















Twin pair analysis uncovers links between DNA methylation, mitochondrial DNA quantity and obesity

Received: 4 April 2024

Accepted: 24 April 2025

Published online: 12 May 2025



Aino Heikkinen ^{1,2}✉, Vivienne F. C. Esser ³, Seung Hyuk T. Lee ⁴, Sara Lundgren¹, Antti Hakkarainen ⁵, Jesper Lundbom^{5,6}, Juho Kuula^{5,7}, Per-Henrik Groop ^{8,9,10}, Sini Heinonen^{11,12}, Sergio Villicaña ¹³, Jordana T. Bell ¹³, Alice Maguolo ¹⁴, Emma Nilsson¹⁴, Charlotte Ling ¹⁴, Allan Vaag^{15,16,17}, Päivi Pajukanta ^{4,18,19}, Jaakko Kaprio ¹, Kirsi H. Pietiläinen ^{11,20}, Shuai Li ^{3,21} & Miina Ollikainen ^{1,2}✉

Alterations in mitochondrial metabolism in obesity may indicate disrupted communication between mitochondria and nucleus, and DNA methylation may influence this interplay. Here, we leverage data from the Finnish Twin Cohort study subcohort ($n = 173$; 86 full twin pairs, 1 singleton), including comprehensive measurements of obesity-related outcomes, mitochondrial DNA quantity and nuclear DNA methylation levels in adipose and muscle tissue, to identify one CpG at *SH3BP4* significantly associated with mitochondrial DNA quantity in adipose tissue (FDR < 0.05). We also show that *SH3BP4* methylation correlates with its gene expression. Additionally, we find that 14 out of the 35 obesity-related traits display significant associations with both *SH3BP4* methylation and mitochondrial DNA quantity in adipose tissue. We use data from TwinsUK and the Scandinavian T2D-discordant monozygotic twin cohort, to validate the observed associations. Further analysis using ICE FALCON suggests that mitochondrial DNA quantity, insulin sensitivity and certain body fat measures are causal to *SH3BP4* methylation. Examining mitochondrial DNA quantity and obesity-related traits suggests causation from mitochondrial DNA quantity to obesity, but unmeasured within-individual confounding cannot be ruled out. Our findings underscore the impact of mitochondrial DNA quantity on DNA methylation and expression of the *SH3BP4* gene within adipose tissue, with potential implications for obesity.

The significance of mitochondria as a primary energy source for cell growth and survival is indisputable. While mitochondria contain their own circular 16.6 kb genome, most of the genes required for mitochondrial function are encoded in the nuclear DNA. Therefore, cells require synchronized communication between the nucleus and mitochondria, known as mitonuclear communication, to adapt to changing metabolic demands. This intricate interplay can be subject to

regulation by epigenetic mechanisms^{1,2}, including DNA methylation³, which may influence the activity of gene expression. Consequently, the importance of DNA methylation for mitochondrial function has been highlighted in the literature⁴.

The existing body of literature supports a bidirectional relationship between mitochondrial function and nuclear DNA methylation. On one hand, different mitochondrial haplotypes and variants can lead

A full list of affiliations appears at the end of the paper. ✉ e-mail: aino.heikkinen@helsinki.fi; miina.ollikainen@helsinki.fi

to differences in DNA methylation^{5–8} via numerous metabolic pathways, such as the methionine cycle and the production of methyl groups⁴. On the other hand, nuclear DNA methylation may impact mitochondrial metabolism by regulating mitochondrial-associated gene expression^{9–12}. The term ‘mitochondrial metabolism’ is used here as an umbrella term for any measurable trait associated with mitochondrial metabolism, such as mtDNA quantity (mtDNAq), copy number, biogenesis, dynamics, and OXPHOS (oxidative phosphorylation) activity.

The vital role of mitochondria in cellular energy metabolism has placed them in the center of interest in human traits and diseases with metabolic symptoms, including obesity.

Obesity is characterized by excess body fat, which can lead to a range of systemic metabolic disturbances, and it has been previously associated with compromised mitochondrial biogenesis and OXPHOS capacity^{13,14}. Furthermore, obesity is a highly heterogeneous trait, and the precise molecular phenotypes contributing to the altered mitochondrial metabolism remain unclear. In addition, the exact tissue-specific roles of mitochondria in obesity are not well documented, although it has been shown that adipose tissue metabolism may be more affected by the acquired weight than muscle tissue¹⁴.

The research on the role of DNA methylation underlying obesity-associated decline in mitochondrial metabolism remains scarce. A recent study showed that mtDNAq influences cardiovascular disease and mortality through changes in DNA methylation¹⁵. DNA methylation profiles of adipocyte progenitor cells of individuals with obesity have also been linked to mitochondrial metabolism¹⁶. Despite these insights, the causal pathways and comprehensive characterization of the orchestrated effects of DNA methylation and mitochondrial decline in obesity and obesity-related phenotypes remain elusive.

Here, we aimed to identify DNA methylation sites associated with differences in mtDNAq in two primary tissues affected by obesity, namely subcutaneous adipose tissue (SAT) and skeletal muscle. In addition, we explored whether these methylation sites are linked to a comprehensive range of obesity-related outcomes and other related phenotypes, including several anthropometric and body composition measures, clinical biomarkers, physical activity, and biological aging¹⁷. Finally, we explored the potential causal relationships between mtDNAq, DNA methylation, and obesity-related measures with a method called Inference about Causation from Examination of FAmilial CONfounding (ICE FALCON)¹⁸.

Results

A total of 173 individuals participated in the study, comprising 81 complete MZ twin pairs, 5 dizygotic (DZ) twin pairs, and 1 MZ singleton without available co-twin data (Table 1 and Supplementary Table 1). By singletons, we refer to twin individuals who lack their co-twin’s data in the dataset or analysis. The DZ pairs and singletons were included in the first part of the study, i.e., the EWAS, whereas the ICE FALCON causation analysis was conducted first for complete pairs of MZ twins only and then using both MZ and DZ twin pairs. The age range of the study cohort spanned from 23 to 70 years old, with females accounting for 59% of the cohort (Table 1). The mean body mass index (BMI) and fat percentage were 29.2 kg/m² and 36.6%, respectively, indicating that the sample is predominantly with overweight. In addition, average fasting glucose levels (5.6 mmol/l) and HOMA-IR measures (2.1 units) were slightly elevated, which suggests a reduced insulin sensitivity. Twenty-two participants were diagnosed with type 2 diabetes (T2D). The main analysis strategy is described in Fig. 1, while Supplementary Table 2 summarizes the epigenetic age acceleration estimates of the study participants. Furthermore, the MZ twin pairs with adipose tissue data (*n* = 71 pairs) used in the causal inference analysis are presented in Supplementary Table 3.

Table 1 | Participant characteristics

N (twin pairs, singletons)	173 (86, 1)
MZ / DZ (full twin pairs)	163 (81) / 10 (5)
Age, mean (range)	45.7 (22.8–69.3)
Female, %	59.0
Smoking, %	Never 39.3 Current 30.1 Former 30.6
Obesity-related measures, Mean (range) SD	
Weight (kg)	83.7 (48.7–143.6) 18.3
BMI (kg/m ²)	29.2 (19.7–45.9) 5.8
Waist circumference (cm)	97.6 (65.2–144.3) 15.7
WHR	0.9 (0.8–1.1) 0.1
Fat (%)	36.6 (8.9–57.3) 9.6
Fat (kg)	31.9 (7.4–66.0) 13.0
Fat-free mass (kg)	49.8 (29.5–76.7) 10.8
Intra-abdominal fat (cm ³)	1015 (152–3950) 870
Subcutaneous fat (cm ³)	4582 (827–15,129) 2830
Adipocyte volume (mm ³)	539 (123–1131) 218
Liver fat (%)	2.7 (0.1–22.4) 4.0
Total cholesterol (mmol/l)	4.8 (2.7–7.7) 0.9
HDL (mmol/l)	1.5 (0.5–3) 0.5
LDL (mmol/l)	3 (1–5.1) 0.8
Triglycerides (mmol/l)	1 (0.3–5.9) 0.6
hsCRP (mg/l)	2.3 (0.1–9.7) 2.3
Adipsin (µg/ml)	1.2 (0.7–1.7) 0.2
Adiponectin (ng/ml)	3320 (1310–7710) 1408
ALAT (U/l)	29.3 (14–127) 20.1
ASAT (U/l)	29.2 (7–131) 14.3
Systolic blood pressure (mmHg)	133 (97–195) 19
Diastolic blood pressure (mmHg)	80 (48–110) 78
Fasting glucose (mmol/l)	5.6 (4.0–16.1) 1.1
Fasting insulin (mU/l)	8.1 (0.9–34.3) 5.9
HOMA-IR	2.1 (0.2–9.9) 1.8
Matsuda index	6.7 (0.7–34.2) 4.9
Physical activity measures (Baecke scale), Mean (SD)	
Leisure-time physical activity	2.9 (0.6)
Sports activity	2.7 (0.9)
Work activity	2.5 (1.0)
Total activity	8.1 (1.5)

ALAT Alanine aminotransferase, ASAT Aspartate aminotransferase, BMI body mass index, DZ dizygotic, HOMA-IR Homeostatic model assessment for insulin resistance, hsCRP high-sensitivity C-reactive protein, MZ monozygotic, WHR waist-to-hip ratio.

Differential methylation analysis

To investigate whether nuclear DNA methylation was associated with mtDNAq, we performed an epigenome-wide association analysis (EWAS). In adipose tissue (*n* = 153 individuals; 71 MZ pairs, 4 DZ pairs, 3 singletons), we identified one CpG site (*cg19998400*) that was inversely associated with mtDNAq (FDR = 0.002) (Fig. 2a). The CpG is in the enhancer region of *SH3BP4* (SH3 domain binding protein 4) gene that codes for a protein involved in intracellular signaling pathways. Another CpG (*cg17468563*) located in the *DHRS3* (Dehydrogenase/Reductase 3) enhancer region showed marginal association with mtDNAq (FDR = 0.078) (Fig. 2a). Muscle tissue CpG methylation (*n* = 155 individuals; 74 MZ pairs, 3 DZ pairs, 1 singleton) was not associated with mtDNAq (Fig. 2b). The association profiles of the two tissues did not show similar patterns, as none of the top 100 CpG were found to be in common between adipose and muscle samples. In addition, there was no strong correlation between the

