Multidimensional Heritability Analysis of Neuroanatomical Shape

Tian Ge[a,b,c], Martin Reuter[a,d], Anderson M. Winkler[e], Avram J. Holmes[a,f,g], Phil H. Lee[b,c], Lee S. Tirrell[a], Joshua L. Roffman[g], Randy L. Buckner[a,g,h], Jordan W. Smoller[b,c,*], and Mert R. Sabuncu[a,d,*]

[a]Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital / Harvard Medical School, Charlestown, MA 02129, USA; [b]Psychiatric and Neurodevelopmental Genetics Unit, Center for Human Genetic Research, Massachusetts General Hospital, Boston, MA 02114, USA; [c]Stanley Center for Psychiatric Research, Broad Institute of MIT and Harvard, Cambridge, MA 02138, USA; [d]Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA 02139, USA; [e]Centre for Functional MRI of the Brain (FMRIB), University of Oxford, Oxford OX3 9DU, UK; [f]Department of Psychology, Yale University, New Haven, CT 06520, USA; [g]Department of Psychiatry, Massachusetts General Hospital / Harvard Medical School, Boston, MA 02114, USA; [h]Department of Psychology and Center for Brain Science, Harvard University, Cambridge, MA 02138, USA.

*MRS and JWS contributed equally.

Correspondance to: msabuncu@nmr.mgh.harvard.edu (MRS) or tge1@mgh.harvard.edu (TG)

Abstract

In the dawning era of large-scale biomedical data, multidimensional phenotype vectors will play an increasing role in examining the genetic underpinnings of brain features, behavior and disease. For example, shape measurements derived from brain MRI scans are multidimensional geometric descriptions of brain structure and provide an alternate class of phenotypes that remains largely unexplored in genetic studies. Here we extend the concept of heritability to multidimensional traits, and present the first comprehensive analysis of the heritability of neuroanatomical shape measurements across an ensemble of brain structures based on genome-wide SNP and MRI data from 1,320 unrelated, young and healthy individuals. We replicate our findings in an extended twin sample from the Human Connectome Project (HCP). Our results demonstrate that neuroanatomical shape can be significantly heritable, above and beyond volume, and can serve as a complementary phenotype to study the genetic determinants and clinical relevance of brain structure.

Introduction

The exponential progress in genomic technologies has accelerated the examination of the genetic underpinnings of complex phenotypes, such as psychiatric and neurological disorders, many of which are highly heritable [1, 2]. For example, large-scale genome-wide association studies (GWAS) have provided insights about common genetic variants linked with a range of clinical conditions [3–6]. Most prior genetic studies have focused on univariate (scalar) phenotypes, such as diagnosis or a quantitative measurement. However, with the emergence of large-scale data collection efforts, such as the Human Connectome Project (HCP; http://www.humanconnectome.org) and the UK Biobank (http://www.ukbiobank.ac.uk), each subject can be linked to a high-dimensional phenotype vector, which might include imaging measurements or electronic health record. Such phenotypically rich studies open up the opportunity to analyze collections of multidimensional phenotypes, which can be more informative than scalar traits.

Brain imaging is playing an increasing role in the study of the relationship between genetic variants, neuroanatomy, behavior and disease susceptibility [7–10]. To date, most structural neuroimaging genetics studies have utilized the size, average cortical thickness, or surface area of a brain region to yield important discoveries about the genetic basis of brain morphology [see e.g., 11–14]. While these measurements capture a few basic dimensions of anatomical variability, they provide a limited description of the underlying geometry.

Neuroanatomical shape measurements — multidimensional geometric descriptions of brain structure — have attracted increasing attention in medical image analysis. Shape measurements characterize isometry-invariant (in particular, independent of location and orientation) geometric attributes of an object, which provide a rich description of an anatomical structure and can encompass volumetric variation. Such measurements may thus offer increased sensitivity and specificity in examining the clinical relevance and genetic underpinnings of brain structure. Recent studies have shown that the shape of subcortical brain regions and cortical folding patterns provide information that is not available in volumetric measurements and is predictive of disease status, onset and progression in schizophrenia [15–17], autism [18, 19], bipolar disorder [20, 21], Alzheimer's disease [22–25], and other mental disorders [26, 27]. There is also increasing evidence that genetic variants may have influences on brain morphology that can be captured by shape measurements [28–32].

This paper makes two major contributions to the investigation of the genetic basis of neuroanatomical shape. First, we extend the theoretical concept of heritability to multidimensional traits, such as the shape descriptor of an object, and propose a novel method to estimate the heritability of multidimensional traits

based on genome-wide single nucleotide polymorphism (SNP) data from unrelated individuals (known as SNP heritability). Our estimation method builds on genome-wide complex trait analysis (GCTA) [33, 34] and phenotype correlation-genetic correlation (PCGC) regression [35], and generalizes these techniques to the multivariate setting. Second, using structural magnetic resonance imaging (MRI) and SNP data from 1,320 unrelated individuals collected as part of the Harvard/Massachusetts General Hospital (MGH) Brain Genomics Superstruct Project (GSP) [36, 37], we present the first comprehensive heritability analysis of the shape of an ensemble of brain structures, quantified by the truncated Laplace-Beltrami Spectrum (LBS) (also known as the "Shape-DNA") [38–40], in this young (18-35 years) and healthy cohort, and devise a strategy to visualize primary modes of shape variation. We also replicate our findings in an extended twin sample (72 monozygotic twin pairs, 69 dizygotic twin pairs, 253 full siblings of twins and 55 singletons) from the Human Connectome Project (HCP) [41].

The truncated LBS is a multidimensional shape descriptor, which can be obtained by solving an eigenvalue problem on the 2D boundary surface representation of an object. It is invariant to the representation of the object including parameterization, location and orientation, and thus does not require spatial alignment with a population template, making it computationally efficient and robust to registration errors. LBS also depends continuously on topology-preserving deformations, and is thus suitable to quantify differences between shapes. Recent empirical evidence suggests that the LBS-based shape descriptor provides a discriminative characterization of brain anatomy and offers state-of-the-art performance for a range of shape retrieval and segmentation applications [42, 43]. A collection of the descriptors of brain structures, known as the *BrainPrint*, can provide an accurate and holistic representation of brain morphology, and has been successfully applied to subject identification, sex and age prediction, brain asymmetry analysis, twin studies, and computer-aided diagnosis of dementia [40, 44]. Our LBS-based heritability analyses demonstrate that neuroanatomical shape can be significantly heritable, above and beyond volume, and yield a complementary phenotype that offers a unique perspective in studying the genetic determinants of brain structure.

Results

Heritability of the volume of neuroanatomical structures. To benchmark our shape results, we first computed SNP heritability estimates for the volumetric measurements of an array of brain regions using the GSP sample. Table 1 lists these heritability estimates after adjusting for intracranial volume (ICV or head size) as a covariate. Point estimates of the heritability of volumetric measurements suggested that several neuroanatomi-

cal structures have moderately heritable volumes. In particular, the caudate, corpus callosum, lateral ventricle, 3rd and 4th ventricles, pallidum, putamen and thalamus all had volume heritability estimates greater than 25%. Table 1 further includes p-values for the statistical significance of the heritability estimates. The parametric (Wald) and non-parametric (permutation-based) p-values were virtually identical, confirming the accuracy of the standard error estimates we computed (see Methods). We observe that none of the volume heritability estimates were statistically significant after correcting for multiple comparisons (false discovery rate or FDR at $q = 0.05$), possibly due to sample size limitations. Only the volumes of the caudate, corpus callosum and 3rd ventricle achieved a heritability that was nominally significant in our sample (uncorrected $p < 0.05$). Table 1 also includes test-retest reliability estimates of volume after regressing out ICV, computed as Lin's concordance correlation coefficient [45] using measurements from 42 subjects with repeated scans on separate days. Almost all the structures had a volume estimate reliability greater than 0.75 except for the pallidum. There was no significant correlation between the reliability and heritability estimates of volume ($p = 0.828$).

Heritability of the shape of neuroanatomical structures. Neuroanatomical shape measurements provide a geometric characterization and a rich description of brain structure. We therefore hypothesize that analyzing the shape variation of neuroanatomical structures can identify genetic influences beyond captured by volumetric measurements. Fig. 1 and Table 2 show the SNP heritability estimates of the shape of an ensemble of brain structures using the GSP sample. These estimates were computed based on LBS descriptors normalized for size and explicitly including the volume of the corresponding structure as a covariate in the analysis to account for potential volume effects. A number of structures showed moderate to high SNP heritability. Specifically, the shape of the caudate, cerebellum, corpus callosum, hippocampus, 3rd ventricle and putamen exhibited heritability estimates greater than 25%. All these estimates were statistically significant after correcting for an FDR at $q = 0.05$. We observe that this is in contrast with the case of volume, where despite a similar heritability range, no estimate reached FDR-corrected significance. The main reason for this discrepancy is the theoretically guaranteed reduced standard errors in SNP heritability estimates of multidimensional traits (see Methods for a theoretical treatment). The shape of the accumbens area was also marginally significantly heritable with an uncorrected p-value less than 0.05. As in the case of volume, the parametric (Wald) p-values were virtually identical to the permutation p-values, suggesting that our standard error estimates are accurate (see Methods).

Table 2 also lists test-retest reliability estimates for the shape of different structures. Analogous to the case of volume, we quantified reliability as the average Lin's concordance correlation coefficient of individual com-

Structure	$\widehat{h}^2_{\mathrm{SNP}}$	SE	Wald p-value	Perm p-value	Reliability
Accumbens Area	0.001	0.281	0.500	1.000	0.797
Amygdala	0.141	0.281	0.308	0.305	0.864
Caudate	**0.657**	0.281	0.010	0.009	0.947
Cerebellum	0.084	0.281	0.383	0.382	0.989
Corpus Callosum	**0.538**	0.281	0.028	0.029	0.882
Hippocampus	0.005	0.281	0.493	0.492	0.939
Lateral Ventricle	0.331	0.281	0.119	0.120	0.995
3rd Ventricle	**0.500**	0.281	0.038	0.040	0.832
4th Ventricle	0.381	0.281	0.087	0.089	0.986
Pallidum	0.300	0.281	0.142	0.142	0.642
Putamen	0.328	0.281	0.121	0.122	0.934
Thalamus	0.252	0.281	0.184	0.186	0.867

Table 1: **SNP heritability estimates $\widehat{h}^2_{\mathrm{SNP}}$ of the volume of brain structures using the GSP sample.** The standard errors (SE) were computed using an approximation, which, given the empirical genetic similarity matrix, only depends on the sample size. p-values were obtained by the Wald test and the permutation inference (based on 10,000 permutations), respectively. The strong agreement between the parametric and non-parametric p-values indicates that the estimated SE values are accurate. Estimates with uncorrected significant p-values (< 0.05) are shown in **bold**. Test-retest reliability of the volumetric measurements was computed as Lin's concordance correlation coefficient using measurements from 42 subjects with repeated scans on separate days.

