nature medicine

Article

https://doi.org/10.1038/s41591-025-03999-8

MRI-based multi-organ clocks for healthy aging and disease assessment

Received: 10 April 2025

Accepted: 4 September 2025

Published online: 16 October 2025

Check for updates

The MULTI Consortium*, Huizi Cao[1], Zhiyuan Song[1], Michael R. Duggan[2], Guray Erus[3], Dhivya Srinivasan[3], Ye Ella Tian[4], Wenjia Bai[5], Michael S. Rafii[6], Paul Aisen[6], Daniel W. Belsky[7], Keenan A. Walker[2], Andrew Zalesky[4], Luigi Ferrucci[8], Christos Davatzikos[3] & Junhao Wen[1,9,10,11,12,13,14]

Biological aging clocks across organ systems and tissues have advanced understanding of human aging and disease. In this study, we expand this framework to develop seven magnetic resonance imaging-based multi-organ biological age gaps (MRIBAGs), including the brain, heart, liver, adipose tissue, spleen, kidney and pancreas. Using data from 313,645 individuals curated by the MULTI Consortium, we link the seven MRIBAGs to 2,923 plasma proteins, 327 metabolites and 6,477,810 common genetic variants. Genome-wide associations identify 53 MRIBAG–locus pairs ($P < 5 \times 10^{-8}$). Genetic correlation and Mendelian randomization analyses support organ-specific and cross-organ interconnection, including 24 non-MRI biological aging clocks and 525 disease endpoints. Through functional gene mapping and Bayesian co-localization multi-omics evidence, we prioritize nine druggable genes as targets for future anti-aging treatments. Furthermore, the seven MRIBAGs are linked to future risk of systemic disease endpoints (for example, diabetes mellitus) and all-cause mortality. Finally, participants with more youthful versus more aged brain profiles exhibited distinct cognitive decline trajectories over 240 weeks of treatment with the Alzheimer's disease drug solanezumab, although this heterogeneity cannot be fully attributed to the drug. In summary, we developed seven MRIBAGs that enhance the existing multi-organ biological aging framework, and we demonstrate their clinical potential to advance aging research.

Magnetic resonance imaging (MRI[1]) provides a non-invasive window into structural and functional changes occurring with age, enabling the development of personalized disease biomarkers. Brain MRI-based aging clocks, often referred to as 'brain age' and derived from artificial intelligence and machine learning (AI/ML) applied to MRI data, have been widely used as clinical biomarkers of neurological aging, cognitive decline and neurodegenerative disease risk[2]. However, although brain imaging has been extensively used for aging research, no studies have systematically extended this concept to other organ systems. With the advent of large-scale, multi-organ MRI datasets, such as those from the UK Biobank (UKBB), we now have an unprecedented opportunity to conduct a more holistic, system-wide investigation of biological aging across multiple organs.

Beyond imaging-based aging clocks, researchers have also developed various omics-based aging clocks to capture different biological dimensions of aging. For example, plasma proteomics-based aging clocks[3,4] have been introduced to enhance the granularity and coverage of multi-organ aging clocks. Plasma proteomics-based aging

A full list of affiliations appears at the end of the paper. *A list of authors and their affiliations appears at the end of the paper.
e-mail: junhao.wen89@gmail.com

clocks capture the biological aging process by analyzing the abundance of proteins in circulation. Because proteins are the functional units of the cell, these clocks provide insights into the downstream effects of genetic regulation and environmental influences on aging. Metabolomics-based aging clocks[5,6] capture biological aging by profiling small-molecule metabolites that represent the biochemical activity of cellular and systemic metabolism. Metabolomics examines the end products of metabolic pathways, making it highly dynamic and sensitive to external factors such as diet, microbiome composition, lifestyle and environmental exposures. Finally, epigenetics-based aging clocks[7] focus on the analysis of DNA methylation patterns, particularly at specific CpG sites across the genome, to estimate biological age. Integrating these aging clocks into the multi-organ and multi-omics concept[8–10] is essential for gaining a comprehensive understanding of aging biology, age-related diseases and longevity.

In our previous efforts, we developed nine phenome-based biological age gaps (BAGs)[11,12] (PhenoBAGs), 11 proteome-based BAGs (ProtBAGs[13]) and five metabolome-based BAGs (MetBAGs[5]). In the present study, we derived six additional MRIBAGs, on top of the brain MRIBAG[14], for the heart, liver, spleen, adipose, kidney and pancreas using data consolidated from the MULTI Consortium (Method 1). We first benchmarked the age prediction performance of the six additional MRIBAGs (Method 2). Next, we associated the seven MRIBAGs with 2,923 plasma proteins for proteome-wide associations (ProWASs) and 327 metabolites for metabolome-wide associations (MetWASs) (Method 3). Through genome-wide association studies (GWASs) and post-GWAS analyses, we depicted their genetic architecture and identified potential druggable genes (Method 4). Finally, we assessed the clinical applicability and predictability of the MRIBAGs and their polygenic risk scores (PRSs) (Method 5). All results and pre-trained AI/ML models are publicly available at the MEDICINE portal: https://labs-laboratory.com/medicine/.

Results

Age prediction performance of the seven MRIBAGs

The brain MRIBAG was initially developed in our previous study[14] using 119 gray matter volumes, and we re-trained the brain MRIBAG here using a consistent nested cross-validation procedure along with a within-distribution, holdout test dataset (Extended Data Fig. 1). Specifically, to rigorously evaluate the performance (that is, overfitting and generalizability) of biological age prediction models, we partitioned the healthy control (CN, without any pathologies) participants into the CN training/validation/test ($3{,}573 < N < 6{,}327$ as the sample sizes vary across organs) and the within-distribution, holdout test ($N = 500$) datasets (Method 2 and Supplementary Table 1).

When fitting the organ-specific MRI features (Fig. 1a, Method 2 and Supplementary File 1), the two AI/ML models (least absolute shrinkage and selection operator (LASSO) regression and support vector regressor) exhibited slight variability in performance, with no single model consistently outperforming the others (Fig. 1b). The selected optimal models for the seven MRIBAGs demonstrated moderate Pearson's r coefficients (ranging from 0.23 to 0.77) and the mean absolute errors (MAEs) of approximately 5 years in the within-distribution, holdout test dataset (Fig. 1c) before applying the age bias correction[15,16]. Supplementary Table 2 presents detailed statistics for the age prediction tasks before the age bias correction (Supplementary Fig. 1). Supplementary Fig. 2 presents the incremental statistical power, beyond age and sex, of the seven MRIBAGs in predicting mortality and disease categories. Model improvements may be possible for abdominal features by employing more advanced modeling approaches, such as deep neural networks that directly process voxel-wise MRI data[17]. Additionally, we found that high collinearity among imaging features from abdominal MRI led to poor generalizability in within-distribution, holdout test datasets (Supplementary Fig. 3). Supplementary Fig. 4 shows potential domain shift while applying the pre-trained model of the brain MRIBAGs to external datasets, exemplified by the A4 study[18,19]. Supplementary Fig. 5 presents a sensitivity analysis for the Alzheimer's disease drug differential responses. Supplementary Fig. 6 discusses sex differences in the seven MRIBAGs. Supplementary Figs. 7 and 8 discuss the feature importance of deriving these MRIBAGs and their biological interpretation. For all subsequent analyses, we used the age bias-corrected[16] MRIBAGs.

The plasma proteomics and metabolomics landscape of the seven MRIBAGs

In our ProWAS analyses (Method 3), we identified 603 protein–MRIBAG significant associations ($P < 0.05/2{,}923/7$). Among these, the kidney MRIBAG exhibited the highest number of significant protein associations ($N = 301$; for example, NPDC1, IGFBP6 and TAFA5), followed by the spleen MRIBAG for 136 associations (for example, VCAM1, PTPRH and C1QA), the liver MRIBAG for 62 associations (for example, NCAN, SEZ6L and LEP), the adipose MRIBAG for 57 signals (for example, GDF15, CHI3L1 and CA14), the pancreas MRIBAG for 21 associations (for example, PLA2G1B, CTRC and CELA2A), the brain MRIBAG for 16 associations (for example, BCAN, NCAN and GDF15) and the heart MRIBAG for only REN (Fig. 2a).

VCAM1 was associated with the spleen MRIBAG ($\beta = -0.89 \pm 0.04$; $P = 7.05 \times 10^{-81}$; $r = -0.38$); PLA2G1B was associated with the pancreas MRIBAG ($\beta = -1.14 \pm 0.07$; $P = 8.26 \times 10^{-56}$; $r = -0.30$) and the adipose and spleen MRIBAGs. The two proteins are organ enriched in respective organ and tissue, as identified by the Human Protein Atlas (HPA; https://www.proteinatlas.org/), emphasizing their relevance to organ-specific aging processes, which has been widely used to define organ-enriched proteins in previous proteome-based aging clock papers[4,13]. Supplementary File 2a presents detailed statistics. Supplementary Fig. 9 and Supplementary File 2b explore the potential replication of UKBB Olink ProWAS signals in Baltimore Longitudinal Study of Aging (BLSA) SomaScan data. Extended Data Figs. 2 and 3 present individual volcano plots and protein set enrichment results via the STRING platform (version 12.0 (ref. 20)) for each MRIBAG (Supplementary Figs. 10–14). Supplementary Fig. 15 presents the sex-specific ProWAS results.

In our MetWAS (Method 3), we identified 758 metabolite–MRIBAG significant associations ($P < 0.05/327/7$) between the seven MRIBAGs and metabolites, showing relatively weak associations (Pearson's $r < 0.3$). The spleen MRIBAG showed the highest number of significant metabolite associations ($N = 199$; for example, acetate, acetoacetate and albumin), followed by the adipose MRIBAG ($N = 196$; for example, L_HDL_PL_pct, L_HDL_C_pct and PUFA_pct), the heart MRIBAG ($N = 139$; for example, PUFA_pct, M_HDL_TG and HDL_TG), the brain MRIBAG ($N = 97$; for example, LA_pct, Omega_6_pct and ApoA1), the pancreas MRIBAG ($N = 57$; for example, GlycA, Glucose_Lactate and HDL_PL_pct), the liver MRIBAG ($N = 51$; for example, Gly, Tyr and L_LDL_PL_pct) and the kidney MRIBAG ($N = 19$; for example, creatinine, citrate and DHA) (Fig. 2b). Descriptions of these metabolites are provided in Supplementary Table 3.

Creatinine was linked to the kidney ($\beta = 21.77 \pm 0.65$; $P = 1.58 \times 10^{-258}$; $r = 0.28$) and adipose ($\beta = -4.37 \pm 0.83$; $P = 1.48 \times 10^{-7}$; $r = -0.18$) MRIBAGs, which may suggest distinct roles of the kidney and adipose tissue in creatinine regulation. In addition to these small-molecule metabolites, other lipid complexes and subclasses also showed significant associations. Total concentration of phospholipids (Total_P) was negatively associated with liver ($\beta = 51.57 \pm 3.59$; $P = 2.52 \times 10^{-46}$; $r = 0.18$). Phospholipids in small HDL (S_HDL_PL) were linked to the brain, heart, spleen and adipose MRIBAGs. Similarly, cholesteryl esters in very large HDL (XL_HDL_CE) associated with the adipose and heart MRIBAGs. Supplementary File 3 presents detailed statistics. Extended Data Fig. 4 presents individual volcano plots for each MRIBAG. Metabolite set enrichment analysis is presented in Extended Data Fig. 5. Supplementary Fig. 16 shows the genetic correlations among the seven MRIBAGs, 2,923 proteins and 327 metabolites. Extended Data Fig. 6 presents

https://doi.org/10.1038/s41591-025-03999-8

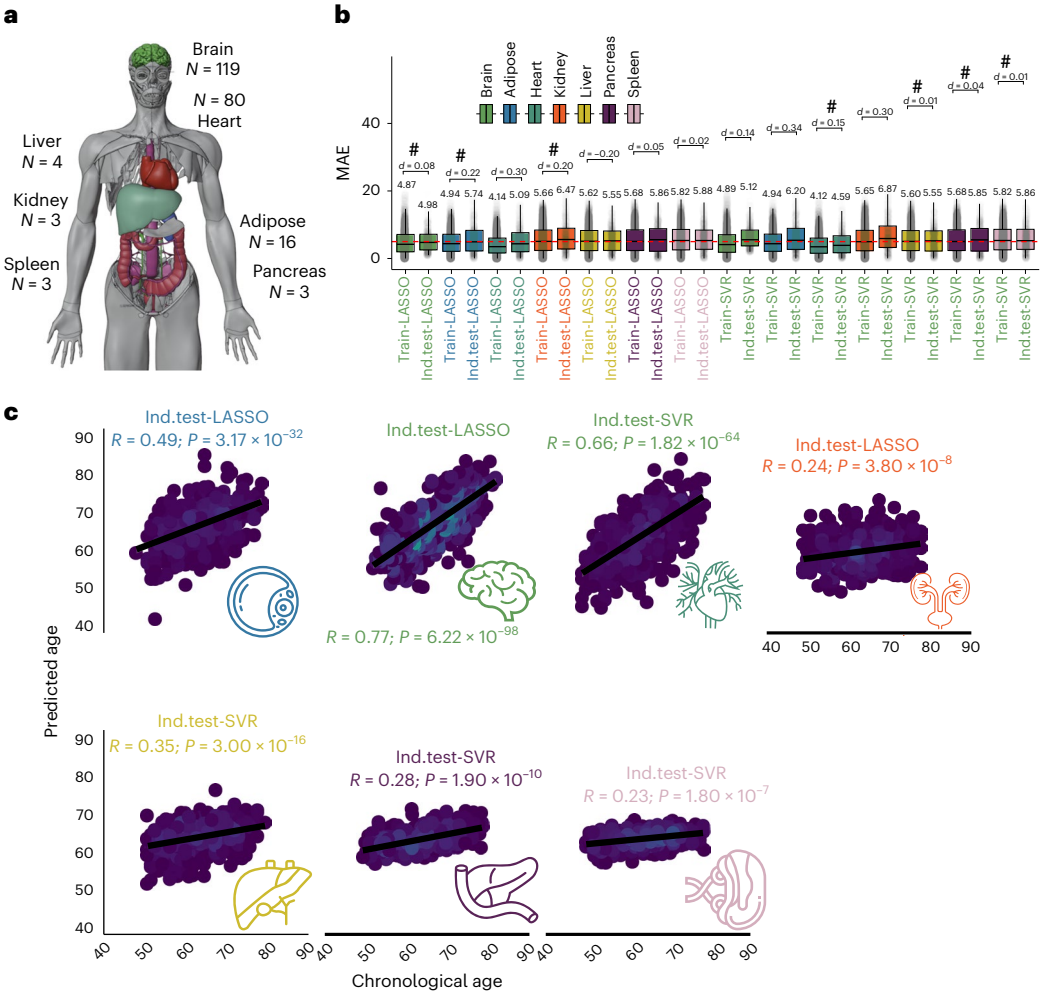


Fig. 1 | Model performance benchmarking for seven MRIBAGs. a, The seven MRIBAGs are derived from various imaging modalities available in the UKBB. The number (*N*) of IDPs used to derive the seven MRIBAGs is displayed. **b**, For the seven MRIBAGs, the MAE of age prediction models using LASSO regression and linear support vector regressor (SVR) is shown for both the training dataset (cross-validated training/validation/test; 3,582 < *N* < 6,325) and the within-distribution, holdout test dataset (*N* = 500). Cohen's *d* quantifies the effect size of the difference between these datasets, reflecting potential model generalizability, assuming similar age and sex distributions. The optimal model (#) for each organ and tissue was selected based on the lower Cohen's *d* value, and these models were used for all subsequent analyses. All MAEs are reported without age bias correction. **c**, Scatter plots display the optimal model for each organ/tissue in the within-distribution, holdout test dataset, with Pearson's *r* and two-sided *P* values indicating the association between chronological age and predicted age. All results are shown before the correction for age bias was applied[16]. To evaluate the generalizability of the brain MRIBAG model, we validated it using two independent studies. In **b**, individual data points are overlaid on box plots summarizing the distribution. The center line indicates the median; the box bounds represent the first and third quartiles; and the whiskers extend to the most extreme values within 1.5 times the interquartile range. Ind., independent.

example associations for the ProWAS and MetWAS. Supplementary Fig. 17 presents the sex-specific MetWAS results.

The genetic architecture of the seven MRIBAGs

We conducted GWAS (Method 4) for the seven MRIBAGs (19,686 < *N* < 31,557 participants with European ancestries) and identified 53 ($P < 5 \times 10^{-8}$) genomic locus–BAG pairs. We denoted the genomic loci using their top lead single-nucleotide polymorphisms (SNPs) defined by FUMA, considering linkage disequilibrium; the genomic loci are presented in Supplementary Table 4. We visually present the shared genomic loci annotated by cytogenetic regions based on the GRCh37 cytoband (Fig. 3a). Supplementary Fig. 18 and Extended Data Fig. 7 detail the robustness of our GWASs and sex-specific GWAS analyses.

We estimated the SNP-based heritability (h^2_{SNP}) for the seven MRIBAGs using GCTA[21] software (Method 4), with values ranging from 0.29 to 0.47 (Fig. 3b). Supplementary Table 5 presents detailed statistics. We validated the GWAS signals through partitioned heritability analyses (Method 4) using LDSC software, revealing strong organ-specific enrichment. Notably, the heart MRIBAG exhibited chromatin state-specific enrichment across four distinct heart tissues (Fig. 3c). For example, significant heritability enrichment was shown in the right atrium in the H3K4me3 region ($P = 1.00 \times 10^{-6}$) and in the left ventricle in the H3K27ac region ($P = 2.00 \times 10^{-5}$) for the heart MRIBAG (Supplementary File 4).

