

1 **Serum proteomic profiling of physical activity reveals CD300LG as a novel exerkin with a**
2 **potential causal link to glucose homeostasis**

3 Sindre Lee-Ødegård^{1,2}, Marit Hjorth^{3*}, Thomas Olsen^{3*}, Gunn-Helen Moen^{2,4,5,7}, Emily
4 Daubney⁴, David M Evans⁴⁻⁶, Andrea Hevener⁸, Aldons Jake Lusic^{9,10}, Mingqi Zhou¹¹, Marcus
5 Michael Seldin¹¹, Hooman Allayee^{12,13}, James R. Hilser^{12,13}, Jonas Krag Viken², Hanne L.
6 Gulseth¹⁴, Frode Norheim³, Christian A. Drevon¹⁵, Kåre I. Birkeland^{1,2}

7 ¹Department of Endocrinology, Morbid obesity and Preventive Medicine, Oslo University Hospital, Norway.

8 ²Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, Norway.

9 ³Department of Nutrition, Institute of Basic Medical Sciences, Faculty of Medicine, University of Oslo, Norway.

10 ⁴Institute for Molecular Bioscience, The University of Queensland, Brisbane, Australia

11 ⁵The Frazer Institute, The University of Queensland, 4102, Woolloongabba, QLD, Australia.

12 ⁶MRC Integrative Epidemiology Unit, University of Bristol, Bristol, UK.

13 ⁷Department of Public Health and Nursing, K.G. Jebsen Center for Genetic Epidemiology, NTNU, Norwegian
14 University of Science and Technology, Trondheim, Norway.

15 ⁸Division of Endocrinology, Department of Medicine, David Geffen School of Medicine, University of California
16 at Los Angeles, Los Angeles, California.

17 ⁹Department of Human Genetics, UCLA, Los Angeles, CA, USA.

18 ¹⁰Division of Cardiology, Department of Medicine, David Geffen School of Medicine at UCLA, Los Angeles, CA,
19 USA.

20 ¹¹Department of Biological Chemistry, University of California, Irvine, Irvine, USA.

21 ¹²Departments of Population & Public Health Sciences, Keck School of Medicine, University of Southern
22 California, Los Angeles, USA.

23 ¹³Department of Biochemistry & Molecular Medicine, Keck School of Medicine, University of Southern
24 California, Los Angeles, USA.

25 ¹⁴Department of Chronic Diseases and Ageing, Norwegian Institute of Public Health, Norway.

26 ¹⁵Vitas Ltd, Oslo Science Park, Oslo, Norway.

27 *Shared second authorship

28

29 Correspondence: Sognsvannsveien 9, Domus Medica, 0372 Oslo, Norway. Email:
30 sindre.lee@medisin.uio.no. Phone: +47 22 85 13 92. Fax: + 47 22 85 03 01.

31

32

33

34

35 **Abstract**

36 **Background:** Physical activity has been associated with preventing the development of type
37 2 diabetes and atherosclerotic cardiovascular disease. However, our understanding of the
38 precise molecular mechanisms underlying these effects remains incomplete and good
39 biomarkers to objectively assess physical activity are lacking.

40 **Methods:** We analyzed 3072 serum proteins in 26 men, normal weight or overweight,
41 undergoing 12 weeks of a combined strength and endurance exercise intervention. We
42 estimated insulin sensitivity with hyperinsulinemic euglycemic clamp, maximum oxygen
43 uptake, muscle strength, and used MRI/MRS to evaluate body composition and organ fat
44 depots. Muscle and subcutaneous adipose tissue biopsies were used for mRNA sequencing.
45 Additional association analyses were performed in samples from up to 47,747 individuals in
46 the UK Biobank, as well as using 2-sample Mendelian randomization and mice models.

47 **Results:** Following 12 weeks of exercise intervention, we observed significant changes in 283
48 serum proteins. Notably, 66 of these proteins were elevated in overweight men and
49 positively associated with liver fat before the exercise regimen, but were normalized after
50 exercise. Furthermore, for 19.7% and 12.1% of the exercise-responsive proteins,
51 corresponding changes in mRNA expression levels in muscle and fat, respectively, were
52 shown. The protein CD300LG displayed consistent alterations in blood, muscle, and fat.
53 Serum CD300LG exhibited positive associations with insulin sensitivity, and to angiogenesis-
54 related gene expression in both muscle and fat. Furthermore, serum CD300LG was positively
55 associated with physical activity and negatively associated with glucose levels in the UK
56 Biobank. In this sample, the association between serum CD300LG and physical activity was
57 significantly stronger in men than in women. Mendelian randomization analysis suggested
58 potential causal relationships between levels of serum CD300LG and fasting glucose, 2-hour
59 glucose after an oral glucose tolerance test, and HbA1c. Additionally, *Cd300lg* responded to
60 exercise in a mouse model, and we observed signs of impaired glucose tolerance in male,
61 but not female, *Cd300lg* knockout mice.

62 **Conclusion:** Our study identified several novel proteins in serum whose levels change in
63 response to prolonged exercise and were significantly associated with body composition,
64 liver fat, and glucose homeostasis. Serum CD300LG increased with physical activity and is a

65 potential causal link to improved glucose levels. CD300LG may be a promising exercise
66 biomarker and a therapeutic target in type 2 diabetes.

67

68 **Introduction**

69 Physical activity is a cornerstone in the prevention and treatment of several chronic diseases
70 like obesity, non-alcoholic fatty liver disease (NAFLD), atherosclerotic vascular disease, and
71 type 2 diabetes mellitus¹. Both acute and long-term exercise may enhance insulin sensitivity
72 and thereby improve glucose tolerance². Both resistance and endurance exercises enhance
73 insulin sensitivity, although the most pronounced effect is observed when combining these
74 training modalities³.

75

76 Metabolic adaptations to exercise encompass intricate inter-organ communication
77 facilitated by molecules referred to as exerkinines⁴⁻⁷. These exerkinines are secreted from
78 various tissues and include a variety of signal molecules released in response to acute
79 and/or long-term exercise with endocrine, paracrine and/or autocrine functions^{4,5}. Although
80 there has been considerable emphasis on exerkinines originating from skeletal muscle (SkM)
81^{8,9}, it is also known that exerkinines can originate from organs such as white^{6,10,11} and brown
82 adipose tissue¹² or the liver¹³. The established *bona fide* exerkinine, interleukin-6 (IL6), is
83 released during muscle contractions, contributing to improved overall glucose homeostasis
84^{4,14}. In addition, a range of other exerkinines are recognized, including IL7¹⁵, 12,13-diHOME¹²,
85 myonectin¹⁶, myostatin^{17,18}, METRNL¹⁹, CSF1⁸, decorin²⁰, SFRP4¹⁰, fetuin-A^{13,21}, and
86 ANGPTL4^{22,23}, among many others⁵.

87

88 Extensive screening aimed at discovering novel exercise responsive blood proteins have
89 faced considerable challenges, primarily due to the technical challenges in quantifying the
90 blood proteome on a large scale. However, recent advances in multi-plex technology, such as
91 the proximity extension assay (PEA), have made it possible to quantify more than 3000
92 proteins in blood samples more reliably than traditional untargeted mass spectrometry
93 (<https://olink.com/application/pea>). Some recent studies have used other proteomic
94 platforms, such as aptamer-based techniques (<https://somalogic.com/somascan-platform/>),
95 to show that acute and long-term aerobic exercise affected several hundred serum proteins

96 ²⁴⁻²⁸, but the downstream causal effects of such changes on clinical phenotypes are not
97 known. Furthermore, no studies have used the PEA technology to identify exerkinases
98 potentially underlying the mechanisms through which long-term physical activity, including
99 strength exercise, enhances glucose homeostasis.

100

101 We performed the ‘physical activity, myokines and glucose metabolism’ (MyoGlu) study ²⁹,
102 which was a controlled clinical trial aiming to identify novel secreted factors (‘exerkinases’)
103 that may serve as links between physical activity and glucose metabolism. We conducted a
104 comprehensive serum screening of 3072 proteins in normal weight and overweight men
105 both before and after combined endurance and strength exercise. Rigorous phenotyping
106 was carried out, including hyperinsulinemic euglycemic clamping, assessments of maximum
107 oxygen uptake, maximum muscle strength, and ankle-to-neck MRI/MRS scans.

108

109 Exerkinases identified with potential effects on glucose homeostasis in the MyoGlu study were
110 subsequently subject to analysis across several external data sets. Using data from 47,747
111 participants in the UK Biobank ³⁰, we assessed correlations between candidate proteins and
112 estimates of physical activity and glucometabolic outcomes. These associations were then
113 tested for causality using Mendelian randomization (MR). Exerkinases of interest were also
114 assessed in a knockout mouse model and in an exercise mouse model to further assess
115 potential links with glucose homeostasis.

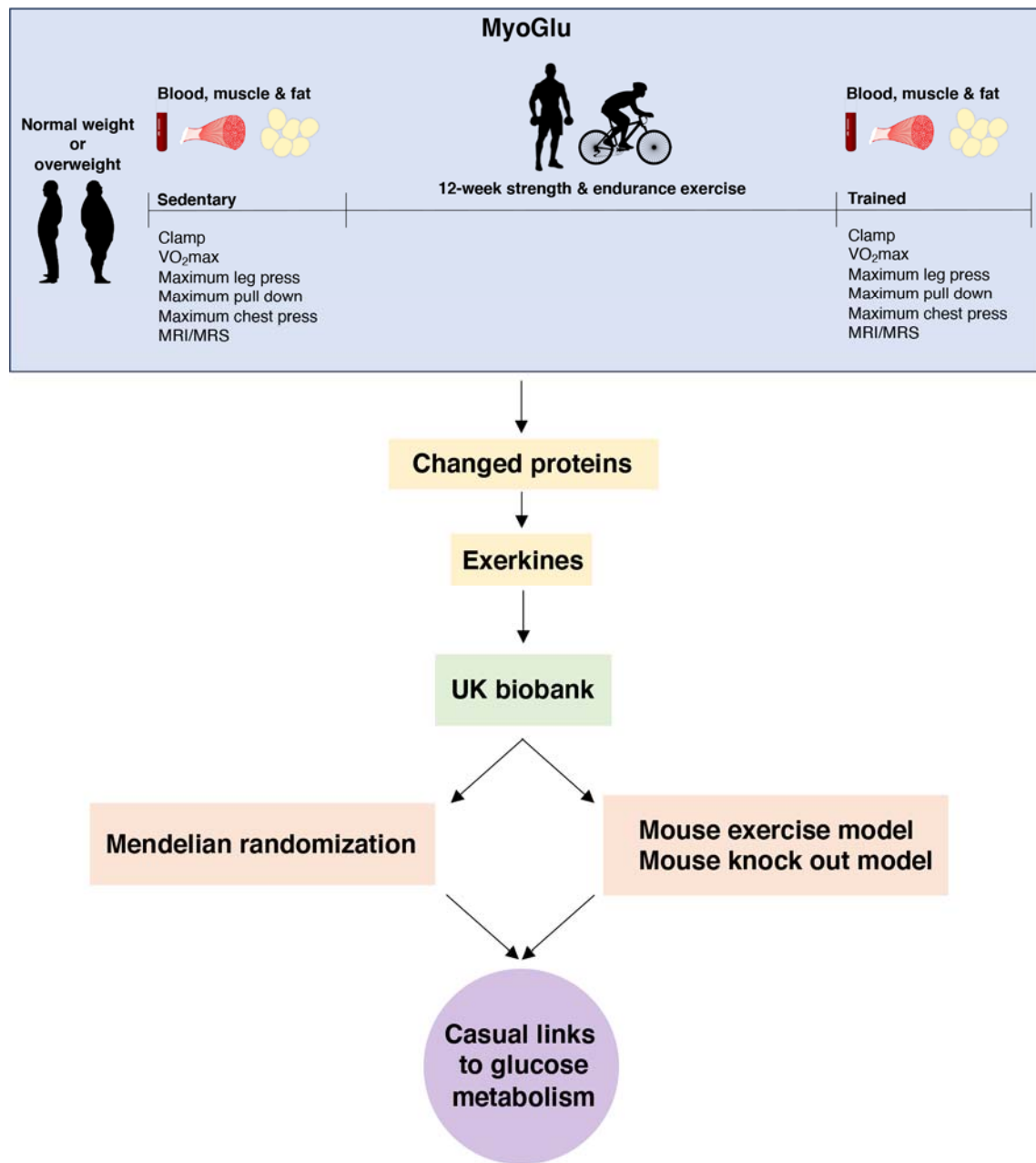
116

117 **Results**

118 **Cohort characteristics**

119 We studied 26 male participants, including 13 with normal weight, and another 13 with
120 overweight, as described previously ²⁹. They were subjected to 12-weeks of high-intensity
121 resistance and endurance exercise (Figure 1). The overweight participants had lower glucose
122 tolerance and insulin sensitivity compared to the normal weight participants
123 (Supplementary Table 1). After the 12-week intervention, body fat mass decreased and lean
124 body mass increased, together with significant improvements in insulin sensitivity (~40%),
125 maximum oxygen uptake and muscle strength in both groups (Supplementary Table 1).