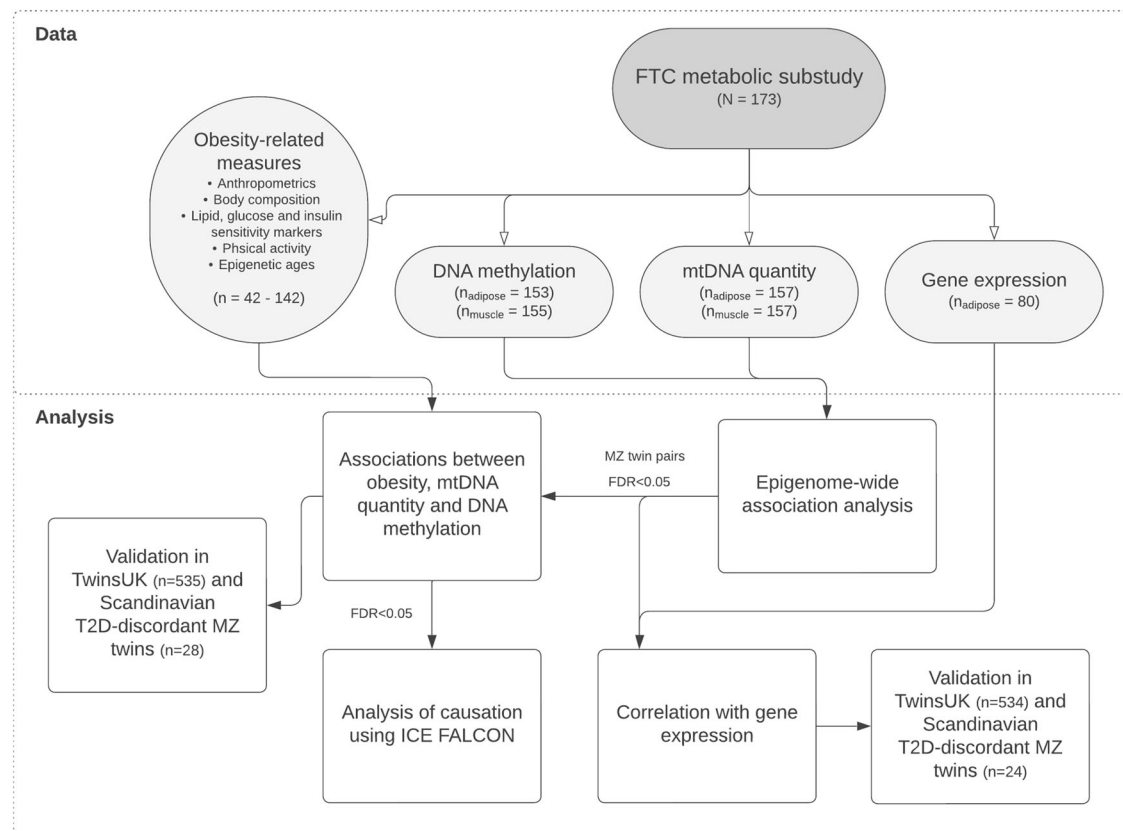


Fig. 1 | Study flowchart. FTC Finnish Twin Cohort, mtDNA mitochondrial DNA, MZ monozygotic, T2D type 2 diabetes.

regression effect sizes (Supplementary Fig. 1). The rest of the manuscript focuses explicitly on the findings in adipose tissue, given that we did not detect any statistically significant mtDNAq-linked DNA methylation sites in muscle.

Causal inference between adipose tissue mitochondrial DNA quantity and identified CpG sites

To investigate the potential evidence for causation underlying the associations between mtDNAq and the two identified CpG sites in adipose tissue, we performed ICE FALCON analysis for the complete MZ twin pairs with data on adipose tissue DNA methylation ($n = 68$ MZ pairs). ICE FALCON is a statistical method that uses familial structure (e.g., twin pairs) to make causal inferences based on changes in regression coefficients¹⁸. We analyzed each CpG separately, given that we cannot assume them to have identical causal pathways. For instance, the *cg17468563* has been reported to be under the genetic control of multiple loci¹⁹, whereas no meQTLs have been identified for *cg19998400*. Our data was consistent with the hypothesis that mtDNAq is causally linked to *cg19998400* (Table 2). This was suggested when we set mtDNAq as the independent variable and observed a cross-twin cross-trait association (Model 2: $B_{\text{cotwin}} = -0.167$, $p = 0.039$) which was attenuated towards null when adjusting for the twin's own mtDNAq. There was no evidence of change to within-individual association between Models 1 and 3. Conversely, when using mtDNAq as the dependent variable, we did not observe any cross-twin cross-trait associations in Models 2 or 3. However, the increased absolute value of the regression coefficient in Model 3 is in line with mtDNAq being causal to methylation at *cg19998400* (Table 2). The change was not statistically significant, which may reflect a low sample size and consequently reduced statistical power. The results from ICE FALCON on *cg17468563* were more ambiguous, with the regression coefficients pointing to either

causation from mtDNAq to methylation at *cg17468563* or the presence of within-individual confounding (Table 2). The results from analysis including also the complete DZ twin pairs ($n = 68$ MZ pairs, 4 DZ pairs) were consistent with the results obtained from MZ twin pairs alone (Supplementary Table 4).

Causal inference between adipose tissue DNA methylation and gene expression

As DNA methylation is a potential mechanism to regulate gene expression, we investigated whether the DNA methylation at *cg19998400* and *cg17468563* were associated with the expression of their respective genes, *SH3BP4* and *DHRS3* in adipose tissue ($n = 80$ individuals; 38 MZ pairs, 4 singletons). There was a positive correlation between the expression and methylation of *SH3BP4* ($r_{\text{Pearson}} = 0.46$, $p < 0.001$) (Fig. 3a), where the methylation explained solely 18% of the variation (marginal R-squared) and with familial confounding 58% (conditional R-squared) of the variation in gene expression (Fig. 3c). For the *DHRS3* expression and methylation, we observed a negative correlation ($r_{\text{Pearson}} = -0.66$, $p < 0.001$) (Fig. 3b), with a marginal R-squared of 42% and conditional R-squared 56% (Fig. 3c).

It is important to acknowledge that changes in gene expression patterns can also directly influence DNA methylation levels. Considering this, we conducted an ICE FALCON analysis for the 38 complete MZ pairs to explore the direction of causality between gene expression and DNA methylation. We observed a marginal cross-twin cross-trait association between *SH3BP4* expression and *cg19998400* methylation (Model 2: $B_{\text{cotwin}} = -0.112$, $p = 0.078$) (Supplementary Table 5) that attenuated towards null when adjusting for twin's own *cg19998400* levels, whereas the twin's own regression coefficient remained approximately the same. This observation is consistent with *cg19998400* methylation influencing

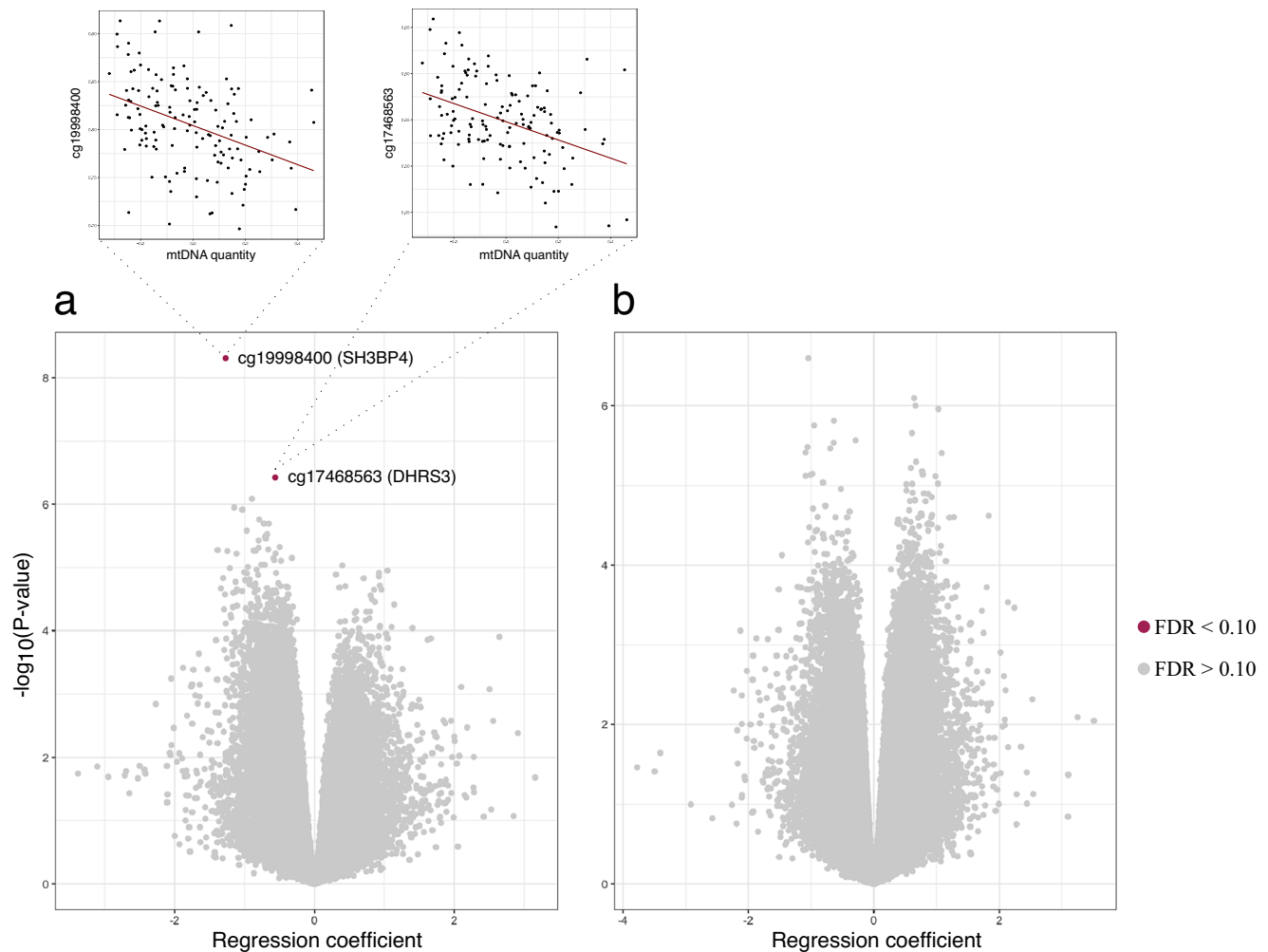


Fig. 2 | Volcano plots of the epigenome-wide association studies on mtDNA quantity. Results from (a) adipose tissue ($n = 153$ individuals) and (b) skeletal muscle ($n = 155$ individuals). Red dots indicate CpGs with $FDR < 0.10$. The correlations between the highlighted CpGs and mtDNA quantity is plotted above with the

red line representing the fitted regression line. The data was analyzed using moderated t-statistics and adjusted for multiple comparisons with false discovery rate (FDR). Source data are provided through Figshare [<https://doi.org/10.6084/m9.figshare.26941927.v3>]. mtDNA mitochondrial DNA.

the *SH3BP4* expression. On the other hand, the possibility for the presence of within-individual confounding cannot be ruled out. In contrast, with *DHRS3* and *cg17468563*, we did not find any compelling evidence for either causation or familial confounding (Supplementary Table 5).

