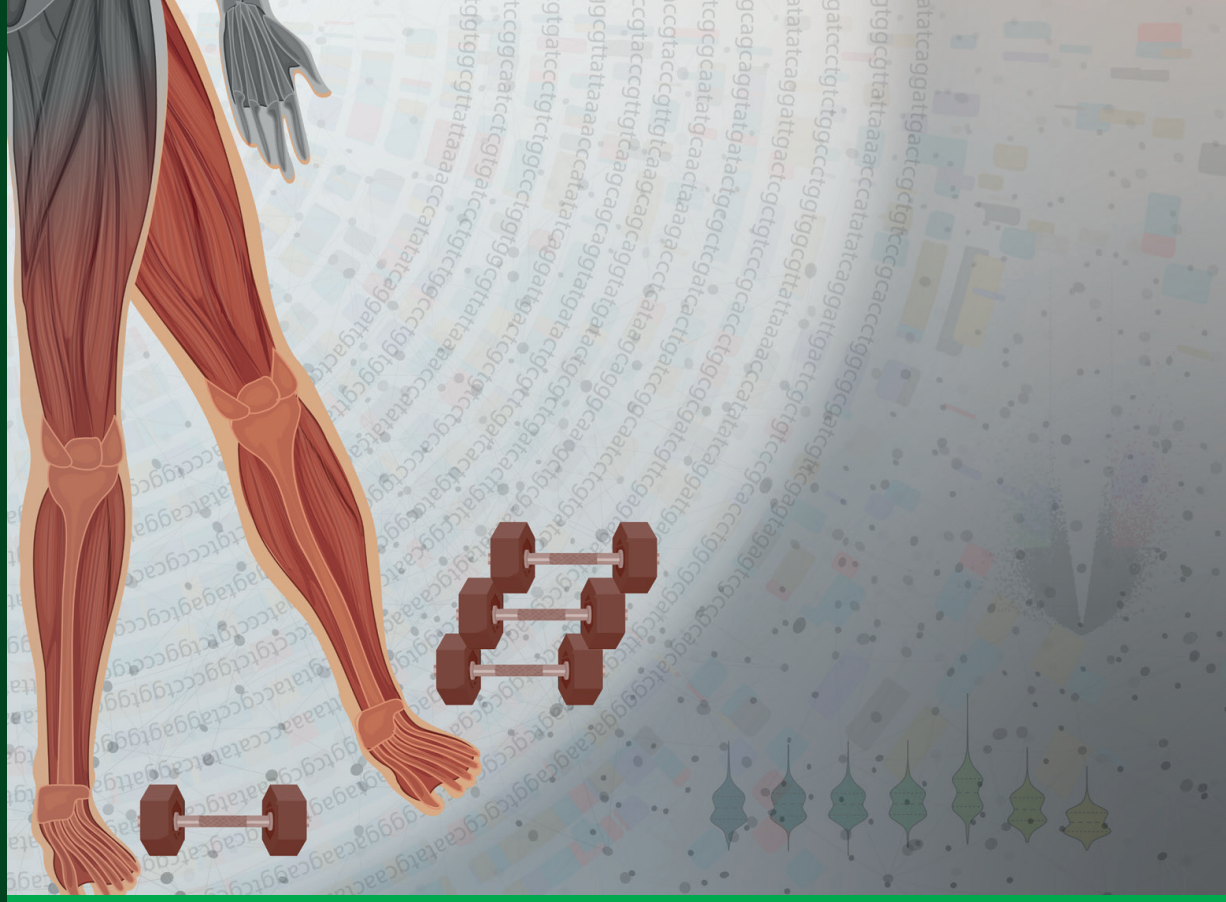
Inland Norway University of Applied Sciences



Faculty of Applied Ecology, Agricultural Sciences and Biotechnology

Yusuf Khan

Computational strategies to study skeletal muscle transcriptome responses to resistance training

PhD Dissertations in Applied Ecology and Biotechnology
2021

Ph.d.-avhandlinger i anvendt økologi og bioteknologi
PhD Dissertations in Applied Ecology and Biotechnology

No. 1 - 2021 **Kauê de Sousa:** Agrobiodiversity and climate adaptation: Insights for risk management in smallscale farming

No. 2 - 2021 **Gjermund Gomo:** The Vertebrate Scavenger Community Along a Boreal Forest-Alpine Gradient: The Importance of Ungulate Management, Small Rodent Cycles and Winter Climate

No. 3 - 2021 **Ana Maria Peris Tamayo:** Adaptive radiation of Arctic charr *(Salvelinus alpinus)* in three Norwegian lakes - niche segregation, phenotypic and genetic variation

No. 4 - 2021 **Anne Elizabeth Loosen:** Spatial ecology of moose *(Alces alces)* in a dynamic world

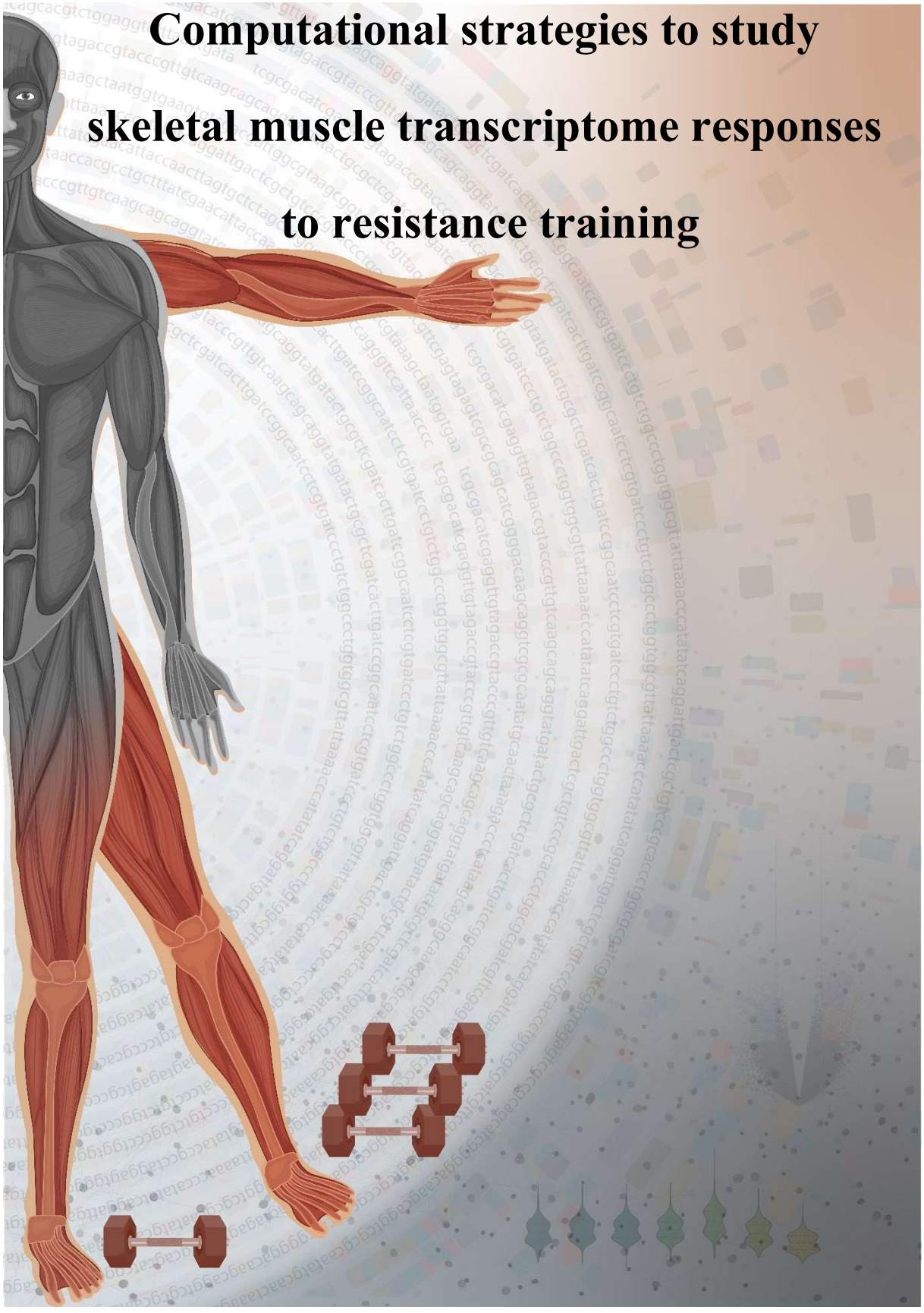
No. 5 - 2021 **Jérémy Monsimet:** Predicting future changes in the distribution of wandering spiders. Habitat requirement, cold tolerance, and dispersal are crucial traits for *Dolomedes* species.

No. 6 - 2021 **Ilaria de Meo:** Ecological mechanisms driving the anti-predator defense response in Crucian carp *(Carassius carassius)* - Variation in morphology, resource use, and life-history strategies along a gradient of predation risk

No. 7 - 2021 **Altamash Bashir:** Non-industrial private forest owners' attitudes and forest management decisions

No. 8 - 2021 **Yusuf Khan:** Computational strategies to study skeletal muscle transcriptome responses to resistance training

Computational strategies to study skeletal muscle transcriptome responses to resistance training



Yusuf Khan

Computational strategies to study skeletal muscle transcriptome responses to resistance training

Doctoral Thesis

2021

Inland Norway University of Applied Sciences

Department of Biotechnology, Hamar

Section for Health and Exercise Physiology, Lillehammer.

Supervisor

Assoc. Prof. Rafi Ahmad

Department of Biotechnology

Faculty of Applied Ecology, Agricultural Sciences, and Biotechnology

Hamar campus

Inland Norway University of Applied Science

Co-supervisors

Prof. Stian Ellefsen

Section for Health and exercise physiology

Faculty of Health and Social Sciences

Lillehammer campus

Inland Norway University of Applied Sciences

Assoc. Prof. Daniel Hammarström

Section for Health and exercise physiology

Faculty of Health and Social Sciences

Lillehammer campus

Inland Norway University of Applied Sciences

Printed by: Flisa Trykkeri A/S
Place of publication: Hamar

Ph.D. Thesis in Applied Ecology and biotechnology no. 8 - 2021
ISBN printed version: 978-82-8380-310-5
ISBN digital version: 978-82-8380-311-2
ISSN printed version: 2703-819X
ISSN digital version: 2703-8181

Table of contents

Acknowledgement

First and foremost, thank you to my supervisors — **Rafi Ahmad** and **Stian Ellefsen**. It has been my privilege to learn from you. Your inspirational leadership, wisdom, gentle guidance, generosity with time, and enthusiastic encouragement has formed both the way I approach scientific research as well as life in general. I am forever grateful to you both.

Thank you to colleagues – **Abdulrahman Khezri, Daniel Hammarström, Ekaterina Avershina** and **Knut Sindre Mølmen** for your active interest in my research, meaningful feedback, and guidance in my maturation as a scientist. I also wish to express my appreciation to the PhD committee members – **Øyvind Gunnar Antonsen, Henriette Wathne Gelink** and **Kaja Johnsen**. It will be remiss of me not to highlight and acknowledge my gratitude to the dedicated staff of the **HSV** and **ALB**, **Inland Norway University of Applied Sciences**. Without your support, this research would not have been possible.

I want to thank **Marit Roland, Håvard Nygard, Mona Ødegaard, Heidi Bråton,** and **Sjur Johansen Øfsteng** for helping in my initial days in Lillehammer, Norway. Many thanks go to my fellow student **Reina Jochems**; and my neighbours and the master students - **Anand Dev, Mithlesh Ray, Ambreen Kausar** and **Nausheen Kiran**, who have provided friendship and support, and with whom I have shared laughter, frustration and companionship, thanks for the pleasant company in Fredrik. I want to thank **Rafi's family,** who constantly supported my family and me during all good and bad times.

I cannot articulate my gratitude to my family for the love, support, encouragement and prayers they have sent my way along this journey. To my **parents** and **siblings**, thank you for being my champions. I hope I have made you proud. To my wife, **Shaheen,** for your unceasing support and encouragement. To my daughter **Aliza** for being there at the end of the day and for being an inspiration to achieve greatness. To all my friends for their support.

Finally, I am grateful to have had the privilege of attending the **Inland Norway University of Applied Sciences**. This experience has allowed me to work with some of the best and brightest and provided the resources for me to achieve great success. Thank you for this opportunity.

Summary in English

Lifestyle-related diseases are widespread, and their incidences are increasing at an alarming rate in the human population. These diseases are associated with ageing, a sedentary lifestyle and/or poor diet and are usually accompanied by loss of muscle strength, muscle mass, and changes in muscle biology. The latter may be treated and prevented by weight-bearing training exercises (resistance training), which stimulates growth and maintain the health of the human skeletal muscle, the largest organ in the body.

RNA-seq is a widely used method to study cellular characteristics and changes thereof (including muscle plasticity) under different conditions at a given time point. It produces a large amount of data, and many tools and different pipelines have been used to facilitate analysis. Several studies have focused on comparing different tools for RNA-seq data analysis. However, currently, there is no pipeline specific for analysing skeletal muscle data or transcriptome response to resistance training. Therefore, it is crucial to identify tools specifically for resistance training-induced growth in skeletal muscle. Moreover, each dataset is different and needs careful evaluation of tools to decide which is best for a given project.

The main aim of the thesis was to develop RNA-seq data analysis strategies to analyse skeletal muscle transcriptome responses to resistance training and to employ these to study two different datasets. YOUNG dataset (18 – 40 years) – RNA-seq data from muscle biopsies collected from participants trained for 12 weeks with two different training volumes. For this dataset, the aim was to investigate the impacts of training volume for transcriptome responses. OLDER dataset (56 - 77 years) – RNA-seq from muscle biopsies collected from participants, both healthy and with chronic obstructive pulmonary disease diagnosis, which were trained for 13 weeks with two different training loads and provided with either vitamin D_3 or placebo supplementation. This dataset aimed to investigate the impact of vitamin D_3 supplementation and COPD diagnosis for transcriptome responses.

The YOUNG dataset was used to develop the standardised RNA-seq pipeline. Bioinformatic tools were selected based on read quality, observed gene counts, methodological variation between paired observations, and correlations between mRNA abundance and protein expression of myosin heavy chain family proteins. Based on exploring the effects of using different normalisation strategies, our results suggest that it may be necessary to account for biological biasness caused by global changes in the total RNA population (per tissue mass) to increase the biological relevance of transcriptome analyses.

This developed pipeline was used on the YOUNG and OLDER datasets to understand better muscle transcriptome responses to the above-mentioned resistance training interventions. In the YOUNG dataset, the two different resistance training volumes were associated with differential expression of 21 genes at week 2, mostly related to extracellular remodelling. In contrast, no readily explainable dose-dependencies were observed at Week 12. In the OLDER dataset, neither vitamin D_3 supplementation nor COPD diagnosis affected transcriptome responses to training, as estimated at the single-gene level. However, enrichment analyses revealed that both conditions were associated with slight effects on response patterns. For example, vitamin D_3 was associated with increased expression of gene sets involved in endothelial proliferation and blood vessel morphogenesis compared to placebo (Week 13). COPD was associated with more pronounced decreases in gene sets relating to oxidative phosphorylation and myogenesis.

The results from this study could potentially help understand transcriptomic changes of training and factors affecting training response, which could help design better training protocols for healthy ageing and life.

Sammendrag på Norsk

Livsstilsrelaterte sykdommer er utbredt, og forekomsten øker i en alarmerende hastighet i den menneskelige befolkningen. Disse sykdommene er forbundet med aldring, en stillesittende livsstil og/eller dårlig kosthold, og er vanligvis ledsaget av tap av muskelstyrke, muskelmasse og endringer i muskelbiologi. Sistnevnte kan behandles og forhindres av vektbærende treningsøvelser (motstandstrening), som stimulerer vekst og opprettholder helsen til menneskelig skjelettmuskel, det største organet i kroppen.

RNA-seq er en mye brukt metode for å studere cellulære egenskaper og endringer derav (inkludert muskelplastisitet) under forskjellige forhold på et gitt tidspunkt. Den produserer en stor mengde data, og mange verktøy har blitt brukt for å lette analysen. Flere studier har fokusert på å sammenligne forskjellige verktøy for RNA-seq analyse. Imidlertid er det for øyeblikket ingen pipeline spesifikk for å analysere skjelettmuskeldata eller transkriptomrespons på motstandstrening. Derfor er det avgjørende å identifisere verktøy spesielt for motstandstreningsindusert vekst i skjelettmuskulatur. I tillegg er hvert datasett annerledes og trenger grundig evaluering av verktøy for å bestemme hvilket som er best for et gitt prosjekt.

Hovedmålet med oppgaven var å utvikle RNA-seq dataanalyse strategier for å analysere transkriptomresponser av skjelettmuskulatur på motstandstrening og å bruke disse til å studere to forskjellige datasett. YOUNG datasett (18 - 40 år) - RNA-seq data fra muskelbiopsier samlet fra deltakere trent i 12 uker med to forskjellige treningsvolumer. For dette datasettet var målet å undersøke virkningene av treningsvolum for transkriptomresponser. OLDER datasett (56 - 77 år) - RNA-seq fra muskelbiopsier samlet fra deltakere, både friske og med kronisk obstruktiv lungesykdomsdiagnose (KOL), som ble trent i 13 uker med to forskjellige treningsbelastninger og utstyrt med enten vitamin D_3 eller placebo tilskudd. For dette datasettet var målet å undersøke virkningen av vitamin D_3-tilskudd og KOL-diagnose for transkriptomresponser.

Det YOUNG datasettet ble brukt til å utvikle den standardiserte RNA-seq pipeline. Bioinformatiske verktøy ble valgt ut fra lesekvalitet, observerte gentall, metodologisk variasjon mellom parede observasjoner og korrelasjoner mellom mRNA-overflod og proteinuttrykk av myosin tungkjede familieproteiner. Basert på å utforske effekten av å bruke forskjellige normaliseringsstrategier, antyder våre resultater at det kan være nødvendig å ta hensyn til biologisk skjevhet forårsaket av globale endringer i den totale RNA (per vevsmasse) for å øke den biologiske relevansen av transkriptomanalyser.

Denne utviklede pipeline ble brukt på de YOUNG og OLDER datasettene for å gi en bedre forståelse av muskeltransskriptomresponser. I det YOUNG datasettet var de to

forskjellige motstandstreningsvolumene assosiert med differensial uttrykk av 21 gener i uke 2, hovedsakelig relatert til ekstracellulær remodellering. I motsetning ble det ikke observert noen lett forklarbare doseavhengigheter i uke 12. I det OLDER datasettet påvirket verken vitamin D_3-tilskudd eller KOL-diagnose transkriptomresponsene på trening, når estimert på enkeltgenivå. Anrikningsanalyser viste imidlertid at begge forholdene var forbundet med små effekter på responsmønstre. For eksempel var vitamin D_3 assosiert med økt uttrykk av gensett som er involvert i endotelproliferasjon og blodkarmorfogenese sammenlignet med placebo (uke 13), og KOL var assosiert med mer markante reduksjoner i gensett knyttet til oksidativ fosforylering og myogenese.

Resultatene fra denne studien kan potensielt bidra til å forstå transkriptomiske endringer i trening og faktorer som påvirker treningsrespons, noe som kan bidra til å designe bedre treningsprotokoller for sunn aldring og liv.

List of Papers

Paper I.

Khan Y, Hammarström D, Rønnestad BR, Ellefsen S, Ahmad R. Increased biological relevance of transcriptome analyses in human skeletal muscle using a model-specific pipeline. BMC Bioinformatics. 2020 Nov 30;21(1):548. doi: 10.1186/s12859-020-03866-y. PMID: 33256614; PMCID: PMC7708234.

Paper II.

Khan Y, Hammarström D, Ellefsen S, Ahmad R. Normalization of gene expression data revisited: the three viewpoints of the transcriptome in human skeletal muscle undergoing load-induced hypertrophy. 2021 (*Manuscript under Review*).

Paper III.

Mølmen KS, Hammarström D, Falch GS, Grundtvig M, Koll L, Hanestadhaugen M, **Khan Y**, Ahmad R, Malerbakken B, Rødølen TJ, Lien R, Rønnestad BR, Raastad T, Ellefsen S. Chronic obstructive pulmonary disease does not impair responses to resistance training. J Transl Med. 2021 Jul 6;19(1):292. doi: 10.1186/s12967-021-02969-1. PMID: 34229714; PMCID: PMC8261934.

Paper IV.

Mølmen KS, Hammarström D, Pedersen K, Lian Lie AC, Steile RB, Nygaard H, **Khan Y**, Hamarsland H, Koll L, Hanestadhaugen M, Eriksen AL, Grindaker E, Whist JE, Buck D, Ahmad R, Strand TA, Rønnestad BR, Ellefsen S. Vitamin D_3 supplementation does not enhance the effects of resistance training in older adults. J Cachexia Sarcopenia Muscle. 2021 Jun;12(3):599-628. doi: 10.1002/jcsm.12688. Epub 2021 Mar 31. PMID: 33788419; PMCID: PMC8200443.

Abbreviations

AMPK	activated protein Kinase
B2M	Beta-2-2microglobulin
BAM	binary alignment map format
Bp	base pair
BWT	Burrows-Wheeler transformation
CAGE	cap analysis of gene expression
COPD	chronic obstructive pulmonary disease
CSA	cross-sectional area
DNA	deoxy ribonucleic acid
e.g.	example gratia (for example)
ECM	extracellular Matrix
FDR	false discovery ratio
FEV	forced expiratory volume
FPKM	fragment Per thousand Million
FVC	forced vital capacity
GFM	graph ferragina marnzini
GOLD	Global Initiative for Chronic Obstructive Lung Disease
HGFM	hierarchical graph ferragina marnzini
HISAT2	hierarchical Indexing for Spliced Alignment of Transcripts 2
ICC	interclass correlation
IL6	interlukin 6
MiR	micro-RNA
MMP	maximal mappable prefix
MPSS	massively parallel signature sequencing
mRNA	messenger ribonucleic acid
mTOR	mammalian target of rapamycin
Myo	myology
ncRNA	non-coding ribonucleic acid
PCR	polymerase chain reaction
PGC-1α	peroxisome proliferator-activated receptor gamma coactivator 1-alpha
qPCR	quantitative polymerase chain reaction

RFT1	required fifty three 1
RNA	ribonucleic acid
RNA-seq	ribonucleic acid sequencing
ROS	reactive oxygen species
rRNA	ribosomal ribonucleic acid
RSEM	RNA-seq by expectation maximisation
RT	resistance training
SAGE	serial analysis of gene expression
SAM	sequence alignment map format
STAR	spliced Transcripts Alignment to a Reference
TC	transcript count
TF	transcription factor
TMM	trimmed Mean of Median
TNFα	tumor necrosis factor alpha
TPM	transcript Per Million
TTS	transcription termination sites
UQ	upper Quantile

1. Background of the study

Lifestyle-related diseases involving dysfunction of metabolic, cardiopulmonary, neuronal and immune systems are increasingly prevalent in the human population. The progression of these diseases tends to be associated with life-long malnutrition or a lack of exercise and leads to personal suffering and substantial socioeconomic costs [1]. A sedentary lifestyle and poor dietary decisions cause muscle mass loss and/or obesity. In order to maintain regular healthy muscle mass, some lifestyle modifications along with the right amount of exercise are required [2].

Skeletal muscle is the voluntary muscle, making up 40% of the body weight [3]. Healthy skeletal muscle mass is associated with longer life expectancy, a healthy life, and healthy ageing [4]. In the last couple of decades, training techniques have been optimised to help rehabilitate people living a sedentary lifestyle or with related diseases to lead a healthy life and have healthy ageing [5, 6]. Muscle mass can be affected by various factors, including training type, training volume, diet, sleep, supplementation of micro and macromolecules (vitamins and proteins) and physiological conditions [7-14]. The same training protocol could lead to a different response in different individuals, which suggests the need for a personalised rehabilitation programme. To design these protocols, we need biomarkers to determine health and training protocols for an individual. During the exercise condition, an individual could experience increased oxidative and inflammatory stress, altered transcription and coordinated stress response due to multiple molecular changes such as phosphocreatine and muscle glycogen breakdown, regulation by transcription factors and DNA hypomethylation [15]. With age, physical activity becomes limited, leading to muscle atrophy. Also, various chronic diseases like chronic obstructive pulmonary disease (COPD), cancer, arthritis, diabetes and obesity might impair physical activity and exacerbate muscle atrophy [16].

COPD is the third most prevalent cause of death worldwide [17] [18]. COPD causes irreversible loss of lung function induced by smoking, air pollution, and direct contact with certain chemicals [19]. Muscle loss and dysfunction during COPD is the compounded result of inflammatory cytokines, oxidative stress, growth and anabolic hormones, nutritional status, and physical activity [20, 21]. Exercise seems to be the effective rehabilitation approach with and without clinical drugs targeting *TNFα, IL-6* and reactive oxygen species (*ROS*) [22]. But the loss of lung function prevails fatigue, causing low motivation towards training. However, specific measures should be taken for a healthy lifestyle, like reducing contact with chemical

smoke and increasing lung function by training. Training is directly associated with muscle growth. Also, it has been stipulated by previous studies that vitamin D_3 affect muscle growth [23-26]. Vitamin D_3 indirectly regulates glucose uptake of skeletal muscle through calcium pathway cascade and elevated plasticity [27]. This suggests that vitamin D_3 supplementation could benefit rehabilitation programmes for the elderly, healthy ageing and the COPD-affected population.

To study muscle growth as a response to training volume, RNA sequencing (RNA-seq) can reveal the complete information about the presence and quantity of RNA in a biological sample at a given time [28]. The basic steps of RNA-seq data analysis include quality filtering, alignment, quantification, differential gene expression and gene ontology. However, the tools and algorithms for RNA-seq analysis keep evolving. Since the beginning of RNA-seq application, data analysis has been performed by different sets of tools with varying performance reports [29-32]. Each data set is different and needs careful evaluation of tools to decide which is best for a given project.

Additionally, normalisation of RNA-seq is a critical, decisive step, which adjusts the data according to several factors (e.g., the run time difference between the samples, position of sample on the plate, sequencing plate usage, handling of samples), which prevents biases in comparison of quantified expression. Therefore, it is critical to compare different perspectives to normalise data and remove technical bias from the data, which can affect downstream analyses [33]. Different normalisation methods are associated with a different set of stable genes and differentially expressed (DE) genes. In addition, it is often ambiguous whether the selected stable genes are stable in expression or not [34]. Therefore, careful consideration is needed to decide on a standardised RNA-seq pipeline. This pipeline will be used to observe molecular changes caused by different training volumes, physiological conditions, and supplementation.

2. Introduction – Biological Background

2.1.Muscle Tissue

Muscles are contractile tissue found in animals, which function to produce motion and are a source of power and as a reservoir for protein. There are two types of muscles, i.e., striated and non-striated. Striated muscles are defined as having visible banding within muscle cells to ensure the constant direction of tension, whereas non-striated muscles lack this banding and are spindle-shaped.

Striated muscles are also categorised into two types: skeletal muscle and cardiac muscle (Table 1). *Skeletal muscle* is the largest organ in the body [3] and contracts only when there is a signal from the somatic motor neuron. Skeletal muscle is connected to bone by a flexible joint called *flexor*/*tendons* to facilitate physical activity. *Cardiac muscle*, also known as the myocardial, comprises two types of contractile cells (99% of atria and ventricles) and generates spontaneous action and myocardial conducting cells make a conduction system of the heart. This tissue account for the unique ability of the heart to contract without any outside signal, instead of from auto-rhythmic cells/pacemaker cells. Cardiac muscle can be further differentiated from skeletal muscle by the presence of intercalated discs, which control the synchronised contraction of cardiac tissues. *Non-striated or smooth muscles* are mainly found in the digestive system, urinary bladder and respiratory system. They play an essential role in regulating the flow of material in and out of the body, for example, in aiding the movement of food through the digestive system via peristalsis.

2.2.Skeletal Muscle

Skeletal muscle is a set of innervated, non-voluntary muscle cells. It exhibits fatigue with high-energy requirements during high physical activity. The role of skeletal muscles includes maintaining posture and all physical activity, maintaining basal energy levels, bone density, helping in recovery during disease, and acting as a reservoir for amino acids and carbohydrates for other organs and longevity [3, 35, 36]. Skeletal muscles consist of several long cylindrical striated muscle fibres (Figure 1A), which are made up of thousands of myofibrils with hundreds of nuclei arranged at their longer axis (Figure 1C). It contains the following: contractile and elastic proteins - *myosin* and *actin*; regulatory proteins - *tropomyosin* and *troponin*; accessory proteins - *titin* and *nebulin* [3]. Every skeletal muscle is constrained by connective tissue –

epimysium and each bundle of fibre are surrounded by perimysium (Figure 1B); within this, every muscle fibre is surrounded by sarcolemma.

Table 1. *Comparison of the different muscle tissue type's characteristics*.

	Smooth Muscle	**Cardiac Muscle**	**Skeletal Muscle**
Appearance	Non-striated	Striated	Striated
Function	Maintain the flow of fluids and solids along hollow structures	Pumping blood through the heart chambers into blood vessels	Body posture and locomotion, joint stability and heat generation
Location	iris, erectors of hairs, walls of hollow organs	Heart	Directly attached to the skeleton by tendons
Cell shape	Short and fusiform	Branching	Long and cylindrical
Nucleus	Single nucleus, centrally located	Single or bi-nucleated centrally located	Multinucleated periphery located under the sarcolemma
Regulation of contraction	Involuntary, autonomic motor neurons, hormones, cytokines	Involuntary, autonomic motor neurons	Voluntary, somatic motor neurons
Stimulus effect	Excitation or inhibition	Excitation or inhibition	Excitation

The ability of the skeletal muscle to change its structural and functional properties to adapt to imposed conditions is known as *plasticity*. Muscle plasticity adapts for continuous mechanical stress imposed on the body and leads to muscle gain known as *muscle hypertrophy.* Disuse of muscle for a more extended time can lead to loss of muscle mass, which is called *muscle atrophy*.

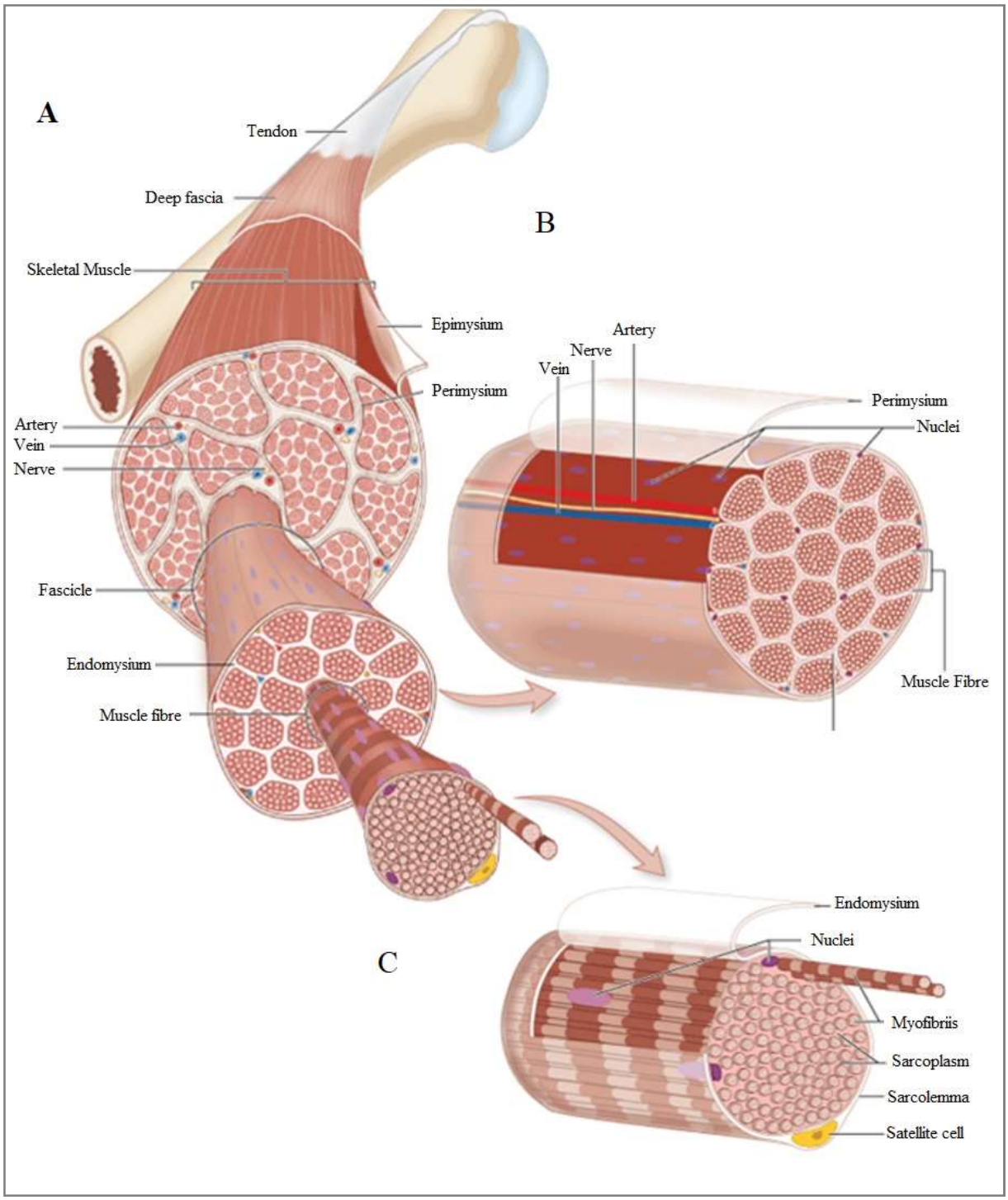


Figure 1. *A) Skeletal muscle is connected to the bones via flexor; each skeletal muscle tissue is covered by epimysium. B) Muscle fibre has nerve, artery and vein to transport glucose and proteins, and a nerve to control the movement of muscle C) Each subunit of muscle fibre is made up of several myofiber bundles. Each bundle is enveloped by endomysium*. (Adapted from Human Physiology, Silverthorn DU, Copyright © Pearson Education, Inc. (reproduced with permission from Pearson Education Inc.) [3].

2.2.1. Skeletal Muscle contraction proteins

Myosin is a motor protein, which helps in the ability to create movement. Different isoforms of myosin are present in different muscle types; myosin decides contraction speed [37]. Myosin structures consist of the long tail and two tadpoles-like heads clustered together to make a thick filament.

Actin is a family of globular multi-functional proteins that form microfilaments in the cytoskeleton and the thin filaments in muscle fibrils. Multiple globular proteins (G-actin) linearly polymerise to make filament proteins (F-actin) and further polymerising into double-stranded beads to make thin filament with troponin, nebulin, and tropomyosin. Under the light microscope, the thick and thin filament has a repeating pattern of alternating dark and light bands (Figure 2A). Most of the thin and thick filament of myofibril are connected by myosin cross-bridges. (Figure 2B). Cross-bridges connect both the thin and thick filament to make sarcomere. Each G-actin has a myosin-binding site and an ATP binding site. Muscle contraction is the process of sliding of actin and myosin filaments. It is initiated by calcium signalling. Ca^{2+} binds to troponin and exposes myosin-binding sites on thick filaments, and the myosin head then binds to the active site of actin. This binding or cross-bridge is broken by adenosine triphosphate (ATP), as ATP hydrolysis changes the orientation of myosin heads. The myosin head again binds to actin and releases adenosine diphosphate (ADP). Myosin repositioning pushes thin filaments to the centre of the sarcomere, causing muscle contraction by transmitting both longitudinal and radial forces, which reinforce the dystrophin-glycoprotein complex and the integrin complex provide mechanical linkage between myofilament and Extracellular Matrix (ECM) [38, 39].

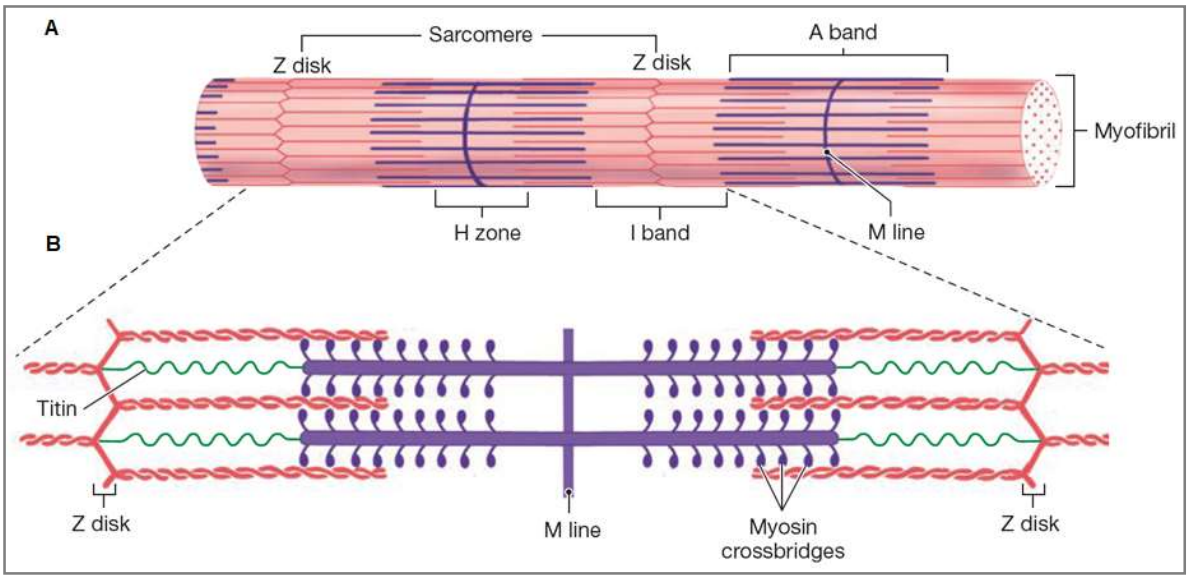


Figure 2. *A) A repeating pattern of thin and thick filament making up the sarcomere can be seen under the microscope. A zigzag pattern where two sarcomeres are connected is called the Z-disk. The region where the Z-disk is present is called the I-band, and in this region, there is no thin filament present. Similarly, the region where the thin filament is present with and without thick filament overlapping is called A-band. If a thick filament is absent, that region is called the H-zone. The linear line where thin filaments are connected to each other is called the M-line. B) Thin filament joined to each other through the M-line and made up of myosin, slides between actin proteins facilitated by titin protein.* (Adapted from Human Physiology, Silverthorn DU, Copyright © Pearson Education, Inc. (reproduced with permission from Pearson Education Inc.) [3].

2.2.2. Skeletal muscle Plasticity and Hypertrophy

The ability of muscle tissue to modify its structure, function and metabolism in response to mechanical stress and numerous physiological changes is known as plasticity [40]. Repetitive resistance training acts as external stimuli or mechanical stress, activating adaptation and increasing muscle mass and strength with endurance or contractile speed. Muscle growth refers to an increase in muscle mass called hypertrophy. Training exercises act as stimuli for an increase in muscle volume. This adaptation takes time and depends upon the fitness goals of the individual. There are many factors affecting muscle hypertrophy, i.e. training type [13], training volume [10, 12, 41], sleep [42, 43], and supplements including macronutrients [9], micronutrients and minerals [14]

- ***Training type***

Endurance exercise signals transcriptional upregulation of many genes involved in the response of structure genes enabling the muscle to function at a higher level of metabolic activity. The Ca^{2+} signalling cascade and high skeletal muscle energy status activate *AMPK* (Activated protein Kinase) by glycogen content. *ROS*/redox signalling sense hypoxia muscle. Exercise-induced epinephrine initiation of angiogenesis takes place through *cAMP* pathways. *PGC-1α* activates multiple transcription factors and nuclear receptors and recruits chromatin remodelling complexes that facilitate transcription and post-transcriptional processes through splice proteins. *PGC-1α* expression is proportional to muscle tissue growth and also activates its antagonist*, RIP140*, a transcriptional corepressor that targets similar transcription factors and nuclear receptors such as *PGC-1*, but oppositely. On the other hand, resistance training enhanced translation. Mechanical stress activates *AKT*-dependent and *AKT*-independent pathways. This modulates *mTOR*, which controls most of the signalling cascade and increases translation and elongation. *mTOR* is negatively regulated by low energy metabolism and myostatin. Significant growth in muscle is not facilitated by stem cell type cell "satellite cells" that present themself in the basement membrane of muscle fibre, whose proliferation and differentiation are led by Notch and *WNT* signalling and *MyoG, MyoD* and *Myf3* regulatory factors. [40]

- ***Training volume***

The total number of lifts and sets per exercise is called *training volume*. High training volume might lead either to high muscle mass growth or to muscle damage [44]. In studies on health and performance improvement studies, inter and intraindividual variability can vary the results. However, to reduce or remove the interindividual effects, studies need to consider it in experimental design [45]. *No previous studies have reported molecular changes at the wholistic level to explain the effect of training volume while accounting for the intraindividual effect.*

- ***Supplementation***

Skeletal muscle hypertrophy is the balance between muscle protein synthesis and breakdown. For good muscle mass growth and maintenance, there is a need for an adequate amount of nutrition. Several supplements are proposed for good or high muscle mass gain and maintenance. Usage of these supplements has grown by tenfold a decade [46]. Different goals need different dietary proportions and a different supplementation mix. Research on Vitamin D_3 has recently received attention from several focus areas. It has been proposed to play a

prominent role in cancers, cardiovascular diseases, diabetes, neurological and psychiatric disorders and autoimmune diseases [47, 48]. Halfon et al. reported that calcium and phosphorus supplementation helps the frail elderly and positively affects muscle growth [49]. However, another study has noted otherwise [23], which throws doubt on the role of vitamin D_3 in muscle hypertrophy.

- ***Loss of skeletal muscle - Muscle atrophy***

Muscle loss or wastage is known as atrophy. It is caused by a sedentary lifestyle, low muscle movement, ageing, malnutrition, alcohol, burns, and diseases, leading to early fatigue and demotivation to perform physical activity. However, these can be treated by medicine and/or rehabilitation programmes. Due to all these reasons, COPD affects the population adversely and leads to mortality in the elderly. *The effects of COPD cannot be reversed but can be stopped by rehabilitation programmes by increasing lung capacity, which is directly related to training and muscle mass.*

- ***Chronic Obstructive Pulmonary Disease (COPD)***

Emphysema and chronic bronchitis are covered under the term COPD. COPD is a chronic lung disease that causes progressive damage to airways and lung parenchyma, causing airflow obstruction. It is estimated that worldwide, 210 million people suffer from COPD, and it was the third leading cause of death (3.23 million) in 2019 [18, 19]. Exposure to cigarette smoking, biomass smoke and chemicals [50-52] contributes to the destruction of parenchymal, which leads to airway lumen narrowing through smooth muscle hypertrophy, inflammation and elastic recoil loss. This results in a loss of alveolar surface area, ventilation-perfusion mismatch, airflow limitation and gas trapping. All these complications contribute to dyspnoea (difficult or laboured breathing) and fatigue. The pathological mechanism of COPD is still not well understood. But it is evident that alteration of the protease or antiprotease balance is mediated by an increase in neutrophils and macrophages, increased oxidative stress, autoimmune dysfunction and dysfunction of lung development and maintenance pathways like the retinoic acid pathway and hedgehog signalling pathway [53-56]. *However, not all smokers get COPD, so genetics plays an important role too, which needs to be better understood.*

COPD causes skeletal muscle weakness, and with age, it becomes more severe, causing mortality in one-quarter of the affected population [57-59]. Weakness is usually caused by the limitation of exercise and physical movement due to fatigue, systemic inflammation, oxidative stress, genetic susceptibility and loss of motivation.

Several studies have found an association between vitamin D_3 and the measure of muscle strength [60, 61], which suggests that vitamin D_3 supplementation could positively affect the rehabilitation of the COPD population [57-59] [57]. In contrast, a few studies show no effect of vitamin D_3 on the COPD diagnosed population [25, 26]. *This indicates that there is a contrasting view of the usefulness of vitamin D_3 supplementation on the COPD population. Moreover, there is a need for a detailed molecular study to understand this issue further.*

2.3.Transcription

All organisms share the foundational mechanism of life: "the central dogma" (Figure 3), which describes the central component of life and how they are related [62]. The first component of the central dogma is deoxyribonucleic acid (DNA). DNA stores the information for an organism to develop, grow, function and reproduce. DNA is transcribed into ribonucleic acid (RNA) by a process called *transcription*. RNA was only known as a transitional step containing protein information until recent years, but RNA itself has many essential functions, such as gene regulation [63]. The stretch of DNA, transcribed into an RNA molecule encoding a protein, is called messenger RNA (mRNA). Other stretches of DNA transcribed into RNA molecules are called non-coding RNAs, including transfer RNA (tRNA), small nucleolar RNA, ribozymes, ribosomal RNA (rRNA), long non-coding RNA, etc.

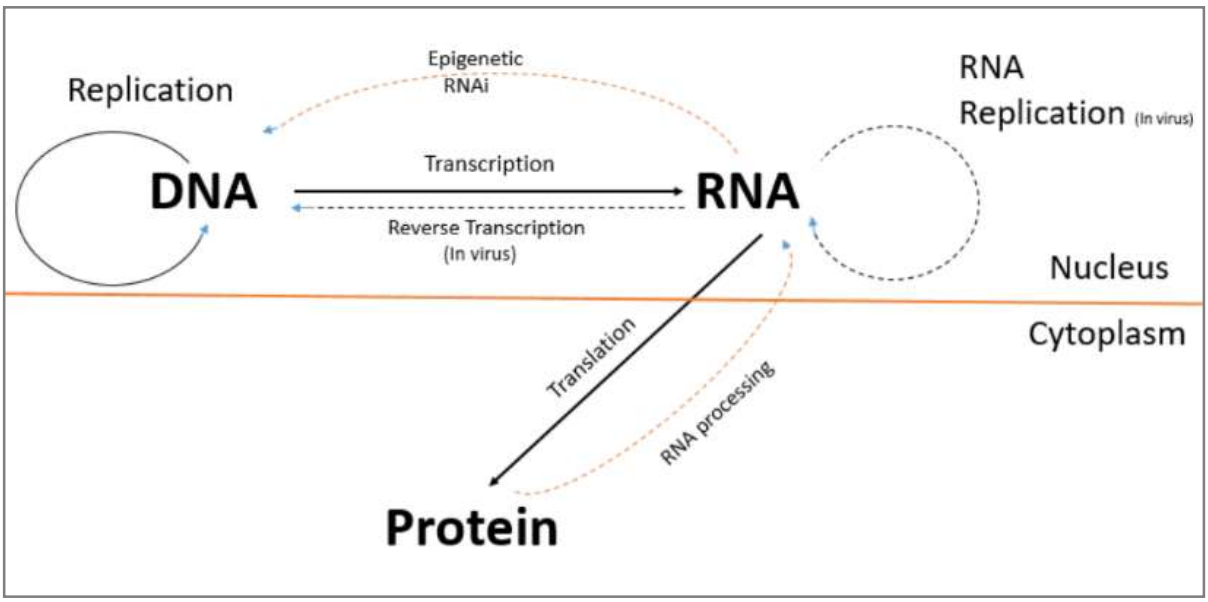


Figure 3. *The central dogma of molecular biology. It is the explanation of the flow of genetic information within a biological system. In this DNA makes RNA, and RNA makes Proteins facilitated by different enzymes.*

The process by which mRNA information is used to make protein is known as *translation*. Proteins act as active components and facilitate the function of an organism. The abundance of mRNA is directly proportional to protein abundance, which leads to different phenotypic and physiological results/outcomes. In other words, transcription of DNA is the key regulatory step in managing functional molecules [64]. Commencement of transcription is mainly regulated by sequence-specific transcription factors (TFs). The promoter region has a transcription start site (TSS) possessing many transcription factor binding sites (TFBS) at the enhancers region, which are present far away from the gene, and are what regulate the transcription process [65].

Furthermore, DNA is divided into multiple loci known as exons, separated by introns [66]. These exons accumulate mRNA by a splicing event or an alternative splicing event, forming different products/mRNAs from a single gene [67]. The final step towards mRNA formation is terminating the transcription, facilitated by transcription termination sites (TTS), a conserved sequence motif in emerging RNA sequences [68]. These mRNA sequences are translated into protein resulting in phenotypic changes. High throughput approaches are used to study overall gene expression at a condition and a given time.

2.4. Training induced changes in expression profile

A single bout of exercise can alter gene expression in both exercised and non-exercised muscles, promoting upregulation of gene expression. This expression profile at different time points can provide a clear picture of molecular changes [69]. Previously, when high-throughput technology was not available, it was challenging to calculate these changes. The only method was PCR and multiplex PCR, which is only for one gene or a set of genes at any one time. With advances in technology including RNA-seq, and microarray, it has become possible to study the expression profile of all the transcripts in one go at a given time point.

3. Introduction – Technological Background

3.1. High-throughput transcriptomic/whole transcriptome technologies

Transcriptome analysis is the study of the transcriptome, which is the complete set of RNA transcripts produced by the genome, under specific circumstances or in a particular cell, using high-throughput methods. Although many methodologies have been developed for transcriptome profiling, two of these methods have revolutionised gene expression profiling, i.e., microarray and RNA-seq technology.

3.1.1. Microarray

Since the 1990s, microarray has been the choice for genome-wide transcriptomic studies. This method enables the simultaneous examination of thousands of transcripts. In this technology, a large set of short (typically 20 to 70 bp long) single-stranded DNA probes are attached to a fixed plate/sheet, complementary to cDNA sequences, followed by labelling it with fluorescent dye. Fluorescence from each probe cDNA complex is then observed by measuring fluorescence with a fluorescent scanner [70]. However, with the emergence of the high throughput next-generation sequencing technology RNA-seq, the usage of microarray has become limited. In microarray, probes need to be complementary to the target sequence, and therefore these have to be defined in advance. In contrast, RNA-seq does not require any prior knowledge about nucleotide sequence. Microarray is a good choice if one is interested in analysing known targets due to its low cost, cheaper data storage, and lower complexity than RNA-seq [71]. RNA-seq, on the other hand, brings an advantage when high throughput data are needed, and unknown targets are considered.

3.1.2. RNA-seq

Sanger sequencing was the very first whole-genome sequencing technology. It is expensive and has low throughput as each reaction gives out only one sequence. To overcome this limitation, a tag-based method was developed, which is used in cap analysis of gene expression (CAGE), serial analysis of gene expression (SAGE) and massively parallel signature sequencing (MPSS) [72-74]. These technologies provide less expression sensitivity than microarrays, and tags can be mapped to several places on the reference genome. High throughput sequencing overcame these limitations.

RNA-seq is a high throughput sequence-based transcriptome analysis method to survey, characterise and quantify the transcriptome at the whole genome level [28]. In contrast to

previous methods, RNA-seq utilises sequencing by synthesis technology to define the nucleotide sequences and to quantify RNA molecules in a sample. RNA-seq can process this data in hours to days with high fidelity, making it the preferred technique for RNA analysis amongst many researchers. Currently, RNA-seq from Illumina is the leading technology for performing high throughput sequencing. RNA sequencing can identify differential expression of genes (DEGs), mutated genes and gene isoforms in disease states.

A typical RNA-seq experiment (Figure 4) starts with isolating the whole RNA and purifying subset RNA from total RNA. The majority of RNA species consist of rRNA. The mRNAs are enriched either by selecting poly-A molecules using poly-T oligo attached magnetic beads or by the rRNA depletion method [75-77]. The enriched mRNA is then fragmented via RNA hydrolysis or nebulization, or it can be performed after cDNA conversion by DNase I treatment or sonication (100-300 bp). The cleaved mRNA fragments are then converted into cDNA using reverse transcriptase and random primers. Following this adaptor, sequences are ligated using ligase enzymes. Adenine bases are added to the 3' end to prepare cDNA for ligation to adaptors, containing a single T base overhang at the 3' end. Later, cDNA molecules are split into different sizes. The desired range of cDNA length is purified using the gel extraction method to select molecules of a similar size, which removes un-ligated adaptors and dimers of adaptors. The process is followed by the amplification of collected DNA fragments by PCR. Double-stranded cDNA is then denatured in single-stranded molecules. These single-stranded molecules are hybridised with forward and reverse primers (adapter complimenting primers) fixed on a sequencing plate. After hybridisation, fragments loop over to the hybridised sequence to the adjacent primer, and DNA polymerase copies the template forming double-stranded DNA bridges, which are denatured to obtain a single-stranded template population. Then, reverse strands are removed by base-specific cleavage, and the 3' end of the forward strand is blocked to prevent interference in the sequencing process. This process is repeated several times until dense clusters containing at least 1000 fragments per cluster are procured/achieved. The final step is sequencing by synthesis, where clusters are amplified in rounds.

In each round, fluorescence-labelled nucleotides are added. Separate dyes are used for dATPs, dCTPs, dTTPs and dGTPs. Additionally, these nucleotides are blocked on the 3'-end, so amplification is stopped after their incorporation into a growing strand. Following this,

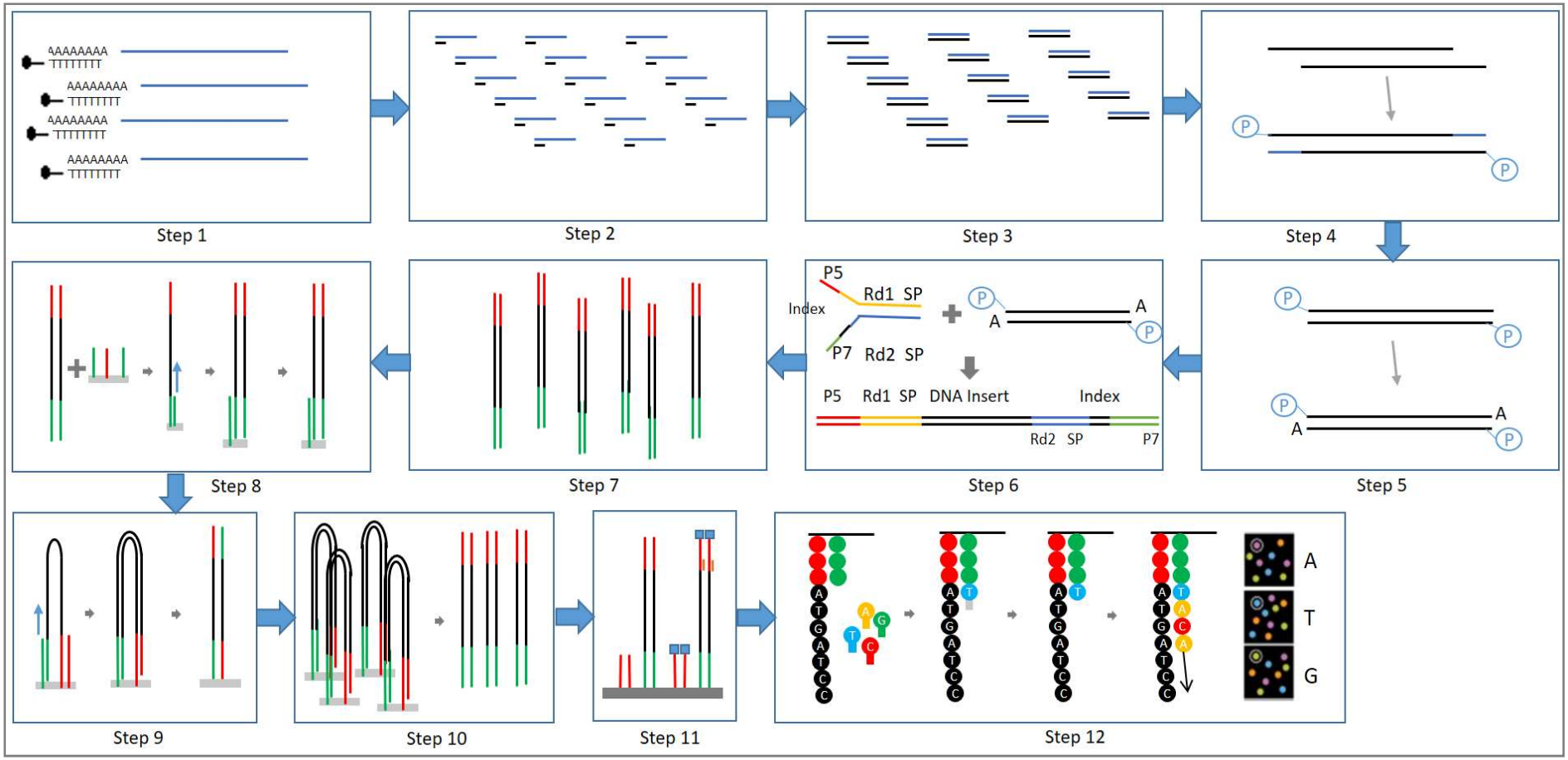

Step 1
Step 2
Step 3
Step 4
Step 5
Step 6
Step 7
Step 8
Step 9
Step 10
Step 11
Step 12
P5
Rd1 SP
Index
P7
Rd2 SP
P5 Rd1 SP DNA Insert Index
Rd2 SP P7
A
T
G

Figure 4. *The first step in RNA sequencing is to select suitable mRNA molecules from the total RNA by poly-A tail, and then it is fragmented into the required length. The RNA of interest is then filtered out by gel electrophoresis. Then it is converted into cDNA by reverse transcriptase in step 2 (primer annealing) and step 3 (elongation). cDNA has blunt ends, so A base is added to it at step 4 and step 5. At step 6, the ends of the cDNA chain are blunted, and A base is added to it. To this product, the adaptor sequence is hybridised with inserted A base. The adaptor-ligated product is purified and amplified using PCR at step 7. Step 8 is to hybridise it to a fixed plate with a complementary adaptor sequence. These strands are amplified using bridge amplification at step 9 and are clustered by repeating this process at step 10. The reverse strand is cleaved and washed away the sequencing primers (complementary to the adaptor sequence attached to the plate) at step 11. This is followed by the sequencing-by-synthesis process at step 12. After the hybridisation of the sequencing primers, four labelled reversible terminators and DNA polymerase are added; one nucleotide is added to the nucleic acid chain, and after laser excitation, the colour of its fluorescent label is observed and recorded. The fluorescent label and the blocking group are then washed away, and the described process is repeated until the length of interest is achieved.* Adapted from Illumina.com [81].

unincorporated nucleotides are washed, and the sequencing plate is excited by a four-channel laser to record to which clusters which nucleotides were added. Next, fluorescence dyes and blocking groups are removed, and the next sequencing round starts. This process is repeated 200-300 times (for small RNAs a lesser number of times). The output of this step is a sequencing image; one each has information of a newly incorporated nucleotide, combining it to form sequences, and each base has its quality score based on the fluorescence levels. This information is saved as text using the fastq file format, which contains a unique ID for each sequence and quality score for each base.

3.2.RNA-seq application in sports and training

Advances in high throughput sequencing technology, algorithms, machine learning and data science have enabled researchers to explore genomic transposition with exercise [78]. This has led to important discoveries related to sports and health genomics. Progression in this field promises to deliver more precise, personalised health care and rehabilitation programmes to benefit the population [79]. However, this progression can only be expected by more coordinated research. As such, meta-analysis of several studies reported by the metamex.eu database has revealed that most studies have statistical ascendency, are oblique towards male

volunteers and have no contemplation towards individual effect. [80]. *Exercise and sports genomics experiments need to be innovative to accommodate inter and intraindividual effects while protecting privacy and avoiding misuse of genomic information. These experiments have the potential to contribute to designing rehabilitation programmes for improving health and for healthy ageing*

3.3.RNA-seq data analysis

3.3.1. RNA isolation and sequencing

RNA isolation can be performed using standard protocol or by a commercial kit. The quality of these samples needs to be observed to select only high-quality samples for sequencing (RNA integrity score RQI >7 is preferred). Sequencing libraries are prepared for each sample (1000 ng RNA) using TruSeq Stranded Total RNA Library Prep (Illumina, San Diego, CA USA). And sequencing is performed using any Illumina Sequences, e.g., HiSeq. The files retrieved from the sequencer are called raw data, and sequences in it are called reads. Although Illumina has high-quality reads, it still needs to be preprocessed followed by basic steps for RNA-seq analysis.

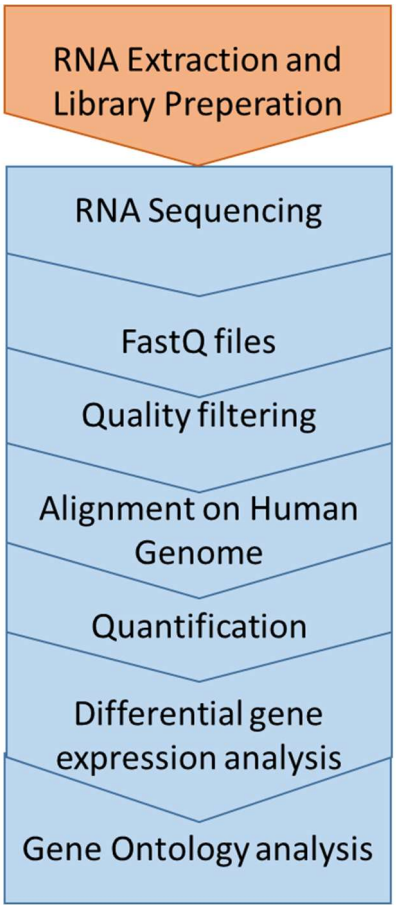


Figure 5. *Overview of RNA-seq data analysis for detecting differential gene expression.*

3.3.2. Quality filtering

The NGS platform provides data in fastq format; it is similar to the fasta format with two additional lines reporting the header or the id of the sequence and starting with the + sign. The third line is left blank in the new platform, and the fourth line represents base quality score values using an ASCII character, one for each base. The quality of sequence is shown by the Phred score, which is the probability of the base being wrong (Q = -10log10P). P is the probability of a nucleotide being incorrectly identified. For example: if Q = 10, then P = 0.1 and estimate correctness is 90%; if Q = 20, then P = 0.01 and estimated correctness is 99%; and if Q = 30, then P = 0.001 and estimated correctness is 99.9%.

Sequencing bias can be caused by calling the wrong base due to poor sequencing quality during the sequencing reaction. In addition, when fragments are smaller than the sequencing length, adaptor sequence contamination can adversely affect downstream analysis [82]. Before downstream analysis, quality control of read sequences is important to remove short and/or low-quality reads/bases and trim out the adaptor sequence. The two most cited tools for these steps are Trim Galore (https://github.com/FelixKrueger/TrimGalore) and trimmomatic [83]. Trim Galore is a wrapper tool for Cutadapt [84] and FastQC [85]. Cutadapt is an adaptor and quality filtering tool, and FastQC calculates and visualises the quality of reads in each fastq file. It uses the first 13 bases of the adaptor sequence for its identification. Users can define a Phred score cut off where any base below the cutoff would be trimmed. The default Phred score threshold is 20. Cutadapt calculates the Phred score for each base, subtracts the default score from each base score and then partially sums all positive values and keeps the negative score as it is. Next, it removes all the bases with a negative Phred score. Cutadapt also removes the sequences, which are smaller than a pre-defined cut-off after filtering. The pre-defined cut off for sequence length is 20 bp.

Trimmomatic uses an alignment-based score and sliding window methods to identify/remove the adaptor sequence and for quality filtering, respectively. In the sliding window method, the average of the base quality for a specific length is taken. If the average is less than the specified threshold (default is 25), it removes the afterwards bases. After that, it selects the sequences on maximal information quality filtering, which is based upon three rules, i.e., length threshold, sequence coverage and error rate. Then these filtered out files were observed for overall quality using FastQC. FastQC produces a report file for each file separately. To collectively tabulate and visualise the quality reports, MultiQC [86] is used, which provides a summary of all files as graphs and tables.

3.3.3. Alignment and quantification

After removing low-quality reads and trimming the adaptor sequences, high-quality reads are aligned on the reference genome (e.g., human genome). The main goal of alignment is to find loci of each read on the reference genome and quantify their expression (i.e., counting the number of reads that align to each gene/transcript [read count]). In recent years, alignment tools have evolved, and reads can be aligned directly to transcripts, and it does not need to be aligned to the genome. Hence, alignment tools are divided into two main groups, i.e. genome-based aligners, e.g. STAR [87] and HISAT2 [88, 89] and transcript based aligners. Transcript based aligners are further divided into two groups: alignment-based, e.g. RSEM [90], and pseudo-aligners, e.g. Kallisto [91] and Salmon [92]. STAR and HISAT2 need a quantification tool like htseq2 [93], while Kallisto and Salmon tools have their own quantification algorithm. Teng et al. suggested using standard deviation between the samples within the group as a measure of comparing better alignment tools, as it is expected to have a similar expression in samples in similar conditions [94].

STAR: Spliced Transcripts Alignment to a Reference (STAR) was designed for non-contiguous sequences. The alignment process using STAR consists of two steps. First is seed search, where Maximal Mappable Prefix (MMP), i.e. the longest matching sequence loci/s, is identified on the reference genome, and this MMP becomes seed1. Then the algorithm will search for an unaligned portion of the sequence to find the next longest match on reference genomes, which is seed2. If it does not find the exact match due to a mismatch or indel, MMPs will be extended (Figure 6). Second, this is followed by clustering, stitching and scoring. In this step, seeds are clustered together by similarity to selected anchor seeds. Then the seeds are stitched together based on alignment score [87].

HISAT2: Hierarchical Indexing for Spliced Alignment of Transcripts 2 (HISAT2) uses a graph-based approach. It uses Burrows-Wheeler transformation (BWT) (Siren et al. 2014) to generate a new Graph Ferragina Marnzini index (GFM). A large set of small GFM indexes collectively covers a whole genome. These small indexes combine, which enables rapid alignment. This strategy is called the Hierarchical Graph Ferragina Marzini index (HGFM) and is used to index the whole genome[89]

HTseq: STAR and HISAT2 provide SAM/BAM files as output, which need to be quantified. HTseq uses GFF (General Feature Format) or GTF (General Transfer Format) files to assign aligned reads to exon or genes [95]. It only counts uniquely assigned reads as count. The tools listed below have their own quantification algorithm.

RSEM: RNA-seq by expectation maximisation (RSEM) consist of two steps. The first is to generate a transcript index, which is created by using the transcript sequences provided. The second step is to align reads on this index to estimate the abundance and their credibility interval. It is based on collecting all alignments and not just the best alignment. It allows two mismatches in the seed sequence of input read, usually the first 25 bases, and then it collects all loci where the sequence is aligned. After alignment, RSEM computes maximum likelihood abundance estimate using an expectation maximisation algorithm [90, 96]. A different aligner generates RSEM process alignment files. Bowtie2 is suitable for both short read alignment and produces multiple alignments with a user-defined gap and mismatches.

Bowtie2: It makes the FM index of reference genome or transcript to align reads. It uses the vectorised Smith-Waterman or Needleman-Wunsch algorithm, which enable it to perform alignment with mismatches and gaps [97].