126



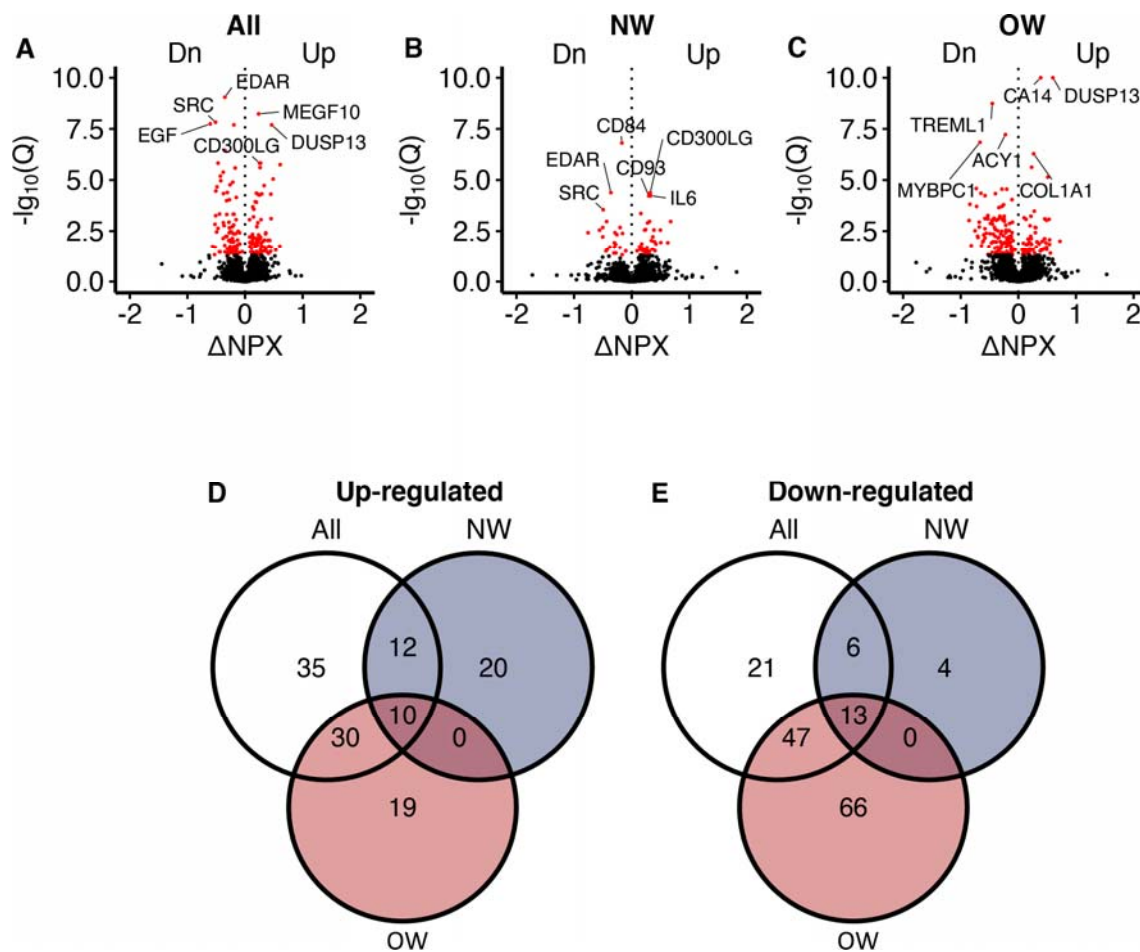
127

128 **Figure 1. Study overview.** We recruited sedentary men with either normal weight or
129 overweight for deep phenotyping before and after a prolonged exercise intervention. Multi-
130 omic analyses, including serum proteomics, clinical traits and muscle and fat transcriptomics
131 identified changed proteins and potential exerkines. Candidate exerkines were subsequently
132 analyzed in serum samples from the UK biobank and tested for associations with physical
133 activity and glucometabolic traits. Top candidates were then subjected to Mendelian
134 randomization and investigated in a mouse exercise model and in a mouse knock-out model
135 to assess casual links between exerkines and glucometabolic traits.

136 **Serum proteome responses to prolonged exercise**

137 Recognizing that circulating proteins could mediate exercise-induced metabolic
138 improvements, we next investigated alterations in the serum proteome in response to the
139 12-week intervention using PEA technology. Of the 3072 proteins quantified, we detected
140 increased serum concentrations of 126 proteins, and decreased serum concentrations of
141 157 proteins following the 12-week intervention, at a false discovery rate (FDR) below 5%
142 (Figure 2 A-C; Supplementary Table 2-4). Among these, 20 proteins increased exclusively in
143 normal weight men, whereas 19 proteins increased exclusively in overweight men (Figure 2
144 D). Four proteins were uniquely reduced in normal weight men, and 66 proteins were
145 uniquely reduced in overweight men (Figure 2 E).

146



147

148 **Figure 2. Serum proteomic responses to prolonged exercise.** (A) A volcano plot showing
 149 responses in all participants. The X-axis shows $\log_2(\text{fold-changes})$ and the Y-axis shows
 150 negative $\log_{10}(Q\text{-values})$. The red dots indicate statistical significance ($Q < 0.05$). Only the top
 151 3 up/down-regulated proteins are annotated. (B-C) Similar to A, but in normal weight and
 152 overweight men only. (D-E) Venn diagrams of the significant change in proteins shown in A-
 153 C. NPX = normalized protein expression. Q = P-values corrected using Benjamini-Hochberg's
 154 method. NW = normal weight. OW = overweight.

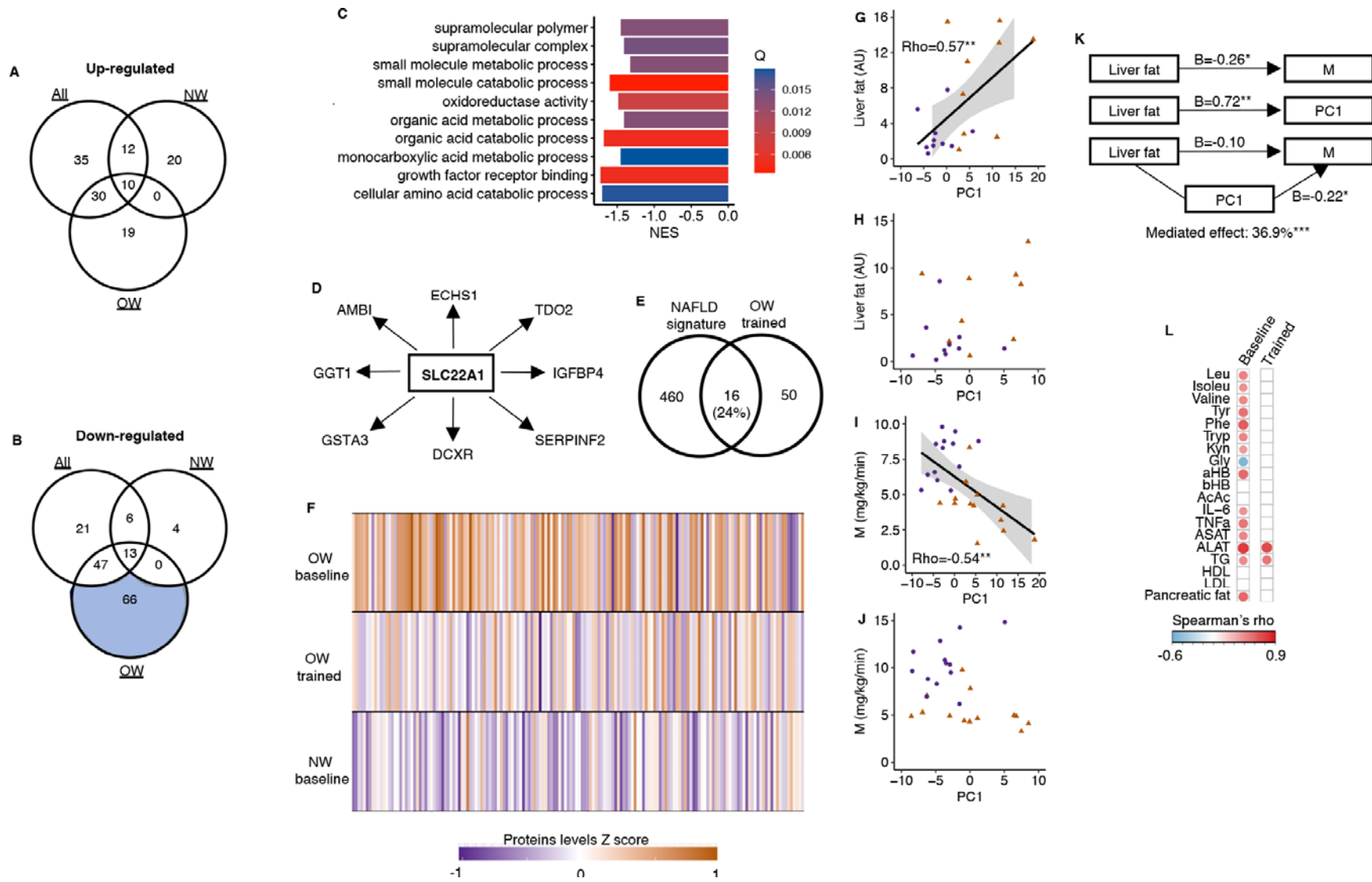
155 Several of the exercise-responsive proteins had potential roles in muscle adaptation and
156 metabolism. For example, platelet-derived growth factor subunit B (PDGFB) and IL7, are
157 both myokines with potential effects on muscle differentiation^{15,31}. Further, fibroblast
158 growth factor-binding protein 3 (FGFBP3) may influence running capacity³² and muscle
159 strength³³. NADH-cytochrome b5 reductase 2 (CYB5R2) can preserve SkM mitochondria
160 function in aging mice³⁴. FGFBP3 and switch-associated protein 70 (SWAP70) may protect
161 against weight gain³⁵ and cardiac hypertrophy³⁶, respectively. Finally, dual specificity
162 protein phosphatase 13 isoform A (DUSP13A) is highly specific to SkM³⁷, making it a
163 potential novel muscle-specific marker for long-term exercise. Detailed results for 2885
164 proteins in response to prolonged exercise are shown in Supplementary Table 2.

165

166 **A proteomic liver fat signature in overweight men**

167 In response to the 12-week exercise intervention, a larger number of serum proteins
168 responded in overweight men than in normal weight men (Figure 2 B-C). In particular, 66
169 proteins decreased in serum after 12 weeks in overweight men (Figure 2 E). Gene ontology
170 analyses revealed known pathways only for the proteins that decreased in overweight men
171 (Figure 3 A-B), and one of the most enriched pathways is related to metabolism of organic
172 acids (Figure 3 C). A key driver analysis of the 66 proteins identified SLC22A1, a hepatocyte
173 transporter related to liver fat content (Figure 3 D). Furthermore, the 66 proteins also
174 displayed a 24% overlap with a known human serum proteomic signature of non-alcoholic
175 fatty liver disease (NAFLD; Figure 3 E)³⁸, but no common proteins with signatures of specific
176 liver cells (The Human Liver Cell Atlas:³⁹). Baseline serum protein concentrations in the
177 identified signature of 66 proteins were higher among men with overweight compared to
178 those with normal weight, but were reduced or normalized in overweight men following
179 prolonged exercise (Figure 3 F). Using principal component analysis of the 66 proteins, the
180 first component correlated positively to liver fat content at baseline (Figure 3 G), but not
181 after prolonged exercise (Figure 3 H). Similarly, the first component also correlated positively
182 with several liver-related markers at baseline (Figure 3 L) and negatively to insulin sensitivity
183 at baseline (Figure 3 I), but not after prolonged exercise (Figure 3 J). The first component
184 mediated 37% of the association between baseline insulin sensitivity and liver fat content
185 (Figure 3 K). We observed no enrichments for the remaining proteins (Figure 3 A-B).

