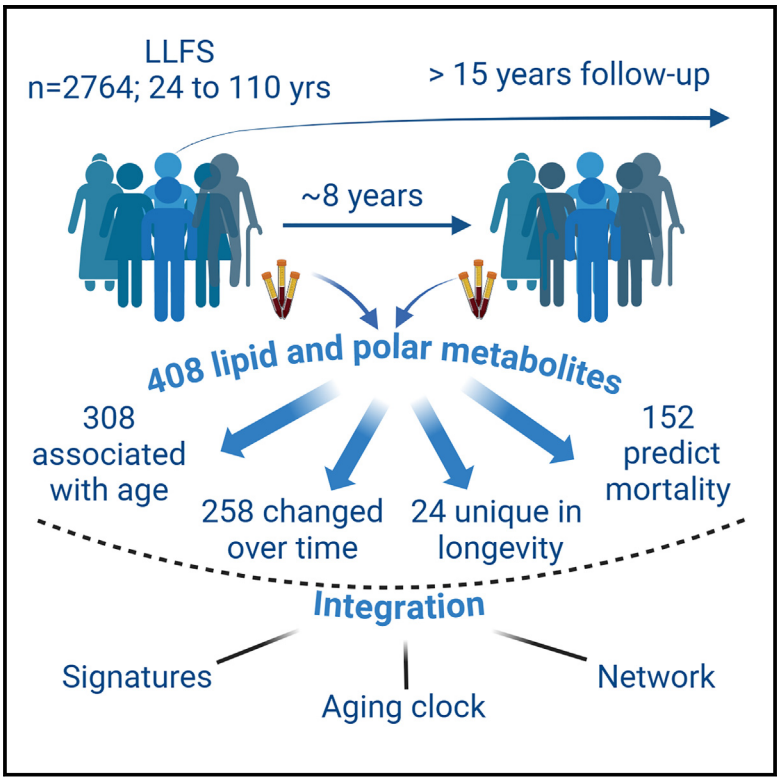
Cell Reports

Article

Metabolite signatures of chronological age, aging, survival, and longevity

Graphical abstract



Authors

Paola Sebastiani, Stefano Monti, Michael S. Lustgarten, ..., Noa Rappaport, Thomas T. Perls, Gary J. Patti

Correspondence

psebastiani@tuftsmedicalcenter.org

In brief

Metabolism plays a crucial role in both aging and aging-related diseases, but the metabolites that serve as pivotal biomarkers of aging are not fully known. Sebastiani et al. analyze metabolites in plasma of Long Life Family Study participants to characterize metabolic markers of chronological age, aging, extreme longevity, and mortality.

Highlights

- Aging correlates with small changes of molecules produced by many metabolic processes
- Metabolites mark compensatory aging mechanisms, cumulative damage, and extreme longevity
- Aging-associated metabolites point to nutrition as source of intervention for healthy aging
- Metabolic profiles of centenarians' plasma highlights APOE2 role in achieving longevity

Sebastiani et al., 2024, Cell Reports *43*, 114913
November 26, 2024
https://doi.org/10.1016/j.celrep.2024.114913

CellPress

Cell Reports

CellPress
OPEN ACCESS

Article

Metabolite signatures of chronological age, aging, survival, and longevity

Paola Sebastiani,[1,2,17,*] Stefano Monti,[3,4] Michael S. Lustgarten,[5] Zeyuan Song,[1] Dylan Ellis,[6] Qu Tian,[7] Michaela Schwaiger-Haber,[8] Ethan Stancliffe,[8] Anastasia Leshchyk,[4] Meghan I. Short,[1,2] Andres V. Ardisson Korat,[5] Anastasia Gurinovich,[1,2] Tanya Karagiannis,[1,2] Mengze Li,[4] Hannah J. Lords,[4] Qingyan Xiang,[1] Megan M. Marron,[9] Harold Bae,[10] Mary F. Feitosa,[11] Mary K. Wojczynski,[11] Jeffrey R. O'Connell,[12] May E. Montasser,[12] Nicole Schupf,[13] Konstantin Arbeev,[14] Anatoliy Yashin,[14] Nicholas Schork,[15] Kaare Christensen,[16] Stacy L. Andersen,[3] Luigi Ferrucci,[7] Noa Rappaport,[6] Thomas T. Perls,[3] and Gary J. Patti[8]

[1]Institute for Clinical Research and Health Policy Studies, Tufts Medical Center, Boston, MA 02111, USA
[2]Department of Medicine, School of Medicine, Tufts University, Boston, MA 02111, USA
[3]Department of Medicine, Chobanian & Avedisian School of Medicine, Boston University, Boston, MA 02118, USA
[4]Bioinformatics Program, Boston University, Boston, MA 02215, USA
[5]Jean Mayer USDA Human Nutrition Research Center on Aging, Tufts University, Boston, MA 02111, USA
[6]Institute for Systems Biology, Seattle, WA 98109, USA
[7]Longitudinal Studies Section, Translational Gerontology Branch, National Institute on Aging, Baltimore, MD 21224, USA
[8]Department of Chemistry, Washington University in St. Louis, St. Louis, MO 63130, USA
[9]Department of Epidemiology, University of Pittsburgh, Pittsburgh, PA 15213, USA
[10]Biostatistics Program, College of Health, Oregon State University, Corvallis, OR 97331, USA
[11]Department of Genetics, Washington University School of Medicine, St. Louis, MO 63130, USA
[12]Department of Medicine, University of Maryland School of Medicine, Baltimore, MD 21201, USA
[13]Department of Epidemiology, Columbia University Medical Center, New York, NY 10032, USA
[14]Social Science Research Institute, Duke University, Durham, NC 27708, USA
[15]The Translational Genomics Research Institute, Phoenix, AZ 85004, USA
[16]Danish Aging Research Center, University of Southern Denmark, 5000 Odense, Denmark
[17]Lead contact
*Correspondence: psebastiani@tuftsmedicalcenter.org
https://doi.org/10.1016/j.celrep.2024.114913

SUMMARY

Metabolites that mark aging are not fully known. We analyze 408 plasma metabolites in Long Life Family Study participants to characterize markers of age, aging, extreme longevity, and mortality. We identify 308 metabolites associated with age, 258 metabolites that change over time, 230 metabolites associated with extreme longevity, and 152 metabolites associated with mortality risk. We replicate many associations in independent studies. By summarizing the results into 19 signatures, we differentiate between metabolites that may mark aging-associated compensatory mechanisms from metabolites that mark cumulative damage of aging and from metabolites that characterize extreme longevity. We generate and validate a metabolomic clock that predicts biological age. Network analysis of the age-associated metabolites reveals a critical role of essential fatty acids to connect lipids with other metabolic processes. These results characterize many metabolites involved in aging and point to nutrition as a source of intervention for healthy aging therapeutics.

INTRODUCTION

Metabolism plays a crucial role in both aging and aging-related disease processes.[1,2] Advances in metabolomics have yielded critical insights,[3–11] but despite the increasing number of studies, identification and characterization of the metabolites that serve as pivotal biomarkers of aging are far from complete.[1] Furthermore, there are considerable gaps in differentiating between metabolites of age that are measured cross-sectionally, metabolites of aging that are measured longitudinally,[3,12] metabolites of human extreme longevity (EL) that are measured in very long-lived individuals,[13] and metabolites that correlate with mortality risk.[7,8,14] Better clarification of metabolites' roles in these processes could help the discovery of aging biomarkers and therapeutic targets.

The Long Life Family Study (LLFS) enrolled approximately 5,000 individuals aged 18–110 years from long-lived families.[15,16] The LLFS collected blood samples from study participants at time of enrollment and approximately 8 years later and generated longitudinal plasma metabolomic profiles. The wide age range of the study participants, the availability of longitudinal phenotypic and metabolomic data, and the large number

Table 1. Metabolites most significantly associated with age

Standardized name[a]	Age effect[b]	EL effect[c]	Offspring effect[d]	HR for death[e]	Change rate effect[f]	Age (Adj_*p* value)[g]	EL (Adj_*p* value)[h]	Offspring (Adj_*p* value)[i]	Metab (Adj_*p* value)[j]	Change rate (Adj_*p* value)[k]
N2,N2-dimethylguanosine	1.01	1.74	0.95	1.37	1.01	3.3E−189	2.8E−40	0.424076	6.8E−09	6.7E−14
ST 19:2;O2;S	0.97	0.41	1.19	0.86	0.97	5.9E−143	8.3E−19	0.229744	0.001895	5.0E−55
Alpha-N-phenylacetylglutamine	1.02	2.21	0.90	1.08	1.03	1.0E−127	1.4E−21	0.424076	0.206718	3.8E−27
Threitol	1.01	1.74	0.98	1.19	1.01	4.3E−120	1.3E−43	0.799935	2.4E−05	1.2E−30
Ribose	1.01	1.94	0.99	1.14	1.02	2.2E−116	6.1E−46	0.850312	0.002278	1.5E−28
Octadecatrienoic acid	0.99	0.63	1.05	0.74	0.99	4.4E−114	4.3E−25	0.490719	4.3E−13	2.9E−05
Sucrose	1.03	4.42	0.94	1.25	1.02	1.2E−87	2.0E−32	0.780733	2.8E−07	2.0E−08
Myo-inositol	1.01	1.56	0.99	1.12	1.01	3.0E−86	6.7E−32	0.800028	0.02616	1.6E−06
1-Methyladenosine	1.01	1.30	0.99	1.06	1.01	3.0E−84	2.6E−30	0.788118	0.397384	1.4E−40
Gluconic acid	1.01	1.55	0.98	1.21	1.01	8.7E−81	5.6E−23	0.757756	5.6E−06	4.5E−07
Glucuronic acid	1.01	1.76	0.97	1.16	1.01	3.2E−71	1.6E−30	0.757756	1.0E−04	8.1E−10
Citrulline	1.01	1.30	0.97	0.98	1.01	4.0E−69	2.7E−16	0.740598	0.868024	2.2E−11
Tryptophan	0.99	0.79	1.02	0.85	0.99	1.5E−65	3.8E−17	0.740598	4.6E−06	1.7E−18
Guanidinosuccinic acid	1.02	1.91	0.91	1.02	1.02	2.1E−65	1.7E−20	0.345947	0.864393	4.8E−30
Tartaric acid	1.02	1.47	0.93	1.42	1.00	4.1E−65	9.8E−06	0.740598	4.3E−13	0.86974
Ergothioneine	0.98	0.52	1.03	0.75	0.98	4.3E−62	1.3E−17	0.796437	9.4E−11	1.6E−25
Esculetin	1.01	1.32	0.98	1.17	1.01	4.4E−62	2.8E−14	0.740598	0.00120	1.5E−15
P-cresol	1.01	1.84	0.97	1.02	1.01	5.1E−62	1.2E−13	0.850312	0.814504	4.5E−9
N-acetyl aspartic acid	1.01	1.39	0.98	1.05	1.01	6.7E−60	6.1E−41	0.740598	0.470733	1.4E−39
3-Methylindole	0.99	0.82	1.02	0.87	0.99	1.0E−59	1.9E−13	0.740598	9.9E−05	1.3E−14

[a]From RefMet (July 2023 release).
[b]FC for 1-year difference in the cohort.
[c]FC comparing EL to controls.
[d]FC comparing EL offspring to controls.
[e]HR for mortality for 1 standard deviation increase from the mean (log scale).
[f]FC for 1-year change.
[g]Adj_*p* value for testing age effect.
[h]Adj_*p* value for testing EL effect.
[i]Adj_*p* value for testing offspring effect.
[j]Adj_*p* value for testing metabolite effect on mortality.
[k]Adj_*p* value for testing longitudinal time effect.

of extremely old participants in the study provide a unique opportunity to characterize metabolites that specifically correlate with age, EL, and mortality risk and to distinguish between cross-sectional and longitudinal age effects. We leveraged collaborations with other studies of healthy aging (Arivale[17] and Baltimore Longitudinal Study of Aging [BLSA][18]) and EL (New England Centenarian Study [NECS][19] and a Chinese longevity cohort [XU][20]) to replicate our findings.

RESULTS

Figure S1 summarizes the analytical workflow and cohorts used in this study. We analyzed 188 lipid and 220 polar molecules (Table S1) in a subset of 2,764 LLFS participants with ages ranging between 24 and 110 years and a median age of 74 years at the first blood draw. The sample was approximately sex balanced.

Age correlates with many metabolites

Discovery in the LLFS

We identified 308 metabolites associated with age at 5% false discovery rate (FDR), including 156 metabolites with a positive age correlation and 152 with a negative age correlation. Table 1 includes the 20 most significant results; all results are shown in Table S1. The volcano plot in Figure 1A highlights some significant metabolites, including N2,N2-dimethylguanosine, which increased by an average of 1.0% for each year difference in the cohort (fold change [FC] = 1.01; adjusted *p* value [Adj_*p*] = 3.3E−189), and kynurenine, which increased by an average of 0.4% (Adj_*p* = 3.7E−15). In contrast, tryptophan decreased by

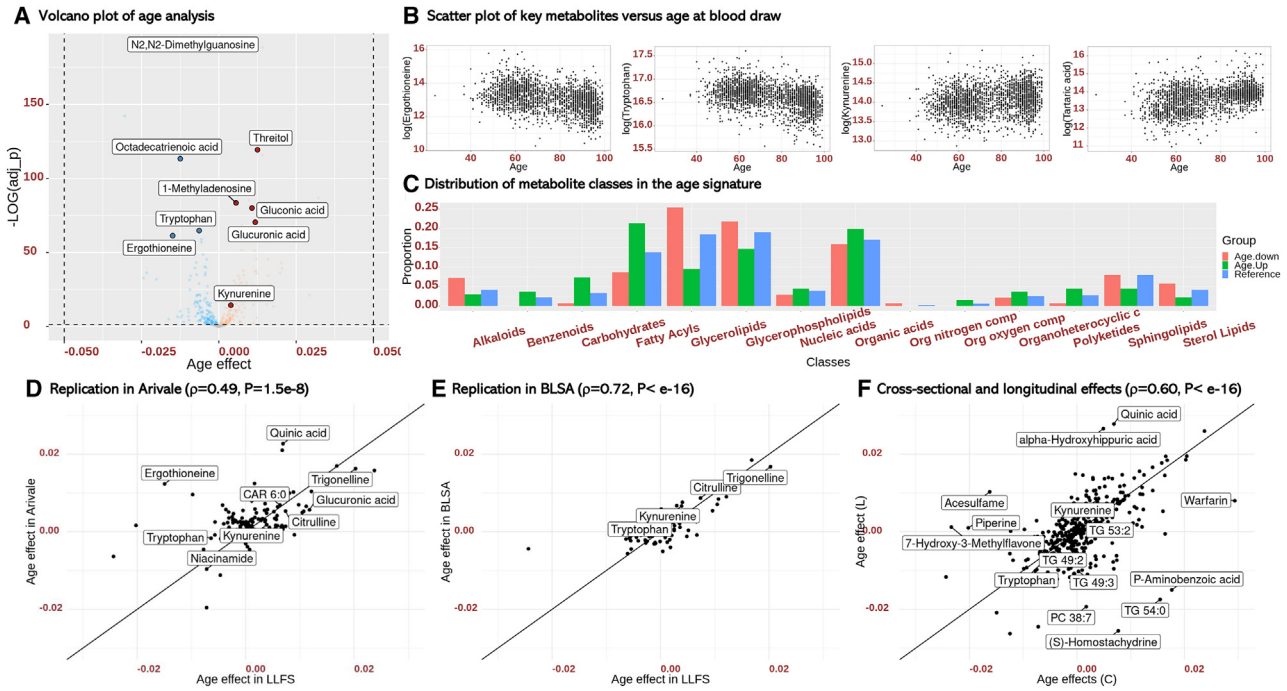


Figure 1. Metabolites correlate with age

(A) Volcano plot of age effect on 408 log-transformed metabolites (x axis) and –log10 Adj_ values (y axes). Age effect: log FC of metabolites for a year older age.
(B) Scatterplots of log-transformed metabolites (y axis) versus age at blood draw (x axis).
(C) Proportion of metabolite main classes (blue) and of metabolites positively correlated (green) and negatively correlated (red) with age.
(D) Age effects on 120 metabolites in the LLFS (x axis) and Arivale (y axis). ρ, Pearson correlation. P, *p* value.
(E) Age effects on 99 metabolites in the LLFS (x axis) and BLSA (y axis).
(F) Cross-sectional (x axis) and longitudinal (y axis) age effects.

an average of 0.6% (Adj_*p* = 1.5E−65), while ergothioneine had a 1.5% decrease (Adj_*p* = 1.5E−62). Sucrose had the largest positive change (FC = 1.03, Adj_*p* = 1.2E−87), and testosterone sulfate had the largest negative change (FC = 0.97, Adj_*p* = 5.9E−143). Some examples are shown in Figure 1B. The bar plot in Figure 1C shows the distribution of metabolite super-classes in our dataset and the distribution of super-classes of metabolites positively and negatively correlated with age. Metabolites positively correlated with age were enriched for benzenoids, carbohydrates, fatty acyls, and organic acids (*p* = 0.015), whereas metabolites negatively correlated with age were enriched for alkaloids, glycerolipids, glycerophospholipids, sphingolipids, and sterol lipids (*p* = 0.001). In agreement with previous results,[21] 24 age-associated lipids were significantly associated with the E2 genotype group (e2e2 or e2e3) of *APOE* (Figure S2).