Kallisto: It is based on the unique idea of pseudo-alignment to determine the match between reads and the transcript [91]. In this approach, an index is created by the transcript sequence provided, making a transcriptome De Brujin graph of k-mers. K-compatibility class lookup is done for each fragment if it is a paired-end read. To quantify the pseudo-alignment, Kallisto uses the following likelihood function (Equation 1). In this, $y_{f,t}$ is the compatibility matrix, f is the set of fragments, t is the set of transcripts, L_t is the effective length, and α is the probability of selecting fragments from the transcript:

$$L(\alpha) \propto \prod \sum y_{f,t} \frac{\alpha_t}{i} = \prod \left(\sum \frac{\alpha_t}{i} \right)^{c_e} \quad \text{(Equation 1)}$$

Salmon: Salmon uses a quasi-alignment approach to quantify reads aligned to the transcript index having a hash table of k-mers. In the quasi-alignment approach, the read is scanned from left to right until matching k-mer is identified in the hash table. Then all suffixes available in the SA matrix are retrieved, and the maximum matching prefix is identified using the next k-mer until the end of the read is reached. The final set is determined by consensus formation by all MMPs in that read. In Salmon, an expectation maximisation model is used to estimate abundance.

3.3.4. Normalisation

The quantified result from the alignment tools provides raw sequences, which need to be normalised for technical and biological bias, reducing redundancy. There are three sources of technical bias: i) normalisation for library size, i.e. TC (Transcript count), Median, UQ (Upper Quantile) and TMM (Trimmed Mean of Median); ii) normalisation for gene length, i.e. TPM (Transcript Per Million) and FPKM (Fragment Per thousand Million) and iii) normalisation for

known and unknown artefacts, i.e. housekeeping/stable genes and spike-in [96, 98-102]. Additionally, it has been shown that data normalisation accounting for the number of cells leads to more accurate analysis (tissue-offset) [33]. Furthermore, a spike-in sequence can also be added during RNA isolation and is subject to the stable expression [103, 104]. Below is a brief of the most commonly used normalisation methods:

TC (Total count): the number of transcript counts remained after removing transcript with zero counts.

Library size: the total number of reads aligned on the reference genome.

UQ (Upper Quantile): Upper quantile is a scaling factor calculated as the ratio of counts of each sample divided by mean 75^{th} percentile in all samples. If it is done for the 50^{th} percentile, it is known as the median.

TMM (Trimmed Mean on M-values): TMM normalisation is based on weighted trimmed mean log gene expression ratios. It is used by edgeR.

TPM (Transcript Per Million): In TPM normalisation, the raw count is the length of a gene in kilobase and then scaled by per million scaling factor.

FPKM (Fragment Per Thousand Million): In FPKM, the raw count is normalised for sequencing depth (per million reads) and then the gene length (in kilobases).

RLE (Relative Log Expression): In this, the median library is calculated from the geometric mean of all columns, and the median ratio of each sample to the median library is taken as the scale factor. It is used by edgeR and DESeq2.

Tissueoffset: It is assumed that each cell produces the same amount of RNA [101], which is not necessarily true [105]. For a biologically relevant count, the raw count needs to be scaled by tissue weight used in RNA isolation [33].

Spike-in normalisation

RNA spike-in is RNA with a known reference sequence, acting as a negative control for the experiment used in a hybridisation assay. It is added to the sample according to the proportion of the total RNA before sequencing library preparation. The sequence and quantity of spike-in are known for the sample. It was initially made to normalize qPCR data, but later it was used in RNA-seq too. In this approach, the read count of the spike-in sequence is observed, and raw counts are scaled accordingly.

3.3.5. Stable transcripts

Reference transcripts or housekeeping transcripts are transcripts that express stably irrespective of condition and tissue. They are also used to normalise data. In some studies, it has been

observed that the expressions of these transcripts are not stable [106-109]. Therefore, it is not viable to call them housekeeping transcripts or reference transcripts, but they can be called stable genes for a specific condition. Little is known about stable genes/transcripts for skeletal muscle in training conditions, which thus need to be investigated. It has been suggested that both B2M and GAPDH can be considered as housekeeping genes by qPCR in resistance training based skeletal muscle. There are different ways of identifying stable transcripts. The classical way to identify stable genes is based on differential gene expression analysis and identification of genes/transcripts with less than 20-40% change in expression with significant p-value and gene ontology [110]. However, this method cannot provide a constant answer if the normalisation approach is changed. Another method is based upon CV value (ratio of mean and standard deviation), calculated from different count files generated from different normalisation approaches. Before that, low and very high expressing transcripts are removed, and transcripts with a fold change of more than 0.2 are filtered out. With the emergence of new statistical algorithms, it is understood that this selection does not account for systemic effect across groups and interclass correlation across samples [33, 111].

3.3.6. Differential gene expression

Genes having altered expression between conditions are known as differential expressed genes. To determine these genes, normalised gene counts are arranged into an expression matrix, and statistical analysis is performed to discover significant differences between expression levels in experimental groups. The significance of DE genes is calculated by p-values which are then corrected for multiple testing using, e.g. FDR. DESeq2 [98] and edgeR [99, 101] are the two most cited tools for differential gene expression. Both approaches are based on the negative binomial approach. These methods are based on sample size calculation and power estimation and have been developed for experiments with independent measurements. RNA-seq experiments that have a complex design with repeated measures need different approaches in identifying DE genes, which can include. Normal pairwise comparison for complex design cannot be relied upon, as they do not consider data arising from such experiments should be analysed by generalised linear mix model approach, which allows for random effect [112, 113]. The selection of DE genes is done by log-fold changes (logFC > 1) and significant value (Adjusted p-value < 0.05).

Few studies are performing the comparison of different tools [29-32]. Each study reports different tools as outperforming others, which indicates that each dataset requires a specific pipeline. An absence of an RNA-seq pipeline for a skeletal muscle dataset motivated

this work. In this work, we aimed to standardize a data analysis pipeline for the correct analysis of the data.

3.3.7. Gene ontology (GO) analysis

The biological domain and knowledge relating to gene and gene product attribute, functions, assimilations and disseminations are annotated into ontologies provided by numbers or IDs. It covers three domains: cellular component, molecular function and biological process [114]. It is vital to know which functions or pathways are changed in differential expression analysis between the condition(s). It provides further insight into genes expression change with the condition and provides associated pathways, co-regulating genes and signals. Therefore, if anyone wants to design a drug target or identify genes affected in a pathway and want to segregate on that basis, it is possible by using this method.

4. Materials and Methods

Only a few transcriptomic studies related to health and sports [80, 115, 116]. This study is undertaken to understand the training responses on muscle growth of low and moderate training volume, vitamin D_3 supplementation and COPD diagnosis. To achieve this, two datasets were used.

4.1. Datasets

4.1.1. Resistance training in young adults (YOUNG dataset)

The YOUNG dataset comprised 16 male and 18 female non-smoking participants between 18 - 40 years enrolled in the study. Participants were excluded who had an intolerance to local anaesthetic, a training history of more than one weekly resistance-exercise session during the last 12 months leading up to the intervention, impaired muscle strength due to any type of injury or an intake of prescribed medication that could affect adaptations to training, were excluded. Seven participants were excluded due to not completing the scheduled training sessions for differing reasons: discomfort, pain or injury. Each participant performed a standardised warm-up routine consisting of 5 mins of ergometer cycling and ten repetitions of bodyweight exercises, including push-ups with individually adjusted leverage, sit-ups, back-extensions and squats. One set of 10 repetitions at ∼50% of the one-repetition maximum (1RM) for each resistance exercise was performed. Leg resistance exercises were performed in the following order: unilateral leg press, leg curl and knee extension. They were performed as either *one set* (*single set*) or *three sets* (*multiple sets*) per exercise for alternate legs selected randomly [117, 118]. Biopsies were collected from the musculus vastus lateralis before and after the training intervention and at the week two training session.

4.1.2. Resistant training in older adults with or without COPD (OLDER dataset)

The OLDER dataset comprised 44 male, and 51 female participants (age 56 – 77) enrolled in the study. Participants had a consumption of < 400 international units (IU) of vitamin D_3 per day for the two months leading up to the study, and either normal lung function or a medical diagnosis of COPD (GOLD grade II or III, FEV predicted between 80% - 30%, FEV/FVC < 70% (Forced Expiratory Volume/Forced Vital Capacity) after reversibility testing with inhalation of salbutamol and ipratropiumbromid Due to incompletion of the training protocol, 17 participants were excluded from the study. All participants performed unilateral leg press, unilateral knee extension, unilateral knee flexion, chest press and leg pulldown. Leg exercises were performed as three sets of 10 repetitions (high-load) and another of 30 repetitions

(low-load) or to exhaustion. All three sets for one leg were conducted before the other leg was exercised [119]. Biopsies were collected from *m. vastus letralis* before supplementation, before and after the training intervention and at 3.5 weeks mark of the training intervention.

4.2. Biopsy collection

Muscle tissue was obtained bilaterally from *m. vastus lateralis* using a 12-gauge needle (Universal-plus, Medax, San Possidonio, Italy) under local anaesthesia (Xylocaine, 10 mg×ml−1mg×ml−1 with adrenaline 5 μg × ml^{-1}, AstraZeneca AS, Oslo, Norge).

4.3. RNA isolation

Approximately 10 – 20 mg of wet muscle tissue was homogenised in a total volume of 1 mL TRIzol reagent (Invitrogen, Carlsbad, CA, USA) using 0.5 mm RNase-free zirconium oxide beads and a bead homogeniser (Bullet Blender, Next Advance, Averill Park, NY, USA), as previously described [117]. To enable an analysis of target gene expression per unit tissue weight, an exogenous RNA control (λ polyA External Standard Kit, Takara Bio Inc., Shiga, Japan) was added at a fixed amount (0.04 ng/mL of Trizol reagent) per extraction before homogenisation [120, 121]. Following phase separation, 450 μL of the upper phase was transferred to a new tube and RNA was precipitated using isopropanol. The resulting RNA pellet was washed three times with 75% ethanol, eluted in 30 μL TE buffer and diluted to 100 ng RNA/μL, following quantification of total RNA concentration using μDrop plate and the Multiskan GO microplate spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA). RNA integrity was assessed using capillary electrophoresis (Experion Automated Electrophoresis Station using RNA StdSens Assay, Bio-Rad, Hercules, CA, USA) with an average integrity score <7.

4.4.RNA-Seq

mRNA sequencing libraries were prepared from 1000 ng of total RNA 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, San Diego, CA, USA) at the Norwegian Sequencing Centre, Oslo, Norway.

4.5.Bioinformatics tools

4.5.1. Data Preprocessing

Trim Galore (version 0.6.5, https://github.com/FelixKrueger/TrimGalore) and Trimmomatic (version 0.39) [122] were used to discard low-quality reads and to trim poor-quality bases

before alignment, using default settings. The quality of filtered files was calculated by FastQC (version 0.11.4) and summarised using MultiQC (version 1.8) [86].

4.5.2. Mapping to reference human genome

Filtered reads were aligned to the human genome (GRCh38 release-97 downloaded from ftp.ensemble.org) using the alignment-based methods HISAT2 (version 2.1.0) [123] , STAR (version 2.7.2) [87] and RSEM (version 1.3.1) [90] and were used together with Bowtie 2 (version 2.3.4.3) [97] and non-alignment methods including Kallisto (version 0.44.0) [91] and Salmon (version 0.13.1) [92]. For HISAT2 and STAR, SAM files generated as alignment results were used by HTSeq to generate a quantified expression of features [93]. RSEM, Kallisto, and Salmon have in-built quantification functions.

4.5.3. Normalisation

Normalisation accounts for bias due to sequencing depth, gene length was performed [124, 125]. It also aims to remove systemic technical effects and ensure that it has no or negligible effect on results [126].

4.5.4. Sequencing for Spike-in sequence

Commercial spike-in control (λ polyA External Standard Kit, Takara Bio Inc., Shiga, Japan) was added before preparing samples for sequencing. For propriety reasons, the sequence of spike-in control was confidential and therefore was not shared by the vendor. Sequencing was performed on spike-in control using Illumina HiSeq 3000. Filtered reads were assembled using Spades v3.15.2 [127]. The scaffolds were blast searched against the NCBI reference genome database (phage genome). The reference genome was identified, and a longer scaffold was made by manually joining the reads aligned on the genome perfectly. Expression for this sequence was identified for all the files by alignment and quantification by RSEM and was used as the normalisation factor.

4.5.5. Differentially expressed (DE) genes

Three different normalisation strategies were compared based on negative binomial and GLMM, i.e. DESeq2 [98], edgeR [101] and the GLMM approach [112].

4.5.6. Gene ontology

Enrichment analyses of gene ontology (GO) gene sets were performed using three different approaches. First, a non-parametric rank test which is described in Yamaguchi et al. [128] and implemented in the tmod package v.0.40 [129], 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 fold-change (FC) when log(FC) > 0 and the negative inverse of the upper 95% CI when log(FC) < 0. This metric has been shown to have lower false-positive rates compared to other metrics applied during enrichment analyses [130]. 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) [131], was used to quantify the directional regulation of the gene set. GSEA was performed using the fgsea package [132] with Log (FC) as the gene level metric. Thirdly, over-representation analysis (ORA) was performed to assess if genes identified as DE (|Log2 fold-change|> 0.5 and adjusted *p*-values < 0.05; DE-genes) belonged to specific gene sets. ORA was performed using the enrichGO function in the clusterProfiler package [113] (version 3.16.0). GO gene sets (biological process, cellular component and molecular function) were retrieved from the molecular signature database (version 7.1)

5. Knowledge gaps and aim of the thesis

Primary aim: *To develop RNA-seq data analysis strategies to analyse skeletal muscle transcriptome responses to resistance training and to employ these to study two different datasets.*

Sub-objective 1

Background: In recent years, several benchmarking studies have shown strong disagreement with each other on which tool or method performs best for the analysis of the RNA-seq data. The reason for this disagreement may lie in the usage of different datasets and/or simulated data to evaluate these tools. It is possible that each dataset type requires its own data analysis pipeline. The normalisation of technical bias is one of the crucial steps in data analysis. However, biological biasness is also needed to be considered.

Aim: *To develop a pipeline for analysis of RNA-seq dataset from skeletal muscle training using the following secondary objectives:*

- To establish a bioinformatic pipeline specific for the analysis of RNA-seq data from skeletal muscles. (Paper I)
- To explore the effects of using different normalisation strategies for analysing and interpreting skeletal muscle transcriptome responses to resistance training. (Paper I and II)

Sub-objective 2

Background: Resistance training is a potent driver of skeletal muscle plasticity, and the degree of plasticity can be affected by training volume, vitamin D_3 supplementation and physiological condition like COPD. However, little is known about the interplay between skeletal muscle plasticity and these variables.

Aim: *The standardised RNA-seq pipeline was employed to analyse skeletal muscle transcriptome data from two human interventions,* including (i) young and healthy adults performing 12 weeks of resistance training with two different volumes; and (ii) older participants with and without COPD condition performing 13 weeks of resistance training, with and without Vitamin D_3 supplementation.

This was achieved using the following secondary objective:

- To explore the effects of resistance training with low and moderate volume for transcriptome responses in skeletal muscle in young adults. (Paper I)

- To explore the impact of vitamin D_3 supplementation and COPD diagnosis for skeletal muscle transcriptome responses to resistance training. (Paper III and IV)

6. Results and Discussion

6.1. Standardisation of RNA-seq pipeline for skeletal muscle

NGS has become a valuable tool for gene expression studies in precision medicine, diagnostics and population studies during the past few decades [133]. The data generated from such studies represent gene expression of the whole mRNA population at a given condition with remarkable detail. Despite a wealth of available data and studies focused on comparing different tools for RNA-seq data analysis [30, 32, 94], there is no pipeline specific for analysis of skeletal muscle data or transcriptome response to resistance training. Moreover, reports on which tools are best for such analyses vary from dataset to dataset. Therefore, it is crucial to identify tools specifically for resistance training-induced growth in skeletal muscle. The YOUNG dataset was used to compare different tools to obtain a skeletal muscle specific pipeline.

6.1.1. Tools for preprocessing of data

Before the analysis, raw sequencing data needs to be quality filtered, and sequencing adaptors should also be removed. Trimming and quality filtering affect downstream analysis by removing low-quality and ambiguous reads before alignments. Such steps, collectively referred to as data pre-processing, apply to all sequencing technologies where sequence quality needs to be monitored [134]. These steps are collectively referred to as data preprocessing. As there is no single answer for the generic question "Which tool is best for preprocessing?"; we decided to test the most frequently cited tools data pre-processing, trimmomatic and trim galore. Trimmomatic and trim galore are cited in more than 27100 and 11600 articles, respectively (listed on Google Scholar on 15 October 2021). Before and after preprocessing, quality was assessed by FASTQC and summarised by multiQC (**Paper I**: Figure 1G). Based on this quality assessment, trim galore failed to filter out low-quality data, whereas trimmomatic was more effective in removing low-quality bases for the YOUNG dataset. This is because adaptor sequence search in trimmomatic is more effective than trim galore [135]. Although trim galore retained a higher number of sequences after filtering, their average quality was lower. In addition, trimmomatic was fast in processing; it took 12 minutes 12 seconds per million reads, while trim galore took 19 minutes 13 seconds per million reads. These run times are similar (trimmomatic – 8 minutes and trim galore 23 minutes) to what has been earlier reported [135]. Using data without preprocessing can lead to alignments with a low score which can be dropped if they have many mismatches due to the presence of adaptor sequences [136]. Therefore, we used trimmomatic to do the preprocessing before performing the read alignment step.

6.1.2. Selection of alignment tool

Filtered reads were aligned on the Human Genome GRCh38 using three different alignment-based methods (STAR, HISAT2 and RSEM) and two non-alignment-based methods (Kallisto and salmon). RSEM, Kallisto and Salmon identified more features than STAR and HISAT2 (Figure 6). Transcript based methods aligned around ten times the number of reads and identified 50% more features as compared to the genome-based methods. The average reads aligned by STAR, HISAT2 were less than two million reads, and less than 15000 features were identified. In contrast, Kallisto and salmon aligned, more than 12 million reads, and RSEM aligned with over 8.5 million reads.

All transcript-based aligners identified more than 22000 features/genes. However, a higher number of genes is not a good criterion for assessing the performance of alignment tools. Instead, variability in replicates can be used as a measure of alignment performance [94]. RSEM had the lowest standard deviation between technical replicates in resting samples before the training protocol, followed by Salmon and Kallisto (**Paper I**: Figure 1J). In addition, alignment performance can be assessed by mRNA to protein expression correlations. To compare alignment performance between tools, we specifically used mRNA and the protein profile of the main myosin heavy chain genes. The myosin heavy chain protein profile and RNA-seq based myosin heavy chain family mRNA show a higher degree of correlation with RSEM count, which was lowest in HISAT and STAR (**Paper I**: Figure 2a).

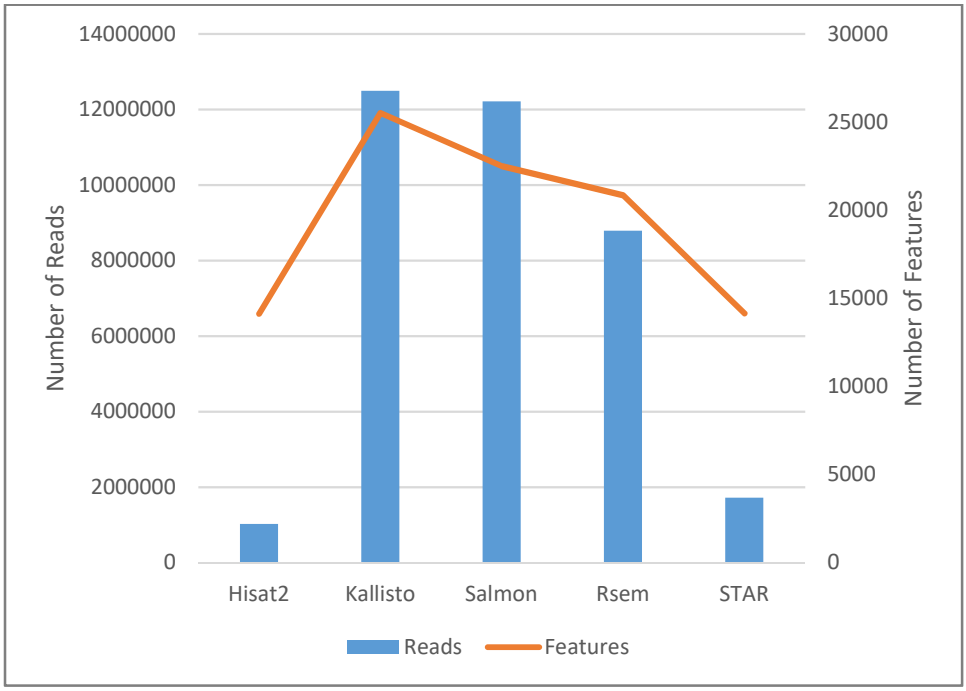


Figure 6. *The average number of reads identified after preprocessing and the average number of features identified after alignment.*

6.1.3. Total RNA concentration changes with resistance training

Before identifying DE genes, count data need to be normalised to reduce the effects of technical bias and variation. TMM and RLE, which are the most used RNA-seq data normalisation strategies are based on the assumption that all cell produces an equal amount of RNA at a given condition [33]. In contrast to this assumption, increased total RNA per unit tissue weight was seen in the samples obtained after resistance training [117]. Samples obtained at weeks 2 and 12 had higher total RNA concentrations than resting samples (week 0). Total RNA concentrations were also affected by training volume. As a fixed amount of total RNA was used in all preparations, these effects resulted in lesser amounts of tissue used in library preparations in trained compared to non-trained samples and in high-volume samples compared to low-volume samples (**Paper I:** Figure 3a and 3b). The mRNA volume has also increased with total RNA, which is evident from larger library sizes during the training intervention and suggests that it is not feasible to ignore total RNA changes. Previous studies by Lin et al. and Nie et al. in tumour cells and stem cells, respectively, found that proto-oncogene *c-Myc*, which plays a role in cell cycle progression, increases the total RNA as an effect of *c-MyC* [105, 137]. This changes the assumption that the cell produces a similar amount of total RNA, which is needed to account for better biologically correct normalisation factors (**Paper I and II**). Moreover, spike-in was added in all the samples in equal ratio with total RNA, acting as a negative control to normalize the count matrix.

6.1.4. Spike-in sequence identifications

Raw sequence files were filtered for quality and de novo assembled by Spades, which resulted in 26 contigs with lengths larger than 200 bp. The maximum contigs size was 837 bp. These selected contigs were blast against the NCBI non-redundant nucleotide database, and the sequence was found to match against the *EA22* CDS of *Enterobacteria* phage lambda genome (NC_001416.1). Three of the contigs aligned with 99% sequence identity. After joining these contigs based on the alignment on the lambda genome resulted in a 954 bp long consensus sequence. This sequence was treated as the spike-in transcript for downstream analysis. Based on our previous analysis of standardising the analysis pipeline alignment tool, RSEM was selected to perform the alignment and quantification. The above selected spike-in transcript was aligned to the RNA-seq data to get the spike-in expression for all the biopsy samples.

6.1.5. Normalisation strategies and their effects on RNA-seq data

Different normalisation methods can lead to different results for identifying stable and DE genes [29, 138-140]. In this study, four different normalization methods were compared, i.e.

mRNA (TMM), total RNA (raw), tissue size (tissue offset) and spike-in. Figure 7 represents the difference in expression of all the expressed genes by the coefficient of variance (CV) value calculated by the ratio of standard deviation and mean. Library size normalization and mRNA (TMM) normalization showed similar CV values, while high CV values are observed in spike-in and tissue offset normalization.

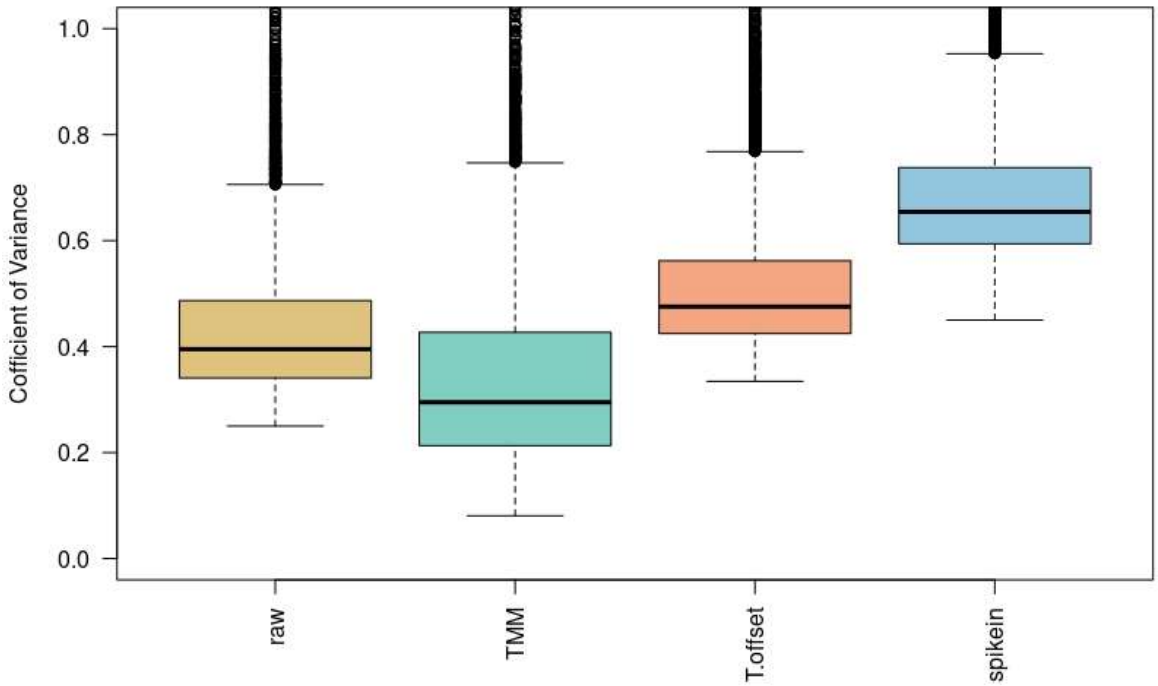


Figure 7. *Coefficient of variance by different normalisation strategies, including - library size, TMM, tissue offset and Spike-in.*

6.1.6. Selection of stable transcripts

In many studies, CV value and logFC have been used to identify stable genes from the normalised count, and to account for differences in various normalisation techniques, common genes were taken using count data generated by different normalisation strategies [111, 141-144]. However, CV value does not account for systemic effect and intraclass correlation, which is defined as the amount of variance attributed to between-participant variation relative to the total variance. A cutoff of 1.5 t-value is used to filter for potential stable genes. These transcripts were ranked by ascending interclass correlation (ICC). TMM and spike-in normalised count data reported 1266 and 1337 transcripts, respectively (Figure 8), and the selected stable genes have high ICC value and rank.

High ICC value transcripts represent the most stable genes across the samples (**Paper II**: Figure 1) and (Figure 8). Total RNA (raw) and tissue-offset normalised count identified 90 and 19 transcripts, respectively; selected transcripts having high ICC scores are coloured (Fig.

8). Transcripts selected by these methods are compared for their expression change by normalizing with normalisation factors calculated by shortlisted transcripts (Figure 9), which shows a slight change in transcript expression (of transcripts listed in Table 2 and **Paper II**: Table 1) while comparing to each other. Comparing the average ratio of fold changes with time and volume between normalisation strategies (Figure 9), the least differences were identified between spike-in and tissue normalised data compared to any other comparison. It shows a 10% fold change difference at week 2 and 5% at week 12 for multiple set data, and single set training has less effect as muscle growth have less difference between fold changes. The similarity between tissue weight and spike-in normalised fold changes is due to the spike-in control sequence being added in accordance with tissue weight. Even if spike-in and TMM normalised counts reduced technical biasness, it could not account for biological biasness of tissue weight usage during RNA isolation. This suggests that tissue weight should be used to normalise the count. One of the advantages of tissue weight normalisation is that there is no additional cost for external spike-in purchase and sequencing.

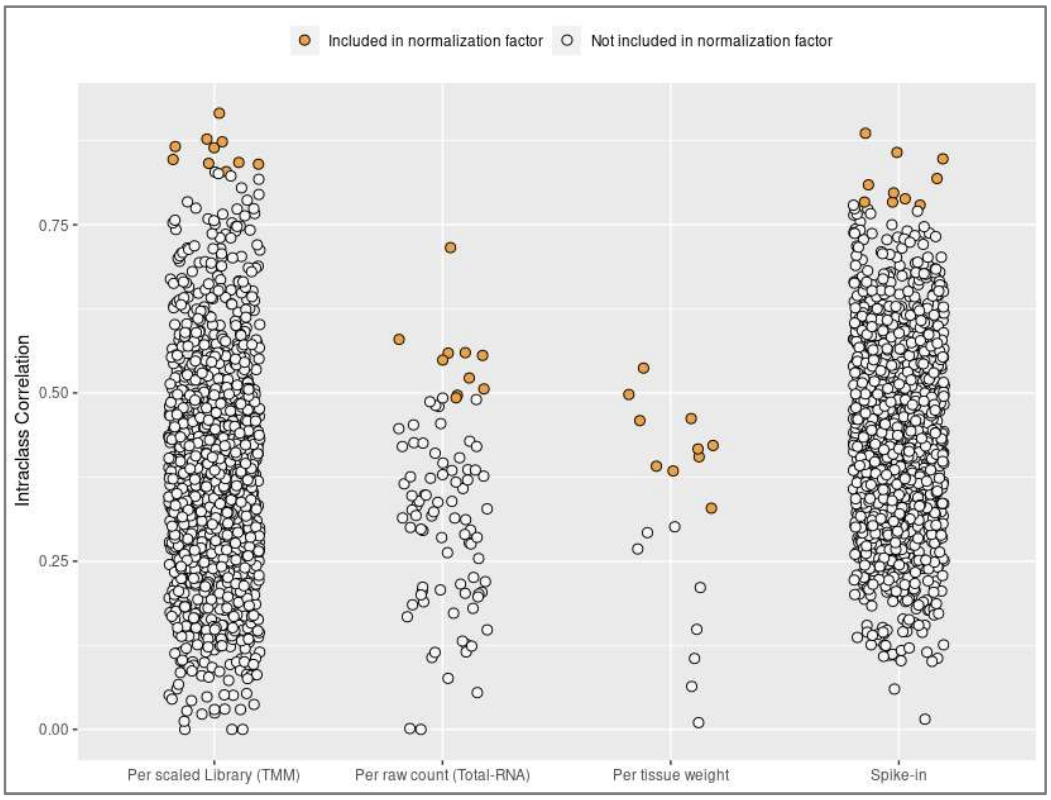


Figure 8. *Genes were selected by different normalisation methods based on the intraclass correlation among genes with systemic effect (< 1.5). Selected stable transcripts are coloured orange.*

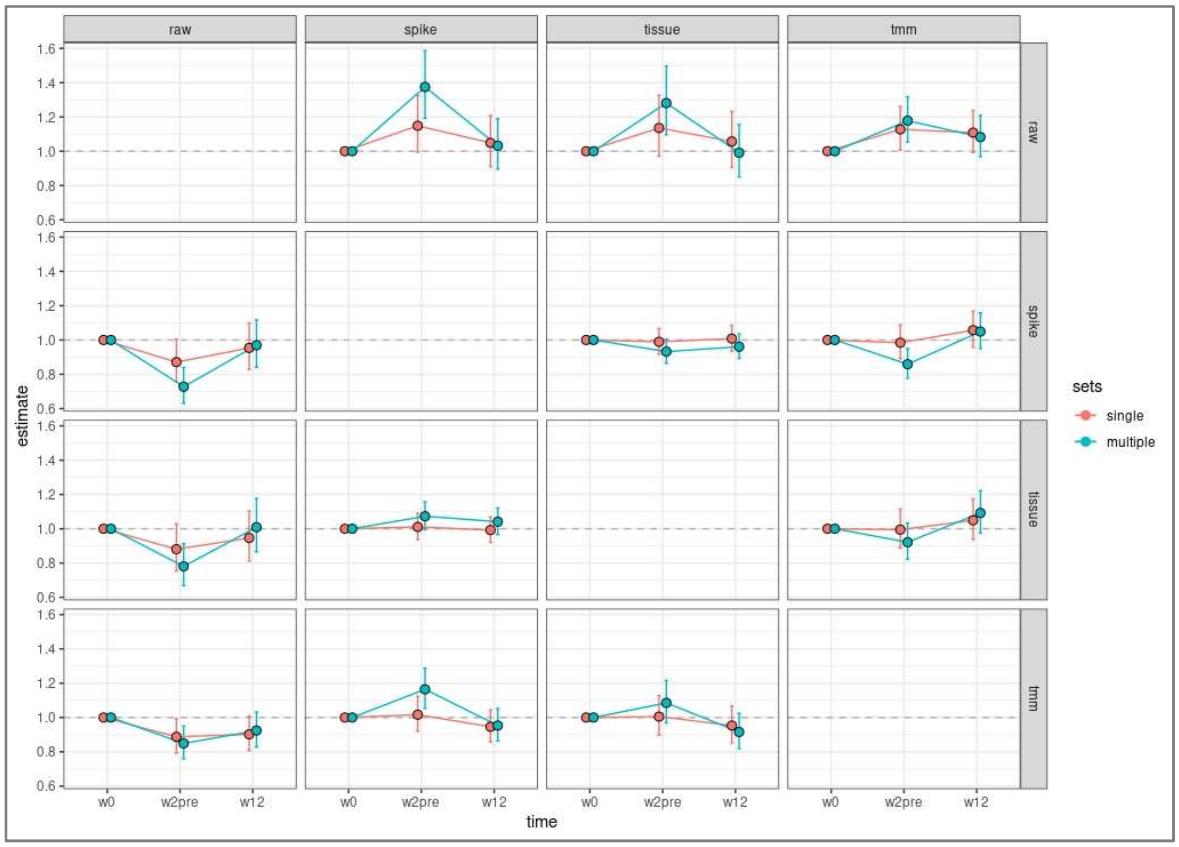


Figure 9. *The ratio of normalisation factors over time and volume condition.*

Table 2. *Genes selected as stable reference genes from each normalisation scenario.*

Transcript ID	Gene Symbol	Gene biotype	Intraclass correlation
ENST00000457540	*MTND2P28*	pseudogene	0.8859148
ENST00000361681	*MTND6*	Protein coding	0.8574938
ENST00000346365	*NUDT2*	Protein coding	0.8479227
ENST00000522543	*ANK1*	Protein coding	0.8186528
ENST00000295955	*RPL9*	Protein coding	0.809337
ENST00000361227	*MTND3*	Protein coding	0.7973469
ENST00000445193	*RPP14*	Protein coding	0.7885435
ENST00000520612	*MTND6P3*	pseudogene	0.7837061
ENST00000342751	*SDHC*	Protein coding	0.783665
ENST00000370474	*SDHAF4*	Protein coding	0.7792214

6.1.7. Differential gene expression by different tools

Two different approaches were tested for DE analysis - negative binomial models (DESeq2 and edgeR) and negative binomial Generalised Linear Mixed Model (GLMM). The advantage of using GLMM is that, unlike negative binomial models, this method allows for incorporating both random and fixed effects, which is not possible in DESeq2 and edgeR. For comparing the

simplest experimental design, a subset of data (weeks 0 and 2) from the YOUNG dataset was used to look for DE genes. EdgeR identified 1689 genes and DESeq2 identified 943 DE genes, having 920 overlapping genes with each other. In comparison, GLMM identified 340 DE genes and 319 genes common in all methods. DESeq2 and edgeR are based on sample size calculation and power estimation at the marginal or dataset level to test DE genes. However, these methods cannot be applied to correlated expression data, as they can lead to biased estimates and fail on error rate control and increase false-positive results [145-147]. To overcome this, the GLMM approach was considered for experimental design with a diverse correlation structure. This (GLMM) procedure can be applied to multifactorial design and allow multiple random effects [112, 147]. So, even though observing a higher number of DE genes, edgeR and DESeq2 were not considered for further analysis.

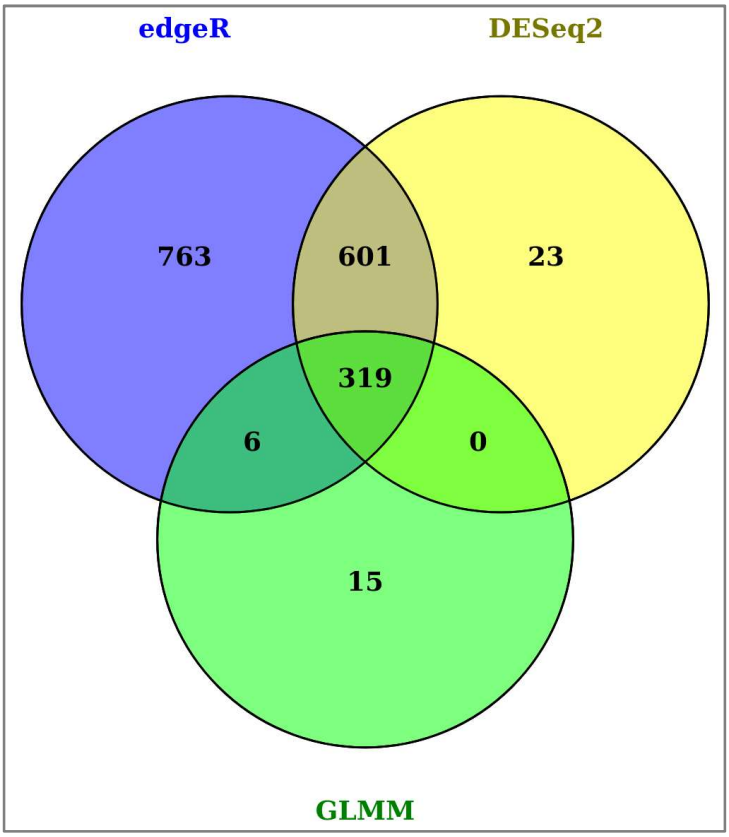


Figure 9. *Venn diagram comparing DE genes identified by each method in comparing week 0 and week 2.*

6.1.8. Transcriptomic changes related to training volume

We found 21 genes with higher expression in the moderate- compared to the low-volume condition at week 2. At weeks 0 and 12, there were no DE genes (**Paper I**: Figure 3D and 3H). Functional analysis of these DE genes revealed a gene ontology relationship with an ECM

(**Paper I:** Figure 3F ad Table 2). Different ECM changes induced by different training volumes have been reported in a few studies ([148] and **Paper I**). During systematic and repeated exercise, cells undergo changes in their mRNA and protein pool, specifically more in ECM related genes [148]. The formation of new fibre tubulin, signalling, cell size is related to ECM related gene ontologies [148]. Genes related to ECM function are upregulated during the early stages of resistance training and during prolonged endurance and resistance training, which has a beneficial effect on high-volume resistance training [149-156]. This is a novel finding, as before this, ECM remodelling gene changes for different training volumes has not been identified. Also, requiring fifty-three 1 (*RFT1)* gene showed a 6-fold increase in expression in moderate vs low volume resistance training. GO analysis of the *RFT1* gene showed that it was linked with lipid transport, carbohydrate transport and is located in the endoplasmic reticulum membrane. Previously, Pillion et al. have reported a decrease in *RFT1* immediately after training [157].

6.1.9. Transcriptomic changes with time during training

3923 genes were upregulated, and 77 genes were downregulated at week 2. In addition, at week 12, 1733 genes were upregulated, and two genes were downregulated (**Paper I:** Figure 4). GO terms associated with these genes belonged to ECM remodelling, which is expected as we have seen an increase in muscle mass and training act as stress [158].

6.2.Application of standardised RNA-seq pipeline on the OLDER dataset

6.2.1. Vitamin D_3 effect on muscle growth

Vitamin D_3 supplementation showed a remarkable increase in (25(OH)D) and ($1,25(OH)_2D$) (**Paper III**: Figure 2 B and C), and with 12 weeks of resistance training showed increased muscle strength (**Paper III**: Figure 6A) and muscle mass (**Paper III**: Figure 7A). Vitamin D_3 itself did not exhibit any effect on muscle mass (**Paper III**: Figure 6B and 7B). However, minor gene expression changes have been observed in the transcriptome profile due to vitamin D_3 supplementation (**Paper III**: Figure 5A) as we identified 27 upregulated and 27 downregulated genes. This included increased expression of B-cell lymphoma 6 (*BCL6*) and prolyl 4-hydroxylase subunit alpha-1 (*P4HA1*), both of which are known to oppose accumulation of ROS, and decreased expression of angiopoietin-like protein 4 (*ANGPTL4*), which is closely correlated with levels of mitochondrial respiration. This was supported by gene enrichment analysis, which showed a general reduction in the expression of genes related to oxidative and

glycolytic metabolism in the vitamin D_3 arm altered expression of genes related to mitochondrial function (**Paper III**: Figure 5B).

Besides, vitamin D_3 supplementation had no effect on training-associated changes in gene expression, either at 3.5 weeks or at 13 weeks (**Paper III**: Figure 11D and E). This suggests that no single gene was differentially affected by combined vitamin D_3 supplementation and resistance training and resistance training-only. In contrast, enrichment analyses showed minor vitamin D_3-sensitive changes in expression at both 3.5 and 13 weeks (**Paper III**: Figure 11F). After 3.5 weeks of training, there was DE of genes involved in cell junctions, blood vessel morphogenesis and muscle cell differentiation. After 13 weeks, the vitamin D_3 arm showed increased expression of genes involved in endothelial proliferation and blood vessel morphogenesis compared with placebo.

However, the role of vitamin D_3 cannot be denied, as it increases total lean mass by 1.4% and reduces lipid level and fat mass (**Paper III**: Table S11). Vitamin D_3 controls calcium metabolism to increase calcium uptake and reduce intestinal fatty acid absorption [123]. Also, our results suggest that vitamin D_3 supplementation reduced lung function by approximately 3% (**Paper III**: Table S2), which is in contrast to a previous study that reported vitamin D_3 deficiency is associated with reduced lung function [159].

6.2.2. Muscle mass change in COPD diagnosed vs healthy individuals

With decreased lung function and sarcopenic condition, COPD diagnosed participants exhibited impaired training capabilities (**Paper IV**: Table 1) [160]. During training, COPD participants exhibited a higher increase in muscle strength than healthy participants (**Paper IV**: Figure 4). However, joining both high and low load training modalities reported similar outcomes. Both groups averagely exhibited similar improvement in muscle mass gain, but overall improvement was better in COPD participants. A baseline baseline-level transcriptome profile between COPD diagnosed participants and healthy participants displayed 227 genes being DE (151 upregulated and 76 downregulated). However, no transcript showed differential responses at the single-gene level at either time point, despite clear differences in transcriptome profile at the baseline. In contrast, the enrichment analysis revealed minor increases in expression of genes relating to oxidative phosphorylation after 3.5 weeks and decreases in genes associated with myogenesis after 13 weeks (**Paper IV**: Figure 3A and B; Table S2).

7. Conclusion

A standardized pipeline for analysis of RNA-seq dataset from human skeletal muscle was developed based on the YOUNG dataset. For pre-processing of data, trimmomatic was found to be a better choice than trim galore. For read alignment and quantification, RSEM proved to be a better tool having less variation between replicates and increased biological relevance. Also, the effects of using different normalisation strategies for analysing and interpreting skeletal muscle transcriptome responses to resistance training was explored. Analysis of data from this model showed that the choice of the normalization model affected downstream data interpretations. While all normalization models arguably remove technical biasness, our results suggest that it may be necessary to account for biological biasness caused by global changes in total RNA population (i.e., total RNA per tissue mass) to increase the biological relevance of transcriptome analyses from resistance-trained skeletal muscle. It is further suggested that assumptions of specific standardised RNA-seq pipelines are explicitly tested in each dataset.

The developed RNA-seq pipeline was employed to analyse skeletal muscle transcriptome data from the YOUNG and the OLDER human interventions. Firstly, in the YOUNG dataset, the effects of resistance training with low and moderate volume for transcriptome responses in skeletal muscle in young adults were explored. At week 2, the two-volume condition was associated with DE of 21 genes, mostly related to ECM remodelling. In contrast, at Week 12, no readily explainable dose-dependencies were observed.

Secondly, in older participants, the impact of vitamin D_3 supplementation and COPD diagnosis for skeletal muscle transcriptome responses to resistance training was explored. Overall, resistance training improved muscle functions and increased muscle mass, but there was no additional effect of vitamin D_3 supplementation. COPD participants showed hitherto largely unrecognized responsiveness to resistance training as they improved muscle functions and increased muscle mass to a similar or larger extent than healthy participants. After 3.5 weeks of resistance training, enrichment analysis showed that COPD participants had a greater increase in expression of genes related to oxidative phosphorylation than healthy participants. After 13 weeks of resistance training, COPD showed more pronounced decreases in genes related to myogenesis.

In conclusion, the developed RNA-seq data analysis protocol to analyse skeletal muscle transcriptome responses to resistance training was successfully employed to study the YOUNG and OLDER datasets.

8. Future work

This study was based on a comparative analysis of different tools and strategies used in RNA-seq. Our results present an RNA-seq data analysis pipeline for resistance training-induced growth skeletal muscles data. It is suggested to perform similar comparison/s using additional tools and models for analysis and to use different datasets.

Transcriptome analysis has been altered by resistance training, but gene expression is not the only aspect that needs investigation. Small ncRNAs (e.g. miRNA, siRNA) also regulated transcription. Small RNA sequencing can be done to find the relationship between small RNA and mRNA. There have been studies for myo-MIRs (micro RNAs). However, a joined experiment is needed to explore mRNA and small RNA changes together. A study design consisting of four different training modules involving no training, low, moderate and very high training is needed. This can be achieved by studying data of these contrast, which could lead to improved training protocols for healthy ageing. Additionally, it can also aid in designing miRNA-based medication for the population with critical muscle mass.

Moreover, personalised training interventions can be designed using transcriptome response as a rehabilitation program for a sedentary lifestyle and age-related loss of muscles.

9. References

1. Fontana L, Kennedy BK, Longo VD, Seals D, Melov S: **Medical research: treat ageing**. *Nature* 2014, **511**(7510):405-407.
2. Cartee GD, Hepple RT, Bamman MM, Zierath JR: **Exercise Promotes Healthy Aging of Skeletal Muscle**. *Cell Metab* 2016, **23**(6):1034-1047.
3. Silverthorn DU, Johnson BR, Ober WC, Ober CE, Impaglizzo A, Silverthorn AC: **Human physiology : an integrated approach**, 8th Edition, 9780134605197 : Pearson; 2018.
4. Booth FW, Laye MJ, Roberts MD: **Lifetime sedentary living accelerates some aspects of secondary aging**. *Journal of Applied Physiology* 2011, **111**(5):1497-1504.
5. Chodzko-Zajko W, Schwingel A, Chae Hee P: **Successful Aging: The Role of Physical Activity**. *American Journal of Lifestyle Medicine* 2008, **3**(1):20-28.
6. Charansonney OL: **Physical activity and aging: a life-long story**. *Discov Med* 2011, **12**(64):177-185.
7. Aube D, Wadhi T, Rauch J, Anand A, Barakat C, Pearson J, Bradshaw J, Zazzo S, Ugrinowitsch C, De Souza EO: **Progressive Resistance Training Volume: Effects on Muscle Thickness, Mass, and Strength Adaptations in Resistance-Trained Individuals**. *J Strength Cond Res* 2020.
8. La Scala Teixeira CV, Motoyama Y, de Azevedo P, Evangelista AL, Steele J, Bocalini DS: **Effect of resistance training set volume on upper body muscle hypertrophy: are more sets really better than less?** *Clin Physiol Funct Imaging* 2018, **38**(5):727-732.
9. Lambert CP, Frank LL, Evans WJ: **Macronutrient considerations for the sport of bodybuilding**. *Sports Med* 2004, **34**(5):317-327.
10. Mangine GT, Hoffman JR, Gonzalez AM, Townsend JR, Wells AJ, Jajtner AR, Beyer KS, Boone CH, Miramonti AA, Wang R *et al*: **The effect of training volume and intensity on improvements in muscular strength and size in resistance-trained men**. *Physiol Rep* 2015, **3**(8).
11. Ralston GW, Kilgore L, Wyatt FB, Dutheil F, Jaekel P, Buchan DS, Baker JS: **Re-examination of 1- vs. 3-Sets of Resistance Exercise for Pre-spaceflight Muscle Conditioning: A Systematic Review and Meta-Analysis**. *Frontiers in Physiology* 2019, **10**:864.
12. Schoenfeld BJ, Contreras B, Krieger J, Grgic J, Delcastillo K, Belliard R, Alto A: **Resistance Training Volume Enhances Muscle Hypertrophy but Not Strength in Trained Men**. *Med Sci Sports Exerc* 2019, **51**(1):94-103.
13. Solberg PA, Kvamme NH, Raastad T, Ommundsen Y, Tomten SE, Halvari H, Loland NW, Hallén J: **Effects of different types of exercise on muscle mass, strength, function and well-being in elderly**. *European Journal of Sport Science* 2013, **13**(1):112-125.
14. Williams MH: **Dietary supplements and sports performance: minerals**. *J Int Soc Sports Nutr* 2005, **2**:43-49.
15. Magherini F, Fiaschi T, Marzocchini R, Mannelli M, Gamberi T, Modesti PA, Modesti A: **Oxidative stress in exercise training: the involvement of inflammation and peripheral signals**. *Free Radic Res* 2019, **53**(11-12):1155-1165.
16. Ebner N, Sliziuk V, Scherbakov N, Sandek A: **Muscle wasting in ageing and chronic illness**. *ESC Heart Fail* 2015, **2**(2):58-68.
17. Lozano R, Naghavi M, Foreman K, Lim S, Shibuya K, Aboyans V, Abraham J, Adair T, Aggarwal R, Ahn SY *et al*: **Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010**. *The Lancet* 2012, **380**(9859):2095-2128.

18. WHO: **The top 10 causes of death– WHO**. *WHO* 2020.
19. WHO: **COPD**. *WHO* 2021.
20. Gea J, Martinez-Llorens J, Barreiro E: **[Nutritional abnormalities in chronic obstructive pulmonary disease]**. *Med Clin (Barc)* 2014, **143**(2):78-84.
21. Maltais F, Decramer M, Casaburi R, Barreiro E, Burelle Y, Debigare R, Dekhuijzen PN, Franssen F, Gayan-Ramirez G, Gea J *et al*: **An official American Thoracic Society/European Respiratory Society statement: update on limb muscle dysfunction in chronic obstructive pulmonary disease**. *Am J Respir Crit Care Med* 2014, **189**(9):e15-62.
22. Passey SL, Hansen MJ, Bozinovski S, McDonald CF, Holland AE, Vlahos R: **Emerging therapies for the treatment of skeletal muscle wasting in chronic obstructive pulmonary disease**. *Pharmacol Ther* 2016, **166**:56-70.
23. Agergaard J, Trøstrup J, Uth J, Iversen JV, Boesen A, Andersen JL, Schjerling P, Langberg H: **Does vitamin-D intake during resistance training improve the skeletal muscle hypertrophic and strength response in young and elderly men? - a randomized controlled trial**. *Nutr Metab (Lond)* 2015, **12**:32-32.
24. Gillespie LD, Robertson MC, Gillespie WJ, Sherrington C, Gates S, Clemson LM, Lamb SE: **Interventions for preventing falls in older people living in the community**. *Cochrane Database Syst Rev* 2012(9):CD007146.
25. Glendenning P, Zhu K, Inderjeeth C, Howat P, Lewis JR, Prince RL: **Effects of three-monthly oral 150,000 IU cholecalciferol supplementation on falls, mobility, and muscle strength in older postmenopausal women: A randomized controlled trial**. *Journal of Bone and Mineral Research* 2011, **27**(1):170-176.
26. Goswami R, Vatsa M, Sreenivas V, Singh U, Gupta N, Lakshmy R, Aggarwal S, Ganapathy A, Joshi P, Bhatia H: **Skeletal Muscle Strength in Young Asian Indian Females after Vitamin D and Calcium Supplementation: A Double-Blind Randomized Controlled Clinical Trial**. *The Journal of Clinical Endocrinology & Metabolism* 2012, **97**(12):4709-4716.
27. Girgis CM, Clifton-Bligh RJ, Hamrick MW, Holick MF, Gunton JE: **The roles of vitamin D in skeletal muscle: form, function, and metabolism**. *Endocr Rev* 2013, **34**(1):33-83.
28. Morin R, Bainbridge M, Fejes A, Hirst M, Krzywinski M, Pugh T, McDonald H, Varhol R, Jones S, Marra M: **Profiling the HeLa S3 transcriptome using randomly primed cDNA and massively parallel short-read sequencing**. *Biotechniques* 2008, **45**(1):81-94.
29. Rapaport F, Khanin R, Liang Y, Pirun M, Krek A, Zumbo P, Mason CE, Socci ND, Betel D: **Comprehensive evaluation of differential gene expression analysis methods for RNA-seq data**. *Genome Biol* 2013, **14**(9):R95.
30. Seyednasrollah F, Laiho A, Elo LL: **Comparison of software packages for detecting differential expression in RNA-seq studies**. *Briefings in bioinformatics* 2015, **16**(1):59-70.
31. Soneson C: **compcodeR--an R package for benchmarking differential expression methods for RNA-seq data**. *Bioinformatics* 2014, **30**(17):2517-2518.
32. Soneson C, Love MI, Robinson MD: **Differential analyses for RNA-seq: transcript-level estimates improve gene-level inferences**. *F1000Res* 2015, **4**:1521-1521.
33. Loven J, Orlando DA, Sigova AA, Lin CY, Rahl PB, Burge CB, Levens DL, Lee TI, Young RA: **Revisiting global gene expression analysis**. *Cell* 2012, **151**(3):476-482.
34. Panina Y, Germond A, Masui S, Watanabe TM: **Validation of Common Housekeeping Genes as Reference for qPCR Gene Expression Analysis During iPS Reprogramming Process**. *Scientific reports* 2018, **8**(1):8716-8716.
35. Srikanthan P, Karlamangla AS: **Muscle mass index as a predictor of longevity in older adults**. *Am J Med* 2014, **127**(6):547-553.

36. Wolfe RR: **The underappreciated role of muscle in health and disease**. *The American Journal of Clinical Nutrition* 2006, **84**(3):475-482.
37. Barany M: **ATPase activity of myosin correlated with speed of muscle shortening**. *J Gen Physiol* 1967, **50**(6):Suppl:197-218.
38. Geeves MA, Holmes KC: **The Molecular Mechanism of Muscle Contraction**. In: *Advances in Protein Chemistry.* vol. 71: Academic Press; 2005: 161-193.
39. Huxley AF: **Muscle structure and theories of contraction**. *Prog Biophys Biophys Chem* 1957, **7**:255-318.
40. Hoppeler H, Baum O, Lurman G, Mueller M: **Molecular mechanisms of muscle plasticity with exercise**. *Compr Physiol* 2011, **1**(3):1383-1412.
41. Ralston GW, Kilgore L, Wyatt FB, Dutheil F, Jaekel P, Buchan DS, Baker JS: **Re-examination of 1- vs. 3-Sets of Resistance Exercise for Pre-spaceflight Muscle Conditioning: A Systematic Review and Meta-Analysis**. *Frontiers in physiology* 2019, **10**:864-864.
42. Dattilo M, Antunes HKM, Medeiros A, Mônico Neto M, Souza HS, Tufik S, de Mello MT: **Sleep and muscle recovery: Endocrinological and molecular basis for a new and promising hypothesis**. *Medical Hypotheses* 2011, **77**(2):220-222.
43. Trommelen J, van Loon LJC: **Pre-Sleep Protein Ingestion to Improve the Skeletal Muscle Adaptive Response to Exercise Training**. *Nutrients* 2016, **8**(12):763.
44. Damas F, Phillips SM, Libardi CA, Vechin FC, Lixandrão ME, Jannig PR, Costa LAR, Bacurau AV, Snijders T, Parise G *et al*: **Resistance training-induced changes in integrated myofibrillar protein synthesis are related to hypertrophy only after attenuation of muscle damage**. *The Journal of physiology* 2016, **594**(18):5209-5222.
45. Könen T, Karbach J: **The benefits of looking at intraindividual dynamics in cognitive training data**. *Front Psychol* 2015, **6**:615-615.
46. Naderi A, de Oliveira EP, Ziegenfuss TN, Willems MT: **Timing, Optimal Dose and Intake Duration of Dietary Supplements with Evidence-Based Use in Sports Nutrition**. *J Exerc Nutrition Biochem* 2016, **20**(4):1-12.
47. Gibbs D: **Rickets and the crippled child: an historical perspective**. *J R Soc Med* 1994, **87**(12):729-732.
48. Holick MF: **Vitamin D: importance in the prevention of cancers, type 1 diabetes, heart disease, and osteoporosis**. *The American Journal of Clinical Nutrition* 2004, **79**(3):362-371.
49. Halfon M, Phan O, Teta D: **Vitamin D: a review on its effects on muscle strength, the risk of fall, and frailty**. *BioMed research international* 2015, **2015**:953241-953241.
50. Blanc PD, Eisner MD, Trupin L, Yelin EH, Katz PP, Balmes JR: **The association between occupational factors and adverse health outcomes in chronic obstructive pulmonary disease**. *Occup Environ Med* 2004, **61**(8):661-667.
51. Døssing M, Khan J, Al-Rabiah F: **Risk factors for chronic obstructive lung disease in Saudi Arabia**. *Respiratory Medicine* 1994, **88**(7):519-522.
52. Özbay B, Uzun K, Arslan H, Zehir Is: **Functional and radiological impairment in women highly exposed to indoor biomass fuels**. *Respirology* 2001, **6**(3):255-258.
53. Cosio MG, Saetta M, Agusti A: **Immunologic Aspects of Chronic Obstructive Pulmonary Disease**. *New England Journal of Medicine* 2009, **360**(23):2445-2454.
54. Djekic UV, Gaggar A, Weathington NM: **Attacking the multi-tiered proteolytic pathology of COPD: new insights from basic and translational studies**. *Pharmacol Ther* 2009, **121**(2):132-146.
55. MacNee W: **Pulmonary and Systemic Oxidant/Antioxidant Imbalance in Chronic Obstructive Pulmonary Disease**. *Proceedings of the American Thoracic Society* 2005, **2**(1):50-60.

56. Shi W, Chen F, Cardoso WV: **Mechanisms of lung development: contribution to adult lung disease and relevance to chronic obstructive pulmonary disease**. *Proceedings of the American Thoracic Society* 2009, **6**(7):558-563.
57. Jackson AS, Shrikrishna D, Kelly JL, Hart N, Moxham J, Polkey MI, Kemp P, Hopkinson NS: **Vitamin D and skeletal muscle strength and endurance in COPD**. *European Respiratory Journal* 2012, **41**(2):309-316.
58. Kalyani RR, Stein B, Valiyil R, Manno R, Maynard JW, Crews DC: **Vitamin D treatment for the prevention of falls in older adults: systematic review and meta-analysis**. *J Am Geriatr Soc* 2010, **58**(7):1299-1310.
59. Roth SM, Zmuda JM, Cauley JA, Shea PR, Ferrell RE: **Vitamin D Receptor Genotype Is Associated With Fat-Free Mass and Sarcopenia in Elderly Men**. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences* 2004, **59**(1):B10-B15.
60. Bischoff-Ferrari HA, Dawson-Hughes B, Willett WC, Staehelin HB, Bazemore MG, Zee RY, Wong JB: **Effect of Vitamin D on falls: a meta-analysis**. *JAMA* 2004, **291**(16):1999-2006.
61. Shrikrishna D, Patel M, Tanner RJ, Seymour JM, Connolly BA, Puthucheary ZA, Walsh SL, Bloch SA, Sidhu PS, Hart N *et al*: **Quadriceps wasting and physical inactivity in patients with COPD**. *Eur Respir J* 2012, **40**(5):1115-1122.
62. Alberts B: **Molecular biology of the cell**; 2015.
63. St Laurent G, Wahlestedt C, Kapranov P: **The Landscape of long noncoding RNA classification**. *Trends Genet* 2015, **31**(5):239-251.
64. Nechaev S, Adelman K: **Pol II waiting in the starting gates: Regulating the transition from transcription initiation into productive elongation**. *Biochim Biophys Acta* 2011, **1809**(1):34-45.
65. Lenhard B, Sandelin A, Carninci P: **Metazoan promoters: emerging characteristics and insights into transcriptional regulation**. *Nature Reviews Genetics* 2012, **13**(4):233-245.
66. Wahl MC, Will CL, Lührmann R: **The Spliceosome: Design Principles of a Dynamic RNP Machine**. *Cell* 2009, **136**(4):701-718.
67. Nilsen TW, Graveley BR: **Expansion of the eukaryotic proteome by alternative splicing**. *Nature* 2010, **463**(7280):457-463.
68. Shi Y: **Alternative polyadenylation: new insights from global analyses**. *RNA* 2012, **18**(12):2105-2117.
69. Catoire M, Mensink M, Boekschoten MV, Hangelbroek R, Muller M, Schrauwen P, Kersten S: **Pronounced effects of acute endurance exercise on gene expression in resting and exercising human skeletal muscle**. *PLoS One* 2012, **7**(11):e51066.
70. Hoheisel JD: **Microarray technology: beyond transcript profiling and genotype analysis**. *Nature Reviews Genetics* 2006, **7**(3):200-210.
71. Zhao S, Fung-Leung W-P, Bittner A, Ngo K, Liu X: **Comparison of RNA-Seq and microarray in transcriptome profiling of activated T cells**. *PloS one* 2014, **9**(1):e78644-e78644.
72. Brenner S, Johnson M, Bridgham J, Golda G, Lloyd DH, Johnson D, Luo S, McCurdy S, Foy M, Ewan M *et al*: **Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays**. *Nat Biotechnol* 2000, **18**(6):630-634.
73. Shiraki T, Kondo S, Katayama S, Waki K, Kasukawa T, Kawaji H, Kodzius R, Watahiki A, Nakamura M, Arakawa T *et al*: **Cap analysis gene expression for high-throughput analysis of transcriptional starting point and identification of promoter usage**. *Proc Natl Acad Sci U S A* 2003, **100**(26):15776-15781.
74. Velculescu VE, Zhang L, Vogelstein B, Kinzler KW: **Serial analysis of gene expression**. *Science* 1995, **270**(5235):484-487.

75. Sooknanan R, Pease J, Doyle K: **Novel methods for rRNA removal and directional, ligation-free RNA-seq library preparation**. *Nature Methods* 2010, **7**(10):i-ii.
76. Tariq MA, Kim HJ, Jejelowo O, Pourmand N: **Whole-transcriptome RNAseq analysis from minute amount of total RNA**. *Nucleic acids research* 2011, **39**(18):e120-e120.
77. Wilhelm BT, Landry J-R: **RNA-Seq—quantitative measurement of expression through massively parallel RNA-sequencing**. *Methods* 2009, **48**(3):249-257.
78. Zierath Juleen R, Wallberg-Henriksson H: **Looking Ahead Perspective: Where Will the Future of Exercise Biology Take Us?** *Cell Metabolism* 2015, **22**(1):25-30.
79. Molster CM, Bowman FL, Bilkey GA, Cho AS, Burns BL, Nowak KJ, Dawkins HJS: **The Evolution of Public Health Genomics: Exploring Its Past, Present, and Future**. *Front Public Health* 2018, **6**:247-247.
80. Pillon NJ, Gabriel BM, Dollet L, Smith JAB, Sardón Puig L, Botella J, Bishop DJ, Krook A, Zierath JR: **Transcriptomic profiling of skeletal muscle adaptations to exercise and inactivity**. *Nature communications* 2020, **11**(1):470-470.
81. Illumina: **An introduction to Next-GenerationSequencing Technology**.
82. Kircher M, Heyn P, Kelso J: **Addressing challenges in the production and analysis of illumina sequencing data**. *BMC genomics* 2011, **12**:382-382.
83. Bolger AM, Lohse M, Usadel B: **Trimmomatic: a flexible trimmer for Illumina sequence data**. *Bioinformatics (Oxford, England)* 2014, **30**(15):2114-2120.
84. Martin M: **Cutadapt removes adapter sequences from high-throughput sequencing reads**. *EMBnetjournal* 2011, **17**(1):10.
85. S A: **FastQC: A Quality Control Tool for High Throughput Sequence Data** 2010.
86. Ewels P, Magnusson M, Lundin S, Kaller M: **MultiQC: summarize analysis results for multiple tools and samples in a single report**. *Bioinformatics* 2016, **32**(19):3047-3048.
87. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR: **STAR: ultrafast universal RNA-seq aligner**. *Bioinformatics (Oxford, England)* 2013, **29**(1):15-21.
88. Kim D, Langmead B, Salzberg SL: **HISAT: a fast spliced aligner with low memory requirements**. *Nature methods* 2015, **12**(4):357-360.
89. 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**(8):907-915.
90. Li B, Dewey CN: **RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome**. *BMC Bioinformatics* 2011, **12**:323.
91. Bray NL, Pimentel H, Melsted P, Pachter L: **Near-optimal probabilistic RNA-seq quantification**. *Nature Biotechnology* 2016, **34**(5):525-527.
92. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C: **Salmon provides fast and bias-aware quantification of transcript expression**. *Nature methods* 2017, **14**(4):417-419.
93. Anders S, Pyl PT, Huber W: **HTSeq--a Python framework to work with high-throughput sequencing data**. *Bioinformatics (Oxford, England)* 2015, **31**(2):166-169.
94. Teng M, Love MI, Davis CA, Djebali S, Dobin A, Graveley BR, Li S, Mason CE, Olson S, Pervouchine D *et al*: **A benchmark for RNA-seq quantification pipelines**. *Genome Biol* 2016, **17**:74.
95. Anders S, Pyl PT, Huber W: **HTSeq - A Python framework to work with high-throughput sequencing data**. In.: Cold Spring Harbor Laboratory; 2014.
96. Li B, Ruotti V, Stewart RM, Thomson JA, Dewey CN: **RNA-Seq gene expression estimation with read mapping uncertainty**. *Bioinformatics (Oxford, England)* 2010, **26**(4):493-500.
97. Langmead B, Salzberg SL: **Fast gapped-read alignment with Bowtie 2**. *Nature methods* 2012, **9**(4):357-359.

