<http://www.diva-portal.org>

Postprint

This is the accepted version of a paper published in *International Journal of Obesity*. This paper has been peer-reviewed but does not include the final publisher proof-corrections or journal pagination.

Citation for the original published paper (version of record):

Karlsson, I K., Ericsson, M., Wang, Y., Jylhävä, J., Hägg, S. et al. (2020)
Replicating associations between DNA methylation and body mass index in a
longitudinal sample of older twins
International Journal of Obesity, 44(6): 1397-1405
<https://doi.org/10.1038/s41366-019-0498-6>

Access to the published version may require subscription.

N.B. When citing this work, cite the original published paper.

Permanent link to this version:

<http://urn.kb.se/resolve?urn=urn:nbn:se:hj:diva-47179>

Replicating associations between DNA methylation and body mass index in a longitudinal sample of older twins

Running title: Methylation and change in body mass index

Ida K. Karlsson^{1,2}, Malin Ericsson², Yunzhang Wang², Juulia Jylhävä², Sara Hägg², Nancy L. Pedersen^{2,3}, Chandra A. Reynolds⁴, Anna K. Dahl Aslan^{1,2}

¹Institute of Gerontology and Aging Research Network – Jönköping (ARN-J), School of Health and Welfare, Jönköping University, Jönköping, Sweden

²Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, Stockholm, Sweden

³Department of Psychology, University of Southern California, Los Angeles, California, USA

⁴Department of Psychology, University of California, Riverside, USA

Corresponding author: Ida K. Karlsson, Hälsohögskolan, Box 1026, SE-551 11 Jönköping, Sweden. Tel: +46 736 939 589. Email: ida.karlsson@ju.se

Competing Interests: The authors declare no competing financial interests.

Abstract

Background: There is important interplay between epigenetic factors and body weight, and previous work has identified ten sites where DNA methylation is robustly associated with body mass index (BMI) cross-sectionally. However, interpretation of the associations is complicated by the substantial changes in BMI often occurring in late-life, and the fact that methylation is often driven by genetic variation. This study therefore investigated the longitudinal association between these ten sites and BMI from midlife to late-life, and whether associations persist after controlling for genetic factors.

Methods: We used data from 535 individuals (mean age 68) in the Swedish Adoption/Twin Study of Aging (SATSA) with longitudinal measures of both DNA methylation from blood samples and BMI, spanning up to 20 years. Methylation levels were measured with the Infinium Human Methylation 450K or Infinium MethylationEpic array, with seven of the ten sites passing quality control. Latent growth curve models were applied to investigate longitudinal associations between methylation and BMI, and between-within models to study associations within twin pairs, thus adjusting for genetic factors.

Results: Baseline DNA methylation levels at five of the seven sites were associated with BMI level at age 65 (cg00574958 [*CPT1A*]; cg11024682 [*SREBF1*]) and/or change (cg06192883 [*MYO5C*]; cg06946797 [*RM12*]; cg08857797 [*VPS25*]). For four of the five sites, the associations remained comparable within twin pairs. However, the effects of cg06192883 were substantially attenuated within pairs. No change in DNA methylation was detected for any of the seven evaluated sites.

Conclusion: Five of the seven sites investigated were associated with late-life level and/or change in BMI. The effects for four of the sites remained similar when examined within twin pairs, indicating that the associations are mainly environmentally driven. However, the substantial attenuation in the association between cg06192883 and late-life BMI within pairs points to the importance of genetic factors in this association.

Introduction

Body weight is subject to change throughout different time-periods of our lives. While most individuals gain weight from young adulthood to late midlife, there is greater heterogeneity in old age, where some continue to gain weight, some maintain a stable weight, and a substantial proportion of older individuals lose weight¹. As changes in weight in old age may be important markers of underlying disease and survival^{2,3}, it is important to unravel the etiology of these changes.

Genetic factors are important determinants of body composition, and the heritability of body mass index (BMI) has been estimated to 45-85% in twin studies⁴. However, several studies have shown that the heritability of BMI as well as the genetic variants associated differ across age groups⁴⁻⁶. Epigenetic mechanisms are important players in the biological aging process and may be a mechanism through which gene-environment interactions work⁷. DNA methylation refers to addition of a methyl group to a cytosine located next to a guanine in the genome, in the context of a dinucleotide called a CpG site⁸. Promoter regions are often rich in CpG sites, where DNA methylation inhibits binding of transcription factors and hence downregulates gene expression. Epigenome-wide association studies (EWASs), where methylation at CpG sites across the entire genome are analyzed, have identified several associations between leukocyte DNA methylation and BMI or obesity⁹⁻¹⁴. A comparison of findings from the three largest EWASs⁹⁻¹¹ revealed ten sites robustly associated with BMI in all three studies¹⁵. Seven of the ten sites have also been identified in associations with other factors related to the metabolic syndrome, such as lipid fractions, glucose levels, and inflammation¹⁵, further highlighting the importance of the sites. Two of the EWASs investigated the association between methylation and change in BMI by modelling the difference in BMI between two time points. One of the studies evaluated change in BMI from age 25 to study baseline at mean age 56⁹, and the other change in BMI and in DNA methylation over seven years from mean age 61 at baseline¹¹. All three studies accounted for genetic factors by studying associations between nearby *cis*-methylation quantitative trait loci (meQTLs) and methylation levels. However, many

meQTLs influence methylation through *trans* effects, i.e. at sites further away in the genome¹⁶, complicating interpretations of the findings. None of the previous studies focused specifically on changes in BMI in late-life and very late-life, and with substantial changes occurring in both BMI¹ and methylation⁷ during aging, this relationship requires specific attention.

The objective of this study was to thoroughly investigate the longitudinal association between methylation at these ten CpG sites and BMI during late-life, by using multiple measures of BMI and DNA methylation spanning up to 20 years while accounting for familial factors (such as meQTLs) through twin designs. Specifically, we investigated: 1) whether DNA methylation at these CpG sites are also associated with the non-linear longitudinal changes in BMI during late-life, 2) whether levels in DNA methylation at these sites change during late-life, and if so whether those changes are associated with changes in BMI, and 3) whether the associations remain when comparing within twin pairs, hence controlling for shared genetic and other familial factors. By focusing on methylation at sites robustly associated with BMI across adulthood and investigating associations with both level and change in BMI during late-life, we thus aimed to promote a better understanding of whether the same methylomic factors are associated with changes in BMI during late-life as for level of BMI earlier in life.