We then assessed the genetic correlations (Method 4) between the seven MRIBAGs and 24 previously developed multi-organ, multi-omics aging clocks, including eight PhenoBAGs[11] (excluding the brain PhenoBAG based on multimodal imaging), 11 ProtBAGs[13] and five MetBAGs[5]. After Bonferroni correction, we found seven within-organ and inter-organ significant genetic correlations ($P < 0.05/11$), exemplified between the kidney MRIBAG and the renal PhenoBAG ($g_c = 0.23 \pm 0.05$; $P < 2.00 \times 10^{-5}$) and between the spleen MRIBAG and the hepatic MetBAG ($g_c = 0.23 \pm 0.08$; $P < 7.00 \times 10^{-4}$). We also performed genetic correlation between the seven MRIBAGs and 525 disease endpoints and found six significant signals ($P < 0.05/525$). For example, within-organ connections were demonstrated between

https://doi.org/10.1038/s41591-025-03999-8

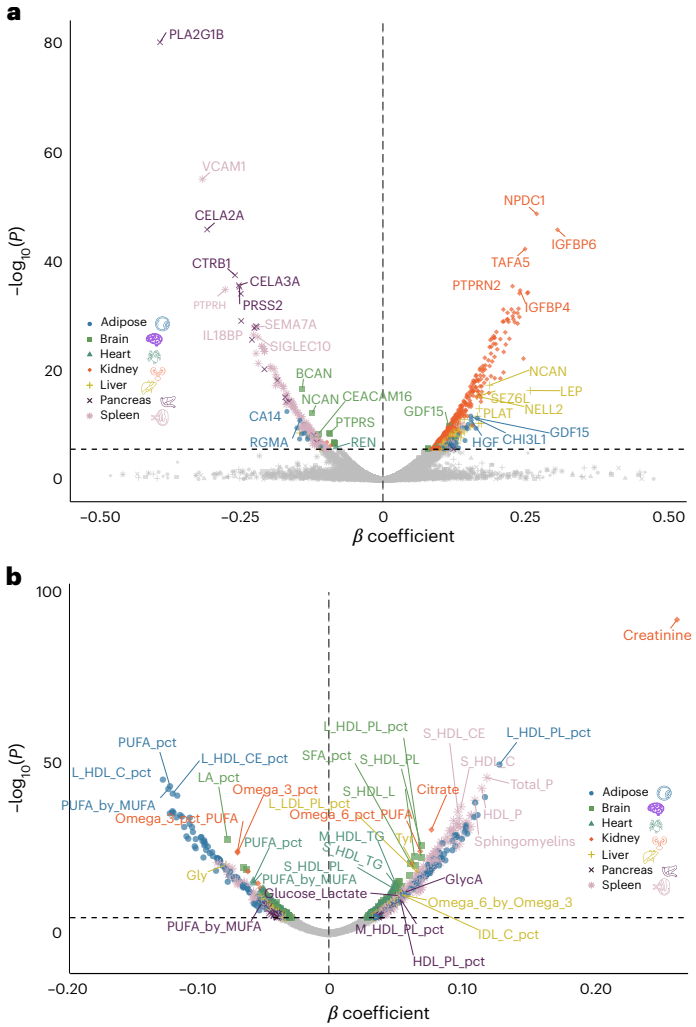


Fig. 2 | ProWAS and MetWAS with the seven MRIBAGs. a, ProWAS between the seven MRIBAGs and 2,923 plasma proteins via a linear regression model, accounting for a full set of covariates ($P < 0.05/2{,}923/7$). An online interactive webpage is available at https://labs-laboratory.com/medicine/mribag_prowas.html to ease visualization. Plots of individual organs are shown in Extended Data Fig. 2. **b**, MetWAS between the seven MRIBAGs and 327 plasma metabolites using a linear regression model while adjusting for a comprehensive set of covariates ($P < 0.05/327/7$). Plots of individual organs are shown in Extended Data Fig. 4. An online interactive webpage is available at https://labs-laboratory.com/medicine/mribag_metwas.html to ease visualization. The colors and shapes of the icons indicate the organ systems involved in the associations. We used the standardized β value to represent the effect size.

the heart MRIBAG and two forms of hypertension, which is a major risk factor for cardiovascular diseases (for example, FinnGen code: I9_HYPTENS: $g_c = 0.28 \pm 0.06$; $P < 1.00 \times 10^{-5}$); cross-organ interactions were also identified between the adipose MRIBAG and substance abuse ($g_c = 0.27 \pm 0.06$; $P < 5.00 \times 10^{-5}$) (Fig. 3d, Supplementary Fig. 19 and Supplementary File 5).

We conducted bidirectional Mendelian randomization analyses (Method 4) between the seven MRIBAGs and 525 disease endpoints, ensuring at least seven valid linkage disequilibrium-considered SNPs as instrumental variables after quality control. Although the MRIBAGs were generally underpowered to meet this threshold, we identified six potential causal relationships. The within-organ interactions between the heart MRIBAG and hypertension were reaffirmed by a potential causal relationship from various forms of hypertension to the heart MRIBAG (for example, I9_HYPTENS: $P = 2.12 \times 10^{-7}$; odds ratio (95% confidence interval) = 1.13 (1.08–1.18); number of instrumental variables = 110). Additionally, we identified a potential causal relationship from Alzheimer's disease to the liver MRIBAG (G6_AD_WIDE: $P = 3.02 \times 10^{-5}$; odds ratio (95% confidence interval) = 0.91 (0.88–0.96); number of instrumental variables = 8), consistent with our previous findings linking Alzheimer's disease to the hepatic PhenoBAG[11]. Finally, we found that the kidney MRIBAG was negatively linked to two forms of type 2 diabetes mellitus (T2D) (for example, T2D: $P = 8.40 \times 10^{-5}$; odds ratio (95% confidence interval) = 0.78 (0.70–0.88); number of instrumental variables = 9) (Fig. 3e). Supplementary File 6 and Supplementary Figs. 20–25 present detailed statistics for our causal analyses and sensitivity check analyses.

Using split-sample GWAS analyses, we also developed PRSs for the seven MRIBAGs, with the first split serving as the training/source dataset and the second split serving as the test/target dataset (Method 4). Overall, these PRSs exhibit limited predictive power. In the second split population, the brain PRS exhibited the highest incremental R^2 (2.18%; $P < 1.00 \times 10^{-10}$), whereas the spleen PRS accounted for an additional 0.70% ($P < 1.00 \times 10^{-10}$) of the variance in the spleen MRIBAG (Fig. 3f and Supplementary Table 6).

Genetic evidence to prioritize nine druggable genes

We conducted multiple post-GWAS analyses (Method 4) to identify potential druggable genes associated with the seven MRIBAGs. Our overarching hypothesis posits that these imaging-derived aging clocks, as AI-driven endophenotypes[22] (see Supplementary Fig. 26 for a conceptual visualization), reside along the causal pathway of human aging and disease, linking genetic predisposition to multi-omics data, such as gene expression, proteomics and metabolomics, leading to disease outcomes and cognitive decline.

Initially, we mapped independent GWAS signals to genes using three approaches—(1) positional mapping, (2) expression quantitative trait locus (eQTL) mapping and (3) chromatin interaction mapping—identifying a total of 1,164 unique genes. To further refine these genes, we prioritized genes that demonstrated genetic evidence from both eQTL mapping and chromatin interactions, resulting in a final set of 246 unique genes linked to the seven MRIBAGs. For instance, *MAPT* was prioritized in the brain MRIBAG through both eQTL and chromatin interaction mapping and was previously identified as a genetically supported druggable gene[23]. Subsequently, to further extend genetic evidence to other omics data types, we conducted Bayesian co-localization analysis on genomic loci associated with the seven MRIBAGs and aging clocks derived from plasma proteomics[13] (ProtBAGs) and metabolomics (MetBAGs[5]). The key hypothesis is that a causal variant positionally or functionally linked to a gene influencing both MRIBAGs and other omics-based aging clocks is more likely to play a causal role in human aging and disease. Our co-localization analysis identified a total of 40 MRIBAG–MetBAG or MRIBAG–ProtBAG co-localization signals (PP.H4.ABF > 0.8) (Fig. 4a). By integrating these findings with the 246 prioritized genes, we identified 62 MRIBAG–gene pairs (comprising 62 unique genes) that share at least one causal SNP with either MetBAGs or ProtBAGs (Fig. 4b, Supplementary Table 7 and Supplementary File 7).

Among the 62 genes with multi-level genetic evidence supporting their potential causal role in human aging and disease, we further explored existing drug–gene interaction data to identify potential drug repurposing opportunities and evidence. To achieve this, we queried the Drug Gene Interaction database (DGIdb) (https://dgidb.org/; query date: 21 March 2025) and identified nine genes (linked to the liver, spleen and pancreas MRIBAGs) with known interactions involving 122 unique drugs (Fig. 5a and Supplementary File 8). Among the 122 unique drugs, many indications are manifested, including immunosuppressant agents (for example, everolimus; DrugBank ID: DB01590), antipsychotic agents (for example, loxapine for the management of the manifestations of psychotic disorders such as schizophrenia; DrugBank ID: DB00408), antidiabetic agents (for example, bromocriptine

https://doi.org/10.1038/s41591-025-03999-8

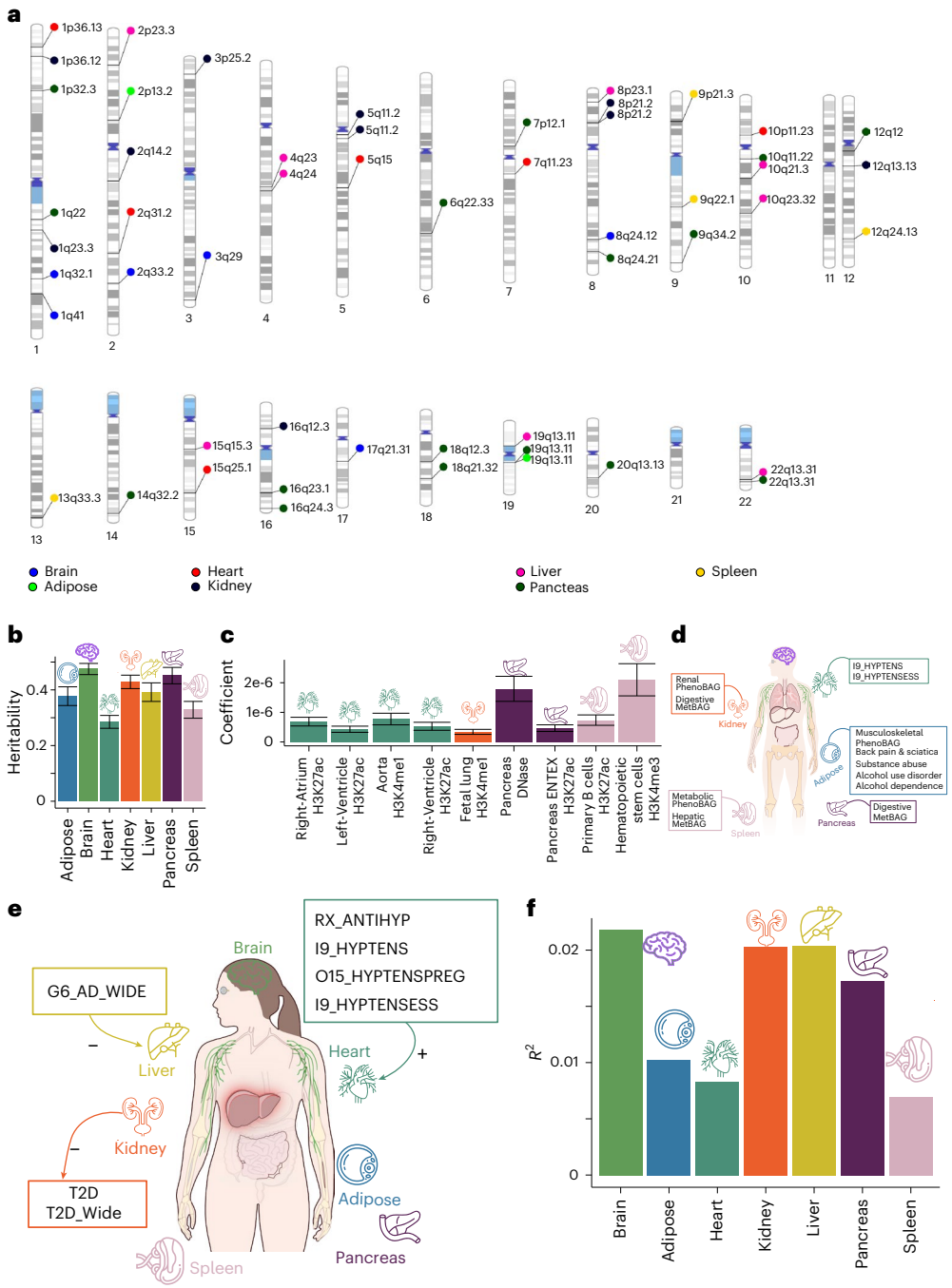


Fig. 3 | The genetics of the seven MRIBAGs. a, Cytogenetic regions associated with the seven MRIBAGs, with significant genomic loci identified using the genome-wide significance threshold (two-sided $P < 5 \times 10^{-8}$). **b**, SNP-based heritability estimates of the seven MRIBAGs (sample size: 18,778 < N < 31,557). **c**, Partitioned heritability enrichment analysis using chromatin-specific multi-tissue chromatin data, multi-tissue gene expression profiles and cell-type-specific datasets. Only significant results (two-sided $P < 0.05/697$) are displayed (sample size: 18,778 < N < 31,557). **d**, Genetic correlation estimates via LDSC between the seven MRIBAGs and eight PhenoBAGs (excluding the multimodal brain PhenoBAG[11]), 11 ProtBAGs[13] and five MetBAGs[5] (two-sided $P < 0.05/11$) as well as 525 disease endpoints from FinnGen and PGC ($P < 0.05/525$). **e**, Potential causal relationships between the seven MRIBAGs (that is, number of instrumental variables > 7) and 525 disease endpoints were examined through two networks: *BAG2DE*, where the seven MRIBAGs serve as exposure variables and the 525 DEs as outcome variables, with two-sided $P < 0.05/525$, and *DE2BAG*, where the 214 disease endpoints serve as effective exposure variables (that is, number of instrumental variables > 7), with two-sided $P < 0.05/214$. Although multiple sensitivity checks were performed to evaluate potential violations of underlying assumptions, these findings should be interpreted with caution. The direction array represents the causal relationship from the exposure to the outcome variables, where '+' denotes an odds ratio > 1 and '−' indicates an odds ratio < 1. **f**, The bar plot shows the incremental R^2 (that is, the R^2 of the alternative model minus that of the null model) for the PRS of each MRIBAG. The PRS was calculated using the split2 target GWAS data, with split1 GWAS data serving as the training set for the PRScs model. Bar plots represent the mean (center of bar), with error bars indicating ±1 s.e.

https://doi.org/10.1038/s41591-025-03999-8

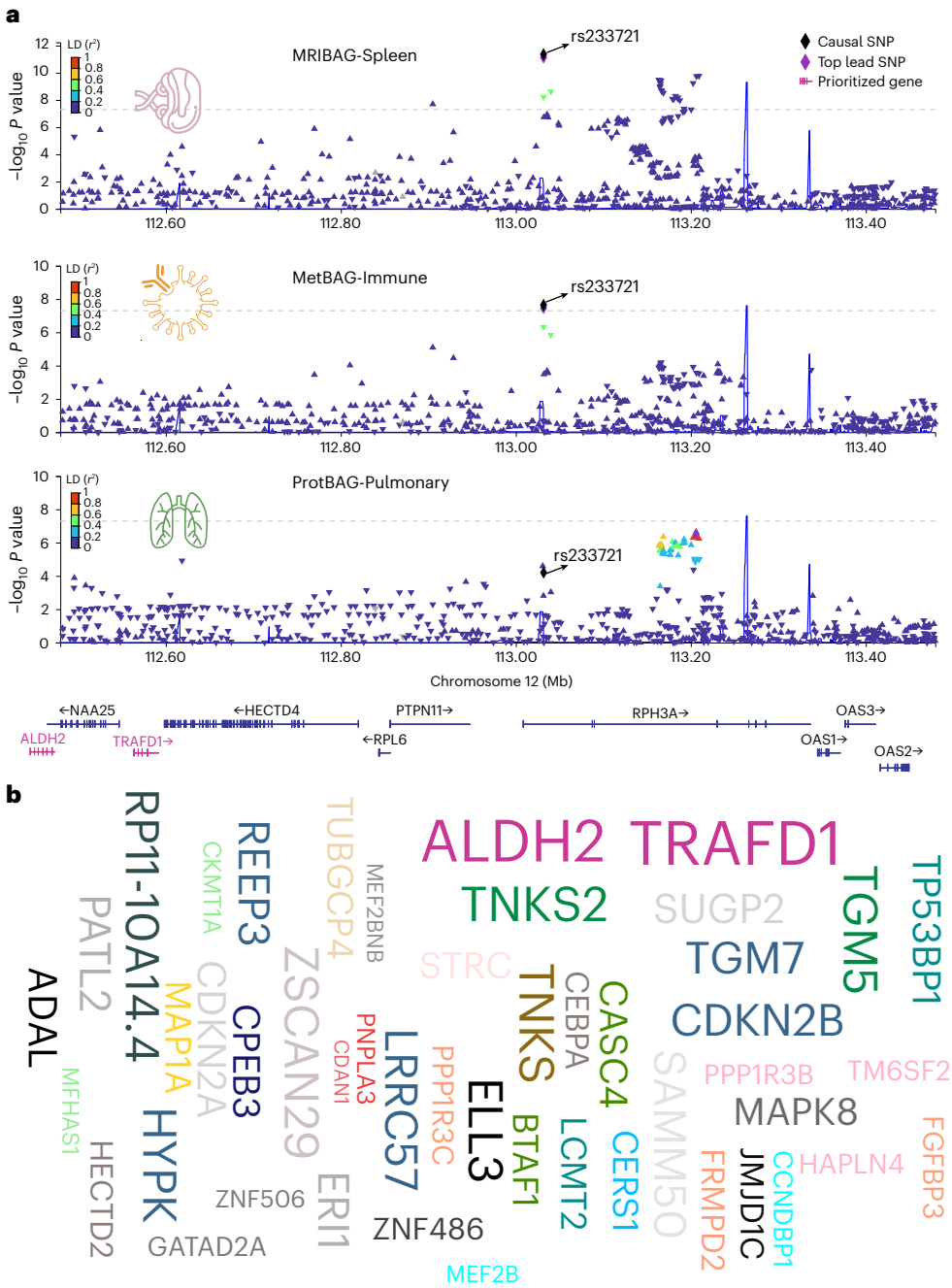


Fig. 4 | Multi-level genetic evidence prioritizes 62 genes as potential drug repurposing genes. a, To prioritize potential druggable genes, leveraging aging clocks derived from plasma proteomics (ProtBAGs) and metabolomics (MetBAGs), we filtered and mapped genes to retain only those showing significant co-localization with any ProtBAG or MetBAG. As an example, we highlighted a specific genomic locus demonstrating co-localization among the spleen MRIBAG, the immune MetBAG and the pulmonary ProtBAG. The shared causal variant (rs233721 on chromosome 12) exhibited a strong co-localization signal, with a posterior probability (PP.H4.ABF) of 0.99 between the spleen MRIBAG and the immune MetBAG, implying a single shared causal variant influencing both traits within this locus. To determine the significance of the H4 hypothesis, we set a threshold of PP.H4.ABF > 0.8. This variant was mapped to the *TRAFD1* and *ALDH2* genes through both eQTL mapping (for example, *ALDH2* in the aorta from GTEx) and chromatin interaction mapping (for example, *TRAFD1* in the left ventricle and liver). **b**, By integrating this multi-level genetic evidence, we identified 62 unique genes associated with the liver, spleen and pancreas MRIBAGs. LD, linkage disequilibrium.

for treating T2D; DrugBank ID: DB01200) and antiparkinson agents (lisuride for the management of Parkinson's disease; DrugBank ID: DB00589), as well as many antineoplastic agents.