Replication

After converting names to RefMet nomenclature,[22] we identified 120 metabolites in both the LLFS and Arivale and 99 metabolites in both the LLFS and BLSA (Figure S1, UpSet plot). Figures 1D and 1E show good correlation between age effects estimated in the three studies: 49% correlation with Arivale (*p* = 1.5E−8), 72% correlation with BLSA (p<E−16), and good agreement of significant results (Tables S2A and S2B). We compared LLFS results with those published previously.[23] Only 54 metabolites were on our target list, including 35 with concordant effects and 26 with concordant and significant effects (ρ = 0.58, *p* < 3.9E−6; Table S2C).

Aging affects many metabolic processes

Discovery in the LLFS

We analyzed the aging-related change of the 408 metabolites in 1,236 LLFS participants with two metabolomic profiles that were generated an average of 8 years apart. We identified 117 metabolites with a positive aging effect and 138 with a negative aging effect (FDR < 5%), including 204 metabolites of the age signature (Table S1; Figure S3). The age effects from the cross-sectional and longitudinal analyses had good concordance (Figure 1F; Table 1). The analysis also identified 30 metabolites with significant longitudinal changes that were not associated with age (Figure S3); e.g., cortisol (−0.25% year-change; Adj_*p* = 0.002) and glycocholic acid (0.8% year-change; Adj_*p* = 0.006).

EL is characterized by unique metabolic profiles

Discovery in the LLFS

We identified 122 metabolites that were higher and 108 metabolites that were lower in EL compared to controls (5% FDR). No metabolite was significantly different in offspring of long-lived families when compared to control spouses (Table S1). The volcano plot in Figure 2A highlights some of the most significant

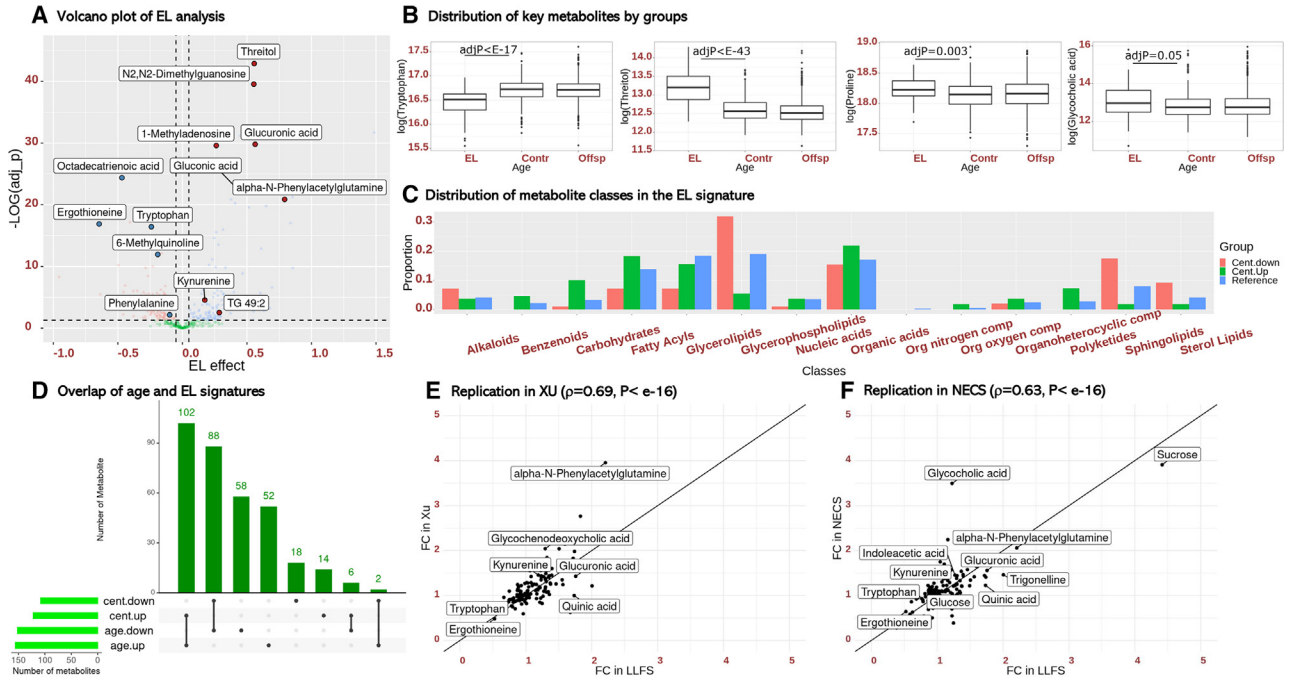


Figure 2. Metabolites correlate with EL

(A) Volcano plot of EL effect on 408 log-transformed metabolites (x axis) and –log10 Adj_*p* values (y axes). EL effect, log FC of metabolite comparing EL to controls.

(B) Boxplots of log-transformed metabolites (y axis) in EL, controls (Contr), and Offspring (Offsp). AdjP, *p* value adjusted for multiple testing.

(C) Proportion of metabolite main classes (blue) and of metabolites higher (green) and lower in EL (red) compared to controls.

(D) UpSet plot of associations with age and EL. Age_up/Age_down, metabolites positively/negatively associated with age; Cent_up/Cent_down, metabolites higher/lower in EL.

(E) EL effects on 121 overlapping metabolites in the LLFS (x axis) and XU (y axis). FC, fold change of metabolites comparing EL to younger controls. ρ, Pearson correlation. P, *p* value.

(F) EL effects on 150 overlapping metabolites in the LLFS (x axis) and NECS (y axis).

results, and some examples are showcased in Figure 2B. The most elevated metabolite in the EL group was sucrose (FC = 4.4, EL vs. controls, Adj_*p* = 2.0E−32), while lenticin was the most downregulated (FC = 0.38, Adj_*p* = 5.1E−09). Figure 2C summarizes the distribution of metabolite super-classes that were higher and lower in EL. Metabolites higher in EL were enriched for benzenoids, carbohydrates, fatty acyls, and organic acids (*p* = 0.026), whereas metabolites lower in EL were enriched for alkaloids, glycerophospholipids, sphingolipids, and sterol lipids (*p* = 0.005). These patterns match those in the age analysis except for the glycerolipid class. The UpSet plot in Figure 2D shows a substantial overlap between metabolomic signatures of age and EL, suggesting that many changes observed in the oldest individuals are a product of their extremely old age rather than drivers of longevity. However, we identified 24 metabolites associated with EL and not with age (FDR > 0.2) or associated with age with opposite direction (Table S3).

Replication

We identified 121 metabolites in both the LLFS and XU and 150 in both the LLFS and NECS. Figures 2E and 2F show substantial agreement between estimated FCs for EL: 69% correlation with XU (*p* < E−16; Table S2D) and 63% correlation with NECS (*p* < E−16; Table S2E).

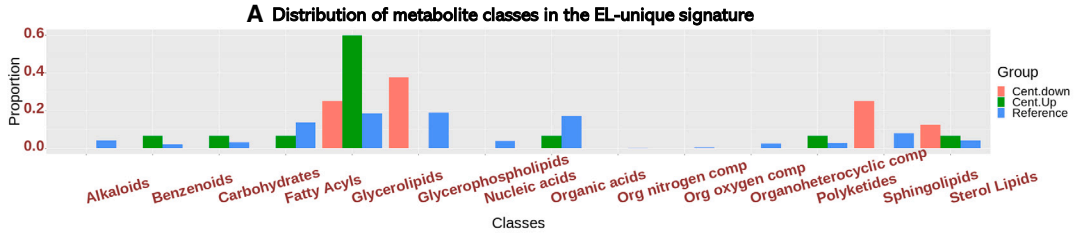
Unique signature for EL

Sixteen metabolites associated with EL were not associated with age (20% FDR), and eight had an opposite age trend (Table S3). Metabolites that were uniquely higher in EL were enriched for glycerolipids, mainly triacylglycerides, while metabolites lower in EL were enriched for glycerophospholipids and sphingolipids (Figure 3A). Triglyceride (TG) 51:0 and TG 53:1 were high in EL and also significantly elevated in carriers of the E2 genotype group of *APOE* compared to E3 carriers, whereas sphingomyelin (SM) 36:1;O2 was low in both EL and E2 carriers (10% FDR) (Figure 3B). The genetic effect was much stronger in homozygote carriers of the e2 allele (Figure 3 C). The association of phosphatidylcholine (PC) 38:6 (measured in NECS only), proline, and glycocholic acid was replicated in NECS and XU (Figure 3D), but the association of phenylalanine with EL in the LLFS (Figure 3E) was in the opposite direction as NECS and XU.

Metabolites predict mortality risk

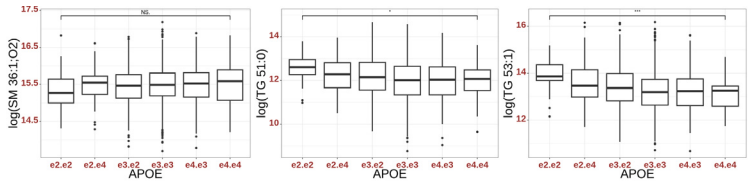
Discovery in the LLFS

Thirty-nine metabolites were positively associated and 113 metabolites were negatively associated with mortality risk (Table 1; Figure 4A; Table S1). Significant findings include ergothioneine, with a hazard ratio (HR) of 0.75 for mortality for 1 standard deviation increase from the mean in log scale (Adj_*p* = 9.4E−11),



B APOE2 effect on the EL-unique signature in LLFS

Standardized Name	Fold change comparing EL to controls	Adjusted p-value	Fold change comparing E2 to E3	Adjusted p-value
SM 36:1;O2	0.86	0.021675	0.75	0.097477
TG 51:0	1.31	0.031913	1.56	0.044617
TG 53:1	1.28	0.029147	1.72	4.38E-05



D Replication of the EL-unique signature

Standardized Name	Super Class	EL-FC (LLFS)	Adjusted p-value	EL-FC (XU)	Adjusted p-value	EL-FC (NECS)	Adjusted p-value
Glucose	Carbohydrates	1.34	0.01698	1.08	1	1.09	0.92525
Glycocholic acid	Sterol Lipids	1.23	0.04705	1.56	1	3.50	0.00020
Proline	Organic acids	1.10	0.00279	1.22	1.72E-06	1.37	0.00251
PC 38:6	Glycerophospholipids	0.86	0.01481			0.89	0.07434
Phenylalanine	Organic acids	0.91	0.00683	1.11	0.002459	1.12	0.02315

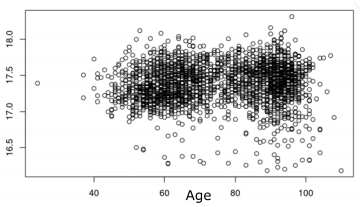


Figure 3. Metabolomic signature of EL independent of age

(A) Proportions of main classes of 24 metabolites (blue): 16 higher (green) and 8 lower in EL (red).
(B) *APOE* effect on EL-associated lipids. E2, e2e2 or e2e3; E3, e3e3.
(C) Distribution of EL-associated lipids by *APOE* genotype groups. *Adj_p < 0.05, **Adj_p < E−04.
(D) EL-FC, FC for EL versus control.
(E) Scatterplot of log-transformed phenylalanine (y axis) and age at blood draw (x axis) in the LLFS.

tryptophan (HR = 0.85, Adj_p = 4.6E−06), octadecatrienoic acid (HR = 0.74, Adj_p = 4.3E−13), N2,N2-dimethylguanosine (HR = 1.37, Adj_p = 6.8E−09), and warfarin (HR = 1.22, Adj_p = 2.3E−14). Tartaric acid was the metabolite with the largest mortality risk (HR = 1.42, Adj_p = 1.9E−17), while 7-hydroxy-3-methylflavone was associated with the most protective effect (HR = 0.72, Adj_p = 1.5E−13). Kaplan-Meier curves in Figure 4B show some example associations. We analyzed the distribution of metabolite super-classes (Figure 4C): metabolites positively correlated with mortality risk were enriched for carbohydrates, nucleic acids, and polyketides (p = 0.0007), while metabolites negatively correlated with mortality risk were enriched for alkaloids, organic acids, fatty acids, sphingolipids, and sterol lipids (p = 0.0007).

Replication

We used BLSA for replication, and Figure 4D shows the scatterplot of the HR estimated for 99 metabolites measured by both studies (Table S2F). Although no metabolite reached significant association with mortality risk in BLSA after multiple-testing correction, the estimated HRs were in good agreement (ρ = 0.39, p = 0.0005). The associations of acetoacetic acid, histidine, lactic acid, leucine, phenylalanine, and valine with mortality in the LLFS replicated those reported previously,[14] but the association of lactic acid did not reach 5% FDR in the LLFS (Table S2G). The associations of ergothioneine, histidine, tryptophan, leucine, valine, methionine, and trigonelline also replicated those reported previously.[24] We found a concordant and significant effect of mortality for 3-dehydroxycarnitine[24] and a concordant but not significant effect of glutamic acid (Table S2H). We also identified 105 metabolites in common with the analysis of longevity described in a prior report[25] and found good agreement between decreased levels of glutamine, histidine, PC 38:6, and threonine and increased mortality risk in the LLFS and decreased odds for living beyond 80 years in the Framingham Heart Study. However, we found discordant and significant results for uridine and aconitate (matched to *trans*-aconitic acid via RefMet[22]) and no association of 3-hydroxybutyric acid, D-malic acid, and glycocholic acid with mortality risk in the LLFS (Table S2K; Figure S4). To confirm that the discordant results were not due to incorrect annotation of uridine in the LLFS, we validated the identification of uridine by matching the retention time of research samples to the retention time of an authentic standard (Figure S4).

Integrative analyses

We integrated the effects of multiple metabolites on age, EL, and mortality risk with signatures, an aging clock, and network analysis.

