

RESEARCH

Open Access



Breaking down causes, consequences, and mediating effects of telomere length variation on human health

Samuel Moix^{1,2*} , Marie C Sadler^{1,2,3}, Zoltán Kutalik^{1,2,3*†} and Chiara Auwerx^{1,2,3,4*†}

[†]Zoltán Kutalik and Chiara Auwerx jointly supervised this work.

*Correspondence: moixsamuel@gmail.com; zoltan.kutalik@unil.ch; cauwerx@gmail.com

¹ Department of Computational Biology, UNIL, Lausanne 1015, Switzerland

² Swiss Institute of Bioinformatics, Lausanne 1015, Switzerland

³ University Center for Primary Care and Public Health, Lausanne 1015, Switzerland

⁴ Center for Integrative Genetics, UNIL, Lausanne 1015, Switzerland

Abstract

Background: Telomeres form repeated DNA sequences at the ends of chromosomes, which shorten with each cell division. Yet, factors modulating telomere attrition and the health consequences thereof are not fully understood. To address this, we leveraged data from 326,363 unrelated UK Biobank participants of European ancestry.

Results: Using linear regression and bidirectional univariable and multivariable Mendelian randomization (MR), we elucidate the relationships between leukocyte telomere length (LTL) and 142 complex traits, including diseases, biomarkers, and lifestyle factors. We confirm that telomeres shorten with age and show a stronger decline in males than in females, with these factors contributing to the majority of the 5.4% of LTL variance explained by the phenome. MR reveals 23 traits modulating LTL. Smoking cessation and high educational attainment associate with longer LTL, while weekly alcohol intake, body mass index, urate levels, and female reproductive events, such as childbirth, associate with shorter LTL. We also identify 24 traits affected by LTL, with risk for cardiovascular, pulmonary, and some autoimmune diseases being increased by short LTL, while longer LTL increased risk for other autoimmune conditions and cancers. Through multivariable MR, we show that LTL may partially mediate the impact of educational attainment, body mass index, and female age at childbirth on proxied lifespan.

Conclusions: Our study sheds light on the modulators, consequences, and the mediatory role of telomeres, portraying an intricate relationship between LTL, diseases, lifestyle, and socio-economic factors.

Keywords: Telomeres, Mendelian randomization, UK Biobank, Complex traits, Lifespan, Aging, Lifestyle, Female reproduction, Mediation

Background

Aging represents a leading risk factor for diseases such as cancer, cardiovascular diseases, and neurodegeneration [1]. Chronological age fails to account for individual differences in aging rates and vulnerability to diseases [2]. Biological age intends to address



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

this limitation by reflecting the physiological state of an individual and accounting for variations in cellular and tissue health. Several biomarkers can be used to estimate biological age, with DNA methylation being particularly popular as it can be measured across different tissues, and is sensitive to both disease states and environmental factors [3–5]. However, given the complex nature of the aging process, additional biomarkers beyond DNA methylation are required to fully understand the underlying causes and mechanisms of aging and age-related diseases [6].

One such biomarker is telomere length. Telomeres are DNA repeats at chromosome ends that act as protective caps against genomic degradation. As organisms age, they undergo an increasing number of cell divisions, leading to an incremental decrease in telomere length. Acting as mitotic clocks, telomeres shorten until they reach a critical length, triggering cellular senescence and/or apoptosis [7]. Consequently, shorter telomeres have been associated with lifestyle factors, including smoking [8], reduced physical activity [9], high processed meat and low fruit consumption [10, 11], as well as a wide range of diseases, from pulmonary [12], renal [13], and metabolic [14] disorders to cancer [15, 16]. Paradoxically, longer telomeres have also been associated with poor health outcomes, especially cancers [17]. However, most studies so far were limited in the number of studied traits, relied on small sample sizes, and did not probe the directionality of the established associations.

Recently, efforts to assess leukocyte telomere length (LTL) in large population biobanks have allowed comprehensive exploration of its relationships with lifestyle factors and health outcomes. Performing an LTL phenome-wide association study in 62,271 participants from the biobank of Vanderbilt University Medical Center (BioVU) and Marshfield Clinic's Personalized Medicine Research Project (PMRP), Allaire et al. identified associations with 67 phenotypes and showed that both shorter and longer telomeres associated with increased mortality [18]. Release of LTL measurements for ~500,000 UK Biobank (UKBB) participants [19] and the companion first large-scale telomere length genome-wide association study (GWAS) [20] prompted investigation of the impact of LTL on hundreds of traits using phenome-wide Mendelian randomization (MR) [20, 21]. These studies showed that longer LTL increases risk for neoplastic and genitourinary diseases while shorter LTL increases risk for respiratory, digestive, and cardiovascular disorders [20, 21], with about 40% of these associations confirmed when using FinnGen disease association summary statistics [21].

Our study builds on this body of work by dissecting observational correlations between LTL and 142 traits into causes and consequences through a robust bidirectional MR causal framework (Fig. 1). Additionally, we used multivariable Mendelian randomization (MVMR) to disentangle the interplay between LTL and various traits, with a particular focus on the mediating role of LTL in longevity. Together, we identify traits influencing LTL, and how in turn the latter impacts the human phenome, contributing to a deeper understanding of telomere biology and its relation to health.

Results

Age and sex are the main predictors of LTL variability

Consistent with previous research [22], LTL significantly associated with both age ($\hat{\beta} = -0.023$; $p < 2.2e-308$) and sex ($\hat{\beta} = 0.091$; $p < 2.2e-308$), with a stronger

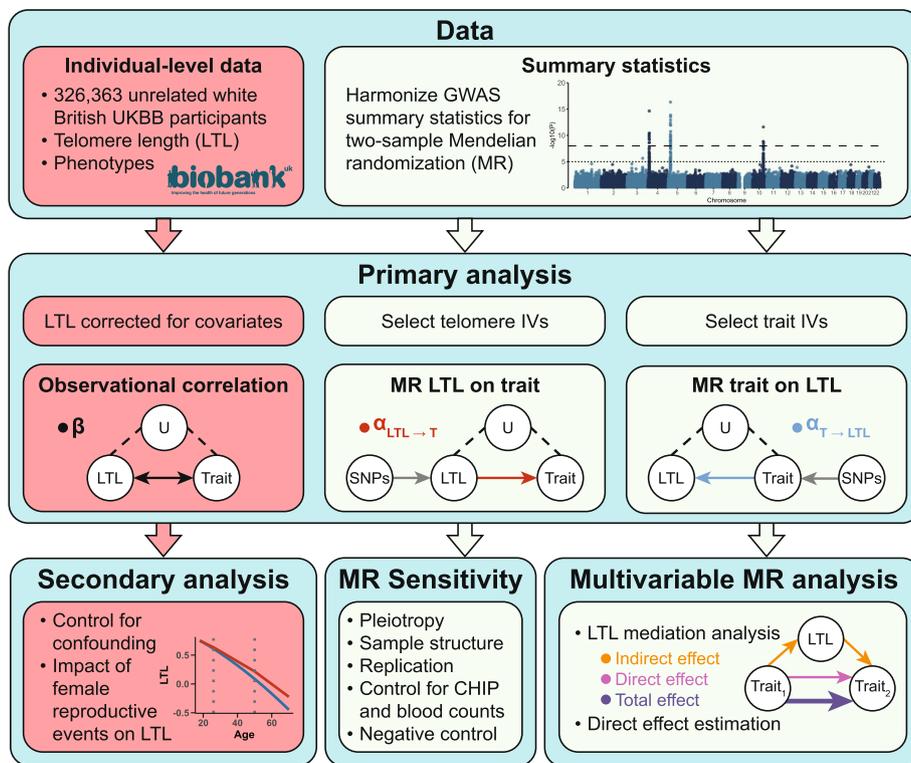


Fig. 1 Schematic representation of the study's workflow. Red and light green boxes denote steps using individual-level phenotypic data from the UK Biobank and genome-wide association studies (GWAS) summary statistics, respectively. Top: Data extraction process. Middle: Analyses focused on leukocyte telomere length (LTL) trait relationships including observational correlation (β ; black), Mendelian randomization (MR) to assess the impact of LTL on traits ($\alpha_{LTL \rightarrow T}$; red), and MR to assess the impact of traits on LTL ($\alpha_{T \rightarrow LTL}$; blue). LTL covariates comprise age, age², genotyping array, sex, and the interaction of the latter with the priors. U = unmeasured confounding factors; IVs = instrumental variables, i.e., genetic variants with genome-wide significant association to the considered trait. SNPs = single nucleotide polymorphisms. Bottom: Follow-up analyses include exploring the association of female reproductive events with LTL, sensitivity analyses (i.e., implementation of seven complementary MR methods, replication using independent LTL summary statistics [18], controlling for confounding by blood-related traits, and evaluating the MR effect of LTL on sex as a negative control), and perform mediation analysis through multivariable MR. Note that for mediation analyses, both exposure and mediator are instrumented. CHIP = clonal hematopoiesis of indeterminate potential

($p_{diff} = 1.4e-25$) decline over time in males ($\hat{\beta}_{males} = -0.025$) than in females ($\hat{\beta}_{females} = -0.021$) (Additional file 1: Fig. S1). To further explore factors contributing to LTL variability, we included 80 traits (Additional file 2: Table S1) with < 7% missingness rates as predictor variables in a Lasso regression model. Traits retained included age, sex, educational attainment (EA), waist-to-hip ratio (WHR), insulin-like growth factor 1 (IGF-1), urate, and cystatin C levels, along with four blood parameters (Table 1; see the “Predictors of LTL variability” section). Among these, LTL was found to be positively associated with female sex, higher EA, and higher IGF-1 levels, while it negatively correlated with the remaining traits. Age and sex accounted for 4.33% of the observed variance in LTL. Incorporating the nine additional above-mentioned traits increased the explained variance to 5.39%. Repeating the analysis with missingness rate thresholds of 5% and 10% retained twelve and

Table 1 Major associations with LTL. Effect sizes with 95% confidence intervals (CI) from a joint linear regression model for traits with less than 7% missing data retained as significant LTL predictors by Lasso analysis. EA = educational attainment; IGF-1 = insulin-like growth factor 1; WHR = waist-to-hip ratio; MCH = mean corpuscular hemoglobin

Trait	Estimate	Lower 95% CI	Upper 95% CI	P-value
Female	0.071	0.066	0.077	2.33×10^{-160}
EA	0.042	0.038	0.046	1.95×10^{-111}
IGF-1	0.028	0.024	0.032	3.57×10^{-49}
WHR	-0.002	-0.007	0.004	5.43×10^{-1}
Urate	-0.009	-0.014	-0.004	1.16×10^{-4}
Monocyte count	-0.017	-0.021	-0.013	4.89×10^{-17}
Eosinophil count	-0.021	-0.025	-0.017	1.00×10^{-28}
Cystatin C	-0.026	-0.030	-0.022	4.93×10^{-33}
Lymphocyte count	-0.039	-0.043	-0.035	7.43×10^{-87}
MCH	-0.058	-0.062	-0.055	4.98×10^{-223}
Age	-0.155	-0.159	-0.151	$< 2.2 \times 10^{-308}$

seven traits in addition to age and sex, which together explain 5.42% and 5.36% of variability in LTL, respectively, confirming the limited predictive power of the phenotype over LTL variability (Additional file 2: Table S1).

LTL broadly associates with complex traits

Considering the strong correlation between age and sex with LTL, we adjusted LTL for age, age², genotyping array, sex, and the interaction of the latter with the priors. We then regressed adjusted LTL (hereafter simply referred to as LTL) on 166 traits through linear regression, identifying 100 significant associations ($p < 0.05/141 = 3.5e-4$; Additional file 2: Table S2). We observed a negative association between the disease burden and LTL ($\hat{\beta} = -0.027$; $p = 1.2e-52$), suggesting that LTL acts as a global health indicator. The largest effect sizes were noted for father's ($\hat{\beta} = 0.094$; $p = 4.4e-144$) and mother's ($\hat{\beta} = 0.088$; $p = 8.5e-216$) age at birth, which positively associated with LTL (Fig. 2a). To test confounding by socio-economic status (SES), we jointly modeled LTL as a function of both parental ages at birth, participant's age, sex, and age-sex interaction, and EA. We found that the associations with parental ages at birth were independent of the participant's education level (Additional file 1: Fig. S2), which likely echoes parental EA [23] and indirectly affects parental age at birth. This suggests that the association is likely not confounded by SES and is genuinely driven by older parental age at birth. As these traits cannot be genetically instrumented, MR is not applicable. As such, the observational nature of our analysis prevents us from further dissecting the effects of paternal versus maternal age at birth on LTL. Next, for the 142 traits with available GWAS summary statistics and at least two instrumental variables (IVs), we inferred bidirectional causal relationships through univariable MR, identifying 23 significant causal effects of traits on LTL ($\hat{\alpha}_{T \rightarrow LTL}$) and 24 significant effects of LTL on traits ($\hat{\alpha}_{LTL \rightarrow T}$) ($p < 0.05/141 = 3.5e-4$; Fig. 2 and Additional file 2: Table S2).