186

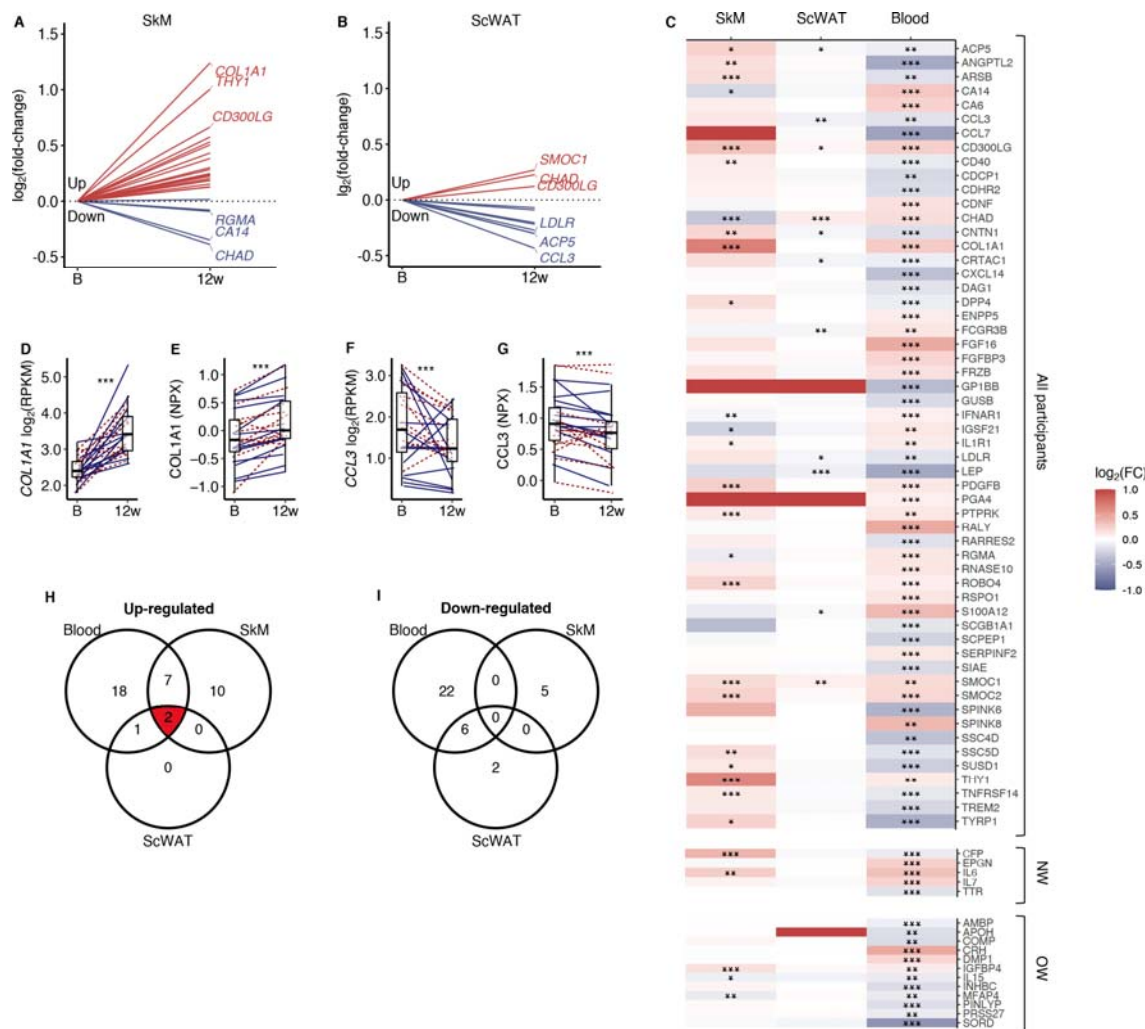


188 **Figure 3. A serum proteomic liver fat signature.** (A) No up-regulated proteins after prolonged exercise overlapped with known pathways. (B)
189 Only the 66 down-regulated proteins in the OW group overlapped with known pathways. (C) Top 10 gene sets overlapping with these 66
190 proteins. (D) SLC22A1 is a key driver among these 66 proteins. (E) These 66 proteins overlapped with a known human serum proteomic non-
191 alcoholic fatty liver disease signature from Govaere et al³⁸. (F) The down-regulated proteins in the OW group were elevated in OW vs. NW at
192 baseline but normalized in the OW group after prolonged exercise. The principal component of these 66 proteins correlated with (G) liver fat
193 content at baseline, but (H) not after prolonged exercise, with (I) the clamp M value at baseline, but (J) not after prolonged exercise. (K) The
194 principal component (PC) of these 66 proteins mediated 36.9% of the association between liver fat and M. (L) The principal component of these
195 66 proteins correlated with several liver-related markers at baseline, but not after prolonged exercise except for ASAT and ALAT (white = non-
196 significant, red/blue = significant). * $p < 0.05$ and ** $p < 0.01$.

197 **Secretory proteins**

198 Among the 96 up-regulated and 110 down-regulated serum proteins responding to the 12-
199 week exercise intervention (Figure 2 D-E), 37 are curated secretory proteins, and another 46
200 proteins are predicted as highly likely secretory proteins (Figure 4 A-C). We assessed the
201 corresponding mRNA responses in SkM and subcutaneous white adipose tissue (ScWAT)
202 following the 12-week intervention (Figure 4 A-C). In total, 19.7 % of the serum secretory
203 proteins displayed a directionally consistent significant change mRNA levels in SkM, whereas
204 12.1 % of the serum secretory proteins exhibited a corresponding mRNA response in ScWAT
205 (Figure 4 C). *COL1A1* was the most responsive SkM mRNA that also had a corresponding
206 increase in serum *COL1A1* after prolonged exercise (Figure 4 D-E). *CCL3* was the most
207 responsive ScWAT mRNA that also had a corresponding decrease in serum after prolonged
208 exercise (Figure 4 F-G). To prioritize proteins for follow-up analyses, we focused on *SMOC1*
209 and *CD300LG*, which had similar exercise responses in blood, SkM and ScWAT (Figure 4 H).
210 *SMOC1* is a known hepatokine with effects on insulin sensitivity in mice⁴⁰, but probably with
211 no causal link to insulin sensitivity in humans^{40,41}. Thus, we focused on *CD300LG* in
212 subsequent analyses.

213



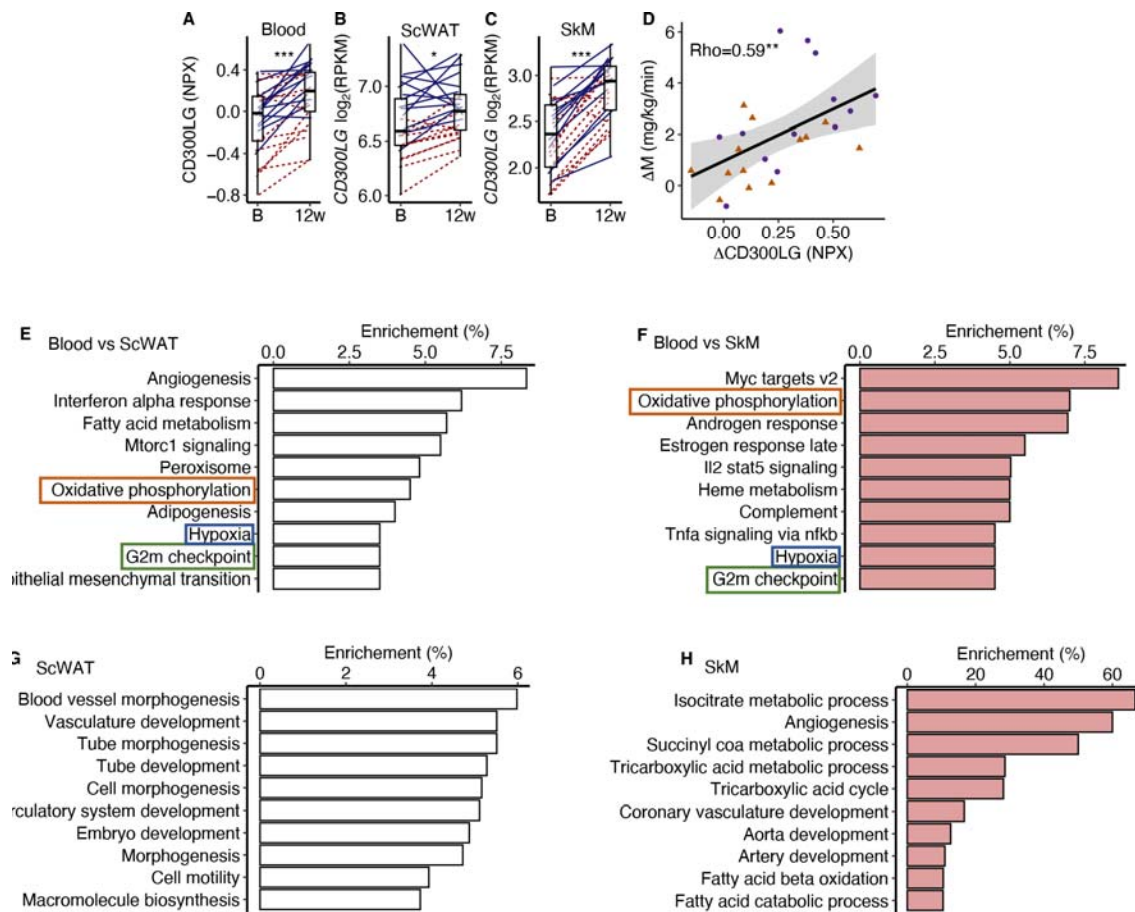
214
215
216
217
218
219
220
221
222
223
224
225

Figure 4. Comparison of secretory protein responses to prolonged exercise in blood with corresponding mRNA levels in skeletal muscle and adipose tissue. (A) mRNA levels in skeletal muscle and (B) adipose tissue for proteins that responded significantly to prolonged exercise. (C) A heatmap of $\log_2(\text{fold-changes})$ in blood, skeletal muscle and adipose tissue. (D) The most responding mRNA in skeletal muscle, and (E) the response in the blood protein. (F) The most responding mRNA in adipose tissue, and (G) the response in the blood. (H-I) Venn diagrams of significant changes in blood, skeletal muscle and adipose tissue. FC = fold-change. SkM = skeletal muscle. ScWAT = subcutaneous adipose tissue. NPX = normalized protein expression. RPKM = reads per kilobase per million mapped read. * $p < 0.05$, ** $p < 0.01$ and *** $p < 0.001$.

226 **CD300LG**

227 CD300LG displayed increased concentration in serum (+20%, $p < 0.001$) together with
228 increased levels in both SkM (+60%, $p < 0.001$) and scWAT (+13%, $p = 0.01$) mRNA following
229 the 12-week exercise intervention (Figure 5 A-C). Changes in serum CD300LG correlated
230 positively with changes in insulin sensitivity after the intervention ($\rho = 0.59$, $p = 0.002$; Figure
231 5 D). In addition, serum CD300LG concentration was lower in overweight than normal
232 weight men (-51%, $p = 0.014$) and positively correlated with insulin sensitivity before as well
233 as after the 12-week intervention (pre-trained: $r = 0.50$, $p = 0.001$, and post-trained: $r = 0.43$,
234 $p = 0.028$).

235



236

237

238 **Figure 5. CD300LG.** (A) The response from baseline to week 12 in serum CD300LG and

239 CD300LG mRNA in (B) subcutaneous adipose tissue (ScWAT) and (C) skeletal muscle (SkM).

240 (D) Correlation between the change from before to after prolonged exercise in serum

241 CD300LG and insulin sensitivity. (E-H) Pathway enrichment analyses were performed on the

242 top 500 most correlated (and $p < 0.05$) genes in (E) ScWAT or (F) SkM to the change in serum

243 CD300LG levels, or to the change in CD300LG mRNA levels in (G) ScWAT or (H) SkM. Only the

top 10 pathways with $Q < 0.05$ are presented. * $p < 0.05$, ** $p < 0.01$ and *** $p < 0.001$.

244

245 To investigate the potential effect of serum CD300LG on SkM and ScWAT, we performed an
246 overrepresentation analysis on the top 500 mRNAs that were positively correlated ($p < 0.05$)
247 with serum CD300LG levels in each tissue (Figure 5 E-H). Pathway analyses revealed that the
248 change in serum CD300LG concentrations correlated with changes in expression of genes
249 involved in oxidative phosphorylation, G2M check point and hypoxia both in ScWAT and SkM
250 (Figure 5 E-F). In ScWAT, serum CD300LG levels also showed the strongest enrichment with
251 angiogenesis pathways (Figure 5 E). In ScWAT, the change in ScWAT *CD300LG* mRNA levels
252 correlated positively with the change in ScWAT mRNA of genes related to
253 angiogenesis/vasculature development (Figure 5 G). Similar correlations between *CD300LG*
254 mRNA and angiogenesis genes were observed in SkM as well (Figure 5 H). For example, 60%
255 of the mRNAs in the angiogenesis pathway correlated with *CD300LG* (Figure 5 H). However,
256 serum CD300LG levels were also correlated positively with pathways related to fatty acid
257 metabolism in both ScWAT (Figure 5 E) and SkM (Figure 5 H).

258

259 To explore tissue specific expression of CD300LG, we utilized data from a publicly available
260 human tissue panel⁴². CD300LG is highly expressed in adipose tissue compared to other
261 tissues (Supplementary Figure 2 A) supporting our observation that ScWAT expression was
262 substantially higher than in SkM (Figure 5 B-C). To further investigate which cells in ScWAT
263 that express CD300LG, we utilized data from a single cell mRNA sequencing atlas of human
264 ScWAT (https://singlecell.broadinstitute.org/single_cell) generated by Emont *et al.*⁴³.
265 *CD300LG* mRNA in ScWAT was primarily expressed in venular endothelial cells, but not
266 adipocytes or other cell types present in ScWAT (Supplementary Figure 2 B-E).

267

268 We next explored whether CD300LG mediates tissue-tissue cross-talk using data from the
269 GD-CAT (Genetically-Derived Correlations Across Tissues) database^{44 45}, which is a tool for
270 analyzing human gene expression correlations in and across multiple tissues. In men, ScWAT
271 *CD300LG* correlated strongly with ScWAT, SkM and aortic gene expression (Supplementary
272 Figure 3). Consistent with our observations in the MyoGlu exercise intervention study, the
273 top network of gene expression in ScWAT related to ScWAT *CD300LG* mRNA was
274 angiogenesis (Supplementary Figure 3 A). Like ScWAT, SkM *CD300LG* also correlated strongly
275 with ScWAT, SkM and aortic gene expression (Supplementary Figure 3 B). The proteasome

276 complex was the top network of gene expression related to SkM *CD300LG* mRNA
277 Supplementary Figure 3 B). In contrast, running the same analyses in women did not reveal
278 associations between *CD300LG* and angiogenesis (Supplementary Figure 4 A-B).

279

280 We then evaluated serum *CD300LG* levels in up to 47,747 samples in the UK biobank (see
281 Methods). Descriptive statistics of the UK biobank cohort are presented in Supplementary
282 Table 5. Serum *CD300LG* levels were positively associated with several measures of physical
283 activity (all metabolic equivalent tasks, results from the international physical activity
284 questionnaire and meeting the recommended amount of weekly physical activity or not;
285 Table 1). Interestingly, serum *CD300LG* levels were most strongly related to vigorous activity
286 (Table 1). Furthermore, the associations between serum *CD300LG* and physical activity were
287 significantly stronger in men than in women (Table 1). Serum *CD300LG* levels were also
288 positively associated with fat mass and fat free mass, and negatively associated with
289 glucometabolic traits including serum glucose levels, Hb1Ac and the risk of having type 2
290 diabetes (Table 1). These associations were independent of body mass index (BMI).

