University Library

MINERVA ACCESS

A gateway to Melbourne's research publications

Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:
Molendijk, J;Blazev, R;Mills, RJ;Ng, YK;Watt, KI;Chau, D;Gregorevic, P;Crouch, PJ;Hilton, JBW;Lisowski, L;Zhang, P;Reue, K;Lusis, AJ;Hudson, JE;James, DE;Seldin, MM;Parker, BL

Title:
Proteome-wide systems genetics identifies UFMylation as a regulator of skeletal muscle function

Date:
2022-01-01

Citation:
Molendijk, J., Blazev, R., Mills, R. J., Ng, Y. K., Watt, K. I., Chau, D., Gregorevic, P., Crouch, P. J., Hilton, J. B. W., Lisowski, L., Zhang, P., Reue, K., Lusis, A. J., Hudson, J. E., James, D. E., Seldin, M. M. & Parker, B. L. (2022). Proteome-wide systems genetics identifies UFMylation as a regulator of skeletal muscle function. Elife, 11, https://doi.org/10.7554/ELIFE.82951.

Persistent Link:
https://hdl.handle.net/11343/327267

License:
CC BY

RESEARCH ARTICLE

Proteome-wide systems genetics identifies UFMylation as a regulator of skeletal muscle function

Jeffrey Molendijk[1,2], Ronnie Blazev[1,2], Richard J Mills[3], Yaan-Kit Ng[1,2], Kevin I Watt[1,2†], Daryn Chau[4], Paul Gregorevic[1,2], Peter J Crouch[5], James BW Hilton[5], Leszek Lisowski[6,7], Peixiang Zhang[8], Karen Reue[8], Aldons J Lusis[8,9], James E Hudson[3], David E James[10], Marcus M Seldin[4], Benjamin L Parker[1,2]*

[1]Department of Anatomy and Physiology, University of Melbourne, Melbourne, Australia; [2]Centre for Muscle Research, University of Melbourne, Melbourne, Australia; [3]QIMR Berghofer Medical Research Institute, Brisbane, Australia; [4]Department of Biological Chemistry and Center for Epigenetics and Metabolism, University of California, Irvine, Irvine, United States; [5]Department of Biochemistry and Pharmacology, University of Melbourne, Melbourne, Australia; [6]Children's Medical Research Institute, University of Sydney, Sydney, Australia; [7]Military Institute of Medicine, Warszawa, Poland; [8]Department of Human Genetics/Medicine, David Geffen School of Medicine, University of California, Los Angeles, Los Angeles, United States; [9]Department of Microbiology, Immunology and Molecular Genetics, University of California, Los Angeles, Los Angeles, United States; [10]Charles Perkins Centre, School of Life and Environmental Science, School of Medical Science, University of Sydney, Sydney, Australia

***For correspondence:** ben.parker@unimelb.edu.au

Present address: †Novo Nordisk Foundation Centre for Stem Cell Medicine, Murdoch Children's Research Institute, Parkville, Australia

Competing interest: See page 18

Funding: See page 19

Preprinted: 22 August 2022
Received: 24 August 2022
Accepted: 29 November 2022
Published: 06 December 2022

Reviewing Editor: Charles Farber, University of Virginia, United States

© Copyright Molendijk *et al.* This article is distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use and redistribution provided that the original author and source are credited.

Abstract Improving muscle function has great potential to improve the quality of life. To identify novel regulators of skeletal muscle metabolism and function, we performed a proteomic analysis of gastrocnemius muscle from 73 genetically distinct inbred mouse strains, and integrated the data with previously acquired genomics and >300 molecular/phenotypic traits via quantitative trait loci mapping and correlation network analysis. These data identified thousands of associations between protein abundance and phenotypes and can be accessed online (https://muscle.coffeeprot.com/) to identify regulators of muscle function. We used this resource to prioritize targets for a functional genomic screen in human bioengineered skeletal muscle. This identified several negative regulators of muscle function including UFC1, an E2 ligase for protein UFMylation. We show UFMylation is up-regulated in a mouse model of amyotrophic lateral sclerosis, a disease that involves muscle atrophy. Furthermore, in vivo knockdown of UFMylation increased contraction force, implicating its role as a negative regulator of skeletal muscle function.

Editor's evaluation

This manuscript will be of broad interest to those working in the genetics of complex diseases, with the results strongly supporting the author's primary claims. Overall, this is an important study that demonstrates the power of proteomics-based systems genetics studies in the mouse.

Introduction

Building and maintaining healthy skeletal muscles is crucial for all stages of life. Skeletal muscle makes up 30–40% of an adult human's body mass and is vital not just for breathing and movement, but also for metabolism and longevity. The maintenance of muscle function is one of the best predictors for overall health, with sarcopenia being the major contributor of age-associated frailty (***McGregor et al., 2014***). Identifying factors that regulate skeletal muscle function has great potential to improve the quality of life for humans and animals.

Systems genetics is a population-based approach that links genetic variation to complex traits (***Baliga et al., 2017***). These forward genetics approaches integrate genomics and other multi-omic data to phenotypic traits using various modelling approaches including genetic mapping and correlation analysis. Advances in high-throughput proteomic techniques have enabled the analysis of genetically diverse populations, and integration of these data with systems genetics has emerged as a powerful approach to identify novel associations between protein abundance and complex phenotypes (for a review, see ***Molendijk and Parker, 2021a***). Recent large-scale plasma proteomic studies in human populations have begun to unravel complex genetic variations and their contribution to proteome diversity and disease-relevant phenotypes (***Suhre et al., 2017***; ***Benson et al., 2018***; ***Sun et al., 2018***; ***Emilsson et al., 2018***). However, unlike studies in humans, the use of genetic reference panels (GRPs) enables accurate control of the environment, breeding patterns, and access to a range of tissues for molecular analysis. Systems genetic analyses incorporating transcriptomics, lipidomics, and/or metabolomics in GRPs have led to the discovery of a range of novel regulators of complex phenotypes ranging from insulin resistance (***Parks et al., 2015***), insulin secretion (***Keller et al., 2019***), atherosclerosis (***Bennett et al., 2015***), lipid metabolism (***Jha et al., 2018a***; ***Jha et al., 2018b***; ***Linke et al., 2020***), cardiac hypertrophy (***Rau et al., 2015***), cardiac diastolic dysfunction (***Cao et al., 2022***), and many more. The inclusion of proteomics into systems genetic analysis provides information on an important biological layer and has been performed in a range of GRPs including yeast (***Foss et al., 2007***; ***Picotti et al., 2013***; ***Parts et al., 2014***), worms (***Singh et al., 2016***), fruit fly (***Okada et al., 2016***), plants such as maize (***Hu et al., 2017***; ***Jiang et al., 2019***), and several livestock such as cattle (***Boudon et al., 2020***) and pig (***Bovo et al., 2018***). The use of proteomics to analyse the liver proteome of mouse GRPs has also gained popularity and been used to analyse the BxD panel (***Wu et al., 2014***; ***Williams et al., 2016***), the Hybrid Mouse Diversity Panel (HMDP) (***Ghazalpour et al., 2011***; ***Parker et al., 2019***), and cohorts of the Collaborative Cross/Diversity Outbred (CC/DO) (***Chick et al., 2016***). More recently, several studies have performed proteomic analysis of additional tissues from cohorts of the BXD (***Williams et al., 2018***) and CC/DO (***Xiao et al., 2022***) and include further phenotypic associations.

Functional screening of genetic perturbations with high-throughput phenotypic measurements have identified novel regulators of muscle biology and disease. These include forward genetic mutagenesis screens to identify regulators of skeletal muscle development and locomotion in zebrafish (***Birely et al., 2005***; ***Horstick et al., 2013***; ***Johnson et al., 2013***; ***Bennett et al., 2018***) and worms (***Beron et al., 2015***), RNAi screening of muscle size and function in fruit fly (***Kao et al., 2021***; ***Graca et al., 2021***), and CRISPR/Cas9 screening of muscle cells to identify regulators of myogenesis and cell survival in dystrophy models (***Bi et al., 2017***; ***Lek et al., 2020***; ***Ashoti et al., 2022***). Here, we combined forward genetics via proteomic analysis of a diverse mouse panel with a targeted reverse genetics screen via AAV6 vector-mediated expression of shRNAs to knock down specific genes in bioengineered skeletal muscle to identify candidate regulators of skeletal muscle function. Our approach identified UFMylation as a regulator of skeletal muscle function that was validated in vivo.

Results

Genetic regulation of the mouse skeletal muscle proteome

To begin to understand how variations in the genome drive changes in the skeletal muscle proteome, we performed a proteomic analysis of gastrocnemius muscle from 73 inbred mouse strains of the HMDP that were fed a chow diet and housed under identical environmental conditions (n=2–4; 161 mice). The proteomic data were integrated with previously acquired genomic and various molecular/phenotypic data via systems genetics analysis (***Figure 1A***). Proteomics was performed with eighteen 10-plex tandem mass tag (TMT) experiments each consisting of nine strains plus a pooled common internal

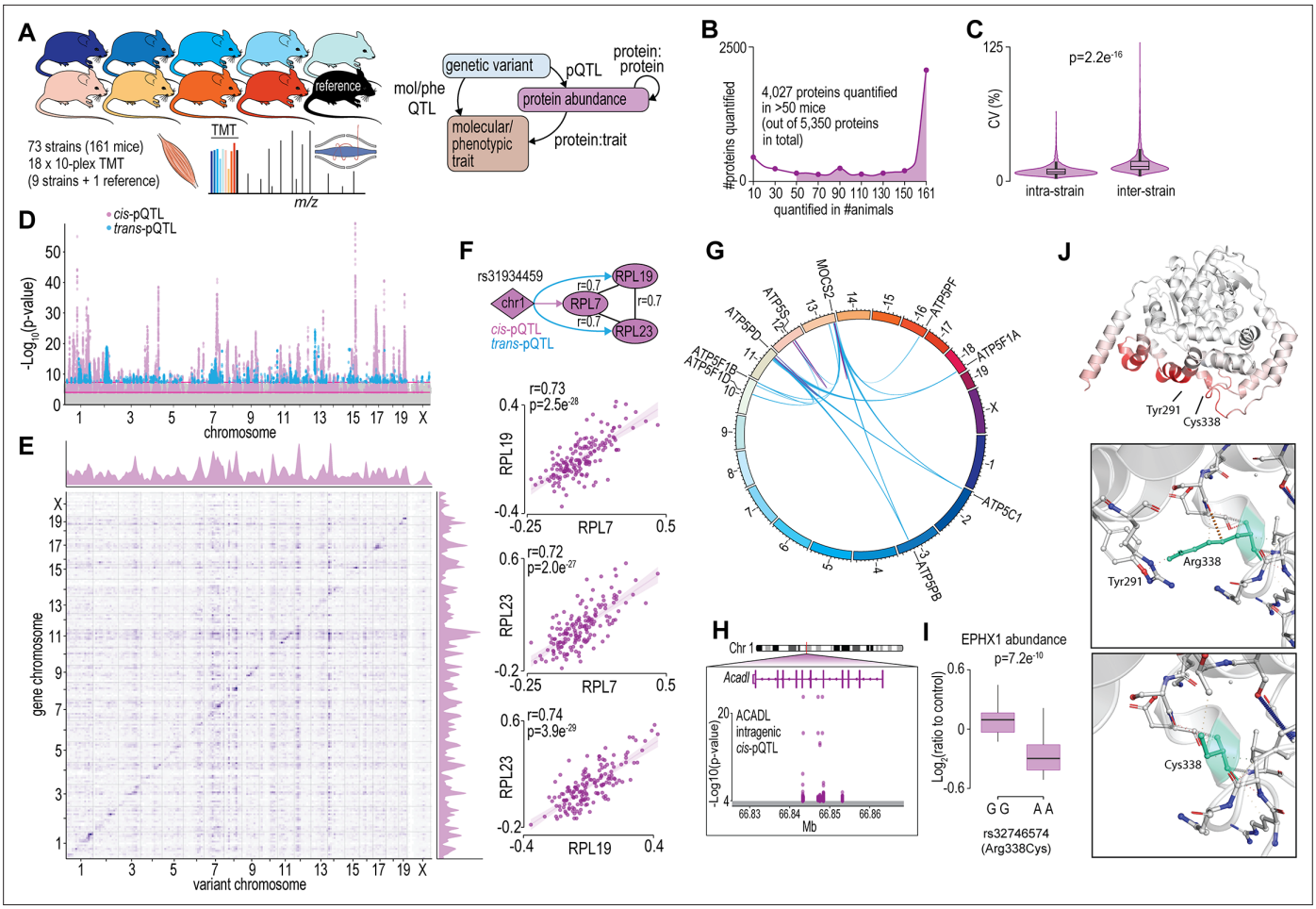


Figure 1. Proteome-wide systems genetics analysis of the mouse skeletal muscle proteome. (**A**) Overview of the experimental design. (**B**) Number of proteins identified. (**C**) Intra- and inter-strain coefficient of variation. (**D**) Protein-quantitative trait loci (pQTL) Manhattan plot. (**E**) pQTL variant and gene location density. (**F**) Ribosomal proteins correlation and variant network (upper), and scatterplots expressed as Log2(ratio to control) showing correlation coefficient calculated using biweight midcorrelation (n=161) (lower). (**G**) Genetic associations of variant hotspot on chromosome 13 associated with mitochondrial complex V subunits in trans. (**H**) Intragenic variants associated to ACADL abundance. (**I**) Boxplot showing variant allele associated to EPHX1 abundance (Student's t-test). (**J**) EPHX1 Arg338Cys mutation DynaMut protein flexibility analysis.

The online version of this article includes the following figure supplement(s) for figure 1:

Figure supplement 1. Hybrid Mouse Diversity Panel (HMDP) mouse sample dendrogram.

Figure supplement 2. EPHX1 Arg338Cys mutation analysis.

reference. Peptides were analysed by 2D-liquid chromatography coupled to tandem mass spectrometry (2D-LC-MS/MS) resulting in the quantification of 5350 proteins with 4027 quantified in >50 mice and 2069 proteins quantified in all 161 animals (***Figure 1B*** and ***Supplementary file 1***). Biological replicates showed similar proteomes, as evidenced by the hierarchical clustering dendrogram, where mice of the same strains are typically neighbouring or located closely (***Figure 1—figure supplement 1***). Median intra-strain coefficient of variation was 9.2% while inter-strain coefficient was significantly larger, suggesting reproducible quantification and genetically driven variation in the proteome was captured in the data (***Figure 1C***).

To identify possible genetic factors regulating protein abundance, we next associated single nucleotide polymorphisms (SNPs) to the abundance of 4027 skeletal muscle proteins quantified in >50 mice via a protein-quantitative trait loci (pQTL) mapping. We identified significant *cis*-regulation of the proteome with local SNPs associated to the abundance of 527 unique proteins (±10 Mb of gene; local adjusted $p<1 \times 10^{-4}$) and *trans*-regulation with distant SNPs (>10 Mb of the gene or on a different chromosome) associated to the abundance of 170 unique proteins (global adjusted p<5 ×

10^{-8}) (***Figure 1D*** and ***Supplementary file 2***). Visualizing SNP densities identified regions with a high number of variants (***Figure 1E***). Of particular interest are the genetic 'hotspots' (observed as vertical 'streaks' on chromosomes 7, 10, 12, and 14) abundant in *trans*-associations affecting genes in various locations. These regions agree with those observed in eQTLs of the HMDP (***Lusis et al., 2016***). We further performed variant effect predictions and annotated the position of variants relative to gene location (within a gene, i.e., intragenic or between genes, i.e., intergenic SNPs) to investigate possible mechanisms regulating protein abundance (***McLaren et al., 2016***). We observed a complex array of pQTLs on chromosome 1 with 46 and 21 proteins regulated in -*cis* and -*trans*, respectively. The majority of these pQTLs are in the distal region of the chromosome previously described as the QTL-rich region on chromosome 1 (Qrr1) with separate linkage disequilibrium (LD) blocks within 1qH2.1, 1qH5, and 1qH3 (***Mozhui et al., 2008***). The various loci in Qrr1 containing these pQTLs have been associated with a range of neural, behavioural, and cardiometabolic phenotypes, and form complex co-regulatory networks modulating a range of pathways such as RNA metabolism and translation. Here, we show a *cis*-pQTL on chromosome 1 (lead SNP rs31934459) is associated to the abundance of RPL7 and is also a *trans*-pQTL for RPL19 and RPL23. These proteins are correlated in abundance and form a co-regulated network with other 60S ribosomal proteins suggesting genetic variants regulating RPL7 abundance subsequently regulate protein complex stability/assembly (***Figure 1F*** and ***Supplementary file 3***). The distal region of chromosome 13 also contains a complex series of *cis*- and *trans*-pQTLs giving rise to 'hotspot' co-regulation. A *cis*-pQTL associated with MOCS2 is also a *trans*-pQTL for eight proteins all members of mitochondrial complex V (***Figure 1G***). MOCS2 is involved in molybdopterin biosynthesis and it is unclear how MOCS2 might regulate mitochondrial complex V abundance. Patients with mutations in *MOCS1* that result in mild molybdenum cofactor deficiency also display reduced mitochondrial respiration suggesting a link between molybdopterin biosynthesis and ATP production (***Grings et al., 2019***). Among the 527 proteins with a *cis*-pQTL association, 212 had an intragenic association. For example, non-coding SNPs in the fifth and sixth intron of *Acadl* were *cis*-pQTLs for ACADL (***Figure 1H***). We also identified 14 missense variants as pQTLs such as the rs32746574 variant (GG >AA; R338C), which was associated with significantly lower abundance of EPHX1 (***Figure 1I***). The R338C mutation was found to be deleterious (PROVEAN score: –5.6) (***Choi and Chan, 2015***), destabilizing (FoldX:+1.1 ΔG) (***Delgado et al., 2019***) and possibly damaging (PolyPhen2 HumDiv: 0.62) (***Adzhubei et al., 2013***; ***Figure 1—figure supplement 2***). R338C increases the flexibility of EPHX1 due to the missing interaction (hydrogen bond) between R338 and Y291 (***Figure 1J***). Aligning human and mouse EPHX1 protein structures revealed that the missense mutation causing R338C (yellow) is in close proximity to the catalytic (blue) and disease-related sites (red) identified by ***Gautheron et al., 2021***; ***Figure 1—figure supplement 2***. Taken together, our proteomic analysis of the HMDP has helped define the genetic factors potentially regulating the abundance of hundreds of skeletal muscle proteins.