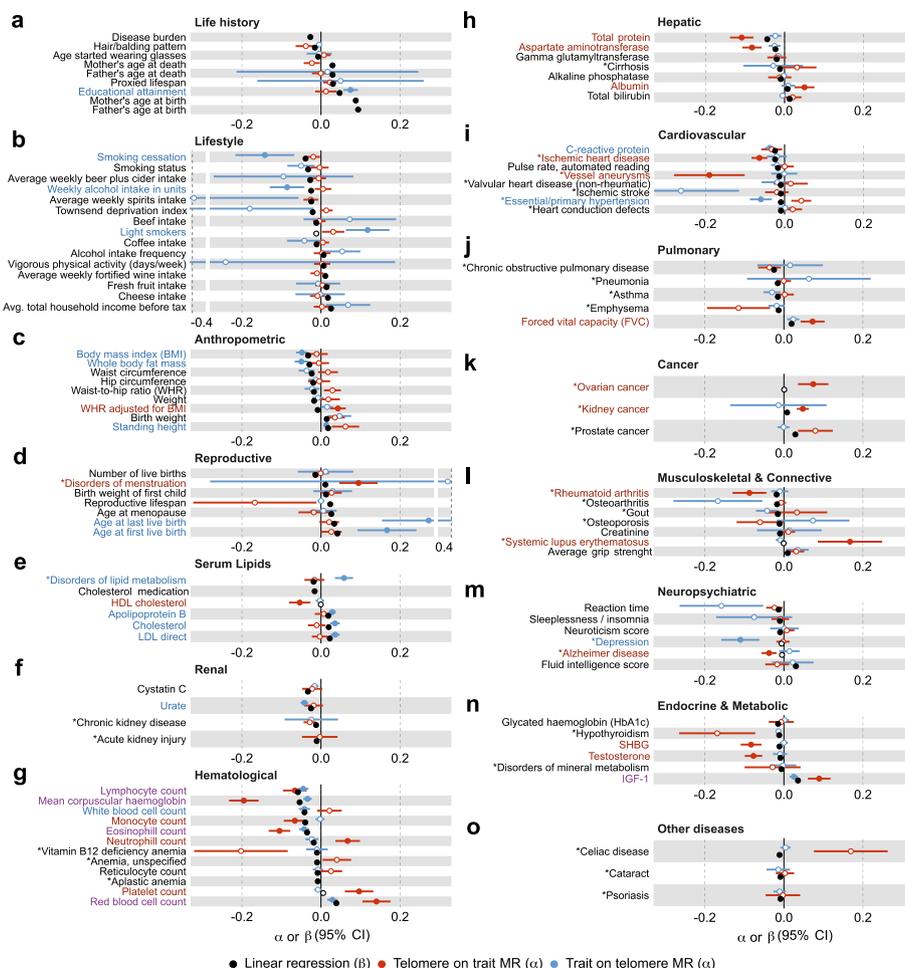


Fig. 2 Observational and causal associations between traits and LTL. Estimates (x -axis) with 95% confidence intervals (CI) for traits (y -axis) with at least one strictly significant ($p < 0.05/141 = 3.5e-4$) association with LTL across the observational correlation (linear regression; β ; black) and inverse-variance weighted (IVW) Mendelian randomization (MR) estimates of LTL on trait (α ; red) and trait on LTL (α ; blue) are shown. Strictly significant effects are shown as full circles; otherwise as empty circles. Traits are colored according to their MR effects, with red, blue, or purple indicating a significant LTL to trait, trait to LTL, or bidirectional effect. For diseases (*), one standard deviation (SD) change in LTL corresponds to one $\log(OR)$ change, implying a scale of $SD_{LTL} / \log(OR)$ for the effects of diseases on LTL, and $\log(OR) / SD_{LTL}$ for the effect of LTL on the disease, so that observational effects and MR effects are not directly comparable (Additional file 2: Table S2)

Sensitivity analyses

To ensure the robustness and reliability of our results, we gauged the reliability of inverse-variance weighted (IVW) significant associations through several approaches (Fig. 1). First, we estimated robustness towards MR assumption violation. We applied four additional MR methods implemented in the TwoSampleMR package (MR Egger, simple mode, weighted median, and weighted mode), as well as MR-PRESSO (Additional file 1: Figs. S3–4). To mitigate pleiotropy bias, we further implemented a custom approach based on Steiger filtering that requires IVs to have a stronger association with the exposure (i.e., LTL) than with any of the other 152 traits analyzed through MR (see the “Methods” section and Additional file 1: Fig. S5). To determine the impact of sample structure, and more notably sample overlap, we ran MR-APSS (which also accounts

for pleiotropy) and replicated IVW results using independent LTL summary statistics from BioVU/PMRP [18]. Estimates were globally consistent across all methods (Additional file 1: Fig. S3–4). Regarding replication in BioVU/PMRP, the smaller sample size ($N = 62,271$) resulted in larger confidence intervals (CI), yet the correlation between effect sizes remained high ($\rho_{LTL \rightarrow T} = 0.930$; $\rho_{T \rightarrow LTL} = 0.874$). Specifically, 9 LTL on trait and 5 trait on LTL effects strictly replicated ($p < 0.05/141 = 3.5e-4$), while 17 and 12 reached nominal significance, respectively. Only the effect of white blood cell (WBC) counts on LTL had a significantly different effect size ($p_{\text{diff}} < 0.05/47 = 0.001$).

Second, we sought to assess if our results could be confounded by hematological factors, given that we use telomere length assessed in leukocytes. We therefore adjusted LTL for eosinophil, lymphocyte, monocyte, neutrophil, platelet, red blood cell, reticulocyte, and WBC counts in addition to core covariates. Regressing this new variable on the same 158 traits (i.e., 166 traits, excluding the 8 blood count traits we corrected for), we obtained highly similar effect sizes ($\rho = 0.98$). Only associations with smoking status ($p_{\text{diff}} = 1.4e-9$), smoking cessation ($p_{\text{diff}} = 5.7e-6$), as well as mean corpuscular hemoglobin (MCH; $p_{\text{diff}} = 3.3e-26$) were significantly reduced ($p_{\text{diff}} < 0.05/141 = 3.5e-4$), yet remained significant. Association with total bilirubin ($p_{\text{diff}} = 1.4e-5$) was lost, while the one with phosphate levels ($p_{\text{diff}} = 8.0e-5$) became significant (Additional file 1: Fig. S6). As a relationship between LTL and clonal hematopoiesis of indeterminate potential (CHIP) has been suggested [24, 25], we conducted bidirectional MR analysis to assess whether CHIP [26] could confound LTL associations. While long LTL had a causal impact on CHIP incidence ($\hat{\alpha}_{LTL \rightarrow T} = 0.147$; $p = 3.0e-9$), we did not identify a reverse effect ($\hat{\alpha}_{T \rightarrow LTL} = 0.040$; $p = 0.618$). This suggests that confounding of our MR analyses by CHIP is an unlikely scenario. MVMR analysis with blood counts (with significant association with LTL), MCH, CHIP, and LTL as exposures against LTL-impacted traits also did not reveal significant effect changes in effect sizes, confirming that neither CHIP nor other hematological parameters biased our results (Additional file 2: Table S3).

Third, we performed a negative control. As sex of an individual is determined prior to adult LTL, we should not observe any causal link from LTL to sex [27]. As expected, we did not find a significant causal IVW MR effect of LTL on sex ($p = 0.656$). To conclude, the broad range of sensitivity analyses we performed showed that our results are globally robust to assumption violation and confounding, allowing their biological interpretation.

Modulators of LTL

Lifestyle and environmental factors

Our results are overall concordant with deleterious lifestyle habits leading to shorter LTL (Fig. 2b). A negative correlation was observed between smoking cessation and LTL ($\hat{\beta} = -0.039$; $p = 9.4e-50$), mirrored by a detrimental causal effect of failure to quit smoking on LTL ($\hat{\alpha}_{T \rightarrow LTL} = -0.142$; $p = 1.8e-4$). Alcohol consumption, measured as total weekly intake of alcohol units, also exhibited a negative causal effect on LTL ($\hat{\alpha}_{T \rightarrow LTL} = -0.086$; $p = 1.3e-4$), while beef consumption showed a mere associative ($\hat{\beta} = -0.012$; $p = 2.4e-11$) but no causal link ($p = 0.223$). Conversely, healthy habits such as high fresh fruit intake ($\hat{\beta} = 0.014$; $p = 6.4e-15$) and physical activity ($\hat{\beta} = 0.007$; $p = 1.7e-4$) displayed positive associations with LTL, as did SES captured by average household income ($\hat{\beta} = 0.025$; $p = 1.1e-40$) or EA ($\hat{\beta} = 0.047$; $p = 1.9e-155$),

even though only the latter showed clear causal evidence ($\hat{\alpha}_{T \rightarrow LTL} = 0.075$; $p = 2.2e-15$). Our data also suggest that the psychological state of an individual can impact LTL as depression causes shorter LTL ($\hat{\alpha}_{T \rightarrow LTL} = -0.110$; $p = 8.0e-6$). We speculate that depression could accelerate LTL shortening by influencing lifestyle factors that promote oxidative stress and inflammation, both critical modulators of LTL [7, 28, 29]. This hypothesis is supported by a negative causal effect of the inflammation marker C-reactive protein (CRP) on LTL ($\hat{\alpha}_{T \rightarrow LTL} = -0.037$; $p = 9.3e-10$). While it is challenging to genetically instrument lifestyle factors, we conducted mediation analyses to explore the extent to which smoking cessation, frequency of alcohol intake, body mass index (BMI), and CRP levels mediate the effect of depression on LTL. We found that only CRP significantly mediated part of the relationship between depression and LTL ($P_M = 14.5\%$; 95% CI = [3.3%; 32.1%]). Overall, these results highlight the significant impact of lifestyle and environmental factors on LTL and support the paradigm that exposures typically considered as deleterious lead to shorter LTL.

Anthropometric traits

We detect several associations with anthropometric traits (Fig. 2c). Body metrics such as BMI and body fat mass (BFM) demonstrated significant negative observational correlation (BMI: $\hat{\beta} = -0.032$; $p = 2.4e-75$; BFM: $\hat{\beta} = -0.029$; $p = 1.2e-60$) and causal effects on LTL (BMI: $\hat{\alpha}_{T \rightarrow LTL} = -0.048$; $p = 4.9e-10$; BFM: $\hat{\alpha}_{T \rightarrow LTL} = -0.050$; $p = 7.6e-9$). Conversely, a positive correlation was observed between LTL and height ($\hat{\beta} = 0.018$; $p = 2.2e-24$), with MR analysis revealing a nominally significant effect of LTL on height ($\hat{\alpha}_{LTL \rightarrow T} = 0.062$; $p = 4.5e-4$) and strictly significant effect of height on LTL ($\hat{\alpha}_{T \rightarrow LTL} = 0.013$; $p = 4.0e-5$).

Female reproductive traits

Observational correlation between LTL and female reproductive traits including age at first (AFB; $\hat{\beta} = 0.042$; $p = 1.4e-54$) and last (ALB; $\hat{\beta} = 0.034$; $p = 2.5e-36$) live birth, reproductive lifespan ($\hat{\beta} = 0.023$; $p = 3.7e-13$), age at menopause ($\hat{\beta} = 0.026$; $p = 1.5e-16$), and menstrual disorders ($\hat{\beta} = 0.011$; $p = 1.7e-5$) were observed (Fig. 2d). Testing for age at menarche showed no association with LTL ($\hat{\beta} = 0.004$; $p = 0.113$). Only the effect of AFB ($p_{\text{diff}} = 6.7e-5$) and ALB ($p_{\text{diff}} = 0.001$) were significantly ($p_{\text{diff}} < 0.05/11 = 4.5e-3$) reduced after accounting for SES, even though they remained significant (Additional file 1: Fig. S7a). Both traits also causally influenced LTL (AFB: $\hat{\alpha}_{T \rightarrow LTL} = 0.167$; $p = 1.2e-5$; ALB: $\hat{\alpha}_{T \rightarrow LTL} = 0.272$; $p = 6.1e-6$), suggesting that timing of female reproductive events could modulate LTL. While these effects showed nominally significant MR-Egger intercepts, potentially indicating directional pleiotropy, the latter did not survive multiple testing correction (Additional file 2: Table S2). To explore this further, we compared LTL in women with and without children, finding shorter LTL in women who had given birth (Welch two-sample t-test: $p = 7.4e-11$). This suggests that childbirth could accelerate LTL shortening. We next divided female participants' age into three reproductive periods: (1) premenopausal before first live birth, (2) premenopausal after first live birth, and (3) postmenopausal, and used the number of years spent in each period as predictors of LTL. LTL shortening accelerated over the course of these periods, with the weakest effect

on LTL found for premenopausal years before childbirth ($\hat{\beta} = -0.014$; $p = 3.6e-120$), followed by premenopausal years after childbirth ($\hat{\beta} = -0.017$; $p = 7.1e-233$), and postmenopausal years ($\hat{\beta} = -0.022$; $p < 2.2e-308$) (Fig. 3), in line with the hypothesis that female reproductive events trigger acceleration in LTL shortening.