291

292
293**Table 1.** Multiple regression analyses between serum CD300LG, and measures of physical activity and glucometabolic traits in the UK biobank.

	Women			Men			Interaction			Description	No. of women	No. of men
	Beta-estimate	SE	P	Beta-estimate	SE	P	Beta-estimate	SE	P			
NPX – physical activity												
MET per week all activity	1,0E-06	1,2E-06	0,384	4,5E-06	9,9E-07	<0.001	4,3E-06	1,5E-06	0,005	MET minutes per week	22527	20726
MET minutes walking	-6,0E-07	2,6E-06	0,818	-1,8E-06	2,6E-06	0,478	1,7E-07	3,7E-06	0,962	MET minutes per week	22527	20726
MET minutes moderate activity	-2,5E-06	2,4E-06	0,294	1,9E-07	2,3E-06	0,932	1,9E-06	3,3E-06	0,554	MET minutes per week	22527	20726
MET minutes vigorous activity	1,0E-05	2,8E-06	<0.001	2,2E-05	2,1E-06	<2e-16	1,5E-05	3,5E-06	<0.001	MET minutes per week	22527	20726
Sedentary overall average	0,077	0,053	0,147	0,042	0,054	0,441	-0,100	0,075	0,185	Proportion sedentary activity.	7430	5825
Light overall average	-0,119	0,062	0,053	-0,420	0,071	<0.001	-0,232	0,094	0,013	Proportion light activity.	7430	5825
Moderate/vigorous overall average	0,633	0,171	<0.001	1,357	0,194	<0.001	1,043	0,254	<0.001	Proportion moderate/vigorous activity.	7430	5825
IPAQ activity group	9,6E-03	3,8E-03	0,012	3,1E-02	3,8E-03	<0.001	2,6E-02	5,4E-03	<0.001	IPAQ category	22527	20726
Summed days activity	1,9E-03	5,9E-04	<0.001	3,8E-03	5,7E-04	<0.001	2,7E-03	8,1E-04	<0.001	Days performing walking, moderate and vigorous activity	23199	21138
Summed minutes activity	-1,8E-05	3,0E-05	0,550	6,4E-05	2,7E-05	0,017	9,9E-05	4,0E-05	0,013	Mins performing walking, moderate and vigorous activity	22527	20726
Moderate/vigorous recommendation1	7,5E-03	5,6E-03	0,186	4,8E-02	5,8E-03	< 2e-16	4,6E-02	8,0E-03	<0.001	Yes/No	22527	20726
Moderate/vigorous walking recommendation1	-1,3E-05	7,2E-03	0,999	4,0E-02	7,4E-03	<0.001	4,6E-02	1,0E-02	<0.001	Yes/No	22521	20723
Trait – NPX												
Body fat percentage impedance	-0,137	0,051	0,007	-0,552	0,052	<0.001	-0,375	0,073	<0.001	Body fat percentage	28099	23802
Whole body fat mass impedance	0,345	0,049	<0.001	0,023	0,051	0,656	-0,209	0,071	0,003	Fat mass (kg)	28108	23866
Whole body fat free mass impedance	0,450	0,050	<0.001	1,045	0,090	<0.001	0,622	0,101	<0.001	Fat free mass (kg)	28107	23870
Body mass index	-	-	-	-	-	-	-	-	-	kg/m2	24810	21509
Glucose	-0,066	0,016	<0.001	-0,041	0,022	0,062	0,033	0,026	0,202	mmol/L	27271	23294
HbA1c	-0,898	0,078	<0.001	-0,911	0,107	<0.001	-0,010	0,130	0,936	mmol/mol	27323	23292
Triglycerides	-0,399	0,011	<0.001	-0,413	0,017	<0.001	-0,058	0,020	0,004	mmol/L	28387	24332
Type 2 diabetes	-0,012	0,002	<0.001	-0,018	0,004	<0.001	0,000	0,004	0,982	Yes/No.	24802	21483
TyG	-2,189	0,074	<0.001	-2,228	0,120	<0.001	-0,203	0,136	0,138	mmol/L x mmol/L	22527	23841

MET = metabolic equivalent of task. NPX = Normalized protein expression. SE = standard error. IPAQ = International physical activity questionnaire. TyG = triglycerid glucose index on insulin resistance.

Model 1 (NPX – physical activity) was a linear regression model of NPX values as a function of a measure of physical activity.

Model 2 (Trait – NPX) the measures of body composition and glucometabolic traits were the outcomes and NPX values were set as the exposure.

Models 1 and 2 were adjusted for age, batch, study centre, storage time and BMI.

¹Indicates whether a person met the 2017 UK Physical activity guidelines of 150 minutes of moderate activity per week or 75 minutes of vigorous activity.

294
295
296
297
298
299

300

301

302 **GWAS of serum CD300LG levels**

303 GWAS analyses of CD300LG levels detected 43 independent genome-wide significant genetic
304 associations across the genome (Supplementary Figure 5). The genomic inflation factor
305 ($\lambda=1.0966$) and LD score intercept (1.039) were consistent with our GWAS being well
306 controlled for population stratification and other possible biases (Supplementary Figure 6).
307 The most significant SNPs lay along chromosome 17 with these SNPs mapping to the
308 genomic region encoding the *CD300LG* gene (Supplementary Figure 5). Follow-up analyses
309 revealed three significant, independent *cis*-pQTLs associated with the protein CD300LG
310 (Supplementary Table 6) and a number of *trans*-pQTLs (Supplementary Table 7).

311

312 *Mendelian randomization (MR) analysis*

313 The independent genome-wide significant SNPs from the CD300LG GWAS were used for
314 two-sample MR (see Methods for details), where 39 SNPs were also available in the
315 outcome GWAS⁴⁶. We first performed Inverse Variance Weighted (IVW) MR analysis to test
316 the causal relationship between CD300LG and fasting glucose, 2-hour post oral glucose
317 tolerance test (OGTT) glucose levels and HbA1c using only *cis*-SNPs (Supplementary Table 8)
318 and all SNPs (Supplementary Table 9). The *cis* IVW MR analysis showed some evidence for a
319 negative causal effect of CD300LG on fasting insulin ($p=0.01$), but due to only three SNPs in
320 these analyses, we could not perform additional sensitivity analyses (except for tests of
321 heterogeneity in estimates of the causal effect across SNPs) and could not determine
322 whether the absence of strong evidence for a causal effect on the glycaemic parameters was
323 genuine or whether our analyses just lacked power. Although some of the analyses involving
324 all the genome-wide significant SNPs indicated a potential causal link between increased
325 serum CD300LG concentration and these outcomes (Supplementary Table 9) the analysis
326 showed significant heterogeneity. We did not detect strong evidence of directional
327 pleiotropy (significant MR Egger intercept, Supplementary Table 9). The heterogeneity in the
328 analysis is possibly due to the fact that many of the SNPs found in the GWAS of CD300LG are
329 associated with related phenotypes which could exert pleiotropic effects on diabetes related
330 outcomes, and so the results should be interpreted with care (Supplementary Table 10). Due
331 to the heterogeneity in our results we therefore performed MR PRESSO to account for
332 outliers. The MR PRESSO analysis showed a significant negative effect of CD300LG on all
333 outcomes (Table 2).

334

335 **Table 2.** Mendelian randomization of serum CD300LG levels and glucose outcomes
336 using MR PRESSO.

Outcome	MR analysis	Number of outliers	Effect	SD	p-value
2-hour post OGTT glucose (mmol/L)	Raw		-0.3722	0.0998	6.2x10 ⁻⁴
2-hour post OGTT glucose (mmol/L)	Outlier-corrected	2	-0.3049	0.0855	1.04x10 ⁻²
Fasting glucose (mmol/L)	Raw		-0.0307	0.0358	0.3963
Fasting glucose (mmol/L)	Outlier-corrected	2	-0.0556	0.0133	1.73x10 ⁻⁴
Fasting insulin (pmol/L)	Raw		-0.0870	0.0558	0.1271
Fasting insulin (pmol/L)	Outlier-corrected	10	-0.0534	0.0252	0.0432
HbA1c (%)	Raw		-0.0485	0.0155	3.28x10 ⁻³
HbA1c (%)	Outlier-corrected	3	-0.0560	0.0155	1.04x10 ⁻⁴

337 For detailed results see Supplementary Table 8 and 9. Fasting glucose adjusted for BMI
338 n=200622, 2-hour post oral glucose tolerance test (OGTT) glucose adjusted for BMI n=63396,
339 fasting insulin adjusted for BMI n=151013, and HbA1c n=146806.

340

341 *Mouse models*

342 To functionally validate association of CD300LG with metabolic homeostasis, we leveraged
343 phenotypic data for exercising mice and for *Cd300lg* deficient (*Cd300lg*^{-/-}) mice that both
344 were publicly available through the MoTrPAC⁴⁷ study and the international mouse
345 phenotyping consortium (PhenoMouse)⁴⁸.

346

347 There is a 51% homology between human *CD300LG* and mouse *Cd300lg*⁴⁹, and also in mice,
348 *Cd300lg* is predominantly expressed in adipose tissue endothelial cells⁴³.

349

350 In MoTrPAC⁴⁷, *Cd300lg* levels in scWAT from n=12-15 male and female mice were increased
351 after exercise for 8 weeks (~30% in both female (p=0.03) and male (p=0.01) mice)
352 (Supplementary Figure 7 A-C). Based on data from n=3050 mice from PhenoMouse male,
353 but not female, mutants for the *Cd300lg*^{tm1a(KOMP)Wtsi} allele displayed impaired glucose
354 tolerance (Supplementary Figure 7 D), but no change in fasting glucose and insulin
355 (Supplementary Figure 7 E-F). Mutant male, but not female, mice also displayed increased
356 lean mass (Supplementary Figure 7 G) and less fat mass (Supplementary Figure 7 H).
357 Detailed PhenoMouse results are presented in Supplementary Table 11.

358

359 Discussion

360 In the present study, we characterized the effects of strength and endurance exercise on the
361 serum proteome of sedentary normal weight and overweight men. We identified significant
362 changes in 283 serum proteins related to many signaling pathways after the 12-week
363 intervention. Some of these proteins were related to the mitochondria, muscle
364 differentiation and exercise capacity. Among known secretory proteins, 19.7% and 12.1%
365 displayed corresponding mRNA changes in SkM and ScWAT, respectively. Although some
366 proteins may be myokines, others may be adipokines or other types of exerkinines. A multi-
367 tissue responding protein was CD300LG, which also correlated positively to insulin
368 sensitivity. CD300LG was particularly interesting because we could replicate the finding in an
369 external cohort, find evidence of a causal link to glucose homeostasis, and perform
370 functionally validation in mice models. Furthermore, the association between CD300LG,
371 physical activity and glycemic traits might display sex dimorphic relationships.

372

373 One of the protein signatures observed in response to exercise was based on strong
374 associations with markers of liver fat content in overweight men. This was related to
375 SLC22A1, which regulates the hepatic glucose-fatty acid cycle affecting gluconeogenesis and
376 lipid metabolism⁵⁰, and may influence liver fat accumulation⁵¹. This signature also shared
377 many common proteins with a known serum NAFLD proteomic signature³⁸. However, we
378 did not detect overlaps between proteins in this signature and specific gene expression
379 patterns of liver cells (e.g. hepatocytes, immune cells etc.)³⁹. This observation suggests that
380 the proteins detected do not relate to liver protein synthesis per se, but may accumulate in
381 serum due to being released in the blood stream as a result of impaired liver protein
382 catabolism or cell damage as a consequence of overweight/obesity. Notably, this protein
383 signature in overweight men normalized after 12 weeks of exercise and resembled the
384 signature observed in normal weight men. These data suggest prolonged exercise leads to
385 improvements of liver function in overweight men.

386

387 Several proteins responding to prolonged exercise had a known signal sequence. These
388 secretory proteins are of particular interest because they could mediate inter-tissue
389 adaptations to exercise. For example, COL1A1 was substantially increased in serum and its
390 corresponding mRNA level was increased in SkM. However, COL1A1 is a collagen peptide

391 that is related to muscle damage, turn-over and extracellular matrix remodeling in response
392 to exercise⁵² and may mostly reflect muscle restructuring and not represent signaling effects
393 to distant tissues. The large overlap between serum proteins and SkM mRNA most likely
394 suggests a similar phenomenon, where tissue restructuring following exercise is reflected in
395 blood. However, there are probably also several myokines with distant signaling effects
396 among the identified proteins. CCL3 was reduced in serum in parallel with a reduction in its
397 mRNA level in ScWAT. CCL3 is a monocyte chemoattractant protein that may be related to
398 immune cell infiltration in adipose tissue⁵³. Hence, this may reflect a positive effect of
399 prolonged exercise on adipose tissue inflammation, which is in line with our previous results
400 showing normalization of adipose tissue inflammation following prolonged exercise¹⁰.