Systems genetics integration of the skeletal muscle proteome with molecular and phenotypic traits

The renewable nature of extensively characterized inbred mouse strains from genetic references panels such as the HMDP and the BxD allow for the integration of data across multiple cohorts (***Lusis et al., 2016***; ***Ashbrook et al., 2021***). We focused our analysis on a subset of 300 molecular or phenotypic traits incorporating various plasma metabolites, lipids, and cytokines; whole body measurements such as glucose/insulin sensitivity and body composition/organ weights; and muscle phenotypes such as cardiac and skeletal muscle function previously quantified in the same strains of mice in the HMDP (***Lusis et al., 2016***). Note that data integration is performed at the strain level, since the proteomic data was not generated from the same mice, as those used in previous studies. ***Supplementary file 4*** summarizes all the phenotypic data integrated in the current study and includes data sources. Cumulatively, these previous systems-level mouse studies identified hundreds of loci associated to molecular or phenotypic traits (mol/pheQTLs) (***Figure 2A*** and ***Supplementary file 5***). We next devised a three-step systems genetic analyses to associate these molecular or phenotypic traits to the skeletal muscle proteome that included: (1) identification of SNPs shared between skeletal muscle *cis*-pQTLs and mol/pheQTLs, (2) correlation and supervised multivariate associations between the abundance of skeletal muscle proteins and each molecular or phenotypic trait, and (3) comparison of protein abundance and molecular or phenotypic differences between allelic variations (***Figure 2B***). All data can be browsed at

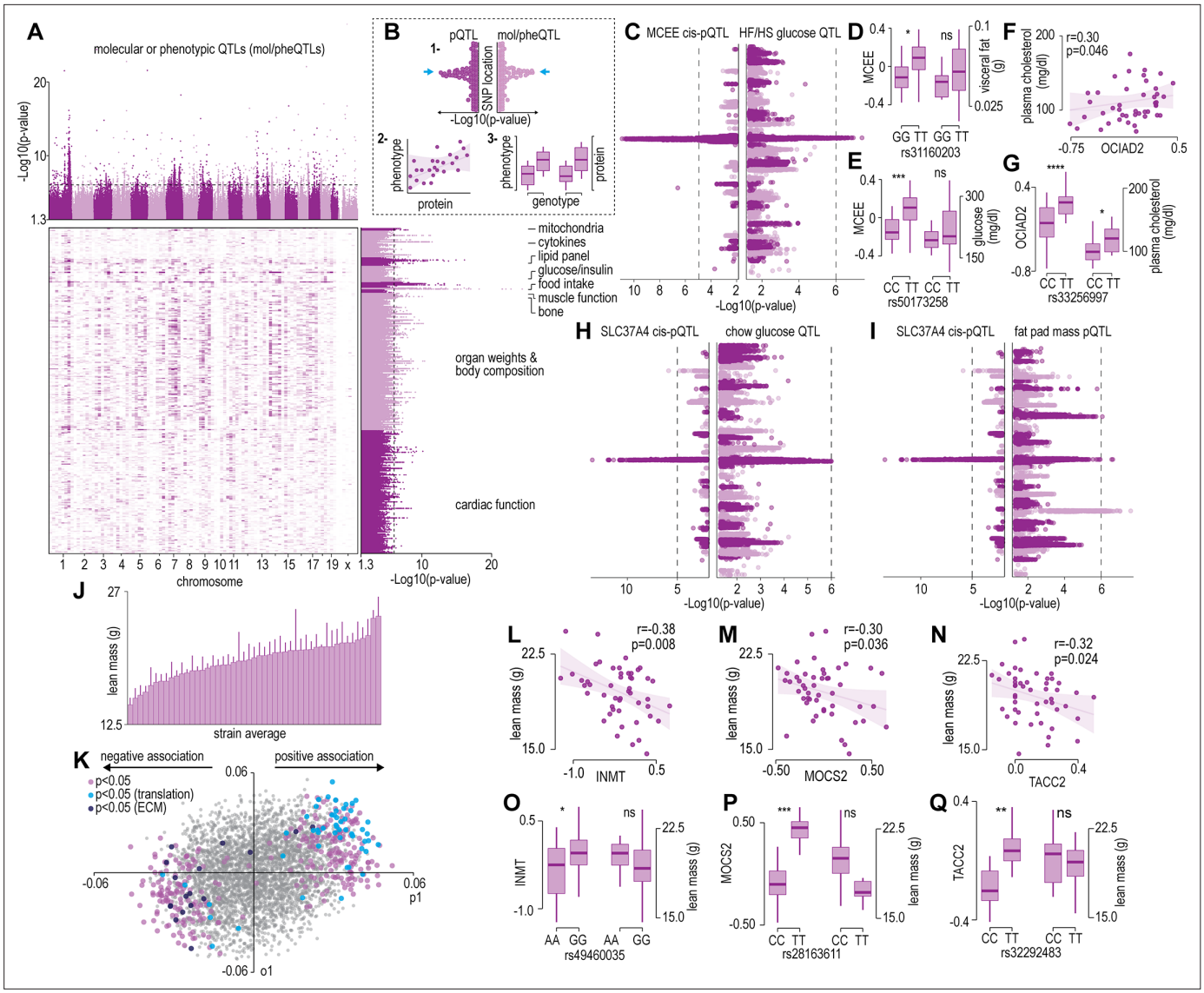


Figure 2. Protein and phenotype quantitative trait locus (QTL) analysis. (**A**) Manhattan plot and genomic location distribution of mol/pheQTLs. (**B**) Overview of the three-step integrative analysis approach. (**C**) Mirrored Manhattan plots of MCEE and glucose QTLs. (**D**) Allelic variant boxplots of rs31160203 for MCEE and visceral fat. (**E**) Allelic variant boxplots of rs50173258 for MCEE and glucose. (**F**) Correlation scatterplot of OCIAD2 abundance expressed as Log2(ratio to control) and plasma cholesterol concentrations. (**G**) Allelic variant boxplots of rs33256997 for OCIAD2 and plasma cholesterol. (**H**) Mirrored Manhattan plots of SLC37A4 and glucose QTLs. (**I**) Mirrored Manhattan plots of SLC37A4 and fat pas mass QTLs. (**J**) Average distribution of lean mass per mouse strain. (**K**) Orthogonal partial least-squares (OPLS) loading plot of proteins explaining the variance related to strain lean mass. Separation on the x-axis shows variation related to the predictive component (p1), whilst the y-axis shows the orthogonal component (o1). Highlighted points reflect Student's correlation p-values for multiple biweight midcorrelations of proteins correlated with lean mass (–0.3< r > 0.6, p<0.05). Correlation of lean mass and the protein abundance expressed as Log2(ratio to control) of INMT (**L**), MOCS2, (**M**) and TACC2 (**N**). Allelic variant boxplots of selected single nucleotide polymorphisms (SNPs) with lean mass and INMT (rs49460035) (**O**), MOCS2 (rs28163611) (**P**), and TACC2 (rs32292483) (**Q**). *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001 by Student's t-test.

muscle.coffeeprot.com and includes querying at the protein- or phenotype-level followed by several interactive visualizations. Analysis of this new resource identified hundreds of associations allowing for the prioritization of potential causal proteins regulating molecular or phenotypic traits. We present a few examples of phenotype-protein associations below. For example, a locus on chromosome 7 contained *cis*-pQTLs associated to MCEE that were shared with fasting glucose in mice subject to high-fat/high-sugar feeding (***Figure 2C***). Homozygous allelic variation of SNPs were associated to MCEE protein abundance and also a trend for greater visceral adiposity and fasting glucose on a chow diet (***Figure 2D–E***). MCEE functions as a methylmalonyl-CoA epimerase important for propionyl-CoA

metabolism. Rare autosomal recessive missense mutations in *MCEE* have been identified in patients with methylmalonic aciduria (***Bikker et al., 2006***). We also identified *cis*-pQTLs associated to the abundance of both OCIAD1 and OCIAD2; two neighbouring genes located on chromosome 5. Intragenic SNPs in both genes were associated with several cardiometabolic parameters of adiposity including percentage body fat assessed by nuclear magnetic resonance, retroperitoneal fat mass, circulating free fatty acids, HDL, and cholesterol. Furthermore, the abundance of OCIAD1 and OCIAD2 were positively correlated to several of these traits, and homozygous allelic variation of SNPs in both the *Ociad1* and *Ociad2* loci were associated to protein abundance and cholesterol (***Figure 2F–G***). Very little is known about the functions of OCIAD1/2 but recent data have revealed a role in mitochondrial complex III assembly (***Le Vasseur et al., 2021***) and human GWAS analysis has identified variants in both the *OCIAD1* and *OCIAD2* loci are associated with susceptibility to type 2 diabetes (***Vujkovic et al., 2020***). We also identified genetic variants associated to the abundance of the ER resident glucose-6-phosphate transporter SLC37A4 that co-localize to fasting glucose and fat pad mass in mice fed a chow diet (***Figure 2H–I***). SLC37A4 has enhanced expression in the liver, gut, and kidney, and plays a role in the regulation of glycogenolysis and gluconeogenesis, however, its role in skeletal muscle metabolism has been comparatively less studied. Patients with mutations in SLC37A4 present with several metabolic complications, particularly hepatomegaly and enlarged kidneys due to the accumulation of glycogen but also often have signs of dyslipidaemia and hypoglycemia.

We next analysed associations of the skeletal muscle proteome to lean mass which showed a range of genetic variation in female mice fed a chow diet (***Figure 2J***). Using supervised multivariate and pairwise correlation analysis, we identified 300 proteins positively or negatively associated to lean

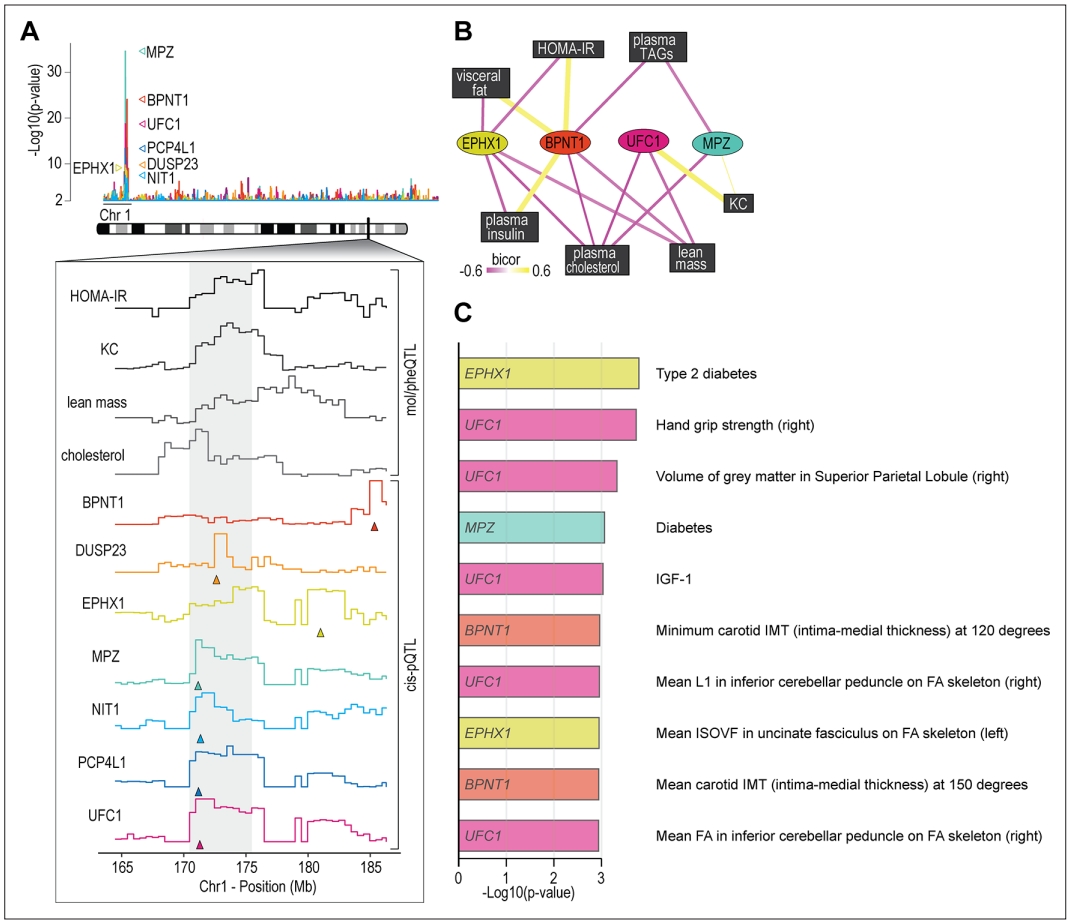


Figure 3. Proteome-phenotype associations of Qrr1 region on chromosome 1. (**A**) Manhattan plot of selected genes located near the Qrr1 region, with corresponding traits. (**B**) Protein-trait correlation network. (**C**) Top 10 GeneBass associations from the 'UK BioBank Assessment Centre' and 'Biological samples' categories, excluding 'Touchscreen', 'Medications', and 'Operations' categories.

mass (–0.3 < r > 0.6, p<0.05) (***Figure 2K***). Proteins involved in translation including ribosomal subunits and elongation initiation factors were positively correlated while we observed several extracellular matrix proteins including collagens and cathepsins to be negatively correlated with lean mass. Allelic variations were integrated to highlight several trends such as negative correlations between INMT, MOCS2, and TACC2 versus lean mass (***Figure 2L–N***), where alleles which reduced protein abundance (rs49460035, rs28162611, and rs32292483) generally displayed lower lean mass (***Figure 2O–Q***). We also investigated the relationships between mol/pheQTL's and skeletal muscle *cis*-pQTLs in or around the Qrr1 region, located distal on chromosome 1. ***Figure 3A*** displays regional association plots of pQTLs (arrows indicating gene location) shared with mol/pheQTLs including HOMA-IR, lean mass, plasma concentrations of keratinocyte-derived growth factor (KC; CXCL1), and plasma cholesterol. Two genomic regions are associated with HOMA-IR which contains *cis*-pQTLs for MPZ, NIT1, PCP4L1, and UFC1 in the first region while EPHX1 is located in the second more distal region. It is important to note that these regions were also associated with plasma insulin concentrations and as such, variations in the expression of these proteins may play greater roles in other tissues such as the pancreas. The QTL for plasma cholesterol and KC concentrations were only shared to the first region of the HOMA-IR blocks while the QTL for lean mass was distributed across both regions. To further investigate these associations, we correlated the abundance of the proteins in skeletal muscle to the molecular or phenotypic traits (***Figure 3B***). Our data identified a negative correlation between UFC1, lean mass, and plasma cholesterol. UFC1 is the E2 ligase for the post-translational modification of ubiquitin-fold modifier 1 (UFM1) to target proteins through UFMylation. We also identified negative correlations between MPZ and plasma cholesterol/triacylglycerols (TAGs). It is unclear if MPZ plays a causative role in the regulation of lipid metabolism, however diabetic peripheral neuropathy is associated with reduced expression of Mpz, myelin abnormalities, and several defects in lipid metabolism (***Cermenati et al., 2012***). Positive correlations were found between BPNT1 and plasma insulin, HOMA-IR, and visceral fat while also negatively correlated with plasma cholesterol, TAGs, and lean mass. BPNT1 is involved in phosphatidylinositol phosphate and adenosine phosphate metabolism. Mice lacking *Bpnt1* display severe liver defects but the role of this enzyme in whole body energy metabolism and insulin sensitivity is currently unknown (***Hudson et al., 2013***). Finally, we observed an overall negative correlation between the abundance of EPHX1 and HOMA-IR, visceral fat, plasma insulin, and cholesterol. Our correlation results are further corroborated by publicly available UK Biobank gene-trait associations accessed through the Genebass webserver (***Karczewski et al., 2022***; ***Figure 3C***). Of particular interest are the associations between *EPHX1*/type 2 diabetes, *UFC1*/hand grip strength, and *UFC1*/IGF-1. The associations of UFC1 with grip strength (UK BioBank, human) and lean mass (HMDP, mouse) indicate a potential role of this protein in skeletal muscle compositions and/or functional capacity.

Targeted functional genomic screen in bioengineered skeletal micro-muscles

To validate potential causal regulators of muscle function, we targeted genes encoding novel skeletal muscle pQTLs and molecular/phenotypic associations and performed a targeted functional genomic screen in human skeletal micro-muscles (hμMs) (***Mills et al., 2019***). We focused on proteins with negative associations to lean mass, grip strength, or other metabolic traits, and generated a total of 27 individual recombinant AAV serotype 6 viral vectors expressing shRNA (rAAV6:shRNAs) to knock down the expression of these proteins in an arrayed fashion. hμMs were grown around flexible pillars to assess contractile force during electrical stimulation, and transduced following differentiation and maturation to limit effects on the myogenic program (***Figure 4A***). Electrical stimulation was applied to induce either a high-frequency tetanic contraction for assessment of maximum force producing capacity or stimulated with sustained lower frequency for assessment of endurance/fatigue. Following this protocol, hμMs were analysed by proteomics which quantified 17/27 targets with 13 targets significantly reduced in abundance by rAAV6:shRNA (***Figure 4B***). Knockdown of UFC1, MCEE, TOM1L2, and SH3BGR was confirmed at the protein level and resulted in significant increases in maximum force production during the tetanic contractions (***Figure 4C***). These effects of knockdown were consistent with the observed negative correlation between lean mass and the abundance of UFC1, MCEE, and SH3BGR in skeletal muscle of the HMDP providing evidence for a causal regulation of muscle function. The fatigue protocol resulted in 20% decline in muscle function in control scramble-treated

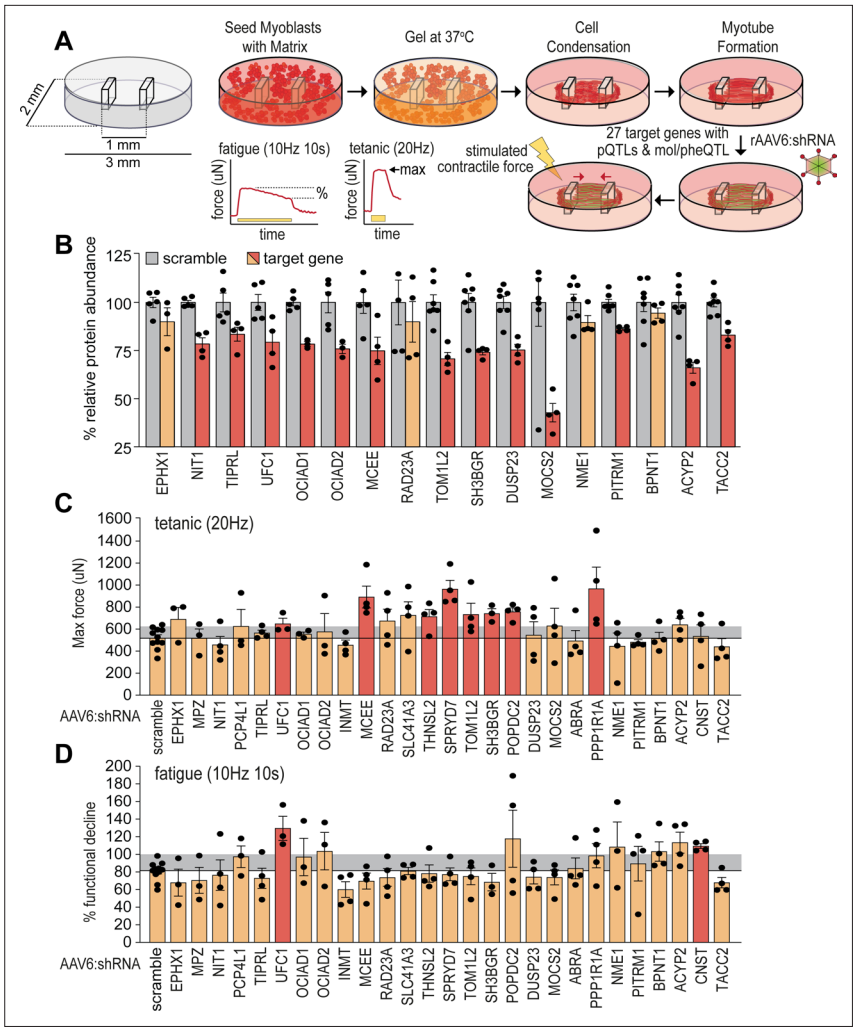


Figure 4. Functional screening of skeletal muscle function. (**A**) Overview of experimental design. (**B**) Knockdown efficiency of target proteins (n=4–10). (**C**) Maximum tetanic force, and (**D**) % fatigue of rAAV6:shScramble and target proteins. Red: q<0.05; yellow: q>0.05 (Student's t-test relative to scramble with Benjamini-Hochberg FDR).

hμMs (***Figure 4D***). Remarkably, knockdown of UFC1 increased force production during the fatigue protocol. Knockdown of CNST also protected against fatigue, but we were unable to identify CNST in the proteomics analyses to confirm knockdown. Taken together, these data identify potential causal negative regulators of muscle function.