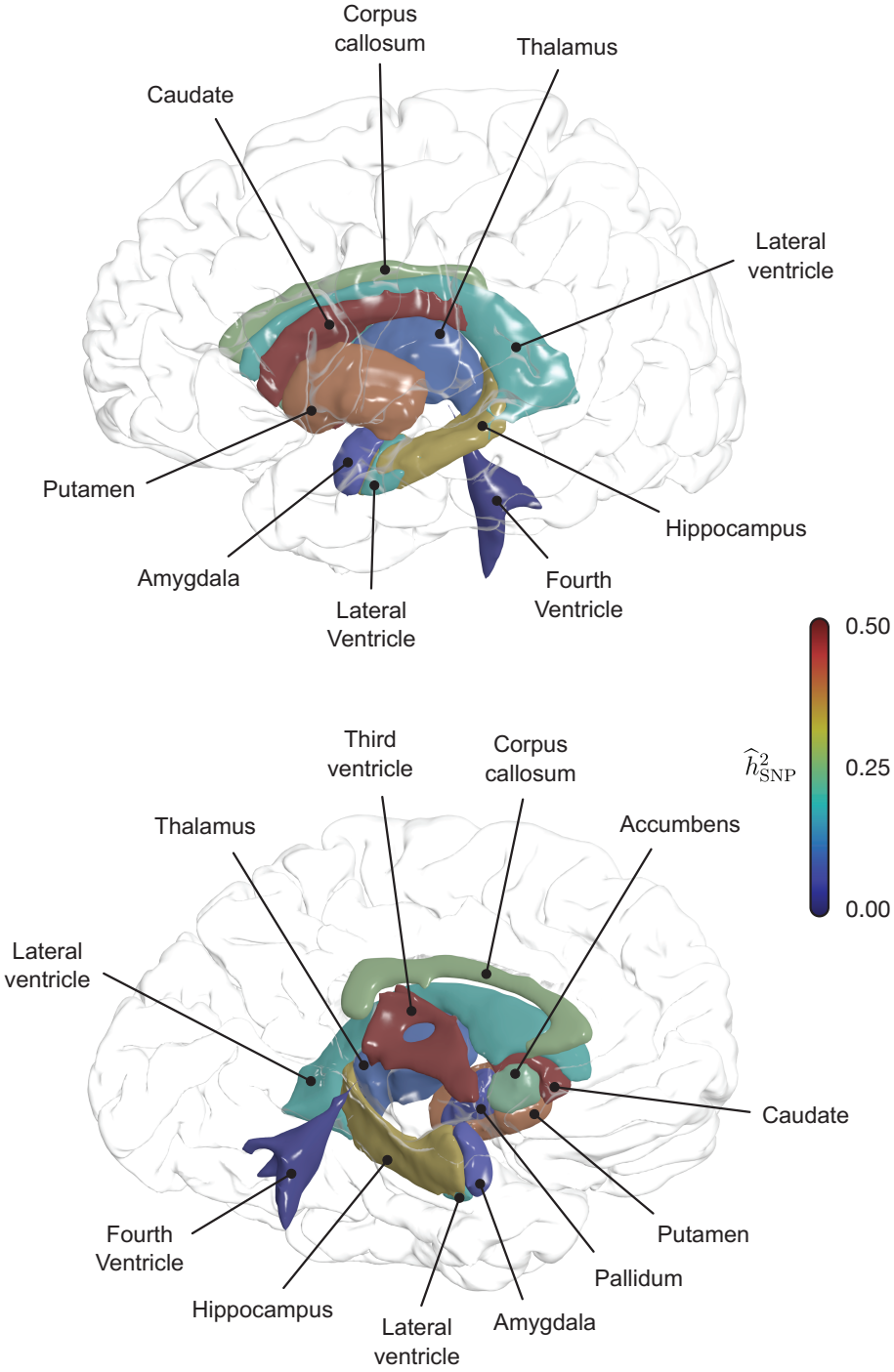

Figure 1: **SNP heritability estimates of the shape of brain structures in the GSP sample.** Top: lateral view. Bottom: medial cross-section.

Structure	$\widehat{h}^2_{\text{SNP}}$	SE	Wald p-value	Perm p-value	Reliability
Accumbens Area	**0.237**	0.135	0.039	0.039	0.418
Amygdala	0.061	0.139	0.330	0.327	0.670
Caudate	**0.499**	0.188	0.004	0.005	0.759
Cerebellum	**0.452**	0.192	0.009	0.009	0.844
Corpus Callosum	**0.264**	0.133	0.023	0.022	0.622
Hippocampus	**0.347**	0.169	0.020	0.019	0.866
Lateral Ventricle	0.190	0.153	0.107	0.105	0.890
3rd Ventricle	**0.500**	0.157	0.001	0.001	0.761
4th Ventricle	0.005	0.208	0.490	0.491	0.633
Pallidum	0.061	0.117	0.299	0.299	0.402
Putamen	**0.413**	0.148	0.003	0.003	0.781
Thalamus Proper	0.086	0.143	0.274	0.276	0.552

Table 2: **SNP heritability estimates $\widehat{h}^2_{\text{SNP}}$ of the shape of brain structures using the GSP sample.** Standard errors (SE) are less than those corresponding to volume heritability. p-values were obtained by the Wald test and the permutation inference (based on 10,000 permutations), respectively. The strong agreement between the parametric and non-parametric p-values indicates that the SE estimates are accurate. Estimates with uncorrected significant p-values (< 0.05) are shown in **bold**. False discovery rate (FDR) corrected significant p-values (< 0.05) are shown in **red**. Test-retest reliability of the shape measurements were computed as the average Lin's concordance correlation coefficient of individual components of the LBS-based shape descriptor from 42 subjects with repeated scans on separate days.

ponents of the multidimensional shape descriptor from 42 subjects with repeat scans on separate days. These results suggest that the LBS-based shape descriptors were overall less reliable than volumetric measurements, with half of the structures exhibiting a shape reliability less than 0.75. This is likely due to the increased sensitivity of shape to segmentation differences relative to the volume. Furthermore, there was a marginally significant correlation between reliability and heritability of shape (Pearson's $r = 0.562$ and $p = 0.057$). We conclude that close to 30% of the variation in shape heritability across structures can be attributed to the reliability of the shape descriptor. This suggests that for structures that exhibited low shape heritability (e.g., amygdala), a more accurate image segmentation and shape analysis pipeline might yield an increased estimate of heritability. We further conducted a sensitivity analysis of shape heritability estimates, with respect to the two free parameters of the LBS-based shape descriptor: number of eigenvalues incorporated and amount of smoothing applied to the surface mesh representing the geometry of the object. Supplementary Fig. 1 shows that the heritability estimates were largely robust to variations in these parameters.

We sought to replicate these findings in the HCP sample, in which subjects are healthy and have a similar age range as the GSP sample. We selected 590 non-Hispanic/Latino Europeans aged between 22 and 35, comprising 72 monozygotic (MZ) twin pairs, 69 dizygotic (DZ) twin pairs, 253 full siblings of twins and 55 singletons (single birth individuals without siblings). We estimated the shape heritability for brain structures that had significantly heritable shapes in the GSP sample using an ACE model (A: additive genetics; C: common environment; E: unique or subject-specific environment), where the additive genetic similarity was derived from pedigree information and the common environment term reflected household sharing between subjects. We also obtained the standard error of the shape heritability estimates using a block bootstrapping procedure (see Methods). All the shapes we analyzed were significantly heritable in the HCP sample: accumbens area 0.309 ± 0.081; caudate 0.583 ± 0.062; cerebellum 0.653 ± 0.060; corpus callosum 0.558 ± 0.068; hippocampus 0.363 ± 0.095; 3rd ventricle 0.536 ± 0.067; putamen 0.483 ± 0.106. We also observe that the HCP shape heritability estimates were consistently larger than the GSP estimates, which is theoretically expected because the SNP heritability estimated from unrelated GSP subjects only captured the genetic variation tagged by common SNPs in the data set, and is thus a lower bound for the classical narrow-sense heritability estimated from familial data such as the HCP sample.

Visualizing the principal mode of shape variation. The LBS-based shape descriptor is suitable to efficiently and accurately extract intrinsic properties of the shape of brain structures from a large number of individuals, but is not designed to visually inspect shape differences. Here, we propose a strategy to visualize the principal

mode of shape variation. Specifically, it can be shown that the first principal component (PC) of the multi-dimensional LBS-based shape descriptor captures the greatest shape variation and has the largest impact on the overall heritability estimate of the shape (see Methods). We thus visualized shape variation along the first PC of the shape descriptor for brain structures with significantly heritable shapes in the GSP sample: right caudate, cerebellum, corpus callosum, right hippocampus, 3rd ventricle, and left putamen. The illustrations of contralateral structures (i.e., left caudate, left hippocampus and right putamen), which showed similar shape variation, are provided in Supplementary Fig. 2. In each panel of Fig. 2, the structure is represented with a sample-specific population average, on which average shapes at the two extremes (± 2 standard deviation or SD) of the principal axis with identical volume (-2 SD, blue; $+2$ SD, red) are depicted. Blue regions indicate where shapes around the -2 SD are larger than shapes around the $+2$ SD, and vice versa for the red regions.

The first PC of the right caudate captured 77% of the shape variation and had a SNP heritability estimate of 0.88. Moving along the principal mode of shape variation, the right caudate had a larger (smaller) head with a corresponding shorter (longer) tail. For the cerebellum, the first PC explained 69% of the shape variation and had a SNP heritability of 0.61. A clear expansion (contraction) of the anterior lobe and a corresponding contraction (expansion) of the posterior lobe can be observed along the principal axis. The first PC of the corpus callosum captured 41% of the shape variation and had a SNP heritability of 0.41. The principal mode captured an expansion (contraction) of the middle corpus callosum along the dorsoventral axis and a corresponding shrinking (enlargement) of the anterior and posterior part of the structure. For the right hippocampus, the first PC explained 69% of the shape variation, had a SNP heritability of 0.47, and exhibited dorsoventral widening (narrowing) of the body and corresponding lateral and anterior-posterior contraction (expansion). The first PC of the 3rd ventricle captured 69% of the shape variation and had a SNP heritability estimate of 0.80. The principal mode captured an enlargement (shrinking) of the posterior protrusions and an expansion (contraction) of the lateral walls, coupled with a corresponding contraction (expansion) of the roof of the cleft. Finally, first PC of the left putamen explained 61% of the shape variation, had a SNP heritability of 0.70, and captured lateral widening (narrowing) and a corresponding contraction (expansion) along the dorsoventral and anterior-posterior axes.

