Qiu *et al. Journal of Translational Medicine* (2025) 23:283
https://doi.org/10.1186/s12967-024-05932-y

Journal of Translational Medicine

RESEARCH Open Access

COVID-19 infection and longevity: an observational and mendelian randomization study

Shizheng Qiu[1†], Jianhua Liu[2†], Jiahe Guo[1], Zhishuai Zhang[1], Yu Guo[1] and Yang Hu[1*]

Abstract

Background Studies have indicated that COVID-19 infection may accelerate the aging process in organisms. However, it remains unknown whether contracting COVID-19 affects life expectancy. Furthermore, the underlying biological mechanisms behind these findings are still unclear.

Methods We conducted a prospective cohort study on 56,504 participants of European ancestry from the UK Biobank who reported the time and number of COVID-19 infection between January 2020 and September 2023. The parental average longevity was used as a proxy for their own longevity. Linear regression was used to assess the relationship between COVID-19 infection and longevity. Furthermore, we investigated the shared genetic basis between COVID-19 and longevity using large-scale genome-wide association studies (GWAS) for COVID-19 (122,616 cases and 2,475,240 controls) and longevity (3,484 cases and 25,483 controls). Mendelian randomization (MR) and mediation analysis were utilized to assess causal relationships and potential mediators between COVID-19 susceptibility and longevity. Shared genetic loci between the two phenotypes were identified using conjunctional false discovery rate (conjFDR) statistical frameworks.

Results After controlling for relevant covariates, COVID-19 infection might not be significantly correlated with longevity. In all MR methods, generalized summary-data-based Mendelian randomization (GSMR) analysis revealed a significant decrease in longevity due to severe COVID-19 infection (OR = 0.91, 95%CI: 0.84–0.98, $P = 0.015$). Mediation analysis identified stroke and myocardial infarction as potential mediators between COVID-19 susceptibility and reduced longevity. At conjFDR < 0.05, we identified rs62062323 (*KANSL1*) and rs9530111 (*PIBF1*) as shared loci between COVID-19 and longevity.

Conclusion Together, our findings provided preliminary evidence for the shared genetic basics between COVID-19 and aging. This discovery may have implications for personalized medicine and preventive strategies, helping identify individuals who may be more vulnerable to severe outcomes from COVID-19 due to their genetic makeup.

Keywords COVID-19, Longevity, Genome-wide association study, Mendelian randomization, Conjunctional false discovery rate, Mediation analysis

†Shizheng Qiu and Jianhua Liu are contributed equally to this work..

*Correspondence:
Yang Hu
huyang@hit.edu.cn

[1]School of Computer Science and Technology, Harbin Institute of Technology, 92 Xidazhi Street, Nangang District, Harbin 150001, China
[2]Beidahuang Industry Group General Hospital, Harbin 150088, China

Introduction

The COVID-19 pandemic has resulted in a significant global death toll, surpassing 6.7 million as of mid-April 2023. Although COVID-19 mortality rates have notably declined, a range of sequelae can manifest in individuals following recovery from the infection. These sequelae may include persistent health issues such as severe fatigue, breathing difficulties, insomnia, heart palpitations, cognitive impairment, and other symptoms [1–5]. This condition, known as long COVID or post-acute sequelae of SARS-CoV-2 infection (PASC), has the potential to significantly impact the long-term health and longevity of patients [1, 2].

While advancements in socioeconomic development and improved medical care have contributed to increased life expectancy over the past few decades, the COVID-19 pandemic is anticipated to have a negative effect on life expectancy [6, 7]. For instance, in the United States, life expectancy dropped by approximately 1.87 years in 2020 [7]. Similarly, in the United Kingdom, the life expectancy of women in 2020 decreased by 0.9 years, and that of men decreased by 1.2 years when compared to the figures from 2019 [6]. The substantial number of deaths among individuals in their working years can have enduring adverse consequences for society, the economy, and overall health [8]. In addition to the direct fatalities resulting from COVID-19, it remains uncertain whether individuals previously infected with COVID-19 may experience accelerated aging and a reduced life expectancy [9].

Long-term follow-up of COVID-19 patients for several decades would be impractical due to the substantial time and financial resources required. However, a recently developed statistical framework based on summary statistics from genome-wide association studies (GWAS) offers a promising avenue for examining genetic overlap among complex traits. These GWAS cohorts typically encompass tens of thousands or even millions of samples, providing robust statistical power. Additionally, longevity exhibits a high heritability, with genetic factors accounting for approximately one-third of the variation in human lifespan [10]. Given these considerations, genetic data can be leveraged to investigate the shared genetic architecture and specific genetic loci between COVID-19 and longevity.

In this study, we analyzed the relationship between participants who reported COVID-19 infection between January 2020 and September 2023 and lifespan in the UK Biobank. In addition, we utilized mendelian randomization (MR) and mediation analysis to assess the causal relationship between COVID-19 susceptibility and longevity, as well as to examine whether COVID-19 impacts longevity through potential mediators. To identify shared genetic loci, we applied conditional and conjunctional false discovery rate (FDR) statistical frameworks to the most recent GWAS data available for COVID-19 and longevity. By employing these approaches, our objective was to enhance our understanding of the shared genetic basis and underlying molecular mechanisms linking COVID-19 and longevity. The conceptual framework outlining our study is depicted in Fig. 1.

Methods

Study design and participants

The UK Biobank is a large-scale biomedical database and research resource containing in-depth genetic and health information from half a million UK participants aged between 40 and 69 years at recruitment. Participants were recruited between 2006 and 2010, providing a wide range of information about their health, lifestyle, and physical measures.

Phenotypes

Participants who were infected with COVID-19 were defined by the mental well-being online question collected in September 2023: "How many times do you think you have had COVID-19?" (Field 29156). Participants who answered 1 or more times were included. The time of COVID-19 infection was defined by the question in Field 29,157: "When do you think you first had or might have had COVID-19". Participants who did not report the time and number of infections were excluded. As the UK Biobank is a general population cohort, the number of deceased participants is relatively small. Therefore, we used the average of Father's age at death (Field 1807) and Mother's age at death (Field 3526) as a proxy for the participants' lifespan. Participants with missing parental lifespan data and those who answered "Do not know" and "Prefer not to answer" were excluded.