401

402 A particularly interesting protein was CD300LG, which responded to prolonged exercise in
403 serum, and, judged by its mRNA levels, in SkM and ScWAT. Serum CD300LG levels were
404 lower in overweight compared to normal weight men. Furthermore, the exercise-induced
405 response in CD300LG correlated positively to improvements in insulin sensitivity, and there
406 was also a significant correlation between serum CD300LG and insulin sensitivity both
407 before and after the intervention. We therefore analyzed CD300LG in an external data set,
408 the UK Biobank, and again we observed positive associations between especially vigorous
409 exercise and serum CD300LG. Moreover, serum CD300LG levels were negatively associated
410 with glucose levels and type 2 diabetes in the UK Biobank, and these associations might be
411 causal based on MR analysis. These findings were functionally corroborated by the
412 alterations in glucose tolerance and parameters related to insulin sensitivity observed in
413 *Cd300lg*^{-/-} mice. Thus, CD300LG may represent an exerkin with a causal link to glucose
414 homeostasis. However, whether CD300LG can mediate tissue-tissue crosstalk is unknown.
415 CD300LG is a cell surface protein with a transmembrane domain, but is also a predicted
416 secretory protein⁵⁴. Whether the protein is released from the cell surface in a regulated
417 manner to mediate cross-tissue signaling needs further investigation. Furthermore, the
418 exact link between CD300LG and glucose metabolism is not clear, but possibly related to the
419 fact that CD300LG is expressed in endothelial cells⁵⁵, linked to blood pressure⁵⁶,
420 lymphocyte binding⁵⁷, blood triacylglycerol levels^{58,59} and molecular traffic across the
421 capillary endothelium⁶⁰. Both MyoGlu and GD-CAT data also suggested that CD300LG may
422 be related to angiogenesis in ScWAT⁶¹ and SkM^{61,62}, at least in men. Hence, we speculate

423 that the link between CD300LG and glucose metabolism is related to improved tissue
424 capillarization/vascular function following prolonged exercise. Furthermore, since vigorous
425 exercise leads to angiogenesis in ScWAT and SkM⁶¹, serum CD300LG may be a marker of
426 exercise intensity.

427

428 *Strengths and limitations*

429 Although MyoGlu included only 26 sedentary men, they were extensively phenotyped with
430 euglycemic hyperinsulinemic clamp, fitness tests, whole body imaging (MRI/MRS) and
431 mRNA sequencing of ScWAT and SkM. We also supplied our study with data from 47747
432 persons in the UK biobank to enhance the validity and generalization of the results.
433 Furthermore, to assess sex differences we stratified analyses for men and women in the UK
434 Biobank, in external data bases (GD-CAT⁴⁵) and analyzed data from both male and female
435 mice. Since correlations with the clamp data only imply a role for a protein with regard to
436 glucose homeostasis, so we also tested associations with related glucometabolic traits in the
437 UK Biobank and tested these associations for causality using MR. We also included data from
438 exercised mice, and mutant mice to further strengthen the results. Our serum proteome
439 study assessed 3072 proteins, and therefore we do not cover the complete human
440 proteome. However, the Olink platform is based on dual recognition of correctly matched
441 DNA-labeled antibodies, and DNA sequence-specific protein-to-DNA conversion to generate
442 a signal. This is a highly scalable method with an exceptional specificity
443 (<https://olink.com/application/pea/>). Previous exercise-proteomic studies has looked at
444 ~600 proteins in overweight men after endurance exercise²⁵, and three papers were
445 published from the HERITAGE study analyzing ~5000 proteins in response to endurance
446 exercise²⁶⁻²⁸. However, our study is the first and largest exercise study using PEA in both
447 overweight and normal weight men, and also including strength exercise.

448

449 However, CD300LG's role related to angiogenesis is only suggested through association
450 analyses in our data, necessitating follow-up studies to confirm any causal role of CD300LG
451 in angiogenesis. Although the open-source cd300lgtm1a(KOMP)Wtsi mice provided
452 interesting indications, a future study should directly phenotype mice with alterations in the
453 CD300LG gene, and measure the effects on circulating CD300LG levels and potential
454 regulatory mechanisms related to angiogenesis and glucose tolerance. Furthermore, it is

455 also unknown if circulating CD300LG is full-length or a cleaved fragment, and the
456 mechanisms for CD300LG secretion should be further studied *in vitro*. Finally, future
457 experiments should also identify the epitope for O-link binding, and confirm its specificity
458 using targeted mass spectrometry or antibody-based validations.

459

460 *Conclusion*

461 Our study provided a detailed analysis of serum proteins responding to three months of
462 strength and endurance exercise in both normal weight and overweight men. Our results
463 identified a novel NAFLD-related serum protein signature in overweight men that was
464 normalized after prolonged exercise. We also identified hundreds of tissue-specific and
465 multi-tissue serum markers of e.g., mitochondrial function, muscle differentiation, exercise
466 capacity and insulin sensitivity. Our results were enriched for secretory proteins (exerkines),
467 such as CD300LG, which may be a marker of exercise intensity especially in men, and may
468 also have causal roles in improved glucose homeostasis after physical activity.

469

470 **Methods**

471 The MyoGlu study was conducted as a controlled clinical trial (clinicaltrials.gov:
472 NCT01803568) and was carried out in adherence to the principles of the Declaration of
473 Helsinki. The study received ethical approval from the National Regional Committee for
474 Medical and Health Research Ethics North in Tromsø, Norway, with the reference number
475 2011/882. All participants provided written informed consent before undergoing any
476 procedures related to the study. The UK biobank has ethical approval from the North West
477 Multi-Centre Research Ethics Committee (MREC), which covers the UK, and all participants
478 provided written informed consent. This particular project from the UK biobank received
479 ethical approval from the Institutional Human Research Ethics committee, University of
480 Queensland (approval number 2019002705).

481

482 *Participants*

483 The MyoGlu study enrolled men aged 40 to 65 years who were healthy but sedentary
484 (having engaged in fewer than one exercise session per week in the previous year)^{29,63}.
485 These participants were divided into two groups based on their BMI and glucose tolerance:
486 overweight (with an average BMI of 29.5 ± 2.3 kg/m²) and normal weight controls (with an

487 average BMI of 23.6 ± 2.0 kg/m²). The overweight men had reduced glucose tolerance
488 and/or insulin sensitivity (Supplementary Table 1). Both groups, consisting of 13 individuals
489 each, underwent a 12-week regimen of combined strength and endurance training.

490

491 *Exercise protocols*

492 This 12-week training intervention included two weekly sessions of 60 minutes each for
493 endurance cycling and two sessions of 60 minutes each for whole-body strength training.
494 Prior to and after the 12-week exercise intervention, a 45-minute bicycle test at 70% of their
495 maximum oxygen uptake (VO₂max) was conducted. Serum, muscle (*m. vastus lateralis*) and
496 subcutaneous white adipose tissue biopsies were taken before, and 48 hours after the last
497 exercise session of the 12-week intervention period^{29,63}.

498

499 *Clinical data*

500 The euglycaemic hyperinsulinemic clamp was performed after an overnight fast^{29,63}. A fixed
501 dose of insulin 40 mU/m²·min⁻¹ was infused, and glucose (200 mg/mL) was infused to
502 maintain euglycaemia (5.0 mmol/L) for 150 min. Insulin sensitivity is reported as glucose
503 infusion rate (GIR) during the last 30 min relative to body weight. Whole blood glucose
504 concentration was measured using a glucose oxidase method (YSI 2300, Yellow Springs, OH)
505 and plasma glucose concentration was calculated as whole blood glucose x 1.119. Magnetic
506 resonance imaging/spectroscopy (MRI/MRS) methods were used to quantify fat and lean
507 mass. The ankle-to-neck MRI protocol included a 3D DIXON acquisition providing water and
508 lipid quantification, data were then analysed using the nordicICE software package
509 (NordicNeuroLab, Bergen, Norway), and the jMRUI workflow. VO₂max tests were performed
510 after standardized warm-up at a workload similar to the final load of an incremental test in
511 which the relationship between workload (watt) and oxygen uptake was established.
512 Participants cycled for one minute followed by a 15 watt increased workload every 30 seconds until
513 exhaustion. Test success was based on O₂ consumption increased <0.5 mL·kg⁻¹·min⁻¹ over a
514 30 watt increase in workload, respiratory exchange ratio values >1.10, and blood
515 lactate >7.0 mmol/L. We obtained scWAT, SkM biopsies and blood samples as described
516 previously²⁹. Biopsies were obtained from the periumbilical subcutaneous tissue and from
517 *m. vastus lateralis*. After sterilization, a lidocaine based local anaesthetic was injected in the
518 skin and sub cutis prior to both SkM and scWAT biopsies. Biopsies were dissected on a cold

519 aluminium plate to remove blood etc. before freezing. For standard serum parameters,
520 measurement were either conducted using standard in-house methods or outsourced to a
521 commercial laboratory (Frst Laboratories, Oslo, Norway).

522

523 *The Olink proteomics explorer 3072 platform*

524 We utilized antibody-based technology (Olink Proteomics AB, Uppsala, Sweden) to conduct
525 profiling of serum samples through the Olink Explore 3072 panel. This PEA technique
526 involves using pairs of DNA oligonucleotide-labeled antibodies to bind to the proteins of
527 interest. When two matching antibodies attach to a target protein, the linked
528 oligonucleotides hybridize and are extended by DNA polymerase, forming a unique DNA
529 "barcode". This barcode is then read using next-generation sequencing. The specificity and
530 sensitivity of the PEA technology are notably high because only accurately matched DNA
531 pairs generated detectable and measurable signals. To refine the dataset, proteins that were
532 not detected or were duplicated were removed, resulting in an analysis of 2886 proteins.
533 Only a single assay was conducted, eliminating inter-assay variability. Data are presented as
534 normalized protein expression (NPX) units, which are logarithmically scaled using a \log_2
535 transformation.

536

537 *Proteomics validations*

538 Duplicate measurements of IL6 and leptin in plasma were conducted using ELISA kits (Leptin;
539 Camarillo, CA and IL6; R&D Systems, Minneapolis, MN) following the manufacturer's
540 instructions. The correlations between PEA or enzyme linked immunosorbent (ELISA) assays
541 were $r=0.94$ ($p=1.4 \times 10^{-11}$), and $r=0.92$ ($p=2.2 \times 10^{-11}$) for IL6 and leptin, respectively
542 (Supplementary Figure 1).

543

544 *mRNA sequencing*

545 Biopsies were frozen in liquid nitrogen, crushed to powder by a pestle in a liquid nitrogen-
546 cooled mortar, transferred into 1 mL QIAzol Lysis Reagent (Qiagen, Hilden, Germany), and
547 homogenized using TissueRuptor (Qiagen) at full speed for 15 sec, twice^{29,63}. Total RNA was
548 isolated from the homogenate using miRNeasy Mini Kit (Qiagen). RNA integrity and
549 concentration were determined using Agilent RNA 6000 Nano Chips on a Bioanalyzer 2100
550 (Agilent Technologies Inc., Santa Clara, CA). RNA was converted to cDNA using High-Capacity

551 cDNA Reverse Transcription Kit (Applied Biosystems, Foster, CA). The cDNA reaction mixture
552 was diluted in water and cDNA equivalent of 25 ng RNA used for each sample. All muscle
553 and scWAT samples were deep-sequenced using the Illumina HiSeq 2000 system with
554 multiplex at the Norwegian Sequencing Centre, University of Oslo. Illumina HiSeq RTA (real-
555 time analysis) v1.17.21.3 was used. Reads passing Illumina's recommended parameters were
556 demultiplexed using CASAVA v1.8.2. For prealignment quality checks, we used the software
557 FastQC v0.10.1. The mean library size was ~44 millions unstranded 51 bp single-ended reads
558 for muscle tissue and ~52 millions for scWAT with no differences between groups or time
559 points. No batch effects were present. cDNA sequenced reads alignment was done using
560 Tophat v2.0.8, Samtools v0.1.18, and Bowtie v2.1.0 with default settings against the UCSC
561 hg19 annotated transcriptome and genome dated 14th of May 2013. Post-alignment quality
562 controls were performed using the Integrative Genome Viewer v2.3 and BED tools v2.19.1.
563 Reads were counted using the intersection strict mode in HTSeq v0.6.1.

564

565 *Statistics and bioinformatics*

566 Olink data were analyzed using the 'AnalyzeOlink' R package for pre-processing, testing using
567 mixed linear regression and annotation. Pathway and gene ontology overrepresentation
568 analyses were performed using MSigDB data sets (Hallmark pathways and biological
569 processes). mRNA sequencing data were normalized as reads per kilobase per million
570 mapped read (RPKM) and analyzed using mixed linear regression from the 'lme4' R package.
571 Normality was determined by quantile-quantile plots. P-values were corrected using the
572 Benjamini-Hochberg approach set at a false discovery rate (FDR) of 5%. For univariate
573 correlations, Pearsons' or Spearman's method were applied as appropriate. Principal
574 component analysis was performed using the 'prcomp' R package. Key driver analysis was
575 performed using the 'Mergeomics' R package. Mediation analysis was performed using the
576 'Mediation' R package with 1000 bootstraps and the *set.seed* function to ensure
577 reproducibility.

578

579 *UK Biobank*

580 The UK biobank is a large prospective population-based cohort containing ~500,000
581 individuals (~273,000 women), with a variety of phenotypic and genome-wide genetic data
582 available ⁶⁴. The UK biobank has ethical approval from the North West Multi-Centre

583 Research Ethics Committee (MREC), which covers the UK, and all participants provided
584 written informed consent.

585

586 We utilized imputed genetic data from the October 2019 (version 3) release of the UK
587 biobank for our analyses (Application ID: 53641). In addition to the quality control metrics
588 performed centrally by the UK biobank⁶⁵, we defined a subset of unrelated “white
589 European” individuals. We excluded those with putative sex chromosome aneuploidy, high
590 heterozygosity or missing rate, or a mismatch between submitted and inferred sex as
591 identified by the UK biobank (total N = 1932). We excluded individuals who we did not
592 identify as ancestrally European using K-means clustering applied to the first four genetic
593 principal components generated from the 1000 Genomes Project⁶⁶. We also excluded
594 individuals who had withdrawn their consent to participate in the study as of February 2021.