Discussion

This work makes two contributions to neuroscience and genetic research. First, we extend the concept of heritability to multidimensional traits and present an analytic strategy that generalizes SNP heritability analysis.

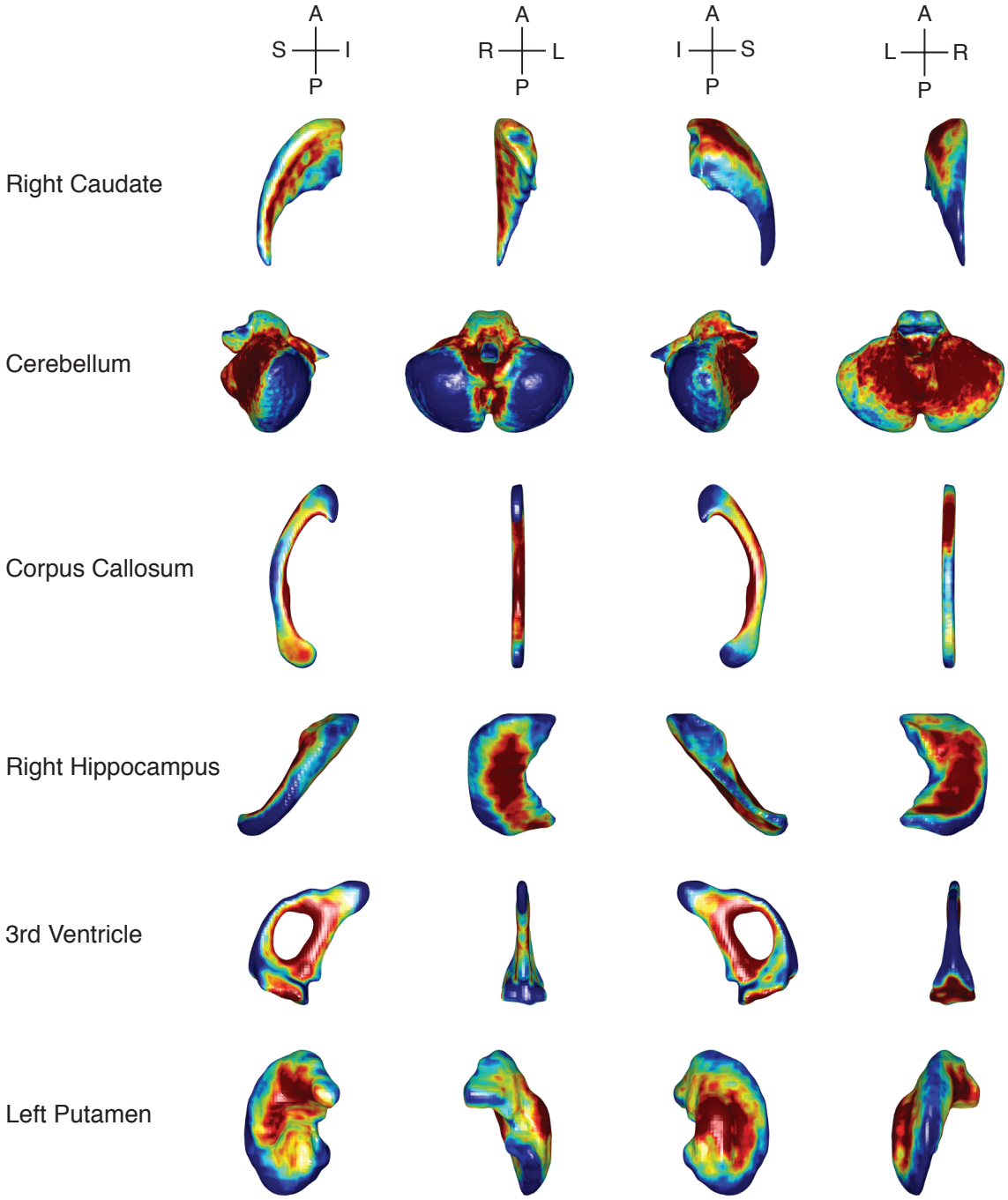

Figure 2: **The principal mode of shape variation for brain structures with significantly heritable shape in the GSP sample.** Each structure is represented with a sample-specific population average, on which average shapes at the two extremes (± 2 standard deviation or SD) along the first principal component (PC) of the shape descriptor (-2 SD, blue; $+2$ SD, red) are depicted. Anatomical orientation is indicated with embedded coordinate axes. I: Inferior, S: Superior, A: Anterior, P: Posterior, L: Left, R: Right.

The heritability estimator we propose for multidimensional traits has reduced uncertainty in its point estimate relative to univariate estimates, and thus offers more statistical power. Our empirical analyses confirmed this theoretical expectation. Moreover, and in the same line, we provide methods that can easily adjust for covariates in multivariate models, and also both parametric and nonparametric inferential tools that can assess the significance of a heritability estimate. Our approach opens the door to the genetic characterization of shape measurements and other multidimensional traits.

Second, we use the proposed approach to quantify the SNP heritability of the shape of an ensemble of anatomical structures spanning the human brain in a group of young and healthy subjects. The shape of caudate, cerebellum, corpus callosum, hippocampus, 3rd ventricle and putamen exhibited moderate to high heritability (i.e., greater than 25%), after controlling for volume. All of these estimates achieved FDR-corrected significance at $q = 0.05$. This is in contrast to the volume heritability estimates of the same set of brain structures on the same sample, none of which reached FDR-corrected significance. Although our heritability analysis of volume, which was used to benchmark the shape analysis, may be less informative compared to more powerful twin studies and large-scale meta-analyses in the literature, the increased statistical power and the additional information shape analysis can provide relative to volumetric analysis demonstrate the usefulness of our methods and underscore the potential of leveraging multidimensional traits when analyzing data sets with moderate sample sizes.

Using the extended twin data from the HCP, we also replicated significant shape heritability estimates observed in the GSP sample. Our HCP estimates were consistently larger than the GSP estimates, which is theoretically expected because SNP heritability estimated from the unrelated sample in GSP does not capture genetic contributions (e.g., from rare variants) that are not tagged by genotyped SNPs, and thus explains a smaller proportion of the phenotypic variation. However, additional factors may contribute to the difference between SNP and familial heritability estimates, which include improper modeling of shared environment, assortative mating, genetic interaction (epistasis), suboptimal statistical methods for heritability estimation, and differences in sample characteristics such as age range, ethnic background and environmental exposures [35, 46–51]. Dissecting the discrepancy in heritability estimates from familial and unrelated data is an area under active investigation. More systematic future work is required to fully disentangle this problem.

A handful of prior neuroimaging studies have explored the shape of certain brain structures as potential phenotypes in examining genetic associations. For example, Qiu et al. [28] and Shi et al. [29] reported influences of the apolipoprotein E (APOE) $\varepsilon 4$ allele on hippocampal morphology in depressive and Alzheimer's disease patients. Variants involved in the regulation of the *FKBP5* gene were recently associated with hip-

pocampal shape [30]. A meta-study [32] identified a GWAS significant SNP that exerts its effect on the shape of putamen bilaterally. Prior studies have also estimated heritability of shape based on familial relatedness. In a recent study, the heritability of the shape of subcortical and limbic structures was estimated using data from multigenerational families with schizophrenia [31]. In other related work, Mamah et al. [52] and Harms et al. [53] revealed shape abnormalities in basal ganglia structures (caudate, putamen and globus pallidus) and the thalamus in siblings of schizophrenia patients. An application of the LBS-based shape descriptor to twin data found increased shape similarity of brain structures in MZ twin pairs relative to DZ twins, indicating genetic influences on brain morphology [40], although heritability was not estimated.

However, to date, outside of these notable exceptions, most structural imaging genetics studies have utilized scalar measurements (e.g., volume, thickness, area) as phenotypes. In the present study, we accounted for potential volume effects in our shape analyses by normalizing the LBS-based shape descriptor for size and additionally including the volumetric measurement of the corresponding structure as a covariate when estimating heritability. Our results show that shape measurements provide a rich and novel set of phenotypes for exploring the genetic basis of brain structure, and may identify novel genetic influences on the brain that are not detectable with conventional analyses based on the volume of structures.

There are several biological mechanisms that might lead to shape differences with minimal effect on the overall size of the structure. These include localized volumetric effects that are confined to subfields, subnuclei or other sub-regions that make up the structure. Shape analysis may provide significant information about neurodevelopmental abnormalities, such as those associated with neuronal migration, synaptogenesis, synaptic pruning and myelination. Shape measurements might for example shed light on morphogenetic mechanisms that involve mechanical tensions along axons, dendrites and glial processes [54]. Thus, shape measurements are particularly promising phenotypes for studying neurodevelopmental disorders. Neurodegenerative processes and other pathologies, many of which are known to be genetically influenced, can also impact neuroanatomical shape by exerting focal and/or selective insults. For example, in Alzheimer's disease, morphological alterations in the hippocampus may only target certain subfields [55].

The shape analysis literature offers an expanding list of methods to quantify and characterize shape [43]. A major advantage of the LBS-based shape descriptor [38] employed in this study is that it is robust to intensity variation across scans and does not require the nonlinear spatial registration of the object with a population template, which can be computationally demanding and prone to error. In this paper, we also present a novel strategy to visualize the principal mode of shape variation across the population. For brain structures with significantly heritable shapes, we demonstrated that the principal mode explains a large portion of the

overall shape variation and is often highly heritable. This approach can thus shed light on the global genetic influences on brain structures, and is complementary to studies that rely on nonlinear group-wise registration to characterize localized genetic influences on shape variation.

In the present study, in light of the similar shape variation of bilateral brain structures as observed in Fig. 2 and Supplementary Fig. 2, and to increase signal-to-noise ratio and statistical power, we combined the left and right structures in our shape heritability analysis. SNP heritability estimates of the shape of bilateral brain structures using the GSP sample are provided in Supplementary Table 1. The results indicate that the genetic influences on several anatomical structures (e.g., caudate) may be lateralized, although with the current sample size we are not able to claim that the lateral difference is significant. It will be interesting to investigate the laterality of brain structures from the genetic perspective when we have a better understanding of the genetic basis of brain morphology and when a data set with larger sample size becomes available.