98. Love MI, Huber W, Anders S: **Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2**. *Genome biology* 2014, **15**(12):550-550.
99. McCarthy DJ, Chen Y, Smyth GK: **Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation**. *Nucleic Acids Res* 2012, **40**(10):4288-4297.
100. Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B: **Mapping and quantifying mammalian transcriptomes by RNA-Seq**. *Nat Methods* 2008, **5**(7):621-628.
101. Robinson MD, McCarthy DJ, Smyth GK: **edgeR: a Bioconductor package for differential expression analysis of digital gene expression data**. *Bioinformatics (Oxford, England)* 2010, **26**(1):139-140.
102. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, Salzberg SL, Wold BJ, Pachter L: **Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation**. *Nat Biotechnol* 2010, **28**(5):511-515.
103. Jiang L, Schlesinger F, Davis CA, Zhang Y, Li R, Salit M, Gingeras TR, Oliver B: **Synthetic spike-in standards for RNA-seq experiments**. *Genome Res* 2011, **21**(9):1543-1551.
104. Risso D, Ngai J, Speed TP, Dudoit S: **Normalization of RNA-seq data using factor analysis of control genes or samples**. *Nature biotechnology* 2014, **32**(9):896-902.
105. Lin CY, Loven J, Rahl PB, Paranal RM, Burge CB, Bradner JE, Lee TI, Young RA: **Transcriptional amplification in tumor cells with elevated c-Myc**. *Cell* 2012, **151**(1):56-67.
106. Alessandra L, Luca P, Adriano M: **Differential gene expression in kernels and silks of maize lines with contrasting levels of ear rot resistance after Fusarium verticillioides infection**. *Journal of Plant Physiology* 2010, **167**(16):1398-1406.
107. Le DT, Aldrich DL, Valliyodan B, Watanabe Y, Ha CV, Nishiyama R, Guttikonda SK, Quach TN, Gutierrez-Gonzalez JJ, Tran LS *et al*: **Evaluation of candidate reference genes for normalization of quantitative RT-PCR in soybean tissues under various abiotic stress conditions**. *PLoS One* 2012, **7**(9):e46487.
108. Müller OA, Grau J, Thieme S, Prochaska H, Adlung N, Sorgatz A, Bonas U: **Genome-Wide Identification and Validation of Reference Genes in Infected Tomato Leaves for Quantitative RT-PCR Analyses**. *PloS one* 2015, **10**(8):e0136499-e0136499.
109. Wan Q, Chen S, Shan Z, Yang Z, Chen L, Zhang C, Yuan S, Hao Q, Zhang X, Qiu D *et al*: **Stability evaluation of reference genes for gene expression analysis by RT-qPCR in soybean under different conditions**. *PloS one* 2017, **12**(12):e0189405-e0189405.
110. Chen CM, Lu YL, Sio CP, Wu GC, Tzou WS, Pai TW: **Gene Ontology based housekeeping gene selection for RNA-seq normalization**. *Methods* 2014, **67**(3):354-363.
111. Wang Z, Lyu Z, Pan L, Zeng G, Randhawa P: **Defining housekeeping genes suitable for RNA-seq analysis of the human allograft kidney biopsy tissue**. *BMC Med Genomics* 2019, **12**(1):86.
112. 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**. *Statistical Applications in Genetics and Molecular Biology* 2016, **15**(2).
113. Yu G, Wang L-G, Han Y, He Q-Y: **clusterProfiler: an R package for comparing biological themes among gene clusters**. *OMICS* 2012, **16**(5):284-287.
114. Gene Ontology C: **The Gene Ontology project in 2008**. *Nucleic acids research* 2008, **36**(Database issue):D440-D444.
115. Chapman MA, Arif M, Emanuelsson EB, Reitzner SM, Lindholm ME, Mardinoglu A, Sundberg CJ: **Skeletal Muscle Transcriptomic Comparison between Long-Term Trained and Untrained Men and Women**. *Cell Rep* 2020, **31**(12):107808.

116. Popov DV, Makhnovskii PA, Shagimardanova EI, Gazizova GR, Lysenko EA, Gusev OA, Vinogradova OL: **Contractile activity-specific transcriptome response to acute endurance exercise and training in human skeletal muscle**. *Am J Physiol Endocrinol Metab* 2019, **316**(4):E605-E614.
117. Hammarstrom D, Ofsteng S, Koll L, Hanestadhaugen M, Hollan I, Apro W, Whist JE, Blomstrand E, Ronnestad BR, Ellefsen S: **Benefits of higher resistance-training volume are related to ribosome biogenesis**. *J Physiol* 2020, **598**(3):543-565.
118. Khan Y, Hammarstrom D, Ronnestad BR, Ellefsen S, Ahmad R: **Increased biological relevance of transcriptome analyses in human skeletal muscle using a model-specific pipeline**. *BMC Bioinformatics* 2020, **21**(1):548.
119. Molmen KS, Hammarstrom D, Falch GS, Grundtvig M, Koll L, Hanestadhaugen M, Khan Y, Ahmad R, Malerbakken B, Rodolen TJ *et al*: **Chronic obstructive pulmonary disease does not impair responses to resistance training**. *J Transl Med* 2021, **19**(1):292.
120. Ellefsen S, Stensløkken K-O, Sandvik GK, Kristensen TA, Nilsson GE: **Improved normalization of real-time reverse transcriptase polymerase chain reaction data using an external RNA control**. *Analytical Biochemistry* 2008, **376**(1):83-93.
121. Ellefsen S, Vikmoen O, Slettaløkken G, Whist JE, Nygaard H, Hollan I, Rauk I, Vegge G, Strand TA, Raastad T *et al*: **Irisin and FNDC5: effects of 12-week strength training, and relations to muscle phenotype and body mass composition in untrained women**. *European Journal of Applied Physiology* 2014, **114**(9):1875-1888.
122. Bolger AM, Lohse M, Usadel B: **Trimmomatic: a flexible trimmer for Illumina sequence data**. *Bioinformatics* 2014, **30**(15):2114-2120.
123. Kim MR, Jeong SJ: **Relationship between Vitamin D Level and Lipid Profile in Non-Obese Children**. *Metabolites* 2019, **9**(7).
124. Oshlack A, Wakefield MJ: **Transcript length bias in RNA-seq data confounds systems biology**. *Biol Direct* 2009, **4**:14.
125. Robinson MD, Oshlack A: **A scaling normalization method for differential expression analysis of RNA-seq data**. *Genome Biol* 2010, **11**(3):R25.
126. Rudnick PA, Wang X, Yan X, Sedransk N, Stein SE: **Improved normalization of systematic biases affecting ion current measurements in label-free proteomics data**. *Mol Cell Proteomics* 2014, **13**(5):1341-1351.
127. Prjibelski A, Antipov D, Meleshko D, Lapidus A, Korobeynikov A: **Using SPAdes De Novo Assembler**. *Curr Protoc Bioinformatics* 2020, **70**(1):e102.
128. Yamaguchi KD, Ruderman DL, Croze E, Wagner TC, Velichko S, Reder AT, Salamon H: **IFN-β-regulated genes show abnormal expression in therapy-naïve relapsing–remitting MS mononuclear cells: Gene expression analysis employing all reported protein–protein interactions**. *Journal of Neuroimmunology* 2008, **195**(1-2):116-120.
129. 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 (Oxford, England)* 2019, **35**(24):5146-5154.
130. Zyla J, Marczyk M, Weiner J, Polanska J: **Ranking metrics in gene set enrichment analysis: do they matter?** *BMC bioinformatics* 2017, **18**(1):256-256.
131. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES *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**(43):15545-15550.
132. Korotkevich G, Sukhov V, Budin N, Shpak B, Artyomov MN, Sergushichev A: **Fast gene set enrichment analysis**. In.: Cold Spring Harbor Laboratory; 2016.

133. Stark R, Grzelak M, Hadfield J: **RNA sequencing: the teenage years**. *Nat Rev Genet* 2019, **20**(11):631-656.
134. 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**(12):e85024.
135. Liao Y, Shi W: **Read trimming is not required for mapping and quantification of RNA-seq reads**. In.: Cold Spring Harbor Laboratory; 2019.
136. Ballouz S, Dobin A, Gingeras TR, Gillis J: **The fractured landscape of RNA-seq alignment: the default in our STARs**. *Nucleic Acids Res* 2018, **46**(10):5125-5138.
137. Nie Z, Hu G, Wei G, Cui K, Yamane A, Resch W, Wang R, Green DR, Tessarollo L, Casellas R *et al*: **c-Myc is a universal amplifier of expressed genes in lymphocytes and embryonic stem cells**. *Cell* 2012, **151**(1):68-79.
138. Bullard JH, Purdom E, Hansen KD, Dudoit S: **Evaluation of statistical methods for normalization and differential expression in mRNA-Seq experiments**. *BMC bioinformatics* 2010, **11**:94-94.
139. Dillies MA, Rau A, Aubert J, Hennequet-Antier C, Jeanmougin M, Servant N, Keime C, Marot G, Castel D, Estelle J *et al*: **A comprehensive evaluation of normalization methods for Illumina high-throughput RNA sequencing data analysis**. *Briefings in Bioinformatics* 2012, **14**(6):671-683.
140. Zyprych-Walczak J, Szabelska A, Handschuh L, Gorczak K, Klamecka K, Figlerowicz M, Siatkowski I: **The Impact of Normalization Methods on RNA-Seq Data Analysis**. *Biomed Res Int* 2015, **2015**:621690.
141. Jureckova K, Raschmanova H, Kolek J, Vasylkivska M, Branska B, Patakova P, Provaznik I, Sedlar K: **Identification and Validation of Reference Genes in Clostridium beijerinckii NRRL B-598 for RT-qPCR Using RNA-Seq Data**. *Front Microbiol* 2021, **12**:640054.
142. Li Y, Zhang L, Li R, Zhang M, Li Y, Wang H, Wang S, Bao Z: **Systematic identification and validation of the reference genes from 60 RNA-Seq libraries in the scallop Mizuhopecten yessoensis**. *BMC Genomics* 2019, **20**(1):288.
143. Liang L, He Z, Yu H, Wang E, Zhang X, Zhang B, Zhang C, Liang Z: **Selection and Validation of Reference Genes for Gene Expression Studies in Codonopsis pilosula Based on Transcriptome Sequence Data**. *Sci Rep* 2020, **10**(1):1362.
144. Zhou Z, Cong P, Tian Y, Zhu Y: **Using RNA-seq data to select reference genes for normalizing gene expression in apple roots**. *PLoS One* 2017, **12**(9):e0185288.
145. Ching T, Huang S, Garmire LX: **Power analysis and sample size estimation for RNA-Seq differential expression**. *RNA* 2014, **20**(11):1684-1696.
146. Li CI, Su PF, Shyr Y: **Sample size calculation based on exact test for assessing differential expression analysis in RNA-seq data**. *BMC Bioinformatics* 2013, **14**:357.
147. Yu L, Fernandez S, Brock G: **Power analysis for RNA-Seq differential expression studies using generalized linear mixed effects models**. *BMC bioinformatics* 2020, **21**(1):198-198.
148. Gumpenberger M, Wessner B, Graf A, Narici MV, Fink C, Braun S, Hoser C, Blazevich AJ, Csapo R: **Remodeling the Skeletal Muscle Extracellular Matrix in Older Age-Effects of Acute Exercise Stimuli on Gene Expression**. *Int J Mol Sci* 2020, **21**(19):7089.
149. Damas F, Ugrinowitsch C, Libardi CA, Jannig PR, Hector AJ, McGlory C, Lixandrao ME, Vechin FC, Montenegro H, Tricoli V *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**(12):2607-2616.
150. Hjorth M, Norheim F, Meen AJ, Pourteymour S, Lee S, Holen T, Jensen J, Birkeland KI, Martinov VN, Langleite TM *et al*: **The effect of acute and long-term physical activity on extracellular matrix and serglycin in human skeletal muscle**. *Physiol Rep* 2015, **3**(8).

151. Holm L, van Hall G, Rose AJ, Miller BF, Doessing S, Richter EA, Kjaer M: **Contraction intensity and feeding affect collagen and myofibrillar protein synthesis rates differently in human skeletal muscle**. *Am J Physiol Endocrinol Metab* 2010, **298**(2):E257-269.
152. Mackey AL, Brandstetter S, Schjerling P, Bojsen-Moller J, Qvortrup K, Pedersen MM, Doessing S, Kjaer M, Magnusson SP, Langberg H: **Sequenced response of extracellular matrix deadhesion and fibrotic regulators after muscle damage is involved in protection against future injury in human skeletal muscle**. *FASEB J* 2011, **25**(6):1943-1959.
153. Makhnovskii PA, Zgoda VG, Bokov RO, Shagimardanova EI, Gazizova GR, Gusev OA, Lysenko EA, Kolpakov FA, Vinogradova OL, Popov DV: **Regulation of Proteins in Human Skeletal Muscle: The Role of Transcription**. *Sci Rep* 2020, **10**(1):3514.
154. 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**(6):E1153-1159.
155. Raue U, Trappe TA, Estrem ST, Qian HR, Helvering LM, Smith RC, Trappe S: **Transcriptome signature of resistance exercise adaptations: mixed muscle and fiber type specific profiles in young and old adults**. *J Appl Physiol (1985)* 2012, **112**(10):1625-1636.
156. 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**(3):354-364.
157. Pillon NJ, Gabriel BM, Dollet L, Smith JAB, Sardon Puig L, Botella J, Bishop DJ, Krook A, Zierath JR: **Transcriptomic profiling of skeletal muscle adaptations to exercise and inactivity**. *Nat Commun* 2020, **11**(1):470.
158. Egan B, O'Connor PL, Zierath JR, O'Gorman DJ: **Time course analysis reveals gene-specific transcript and protein kinetics of adaptation to short-term aerobic exercise training in human skeletal muscle**. *PLoS One* 2013, **8**(9):e74098.
159. Zosky GR, Berry LJ, Elliot JG, James AL, Gorman S, Hart PH: **Vitamin D deficiency causes deficits in lung function and alters lung structure**. *Am J Respir Crit Care Med* 2011, **183**(10):1336-1343.
160. Pitta F, Troosters T, Spruit MA, Probst VS, Decramer M, Gosselink R: **Characteristics of physical activities in daily life in chronic obstructive pulmonary disease**. *Am J Respir Crit Care Med* 2005, **171**(9):972-977.

Papers

1

BMC Bioinformatics

RESEARCH ARTICLE

Open Access

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. Rønnestad[2], Stian Ellefsen[2,4] and Rafi Ahmad[1,5*]

*Correspondence: rafi.ahmad@inn.no
†Yusuf Khan and Daniel Hammarström have contributed equally to this work
[1] Department of Biotechnology, Inland Norway University of Applied Sciences, Holsetgata 22, 2317 Hamar, Norway
Full list of author information is available at the end of the article

Abstract

Background: Human skeletal muscle responds to weight-bearing exercise with significant inter-individual differences. Investigation of transcriptome responses could improve our understanding of this variation. However, this requires bioinformatic pipelines to be established and evaluated in study-specific contexts. Skeletal muscle subjected to mechanical stress, such as through resistance training (RT), accumulates RNA due to increased ribosomal biogenesis. When a fixed amount of total-RNA is used for RNA-seq library preparations, mRNA counts are thus assessed in different amounts of tissue, potentially invalidating subsequent conclusions. The purpose of this study was to establish a bioinformatic pipeline specific for analysis of RNA-seq data from skeletal muscles, to explore the effects of different normalization strategies and to identify genes responding to RT in a volume-dependent manner (moderate vs. low volume). To this end, we analyzed RNA-seq data derived from a twelve-week RT intervention, wherein 25 participants performed both low- and moderate-volume leg RT, allocated to the two legs in a randomized manner. Bilateral muscle biopsies were sampled from *m. vastus lateralis* before and after the intervention, as well as before and after the fifth training session (Week 2).

Result: Bioinformatic tools were selected based on read quality, observed gene counts, methodological variation between paired observations, and correlations between mRNA abundance and protein expression of myosin heavy chain family proteins. Different normalization strategies were compared to account for global changes in RNA to tissue ratio. After accounting for the amounts of muscle tissue used in library preparation, global mRNA expression increased by 43–53%. At Week 2, this was accompanied by dose-dependent increases for 21 genes in rested-state muscle, most of which were related to the extracellular matrix. In contrast, at Week 12, no readily explainable dose-dependencies were observed. Instead, traditional normalization and non-normalized models resulted in counterintuitive reverse dose-dependency for many genes. Overall, training led to robust transcriptome changes, with the number of differentially expressed genes ranging from 603 to 5110, varying with time point and normalization strategy.

Conclusion: Optimized selection of bioinformatic tools increases the biological relevance of transcriptome analyses from resistance-trained skeletal muscle. Moreover,

normalization procedures need to account for global changes in rRNA and mRNA abundance.

Keywords: RNA-seq, Skeletal muscle, Bioinformatics pipeline, Normalization

Background

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 beneficial outcomes of such interventions for muscle function and overall health [4, 5]. Currently, little is known about the etiology of this variation. However, it is usually associated with phenotypic traits of skeletal muscle [6–8], which implies interactions with environmental factors, genetics, epigenetics, and composites of the intra 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 PGC1α [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. First, biologically founded variability can be attributed to differences in exercise protocols (e.g., differences in exercise-volume or intensity). This makes it difficult to discern a general transcriptome exercise response, as training variables are not standardized between studies. Biological heterogeneity is also caused by differences between research participants, affecting signal-to-noise ratios and making it difficult to discern the effects of single independent factors such as training variables. Design stage decisions such as the use of within-participant designs [3, 23] are likely to reduce this variation and to provide transcriptome data with increased biological meaningfulness. Second, technical variability can be attributed to decisions made during the bioinformatical treatment of data. As described by Concea et al. [24], there is no optimal pipeline for sequencing technology as new tools keep evolving and emerging, different tools should be explored to an optimum pipeline for the specific type of data. To exploit the potential of any study design, there is a need for identifying an appropriate pipeline for transcriptome analyses to ensure a biologically valid interpretation of data. This entails identifying potential violations of common assumptions caused by the experimental model at hand, relating to, for example, data normalization [25, 26].

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 [24, 27]. Of these steps, data normalization is particularly decisive [26], as it aims to transform naïve transcript counts into biologically meaningful results. This essentially means expressing them as *per*-cell abundances [28]. For most experimental models, this is equivalent to providing transcript-to-total RNA ratios, given the fulfillment of the assumption that total RNA levels remain stable between conditions on a per-unit-cell or per-unit-tissue basis [28]. In cell models that exhibit high degrees of plasticity, gene expression events result in increased amounts of total RNA and mRNA transcripts per cell [29], specifically violating the assumption that most genes are not differentially expressed [25, 28]. We are not aware of any study that has addressed the need to account for such perspectives during transcriptome analyses of skeletal muscle subjected to mechanical stress, such as resistance training. Indeed, this assumption can be expected to be violated, as total RNA content increases markedly on a per-unit-weight basis [3], with potential global changes also occurring for the mRNA pool, though this remains unknown. 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 (1) establish a bioinformatic pipeline specific for analysis of RNA-seq data from skeletal muscles, (2) explore the effects of using different normalization strategies for analyzing skeletal muscle tissue subjected to resistance training, and (3) identify genes responding to moderate versus low resistance exercise volume, while simultaneously identifying genes whose expression changes with time. To achieve these aims, we utilized RNA-seq data generated from a within-participant study, comparing the effects of low and moderate resistance training volume, as previously described [3]. Also, myosin heavy chain protein expression, quantified using immunohistochemistry, was used to validate RNA quantification tools.

Results

For the RNA-seq analyses presented here, a subset of participants was selected based on RNA quality measurements from a previously reported study comparing the functional and biological efficacy of low- and moderate-volume resistance training [3] (Fig. 1a). Twenty-five participants (out of 34) had a full set of RNA-samples with RNA quality indicator (RQI) scores ≥ 7, which were subjected to bioinformatic data analysis (Fig. 1b). RQI scores were not associated with RNA yield (Fig. 1c). In these participants, twelve weeks of training with moderate volume led to greater increases in limb lean-mass than low volume (3.5% vs. 2.0%, pre-training MOD mean (SD) 8.9 (2.2), to post-training 9.2 (2.3) kg; pre-training LOW 8.9 (2.2), to post-training 9.0 (2.2) kg, Fig. 1d), corresponding well with MRI-based muscle cross-sectional area data from the full data set [3]. Similar responses were seen in the excluded participants (Fig. 1d). This coincided with greater strength gains (~ 25% vs. ~ 19%, Fig. 1f), which also agrees with observations made in the full cohort, accompanied by greater changes in muscle fiber proportions (type IIX fibers ↓) [3].

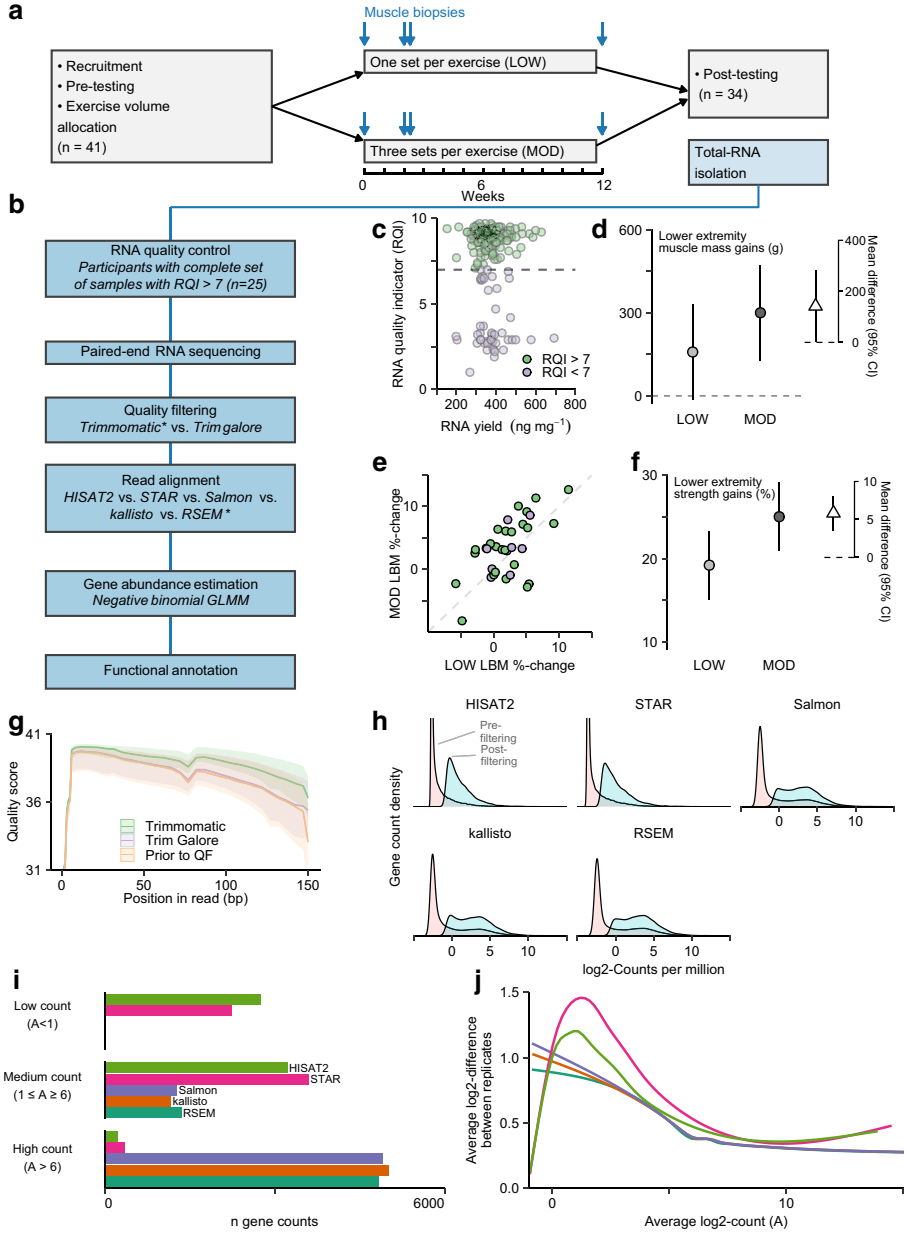


Fig. 1 Study overview and RNA-seq analysis pipeline. Forty-one participants performed twelve weeks of resistance training with low- (one set per exercise, LOW) and moderate-volume (three sets per exercise, MOD) in a contralateral manner (2–3 sessions week-1) (**a**). Pre- and post-training testing included strength and muscle lean-mass assessments. Muscle biopsies were collected from *m. vastus lateralis* 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 participants with RNA quality indicator (RQI) scores > 7 were submitted for RNA-seq ($n = 25$). RNA quality was not associated with muscle tissue weight (**c**), and participants included in RNA-seq experiments did not differ from excluded in terms of limb lean-mass gains (**d**). Higher training volume led to greater gains in limb lean mass (**e**) and strength (**f**) in the lower extremities ($n = 25$). RNA-seq data were 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**)

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

To select the most appropriate tools for bioinformatic analyses, we first compared quality filtering using Trimmomatic and Trim Galore, both of which are commonly used [30, 31]. Quality scores were generally better with Trimmomatic (Fig. 1g). Trim Galore did not improve scores over non-filtered data (Fig. 1g). Filtered reads were then aligned to the human genome using three alignment-based methods (including HISAT2, STAR, RSEM, all used together 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, resulting in the expected bimodal distribution of counts and a larger subset of detected genes after expression filtering compared to STAR and HISAT2 (Fig. 1h). For a selection of genes with known robust expression across tissues [32], Salmon, kallisto, and RSEM resulted in higher proportions of genes with high count numbers (Fig. 1i). RSEM was found to be associated with lower technical variation than Salmon, kallisto, HISAT2, and STAR expressed as log-differences in the expression of these genes between bilateral biopsies sampled prior to the intervention (Fig. 1j). For HISAT2 and STAR, this distorted correlations between RNA-seq based myosin heavy chain family mRNA and myosin heavy chain protein profiles (Fig. 2a and b), with the latter identified in Hammarström et al. [3], which are established hallmarks of skeletal muscle biology [33–35]. Overall, RSEM, kallisto, and Salmon thus displayed superior technical performance than HISAT2 and STAR, resulting in data with lower degrees of technical variation and higher degrees of biological validity. RSEM displayed slightly lower average variation between paired samples and was thus chosen for downstream analyses.

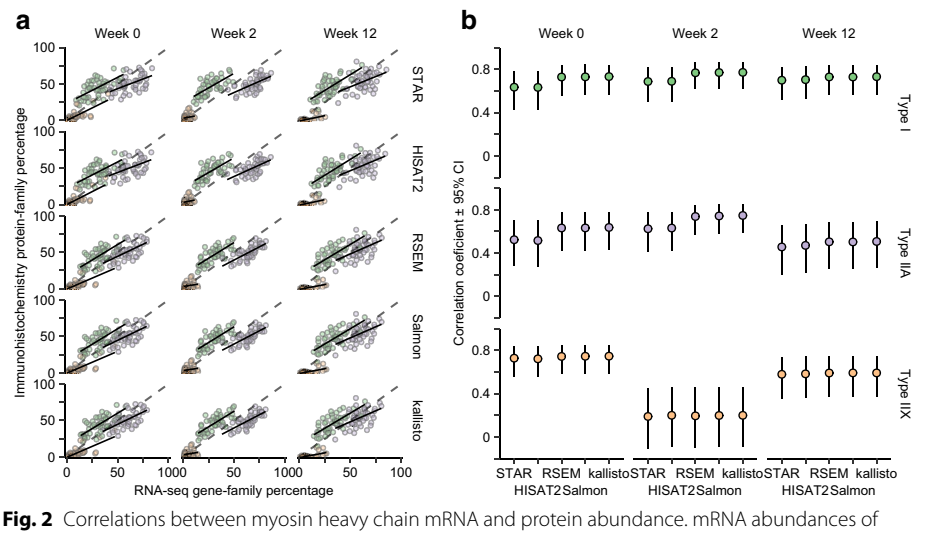


Fig. 2 Correlations between myosin heavy chain mRNA and protein abundance. mRNA abundances of myosin-heavy chains in *m. vastus lateralis* estimated using RSEM, kallisto, and Salmon showed stronger correlations with immunohistochemistry-determined protein expression than HISAT2 and STAR (**a**, **b**). mRNA and protein abundances of *MYH7*/Type I, *MYH2*/Type IIA, and *MYH1*/Type IIX were calculated as percentages of overall myosin-heavy chain mRNA and protein expression, analyses unbiased by normalization [34, 46]

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

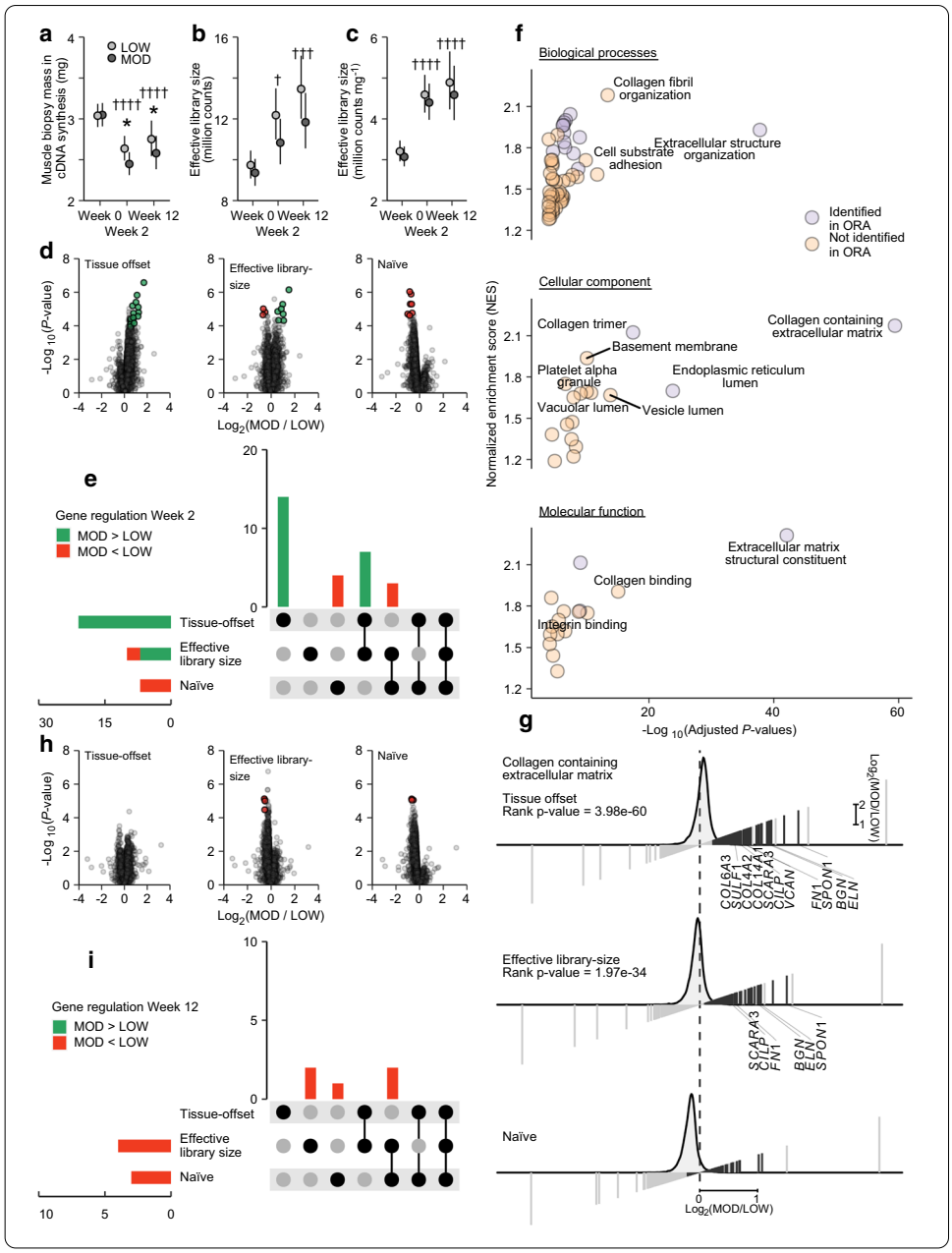
In the selected participants, similarly to what has been reported in the full cohort [3], resistance training led to an increase in total RNA per-unit tissue weight that was larger in response to moderate- versus low-volume training (Week 2, low 15% vs. moderate-volume 24%, mean difference 7.7%, 95% CI [1.1, 14.8]; Week 12, low 15% vs. moderate-volume 24%, mean difference 7.7%, 95% CI [1.1, 14.8]). As equal amounts of total RNA were used for preparing RNA-seq libraries, the amounts of muscle tissue used for library preparations decreased from baseline to Week 2 and 12 in both legs (low volume, − 13% and − 9%; moderate volume − 20% and − 15%). This decrease was subsequently more pronounced in the moderate volume condition, resulting in lower amounts of tissue used in cDNA synthesis (− 7.1%, 95% CI [− 12.9, − 1.0]; − 6.3%, [− 11.8, − 0.4]; Fig. 3a). Despite the utilization of less muscle tissue during library preparations in the trained state, effective library sizes increased compared to baseline levels (low volume, 25%, and 38% at Week 2 and 12, respectively; moderate volume, 16%, and 26%; Fig. 3b). Initially, this increase was numerically less pronounced in the moderate volume condition (− 11%, [− 22, 1.7]; − 12%, [− 24, 2.2]; Fig. 3b), but after normalization to tissue weight, the two training modalities resulted in similar increases in effective library size (low volume, 43%, and 53% at Week 2 and 12, respectively; moderate volume, 43%, and 49%; Fig. 3c). Overall, this suggests marked increases in global mRNA expression in response to both low- and moderate-volume resistance training.

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

During subsequent identification of differentially expressed (DE) genes in response to low- and moderate-volume resistance training, three normalization models were used and compared. The first model contained effective library sizes as a covariate, as

(See figure on next page.)

Fig. 3 Global mRNA expression and transcriptome profiles in response to low and moderate volume resistance training. The amounts of muscle tissue used during cDNA synthesis varied over the course of the study and between volume conditions (**a** low-volume, LOW; moderate-volume, MOD). Library sizes increased during the course of the intervention, with a tendency towards a greater increase in the low-volume condition (**b**). Difference in library sizes between volume conditions when expressed per-unit tissue weight were diminished, though increases from baseline were maintained (**c**). The tissue offset-normalized model identified 21 genes with higher expression in the moderate volume condition (**d**, **e**), ten of which was shared with the effective library-size normalized model at week 2 (**e**), and none of which was shared with the naïve model. No volume-dependent differences were found at Week 12 using the tissue-offset model. At this time point, library-size and naïve models both showed a marked skew towards augmented expression in the low-volume condition. At Week 2, functional annotation identified gene sets relating to extracellular matrix in response to higher training volume (tissue-offset model, orange and purple circles, **f**), all of which were more highly expressed in MOD, indicated by the positive enrichment score. Orange circles denote gene sets that were identified from rank-based enrichment tests based on the full data set. Purple circles denote gene categories that were also identified using over-representation analysis (ORA). Normalization strategies had global effects on enrichment analyses using rank tests, assessed using fold-changes and minimum significant differences scores (not shown), illustrated with the tissue-offset model leading to marked increases in genes associated with the "Collagen containing extracellular matrix" gene set (**g**) as well as a shift in the full distribution of Log2 fold-changes between volume conditions towards MOD (shown as density curves). Black bars represent genes that belong to the gene set identified as enriched (**g**). Genes symbols indicate genes identified as differentially expressed in each normalization scenario



previously suggested [36], while also containing tissue weight as an offset to account for amounts of tissue used during RNA-seq library preparation (tissue-offset model). The second model contained effective library sizes as a covariate only (library-size normalization), thus representing an effort to compare expression levels across training modalities while accounting for technical variation during library preparation [25, 36]. The third model was a non-normalized model (naïve model, included for comparison) (Table 1).

At Week 2, 21 genes were identified as DE between low and moderate volume using the tissue-offset model, with all genes showing higher expression in the moderate volume condition (Fig. 3d; rested-state biopsies sampled after four training sessions).

Table 1 Participant characteristics

			Mean	SD
Female	n = 11	Age (years)	22.6	0.9
		Body mass (kg)	166.2	6.2
		Stature (cm)	61.5	7.4
		Lean mass (%)	63.7	5.6
		Fat mass (%)	32.2	5.7
Male	n = 14	Age (years)	23.9	4.2
		Body mass (kg)	183.7	5.6
		Stature (cm)	77.4	10.4
		Lean mass (%)	75.4	5.5
		Fat mass (%)	20.1	5.7

Values from pre-intervention assessments. Relative lean and fat mass from whole-body data

Similarly, 10 and seven genes were identified as DE between volume conditions using the library-size model and the naïve model, respectively (Fig. 3d and e). For the library-size model, seven of these DE genes showed higher expression in the moderate-volume condition, overlapping completely with genes found using the tissue-offset model (Fig. 3e). For the naïve model, the seven DE genes all showed decreased expression in the moderate volume condition, with 3 genes still overlapping with the library-size model, thus resulting in a contra-intuitive decrease in expression. Using tissue-offset model-derived estimates for functional analyses (Rank-based enrichment tests of minimum significant differences, MSD), revealed enrichment of genes associated with extracellular matrix (ECM) gene ontology (GO) sets (Fig. 3f, Table 2). The top-ranked GO terms were also identified by over-representation tests (ORA) using DE genes (Fig. 3f, Table 2, detailed table in Additional files 1 and 2). Using library-size model estimates similar top-ranked GO sets were identified as with the tissue-offset model albeit with lower levels of significance and lower degrees of agreement between methods (Table 2). The naïve model generally identified GO sets with negative enrichment scores indicating gene sets with lower expression in moderate volume compared to low volume, with a weak agreement between enrichment methods (Table 2). The analytical consequences of using the different normalization strategies were particularly apparent in comparisons of rank metrics, such as fold-changes. Importantly, this analytical approach uses the entire gene set to identify enriched gene sets, rather than being confined to DE genes. After controlling for amounts of tissue used during preparation of RNA-seq libraries, the distribution of Log2 differences between volume conditions shifted markedly in favor of higher training volume (Fig. 3g), and robust gene sets appeared with higher expression in the moderate volume condition, such as genes belonging to the Collagen-containing ECM GO set (Fig. 3g). Accordingly, the number of DE genes identified to this GO set was highest using the tissue-offset model ($n = 11$), followed by the library-size model ($n = 6$), with no genes identified using the naïve model (Fig. 3g).

At Week 12, no genes were identified as differentially expressed between resistance training with low and moderate volume using the tissue-offset model (Fig. 3h; rested-state biopsies sampled after finalization of the intervention). In contrast, a small number of genes were identified as DE between volume conditions using library-size and naïve models (n = 4 and n = 3, respectively; Fig. 3h and i), with all genes showing lower

Table 2 Functional annotation analysis comparing moderate- and low-volume training

Comparison	Normalization model	Gene ontology category	ID	Description	Rank *P*-value[a]	Gene-set enrichment analysis (GSEA)		ORA *P*-value[c]
						GSEA *P*-value[b]	NES	
Week 2 MOD versus LOW	Tissue offset	Biological process	GO:0043062	Extracellular structure organization	6.19e−39	6.22e−24	1.93	4.70e−06
			GO:0030199	Collagen fibril organization	1.46e−14	4.72e−11	2.22	NA
		Cellular component	GO:0062023	Collagen containing extracellular matrix	1.39e−60	7.90e−44	2.17	6.63e−12
			GO:0031012	Extracellular matrix	1.01e−58	6.10e−44	2.08	2.88e−11
			GO:0005788	Endoplasmic reticulum lumen	3.58e−25	7.82e−11	1.70	7.80e−07
			GO:0005581	Collagen trimer	9.38e−19	7.20e−12	2.14	3.89e−05
			GO:0031983	Vesicle lumen	4.15e−15	1.52e−10	1.67	NA
		Molecular function	GO:0005201	Extracellular matrix structural constituent	2.23e−43	3.67e−29	2.28	2.86e−11
			GO:0005198	Structural molecule activity	9.12e−31	3.39e−14	1.58	NA
			GO:0005518	Collagen binding	1.49e−16	5.96e−08	1.97	NA
	Effective library size	Biological process	GO:0043062	Extracellular structure organization	7.50e−21	2.68e−27	2.63	NA
			GO:0030199	Collagen fibril organization	1.56e−12	3.25e−08	2.46	NA
		Cellular component	GO:0062023	Collagen containing extracellular matrix	6.85e−35	4.30e−45	2.89	7.93e−08
			GO:0031012	Extracellular matrix	1.77e−32	3.26e−41	2.73	1.29e−07
			GO:0005788	Endoplasmic reticulum lumen	3.16e−14	1.56e−13	2.28	NA
			GO:0005581	Collagen trimer	1.88e−13	7.84e−11	2.50	NA
		Molecular function	GO:0005201	Extracellular matrix structural constituent	2.67e−31	5.53e−26	2.83	1.67e−07
			GO:0005198	Structural molecule activity	1.21e−18	7.39e−21	2.22	NA
			GO:0005518	Collagen binding	2.24e−10	1.83e−08	2.37	NA
			GO:0030020	Extracellular matrix structural constituent conferring tensile strength	8.51e−09	2.52e−06	2.26	NA
	Naïve	Biological process	GO:0006397	mRNA processing	1.76e−17	5.03e−04	−1.48	NA
			GO:0008380	RNA splicing	3.19e−17	0.002	−1.47	NA
			GO:0000375	RNA splicing via transesterification reactions	3.41e−15	0.005	−1.49	NA

Table 2 (continued)

Comparison	Normalization model	Gene ontology category	ID	Description	Rank *P*-value[a]	Gene-set enrichment analysis (GSEA)		ORA *P*-value[c]
						GSEA *P*-value[b]	NES	
			GO:1903311	Regulation of mRNA metabolic process	7.48e−10	0.003	−1.51	NA
			GO:0050684	Regulation of mRNA processing	1.18e−07	0.015	−1.64	NA
			GO:0043484	Regulation of RNA splicing	1.76e−07	0.049	−1.57	NA
			GO:0048024	Regulation of mRNA splicing via spliceosome	3.48e−07	0.009	−1.75	NA
			GO:0000380	Alternative mRNA splicing via spliceosome	8.32e−07	0.027	−1.74	NA
		Cellular component	GO:0005681	Spliceosomal complex	2.65e−11	0.007	−1.56	NA
			GO:0016607	Nuclear speck	3.10e−08	0.010	−1.40	NA
Week 12 MOD versus LOW	Tissue offset	Biological process	GO:0010498	Proteasomal protein catabolic process	0.046	0.685	1.00	NA
			GO:0006401	RNA catabolic process	0.046	0.737	−0.87	NA
			GO:0006397	mRNA processing	0.046	0.904	0.88	NA
			GO:0000209	Protein polyubiquitination	0.046	0.579	1.05	NA
		Molecular function	GO:0003729	mRNA binding	0.003	0.844	0.99	NA
			GO:0019783	Ubiquitin like protein specific protease activity	0.021	0.286	1.37	NA
			GO:0019787	Ubiquitin like protein transferase activity	0.021	0.796	1.01	NA
			GO:0008234	Cysteine type peptidase activity	0.021	0.247	1.40	NA
			GO:0016874	Ligase activity	0.039	0.691	1.10	NA
			GO:0003730	mRNA 3 utr binding	0.043	0.775	1.01	NA
	Effective library size	Biological process	GO:0006613	Cotranslational protein targeting to membrane	1.18e−36	1.29e−07	−2.11	NA
			GO:0072599	Establishment of protein localization to endoplasmic reticulum	3.52e−36	6.44e−09	−2.12	NA
			GO:0070972	Protein localization to endoplasmic reticulum	1.52e−33	2.30e−08	−2.04	NA
			GO:0019080	Viral gene expression	1.14e−30	1.75e−07	−1.86	NA

Table 2 (continued)

Comparison	Normalization model	Gene ontology category	ID	Description	Rank *P*-value[a]	Gene-set enrichment analysis (GSEA)		ORA *P*-value[c]
						GSEA *P*-value[b]	NES	
		Cellular component	GO:0005840	Ribosome	6.28e−42	6.81e−12	−2.01	NA
			GO:0044391	Ribosomal subunit	6.74e−42	4.05e−12	−2.09	NA
			GO:0022626	Cytosolic ribosome	3.40e−36	2.36e−09	−2.15	NA
			GO:0098798	Mitochondrial protein complex	1.20e−35	7.52e−10	−1.89	NA
			GO:0019866	Organelle inner membrane	2.88e−35	7.52e−10	−1.67	NA
		Molecular function	GO:0003735	Structural constituent of ribosome	7.79e−43	7.81e−12	−2.13	NA
	Naïve	Biological process	GO:0006613	Cotranslational protein targeting to membrane	7.73e−30	3.48e−08	−2.14	NA
			GO:0072599	Establishment of protein localization to endoplasmic reticulum	7.73e−30	1.65e−08	−2.10	NA
			GO:0070972	Protein localization to endoplasmic reticulum	7.58e−28	6.14e−07	−1.92	NA
		Cellular component	GO:0019866	Organelle inner membrane	4.52e−43	3.70e−11	−1.68	NA
			GO:0098798	Mitochondrial protein complex	2.31e−40	8.47e−10	−1.83	NA
			GO:0005840	Ribosome	4.35e−39	1.75e−10	−1.91	NA
			GO:0044391	Ribosomal subunit	5.29e−39	2.47e−11	−2.03	NA
			GO:0098800	Inner mitochondrial membrane protein complex	1.13e−30	6.41e−09	−2.06	NA
			GO:0022626	Cytosolic ribosome	2.07e−30	1.75e−10	−2.23	NA
		Molecular function	GO:0003735	Structural constituent of ribosome	2.12e−39	3.20e−12	−2.17	NA

[a] Rank-based enrichment test based on minimum significant difference identifies gene-sets that are over-represented among top-ranked genes without a directional hypothesis

[b] Gene-set enrichment analysis (GSEA) tests for over-representation in among top and bottom genes based on Log_2 fold-changes in comparing moderate- (MOD) versus low-volume (LOW) conditions. Positive normalized enrichment scores (NES) indicates genes with higher expression in MOD compared to LOW, negative NES indicates genes with higher expression in LOW compared to MOD

[c] Over-representation tests based on differentially expressed genes. *P*-values are adjusted for FDR

expression in the moderate volume condition (Fig. 3h and i). Of these genes, two were shared between models (Fig. 3i). Using tissue-offset model-derived estimates for functional annotation analyses, revealed no consensus GO sets, using either of the two enrichment approaches (Table 2). In contrast, functional annotation based on estimates from the library-size and naïve model revealed GO terms related to cellular respiration and protein translation with enrichment scores (NES) indicating higher expression of

genes in the low-volume condition (Table 2). No DE genes contributed to ORA-identified GO terms among the top-ranked GO terms from the library-size model. In the naïve model, genes related to cellular respiration were identified as DE with subsequent contribution in ORA-identified GO terms (Table 2).

Identification of genes with altered expression over time (0, 2 and 12 weeks)

At Week 2 and 12 (rested-state biopsies), we also investigated the overall effects of resistance training on transcriptome profiles: i.e. the time effect, assessed by combining data from the two training modalities. At Week 2, resistance training led to increased expression for 3923, 1609, and 3875 genes and decreased expression of 77, 289, and 100 genes using the tissue-offset, the library-size, and the naïve model, respectively (Fig. 4a). The majority of these DE genes were found in the intersection between all models (Fig. 4a lower panel). At Week 12, resistance training led to increased expression of 1733, 584, and 5108 genes and decreased expression of 2, 19, and 2 genes using the tissue-offset, the library-size, and the naïve model, respectively (Fig. 4b). Here, the majority number of DE genes were found in the intersection between the tissue-offset model and the naïve model (Fig. 4b lower panel). At both Week 2 and 12 (and using any normalization model), functional analyses of DE genes revealed enrichment of GO terms associated with ECM structure, organization, and synthesis, as well as stress responses (Table 3).

Effects of acute exercise on transcriptome profiles (pre- to post-exercise in Week 2)

At Week 2, we also investigated the effects of acute bouts of resistance exercise with low and moderate volume on transcriptome profiles. As we did not expect changes in the total RNA-to-muscle mass ratio in this short time span [37] and rather fluid shifts [38] may have affected tissue weight and hence downstream analyses in an undesirable manner, transcriptome analyses were performed using the library-size model. First, we performed an analysis of the effects of resistance training per se on transcriptome profiles (combining data from the two training modalities). These analyses identified 1736 DE genes after acute resistance exercise, 707 of which showed increased expression and 1029 of which showed decreased expression (Fig. 5a). Genes that showed increased expression were generally associated with stress-related GO terms, including immune response (Table 4). Genes that showed decreased expression were associated with ECM-related GO terms (Table 4), contrasting observations made in rested-state biopsies at Week 2 and 12 (Table 3; detailed table in Additional files 1 and 2). We then compared the effects of low and moderate-volume conditions. These analyses identified one single DE gene (RFT1, Fig. 5b), which decreased to a greater extent in the moderate-volume condition. Despite this, rank-based enrichment tests with MSD identified five GO terms with significant enrichment. Among these five categories, three had genes with MSD > 0 (RNA splicing, RNA localization, and Covalent chromatin modification), indicating that the lower bound of 95% CI did not overlap no-change. However, as differences between volume conditions were both negative and positive, as indicated by the rug-plot in Fig. 5c, these categories were not identified in gene-set enrichment analysis based on fold-changes. Overall, these analyses do not provide evidence for pronounced

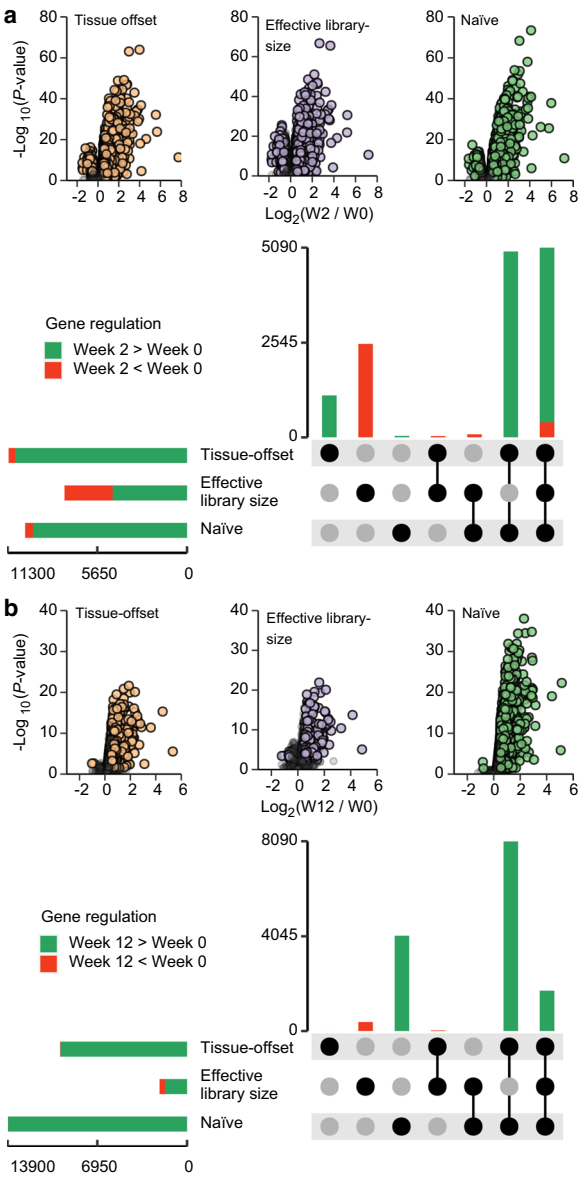


Fig. 4 Comparing the effects of resistance training per se on transcriptome profiles using different normalization models. Volcano plot identifies differentially expressed genes at Week 2 (**a**) and Week 12 (**b**) (adjusted P-values < 0.05 and Log2 fold-changes >|0.5|, filled circles). Bar-plots shows the total number of DE genes (horizontal bars) and sets exclusively found in each model or shared among models (vertical bars). The majority of differentially expressed genes were identified by all three normalization models, though the effective library-size model identified a larger number of genes with decreased expression (**a**)

volume-dependent regulation of mRNA expression in the acute recovery phase after resistance exercise (1-h).

Discussion

In the present study, we used a within-participant model to study the effects of low and moderate resistance training volumes on transcriptome responses. For these analyses, we used a subset of muscle biopsy material from a previously reported

Table 3 Functional annotation analysis of time-dependent effects of resistance training

Comparison	Normalization model	Gene ontology category	ID	Description	Rank P-value	Gene-set enrichment analysis (GSEA)		ORA P-value
						GSEA P-value	NES	
Week 2	Tissue offset	Biological process	GO:0043062	Extracellular structure organization	7.28e−41	5.92e−25	1.96	9.04e−28
			GO:0006954	Inflammatory response	1.28e−30	5.92e−25	1.79	9.04e−28
			GO:0002274	Myeloid leukocyte activation	4.48e−24	3.87e−14	1.59	2.40e−16
			GO:0050900	Leukocyte migration	4.79e−22	4.62e−18	1.82	2.20e−17
			GO:0002444	Myeloid leukocyte mediated immunity	3.46e−20	4.71e−12	1.57	8.98e−15
		Cellular component	GO:0031012	Extracellular matrix	5.85e−70	5.45e−51	2.17	3.58e−52
			GO:0062023	Collagen containing extracellular matrix	1.19e−68	3.15e−47	2.20	5.88e−53
			GO:0005788	Endoplasmic reticulum lumen	6.10e−24	2.71e−12	1.75	5.21e−17
			GO:0005581	Collagen trimer	1.29e−21	9.26e−14	2.24	1.29e−10
		Molecular function	GO:0005201	Extracellular matrix structural constituent	1.70e−40	3.56e−24	2.23	8.04e−30
	Effective library size	Biological process	GO:0043062	Extracellular structure organization	1.28e−34	6.94e−23	2.03	4.27e−29
			GO:0006954	Inflammatory response	3.11e−26	3.15e−21	1.82	9.15e−28
			GO:0050900	Leukocyte migration	5.50e−18	1.58e−13	1.83	8.59e−15
			GO:0030199	Collagen fibril organization	1.01e−17	2.75e−11	2.28	2.49e−14
			GO:0042330	Taxis	5.63e−17	1.74e−12	1.69	2.15e−16
		Cellular component	GO:0031012	Extracellular matrix	2.64e−63	7.23e−43	2.17	1.48e−53
			GO:0062023	Collagen containing extracellular matrix	2.38e−61	4.14e−38	2.20	4.43e−52
			GO:0005581	Collagen trimer	3.52e−20	8.15e−12	2.23	1.18e−14
			GO:0005788	Endoplasmic reticulum lumen	8.79e−18	2.51e−12	1.87	1.50e−10
		Molecular function	GO:0005201	Extracellular matrix structural constituent	9.00e−38	6.36e−18	2.18	2.82e−28

Table 3 (continued)

Comparison	Normalization model	Gene ontology category	ID	Description	Rank P-value	Gene-set enrichment analysis (GSEA)		ORA P-value
						GSEA P-value	NES	
	Naïve	Biological process	GO:0043062	Extracellular structure organization	1.38e−40	2.25e−23	1.93	3.36e−26
			GO:0006954	Inflammatory response	1.30e−31	2.61e−26	1.78	5.70e−28
			GO:0002274	Myeloid leukocyte activation	3.44e−23	4.99e−14	1.57	6.42e−16
			GO:0050900	Leukocyte migration	7.26e−23	8.32e−17	1.79	7.78e−20
			GO:0042330	Taxis	2.04e−19	4.63e−18	1.69	5.41e−17
		Cellular component	GO:0031012	Extracellular matrix	4.84e−71	2.20e−54	2.15	7.81e−53
			GO:0062023	Collagen containing extracellular matrix	9.15e−70	7.75e−49	2.19	1.01e−53
			GO:0005788	Endoplasmic reticulum lumen	4.50e−24	2.56e−12	1.72	4.68e−18
			GO:0005581	Collagen trimer	8.56e−22	1.06e−13	2.20	2.17e−11
		Molecular function	GO:0005201	Extracellular matrix structural constituent	1.98e−40	4.68e−25	2.20	6.38e−29
Week 12	Tissue offset	Biological process	GO:0043062	Extracellular structure organization	1.86e−49	2.01e−28	2.22	1.09e−37
			GO:0001501	Skeletal system development	7.13e−21	6.49e−14	1.77	1.09e−16
			GO:0030199	Collagen fibril organization	2.46e−20	4.83e−12	2.48	1.29e−15
		Cellular component	GO:0031012	Extracellular matrix	8.63e−72	3.48e−58	2.47	2.23e−69
			GO:0062023	Collagen containing extracellular matrix	4.96e−69	4.97e−54	2.52	8.25e−67
			GO:0005581	Collagen trimer	3.65e−25	1.27e−19	2.64	2.40e−25
			GO:0005788	Endoplasmic reticulum lumen	8.11e−20	1.85e−10	1.84	1.87e−13
		Molecular function	GO:0005201	Extracellular matrix structural constituent	2.46e−47	1.34e−34	2.63	4.24e−46
			GO:0005539	Glycosaminoglycan binding	1.48e−20	3.73e−15	2.15	2.01e−17
			GO:0008201	Heparin binding	5.72e−19	1.03e−15	2.25	9.82e−16

Table 3 (continued)

Comparison	Normalization model	Gene ontology category	ID	Description	Rank P-value	Gene-set enrichment analysis (GSEA)		ORA P-value
						GSEA P-value	NES	
	Effective library size	Biological process	GO:0043062	Extracellular structure organization	1.79e−44	3.76e−21	1.90	8.48e−33
			GO:0030199	Collagen fibril organization	4.77e−19	3.05e−08	2.10	2.69e−10
		Cellular component	GO:0031012	Extracellular matrix	6.91e−67	6.04e−41	2.03	2.38e−61
			GO:0062023	Collagen containing extracellular matrix	1.43e−63	2.57e−38	2.07	2.59e−56
			GO:0005581	Collagen trimer	1.74e−24	6.46e−13	2.17	1.11e−26
			GO:0005788	Endoplasmic reticulum lumen	1.45e−17	8.84e−09	1.66	1.58e−13
		Molecular function	GO:0005201	Extracellular matrix structural constituent	3.54e−45	1.48e−21	2.13	4.31e−42
			GO:0005198	Structural molecule activity	3.18e−20	2.05e−14	1.61	NA
			GO:0005539	Glycosaminoglycan binding	1.36e−18	2.43e−12	1.90	1.05e−15
			GO:0008201	Heparin binding	2.49e−18	7.91e−11	1.95	6.20e−16
	Naïve	Biological process	GO:0043062	Extracellular structure organization	2.90e−52	1.61e−39	2.94	1.06e−28
			GO:0001501	Skeletal system development	7.74e−23	3.14e−18	2.21	1.05e−11
			GO:0030199	Collagen fibril organization	2.99e−21	1.57e−15	3.14	5.82e−08
		Cellular component	GO:0031012	Extracellular matrix	1.49e−81	7.09e−79	3.31	2.66e−44
			GO:0062023	Collagen containing extracellular matrix	4.03e−79	1.04e−69	3.39	4.01e−46
			GO:0005581	Collagen trimer	5.29e−27	6.03e−27	3.39	4.83e−12
			GO:0005788	Endoplasmic reticulum lumen	8.14e−24	5.17e−16	2.37	1.45e−14

Table 3 (continued)

Comparison	Normalization model	Gene ontology category	ID	Description	Rank P-value	Gene-set enrichment analysis (GSEA)		ORA P-value
						GSEA P-value	NES	
		Molecular function	GO:0005201	Extracellular matrix structural constituent	1.22e−50	7.93e−48	3.59	6.42e−28
			GO:0005539	Glycosaminoglycan binding	7.61e−24	1.07e−19	2.78	6.84e−13
			GO:0008201	Heparin binding	1.83e−21	1.44e−18	2.90	1.97e−09

[a] Rank-based enrichment test based on minimum significant difference identifies gene−sets that are over-represented among top-ranked genes without a directional hypothesis

[b] Gene−set enrichment analysis (GSEA) tests for over-representation in among top and bottom genes based on Log_2 fold-changes in comparing time−points (Week 2 vs. Week 0 and Week 12 vs. Week 0). Positive normalized enrichment scores (NES) indicate genes with higher expression in Week 2/12 compared to Week 0; negative NES indicates genes with higher expression in Week 0 compared to Week 2/12

[c] Over-representation tests based on differentially expressed genes. *P*-values are adjusted for FDR

study [3]. Training volume led to robust increases in muscle strength and limb lean mass, resembling observations made in the full study cohort [3], and previous studies [39]. Despite these benefits, few differences were detected in transcriptome profiles between volume conditions, with the most prominent exception being a selection of genes involved in extracellular matrix organization and biology in the early stages of resistance training. We have shown that disclosure of these differences was made possible by our systematic selection of bioinformatic tools.

Identification of a model-specific bioinformatics pipeline

The continued development of bioinformatic tools for RNA-seq analyses require continuous optimization of analytic pipelines to any specific study conditions [24, 40]. To this end, we first sought to select a suitable read-trimming method. This is necessary to provide high-quality downstream alignment and k-mer search in reads [30]. Two commonly used algorithms were compared [30, 31]. Trimmomatic provided data with higher quality than Trimgalore (Fig. 1G), which did not improve quality scores compared to non-filtered data. Second, we compared five mapping tools for performing transcript quantification of trimmed reads: two genome-based mapping tools (STAR [41] and HISAT2 [42]) and three transcript based mapping tools (RSEM [43], kallisto [44] and Salmon [45]). Transcript-based mapping tools resulted in stronger correlations between mRNA and protein profiles, measured as relationships between myosin heavy chain mRNA profiles and protein abundances in rested state biopsies [3] (Fig. 2a), which is known to correlate in resting human skeletal muscle [33–35]. This comparison was performed using gene-/protein-family normalization [34, 46], allowing the deduction of mRNA-to-protein relationships without the need for other normalization assumptions. Notably, a marked skew was observed in the relationship between *MYH1* proportions and its corresponding Type IIX fiber following the initial part of the training intervention. This coincided with robust changes in *MYH1*

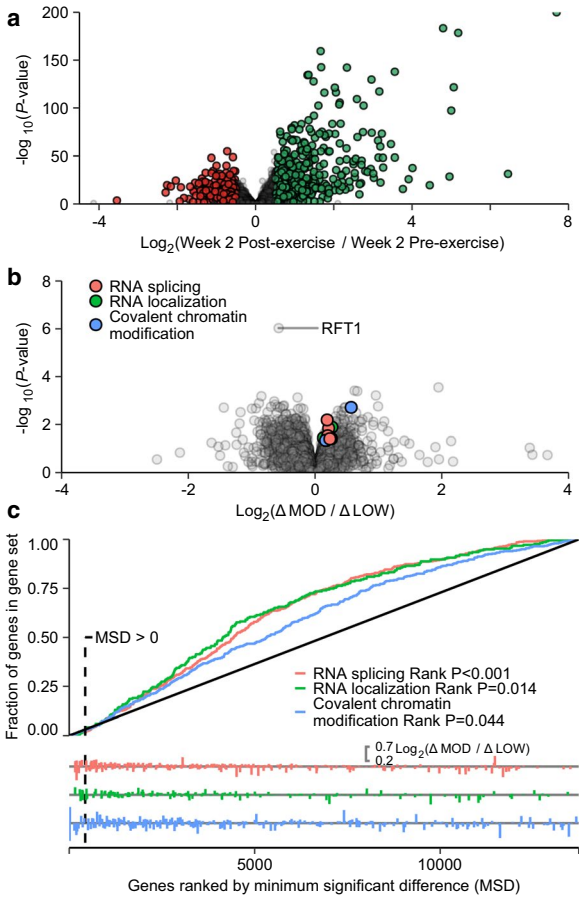


Fig. 5 Comparing the effects of an acute bout of resistance training with low and moderate volume on transcriptome profiles in muscle biopsies. Overall, an acute bout of resistance training led to large-scale alterations in gene expression (volume-conditions combined) (**a**). Comparing differentially expressed events between volume conditions (low-volume, LOW; moderate-volume, MOD) identified one gene with volume-dependent changes in expression (RFT1, **b**). Three gene ontology categories were identified as significantly enriched using a rank-based test (minimum significant difference, MSD as the rank metric) as well as containing genes with unadjusted *P*-values < 0.05 (MSD > 0). Traces from the rank test are displayed in **c**

expression, as is typically seen in response to mechanical loading in such short time frames [47]. Furthermore, the transcript-based mapping tools (RSEM, kallisto, and Salmon) resulted in transcriptome profiles with an expected bimodal distribution of counts and a larger subset of detected genes compared to genome-based tools (Fig. 1h). They were also associated with the less technical variation, evident as lower Log2-fold differences in expression for a selection of highly expressed genes between the two legs at baseline [23], assuming minimal biological variation between such paired samples. In these analyses, RSEM displayed slightly lower average variation between paired samples, thus outperforming kallisto and Salmon (Fig. 1j).

Table 4 Functional annotation analysis of time−dependent effects of acute resistance exercise

Gene ontology category	ID	Description	Rank P-value	Gene−set enrichment analysis (GSEA)		ORA P-value
				GSEA P-value	NES	
Biological process	GO:0050900	Leukocyte migration	2.31e−16	0.010	1.58	1.04e−04
	GO:0009617	Response to bacterium	6.61e−14	1.42e−04	1.81	8.54e−06
	GO:0060326	Cell chemotaxis	6.69e−14	0.019	1.61	1.15e−04
	GO:0006954	Inflammatory response	2.08e−12	0.019	1.45	0.048
	GO:0002237	Response to molecule of bacterial origin	2.08e−12	4.61e−05	2.04	2.29e−05
	GO:0042330	Taxis	1.11e−11	9.24e−04	1.68	3.05e−05
	GO:0007159	Leukocyte cell cell adhesion	1.17e−11	0.003	1.78	4.02e−04
	GO:0030595	Leukocyte chemotaxis	8.29e−11	0.038	1.54	0.004
	GO:0048514	Blood vessel morphogenesis	8.72e−11	1.20e−05	1.89	1.78e−11
	GO:0042110	T cell activation	1.25e−09	0.022	1.46	6.84e−04
Cellular component	GO:0072562	Blood microparticle	3.23e−04	0.726	1.20	NA
	GO:0098589	Membrane region	0.004	0.162	1.39	NA
	GO:0042581	Specific granule	0.009	0.644	1.18	NA
	GO:0070820	Tertiary granule	0.055	0.444	1.28	NA
	GO:0005667	Transcription factor complex	0.070	1.93e−04	1.88	NA
	GO:0000932	P body	0.082	0.243	1.49	0.049
	GO:1903293	Phosphatase complex	0.095	0.012	1.97	0.049
	GO:0030055	Cell substrate junction	0.110	0.033	1.49	NA
	GO:0051233	Spindle midzone	0.116	0.647	1.29	NA
	GO:1904724	Tertiary granule lumen	0.180	0.091	1.81	NA
Molecular function	GO:0001216	DNA binding transcription activator activity	8.87e−15	1.50e−14	2.56	NA
	GO:0035326	Cis regulatory region binding	7.65e−10	3.00e−15	2.38	0.035
	GO:0030545	Receptor regulator activity	3.34e−07	0.004	1.75	0.003
	GO:0005125	Cytokine activity	8.04e−07	0.002	2.00	0.004
	GO:0001217	DNA binding transcription repressor activity	2.71e−06	6.07e−04	1.89	NA
	GO:0001968	Fibronectin binding	3.90e−05	0.411	1.30	0.024
	GO:0001664	G protein coupled receptor binding	3.90e−05	0.078	1.52	0.047
	GO:0008083	Growth factor activity	4.22e−05	0.001	2.03	0.004

Table 4 (continued)

Gene ontology category	ID	Description	Rank P-value	Gene—set enrichment analysis (GSEA)		ORA P-value
				GSEA P-value	NES	
	GO:0005126	Cytokine receptor binding	2.11e−04	9.14e−04	1.94	0.047
	GO:0140272	Exogenous protein binding	3.80e−04	0.198	1.54	NA
Biological process	GO:0043062	Extracellular structure organization	1.21e−11	0.006	−1.57	0.046
	GO:0001501	Skeletal system development	6.32e−09	0.057	−1.32	0.008
	GO:0072676	Lymphocyte migration	3.54e−07	0.126	−1.42	NA
	GO:0032963	Collagen metabolic process	1.43e−06	0.032	−1.61	NA
	GO:0002697	Regulation of immune effector process	1.60e−06	0.115	−1.31	NA
	GO:0002250	Adaptive immune response	6.08e−06	0.023	−1.43	NA
	GO:0070661	Leukocyte proliferation	6.23e−06	0.055	−1.40	NA
	GO:0060348	Bone development	7.92e−06	0.031	−1.49	NA
	GO:0042098	T cell proliferation	1.68e−05	0.127	−1.34	NA
	GO:0033627	Cell adhesion mediated by integrin	1.69e−05	0.288	−1.27	0.007
Cellular component	GO:0031012	Extracellular matrix	8.06e−15	7.19e−11	−2.12	NA
	GO:0062023	Collagen containing extracellular matrix	8.74e−12	3.62e−13	−2.29	NA
	GO:0005581	Collagen trimer	3.23e−04	7.19e−11	−2.56	NA
	GO:0009897	ExteRNAl side of plasma membrane	0.001	0.023	−1.57	NA
	GO:0005788	Endoplasmic reticulum lumen	0.008	0.003	−1.69	NA
	GO:0098552	Side of membrane	0.012	0.460	−1.18	NA
	GO:0043235	Receptor complex	0.014	0.210	−1.29	NA
	GO:0043202	Lysosomal lumen	0.032	0.002	−1.91	NA
	GO:0035579	Specific granule membrane	0.065	0.839	−1.09	NA
	GO:0098802	Plasma membrane signaling receptor complex	0.092	0.263	−1.37	NA
Molecular function	GO:0005201	Extracellular matrix structural constituent	2.40e−07	6.68e−10	−2.38	NA
	GO:0005539	Glycosaminoglycan binding	1.38e−06	0.251	−1.27	0.048
	GO:0008201	Heparin binding	7.60e−06	0.224	−1.33	NA
	GO:0005178	Integrin binding	7.04e−04	0.103	−1.42	NA

Table 4 (continued)

Gene ontology category	ID	Description	Rank P-value	Gene—set enrichment analysis (GSEA)		ORA P-value
				GSEA P-value	NES	
	GO:1901681	Sulfur compound binding	8.19e—04	0.098	— 1.35	NA
	GO:0030246	Carbohydrate binding	8.22e—04	0.230	— 1.27	NA
	GO:0005518	Collagen binding	0.004	0.050	— 1.66	NA
	GO:0043394	Proteoglycan binding	0.010	0.185	— 1.53	NA
	GO:0050840	Extracellular matrix binding	0.011	0.351	— 1.27	NA
	GO:0019838	Growth factor binding	0.016	0.195	— 1.35	NA

[a] Rank-based enrichment test based on minimum significant difference identifies gene—sets that are over-represented among top-ranked genes without a directional hypothesis

[b] Gene—set enrichment analysis (GSEA) tests for over-representation in among top and bottom genes based on Log_2 fold-changes in comparing time—points (Post- vs. Pre—exercise). Positive normalized enrichment scores (NES) indicate genes with higher expression Post- compared to Pre—exercise, negative NES indicates genes with higher expression Pre—compared to Post-exercise

[c] Over-representation tests based on differentially expressed genes. *P*-values are adjusted for FDR

Comparison of normalization strategies

Transcriptome analyses often rely on the assumption that gene expression is counted and compared between conditions on a per-cell level [28]. This is 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, 48].

