bioRxiv preprint doi: https://doi.org/10.1101/2021.10.22.465437; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

General dimensions of human brain morphometry inferred from genome-wide association data

Anna E. Fürtjes[1], Ryan Arathimos[1,2], Jonathan R. I. Coleman[1,2], James H. Cole[3,4,5], Simon R. Cox[6,7], Ian J. Deary[6,7], Javier de la Fuente[8,9], James W. Madole[8], Elliot M. Tucker-Drob[8,9], Stuart J. Ritchie[1]

The names of the co-authors are listed in alphabetical order and grouped by affiliation.

[1] Social, Genetic and Developmental Psychiatry (SGDP) Centre, Institute of Psychiatry, Psychology & Neuroscience, King's College London, SE5 8AF, UK

[2] National Institutes for Health Research Maudsley Biomedical Research Centre, South London and Maudsley NHS Trust, London, SE5 8AF, UK

[3] Department of Neuroimaging, Institute of Psychiatry, Psychology & Neuroscience, King's College London, London, SE5 8AF, UK

[4] Centre for Medical Image Computing, Department of Computer Science, University College London, London, WC1V 6LJ, UK

[5] Dementia Research Centre, Institute of Neurology, University College London, London, WC1N 3BG, UK

[6] Department of Psychology, The University of Edinburgh, Edinburgh, EH8 9JZ, UK

[7] Lothian Birth Cohorts, University of Edinburgh, Edinburgh, EH8 9JZ, UK

[8] Department of Psychology, University of Texas at Austin, Austin, TX 78712-1043, USA

[9] Population Research Center and Center on Aging and Population Sciences, University of Texas at Austin, Austin, TX 78712-1043, USA

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.22.465437; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Correspondence:

Anna E. Fürtjes

anna.furtjes@kcl.ac.uk

.

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.22.465437; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Summary

The human brain is organised into networks of interconnected regions that have highly correlated volumes. In this study, we aim to triangulate insights into brain organisation and its relationship with cognitive ability and ageing, by analysing genetic data. We estimated general genetic dimensions of human brain morphometry within the whole brain, and nine predefined canonical brain networks of interest. We did so based on principal components analysis (PCA) of genetic correlations among grey-matter volumes for 83 cortical and subcortical regions ($N_{participants}$ = 36,778). We found that the corresponding general dimension of brain morphometry accounts for 40% of the genetic variance in the individual brain regions across the whole brain, and 47-65% within each network of interest. This genetic correlation structure of regional brain morphometry closely resembled the phenotypic correlation structure of the same regions. Applying a novel multivariate methodology for calculating SNP effects for each of the general dimensions identified, we find that general genetic dimensions of morphometry within networks are negatively associated with brain age (r_g = -0.34) and profiles characteristic of age-related neurodegeneration, as indexed by cross-sectional age-volume correlations (r = -0.27). The same genetic dimensions were positively associated with a genetic general factor of cognitive ability (r_g = 0.17-0.21 for different networks). We have provided a statistical framework to index general dimensions of shared genetic morphometry that vary between brain networks, and report evidence for a shared biological basis underlying brain morphometry, cognitive ability, and brain ageing, that are underpinned by general genetic factors.

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.22.465437; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Introduction

Cheverud's Conjecture holds that phenotypic correlations correspond closely to genetic correlations, meaning that they can serve as proxies for one another. Cheverud originally speculated that "If genetically and environmentally based phenotypic variations are produced by similar disruptions of developmental pathways, genetic and environmental correlations should be similar." [1]. This relationship is assumed to apply across traits and species, including humans. For example, a 2018 study of a wide range of morphometric human traits (for example, height and body mass index) in the UK Biobank cohort reported a strong correspondence between phenotypic and genetic correlations, with slightly larger average genetic than phenotypic correlations [2]. Here we test whether a similar level of close correspondence between genetic and phenotypic correlations applies to magnetic resonance imaging (MRI)-derived indices of human brain morphometry.

The human brain is organised into complex networks of interconnected regions containing a variety of well-studied macroscopic subnetworks [3-5]. The specific subnetworks considered in this study have been characterised on the basis of synchronised (i.e., correlated) regional activity in functional MRI data [6], in addition to converging evidence from other modalities (i.e., structural MRI and lesion-based mapping [7-9]). These characterisations allow for a whole-brain approach including cortical and subcortical brain volumes (Figure 1). Regions within networks have been found to share functional specialisation [4] and to be internally connected through white matter tracts [10-12]. Among the most reported, 'canonical', networks are the central executive, default mode, salience, and multiple demand networks.

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.22.465437; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

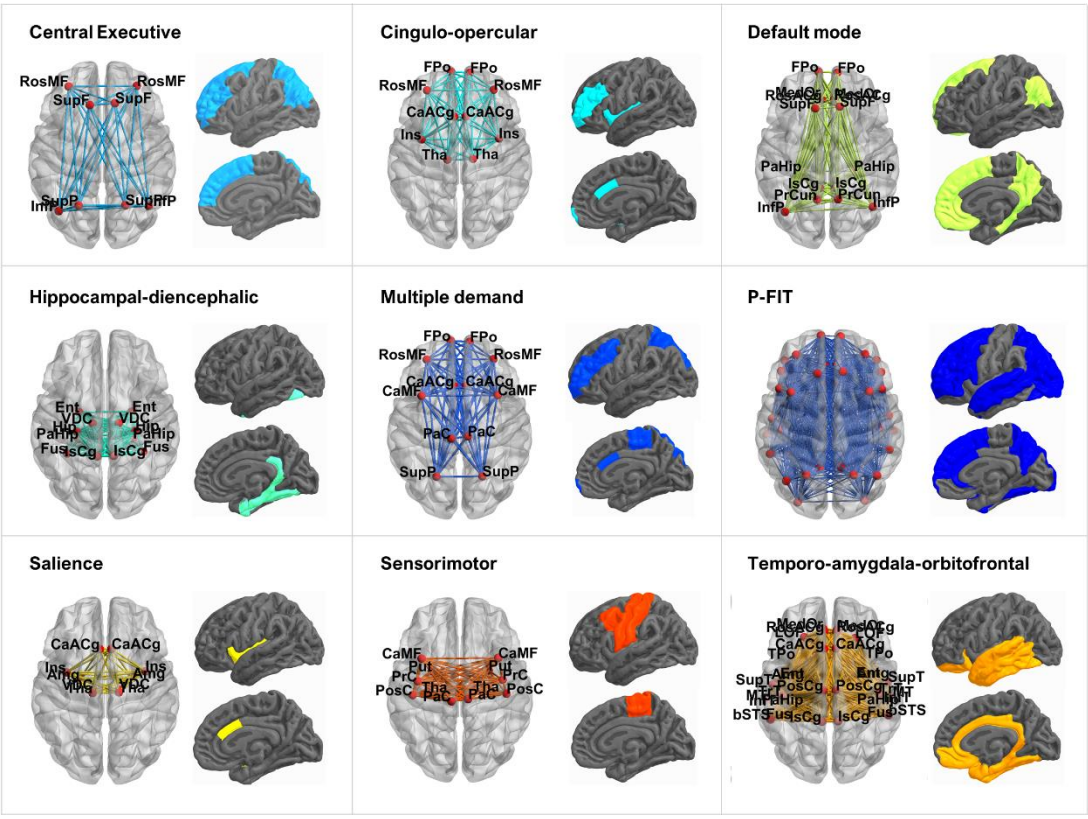

Figure 1. Nine canonical brain subnetworks. The networks were visualized with the BrainNet Viewer (http://www.nitrc.org/projects/bnv/) [13]. Regions of interest were visualised using scripts by Dr. Colin Buchanan (University of Edinburgh). Included brain regions and their abbreviations are listed in STable 2.

Brain networks are theorised to integrate information across the brain and, collectively, to give rise to cognitive functions. The central executive network is thought to underpin higher-level cognitive functions, including attention and working memory processes [9, 14]; whereas the default mode network is associated with internally directed and abstract thought [15]. The salience network is thought to detect salient sensory cues [16], helping to integrate executive and default functions [14, 17]. Mental processes that organise multiple cognitive requirements into a series of successive cognitive tasks are thought to be associated with the multiple demand network [18].

Because brain volumes and networks are highly heritable [19-21], under Cheverud's Conjecture we predict a close correspondence between phenotypic network organisation and

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.22.465437; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

its genetic equivalent. Here, we aim to compare phenotypic and genetic brain organisation, by focusing on *structural* characteristics measured using MRI. Structural neuroimaging is independent of mental processes during MRI scanning, and substantially outperforms functional imaging in terms of test-retest reliability [22]. Structural neuroimaging is reliably correlated with cognitive performance [23], and it reflects atrophy; an important indicator of health outcomes [24]. We focus on specific network characterisations prominent in the literature, using common, but not indisputable, definitions of the exact regions comprising them [5].

Overall, this study aims to triangulate previous studies describing brain organisation and its relationships with ageing and cognitive ability, by analysing genetic data. According to Cheverud's Conjecture [1], morphometric correlations across phenotypic brain networks should mirror the structure of genetic correlations within the same networks. A dissimilar organisation of phenotypic and genetic brain architecture would contradict the neurobiological validity of canonical brain networks (a similar organisation would be consistent with a measurable genetic foundation of brain networks). In this pre-registered study (https://osf.io/7n4qj), we used genome-wide association data to identify broad genetic dimensions underlying morphometric measures across human brain regions. We do so within canonical brain networks that have been implicated in cognitive functioning and cognitive ageing [6]. We compared these genetic dimensions to phenotypic dimensions within the same networks identified in a large recent study (N = 8,185) which examined phenotypic correlations between structural brain networks and cognitive abilities [6]. Here, we estimate genetic associations between these general dimensions of human brain morphometry and both cognitive abilities and 'brain age' [25].

More specifically, our study modelled shared genetic variance between grey-matter volumes across the whole brain, as well as nine canonical subnetworks [e.g., 3, 4]. We

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.22.465437; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

calculated genome-wide association study (GWAS) summary statistics (N = 36,778) for 83 regional volumes, and built genetic correlation matrices, indicating genetic overlap between brain volumes, through multivariate linkage disequilibrium score regression (LDSC) [26]. Using principal components analysis (PCA), we modelled the first whole-brain and network-specific principal components (PCs) from genetic and phenotypic correlations between volumes matrices, to index general dimensions of shared morphometry between brain volumes. We described those general dimensions extracted on a genetic level of analysis, and contrasted them with the same general dimensions extracted on a phenotypic level of analysis. By testing whether the relative magnitudes of brain regions' loadings on genetic dimensions of shared morphometry correspond to their sensitivity to advancing age, we can evaluate whether the pathways giving rise to genetic variation in brain morphometry correspond with those on which ageing-related neurodegeneration occurs. We present a novel statistical procedure to extract genetic PCs in univariate per-SNP (single nucleotide polymorphism) summary statistics, to represent genetic network structure. They can be used in a genomic structural equation framework [27] to understand whether there is genetic covariation between dimensions of shared morphometry within networks and cognitive abilities, brain age, and cross-sectional patterning of age-related neurodegeneration [6].