Serum lipids

We found predominantly positive associations between LTL and serum lipid levels, i.e., apolipoprotein B (ApoB; $\hat{\beta} = 0.019$; $p = 9.4e-25$), total cholesterol ($\hat{\beta} = 0.019$; $p = 2.9e-27$), and low-density lipoprotein (LDL)-cholesterol ($\hat{\beta} = 0.022$; $p = 4.2e-35$) (Fig. 2e). After adjusting for cholesterol-lowering drug use, the positive relation between LTL and both total and LDL-cholesterol decreased but remained significant (Additional file 1: Fig. S7b). ApoB ($\hat{\alpha}_{T \rightarrow LTL} = 0.029$; $p = 2.6e-10$), total cholesterol ($\hat{\alpha}_{T \rightarrow LTL} = 0.035$; $p = 1.5e-08$), and LDL-cholesterol ($\hat{\alpha}_{T \rightarrow LTL} = 0.036$; $p = 4.7e-10$) levels also causally influenced LTL. Consistently, our findings suggest that disorders of lipid metabolism contribute to longer LTL ($\hat{\alpha}_{T \rightarrow LTL} = 0.058$; $p = 5.8e-7$), reiterating the association between increased LTL and high serum lipid levels. Due to their correlated nature, MVMR including levels of LDL-cholesterol, ApoB, and triglycerides as exposures could not disentangle their individual contribution to LTL (Additional file 1: Fig. S8).

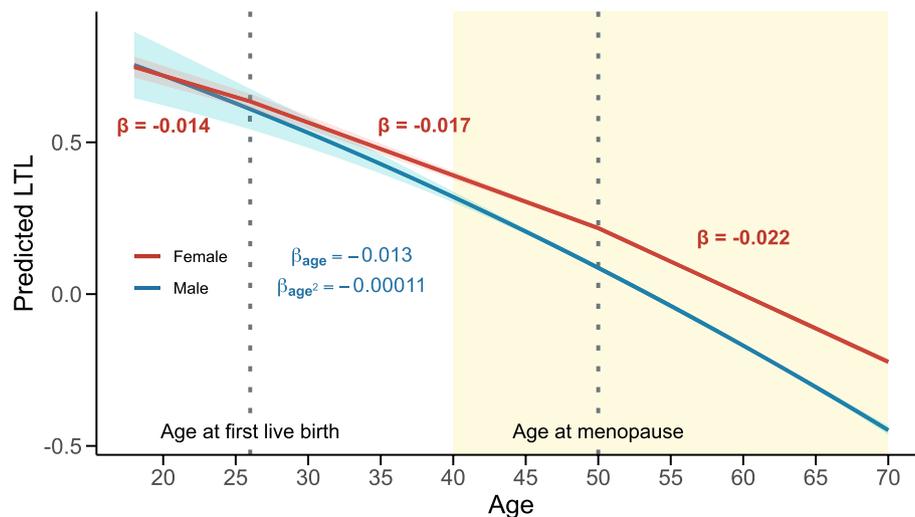


Fig. 3 Schematic representation of LTL shortening across different female reproductive life phases. Relation (β) between predicted (i.e., regression model fit) standardized LTL (y-axis) and age (x-axis) across three female reproductive life periods (red). Dotted vertical lines indicate mean age at first live birth (26 years) and mean age at menopause (50 years). As a comparison, we depict the quadratic LTL regression in males (β_{age} ; β_{age^2} ; blue). 95% confidence intervals are shown for the predictions. Yellow background indicates the age range for which data are available (40–70 years) and used to build predictions; regions outside this range are extrapolated for males and estimated from age at first live birth and age at menopause information for females. The x-axis was set to begin at age 18, reflecting the shift in the rate of LTL decline after puberty [30]. This change cannot be accurately captured by our linear models, which are based on LTL measurements at older ages

Urate

Urate levels, also retained as a relevant predictor of LTL by the Lasso regression analysis, displayed a negative association with LTL ($\hat{\beta} = -0.025$; $p = 2.1e-44$). As previously reported [31], MR analyses showed that elevated urate levels decreased LTL ($\hat{\alpha}_{T \rightarrow LTL} = -0.042$; $p = 9.4e-18$), possibly due to increased cellular stress and reactive oxygen species production [32] (Fig. 2f). The urate-LTL association was significantly mediated by CRP, confirming the role of inflammation in this process ($P_M = 34.7\%$; 95% CI = [17.1%; 55.6%]; Additional file 2: Table S4).

Consequences of altered LTL

Blood cell counts

Hematological traits (e.g., WBC count: $\hat{\beta} = -0.042$; $p = 2.9e-120$; and MCH: $\hat{\beta} = -0.054$; $p = 1.7e-200$) are among the ones showcasing the strongest observational correlation with LTL (Fig. 2g). For four out of eleven significantly correlated blood traits, we identified bidirectional causal relationships with LTL, with less pronounced effects from traits on LTL (e.g., MCH: $\hat{\alpha}_{LTL \rightarrow T} = -0.195$; $p = 2.2e-24$; $\hat{\alpha}_{T \rightarrow LTL} = -0.034$; $p = 5.2e-10$). While effects of LTL on MCH, eosinophil, platelet, and red blood cell counts were robust across multiple MR methods (Additional file 1: Fig. S3), the effects of blood traits on LTL did not necessarily pass Bonferroni correction in all sensitivity analyses (Additional file 1: Fig. S4). Given the previously described analyses (see the “Sensitivity analyses” section), it appears that these blood traits do not confound the other observed relationships with LTL.

Hepatic biomarkers

LTL associated with levels of the hepatic biomarkers aspartate aminotransferase (AST; $\hat{\beta} = -0.023$; $p = 1.6e-37$) and albumin ($\hat{\beta} = 0.007$; $p = 7.2e-5$) (Fig. 2h). Accordingly, finding that shorter LTL causally associated with higher AST ($\hat{\alpha}_{LTL \rightarrow T} = -0.082$; $p = 3.7e-11$) and lower albumin levels ($\hat{\alpha}_{LTL \rightarrow T} = 0.050$; $p = 9.1e-5$), telltales of underlying liver or inflammatory conditions. Hepatic disorders, which can lead to altered levels of AST [33], are a feature of telomere biology disorders [34]. Accordingly, we observe an association between LTL and liver fibrosis/cirrhosis ($\hat{\beta} = -0.012$; $p = 1.2e-10$). However, nonalcoholic fatty liver disease, which reflects the early stages of liver disease and has GWASs with larger sample sizes, does not demonstrate any significant MR effects ($p = 0.212$). Nevertheless, these results underscore the potential role of telomere-driven cellular aging in hepatic function and/or inflammatory processes.

Diseases

Longer LTL correlated with decreased risk for cardiovascular and pulmonary conditions, reflecting previous findings [12, 35]. For instance, LTL had a negative causal impact on aneurysm risk ($\hat{\alpha}_{LTL \rightarrow T} = -0.190$; $p = 3.0e-5$) and a positive one on forced vital capacity ($\hat{\alpha}_{LTL \rightarrow T} = 0.072$; $p = 3.2e-6$). In line with that, we observed a negative correlation with risk for pulmonary diseases such as emphysema ($\hat{\beta} = -0.013$; $p = 3.5e-12$) or chronic obstructive pulmonary

disease (COPD; $\hat{\beta} = -0.025$; $p = 9.0e-40$). While the MR effects of LTL on emphysema ($\hat{\alpha}_{LTL \rightarrow T} = -0.115$; $p = 0.005$) or COPD ($\hat{\alpha}_{LTL \rightarrow T} = -0.037$; $p = 0.012$) were concordant, they did not survive multiple testing correction. In addition to replicating a previously established correlation between short LTL and increased risk for ischemic heart disease ($\hat{\beta} = -0.024$; $p = 6.6e-41$) [35], we also found causal evidence for the effect of LTL on ischemic heart disease ($\hat{\alpha}_{LTL \rightarrow T} = -0.064$; $p = 3.5e-10$) (Fig. 2i–j). Hematological cancer risk negatively correlated with LTL ($\hat{\beta} = -0.015$; $p = 5.8e-14$), while longer LTL correlated with kidney ($\hat{\beta} = 0.008$; $p = 9.4e-05$) and prostate ($\hat{\beta} = 0.029$; $p = 2.1e-23$) cancer risk. While we do not have causal estimates for the former, MR confirmed that LTL causally increased risk for kidney cancer ($\hat{\alpha}_{LTL \rightarrow T} = 0.048$; $p = 8.1e-10$) and we found a near-significant trend for prostate cancer ($\hat{\alpha}_{LTL \rightarrow T} = 0.080$; $p = 4.0e-4$) (Fig. 2k), aligning with previous findings [17, 20, 36, 37]. This paradox, in which both longer and shorter LTL impact disease risk, was also observed in disorders with an autoimmune component, where shorter LTL is a risk factor for rheumatoid arthritis ($\hat{\alpha}_{LTL \rightarrow T} = -0.087$; $p = 6.1e-5$) and Alzheimer's disease ($\hat{\alpha}_{LTL \rightarrow T} = -0.038$; $p = 1.2e-4$), while longer LTL increased risk for systemic lupus erythematosus ($\hat{\alpha}_{LTL \rightarrow T} = 0.167$; $p = 5.8e-5$) (Fig. 2l–o). Overall, these results highlight the disease-promoting role of both long and short LTL, aligning with previous findings that both shorter and longer telomeres are associated with premature death [18].

Mediating role of LTL

Analogously to DNA methylation, LTL represents a marker of biological age that can be viewed as a clock integrating a broad range of lifestyle and health parameters [38]. This raises the question of whether LTL mediates the relation between complex traits and lifespan. We tested the mediating role of LTL for the relation between 18 non-hematological LTL-modulating traits and lifespan, the latter being affected by LTL at nominal significance ($\hat{\alpha}_{LTL \rightarrow T} = 0.023$; $p = 0.008$). We identified six significant indirect effects ($p_{\text{indirect}} < 0.05/323 = 1.5e-4$), i.e., mediated through LTL (Fig. 4a and Additional file 2: Table S4). For instance, the negative impact of BMI ($P_M = 7.2\%$; 95% CI = [3.9%; 10.6%]) or the positive effect of EA ($P_M = 18.8\%$; 95% CI = [12.3%; 25.7%]) on lifespan were partially mediated by LTL. Given the considerable mediation of AFB ($P_M = 80.8\%$; 95% CI = [39.4%; 100%]) and ALB ($P_M = 100\%$; 95% CI = [70.1%; 100%]) on lifespan by LTL, we further investigated these traits through an iterative MVMR approach to build a causal network (Fig. 4b and Additional file 2: Table S3). Results emphasized the partial mediating role of LTL and EA on the effect of AFB on lifespan.

Given that lifestyle factors were found to affect LTL, which in turn influences risk for many diseases, we next used MVMR to assess the LTL mediatory effect for all pairs of 18 LTL modulators and 17 LTL-affected traits. We identified 24 significant ($p_{\text{indirect}} < 0.05/323 = 1.5e-4$) LTL-mediated relationships (Fig. 4c and Additional file 2: Table S4). Effects on ischemic heart disease, total protein levels, and forced vital capacity were the most frequently mediated by LTL, whereas urate levels, CRP, BMI, BFM, and EA were the most common exposures. Overall, while we do detect a substantial number of significant mediations through LTL, the average mediation proportion is 5.71%, only accounting for a fraction of these relations.