Validation of the associations between methylation and expression of *SH3BP4*, and mitochondrial DNA quantity

To validate the association between *cg19998400* methylation and mtDNAq, we analyzed adipose tissue DNA methylation and mtDNAq data from T2D-discordant MZ twin pairs from Scandinavian Twin Registries^{20,21} (Supplementary Table 8.). We found a negative association between mtDNAq and *SH3BP4* methylation at *cg19998400* ($p < 0.001$; $n = 14$ MZ pairs), replicating our findings (Supplementary Table 9.). To validate the positive correlation between *cg19998400* methylation and *SH3BP4* expression, we used data from both the Scandinavian T2D-discordant MZ twin cohort ($n = 12$ MZ pairs) and TwinsUK²² ($n = 79$ MZ pairs, 111 DZ pairs, and 154 singletons) (Supplementary Table 8.). The TwinsUK data showed a significant positive correlation between *cg19998400* methylation and *SH3BP4* expression ($r_{\text{Pearson}} = 0.20$, $p < 0.001$). While the Scandinavian T2D-discordant MZ twins also showed a positive effect size, it did not reach statistical significance ($r_{\text{Pearson}} = 0.09$, $p = 0.66$).

Obesity-related variables associated with adipose tissue *SH3BP4* methylation and mitochondrial DNA quantity

Considering the well-established connection between mitochondrial dysfunction and obesity, we were interested to study whether the mtDNAq-associated CpG sites would be also associated with different obesity-related variables. We specifically focused on the DNA methylation of *cg19998400* at the *SH3BP4* locus (herein referred to as *SH3BP4* methylation), as our data suggested a potential causal relationship with mtDNAq, in contrast to *cg17468563* methylation at *DHRS3*. Among a set of 35 obesity-related traits, six showed significant association only with mtDNAq, four with *SH3BP4* methylation, and 14 with both mtDNAq and *SH3BP4* methylation ($FDR < 0.05$) (Fig. 4 and Supplementary Tables 6, 7).

Three distinct measures of epigenetic age acceleration (EAA), namely Horvath, Hannum, and PhenoAge, displayed a negative association with mtDNAq but not with *SH3BP4* methylation. In addition to EAA, mtDNAq showed negative associations with fat-free mass and adipon levels, and positive with sports activity. Obesity-related variables that associated exclusively with *SH3BP4* methylation included elevated blood triglyceride levels, hsCRP, fasting glucose levels, and systolic blood pressure. The 14 shared associations between mtDNAq and *SH3BP4* methylation included parameters mainly related to body fat composition, insulin sensitivity, and HDL cholesterol levels. Consistent with existing

Table 2 | Results from ICE FALCON analysis between mtDNA quantity and the identified CpG sites ($n = 68$ monozygotic twin pairs) in adipose tissue

Formula	Coef*	Model 1			Model 2			Model 3			Change in coefficients		
		Est	SE	P	Est	SE	P	Est	SE	P	Est	P	
mtDNAq (Y) cg19998400 (X)	βself	- 0.350	0.061	9.2E-09				- 0.358	0.060	2.1E-09	- 0.008	0.728	
	βcortwin				0.006	0.070	0.930	- 0.067	0.059	0.257	- 0.073	0.426	
cg19998400 (Y) mtDNAq (X)	βself	- 0.487	0.098	7.3E-07				- 0.475	0.102	3.6E-06	0.013	0.643	
	βcortwin				- 0.167	0.081	0.039	- 0.055	0.072	0.444	0.112	0.351	
mtDNAq (Y) cg17468563 (X)	βself	- 0.431	0.092	2.9E-06				- 0.428	0.092	3.6E-06	0.003	0.942	
	βcortwin				0.163	0.080	0.043	0.143	0.073	0.051	- 0.020	0.788	
cg17468563 (Y) mtDNAq (X)	βself	- 0.342	0.082	2.9E-05				- 0.331	0.082	5.0E-05	0.011	0.664	
	βcortwin				0.132	0.078	0.090	0.050	0.067	0.460	- 0.082	0.289	

*Standardized regression coefficient; β_{self} represents the association between twin's own X and Y variables whereas β_{cotwin} is the cross-twin cross-trait association i.e., the association between twin's own X variable with their co-twin's Y variable. X and Y in brackets indicate whether the variable was used as a predictor or an outcome, respectively. The p-values for models 1–3 were calculated from regression coefficients and standard errors using two-sided z-statistics. The change in coefficients was analyzed using non-parametric bootstrapping. No multiple comparison adjustment was applied. Regression models were adjusted for age, sex, smoking, BMI, and methylation beadchip and row. P-values < 0.05 are bolded.

literature, higher mtDNAq were correlated with lower body fat, and higher insulin sensitivity, HDL cholesterol levels, and adiponectin levels.

Validation of the associations between methylation and expression of *SH3BP4* and body mass index

To validate the observed associations between *SH3BP4* methylation and BMI, we again used data from T2D-discordant MZ twin pairs from Scandinavian Twin Registries and from TwinsUK. With TwinsUK, we replicated the positive association between *SH3BP4* methylation and BMI ($B = 0.061$, $p < 0.001$; $n = 79$ MZ pairs, 112 DZ pairs, and 153 singletons). In line with these findings, the expression of *SH3BP4* was also positively associated with BMI ($B = 0.072$, $p < 0.001$; $n = 131$ MZ pairs, 186 DZ pairs, and 131 singletons). Similar analyses from the Scandinavian T2D-discordant MZ twin cohort ($n = 14$ MZ pairs) produced effect sizes in the same direction but without statistically significant p-values (Supplementary Table 9.). While we initially omitted smoking as a covariate in the Scandinavian cohort analysis due to missing information for ten (35.7%) of the participants, additional analysis with the subsample that included smoking data showed an increased effect size when adjusting for smoking (Supplementary Table 9.). Consistent with our primary and validation findings, Rönn et al. also reported a positive association between BMI and DNA methylation of *cg19998400* at *SH3BP4* in adipose tissue of unrelated female participants²³.

Causal inference between obesity-related traits, and adipose tissue *SH3BP4* methylation and mitochondrial DNA quantity

To discern the potential causal relationship between the 14 obesity-related variables and both *SH3BP4* methylation and mtDNAq, we employed an ICE FALCON analysis for the complete MZ twin pairs in the cohort (Supplementary Table 3). Our findings indicate that certain variables related to insulin resistance and ectopic fat may exert causal influence on *SH3BP4* methylation, as suggested by significant cross-twin cross-trait association in Model 2 (body fat percentage $B_{\text{cotwin}} = 0.170$, $p = 0.074$; intra-abdominal fat $B_{\text{cotwin}} = 0.235$, $p = 0.031$; HOMA-IR $B_{\text{cotwin}} = 0.278$, $p = 0.002$; Matsuda $B_{\text{cotwin}} = -0.253$, $p = 0.006$; fasting insulin $B_{\text{cotwin}} = 0.235$, $p = 0.004$) that attenuated towards null after conditioning on twin's own corresponding measures (Fig. 5b). The changes in regression coefficients were significant in fat percentage ($p = 0.047$) and marginally significant in Matsuda ($p = 0.062$). Reversing the regression and using *SH3BP4* methylation as a predictor variable also suggested causality from these obesity-related traits to methylation, based on the behavior of the regression coefficients (i.e., the cross-twin cross-trait association increased substantially while the B_{self} remained relatively stable) (Fig. 5c). Results of other obesity-related traits, specifically those measuring body size, liver fat percentage, subcutaneous fat, and HDL cholesterol, were less clear. While the behavior of co-twin's regression coefficients in ICE FALCON may indicate that these variables are a consequence of *SH3BP4* methylation (Fig. 5b, c), the opposing signs between Model 1 and Model 2 coefficients, when using *SH3BP4* methylation as predictor variable X, can also suggest the presence of within-individual confounding (Fig. 5c). Similarly, the ICE FALCON analysis on the association between mtDNAq and obesity-related variables suggested either mtDNAq being causal to most of the obesity-related variables or being subject to unmeasured within-individual confounding (Fig. 5d, e). The ICE FALCON analysis, including both the complete MZ and DZ pairs, replicated the results discussed in this paragraph (Supplementary Fig. 2)

Discussion

In this study, we identified a significant association between mtDNAq and DNA methylation at *SH3BP4*, which correlated with its gene expression levels. This methylation site was also linked to numerous obesity-related traits, specifically those that measure body fat composition and insulin sensitivity. We further validated the association

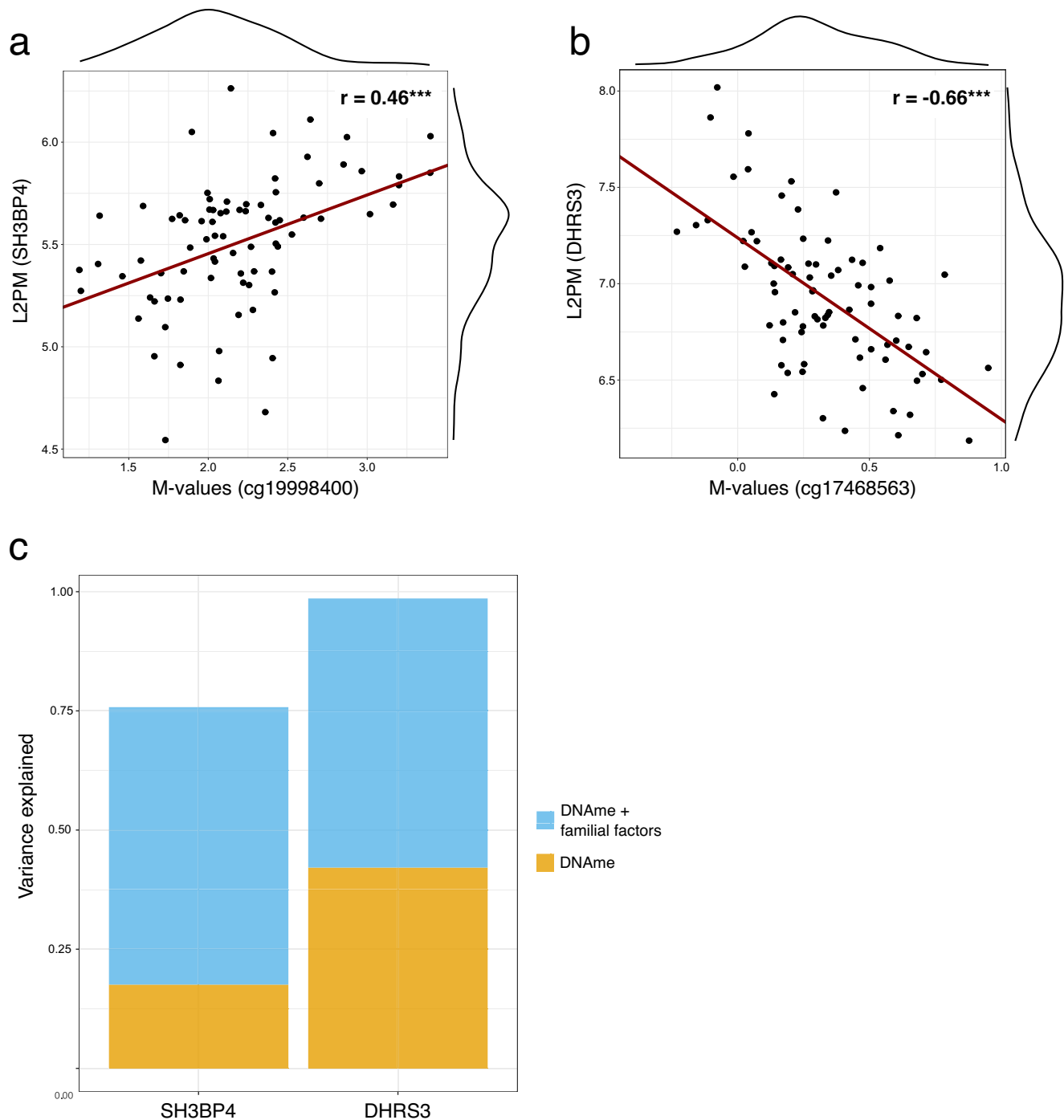


Fig. 3 | Relationship between DNA methylation and gene expression at two genomic loci in adipose tissue. a Correlation between methylation at *cg19998400* and *SH3BP4* expression ($n = 80$ individuals), showing the distribution of *cg19998400* on top and *SH3BP4* expression on right. **b** Correlation between methylation at *cg17468563* and *DHRS3* expression ($n = 80$ individuals), showing the distribution of *cg17468563* on top and *DHRS3* expression on right. Red lines in

indicate fitted linear regression lines. **c** Variation in *SH3BP4* and *DHRS3* expression explained by methylation at *cg19998400* and *cg17468563* (yellow bars), respectively, and combined with common familial factors (genetic or environmental) (blue bars) ($n = 80$ individuals). Source data are provided in the Source Data file. DNAme DNA methylation, L2PM = \log_2 -counts per million; r = Pearson correlation coefficient; *** $p < 0.001$.

between mtDNAq and *SH3BP4* methylation in the Scandinavian T2D-discordant MZ twin cohort, as well as the association between *SH3BP4* methylation, expression, and BMI in TwinsUK. Leveraging our data on MZ twin pairs, we identified potentially causal associations from mtDNAq and obesity-related outcomes to *SH3BP4* methylation in adipose tissue.