Metabolomic signatures

The UpSet plots in Figures 4E and 4F show that metabolites associated with age and/or EL in the LLFS were not always associated with mortality risk and vice versa. We grouped the

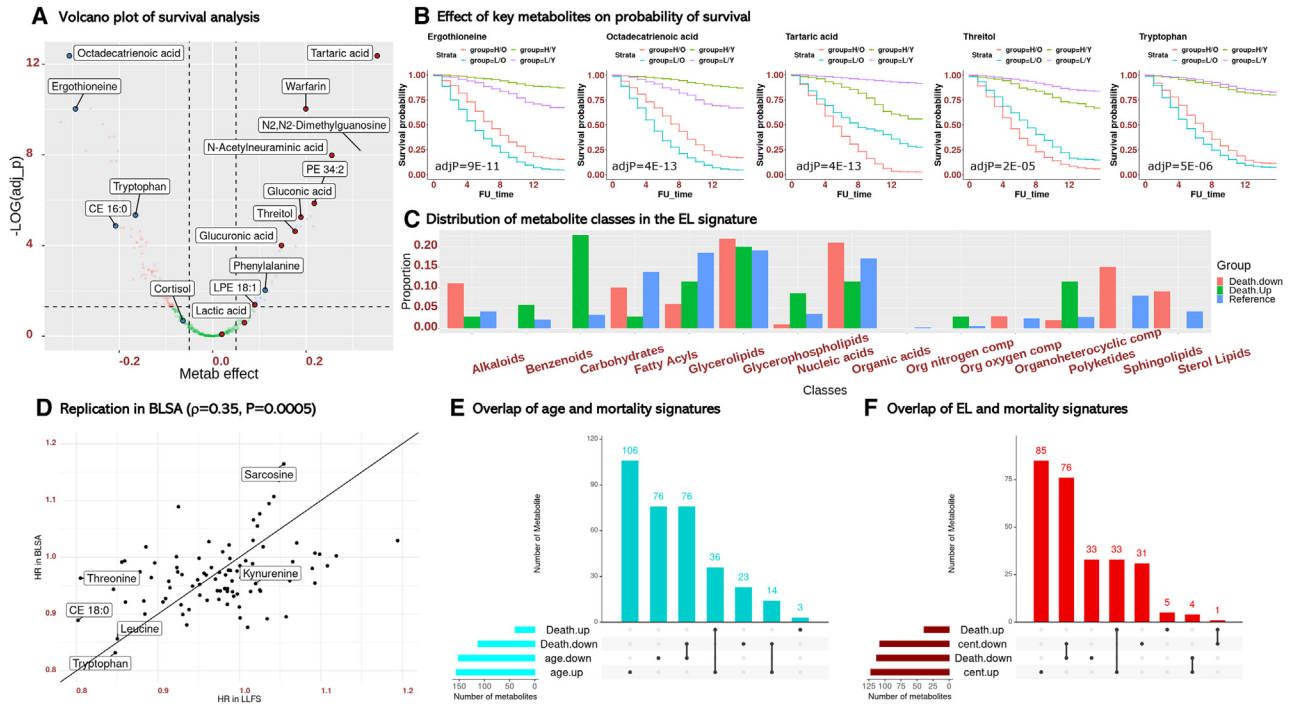


Figure 4. Metabolites predict mortality risk

(A) Volcano plot of effects of 408 log-transformed metabolites on hazard for mortality (x axis) and –log10 Adj_*p* values (y axis). Metab effect, effect of 1 standard deviation from the mean of the log-transformed metabolite on hazard for mortality.

(B) Generation-stratified Kaplan-Meier survival curves. H/O, above-the-mean metabolite in individuals born before 1935. H/Y, above-the-mean metabolite in individuals born in 1935 or later; L/O, below-the-mean mean metabolite in individuals born before 1935; L/Y, below-the-mean mean metabolite in individuals born in 1935 or later; FU_time, years after enrollment in the LLFS. AdjP, *p* value adjusted for multiple testing.

(C) Proportion of metabolite main classes (blue) and of metabolites positively correlated (green) and negatively correlated with mortality risk (red).

(D) HR for mortality associated with 99 metabolites in the LLFS (x axis) and BLSA (y axis). ρ, Pearson correlation. P, *p* value

(E and F) UpSet plots of metabolites associated with age and mortality and EL and mortality. Age_up/Age_down, metabolites positively/negatively associated with age; Cent_up/Cent_down, metabolites higher/lower in EL; Death_up/Death_down, metabolites positively/negatively correlated with mortality risk.

metabolites into 19 signatures based on their directions of effect and 5% FDR association with age, EL, and mortality risk (Figure 5; Table S4). We did not include the longitudinal age effects, which were based on a much smaller sample size. In this analysis, we considered FDR > 20% indicative of null association, and we excluded associations with 0.05 < FDR < 0.20. These signatures show that, among the 308 age-associated metabolites, 79 were associated with EL but not mortality risk (Figure 5A), 101 were associated with both EL and mortality risk (Figure 5B), 15 had a significant effect on mortality risk but not EL (Figures 5C1, 5C4, and 5C5), and 44 were only associated with age (Figures 5C2, and C3). In addition, we identified metabolites that correlated with EL and/or mortality but not with age (Figures 5D and 5E).

Signatures A1 and A2 included 73 metabolites with consistent effects on age and EL, while signature A3 included 6 metabolites negatively correlated with age but elevated in EL, thus suggesting a possible protective effect for EL. Of the 38 metabolites that predicted increasing mortality risk, 33 increased with age and EL (signature B2), two increased with age but not EL (signature C4), two were not associated with age or EL (signature E2), and one was lower in EL (signature D1). Similarly, of the 90 metabolites that predicted decreasing mortality risk, 72 were lower in older age (signatures B1 and C1), 9 did not correlate with age (signatures D4, D5, and E1), and 9 were increased with older age (signatures B3, B4, and C5). Among the mortality-associated metabolites that did not correlate with age, lysophosphatidylcholine (LPC) 22:6/0:0 and TG 58:8 (signature D5) had decreasing levels associate with both EL and increased mortality risk, which points to markers of extremely old age rather than successful aging. Interestingly, both lipids changed significantly in the longitudinal analysis, thus suggesting that these metabolites may show substantial changes only in older ages.

A few signatures stand out as comprising potentially protective metabolites. Metabolites in signature D4 were elevated in EL, and higher levels predicted a decreased hazard for mortality, suggesting that these metabolites could be markers of longevity. Signature B3 included citric acid and *trans*-aconitic acid, which increased with older age and EL. Higher levels correlated with a decreased mortality risk, thus suggesting a protective role during old age. Citric acid levels increased by 0.3% for a year difference (Adj_*p* = 3.3E−13), were higher in EL compared to younger individuals (FC = 1.16, Adj_*p* = 2.7E−06), and a standard deviation increased from the mean in log-scale was associated with a

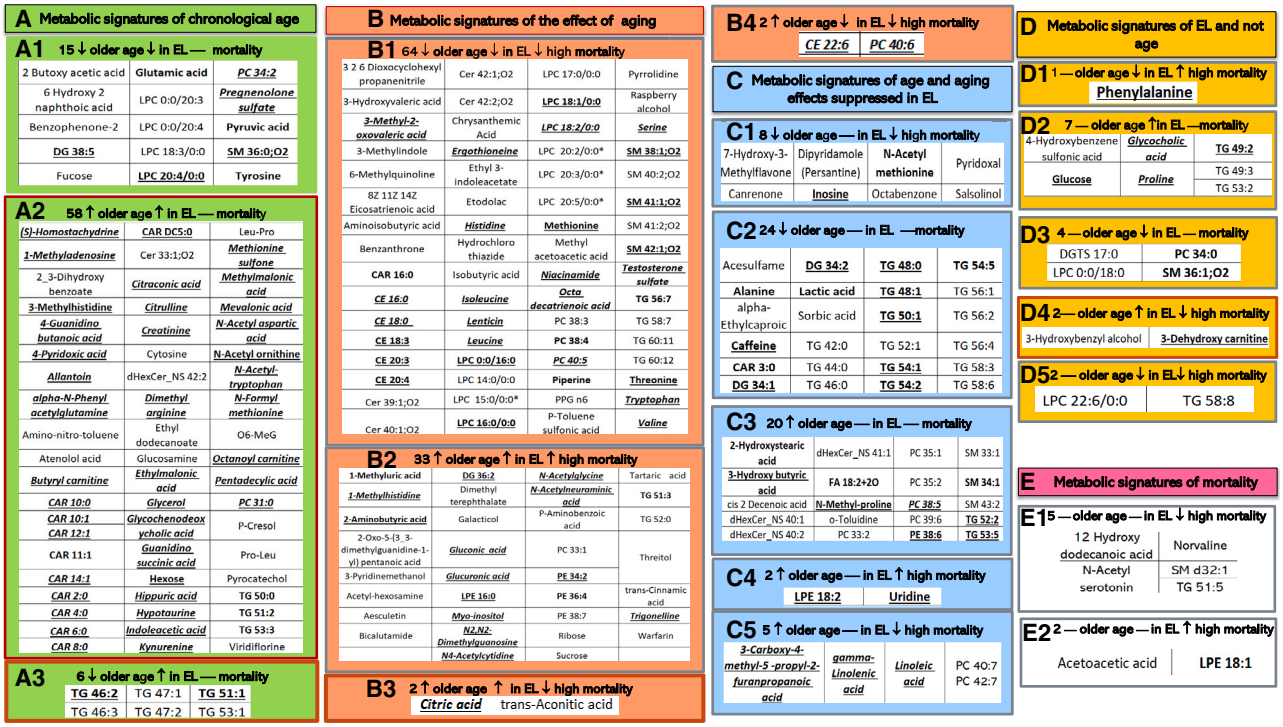


Figure 5. Summary of the metabolomic signatures of age, EL, and mortality

(A) metabolites that correlate with age and EL but not mortality. (B) Metabolites that correlate with age, EL, and mortality. (C) metabolites that correlate with age and/or mortality but not EL. (D) metabolites that correlate with EL and/or mortality but not age. (E) metabolites that correlate with mortality but not age and EL. Upward arrow: positive correlation. Downward arrow: negative correlation. Dash: no correlation. Underlined, validated metabolites with concordant effect; italics, validated metabolites with significant effect; boldface, metabolite found in one or more independent sets; AKOS040738329, NCGC00380721-01_C20H30O4_2-hydroxy-4_5′_8a′-trimethyl-1′-oxo-4-vinyloctahydro-1′H-spiro[cyclopentane-1_2′-naphthalene]-5′-carboxylic acid.

12% reduction of the HR for death (Adj_p = 0.0007). *trans*-Aconitic acid levels increased by 0.4% for a year difference (Adj_p = 5.3E−20), were higher in EL compared to younger individuals (FC = 1.26, Adj_p = 7.2E−12), and a standard deviation increase from the mean was associated with a 12% reduction of mortality risk (Adj_p = 0.00376).

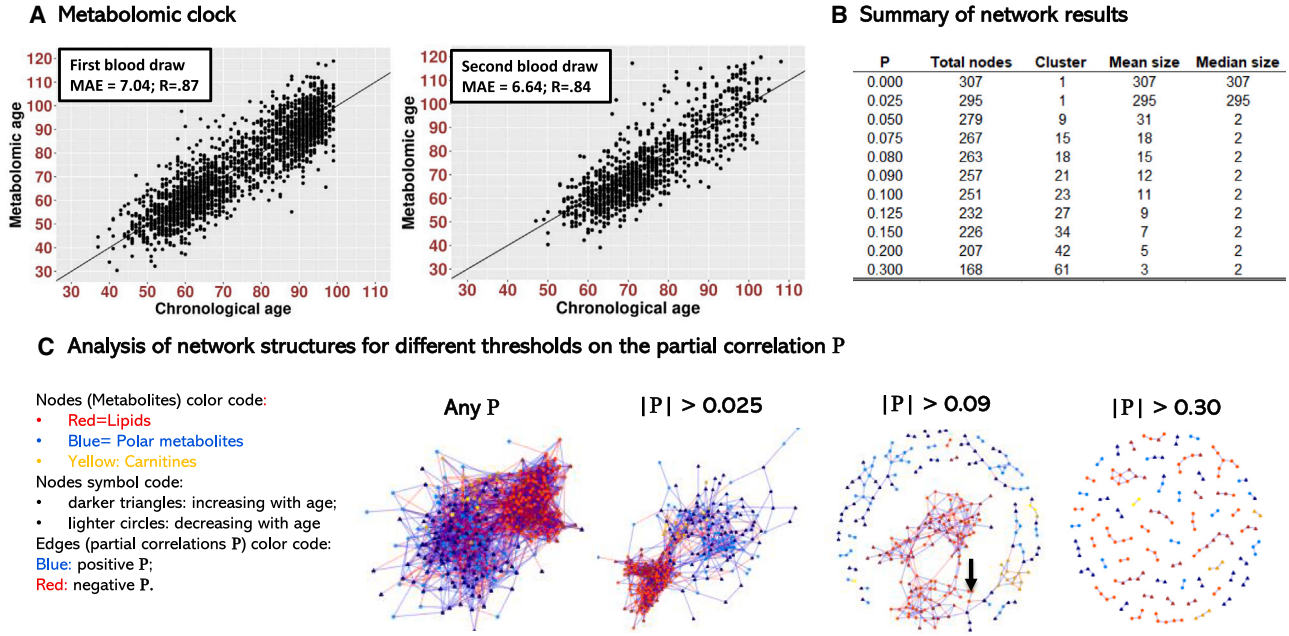
Cholesteryl ester (CE) 22:6 and PC 40:6 in signature B4 positively correlated with age, and higher levels were associated with a significant reduction in mortality risk, thus suggesting a protective mechanism. However, these metabolites were lower in EL. The five metabolites in signature C5 were higher in older ages, and higher levels appeared to reduce mortality risk, but we did not detect a significant difference in EL compared to younger age groups. This set included, for example, linoleic acid, which increased by ∼0.3%/year (Adj_p = 0.00001), and 1 standard deviation increase from the mean in log-scale was associated with an 8% reduced death risk (Adj_p = 0.04156). Other examples are listed in Table S1.

Metabolomic clocks

We generated a clock from the 308 age-associated metabolites. The clock included 137 metabolites (Table S5) and correlated well with age after using the bias correction reparameterization introduced by de Lange and Cole[26] (Figure 6A). Metabolites with the greatest effects on the clock included octadecatrienoic acid, testosterone sulfate, and dHexCer_NS 40:2, with predicted changes ranging between 1.5 and 2.5 years for log FCs of 1 standard deviation. To investigate the predictive value of the clock, we analyzed the association between metabolic age acceleration and time to death in 2,250 LLFS participants. Age acceleration was associated with increased mortality risk (HR = 1.05, p < 2E−16), and the distribution of martingale residuals showed no evidence of bias (Figure S5C). We used the clock to calculate the metabolic age of LLFS participants at the second blood draw and compared it to the chronological age (Figure 6A). The 84% correlation between metabolic and chronological age suggests that the clock works well in new data.