Material and methods

Study population

The study population originated from the Swedish Adoption/Twin Study of Aging (SATSA)¹⁷, a sub-study of the Swedish Twin Registry (STR)¹⁸. SATSA is a longitudinal study of same-sex twin pairs with up to 10 in-person testing occasions between 1984 and 2014, consisting of a health examination, cognitive tests, and collection of blood samples. In total, 859 individuals participated in at least one testing occasion, out of whom 535 had information on DNA methylation and hence were included in this study. Baseline was

set at first available blood sample with methylation information. All participants provided informed consent, and the study was approved by the Regional Ethics board at Karolinska Institutet, Stockholm.

DNA methylation measurements and pre-processing

DNA methylation was analyzed in blood samples collected during the third (1992-1994), fifth (1999-2002), sixth (2002-2005), eighth (2008-2010), ninth (2010-2012), and tenth (2012-2014) in-person testing occasions. Up to six measurements were thus available from the 535 individuals, leading to a total of 1402 samples. Table S2 shows the number of samples with methylation information and their mean age at each in-person testing occasion.

DNA was extracted and bisulfite converted using EZ-96 DNA MagPrep methylation kit (Zymo Research Corp., Orange, CA, USA) according to the manufacturer's protocol. The DNA was then hybridized onto the Infinium Human Methylation 450K Bead Chip (n=385), or the Infinium MethylationEPIC BeadChip (n=150, both from Illumina Inc., San Diego, CA, USA). The obtained methylation data then went through the following rigorous multi-step quality control pipeline: 1) removal of samples with poor correlation to genotype controls; 2) removal of samples with the wrong predicted sex based on signal ratio of probes on sex chromosomes; 3) removal of probes overlapping a SNP site; 4) removal of probes residing on sex chromosomes; 5) removal of probes with detection p-value above 0.05. Seven out of the ten CpG sites of interest remained after these quality control steps (cg06192883, cg11024682, cg00574958, cg08857797, cg06946797, cg13123009, and cg09664445). Processing was done in R with background correction with methylumi.noob¹⁹, and normalization with wateRmelon.dasen²⁰. Cell counts were adjusted for using the Houseman method²¹, and batch effects using the ComBat function in the sva package²². Methylation levels at each of the seven sites were transformed to M-values by calculating the logit2-transformed ratios of methylated to total (methylated plus unmethylated) probe intensity²³. The M-values were then standardized to mean zero and standard deviation of one for easier interpretation.

BMI measurements

The participant's height and weight were measured at each testing occasion by trained research nurses, and BMI was calculated as kg/m^2 . All available BMI measurements from the first in-person testing occasion with methylation information and onwards were included. Table S2 shows the number of samples with BMI measurement and their mean age at each occasion.

Covariates

Smoking status and diagnosis of type 2 diabetes was self-reported at the testing occasions. Smoking was categorized into current smoking or not smoking at baseline. Type 2 diabetes was categorized into lifetime presence or absence of disease. Blood levels of glucose, total cholesterol, and triglycerides were measured at each testing occasion together with fasting status, and levels at first blood sample with methylation data were included in the current study.

Statistical analyses

All analyses were performed in R version 3.5.2. To investigate the temporal association between DNA methylation at the sites of relevance and BMI, the following three steps were taken.

1. Study the cross-sectional association between DNA methylation and BMI at baseline

The cross-sectional association between baseline measures of methylation levels at each CpG site and BMI was analyzed first using linear regression in the full sample, with robust sandwich estimators to account for the relatedness between twins.

To obtain estimates adjusted for genetic factors, we applied between-within analyses²⁴. Here, the between-pair estimate examines the exposure-outcome association across all twin pairs and represents the average effect in the population. The within-pair estimate examines the exposure-outcome association within twin pairs, and thus represents the effect not attributable to shared genetic and other

familial factors. This was done using linear mixed models (lme4 package) with both the twin-pair mean methylation (between-pair effect) and individual deviation from the twin-pair mean (within-pair effect) as predictors of BMI. Individual IDs nested within twin pair IDs were included as random effects.

All models were adjusted for age, sex, smoking status at baseline, and type of methylation array (Infinium Human Methylation 450K Bead Chip or Infinium MethylationEPIC BeadChip).

To study potential environmental drivers of the associations, mediation analyses were carried out investigating the effect of type 2 diabetes, blood glucose, total cholesterol, and triglycerides. This was done by 1) modelling the effect of DNA methylation on BMI, 2) modelling the effect of DNA methylation on the mediator, and 3) modelling the effect of DNA methylation on BMI, adjusting for the mediator. The significance of the mediation was tested through bootstrapping with 1000 repetitions (mediate package). All models, except that adjusting for type 2 diabetes, were also adjusted for fasting status at time of blood sample. All were adjusted for age, sex, smoking and type of methylation array, as above.

2. Study the association between baseline DNA methylation and longitudinal change in BMI

We investigated if baseline DNA methylation at each CpG site was associated with longitudinal change in BMI by applying latent growth curve models²⁵ with age as the time scale. The models included fixed effects, linear, and quadratic trends, with age centered at 65 years (as a turning point in the BMI curve has been shown to occur at age 65⁴). Methylation level at baseline was included as a fixed effect on BMI at the intercept (reflecting BMI at the centering age of 65), and in interaction with linear and quadratic age to investigate the effect of baseline methylation levels on longitudinal BMI trajectory features. Individuals nested within twin pairs were modelled as random effects to account for relatedness of twins.

To adjust for shared familial factors, between-within analyses were applied by modelling the baseline twin-pair mean methylation level (between-pair effect) and individual deviation from the twin pair mean (within-pair effect) as fixed effects and in interaction with linear and quadratic age.

Sex, smoking status at baseline, and methylation array were included as fixed effects in all the models described above. To deal with convergence issues, the twin pair ID was omitted from the random quadratic age effect in latent growth curve models of cg06192883 and cg11024682.