The dynamic interplay between mitochondria and the nucleus plays a pivotal role in responding to diverse extracellular signals and

metabolic conditions, as suggested by our findings: modifications in mtDNAq may precede changes in *SH3BP4* methylation, which would indicate a retrograde signaling from mitochondria to nuclear DNA methylation. This observation aligns with prior research, although limited, showing the impact of reduced mtDNAq on nuclear DNA methylation in human embryonic kidney cell lines⁵. One of the main hypotheses for retrograde signaling includes the importance of mitochondria in the methionine cycle and in the production of

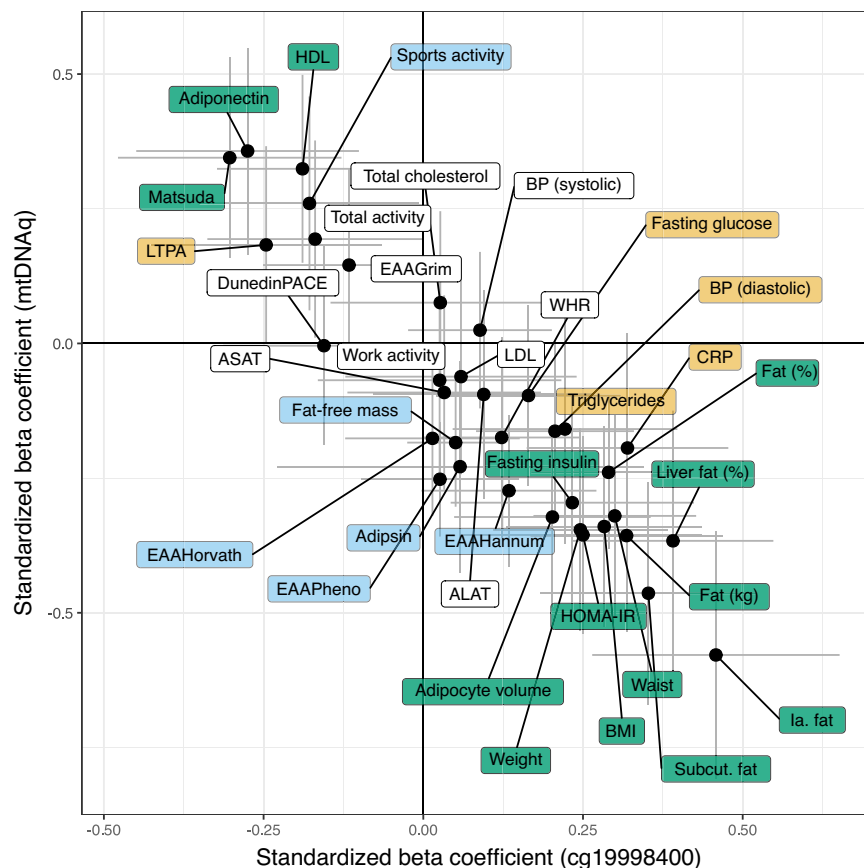


Fig. 4 | Standardized beta coefficients and their 95% confidence intervals for the associations between obesity-related outcomes and methylation at *cg19998400* (X-axis) or mtDNA quantity (Y-axis) ($n = 42\text{--}142$ individuals) in adipose tissue. Yellow background indicates variables with $FDR < 0.05$ associated with *cg19998400* methylation only, blue indicates variables associated with mtDNA quantity only, and green indicates variables associated with both. Black dots

present estimates for standardized beta coefficients, and gray lines represent 95 % confidence intervals. Source data are provided in the Source Data file. ALAT Alanine aminotransferase, ASAT Aspartate aminotransferase, CRP high-sensitivity C-reactive protein, EAA epigenetic age acceleration, BP blood pressure, HDL high-density, LTPA leisure-time physical activity, mtDNAq mitochondrial DNA quantity, Subcut. subcutaneous.

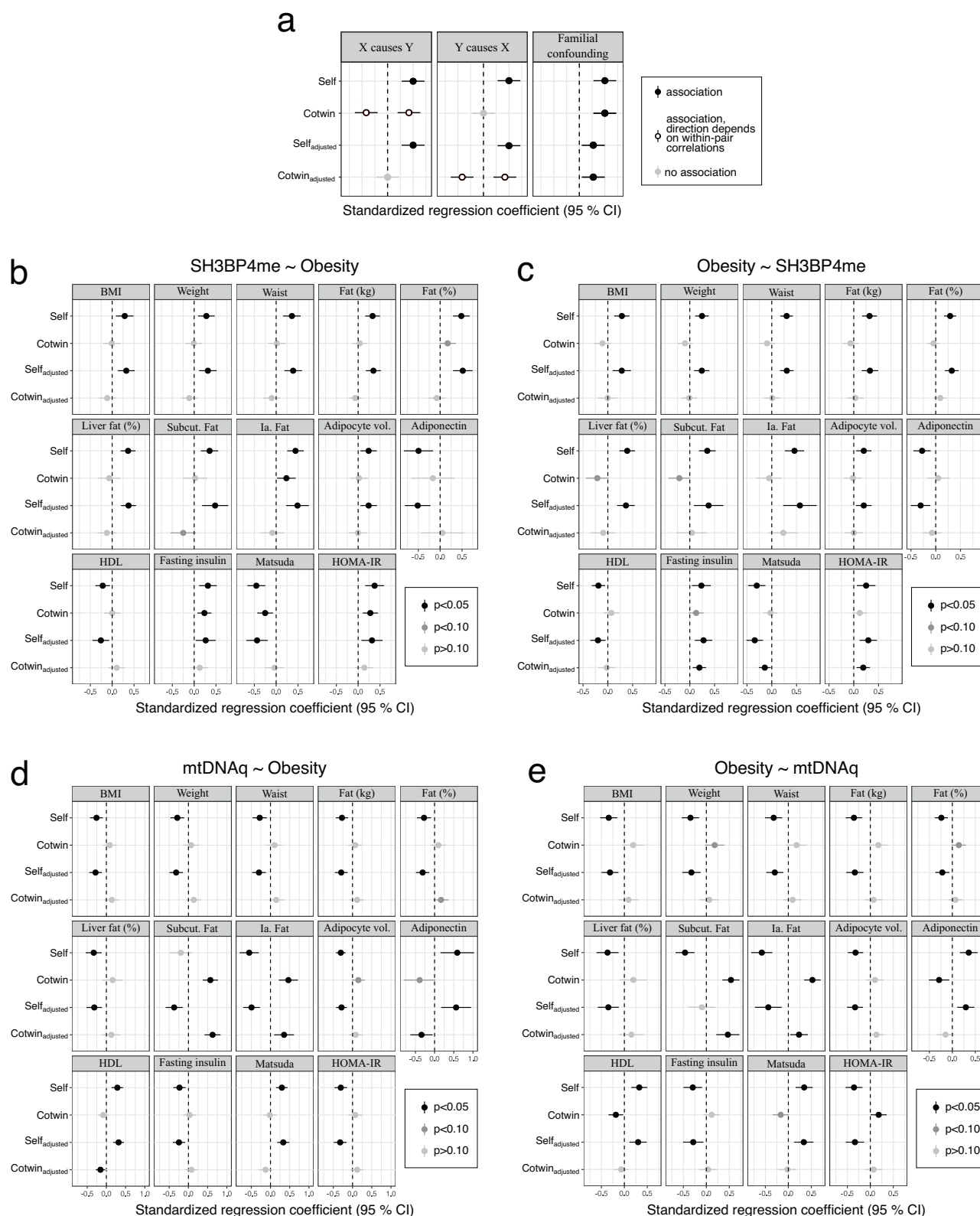
S-adenosylmethionine. S-adenosylmethionine is a primary methyl donor in cells, which interacts with DNA methyltransferases and therefore can affect DNA methylation²⁴. However, the targeted nature of the mitochondria-dependent methylation at specific genomic loci, such as *SH3BP4*, requires further investigation, along with its potential functional implications.

Our study indicated that the DNA methylation of *SH3BP4* and *DHRS3* associated with mtDNAq may be reflected at the gene expression level in adipose tissue. Specifically, methylation of *cg19998400* located at the 5'UTR region of *SH3BP4* correlated positively with *SH3BP4* expression, whereas methylation at the gene body (*cg17468563*) of *DHRS3* displayed a negative correlation, which is contrary to the previously observed general pattern of positive association between gene body methylation and expression^{25,26}. However, recent research indicates a far more complex relationship between these two factors, heavily influenced by the underlying genomic context^{27–29}. Although our study does not pinpoint the molecular mechanisms underlying the association, we underscore the potential implication of these genes and their methylation status in relation to varying mtDNAq in obesity. Moreover, we demonstrated that DNA methylation of both *SH3BP4* and *DHRS3* accounted for a substantial proportion of the variation in the expression of these genes, alongside shared familial factors within the twin pairs. These factors encompass both genetic elements, such as eQTLs, and environmental factors, like age and lifestyle, which cannot be distinguished apart using MZ twin pairs only. Consequently, future studies with the inclusion of a cohort

of DZ twins or other family members could provide insights into the relative significance of genes and environment.

The link we identified between the mtDNAq-associated methylation site in *SH3BP4* and various obesity-related outcomes, particularly those assessing insulin sensitivity and body fat composition, implies a potential role of mtDNAq-induced methylation in the etiology of obesity. *SH3BP4* acts as a negative regulator in many signaling pathways such as mTORC1 (mammalian target of rapamycin complex 1), a key promoter of cell growth and proliferation³⁰, and has been observed to promote adipogenesis, possibly by regulating mitochondrial functions^{31–33}. Previous research has demonstrated that mTOR signaling is compromised in obesity, potentially influenced by factors such as diet quality^{34–36} and oxidative stress³⁷. Still, the precise contribution of *SH3BP4* methylation and gene expression to cellular function and disease development remains unclear.

