Supporting Information

Horvath et al. 10.1073/pnas.1412759111

SI Text

Detailed Description of Datasets. ***Dataset 1.*** Liver data come from Gene Expression Omnibus (GEO) series GSE48325 and were described in Ahrens et al. (1). Bisulphite converted DNA from these samples were hybridized to the Illumina Infinium 450K Human Methylation Beadchip. In Fig. 1*A*, we only considered the 62 samples taken before bariatric surgery. The clinical variables are explained a separate section.
Dataset 2. Subcutaneous adipose tissue data from Grundberg et al. (2). The Illumina 450K data come from the Multiple Tissue Human Expression Resource (MuTHER) and were generated from adipose tissue data from 648 twins. Array Express ID E-MTAB1866.
Dataset 3. Skeletal muscle data from GSE50498 (3).
Dataset 4. Blood Illumina 27K from GSE49909. Blood data (buffy coat) from Day et al. (4).
Dataset 5. Peripheral blood mononuclear cells from GSE37008 (5). Using these data, we also found that age acceleration (AA, measured using DNAm age) is independent of blood cell counts.
Dataset 6. Whole blood data from GSE53840.
Dataset 7. Novel liver tissue Illumina 450K dataset (analogous to dataset 1). These data are publicly available from GEO superseries GSE61256.
Dataset 8. Novel adipose tissue measured on the Illumina 450K array. These data are publicly available from GEO superseries GSE61256.
Dataset 9. Novel muscle tissue Illumina 450K data. These data are publicly available from GEO superseries GSE61256.

Clinical Variables in the Liver Data. The following variables relate to datasets 1, 7, 8, and 9.
BMI. The body mass index (BMI) is a measure of relative weight based on an individual's mass and height. It is defined as the weight (in kilograms) divided by the square of the height (in meters). A value between 18.5 and 25 is considered as normal. Moderately obese people have a BMI between 30 and 35, severely obese between 35 and 40, and very severely obese over 40.
Disease status. This variable indicates the disease status. It takes on the following values: NAFLD (nonalcoholic fatty liver disease), NASH (nonalcoholic steatohepatitis), healthy obese, control (for healthy control subjects), primary biliary cirrhosis), and primary sclerosing cholangitis.
Total NAFLD activity score. The total NAFLD Activity Score (NAS) represents the sum of scores for steatosis, lobular inflammation, and hepatocyte ballooning, and ranges from 0 to 8. After diagnosis, NASH or fatty liver not diagnostic of NASH, the total NAS is used to grade activity. NAS scores of 0–2 typically occur in cases largely considered not diagnostic of NASH, whereas scores of 5–8 usually occurs in cases that are considered diagnostic of NASH (6).
Steatosis. Ordinal variable that relates to the amount of surface area involved by steatosis as evaluated on medium power examination. Minimal steatosis (<5%) receives a score of 0. 5–33% (score of 1), 33–66% (score 2), and >66% (score 3).
Liver inflammation. Ordinal variable: 0 corresponds to no foci, 1 (<2 foci/200×), 2 (2–4 foci/200×), 3 (>4 foci/200×).
Fibrosis. Ordinal variable that takes on (half) integer values between 0 and 4: 0 (none), 2 (perisinusoidal and portal/periportal), 3 (bridging fibrosis), 4 (cirrhosis). The fibrosis stage is evaluated separately from the total NAFLD score.
Hepatocye ballooning. This ordinal variable measures hepatocyte ballooning: 0 (none), 1 (few balloon cells). Here "few" means rare but definite ballooned hepatocytes as well as cases that are diagnostically borderline, 2 (many cells/prominent ballooning).

Patient Samples. Datasets 1, 7, 8, and 9 involved only Caucasians collected in Germany; that is, our results regarding the relationship between BMI and age acceleration are not confounded by ethnicity. Standardized histological scoring by a single pathologist were applied to our liver datasets 1 and 7.

Liver samples were obtained percutaneously for patients undergoing liver biopsy for suspected NAFLD or intraoperatively for assessment of liver histology. Normal control samples were recruited from samples obtained for exclusion of liver malignancy during major oncological surgery. None of the normal control individuals underwent preoperative chemotherapy, and liver histology demonstrated absence of both cirrhosis and malignancy. Consenting patients underwent a routine liver biopsy during bariatric surgery for assessment of liver affection. Biopsies were immediately frozen in liquid nitrogen, ensuring an ex vivo time of less than 40 s in all cases. A percutaneous follow-up biopsy was obtained in consenting bariatric patients 5–9 mo after surgery. Patients with evidence of viral hepatitis, hemochromatosis, or alcohol consumption greater than 20 g/d for women and 30 g/d for men were excluded. All patients provided written, informed consent. The study protocol was approved by the institutional review board ("Ethics commission of the Medical Faculty, University of Kiel" D425/07, A111/99) before the commencement of the study. Standardized histopathological assessment (6) was performed by a single pathologist blinded to the molecular analyses. Of the 23 postbariatric patients, 20 underwent a sleeve gastrectomy and 3 a gastric bypass procedure.