595

596 The Olink proteomics explorer 1536 platform

597 All analysis were done using the UK Biobank Olink data containing a total of 58699 samples
598 and 54309 individuals, after excluding individuals as mentioned above we had 47747
599 samples with measured serum CD300LG levels. Data was generated according to Olink’s
600 standard procedures.

601

602 Observational analyses

603 For the physical activity measurements we investigated if the degree of physical activity was
604 associated with serum levels of protein (serum levels of protein regressed on physical
605 activity), alternatively for the metabolic measurements we investigated if the protein
606 expression affected the metabolic measurements (trait regressed on serum levels of
607 protein), for both we used a linear regression model. We performed analyses stratified by
608 sex and adjusting for age, protein batch, UK Biobank assessment centre, the time the sample
609 was stored and BMI. All analyses were performed in R version 3.4.3.

610

611 Genome-wide association analysis

612 A GWAS of serum CD300LG levels (\log_2 transformed) measured in the UK Biobank was
613 performed using BOLT-LMM⁶⁷ on individuals of European descent who had proteomic data
614 available (N=45788). We included sex, year of birth, protein and genotyping batch, time

615 from sample collection to processing time (in weeks) and five ancestry informative principal
616 components as covariates in the analysis.

617

618 Post GWAS quality control included the removal of SNPs with MAF ≤ 0.05 and info score ≤ 0.4
619 ($n_{\text{SNPs}}=6,945,819$). The previously generated LD reference panel for clumping consisted of a
620 random sample of 47674 unrelated British UK biobank individuals identified using GCTA⁶⁸
621 with identity by state (IBS) <0.025 and identity by descent (IBD) sharing of $\pi < 0.1$. LD score
622 regression analysis⁶⁹ was used to investigate whether genomic inflation was likely due to
623 polygenicity or population stratification/cryptic relatedness.

624

625 Prior to gene annotation palindromic SNPs were excluded ($n_{\text{SNPs}}=6,882,889$ remaining).
626 Variants were classified as either *cis*- or *trans*-pQTLs based on SNP proximity to the protein-
627 encoding gene (CD300LG) of interest. Variant annotation was performed using ANNOVAR⁷⁰,
628 labelling genes +/-500kb from variants. A pQTL was considered a *cis*-pQTL if the gene
629 annotation in the 1Mb window matched the protein name, all remaining variants were
630 considered *trans*-pQTLs.

631

632 To extract independent genome-wide significant pQTLs ($p < 5 \times 10^{-8}$) clumping was
633 performed using the PLINK v1.90b3.31 software package⁷¹; variants with $r^2 > 0.001$ with the
634 index SNP were removed using a 1 Mb window. Variants which lied within the human major
635 histocompatibility complex (MHC) region were removed, excluding pQTLs on chromosome 6
636 from 26Mb to 34Mb.

637

638 *Mendelian Randomization (MR)*

639 To obtain valid instrumental variables (SNPs) for our analysis we assessed them against the
640 three core assumptions for MR analysis: 1) That the SNPs were robustly associated with the
641 exposure of interest. For that we obtained summary result statistics on genome-wide
642 significant SNPs from our own GWAS. 2) That the SNPs were not associated with any known
643 or unknown confounders. This is not an assumption that can be fully tested, however we
644 used PhenoScanner^{72,73} to assess whether any SNPs were associated with known
645 confounders (described below). 3) That the SNPs were not associated with the outcomes
646 through any other path than through the exposure. To test this assumption, we searched

647 PhenoScanner^{72,73} (detailed below) to see if our exposures of interest were associated with
648 other potentially pleiotropic phenotypes.

649

650 MR statistical analysis

651 We used the TwoSampleMR package⁷⁴ (<https://github.com/MRCIEU/TwoSampleMR>) in R
652 version 4.2.2 (<https://cran.r-project.org/>). The outcome studies were obtained from
653 <https://www.magicinvestigators.org/>⁴⁶ and were external to the UK biobank. Specifically,
654 we used the outcomes “fasting glucose adjusted for BMI” (mmol/L, n=200622), “2-hour post
655 OGTT glucose adjusted for BMI” (mmol/L, n=63396), “fasting insulin adjusted for BMI”
656 (pmol/L, n=151013) and “HbA1c” (% , n=146806)⁴⁶.

657

658 We performed a two-sample inverse variance weighted (IVW) analysis to assess the causal
659 effect of CD300LG on metabolic factors (Supplementary Table 8 and 9). To explore potential
660 violations of the core assumptions when using the full set of SNPs, we performed a
661 heterogeneity test using Cochran’s Q, and a test for directional pleiotropy was conducted by
662 assessing the degree to which the MR Egger intercept differed from zero⁷⁵. We also
663 performed additional sensitivity analyses using MR Egger regression⁷⁵, weighted median⁷⁶,
664 simple and weighted mode estimation methods⁷⁷. Effect estimates from the different
665 sensitivity analysis were compared as a way of assessing the robustness of the results. To
666 assess potential heterogeneity in the MR estimates we further performed MR-PRESSO^{46,78}
667 to detect (MR-PRESSO global test) and correct for horizontal pleiotropy via outlier removal
668 (MR-PRESSO outlier test).

669

670 Investigation of potentially pleiotropic SNPs

671 SNPs robustly associated with the exposure investigated in the MR analyses (serum CD300LG
672 levels) were checked for other possible associations (PhenoScanner v2^{72,73},
673 <http://www.phenoscanter.medschl.cam.ac.uk/>) which may contribute to a pleiotropic effect
674 on the metabolic outcomes. Supplementary Table 10 lists the SNPs used in our analysis that
675 many influence related phenotypes. Phenotypes from PhenoScanner were listed if they
676 were associated with the SNPs or nearby variants in high LD ($r^2=0.8$) at p-value level $<1 \times 10^{-5}$
677 and could have potential pleiotropic effects in the analysis.

678

679 *Data availability*

680 mRNA sequencing data from MyoGlu can be found at
681 <https://exchmdpmsg.medsch.ucla.edu/app/> as well as in GEO:GSE227419. Secretory proteins
682 are available in the MetazSecKB data base at
683 <http://proteomics.yzu.edu/secretomes/animal/>. The human serum proteomic NAFLD
684 signature is available in the study of Govaere *et al*³⁸. Expression profiles in human liver cells
685 are available in the Human Liver Cell Atlas³⁹. Data obtained from the UK biobank (Olink
686 explore 1536 and measures of physical activity⁷⁹ can be found at
687 <https://biobank.ndph.ox.ac.uk/ukb/>. Glucometabolic outcomes used in MR analyses are
688 available at: <https://www.magicinvestigators.org/>⁴⁶. Data from the GD-CAT database⁴⁵ is
689 available from: <https://pipeline.biochem.uci.edu/gtex/demo2/>. Mice exercise data are
690 available at <https://motrpcac-data.org/> and knock-out data at
691 <https://www.mousephenotype.org/>. CD300LG expression values from a human tissue panel
692 were obtained from Uhlén *et al.*⁴². The single nuclei mRNA sequencing data from human
693 adipose tissue was plotted in Seurat v. 4 by downloading processed data from the Single Cell
694 Portal⁴³. The data can also be explored at:
695 [https://singlecell.broadinstitute.org/single_cell/study/SCP1376/a-single-cell-atlas-of-human-](https://singlecell.broadinstitute.org/single_cell/study/SCP1376/a-single-cell-atlas-of-human-and-mouse-white-adipose-tissue)
696 [and-mouse-white-adipose-tissue](https://singlecell.broadinstitute.org/single_cell/study/SCP1376/a-single-cell-atlas-of-human-and-mouse-white-adipose-tissue)). UK Biobank (<https://www.ukbiobank.ac.uk/>) data are
697 available to researchers upon application to the individual cohorts via their websites. All
698 other data used are publicly available and referenced according in the main text. For
699 additional details and data inquiries, please contact Sindre Lee-Ødegård.

700

701 **Acknowledgements**

702 South-Eastern Norway Regional Health Authority, Simon Fougner's fund, Diabetesforbundet,
703 Johan Selmer Kvanes' legat til forskning og bekjempelse av sukkersyke. This research has
704 been conducted using the UK Biobank resource (Reference 53641). DME is funded by an
705 Australian National Health and Medical Research Council Investigator Grant (APP2017942).
706 GHM is the recipient of an Australian Research Council Discovery Early Career Award
707 (Project number: DE220101226) funded by the Australian Government and supported by
708 the Research Council of Norway (Project grant: 325640 & Mobility grant: 287198). JKV is
709 supported by The Medical Student Research Program at the University of Oslo. SL is

710 supported by the Novo Nordisk Fonden Excellence Emerging Grant in Endocrinology and
711 Metabolism 2023 (NNF23OC0082123).

712

713 **Author contributions**

714 Conceptualization: SLØ, CAD, KIB. Methodology: SLØ, GHM, ED, DME. Data Collection: SLØ,
715 FN. Data analysis: SLØ, TO, MH, GHM, ED. Visualization: SLØ, GHM and MH. Writing original
716 draft: SLØ. Writing, review and editing: SLØ, MH, TO, GHM, ED, DME, JKV, HLG, FN, CAD, KIB.
717 Supervision: CAD, KIB. Funding acquisition : SLØ, CAD, KIB. Project administration: KIB.

718

719 **Competing interests**

720 The authors declare no conflict of interest.

721

722 **References**

- 723 1. Piercy, K.L., Troiano, R.P., Ballard, R.M., Carlson, S.A., Fulton, J.E., Galuska, D.A.,
724 George, S.M., and Olson, R.D. (2018). The Physical Activity Guidelines for Americans.
725 *Jama* 320, 2020-2028. 10.1001/jama.2018.14854.
- 726 2. Hawley, J.A., and Lessard, S.J. (2008). Exercise training-induced improvements in
727 insulin action. *Acta Physiol (Oxf)* 192, 127-135. 10.1111/j.1748-1716.2007.01783.x.
- 728 3. Bacchi, E., Negri, C., Zanolin, M.E., Milanese, C., Faccioli, N., Trombetta, M., Zoppini,
729 G., Cevese, A., Bonadonna, R.C., Schena, F., et al. (2012). Metabolic effects of aerobic
730 training and resistance training in type 2 diabetic subjects: a randomized controlled
731 trial (the RAED2 study). *Diabetes care* 35, 676-682. 10.2337/dc11-1655.
- 732 4. Pedersen, B.K., Akerström, T.C., Nielsen, A.R., and Fischer, C.P. (2007). Role of
733 myokines in exercise and metabolism. *Journal of applied physiology* (Bethesda, Md. :
734 1985) 103, 1093-1098. 10.1152/jappphysiol.00080.2007.
- 735 5. Chow, L.S., Gerszten, R.E., Taylor, J.M., Pedersen, B.K., van Praag, H., Trappe, S.,
736 Febbraio, M.A., Galis, Z.S., Gao, Y., Haus, J.M., et al. (2022). Exerkines in health,
737 resilience and disease. *Nat Rev Endocrinol* 18, 273-289. 10.1038/s41574-022-00641-
738 2.
- 739 6. Görgens, S.W., Eckardt, K., Jensen, J., Drevon, C.A., and Eckel, J. (2015). Exercise and
740 Regulation of Adipokine and Myokine Production. *Progress in molecular biology and
741 translational science* 135, 313-336. 10.1016/bs.pmbts.2015.07.002.
- 742 7. Lee-Ødegård, S., Olsen, T., Norheim, F., Drevon, C.A., and Birkeland, K.I. (2022).
743 Potential Mechanisms for How Long-Term Physical Activity May Reduce Insulin
744 Resistance. *Metabolites* 12. 10.3390/metabo12030208.
- 745 8. Pourteymour, S., Eckardt, K., Holen, T., Langleite, T., Lee, S., Jensen, J., Birkeland, K.I.,
746 Drevon, C.A., and Hjorth, M. (2017). Global mRNA sequencing of human skeletal
747 muscle: Search for novel exercise-regulated myokines. *Molecular metabolism* 6, 352-
748 365. 10.1016/j.molmet.2017.01.007.