Our analysis indicates that changes in DNA methylation at *SH3BP4* may result from alterations in body insulin sensitivity and intra-abdominal fat accumulation. Previous studies conducted on blood have also suggested that methylation may be a consequence rather than a cause of obesity^{38,39}. Intriguingly, our findings suggest that only certain obesity-related outcomes, perhaps those more closely tied to metabolic disruptions such as insulin resistance and ectopic fat, are causally associated to *SH3BP4* methylation, while others closely related, such as BMI, are not. This may reflect the high metabolic heterogeneity often observed among people with similar BMI⁴⁰. Nevertheless, the discrepancy warrants further investigation into the



pathways that result in changes in *SH3BP4* methylation in the context of obesity.

The connection between mtDNAq and obesity has been established^{13,41,42}, yet the exploration of causal pathways has been limited. We identified a dualistic relationship between mtDNAq and obesity-related outcomes, which pointed to either causation from mtDNAq to several obesity-related outcomes or the presence of

unmeasured within-individual confounding. Excessive nutrient intake is a plausible factor influencing this association, as it is known to impair mitochondrial function^{43–45} and contribute to the excess body weight. Whether nutrient intake serves as a confounder in the association or initiates a pathway mediated by mitochondria leading to obesity remains uncertain. It is also possible that there is a circular relationship between mtDNAq and obesity. In addition, changes in mtDNAq may

Fig. 5 | Results from the ICE FALCON analysis between adipose tissue mtDNA quantity, *SH3BP4* methylation and obesity-related outcomes. **a** Conceptual figure illustrating the behavior of the ICE FALCON regression coefficients of Models 1–3 in each causal scenario. **b–e** Point estimates for the standardized regression coefficients and their 95% confidence intervals for the ICE FALCON analysis of (**b**) obesity-related outcomes regressed against *SH3BP4* methylation, (**c**) *SH3BP4* methylation regressed against obesity-related outcomes, (**d**) obesity-related outcomes regressed against mtDNA quantity and (**e**) mtDNA quantity regressed against obesity-related outcomes. ‘Self’ represents the association between twin’s own *X* and *Y* variables whereas ‘Cotwin’ is the cross-twin cross-trait association, i.e., the association between twin’s own *X* variable with their co-twin’s *Y* variable. ‘Adjusted’ refers to the regression coefficients derived from the ICE FALCON Model

manifest only after changes in other mitochondrial parameters, which were not covered in this study, thereby limiting the identification of causal associations between mitochondrial function and obesity and DNA methylation. Despite this, we demonstrate that both mtDNAq as well as specific obesity-related outcomes may be causal to DNA methylation at *SH3BP4*, via shared or independent molecular pathways.

Aging is widely linked with a decline in mitochondrial metabolism, including reduced mtDNAq^{46,47}. We revealed an association between mtDNAq and EAA, measured with Hannum, Horvath, and PhenoAge, in adipose tissue, which can indicate that mitochondrial metabolism is one of the key components in driving biological aging. These clocks have been reported to exhibit similar transcriptional signals with one another⁴⁸. While many of the epigenetic clocks were originally developed for whole blood (except for Horvath, that is a multi-tissue clock), it has been shown that some of the clocks, including PhenoAge and Hannum, work fairly robustly in other tissues too⁴⁸. Moreover, PhenoAge shows increased age acceleration in cells with depleted mitochondria⁴⁸.

Our study underscores the importance of investigating disease-affected tissues beyond readily available blood samples. Exploring two primary tissues affected by excess weight, we discovered that the associations between DNA methylation and mtDNAq in obesity are not uniform. Only adipose tissue DNA methylation, but not muscle, was found to be associated with mtDNAq, which may reflect the different roles of these two tissues in obesity. For instance, in obesity, alterations in adipose tissue seem to be more profoundly associated with metabolic health than those in muscle¹⁴. It may be that changes in muscle emerge only after systemic changes, such as insulin resistance, that are followed by lipid accumulation. Furthermore, mtDNAq as a proxy for mitochondrial biogenesis may not be directly comparable between the two tissues.

This study encompasses several strengths. First, the carefully phenotyped twin cohort for a comprehensive range of obesity-related outcomes enables disentangling the most significant molecular phenotypes of obesity to mtDNAq. Second, the inclusion of adipose and skeletal muscle tissues broadens the examination of the impact of excess weight and other obesity-related outcomes across tissues. In addition, using MZ twins, we can use statistical methods such as ICE FALCON to investigate the causality of the observed associations. The ICE FALCON offers a robust approach to explore the causal relationship in the observational data using related individuals, including twins^{18,49–51}, without the need for genetic instrumental variables. To our knowledge, there is a lack of established genetic variants or polygenic scores to estimate mtDNAq specifically in adipose and muscle tissues. This absence of genetic information on mtDNAq prevented us from using statistical methods such as MR-DoC⁵², which depend on genetic data.

However, there are limitations to consider that include the cross-sectional nature of the study, as well as the modest sample size, which substantially limits statistical power. In addition, we lacked precise data on some potential confounding variables, such as dietary habits,

3 that includes both twin’s own and their cotwin’s *X* variables. The *p*-values were calculated from regression coefficients and standard errors using two-sided *z*-statistics. No multiple comparison adjustment was applied. Source data with the exact *p*-values are provided as a Source Data file. **b, c** N (BMI, Fat %, Fat kg, HDL, Weight) = 68 pairs, N (Fasting insulin) = 64 pairs, N (HOMA-IR, Waist) = 61 pairs, N (Adipocyte vol.) = 57 pairs, N (Matsuda) = 56 pairs, N (Liver fat %) = 41 pairs, N (Adiponectin, la. Fat, Subcut. Fat) = 21 pairs d-e N (BMI, Fat %, Fat kg, HDL, Weight) = 71 pairs, N (Fasting insulin) = 67 pairs, N (HOMA-IR, Waist) = 64 pairs, N (Adipocyte vol.) = 60 pairs, N (Matsuda) = 59 pairs, N (Liver fat %) = 42 pairs, N (Adiponectin, la. Fat, Subcut. Fat) = 21 pairs HDL high-density lipoprotein, la. intra-abdominal, mtDNAq mitochondrial DNA quantity, Subcut. subcutaneous.

which could potentially affect the observed associations. It is important to note that our cohort comprised mostly participants without severe health conditions, so our findings may not generalize to more serious conditions like metabolic syndrome or T2D. For instance, while the replication analysis between *SH3BP4* methylation and BMI in the Scandinavian MZ T2D-discordant cohort validated the direction of effects, it was statistically nonsignificant. This may be due to low statistical power (only 7 twin pairs were available for the analysis, adjusting for smoking, a known important biological covariate), or cohort-specific differences such as age range and particularly, the higher prevalence of T2D compared to FTC or TwinsUK. Indeed, the link could be validated in a larger Scandinavian cohort, including adipose tissue from 94 unrelated female participants, which also found a positive association between BMI and *SH3BP4* methylation²³.

The characteristics of our dataset serves as both a strength and a limitation: while its distinctive integration of DNA methylation with mtDNAq in adipose and muscle tissues from twins with obesity-related measures represents a comprehensive approach, yet it restricts our ability to extensively validate all our findings in comparable independent datasets, which are very scarce.

Overall, we demonstrate a potential causal link from adipose tissue mitochondrial metabolism to DNA methylation and expression of *SH3BP4*. In addition, this connection holds significance in obesity, where certain outcomes related to insulin sensitivity and intra-abdominal fat were seen to contribute to *SH3BP4* methylation, influenced either by mtDNAq or through alternative pathways. We propose the existence of a complex network interconnecting DNA methylation and mitochondrial metabolism in obesity, contributing to the multifaceted nature of obesity as a phenotype. Comprehensive examination of their interplay with various metabolic parameters holds promise for advancing our understanding of the intricate metabolic landscape in obesity.

Methods

The Ethics Committee of the Hospital District of Helsinki and Uusimaa approved the protocols of this FTC substudy data collection (270/13/03/01/2008), and all participants provided their written informed consent. The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Declaration of Helsinki.

Study cohort

The study participants originate from the metabolic substudy from the larger population-based Finnish Twin Cohort (FTC; FinnTwin12⁵³, FinnTwin16⁵⁴, and Older Finnish Twin Cohort⁵⁵). The present substudy was originally designed to study obesity-related metabolism by targeting MZ twin pairs. Twins were invited to participate based on their self-reported weight and height to ensure the presence of twins with varying discordance for BMI. Zygosity was first defined by questionnaires⁵⁶, and later confirmed using genotyping, revealing some of the expected MZ pairs to be DZ instead. Participants who

consented to adipose and/or muscle biopsies were included in the study with no other specific inclusion or exclusion criteria beyond the above described. Our data consisted of 173 individuals (81 MZ twin pairs, 5 DZ twin pairs, and 1 MZ singleton without available co-twin data) from this substudy (Supplementary Table 1).

Clinical data

The selected study participants were deeply phenotyped for obesity-related clinical measures (Table 1) as described more in detail in van der Kolk et al.¹⁴. Briefly, anthropometric and body composition were measured after overnight fasting. Body mass index (BMI) was calculated from weight and height (kg/m^2), measured during on-site visits. In addition, waist circumference was measured and the waist-to-hip ratio (WHR) calculated. Fat mass, fat percentage, and lean mass were quantified using dual-energy X-ray absorptiometry. Intra-abdominal and subcutaneous fat volumes were measured using magnetic resonance imaging (MRI), and liver fat content using magnetic resonance spectroscopy (MRS). Supine blood pressure measurements were also taken (mean of three measurements).

Blood samples were obtained after overnight fasting, and concentrations of plasma glucose, serum insulin, plasma total cholesterol, low-density lipoprotein (LDL), high-density lipoprotein (HDL), triglycerides, high sensitivity C-reactive protein (hsCRP), alanine aminotransferase (ALAT) and aspartate aminotransferase (ASAT) were measured using standard HUSLAB clinical laboratory procedures. Homeostatic model assessment-insulin resistance index (HOMA-IR) and the Matsuda index for insulin sensitivity were calculated from the standard 4-point oral glucose tolerance test (OGTT).

Levels of different forms of physical activity (sport, work, leisure and total) were assessed using the Baecke questionnaire^{57,58}.

Sample collection

The samples of this study have been used in previous research^{14,59,60}. Briefly, the adipose and muscle tissue samples were collected from subcutaneous abdominal adipose tissue and vastus lateralis muscle, respectively, under local anesthesia (lidocaine). Adipose tissue samples were taken using a surgical technique or needle biopsy, and muscle samples through Bergström needle biopsy. Determination of adipocyte volume (dm^3) was determined with the custom algorithm for ImageJ¹⁴ and has been previously described in Lapatto et al.⁶⁰. The tissue samples for DNA/RNA extraction were snap-frozen in liquid nitrogen.