Evaluation of Causal Models Using Structural Equation Models. In the following, we discuss several causal models that could explain the relationship between BMI, epigenetic age acceleration, and liver traits (NAS, steatosis, fat percentage, which were significantly correlated with age acceleration) (Fig. S7). As in Figs. 1 and 2, age acceleration was defined as residual resulting from regressing DNAm age on chronological age.

As indicated in the last column of Table S2, model 3: (Traits←BMI→AA) fits the data well for NAS, hepatocye ballooning, liver inflammation, and fibrosis.

Model 3 is the independence model, in which high BMI leads to liver traits (e.g., NAS) and age acceleration independently. In other words, in this model BMI confounds the relationship between NAS and age acceleration.

We caution the reader that structural equation models make several assumptions (reviewed in refs. 7–12), which may not be satisfied in our data. We used the sem R package (10) to evaluate the fit of causal models. The sem R function provides the model fitting χ^2 statistic which was used to compute a model fitting P value for each causal model. For example, P(data|BMI→Steatosis→AA) denotes the P value for the model in which BMI leads to steatosis, which in turn leads to epigenetic age acceleration. The model fitting χ^2 statistic tests the null hypothesis that the model is correct. A small model P value (say $P < 0.05$) indicates that the causal model does not fit well. Following the logic of an "accept-support" context (12, 13), where the null hypothesis represents the researcher's belief, it is the failure to reject the null hypothesis that supports the causal model. The model fit statistic and the corresponding model

P value have several limitations; for example, they are sensitive to the size of correlations and they depend on the sample size *n* (12). However, the model fitting *P* value is the key ingredient of most, if not all, alternative model-fitting indices.

1. Ahrens M, et al. (2013) DNA methylation analysis in nonalcoholic fatty liver disease suggests distinct disease-specific and remodeling signatures after bariatric surgery. *Cell Metab* 18(2):296–302.
2. Grundberg E, et al.; Multiple Tissue Human Expression Resource Consortium (2013) 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(5):876–890.
3. Zykovich A, et al. (2014) Genome-wide DNA methylation changes with age in disease-free human skeletal muscle. *Aging Cell* 13(2):360–366.
4. Day K, et al. (2013) Differential DNA methylation with age displays both common and dynamic features across human tissues that are influenced by CpG landscape. *Genome Biol* 14(9):R102.
5. Lam LL, et al. (2012) Factors underlying variable DNA methylation in a human community cohort. *Proc Natl Acad Sci USA* 109(Suppl 2):17253–17260.
6. Kleiner DE, et al.; Nonalcoholic Steatohepatitis Clinical Research Network (2005) Design and validation of a histological scoring system for nonalcoholic fatty liver disease. *Hepatology* 41(6):1313–1321.
7. Shipley B (2000) *Cause and Correlation in Biology* (Cambridge University Press, Cambridge, United Kingdom), 2 Ed.
8. Pearl J (1988) *Probabilistic reasoning in intelligent systems: networks of plausible inference* (Morgan Kaufmann Publishers, Inc., San Francisco), 2 Ed.
9. Fox J (1984) Linear structural-equation models. *Linear Statistical Models and Related Methods* (Wiley, New York), Chap 4.
10. Fox J (2006) Structural equation modeling with the sem package in R. *Structural Equation Modeling* 13:465–486.
11. Aten J, Fuller T, Lusis A, Horvath S (2008) Using genetic markers to orient the edges in quantitative trait networks: The NEO software. *BMC Syst Biol* 2(1):34.
12. Kline R (2005) *Principles and Practice of Structural Equation Modeling* (The Guilford Press, New York).
13. Steiger JH, Fouladi RT (1997) Noncentrality interval estimation and the evaluation of statistical models. *What If There Were No Significance Tests?*, eds LL Harlow, SA Mulaik, JH Steiger (Lawrence Erlbaum Associates, Inc., Mahwah, NJ), pp 221–257.