Covariates

We included gender (male as 1, female as 0), age, body mass index (BMI), Townsend deprivation index, education level (having received college/university education or above as 1, others as 0), smoking (answering 'No' as 0, 'Yes, on most or all days' and 'Only occasionally' as 1), alcohol consumption (answering 'Current' and 'Previous' as 1, 'Never' as 0), and diabetes history (answering 'No' as 0, 'Yes' as 1) as covariates. Participants who answered 'Prefer not to answer' or 'Do not know' were excluded.

Statistical analysis

The baseline information of participants was summarized. For continuous traits, the median[minimum to maximum] and mean with standard deviation was calculated, while for categorical variables, the number (percentage) was reported. Linear regression models were used to analyze the correlation between COVID-19 infection and expected lifespan. The first model

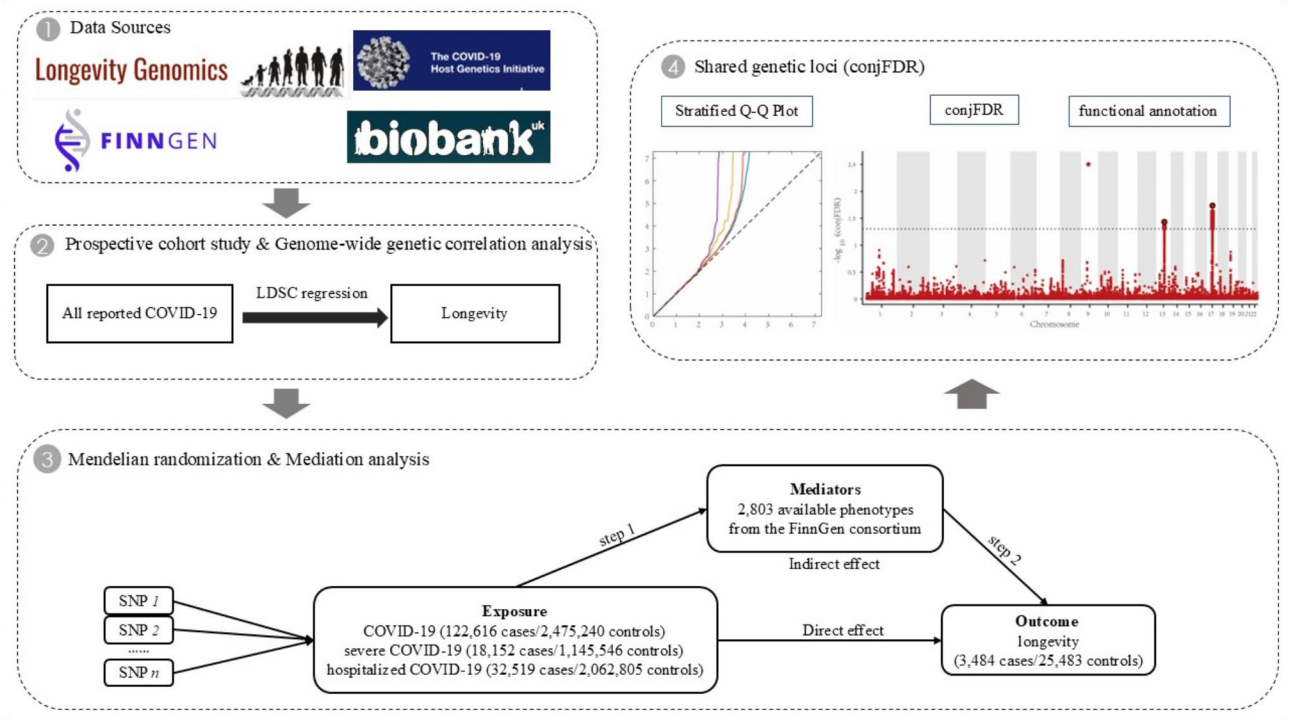

Fig. 1 Study design

considered age and gender as covariates, while the second model considered age, gender, BMI, Townsend deprivation index, education level, smoking, alcohol consumption, and diabetes history as covariates. All statistical analyses were performed using R (version 4.3) and the UK Biobank Research Analysis Platform. A P value < 0.05 (two-sided) was considered statistically significant.

GWAS summary statistics

All GWAS summary statistics in this study was publicly available, and the ethical approval and informed consent were obtained in all original studies.

GWAS summary statistics for COVID-19 (122,616 cases and 2,475,240 controls), very severe respiratory confirmed COVID-19 (18,152 cases and 1,145,546 controls), and hospitalized COVID-19 (32,519 cases and 2,062,805 controls) from the COVID-19 host genetics initiative [11]. The COVID-19 infection cases included individuals with laboratory-confirmed COVID-19 infection or electronic health record, ICD coding or clinically confirmed COVID-19, or self-reported COVID-19 [11, 12]. All participants were of European descent. GWAS for longevity was derived from a GWAS meta-analysis for age of survival of participants from 20 cohorts of European, East Asian, and African American populations [11, 13]. Cases were participants who lived to an age above the 99th percentile (3,484 cases and 25,483 controls) based on cohort life Table [13]. Controls were participants who died at or before the age at the 60th percentile [13]. For instance, the 60th and 99th percentile correspond to ages of 83 and 102 years for women in the United States cohort [13]. We performed quality control on these GWAS data. We focused only on autosomal genetic variants and removed the rare variants with minor allele frequency (MAF) < 0.01 and variants in the HLA region (chr6:26000000–34000000).

GWAS for potential mediators

We obtained GWAS data for potential mediating factors of the effect of COVID-19 on longevity from the FinnGen consortium [14]. The FinnGen consortium was composed of an isolated population with a good phenotype, the Finnish population, and included more than 500,000 participants [14]. We selected 2803 available phenotypes from the FinnGen database to perform a mediation analysis.