Partial correlation networks of age associated metabolites

The network without restriction on partial correlation connected all metabolites (Figures 6B and 6C1), with an average of 16 edges per node (Table S6; Figures S6). The most connected nodes were the diacylglycerol DG 34:1 (38 connections) and PC 33:1 (36 connections). As we imposed more stringent thresholds on partial correlations, the dependency structure of the networks was simplified. Figure 6C shows that the network with absolute partial correlations significantly greater than 0.025 began to separate into two highly connected components enriched for lipids and polar metabolites. This separation increased in networks with absolute partial correlations greater than 0.05 (Figure S6). When we increased the threshold to 0.09, only 257



P	Total nodes	Cluster	Mean size	Median size
0.000	307	1	307	307
0.025	295	1	295	295
0.050	279	9	31	2
0.075	267	15	18	2
0.080	263	18	15	2
0.090	257	21	12	2
0.100	251	23	11	2
0.125	232	27	9	2
0.150	226	34	7	2
0.200	207	42	5	2
0.300	168	61	3	2

Figure 6. Summary of integrative analyses

(A) Left: metabolic (y axis) against chronological (x axis) age using a bias-corrected metabolomic clock. Right: metabolic versus chronological age at follow-up. MAE, mean absolute error; R, Pearson correlation.

(B) Summary table of the networks generated with different thresholds on the absolute partial correlation P.

(C) Markov graphs of the networks inferred with different thresholds on the absolute partial correlation. The black arrow in the third graph points to linoleic acid and gamma-linolenic acid, which separate the majority of lipid and polar metabolites.

metabolites remained connected and clustered into two large networks and several small subnetworks (Figures 6C and S6). One large network consisted of lipids and polar metabolites that were conditionally independent, given linoleic and gamma linolenic acids. More conservative thresholds produced sparser networks (Figures 6 and S6). Figure S7 shows that, with absolute partial correlation greater than 0.3, metabolites clustered into many small networks of highly correlated metabolites including the network of tryptophan, 3-methylindole, and 6-methylquinoline or the cluster of gamma-linolenic, linoleic, and *cis*-vaccenic acid. Table S6 includes degree analysis, top hubs, and vulnerability nodes of each network. Interactive versions of these networks are available as supplemental information.

DISCUSSION

Overview

Metabolism plays a major role in aging and aging-related diseases,[2] but key metabolic markers of aging are not fully known. We identified 308 age-associated metabolites, 230 EL-associated metabolites, and 152 metabolites that predicted mortality risk. We replicated the associations of 87 metabolites with age in at least one study and the associations of 100 metabolites with EL in at least one study (Tables S2 and S7). We did not replicate results for survival analysis in the BLSA, but the estimated effects showed high concordance. In addition, our results agreed with metabolites described as predictors of mortality risk in other studies (Table S2).[14,24,25] We also identified 255 metabolites that changed over an average of 8 years, thus implicating many metabolites in human aging. This set included 204 age-associated metabolites. Both cross-sectional and longitudinal analyses suggest that yearly FCs of most metabolites are small, within 2%–3%. The concordant estimates from the cross-sectional and longitudinal analyses may result from the relatively short follow-up, and longer longitudinal studies are needed to capture potentially non-linear changes over time. We identified metabolites with longitudinal changes that did not correlate with age and could be biomarkers of underlying conditions unrelated to aging (Figure S3). We integrated the results via signatures, a metabolomic clock, and partial correlation networks.

Metabolomic signatures

The intersection of the results from the various analyses suggests that there is not a single metabolomic signature of age, EL, or mortality but, rather, a variety of signatures with different clinical and biological interpretations and translational potential. In Figure 5, we summarized five groups of signatures that distinguish between biological mechanisms of age, mortality, and EL.

Metabolic signatures of age

We identified 73 metabolites with higher or lower levels in older ages and an age-consistent pattern in EL that did not predict mortality (Figures 5A1 and 5A2), and six lipids that decreased with older age were higher in EL compared to younger individuals and did not predict mortality (Figure 5A3). Signatures A1 and A2 comprised several LPC and sphingolipids (Ceramydes

[Cer] and SM) that were lower in older ages and EL, TGs that were higher in older ages and EL, and fatty acids that had a mixed pattern. Our results agree with lower levels of LPC 20:4 identified previously in centenarians compared with younger subjects[13] and the age-related decline of LPC 20:4 shown in the Genetics of Lipid Lowering Drugs and Diet Network (GOLDN) study.[27] Several acylcarnitines were increased in older age and EL. Acylcarnitine formation is required to generate energy from fatty acids via mitochondrial oxidation,[28] and the age-related increase for acylcarnitines may reflect decreased mitochondrial function.

Several metabolites in signature A2 are elevated in the presence of kidney disease, including butyrylcarnitine, citrulline, creatinine, dimethylarginine, the gut-derived metabolite hippuric acid, and alpha-N-phenylacetylglutamine.[29–31] Consistent with other studies, kynurenine levels were higher in older ages and EL, but the effect of kynurenine on mortality risk was insignificant in the LLFS (HR = 1.02, Adj_p = 0.82). Signature A1 included tyrosine, which is involved in the production of dopamine and may correlate with cognitive function.[32] This signature also included some markers of medications; for example, the beta-blocker atenolol acid.

Signature A3 included 6 lipids that decreased with older age but were higher in EL compared to younger individuals. This signature included TG 53:1, which was elevated in carriers of the E2 allele of *APOE* (Figures 3B and 4C), thus suggesting that this signature could be a specific driver of EL under genetic regulation.

Metabolic signatures of the effect of aging

We identified four groups of metabolites (Figure 5B) that were associated with age, EL, and mortality risk. Signatures B1 and B2 showed patterns consistent with a damaging effect of aging (concordant directions of effect with age, EL, and mortality risk), while B3 and B4 point to possibly longevity-enabling processes (associations with age and mortality risk have opposite directions of effect).

Signature B3 included citric acid and *trans*-aconitic acid (aconitate) that were higher at older ages and EL, and were associated with a decreased mortality risk, thus suggesting an age-related protective/compensatory mechanism. This hypothesis is supported by other studies: citric acid increased during aging,[33] and relatively higher levels of citric acid were associated with a younger biological age.[34] Plasma levels of citric acid were significantly higher in centenarians when compared with non-centenarians,[35] and dietary supplements of citric acid increased metabolic health and extended the lifespan in *Drosophila melanogaster*.[36] *Trans*-Aconitic acid increased during aging for people in the age range of 40–65 years,[37] but higher levels of this metabolite correlated with high mortality risk in both the Framingham Heart Study[25] and Women's Health Initiative Study.[38] Both metabolites are part of the tricarboxylic acid cycle, an essential component of aerobic metabolism that occurs in mitochondria. The importance of the tricarboxylic acid cycle to longevity in lower organisms is well understood,[39] and our data suggest that its activity might facilitate EL in humans.

Signature B4 included CE 22:6 and PC 40:6. These lipids were higher at older ages, and lower levels were associated with an increased mortality risk, thus suggesting a protective role for high levels of these two metabolites. Consistently with this hypothesis, Carrard et al. recently reported on the favorable role of circulating CE 22:6.[40]

Thirty-two of the 64 metabolites in signature B1 were lipids. Low plasma levels of LPC 16:0, LPC 17:0, LPC 18:1, LPC 18:2, and LPC 20:3 have been associated with impaired mitochondrial oxidative capacity.[41] In agreement with our findings, LPC 18:2, 20:3, and 20:5 decreased during aging in the GOLDN study,[27] and LPC 18:2 was lower in centenarians than younger subjects.[13] In addition, LPC 14:0, 15:0, 18:1, 18:2, and 20:2 were reduced in obese subjects,[42] and lower levels of LPC 16:0, 17:0, and 18:1 predicted memory and gait speed decline in older adults.[43] Several glycerophospholipids, including LPC 16:0, LPC 18:2, PC 38:3, and PC 38:4, have been associated previously with mortality risk.[44] Lower levels of some ceramides and sphingomyelins in centenarians have been reported previously,[45] and low levels of SM 38:1, 41:1, and 42:1 are predictive of mortality.[46] Signature B1 also included tryptophan, which has a critical role in aging.[47] Our results are consistent with several studies showing a decline of plasma tryptophan in older individuals and centenarians.[13,48] Ergothioneine is considered one of the most promising "longevity vitamins."[49,50] The strong association of ergothioneine with aging and mortality risk in the LLFS is consistent with evidence showing that low ergothioneine increases the risk of coronary artery disease, CVD mortality, and overall mortality.[51] As in the LLFS, low levels of ergothioneine, aminoisobutyric acid, histidine, leucine, methionine, tryptophan, and valine were associated with an increased mortality risk.[52] Higher values of isoleucine were associated with a higher mortality risk,[14] but in the LLFS, we observed higher levels of predicting lower mortality risk (HR = 0.88, Adj_p = 0.0087). Increasing mortality risk with lower glutamine, histidine, and threonine levels have also been reported in the Framingham Heart Study[25] and Women's Health Initiative Study,[38] although we did not detect a significant association between glutamine levels and mortality risk in the LLFS.

Signature B2 included well-known markers of age, disease, and overall mortality, including 1-methylhistidine,[53] 1-methyluric acid,[54] N2,N2-dimethylguanosine, glucuronic acid,[38,55] N4-acetylcytidine,[38] N4-acetylneuramininic acid,[56] and trigonelline.[24,55] Consistent with our results, circulating levels of galactitol, gluconic acid, glucuronic acid, myo-inositol, ribose, threitol, and sucrose were higher in people with chronic liver disease and/or hepatocellular carcinoma when compared with controls.[57] Gluconic acid is elevated in the presence of hyperglycemia and could mark oxidative stress.[58] High plasma levels of threitol and myo-inositol were associated with reduced kidney function across 3 separate studies.[59] The phospholipid LPE 16:0 was positively associated with age in the GOLDN study,[27] and PE 34:2, PE 38:7, and PE 36:4 were associated with an increased mortality risk in other work.[44] A novel finding was the association between increased levels of tartaric acid and mortality risk (Figures 1B and 4B). We validated the identification of this metabolite with an authentic standard but did not find data for independent replication of the association with mortality. The common use of tartaric acid in food warrants additional studies to assess the toxicity level of this metabolite, considering the hypothetical role of this metabolite in colorectal cancer.[60]

Metabolic signatures of age and aging effects absent in EL

We identified 59 metabolites that were associated with age, with or without affecting the risk for mortality, and did not associate with EL (Figures 5C1–5C5). Signature C1 included several metabolites that were lower in older age groups, with lower levels associated with increased mortality risk, and may mark specific nutrient intake at older age. 7-Hydroxy-3-methylflavone is a plant-derived metabolite whose intake is associated with reduced cardiovascular disease (CVD) risk[61] and anti-aging and senolytic effects in lower organisms.[62] Salsolinol is a compound found in chocolate,[63] and chocolate consumption is associated with a reduced all-cause mortality risk.[64] Pyridoxal (vitamin B_6) is derived via diet, and a higher B_6 intake is associated with a reduced all-cause mortality risk.[65] The association of N-acetylmethionine with age and mortality risk in the LLFS did not replicate in NECS and XU, but increasing levels of N-acetylmethionine predicted high risk of mortality in men in other studies.[66]

Signature C4 included uridine, which was higher in older age groups. Higher uridine levels were associated with an increased mortality risk, but they did not differ in the EL group. In disagreement with our results, uridine was a predictor of longevity in other studies,[25,67] and levels of this metabolite were significantly lower in XU centenarians compared to younger controls.[20] Notably, there was substantial heterogeneity in the levels of uridine at younger ages in the LLFS (Figure S4), and some of the discordant results across studies may depend on the sample selection.

Signature C5 included PC 40:7, PC 42:7, and linoleic and gamma-linolenic acids. Consistent with our results, these lipid metabolites were positively correlated with age in the GOLDN study.[27] Dietary intake of linoleic acid has increased substantially in Western diets, but there are questions about impacts of high linoleic acid in brain aging.[68] The increased levels of these metabolites and associated reduction in mortality risk may represent protective compensatory mechanisms of older age. The lack of change of these markers in the EL group may indicate no need to compensate for these aging-related mechanisms in individuals who reach EL.

Metabolomic signatures of EL but not age

Sixteen metabolites were not associated with age but had unusually high or low values in the EL group (Figures 5D1–5D5). Phenylalanine was not associated with age but was lower in EL compared to younger individuals, and higher levels predicted increased mortality risk. These age and EL associations in LLFS did not replicate in Arivale, BLSA, NECS, and XU (Table S2) or other reports.[69] In the LLFS, we observed several individuals with low plasma levels of phenylalanine that may explain the discordant trend (Figure 3E).

Signatures D2, D3, and D4 included potential pro-longevity metabolites. The lack of association with age of the TG species in signature D2 was consistent with the GOLDN study.[27] Association of SM 36:1;O2 with the E2 allele of *APOE* (Figures 3B and 4C) suggests that this lipid profile could be a specific driver of EL. The higher plasma levels of glycocholic acid and proline in EL were replicated in the NECS and XU studies (Figure 3D), and glycocholic acid was higher in individuals with normal cognition compared to dementia.[70] 3-Hydroxybenzyl alcohol and 3-dehydroxycarnitine were elevated in centenarians, and higher values predicted decreased mortality risk, thus pointing to protective mechanisms for EL. 3-Dehydroxycarnitine is a gut bacterium-derived metabolite that is higher in non-diabetic patients compared to diabetic patients.[71] Elevated 3-hydroxylbenzyl alcohol could be a marker of elevated nicotinamide adenine dinucleotide (NAD) levels and reduced oxidative stress in EL.[72]

Signature D5 included LPC 22:6/0:0 and TG 58:8, which were lower in EL, and low values were predictive of higher mortality, which suggests that these metabolites may be biomarkers of mortality in extremely old age. Supporting this hypothesis, these two metabolites changed in the longitudinal analysis but did not correlate with age.

Metabolomic signatures of age-independent mortality

We identified seven biomarkers that predicted mortality that were not correlated with age or EL (Figures 5E1 and 5E2). In cell and animal models, norvaline alleviated hypertension, and it was anti-inflammatory and reversed cognitive decline.[73] In animal models, N-acetylserotonin was neuroprotective and had anti-depressant and cognition-enhancing effects.[74] Lysophosphatidylethanolamine (LPE) 18:1 induced monocytes and epithelial cells to release pro-inflammatory cytokines[75] and could be a biomarker of Alzheimer's disease.[76] Several of these metabolites declined longitudinally (Figure S3) and may be markers of underlying disease.

Metabolomic clock

We built an aging clock that aggregated the effects of 137 age-associated metabolites. The clock included many metabolites with small effects, providing evidence that aging is characterized by small changes of many metabolites. We showed that this clock can be used with new metabolite data and still predicts reasonably well. The metabolomic age acceleration derived from this clock appears to be an informative marker of mortality, and younger metabolomic age correlated with a 5% reduced risk for mortality. These results are consistent with other metabolomic clocks.[3,77]

Age-associated metabolomic networks

The network analysis showed that each age-associated metabolite is correlated with an average of 16 metabolites, but these correlations are not very strong (Figure 6C). While the large number of connections supports grouping multiple metabolites into signatures with similar associations with age, the weak correlations corroborate the need to include many metabolites to predict age and suggest that there is little redundancy in the elements of the various signatures. This finding is consistent with the heterogeneity of aging. The analysis also showed that age-associated lipids are more strongly correlated than age-associated metabolites, and correlations of species are stronger within the same class than between different classes. However, not all species from the same classes were strongly correlated. For example, TG species clustered into several small networks that were consistent with their grouping into signatures with different effects on age, EL, and mortality risk. The analysis also suggests an important role for linoleic and gamma-linolenic acids as connections between lipids and other metabolites. While it is tempting to put forward possible mechanistic relations based on the network analysis, a current limitation of the data is the

lack of serum proteomics and, therefore, important information about enzyme regulation. Additional data and analyses are needed to help translate these findings into hypotheses that can be tested experimentally.

