gene expression responses to resistance training.

The GO method, over-representation analysis (ORA), with significance level of adjusted p-value set to 0.05, log2fold-change to 0.5, was conducted on all the lncRNAs with Entrezgene id id using enrichGO from the clusterProfiler package made for R (Yu et al., 2012). ORA was used to assess if DE lncRNAs were affiliated to specific gene clusters. The latter analysis used pooled data (both legs combined) to explore time effects on adaptations to resistance training. Of the 1400 identified lncRNAs, 507 had been assigned Entrezgene id id, enabling GO-assessment, of which only 2 genes were annotated with description of cellular functions. As an alternative approach, the top five lncRNAs at each time point were chosen, based on adjusted p-value and log2FC, and explored further using literature survey (Table 2). Of the top 15 lncRNAs (5 from w2pre, w12 and w2post respectively), 11 were assigned with hgnc symbol of which 9 returned information after a literature survey. Of the top 15 lncRNAs from all three timepoints, only one

transcript from w2post (ENSG00000270605), was downregulated. Literature survey was also used to explore the 17 DE lncRNAs that were shared between the three timepoints (Table 3). Many of the lncRNAs have previously been associated with tumor growth or cancer cell proliferation. Others are known to accelerate muscle differentiation (H19), stimulate p53 (MEG3), β -catenin signaling (MIR4435-2HG) or activate the PI3K/AKT pathway (LINC00963). The shared 17 DE lncRNAs across all three timepoints show two distinct patterns. One is up-down-up, w2pre, w2post and w12 respectively. The other is up at all the three timepoints.

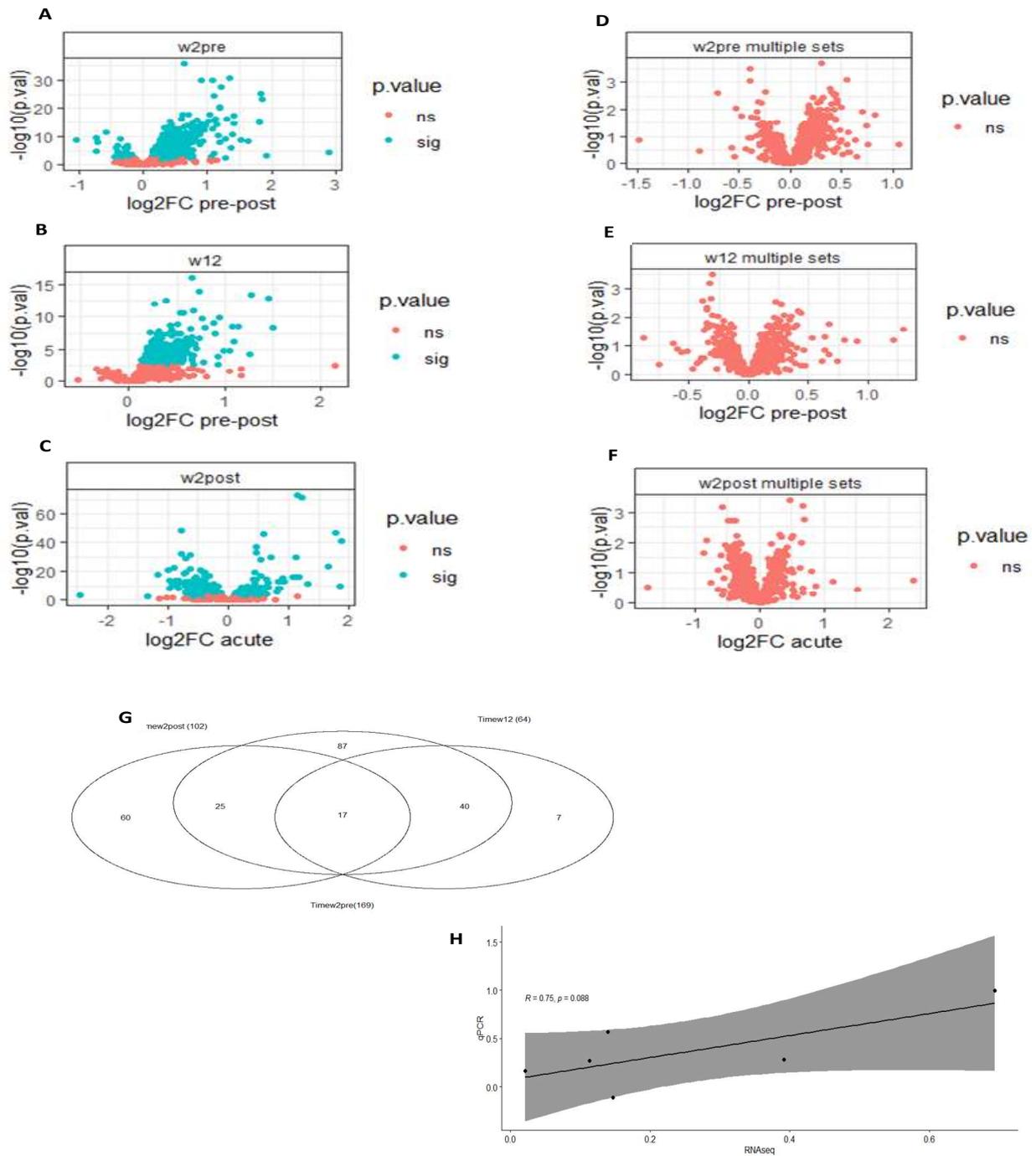


Figure 1. Results from pooled (both legs combined) analysis. Differentially expressed (DE) lncRNAs. A, B & C: shows DE lncRNAs that are DE log₂fold change with single set at timepoint w2pre, w12 and w2post respectively. D, E & F: shows DE lncRNAs that are DE log₂fold change with multiple set at timepoint w2pre, w12 and w2post respectively. G: A Venn diagram that visualizes shared DE lncRNAs across different timepoints. 17 transcripts are shared between all three timepoints. H: Correlation between RNA-seq and qPCR data based on data from timepoint w12.

Table 2. Top 5 lncRNAs, from all timepoints sorted on log2foldchange and P.adj. P.adj is adjusted p-value with “fdr” method in R. Not one of the lncRNAs were assigned with entrezgene id. “Function” is known functions described in the referenced article. The arrows in “Expected consequence for muscle cell signaling” means up or downregulating, stimulating or suppressing. Sp = species, Ref = references. NA = not available.

	<i>Ensemble gene id (Hgnc-symbol)</i>	<i>log2FC</i>	<i>P.adj</i>	<i>Function</i>	<i>Mode of action</i>	<i>Expected consequence for muscle cell signaling</i>	<i>Sp.</i>	<i>Ref.</i>
<i>Resistance training-related responses in rested-state muscle</i>	W2pre							
	ENSG00000214548 (MEG3)	0.64	<i>2.00E-33</i>	Tumor supressor.	Stimulates expression of growth differentiation factor 15 (GDF15)	Stimulates p53	Human	Zhou et al. (2007)
	ENSG00000172965 (MIR4435-2HG)	1.33	<i>8.00E-29</i>	Lung cancer cells proliferation.	Proteasome system	β-catenin ↑ signaling	Human	Qian et al. (2018)
	ENSG00000204054 (LINC00963)	1.08	<i>4.40E-28</i>	Promoting the proliferative ability of HCC cells	Promote protein expression	Activating PI3K/AKT pathway,	Human	Wu, Tian, An, Guan, and Hao (2018)
	ENSG00000230630 (DNM3OS)	0.9	<i>4.40E-28</i>	Growth and skeletal development in mice	Antisense	miR-199a, miR-199a, and miR-214 ↓	Mice	Watanabe et al. (2008)
ENSG00000130600 (H19)	1.2	<i>8.70E-26</i>	Acelerate muscle differentiation.	Molecular sponge	Regulates major let-7family of microRNAs.	Human and Mice	Kallen et al. (2013)	
<i>Resistance training-related responses in acute-state muscle</i>	w12							
	ENSG00000286214	0.66	<i>1.40E-13</i>	NA	NA	NA	NA	NA
	ENSG00000250208 (FZD10-AS1)	0.73	<i>7.40E-12</i>	Non-small-cell lung carcinomas ?	Antisense		Human	Yu et al. (2015)
	ENSG00000260807 (CEROX1)	1.27	<i>1.80E-11</i>	Regulates catalytic activity mitochondrial complex 1	Blocks effect of microRNA	miR-488-3p	Human	Sirey et al. (2019)
	ENSG00000286191	1.45	<i>4.10E-11</i>	NA	NA	NA	NA	NA
ENSG00000272168 (CASC15)	0.67	<i>2.30E-09</i>	Tumor supressor	Tumor promoting properties	siRNAs	Human	Lessard et al. (2015)	
w2post								
ENSG00000259820	1.14	<i>1.00E-70</i>	NA	NA	NA	NA	NA	
ENSG00000221817 (PPP3CB-AS1)	1.23	<i>1.80E-69</i>	May regulate Hutingtons disease and Parkinsons disease	Antisense	NA	Human	B. Hu et al. (2016)	
ENSG00000270605	-0.78	<i>4.80E-46</i>	NA	NA	NA	NA	NA	
ENSG00000242902 (FLNC-AS1)	1.78	<i>3.10E-45</i>	NA	Antisense	NA	NA	NA	
ENSG00000212719 (LINC02693)	0.6	<i>2.30E-44</i>	NA	NA	NA	NA	NA	

Table 3. Shared DE lncRNAs across all three timepoints with estimate and p-value adjusted. Shared across all three timepoints means that they are differentially expressed at all three timepoints. P.adj is adjusted p-value with “fdr” method in R. Not one of the lncRNAs were assigned with entrezgene id. “Function” is known functions described in the referenced article. The arrows in “Expected consequence for muscle cell signaling” means up or downregulating, stimulating or suppressing. Sp = species, Ref = references

<i>Ensemble gene id (Hgnc symbol)</i>	<i>W2pre log2FC (p.adjust)</i>	<i>W2post log2FC (p.adjust)</i>	<i>W12 log2FC (p.adjust)</i>	<i>Function</i>	<i>Mode of action</i>	<i>Expected consequence for muscle cell signaling</i>	<i>Sp.</i>	<i>Ref.</i>
ENSG00000116652 (DLEU2L)	1.17 (9.20E-10)	-0.71 (7.10E-06)	1.15 (2.90E-07)	Accelerates hepato-cellular carcinoma (HCC)	Binds to EZH2	Aggravates proliferation and migration in HCC	Human	Guo et al. (2019)
ENSG00000130600 (H19)	1.2 (8.70E-26)	-0.61 (3.70E-06)	0.58 (1.30E-05)	Accelerate muscle differentiation.	Molecular sponge	Regulates major let-7family of microRNAs.	Human and Mice	Kallen et al. (2013)
ENSG00000205056 (LINC02397)	1.3 (6.80E-15)	-0.97 (4.90E-13)	0.9 (1.90E-06)	Increases survival of melanoma metastasis	NA	NA	Human	L. Sun et al. (2019)
ENSG00000224361	1.42 (1.40E-13)	-0.5 (6.70E-05)	1.08 (3.20E-07)	NA	NA	NA	NA	NA
ENSG00000224609	1.11 (9.00E-15)	-0.89 (6.70E-22)	0.94 (1.90E-08)	NA	NA	NA	NA	NA
ENSG00000225613 (LINCMD1)	1.19 (1.10E-07)	-0.82 (4.20E-08)	0.1 (2.20E-04)	Controls muscle differentiation	Competing endogenous RNA	Controls miR-133 and miR-135	Human	Cesana et al. (2011)
ENSG00000229821	1.27 (8.80E-09)	0.64 (2.30E-04)	0.67 (7.30E-03)	NA	NA	NA	NA	NA
ENSG00000230438 (SERPINB9P1)	0.75 (1.00E-07)	-0.61 (7.90E-10)	0.69 (2.90E-05)	Affects post-menopausal osteoporosis.	Down-regulated	Interaction with mRNA	Human	S. Wang (2020)
ENSG00000237499	1.14 (5.30E-15)	-0.83 (4.90E-13)	0.88 (1.20E-07)	NA	NA	NA	NA	NA
ENSG00000260793	0.82 (1.10E-12)	-0.5 (3.90E-08)	0.59 (9.00E-06)	NA	NA	NA	NA	NA
ENSG00000260807 (CEROX1)	1.42 (2.60E-16)	-0.66 (3.60E-05)	1.28 (1.80E-11)	Regulates catalytic activity mitochondrial complex 1	Blocks effect of microRNA	miR-488-3p	Human	Sirey et al. (2019)
ENSG00000260966	0.76 (4.50E-12)	-0.53 (3.70E-09)	0.51 (1.40E-04)	NA	NA	NA	NA	NA
ENSG00000263873	1.8 (3.50E-14)	-0.7 (3.90E-05)	1.05 (1.90E-04)	NA	NA	NA	NA	NA
ENSG00000265206	1.15 (1.90E-13)	-0.54 (7.70E-04)	0.64 (4.50E-04)	NA	NA	NA	NA	NA
ENSG00000266923	1.4 (4.80E-10)	-0.97 (7.00E-08)	1.09 (2.00E-05)	NA	NA	NA	NA	NA
ENSG00000273812	0.78 (1.70E-05)	0.84 (5.00E-08)	0.65 (2.00E-03)	NA	NA	NA	NA	NA
ENSG00000286191	1.86 (6.20E-22)	-1.16 (4.60E-17)	1.45 (4.10E-11)	NA	NA	NA	NA	NA

4. Discussion

To the best of the authors' knowledge, this is one of the first studies to investigate global lncRNA expression in skeletal muscle in humans, and the first study to investigate how their expression is affected by resistance training. RNA-seq analysis identified 1400 lncRNAs in muscle tissue and somewhere between 10 and 15 percent were differentially expressed. The total number of DE lncRNAs found are lower than others have found (X. Sun et al., 2016). Yang et al. (2016) discovered a higher number of lncRNAs both up and downregulated in muscle cells, but without the resistance training intervention. 169 lncRNAs were DE at twelve weeks. Most of the DE lncRNAs detected were upregulated. Many of the lncRNAs were differentially expressed at more than one timepoint. As many as 17 lncRNAs were DE across all three timepoints, suggesting a role in muscle adaptations to resistance training. Indeed, some of these genes have either previously been shown to play a role in muscle differentiation and growth in animal (and in vitro) or to be involved in regulating pathways that have previously ascribed central roles in muscle plasticity. Little is known of lncRNAs function in muscle tissue, and few, if any, articles have studied DE across different timepoints after a training intervention period. Gene counts increased by 43-53% throughout the whole dataframe (Khan et al., 2020). Suggesting that the transcriptome increased significantly, the collected poly-RNA expression increased by 43-53%. No significant difference between low and moderate volume training was detected in the data analysis, which was in accordance with the findings in Khan et al. (2020). The study of Khan et al. (2020) is based on the same RNA-seq data as this study. They describe very similar responses to low and moderate volume. The resistance training intervention resulted in the expected increase in muscle strength and muscle mass. Higher training volume was associated with increased adaptations (Khan et al., 2020).

The lncRNAs have not been mapped with a GO analysis, thus the information that can be found are based on individual articles. Several of the top 5 genes are indicated to influence cellular plasticity in human and animal models through signaling pathways. Such as p53, β -catenin and PI3K/AKT pathway. These signaling pathways are also important in muscle-plasticity (Qian et al., 2018; Wu et al., 2018; Zhou et al., 2007) Only 6 of the 17 DE lncRNAs that can be found at

all timepoints, are annotated with hgnc symbol. Thus, literature survey is possible on 6 of them. LincMD1 may be important for muscle biology and adaptation to resistance training Cesana et al. (2011). H19 is another interesting lncRNA, and the latter is part of the top 5 lncRNAs and the 17 DE lncRNAs that can be found at all timepoints. H19 is known to accelerate muscle differentiation in mice and probably also in human (Kallen et al., 2013), thus the rested state expression found in the RNA-seq data. The lncRNA MEG3 is interesting. It regulates p53, and thus indirectly regulates cellular homeostasis by the Mdm2 -p53 stress response pathway (Bartlett et al., 2014). Thus, MEG3 could be important to regulate cell growth. Two distinct patterns can be found in the 17 shared DE lncRNAs. Most of them have an up-down-up expression profile at timepoints w2pre, w2post & w12, respectively (table 3). Some of them have an up-up-up profile. The former pattern is probably due to negative feedback. The acute training responses blocks cell growth to allow the cell to repair before growth. The latter pattern is interesting and unexpected. It could be explained by some sort of positive feedback mechanism. Both of the patterns found indicates that the lncRNAs are exercise responsive. Many of the DE lncRNAs from (top 5 and shared 17) are novel, and not previously studied or annotated important functions in eukaryote cells. Much are still unknown. Rested state samples from timepoint w2pre and w12 expressed most of the DE lncRNAs with increased expression. This is according to Khan et al. (2020), Increased expression at rested state could be linked to general increase in mRNA expression (Khan et al., 2020).

Only six genes at timepoint w12 were used in the correlation test between qPCR and RNA-seq. Why the correlation coefficient was not higher, is probably due to the normalizing discrepancies between qPCR and RNA-seq. qPCR results were normalized with a tissue offset normalizing factor based on amount of tissue used in cDNA synthesis. RNA-seq (rested state results, from w2pre and w12) were normalized with the tissue offset normalizing factor. The acute results (from w2post) was normalized to lib size. Khan et al. (2020) compared different RNA-seq data normalizing approaches. They advocate tissue weight normalizing approach, as done in this study. Except from when normalizing acute data. Minimal muscle growth is expected after only one training bout, and lib size normalizing of the data are therefore sufficient (Khan et al., 2020).

Ten different lncRNAs were analyzed with qPCR, but only 6 (LincMD1, GAS5, LINP1, PVT1, LINC01405 and RMRP) were found in the RNA-seq data. Of them all were differentially expressed, at one timepoint, except PVT1. To rule out coding RNA, the ensemble database was used. A search was made, and only genes annotated as lncRNAs were chosen. This method may be biased, and important lncRNAs that are not annotated, may be wrongfully sorted out. Few lncRNAs with entrezgene id ID made it difficult to conduct a proper GO analysis. In this study the ORA method was used to assess time-based effect on pooled (both legs combined) resistance training data. Another option was to run a rank-based GO analysis to explore the volume effect on muscle adaptations after resistance training (Khan et al., (2020)). The ORA test was impossible to implement due to lncRNAs not annotated. The same applied to the rank-based analysis. Most of the lncRNAs have hgnc numbers and literature surveys are possible to conduct, but little is known of their function and the studies are scarce at best. The lncRNAs functions in muscle adaptations are probably much more complex, but difficult to explore. No difference was detected in the statistical analysis between low and moderate training volume. That is probably due to measure difficulties as mentioned in Khan et al. (2020). The small bashful changes in the microbiology are difficult to measure. Sequencing depth determines how many genes that are recognized. One may stipulate that more lncRNAs could be discovered if the sequencing depth was adjusted. The correlation coefficient between qPCR and RNA-seq data was significant with a p-value of 0.088. Thus, the RNA-seq and the qPCR results yielded mostly comparable results. qPCR analysis is considered the gold standard when analyzing gene expression (Adamski et al., 2014) thus the quality on the RNA-seq data was high. The RNA sequencing method used, only recognizes polyA lncRNAs. Thus, important lncRNAs could be eluded from the analysis. This study utilized only a part of the whole RNA-seq data frame. DE analysis was performed on lncRNAs only, thus the analysis could be biased.

Many of the DE lncRNAs have been indicated to be important in muscle plasticity, but the mechanisms are largely unknown. Other DE lncRNAs remain completely uncharacterized, thus future studies are needed. LncRNAs are gaining status, and more research are conducted on their different functions. In cancer research, knockout studies of different lncRNAs shows promising treatment paths. The expression of different lncRNAs may also be used for predictive purposes. It

may be possible to predict the adaptation to specific strength training based on expression of lncRNAs that induces muscle growth or muscle differentiation. Knockout studies are difficult to conduct on humans, and therefore many studies are conducted on mice.