The heritability analysis of multidimensional traits developed here can be applied to phenotypes other than shape that are intrinsically multivariate. Another application might involve heritability or genetic association analyses combining related traits to obtain more stable effect estimates. For example, it can be used as an alternative to principal component analysis (PCA) and factor analysis when investigating the genetic basis of various psychometric or behavioral traits. Also, voxel- or vertex-level neuroimaging measurements are often noisy, and analyzing these measurements in homogeneous brain regions in a multivariate fashion may increase the reliability and reproducibility of the results. Finally, the empirical genetic similarity matrix can be computed with other SNP grouping strategies (e.g., based on genes, pathways, functional annotations and previous GWAS findings) to model the genetic influences from a specific genomic region or partition the heritability of multidimensional traits, as in Yang et al. [56].

Methods

Variance component models. We start with a brief review of variance component models (also known as random effects models) for the heritability analysis of univariate (scalar) traits, which provide a general statistical framework that can handle both familial designs and unrelated individuals randomly sampled from the population. Assuming, for the moment, no covariate needs to be adjusted, and a trait can be partitioned into the sum of additive genetic effect $\boldsymbol{g}$, common (or shared) environment $\boldsymbol{c}$, and unique (subject-specific) environment $\boldsymbol{e}$, the variance component model takes the following form:

$$\boldsymbol{y} = \boldsymbol{g} + \boldsymbol{c} + \boldsymbol{e}, \quad \boldsymbol{g} \sim \mathrm{N}(\boldsymbol{0}, \sigma_A^2 \boldsymbol{K}), \quad \boldsymbol{c} \sim \mathrm{N}(\boldsymbol{0}, \sigma_C^2 \boldsymbol{\Lambda}), \quad \boldsymbol{e} \sim \mathrm{N}(\boldsymbol{0}, \sigma_E^2 \boldsymbol{I}), \tag{1}$$

where $\boldsymbol{y} = [y_1, \cdots, y_N]^\top$ is a vector comprising quantitative traits from N individuals, σ_A^2, σ_C^2 and σ_E^2 are the additive genetic variance, common environmental variance and unique environmental variance, respectively, $\boldsymbol{K}$ is a genetic similarity matrix, $\boldsymbol{\Lambda}$ quantifies shared environment between pairs of individuals, and $\boldsymbol{I}$ is an identity matrix.

In familial studies, $\boldsymbol{K}$ is twice the kinship matrix, $\boldsymbol{K} = 2\boldsymbol{\Phi}$, and indicates expected additive genetic covariance among individuals. The ij-th entry of the kinship matrix, ϕ_{ij}, known as the kinship coefficient, defines genetic relatedness for subjects i and j, and in general can be derived from pedigree information [57, 58]. $\boldsymbol{\Lambda}$ is a matrix that usually reflects household sharing between pairs of individuals. For example, in the present study, $\phi_{ij} = 1/2$ for MZ twins and $\phi_{ij} = 1/4$ for DZ twins and full siblings, and we assume that twin pairs and their non-twin siblings share the same environment and the corresponding elements in $\boldsymbol{\Lambda}$ are 1.

When modeling unrelated individuals randomly sampled from the population, $\boldsymbol{K}$ is the empirical genetic similarity matrix for each pair of individuals estimated from genome-wide SNP data, and the corresponding variance component parameter σ_A^2 is the total additive genetic variance tagged by common SNPs spanning the genome. We note that in unrelated subject studies σ_A^2 does not capture contributions (e.g., from rare variants) that are not assayed by the genotyping microarray, and thus needs to be interpreted differently from σ_A^2 in familial studies, although we use the same notation here for simplicity. In addition, the common environmental matrix $\boldsymbol{\Lambda}$ is often assumed to vanish when analyzing unrelated individuals, in which case Eq. (1) becomes the classical model used in genome-wide complex trait analysis (GCTA) [33, 34].

The heritability of a univariate (scalar) trait is defined as

$$h^2 = \frac{\sigma_A^2}{\sigma_P^2} := \frac{\sigma_A^2}{\sigma_A^2 + \sigma_C^2 + \sigma_E^2}, \tag{2}$$

where σ_P^2 is the phenotypic variance. In familial studies, h^2 measures the narrow-sense heritability of a trait, while in unrelated subject studies, h^2 measures additive heritability attributable to common genetic variants (known as SNP heritability and often denoted as h^2_{SNP}), and provides a lower bound for the narrow-sense heritability estimated by familial studies.

Heritability of multidimensional traits. We now consider an M-dimensional trait $\boldsymbol{Y} = [\boldsymbol{y}_1, \cdots, \boldsymbol{y}_M] = [y_{im}]_{N \times M}$. We model $\boldsymbol{Y}$ by a multivariate variance component model:

$$\boldsymbol{Y} = \boldsymbol{G} + \boldsymbol{C} + \boldsymbol{E}, \tag{3}$$

where $\boldsymbol{G}$, $\boldsymbol{C}$ and $\boldsymbol{E}$ are $N \times M$ matrices, and represent additive genetic effects, common environmental factors

and unique environmental factors, respectively. We have the following distributional assumptions:

$$\mathbf{vec}(\boldsymbol{G}) \sim \mathrm{N}(\mathbf{0}, \boldsymbol{\Sigma}_A \otimes \boldsymbol{K}), \quad \mathbf{vec}(\boldsymbol{C}) \sim \mathrm{N}(\mathbf{0}, \boldsymbol{\Sigma}_C \otimes \boldsymbol{\Lambda}), \quad \mathbf{vec}(\boldsymbol{E}) \sim \mathrm{N}(\mathbf{0}, \boldsymbol{\Sigma}_E \otimes \boldsymbol{I}), \tag{4}$$

where $\mathbf{vec}(\cdot)$ is the matrix vectorization operator which converts a matrix into a vector by stacking its columns, $\otimes$ is the Kronecker product of matrices, $\boldsymbol{\Sigma}_A = [\sigma_{A_{rs}}]_{M\times M}$ is the genetic covariance matrix, whose rs-th element $\sigma_{A_{rs}}$ is the genetic covariance between $\boldsymbol{y}_r$ and $\boldsymbol{y}_s$, $\boldsymbol{\Sigma}_C = [\sigma_{C_{rs}}]_{M\times M}$ is the common environmental covariance matrix, and $\boldsymbol{\Sigma}_E = [\sigma_{E_{rs}}]_{M\times M}$ is the unique environmental covariance matrix. The distributional assumptions in Eq. (4) indicate that both the genetic effects and the environmental factors can be correlated across trait dimensions. When the trait is a scalar, the multivariate model (3) degenerates to the univariate model specified in Eq. (1). Analogous to the discussion above, $\boldsymbol{K}$ is derived from pedigree information in familial studies and empirically estimated from genome-wide SNP data in unrelated subject studies. As a result, $\boldsymbol{\Sigma}_A$ denotes the genetic covariance due to common SNPs when analyzing unrelated subjects.

We define the heritability of a multidimensional trait as

$$h^2 = \frac{\mathbf{tr}[\boldsymbol{\Sigma}_A]}{\mathbf{tr}[\boldsymbol{\Sigma}_P]} := \frac{\mathbf{tr}[\boldsymbol{\Sigma}_A]}{\mathbf{tr}[\boldsymbol{\Sigma}_A] + \mathbf{tr}[\boldsymbol{\Sigma}_C] + \mathbf{tr}[\boldsymbol{\Sigma}_E]}, \tag{5}$$

where $\boldsymbol{\Sigma}_P = [\sigma_{P_{rs}}]_{M\times M}$ is the phenotypic covariance matrix and $\mathbf{tr}[\cdot]$ is the trace operator of a matrix. This definition measures the proportion of the total phenotypic variance $\mathbf{tr}[\boldsymbol{\Sigma}_P]$ that can be explained by the total additive genetic variance $\mathbf{tr}[\boldsymbol{\Sigma}_A]$, and yields a heritability measure that is bounded between 0 and 1. When the trait is univariate, $\boldsymbol{\Sigma}_A$, $\boldsymbol{\Sigma}_C$, and $\boldsymbol{\Sigma}_E$ become scalars, and Eq. (5) reduces to the classical definition of heritability in Eq. (2). We use h^2_{SNP} in place of h^2 in unrelated subject studies to emphasize that it only captures genetic influences due to common genetic variants and is a lower bound for the narrow-sense heritability estimated using familial designs.

Properties of multidimensional heritability. Our definition of heritability is invariant to rotations of the data. For a linear transformation $\boldsymbol{T}$ applied to the trait dimensions in model (3), i.e.,

$$\boldsymbol{YT} = \boldsymbol{GT} + \boldsymbol{CT} + \boldsymbol{ET}, \tag{6}$$

using the properties of vectorization and the Kronecker product, the covariance structure of the transformed

trait $\boldsymbol{YT}$ can be computed as follows:

$$\begin{aligned}
\mathbf{cov}\big[\mathbf{vec}(\boldsymbol{YT})\big] &= \mathbf{cov}\big[\mathbf{vec}(\boldsymbol{GT})\big] + \mathbf{cov}\big[\mathbf{vec}(\boldsymbol{CT})\big] + \mathbf{cov}\big[\mathbf{vec}(\boldsymbol{ET})\big] \\
&= \mathbf{cov}\big[(\boldsymbol{T}^\top \otimes \boldsymbol{I})\mathbf{vec}(\boldsymbol{G})\big] + \mathbf{cov}\big[(\boldsymbol{T}^\top \otimes \boldsymbol{I})\mathbf{vec}(\boldsymbol{C})\big] + \mathbf{cov}\big[(\boldsymbol{T}^\top \otimes \boldsymbol{I})\mathbf{vec}(\boldsymbol{E})\big] \\
&= (\boldsymbol{T}^\top \otimes \boldsymbol{I})(\boldsymbol{\Sigma}_A \otimes \boldsymbol{K})(\boldsymbol{T} \otimes \boldsymbol{I}) + (\boldsymbol{T}^\top \otimes \boldsymbol{I})(\boldsymbol{\Sigma}_C \otimes \boldsymbol{\Lambda})(\boldsymbol{T} \otimes \boldsymbol{I}) + (\boldsymbol{T}^\top \otimes \boldsymbol{I})(\boldsymbol{\Sigma}_E \otimes \boldsymbol{I})(\boldsymbol{T} \otimes \boldsymbol{I}) \\
&= (\boldsymbol{T}^\top \boldsymbol{\Sigma}_A \boldsymbol{T}) \otimes \boldsymbol{K} + (\boldsymbol{T}^\top \boldsymbol{\Sigma}_C \boldsymbol{T}) \otimes \boldsymbol{\Lambda} + (\boldsymbol{T}^\top \boldsymbol{\Sigma}_E \boldsymbol{T}) \otimes \boldsymbol{I}.
\end{aligned} \tag{7}$$

Therefore, the transformed heritability is

$$h_T^2 = \frac{\mathbf{tr}[\boldsymbol{T}^\top \boldsymbol{\Sigma}_A \boldsymbol{T}]}{\mathbf{tr}[\boldsymbol{T}^\top \boldsymbol{\Sigma}_A \boldsymbol{T}] + \mathbf{tr}[\boldsymbol{T}^\top \boldsymbol{\Sigma}_C \boldsymbol{T}] + \mathbf{tr}[\boldsymbol{T}^\top \boldsymbol{\Sigma}_E \boldsymbol{T}]} = \frac{\mathbf{tr}[\boldsymbol{\Sigma}_A(\boldsymbol{T}\boldsymbol{T}^\top)]}{\mathbf{tr}[\boldsymbol{\Sigma}_P(\boldsymbol{T}\boldsymbol{T}^\top)]}. \tag{8}$$

When $\boldsymbol{T}$ is an orthogonal matrix satisfying $\boldsymbol{T}\boldsymbol{T}^\top = \boldsymbol{T}^\top\boldsymbol{T} = \boldsymbol{I}$, we have $h_T^2 = h^2$.