Genetic correlation analysis

Based on the pre-computed linkage disequilibrium (LD) scores from the 1000 Genomes projects, we estimated the genetic correlation between COVID-19 and longevity without constraining the intercept using bivariate linkage disequilibrium score (LDSC) regression [15, 16]. The statistically significant association after multiple testing is defined to be $P < 0.05/3 = 0.017$.

MR analysis

MR used genetic variants as instrumental variables (IVs) to assess the causal relationship between exposure and outcome. The randomly allocated process of alleles of genetic variants is used to simulate randomized controlled trials (RCTs) [17, 18]. The alleles of these instruments are determined at the time of meiosis and fertilization, thereby minimizing issues of confounding factors and reverse causality [17]. MR model should meet three assumptions: (1) the IVs are significantly associated with exposure, (2) the IVs are not associated with any potential confounder, and (3) the IVs do not affect outcome independently of exposure [17, 19, 20].

Here, we selected the single nucleotide polymorphisms (SNPs) that were significantly associated with COVID-19 (P-value $< 5 \times 10^{-8}$) as IVs and clumped them by ruling out linkage disequilibrium (LD) with $r^2 < 0.001$ and physical distance within 10 000 kb. We performed a univariate MR analysis to calculate the causal estimates of genetically predicted COVID-19 on longevity using inverse variance weighted (IVW), weighted median, and generalized summary-data-based Mendelian randomization (GSMR) methods [17, 21–25]. We utilized MR-Egger to provide a causal effect through the slope coefficient from Egger regression and test whether IVs had pleiotropy [26–28].

Mediation analysis

We applied a two-step MR to calculate the mediation effect of 2803 potential mediators from the FinnGen consortium [14]. In the first step, we used IVs for COVID-19 to estimate the causal effect of COVID-19 on potential mediators. In the second step, we used IVs for potential mediators to estimate the causal effect of potential mediators on longevity. We assessed the indirect effect of COVID-19 on longevity via each mediator using product of coefficients method [29–31]. The standard error for the indirect effect was derived by using the delta method [32]. Specifically, the product of coefficients method is to test whether the indirect effect is significant: $H_0 = \beta_a \beta_b$. The standard error of indirect effect is represented as $S_{ab} = \sqrt{\beta_a^2 S_b + \beta_b^2 S_a}$. Where β_a represents the causal effect of step 1, β_b represents the causal effect of step 2, S_a is the standard error of β_a, S_b is the standard error of β_b. The statistically significant pathway is defined to be $P < 0.05$.

Polygenic overlap

We constructed conditional quantile–quantile (Q-Q) plots to visualize the cross-trait pleiotropic enrichment in SNPs associations between COVID-19 and longevity [33–35]. The conditional Q-Q plots compare the association with primary phenotype (e.g. longevity) within SNPs strata determined by the strength of association with a secondary trait (e.g. COVID-19) [36]. As the proportion of SNPs associated with the primary phenotype increased in association with the secondary phenotype, this cross-trait enrichment shifted continuously upward from the null line in the conditional Q-Q plots [36, 37].

conjFDR statistical framework

We applied condFDR and conjFDR statistical framework to identify the shared genetic loci of COVID-19 and longevity [33, 34, 37]. FDR is the probability that a SNP is null given that its P-value is smaller than its observed P-value [38]. The condFDR is an extension of standard FDR, which re-ranks the test-statistics of the primary phenotype (e.g. longevity) based on a conditional variable, in this case the strength of the association with the secondary phenotype (e.g. COVID-19) [33, 39]. The conjFDR is an extension of condFDR, defined as the maximum of the two condFDR statistics for a specific SNP [33, 40]. Under the premise that the P value of both traits is less than the significance threshold, condFDR statistics estimates the posterior probability that a SNP is null for either trait or both [33, 34]. P-values are adjusted for inflation using a genomic inflation control procedure. The statistically significant loci are defined to conjFDR < 0.05. Further details of the methods could be found in Supplementary Methods.

Genomic loci definition and directions of effect sizes

Independent significant SNPs were identified as SNPs with LD $r^2 < 0.6$ and LD blocks < 250 kb using the 1000 Genomes Project reference panel [41, 42]. We assessed the effect directions of the shared loci between COVID-19 and longevity by comparing their z-scores from the original GWAS. Where i represents the shared loci, β_i represents the effect coefficient of i, se_i is the standard error of i.

$$Z_i = \beta_i \Big/ se_i$$

Functional annotation

We identified genes whose expression were associated with the shared SNPs at conjFDR < 0.05 in expression quantitative trait loci (eQTL) from the GTEx version 8 (http://gtexportal.org) [43]. P < 1E-05 was considered as significant associations after multiple testing.

Results

Baseline characteristics

A total of 56,504 participants of European ancestry were included in the study, of which 28,309 (56.05%) were women (Table 1). The median BMI was 26.3 kg/m² (13.7–59.1), and the median Townsend deprivation index

Table 1 Characteristics of study participants

Variable	Count (%)	Median [Min-Max]
Sex		
Female	28,309 (56.05%)	-
Male	22,195 (43.95%)	-
Age	-	60 [39 to 72]
BMI	-	26.3 [13.7 to 59.1]
Townsend deprivation index	-	-2.69 [-6.26 to 10.46]
University/college education		
No	29,597 (58.6%)	-
Yes	20,907 (41.4%)	-
Alcohol consumption		
No	1278 (2.53%)	-
Yes	49,226 (97.47%)	-
Smoking		
No	47,730 (94.51%)	-
Yes	2774 (5.49%)	-
Diabetes		
No	48,859 (96.74%)	-
Yes	1645 (3.26%)	-
COVID-19 infection		
No	15,574 (30.84%)	-
Yes	34,930 (69.16%)	-
Longevity	-	75 [24.5 to 101]

was – 2.69 (-6.26 to 10.46). Regarding education, 58.6% (29,597) did not have a university or college degree, while 41.4% (20,907) did. Most participants reported current or ever alcohol consumption (97.47%, 49,226), and 5.49% (2,774) were current or ever smokers. Diabetes was present in 3.26% (1,645) of participants. Regarding COVID-19 infection, 69.16% (34,930) had been infected (until September 2023). The median parental longevity was 75 years (24.5–101).