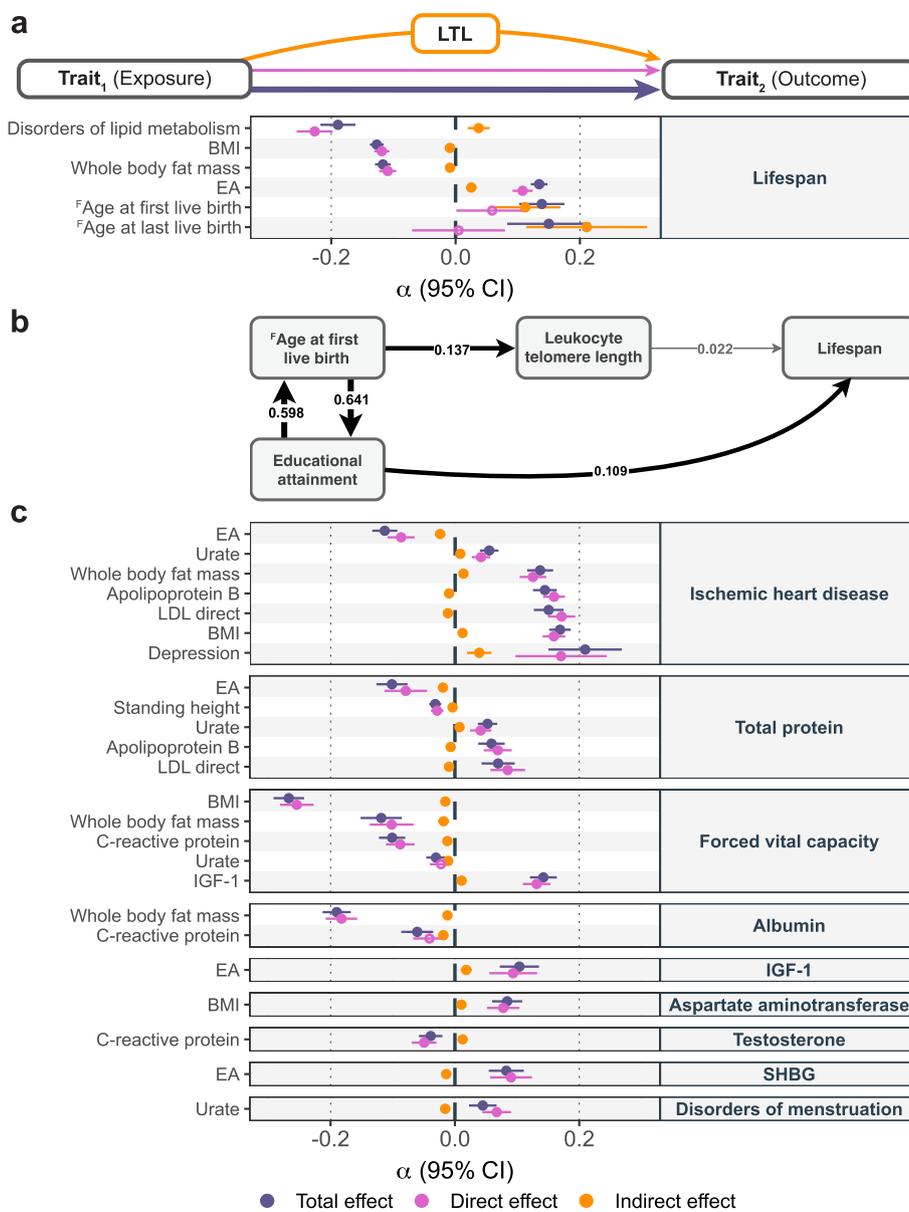


Fig. 4 Mediating role of LTL. **a** Mediation analysis of 18 LTL-affecting exposures (y-axis; left) on lifespan (y-axis; right) through LTL with effect size estimates and 95% confidence intervals (CI; x-axis) of the total effect (i.e., IVW MR estimate of exposure on outcome; purple), direct effect (i.e., not mediated by LTL; MVMR estimate; pink) and indirect effect (i.e., LTL mediation by product method; orange) as displayed in the scheme on top of the figure. Displayed are relationships with significant ($p < 0.05/323 = 1.5e-4$) total and indirect effects. **b** Schematic illustration of the magnitude and direction of nominally significant MVMR effects ($p < 0.05$; gray arrow). Strictly significant ($p < 0.05/141 = 3.5e-4$) effects are shown as black arrows. Arrow thickness is proportional to the effect size. Nominally significant effect from lifespan to EA is not displayed. **c** Mediation analysis of 18 LTL-affecting exposures (y-axis; left) on 17 LTL-affected outcomes (y-axis; right) through LTL. Legend as in **a**. EA = educational attainment; LDL = low-density lipoprotein; BMI = body mass index; IGF-1 = insulin-like growth factor 1; SHBG = sex hormone binding globulin; LTL = leukocyte telomere length. Labels preceded by an uppercase F denote female-specific traits (i.e., age of first and last live birth)

Discussion

In this study, we comprehensively examined the bidirectional causal relationships between LTL and complex traits, diseases, and lifestyle factors and used MVMR to examine causal effect mediation. Our study reiterates age and sex as major determinants of LTL variability [18, 22] and confirms the causal effects of lifestyle factors on LTL. Furthermore, we provide robust evidence for a causal role of abnormal LTL in a broad spectrum of clinically relevant traits, including cancer, autoimmune disorders, lung diseases, and cardiovascular conditions. Lastly, our results show that LTL partially mediates the effect of BMI, EA, and reproductive traits on lifespan.

Others [20, 21, 39] have used MR to estimate the impact of LTL on the human phenotype. In contrast to Codd et al. [20], we used summary statistics originating from large consortia, which offer the double advantage of reducing sample overlap between the exposure and outcome sample and have much higher case numbers (and thus power) than in the relatively healthy UKBB. We hypothesize that this explains some of the nine novel findings that are unique to our study, five of which were tested but did not yield significant results in Codd et al. (Additional file 2: Table S5). For instance, our results support the Alzheimer's disease risk-increasing effect of shorter LTL, which was proposed to be driven by promotion of cellular senescence [40–45]. In line with previous literature [46, 47], our results also suggest that longer LTL increases the risk of systemic lupus erythematosus. Combined with our confirmation that shorter LTL raises rheumatoid arthritis risk, these findings highlight the implication of LTL in autoimmune conditions. Finally, we report a coherent impact of longer LTL on risk for menstrual disorders [21], once again linking LTL to female reproductive health (Fig. 3). Overall, this supports the deleterious role of both long and short telomeres in human health.

Unlike previous large-scale studies investigating the link between LTL and the human phenotype [18, 20, 37, 39], we also estimated the causal effects of phenotypes on LTL. In line with prior research, alcohol consumption [48], smoking [8, 49], obesity [50], and socio-economic disadvantages [51, 52] emerged as significant contributors to telomere shortening, underscoring the potential benefits of lifestyle modifications. Some of these factors, such as BMI and EA, were found to exert a small, albeit significant proportion of their impact on longevity through LTL. Our findings further indicate that the impacts of both depression and urate levels on LTL are partly exerted through CRP, highlighting inflammation's role as a mediating factor of LTL shortening [29]. Surprisingly, the positive influence of serum lipid levels on LTL often attenuated the total effect of lipid-trait relationships. These results are unexpected as high cholesterol levels promote oxidative stress [53], which in turn accelerates LTL shortening [54]. Yet they are coherent with recent literature [55] suggesting a nuanced protective role of specific lipids and lipoproteins in maintaining telomere length. However, our inability to replicate these effects when using independent LTL summary statistics, possibly due to lack of statistical power, warrants further investigation into the mechanisms through which elevated lipid levels may support telomere preservation.

One of the most intriguing findings of our study is the causal relationship between delayed AFB and ALB and extended LTL in females, which was only partially confounded by SES. These results align with the accelerated LTL shortening rate we observed after childbirth, which is further exacerbated post-menopause. Although we

did not observe an association between oestradiol levels — measured only in ~49,000 UKBB female participants — and LTL, we hypothesize that hormonal shifts following pregnancy and menopause could accelerate LTL shortening [56]. An alternative explanation is that LTL shortening is driven by the stress imposed by such events on the body, aligning with literature that posits pregnancy as a significant accelerator of biological aging, measured per methylation [57]. Notably, no significant association between age at menarche and LTL was found. While this could reflect a genuine absence of association, inaccurate reporting of age at menarche and inability to assess shifts in telomere shortening rate before adulthood [30] could also contribute to this negative result. Extending prior literature [58], our findings suggest that delayed female age at childbirth increases longevity partly through telomere length, even though we cannot exclude partial confounding by SES, which also significantly influences this relationship. While further research is required to test these hypotheses, our results highlight the prominent role of life history events in LTL shortening rates.

Our study is subject to several limitations. First, the use of cross-sectional bulk LTL limits our capacity to analyze individual telomere shortening rates, which might be a critical factor in disease prediction [38]. Second, although LTL and telomere length in other tissues are correlated [59], this proxy might miss more subtle and tissue-specific relations between telomere length and the phenome. Importantly, the causal relevance of LTL in non-blood-related diseases may be limited, as observed associations could reflect shared genetic effects on telomere length across different tissue types, rather than direct effects of telomere length in leukocytes. Furthermore, we cannot exclude that since telomere length was measured in leukocytes, it made finding associations with hematological traits more likely, despite our sensitivity analyses showing robustness against confounding by hematological factors. Additionally, our study focused on White-British ancestry, meaning the results may not translate to other ancestral groups. In the future, single-cell telomere length measured at chromosomal resolution through long-read sequencing approaches across various tissues, time points, and ancestral groups should provide a more refined view of telomere dynamics. Third, Lasso regression in our study was limited to data without missing values. Comparisons between cases with and without missing data showed shorter mean LTL in the former, indicating that data missingness, previously correlated with several health traits [60], constitutes an additional limitation to our findings. Fourth, MR presents with inherent limitations, notably susceptibility to horizontal pleiotropy violations, especially given the considerable heterogeneity across our IVs. While MVMR analyses can mitigate biases introduced by pleiotropy and elucidate direct causal effects, these analyses are also more likely to be subject to weak instrument bias, which is indicated by several conditional F-statistics falling below 10 (Additional file 2: Table S3) [61]. We used a broad range of sensitivity analyses and focused on results robust across these various methods. Another limitation of MR is that detection power is bound by the number of available IVs, so that our power to detect causal relations between traits and LTL is variable across phenotypes and might be lower or larger than for the reverse LTL on trait relation, depending on whether the trait has less or more IVs than LTL, respectively. Moreover, we derive our causal estimates from two unidirectional models, each fitting a different causal direction. Not using an explicit bidirectional model may marginally overestimate the effect

sizes for the five hematological traits with bidirectional MR effects, but the expected bias is minimal [62]. Fifth, treating diseases as binary exposures may violate the exclusion restriction assumption [63], prompting careful interpretation of the effects of hypertension, lipid disorders, and depression on LTL. Finally, MR does not account for dynamic spatiotemporal changes in LTL that occur over lifetime and/or in the context of some diseases such as cancer.

Conclusions

In conclusion, through usage of rigorous univariable and multivariable bidirectional Mendelian randomization, we identify a complex network of causal relations wherein both exogenous (i.e., lifestyle or environmental) and endogenous (i.e., physiological) factors modulate LTL, which in turn influences the risk for numerous diseases and mediates the impact of some of these traits on lifespan. Still, based on currently available data, the mediating role of LTL between lifestyle and disease risk is estimated to be modest, and further research is needed to explore the relation between LTL and other aging biomarkers, such as DNA methylation, to understand its clinical value as a proxy of biological age.

Methods

Software

All analyses were conducted using R v4.2.1 and Python v3.11.3. PLINK v1.90b7 [64] was used. Workflow management was facilitated by Snakemake v7.25.3 [65].

Data

Individual-level UK Biobank data

Observational analyses were carried out in the UK Biobank (UKBB), a cohort of ~500,000 volunteers from the general UK population aged between 40 and 69 years at recruitment [66]. Analyses were conducted on 326,363 participants with known sex, age, and LTL after the exclusion of individuals of non-white and non-British ancestry (self-reported + genetically defined), relatives ($\leq 3^{rd}$ degree), and gender mismatches (see UKBB Resource 531), as well as those who retracted their participation. Given that LTL measurements are derived from blood, we further excluded 4,376 individuals with blood malignancies, based on self-reports (UKBB field #20001 codes 1047, 1048, 1050, 1051, 1052, 1053, 1055, 1056, 1058) or hospital diagnoses (#41270; International Classification of Diseases 10th Revision [ICD10] codes mapping to the Phecode “cancer of lymphatic and hematopoietic tissue” [67]).