In the current study population, we previously showed that total RNA increases per-unit-muscle tissue in a volume-dependent manner following initiation of resistance training [3]. Consequently, the preparation of cDNA libraries for RNA-seq experiments was unavoidably based on different amounts of tissue, as fixed amounts of total RNA (1000 ng) were used for this purpose. If unaccounted for, this would lead to a comparison of transcriptome data originating from different amounts of muscle tissue for the two volume conditions. We show that this leads to a contra-intuitive larger increase in global transcript counts in the low-volume condition compared to the moderate-volume condition (though without reaching statistical significance). In contrast, after adjusting for the difference in amounts of muscle tissue fed into RNA-seq experiments (i.e., tissue-offset normalization), this apparent difference in average library size disappeared. Overall, we thus observed an increase in global mRNA expression per-unit-muscle weight (43–53%) in response to resistance training that did not depend on training volume, contrasting the observed volume-dependency of total RNA expression [3]. This global change in mRNA expression was associated with substantial alterations in the expression of a multitude of genes, with as many as 26 and 12% of the total read-count pool showing increased expression using the tissue-offset model at Week 2 and Week 12, respectively. These genes were associated with biological processes such as ECM synthesis and organization corroborating with previous studies [13, 18, 22].

As the volume-dependent changes in muscle growth and total RNA levels in *m. vastus lateralis* arguably will affect downstream bioinformatics analyses and identification of DE genes, we aimed to compare three different normalization strategies during the subsequent analyses: tissue-offset, library-size, and naïve. At Weeks 2 and 12, the different normalization strategies resulted in marked shifts in global mRNA responses between the two volume conditions, having pronounced effects on identification of GO terms during enrichment analyses (Fig. 3g). In general, tissue-offset normalization (providing mRNA expression per-mg-muscle weight) resulted in a global shift in mRNA responses towards moderate volume. At Week 2, this was evident as more robust increases in mRNA expression in response to higher training volume for most genes, contrasting findings in library-size and naïve analyses. At Week 12, this was evident as a counterbalancing of mRNA expression profiles, with tissue-offset providing rather normally distributed responses to the two volume conditions, contrasting the skew towards larger mRNA expression in response to low training volume in library-size and naïve analyses. The utilization of generalized linear mixed models (GLMM) allowed convenient comparisons of normalization models, which could then be fitted using in the same statistical framework, as previously suggested [36]. GLMM also allowed the incorporation of random effects into the model to account for the repeated measures design. Although there are approaches to account for correlated observations in commonly used RNA-seq modeling frameworks [49], GLMMs provides a more robust and potentially more powerful framework for dealing with correlated data [36].

Training volume-dependent changes in transcriptome profiles.

Using the tissue-offset model, we were able to identify genes relating to ECM functions as volume-sensitive during the early stages of resistance training. This may indicate a role for ECM remodeling in the beneficial effects of higher training volumes on muscular adaptations and strength. As such, previous research has shown that ECM remodeling is induced by exercise training, both acutely [50–53], and after prolonged endurance and resistance training [13, 18, 54, 55], at both mRNA and protein levels [13, 18, 54, 55]. However, none of these previous studies have found ECM remodeling to be differentially affected by different exercise modalities. Rather, different resistance training modalities such as low- and high-load training have been associated with similar responses, measured as collagen synthesis [52]. Importantly, there seems to be a close association between training-induced changes in abundances of ECM-related mRNAs and their respective proteins, including collagen-organization proteins [54]. This suggests that ECM remodeling is primarily controlled at the transcriptional level, which arguably increases the biological relevance of the herein presented transcriptome analyses. However, this relationship seems to involve a complex time course dependency. For example, transcriptional regulation of COL1A2 shows a considerable lag from stimuli to transcription, as shown in fibroblasts [56]. In line with this, our data on the effects of acute resistance exercise on transcriptome profiles, suggests a counterintuitive reduction in the expression of e.g., collagen mRNA immediately after exercise, as has also been found by others [18]. Indeed, enrichment analyses confirmed these negative changes in ECM-related transcripts in response to acute exercise, contrasting the effects of chronic resistance training [13, 18].

The physical properties of ECM are distorted by disuse and aging, resulting in increased stiffness and potentially decreased force transmission and muscle efficiency [57, 58]. Training-induced ECM remodeling may thus constitute an effective measure to reverse these adversities [50–53], and has been suggested to exert a protective role against injury [51]. However, available studies are ambiguous in their conclusions [52, 53], and the link between observed changes in ECM-related gene expression and muscle biology and functionality remain uncertain. Adding to this, ECM remodeling-responses to training seems to be age-dependent [50, 59], and also shows a clear dependency of time. Indeed, in the present study, the volume-associated differences in ECM-related transcriptome profiles disappeared entirely after twelve weeks of resistance training, whereby no genes were identified as showing volume-dependent responses. From a general perspective, this indicates that after prolonged training, the biological state of the muscle may have reached a new equilibrium, with low- and moderate-volume training having led to similar muscle phenotypic traits. This would imply that the benefits of higher training volume are restricted to augmented increases in muscle mass, perhaps facilitated by increases in ribosomal biogenesis [3]. Notably, this is likely to be an oversimplification. Indeed, at the protein level, in the current study, higher training volume led to a more robust phenotypic switch from type IIX → type IIA fiber also after twelve weeks of training [3]. Taken together, our data provide valuable directions for future research. It suggests that ECM remodeling is volume-dependent, at least during the initial part of a training program. This needs to be confirmed by studies in other populations, and its biological and functional significance needs clarification. Such studies should take advantage of the increased biological resolution of within-subject contralateral models.

Analyses of transcriptome responses to acute bouts of low and moderate resistance training volume revealed one single gene with volume-dependent changes in expression (RFT1). RFT1 is associated with the GO terms lipid transport, carbohydrate transport, and endoplasmic reticulum membrane. RFT1 expression has previously been shown to decrease in muscle immediately after training [22]. Although this warrants more research, one single gene arguably provides limited information. As such, rank-based enrichment tests identified three GO terms with volume-dependent changes in gene expression in the acute data set (RNA splicing, RNA localization, and covalent chromatin modification). Upon closer examination, these gene sets consisted of genes with both increased and decreased expression in the moderate- compared to the low-volume condition. Only a small fraction of these transcripts showed actual positive MSDs, indicating changes with unadjusted P-values < 0.05. While this enrichment analysis supports acute volume-dependent regulation of gene sets after resistance training at the selected time point of biopsy sampling (1 h after sessions), it remains plausible that such regulation would have been more pronounced at later time points.

Conclusions

Transcriptomic analyses of skeletal muscle subjected to altered growth conditions should account for global changes in mRNA to total RNA and cell density to accurately reflect biologically meaningful events. When accounting for such aspects, ECM

remodeling showed volume-dependent responses to resistance training. These recommendations could be applicable to studies of other cell types and model systems undergoing increased or arrested growth. Also, the optimized selection of bioinformatic tools increases the biological relevance of transcriptome analyses from resistance-trained skeletal muscle.

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) or moderate-volume (three sets per exercise) training (Fig. 1a). Muscle biopsies were obtained from each leg prior (Week 0) to and after the intervention (Week 12), as well as prior to (Week 2 Pre-ex) and 60-min after (Week 2 Post-ex) the last training session of week 2, as previously described [3]. Participants with a complete set of high-quality RNA samples (RQI ≥ 7, $n = 25$) were selected for RNA-seq analyses (Fig. 1b). Training-induced changes in muscle size and strength were estimated for each leg using several methods (for a complete overview, see [3]). Herein, we present dual-energy X-ray absorptiometry (DXA) 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) (Table 1).

Training protocol

The training protocol consisted of unilateral lower body exercises (leg-press, leg-curl, and knee-extension). Each participant's leg was randomly assigned to perform either one or three sets per exercise (low- vs. moderate-volume), ensuring within-participant comparisons. Rest periods between sets were 90–180 s. The single-set leg was always trained in the rest period between the second and third sets 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 was 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 lean mass assessments

Muscle strength was assessed twice before and once after the intervention. A detailed description of strength outcomes resulting from the study has been previously reported [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° $\times$ s^{-1}) knee extension, and one-repetition maximum (1RM) in unilateral knee extension and leg press, as previously reported [3]. 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. Limb lean-mass 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. Limb lean-mass was derived from a segment 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, immunohistochemistry 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 mg $\times$ ml^{-1} with adrenaline 5 µg $\times$ ml^{-1}, AstraZeneca AS, Oslo, Norge). Samples were obtained from the two legs within 10 min of each other at all time-points. All rested state samples were obtained in the morning after a standardized breakfast. Resting samples obtained in Week 2 were sampled approximately 48 h after the fourth session. After the training period (Week 12), samples were obtained six days after the last training session and three days after the last strength assessment. Participants were instructed to ingest standardized meals during the 24 h leading up to the sampling event, and to refrain from strenuous physical activity during the last 48 h. Samples for immunohistochemistry (~ 15 mg) were transferred to a 4% formalin solution for fixation 24 − 72 h, before further preparation. Samples for RNA analyses (~ 25 mg) were dissected in ice-cold sterile saline solution (0.9% NaCl), blotted dry and snap-frozen using −80 °C isopentane, before storage at −80 °C until further processing.

Immunohistochemistry was utilized to quantify myosin heavy chain abundance in formalin-fixed muscle biopsy cross-sections, performed as previously described and reported [3]. Briefly, 4 µm transverse sections were incubated with (1) a primary antibody that detects all three adult myosin heavy chain isoforms but type IIX (BF-35, 5 µg $\times$ ml^{-1}, Developmental Studies Hybridoma Bank, deposited by Schiaffino, S.) and (2) type I myosin (MyHCSlow, 1:4000, catalog 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 Fig. 3 in [3]). Hybrid fibers were analyzed as 0.5 $\times$ Type IIA and 0.5 $\times$ Type IIX.

For RNA extraction, the frozen muscle was 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 µl 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 a 260/280 nm ratio > 1.95 assessed by a spectrophotometer (NanoDrop 2000, ThermoFisher Scientific, Oslo, Norway). RNA integrity scores (RQI) were determined by capillary electrophoresis (Experion Automated Electrophoresis Station using RNA StdSens Assay, Bio-Rad). For each participant, all samples were extracted in the same extraction session in a randomized order. 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] (Fig. 1c). Notably, to achieve accurate normalization of qPCR data (and potentially also RNA-seq data), a

commercially available exogenous RNA control (λ polyA External Standard Kit, Takara Bio Inc., Shiga, Japan) was added at a fixed amount per extraction prior to homogenization (0.04 ng ml^{-1} of Trizol reagent), as previously described [3, 60]. Unfortunately, at present, we do not have access to the sequence of this spike-in, prohibiting its identification in RNA-seq data and rendering its subsequent utilization for normalization purposes difficult.

Illumina library preparation and sequencing

For each muscle sample, 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, https://github.com/FelixKrueger/TrimGalore) and Trimmomatic (version 0.39) [31] 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) and summarized using MultiQC (version 1.8) [61].

Read alignment

Filtered reads were aligned to the Human genome (GRCh38 release-97 downloaded from ftp.ensemble.org) using the alignment-based methods HISAT2 (version 2.1.0) [42], STAR (version 2.7.2) [41], and RSEM (version 1.3.1) [43], used together with Bowtie 2 (version 2.3.4.3) [62], and non-alignment methods including kallisto (version 0.44.0) [44] and Salmon (version 0.13.1) [45]. For HISAT2 and STAR, HTSeq was used for quantification as previously described [63]. 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 [36], 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 (pre- to post-exercise in the last session of week 2), only the library size normalized model was used as we expected that fluid shifts [38] could influence the muscle weight measurement, and changes in total RNA were unlikely to occur in this short time span [37]. The effective library size was calculated by multiplying the total library size with the RNA composition normalization factor, calculated using the trimmed mean method [25], followed by dividing the value

by the median effective library size, as suggested by Cui et al. [36]. The effect of resistance training on gene counts was assessed as (1) the effect of exercise volume and (2) the effect of time. For analyses of the effect of exercise volume, differential expression was evaluated using GLMMs containing the interaction between time and exercise volume. For analyses of the effect of time, differential expression was evaluated using GLMMs 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 [64]. 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 primarily below 0.05 (0.37% of models showed $P > 0.05$). Heteroscedasticity was assessed using the uniformity test in the DHARMa package [65], 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 Log2 fold-change was greater than 0.5, and the adjusted *P*-value was below 5%. *P*-values were adjusted per-model coefficient to control the false discovery rate [66].

Functional annotation

Enrichment analyses of gene ontology (GO) gene sets were performed using three approaches. First, a non-parametric rank test (described in [67] and implemented in the tmod package [68], 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 fold-change (FC) when $\text{Log(FC)} > 0$ and the negative inverse of the upper 95% CI when $\text{Log(FC)} < 0$. This metric has been shown to have lower false-positive rates compared to other metrics applied during enrichment analyses [69]. 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 down-regulated gene sets. A second approach, gene set enrichment analysis (GSEA) [70], was used to quantify the directional regulation of the gene set. GSEA was performed using the fgsea package [71] with Log(FC) as the gene level metric. Thirdly, over-representation analysis (ORA) was performed to assess if genes identified as differentially expressed ($|\text{Log2 fold-change}| > 0.5$ and adjusted $P\text{-values} < 0.05$; DE-genes) belonged to specific gene sets. ORA was performed using the enrichGO function in the clusterProfiler package [72], (version 3.16.0). GO gene sets (biological process, cellular component, and molecular function) were retrieved from the molecular signature database (version 7.1) [73].

Statistical analysis

Descriptive data are presented as mean and standard deviation (SD). Changes in muscle strength and mass were estimated using linear mixed models on change scores with baseline values as covariates. The performance of alignment tools was assessed by comparing log-differences between biological replicates, as suggested by Teng et al. [74], with modifications. Briefly, a subset of genes previously shown to be stably expressed between tissues was selected [32], whereupon log-differences between paired biopsy samples were 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 [34]. Notably, these data also provided insight into the overall biological validity of the RNA-seq data, linking gene counts to protein phenotypes.

Supplementary information

Supplementary information accompanies this paper at https://doi.org/10.1186/s12859-020-03866-y.

Additional file 1: Gene count estimates. Model (negative binomial generalized linear mixed models) based estimates of gene counts from models using different normalization strategies. ensemblid: Ensembl gene identifiers; normalization_model: Indicates normalization model: lib_size_normalized, gene counts are modeled with study conditions as fixed effects and participant id as random effects and with the addition of library size as a fixed effect. tissue_offset_lib_size_normalized, same as lib_size_normalized but with the addition of tissue weight used in library prep as an offset. The non-normalized model is not included in the data set; interaction_model: Indicates if the coefficient is estimated in a model containing the interaction between volume condition and time. If FALSE, the model only contains time as a fixed effect (in addition to any normalization, see above), coefficients should in this case be interpreted as averages over volume conditions; coefficient: Names of coefficients (fixed effects, time and volume conditions). timew2pre and timew12 indicate differences from timew0 (intercept) in rested state models. timew2post indicates differences from timew2pre (intercept) in acute exercise models. setsmultiple indicates interaction effects, the difference between setssingle (reference level) and setsmultiple; estimate: Estimates on the natural log scale; se: Standard errors (SE) on the natural log scale; zvalue: Z-values; pvalue: Un-adjusted P-values; pvalue_adjust: Adjusted P-values (FDR) per model/normalization method and coefficient.

Additional file 2: Functional annotation using gene ontology terms. Significance tests for functional annotation using rank-based and over-representation based analysis. ID: Gene ontology (GO) id; go_category: GO category, bp, Biological process; cc, cellular component; mf, molecular function; name: Descriptive GO name; normalization_model: Indicate normalization model used to test enrichment: lib_size_normalized, gene counts are modeled with study conditions as fixed effects and participant id as random effects and with the addition of library size as a fixed effect, tissue_offset_lib_size_normalized, same as lib_size_normalized but with the addition of tissue weight used in library prep as an offset. The non-normalized model is not included; coefficient: Names of coefficients used to test enrichment (fixed effects, time and volume conditions). timew2pre and timew12 indicates differences from timew0 (intercept) in rested state models. timew2post indicates differences from timew2pre (intercept) in acute exercise models. setsmultiple indicates interaction effects, the difference between setssingle (reference level) and setsmultiple; cerno_statistic: Test statistic from the rank-based cerno-test; cerno_auc: Area under the curve from the Cerno test (see tmod documentation for details, https://CRAN.R-project.org/package=tmod); cerno_pval: Un-adjusted P-values from cerno test, cerno_padj: Adjusted P-values (default settings in tmodCERNOtest, see tmod documentation); fgsea_pval: Un-adjusted P-values from gene set enrichment tests performed with the fgsea package; fgsea_padj: Adjusted fgsea P-values; NES: Normalized enrichment scores from gene set enrichment tests; set_size: Size of gene sets using genes expressed in the present data set; ora_geneRatio: Gene ratio in over-representation analysis of genes identified as differentially expressed; ora_padj: Adjusted p-values from over-representation analysis.

Abbreviations

DXA: Dual-energy X-ray absorptiometry; ECM: Extracellular matrix; DE: Differentially expressed; MSD: Minimum significant differences; RNA-seq: Ribonucleic acid sequencing; RQI: RNA quality indicator; ORA: Over-representation analysis.

Acknowledgements

The authors would like to acknowledge the skilled and hard work of Johanne Seeberg, Stine Dahl, Marianne Bratlien, Martin Nordseth, Erlend Hakestad, Ole-Martin Hveem and Mari Skifjeld in performing the training intervention, and Marita Hanestadhaugen and Lise Koll in establishing the immunohistochemistry protocol. The authors are also thankful for the dedicated contributions made by all participants.

Authors' contributions

DH, SE and BR planned and supervised the training intervention, and sampled muscle biopsies. DH performed immunohistochemistry analyses and RNA extraction. YK, DH, SE, and RA planned RNA-seq analyses. YK and DH performed bioinformatics analyses, with inputs from RA. YK and DH performed the statistical analyses. All authors provided useful input during the interpretation of data, and contributed to drafting and finalizing the manuscript. All authors have read and approved the manuscript.

Funding

The study was supported by grants from Innlandet Hospital Trust (Grant No. 150282, running costs) and Inland Norway University of Applied (running and wage costs). The funding bodies were not involved in the design of the study and collection, analysis, and interpretation of data and in writing the manuscript.

Availability of data and materials
Gene count estimates are available as Additional files 1 (see description below). Functional annotation using gene ontology are available as Additional files 2 (see description below). Additional datasets and code used in analysis during the current study are available from the corresponding author on reasonable request.

Ethics approval and consent to participate
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 preregistered at ClinicalTrials.gov (Identifier: NCT02179307). The study was conducted in accordance with the ***Declaration of Helsinki***.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1] Department of Biotechnology, Inland Norway University of Applied Sciences, Holsetgata 22, 2317 Hamar, Norway. [2] Section for Health and Exercise Physiology, Department of Public Health and Sport Sciences, Inland Norway University of Applied Sciences, Lillehammer, Norway. [3] Swedish School of Sport and Health Sciences, Stockholm, Sweden. [4] Innlandet Hospital Trust, Lillehammer, Norway. [5] Faculty of Health Sciences, Institute of Clinical Medicine, UiT - The Arctic University of Norway, Hansine Hansens veg 18, 9019 Tromsø, Norway.

Received: 19 August 2020 Accepted: 9 November 2020
Published online: 30 November 2020

References

1. Egan B, Zierath JR. Exercise metabolism and the molecular regulation of skeletal muscle adaptation. Cell Metab. 2013;17:162–84. https://doi.org/10.1016/j.cmet.2012.12.012.
2. Dirks ML, Wall BT, van de Valk B, 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. 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. J Physiol. 2020;598:543–65. 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.
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. 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. https://doi.org/10.1152/japplphysiol.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. 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. https://doi.org/10.1152/japplphysiol.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. 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. 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. https://doi.org/10.1152/japplphysiol.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. 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-κPB following muscle eccentric contractions in humans is localized primarily to skeletal muscle-residing pericytes. FASEB J. 2011;25:2956–66. 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 Metab. 2017;25:581–92. 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. 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. https://doi.org/10.1152/japplphysiol.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. https://doi.org/10.1152/japplphysiol.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. https://doi.org/10.1371/journal.pgen.1003389.
20. Hangelbroek RWJ, Fazelzadeh P, Tieland M, Boekschoten MV, Hooiveld GJEJ, van Duynhoven JPM, 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. 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. Med Sci Sports Exerc. 2009;41:546–65. 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. Nat Commun. 2020;11:470. 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. J Gerontol Ser A. 2007;62:1088–95. https://doi.org/10.1093/gerona/62.10.1088.
24. 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 Biol. 2016;17:13–23. https://doi.org/10.1186/s13059-016-0881-8.
25. Robinson MD, Oshlack A. A scaling normalization method for differential expression analysis of rna-seq data. Genome Biol. 2010;11:R25. https://doi.org/10.1186/gb-2010-11-3-r25.
26. 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. Brief Bioinform. 2013;14:671–83. https://doi.org/10.1093/bib/bbs046.
27. Arora S, Pattwell SS, Holland EC, Bolouri H. Variability in estimated gene expression among commonly used RNA-seq pipelines. Sci Rep. 2020;10:1. https://doi.org/10.1038/s41598-020-59516-z.
28. Lovén J, Orlando DA, Sigova AA, Lin CY, Rahl PB, Burge CB, et al. Revisiting global gene expression analysis. Cell. 2012;151:476–82. https://doi.org/10.1016/j.cell.2012.10.012.
29. 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. https://doi.org/10.1016/j.cell.2012.08.026.
30. 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. https://doi.org/10.1371/journal.pone.0085024.
31. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for illumina sequence data. Bioinformatics (Oxford, England). 2014;30:2114–20. https://doi.org/10.1093/bioinformatics/btu170.
32. Eisenberg E, Levanon EY. Human housekeeping genes, revisited. Trends Genet. 2013;29:569–74. https://doi.org/10.1016/j.tig.2013.05.010.
33. 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. https://doi.org/10.1046/j.1469-7580.2001.19930329.x.
34. 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. https://doi.org/10.1111/sms.12185.
35. 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. https://doi.org/10.1093/gerona/57.6.b232.
36. 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. https://doi.org/10.1515/sagmb-2015-0011.
37. Figueiredo VC, McCarthy JJ. Regulation of ribosome biogenesis in skeletal muscle hypertrophy. Physiology (Bethesda). 2019;34:30–42.
38. 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.
39. 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. https://doi.org/10.1080/02640414.2016.1210197.
40. Payá-Milans M, Olmstead JW, Nunez G, Rinehart TA, Staton M. Comprehensive evaluation of RNA-seq analysis pipelines in diploid and polyploid species. Gigascience. 2018;7:giy132.
41. 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. https://doi.org/10.1093/bioinformatics/bts635.
42. Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. Graph-based genome alignment and genotyping with hisat2 and hisat-genotype. Nat Biotechnol. 2019;37:907–15. https://doi.org/10.1038/s41587-019-0201-4.
43. Li B, Dewey CN. RSEM: Accurate transcript quantification from rna-seq data with or without a reference genome. BMC Bioinformatics. 2011;12:323. https://doi.org/10.1186/1471-2105-12-323.
44. Bray NL, Pimentel H, Melsted P, Pachter L. Near-optimal probabilistic rna-seq quantification. Nat Biotechnol. 2016;34:525–7. https://doi.org/10.1038/nbt.3519.

45. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression. Nat Methods. 2017;14:417–9. https://doi.org/10.1038/nmeth.4197.
46. 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. https://doi.org/10.1152/physiolgenomics.00196.2009.
47. 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.
48. 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. Sci Rep. 2017;7:15134. https://doi.org/10.1038/s41598-017-15420-7.
49. McCarthy DJ, Chen Y, Smyth GK. Differential expression analysis of multifactor rna-seq experiments with respect to biological variation. Nucleic Acids Res. 2012;40:4288–97. https://doi.org/10.1093/nar/gks042.
50. 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. https://doi.org/10.1080/17461391.2018.1526974.
51. Mackey AL, Brandstetter S, Schjerling P, Bojsen-Moller J, Qvortrup K, Pedersen MM, et al. Sequenced response of extracellular matrix deadhesion and fibrotic regulators after muscle damage is involved in protection against future injury in human skeletal muscle. Faseb j. 2011;25:1943–59. https://doi.org/10.1096/fj.10-176487.
52. Holm L, van Hall G, 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.
53. 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.
54. 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. Sci Rep. 2020;10:3514. https://doi.org/10.1038/s41598-020-60578-2.
55. 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. Phys Rep. 2015;3:e12473. https://doi.org/10.14814/phy2.12473.
56. 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. https://doi.org/10.1016/j.bbrep.2015.07.007.
57. Csapo R, Gumpenberger M, Wessner B. Skeletal muscle extracellular matrix - what do we know about its composition, regulation, and physiological roles? A narrative review. Front Physiol. 2020;11:253–63. https://doi.org/10.3389/fphys.2020.00253.
58. Azizi E, Deslauriers AR, Holt NC, Eaton CE. Resistance to radial expansion limits muscle strain and work. Biomech Model Mechanobiol. 2017;16:1633–43. https://doi.org/10.1007/s10237-017-0909-3.
59. 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. Exp Gerontol. 2018;106:28–38. https://doi.org/10.1016/j.exger.2018.02.013.
60. Ellefsen S, Stensløkken K-O, Sandvik GK, Kristensen TA, Nilsson GE. Improved normalization of real-time reverse transcriptase polymerase chain reaction data using an external RNA control. Anal Biochem. 2008;376:83–93. https://doi.org/10.1016/j.ab.2008.01.028.
61. Ewels P, Magnusson M, Lundin S, Käller M. MultiQC: Summarize analysis results for multiple tools and samples in a single report. Bioinformatics. 2016;32:3047–8. https://doi.org/10.1093/bioinformatics/btw354.
62. Langmead B, Salzberg SL. Fast gapped-read alignment with bowtie 2. Nat Methods. 2012;9:357–9. https://doi.org/10.1038/nmeth.1923.
63. Anders S, Pyl PT, Huber W. HTSeq–a python framework to work with high-throughput sequencing data. Bioinformatics. 2015;31:166–9. https://doi.org/10.1093/bioinformatics/btu638.
64. Brooks ME, Kristensen K, van Benthem KJ, Magnusson A, Berg CW, Nielsen A, et al. glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. R J. 2017;9:378–400. https://doi.org/10.32614/RJ-2017-066.
65. Hartig F. DHARMa: residual diagnostics for hierarchical (multi-level / mixed) regression models. 2020. https://florianhartig.github.io/DHARMa/.
66. Benjamini Y, Hochberg Y. Controlling the false discovery rate: A practical and powerful approach to multiple testing. J R Stat Soc Ser B (Methodol). 1995;57:289–300.
67. 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.
68. 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. https://doi.org/10.1093/bioinformatics/btz447.
69. Zyla J, Marczyk M, Weiner J, Polanska J. Ranking metrics in gene set enrichment analysis: do they matter? BMC Bioinformatics. 2017;18:256.
70. 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. https://doi.org/10.1073/pnas.0506580102.
71. Korotkevich G, Sukhov V, Sergushichev A. Fast gene set enrichment analysis. bioRxiv. 2019. https://doi.org/10.1101/060012.
72. Yu G, Wang L-G, Han Y, He Q-Y. ClusterProfiler: an r package for comparing biological themes among gene clusters. OMICS. 2012;16:284–7.
73. Liberzon A, Subramanian A, Pinchback R, Thorvaldsdóttir H, Tamayo P, Mesirov JP. Molecular signatures database (msigdb) 3.0. Bioinformatics. 2011;27:1739–40. https://doi.org/10.1093/bioinformatics/btr260.

74. Teng M, Love MI, Davis CA, Djebali S, Dobin A, Graveley BR, et al. A benchmark for rna-seq quantification pipelines. Genome Biol. 2016;17:74. https://doi.org/10.1186/s13059-016-0940-1.

Publisher's Note

2

Normalization of gene expression data revisited: the three viewpoints of transcriptome in human skeletal muscle undergoing load-induced hypertrophy

Yusuf Khan [1,2], Daniel Hammarström[2], Stian Ellefsen[2,3], Rafi Ahmad[1,4,*]

[1] Department of Biotechnology, Inland Norway University of Applied Sciences, Hamar, Norway.

[2] Section for Health and Exercise Physiology, Department of Public Health and Sport Sciences, Inland Norway University of Applied Sciences, Lillehammer, Norway.

[3] Innlandet Hospital Trust, Lillehammer, Norway.

[4] Institute of Clinical Medicine, Faculty of Health Sciences, UiT - The Arctic University of Norway, Hansine Hansens veg 18, 9019, Tromsø, Norway.

* Correspondance

Rafi Ahmad, Department of Biotechnology, Inland Norway University of Applied Sciences, Holsetgata 22, 2317 Hamar, Norway. Email: rafi.ahmad@inn.no

Running title: The three viewpoints of the transcriptome in human skeletal muscle

Abstract

The biological relevance and accuracy of gene expression data depend on the adequacy of data normalization. This is due both to its role in resolving and accounting for technical variation and errors, and to its defining role in shaping the biological viewpoint of data interpretations. Yet, normalization is often treated in serendipitous manners. This is especially true for its impact on the viewpoint of biological interpretations, and may have exacerbated consequences for conclusions in studies involving pronounced cellular plasticity.

In this case report, we highlight the consequences of using three fundamentally different normalization approaches for interpreting RNA-seq data from human skeletal muscle undergoing exercise-training induced growth. Briefly, 25 participants conducted 12 weeks of high-load resistance training. Muscle biopsy specimens were sampled from m. vastus lateralis before, after two weeks of training (week 2) and after the intervention (week 12), and were subsequently analyzed using RNA-seq. Transcript counts were modeled as i) per-library-size, ii) per-total-RNA, and iii) per-sample-size (per-mg-tissue-mass). Initially, this led to identification of three sets of stable genes, which were subsequently used for calculation of normalization factors and modelling of training-associated changes in whole-transcriptome transcript expression.

The three modes of transcript counts led to identification of different sets of stable genes. In general, genes that showed stable expression in the per-library-size dataset, showed increased expression in per-total-RNA and per-sample-size datasets. This led to three deviating sets of normalization factors, eventually leading to differential estimates of training-associated changes in transcript expression. Briefly, while 33% and 26% of transcripts showed no differences between normalization modes at weeks 2 and 12, respectively, 28 and 24% of transcripts showed changes in per-total-RNA and per-sample-size scenarios only. At week 2, this was manifested as opposite conclusions for 5% of the transcripts between per-library-size and per-sample-size data (↑ and ↓, respectively). Most down-regulated transcripts (7% and 4%) were found exclusively using per-library-size normalization.

Scientists should be explicit with their choice of normalization strategies and should interpret results of gene expression analyses with caution. This is particularly important for data sets involving a limited number of genes and involving growing or differentiating cellular models, where the risk of biased conclusions are likely elevated.

Keywords: RNA-seq, skeletal muscle, normalization, resistance training

Introduction

In gene expression analyses, data normalization can be performed using many different approaches, acting as a major determinant of the validity and reliability of interpretations [1–4]. For any data set, the available normalization strategies are partially predetermined by the technique used for data acquisition. Still, normalization always involves a myriad of explicit choices that may profoundly affect the analytical outcome. For example, for studies utilizing quantitative PCR (qPCR), the selection of internal reference genes will largely define downstream analyses and conclusions, and the utilization of non-validated reference genes will lead to substantial bias [2]. Analogous to this, for studies involving RNA sequencing, appropriate library size scaling will determine samples' comparability in downstream statistical analysis [3, 4]. Overall, data normalization essentially targets sources of technique-specific artifacts and non-biological variation. In addition, it also defines the biological perspective from which data are interpreted [1, 5]. This means that the choice of normalization mode will define the biological output of the experiment. Indeed, the expression of mRNA transcripts can be quantified using three distinctly different approaches, relative to the overall mRNA pool (i, i.e. using geometric averaging), the total amount of RNA (ii), or the amount of tissue or number of cells used in the experiment (iii) [1, 5].

While each of these perspectives holds biological merit, providing gene expression data that can be interpreted and compared between samples (e.g., changes from before to after a specific treatment), they do so in different manners. First, the per-library-size approach provides data that inform on the relative abundances of transcripts relative to all other transcripts, arguably enabling assessment of transcript expression that compares their ability to compete for slots on ribosomes. Second, the per-total-RNA approach provides data that inform on the absolute level of transcripts relative to the entire pool of RNA, enabling assessment of transcript expression that compares their ability to recognize and bind to ribosomes. Third, the per-sample-size approach provides data that inform on the overall abundances of transcripts in the biological sample at hand, and thus their content per-cell or per-tissue weight. Therefore, the three different normalization scenarios provide different biological perspectives, setting the stage for data interpretations. These differences will be exacerbated in experimental models and designs involving large cellular perturbations, accompanying changes in the overall mRNA and total RNA expression [5–7]. Despite this, the consequence of a specific normalization strategy for

data interpretation is rarely explicitly addressed in the biomedical literature, even though it represents an old and ever-present issue in mRNA-based analyses [1, 5].

The present study aimed to investigate the consequences of using each of the three modes of normalization (per-library-size, per-total-RNA, and per-sample-size) for transcriptome profiling of RNA-seq data in a highly plastic model of human biology. Briefly, twenty-five human participants conducted twelve weeks of high-load resistance training using a within-participant study design. Each participant performed exercise training with either low or moderate volume, allocated to either leg [8]. Overall, both study protocols led to substantial changes in muscle strength, mass, and biology, with the latter being evaluated from bilateral muscle biopsies (m. vastus lateralis) sampled at baseline and after two and twelve weeks of training. Biopsy samples showed marked increases in overall total RNA and mRNA abundances, arguably making it an adequate experimental system for the proposed comparisons [8, 9]. In the current analyses, the first objective was to identify a subset of gene transcripts that show relative stability within participants across all time points, measured as transcript abundances per-library-size, per-total-RNA, and per-sample-size, respectively. Secondly, we used the resulting reference gene sets to normalize the entire RNA-seq dataset, ultimately providing estimates of transcript abundances corresponding to each of the three perspectives of normalization.

Methods

Study overview

Thirty-four participants completed a 12-week progressive resistance training intervention with legs randomly allocated to either low (one set per exercise) or moderate-volume (three sets per exercise) training, as previously described [8]. The training intervention consisted of leg-press, leg- curl, and knee-extension. Bilateral muscle biopsies were obtained before the intervention and after two and twelve weeks of training. Total RNA was extracted from the biopsy material (TRIzol, ThermoFisher Scientific, Oslo, Norway)[8], and samples were selected for analysis based on RNA integrity scores. Twenty-five participants had a complete set of samples with integrity scores ≥ 7 (Average RQI 9, SD: 0.4; Experion Automated Electrophoresis Station using RNA StdSens Assay, Bio-Rad, Norway) and were selected for the RNA-sequencing experiment [9]. A fixed amount of total RNA (1000 ng) was used for RNA-seq library preparations and subjected to Paired-end sequencing (Illumina HiSeq 3000, Illumina, San Diego, CA USA), as detailed elsewhere [9].

Filtering and read alignment

Trimmomatic (version 0.39) [10] was used before alignment to filter out low-quality reads, as well as to remove poor-quality bases and adaptor sequence using default settings. The quality of filtered files was calculated using FastQC (version 0.11.4). After quality filtering, reads were aligned to the human genome and quantified on the level of transcripts using RSEM (version 1.3.1) [11] and GRCh38 release-97 (downloaded from ftp.ensemble.org).

Identification of stable reference genes and modeling of transcript counts

We assumed that expressing transcripts to crude estimates of each biological denominator (mRNA, total RNA, or sample size) would introduce bias in downstream analyses due to incomparable scaling and measurement errors between normalization modes. Therefore, we selected internal reference genes to create comparable normalization factors between normalization modes for subsequent analyses. Reference genes were selected from a subset of transcripts with robust expression across all samples, filtered with the minimum count per sample set to 30. After filtering, 5687 genes remained in the data set for assessment of within-participant stability. Identification of stable transcripts in each of the normalization scenarios (per library size, per total RNA, and per sample size) was made using transcript counts transformed to counts per million (CPM). CPM values were calculated after expressing counts per scaled library size (total counts scaled by trimmed mean of M-values [4]), amount of tissue (mg of tissue $\times 10^6$), or total RNA (1×10^6, assuming equal total RNA in each reaction) used in library preparation. CPM values were log-transformed before being fitted to linear mixed-effects models on a target-by-target basis. Models were subsequently used to (1) assess the effects of the intervention on transcript abundances over time and between volume conditions and to (2) determine the intra class correlation coefficient (ICC) defined as the amount of variance attributed to between-participant variation relative to the total variance. For each normalization mode, maximal t-values calculated from model coefficients representing study conditions (time and exercise volume) were used to remove transcripts that showed indices of systematic effects of the intervention. Transcripts with absolute t-values 1.5 were subjected to subsequent ranking based on ICC values. The top ten transcripts from each normalization mode were defined as stable reference genes, suitable for the calculation of normalization denominators. Selected transcripts were scaled ($x_1/max(x)$) and averaged per sample to form the sample reference. To compare sample references from each normalization mode, ratios were evaluated over time from estimates obtained from linear mixed-effects models.

The full set of transcripts (excluding reference genes, filtered with minimum count = 1, n = 12066) was modeled on a target-by-target basis using negative binomial generalized linear mixed models (GLMM)[9, 12] with normalization mode-specific reference genes used as offsets in each model fit to express gene counts per-library-size, per-total-RNA and per-sample-size. Model fits were used to assess the effects of study conditions on relative gene counts. For the sake of this analysis, samples from each leg were considered biological replicates to determine the impact of exercise training per se (time-effects). Differentially expressed genes were defined if significantly different from baseline on a target-by-target basis (unadjusted $p < 0.05$).

After filtering of low abundance transcripts (minimum count per sample = 30), 5687 remained in the data set for further assessment of within participant stability. Gene counts were transformed to express gene counts per the total scaled library size, per input total RNA, and per tissue mass. As a fixed amount of total RNA was used for library preparation, the raw counts were used to express counts per total RNA. The amount of tissue used in library preparations was used as the denominator in transforming transcript counts to counts per tissue mass.

Transformed transcript counts were subsequently fitted to linear mixed-effects models on a target-by-target basis to (1) assess the effects of the intervention over time and between volume conditions, and (2) to determine the intraclass correlation coefficient (ICC) defined as the amount of variation attributed to between-participant variation. Maximal *t*-values from each model coefficients representing time and time/volume-condition interactions were used to filter transcripts with evidence of systematic effects. Transcripts with absolute *t*-values ≤ 1.5 were further considered for ICC. Transcripts were subsequently ranked based on their ICC and the top ten transcript from each normalization scenario were used as internal references. Selected transcripts were scaled ($x_1/max(x)$) and averaged per sample to form the sample reference.

Sample references from each normalization scenario were used as an offset in negative binomial generalized linear mixed models (GLMM) to assess the effects of study conditions of gene counts of non-reference genes from the pool of high abundance transcripts (n = 5687).

Results

The training intervention led to robust increases in muscle mass and strength (on average 4% and 25%, respectively [8]). This was accompanied by an increase in total RNA from baseline to weeks two and twelve (on average 27% and 17%, respectively) and an increase in the

sequenced library size, despite a lower amount of tissue being used during library preparations [8, 9].

In the present RNA-seq data set, the three modes of normalization led to the identification of three different clusters of stable gene transcripts across untrained and resistance-trained muscle specimens. Each of the three gene clusters consisted of genes that showed unaltered expression across muscle biopsies sampled from each participant (and as such was not affected by the resistance training). The number of genes in the three clusters varied substantially (per-library-size, n=1266; per-total-RNA, n=90; per-sample-size, n=18), with per-library-size normalization being associated with higher ICC estimates, suggesting higher degrees of consistency between samples. The top ten most stable transcripts from each normalization scenario were then identified based on ICC estimates (Table 1). There was no overlap between the per-library-size cluster and either of the two other clusters among these transcripts. At the same time, per-total-RNA and per-samples-size data sets displayed two overlapping transcripts (Table 1).

Sample references, computed as the scaled average of the ten most stable transcripts per normalization mode, were analyzed as ratios to explore dissimilarities over time. Stable transcripts identified in the per-library-size dataset, showed clearly increased expression estimates in per-total-RNA and per-sample-size datasets (Figure 1A, right panels). Similarly, stable genes identified in the per-sample-size dataset, showed decreased expression in the per-total-RNA dataset (Figure 1A, upper left panel). These differences were more pronounced in samples obtained after two weeks of resistance training (Figure 1A).

We further used sample references to identify differentially expressed transcripts from baseline to after training for each of the three normalization modes. Most differentially expressed transcripts were shared among all three normalization modes (33% and 26% from baseline to week two and twelve, respectively, Figure 1B). However, a substantial fraction of differentially expressed targets were exclusively identified in per-total-RNA and per-sample-size datasets (Figure 1B). A smaller fraction of differentially expressed transcripts were identified in the per-sample-size mode only. Most transcripts identified as differentially expressed were identified with higher expression at weeks two and twelve. However, transcripts identified with lower expression were predominantly identified on a per-library-size basis. At week 2, 5% of all transcripts were identified as down-regulated using per-library-size normalization but up-regulated when expressed per-sample-size (Figure 1B).

208 *Table 1. Genes selected as stable reference genes from each normalization scenario.*

Normalization strategy	Transcript ID	Gene Symbol	Gene biotype	Intraclass correlation
Per-library-size	ENST00000643905			0.9157487
	ENST00000439211	DHFR	protein coding	0.8773145
	ENST00000582787	SP2-DT	lncRNA	0.8733079
	ENST00000342992	TTN	protein coding	0.8661879
	ENST00000361681	MT-ND6	protein coding	0.8643855
	ENST00000371470	MAGOH	protein coding	0.8469231
	ENST00000234256	SLC1A4	protein coding	0.8425609
	ENST00000341162	FCF1	protein coding	0.8411509
	ENST00000480046	METTL2B	protein coding	0.8399730
	ENST00000295955	RPL9	protein coding	0.8289768
Per-total-RNA)	ENST00000445125		processed pseudogene	0.7159682
	ENST00000312184	TMEM70	protein coding	0.5795889
	ENST00000552002	CHURC1	protein coding	0.5598877
	ENST00000357033	DMD	protein coding	0.5592342
	ENST00000275300	SLC22A3	protein coding	0.5557768
	ENST00000496823	BCL6	protein coding	0.5488948
	ENST00000546248	TRDN	protein coding	0.5222352
	ENST00000309881	CD36	protein coding	0.5059408
	ENST00000005178	PDK4	protein coding	0.4966191
	ENST00000522603	ASPH	protein coding	0.4927794
Per-sample-size	ENST00000496823	BCL6	protein coding	0.5369798
	ENST00000546248	TRDN	protein coding	0.4976999
	ENST00000216019	DDX17	protein coding	0.4619666
	ENST00000005178	PDK4	protein coding	0.4589237
	ENST00000361915	AGL	protein coding	0.4219960
	ENST00000418381			0.4167778
	ENST00000294724	AGL	protein coding	0.4050706
	ENST00000366645	EXOC8	protein coding	0.3911603
	ENST00000261349	LRP6	protein coding	0.3841484
	ENST00000306270	IBTK	protein coding	0.3287734

Discussion

The present study shows that the choice of normalization modality will affect the outcome of gene expression analyses in models of load-induced skeletal muscle plasticity in humans. The three modes of normalization (per-library-size, per-total-RNA, and per-sample size) were associated with different patterns of differentially expressed transcripts in response to resistance training. In general, per-library-size-based normalization underestimated increases in mRNA abundances compared to the other two approaches. Such underestimation is an effect of the assumption that most transcripts/genes remain unchanged in response to treatment, an assumption corresponding to using the library size as the point of reference in gene expression studies [4, 13]. Indeed, in the extreme case of global scale transcription amplification in response to a treatment [6], the assumption of relative transcription stability between samples leads to underestimation of transcript expression, and thus potentially erroneous conclusions. This is the case in studies of resistance training-induced muscle hypertrophy, where ribosome biogenesis, and thus increased ribosomal RNA transcription and total RNA expression, are known to occur [8, 14, 15]. In such studies, neglecting to account for changes in total RNA expression during analyses of transcript data will lead to underestimation of transcript abundances per tissue mass [14].

While the results from the present study indicate the impact of using different normalization strategies, they do not inform us on the rights and wrongs of normalization choices. The correct use of a specific normalization strategy depends on the research questions and model systems under study. For example, in systems where relative transcript abundances are of primary interest or in models where whole-transcriptome expression remain stable, it seems prudent to assume that transcript abundances will both relative to other transcripts (per-library-size) and in absolute terms (per-sample-size) [5]. However, as exemplified in the present study, in systems where these assumptions are not true, and global changes occur in variables such as total RNA and global mRNA expression, absolute expression of transcript may be a more suitable outcome. In such studies, the selection of normalization strategy that violate these assumptions will lead to biased conclusions [5]. In general, studies involving gene expression seldom account for these perspectives during data analyses [5], and their consequences for data interpretations remain underappreciated, reducing the ecological validity of analyses and making comparisons between studies difficult. To alleviate this, researchers should explicitly state and ideally test their assumptions regarding choices of normalization strategies.

Conclusion

In the present study, we show that the choice of normalization modality (per-library-size vs per-total-RNA vs per-sample-size) has pronounced effects on interpretations of transcriptome responses in a human model of load-induced skeletal muscle plasticity. For any study of gene expression, data normalization should be conducted and evaluated with care and intent, ensuring the stability and representativity of normalization denominators, and importantly also the biological relevance of outcome measures.

References

1. Huggett J, Dheda K, Bustin S, Zumla A. Real-time RT-PCR normalisation; strategies and considerations. Genes Immun. 2005;6:279–84. doi:10.1038/sj.gene.6364190.

2. Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, et al. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol. 2002;3:RESEARCH0034. http://www.ncbi.nlm.nih.gov/pubmed/12184808.

3. Anders S, Huber W. Differential expression analysis for sequence count data. Genome Biology. 2010;11:R106. doi:10.1186/gb-2010-11-10-r106.

4. 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.

5. Coate JE, Doyle JJ. Variation in transcriptome size: Are we getting the message? Chromosoma. 2015;124:27–43. doi:10.1007/s00412-014-0496-3.

6. 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.

7. Hansen MC, Nielsen AK, Molin S, Hammer K, Kilstrup M. Changes in rRNA levels during stress invalidates results from mRNA blotting: Fluorescence in situ rRNA hybridization permits renormalization for estimation of cellular mRNA levels. J Bacteriol. 2001;183:4747–51. doi:10.1128/jb.183.16.4747-4751.2001.

8. 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.

9. Khan Y, Hammarström D, Rønnestad BR, Ellefsen S, Ahmad R. Increased biological relevance of transcriptome analyses in human skeletal muscle using a model-specific pipeline. BMC Bioinformatics. 2020;21:548. doi:10.1186/s12859-020-03866-y.

10. 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.

11. 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.

12. 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.

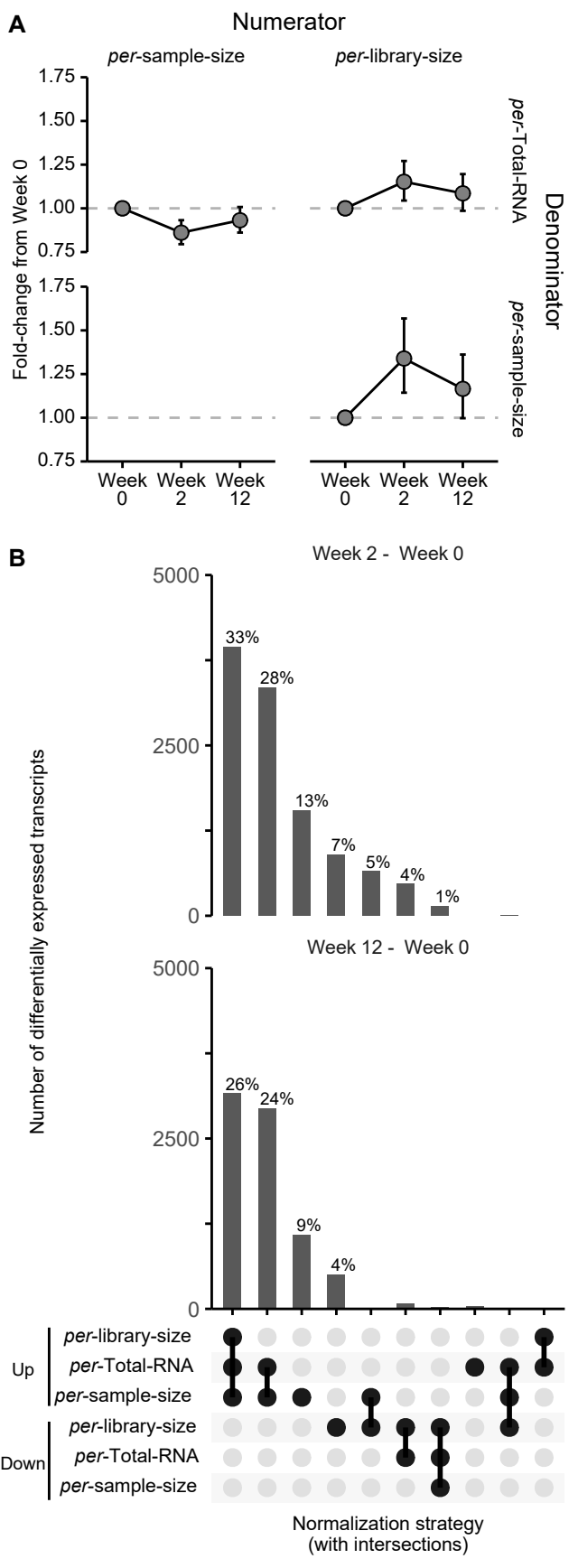
13. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology. 2014;15:550. doi:10.1186/s13059-014-0550-8.

14. Chaillou T, Malgoyre A, Banzet S, Chapot R, Koulmann N, Pugniere P, et al. Pitfalls in target mRNA quantification for real-time quantitative RT-PCR in overload-induced skeletal muscle hypertrophy. Physiol Genomics. 2011;43:228–35. doi:10.1152/physiolgenomics.00109.2010.

15. Figueiredo VC, McCarthy JJ. Regulation of ribosome biogenesis in skeletal muscle hypertrophy. Physiology (Bethesda). 2019;34:30–42.

Figure captions

Figure 1. (A) Fold-changes of sample references (average of the top ten stable genes per normalization mode) ratios with numerators plotted over columns and denominators over rows. Error bars represent 95% CI. (B) Genes identified as differentially up and down regulated over time (differences from Week 0 to Week 2 and 12 respectively) from generalized linear models with each normalization factor used as an model offset. Up- and down-regulation determined from un-adjusted *p*-values ($p < 0.05$).


A
Numerator
per-sample-size
per-library-size
Denominator
per-Total-RNA
per-sample-size
Fold-change from Week 0
1.75
1.50
1.25
1.00
0.75
Week 0
Week 2
Week 12
B
Week 2 - Week 0
Week 12 - Week 0
Number of differentially expressed transcripts
5000
2500
0
33%
28%
13%
7%
5%
4%
1%
26%
24%
9%
4%
Up
Down
per-library-size
per-Total-RNA
per-sample-size
Normalization strategy
(with intersections)

3

Mølmen *et al. J Transl Med* (2021) 19:292
https://doi.org/10.1186/s12967-021-02969-1

Journal of Translational Medicine

RESEARCH Open Access

Chronic obstructive pulmonary disease does not impair responses to resistance training

Knut Sindre Mølmen[1*], Daniel Hammarström[1], Gunnar Slettaløkken Falch[1], Morten Grundtvig[2], Lise Koll[3], Marita Hanestadhaugen[3], Yusuf Khan[1,4], Rafi Ahmad[4,5], Bente Malerbakken[6], Tore Jørgen Rødølen[7], Roger Lien[7], Bent R. Rønnestad[1], Truls Raastad[8] and Stian Ellefsen[1,9]

Abstract

Background: Subjects with chronic obstructive pulmonary disease (COPD) are prone to accelerated decay of muscle strength and mass with advancing age. This is believed to be driven by disease-inherent systemic pathophysiologies, which are also assumed to drive muscle cells into a state of anabolic resistance, leading to impaired abilities to adapt to resistance exercise training. Currently, this phenomenon remains largely unstudied. In this study, we aimed to investigate the assumed negative effects of COPD for health- and muscle-related responsiveness to resistance training using a healthy control-based translational approach.

Methods: Subjects with COPD (n = 20, GOLD II-III, $FEV_{1predicted}$ 57 ± 11%, age 69 ± 5) and healthy controls (Healthy, n = 58, $FEV_{1predicted}$ 112 ± 16%, age 67 ± 4) conducted identical whole-body resistance training interventions for 13 weeks, consisting of two weekly supervised training sessions. Leg exercises were performed unilaterally, with one leg conducting high-load training (10RM) and the contralateral leg conducting low-load training (30RM). Measurements included muscle strength ($n_{variables}$ = 7), endurance performance ($n_{variables}$ = 6), muscle mass ($n_{variables}$ = 3), muscle quality, muscle biology (*m. vastus lateralis*; muscle fiber characteristics, RNA content including transcriptome) and health variables (body composition, blood). For core outcome domains, weighted combined factors were calculated from the range of singular assessments.

Results: COPD displayed well-known pathophysiologies at baseline, including elevated levels of systemic low-grade inflammation ($[c\text{-reactive protein}]_{serum}$), reduced muscle mass and functionality, and muscle biological aberrancies. Despite this, resistance training led to improved lower-limb muscle strength (15 ± 8%), muscle mass (7 ± 5%), muscle quality (8 ± 8%) and lower-limb/whole-body endurance performance (26 ± 12%/8 ± 9%) in COPD, resembling or exceeding responses in Healthy, measured in both relative and numeric change terms. Within the COPD cluster, lower $FEV_{1predicted}$ was associated with larger numeric and relative increases in muscle mass and superior relative improvements in maximal muscle strength. This was accompanied by similar changes in hallmarks of muscle biology such as rRNA-content↑, muscle fiber cross-sectional area↑, type IIX proportions↓, and changes in mRNA transcriptomics. Neither of the core outcome domains were differentially affected by resistance training load.

Conclusions: COPD showed hitherto largely unrecognized responsiveness to resistance training, rejecting the notion of disease-related impairments and rather advocating such training as a potent measure to relieve pathophysiologies.

*Correspondence: knut.sindre.molmen@inn.no
[1] Section for Health and Exercise Physiology, Inland Norway University of Applied Sciences, P.O. Box 422, 2604 Lillehammer, Norway
Full list of author information is available at the end of the article

Trial registration: ClinicalTrials.gov ID: NCT02598830. Registered November 6th 2015, https://clinicaltrials.gov/ct2/show/NCT02598830

Keywords: Anabolic resistance, COPD, Pathophysiology, Skeletal muscle, Strength training, Training load

Introduction

Chronic obstructive pulmonary disease (COPD) is associated with impaired cardiorespiratory fitness and decreased skeletal muscle mass and strength [1], leading to reduced levels of daily activity and reduced quality of life [2, 3]. This deterioration is accompanied by systemic co-morbidities such as reduced levels of testosterone [4], vitamin D [5, 6] and oxygen saturation levels [7], and elevated levels of low-grade inflammation [8], which arguably leaves COPD subjects in a state of anabolic resistance [9], resulting in impaired abilities to adapt to exercise training [10–12]. In particular, these pathophysiologies are believed to impair adaptations to resistance training, which represent the most potent intervention for improving muscle functions [13–16] and preventing escalation into late-stage morbidities such as pulmonary cachexia [17]. Despite this general belief, the presence of anabolic resistance in COPD subjects and its consequences for responses to resistance training remain circumstantial. A mere single study has compared functional and biological adaptations to resistance training between COPD and healthy controls (ISRCTN ID: 22764439) [18–20], and as such was limited by a relatively short training intervention (8 weeks), a rather untraditional training protocol with little clinical and practical relevance, and a limited selection of outcome variables. Whereas the study failed to disclose COPD-related impairments in muscle strength and growth responses, it seems premature to dismiss the notion that COPD pathophysiologies may impair training responsiveness [21], and there is clearly need for further study.

The primary aim of the present study was to investigate the assumed negative effects of COPD pathophysiologies on physiological responses to 13 weeks of resistance training, with emphasis on a broad range of muscle functional and biological outcome measures. The secondary aim was to investigate inherent differences between COPD and Healthy, and to investigate the interaction between two different resistance training modalities and training responsiveness (high-load vs. low-load resistance training; 10 vs 30 repetitions maximum, RM).

Methods

For in-depth description of study protocols and methods, including description of a placebo-controlled vitamin D_3 supplementation protocol (randomized clinical trial), see Figs. 1, 2 and clinicaltrial.gov (ClinicalTrials.gov Identifier: NCT02598830). The study was designed and scaled to allow elucidation of the effects of vitamin D_3 supplementation for adaptations to resistance training, as well as to compare training responsiveness between COPD and Healthy. The vitamin D_3 perspective is covered in detail elsewhere [22].

Study ethics and participants

The study was approved by the Regional Committee for Medical and Health Research Ethics (reference no. 2013/1094), preregistered at clinicaltrials.gov (NCT02598830), and conducted according to the Declaration of Helsinki. All participants were informed about the potential risks and discomforts associated with the study and gave their informed consent prior to study enrolment.

Persons with either medical diagnosis of stable COPD (GOLD grade II-III [23], predicted forced expiratory volume in first second (FEV_1) between 80%-30%, FEV_1/forced vital capacity (FVC) < 70% after reversibility testing, n = 24, age 70 ± 5) or normal lung function (n = 70, age 67 ± 5) received the study intervention. For study flow chart, see Fig. 1. For baseline characteristics of the participants completing the study, see Table 1.

Study conduct

COPD and Healthy conducted identical 13-week resistance training protocols, consisting of two weekly full-body training sessions (Fig. 2) with primary focus on leg exercises. The leg exercises, i.e. leg press, knee extension and knee flexion, were performed unilaterally in that consecutive order, with one of the legs of each participant being randomly assigned to perform three sets of 10RM (high-load) and the contralateral leg to perform three sets of 30RM (low-load). For each exercise, all three sets for one leg were conducted before the other leg was exercised. This unilateral training protocol served two purposes: i) to circumvent issues relating to conduction of training with two-legged exercises in COPD [24] and ii) to investigate the relative efficacy of two different training modalities (10RM vs 30RM). Exercises and sets were separated by ~ 2 min of rest, with individual adjustments being made whenever participants needed a longer rest period. All sessions were supervised by qualified personnel and lasted for ~ 60 min. The effectiveness of the training intervention was assessed as a wide range of outcome measures (Fig. 2), including multiple assessments

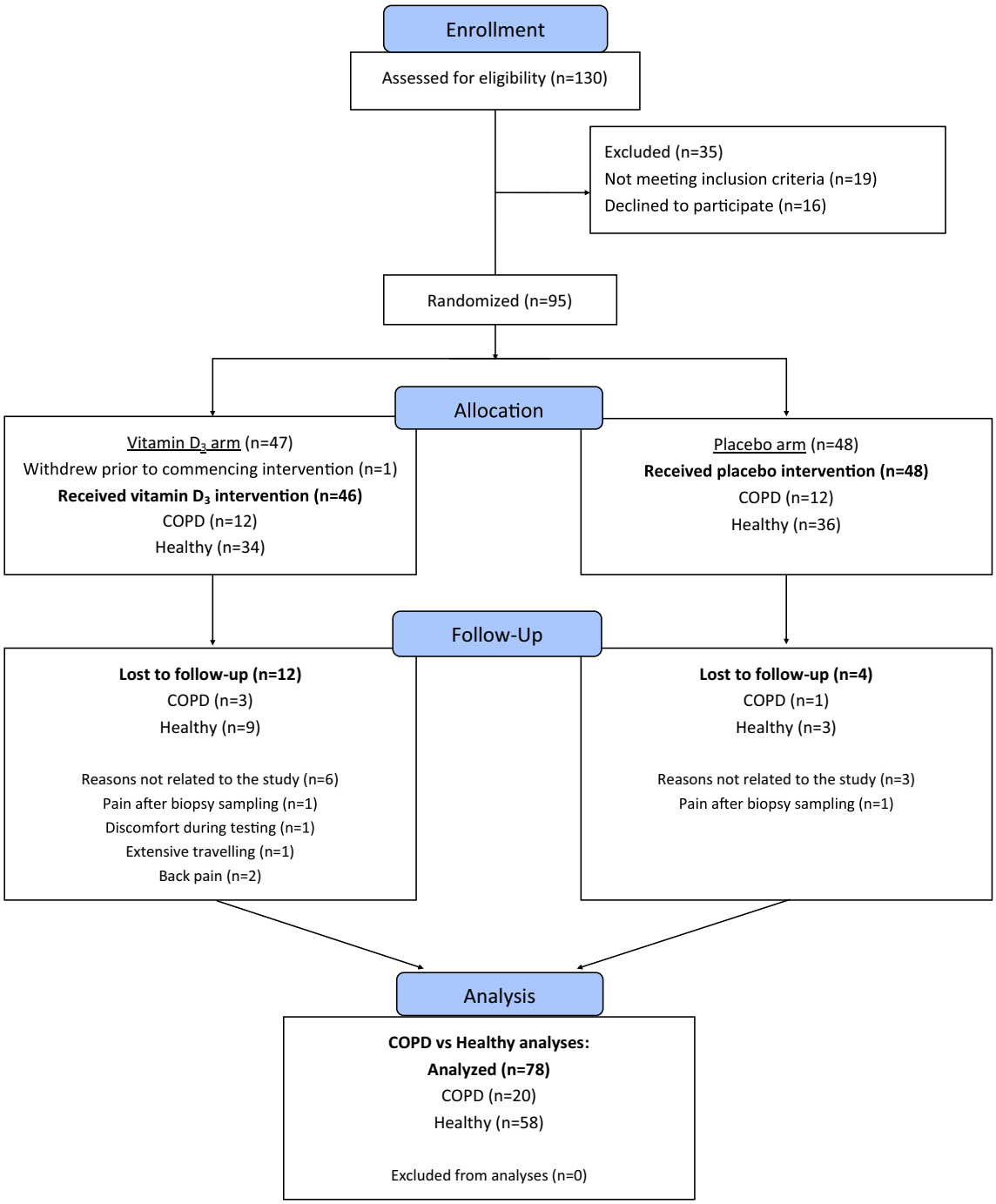


Fig. 1 CONSORT flow chart of the study. The study was conducted as a double-blind randomized clinical trial, with the primary aim of investigating the effects of vitamin D_3 supplementation on resistance training-associated adaptations in a mixed population of older subjects, including both COPD and healthy control subjects (COPD and Healthy, respectively) (ClinicalTrials.gov Identifier: NCT02598830). Vitamin D_3 supplementation did not affect any primary or secondary outcome, and no conditional effects were observed for COPD vs Healthy in that context [22]. In the present study, the main purpose was to compare the effects of resistance training between COPD and Healthy participants (number of participants completing the study protocol; n COPD = 20; n Healthy = 58)

of endurance performance, muscle strength and mass, measures of work economy/efficiency, and collection of blood and *vastus lateralis* biopsies (both legs) (Fig. 2).

Blood and muscle measurements

Prior to collection of blood and muscle biopsies, participants were instructed to attend an overnight fast and to avoid heavy physical activity for the last 48 h. Blood samples were analyzed for serum concentrations of hormones, lipids, and markers of iron metabolism and tissue damage, as previously described [22]. Muscle biopsies were analyzed for muscle fiber type proportions, myonuclei content, muscle fiber cross-sectional area (CSA), and rRNA and mRNA content (total RNA, rRNA subspecies, myosin heavy chain isoforms I, IIA and IIX, and whole-genome transcriptome), as previously described [22, 25, 26]. Transcriptome analysis was restricted to a subset of participants (COPD, n = 19 (n prior to resistance training, 19; n after 3 ½ week of training, 17; n post resistance training, 19); Healthy, n = 34 for all time points), selected based on quality of total RNA samples (RNA Quality Indicator > 7.0, avg 9.0 ± 0.5), with participants with COPD and participants with complete sets of muscle biopsies being prioritized.