The definition of heritability in Eq. (5) can also been written as

$$h^2 = \frac{\mathbf{tr}[\boldsymbol{\Sigma}_A]}{\mathbf{tr}[\boldsymbol{\Sigma}_P]} = \frac{\sum_m \sigma_{A_{mm}}}{\sum_m \sigma_{P_{mm}}} = \sum_m \gamma_m h_m^2, \tag{9}$$

where $\gamma_m = \sigma_{P_{mm}} / \sum_m \sigma_{P_{mm}}$ with $\sum_m \gamma_m = 1$, and $h_m^2 = \sigma_{G_{mm}}/\sigma_{P_{mm}}$ is the heritability of the m-th component of the trait. Therefore, our definition of the heritability of a multidimensional trait is essentially a weighted average of the heritability of its individual components.

A moment-matching estimator for unrelated subject studies. The model (3) can in principle be fitted using likelihood-based methods. However, this can be computationally expensive when the dimension of the trait is moderate. Here we derive an alternative moment-matching estimator for unrelated subject studies where the common environmental matrix $\boldsymbol{\Lambda}$ vanishes. Specifically, the multivariate model $\boldsymbol{Y} = \boldsymbol{G} + \boldsymbol{E}$ and its distributional assumptions $\mathbf{vec}(\boldsymbol{G}) \sim \mathrm{N}(\mathbf{0}, \boldsymbol{\Sigma}_A \otimes \boldsymbol{K})$ and $\mathbf{vec}(\boldsymbol{E}) \sim \mathrm{N}(\mathbf{0}, \boldsymbol{\Sigma}_E \otimes \boldsymbol{I})$ lead to the following relationship:

$$\mathbf{cov}[\boldsymbol{y}_r, \boldsymbol{y}_s] = \sigma_{A_{rs}}\boldsymbol{K} + \sigma_{E_{rs}}\boldsymbol{I}, \quad 1 \leqslant r, s \leqslant M. \tag{10}$$

Therefore, an unbiased estimator of $\sigma_{A_{rs}}$ and $\sigma_{E_{rs}}$ can be obtained by regressing $\boldsymbol{y}_r\boldsymbol{y}_s^\top$, the empirical estimate of the phenotypic covariance matrix $\mathbf{cov}[\boldsymbol{y}_r, \boldsymbol{y}_s]$, onto the the empirical genetic similarity matrix $\boldsymbol{K}$ and identity matrix $\boldsymbol{I}$. In particular, we consider the following multiple regression problem:

$$\mathbf{vec}(\boldsymbol{y}_r\boldsymbol{y}_s^\top) = \sigma_{A_{rs}}\mathbf{vec}(\boldsymbol{K}) + \sigma_{E_{rs}}\mathbf{vec}(\boldsymbol{I}) + \boldsymbol{\epsilon}_{rs}, \tag{11}$$

where $\boldsymbol{\epsilon}_{rs}$ is the residual of this regression. This approach is essentially the Haseman-Elston regression for the classical heritability analysis [59, 60], and has been extended recently to handle various study designs including case-control studies, and more generally termed as phenotype correlation-genetic correlation (PCGC)

regression [35]. The ordinary least squares (OLS) estimator of the multiple regression problem (11) satisfies the linear system:

$$\begin{bmatrix} \mathbf{tr}[\boldsymbol{K}^2] & \mathbf{tr}[\boldsymbol{K}] \\ \mathbf{tr}[\boldsymbol{K}] & \mathbf{tr}[\boldsymbol{I}] \end{bmatrix} \begin{bmatrix} \sigma_{A_{rs}} \\ \sigma_{E_{rs}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{y}_r^\top \boldsymbol{K} \boldsymbol{y}_s \\ \boldsymbol{y}_r^\top \boldsymbol{y}_s \end{bmatrix}, \tag{12}$$

and can be explicitly written as

$$\begin{aligned} \widehat{\sigma}_{A_{rs}} &= \frac{\boldsymbol{y}_r^\top (N\boldsymbol{K} - \mathbf{tr}[\boldsymbol{K}]\boldsymbol{I})\boldsymbol{y}_s}{N\mathbf{tr}[\boldsymbol{K}^2] - \mathbf{tr}^2[\boldsymbol{K}]} := \frac{1}{v_K}\boldsymbol{y}_r^\top (\boldsymbol{K} - \tau\boldsymbol{I})\boldsymbol{y}_s, \\ \widehat{\sigma}_{E_{rs}} &= \frac{\boldsymbol{y}_r^\top (\mathbf{tr}[\boldsymbol{K}^2]\boldsymbol{I} - \mathbf{tr}[\boldsymbol{K}]\boldsymbol{K})\boldsymbol{y}_s}{N\mathbf{tr}[\boldsymbol{K}^2] - \mathbf{tr}^2[\boldsymbol{K}]} := \frac{1}{v_K}\boldsymbol{y}_r^\top (\kappa\boldsymbol{I} - \tau\boldsymbol{K})\boldsymbol{y}_s, \end{aligned} \tag{13}$$

where we have defined $\tau = \mathbf{tr}[\boldsymbol{K}]/N$, $\kappa = \mathbf{tr}[\boldsymbol{K}^2]/N$ and $v_K = \mathbf{tr}[\boldsymbol{K}^2] - \mathbf{tr}^2[\boldsymbol{K}]/N = N(\kappa - \tau^2)$. Therefore, it can be seen that unbiased estimates of the genetic and environmental covariance matrices are as follows:

$$\widehat{\boldsymbol{\Sigma}}_A = \frac{1}{v_K}\boldsymbol{Y}^\top (\boldsymbol{K} - \tau\boldsymbol{I})\boldsymbol{Y}, \quad \widehat{\boldsymbol{\Sigma}}_E = \frac{1}{v_K}\boldsymbol{Y}^\top (\kappa\boldsymbol{I} - \tau\boldsymbol{K})\boldsymbol{Y}. \tag{14}$$

Let $\widehat{\boldsymbol{\Sigma}}_P = \widehat{\boldsymbol{\Sigma}}_A + \widehat{\boldsymbol{\Sigma}}_E$, the SNP heritability of a multidimensional trait can then be estimated as

$$\widehat{h}^2_{\text{SNP}} = \frac{\mathbf{tr}[\widehat{\boldsymbol{\Sigma}}_A]}{\mathbf{tr}[\widehat{\boldsymbol{\Sigma}}_P]} = \frac{\mathbf{tr}[\widehat{\boldsymbol{\Sigma}}_A]}{\mathbf{tr}[\widehat{\boldsymbol{\Sigma}}_A] + \mathbf{tr}[\widehat{\boldsymbol{\Sigma}}_E]}. \tag{15}$$

For scalar traits, Eq. (15) degenerates to the classical Haseman-Elston regression estimator.

Sampling variance of the point estimator. We now derive the variance of $\widehat{h}^2_{\text{SNP}}$. For notational simplicity, we denote $\boldsymbol{Q}_A = (\boldsymbol{K} - \tau\boldsymbol{I})/v_K$, $\boldsymbol{Q}_E = (\kappa\boldsymbol{I} - \tau\boldsymbol{K})/v_K$, and also $t_A = \mathbf{tr}[\widehat{\boldsymbol{\Sigma}}_A] = \mathbf{tr}[\boldsymbol{Y}^\top \boldsymbol{Q}_A \boldsymbol{Y}]$, $t_E = \mathbf{tr}[\widehat{\boldsymbol{\Sigma}}_E] = \mathbf{tr}[\boldsymbol{Y}^\top \boldsymbol{Q}_E \boldsymbol{Y}]$, $t_P = t_A + t_E$. Let $\boldsymbol{t} = (t_A, t_E)^\top$ and define $f(\boldsymbol{t}) = t_A/(t_A + t_E) = t_A/t_P$. Using a Taylor expansion, we can approximate the variance of the function f as follows:

$$\mathbf{var}[\widehat{h}^2_{\text{SNP}}] = \mathbf{var}[f(\boldsymbol{t})] \approx \frac{\partial f(\boldsymbol{t})}{\partial \boldsymbol{t}} \mathbf{cov}[\boldsymbol{t}] \frac{\partial f(\boldsymbol{t})}{\partial \boldsymbol{t}^\top}, \tag{16}$$

where

$$\frac{\partial f(\boldsymbol{t})}{\partial \boldsymbol{t}} = \left(\frac{\partial f(\boldsymbol{t})}{\partial t_A}, \frac{\partial f(\boldsymbol{t})}{\partial t_E} \right) = \left(\frac{t_E}{t_P^2}, \frac{-t_A}{t_P^2} \right). \tag{17}$$

To compute $\mathbf{cov}[\boldsymbol{t}]$, we define $\boldsymbol{V}_{rs} = \mathbf{cov}[\boldsymbol{y}_r, \boldsymbol{y}_s] = \sigma_{A_{rs}}\boldsymbol{K} + \sigma_{E_{rs}}\boldsymbol{I}$, and notice that for any symmetric matrices $\boldsymbol{Q}_\alpha$ and $\boldsymbol{Q}_\beta$,

$$\begin{aligned} \mathbf{cov}\left\{\mathbf{tr}[\boldsymbol{Y}^\top \boldsymbol{Q}_\alpha \boldsymbol{Y}], \mathbf{tr}[\boldsymbol{Y}^\top \boldsymbol{Q}_\beta \boldsymbol{Y}]\right\} &= \textstyle\sum_{r,s=1}^M \mathbf{cov}\left\{\boldsymbol{y}_r^\top \boldsymbol{Q}_\alpha \boldsymbol{y}_r, \boldsymbol{y}_s^\top \boldsymbol{Q}_\beta \boldsymbol{y}_s\right\} \\ &= 2\textstyle\sum_{r,s=1}^M \mathbf{tr}[\boldsymbol{Q}_\alpha \boldsymbol{V}_{rs} \boldsymbol{Q}_\beta \boldsymbol{V}_{rs}]. \end{aligned} \tag{18}$$