UFMylation regulates skeletal muscle function

Our data suggest that the regulation of UFC1 and subsequent changes in UFMylation may play a key role in muscle function. We first investigated the regulation of UFMylation in a mouse model of amyotrophic lateral sclerosis (ALS), a rapidly progressive adult-onset disease that involves substantial muscle atrophy (***Pansarasa et al., 2014***). The model involves progressive muscle atrophy from 11 weeks of age due to transgenic expression of an ALS-causing mutation in superoxide dismutase 1 (SOD1(G37R)) (***Wong et al., 1995***). At 25 weeks of age, we observed a significant increase in conjugated and free UFM1, UFC1, and UFSP2 (a deUFMylase) in gastrocnemius skeletal muscles of SOD1(G37R) mice, which was independent of any changes in the abundance of BiP chaperone as a marker of ER stress (***Figure 5A–B***). These data suggest the overall pool of UFM1 increases and there is an increase in UFMylation flux. We provide the first evidence of changes in UFMylation following muscle atrophy in vivo. We next manipulated in vivo UFMylation levels by injecting rAAV6:shRNA into tibialis anterior (TA) and extensor digitorum longus (EDL) skeletal muscles of 8-week-old C57BL/6J

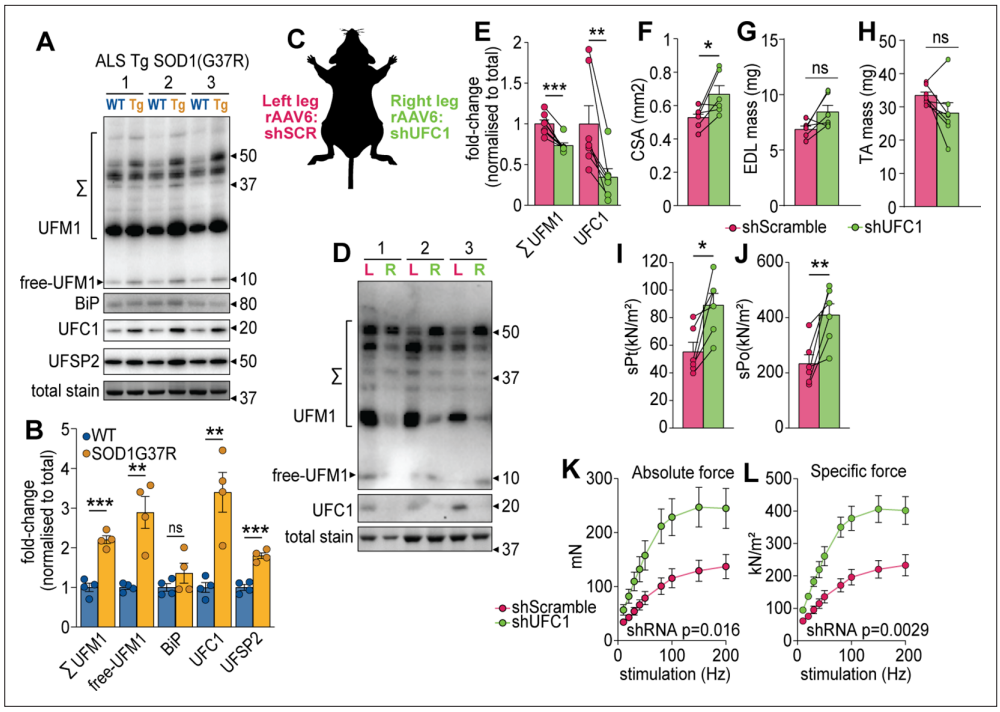


Figure 5. UFMylation is regulated in atrophy and influences skeletal muscle function. (**A**) Western blot and (**B**) densitometry of UFMylation and BiP chaperone in a gastrocnemius muscle of a mouse model of amyotrophic lateral sclerosis (ALS). (**C**) Overview of the experimental design. (**D**) Western blot of extensor digitorum long (EDL) muscles treated with rAAV6:shScramble (red, left leg (L)) and rAAV:shUFC1 (green, right leg (R)). (**E**) Densitometry of western blot (n=6). (**F**) Muscle cross-sectional area (CSA) (n=6). (**G**) EDL mass and (**H**) tibialis anterior (TA) mass (n=6). Ex vivo analysis of contraction force in EDL muscles showing (**I**) single twitch contraction force normalized to CSA (sPt), (**J**) tetanic contraction force normalized to CSA (sPt), and (**K**) absolute, and (**L**) specific force normalized to CSA following shUFC1 or scrambled control. *p/q-value<0.05; **p/q-value<0.01; ***p/q-value<0.005; (B–C) paired Student's t-test; (E–J) paired Student's t-test; (K–L) two-way ANOVA.

The online version of this article includes the following source data for figure 5:

Source data 1. Zip file containing uncropped western blot image files as Image Lab Documents, tiff files, and a summarized.pdf highlighting the lane identifications, highlighted bands used to create ***Figure 5A***, antibody information, and all densitometry results for each individual sample.

Source data 2. Zip file containing uncropped western blot image files as Image Lab Documents, tiff files, and a summarized.pdf highlighting the lane identifications, highlighted bands used to create ***Figure 5D***, antibody information, and all densitometry results for each individual sample.

mice in a paired experimental design, where muscles of the left leg received rAAV6:shScramble and muscles of the right contralateral leg received rAAV6:shUFC1 (***Figure 5C***). Following 12 weeks of transduction, we observed significant reductions in the abundance of UFC1 and conjugated UFM1 (***Figure 5D–E***). There was a subtle but significant increase in whole muscle cross-sectional area (CSA) but no change in muscle mass (***Figure 5G–H***). EDL muscles were subjected to ex vivo assessments of muscle function using electrically induced contractions. Knockdown of UFC1 increased specific force production in response to a single twitch contraction (sPt) and generated almost a doubling of peak force (sPo), normalized to CSA (***Figure 5I–J***). Furthermore, both the absolute and specific tetanic force increased across all stimulation frequencies tested following knockdown of UFC1 (***Figure 5K–L***).

We next performed a more detailed cellular and molecular analysis. First, we analysed fiber-type composition of TA muscles by immunofluorescence microscopy of which revealed no differences in the abundance of myosin isoforms but there was a trend for a decrease in the total number of muscle fibers following knockdown of UFC1 (***Figure 6A–C***). Next, we performed a proteomic analysis of EDL muscles which quantified 5909 proteins of which 573 were regulated in abundance following knockdown of UFC1 (***Figure 6D*** and ***Supplementary file 6***). The top up-regulated pathways included proteins associated with translation, muscle contraction, and signal recognition particle

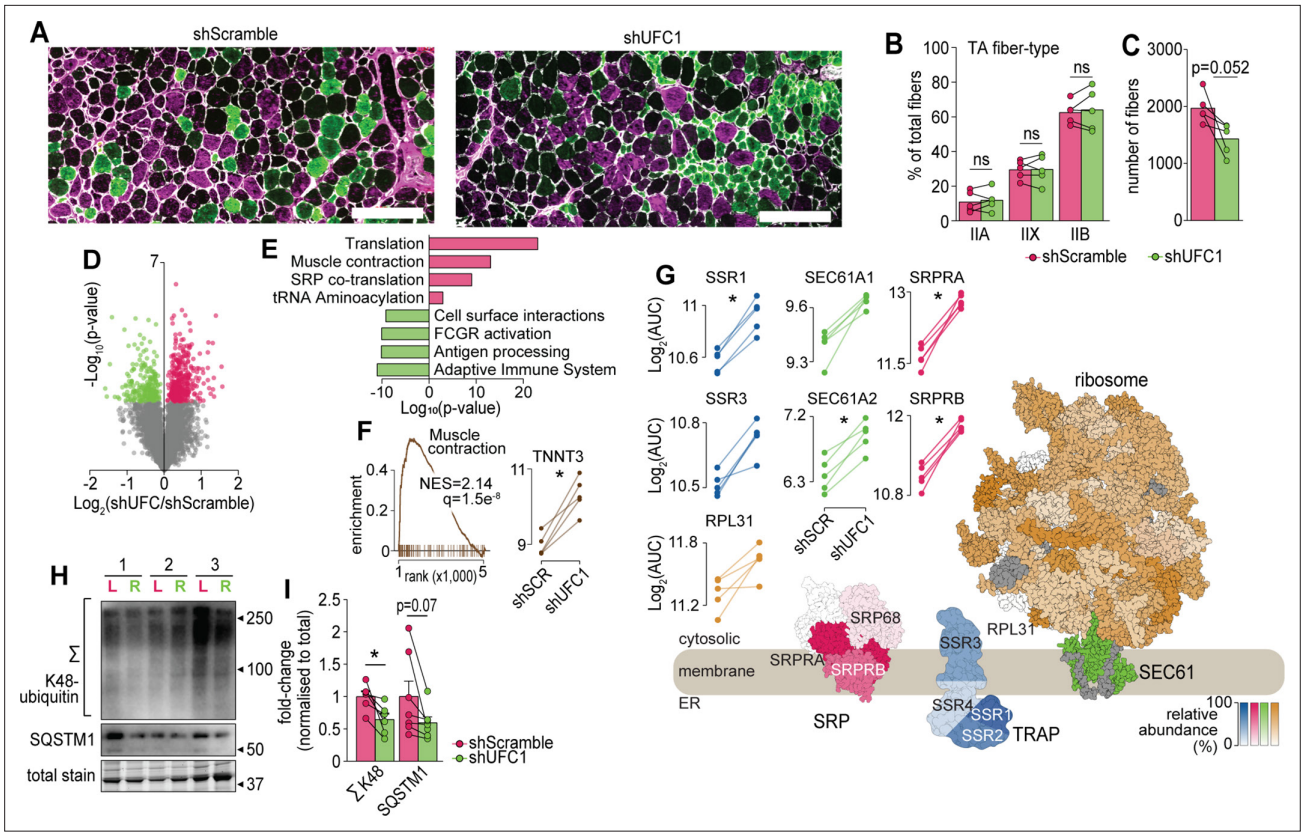


Figure 6. Characterization of skeletal muscles following UFC1 knockdown. (**A**) Representative immunofluorescence microscopy of fiber-type composition in tibialis anterior (TA). Myosin heavy chain isoforms (MYH2, green, type IIa; MYH4, purple, type IIb; MYH1, unstained, type IIx) while laminin is white. Scale bar = 200 µm. (**B**) TA fiber-type distribution, and (**C**) TA total fiber number (n=5). (**D**) Volcano plot and (**E**) gene set enrichment analysis of proteins affected by UFC1 knockdown. (**F**) Enrichment plot of the muscle contraction gene set (REACTOME_MUSCLE_CONTRACTION, MSigDB C2 collection) and paired analysis of TNNT3. (**G**) Knockdown of UFC1 up-regulates the ribosome-SEC61 complex, signal recognition particle, and translocon-associated protein. Protein constituents of each structure were coloured based on the relative increased abundance following shUFC1, where the colours are scaled based on the relative fold change per complex (signal recognition particle [SRP] – red; translocon-associated protein (TRAP) – blue; SEC61 – green; ribosome – yellow/orange, grey – not measured). (**H**) Western blot of extensor digitorum longus (EDL) muscles treated with rAAV6:shScramble (red, left leg (L)) and rAAV:shUFC1 (green, right leg (R)). (**I**) Densitometry of western-blot (n=6). *p/q-value<0.05; (B, C, I): paired Student's t-test; (F–G): paired Student's t-test with Benjamini-Hochberg FDR.

The online version of this article includes the following source data for figure 6:

Source data 1. Zip file containing uncropped western blot image files as Image Lab Documents, tiff files, and a summarized.pdf highlighting the lane identifications, highlighted bands used to create ***Figure 6H***, antibody information and all densitometry results for each individual sample.

(SRP)-mediated translocation to the ER, while down-regulation was observed in immune-related pathways and cell surface interactions (***Figure 6E***). The up-regulation of contractile proteins following knockdown of UFC1 includes fast-type Troponin T (TNNT3), which binds tropomyosin, and the protein with the largest fold change (>2.5-fold) (***Figure 6F***). The most up-regulated proteins in the translational machinery were SRPA/B of the SRP complex, SSR1/2/3/4 of the TRAP complex, and SEC61A/B (***Figure 6G***). We also observed a trend for increased RPL31 which mediates binding of SRP to the ribosome (***Pech et al., 2010***). The core subunits of the proteasome, and atrogenes including TRIM63 and ASB2 were not regulated following knockdown of UFC1 (***Supplementary file 6***). However, the proteasome activating complex PSME1/2 was down-regulated with UFC1 knockdown and western blotting revealed a decrease in K48-linked ubiquitination (***Figure 6H–I***). We also observed a complex regulation of autophagy-associated proteins such as up-regulation of GABARAP and ATG3, and down-regulation of ATG4A, WIPI1, and SQSTM1, the latter which trended to decrease by >1.8-fold with both western blotting and proteomics (***Figure 6H–I*** and ***Supplementary file 6***). Taken together, our data reveal that UFMylation is regulated during atrophy and plays a role in skeletal muscle function via modulating proteostasis mechanisms and contractile proteins.

Discussion

Genetic determinants of diversity in skeletal muscle attributes remain incompletely defined, and thus an opportunity to transform our understanding of muscle biology. We performed a proteomic analysis of skeletal muscle in a genetically diverse mouse panel and integrated the data with a variety of molecular and phenotypic traits that can be browsed at https://muscle.coffeeprot.com to potentially uncover new biology relevant for human disease. We demonstrate the utility of this resource by performing a targeted functional screen in human bioengineered skeletal muscle which revealed UFMylation as a negative regulator of muscle function that was further validated in vivo. Targets of UFMylation include nuclear proteins Histone H4 (***Qin et al., 2019***), MRE11 (***Lee et al., 2021***), and ACS1 (***Yoo et al., 2014***) with the latter particularly relevant to the current study given mutations in ACS1 are associated with muscular atrophy (***Davignon et al., 2016***). However, the most abundant target of UFMylation in mammalian cell culture and validated by two independent studies is RPL26, a ribosomal-associated protein in close proximity to the interaction sites of RPL31 and the SEC61 complex (***Walczak et al., 2019***; ***Wang et al., 2020***). Indeed, the UFM1 E3 ligase, UFL1, is targeted to the ER via UFBP1 (DDRGK1) (***Wang et al., 2020***), and UFMylation is required for ER-phagy (***Liang et al., 2020***) and ER-associated protein degradation (ERAD) (***Walczak et al., 2019***). Mechanistically, RPL26 UFMylation promotes the degradation of a translocation-arrested ER protein and up-regulation may play a role in enhanced secretory flux (***Wang et al., 2020***). Hence, UFMylation plays a positive role in ER stress and cellular protection, and it is potentially paradoxical that reducing UFMylation can improve muscle contractile function. One potential explanation is an acute reduction in UFMylation leads to proteome compensatory remodelling and up-regulation of translational machinery which provide later benefits to muscle function. A limitation of our approach is that the abundance of UFMylation enzymes was measured, but not the degree of UFMylation among proteins. While RPL26 was not significantly regulated following UFC1 knockdown, it is possible that the UFMylation status of RPL26 was altered. The positive enrichment of ribosomal subunits including RPL31, and SRP co-translational factors (SSR1/3, SEC61α, SRPRB) located in close proximity to RPL26 suggests an involvement of UFMylation in this process. In further support of a negative association of UFMylation and muscle function, mutations in UFSP2 have been identified which increase UFMylation and result in musculoskeletal dysplasia (***Watson et al., 2015***; ***Zhang et al., 2020***). Furthermore, the expression of UFC1 is also up-regulated in skeletal myocytes differentiated from induced pluripotent stem cells derived from familial ALS (C9ORF72 mutations) (***Lynch et al., 2021***), and the expression of UBA5, the E1 ligase for UFMylation has recently been associated with Becker muscular dystrophy (***Xu et al., 2021***). Interestingly, genetic ablation of UFMylation in cell culture reduces viral-mediated interferon production and correlates with our observed down-regulation of proteins involved in the innate immune response following UFC1 knockdown suggesting an overall reduction in inflammation (***Snider et al., 2022***).

Ideas and speculation

A major question arising from our data is how does a reduction in UFMylation lead to the up-regulation of contractile proteins? Is it driven by a reduction in K48-linked ubiquitination and potential down-regulation of protein degradation or an increase in protein synthesis despite the major translational machinery up-regulated being associated with ER targeting and SRP/SEC61, or a combination of both? Furthermore, how does the regulation of UFMylation change components of the autophagy system? It is likely that the identification of additional UFMylation substrates may further unravel these mechanisms. Collectively, our systems genetics and functional screening strategy provides a rich resource to further explore mechanisms governing skeletal muscle metabolic function, and we demonstrate its use by identifying UFMylation as an important modification for muscle biology.