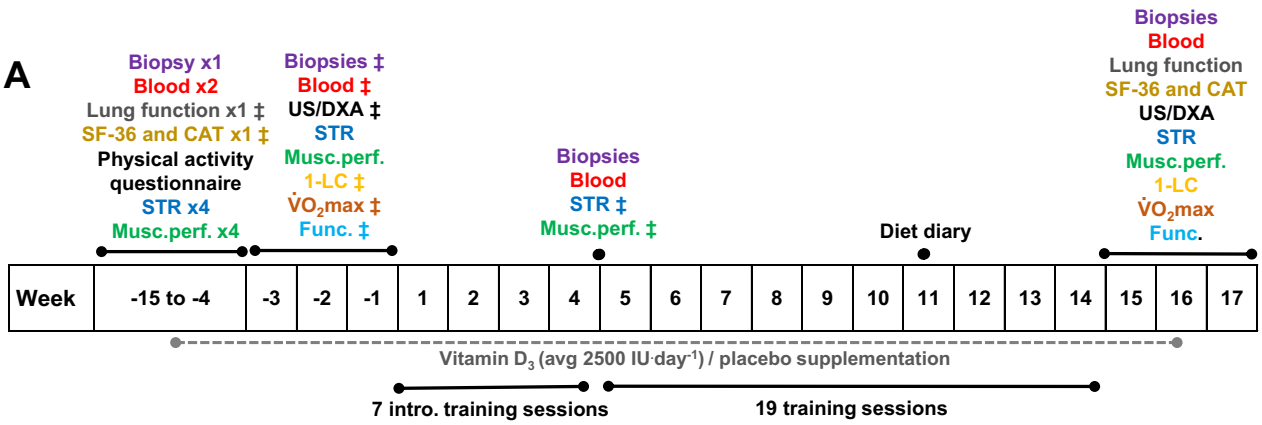
Data analyses and statistics

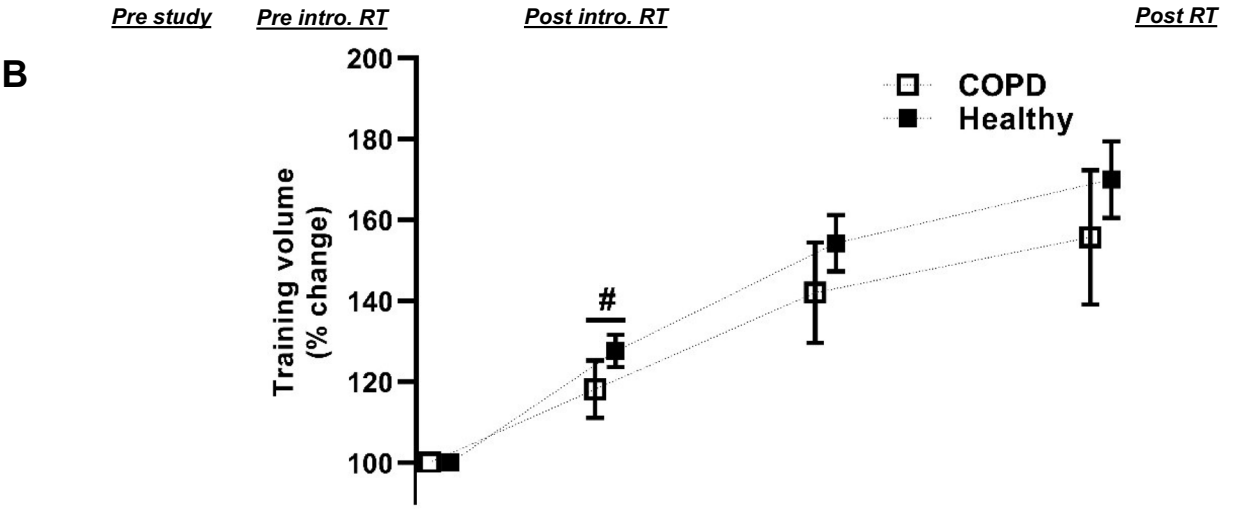
Analyses were conducted per-protocol, due to the translational approach of the study. For continuous variables, linear mixed-effects models were used to examine differences between COPD and Healthy, both at baseline and as responses to resistance training. For the latter, relative and numeric changes from baseline were defined as dependent variables, with COPD/Healthy being defined as the fixed effect. Effects of sex were implemented into the models. Analyses included evaluation of interaction effects with training load (repeated measures/observations from high- and low-load training legs were added to the model for unilateral outcome measures) and sex. Time effects were examined using mixed modelling, with the dependent variable and time points being defined as repeated measures/observations. To describe the relationship between COPD severity and training responses, simple linear regression analyses of core outcome domains change scores and predicted FEV_1 were performed.

For non-continuous variables, generalized linear mixed-effects models (GLMMs) were used (binomial GLMMs, immunohistochemical fiber type proportion analyses; negative binomial GLMMs, rRNA/mRNA content in quantitative real-time polymerase chain reaction (qPCR) and transcriptome analyses). For transcriptome analyses, gene counts were modelled using the total library size as a fixed effect [27], together with sex and study conditions (time points and COPD/Healthy). Models were iteratively fitted using glmmTMB [28]. Genes were regarded as differentially expressed when the numeric $\log_2$ fold-change/difference were greater than 0.5 and the adjusted *p*-value (false discovery rate adjusted *per* model coefficient) was below 5% [25]. Moreover, enrichment analyses were performed on Hallmark,

(See figure on next page.)

Fig. 2 Schematic overview of the study protocol, including its time line (**A**; ‡indicates the defined baseline measurement for the specific outcome measure), training volumes during the resistance training (RT) intervention (**B**), perceived exertion (Borg RPE, 6–20) reported after training sessions (**C**), and relative training loads (% of 1RM) during the training period (**D**). Training volume is presented as average increases in per-session for lower-body appendices from the first week of training (kg · repetitions; high-load (10RM) and low-load (30RM) leg press and knee extension combined). Training loads in numeric values (kg) during the resistance training intervention are provided in Additional file 1: Fig. S1. COPD, participants diagnosed with chronic obstructive pulmonary disease; Healthy, healthy control participants; *statistical different from 1th training week; #statistical difference between COPD and Healthy. Data are presented as means with 95% confidence limits. Methodological notes on retrieval of outcome measures: **i)** Lung function. Spirometry testing was performed following the guidelines from the American Thoracic Society and the European Respiratory Society [72]. Participants with COPD were tested before and after inhalation of two bronchodilators (salbutamol/ipratropiumbromid). **ii)** Muscle strength and performance (STR and Musc. perf). Muscle strength was assessed as one-repetition maximum (1RM) in unilateral knee extension and leg press, bilateral chest press, and handgrip. Muscle performance was defined as the number of repetitions achieved at 50% of pre-study 1RM and was assessed using unilateral knee extension and bilateral chest press. Isokinetic unilateral knee-extension torque was tested at three angular speeds (60°, 120° and 240°· sec^{-1}; Humac Norm, CSMi, Stoughton, MA, USA). **iii)** One-legged cycling and bicycling performance (1-LC and VO_2max). Participants conducted one-legged cycling tests (Excalibur Sport, Lode BV, Groningen, the Netherlands) to assess O_2-costs and mechanical efficiency [73] during submaximal cycling, and maximal one-legged oxygen consumption ($\dot{V}O_2$max) and maximal workload. Maximal two-legged cycling $\dot{V}O_2$max and workload were tested on a separate day. Oxygen consumption was measured using the JAEGER Oxycon Pro™ system (Carefusion GmbH, Höchberg, Germany). **iv)** Functional performance (Func.). Functional tests were conducted as the maximal number of sit-to-stands during one minute (seat height: 45 cm) and as the number of steps onto a 20 cm step box during 6 min. **v)** Health-related quality of life (SF-36 and CAT). All participants completed the Short Form (36-item) Health Survey (SF-36). COPD participants also completed the COPD Assessment Test (CAT) questionnaire. **vi)** Muscle thickness and body mass composition (US/DXA). Muscle thickness of *m. vastus lateralis* and *m. rectus femoris* were measured using B-mode ultrasonography (SmartUs EXT-1 M, Telemed, Vilnius, Lithuania). Body mass composition was measured using dual-energy X-ray absorptiometry (DXA; Lunar Prodigy, GE Healthcare, Madison, WI, USA). At pre study, all participants completed a questionnaire regarding regular weekly activity habits. The results (time spent for different activities) were translated into energy expenditure ($kcals{\cdot}week^{-1}$) during activities using number of metabolic equivalents provided in Jetté et al. [74]. During week 11, all participants conducted a dietary registration, in which they logged their dietary intake for three days, including one weekend day





C

Perceived exertion (6-20)				
COPD	15.2 ± 1.0	15.7 ± 1.1	16.4 ± 1.0 *	16.8 ± 1.2 *
Healthy	15.5 ± 1.6	16.1 ± 1.4 *	16.6 ± 1.5 *	17.1 ± 1.2 *

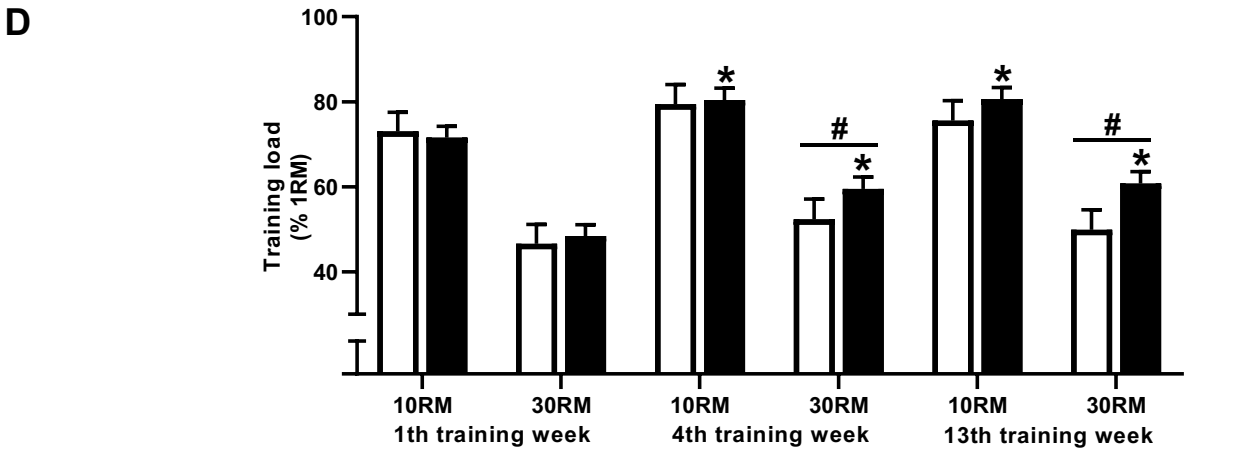


Fig. 2 (See legend on previous page.)

Table 1 Characteristics of the participants completing the study

	COPD	Healthy	Sex-adjusted estimated difference	
			COPD – Healthy (95% CI)	*P*-value
General				
Participants, completing (no. ♂/♀) / dropouts† (no.)	20 (12/8) / 2	58 (21/37) / 2	–	–
Age (years)	69 ± 5 (range, 60–79)	67 ± 4 (range, 57–78)	2 (0, 5)	0.049*
Height (cm)	171 (10)	170 (10)	−3 (−6, 0)	0.056
Body mass (kg)	73 (18)	76 (16)	−7 (−14, 0)	0.061
Body mass index (kg· m^2)	25 (5)	26 (5)	−2 (−4, 1)	0.237
Pack-years (no.)	30 (16)	6 (10)	23 (17, 29)	< 0.001*
GOLD grade (no. of grade II/III)	15/5	–	–	–
COPD Assessment Test™ score (0–40)	16.6 (6.8)	–	–	–
Self-reported conception of health (0–10)	4.9 (1.2)	6.7 (1.6)	− 1.7 (− 2.5, − 0.7)	0.001*
Physical activity level				
Household work (kcals·$week^{-1}$)	1754 (2062)	1866 (2201)	− 164 (− 1322, 995)	0.779
Recreational activities (kcals·$week^{-1}$)	2512 (2619)	2654 (1841)	188 (− 862, 1237)	0.723
Total activity (kcals·$week^{-1}$)	4266 (4036)	4520 (2837)	24 (− 1657, 1704)	0.978
Pulmonary function				
FVC (L)	3.2 (0.9)	3.6 (0.9)	− 0.7 (− 1.0, − 0.4)	< 0.001*
FVC (% predicted)	97 (19)	112 (16)	− 13 (− 22, − 4)	0.003*
FEV_1 (L· sec^{-1})	1.5 (0.4)	2.7 (0.7)	− 1.4 (− 1.6, − 1.2)	< 0.001*
FEV_1 (% predicted)	57 (11)	104 (16)	− 47 (− 55, − 39)	< 0.001*
FEV_1/FVC (%)	47 (8)	75 (6)	− 28 (− 31, − 24)	< 0.001*
PEF (L· sec^{-1})	5.0 (1.6)	7.7 (2.1)	− 3.4 (− 4.1, − 2.7)	< 0.001*
Pulmonary medication				
B_2-agonists (no.)	17/20	–	–	–
Muscarinic agonists (no.)	15/20	–	–	–
Medication containing both b_2-agonist and glucocorticoid (no.)	10/20	–	–	–
Body composition				
Total lean mass (kg)	♂, 53 (4); ♀, 36 (6)	♂, 60 (5); ♀, 41 (4)	− 6 (− 9, − 4)	< 0.001*
Whole-body bone mineral density (g· cm^2)	♂, 1.2 (0.1); ♀, 1.0 (0.2)	♂, 1.3 (0.1); ♀, 1.1 (0.1)	− 0.1 (− 0.2, − 0.0)	0.007*
Total fat mass (kg)	♂, 26 (10); ♀, 27 (15)	♂, 26 (9); ♀, 25 (10)	1 (− 5, 7)	0.703
Visceral fat (kg)	♂, 1.9 (1.3); ♀, 1.0 (0.7)	♂, 1.7 (1.0); ♀, 0.8 (0.7)	0.2 (− 0.3, 0.7)	0.412
Lower—-body muscle strength				
1RM leg press (kg)	♂, 121 (35); ♀, 82 (21)	♂, 152 (27); ♀, 124 (25)	− 36 (− 47, − 26)	< 0.001*
1RM knee extension (kg)	♂, 21 (4); ♀, 11 (4)	♂, 31 (5); ♀, 16 (3)	− 7 (− 9, − 5)	< 0.001*
Peak torque knee extension 60° · sec^{-1} (Nm)	♂, 127 (34); ♀, 80 (25)	♂, 160 32); ♀, 101 (16)	− 27 (− 36, − 17)	< 0.001*
Peak torque knee extension 180° · sec^{-1} (Nm)	♂, 83 (25); ♀, 47 (17)	♂, 102 (23); ♀, 62 (11)	− 19 (− 28, − 9)	< 0.001*
Peak torque knee extension 240° · sec^{-1} (Nm)	♂, 68 (20); ♀, 38 (14)	♂, 84 (20); ♀, 50 (9)	− 15 (− 20, − 9)	< 0.001*
Lower-body muscle strength factor (AU)	♂, 0.5 (0.1); ♀, 0.3 (0.1)	♂, 0.6 (0.1); ♀, 0.4 (0.1)	− 0.1 (− 0.2, − 0.1)	< 0.001*
Lower-body muscle mass measures				
Leg lean mass (kg)	♂, 18 (2); ♀, 12 (3)	♂, 20 (2); ♀, 14 (2)	− 3 (− 4, − 2)	< 0.001*
M. vastus lateralis thickness (mm)	♂, 20 (3); ♀, 18 (5)	♂, 22 (3); ♀, 20 (3)	− 2 (− 3, − 1)	0.002*
M. rectus femoris thickness (mm)	♂, 13 (4); ♀, 10 (3)	♂, 16 (4); ♀, 15 (4)	− 4 (− 5, − 2)	< 0.001*
Lower-body muscle mass factor (AU)	♂, 0.6 (0.1); ♀, 0.5 (0.1)	♂, 0.7 (0.1); ♀, 0.6 (0.1)	− 0.1 (− 0.2, − 0.1)	< 0.001*
Endurance measures				
Maximal power output one-legged cycling (W)	♂, 73 (13); ♀, 48 (17)	♂, 148 (28); ♀, 108 (21)	− 67 (− 77, − 58)	< 0.001*
Maximal power output two-legged cycling (W)	♂, 118 (38); ♀, 75 (32)	♂, 252 (48); ♀, 167 (32)	− 113 (− 134, − 92)	< 0.001*
Maximal oxygen consumption (mL O_2· kg^{-1}· min^{-1})	♂, 20 (5); ♀, 16 (5)	♂, 35 (7); ♀, 28 (6)	− 14 (− 18, − 10)	< 0.001*
6 min step test (maximal number of steps)	♂, 123 (35); ♀, 115 (44)	♂, 208 (41); ♀, 196 (38)	− 83 (− 105, − 61)	< 0.001*

Table 1 (continued)

	COPD	Healthy	Sex-adjusted estimated difference	
			COPD – Healthy (95% CI)	*P*-value
1 min sit-to-stand test (maximal number)	♂, 21 (5); ♀, 21 (6)	♂, 30 (5); ♀, 29 (5)	− 9 (− 12, − 6)	< 0.001*
$n_{repetitions}$ at 50% of 1RM knee $extension_{pre\ study}$	♂, 19 (5); ♀, 17 (5)	♂, 23 (6); ♀, 20 (7)	− 4 (− 6, − 1)	0.005*
One-legged endurance performance factor (AU)	♂, 0.2 (0.0); ♀, 0.2 (0.0)	♂, 0.4 (0.1); ♀, 0.3 (0.1)	− 0.2 (− 0.2, − 0.1)	< 0.001*
Whole-body endurance performance factor (AU)	♂, 0.4 (0.1); ♀, 0.3 (0.1)	♂, 0.7 (0.1); ♀, 0.6 (0.1)	− 0.3 (− 0.3, − 0.2)	< 0.001*

COPD, participants diagnosed with chronic obstructive pulmonary disease; Healthy, healthy control participants; ♂, males; ♀, females; †, dropouts during the training period; *study clusters are significantly different from each other ($p < 0.05$); GOLD, Global Initiative for Chronic Obstructive Lung Disease; pack-years, (number of cigarettes smoked per day/20) × number of years smoked; FVC, forced vital capacity; FEV_1, forced expiratory volume in one second; PEF, peak expiratory flow; 1RM, one repetition maximum; Nm, newton-meter; AU, arbitrary units. Data mainly presented as mean (SD), and sex-adjusted estimated mean differences between study clusters (95% CI). For core outcome domains, i.e. lower-body muscle strength, lower-body muscle mass, one-legged endurance performance and whole-body endurance performance, factors were calculated. Briefly, each factor was calculated using multiple singular outcome measures, where each of these variables were normalized to the participant with the highest value recorded during the study, resulting in individual scores ≤ 1. Thereafter, outcome domain factors were calculated as the mean of the normalized values for each variable for each subject (see Additional file 1: Table S1 for complete overview over calculations and composition of each factor)

Kyoto encyclopedia of Genes and Genomes (KEGG) and Gene Ontology gene sets, using two approaches. First, a non-parametric rank test was performed based on gene-specific minimum significant differences. Second, a gene set enrichment analysis (GSEA) was performed to quantify directional regulation of the gene set. GSEA was performed using the fgsea package [29]. Consensus results (i.e. when both the non-directional rank test and the directional GSEA turned out significant) were interpreted as having greater biological meaning, while Hallmark was interpreted as contributing with the most meaningful stand-alone interpretation, as it reduces the analytical noise by taking into account genes that overlap between gene sets [30]. All gene sets were retrieved using the molecular signature database (version 7.1.) [31]. Overview of gene enrichment analyses with exact *p*-values are presented in Additional file 1: Table S3. A repository containing all transcriptome data and scripts used for transcriptome and enrichment analyses are available at https://github.com/dhammarstrom/rnaseq-copd.

For all immunohistochemical variables, statistical models were weighted for numbers of counted fibers *per* biopsy. This was done to account for the reduced reliability accompanying fewer observations/fibers [22].

To achieve reliable assessment of core outcome domains, and thus to lower the risk of statistical errors, combined factors were calculated for outcome measures relating to *lower-body muscle strength* (composed of values from the variables 1RM knee extension and leg press (I), and peak torque for knee extension at 60, 180 and 240°/sec (II)), *lower-body muscle mass* (leg lean mass (I) and *vastus lateralis* and *rectus femoris* thickness (II)), *one-legged endurance performance* (maximal workload achieved during one-legged cycling (I) and number of repetitions at 50% of 1RM knee extension at pre-study (II)) and *whole-body endurance performance* (maximal workload achieved during bicycling (I), maximal number of steps achieved in a 6-min test (II), and maximal number of sit-to-stands in a 1-min test (III)), as previously described [22]. During factor calculation, each of the underlying variables were normalized to the participant with the highest value recorded during the RCT, resulting in individual scores ≤ 1. Thereafter, outcome domain factors were calculated as the mean of the normalized values for each variable for each participant. For details, see Additional file 1: Table S1.

In all mixed-effects models, a single random effect was used, giving each participant an individual intercept. Statistical significance was set to $p < 0.05$. In both text and figures, data are presented as adjusted, marginal means, with or without 95% confidence intervals, unless otherwise stated. Statistical analyses were performed using SPSS Statistics package version 24 (IBM, Chicago, IL, USA) (statistical models with continuous variables, as well as immunohistochemical fiber type proportions) and R software [32] (statistical analyses of rRNA/mRNA content). Figures were made using Prism Software (GraphPad 8, San Diego, CA, USA) and R software [32].

Results and discussion

Baseline characteristics: COPD vs Healthy

Exercise capacity, body composition and muscle and blood biology

At baseline (prior to onset of training), COPD displayed impaired exercise capacity compared to Healthy, as expected from previous studies [3, 18, 20, 33]. This was evident as impaired whole-body performance (range: − 41 − 54%, Table 1), and lower-body unilateral muscle strength and endurance performance (− 17 − 30%, Table 1), presumably reflecting the cardiorespiratory and muscular limitations inherent to the condition [21], and likely being decoupled from levels of habitual physical

Table 2 Baseline characteristics of *m. vastus lateralis* for COPD and Healthy

	COPD	Healthy	Sex-adjusted estimated difference	
			COPD – Healthy (95% CI)	*P*-value
Cross-sectional area (μm^2)				
Type I	4614 (1088)	3720 (951)	449 (70, 827)	0.020*
Type II	3639 (1235)	3059 (1121)	182 (− 118, 482)	0.232
Myonuclei per fiber				
Type I	2.2 (0.9)	2.1 (0.9)	− 0.1 (− 0.4, 0.2)	0.357
Type II	2.1 (0.7)	1.9 (0.7)	− 0.1 (− 0.3, 0.2)	0.504
Myonuclear domain (cross sectional area/nuclei per fiber)				
Type I	2292 (585)	1928 (1030)	360 (107, 613)	0.006*
Type II	1775 (529)	1740 (1049)	− 62 (− 316, 191)	0.628
Fiber type proportion (%)				
Type I	52 (15)	65 (14)	− 16 (− 24, − 9)	< 0.001*
Type IIA	32 (12)	23 (11)	10 (4, 16)	0.001*
Type IIX	13 (7)	9 (6)	5 (1, 9)	0.007*
Type IIA/IIX	3 (2)	2 (2)	0.7 (− 0.4, 1.9)	0.159
Total RNA (ng / ml)	477 (103)	504 (106)	− 20 (-59, 18)	0.302

COPD, participants diagnosed with chronic obstructive pulmonary disease; Healthy, healthy control participants. Data presented as mean (SD), and sex-adjusted estimated mean differences between study clusters (95% CI). Alpha level at $p < 0.05$

activity, as no difference was observed between study clusters prior to onset of the study (Table 1). Alongside the reduced exercise capacity, COPD had less lean body mass than Healthy (− 13%, Table 1), with 45% of COPD showing signs of sarcopenia, as defined by Baumgartner et al. [34]. In the legs, this was manifested as -16% reductions in leg-specific lean mass and − 9/− 24% smaller *vastus lateralis/rectus femoris* thicknesses (Table 1), offering potential explanations for the impaired maximal leg muscle strength. Of note, for markers of muscle mass the difference between study clusters was likely related to traits inherent to the COPD condition rather than to the small age difference between COPD and Healthy (− 2 years; Table 1), as the magnitude of the difference would have implied an annual loss of ~ 2.6 kg lean mass *per* year in the COPD cluster, deviating substantially from the expected loss in this age group (~ 0.5 kg *per* year) [35].

For muscle biological variables, the COPD cluster showed lowered proportions of type I fibers and greater proportions of type IIA and IIX muscle fibers in *vastus lateralis* compared to Healthy (32/23% vs 13/9%, respectively), corroborating with previous studies [36, 37]. For type I fibers, COPD showed larger CSA (12%, Table 2) and larger myonuclear domain (CSA *per* myonuclei; 20%, Table 2), with no such differences being observed for type II fibers. This contrasts previous studies, who have reported smaller or similar CSA in type I fibers in COPD compared to Healthy [33, 38, 39], and may point to a compensatory mechanism for the likely loss of motor units in COPD subjects [40], whereby reduced quantities of muscle fibers are compensated for by increased sizes of remaining fibers, as previously reported in rodents [41]. These observed differences in muscle fiber characteristics were accompanied by differences in RNA expression. Although COPD and Healthy showed similar levels of total RNA and rRNA expression *per* amount of muscle tissue at baseline (Table 2), COPD displayed distinct whole-genome transcriptome profiles, with 227 genes being differentially expressed compared to Healthy (151↑ and 76↓; Fig. 3A and Additional file 1: Table S2). Hallmark enrichment analysis revealed lower expression of genes involved in *oxidative phosphorylation* (consensus, i.e. agreement between GSEA and rank-based analyses), corroborating with the lower type I proportion, as well as greater expression of genes involved in regulation of *myogenesis* (Rank) (Fig. 3A, B, Table 3; confirmed by gene ontology analysis, Additional file 1: Table S3), which may be related to the pathophysiological elevation of protein turnover in COPD [42, 43].

For blood variables, the COPD cluster showed elevated levels of low-grade inflammation compared to Healthy, measured as levels of c-reactive protein prior to the study (5.0 vs 1.6 $mg{\cdot}L^{-1}$, $p = 0.001$, data not shown; baseline (i.e. prior to resistance training), 5.0 vs 1.6 $mg{\cdot}L^{-1}$, $p = 0.053$, Table 4), as expected from previous studies [8]. For other characteristics, including hormonal status in blood (e.g. testosterone), no differences were observed between COPD and Healthy (Table 4).

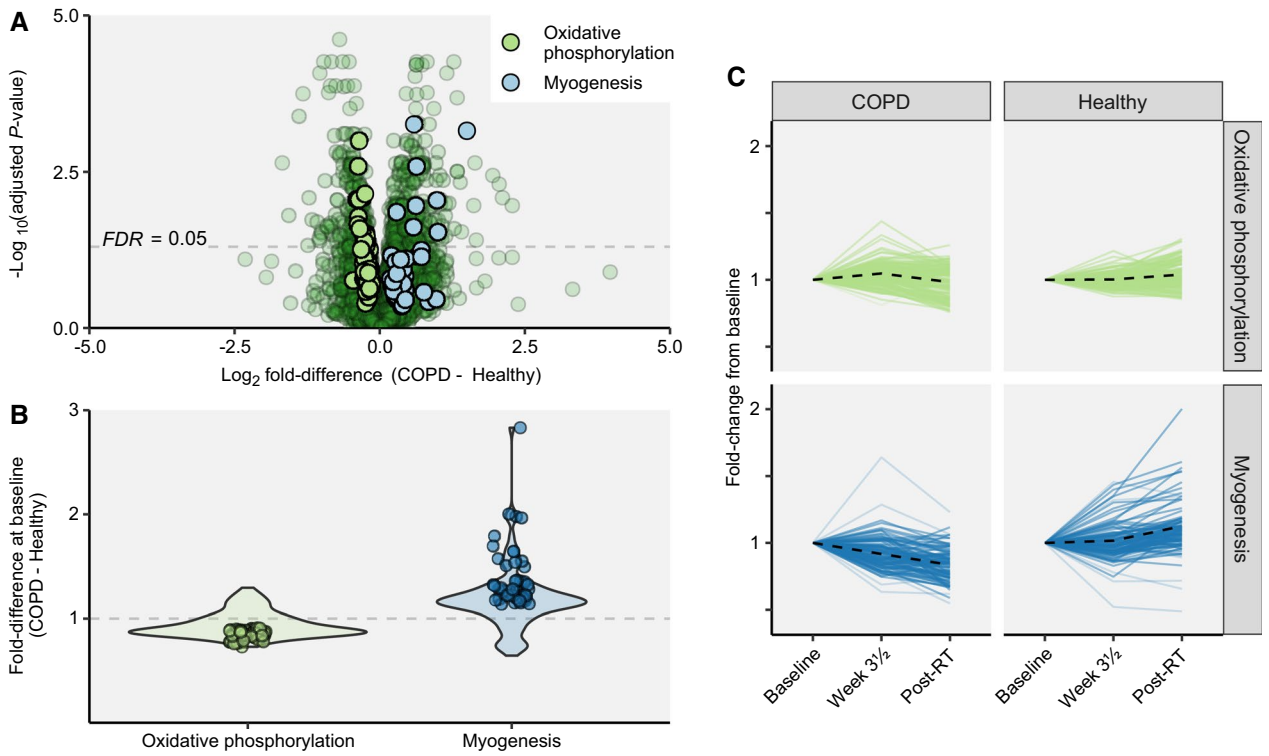


Fig. 3 Whole-genome transcriptome analyses of *m. vastus lateralis* in COPD and Healthy (COPD, n = 19; Healthy, n = 34). At baseline, numerous genes were differentially expressed between COPD and Healthy. In **A**, differences in gene expression between COPD and Healthy are presented with leading edge genes (i.e. genes that contributes to the enrichment score) from two gene sets identified as differentially expressed between COPD and Healthy from gene enrichment analyses (*oxidative phosphorylation* and *myogenesis*; see Table 3). In **B**, average fold differences (COPD-Healthy) of genes contributing to baseline differences in *oxidative phosphorylation* and *myogenesis* gene sets are shown as individual data points, and violin plots shows the distribution of all leading edge genes from each gene set. **C** displays the average development of each gene set over time, where the dotted line indicates the mean fold change of all genes contributing to the differential change over time between COPD and Healthy. COPD displayed larger increases in expression of genes relating to *oxidative phosphorylation* after 3½ weeks of training, and more pronounced decreases in genes associated with *myogenesis* to after the training intervention (Post-RT; see Table 3). FDR, false discovery rate-adjusted *p*-value

The efficacy of the resistance training intervention: COPD vs Healthy

For both COPD and Healthy, the training intervention was associated with low drop-out rates (n = 4, ~ 5%; COPD, n = 2), high adherence to the protocol (COPD, 97%; Healthy, 98%; measured as the average number of training sessions completed), progressive increases in training volume (Fig. 2), and robust increases in muscle strength *per* training session (e.g. 1RM knee extension, 0.9% · session^{-1}/0.8% · session^{-1}, COPD/Healthy; 1RM leg press, 1.4% · session^{-1}/1.3% · session^{-1}). The habitual dietary intake was similar between COPD and Healthy, with protein intake being 1.2 (0.3) and 1.3 (0.4) g · kg^{-1} · day^{-1}, respectively, complying with current guidelines [44]. The vitamin D_3 supplementation RCT of the project did not enhance or affect training-associated changes for any of the primary or secondary outcome measures [22].

Muscle strength, muscle mass, muscle quality and one-legged endurance performance

Overall, COPD showed larger training-associated increases in lower-body muscle strength and mass compared to Healthy (the two legs/training modalities combined), measured as relative changes in combined factors from baseline (Fig. 4A), with no difference being observed for numeric changes (Fig. 4A). COPD and Healthy showed similarly scaled improvements in muscle quality and one-legged endurance performance (Fig. 4A). Within the COPD cluster, worsening of lung function (i.e. decreasing predicted FEV_1 values) was associated with larger numeric and relative increases in muscle mass, as well as larger relative improvements in maximal muscle strength, with no such relationship being observed for muscle quality or one-legged endurance performance (Table 5). Neither of the four core outcome domains (muscle strength/mass/quality or one-legged endurance

Table 3 Comparison of Hallmark gene sets identified in whole-genome transcriptome data between COPD (n = 19) and Healthy (n = 34), assessed at baseline and as resistance training-associated changes

Comparison	Gene set	Significance category*	Set size†	Rank *P*-value‡	% MSD > 0§	GSEA *P*-value‖	NES	LE**	Log_2 fold difference in LE (95% CI)
Baseline: COPD vs Healthy	Oxidative phosphorylation	Consensus	190 (200)	0.007	36.8%	< 0.001	− 2.10	70 (94.3%)	− 0.24 (− 0.45, − 0.13)
	Myogenesis	Rank	163 (200)	< 0.001	33.7%	0.417	1.21	45 (75.6%)	0.46 (0.19, 1.5)
3½ weeks of training: ΔCOPD vs ΔHealthy	Allograft rejection	GSEA	115 (200)	0.956	7.8%	0.014	1.71	20 (35%)	0.39 (0.13, 0.76)
	Oxidative phosphorylation	GSEA	190 (200)	0.999	1.1%	0.009	1.69	83 (2.4%)	0.11 (0.05, 0.39)
	Pancreas beta cells	GSEA	15 (40)	0.969	6.7%	0.028	1.71	3 (33.3%)	0.35 (0.08, 0.54)
Post− RT (13 weeks of training): ΔCOPD vs ΔHealthy	Myogenesis	Consensus	163 (200)	< 0.001	42.3%	< 0.001	− 1.52	68 (85.3%)	− 0.5 (− 1.13, − 0.26)

*Consensus significance indicates agreement between directional (GSEA) and non-directional (Rank) hypothesis test of overrepresentation (see methods for details). †Indicates number of identified genes in the gene set and total number of genes in the gene set in parentheses. ‡Rank-based enrichment test, based on minimum significant difference (MSD), identifies gene sets that are overrepresented among top-ranked genes without a directional hypothesis. § Fraction of genes in gene set with unadjusted 95% CI not spanning zero, i.e. MSD > 0. ‖ Gene-set enrichment analysis (GSEA) tests for overrepresentation among top and bottom genes based on Log_2 fold differences or changes × $-log_{10}$(*P*-values) in comparing differences at baseline or changes from baseline between COPD and Healthy. A positive normalized enrichment score (NES) indicate gene set with higher expression in COPD than Healthy; negative NES indicate gene set with lower expression at respective time-points. ** Number of genes in leading edge (LE, genes that contributes to the enrichment score) with the fraction of leading edge genes with unadjusted 95% CI not spanning zero. *Δ* change score

performance) were differentially affected by resistance training load (neither in COPD nor in Healthy), suggesting that 30RM training is an effective alternative to 10RM training in older individuals (Fig. 4B, C). Of note, the comparisons between 10 and 30RM training responses may have been confounded by the so-called *cross-education effect*, whereby training of one limb affects functional and biological characteristics of the contralateral limb. However, the true existence of such cross-education effects remains disputed, and if it does exist, its impact is likely restricted to neuromuscular functionality [45, 46], with no apparent effects on muscle biological measures such as mRNA abundance [45], mitochondrial content [47, 48], capillarization [49], muscle protein synthesis [50] or muscle hypertrophy [51, 52]. In accordance with this, the cross-education effect may have affected measures of muscle strength and one-legged endurance performance in the present study. Importantly, however, several measures were implemented into the study protocol to minimize its impact, including extensive familiarization to training and physical testing (e.g. baseline muscle strength was measured after 3 ½ weeks of introduction to training and was preceded by 3–5 familiarization sessions to muscle strength testing) [22].

Overall, COPD showed marked and hitherto unrecognized responsiveness to resistance training in respect of improvements in muscle strength, muscle mass, muscle quality and one-legged endurance performance, contradicting previous suggestions of a negative impact of co-morbidities such as low cardiorespiratory fitness and chronic low-grade systemic inflammation [8, 24]. Indeed, a more severe COPD diagnosis was associated with larger increases in muscle mass and muscle strength improvements. This observation cannot be readily explained by baseline differences between the COPD participants (e.g. baseline muscle mass vs predicted FEV_1, $p = 0.998$; baseline muscle strength vs predicted FEV_1, $p = 0.646$). The marked training responsiveness in COPD was presumably also decoupled from initial differences in physical activity habits between study clusters, as COPD and Healthy showed similar characteristics regarding these measures (Table 1), though some caution is warranted for interpretation of such self-reported recall questionnaire results [53].

Cycling and functional performance

COPD and Healthy showed pronounced and similarly scaled training-associated improvements in whole-body endurance performance, measured as changes from baseline, including 6-min step test performance, 1-min sit-to-stand performance and maximal workload achieved during two-legged cycling (Fig. 5). Surprisingly, COPD and Healthy also showed similar changes in performance for these outcome measures in numeric terms, with exception of 6-min step test performance, for which Healthy showed larger improvements

Table 4 Effects of the training intervention on body composition and blood variables in COPD and Healthy, assessed as changes from baseline to after completion of the study (per study cluster) and as differential changes between study clusters

	COPD			Healthy			Δ COPD vsΔ Healthy (*P* value)
	Baseline	Post RT	Time effect (*P* < 0.05)	Baseline	Post RT	Time effect (*P* < 0.05)	
Dual-energy x-ray absorptiometry							
Whole-body bone mineral density ($g \cdot cm^2$)	1.13 (0.21)	1.13 (0.21)	No	1.15 (0.16)	1.14 (0.15)	No	0.119
Total lean mass (kg)	46.7 (9.9)	47.6 (10.2)	Yes ↑	48.1 (10.0)	48.6 (10.0)	Yes ↑	0.395
Appendicular lean mass (kg)	20.3 (5.3)	20.9 (5.5)	Yes ↑	21.6 (5.0)	21.9 (5.0)	Yes ↑	0.166
Total fat mass (kg)	26.4 (11.7)	26.3 (11.5)	No	25.3 (9.3)	24.4 (9.2)	Yes ↓	0.068
Visceral fat (kg)	1.59 (1.18)	1.56 (1.21)	No	1.12 (0.98)	1.01 (0.81)	Yes ↓	0.138
Inflammation							
C-reactive protein ($mg \cdot L^{-1}$)	3.4 (5.0)	3.6 (4.0)	No	1.7 (2.5)	1.8 (3.5)	No	0.934
Hormones							
Cortisol ($nmol \cdot L^{-1}$)	307 (130)*	310 (109)	No	369 (88)	372 (99)	No	0.861
Growth hormone ($\mu g \cdot L^{-1}$)	1.4 (2.8)	1.4 (3.1)	No	1.1 (1.7)	1.3 (1.6)	No	0.837
IGF-1 ($nmol \cdot L^{-1}$)	15.7 (4.2)	15.0 (4.5)	No	14.4 (3.2)	13.6 (3.1)	Yes ↓	0.977
Testosterone ($nmol \cdot L^{-1}$)†	11.2 (4.4)	11.4 (4.2)	No	11.9 (3.3)	12.4 (4.2)	No	0.938
Sex-hormone binding globulin ($nmol \cdot L^{-1}$)	60 (33)	60 (34)	No	60 (22)	60 (21)	No	0.488
Androstenedione ($nmol \cdot L^{-1}$)	3.3 (2.4)	3.3 (2.4)	No	3.8 (2.7)	3.8 (2.4)	No	0.984
Parathyroid hormone ($pmol \cdot L^{-1}$)	5.7 (2.6)	6.0 (3.3)	No	5.0 (2.2)	5.2 (1.9)	No	0.870
Lipid profile variables							
Triglycerides ($mmol \cdot L^{-1}$)	1.2 (0.5)	1.1 (0.5)	No	1.2 (0.5)	1.1 (0.6)	Yes ↓	0.661
HDL ($mmol \cdot L^{-1}$)	1.6 (0.6)	1.7 (0.5)	No	1.7 (0.5)	1.7 (0.5)	No	0.523
LDL ($mmol \cdot L^{-1}$)	2.8 (1.0)*	2.8 (1.0)	No	3.4 (1.0)	3.3 (0.8)	No	0.775
Iron biology variables							
Fe^{2+} ($\mu mol \cdot l\, L^{-1}$)	18 (7)	18 (6)	No	18 (6)	18 (5)	No	0.410
Transferrin ($g \cdot L^{-1}$)	2.66 (0.44)*	2.67 (0.45)	No	2.41 (0.27)	2.38 (0.29)	No	0.563
Ferritin ($\mu g \cdot L^{-1}$)	113 (92)	90 (81)	Yes ↓	139 (79)	133 (68)	No	0.089
Calcium status							
Calcium ($mmol \cdot L^{-1}$)	2.4 (0.1)	2.4 (0.1)	No	2.4 (0.1)	2.4 (0.1)	No	0.865
Albumin-corrected calcium ($mmol \cdot L^{-1}$)	2.3 (0.1)	2.3 (0.1)	No	2.3 (0.1)	2.3 (0.1)	No	0.802
Tissue damage variables							
Aspartate transaminase ($units \cdot L^{-1}$)	27 (9)	24 (6)	No	26 (21)	26 (7)	No	0.807
Creatine kinase ($units \cdot L^{-1}$)	112 (69)	123 (71)	No	95 (47)	125 (72)	Yes ↑	0.523

Body composition analyses, n COPD = 19, n Healthy = 48; blood analyses, n COPD = 20, n Healthy = 58. *significant difference between COPD and Healthy at baseline; †only men were included in testosterone analysis; ↓significant decrease from baseline to post RT (after 13 weeks of resistance training); ↑significant increase from baseline to post RT. Alpha level at $p < 0.05$. Data are presented as means (SD)

(COPD, 6 steps; Healthy, 17 steps; Δ11 steps, $p = 0.009$; Fig. 5), arguably related to the considerable cardiorespiratory demand of this test, leaving COPD with morbidity-specific restraints. Corroborating with this, within the COPD cluster, there was no association between the severity of the COPD diagnosis and resistance training-induced changes in whole-body endurance performance (Table 5). For other performance indices such as cycling economy and gross efficiency, which were measured using a one-legged cycling protocol, COPD showed larger relative improvements compared to Healthy (Δ4% (COPD – Healthy) for both cycling economy and gross efficiency, Fig. 5). For these indices of cycling performance, COPD, but not Healthy, displayed benefits of 10RM compared to 30RM training (Fig. 5), corresponding to previously observed effects of heavy resistance training in healthy, young individuals [54].

Together, these observations reiterate on the substantial benefits of resistance training for subjects with COPD, even for performance measures that pose large whole-body metabolic demands, which has previously

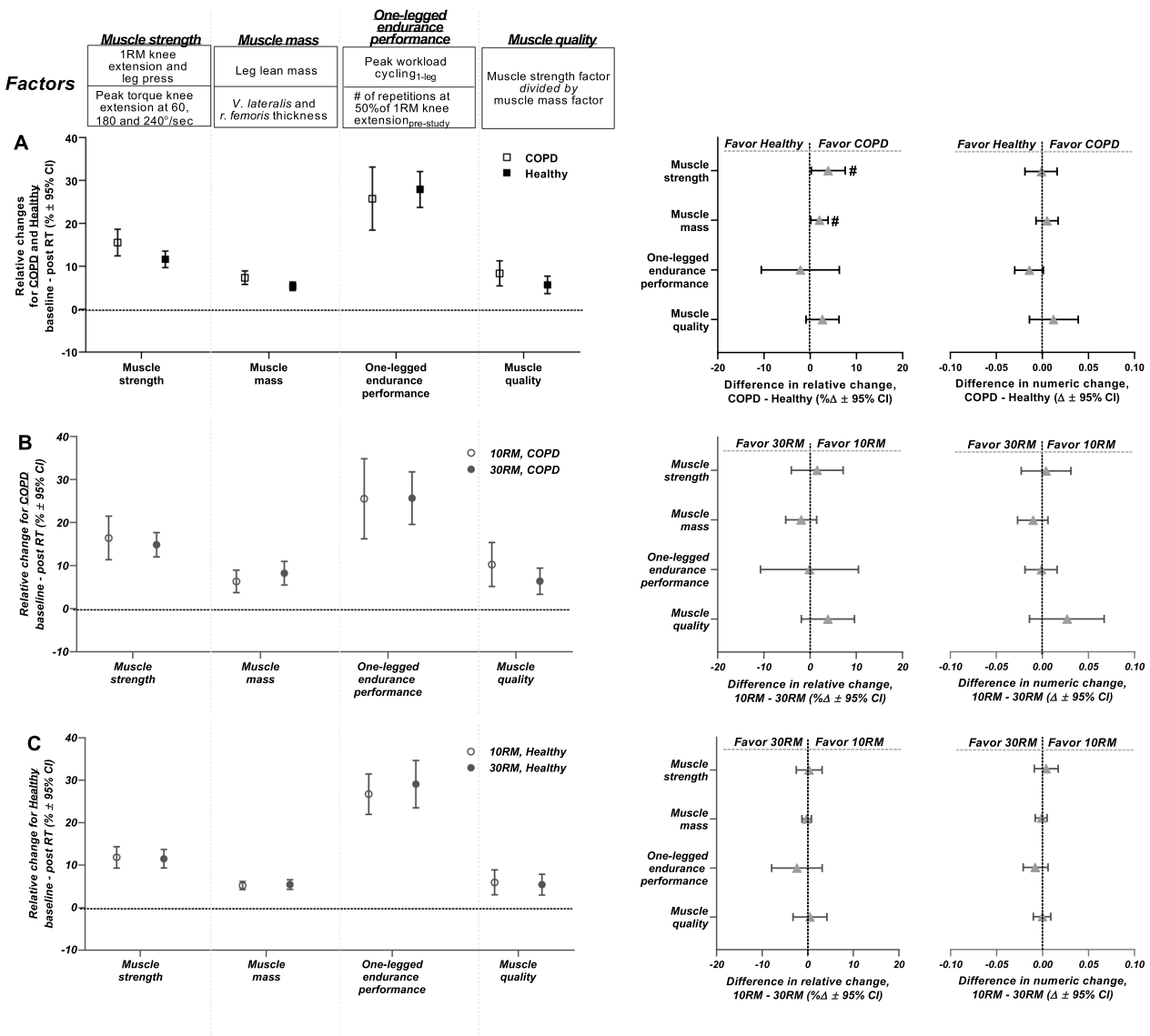


Fig. 4 Effects of the resistance training intervention on lower-body muscle strength (COPD, n = 18; Healthy, n = 50), lower-body muscle mass (COPD, n = 19; Healthy, n = 47), one-legged endurance performance (COPD, n = 15; Healthy, n = 49) and lower-body muscle quality (COPD, n = 18; Healthy, n = 38) in COPD and Healthy. Each outcome domain is represented by a combined factor, computed from various performance assessments, as defined in the upper panel of the figure and previously described [22]. **A** presents comparison of overall training effects between COPD and Healthy, measured as relative changes from baseline to after the resistance training intervention (per study cluster; left panel) and as relative and numeric differences in change scores between study clusters (right panels). In these analyses, high- and low-load resistance training (10RM and 30RM, respectively) were combined, warranted by the lack of differences between training load conditions in (B, C). COPD showed greater relative changes in muscle strength and muscle mass than Healthy. **B**, **C** presents comparison of effects of 10RM and 30RM resistance training in COPD (**B**) and Healthy (**C**) (i.e. per study cluster), measured as relative changes from baseline to after the intervention (left panels) and as relative and numeric differences in change scores between load conditions (right panels). #statistically different effects of resistance training between COPD and Healthy. Data are presented as means with 95% confidence limits

been suggested to be irresponsive to such training [55]. As such, it seems plausible that the observed improvements in 6-min step test performance, 1-min sit-to-stand performance and two-legged cycling were associated with improvements in work economy/gross efficiency and muscle strength, as neither COPD nor Healthy showed training-associated changes in maximal oxygen consumption (Fig. 5), with improvements in anaerobic capacity being a potential contributor (not measured).

Table 5 Simple linear regression analyses on the relationship between training response and lung function in COPD participants

Analysis	n	Slope (95% CI)	Intercept (95% CI)	*r*	*P*
Change in muscle strength vs $FEV_{1predicted}$					
% change	18	− 0.3 (− 0.6, 0.0)	34.8 (16.8, 52.9)	− 0.504	0.033
Numeric change	18	− 0.001 (− 0.003, 0.001)	0.121 (0.017, 0.225)	− 0.303	0.222
Change in muscle mass vs $FEV_{1predicted}$					
% change	19	− 0.3 (− 0.4, − 0.1)	21.4 (12.1, 30.7)	− 0.624	0.004
Numeric change	19	− 0.002 (− 0.003, 0.000)	0.127 (0.068, 0.186)	− 0.603	0.006
Change in muscle quality vs $FEV_{1predicted}$					
% change	18	− 0.1 (− 0.4, 0.2)	12.6 (− 4.2, 29.4)	− 0.141	0.577
Numeric change	18	0.000 (− 0.002, 0.002)	0.063 (− 0.060, 0.186)	− 0.038	0.881
Change in one-legged endurance performance vs $FEV_{1predicted}$					
% change	15	0.3 (− 0.4, 1.0)	8.5 (− 32.8, 49.7)	0.249	0.371
Numeric change	15	0.001 (− 0.001, 0.002)	0.006 (− 0.066, 0.079)	0.282	0.308
Change in whole-body endurance performance vs $FEV_{1predicted}$					
% change	17	− 0.2 (− 0.6, 0.3)	17.7 (− 7.8, 43.1)	− 0.211	0.416
Numeric change	17	0.000 (− 0.001, 0.001)	0.023 (− 0.042, 0.089)	0.012	0.963

$FEV_{1predicted}$, predicted forced expiratory volume in first second; *r*, Pearson's *r*; *P*, *P*-value

Muscle fiber characteristics

Whereas COPD and Healthy displayed similar increases in type II fiber CSA in *m. vastus lateralis* in response to resistance training (COPD, 18%; Healthy, 24%; Δ-6%, $p = 0.438$; Fig. 6, upper panel), only Healthy showed increases in type I fiber CSA (16%), with no statistical difference being observed between study clusters. For Healthy, the increase in CSA was accompanied by increased myonuclei · $fiber^{-1}$ in both fiber types (36%/25% for type I/II; Fig. 7), leading to decreased myonuclear domain size estimates in type I fibers (-10%, Fig. 7). For COPD, no such effects were observed (Fig. 7). Despite the lack of difference between the two study clusters for these variables, the data hints at blunted plasticity of type I muscle fibers in COPD only, potentially relating to their altered biological characteristics at baseline or to blunted myonuclear accretion. Interestingly, in sub-analyses, the blunted type I responses in COPD seemed to be specific to 10RM training, with a tendency towards superior responses to 30RM training (10RM, -3%; 30RM, 19%; Δ22%, $p = 0.060$; Fig. 6, middle panel). Such a phenomenon is supported by previous observations in responses to blood-flow-restricted low-load training [56], which arguably is mimicked by COPD subjects during low-load training, as they display inherent lowering of oxygen saturation in blood.

Both COPD and Healthy displayed training-associated reductions in type IIX muscle fiber proportions (Fig. 7). While this reduction was more pronounced in COPD when measured at the protein level (immunohistochemistry), it was more pronounced in Healthy when measured at the mRNA level, suggesting differential orchestration of muscle fiber shifts between study clusters, possibly relating to their inherently different muscle fiber proportions at baseline.

Muscle RNA content

In general, COPD and Healthy showed similar increases in ribosomal RNA abundance *per* unit muscle tissue weight, measured as both total RNA and rRNA expression, and measured after both 3½ week (1.19/1.29 and 1.16/1.16 fold increases, total RNA/rRNA abundances) and after finalization of the training intervention (1.12/1.05 and 1.19/1.17 fold increases) (Fig. 8). While these changes in ribosomal RNA content were generally similar between COPD and Healthy, a few noteworthy differences were evident, including a more robust early increase in 45s pre-rRNA abundance in COPD (Fig. 8) and a trend towards reduced changes in response to 13 weeks training in COPD, which led to the absence of time effects for all rRNA species. The early increases in ribosomal content seen in both COPD and Healthy resemble those typically seen after similar interventions in untrained young individuals [26], and may be important for muscle growth capabilities over the entirety of the study period [26, 57], accommodating increases in protein synthesis capacity, thus potentially contributing to the pronounced muscular responses to resistance training seen in both study clusters.

In both COPD and Healthy, resistance training led to marked changes in mRNA transcriptome profiles, with 499 and 312 differentially expressed genes being observed after 3½ and 13 weeks of resistance training,

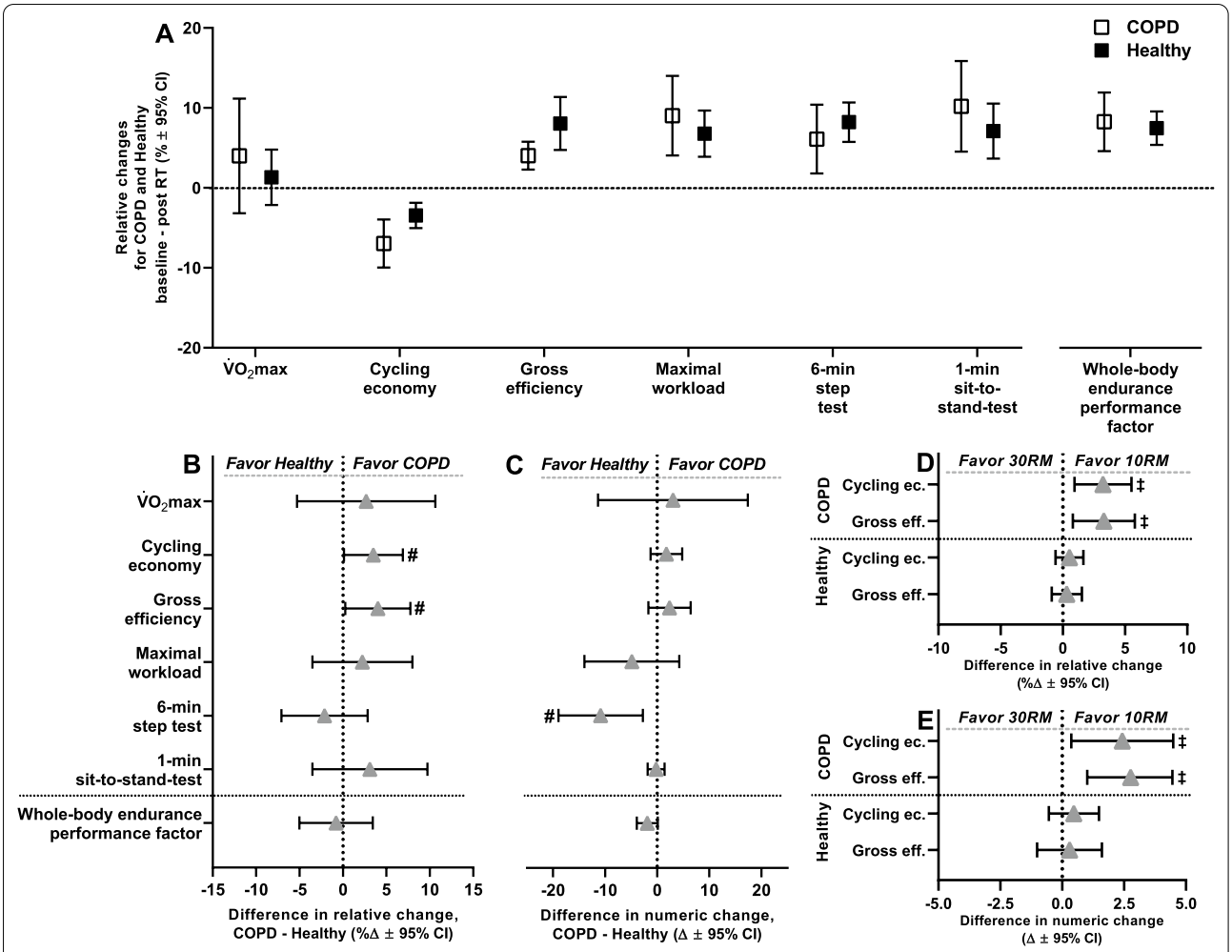


Fig. 5 Comparison of the effects of the resistance training intervention on whole-body endurance performance in COPD and Healthy, presented as relative changes from baseline (per study cluster; **A**) and as relative and numeric differences in change scores between study clusters (**B** and **C**, respectively). Endurance measures included maximal oxygen consumption ($\dot{V}O_2$max, cl· min^{-1}; COPD, n = 15; Healthy, n = 54) and maximal workload (watts; COPD, n = 18; Healthy, n = 55) achieved during two-legged cycling, cycling economy (cl· min^{-1}; COPD, n = 15; Healthy, n = 54) and gross efficiency measured during submaximal one-legged cycling, the number of steps achieved during 6-min step test (COPD, n = 18; Healthy, n = 57) and the number of sit-to-stands achieved during a 1-min sit-to-stand test (COPD, n = 19; Healthy, n = 56). COPD showed greater relative improvements in cycling economy and gross efficiency. For these outcome measures, COPD, but not Healthy, displayed benefits of high-load training (10RM) compared to low-load training (30RM) (**D** and **E**). Healthy showed greater numeric improvement in the number of steps achieved during the 6-min step test. COPD and Healthy showed similar relative and numeric training-associated changes in the whole-body endurance performance factor. #statistically different response to resistance training between study clusters. ‡statistically different response to 10RM and 30RM resistance training in study cluster. Data are presented as means with 95% confidence limits

(See figure on next page.)

Fig. 6 Effects of the resistance training intervention on cross-sectional area of muscle fiber types I and II in *m. vastus lateralis* in COPD (n = 18) and Healthy (n = 55). **A** presents comparison of overall training effects on fiber CSA between COPD and Healthy, measured as relative changes from baseline to after the training intervention (per study cluster; left panel) and as relative differences in change scores between study clusters (right panel). In these analyses, high- and low-load resistance training (10RM and 30RM, respectively) were combined, warranted by the lack of significant differences between training load conditions in (**B**, **C**), though COPD tended to show higher efficacy of 30RM resistance training for changes in fiber type I CSA. **B**, **C** presents comparisons of effects of 10RM and 30RM resistance training on fiber CSA in COPD (**B**) and Healthy (**C**) (i.e. per study cluster), measured as relative changes from baseline to after the training intervention (left panels) and as relative and numeric differences in change scores between load conditions (right panels). Data are presented as means with 95% confidence limits

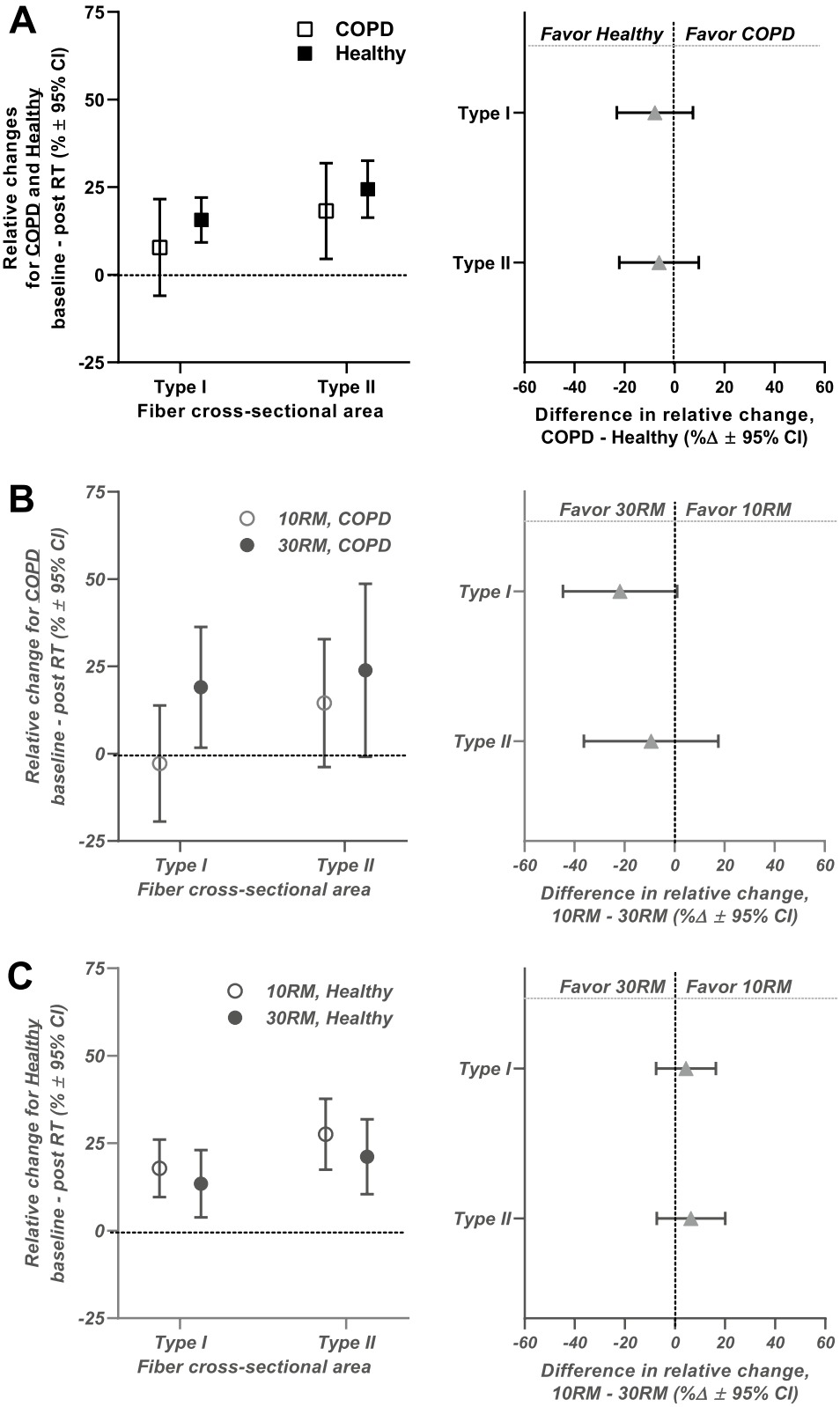


Fig. 6 (See legend on previous page.)

respectively (for general information about transcriptomic responses, see Mølmen et al. [22]). Overall, at the single-gene level, no transcripts showed differential responses to training between the two study clusters, neither at 3½ weeks nor at 13 weeks, despite clear differences in transcriptome profiles at baseline (Fig. 3A and Additional file 1: Table S2). In contrast, enrichment analyses revealed traces of differential changes (Fig. 3C, Table 3 and Additional file 1: TableS 3), with COPD showing more pronounces increases in expression of genes relating to *oxidative phosphorylation* after 3½ weeks (GSEA), and, in particular, more pronounced decreases in genes associated with *myogenesis* after 13 weeks (consensus) (Fig. 3C, Table 3). Interestingly, as these two gene sets represented the most prominent differences between COPD and Healthy at baseline (Fig. 3A, B), and as resistance training led to directional changes that mitigated these differences, training arguably shifted the COPD phenotype in a healthy direction.

Blood and health-related outcomes

Overall, COPD and Healthy showed similar training-associated increases in whole-body and appendicular lean mass (Table 4). This was accompanied by increased appendicular skeletal muscle mass index relative to the sex-specific mean of young, healthy adults (COPD, from 84 to 86%; Healthy, from 95 to 97%), suggesting that the intervention was effective for reversing age-related decline in muscle mass. For blood variables such as markers of systemic inflammation and hormone, lipid and iron biology, no noteworthy effects were observed of the intervention, nor were any differential changes observed between COPD and Healthy (Table 4).

Lung function

For COPD, the training intervention did not affect any of the lung function variables (Table 6), implying no effects on this core epidemiological trait. This seems reasonable given the irreversible nature of the respiratory impairments of COPD, contradicting the beneficial effects observed in Hoff et al. [14] In contrast, for Healthy, the intervention was associated with reduced FVC and FEV_1 (− 2.7% and − 1.5%, respectively). Rather than being a consequence of the intervention protocol per se, this may be due to a general age-related decline, as the magnitude of the changes resemble those seen in corresponding age cohorts over a similar time frame [58].

Health-related quality of life

For COPD, the intervention was associated with marked improvements in several aspects of health-related quality of life (Table 7). These included reduced experience of limitations of physical functioning and improved social function and mental health, with only marginal effects being seen in Healthy. While these changes of course may be directly related to the resistance training intervention, they may also be related to other aspects of the study protocol, such as performing training sessions in a social setting and the close follow-up each participant received from study personnel. As the intervention was conducted without a control group (not receiving the intervention protocol), caution is warranted for interpretation of these data.

Concluding remarks

COPD-related pathophysiologies, such as reduced testosterone [4], vitamin D [5] and oxygen saturation levels [7, 59] in blood, and elevated levels of low-grade inflammation [8], are generally believed to drive metabolism into a chronic catabolic state [4, 7, 9]. This has been suggested to lead to impaired responses to lifestyle interventions such as resistance training [7, 60], which are essential measures for preventing and treating disease-related reductions in skeletal muscle mass and strength, counteracting escalation into serious conditions such as pulmonary cachexia [17]. Despite this general belief, the

(See figure on next page.)