5. Conclusion

Despite the fact that GO-analyses could not be properly performed on these genes, making it difficult to decipher their biological role, several of the response lncRNAs have previously been ascribed roles as regulators of muscle plasticity. Between timepoint w2pre and w12, 169 lncRNAs were differentially expressed. 17 DE lncRNAs were found on all three timepoints, suggesting that they are important in muscle adaptations to resistance training. Resistance training with low and moderate volume resulted in similar changes in lncRNA expression, reiterating on the fact that the different volume conditions do not lead to substantial differences in cellular phenotypes measured per unit muscle tissue (though higher volume is associated with larger increases in muscle mass). More research is needed to expand the entrezgene id database and allocate gene annotations.

References

- Adamski, M. G., Gumann, P., & Baird, A. E. (2014). A Method for Quantitative Analysis of Standard and High-Throughput qPCR Expression Data Based on Input Sample Quantity. *PLoS One*, *9*(8), e103917. doi:10.1371/journal.pone.0103917
- Bartlett, J. D., Close, G. L., Drust, B., & Morton, J. P. (2014). The emerging role of p53 in exercise metabolism. *Sports Med*, *44*(3), 303-309. doi:10.1007/s40279-013-0127-9
- Bodine, S. C., Stitt, T. N., Gonzalez, M., Kline, W. O., Stover, G. L., Bauerlein, R., . . . Yancopoulos, G. D. (2001). Akt/mTOR pathway is a crucial regulator of skeletal muscle hypertrophy and can prevent muscle atrophy in vivo. *Nature Cell Biology*, *3*(11), 1014-1019. doi:10.1038/ncb1101-1014
- Brannan, C. I., Dees, E. C., Ingram, R. S., & Tilghman, S. M. (1990). The product of the H19 gene may function as an RNA. *10*(1), 28-36. doi:10.1128/mcb.10.1.28
- Cabili, M. N., Trapnell, C., Goff, L., Koziol, M., Tazon-Vega, B., Regev, A., & Rinn, J. L. (2011). Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes & Development*, *25*(18), 1915-1927. doi:10.1101/gad.17446611
- Cesana, M., Cacchiarelli, D., Legnini, I., Santini, T., Sthandier, O., Chinappi, M., . . . Bozzoni, I. (2011). A Long Noncoding RNA Controls Muscle Differentiation by Functioning as a Competing Endogenous RNA. *Cell*, *147*(2), 358-369. doi:10.1016/j.cell.2011.09.028
- Chaillou, T., Kirby, T. J., & McCarthy, J. J. (2014). Ribosome Biogenesis: Emerging Evidence for a Central Role in the Regulation of Skeletal Muscle Mass. *Journal of Cellular Physiology*, *229*(11), 1584-1594. doi:10.1002/jcp.24604
- Chen, X., Sun, Y., Cai, R., Wang, G., Shu, X., & Pang, W. (2018). Long noncoding RNA: multiple players in gene expression. *BMB reports*, *51*(6), 280-289. doi:10.5483/bmbrep.2018.51.6.025
- Damas, F., Phillips, S. M., Lixandrão, M. E., Vechin, F. C., Libardi, C. A., Roschel, H., . . . Ugrinowitsch, C. (2016). Early resistance training-induced increases in muscle cross-sectional area are concomitant with edema-induced muscle swelling. *European Journal of Applied Physiology*, *116*(1), 49-56. doi:10.1007/s00421-015-3243-4
- Douglas, Kelly, Chang, C.-L., Catherine, Benjamin, John, . . . Eric. (2015). A Micropeptide Encoded by a Putative Long Noncoding RNA Regulates Muscle Performance. *Cell*, *160*(4), 595-606. doi:10.1016/j.cell.2015.01.009
- Egan, B., & Zierath, J. R. (2013). Exercise metabolism and the molecular regulation of skeletal muscle adaptation. *Cell Metab*, *17*(2), 162-184. doi:10.1016/j.cmet.2012.12.012
- Folland, J. P., & Williams, A. G. (2007). The Adaptations to Strength Training. *Sports Medicine*, *37*(2), 145-168. doi:10.2165/00007256-200737020-00004
- Fry, A. C. (2004). The Role of Resistance Exercise Intensity on Muscle Fibre Adaptations. *Sports Medicine*, *34*(10), 663-679. doi:10.2165/00007256-200434100-00004
- Guo, Y., Bai, M., Lin, L., Huang, J., An, Y., Liang, L., . . . Huang, W. (2019). LncRNA DLEU2 aggravates the progression of hepatocellular carcinoma through binding to EZH2. *Biomedicine & Pharmacotherapy*, *118*, 109272. doi:10.1016/j.biopha.2019.109272

- Hammarström, D., Øfsteng, S., Koll, L., Hanestadhaugen, M., Hollan, I., Apró, W., . . . Ellefsen, S. (2020). Benefits of higher resistance-training volume are related to ribosome biogenesis. *The Journal of Physiology*. doi:10.1113/jp278455
- Han, P., Li, W., Lin, C.-H., Yang, J., Shang, C., Nurnberg, S. T., . . . Chang, C.-P. (2014). A long noncoding RNA protects the heart from pathological hypertrophy. *Nature*, *514*(7520), 102-106. doi:10.1038/nature13596
- Hester, R. L., Iliescu, R., Summers, R., & Coleman, T. G. (2011). Systems biology and integrative physiological modelling. *The Journal of Physiology*, *589*(5), 1053-1060. doi:10.1113/jphysiol.2010.201558
- Hoppeler, H. (2016). Molecular networks in skeletal muscle plasticity. *J Exp Biol*, *219*(Pt 2), 205-213. doi:10.1242/jeb.128207
- Hu, B., Huo, Y., Chen, G., Yang, L., Wu, D., & Zhou, J. (2016). Functional prediction of differentially expressed lncRNAs in HSV-1 infected human foreskin fibroblasts. *13*(1). doi:10.1186/s12985-016-0592-5
- Hu, G., Niu, F., Humburg, B. A., Liao, K., Bendi, S., Callen, S., . . . Buch, S. (2018). Molecular mechanisms of long noncoding RNAs and their role in disease pathogenesis. *Oncotarget*, *9*(26), 18648-18663. doi:10.18632/oncotarget.24307
- Hughes, D. C., Ellefsen, S., & Baar, K. (2018). Adaptations to Endurance and Strength Training. *Cold Spring Harb Perspect Med*, *8*(6). doi:10.1101/cshperspect.a029769
- Jason, Spacek, D. V., & Michael. (2015). High-Throughput Sequencing Technologies. *Molecular Cell*, *58*(4), 586-597. doi:10.1016/j.molcel.2015.05.004
- Jin, J. J., Lv, W., Xia, P., Xu, Z. Y., Zheng, A. D., Wang, X. J., . . . Zuo, B. (2018). Long noncoding RNA SYISL regulates myogenesis by interacting with polycomb repressive complex 2. *Proceedings of the National Academy of Sciences*, *115*(42), E9802-E9811. doi:10.1073/pnas.1801471115
- Kallen, A., Zhou, X.-B., Xu, J., Qiao, C., Ma, J., Yan, L., . . . Huang, Y. (2013). The Imprinted H19 LncRNA Antagonizes Let-7 MicroRNAs. *Molecular Cell*, *52*(1), 101-112. doi:10.1016/j.molcel.2013.08.027
- Khan, Y., Hammarström, D., Rønnestad, B. R., Ellefsen, S. & Ahmad, R. (2020). Increased biological relevance of transcriptome analyses in human skeletal muscle using a model-specific pipeline. *Not published*.
- Kraemer, W. J., Ratamess, N. A., & French, D. N. (2002). Resistance training for health and performance. *Current sports medicine reports*, *1*(3), 165-171.
- Kuang, S., Kuroda, K., Le Grand, F., & Rudnicki, M. A. (2007). Asymmetric Self-Renewal and Commitment of Satellite Stem Cells in Muscle. *Cell*, *129*(5), 999-1010. doi:10.1016/j.cell.2007.03.044
- Kung, J. T. Y., Colognori, D., & Lee, J. T. (2013). Long Noncoding RNAs: Past, Present, and Future. *Genetics*, *193*(3), 651-669. doi:10.1534/genetics.112.146704
- Li, Y., Chen, X., Sun, H., & Wang, H. (2018). Long non-coding RNAs in the regulation of skeletal myogenesis and muscle diseases. *Cancer Letters*, *417*, 58-64. doi:10.1016/j.canlet.2017.12.015
- Maglott, D., Ostell, J., Pruitt, K. D., & Tatusova, T. (2011). Entrez Gene: gene-centered information at NCBI. *Nucleic Acids Research*, *39*(Database), D52-D57. doi:10.1093/nar/gkq1237

- McCarthy, J. J., & Esser, K. A. (2010). Anabolic and catabolic pathways regulating skeletal muscle mass. *Current Opinion in Clinical Nutrition and Metabolic Care*, 13(3), 230-235. doi:10.1097/mco.0b013e32833781b5
- McCroskery, S., Thomas, M., Maxwell, L., Sharma, M., & Kambadur, R. (2003). Myostatin negatively regulates satellite cell activation and self-renewal. *The Journal of Cell Biology*, 162(6), 1135-1147. doi:10.1083/jcb.200207056
- Mirzoev, T. M., & Shenkman, B. S. (2018). Regulation of Protein Synthesis in Inactivated Skeletal Muscle: Signal Inputs, Protein Kinase Cascades, and Ribosome Biogenesis. *Biochemistry (Moscow)*, 83(11), 1299-1317. doi:10.1134/S0006297918110020
- Morgan, J. E., & Partridge, T. A. (2003). Muscle satellite cells. *The International Journal of Biochemistry & Cell Biology*, 35(8), 1151-1156. doi:10.1016/s1357-2725(03)00042-6
- Munn, J., Herbert, R. D., & Gandevia, S. C. (2004). Contralateral effects of unilateral resistance training: a meta-analysis. *Journal of Applied Physiology*, 96(5), 1861-1866. doi:10.1152/jappphysiol.00541.2003
- O'Neil, T. K., Duffy, L. R., Frey, J. W., & Hornberger, T. A. (2009). The role of phosphoinositide 3-kinase and phosphatidic acid in the regulation of mammalian target of rapamycin following eccentric contractions. *The Journal of Physiology*, 587(14), 3691-3701. doi:10.1113/jphysiol.2009.173609
- Pickard, M. R., Mourtada-Maarabouni, M., & Williams, G. T. (2013). Long non-coding RNA GAS5 regulates apoptosis in prostate cancer cell lines. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, 1832(10), 1613-1623. doi:10.1016/j.bbadis.2013.05.005
- Ponting, C. P., Oliver, P. L., & Reik, W. (2009). Evolution and functions of long noncoding RNAs. *Cell*, 136(4), 629-641. doi:10.1016/j.cell.2009.02.006
- Qian, H., Chen, L., Huang, J., Wang, X., Ma, S., Cui, F., . . . Zheng, G. (2018). The lncRNA MIR4435-2HG promotes lung cancer progression by activating β -catenin signalling. *Journal of Molecular Medicine*, 96(8), 753-764. doi:10.1007/s00109-018-1654-5
- Ralston, G. W., Kilgore, L., Wyatt, F. B., & Baker, J. S. (2017). The Effect of Weekly Set Volume on Strength Gain: A Meta-Analysis. *Sports Medicine*, 47(12), 2585-2601. doi:10.1007/s40279-017-0762-7
- Ritz, C., & Spiess, A.-N. (2008). qpcR: an R package for sigmoidal model selection in quantitative real-time polymerase chain reaction analysis. *Bioinformatics*, 24(13), 1549-1551.
- RStudio Team (Producer). (2016). Studio: Integrated Development for R. Retrieved from <http://www.rstudio.com/>
- Sun, L., Guan, Z., Wei, S., Tan, R., Li, P., & Yan, L. (2019). Identification of long non-coding and messenger RNAs differentially expressed between primary and metastatic melanoma. *Frontiers in genetics*, 10, 292.
- Sun, X., Li, M., Sun, Y., Cai, H., Lan, X., Huang, Y., . . . Chen, H. (2016). The developmental transcriptome sequencing of bovine skeletal muscle reveals a long noncoding RNA, lncMD, promotes muscle differentiation by sponging miR-125b. *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, 1863(11), 2835-2845. doi:10.1016/j.bbamcr.2016.08.014

- Team, R. C. (2013). R: A language and environment for statistical computing. In: Vienna, Austria.
- Tseng, Y.-Y., Moriarity, B. S., Gong, W., Akiyama, R., Tiwari, A., Kawakami, H., . . . Bagchi, A. (2014). PVT1 dependence in cancer with MYC copy-number increase. *Nature*, *512*(7512), 82-86. doi:10.1038/nature13311
- Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B. C., Remm, M., & Rozen, S. G. (2012). Primer3—new capabilities and interfaces. *Nucleic Acids Research*, *40*(15), e115-e115. doi:10.1093/nar/gks596
- Wang, G.-Q., Wang, Y., Xiong, Y., Chen, X.-C., Ma, M.-L., Cai, R., . . . Pang, W.-J. (2016). Sirt1 AS lncRNA interacts with its mRNA to inhibit muscle formation by attenuating function of miR-34a. *6*, 21865. doi:10.1038/srep21865
- Wang, S. (2020). Investigation of long non-coding RNA expression profiles in patients with post-menopausal osteoporosis by RNA sequencing. *Experimental and Therapeutic Medicine*. doi:10.3892/etm.2020.8881
- Wang, X., Peng, L., Gong, X., Zhang, X., Sun, R., & Du, J. (2018). LncRNA-RMRP promotes nucleus pulposus cell proliferation through regulating miR-206 expression. *Journal of Cellular and Molecular Medicine*. doi:10.1111/jcmm.13817
- Wu, J. H., Tian, X. Y., An, Q. M., Guan, X. Y., & Hao, C. Y. (2018). LINC00963 promotes hepatocellular carcinoma progression by activating PI3K/AKT pathway. *Eur Rev Med Pharmacol Sci*, *22*(6), 1645-1652.
- Yamaguchi, K. D., Ruderman, D. L., Croze, E., Wagner, T. C., Velichko, S., Reder, A. T., & Salamon, H. (2008). IFN- β -regulated genes show abnormal expression in therapy-naïve relapsing–remitting MS mononuclear cells: Gene expression analysis employing all reported protein–protein interactions. *195*(1-2), 116-120. doi:10.1016/j.jneuroim.2007.12.007
- Yang, F., Lyu, S., Dong, S., Liu, Y., Wang, O., & Zhang, X. (2016). Expression profile analysis of long noncoding RNA in HER-2-enriched subtype breast cancer by next-generation sequencing and bioinformatics. *761*. doi:10.2147/ott.s97664
- Yu, G., Wang, L.-G., Han, Y., & He, Q.-Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics: a journal of integrative biology*, *16*(5), 284-287.
- Zhang, Y., He, Q., Hu, Z., Feng, Y., Fan, L., Tang, Z., . . . Zhang, L. (2016). Long noncoding RNA LINP1 regulates repair of DNA double-strand breaks in triple-negative breast cancer. *Nature Structural & Molecular Biology*, *23*(6), 522-530. doi:10.1038/nsmb.3211
- Zhou, Y., Zhong, Y., Wang, Y., Zhang, X., Batista, D. L., Gejman, R., . . . Klibanski, A. (2007). Activation of p53 by MEG3 Non-coding RNA. *Journal of Biological Chemistry*, *282*(34), 24731-24742. doi:10.1074/jbc.m702029200

Appendix 1, Khan et al. (2020)

Increased biological relevance of transcriptome analyses in human skeletal muscle using a model-specific pipeline

Yusuf Khan^{1,2,£}, Daniel Hammarström^{2,3,£}, Bent R. Rønnestad², Stian Ellefsen^{2,4}, Rafi Ahmad^{1*}

¹Department of Biotechnology, Inland Norway University of Applied Sciences, Holsetgata 22, 2317 Hamar, Norway.

²Section for Health and Exercise Physiology, Department of Public Health and Sport Sciences, Inland Norway.

³Swedish School of Sport and Health Sciences, Box 5626, SE-114 86 Stockholm, Sweden.

⁴Innlandet Hospital Trust, Postboks 990, 2629 Lillehammer, Norway.

£ Authors contributed equally

*Corresponding author rafi.ahmad@inn.no

Abstract

Background: Human skeletal muscle responds to weight-bearing exercise with large inter-individual differences. Investigation of transcriptome responses could improve our understanding of this variation. Interpretation of such data relies on appropriate selection of analytical tools. Here, we established a skeletal muscle-specific bioinformatic pipeline for transcriptome analyses. We then used it to assess dose-dependent changes in transcriptome responses to resistance training in m. vastus lateralis, using three different normalization strategies (tissue-offset, effective library size and naïve). Briefly, 25 young participants performed low- and moderate-volume resistance exercise for twelve weeks (31 sessions), allocated to the two lower limbs in a randomized manner. Bilateral muscle biopsies were sampled before and after the intervention (in a rested state), as well as before and after the fifth training session. Total RNA was extracted and subjected to RNA sequencing.

Results: Bioinformatic tools were selected based on read quality, observed gene counts, methodological variation between paired observations, and correlations between RNA abundance and protein expression of myosin heavy chain family proteins. Overall, training led to robust transcriptome changes, with the number of differentially expressed (DE) genes ranging from 603-5110, varying with time point and normalization strategy. In all models, >82% of DE genes increased in rested-state muscle (range 82-99%). After accounting for the amounts of muscle tissue used in library preparation (tissue offset), DE analysis revealed dose-dependent increases for 21 genes in the early phase of resistance training. Many of these were related to extracellular matrix function (BGN, CILP, COL6A3, COL4A2, COL14A1, ELN, FN1, SCARA3, SPON1, SULF1), genes that are involved in elasticity, growth and maturation in muscle. No difference was seen between volume conditions at twelve weeks. In contrast, normalization to effective library size showed a counterintuitive reversed dose-dependence for many genes after the training period, resembling a non-normalized model.

Conclusions: To achieve biologically meaningful data from transcriptome analyses of skeletal muscle subjected to altered growth conditions, normalization procedures need to account for global changes in rRNA and mRNA expression. These recommendations are likely applicable to studies

of other cell types and model systems undergoing increased or arrested growth. **Keywords:** RNA-seq, skeletal muscle, bioinformatics pipeline, normalization

Introduction

Skeletal muscle is a highly adaptable tissue that responds to environmental stress by altering growth rates and differentiation processes. During resistance training, signaling cascades that stimulate muscle plasticity are triggered. Upon repeated exposures, this facilitates growth and a phenotypic shift in a metabolically active direction [1], with the opposite happening during inactivity [2]. Despite this generalized view, muscle responsiveness and plasticity vary, both in response to different resistance-training protocols [3] and, perhaps more importantly, between individuals [4, 5]. Selected individuals show a near-complete absence of muscle growth after prolonged resistance training, which markedly reduces the functional and health-beneficial outcomes of such interventions [4, 5]. Currently, little is known about the etiology of this variation. However, it is usually associated with muscle phenotypic traits [6–8], which implies interactions with environmental factors, genetics, epigenetics, and composites of the inner physiological milieu [9, 10]. This multifaceted origin makes the training-response-spectrum difficult to study directly, with each of the underlying factors offering limited explanatory value alone [11]. Instead, a more indirect approach is necessary, whereby the combined effects of the factors are targeted by studying global patterns of mRNA, protein expression, and skeletal muscle biology.