Limitations of study

We only performed proteomics on female mice from the HMDP, whereas phenotypic analyses were performed on several separate cohorts using both sexes. Clearly, sex interacts with common genetic variation to influence many outcomes. In this light, an advantage of renewable resources such as the HMDP is to allow the same genetic background to be assayed across multiple studies. Thus, as more data becomes available, which aspects of this study are either female-specific or penetrant across both sexes will be easily addressed. Another potential limitation of our study is that the associations between genetic variants and proteins/phenotypes may be either causal or reactive in nature. That

is, a genetic variant may drive a change in protein abundance which then regulates a phenotype, or the genetic variant changes the phenotype which subsequently changes the abundance of the protein. Furthermore, associations may have pleiotropic affects where a genetic variant may regulate the abundance of multiple proteins which modulate a phenotype either via horizontal or vertical pleiotropy. When interpreting genetic associations, it is also relevant to consider the overall heritability of a given trait. Specifically, broad sense heritability measures can inform the overall confidence in linking genotype to phenotype and inferring genetic interactions with environment and sex (***Andreux et al., 2012***; ***Seldin et al., 2019***; ***Ashbrook et al., 2021***). The genetic repeatability (R) for each trait, as determined using the rptR workflow, is reported in ***Supplementary file 4*** (***Stoffel et al., 2017***). For traits which exhibit a high degree of technical variability such as cardiac function or grip strength, these estimates provide a quantitative metric with which to guide genetic contributions. To further prioritize causal associations focusing on muscle function, we targeted genes containing skeletal muscle *cis*-pQTLs that were also associated to molecular or phenotypic traits and performed a knockdown screen in hμM. Here, our goal was to identify negative regulators of adult muscle function and we utilized rAAV6:shRNA vectors applied to the micro-muscles post-differentiation to model mature skeletal muscle. Hence, the results of our screen may not be relevant for the study of myogenesis and muscle regeneration. For the first time, we show that UFMylation is regulated in an atrophy model of ALS motor neuron disease and further studies are required to investigate the regulation of UFMylation in other diseases of muscle wasting such as dystrophy, cachexia, and sarcopenia. Furthermore, additional studies are warranted to investigate if the regulation of UFMylation can provide functional benefits during atrophy. It is also important to note that we used rAAV6 vectors which provides high tropism to terminally differentiated myofibers (***Blankinship et al., 2004***) and hence, additional methodologies are required to investigate the role of UFMylation in other cell types including defective neuronal cell populations.

Materials and methods

HMDP animals

All mice were from The Jackson Laboratory and were subsequently bred and housed at University of California, Los Angeles, to generate offspring used in this study as previously described (***Parks et al., 2015***; ***Parks et al., 2013***). Only female mice were used and housed at 22°C (±1°C) on a 12 hr light/dark cycle and ad libitum access to food and water with a chow diet (Ralston Purina Company – 5001) until 8–10 weeks of age before being fasted for 14 hr in a fresh cage. Animals were anaesthetized, exsanguinated, and gastrocnemius muscles immediately removed and snap-frozen. All protocols for these studies were approved by the Institutional Care and Use Committee (IACUC) at University of California, Los Angeles. A list of mice used in this study is shown in the 'Key resources table'.

Proteomics sample preparation

Muscle tissue from the HMDP were lysed in 6 M guanidine HCL (Sigma; #G4505), 100 mM Tris pH 8.5 containing 10 mM tris(2-carboxyethyl)phosphine (Sigma; #75259) and 40 mM 2-chloroacetamide (Sigma; #22790) by tip-probe sonication. The lysate was heated at 95°C for 5 min and centrifuged at 20,000 × *g* for 10 min at 4°C. The supernatant was diluted 1:1 with water and precipitated overnight with five volumes of acetone at –20°C. The lysate was centrifuged at 4000 × *g* for 5 min at 4°C and the protein pellet was washed with 80% acetone. The lysate was centrifuged at 4000 × *g* for 5 min at 4°C and the protein pellet was resuspended in Digestion Buffer (10% 2,2,2-trifluoroethanol [Sigma; #96924]) in 100 mM HEPES pH 7.5. Protein was quantified with BCA (Thermo Fisher Scientific) and normalized in Digestion Buffer to a final concentration of 2 μg/μl. Protein was digested with sequencing grade trypsin (Sigma; #T6567) and sequencing grade LysC (Wako; #129-02541) at a 1:50 enzyme:substrate ratio overnight at 37°C with shaking at 2000× rpm. Eight μg of peptide was directly labelled with 32 μg of 10-plex TMT (lot #QB211242) in 20 μl at a final concentration of 50% acetonitrile for 1.5 hr at room temperature. The reaction was de-acylated with a final concentration of 0.3% (w/v) hydroxylamine and quenched with a final concentration of 1% trifluoroacetic acid (TFA). Each 10-plex experiment contained nine different strains with a tenth reference label (131 isobaric label) made up of the same peptide digest from pooled mix of C57BL/6J muscles. The sample identity and labelling channels have been uploaded as a table with the.raw proteomic data to the PRIDE

ProteomeXchange (see Data availability section). Following labelling, the peptides from each of the 18 TMT 10-plex batches were pooled and purified directly by styrene divinylbenzene reversed-phase sulfonate (SDB-RPS) microcolumns, washed with 99% isopropanol containing 1% TFA and eluted with 80% acetonitrile containing 2% ammonium hydroxide followed by vacuum concentration. Peptides were resuspended in 2% acetonitrile containing 0.1% TFA and 30 µg of peptide was fractionated on an in-house fabricated 25 cm × 320 µm column packed with C18BEH particles (3 µm, Waters). Peptides were separated on a gradient of 0–30% acetonitrile containing 10 mM ammonium formate (pH 7.9) over 60 min at 6 µl/min using an Agilent 1260 HPLC and detection at 210 nm with a total of 48 fractions collected and concatenated down to 12 fractions. Skeletal muscle micro-muscles were lysed in 4% sodium deoxycholate in 100 mM Tris pH 8.5 containing 10 mM tris(2-carboxyethyl)phosphine and 40 mM 2-chloroacetamide by tip-probe sonication. The lysate was heated at 95°C for 5 min and centrifuged at 18,000 × *g* for 10 min at 4°C. Protein was digested with 0.2 µg of sequencing grade trypsin and 0.2 µg of sequencing grade LysC overnight at 37°C. Peptides were first diluted with 100% isopropanol, mixed and then acidified with TFA to a final concentration of 50% isopropanol, 0.1% TFA. Peptides were desalted with SDB-RPS microcolumns, washed with 99% isopropanol containing 1% TFA and eluted with 80% acetonitrile containing 2% ammonium hydroxide followed by vacuum concentration. Peptides were resuspended in 2% acetonitrile containing 0.1% TFA and a 5% aliquot of each sample pooled and fractionated into 12 fractions as described above to generate a spectral library. The sample and MS file identifies have been uploaded as a table with the .raw proteomic data to the PRIDE ProteomeXchange (see Data availability section). Muscle tissue from rAAV6-treated mice were processed using the identical procedure described above with only minor modifications including the 10-plex TMT lot #WC306775 was used. The sample identity and labelling channels have been uploaded as a table with the .raw proteomic data to the PRIDE ProteomeXchange (see Data availability section). Peptides were fractionated using a separate in-house fabricated column with identical dimensions and particles, but a Dionex 3500 HPLC was used with the same detection at 210 nm and a total of 48 fractions collected and concatenated down to 12 fractions.

Mass spectrometry and data processing

Peptide fractions from skeletal muscle of the HMDP were resuspended in 2% acetonitrile containing 0.1% TFA and analyzed on a Dionex ultra-high pressure liquid chromatography system coupled to an Orbitrap Lumos mass spectrometer. Briefly, peptides were separated on 40 cm × 75 µm column containing 1.9 um C18AQ Reprosil particles on a linear gradient of 2–30% acetonitrile over 2 hr. Electrospray ionization was performed at 2.3 kV with 40% RF lens and positively charged peptides detected via a full-scan MS (350–1550 m/z, 1e6 AGC, 60 K resolution, 50 ms injection time) followed by data-dependent MS/MS analysis performed with CID of 35% normalized collision energy (NCE) (rapid scan rate, 2e4 AGC, 50 ms injection time, 10 ms activation time, 0.7 m/z isolation) of the top 10 most abundant peptides. Synchronous-precursor selection with MS3 (SPS-MS3) analysis was enabled with HCD of 60 NCE (100–500 m/z, 50 K resolution, 1e5 AGC, 105 ms injection time) (***McAlister et al., 2014***). Dynamic exclusion was enabled for 60 s. Data were processed with Proteome Discoverer v2.3 and searched against the Mouse UniProt database (November 2018) using SEQUEST (***Eng et al., 1994***). The precursor MS tolerance was set to 20 ppm and the MS/MS tolerance was set to 0.8 Da with a maximum of two miss-cleavage. The peptides were searched with oxidation of methionine set as variable modification, and TMT on peptide N-terminus/lysine and carbamidomethylation of cysteine set as a fixed modification. All data was searched as a single batch and the peptide spectral matches (PSMs) of each database search filtered to 1% FDR using a target/decoy approach with Percolator (***Käll et al., 2007***). The filtered PSMs from each database search were grouped and q-values generated at the peptide level with the Qvality algorithm (***Käll et al., 2009***). Finally, the grouped peptide data was further filtered to 1% protein FDR using Protein Validator. Quantification was performed with the reporter ion quantification node for TMT quantification based on MS3 scans in Proteome Discoverer. TMT precision was set to 20 ppm and corrected for isotopic impurities. Only spectra with <50% co-isolation interference were used for quantification with an average signal-to-noise filter of >10. The data was filtered to retain Master proteins that were measured in at least 50 mice. Peptides from skeletal muscle micro-muscles were resuspended in 2% acetonitrile containing 0.1% TFA and analyzed on a Dionex ultra-high pressure liquid chromatography system coupled to an Orbitrap Exploris 480 mass spectrometer. Briefly, peptides were separated on 40 cm × 75 µm column containing 1.9 µm C18AQ

Reprosil particles on a linear gradient of 2–30% acetonitrile over 70 min. Electrospray ionization was performed at 1.9 kV with 40% RF lens and positively charged peptides detected via a full-scan MS (350–950 m/z, 2.5e6 AGC, 60 K resolution, 50 ms injection time) followed by data-independent MS/MS analysis performed with HCD of 28% NCE (16 m/z isolation, 38 windows with 1 m/z overlap, 2e6 AGC, 30 K resolution, auto injection time). The pooled and fractionated samples were used to generate a spectral library using data-dependent acquisition acquired in the same batch using the identical liquid chromatography and column. Each of the 12 fractions were injected twice using two-step gas-phase fraction to generate a spectral library. A full-scan MS from 350 to 651 m/z or 650 to 950 m/z was performed for each of the two injections (2.5e6 AGC, 60 K resolution, 50 ms injection time) followed by data-dependent MS/MS analysis performed with HCD of 28% NCE (1.2 m/z isolation, 5e4 AGC, 15 K resolution, auto injection time). Data were processed with Spectronaut v15.0.210615.50606 and the DDA data were searched against the Human UniProt database (June 2021) using Pulsar. The minimum peptide length set to seven amino acids with specific trypsin cleavage and search criteria included oxidation of methionine and protein N-terminal acetylation set as variable modifications, and carbamidomethylation set as a fixed modification. Data were filtered to 1% FDR at the peptide and protein level (q-value cut-off <0.01). The DIA data were searched within Spectronaut using the project-specific library and peptide quantification was performed at MS2 level using three to six fragment ions which included automated interference fragment ion removal as previously described (***Bruderer et al., 2015***). Dynamic mass MS1 and MS2 mass tolerance was enabled, and local (non-linear) regression was performed for retention time calibration. A dynamic extracted ion chromatogram window size was performed, and protein quantification performed with weighted peptide average values. Peptide fractions from skeletal muscle treated with rAAV6 were analyzed as described above for muscle of the HMDP with minor modifications. Briefly, peptides were separated using a Dionex ultra-high pressure liquid chromatography system using the identical chromatography configuration, but detection was achieved with an Orbitrap Eclipse mass spectrometer. Electrospray ionization was performed at 1.9 kV with 30% RF lens and positively charged peptides detected via a full-scan MS (350–1550 m/z, 2e6 AGC, 60 K resolution, 50 ms injection time) followed by data-dependent MS/MS analysis performed with HCD of 36% NCE (1e5 AGC, 86 ms injection time, 0.7 m/z isolation) with a 2.5 s cycle time and dynamic exclusion was enabled for 60 s. Data were processed with Proteome Discoverer v2.3 and searched against the Mouse UniProt database (February 2022) using SEQUEST (***Eng et al., 1994***). The precursor MS tolerance was set to 20 ppm and the MS/MS tolerance was set to 0.02 Da with a maximum of two miss-cleavage. The peptides were searched with oxidation of methionine set as variable modification, and TMT on peptide N-terminus/lysine and carbamidomethylation of cysteine set as a fixed modification. All data was searched as a single batch and the PSMs of each database search filtered to 1% FDR using a target/decoy approach with Percolator (***Käll et al., 2007***). The filtered PSMs from each database search were grouped and q-values generated at the peptide level with the Qvality algorithm (***Käll et al., 2009***). Finally, the grouped peptide data was further filtered to 1% protein FDR using Protein Validator. Quantification was performed with the reporter ion quantification node for TMT quantification based on MS2 scans in Proteome Discoverer. TMT precision was set to 20 ppm and corrected for isotopic impurities. Only spectra with <50% co-isolation interference were used for quantification with an average signal-to-noise filter of >10. The data was filtered to retain Master proteins that were measured in at least 50 mice.

Protein-protein and protein-trait correlations

Data analyses were performed using R (version 4.1.1). Circos plots and circular dendrograms were created using the circlize R package (***Gu et al., 2014***). CoffeeProt was used to assess protein-protein correlations and to produce network plots (***Molendijk et al., 2021b***). TeaProt was used to perform functional enrichment analyses (***Molendijk et al., 2022***). The packages Gviz and TxDb.Mmusculus.UCSC.mm10.knownGene were used to produce genomic tracks (***Hahne and Ivanek, 2016***). Biweight midcorrelation (bicor) was performed using the WGCNA package (***Langfelder and Horvath, 2008***). Both the proteomics and trait datasets were summarized (mean) at the strain level prior to performing protein-trait correlations. The correlation coefficient and p-value were reported for each protein-trait pair, followed by the calculation of adjusted p-values (q-value) using the Benjamini-Hochberg procedure. Orthogonal partial least-squares (OPLS) modelling was performed using the ropls package (***Thévenot et al., 2015***) where the summarized proteomics data represents the input numerical matrix

(x), and the measurement of a trait of interest represents the response to be modelled (y). Models were created with a single predictive component (predI) and a single orthogonal component (orthoI). For OPLS and protein-trait correlations, proteomic data was quantile normalized and the biological replicates within each strain averaged. MuscleProt can be used to export the metrics, loadings, and scores tables from each model.

Protein and molecular/phenotypic quantitative trait locus (QTL) mapping The identification of SNPs associated to protein abundance was performed using an efficient mixed-model association (fast-lmm) (***Kang et al., 2008***) as described below where the model was adjusted for population structure (***Flint and Eskin, 2012***):

$$y = 1n\mu + x\beta + u + e \quad (1)$$

in which n is the number of individuals; μ is the mean; β is the allele effect of the SNP; x is the (n×1) vector of observed genotypes of the SNP. This model takes population structure into account, as u is the random effects due to genetic relatedness with var.(u)=σ2uK, and e denotes the random noise with var.(e)=σ2eI. Here, K indicates the identity-by-state kinship matrix estimated using all SNPs; I represents the (n×n) identity matrix; and 1n is the (n×1) vector of ones. σ2u and σ2e were estimated using restricted maximum likelihood and computed p-values using the standard F-test to test the null hypothesis in which β = 0. Genome-wide significance threshold and genome-wide association mapping were determined as the family-wise error rate as the probability of observing one or more false positives across all SNPs for a given phenotype. To correct for false discovery, q-values were estimated from the distribution of p-values using the linear mixed model from the R package 'q value'. Significance was calculated at q-value <0.1 (*cis*-pQTL = ±10 Mb of the gene, approximated local adjusted $p<1 \times 10^{-4}$, and *trans*-pQTL=approximated global adjusted $p<5 \times 10^{-8}$) as described previously (***Chick et al., 2016***; ***Parker et al., 2019***). SNP locations and variant effects were retrieved from the Ensembl Variant database (release 102, GRCm38). Molecular/phenotypicQTL data were obtained from previously published studies and processed using fast-lmm as described above (***Ghazalpour et al., 2012***; ***Ghazalpour et al., 2014***; ***Parks et al., 2015***; ***Norheim et al., 2017***; ***Rau et al., 2017***; ***Norheim et al., 2019***; ***Parker et al., 2019***; ***Tuominen et al., 2021***; ***Norheim et al., 2021***). A summary of phenotypic data and sources is described in ***Supplementary file 4***.

Structural biology

The PROVEAN (Protein Variant Effect Analyzer) tool was used to predict the functional effects of missense mutations in our dataset (***Choi and Chan, 2015***). The Colabfold platform utilizing Alphafold to predict the protein structures was used generate the EPHX1 structure (***Jumper et al., 2021***; ***Mirdita et al., 2022***). The mmseq2 method was used for the multiple sequence alignment step. Alphafold models were ranked by pLDDT and further assessed using the sequence coverage, sequence identity, and predicted alignment error metrics generated in Colabfold (***Mirdita et al., 2022***). Generated models were visualized using PyMol. The PROVEAN (***Choi et al., 2012***) and PolyPhen-2 (***Adzhubei et al., 2013***) servers were used to predict the functional effect of mutations. FoldX was used to determine the effects of mutations on protein stability (***Delgado et al., 2019***). DynaMut was used to determine the protein flexibility as a result of mutation and predict the interactomic interactions (***Rodrigues et al., 2018***). Mol* (Molstar) was used to display PDB format structures and colour protein complex constituents according to relative fold change values (***Sehnal et al., 2021***). Structural models for the SRP (7OBQ) and ribosome-SEC61 complex (3J7R) were retrieved from the RCSB Protein Data Bank (PDB). Models were edited to remove RNA and small molecule entities. The diagram of the TRAP complex was based on the electron microscopy density model EMD-3068. The R functions colorRampPalette and col2rgb were used to generate gradients of hex colour codes and corresponding RGB colour values used in Mol*.