Fig. 7 Comparisons of the effects of the resistance training intervention on changes in myonuclei *per* fiber and myonuclei domain in muscle fiber types I and II (**A**, **B**; COPD, n = 11; Healthy, n = 34), and on changes in muscle fiber type proportions in COPD and Healthy, measured using immunohistochemistry (**C**–**E**; COPD, n = 17; Healthy, n = 51)) and qPCR (gene family profiling-normalized myosin heavy chain mRNA expression, **F**–**H**; COPD, n = 19; Healthy, n = 55), as previously described [26, 75]. Myonuclei domain was calculated as mean fiber cross-sectional area divided by myonuclei *per* fiber. For myonuclei *per* fiber and myonuclei domain in muscle fiber types I and II, comparisons are presented as relative changes from baseline to after the training intervention (per study cluster; **A**) and as relative differences in change scores between study clusters (**B**). For muscle fiber type proportions, data are presented as adjusted values at baseline and after the training intervention (Post RT), and results are presented as the effect of the training intervention for the study clusters combined and its interaction with study clusters (**C**–**H**). For myonuclei variables, no training-associated differences were observed between study clusters. Both COPD and Healthy displayed training-associated reductions in proportions of type IIX muscle fibers, measured using both immunohistochemistry and qPCR. Intriguingly, while this reduction was greater in COPD when measured at the protein level (immunohistochemistry), it was greater in Healthy when measured at the mRNA level (qPCR), indicating differentially regulated muscle fiber shifting in COPD and Healthy. Data are presented as means with 95% confidence limits

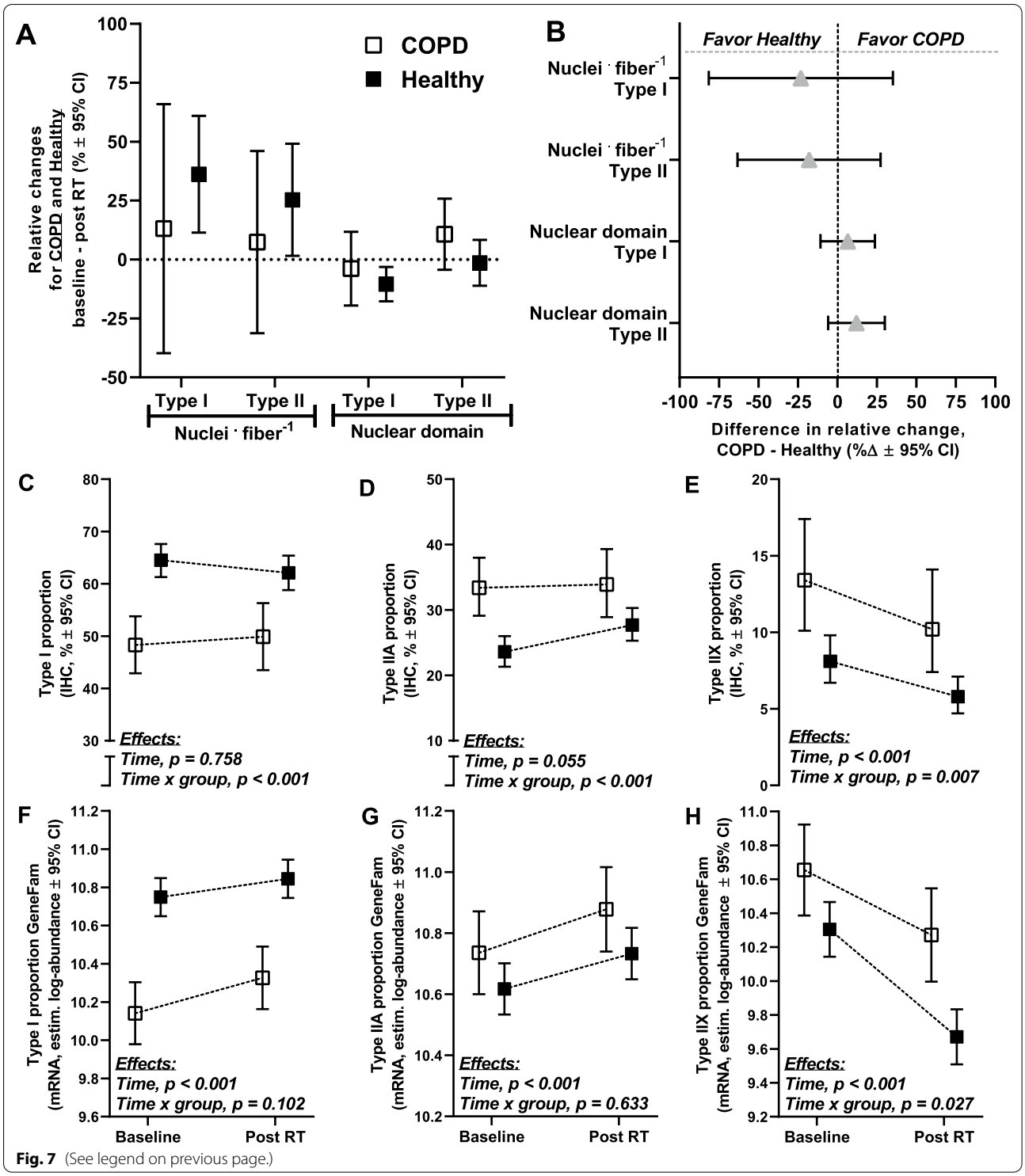


Fig. 7 (See legend on previous page.)

presence of impaired training responsiveness in COPD is not backed by experimental data, and there is limited de facto evidence for such impairments. To date, a mere single study has compared responses between COPD and healthy control subjects [18–20], and as such failing to lend support to the prevailing view, though being limited by a relatively short time span (8 weeks) and a restricted selection of outcome variables. In the present study, we largely disavow the myth of impaired responsiveness to training in COPD, measured as responses to a 13-week

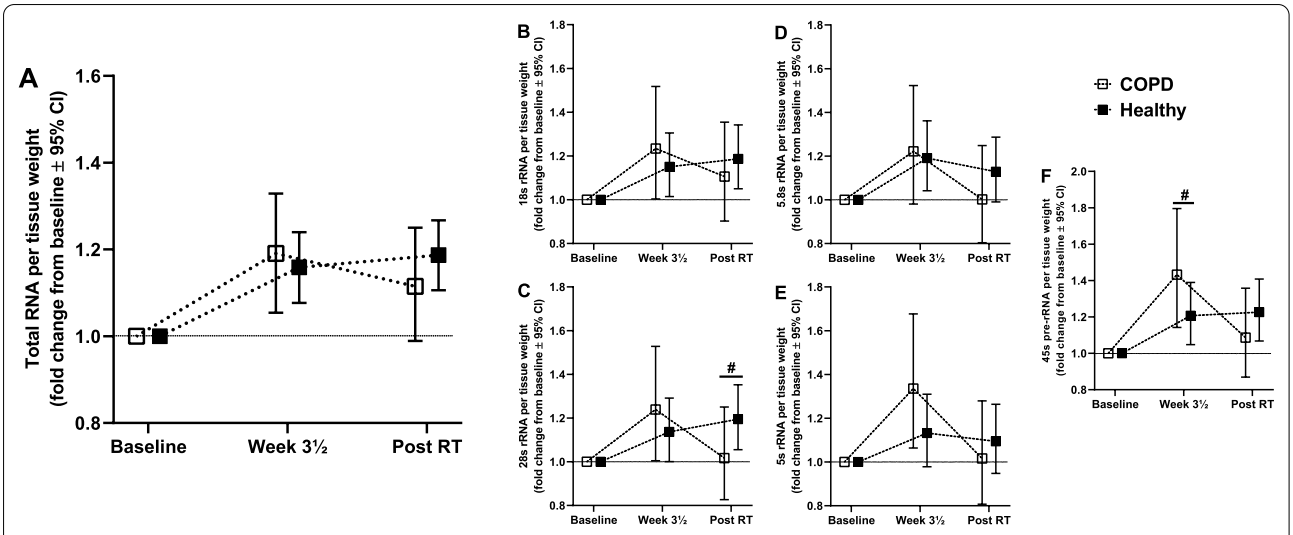


Fig. 8 Effects of the resistance training intervention on total RNA content (**A**) and rRNA expression (**B**–**F**) in *m. vastus lateralis* of COPD (n = 19) and Healthy (n = 55). Data are presented as fold changes from baseline to Week 3½ (Post-intro RT; seven training sessions) and to after the training intervention (Post RT; 26 training sessions). Total RNA (**A**), 18s rRNA (**B**), 28s rRNA (**C**), 5.8s rRNA (**D**) 5s rRNA (**E**) and 45s pre-rRNA (**F**) abundances. Total RNA- and qPCR-analyses were assessed as per-amounts of tissue weight, as previously described [22, 26]. #statistical difference in fold change between COPD and Healthy (alpha level, $p < 0.05$). Data are presented as means with 95% confidence limits

Table 6 Effects of the training intervention on lung function in COPD (n = 20) and Healthy (n = 58), assessed as changes from baseline to after completion of the study (per study cluster) and as differential changes between study clusters

	COPD			Healthy			
	Baseline	Post RT	Time effect p < 0.05)	Baseline	Post RT	Time effect (p < 0.05)	Δ COPD vs Δ healthy (*p*-value)
FVC (L)	3.3 ± 0.9	3.2 ± 0.9	No	3.6 ± 0.9	3.5 ± 0.8	Yes↓	0.189
FEV_1 (L · sec^{-1})	1.5 ± 0.4	1.5 ± 0.4	No	2.7 ± 0.7	2.7 ± 0.6	Yes↓	0.243
FEV_1 (% predicted)	56 ± 11	58 ± 13	No	103 ± 16	103 ± 16	No	0.138
FEV_1/FVC (%)	47 ± 8	48 ± 10	No	75 ± 6	76 ± 6	No	0.714
PEF (L · sec^{-1})	5.0 ± 1.6	5.1 ± 1.6	No	7.8 ± 2.1	7.6 ± 2.2	No	0.238

FVC forced vital capacity, FEV_1 forced expiratory volume in one second, *PEF* peak expiratory flow, *Δ* change score, ↓significant decrease from baseline to post RT (after 13 weeks of resistance training). Alpha level at $p < 0.05$. Values are means with standard deviation

whole-body resistance training intervention, conducted using an exhaustive follow-up and testing protocol, which included extensive test–retest validations (for details, see Mølmen et al. [22]). Whereas COPD participants displayed clear and well-known disease-related aberrancies compared to Healthy at baseline, including altered skeletal muscle characteristics and elevated levels of systemic inflammation, they showed similar or superior improvements for virtually every measure of health, performance and biology. Specifically, COPD showed greater relative improvements in core outcome domains such as lower-body muscle strength and mass, and similar relative improvements in muscle quality, one-legged endurance performance and whole-body endurance performance. These similarities were also evident in numeric change terms, suggesting that the improvements seen in COPD was decoupled from the compromised levels at baseline. Indeed, within the COPD cluster, worsening of lung function was associated with larger numeric and relative increases in muscle mass, as well as larger relative improvements in maximal muscle strength. These observations were accompanied by similar alterations in muscle biology, including changes in hallmark traits such as muscle fiber characteristics, rRNA content and transcriptome profiles. Together, these data suggest that COPD-related etiologies and pathophysiologies do not impair responsiveness to resistance training, at least not for skeletal muscle characteristics, and at least not in the

Table 7 Effects of the training intervention on health-related quality of life in COPD and Healthy, measured using COPD Assessment Test (CAT; COPD-only, n = 20) and the 36-item Short Form Health Survey (SF-36; all participants; n COPD = 20, n Healthy = 57), and assessed as changes from baseline to after completion of the study (per study cluster; CAT and SF-36) and as differential changes between study clusters (SF-36)

	COPD			Healthy			Δ COPD vs Δ Healthy (*P* value)
	Baseline	Post RT	Time effect ($P < 0.05$)	Baseline	Post RT	Time effect ($P < 0.05$)	
COPD assessment Test™ score (0–40)	16.6 ± 6.8	16.4 ± 6.8	No	–	–	–	–
Short Form (36) Health Survey (0–100)							
Physical function*	63 ± 19	67 ± 18	No	90 ± 14	92 ± 12	No	0.321
Role physical*	43 ± 34	59 ± 37	Yes↑	87 ± 25	94 ± 18	No	0.226
Bodily pain	71 ± 27	82 ± 19	Yes↑	79 ± 21	80 ± 19	No	0.070
General health*	48 ± 20	56 ± 19	No	75 ± 18	80 ± 12	No	0.208
Vitality*	52 ± 16	57 ± 13	No	72 ± 18	78 ± 11	Yes↑	0.509
Social function*	74 ± 23	84 ± 16	Yes↑	90 ± 18	94 ± 13	No	0.280
Role emotional*	65 ± 39	84 ± 26	Yes↑	93 ± 19	96 ± 15	No	0.059
Mental health*	77 ± 13	84 ± 13	Yes↑	86 ± 11	89 ± 8	Yes↑	0.196

*difference between COPD and Healthy at baseline; ↑significant increase from baseline to after the training intervention (Post RT). Alpha level at $p < 0.05$. Values are means with standard deviation

enrolled cluster of COPD participants (GOLD grade II-III) and within the time frame of the current study.

During planning of the study protocol, two strategies were implemented to resolve the hypothesized, albeit rejected, negative impact of COPD-specific pathophysiologies for the efficacy of resistance training. *First*, as vitamin D insufficiency is common among COPD subjects [5], and has been suggested to contribute to development of anabolic resistance [61], dietary habits were manipulated to investigate the effects of vitamin D_3 supplementation. Contrary to our hypothesis, vitamin D_3 did not enhance responses to resistance training for any of the outcome variables [22]. *Second*, the resistance training protocol was conducted using two different training modalities, 10RM and 30RM resistance training, performed in a contralateral manner. The efficacies of these training modalities were initially hypothesized to be dissimilarly affected by COPD-related pathophysiologies, as they convey muscular adaptations through different signaling cues in the cellular environment (i.e. mechanical tension vs metabolic perturbation) [62], and may thus well be differentially affected by extracellular signaling such as inflammation and oxygen availability. While this hypothesis was rejected for all core outcome domains, with no differences being observed between training modalities and no evidence being found for the presence of impaired training responsiveness, a noteworthy observation was made for muscle fiber-specific traits. Specifically, in COPD, 10RM training was associated with blunted growth of type I muscle fiber CSA, a phenomenon that was not observed for responses to 30RM training, suggesting that 30RM offers benefits for muscle fiber type I hypertrophy. In addition to this, 10RM was associated with greater improvements in cycling economy and gross efficiency in COPD. These observations warrant further study. Of note, the unilateral resistance-training design was arguably supportive for the pronounced resistance-training effects in COPD participants. By reducing cardiorespiratory demand, and thus facilitating higher degrees of muscle activation and muscle mass-specific intensities during exercise compared to conventional two-legged resistance exercise [24] this seems to translate into larger functional improvements for this population [16].

Study limitations. Functional and physiological responses to resistance training is not uniform in the human population, and covary with individual characteristics such as genetics, epigenetics and composites of the inner physiological milieu [63–65]. For any research project that aim to understand the aetiology of such training, the interpretation of outcome data is thus a complicated task, which is further complicated by our present crude understanding of the interplay between the characteristics in question and their associated response patterns. While these limitations need to be acknowledged also in the current analyses, their presence underlines the importance of making study-design decisions that contribute to increase the ecological validity of the research project. As an example, in the present study, the advent of contralateral training protocols arguably increased the resolution of 10RM vs 30RM comparisons by removing genetic variability as a source of variation, albeit even

this perspective may have been affected by additional complications, such as the previously discussed cross-education effect. Furthermore, in any study, it is prudent to monitor, and ideally also account for, exogenous factors that may have impacted the physiological milieu, and therefore also training responses. In the present study, these included surveillance of lifestyle characteristics such as habitual dietary intake, activities of daily living and intake of pulmonary medication. For habitual patterns of dietary intake and activities of daily living, we observed no difference between study clusters (COPD vs Healthy), though it should be acknowledged that the collection of these data were performed only once during the study, and as such were performed using diary/questionnaire, making them prone to validity issues and warranting caution upon their interpretation [53, 66]. For pulmonary medication, the COPD and Healthy clusters deviated from each other for disease-related reasons, with 19 out of 20 COPD participants reporting intake of either $beta_2$-agonists, muscarinic agonists, or drugs containing a mixture of $beta_2$-agonists and glucocorticoids, as detailed in Table 1. These drugs are known to affect muscle biology and functionality in humans [67–70], and as such may have influenced the outcome of the study. However, the reported medication status of the COPD participants corresponds to what is normal for COPD subjects in general [71], and as such reflects the population intended to be studied.

In conclusion, 13-week resistance training program was well-tolerated by subjects with COPD and led to pronounced improvements for a range of health and muscle functional and biological variables, resembling or exceeding those seen in Healthy, with some outcome measures even showing indices of more beneficial adaptations in COPD participants with a more severe diagnosis. COPD was thus not associated with impaired responsiveness to resistance exercise training, which rather posed a potent measure to relieve disease-related pathophysiologies.

Abbreviations

COPD: Chronic obstructive pulmonary disease; RM: Repetition(s) maximum; FEV_1: Forced expiratory volume in one second; FVC: Forced vital capacity; CSA: Cross-sectional area; GSEA: Gene set enrichment analysis.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12967-021-02969-1.

Additional file 1. Additional tables and figure.

Acknowledgements

The authors would like to express their gratitude to all students involved in the study for invaluable assistance during intervention follow-up and data sampling. The authors also acknowledge the contributions to the study from MD Bjørn S. Svensgaard (Innlandet Hospital Trust), conducting the pre-inclusion consultations for the participants with COPD, biomedical laboratory technician Randi Sivesind (Innlandet Hospital Trust), performing the blood analyzes, Prof. Olivier Seynnes (Norwegian School of Sport Sciences) for instructions and education in ultrasound assessments, and Thomas Urianstad, Peter Nore Bengtsson, Gudmund Storlien, Joar Hansen and Anne Sofie Lofthus for valuable support. Finally, yet importantly, we would like to thank all study participants for their effortful and dedicated contributions.

This article was first published as a preprint: Mølmen KS, Hammarström D, Falch GS, Grundtvig M, Koll L, Hanestadhaugen M, Khan Y, Ahmad R, Malerbakken B, Rødølen TJ, Lien R, Rønnestad BR, Raastad T, Ellefsen S. Chronic Obstructive Pulmonary Disease Does Not Impair Responses to Resistance Training. *medRxiv*. https://doi.org/10.1101/2021.02.06.21251254

Authors' contributions

KSM and SE developed the project, with input from GSF, TJR, BRR and TR. KSM led the study intervention, including coordination and conduction of exercise training and testing, with aid from DH, GSF, BRR and SE. MG and TJR planned, organized and conducted participant recruitment and performed medical screening. BM and RL planned, organized and conducted lung spirometry and DXA measurements. KSM, DH, LK, MH, YK, RA and SE planned and performed muscle biological analyses. KSM, DH and SE planned and performed data analyses, with input from YK and RA. KSM, TR and SE drafted the manuscript. All authors provided useful input to data interpretation and contributed to drafting and finalizing the manuscript. All authors read and approved the final manuscript.

Funding

The study was funded by Inland Norway University of Applied Sciences, Innlandet Hospital Trust (grant number 150339, SE) and Regional Research Fund Inland Norway (grant number 298419, SE).

Availability of data and materials

A repository containing all transcriptome data and scripts used for transcriptome and enrichment analyses are available at https://github.com/dhammarstrom/rnaseq-copd. For other outcome measures, data are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

The study was approved by the Regional Committee for Medical and Health Research Ethics (reference no. 2013/1094) and all participants signed the informed consent prior to study enrolment.

Consent for publication

Not applicable.

Competing interests

The authors have no conflicts of interest to disclose.

Author details

[1]Section for Health and Exercise Physiology, Inland Norway University of Applied Sciences, P.O. Box 422, 2604 Lillehammer, Norway. [2]Department of Medicine, Innlandet Hospital Trust, Lillehammer, Norway. [3]Department of Pathology, Innlandet Hospital Trust, Lillehammer, Norway. [4]Department of Biotechnology, Inland Norway University of Applied Sciences, Hamar, Norway. [5]Institute of Clinical Medicine, Faculty of Health Sciences, UiT – The Arctic University of Norway, Tromsø, Norway. [6]Lillehammer Hospital for Rheumatic Diseases, Lillehammer, Norway. [7]Innlandet Hospital Trust, Granheim Lung Hospital, Follebu, Norway. [8]Department of Physical Performance, Norwegian School of Sport Sciences, Oslo, Norway. [9]Innlandet Hospital Trust, Lillehammer, Norway.

Received: 4 March 2021 Accepted: 28 June 2021
Published online: 06 July 2021

References

1. Spruit MA, Singh SJ, Garvey C, Zu Wallack R, Nici L, Rochester C, et al. An official American thoracic society/European respiratory society statement: key concepts and advances in pulmonary rehabilitation. Am J Respir Crit Care Med. 2013;188(8):e13-64.
2. Hajiro T, Nishimura K, Tsukino M, Ikeda A, Toru O, Izumi T. A comparison of the level of dyspnea vs disease severity in indicating the health-related quality of life of patients with COPD. Chest. 1999;116(6):1632–7. https://doi.org/10.1378/chest.116.6.1632.
3. Pitta F, Troosters T, Spruit MA, Probst VS, Decramer M, Gosselink R. Characteristics of physical activities in daily life in chronic obstructive pulmonary disease. Am J Respir Crit Care Med. 2005;171(9):972–7.
4. Debigaré R, Marquis K, Côté CH, Tremblay RR, Michaud A, LeBlanc P, et al. Catabolic/anabolic balance and muscle wasting in patients with COPD. Chest. 2003;124(1):83–9.
5. Janssens W, Bouillon R, Claes B, Carremans C, Lehouck A, Buysschaert I, et al. Vitamin D deficiency is highly prevalent in COPD and correlates with variants in the vitamin D-binding gene. Thorax. 2010;65(3):215–20. https://doi.org/10.1136/thx.2009.120659.
6. Dawson-Hughes B. Vitamin D and muscle function. J Steroid Biochem Mol Biol. 2017;173(March):313–6.
7. Wüst RCI, Degens H. Factors contributing to muscle wasting and dysfunction in COPD patients. Int J COPD. 2007;2(3):289–300.
8. Gan WQ, Man SFP, Senthilselvan A, Sin DD. Association between chronic obstructive pulmonary disease and systemic inflammation: a systematic review and a meta-analysis. Thorax. 2004;59(7):574–80.
9. Van De Bool C, Steiner MC, Schols AMWJ. Nutritional targets to enhance exercise performance in chronic obstructive pulmonary disease. Curr Opin Clin Nutr Metab Care. 2012;15(6):553–60.
10. Fisher G, Scott Bickel C, Hunter GR. Elevated circulating TNF-α in fat-free mass non-responder compared to responders following exercise training in older women. Biology (Basel). 2014;3(3):551–9.
11. Antoniak AE, Greig CA. The effect of combined resistance exercise training and Vitamin D3 supplementation on musculoskeletal health and function in older adults: A systematic review and meta-analysis. BMJ Open. 2017;7(7):1–16.
12. Rønnestad BR, Nygaard H, Raastad T. Physiological elevation of endogenous hormones results in superior strength training adaptation. Eur J Appl Physiol. 2011;111(9):2249–59.
13. Vonbank K, Strasser B, Mondrzyk J, Marzluf BA, Richter B, Losch S, et al. Strength training increases maximum working capacity in patients with COPD - Randomized clinical trial comparing three training modalities. Respir Med. 2012;106(4):557–63. https://doi.org/10.1016/j.rmed.2011.11.005.
14. Hoff J, Tjønna AE, Steinshamn S, Høydal M, Richardson RS, Helgerud J. Maximal strength training of the legs in COPD: a therapy for mechanical inefficiency. Med Sci Sports Exerc. 2007;39(2):220–6.
15. Kongsgaard M, Backer V, Jørgensen K, Kjær M, Beyer N. Heavy resistance training increases muscle size, strength and physical function in elderly male COPD-patients—a pilot study. Respir Med. 2004;98(10):1000–7.
16. Nyberg A, Martin M, Saey D, Milad N, Patoine D, Morissette MC, et al. Effects of low-load/high-repetition resistance training on exercise capacity, health status, and limb muscle adaptation in patients with severe COPD: a randomized controlled trial. Chest. 2021;159(5):1821–32. https://doi.org/10.1016/j.jhazmat.2020.124187.
17. Remels AHV, Gosker HR, Langen RCJ, Schols AMWJ. The mechanisms of cachexia underlying muscle dysfunction in COPD. J Appl Physiol. 2013;114(9):1253–62.
18. Constantin D, Menon MKM, Houchen-Wolloff L, Morgan MD, Singh SJ, Greenhaff P, et al. Skeletal muscle molecular responses to resistance training and dietary supplementation in COPD. Thorax. 2013;68(7):1–19.
19. Menon MK, Houchen L, Harrison S, Singh SJ, Morgan MD, Steiner MC. Ultrasound assessment of lower limb muscle mass in response to resistance training in COPD. Respir Res. 2012;13(1):119.
20. Menon MK, Houchen L, Singh SJ, Morgan MD, Bradding P, Steiner MC. Inflammatory and satellite cells in the quadriceps of patients with COPD and response to resistance training. Chest. 2012;142(5):1134–42.
21. Sanders KJC, Kneppers AEM, van de Bool C, Langen RCJ, Schols AMWJ. Cachexia in chronic obstructive pulmonary disease: New insights and therapeutic perspective. J Cachexia Sarcopenia Muscle. 2016;7(1):5–22.
22. Mølmen KS, Hammarström D, Pedersen K, Lian Lie AC, Steile RB, Nygaard H, et al. Vitamin D 3 supplementation does not enhance the effects of resistance training in older adults. J Cachexia Sarcopenia Muscle. 2021;12(3):599–628.
23. Global initiative for chronic obstructive lung disease: Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease (2020 report). 2020. Available from: https://goldcopd.org/wp-content/uploads/2019/12/GOLD-2020-FINAL-ver1.2-03Dec19_WMV.pdf
24. Mølmen KS, Evensen Thy J, Thallaug Dalane S, Ellefsen S, Falch GS. Muscular performance decreases with increasing complexity of resistance exercises in subjects with chronic obstructive pulmonary disease. Transl Sport Med. 2020;3(1):26–33. https://doi.org/10.1002/tsm2.118.
25. Khan Y, Hammarström D, Rønnestad BR, Ellefsen S, Ahmad R. Increased biological relevance of transcriptome analyses in human skeletal muscle using a model-specific pipeline. BMC Bioinformatics. 2020;21(1):1–32. https://doi.org/10.1186/s12859-020-03866-y.
26. 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. J Physiol. 2020;598(3):543–65.
27. 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(2):87–105.
28. Brooks ME, Kristensen K, Van Benthem KJ, Magnusson A, Berg CW, Nielsen A, et al. glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. RJ. 2017;9(2):378.
29. Korotkevich G, Sukhov V. Fast gene set enrichment analysis. bioRxiv. 2019. https://doi.org/10.1101/060012
30. Liberzon A, Birger C, Thorvaldsdóttir H, Ghandi M, Mesirov JP, Tamayo P. The molecular signatures database hallmark gene set collection. Cell Syst. 2015;1(6):417–25.
31. Liberzon A, Subramanian A, Pinchback R, Thorvaldsdóttir H, Tamayo P, Mesirov JP. Molecular signatures database (MSigDB) 3.0. Bioinformatics. 2011;27(12):1739–40.
32. R Core Team. R: a language and environment for statistical computing. R foundation for statistical computing. Vienna: R Core Team; 2018.
33. Eliason G, Abdel-Halim S, Arvidsson B, Kadi F, Piehl-Aulin K. Physical performance and muscular characteristics in different stages of COPD. Scand J Med Sci Sport. 2009;19(6):865–70.
34. Baumgartner RN, Koehler KM, Gallagher D, Romero L, Heymsfield SB, Ross RR, et al. Epidemiology of sarcopenia among the elderly in New Mexico. Am J Epidemiol. 1998;147(8):755–63.
35. Flynn MA, Nolph GB, Baker AS, Krause G. Aging in humans: a continuous 20-year study of physiologic and dietary parameters. J Am Coll Nutr. 1992;11(6):660–72.
36. Sharanya A, Ciano M, Withana S, Kemp PR, Polkey MI, Sathyapala SA. Sex differences in COPD-related quadriceps muscle dysfunction and fibre abnormalities. Chron Respir Dis. 2019. https://doi.org/10.1177/1479973119843650.
37. Gosker HR, van Mameren H, van Dijk PJ, Engelen MPKJ, van der Vusse GJ, Wouters EFM, et al. Skeletal muscle fibre-type shifting and metabolic profile in patients with chronic obstructive pulmonary disease. Eur Respir J. 2002;19(4):617–25.
38. Eliason G, Abdel-Halim SM, Piehl-Aulin K, Kadi F. Alterations in the muscle-to-capillary interface in patients with different degrees of chronic obstructive pulmonary disease. Respir Res. 2010;11:97.
39. Whittom F, Jobin J, Simard P-M, LeBlanc P, Simard C, Bernard S, et al. Histochemical and morphological characteristics of the vastus lateralis muscle in patients with chronic obstructive pulmonary disease. Med Sci Sport Exerc. 1998;30(10):1467–74.
40. Aagaard P, Suetta C, Caserotti P, Magnusson SP, Kjær M. Role of the nervous system in sarcopenia and muscle atrophy with aging: Strength training as a countermeasure. Scand J Med Sci Sport. 2010;20(1):49–64.
41. Hepple RT, Ross KD, Rempfer AB. Fiber Atrophy and Hypertrophy in Skeletal Muscles of Late Middle-Aged Fischer 344 x Brown Norway F1-Hybrid Rats. Journals Gerontol Ser A Biol Sci Med Sci. 2004;59(2):B108–17. https://doi.org/10.1093/gerona/59.2.B108.
42. Engelen MPKJ, Deutz NEP, Wouters EFM, Schols AMWJ. Enhanced levels of whole-body protein turnover in patients with chronic obstructive pulmonary disease. Am J Respir Crit Care Med. 2000;162:1488–92.

43. Kneppers AEM, Langen RCJ, Gosker HR, Verdijk LB, Cebron Lipovec N, Leermakers PA, et al. Increased myogenic and protein turnover signaling in skeletal muscle of chronic obstructive pulmonary disease patients with sarcopenia. J Am Med Dir Assoc. 2017;18(7):637.e1-637.e11. https://doi.org/10.1016/j.jamda.2017.04.016.
44. Baum JI, Kim IY, Wolfe RR. Protein consumption and the elderly: What is the optimal level of intake? Nutrients. 2016;8(6):1–9.
45. MacInnis MJ, McGlory C, Gibala MJ, Phillips SM. Investigating human skeletal muscle physiology with unilateral exercise models: when one limb is more powerful than two. Appl Physiol Nutr Metab. 2017;42(6):563–70.
46. Hendy AM, Lamon S. The cross-education phenomenon: Brain and beyond. Front Physiol. 2017;8(MAY):1–9.
47. Kiens B, Essen-Gustavsson B, Christensen NJ, Saltin B. Skeletal muscle substrate utilization during submaximal exercise in man: effect of endurance training. J Physiol. 1993;469:459–78.
48. Rud B, Foss O, Krustrup P, Secher NH, Hallén J. One-legged endurance training: leg blood flow and oxygen extraction during cycling exercise. Acta Physiol (Oxf). 2012;205(1):177–85.
49. Jensen L, Bangsbo J, Hellsten Y. Effect of high intensity training on capillarization and presence of angiogenic factors in human skeletal muscle. J Physiol. 2004;557(Pt 2):571–82.
50. Miller BF, Olesen JL, Hansen M, Døssing S, Crameri RM, Welling RJ, et al. Coordinated collagen and muscle protein synthesis in human patella tendon and quadriceps muscle after exercise. J Physiol. 2005;567(3):1021–33.
51. Wilkinson SB, Tarnopolsky MA, Grant EJ, Correia CE, Phillips SM. Hypertrophy with unilateral resistance exercise occurs without increases in endogenous anabolic hormone concentration. Eur J Appl Physiol. 2006;98(6):546–55.
52. Houston ME, Froese EA, Valeriote SP, Green HJ, Ranney DA. Muscle performance, morphology and metabolic capacity during strength training and detraining: A one leg model. Eur J Appl Physiol Occup Physiol. 1983;51(1):25–35.
53. Sember V, Meh K, Sorić M, Jurak G, Starc G, Rocha P. Validity and reliability of international physical activity questionnaires for adults across eu countries: Systematic review and meta analysis. Int J Environ Res Public Health. 2020;17(19):1–23.
54. Rønnestad BR, Mujika I. Optimizing strength training for running and cycling endurance performance: A review. Scand J Med Sci Sport. 2014;24(4):603–12.
55. Liao W-H, Chen J-W, Chen X, Lin L, Yan H-Y, Zhou Y-Q, et al. Impact of resistance training in subjects with COPD: a systematic review and meta-analysis. Respir Care. 2015;60(8):1130–45.
56. Bjørnsen T, Wernbom M, Kirketeig A, Paulsen G, Samnøy L, Bækken L, et al. Type 1 muscle fiber hypertrophy after blood flow-restricted training in powerlifters. Med Sci Sports Exerc. 2019;51(2):288–98.
57. 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(8):E652–61.
58. Luoto J, Pihlsgård M, Wollmer P, Elmståhl S. Relative and absolute lung function change in a general population aged 60–102 years. Eur Respir J. 2019. https://doi.org/10.1183/13993003.01812-2017.
59. Costes F, Gosker H, Feasson L, Desgeorges M, Kelders M, Castells J, et al. Impaired exercise training-induced muscle fiber hypertrophy and Akt/mTOR pathway activation in hypoxemic patients with COPD. J Appl Physiol. 2015;118(8):1040–9.
60. Spruit MA, Gosselink R, Troosters T, Kasran A, Van Vliet M, Decramer M. Low-grade systemic inflammation and the response to exercise training in patients with advanced COPD. Chest. 2005;128(5):3183–90. https://doi.org/10.1378/chest.128.5.3183.
61. Dzik KP, Kaczor JJ. Mechanisms of vitamin D on skeletal muscle function: oxidative stress, energy metabolism and anabolic state. Eur J Appl Physiol. 2019;119(4):825–39.
62. American College of Sports Medicine. American College of Sports Medicine position stand. Progression models in resistance training for healthy adults. Med Sci Sports Exerc. 2009;41(3):687–708.
63. Rea IM. Towards ageing well: Use it or lose it: Exercise, epigenetics and cognition. Biogerontology. 2017;18(4):679–91.
64. Thomaes T, Thomis M, Onkelinx S, Goetschalckx K, Fagard R, Lambrechts D, et al. Genetic predisposition scores associate with muscular strength, size, and trainability. Med Sci Sports Exerc. 2013;45(8):1451–9.
65. Kraemer WJ, Ratamess NA. Hormonal responses and adaptations to resistance exercise and training. Sports Med. 2005;35(4):339–61.
66. Burrows TL, Ho YY, Rollo ME, Collins CE. Validity of dietary assessment methods when compared to the method of doubly labeled water: a systematic review in adults. Front Endocrinol. 2019. https://doi.org/10.3389/fendo.2019.00850.
67. Hostrup M, Jacobson GA, Jessen S, Lemminger AK. Anabolic and lipolytic actions of beta2-agonists in humans and antidoping challenges. Drug Test Anal. 2020;12(5):597–609. https://doi.org/10.1002/dta.2728.
68. Caruso JF, Hamill JL, De Garmo N. Oral albuterol dosing during the latter stages of a resistance exercise program. J strength Cond Res. 2005;19(1):102–7.
69. Sato AY, Peacock M, Bellido T. Glucocorticoid Excess in Bone and Muscle. Clin Rev Bone Miner Metab. 2018;16(1):33–47.
70. Arlettaz A, Portier H, Lecoq AM, Rieth N, De Ceaurriz J, Collomp K. Effects of short-term prednisolone intake during submaximal exercise. Med Sci Sports Exerc. 2007;39(9):1672–8.
71. Raluy-Callado M, Lambrelli D, Maclachlan S, Khalid JM. Epidemiology, severity, and treatment of chronic obstructive pulmonary disease in the United Kingdom by GOLD 2013. Int J COPD. 2015;10:925–37.
72. Miller MR, Hankinson J, Brusasco V, Burgos F, Casaburi R, Coates A, et al. Standardisation of spirometry. Eur Respir J . 2005;26(2):319–38. https://doi.org/10.1183/09031936.05.00034805.
73. Péronnet F, Massicotte D. Table of nonprotein respiratory quotient: an update. Can J Sport Sci. 1991;16(1):23–9.
74. Jetté M, Sidney K, Blümchen G. Metabolic equivalents (METS) in exercise testing, exercise prescription, and evaluation of functional capacity. Clin Cardiol. 1990;13(8):555–65.
75. Ellefsen S, Vikmoen O, Zacharoff E, Rauk I, Slettaløkken G, Hammarström 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(5):e332–42. https://doi.org/10.1111/sms.12185.

Publisher's Note

4

ORIGINAL ARTICLE

Journal of Cachexia, Sarcopenia and Muscle 2021; **12**: 599–628
Published online 31 March 2021 in Wiley Online Library (wileyonlinelibrary.com) **DOI:** 10.1002/jcsm.12688

Vitamin D_3 supplementation does not enhance the effects of resistance training in older adults

Knut Sindre Mølmen[1*], Daniel Hammarström[1], Karianne Pedersen[1], Anne Cecilie Lian Lie[1], Ragnvald B. Steile[1], Håvard Nygaard[1], Yusuf Khan[1,2], Håvard Hamarsland[1], Lise Koll[3], Marita Hanestadhaugen[3], Atle Lie Eriksen[3], Eirik Grindaker[1], Jon Elling Whist[3], Daniel Buck[1], Rafi Ahmad[2,4], Tor A. Strand[3,5], Bent R. Rønnestad[1] & Stian Ellefsen[1,3]

[1]*Section for Health and Exercise Physiology, Inland Norway University of Applied Sciences, Lillehammer, Norway;* [2]*Department of Biotechnology, Inland Norway University of Applied Sciences, Hamar, Norway;* [3]*Innlandet Hospital Trust, Lillehammer, Norway;* [4]*Institute of Clinical Medicine, Faculty of Health Sciences, UiT – The Arctic University of Norway, Tromsø, Norway;* [5]*Centre for International Health, University of Bergen, Bergen, Norway*

Abstract

Background Lifestyle therapy with resistance training is a potent measure to counteract age-related loss in muscle strength and mass. Unfortunately, many individuals fail to respond in the expected manner. This phenomenon is particularly common among older adults and those with chronic diseases (e.g. chronic obstructive pulmonary disease, COPD) and may involve endocrine variables such as vitamin D. At present, the effects of vitamin D supplementation on responses to resistance training remain largely unexplored.

Methods Ninety-five male and female participants (healthy, n = 71; COPD, n = 24; age 68 ± 5 years) were randomly assigned to receive either vitamin D_3 or placebo supplementation for 28 weeks in a double-blinded manner (latitude 61°N, September–May). Seventy-eight participants completed the RCT, which was initiated by 12 weeks of supplementation-only (two weeks with 10 000 IU/day, followed by 2000 IU/day), followed by 13 weeks of combined supplementation (2000 IU/day) and supervised whole-body resistance training (twice weekly), interspersed with testing and measurements. Outcome measures included multiple assessments of muscle strength ($n_{variables}$ = 7), endurance performance (n = 6), and muscle mass (n = 3, legs, primary), as well as muscle quality (legs), muscle biology (*m. vastus lateralis*; muscle fibre characteristics, transcriptome), and health-related variables (e.g. visceral fat mass and blood lipid profile). For main outcome domains such as muscle strength and muscle mass, weighted combined factors were calculated from the range of singular assessments.

Results Overall, 13 weeks of resistance training increased muscle strength (13% ± 8%), muscle mass (9% ± 8%), and endurance performance (one-legged, 23% ± 15%; whole-body, 8% ± 7%), assessed as weighted combined factors, and were associated with changes in health variables (e.g. visceral fat, −6% ± 21%; $[LDL]_{serum}$, −4% ± 14%) and muscle tissue characteristics such as fibre type proportions (e.g. IIX, −3% points), myonuclei per fibre (30% ± 65%), total RNA/rRNA abundances (15%/6–19%), and transcriptome profiles (e.g. 312 differentially expressed genes). Vitamin D_3 supplementation did not affect training-associated changes for any of the main outcome domains, despite robust increases in $[25(OH)D]_{serum}$ (Δ49% vs. placebo). No conditional effects were observed for COPD vs. healthy or pre-RCT $[25(OH)D]_{serum}$. In secondary analyses, vitamin D_3 affected expression of gene sets involved in vascular functions in muscle tissue and strength gains in participants with high fat mass, which advocates further study.

Conclusions Vitamin D_3 supplementation did not affect muscular responses to resistance training in older adults with or without COPD.

Keywords Strength training; Cholecalciferol; Muscle plasticity

Received: 10 July 2020; Revised: 30 December 2020; Accepted: 26 January 2021
**Correspondence to: Knut Sindre Mølmen, Section for Health and Exercise Physiology, Inland Norway University of Applied Sciences, P.O. Box. 422, 2604 Lillehammer, Norway. Email:knut.sindre.molmen@inn.no*
This article was first published as a preprint: Mølmen KS, Hammarström D, Pedersen K, Lian Lie AC, Steile RB, Nygaard H, Khan Y, Hamarsland H, Koll L, Hanestadhaugen M,

Lie Eriksen A, Grindaker E, Whist JE, Buck D, Ahmad R, Strand TA, Rønnestad BR, Ellefsen S. 2020. No benefits of prolonged vitamin D_3 supplementation for adaptations to resistance training in old adults. medRxiv. https://doi.org/10.1101/2020.07.10.20150037.

Introduction

Aging is associated with progressive loss of muscle strength and mass, accompanied by declines in physical performance. In 2016, this had escalated into ~11 million Europeans (>65 years of age) suffering from sarcopenia,[1] a formally recognized disease characterized by severe loss of muscle quantity and quality.[1] Sarcopenia increases the likelihood of adverse events such as falling, fractures, physical disability, morbidity and mortality,[2,3] further fuelling muscle deterioration, resulting in a spiralling decrease in overall health and health-related quality of life.[4–6] In Europe, the prevalence of sarcopenia is expected to increase to at least ~19 million by 2045,[1] coinciding with increasing proportions of older adults, potentiated by suboptimal nutrition and increasing incidences of causal morbidities such as systemic inflammatory diseases.[7,8] For elderly to stay healthy, active and independent, efficient interventions are warranted for its prevention, treatment and reversal.[7,8] To this end, lifestyle therapy with resistance training is an attractive, low-cost and potent intervention.[9,10] Unfortunately, the benefits of such interventions are not always consistent, especially in the older population, with selected individuals and populations showing impaired abilities to increase muscle strength and mass.[11,12] At present, this training-response-spectrum has an unknown causality, although it interdepends on factors such as genetics,[13,14] epigenetics,[14] and composites of the inner physiological milieu, including nutrition,[15,16] endocrine variables (e.g. vitamin D),[17,18] and hallmarks of health such as low-grade chronic inflammation.[19] There is thus a need for development of combinatorial lifestyle protocols that target and correct these factors alongside resistance training, thereby allowing adequate muscle adaptations to occur.

Over the last two decades, vitamin D has emerged as a potential determinant of muscle functionality and biology.[20] There seems to be a robust relationship between heterogeneity in vitamin D status and traits such as physical performance[21–23] and susceptibility to falling,[24] suggesting a causal association between vitamin D and increased risk of sarcopenia.[25] As such, vitamin D status varies substantially in the human population, both in an annual cycle, and between individuals and groups of individuals.[26,27] Vitamin D insufficiency is particularly prevalent in older adults, measured as 25-hydroxyvitamin D (25(OH)D) levels <50 nmol/L, and especially in older adults living in the Northern Hemisphere,[27,28] where cutaneous vitamin D synthesis is miniscule or absent during winter months.[29] In accordance with this, exogenous vitamin D supplementation is gaining momentum as a potential ergogenic aid for preventing and treating sarcopenia.[25] Unfortunately, the presumed benefits of vitamin D supplementation deduced from crossover studies are not necessary supported by data from interventional studies. While some studies and meta-analyses report favourable effects of vitamin D supplementation *per se* on muscle strength[30–32] and falling,[33,34] with benefits being more pronounced in subjects with low baseline values (<30 nmol/L)[35] and in older subjects,[35] others do not.[36–39] These discrepancies may not be surprising, as resistance training is arguably necessary to provoke changes in muscle functions.[40] However, a similar ambiguity is present in the few studies that have assessed the effects of vitamin D supplementation on outcomes of resistance training.[41–44] While none of these studies report clear benefits of vitamin D supplementation for alterations in muscle strength,[41–44] muscle mass,[42–44] or incidences of falling,[41,43] a recent meta-analysis still concluded that it provides benefits for training-associated changes in lower body muscle strength.[40]

Consequently, we have limited and conflicting knowledge about the combined effects of vitamin D supplementation and resistance training for muscle functions and biology in humans. The present confusion may partly be attributed to methodological uncertainties in available studies, potentially lowering their ecological validity and explaining their lack of coherence with the resulting meta-analysis data. This includes heterogeneous study populations (varying from young adults[42,44] to older adults[44] to elderly[41,43]) with large differences in baseline 25(OH)D levels (average 31 nmol/L[43]–71 nmol/L[44]), large variation in vitamin D dosage (from 400 IU/day[43]–4000 IU/day[42]), lack of familiarization to strength tests,[41,43] suboptimal training protocols[41,43] (failing to comply to current guidelines, advocating resistance training with controlled maximal effort[45,46]), low compliance to training,[41,43] and a lack of dietary assessment during the intervention.[41,43,44] Also, neither of the studies included a period of vitamin D supplementation prior to resistance training, which may be necessary to prime muscle cells for adaptations, potentially acting by changing epigenetic traits, which has been observed in other cell types, such as T-cells[47] and oral squamous cell carcinoma cells.[48] Furthermore, the effects of vitamin D supplementation on muscle fibre characteristics and biology remain poorly understood and unclear.[49] In theory, vitamin D may potentiate muscle fibre responsiveness in two ways. Either directly by acting through vitamin D receptors in muscle fibres or progenitor cells, perhaps inducing intramuscular signalling pathways such as the p38 mitogen-activated protein kinase pathway,[50,51] or indirectly by interacting with systemic signalling event, perhaps inducing testosterone signalling[52] and thereby facilitating muscle plasticity. Our lack of insight is underlined by the longstanding uncertainty of the presence

of vitamin D receptors in muscle tissue,[53] although several indications advocate its expression. First, there seems to be associations between mutations in the vitamin D receptor and muscle weakness in both humans and mice.[54,55] Second, muscle-specific knock-out of the vitamin D receptor in mice deteriorates muscle strength and mass in a manner that resemble sarcopenia.[56,57] The prevailing uncertainty is fuelled by a seeming lack of effects of vitamin D supplementation *per se* on the muscle transcriptome in vitamin D-insufficient frail elderly, although also in that study the vitamin D dosage was relatively low (400 IU/day).[58] To date, a mere single study has assessed the effects of vitamin D supplementations on resistance training-induced muscle biological adaptations in humans, and as such assessing only a limited selection of traits and failing to disclose conclusive findings.[44]

The aim of the present study was to investigate the effects of 12 weeks of vitamin D_3 supplementation only (the initial two weeks with 10 000 international units (IU)/day, succeeded by 10 weeks with 2000 IU/day), followed by 13 weeks of combined vitamin D_3 supplementation (2000 IU/day) and resistance training, on training-associated adaptations in a mixed population of older subjects. The RCT thus allowed assessment of responses to both vitamin D_3 supplementation-only and combined vitamin D_3 supplementation and resistance training. The study population included individuals that were either at risk of developing sarcopenia (age or disease, i.e. COPD patients)[59,60] or showed diagnostic indications of sarcopenia (16.4% of the participants had appendicular lean mass (kg)/m^2 greater than two standard deviations below the sex-specific means of young adults).[61] Outcome measures included a large range of muscle strength and endurance performance tests, multiple assessments of muscle mass, muscle quality, in-depth analyses of muscle biology including muscle fibre characteristics and analyses of the muscle transcriptome, and a range of health-related measures including body composition, blood variables and self-reported health variables.

Methods

Study ethics and participants

The study was approved by the Regional Committee for Medical and Health Research Ethics - South-East Norway (reference no: 2013/1094) and was preregistered at ClinicalTrials.gov (ClinicalTrials.gov Identifier: NCT02598830). All participants were informed about the potential risks and discomforts associated with the study and gave their informed consent prior to study enrolment. The study was conducted according to the Declaration of Helsinki.

Ninety-five male and female participants (age 68 ± 5 years, range 56–77) were enrolled into the study (*Figure* 1). Eligibility criteria were consumption of less than 400 international units (IU) of vitamin D_3 per day for the two months leading up to the study, and either normal lung function or medical diagnosis of COPD (GOLD[62] grade II or III, FEV_1 predicted between 80% - 30%, $FEV_1/FVC < 70\%$ after reversibility testing with inhalation of salbutamol and ipratropiumbromid). Exclusion criteria were unstable cardiovascular disease, chronic granulomatous disease, known active malignancy within the last five years, serious psychiatric comorbidity, steroid use the previous two months and musculoskeletal disorders preventing the participant from participating in the resistance training programme. Initially, all participants were screened using spirometry and a medical questionnaire. For healthy participants, this formed the basis for inclusion. For COPD participants and participants with unclear disease status, the initial screening was followed by consultation with a medical doctor to ensure that they met diagnostic criteria corresponding to GOLD grade II or III, followed by inclusion. All participants were recreationally active, but none had partaken in systematic resistance training for the 12 months leading up to the study. During study conduct, all participants were instructed to restrict vitamin D intake from food sources to <400 IU/day and to abstain from solarium and travels to southern and/or sunny areas.

Participants were randomly assigned into one of the two study arms (vitamin D_3 vs. placebo) using concealed allocation, stratified by sex and health status (COPD vs. non-COPD) (*Figure* 1 and *Table* 1). An off-site third party performed the randomization. During the initial two weeks of the study, the vitamin D_3 arm consumed 10 000 IU vitamin D_3/day, followed by 2000 IU/day for the remainder of the study period. Placebo capsules contained cold-pressed olive oil and were identical in appearance to vitamin D_3 capsules. Pharma Nord ApS (Vejle, Denmark) produced the two supplements, complying with Good Manufacturing Practice requirements. All participants consumed 500 mg calcium/day (Nycoplus, Takeda AS, Asker, Norway). Vitamin D status was primarily assessed as 25(OH)D levels in blood (*Figure* 2), corroborating with previous studies,[63] and secondarily as 1,25 dihydroxycholecalciferol ($1,25(OH)_2D$; the biologically active form). 25(OH)D is accepted to be the most reliable measure of vitamin D status,[64] as it is unaffected by parathyroid hormone (PTH) activity, and is more stable and represents more accurate measurements compared with $1,25(OH)_2D$.[64]

Of the 95 participants included in the study, one withdrew from the study prior to onset on supplementation, 12 withdrew prior to onset of resistance training (vitamin D_3 arm, n = 9; placebo arm, n = 3), and 4 participants withdrew during the resistance training period (vitamin D_3 arm, n = 3; placebo arm, n = 1) (*Figure* 1). In summary, 78 participants completed the study; 58 healthy participants and 20 COPD participants. For participant characteristics, see *Table* 1.

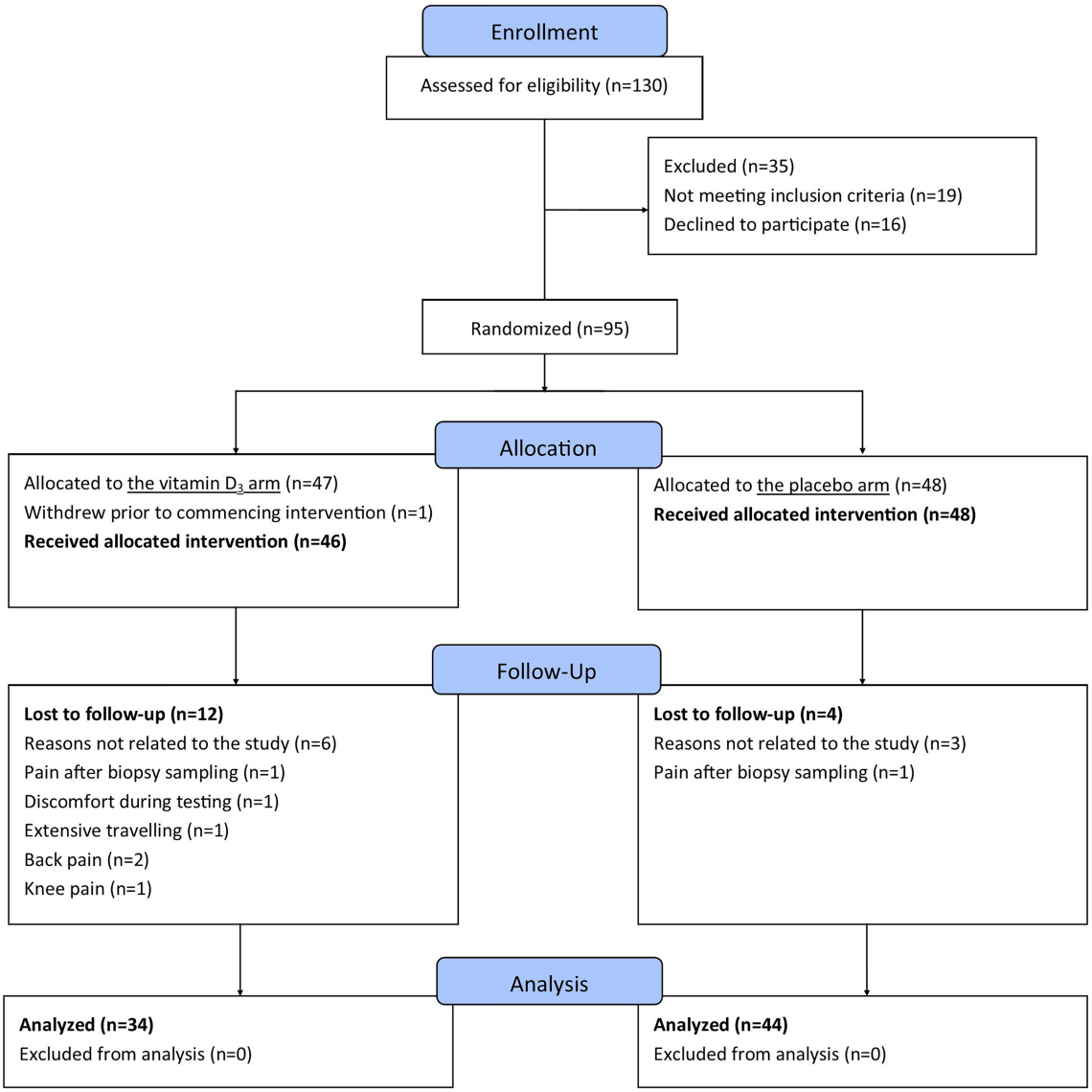


Figure 1 CONSORT flow chart of the study.

Study conduct

The study was conducted as a double-blind randomized clinical trial (RCT), consisting of an initial 12 weeks of supplementation-only (in average, 3333 IU vitamin D_3/day or placebo; 14 days of 10 000 IU vitamin D_3/day, 10 weeks of 2000 IU/day), followed by 13 weeks of combined supplementation (2000 IU vitamin D_3/day or placebo) and resistance training (*Figure* 2). During study conduct, supplement allocation was blinded for both participants and investigators. Unblinding was performed after completion of primary outcome measure clean-up and analyses. The intervention was conducted at Lillehammer, Norway (latitude 61°N) from September to May, ensuring low or no natural vitamin D synthesis by the skin from sunlight UVB radiation.[29] Prior to onset of the supplementation protocol (i.e. pre-RCT), participants undertook two weeks of baseline testing and tissue/blood sampling (*Figure* 2, Weeks −2 and −1), including testing of unilateral strength and muscle performance (tested twice, separated by at least 48 h; the first test was performed at ~95% of maximal effort), lung function, and collection of fasting blood and rested-state muscle biopsy, sampled from *m. vastus lateralis* of the dominant leg using the microbiopsy technique (Bard Magnum, Bard, Covington, GA, USA). Thereafter, participants were randomized to the two supplementation arms. After two weeks of supplementation, a second blood sample was collected (*Figure* 2, Week 2) to validate the efficacy of vitamin D_3 supplementation for blood 25(OH)D and $1,25(OH)_2D$ levels. Prior to introduction to resistance training, the participants conducted repeated

Table 1 Participant characteristics

	Vitamin D_3 arm	Placebo arm
Participants (*n*)	46	48
Females (*n*)	24	27
COPD subjects (*n*)	12	12
Age (years ± SD)	69 ± 5	67 ± 4
Weight (kg ± SD)	75 ± 17	75 ± 16
Lean mass (kg ± SD)	48 ± 11	48 ± 9
Fat percentage (% ± SD)	35 ± 6	34 ± 9
Body mass index (kg/m^2 ± SD)	26 ± 5	26 ± 5
1RM knee extension (kg ± SD)	18 ± 8	18 ± 7
1RM chest press (kg ± SD)	47 ± 17	45 ± 16
Withdrawn prior to intro. RT (*n*)	9	3
Withdrawn after intro. RT (*n*)	3	1
Renal function		
Creatinine (μmol/L)	78 ± 18	80 ± 22
Est. GFR (mL/min/1.73 m^2)	80 ± 15	79 ± 15
CKD stage 3, i.e. est. GFR of 30–59 (*n*)	2	3
Lung function		
FVC (L ± SD)	3.4 ± 0.8	3.6 ± 0.9
FEV_1/FVC (% ± SD)	67 ± 15	69 ± 14
FEV_1 (% predicted ± SD)	87 ± 24	94 ± 26
PEF (L/s ± SD)	6.9 ± 2.4	7.1 ± 2.1
Habitual dietary data		
Kilocalories/day ± SD	1777 ± 529	1985 ± 611
Protein (g/kg/day ± SD)	1.26 ± 0.40	1.27 ± 0.36
Fat (g/kg/day ± SD)	0.99 ± 0.47	1.05 ± 0.38
Carbohydrates (g/kg/day ± SD)	2.46 ± 1.05	2.88 ± 1.03
Alcohol (units/day ± SD)	0.76 ± 0.92	0.67 ± 1.04
Vitamin D (IU/day ± SD)	281 ± 235	331 ± 260
Other vitamin D exposures		
Number of hours outdoors per week	8.8 ± 6.0	8.9 ± 6.4
Fish for dinner per week	1.9 ± 0.8	1.8 ± 0.7
Fish for other meals per week	2.0 ± 1.7	1.6 ± 1.1
Cod liver oil (teaspoons per week)	1.2 ± 3.8	1.6 ± 3.4
Cod liver oil (capsules per week)	1.5 ± 3.8	2.0 ± 3.8
Number of eggs eaten per week	3.2 ± 1.8	2.9 ± 2.2
Adherence		
Adherence to supplementation plan (%)	99 (91–100)	99 (93–100)
Adherence to the training protocol (%)	98 (81–100)	98 (81–100)

Training volume (kg x repetitions)	Leg press	Knee extension	RPE	Leg press	Knee extension	RPE
Training week 1 (Introduction period, week 1)	4074 (1741)	298 (143)	15.4 (1.4)	4307 (1737)	360 (206)	15.4 (1.5)
Training week 4 (Training period, week 1)	5117 (2199)	364 (187)	15.9 (1.4)	5393 (2247)	407 (201)	16.0 (1.3)
Training week 8 (Training period, week 5)	6071 (2710)	446 (233)	16.5 (1.5)	6200 (2638)	495 (255)	16.6 (1.3)
Training week 13 (Training period, week 10)	6698 (3183)	489 (255)	17.0 (1.3)	6706 (2598)	550 (293)	17.1 (1.2)

1RM, one repetition maximum; CKD, chronic kidney disease; FEV_1, forced expiratory volume in 1 s; FVC, forced vital capacity; GFR, glomerular filtration rate (calculated using the *Modification of Diet in Renal Disease study* equation; IU, international units; PEF, peak expiratory flow; RT, resistance training; RPE, rating of perceived exertion (6–20).

performance tests at several occasions (*Figure* 2, Week −2–Week 13), including unilateral maximal strength and muscular performance, isokinetic unilateral knee-extension torque, measures of functional capacity (i.e. 6-min step and 1-min sit-to-stand test), submaximal and maximal one-legged cycling, and maximal bicycling. During the last week before introduction to resistance training (*Figure* 2, Week 13), bilateral rested-state biopsies and a fasted blood sample were collected, muscle thickness of *m. vastus lateralis* and *m. rectus femoris* were measured using ultrasound (SmartUs EXT-1 M; Telemed, Vilnius, Lithuania), and body composition was measured using dual-energy X-ray absorptiometry scan (DXA; Lunar Prodigy, GE Healthcare, Chicago, IL, USA).

The training intervention consisted of 13 weeks of two weekly whole-body resistance training sessions (*Figure* 2, Week 14–27). Leg exercises were performed unilaterally to allow within-participant differentiation of resistance training load. Accordingly, for each participant, the two legs were randomly assigned to perform either three sets with 10 repetitions to exhaustion (high-load resistance exercise) or three sets with 30 repetitions to exhaustion (low-load resistance exercise); that is, each participant performed both protocols

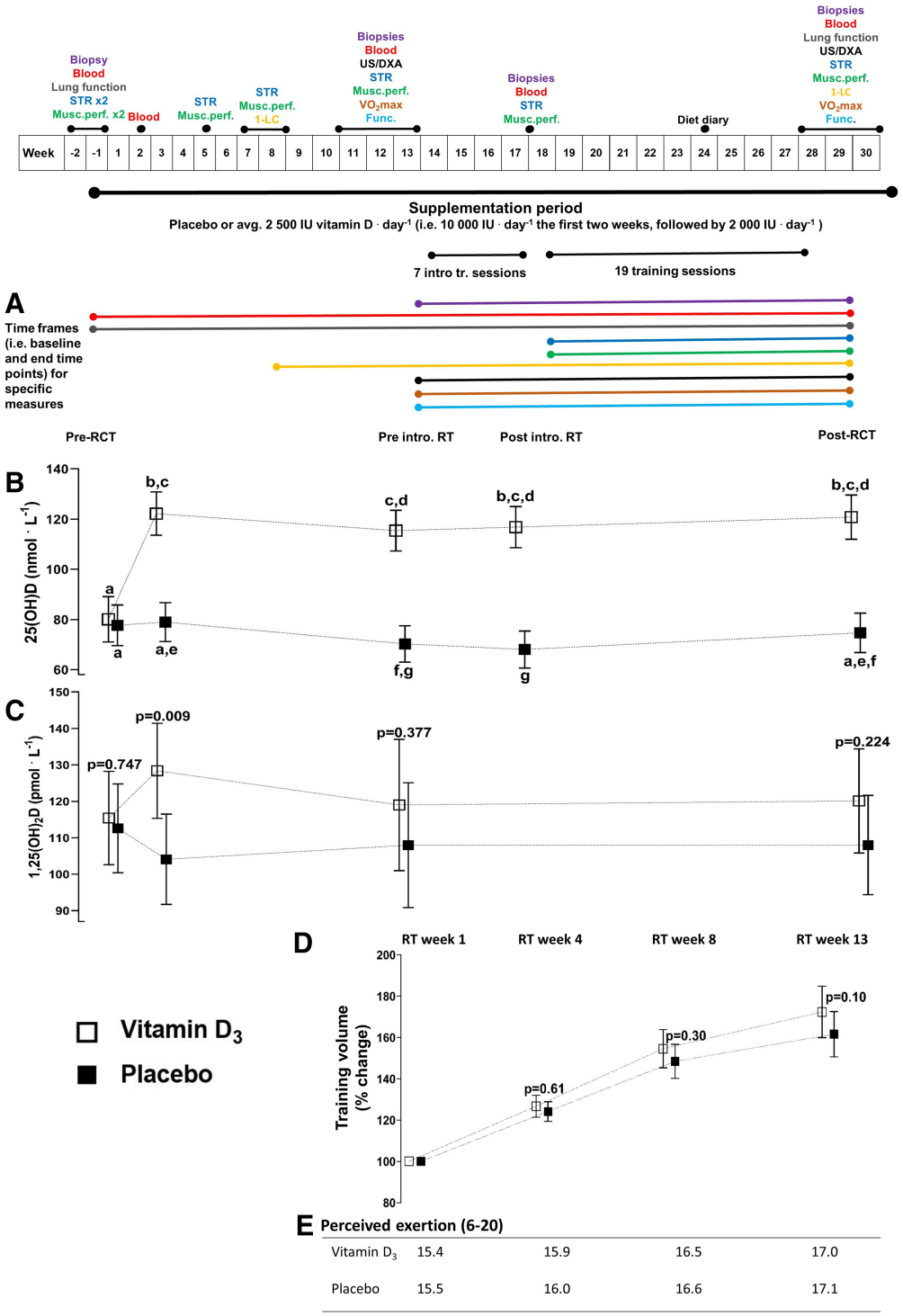


Figure 2 Schematic overview of the study protocol. Pre-defined main time frames (baseline and end time points) for specific outcome measures (the color lines represents the measurement marked with the same color at the top of the figure; *(A)*, vitamin D-status (25-hydroxyvitamin D levels, *(B)* and 1,25 dihydroxyvitamin D levels *(C)* during the RCT, training volume during the resistance training intervention *(D)*, and perceived exertion (Borg RPE, 6–20) reported after training sessions *(E)*. The training volume was calculated as average increase in volume (kg · repetitions) in leg press and knee extension from the first week of training. STR, maximal strength test; Musc.perf., test of muscular performance; 1-LC, one-legged cycling test; Func., test of functional capacity (6-min step test and 1-min sit-to-stand test); US, ultrasound measures of muscle thickness; DXA, Dual-energy X-ray Absorptiometry; $\dot{V}O_{2max}$, maximal oxygen consumption; IU, international units; RT, resistance training; 25(OH)D, 25-hydroxyvitamin D. In *(B)*, statistical differences between time points and supplementation arms are denoted by letters: different letter indicates $P < 0.05$, that is, all time point measures denoted with the same letter are statistically similar ($P > 0.05$). Data for 25(OH)D and training volume are presented as means with 95% confidence intervals.

in each session. For the upper-body, resistance exercises were performed bilaterally, consisting of two sets of 10 repetitions to exhaustion. After seven training sessions (i.e. after 3.5 weeks of training; post-introduction to resistance training), participants performed a selected battery of tests and measurements (*Figure* 2, i.e. Week 17–18), including rested-state bilateral muscle biopsies, a fasted blood sample, and measures of muscle strength, performance and torque. These tests were conducted for two reasons i) to assess the initial response to resistance training and ii) to reduce the impact of neural adaptations for training-associated increases in performance (i.e. Week 17–18 was defined as baseline for these performance measurements). After the training intervention (post-RCT), the complete battery of tests and measurements were repeated (*Figure* 2, i.e. Week 28–30). During week 24, participants conducted a dietary registration, in which they logged their dietary intake for 3 days, including one weekend day (*Table* 1). Throughout the entirety of the study, participants completed a weekly health survey every Sunday evening, which included information about supplementation compliance, self-reported health and potential discomforts caused by the nutritional supplement, such as digestive issues, sleep issues, issues with the urinary system, issues with the vestibular system, and dermal irritations. Moderate verbal motivation was given to all participants during all performance tests.

Resistance-exercise training protocol

All participants performed the same whole-body resistance-exercise training programme, consisting of the following exercises (listed in order of conductance): unilateral leg press, unilateral knee extension, unilateral knee flexion, chest press, and lat pulldown. Leg exercises were performed as three series of 10 repetitions (high-load) and 30 repetitions (low-load) to exhaustion (10RM and 30RM, respectively), and upper-body exercises were performed as two series of 10 repetitions (high-load) to exhaustion, as previously described. Exercises and sets were separated by 2 min of rest. For leg exercises, all three sets for one leg were conducted before the other leg was exercised. The order in which the two legs were exercised was switched between each session. For all exercises, training loads were adjusted from session to session, i.e. when participants managed to perform more than 12 or 35 repetitions per set for high- and low-load training, respectively. All sessions were supervised by qualified personnel to ensure correct technical execution and to ensure maximal efforts through verbal encouragement. To aid recovery and to ensure adequate protein intake after training, participants ingested half a protein bar immediately after each training session (~15 g protein; Big 100, Proteinfabrikken, Sandefjord, Norway).

Spirometry

Spirometry testing was performed using either the Oxycon Pro™ with the TripleV digital volume sensor (Carefusion GmbH, Höchberg, Germany) or the Spirare SPS320 ultrasonic spirometer (Diagnostica AS, Oslo, Norway) following guidelines from the American Thoracic Society and the European Respiratory Society.[65] Importantly, for each particular participant, all spirometry tests were performed using the same system. Participants with COPD were tested before and after inhalation of two bronchodilators (salbutamol, 0.2 mg and ipratropiumbromid, 20 μg).