We highlighted the spleen MRIBAG-linked gene *ALDH2* as exemplified in our co-localization analyses (Fig. 4a). Multi-level genetic evidence links the spleen BAG, *ALDH2* and 15 drugs (loxapine, clozapine, rotigotine, sulpiride, prochlorperazine, terguride, sertindole, fenoldopam, pergolide, trifluoperazine, naltrexone, cabergoline, apomorphine, lisuride and bromocriptine) with multiple clinical indications. These indications include antipsychotic (for example, clozapine for treating severe forms of psychiatric conditions such as schizophrenia and for preventing suicide in patients who have this condition); antiparkinson (for example, rotigotine); antidepressive (for example, sulpiride, used to treat acute and chronic schizophrenia

https://doi.org/10.1038/s41591-025-03999-8

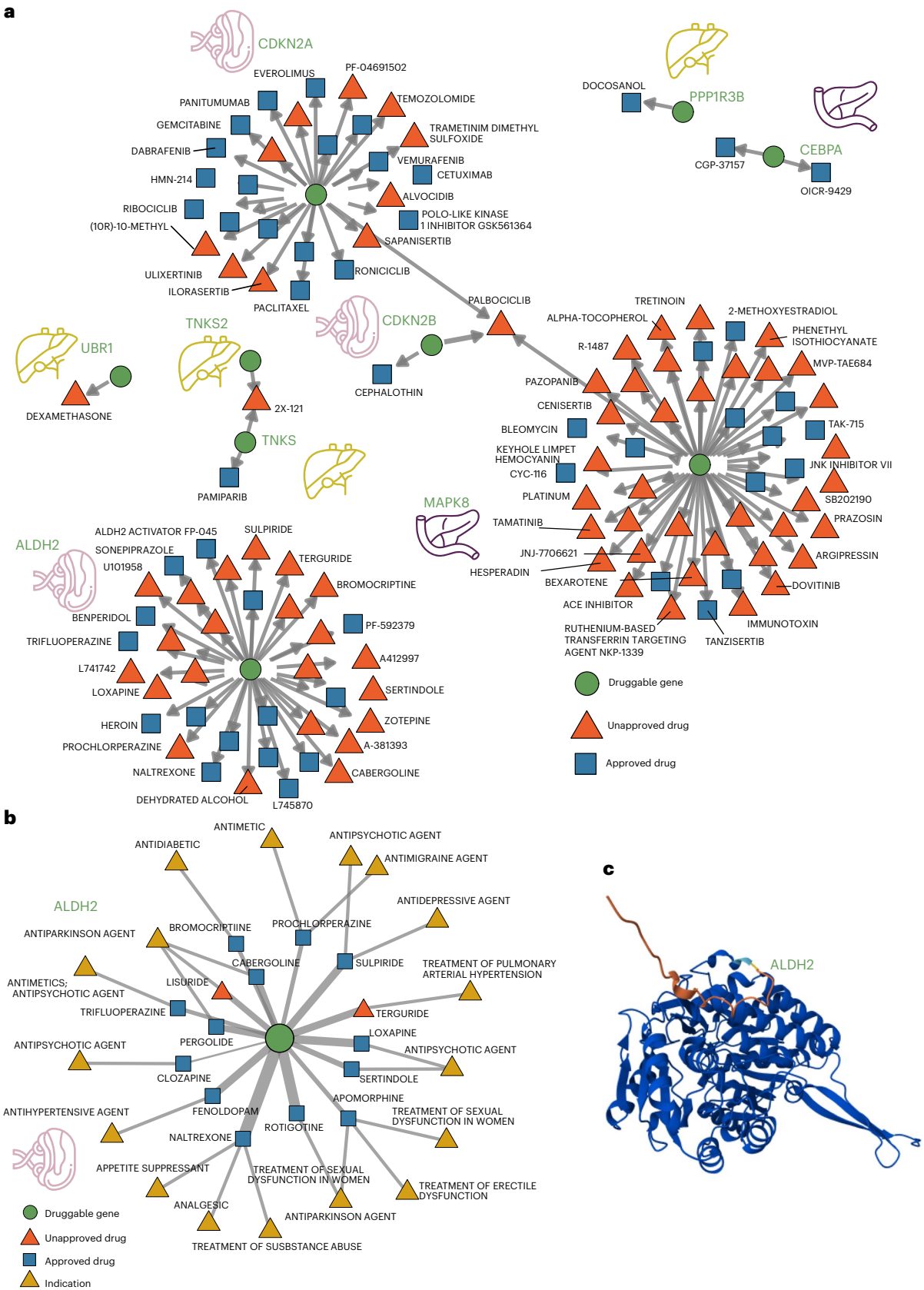


Fig. 5 | Drug–gene interactions for nine potential drug repurposing genes. **a**, We queried the 62 prioritized MRIBAG genes in the DGIdb platform (https://dgidb.org/) to identify drug–gene interactions and potential drug repurposing opportunities. We show only the nine MRIBAG genes linked to existing drugs, whether approved or unapproved, using curated data via the DGIdb. For visualization, we annotated representative drug names; the complete results can be found at an online interactive webpage at https://labs-laboratory.com/medicine/mribag_gdi.html. **b**, Drug indications for existing drug–gene interactions related to the spleen MRIBAG-linked *ALDH2* gene. **c**, The three-dimensional structure of the ALDH2 protein predicted by the AlphaFold model.

https://doi.org/10.1038/s41591-025-03999-8

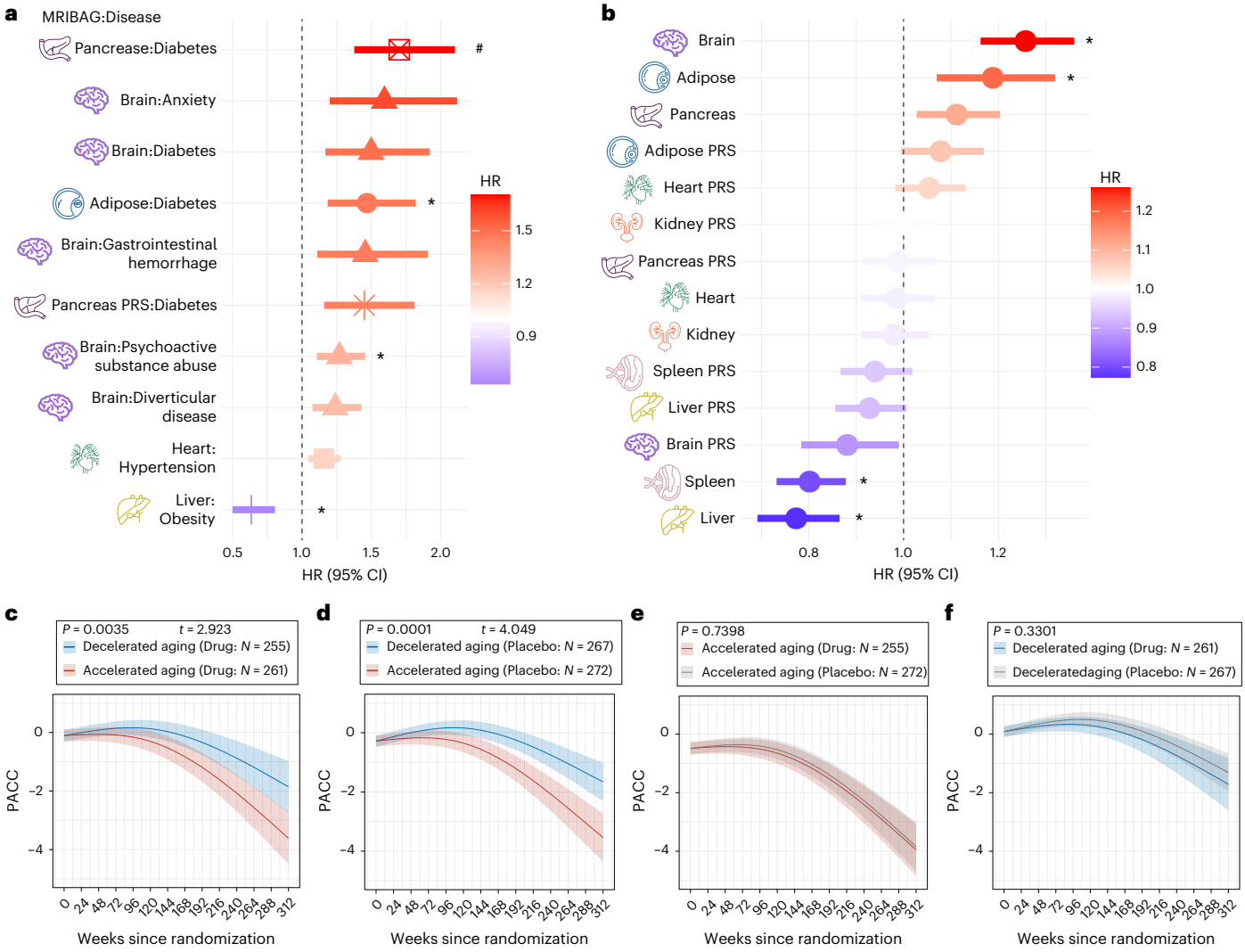


Fig. 6 | Clinical utility of the seven MRIBAGs and their PRSs. a, The seven MRIBAGs and their PRSs show significant associations with the incidence of single disease entities (for example, I10 for hypertension). We only included disease entities with more than 50 cases (two-sided $P < 0.05/7 = 0.0071$ corrected for the number of organ systems), and non-cases are participants who do not show any disease diagnosis (disease free) after enrolling in the UKBB, so that all non-case populations remained the same reference population across all tasks. Age and sex were included as covariates in the Cox proportional hazard model. An asterisk (*) denotes results that remain significant after applying the Bonferroni correction for multiple comparisons (two-sided $P < 0.05/53 = 0.00094$, accounting for the number of disease endpoints), whereas a hash symbol (#) indicates significance under a more stringent threshold (two-sided $P < 0.05/53/7 = 0.00013$). **b**, The brain, adipose, spleen and liver MRIBAGs showed significant associations with all-cause mortality risk. To further validate the negative associations of the spleen and liver MRIBAGs with mortality, we conducted a disease-free analysis, including only participants without any prior disease diagnoses, to rule out potential confounders. **c**, At week 240, PACC scores differed between the accelerated and decelerated aging groups—stratified by brain MRIBAG expression—among participants originally assigned to the drug group in the solanezumab trial. **d**, At week 240, PACC scores differed between the accelerated and decelerated aging groups among participants originally assigned to the placebo group in the solanezumab trial. **e**, At week 240 of the solanezumab trial, PACC scores in the accelerated aging group showed no difference between participants receiving the drug and those receiving a placebo. **f**, At week 240 of the solanezumab trial, PACC scores in the decelerated aging group showed no difference between participants receiving the drug and those receiving a placebo. Supplementary File 9 provides all statistical details, including sample size, *P* values, hazard ratios (HRs) and confidence intervals (CIs). In **c**–**f**, the shaded regions represent the 95% CI for the estimated PACC score. For **a** and **b**, each dot or shape represents the estimated HR for a given disease or mortality. The line extending from the dots or shapes shows the 95% CI for each HR. For **c**–**f**, all *P* values are two-sided, and the corresponding *t* values are reported for signals with $P < 0.05$.

and major depressive disorder but potentially toxic with a high dose, with side effects for depression); antimigraine (for example, prochlorperazine, used to treat psychotic symptoms such as schizophrenia and non-psychotic anxiety but has off-label uses, including emergency use to treat migraine in adults and children); antihypertensive (for example, fenoldopam, a medication used to temporarily treat high blood pressure); antiemetics (for example, trifluoperazine, a medication used as an antipsychotic and an antiemetic); appetite suppressant; treatment of erectile dysfunction; antidiabetic; treatment of pulmonary arterial hypertension; treatment of substance abuse; analgesic; and treatment of sexual dysfunction in women (Fig. 5b). The 15 drugs linked to *ALDH2* encompass a broad range of therapeutic areas, many of which are relevant to aging-related diseases such as neurodegeneration, metabolic disorders and immune system aging. As a central immune organ, the spleen's aging may indicate immune dysfunction that contributes to neuroinflammation. *ALDH2* could influence both immune cell activity in the spleen and neuroinflammation in the brain, especially concerning drugs for neurodegenerative and

psychiatric disorders. Drugs targeting *ALDH2* (or related pathways) for neurodegenerative diseases may also affect systemic aging, positioning them as potential candidates for modulating spleen or immune aging as well (Supplementary Fig. 27).

Potential clinical relevance of the seven MRIBAGs

We highlight the clinical potential of the seven MRIBAGs and their PRSs in predicting (1) the incidence of 53 individual disease conditions (with at least 50 patients) using International Classification of Diseases, Tenth Revision (ICD-10) codes; (2) all-cause mortality risk through survival analyses; and (3) differentiation in cognitive decline trajectories in Alzheimer's disease drug solanezumab (Method 5).

We found that the brain, adipose and pancreas MRIBAGs and the pancreas PRS could significantly ($P < 0.05/7$) predict non-insulin-dependent diabetes mellitus (E119; 1.44 < hazard ratio < 1.70). The brain MRIBAG also predicted personal history of psychoactive substance abuse (Z864; hazard ratio = 1.27 (1.11–1.46); $P = 6.70 \times 10^{-4}$), anxiety disorders (F419; hazard ratio = 1.60 (1.20–2.11); $P = 1.31 \times 10^{-3}$) and gastrointestinal hemorrhage (K922; hazard ratio = 1.46 (1.11–1.91); $P = 6.63 \times 10^{-3}$). The heart MRIBAG was associated with an increased risk of hypertension (I10; hazard ratio = 1.16 (1.05–1.28); $P = 4.80 \times 10^{-3}$), reinforcing findings from genetic correlation and Mendelian randomization analyses (Fig. 6a and Supplementary File 9).

We further found that several MRIBAGs significantly ($P < 0.05/7$) predicted all-cause mortality (Fig. 6b). Among these, the brain and adipose MRIBAGs were identified as risk biomarkers (hazard ratio > 1) for all-cause mortality, whereas the liver and spleen MRIBAGs (hazard ratio < 1) appeared to have a protective effect. We conducted additional sensitivity analyses by excluding participants with any disease diagnoses to assess disease-free survival. Despite the reduced sample size, the protective effect remained evident. For instance, the spleen MRIBAG retained nominal significance under this approach ($P = 0.04$; hazard ratio = 0.62 (0.40–0.99)). Supplementary Table 8a,b and Supplementary Figs. 28 and 29 present a detailed discussion on additional sensitivity check results.

Finally, we assessed the use of the brain MRIBAG for stratifying clinical trial participants into subgroups to (1) demonstrate differential cognitive decline trajectories across subgroups and (2) provide insights for future clinical trial design, using data from the A4 trial. As shown in Fig. 6c,d, participants exhibiting decelerated aging signatures (that is, youthful brain: MRIBAG below the median) showed a higher Preclinical Alzheimer Cognitive Composite (PACC) score than participants within the accelerated aging group (that is, aged brain), both in the drug group ($P = 0.003$ and $t = 2.923$) and the placebo group ($P = 0.001$ and $t = 4.049$). A permutation test showed that the cognitive differences between accelerated and decelerated agers, quantified by t values, were significantly more pronounced in the placebo group compared to the drug group (one-sided $P = 0.03$ based on 1,000 permutations). A one-sided P value was used, as our hypothesis was directional; we specifically aimed to test whether the placebo group exhibited stronger cognitive separation between aging subgroups, consistent with the observed pattern in cognitive trajectories (Fig. 6c,d). Furthermore, neither the decelerated nor the accelerated aging group exhibited significant drug effects (Fig. 6e,f) in preventing cognitive decline.

Discussion

This study introduces seven MRI-based aging clocks by leveraging large-scale multi-organ and multi-omics data from 313,645 participants curated by the MULTI Consortium. Collectively, this work expands existing biological aging clocks and highlights their clinical potential for future research.

Overall, the age prediction models highlight limitations in relying on abdominal MRI features, including a limited number of informative features and substantial multicollinearity, which reduced model performance and hindered generalizability to unseen data. Compared to the brain imaging community[24], the development of imaging protocols and standardized imaging-derived phenotypes (IDPs) for other organs may be relatively underdeveloped. This disparity is reflected in large-scale biobanks such as the UKBB, where the majority of well-characterized IDPs are derived from brain MRI ($N > 4,000$) and heart MRI, whereas imaging-derived features, returned and made available to the community, for other organs are limited in both number and diversity. Several factors contribute to this imbalance, including the historical focus of neuroimaging research, the availability of standardized brain imaging pipelines[25] and the widespread adoption of advanced computational tools tailored for brain MRI analysis. Imaging data from other organs should also be incorporated into this framework. For instance, integrating retina imaging[26] with brain MRI, genetic data and systemic biomarkers could reveal mechanisms underlying aging and disease.