Database

SNP locations and variant effects were retrieved from the mus_musculus_incl_consequences.vcf.gz (03 August, 2020) file in the Ensembl Variant database (release 102) (***McLaren et al., 2016***). Gene information and UniProt accession mappings were retrieved from the Ensembl project, release 102 (***Howe et al., 2021***) Mus_musculus.GRCm38.102.gtf.gz (27 October 2020) and Mus_musculus.GRCm38.102.

uniprot.tsv.gz (26 October 2020). UK Biobank exome sequencing data was accessed through the Genebass webserver (***Karczewski et al., 2022***).

shRNA:rAAV6 production

All shRNA sequences are shown in ***Supplementary file 7*** and were designed using SplashRNA including optimized stem-loop design (5'TAGTGAAGCCACAGATGTA) as previously described (***Pelossof et al., 2017***). The scramble shRNA sequence was 5'GATCGAATGTGTACTTCGA and selected based on a previously described screen for low toxicity (***Grimm et al., 2006***). All AAV vectors were produced by the Vector and Genome Engineering Facility (VGEF) at Children's Medical Research Institute (CMRI). Vectors were produced by standard transient transfection of 5×15 cm plates of HEK293 (ATCC# CRL-1573) cells using PEI (polyethylenimine, PolyPlus, Cat# 115-100) with a 1:1:2 molar ratio of pTransgene:pRep2CapXHelper:pAd5Helper. Vectors were purified using iodixanol gradient ultracentrifugation as previously described (***Khan et al., 2011***). Amicon Ultra-4 Centrifuge Filter Units (Ultracel-100 kDa membrane, EMD Millipore, Cat# UFC810024) were used to perform buffer exchange (phosphate-buffered saline [PBS, Gibco, Cat# 14190], 50 mM NaCl [Sigma-Aldrich, Cat# S5150-1L], 0.001%, Pluronic F68 [v/v] [Gibco, Cat# 24040]) and the final concentration step. Iodixanol-purified AAVs were quantified using droplet digital PCR (ddPCR [Bio-Rad, Berkeley]) using QX200 ddPCR EvaGreen Supermix (Cat# 1864034; Bio-Rad) with eGFP primers (5' TCAAGATCCGCC ACAACATC and 5' TTCTCGTTGGGGTCTTTGCT). All cell stocks were regularly checked for absence of mycoplasma with the Mycoplasma Detection Kit (Jena Bioscience; # PP-401).

hμM production and functional assessment

hμMs were generated as described previously (***Mills et al., 2019***). Briefly, primary human skeletal muscle myoblasts from a male, 20 years of age (Lonza, lot #18TL269121) were mixed with collagen I gel to make a 3.5 μl final solution containing 3.3 mg/mL collagen I and 22% (v/v) Matrigel (52,500 cells per hμM). The bovine acid-solubilized collagen I (Devro) was first salt balanced and pH neutralized using 10× DMEM and 0.1 M NaOH, respectively, prior to mixing with Matrigel and then combined with the cells. The mixture was prepared on ice and pipetted into the cell-culture inserts (***Mills et al., 2017***). The mixture was then gelled at 37°C for 30 min. After 1 day of formation, media was switched to containing MEM α (Thermo Fisher Scientific) with 1% P/S (Thermo Fisher Scientific), and 1% B-27 supplement (Thermo Fisher Scientific), with 10 μM DAPT (Stem Cell Technologies) and 1 μM Dabrafenib (Stem Cell Technologies) to induce differentiation. On day 8, media was switched to maintenance media containing MEM α (Thermo Fisher Scientific) with 1% P/S (Thermo Fisher Scientific), and 1% B-27 supplement (Thermo Fisher Scientific). Media was changed every 2–3 days. On day 12, hμMs were treated with AAV6 encoding shRNA for genes of interest or a scrambled control at 6e7 vg/hμM. After 72 hr of treatment, hμM were analysed for their function via electrically stimulation at 10 or 20 Hz; 5 ms square pulses with 20 mA current using a Panlab/Harvard Apparatus Digital Stimulator. During stimulation, a Leica DMi8 inverted high content Imager was used to capture a 5 or 15 s time-lapse of each hμM contracting in real time at 37°C. Pole deflection was used to approximate the force of contraction as per ***Mills et al., 2017***. Custom batch processing files were written in Matlab R2013a (Mathworks) to convert the stacked TIFF files to AVI, track the pole movement (using vision. PointTracker), produce a force-time figure, and export the batch data to an Excel (Microsoft) spreadsheet. hμM max force was assessed as the peak force of contraction during a 20 Hz stimulation for 1 s. Whilst, hμM endurance/fatigue was assessed as the change in contraction force in response to a 10 Hz stimulation for 10 s (force at the end of stimulation [10 s] compared to peak force).

Mouse housing and rAAV6 intramuscular injection

All mouse experiments were approved by The University of Melbourne Animal Ethics Committee (AEC ID1914940) and conformed to the National Health and Medical Research Council of Australia guidelines regarding the care and use of experimental animals. C57BL/6J mice (JAX 000664) were obtained from Animal Resource Centre (WA, Australia). Mice were housed at 22°C (±1°C) in groups of five/cage and maintained on a Standard Chow diet (Specialty Feeds, Australia) with a 12 hr light/dark cycle and ad libitum access to food and water. For intramuscular injections of rAAV6, mice were anaesthetized with isoflurane (4% in oxygen at 1 l/min) and then transferred to a dissecting microscope stage with heat pad and isoflurane inhalation nose piece (2% in oxygen at 1 l/min). Unconsciousness

was assessed via the lack of leg and optical reflexes for at least 1 min to ensure head position does not affect normal breathing. Mice received subcutaneous analgesic injection of meloxicam between the shoulder blades (5 mg/kg) and the surface of the hindlimbs were sterilized with 80% ethanol. The TA/EDL muscles were injected with 2×1010 vector genomes/30 µl of rAAV6 using a 32 G needle. Mice were returned to cages and body weights monitored daily for the first 3 days and then weekly.

Ex vivo muscle function testing

Mice were anaesthetized in the non-fasted state with isoflurane (4% in oxygen at 1 l/min; muscle contraction experiments) and transferred to a dissecting microscope stage with isoflurane inhalation nose piece (2% in oxygen at 1 l/min). Depth of anaesthesia was assessed via the lack of leg and optical reflexes for at least 1 min to ensure head position did not affect normal breathing. After confirming anaesthesia, skin from the hind legs was removed, and EDL muscles were sutured using 5.0 braided suture at both proximal and distal ends at the tendomuscular junction. Muscles were excised and incubated in Modified Krebs Buffer (116 mM NaCl, 4.6 mM KCl, 1.16 mM KH_2PO_4, 25.3 mM $NaHCO_3$, 2.5 mM $CaCl_2$, 1.16 mM $MgSO_4$) in a myograph (DMT, Denmark; #820 MS) at 30°C with constant gentle bubbling of 5% medical carbon dioxide in oxygen. For contractile function experiments, a DMT CS4 stimulator was used to deliver 0.2 ms supramaximal (26 V) pulses via stimulation electrodes (DMT 300145) placed over the mid-belly of the muscle. Successive twitch stimulations, with at least 30 s rest, were used to determine muscle optimal length by very carefully stretching the muscle to optimum length when maximal twitch force was obtained. The frequency-force relationship was determined by stimulating muscles at different frequencies (10–200 Hz, 350 ms duration) with muscles rested for 2 min between stimuli to avoid fatigue. Where appropriate, force values were normalized to muscle CSA (i.e. to calculate specific force) by diving muscle mass by the product of muscle length and muscle density (1.06 mg/mm^3). A PowerLab 8/35 unit (ADInstruments) was used to digitize all force recordings and the Peak Parameters module in LabChart Pro (v8.1.16, ADInstruments) used for analysis of force responses.

ALS model mice

SOD1G37R mice were sourced from The Jackson Laboratory and were subsequently bred and housed at the University of Melbourne, Melbourne, as previously described (***Roberts et al., 2014***) to generate transgenic mice and non-transgenic littermates used in this study. Animals were sacrificed at 25 weeks of age and tissues collected using previously described protocols in the non-fasted state (***Hilton et al., 2017***). All studies involving the use of SOD1G37R mice and non-transgenic littermates were approved by a University of Melbourne Animal Experimentation Ethics Committee (approval #2015124) and conformed with guidelines of the Australian National Health and Medical Research Council.

Western blotting

Proteins were separated on NuPAGE 4–12% Bis-Tris protein gels (Thermo Fisher Scientific) in MOPS SDS Running Buffer at 160 V for 1 hr at room temperature. The protein was transferred to PVDF membranes (Millipore; #IPFL00010) in NuPAGE Transfer Buffer at 20 V for 1 hr at room temperature and blocked with 5% skim milk in Tris-buffered saline containing 0.1% Tween-20 (TBST) for at least 30 min at room temperature with gentle shaking. The membranes were incubated overnight in primary antibody with 5% BSA in TBST with gentle shaking at 4°C and washed three times in TBST at room temperature. Anti-UFC1 (EPR15014-102, ab189252), anti-UFSP2 (EP13424-49, ab192597), and anti-UFM1 (EPR4264(2), ab109305) were ordered from Abcam. The membranes were incubated with HRP-secondary antibody in 5% skim milk in TBST for 45 min at room temperature and washed three times with TBST. HRP-Donkey Anti-Rabbit (711-035-152, RRID:AB_10015282) was ordered from Jackson ImmunoResearch. Protein was visualized with Immobilon Western Chemiluminescent HRP Substrate (Millipore; #WBKLS0500) and imaged on a ChemiDoc (Bio-Rad). Densitometry was performed in ImageJ (***Schneider et al., 2012***).

Immunostaining and microscopy

TA muscles were embedded in optimal cutting temperature compound (Tissue-Tek) and frozen in 2-methylbutane (Sigma-Aldrich; #320404) cooled in liquid nitrogen. Eight µm serial transverse sections were cut from the middle of the TA in a cryostat then mounted on uncoated, pre-cleaned glass slides.

Sections were fixed in 4% PFA for 10 min then blocked for 1 hr at room temperature in goat serum solution (5% goat serum, 2% BSA, 0.1% Triton in PBS). Sections were then air-dried and incubated in a humidity chamber overnight at room temperature in a primary antibody cocktail solution comprised of 1:25 SC-71 (DSHB; mouse IgG1), 1:10 BF-F3 (DSHB; mouse IgM), and 1:250 Laminin (Sigma; L9393; rabbit IgG) in 0.05% PBSTween-20 to differentiate MHC type I, MHC type IIA, MHC type IIB fibers, and laminin regions, respectively. All non-reactive fibers were assumed to be MHC type IIX fibers. After primary incubation, sections were washed three times with PBS then incubated in a humidity chamber for 1.5 hr at room temperature in a secondary antibody cocktail solution comprised of 1:250 Alexa Fluor 555 (Goat anti-mouse IgG1), 1:250 Alexa Fluor 350 (Goat anti-mouse IgM), and 1:250 Alexa Fluor 647 (Goat anti-rabbit IgG) in 0.05% PBSTween-20. After secondary incubation, sections were washed three times with PBS, air-dried and mounted with Fluoro-Gel (ProSciTech IM030) under a coverslip. Fluorescence imaging of the whole section was captured with an upright microscope with a camera (Axio Imager M2, Carl Zeiss, Wrek, Göttingen, Germany). Pseudo colouring of each fluorescence channel was performed with ZEN 3.3 (Blue edition, Carl Zeiss, Wrek, Göttingen, Germany). Quantification was performed with Fiji (ImageJ; NIH).

MuscleProt web server implementation

MuscleProt was developed using the R programming language for the backend and relies on the shiny package for the web server front-end in addition to HTML, CSS, and JavaScript. The WGCNA package is used to perform biweight midcorrelation (bicor) analyses (***Langfelder and Horvath, 2008***) and interactive network plots are created using networkD3. Multivariate analyses are performed using the ropls package (***Thévenot et al., 2015***). Tables are created using the DT package and all other visualizations use a combination of ggplot2 and plotly. MuscleProt is deployed on the Melbourne Research Cloud, running Ubuntu 18.04 and utilizing hypervisors built on AMD EPYC 2 (base CPU clock speed 2.0 GHz, burst clock speed 3.35 GHz). The Melbourne Research Cloud is based on the OpenStack open-source cloud platform.

Statistical analysis

Statistics on densitometry, muscle function, and histology were performed in GraphPad Prism (Version 9.0.0). T-tests (unpaired for ALS analysis or paired for rAAV6 analysis) or two-way ANOVA (rAAV6 analysis) were used with a significance level of $p<0.05$. All mentions of sample size (n) refer to biological replicates.

Acknowledgements

We would like to thank the Melbourne Mass Spectrometry and Proteomics Facility in The Bio21 Molecular Science and Biotechnology Institute at The University of Melbourne, and the Sydney Mass Spectrometry Facility in the Charles Perkins Centre at The University of Sydney for mass spectrometry support. We also thank the Melbourne Mouse Metabolic Phenotyping Platform at the University of Melbourne, and facilities at the University of California, Los Angles for mouse support. This research was supported by use of the Nectar Research Cloud and by the University of Melbourne Research Platform Services. The Nectar Research Cloud is a collaborative Australian research platform supported by the National Collaborative Research Infrastructure Strategy.

Additional information

Competing interests

David E James: Senior editor, *eLife*. The other authors declare that no competing interests exist.

Funding

Funder	Grant reference number	Author
National Health and Medical Research Council	APP1184363	Karen Reue Marcus M Seldin Benjamin L Parker
National Health and Medical Research Council	APP2009642	Benjamin L Parker
National Health and Medical Research Council	APP2013189	Richard J Mills
National Health and Medical Research Council	APP1156562	Paul Gregorevic Benjamin L Parker
National Institutes of Health	HL138193	Marcus M Seldin
National Institutes of Health	DK130640	Marcus M Seldin
National Institutes of Health	DK097771	Marcus M Seldin
National Institutes of Health	GM115318	Karen Reue
National Institutes of Health	AG070959	Aldons J Lusis
National Institutes of Health	HL147883	Aldons J Lusis
National Institutes of Health	DK117850	Aldons J Lusis
Weary Dunlop Foundation		Benjamin L Parker
The ALS Association	21-DDC-574	Paul Gregorevic Peter J Crouch

The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

Author contributions

Jeffrey Molendijk, Software, Formal analysis, Visualization, Writing – original draft, Writing – review and editing; Ronnie Blazev, Formal analysis, Investigation, Methodology, Writing – review and editing; Richard J Mills, Formal analysis, Investigation, Writing – review and editing; Yaan-Kit Ng, Formal analysis, Writing – review and editing; Kevin I Watt, Peter J Crouch, Investigation, Writing – review and editing; Daryn Chau, Paul Gregorevic, James BW Hilton, Leszek Lisowski, Peixiang Zhang, Karen Reue, James E Hudson, Methodology, Writing – review and editing; Aldons J Lusis, David E James, Conceptualization, Writing – review and editing; Marcus M Seldin, Conceptualization, Formal analysis, Investigation, Methodology, Writing – review and editing; Benjamin L Parker, Conceptualization, Formal analysis, Investigation, Visualization, Methodology, Writing – original draft, Writing – review and editing

Author ORCIDs

Jeffrey Molendijk http://orcid.org/0000-0001-6575-504X
Peter J Crouch http://orcid.org/0000-0002-7777-4747
David E James http://orcid.org/0000-0001-5946-5257
Marcus M Seldin http://orcid.org/0000-0001-8026-4759
Benjamin L Parker http://orcid.org/0000-0003-1818-2183

Ethics

All rAAV6 intramuscular injection mouse experiments were approved by The University of Melbourne Animal Ethics Committee (AEC ID1914940) and conformed to the National Health and Medical Research Council of Australia guidelines regarding the care and use of experimental animals. All studies involving the use of SOD1G37R mice and non-transgenic littermates were approved by a University of Melbourne Animal Experimentation Ethics Committee (approval #2015124) and conformed with guidelines of the Australian National Health and Medical Research Council.

Decision letter and Author response

Decision letter https://doi.org/10.7554/eLife.82951.sa1
Author response https://doi.org/10.7554/eLife.82951.sa2

Additional files

Supplementary files

- Supplementary file 1. Hybrid Mouse Diversity Panel (HMDP) skeletal muscle proteomics. Proteomics data of gastrocnemius muscle displaying quantification ratios of each sample compared to its corresponding pooled tandem mass tag (TMT) control. PEP: posterior error probability. Related to ***Figures 1–3***, ***Figure 1—figure supplement 1***.
- Supplementary file 2. Hybrid Mouse Diversity Panel (HMDP) skeletal muscle protein-quantitative trait loci (pQTLs). Protein quantitative trait loci from 161 HMDP cohort mice. Table contains *cis*-pQTLs ($p < 1\times10^{-4}$) and *trans*-pQTLs ($p < 5 \times 10^{-8}$), including genomic locations of the single nucleotide polymorphism (SNP) and associated gene. pQTLs are annotated with proxy (*cis*/*trans*), intragenic variants, known LD blocks, variant effect, variant impact, and target screen prioritization columns. Related to ***Figures 1–3***, and ***Figure 1—figure supplement 2***.
- Supplementary file 3. Hybrid Mouse Diversity Panel (HMDP) skeletal muscle pairwise protein-protein correlations. Protein-protein correlation as determined using biweight midcorrelation. p-Values and q-values derived using the Benjamini-Hochberg procedure, and only positive correlations are shown (cor > 0.3 and q < 0.05). Correlated protein pairs are annotated with protein:protein interactions from the CORUM and BioPlex databases, and subcellar. Related to ***Figures 1 and 2***.
- Supplementary file 4. Hybrid Mouse Diversity Panel (HMDP) molecular or phenotypic traits. Summary of traits integrated into the current study including Pubmed ID sources. Related to ***Figures 2 and 3***.
- Supplementary file 5. Hybrid Mouse Diversity Panel (HMDP) molecular or phenotypic quantitative trait loci (QTLs). Table contains QTLs ($p < 1\times10^{-4}$) including chromosome, genomic location, including Pubmed ID sources. Related to ***Figures 2 and 3***.
- Supplementary file 6. Proteomics of skeletal muscle treated with either rAAV6:shScramble or AAV6:shUFC1. Proteomics of extensor digitorum long (EDL) muscles displaying tandem mass tag (TMT) quantification expressed as Log2(area under the curve). Significance was calculated using paired Student's t-test with Benjamini-Hochberg FDR. PEP: posterior error probability. Related to ***Figure 4***.
- Supplementary file 7. shRNA sequences used in the human micro-muscle screen and mouse shUFC1 experiments. Related to ***Figure 4***.
- MDAR checklist
- Source data 1. Genebass datasets for UFC1, EPHX1, DUSP23, NIT1, MPZ, BPNT1 and PCP4L1.

Data availability

The proteomics data generated in this study are deposited to the ProteomeXchange Consortium via the PRIDE (***Perez-Riverol et al., 2019***) under the identifiers PXD032729, PXD034913 and PXD035170. The code used for downstream analysis of proteomic data can be found at: https://github.com/JeffreyMolendijk/skeletal_muscle, (copy archived at swh:1:rev:9311d7bfb59979d80e18612879631dc78f2f0902; ***Molendijk, 2022***). The following entries

from Genebass (***Karczewski et al., 2022***) were used: UFC1, EPHX1, DUSP23, NIT1, MPZ, BPNT1 and PCP4L1.