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.22.465437; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Results

Results presented in the following sections are structured as displayed in Figure 2. (1) We calculated 83 GWAS summary statistics that served as input data, and (2) calculated genetic correlation matrices indicating genetic overlap between brain volumes using genomic structural equation modelling (GenomicSEM) [27] . (3) We performed PCA to extract general dimensions of shared genetic morphometry between regions and described their characteristics. We compared genetic correlation structures with phenotypic correlation structures between the 83 regional volumes and tested whether the relative ordering of phenotypic and genetic PC loadings correlated with indices of age sensitivity. Finally, (4) we summarised shared genetic morphometric variance within brain networks in sets of univariate summary statistics (i.e., genetic PCs underlying multiple brain volumes; see Methods). These univariate summary statistics can be viewed as a summary-based method of computing GWAS summary statistics that would be obtained from a GWAS on individuals' scores on the underlying genetic PCs (see Methods; Figure 2). We used this method to create summary statistics of general dimensions of shared morphometry across the whole brain and nine canonical networks, and then calculated genetic correlations between brain networks and both cognitive ability and brain age.

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.22.465437; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

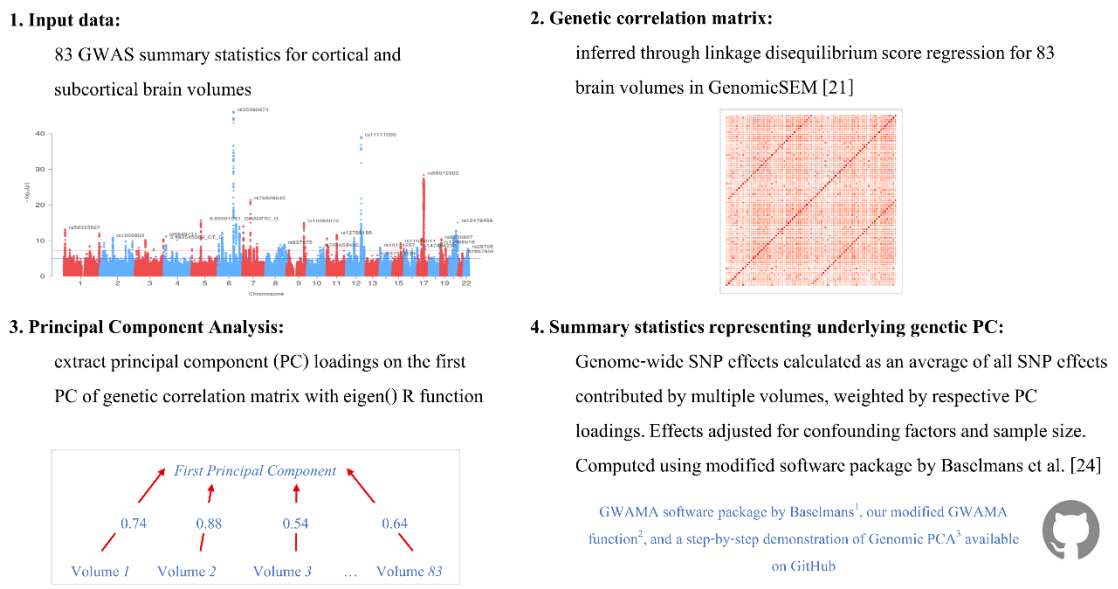

Figure 2. Four-step procedure to obtain statistical representation of genetic brain network structure from GWAS summary statistics. (1) GWAS summary statistics for 83 grey-matter volumes in UK Biobank from European ancestry were used as input data (N = 36,778). They were calculated as described in Methods and are publicly available. (2) LDSC was used to infer genetic correlations between 83 brain volumes. (3) Genetic correlations are analysed using PCA to derive PC loadings on the first PC, representing an underlying dimension of shared morphometry. (4) We developed a method to derive univariate summary statistics for a genetic PC of multiple GWAS phenotypes (derived from samples of unknown degrees of overlap). We have validated this novel approach in an independent set of GWAS summary statistics [28]. All software we used is available on https://github.com/. [1] The software by Baselmans et al. [29], containing the GWAMA function is available at https://github.com/baselmans/multivariate_GWAMA/. [2] Our modified version of the GWAMA function is at https://github.com/AnnaFurtjes/Genetic_networks_project/blob/main/my_GWAMA_26032020.R and [3] a step-by-step demonstration of genomic PCA is at https://annafurtjes.github.io/genomicPCA/.

Genetic correlations between brain-wide volumes recapitulated phenotypic correlations

On a phenotypic level of analysis, correlations between the 83 brain volumes were obtained using Pearson's correlations from volumetric phenotypes that were residualised for age (mean = 63.3, range = 40.0-81.8 years) and sex (54% females). Phenotypic correlations

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.22.465437; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

were exclusively positive (Figure 3C). On a genetic level of analysis, we calculated genome-wide association study (GWAS) summary statistics to get per-SNP associations between participants' imputed genotype data (N = 36,778) and 83 cortical and subcortical grey-matter volumes (Figure 2.1). SNP-heritability estimates ranged between 7% (SE = 0.07) for the frontal poles and 42% (SE = 0.04) for the brain stem (mean = 0.23, SD = 0.07; Figure 3A).

These 83 GWAS summary statistics were used to infer genetic correlations between 83 volumes through LDSC [26] ($\frac{83(83-1)}{2}$ = 3403 between-region correlations; Figure 2.2). All bilateral regions were almost perfectly correlated with the corresponding contralateral region. Genetic correlations between regions ranged from r_g = -0.08 (SE = 0.09) between right frontal pole and left pallidum to r_g = 0.87 (SE = 0.08) between left middle temporal and left inferior temporal (distribution of genetic correlations in Figure 3B, SFigure 1). Standard errors of the genetic correlations ranged between 0.01 and 0.03 (mean = 0.014; SD = 0.002). Genetic correlations between regional volumes, corresponding to the canonical networks that are displayed in Figure 1, are provided in SFigures 2-10.

A linear regression was performed to assess the relationship between the 3403 phenotypic and the 3403 genetic correlations (Figure 3&5A). The association was positive, and large (r = .84; b = 0.60; SE = 0.007, $p < 2 \times 10^{-16}$, R^2 = 70%), indicating that the same regions, that had strongly correlated phenotypic volumes, were also genetically correlated. Phenotypic correlations were exclusively positive, as were 3,392 of 3,403 genetic correlations; the 11 (0.32%) negative genetic correlations were close to zero (smallest r_g -0.083).

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.22.465437; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

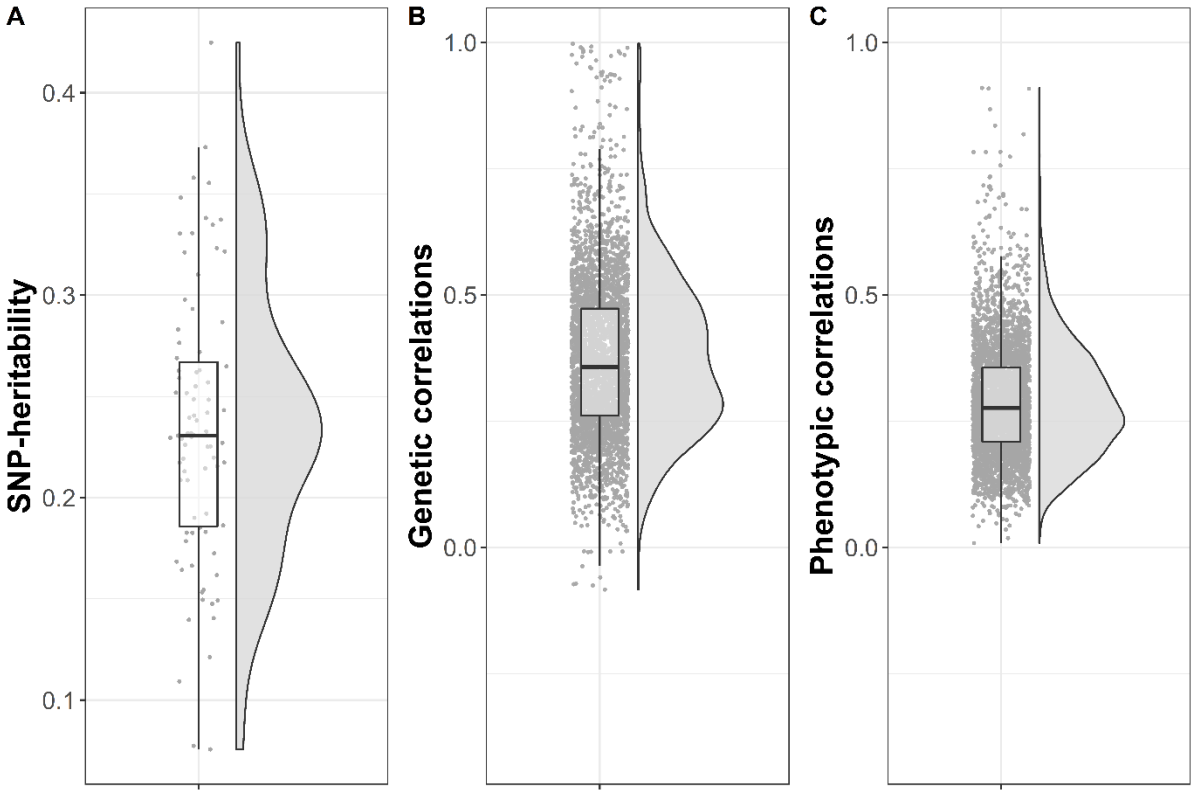

Figure 3. **(A)** Distribution of SNP-heritability estimates for 83 regional grey-matter volumes inferred through univariate LDSC. **(B)** Distribution of genetic correlations among 83 regional grey-matter volumes inferred through between-region LDSC. This figure only depicts between-region correlations but not the very high genetic inter-region correlations between regions and their homologous counterpart in the opposite hemisphere (excluding brain stem). **(C)** Distribution of phenotypic correlations among 83 regional grey-matter volumes inferred through Pearson's correlations. The raincloud plots were created based on code adapted from Allen et al. [30].

PCs of shared genetic variance across the whole-brain and canonical networks

Distributions of phenotypic PC loadings in Figure 4A (descriptive statistics for phenotypic shared morphometry in STable 3). On a genetic level of analysis, we extracted PCs from genetic correlation matrices. The first genetic whole-brain PC explained 40% of the genetic variance across 83 regional volumes - slightly larger than the 31% explained by the first phenotypic whole-brain PC. The second genetic whole-brain PC accounted for 6.7% of

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.22.465437; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

the total genetic variance; that is, 17% of the variance explained by the 1st genetic PC (SFigure 20). We obtained loadings on this first genetic PC for each regional volume, quantifying how well an individual volume mapped onto the underlying dimension of shared morphometry across the whole brain. The distribution of the whole-brain PC loadings ranged between 0.30 and 0.81 (mean = 0.62, *SD* = 0.13, median = 0.65; Figure 4, STable 3a).