Mitochondrial DNA quantity

DNA was extracted from adipose and muscle tissue biopsies using AllPrep DNA/RNA/miRNA Universal Kit (Qiagen). The amount of mtDNA was quantified using quantitative PCR (qPCR), targeting for two mitochondrial encoded genes, *ND5* (NADH dehydrogenase 5) and *CYTB* (cytochrome b), and normalized to genomic DNA as measured from *APP* (amyloid-beta precursor protein) and *B2M* (beta-2-microglobulin). Details on primer sequences is provided in Supplementary Data 1. Data was processed using the $2^{-\Delta\Delta C_t}$ method with qbase+ software (Biogazelle) to obtain calibrated normalized relative quantities (CNRQ) for the mtDNA quantity. In cases of missing data (due to poor sample quality) for either *ND5* or *CYTB*, we applied a stochastic regression method to input missing values, using the other gene as a reference. Subsequently, we calculated the mean of *ND5* and *CYTB*, which served as the metric for mtDNAq.

DNA methylation data

High molecular weight DNA was extracted from adipose and muscle biopsies with AllPrep DNA/RNA/miRNA Universal Kit (Qiagen) or QIAmp DNA Mini Kit (Qiagen) and bisulfite converted with an EZ DNA Methylation Kit (ZYMO Research) according to the manufacturers' protocols. DNA methylation was quantified using Illumina Infinium HumanMethylation450K (450K; adipose tissue) or

Illumina HumanMethylationEPIC BeadChip arrays (EPIC; adipose and muscle tissues).

DNA methylation data was preprocessed and normalized with R package meffil⁶¹. Due to the modest sample size, adipose tissue 450K and EPIC data were preprocessed together, omitting the platform-specific probes from the analysis. After background and bias correction, we excluded bad quality samples with following criteria: (i) Median difference in X and Y chromosome intensities > 3 standard deviations (SDs), (ii) Median methylated vs. unmethylated intensity > 3 SDs, (iii) unreliable control probes, (iv) detection p -value > 0.01 in more than 20% of probes and (v) bead number < 3 in more than 20% of the probes.

We then applied quantile normalization to reduce technical variation by adjusting for methylation sample, slide, and control probe PCs. Number of PCs included was estimated from a scree plot separately for adipose and muscle data (Supplementary Fig. 3). Bad quality probes were removed with following criteria: (i) Detection p -value > 0.01 in more than 20% samples, (ii) bead number < 3 in more than 20% samples, (iii) SNP probes and (iv) ambiguous mapping probes^{62,63}. After QC, the number of samples and probes were 153/411,585 and 155/765,201 CpG sites for adipose and muscle data, respectively. The data was then beta mixture quantile (BMIQ) normalized to adjust for type2 probe bias. Because a considerable amount of batch effect remained from 450K and EPIC platforms in adipose tissue, we applied ComBat to minimize the effect of a platform (Supplementary Fig. 4). Methylation M-values, calculated as the \log_2 ratio between the methylated versus unmethylated probe intensities, were used in the statistical analysis⁶⁴.

Epigenetic age acceleration estimates. Epigenetic age for each individual was calculated from the preprocessed DNA methylation data. We used the principal component (PC) versions of the original Hannum⁶⁵, Horvath⁶⁶, GrimAge⁶⁷, and PhenoAge⁶⁸ clocks, as that has been shown to remove bias caused by technical variation in certain CpGs⁶⁹. Other epigenetic clocks applied were DunedinPACE⁷⁰, which measures the pace of aging, as well as muscle-specific epigenetic clock MEAT⁷¹ that was calculated for skeletal muscle tissue only. Epigenetic age acceleration measures, defined as the residuals from regressing an epigenetic age estimate on chronological age, were used in the statistical analyses.

RNA sequencing data

Adipose tissue RNA sequencing data was available for a subset of the twins ($n = 80$ individuals). The generation and preprocessing of the RNA-seq data has been described in detail elsewhere¹⁴. Briefly, RNA was extracted using AllPrep DNA/RNA/miRNA Universal Kit (Qiagen) with DNase I (Qiagen) digestion according to manufacturers' protocol. After the calculation of RNA integrity numbers, the libraries were prepared with Illumina Stranded mRNA preparation and sequenced with the Illumina HiSeq2000 platform. The data was mapped against the human reference genome 38, the quality was calculated, and read counts generated.

Statistical analysis

Epigenome-wide association analysis on mitochondrial DNA quantity. To identify individual CpG sites associated with mtDNAq in adipose and muscle tissue, we performed an EWAS using the R package limma⁷² that fits a linear model for each probe and computes moderated Bayes t-statistics. The models were adjusted for known biological and behavioral (age, self-reported sex, smoking status), and technical (beadchip date and row) covariates, as well as cell type proportions. In the absence of reference-based cell type deconvolution methods developed specifically for either adipose or muscle tissues, we applied EpiDISH⁷³ to adjust for key known cell types: fibroblasts and epithelial cells. In addition, the adipose tissue model was adjusted for the proportion of fat cells derived from EpiDISH. The fraction of immune cells

was omitted from the model due to high correlation with other cell types. The relatedness of twins in a pair was accounted for as a random effect in the model.

Associations between DNA methylation and gene expression in adipose tissue. DNA methylation is known to influence gene expression, due to which we investigated whether the identified DNA methylation sites correlated with the expression of their closest genes. We applied Pearson correlation to see to what extent DNA methylation and the expression of their respective genes were related. In addition, using linear mixed effects modeling between gene expression (outcome) and DNA methylation (predictor), and adjusting for the relatedness of the twins, we investigated how much of the variation in gene expression is explained (i) by the changes in DNA methylation solely (marginal R squared) and (ii) together with common familial factors shared within the twin pairs (conditional R squared).

Associations between obesity-related outcomes, mitochondrial DNA quantity and DNA methylation. Given that there is a well-established link between mitochondria and obesity, we were interested to explore whether the mtDNAq-associated CpG methylation was further linked to various obesity-related measures. We performed a generalized equation estimation (gee) regression using the R package *geepack*⁷⁴ with ‘exchangeable’ correlation structure to account for within-twin pair similarities. The CpG methylation was used as an outcome variable and each obesity-related outcome separately as a predictor variable. Those obesity-related variables that did not follow normal distribution were \log_{10} transformed before the analysis (Supplementary Tables 6, 7). The models were adjusted for age, sex, and smoking status, beadchip date, and row. We determined the associations between the obesity-related outcomes and mtDNAq and chose variables that exhibited statistically significant (FDR < 0.05) links with both mtDNAq and CpG methylation to the causal inference analysis.

Causal inference. To assess the potential evidence for a causal relationship or common familial confounding underlying the identified associations between DNA methylation, mtDNAq, and obesity-related outcomes, we applied a statistical method called ICE FALCON (Inference about Causation from Examination of FAMILial CONfounding)¹⁸. The method is based on regression models of observational data of related individuals, specifically twins, and assesses the changes in twin’s own and co-twin’s regression coefficient from with and without adjusting the counterparts’ predictor variables (Eqs. 1–3 below). As the power of the ICE FALCON depends on the within-pair correlation, we originally restricted the ICE FALCON analysis for the complete MZ pairs, which are expected to be more correlated than DZ pairs. However, we performed additional ICE FALCON analyses that also included the DZ pairs.

$$Y_{self} \sim \alpha + \beta_{self} \times X_{self} \quad (1)$$

$$Y_{self} \sim \alpha + \beta_{cotwin} \times X_{cotwin} \quad (2)$$

$$Y_{self} \sim \alpha + \beta_{self} \times X_{self} + \beta_{cotwin} \times X_{cotwin} \quad (3)$$

Briefly, if the observed cross-twin cross-trait association (β_{cotwin}) diminishes after adjusting for within-individual association, it suggests a causal effect from the predictor variable to the outcome (Fig. 5a). Conversely, if the β_{cotwin} appears only when conditioning on the within-individual association, it indicates a causal relationship from the outcome to the predictor. If both the β_{cotwin} and the β_{self} decrease after conditioning on each other, it indicates familial confounding (genetic or environmental) between the predictor and outcome. We examined β_{cotwin} with and without conditioning on within-individual association

using generalized equation estimation in the R package *geepack*⁶² using ‘exchangeable’ correlation structure for the twin pairs. All analyses were adjusted for age, sex, and smoking. Technical covariates of beadchip date and row were added as covariates to adjust analysis which included DNA methylation data. In addition, the ICE FALCON between mtDNAq and DNA methylation sites were further adjusted for BMI. We then calculated the changes in regression coefficients, for which we estimated standard errors of changes with non-parametric bootstrapping, generating 1000 datasets having the original sample size. The models were then reversed, e.g., using the previous X variable as the Y variable to gain additional evidence on the causal pathway.

Validation cohorts and statistical analysis. From the TwinsUK²² 535 study, participants with available adipose tissue DNA methylation and 765 with gene expression data were selected (Supplementary Table 9) of which 534 had both DNA methylation and gene expression data. Ethical approval for TwinsUK was obtained by the National Research Ethics Service, London-Westminster, the St Thomas’ Hospital Research Ethics Committee (EC04/015 and 07/H0802/84). DNA methylation profiles were generated from SAT biopsies using the Illumina Infinium HumanMethylation450K platform⁷⁵. The data was pre-processed for background and dye bias correction, quantile normalized, and probes with detection p -value < 0.000001 and bead number < 3 set as missing⁷⁶. The generation of SAT RNA-seq data is described in Christiansen et al.⁷⁶. Briefly, reads were aligned to the hg19 reference genome, and samples with < 10 million reads excluded. Gene counts were inverse-normalized. The associations between *cg19998400* methylation, *SH3BP4* expression, and BMI were performed in R using linear mixed effect modeling, adjusting for age, sex, and three-category smoking status (never, former, and current), technical covariates, and accounting for twin pair similarity as a random effect.

The Scandinavian discordant MZ twin cohort includes fourteen MZ pairs, discordant for T2D, recruited from the Scandinavian twin registries^{20,21} (Supplementary Table 8). Ethical approvals were granted by local ethics committees in Fyn (96/253) and Lund (520/2008). Genomic DNA was extracted with a DNeasy Blood and Tissue kit (Qiagen) and subjected to DNA methylation analysis using the Illumina Infinium HumanMethylation450K platform. The methylation data was background corrected, quantile normalized, and probes with detection p -value < 0.01 were removed²¹. Total RNA was extracted from frozen SAT biopsies using the miRNeasy kit (Qiagen, Hilden, Germany) and RNA expression analyzed with GeneChip Human Gene 1.0 ST arrays (Affymetrix, Santa Clara, CA), and Robust Multichip Average expression measures were computed^{20,21}. Data on SAT mtDNAq was determined with the ViiA 7 system (Applied Biosystems) as the mean value of the quantity of three mitochondrial genes, *ND6*, *RNR2*, and *CYTb*, normalized to the nuclear gene *RNaseP* quantity²¹. The associations between *cg19998400* methylation, *SH3BP4* expression, BMI, and mtDNAq were determined using gee, adjusting for age, sex, and two-category smoking status (yes/no) with ‘exchangeable’ correlation structure to account for within twin pair similarities. Pearson correlation coefficient was used to determine the correlation between *SH3BP4* methylation and expression.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The FTC omics data (DNA methylation and RNA sequencing) is part of the ‘Twin Study’ and deposited with the Biobank of the Finnish Institute for Health and Welfare [<https://thl.fi/en/research-and-development/thl-biobank/for-researchers/sample-collections/twin->

study]. For details on accessing the data, see: <https://thl.fi/en/research-and-development/thl-biobank/for-researchers/application-process>. The TwinsUK methylation dataset analyzed in the current study is available under ArrayExpress accession number [E-MTAB-1866](#). The expression dataset analyzed in the current study is available under EGA accession number [EGAS00001000805](#). All additional data access requests are overseen by the TwinsUK Resource Executive Committee (TREC). For information on access to these data and how to apply, see: <https://twinsuk.ac.uk/researchers/access-data-and-samples/request-access/>. The Scandinavian T2D-discordant cohort of MZ twins data (accession number LUDC2020.08.14 [<https://www.ludc.lu.se/resources/repository>]) are deposited in the Lund University Diabetes Center repository and while summary data and look-ups are available to academic researchers upon request through the repository portal, individual level data are not available due to ethical and legal restrictions related to Swedish Biobanks in Medical Care Act and European GDPR legislation. The EWAS summary statistics on mtDNA quantity in adipose and muscle tissues produced in this study are deposited to Figshare [<https://doi.org/10.6084/m9.figshare.26941927.v3>]. Source data are provided in this paper.