Therefore,

$$\mathbf{cov}[\boldsymbol{t}] = 2\sum_{r,s=1}^{M}\begin{bmatrix} \mathbf{tr}[\boldsymbol{Q}_A\boldsymbol{V}_{rs}\boldsymbol{Q}_A\boldsymbol{V}_{rs}] & \mathbf{tr}[\boldsymbol{Q}_A\boldsymbol{V}_{rs}\boldsymbol{Q}_E\boldsymbol{V}_{rs}] \\ \mathbf{tr}[\boldsymbol{Q}_E\boldsymbol{V}_{rs}\boldsymbol{Q}_A\boldsymbol{V}_{rs}] & \mathbf{tr}[\boldsymbol{Q}_E\boldsymbol{V}_{rs}\boldsymbol{Q}_E\boldsymbol{V}_{rs}] \end{bmatrix}. \tag{19}$$

Eq. (19) can be computationally expensive for a moderate or large M. We approximate Eq. (19) by using two assumptions that are often made when estimating the sampling variance of the heritability estimator in the study of unrelated individuals [61]: (1) the off-diagonal elements in the empirical genetic similarity matrix $\boldsymbol{K}$ are small, such that $\boldsymbol{K} \approx \boldsymbol{I}$ and $\boldsymbol{V}_{rs} = \sigma_{A_{rs}}\boldsymbol{K} + \sigma_{E_{rs}}\boldsymbol{I} \approx \sigma_{A_{rs}}\boldsymbol{I} + \sigma_{E_{rs}}\boldsymbol{I} = \sigma_{P_{rs}}\boldsymbol{I}$; and (2) the phenotypic covariance $\boldsymbol{\Sigma}_P$ is known or can be estimated with very high precision. Using assumption (1), the covariance of $\boldsymbol{t}$ can be simplified as follows:

$$\begin{aligned} \mathbf{cov}[\boldsymbol{t}] &\approx 2\textstyle\sum_{r,s=1}^{M}\sigma_{P_{rs}}^2 \begin{bmatrix} \mathbf{tr}[\boldsymbol{Q}_A^2] & \mathbf{tr}[\boldsymbol{Q}_A\boldsymbol{Q}_E] \\ \mathbf{tr}[\boldsymbol{Q}_E\boldsymbol{Q}_A] & \mathbf{tr}[\boldsymbol{Q}_E^2] \end{bmatrix} \\ &= 2\mathbf{tr}[\boldsymbol{\Sigma}_P^2] \begin{bmatrix} \mathbf{tr}[\boldsymbol{Q}_A^2] & \mathbf{tr}[\boldsymbol{Q}_A\boldsymbol{Q}_E] \\ \mathbf{tr}[\boldsymbol{Q}_E\boldsymbol{Q}_A] & \mathbf{tr}[\boldsymbol{Q}_E^2] \end{bmatrix} \\ &= \frac{2\mathbf{tr}[\boldsymbol{\Sigma}_P^2]}{v_K}\begin{bmatrix} 1 & -\tau \\ -\tau & \kappa \end{bmatrix} \approx \frac{2\mathbf{tr}[\boldsymbol{\Sigma}_P^2]}{v_K}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}. \end{aligned} \tag{20}$$

Therefore,

$$\begin{aligned} \mathbf{var}[\widehat{h}_{\mathrm{SNP}}^2] = \mathbf{var}[f(\boldsymbol{t})] &\approx \frac{\partial f(\boldsymbol{t})}{\partial \boldsymbol{t}}\mathbf{cov}[\boldsymbol{t}]\frac{\partial f(\boldsymbol{t})}{\partial \boldsymbol{t}^\top} \\ &\approx \frac{2\mathbf{tr}[\boldsymbol{\Sigma}_P^2]}{v_K t_P^4}\cdot(t_E, -t_A)\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} t_E \\ -t_A \end{bmatrix} \\ &= \frac{2\mathbf{tr}[\boldsymbol{\Sigma}_P^2]}{v_K t_P^4}(t_A + t_E)^2 = \frac{2\mathbf{tr}[\boldsymbol{\Sigma}_P^2]}{v_K t_P^2} = \frac{2}{v_K}\cdot\frac{\mathbf{tr}[\boldsymbol{\Sigma}_P^2]}{\mathbf{tr}^2[\widehat{\boldsymbol{\Sigma}}_P]} \\ &\approx \frac{2}{v_K}\cdot\frac{\mathbf{tr}[\widehat{\boldsymbol{\Sigma}}_P^2]}{\mathbf{tr}^2[\widehat{\boldsymbol{\Sigma}}_P]}, \end{aligned} \tag{21}$$

where in the last approximation we have used assumption (2) and replaced $\boldsymbol{\Sigma}_P$ with its empirical estimate $\widehat{\boldsymbol{\Sigma}}_P = \widehat{\boldsymbol{\Sigma}}_A + \widehat{\boldsymbol{\Sigma}}_E$. We note that given the empirical genetic similarity matrix $\boldsymbol{K}$, the estimator (21) only depends on the sample size and the phenotypic covariance structure.

For scalar traits, $\mathbf{tr}[\widehat{\boldsymbol{\Sigma}}_P^2] = \mathbf{tr}^2[\widehat{\boldsymbol{\Sigma}}_P]$, and the estimator (21) degenerates to $\mathbf{var}[\widehat{h}_{\mathrm{SNP}}^2] \approx 2/v_K$, which coincides with existing results in the literature [61]. In general, the covariance matrix $\widehat{\boldsymbol{\Sigma}}_P$ is non-negative

definite. Let $\mu_1 \geqslant \mu_2 \geqslant \cdots \geqslant \mu_M \geqslant 0$ denote its eigenvalues, we have

$$\frac{\mathbf{tr}[\widehat{\boldsymbol{\Sigma}}_P^2]}{\mathbf{tr}^2[\widehat{\boldsymbol{\Sigma}}_P]} = \frac{\sum_{i=1}^M \mu_i^2}{\left(\sum_{i=1}^M \mu_i\right)^2} \leqslant 1. \tag{22}$$

This inequality becomes an equality if and only if $\mathbf{rank}[\widehat{\boldsymbol{\Sigma}}_P] = 1$, i.e., the M traits are perfectly correlated. Therefore, combining multiple traits reduces the variability of heritability estimates relative to analyzing each trait individually.

Statistic inference. To measure the significance of a heritability estimate, a p-value can be computed by conducting a Wald test. Since the null hypothesis, $\mathcal{H}_0 : h^2_{\text{SNP}} = 0$, lies on the boundary of the parameter space, the Wald test statistic is distributed as

$$\frac{\widehat{h}^4_{\text{SNP}}}{\mathbf{var}[\widehat{h}^2_{\text{SNP}}]} \sim \frac{1}{2}\chi_0^2 + \frac{1}{2}\chi_1^2, \tag{23}$$

a half-half mixture of χ_0^2, a chi-squared distribution with all probability mass at zero, and χ_1^2, a chi-squared distribution with 1 degrees of freedom [62].

Alternatively, permutation inference can be used by shuffling the rows and columns of the empirical genetic similarity matrix $\boldsymbol{K}$. For each permutation $r = 1, 2, \cdots, N_{\text{perm}}$, we record the heritability estimate $\widehat{h}^2_{\text{SNP}}(r)$ computed from the permuted data. Then for an observed heritability estimate $\widehat{h}^2_{\text{SNP}}$, the permutation p-value can be computed as [63]

$$p_{\text{perm}} = \frac{\#\{\widehat{h}^2_{\text{SNP}}(r) \geqslant \widehat{h}^2_{\text{SNP}}\}}{N_{\text{perm}}}. \tag{24}$$

Heritability estimation in familial studies. A moment-matching estimator can be analogously derived for familial data analysis but has low statistical efficiency due to the strong correlation between the genetic similarity matrix $\boldsymbol{K}$ and the common environmental matrix $\boldsymbol{\Lambda}$. Therefore, when analyzing the HCP data, we estimate the heritability of each individual component of a multidimensional trait using the restricted maximum likelihood (ReML) algorithm and combine these estimates using a variance-weighted average as derived in Eq. (9).

To estimate the variance of the heritability estimate of a multidimensional trait, we employ a block bootstrapping procedure whereby families are randomly resampled with replacement to produce a bootstrap sample and the heritability is re-estimated. This procedure is repeated for N_{boots} times ($N_{\text{boots}} = 1,000$ in the present study) to yield bootstrap heritability estimates $\widehat{h}^2(b)$, $b = 1, 2, \cdots, N_{\text{boots}}$. The variance of the heri-

tability estimate is then estimated as [64]

$$\mathbf{var}[\widehat{h}^2] = \frac{1}{N_{\text{boots}} - 1} \sum_{b=1}^{N_{\text{boots}}} \left(\widehat{h}^2(b) - \widehat{h}^2(\cdot) \right)^2, \tag{25}$$

where $\widehat{h}^2(\cdot) = \sum_{b=1}^{N_{\text{boots}}} \widehat{h}^2(b)/N_{\text{boots}}$.