- 749 9. Pedersen, B.K., and Febbraio, M.A. (2012). Muscles, exercise and obesity: skeletal
750 muscle as a secretory organ. *Nat Rev Endocrinol* 8, 457-465.
751 10.1038/nrendo.2012.49.
- 752 10. Lee, S., Norheim, F., Langleite, T.M., Gulseth, H.L., Birkeland, K.I., and Drevon, C.A.
753 (2019). Effects of long-term exercise on plasma adipokine levels and inflammation-
754 related gene expression in subcutaneous adipose tissue in sedentary dysglycaemic,
755 overweight men and sedentary normoglycaemic men of healthy weight. *Diabetologia*
756 62, 1048-1064. 10.1007/s00125-019-4866-5.
- 757 11. Bouassida, A., Chamari, K., Zaouali, M., Feki, Y., Zbidi, A., and Tabka, Z. (2010). Review
758 on leptin and adiponectin responses and adaptations to acute and chronic exercise.
759 *Br J Sports Med* 44, 620-630. 10.1136/bjism.2008.046151.
- 760 12. Stanford, K.I., Lynes, M.D., Takahashi, H., Baer, L.A., Arts, P.J., May, F.J., Lehnig, A.C.,
761 Middelbeek, R.J.W., Richard, J.J., So, K., et al. (2018). 12,13-diHOME: An Exercise-
762 Induced Lipokine that Increases Skeletal Muscle Fatty Acid Uptake. *Cell metabolism*
763 27, 1111-1120.e1113. 10.1016/j.cmet.2018.03.020.
- 764 13. Lee, S., Norheim, F., Gulseth, H.L., Langleite, T.M., Kolnes, K.J., Tangen, D.S.,
765 Stadheim, H.K., Gilfillan, G.D., Holen, T., Birkeland, K.I., et al. (2017). Interaction
766 between plasma fetuin-A and free fatty acids predicts changes in insulin sensitivity in
767 response to long-term exercise. *Physiological reports* 5. 10.14814/phy2.13183.
- 768 14. Kistner, T.M., Pedersen, B.K., and Lieberman, D.E. (2022). Interleukin 6 as an energy
769 allocator in muscle tissue. *Nature Metabolism* 4, 170-179. 10.1038/s42255-022-
770 00538-4.
- 771 15. Haugen, F., Norheim, F., Lian, H., Wensaas, A.J., Dueland, S., Berg, O., Funderud, A.,
772 Skålhegg, B.S., Raastad, T., and Drevon, C.A. (2010). IL-7 is expressed and secreted by
773 human skeletal muscle cells. *Am J Physiol Cell Physiol* 298, C807-816.
774 10.1152/ajpcell.00094.2009.
- 775 16. Otaka, N., Shibata, R., Ohashi, K., Uemura, Y., Kambara, T., Enomoto, T., Ogawa, H.,
776 Ito, M., Kawanishi, H., Maruyama, S., et al. (2018). Myonectin Is an Exercise-Induced
777 Myokine That Protects the Heart From Ischemia-Reperfusion Injury. *Circ Res* 123,
778 1326-1338. 10.1161/circresaha.118.313777.
- 779 17. Hjorth, M., Pourteymour, S., Görgens, S.W., Langleite, T.M., Lee, S., Holen, T., Gulseth,
780 H.L., Birkeland, K.I., Jensen, J., Drevon, C.A., and Norheim, F. (2016). Myostatin in
781 relation to physical activity and dysglycaemia and its effect on energy metabolism in
782 human skeletal muscle cells. *Acta Physiol (Oxf)* 217, 45-60. 10.1111/apha.12631.
- 783 18. McPherron, A.C., Lawler, A.M., and Lee, S.J. (1997). Regulation of skeletal muscle
784 mass in mice by a new TGF-beta superfamily member. *Nature* 387, 83-90.
785 10.1038/387083a0.
- 786 19. Rao, R.R., Long, J.Z., White, J.P., Svensson, K.J., Lou, J., Lokurkar, I., Jedrychowski,
787 M.P., Ruas, J.L., Wrann, C.D., Lo, J.C., et al. (2014). Meteorin-like is a hormone that
788 regulates immune-adipose interactions to increase beige fat thermogenesis. *Cell* 157,
789 1279-1291. 10.1016/j.cell.2014.03.065.
- 790 20. Kanzleiter, T., Rath, M., Gorgens, S.W., Jensen, J., Tangen, D.S., Kolnes, A.J., Kolnes,
791 K.J., Lee, S., Eckel, J., Schurmann, A., and Eckardt, K. (2014). The myokine decorin is
792 regulated by contraction and involved in muscle hypertrophy. *Biochemical and*
793 *biophysical research communications* 450, 1089-1094. 10.1016/j.bbrc.2014.06.123.

- 794 21. Malin, S.K., del Rincon, J.P., Huang, H., and Kirwan, J.P. (2014). Exercise-induced
795 lowering of fetuin-A may increase hepatic insulin sensitivity. *Medicine and science in*
796 *sports and exercise* 46, 2085-2090. 10.1249/mss.0000000000000338.
- 797 22. Catoire, M., Alex, S., Paraskevopoulos, N., Mattijssen, F., Evers-van Gogh, I., Schaart,
798 G., Jeppesen, J., Kneppers, A., Mensink, M., Voshol, P.J., et al. (2014). Fatty acid-
799 inducible ANGPTL4 governs lipid metabolic response to exercise. *Proceedings of the*
800 *National Academy of Sciences of the United States of America* 111, E1043-1052.
801 10.1073/pnas.1400889111.
- 802 23. Norheim, F., Hjorth, M., Langleite, T.M., Lee, S., Holen, T., Bindesboll, C., Stadheim,
803 H.K., Gulseth, H.L., Birkeland, K.I., Kielland, A., et al. (2014). Regulation of
804 angiotensin-like protein 4 production during and after exercise. *Physiological reports*
805 2. 10.14814/phy2.12109.
- 806 24. Contrepois, K., Wu, S., Moneghetti, K.J., Hornburg, D., Ahadi, S., Tsai, M.S., Metwally,
807 A.A., Wei, E., Lee-McMullen, B., Quijada, J.V., et al. (2020). Molecular Choreography
808 of Acute Exercise. *Cell* 181, 1112-1130.e1116. 10.1016/j.cell.2020.04.043.
- 809 25. Diaz-Canestro, C., Chen, J., Liu, Y., Han, H., Wang, Y., Honoré, E., Lee, C.H., Lam, K.S.L.,
810 Tse, M.A., and Xu, A. (2023). A machine-learning algorithm integrating baseline
811 serum proteomic signatures predicts exercise responsiveness in overweight males
812 with prediabetes. *Cell Rep Med* 4, 100944. 10.1016/j.xcrm.2023.100944.
- 813 26. Robbins, J.M., Peterson, B., Schraner, D., Tahir, U.A., Rienmüller, T., Deng, S., Keyes,
814 M.J., Katz, D.H., Beltran, P.M.J., Barber, J.L., et al. (2021). Human plasma proteomic
815 profiles indicative of cardiorespiratory fitness. *Nat Metab* 3, 786-797.
816 10.1038/s42255-021-00400-z.
- 817 27. Robbins, J.M., Rao, P., Deng, S., Keyes, M.J., Tahir, U.A., Katz, D.H., Beltran, P.M.J.,
818 Marchildon, F., Barber, J.L., Peterson, B., et al. (2023). Plasma proteomic changes in
819 response to exercise training are associated with cardiorespiratory fitness
820 adaptations. *JCI Insight* 8. 10.1172/jci.insight.165867.
- 821 28. Mi, M.Y., Barber, J.L., Rao, P., Farrell, L.A., Sarzynski, M.A., Bouchard, C., Robbins, J.M.,
822 and Gerszten, R.E. (2023). Plasma Proteomic Kinetics in Response to Acute Exercise.
823 *Mol Cell Proteomics* 22, 100601. 10.1016/j.mcpro.2023.100601.
- 824 29. Langleite, T.M., Jensen, J., Norheim, F., Gulseth, H.L., Tangen, D.S., Kolnes, K.J., Heck,
825 A., Storas, T., Grothe, G., Dahl, M.A., et al. (2016). Insulin sensitivity, body
826 composition and adipose depots following 12 w combined endurance and strength
827 training in dysglycemic and normoglycemic sedentary men. *Archives of physiology*
828 *and biochemistry* 122, 167-179. 10.1080/13813455.2016.1202985.
- 829 30. Sudlow, C., Gallacher, J., Allen, N., Beral, V., Burton, P., Danesh, J., Downey, P., Elliott,
830 P., Green, J., Landray, M., et al. (2015). UK Biobank: An Open Access Resource for
831 Identifying the Causes of a Wide Range of Complex Diseases of Middle and Old Age.
832 *PLOS Medicine* 12, e1001779. 10.1371/journal.pmed.1001779.
- 833 31. Hamaguchi, H., Dohi, K., Sakai, T., Taoka, M., Isobe, T., Matsui, T.S., Deguchi, S.,
834 Furuichi, Y., Fujii, N.L., and Manabe, Y. (2023). PDGF-B secreted from skeletal muscle
835 enhances myoblast proliferation and myotube maturation via activation of the
836 PDGFR signaling cascade. *Biochemical and biophysical research communications* 639,
837 169-175. <https://doi.org/10.1016/j.bbrc.2022.11.085>.
- 838 32. Lories, R.J., Peeters, J., Szulc, K., Hespel, P., and Luyten, F.P. (2009). Deletion of
839 frizzled-related protein reduces voluntary running exercise performance in mice.
840 *Osteoarthritis Cartilage* 17, 390-396. 10.1016/j.joca.2008.07.018.

- 841 33. Casas-Fraile, L., Cornelis, F.M., Costamagna, D., Rico, A., Duelen, R., Sampaolesi,
842 M.M., López de Munain, A., Lories, R.J., and Sáenz, A. (2020). Frizzled related protein
843 deficiency impairs muscle strength, gait and calpain 3 levels. *Orphanet Journal of*
844 *Rare Diseases* 15, 119. 10.1186/s13023-020-01372-1.
- 845 34. López-Bellón, S., Rodríguez-López, S., González-Reyes, J.A., Burón, M.I., de Cabo, R.,
846 and Villalba, J.M. (2022). CY5R3 overexpression preserves skeletal muscle
847 mitochondria and autophagic signaling in aged transgenic mice. *Geroscience* 44,
848 2223-2241. 10.1007/s11357-022-00574-8.
- 849 35. Tassi, E., Garman, K.A., Schmidt, M.O., Ma, X., Kabbara, K.W., Uren, A., Tomita, Y.,
850 Goetz, R., Mohammadi, M., Wilcox, C.S., et al. (2018). Fibroblast Growth Factor
851 Binding Protein 3 (FGFBP3) impacts carbohydrate and lipid metabolism. *Sci Rep* 8,
852 15973. 10.1038/s41598-018-34238-5.
- 853 36. Qian, Q., Hu, F., Yu, W., Leng, D., Li, Y., Shi, H., Deng, D., Ding, K., Liang, C., and Liu, J.
854 (2023). SWAP70 Overexpression Protects Against Pathological Cardiac Hypertrophy in
855 a TAK1-Dependent Manner. *J Am Heart Assoc* 12, e028628.
856 10.1161/jaha.122.028628.
- 857 37. Chen, H.H., Luche, R., Wei, B., and Tonks, N.K. (2004). Characterization of two distinct
858 dual specificity phosphatases encoded in alternative open reading frames of a single
859 gene located on human chromosome 10q22.2. *The Journal of biological chemistry*
860 279, 41404-41413. 10.1074/jbc.M405286200.
- 861 38. Govaere, O., Hasoon, M., Alexander, L., Cockell, S., Tiniakos, D., Ekstedt, M.,
862 Schattenberg, J.M., Boursier, J., Bugianesi, E., Ratziu, V., et al. (2023). A proteo-
863 transcriptomic map of non-alcoholic fatty liver disease signatures. *Nature*
864 *Metabolism* 5, 572-578. 10.1038/s42255-023-00775-1.
- 865 39. Aizarani, N., Saviano, A., Sagar, Mailly, L., Durand, S., Herman, J.S., Pessaux, P.,
866 Baumert, T.F., and Grün, D. (2019). A human liver cell atlas reveals heterogeneity and
867 epithelial progenitors. *Nature* 572, 199-204. 10.1038/s41586-019-1373-2.
- 868 40. Montgomery, M.K., Bayliss, J., Devereux, C., Bezawork-Geleta, A., Roberts, D., Huang,
869 C., Schittenhelm, R.B., Ryan, A., Townley, S.L., Selth, L.A., et al. (2020). SMOC1 is a
870 glucose-responsive hepatokine and therapeutic target for glycemic control. *Sci Transl*
871 *Med* 12. 10.1126/scitranslmed.aaz8048.
- 872 41. Ghodsian, N., Gagnon, E., Bourgault, J., Gobeil, É., Manikpurage, H.D., Perrot, N.,
873 Girard, A., Mitchell, P.L., and Arsenault, B.J. (2021). Blood Levels of the SMOC1
874 Hepatokine Are Not Causally Linked with Type 2 Diabetes: A Bidirectional Mendelian
875 Randomization Study. *Nutrients* 13. 10.3390/nu13124208.
- 876 42. Uhlén, M., Fagerberg, L., Hallström, B.M., Lindskog, C., Oksvold, P., Mardinoglu, A.,
877 Sivertsson, Å., Kampf, C., Sjöstedt, E., Asplund, A., et al. (2015). Proteomics. Tissue-
878 based map of the human proteome. *Science (New York, N.Y.)* 347, 1260419.
879 10.1126/science.1260419.
- 880 43. Emont, M.P., Jacobs, C., Essene, A.L., Pant, D., Tenen, D., Colletuori, G., Di Vincenzo,
881 A., Jørgensen, A.M., Dashti, H., Stefek, A., et al. (2022). A single-cell atlas of human
882 and mouse white adipose tissue. *Nature* 603, 926-933. 10.1038/s41586-022-04518-
883 2.
- 884 44. Battle, A., Brown, C.D., Engelhardt, B.E., and Montgomery, S.B. (2017). Genetic
885 effects on gene expression across human tissues. *Nature* 550, 204-213.
886 10.1038/nature24277.