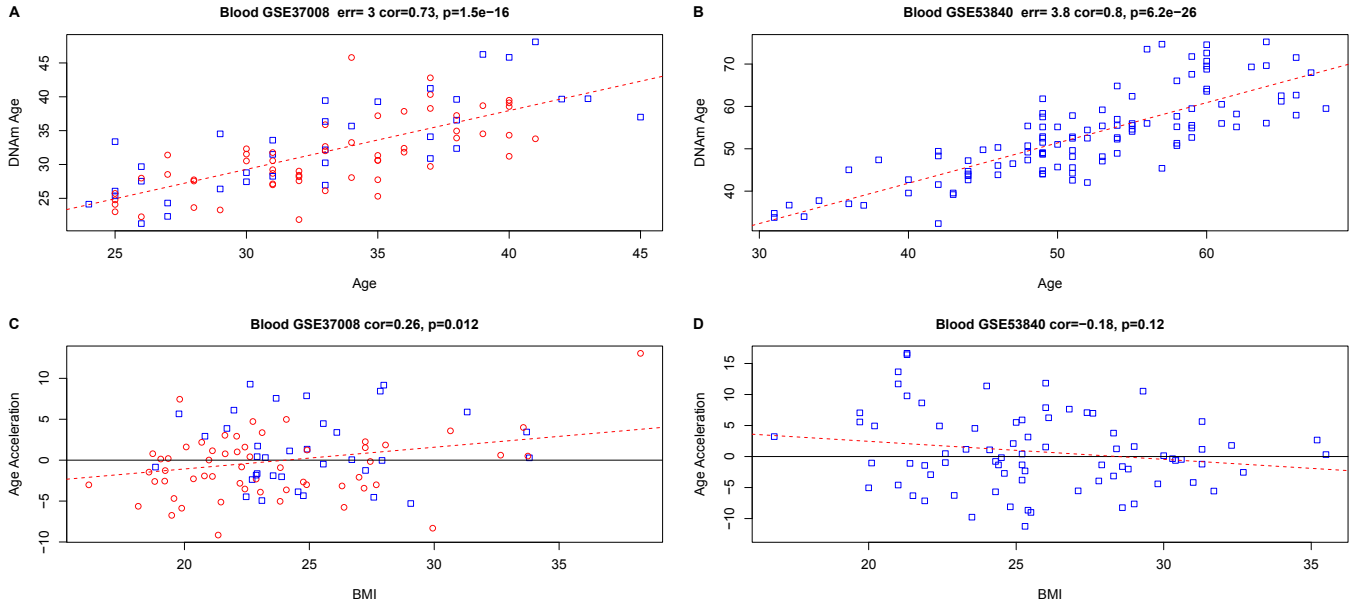

Fig. S1. Testing for age acceleration effects in other blood datasets. (*A* and *B*) DNAm age (*y* axis) versus chronological age in blood datasets 5 and 6, respectively. (*C* and *D*) Age acceleration (*y* axis) versus BMI in the respective datasets. Points are colored by sex (male samples correspond to blue squares). The dashed line corresponds to a linear regression line.