Muscle strength and performance

Maximal muscle strength was assessed as one repetition maximum (1RM) in unilateral knee extension and leg press (Technogym, Cesena, Italy) and bilateral chest press (Panatta, Apiro, Italy). Each test started with specific warm-up, consisting of 10, 6, and 3 repetitions at 40%, 70%, and 85% of the anticipated maximum. Thereafter, 1RM was found by increasing the resistance progressively until the weight could not be lifted through the full range of motion. Loads were increased in intervals of 1.25, 2.5, and 1.25 kg for knee extension, leg press, and chest press, respectively. Two minutes of rest was provided between attempts. Maximal handgrip strength was measured for the dominant hand using a hand-held dynamometer (Baseline®, Fabrication Enterprises, Inc., Elmsford, NY, USA). Each test session consisted of three attempts, and the average score was used in further analyses.

Muscle performance was defined as the maximal number of repetitions achieved at 50% of pre-RCT 1RM and was assessed in unilateral knee extension and bilateral chest press. Participants were instructed to lift at a composed and controlled pace, with <1 s breaks in the lower and upper position. Whenever this requirement was not met, or participants failed to lift the weight through the full range of motion, the test was aborted.

Isokinetic unilateral knee-extension torque was assessed using a dynamometer (Humac Norm, CSMi, Stoughton, MA, USA). Participants were seated and secured with the knee joint aligned with the rotation axis of the dynamometer. Maximal isokinetic torque was tested at three angular speeds (60°, 120°, and 240° per second) with 2 min of rest provided between each of them. Prior to each test session, participants were familiarized with the test protocol by performing three submaximal efforts at each angular speed. Participants were given three attempts performed in immediate succession. The highest value was used in further analyses.

For all tests of unilateral strength and performance, the dominant leg was tested first. Seat position and general

settings for each test were noted for each participant and reproduced at each time-point.

One-legged cycling and bicycling performance

Participants conducted one-legged cycling tests (Excalibur Sport, Lode BV, Groningen, the Netherlands) to assess O_2-costs of submaximal cycling, and maximal one-legged oxygen consumption ($\dot{V}O_{2max}$) and power output (W_{max}). Each test was initiated by 2 × 5 min submaximal workloads at 30 and 40 watts (healthy), respectively, or 20 and 30 watts (COPD) with a cadence of 60 revolutions per minute. Loads were individually adjusted if the predefined workload was higher than 50% of the W_{max} achieved during the familiarization session. Thereafter, a maximal step-wise incremental protocol was conducted (10 and 5 watts/min for healthy and COPD participants, respectively). Starting loads were individually adjusted to elicit exhaustion after 6–10 min of cycling, based on results from the familiarization session. The cadence was freely chosen (>50 rpm). The test was terminated when cadence fell below 50 rpm. For all participants, submaximal and maximal performance on the dominant leg was tested first. After testing of the first leg, participants were allowed 20 min rest and/or low-intensity cycling, before testing of the other leg. During one-legged cycling tests, a 10 kg counterweight was attached to the contralateral ergometer crank to facilitate smooth cycling. The foot of the non-exercising leg was rested on a chair placed in front of the subject. Breath-to-breath measurements of pulmonary oxygen consumption and ventilation (JAEGER Oxycon PRO™; Carefusion GmbH, Höchberg, Germany) and heart rate (Polar Electro Oy, Kempele, Finland) was monitored continuously during all tests. The average oxygen consumption during the last 2 min of each submaximal workload was defined as the O_2-cost, while $\dot{V}O_{2max}$ was defined as the highest average oxygen consumption measured over a period of 30-s. Measurement of capillary lactate concentration (Biosen C-line, EKF Diagnostics, Barleben, Germany) was performed after finalization of tests.

Testing of maximal bilateral cycling $\dot{V}O_{2max}$ and W_{max} was performed on a separate day. A step-wise incremental protocol (20 and 15 watts/min for healthy men and women, respectively; 10 watts/min for participants with COPD) was conducted. Oxygen consumption was measured continuously using a computerized metabolic system with mixing chamber (JAEGER Oxycon PRO™; Carefusion GmbH, Höchberg, Germany). Prior to each cycling test, the gas analyser was calibrated using certified calibration gases with known concentrations, and the flow turbine (TripleV; JAEGER, Carefusion GmbH, Höchberg, Germany) was calibrated using the metabolic system's automatic volume calibration, or a 3 L, 5530 series calibration syringe (Hans Rudolph Inc., Kansas City, MO, USA), for one-legged and bicycling tests, respectively.

Functional performance

One-minute sit-to-stand and 6-min step tests were conducted in consecutive order on the same test day. Each test session was initiated with 10 min warm-up of low-intensity bicycling. Briefly, during the 1-min sit-to-stand tests, participants were instructed to fold their arms and sit/stand up for as many times possible during a 1-min period. The seat was 45 cm from the floor. Sit-to-stand repetitions were approved if both knees and hip joints were fully extended after each seating. Three minutes after the 1-min sit-to-stand test, the 6-min step test was conducted. Briefly, participants were instructed to perform as many steps as possible onto a 20 cm high step box with a non-slip rubber surface within 6 min (Reebok Step; Boston, MA, USA). During each step, participants were instructed to place both legs on the box, with the hip fully extended.

Muscle thickness by ultrasound and dual-energy X-ray absorptiometry-derived body mass measures

Prior to measurements of muscle thickness and DXA measurements, the participants were instructed to attend an overnight fast and avoid heavy physical activity for the last 24 h leading up to the event.

Muscle thickness of *m. vastus lateralis* and *m. rectus femoris* were measured using B-mode ultrasonography (SmartUs EXT-1 M, Telemed, Vilnius, Lithuania) with a 39 mm 12 MHz, linear array probe. Transverse images were obtained ~60% distally from the *trochanter major* towards the femoral lateral epicondyle. Three images were captured for each muscle, where the probe was relocated to the same position between each image. The position of the probe was marked on the skin and subsequently marked on a soft transparent plastic sheet superimposed on the thigh. Landmarks such as moles and scars were also marked on the plastic sheet for relocation of the scanned areas during post-training measurements. During analysis, pre and post images from the same participant were analysed consecutively using the Fiji software[66] and by two independent researchers. The average muscle thickness of the three images captured per muscle was used for further analyses.

Body composition was determined using DXA (Lunar Prodigy, GE Healthcare, Madison, WI, USA) and was analysed using the manufacturer's software, in accordance with the manufacturer's protocol. Leg lean mass was defined as the region distally of *collum femoris*. Care was taken to match the region of interest on pre and post images. Analyses of both muscle thickness and body composition were performed in a blinded manner regarding participant identity and time point of the measurement.

Blood sampling and measurements, and muscle biopsy sampling

Prior to collection of blood and muscle biopsies, participants were instructed to attend an overnight fast and to avoid heavy physical activity for the last 48 h leading up to the event. All blood samples and muscle biopsies were collected between 08:00 and 11:00 a.m. Blood samples were collected from an antecubital vein into serum-separating tubes and kept at room temperature for 30 min before centrifugation (2600 *g*, 15 min). Serum was aliquoted and stored at −80°C until further processing. Serum concentrations of total testosterone, cortisol, growth hormone, insulin-like growth-factor 1 (IGF-1), sex-hormone binding globulin (SHBG) and androstenedione were measured using an Immulite 2000 analyser with kits from the Immulite Immunoassay System menu (Siemens Medical Solutions Diagnostics, Malvern, PA, USA). Serum 25(OH)D, parathyroid hormone, calcium, albumin, creatinine, creatine kinase, aspartate aminotransferase, C-reactive protein, triglycerides, low-density lipoprotein, high-density lipoprotein, thyroid hormones and iron metabolism variables were measured using a Roche Cobas 6000 analyser and kits from Roche (Roche Diagnostics, Rotkreuz, Switzerland). In a subset of participants, $1,25(OH)_2D$ levels in serum were measured at Week −1, Week 2, Week 13 and Week 28 (vitamin D_3 arm, n = 19; placebo arm, n = 21) using enzyme immunoassays with kits from Immunodiagnostic Systems (IDS, Boldon, Tyne & Wear, UK).

Muscle biopsies were sampled from *m. vastus lateralis* under local anaesthesia (Lidocaine, 10 mg/mL, AstraZenaca AS, Oslo, Norway) using a 12-gauge needle (Universal Plus, Medax, San Possidonio, Italy) operated with a spring-loaded biopsy instrument (Bard Magnum, Bard, Covington, GA, USA), as previously described.[67] Biopsies were sampled at 1/3 of the distance from the patella to the *anterior superior iliac spine*. The tissue was quickly dissected free of blood and visible connective tissue in ice-cold sterile saline solution (0.9% NaCl). Samples for immunohistochemistry were transferred to a 4% formalin solution for fixation for 24–72 h, before further preparation. Samples for RNA analyses were blotted dry, snap-frozen in isopentane (−80°C) and stored at −80°C until further processing.

Immunohistochemistry

Formalin-fixed muscle biopsies were processed rapidly using a Shandon Excelsior ES (Thermo Fisher Scientific, Waltham, MA, USA), whereupon biopsies were paraffin-embedded and sectioned into transverse sections (4 μm). Antigen retrieval was performed at 97°C for 20 min in a target retrieval solution (cat. no. DM828, Agilent Dako, Santa Clara, CA, USA) using a PT link (PT 200, Agilent Dako, Santa Clara, CA, USA). Staining was performed using a DAKO Autostainer Link 48 (Agilent Dako, Santa Clara, CA, USA). For determination of muscle fibre types, cross-sections were first treated with protease 2 (cat. no. 760–2019, Roche Diagnostics, Rotkreuz, Switzerland), before they were triple-stained using 2.5 μg/mL BA-F8, BF-35 and 6H1 (all from Developmental Studies Hybridoma Bank, University of Iowa, Iowa City, IA, USA; BA-F8 and BF-35 deposited by Schiaffino, S., Uni. of Padova, Italy; 6H1 deposited by Lucas, C., Uni. of Sydney, Australia). Visualization of the primary antibodies was achieved by incubation of appropriate secondary antibodies, diluted 1:400: goat anti-mouse Alexa Fluor (Thermo Fisher Scientific, Waltham, MA, USA) 350 (IgGγ2b, cat. no. A21140), 488 (IgGγ1, cat. no. A21121) and 594 (IgM H + L, cat. no. A21044) for BA-F8, BF-35 and 6H1, respectively.

For determination of muscle fibre cross-sectional area (CSA) and numbers of myonuclei per muscle fibre type, a different tissue cross-section was double-stained using primary antibodies against muscle fibre membrane (dystrophin, diluted 1:100, cat. no. PA1–21011; Thermo Fisher Scientific, Waltham, MA, USA) and myosin heavy chain I (diluted 1:2000, cat. no. M8421, Sigma-Aldrich, Saint-Louis, MO, USA). Visualization was achieved using the secondary antibodies Alexa Fluor 594 (IgG H + L, diluted 1:400, cat. no. A11037) and 488 (IgG1γ1, diluted 1:400, cat. no. A21121), respectively (Thermo Fisher Scientific, Waltham, MA, USA). Muscle sections were then covered with a coverslip and glued with EverBrite™ Hardset Mounting Medium with DAPI (cat. no. 23004, Biotium Inc., Fremont, CA, USA), to visualize cell nuclei.

Images of stained cross-sections were captured using a high-resolution camera (Axiocam, Zeiss, Oberkochen, Germany) mounted on a light microscope (Axioskop-2, Zeiss, Oberkochen, Germany), with a fluorescent light source (X-Cite 120, EXFO Photonic Solutions Inc., Mississauga, Canada). Multiple images were taken using 20× objectives to capture the entirety of each cross-section. For representative images, see *Figure* 3. All analyses of muscle fibre characteristics were performed using automated procedures, ensuring unbiased quantification.

Analyses of muscle fibre type proportions were performed using the Cell Counter function in the Fiji software,[66] whereby muscle fibres were categorized as either type I, type IIA, type IIX or hybrid fibres type IIA/IIX. Sections and/or images with insufficient staining to distinguish between fibre types were excluded. Muscle fibre type-specific CSA (type I or type II) were calculated using the TEMA software (CheckVision, Hadsund, Denmark). Myonuclei were counted using the CellProfiler software.[68]

Total RNA extraction and qPCR

Approximately 10–20 mg of wet muscle tissue (average 13 ± 4 mg, range 3–26 mg) was homogenized in a total

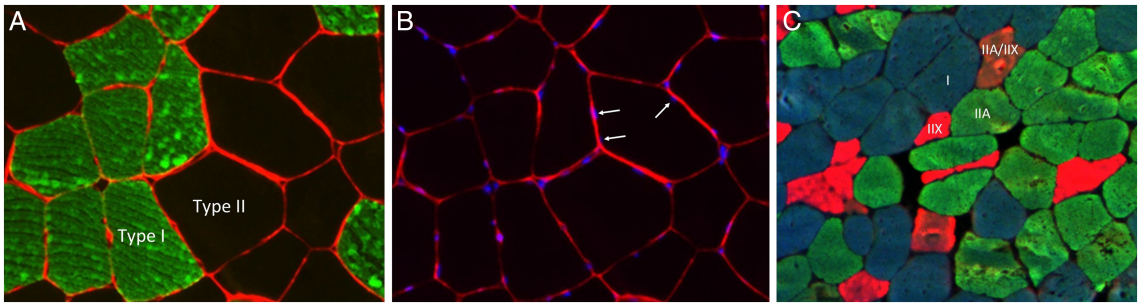


Figure 3 Representative immunohistochemistry images of *(A)* myosin heavy chain I (green) and cell membrane (red), *(B)* myonuclei (blue) and cell membrane (dystrophin, red), and *(C)* myosin heavy chain I (blue), IIA (green), IIX (red), and IIA/IIX hybrids (orange). Images in *(A)* and *(B)* are from the same tissue cross-section: triple-staining myosin heavy chain I, dystrophin and cell nuclei.

volume of 1 mL TRIzol reagent (Invitrogen, Carlsbad, CA, USA) using 0.5 mm RNase-free zirconium oxide beads and a bead homogenizer (Bullet Blender, Next Advance, Averill Park, NY, USA), as previously described.[67] To enable analysis of target gene expression per unit tissue weight, an exogenous RNA control (λ polyA External Standard Kit, Takara Bio Inc., Shiga, Japan) was added at a fixed amount (0.04 ng/mL of Trizol reagent) per extraction prior to homogenization, as previously described.[69,70] Following phase separation, 450 μL of the upper phase was transferred to a new tube and RNA was precipitated using isopropanol. The resulting RNA pellet was washed three times with 75% ethanol, eluted in 30 μL TE buffer, and diluted to 100 ng RNA/μL, following quantification of total RNA concentration using μDrop plate and the Multiskan GO microplate spectrophotometer (Thermo Fisher Scientific, Waltham, MA, USA). RNA integrity was assessed using capillary electrophoresis (Experion Automated Electrophoresis Station using RNA StdSens Assay, Bio-Rad, Hercules, CA, USA) with average integrity score (RNA quality indicator; RQI): 8.9 ± 0.8.

Five hundred nanograms of RNA were reverse transcribed using anchored oligo-dT (Thermo Fisher Scientific, Waltham, MA, USA), random hexamer primers (Thermo Fisher Scientific, Waltham, MA, USA) and Super-Script IV Reverse Transcriptase (Invitrogen, Carlsbad, CA, USA), according to manufacturers' instructions. All samples were reverse transcribed in duplicates and diluted 1:50 prior to quantitative real-time polymerase chain reaction (qPCR). qPCR reactions were conducted using a fast-cycling real-time detection system (Applied Biosystems 7500 fast Real-Time PCR Systems, Life Technologies AS), with total volumes of 10 μL, containing 2 μL cDNA (1:25 dilutions), target gene-specific primers (final concentration 0.5 μM) and a commercial master mix (2× SYBR Select Master Mix, Applied Biosystems, Life Technologies Corp., Carlsbad, CA, USA). qPCR reactions consisted of 40 cycles (3 s 95°C denaturing and 30 s 60°C annealing). Melt-curve analyses were performed for all reactions to verify single-product amplification. Gene-specific primers were designed using Primer3Plus[71] and synthesized by Thermo Scientific, except for the external RNA control, for which primers were supplied with the kit (λ polyA External Standard Kit, Takara Bio Inc., Shiga, Japan). Raw fluorescence data were exported from the platform-specific software and amplification curves were modelled using a best-fit sigmoidal model using the qpcR-package[72] written for R.[73] Threshold cycles (Ct) were estimated from the models by the second-derivate maximum method with technical duplicates modelled independently. Amplification efficiencies were estimated for every reaction.[74] For every primer pair, mean amplification efficiencies (E) were utilized to transform data to the linear scale using E^{-Ct}. Primer sequences and primer characteristics (i.e. average primer efficiencies and Ct values) are presented in Supporting Information, *Table* S1. Gene expression data were log-transformed prior to statistical analysis. As Ct values, but not primer efficiencies depend on RNA integrity,[75] RQI scores were used as a random variable on a per-target basis to control for potential degradation during statistical analyses (see below).

RNA sequencing

RNA sequencing was performed on pairwise muscle samples collected before the RCT (vitamin D_3, n = 11; placebo, n = 13), after 12 weeks of supplementation-only (vitamin D_3, n = 24; placebo, n = 29), after 3.5 weeks of introduction to resistance training (vitamin D_3, n = 23; placebo, n = 28), and after 13 weeks of resistance training (vitamin D_3 arm, n = 24; placebo arm, n = 29). Samples was selected based on quality of total RNA samples (RQI $>$ 7.0, avg 9.0 ± 0.5). Participants with complete sets of muscle biopsies were prioritized. For each muscle sample, mRNA sequencing libraries were prepared from 1000 ng of total RNA 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, San Diego, CA, USA) at the Norwegian Sequencing Centre, Oslo, Norway.

Data analyses and statistics

As defined in the pre-registration of the study protocol (ClinicalTrials.gov Identifier: NCT02598830), the effects of vitamin D_3 supplementation for different outcome measures were evaluated using different baseline time points (outlined in *Figure* 2). For transparency, statistical comparisons of all outcome measures and all relevant time points are presented in Supporting Information, *Tables* S2 and S3. These tables also specify the statistical models used for each specific variable and analysis. In general, for continuous variables, the effects of vitamin D_3 supplementation (compared with placebo) were investigated using linear mixed-effects models with the relative change from baseline being defined as the dependent variable and the supplementation arms being defined as the fixed effect. The two different training loads (high- and low-load) were added to the model as repeated measures/observations (for unilateral outcome measures), and baseline values were used as co-variates. For all participants, random intercepts were specified. For all unilateral leg variables, interaction effects were explored between the fixed effect and health status (COPD vs. non-COPD) and training loads. For other variables, interactions were investigated between the fixed effect (vitamin D_3 vs. placebo) and health status, with the exception for blood variables, for which the interaction with sex was also examined. For all statistical analyses of immunohistochemical variables (muscle fibre CSA, fibre type proportion, and myonuclei per fibre), the models were weighted for the number of counted fibres per biopsy. This was carried out to account for the reduced reliability accompanying fewer observations/fibres (see Supporting Information, *Figure* S2). For non-continuous variables, a different statistical approach was used to investigate the effects of the vitamin D_3 supplementation. For fibre type proportions (immunohistochemistry) and variables from the weekly health survey, a generalized linear mixed model (GLMM) with binomial error distribution and link function was used to examine differences in changes between supplementation arms (time*supplementation arm interactions). For gene family-based analyses of myosin heavy-chain mRNA data,[76] a GLMM with negative binomial distribution/link function (log-link) was used following transformation to transcript counts.[77] Target gene mRNA abundance, expressed as per unit muscle weight using the external reference gene, were analysed using mixed linear models with within-model normalization through the addition of random effects of technical replicates. To allow for gene-specific variances, variance functions were specified per strata (per gene). RQI scores were included in the model on a per target basis to control for RNA degradation. The number of observations per statistical analysis is presented in Supporting Information, *Table* S2. For most outcome measures, the main effect of time was examined using mixed modelling, using absolute values for the dependent variable and time points as repeated measures/observations with random intercepts for each subject (Supporting Information, *Table* S2 for complete overview).

During transcriptome analyses, gene counts were modelled using negative binomial GLMM with the total library size modelled as a fixed effect[78] together with sex and study conditions (time point and supplementation arms). The effect of resistance training on gene counts was assessed as i) the effect of time and ii) its interaction with supplementation arm (vitamin D_3 and placebo supplementation). For analyses of the effect of time, differential expression was evaluated using GLMMs containing only the time factor, combining all data irrespective of supplementation arm. For analyses of the effect of supplementation over time, differential expression was evaluated using GLMMs containing the interaction between time and supplementation arm. The supplementation-only period was modelled independently of the training period. In all models, a single random effect was used, giving each participant an individual intercept. Models were iteratively fitted using glmmTMB.[79] Model adequacy was tested for each model fit by assessing uniformity of simulated residuals.[80] A total of 15 093 genes were included in the RNA-seq data set after initial filtering, and 0.4–3.7% of these were subsequently removed due to violation of the uniformity assumption ($P < 0.05$). Genes were identified as differentially expressed when the absolute $\log_2$ fold-change was greater than 0.5 and the adjusted *P*-value (false discovery rate adjusted per model coefficient) was below 5%. Enrichment analyses of gene ontology (GO) gene sets were performed using two approaches. First, a non-parametric rank test[81,82] 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 fold-change (FC) when log (FC) > 0 and the negative inverse of the upper 95% CI when log (FC) < 0. Genes with MSD < 0 were further ranked based on *P*-values. The rank test assessed non-directional changes in gene sets. Second, gene set enrichment analysis (GSEA)[83] was performed to quantify directional regulation of the gene set. GSEA was performed using the fgsea package,[84] with $-\log_{10}$(*P*-values) *$\log_2$(fold-change) acting as the gene level metric.[85] Consensus results between the two analyses were given higher importance. GO gene sets (biological process, cellular component and molecular function), as well as Hallmark and KEGG gene sets were retrieved from the molecular signature database (version 7.1).[86] Overview of enrichment analyses with exact *P*-values are presented in Supporting Information, *Tables* S5, S6, and S8–S10.

To achieve reliable assessment of the main outcome domains muscle strength, muscle mass, one-legged endurance

performance and whole-body endurance performance, and thus to lower the risk of statistical errors, combined factors were calculated for outcome measures. For complete overview over the composition of each factor, see Supporting Information, *Table* S4. During factor calculation, each of the underlying variables were normalized to the participant with the highest value recorded during the RCT, resulting in individual scores ≤ 1. Thereafter, outcome domain factors were calculated as the mean of the normalized values for each variable for each subject (e.g. the muscle mass factor of the legs included muscle thickness, leg lean mass, and muscle fibre CSA). To evaluate the biological coherence of these factors, a factor analysis was performed to ensure correlation between the combined factors and their underlying outcome variables (Supporting Information, *Table* S4).[87] To assess the effect of vitamin D_3 supplementation for changes in these combined factors, linear mixed-effects models were used, as previously described. In addition, these factors were used to investigate the influence of pre-RCT levels of 25(OH)D, body fat proportions and body mass index on the effects of vitamin D_3 supplementation. To perform these analyses, each of the two supplementation arms were divided into quartiles, defined by baseline 25(OH)D, body fat percentage and body mass index levels, respectively (quartile 1, lowest, … quartile 4, highest). For each of the calculated factors, the effect of quartile and the interaction between quartile and supplementation arm was examined using mixed modelling.

Statistical significance was set to $P < 0.05$. In the text, data are presented as means ± standard deviation. In figures, data are shown as adjusted, estimated marginal means of relative changes and differences in relative changes between supplementation arms, with 95% confidence intervals, unless otherwise stated. Statistical analyses were performed using SPSS Statistics package version 24 (IBM, Chicago, IL, USA) and R software.[73] Figures were made using Prism Software (GraphPad 8, San Diego, CA, USA) and R software.[73]

Results and discussion

Effects of vitamin D_3 supplementation on 25(OH)D and 1,25(OH)$_2$D in blood

At pre-RCT, participants in vitamin D_3 and placebo intervention arms had similar [25(OH)D] levels in serum (80 nmol/L vs. 78 nmol/L, range: 24–144 nmol/L, *Figure* 2). [25(OH)D] levels did not differ between participants with different health status (i.e. with or without COPD diagnosis). In the vitamin D_3 arm, the study was initiated by 14 days of high-dosage vitamin D_3 intake (10 000 IU per day), which led to 42 nmol/L increases in [25(OH)D] (to 122 ± 24 nmol/L; range = 82–175 nmol/L; $P < 0.001$), with no change in the placebo arm (79 ± 31 nmol/L; range = 36–167 nmol/L) (*Figure* 2). During the remainder of the study (weeks 3–30), vitamin D_3 was ingested at 2000 IU per day, which led to stabilization of [25(OH)D] at elevated levels compared with the placebo arm (Week 13, Δ45 nmol/L; Week 17, Δ49 nmol/L; Week 29, Δ46 nmol/L; *Figure* 2), resembling the efficacy of previous studies with comparable study protocols (~2500 IU per day).[88,89] Conversely, in the placebo arm, [25(OH)D] either declined or was similar to pre-RCT levels (Week 13, $-$8 nmol/L; Week 17, $-$11 nmol/L; Week 29, $-$6 nmol/L; *Figure* 2), corroborating with changes typically seen in Northern populations during winter months,[27] with the notable observation that values were slightly higher than expected.[28]

After the initial 14 days of supplementation-only, the marked increases in 25(OH)D in the vitamin D_3 arm were accompanied by robust increases in [1,25(OH)$_2$D] compared with the placebo arm (vitamin D_3, +17 pmol/L; placebo, $-$7 pmol/L; Δ24 pmol/L, $P = 0.004$; *Figure* 2). During this time frame, change scores for [1,25(OH)$_2$D] were correlated with change scores for [25(OH)D] ($r = 0.429$, $P = 0.006$; data not shown). At Week 13 and 29, the statistical difference in changes in [1,25(OH)$_2$D] between supplementation arms had disappeared (Δ11 pmol/L, $P = 0.377$, and Δ12 pmol/L, $P = 0.224$; *Figure* 2), and the correlation between changes in [1,25(OH)$_2$D] and [25(OH)D] was no longer evident ($r = 0.169$–0.243, $P = 0.131$–0.298; data not shown). The initial period of high-dosage vitamin D_3 supplementation thus led to rapid elevations in 1,25(OH)$_2$D levels, which was subsequently reversed towards baseline levels during the follow-up period with maintenance intake (2000 IU/day), although vitamin D_3 supplementation was still associated with increased numerically values and the levels of individual variation was large. In all but three samples, measures of [1,25(OH)$_2$D] were within the normal range for adults (39–193 pmol/L), as defined by the manufacturer,[90] with all deviating samples being $>$193 pmol/L (vitamin D_3, $n = 2$; placebo, $n = 1$).

At the onset of introduction to training (Week 13) and throughout the training intervention (Week 17, Week 29), participants in the vitamin D_3 arm were all vitamin D-sufficient, as classified by the National Academy of Medicine ([25(OH)D] $>$ 50 nmol/L),[27] while in the placebo arm, 13 (Week 13), 12 (Week 17) and 5 (Week 29) participants were vitamin D-insufficient. In both supplementation arms, calcium was ingested at 500 mg/day throughout the intervention. Despite this, no changes were seen in calcium or albumin-corrected calcium levels in blood at any time point (Supporting Information, *Table* S11). Levels of the parathyroid hormone decreased throughout the intervention ($P = 0.035$; Supporting Information, *Table* S11), most likely caused by an autoregulatory response to increased calcium intake.[91] Vitamin D_3 supplementation did not alter this response. Compliance to the supplementation protocol was high in both intervention arms (vitamin D_3, 99.3%; placebo, 99.3%;

P = 0.998). Together, these observations suggest that vitamin D_3 supplementation led to improved vitamin D-status during the intervention, measured as 25(OH)D, whereas placebo led to reduced or maintained levels, with approximately 1/3rd of placebo-receiving participants showing levels associated with impaired muscle functionality (<50 nmol/L) at the onset of resistance training.[21,22,92]

Effects of vitamin D_3 supplementation on resistance training-associated changes in myofibre cross-sectional area and proportions (primary objectives)

In contrast to our main hypotheses, vitamin D_3 supplementation did not enhance resistance training-associated increases in muscle fibre cross-sectional area or changes in muscle fibre proportions (*Figure* 4; pre-defined as primary objectives of the study), despite clear improvements in vitamin D status (25(OH)D). The results are presented in more detail in later sections (*Effects of vitamin D_3 supplementation on training-associated changes in maximal muscle strength and lower-limb muscle mass* and *Effects of vitamin D_3 supplementation on training-associated changes in muscle fibre characteristics and transcriptomics)*.

Effects of 12 weeks of vitamin D_3-supplementation only (weeks 1–12) on muscle strength, performance and characteristics

The main purpose of the initial 12 weeks of vitamin D_3 supplementation-only was to ensure physiologically elevated [25(OH)D] for a prolonged period prior to onset of resistance training, thus potentially priming muscle cells for plasticity. Vitamin D_3 supplementation itself had no effect on upper- and lower-body muscle strength and performance, muscle fibre area and characteristics (*m. vastus lateralis*), or hormone concentrations in blood compared with placebo (Supporting Information, *Figure* S1 and *Table* S2), showing no interaction with health status. Surprisingly, the only exception was 1RM knee extension, for which vitamin D_3 led to negative changes compared with placebo (Δ−8.4%; P = 0.008), opposing the seemingly accepted dogma that vitamin D supplementation *per se* exerts positive effects on leg muscle strength.[35,93] Notably, for all muscle strength and muscular performance variables, the initial 12 week supplementation period was

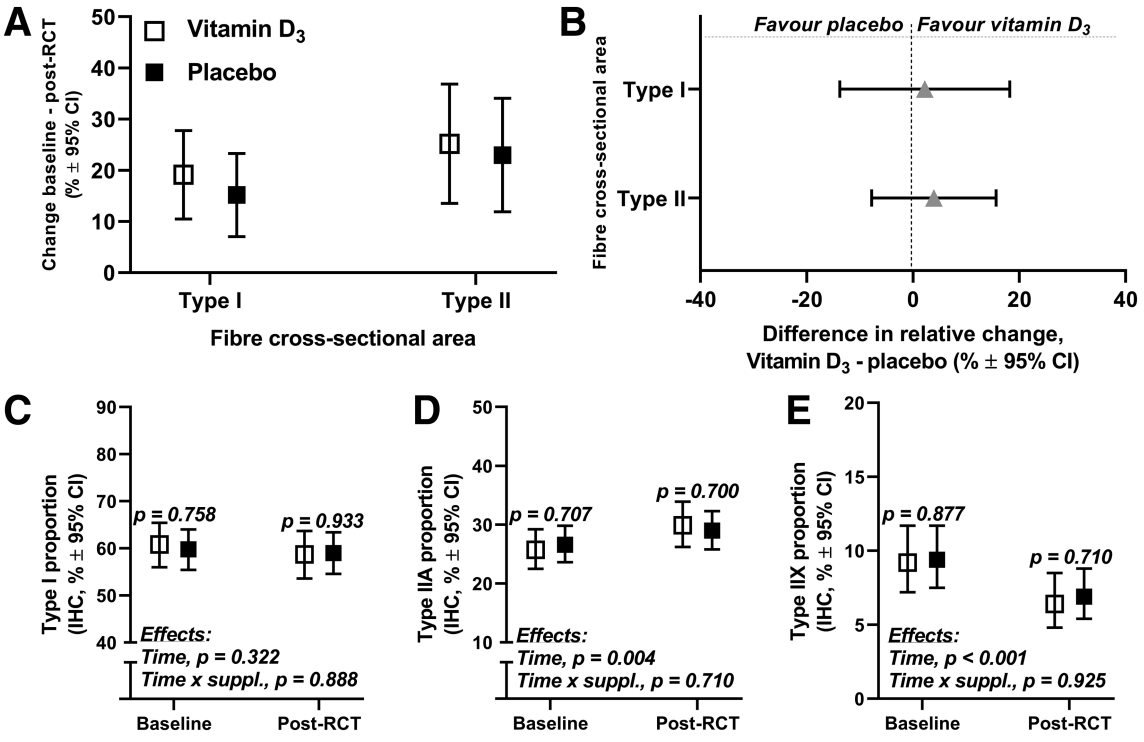


Figure 4 Primary outcome objectives of the study; effects of combined vitamin D_3 supplementation and resistance training on changes in muscle fibre cross-sectional area *(A, B)* and fibre type proportions *(C–E)* in older adults. Alpha level at $P < 0.05$. Data are presented as means with 95% confidence intervals.

associated with improved performance in all performance tests (5–71%; for details, see Supporting Information, *Figure* S1). These improvements occurred without any apparent changes in muscle cell characteristics in thigh muscle, including muscle fibre CSA (type I, 4%, $P = 0.573$; type II, 9%, $P = 0.312$), muscle fibre type proportions ($P = 0.127$–0.901), and total RNA/rRNA expression ($P = 0.604$–1.000) (Supporting Information, *Figure* S1). They were hence likely caused by technical, psychological and neural learning effects,[94] effectuated by repeated exposure to testing prior to and during the supplementation period (Supporting Information, *Figure* S1), as is typically seen in older subjects.[95] Indeed, dynamic exercises like knee extension and chest press are associated with lower intra-rater reliability than the grip strength test,[94] which remains unaffected by test–retest,[94] as was likely the case in the present study.

Overall, the 12-weeks supplementation-only period did not lead to marked changes in mRNA transcriptome profiles in the two supplementation arms combined (vitamin D_3, $n = 11$; placebo, $n = 13$). Vitamin D_3 supplementation was, however, associated with differential changes in the expression of a selected genes compared with placebo; 27 genes ↑ and 27 genes ↓ (*Figure* 5A and Supporting Information, *Table* S7). This included increased expression of B-cell lymphoma 6 and prolyl 4-hydroxylase subunit alpha-1 (*BCL6* and *P4HA1*; *Figure* 5A), both of which are known to oppose accumulation of reactive oxygen species (ROS),[96–98] and decreased expression of angiopoietin-like protein 4 (*ANGPTL4*; *Figure* 5A), which is closely correlated with levels of mitochondrial respiration.[99] These findings were reaffirmed by gene enrichment analyses, which showed a general reduction in the expression of gene sets relating to both oxidative and glycolytic metabolism in the vitamin D_3 arm (*Figure* 5B and Supporting Information, *Tables* S5–S6). This is in line with previous observations whereby vitamin D has been shown to counteract ROS and mitochondrial oxidative stress.[100] The seemingly negative effect of vitamin D_3 supplementation for expression of mitochondrial genes may thus be due to reduced mitochondrial turnover. Of note, expression of the vitamin D receptor (VDR) was observed in the data set, but was not affected by supplementation.

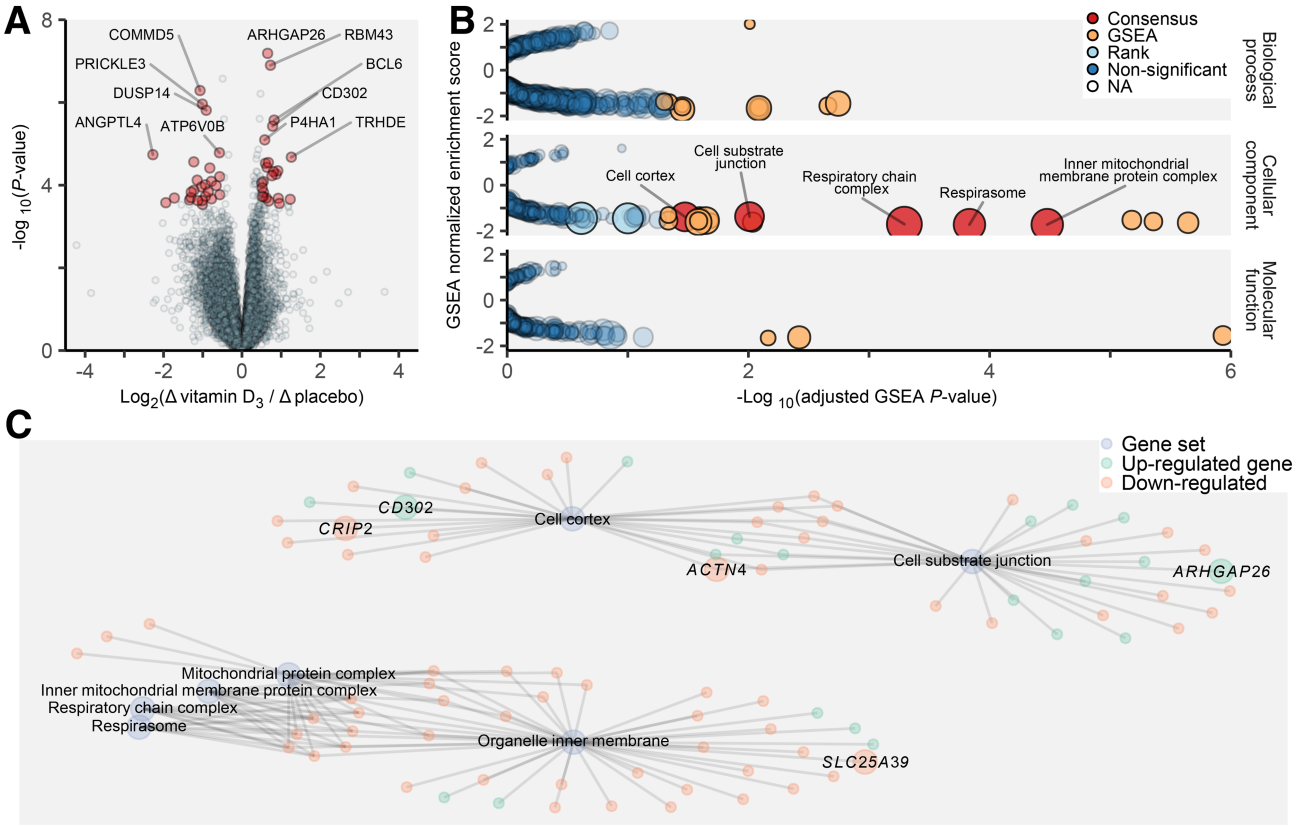


Figure 5 Effects of 12 weeks of vitamin D_3 supplementation-only on whole-genome transcriptome profiles in *m. vastus lateralis* of older adults. After 12 weeks of supplementation-only, numerous genes were differentially expressed between the vitamin D_3 and the placebo arm *(A)*; Δ, pre-introduction to resistance training/pre-RCT). Gene ontology (GO) enrichment analyses showed that these genes were primarily related to mitochondrial function and cell cortex/cell-substrate junction *(B)*; positive/negative GSEA-normalized enrichment scores indicates higher/lower expression of gene sets in the vitamin D_3 arm compared with the placebo arm). The seven differentially expressed gene sets were clustered into two distinct groups of genes *(C)*.

Introductory observations on the quality and general efficacy of the resistance training protocol (weeks 13–28)

Before assessing the effects of combined vitamin D_3 supplementation and resistance training, it is vital to reaffirm that the protocols and methods held sufficient validity and reliability, including a general assessment of the efficacy of the resistance training intervention. All training sessions were supervised by qualified personnel, as suggested by others,[46] which likely contributed to the very low drop-out rate (n = 4 during the training period, ~5%, *Table* 1), and ensured high adherence to the protocol (98%, range 81–100%, *Table* 1) and appropriate training progression throughout the intervention (*Figure* 2). Training volume (repetitions x kg) increased by 20% (knee extension) and 30% (leg press) from Week 14 (the first week of training) to Week 18 (the 4th week of training), by 48% and 54% to Week 22 (the 8th week of training) and by 65% and 68% to Week 27 (the last week of training) (*Figure* 2). This resembles or exceeds training progression seen in similar studies on previously untrained participants[101,102] and was accompanied by progressive increases in perceived exercise intensities (using the Borg RPE-scale[103]) (*Figure* 2). For these training characteristics, no differences were observed between supplementation arms (P = 0.897–0.980). The arguably successful completion of the resistance training intervention was accompanied by marked functional and biological adaptations in the participants, including increased muscle strength and performance (e.g. 22% and 72% increases in 1RM and muscular performance in knee extension, respectively, $P < 0.05$, Supporting Information, *Figure* S1), increased muscle mass (e.g. 16–24% increases in muscle fibre CSA for *m. vastus lateralis*, $P < 0.05$, Supporting Information, *Figure* S1), increases in myonuclei number per fibre (30–37%, $P < 0.05$, Supporting Information, *Figure* S1), alterations in muscle fibre proportions (e.g. type IIX fibre proportions changed from 10% to 7%, $P < 0.05$, Supporting Information, *Figure* S1), and robust alterations in muscle transcriptome profiles (499 and 312 differentially expressed genes at post-introduction resistance training and post-RCT, compared with pre-introduction to resistance training, *Figure* 11A,B). Importantly, neither of these muscle fibre characteristics changed from pre-RCT to before onset of resistance training (Week 13), suggesting that muscle biopsies sampled before and after the supplement-only period could be regarded as a sampling-resampling event (Supporting Information, *Figure* S1). For muscle strength, the intervention had relative efficiencies of 0.86% (knee extension) and 1.43% (leg press) increase per session, which resemble or exceeds expectations based on previous studies of untrained older adults (0.5–1.0% per session).[104–106]

Analytical measures to increase the validity of vitamin D_3-based analyses

To ensure valid analyses of the effects of vitamin D_3 supplementation on muscle-related features, two precautionary measures were deemed to be necessary. First, for muscle strength and muscle performance (apparatus exercises), we defined baseline levels to be equivalent to values collected after 3.5 weeks of introduction to resistance training (main analyses, *Figure* 2), rather than values collected before its onset, as noted in the preregistration of the study (NCT02598830). At this time point, initial adaptations to training were likely to have occurred, preferably non-hypertrophic effects relating to technical, psychological and neural learning effects,[94] phenomena that are particularly prominent in older subjects.[95] Using this time point as baseline arguably strengthens the association between changes in muscle strength and muscle mass, which was the main perspective of our vitamin D_3-based analyses. For other outcome measures, baseline levels were either defined as values obtained at the onset of introduction to resistance training (*Figure* 2, Week 13; muscle biological data, muscle thickness, body composition, endurance-related outcome measures) or as values obtained pre-RCT (Week −1, *Figure* 2; self-reported health, blood variables, lung function).

To further minimize the confounding effects of non-hypertrophic increases in strength and performance, all participants conducted a series of repeated tests prior to baseline tests, including five repeated 1RM and muscular performance tests in knee extension and chest press (Supporting Information, *Figure* S1a,b,e,f), respectively, four of which was conducted prior to onset of introduction to training. As expected, this led to marked and progressive increases in strength/performance levels for all test procedures compared with pre-RCT values (e.g. 4–8 – 14% for 1RM knee extension, 3–5 – 13% for 1RM bench press; the first test was conducted at ~95% of maximal effort and was thus removed from analyses) (Supporting Information, *Figure* S1). For leg press, three tests were performed prior to the defined baseline test at post-introduction to resistance training, resulting in similarly scaled improvements as observed for knee extension and chest press (Supporting Information, *Figure* S1, 14%; the first test was conducted at ~95% of maximal effort and was thus removed from analyses). These improvements occurred without any apparent hypertrophy in *m. vastus lateralis* of the dominant leg, measured as muscle fibre CSA (pre-RCT vs. pre-introduction to resistance training; type I, P = 0.573; type II, P = 0.312), as previously presented (Supporting Information, *Figure* S1g), strengthening the notion that the improvements were due to other factors. After adopting the post-introduction-to-training time point as baseline for the strength outcome measures, the efficiency of the intervention

on muscle strength was still somewhat higher than expected based on previous observations[104–106] (1RM knee extension, 0.8% per session; 1RM leg press, 1.3% per session). Notably, while these former studies contained less extensive measures to ensure reproducibility, they reported low test–retest variability, which does not concur with our results.[104–106]

Second, for analyses of the effects of vitamin D_3 supplementation on changes in muscle mass, we found it necessary to reconsider our choice of using changes in muscle fibre CSA and fibre type proportions in *m. vastus lateralis* as the primary objective of the study. These data were associated with large degrees of sampling-to-resampling variation, as evaluated using repeated muscle biopsies from the dominant leg, sampled at weeks −1 and 13, i.e. prior to introduction to resistance training (Supporting Information, *Figure* S2). Similar issues have been previously reported for such analyses,[107] although not in all studies[108,109] and are likely exacerbated in older adults, for whom larger spatial heterogeneity are present in muscle fibre characteristics compared with young adults,[110] possibly relating to the age-related remodeling of motor units.[111] Despite these issues, the data provided sufficient resolution to disclose marked increases in muscle fibre CSA and changes in muscle fibre proportions over the entirety of the training intervention, as previously presented (*Figure* 4 and Supporting Information, *Figure* S1).

In order to achieve reliable assessment of changes in muscle mass, we thus had to take on a different approach. Instead of relying on muscle fibre CSA data alone, we developed a combined muscle mass factor, in which change scores from a collection of muscle mass-related outcome measures were combined in a weighted manner (Supporting Information, *Table* S4). This factor included data on muscle fibre CSA, leg lean mass (DXA) and muscle thickness (*m. rectus* femoris, *m. vastus lateralis*; ultrasound), all of which are known to correlate.[112–114] Careful investigation of the computed muscle mass factor suggested that it increased the biological value of muscle mass-related analyses (for more information, see Supporting Information, *Table* S4). As such, it changed markedly from baseline to post-RCT (9%, $P < 0.001$, Supporting Information, *Table* S4). Following this logic, combined factors were also computed for other core outcome domains, including maximal muscle strength and one-legged and whole-body endurance performance (Supporting Information, *Table* S4).

Effects of vitamin D_3 supplementation on training-associated changes in maximal muscle strength and lower-limb muscle mass

Participants in both vitamin D_3 and placebo arms showed increases for every measure of muscle strength and mass, assessed from baseline to after finalization of the resistance training intervention: 12–25% for upper- and lower body 1RM muscle strength, 6–11% for leg muscle torque, 7–26% for muscle fibre CSA and muscle thickness and 1–3% for leg lean mass (*Figures* 6 and 7). Unsurprisingly, after combining these measures into weighted muscle strength and muscle mass factors, similarly scaled increases were observed (13% ± 8% and 9% ± 8%, respectively; *Figures* 6 and 7), which was also the case for a calculated score of relative muscle quality (Δmuscle strength factor/Δmuscle mass factor; 4% ± 10%, *Figure* 7).

Overall, vitamin D_3 supplementation did not affect these outcome measures compared with placebo in the participants, primarily evaluated as changes in muscle strength and muscle mass factors (strength, Δ2.5% (95% CI, −1.0, 6.0), $P = 0.194$; mass, Δ0.4% (95% CI, −3.5, 4.3), $P = 0.940$, *Figures* 6 and 7), and secondarily as changes in each of the underlying outcome measures (i.e. seven measures of muscle strength and three measures of muscle mass; *Figures* 6 and 7). This lack of a beneficial effect was also evident for changes in relative muscle quality (Δ1.9% (95% CI, −3.0, 6.8), $P = 0.415$; *Figure* 7). Vitamin D_3 supplementation thus had no main effect on training-associated changes in muscle functionality or gross muscle biology. While this conclusion coheres with the few comparable studies assessing the effect of combined vitamin D_3 intake and resistance training,[40,42–44] it contrasts the conclusion drawn in the only available meta-analysis on this subject, wherein vitamin D_3 supplementation was associated with augmented increases in muscle strength in older adults.[41] Notably, among the selection of ten specific outcome measures, two did not conform with the main finding. Vitamin D_3 was associated with beneficial effects for changes in 1RM knee extension (Δ6.8% (95% CI, 1.3, 12.3), $P = 0.016$; *Figure* 6) and muscle thickness of *m. rectus femoris* (Δ7.5% (95% CI, 1.8, 13.2), $P = 0.011$; *Figure* 7). For 1RM knee extension, the effect was interrelated with the negative development seen from pre-RCT to pre-introduction to training in the vitamin D_3 arm (Supporting Information, *Figure* S1). Indeed, when assessing the effect of vitamin D_3 on 1RM knee extension from pre- to post-RCT (rather than from baseline at post-introduction to training), no beneficial effect was observed compared with placebo (Δ−2% (95% CI, −12, 7), $P = 0.628$; Supporting Information, *Table* S2). As for muscle thickness in *m. rectus femoris*, we did not collect data pre-RCT and can thus not deduce if this variable followed the same pattern as 1RM knee extension. The observed benefits of vitamin D_3 supplementation for changes in *m. rectus femoris* thickness contrasts observations made for *m. vastus lateralis* thickness (Δ−0.3%, $P = 0.838$), and even oppose those made for lean mass of the legs, which tended to increase less in the vitamin D_3 arm compared with the placebo arm (Δ−1.8%, $P = 0.090$).

So far, analyses have focused on the main effect of vitamin D_3 supplementation for training-induced development of muscle strength and mass, and have thus neglected potential interactions with other independent variables such

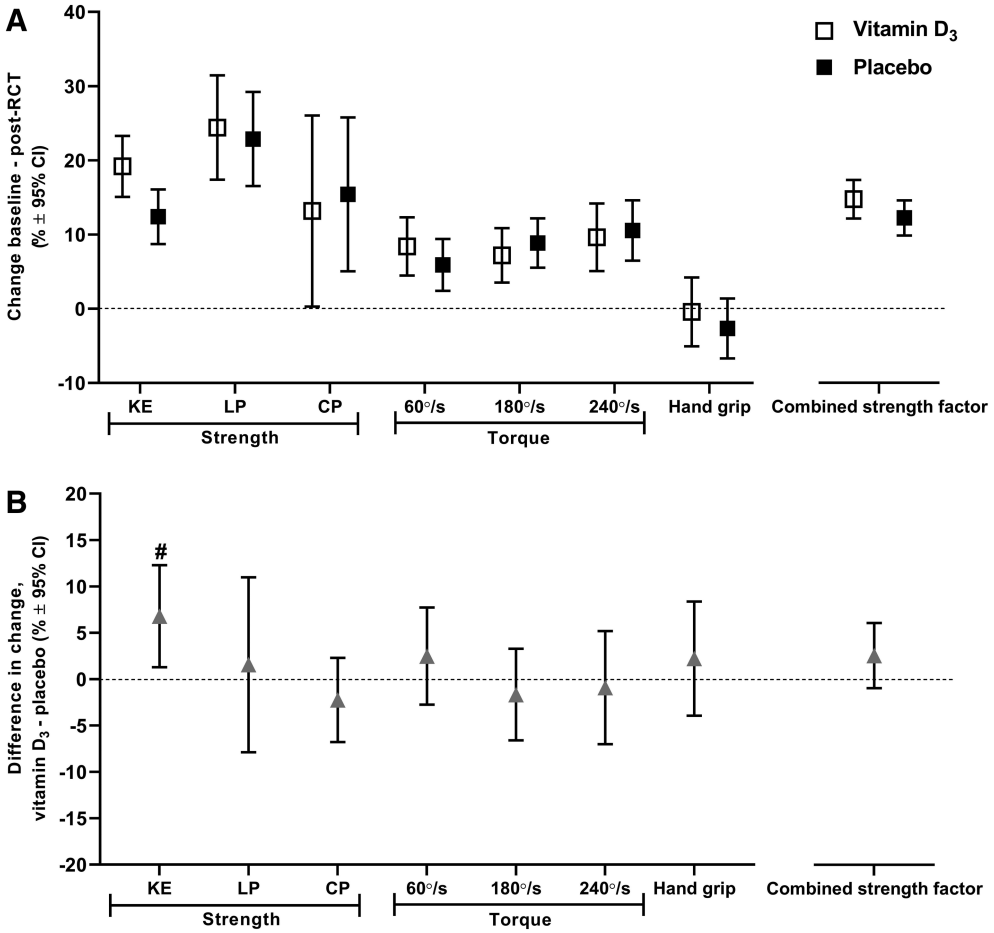


Figure 6 Effects of combined vitamin D_3 supplementation and resistance training on maximal muscle strength in older adults. Changes in muscle strength from baseline (after three weeks of introduction to resistance training) to post-RCT *(A)*, and differences in changes between vitamin D_3 and placebo arms *(B)*. KE, one-legged knee extension; LP, one-legged leg press; CP, chest press; maximal torque measured using one-legged knee extension at three velocities; 60, 180, and 240° per second; #, significant difference between vitamin D_3 and placebo arms; combined strength factor, weighted combined strength factor of unilateral strength measures (one-repetition maximum in KE and LP, and KE torque at 60, 180, and 240° per second). Alpha level at $P < 0.05$. Data are presented as means with 95% confidence intervals.

as pre-RCT levels of 25(OH)D, health status (COPD vs. non-COPD) or training modality (high-load, 10RM, vs. low-load, 30RM). The benefits of vitamin D_3 supplementation were expected to be more pronounced in participants with low baseline levels of 25(OH)D (ClinicalTrials.gov Identifier: NCT02598830). This hypothesis was based on observations made in cohort studies, wherein subjects with levels <30–50 nmol/L are more likely to show adverse muscle phenotypes.[21–23] To investigate this perspective, participants in each supplementation arm were divided into quartiles based on pre-RCT 25(OH)D levels in blood (Supporting Information, *Figure* S3). This resulted in two lower quartiles, one for the vitamin D_3 arm (vitamin $D3_{low}$, $[25(OH)D]_{mean}$ = 49.5 nmol/L, n = 8), and one for the placebo arm ($placebo_{low}$, $[25(OH)D]_{mean}$ = 47.4 nmol/L, n = 12) (Supporting Information, *Figure* S3). At the onset of introduction to resistance training, 25(OH)D levels in vitamin $D3_{low}$ had increased to 103.3 nmol/L (range 76–138), with all participants being classified as sufficient (>50 nmol/L),[17] whereas 25(OH)D levels in $placebo_{low}$ remained unchanged (45.5 nmol/L, range 22–71), with 9 out of 12 participants being classified as insufficient (<50 nmol/L). Within each of the pre-RCT 25(OH)D quartiles, the effect of vitamin D_3 and placebo supplementation on training-induced changes in muscle strength and mass (using the combined factors) were assessed. With exception of one quartile (muscle strength factor, quartile 3, P = 0.048; Supporting Information, *Figure* S3), no beneficial effects of vitamin D_3 supplementation were observed in any quartile (e.g. vitamin $D3_{low}$ vs. $placebo_{low}$, muscle strength, $\Delta -2.0\%$ (95% CI, −8.0, 3.9, P = 0.496) (Supporting Information, *Figure* S3). Instead, in vitamin $D3_{low}$, training-associated changes in muscle mass were reduced compared with $placebo_{low}$ ($\Delta -6.5\%$ (95% CI, −12.7, −0.27), P = 0.041; Supporting Information, *Figure* S3), suggesting that vitamin

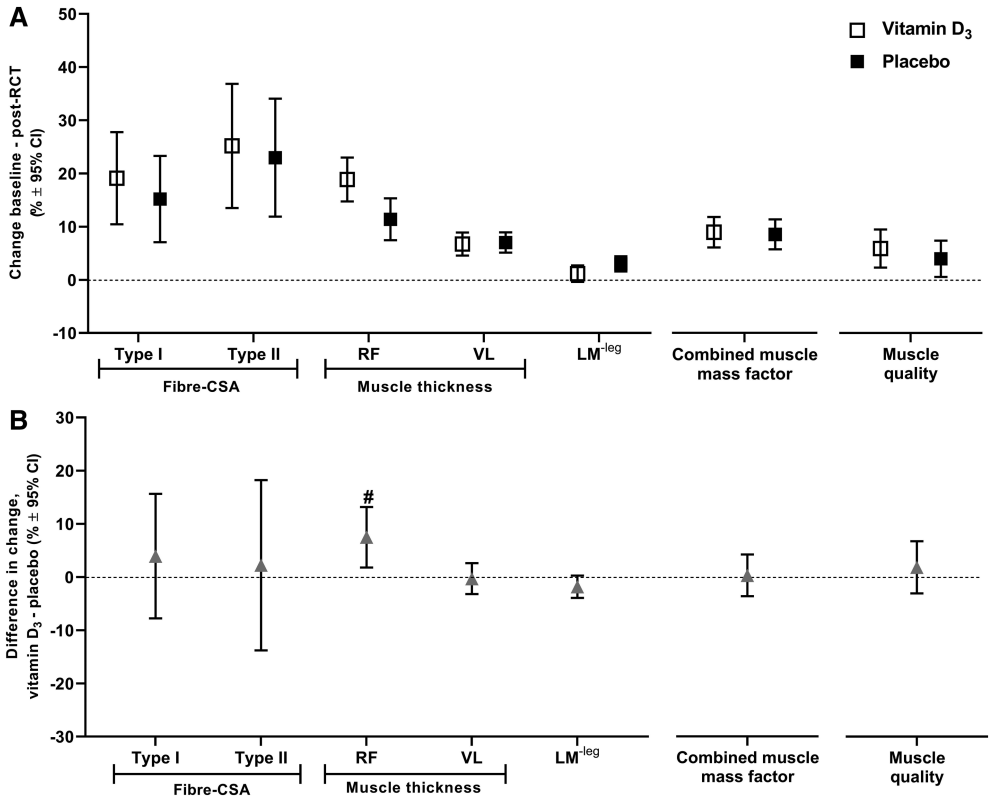


Figure 7 Effects of combined vitamin D_3 supplementation and resistance training on lower-limb muscle mass in older adults. Changes in lower-limb muscle mass from baseline (before introduction to resistance training) to post-RCT *(A)*, and differences in changes between vitamin D_3 and placebo arms *(B)*. CSA, cross-sectional area (also presented in Figure 4); RF, *m. rectusfemoris*; VL, *m. vastus lateralis*; LM per leg, leg lean mass per leg; #, significant difference between vitamin D_3 and placebo arms; combined muscle mass factor, weighted combined muscle mass factor including fibre cross-sectional area (type I and type II), muscle thickness (RF and VL) and LM per leg; muscle quality, muscle strength factor/muscle mass factor. Alpha level at $P < 0.05$. Data are presented as means with 95% confidence intervals.

D_3 supplementation may even have compromised training adaptations in subjects with low pre-RCT 25(OH)D levels. Adding to this, participants in the entire spectre of quartiles responded quite similarly to resistance training, irrespective of supplementation arms, evident as no interaction between 25(OH)D quartiles/supplementation arm and changes in muscle strength ($P = 0.237$) or muscle mass ($P = 0.159$). Arguably, the statistical power of these analyses were not sufficiently high to conclude on this perspective.

The impact of vitamin D_3 supplementation for training-associated changes in muscle strength and muscle mass factors did not interact with health status (COPD vs. non-COPD) or training modality (10RM vs. 30RM) (Supporting Information, *Table* S2). However, it should be noted that for selected specific outcome measures, interactions were found with both of these independent variables (summarized in Supporting Information, *Table* S2), including an interaction between changes in type II-fibre CSA and COPD/non-COPD, and between changes in 1RM knee extension/vastus lateralis thickness and 10RM/30RM. In addition to these interaction analyses, we also investigated the potential relation between the effects of vitamin D_3 supplementation and baseline body fat proportions, as overweight and obese have been shown to have decreased bioavailability of vitamin D due to deposition of 25(OH)D in body fat compartments (while concomitantly showing attenuated anabolic response to resistance exercise[115]).[116] To this end, we performed quartile-based analyses, as previously described. These analyses did not reveal an effect of baseline body fat proportions for changes in [25(OH)D] (fat percentage, $P = 0.432$; BMI, $P = 0.369$) or muscle mass factor (fat percentage, $P = 0.355$; BMI, $P = 0.293$) (Supporting Information, *Figure* S4). However, it did have an effect on changes in the muscle strength factor (fat percentage, $P = 0.016$; BMI, $P = 0.706$), that is, in $\text{quartile}_{\text{high fat percentage}}$, vitamin D_3 supplementation was associated with larger increases in muscle strength compared with placebo (fat percentage, Δ 5.8% (95% CI, 0.5, 11.0), $P = 0.032$; BMI, Δ7.8% (95% CI, 2.5, 13.1), $P = 0.005$; Supporting Information, *Figure* S4 and *Table* S2), suggesting beneficial effects of vitamin D_3 supplementations in subjects with high proportions of body fat, opposing our initial expectations.

Effects of vitamin D_3 supplementation on training-associated changes in one-legged and whole-body endurance performance

Participants in both vitamin D_3 and placebo arms showed improvements in one-legged and whole-body endurance performance over the course of the resistance training intervention: 42–74% increases in one-legged muscular performance (*Figure* 8), 7–9% increases in peak power output (W_{max}) in one- and two-legged cycling (*Figure* 8), 3–5% reductions in O_2 costs of submaximal one-legged cycling (Supporting Information, *Table* S2), and 6–10% increases in functional performance (1-min sit-to-stand test and 6-min step test, *Figure* 8). In accordance with this, marked increases were observed in weighted one-legged and whole-body endurance performance factors (one-legged, vitamin D_3 25% ± 19%, placebo 22% ± 11%; whole-body, vitamin D_3 9% ± 8%, placebo 7% ± 6%; *Figure* 8). These effects cohere well with previously observed benefits of resistance training for endurance variables in older adults.[117–119]

Vitamin D_3 supplementation had no effect for any of these outcome measures compared with placebo, neither for weighted endurance performance factors (one-legged, Δ2% (95% CI, −5, 10), P = 0.773; two-legged, Δ2% (95% CI, −2, 6), P = 0.636; *Figure* 8), nor for any of the specific outcome measures (*Figure* 8). For combined endurance factors, there was no interaction between baseline 25(OH)D quartiles and effects of vitamin D_3 supplementation (one-legged, P = 0.950; whole-body, P = 0.266; Supporting Information, *Figure* S3 and *Table* S2), nor was there any interactions with health status (one-legged, P = 0.747, whole-body, P = 0.129, Supporting Information, *Table* S2) or training modality (one-legged, P = 0.719, Supporting Information, *Table* S2).

Effects of vitamin D_3 supplementation on training-associated changes in muscle fibre characteristics and transcriptomics

Participants in both vitamin D_3 and placebo arms showed marked changes in muscle fibre characteristics over the course of the training intervention. These included decreased type IIX muscle fibre proportions from 10% to 7% (*Figure* 9), increased type IIA proportions from 26% to 29% (*Figure* 9),

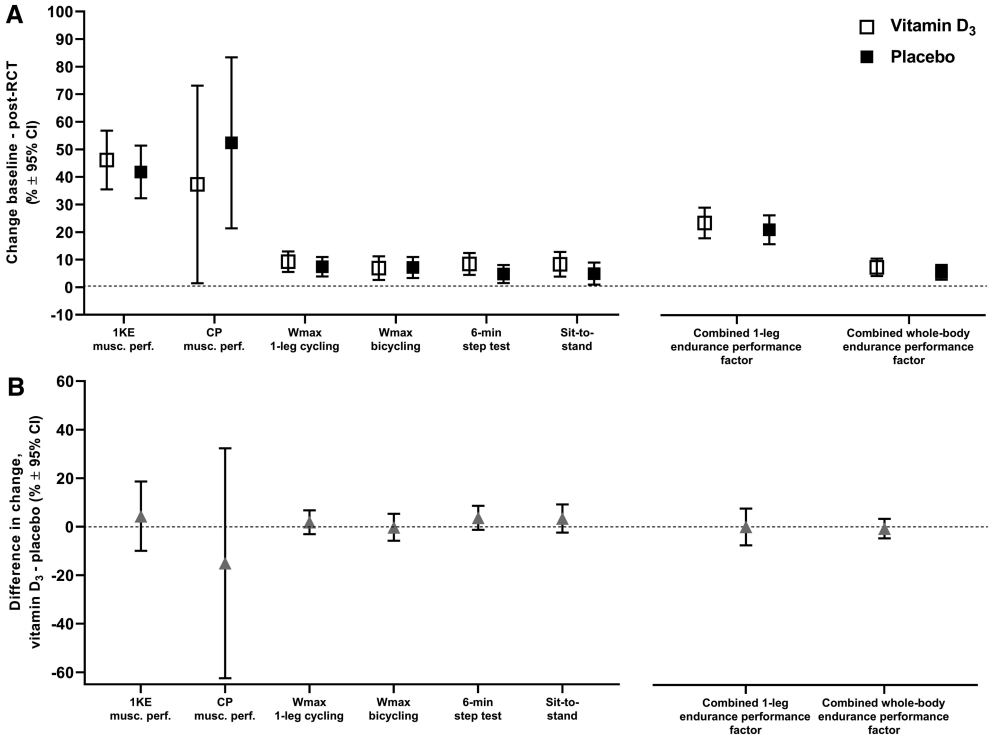


Figure 8 Effects of combined vitamin D_3 supplementation and resistance training on one-legged and whole-body endurance performance in older adults. Changes in endurance performance from baseline (before introduction to resistance training) to post-RCT *(A)*, and differences in changes between vitamin D_3 and placebo arms *(B)*. 1KE, repetitions to failure in one-legged knee extension (50% of pre-intervention 1RM); CP, repetitions to failure in chest press (50% of pre-intervention 1RM); W_{max}, maximal power output; 6-min step test, maximal number of steps achieved during 6 min; Sit-to-stand, maximal number of sit-to-stands achieved during 1 min; combined 1-leg endurance performance factor, weighted combined one-legged endurance factor including 1KE muscular performance and one-legged cycling W_{max}; weighted combined whole-body endurance factor including W_{max} bicycling, 6-min step test and sit-to-stand test. Alpha level at $P < 0.05$. Data are presented as means with 95% confidence intervals.