The following datasets were generated:

Author(s)	Year	Dataset title	Dataset URL	Database and Identifier
Parker BL	2023	Proteomic analysis of a targeted functional genomic screen in human skeletal muscle organoids	https://www.ebi.ac.uk/pride/archive/projects/PXD034913	PRIDE, PXD034913
Parker BL	2023	Proteomic analysis of UFC1 knockdown in mouse skeletal muscle	https://www.ebi.ac.uk/pride/archive/projects/PXD035170	PRIDE, PXD035170
Parker BL	2023	Proteomic analysis of skeletal muscle from the Hybrid Mouse Diversity Panel	https://www.ebi.ac.uk/pride/archive/projects/PXD032729	PRIDE, PXD032729

References

Adzhubei I, Jordan DM, Sunyaev SR. 2013. Predicting functional effect of human missense mutations using polyphen-2. *Current Protocols in Human Genetics* **Chapter 7**:Unit7. DOI: https://doi.org/10.1002/0471142905.hg0720s76, PMID: 23315928

Andreux PA, Williams EG, Koutnikova H, Houtkooper RH, Champy MF, Henry H, Schoonjans K, Williams RW, Auwerx J. 2012. Systems genetics of metabolism: the use of the bxd murine reference panel for multiscalar integration of traits. *Cell* **150**:1287–1299. DOI: https://doi.org/10.1016/j.cell.2012.08.012, PMID: 22939713

Ashbrook DG, Arends D, Prins P, Mulligan MK, Roy S, Williams EG, Lutz CM, Valenzuela A, Bohl CJ, Ingels JF, McCarty MS, Centeno AG, Hager R, Auwerx J, Lu L, Williams RW. 2021. A platform for experimental precision medicine: the extended bxd mouse family. *Cell Systems* **12**:235–247. DOI: https://doi.org/10.1016/j.cels.2020.12.002, PMID: 33472028

Ashoti A, Limone F, van Kranenburg M, Alemany A, Baak M, Vivié J, Piccioni F, Dijkers PF, Creyghton M, Eggan K, Geijsen N. 2022. Considerations and practical implications of performing a phenotypic CRISPR/cas survival screen. *PLOS ONE* **17**:e0263262. DOI: https://doi.org/10.1371/journal.pone.0263262, PMID: 35176052

Baliga NS, Björkegren JLM, Boeke JD, Boutros M, Crawford NPS, Dudley AM, Farber CR, Jones A, Levey AI, Lusis AJ, Mak HC, Nadeau JH, Noyes MB, Petretto E, Seyfried NT, Steinmetz LM, Vonesch SC. 2017. The state of systems genetics in 2017. *Cell Systems* **4**:7–15. DOI: https://doi.org/10.1016/j.cels.2017.01.005, PMID: 28125793

Bennett BJ, Davis RC, Civelek M, Orozco L, Wu J, Qi H, Pan C, Packard RRS, Eskin E, Yan M, Kirchgessner T, Wang Z, Li X, Gregory JC, Hazen SL, Gargalovic PS, Lusis AJ. 2015. Genetic architecture of atherosclerosis in mice: a systems genetics analysis of common inbred strains. *PLOS Genetics* **11**:e1005711. DOI: https://doi.org/10.1371/journal.pgen.1005711, PMID: 26694027

Bennett AH, O'Donohue M-F, Gundry SR, Chan AT, Widrick J, Draper I, Chakraborty A, Zhou Y, Zon LI, Gleizes P-E, Beggs AH, Gupta VA. 2018. RNA helicase, ddx27 regulates skeletal muscle growth and regeneration by modulation of translational processes. *PLOS Genetics* **14**:e1007226. DOI: https://doi.org/10.1371/journal.pgen.1007226, PMID: 29518074

Benson MD, Yang Q, Ngo D, Zhu Y, Shen D, Farrell LA, Sinha S, Keyes MJ, Vasan RS, Larson MG, Smith JG, Wang TJ, Gerszten RE. 2018. Genetic architecture of the cardiovascular risk proteome. *Circulation* **137**:1158–1172. DOI: https://doi.org/10.1161/CIRCULATIONAHA.117.029536, PMID: 29258991

Beron C, Vidal-Gadea AG, Cohn J, Parikh A, Hwang G, Pierce-Shimomura JT. 2015. The burrowing behavior of the nematode *Caenorhabditis elegans*: a new assay for the study of neuromuscular disorders. *Genes, Brain, and Behavior* **14**:357–368. DOI: https://doi.org/10.1111/gbb.12217, PMID: 25868909

Bi P, Ramirez-Martinez A, Li H, Cannavino J, McAnally JR, Shelton JM, Sánchez-Ortiz E, Bassel-Duby R, Olson EN. 2017. Control of muscle formation by the fusogenic micropeptide Myomixer. Science **356**:323–327. DOI: https://doi.org/10.1126/science.aam9361, PMID: 28386024

Bikker H, Bakker HD, Abeling NGGM, Poll-The BT, Kleijer WJ, Rosenblatt DS, Waterham HR, Wanders RJA, Duran M. 2006. A homozygous nonsense mutation in the methylmalonyl-CoA epimerase gene (mcee) results in mild methylmalonic aciduria. *Human Mutation* **27**:640–643. DOI: https://doi.org/10.1002/humu.20373, PMID: 16752391

Birely J, Schneider VA, Santana E, Dosch R, Wagner DS, Mullins MC, Granato M. 2005. Genetic screens for genes controlling motor nerve-muscle development and interactions. *Developmental Biology* **280**:162–176. DOI: https://doi.org/10.1016/j.ydbio.2005.01.012, PMID: 15766756

Blankinship MJ, Gregorevic P, Allen JM, Harper SQ, Harper H, Halbert CL, Miller AD, Chamberlain JS. 2004. Efficient transduction of skeletal muscle using vectors based on adeno-associated virus serotype 6. *Molecular Therapy* **10**:671–678. DOI: https://doi.org/10.1016/j.ymthe.2004.07.016, PMID: 15451451

Boudon S, Ounaissi D, Viala D, Monteils V, Picard B, Cassar-Malek I. 2020. Label free shotgun proteomics for the identification of protein biomarkers for beef tenderness in muscle and plasma of heifers. *Journal of Proteomics* **217**:103685. DOI: https://doi.org/10.1016/j.jprot.2020.103685, PMID: 32058039

Bovo S, Di Luca A, Galimberti G, Dall'Olio S, Fontanesi L. 2018. A comparative analysis of label-free liquid chromatography-mass spectrometry liver proteomic profiles highlights metabolic differences between pig breeds. *PLOS ONE* **13**:e0199649. DOI: https://doi.org/10.1371/journal.pone.0199649, PMID: 30208024

Bruderer R, Bernhardt OM, Gandhi T, Miladinović SM, Cheng L-Y, Messner S, Ehrenberger T, Zanotelli V, Butscheid Y, Escher C, Vitek O, Rinner O, Reiter L. 2015. Extending the limits of quantitative proteome profiling with data-independent acquisition and application to acetaminophen-treated three-dimensional liver microtissues. *Molecular & Cellular Proteomics* **14**:1400–1410. DOI: https://doi.org/10.1074/mcp.M114.044305, PMID: 25724911

Cao Y, Vergnes L, Wang YC, Pan C, Chella Krishnan K, Moore TM, Rosa-Garrido M, Kimball TH, Zhou Z, Charugundla S, Rau CD, Seldin MM, Wang J, Wang Y, Vondriska TM, Reue K, Lusis AJ. 2022. Sex differences in heart mitochondria regulate diastolic dysfunction. *Nature Communications* **13**:3850. DOI: https://doi.org/10.1038/s41467-022-31544-5, PMID: 35787630

Cermenati G, Abbiati F, Cermenati S, Brioschi E, Volonterio A, Cavaletti G, Saez E, De Fabiani E, Crestani M, Garcia-Segura LM, Melcangi RC, Caruso D, Mitro N. 2012. Diabetes-Induced myelin abnormalities are associated with an altered lipid pattern: protective effects of LXR activation. *Journal of Lipid Research* **53**:300–310. DOI: https://doi.org/10.1194/jlr.M021188, PMID: 22158827

Chick JM, Munger SC, Simecek P, Huttlin EL, Choi K, Gatti DM, Raghupathy N, Svenson KL, Churchill GA, Gygi SP. 2016. Defining the consequences of genetic variation on a proteome-wide scale. *Nature* **534**:500–505. DOI: https://doi.org/10.1038/nature18270, PMID: 27309819

Choi Y, Sims GE, Murphy S, Miller JR, Chan AP. 2012. Predicting the functional effect of amino acid substitutions and indels. *PLOS ONE* **7**:e46688. DOI: https://doi.org/10.1371/journal.pone.0046688, PMID: 23056405

Choi Y, Chan AP. 2015. PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics* **31**:2745–2747. DOI: https://doi.org/10.1093/bioinformatics/btv195, PMID: 25851949

Davignon L, Chauveau C, Julien C, Dill C, Duband-Goulet I, Cabet E, Buendia B, Lilienbaum A, Rendu J, Minot MC, Guichet A, Allamand V, Vadrot N, Fauré J, Odent S, Lazaro L, Leroy JP, Marcorelles P, Dubourg O, Ferreiro A. 2016. The transcription coactivator ASC-1 is a regulator of skeletal myogenesis, and its deficiency causes a novel form of congenital muscle disease. *Human Molecular Genetics* **25**:1559–1573. DOI: https://doi.org/10.1093/hmg/ddw033, PMID: 27008887

Delgado J, Radusky LG, Cianferoni D, Serrano L. 2019. FoldX 5.0: working with RNA, small molecules and a new graphical interface. *Bioinformatics* **35**:4168–4169. DOI: https://doi.org/10.1093/bioinformatics/btz184, PMID: 30874800

Emilsson V, Ilkov M, Lamb JR, Finkel N, Gudmundsson EF, Pitts R, Hoover H, Gudmundsdottir V, Horman SR, Aspelund T, Shu L, Trifonov V, Sigurdsson S, Manolescu A, Zhu J, Olafsson Ö, Jakobsdottir J, Lesley SA, To J, Zhang J, et al. 2018. Co-Regulatory networks of human serum proteins link genetics to disease. Science **361**:769–773. DOI: https://doi.org/10.1126/science.aaq1327, PMID: 30072576

Eng JK, McCormack AL, Yates JR. 1994. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *Journal of the American Society for Mass Spectrometry* **5**:976–989. DOI: https://doi.org/10.1016/1044-0305(94)80016-2, PMID: 24226387

Flint J, Eskin E. 2012. Genome-Wide association studies in mice. *Nature Reviews. Genetics* **13**:807–817. DOI: https://doi.org/10.1038/nrg3335, PMID: 23044826

Foss EJ, Radulovic D, Shaffer SA, Ruderfer DM, Bedalov A, Goodlett DR, Kruglyak L. 2007. Genetic basis of proteome variation in yeast. *Nature Genetics* **39**:1369–1375. DOI: https://doi.org/10.1038/ng.2007.22, PMID: 17952072

Gautheron J, Morisseau C, Chung WK, Zammouri J, Auclair M, Baujat G, Capel E, Moulin C, Wang Y, Yang J, Hammock BD, Cerame B, Phan F, Fève B, Vigouroux C, Andreelli F, Jeru I. 2021. *ephx1* mutations cause a lipoatrophic diabetes syndrome due to impaired epoxide hydrolysis and increased cellular senescence. *eLife* **10**:e68445. DOI: https://doi.org/10.7554/eLife.68445, PMID: 34342583

Ghazalpour A, Bennett B, Petyuk VA, Orozco L, Hagopian R, Mungrue IN, Farber CR, Sinsheimer J, Kang HM, Furlotte N, Park CC, Wen PZ, Brewer H, Weitz K, Camp DG, Pan C, Yordanova R, Neuhaus I, Tilford C, Siemers N, et al. 2011. Comparative analysis of proteome and transcriptome variation in mouse. *PLOS Genetics* **7**:e1001393. DOI: https://doi.org/10.1371/journal.pgen.1001393, PMID: 21695224

Ghazalpour A, Rau CD, Farber CR, Bennett BJ, Orozco LD, Nas A, Pan C, Allayee H, Beaven SW, Civelek M, Davis RC, Drake TA, Friedman RA, Furlotte N, Hui ST, Jentsch JD, Kostem E, Kang HM, Kang EY, Joo JW, et al. 2012. Hybrid mouse diversity panel: a panel of inbred mouse strains suitable for analysis of complex genetic traits. *Mamm Genome* **23**:680–692. DOI: https://doi.org/10.1007/s00335-012-9411-5

Ghazalpour A, Bennett BJ, Shih D, Che N, Orozco L, Pan C, Hagopian R, He A, Kayne P, Yang W, Kirchgessner T, Lusis AJ. 2014. Genetic regulation of mouse liver metabolite levels. *Molecular Systems Biology* **10**:730. DOI: https://doi.org/10.15252/msb.20135004, PMID: 24860088

Graca FA, Sheffield N, Puppa M, Finkelstein D, Hunt LC, Demontis F. 2021. A large-scale transgenic rnai screen identifies transcription factors that modulate myofiber size in *Drosophila*. *PLOS Genetics* **17**:e1009926. DOI: https://doi.org/10.1371/journal.pgen.1009926, PMID: 34780463

Grimm D, Streetz KL, Jopling CL, Storm TA, Pandey K, Davis CR, Marion P, Salazar F, Kay MA. 2006. Fatality in mice due to oversaturation of cellular microrna/short hairpin RNA pathways. *Nature* **441**:537–541. DOI: https://doi.org/10.1038/nature04791, PMID: 16724069

Grings M, Seminotti B, Karunanidhi A, Ghaloul-Gonzalez L, Mohsen AW, Wipf P, Palmfeldt J, Vockley J, Leipnitz G. 2019. ETHE1 and mocs1 deficiencies: disruption of mitochondrial bioenergetics, dynamics, redox homeostasis and endoplasmic reticulum-mitochondria crosstalk in patient fibroblasts. *Scientific Reports* **9**:12651. DOI: https://doi.org/10.1038/s41598-019-49014-2, PMID: 31477743

Gu Z, Gu L, Eils R, Schlesner M, Brors B. 2014. Circlize implements and enhances circular visualization in R. *Bioinformatics* **30**:2811–2812. DOI: https://doi.org/10.1093/bioinformatics/btu393, PMID: 24930139

Hahne F, Ivanek R. 2016. Visualizing genomic data using gviz and bioconductor. *Methods in Molecular Biology* **1418**:335–351. DOI: https://doi.org/10.1007/978-1-4939-3578-9_16, PMID: 27008022

Hilton JB, Mercer SW, Lim NKH, Faux NG, Buncic G, Beckman JS, Roberts BR, Donnelly PS, White AR, Crouch PJ. 2017. Cuii (atsm) improves the neurological phenotype and survival of SOD1G93A mice and selectively increases enzymatically active SOD1 in the spinal cord. *Scientific Reports* **7**:42292. DOI: https://doi.org/10.1038/srep42292, PMID: 28205575

Horstick EJ, Linsley JW, Dowling JJ, Hauser MA, McDonald KK, Ashley-Koch A, Saint-Amant L, Satish A, Cui WW, Zhou W, Sprague SM, Stamm DS, Powell CM, Speer MC, Franzini-Armstrong C, Hirata H, Kuwada JY. 2013. Stac3 is a component of the excitation-contraction coupling machinery and mutated in native american myopathy. *Nature Communications* **4**:1952. DOI: https://doi.org/10.1038/ncomms2952, PMID: 23736855

Howe KL, Achuthan P, Allen J, Allen J, Alvarez-Jarreta J, Amode MR, Armean IM, Azov AG, Bennett R, Bhai J, Billis K, Boddu S, Charkhchi M, Cummins C, Da Rin Fioretto L, Davidson C, Dodiya K, El Houdaigui B, Fatima R, Gall A, et al. 2021. Ensembl 2021. *Nucleic Acids Research* **49**:D884–D891. DOI: https://doi.org/10.1093/nar/gkaa942, PMID: 33137190

Hu X, Wang H, Li K, Wu Y, Liu Z, Huang C. 2017. Genome-wide proteomic profiling reveals the role of dominance protein expression in heterosis in immature maize ears. *Scientific Reports* **7**:16130. DOI: https://doi.org/10.1038/s41598-017-15985-3, PMID: 29170427

Hudson BH, Frederick JP, Drake LY, Megosh LC, Irving RP, York JD. 2013. Role for cytoplasmic nucleotide hydrolysis in hepatic function and protein synthesis. *PNAS* **110**:5040–5045. DOI: https://doi.org/10.1073/pnas.1205001110, PMID: 23479625

Jha P, McDevitt MT, Gupta R, Quiros PM, Williams EG, Gariani K, Sleiman MB, Diserens L, Jochem A, Ulbrich A, Coon JJ, Auwerx J, Pagliarini DJ. 2018a. Systems analyses reveal physiological roles and genetic regulators of liver lipid species. *Cell Systems* **6**:722–733. DOI: https://doi.org/10.1016/j.cels.2018.05.016, PMID: 29909277

Jha P, McDevitt MT, Halilbasic E, Williams EG, Quiros PM, Gariani K, Sleiman MB, Gupta R, Ulbrich A, Jochem A, Coon JJ, Trauner M, Pagliarini DJ, Auwerx J. 2018b. Genetic regulation of plasma lipid species and their association with metabolic phenotypes. *Cell Systems* **6**:709–721. DOI: https://doi.org/10.1016/j.cels.2018.05.009, PMID: 29909275

Jiang LG, Li B, Liu SX, Wang HW, Li CP, Song SH, Beatty M, Zastrow-Hayes G, Yang XH, Qin F, He Y. 2019. Characterization of proteome variation during modern maize breeding. *Molecular & Cellular Proteomics* **18**:263–276. DOI: https://doi.org/10.1074/mcp.RA118.001021, PMID: 30409858

Johnson AN, Mokalled MH, Valera JM, Poss KD, Olson EN. 2013. Post-transcriptional regulation of myotube elongation and myogenesis by hoi polloi. *Development* **140**:3645–3656. DOI: https://doi.org/10.1242/dev.095596, PMID: 23942517

Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, Tunyasuvunakool K, Bates R, Žídek A, Potapenko A, Bridgland A, Meyer C, Kohl SAA, Ballard AJ, Cowie A, Romera-Paredes B, Nikolov S, Jain R, Adler J, Back T, et al. 2021. Highly accurate protein structure prediction with alphafold. *Nature* **596**:583–589. DOI: https://doi.org/10.1038/s41586-021-03819-2, PMID: 34265844

Käll L, Canterbury JD, Weston J, Noble WS, MacCoss MJ. 2007. Semi-supervised learning for peptide identification from shotgun proteomics datasets. *Nature Methods* **4**:923–925. DOI: https://doi.org/10.1038/nmeth1113, PMID: 17952086

Käll L, Storey JD, Noble WS. 2009. QVALITY: non-parametric estimation of q-values and posterior error probabilities. *Bioinformatics* **25**:964–966. DOI: https://doi.org/10.1093/bioinformatics/btp021, PMID: 19193729

Kang HM, Zaitlen NA, Wade CM, Kirby A, Heckerman D, Daly MJ, Eskin E. 2008. Efficient control of population structure in model organism association mapping. *Genetics* **178**:1709–1723. DOI: https://doi.org/10.1534/genetics.107.080101, PMID: 18385116

Kao SY, Nikonova E, Chaabane S, Sabani A, Martitz A, Wittner A, Heemken J, Straub T, Spletter ML. 2021. A candidate rnai screen reveals diverse rna-binding protein phenotypes in *Drosophila* flight muscle. *Cells* **10**:2505. DOI: https://doi.org/10.3390/cells10102505

Karczewski KJ, Solomonson M, Chao KR, Goodrich JK, Tiao G, Lu W, Riley-Gillis BM. 2022. Systematic Single-Variant and Gene-Based Association Testing of Thousands of Phenotypes in 426,370 UK Biobank Exomes. *medRxiv*. DOI: https://doi.org/10.1101/2021.06.19.21259117

Keller MP, Rabaglia ME, Schueler KL, Stapleton DS, Gatti DM, Vincent M, Mitok KA, Wang Z, Ishimura T, Simonett SP, Emfinger CH, Das R, Beck T, Kendziorski C, Broman KW, Yandell BS, Churchill GA, Attie AD. 2019.