The ProWAS results highlight both organ-specific aging mechanisms and systemic inter-organ communication, providing a foundation for understanding multi-tissue aging dynamics. Such organ specificity of the brain MRIBAG was also reinforced by two previous studies that used plasma proteomics data to develop brain tissue-enriched aging clocks (that is, the brain ProtBAG) using both SomaScan and Olink platforms. Oh et al.[4] identified the three brain tissue-enriched proteins based on Genotype-Tissue Expression (GTEx) data to construct the brain ProtBAG, whereas Wen et al.[13] leveraged the HPA platform, which integrates more extensive curation at both RNA and protein levels, to define brain tissue-specific proteins. Other proteins include CHI3L1, a well-established neuroinflammatory marker elevated in Alzheimer's disease and multiple sclerosis[27], and GDF15, a stress-induced cytokine linked to cognitive decline, Alzheimer's disease and aging[28,29]. The significant metabolites identified in the MetWAS analysis further reinforce both cross-organ and within-organ associations. Although certain proteins or metabolites, such as GDF15 and FABP4, serve as systemic aging markers, others, such as NCAN and IGFBP6, exhibit tissue-specific roles. Shared underlying biological pathways may include inflammation[30], extracellular matrix dysregulation[31] and metabolic dysfunction[32].

The results of our genetic analyses reveal that, although the MRIBAGs share certain genetic underpinnings, each organ also demonstrates a distinct genetic architecture shaped by its unique biological processes. Genetic correlations between the MRIBAGs and other multi-omics aging clocks, as well as the MRIBAGs' associations with various disease endpoints, strengthen the notion of both shared and organ-specific genetic influences. For example, the correlation between the kidney MRIBAG and the renal PhenoBAG highlights how kidney-specific genetic factors can interact with broader multi-organ aging mechanisms at multiple omics levels[33]. Notably, although our analyses show that these organs share common genetic influences that link them to broader disease pathways, they also exhibit unique, organ-specific contributions that reflect their individual physiological roles and functions. Together, these findings provide strong support for the hypothesis that the MRIBAGs share genetic influences across organs while also demonstrating distinct genetic signatures that contribute to the specificity of each organ's role in aging and disease. Supplementary Fig. 30 and Supplementary Table 9 performed mediation analyses to link the MRIBAGs to ProtBAGs and late-life depression[34].

The broad spectrum of drug indications linked to the prioritized genes highlights the opportunity to repurpose existing therapies for targeting aging and age-related diseases in different organs, providing a promising direction for future therapeutic advancements. For example, everolimus, a rapamycin analog, has shown promise in extending lifespan and improving lifespan in humans[35,36], particularly in delaying age-related diseases and improving immune function. It is known for its ability to suppress immune system activity, which can reduce chronic inflammation, a hallmark of aging. Studies have

suggested that chronic low-grade inflammation, often referred to as 'inflammaging'[37], contributes to various age-related diseases, such as cardiovascular disease[38] and Alzheimer's disease[39]. Immunosuppressants may mitigate this process by modulating immune responses and promoting tissue repair and regeneration. Notably, the mTOR pathway[35], targeted by drugs such as everolimus, has been shown to regulate cellular processes such as growth, survival and metabolism, and its inhibition has been linked to lifespan extension in model organisms, including yeast, worms and mice[40].

The counterintuitive finding that higher values of the liver and spleen MRIBAGs are associated with a lower risk of mortality could reflect their role in preserving organ function and overall physiological balance as people age. Potentially, the higher MRIBAG values likely reflect healthier, more resilient organs that maintain their metabolic and immune functions more effectively over time. Specifically, a higher liver MRIBAG may signal better metabolic health, efficient detoxification and regulation of systemic inflammation, which are crucial for maintaining overall homeostasis. Similarly, a higher spleen MRIBAG could indicate better immune function, including more effective immune surveillance and response to infections and other stressors. In the context of aging, these organs might act as 'reserve systems'[41], buffering the body against environmental insults and chronic diseases, which would reduce the risk of mortality.

This study has several limitations. First, the analysis is based on cross-sectional data. Longitudinal studies will be crucial to better investigate the dynamic trajectories of MRIBAGs over time and their temporal relationships with health outcomes. Second, although the analysis focused primarily on genetic and imaging data, environmental factors such as lifestyle, diet and pollution were not considered, yet these may play important roles in aging and disease[42]. Third, the imaging features for the five abdominal organs and tissues were relatively underdeveloped, underscoring the need to implement more advanced feature engineering techniques and include external independent studies. Furthermore, the genetic analysis was focused primarily on individuals of European ancestry, and future studies should extend these findings to underrepresented populations to enhance their generalizability. Future refinement through larger, multi-ethnic GWASs and the inclusion of rare variants may improve their utility and biological interpretability. In addition, the current results derived from the UKBB study would benefit from confirmation in other large-scale cohorts to ensure their generalizability across populations and study designs. Moreover, as a future avenue, studies incorporating more Alzheimer's disease-oriented and AI-derived imaging biomarkers (for example, subtypes informed by tau or amyloid positron emission tomography (PET) data) may enhance statistical power and clinical relevance to benefit future clinical trials. Finally, the observed protective associations for the liver and spleen MRIBAGs are biologically ambiguous and complex. As such, these findings should be interpreted cautiously.

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41591-025-03999-8.

References

1. Eisenstein, M. Pushing the limits of MRI brain imaging. *Nat. Methods* **21**, 1975–1979 (2024).
2. Cole, J. H., Marioni, R. E., Harris, S. E. & Deary, I. J. Brain age and other bodily 'ages': implications for neuropsychiatry. *Mol. Psychiatry* **24**, 266–281 (2019).
3. Argentieri, M. A. et al. Proteomic aging clock predicts mortality and risk of common age-related diseases in diverse populations. *Nat. Med.* **30**, 2450–2460 (2024).
4. Oh, H. S.-H. et al. Organ aging signatures in the plasma proteome track health and disease. *Nature* **624**, 164–172 (2023).
5. Anagnostakis, F. et al. Multi-organ metabolome biological age implicates cardiometabolic conditions and mortality risk. *Nat. Commun.* **16**, 4871 (2025).
6. Mutz, J., Iniesta, R. & Lewis, C. M. Metabolomic age (MileAge) predicts health and life span: A comparison of multiple machine learning algorithms. *Sci. Adv.* **10**, eadp3743 (2024).
7. Teschendorff, A. E. & Horvath, S. Epigenetic ageing clocks: statistical methods and emerging computational challenges. *Nat. Rev. Genet.* **26**, 350–368 (2025).
8. Boquet-Pujadas, A. et al. MUTATE: a human genetic atlas of multiorgan artificial intelligence endophenotypes using genome-wide association summary statistics. *Brief. Bioinform.* **26**, bbaf125 (2025).
9. The MULTI consortium et al. Brain-heart-eye axis revealed by multi-organ imaging genetics and proteomics. *Nat. Biomed. Eng.* https://doi.org/10.1038/s41551-025-01506-5 (2025).
10. Wen, J. Towards a multi-organ and multi-omics medical digital twin. *Nat. Biomed. Eng.* **9**, 1386–1389 (2025).
11. Wen, J. et al. The genetic architecture of biological age in nine human organ systems. *Nat. Aging* **4**, 1290–1307 (2024).
12. Tian, Y. E. et al. Heterogeneous aging across multiple organ systems and prediction of chronic disease and mortality. *Nat. Med.* **29**, 1221–1231 (2023).
13. Wen, J. Refining the generation, interpretation and application of multi-organ, multi-omics biological aging clocks. *Nat. Aging* **5**, 1897–1913 (2025).
14. Wen, J. et al. The genetic architecture of multimodal human brain age. *Nat. Commun.* **15**, 2604 (2024).
15. de Lange, A.-M. G. & Cole, J. H. Commentary: Correction procedures in brain-age prediction. *Neuroimage Clin.* **26**, 102229 (2020).
16. Beheshti, I., Nugent, S., Potvin, O. & Duchesne, S. Bias-adjustment in neuroimaging-based brain age frameworks: a robust scheme. *Neuroimage Clin.* **24**, 102063 (2019).
17. Bashyam, V. M. et al. MRI signatures of brain age and disease over the lifespan based on a deep brain network and 14 468 individuals worldwide. *Brain* **143**, 2312–2324 (2020).
18. Sperling, R. A. et al. Trial of solanezumab in preclinical Alzheimer's disease. *N. Engl. J. Med.* **389**, 1096–1107 (2023).
19. Sperling, R. A. et al. Association of factors with elevated amyloid burden in clinically normal older individuals. *JAMA Neurol.* **77**, 735–745 (2020).
20. Szklarczyk, D. et al. The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. *Nucleic Acids Res.* **51**, D638–D646 (2023).
21. Yang, J., Lee, S. H., Goddard, M. E. & Visscher, P. M. GCTA: a tool for genome-wide complex trait analysis. *Am. J. Hum. Genet.* **88**, 76–82 (2011).
22. Kendler, K. & Neale, M. Endophenotype: a conceptual analysis. *Mol. Psychiatry* **15**, 789–797 (2010).
23. Yi, F. et al. Genetically supported targets and drug repurposing for brain aging: a systematic study in the UK Biobank. *Sci. Adv.* **11**, eadr3757 (2025).
24. Markiewicz, C. J. et al. The OpenNeuro resource for sharing of neuroscience data. *eLife* **10**, e71774 (2021).
25. Gorgolewski, K. et al. Nipype: a flexible, lightweight and extensible neuroimaging data processing framework in Python. *Front. Neuroinform.* **5**, 13 (2011).
26. MacGillivray, T. J. et al. Suitability of UK Biobank retinal images for automatic analysis of morphometric properties of the vasculature. *PLoS ONE* **10**, e0127914 (2015).

27. Connolly, K. et al. Potential role of chitinase-3-like protein 1 (CHI3L1/YKL-40) in neurodegeneration and Alzheimer's disease. *Alzheimers Dement.* **19**, 9–24 (2023).
28. Chiariello, A. et al. The expression pattern of GDF15 in human brain changes during aging and in Alzheimer's disease. *Front. Aging Neurosci.* **14**, 1058665 (2023).
29. Beydoun, M. A. et al. GDF15 and its association with cognitive performance over time in a longitudinal study of middle-aged urban adults. *Brain Behav. Immun.* **108**, 340–349 (2023).
30. Li, X. et al. Inflammation and aging: signaling pathways and intervention therapies. *Signal Transduct. Target. Ther.* **8**, 239 (2023).
31. Statzer, C., Park, J. Y. C. & Ewald, C. Y. Extracellular matrix dynamics as an emerging yet understudied hallmark of aging and longevity. *Aging Dis.* **14**, 670–693 (2023).
32. Amorim, J. A. et al. Mitochondrial and metabolic dysfunction in ageing and age-related diseases. *Nat. Rev. Endocrinol.* **18**, 243–258 (2022).
33. Rowland, J. et al. Uncovering genetic mechanisms of kidney aging through transcriptomics, genomics, and epigenomics. *Kidney Int.* **95**, 624–635 (2019).
34. Wen, J. et al. Characterizing heterogeneity in neuroimaging, cognition, clinical symptoms, and genetics among patients with late-life depression. *JAMA Psychiatry* **79**, 464–474 (2022).
35. Mannick, J. B. et al. mTOR inhibition improves immune function in the elderly. *Sci. Transl. Med.* **6**, 268ra179 (2014).
36. Blagosklonny, M. V. Rejuvenating immunity: 'anti-aging drug today' eight years later. *Oncotarget* **6**, 19405–19412 (2015).
37. Franceschi, C., Garagnani, P., Parini, P., Giuliani, C. & Santoro, A. Inflammaging: a new immune–metabolic viewpoint for age-related diseases. *Nat. Rev. Endocrinol.* **14**, 576–590 (2018).
38. Barcena, M. L., Aslam, M., Pozdniakova, S., Norman, K. & Ladilov, Y. Cardiovascular inflammaging: mechanisms and translational aspects. *Cells* **11**, 1010 (2022).
39. Kosyreva, A. M., Sentyabreva, A. V., Tsvetkov, I. S. & Makarova, O. V. Alzheimer's disease and inflammaging. *Brain Sci.* **12**, 1237 (2022).
40. Harrison, D. E. et al. Rapamycin fed late in life extends lifespan in genetically heterogeneous mice. *Nature* **460**, 392–395 (2009).
41. Atamna, H., Tenore, A. & Lui, F. & Dhahbi, J. M. Organ reserve, excess metabolic capacity, and aging. *Biogerontology* **19**, 171–184 (2018).
42. Argentieri, M. A. et al. Integrating the environmental and genetic architectures of aging and mortality. *Nat. Med.* **31**, 1016–1025 (2025).

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

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, 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 you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. 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-nc-nd/4.0/.

© The Author(s) 2025

[1]Laboratory of AI and Biomedical Science (LABS), Columbia University, New York, NY, USA. [2]Laboratory of Behavioral Neuroscience, National Institute on Aging, National Institutes of Health, Bethesda, MD, USA. [3]Artificial Intelligence in Biomedical Imaging Laboratory (AIBIL), Center for AI and Data Science for Integrated Diagnostics (AI2D), Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA. [4]Systems Lab, Department of Psychiatry, Melbourne Medical School, The University of Melbourne, Melbourne, Victoria, Australia. [5]Department of Brain Sciences and Department of Computing, Imperial College London, London, UK. [6]Alzheimer's Therapeutic Research Institute, Keck School of Medicine of the University of Southern California, San Diego, CA, USA. [7]Robert N Butler Columbia Aging Center and Department of Epidemiology, Columbia University Mailman School of Public Health, New York, NY, USA. [8]National Institute on Aging, National Institutes of Health, Bethesda, MD, USA. [9]Center for Innovation in Imaging Biomarkers and Integrated Diagnostics (CIMBID), Department of Radiology, Columbia University, New York, NY, USA. [10]Department of Radiology, Columbia University, New York, NY, USA. [11]Department of Biomedical Engineering, Columbia University, New York, NY, USA. [12]Zuckerman Institute, Columbia University, New York, NY, USA. [13]New York Genome Center (NYGC), New York, NY, USA. [14]Data Science Institute (DSI), Columbia University, New York, NY, USA. *A list of authors and their affiliations appears at the end of the paper. ✉e-mail: junhao.wen89@gmail.com

The MULTI Consortium

Junhao Wen[1,9,10,11,12,13,14], Christos Davatzikos[3], Ye Ella Tian[4], Wenjia Bai[5], Michael S. Rafii[6], Paul Aisen[6], Daniel W. Belsky[7], Keenan A. Walker[2], Andrew Zalesky[4] & Luigi Ferrucci[8]

Methods

Method 1: The MULTI Consortium

The MULTI Consortium is an ongoing initiative to integrate and consolidate multi-organ data and multi-omics data, including imaging, genetics and proteomics. Building upon existing consortia and studies, such as those listed below, MULTI aims to curate and harmonize the data to model human aging and disease across the lifespan. In the present study, in total, individual-level data for 313,645 participants were analyzed, including multi-organ MRI data across seven organs and tissues (Category ID: 100003), genetics, plasma proteomics data and metabolomics data. GWAS summary statistics from FinnGen and the Psychiatric Genomics Consortium (PGC) were downloaded and harmonized for our post-GWAS analyses. Refer to Supplementary Table 1 for comprehensive information, including the complete list of data analyzed and their respective sample characteristics.

The MULTI Consortium received approval from the institutional review board at Columbia University (AAAV6751). Participants provided written informed consent for each study that included individual-level data, whereas FinnGen and PGC provided only summary-level statistics. Specifically, the UKBB obtained research ethics approval and received guidance from an Ethics Advisory Committee, with additional ethical input provided to the Access Committee, which oversees decisions regarding access to data and biological samples. The BLSA protocol was approved by the institutional review board of the National Institute of Environmental Health Science, National Institutes of Health (03AG0325). For the A4 trial, institutional review board approval was obtained at each trial site, and all participants provided written informed consent. The study was conducted in accordance with International Council for Harmonization Good Clinical Practice guidelines.

UKBB. The UKBB[43] is a population-based research initiative comprising approximtely 500,000 individuals from the United Kingdom between 2006 and 2010. Ethical approval for the UKBB study was secured, and information about the ethics committee can be found here: https://www.ukbiobank.ac.uk/learn-more-about-uk-biobank/governance/ethics-advisory-committee.

This study collectively analyzed multi-organ MRI data from seven organ systems and tissues, including the brain, heart, liver, pancreas, spleen, adipose and kidney. For the genetic data, we conducted a quality check on the imputed genotype data[43] for the entire UKBB population (approximately 500,000 individuals). Subsequently, we merged the processed data with the organ-specific populations for all genetic analyses. Finally, we also included Olink plasma proteomics data released by the UK Biobank Pharma Proteomics Project (UKB-PPP) and metabolomics data from the UKBB, which are detailed in the following sections.

BLSA. The main goal of the BLSA is to understand the normal aging process. Tracking physiological and cognitive changes over time aims to identify risk factors for age-related diseases, study patterns of decline and discover predictors of healthy aging. BLSA[44,45] brain MRI and SomaScan proteomics data (https://www.blsa.nih.gov/) were used to compare and replicate the ProWAS results from the UKBB Olink data. After quality checks in this study, we included 1,114 brain MRI scans at baseline and measurements of 7,268 plasma proteins from 909 participants quantified with the SomaScan version 4.1 platform. Age (years), sex (male/female), race (white/non-white) and education level (years) were defined based on participant self-reports.

Anti-amyloid treatment in asymptomatic Alzheimer's disease. The A4 study[18,19] (https://atri.usc.edu/study/a4-study/) is a clinical trial study to test a specific way to prevent memory loss associated with Alzheimer's disease (clinical trial number: NCT02008357). The A4 study focused on symptom-free adults at higher risk for Alzheimer's disease to assess whether an investigational drug (that is, solanezumab) could slow memory decline linked to amyloid plaques in the brain. It also examined whether solanezumab could delay Alzheimer's disease progression, measuring related brain changes using imaging, blood biomarkers and baseline PET scans to assess amyloid levels. This study analyzed 1,055 participants at baseline with brain MRI scans to derive the brain MRIBAG. Longitudinal outcomes from the clinical trial, with the PACC score as the primary measure over 312 weeks, were included. The PACC scores between groups were evaluated at week 240.