3. Explore characteristics of the CpG sites

To identify late-life changes in DNA methylation levels at the seven sites of interest, latent growth curve models were applied, with age centered at 65. The model included sex, smoking at the time of blood sample, and methylation array as fixed effects on methylation levels at the intercept age, and a linear age effect was subsequently added to the model. Individual ID nested within twin pair ID was included as a random effect. Likelihood ratio tests were applied to investigate the difference in model fit when introducing the linear growth term, i.e. the evidence of change in methylation levels over time.

In order to identify evidence of genetic factors driving methylation levels at the seven sites, twin-pair correlations were first calculated for monozygotic and dizygotic twin pairs separately, and the estimates compared. The Falconer heritability, a crude measure of the unadjusted broad-sense heritability, was calculated by doubling the difference between the monozygotic and dizygotic twin-pair correlations²⁶. To identify *cis*-meQTLs (within 1 million base pairs) driving methylation levels at these sites, results were extracted from a previous study by Wang et al.²⁷, specifically modelling *cis*-meQTLs in this sample. To confirm the presence or absence of *cis*-meQTLs and to identify *trans*-meQTLs, lookup was performed in the mQTL database²⁸, filtering on middle-age individuals.

For all models, nominal significance level was set at the $\alpha=0.05$ level and a conservative Bonferroni corrected significance level adjusted for the seven sites tested at $\alpha=0.007$ ($0.05/7$).

Results

Study population

Among the 535 individuals in the sample, 313 (58.5%) were women and 222 (41.5%) were men. At baseline (first blood sample collection), mean age was 68.2 years (SD=9.5, range 48-94), and mean BMI was 26.4 (SD=4.2, range 17.8-50.1). The average follow-up time for BMI measurements was 9.7 years (SD=6.9, range 0-20.6), during which BMI was measured on average 4.0 times (SD=2.3, range 1-8). A total of 161 individuals had BMI measures both before and after age 65. The average follow-up time for methylation samples was 8.4 years (SD=6.8, range 0-20.2), with a mean of 2.6 samples per individual (SD=1.3, range 1-6). Ninety-five individuals were smokers at baseline (17.8%). The sample included 238 complete twin pairs, out of which 82 were monozygotic pairs and 156 were dizygotic.

The cross-sectional association between DNA methylation and BMI at baseline

At baseline, linear regression models demonstrated that DNA methylation at cg06192883 and cg11024682 were significantly associated with BMI in the full sample after conservative Bonferroni correction (Table 1). When comparing within twin pairs, the effect of cg06192883 was substantially reduced and the association was no longer significant (Table 1). When DZ and MZ pairs were analyzed separately, the association was similar to that in the full sample within DZ twins ($\beta=0.71$, $p=0.03$), but was substantially diminished within MZ twin pairs ($\beta=-0.18$, $p=0.56$, Table S3). For cg11024682, the effect size within twin pairs remained similar to that of the full sample and the association was nominally significant (Table 1).

The association between BMI, cg00574958 and cg08857797 reached nominal significance in the full sample, with comparable estimates within twin pairs (Table 1). None of the remaining three CpG sites (cg06946797, cg13123009, and cg09664445) were associated with BMI at baseline. Compared to the

three previous studies¹⁵, all estimates were in the same direction, except cg13123009 which had an effect size very close to zero.

Follow-up analyses studying mediating effects of type 2 diabetes, blood glucose levels, total cholesterol, and triglycerides revealed that the associations with both cg06192883 and cg11024682 were significantly mediated through blood glucose levels ($p=0.006$ and $p=0.01$, respectively) and triglycerides ($p<0.001$ and $p=0.006$, respectively, Table S5). Triglycerides also mediated the association between cg00574958 and BMI ($p=0.002$). No mediating effects were found for cg08857797.

Baseline DNA methylation and longitudinal change in BMI

Figure 1 visualizes the longitudinal change in BMI, stratified by tertiles of methylation levels at each of the seven sites. DNA methylation at cg06192883 was significantly associated with a higher BMI at age 65 as well as with linear and quadratic change in BMI (Figure 1). Again, the effect was substantially reduced when comparing within twin pairs (Table 2). DNA methylation at cg11024682 was significantly associated with higher BMI at the intercept age of 65 but not with change in BMI (Figure 1), and as in the baseline models the effect size remained similar when comparing within twin pairs (Table 2).

Nominally significant associations were found between cg00574958 and lower BMI at age 65, cg08857797 with higher BMI at age 65 and a steeper linear decline, and for cg06946797 and lessened quadratic change in BMI (Figure 1). When comparing within twin pairs, the effects of cg00574958 and cg06946797 were similar to that in the full sample (Table 2). The within-pair effect of cg08857797 was similar for BMI at the intercept age, but stronger for change in BMI.

Characteristics of the CpG sites

While there was significant linear change in the methylation levels at cg11024682, cg00574958, and cg09664445, the estimates for the linear slope were negligible ($\beta=-0.01$ for all) and the random slope did

not significantly improve model fit (Table S4). Hence, investigations of change in methylation at these sites would not be of relevance, and therefore no further models of longitudinal change in DNA methylation in relation to BMI were applied.

Comparison of twin-pair correlations between monozygotic and dizygotic twin pairs indicated a broad-sense heritability of 13-21 % for cg06192883, cg08857797, cg13123009, and cg09664445 (Table S6). cg06946797 showed a high heritability estimate of 54%, and the correlation pattern indicates the presence of dominance effects. Heritability could not be estimated for cg11024682 as the twin-pair correlation was higher among dizygotic twins than among monozygotic twins. This is likely a result of imprecision in the estimate, as indicated by the wide confidence intervals. For cg06192883 and cg13123009, the twin-pair correlation among monozygotic pairs exceeded the heritability estimate, indicating presence of additional familial factors (shared environment). No evidence of *cis*-meQTLs affecting any of the sites were identified in this sample. Lookup in the mQTL database²⁸ only revealed *trans*-meQTL for cg09664445 (rs13182794 and rs115422955, both located in the same region on chromosome 5), which was not associated with BMI in any of the analyses. Further investigation of meQTLs were therefore not carried out.