- 887 45. Zhou, M., Tamburini, I., Van, C., Molendijk, J., Nguyen, C.M., Chang, I.Y.-Y., Johnson,
888 C., Velez, L.M., Cheon, Y., Yeo, R., et al. (2024). Leveraging inter-individual
889 transcriptional correlation structure to infer discrete signaling mechanisms across
890 metabolic tissues. *eLife* *12*, RP88863. [10.7554/eLife.88863](https://doi.org/10.7554/eLife.88863).
- 891 46. Chen, J., Spracklen, C.N., Marenne, G., Varshney, A., Corbin, L.J., Luan, J., Willems,
892 S.M., Wu, Y., Zhang, X., Horikoshi, M., et al. (2021). The trans-ancestral genomic
893 architecture of glycemic traits. *Nat Genet* *53*, 840-860. [10.1038/s41588-021-00852-9](https://doi.org/10.1038/s41588-021-00852-9).
- 894 47. Sanford, J.A., Nogiec, C.D., Lindholm, M.E., Adkins, J.N., Amar, D., Dasari, S., Drugan,
895 J.K., Fernández, F.M., Radom-Aizik, S., Schenk, S., et al. (2020). Molecular Transducers
896 of Physical Activity Consortium (MoTrPAC): Mapping the Dynamic Responses to
897 Exercise. *Cell* *181*, 1464-1474. [10.1016/j.cell.2020.06.004](https://doi.org/10.1016/j.cell.2020.06.004).
- 898 48. Dickinson, M.E., Flenniken, A.M., Ji, X., Teboul, L., Wong, M.D., White, J.K., Meehan,
899 T.F., Weninger, W.J., Westerberg, H., Adissu, H., et al. (2016). High-throughput
900 discovery of novel developmental phenotypes. *Nature* *537*, 508-514.
901 [10.1038/nature19356](https://doi.org/10.1038/nature19356).
- 902 49. Takatsu, H., Hase, K., Ohmae, M., Ohshima, S., Hashimoto, K., Taniura, N., Yamamoto,
903 A., and Ohno, H. (2006). CD300 antigen like family member G: A novel Ig receptor like
904 protein exclusively expressed on capillary endothelium. *Biochemical and biophysical
905 research communications* *348*, 183-191. <https://doi.org/10.1016/j.bbrc.2006.07.047>.
- 906 50. Liang, X., Yee, S.W., Chien, H.C., Chen, E.C., Luo, Q., Zou, L., Piao, M., Mifune, A.,
907 Chen, L., Calvert, M.E., et al. (2018). Organic cation transporter 1 (OCT1) modulates
908 multiple cardiometabolic traits through effects on hepatic thiamine content. *PLoS
909 Biol* *16*, e2002907. [10.1371/journal.pbio.2002907](https://doi.org/10.1371/journal.pbio.2002907).
- 910 51. Chen, L., Shu, Y., Liang, X., Chen, E.C., Yee, S.W., Zur, A.A., Li, S., Xu, L., Keshari, K.R.,
911 Lin, M.J., et al. (2014). OCT1 is a high-capacity thiamine transporter that regulates
912 hepatic steatosis and is a target of metformin. *Proceedings of the National Academy
913 of Sciences of the United States of America* *111*, 9983-9988.
914 [10.1073/pnas.1314939111](https://doi.org/10.1073/pnas.1314939111).
- 915 52. Jacob, Y., Anderton, R.S., Cochrane Wilkie, J.L., Rogalski, B., Laws, S.M., Jones, A.,
916 Spiteri, T., Hince, D., and Hart, N.H. (2022). Genetic Variants within NOGGIN, COL1A1,
917 COL5A1, and IGF2 are Associated with Musculoskeletal Injuries in Elite Male
918 Australian Football League Players: A Preliminary Study. *Sports Medicine - Open* *8*,
919 126. [10.1186/s40798-022-00522-y](https://doi.org/10.1186/s40798-022-00522-y).
- 920 53. Barry, J.C., Simtchouk, S., Durrer, C., Jung, M.E., and Little, J.P. (2017). Short-Term
921 Exercise Training Alters Leukocyte Chemokine Receptors in Obese Adults. *Medicine
922 and science in sports and exercise* *49*, 1631-1640. [10.1249/mss.0000000000001261](https://doi.org/10.1249/mss.0000000000001261).
- 923 54. Meinken, J., Walker, G., Cooper, C.R., and Min, X.J. (2015). MetazSecKB: the human
924 and animal secretome and subcellular proteome knowledgebase. *Database (Oxford)*
925 *2015*. [10.1093/database/bav077](https://doi.org/10.1093/database/bav077).
- 926 55. Umemoto, E., Takeda, A., Jin, S., Luo, Z., Nakahogi, N., Hayasaka, H., Lee, C.M.,
927 Tanaka, T., and Miyasaka, M. (2013). Dynamic changes in endothelial cell adhesion
928 molecule nepmucin/CD300LG expression under physiological and pathological
929 conditions. *PloS one* *8*, e83681. [10.1371/journal.pone.0083681](https://doi.org/10.1371/journal.pone.0083681).
- 930 56. Støy, J., Grarup, N., Hørlyck, A., Ibsen, L., Rungby, J., Poulsen, P.L., Brandslund, I.,
931 Christensen, C., Hansen, T., Pedersen, O., et al. (2014). Blood pressure levels in male
932 carriers of Arg82Cys in CD300LG. *PloS one* *9*, e109646.
933 [10.1371/journal.pone.0109646](https://doi.org/10.1371/journal.pone.0109646).

- 934 57. Umemoto, E., Tanaka, T., Kanda, H., Jin, S., Tohya, K., Otani, K., Matsutani, T.,
935 Matsumoto, M., Ebisuno, Y., Jang, M.H., et al. (2006). Nepmucin, a novel HEV
936 sialomucin, mediates L-selectin-dependent lymphocyte rolling and promotes
937 lymphocyte adhesion under flow. *J Exp Med* 203, 1603-1614.
938 10.1084/jem.20052543.
- 939 58. Surakka, I., Horikoshi, M., Mägi, R., Sarin, A.-P., Mahajan, A., Lagou, V., Marullo, L.,
940 Ferreira, T., Miraglio, B., Timonen, S., et al. (2015). The impact of low-frequency and
941 rare variants on lipid levels. *Nature Genetics* 47, 589-597. 10.1038/ng.3300.
- 942 59. Støy, J., Kampmann, U., Mengel, A., Magnusson, N.E., Jessen, N., Grarup, N., Rungby,
943 J., Stødkilde-Jørgensen, H., Brandslund, I., Christensen, C., et al. (2015). Reduced
944 CD300LG mRNA tissue expression, increased intramyocellular lipid content and
945 impaired glucose metabolism in healthy male carriers of Arg82Cys in CD300LG: a
946 novel genometic cross-link between CD300LG and common metabolic
947 phenotypes. *BMJ Open Diabetes Res Care* 3, e000095. 10.1136/bmjdr-2015-
948 000095.
- 949 60. Takatsu, H., Hase, K., Ohmae, M., Ohshima, S., Hashimoto, K., Taniura, N., Yamamoto,
950 A., and Ohno, H. (2006). CD300 antigen like family member G: A novel Ig receptor like
951 protein exclusively expressed on capillary endothelium. *Biochemical and biophysical
952 research communications* 348, 183-191. 10.1016/j.bbrc.2006.07.047.
- 953 61. Van Pelt, D.W., Guth, L.M., and Horowitz, J.F. (2017). Aerobic exercise elevates
954 markers of angiogenesis and macrophage IL-6 gene expression in the subcutaneous
955 adipose tissue of overweight-to-obese adults. *Journal of applied physiology* 123,
956 1150-1159. 10.1152/jappphysiol.00614.2017.
- 957 62. Ross, M., Kargl, C.K., Ferguson, R., Gavin, T.P., and Hellsten, Y. (2023). Exercise-
958 induced skeletal muscle angiogenesis: impact of age, sex, angiocrines and cellular
959 mediators. *European Journal of Applied Physiology* 123, 1415-1432. 10.1007/s00421-
960 022-05128-6.
- 961 63. Lee, S., Gulseth, H.L., Langleite, T.M., Norheim, F., Olsen, T., Refsum, H., Jensen, J.,
962 Birkeland, K.I., and Drevon, C.A. (2021). Branched-chain amino acid metabolism,
963 insulin sensitivity and liver fat response to exercise training in sedentary
964 dysglycaemic and normoglycaemic men. *Diabetologia* 64, 410-423. 10.1007/s00125-
965 020-05296-0.
- 966 64. Sudlow, C., Gallacher, J., Allen, N., Beral, V., Burton, P., Danesh, J., Downey, P., Elliott,
967 P., Green, J., Landray, M., et al. (2015). UK biobank: an open access resource for
968 identifying the causes of a wide range of complex diseases of middle and old age.
969 *PLoS Med* 12, e1001779. 10.1371/journal.pmed.1001779.
- 970 65. Bycroft, C., Freeman, C., Petkova, D., Band, G., Elliott, L.T., Sharp, K., Motyer, A.,
971 Vukcevic, D., Delaneau, O., O'Connell, J., et al. (2018). The UK Biobank resource with
972 deep phenotyping and genomic data. *Nature* 562, 203-209. 10.1038/s41586-018-
973 0579-z.
- 974 66. Auton, A., Brooks, L.D., Durbin, R.M., Garrison, E.P., Kang, H.M., Korbel, J.O.,
975 Marchini, J.L., McCarthy, S., McVean, G.A., and Abecasis, G.R. (2015). A global
976 reference for human genetic variation. *Nature* 526, 68-74. 10.1038/nature15393.
- 977 67. Loh, P.-R., Tucker, G., Bulik-Sullivan, B.K., Vilhjálmsson, B.J., Finucane, H.K., Salem,
978 R.M., Chasman, D.I., Ridker, P.M., Neale, B.M., Berger, B., et al. (2015). Efficient
979 Bayesian mixed-model analysis increases association power in large cohorts. *Nature
980 Genetics* 47, 284-290. 10.1038/ng.3190.

- 981 68. Wu, Y., Burch, K.S., Ganna, A., Pajukanta, P., Pasaniuc, B., and Sankararaman, S.
982 (2022). Fast estimation of genetic correlation for biobank-scale data. *The American*
983 *Journal of Human Genetics* 109, 24-32. <https://doi.org/10.1016/j.ajhg.2021.11.015>.
- 984 69. Lee, J.J., McGue, M., Iacono, W.G., and Chow, C.C. (2018). The accuracy of LD Score
985 regression as an estimator of confounding and genetic correlations in genome-wide
986 association studies. *Genet Epidemiol* 42, 783-795. 10.1002/gepi.22161.
- 987 70. Wang, K., Li, M., and Hakonarson, H. (2010). ANNOVAR: functional annotation of
988 genetic variants from high-throughput sequencing data. *Nucleic Acids Res* 38, e164.
989 10.1093/nar/gkq603.
- 990 71. Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J.,
991 Sklar, P., de Bakker, P.I., Daly, M.J., and Sham, P.C. (2007). PLINK: a tool set for whole-
992 genome association and population-based linkage analyses. *American journal of*
993 *human genetics* 81, 559-575. 10.1086/519795.
- 994 72. Staley, J.R., Blackshaw, J., Kamat, M.A., Ellis, S., Surendran, P., Sun, B.B., Paul, D.S.,
995 Freitag, D., Burgess, S., Danesh, J., et al. (2016). PhenoScanner: a database of human
996 genotype-phenotype associations. *Bioinformatics (Oxford, England)* 32, 3207-3209.
997 10.1093/bioinformatics/btw373.
- 998 73. Kamat, M.A., Blackshaw, J.A., Young, R., Surendran, P., Burgess, S., Danesh, J.,
999 Butterworth, A.S., and Staley, J.R. (2019). PhenoScanner V2: an expanded tool for
1000 searching human genotype-phenotype associations. *Bioinformatics (Oxford, England)*
1001 35, 4851-4853. 10.1093/bioinformatics/btz469.
- 1002 74. Hemani, G., Zheng, J., Elsworth, B., Wade, K.H., Haberland, V., Baird, D., Laurin, C.,
1003 Burgess, S., Bowden, J., Langdon, R., et al. (2018). The MR-Base platform supports
1004 systematic causal inference across the human phenome. *eLife* 7, e34408.
1005 10.7554/eLife.34408.
- 1006 75. Bowden, J., Davey Smith, G., and Burgess, S. (2015). Mendelian randomization with
1007 invalid instruments: effect estimation and bias detection through Egger regression.
1008 *Int J Epidemiol* 44, 512-525. 10.1093/ije/dyv080.
- 1009 76. Bowden, J., Del Greco, M.F., Minelli, C., Davey Smith, G., Sheehan, N.A., and
1010 Thompson, J.R. (2016). Assessing the suitability of summary data for two-sample
1011 Mendelian randomization analyses using MR-Egger regression: the role of the I²
1012 statistic. *Int J Epidemiol* 45, 1961-1974. 10.1093/ije/dyw220.
- 1013 77. Hartwig, F.P., Davey Smith, G., and Bowden, J. (2017). Robust inference in summary
1014 data Mendelian randomization via the zero modal pleiotropy assumption. *Int J*
1015 *Epidemiol* 46, 1985-1998. 10.1093/ije/dyx102.
- 1016 78. Verbanck, M., Chen, C.-Y., Neale, B., and Do, R. (2018). Detection of widespread
1017 horizontal pleiotropy in causal relationships inferred from Mendelian randomization
1018 between complex traits and diseases. *Nature Genetics* 50, 693-698. 10.1038/s41588-
1019 018-0099-7.
- 1020 79. Sophie, C., Josephine, Y.C., Michael, C., Adrian, B., and Michael, I.T. (2016). Cross-
1021 sectional study of diet, physical activity, television viewing and sleep duration in 233
1022 110 adults from the UK Biobank; the behavioural phenotype of cardiovascular
1023 disease and type 2 diabetes. *BMJ Open* 6, e010038. 10.1136/bmjopen-2015-010038.
1024

MyoGlu