FinnGen. The FinnGen[46] study is a large-scale genomics initiative that has analyzed more than 500,000 Finnish biobank samples and correlated genetic variation with health data to understand disease mechanisms and predispositions. The project is a collaboration between research organizations and biobanks within Finland and international industry partners. For the benefit of research, FinnGen generously made their GWAS findings accessible to the wider scientific community (https://www.finngen.fi/en/access_results). This research used the publicly released GWAS summary statistics (version R9), which became available on 11 May 2022, after harmonization by the consortium. No individual data were used in the current study.

FinnGen published the R9 version of GWAS summary statistics via REGENIE software (version 2.2.4)[47], covering 2,272 disease endpoints, including 2,269 binary traits and three quantitative traits. The GWAS model encompassed covariates such as age, sex, the initial 10 genetic principal components and the genotyping batch. Genotype imputation was referenced on the population-specific SISu version 4.0 panel. We included GWAS summary statistics for 521 FinnGen disease endpoints in our analyses.

PGC. PGC[48] is an international collaboration of researchers studying the genetic basis of psychiatric disorders. PGC aims to identify and understand the genetic factors contributing to various psychiatric disorders, such as schizophrenia, bipolar disorder, major depressive disorder and others. The GWAS summary statistics were acquired from the PGC website (https://pgc.unc.edu/for-researchers/download-results/), underwent quality checks and were harmonized to ensure seamless integration into our analysis. No individual data were used from PGC. Each study detailed its specific GWAS models and methodologies, and the consortium consolidated the release of GWAS summary statistics derived from individual studies. In the present study, we included summary data for four brain diseases for which allele frequencies were present.

Method 2: Multi-organ MRI analyses to derive the seven MRIBAGs

In our previous analysis of raw brain MRI data from the UKBB and other studies included in this work, we extracted 119 gray matter regions of interest from T1-weighted MRI to generate brain MRIBAGs through the iSTAGING consortium[49]. For the heart MRIBAG, we used 80 heart MRI traits from Bai et al.[50] and used these imaging features in a previous study[9]. To develop the remaining five MRIBAGs based on abdominal MRI (Category ID: 105 (refs. 51–56)), we incorporated 16 imaging features for the adipose MRIBAG, four for the liver MRIBAG, three for the kidney MRIBAG, three for the spleen MRIBAG and three for the pancreas MRIBAG. Notably, in the development of adipose and kidney MRIBAGs, we observed a high degree of collinearity among certain imaging features, which led to overfitting. To mitigate this issue, we removed highly correlated features to improve model robustness.

Additional quality checks for the abdominal MRI. The abdominal MRI data underwent initial quality control procedures prior to public release, including the removal of low-quality images and biologically implausible values. As noted on the UKBB website (https://biobank.ndph.ox.ac.uk/ukb/label.cgi?id=158), different pipelines were used to generate imaging features from abdominal MRI, and data should not be combined without careful consideration. To account for this, we ensured that each

abdominal feature was derived from a consistent pipeline, selecting the one with the largest available sample size across different pipelines. For liver MRI metrics specifically, we included only images acquired using the IDEAL protocol (Data-Field 40063, acquisition protocol 2). Following these additional quality control steps, our analyses incorporated abdominal MRI biomarkers from all available UKBB participants for the liver, kidney, spleen, pancreas and adipose MRIBAGs.

AI/ML models. Our previous study systematically evaluated age prediction performance across various AI/ML models using multimodal brain MRI features[14] for the brain MRIBAG. Applying the same framework, we assessed the performance of models in deriving the six additional MRIBAGs using several ML methods. Hyperparameter tuning was performed through nested, repeated holdout cross-validation[57] with 50 repetitions (80% training/validation and 20% testing). Specifically, the hyperparameters of each model were tuned using grid search: LASSO regression's alpha (α), linear support vector regressor's C and both α and L1-ratio for the elastic net. Due to the flexibility and large hyperparameter space of neural networks (for example, number of layers and neurons), they were not included in the nested cross-validation pipeline. Instead, the neural network was manually tuned and evaluated on the cross-validation test dataset, and the final model was subsequently assessed on the within-distribution, holdout test set. The within-distribution, holdout test dataset was held out to unbiasedly evaluate model performance (Supplementary Fig. 31).

Dataset split. To rigorously train the AI/ML models, we first defined participants without any pathologies based on ICD code and clinical history as CN (N = 5,291 for the heart MRIBAG, varying across the seven organs/tissues). We further split the CN into the following datasets:

- CN within-distribution, holdout test dataset: 500 participants were randomly drawn from the CN population. Within-distribution, holdout test datasets are ideal for objectively evaluating ML model performance, especially in studies with large sample sizes, such as the present one.
- CN training/validation dataset: 80% of the remaining 4,791 CN (taking the heart MRIBAG for illustration) were used for the inner loop 10-fold cross-validation for hyperparameter selection.
- CN cross-validated test dataset: 20% of the remaining 4,791 CN were used for the outer loop 50 repetitions.
- Patient dataset: 29,785 patients who have at least one ICD-10-based diagnosis.

The CN training/validation/test datasets were used for model development and were employed with a nested cross-validation procedure for four AI/ML models (LASSO regression and support vector regressor, elastic net and neural network), whereas the within-distribution, holdout test set provided an unbiased assessment of model performance. Model evaluation metrics included MAE and Pearson's r. Notably, consistent with our previous studies, only CN participants were included in the training/validation dataset for modeling and clinical interpretation considerations[13]. The BAG was determined by subtracting the participant's chronological age from the AI/ML-predicted age. Age bias correction was applied using the approach outlined by Beheshti et al.[16] and was systematically discussed in our previous comment[13]. To summarize, we applied an age bias correction method commonly used in neuroimaging studies, where a linear regression of the BAG on chronological age is used to estimate and remove systematic age-related bias via the model's slope and intercept. Notably, to address potential domain shift in the patient data compared to the CN training set, we computed the regression parameters directly from the patient set to more accurately correct for this bias (refer to our comment[13] for further details). Supplementary Fig. 32 details the feature importance across UKBB, BLSA and A4.

Method 3: ProWAS and MetWAS with the seven MRIBAGs

ProWAS. We used the original dataset (Category ID: 1838), which was analyzed and shared with the research community by the UKB-PPP[58]. The initial quality control procedures were described in the original study[59], and we implemented additional quality control steps as outlined below. Our analysis focused on the first instance of the proteomics data ('instance' = 0). We then integrated Olink files containing coding information, batch numbers, assay details and limit of detection (LOD) data (Category ID: 1839) by matching them to the proteomics dataset ID. Finally, we excluded Normalized Protein eXpression (NPX) values that fell below the protein-specific LOD.

We conducted ProWASs by linking the seven MRIBAGs to 2,923 unique plasma proteins measured in 53,016 participants (with sample sizes ranging from 10,018 to 39,489 per protein after quality control) using the Olink platform. The linear regression model adjusted for common covariates, including age (Field ID: 21003), sex (Field ID: 31), body weight (Field ID: 21002), height (Field ID: 50), waist circumference (Field ID: 48), body mass index (BMI) (Field ID: 23104), assessment center (Field ID: 54), protein batch number (Category ID: 1839), LOD (Category ID: 1839) and the first 40 genetic principal components. Additionally, organ-specific covariates were incorporated, such as brain positioning in the scanner (lateral, transverse and longitudinal; Field ID: 25756–25758), head motion (Field ID: 25741) and intracranial volume for the brain MRIBAG as well as diastolic (Field ID: 4079) and systolic (Field ID: 4080) blood pressure for the heart MRIBAG, among others. Multiple testing corrections were applied using Bonferroni adjustment ($P < 0.05/2{,}923/7$). To identify and exclude extreme outliers, we defined an upper threshold as the mean plus six times the s.d. for each protein/metabolite. Finally, ProWAS signals identified in the UKBB Olink dataset were compared with SomaScan data from the BLSA.

MetWAS. We used the original data (Category ID: 220), which were analyzed and made available to the community by Nightingale Health Plc. The original data (1) were calibrated absolute concentrations (or ratios) and not raw NMR spectra and (2), before release, had already been subjected to quality control procedures by Nightingale Health Plc[60]. Following the additional procedures described in Ritchie et al.[61], we performed additional quality check steps to remove a range of unwanted technical variations, including shipping batch, 96-well plate, well position, aliquoting robo and aliquot tip. We focused our analysis on the first instance of the metabolomics data ('instance' = 0). The analysis included 327 metabolites (comprising both small molecules and lipoprotein measures), of which 107 were non-derived metabolites and the remainder were composite metabolites, across 274,247 participants. Descriptions of these metabolites are provided in Supplementary Table 3.

We conducted MetWASs by linking the seven MRIBAGs to 327 plasma metabolites. The linear regression model adjusted for common covariates, including age (Field ID: 21003), sex (Field ID: 31), body weight (Field ID: 21002), height (Field ID: 50), waist circumference (Field ID: 48), BMI (Field ID: 23104), assessment center (Field ID: 54) and the first 40 genetic principal components. Additionally, organ-specific covariates were incorporated, such as brain positioning in the scanner (lateral, transverse and longitudinal; Field ID: 25756–25758), head motion (Field ID: 25741) and intracranial volume for the brain MRIBAG, as well as diastolic (Field ID: 4079) and systolic (Field ID: 4080) blood pressure for the heart MRIBAG, among others. Multiple testing corrections were applied using Bonferroni adjustment ($P < 0.05/327/7$).

Method 4: Genetic analyses

We used the imputed genotype data from the UKBB for all genetic analyses. Our quality check pipeline focused on European ancestry in the UKBB (6,477,810 SNPs passing quality checks), and the quality-checked genetic data were merged with respective organ-specific populations for GWAS. We summarize our genetic quality check steps. First, we

skipped the step for family relationship inference[62] because the linear mixed model via fastGWA[63] inherently addresses population stratification, encompassing additional cryptic population stratification factors. We then removed duplicated variants from all 22 autosomal chromosomes. Individuals whose genetically identified sex did not match their self-acknowledged sex were removed. Other excluding criteria included (1) individuals with more than 3% of missing genotypes; (2) variants with minor allele frequency (MAF; dosage mode) of less than 1%; (3) variants with more than 3% missing genotyping rate; and (4) variants that failed the Hardy–Weinberg test at 1×10^{-10}. To further adjust for population stratification,[64] we derived the first 40 genetic principal components using FlashPCA software[65]. Details of the genetic quality check protocol are described elsewhere[11,14,34,49,66,67].

GWAS. We applied a linear mixed model regression to the European ancestry populations using fastGWA[63] implemented in GCTA[21]. We used fastGWA to perform the seven MRIBAG GWASs, adjusting common variates, including age, dataset status (training/validation/test or within-distribution, holdout test), age-squared, sex, interactions of age with sex, BMI, waist circumference, standing height, weight and the first 40 genetic principal components, as well as organ-specific covariates, including the brain scan positions for the brain MRIBAG and systolic/diastolic blood pressure for the heart MRIBAG. We applied a genome-wide significance threshold (5×10^{-8}) to annotate the significant independent genomic loci. We previously conducted GWAS of 2,923 plasma proteins and 327 metabolites using fastGWA[5,68].

Annotation of genomic loci. For all GWASs, genomic loci were annotated using FUMA[69]. For genomic loci annotation, FUMA initially identified lead SNPs (correlation $r^2 \leq 0.1$, distance < 250 kb) and assigned them to non-overlapping genomic loci. The lead SNP with the lowest *P* value (that is, the top lead SNP) represented the genomic locus. Further details on the definitions of top lead SNP, lead SNP, independent significant SNP and candidate SNP can be found in the FUMA documentation (https://fuma.ctglab.nl/). For visualization purposes in Fig. 3a, we mapped the top lead SNP of each locus to the cytogenetic regions based on the GRCh37 cytoband.

SNP-based heritability. We estimated the SNP-based heritability (h^2) using GCTA[21] with the same covariates as in GWAS. GCTA estimates the SNP-based heritability using a method called restricted maximum likelihood (REML) to quantify the proportion of phenotypic variance in a trait that the additive effects of all common SNPs can explain. The main steps involved in GCTA include constructing the genetic relationship matrix, modeling phenotypic variance and using REML to estimate the h^2.

Partitioned heritability estimate. The partitioned heritability analysis via stratified linkage disequilibrium score regression calculates the extent to which heritability enrichment can be attributed to predefined and annotated genome regions and categories[70]. Three sets of functional categories and cell-specific and tissue-specific types were considered. First, the partitioned heritability was calculated for 53 general functional categories (one including the entire set of SNPs). The 53 functional categories are not specific to any cell type and include coding regions, untranslated regions, promoter regions and intronic regions. The details of the 53 categories are described elsewhere[70]. Subsequently, cell-type-specific and tissue-type-specific partitioned heritability was estimated using gene sets from Cahoy et al.[71] for three main cell types (that is, astrocyte, neuron and oligodendrocyte), multi-tissue chromatin state-specific data (Roadmap[72] and ENTEx[73]) and multi-tissue gene expression data (GTEx version 8 (ref. 74)). Bonferroni correction was performed for all tested annotations and categories. The detailed methodologies for the stratified linkage disequilibrium score regression are presented in the original work[70]. The linkage disequilibrium scores and allele frequencies for the European ancestry were obtained from a predefined version based on data from the 1000 Genomes project.

Genetic correlation. We estimated the genetic correlation (g_c) using LDSC[75] software. We employed precomputed linkage disequilibrium scores from the 1000 Genomes of European ancestry, maintaining default settings for other parameters in LDSC. It is worth noting that LDSC corrects for sample overlap, ensuring an unbiased genetic correlation estimate[76]. Statistical significance was determined using Bonferroni correction.

Two-sample bidirectional Mendelian randomization. We constructed two-sample bidirectional Mendelian randomization by linking the seven MRIBAGs and 525 disease endpoints from FinnGen[46] and PGC[48]. In total, two networks were established: *MRIBAG2DE* and *DE2MRIBAG*. The systematic quality-checking procedures to ensure unbiased exposure/outcome variable and instrumental variable selection are detailed below.

We used a two-sample Mendelian randomization approach implemented in the TwoSampleMR package[77] to infer the causal relationships within these networks. We employed five distinct Mendelian randomization methods, including the inverse variance weighted (IVW) method, Egger, weighted median and simple mode and weighted mode estimators. The STROBE-MR Statement[78] guided our analyses to increase transparency and reproducibility, encompassing the selection of exposure and outcome variables, reporting statistics and implementing sensitivity checks to identify potential violations of underlying assumptions. First, we performed an unbiased quality check on the GWAS summary statistics. Notably, the absence of population overlapping bias[79] was confirmed, given that FinnGen and UKBB participants largely represent populations of European ancestry without explicit overlap with the UKBB. PGC GWAS summary data were ensured to exclude UKBB participants. Furthermore, the GWAS summary statistics from all consortia were based on or lifted to GRCh37. Subsequently, we selected the effective exposure variables by assessing the statistical power of the exposure GWAS summary statistics in terms of instrumental variables, ensuring that the number of instrumental variables exceeded seven before harmonizing the data. Crucially, the function 'clump_data' was applied to the exposure GWAS data, considering linkage disequilibrium. The function 'harmonise_data' was then used to harmonize the GWAS summary statistics of the exposure and outcome variables. Bonferroni correction was applied to all tested traits based on the number of effective disease endpoints.

Finally, we conducted multiple sensitivity analyses. First, we conducted a heterogeneity test to scrutinize potential violations of the instrumental variables' assumptions. To assess horizontal pleiotropy, which indicates the instrumental variables' exclusivity assumption[80], we used a funnel plot, single-SNP Mendelian randomization methods and the Egger estimator. Furthermore, we performed a leave-one-out analysis, systematically excluding one instrument (SNP/instrumental variable) at a time, to gauge the sensitivity of the results to individual SNPs.

PRS calculation. PRS was computed using split-sample GWASs (split1 and split2) for the seven MRIBAG GWASs. The PRS weights were established using split1/discovery GWAS data as the base/training set, whereas the split2/replication GWAS summary statistics served as the target/testing data. Both base and target data underwent rigorous quality control procedures involving several steps: (1) excluding duplicated and ambiguous SNPs in the base data and (2) excluding high heterozygosity samples in the target data and (3) eliminating duplicated, mismatching and ambiguous SNPs in the target data.

After completing the quality control procedures, PRS for the split2 and split1 groups was calculated using the PRS-CS[81] method. PRS-CS applies a continuous shrinkage prior, which adjusts the SNP effect

sizes based on their linkage disequilibrium structure. SNPs with weaker evidence are 'shrunk' toward zero, whereas those with stronger evidence retain larger effect sizes. This avoids overfitting and improves prediction performance. The shrinkage parameter was not set, and the algorithm learned it via a fully Bayesian approach.

Genetic evidence for prioritizing potential druggable genes. We conducted multiple analyses to establish genetic evidence for prioritizing potential druggable genes in future drug repurposing. Our central hypothesis is that genes with functional implications and causal roles validated across multiple omics layers, such as genomics, transcriptomics and proteomics, are more likely to be actionable for drug repurposing, as they offer a stronger and more robust foundation for identifying therapeutic targets[82,83].

First, we performed SNP-to-gene mapping via three approaches: positional mapping, eQTL mapping and chromatin interaction mapping. For eQTL and chromatin interaction mapping, we used data from various resources, including GTEx, PsychENCODE, EyeGEx, TIGER, DICE, eQTLGen, Blood eQTL, MuTHER, CommonMind Consortium, BRAINEAC, FANTOM5, the Brain Open Chromatin Atlas and the Roadmap 111 epigenomes, all consolidated through the FUMA platform. By requiring both eQTL and chromatin interaction support, we enhance confidence in selecting functionally relevant genes because eQTL mapping identifies genetic variants that regulate gene expression. Chromatin interaction mapping links distal regulatory elements (for example, enhancers) to target genes, offering spatial genomic context. Together, these criteria ensure that the prioritized genes are not only associated with GWAS loci but also have strong regulatory and functional evidence, making them more biologically plausible as druggable targets.