We used the same approach – extracting the first genetic PC and its genetic PC loadings – to examine nine predefined genetic brain subnetworks (Figure 1). The brain volumes included in those networks are listed in STable 2. The percentage of genetic variance accounted for by the first network-specific PCs ranged between 65% for the central executive network and 47% for the temporo-amygdala-orbitofrontal network. While the central executive and the hippocampal-diencephalic networks had a narrow, unimodal distribution of PC loadings, the temporo-amygdala-orbitofrontal and cingulo-opercular networks had a wider, and bimodal distribution. That is, volumes included in the central executive network, for example, were more homogeneous and indexed more similar genetic variation, compared with the temporo-amygdala-orbitofrontal network (Figure 4). Overall, percentages of explained variances were larger for networks including fewer volumes, potentially because larger networks tend to be more heterogeneous.

To test whether data-derived PCs explained more genetic variance than could be expected by chance, we present a version of parallel analysis which simulates PCs for uncorrelated elements with matched genetic sampling variance (see Methods). Parallel analysis confirmed that genetic PCs of the whole brain and the nine canonical subnetworks explained significantly and substantially more variance than expected by chance (Scree Plots SFigures 11-20). Furthermore, we demonstrated that PCs extracted from 800 networks with randomly included brain volumes explained substantially less averaged variance than empirical canonical networks. Mean explained variances by the first PCs (randomly including

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.22.465437; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

volumes) and a 95% confidence interval around the mean are presented in STable 5. In summary, these results illustrate that genetic dimensions of shared morphometry are well represented by the first underlying PC (i.e., accounts for the majority of genetic variance); the dimensions differ between networks, and that they explain similar magnitudes of variance as their corresponding phenotypes.

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.22.465437; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

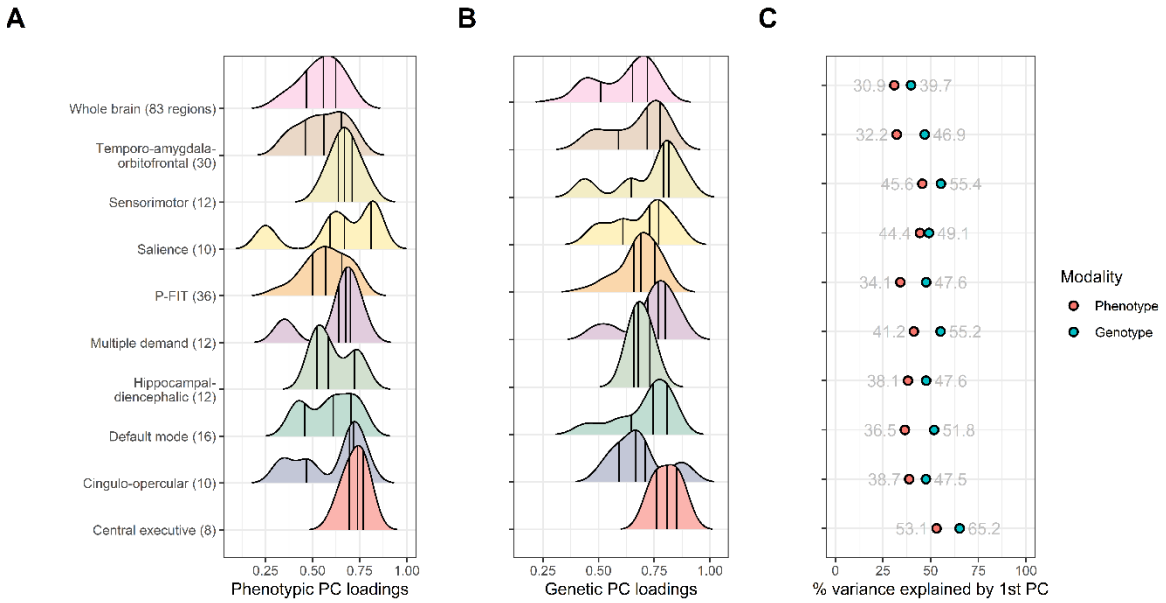

Figure 4. (A) Density distributions of principal component (PC) loadings on the first PC underlying phenotypic and (B) genetic networks. Vertical lines indicate quantiles. (C) Variance explained by phenotypic and genetic first PC in each network.

General dimensions of phenotypic and genetic shared morphometry across brain volumes were similarly organised

To quantify how closely patterns of shared variance between phenotypic and genetic brain volumes resemble each other, we calculated a linear regression between sets of 83 phenotypic and 83 genetic PC loadings. PC loadings indicate relative magnitudes of brain regions' loadings on either phenotypic or genetic dimensions of shared morphometry, and serve as an index of how well a volume represents trends across the brain (or the network). The association between phenotypic PC loadings and genetic PC loadings was large and statistically significant (b = 0.65, SE = 0.06, p = 5.07 x10^{-17}, R^2 = 58%), indicating that an increase in one unit in the genetic PC loadings is associated with an increase of .65 units in

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.22.465437; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

the phenotypic PC loadings (intercept = 0.15). This approach considers ordering relative to the mean. The mean genetic loading was 0.62 (*SD* = 0.13), and the mean phenotypic loading was 0.55 (*SD* = 0.11).

The Tucker congruence coefficient was used to index the degree of similarity of genetic and phenotypic PC loadings, taking into account both their relative ordering and their absolute magnitudes [31]. The Tucker coefficient revealed very high congruence in the deviation from zero between phenotypic and genetic PC loadings for the 83 volumes (Tucker coefficient = 0.99). These results illustrate a close correspondence and an equivalent organisation of phenotypic and genetic dimensions of shared morphometry; a finding that aligns with Cheverud's Conjecture.

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.22.465437; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

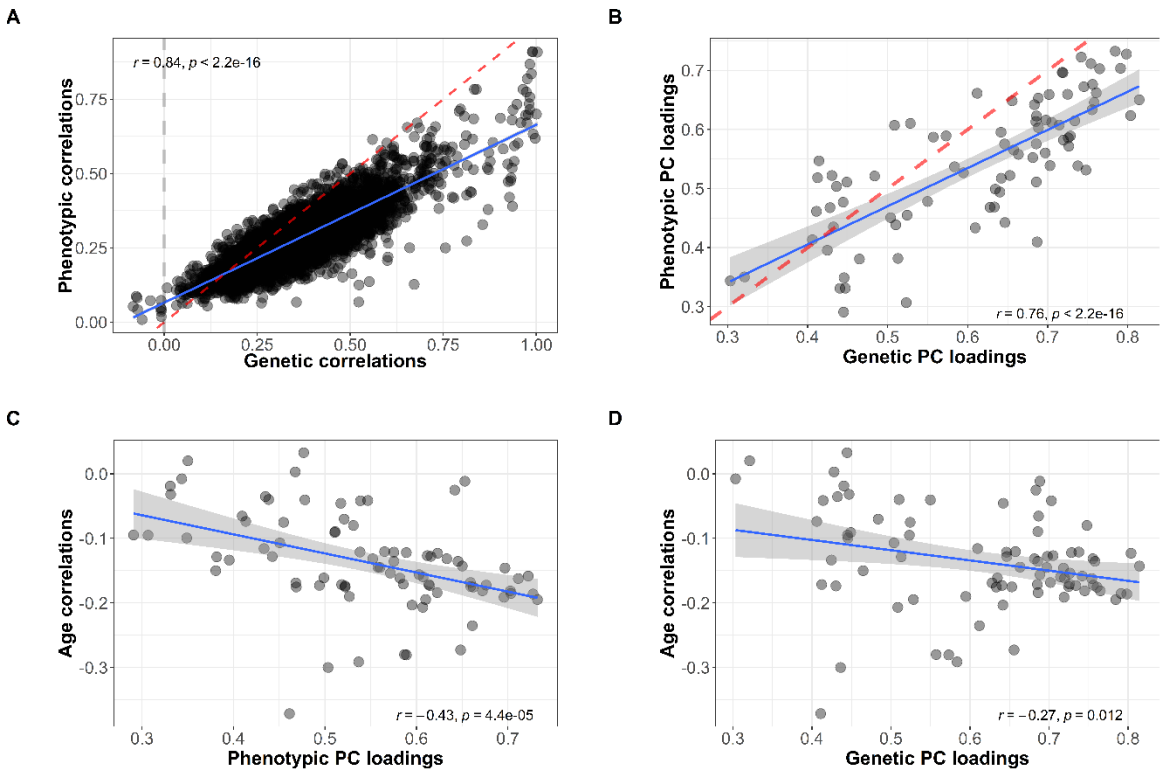

Figure 5. **(A)** Association between phenotypic and genetic between-region correlations of 83 regional grey-matter volumes. The dashed red line is the line of identity, with a slope of 1 and an intercept of 0. The dashed grey line indicates $r_g = 0$. **(B)** Correlation between phenotypic and genetic PC loadings on the first PC underlying 83 regional volumes. The dashed red line is the line of identity. **(C)** Correlation between *phenotypic* PC loadings and age sensitivity as indexed by phenotypic cross-sectional age-volume correlations. **(D)** Correlation between *genetic* PC loadings and age sensitivity as indexed by phenotypic cross-sectional age-volume correlations.

Genetic dimensions of shared morphometry were associated with age sensitivity

Previous work demonstrated an association between phenotypic dimensions of shared morphometry across the whole brain, represented by phenotypic PC loadings, and indices of *age sensitivity* [6]. Age sensitivity is approximated by a correlation of a regional brain volume with age across the sample, which is typically negative in adult populations. Here, we replicated this association between phenotypic shared morphometry (i.e., phenotypic PC

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.22.465437; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

loadings) and age sensitivity (r = -0.43, p = 4.4 x10^{-5}; Figure 5c), and we found a significant, but smaller association for genetic PC loadings (r = -0.27, p = 0.012; Figure 5d). That is, we demonstrated that the more the genetic variation of a brain volume resembles general morpjometric trends across the brain (larger genetic PC loading), the more negatively correlated this volume is with age. These are notable results because the PC loadings were extracted from brain volume variables *residualised* for age, and nevertheless, demonstrated a significant association with age sensitivity. In summary, these results show that phenotypic PC loadings and genetic PC loadings both display associations with patterns characteristic of ageing-related neurodegeneration, as indexed by cross-sectional age-volume correlations.

General dimensions of shared morphometry were genetically correlated with general cognitive ability

To quantify genetic correlations between general dimensions of network morphometry and general cognitive ability, we indexed shared genetic variance across brain networks, by summarising genome-wide per-SNP effects across multiple grey-matter volume GWAS summary statistics, weighted by volume- and network-specific PC loadings (novel method presented in Figure 2.4). Using the GenomicSEM software [27], we calculated genetic correlations between the brain networks and seven cognitive traits [27] (SFigure 21). The cognitive traits mostly had high loadings on a genetic general cognitive ability factor (median = 0.81, range = 0.30-0.95); the Reaction Time task had the lowest loading on the factor (SFigure 22). High genetic correlations between the brain networks indicated that they indexed very similar polygenic signal (r_g ranging from 0.63 between multiple demand and hippocampal-diencephalic to 0.97 between P-FIT and whole brain). All networks were significantly genetically associated with the general cognitive ability factor; correlation

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.22.465437; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

magnitudes across all networks ranged between r_g = 0.17 and 0.21 (Table 1). According to commonly-used rules of thumb from Hu and Bentler [32](CFI > 0.95, RMSEA < 0.08), all models showed good model fit (STable 4).