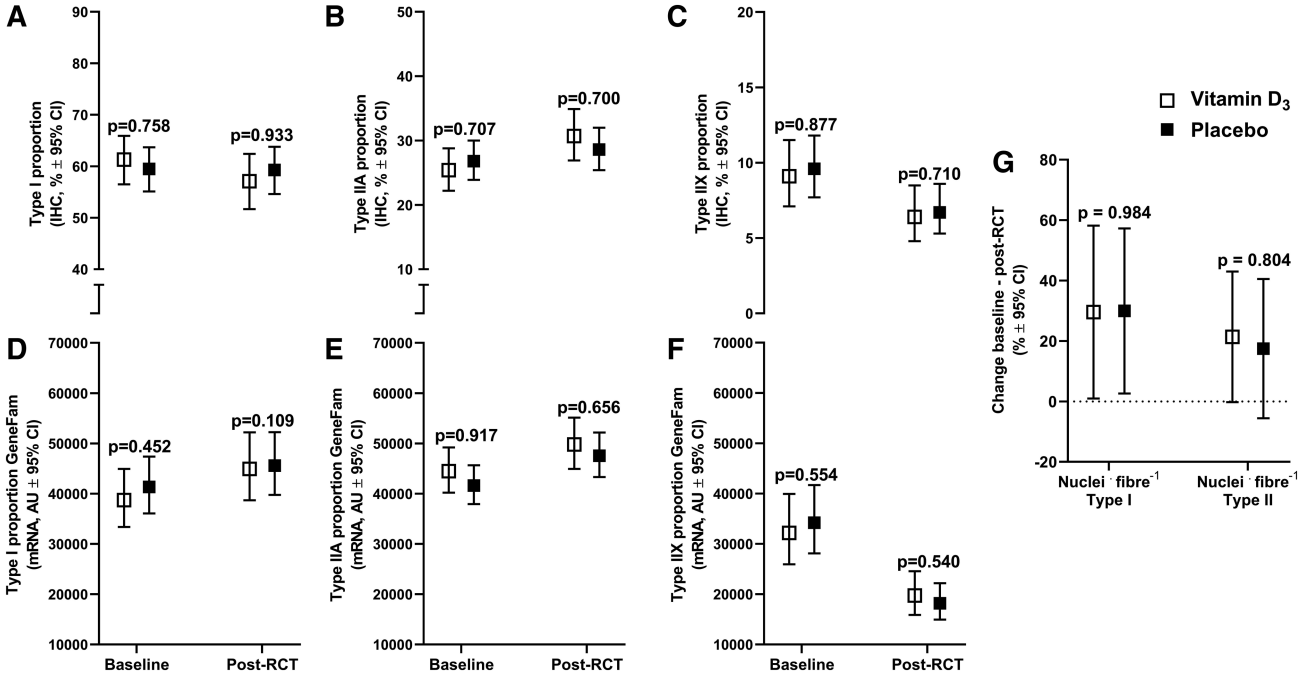


Figure 9 Effects of combined vitamin D_3 supplementation and resistance training on muscle fibre type proportions and myonuclei per fibre in *m. vastus lateralis* of older adults. Muscle fibre type proportions *(A–F)* at baseline (before introduction to resistance training) and post-RCT measured using immunohistochemistry *(A–C)* and qPCR (gene family profiling (GeneFam)-normalized myosin heavy chain mRNA expression, *(D–F)*, and changes in myonuclei count per type I and type II fibre from baseline to post-RCT *(G)*. Significant changes were observed for fibre type IIA and IIX using both methods (significant increase and decrease, respectively; $P < 0.05$). For fibre type I, an increased expression was present using qPCR ($P < 0.05$), but no change was observed for immunohistochemistry ($P = 0.322$). *P*-values denotes the statistical difference between the supplementation arms. RT, resistance training. Data are presented as means with 95% confidence intervals.

increased type IIA/IIX hybrid fibres abundances from 2.6% to 3.2% (Supporting Information, *Table* S2), and 25–48% increases in myonuclei number per muscle fibre (*Figure* 9). Changes in IIX and IIA proportions were verified using qPCR, showing decreased levels of type IIX mRNA abundance and increased levels of type IIA (*Figure* 9), calculated using the gene family-profiling approach.[76] These analyses also revealed increased proportions of type I mRNA after the training intervention (*Figure* 9), potentially caused by increased type I protein turnover. The observed changes in muscle fibre-type characteristics corroborate well with previous studies in older adults,[120–122] although increased numbers of myonuclei per muscle fibre are not consistently reported.[123] Vitamin D_3 supplementation had no effect on training-associated changes in muscle fibre proportions or myonuclei content compared with placebo (*Figure* 9).

The training intervention resulted in 1.14- to 1.16-fold increases in total RNA per unit muscle tissue weight (*Figure* 10), a proxy marker for ribosomal RNA content that has previously been associated with training-induced changes in muscle growth and strength.[67,124] Similar increases were found for the mature ribosomal species 18 s (1.18-fold) and 28 s (1.16-fold), in addition to the 45 s pre-ribosomal rRNA (1.19-fold) using qPCR (*Figure* 10). No changes were observed for 5.8 s (1.07-fold, $P = 0.722$) or 5 s (1.06, $P = 0.940$) following the entire training intervention. Notably, for analyses of total RNA and ribosomal RNA, an additional time point were included in main analyses, i.e. in muscle biopsies sampled after introduction to training (3.5 weeks, 7 sessions), as early increases in total RNA seem to associate with long-term chronic responses to training, making it a potential hallmark of muscle plasticity.[67] As expected, 3.5 weeks of training led to marked increases in total RNA (1.10- to 1.21-fold) and expression of all ribosomal RNA species (1.13- to 1.27-fold) (*Figure* 10). Whereas these changes corroborates quite well with changes observed in healthy, young subjects,[67] although with a notable reduction in the relative increase, they contradict previous observations of no resistance training-associated increases in total RNA per unit muscle tissue weight in older subjects.[125] Vitamin D_3 supplementation had no effect on training-associated changes in total RNA or rRNA expression compared with placebo.

The training intervention led to marked changes in muscle mRNA transcriptome profiles in the two supplementation arms combined, with 499 genes being differentially expressed (DE) after 3.5 weeks of resistance training

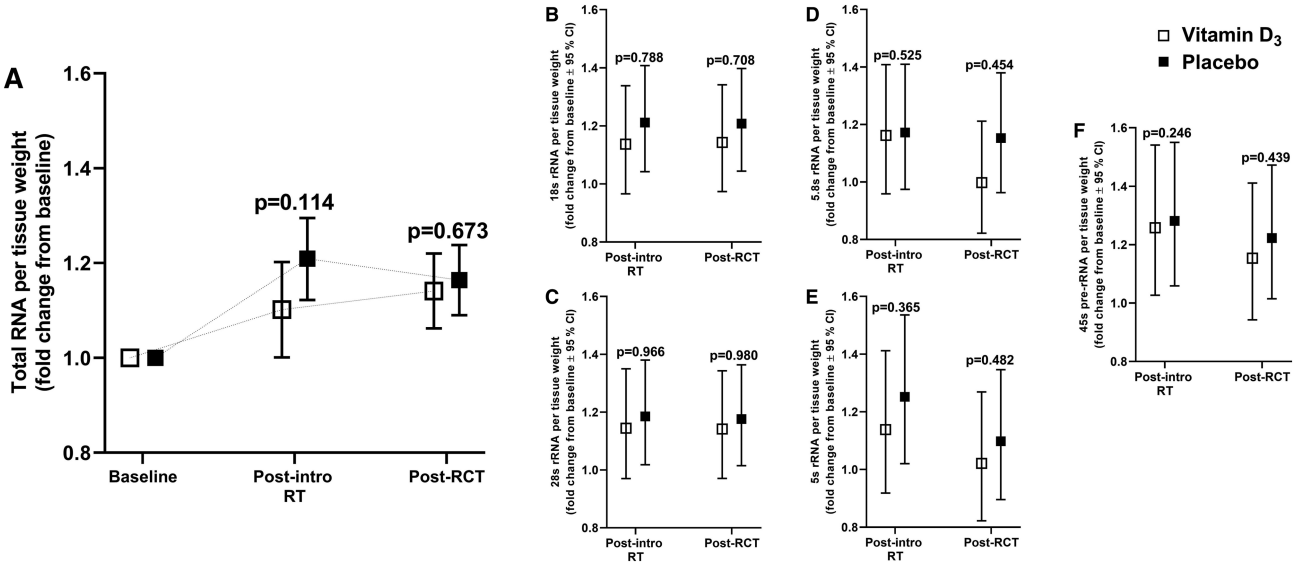


Figure 10 Effects of combined vitamin D_3 supplementation and resistance training on total RNA abundances and rRNA expression in *m. vastus lateralis* of older adults. Total RNA *(A)*, 18 s rRNA *(B)*, 28 s rRNA *(C)*, 5.8 s rRNA *(D)*, 5 s rRNA *(E)*, and 45 s pre-rRNA *(F)* abundances at baseline (before introduction to resistance training) and post-RCT. Significant increases from baseline–post-introduction to resistance training were present for all variables ($P < 0.05$). From baseline–post-RCT significant increases were present for all variables ($P < 0.05$), with the exception of 5.8 s rRNA ($P = 0.722$) and 5 s rRNA ($P = 0.940$). RT, resistance training. *P*-values denotes the statistical difference between the supplementation arms. Alpha level at $P < 0.05$. Data are presented relative to amounts of tissue weight. Data are presented as means with 95% confidence intervals.

(post-intro RT; 436 genes ↑, 63 genes ↓, *Figure* 11A) and 312 genes being DE after 13 weeks of resistance training (post-RCT; 255 genes ↑, 57 genes ↓) (*Figure* 11A,B). VDR was expressed, but unaffected by combined vitamin D_3 supplementation and resistance training, contradicting previous observations of a positive association between supplementation-induced improvements in 25(OH)D status and leukocyte,[126] myoblast/myotube[127] and skeletal muscle[128] VDR expression. GO enrichment analyses revealed increased expression of gene sets associated with extracellular matrix, blood vessel morphogenesis and leukocyte migration at both 3.5 and 13 weeks (*Figure* 11C, Supporting Information, *Table* S8), as well as increased expression of the inflammatory response gene set at 3.5 weeks (Supporting Information, *Table* S8). Conversely, decreased expression was observed for gene sets involved in ribosomal functions at both 3.5 and 13 weeks (*Figure* 11C). This could be interpreted as contradicting the likely important role of *de novo* ribosomal biogenesis for training-associated muscular adaptations.[67,124] Notably, as these analyses were performed using traditional library size-based normalization, which basically provided target gene expression relative to the expression of all other genes.[129] In an alternative set of transcriptome analyses, which rather included normalization that corrected for muscle sample weight and thus provided gene expression analyses per sample size (tissue-offset normalization),[129] the negative effects of resistance training on ribosomal gene expression was not evident (data not shown). This was the only major difference between library size and tissue-offset normalization in the present study setting.

Vitamin D_3 supplementation had no effect on training-associated changes in gene expression, neither at 3.5 weeks (*Figure* 11D) nor at 13 weeks (*Figure* 11E), suggesting that no single gene was differentially affected by combined vitamin D_3 supplementation and resistance training and resistance training-only. In contrast to this, enrichment analyses showed traces of vitamin D_3-sensitive changes in expression at both 3.5 and 13 weeks of resistance training (*Figure* 11F and Supporting Information, *Tables* S9–S10). After 3.5 weeks of training, there was differential expression of gene sets involved in cell junctions, blood vessel morphogenesis and muscle cell differentiation. These initial responses to resistance training should be interpreted with caution, as they were only evident in one of the two analyses (GSEA or rank-based analyses; *Figure* 11F and Supporting Information, *Tables* S9–S10). After 13 weeks of resistance training, the vitamin D_3 arm showed increased expression of gene sets involved in endothelial proliferation and blood vessel morphogenesis compared with placebo (*Figure* 11F). This agrees with the previously observed positive relationship between 25(OH)D-status and endothelial function, potentially interacting through the endothelium-derived vasodilator, nitric oxide.[100] Indeed, this coheres well with a recent study, which showed favorable effects of combined vitamin D_3 supplementation and resistance

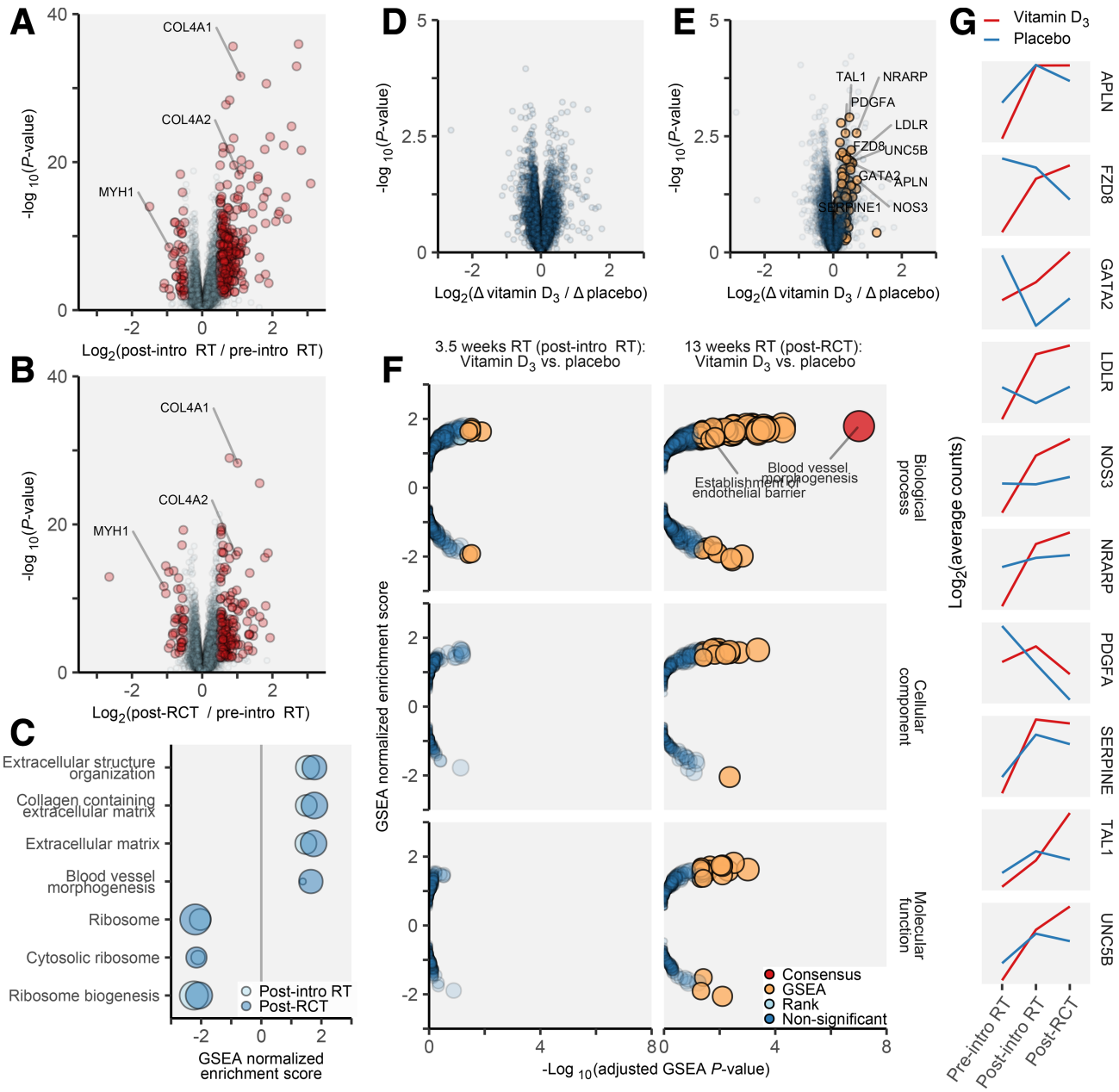


Figure 11 Effects of 3.5/13 weeks of resistance training-only *(A–C)* and 3.5/13 weeks of combined vitamin D_3 supplementation and resistance training *(D–G)* on mRNA transcriptome profiles in *m. vastus lateralis* of older adults. Resistance training-only led to robust changes in gene expression at both 3.5 weeks (*A*; post-intro resistance training – pre-intro resistance training) and 13 weeks (*B*; post-RCT – pre-intro resistance training), including increased expression of collagen type IV α1 and α2 genes (*COL4A1* and *COL4A2*, respectively) and decreased expression of the myosin heavy chain IIX gene (*MYH1*). The three most enriched gene sets with increased and decreased expression, in addition to the 'blood vessel morphogenesis' gene set are shown in *C* (light blue, 3.5 weeks; dark blue, 13 weeks; according to the GSEA enrichment score). Combined vitamin D_3 supplementation and resistance training did not lead to differential changes in expression for a singular gene compared with placebo at neither 3.5 weeks (*D*; Δ, post-introduction to resistance training - pre-introduction to resistance training) nor 13 weeks of resistance training (*E*; Δ, post-RCT - pre-introduction to resistance training; orange dots/genes denotes leading edge genes from the 'blood vessel morphogenesis' GO gene set, that is, the most highly enriched gene set between supplementation arms after 13 weeks of resistance training). GO enrichment analyses of differentially regulated gene sets between the vitamin D_3 and the placebo arms following 3.5 weeks (left panel, *F*) and 13 weeks of resistance training (right panel, *F*; positive/negative GSEA-normalized enrichment scores indicates higher/lower expression of gene sets in the vitamin D_3 arm compared with the placebo arm). (*G*) Timeline for the 10 most affected genes between vitamin D_3 and placebo arms belonging to the 'blood vessel morphogenesis' GO gene set. RT, resistance training; Consensus, when both the non-directional rank-based enrichment test and the directional gene-set enrichment analysis (GSEA) turned out significant. In *Figure* 11*C,F*, circle sizes of gene sets are relative to *P*-values, i.e. larger circles indicate lower *P*-values (see Supporting Information, *Tables S5–S10* for exact *P*-values).

training on flow-mediated dilation of blood vessels and blood pressure in postmenopausal women.[130] Unfortunately, endothelial function was not assessed in the current study.

Effects of vitamin D_3 on hormones in blood and health-related outcome measures

In general, the intervention was associated with beneficial changes for several health-related variables, including reduced levels of lipids (triglycerides and low-density lipoprotein/LDL), reduced levels of fat mass (total and visceral fat) and improved self-reported health (Supporting Information, *Table* S11). Conversely, a small but undesirable decrease was observed in lung capacity, measured as forced ventilatory capacity (FVC) (Supporting Information, *Table* S2). The intervention was not associated with changes in whole-body bone mineral density or changes in serum levels of hormones, except for decreased levels of parathyroid hormone (Supporting Information, *Table* S11), as previously presented. For most of the health variables, there was no effect of vitamin D_3 supplementation (Supporting Information, *Tables* S2 and S11), with exception of cortisol levels in blood, which increased more in the vitamin D_3 arm (*Table* S11), and lung function measured as FEV_1/FVC-ratios, which declined in subjects with COPD in the vitamin D_3 arm (Supporting Information, *Table* S2).

Sarcopenia

The intervention proved effective for treating age-related loss in muscle mass, leading to 1.4% increases in total lean body mass ($P < 0.001$) (Supporting Information, *Table* S11). This reduced the number of participants that could be defined as sarcopenic from 16% (11 subjects) to 12% (8 subjects), with sarcopenia being defined as appendicular lean mass $(kg)/m^2$ greater than two standard deviations below the sex-specific means of young adults.[61] Speculatively, the increase in total lean mass was supported by increased levels of serum creatinine in both supplementation arms (+6%; Supporting Information, *Table* S11). Although serum creatinine is generally used for evaluation of renal function,[131] creatinine production and levels also increases with increases in total muscle mass.[131,132]

Steroid hormones

Vitamin D_3 supplementation did not affect levels of anabolic steroid hormones such as testosterone. This was in discordance with our initial hypothesis, as we presumed a positive association between vitamin D levels (measured as 25(OH)D) and testosterone levels, based on previous observations from vitamin D_3 supplementation studies[52] and cohort studies.[133] Despite this, our finding is in line with several other vitamin D supplementation studies, which has reported no effect on testosterone in blood.[134,135] Conversely, vitamin D_3 supplementation seemed to affect serum cortisol levels compared with placebo (Δ48 nmol/L, $P = 0.038$; Supporting Information, *Table* S11), although no main effect of time was observed (i.e. the observed increase in the vitamin D_3 arm was not statistically significant, $P = 0.374$) and there was no statistical difference between supplementation arms at the end of the intervention ($P = 0.053$).

Lung function

The small -1.95% reduction in FVC seen after the 28 week long RCT ($P = 0.006$; Supporting Information, *Table* S2) was surprising, as exercise is generally accepted to be beneficial for lung functionality, including resistance training.[136,137] Notably, other measures of lung function, such as forced ventilatory volume in one second (FEV_1 and predicted FEV_1) and FEV_1/FVC, were not affected by the intervention (Supporting Information, *Table* S2).

The negative effects of vitamin D_3 on lung function, measured as FEV_1/FVC ($\Delta -2.9\%$ points, $P = 0.012$; Supporting Information, *Table* S2), were also surprising. This effect showed a clear interaction with health status, and as such was only evident in COPD patients in the vitamin D_3 arm, which showed $\Delta -8.4\%$ reductions compared with placebo (Supporting Information, *Table* S2). This subgroup analysis was however clearly weakened by the small sample size (COPD, $n = 9$ vs. $n = 11$, vitamin D_3 vs. placebo). The negative effect of vitamin D_3 on FEV_1/FVC did not interact with pre-RCT levels of FEV_1/FVC, but surprisingly, in another subgroup-analysis, the pre-RCT 25(OH)D vitamin $D3_{low}$ quartile was associated with larger decrement in FEV_1/FVC than $placebo_{low}$ ($\Delta -5.4\%$ points, $P = 0.009$; data not shown). This observation is difficult to explain, as it indirectly opposes the notion that vitamin D deficiency leads to impaired lung functions.[138] More research is clearly needed to elucidate on the consequences of resistance training and vitamin D_3 supplementation for lung functionality.

Adverse effects of the intervention

Overall, neither vitamin D_3 supplementation nor resistance training was associated with adverse effects or events during the intervention, with potential exception of certain aspects of lung function, as previously discussed, and iron biology (see Supporting Information, *Table* S11).

Primarily, a health survey was administered to the participants on a weekly basis. This included rating of 11 potential discomforts relating to digestion problems, sleep problems, issues with the urinary system, issues with the vestibular system and dermal irritations (Supporting Information, *Table* S2). No effect of vitamin D_3 supplementation was found for any of these variables. In the health survey, participants were also asked to rate their experienced health on a point-scale from 0–10. This self-reported conception of health improved from 6.3 ± 1.6 to 7.1 ± 1.6 ($P < 0.001$, Supporting Information, *Table* S2), with no difference between

supplementation arms (P = 0.433, Supporting Information, *Table* S11).

The intervention was not associated with training-associated injuries, with only five participants (6%) reporting discomforts with training towards the end of the intervention and only four participants (5%) withdrawing from study during the resistance training intervention, neither of which were due to injuries associated with the training. As such, serum levels of markers of muscle tissue damage (creatine kinase and aspartate aminotransferase) even decreased during the intervention, with no effects of vitamin D_3 supplementation (Supporting Information, *Table* S11). Supervised resistance training can safely be advocated for both healthy older adults and persons with COPD.

Concluding remarks

The study was conducted as a double-blinded RCT, addressing the effects of 12 weeks of vitamin D_3 supplementation only (i.e. two weeks of 10 000 IU/day, followed by ten weeks of 2000 IU/day), and 13 weeks of combined vitamin D_3 (2000 IU/day) and resistance training on functional measures, health markers and muscle biology in a mixed population of older adults. Vitamin D_3 supplementation is often hailed as an ergogenic aid for optimizing the outcome of resistance training, and is recommended for a variety of human populations, ranging from healthy subjects to athletes and chronically diseased subjects.[7,20] Vitamin D is thus presumed to play an important role in training-associated muscle plasticity. Despite this, its importance for humans remains largely elusive, with current knowledge stemming predominantly from animal research,[55] and the few existing human studies providing limited, uncertain and contradicting results.[41–44] Indeed, the present data do not support a role for vitamin D in training-associated muscle plasticity and functionality, at least not in older adults (with and without moderate COPD) with suboptimal to adequate baseline levels of 25(OH)D. More precisely, vitamin D_3 supplementation had no effect on core outcome domains such as changes in muscle strength, muscle mass, endurance performance and general muscle cell characteristics, and its effects on the muscle transcriptome was largely limited to gene sets relating to endothelial and cardiovascular functions. The validity of this insight is fortified by the thorough methodological and analytical approach. This included accounting for previous methodological issues such as a lack of a pre-training supplementation period, low vitamin D dosages, and neglecting to standardize test/training routines such as supervision of training sessions, test–retest analyses of functional and biological outcome measures, familiarization to training and a low reproducibility of singular outcome measures. The analytical approach also accounted for the potential confounding effects of the heterogeneity of the study population, as no interaction was found between effects of vitamin D_3 supplementation and disease status (healthy vs. COPD), or differences in pre-RCT vitamin D status, as all $[25(OH)D]_{baseline}$ quartiles responded in similar manners.

Despite our substantial efforts to strengthen the ecological value of the data set, there are aspects of vitamin D biology that remain unresolved, and that may have affected the conclusions and outcomes of the study. First, in skeletal muscle, adequate vitamin D signaling may occur at 25(OH)D levels lower than the defined cutoff (insufficient, $<$50 nmol/L).[27] Speculatively, all participants in the placebo arm may thus have been vitamin D-sufficient at the onset of resistance training, leaving our quartile-based analyses with limited biological value. Indeed, studies have suggested that vitamin D insufficiency will affect human muscle in an adverse manner only at concentrations $<$30 nmol/L.[139] Second, although serum 25(OH)D level is widely regarded as an adequate measure of vitamin D status,[63] it may be a poor proxy marker for vitamin D biology, as it largely fails to reflect $1,25(OH)_2D$ levels, the metabolically active form of vitamin D.[140] In line with this, in the present study, $[25(OH)D]$ was not correlated with $[1,25(OH)_2D]$ at baseline (data not shown) and was not increased by long-term vitamin D_3 supplementation (at weeks 13 and 29). Such decoupling of 25(OH)D and $1,25(OH)_2D$ levels have several potential explanations. These include feedback-mediated regulation of vitamin D biology, which is largely affected by PTH levels,[141] as well as impaired $25(OH)D \rightarrow 1,25(OH)_2D$ conversion in individuals with pathophysiological indications such as renal dysfunction.[142] The latter is unlikely to explain the lack of increases in $[1,25(OH)_2D]$ in the present study, as only two participants were indicated with renal dysfunction (estimated based on levels of creatinine in serum; *Table* 1). Rather, the initial two weeks of high-dosage vitamin D_3 supplementation did lead to marked increases in $[1,25(OH)_2D]$, emphasizing that supplementation is indeed capable of increasing levels of metabolically active vitamin D, at least at high doses and within a short time frame. At weeks 13 and 29 were the PTH levels suppressed for both supplementation arms compared with pre-RCT levels. This was possibly related to the calcium supplement, and may have contributed to the unaltered $1,25(OH)_2D$ levels at these time points. Third, muscle cells may themselves possess the apparatus to convert 25(OH)D into $1,25(OH)_2D$, as they express the 25-Hydroxyvitamin D 1-alpha-hydroxylase (*CYP27B1*) protein. Indeed, in *in vitro* experiments on murine myoblast and myotubes, 25(OH)D and $1,25(OH)_2D$ treatment seem to lead to similar increases in the expression of vitamin D markers such as VDR, suggesting that peripheral regulation of vitamin D biology is a biological opportunity.[127] Fourth, while 25(OH)D was assessed as $[25(OH)D]_{total}$ in the present study, levels of unbound 25(OH)D (i.e. not bound to vitamin D binding protein or albumin; ~0.03%) may represent a more accurate measure of vitamin D status in a clinical setting.[143] Indeed,

in mice lacking vitamin D binding protein, and therefore displaying very low $[25(OH)D]_{total}$ (~8 nmol/L), no signs of vitamin D deficiency are seen unless they are put on a vitamin D deficient diet.[144] Fifth, in the present study, the resistance training intervention lasted for only 13 weeks. Speculatively, this may have been too short for vitamin D_3 supplementation to manifest its potential benefits for muscle plasticity, despite the presence of a 12-week lead-in supplementation period. Arguably, however, if vitamin D status and signaling is indeed important of muscle biological adaptations to training, even shorter interventions should lead to detectable changes in muscle biology, such as its transcriptome. This was not observed, neither in general, nor for specific vitamin D-responsive genes such as VDR.[128] Sixth, the study protocol was unavoidably associated with large interindividual variation in responses. This variation may have been related to vitamin D_3 supplementation *per se*, resistance training *per se* or to a combination of both, and may have affected groupwise comparisons. More research is clearly needed to elucidate on these perspectives.

Despite these uncertainties, it seems clear that vitamin D_3 supplementation did not affect muscle biological characteristics in the present study, particularly those measured using RNA-seq. Indeed, in our transcriptome analyses, not a single gene was found to be vitamin D_3-sensitive after a period of resistance training, which is surprising given the accepted dogma that vitamin D primarily acts as a transcriptional regulator,[55] and that the VDR was rather highly expressed in the data set, although it did not change with vitamin D_3 supplementation. Moreover, gene sets that were identified as vitamin D_3-sensitive in gene enrichment analyses were largely associated with vascular function rather than muscle cell biology.

Despite the general lack of effects of vitamin D_3 supplementation on muscle mass and phenotype (primary objectives of the study), as well as the lack of effects on other muscle functional and biological traits, the data set contained a couple of interesting observations. First, in the muscle transcriptome data, the effects of vitamin D_3 supplementation *per se* on expression of mitochondrial genes and the effects of combined vitamin D_3 supplementation and resistance training on biomarkers of endothelial and vascular biology calls for further study. Arguably, these biological features would be more decisive for adaptations to endurance-like training, posing the intriguing possibility that vitamin D_3 supplementation may be beneficial for the outcome of such training. Second, in participants with high baseline fat proportions/high BMI, vitamin D_3 supplementation led to increased training-associated changes in muscle strength. In these participants, the bioavailability of vitamin D may have been compromised by the high fat content (in the placebo arm, although they did not exhibit lowered 25(OH)D levels), corroborating with previous observation of interactions between vitamin D biology and fat mass.[116] While this may indicate that vitamin D exerts direct effects on muscle biology, as muscle strength is predominately defined by muscle mass,[145] this still seems unlikely as no such vitamin D_3-effect was seen for other muscle-specific outcome measures (e.g. muscle mass and phenotype). The causality may thus involve other physiological adaptations such as motoneuron function,[146] which has indeed been suggested to be affected by vitamin D supplementation in rodents.[147]

In retrospect, the pre-identified primary objectives of the current study were not ideal (i.e. the effects of vitamin D_3 supplementation on muscle fibre CSA and proportions). The underlying rationale behind this choice was to investigate the effects of vitamin D_3 supplementation on a set of unbiased biological variables (not prone to test–retest fluctuations), adhering to the existing notion that vitamin D may affect muscle fibre size and fibre type proportions (e.g. elucidated in the review from Ceglia, 2009[148]). This clearly underestimating the reliability issues associated with histological measures, which were indeed evident in the data set (Supporting Information, *Figure* S2). Importantly, vitamin D_3 supplementation was not associated with beneficial effects for any of the investigated primary or secondary outcome measures, hence leaving the overall conclusion as unambiguous.

In conclusion, in older adults with or without COPD, vitamin D_3 supplementation efficiently improved vitamin D-status without any adverse effects, but did not lead to beneficial effects in resistance training-associated changes in muscle function or characteristics. This rejects the notion that vitamin D_3 supplementation is necessary to obtain adequate muscular responses to resistance training in the general older population. Secondary analyses revealed positive effects of vitamin D_3 supplementation for participants with high proportions of fat mass and for gene sets involved in vascular functions, advocating further research to elucidate on these specific biological characteristics. Finally, the training programme was well-tolerated and associated with pronounced effects for a variety of health variables, emphasizing the potency of resistance training for relieving sarcopenia and maintaining functional capacity in older adults with and without COPD.

Acknowledgements

The authors would like to express their gratitude to students Katarina Alsvik, Karianne Pedersen, Nicolai Blindheim Jacobsen, Jan Christian Fjelltun, Håvard Berge Bredesen, Karoline Michalsen, Håkon Fure Jerpseth, Karin Heggeli Moen, Krister Flobergseter, Mads-Henrik Hafsmo, Sondre Bjerke, Joachim Skjeseth Fjeller, Vilde Bakkhaug, Vetle Olsen, Hilde Juvik, Hallvard Underdal, Anniken Braaten, Anette

Gårderløkken, Malene Wilhelmsen, Vemund Lien, Malene Grahl-Madsen, Pauline Forren, Anders Kristoffersen, Jørn Klepp Thorgersen, Berit Hauge Aakvik, Ole-Jørgen Folkestad, Simen Bratberg Ramstad, Lasse Løwstrøm Aulin, Simen Næss Berge, Marius Midtmageli Bekkemellem, Martine Pedersen, Kristian Lian, Even Hovland Rosenlund, Henrik Eckhoff, Synne Skogstad, Marte Johannson, Simen Longva Hedlund, Jon Sindre Aas, Marte Fosvold Løtveit, Stine Studsrød, and Marius Fagerås for invaluable assistance during intervention follow-up and data sampling. The authors are grateful for the contribution from Tore J. Rødølen (Granheim Lung Hospital, Innlandet Hospital Trust) and Anne Sofie Lofthus, and the technical support from Bente Malerbakken (Lillehammer Hospital for Rheumatic Diseases) and Randi Christiansen. We also acknowledge Pharma Nord ApS for a good cooperation and sponsoring the vitamin D_3 and placebo supplements in the study. Further, the authors wish to thank Proteinfabrikken A/S for sponsoring the project with protein chocolate bars and Johanne Haugen, University of Bergen, for taking care of the third-party randomization. Last but not least, thanks to all participants for your cooperation. Without your effort and dedication, this study had not been possible. The authors of this manuscript certify that they comply with the ethical guidelines for authorship and publishing in the *Journal of Cachexia, Sarcopenia and Muscle*.[149]

Online supplementary material

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Table S1. qPCR primer sequences and performance.
Table S2. Statistical summary table.
Table S3. Statistical summary table, qPCR data.
Table S4. Computed factors for main outcome domains.
Table S5. Gene ontology analyses, effects of vitamin D_3 supplementation.
Table S6. KEGG and Hallmark analyses, effects of vitamin D_3 supplementation.
Table S7. Differentially expressed genes, effects of vitamin D_3 supplementation.
Table S8. Gene ontology analyses, effects of resistance training.
Table S9. Gene ontology analyses, effects of combined vitamin D_3 supplementation and resistance training.
Table S10. KEGG and Hallmark analyses, effects of combined vitamin D_3 supplementation and resistance training.
Table S11. Blood and health variables.
Figure S1. General efficacy of the RCT.
Figure S2. Sample-resample reliability measures of immunohistochemical assessments.
Figure S3. Baseline vitamin D-status and the interaction with the study's main outcomes.
Figure S4. Baseline body fat proportions/body mass index and the interaction with the study's main outcomes.

Funding

The study was funded by Inland Norway University of Applied Sciences, Innlandet Hospital Trust (grant number 150339, S. E.) and Regional Research Fund Inland Norway (grant number 298419, S. E.).

Conflict of interest

None declared. Pharma Nord ApS procured supplements but was not in any way involved in data collection, analyses or interpretations.

References

1. Ethgen O, Beaudart C, Buckinx F, Bruyère O, Reginster JY. The future prevalence of sarcopenia in Europe: a claim for public health action. *Calcif Tissue Int* 2017;**100**:229–234.
2. Beaudart C, Zaaria M, Pasleau F, Reginster JY, Bruyère O. Health outcomes of sarcopenia: a systematic review and meta-analysis. *PLoS One* 2017;**12**:1–16.
3. Evans WJ, Wayne A. Symposium: aging and body composition: technological advances and physiological interrelationships sarcopenia and age-related changes in body composition and functional capacity. *J Nutr* 1993;**123**:465–468.
4. Grimby G, Saltin B. The ageing muscle. *Clin Physiol* 1983;**3**:209–218.
5. Sobestiansky S, Michaelsson K, Cederholm T. Sarcopenia prevalence and associations with mortality and hospitalisation by various sarcopenia definitions in 85-89 year old community-dwelling men: a report from the ULSAM study. *BMC Geriatr* 2019;**19**:1–13.
6. Cawthon PM, Fox KM, Gandra SR, Delmonico MJ, Chiou CF, Anthony MS, et al. Do muscle mass, muscle density, strength, and physical function similarly influence risk of hospitalization in older adults? *J Am Geriatr Soc* 2009;**57**:1411–1419.
7. Robinson SM, Reginster JY, Rizzoli R, Shaw SC, Kanis JA, Bautmans I, et al. Does nutrition play a role in the prevention and management of sarcopenia? *Clin Nutr* 2018;**37**:1121–1132.
8. Dalle S, Rossmeislova L, Koppo K. The role of inflammation in age-related sarcopenia. *Front Physiol* 2017;**8**: (DEC).
9. Hunter GR, McCarthy JP, Bamman MM. Effects of resistance training on older adults. *Sports Med* 2004;**34**:329–348.
10. Suetta C, Magnusson SP, Beyer N, Kjaer M. Effect of strength training on muscle function in elderly hospitalized patients: review. *Scand J Med Sci Sports* 2007;**17**:464–472.
11. Nunes JP, Pina F, Ribeiro A, Cunha P, Kassiano W, Costa D. Responsiveness to muscle mass gain following 12 and

24 weeks of resistance training in older women. *Aging Clin Exp Res* 2020;(April). https://doi.org/10.1007/s40520-020-01587-z
12. Chmelo EA, Crotts CI, Newman JC, Brinkley TE, Lyles MF, Leng X, et al. Heterogeneity of physical function responses to exercise training in older adults. *J Am Geriatr Soc* 2015;**63**: 462–469.
13. Thomaes T, Thomis M, Onkelinx S, Goetschalckx K, Fagard R, Lambrechts D, et al. Genetic predisposition scores associate with muscular strength, size, and trainability. *Med Sci Sports Exerc* 2013;**45**:1451–1459.
14. Rea IM. Towards ageing well: use it or lose it: exercise, epigenetics and cognition. *Biogerontology* 2017;**18**: 679–691.
15. Close G, Hamilton L, Philp A, Burke L, Morton J. New strategies in sport nutrition to increase exercise performance. *Free Radic Biol Med* 2016;1–15.
16. Koopman R. Dietary protein and exercise training in ageing. *Proc Nutr Soc* 2011;**70**:104–113.
17. Kraemer WJ, Ratamess NA. Hormonal responses and adaptations to resistance exercise and training. *Sports Med* 2005;**35**:339–361.
18. Dahlquist DT, Dieter BP, Koehle MS. Plausible ergogenic effects of vitamin D on athletic performance and recovery. *J Int Soc Sports Nutr* 2015;**12**:33.
19. Calle MC, Fernandez ML. Effects of resistance training on the inflammatory response. *Nutr Res Pract* 2010;**4**:259–269.
20. Owens DJ, Allison R, Close GL. Vitamin D and the athlete: current perspectives and new challenges. *Sports Med* 2018;**48**:3–16.
21. Houston DK, Cesari M, Ferrucci L, et al. Association between vitamin D status and physical performance: the InCHIANTI study. *J Gerontol A Biol Sci Med Sci* 2007;**62**:440–446.
22. Mastaglia SR, Seijo M, Muzio D, Somoza J, Nuñez M, Oliveri B. Effect of vitamin D nutritional status on muscle function and strength in healthy women aged over sixty-five years. *J Nutr Health Aging* 2011;**15**:349–354.
23. Tieland M, Brouwer-Brolsma EM, Nienaber-Rousseau C, van Loon LJC, De Groot LCPGM. Low vitamin D status is associated with reduced muscle mass and impaired physical performance in frail elderly people. *Eur J Clin Nutr* 2013;**67**: 1050–1055.
24. Dawson-Hughes B. Vitamin D and muscle function. *J Steroid Biochem Mol Biol* 2017;**173**:313–316.
25. Remelli F, Vitali A, Zurlo A, Volpato S. Vitamin D deficiency and sarcopenia in older persons. *Nutrients* 2019;**11**:1–14.
26. Klingberg E, Oleröd G, Konar J, Petzold M, Hammarsten O. Seasonal variations in serum 25-hydroxy vitamin D levels in a Swedish cohort. *Endocrine* 2015;**49**: 800–808.
27. *Dietary Reference Intakes for Calcium and Vitamin D*. Washington, D.C: National Academies Press; 2011.
28. Huotari A, Herzig KH. Vitamin D and living in northern latitudes—an endemic risk area for vitamin D deficiency. *Int J Circumpolar Health* 2008;**67**:164–178.
29. Webb AR, Kline L, Holick MF. Influence of season and latitude on the cutaneous synthesis of vitamin D3: exposure to winter sunlight in Boston and Edmonton will not promote vitamin D3 synthesis in human skin. *J Clin Endocrinol Metab* 1988;**67**:373–378.
30. Chiang C-M, Ismaeel A, Griffis RB, Weems S. Effects of vitamin D supplementation on muscle strength in athletes: a systematic review. *J Strength Cond Res* 2017;**31**:566–574.
31. Tomlinson PB, Joseph C, Angioi M. Effects of vitamin D supplementation on upper and lower body muscle strength levels in healthy individuals. A systematic review with meta-analysis. *J Sci Med Sport* 2015;**18**:575–580.
32. Muir SW, Montero-Odasso M. Effect of vitamin D supplementation on muscle strength, gait and balance in older adults: A systematic review and meta-analysis. *J Am Geriatr Soc* 2011;**59**:2291–2300.
33. Bischoff-Ferrari HA, Dawson-Hughes B, Willett WC, et al. Effect of vitamin D on falls. *JAMA* 2004;**291**:1999–2006.
34. Kalyani RR, Stein B, Valiyil R, Manno R, Maynard JW, Crews DC. Vitamin D treatment for the prevention of falls in older adults: systematic review and meta-analysis. *J Am Geriatr Soc* 2010;**58**:1299–1310.
35. Beaudart C, Buckinx F, Rabenda V, Gillain S, Cavalier E, Slomian J, et al. The effects of vitamin D on skeletal muscle strength, muscle mass, and muscle power: a systematic review and meta-analysis of randomized controlled trials. *J Clin Endocrinol Metab* 2014;**99**:4336–4345.
36. Knutsen KV, Madar AA, Lagerløv P, Brekke M, Raastad T, Stene LC, et al. Does vitamin D improve muscle strength in adults? A randomized, double-blind, placebo-controlled trial among ethnic minorities in Norway. *J Clin Endocrinol Metab* 2014;**99**:194–202.
37. Stockton KA, Mengersen K, Paratz JD, Kandiah D, Bennell KL. Effect of vitamin D supplementation on muscle strength: a systematic review and meta-analysis. *Osteoporos Int* 2011;**22**:859–871.
38. Kenny AM, Biskup B, Robbins B, Marcella G, Burleson JA. Effects of vitamin D supplementation on strength, physical function, and health perception in older, community-dwelling men. *J Am Geriatr Soc* 2003;**51**:1762–1767.
39. Rosendahl-Riise H, Spielau U, Ranhoff AH, Gudbrandsen OA, Dierkes J. Vitamin D supplementation and its influence on muscle strength and mobility in community-dwelling older persons: a systematic review and meta-analysis. *J Hum Nutr Diet* 2017;**30**:3–15.
40. Antoniak AE, Greig CA. The effect of combined resistance exercise training and vitamin D3 supplementation on musculoskeletal health and function in older adults: a systematic review and meta-analysis. *BMJ Open* 2017;**7**:1–16.
41. Uusi-Rasi K, Patil R, Karinkanta S, et al. Exercise and vitamin din fall prevention among older women a randomized clinical trial. *JAMA Intern Med* 2015;**175**: 703–711.
42. Carrillo AE, Flynn MG, Pinkston C, Markofski MM, Jiang Y, Donkin SS, et al. Impact of vitamin D supplementation during a resistance training intervention on body composition, muscle function, and glucose tolerance in overweight and obese adults. *Clin Nutr* 2013;**32**:375–381.
43. Bunout D, Barrera G, Leiva L, Gattas V, de la Maza MP, Avendaño M, et al. Effects of vitamin D supplementation and exercise training on physical performance in Chilean vitamin D deficient elderly subjects. *Exp Gerontol* 2006;**41**:746–752.
44. Agergaard J, Trøstrup J, Uth J, Iversen JV, Boesen A, Andersen JL, et al. Does vitamin-D intake during resistance training improve the skeletal muscle hypertrophic and strength response in young and elderly men? - a randomized controlled trial. *Nutr Metab* 2015;**12**:1–14.
45. American College of Sports Medicine. American College of Sports Medicine position stand. Progression models in resistance training for healthy adults. *Med Sci Sports Exerc* 2009;**41**:687–708.
46. Mazzetti SA, Kraemer WJ, Volek JS, Duncan ND, Ratamess NA, Gómez AL, et al. The influence of direct supervision of resistance training on strength performance. *Med Sci Sports Exerc* 2000;**32**: 1175–1184.
47. Zeitelhofer M, Adzemovic MZ, Gomez-Cabrero D, Bergman P, Hochmeister S, N'diaye M, et al. Functional genomics analysis of vitamin D effects on CD4+ T cells in vivo in experimental autoimmune encephalomyelitis. *Proc Natl Acad Sci* 2017;**114**:E1678–E1687.
48. Huang Z, Zhang Y, Li H, Zhou Y, Zhang Q, Chen R, et al. Vitamin D promotes the cisplatin sensitivity of oral squamous cell carcinoma by inhibiting LCN2-modulated NF-κB pathway activation through RPS3. *Cell Death Dis* 2019;**10**.
49. Książek A, Zagrodna A, Słowińska-Lisowska M. Vitamin D, skeletal muscle function and athletic performance in athletes—a narrative review. *Nutrients* 2019;**11**:1–12.
50. Buitrago CG, Arango NS, Boland RL. 1α,25 (OH) 2D 3-dependent modulation of Akt in proliferating and differentiating C2C12 skeletal muscle cells. *J Cell Biochem* 2012;**113**:1170–1181.
51. Zhang X, Zanello LP. Vitamin D receptor-dependent 1α,25(OH)2 vitamin D 3-induced anti-apoptotic PI3K/AKT signaling in osteoblasts. *J Bone Miner Res* 2008;**23**:1238–1248.
52. Pilz S, Frisch S, Koertke H, Kuhn J, Dreier J, Obermayer-Pietsch B, et al. Effect of vitamin D supplementation on testosterone

levels in men. *Horm Metab Res* 2011;**43**:223–225.
53. Todd JJ, Pourshahidi LK, McSorley EM, Madigan SM, Magee PJ. Vitamin D: recent advances and implications for athletes. *Sports Med* 2015;**45**.
54. Roth SM, Zmuda JM, Cauley JA, Shea PR, Ferrell RE. Vitamin D receptor genotype is associated with fat-free mass and sarcopenia in elderly men. *Journals Gerontol Ser A Biol Sci Med Sci* 2004;**59**: B10–B15.
55. Girgis CM, Clifton-Bligh RJ, Hamrick MW, Holick MF, Gunton JE. The roles of vitamin D in skeletal muscle: form, function, and metabolism. *Endocr Rev* 2013;**34**:33–83.
56. Girgis CM, Cha KM, So B, Tsang M, Chen J, Houweling PJ, et al. Mice with myocyte deletion of vitamin D receptor have sarcopenia and impaired muscle function. *J Cachexia Sarcopenia Muscle* 2019;**10**:1228–1240.
57. Bass JJ, Kazi AA, Deane CS, Nakhuda A, Ashcroft SP, Brook MS, et al. The mechanisms of skeletal muscle atrophy in response to transient knockdown of the vitamin D receptor (VDR) in vivo. *J Physiol* December 2020;**599**:963–979.
58. Hangelbroek RWJ, Vaes AMM, Boekschoten MV, Verdijk LB, Hooiveld GJEJ, van Loon LJC, et al. No effect of 25-hydroxyvitamin D supplementation on the skeletal muscle transcriptome in vitamin D deficient frail older adults. *BMC Geriatr* 2019;**19**:1–8.
59. Ogawa S, Yakabe M, Akishita M. Age-related sarcopenia and its pathophysiological bases. *Inflamm Regen* 2016;**36**:1–6.
60. Sanders KJC, Kneppers AEM, van de Bool C, Langen RCJ, Schols AMWJ. Cachexia in chronic obstructive pulmonary disease: new insights and therapeutic perspective. *J Cachexia Sarcopenia Muscle* 2016;**7**: 5–22.
61. Baumgartner RN, Koehler KM, Gallagher D, Romero L, Heymsfield SB, Ross RR, et al. Epidemiology of sarcopenia among the elderly in New Mexico. *Am J Epidemiol* 1998;**147**:755–763.
62. Global Initiative for Chronic Obstructive Lung Disease: Global Strategy for the Diagnosis, Management, and Prevention of Chronic Obstructive Pulmonary Disease (2020 Report); 2020.
63. Holick MF, Vitamin D. Status: measurement, interpretation, and clinical application. *Ann Epidemiol* 2009;**19**:73–78.
64. Dastani Z, Berger C, Langsetmo L, Fu L, Wong BYL, Malik S, et al. In healthy adults, biological activity of vitamin D, as assessed by serum pth, is largely independent of DBP concentrations. *J Bone Miner Res* 2014;**29**:494–499.
65. Miller MR, Hankinson J, Brusasco V, Burgos F, Casaburi R, Coates A, et al. Standardisation of spirometry. *Eur Respir J* 2005;**26**:319–338.
66. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, et al. Fiji: an open-source platform for biological-image analysis. *Nat Methods* 2012;**28**:676–682.
67. 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. *J Physiol* 2020;**598**:543–565.
68. McQuin C, Goodman A, Chernyshev V, Kamentsky L, Cimini BA, Karhohs KW, et al. CellProfiler 3.0: next-generation image processing for biology. *PLoS Biol* 2018;**16**:1–17.
69. Ellefsen S, Stensløkken KO, Sandvik GK, Kristensen TA, Nilsson GE. Improved normalization of real-time reverse transcriptase polymerase chain reaction data using an external RNA control. *Anal Biochem* 2008;**376**:83–93.
70. Ellefsen S, Vikmoen O, Slettaløkken G, Whist JE, Nygaard H, Hollan I, et al. Irisin and FNDC5: effects of 12-week strength training, and relations to muscle phenotype and body mass composition in untrained women. *Eur J Appl Physiol* 2014;**114**:1875–1888.
71. Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, et al. Primer3-new capabilities and interfaces. *Nucleic Acids Res* 2012;**40**:1–12.
72. Ritz C, Spiess AN. qpcR: an R package for sigmoidal model selection in quantitative real-time polymerase chain reaction analysis. *Bioinformatics* 2008; **24**:1549–1551.
73. R Core Team. R. *A language and environment for statistical computing. R Foundation for Statistical Computing*. Vienna, Austria.
74. Tichopad A, Dilger M, Schwarz G, Pfaffl MW. Standardized determination of real-time PCR efficiency from a single reaction set-up. *Nucleic Acids Res* 2003;**31**: e122:122e–1122e.
75. Fleige S, Pfaffl MW. RNA integrity and the effect on the real-time qRT-PCR performance. *Mol Aspects Med* 2006;**27**: 126–139.
76. Ellefsen S, Vikmoen O, Zacharoff E, Rauk I, Slettaløkken G, Hammarström 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–e342.
77. Matz MV, Wright RM, Scott JG. No control genes required: Bayesian analysis of qRT-PCR data. *PLoS One* 2013;**8**:1–12.
78. 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.
79. Brooks ME, Kristensen K, Van Benthem KJ, Magnusson A, Berg CW, Nielsen A, et al. glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *R I Dent J* 2017;**9**:378.
80. Hartig F. DHARMa: residual diagnostics for hierarchical (multi-level/mixed) regression models. 2020. https://cran.r-project.org/package=DHARMa
81. Yamaguchi KD, Ruderman DL, Croze E, Wagner TC, Velichko S, Reder AT, et al. IFN-β-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**:116–120.
82. Zyla J, Marczyk, Domaszewska T, Kaufmann SHE, Polanska J, Weiner J. Gene set enrichment for reproducible science: comparison of CERNO and eight other algorithms. *Bioinformatics*. 2019;**35**: 5146–5154.
83. 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**: 15545–15550.
84. Korotkevich G, Sukhov V. *Fast gene set enri analy bioRxiv* 2019;https://doi.org/10.1101/060012
85. Xiao Y, Hsiao T-H, Suresh U, Chen H-IH, Xiaowu W, Wolf SE, et al. A novel significance score for gene selection and ranking. *Bioinformatics* 2014;**30**: 801–807.
86. Liberzon A, Subramanian A, Pinchback R, Thorvaldsdóttir H, Tamayo P, Mesirov JP. Molecular signatures database (MSigDB) 3.0. *Bioinformatics* 2011;**27**:1739–1740.
87. Gerber NL, Price JK. *Measures of function and health-related quality of life*. Elsevier Inc.; 2018.
88. Gepner AD, Ramamurthy R, Krueger DC, Korcarz CE, Binkley N, Stein JH. A prospective randomized controlled trial of the effects of vitamin D supplementation on cardiovascular disease risk. *PLoS One* 2012;**7**:e36617.
89. Kimball SM, Mirhosseini N, Holick MF. Evaluation of vitamin D3 intakes up to 15,000 international units/day and serum 25-hydroxyvitamin D concentrations up to 300 nmol/L on calcium metabolism in a community setting. *Dermatoendocri* 2017;**9**:e1300213.
90. Immunodiagnostic Systems. *1,25-dihydroxy vitamin D EIA*. Boldon; 2014. http://peramed.com/peramed/docs/AC-62F1_EN.pdf
91. Kochersberger G, Bales C, Lobaugh B, Lyles KW. Calcium supplementation lowers serum parathyroid hormone levels in elderly subjects. *J Gerontol* 1990;**45**: M159–M162.
92. Hirani V, Cumming RG, Naganathan V, Blyth F, Le Conteur DG, Handelsman DJ, et al. Associations between serum 25-hydroxyvitamin D concentrations and multiple health conditions, physical performance measures, disability, and all-cause mortality: the concord health and ageing in men project. *J Am Geriatr Soc* 2014;**62**:417–425.
93. Rejnmark L. Effects of vitamin D on muscle function and performance: a review

of evidence from randomized controlled trials. *Ther Adv Chronic Dis* 2011;**2**:25–37.
94. Nuzzo JL, Taylor JL, Gandevia SC. CORP: measurement of upper and lower limb muscle strength and voluntary activation. *J Appl Physiol* 2018;**126**:513–543.
95. Ploutz-Snyder LL, Giamis EL. Orientation and familiarization to 1RM strength testing in old and young women. *J Strength Cond Res* 2001;**15**:519–523.
96. Kurosu T, Fukuda T, Miki T, Miura O. BCL6 overexpression prevents increase in reactive oxygen species and inhibits apoptosis induced by chemotherapeutic reagents in B-cell lymphoma cells. *Oncogene* 2003;**22**:4459–4468.
97. Chen D, Zang Y-H, Qiu Y, Zhang F, Chen A-D, Wang J-J, et al. BCL6 Attenuates proliferation and oxidative stress of vascular smooth muscle cells in hypertension. *Oxid Med Cell Longev* January 2019;1–9.
98. Xiong G, Stewart RL, Chen J, Gao T, Scott TL, Samayoa LM, et al. Collagen prolyl 4-hydroxylase 1 is essential for HIF-1α stabilization and TNBC chemoresistance. *Nat Commun* 2018;**9**.
99. Wang Y, Lam KSL, Lam JBB, Lam MC, Leung PTY, Zhou M, et al. Overexpression of angiopoietin-like protein 4 alters mitochondria activities and modulates methionine metabolic cycle in the liver tissues of db/db diabetic mice. *Mol Endocrinol* 2007;**21**:972–986.
100. Kim DH, Meza CA, Clarke H, Kim JS, Hickner RC. Vitamin D and endothelial function. *Nutrients* 2020;**12**:1–17.
101. Mitchell CJ, Churchward-Venne TA, West DWD, Burd NA, Breen L, Baker SK, et al. Resistance exercise load does not determine training-mediated hypertrophic gains in young men. *J Appl Physiol* 2012;**113**:71–77.
102. Ellefsen S, Hammarström D, Strand TA, Zacharoff E, Whist JE, Rauk I, et al. Blood flow-restricted strength training displays high functional and biological efficacy in women: a within-subject comparison with high-load strength training. *Am J Physiol Regul Integr Comp Physiol* 2015;767–779.
103. Borg GA. Psychophysical bases of perceived exertion. *Med Sci Sports Exerc* 1982;**14**:377–381.
104. Hamarsland H, Johansen MK, Seeberg F, Brochmann M, Garthe I, Benestad HB, et al. Native whey induces similar adaptation to strength training as milk, despite higher levels of leucine, in elderly individuals. *Nutrients* 2019;**11**:2094.
105. Häkkinen K, Kallinen M, Linnamo V, Pastinen U-M, Newton RU, Kraemer WJ. Neuromuscular adaptations during bilateral versus unilateral strength training in middle aged and elderly men and women. *Acta Physiol Scand* 1996;**158**:77–88.
106. Ahtiainen JP, Walker S, Peltonen H, Holviala J, Sillanpää E, Karavirta L, et al. Heterogeneity in resistance training-induced muscle strength and mass responses in men and women of different ages. *Age* 2016;**38**:1–13.
107. Haun CT, Vann CG, Roberts BM, Vigotsky AD, Schoenfeld BJ, Roberts MD. A critical evaluation of the biological construct skeletal muscle hypertrophy: size matters but so does the measurement. *Front Physiol* 2019;**10**(MAR):1–23.
108. Nederveen JP, Ibrahim G, Fortino SA, Snijders T, Kumbhare D, Parise G. Variability in skeletal muscle fibre characteristics during repeated muscle biopsy sampling in human vastus lateralis. *Appl Physiol Nutr Metab* 2020;**8**:1–8.
109. McCall GE, Byrnes WC, Dickinson AL, Fleck SJ. Sample size required for the accurate determination of fiber area and capillarity of human skeletal muscle. *Can J Appl Physiol* 1998;**23**:594–599.
110. Lexell J, Taylor CC. Variability in muscle fibre areas in whole human quadriceps muscle. How much and why? *Acta Physiol Scand* 1989;**136**:561–568.
111. Miljkovic N, Lim JY, Miljkovic I, Frontera WR. Aging of skeletal muscle fibers. *Ann Rehabil Med* 2015;**39**:155–162.
112. Campbell IT, Watt T, Withers D, England R, Sukumar S, Keegan MA, et al. Muscle thickness, measured with ultrasound, may be an indicator of lean tissue wasting in multiple organ failure in the presence of edema. *Am J Clin Nutr* 1995;**62**:533–539.
113. Verdijk LB, Snijders T, Beelen M, Savelberg HHCM, Meijer K, Kuipers H, et al. Characteristics of muscle fiber type are predictive of skeletal muscle mass and strength in elderly men. *J Am Geriatr Soc* 2010;**58**:2069–2075.
114. Lexell J, Taylor CC, Sjöström M. What is the cause of the ageing atrophy?. Total number, size and proportion of different fiber types studied in whole vastus lateralis muscle from 15- to 83-year-old men. *J Neurol Sci* 1988;**84**:275–294.
115. Beals JW, Burd NA, Moore DR, van Vliet S. Obesity alters the muscle protein synthetic response to nutrition and exercise. *Front Nutr* 2019;**6**:1–14.
116. Wortsman J, Matsuoka LY, Chen TC, Lu Z, Holick MF. Decreased bioavailability of vitamin D in obesity. *Am J Clin Nutr* 2000;**72**:690–693.
117. Sipilä S, Multanen J, Kallinen M, Era P, Suominen H. Effects of strength and endurance training on isometric muscle strength and walking speed in elderly women. *Acta Physiol Scand* 1996;**156**:457–464.
118. de Vos NJ, Singh NAF, Ross DA, Stavrinos TM, Orr R, Fiatarone Singh MA. Optimal load for increasing muscle power during explosive resistance training in older adults. *Journals Gerontol Ser A Biol Sci Med Sci* 2005;**60**:638–647.
119. Solberg PA, Kvamme NH, Raastad T, Ommundsen Y, Tomten SE, Halvari H, et al. Effects of different types of exercise on muscle mass, strength, function and well-being in elderly. *Eur J Sport Sci* 2013;**13**:112–125.
120. Hikida RS, Staron RS, Hagerman FC, Walsh S, Kaiser E, Shell S, et al. Effects of high-intensity resistance training on untrained older men. II. Muscle fiber characteristics and nucleo-cytoplasmic relationships. *Journals Gerontol Ser A Biol Sci Med Sci* 2000;**55**:B347–B354.
121. Sharman MJ, Newton RU, Triplett-McBride T, McGuigan MRM, McBride JM, Häkkinen A, et al. Changes in myosin heavy chain composition with heavy resistance training in 60- to 75-year-old men and women. *Eur J Appl Physiol* 2001;**84**:127–132.
122. Moro T, Brightwell CR, Volpi E, Rasmussen BB, Fry CS. Resistance exercise training promotes fiber type-specific myonuclear adaptations in older adults. *J Appl Physiol* 2020;**128**:795–804.
123. Karlsen A, Bechshøft RL, Malmgaard-Clausen NM, Andersen JL, Schjerling P, Kjaer M, et al. Lack of muscle fibre hypertrophy, myonuclear addition, and satellite cell pool expansion with resistance training in 83-94-year-old men and women. *Acta Physiol* 2019;**227**:1–19.
124. 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–E661.
125. Brook MS, Wilkinson DJ, Mitchell WK, Lund JN, Phillips BE, Szewczyk NJ, et al. Synchronous deficits in cumulative muscle protein synthesis and ribosomal biogenesis underlie age-related anabolic resistance to exercise in humans. *J Physiol* 2016;**594**:7399–7417.
126. Medeiros JFP, de Oliveira Borges MV, Soares AA, dos Santos JC, de Oliveira ABB, da Costa CHB, et al. The impact of vitamin D supplementation on VDR gene expression and body composition in monozygotic twins: randomized controlled trial. *Sci Rep* 2020;**10**:1–10.
127. van der Meijden K, Bravenboer N, Dirks NF, Heijboer AC, den Heijer M, de Wit GMJ, et al. Effects of 1,25(OH)2D3 and 25(OH)D3 on C2C12 myoblast proliferation, differentiation, and myotube hypertrophy. *J Cell Physiol* 2016;**231**:2517–2528.
128. Pojednic RM, Ceglia L, Olsson K, Gustafsson T, Lichtenstein AH, Dawson-Hughes B, et al. Effects of 1,25-dihydroxyvitamin D3 and vitamin D3 on the expression of the vitamin D receptor in human skeletal muscle cells. *Calcif Tissue Int* 2015;**96**:256–263.
129. Khan Y, Hammarström D, Rønnestad BR, Ellefsen S, Ahmad R. Increased biological relevance of transcriptome analyses in human skeletal muscle using a model-specific pipeline. *BMC Bioinformatics* 2020;**21**:1–32.
130. Kaviani M, Sharabiyani S, Rajabi H. Effect of resistance training combined with vitamin D supplementation on adropin, NO and eNOS in stage-1 hypertensive postmenopausal women. *Proc Cana Soc Exer Physio 51st An Ge Mee - Health in Motion, Sci Exe* 2018;**43**:S68.

131. Thongprayoon C, Cheungpasitporn W, Kashani K. Serum creatinine level, a surrogate of muscle mass, predicts mortality in critically ill patients. *J Thorac Dis* 2016;**8**: E305–E311.
132. Baxmann AC, Ahmed MS, Marques NC, Menon VB, Pereira AB, Kirsztajn GM, et al. Influence of muscle mass and physical activity on serum and urinary creatinine and serum cystatin C. *Clin J Am Soc Nephrol* 2008;**3**:348–354.
133. Wehr E, Pilz S, Boehm BO, März W, Obermayer-Pietsch B. Association of vitamin D status with serum androgen levels in men. *Clin Endocrinol (Oxf)* 2010;**73**:243–248.
134. Heijboer AC, Oosterwerff M, Schroten NF, Eekhoff EMW, Chel VGM, de Boer RA, et al. Vitamin D supplementation and testosterone concentrations in male human subjects. *Clin Endocrinol (Oxf)* 2015;**83**:105–110.
135. Lerchbaum E, Trummer C, Theiler-Schwetz V, Kollmann M, Wölfler M, Heijboer AM, et al. Effects of vitamin D supplementation on androgens in men with low testosterone levels: a randomized controlled trial. *Eur J Nutr* 2019;**58**: 3135–3146.
136. Luzak A, Karrasch S, Thorand B, Nowak D, Holle R, Peters A, et al. Association of physical activity with lung function in lung-healthy German adults: results from the KORA FF4 study. *BMC Pulm Med* 2017;**17**:1–9.
137. Hoff J, Tjønna AE, Steinshamn S, Høydal M, Richardson RS, Helgerud J. Maximal strength training of the legs in COPD: a therapy for mechanical inefficiency. *Med Sci Sports Exerc* 2007;**39**:220–226.
138. Zosky GR, Berry LJ, Elliot JG, James AL, Gorman S, Hart PH. Vitamin D deficiency causes deficits in lung function and alters lung structure. *Am J Respir Crit Care Med* 2011;**183**:1336–1343.
139. Aspell N, Laird E, Healy M, Lawlor B, O'Sullivan M. Vitamin D deficiency is associated with impaired muscle strength and physical performance in community-dwelling older adults: findings from the English longitudinal study of ageing. *Clin Interv Aging* 2019;**14**:1751–1761.
140. Tang JCY, Jackson S, Walsh NP, Greeves J, Fraser WD, Bioanalytical Facility team. The dynamic relationships between the active and catabolic vitamin D metabolites, their ratios, and associations with PTH. *Scientific Reports*. 2019;**9**:1–10.
141. Goltzman D, Mannstadt M, Marcocci C. Physiology of the calcium-parathyroid hormone-vitamin D axis. *Front Horm Res* 2018;**50**:1–13.
142. Dusso AS. Kidney disease and vitamin D levels: 25-hydroxyvitamin D, 1,25-dihydroxyvitamin D, and VDR activation. *Kidney Int Suppl* 2011;**1**:136–141.
143. Bikle DD, Schwartz J. Vitamin D binding protein, total and free vitamin D levels in different physiological and pathophysiological conditions. *Front Endoc (Lausanne)* 2019;**10**(MAY):1–12.
144. Safadi FF, Thornton P, Magiera H, Hollis BW, Gentile M, Haddad JG, et al. Osteopathy and resistance to vitamin D toxicity in mice null for vitamin D binding protein. *J Clin Invest* 1999;**103**:239–251.
145. Brechue WF, Abe T. The role of FFM accumulation and skeletal muscle architecture in powerlifting performance. *Eur J Appl Physiol* 2002;**86**:327–336.
146. Netreba AI, Bravyy YR, Makarov VA, Ustyuzhanin DV, Vinogradova OL. Evaluation of training effectiveness for improving maximal voluntary contraction without noticeable hypertrophy of muscles. *Hum Physiol* 2011;**37**:719–725.
147. Solomon JA, Gianforcaro A, Hamadeh MJ. Vitamin D 3 deficiency differentially affects functional and disease outcomes in the G93A mouse model of amyotrophic lateral sclerosis. *PLoS One* 2011;**6**.
148. Ceglia L. Vitamin D and its role in skeletal muscle. *Curr Opin Clin Nutr Metab Care* 2009;**12**:628–633.
149. von Haehling S, Morley JE, Coats AJS, Anker SD. Ethical guidelines for publishing in the Journal of Cachexia, Sarcopenia and Muscle: update 2019. *J Cachexia Sarcopenia Muscle* 2019;**10**:1143–1145.

A standardized pipeline for analysis of RNA-seq dataset from human skeletal muscle was developed. For pre-processing of data, trimmomatic was found to be a better choice than trim galore. For read alignment and quantification, RSEM proved to be a better tool having less variation between replicates and increased biological relevance. In addition, the effects of using different normalization strategies for analysing and interpreting skeletal muscle transcriptome responses to resistance training was explored. Analysis of data from this model showed that the choice of the normalization model affected downstream data interpretations. While all normalization models arguably remove technical biasness, our results suggest that it may be necessary to account for biological biasness caused by global changes in total RNA population (i.e., total RNA per tissue mass) to increase the biological relevance of transcriptome analyses from resistance-trained skeletal muscle. This perspective and considerations is applicable to all sequencing technologies to have results with better biological relevance. In addition, the vitamin D3 supplementation and chronic obstructive pulmonary disease diagnosis during training linked to some molecular change in skeletal muscle seen in transcriptome responses with improvement in lung function but does not account for significant phenotypic change as muscle growth.

Inland Norway
University of
Applied Sciences