Linear regression analysis

Two linear regression models were constructed to assess the association between COVID-19 infection and longevity, adjusting for various covariates. COVID-19 infection was not significantly associated with longevity in model 1 (β = 0.091, SE = 0.090, $p = 0.312$) (Table 2). In model 2, COVID-19 infection was also not significantly associated with longevity after adjusting for covariates (β = 0.102, SE = 0.089, $p = 0.255$). BMI (β = -0.118, SE = 0.010, $p < 0.001$), Townsend deprivation index (β = -0.065, SE = 0.016, $p < 0.001$), current or ever smoking (β = -0.374, SE = 0.181, $p = 0.0385$), and diabetes (β = -1.146, SE = 0.233, $p < 0.001$) were significantly negatively associated with longevity, and educational attainment was positively associated with longevity (β = 1.400, SE = 0.084, $p < 0.001$).

Estimating total genetic overlap between COVID-19 and longevity

We utilized bivariate LDSC regression to estimate the genetic correlation between three COVID-19 subtypes and longevity. The critically ill COVID-19 group (r_g = -0.34, $P = 0.0003$) and hospitalized COVID-19 group (r_g = -0.35, $P = 1.81E-05$) were significantly associated with longevity (Supplementary Table 1). The reported COVID-19 infection group was suggestively associated with longevity (r_g = -0.22, $P = 0.023$).

Direct effect and indirect effect of COVID-19 on longevity

Given the substantial genetic overlap between COVID-19 and longevity, we evaluated the causal effects of genetically predicted COVID-19 susceptibility on longevity. Univariate MR analysis showed a non-significant causal association between genetically predicted COVID-19 and longevity ($P > 0.05$) (Table 3). The *P*-values for the causal effects of both hospitalized COVID-19 (OR = 0.86, 95%CI: 0.74-1.00, $P = 0.055$) and severe COVID-19 (OR = 0.94, 95%CI: 0.88–1.01, $P = 0.077$) on longevity were close to significant. Importantly, when incorporating additional single nucleotide polymorphisms (SNPs) as genetic instruments, GSMR analysis revealed

Table 2 Linear regression models of the association between COVID-19 infection and longevity

	Model 1			Model 2		
Variable	**Estimate**	**std.error**	***P* value**	**Estimate**	**std.error**	***P* value**
Sex	0.032	0.083	0.701	0.033	0.084	0.693
Age	0.388	0.007	< 1E-100	0.387	0.007	< 1E-100
BMI	-	-	-	-0.118	0.010	1.55E-34
Townsend deprivation index	-	-	-	-0.065	0.016	3.11E-05
University/college education	-	-	-	1.400	0.084	1.86E-62
Alcohol Consumption	-	-	-	0.226	0.261	0.388
Smoking	-	-	-	-0.374	0.181	0.0385
Diabetes	-	-	-	-1.146	0.233	9.16E-07
COVID19 Infection	0.091	0.090	0.312	0.102	0.089	0.255
Intercept	50.741	0.435	< 1E-100	53.067	0.582	< 1E-100

Model 1 were adjusted for age and gender. Model 2 were adjusted for age, gender, BMI, Townsend deprivation index, education level, smoking, alcohol consumption, and diabetes history as covariates

Table 3 Causal effects of COVID-19 on longevity in univariate MR study

Exposure	Method	No. SNP	OR (95%CI)	*P* value
COVID-19	IVW	14	0.78 (0.55-)1.11	0.17
	Weighted median	14	0.71 (0.45–1.13)	0.15
	GSMR	29	0.78 (0.57–1.06)	0.12
	MR Egger	14	0.94 (0.49–1.79)	0.85
very severe COVID-19	IVW	28	0.94 (0.88–1.01)	0.077
	Weighted median	28	0.94 (0.86–1.02)	0.14
	GSMR	49	0.91 (0.84–0.98)	0.015
	MR Egger	28	0.92 (0.82–1.04)	0.20
hospitalized COVID-19	IVW	32	0.86 (0.74-1.00)	0.055
	Weighted median	32	0.88 (0.73–1.07)	0.21
	GSMR	46	0.91 (0.81–1.02)	0.11
	MR Egger	32	0.89 (0.67–1.18)	0.43

The statistically significant association is defined to be $P < 0.05/3 = 0.0167$ after multiple testing

Table 4 The mediation effect of COVID-19 on longevity via potential mediators

Mediator	Indirect effect β (95% CI)	Mediated proportion (%)	*P* value
Major coronary heart disease event	-0.042 (-0.088 to 0.0030)	17.33 (0 to 35.90)	0.067
Stroke	-0.17 (-0.33 to -0.014)	70.28 (5.69 to 100)	0.033
Hard cardiovascular diseases	-0.080 (-0.17 to 0.014)	32.71 (0 to 71.06)	0.095
Heart fail and coronary heart disease	-0.12 (-0.25 to 0.0048)	50.57 (0- 100)	0.059
Myocardial infarction (no controls excluded)	-0.076 (-0.17 to 0.014)	31.01 (0 to 67.61)	0.097
Myocardial infarction	-0.084 (-0.16 to -0.0040)	34.41 (1.64–67.17)	0.040
Myocardial infarction (strict)	-0.068 (-0.14 to 0.0038)	27.76 (0 to 57.09)	0.064

Table 5 Shared loci between COVID-19 and longevity

Phenotypic pairing	SNP	Gene	CHR	conjFDR
COVID-19 & longevity	rs62062323	*KANSL1*	17	0.038
critically ill COVID-19 & longevity	rs62062323	*KANSL1*	17	0.0198
hospitalized COVID-19 & longevity	rs62062323	*KANSL1*	17	0.0184
	rs9530111	*PIBF1*	13	0.037

a significant decrease in longevity due to severe COVID-19 infection (OR = 0.91, 95%CI: 0.84–0.98, $P = 0.015$). MR-Egger intercept analysis indicated no evidence of pleiotropy.