Discussion

Out of the seven CpG sites analyzed, all robustly associated with BMI in the three largest EWASs to date, we demonstrated that five are also of importance to late-life BMI level (cg11024682 and cg00574958) and/or change in BMI (cg06192883, cg08857797, and cg06946797). Among four of the five sites, the associations remained significant within twin pairs. However, for cg06192883, which was strongly associated with both higher BMI level at age 65 and with steeper decline in BMI, the associations were substantially reduced within twin pairs, indicating they are largely driven by genetic or other factors

shared within families. No evidence of longitudinal change in methylation levels was detected for any of the CpG sites, and therefore no tests of change in methylation as a function of BMI were carried out.

The three EWASs included in the review by Bell¹⁵ were all of substantial size, with both discovery and replication samples of more than 2000 individuals⁹⁻¹¹. Since then, another three EWASs of BMI of substantial size have been published, but based on smaller samples¹²⁻¹⁴. Interestingly, three of the ten sites were also identified in each of these studies, namely cg11024682, cg00574958, and cg06500161.

Unfortunately, the third site was lost during quality control in the current study, but the other two were both associated with BMI at age 65, cg11024682 being a top hit remaining significantly associated with higher BMI at age 65 after conservative Bonferroni correction. Further strengthening the importance of the three sites, one of the earlier EWASs identified them in three-way interactions between BMI, DNA methylation, and gene expression¹⁰. The same study demonstrated a causal effect of cg11024682 on BMI, and a causal effect of BMI on cg06500161. The reason cg06192883 was not identified in these more recent EWASs may be that the association is age-dependent, as the site was strongly associated not only with level but also with change in BMI during late-life in this study. The site is located within the myosin 5C (*MYO5C*) gene, coding for an intracellular transporter protein mainly expressed in glandular tissues²⁹.

Little is known of the biological mechanisms around MYO5C, but as a key protein in intracellular transport in pancreatic cells, it has been implicated in insulin excretion³⁰, possibly explaining its association with BMI. Moreover, methylation at the same site has been found to be associated with levels of the inflammatory marker C-reactive protein (CRP) in blood, plausibly stemming from the chronic inflammation associated with adiposity¹⁵. Our follow-up analyses indicated that the association between cg06192883 and BMI may be mediated through blood glucose levels and triglycerides. The other four sites have all been identified in associations with lipid fractions¹⁵. Notably, cg11024682 resides in the Sterol Regulatory Element Binding Transcription Factor 1 (*SREBF1*) gene, which is involved in energy homeostasis and lipid metabolism¹⁰. *SREBF1* has been identified in EWASs of type-2 diabetes³¹, and both

SREBF1 and *CPT1A* (where cg00574958 resides) in EWASs of dyslipidemia³². In line with these findings, the current study indicated that blood glucose levels may mediate the association between cg11024682 and BMI, and that triglyceride levels may mediate the associations with both cg11024682 and cg00574958.

One recent study utilized twin data to assess the bi-directional causality between methylation and BMI³³. They could thereby confirm the results from Wahl et al.¹¹, who through Mendelian randomization methods showed that, for a majority of the CpG sites investigated, BMI causally influences methylation levels rather than the opposite. The associations between baseline DNA methylation levels and prospective longitudinal change in BMI might be interpreted as evidence of methylation driving late-life changes in BMI. However, as both BMI and methylation are complex biological mechanisms, influenced by a variety of not only genetic factors but also environmental factors throughout life, any interpretations of directionality should be made with caution. The lack of change in methylation levels in the current study does not necessary speak against BMI causally influencing methylation at these sites. It is plausible that past levels of BMI levels does determine the methylation levels at these sites, and that they then remain stable throughout late-life. Moreover, obesity is one of the major players in the metabolic syndrome, defined as having a combination of three or more of the factors high blood pressure, abdominal obesity, high triglyceride levels, low levels of high-density lipoproteins, and elevated fasting blood glucose³⁴. The metabolic syndrome increases the risk of several chronic diseases, such as cardiovascular disease and type 2 diabetes, and it is plausible that this complex syndrome influences blood methylation levels in a long-term manner. Further work on the topic is warranted to address this complex issue.

While the moderate sample size is a limitation, this study is based on a well-established twin-cohort with longitudinal measures of both BMI and methylation levels. Much evidence indicates that the causes and consequences of BMI in late-life may differ from those in midlife, and the factors influencing late-life

changes in BMI are poorly understood. It would hence be plausible that the associations between DNA methylation and BMI are age specific. By using BMI measured at several time points throughout late-life we could further investigate associations between DNA methylation and both level and changes in BMI from midlife to late-life. Studying older individuals does entail methodological considerations, such as attrition rate and survival bias³⁵. Studies including older individuals are always limited in the sense that there is a selection bias for healthier individuals, and our sample with a mean age of 68 years at baseline is no exception. As seen in Table S2, the increase in mean age across testing occasion does not correspond to the number of years between waves, indicating that older individuals drop out to a higher extent than the younger individuals. The attrition rates between the last IPTs is low, around 5%, and predominantly due to mortality or poor health. The selection bias based on BMI are most often in favor for those with higher BMI in late-life compared to those with normal BMI, due to the previously mentioned obesity paradox. However, it is important to consider that the selected CpG sites are mainly identified in younger samples, and it is possible that individuals who were overweight or obese in midlife never entered the study, as they might have had worse health earlier in life. This might reduce our ability to identify true effects. The interplay between genetic variants and methylation adds a substantial complexity level to epigenetic studies. Almost 20% of the variance in DNA methylation in blood is driven by genetic factors, the majority acting through *trans* effects where the controlling genetic variant resides distant to the CpG site in the genome²⁸. Despite indications of heritable influences on methylation at several sites investigated in this study only one meQTL could be identified, further highlighting this issue. The twin design is hence highly valuable in methylation studies, especially in monozygotic twins where it offers complete adjustment for genetic influences³⁶. However, twin designs also have caveats worth considering. Although within-pair estimates are naturally adjusted for factors shared between the twins, the estimates may still be affected by bias from non-shared confounders and measurement errors³⁷. While DNA methylation levels in adipose tissue certainly would have been of interest to study, blood

samples have the advantage, in addition to being easier to access, that it is a metabolically active tissue. It hence has important roles in the inflammatory and vascular processes associated with adiposity. Unfortunately, cell type composition was not available, but adjusted for using the well-established Houseman method²¹. While the method is widely used and works very well in blood samples, it is important to note that the method only corrects for the major cell types present in whole blood and it is possible that effects stemming from differences in smaller cell type fractions remain³⁸.