Table 1. Genetic correlations between general cognitive ability and nine canonical brain networks

Network	***Included volumes***	r_g	**95% CI**	***p*-value**	**FDR *q*-value**
Whole brain	83	0.21	0.13-0.29	1.00×10^{-7}	3.00×10^{-7}
Central executive	8	0.20	0.12-0.27	1.00×10^{-7}	3.00×10^{-7}
Cingulo-opercular	10	0.20	0.13-0.27	1.00×10^{-7}	3.00×10^{-7}
Default Mode	16	0.19	0.12-0.26	2.00×10^{-7}	3.00×10^{-7}
Hippocampal-Diencephalic	12	0.17	0.09-0.24	2.66×10^{-5}	2.66×10^{-5}
Multiple Demand	12	0.19	0.12-0.27	7.00×10^{-7}	9.00×10^{-7}
P-FIT	36	0.20	0.12-0.27	2.00×10^{-7}	3.00×10^{-7}
Salience	10	0.19	0.12-0.26	3.00×10^{-7}	4.00×10^{-7}
Sensorimotor	12	0.19	0.11-0.27	1.20×10^{-7}	1.30×10^{-6}
Temporo-amygdala-orbitofrontal	30	0.20	0.12-0.27	2.00×10^{-7}	4.00×10^{-7}

r_g = genetic correlation between brain network and a factor of general cognitive ability modelled from seven cognitive traits, SE = standard error, 95% CI = 95% confidence interval, *p*-value = original *p*-value as indicated by the GenomicSEM model, false discovery rate (FDR) *q*-value = *p*-value corrected using 5% false discovery rate.

Based on phenotypic findings that have highlighted the importance of the central executive network to general cognitive function [6], we hypothesised that we would find a stronger genetic association between general cognitive ability and the central executive network relative to other subnetworks (see pre-registered plan; https://osf.io/7n4qj). There was no evidence for significant differences in total correlation magnitudes between the central executive network and general cognitive ability compared with other brain networks, even after accounting for network sizes. We accounted for network sizes by comparing

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.22.465437; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

relative genetic associations between networks and general cognitive ability that had been divided by the number of regions included in a network (see Methods; STable 6). This was examined using GenimicSEMs [27] in which total and relative genetic correlation magnitudes were either estimated freely or constrained to be equal (SFigure 23). The two nested models did not significantly differ in model fit ($\Delta\chi^2$ *p*-values ranged between .072 and 1.00; STable 5).

We also investigated whether genetic associations were driven by specific cognitive traits. We obtained non-significant Q_{trait} heterogeneity indices [33] for all brain networks, demonstrating that the general cognitive ability factor accounted well for the patterns of association between specific cognitive abilities and the brain networks (SFigure 24). Genetic correlations were mostly significant between brain networks and three representative specific cognitive abilities with high test-retest reliabilities [34]: (1) *Matrix Pattern Completion task*, (2) *Memory – Pairs Matching Test*, and (3) *Symbol Digit Substitution Task*. Of 30 genetic correlations, 23 survived correction for multiple testing (5% FDR; STable 7). The fact that the general cognitive ability factor accounted well for specific abilities, and that the specific abilities were mostly significantly associated with the networks, confirms that the genetic associations between specific cognitive abilities and brain networks are likely general and act through a factor of general cognitive ability.

General dimensions of shared morphometry were genetically correlated with brain age

Finally, we calculated a genetic correlation between shared morphometry across the whole brain and *brain age*. Brain age is based on individual-level predictions of how much older (or younger) an individual's brain appears from structural MRI measures, relative to their chronological age [35] (see Methods). We found a moderate negative genetic association (r_g = -0.34; *SE* = 0.06) between general dimensions of shared morphometry

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.22.465437; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

across the whole-brain and brain age, suggesting that consistently larger volumes across the whole brain indicate younger brain age.

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.22.465437; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Discussion

This study provides a new line of evidence on brain organisation and its relations with ageing and cognitive ability, by analysing genetic data. We demonstrated using a novel statistical approach that both phenotypic and genetic data can be used reliably to extract general dimensions of shared morphometry across canonical brain networks. We modelled the genetic architecture of brain networks by extracting general dimensions of human brain morphometry, based on PCA of genetic overlap (i.e., shared polygenic signal) between brain volumes. These general dimensions accounted for substantial amounts of the total genetic variance in brain regional morphometry and displayed distinct descriptive characteristics.

As predicted by Cheverud's Conjecture, phenotypic and genetic dimensions of shared morphometry were similarly organised and may therefore share developmental pathways, which is consistent with a genetic foundation of canonical brain networks. We found that general dimensions of brain morphometry genetically correlated with general cognitive ability, brain age, and cross-sectional profiles characteristic of ageing-related neurodegeneration. These findings indicate that pathways giving rise to genetic variation in brain morphometry covary between individuals alongside pathways underlying cognitive ability and vulnerability towards ageing.

Characteristics of genetic brain network organisation

We demonstrated that genetic dimensions of shared morphometry underlying brain networks are quantifiable through the first genetic PC, which accounted for most of the systematic variance shared between brain volumes (e.g., 40% of the variance across the whole brain explained by 1st genetic PC; only 7% explained by the 2nd PC). Such major genetic dimensions even explained more variance than their phenotypic analogue (e.g., 40% of the variance across the whole brain explained by first genetic PC, 31% explained by first phenotypic PC). All genetic networks explained substantially more variance than was

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.22.465437; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

expected by chance, which was supported by parallel analysis modelled using empirical sampling variances, as well as simulations of networks with randomly included brain volumes. These findings provide a new line of evidence characterising and underpinning the existence of a genetic foundation for canonical brain networks that have featured prominently in neuroscientific studies [e.g., 7].

Analogous organisation of phenotypic and genetic dimensions of shared morphometry

We discovered a high degree of similarity between phenotypic and genetic features of brain network organisation. This was supported by highly similar correlation structures within the whole brain and the subnetworks ($r_{\text{genetic vs. phenotypic correlations}}$ = 0.84, *p*-value = 2.2 x10^{-16}). Phenotypic and genetic PC loadings on their respective major dimensions of shared morphometry were comparable both in their absolute magnitudes and their relative ordering to the mean (Tucker congruence = 0.99; $r_{\textit{phenotypic vs. genetic PC loadings}}$ = 0.76, *p*-value = 2.2 x10^{-16}). According to Cheverud's Conjecture [1], this indicates that brain organisation, indexed through both phenotypic and genetic variance in brain-wide volumes, seems to be underpinned by similar, overlapping developmental pathways.

Dimensions of regionally shared morphometry as an index of cross-sectional profiles characteristic of neurodegeneration

A previous study demonstrated that phenotypic PC loadings onto an underlying dimension of shared morphometry across the whole brain demonstrated patterning similar to volumetric vulnerability towards ageing (i.e., cross-sectional age-volume correlations). That is, the more representative a brain volume's variance was for general morphometric trends across the whole brain, the more strongly its volume was negatively correlated with age. Here, we replicated this negative association between phenotypic PC loadings and age-volume correlations, and showed that this association also exists, albeit to a lesser degree, with genetic instead of phenotypic PC loadings. This suggests that dimensions along which

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.22.465437; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

brain regions share morphometric variance (i.e., generally larger volumes across an individuals' brain) are structured similarly to patterns by which brain regions display increased vulnerability to ageing. Future longitudinal studies and cross-sectional studies modelling within-person atrophy by incorporating information on prior brain size (e.g., intracranial volume) are needed to triangulate this finding.

One potential explanation for this association is that brain regions that: are genetically predisposed to be large volumes, share higher levels of morphometric variance with the rest of the brain, and are more central to heavily-demanding cognitive processes, might come under more strenuous developmental and environmental pressure, perhaps through increased metabolic burden, compared with other, less central regions. Thus, the embedding of a brain volume within the whole brain's organisation, and the genetic foundation of its positioning in the brain, could govern the functional stresses and other influences to which certain areas are exposed. This might alter disproportionately the speed at which some regions atrophy with advancing age.

That dimensions of shared morphometry resemble patterns of age sensitivity is of particular interest, because it emerged from shared variance among brain volumetric phenotypes that had been residualised for age. Consequently, we suggest that patterns of brain structural ageing, a construct labelled *brain age* [24], might not capture how quickly an individual's regional volumes decline compared to their peers, but rather, general healthy morphometry across the brain. Previous research showed that a younger-appearing brain, relative to the individual's chronological age, predicted better physical fitness, better fluid intelligence and longevity [24]. Healthy brain morphometry could vary between people for many non-age-related reasons, including genetic predisposition. Individuals that are genetically predisposed towards consistently larger brain volumes across the whole brain

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.22.465437; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

might have generally healthier, better-integrated brains, which could be more resilient towards harmful environmental factors.

To test this theory, we calculated the genetic correlation between general dimensions of shared morphometry across the whole-brain and brain age. We found that younger brain age was genetically associated with greater overall brain volume as indexed by the general dimension of shared morphometry (r_g = -0.34; SE = 0.06). Thus, consistently larger volumes across the brain indicate a younger structural brain organisation, and this is the first study to quantify the degree to which brain age and a major dimension of shared morphometry between brain volumes are overlapping concepts. It motivates further investigation into the possibility that they are partly underpinned by general shared biological pathways.

Dimensions of shared morphometry as an index of cognitive performance

This study demonstrated that general cognitive ability is associated with shared genetic morphometric variance across the whole brain, and across smaller canonical networks. This was investigated using seven cognitive ability GWAS summary statistics , which were used to model a genetic factor of general cognitive ability using GenomicSEM [27]. We then calculated the genetic correlation between this factor of general cognitive ability and general dimensions of shared morphometry across the whole brain, and nine canonical subnetworks. The genetic networks were represented by univariate summary statistics of a genetic PC underlying multiple brain volume GWAS phenotypes. The whole brain and all nine networks were significantly genetically correlated with general cognitive ability at magnitudes between 0.17 and 0.21. This was the same level of genetic association with general cognitive ability that was previously found for broad measures of total brain volume [36]. There was no evidence to suggest that those magnitudes statistically differed between the networks; probably due to the fact that the polygenic signal indexed by the

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.22.465437; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

network-morphometry summary statistics was highly similar between brain networks (mean r_g between networks 0.83, SD = 0.09).

This indicates that the genetic association between brain morphometry and cognitive ability was not driven by specific network configurations. Instead, dimensions of shared genetic morphometry in general indexed genetic variance relevant to larger brain volumes and a brain organisation that is advantageous for better cognitive performance. This was regardless of how many brain regions and from which regions the measure of shared genetic morphometry was extracted. This lack of differentiation between networks, in how strongly they correlate with cognitive ability, is in line with the suggestion that the total number of neurons in the mammalian cortex, which should at least partly correspond to its volume, is a major predictor of higher cognitive ability [37]. These findings suggest that highly shared brain morphometry between regions, and its genetic analogue, indicate a generally bigger, and cognitively better-functioning brain.