We used technically adjusted and standardized LTL (#22192) [19] and assessed its relation to 166 complex traits (Additional file 2: Table S1). These include 60 common diseases defined based on hospital diagnoses (#41270; last diagnosis September 2021), while excluding from controls individuals with self-reported (#20001, #20002) or hospital-diagnosed (#41270) conditions related to the investigated disease [68]. Disease phenotypes were used to calculate a disease burden phenotype, i.e., the total number of diseases diagnosed in an individual among the 60 considered ones. The remaining 105 traits include 11 anthropometric traits (e.g., weight), 41 biomarkers (e.g., serum lipids), 18 life events (e.g., age at menarche and menopause), 26 lifestyles

(e.g., beef intake) and socio-economic factors (e.g., Townsend deprivation index), and 9 miscellaneous traits. Definitions of composite phenotypes are described in the Additional file 1. Briefly, continuous traits with multiple instances were averaged, while the first instance was used for integers or factors. To minimize noise, outliers (mean ± 5 standard deviations [SD]) in continuous traits were removed. Factorial variables were numerically converted for efficient integration into the regression model. All traits, including binary predictors, were then scaled to have zero mean and unit variance to obtain more comparable effect sizes. As the 167 assessed traits (i.e., 166 above-mentioned + blood cancer) were partially correlated we estimated the number of effective tests [69], i.e., the number of tests needed to explain 99.5% of the variance in our phenotypic dataset, to 141, resulting in a significance threshold of $p < 0.05/141 = 3.5e - 4$ for observational correlation and MR analyses.

GWAS summary statistics

When available (i.e., for non-composite traits), genome-wide association study (GWAS) summary statistics originate from the Neale group (file release July 2018 [70]; Additional file 2: Table S6). Summary statistics for reproductive lifespan were derived from GWAS on age at menopause and menarche by first back-transforming the effects on year-scale and then computing their difference:

$$\beta = \frac{SD_{menopause} \times \beta_{menopause} - SD_{menarche} \times \beta_{menarche}}{SE = \sqrt{SD_{menopause}^2 \times SE_{menopause}^2 + SD_{menarche}^2 \times SE_{menarche}^2}}$$

The sample size for the resulting summary statistic was set to the lowest of the two (i.e., age at menopause; $N = 111,593$) and p -values were computed with a two-sided test based on a t -statistic obtained by dividing the effect size by its standard error. For diseases, a set of previously compiled GWAS summary statistics [71] of predominantly European-descent consortia meta-analyses was used (Additional file 2: Table S6). CHIP summary statistics, originally in build 38, were mapped to human genome build 37 using UCSC LiftOver [72]. Summary statistics were harmonized with the UK10K reference panel [73] and restricted to autosomal chromosomes. After excluding palindromic single-nucleotide polymorphisms (SNP) and adjusting strand-flipped SNPs, effect sizes were standardized to represent the square root of the explained variance [74].

Observational correlation

Predictors of LTL variability

To estimate the fraction of LTL variability explained by the human phenome, we used Lasso regression (glmnet package in R [75]) with unadjusted normalized LTL as the outcome variable and traits with less than 5%, 7%, and 10% missing data as possible predictors in a joint model. Given the non-deterministic choice of the optimized regularization parameter (one SE rule lambda), 50 regressions were fitted and traits that were selected in at least 95% of the cases were considered as predictors.

Single trait linear regression

We adjusted LTL by regressing out age, age², genotyping array, sex, and the interaction of the latter with the priors as fixed effects and used this variable as the outcome in 166 linear regression models with the traits described in Additional file 2: Table S1 as explanatory variables. Effect sizes reported in text are in SD_{LTL}/SD_{Trait} , except for the effect of age, in which case effects are reported in $SD_{LTL}/year$. We followed up on specific associations with sensitivity analyses to identify possible confounders:

- In individuals using cholesterol-lowering drugs (#6177 and #6153), serum lipid levels were corrected for average simvastatin effect, i.e., + 1.6 mmol/L, 1.4 mmol/L, 0.4 mmol/L, – 0.1 mmol/L of total cholesterol, low-density lipoprotein (LDL), triglycerides and high-density lipoprotein (HDL), respectively [76].
- Reproductive traits showing significant ($p < 0.05/141 = 3.5e-4$) association with LTL were corrected for socio-economic status (i.e., Townsend deprivation index (#189), average total household income before tax (#738), and educational attainment (see Additional file 1)).
- In addition to age, sex, and array, LTL was corrected for eosinophil (#30150), lymphocyte (#30120), monocyte (#30130), neutrophil (#30140), platelet (#30080), red blood cell (#30010), reticulocyte (#30250), and white blood cell (#30000) counts and linear regressions with non-blood trait count traits were performed anew to ensure the LTL associations were unbiased. As a result, the available sample size was reduced to $N = 308,346$.

Female reproductive phases

To assess the impact of childbearing and menopause on LTL, we identified three distinct female reproductive phases: (1) years before first live birth, (2) premenopausal years after first live birth, and (3) postmenopausal years. Number of years spent in each phase was derived from current age (#21003), age at first live birth (#2754), and age at menopause onset (#3581). Phases (2) and (3) were set to 0 for females with no children (#2734: number of live births = 0) and premenopausal women, respectively. The joint linear regression model included time spent in each phase and two indicator variables for whether the women carried a pregnancy to term and experienced menopause. Female participants who had their first child post-menopause, lacked a menopausal status (#2724) or age at menopause (#3581), or did not specify childbirth events (#2734) or age at first childbirth (#2754) were excluded from this analysis.

Mendelian randomization

Bidirectional univariable Mendelian randomization

GWAS summary statistics were used to conduct bidirectional two-sample MR, with $\hat{\alpha}_{LTL \rightarrow T}$ representing the causal impact of LTL (exposure) on complex traits (outcome) and $\hat{\alpha}_{T \rightarrow LTL}$ the causal impact of complex traits (exposure) on LTL (outcome) (Fig. 1). Harmonized SNPs significantly associated ($p < 5e-8$) with the exposure were clumped ($p1 = 0.0001$, $p2 = 0.01$, $kb = 250$, and $r2 = 0.01$) with PLINK v1.9 [64] and

retained as instrumental variables (IVs). As the *HBB* gene was used as a control for the LTL measurements, SNPs in this gene (chr11:5,246,696–5,248,301; GRCh37/hg19) that associated with LTL were removed to prevent spurious associations [20]. For LTL IVs, this led to the exclusion of a single variant, rs1609812. Due to the complex long-range linkage disequilibrium (LD) structure of the HLA locus, SNPs mapping to that region (chr6:25,000,000–37,000,000; GRCh37/hg19) were also excluded from our IVs [77]. For each exposure-outcome pair, further IVs were removed based on difference in allele frequency (≥ 0.05) and Steiger filtering ($Z \leq -1.96$). Bidirectional MR analyses were carried out with the TwoSampleMR R package (v0.5.7) [78], primarily through the IVW method. LTL on trait and trait on LTL MR effects were computed for 152 and 142 traits, respectively, with at least two IVs.

Sensitivity analyses for relationships with significant IVW MR effects were conducted using additional MR methods, i.e., MR Egger, simple mode, weighted median, and weighted mode, to ensure robustness of the results. Heterogeneity was assessed using Cochran's Q-statistics. Given a high proportion of elevated Q-statistics, we additionally run MR-PRESSO [79] for relationships with significant IVW MR effects. To further ensure that our results are not biased by pleiotropy – which violates the MR assumption that IVs only affect the outcome through the exposure [80] – we first filtered genome-wide significant exposure SNPs and harmonized these SNPs across all 153 traits with available GWAS summary statistics, i.e., verified that SNPs are present across the 152 traits + LTL summary statistics. This step is carried out before clumping to guarantee that the identified IVs are consistently present across all outcomes, enabling subsequent comparisons. After clumping, Steiger filtering was applied between the exposure and all other traits to ensure that the selected SNPs are more strongly associated with the exposure than with any of the other included traits. SNPs that passed filtering for all traits were retained as IVs and MR analyses were conducted on these. This approach serves as a reasonable pleiotropy filter due to the diverse nature of our phenotypes. While sample overlap in two-sample MR may bias results toward observational effects, no overlap may also lead to 'winner's curse' bias [81, 82]. Weak instruments can further exacerbate these biases. Although our extensive simulations have demonstrated that these issues lead to mild biases [82], we used MR-APSS [83] (default parameters and LD scores from 1000 Genomes Data [84]), which addresses both sample overlap and pleiotropy. Finally, we replicated the MR analyses using LTL summary statistics generated based on an independent sample ($N = 62,271$) from the biobank of Vanderbilt University Medical Center (BioVU) and Marshfield Clinic's Personalized Medicine Research Project (PMRP) [18, 85, 86].

LTL mediation analysis

Excluding hematological traits due to potential confounding, we used a two-exposure multivariable MR (MVMR) framework to individually assess the mediating role of LTL across 18 LTL-affecting traits ($p < 0.05/141 = 3.5e-4$) on lifespan, proxied through parental lifespan [87]. We further examined the global mediatory role of LTL between each of these 18 LTL-affecting traits and 17 traits causally impacted by LTL ($p < 0.05/141 = 3.5e-4$). This corresponded to 323 pairs (18 * 18 (i.e., 17 traits + lifespan), excluding one pair as IGF-1 associated with LTL as both exposure and outcome). This sets our significance threshold for

the total and indirect effects at $p < 0.05/323 = 1.5e-4$. IVs for mediation analyses were selected from summary statistics through a two-phase clumping process (see the “[Bidirectional univariable Mendelian randomization](#)” section) in which harmonized exposure IVs (i.e., trait IVs) were first independently clumped. In the second phase, exposure and mediator IVs (i.e., LTL IVs) were clumped together, prioritizing the former over the latter (i.e., retaining exposure IVs over mediator IVs). Providing MR assumptions hold, by instrumenting both the exposure and mediator we also reduce confounding bias between mediator and outcome. Steiger filtering was applied to both exposure IVs with respect to outcome and mediator and to mediator IVs with respect to the outcome. Indirect effects were determined through two strategies: difference in coefficients and product of coefficients [88]. The former subtracts the direct effect (MVMR) from the total effect (IVW), while the latter multiplies the univariable MR estimates from the exposure on the mediator by the MVMR effect of the mediator on the outcome. Both approaches generated consistent results (Additional file 1: Fig. S9) and we present the product of coefficients method due to easier interpretability in the main text. We further corrected these estimates for regression dilution bias [74]. Mediation proportions (P_M) represent the ratio of the indirect (α_{indirect}) to total (α_{total}) effect with 95% confidence intervals (upper limit capped at 100%) estimated from the 2.5th and 97.5th quantiles of the distribution of 10,000 simulated ratios drawn from $\tilde{\alpha}_{\text{indirect}} \sim N(\hat{\alpha}_{\text{indirect}}, \hat{SE}(\hat{\alpha}_{\text{indirect}}))$ and $\tilde{\alpha}_{\text{total}} \sim N(\hat{\alpha}_{\text{total}}, \hat{SE}(\hat{\alpha}_{\text{total}}))$.

Multi-trait analysis for direct effect estimation

For MVMR with multiple exposures and no predefined mediator, IVs were selected through a two-step process [89]. First, SNPs for each exposure were ranked according to their p -values (more significant p -values receiving lower ranks) and minimum rank across all exposures was determined for each SNP. This minimum rank was used to prioritize SNPs in a subsequent clumping process. IVs were filtered as previously described. Finally, MVMR regression estimates were compared to univariable MR estimates (see the “[Effect size comparison](#)” section). For the univariable MR, we either used the same IVs as in the MVMR or employed a subset of IVs, which were retained after Steiger filtering between both the outcome and the exposure of interest, as well as between the exposure of interest and the other exposures. We report weak instrument bias via conditional F-statistics [61] and heterogeneity through Cochran’s Q-statistic [90] (Additional file 2: Table S3).

Effect size comparison

Significant differences between two estimated effect sizes $\hat{\beta}_X$ and $\hat{\beta}_Y$ were assessed with a two-sided p -value (p_{diff}) derived from:

$$t = \frac{\hat{\beta}_X - \hat{\beta}_Y}{\sqrt{SE_X^2 + SE_Y^2}}$$

which assumes that the two estimates are uncorrelated. Often these estimates have a positive correlation (as estimated from the same data) and hence the t -statistic has a variance smaller than one, thus the test is conservative. This approach was used throughout the study to assess the effect of sensitivity analyses and compare univariable MR and MVMR results.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s13059-024-03269-9>.