Modeling covariates. When covariates or nuisance variables need to be adjusted, model (3) becomes a multivariate linear mixed effects model:

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{B} + \boldsymbol{G} + \boldsymbol{C} + \boldsymbol{E}, \tag{26}$$

with the distributional assumptions $\mathbf{vec}(\boldsymbol{G}) \sim \mathrm{N}(\mathbf{0}, \boldsymbol{\Sigma}_A \otimes \boldsymbol{K})$, $\mathbf{vec}(\boldsymbol{C}) \sim \mathrm{N}(\mathbf{0}, \boldsymbol{\Sigma}_C \otimes \boldsymbol{\Lambda})$ and $\mathbf{vec}(\boldsymbol{E}) \sim \mathrm{N}(\mathbf{0}, \boldsymbol{\Sigma}_E \otimes \boldsymbol{I})$, where $\boldsymbol{X}$ is an $N \times q$ matrix of covariates, and $\boldsymbol{B}$ is a $q \times M$ matrix of fixed effects. We employ a strategy described in [65] to removes the covariate matrix from the model. Specifically, there exists an $N \times (N-q)$ matrix $\boldsymbol{U}$, satisfying $\boldsymbol{U}^\top \boldsymbol{U} = \boldsymbol{I}_{(N-q)\times(N-q)}$, $\boldsymbol{U}\boldsymbol{U}^\top = \boldsymbol{P}_0$, and $\boldsymbol{U}^\top \boldsymbol{X} = \mathbf{0}$, where $\boldsymbol{P}_0 = \boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}^\top \boldsymbol{X})^{-1}\boldsymbol{X}^\top$. The matrix $\boldsymbol{U}^\top$ projects the data from the N dimensional space onto an $N-q$ dimensional subspace:

$$\widetilde{\boldsymbol{Y}} := \boldsymbol{U}^\top \boldsymbol{Y} = \boldsymbol{U}^\top \boldsymbol{G} + \boldsymbol{U}^\top \boldsymbol{C} + \boldsymbol{U}^\top \boldsymbol{E} := \widetilde{\boldsymbol{G}} + \widetilde{\boldsymbol{C}} + \widetilde{\boldsymbol{E}}, \tag{27}$$

where $\mathbf{vec}(\widetilde{\boldsymbol{G}}) \sim \mathrm{N}\big(\mathbf{0}, \boldsymbol{\Sigma}_A \otimes (\boldsymbol{U}^\top \boldsymbol{K} \boldsymbol{U})\big)$, $\mathbf{vec}(\widetilde{\boldsymbol{C}}) \sim \mathrm{N}\big(\mathbf{0}, \boldsymbol{\Sigma}_C \otimes (\boldsymbol{U}^\top \boldsymbol{\Lambda} \boldsymbol{U})\big)$, and $\mathbf{vec}(\widetilde{\boldsymbol{E}}) \sim \mathrm{N}(\mathbf{0}, \boldsymbol{\Sigma}_E \otimes \boldsymbol{I})$. The transformed model is the same as model (3) and thus all estimation and inferential methods developed above can be applied.

The Brain Genomics Superstruct Project (GSP). The Harvard/Massachusetts General Hospital (MGH) Brain Genomics Superstruct Project (GSP) is a neuroimaging and genetics study of brain and behavioral phenotypes. More than 3,500 native English-speaking adults with normal or corrected-to-normal vision were recruited from Harvard University, MGH, and the surrounding Boston communities. To avoid spurious effects resulting from population stratification, we restricted our analyses to 1,320 young adults (18-35 years) of non-Hispanic European ancestry with no history of psychiatric illnesses or major health problems (age, 21.54±3.19 years; female, 53.18%; right-handedness, 91.74%). All participants provided written informed consent in accordance with guidelines set by the Partners Health Care Institutional Review Board or the Harvard University Committee on the Use of Human Subjects in Research. For further details about the recruitment process, participants, and imaging data acquisition, we refer the reader to Holmes et al. [36, 37].

The Human Connectome Project (HCP). The Human Connectome Project (HCP) collects imaging, behavioral, and demographic data from a large population of healthy adults and aims to shed light on anatomical and functional connectivity within the healthy human brain. We used preprocessed structural MRI data from the WU-Minn HCP consortium and selected subjects that have a similar age range (22-35 years) and ancestry (non-Hispanic/Latino European) as the GSP sample. The 590 subjects we analyzed (age, 29.21±3.45 years; female 55.93%) come from 249 families and comprise 72 monozygotic (MZ) twin pairs, 69 dizygotic (DZ) twin pairs, 253 full siblings of twins and 55 singletons (single birth individuals without siblings). Further details about the recruitment process, imaging data acquisition and MRI data preprocessing can be found in [41, 66].

Genetic analysis. We used PLINK 1.90 (https://www.cog-genomics.org/plink2) [67], to preprocess the GSP genome-wide SNP data. Major procedures included sex discrepancy check, removing population outliers, spuriously related subjects and subjects with low genotype call rate ($< 97\%$). Individual markers that contained an ambiguous strand assignment and that did not satisfy the following quality control criteria were excluded from the analyses: genotype call rate $\geqslant 97\%$, minor allele frequency (MAF) $\geqslant 1\%$, and Hardy-Weinberg equilibrium $p \geqslant 1 \times 10^{-6}$. 574,632 SNPs remained for analysis after quality control. We performed a multidimensional scaling (MDS) analysis to ensure that no clear population stratification and outliers exist in the sample (Supplementary Fig. 3). The genetic similarity matrix was estimated from all genotyped autosomal SNPs.

Laplace-Beltrami Spectrum based shape descriptor. The intrinsic geometry of any 2D or 3D manifold can be characterized by its Laplace-Beltrami Spectrum (LBS) [38, 39], which is obtained by solving the following Laplacian eigenvalue problem (or Helmoltz equation):

$$\Delta f = -\lambda f, \tag{28}$$

where Δ is the Laplace-Beltrami operator, a generalization of the Laplacian in the Euclidean space to manifolds, f is a real-valued eigenfunction defined on a Riemannian manifold, and λ is the corresponding eigenvalue. Eq. (28) can be solved by the finite element method, yielding a diverging sequence of eigenvalues $0 \leqslant \lambda_1 \leqslant \lambda_2 \leqslant \cdots \uparrow +\infty$. An implementation of the algorithm is freely available (http://reuter.mit.edu/software/shapedna). The first M eigenvalues of the LBS can be used to define a description of the object, which provides a numerical fingerprint or signature of the shape, and is thus known as (length-M) "Shape-DNA".

Shape analysis pipeline. We used FreeSurfer (http://freesurfer.net) [68], version 4.5.0, a freely available, widely used, and extensively validated brain MRI analysis software package, to process the GSP structural brain MRI scans and label subcortical brain structures. HCP MRI scans were preprocessed by the WU-Minn HCP consortium, and the label files of subcortical structures have been made available [41, 66]. Surface meshes of brain structures were obtained via marching cubes from subcortical segmentations. We created triangular meshes on the boundary surfaces for 20 structures. We then geometrically smoothed these meshes and solved the eigenvalue problem of the 2D Laplace-Beltrami operator on each of these representations, yielding the LBS-based shape descriptor [40]. A python implementation of this pipeline is freely available (http://reuter.mit.edu/software/brainprint).

Heritability analyses of neuroanatomical shape. We treated the length-M LBS-based shape descriptor of each structure as a multidimensional trait and quantified its heritability. In the case of a closed manifold without a boundary, the first eigenvalue is always zero and was thus removed from analysis. Theoretical and empirical evidence have confirmed that the eigenvalues grow linearly and their variance grows quadratically [38, 40]. To avoid that higher eigenvalues dominate the phenotypic covariance, we re-weighted the m-th eigenvalue for the i-th subject as [38]:

$$\widetilde{\lambda}_{i,m} = \lambda_{i,m}/m, \quad i = 1, 2, \cdots, N, \quad m = 1, 2, \cdots, M. \tag{29}$$

This ensures a balanced contribution of lower and higher eigenvalues on the phenotypic covariance. The LBS also depends on the overall size of the structure. To measure the genetic influences on the shape that are complementary to volume, we further scaled the eigenvalues as:

$$\widetilde{\widetilde{\lambda}}_{i,m} = \widetilde{\lambda}_{i,m} \cdot V_i^{2/3}, \quad i = 1, 2, \cdots, N, \quad m = 1, 2, \cdots, M, \tag{30}$$

where V_i is the volume of the structure for the i-th subject. Since scaling the eigenvalues by a factor η results in scaling the underlying manifold by a factor $\eta^{-1/2}$ [38], the normalization (30) ensures that the volumes of the structure are identical across individuals.

We combined the left and right structures in our heritability analyses by averaging their volumetric measurements and concatenating their re-weighted and scaled shape descriptors into one multidimensional trait. We included age, gender, handedness, scanner group, console group, and the top ten principal components of the empirical genetic similarity matrix as covariates when analyzing the GSP sample, and included age, gender and handedness as covariates when analyzing the HCP sample. To remove potential size effect, we always explicitly included the volume of the corresponding structure as a covariate in our shape analyses.

The number of eigenvalues incorporated in the LBS-based shape descriptor and the amount of smoothing applied to the surface mesh are crucial study designs, which might have an impact on heritability estimates. In particular, incorporating a very small number of eigenvalues may be insufficient to characterize the shape of a structure, while very large eigenvalues typically capture fine-scale details, which can be noise and thus might reduce sensitivity. In this study, we reported results obtained by incorporating 50 eigenvalues in the shape descriptor and applying 3 iterations of geometric smoothing to the surface mesh. We conducted sensitivity analyses and confirmed that in the present shape analysis the results were largely robust to different parameter settings (Supplementary Fig. 1).

Visualizing the principal mode of shape variation. We note that, as shown above, our definition of the heritability of a multidimensional trait is a variance-weighted average of individual components, and is invariant to the rotation of the trait vector. Therefore, an equivalent definition of the heritability of a length-M LBS-based shape descriptor is the variance-weighted average of the heritability of the first M principal components (PCs) of the descriptor, because principal component analysis (PCA) is essentially a rotation of the data. The first PC thus explains the greatest shape variation and has the largest impact on the overall heritability estimate of the shape.

To visualize shape variation along the first PC of the shape descriptor for a given structure, we first aligned the structures from all subjects to a template, *fsaverage*, which is a population average distributed with FreeSurfer [68], using a 7-parameter (global scaling plus 6-parameter rigid body transformation) registration with linear interpolation. Both individual structures and the template were represented with binary label maps, where voxels within the corresponding segmentation label had one and the remainder of the volume had zero values. The registration algorithm maximized the overlap, measured with the Dice score [69], between the corresponding label maps (the fixed template and moving subject which was interpolated and thresholded at 0.5). Note that LBS is invariant to the spatial position and orientation of an object, and we had normalized the shape descriptor for volume in all the analyses. Thus this registration has no impact on the results of our heritability analyses. We then created a sample-specific population average of the structure by computing a weighted average of the interpolated subject images. In particular, each subject was associated with a weight equal to a Gaussian kernel centered around the mean of the first PC and evaluated at the subject's first PC of the shape descriptor. The width of the kernel was selected such that 500 shapes received non-zero weights. The isosurface of the resulting probability map at 0.5 was used to represent the average shape of the structure, and all visualizations were presented on this surface.

The same Gaussian kernel was used to generate average probability images for shapes centered at the two extremes (± 2 standard deviation or SD) of the principal axis. These average probability images were offset to achieve identical volumes when thresholded at 0.5. The difference of the two extreme shapes were depicted on the sample-specific population average, by visualizing the difference in the probability values. Blue indicated that the average shape at -2 SD achieved a higher probability value and thus was larger in those regions than the average shape at the $+2$ SD. For red regions, the opposite was true.