In conclusion, five of the seven sites investigated were associated with levels and/or change in BMI during the second half of the life span. The effect estimates remained similar when comparing within twin pairs for four of the sites, indicating the associations are mainly driven by environmental factors. However, the association between cg06192883 and late-life BMI was substantially reduced within twin pairs, pointing to the importance of familial factors in this association.

Acknowledgements

SATSA was supported by National Institutes of Health (NIH; grants AG04563 and AG10175), the MacArthur Foundation Research Network on Successful Aging, the Swedish Research Council for Working Life and Social Research (FAS; Grants 97:0147:1B, 2009-0795), and the Swedish Research Council (825-2007-7460 and 825-2009-6141). This work was supported by the Swedish Research Council Grant (2016-03081) and the Swedish Research Council for Health, Working Life and Welfare (2018-01201).

We acknowledge The Swedish Twin Registry for access to data. The Swedish Twin Registry is managed by Karolinska Institutet and receives funding through the Swedish Research Council under the grant no 2017-00641.

Methylation profiling on the Infinium MethylationEPIC BeadChip was performed by the SNP&SEQ Technology Platform in Uppsala (www.genotyping.se). The facility is part of the National Genomics Infrastructure (NGI) Sweden and Science for Life Laboratory. The SNP&SEQ Platform is also supported by the Swedish Research Council and the Knut and Alice Wallenberg Foundation.

Competing Interests

The authors declare no competing financial interests.

References

1. Dahl AK, Reynolds CA, Fall T, Magnusson PK, Pedersen NL. Multifactorial analysis of changes in body mass index across the adult life course: a study with 65 years of follow-up. *Int. J. Obes. (Lond.)* 2014; **38**(8): 1133-41.
2. Dahl AK, Fauth EB, Ernst-Bravell M, Hassing LB, Ram N, Gerstoft D. Body mass index, change in body mass index, and survival in old and very old persons. *J Am Geriatr Soc* 2013; **61**(4): 512-8.
3. Stokes A, Preston SH. Revealing the burden of obesity using weight histories. *Proc Natl Acad Sci U S A* 2016; **113**(3): 572-7.
4. Min J, Chiu DT, Wang Y. Variation in the heritability of body mass index based on diverse twin studies: a systematic review. *Obes Rev* 2013; **14**(11): 871-82.
5. Winkler TW, Justice AE, Graff M, Barata L, Feitosa MF, Chu S *et al.* The Influence of Age and Sex on Genetic Associations with Adult Body Size and Shape: A Large-Scale Genome-Wide Interaction Study. *PLoS Genet* 2015; **11**(10): e1005378.
6. Song M, Zheng Y, Qi L, Hu FB, Chan AT, Giovannucci EL. Longitudinal Analysis of Genetic Susceptibility and BMI Throughout Adult Life. *Diabetes* 2018; **67**(2): 248-255.
7. Benayoun BA, Pollina EA, Brunet A. Epigenetic regulation of ageing: linking environmental inputs to genomic stability. *Nat Rev Mol Cell Biol* 2015; **16**(10): 593-610.
8. Luo C, Hajkova P, Ecker JR. Dynamic DNA methylation: In the right place at the right time. *Science* 2018; **361**(6409): 1336-1340.
9. Demerath EW, Guan W, Grove ML, Aslibekyan S, Mendelson M, Zhou YH *et al.* Epigenome-wide association study (EWAS) of BMI, BMI change and waist circumference in African American adults identifies multiple replicated loci. *Hum. Mol. Genet.* 2015; **24**(15): 4464-79.
10. Mendelson MM, Marioni RE, Joehanes R, Liu C, Hedman AK, Aslibekyan S *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.* 2017; **14**(1): e1002215.
11. Wahl S, Drong A, Lehne B, Loh M, Scott WR, Kunze S *et al.* Epigenome-wide association study of body mass index, and the adverse outcomes of adiposity. *Nature* 2017; **541**(7635): 81-86.

12. Sayols-Baixeras S, Subirana I, Fernandez-Sanles A, Senti M, Lluís-Ganella C, Marrugat J *et al.* DNA methylation and obesity traits: An epigenome-wide association study. The REGICOR study. *Epigenetics* 2017; **12**(10): 909-916.
13. Campanella G, Gunter MJ, Polidoro S, Krogh V, Palli D, Panico S *et al.* Epigenome-wide association study of adiposity and future risk of obesity-related diseases. *Int. J. Obes. (Lond.)* 2018; **42**(12): 2022-2035.
14. Dhana K, Braun KVE, Nano J, Voortman T, Demerath EW, Guan W *et al.* An Epigenome-Wide Association Study of Obesity-Related Traits. *Am. J. Epidemiol.* 2018; **187**(8): 1662-1669.
15. Bell CG. The Epigenomic Analysis of Human Obesity. *Obesity (Silver Spring)* 2017; **25**(9): 1471-1481.
16. McRae AF, Marioni RE, Shah S, Yang J, Powell JE, Harris SE *et al.* Identification of 55,000 Replicated DNA Methylation QTL. *Sci. Rep.* 2018; **8**(1): 17605.
17. Finkel D, Pedersen N. Processing Speed and Longitudinal Trajectories of Change for Cognitive Abilities: The Swedish Adoption/Twin Study of Aging. *Neuropsychology, development, and cognition. Section B, Aging, neuropsychology and cognition* 2004; **11**(2): 325-345.
18. Magnusson PK, Almqvist C, Rahman I, Ganna A, Viktorin A, Walum H *et al.* The Swedish Twin Registry: Establishment of a Biobank and Other Recent Developments. *Twin Res Hum Genet* 2012; **16**(1): 317-329.
19. Davis S DP, Bilke S, Triche, Jr. T and Bootwalla M. methylumi: Handle Illumina methylation data. R package version 2.18.2. 2015.
20. Pidsley R, CC YW, Volta M, Lunnon K, Mill J, Schalkwyk LC. A data-driven approach to preprocessing Illumina 450K methylation array data. *BMC Genomics* 2013; **14**: 293.
21. Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH *et al.* DNA methylation arrays as surrogate measures of cell mixture distribution. *BMC Bioinformatics* 2012; **13**: 86.
22. Leek JT JW, Parker HS, Fertig EJ, Jaffe AE and Storey JD. sva: Surrogate Variable Analysis. R package version 3.20.0. <https://bioconductor.org/packages/release/bioc/html/sva.html> 2016.
23. Du P, Zhang X, Huang CC, Jafari N, Kibbe WA, Hou L *et al.* Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. *BMC Bioinformatics* 2010; **11**: 587.