Previous studies have investigated transcriptome responses to acute resistance exercise [12–14] and chronic resistance training [12, 13, 15–18], as well as described associations between transcriptome characteristics and degrees of muscle growth [18, 19], and function [20, 21]. Whereas these studies have merited interesting findings, they lack clear coherences in terms of differential expression events, even for classical exercise-inducible genes such as PGC1 α [22]. This lack of clear coherence is potentially due to a combination of issues such as differences in study design and methods for synthesis and analysis of transcriptome data. Variability in transcriptome responses to exercise can be attributed to different exercise protocols (e.g. differences in exercise-volume or intensity). This makes it difficult to discern a general transcriptome exercise response, as training variables (such as volume or intensity) are not standardized between studies. Additionally, biological heterogeneity between research participants affects the signal-to-noise ratio, making it difficult to discern effects of single independent factors such as training variables. Design stage decisions such as the use of within-participant designs (discussed elsewhere, e.g. [3, 23]) are likely to reduce this variation and to provide transcriptome data with increased biological

meaningfulness. However, to fully exploit the potential of any study design we need to identify an appropriate pipeline for performing transcriptome analyses to ensure biologically valid interpretation of data. This entails identifying potential violations of common assumptions caused by the experimental model at hand, such as relating to data normalization [24, 25]. In cell models that exhibit high degrees of plasticity, gene expression events result in increased amounts of total RNA and mRNA transcripts per cell [26], specifically violating the assumption that most genes are not differentially expressed [24, 27]. At present, this perspective remains understudied in skeletal muscle subjected to increased mechanical stress such as resistance training, with no study addressing the need to account for such perspectives during transcriptome analyses. For transcriptome data to provide adequate biological information about a given experimental set-up, numerous bioinformatic steps need to be adopted in a customized manner [28]. Of these steps, data normalization is particularly decisive [25] to study gene expression, as it aims to transform naïve transcript counts into biologically meaningful results. This essentially means expressing them as per-cell abundances [27]. For most experimental models, this is equivalent to providing transcript-to-total RNA ratios, given the accuracy of the assumption that total RNA levels remain stable between conditions on a per-unit-cell or per-unit-tissue basis [27]. In skeletal muscle, this assumption is violated after resistance training, as total RNA content increases markedly on a per-unit-weight basis [3], with potential global changes also occurring in the mRNA pool, though this remains unaccounted. The extent to which total RNA, and therefore ribosomal RNA, increases, coincides with the increase in muscle mass [3, 7], underlining its importance for cellular growth but also its inevitable presence as a potential confounding factor in RNA sequencing experiments.

In this study, we aimed to (i) establish a bioinformatic pipeline specific for analysis of RNA-seq data from skeletal muscles, to (ii) explore the effects of using different normalization strategies for analyzing skeletal muscle tissue subjected to resistance training, and to (iii) identify genes responding to moderate, compared to low exercise volume. To achieve these aims, we utilized RNA-seq data generated from a within-participant study, comparing the effects of low and moderate volume training, as previously described [3]. Also, myosin heavy chain protein expression, quantified by immunohistochemistry was used to validate RNA quantification tools.

Results

For the RNA-seq analyses present here, a subset of participants was selected based on RNA quality measurements. Participants having functional and molecular responses were selected [3]. Twenty-five participants had a full set of RNA-samples with RNA quality indicator scores of ≥ 7 . RNA quality scores were not associated with RNA yield (Figure 1C). Also, the subset of participants selected for RNA-seq analyses did not differ from excluded participants concerning training-induced changes in muscle mass, illustrated as the lean body-mass change in the two-volume conditions used in the study (Figure 1D). Twelve weeks of training with low- and moderate-volume led to greater muscle hypertrophy in response to moderate- compared to low-volume training ($\sim 3.5\%$ vs. $\sim 2.0\%$, Figure 1E). Greater muscle hypertrophy in the multiple-set condition coincided with greater strength gains ($\sim 25\%$ vs. $\sim 19\%$, Figure 1F). This is in agreement with what we have previously reported from the full cohort [3].

Bioinformatic pipeline for analysis of RNA-seq data from skeletal muscles

To select the most appropriate tools for downstream bioinformatic analyses, we first proceeded with comparing Trimmomatic and Trim Galore, which are the two commonly used tools for quality filtering. [29,33,34] Quality scores were generally better with Trimmomatic compared to Trim Galore, which did not improve scores over non-filtered data (Figure 1G). Filtered reads were aligned to the human genome using three alignment-based methods (including HISAT2, STAR, RSEM, all used with Bowtie 2) and two non-alignment-based methods (kallisto and Salmon). RSEM, Salmon, and kallisto all showed similar characteristics in terms of gene counts with a bimodal distribution of counts resulting in a larger subset of detected genes after expression filtering compared to STAR and HISAT2 (Figure 1H). Using a selection of genes with known robust expression across tissues [30] it was confirmed that the differences between methods in genes grouped per detected counts with Salmon, kallisto, and RSEM had a higher proportion of genes with high counts (Figure 3C). Using this selection of genes, RSEM performed better in terms of technical variation expressed as a typical log fold difference between bilateral biopsies sampled prior to the intervention (Figure 1I). Also, a higher variation in HISAT2 and STAR coincided with lower correlations between myosin heavy chain family RNA and protein (Figure 2A and B).

Overall, these comparisons showed similar technical performance of RSEM, kallisto, and Salmon in terms of variability and biological validity. The slightly lower average variation between paired samples in RSEM led us to proceed further using this method.

Effects of normalization strategies on transcriptomic data analysis from skeletal muscle under hypertrophic stress

As previously reported [3], resistance training led to an increase in total RNA per-unit tissue weight. As an equal amount of total RNA was used for preparing RNA-seq libraries, the amount of muscle tissue used in library preparations decreased at 2 and 12 weeks of single-set training by 13% and 9%, respectively. The decrease was more pronounced in response to multiple-set training (-7.1%, 95% CI: [-12.9, -1.0] and -6.3%, 95% CI: [-11.8, -0.4] for multiple-sets vs. single-set at Week 2 and 12 respectively, Figure 3A). Although a smaller amount of muscle tissue was used to prepare libraries, the total effective library size was increased by 25% and 38% from before to 2 and 12 weeks of training, respectively (Figure 3B). The increase was not as pronounced in the moderate volume condition (-11%, 95% CI: [-22, 1.7] and -12%, 95% CI: [-24, 2.2] for multiple-set vs. single-set at week 2 and 12 respectively, Figure 4B). The differences were less pronounced when the effective library size was normalized to tissue weight (-4%, 95% CI: [-16, 9.7] and -6%, 95% CI: [-23, 14.1] for multiple-set vs. single-set at week 2 and 12 respectively, Figure 3C).

Identification of genes responding to moderate, compared to low exercise volume

Three models were used to compare normalization strategies. To account for the amount of tissue used in RNA-seq library preparation, tissue weight was included as an offset in the first model (tissue offset model) in addition to having the effective library size as a covariate in the model as previously suggested [31]. A second model only contained the effective library-size as a covariate (library-size normalization) and represented a scenario where normalization aimed to compare expression levels while accounting for technical variation in library preparation [24, 31]. Lastly a non-normalized model was used for comparison (naïve model).

In Week 2, when accounting for the amount of tissue used in library preparation, 21 genes were identified as differentially expressed (\log_2 fold-change $> |0.5|$ and FDR < 0.05), having higher expression in multiple-set compared to single-set (Figure 3D). Seven genes were identified as having higher expression in multiple-set compared to single-set when normalizing only to effective library size (Figure 3D). The gene set identified as up-regulated in multiple-set vs. single-set at week 2 in the effective library-size normalized model overlapped completely with the tissue offset model (Figure 3E). Rank based enrichment tests, using the minimum significant difference (MSD) identified gene ontology (GO) sets associated with the extracellular matrix (ECM) (Figure 3F, Table 2). The top ranked GO terms were also identified in over-representation tests (ORA) using DE (Figure 3F, Table 2). In contrast to the tissue offset model, seven genes were identified as more highly expressed in single-set vs. multiple-set in the non-normalized naïve model; among these three were shared with the library-size normalized model (Figure 3E). No significantly enriched GO term was identified among genes identified as highly expressed in single-set vs. multiple-set in either model. When comparing rank metrics, like fold-change, between models the effect of the different normalization scenarios became apparent. Controlling for the amount of tissue shifted the distribution of fold-changes in favor of multiple set (Figure 3G). Subsequently, gene sets appeared with higher expression in multiple- vs. single-set, exemplified with the GO set Collagen containing ECM in Figure 3G. The number of genes identified as DE from this GO term were higher in the tissue offset model, followed by the effective library size normalized model and the naïve model, where no genes from the category were identified as DE (Figure 3G).

At Week 12, no genes were identified as differentially expressed between single- vs. multiple- set training in the tissue offset model (Figure 3H). However, a small number of genes ($n = 4$) were identified as having higher expression in the single-set vs. multiple-set condition in the effective library size model of which two genes were shared with the naïve model (Figure 3I).

The effects of acute exercise were examined as changes pre- to post-exercise in the fifth training session. Only the effective library size model was used as we did not expect changes in total RNA to muscle mass in this short time span [Figueriedo rev]. A total of 707 and 1029 genes were identified as DE with higher and lower expression post-exercise respectively when both conditions were analyzed collectively. Up-regulated genes were associated with stress related GO terms (Figure 4B, Table 3). In contrast to rested state biopsies, ECM related GO terms were

identified as down-regulated following acute exercise (Figure 4B, Table 3). When comparing multiple- and single-set acute exercise. A single gene was identified as DE (RFT1, Figure 4C) that was reduced to a greater extent in response to multiple-set vs. single-set exercise. Although only this single gene showed DE between conditions, five GO terms were identified as significantly enriched with top ranked genes based on MSD. Among these five categories, three had genes with $MSD > 0$ indicating that the lower bound of 95% CI did not overlap no change. However these categories were not identified in gene set enrichment analysis based on fold-changes as differences between volume conditions were both negative and positive as indicated by the rug-plot in Figure 4D. Overall these analyses gives no strong indications of volume dependent regulation in the acute phase (1-h) after exercise.

When examining the overall effects of training, 3923, 1609 and 3875 genes were identified as having higher expression and 77, 289 and 100 genes were identified as having lower expression at Week 2 compared to pre-training in the tissue offset, library-size normalized and naïve models respectively (Figure 5A). Majority of identified DE genes found in the intersection between all models (Figure 5A lower panel). When comparing Week 12 to pre-training, 1733, 584 and 5108 genes were identified as more highly expressed and 2, 19 and 2 genes identified as having lower expression in the tissue offset, library-size normalized and naïve model respectively (Figure 5B). Here, the largest number of DE genes identified with higher expression post-training were found in the intersection between the tissue offset model and the naïve model (Figure 5B lower panel).

Enrichment analysis of time-effects in rested state samples (Week 2 vs. Week 0 and Week12 vs. Week 0) identified similar top-ranked GO terms associated with ECM structure, organization and synthesis as well as stress response (Table 2).

Discussion

In the present study, within participant model to study the effects of different training volumes on transcriptome responses. Similar to previous study [3] and here, shown that this protocol resulted in robust differences in training outcomes, in line with previous studies [32]. Despite changes in muscle mass and strength, only small differences were detected in transcriptome profiles between conditions. Arguably, identification of these differences was made possible by systematic selection of analytic tools in establishing our bioinformatic pipeline. As a first step we sought to select a

suitable read trimming method and compared two commonly used algorithms [29, 33, 34]. Read trimming is known to affect downstream alignment and k-mer search in reads [29]. Trimmomatic and TrimGalore were compared and Trimmomatic was found to provide better quality than Trimgalore (Figure 1G). Subsequent transcript quantification is based on genome or transcriptome mapping of trimmed reads. For mapping two genome based mapping tools (STAR [35] and HISAT2 [36]) and three transcript based mapping tools (RSEM [37], kallisto [38] and Salmon [39]) were used. To select the most suitable alignment tool, the relationship between myosin heavy chain mRNA and protein abundances were evaluated, which are known to correlate in resting human skeletal muscle [40–42]. The use of gene/protein-family normalization [41, 43] allowed us to specifically interrogate mRNA to protein relationships without the need of other normalization assumptions. With the assumption that improved biological resolution would manifest in stronger correlations between relative mRNA and protein abundances we noted that STAR and HISAT performed worse than their transcriptome-mapping counterparts (RSEM, kallisto and Salmon). In the same analysis, a clear attenuation in the relationship between of *MYH1* and corresponding Type IIX fibers was observed following initiation of training as an expected resulting from changes in *MYH1* gene regulation in response to mechanical loading [44]. We further utilized the fact that collected baseline samples from both legs prior to any training and between leg variations can be expected to be negligible [23]. Here we assumed that maximal biological resolution would be achieved when technical variation between paired samples were minimized. Following suggestions from Teng et al. [45] with modifications to suit the within-participant design of the present study, we calculated average log₂-differences between replicates. Overall, these analyses showed that transcriptome- outperformed genome-mapping tools and that RSEM showed slightly lower average variation compared to Salmon and kallisto in the present data set.

A basic assumption in many transcriptome studies is that transcripts are counted and compared between conditions at a per-cell level [27]. This is likely also often implicitly assumed to be equivalent to measuring transcriptome data as ratios between mRNA and total RNA as the input in sequencing or hybridization experiments usually is total RNA [13, 15, 46]. Using the present data set, we have previously reported that total RNA increases per-unit-tissue in a volume-dependent manner following initiation of resistance training [3]. This in turn resulted in different amounts of

tissue used for preparing cDNA libraries (Figure 3A), as a fixed amount of total RNA was used for this purpose. When unaccounted for, this arguably lead to comparison of transcript counts between different numbers of muscle tissue cells when comparing the two volume conditions. Interestingly, when normalizing the effective library size to the amount of muscle tissue used to prepare libraries (Figure 3B and C), the apparent difference between conditions in average library size was diminished. This suggests that the amount of mRNA increases similarly in response to resistance training per-unit-tissue weight regardless of training volume. This is in contrast to increases in total RNA, presumably relating to ribosomal RNA (rRNA) induction indicated by targeted analysis of rRNA species [3].

In the light of the above mentioned volume-dependent differences and with the aim to compare transcriptome perturbations between volume conditions, we decided to use three different normalization scenarios when examining condition differences. A first scenario was formulated to account for the amount of tissue used in the experiment together with the resulting library size. A second scenario accounted for effective library sizes and third, naïve scenario, represented an analysis without normalization. As any normalization strategy was included to affect all genes similarly in the analysis, this resulted in global shifts in differences between volume conditions between normalizations scenarios at Week 2. This was evident when examining the full distribution of fold-changes between conditions and specific subsets of genes related to specific GO terms (Figure 3G). The tissue offset model shifted fold-changes in the direction of multiple-set compared to other scenarios as when transcript counts were expressed as a rate per tissue-weight the lower amount of tissue used to prepare sequencing reactions pushed fold-changes in the direction of multiple-set. The naïve scenario showed a reversed shift compared to the tissue offset model as not accounting for the amount of tissue led to less counts in the multiple-set condition. Results from the effective library size model showed smaller differences between volume conditions, here differences in counts between conditions are differences at the average library size. These scenarios could conveniently be compared as corresponding models were fitted in the same statistical framework, i.e. using generalized linear mixed models (GLMM) as previously suggested [31]. An additional benefit of using GLMM was the incorporation of random effects to account for the present repeated measures design. Although approaches exists to account for correlated

observations in commonly used RNA-seq modelling frameworks [47], GLMMs provides a more robust and potentially more powerful framework for dealing with correlated data [31].

On a global scale, comparing these scenarios showed that accounting for the amount of tissue used in library preparations affected the interpretation of the data to some degree. More specifically this led to the identification volume dependence in extracellular matrix related genes in the early stage of resistance training. Extracellular matrix (ECM) remodeling has been shown to be induced by exercise training evident from acute phase, collagen synthesis studies [48, 49], studies of long term endurance and resistance training examining mRNA and protein abundances [13, 18, 50, 51] and single bout damaging muscle contractions [52] affecting ECM related proteins and mRNAs as well single bout non-damaging exercise affecting ECM related genes [53]. Although thought to have an important function in protecting skeletal muscle from injury [52], studies comparing eccentric exercise, known induce more muscle damage, to concentric exercise have shown diverging results regarding acute collagen synthesis [49, 54]. Additionally, light vs. heavy loadings were not shown to affect collagen synthesis differently when total load lifted were equated between conditions [48]. Overall, these limited data from human exercise studies gives no clear indication which exercise modalities affects ECM remodeling. Studying ECM remodeling in response to exercise could prove important for the understanding of exercise induced adaptations in specific populations. Ageing and disuse affects the ECM leading to increased stiffness and potentially decreased force transmission and muscle efficiency [55, 56]. Although exercise-training generally affects muscle function and specifically ECM remodeling [13, 18, 50, 51], there are indications that aged muscle responds differently with regard to ECM remodeling [53, 57]. In order to study e.g. effects of ageing on ECM remodeling, robust exercise models should be utilized. Our data provides a valuable direction showing dose dependence of ECM related genes in response to resistance training. Recently, increases in ECM and collagen fibril organization proteins in response to training was shown to scale to their respective mRNAs [50]. This suggests transcriptional regulation of ECM, however, the time course of this relationship could however be more complex as transcriptional regulation of e.g. *COL1A2* shows a a considerable lag from stimuli to transcription as shown in fibroblasts [58]. Indeed, acute exercise counterintuitively leads to reduced expression of e.g. collagen mRNA evident from our study and others [18]. In contrasts chronic resistance training leads to increased expression [13, 18]. Together this indicates that the time-point

selected in the initial part of the present study was well suited to study specifically ECM related gene expression. After twelve weeks of training, differences between sets were diminished indicating that the initial phase training response provides relevant information with regards to dose effects.

In contrast to initial resting state biopsies, the acute phase comparison between volume conditions did not reveal any apparent volume-dependent effects. Only a single gene, *RFT1* was shown to be differentially expressed as it was down-regulated 1-h after acute exercise (~ 0.8-fold, corroborating previous indications [22]) and more so in response to multiple- compared to single-set training (~ 0.7-fold). *RFT1* is associated with GO terms lipid transport, carbohydrate transport and endoplasmic reticulum membrane. The possible importance of this and previous estimates [22] of *RFT1* regulation in response to acute exercise warrants more research.

As single genes provides limited information, gene set enrichment analysis could provide more valuable insights. In the acute phase three gene GO sets were identified has more highly ranked among equally sized gene sets with regard to their minimum significant difference. Closer examination of these gene sets showed genes both up- and down-regulated in multiple- compared to single-set exercise and there by missed by gene set enrichment analysis based on fold-changes. Additionally only a small fraction of these transcripts actually showed positive MSD, indicating changes with unadjusted P-values < 0.05. Overall, these results did support robust volume-dependent regulation of these gene sets (RNA splicing, RNA localization and covalent chromatin modification). This also underlines the fact that the chosen sampling time-point was not insufficient to provide potentially valuable information on volume-dependent regulation of gene expression in the acute phase.

Conclusions

Transcriptomic analyses of skeletal muscle subjected to altered growth condition should account for global changes in mRNA to total RNA and cell density to accurately reflect biologically meaningful events. When accounting for such aspects in the present study, ECM remodeling in response to resistance training was identified as volume-dependent. Recommendations regarding normalization assumptions could be applicable to the study of other cell types and model systems undergoing increased or arrested growth.

Methods

Participants and study overview

The full study design has been previously described in detail [3]. Thirty-four participants completed a 12-week training-intervention with legs allocated to either low- (one set per exercise, single-set) or moderate-volume (three sets per exercise, multiple-set) training (Figure 1A). Muscle biopsies were obtained from each leg prior to and after the intervention, as well as prior to and 60-min after the fifth training session. Participants with a complete set of high-quality RNA samples ($RQI \geq 7$, $n=25$) were selected for RNA-seq (Figure 1B). Training-induced changes in muscle size and strength were estimated for each leg using several methods (for complete overview, see[3]). Herein, we present DXA-based measurement of lean mass for the 25 participants eligible for RNA-seq, as well as a weighted combined measure of strength (combining data from different strength tests).