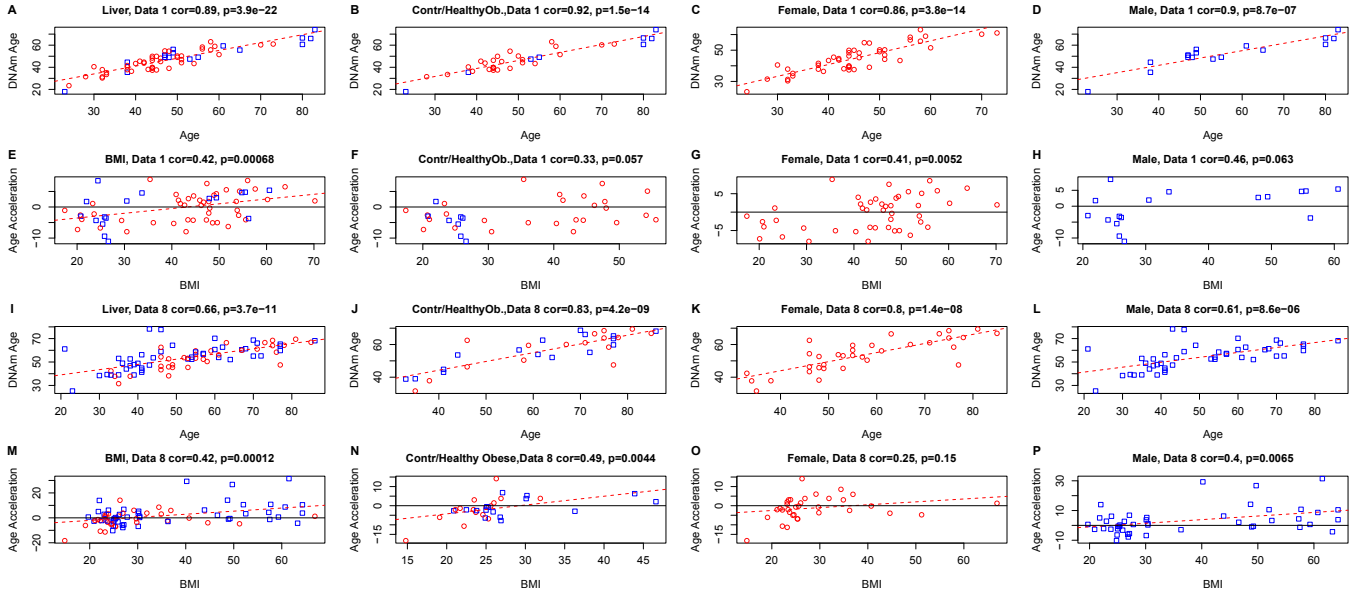

Fig. S2. Subgroup-analysis in the first and second liver datasets. Panels in the first row (*A*–*D*): DNAm age (*y* axis) versus chronological age in (*A*) all liver samples, (*B*) controls/healthy obese subjects, (*C*) female subjects only, and (*D*) male subjects from liver dataset 1. Panels in the second row (*E*–*H*): Age acceleration (*y* axis) versus BMI in (*E*) all samples from liver dataset 1, (*F*) controls/healthy obese subjects, (*G*) female subjects only, and (*H*) male subjects from liver dataset 1. Panels in the third row (*I*–*L*) and fourth row (*M*–*P*) are analogous to those in the first and second rows but involve samples from the second liver dataset (dataset 8 in Table 1). Points are colored by sex (male samples correspond to blue squares). The dashed line corresponds to a linear regression line. The solid black line corresponds to an age acceleration of zero.

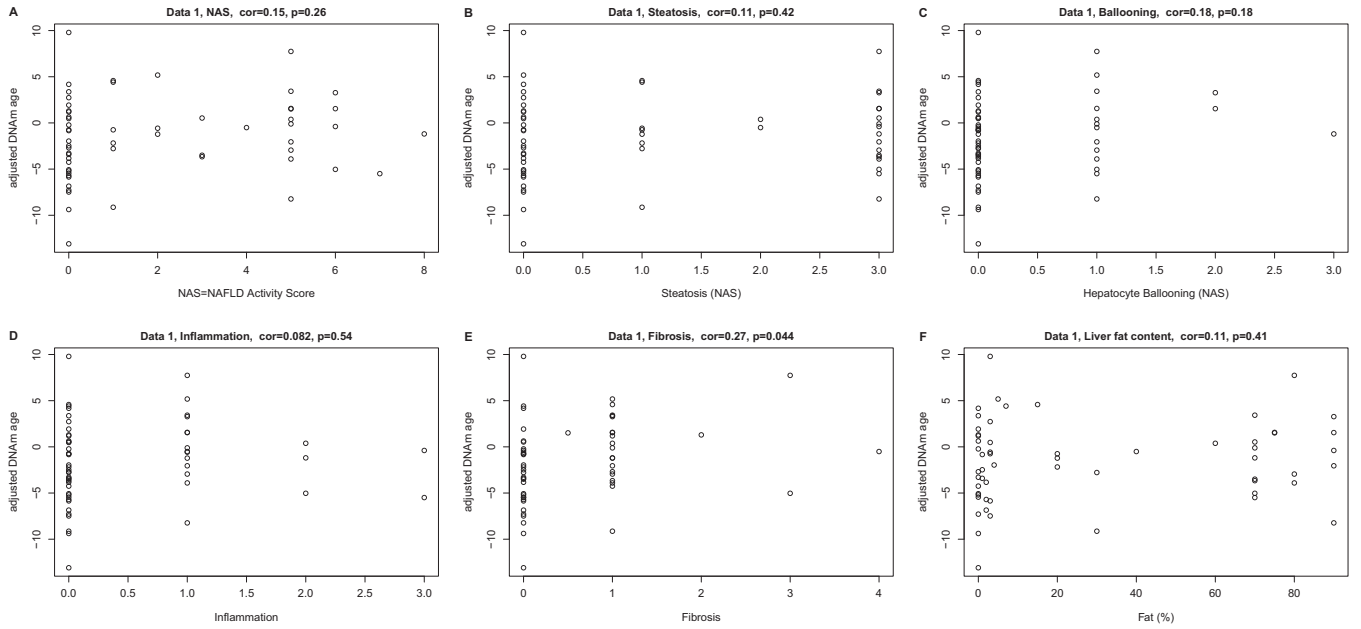

Fig. S3. Post hoc analyses of parameters of liver histology in the first liver dataset. An adjusted measure of DNAm age (*y* axis) is related to various measures of liver pathology. The adjusted measure of DNAm age acceleration was defined as residual from a regression model that regressed DNAm age on chronological age+BMI+sex. In the first liver dataset, this measure of age acceleration has a marginally significant correlation with fibrosis (*E*) but not with the (*A*) NAS, (*B*) steatosis, (*C*) hepatocyte ballooning, (*D*) inflammation, or (*F*) fat percentage.