Unexpectedly, genetic correlations between networks and cognitive ability did not suggest any prominent role of the central executive network (a previous *phenotypic* study [6] demonstrated that the central executive network was disproportionately predictive of cognitive abilities relative to its few included volumes). On a genetic level of analysis, we expected a stronger correlation with cognitive ability for the central executive network compared with the other networks, which would have indicated that genetic variation relevant to larger volumes, specifically within the central executive network, are associated with higher general cognitive ability. The fact that we found that larger volumes across the whole brain, and all other networks were just as closely correlated with cognitive ability as the central executive network, taken together with previous phenotypic evidence for such a disproportionately large association between the two, suggests nongenetic mechanisms to

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.22.465437; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

play important roles, perhaps developmental and environmental influences, through which the central executive network matures, and specialises for cognitive performance.

Limitations

Analyses in this study come with limitations. Genetic correlations were estimated based on polygenic biology, assuming that they are representative for genetic associations across the entire genome. As genetic correlations were calculated using LDSC, the limitations that apply to LDSC methodology are relevant to our study (discussion in Supplementary Note). We conclude based on heritability estimates, indexing signal-to-noise ratios in GWAS, that there was sufficient polygenic signal to warrant LDSC analysis (heritability ranged 7-42%). LDSC intercepts were perfectly associated with phenotypic correlations, indicating that the analyses successfully separated confounding signal (including environmental factors) from the estimates of genetic correlations.

This study was conducted in the UK Biobank sample, which is not fully representative of the general population: its participants are more wealthy, healthy and educated than average [38]. Cohort effects may affect the degree to which differential cortical regional susceptibility to ageing can be inferred from cross-sectional data. It remains to be tested whether our results can be extrapolated to socio-economically poorer subpopulations, or outside European ancestry. Results were also dependent on the choice of brain parcellation to divide the cortex into separate regions.

Conclusion

This pre-registered study modelled genetic brain organisation of grey-matter volumes in the human whole brain and nine canonical subnetworks, using the largest available, high-quality brain volume GWAS summary statistics. We presented a novel methodology to summarise genetic patterns of shared morphometry across brain volumes. We demonstrated

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.22.465437; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

that general dimensions of shared genetic morphometric variance differ between canonical brain networks and are very similarly organised at both phenotypic and genetic levels of analysis.

This study provides a new methodological framework which we used to interrogate the overlapping molecular genetic factors underlying brain morphometry, cognitive ability, and brain ageing. We found that genetic dimensions of shared morphometry across the whole brain, and nine smaller canonical brain networks, were negatively associated with cross-sectional profiles characteristic of ageing-related neurodegeneration and brain age. We also found that such dimensions of shared genetic morphometry were positively associated with cognitive ability. These findings provide evidence that biological pathways giving rise to genetic variation in brain morphometry correspond to pathways underlying vulnerability towards rapid adult ageing and cognitive ability. The methodology and resulting insights provide a basis for future investigations that aim to disentangle genetic and environmental causes of ageing and cognitive decline.

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.22.465437; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

STAR methods

UK Biobank data

Magnetic resonance imaging **(**MRI) data was collected by the UK Biobank study with identical hardware and software in Manchester, Newcastle, and Reading. Brain volumetric phenotypes were pre-processed by an imaging-pipeline developed and executed on behalf of UK Biobank [39]. More information on T1 processing can be found in the UK Biobank online documentation [40]. Briefly, cortical surfaces were modelled using FreeSurfer, and volumes were extracted based on Desikan-Killiany surface templates [41]; subcortical areas were derived using FreeSurfer aeseg tools [42]. Volumetric measures (mm^3) have been generated in each participant's native space. We used 83 available imaging-derived phenotypes (IDPs) of cortical and subcortical grey-matter volumes in regions of interest spanning the whole brain (UK Biobank category 192 & 190; STable 1). We assume the IDPs to be normally-distributed.

Phenotypic quality control

Excluding participants who withdrew consent, we considered 41,776 participants with non-missing T1-weighted IDPs that had been processed in conjunction with T2-weighted FLAIR (UK Biobank field ID 26500) where available. Using both T1 and T2 measures ensures more precise cortical segmentation [43]. Extreme outliers outside of 4 standard deviations from the mean were excluded, which resulted in between 41,686 to 41,769 available participants depending on the IDP. 381 participants were excluded as they self-reported non-European ethnicity. Across the 83 brain volumes variables and the covariates, this phenotypic quality control resulted in 39,947 complete cases, for whom the following genetic quality control steps were performed.

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.22.465437; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Genetic quality control

Out of the 39,947 UK Biobank participants, genetic data were available for 38,957 participants. Genetic data was quality controlled on by UK Biobank and were downloaded from the full release [44]. We applied additional quality control as previously described in Coleman et al. [45] using PLINK2 [46]. 38,038 participants were of European ancestry according to 4-means clustering on the first two genetic principal components available through UK Biobank [47]. Of those participants, we removed 72 due to quality assurance provided by UK Biobank and 204 participants due to high rates of missingness (2% missingness). To obtain a sample of unrelated individuals, 956 participants were removed using the greedyRelated algorithm (KING $r < 0.044$ [48]). The algorithm is "greedy" because it maximises sample size; for example, it removes the child in a parent-child-trio. Finally, 28 participants were removed because genetic sex did not align with self-reported sex, resulting in a total of 36,778 participants (STable 10). Genetic sex was identified based on measures of X-chromosome homozygosity (F_X ; removal of participants with $F_X < 0.9$ for phenotypic males, $F_X > 0.5$ for phenotypic females). The final sample (N = 36,778) included 19,888 females (54 %) and had an average age of 63.3 years at the neuroimaging visit (range from 40.0 to 81.8 years).

Out of 805,426 available directly genotyped variants, 104,771 were removed for high rates of missing genotype data (> 98%). 103,137 variants were removed due to a minimum allele frequency of 0.01, and 9,935 variants were removed as they failed the Hardy-Weinberg exact test (p-value = 10^{-8}). After excluding 16,326 variants on the sex chromosomes and those with chromosome labels larger than 22, we obtained a final sample of 571,257 directly genotyped SNPs. Imputed genotype data was obtained by UK Biobank with reference to the Haplotype Reference Consortium [49], and we filtered them for a minor allele frequency of above 0.01 and an IMPUTE INFO metric of above 0.4.

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.22.465437; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Cognitive performance. UK Biobank collected cognitive performance data using assessment on a touchscreen computer. The following seven tests were implemented: *Matrix Pattern Completion task* for nonverbal reasoning, *Memory – Pairs Matching Test* for memory, *Reaction Time* for perceptual motor speed, *Symbol Digit Substitution Task* for information processing speed, *Trail Making Test – B* and *Tower Rearranging Task* for executive functioning, and *Verbal Numerical Reasoning Test* for verbal and numeric problem solving, or fluid intelligence. Despite the non-standard and unsupervised delivery of assessment, these cognitive tests demonstrate strong concurrent validity compared with standard reference tests (r = .83) and good test-retest reliability (Pearson r range for different cognitive tests = 0.4–0.78) [34].

In this study, we considered GWAS summary statistics of performance in these seven cognitive tests provided by de la Fuente et al. [50] that had been calculated with varying numbers of participants for each test, ranging from 11,263 to 331,679 participants. Here, we consider the HapMap 3 reference SNPs with the MHC regions removed.

Statistical analysis

Genome-wide association study (GWAS) summary statistics calculation. GWAS summary statistics for the 83 regional brain volumes (continuous variables) were calculated using the REGENIE software [51], which fits polygenic effects in a computationally efficient linear mixed model using Ridge regression. The REGENIE pipeline is split into two steps: In the first step, blocks of directly genotyped SNPs are used to fit a cross-validated whole-genome regression model using Ridge regression, to determine the amount of phenotypic variance explained by genetic effects. In a second step, the association between the phenotype and imputed genetic variants is calculated conditional upon Ridge regression

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.22.465437; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

predictions from the first step. Proximal contamination is circumvented by using a leave-one-chromosome-out scheme.

Covariates included in the GWAS analyses were *age at neuroimaging visit*, *sex*, *genotyping batch*, and *40 genetic principal components* as provided by UK Biobank. We also derived the variables *time of year*, *head position*, and *acquisition site*, but excluded them from our set of GWAS covariates because they were not associated with the brain volumes at the pre-registered arbitrary cut-off of $r \leq .10$ (STable 9), and therefore explained less than 1% of the phenotype variance. Note that, in contrast to other existing brain-volume GWAS in UK Biobank [e.g., 52], our analyses were conducted *without* controlling for brain size (or any other global brain measure such as total grey-matter volume or intracranial volume). Genetic correlations calculated relative to such global measures are known to attenuate genetic correlations among volumes, as well as with other traits such as cognitive abilities [21]. In the context of this study, attenuated genetic correlations would not allow us to model genetic brain networks, because most of the variance shared between volumes overlaps with variance indexed by brain size and would therefore not tag general dimensions of shared genetic variance between brain volumes.

Genetic and phenotypic correlation matrices between brain volumes. To derive dimensions of shared morphometry across brain volumes, we calculated both a phenotypic and a genetic correlation matrix from 83 grey-matter volume variables. Phenotypic regional brain volumes were residualised for age at neuroimaging visit and sex, and then used to estimate a phenotypic correlation matrix through Pearson's correlations with complete pairwise observations. The genetic correlation matrix was inferred through LDSC, a technique quantifying shared polygenic effects between traits using GWAS summary statistics. Cross-trait LDSC regresses the product of effect sizes in two GWAS onto linkage disequilibrium scores, indicating how correlated a genetic variant is with its neighbouring

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.22.465437; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

variants [26]. The slope indexes the genetic correlation, while the intercept captures signal uncorrelated with LD, such as population stratification, environmental confounding, and sample overlap.

To quantify the relationship between phenotypic and genetic correlations, we estimated the correlation between 3403 phenotypic and genetic between-region correlations ($\frac{83(83-1)}{2}$ = 3403 correlations between 83 volumes). Additionally, we calculated genetic correlation matrices for smaller canonical networks including fewer brain volumes than the whole brain. For example, the central executive network included eight regional volumes (STable 2 lists volumes included in the nine canonical networks). We reported SNP-heritability estimates for each brain volume inferred through LDSC.

Principal component analysis of genetic and phenotypic correlation matrices. Principial component analysis (PCA) was applied to the phenotypic and genetic correlation matrices indicating genetic overlap between brain volumes described above to obtain their respective first principal component (PC). The first PC represents an underlying dimension of common structural sharing across regional volumes, which we refer to as general dimensions of shared morphometry throughout this manuscript. PC loadings were calculated for all volumes in the whole brain, as well as volumes in smaller canonical networks to quantify contributions of regional volumes to this either brain-wide, or network-specific dimension of shared morphometry.

Parallel analysis. We tested whether genetic PCs explained more variance than expected by chance, that is, whether they explained more than 95% of their corresponding PCs generated under a simulated null correlation matrix. We developed a version of parallel analysis to generate null distributions of eigenvalues by simulating null correlation matrices sampled from a diagonal population correlation matrix, where the multivariate sampling

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.22.465437; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

distribution is specified to take the form of the sampling distribution of the standardised empirical genetic correlation matrix (the V_{STD} matrix, as estimated using GenomicSEM [27]). This sampling correlation matrix serves as an index of the precision of the elements in the empirical genetic covariance matrix (i.e., heritabilities and co-heritabilities across traits) and the sampling dependencies among these when generating the random null models. We specified 1,000 replications to simulate the null correlation matrices and use a 95% threshold for distinguishing true eigenvalues from noise.