In order to explore whether COVID-19 susceptibility would reduce longevity via potential mediators, we conducted a two-step MR. Among the 2803 available phenotypes from the FinnGen consortium, we identified seven phenotypes that were significant in two-step MR, i.e., COVID-19 had a causal effect on longevity via these phenotypes (Table 4). In mediation analysis, we found that stroke and myocardial infarction were significant mediating factors and the mediation effects of stroke and myocardial infarction were − 0.17 (95% CI, -0.33 to -0.014; $P = 0.033$) and − 0.084 (95% CI, -0.16 to -0.0040; $P = 0.040$) with a mediated proportion of 70.28% (95% CI, 5.69 to 100%) and 34.41% (95% CI, 1.64 to 67.17%), respectively (Table 4).

Visualization of cross-trait enrichment

We generated conditional Q-Q plots to visualize pleiotropic enrichment between COVID-19 and longevity. We observed successive increments of SNP enrichment for COVID-19 as a function of the significance of the associations with longevity, and vice versa, suggesting polygenic overlap between them (Supplementary Fig. 1).

Identifying shared genetic loci

At conjFDR < 0.05, we identified 1, 1, and 2 genomic loci jointly associated with COVID-19 infection, critically ill COVID-19, hospitalized COVID-19 and longevity, respectively (Table 5; Fig. 2). rs62062323 (*KANSL1*) was identified across the three phenotypic pairings simultaneously, and rs9530111 (*PIBF1*) was unique shared loci of hospitalized COVID-19 and longevity. *KANSL1* and *PIBF1* were novel for both COVID-19 and longevity, suggesting the ability of conjFDR statistical framework to boost the power to discover novel loci. Furthermore, all these lead SNPs had the opposite effect direction on COVID-19 and longevity (Supplementary Tables 2–4).

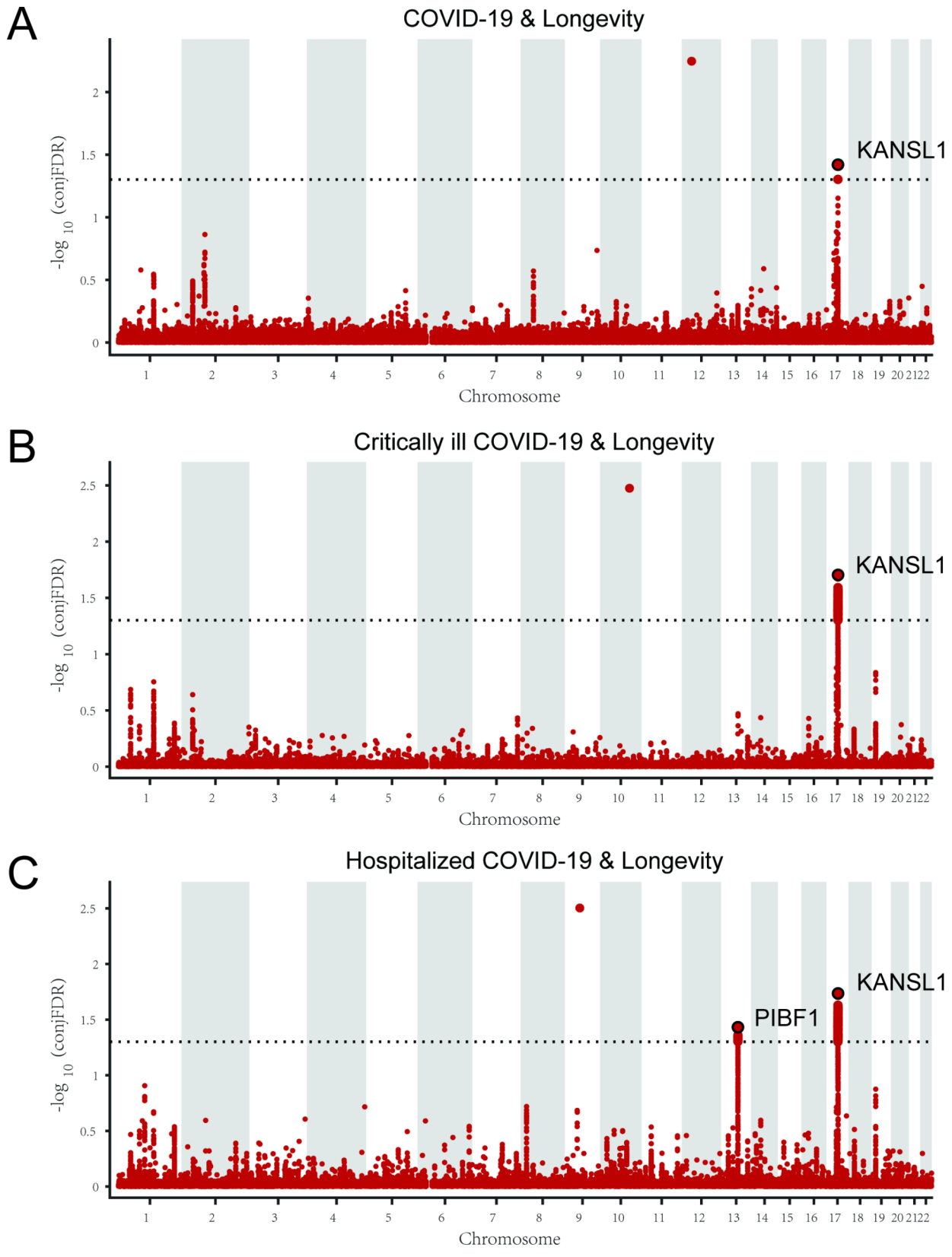

Fig. 2 Manhattan plots of shared genetic variants associated with both COVID-19 and longevity at conjFDR < 0.05. A black circle around the enlarged data points indicates the lead-SNPs in each LD block

The opposite effect directions were consistent with the negative genetic correlations.

Functional annotation

The genetic loci shared by COVID-19 and longevity (rs62062323 and rs9530111) were both located in intron, so we further annotated them in the Genotype-Tissue Expression (GTEx). rs62062323 was significantly associated with altered expression of 41 genes in 49 human tissues, including brain, lung, and whole blood (Supplementary Table 5). rs9530111 was significantly associated with altered expression of *KLF5*, *PIBF1*, and *DIS3* in cultured fibroblasts, spleen, and thyroid (Supplementary Table 5).