Conclusions

We identified several groups of metabolites with specific roles in age, mortality, and EL that could provide targets for healthy aging therapeutics. Levels of many metabolites changed with age and maintained consistent patterns in EL without affecting mortality risk. These findings suggest that not everything that changes with age is "bad" and should be reversed. In contrast, the levels of metabolites in Figure 5B were different in different age groups, maintained a consistent pattern in EL, and affected mortality risk. Many of these metabolites likely represent cumulative damage due to aging and include many well-known biomarkers of aging-related diseases. It is surprising that some individuals can reach extreme ages with these unhealthy metabolomic profiles, but they may have protective factors that determine their resiliency to age-related stressors. We also identified groups of metabolites that should represent the effect of age and aging-related disease but were unchanged in EL compared to younger controls (Figure 5C). While some of these metabolites may mark poor health and/or eating habits that are absent in successful aging, others may represent compensatory mechanisms that are triggered in old age.

It is well established that lipids play an important role in aging. Their metabolism is intertwined with pro-longevity pathways such as insulin/insulin growth factor 1, mammalian target of rapamycin (mTORC) signaling, and germline endocrine signaling, which have been targets of aging therapeutics and lifespan extension in non-human species.[78] We observed major changes of all lipid classes in almost every signature, although we did not detect major themes (e.g., "TG decrease with age") but, rather, specific effects of lipid species. The network analysis provided further evidence for these results. One novel theme was the increased levels of some glycerolipids in EL and in carriers of the E2 allele of *APOE*, which points to potential targets for healthy aging therapeutics. We also noticed some discordant results in patterns of lipids with older age in the LLFS compared to the literature. While incorrect identification of lipids and metabolites may be one explanation, an alternative hypothesis is that many metabolic changes happen in concert, and groups of metabolites should be analyzed together rather than one at a time.

We observed strong effects of many gut-produced metabolites on aging and longevity, and while many of these metabolites are toxic, others are protective. Integrating plasma metabolomics with the gut microbiome will be important to decipher these mechanisms and identify interventions for healthy aging.

Both tryptophan metabolism and the tricarboxylic acid cycle emerged as important pathways for age, EL, and mortality. The role of these pathways in aging and longevity in humans and other organisms has been studied extensively,[39] but the heterogeneity of some key intermediates among LLFS participants suggests that uncharacterized regulatory mechanisms may be involved.

Many essential amino acids changed with age and may represent different dietary patterns due to age or alterations in the absorption/degradation of many nutrients, which could be important targets for nutritional intervention.[79] High levels of citric acid (present in fruits and vegetables) reduced the risk for mortality in the LLFS. Flavones (plant-derived metabolites present in celery and parsley[80]) and salsolinol (a marker of chocolate consumption) were maintained at youthful levels in centenarians, suggesting a protective role of these metabolites in healthy aging. Associations of phospholipids and TGs with aging, EL, and mortality further support the hypothesis that improved dietary quality may be related to improved survival in the LLFS. For example, LPC 20:5 was increased in young adults who supplemented with fish oil,[81] and plasma TG 60:12 and LPC 20:5 levels were increased following a 12-week randomized control trial for people on a "healthy diet."[82] Plasma LPC 15:0, 16:0, 18:1, 18:2, and 20:5 were decreased in mice on a high-fat diet.[42] Higher levels of these metabolites were found in youths and were associated with a lower all-cause mortality risk in the LLFS. Ergothioneine association with age and mortality also suggests a role of dietary intake in circulating levels, as it is abundant in mushrooms and asparagus.[83] All these results point to nutrients as targets for powerful healthy aging interventions, but they also point to the need for carefully designed intervention studies that assess dosage and possible interactions of multiple nutrients on healthy aging and longevity.

Limitations

When performing metabolomics with liquid chromatography/mass spectrometry, compounds are identified by matching data from research samples to reference data from authentic standards. The LLFS used state-of-the-art computational pipelines for identification of metabolites, but unambiguous identification is not always possible, and further independent validation of these metabolite findings should be performed. We replicated many results in independent studies, but the lack of standard metabolite nomenclature made mapping between studies imperfect and limited the number of metabolites that we could potentially replicate. In addition, the results from the longitudinal analysis will need to be replicated in independent and possibly longer follow-up studies. We adopted robust statistical modeling techniques to account for the relatedness of individuals in the LLFS, but both the survival analysis results and the change analysis of longitudinally collected samples may be affected by informative dropouts or mortality selection. These are challenging statistical problems that do not yet have a robust solution. Given the limited sample size of the LLFS, we did not investigate sex- and race-specific effects of metabolites on age, EL, and mortality risk. Some of the age-related changes may represent medications and underlying conditions that correlate with age rather than age effects. We addressed this question by adjusting our analyses for dyslipidemia, hypertension, type II diabetes, and heart disease-related medications that are well curated in the LLFS, but we could not include medications for kidney disease or other conditions. These critically important questions need to be addressed with larger studies. We could not replicate the clock analyses in truly independent data, and validation is needed before the clock can be used in other studies. Finally, a limitation of the LLFS is the lack of diet data, which could help characterize natural interventions contributing to healthy aging.

RESOURCE AVAILABILITY

Lead contact

Requests for further information, resources, and reagents should be directed to and will be fulfilled by the lead contact, Paola Sebastiani (psebastiani@tuftsmedicalcenter.org).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- Metabolomics data in the LLFS available from the Exceptional Longevity Translational Resources (ELITE) portal: https://eliteportal.synapse.org/Explore/Studies/DetailsPage/StudyDetails?studyKey=LLFS_Metabolomics.
- All additional data generated for the LLFS until 2020 are available from the database of Genotypes and Phenotypes: dbGaP Study (nih.gov). Data from the BLSA can be requested from L.F.. Qualified researchers can further access the full Arivale deidentified dataset supporting the findings in this study for research purposes by signing a data use agreement. Inquiries to access the data can be made at data-access@isbscience.org and will be responded to within 7 business days. NECS data will be available from the ELITE portal.
- All of the code we used to identify and harmonize the metabolomics data is available at https://doi.org/10.5281/zenodo.13224269. All downstream analyses in the LLFS were conducted using R, and the scripts are available from the GitHub repository: QM-DS-Tufts-Medical-Center/Metabolite-signatures-of-chronological-age-aging-survival-and-longevity: v1.0.0 (zenodo.org). https://doi.org/10.5281/zenodo.13155184. Scripts used for analyses of Arivale data are available upon request.
- Any additional information required to reanalyze the data reported in this work is available from the lead contact upon request.

ACKNOWLEDGMENTS

This work was supported by NIA U19-AG063893 (to G.J.P. and T.T.P.), NIA UH2-AG064704 (to T.T.P. and P.S.), NIA UH3-AG064706 (to N.S.), NIA R01-AG061844 (to P.S. and T.T.P.), NIA U19-AG023122 (to N.R., T.T.P., P.S., and N.S.), U19-AG065169-01A1 (to N.J.S.), and 1KL2-TR002545 (to A.A.K.). This study was supported by the USDA Agricultural Research Service under Cooperative Agreement 58-8050-9-004. The graphical abstract was generated using BioRender.

AUTHOR CONTRIBUTIONS

Conceptualization, P.S. and S.M.; data curation, P.S., S.M., M.S.-H., E.S., and G.J.P.; formal analysis, P.S., S.M., Z.S., D.E., Q.T., M.S.-H., E.S., A.G., T.K., A.L., M.L., H.J.L., M.I.S., Q.X., M.M.M., H.B., and N.R.; funding acquisition, P.S., K.C., L.F., N.R., T.T.P., and G.J.P.; methodology, P.S., S.M., M.S.-H., E.S., and G.J.P.; resources, P.S., S.M., L.F., N.R., T.T.P., and G.J.P.; software, P.S., S.M., and Z.S.; validation, P.S., S.M., D.E., Q.T., N.R., L.F., and G.J.P.; writing – original draft, P.S. and M.S.L.); writing – review & editing, all authors.

DECLARATION OF INTERESTS

M.E.M. receives research funding unrelated to this work from Regeneron Pharmaceutical Inc.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.celrep.2024.114913.

Received: September 22, 2023
Revised: July 5, 2024
Accepted: October 10, 2024
Published: November 5, 2024

REFERENCES

1. Srivastava, S. (2019). Emerging Insights into the Metabolic Alterations in Aging Using Metabolomics. Metabolites *9*, 301. https://doi.org/10.3390/metabo9120301.
2. López-Otín, C., Blasco, M.A., Partridge, L., Serrano, M., and Kroemer, G. (2023). Hallmarks of aging: An expanding universe. Cell *186*, 243–278. https://doi.org/10.1016/j.cell.2022.11.001.
3. Panyard, D.J., Yu, B., and Snyder, M.P. (2022). The metabolomics of human aging: Advances, challenges, and opportunities. Sci. Adv. *8*, eadd6155. https://doi.org/10.1126/sciadv.add6155.
4. Adav, S.S., and Wang, Y. (2021). Metabolomics Signatures of Aging: Recent Advances. Aging Dis. *12*, 646–661. https://doi.org/10.14336/ad.2020.0909.
5. Stancliffe, E., Schwaiger-Haber, M., Sindelar, M., Murphy, M.J., Soerensen, M., and Patti, G.J. (2022). An Untargeted Metabolomics Workflow that Scales to Thousands of Samples for Population-Based Studies. Anal. Chem. *94*, 17370–17378. https://doi.org/10.1021/acs.analchem.2c01270.
6. Moaddel, R., Ubaida-Mohien, C., Tanaka, T., Tian, Q., Candia, J., Moore, A.Z., Lovett, J., Fantoni, G., Shehadeh, N., Turek, L., et al. (2024). Cross-sectional analysis of healthy individuals across decades: Aging signatures across multiple physiological compartments. Aging Cell *23*, e13902. https://doi.org/10.1111/acel.13902.
7. Kuiper, L.M., Polinder-Bos, H.A., Bizzarri, D., Vojinovic, D., Vallerga, C.L., Beekman, M., Dollé, E.T., Ghanbari, M., Voortman, T., Reinders, M.J.T., et al. (2023). Epigenetic and Metabolomic Biomarkers for Biological Age: A Comparative Analysis of Mortality and Frailty Risk. J. Gerontol. A Biol. Sci. Med. Sci. *78*, 1753–1762. https://doi.org/10.1093/gerona/glad137.
8. Janssens, G.E., Grevendonk, L., Schomakers, B.V., Perez, R.Z., van Weeghel, M., Schrauwen, P., Hoeks, J., and Houtkooper, R.H. (2023). A metabolomic signature of decelerated physiological aging in human plasma. Geroscience *45*, 3147–3164. https://doi.org/10.1007/s11357-023-00827-0.
9. Tian, Q., Adam, M.G., Ozcariz, E., Fantoni, G., Shehadeh, N.M., Turek, L.M., Collingham, V.L., Kaileh, M., Moaddel, R., and Ferrucci, L. (2023). Human Metabolome Reference Database in a Biracial Cohort across the Adult Lifespan. Metabolites *13*, 591. https://doi.org/10.3390/metabo13050591.
10. Wang, F., Tessier, A.J., Liang, L., Wittenbecher, C., Haslam, D.E., Fernández-Duval, G., Heather Eliassen, A., Rexrode, K.M., Tobias, D.K., Li, J., et al. (2023). Plasma metabolomic profiles associated with mortality and longevity in a prospective analysis of 13,512 individuals. Nat. Commun. *14*, 5744. https://doi.org/10.1038/s41467-023-41515-z.
11. Hornburg, D., Wu, S., Moqri, M., Zhou, X., Contrepois, K., Bararpour, N., Traber, G.M., Su, B., Metwally, A.A., Avina, M., et al. (2023). Dynamic

lipidome alterations associated with human health, disease and ageing. Nat. Metab. *5*, 1578–1594. https://doi.org/10.1038/s42255-023-00880-1.