Information about potential risks and discomforts associated with the study was given to participants prior to enrollment and all participants gave their written informed consent prior to inclusion. All procedures were approved by the local ethics committee at Inland Norway University of Applied Sciences (nr 2013-11-22:2) and the study design was pre-registered at ClinicalTrials.gov (Identifier: NCT02179307). The study was conducted in accordance with the *Declaration of Helsinki*.

Training protocol

The training protocol consisted of unilateral lower body exercises (leg-press, leg-curl and knee-extension). Each participant leg was randomly assigned to perform either one or three sets per exercise, ensuring within-subject comparisons. Rest periods between sets were 90-180 sec. The single-set leg was always trained in the rest period between the second and third set of the multiple-set protocol. Training protocols were performed in a progressive manner, whereby resistance was continuously adjusted to ensure that the targeted number of repetitions were reached at volatile fatigue. This was equivalent to 10 repetitions maximum (RM) in weeks one and two, followed by 8RM in weeks three to five and 6RM in weeks six to twelve. Each week consisted of either 2 or 3 training sessions. From week four, weeks with three sessions contained one session at a sub-maximal load (90% of previous session load). All sessions commenced with a standardized warm-up. After each session, participants were given a standardized milk-based drink [3].

Muscle strength and hypertrophy assessments

Muscle strength was assessed twice before and once after the intervention. A detailed description of strength outcomes resulting from the study has been reported previously [3]. For the purpose of the present analyses, we present a weighted average of strength gains for the 25 participants eligible for RNA-seq, based on data from unilateral isometric and isokinetic (60° , 120° and $240^\circ \times \text{sec}^{-1}$) knee extension, and one-repetition maximum (1RM) in unilateral knee extension and leg press. Isometric and isokinetic strength was assessed using an individually adjusted dynamometer (Cybex 6000, Cybex International, Medway USA). 1RM was defined as the maximum load lifted through the full range of motion. From pre-intervention tests, the highest values were used for change score calculations.

Muscle hypertrophy was assessed from full-body dual-energy X-ray absorptiometry (DXA; Lunar prodigy, GE Healthcare, Oslo, Norway) scans performed prior to and after the intervention. Leg lean-mass was derived from region of interests covering the full leg from collum femoris to the distal end of the foot defined in the analysis software (enCore, GE Healthcare, Oslo, Norway).

Muscle tissue sampling and RNA extraction

Muscle tissue was obtained bilaterally from m. vastus lateralis using a 12-gauge needle (Universal-plus, Medax, San Possidonio, Italy) under local anesthesia (Xylocaine, $10 \text{ mg} \times \text{ml}^{-1}$ with adrenaline $5 \mu\text{g} \times \text{ml}^{-1}$, AstraZeneca AS, Oslo, Norge). Samples were obtained from the two legs within 10 minutes of each other at all time-points. All rested state samples were obtained in the morning after a standardized breakfast. Participants were instructed to ingest standardized meals during the last 24 h leading up to the sampling event, and to refrain from strenuous physical activity the last 48 h. Samples were dissected in ice-cold sterile saline solution (0.9% NaCl), blotted dry, weighed and snap-frozen in isopentane, before storage at -80°C until further processing. For RNA extraction, frozen muscle samples were homogenized in 1 ml of TRIzol reagent (Invitrogen, Life technologies AS, Oslo, Norway) using a bead homogenizer (Bullet Blender, Next Advanced, Averill Park, NY, USA). After phase separation, $400 \mu\text{l}$ of the aqueous phase was used in isopropanol precipitation of RNA, and after three washing steps (70% ethanol) the pellet was eluted in TE buffer. All samples showed 260/280 *nm* ratio > 1.95 assessed using a spectrophotometer (NanoDrop 2000, ThermoFisher Scientific, Oslo, Norway). RNA integrity scores (RQI) were

determined using capillary electrophoresis (Experion Automated Electrophoresis Station using RNA StdSens Assay, Bio-Rad). Participants with complete sets of high quality RNA samples had an average RQI score of 9.0 (0.4) (full data set, 8.1 (2.1), range: 1-9.7) (Figure 1C).

Illumina library preparation and sequencing

For each participant, mRNA sequencing libraries were prepared from the same amount of RNA (1000 ng, depending on the minimum amount available) using TruSeq Stranded Total RNA Library Prep (Illumina, San Diego, CA USA). Paired-end sequencing (150 bp) was performed using an Illumina HiSeq 3000 (Illumina) at the Norwegian Sequencing Centre.

Bioinformatic analysis

Pre-alignment filtering

Trim Galore (version 0.6.5) [34] and Trimmomatic (version 0.39) [33] were used to discard low-quality reads and trim poor-quality bases before alignment, using default settings. The quality of filtered files was calculated by FastQC (version 0.11.4) [59] and summarized using MultiQC (version 1.8) [60].

Read alignment and quantification

Filtered reads were aligned to the Human genome (GRCh38 release-97 downloaded from <ftp.ensembl.org>) using the alignment-based methods HISAT2 (version 2.1.0) [36], STAR (version 2.7.2) [35], and RSEM (version 1.3.1) [37], used together with Bowtie 2 (version 2.3.4.3) [61], and the non-alignment methods Kallisto (version 0.44.0) [38] and Salmon (version 0.13.1) [39]. For HISAT2 and STAR, HTSeq was used for quantification as previously described [62]. RSEM, kallisto, and Salmon have in-built quantification functions.

Modeling of gene counts

Gene counts were modeled using negative binomial generalized linear mixed models (GLMM), as suggested in [31], with modifications. Three model formulations were used to represent three different normalization scenarios. First, to account for fluctuations in RNA-to-tissue ratios, the amount of tissue used in cDNA synthesis was included as an offset term together with the effective library size and study conditions (time and volume condition), added as a fixed effects in the model

(tissue offset model). A simplified model contained only the effective library size together with study conditions, included as fixed effects (Effective library-size model). And finally, a naïve model formulation, without any form of normalization term was used for comparisons. For acute exercise effects (fifth session pre- to post-exercise), only the library size normalized model was used as we expected that fluid shifts [63] could influence the muscle weight measurement and changes in Total-RNA were unlikely to occur in this short time span [64]. The effective library size was calculated by multiplying the total library size with the RNA composition normalization factor, calculated using the trimmed mean method [24], followed by dividing the value by the median effective library size, as suggested by Cui et al. [31]. The effect of resistance training on gene counts was assessed as i) the effect of exercise volume and ii) the effect of time. For analyses of the effect of exercise volume, differential expression was evaluated using models containing the interaction between time and exercise volume. For analyses of the effect of time, differential expression was evaluated using models containing only the time factor, combining all data irrespective of volume condition. In all models, a single random effect was used, giving each participant an individual intercept. Models were iteratively fitted using glmmTMB [65]. Utilization of the negative binomial distribution was supported by comparing the full model with a Poisson model (not containing the dispersion term), providing likelihood-ratio tests p -values that were distributed of p -primarily below $p=0.05$ (0.37% of models showed $p > 0.05$). Heteroscedasticity was assessed using the uniformity test in the DHARMa package [66], which generally showed good agreement with model assumptions, providing p -values concentrated near 1 (98.51% of models showed $p > 0.05$).

Genes were identified as differentially expressed when the absolute \log_2 fold-change was greater than 0.5 and the adjusted p -value was below 5%. P -values were adjusted *per-model* coefficient to control for the false discovery rate [67].

Functional annotation

Enrichment analyses of gene ontology (GO) gene sets were performed using three approaches. First, a non-parametric rank test (described in [68] and implemented in the tmod package [69], version 0.40) was performed based on gene specific minimum significant differences (MSD). MSD was defined as the lower limit of the 95% confidence interval (CI, based on estimated standard

errors) around the $\log(\text{FC})$ when $\log(\text{FC}) > 0$ and the negative inverse of the upper 95% CI when $\log(\text{FC}) < 0$. This metric has been shown to have lower false positive rates compared to other metrics applied to enrichment analysis [70]. As the MSD metric is positive when the CI does not overlap 0 and negative when overlap occurs, the rank test does not discern between up and downregulated gene sets. A second approach, gene set enrichment analysis (GSEA) [71], was used to quantify directional regulation of the gene set. GSEA was performed using the `fgsea` package [72] with \log_2 fold-change as the gene level metric. Thirdly, over-representation analysis (ORA) was performed to assess if genes identified as differentially expressed ($|\log_2 \text{fold-change}| > 0.5$ and adjusted p-values < 0.05) belonged to specific gene sets. ORA was performed using the `enrichGO` function in the `clusterProfiler` package [73], version 3.16.0. GO gene sets (biological process, cellular component and molecular function) were retrieved from the molecular signature database (version 7.1) [74].

Statistical analysis

Descriptive data are presented as mean and standard deviation (SD). Changes in muscle strength and CSA were estimated using linear mixed models on change scores with baseline values as covariates. Alignment tools were assessed by comparing \log_2 -differences between biological replicates, as suggested by Teng et al. [39], with modifications. Briefly, a subset of genes previously shown to be stably expressed between tissues was selected [45], whereupon \log fold-differences between paired biopsy samples was calculated (i.e. using biopsies collected from each of the two legs prior to the training intervention). In addition, alignment tools were assessed by comparing relationships (Pearson's correlation coefficient) between gene family profiling of myosin heavy chains (*MYH1*, *MYH2* and *MYH7*; muscle-specific) and the corresponding myosin heavy chain protein expression (measured using immunohistochemistry as fiber types IIX, IIA and I).. These mRNA and protein profiles were expressed as a fraction of total counts, thus removing the need for normalization of the RNA-seq data, as previously described for qPCR data [41]. Notably, these data also provided insight into the overall biological validity of the RNA-seq data, linking gene counts to protein phenotypes.

Immunohistochemistry

Quantification of myosin heavy chain abundance from formalin-fixed muscle biopsy cross-sections was performed as previously described and reported [3]. Briefly, 4 μm transverse sections were incubated with primary antibodies detecting all myosin isoforms but type IIX (BF-35, 5 $\mu\text{g} \times \text{ml}^{-1}$, Developmental Studies Hybridoma Bank, deposited by Schiaffino, S.) and type I myosin (MyHCSlow, 1:4000, catalogue M8421L, Sigma-Aldrich Norway AS, Oslo, Norway). Primary antibodies were visualized using BMU UltraView DAB and UltraView Red (Ventana Medical Systems, Inc. Tucson, USA). Muscle fibers were identified as either Type I (red), Type IIA (brown), Type IIX (unstained) or hybrid fibers Type IIA/IIX (light brown) (for representative images, see figure 3 in [3]). Hybrid fibers were analyzed as 0.5 \times Type IIA and 0.5 \times Type IIX.

List of abbreviations

If abbreviations are used in the text they should be defined in the text at first use, and a list of abbreviations can be provided.

Declarations

Acknowledgements

Not applicable

Funding

The work presented here was supported by a grant from Inland Hospital Trust (grant nr. 150282).

Availability of data and materials

RNA sequencing data will be submitted to the Gene Expression Omnibus (identifier: XXXXX). Additional data and code is available at <https://github.com/trainome/rnaseq-pipeline>.

Authors' contribution

DH, SE and BR planned and supervised the training intervention. YK, DH, SE, and RA were involved in the planning and writing of the manuscript. DH and SE performed the biopsies and collected the samples. DH performed Muscle tissue sampling and RNA extraction. YK performed bioinformatics analyses with inputs from RA. YK and DH performed the statistical analyses. DH performed immunohistochemistry. All authors provided useful inputs, interpreted the data, read and approved the manuscript.

Ethics approval and consent to participate

All procedures were approved by the local ethics committee at Lillehammer University College, Department of Sport Science (nr 2013-11-22:2) and participants gave their written informed consent prior to enrolment. The study design was pre-registered at ClinicalTrials.gov (Identifier: NCT02179307).

Consent for publication

Not applicable

Competing interests

The authors declare that they have no competing interests.

References

1. Egan B, Zierath JR. Exercise metabolism and the molecular regulation of skeletal muscle adaptation. *Cell metabolism*. 2013;17:162–84. doi:[10.1016/j.cmet.2012.12.012](https://doi.org/10.1016/j.cmet.2012.12.012).
2. Dirks ML, Wall BT, Valk B van de, Holloway TM, Holloway GP, Chabowski A, et al. One week of bed rest leads to substantial muscle atrophy and induces whole-body insulin resistance in the absence of skeletal muscle lipid accumulation. *Diabetes*. 2016;65:2862–75. doi:[10.2337/db15-1661](https://doi.org/10.2337/db15-1661).
3. Hammarström D, Øfsteng S, Koll L, Hanestadhaugen M, Hollan I, Apró W, et al. Benefits of higher resistance-training volume are related to ribosome biogenesis. *The Journal of physiology*. 2020;598:543–65. doi:[10.1113/JP278455](https://doi.org/10.1113/JP278455).
4. Hubal MJ, Gordish-Dressman H, Thompson PD, Price TB, Hoffman EP, Angelopoulos TJ, et al. Variability in muscle size and strength gain after unilateral resistance training. *Med Sci Sports Exerc*. 2005;37:964–72. <http://www.ncbi.nlm.nih.gov/pubmed/15947721>.

5. Ahtiainen JP, Walker S, Peltonen H, Holviala J, Sillanpaa E, Karavirta L, et al. Heterogeneity in resistance training-induced muscle strength and mass responses in men and women of different ages. *Age (Dordr)*. 2016;38:10. doi:[10.1007/s11357-015-9870-1](https://doi.org/10.1007/s11357-015-9870-1).
6. Bamman MM, Petrella JK, Kim JS, Mayhew DL, Cross JM. Cluster analysis tests the importance of myogenic gene expression during myofiber hypertrophy in humans. *J Appl Physiol* (1985). 2007;102:2232–9. doi:[10.1152/jappphysiol.00024.2007](https://doi.org/10.1152/jappphysiol.00024.2007).
7. Stec MJ, Kelly NA, Many GM, Windham ST, Tuggle SC, Bamman MM. Ribosome biogenesis may augment resistance training-induced myofiber hypertrophy and is required for myotube growth in vitro. *Am J Physiol Endocrinol Metab*. 2016;310:E652–61. doi:[10.1152/ajpendo.00486.2015](https://doi.org/10.1152/ajpendo.00486.2015).
8. Davidsen PK, Gallagher IJ, Hartman JW, Tarnopolsky MA, Dela F, Helge JW, et al. High responders to resistance exercise training demonstrate differential regulation of skeletal muscle microRNA expression. *J Appl Physiol* (1985). 2011;110:309–17. doi:[10.1152/jappphysiol.00901.2010](https://doi.org/10.1152/jappphysiol.00901.2010).
9. Morton RW, Murphy KT, McKellar SR, Schoenfeld BJ, Henselmans M, Helms E, et al. A systematic review, meta-analysis and meta-regression of the effect of protein supplementation on resistance training-induced gains in muscle mass and strength in healthy adults. *Br J Sports Med*. 2018;52:376–84. doi:[10.1136/bjsports-2017-097608](https://doi.org/10.1136/bjsports-2017-097608).
10. Brook MS, Wilkinson DJ, Phillips BE, Perez-Schindler J, Philp A, Smith K, et al. Skeletal muscle homeostasis and plasticity in youth and ageing: Impact of nutrition and exercise. *Acta Physiol (Oxf)*. 2016;216:15–41. doi:[10.1111/apha.12532](https://doi.org/10.1111/apha.12532).
11. Timmons JA. Variability in training-induced skeletal muscle adaptation. *J Appl Physiol* (1985). 2011;110:846–53. doi:[10.1152/jappphysiol.00934.2010](https://doi.org/10.1152/jappphysiol.00934.2010).
12. Gordon PM, Liu D, Sartor MA, IglayReger HB, Pistilli EE, Gutmann L, et al. Resistance exercise training influences skeletal muscle immune activation: A microarray analysis. *J Appl Physiol* (1985). 2012;112:443–53.
13. Damas F, Ugrinowitsch C, Libardi CA, Jannig PR, Hector AJ, McGlory C, et al. Resistance training in young men induces muscle transcriptome-wide changes associated with muscle structure and metabolism refining the response to exercise-induced stress. *Eur J Appl Physiol*. 2018;118:2607–16. doi:[10.1007/s00421-018-3984-y](https://doi.org/10.1007/s00421-018-3984-y).
14. Hyldahl RD, Xin L, Hubal MJ, Moeckel-Cole S, Chipkin S, Clarkson PM. Activation of nuclear factor- κ PB following muscle eccentric contractions in humans is localized primarily to skeletal muscle-residing pericytes. *The FASEB Journal*. 2011;25:2956–66. doi:[10.1096/fj.10-177105](https://doi.org/10.1096/fj.10-177105).
15. Robinson MM, Dasari S, Konopka AR, Johnson ML, Manjunatha S, Esponda RR, et al. Enhanced protein translation underlies improved metabolic and physical adaptations to different exercise training modes in young and old humans. *Cell Metabolism*. 2017;25:581–92. doi:[10.1016/j.cmet.2017.02.009](https://doi.org/10.1016/j.cmet.2017.02.009).

16. Melov S, Tarnopolsky MA, Beckman K, Felkey K, Hubbard A. Resistance exercise reverses aging in human skeletal muscle. *PLOS ONE*. 2007;2:e465. doi:[10.1371/journal.pone.0000465](https://doi.org/10.1371/journal.pone.0000465).
17. Murton AJ, Billeter R, Stephens FB, Des Etages SG, Graber F, Hill RJ, et al. Transient transcriptional events in human skeletal muscle at the outset of concentric resistance exercise training. *J Appl Physiol (1985)*. 2014;116:113–25. doi:[10.1152/jappphysiol.00426.2013](https://doi.org/10.1152/jappphysiol.00426.2013).
18. Raue U, Trappe TA, Estrem ST, Qian HR, Helvering LM, Smith RC, et al. Transcriptome signature of resistance exercise adaptations: Mixed muscle and fiber type specific profiles in young and old adults. *J Appl Physiol (1985)*. 2012;112:1625–36. doi:[10.1152/jappphysiol.00435.2011](https://doi.org/10.1152/jappphysiol.00435.2011).
19. Phillips BE, Williams JP, Gustafsson T, Bouchard C, Rankinen T, Knudsen S, et al. Molecular networks of human muscle adaptation to exercise and age. *PLoS Genet*. 2013;9:e1003389. doi:[10.1371/journal.pgen.1003389](https://doi.org/10.1371/journal.pgen.1003389).
20. Hangelbroek RWJ, Fazelzadeh P, Tieland M, Boekschoten MV, Hooiveld GJEJ, Duynhoven JPM van, et al. Expression of protocadherin gamma in skeletal muscle tissue is associated with age and muscle weakness. *Journal of Cachexia, Sarcopenia and Muscle*. 2016;7:604–14. doi:[10.1002/jcsm.12099](https://doi.org/10.1002/jcsm.12099).
21. STEPTO NK, COFFEY VG, CAREY AL, PONNAMPALAM AP, CANNY BJ, POWELL D, et al. Global gene expression in skeletal muscle from well-trained strength and endurance athletes. *Medicine & Science in Sports & Exercise*. 2009;41:546–65. doi:[10.1249/MSS.0b013e31818c6be9](https://doi.org/10.1249/MSS.0b013e31818c6be9).
22. Pillon NJ, Gabriel BM, Dollet L, Smith JAB, Sardón Puig L, Botella J, et al. Transcriptomic profiling of skeletal muscle adaptations to exercise and inactivity. *Nature communications*. 2020;11:470–0. doi:[10.1038/s41467-019-13869-w](https://doi.org/10.1038/s41467-019-13869-w).
23. Tarnopolsky M, Phillips S, Parise G, Varbanov A, DeMuth J, Stevens P, et al. Gene expression, fiber type, and strength are similar between left and right legs in older adults. *The Journals of Gerontology: Series A*. 2007;62:1088–95. doi:[10.1093/gerona/62.10.1088](https://doi.org/10.1093/gerona/62.10.1088).
24. Robinson MD, Oshlack A. A scaling normalization method for differential expression analysis of rna-seq data. *Genome Biology*. 2010;11:R25. doi:[10.1186/gb-2010-11-3-r25](https://doi.org/10.1186/gb-2010-11-3-r25).
25. Dillies M-A, Rau A, Aubert J, Hennequet-Antier C, Jeanmougin M, Servant N, et al. A comprehensive evaluation of normalization methods for illumina high-throughput rna sequencing data analysis. *Briefings in bioinformatics*. 2013;14:671–83. doi:[10.1093/bib/bbs046](https://doi.org/10.1093/bib/bbs046).
26. Lin CY, Lovén J, Rahl PB, Paranal RM, Burge CB, Bradner JE, et al. Transcriptional amplification in tumor cells with elevated c-myc. *Cell*. 2012;151:56–67. doi:[10.1016/j.cell.2012.08.026](https://doi.org/10.1016/j.cell.2012.08.026).
27. Lovén J, Orlando DA, Sigova AA, Lin CY, Rahl PB, Burge CB, et al. Revisiting global gene expression analysis. *Cell*. 2012;151:476–82. doi:[10.1016/j.cell.2012.10.012](https://doi.org/10.1016/j.cell.2012.10.012).
28. Conesa A, Madrigal P, Tarazona S, Gomez-Cabrero D, Cervera A, McPherson A, et al. A survey of best practices for rna-seq data analysis. *Genome biology*. 2016;17:13–3. doi:[10.1186/s13059-016-0881-8](https://doi.org/10.1186/s13059-016-0881-8).