Subsequently, we connected MRIBAG, as an imaging-derived endophenotype, to aging clocks derived from plasma proteomics and metabolomics to further strengthen the genetic evidence for prioritizing these druggable genes. To achieve this, we used a Bayesian co-localization approach to assess whether the genomic loci associated with the seven MRIBAGs share common causal variants with the 11 ProtBAGs or the five MetBAGs. We used the R package 'coloc' to investigate the genetic co-localization signals between two traits (for example, the brain MRIBAG versus the brain ProtBAG) at each genomic locus associated with the MRIBAG. This method ('coloc.abf') examines the posterior probability to evaluate hypothesis H4 (PP.H4.ABF), which suggests the presence of a single shared causal variant associated with both traits within a specific genomic locus. To determine the significance of the H4 hypothesis, we set a threshold of PP.H4.ABF > 0.8 (ref. 84). All other parameters (for example, the prior probability of p_{12}) were set as the default. For each pair of traits, the genomic locus was defined by default from FUMA for one trait, and then the 'coloc' package extracted and harmonized the GWAS summary statistics within this locus for the other trait.

Finally, we searched the DGIdb platform (https://dgidb.org/) for genes that demonstrated both functional mapping evidence and causal associations to investigate their drug–gene interactions, revealing existing drugs and their therapeutic indications.

Method 5: Prediction analyses for the risk of mortality and single disease endpoints, and cognitive decline in the solanezumab drug

We investigated the clinical promise of the seven MRIBAGs and their PRSs (MRIBAG-PRS) in three sets of prediction analyses: (1) survival analysis for the incidence of single disease entities based on the ICD-10 code, (2) survival analysis for the risk of all-cause mortality and (3) differential response to the Alzheimer's disease drug (solanezumab).

Survival analysis for ICD-based single disease endpoint. We employed a Cox proportional hazard model while adjusting for covariates (that is, age and sex) to test the associations of the seven MRIBAGs with the incidence of ICD-based single disease entities. The covariates were included as additional right-side variables in the model. To train the model, the 'time' variable was determined by calculating the difference between the date of death (Field ID: 40000) for cases (or the censoring date for non-cases) and the date attending the assessment center (Field ID: 53). Participants who were diagnosed for a specific disease of interest after enrolling in the study were classified as cases; non-cases were defined by participants without any disease diagnoses.

Survival analysis for mortality risk. We employed a Cox proportional hazard model while adjusting for covariates (that is, age and sex) to test the associations of the seven MRIBAGs and seven MRIBAG-PRSs with all-cause mortality. The covariates were included as additional right-side variables in the model. The hazard ratio, $\exp(\beta_R)$, was calculated and reported as the effect size measure that indicates the influence of each biomarker on the risk of mortality. To train the model, the 'time' variable was determined by calculating the difference between the date of death (Field ID: 40000) for cases (or the censoring date for non-cases) and the date attending the assessment center (Field ID: 53). Participants who died after enrolling in the study were classified as cases. We also conducted a disease-free survival analysis, including only participants without any diagnosed conditions, to eliminate potential confounding effects from disease pathology.

Differential trajectories of cognitive decline based on brain MRIBAG-defined decelerated (youthful brain) and accelerated (aged brain) aging groups from drug outcomes in the A4 study. To evaluate our hypothesis that individuals with varying aging clock paces may progress differently to cognitive decline, we analyzed clinical trial data from the A4 (ref. 18) study. Specifically, we investigated whether brain MRIBAG could be used to stratify participants into decelerated and accelerated aging groups and demonstrated different cognitive profiles at 240 weeks by taking solanezumab. The original trial did not demonstrate cognitive decline slowing compared to the placebo over 240 weeks, with the treatment group even showing a slight worsening in the PACC score. To evaluate this, we performed four comparisons using natural cubic spline modeling: (1) decelerated aging (for example, below the brain MRIBAG median) versus accelerated aging (for example, above the median) within the drug group and (2) within the placebo group; (3) between drug and placebo groups among participants with accelerated aging; and (4) between drug and placebo groups among those with decelerated aging.

From a statistical perspective, this is to model repeated measures (that is, PACC as the primary trial outcome for global cognition) as a continuous outcome in a mixed-effect model. We used the same method proposed by the original work from Donohue et al.[85], in which the authors proposed a constrained longitudinal data analysis with natural cubic splines that treated time as continuous and used test version effects to model the mean over time for PACC. Fixed effects included the following terms: (1) spline basis expansion terms (two terms); (2) interaction of the spline basis expansion terms with treatment (two terms); (3) the version of the PACC test implemented; (4) baseline age; (5) education; (6) *APOE4* carrier status (yes/no); and (7) baseline florbetapir cortical standardized uptake value ratio (SUVR) value. At week 240, we compared the mean PACC values between each pair of the two groups.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The GWAS summary statistics and pre-trained AI models from this study are publicly accessible via the MEDICINE Knowledge Portal (https://labs-laboratory.com/medicine/) and Synapse (https://www.synapse.

org/Synapse:syn64923248/wiki/630992 (ref. 86)). Our study used data generated by the HPA (https://www.proteinatlas.org). GWAS summary data for the disease endpoints were downloaded from the official websites of FinnGen (R9: https://www.finngen.fi/en/access_results) and PGC (https://pgc.unc.edu/for-researchers/download-results/). Individual data from the UKBB can be requested with proper registration at https://www.ukbiobank.ac.uk/. The A4 study data can be requested at https://www.a4studydata.org/. The NIA-supported BLSA data can be requested at https://www.nia.nih.gov/research/labs/blsa. Due to privacy considerations, allele frequency information is available from J.W. (junhao.wen89@gmail.com) upon request for research purposes only, with responses typically provided within 1 week.

Code availability

The software and resources used in this study are all publicly available: MLNI[87] (version 0.1.2 and version 0.1.5): https://github.com/anbai106/mlni, MRIBAG generation; FUMA (version 1.5.0): https://fuma.ctglab.nl/, gene mapping and genomic locus annotation; GCTA (version 1.94.1): https://yanglab.westlake.edu.cn/software/gcta/#Overview, fastGWA; LDSC (git version aa33296): https://github.com/bulik/ldsc, genetic correlation; TwoSampleMR (version 0.5.6): https://mrcieu.github.io/TwoSampleMR/index.html, Mendelian randomization; PRScs (release date: 10 August 2023): https://github.com/getian107/PRScs, PRS calculation; Lifelines (version 0.27.8): https://lifelines.readthedocs.io/en/latest/, survival analysis; String (version 12.0): https://string-db.org/, protein analyses; and coloc (version 5): https://github.com/chr1swallace/coloc; Bayesian co-localization. All main figures, extended data figures, the code used to derive the seven MRIBAGs and the supplementary files are available on Zenodo: https://zenodo.org/records/15857817 (ref. 88).

References

43. Bycroft, C. et al. The UK Biobank resource with deep phenotyping and genomic data. *Nature* **562**, 203–209 (2018).
44. Normal Human Aging: The Baltimore Longitudinal Study of Aging. *J. Gerontol.* **40**, 767 (1985).
45. Duggan, M. R. et al. Proteomic analyses reveal plasma EFEMP1 and CXCL12 as biomarkers and determinants of neurodegeneration. *Alzheimers Dement.* **20**, 6486–6505 (2024).
46. Kurki, M. I. et al. FinnGen provides genetic insights from a well-phenotyped isolated population. *Nature* **613**, 508–518 (2023).
47. Mbatchou, J. et al. Computationally efficient whole-genome regression for quantitative and binary traits. *Nat. Genet.* **53**, 1097–1103 (2021).
48. O'Donovan, M. C. What have we learned from the Psychiatric Genomics Consortium. *World Psychiatry* **14**, 291–293 (2015).
49. Wen, J. et al. Genomic loci influence patterns of structural covariance in the human brain. *Proc. Natl Acad. Sci. USA* **120**, e2300842120 (2023).
50. Bai, W. et al. A population-based phenome-wide association study of cardiac and aortic structure and function. *Nat. Med.* **26**, 1654–1662 (2020).
51. Mojtahed, A. et al. Reference range of liver corrected T1 values in a population at low risk for fatty liver disease—a UK Biobank sub-study, with an appendix of interesting cases. *Abdom. Radiol. (NY)* **44**, 72–84 (2019).
52. Whitcher, B. et al. Precision MRI phenotyping enables detection of small changes in body composition for longitudinal cohorts. *Sci. Rep.* **12**, 3748 (2022).
53. Parisinos, C. A. et al. Genome-wide and Mendelian randomisation studies of liver MRI yield insights into the pathogenesis of steatohepatitis. *J. Hepatol.* **73**, 241–251 (2020).
54. Karlsson, A. et al. Automatic and quantitative assessment of regional muscle volume by multi-atlas segmentation using whole-body water-fat MRI. *J. Magn. Reson. Imaging* **41**, 1558–1569 (2015).
55. Borga, M. et al. Validation of a fast method for quantification of intra-abdominal and subcutaneous adipose tissue for large-scale human studies. *NMR Biomed.* **28**, 1747–1753 (2015).
56. Liu, Y. et al. Genetic architecture of 11 organ traits derived from abdominal MRI using deep learning. *eLife* **10**, e65554 (2021).
57. Wen, J. et al. Convolutional neural networks for classification of Alzheimer's disease: overview and reproducible evaluation. *Med. Image Anal.* **63**, 101694 (2020).
58. Li, W. U. K. Biobank pharma proteomics resource. *Nat. Genet.* **55**, 1781 (2023).
59. Sun, B. B. et al. Plasma proteomic associations with genetics and health in the UK Biobank. *Nature* **622**, 329–338 (2023).
60. Würtz, P. et al. Quantitative serum nuclear magnetic resonance metabolomics in large-scale epidemiology: a primer on -omic technologies. *Am. J. Epidemiol.* **186**, 1084–1096 (2017).
61. Ritchie, S. C. et al. Quality control and removal of technical variation of NMR metabolic biomarker data in ~120,000 UK Biobank participants. *Sci. Data* **10**, 64 (2023).
62. Manichaikul, A. et al. Robust relationship inference in genome-wide association studies. *Bioinformatics* **26**, 2867–2873 (2010).
63. Jiang, L. et al. A resource-efficient tool for mixed model association analysis of large-scale data. *Nat. Genet.* **51**, 1749–1755 (2019).
64. Price, A. L., Zaitlen, N. A., Reich, D. & Patterson, N. New approaches to population stratification in genome-wide association studies. *Nat. Rev. Genet.* **11**, 459–463 (2010).
65. Abraham, G., Qiu, Y. & Inouye, M. FlashPCA2: principal component analysis of Biobank-scale genotype datasets. *Bioinformatics* **33**, 2776–2778 (2017).
66. Wen, J. et al. Genetic and clinical correlates of two neuroanatomical AI dimensions in the Alzheimer's disease continuum. *Transl. Psychiatry* **14**, 420 (2024).
67. Wen, J. et al. Neuroimaging endophenotypes reveal underlying mechanisms and genetic factors contributing to progression and development of four brain disorders. *Nat. Biomed. Eng.* https://doi.org/10.1038/s41551-025-01412-w (2025).
68. Multiorgan biological age shows that no organ system is an island. *Nat. Aging* **4**, 1182–1183 (2024).
69. Watanabe, K., Taskesen, E., van Bochoven, A. & Posthuma, D. Functional mapping and annotation of genetic associations with FUMA. *Nat. Commun.* **8**, 1826 (2017).
70. Finucane, H. K. et al. Partitioning heritability by functional annotation using genome-wide association summary statistics. *Nat. Genet.* **47**, 1228–1235 (2015).
71. Cahoy, J. D. et al. A transcriptome database for astrocytes, neurons, and oligodendrocytes: a new resource for understanding brain development and function. *J. Neurosci.* **28**, 264–278 (2008).
72. Bernstein, B. E. et al. The NIH Roadmap Epigenomics Mapping Consortium. *Nat. Biotechnol.* **28**, 1045–1048 (2010).
73. Dunham, I. et al. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**, 57–74 (2012).
74. GTEx Consortium. The Genotype-Tissue Expression (GTEx) project. *Nat. Genet.* **45**, 580–585 (2013).
75. Bulik-Sullivan, B. K. et al. LD Score regression distinguishes confounding from polygenicity in genome-wide association studies. *Nat. Genet.* **47**, 291–295 (2015).
76. Bulik-Sullivan, B. et al. An atlas of genetic correlations across human diseases and traits. *Nat. Genet.* **47**, 1236–1241 (2015).
77. Hemani, G. et al. The MR-Base platform supports systematic causal inference across the human phenome. *eLife* **7**, e34408 (2018).
78. Skrivankova, V. W. et al. Strengthening the Reporting of Observational Studies in Epidemiology Using Mendelian Randomization: the STROBE-MR Statement. *JAMA* **326**, 1614–1621 (2021).

79. Sanderson, E. et al. Mendelian randomization. *Nat. Rev. Methods Primers* **2**, 6 (2022).
80. Bowden, J. et al. A framework for the investigation of pleiotropy in two-sample summary data Mendelian randomization. *Stat. Med.* **36**, 1783–1802 (2017).
81. Ge, T., Chen, C.-Y., Ni, Y., Feng, Y.-C. A. & Smoller, J. W. Polygenic prediction via Bayesian regression and continuous shrinkage priors. *Nat. Commun.* **10**, 1776 (2019).
82. Qi, T., Song, L., Guo, Y., Chen, C. & Yang, J. From genetic associations to genes: methods, applications, and challenges. *Trends Genet.* **40**, 642–667 (2024).
83. Allesøe, R. L. et al. Discovery of drug–omics associations in type 2 diabetes with generative deep-learning models. *Nat. Biotechnol.* **41**, 399–408 (2023).
84. Giambartolomei, C. et al. Bayesian test for colocalisation between pairs of genetic association studies using summary statistics. *PLoS Genet.* **10**, e1004383 (2014).
85. Donohue, M. C. et al. Natural cubic splines for the analysis of Alzheimer's clinical trials. *Pharm. Stat.* **22**, 508–519 (2023).
86. Wen, J. MEDICINE: AI-derived GWAS summary statistics sharing portal. https://doi.org/10.7303/SYN64923248 (2025).
87. Wen, J. & Abdulkadir, A. Machine Learning in NeuroImaging (MLNI), version 0.1.4. https://github.com/anbai106/mlni (2025).
88. Wen, J. MRIBAG_Figures_Scripts_Dataset. https://doi.org/10.5281/zenodo.15857816 (2025).

Acknowledgements

This work is supported by the National Institute on Aging (NIA), National Institutes of Health (NIH) (RF1AG092412). The MULTI Consortium (J.W.; UKBB application number: 647044) aims to integrate multi-organ imaging and multi-omics data to advance understanding of human aging and disease mechanisms. This study used the UKBB resource under application numbers 647044 (J.W.; NIA grant number: RF1AG092412) and 35148 (D.C.; NIA grant number: RF1 AG054409). We would like to express our sincere gratitude to the UKBB team for their invaluable contribution to advancing clinical research in our field (https://www.ukbiobank.ac.uk/). We would like to acknowledge the participants and investigators of the FinnGen study and the PGC, and we thank FinnGen (https://www.finngen.fi/en) and PGC (https://pgc.unc.edu/) for their generosity in sharing the GWAS summary statistics with the scientific community. We thank the BLSA participants and staff for their participation and continued dedication. The BLSA protocol was approved by the institutional review board of the National Institute of Environmental Health Science, NIH (03AG0325). The A4 study was a secondary prevention trial in preclinical Alzheimer's disease, aiming to slow cognitive decline associated with brain amyloid accumulation in clinically normal older individuals. The A4 study was funded by a public–private–philanthropic partnership, including funding from the NIA/NIH, Eli Lilly and Company, the Alzheimer's Association, the Accelerating Medicines Partnership, the GHR Foundation, an anonymous foundation and additional private donors, with in-kind support from Avid Radiopharmaceuticals, Cogstate, Albert Einstein College of Medicine and the Foundation for Neurologic Diseases. The companion observational Longitudinal Evaluation of Amyloid Risk and Neurodegeneration (LEARN) study was funded by the Alzheimer's Association and the GHR Foundation. The A4 and LEARN studies were led by R. Sperling at Brigham and Women's Hospital, Harvard Medical School, and by P. Aisen at the Alzheimer's Therapeutic Research Institute at the University of Southern California. The A4 and LEARN studies were coordinated by the Alzheimer's Therapeutic Research Institute at the University of Southern California, and the data are made available under the auspices of the Alzheimer's Clinical Trial Consortium through the Global Research & Imaging Platform. The complete A4 study team list is available at https://www.actcinfo.org/a4-study-team-lists/. We would like to acknowledge the dedication of the study participants and their partners who made the A4 and LEARN studies possible. D.R.M., W.A.K. and F.A. were supported by the Intramural Research Program of the NIA, NIH. We acknowledge the leadership of the Brain Imaging Genetics workgroup, led by T. Evans, N. Vilor-Tejedor and J.W., within the International Society to Advance Alzheimer's Research and Treatment (ISTAART) community, for advocating brain imaging genetics in Alzheimer's disease and aging research.

Author contributions

J.W. has full access to all the study data and is responsible for integrity and accuracy. Study concept and design: J.W. Drafting of the manuscript: J.W. Critical revision of the manuscript for important intellectual content: all authors. Statistical analysis: J.W. ran all main analyses. C.H. and S.Z. helped with the development of the MEDICINE portal. D.M. ran the ProWAS for the SomaScan proteomics data in the BLSA sample. Detailed contributions: J.W. leads the MULTI Consortium. D.C. leads the iSTAGING Consortium. W.A.K. represents the BLSA study. R.S.M. and A.P. represent the A4 study. B.W. derived the original heart IDPs. Z.A. and T.Y. derived the nine original PhenoBAGs, and we only used the derived GWAS summary statistics in this study.

Competing interests

B.W.D. is an inventor of the DunedinPACE epigenetic clock, which is licensed by Duke University and the University of Otago to TruDiagnostic, from which he receives royalties (DunedinPACE is freely available to researchers). All other authors declare no competing interests.