Discussion

In this study, we employed a novel statistical framework to investigate the genetic overlap between COVID-19 susceptibility and longevity, and to identify the shared genetic loci underlying this overlap. Our analysis indicated that COVID-19 might indirectly influence the likelihood of attaining longevity by increasing the risk of stroke and myocardial infarction. Moreover, we identified two novel genomic loci that exhibited joint associations with COVID-19 susceptibility and longevity at conjFDR < 0.05. By shedding light on how COVID-19 susceptibility affects longevity, our study provides valuable insights into the underlying mechanisms and genetic factors involved in this relationship.

Previous studies have primarily focused on assessing the impact of the COVID-19 pandemic on life expectancy by examining population-level mortality resulting directly or indirectly from the virus. However, it is important to consider the long-term effects experienced by COVID-19 patients, particularly those who have been hospitalized or critically ill. These individuals may suffer from Long COVID. Some of the long-term consequences of COVID-19, such as cognitive impairment and cardiovascular diseases, have the potential to shorten life expectancy. For instance, MR analyses utilizing genetic data have indicated that COVID-19 infection raises the risk of developing dementia and psychiatric disorders and dementia has been associated with a reduction in longevity [44, 45]. Additionally, there is evidence suggesting that individuals who have been infected with COVID-19 may experience accelerated aging processes [9]. Consequently, in addition to the direct impact of COVID-19-related deaths on life expectancy, COVID-19 can also contribute to a decrease in life expectancy through the long-term health effects.

For the first time, we assess the shared genetic basis of COVID-19 and longevity based on a genetic framework. We conducted a MR analysis to examine the causal relationship between susceptibility to COVID-19 and longevity and identify the potential mediating factors within a separate cohort, the Finnish population. Our findings revealed that while susceptibility to COVID-19 might not exhibit a direct causal effect on longevity (only significant in GSMR), it did exert an indirect influence by increasing the risk of stroke and myocardial infarction. These cardiovascular conditions were identified as mediating factors that contribute to the impact of COVID-19 on longevity. The impact of COVID-19 on the cardiovascular system has been widely reported [46–50]. The underlying mechanisms are primarily attributed to inflammatory responses, coagulation abnormalities, and vascular endothelial damage [51–56]. The SARS-CoV-2 virus enters host cells via the ACE2 receptor, which is widely expressed in vascular endothelial cells. SARS-CoV-2 infection and the associated inflammatory response can damage endothelial cells, impairing endothelium-dependent vasodilation while promoting vasoconstriction, inflammation, and thrombosis. These effects may significantly increase the risk of stroke and myocardial infarction.

Of the two novel loci found to be shared between COVID-19 and longevity, one is *PIBF* (progesterone-induced blocking factor), a protein that plays a role in regulating multiple immune responses [57, 58]. *PIBF* has been shown to modulate the cytokine IL-6, which has been observed to be increased in COVID-19 patients [59]. Interestingly, inhibiting PIBF through immunotherapy has been associated with extended longevity in tumor patients [60, 61]. The second shared locus is *KANSL1* (KAT8 regulatory NSL complex subunit 1), which encodes a protein involved in chromatin remodeling. This locus is also considered a susceptibility locus for several age-related neurodegenerative diseases, including Alzheimer's disease and Parkinson's disease [62–64]. While there is no known direct association between *KANSL1*, COVID-19, and longevity, previous studies have indicated that COVID-19 infection may contribute to aging processes or cognitive decline in the human brain [65, 66]. These findings highlight the potential relevance of the identified loci, particularly *PIBF* and *KANSL1*, in understanding the shared genetic mechanisms underlying COVID-19 susceptibility and longevity.

However, it is important to acknowledge the limitations of our study. Firstly, since the UK Biobank is a general population cohort and most participants are still alive, we can only use the parental average longevity as a proxy for the participants' own longevity. The lifespan of participants is highly heritable with respect to their parents' lifespan. Using parental lifespan as a proxy for an individual's lifespan is common in previous studies [67–70]. If COVID-19 infection has a genetic component, then the genetic variants or genes that make participants more susceptible to COVID-19 infection

might also be highly correlated with lifespan. However, our study focused solely on genetic variants identified through GWAS data, without considering other potential factors that could contribute to COVID-19 susceptibility or longevity, such as environmental or lifestyle factors. Genetic factors may account for only a small portion of all factors related to lifespan. Secondly, there is a certain degree of sample overlap between the COVID-19 and longevity cohorts, specifically in the UK Biobank cohort. Fortunately, this sample overlap had minimal impact on the results of our study. Thirdly, our study primarily investigated the genetic overlap between COVID-19 susceptibility and longevity in individuals of European ancestry. Therefore, the findings may not generalize to individuals of other ancestries. Lastly, the two shared loci we identified, which influence the longevity of COVID-19 patients, require further validation through biological experiments. Notably, rs62062323 regulates gene expression in 49 human tissues, suggesting that COVID-19 may accelerate biological aging across multiple human tissues or organs.

In conclusion, while susceptibility to COVID-19 might not exhibit a direct effect on longevity, it might exert an indirect influence by increasing the risk of stroke and myocardial infarction. The identification of two shared loci holds promise as potential targets for interventions aimed at addressing the long-term health consequences of "Long COVID". Individuals with genetic variants associated with higher risk could be prioritized for targeted interventions. We reiterate that the influence of genetic factors on lifespan in this study accounts for only a small portion of the overall effect, and genetic determinism should be avoided.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12967-024-05932-y.

Supplementary Material 1

Supplementary Material 2

Supplementary Material 3

Acknowledgements

We thank the UK Biobank, FinnGen, the Longevity Genomics research group and the COVID-19 host genetics initiative for providing data.

Author contributions

YH, JHL, ZSZ, and YG designed the study. SZQ and JHG analyzed the data. SZQ wrote the first draft of the manuscript. All authors had full access to all the data in the study and had final responsibility for the decision to submit for publication.