Additional file 1: Supplementary note. Definition of composite traits. Fig. S1. Linear regression of LTL against age, stratified by sex. Fig. S2. Decomposition of the effects of mother's and father's age at birth on LTL. Fig. S3. MR estimates of LTL on traits across different methods. Fig. S4. MR estimates of trait on LTL effects across different methods. Fig. S5. Stringent Steiger pleiotropy-sensitivity analysis. Fig. S6. Linear regression estimates adjusted for blood counts. Fig. S7. LTL associations adjusted for potential confounding variables. Fig. S8. MVMR effect estimation of traits on LTL. Fig. S9. Mediating role of LTL on complex trait pair relations.

Additional file 2: Table S1. Description of analyzed complex traits. Table S2. Associations of traits with LTL: observational and MR estimates. Table S3. Complementary LTL association analyses. Table S4. Mediation analysis results. Table S5. Comparative analysis of LTL on traits MR findings. Table S6. GWAS summary statistics sources.

Additional file 3: Review history.

Acknowledgements

We thank all UK biobank participants for sharing their data, as well as the authors of Allaire et al. 2023 for granting us early access to their telomere length GWAS summary statistics. Computations were performed on the Urblauna servers from the University of Lausanne.

Review history

The review history is available as Additional file 3.

Peer review information

Andrew Cosgrove was the primary editor of this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.

Authors' contributions

SM, CA, and ZK conceived the study; SM carried out the analyses with contributions from MCS; ZK supervised statistical analyses; SM generated the figures; SM and CA drafted the manuscript and ZK made critical revisions; All authors read, approved, and provided feedback on the final manuscript.

Authors' X handles

X handles: @SamuelMoix (Samuel Moix); @smarie_smarie (Marie Sadler); @zkutalik (Zoltán Kutalik); @CAuwerx (Chiara Auwerx).

Funding

Open access funding provided by University of Lausanne The study was funded by the Swiss National Science Foundation (310030_189147, ZK) and the Department of Computational Biology of the University of Lausanne (ZK).

Availability of data and materials

UKBB data are available for registered users. UK10K reference panel is available upon request at https://www.uk10k.org/data_access.html [73]. European LD scores from 1000 Genomes Data are freely accessible [84, 91]. GWAS summary statistics originate from various sources: Alzheimer's disease [92], lifestyles [93], sleep apnea, celiac disease, endometriosis, pneumonia, psoriasis, and valvular heart disease [94], asthma [95], balding [96], breast cancer [97], bipolar disorder [98], cataract [99], CHIP [26], liver fibrosis [100], CKD [101], atrial fibrillation [102], colorectal cancer, kidney cancer [103], depression [104], epilepsy [105], glaucoma [106], inflammatory bowel disease [107], coronary artery disease [108], kidney stones [109], multiple sclerosis [110], osteoarthritis [111], ovarian cancer [112], prostate cancer [113], Parkinson's disease [114], lifespan [87], rheumatoid arthritis [115], schizophrenia [116], sex [117], smoking cessation [118], stroke [119], type 1 diabetes [120], WHR [121], LTL [18, 20], and else from the Neale Lab [70] or Pan-UK Biobank [122]. Links to download the summary statistics are provided in Additional file 2: Table S6. Code used in this study is available under the Creative Commons Attribution 4.0 International License (CC BY 4.0) on GitHub [123] or on Zenodo [124].

Declarations

Ethics approval and consent to participate

UKBB data were accessed through application #16389. UKBB has approval from the North West Multi-centre Research Ethics Committee as a Research Tissue Bank and all participants signed a broad informed consent form.

Competing interests

The authors declare no competing interests.

Received: 1 November 2023 Accepted: 7 May 2024

Published online: 17 May 2024

References

- Niccoli T, Partridge L. Ageing as a Risk Factor for Disease. *Curr Biol*. 2012;22(17):R741–52. <https://doi.org/10.1016/j.cub.2012.07.024>.

2. Diebel LWM, Rockwood K. Determination of Biological Age: Geriatric Assessment vs Biological Biomarkers. *Curr Oncol Rep.* 2021;23(9):104–8. <https://doi.org/10.1007/s11912-021-01097-9>.
3. Horvath S. DNA Methylation Age of Human Tissues and Cell Types. *Genome Biol.* 2013;14(10):3156. <https://doi.org/10.1186/gb-2013-14-10-r115>.
4. Hannum G, Guinney J, Zhao L, Zhang L, Hughes G, Sada S, et al. Genome-wide Methylation Profiles Reveal Quantitative Views of Human Aging Rates. *Mol Cell.* 2013;49(2):359–67. <https://doi.org/10.1016/j.molcel.2012.10.016>.
5. Salameh Y, Bejaoui Y, El Hajj N. DNA Methylation Biomarkers in Aging and Age-Related Diseases. *Front Genet.* 2020;11:480672. <https://doi.org/10.3389/fgene.2020.00171>.
6. Jylhävä J, Pedersen NL, Hägg S. Biological Age Predictors. *EBioMedicine.* 2017;21:29–36. <https://doi.org/10.1016/j.ebiom.2017.03.046>.
7. Shammas MA. Telomeres, lifestyle, cancer, and aging. *Curr Opin Clin Nutr Metab Care.* 2011;14(1):28. <https://doi.org/10.1097/MCO.0b013e32834121b1>.
8. Astuti Y, Wardhana A, Watkins J, Wulaningsih W, For the Pilar Research Network. Cigarette smoking and telomere length: A systematic review of 84 studies and meta-analysis. *Environ Res.* 2017;158:480. <https://doi.org/10.1016/j.envres.2017.06.038>.
9. Song S, Lee E, Kim H. Does Exercise Affect Telomere Length? A Systematic Review and Meta-Analysis of Randomized Controlled Trials. *Medicina.* 2022;58(2). <https://doi.org/10.3390/medicina58020242>.
10. Paul L. Diet, nutrition and telomere length. *J Nutr Biochem.* 2011;22(10):895–901. <https://doi.org/10.1016/j.jnutbio.2010.12.001>.
11. Vidaček NŠ, Nanić L, Ravlić S, Sopta M, Gerić M, Gajski G, et al. Telomeres, Nutrition, and Longevity: Can We Really Navigate Our Aging? *J Gerontol Ser A.* 2018;73(1):39–47. <https://doi.org/10.1093/gerona/glx082>.
12. Stanley SE, Merck SJ, Armanios M. Telomerase and the Genetics of Emphysema Susceptibility. Implications for Pathogenesis Paradigms and Patient Care. *Ann Am Thorac Soc.* 2016;13(Suppl 5):S447. <https://doi.org/10.1513/AnnalsATS.201609-718AW>.
13. Ameh OI, Okpechi IG, Dandara C, Kengne AP. Association Between Telomere Length, Chronic Kidney Disease, and Renal Traits: A Systematic Review. *OMICS J Integr Biol.* 2017;21(3):143–55. <https://doi.org/10.1089/omi.2016.0180>.
14. Rossiello F, Jurk D, Passos JF, D'Adda di Fagnana F. Telomere dysfunction in ageing and age-related diseases. *Nat Cell Biol.* 2022;24(2):135–47. <https://doi.org/10.1038/s41556-022-00842-x>.
15. Maciejowski J, de Lange T. Telomeres in cancer: tumour suppression and genome instability. *Nat Rev Mol Cell Biol.* 2017;18(3):175–86. <https://doi.org/10.1038/nrm.2016.171>.
16. Zhang JM, Zou L. Alternative lengthening of telomeres: from molecular mechanisms to therapeutic outlooks. *Cell Biosci.* 2020;10(1):1–9. <https://doi.org/10.1186/s13578-020-00391-6>.
17. Aviv A, Anderson JJ, Shay JW. Mutations, Cancer and the Telomere Length Paradox. *Trends Cancer.* 2017;3(4):253–8. <https://doi.org/10.1016/j.trecan.2017.02.005>.
18. Allaire P, He J, Mayer J, Moat L, Gerstenberger P, Wilhorn R, et al. Genetic and clinical determinants of telomere length. *Hum Genet Genomics Adv.* 2023;4(3). <https://doi.org/10.1016/j.xhgg.2023.100201>.
19. Codd V, Denniff M, Swinfield C, Warner SC, Papakonstantinou M, Sheth S, et al. Measurement and initial characterization of leukocyte telomere length in 474,074 participants in UK Biobank. *Nat Aging.* 2022;2(2):170–9. <https://doi.org/10.1038/s43587-021-00166-9>.
20. Codd V, Wang Q, Allara E, Musicha C, Kaptoge S, Stoma S, et al. Polygenic basis and biomedical consequences of telomere length variation. *Nat Genet.* 2021;53(10):1425–33. <https://doi.org/10.1038/s41588-021-00944-6>.
21. Wang W, Huang N, Zhuang Z, Song Z, Li Y, Dong X, et al. Identifying Potential Causal Effects of Telomere Length on Health Outcomes: A Phenome-Wide Investigation and Mendelian Randomization Study. *J Gerontol Ser A.* 2023;glad128. <https://doi.org/10.1093/gerona/glad128>.
22. Gardner M, Bann D, Wiley L, Cooper R, Hardy R, Nitsch D, et al. Gender and telomere length: Systematic review and meta-analysis. *Exp Gerontol.* 2014;51:15–27. <https://doi.org/10.1016/j.exger.2013.12.004>.
23. Davis-Kean PE, Tighe LA, Waters NE. The Role of Parent Educational Attainment in Parenting and Children's Development. *Curr Dir Psychol Sci.* 2021;30(2):186–92. <https://doi.org/10.1177/0963721421993116>.
24. Nakao T, Bick AG, Taub MA, Zekavat SM, Uddin MM, Niroula A, et al. Mendelian randomization supports bidirectional causality between telomere length and clonal hematopoiesis of indeterminate potential. *Sci Adv.* 2022;8(14). <https://doi.org/10.1126/sciadv.abl6579>.
25. DeBoy EA, Tassia MG, Schratz KE, Yan SM, Cosner ZL, McNally EJ, et al. Familial Clonal Hematopoiesis in a Long Telomere Syndrome. *N Engl J Med.* 2023;388(26):2422–33. <https://doi.org/10.1056/NEJMoa2300503>.
26. Kessler MD, Damask A, O'Keefe S, Banerjee N, Li D, Watanabe K, et al. Common and rare variant associations with clonal haematopoiesis phenotypes. *Nature.* 2022;612(7939):301–9. <https://doi.org/10.1038/s41586-022-05448-9>.
27. Sanderson E, Richardson TG, Hemani G, Davey Smith G. The use of negative control outcomes in Mendelian randomization to detect potential population stratification. *Int J Epidemiol.* 2021;50(4):1350–61. <https://doi.org/10.1093/ije/dyaa288>.
28. Chamberlain SR, Cavanagh J, de Boer P, Mondelli V, Jones DNC, Drevets WC, et al. Treatment-resistant depression and peripheral C-reactive protein. *Br J Psychiatr.* 2019;214(1):11–9. <https://doi.org/10.1192/bjp.2018.66>.
29. Pousa PA, Souza RM, Melo PHM, Correa BHM, Mendonça TSC, Simões-e Silva AC, et al. Telomere Shortening and Psychiatric Disorders: A Systematic Review. *Cells.* 2021;10(6):1423. <https://doi.org/10.3390/cells10061423>.
30. Aubert G, Baerlocher GM, Vulto I, Poon SS, Lansdorp PM. Collapse of Telomere Homeostasis in Hematopoietic Cells Caused by Heterozygous Mutations in Telomerase Genes. *PLoS Genet.* 2012;8(5):e1002696. <https://doi.org/10.1371/journal.pgen.1002696>.
31. Lv Z, Cui J, Zhang J. Associations between serum urate and telomere length and inflammation markers: Evidence from UK Biobank cohort. *Front Immunol.* 2022;13:1065739. <https://doi.org/10.3389/fimmu.2022.1065739>.
32. Kurajoh M, Fukumoto S, Yoshida S, Akari S, Murase T, Nakamura T, et al. Uric acid shown to contribute to increased oxidative stress level independent of xanthine oxidoreductase activity in MedCity21 health examination registry. *Sci Rep.* 2021;11(7378):1–9. <https://doi.org/10.1038/s41598-021-86962-0>.