Additional information

Extended data is available for this paper at https://doi.org/10.1038/s41591-025-03999-8.

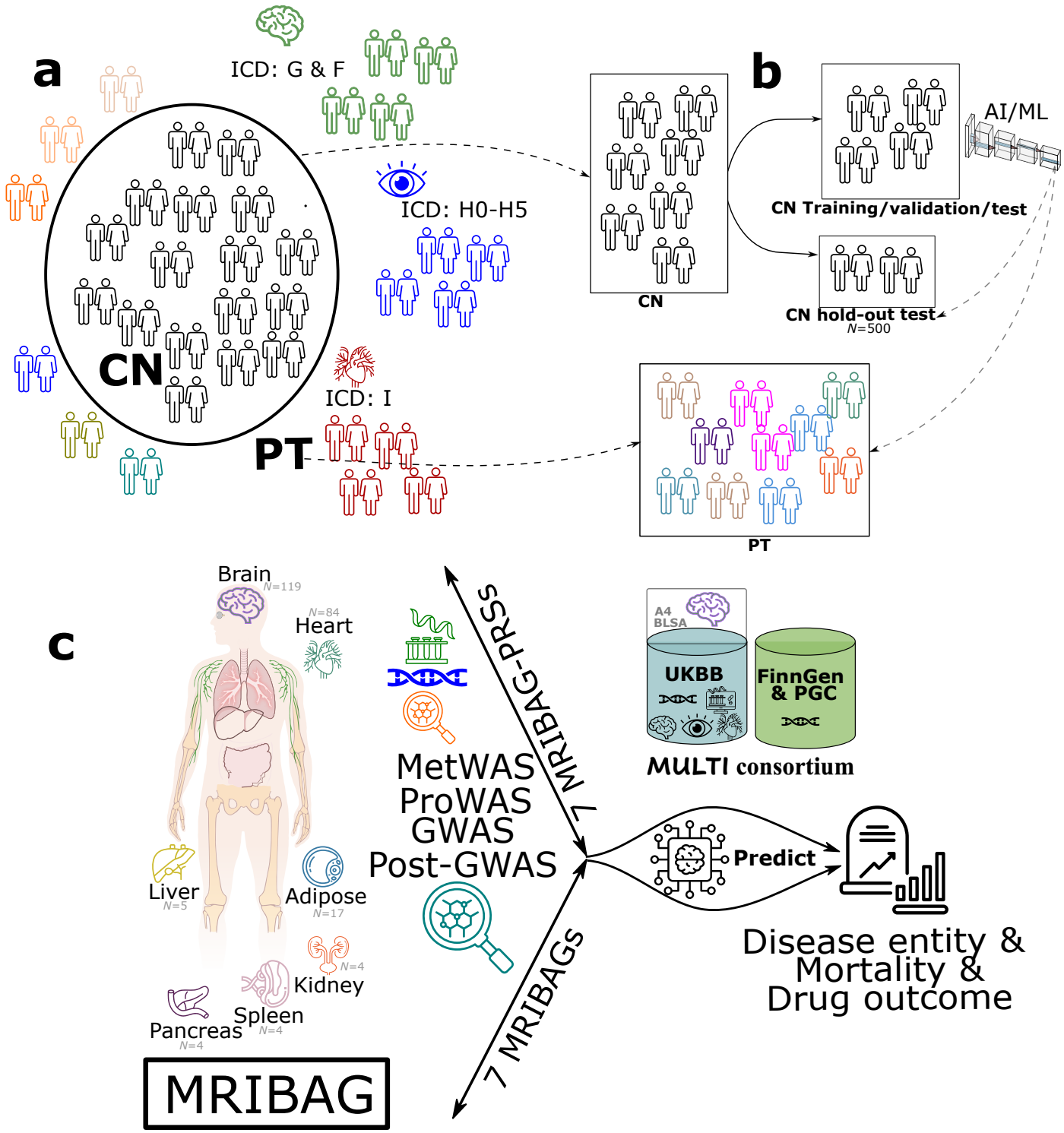
Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41591-025-03999-8.

Correspondence and requests for materials should be addressed to Junhao Wen.

Peer review information *Nature Medicine* thanks David Furman and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editor: Anna Maria Ranzoni in collaboration with the *Nature Medicine* team.

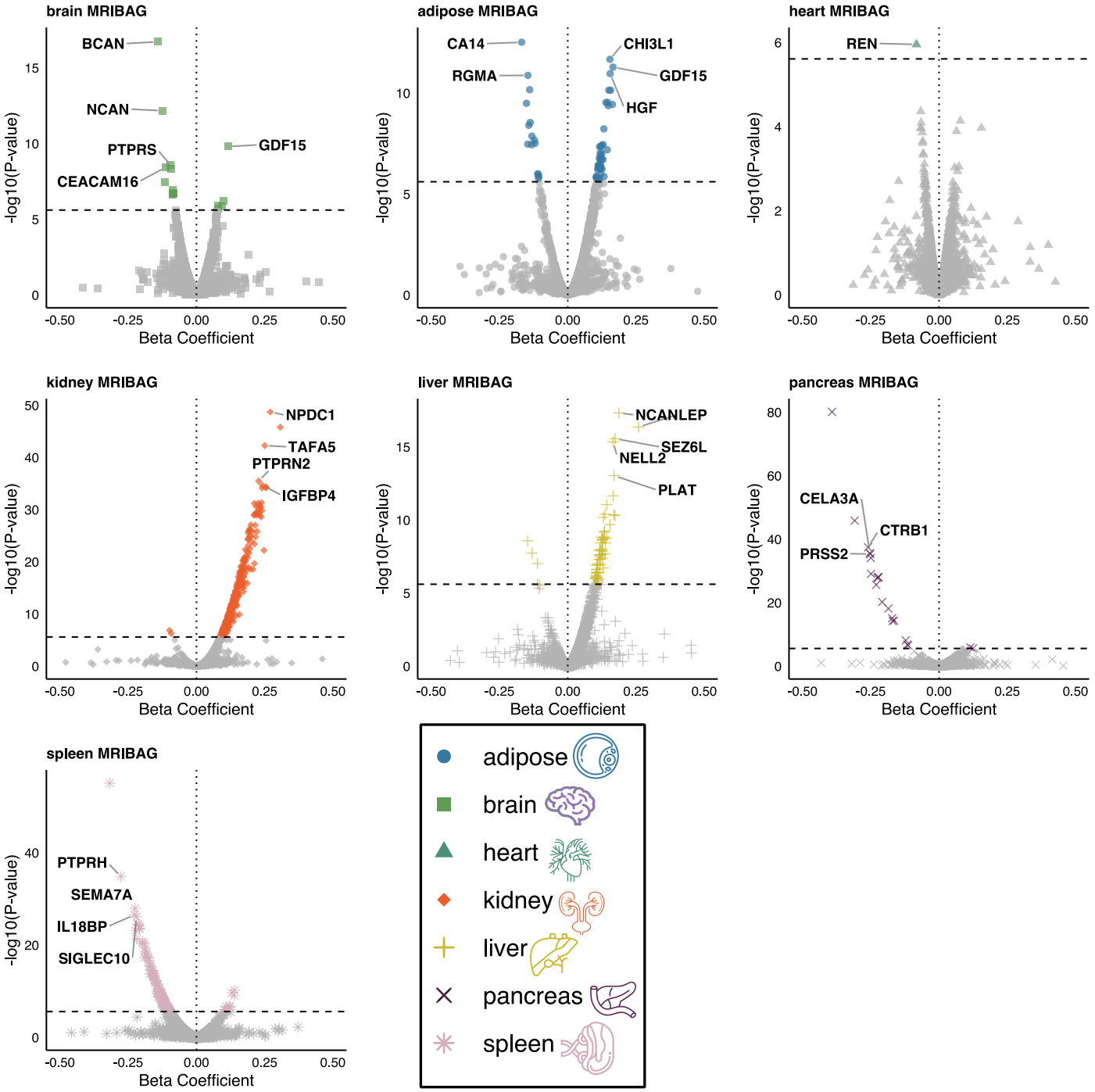
Reprints and permissions information is available at www.nature.com/reprints.

https://doi.org/10.1038/s41591-025-03999-8



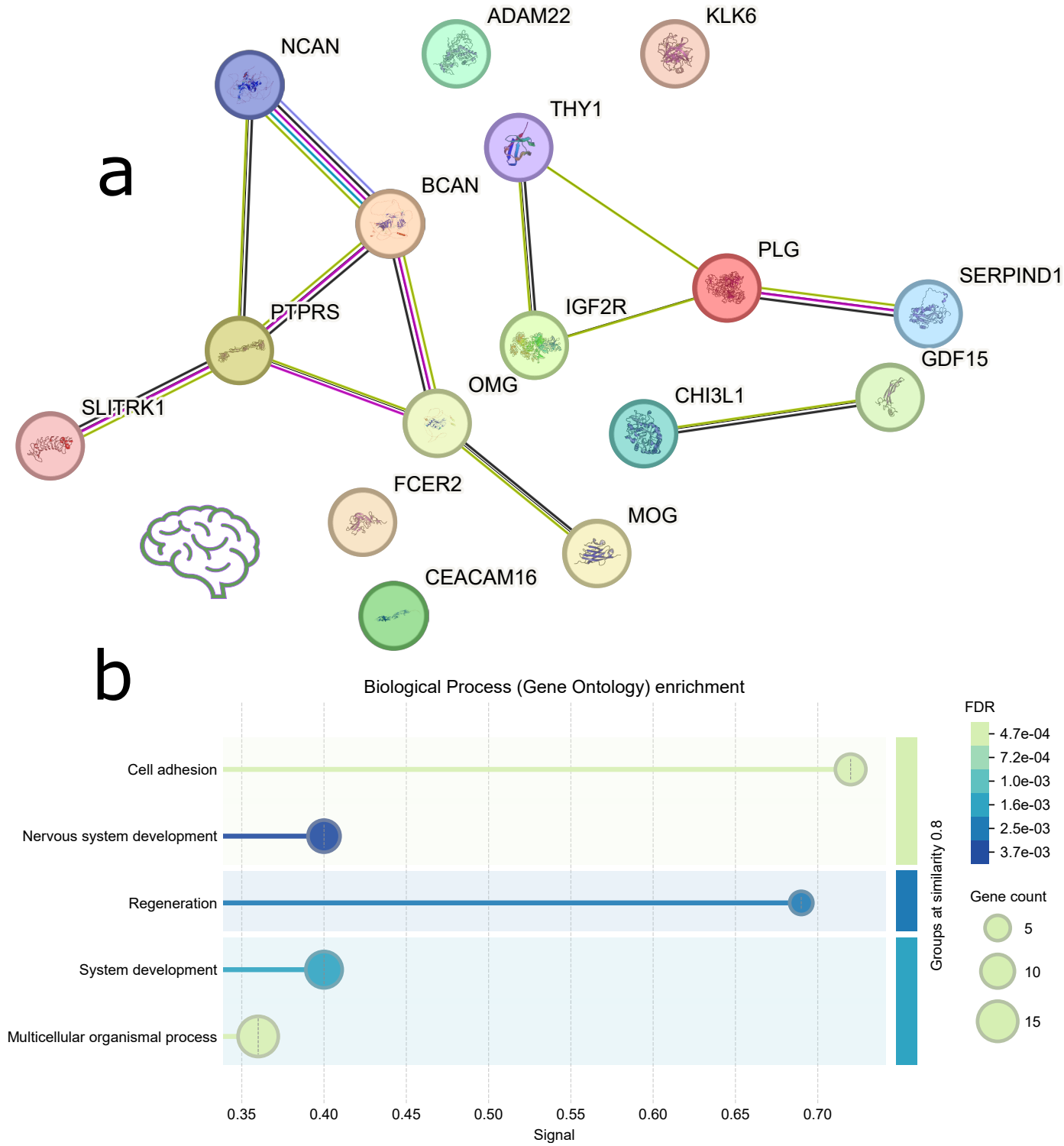
Extended Data Fig. 1 | Schematic figure of the population definition and workflow used to derive MRIBAG. a) For each organ-specific MRIBAG, we split the UK Biobank populations into healthy controls (CN, without any pathologies) and patients (PT) based on ICD-10 codes and their clinical history. **b**) To develop the 7 MRIBAGs, we trained two AI/ML models exclusively on the CN training/validation/test population using a nested cross-validation procedure to identify the optimal model. Hold-out datasets included the CN within-distribution, hold-out test dataset ($N = 500$), and the PT population. **c**) The study's analytical workflow encompassed deriving the 7 MRIBAGs and their PRSs, performing ProWAS, MetWAS, GWAS, and post-GWAS analyses, and assessing the predictive performance of the MRIBAG and MRIBAG-PRSs for single disease entities, mortality, and drug outcomes. Data from the BLSA and A4 studies were used to generate the brain MRIBAG in independent cohorts.

https://doi.org/10.1038/s41591-025-03999-8



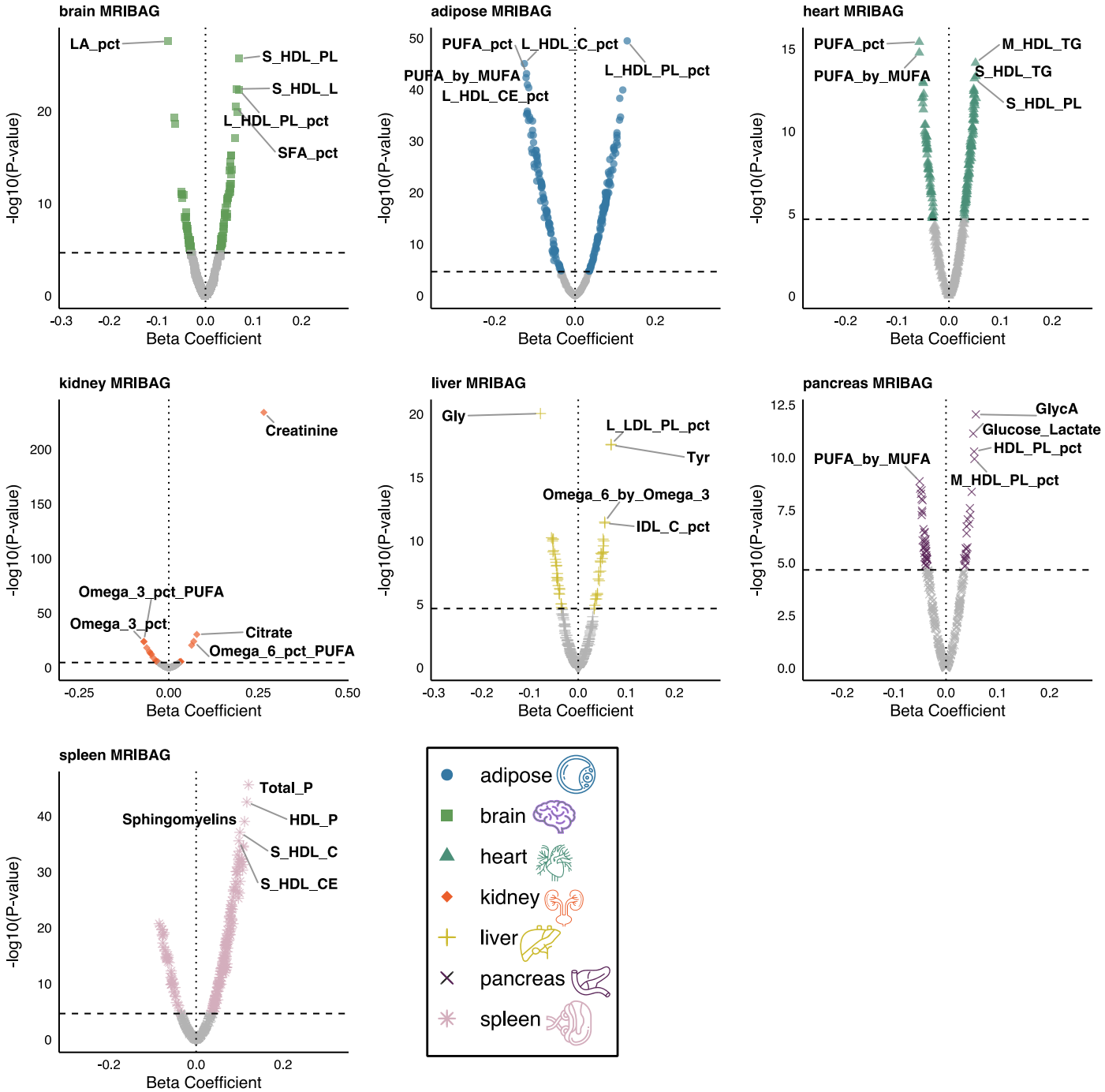
Extended Data Fig. 2 | Individual volcano plots for the 7 MRIBAGs in ProWAS. Proteome-wide association studies (ProWAS) between the 7 MRIBAGs and 2923 plasma proteins via a linear regression model, accounting for a full set of covariates. Representative significant (P-value < 0.05/2923/7, indicated by the dotted horizontal line) MRIBAG-protein associations are annotated by the respective gene symbol. We used the standardized β value as an effect size measurement.