Funding

This work was supported by the National Natural Science Foundation of China (62371161), Heilongjiang Provincial Science and Technology Tackling Project (GNCMSSJH2024), and 0-1 Original Exploration Category: Fundamental Research Funds for the Central Universities Project (2022FRFK030025).

Data availability

Access to the UK Biobank data can be requested through a standard protocol (https://www.ukbiobank.ac.uk/register-apply/). Data used in this study are available in the UK Biobank under application number 249728.
GWAS for COVID-19: https://www.covid19hg.org/results/r7/.
GWAS for longevity: https://www.longevitygenomics.org/downloads.

Declarations

Ethics approval and consent to participate

Previous datasets have been approved by their respective ethical approval committees.

Consent for publication

Not applicable.

Competing interests

NA.

Received: 4 December 2023 / Accepted: 30 November 2024
Published online: 06 March 2025

References

1. Davis HE, et al. Long COVID: major findings, mechanisms and recommendations. Nat Rev Microbiol. 2023;21(3):133–46.
2. Akbarialiabad H, et al. Long COVID, a comprehensive systematic scoping review. Infection. 2021;49(6):1163–86.
3. Monje M, Iwasaki A. The neurobiology of long COVID. Neuron. 2022;110(21):3484–96.
4. Yong SJ. Long COVID or post-COVID-19 syndrome: putative pathophysiology, risk factors, and treatments. Infect Dis (Lond). 2021;53(10):737–54.
5. Koc HC, et al. Long COVID and its management. Int J Biol Sci. 2022;18(12):4768–80.
6. Menassa M, Vriend EM, Franco OH. Healthy longevity in the time of COVID-19: a conceptual framework. Lancet Healthy Longev. 2021;2(5):e243–4.
7. Barbieri M. Covid-19 and the growing disadvantage in US life expectancy. BMJ. 2021;373:n1530.
8. Aburto JM, et al. Quantifying impacts of the COVID-19 pandemic through life-expectancy losses: a population-level study of 29 countries. Int J Epidemiol. 2022;51(1):63–74.
9. Cao X, et al. Accelerated biological aging in COVID-19 patients. Nat Commun. 2022;13(1):2135.
10. Brooks-Wilson AR. Genetics of healthy aging and longevity. Hum Genet. 2013;132(12):1323–38.
11. Initiative C-HG. The COVID-19 Host Genetics Initiative, a global initiative to elucidate the role of host genetic factors in susceptibility and severity of the SARS-CoV-2 virus pandemic. Eur J Hum Genet. 2020;28(6):715–8.
12. Initiative C-HG. Mapping the human genetic architecture of COVID-19. Nature. 2021;600(7889):472–7.
13. Deelen J, et al. A meta-analysis of genome-wide association studies identifies multiple longevity genes. Nat Commun. 2019;10(1):3669.
14. Kurki MI, et al. FinnGen provides genetic insights from a well-phenotyped isolated population. Nature. 2023;613(7944):508–18.
15. Bulik-Sullivan B, et al. An atlas of genetic correlations across human diseases and traits. Nat Genet. 2015;47(11):1236–41.
16. Bulik-Sullivan BK, et al. LD score regression distinguishes confounding from polygenicity in genome-wide association studies. Nat Genet. 2015;47(3):291–5.
17. Hemani G et al. The MR-Base platform supports systematic causal inference across the human phenome. Elife, 2018. 7.
18. Qiu S, et al. Genetic correlation, causal relationship, and shared loci between vitamin D and COVID-19: a genome-wide cross-trait analysis. J Med Virol. 2023;95(5):e28780.
19. Qiu S, et al. Exploring the causality between hypothyroidism and non-alcoholic fatty liver: a mendelian randomization study. Front Cell Dev Biol. 2021;9:643582.
20. Qiu S, et al. Mendelian randomization reveals potential causal candidates for COVID-19 in 123 blood metabolites. J Infect. 2022;84(2):248–88.