Gene loci associated with insulin secretion in islets from non-diabetic mice. *The Journal of Clinical Investigation* **129**:4419–4432. DOI: https://doi.org/10.1172/JCI129143, PMID: 31343992

Khan IF, Hirata RK, Russell DW. 2011. Aav-Mediated gene targeting methods for human cells. *Nature Protocols* **6**:482–501. DOI: https://doi.org/10.1038/nprot.2011.301, PMID: 21455185

Langfelder P, Horvath S. 2008. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* **9**:559. DOI: https://doi.org/10.1186/1471-2105-9-559, PMID: 19114008

Lek A, Zhang Y, Woodman KG, Huang S, DeSimone AM, Cohen J, Ho V, Conner J, Mead L, Kodani A, Pakula A, Sanjana N, King OD, Jones PL, Wagner KR, Lek M, Kunkel LM. 2020. Applying genome-wide CRISPR-cas9 screens for therapeutic discovery in facioscapulohumeral muscular dystrophy. *Science Translational Medicine* **12**:536. DOI: https://doi.org/10.1126/scitranslmed.aay0271, PMID: 32213627

Le Vasseur M, Friedman J, Jost M, Xu J, Yamada J, Kampmann M, Horlbeck MA, Salemi MR, Phinney BS, Weissman JS, Nunnari J. 2021. Genome-Wide crispri screening identifies ociad1 as a prohibitin client and regulatory determinant of mitochondrial complex III assembly in human cells. *eLife* **10**:e67624. DOI: https://doi.org/10.7554/eLife.67624, PMID: 34034859

Lee L, Perez Oliva AB, Martinez-Balsalobre E, Churikov D, Peter J, Rahmouni D, Audoly G, Azzoni V, Audebert S, Camoin L, Mulero V, Cayuela ML, Kulathu Y, Geli V, Lachaud C. 2021. UFMylation of Mre11 is essential for telomere length maintenance and hematopoietic stem cell survival. *Science Advances* **7**:eabc7371. DOI: https://doi.org/10.1126/sciadv.abc7371, PMID: 34559557

Liang JR, Lingeman E, Luong T, Ahmed S, Muhar M, Nguyen T, Olzmann JA, Corn JE. 2020. A genome-wide er-phagy screen highlights key roles of mitochondrial metabolism and ER-resident ufmylation. *Cell* **180**:1160–1177.. DOI: https://doi.org/10.1016/j.cell.2020.02.017, PMID: 32160526

Linke V, Overmyer KA, Miller IJ, Brademan DR, Hutchins PD, Trujillo EA, Reddy TR, Russell JD, Cushing EM, Schueler KL, Stapleton DS, Rabaglia ME, Keller MP, Gatti DM, Keele GR, Pham D, Broman KW, Churchill GA, Attie AD, Coon JJ. 2020. A large-scale genome-lipid association MAP guides lipid identification. *Nature Metabolism* **2**:1149–1162. DOI: https://doi.org/10.1038/s42255-020-00278-3, PMID: 32958938

Lusis AJ, Seldin MM, Allayee H, Bennett BJ, Civelek M, Davis RC, Eskin E, Farber CR, Hui S, Mehrabian M, Norheim F, Pan C, Parks B, Rau CD, Smith DJ, Vallim T, Wang Y, Wang J. 2016. The hybrid mouse diversity panel: a resource for systems genetics analyses of metabolic and cardiovascular traits. *Journal of Lipid Research* **57**:925–942. DOI: https://doi.org/10.1194/jlr.R066944, PMID: 27099397

Lynch EM, Robertson S, FitzGibbons C, Reilly M, Switalski C, Eckardt A, Tey SR, Hayakawa K, Suzuki M. 2021. Transcriptome analysis using patient ipsc-derived skeletal myocytes: bet1l as a new molecule possibly linked to neuromuscular junction degeneration in ALS. *Experimental Neurology* **345**:113815. DOI: https://doi.org/10.1016/j.expneurol.2021.113815, PMID: 34310943

McAlister GC, Nusinow DP, Jedrychowski MP, Wühr M, Huttlin EL, Erickson BK, Rad R, Haas W, Gygi SP. 2014. MultiNotch ms3 enables accurate, sensitive, and multiplexed detection of differential expression across cancer cell line proteomes. *Analytical Chemistry* **86**:7150–7158. DOI: https://doi.org/10.1021/ac502040v, PMID: 24927332

McGregor RA, Cameron-Smith D, Poppitt SD. 2014. It is not just muscle mass: a review of muscle quality, composition and metabolism during ageing as determinants of muscle function and mobility in later life. *Longevity & Healthspan* **3**:9. DOI: https://doi.org/10.1186/2046-2395-3-9, PMID: 25520782

McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GRS, Thormann A, Flicek P, Cunningham F. 2016. The ensembl variant effect predictor. *Genome Biology* **17**:122. DOI: https://doi.org/10.1186/s13059-016-0974-4, PMID: 27268795

Mills RJ, Titmarsh DM, Koenig X, Parker BL, Ryall JG, Quaife-Ryan GA, Voges HK, Hodson MP, Ferguson C, Drowley L, Plowright AT, Needham EJ, Wang QD, Gregorevic P, Xin M, Thomas WG, Parton RG, Nielsen LK, Launikonis BS, James DE, et al. 2017. Functional screening in human cardiac organoids reveals a metabolic mechanism for cardiomyocyte cell cycle arrest. *PNAS* **114**:E8372–E8381. DOI: https://doi.org/10.1073/pnas.1707316114, PMID: 28916735

Mills RJ, Parker BL, Monnot P, Needham EJ, Vivien CJ, Ferguson C, Parton RG, James DE, Porrello ER, Hudson JE. 2019. Development of a human skeletal micro muscle platform with pacing capabilities. *Biomaterials* **198**:217–227. DOI: https://doi.org/10.1016/j.biomaterials.2018.11.030, PMID: 30527761

Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. 2022. ColabFold: making protein folding accessible to all. *Nature Methods* **19**:679–682. DOI: https://doi.org/10.1038/s41592-022-01488-1, PMID: 35637307

Molendijk J, Parker BL. 2021a. Proteome-wide systems genetics to identify functional regulators of complex traits. *Cell Systems* **12**:5–22. DOI: https://doi.org/10.1016/j.cels.2020.10.005, PMID: 33476553

Molendijk J, Seldin MM, Parker BL. 2021b. CoffeeProt: an online tool for correlation and functional enrichment of systems genetics data. *Nucleic Acids Research* **49**:W104–W113. DOI: https://doi.org/10.1093/nar/gkab352, PMID: 33978718

Molendijk J. 2022. Skeletal_muscle. swh:1:rev:9311d7bfb59979d80e18612879631dc78f2f0902. Software Heritage. https://archive.softwareheritage.org/swh:1:dir:878f0ef2fdf3015f817bce6df08d38bc0586bc7a;origin=https://github.com/JeffreyMolendijk/skeletal_muscle;visit=swh:1:snp:1fadaad91ee3bf57f7743ddb30daa0cabfb0c0f7;anchor=swh:1:rev:9311d7bfb59979d80e18612879631dc78f2f0902

Molendijk J, Yip R, Parker BL. 2022. UrPTMdb/teaprot: upstream and downstream proteomics analysis. *Journal of Proteome Research* **1**:2c00048. DOI: https://doi.org/10.1021/acs.jproteome.2c00048, PMID: 35759515

Mozhui K, Ciobanu DC, Schikorski T, Wang X, Lu L, Williams RW. 2008. Dissection of a QTL hotspot on mouse distal chromosome 1 that modulates neurobehavioral phenotypes and gene expression. *PLOS Genetics* **4**:e1000260. DOI: https://doi.org/10.1371/journal.pgen.1000260, PMID: 19008955

Norheim F, Hui ST, Kulahcioglu E, Mehrabian M, Cantor RM, Pan C, Parks BW, Lusis AJ. 2017. Genetic and hormonal control of hepatic steatosis in female and male mice. *Journal of Lipid Research* **58**:178–187. DOI: https://doi.org/10.1194/jlr.M071522, PMID: 27811231

Norheim F, Hasin-Brumshtein Y, Vergnes L, Chella Krishnan K, Pan C, Seldin MM, Hui ST, Mehrabian M, Zhou Z, Gupta S, Parks BW, Walch A, Reue K, Hofmann SM, Arnold AP, Lusis AJ. 2019. Gene-by-sex interactions in mitochondrial functions and cardio-metabolic traits. *Cell Metabolism* **29**:932–949.. DOI: https://doi.org/10.1016/j.cmet.2018.12.013, PMID: 30639359

Norheim F, Chella Krishnan K, Bjellaas T, Vergnes L, Pan C, Parks BW, Meng Y, Lang J, Ward JA, Reue K, Mehrabian M, Gundersen TE, Péterfy M, Dalen KT, Drevon CA, Hui ST, Lusis AJ, Seldin MM. 2021. Genetic regulation of liver lipids in a mouse model of insulin resistance and hepatic steatosis. *Molecular Systems Biology* **17**:e9684. DOI: https://doi.org/10.15252/msb.20209684, PMID: 33417276

Okada H, Ebhardt HA, Vonesch SC, Aebersold R, Hafen E. 2016. Proteome-wide association studies identify biochemical modules associated with a wing-size phenotype in *Drosophila melanogaster*. *Nature Communications* **7**:12649. DOI: https://doi.org/10.1038/ncomms12649, PMID: 27582081

Pansarasa O, Rossi D, Berardinelli A, Cereda C. 2014. Amyotrophic lateral sclerosis and skeletal muscle: an update. *Molecular Neurobiology* **49**:984–990. DOI: https://doi.org/10.1007/s12035-013-8578-4, PMID: 24198230

Parker BL, Calkin AC, Seldin MM, Keating MF, Tarling EJ, Yang P, Moody SC, Liu Y, Zerenturk EJ, Needham EJ, Miller ML, Clifford BL, Morand P, Watt MJ, Meex RCR, Peng K-Y, Lee R, Jayawardana K, Pan C, Mellett NA, et al. 2019. An integrative systems genetic analysis of mammalian lipid metabolism. *Nature* **567**:187–193. DOI: https://doi.org/10.1038/s41586-019-0984-y, PMID: 30814737

Parks BW, Nam E, Org E, Kostem E, Norheim F, Hui ST, Pan C, Civelek M, Rau CD, Bennett BJ, Mehrabian M, Ursell LK, He A, Castellani LW, Zinker B, Kirby M, Drake TA, Drevon CA, Knight R, Gargalovic P, et al. 2013. Genetic control of obesity and gut microbiota composition in response to high-fat, high-sucrose diet in mice. *Cell Metabolism* **17**:141–152. DOI: https://doi.org/10.1016/j.cmet.2012.12.007, PMID: 23312289

Parks BW, Sallam T, Mehrabian M, Psychogios N, Hui ST, Norheim F, Castellani LW, Rau CD, Pan C, Phun J, Zhou Z, Yang WP, Neuhaus I, Gargalovic PS, Kirchgessner TG, Graham M, Lee R, Tontonoz P, Gerszten RE, Hevener AL, et al. 2015. Genetic architecture of insulin resistance in the mouse. *Cell Metabolism* **21**:334–347. DOI: https://doi.org/10.1016/j.cmet.2015.01.002, PMID: 25651185

Parts L, Liu YC, Tekkedil MM, Steinmetz LM, Caudy AA, Fraser AG, Boone C, Andrews BJ, Rosebrock AP. 2014. Heritability and genetic basis of protein level variation in an outbred population. *Genome Research* **24**:1363–1370. DOI: https://doi.org/10.1101/gr.170506.113, PMID: 24823668

Pech M, Spreter T, Beckmann R, Beatrix B. 2010. Dual binding mode of the nascent polypeptide-associated complex reveals a novel universal adapter site on the ribosome. *The Journal of Biological Chemistry* **285**:19679–19687. DOI: https://doi.org/10.1074/jbc.M109.092536, PMID: 20410297

Pelossof R, Fairchild L, Huang CH, Widmer C, Sreedharan VT, Sinha N, Lai DY, Guan Y, Premsrirut PK, Tschaharganeh DF, Hoffmann T, Thapar V, Xiang Q, Garippa RJ, Rätsch G, Zuber J, Lowe SW, Leslie CS, Fellmann C. 2017. Prediction of potent shRNAs with a sequential classification algorithm. *Nature Biotechnology* **35**:350–353. DOI: https://doi.org/10.1038/nbt.3807, PMID: 28263295

Perez-Riverol Y, Csordas A, Bai J, Bernal-Llinares M, Hewapathirana S, Kundu DJ, Inuganti A, Griss J, Mayer G, Eisenacher M, Pérez E, Uszkoreit J, Pfeuffer J, Sachsenberg T, Yilmaz S, Tiwary S, Cox J, Audain E, Walzer M, Jarnuczak AF, et al. 2019. The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Research* **47**:D442–D450. DOI: https://doi.org/10.1093/nar/gky1106, PMID: 30395289

Picotti P, Clément-Ziza M, Lam H, Campbell DS, Schmidt A, Deutsch EW, Röst H, Sun Z, Rinner O, Reiter L, Shen Q, Michaelson JJ, Frei A, Alberti S, Kusebauch U, Wollscheid B, Moritz RL, Beyer A, Aebersold R. 2013. A complete mass-spectrometric map of the yeast proteome applied to quantitative trait analysis. *Nature* **494**:266–270. DOI: https://doi.org/10.1038/nature11835, PMID: 23334424

Qin B, Yu J, Nowsheen S, Wang M, Tu X, Liu T, Li H, Wang L, Lou Z. 2019. UFL1 promotes histone H4 ufmylation and ATM activation. *Nature Communications* **10**:1242. DOI: https://doi.org/10.1038/s41467-019-09175-0, PMID: 30886146

Rau CD, Wang J, Avetisyan R, Romay MC, Martin L, Ren S, Wang Y, Lusis AJ. 2015. Mapping genetic contributions to cardiac pathology induced by beta-adrenergic stimulation in mice. *Circulation. Cardiovascular Genetics* **8**:40–49. DOI: https://doi.org/10.1161/CIRCGENETICS.113.000732, PMID: 25480693

Rau CD, Romay MC, Tuteryan M, Wang JJC, Santolini M, Ren S, Karma A, Weiss JN, Wang Y, Lusis AJ. 2017. Systems genetics approach identifies gene pathways and adamts2 as drivers of isoproterenol-induced cardiac hypertrophy and cardiomyopathy in mice. *Cell Systems* **4**:121–128. DOI: https://doi.org/10.1016/j.cels.2016.10.016, PMID: 27866946

Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, Smyth GK. 2015. Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Research* **43**:e47. DOI: https://doi.org/10.1093/nar/gkv007, PMID: 25605792

Roberts BR, Lim NKH, McAllum EJ, Donnelly PS, Hare DJ, Doble PA, Turner BJ, Price KA, Lim SC, Paterson BM, Hickey JL, Rhoads TW, Williams JR, Kanninen KM, Hung LW, Liddell JR, Grubman A, Monty J-F, Llanos RM, Kramer DR, et al. 2014. Oral treatment with Cu (II) (atsm) increases mutant SOD1 in vivo but protects motor

neurons and improves the phenotype of a transgenic mouse model of amyotrophic lateral sclerosis. *The Journal of Neuroscience* **34**:8021–8031. DOI: https://doi.org/10.1523/JNEUROSCI.4196-13.2014, PMID: 24899723

Rodrigues CH, Pires DE, Ascher DB. 2018. DynaMut: predicting the impact of mutations on protein conformation, flexibility and stability. *Nucleic Acids Research* **46**:W350–W355. DOI: https://doi.org/10.1093/nar/gky300, PMID: 29718330

Schneider CA, Rasband WS, Eliceiri KW. 2012. Nih image to imagej: 25 years of image analysis. *Nature Methods* **9**:671–675. DOI: https://doi.org/10.1038/nmeth.2089, PMID: 22930834

Sehnal D, Bittrich S, Deshpande M, Svobodová R, Berka K, Bazgier V, Velankar S, Burley SK, Koča J, Rose AS. 2021. Mol* viewer: modern web APP for 3D visualization and analysis of large biomolecular structures. *Nucleic Acids Research* **49**:W431–W437. DOI: https://doi.org/10.1093/nar/gkab314, PMID: 33956157

Seldin M, Yang X, Lusis AJ. 2019. Systems genetics applications in metabolism research. *Nature Metabolism* **1**:1038–1050. DOI: https://doi.org/10.1038/s42255-019-0132-x, PMID: 32259026

Singh KD, Roschitzki B, Snoek LB, Grossmann J, Zheng X, Elvin M, Kamkina P, Schrimpf SP, Poulin GB, Kammenga JE, Hengartner MO. 2016. Natural genetic variation influences protein abundances in *C. elegans* developmental signalling pathways. *PLOS ONE* **11**:e0149418. DOI: https://doi.org/10.1371/journal.pone.0149418, PMID: 26985669

Snider DL, Park M, Murphy KA, Beachboard DC, Horner SM. 2022. Signaling from the RNA sensor RIG-I is regulated by ufmylation. *PNAS* **119**:e2119531119. DOI: https://doi.org/10.1073/pnas.2119531119, PMID: 35394863