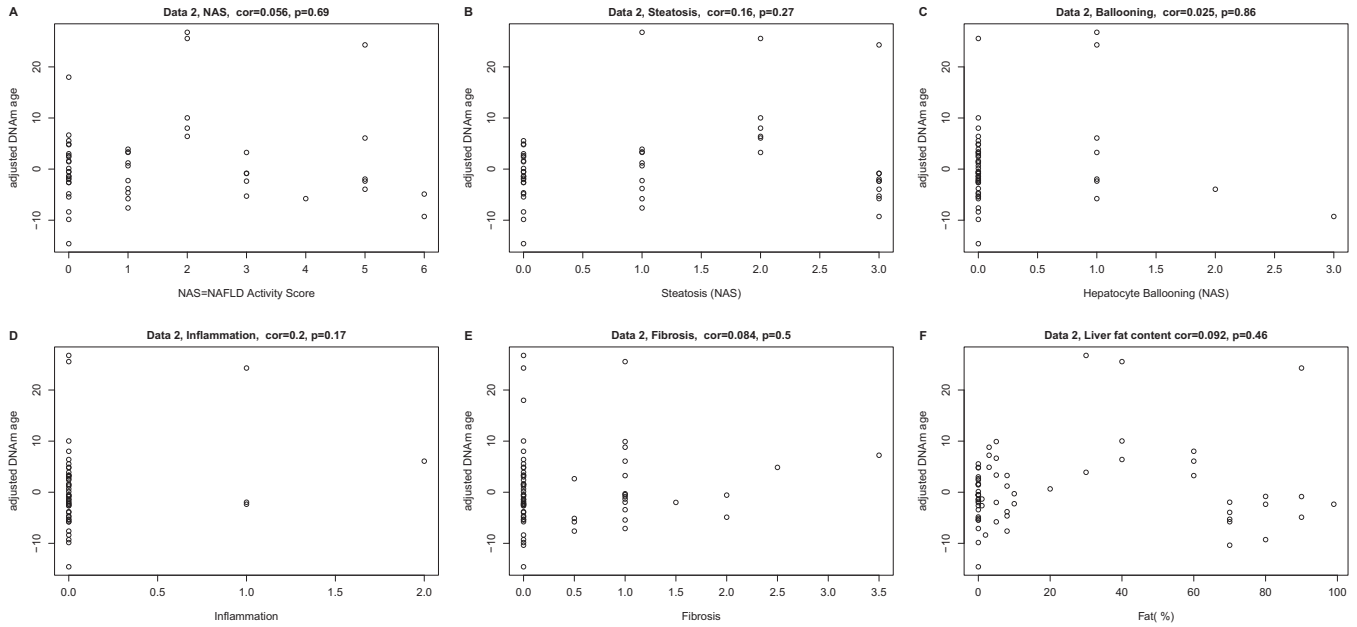

Fig. S4. Adjusted DNAm age versus liver histology in the replication data set. Post hoc analyses of parameters of liver histology in the second liver dataset (dataset 8 in Table 1). An adjusted measure of DNAm age (*y* axis) is related to various measures of liver pathology. The adjusted measure of DNAm age acceleration was defined as residual from a regression model that regressed DNAm age on chronological age+BMI+sex. In the second liver dataset, this measure of age acceleration does not have a significant correlation with (*A*) NAS, (*B*) steatosis, (*C*) hepatocyte ballooning, (*D*) inflammation, (*E*) fibrosis, and (*F*) fat percentage.