24. Sjölander A, Frisell T, Öberg S. Causal Interpretation of Between-Within Models for Twin Research. In: *Epidemiologic Methods*, 2012. p 217.
25. McArdle JJ. Latent variable modeling of differences and changes with longitudinal data. *Annu Rev Psychol* 2009; **60**: 577-605.
26. Falconer DS, Mackay TFC. *Introduction to quantitative genetics*, 4th ed. edn Longman: Harlow, 1996.
27. Wang Y, Karlsson R, Lampa E, Zhang Q, Hedman AK, Almgren M *et al.* Epigenetic influences on aging: a longitudinal genome-wide methylation study in old Swedish twins. *Epigenetics* 2018; **13**(9): 975-987.
28. Gaunt TR, Shihab HA, Hemani G, Min JL, Woodward G, Lyttleton O *et al.* Systematic identification of genetic influences on methylation across the human life course. *Genome Biol* 2016; **17**: 61.
29. Kremntsova EB, Furuta K, Oiwa K, Trybus KM, Ali MY. Small teams of myosin Vc motors coordinate their stepping for efficient cargo transport on actin bundles. *J Biol Chem* 2017; **292**(26): 10998-11008.
30. La Rosa S, Lloyd RV, Capella C. Expression and Role of Myosins in Pancreatic Endocrine Cells and Related Tumors. In: Loft MA (ed) *Trends in Pancreatic Cancer Research*. Nova Publishers, 2005, pp 23-39.
31. Willmer T, Johnson R, Louw J, Pfeiffer C. Blood-Based DNA Methylation Biomarkers for Type 2 Diabetes: Potential for Clinical Applications. *Front Endocrinol (Lausanne)* 2018; **9**: 744.
32. Braun KV, Voortman T, Dhana K, Troup J, Bramer WM, Chowdhury R *et al.* The role of DNA methylation in dyslipidaemia: A systematic review. *Prog Lipid Res* 2016; **64**: 178-191.
33. Li S, Wong EM, Bui M, Nguyen TL, Joo JE, Stone J *et al.* Inference about causation between body mass index and DNA methylation in blood from a twin family study. *Int. J. Obes. (Lond.)* 2019; **43**(2): 243-252.
34. Sherling DH, Perumareddi P, Hennekens CH. Metabolic Syndrome: Clinical and Policy Implications of the New Silent Killer. *Journal of Cardiovascular Pharmacology and Therapeutics* 2017; **22**(4): 365-367.
35. Kelfve S, Fors S, Lennartsson C. Getting better all the time? Selective attrition and compositional changes in longitudinal and life-course studies. *Longitudinal and Life Course Studies* 2017; **8**(1).

36. Li W, Christiansen L, Hjelmborg J, Baumbach J, Tan Q. On the power of epigenome-wide association studies using a disease-discordant twin design. *Bioinformatics* 2018; **34**(23): 4073-4078.
37. Frisell T, Oberg S, Kuja-Halkola R, Sjolander A. Sibling comparison designs: bias from non-shared confounders and measurement error. *Epidemiology* 2012; **23**(5): 713-20.
38. Teschendorff AE, Zheng SC. Cell-type deconvolution in epigenome-wide association studies: a review and recommendations. *Epigenomics* 2017; **9**(5): 757-768.

Figure 1: The association between baseline DNA methylation and longitudinal trajectories of body mass index

Predicted values from latent growth curve models of longitudinal change in body mass index, stratified into tertiles of DNA methylation level at baseline at each of the seven CpG sites. Beta values with 95% CI and p-values are presented for the effect of DNA methylation at baseline on body mass index level at the intercept (age 65), and linear and quadratic change. Models are adjusted for methylation array, sex, and smoking.

Table 1: Association between DNA methylation and body mass index at baseline in the total sample, and between / within twin pairs

	Total_sample			Between-pair effect			Within-pair effect		
	β	95% CI	P-value	β	95% CI	P-value	β	95% CI	P-value
cg06192883	0.72	0.37,1.07	<0.001**	0.83	0.35,1.31	<0.001**	0.40	-0.06,0.86	0.09
cg11024682	0.48	0.15,0.82	0.0047**	0.47	-0.01,0.95	0.053	0.50	0.08,0.93	0.021*
cg00574958	-0.37	-0.72,-0.03	0.035*	-0.44	-0.97,0.10	0.11	-0.24	-0.65,0.18	0.26
cg08857797	0.36	0.01,0.70	0.041*	0.40	-0.10,0.90	0.113	0.40	-0.02,0.83	0.064
cg06946797	-0.07	-0.39,0.25	0.67	-0.34	-0.84,0.15	0.173	0.43	-0.01,0.86	0.055
cg13123009	0.00	-0.36,0.36	1	0.01	-0.47,0.49	0.97	-0.02	-0.49,0.45	0.94
cg09664445	0.22	-0.15,0.59	0.25	0.25	-0.25,0.76	0.33	0.15	-0.26,0.55	0.47

Beta values, 95% confidence intervals, and p-values of the association between DNA methylation at seven CpG sites and body mass index. Linear regression was applied to the total sample, and between-within models to compare estimates between and within twin pairs. All models are adjusted for age, sex, smoking, and methylation array. *Significance at the $\alpha=0.05$ level. ** Significance after conservative Bonferroni correction at $\alpha=0.007$

Table 2: Between-within models of baseline DNA methylation and longitudinal trajectories of body mass index