33. Lala V, Zubair M, Minter DA. Liver Function Tests. *PubMed*. 2024. <https://pubmed.ncbi.nlm.nih.gov/29494096>. Accessed 13 May 2024.
34. Patnaik MM, Kamath PS, Simonetto DA. Hepatic manifestations of telomere biology disorders. *J Hepatol*. 2018;69(3):736–43. <https://doi.org/10.1016/j.jhep.2018.05.006>.
35. Scheller Madrid A, Rode L, Nordestgaard BG, Bojesen SE. Short Telomere Length and Ischemic Heart Disease: Observational and Genetic Studies in 290,022 Individuals. *Clin Chem*. 2016;62(8):1140–9. <https://doi.org/10.1373/clinchem.2016.258566>.
36. Rode L, Nordestgaard BG, Bojesen SE. Long telomeres and cancer risk among 95,568 individuals from the general population. *Int J Epidemiol*. 2016;45(5):1634–43. <https://doi.org/10.1093/ije/dyw179>.
37. Wan B, Lu L, Lv C. Mendelian randomization study on the causal relationship between leukocyte telomere length and prostate cancer. *PLoS ONE*. 2023;18(6):e0286219. <https://doi.org/10.1371/journal.pone.0286219>.
38. Vaiserman A, Krasnienkov D. Telomere Length as a Marker of Biological Age: State-of-the-Art, Open Issues, and Future Perspectives. *Front Genet*. 2021;11:630186. <https://doi.org/10.3389/fgene.2020.630186>.
39. Demanelis K, Tong L, Pierce BL. Genetically Increased Telomere Length and Aging-Related Traits in the U.K. Biobank. *J Gerontol: Series A*. 2021;76(1):15–22. <https://doi.org/10.1093/gerona/gz240>.
40. Cai Z, Yan LJ, Ratka A. Telomere Shortening and Alzheimer's Disease. *Neuromol Med*. 2013;15(1):25–48. <https://doi.org/10.1007/s12017-012-8207-9>.
41. Gao K, Wang X, Yue W, Yu H. Exploring the Causal Pathway From Telomere Length to Alzheimer's Disease: An Update Mendelian Randomization Study. *Front Psychiatry*. 2019;10:489035. <https://doi.org/10.3389/fpsy.2019.00843>.
42. Levstek T, Kozjek E, Dolžan V, Trebušak Podkrajšek K. Telomere Attrition in Neurodegenerative Disorders. *Front Cell Neurosci*. 2020;14:556488. <https://doi.org/10.3389/fncel.2020.00219>.
43. Fani L, Hilal S, Sedaghat S, Broer L, Licher S, Arp PP, et al. Telomere Length and the Risk of Alzheimer's Disease: The Rotterdam Study. *J Alzheimers Dis*. 2020;73(2):707–14. <https://doi.org/10.3233/JAD-190759>.
44. Hackenhaar FS, Josefsson M, Adolfsson AN, Landfors M, Kauppi K, Hultdin M, et al. Short leukocyte telomeres predict 25-year Alzheimer's disease incidence in non-APOE ϵ 4-carriers. *Alzheimers Res Ther*. 2021;13(1):1–13. <https://doi.org/10.1186/s13195-021-00871-y>.
45. Yu G, Lu L, Ma Z, Wu S. Genetically Predicted Telomere Length and Its Relationship With Alzheimer's Disease. *Front Genet*. 2021;12:595864. <https://doi.org/10.3389/fgene.2021.595864>.
46. Wang XF, Xu WJ, Wang FF, Leng R, Yang XK, Ling HZ, et al. Telomere Length and Development of Systemic Lupus Erythematosus: A Mendelian Randomization Study. *Arthritis Rheumatol*. 2022;74(12):1984–90. <https://doi.org/10.1002/art.42304>.
47. Liu M, Luo P, Liu L, Wei X, Bai X, Li J, et al. Immune-mediated inflammatory diseases and leukocyte telomere length: A Mendelian randomization study. *Front Genet*. 2023;14:1129247. <https://doi.org/10.3389/fgene.2023.1129247>.
48. Topiwala A, Taschler B, Ebmeier KP, Smith S, Zhou H, Levey DF, et al. Alcohol consumption and telomere length: Mendelian randomization clarifies alcohol's effects. *Mol Psychiatry*. 2022;27(10):4001–8. <https://doi.org/10.1038/s41380-022-01690-9>.
49. Park S, Kim SG, Lee S, Kim Y, Cho S, Kim K, et al. Causal linkage of tobacco smoking with ageing: Mendelian randomization analysis towards telomere attrition and sarcopenia. *J Cachex Sarcopenia Muscle*. 2023;14(2):955–63. <https://doi.org/10.1002/jcsm.13174>.
50. Wan B, Ma N, Lv C. Identifying effects of genetic obesity exposure on leukocyte telomere length using Mendelian randomization. *PeerJ*. 2023;11:e15085. <https://doi.org/10.7717/peerj.15085>.
51. Needham BL, Straight B, Hilton CE, Olungha CO, Lin J. Family socioeconomic status and child telomere length among the Samburu of Kenya. *Soc Sci Med*. 2021;283:114182. <https://doi.org/10.1016/j.socscimed.2021.114182>.
52. Amin V, Fletcher JM, Sun Z, Lu Q. Higher educational attainment is associated with longer telomeres in midlife: Evidence from sibling comparisons in the UK Biobank. *SSM Popul Health*. 2022;17:101018. <https://doi.org/10.1016/j.ssmph.2021.101018>.
53. Rauchbach E, Zeigerman H, Abu-Halaka D, Tirosh O. Cholesterol Induces Oxidative Stress, Mitochondrial Damage and Death in Hepatic Stellate Cells to Mitigate Liver Fibrosis in Mice Model of NASH. *Antioxidants*. 2022;11(3):536. <https://doi.org/10.3390/antiox11030536>.
54. von Zglinicki T. Oxidative stress shortens telomeres. *Trends Biochem Sci*. 2002;27(7):339–44. [https://doi.org/10.1016/S0968-0004\(02\)02110-2](https://doi.org/10.1016/S0968-0004(02)02110-2).
55. Zhu G, Xu J, Guo G, Zhu F. Association between Lipids, Apolipoproteins and Telomere Length: A Mendelian Randomization Study. *Nutrients*. 2023;15(21):4497. <https://doi.org/10.3390/nu15214497>.
56. Aviv A, Shay J, Christensen K, Wright W. The Longevity Gender Gap: Are Telomeres the Explanation? *Sci Aging Knowl Environ*. 2005;2005(23):pe16. <https://doi.org/10.1126/sageke.2005.23.pe16>.
57. Pham H, Thompson-Felix T, Czamara D, Rasmussen JM, Lombroso A, Entringer S, et al. The effects of pregnancy, its progression, and its cessation on human (maternal) biological aging. *Cell Metab*. 2024. <https://doi.org/10.1016/j.cmet.2024.02.016>.
58. Shadyab AH, Gass MLS, Stefanick ML, Waring ME, Macera CA, Gallo LC, et al. Maternal Age at Childbirth and Parity as Predictors of Longevity Among Women in the United States: The Women's Health Initiative. *Am J Public Health*. 2016. <https://doi.org/10.2105/AJPH.2016.303503>.
59. Demanelis K, Jasmine F, Chen LS, Chernoff M, Tong L, Delgado D, et al. Determinants of telomere length across human tissues. *Science (New York, NY)*. 2020;369(6509):eaaz6876. <https://doi.org/10.1126/science.aaz6876>.
60. Tsiampalis T, Panagiotakos DB. Missing-data analysis: socio- demographic, clinical and lifestyle determinants of low response rate on self- reported psychological and nutrition related multi- item instruments in the context of the ATTICA epidemiological study. *BMC Med Res Methodol*. 2020;20(1):1–13. <https://doi.org/10.1186/s12874-020-01038-3>.
61. Sanderson E, Spiller W, Bowden J. Testing and correcting for weak and pleiotropic instruments in two-sample multivariable Mendelian randomization. *Stat Med*. 2021;40(25):5434–52. <https://doi.org/10.1002/sim.9133>.

62. Darrow L, Mounier N, Kutalik Z. Simultaneous estimation of bi-directional causal effects and heritable confounding from GWAS summary statistics. *Nat Commun.* 2021;12(7274):1–15. <https://doi.org/10.1038/s41467-021-26970-w>.
63. Burgess S, Labrecque JA. Mendelian randomization with a binary exposure variable: interpretation and presentation of causal estimates. *Eur J Epidemiol.* 2018;33(10):947–52. <https://doi.org/10.1007/s10654-018-0424-6>.
64. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience.* 2015;4(1):13742–015. <https://doi.org/10.1186/s13742-015-0047-8>.
65. Mölder F, Jablonski KP, Letcher B, Hall MB, Tomkins-Tinch CH, Sochat V, et al. Sustainable data analysis with Snake-make. *F1000Research.* 2021;10(33):33. <https://doi.org/10.12688/f1000research.29032.2>.
66. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature.* 2018;562(7726):203–9. <https://doi.org/10.1038/s41586-018-0579-z>.
67. Wu P, Gifford A, Meng X, Li X, Campbell H, Varley T, et al. Mapping ICD-10 and ICD-10-CM Codes to Phecodes: Workflow Development and Initial Evaluation. *JMIR Med Inform.* 2019;7(4):e14325. <https://doi.org/10.2196/14325>.
68. Auwerx C, Jöeloo M, Sadler MC, Tesio N, Ojavee S, Clark CJ, et al. Rare copy-number variants as modulators of common disease susceptibility. *Genome Med.* 2024;16(1):1–24. <https://doi.org/10.1186/s13073-023-01265-5>.
69. Gao X, Starmer J, Martin ER. A multiple testing correction method for genetic association studies using correlated single nucleotide polymorphisms. *Genet Epidemiol.* 2008;32(4):361–9. <https://doi.org/10.1002/gepi.20310>.
70. Neale Lab UKBB summary statistics. 2024. <http://www.nealelab.is/uk-biobank>. Accessed 13 May 2024.
71. Sadler MC, Auwerx C, Deelen P, Kutalik Z. Multi-layered genetic approaches to identify approved drug targets. *Cell Genomics.* 2023;3(7):100341. <https://doi.org/10.1016/j.xgen.2023.100341>.
72. Kuhn RM, Haussler D, Kent WJ. The UCSC genome browser and associated tools. *Brief Bioinforma.* 2013;14(2):144–61. <https://doi.org/10.1093/bib/bbs038>.
73. The UK10K Consortium. The UK10K project identifies rare variants in health and disease. *Nature.* 2015;526(7571):82–90. <https://doi.org/10.1038/nature14962>.
74. Sadler MC, Auwerx C, Lepik K, Porcu E, Kutalik Z. Quantifying the role of transcript levels in mediating DNA methylation effects on complex traits and diseases. *Nat Commun.* 2022;13(7559):1–14. <https://doi.org/10.1038/s41467-022-35196-3>.
75. Friedman JH, Hastie T, Tibshirani R. Regularization Paths for Generalized Linear Models via Coordinate Descent. *J Stat Softw.* 2010;33(1):1–22. <https://doi.org/10.18637/jss.v033.i01>.
76. Edwards JE, Moore RA. Statins in hypercholesterolaemia: A dose-specific meta-analysis of lipid changes in randomised, double blind trials. *BMC Fam Pract.* 2003;4(1):1–19. <https://doi.org/10.1186/1471-2296-4-18>.
77. van der Graaf A, Zorro MM, Claringbould A, Vöisa U, Aguirre-Gamboa R, Li C, et al. Systematic Prioritization of Candidate Genes in Disease Loci Identifies TRAFD1 as a Master Regulator of IFN γ Signaling in Celiac Disease. *Front Genet.* 2021;11:562434. <https://doi.org/10.3389/fgene.2020.562434>.
78. Hemani G, Zheng J, Elsworth B, Wade KH, Haberland V, Baird D, et al. The MR-Base platform supports systematic causal inference across the human phenome. *eLife.* 2018. <https://doi.org/10.7554/eLife.34408>.
79. Verbanck M, Chen CY, Neale B, Do R. Detection of widespread horizontal pleiotropy in causal relationships inferred from Mendelian randomization between complex traits and diseases. *Nat Genet.* 2018;50(5):693–8. <https://doi.org/10.1038/s41588-018-0099-7>.
80. Sanderson E, Glymour MM, Holmes MV, Kang H, Morrison J, Munafò MR, et al. Mendelian randomization. *Nat Rev Methods Prim.* 2022;2(6):1–21. <https://doi.org/10.1038/s43586-021-00092-5>.
81. Burgess S, Davies NM, Thompson SG. Bias due to participant overlap in two-sample Mendelian randomization. *Genet Epidemiol.* 2016;40(7):597–608. <https://doi.org/10.1002/gepi.21998>.
82. Mounier N, Kutalik Z. Bias correction for inverse variance weighting Mendelian randomization. *Genet Epidemiol.* 2023;47(4):314–31. <https://doi.org/10.1002/gepi.22522>.
83. Hu X, Zhao J, Lin Z, Wang Y, Peng H, Zhao H, et al. Mendelian randomization for causal inference accounting for pleiotropy and sample structure using genome-wide summary statistics. *Proc Natl Acad Sci.* 2022;119(28):e2106858119. <https://doi.org/10.1073/pnas.2106858119>.
84. Zheng J, Erzurumluoglu AM, Elsworth BL, Kemp JP, Howe L, Haycock PC, et al. LD Hub: a centralized database and web interface to perform LD score regression that maximizes the potential of summary level GWAS data for SNP heritability and genetic correlation analysis. *Bioinformatics.* 2017;33(2):272–9. <https://doi.org/10.1093/bioinformatics/btw613>.
85. McCarty CA, Wilke RA, Giampietro PF, Westbrook SD, Caldwell MD. Marshfield Clinic Personalized Medicine Research Project (PMRP): design, methods and recruitment for a large population-based biobank. *Personalized Med.* 2005. <https://doi.org/10.1517/17410541.2.1.49>.
86. Pulley J, Clayton E, Bernard GR, Roden DM, Masys DR. Principles of Human Subjects Protections Applied in an Opt-Out, De-identified Biobank. *Clin Transl Science.* 2010;3(1):42–8. <https://doi.org/10.1111/j.1752-8062.2010.00175.x>.
87. Timmers PR, Mounier N, Lall K, Fischer K, Ning Z, Feng X, et al. Genomics of 1 million parent lifespans implicates novel pathways and common diseases and distinguishes survival chances. *eLife.* 2019. <https://doi.org/10.7554/eLife.39856>.
88. Carter AR, Sanderson E, Hammerton G, Richmond RC, Davey Smith G, Heron J, et al. Mendelian randomisation for mediation analysis: current methods and challenges for implementation. *Eur J Epidemiol.* 2021;36(5):465–78. <https://doi.org/10.1007/s10654-021-00757-1>.
89. Sulc J, Sonrel A, Mounier N, Auwerx C, Marouli E, Darrow L, et al. Composite trait Mendelian randomization reveals distinct metabolic and lifestyle consequences of differences in body shape. *Commun Biol.* 2021;4(1064):1–13. <https://doi.org/10.1038/s42003-021-02550-y>.
90. Bowden J, Hemani G, Davey Smith G. Invited Commentary: Detecting Individual and Global Horizontal Pleiotropy in Mendelian Randomization—A Job for the Humble Heterogeneity Statistic? *Am J Epidemiol.* 2018;187(12):2681–5. <https://doi.org/10.1093/aje/kwy185>.
91. Pain O. European LD scores from 1000 Genomes. Zenodo. 2023. <https://doi.org/10.5281/zenodo.8182036>.