Simulation of networks with randomly included brain volumes. We performed an additional sensitivity analysis simulating networks with randomly included brain volumes, to determine whether shared structural variance relied on network membership, or arose through phenotypic properties common to all regional brain volumes. To compare explained variances between canonical networks and random networks, we quantified the expected explained variance in random networks by randomly sampling regions 800 times each, for different numbers of included volumes (because networks including fewer volumes generally tend to explain a larger percentage of variance, as larger networks are more heterogeneous). That is, simulations were run for 8, 10, 12, 16, 30, and 36 included regions, to obtain a distribution for each networks size to compare the corresponding network's explained variance to. We reported the mean explained variance by PCs for networks with randomly included volumes and a 95% confidence interval. Comparisons between explained variances for random and canonical networks were done for the same number of included volumes.

Correlation between phenotypic and genetic PC loadings. To compare whether genetic correlations structures of regional brain morphometry resembled the phenotypic correlation structure of the same regions, we calculated an un-standardised linear regression with a vector of 83 phenotypic whole-brain PC loadings as the dependent variable, and a vector containing 83 genetic whole-brain PC loadings as the independent variable. We

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.22.465437; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

calculated the Tucker congruence coefficient to quantify the relative similarity between the two sets of PC loadings independent of their absolute magnitude. The coefficient is insensitive to scalar multiplication [53].

Correlation between genetic PC loadings with age sensitivity. Pearson's correlations between 83 phenotypic grey-matter volumes and age at neuroimaging visit were calculated to quantify cross-sectional age-volume-correlations for each of the 83 brain volumes. These age-volume correlations are referred to as *age sensitivity* throughout the rest of the manuscript. We estimated the correlation between a vector containing indices of age sensitivity and (1) a vector of *genetic* whole-brain PC loadings, and for comparison (2) a vector of *phenotypic* whole-brain PC loadings.

Genome-wide shared genetic variance of morphometry across the whole brain and canonical networks. To statistically represent genome-wide shared morphometric variance across brain volumes (i.e., genetic PCs), we developed a novel method summarising genome-wide by-variant effects contained in the grey-matter volume GWAS summary statistics, which were weighted by their respective (region-specific) PC loadings obtained through PCA. We derived GWAS summary statistics for a genetic principal component of multiple GWAS phenotypes derived from samples of unknown degrees of overlap by adapting existing software for genome-wide multivariate meta-analysis by Baselmans, Jansen [29] and using GenomicSEM [27]. Figure 2 illustrates this approach in a four-step procedure. The input data for our approach are GWAS summary statistics for 83 cortical and subcortical brain volumes (step 1). We have made them publicly available online. Using the GenomicSEM software [27], we obtained a genetic correlation matrix indicating genetic overlap between these 83 brain volumes (step 2). We extracted PC loadings on the underlying general dimension of shared genetic variance for each of the 83 regions (step 3). Finally, we modified the existing genome-wide multivariate meta-analysis software package [29], in

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.22.465437; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

order to create summary statistics for an underlying genetic PC. Genome-wide SNP effects were calculated as an average of all SNP effects contributed by the 83 GWAS phenotypes, weighted by their respective PC loading and sample size. SNP effects are also adjusted for confounding factors such as sample overlap and population stratification using LDSC intercepts (step 4). We used this approach to calculate univariate summary statistics to represent general dimensions of shared morphometry between regional volumes across the whole brain (83 GWAS phenotypes), as well as nine smaller canonical networks.

We had tested and validated this novel approach in an independent set of GWAS summary statistics of four risky behaviours [28]. In addition to the risky behaviour GWAS, another set of summary statistics is available for a phenotypic PC underlying these risky behaviour phenotypes that the authors had calculated phenotypically before running GWAS analyses. We compared these phenotypic PC GWAS summary statistics by Linnér, Biroli [28] with summary statistics for a *genetic* PC underlying the four risky behaviours GWAS that we calculated using our novel method outlined above (Figure 2). We found that they correlated at a magnitude of $r_g = 0.99$ ($SE = 0.037$) confirming that our method captures the same signal as can be obtained from phenotypic PCs, by simply relying on publicly available GWAS data. For details of the analysis and code refer to: https://annafurtjes.github.io/genomicPCA/ .

Genetic correlation between general dimensions of shared morphometry across the whole-brain and brain age. Using LDSC [26], we calculated a genetic correlation between genetic morphometric sharing across the whole brain and *brain age*. The summary statistics indexing dimensions of shared morphometry across brain volumes were created using the novel method presented above (Figure 2). We downloaded the brain age GWAS summary statistics online [35]. Brain age is a phenotype based on individual-level predictions of how much older (or younger) an individual's brain appears, relative to their chronological

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.22.465437; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

age. It is estimated using parameters characterising the relationship between age and structural neuroimaging measures (volume, thickness, and surface area) that were tuned using machine learning in an independent sample. The final brain age phenotype indexed in the GWAS was calculated as the difference between participants chronological age and their age as predicted based on structural brain characteristics.

Genetic correlations between brain networks and a factor of general cognitive ability. We assessed genetic correlations between brain networks and general cognitive ability using GenomicSEM [27]. Using univariate network-specific summary statistics (as describe above; Figure 2) and a genetic general cognitive ability factor modelled from seven cognitive ability GWAS summary statistics , the GenomicSEM software [27] was used to model general cognitive ability and perform multivariate LDSC using diagonally weighted least squares. To quantify model fit, we reported default fit indices calculated by the GenomicSEM package: χ^2 values, the Akaike Information Criterion (AIC), the Comparative Fit Index (CFI) and the Standardised Root Mean Square Residuals (SRMR). The multiple testing burden was addressed by correcting *p*-values from the genetic correlations for multiple testing with a false-positive discovery rate of 5% [54].

We preregistered that we would test for significant differences in correlation magnitudes between the networks that yielded a significant association with general cognitive abilities. Because we hypothesised a particularly strong association for the central executive network, we planned to perform this comparison between the central executive and all other networks, to reduce the multiple testing burden. We fitted two GenomicSEM models in which correlation magnitudes between general cognitive ability and both the central executive and another network were either freely estimated, or they were forced to be the same. A significant decrease in model fit between the freely estimated model and the constrained model ($df = 1$) would indicate that there likely are differences in correlation

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.22.465437; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

magnitudes between the networks in how strongly they correlate with general cognitive ability (SFigure 23).

Additionally, we assessed whether the central executive network was disproportionately genetically correlated with general cognitive ability considering its small size (i.e., few included volumes). Similar to the approach described above, we fitted two models: One, in which we freely estimate the correlation between the central executive and general cognitive ability, and the correlation between another network and general cognitive ability. We then divided the correlation magnitude by the number of regions included in the network (i.e., magnitude was divided by 8 for the central executive network, it was divided by 16 for the default mode, by 36 for the P-FIT etc.). The second model had the same set up, but we forced the adjusted correlations for the two networks to be equal (e.g., $r_{\text{central executive}} / 8 == r_{\text{default}} / 16$). We assessed whether there was a significant difference in χ^2 model fit between these two models. As above, a significant decrease in model fit between the freely estimated model and the constrained model ($df = 1$) would indicate that there likely are differences in relative correlation magnitudes (i.e., magnitudes adjusted for network sizes). Based on previous findings, we expected the relative magnitude for the central executive network to be significantly larger than the relative magnitude for any other network.

To probe whether any specific cognitive ability might have driven the genetic associations between brain networks and general cognitive ability, we reported genetic correlations between the significant networks and three specific cognitive abilities: (1) *Matrix Pattern Completion task* to represent nonverbal reasoning, (2) *Memory – Pairs Matching Test* to represent memory, and (3) *Symbol Digit Substitution Task* to represent information processing speed. Reducing the analyses to only three consistent and representative cognitive measures reduced the burden of multiple testing.

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.22.465437; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

We calculated Q_{trait} heterogeneity indices [33] to evaluate whether the general cognitive ability factor that we fit in the models above accounts well for the specific cognitive abilities. To this end, we compared the fit of two models for each network as displayed in SFigure 24. One model allows for independent associations between the seven cognitive traits, and both general cognitive ability and the brain network. The second model forces the association between the seven cognitive traits and the brain network to go through the general cognitive ability factor. We obtained χ^2 fit statistics for both models and tested their difference for statistical significance ($\Delta\chi^2 \neq 0$; $df = 6$). Non-significant results ($p > 0.05/10$) would suggest that genetic associations between cognitive abilities and brain networks are likely general and act through a factor of general cognitive ability.

Data and code availability. Access to phenotypic and genetic UK Biobank data was granted through the approved application 18177. We have made the 83 GWAS summary statistics of regional volumes available at the GWAS catalogue (https://www.ebi.ac.uk/gwas/deposition). GWAS summary statistics for the seven cognitive traits by de la Fuente, Davies [50] were downloaded at https://datashare.ed.ac.uk/handle/10283/3756. The pre-registration for this analysis can be found online (https://osf.io/7n4qj). Full analysis code including results for this study are available at https://annafurtjes.github.io/Genetic_networks_project/index.html.

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.22.465437; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Acknowledgements

AEF is funded by the Social, Genetic and Developmental Psychiatry Centre, King's College London and the National Institute of Health (NIH) grant R01AG054628. SJR is funded by the Jacobs Foundation. JHC is funded by a UK Research & Innovation(UKRI) Innovation Fellowship (MR/R024790/1; MR/R024790/2). JF is funded by the National Institutes of Health (NIH) grant R01AG054628. JF and EMTD are members of the Population Research Center (PRC) and Center on Aging and Population Sciences (CAPS) at The University of Texas at Austin, which are supported by NIH grants P2CHD042849 and P30AG066614. JD, JWM, and EMTD were supported by NIH R01AG054628. IJD is with the Lothian Birth Cohorts group, which is funded by Age UK (Disconnected Mind grant), the Medical Research Council (grant no. MR/R024065/1) and the University of Edinburgh's School of Philosophy, Psychology and Language Sciences. The contribution by RA represents independent research part-funded by the National Institute for Health Research (NIHR) Maudsley Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care. The contribution by JRIC represents independent research part-funded by the National Institute for Health Research (NIHR) Maudsley Biomedical Research Centre at South London and Maudsley NHS Foundation Trust and King's College London. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health and Social Care. SRC is supported by a Sir Henry Dale Fellowship jointly funded by the Wellcome Trust and the Royal Society (Grant Number 221890/Z/20/Z). This research was funded in part by the Wellcome Trust [221890/Z/20/Z]. For the purpose of open access, the author has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission.
The authors gratefully acknowledge the UK Biobank resource (https://www.ukbiobank.ac.uk/) and its research team, who have made this work possible (project number 18177). The authors acknowledge use of the research computing facility at King's College London, Rosalind (https://rosalind.kcl.ac.uk), which is delivered in partnership with the National Institute for Health Research (NIHR) Biomedical Research Centres at South London & Maudsley and Guy's & St. Thomas' NHS Foundation Trusts, and part-funded by capital equipment grants from the Maudsley Charity (award 980) and Guy's & St. Thomas' Charity (TR130505).