29. Del Fabbro C, Scalabrin S, Morgante M, Giorgi FM. An extensive evaluation of read trimming effects on illumina ngs data analysis. *PLOS ONE*. 2013;8:e85024. doi:[10.1371/journal.pone.0085024](https://doi.org/10.1371/journal.pone.0085024).
30. Eisenberg E, Levanon EY. Human housekeeping genes, revisited. *Trends Genet*. 2013;29:569–74. doi:[10.1016/j.tig.2013.05.010](https://doi.org/10.1016/j.tig.2013.05.010).
31. Cui S, Ji T, Li J, Cheng J, Qiu J. What if we ignore the random effects when analyzing rna-seq data in a multifactor experiment. *Stat Appl Genet Mol Biol*. 2016;15:87–105. doi:[10.1515/sagmb-2015-0011](https://doi.org/10.1515/sagmb-2015-0011).
32. Schoenfeld BJ, Ogborn D, Krieger JW. Dose-response relationship between weekly resistance training volume and increases in muscle mass: A systematic review and meta-analysis. *J Sports Sci*. 2016;1–10. doi:[10.1080/02640414.2016.1210197](https://doi.org/10.1080/02640414.2016.1210197).
33. Bolger AM, Lohse M, Usadel B. Trimmomatic: A flexible trimmer for illumina sequence data. *Bioinformatics* (Oxford, England). 2014;30:2114–20. doi:[10.1093/bioinformatics/btu170](https://doi.org/10.1093/bioinformatics/btu170).
34. 2019. <https://github.com/FelixKrueger/TrimGalore>.
35. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: Ultrafast universal rna-seq aligner. *Bioinformatics* (Oxford, England). 2013;29:15–21. doi:[10.1093/bioinformatics/bts635](https://doi.org/10.1093/bioinformatics/bts635).
36. Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. Graph-based genome alignment and genotyping with hisat2 and hisat-genotype. *Nature Biotechnology*. 2019;37:907–15. doi:[10.1038/s41587-019-0201-4](https://doi.org/10.1038/s41587-019-0201-4).
37. Li B, Dewey CN. RSEM: Accurate transcript quantification from rna-seq data with or without a reference genome. *BMC Bioinformatics*. 2011;12:323. doi:[10.1186/1471-2105-12-323](https://doi.org/10.1186/1471-2105-12-323).
38. Bray NL, Pimentel H, Melsted P, Pachter L. Near-optimal probabilistic rna-seq quantification. *Nature Biotechnology*. 2016;34:525–7. doi:[10.1038/nbt.3519](https://doi.org/10.1038/nbt.3519).
39. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression. *Nature Methods*. 2017;14:417–9. doi:[10.1038/nmeth.4197](https://doi.org/10.1038/nmeth.4197).
40. Serrano AL, Perez M, Lucia A, Chicharro JL, Quiroz-Rothe E, Rivero JL. Immunolabelling, histochemistry and in situ hybridisation in human skeletal muscle fibres to detect myosin heavy chain expression at the protein and mRNA level. *J Anat*. 2001;199 Pt 3:329–37. doi:[10.1046/j.1469-7580.2001.19930329.x](https://doi.org/10.1046/j.1469-7580.2001.19930329.x).
41. Ellefsen S, Vikmoen O, Zacharoff E, Rauk I, Slettalokken G, Hammarstrom D, et al. Reliable determination of training-induced alterations in muscle fiber composition in human skeletal muscle using quantitative polymerase chain reaction. *Scand J Med Sci Sports*. 2014;24:e332–42. doi:[10.1111/sms.12185](https://doi.org/10.1111/sms.12185).

42. Marx JO, Kraemer WJ, Nindl BC, Larsson L. Effects of aging on human skeletal muscle myosin heavy-chain mRNA content and protein isoform expression. *J Gerontol A Biol Sci Med Sci*. 2002;57:B232–8. doi:[10.1093/gerona/57.6.b232](https://doi.org/10.1093/gerona/57.6.b232).
43. Ellefsen S, Stenslokken KO. Gene-family profiling: A normalization-free real-time rt-pcr approach with increased physiological resolution. *Physiol Genomics*. 2010;42:1–4. doi:[10.1152/physiolgenomics.00196.2009](https://doi.org/10.1152/physiolgenomics.00196.2009).
44. Andersen JL, Gruschy-Knudsen T. Rapid switch-off of the human myosin heavy chain iix gene after heavy load muscle contractions is sustained for at least four days. *Scand J Med Sci Sports*. 2018;28:371–80.
45. Teng M, Love MI, Davis CA, Djebali S, Dobin A, Graveley BR, et al. A benchmark for rna-seq quantification pipelines. *Genome Biology*. 2016;17:74. doi:[10.1186/s13059-016-0940-1](https://doi.org/10.1186/s13059-016-0940-1).
46. Laker RC, Garde C, Camera DM, Smiles WJ, Zierath JR, Hawley JA, et al. Transcriptomic and epigenetic responses to short-term nutrient-exercise stress in humans. *Scientific Reports*. 2017;7:15134. doi:[10.1038/s41598-017-15420-7](https://doi.org/10.1038/s41598-017-15420-7).
47. McCarthy DJ, Chen Y, Smyth GK. Differential expression analysis of multifactor rna-seq experiments with respect to biological variation. *Nucleic Acids Research*. 2012;40:4288–97. doi:[10.1093/nar/gks042](https://doi.org/10.1093/nar/gks042).
48. Holm L, Hall G van, Rose AJ, Miller BF, Doessing S, Richter EA, et al. Contraction intensity and feeding affect collagen and myofibrillar protein synthesis rates differently in human skeletal muscle. *Am J Physiol Endocrinol Metab*. 2010;298:E257–69.
49. Moore DR, Phillips SM, Babraj JA, Smith K, Rennie MJ. Myofibrillar and collagen protein synthesis in human skeletal muscle in young men after maximal shortening and lengthening contractions. *Am J Physiol Endocrinol Metab*. 2005;288:E1153–9.
50. Makhnovskii PA, Zgoda VG, Bokov RO, Shagimardanova EI, Gazizova GR, Gusev OA, et al. Regulation of proteins in human skeletal muscle: The role of transcription. *Scientific Reports*. 2020;10:3514. doi:[10.1038/s41598-020-60578-2](https://doi.org/10.1038/s41598-020-60578-2).
51. Hjorth M, Norheim F, Meen AJ, Pourteymour S, Lee S, Holen T, et al. The effect of acute and long-term physical activity on extracellular matrix and serglycin in human skeletal muscle. *Physiological reports*. 2015;3:e12473. doi:[10.14814/phy2.12473](https://doi.org/10.14814/phy2.12473).
52. Mackey AL, Brandstetter S, Schjerling P, Bojsen-Moller J, Qvortrup K, Pedersen MM, et al. Sequenced response of extracellular matrix adhesion and fibrotic regulators after muscle damage is involved in protection against future injury in human skeletal muscle. *Faseb j*. 2011;25:1943–59. doi:[10.1096/fj.10-176487](https://doi.org/10.1096/fj.10-176487).
53. Wessner B, Liebensteiner M, Nachbauer W, Csapo R. Age-specific response of skeletal muscle extracellular matrix to acute resistance exercise: A pilot study. *Eur J Sport Sci*. 2019;19:354–64. doi:[10.1080/17461391.2018.1526974](https://doi.org/10.1080/17461391.2018.1526974).

54. Holm L, Rahbek SK, Farup J, Vendelbo MH, Vissing K. Contraction mode and whey protein intake affect the synthesis rate of intramuscular connective tissue. *Muscle Nerve*. 2017;55:128–30.
55. Csapo R, Gumpfenberger M, Wessner B. Skeletal muscle extracellular matrix - what do we know about its composition, regulation, and physiological roles? A narrative review. *Frontiers in physiology*. 2020;11:253–3. doi:[10.3389/fphys.2020.00253](https://doi.org/10.3389/fphys.2020.00253).
56. Azizi E, Deslauriers AR, Holt NC, Eaton CE. Resistance to radial expansion limits muscle strain and work. *Biomech Model Mechanobiol*. 2017;16:1633–43. doi:[10.1007/s10237-017-0909-3](https://doi.org/10.1007/s10237-017-0909-3).
57. Sorensen JR, Skousen C, Holland A, Williams K, Hyldahl RD. Acute extracellular matrix, inflammatory and mapk response to lengthening contractions in elderly human skeletal muscle. *Experimental Gerontology*. 2018;106:28–38. doi:<https://doi.org/10.1016/j.exger.2018.02.013>.
58. Schwarz RI. Collagen i and the fibroblast: High protein expression requires a new paradigm of post-transcriptional, feedback regulation. *Biochem Biophys Rep*. 2015;3:38–44. doi:[10.1016/j.bbrep.2015.07.007](https://doi.org/10.1016/j.bbrep.2015.07.007).
59. 2019. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.
60. Ewels P, Magnusson M, Lundin S, Källner M. MultiQC: Summarize analysis results for multiple tools and samples in a single report. *Bioinformatics*. 2016;32:3047–8. doi:[10.1093/bioinformatics/btw354](https://doi.org/10.1093/bioinformatics/btw354).
61. Langmead B, Salzberg SL. Fast gapped-read alignment with bowtie 2. *Nature Methods*. 2012;9:357–9. doi:[10.1038/nmeth.1923](https://doi.org/10.1038/nmeth.1923).
62. Anders S, Pyl PT, Huber W. HTSeq—a python framework to work with high-throughput sequencing data. *Bioinformatics*. 2015;31:166–9. doi:[10.1093/bioinformatics/btu638](https://doi.org/10.1093/bioinformatics/btu638).
63. Ploutz-Snyder LL, Convertino VA, Dudley GA. Resistance exercise-induced fluid shifts: Change in active muscle size and plasma volume. *Am J Physiol*. 1995;269 3 Pt 2:R536–43.
64. Figueiredo VC, McCarthy JJ. Regulation of ribosome biogenesis in skeletal muscle hypertrophy. *Physiology (Bethesda)*. 2019;34:30–42.
65. Brooks ME, Kristensen K, Benthem KJ van, Magnusson A, Berg CW, Nielsen A, et al. glmmTMB Balances Speed and Flexibility Among Packages for Zero-inflated Generalized Linear Mixed Modeling. *The R Journal*. 2017;9:378–400. doi:[10.32614/RJ-2017-066](https://doi.org/10.32614/RJ-2017-066).
66. Hartig F. DHARMA: Residual diagnostics for hierarchical (multi-level / mixed) regression models. 2020. <http://florianhartig.github.io/DHARMA/>.
67. Benjamini Y, Hochberg Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society Series B (Methodological)*. 1995;57:289–300. www.jstor.org/stable/2346101.

68. Yamaguchi KD, Ruderman DL, Croze E, Wagner TC, Velichko S, Reder AT, et al. IFN-beta-regulated genes show abnormal expression in therapy-naïve relapsing-remitting ms mononuclear cells: Gene expression analysis employing all reported protein-protein interactions. *J Neuroimmunol.* 2008;195:116–20.
69. Zyla J, Marczyk M, Domaszewska T, Kaufmann SHE, Polanska J, Weiner J. Gene set enrichment for reproducible science: Comparison of cerno and eight other algorithms. *Bioinformatics.* 2019;35:5146–54. doi:[10.1093/bioinformatics/btz447](https://doi.org/10.1093/bioinformatics/btz447).
70. 2017;18:256. doi:[10.1186/s12859-017-1674-0](https://doi.org/10.1186/s12859-017-1674-0).
71. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A.* 2005;102:15545–50. doi:[10.1073/pnas.0506580102](https://doi.org/10.1073/pnas.0506580102).
72. Korotkevich G, Sukhov V, Sergushichev A. Fast gene set enrichment analysis. *bioRxiv.* 2019;060012. doi:[10.1101/060012](https://doi.org/10.1101/060012).
73. Yu G, Wang L-G, Han Y, He Q-Y. ClusterProfiler: An r package for comparing biological themes among gene clusters. *Omics : a journal of integrative biology.* 2012;16:284–7.
74. Liberzon A, Subramanian A, Pinchback R, Thorvaldsdóttir H, Tamayo P, Mesirov JP. Molecular signatures database (msigdb) 3.0. *Bioinformatics.* 2011;27:1739–40. doi:[10.1093/bioinformatics/btr260](https://doi.org/10.1093/bioinformatics/btr260).

Figure legends

Figure 1: Study overview and RNA-seq analysis pipeline. Forty-one participants were recruited and had their legs randomized to either single- (one set per exercise) or multiple-set (three set per exercise) training for the duration of twelve weeks (2-3 sessions week⁻¹) (A). Pre- and post-training testing included strength and muscle lean-mass assessments. *M. vastus lateralis* muscle biopsies were collected at four time-points, prior to and after the intervention (Week 0 and 12) and before and after the fifth training session (Week 2). Biopsies from participants who completed > 85% of prescribed sessions were used for RNA extraction (n=34; A). RNA quality was assessed (B) and a subset of participants with RNA quality indicator (RQI) scores > 7 were included in the RNA-seq experiment. RNA quality was not associated with muscle tissue weight (C) and participants included in RNA-seq experiments (n=25) did not differ from excluded in terms of muscle lean-mass gains (D). Multiple-set training led to greater gains in lean-mass (E) and lower extremity strength (F) compared to single-set training in the subset of participants included in the RNA-seq experiment. RNA-seq data was quality filtered using trimgalore and trimmomatic and reads were

compared to unfiltered reads (G). Read alignment was performed using five tools of which RSEM, kallisto and Salmon showed greater fractions of genes with robust expression after removing low-abundance genes (expression filtering; H) compared to HISAT2 and STAR. RSEM, kallisto and Salmon also showed less log₂-differences between biological replicates in a subset of genes with known robust expression (see text for details, I).

Figure 2: Correlations between myosin heavy chain mRNA and protein abundance. mRNA abundances estimated with RSEM, kallisto and Salmon showed stronger correlations with the corresponding protein expression (B). Relative abundances of mRNA and protein were calculated as a percentage of the whole mRNA and protein family respectively (*MYH1*, *MYH2* and *MYH7* for mRNA and Type IIX, IIA and I for protein).

Figure 3: Effects of training on muscle tissue used in cDNA synthesis and comparison between exercise volume conditions in rested state biopsies between normalization methods.

Biopsy tissue mass used in cDNA synthesis varied over the course of the study and between volume conditions (A) as a result of varied RNA to tissue weight ratios [3]. Despite lower amounts of tissue, effective library sizes increased after the onset of training with an tendency towards greater increase in the Single-set condition (B). When expressing library sizes per-unit tissue weight differences between volume conditions were diminished but increases from baseline were maintained (C). Between volume-condition comparisons in three different modeling scenarios resulted in different sets of differentially expressed (DE) genes. The naive model shared DE genes with higher expression in the single-set condition with the effective-library size normalized model but no genes with the Tissue-offset model at week 2 (D, E). The tissue offset-normalized model shared ten genes with higher expression in multiple-sets with the effective library-size normalized model at week 2 (D, E). The naive model shared genes showing higher expression in the single-set condition at week 12 (H, I). No genes were identified as more highly expressed in multiple-set at week 12. Enrichment analysis revealed gene sets related to extracellular matrix as more highly expressed in multiple-set at Week 2 in the tissue offset model (F). All gene sets identified in the tissue offset model were more highly expressed in multiple-set, indicated by a positive enrichment score in F. Purple dots represents gene categories also identified from over-representation analysis (ORA) from DE genes. Normalization strategies had global effects on rank tests as fold-changes and minimum significant differences scores (not shown) shifted as exemplified by the “Collagen

containing extracellular matrix” gene set in G and the full distribution of log₂ fold-changes shown as density curves. Gray bars represents genes not contained in the gene set, black bars represents genes contained in the gene set. Genes symbols indicate genes identified as differentially expressed in each normalization scenario (log₂ fold-change > 0.5 and adjusted P-values < 0.05).

Figure 4: Effects of acute exercise on gene expression. Many genes changed in both directions in response to acute exercise when both volume conditions were combined (A). Functional annotation revealed increased expression of gene ontology categories associated with stress response and transcription and decreased response of categories related to extracellular matrix (B). Categories also identified in over-representation analysis (ORA) are highlighted in B. Comparing expression perturbations between volume conditions identified a single differentially expressed gene (RFT1, C). Three gene ontology categories were identified as enriched based on minimum significant difference (MSD) ranking, genes from these categories with MSD > 0 are identified in (C) and traces from rank tests are displayed in D.

Figure 5: Comparison of differential expression over time between normalization scenarios. Volcano plot identifies differentially expressed genes (adjusted P-values < 0.05 and log₂ fold-changes > 0.5, filled circles). Bar-plots shows total number of differentially expressed genes (horizontal bars) and sets exclusively found in each model or shared among models (vertical bars). Comparing Week 2 to pre-training between model showed that the largest fraction of genes were shared among normalization scenarios although downregulated genes were to a large extent identified in the Effective library size normalized model (A). The majority of up-regulated genes from pre-training to Week 12 were found in the tissue-offset and naïve models (B).