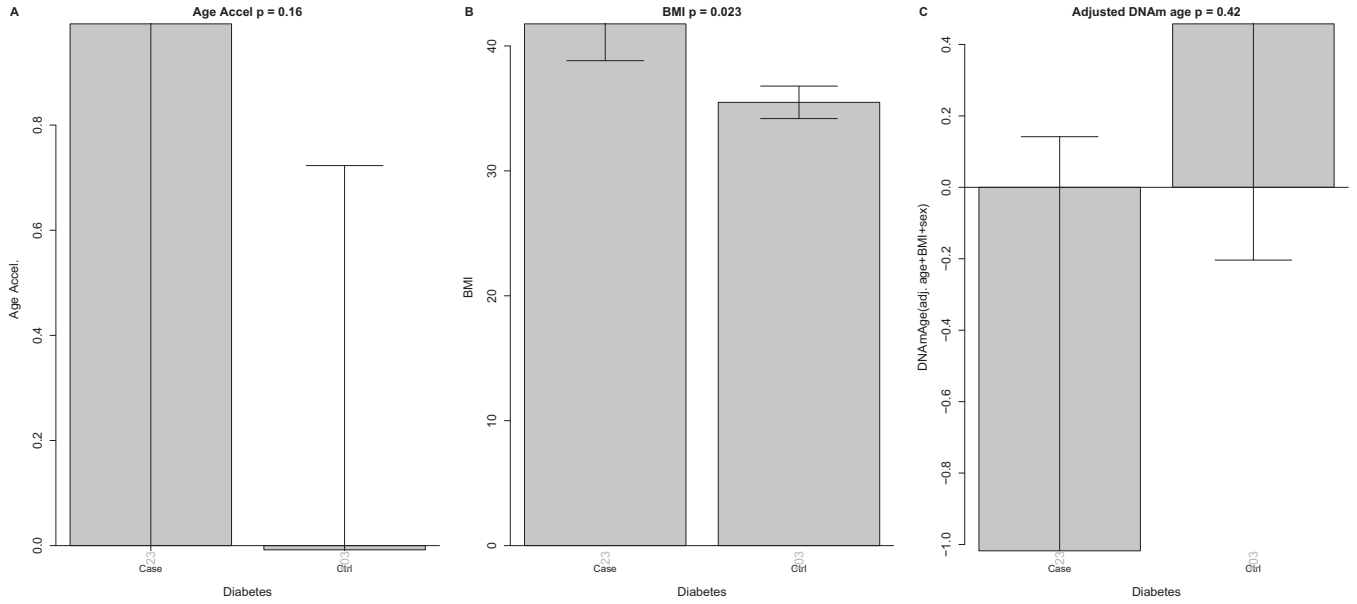

Fig. S5. Age acceleration by diabetes status across both liver datasets. Because of the low number ($n = 23$) of diabetes cases, we combined both liver datasets in this analysis. Diabetes status (*x* axis) has a marginally significant relationship with BMI ($P = 0.023$) (*B*). However, diabetes status does not have a significant relationship with neither (*A*) the age acceleration measure used in Figs. 1*E* or 2*G* nor (*C*) the adjusted measure of DNAm age (defined as raw residual from a linear regression model that regressed DNAm age on chronological age, BMI, and sex). Each bar plot reports the mean value and 1 SE.