12. Mäkinen, V.P., and Mika, A.K. (2016). Metabolomics of aging requires large-scale longitudinal studies with replication. Proc. Natl. Acad. Sci. USA *113*, E3470. https://doi.org/10.1073/pnas.1607062113.
13. Collino, S., Montoliu, I., Martin, F.P.J., Scherer, M., Mari, D., Salvioli, S., Bucci, L., Ostan, R., Monti, D., Biagi, E., et al. (2013). Metabolic signatures of extreme longevity in northern Italian centenarians reveal a complex remodeling of lipids, amino acids, and gut microbiota metabolism. PLoS One *8*, e56564. https://doi.org/10.1371/journal.pone.0056564.
14. Deelen, J., Kettunen, J., Fischer, K., van der Spek, A., Trompet, S., Kastenmüller, G., Boyd, A., Zierer, J., van den Akker, E.B., Ala-Korpela, M., et al. (2019). A metabolic profile of all-cause mortality risk identified in an observational study of 44,168 individuals. Nat. Commun. *10*, 3346. https://doi.org/10.1038/s41467-019-11311-9.
15. Wojczynski, M.K., Jiuan Lin, S., Sebastiani, P., Perls, T.T., Lee, J., Kulminski, A., Newman, A., Zmuda, J.M., Christensen, K., and Province, M.A. (2022). NIA Long Life Family Study: Objectives, Design, and Heritability of Cross-Sectional and Longitudinal Phenotypes. J. Gerontol. A Biol. Sci. Med. Sci. *77*, 717–727. https://doi.org/10.1093/gerona/glab333.
16. Sebastiani, P., Hadley, E.C., Province, M., Christensen, K., Rossi, W., Perls, T.T., and Ash, A.S. (2009). A family longevity selection score: ranking sibships by their longevity, size, and availability for study. Am. J. Epidemiol. *170*, 1555–1562. https://doi.org/10.1093/aje/kwp309.
17. Earls, J.C., Rappaport, N., Heath, L., Wilmanski, T., Magis, A.T., Schork, N.J., Omenn, G.S., Lovejoy, J., Hood, L., and Price, N.D. (2019). Multi-Omic Biological Age Estimation and Its Correlation With Wellness and Disease Phenotypes: A Longitudinal Study of 3,558 Individuals. J. Gerontol. A Biol. Sci. Med. Sci. *74*, S52–S60. https://doi.org/10.1093/gerona/glz220.
18. Ferrucci, L. (2008). The Baltimore Longitudinal Study of Aging (BLSA): a 50-year-long journey and plans for the future. J. Gerontol. A Biol. Sci. Med. Sci. *63*, 1416–1419.
19. Sebastiani, P., and Perls, T.T. (2012). The Genetics of Extreme Longevity: Lessons from the New England Centenarian Study. Front. Genet. *3*, 277. https://doi.org/10.3389/fgene.2012.00277.
20. Xu, Q., Wu, C., Zhu, Q., Gao, R., Lu, J., Valles-Colomer, M., Zhu, J., Yin, F., Huang, L., Ding, L., et al. (2022). Metagenomic and metabolomic remodeling in nonagenarians and centenarians and its association with genetic and socioeconomic factors. Nat. Aging *2*, 438–452.
21. Sebastiani, P., Song, Z., Ellis, D., Tian, Q., Schwaiger-Haber, M., Stancliffe, E., Lustgarten, M.S., Funk, C.C., Baloni, P., Yao, C.H., et al. (2023). A metabolomic signature of the APOE2 allele. Geroscience *45*, 415–426. https://doi.org/10.1007/s11357-022-00646-9.
22. Fahy, E., and Subramaniam, S. (2020). RefMet: a reference nomenclature for metabolomics. Nat. Methods *17*, 1173–1174. https://doi.org/10.1038/s41592-020-01009-y.
23. Verri Hernandes, V., Dordevic, N., Hantikainen, E.M., Sigurdsson, B.B., Smárason, S.V., Garcia-Larsen, V., Gögele, M., Caprioli, G., Bozzolan, I., Pramstaller, P.P., and Rainer, J. (2022). Age, Sex, Body Mass Index, Diet and Menopause Related Metabolites in a Large Homogeneous Alpine Cohort. Metabolites *12*, 205. https://doi.org/10.3390/metabo12030205.
24. Yan, Y., Smith, E., Melander, O., and Ottosson, F. (2022). The association between plasma metabolites and future risk of all-cause mortality. J. Intern. Med. *292*, 804–815. https://doi.org/10.1111/joim.13540.
25. Cheng, S., Larson, M.G., McCabe, E.L., Murabito, J.M., Rhee, E.P., Ho, J.E., Jacques, P.F., Ghorbani, A., Magnusson, M., Souza, A.L., et al. (2015). Distinct metabolomic signatures are associated with longevity in humans. Nat. Commun. *6*, 6791. https://doi.org/10.1038/ncomms7791.
26. de Lange, A.M.G., and Cole, J.H. (2020). Commentary: Correction procedures in brain-age prediction. Neuroimage. Clin. *26*, 102229. https://doi.org/10.1016/j.nicl.2020.102229.
27. Slade, E., Irvin, M.R., Xie, K., Arnett, D.K., Claas, S.A., Kind, T., Fardo, D.W., and Graf, G.A. (2021). Age and sex are associated with the plasma lipidome: findings from the GOLDN study. Lipids Health Dis. *20*, 30. https://doi.org/10.1186/s12944-021-01456-2.
28. Yao, C.H., Liu, G.Y., Wang, R., Moon, S.H., Gross, R.W., and Patti, G.J. (2018). Identifying off-target effects of etomoxir reveals that carnitine palmitoyltransferase I is essential for cancer cell proliferation independent of β-oxidation. PLoS Biol. *16*, e2003782. https://doi.org/10.1371/journal.pbio.2003782.
29. Mair, R.D., Sirich, T.L., Plummer, N.S., and Meyer, T.W. (2018). Characteristics of Colon-Derived Uremic Solutes. Clin. J. Am. Soc. Nephrol. *13*, 1398–1404. https://doi.org/10.2215/cjn.03150318.
30. Yamaguchi, Y., Zampino, M., Moaddel, R., Chen, T.K., Tian, Q., Ferrucci, L., and Semba, R.D. (2021). Plasma metabolites associated with chronic kidney disease and renal function in adults from the Baltimore Longitudinal Study of Aging. Metabolomics *17*, 9. https://doi.org/10.1007/s11306-020-01762-3.
31. Zhu, Y., Dwidar, M., Nemet, I., Buffa, J.A., Sangwan, N., Li, X.S., Anderson, J.T., Romano, K.A., Fu, X., Funabashi, M., et al. (2023). Two distinct gut microbial pathways contribute to meta-organismal production of phenylacetylglutamine with links to cardiovascular disease. Cell Host Microbe *31*, 18–32.e9. https://doi.org/10.1016/j.chom.2022.11.015.
32. Kühn, S., Düzel, S., Colzato, L., Norman, K., Gallinat, J., Brandmaier, A.M., Lindenberger, U., and Widaman, K.F. (2019). Food for thought: association between dietary tyrosine and cognitive performance in younger and older adults. Psychol. Res. *83*, 1097–1106. https://doi.org/10.1007/s00426-017-0957-4.
33. Menni, C., Kastenmüller, G., Petersen, A.K., Bell, J.T., Psatha, M., Tsai, P.C., Gieger, C., Schulz, H., Erte, I., John, S., et al. (2013). Metabolomic markers reveal novel pathways of ageing and early development in human populations. Int. J. Epidemiol. *42*, 1111–1119. https://doi.org/10.1093/ije/dyt094.
34. Johnson, L.C., Parker, K., Aguirre, B.F., Nemkov, T.G., D'Alessandro, A., Johnson, S.A., Seals, D.R., and Martens, C.R. (2019). The plasma metabolome as a predictor of biological aging in humans. Geroscience *41*, 895–906. https://doi.org/10.1007/s11357-019-00123-w.
35. Montoliu, I., Scherer, M., Beguelin, F., DaSilva, L., Mari, D., Salvioli, S., Martin, F.P.J., Capri, M., Bucci, L., Ostan, R., et al. (2014). Serum profiling of healthy aging identifies phospho- and sphingolipid species as markers of human longevity. Aging (Albany NY) *6*, 9–25. https://doi.org/10.18632/aging.100630.
36. Fan, S.Z., Lin, C.S., Wei, Y.W., Yeh, S.R., Tsai, Y.H., Lee, A.C., Lin, W.S., and Wang, P.Y. (2021). Dietary citrate supplementation enhances longevity, metabolic health, and memory performance through promoting ketogenesis. Aging Cell *20*, e13510. https://doi.org/10.1111/acel.13510.
37. Darst, B.F., Koscik, R.L., Hogan, K.J., Johnson, S.C., and Engelman, C.D. (2019). Longitudinal plasma metabolomics of aging and sex. Aging (Albany NY) *11*, 1262–1282. https://doi.org/10.18632/aging.101837.
38. Balasubramanian, R., Paynter, N.P., Giulianini, F., Manson, J.E., Zhao, Y., Chen, J.C., Vitolins, M.Z., Albert, C.A., Clish, C., and Rexrode, K.M. (2020). Metabolomic profiles associated with all-cause mortality in the Women's Health Initiative. Int. J. Epidemiol. *49*, 289–300. https://doi.org/10.1093/ije/dyz211.
39. Longo, V.D., and Anderson, R.M. (2022). Nutrition, longevity and disease: From molecular mechanisms to interventions. Cell *185*, 1455–1470. https://doi.org/10.1016/j.cell.2022.04.002.
40. Carrard, J., Gallart-Ayala, H., Infanger, D., Teav, T., Wagner, J., Knaier, R., Colledge, F., Streese, L., Königstein, K., Hinrichs, T., et al. (2021). Metabolic View on Human Healthspan: A Lipidome-Wide Association Study. Metabolites *11*, 287. https://doi.org/10.3390/metabo11050287.
41. Semba, R.D., Zhang, P., Adelnia, F., Sun, K., Gonzalez-Freire, M., Salem, N., Jr., Brennan, N., Spencer, R.G., Fishbein, K., Khadeer, M., et al. (2019). Low plasma lysophosphatidylcholines are associated with impaired mitochondrial oxidative capacity in adults in the Baltimore Longitudinal Study of Aging. Aging Cell *18*, e12915. https://doi.org/10.1111/acel.12915.

42. Barber, M.N., Risis, S., Yang, C., Meikle, P.J., Staples, M., Febbraio, M.A., and Bruce, C.R. (2012). Plasma lysophosphatidylcholine levels are reduced in obesity and type 2 diabetes. PLoS One *7*, e41456. https://doi.org/10.1371/journal.pone.0041456.

43. Tian, Q., Shardell, M.D., Kuo, P.L., Tanaka, T., Simonsick, E.M., Moaddel, R., Resnick, S.M., and Ferrucci, L. (2023). Plasma metabolomic signatures of dual decline in memory and gait in older adults. Geroscience *45*, 2659–2667. https://doi.org/10.1007/s11357-023-00792-8.

44. Sigruener, A., Kleber, M.E., Heimerl, S., Liebisch, G., Schmitz, G., and Maerz, W. (2014). Glycerophospholipid and sphingolipid species and mortality: the Ludwigshafen Risk and Cardiovascular Health (LURIC) study. PLoS One *9*, e85724. https://doi.org/10.1371/journal.pone.0085724.

45. Pradas, I., Jové, M., Huynh, K., Ingles, M., Borras, C., Mota-Martorell, N., Galo-Licona, J.D., Puig, J., Viña, J., Meikle, P.J., and Pamplona, R. (2022). Long-lived Humans Have a Unique Plasma Sphingolipidome. J. Gerontol. A Biol. Sci. Med. Sci. *77*, 728–735. https://doi.org/10.1093/gerona/glab360.

46. Tofte, N., Suvitaival, T., Ahonen, L., Winther, S.A., Theilade, S., Frimodt-Møller, M., Ahluwalia, T.S., and Rossing, P. (2019). Lipidomic analysis reveals sphingomyelin and phosphatidylcholine species associated with renal impairment and all-cause mortality in type 1 diabetes. Sci. Rep. *9*, 16398. https://doi.org/10.1038/s41598-019-52916-w.

47. Sorgdrager, F.J.H., Naudé, P.J.W., Kema, I.P., Nollen, E.A., and Deyn, P.P.D. (2019). Tryptophan Metabolism in Inflammaging: From Biomarker to Therapeutic Target. Front. Immunol. *10*, 2565. https://doi.org/10.3389/fimmu.2019.02565.

48. Badawy, A.A.-B. (2022). Tryptophan metabolism and disposition in cancer biology and immunotherapy. Biosci. Rep. *42*, BSR20221682. https://doi.org/10.1042/bsr20221682.

49. Beelman, R.B., Kalaras, M.D., Phillips, A.T., and Richie, J.P., Jr. (2020). Is ergothioneine a 'longevity vitamin' limited in the American diet? J. Nutr. Sci. *9*, e52. https://doi.org/10.1017/jns.2020.44.

50. Tian, X., Thorne, J.L., and Moore, J.B. (2023). Ergothioneine: an underrecognised dietary micronutrient required for healthy ageing? Br. J. Nutr. *129*, 104–114. https://doi.org/10.1017/s0007114522003592.

51. Smith, E., Ottosson, F., Hellstrand, S., Ericson, U., Orho-Melander, M., Fernandez, C., and Melander, O. (2020). Ergothioneine is associated with reduced mortality and decreased risk of cardiovascular disease. Heart *106*, 691–697. https://doi.org/10.1136/heartjnl-2019-315485.

52. Bellenguez, C., Küçükali, F., Jansen, I.E., Kleineidam, L., Moreno-Grau, S., Amin, N., Naj, A.C., Campos-Martin, R., Grenier-Boley, B., Andrade, V., et al. (2022). New insights into the genetic etiology of Alzheimer's disease and related dementias. Nat. Genet. *54*, 412–436. https://doi.org/10.1038/s41588-022-01024-z.

53. Said, M.Y., Rodriguez-Niño, A., Post, A., Schutten, J.C., Kieneker, L.M., Gomes-Neto, A.W., van Londen, M., Osté, M.C., Borgonjen-van den Berg, K.J., Nolte, I.M., et al. (2021). Meat intake and risk of mortality and graft failure in kidney transplant recipients. Am. J. Clin. Nutr. *114*, 1505–1517. https://doi.org/10.1093/ajcn/nqab185.

54. Zaworski, J., Bouderlique, E., Anglicheau, D., Duong Van Huyen, J.P., Gnemmi, V., Gibier, J.B., Neugebauer, Y., Haymann, J.P., Bazin, D., Frochot, V., et al. (2020). 1-Methyluric Acid Nephropathy. Kidney Int. Rep. *5*, 737–741. https://doi.org/10.1016/j.ekir.2020.02.1026.

55. Ottosson, F., Smith, E., Fernandez, C., and Melander, O. (2020). Plasma Metabolites Associate with All-Cause Mortality in Individuals with Type 2 Diabetes. Metabolites *10*, 315.

56. Chenze, L., Zhao, M., Xiao, L., Wei, H., Wen, Z., Hu, D., Yu, B., Sun, Y., Gao, J., Shen, X., et al. (2021). Prognostic Value of Elevated Levels of Plasma N-Acetylneuraminic Acid in Patients With Heart Failure. Circ Heart Fail. *14*, e008459. https://doi.org/10.1161/CIRCHEARTFAILURE.121.008459.

57. Ismail, I.T., Fiehn, O., Elfert, A., Helal, M., Salama, I., and El-Said, H. (2020). Sugar Alcohols Have a Key Role in Pathogenesis of Chronic Liver Disease and Hepatocellular Carcinoma in Whole Blood and Liver Tissues. Cancers *12*, 484. https://doi.org/10.3390/cancers12020484.

58. Ament, Z., Bevers, M.B., Wolcott, Z., Kimberly, W.T., and Acharjee, A. (2021). Uric Acid and Gluconic Acid as Predictors of Hyperglycemia and Cytotoxic Injury after Stroke. Transl. Stroke Res. *12*, 293–302. https://doi.org/10.1007/s12975-020-00862-5.

59. Titan, S.M., Venturini, G., Padilha, K., Tavares, G., Zatz, R., Bensenor, I., Lotufo, P.A., Rhee, E.P., Thadhani, R.I., and Pereira, A.C. (2019). Metabolites related to eGFR: Evaluation of candidate molecules for GFR estimation using untargeted metabolomics. Clin. Chim. Acta *489*, 242–248. https://doi.org/10.1016/j.cca.2018.08.037.

60. Wang, Q., Li, L., and Xu, R. (2018). A systems biology approach to predict and characterize human gut microbial metabolites in colorectal cancer. Sci. Rep. *8*, 6225. https://doi.org/10.1038/s41598-018-24315-0.

61. Hejazi, J., Ghanavati, M., Hejazi, E., Poustchi, H., Sepanlou, S.G., Khoshnia, M., Gharavi, A., Sohrabpour, A.A., Sotoudeh, M., Dawsey, S.M., et al. (2020). Habitual dietary intake of flavonoids and all-cause and cause-specific mortality: Golestan cohort study. Nutr. J. *19*, 108. https://doi.org/10.1186/s12937-020-00627-8.

62. Fan, X., Fan, Z., Yang, Z., Huang, T., Tong, Y., Yang, D., Mao, X., and Yang, M. (2022). Flavonoids-Natural Gifts to Promote Health and Longevity. Int. J. Mol. Sci. *23*, 2176. https://doi.org/10.3390/ijms23042176.

63. Melzig, M.F., Putscher, I., Henklein, P., and Haber, H. (2000). In vitro pharmacological activity of the tetrahydroisoquinoline salsolinol present in products from Theobroma cacao L. like cocoa and chocolate. J. Ethnopharmacol. *73*, 153–159. https://doi.org/10.1016/s0378-8741(00)00291-9.

64. Zhong, G.C., Hu, T.Y., Yang, P.F., Peng, Y., Wu, J.J., Sun, W.P., Cheng, L., and Wang, C.R. (2021). Chocolate consumption and all-cause and cause-specific mortality in a US population: a post hoc analysis of the PLCO cancer screening trial. Aging (Albany NY) *13*, 18564–18585. https://doi.org/10.18632/aging.203302.

65. Huang, Y.C., Lee, M.S., and Wahlqvist, M.L. (2012). Prediction of all-cause mortality by B group vitamin status in the elderly. Clin. Nutr. *31*, 191–198. https://doi.org/10.1016/j.clnu.2011.10.010.

66. Huang, J., Weinstein, S.J., Moore, S.C., Derkach, A., Hua, X., Liao, L.M., Gu, F., Mondul, A.M., Sampson, J.N., and Albanes, D. (2018). Serum Metabolomic Profiling of All-Cause Mortality: A Prospective Analysis in the Alpha-Tocopherol, Beta-Carotene Cancer Prevention (ATBC) Study Cohort. Am. J. Epidemiol. *187*, 1721–1732. https://doi.org/10.1093/aje/kwy017.

67. Liu, Z., Li, W., Geng, L., Sun, L., Wang, Q., Yu, Y., Yan, P., Liang, C., Ren, J., Song, M., et al. (2022). Cross-species metabolomic analysis identifies uridine as a potent regeneration promoting factor. Cell Discov. *8*, 6. https://doi.org/10.1038/s41421-021-00361-3.

68. Taha, A.Y. (2020). Linoleic acid–good or bad for the brain? NPJ Sci. Food *4*, 1. https://doi.org/10.1038/s41538-019-0061-9.