Author contributions

Conceptualisation and methodology: SJR, EMTD, JHC, AEF, SRC

Supervision: SJR, EMTD, JHC

Network characterisation: SRC

Idea to investigate genetic brain age – shared morphometry correlation: JWM

Script used to perform genetic parallel analysis: JF

Data access: CML

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.22.465437; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Genetic quality control: AEF, JRIC

GWAS calculation: AEF, RA

Data analysis: AEF

Writing: AEF

Visualisations: AEF

Reviewed draft: all authors

Declaration of Interests

Ian Deary is a participant in UK Biobank.

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.22.465437; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Figures

Figure 1. Nine canonical brain networks

Figure 2. Four-step procedure to obtain statistical representation of genetic brain network structure from GWAS summary statistics. (1) GWAS summary statistics for 83 grey-matter volumes in UK Biobank from European ancestry were used as input data (*N* = 36,778). They were calculated using linear mixed models as described in Methods and are publicly available. (2) LDSC was used to infer genetic correlations between the 83 brain volumes. (3) This genetic correlation matrix is analysed using principal component analysis to derive PC loadings on the first PC, representing an underlying dimension of shared morphometry. (4) We have developed a method to derive GWAS summary statistics for a genetic principal component of multiple GWAS phenotypes derived from samples of unknown degrees of overlap using univariate summary statistics for the individual phenotypes. We have tested and validated this novel approach in an independent set of GWAS summary statistics [28]. Details of the analysis and code can be found at https://annafurtjes.github.io/genomicPCA/.

Figure 3. **(A)** Distribution of SNP-heritability estimates for 83 regional grey-matter volumes inferred through univariate LDSC. **(B)** Distribution of genetic correlations among 83 regional grey-matter volumes inferred through between-region LDSC. This figure only depicts between-region correlations but not the very high genetic inter-region correlations between regions and their homologous counterpart in the opposite hemisphere (excluding brain stem). **(C)** Distribution of phenotypic correlations among 83 regional grey-matter volumes inferred through Pearson's correlations. The raincloud plots were created based on code adapted from Allen, Poggiali [30].

Figure 4. Density distributions of principal component (PC) loadings on the first PC underlying (A) phenotypic and (B) genetic networks.

Figure 5. **(A)** Association between phenotypic and genetic between-region correlations of 83 regional grey-matter volumes. The dashed red line is the line of identity, with a slope of 1 and an intercept of 0. **(B)** Correlation between phenotypic and genetic PC loadings on the first PC underlying 83 regional grey-matter volumes. The dashed red line is the line of identity. **(C)** Correlation between *phenotypic* PC loadings and age sensitivity as indexed by phenotypic cross-sectional age-volume correlations. **(D)** Correlation between *genetic* PC loadings and age sensitivity as indexed by phenotypic cross-sectional age-volume correlations.

Tables

Table 1. Genetic correlations between general cognitive ability and nine canonical brain networks

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.22.465437; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Supplemental Information titles and legends

Supplementary Table 1. 83 cortical and subcortical grey-matter regions of interest

Supplementary Table 2. Network characterisation

Supplementary Table 3. Explained variance and descriptive statistics of PC loadings within phenotypic canonical networks

Supplementary Table 4. Model fit for genetic correlations between genetic general cognitive ability and each canonical network

Supplementary Table 5. Fit indices for the comparison between freely-varying or constrained correlations with general cognitive ability between central executive and other networks

Supplementary Table 6. Fit indices for the *adjusted* comparison between freely-varying or constrained correlations with general cognitive ability between central executive and other networks

Supplementary Table 7. Genetic correlations between three cognitive abilities and brain networks

Supplementary Table 8. Canonical networks explain more variance than networks with randomly included volumes

Supplementary Table 9. Associations between brain volumes and potential covariates

Supplementary Table 10. Genetic quality control exclusion criteria resulting in a total GWAS sample of 36,778 out of 39,947 participants

Supplementary Figure 1. Genetic correlation matrix inferred through LDSC across the whole brain (83 volumes).

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.22.465437; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Supplementary Figure 2. Genetic correlations inferred through LDSC among the central executive network (8 volumes).

Supplementary Figure 3. Genetic correlations inferred through LDSC among the cingulo-opercular network (10 volumes).

Supplementary Figure 4. Genetic correlations inferred through LDSC among the default mode network (16 volumes).

Supplementary Figure 5. Genetic correlations inferred through LDSC among the hippocampal-diencephalic network (12 volumes).

Supplementary Figure 6. Genetic correlations inferred through LDSC among the multiple demand network (12 volumes).

Supplementary Figure 7. Genetic correlations inferred through LDSC among the P-FIT network (36 volumes).

Supplementary Figure 8. Genetic correlations inferred through LDSC among the salience network (10 volumes).

Supplementary Figure 9. Genetic correlations inferred through LDSC among the sensorimotor network (12 volumes).

Supplementary Figure 10. Genetic correlations inferred through LDSC among the temporo-amygdala-orbitofrontal network (30 volumes).

Supplementary Figure 11. Parallel analysis in the central executive network

Supplementary Figure 12. Parallel analysis in the cingulo-operular network

Supplementary Figure 13. Parallel analysis in the default mode network

Supplementary Figure 14. Parallel analysis in the hippocampal-diencephalic network

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.22.465437; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Supplementary Figure 15. Parallel analysis in the multiple demand network

Supplementary Figure 16. Parallel analysis in the P-FIT network

Supplementary Figure 17. Parallel analysis in the salience network

Supplementary Figure 18. Parallel analysis in the sensorimotor network

Supplementary Figure 19. Parallel analysis in the temporo-amygdala-orbitofrontal network

Supplementary Figure 20. Parallel analysis in the whole brain with 83 nodes

Supplementary Figure 21. Genetic correlations between seven cognitive traits and brain networks. Descriptively, performance in the Tower Rearranging Task has the largest association with brain networks in comparison with other cognitive tasks. Abbreviations: Matrix = Matrix Pattern Completion task; Memory = Memory – Pairs Matching Test; RT = Reaction Time; Symbol Digit = Symbol Digit Substitution Task; Trails-B = Trail Making Test – B; Tower = Tower Rearranging Task; VNR = Verbal Numerical Reasoning Test; central exec = central executive; cingulo = cingulo-opercular; hippocampal = hippocampal-diencephalic; multiple = multiple demand; p fit = parieto-frontal integration theory; sensori = sensorimotor; temporo = temporo-amygdala-orbitofrontal

Supplementary Figure 22. Genetic correlation between the central executive network and factor *g* modelled for correlation structure of seven cognitive traits. The seven cognitive traits and the network are inferred through LDSC, and the factor through factor analysis. Matrix = Matrix Pattern Completion task; Memory = Memory – Pairs Matching Test; RT = Reaction Time; Symbol Digit = Symbol Digit Substitution Task; Trails-B = Trail Making Test – B; Tower = Tower Rearranging Task; VNR = Verbal Numerical Reasoning Test. Model fit: $\chi^2 = 124.04$, df = 20, *p*-value = 2.1×10^{-20}, AIC = 174.04, CFI = 0.97, SRMR = 0.079

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.22.465437; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

Supplementary Figure 23. Illustration of the genomic structural equation models used to test whether correlation magnitudes with genetic general cognitive ability differ between the central executive network and other significantly associated networks. The model on the right freely estimates correlation parameters between two networks and genetic g while allowing for correlations between the networks. In the left model, we force the correlation magnitudes to be the same, and assess whether model fit deteriorates significantly, to conclude whether correlation magnitudes between networks are likely different from each other.

Supplementary Figure 24. Structural equation models to calculate Q_{trait} heterogeneity indices

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.22.465437; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

References

1. Cheverud JM. (1988). A Comparison of Genetic and Phenotypic Correlations. Evolution, *42*(5), 958-68. doi:10.2307/2408911
2. Sodini SM, Kemper KE, Wray NR, Trzaskowski M. (2018). Comparison of Genotypic and Phenotypic Correlations: Cheverud's Conjecture in Humans. Genetics, *209*(3), 941-8. doi:10.1534/genetics.117.300630
3. Power Jonathan D, Cohen Alexander L, Nelson Steven M, Wig Gagan S, Barnes Kelly A, Church Jessica A, et al. (2011). Functional Network Organization of the Human Brain. Neuron, *72*(4), 665-78. doi:10.1016/j.neuron.2011.09.006
4. Smith SM, Fox PT, Miller KL, Glahn DC, Fox PM, Mackay CE, et al. (2009). Correspondence of the brain's functional architecture during activation and rest. Proceedings of the National Academy of Sciences, *106*(31), 13040-5. doi:10.1073/pnas.0905267106
5. Uddin LQ, Yeo BTT, Spreng RN. (2019). Towards a Universal Taxonomy of Macro-scale Functional Human Brain Networks. Brain Topography, *32*(6), 926-42. doi:10.1007/s10548-019-00744-6
6. Madole JW, Ritchie SJ, Cox SR, Buchanan CR, Hernández MV, Maniega SM, et al. (2021). Aging-Sensitive Networks Within the Human Structural Connectome Are Implicated in Late-Life Cognitive Declines. Biological Psychiatry, *89*(8), 795-806. doi:10.1016/j.biopsych.2020.06.010
7. Bressler SL, Menon V. (2010). Large-scale brain networks in cognition: emerging methods and principles. Trends in Cognitive Sciences, *14*(6), 277-90. doi:10.1016/j.tics.2010.04.004
8. Jung RE, Haier RJ. (2007). The Parieto-Frontal Integration Theory (P-FIT) of intelligence: Converging neuroimaging evidence. Behavioral and Brain Sciences, *30*(2), 135-54. doi:10.1017/S0140525X07001185
9. Menon V, Uddin LQ. (2010). Saliency, switching, attention and control: a network model of insula function. Brain Structure and Function, *214*(5), 655-67. doi:10.1007/s00429-010-0262-0
10. Lu J, Liu H, Zhang M, Wang D, Cao Y, Ma Q, et al. (2011). Focal Pontine Lesions Provide Evidence That Intrinsic Functional Connectivity Reflects Polysynaptic Anatomical Pathways. The Journal of Neuroscience, *31*(42), 15065-71. doi:10.1523/jneurosci.2364-11.2011
11. van den Heuvel MP, Mandl RCW, Kahn RS, Hulshoff Pol HE. (2009). Functionally linked resting-state networks reflect the underlying structural connectivity architecture of the human brain. Human Brain Mapping, *30*(10), 3127-41. doi:10.1002/hbm.20737
12. Hermundstad AM, Bassett DS, Brown KS, Aminoff EM, Clewett D, Freeman S, et al. (2013). Structural foundations of resting-state and task-based functional connectivity in the human brain. Proceedings of the National Academy of Sciences, *110*(15), 6169-74. doi:10.1073/pnas.1219562110
13. Xia M, Wang J, He Y. (2013). BrainNet Viewer: A Network Visualization Tool for Human Brain Connectomics. PLOS ONE, *8*(7), e68910. doi:10.1371/journal.pone.0068910
14. Sridharan D, Levitin DJ, Menon V. (2008). A critical role for the right fronto-insular cortex in switching between central-executive and default-mode networks. Proceedings of the National Academy of Sciences, *105*(34), 12569-74. doi:10.1073/pnas.0800005105
15. Buckner RL, DiNicola LM. (2019). The brain's default network: updated anatomy, physiology and evolving insights. Nature Reviews Neuroscience, *20*(10), 593-608. doi:10.1038/s41583-019-0212-7
16. Downar J, Crawley AP, Mikulis DJ, Davis KD. (2002). A Cortical Network Sensitive to Stimulus Salience in a Neutral Behavioral Context Across Multiple Sensory Modalities. Journal of Neurophysiology, *87*(1), 615-20. doi:10.1152/jn.00636.2001
17. Li R, Zhang S, Yin S, Ren W, He R, Li J. (2018). The fronto-insular cortex causally mediates the default-mode and central-executive networks to contribute to individual cognitive performance in healthy elderly. Hum Brain Mapp, *39*(11), 4302-11. doi:10.1002/hbm.24247