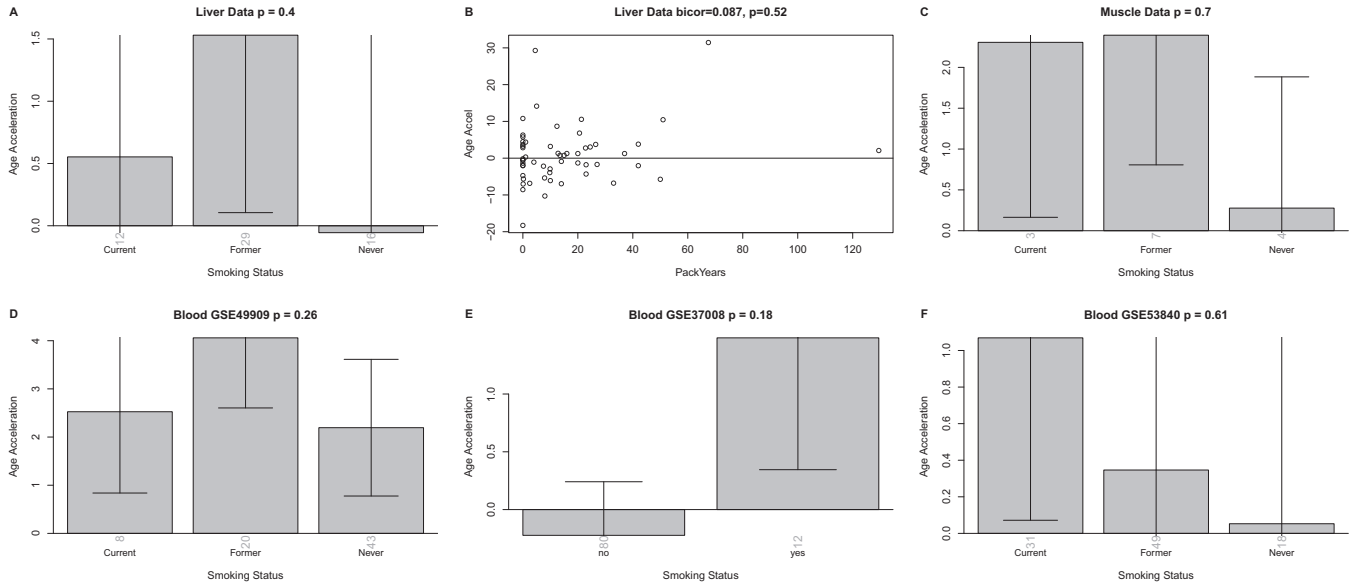

Fig. S6. Age acceleration versus smoking status. Here age acceleration is defined as the residual of a regression model where DNAm age is regressed on chronological age (analogous to our definition in Figs. 1 and 2). In the combined liver dataset, age acceleration is not related to smoking status (*A*) or smoking pack years (*B*). Similarly, there is no significant relationship between age acceleration and smoking status in (*C*) muscle dataset 10 or blood datasets 4 (*D*), 5 (*E*), and 6 (*F*).

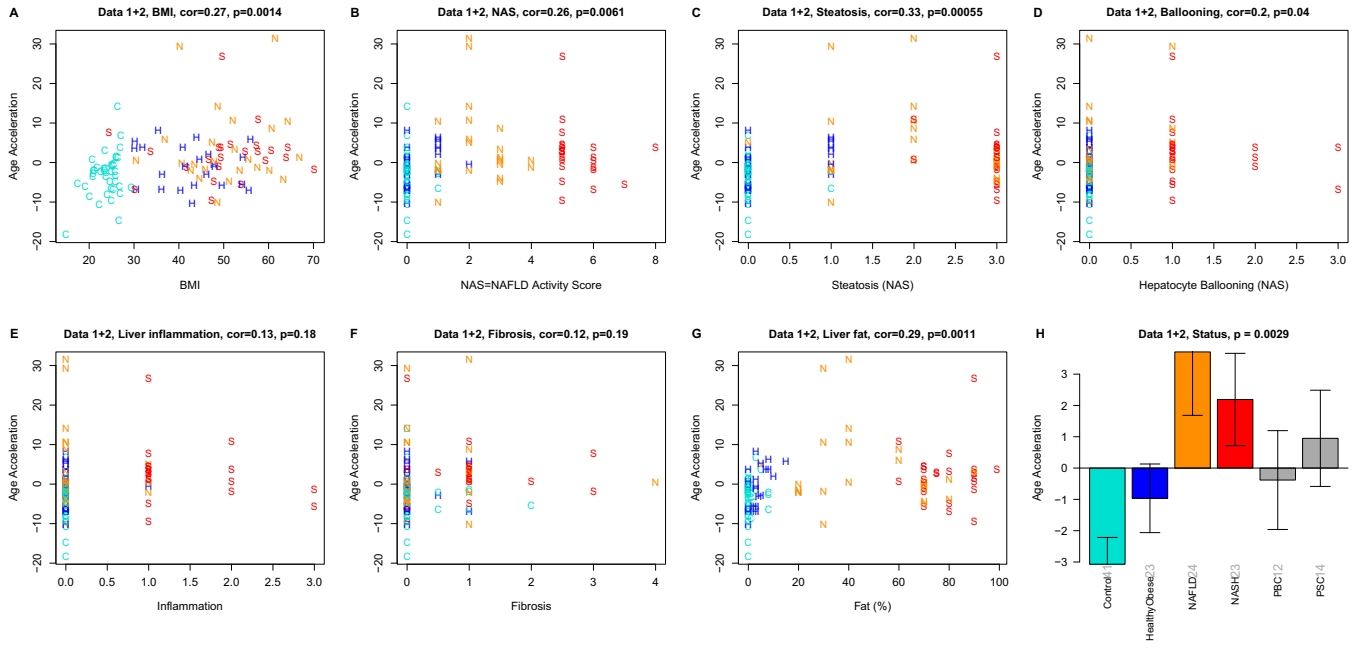

Fig. S7. Age acceleration versus liver traits and disease status. Here age acceleration is defined as the residual of a regression model where DNAm age is regressed on chronological age in both datasets (analogous to our definition in Figs. 1 and 2). In contrast to Fig. 3, this measure of age acceleration is not adjusted for BMI or sex. Age acceleration versus (*A*) BMI, (*B*) NAS, (*C*) steatosis, (*D*) hepatocyte ballooning, (*E*) liver inflammation, (*F*) fibrosis, and (*G*) fat content. The points in the scatter plots are colored and labeled by disease status (as indicated in *H*). Our structural equation model analysis suggests that most of the significant correlations probably reflect confounding due to BMI. (*H*) Age acceleration versus disease status. The bar plots show the mean age acceleration and one SE. As expected, the *P* value from a nonparametric group comparison test (Kruskal–Wallis test) is highly significant.