	Between-pair effect			Within-pair effect		
	β (CpG)	95% CI	P-val	β (CpG)	95% CI	P-val
cg06192883						
BMI level (age 65)	0.88	0.36,1.40	<0.001**	0.40	-0.10,0.91	0.12
Linear change in BMI	-0.63	-0.92,-0.34	<0.001**	-0.28	-0.58,0.02	0.07
Quadratic change in BMI	0.24	0.09,0.38	0.001**	-0.04	-0.23,0.15	0.68
cg11024682						
BMI level (age 65)	0.50	-0.02,1.02	0.06	0.49	0.01,0.96	0.05*
Linear change in BMI	-0.24	-0.52,0.05	0.10	-0.02	-0.34,0.29	0.88
Quadratic change in BMI	0.08	-0.07,0.22	0.30	0.03	-0.15,0.22	0.72
cg00574958						
BMI level (age 65)	-0.52	-1.09,0.06	0.08	-0.37	-0.81,0.07	0.10
Linear change in BMI	0.18	-0.16,0.52	0.30	-0.06	-0.35,0.22	0.66
Quadratic change in BMI	0.06	-0.11,0.23	0.50	0.12	-0.04,0.29	0.15
cg08857797						
BMI level (age 65)	0.43	-0.11,0.98	0.12	0.40	-0.06,0.87	0.09
Linear change in BMI	-0.08	-0.41,0.25	0.63	-0.43	-0.74,-0.13	0.006**
Quadratic change in BMI	0.01	-0.15,0.17	0.89	0.23	0.05,0.41	0.01*
cg06946797						
BMI level (age 65)	-0.47	-1.00,0.06	0.08	0.43	-0.05,0.91	0.08
Linear change in BMI	0.07	-0.24,0.38	0.66	-0.25	-0.60,0.10	0.16
Quadratic change in BMI	0.10	-0.06,0.26	0.23	0.17	-0.01,0.36	0.06
cg13123009						
BMI level (age 65)	-0.12	-0.67,0.42	0.65	-0.05	-0.56,0.47	0.86
Linear change in BMI	-0.03	-0.35,0.29	0.86	-0.14	-0.50,0.21	0.42
Quadratic change in BMI	0.07	-0.07,0.22	0.30	0.13	-0.05,0.32	0.16
cg09664445						
BMI level (age 65)	0.29	-0.26,0.83	0.30	0.25	-0.19,0.69	0.27
Linear change in BMI	-0.08	-0.39,0.23	0.60	-0.14	-0.47,0.19	0.40
Quadratic change in BMI	0.05	-0.10,0.20	0.50	-0.06	-0.24,0.12	0.54

Beta values, 95% confidence intervals, and p-values for the estimates of baseline DNA methylation on BMI at the intercept (age 65), linear change, and quadratic change in BMI. The between-pair estimate represents the average effect in the population, and the within-pair estimate the effect not attributable to shared genetic and other familial factors. Models are adjusted for methylation array, sex, and smoking. *Significance at the $\alpha=0.05$ level. **Significance after conservative Bonferroni correction at $\alpha=0.007$.

Supplementary information

Table S1: Information about each of the seven CpG sites from the Illumina annotation file

IlumID	CHR	MAPINFO	UCSC_RefGene_Name
cg06192883	15	52554171	<i>MYO5C</i>
cg11024682	17	17730094	<i>SREBF1;SREBF1</i>
cg00574958	11	68607622	<i>CPT1A;CPT1A</i>
cg08857797	17	40927699	<i>VPS25</i>
cg06946797	16	11422409	
cg13123009	6	31681882	<i>LY6G6E;LY6G6D;LY6G6E</i>
cg09664445	17	2612406	<i>KIAA0664</i>

Table S2: Number of samples and mean age at each in-person testing occasion

	Individuals with BMI data		Individuals with methylation data	
Testing occasion	N	Mean age	N	Mean age
IPT3 (1992-1994)	328	68.5	328	68.5
IPT4 (1995-1997)	24	70.9	-	-
IPT5 (1999-2002)	386	70.5	336	70.0
IPT6 (2002-2005)	348	71.8	247	71.4
IPT7 (2006-2008)	297	73.9	-	-
IPT8 (2008-2010)	280	75.5	248	75.6
IPT9 (2010-2012)	242	76.5	201	77.3
IPT10 (2012-2014)	224	77.6	42	76.4

Table S3: Association between DNA methylation and body mass index at baseline between and within twin pairs

	Between-pair estimate			Within-pair estimate			Within-pair estimate Dizygotic twin pairs			Within-pair estimate Monozygotic twin pairs		
CpG	β	95% CI	P-val	β	95% CI	P-val	β	95% CI	P-val	β	95% CI	P-val
cg06192883	0.83	0.35,1.31	<0.001**	0.40	-0.06,0.86	0.09	0.71	0.09,1.33	0.026*	-0.18	-0.77,0.42	0.56
cg11024682	0.47	-0.01,0.95	0.053	0.50	0.08,0.93	0.021*	0.55	-0.04,1.15	0.068	0.38	-0.14,0.90	0.15
cg00574958	-0.44	-0.97,0.10	0.11	-0.24	-0.65,0.18	0.26	-0.27	-0.85,0.31	0.36	-0.22	-0.73,0.29	0.39
cg08857797	0.40	-0.10,0.90	0.113	0.40	-0.02,0.83	0.064	0.47	-0.12,1.07	0.12	0.22	-0.31,0.76	0.41
cg06946797	-0.34	-0.84,0.15	0.173	0.43	-0.01,0.86	0.055	0.62	0.05,1.19	0.034*	0.06	-0.55,0.66	0.86
cg13123009	0.01	-0.47,0.49	0.97	-0.02	-0.49,0.45	0.94	-0.06	-0.67,0.56	0.86	0.02	-0.66,0.69	0.96
cg09664445	0.25	-0.25,0.76	0.33	0.15	-0.26,0.55	0.47	0.13	-0.43,0.69	0.64	0.22	-0.29,0.73	0.4

Beta values, 95% confidence intervals, and p-values of the association between DNA methylation at seven different sites and body mass index. Between-within models were applied to compare estimates between and within twin pairs. All models are adjusted for age, sex, smoking, and methylation array.