Code availability

The codes used in the study to run the EWAS and to perform the ICE FALCON analyses have been deposited to Figshare [<https://doi.org/10.6084/m9.figshare.26941927.v3>] public repository and are available to download.

References

- Matilainen, O., Quirós, P. M. & Auwerx, J. Mitochondria and epigenetics – crosstalk in homeostasis and stress. *Trends Cell Biol.* **27**, 453–463 (2017).
- Zhu, D., Li, X. & Tian, Y. Mitochondrial-to-nuclear communication in aging: an epigenetic perspective. *Trends Biochem. Sci.* **47**, 645–659 (2022).
- Oh, S.-Y. et al. Alternative usages of multiple promoters of the Acetyl-CoA carboxylase β gene are related to differential transcriptional regulation in human and rodent tissues*. *J. Biol. Chem.* **280**, 5909–5916 (2005).
- F. C. Lopes, A. Mitochondrial metabolism and DNA methylation: a review of the interaction between two genomes. *Clin. Epigenetics* **12**, 182 (2020).
- Atilano, S. R. et al. Mitochondrial DNA variants can mediate methylation status of inflammation, angiogenesis and signaling genes. *Hum. Mol. Genet.* **24**, 4491–4503 (2015).
- Lee, W. T. et al. Mitochondrial DNA haplotypes induce differential patterns of DNA methylation that result in differential chromosomal gene expression patterns. *Cell Death Discov.* **3**, 17062 (2017).
- Cortés-Pereira, E. et al. Differential association of mitochondrial DNA haplogroups J and H with the methylation status of articular cartilage: Potential role in apoptosis and metabolic and developmental processes. *Arthritis Rheumatol.* **71**, 1191–1200 (2019).
- Laaksonen, J. et al. Mitochondrial genome-wide analysis of nuclear DNA methylation quantitative trait loci. *Hum. Mol. Genet.* **31**, 1720–1732 (2022).
- Gong, Y. et al. Hypermethylation of Cox5a promoter is associated with mitochondrial dysfunction in skeletal muscle of high fat diet-induced insulin resistant rats. *PLoS ONE* **9**, e113784 (2014).
- Lee, W. et al. Mitochondrial DNA copy number is regulated by DNA methylation and demethylation of POLGA in stem and cancer cells and their differentiated progeny. *Cell Death Dis.* **6**, e1664–e1664 (2015).
- Ling, C. et al. Genetic and epigenetic factors are associated with expression of respiratory chain component NDUFB6 in human skeletal muscle. *J. Clin. Invest.* **117**, 3427–3435 (2007).
- Rönn, T. et al. Age influences DNA methylation and gene expression of COX7A1 in human skeletal muscle. *Diabetologia* **51**, 1159–1168 (2008).
- Heinonen, S. et al. Impaired mitochondrial biogenesis in adipose tissue in acquired obesity. *Diabetes* **64**, 3135–3145 (2015).
- van der Kolk, B. W. et al. Molecular pathways behind acquired obesity: Adipose tissue and skeletal muscle multiomics in monozygotic twin pairs discordant for BMI. *Cell Reports Med.* **2**, 100–226 (2021).
- Castellani, C. A. et al. Mitochondrial DNA copy number can influence mortality and cardiovascular disease via methylation of nuclear DNA CpGs. *Genome Med.* **12**, 84 (2020).
- Ejarque, M. et al. Adipose tissue mitochondrial dysfunction in human obesity is linked to a specific DNA methylation signature in adipose-derived stem cells. *Int. J. Obes.* **43**, 1256–1268 (2019).
- Lundgren, S. et al. BMI is positively associated with accelerated epigenetic aging in twin pairs discordant for body mass index. *J. Intern. Med.* **292**, 627–640 (2022).
- Li, S., Bui, M. & Hopper, J. L. Inference about causation from examination of familial confounding (ICE FALCON): a model for assessing causation analogous to Mendelian randomization. *Int. J. Epidemiol.* **49**, 1259–1269 (2020).
- Min, J. L. et al. Genomic and phenotypic insights from an atlas of genetic effects on DNA methylation. *Nat. Genet.* **53**, 1311–1321 (2021).
- Nilsson, E. et al. Differential DNA methylation and expression of miRNAs in adipose tissue from twin pairs discordant for type 2 diabetes. *Diabetes* **70**, 2402–2418 (2021).
- Nilsson, E. et al. Altered DNA methylation and differential expression of genes influencing metabolism and inflammation in adipose tissue from subjects with type 2 diabetes. *Diabetes* **63**, 2962–2976 (2014).
- Verdi, S. et al. TwinsUK: The UK adult twin registry update. *Twin Res. Hum. Genet.* **22**, 523–529 (2019).
- Rönn, T. et al. Impact of age, BMI and HbA1c levels on the genome-wide DNA methylation and mRNA expression patterns in human adipose tissue and identification of epigenetic biomarkers in blood. *Hum. Mol. Genet.* **24**, 3792–3813 (2015).
- Maddocks, O. D. K., Labuschagne, C. F., Adams, P. D. & Vousden, K. H. Serine metabolism supports the methionine cycle and DNA/RNA methylation through De Novo ATP synthesis in cancer cells. *Mol. Cell* **61**, 210–221 (2016).
- Lister, R. et al. Human DNA methylomes at base resolution show widespread epigenomic differences. *Nature* **462**, 315–322 (2009).
- Ball, M. P. et al. Targeted and genome-scale strategies reveal genobody methylation signatures in human cells. *Nat. Biotechnol.* **27**, 361–368 (2009).
- Shakhbazov, K. et al. Shared genetic control of expression and methylation in peripheral blood. *BMC Genomics* **17**, 278 (2016).
- Hemani, G., Tilling, K. & Davey Smith, G. Orienting the causal relationship between imprecisely measured traits using GWAS summary data. *PLOS Genet.* **13**, e1007081 (2017).
- Kreibich, E., Kleinendorst, R., Barzaghi, G., Kaspar, S. & Krebs, A. R. Single-molecule footprinting identifies context-dependent regulation of enhancers by DNA methylation. *Mol. Cell* **83**, 787–802 (2023).
- Kim, Y.-M. et al. SH3BP4 is a negative regulator of amino acid-rag GTPase-mTORC1 signaling. *Mol. Cell* **46**, 833–846 (2012).
- Kim, J. E. & Chen, J. Regulation of peroxisome proliferator-activated receptor- γ activity by mammalian target of rapamycin and amino acids in adipogenesis. *Diabetes* **53**, 2748–2756 (2004).
- Polak, P. et al. Adipose-specific knockout of $\langle \text{em} \rangle \text{raptor} \langle / \text{em} \rangle$ results in lean mice with enhanced mitochondrial respiration. *Cell Metab.* **8**, 399–410 (2008).
- Huwaitbieke, B. et al. Mammalian target of rapamycin signaling pathway regulates mitochondrial quality control of brown adipocytes in mice. *Frontiers in Physiology* **12**, 638352 (2021).