69. Kouchiwa, T., Wada, K., Uchiyama, M., Kasezawa, N., Niisato, M., Murakami, H., Fukuyama, K., and Yokogoshi, H. (2012). Age-related changes in serum amino acids concentrations in healthy individuals. Clin. Chem. Lab. Med. *50*, 861–870. https://doi.org/10.1515/cclm-2011-0846.

70. Amick, K.A., Mahapatra, G., Gao, Z., Dewitt, A., Craft, S., Jain, M., and Molina, A.J.A. (2022). Plasma glycocholic acid and linoleic acid identified as potential mediators of mitochondrial bioenergetics in Alzheimer's dementia. Front. Aging Neurosci. *14*, 954090. https://doi.org/10.3389/fnagi.2022.954090.

71. Sun, L., Liang, L., Gao, X., Zhang, H., Yao, P., Hu, Y., Ma, Y., Wang, F., Jin, Q., Li, H., et al. (2016). Early Prediction of Developing Type 2 Diabetes by Plasma Acylcarnitines: A Population-Based Study. Diabetes Care *39*, 1563–1570. https://doi.org/10.2337/dc16-0232.

72. Camacho-Pereira, J., Tarragó, M.G., Chini, C.C., Nin, V., Escande, C., Warner, G.M., Puranik, A.S., Schoon, R.A., Reid, J.M., Galina, A., and

Chini, E.N. (2016). CD38 Dictates Age-Related NAD Decline and Mitochondrial Dysfunction through an SIRT3-Dependent Mechanism. Cell Metabol. *23*, 1127–1139. https://doi.org/10.1016/j.cmet.2016.05.006.

73. Polis, B., Srikanth, K.D., Elliott, E., Gil-Henn, H., and Samson, A.O. (2018). L-Norvaline Reverses Cognitive Decline and Synaptic Loss in a Murine Model of Alzheimer's Disease. Neurotherapeutics *15*, 1036–1054. https://doi.org/10.1007/s13311-018-0669-5.
74. Luo, C., Yang, Q., Liu, Y., Zhou, S., Jiang, J., Reiter, R.J., Bhattacharya, P., Cui, Y., Yang, H., Ma, H., et al. (2019). The multiple protective roles and molecular mechanisms of melatonin and its precursor N-acetylserotonin in targeting brain injury and liver damage and in maintaining bone health. Free Radic. Biol. Med. *130*, 215–233. https://doi.org/10.1016/j.freeradbiomed.2018.10.402.
75. Wei, J., Liu, X., Xiao, W., Lu, J., Guan, L., Fang, Z., Chen, J., Sun, B., Cai, Z., Sun, X., et al. (2023). Phospholipid remodeling and its derivatives are associated with COVID-19 severity. J. Allergy Clin. Immunol. *151*, 1259–1268. https://doi.org/10.1016/j.jaci.2022.11.032.
76. Peña-Bautista, C., Álvarez-Sánchez, L., Roca, M., García-Vallés, L., Baquero, M., and Cháfer-Pericás, C. (2022). Plasma Lipidomics Approach in Early and Specific Alzheimer's Disease Diagnosis. J. Clin. Med. *11*, 5030. https://doi.org/10.3390/jcm11175030.
77. Robinson, O., Chadeau Hyam, M., Karaman, I., Climaco Pinto, R., Ala-Korpela, M., Handakas, E., Fiorito, G., Gao, H., Heard, A., Jarvelin, M.R., et al. (2020). Determinants of accelerated metabolomic and epigenetic aging in a UK cohort. Aging Cell *19*, e13149. https://doi.org/10.1111/acel.13149.
78. Mutlu, A.S., Duffy, J., and Wang, M.C. (2021). Lipid metabolism and lipid signals in aging and longevity. Dev. Cell *56*, 1394–1407. https://doi.org/10.1016/j.devcel.2021.03.034.
79. Cagigas, M.L., Twigg, S.M., and Fontana, L. (2024). Ten tips for promoting cardiometabolic health and slowing cardiovascular aging. Eur. Heart J. *45*, 1094–1097. https://doi.org/10.1093/eurheartj/ehad853.
80. Hostetler, G.L., Ralston, R.A., and Schwartz, S.J. (2017). Flavones: Food Sources, Bioavailability, Metabolism, and Bioactivity. Adv. Nutr. *8*, 423–435. https://doi.org/10.3945/an.116.012948.
81. Ottestad, I., Hassani, S., Borge, G.I., Kohler, A., Vogt, G., Hyötyläinen, T., Orešič, M., Brønner, K.W., Holven, K.B., Ulven, S.M., and Myhrstad, M.C.W. (2012). Fish oil supplementation alters the plasma lipidomic profile and increases long-chain PUFAs of phospholipids and triglycerides in healthy subjects. PLoS One *7*, e42550. https://doi.org/10.1371/journal.pone.0042550.
82. Lankinen, M., Schwab, U., Kolehmainen, M., Paananen, J., Poutanen, K., Mykkänen, H., Seppänen-Laakso, T., Gylling, H., Uusitupa, M., and Orešič, M. (2011). Whole grain products, fish and bilberries alter glucose and lipid metabolism in a randomized, controlled trial: the Sysdimet study. PLoS One *6*, e22646. https://doi.org/10.1371/journal.pone.0022646.
83. Ey, J., Schömig, E., and Taubert, D. (2007). Dietary sources and antioxidant effects of ergothioneine. J. Agric. Food Chem. *55*, 6466–6474. https://doi.org/10.1021/jf071328f.
84. Mahieu, N.G., Spalding, J.L., Gelman, S.J., and Patti, G.J. (2016). Defining and Detecting Complex Peak Relationships in Mass Spectral Data: The Mz.unity Algorithm. Anal. Chem. *88*, 9037–9046. https://doi.org/10.1021/acs.analchem.6b01702.
85. Cho, K., Schwaiger-Haber, M., Naser, F.J., Stancliffe, E., Sindelar, M., and Patti, G.J. (2021). Targeting unique biological signals on the fly to improve MS/MS coverage and identification efficiency in metabolomics. Anal. Chim. Acta *1149*, 338210. https://doi.org/10.1016/j.aca.2021.338210.
86. Stancliffe, E., Schwaiger-Haber, M., Sindelar, M., and Patti, G.J. (2021). DecoID improves identification rates in metabolomics through database-assisted MS/MS deconvolution. Nat. Methods *18*, 779–787. https://doi.org/10.1038/s41592-021-01195-3.
87. Fan, S., Kind, T., Cajka, T., Hazen, S.L., Tang, W.H.W., Kaddurah-Daouk, R., Irvin, M.R., Arnett, D.K., Barupal, D.K., and Fiehn, O. (2019). Systematic Error Removal Using Random Forest for Normalizing Large-Scale Untargeted Lipidomics Data. Anal. Chem. *91*, 3590–3596. https://doi.org/10.1021/acs.analchem.8b05592.
88. Sumner, L.W., Amberg, A., Barrett, D., Beale, M.H., Beger, R., Daykin, C.A., Fan, T.W.M., Fiehn, O., Goodacre, R., Griffin, J.L., et al. (2007). Proposed minimum reporting standards for chemical analysis Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). Metabolomics *3*, 211–221. https://doi.org/10.1007/s11306-007-0082-2.
89. Watanabe, K., Wilmanski, T., Diener, C., Earls, J.C., Zimmer, A., Lincoln, B., Hadlock, J.J., Lovejoy, J.C., Gibbons, S.M., Magis, A.T., et al. (2023). Multiomic signatures of body mass index identify heterogeneous health phenotypes and responses to a lifestyle intervention. Nat. Med. *29*, 996–1008. https://doi.org/10.1038/s41591-023-02248-0.
90. Wilmanski, T., Rappaport, N., Earls, J.C., Magis, A.T., Manor, O., Lovejoy, J., Omenn, G.S., Hood, L., Gibbons, S.M., and Price, N.D. (2019). Blood metabolome predicts gut microbiome α-diversity in humans. Nat. Biotechnol. *37*, 1217–1228. https://doi.org/10.1038/s41587-019-0233-9.
91. Sebastiani, P., Gurinovich, A., Nygaard, M., Sasaki, T., Sweigart, B., Bae, H., Andersen, S.L., Villa, F., Atzmon, G., Christensen, K., et al. (2019). APOE Alleles and Extreme Human Longevity. J. Gerontol. A Biol. Sci. Med. Sci. *74*, 44–51. https://doi.org/10.1093/gerona/gly174.
92. Song, Z., Gurinovich, A., Federico, A., Monti, S., and Sebastiani, P. (2021). nf-gwas-pipeline: A Nextflow Genome-Wide Association Study Pipeline. J. Open Source Softw. *6*, 2957. https://doi.org/10.21105/joss.02957.
93. Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate - a practical and powerful approach to multiple testing. J. Roy. Stat. Soc. B *57*, 289–300.
94. Rutledge, J., Oh, H., and Wyss-Coray, T. (2022). Measuring biological age using omics data. Nat. Rev. Genet. *23*, 715–727. https://doi.org/10.1038/s41576-022-00511-7.
95. Zou, H., and Hastie, T. (2005). Regularization and Variable Selection Via the Elastic Net. J. Roy. Stat. Soc. B Stat. Methodol. *67*, 301–320. https://doi.org/10.1111/j.1467-9868.2005.00503.x.
96. Therneau, T.M., and Grambsch, P.M. (2000). Modeling Surivival Data: Extending the Cox Model (Springer-Verlag).
97. Hawe, J.S., Theis, F.J., and Heinig, M. (2019). Inferring Interaction Networks From Multi-Omics Data. Front. Genet. *10*, 535. https://doi.org/10.3389/fgene.2019.00535.
98. Song, Z., Gunn, S., Monti, S., Peloso, G.M., Liu, C.-T., Lunetta, K., and Sebastiani, P. (2024). Learning Gaussian Graphical Models from Correlated Data. Preprint at bioRxiv. https://doi.org/10.1101/2024.04.03.587948.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological samples		
Human plasma samples	University of Minnesota	NA
Chemicals, peptides, and recombinant proteins		
Acetonitrile LC/MS OPTIMA	Fisher Scientific	Cat#A955 4, CAS 75-05-8
Formic Acid, Optima™ LC/MS Grade	Fisher Scientific	Cat#A117, CAS 64-18-6
Methanol LC/MS OPTIMA	Fisher Scientific	Cat#A456 4, CAS 67-56-1
Methyl *tert*-butyl ether for HPLC, ≥99.8%,	Millipore Sigma	Cat#34875, CAS 1634-04-4
2-propanol (IPA), OPTIMA LC/MS	Fisher Scientific	Cat#A461 4, CAS 67-63-0
Water Optima™ for HPLC	Fisher Scientific	Cat#W7, CAS 7732-18-5
Ammonium hydroxide, eluent additive for LC-MS	Honeywell Fluka	Cat#44273, CAS 133-21-6
Ammonium Bicarbonate, eluent additive for LC-MS	Honeywell Fluka	Cat#40867, CAS 1066-33-7
Ammonium formate, eluent additive for LC-MS	Millipore Sigma	Cat#70221, CAS 540-69-2
Medronic Acid	Millipore Sigma	Cat#64255, CAS 1984-15-2
Deposited data		
Metabolomics, genetic and phenotypic data	This paper	https://eliteportal.synapse.org/Explore/Projects/DetailsPage?shortName=LLFS
Software and algorithms		
Lipid Annotator v1 B1.0.54.0	Agilent Technologies	N/A
MassHunter Workstation LC/MS Data Acquisition v10.1 B10.1.48	Agilent Technologies	N/A
Thermo Scientific Xcalibur v4.2.28.14, AcquireX workflow	Thermo Fisher Scientific	–
Python v3.7	python	RRID:SCR_008394
Skyline 64-bit v20.2	MacCoss Lab Software	RRID:SCR_014080
Other		
Agilent 6545 Q-TOF (quadrupole time-of-flight mass spectrometer)	Agilent Technologies	N/A
Agilent 1290 Infinity II UHPLC	Agilent Technologies	N/A

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Human subjects

Long Life Family Study (LLFS)

The LLFS is a family-based, longitudinal study of healthy aging and longevity that enrolled 4,953 family members from 539 families selected for familial longevity at three American field centers (Boston, Pittsburgh and New York), and a Danish field center between 2006 and 2009.[15,16] Study participants have been comprehensively phenotyped using a combination of three in-person visits, questionnaire-based assessments, and annual follow-ups. Medications were recoded into ATC codes and medications for heart disease, lipid lowering, Type 2 diabetes (T2D), and hypertension were defined based on ATC codes for the common medications used for these disorders in conjunction with physicians and pharmacists. LLFS participants included 1496 females and 1268 males with an average age of 74 years at enrollment, and range 24–110 years. All participants included in this analysis identified as White. All subjects provided informed consent, and genetic and phenotypic data are available via dbGaP (dbGaP Study Accession: phs000397.v1.p1).

New England Centenarian Study (NECS)

We included for replication the metabolomic profiles of 195 NECS participants comprising centenarians, centenarians' offspring and unrelated controls.[19] We selected participants who were at least one year away from major clinical events, not taking major medications and survived at least one year after the blood draw. Trained phlebotomists collected blood samples from NECS participants

at enrollment, and transferred the samples to the Boston University Molecular Genetic Core biobank for DNA, serum, and plasma extractions. Samples were maintained at -80^0C. NECS participants included 121 females and 74 males with an average age of 88 years at enrollment, and range 56–115 years. Among these participants, 13 identified as African American and the others as White. The NECS protocol was approved by the Boston Medical Center Institutional Review Board. Participants provided written informed consent at enrollment.

Arivale

The Arivale cohort includes adults who self-enrolled in a consumer-facing wellness program (Arivale, Inc. 2015–2019).[17] The participants were required to be over the age of 18 and not pregnant, with no additional screening of participants. Certified phlebotomists collected blood samples from fasting participants. We included a subset of 634 participants from the Arivale cohort with plasma metabolomic and other covariate data, including 402 females and 232 males, with an average age of 49 years (range 22–87 years). Among participants included in this analysis, 82% identified as White, 1.9% as Black or African American, 5.0% as Asian, 5.2% as Hispanic, and 5.9% as other groups. The current study was conducted with de-identified data of the participants who had consented to the sharing of their deidentified data. Procedures were approved by the Western Institutional Review Board.

Baltimore Longitudinal study of aging (BLSA)

The BLSA is a longitudinal study of community-dwelling adults that continuously enrolls participants free of cognitive impairment, functional limitations, chronic diseases, and cancer for the previous 10 years.[18] Details of blood collections have been previously described.[9,30] We selected a subset of 1,135 BLSA participants with plasma metabolomics comprising 545 males and 590 females. The mean age of this sample was 65 years of age (SD = 14.5), ranging from 22 to 96 years. Among these participants, 67% (n = 757) identified as White, 28% (n = 315) as African Americans, and 1% (n = 12) as Asians. To examine metabolite associations with mortality, we analyzed data from a subset of 926 participants aged 50 years or older with a mean follow-up of 9.3 (SD = 3.5) years (ranging 0.6 to 16 years). Among this subset, 68% (n = 634) identified as White, 25.8% (n = 239) as African Americans, 1.3% (n = 12) as Asians. The BLSA protocol was approved by the Institutional Review Board of the National Institutes of Health. Participants provided written informed consent at each BLSA visit.