21. Qiu S, et al. Rheumatoid arthritis and Cardio-Cerebrovascular Disease: a mendelian randomization study. Front Genet. 2021;12:745224.
22. Smith GD, Ebrahim S. Mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? Int J Epidemiol. 2003;32(1):1–22.
23. Hu Y, et al. Mendelian randomization highlights causal association between genetically increased C-reactive protein levels and reduced Alzheimer's disease risk. Alzheimers Dement. 2022;18(10):2003–6.
24. Zhu ZH et al. Causal associations between risk factors and common diseases inferred from GWAS summary data. Nat Commun, 2018. 9.
25. Zhu Z, et al. Integration of summary data from GWAS and eQTL studies predicts complex trait gene targets. Nat Genet. 2016;48(5):481–7.
26. Bowden J, Davey Smith G, Burgess S. Mendelian randomization with invalid instruments: effect estimation and bias detection through Egger regression. Int J Epidemiol. 2015;44(2):512–25.
27. Yoshikawa M, Asaba K, Nakayama T. Estimating causal effects of atherogenic lipid-related traits on COVID-19 susceptibility and severity using a two-sample mendelian randomization approach. BMC Med Genomics. 2021;14(1):269.
28. Bowden J, et al. Improving the visualization, interpretation and analysis of two-sample summary data mendelian randomization via the Radial plot and radial regression. Int J Epidemiol. 2018;47(4):1264–78.
29. Burgess S, et al. Network mendelian randomization: using genetic variants as instrumental variables to investigate mediation in causal pathways. Int J Epidemiol. 2015;44(2):484–95.
30. Yao S, et al. Bidirectional two-sample mendelian randomization analysis identifies causal associations between relative carbohydrate intake and depression. Nat Hum Behav. 2022;6(11):1569–76.
31. Carter AR, et al. Mendelian randomisation for mediation analysis: current methods and challenges for implementation. Eur J Epidemiol. 2021;36(5):465–78.
32. Cheung MW. Comparison of methods for constructing confidence intervals of standardized indirect effects. Behav Res Methods. 2009;41(2):425–38.
33. Andreassen OA, et al. Improved detection of common variants associated with schizophrenia by leveraging pleiotropy with cardiovascular-disease risk factors. Am J Hum Genet. 2013;92(2):197–209.
34. Smeland OB, et al. Discovery of shared genomic loci using the conditional false discovery rate approach. Hum Genet. 2020;139(1):85–94.
35. Schork AJ, et al. New statistical approaches exploit the polygenic architecture of schizophrenia–implications for the underlying neurobiology. Curr Opin Neurobiol. 2016;36:89–98.
36. Torgersen K, et al. Shared genetic loci between depression and cardiometabolic traits. PLoS Genet. 2022;18(5):e1010161.
37. Rodevand L, et al. Extensive bidirectional genetic overlap between bipolar disorder and cardiovascular disease phenotypes. Transl Psychiatry. 2021;11(1):407.
38. Benjamini Y, Hochberg Y. Controlling the false Discovery Rate - a practical and powerful Approach to multiple testing. J Royal Stat Soc Ser B-Statistical Methodol. 1995;57(1):289–300.
39. Smeland OB, et al. Genome-wide analysis reveals extensive genetic overlap between schizophrenia, bipolar disorder, and intelligence. Mol Psychiatry. 2020;25(4):844–53.
40. Karadag N et al. Identification of novel genomic risk loci shared between common epilepsies and psychiatric disorders. Brain, 2023.
41. Watanabe K, et al. Functional mapping and annotation of genetic associations with FUMA. Nat Commun. 2017;8(1):1826.
42. Genomes Project C, et al. A global reference for human genetic variation. Nature. 2015;526(7571):68–74.
43. Consortium GT, et al. Genetic effects on gene expression across human tissues. Nature. 2017;550(7675):204–13.
44. Huang SY, et al. Investigating causal relationships between exposome and human longevity: a mendelian randomization analysis. BMC Med. 2021;19(1):150.
45. Li C, et al. COVID-19 and risk of neurodegenerative disorders: a mendelian randomization study. Transl Psychiatry. 2022;12(1):283.
46. Abdel Moneim A, Radwan MA, Yousef AI. COVID-19 and cardiovascular disease: manifestations, pathophysiology, vaccination, and long-term implication. Curr Med Res Opin. 2022;38(7):1071–9.
47. Fauvel C, et al. Cardiovascular manifestations secondary to COVID-19: a narrative review. Respir Med Res. 2022;81:100904.
48. Hatab T, et al. COVID-19: cardiovascular manifestations-a review of the cardiac effects. J Geriatr Cardiol. 2022;19(3):245–50.
49. Louis DW, et al. The Cardiovascular manifestations of COVID-19. Cardiol Clin. 2022;40(3):277–85.
50. Parvu S, et al. COVID-19 and cardiovascular manifestations. Eur Rev Med Pharmacol Sci. 2022;26(12):4509–19.
51. Yarlagadda LC, et al. Post-COVID-19 Cardiovascular Sequelae and Myocarditis. J Assoc Physicians India. 2023;71(6):11–2.
52. Philip B, et al. COVID-19 and its long-term impact on the cardiovascular system. Expert Rev Cardiovasc Ther. 2023;21(3):211–8.
53. Del Turco S, et al. COVID-19 and cardiovascular consequences: is the endothelial dysfunction the hardest challenge? Thromb Res. 2020;196:143–51.
54. Iqubal A, et al. COVID-19 and cardiovascular complications: an update from the underlying mechanism to consequences and possible clinical intervention. Expert Rev Anti Infect Ther. 2021;19(9):1083–92.
55. South AM, Diz DI, Chappell MC. COVID-19, ACE2, and the cardiovascular consequences. Am J Physiol Heart Circ Physiol. 2020;318(5):pH1084–H1090.
56. Zhu Y, Zhang X, Peng Z. Consequences of COVID-19 on the cardiovascular and renal systems. Sleep Med. 2022;100:31–8.
57. Huang B, et al. Interleukin-33-induced expression of PIBF1 by decidual B cells protects against preterm labor. Nat Med. 2017;23(1):128–35.
58. Druckmann R, Druckmann MA. Progesterone and the immunology of pregnancy. J Steroid Biochem Mol Biol. 2005;97(5):389–96.
59. Orru B, et al. Inhibitory effects of vitamin D on inflammation and IL-6 release. A further support for COVID-19 management? Eur Rev Med Pharmacol Sci. 2020;24(15):8187–93.
60. Check JH, Check D. Therapy aimed to suppress the production of the immunosuppressive protein Progesterone Induced blocking factor (PIBF) may provide palliation and/or increased longevity for patients with a Variety of different Advanced cancers - A Review. Anticancer Res. 2019;39(7):3365–72.
61. Szekeres-Bartho J, Polgar B. PIBF: the double edged sword. Pregnancy and tumor. Am J Reprod Immunol. 2010;64(2):77–86.
62. Cooper YA, et al. Functional regulatory variants implicate distinct transcriptional networks in dementia. Science. 2022;377(6608):eabi8654.
63. Gouveia C, et al. Genome-wide association of polygenic risk extremes for Alzheimer's disease in the UK Biobank. Sci Rep. 2022;12(1):8404.
64. Witoelar A, et al. Genome-wide Pleiotropy between Parkinson Disease and Autoimmune diseases. JAMA Neurol. 2017;74(7):780–92.
65. Crivelli L, et al. Changes in cognitive functioning after COVID-19: a systematic review and meta-analysis. Alzheimers Dement. 2022;18(5):1047–66.
66. Welberg L. COVID-19 and brain aging. Nat Neurosci. 2023;26(1):1.
67. Gagnon E, et al. Impact of the gut microbiota and associated metabolites on cardiometabolic traits, chronic diseases and human longevity: a mendelian randomization study. J Transl Med. 2023;21(1):60.
68. He D, et al. Association between gut microbiota and longevity: a genetic correlation and mendelian randomization study. BMC Microbiol. 2022;22(1):302.
69. Ye CJ, et al. Mendelian randomization evidence for the causal effects of socio-economic inequality on human longevity among europeans. Nat Hum Behav. 2023;7(8):1357–70.
70. Tian Q, et al. The relationship of parental longevity with the aging brain-results from UK Biobank. Geroscience. 2020;42(5):1377–85.

Publisher's note