34. Hara, K. et al. Amino acid sufficiency and mTOR regulate p70 S6 kinase and eIF-4E BP1 through a common effector mechanism *. *J. Biol. Chem.* **273**, 14484–14494 (1998).
35. Sancak, Y. et al. The rag GTPases bind raptor and mediate amino acid signaling to mTORC1. *Science* **320**, 1496–1501 (2008).
36. Kim, E., Goraksha-Hicks, P., Li, L., Neufeld, T. P. & Guan, K.-L. Regulation of TORC1 by Rag GTPases in nutrient response. *Nat. Cell Biol.* **10**, 935–945 (2008).
37. DeYoung, M. P., Horak, P., Sofer, A., Sgroi, D. & Ellisen, L. W. Hypoxia regulates TSC1/2–mTOR signaling and tumor suppression through REDD1-mediated 14–3–3 shuttling. *Genes Dev.* **22**, 239–251 (2008).
38. Mendelson, M. M. et al. Association of body mass index with DNA methylation and gene expression in blood cells and relations to cardiometabolic disease: A mendelian randomization approach. *PLoS Med.* **14**, e1002215–e1002215 (2017).
39. Wahl, S. et al. Epigenome-wide association study of body mass index, and the adverse outcomes of adiposity. *Nature* **541**, 81–86 (2017).
40. Neeland, I. J., Poirier, P. & Després, J.-P. Cardiovascular and metabolic heterogeneity of obesity. *Circulation* **137**, 1391–1406 (2018).
41. Kaaman, M. et al. Strong association between mitochondrial DNA copy number and lipogenesis in human white adipose tissue. *Diabetologia* **50**, 2526–2533 (2007).
42. Agius, R., Pace, N. P. & Fava, S. Reduced leukocyte mitochondrial copy number in metabolic syndrome and metabolically healthy obesity. *Front. Endocrinol.* **13**, 886957 (2022).
43. Liesa, M. & Shirihai, O. S. Mitochondrial dynamics in the regulation of nutrient utilization and energy expenditure. *Cell Metab.* **17**, 491–506 (2013).
44. Zhang, F. F. et al. Dietary patterns are associated with levels of global genomic DNA methylation in a cancer-free population. *J. Nutr.* **141**, 1165–1171 (2011).
45. Perfilov, A. et al. Impact of polyunsaturated and saturated fat overfeeding on the DNA-methylation pattern in human adipose tissue: a randomized controlled trial1, 2, 3. *Am. J. Clin. Nutr.* **105**, 991–1000 (2017).
46. Mengel-From, J. et al. Mitochondrial DNA copy number in peripheral blood cells declines with age and is associated with general health among elderly. *Hum. Genet.* **133**, 1149–1159 (2014).
47. Zhang, R., Wang, Y., Ye, K., Picard, M. & Gu, Z. Independent impacts of aging on mitochondrial DNA quantity and quality in humans. *BMC Genomics* **18**, 890 (2017).
48. Liu, Z. et al. Underlying features of epigenetic aging clocks in vivo and in vitro. *Aging Cell* **19**, e13229 (2020).
49. Li, S. et al. Causal effect of smoking on DNA methylation in peripheral blood: a twin and family study. *Clin. Epigenetics* **10**, 18 (2018).
50. Nissen, F. I. et al. Is there a causal relationship between physical activity and bone microarchitecture? A study of adult female twin pairs. *J. Bone Miner. Res.* **38**, 951–957 (2023).
51. Bui, M. et al. Population-based cohort study of oral contraceptive use and risk of depression. *Epidemiol. Psychiatr. Sci.* **32**, e39 (2023).
52. Minică, C. C., Dolan, C. V., Boomsma, D. I., de Geus, E. & Neale, M. C. Extending causality tests with genetic instruments: An integration of mendelian randomization with the classical twin design. *Behav. Genet.* **48**, 337–349 (2018).
53. Rose, R. J. et al. FinnTwin12 cohort: An updated review. *Twin Res. Hum. Genet.* **22**, 302–311 (2019).
54. Kaidesoja, M. et al. FinnTwin16: A longitudinal study from age 16 of a population-based finnish twin cohort. *Twin Res. Hum. Genet.* **22**, 530–539 (2019).
55. Kaprio, J. et al. The older finnish twin cohort — 45 years of follow-up. *Twin Res. Hum. Genet.* **22**, 240–254 (2019).
56. Sarna, S., Kaprio, J., Sistonen, P. & Koskenvuo, M. Diagnosis of twin zygosity by mailed questionnaire. *Hum. Hered.* **28**, 241–254 (1978).
57. Baecke, J. A., Burema, J. & Frijters, J. E. R. A short questionnaire for the measurement of habitual physical activity in epidemiological studies. *Am. J. Clin. Nutr.* **36**, 936–942 (1982).
58. Mustelin, L. et al. Genetic influences on physical activity in young adults: A twin study. *Med. Sci. Sport. Exerc.* **44**, 1293–1301 (2012).
59. Sillanpää, E. et al. Blood and skeletal muscle ageing determined by epigenetic clocks and their associations with physical activity and functioning. *Clin. Epigenetics* **13**, 110 (2021).
60. Lapatto, H. A. K. et al. Nicotinamide riboside improves muscle mitochondrial biogenesis, satellite cell differentiation, and gut microbiota in a twin study. *Sci. Adv.* **9**, eadd5163 (2023).
61. Min, J. L., Hemani, G., Davey Smith, G., Relton, C. & Suderman, M. Meffil: efficient normalization and analysis of very large DNA methylation datasets. *Bioinformatics* **34**, 3983–3989 (2018).
62. Chen, Y. et al. Discovery of cross-reactive probes and polymorphic CpGs in the Illumina Infinium HumanMethylation450 microarray. *Epigenetics* **8**, 203–209 (2013).
63. Zhou, W., Laird, P. W. & Shen, H. Comprehensive characterization, annotation and innovative use of Infinium DNA methylation Bead-Chip probes. *Nucleic Acids Res.* **45**, e22–e22 (2017).
64. Du, P. et al. Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. *BMC Bioinform.* **11**, 587 (2010).
65. Hannum, G. et al. Genome-wide methylation profiles reveal quantitative views of human aging rates. *Mol. Cell* **49**, 359–367 (2013).
66. Horvath, S. DNA methylation age of human tissues and cell types. *Genome Biol.* **14**, R115 (2013).
67. Lu, A. T. et al. DNA methylation GrimAge strongly predicts lifespan and healthspan. *Aging* **11**, 303–327 (2019).
68. Levine, M. E. et al. An epigenetic biomarker of aging for lifespan and healthspan. *Aging* **10**, 573–591 (2018).
69. Higgins-Chen, A. T. et al. A computational solution for bolstering reliability of epigenetic clocks: implications for clinical trials and longitudinal tracking. *Nat. Aging* **2**, 644–661 (2022).
70. Belsky, D. W. et al. DunedinPACE, a DNA methylation biomarker of the pace of aging. *Elife* **11**, e73420 (2022).
71. Voisin, S. et al. An epigenetic clock for human skeletal muscle. *J. Cachexia. Sarcopenia Muscle* **11**, 887–898 (2020).
72. Ritchie, M. E. et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **43**, e47–e47 (2015).
73. Teschendorff, A. E., Breeze, C. E., Zheng, S. C. & Beck, S. A comparison of reference-based algorithms for correcting cell-type heterogeneity in Epigenome-Wide Association Studies. *BMC Bioinform.* **18**, 105 (2017).
74. Højsgaard, S., Halekoh, U. & Yan, J. The R package geepack for generalized estimating equations. *J. Stat. Softw.* **15**, 1–11 (2005).
75. Grundberg, E. et al. Global analysis of DNA methylation variation in adipose tissue from twins reveals links to disease-associated variants in distal regulatory elements. *Am. J. Hum. Genet.* **93**, 876–890 (2013).
76. Christiansen, C. et al. Adipose methylome integrative-omic analyses reveal genetic and dietary metabolic health drivers and insulin resistance classifiers. *Genome Med.* **14**, 75 (2022).

Acknowledgements

The authors thank all the study participants. We also acknowledge the computational resources of the Institute for Molecular Medicine Finland (FIMM) Technology Center. Open access was facilitated by Helsinki University Library. S.Lu. was involved in this research while affiliated with the Institute for Molecular Medicine Finland (FIMM) and is currently affiliated with Nightingale Health Plc. This study is supported by the following funds: University of Helsinki, Faculty of Medicine, Doctoral School of Population Health (DOCPop) (A.He.), an Australian Government Research Training Program (RTP) Scholarship (V.F.C.E.), Academy

of Finland (#328685, #307339, #297908 and #251316, M.O.; #338417, S.H.; #335443, #314383, #266286, K.H.P.), Academy of Finland Center of Excellence in Complex Disease Genetics (#352792) (J.K.) and Center of Excellence in Research on Mitochondria, Metabolism and Disease (FinMIT) (#272376) (K.H.P.), Sigrid Juselius Foundation (M.O.), Liv o Hälsa society (M.O.), Minerva Foundation (M.O.), Novo Nordisk Foundation (#NNF20OC0060547, #NNF17OC0027232, #NNF10OC1013354, K.H.P.; #NNF23SA0083953, S.H.), National Health and Medical Research Council Investigator Grant (GNT2017373) (S.Li.), Diabetes Research Foundation (S.H., K.H.P.), Paulo Foundation (S.H., K.H.P.), Gyllenberg Foundation (K.H.P.) Finnish Medical Foundation (K.H.P.), University of Helsinki, and Helsinki University Hospital (K.H.P., S.H.), and Government Research Funds (K.H.P.). TwinsUK is funded by the Wellcome Trust, European Community's Seventh Framework Program (FP7/2007-2013), Medical Research Council, Versus Arthritis, European Union Horizon 2020, Chronic Disease Research Foundation (CDRF), Zoe Ltd and the National Institute for Health Research (NIHR)-funded BioResource, Clinical Research Facility and Biomedical Research Center based at Guy's and St Thomas' NHS Foundation Trust in partnership with King's College London. This project also received support from the JPI ERA-HDHL DIMENSION project and UK Biological Sciences Research Council (BBSRC, BB/S020845/1 and BB/T019980/1 to J.T.B.). The data generated in the Scandinavian twin cohort were supported by the Swedish Research Council, Region Skåne (ALF), Strategic Research Area Exodiab (Dnr 2009-1039), the Novo Nordisk Foundation, the Swedish Foundation for Strategic Research (Dnr IRC15-0067), the Crafoord Foundation, and the Swedish Diabetes Foundation.

Author contributions

A. He. and M.O. designed the study. A. He. performed the data analyses, visualized the data, and wrote the first draft of the manuscript. K.H.P., S.H., M.O., and J.Ka. collected and generated the FTC data. A. He. and S.H.T.L. preprocessed the omics data. J.Ku., A.Ha., and J.L. performed the MRI and P-H.G. the DEXA imaging. P.P. generated the RNA sequencing data. V.F.C.E. and S.Li. developed and assisted on the ICE FALCON analysis. S.V. and J.T.B. replicated the findings in TwinsUK, and A.M., C.L., E.N., and A.V. in the Scandinavian T2D-discordant MZ cohort. M.O. and S.Lu. Supervised the work. All authors critically commented and edited the final version of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-025-59576-7>.

Correspondence and requests for materials should be addressed to Aino Heikkinen or Miina Ollikainen.

Peer review information *Nature Communications* thanks Christopher Mason, Afshin Beheshti, and the other anonymous reviewers for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2025

¹Institute for Molecular Medicine Finland (FIMM), HiLIFE, University of Helsinki, Helsinki, Finland. ²Minerva Foundation Institute for Medical Research, Helsinki, Finland. ³Centre for Epidemiology and Biostatistics, Melbourne School of Population and Global Health, University of Melbourne, Melbourne, VIC, Australia. ⁴Department of Human Genetics, David Geffen School of Medicine at UCLA, Los Angeles, CA, USA. ⁵HUS Medical Imaging Center, Radiology, University of Helsinki and Helsinki University Hospital, Helsinki, Finland. ⁶Institute for Clinical Diabetology, German Diabetes Center, Leibniz Center for Diabetes Research, Heinrich Heine University, Düsseldorf, Germany. ⁷Public Health Promotion Unit, National Institute for Health and Welfare, Helsinki, Finland. ⁸Folkhälsan Institute of Genetics, Folkhälsan Research Center, Helsinki, Finland. ⁹Research Program for Clinical and Molecular Metabolism, Faculty of Medicine, University of Helsinki, Helsinki, Finland. ¹⁰Abdominal Center, Nephrology, University of Helsinki and Helsinki University Hospital, Helsinki, Finland. ¹¹Obesity Research Unit, Research Program for Clinical and Molecular Metabolism, Faculty of Medicine, University of Helsinki, Helsinki, Finland. ¹²Department of Internal Medicine, Helsinki University Hospital, Helsinki, Finland. ¹³Department of Twin Research and Genetic Epidemiology, King's College London, London, UK. ¹⁴Epigenetics and Diabetes Unit, Department of Clinical Sciences in Malmö, Lund University Diabetes Centre, Scania University Hospital, Malmö, Sweden. ¹⁵Department of Clinical Sciences in Malmö, Lund University Diabetes Centre, Scania University Hospital, Malmö, Sweden. ¹⁶Copenhagen University Hospital, Steno Diabetes Center Copenhagen, Herlev, Denmark. ¹⁷Department of Endocrinology, Skåne University Hospital, Malmö, Sweden. ¹⁸Bioinformatics Interdepartmental Program, UCLA, Los Angeles, CA, USA. ¹⁹Institute for Precision Health, David Geffen School of Medicine at UCLA, Los Angeles, CA, USA. ²⁰HealthyWeightHub, Endocrinology, Abdominal Center, Helsinki University Central Hospital and University of Helsinki, Helsinki, Finland. ²¹Precision Medicine, School of Clinical Sciences at Monash Health, Monash University, Clayton, Victoria, Australia. ✉ e-mail: aino.heikkinen@helsinki.fi; miina.ollikainen@helsinki.fi