			Mean	SD
Female	n = 11	Age (years)	22.6	0.9
		Body mass (kg)	166.2	6.2
		Stature (cm)	61.5	7.4

Male	n = 14	Age (years)	23.9	4.2
		Body mass (kg)	183.7	5.6
		Stature (cm)	77.4	10.4

Table 2. Functional enrichment analysis

Comparison	Normalization model	Gene ontology category	ID	Description	GSEA			
					Rank P-value	GSEA P-value	NESORA P-value	
Week 2	Tissue offset	Biological process	GO:0043062	Extracellular structure organization	2.04e-40	2.98e-24	1.939.04e-28	
			GO:0030199	Collagen fibril organization	3.33e-18	9.99e-13	2.311.94e-08	
			GO:0060326	Cell chemotaxis	2.36e-15	1.34e-15	1.912.20e-11	
			GO:0062023	Collagen containing extracellular matrix	6.84e-68	4.34e-46	2.205.88e-53	
			GO:0005788	Endoplasmic reticulum lumen	2.62e-23	9.11e-13	1.755.21e-17	
		Cellular component	GO:0005581	Collagen trimer	4.92e-21	4.38e-14	2.251.29e-10	
			GO:0031983	Vesicle lumen	1.82e-15	1.09e-09	1.631.37e-12	
			GO:0005201	Extracellular matrix structural constituent	6.52e-40	1.75e-24	2.218.04e-30	
			Molecular function	GO:0005539	Glycosaminoglycan binding	7.40e-16	5.06e-12	1.881.87e-09
				GO:0005178	Integrin binding	2.83e-14	1.60e-08	1.821.17e-12
Effective library size	Biological process	GO:0043062	Extracellular structure organization	3.58e-34	2.22e-23	2.024.27e-29		

		GO:0030199	Collagen fibril organization	5.66e-17	1.75e-10	2.272.49e-14
		GO:0061448	Connective tissue development	4.04e-14	1.60e-07	1.731.88e-10
		GO:0060326	Cell chemotaxis	4.99e-14	2.03e-12	1.957.75e-15
		GO:0051216	Cartilage development	3.04e-12	5.19e-07	1.802.59e-10
		GO:0062023	Collagen containing extracellular matrix	1.37e-60	2.50e-37	2.194.43e-52
	Cellular component	GO:0005581	Collagen trimer	1.51e-19	5.35e-12	2.231.18e-14
		GO:0005788	Endoplasmic reticulum lumen	3.36e-17	2.34e-12	1.871.50e-10
	Molecular function	GO:0005201	Extracellular matrix structural constituent	3.46e-37	7.63e-19	2.202.82e-28
		GO:0005539	Glycosaminoglycan binding	1.60e-13	1.89e-10	1.972.86e-12
		GO:0043062	Extracellular structure organization	3.86e-40	9.85e-24	1.893.36e-26
	Biological process	GO:0030199	Collagen fibril organization	6.38e-18	3.18e-12	2.271.67e-08
Naïve		GO:0060326	Cell chemotaxis	2.66e-16	2.57e-16	1.895.00e-11
	Cellular component	GO:0062023	Collagen containing extracellular matrix	5.25e-69	5.33e-48	2.181.01e-53
		GO:0005788	Endoplasmic reticulum lumen	1.94e-23	2.76e-12	1.724.68e-18

		GO:0005581	Collagen trimer	3.27e-21	2.86e-14	2.192.17e-11
		GO:0031983	Vesicle lumen	4.65e-15	4.18e-10	1.606.31e-13
		GO:0005201	Extracellular matrix structural constituent	7.62e-40	1.79e-24	2.226.38e-29
	Molecular function	GO:0005539	Glycosaminoglycan binding	1.28e-16	4.15e-13	1.925.04e-10
		GO:0005178	Integrin binding	1.11e-14	6.29e-08	1.814.81e-12
	Biological process	GO:0043062	Extracellular structure organization	5.20e-49	5.18e-29	2.191.09e-37
		GO:0030199	Collagen fibril organization	1.03e-19	2.50e-12	2.541.29e-15
		GO:0062023	Collagen containing extracellular matrix	2.84e-68	7.31e-53	2.508.25e-67
	Cellular component	GO:0005581	Collagen trimer	1.57e-24	9.48e-21	2.632.40e-25
Week 12	Tissue offset	GO:0005788	Endoplasmic reticulum lumen	3.10e-19	2.35e-10	1.831.87e-13
		GO:0005604	Basement membrane	7.34e-16	9.58e-13	2.281.26e-14
		GO:0005201	Extracellular matrix structural constituent	9.47e-47	5.58e-34	2.624.24e-46
	Molecular function	GO:0005539	Glycosaminoglycan binding	5.69e-20	2.79e-14	2.132.01e-17
		GO:0008201	Heparin binding	2.20e-18	3.75e-14	2.239.82e-16

Effective library size	Biological process	GO:0030020	Extracellular matrix structural constituent conferring tensile strength	1.71e-16	2.18e-12	2.491.60e-13
		GO:0043062	Extracellular structure organization	5.01e-44	3.34e-22	1.878.48e-33
		GO:0030199	Collagen fibril organization	1.33e-18	2.37e-08	2.092.69e-10
	Cellular component	GO:0062023	Collagen containing extracellular matrix	8.19e-63	9.42e-39	2.072.59e-56
		GO:0005581	Collagen trimer	7.47e-24	3.20e-13	2.171.11e-26
		GO:0005788	Endoplasmic reticulum lumen	5.55e-17	5.58e-09	1.671.58e-13
		GO:0005201	Extracellular matrix structural constituent	1.36e-44	1.14e-21	2.144.31e-42
	Molecular function	GO:0005539	Glycosaminoglycan binding	7.82e-18	2.72e-12	1.891.05e-15
		GO:0008201	Heparin binding	1.27e-17	1.38e-10	1.946.20e-16
		GO:0030020	Extracellular matrix structural constituent conferring tensile strength	1.85e-16	8.08e-08	2.081.35e-16
GO:0005518		Collagen binding	1.77e-15	2.43e-06	1.921.12e-07	
Naïve	Biological process	GO:0043062	Extracellular structure organization	8.12e-52	1.27e-38	2.851.06e-28
		GO:0030199	Collagen fibril organization	1.26e-20	1.26e-14	3.115.82e-08

	GO:0062023	Collagen containing extracellular matrix	2.31e-78	3.28e-71	3.354.01e-46
Cellular component	GO:0005581	Collagen trimer	2.27e-26	4.96e-27	3.374.83e-12
	GO:0005788	Endoplasmic reticulum lumen	3.11e-23	7.77e-16	2.361.45e-14
	GO:0005201	Extracellular matrix structural constituent	4.70e-50	6.57e-48	3.546.42e-28
	GO:0005539	Glycosaminoglycan binding	2.92e-23	2.16e-19	2.756.84e-13
Molecular function	GO:0008201	Heparin binding	7.04e-21	3.05e-18	2.831.97e-09
	GO:0005178	Integrin binding	7.84e-17	3.77e-12	2.533.06e-09
	GO:0030020	Extracellular matrix structural constituent conferring tensile strength	1.98e-16	2.90e-15	3.163.77e-06

Appendix 2, All lncRNA identified in RNA-seq muscle biopsies data.

ENSEMBLE GENE ID	ENTREZGENE ID ID	HGNC - SYMBOL	ENSEMBLE GENE ID	ENTREZGENE ID ID	HGNC - SYMBOL	ENSEMBLE GENE ID	ENTREZGENE ID ID	HGNC - SYMBOL
ENSG00000093100	NA		ENSG00000176124	100874074	DLEU1	ENSG00000188004	284677	SNHG28
ENSG00000116652	NA	DLEU2L	ENSG00000176593	100128398		ENSG00000188185	NA	LINC00265
ENSG00000117242	100861548	PINK1-AS	ENSG00000176659	284756	C20orf197	ENSG00000188242	25845	
ENSG00000130600	283120	H19	ENSG00000176728	83869	TTY14	ENSG00000188825	NA	LINC00910
ENSG00000130600	102724852	H19	ENSG00000177337	NA	DLGAP1-AS1	ENSG00000189223	654433	PAX8-AS1
ENSG00000145075	NA	CCDC39	ENSG00000177406	100049716	NINJ2-AS1	ENSG00000189316	441239	
ENSG00000151303	NA		ENSG00000177410	441951	ZFAS1	ENSG00000196167	399948	COLCA1
ENSG00000153363	NA	LINC00467	ENSG00000177738	648987		ENSG00000196204	441191	RNF216P1
ENSG00000157306	NA	ZFH2-AS1	ENSG00000178977	284029	LINC00324	ENSG00000196295	NA	GARS1-DT
ENSG00000163364	NA	LINC01116	ENSG00000179406	285908	LINC00174	ENSG00000196696	283970	
ENSG00000163597	100507246	SNHG16	ENSG00000179523	645212	EIF3J-DT	ENSG00000196741	NA	LINC01560
ENSG00000164385	154386	LINC01600	ENSG00000179743	729614		ENSG00000196756	388796	SNHG17
ENSG00000165511	220979	ZNF22-AS1	ENSG00000179818	400960	PCBP1-AS1	ENSG00000196810	NA	CTBP1-DT
ENSG00000166770	NA	ZNF667-AS1	ENSG00000179935	NA	LINC00652	ENSG00000196951	100129858	SCOC-AS1
ENSG00000167920	147184	TMEM99	ENSG00000180139	NA	ACTA2-AS1	ENSG00000197180	158960	
ENSG00000170161	554249		ENSG00000180525	414235	PRR26	ENSG00000197182	400931	MIRLET7BHG
ENSG00000170919	NA	TPT1-AS1	ENSG00000180769	404201	WDFY3-AS2	ENSG00000197291	100190938	RAMP2-AS1
ENSG00000172965	541471	MIR4435-2HG	ENSG00000181798	151477	LINC00471	ENSG00000197536	NA	IRF1-AS1
ENSG00000174365	128439	SNHG11	ENSG00000182165	NA	TP53TG1	ENSG00000197815	NA	
ENSG00000174403	NA	MIR1-1HG-AS1	ENSG00000182257	NA	PRR34	ENSG00000197989	85028	SNHG12
ENSG00000174407	128826	MIR1-1HG	ENSG00000182648	100506380	LINC01006	ENSG00000198358	101928372	
ENSG00000175061	125144	SNHG29	ENSG00000183154	102723701		ENSG00000198468	642946	FLVCR1-DT
ENSG00000175611	100128782	LINC00476	ENSG00000184068	112637020	SREBF2-AS1	ENSG00000198496	10230	NBR2
ENSG00000175772	NA	LINC01106	ENSG00000184224	100505621	C11orf72	ENSG00000203280	100128531	KIAA1671-AS1
ENSG00000176124	10301	DLEU1	ENSG00000185847	100131138	LINC01405	ENSG00000203288	109729141	TDRKH-AS1
			ENSG00000186019	NA		ENSG00000203392	NA	
			ENSG00000186594	84981	MIR22HG	ENSG00000203506	NA	RBMS3-AS2
			ENSG00000186615	100129075	KTN1-AS1	ENSG00000203620	NA	
			ENSG00000187951	100288637		ENSG00000203644	NA	

ENSG00000203709	NA	MIR29B2CHG	ENSG00000212719	339263	LINC02693	ENSG00000215769	109286553	ARHGAP27P1-BPTFP1-KPNA2P3
ENSG00000203804	574406	ADAMTSL4-AS1	ENSG00000212978	339803		ENSG00000216895	100506302	
ENSG00000203808	154442	BVES-AS1	ENSG00000213121	NA		ENSG00000218018	100291105	RBM38-AS1
ENSG00000203875	NA	SNHG5	ENSG00000213599	100526830	SLX1A-SULT1A3	ENSG00000218510	29092	LINC00339
ENSG00000203930	NA	LINC00632	ENSG00000213742	102724826	ZNF337-AS1	ENSG00000219410	NA	
ENSG00000203993	85026	ARRDC1-AS1	ENSG00000213888	NA	LINC01521	ENSG00000219665	101928464	ZNF433-AS1
ENSG00000203999	284751	LINC01270	ENSG00000213904	100996307	LIPE-AS1	ENSG00000221817	101929145	PPP3CB-AS1
ENSG00000204054	NA	LINC00963	ENSG00000213904	101930071	LIPE-AS1	ENSG00000221990	NA	EXOC3-AS1
ENSG00000204261	100507463	PSMB8-AS1	ENSG00000213963	100130691		ENSG00000222041	112597	CYTOR
ENSG00000204282	100131096	TNRC6C-AS1	ENSG00000214106	NA	PAXIP1-AS2	ENSG00000223403	100507257	MEG9
ENSG00000204387	50854	SNHG32	ENSG00000214145	NA	LINC00887	ENSG00000223482	728190	NUTM2A-AS1
ENSG00000204460	151121	LINC01854	ENSG00000214293	100505854	APTR	ENSG00000223704	NA	LINC01422
ENSG00000204588	440894	LINC01123	ENSG00000214401	644246	KANSL1-AS1	ENSG00000223745	NA	CCDC18-AS1
ENSG00000204677	653316	FAM153CP	ENSG00000214401	107984142	KANSL1-AS1	ENSG00000223768	642852	LINC00205
ENSG00000204685	NA	STARD7-AS1	ENSG00000214548	55384	MEG3	ENSG00000223784	NA	LINP1
ENSG00000205056	NA	LINC02397	ENSG00000214708	105371730		ENSG00000223797	285266	ENTPD3-AS1
ENSG00000205106	374387	LINC02716	ENSG00000214719	NA		ENSG00000223799	NA	IL10RB-DT
ENSG00000205181	149837	LINC00654	ENSG00000214783	84820	POLR2J4	ENSG00000223891	100505783	OSER1-DT
ENSG00000205500	100129724	MAPRE3-AS1	ENSG00000214900	283551	LINC01588	ENSG00000223960	101927027	CHROMR
ENSG00000205740	107984285		ENSG00000214942	NA		ENSG00000224023	399821	EDRF1-DT
ENSG00000205791	503693	LOH12CR2	ENSG00000214970	NA		ENSG00000224032	NA	EPB41L4A-AS1
ENSG00000205885	283314	C1RL-AS1	ENSG00000215039	678655	CD27-AS1	ENSG00000224078	NA	SNHG14
ENSG00000205959	NA		ENSG00000215067	100506713	ALOX12-AS1	ENSG00000224086	NA	PPM1F-AS1
ENSG00000206195	503637	DUXAP8	ENSG00000215068	153684		ENSG00000224152	NA	
ENSG00000206337	10866	HCP5	ENSG00000215244	399715	LINC02649	ENSG00000224165	729723	DNAJC27-AS1
ENSG00000206344	253018	HCG27	ENSG00000215256	55449	DHRS4-AS1	ENSG00000224189	401022	HAGLR
ENSG00000206567	NA		ENSG00000215386	NA	MIR99AHG	ENSG00000224259	NA	LINC01133
ENSG00000206573	440944	THUMPD3-AS1	ENSG00000215417	407975	MIR17HG	ENSG00000224272	NA	
ENSG00000212694	338799	LINC01089	ENSG00000215424	114044	MCM3AP-AS1			

ENSG00000224281	100303728	SLC25A5-AS1	ENSG00000225746	NA	MEG8	ENSG00000227354	100505538	RBM26-AS1
ENSG00000224361	NA		ENSG00000225778	219731	PROSER2-AS1	ENSG00000227398	NA	KIF9-AS1
ENSG00000224424	100506637	PRKAR2A-AS1	ENSG00000225783	440823	MIAT	ENSG00000227456	114036	LINC00310
ENSG00000224568	105373537	LINC01886	ENSG00000225791	401264	TRAM2-AS1	ENSG00000227467	101928555	LINC01537
ENSG00000224609	729467		ENSG00000225855	284618	RUSC1-AS1	ENSG00000227496	NA	
ENSG00000224660	100505696	SH3BP5-AS1	ENSG00000225889	NA		ENSG00000227518	NA	
ENSG00000224699	NA	LAMTOR5-AS1	ENSG00000225914	414764	TSBP1-AS1	ENSG00000227543	NA	SPAG5-AS1
ENSG00000224713	NA		ENSG00000226029	107984921	LINC01772	ENSG00000227591	101930114	HSD11B1-AS1
ENSG00000224843	100133205	LINC00240	ENSG00000226051	253264	ZNF503-AS1	ENSG00000227617	100861402	CERS6-AS1
ENSG00000224870	148413	MRPL20-AS1	ENSG00000226137	440465	BAIAP2-DT	ENSG00000227627	NA	
ENSG00000224914	439994	LINC00863	ENSG00000226167	100287722	AP4B1-AS1	ENSG00000227811	NA	INKA2-AS1
ENSG00000224934	NA		ENSG00000226200	NA	SGMS1-AS1	ENSG00000227946	NA	
ENSG00000224958	572558	PGM5-AS1	ENSG00000226239	NA		ENSG00000227953	NA	LINC01341
ENSG00000224975	8552	INE1	ENSG00000226312	65072	CFLAR-AS1	ENSG00000228013	101928101	IL6R-AS1
ENSG00000224985	NA		ENSG00000226328	100506714	NUP50-DT	ENSG00000228113	NA	
ENSG00000225032	102723566		ENSG00000226380	NA		ENSG00000228223	493812	HCG11
ENSG00000225083	NA	GRTP1-AS1	ENSG00000226416	100133545	MRPL23-AS1	ENSG00000228242	NA	
ENSG00000225138	NA	SLC9A3-AS1	ENSG00000226419	100506392	SLC16A1-AS1	ENSG00000228274	NA	
ENSG00000225313	NA		ENSG00000226476	105378763	LINC01748	ENSG00000228315	91316	GUSBP11
ENSG00000225339	NA		ENSG00000226688	728558	ENTPD1-AS1	ENSG00000228340	284757	MIR646HG
ENSG00000225377	100507459	NRSN2-AS1	ENSG00000226696	104355426	LENG8-AS1	ENSG00000228393	NA	LINC01004
ENSG00000225439	100507171	BOLA3-AS1	ENSG00000226715	NA	LINC01709	ENSG00000228434	NA	
ENSG00000225470	554203	JPX	ENSG00000226833	112267877		ENSG00000228506	NA	
ENSG00000225472	NA		ENSG00000226853	NA		ENSG00000228526	106614088	MIR34AHG
ENSG00000225484	101060691	NUTM2B-AS1	ENSG00000226891	NA	LINC01359	ENSG00000228587	NA	
ENSG00000225613	NA	LINCMD1	ENSG00000226950	57291	DANCR	ENSG00000228606	100287049	
ENSG00000225655	NA		ENSG00000227107	NA		ENSG00000228649	109729180	SNHG26
ENSG00000225670	NA	CADM3-AS1	ENSG00000227128	NA	LBX1-AS1	ENSG00000228748	NA	
ENSG00000225733	100505641	FGD5-AS1	ENSG00000227252	NA		ENSG00000228775	285962	WEE2-AS1

ENSG00000228794	643837	LINC01128	ENSG00000230438	NA	SERPINB9P1	ENSG00000231628	NA
ENSG00000228801	102724330		ENSG00000230454	NA		ENSG00000231663	101927765 COA6-AS1
ENSG00000228830	NA		ENSG00000230479	NA		ENSG00000231711	NA LINC00899
ENSG00000228843	NA		ENSG00000230487	114796	PSMG3-AS1	ENSG00000231721	378805 LINC-PINT
ENSG00000228863	NA		ENSG00000230530	NA	LIMD1-AS1	ENSG00000231742	101927541 LINC01273
ENSG00000228878	NA	SEPTIN7-DT	ENSG00000230537	NA		ENSG00000231768	100506795 LINC01354
ENSG00000229043	NA		ENSG00000230551	NA		ENSG00000231806	NA PCAT7
ENSG00000229047	NA		ENSG00000230555	NA		ENSG00000231856	NA
ENSG00000229108	NA	LINC02587	ENSG00000230590	100302692	FTX	ENSG00000231889	NA TRAF3IP2-AS1
ENSG00000229140	137196	CCDC26	ENSG00000230606	NA		ENSG00000231890	NA DARS-AS1
ENSG00000229140	728724	CCDC26	ENSG00000230630	NA	DNM3OS	ENSG00000231969	NA MMADHC-DT
ENSG00000229140	106144608	CCDC26	ENSG00000230724	NA	LINC01001	ENSG00000232063	NA
ENSG00000229152	NA	ANKRD10-IT1	ENSG00000230733	NA		ENSG00000232079	284825 LINC01697
ENSG00000229388	105378616	LINC01715	ENSG00000230844	NA	ZNF674-AS1	ENSG00000232098	NA
ENSG00000229425	101927745		ENSG00000230896	NA		ENSG00000232160	101928578 RAP2C-AS1
ENSG00000229425	105369302		ENSG00000230910	NA		ENSG00000232164	729348 LINC01873
ENSG00000229444	101929592		ENSG00000230943	101927686	LINC02541	ENSG00000232233	102724699 LINC02043
ENSG00000229589	100128640	ACVR2B-AS1	ENSG00000231013	107105282	SCTR-AS1	ENSG00000232300	NA FAM215B
ENSG00000229619	401093	MBNL1-AS1	ENSG00000231064	NA		ENSG00000232442	100505771 MHENCR
ENSG00000229807	7503	XIST	ENSG00000231074	414777	HCG18	ENSG00000232611	NA
ENSG00000229821	NA		ENSG00000231160	NA	KLF3-AS1	ENSG00000232656	55853 IDI2-AS1
ENSG00000229847	196047	EMX2OS	ENSG00000231312	NA	MAP4K3-DT	ENSG00000232677	100506930 LINC00665
ENSG00000229852	NA		ENSG00000231365	101929147	WARS2-AS1	ENSG00000232774	NA
ENSG00000229951	403150		ENSG00000231367	101929596	LINC02613	ENSG00000232807	NA
ENSG00000229980	400604	TOB1-AS1	ENSG00000231419	154822	LINC00689	ENSG00000232850	389791
ENSG00000230082	100874032	PRRT3-AS1	ENSG00000231527	105379444	FAM27C	ENSG00000232931	NA LINC00342
ENSG00000230091	NA	TMEM254-AS1	ENSG00000231536	NA		ENSG00000232940	414765 HCG25
ENSG00000230148	100874362	HOXB-AS1	ENSG00000231560	400002	CLEC12A-AS1	ENSG00000232956	285958 SNHG15
ENSG00000230387	100505664		ENSG00000231607	8847	DLEU2	ENSG00000232973	285154 CYP1B1-AS1