Stoffel MA, Nakagawa S, Schielzeth H, Goslee S. 2017. RptR: repeatability estimation and variance decomposition by generalized linear mixed-effects models. *Methods in Ecology and Evolution* **8**:1639–1644. DOI: https://doi.org/10.1111/2041-210X.12797

Suhre K, Arnold M, Bhagwat AM, Cotton RJ, Engelke R, Raffler J, Sarwath H, Thareja G, Wahl A, DeLisle RK, Gold L, Pezer M, Lauc G, El-Din Selim MA, Mook-Kanamori DO, Al-Dous EK, Mohamoud YA, Malek J, Strauch K, Grallert H, et al. 2017. Connecting genetic risk to disease end points through the human blood plasma proteome. *Nature Communications* **8**:14357. DOI: https://doi.org/10.1038/ncomms14357, PMID: 28240269

Sun BB, Maranville JC, Peters JE, Stacey D, Staley JR, Blackshaw J, Burgess S, Jiang T, Paige E, Surendran P, Oliver-Williams C, Kamat MA, Prins BP, Wilcox SK, Zimmerman ES, Chi A, Bansal N, Spain SL, Wood AM, Morrell NW, et al. 2018. Genomic atlas of the human plasma proteome. *Nature* **558**:73–79. DOI: https://doi.org/10.1038/s41586-018-0175-2, PMID: 29875488

Thévenot EA, Roux A, Xu Y, Ezan E, Junot C. 2015. Analysis of the human adult urinary metabolome variations with age, body mass index, and gender by implementing a comprehensive workflow for univariate and OPLS statistical analyses. *Journal of Proteome Research* **14**:3322–3335. DOI: https://doi.org/10.1021/acs.jproteome.5b00354, PMID: 26088811

Tuominen I, Fuqua BK, Pan C, Renaud N, Wroblewski K, Civelek M, Clerkin K, Asaryan A, Haroutunian SG, Loureiro J, Borawski J, Roma G, Knehr J, Carbone W, French S, Parks BW, Hui ST, Mehrabian M, Magyar C, Cantor RM, et al. 2021. The genetic architecture of carbon tetrachloride-induced liver fibrosis in mice. *Cellular and Molecular Gastroenterology and Hepatology* **11**:199–220. DOI: https://doi.org/10.1016/j.jcmgh.2020.08.010, PMID: 32866618

Vujkovic M, Keaton JM, Lynch JA, Miller DR, Zhou J, Tcheandjieu C, Huffman JE, Assimes TL, Lorenz K, Zhu X, Hilliard AT, Judy RL, Huang J, Lee KM, Klarin D, Pyarajan S, Danesh J, Melander O, Rasheed A, Mallick NH, et al. 2020. Discovery of 318 new risk loci for type 2 diabetes and related vascular outcomes among 1.4 million participants in a multi-ancestry meta-analysis. *Nature Genetics* **52**:680–691. DOI: https://doi.org/10.1038/s41588-020-0637-y, PMID: 32541925

Walczak CP, Leto DE, Zhang L, Riepe C, Muller RY, DaRosa PA, Ingolia NT, Elias JE, Kopito RR. 2019. Ribosomal protein rpl26 is the principal target of ufmylation. *PNAS* **116**:1299–1308. DOI: https://doi.org/10.1073/pnas.1816202116, PMID: 30626644

Wang L, Xu Y, Rogers H, Saidi L, Noguchi CT, Li H, Yewdell JW, Guydosh NR, Ye Y. 2020. UFMylation of RPL26 links translocation-associated quality control to endoplasmic reticulum protein homeostasis. *Cell Research* **30**:5–20. DOI: https://doi.org/10.1038/s41422-019-0236-6, PMID: 31595041

Watson CM, Crinnion LA, Gleghorn L, Newman WG, Ramesar R, Beighton P, Wallis GA. 2015. Identification of a mutation in the ubiquitin-fold modifier 1-specific peptidase 2 gene, ufsp2, in an extended south african family with beukes hip dysplasia. *South African Medical Journal* **105**:558. DOI: https://doi.org/10.7196/SAMJnew.7917, PMID: 26428751

Williams EG, Wu Y, Jha P, Dubuis S, Blattmann P, Argmann CA, Houten SM, Amariuta T, Wolski W, Zamboni N, Aebersold R, Auwerx J. 2016. Systems proteomics of liver mitochondria function. Science **352**:aad0189. DOI: https://doi.org/10.1126/science.aad0189, PMID: 27284200

Williams EG, Wu Y, Wolski W, Kim JY, Lan J, Hasan M, Halter C, Jha P, Ryu D, Auwerx J, Aebersold R. 2018. Quantifying and localizing the mitochondrial proteome across five tissues in a mouse population. *Molecular & Cellular Proteomics* **17**:1766–1777. DOI: https://doi.org/10.1074/mcp.RA118.000554, PMID: 29945935

Wong PC, Pardo CA, Borchelt DR, Lee MK, Copeland NG, Jenkins NA, Sisodia SS, Cleveland DW, Price DL. 1995. An adverse property of a familial ALS-linked SOD1 mutation causes motor neuron disease characterized by vacuolar degeneration of mitochondria. *Neuron* **14**:1105–1116. DOI: https://doi.org/10.1016/0896-6273(95)90259-7, PMID: 7605627

Wu Y, Williams EG, Dubuis S, Mottis A, Jovaisaite V, Houten SM, Argmann CA, Faridi P, Wolski W, Kutalik Z, Zamboni N, Auwerx J, Aebersold R. 2014. Multilayered genetic and omics dissection of mitochondrial activity in a mouse reference population. *Cell* **158**:1415–1430. DOI: https://doi.org/10.1016/j.cell.2014.07.039, PMID: 25215496

Xiao H, Bozi LHM, Sun Y, Riley CL, Philip VM, Chen M, Li J, Zhang T, Mills EL, Emont MP, Sun W, Reddy A, Garrity R, Long J, Becher T, Vitas LP, Laznik-Bogoslavski D, Ordonez M, Liu X, Chen X, et al. 2022. Architecture of the outbred brown fat proteome defines regulators of metabolic physiology. *Cell* **185**:4654–4673.. DOI: https://doi.org/10.1016/j.cell.2022.10.003

Xu X, Hao Y, Wu J, Zhao J, Xiong S. 2021. Assessment of weighted gene co-expression network analysis to explore key pathways and novel biomarkers in muscular dystrophy. *Pharmacogenomics and Personalized Medicine* **14**:431–444. DOI: https://doi.org/10.2147/PGPM.S301098, PMID: 33883925

Yoo HM, Kang SH, Kim JY, Lee JE, Seong MW, Lee SW, Ka SH, Sou YS, Komatsu M, Tanaka K, Lee ST, Noh DY, Baek SH, Jeon YJ, Chung CH. 2014. Modification of Asc1 by Ufm1 is crucial for ERα transactivation and breast cancer development. *Molecular Cell* **56**:261–274. DOI: https://doi.org/10.1016/j.molcel.2014.08.007, PMID: 25219498

Zhang G, Tang S, Wang H, Pan H, Zhang W, Huang Y, Kong J, Wang Y, Gu J, Wang Y. 2020. Corrigendum to ufsp2-related spondyloepimetaphyseal dysplasia: A confirmatory report. *European Journal of Medical Genetics* **63**:104021. DOI: https://doi.org/10.1016/j.ejmg.2021.104143, PMID: 32755715

Appendix 1

Appendix 1—key resources table

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
Antibody	Rabbit monoclonal anti-UFC1	Abcam	EPR15014-102 (ab189252)	(1:1000)
Antibody	Rabbit monoclonal anti-UFSP2	Abcam	EP13424-49 (ab192597)	(1:1000)
Antibody	Rabbit monoclonal anti-UFM1	Abcam	EPR4264(2) (ab109305)	(1:1000)
Antibody	Rabbit monoclonal antiBiP	Cell Signaling Technologies	3177	(1:1000)
Antibody	Rabbit monoclonal SQSTM1/p62 (D1Q5S)	Cell Signaling Technologies	39749	(1:1000)
Antibody	Rabbit monoclonal K48-linkage Specific Polyubiquitin (D9D5)	Cell Signaling Technologies	8081	(1:1000)
Antibody	Donkey polyclonal Anti-Rabbit-HRP	Jackson ImmunoResearch	711-035-152 (RRID:AB_10015282)	(1:10000)
Genetic reagent (*Homo sapiens*)	Human Skeletal Myoblasts	Lonza	CC-2580 (lot #18TL269121)	
Genetic reagent (*Homo sapiens*)	Human embryonic kidney 293 cells expressing SV40 large T antigen	ATCC	CRL-1573	
Strain, strain background (*Mus musculus*)	A/J	JAX	RRID:IMSR_JAX:000646	
Strain, strain background (*Mus musculus*)	AXB10/PgnJ	JAX	RRID:IMSR_JAX:001681	
Strain, strain background (*Mus musculus*)	AXB13/PgnJ	JAX	RRID:IMSR_JAX:001684	
Strain, strain background (*Mus musculus*)	AXB15/PgnJ	JAX	RRID:IMSR_JAX:001685	
Strain, strain background (*Mus musculus*)	AXB19a/PgnJ	JAX	RRID:IMSR_JAX:001686	
Strain, strain background (*Mus musculus*)	AXB4/PgnJ	JAX	RRID:IMSR_JAX:001676	
Strain, strain background (*Mus musculus*)	AXB8/PgnJ	JAX	RRID:IMSR_JAX:001679	
Strain, strain background (*Mus musculus*)	B6.Cg-Tg(SOD1*G37R)42Dpr/J	JAX	RRID:IMSR_JAX:008342	
Strain, strain background (*Mus musculus*)	BALB/cByJ	JAX	RRID:IMSR_JAX:001026	
Strain, strain background (*Mus musculus*)	BTBR T+tf/J	NA	NA	
Strain, strain background (*Mus musculus*)	BUB/BnJ	JAX	RRID:IMSR_JAX:000653	

Appendix 1 Continued on next page

Appendix 1 Continued

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
Strain, strain background (*Mus musculus*)	BXA12/PgnJ	JAX	RRID:IMSR_JAX:001700	
Strain, strain background (*Mus musculus*)	BXA13/PgnJ	JAX	RRID:IMSR_JAX:001826	
Strain, strain background (*Mus musculus*)	BXA14/PgnJ	JAX	RRID:IMSR_JAX:001702	
Strain, strain background (*Mus musculus*)	BXA16/PgnJ	JAX	RRID:IMSR_JAX:001703	
Strain, strain background (*Mus musculus*)	BXA2/PgnJ	JAX	RRID:IMSR_JAX:001693	
Strain, strain background (*Mus musculus*)	BXA4/PgnJ	JAX	RRID:IMSR_JAX:001694	
Strain, strain background (*Mus musculus*)	BXD100	NA	NA	
Strain, strain background (*Mus musculus*)	BXD100/RwwJ	JAX	RRID:IMSR_JAX:007143	
Strain, strain background (*Mus musculus*)	BXD12/TyJ	JAX	RRID:IMSR_JAX:000045	
Strain, strain background (*Mus musculus*)	BXD14/TyJ	JAX	RRID:IMSR_JAX:000329	
Strain, strain background (*Mus musculus*)	BXD19/TyJ	JAX	RRID:IMSR_JAX:000010	
Strain, strain background (*Mus musculus*)	BXD21/TyJ	JAX	RRID:IMSR_JAX:000077	
Strain, strain background (*Mus musculus*)	BXD22/TyJ	JAX	RRID:IMSR_JAX:000043	
Strain, strain background (*Mus musculus*)	BXD27/TyJ	JAX	RRID:IMSR_JAX:000041	
Strain, strain background (*Mus musculus*)	BXD28/TyJ	JAX	RRID:IMSR_JAX:000047	
Strain, strain background (*Mus musculus*)	BXD29/TyJ	NA	NA	
Strain, strain background (*Mus musculus*)	BXD31/TyJ	JAX	RRID:IMSR_JAX:000083	
Strain, strain background (*Mus musculus*)	BXD32/TyJ	JAX	RRID:IMSR_JAX:000078	

Appendix 1 Continued on next page

Appendix 1 Continued

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
Strain, strain background (*Mus musculus*)	BXD33/TyJ	JAX	RRID:IMSR_JAX:003222	
Strain, strain background (*Mus musculus*)	BXD34/TyJ	JAX	RRID:IMSR_JAX:003223	
Strain, strain background (*Mus musculus*)	BXD39/TyJ	JAX	RRID:IMSR_JAX:003228	
Strain, strain background (*Mus musculus*)	BXD40/TyJ	JAX	RRID:IMSR_JAX:003229	
Strain, strain background (*Mus musculus*)	BXD44/RwwJ	JAX	RRID:IMSR_JAX:007094	
Strain, strain background (*Mus musculus*)	BXD45/RwwJ	JAX	RRID:IMSR_JAX:007096	
Strain, strain background (*Mus musculus*)	BXD48/RwwJ	JAX	RRID:IMSR_JAX:007097	
Strain, strain background (*Mus musculus*)	BXD48A	NA	NA	
Strain, strain background (*Mus musculus*)	BXD5/TyJ	JAX	RRID:IMSR_JAX:000037	
Strain, strain background (*Mus musculus*)	BXD50/RwwJ	JAX	RRID:IMSR_JAX:007099	
Strain, strain background (*Mus musculus*)	BXD51/RwwJ	JAX	RRID:IMSR_JAX:007100	
Strain, strain background (*Mus musculus*)	BXD55/RwwJ	JAX	RRID:IMSR_JAX:007103	
Strain, strain background (*Mus musculus*)	BXD60/RwwJ	JAX	RRID:IMSR_JAX:007105	
Strain, strain background (*Mus musculus*)	BXD61/RwwJ	JAX	RRID:IMSR_JAX:007106	
Strain, strain background (*Mus musculus*)	BXD62/RwwJ	JAX	RRID:IMSR_JAX:007107	
Strain, strain background (*Mus musculus*)	BXD63	NA	NA	
Strain, strain background (*Mus musculus*)	BXD65	NA	NA	
Strain, strain background (*Mus musculus*)	BXD66/RwwJ	JAX	RRID:IMSR_JAX:007111	

Appendix 1 Continued on next page

Appendix 1 Continued

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
Strain, strain background (*Mus musculus*)	BXD67/RwwJ	JAX	RRID:IMSR_JAX:007112	
Strain, strain background (*Mus musculus*)	BXD68/RwwJ	JAX	RRID:IMSR_JAX:007113	
Strain, strain background (*Mus musculus*)	BXD69/RwwJ	JAX	RRID:IMSR_JAX:007114	
Strain, strain background (*Mus musculus*)	BXD73/RwwJ	JAX	RRID:IMSR_JAX:007117	
Strain, strain background (*Mus musculus*)	BXD75/RwwJ	JAX	RRID:IMSR_JAX:007119	
Strain, strain background (*Mus musculus*)	BXD86/RwwJ	JAX	RRID:IMSR_JAX:007129	
Strain, strain background (*Mus musculus*)	BXD87/RwwJ	JAX	RRID:IMSR_JAX:007130	
Strain, strain background (*Mus musculus*)	BXH10/TyJ	JAX	RRID:IMSR_JAX:000032	
Strain, strain background (*Mus musculus*)	BXH14/TyJ	JAX	RRID:IMSR_JAX:000009	
Strain, strain background (*Mus musculus*)	BXH8/TyJ	JAX	RRID:IMSR_JAX:000076	
Strain, strain background (*Mus musculus*)	C3H/HeJ	JAX	RRID:IMSR_JAX:000659	
Strain, strain background (*Mus musculus*)	C57BL/6J	JAX	RRID:IMSR_JAX:000664	
Strain, strain background (*Mus musculus*)	C58/J	JAX	RRID:IMSR_JAX:000669	
Strain, strain background (*Mus musculus*)	CBA/J	JAX	RRID:IMSR_JAX:000656	
Strain, strain background (*Mus musculus*)	CE/J	JAX	RRID:IMSR_JAX:000657	
Strain, strain background (*Mus musculus*)	CXB12/HiAJ	JAX	RRID:IMSR_JAX:001633	
Strain, strain background (*Mus musculus*)	CXB2/ByJ	JAX	RRID:IMSR_JAX:000352	
Strain, strain background (*Mus musculus*)	DBA/2J	JAX	RRID:IMSR_JAX:000671	

Appendix 1 Continued on next page

Appendix 1 Continued

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
Strain, strain background (*Mus musculus*)	FVB/NJ	JAX	RRID:IMSR_JAX:001800	
Strain, strain background (*Mus musculus*)	LG/J	JAX	RRID:IMSR_JAX:000675	
Strain, strain background (*Mus musculus*)	LP/J	JAX	RRID:IMSR_JAX:000676	
Strain, strain background (*Mus musculus*)	MRL/MpJ	JAX	RRID:IMSR_JAX:000486	
Strain, strain background (*Mus musculus*)	NON/ShiLtJ	JAX	RRID:IMSR_JAX:002423	
Strain, strain background (*Mus musculus*)	NOR/LtJ	JAX	RRID:IMSR_JAX:002050	
Strain, strain background (*Mus musculus*)	NZB/BINJ	JAX	RRID:IMSR_JAX:000684	
Strain, strain background (*Mus musculus*)	PL/J	JAX	RRID:IMSR_JAX:000680	
Strain, strain background (*Mus musculus*)	SJL/J	JAX	RRID:IMSR_JAX:000686	
Software, algorithm	R version 4.1.1	R Development Core Team, 2016	https://www.R-project.org/	
Software, algorithm	Limma 3.32.2	***Ritchie et al., 2015***	https://bioconductor.org/packages/release/bioc/html/limma.html	
Software, algorithm	CoffeeProt	***Molendijk and Parker, 2021a***	https://www.coffeeprot.com	
Software, algorithm	TeaProt	***Molendijk et al., 2022***	https://tea.coffeeprot.com	
Software, algorithm	Mol* (Molstar)	***Sehnal et al., 2021***	https://molstar.org/	
Software, algorithm	WGCNA	***Langfelder and Horvath, 2008***	https://cran.r-project.org/web/packages/WGCNA/	
Software, algorithm	ColabFold (Alphafold2)	***Mirdita et al., 2022***	https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/AlphaFold2.ipynb	
Software, algorithm	Genebass	***Karczewski et al., 2022***	https://app.genebass.org/	
Software, algorithm	FoldX	***Delgado et al., 2019***	http://foldxsuite.crg.eu/	
Software, algorithm	PROVEAN	***Choi and Chan, 2015***	http://provean.jcvi.org/index.php	
Software, algorithm	DynaMut	***Rodrigues et al., 2018***	http://biosig.unimelb.edu.au/dynamut/	