Xu et al. Cohort (XU)

Cohort description. Xu et al. described multi-omic profiles for 116 nonagenarians and centenarians and 232 of their offspring from Han Chinese longevous families.[20] **Metabolomic data.** Data for 821 lipid and polar metabolites were generated using non-targeted metabolomic analysis and 121 metabolites from the 408 target list (Table S1) were found here using RefMet nomenclature. Additional experimental details are described in the original study, and individual-level data are available in Table S6 of their supplemental information.[20]

METHOD DETAILS

Metabolomic library preparation and mass spectrometry analysis

LLFS

We included data for 408 metabolites (188 lipids and 220 polar metabolites) that were profiled in the plasma of 2,764 LLFS participants collected at enrollment and, for a subset of 1,246 participants, in a second blood draw approximately 8 years later. Our experimental workflows have been described in detail previously.[5,21] In brief, we isolated lipid and polar metabolites from plasma using solid-phase extraction kits. Prior to mass spectrometry analysis, samples were separated using reversed-phase chromatography and hydrophilic interaction liquid chromatography (HILIC) for lipid and polar metabolites, respectively. Mass spectrometric detection was performed with an Agilent 6545 quadrupole time-of-flight mass spectrometer at Washington University in St Louis. We processed the data with a combination of XCMS, DecoID, and Skyline to facilitate removal of background, annotation of adducts, and compound identification.[84–86] We analyzed lipid and polar metabolites in batches of approximately 90 samples each and normalized data by using a random forest-based approach for batch correction,[87] which we previously found to be the optimal approach for this data.[5] We focused on 408 metabolites that we identified from the plasma samples based on retention time and fragmentation pattern matches to both in-house and public databases formed from analysis of authentic standards. For retention time matches, a 30 s tolerance between reference and observed retention times was allowed. For MS/MS matches, a normalized dot-product score of >80 was required. All identifications were manually reviewed after applying the cutoffs listed previously to ensure accurate matches. According to the Metabolomics Standard Initiative,[88] these 408 identifications correspond to high-confidence Level 1 and Level 2 identifications. Three acylcarnitines (decanoylcarnitine, butyrylcarnitine, and octanoylcarnitine) were detected among both the lipid and the polar metabolites. We included these in the analysis as independent validation of the results. We named the lipid and polar metabolites using standardized names from RefMet, version of 07/2023[22] (Table S1).

NECS

Metabolon, Inc. (North Carolina, USA) generated the serum metabolomic profiles that included a total of 1,495 chemicals with 1,213 compounds of known identity and 282 compounds of unknown structural identity. A detailed description of sample preparation and liquid chromatography-tandem/mass spectrometry (LC/MS) was described previously.[21] After renaming the lipid and polar metabolites by using standardized nomenclature from RefMet,[22] 150 metabolites from the target list of 408 (Table S1) were found here.

Arivale

Metabolon, Inc. (North Carolina) generated plasma metabolomic profiles by using LC/MS. Sample handling, quality control, data extraction, biochemical identification, data curation, quantification, data normalizations procedures, and batch correction have been previously described.[89,90] After renaming the lipid and polar metabolites using standardized names from RefMet,[22] 120 metabolites from the target list of 408 (Table S1) were found here.

BLSA

Metabolite concentrations were measured via the Biocrates p500 kit with LC/MS (Biocrates Life Sciences AG, Innsbruck, Austria), and following the manufacturer's protocol for a 5500 QTrap (Sciex, Framingham, MA, USA). Using RefMet nomenclature, 99 metabolites from the target list of 408 (Table S1) were found here.

XU

Data for 821 lipid and polar metabolites were generated using non-targeted metabolomic analysis and 121 metabolites from the target list of 408 (Table S1) were found here, using RefMet nomenclature.

Genetic data

LLFS genetic data

We determined APOE alleles from SNPs rs7412 and rs429358 (e2: rs7412 = T, rs429358 = T; e3: rs7412 = C, rs429358 = T; e4: rs7412 = C, rs429358 = C) that were genotyped using real-time PCR or imputed using IMPUTE2 in participants with genome-wide genotype data as described in.[91]

QUANTIFICATION AND STATISTICAL ANALYSIS

Statistical analysis

LLFS

Quality control. We replaced missing metabolite values with half of the minimum detected value. We transformed metabolites using the natural log transformation, followed by principal component analysis of the metabolites to identify 23 outlier samples (samples that did not cluster with most of the data) that we removed from the dataset. We then searched for metabolite-specific outliers that we defined as values exceeding the mean by more than four standard deviations and replaced the outliers with the mean value. Approximately 50% of metabolites had outliers and the average proportion of outliers per metabolite was 0.1%. All subsequent analyses used natural log transformed metabolite intensities.

Age signature. To identify metabolites that correlate with age at blood draw, we analyzed each log-transformed metabolite using a mixed-effect linear regression model with a random intercept term to adjust for within-family correlation, age at blood draw as the independent predictor, and adjusted the analysis for sex, education, whether participants were enrolled in Denmark or in the USA, medications for type 2 diabetes (T2D), dyslipidemia, heart disease, and hypertension (HTN). We included four genome-wide principal components to adjust for genetic heterogeneity. We modeled the variance-covariance matrix of the random intercept using the genetic relation matrix that we estimated using whole-genome sequence data and the PC-relate R package implemented in our Nextflow-based pipeline for genetic analysis.[92] We calculated adjusted *p*-values using a Benjamini-Hochberg correction,[93] and selected the metabolites that were associated with age at a 5% false discovery rate (FDR) as significant. We compared the distributions of super-classes of metabolites that were positively or negatively correlated with age relative to the overall distribution of the 408 metabolites across super-classes using the Kolmogorov-Smirnov test.

In this analysis, we excluded data from 84 EL individuals who had reached the oldest 1% survival age for their sex and birth year cohort at the time of the first blood draw. All the results are in Table S1, and Figures 1A–1C and we include estimates of the regression coefficients and effects expressed as fold changes (exp(regression coefficient)).

Aging signature. We analyzed the longitudinal fold changes of the metabolites in a subset of LLFS participants with two repeated blood draws using a linear regression model of the repeated, log-transformed, metabolite measurements y_{ij}. The regression analysis included Age_{i1}: the age at first assessment and $(Age_{ij} - Age_{i1})$: the time between the two assessments of participant *i*

$$y_{ij} = \beta_0 + \beta_c Age_{i1} + \beta_l (Age_{ij} - Age_{i1}) + \sum \beta_k X_{ik} \ldots j = 1,2$$

The regression coefficient β_l represents the longitudinal effect of aging, or the change rate. We adjusted the analysis for sex, education, whether participants were enrolled in Denmark or in the USA, medication for T2D, lipids, heart disease, HTN, and four genome-wide principal components to adjust for genetic heterogeneity (these covariates are described by the index *k* in the formula above). We estimated the regression coefficients and standard error using generalized estimating equations, with an exchangeable correlation structure to accommodate the repeated measures and selected the metabolites that correlated with the age difference at a 5% FDR. We compared the cross-sectional and longitudinal effects of age in Figure 1F and calculated Pearson correlation coefficient after removing Lauryl sulfate, PS 38:4, inosine, and hypoxanthine that had a substantially different aging rate. We used the gee package in R for this analysis.

Extreme longevity signature. We identified 84 LLFS participants who had reached the oldest 1% survival age for their sex and birth year cohort (EL) at the time of the first blood draw, 924 offspring of LLFS participants, and 368 spouse controls who were <85 years at

the time of the first blood draw. We used the grouping EL, offspring and spouse controls as the independent predictor in a mixed effect linear regression analysis of each metabolite, using the same mixed effect model and covariate adjustment as in the age analysis. We similarly compared the distribution of super classes of metabolites that were significantly increased or decreased in the EL group at a 5% FDR relative to the main classes represented by the 408 metabolites using the Kolmogorov-Smirnov test (KS-test).

Survival signature. Between enrollment and November 2022, we observed 54% mortality in 2,367 LLFS subjects with metabolomic data and no missing values in covariates. We used these data to identify metabolites that predict the mortality risk. To identify metabolites that predict survival, we analyzed time to death after enrollment using Cox proportional hazards regression models, with each standardized log-transformed metabolite as an independent predictor and adjusting for age at enrollment and the same additional covariates as in the cross-sectional age analyses. Age was censored at last contact for individuals alive in November 2022. The proportional hazards assumption was assessed by graphical inspection of the data. We used the sandwich estimator implemented in the R function *coxph* of the *survival* package to adjust for family correlation. The "metab effect" estimated with this analysis represents the adjusted effect of one standard deviation from the mean of the log-transformed metabolite values on the hazard for mortality in the LLFS. We calculated adjusted *p*-values using the Benjamini-Hochberg correction,[93] and selected the metabolites that were correlated with the hazard for mortality at a 5% FDR as significant. We compared the distribution of super-classes of metabolites that were positively or negatively associated with the hazard for mortality relative to the overall distribution of super-classes represented by the 408 metabolites, using the Kolmogorov-Smirnov test.

Genetic analysis. We investigated the association between each lipid and polar metabolite and *APOE* genotype groups (E2 = e2e2, e2e3, E3 = e3e3, E4 = e3e4, e4e4) using a mixed-effects linear regression model for each metabolite, the APOE genotype group as independent predictor, and the same covariate adjustment used in the other models. All analyses were conducted using custom-made scripts in the statistical software R.

Generation of signatures

We grouped metabolites and lipids based on their significant, directional associations with age, EL, and mortality risk. We ignored borderline results with 0.05 < FDR <0.20.

Metabolomic clock

We used elastic net regularized linear regression to generate a metabolomic clock using the set of 308 metabolites associated with chronological age.[94,95] We standardized the 308 log-transformed metabolite data and adjusted the regression by sex, years of education, whether participants were enrolled in Denmark or in the USA, medications for type 2 diabetes (T2D), dyslipidemia, heart disease, and hypertension (HTN), and four genome-wide principal components. We divided the dataset into 80% training and 20% test set, and used the training set to build the clock and validate it. We determined the penalty parameters α and λ through 10-fold cross-validation to minimize the squared error. Figure S5A, plots the predicted age (=metabolomic age) versus chronological age using the initial clock in the full dataset. While the plot shows 89% correlation between predicted and actual chronological age, the clock exhibited the bias reported in other studies: it predicted younger individuals to be, on average, older than their chronological age, and older individuals to be, on average, younger than their chronological age.[26] Therefore, we adopted the bias-correction reparameterization suggested in ref. [26]: we regressed the predicted age against chronological age using the linear regression

$$predicted\ age\ =\ \beta + \alpha\ chronological\ age$$

to estimate the intercept and slope parameters β, α and calculated the bias-corrected clock as

$$bias\ -\ corrected\ predicted\ age\ = \frac{predicted\ age\ -\ \beta}{\alpha}$$

This transformation indeed removed the initial bias as shown in Figures 6A and S5B. We calculated age acceleration as the difference between the bias-corrected predicted and chronological age and the plot of age acceleration against predicted age showed no bias (Figure S5C). To verify this, we estimated the dependency of age acceleration on predicted age using linear regression and the analysis shows no association (see inset in Figure S5C). To assess the impact of age acceleration on mortality risk, we fit a Cox proportional hazards regression for time to death (censored at last contact for alive participants), stratifying for chronological age in decades and adjusting for sex and years of education. We assessed the bias of the fitted model by analyzing martingale residuals by strata of age acceleration (Figure S5, panels D, E). The martingale residuals are the difference between observed and predicted deaths.[96] A good fit model would produce martingale residuals that are evenly distributed around 0, while an enrichment of negative residuals shows that the model overestimates mortality and an enrichment of positive residuals shows that the model underestimates mortality. All these analyses were conducted using the R packages 'caret', 'glmnet', 'survival' for elastic nets and time to event analyses. While the martingale biased-corrected age acceleration showed no bias, using the uncorrected age acceleration in the same model produces a bias in the extreme old ages that is consistent with the overly optimistic age prediction.

As a validation, we calculated the predicted metabolomic age of LLFS participants based on the values of metabolites at visit 2 and compared it to the chronological age (Figure 6A).

Partial correlation networks of age-associated metabolites

We generated partial correlation networks of the age-associated metabolites. Partial correlation networks are probabilistic networks in which the nodes represent random variables (in this case the metabolites) that follow normal distributions; edges represent linear relations between nodes, and absence of an edge between two nodes represents independence of the two variables, conditionally

on all the other variables in the network. The partial correlation between two nodes is interpreted as their correlation, adjusted for the values of all the other variables in the network. There are well-established algorithms to generate partial correlation networks based on the analysis of the partial correlation matrix estimated from a dataset of independent observations.[97] We developed a bootstrap-based method to learn structure and parameters of partial correlation networks from correlated data in the LLFS. In addition to accounting for correlations on the observations, this method also allows the user to set a threshold on the magnitude of the partial correlations to retain in the model using significance testing. We used extensive simulations of correlated data from family-based studies to show that the bootstrap method does not inflate the Type I error while retaining statistical power compared to alternative solutions.[98] We applied this algorithm to infer the networks of dependency among the 307 metabolites, after log-transformation of the data and removal of one metabolite with a skewed distribution. We generated networks by linking the pairs with absolute partial correlation that was significantly greater than thresholds varying between 0 and 0.30 and employing 1,000x bootstrap resampling for each threshold. All the results are in Figures S6 and S7. We analyzed the network topology and generated interactive graphical displays using the visNetworks package in R that are distributed as html files.

Replication

NECS

Quality control. We transformed metabolites using the natural log transformation and visually inspected the data using principal component analysis. There were 701 metabolites with less than 20% missing values that we included in the final analysis.

EL signature. The NECS data included centenarians, their offspring and controls and we used this group definition as independent predictor in the linear regression model of each metabolite, adjusting for sex and education. We selected metabolites that were significantly different in centenarians at 5% FDR. We conducted the analysis using custom-made scripts in the statistical software R.

Arivale

Quality control. We used an in-house pipeline to batch-correct the metabolomic data that we natural log-transformed to remove right skewness and outliers. We removed metabolites that were missing in more than 8% of the subjects, and subjects with less than 10 detected metabolites. The final filtered metabolomic dataset consisted of 722 metabolites in 634 participants.

Age signature. To analyze the association between each individual metabolite and age at blood draw, we used a linear regression model in which age at blood draw was the independent predictor, and adjusted for sex, body mass index (BMI), statin use, and education. Education was coded according to the following categories: 1: High school diploma or GED; 2: Some college and/or technical school; 3: Associate and/or bachelor's degree; and 4: Master's and/or doctorate and/or professional degree. We calculated adjusted *p*-values using Benjamini-Hochberg correction.[93] We conducted the analysis using the statsmodels package version 0.13.0 in Python version 3.9.7.

BLSA

Quality control. We excluded metabolites with concentrations below the limit of detection (LOD) in more than 20% of participants and imputed the missing values of the remaining metabolites using half of the LOD. We transformed the metabolites using the natural log transformation, and then standardized to Z-scores.

Age signature. We used linear regression to examine the associations of 460 metabolites with age at blood draw, adjusted for sex, race (coded as Black vs. non-Black), education, and batch. *Survival analysis signature*. We used Cox proportional hazards regression to examine the association of each standardized metabolite with the hazard for mortality, adjusted for age, sex, and race, and used a 5% FDR for significance threshold. We performed all analyses using R version 4.0.2 (Boston, MA).

XU

Quality control. We transformed metabolites using the natural log transformation and visually inspected the data using principal component analysis. No outlier samples were detected in the data. There were 666 metabolites with less than 20% missing values that we included in the final analysis.

EL signature. This dataset included trios comprising one centenarian and two of their offspring. We used this group definition as the independent predictor in the linear regression model for each metabolite, adjusting for sex, BMI, medications for T2D, high cholesterol, heart disease, and hypertension. Given the family-based design, we estimated the regression coefficients using generalized estimating equations, with clustering based on family membership and exchangeable correlation structure. We selected metabolites that were significantly different in centenarians at 5% FDR. We conducted the analysis using custom-made scripts in the statistical software R.