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.22.465437; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

18. Duncan J. (2010). The multiple-demand (MD) system of the primate brain: mental programs for intelligent behaviour. Trends in Cognitive Sciences, *14*(4), 172-9. doi:10.1016/j.tics.2010.01.004

19. Zhao B, Ibrahim JG, Li Y, Li T, Wang Y, Shan Y, et al. (2019). Heritability of Regional Brain Volumes in Large-Scale Neuroimaging and Genetic Studies. Cerebral Cortex, *29*(7), 2904-14. doi:10.1093/cercor/bhy157

20. Anderson KM, Ge T, Kong R, Patrick LM, Spreng RN, Sabuncu MR, et al. (2020). Heritability of individualized cortical network topography. bioRxiv, 2020.07.30.229427. doi:10.1101/2020.07.30.229427

21. de Vlaming R, Slob EAW, Jansen PR, Dagher A, Koellinger PD, Groenen PJF, et al. (2021). Multivariate analysis reveals shared genetic architecture of brain morphology and human behavior. bioRxiv, 2021.04.19.440478. doi:10.1101/2021.04.19.440478

22. Elliott ML, Knodt AR, Ireland D, Morris ML, Poulton R, Ramrakha S, et al. (2020). What Is the Test-Retest Reliability of Common Task-Functional MRI Measures? New Empirical Evidence and a Meta-Analysis. Psychological Science, *31*(7), 792-806. doi:10.1177/0956797620916786

23. Cox SR, Ritchie SJ, Fawns-Ritchie C, Tucker-Drob EM, Deary IJ. (2019). Structural brain imaging correlates of general intelligence in UK Biobank. Intelligence, *76*, 101376. doi:10.1016/j.intell.2019.101376

24. Cole JH, Ritchie SJ, Bastin ME, Valdés Hernández MC, Muñoz Maniega S, Royle N, et al. (2018). Brain age predicts mortality. Molecular Psychiatry, *23*(5), 1385-92. doi:10.1038/mp.2017.62

25. Cole JH, Franke K. (2017). Predicting Age Using Neuroimaging: Innovative Brain Ageing Biomarkers. Trends Neurosci, *40*(12), 681-90. doi:10.1016/j.tins.2017.10.001

26. Bulik-Sullivan B, Finucane HK, Anttila V, Gusev A, Day FR, Loh P-R, et al. (2015). An atlas of genetic correlations across human diseases and traits. Nature genetics, *47*(11), 1236. doi:10.1038/ng.3406

27. Grotzinger AD, Rhemtulla M, de Vlaming R, Ritchie SJ, Mallard TT, Hill WD, et al. (2019). Genomic structural equation modelling provides insights into the multivariate genetic architecture of complex traits. Nature Human Behaviour, *3*(5), 513-25. doi:10.1038/s41562-019-0566-x

28. Linnér RK, Biroli P, Kong E, Meddens SFW, Wedow R, Fontana MA, et al. (2019). Genome-wide association analyses of risk tolerance and risky behaviors in over 1 million individuals identify hundreds of loci and shared genetic influences. Nature Genetics, *51*(2), 245-57. doi:10.1038/s41588-018-0309-3

29. Baselmans BML, Jansen R, Ip HF, van Dongen J, Abdellaoui A, van de Weijer MP, et al. (2019). Multivariate genome-wide analyses of the well-being spectrum. Nature Genetics, *51*(3), 445-51. doi:10.1038/s41588-018-0320-8

30. Allen M, Poggiali D, Whitaker K, Marshall T, Kievit R. (2019). Raincloud plots: a multi-platform tool for robust data visualization [version 1; peer review: 2 approved]. Wellcome Open Research, *4*(63). doi:10.12688/wellcomeopenres.15191.1

31. Lorenzo-Seva U, Berge JMFt. (2006). Tucker's Congruence Coefficient as a Meaningful Index of Factor Similarity. Methodology, *2*(2), 57-64. doi:10.1027/1614-2241.2.2.57

32. Hu L-t, Bentler PM. (1998). Fit indices in covariance structure modeling: Sensitivity to underparameterized model misspecification. Psychological methods, *3*(4), 424. doi:10.1037/1082-989X.3.4.424

33. Grotzinger AD, Mallard TT, Akingbuwa WA, Ip HF, Adams MJ, Lewis CM, et al. (2020). Genetic Architecture of 11 Major Psychiatric Disorders at Biobehavioral, Functional Genomic, and Molecular Genetic Levels of Analysis. medRxiv, 2020.09.22.20196089. doi:10.1101/2020.09.22.20196089

34. Fawns-Ritchie C, Deary IJ. (2020). Reliability and validity of the UK Biobank cognitive tests. PLOS ONE, *15*(4), e0231627. doi:10.1371/journal.pone.0231627

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.22.465437; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

35. Kaufmann T, van der Meer D, Doan NT, Schwarz E, Lund MJ, Agartz I, et al. (2019). Common brain disorders are associated with heritable patterns of apparent aging of the brain. Nature Neuroscience, *22*(10), 1617-23. doi:10.1038/s41593-019-0471-7
36. Jansen PR, Nagel M, Watanabe K, Wei Y, Savage JE, de Leeuw CA, et al. (2020). Genome-wide meta-analysis of brain volume identifies genomic loci and genes shared with intelligence. Nature Communications, *11*(1), 5606. doi:10.1038/s41467-020-19378-5
37. Herculano-Houzel S. (2017). Numbers of neurons as biological correlates of cognitive capability. Current Opinion in Behavioral Sciences, *16*, 1-7. doi:10.1016/j.cobeha.2017.02.004
38. Fry A, Littlejohns TJ, Sudlow C, Doherty N, Adamska L, Sprosen T, et al. (2017). Comparison of Sociodemographic and Health-Related Characteristics of UK Biobank Participants With Those of the General Population. American Journal of Epidemiology, *186*(9), 1026-34. doi:10.1093/aje/kwx246
39. Alfaro-Almagro F, Jenkinson M, Bangerter NK, Andersson JLR, Griffanti L, Douaud G, et al. (2018). Image processing and Quality Control for the first 10,000 brain imaging datasets from UK Biobank. NeuroImage, *166*, 400-24. doi:10.1016/j.neuroimage.2017.10.034
40. Smith S, Alfaro-Almagro F, Miller K. UK Biobank Brain Imaging Documentation. 2020 [Available from: https://biobank.ctsu.ox.ac.uk/crystal/crystal/docs/brain_mri.pdf.
41. Desikan RS, Ségonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, et al. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. Neuroimage, *31*(3), 968-80. doi:10.1016/j.neuroimage.2006.01.021
42. Fischl B, Salat DH, Busa E, Albert M, Dieterich M, Haselgrove C, et al. (2002). Whole Brain Segmentation: Automated Labeling of Neuroanatomical Structures in the Human Brain. Neuron, *33*(3), 341-55. doi:10.1016/S0896-6273(02)00569-X
43. Lindroth H, Nair VA, Stanfield C, Casey C, Mohanty R, Wayer D, et al. (2019). Examining the identification of age-related atrophy between T1 and T1 + T2-FLAIR cortical thickness measurements. Scientific Reports, *9*(1), 11288. doi:10.1038/s41598-019-47294-2
44. Bycroft C, Freeman C, Petkova D, Band G, Elliott LT, Sharp K, et al. (2018). The UK Biobank resource with deep phenotyping and genomic data. Nature, *562*(7726), 203-9. doi:10.1038/s41586-018-0579-z
45. Coleman JRI, Peyrot WJ, Purves KL, Davis KAS, Rayner C, Choi SW, et al. (2020). Genome-wide gene-environment analyses of major depressive disorder and reported lifetime traumatic experiences in UK Biobank. Molecular Psychiatry, *25*(7), 1430-46. doi:10.1038/s41380-019-0546-6
46. Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. GigaScience, *4*(1). doi:10.1186/s13742-015-0047-8
47. Warren HR, Evangelou E, Cabrera CP, Gao H, Ren M, Mifsud B, et al. (2017). Genome-wide association analysis identifies novel blood pressure loci and offers biological insights into cardiovascular risk. Nature Genetics, *49*(3), 403-15. doi:10.1038/ng.3768
48. Manichaikul A, Mychaleckyj JC, Rich SS, Daly K, Sale M, Chen W-M. (2010). Robust relationship inference in genome-wide association studies. Bioinformatics, *26*(22), 2867-73. doi:10.1093/bioinformatics/btq559
49. McCarthy S, Das S, Kretzschmar W, Delaneau O, Wood AR, Teumer A, et al. (2016). A reference panel of 64,976 haplotypes for genotype imputation. Nature Genetics, *48*(10), 1279-83. doi:10.1038/ng.3643
50. de la Fuente J, Davies G, Grotzinger AD, Tucker-Drob EM, Deary IJ. (2021). A general dimension of genetic sharing across diverse cognitive traits inferred from molecular data. Nature Human Behaviour, *5*(1), 49-58. doi:10.1038/s41562-020-00936-2
51. Mbatchou J, Barnard L, Backman J, Marcketta A, Kosmicki JA, Ziyatdinov A, et al. (2021). Computationally efficient whole-genome regression for quantitative and binary traits. Nature Genetics. doi:10.1038/s41588-021-00870-7

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.22.465437; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.

52. Smith SM, Douaud G, Chen W, Hanayik T, Alfaro-Almagro F, Sharp K, et al. (2020). Enhanced Brain Imaging Genetics in UK Biobank. bioRxiv, 2020.07.27.223545. doi:10.1101/2020.07.27.223545

53. Tucker LR. A method for synthesis of factor analysis studies. Educational Testing Service Princeton Nj; 1951.

54. Benjamini Y, Hochberg Y. (1995). Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society: Series B (Methodological)*, 57*(1), 289-300. doi:10.1111/j.2517-6161.1995.tb02031.x

bioRxiv preprint doi: https://doi.org/10.1101/2021.10.22.465437; this version posted October 25, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY 4.0 International license.