92. Bellenguez C, Küçükali F, Jansen IE, Kleiheidam L, Moreno-Grau S, Amin N, et al. New insights into the genetic etiology of Alzheimer's disease and related dementias. *Nat Genet.* 2022;54(4):412–36. <https://doi.org/10.1038/s41588-022-01024-z>.
93. Schoeler T, Speed D, Porcu E, Pirastu N, Pingault JB, Kutalik Z. Participation bias in the UK Biobank distorts genetic associations and downstream analyses. *Nat Hum Behav.* 2023;7(7):1216–27. <https://doi.org/10.1038/s41562-023-01579-9>.
94. Kurki MI, Karjalainen J, Palta P, Sipilä TP, Kristiansson K, Donner KM, et al. FinnGen provides genetic insights from a well-phenotyped isolated population. *Nature.* 2023;613(7944):508–18. <https://doi.org/10.1038/s41586-022-05473-8>.
95. Han Y, Jia Q, Jahani PS, Hurrell BP, Pan C, Huang P, et al. Genome-wide analysis highlights contribution of immune system pathways to the genetic architecture of asthma. *Nat Commun.* 2020;11(1776):1–13. <https://doi.org/10.1038/s41467-020-15649-3>.
96. Yap CX, Sidorenko J, Wu Y, Kemper KE, Yang J, Wray NR, et al. Dissection of genetic variation and evidence for pleiotropy in male pattern baldness. *Nat Commun.* 2018;9(5407):1–12. <https://doi.org/10.1038/s41467-018-07862-y>.
97. Michailidou K, Lindström S, Dennis J, Beesley J, Hui S, Kar S, et al. Association analysis identifies 65 new breast cancer risk loci. *Nature.* 2017;551(7678):92–4. <https://doi.org/10.1038/nature24284>.
98. Mullins N, Forstner AJ, O'Connell KS, Coombes B, Coleman JRI, Qiao Z, et al. Genome-wide association study of more than 40,000 bipolar disorder cases provides new insights into the underlying biology. *Nat Genet.* 2021;53(6):817–29. <https://doi.org/10.1038/s41588-021-00857-4>.
99. Choquet H, Melles RB, Anand D, Yin J, Cuellar-Partida G, Wang W, et al. A large multiethnic GWAS meta-analysis of cataract identifies new risk loci and sex-specific effects. *Nat Commun.* 2021;12(3595):1–12. <https://doi.org/10.1038/s41467-021-23873-8>.
100. Ghodsian N, Abner E, Emdin CA, Gobeil É, Taba N, Haas ME, et al. Electronic health record-based genome-wide meta-analysis provides insights on the genetic architecture of non-alcoholic fatty liver disease. *Cell Rep Med.* 2021;2(11):100437. <https://doi.org/10.1016/j.xcrm.2021.100437>.
101. Wuttke M, Li Y, Li M, Sieber KB, Feitosa MF, Gorski M, et al. A catalog of genetic loci associated with kidney function from analyses of a million individuals. *Nat Genet.* 2019;51(6):957–72. <https://doi.org/10.1038/s41588-019-0407-x>.
102. Nielsen JB, Thorolfsdóttir RB, Fritsche LG, Zhou W, Skov MW, Graham SE, et al. Biobank-driven genomic discovery yields new insight into atrial fibrillation biology. *Nat Genet.* 2018;50(9):1234–9. <https://doi.org/10.1038/s41588-018-0171-3>.
103. Rashkin SR, Graff RE, Kachuri L, Thai KK, Alexeeff SE, Blatchins MA, et al. Pan-cancer study detects genetic risk variants and shared genetic basis in two large cohorts. *Nat Commun.* 2020;11(4423):1–14. <https://doi.org/10.1038/s41467-020-18246-6>.
104. Howard DM, Adams MJ, Clarke TK, Hafferty JD, Gibson J, Shirali M, et al. Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. *Nat Neurosci.* 2019;22(3):343–52. <https://doi.org/10.1038/s41593-018-0326-7>.
105. Stevelink R, Campbell C, Chen S, Abou-Khalil B, Adesoji OM, Afawi Z, et al. GWAS meta-analysis of over 29,000 people with epilepsy identifies 26 risk loci and subtype-specific genetic architecture. *Nat Genet.* 2023;55(9):1471–82. <https://doi.org/10.1038/s41588-023-01485-w>.
106. Gharahkhani P, Jorgenson E, Hysi P, Khawaja AP, Pendergrass S, Han X, et al. Genome-wide meta-analysis identifies 127 open-angle glaucoma loci with consistent effect across ancestries. *Nat Commun.* 2021;12(1258):1–16. <https://doi.org/10.1038/s41467-020-20851-4>.
107. de Lange KM, Moutsianas L, Lee JC, Lamb CA, Luo Y, Kennedy NA, et al. Genome-wide association study implicates immune activation of multiple integrin genes in inflammatory bowel disease. *Nat Genet.* 2017;49(2):256–61. <https://doi.org/10.1038/ng.3760>.
108. van der Harst P, Verweij N. Identification of 64 Novel Genetic Loci Provides an Expanded View on the Genetic Architecture of Coronary Artery Disease. *Circ Res.* 2018. <https://www.ahajournals.org/doi/10.1161/CIRCRESAHA.117.312086>.
109. Howles SA, Wiberg A, Goldsworthy M, Bayliss AL, Gluck AK, Ng M, et al. Genetic variants of calcium and vitamin D metabolism in kidney stone disease. *Nat Commun.* 2019;10(5175):1–10. <https://doi.org/10.1038/s41467-019-13145-x>.
110. Patsopoulos NA, Baranzini SE, Santaniello A, Shoostari P, Cotsapas C, Wong G, et al. Multiple sclerosis genomic map implicates peripheral immune cells and microglia in susceptibility. *Science.* 2019;365(6460). <https://doi.org/10.1126/science.aav7188>.
111. Boer G, Hatzikotoulas K, Southam L, Stefánisdóttir L, Zhang Y, Coutinho de Almeida R, et al. Deciphering osteoarthritis genetics across 826,690 individuals from 9 populations. *Cell.* 2021;184(24):6003–5. <https://doi.org/10.1016/j.cell.2021.11.003>.
112. Phelan CM, Kuchenbaecker KB, Tyrer JP, Kar SP, Lawrenson K, Winham SJ, et al. Identification of 12 new susceptibility loci for different histotypes of epithelial ovarian cancer. *Nat Genet.* 2017;49(5):680–91. <https://doi.org/10.1038/ng.3826>.
113. Schumacher FR, Al Olama AA, Berndt SI, Benlloch S, Ahmed M, Saunders EJ, et al. Association analyses of more than 140,000 men identify 63 new prostate cancer susceptibility loci. *Nat Genet.* 2018;50(7):928–36. <https://doi.org/10.1038/s41588-018-0142-8>.
114. Nalls MA, Blauwendraat C, Vallerga CL, Heilbron K, Bandres-Ciga S, Chang D, et al. Identification of novel risk loci, causal insights, and heritable risk for Parkinson's disease: a meta-analysis of genome-wide association studies. *Lancet Neurol.* 2019;18(12):1091–102. [https://doi.org/10.1016/S1474-4422\(19\)30320-5](https://doi.org/10.1016/S1474-4422(19)30320-5).
115. Okada Y, Wu D, Trynka G, Raj T, Terao C, Ikari K, et al. Genetics of rheumatoid arthritis contributes to biology and drug discovery. *Nature.* 2014;506(7488):376–81. <https://doi.org/10.1038/nature12873>.
116. Pardiñas AF, Holmans P, Pocklington AJ, Escott-Price V, Ripke S, Carrera N, et al. Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. *Nat Genet.* 2018;50(3):381–9. <https://doi.org/10.1038/s41588-018-0059-2>.

117. Pirastu N, Cordioli M, Nandakumar P, Mignogna G, Abdellaoui A, Hollis B, et al. Genetic analyses identify widespread sex-differential participation bias. *Nat Genet.* 2021;53(5):663–71. <https://doi.org/10.1038/s41588-021-00846-7>.
118. Liu M, Jiang Y, Wedow R, Li Y, Brazel DM, Chen F, et al. Association studies of up to 1.2 million individuals yield new insights into the genetic etiology of tobacco and alcohol use. *Nat Genet.* 2019;51(2):237–244. <https://doi.org/10.1038/s41588-018-0307-5>.
119. Malik R, Chauhan G, Traylor M, Sargurupremraj M, Okada Y, Mishra A, et al. Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. *Nat Genet.* 2018;50(4):524–37. <https://doi.org/10.1038/s41588-018-0058-3>.
120. Crouch DJM, Inshaw JRJ, Robertson CC, Zhang JY, Chen WM, Onengut-Gumuscu S, et al. Enhanced genetic analysis of type 1 diabetes by selecting variants on both effect size and significance, and by integration with autoimmune thyroid disease. *bioRxiv.* 2022;2021.02.05.429962. <https://doi.org/10.1101/2021.02.05.429962>.
121. Pulit SL, Stoneman C, Morris AP, Wood AR, Glastonbury CA, Tyrrell J, et al. Meta-analysis of genome-wide association studies. *Hum Mol Genet.* 2019;28(1):166–74. <https://doi.org/10.1093/hmg/ddy327>.
122. Karczewski KJ, Gupta R, Kanai M, et al. Pan-UK Biobank GWAS improves discovery, analysis of genetic architecture, and resolution into ancestry-enriched effects. *medRxiv.* 2024. <https://pan.ukbb.broadinstitute.org/>. <https://doi.org/10.1101/2024.03.13.24303864>. Accessed 13 May 2024.
123. Moix S. GitHub Repository for "Breaking down causes, consequences, and mediating effects of telomere length variation on human health". GitHub. 2024. https://github.com/cChiiper/UNIL_SGG_MR_LTL. Accessed 13 May 2024.
124. Moix S. cChiiper/UNIL_SGG_MR_LTL: v1.0.0. Zenodo. 2024. <https://doi.org/10.5281/zenodo.11089964>. Accessed 13 May 2024.

Publisher's Note

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