Table S1. Multivariate regression model

	Combined dataset	
Variable	Estimate (SE)	*P*
Chronological age	0.64645 (0.04927)	<2E-16
BMI	0.23373 (0.06836)	0.000927
NAS, steatosis	−0.70172 (2.29728)	0.760687
NAS, ballooning	−0.52391 (1.48895)	0.725720
NAS, inflammation	−0.46001 (1.38150)	0.739884
Fibrosis	0.96076 (1.07804)	0.375072
Fat percentage	0.03844 (0.08641)	0.657464
Sex (female)	−5.10370 (1.40674)	0.000462
R^2	0.68	
Age accelerated 10-point increase in BMI	3.6 y	

DNAm age was regressed on age, BMI, and various other covariates in the combined liver dataset. Note that BMI remains a highly significant covariate even after adjusting for NAS component traits. According to this multivariate model, age acceleration associated with a 10-point increase in BMI equals $10 \times 0.23373/0.64645 = 3.6$ y.

Table S2. Results of the structural equation model analysis

Model No.	Trait	Causal Model	*P* value (model chisq)	Model fits
M1	Steatosis	BMI→AA→Trait	2.2e-16*	No
M1	Liver fat	BMI→AA→Trait	1.1e-16*	No
M1	NAS	BMI→AA→Trait	4.4e-13*	No
M1	hepatocyte ballooning	BMI→AA→Trait	0.00016*	No
M1	Liver inflammation	BMI→AA→Trait	6.8e-06*	No
M1	Fibrosis	BMI→AA→Trait	0.53†	Yes
M2	Steatosis	BMI→Trait→AA	0.35†	Yes
M2	Liver fat	BMI→Trait→AA	0.18†	Yes
M2	NAS	BMI→Trait→AA	0.077	Weakly
M2	hepatocyte ballooning	BMI→Trait→AA	0.0093*	No
M2	Liver inflammation	BMI→Trait→AA	0.0042*	No
M2	Fibrosis	BMI→Trait→AA	0.0016*	No
M3	Steatosis	Trait←BMI→AA	0.016*	No
M3	Liver fat	Trait←BMI→AA	0.081	Weakly
M3	NAS	Trait←BMI→AA	0.14†	Yes
M3	hepatocyte ballooning	Trait←BMI→AA	0.18†	Weakly
M3	Liver inflammation	Trait←BMI→AA	0.74†	Yes
M3	Fibrosis	Trait←BMI→AA	0.26†	Yes
M4	Steatosis	BMI→Trait←AA	0*	No
M4	Liver fat	BMI→Trait←AA	0*	No
M4	NAS	BMI→Trait←AA	1.1e-14*	No
M4	hepatocyte ballooning	BMI→Trait←AA	2.2e-05*	No
M4	Liver inflammation	BMI→Trait←AA	2.1e-06*	No
M4	Fibrosis	BMI→Trait←AA	0.33†	Yes
M5	Steatosis	BMI→AA←Trait	0.0012*	No
M5	Liver fat	BMI→AA←Trait	0.0012*	No
M5	NAS	BMI→AA←Trait	0.0012*	No
M5	hepatocyte ballooning	BMI→AA←Trait	0.0012*	No
M5	Liver inflammation	BMI→AA←Trait	0.0012*	No
M5	Fibrosis	BMI→AA←Trait	0.0012*	No

The table reports model fitting *P* values based on an SEM analysis. Here age acceleration is again defined as residual from regressing DNAm age on chronological age in both liver datasets. The *P* values correspond to a likelihood ratio test of the null hypothesis that the data fit the causal model. Model fitting *P* values that are significant at a 0.05 threshold (marked by an asterisk) correspond to causal models that do not fit the data. In contrast, model fitting *P* values larger than 0.15 (marked by a dagger) correspond to models that fit the data.

Dataset S1. Transcriptional data correlated with age acceleration in liver

Dataset S1

Robust correlation test results between genes and age acceleration based on the R function *standardScreeningNumericTrait* in the WGCNA R package (1).

1. Langfelder P, Horvath S (2008) WGCNA: An R package for weighted correlation network analysis. *BMC Bioinformatics* 9(1):559.