*Significance at the $\alpha=0.05$ level. ** Significance after strict Bonferroni correction at $\alpha=0.007$

Table S4: Longitudinal trajectories of DNA methylation

Outcome	Random intercept model						Random intercept and slope model						
	Intercept			Slope			Intercept			Slope			P-val
	β	95% CI	P-val	β	95% CI	P-val	β	95% CI	P-val	β	95% CI	P-val	LRT
cg06192883	0.448	0.177,0.719	0.001**	0.005	-0.001,0.011	0.13	0.460	0.191,0.728	<0.001**	0.005	-0.002,0.011	0.15	0.55
cg11024682	0.465	0.185,0.745	0.001**	0.013	0.007,0.019	< 0.001**	0.461	0.179,0.742	0.001**	0.013	0.007,0.019	< 0.001**	0.22
cg00574958	0.104	-0.127,0.334	0.38	0.009	-0.015,-0.003	0.004**	0.108	-0.124,0.339	0.36	0.009	-0.015,-0.003	0.005**	0.90
cg08857797	0.389	0.141,0.638	0.002**	0.003	-0.009,0.004	0.42	0.389	0.141,0.638	0.002**	0.002	-0.009,0.004	0.50	0.15
cg06946797	0.392	0.118,0.665	0.005**	0.005	-0.012,0.001	0.08	0.388	0.113,0.663	0.006**	0.006	-0.012,0.001	0.08	0.61
cg13123009	0.546	0.283,0.809	<0.001**	0.000	-0.006,0.007	0.889	0.553	0.290,0.816	<0.001**	0.000	-0.006,0.006	0.98	0.41
cg09664445	0.180	-0.067,0.426	0.15	0.013	-0.020,-0.007	<0.001**	0.164	-0.082,0.410	0.19	0.014	-0.020,-0.007	<0.001**	0.24

Beta values, 95% confidence intervals, and p-values from latent growth curve of longitudinal change in methylation levels at the seven CpG sites. All models are adjusted for age, sex, smoking, and methylation array. *Significance at the $\alpha=0.05$ level. ** Significance after strict Bonferroni correction at $\alpha=0.007$

Table S5: Association between DNA methylation, body mass index, and potential mediators at baseline

CpG and mediator	a) CpG --> Mediator β (95% CI)	b) CpG --> BMI, adjusted for mediator β (95% CI)	c) P-value, mediation
<u>cg06192883</u>			
No mediator		0.72 (0.37,1.07)	
Type 2 diabetes	0.08 (-0.11,0.27) ^a	0.72 (0.37,1.07)	0.93
Blood glucose	0.21 (0.06,0.36)	0.68 (0.32,1.03)	0.006
Total cholesterol	0.01 (-0.09,0.12)	0.69 (0.33,1.05)	0.78
Triglycerides	0.07 (0.03,0.12)	0.56 (0.20,0.91)	<0.001
<u>cg11024682</u>			
No mediator		0.48 (0.15,0.82)	
Type 2 diabetes	0.06 (-0.12,0.24) ^a	0.48 (0.14,0.82)	0.96
Blood glucose	0.22 (0.07,0.36)	0.41 (0.07,0.76)	0.01
Total cholesterol	-0.03 (-0.13,0.07)	0.56 (0.21,0.90)	0.54
Triglycerides	0.07 (0.02,0.11)	0.44 (0.11,0.78)	0.006
<u>cg00574958</u>			
No mediator		-0.37 (-0.73,-0.02)	
Type 2 diabetes	-0.10 (-0.29,0.09) ^a	-0.37 (-0.73,-0.01)	0.90
Blood glucose	-0.10 (-0.25,0.05)	-0.35 (-0.71,0.01)	0.15
Total cholesterol	-0.04 (-0.14,0.06)	-0.39 (-0.76,-0.03)	0.43
Triglycerides	-0.09 (-0.13,-0.04)	-0.22 (-0.58,0.14)	0.002
<u>cg08857797</u>			
No mediator		0.36 (0.01,0.71)	
Type 2 diabetes	0.02 (-0.17,0.20) ^a	0.36 (0.01,0.70)	0.94
Blood glucose	0.09 (-0.06,0.24)	0.35 (-0.01,0.70)	0.22
Total cholesterol	-0.00 (-0.10,0.10)	0.32 (-0.04,0.67)	0.97
Triglycerides	0.02 (-0.02,0.07)	0.28 (-0.07,0.62)	0.40

Beta values and 95% confidence intervals from linear regression models of DNA methylation at four CpG sites as predictors of a) potential mediators, and b) BMI, with and without adjusting for potential mediators. The p-value for the mediating effect of each potential mediator was obtained through

bootstrapping with 1000 repetitions. All models were adjusted for age, sex, smoking, and methylation array. Models including blood glucose, total cholesterol, or triglycerides were further adjusted for fasting at time of blood sample.

^a The association between DNA methylation and type 2 diabetes was modelled with logistic regression.

Table S6: Comparison of twin-pair correlation between monozygotic and dizygotic twin-pairs

CpG	rMZ (95% CI)	rDZ (95% CI)	Falconer heritability
cg06192883	0.34 (0.14-0.52)	0.24 (0.09-0.38)	0.21
cg11024682*	0.18 (-0.03-0.39)	0.32 (0.17-0.45)	--
cg00574958	0.11 (-0.11-0.32)	0.11 (-0.05-0.26)	0.00
cg08857797	0.18 (-0.04-0.38)	0.10 (-0.06-0.25)	0.16
cg06946797**	0.38 (0.17-0.55)	0.10 (-0.05-0.26)	0.54
cg13123009	0.36 (0.16-0.54)	0.26 (0.10-0.40)	0.21
cg09664445	0.16 (-0.06-0.37)	0.10 (-0.06-0.25)	0.13

Twin-pair correlations and 95% confidence intervals at each CpG site for monozygotic (rMZ) and dizygotic (rDZ) twin pairs. The Falconer heritability is a measure of the unadjusted broad-sense heritability, and was calculated using the Falconer formula: $2 \times (rMZ - rDZ)(1)$.

* Heritability could not be calculated due to rDZ being higher than rMZ

** The low rDZ and high rMZ indicates presence of non-additive/dominance genetic effects