https://doi.org/10.1038/s41591-025-03999-8



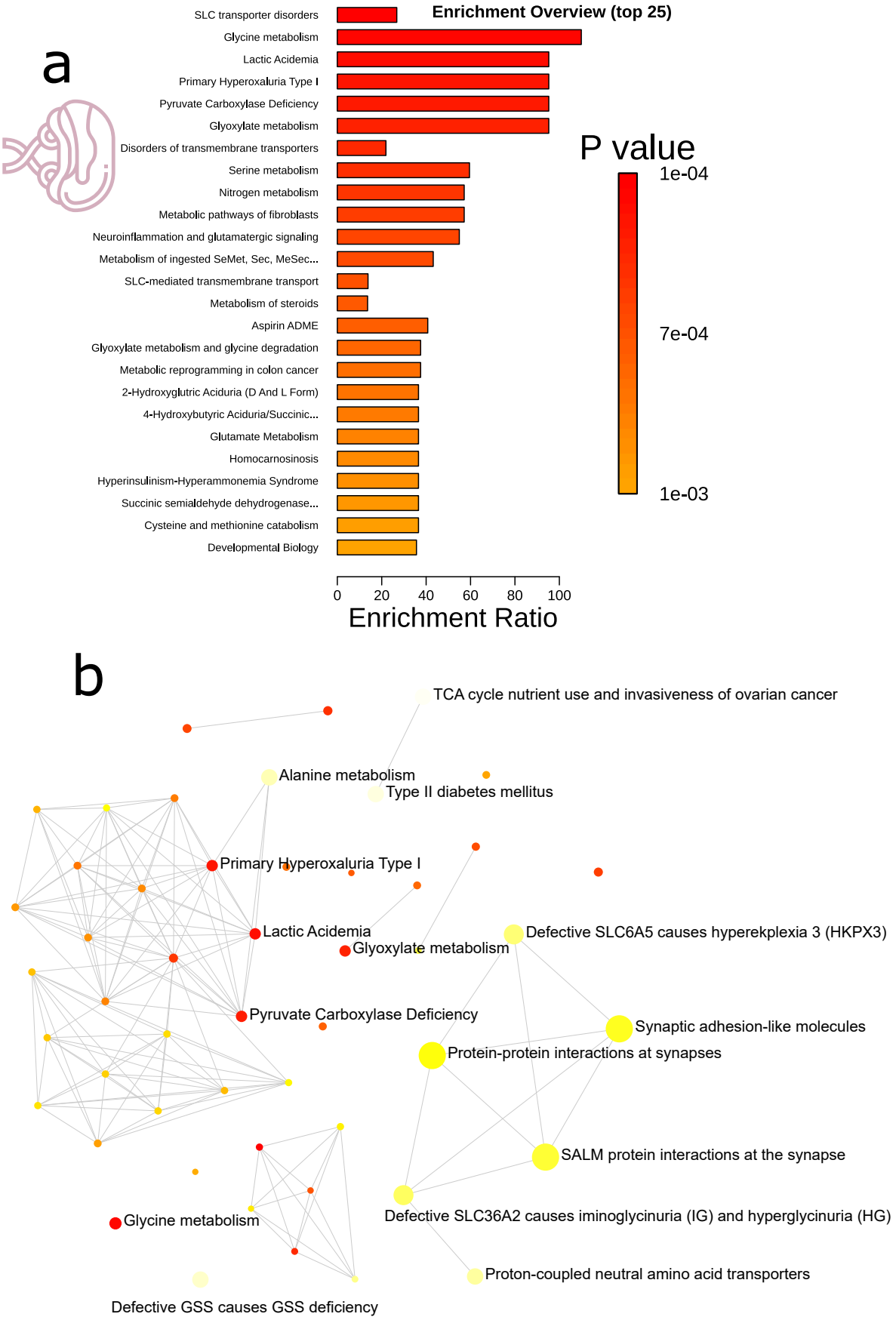
Extended Data Fig. 3 | Functional protein-set enrichment analysis (PSEA) and protein-protein interaction (PPI) network for the brain MRIBAG-linked proteins. STRING conducts protein-protein interaction (PPI) analysis by combining various sources of interaction evidence, such as experimental data, computational predictions, and text mining. In addition to interactions, STRING can carry out functional enrichment analysis to identify Gene Ontology (GO) terms (for example, biological processes, molecular functions), KEGG pathways (for example, metabolic and signaling pathways), and protein domains (for example, shared structural motifs). **a**) PPI network of the brain MRIBAG-related proteins. **b**) Functional PSEA of the brain MRIBAG-related proteins. Results and figures are generated using STRING v12.0 (https://string-db.org/).

https://doi.org/10.1038/s41591-025-03999-8



Extended Data Fig. 4 | Individual volcano plots for the 7 MRIBAGs in MetWAS. Metabolome-wide association studies (MetWAS) between the 7 MRIBAGs and 327 plasma metabolites via a linear regression model, accounting for a full set of covariates. Representative significant (P-value < 0.05/327/7, indicated by the dotted horizontal line) MRIBAG-metabolite associations are annotated. We used the standardized β value as an effect size measurement.

https://doi.org/10.1038/s41591-025-03999-8

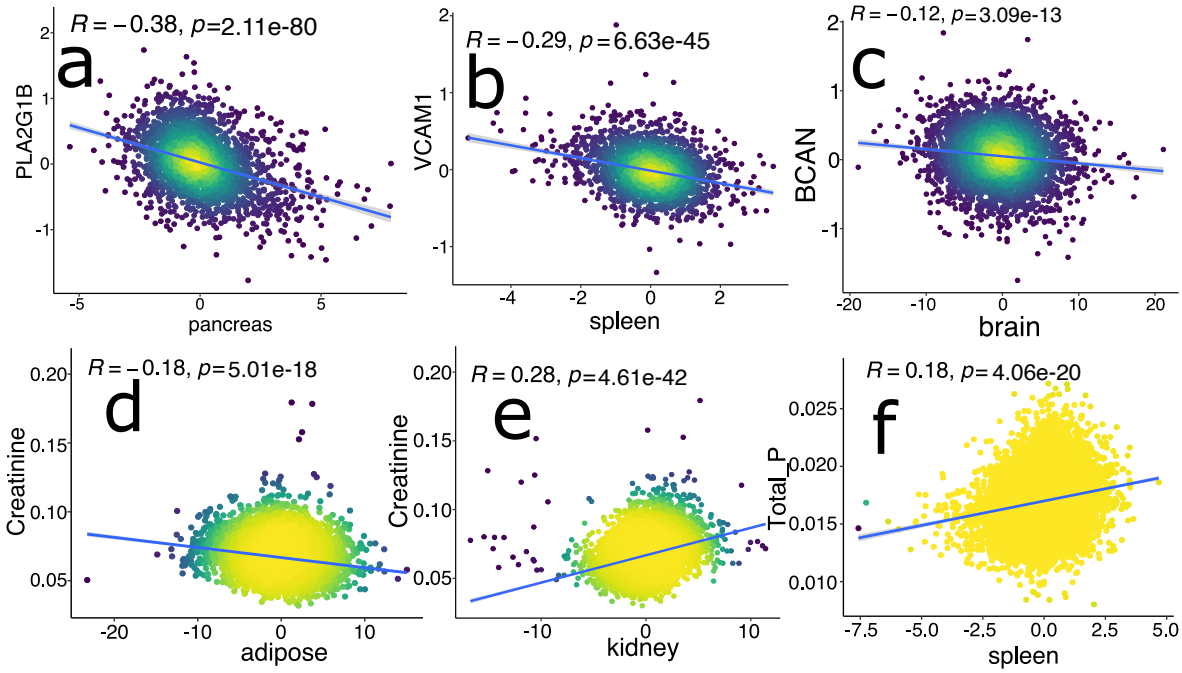


Extended Data Fig. 5 | See next page for caption.

https://doi.org/10.1038/s41591-025-03999-8

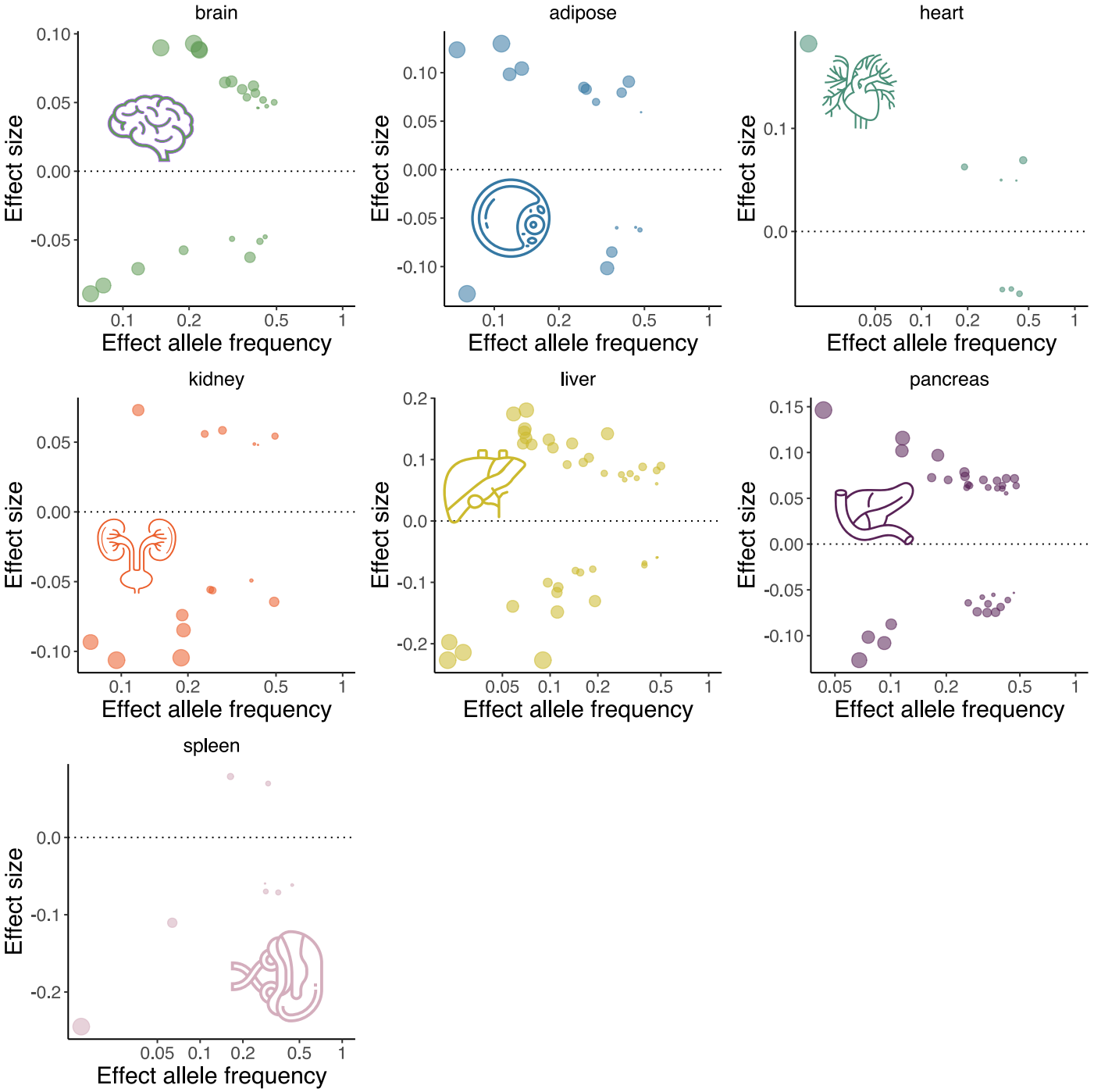
Extended Data Fig. 5 | Functional metabolite-set enrichment analysis (MSEA) and metabolite-metabolite interaction (MMI) network for the 6 spleen MRIBAG-linked small-molecular metabolites. a) Top 25 MSEA results for the spleen MRIBAG-related small molecule metabolites. **b**) The MMI network results for these metabolite sets are displayed as follows: each node represents a metabolite set, with color indicating the enrichment P-value and size reflecting the fold enrichment ratio. Two metabolite sets are connected by an edge if the proportion of shared metabolites between them exceeds 25% of the total combined metabolites. Results and figures are generated using MetaboAnalyst v6.0 (https://www.metaboanalyst.ca/). We used 3694 metabolite and lipid pathways from RaMP-DB (which integrates KEGG via HMDB, Reactome, and WikiPathways, curated via MetaboAnalyst) as the background metabolite dataset. Only metabolite sets with at least two entries were included.

https://doi.org/10.1038/s41591-025-03999-8



Extended Data Fig. 6 | Example associations for the ProWAS and MetWAS results. a) ProWAS results for the PLA2G1B protein and the pancreas MRIBAG. **b**) ProWAS results for the VCAM1 protein and the spleen MRIBAG. **c**) ProWAS results for the BCAN protein and the brain MRIBAG. **d**) MetWAS results for the creatinine metabolite and the adipose MRIBAG. **e**) MetWAS results for the creatinine metabolite and the kidney MRIBAG. **f**) MetWAS results for total concentration of phospholipids (Total_P) and the spleen MRIBAG. A linear regression model was fitted, and the resulting Pearson's *r* and P-value are reported.

https://doi.org/10.1038/s41591-025-03999-8



Extended Data Fig. 7 | Trumpet plots of the effect allele frequency vs. the *β* coefficient of the 7 MRIBAG GWASs. The trumpet plots display the inverse relationship between the alternative (effect) allele frequency and the effect size (*β* coefficient) for the 7 MRIBAGs. We present the independent, significant SNPs defined in FUMA. The dot size corresponds to the effect size, while the transparency of the dot is proportional to its statistical significance.

nature portfolio

Corresponding author(s): Junhao Wen

Last updated by author(s): Aug 13, 2025

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our Editorial Policies and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☐	☒	A description of all covariates tested
☐	☒	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☐	☒	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☐	☒	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	No software was used for data collection
Data analysis	• MLNI (v0.1.2 and v0.1.5): https://github.com/anbai106/mlni, MRIBAG generation • FUMA (v1.5.0): https://fuma.ctglab.nl/, Gene mapping, genomic locus annotation; • GCTA (v1.94.1): https://yanglab.westlake.edu.cn/software/gcta/#Overview, fastGWA; • LDSC (git version aa33296): https://github.com/bulik/ldsc, genetic correlation • TwoSampleMR (v0.5.6): https://mrcieu.github.io/TwoSampleMR/index.html, Mendelian randomization; • PRScs (release date: Aug 10, 2023): https://github.com/getian107/PRScs, PRS calculation; • Lifelines (v0.27.8): https://lifelines.readthedocs.io/en/latest/, Survival analysis; • String (V12.0): https://string-db.org/, protein analyses. • coloc (v5): https://github.com/chr1swallace/coloc; Bayesian colocalization.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our policy

The GWAS summary statistics and pre-trained AI models from this study are publicly accessible via the MEDICINE Knowledge Portal (https://labs-laboratory.com/medicine/) and Synapse (https://www.synapse.org/Synapse:syn64923248/wiki/63099290). Our study used data generated by the Human Protein Atlas (https://www.proteinatlas.org). GWAS summary data for the DEs were downloaded from the official websites of FinnGen (R9: https://www.finngen.fi/en/access_results) and PGC (https://pgc.unc.edu/for-researchers/download-results/). Individual data from UKBB can be requested with proper registration at https://www.ukbiobank.ac.uk/. The A4 study data can be requested at: https://www.a4studydata.org/. The NIA-supported BLSA data can be requested at: https://www.nia.nih.gov/research/labs/blsa. All main figures, Extended Data Figures, the code used to derive the 7 MRIBAGs, and the Supplementary Files are available on Zenodo: https://zenodo.org/records/1585781791.

Research involving human participants, their data, or biological material

Policy information about studies with human participants or human data. See also policy information about sex, gender (identity/presentation), and sexual orientation and race, ethnicity and racism.

Reporting on sex and gender	We included sex as a covariate in all statistical models, and provided sex-specific analyses
Reporting on race, ethnicity, or other socially relevant groupings	Whenever appropriate, we report ethnic groups. All GWAS were conducted in people of European ancestry.
Population characteristics	Study population is detailed in Method 1 and Supplementary eTable 1
Recruitment	All individual-level data were initially recruited by respective studies, underwent ethnic scrutinization. In addition, we performed standard quality check for all imaging, genetics, and proteomics to ensure the data quality.
Ethics oversight	Respective study committee oversees the echics considerations of each study. The MULTI consortium is approved by the Institutional Review Board at Columbia University (AAAV6751).

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	No statistical methods were used to determine the final sample size used. All data from all participants were subjected to a quality check procedure to exclude unqualified data. For example, we have an automated imaging quality check pipeline to rule out low quality brain imaging scans; we also detailed the quality check for genetic data to rule out low quality participants. This resulted in a sufficiently large sample size for most of these model training and genetic analyses (>20k).
Data exclusions	All imaging, genetic, and proteomics were quality-checked to exclude data that did not qualify for this study.
Replication	We used external data, A4 and BLSA, to test whether the brain MRIBAG can be replicated.
Randomization	No randomization was performed to allocate participants into different groups. We included a wide range of covariates that have been commonly used in previous large-scale GWAS.
Blinding	Blinding was not applicable to this study because it involved secondary analyses of pre-existing, de-identified population datasets (e.g., UK Biobank), where group allocation was not manipulated by the investigators.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology and archaeology
☒	☐ Animals and other organisms
☒	☐ Clinical data
☒	☐ Dual use research of concern
☒	☐ Plants

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☐	☒ MRI-based neuroimaging

Magnetic resonance imaging

Experimental design

Design type	NA
Design specifications	NA
Behavioral performance measures	NA

Acquisition

Imaging type(s)	Structural MRI (T1-weighted MRI)
Field strength	3.5T
Sequence & imaging parameters	This can be found at the UK Biobank website for detail.
Area of acquisition	Brain

Diffusion MRI ☐ Used ☒ Not used

Preprocessing

Preprocessing software	All software used in this study is detailed in the Code Availability section.
Normalization	For brain MRI, we used a population-level template for the brain MRI for registration purposes.
Normalization template	For brain MRI, we used a population-level template for the brain MRI for registration purposes.
Noise and artifact removal	All data underwent standard quality check pipelines, such as for brain MRI for remove noise and artifact signals.
Volume censoring	NA

Statistical modeling & inference

Model type and settings	Multivariate analyses (machine learning and deep learning) for deriving the 7 MRIBAGs and univariate for GWAS (linear regression, and mixed effect model)
Effect(s) tested	Beta coefficient are reported for GWAS, Odds ratio is reported for Mendelian randomization.

Specify type of analysis: ☐ Whole brain ☒ ROI-based ☐ Both

Anatomical location(s): Using the MUSE atlas ROIs.

Statistic type for inference (See Eklund et al. 2016)	*Specify voxel-wise or cluster-wise and report all relevant parameters for cluster-wise methods.*
Correction	Bonferroni correction; we detailed all this in each result and method section.

Models & analysis

n/a	Involved in the study
☒	☐ Functional and/or effective connectivity
☒	☐ Graph analysis
☐	☒ Multivariate modeling or predictive analysis

Multivariate modeling and predictive analysis: Age prediction model via Lasso regression, linear SVR, Elastic Net, and neural network.