ENSG00000232977	100506697	LINC00327	ENSG00000234665	NA	LERFS	ENSG00000235888	NA
ENSG00000233006	553103	MIR3936HG	ENSG00000234678	NA	ELF3-AS1	ENSG00000235904	NA
ENSG00000233016	84973	SNHG7	ENSG00000234684	100507495	SDCBP2-AS1	ENSG00000235919	645676
ENSG00000233117	NA	LINC00702	ENSG00000234722	103724390	LINC01287	ENSG00000235927	374987
ENSG00000233178	NA		ENSG00000234741	60674	GAS5	ENSG00000235954	284900
ENSG00000233183	NA		ENSG00000234771	NA	SLC25A25-AS1	ENSG00000236008	101929567
ENSG00000233184	NA		ENSG00000234899	400618	SOX9-AS1	ENSG00000236017	NA
ENSG00000233223	100996842		ENSG00000234912	654434	SNHG20	ENSG00000236088	100874058
ENSG00000233237	79940	LINC00472	ENSG00000235016	100129060	SEMA3F-AS1	ENSG00000236144	100506469
ENSG00000233251	100129434		ENSG00000235027	NA		ENSG00000236200	NA
ENSG00000233396	NA	LINC01719	ENSG00000235033	100505635	DAAM2-AS1	ENSG00000236208	NA
ENSG00000233429	NA	HOTAIRM1	ENSG00000235070	NA		ENSG00000236255	NA
ENSG00000233461	NA		ENSG00000235257	101928153	ITGA9-AS1	ENSG00000236333	NA
ENSG00000233593	105378853	LINC02609	ENSG00000235288	NA		ENSG00000236404	401491
ENSG00000233621	728431	LINC01137	ENSG00000235295	100192420	LINC01634	ENSG00000236540	NA
ENSG00000233695	NA	GAS6-AS1	ENSG00000235314	NA	LINC00957	ENSG00000236581	NA
ENSG00000233871	NA	DLG5-AS1	ENSG00000235381	NA		ENSG00000236618	100306951
ENSG00000233912	NA		ENSG00000235437	92249	LINC01278	ENSG00000236682	100506922
ENSG00000233937	101928649		ENSG00000235513	NA	L3MBTL2-AS1	ENSG00000236753	100506881
ENSG00000234072	NA		ENSG00000235531	100132891	MSC-AS1	ENSG00000236778	NA
ENSG00000234171	NA	RNASEH1-AS1	ENSG00000235535	NA	TRDN-AS1	ENSG00000236810	100506963
ENSG00000234281	NA	LANCL1-AS1	ENSG00000235560	107984875		ENSG00000236819	101060544
ENSG00000234290	NA		ENSG00000235609	NA		ENSG00000236830	100506428
ENSG00000234323	100996590	LINC01505	ENSG00000235652	NA	FBXO30-DT	ENSG00000236833	NA
ENSG00000234327	101928000		ENSG00000235703	NA	LINC00894	ENSG00000236859	254128
ENSG00000234431	NA		ENSG00000235706	400242	DICER1-AS1	ENSG00000236871	751580
ENSG00000234456	100505881	MAGI2-AS3	ENSG00000235823	90271	OLMALINC	ENSG00000236901	81571
ENSG00000234608	51275	MAPKAPK5-AS1	ENSG00000235831	100507582	BHLHE40-AS1	ENSG00000237036	220930
ENSG00000234636	100873985	MED14OS	ENSG00000235865	57000	GSN-AS1	ENSG00000237037	NA

ENSG00000237188	NA		ENSG00000239569	NA	KMT2E-AS1	ENSG00000243155	NA	
ENSG00000237248	100499405	LINC00987	ENSG00000239653	100507062	PSMD6-AS2	ENSG00000243368	NA	MCCC1-AS1
ENSG00000237298	100506866	TTN-AS1	ENSG00000239665	NA		ENSG00000243701	344595	DUBR
ENSG00000237298	101927055	TTN-AS1	ENSG00000239677	NA	PDZRN3-AS1	ENSG00000243926	NA	TIPARP-AS1
ENSG00000237352	NA	LINC01358	ENSG00000240288	100126793	GHRLOS	ENSG00000243960	NA	
ENSG00000237399	100507034	PITRM1-AS1	ENSG00000240291	NA		ENSG00000244041	401232	LINC01011
ENSG00000237491	105378580	LINC01409	ENSG00000240401	NA		ENSG00000244198	NA	ARHGEF35-AS1
ENSG00000237499	100130476	WAKMAR2	ENSG00000240731	NA		ENSG00000244513	NA	
ENSG00000237505	101927891	PKN2-AS1	ENSG00000240801	NA		ENSG00000244625	NA	MIATNB
ENSG00000237560	102723487	LINC01497	ENSG00000240859	100507642		ENSG00000244733	NA	
ENSG00000237686	101929705		ENSG00000240875	730091	LINC00886	ENSG00000244879	NA	GABPB1-AS1
ENSG00000237742	NA		ENSG00000240990	221883	HOXA11-AS	ENSG00000244945	101928445	RUFY1-AS1
ENSG00000237753	NA		ENSG00000241168	NA		ENSG00000244968	100506495	LIFR-AS1
ENSG00000237775	NA	DDR1-DT	ENSG00000241288	101927056	LINC02614	ENSG00000244998	NA	
ENSG00000237943	439949	PRKCQ-AS1	ENSG00000241316	101927111	SUCLG2-AS1	ENSG00000245025	NA	
ENSG00000237945	100506334	LINC00649	ENSG00000241684	NA	ADAMTS9-AS2	ENSG00000245060	729678	LINC00847
ENSG00000237949	NA	LINC00844	ENSG00000241769	100131434	LINC00893	ENSG00000245105	144571	A2M-AS1
ENSG00000238009	NA		ENSG00000241860	NA		ENSG00000245146	NA	MALINC1
ENSG00000238035	NA		ENSG00000241956	102546299		ENSG00000245149	101927612	RNF139-AS1
ENSG00000238045	NA		ENSG00000241990	NA	PRR34-AS1	ENSG00000245156	NA	
ENSG00000238142	105376805		ENSG00000242086	NA	MUC20-OT1	ENSG00000245281	101929066	
ENSG00000238164	NA	TNFRSF14-AS1	ENSG00000242125	8420	SNHG3	ENSG00000245317	100996419	
ENSG00000238197	100506215	PAXB1-AS1	ENSG00000242282	NA		ENSG00000245466	101928075	
ENSG00000238198	100996251	LRIG2-DT	ENSG00000242288	113939925		ENSG00000245498	100507283	
ENSG00000238266	100507127	LINC00707	ENSG00000242539	NA		ENSG00000245532	283131	NEAT1
ENSG00000238273	NA		ENSG00000242588	NA		ENSG00000245556	728769	SCAMP1-AS1
ENSG00000239213	NA	NCK1-DT	ENSG00000242759	NA	LINC00882	ENSG00000245573	497258	BDNF-AS
ENSG00000239219	100128164		ENSG00000242902	110806300	FLNC-AS1	ENSG00000245694	643911	CRNDE
ENSG00000239415	NA		ENSG00000243069	100507524	ARHGEF26-AS1	ENSG00000245694	101927480	CRNDE

ENSG00000245812	101927740	LINC02202	ENSG00000247373	NA	TMED2-DT	ENSG00000249087	NA	ZNF436-AS1
ENSG00000245849	100505648	RAD51-AS1	ENSG00000247400	100289274	DNAJC3-DT	ENSG00000249102	NA	
ENSG00000245910	641638	SNHG6	ENSG00000247516	100505738	MIR4458HG	ENSG00000249249	NA	
ENSG00000245937	644873	LINC01184	ENSG00000247556	729082	OIP5-AS1	ENSG00000249348	100885776	UGDH-AS1
ENSG00000245970	NA		ENSG00000247572	100131067	CKMT2-AS1	ENSG00000249456	NA	
ENSG00000245975	101928725		ENSG00000247679	NA		ENSG00000249464	285419	LINC01091
ENSG00000246022	100862662	ALDH1L1-AS2	ENSG00000247728	NA		ENSG00000249614	NA	LINC02503
ENSG00000246067	NA	RAB30-DT	ENSG00000247765	NA		ENSG00000249669	NA	CARMN
ENSG00000246090	100507053		ENSG00000247796	257396		ENSG00000249673	NA	NOP14-AS1
ENSG00000246174	100289388	KCTD21-AS1	ENSG00000247828	100505894	TMEM161B-AS1	ENSG00000249731	NA	
ENSG00000246263	NA	UBR5-AS1	ENSG00000247903	NA		ENSG00000250041	NA	
ENSG00000246273	283104	SBF2-AS1	ENSG00000247982	283663	LINC00926	ENSG00000250056	NA	LINC01018
ENSG00000246308	101928053		ENSG00000248008	100506668	NRAV	ENSG00000250069	NA	
ENSG00000246339	101929402	EXTL3-AS1	ENSG00000248015	NA		ENSG00000250091	NA	DNAH100S
ENSG00000246430	100507632	LINC00968	ENSG00000248019	285512	FAM13A-AS1	ENSG00000250132	NA	
ENSG00000246451	NA		ENSG00000248049	NA	UBA6-AS1	ENSG00000250159	NA	
ENSG00000246465	NA		ENSG00000248092	NA	NNT-AS1	ENSG00000250208	440119	FZD10-AS1
ENSG00000246523	100506368	FZD4-DT	ENSG00000248275	100507602	TRIM52-AS1	ENSG00000250303	283140	LINC02762
ENSG00000246528	101929759		ENSG00000248309	101929423	MEF2C-AS1	ENSG00000250333	NA	
ENSG00000246560	105377348	UBE2D3-AS1	ENSG00000248323	NA	LUCAT1	ENSG00000250392	100996694	LINC02502
ENSG00000246695	NA	RASSF8-AS1	ENSG00000248360	201853	LINC00504	ENSG00000250397	NA	
ENSG00000246859	100505678	STARD4-AS1	ENSG00000248429	285505	FAM198B-AS1	ENSG00000250451	100874363	HOXC-AS1
ENSG00000246982	NA		ENSG00000248445	NA	SEMA6A-AS1	ENSG00000250497	NA	
ENSG00000246985	144481	SOCS2-AS1	ENSG00000248508	100131089	SRP14-AS1	ENSG00000250616	NA	
ENSG00000247092	NA	SNHG10	ENSG00000248514	NA		ENSG00000250742	400043	LINC02381
ENSG00000247121	NA		ENSG00000248538	157273		ENSG00000250802	NA	ZBED3-AS1
ENSG00000247137	NA		ENSG00000248587	NA	GDNF-AS1	ENSG00000250899	NA	
ENSG00000247240	440288	UBL7-AS1	ENSG00000248636	NA		ENSG00000250900	NA	
ENSG00000247271	729013	ZBED5-AS1	ENSG00000248738	101929237		ENSG00000250903	NA	GMDS-DT

ENSG00000250978	NA	ENSG00000253712	NA	ENSG00000255039	NA	LINC02553
ENSG00000250988	100505616 SNHG21	ENSG00000253716	100507316 MINCR	ENSG00000255135	NA	
ENSG00000251022	NA THAP9-AS1	ENSG00000253738	NA OTUD6B-AS1	ENSG00000255182	NA	
ENSG00000251034	NA	ENSG00000253741	100288181 LNCOC1	ENSG00000255198	NA	SNHG9
ENSG00000251136	101929709	ENSG00000253948	NA VPS13B-DT	ENSG00000255234	NA	
ENSG00000251141	NA MRPS30-DT	ENSG00000253982	NA	ENSG00000255248	399959	MIR100HG
ENSG00000251143	100128494	ENSG00000254154	NA CRYZL2P-SEC16B	ENSG00000255284	171391	
ENSG00000251151	100874365 HOXC-AS3	ENSG00000254162	NA	ENSG00000255310	NA	
ENSG00000251209	91948 LINC00923	ENSG00000254231	105375624	ENSG00000255389	NA	
ENSG00000251257	NA	ENSG00000254258	NA	ENSG00000255426	NA	
ENSG00000251364	100506258	ENSG00000254343	NA	ENSG00000255455	103611081	
ENSG00000251379	NA	ENSG00000254363	101929719	ENSG00000255458	NA	
ENSG00000251417	NA	ENSG00000254369	100133311 HOXA-AS3	ENSG00000255468	102724064	
ENSG00000251562	378938 MALAT1	ENSG00000254428	NA	ENSG00000255471	NA	
ENSG00000251580	NA LINC02482	ENSG00000254473	NA	ENSG00000255495	NA	
ENSG00000251602	100507437	ENSG00000254539	NA	ENSG00000255650	84983	FAM222A-AS1
ENSG00000251615	NA	ENSG00000254614	728975	ENSG00000255717	23642	SNHG1
ENSG00000251665	NA	ENSG00000254635	NA WAC-AS1	ENSG00000255727	NA	LINC01489
ENSG00000251867	NA	ENSG00000254682	NA	ENSG00000255772	101927922	LINC01479
ENSG00000252690	NA	ENSG00000254721	NA	ENSG00000255857	NA	PXN-AS1
ENSG00000253106	NA	ENSG00000254837	100287896	ENSG00000255874	283487	PRECSIT
ENSG00000253115	NA	ENSG00000254860	493900 TMEM9B-AS1	ENSG00000255970	101927901	LINC02421
ENSG00000253200	NA	ENSG00000254873	NA	ENSG00000256028	NA	
ENSG00000253320	NA AZIN1-AS1	ENSG00000254876	100499484 SUGT1P4-STRAGLP	ENSG00000256073	84996	URB1-AS1
ENSG00000253389	NA	ENSG00000254911	619383 SCARNA9	ENSG00000256092	NA	SBNO1-AS1
ENSG00000253552	NA HOXA-AS2	ENSG00000254911	100158262 SCARNA9	ENSG00000256193	100862680	LINC00507
ENSG00000253636	NA	ENSG00000254929	NA	ENSG00000256364	NA	
ENSG00000253645	NA	ENSG00000255036	57653 STRA6LP	ENSG00000256628	100009676	ZBTB11-AS1
ENSG00000253661	NA ZFH4-AS1			ENSG00000257097	NA	CLIP1-AS1

ENSG00000257135	NA	ODC1-DT	ENSG00000258584	283592	FAM181A-AS1	ENSG00000259661	NA
ENSG00000257252	NA		ENSG00000258603	NA		ENSG00000259673	100506686 IQCH-AS1
ENSG00000257261	NA		ENSG00000258604	NA		ENSG00000259768	NA
ENSG00000257298	NA		ENSG00000258610	NA		ENSG00000259820	NA
ENSG00000257303	NA		ENSG00000258634	NA		ENSG00000259828	NA
ENSG00000257337	283335		ENSG00000258655	NA	ARHGAP5-AS1	ENSG00000259865	NA
ENSG00000257354	NA		ENSG00000258675	NA	LINC02308	ENSG00000259877	NA
ENSG00000257379	NA		ENSG00000258711	NA		ENSG00000259881	101927793
ENSG00000257599	101055625	OVCH1-AS1	ENSG00000258727	102724814		ENSG00000259884	NA NR4A1AS
ENSG00000257607	NA		ENSG00000258733	NA	LINC02328	ENSG00000259891	NA
ENSG00000257621	379025	PSMA3-AS1	ENSG00000258768	NA		ENSG00000259943	NA
ENSG00000257663	NA		ENSG00000258938	NA		ENSG00000259953	NA
ENSG00000257698	NA	GIHCG	ENSG00000258952	NA	SALRNA1	ENSG00000259959	NA
ENSG00000257894	NA		ENSG00000259049	NA		ENSG00000259972	NA
ENSG00000257900	NA		ENSG00000259065	NA		ENSG00000259976	NA
ENSG00000257913	105369758	DDN-AS1	ENSG00000259172	NA		ENSG00000259994	NA
ENSG00000258056	105369779		ENSG00000259248	NA	USP3-AS1	ENSG00000260000	NA
ENSG00000258057	100286844	BCDIN3D-AS1	ENSG00000259291	109729181	ZNF710-AS1	ENSG00000260032	647979 NORAD
ENSG00000258168	NA		ENSG00000259319	NA		ENSG00000260035	NA
ENSG00000258232	NA		ENSG00000259366	NA		ENSG00000260077	NA
ENSG00000258283	NA		ENSG00000259380	NA		ENSG00000260083	101928736 MIR762HG
ENSG00000258301	100506603	VASH1-AS1	ENSG00000259456	101927631	ADNP-AS1	ENSG00000260197	NA
ENSG00000258377	NA		ENSG00000259488	NA		ENSG00000260231	100134229 KDM7A-DT
ENSG00000258430	NA		ENSG00000259495	NA		ENSG00000260233	NA ZNRD2-AS1
ENSG00000258441	283624	LINC00641	ENSG00000259498	NA	TPM1-AS	ENSG00000260236	NA
ENSG00000258498	64150	DIO3OS	ENSG00000259583	101927751		ENSG00000260244	NA
ENSG00000258515	NA		ENSG00000259623	NA		ENSG00000260257	NA
ENSG00000258545	101928969	RHOXF1-AS1	ENSG00000259642	NA	ST20-AS1	ENSG00000260260	100507303 SNHG19
ENSG00000258559	NA		ENSG00000259659	NA		ENSG00000260267	NA

ENSG00000260274	NA		ENSG00000260774	NA		ENSG00000261211	NA
ENSG00000260280	100526831	SLX1B-SULT1A4	ENSG00000260793	NA		ENSG00000261215	NA
ENSG00000260285	NA		ENSG00000260804	150967	LINC01963	ENSG00000261220	NA
ENSG00000260296	NA		ENSG00000260805	NA		ENSG00000261253	100287036
ENSG00000260317	NA		ENSG00000260807	115804232	CEROX1	ENSG00000261269	NA
ENSG00000260329	NA		ENSG00000260834	NA		ENSG00000261324	NA
ENSG00000260337	NA		ENSG00000260852	283932	FBXL19-AS1	ENSG00000261326	NA LINC01355
ENSG00000260391	NA		ENSG00000260853	NA		ENSG00000261338	NA
ENSG00000260398	NA		ENSG00000260855	NA		ENSG00000261386	NA
ENSG00000260400	NA		ENSG00000260912	NA		ENSG00000261423	105370888 TMEM202-AS1
ENSG00000260401	NA		ENSG00000260917	103344931		ENSG00000261434	NA
ENSG00000260425	NA		ENSG00000260923	NA	LINC02193	ENSG00000261455	NA LINC01003
ENSG00000260442	100289092	ATP2A1-AS1	ENSG00000260942	NA	CAPN10-DT	ENSG00000261460	NA
ENSG00000260448	NA	LCMT1-AS1	ENSG00000260948	NA		ENSG00000261485	NA PAN3-AS1
ENSG00000260465	NA		ENSG00000260966	NA		ENSG00000261490	NA
ENSG00000260526	NA		ENSG00000260971	NA		ENSG00000261505	NA
ENSG00000260528	NA	FAM157C	ENSG00000261054	NA		ENSG00000261512	NA
ENSG00000260563	NA		ENSG00000261067	NA		ENSG00000261526	NA
ENSG00000260565	NA	ERVK13-1	ENSG00000261069	NA		ENSG00000261534	NA
ENSG00000260566	NA		ENSG00000261087	NA	ZNNT1	ENSG00000261553	NA
ENSG00000260604	NA		ENSG00000261094	NA		ENSG00000261584	NA
ENSG00000260618	NA		ENSG00000261098	NA		ENSG00000261613	NA
ENSG00000260630	197187	SNAI3-AS1	ENSG00000261123	NA		ENSG00000261616	NA
ENSG00000260641	NA		ENSG00000261126	NA	RBFADN	ENSG00000261659	NA
ENSG00000260669	NA		ENSG00000261167	NA		ENSG00000261663	NA
ENSG00000260708	NA		ENSG00000261168	NA		ENSG00000261684	NA
ENSG00000260711	NA		ENSG00000261175	NA	LINC02188	ENSG00000261759	NA
ENSG00000260751	NA		ENSG00000261188	NA		ENSG00000261799	NA
ENSG00000260772	NA		ENSG00000261200	NA		ENSG00000261804	NA

ENSG00000261824	148189	LINC00662	ENSG00000263489	NA	ENSG00000266947	NA		
ENSG00000261845	NA		ENSG00000263731	NA	ENSG00000266962	108783654		
ENSG00000261879	100130950		ENSG00000263753	339290	LINC00667	ENSG00000266993	NA	
ENSG00000261971	NA	MMP25-AS1	ENSG00000263873	NA	THY1-AS1	ENSG00000267002	NA	
ENSG00000262061	NA		ENSG00000264112	NA		ENSG00000267040	100505549	
ENSG00000262160	NA		ENSG00000264151	NA		ENSG00000267058	100505715	
ENSG00000262370	NA		ENSG00000264247	NA	LINC00909	ENSG00000267075	NA	
ENSG00000262410	NA		ENSG00000264456	NA		ENSG00000267080	339201	ASB16-AS1
ENSG00000262420	NA		ENSG00000264490	NA		ENSG00000267100	147727	ILF3-DT
ENSG00000262454	NA	MIR193BHG	ENSG00000264575	NA	LINC00526	ENSG00000267106	NA	ZNF561-AS1
ENSG00000262468	NA	LINC01569	ENSG00000264772	NA		ENSG00000267107	100505495	PCAT19
ENSG00000262533	NA		ENSG00000264920	102724532		ENSG00000267121	339192	
ENSG00000262580	NA		ENSG00000265142	102723167	MIR133A1HG	ENSG00000267152	NA	
ENSG00000262691	NA		ENSG00000265206	NA		ENSG00000267160	NA	
ENSG00000262879	NA		ENSG00000265287	NA		ENSG00000267165	NA	CHMP1B-AS1
ENSG00000262967	NA		ENSG00000265399	NA		ENSG00000267169	100507373	
ENSG00000263004	NA		ENSG00000265413	NA		ENSG00000267199	NA	
ENSG00000263069	100294362	RNF213-AS1	ENSG00000265479	441263	DTX2P1- UPK3BP1- PMS2P11	ENSG00000267247	NA	
ENSG00000263072	NA	ZNF213-AS1	ENSG00000265751	NA		ENSG00000267257	NA	
ENSG00000263089	NA		ENSG00000265778	101927989		ENSG00000267265	NA	
ENSG00000263126	NA		ENSG00000265992	790952	ESRG	ENSG00000267272	339524	LINC01140
ENSG00000263165	NA		ENSG00000266208	NA		ENSG00000267296	80054	CEBPA-DT
ENSG00000263244	NA		ENSG00000266340	NA		ENSG00000267302	101927755	RNFT1-DT
ENSG00000263272	NA		ENSG00000266718	NA		ENSG00000267309	728752	
ENSG00000263276	NA		ENSG00000266872	NA		ENSG00000267317	NA	
ENSG00000263345	NA		ENSG00000266896	NA		ENSG00000267321	NA	SNHG30
ENSG00000263370	NA		ENSG00000266904	NA	LINC00663	ENSG00000267322	677769	SNHG22
ENSG00000263400	101101775	TMEM220-AS1	ENSG00000266923	NA		ENSG00000267322	103091864	SNHG22
ENSG00000263412	NA	NFE2L1-DT				ENSG00000267325	NA	LINC01415

ENSG00000267328	NA	ENSG00000268119	NA	ENSG00000269609	100505761	RPARP-AS1	
ENSG00000267364	NA	ENSG00000268129	NA	ENSG00000269696	NA		
ENSG00000267365	400617	KCNJ2-AS1	ENSG00000268199	NA	ENSG00000269821	10984	KCNQ10T1
ENSG00000267383	NA	ENSG00000268205	NA	ENSG00000269825	NA		
ENSG00000267390	NA	ENSG00000268362	NA	ENSG00000269834	NA	ZNF528-AS1	
ENSG00000267414	101927943	SETBP1-DT	ENSG00000268403	644656	ENSG00000269867	NA	
ENSG00000267423	NA	ENSG00000268471	54553	MIR4453HG	ENSG00000269893	100093630	SNHG8
ENSG00000267469	NA	ENSG00000268516	105372482		ENSG00000269900	6023	RMRP
ENSG00000267470	100507433	ZNF571-AS1	ENSG00000268518	NA	ENSG00000269910	NA	
ENSG00000267480	NA	ENSG00000268555	NA	ENSG00000269918	NA		
ENSG00000267481	NA	ENSG00000268575	NA	ENSG00000269929	NA	MIRLET7A1HG	
ENSG00000267519	NA	ENSG00000268628	NA	ENSG00000269930	NA		
ENSG00000267520	NA	ENSG00000268713	NA	ENSG00000269934	NA		
ENSG00000267532	100506755	MIR497HG	ENSG00000268858	112268269	ENSG00000269937	NA	
ENSG00000267575	101927151		ENSG00000268912	NA	ENSG00000269940	NA	
ENSG00000267632	NA	ENSG00000268947	NA	ENSG00000269945	NA		
ENSG00000267633	NA	ENSG00000268996	100289341	MAN1B1-DT	ENSG00000269958	NA	
ENSG00000267784	NA	ENSG00000269044	NA	ENSG00000269973	NA		
ENSG00000267801	NA	ENSG00000269176	NA	ENSG00000269982	NA		
ENSG00000267834	NA	ENSG00000269243	NA	ENSG00000269983	NA		
ENSG00000267858	100131691	MZF1-AS1	ENSG00000269293	100129195	ZSCAN16-AS1	ENSG00000269994	440173
ENSG00000267871	105372476	ZNF460-AS1	ENSG00000269352	NA	PTOV1-AS2	ENSG00000270012	NA
ENSG00000267904	NA	ENSG00000269386	100507567	RAB11B-AS1	ENSG00000270021	NA	
ENSG00000267934	NA	ENSG00000269399	NA	ENSG00000270039	NA		
ENSG00000268001	100505812	CARD8-AS1	ENSG00000269439	100507551		ENSG00000270049	101927837
ENSG00000268006	100506033	PTOV1-AS1	ENSG00000269473	NA	ENSG00000270055	NA	
ENSG00000268030	NA	ENSG00000269486	NA	ERVK9-11	ENSG00000270060	NA	
ENSG00000268061	100505681	NAPA-AS1	ENSG00000269514	NA	ENSG00000270074	NA	
ENSG00000268108	NA	ENSG00000269604	NA	ENSG00000270110	NA		

ENSG00000270112	NA	ENSG00000271795	NA	ENSG00000272273	NA	IER3-AS1	
ENSG00000270127	NA	ENSG00000271811	NA	ENSG00000272288	101929243		
ENSG00000270175	NA	ENSG00000271816	729096	ENSG00000272316	NA		
ENSG00000270194	152048	ENSG00000271848	NA	ENSG00000272335	NA		
ENSG00000270344	NA	POC1B-AS1	ENSG00000271851	NA	ENSG00000272341	NA	
ENSG00000270362	100288198	HMG3-AS1	ENSG00000271880	119385	AGAP11	ENSG00000272356	NA
ENSG00000270412	NA	ENSG00000271894	NA	ENSG00000272374	NA		
ENSG00000270419	100526820	CAHM	ENSG00000271895	NA	ENSG00000272447	642361	
ENSG00000270580	105369154	ENSG00000271918	NA	ENSG00000272455	NA	MRPL20-DT	
ENSG00000270605	NA	ENSG00000271959	NA	ENSG00000272462	NA		
ENSG00000270638	NA	ENSG00000271976	NA	ENSG00000272501	NA		
ENSG00000270641	9383	TSIX	ENSG00000272010	NA	ENSG00000272505	NA	
ENSG00000270659	NA	ENSG00000272030	NA	ENSG00000272512	NA		
ENSG00000270681	NA	ENSG00000272054	NA	ENSG00000272540	NA		
ENSG00000270696	NA	ENSG00000272070	NA	ENSG00000272599	NA		
ENSG00000270820	NA	ENSG00000272077	NA	ENSG00000272604	NA		
ENSG00000270959	339929	LPP-AS2	ENSG00000272078	NA	ENSG00000272630	NA	
ENSG00000271009	NA	ENSG00000272079	NA	ENSG00000272631	NA		
ENSG00000271122	101930085	ENSG00000272086	NA	ENSG00000272638	NA		
ENSG00000271147	NA	ARMCX5-GPRASP2	ENSG00000272106	NA	ENSG00000272654	NA	
ENSG00000271270	100507032	TMCC1-AS1	ENSG00000272114	NA	ENSG00000272668	107985216	
ENSG00000271344	NA	ENSG00000272129	NA	ENSG00000272686	NA	WASL-DT	
ENSG00000271420	NA	ENSG00000272140	NA	ENSG00000272688	NA		
ENSG00000271533	NA	ENSG00000272144	NA	ENSG00000272695	100506394	GAS6-DT	
ENSG00000271553	NA	ENSG00000272145	NA	NFYC-AS1	ENSG00000272720	NA	
ENSG00000271576	NA	ENSG00000272168	NA	CASC15	ENSG00000272732	NA	
ENSG00000271614	NA	ATP2B1-AS1	ENSG00000272173	NA	ENSG00000272734	NA	ADIRF-AS1
ENSG00000271643	NA	ENSG00000272195	NA	ENSG00000272752	101752399	STAG3L5P-PVRIG2P-PILRB	
ENSG00000271646	NA	ENSG00000272221	NA	ENSG00000272758	NA		

ENSG00000272760	NA	ENSG00000273275	NA	ENSG00000274220	NA
ENSG00000272789	NA	ENSG00000273295	NA	ENSG00000274265	NA
ENSG00000272848	NA	ENSG00000273301	NA	ENSG00000274281	NA
ENSG00000272884	NA	ENSG00000273314	NA	ENSG00000274292	NA
ENSG00000272888	NA LINC01578	ENSG00000273319	NA	ENSG00000274333	102724219
ENSG00000272894	NA	ENSG00000273329	NA	ENSG00000274422	NA
ENSG00000272909	NA	ENSG00000273344	202781 PAXIP1-AS1	ENSG00000274460	NA
ENSG00000272918	NA	ENSG00000273345	NA	ENSG00000274536	NA MIR223HG
ENSG00000272941	NA	ENSG00000273356	NA LINC02019	ENSG00000274565	NA
ENSG00000272975	NA MYHAS	ENSG00000273372	NA SFTPD-AS1	ENSG00000274605	105370333 PCCA-DT
ENSG00000272977	NA	ENSG00000273373	NA	ENSG00000274776	NA
ENSG00000272983	NA	ENSG00000273374	NA	ENSG00000274828	NA
ENSG00000272990	NA	ENSG00000273381	NA	ENSG00000274925	NA ZKSCAN2-DT
ENSG00000272994	NA	ENSG00000273448	NA	ENSG00000275120	NA
ENSG00000273015	NA	ENSG00000273466	NA	ENSG00000275155	NA
ENSG00000273018	80039 FAM106A	ENSG00000273486	NA	ENSG00000275198	NA
ENSG00000273033	100129550 LINC02035	ENSG00000273487	NA	ENSG00000275216	NA
ENSG00000273038	NA	ENSG00000273489	NA	ENSG00000275329	NA
ENSG00000273066	NA	ENSG00000273576	NA	ENSG00000275367	NA
ENSG00000273108	NA	ENSG00000273599	NA	ENSG00000275383	NA
ENSG00000273142	NA LINC02604	ENSG00000273702	NA	ENSG00000275494	NA
ENSG00000273148	NA LINC00653	ENSG00000273711	NA	ENSG00000275496	102724701
ENSG00000273151	NA	ENSG00000273723	NA SUGT1-DT	ENSG00000275512	NA
ENSG00000273156	NA	ENSG00000273729	NA	ENSG00000275672	NA
ENSG00000273221	NA	ENSG00000273812	NA	ENSG00000275734	NA
ENSG00000273230	NA	ENSG00000273906	NA	ENSG00000275759	NA
ENSG00000273247	NA	ENSG00000274020	NA LINC01138	ENSG00000275764	NA
ENSG00000273267	NA	ENSG00000274080	NA	ENSG00000275765	NA
ENSG00000273270	NA	ENSG00000274104	NA	ENSG00000275897	NA

ENSG00000276077	102724951	ENSG00000277938	NA	ENSG00000279168	NA
ENSG00000276232	NA	ENSG00000277954	NA	ENSG00000279175	NA
ENSG00000276248	NA	ENSG00000277969	NA	ENSG00000279232	NA
ENSG00000276278	NA	ENSG00000277991	102723360	ENSG00000279278	NA
ENSG00000276308	NA	ENSG00000278058	NA	ENSG00000279442	NA
ENSG00000276445	NA	ENSG00000278126	NA	ENSG00000279484	NA KLHL30-AS1
ENSG00000276488	NA	ENSG00000278133	NA	ENSG00000279529	NA
ENSG00000276564	NA	ENSG00000278156	641467 TSC22D1-AS1	ENSG00000279668	NA
ENSG00000276570	NA	ENSG00000278175	389741 GLIDR	ENSG00000279738	NA
ENSG00000276698	NA	ENSG00000278291	NA	ENSG00000279833	NA
ENSG00000276724	NA	ENSG00000278376	NA	ENSG00000280007	NA
ENSG00000276728	NA	ENSG00000278390	101929140	ENSG00000280018	NA
ENSG00000276791	NA	ENSG00000278445	NA	ENSG00000280109	191585 PLAC4
ENSG00000276900	NA	ENSG00000278464	NA	ENSG00000280145	NA
ENSG00000276997	NA	ENSG00000278576	NA	ENSG00000280213	100113386 UCKL1-AS1
ENSG00000277007	NA	ENSG00000278600	NA	ENSG00000280279	NA
ENSG00000277067	102724843	ENSG00000278709	105416157 NKILA	ENSG00000280341	NA
ENSG00000277142	NA LINC00235	ENSG00000278730	NA	ENSG00000280383	NA
ENSG00000277200	NA	ENSG00000278768	NA BACE1-AS	ENSG00000280434	NA
ENSG00000277283	NA	ENSG00000278784	NA	ENSG00000280441	NA
ENSG00000277449	NA CEBPB-AS1	ENSG00000278834	NA	ENSG00000280474	NA
ENSG00000277476	NA	ENSG00000278878	NA	ENSG00000280614	NA
ENSG00000277543	NA	ENSG00000278903	NA	ENSG00000280634	102659353 THRIL
ENSG00000277701	NA	ENSG00000278970	100859930 HEIH	ENSG00000280734	NA LINC01232
ENSG00000277715	NA	ENSG00000279066	NA HEXD-IT1	ENSG00000280739	440952 EIF1B-AS1
ENSG00000277767	NA	ENSG00000279080	NA	ENSG00000280798	283267 LINC00294
ENSG00000277801	NA	ENSG00000279094	105379487 LINC01670	ENSG00000280800	NA
ENSG00000277870	653203 FAM230A	ENSG00000279145	NA	ENSG00000280927	285463 CTBP1-AS
ENSG00000277879	NA	ENSG00000279159	NA	ENSG00000281026	116828 N4BP2L2-IT2

ENSG00000281181	NA		ENSG00000283696	NA		ENSG00000285184	NA	
ENSG00000281183	NA	NPTN-IT1	ENSG00000283828	102724474		ENSG00000285219	100506207	HULC
ENSG00000281332	NA	LINC00997	ENSG00000283907	NA		ENSG00000285331	145783	
ENSG00000281344	NA	HELLPAR	ENSG00000284116	NA		ENSG00000285336	101928882	
ENSG00000281383	NA		ENSG00000284196	NA		ENSG00000285410	55056	GABPB1-IT1
ENSG00000281392	100846978		ENSG00000284237	NA	LINC02767	ENSG00000285533	105369347	RELA-DT
ENSG00000281398	NA	SNHG4	ENSG00000284294	NA		ENSG00000285540	NA	
ENSG00000281453	103611157	TGFB2-OT1	ENSG00000284391	105373244		ENSG00000285541	NA	
ENSG00000281501	NA	SEPSECS-AS1	ENSG00000284428	NA		ENSG00000285564	NA	
ENSG00000281649	100506710	EBLN3P	ENSG00000284602	NA		ENSG00000285596	NA	
ENSG00000281691	100775107	RBM5-AS1	ENSG00000284606	NA		ENSG00000285608	NA	
ENSG00000281852	NA	LINC00891	ENSG00000284624	NA		ENSG00000285628	NA	
ENSG00000281904	NA		ENSG00000284633	NA		ENSG00000285667	NA	
ENSG00000282024	NA		ENSG00000284642	NA		ENSG00000285669	NA	
ENSG00000282033	105373553		ENSG00000284650	NA		ENSG00000285725	NA	
ENSG00000282393	NA		ENSG00000284669	NA		ENSG00000285752	NA	CDC42-AS1
ENSG00000282508	NA	LINC01002	ENSG00000284672	NA		ENSG00000285756	NA	
ENSG00000282556	101927825		ENSG00000284693	100506022	LINC02606	ENSG00000285774	NA	
ENSG00000282840	NA		ENSG00000284707	NA		ENSG00000285793	285074	
ENSG00000282851	105221694	BISPR	ENSG00000284734	NA		ENSG00000285796	NA	
ENSG00000283078	NA		ENSG00000284735	NA		ENSG00000285803	NA	
ENSG00000283103	NA		ENSG00000284879	NA		ENSG00000285813	NA	
ENSG00000283175	101929130		ENSG00000284959	NA		ENSG00000285830	NA	
ENSG00000283294	NA		ENSG00000284966	NA		ENSG00000285844	NA	
ENSG00000283341	NA		ENSG00000284968	NA		ENSG00000285864	NA	
ENSG00000283445	NA		ENSG00000285051	101928583	SLC7A14-AS1	ENSG00000285886	NA	
ENSG00000283633	NA		ENSG00000285103	NA		ENSG00000285928	NA	
ENSG00000283662	NA		ENSG00000285106	NA		ENSG00000285967	646719	NIPBL-DT
ENSG00000283674	729732		ENSG00000285155	NA		ENSG00000285979	NA	

ENSG00000286004	NA
ENSG00000286044	NA
ENSG00000286067	NA
ENSG00000286078	NA
ENSG00000286104	NA
ENSG00000286122	NA
ENSG00000286125	NA
ENSG00000286128	NA
ENSG00000286129	NA
ENSG00000286130	NA
ENSG00000286147	NA
ENSG00000286162	NA
ENSG00000286177	105372401
ENSG00000286191	NA
ENSG00000286214	NA