Data availability. The Brain Genomics Superstruct Project (GSP) data analyzed during the current study are publicly available at http://neuroinformatics.harvard.edu/gsp/. The Human Connectome Project (HCP) data analyzed during the current study are publicly available at http://www.humanconnectome.org. Other data are available from the corresponding author on reasonable request.

References

[1] T.A. Manolio, F.S. Collins, N.J. Cox, D.B. Goldstein, L.A. Hindorff, et al. Finding the missing heritability of complex diseases. *Nature*, 461(7265):747–753, 2009.

[2] P.F. Sullivan, M.J. Daly, and M. O'Donovan. Genetic architectures of psychiatric disorders: the emerging picture and its implications. *Nature Reviews Genetics*, 13(8):537–551, 2012.

[3] P.M. Visscher, M.A. Brown, M.I. McCarthy, and J. Yang. Five years of GWAS discovery. *The American Journal of Human Genetics*, 90(1):7–24, 2012.

[4] Psychiatric Genomics Consortium. Biological insights from 108 schizophrenia-associated genetic loci. *Nature*, 511(7510):421–427, 2014.

[5] Psychiatric Genomics Consortium. Identification of risk loci with shared effects on five major psychiatric disorders: a genome-wide analysis. *The Lancet*, 381(9875):1371–1379, 2013.

[6] J. Gratten, N.R. Wray, M.C. Keller, and P.M. Visscher. Large-scale genomics unveils the genetic architecture of psychiatric disorders. *Nature Neuroscience*, 17(6):782–790, 2014.

[7] I. I. Gottesman and J. Shields. *Schizophrenia genetics: A twin study vantage point*. New York: Academic Press, 1972.

[8] I.I. Gottesman and T.D. Gould. The endophenotype concept in psychiatry: etymology and strategic intentions. *American Journal of Psychiatry*, 160(4):636, 2003.

[9] A. Meyer-Lindenberg and D.R. Weinberger. Intermediate phenotypes and genetic mechanisms of psychiatric disorders. *Nature Reviews Neuroscience*, 7(10):818–827, 2006.

[10] P.M. Thompson, T. Ge, D.C. Glahn, N. Jahanshad, and T.E. Nichols. Genetics of the connectome. *NeuroImage*, 80:475–488, 2013.

[11] S.G. Potkin, G. Guffanti, A. Lakatos, J.A. Turner, F. Kruggel, et al. Hippocampal atrophy as a quantitative trait in a genome-wide association study identifying novel susceptibility genes for Alzheimer's disease. *PLoS ONE*, 4(8):e6501, 2009.

[12] L. Shen, S. Kim, S.L. Risacher, K. Nho, S. Swaminathan, et al. Whole genome association study of brain-wide imaging phenotypes for identifying quantitative trait loci in MCI and AD: A study of the ADNI cohort. *NeuroImage*, 53(3):1051–1063, 2010.

[13] J.L. Stein, S.E. Medland, A.A. Vasquez, D.P. Hibar, R.E. Senstad, et al. Identification of common variants associated with human hippocampal and intracranial volumes. *Nature Genetics*, 44(5):552–561, 2012.

[14] M.R. Sabuncu, R.L. Buckner, J.W Smoller, P.H. Lee, B. Fischl, et al. The association between a polygenic alzheimer score and cortical thickness in clinically normal subjects. *Cerebral Cortex*, 22(11): 2653–2661, 2012.

[15] J.G. Csernansky, S. Joshi, L. Wang, J.W. Haller, M. Gado, et al. Hippocampal morphometry in schizophrenia by high dimensional brain mapping. *Proceedings of the National Academy of Sciences*, 95(19):11406–11411, 1998.

[16] J.G. Csernansky, M.K. Schindler, N.R. Splinter, L. Wang, M. Gado, et al. Abnormalities of thalamic volume and shape in schizophrenia. *The American Journal of Psychiatry*, 161(5):896–902, 2004.

[17] S.H. Joshi, K.L. Narr, O.R. Philips, K.H. Nuechterlein, R.F. Asarnow, et al. Statistical shape analysis of the corpus callosum in schizophrenia. *NeuroImage*, 64:547–559, 2013.

[18] C.N. Vidal, R. Nicolson, T.J. DeVito, K.M. Hayashi, J.A. Geaga, et al. Mapping corpus callosum deficits in autism: an index of aberrant cortical connectivity. *Biological Psychiatry*, 60(3):218–225, 2006.

[19] D.L. Dierker, E. Feczko, J.R. Pruett Jr, S.E. Petersen, B.L. Schlaggar, et al. Analysis of cortical shape in children with simplex autism. *Cerebral Cortex*, 25(4):1042–1051, 2015.

[20] J. Hwang, I.K. Lyoo, S.R. Dager, S.D. Friedman, J.S. Oh, et al. Basal ganglia shape alterations in bipolar disorder. *The American Journal of Psychiatry*, 163(2):276–285, 2006.

[21] C.E. Bearden, P.M. Thompson, R.A. Dutton, B.N. Frey, M.A.M Peluso, et al. Three-dimensional mapping of hippocampal anatomy in unmedicated and lithium-treated patients with bipolar disorder. *Neuropsychopharmacology*, 33:1229–1238, 2008.

[22] J.G. Csernansky, L. Wang, J. Swank, J.P. Miller, M. Gado, et al. Preclinical detection of Alzheimer's disease: hippocampal shape and volume predict dementia onset in the elderly. *NeuroImage*, 25(3):783–792, 2005.

[23] E. Gerard, G. Chetelat, M. Chupin, R. Cuingnet, B. Desgranges, et al. Multidimensional classification of hippocampal shape features discriminates Alzheimer's disease and mild cognitive impairment from normal aging. *NeuroImage*, 47(4):1476–1486, 2009.

[24] S.G. Costarred, I.D. Dinov, Z. Tu, Y. Shi, C. Liu, et al. Automated hippocampal shape analysis predicts the onset of dementia in mild cognitive impairment. *NeuroImage*, 56(1):212–219, 2011.

[25] X. Tang, D. Holland, A.M. Dale, L. Younes, M.I. Miller, et al. Shape abnormalities of subcortical and ventricular structures in mild cognitive impairment and Alzheimer's disease: Detecting, quantifying, and predicting. *Human Brain Mapping*, 35(8):3701–3725, 2014.

[26] J.A. Posener, L. Wang, J.L. Price, M.H. Gado, M.A. Province, et al. High-dimensional mapping of the hippocampus in depression. *The American Journal of Psychiatry*, 160(1):83–89, 2003.

[27] A. Qiu, D. Crocetti, M. Adler, E.M. Mahone, M.B. Denckla, et al. Basal ganglia volume and shape in children with attention deficit hyperactivity disorder. *The American Journal of Psychiatry*, 166(1): 74–82, 2009.

[28] A. Qiu, W.D. Taylor, Z. Zhao, J.R. MacFall, M.I. Miller, et al. APOE related hippocampal shape alteration in geriatric depression. *NeuroImage*, 44(3):620–626, 2009.

[29] J. Shi, N. Lepore, B.A. Gutman, P.M. Thompson, L.C. Baxter, et al. Genetic influence of apolipoprotein E4 genotype on hippocampal morphometry: An $n = 725$ surface-based Alzheimer's disease neuroimaging initiative study. *Human Brain Mapping*, 35(8):3903–3918, 2014.

[30] N. Fani, D. Gutman, E.B. Tone, L. Almli, K.B. Mercer, et al. FKBP5 and attention bias for threat associations with hippocampal function and shape. *JAMA Psychiatry*, 70(4):392–400, 2013.

[31] D.R. Roalf, S.N. Vandekar, L. Almasy, K. Ruparel, T.D. Satterthwaite, et al. Heritability of subcortical and limbic brain volume and shape in multiplex-multigenerational families with schizophrenia. *Biological Psychiatry*, 77(2):137–146, 2015.

[32] D.P. Hibar, J.L. Stein, M.E. Renteria, A. Arias-Vasquez, S. Desrivières, et al. Common genetic variants influence human subcortical brain structures. *Nature*, 520(7546):224–229, 2015.

[33] J. Yang, B. Benyamin, B.P. McEvoy, S. Gordon, A.K. Henders, et al. Common SNPs explain a large proportion of the heritability for human height. *Nature Genetics*, 42(7):565–569, 2010.

[34] J. Yang, S.H. Lee, M.E. Goddard, and P.M. Visscher. GCTA: a tool for genome-wide complex trait analysis. *The American Journal of Human Genetics*, 88(1):76–82, 2011.

[35] D. Golan, E.S. Lander, and S. Rosset. Measuring missing heritability: Inferring the contribution of common variants. *Proceedings of the National Academy of Sciences*, 111(49):E5272–E5281, 2014.

[36] A.J. Holmes, P.H. Lee, M.O. Hollinshead, L. Bakst, J.L. Roffman, et al. Individual differences in amygdala-medial prefrontal anatomy link negative affect, impaired social functioning, and polygenic depression risk. *The Journal of Neuroscience*, 32(50):18087–18100, 2012.

[37] A.J. Holmes, M.O. Hollinshead, T.M. OKeefe, V.I. Petrov, G.R. Fariello, et al. Brain Genomics Superstruct Project initial data release with structural, functional, and behavioral measures. *Scientific Data*, 2, 2015.

[38] M. Reuter, F.E. Wolter, and N. Peinecke. Laplace-Beltrami spectra as "Shape-DNA" of surfaces and solids. *Computer-Aided Design*, 38(4):342–366, 2006.

[39] M. Reuter, F.E. Wolter, M. Shenton, and M. Niethammer. Laplace-Beltrami eigenvalues and topological features of eigenfunctions for statistical shape analysis. *Computer-Aided Design*, 41(10):739–755, 2009.

[40] C. Wachinger, P. Golland, W. Kremen, B. Fischl, M. Reuter, et al. BrainPrint: a discriminative characterization of brain morphology. *NeuroImage*, 109:232–248, 2015.

[41] D.C. Van Essen, S.M. Smith, D.M. Barch, T.E.J. Behrens, E. Yacoub, et al. The WU-Minn human connectome project: an overview. *NeuroImage*, 80:62–79, 2013.

[42] M. Reuter. Hierarchical shape segmentation and registration via topological features of Laplace-Beltrami eigenfunctions. *International Journal of Computer Vision*, 89(2-3):287–308, 2009.

[43] Z. Lian, A. Godil, B. Bustos, M. Daoudi, J. Hermans, et al. A comparison of methods for non-rigid 3D shape retrieval. *Pattern Recognition*, 46(1):449–461, 2013.

[44] E.E. Bron, M. Smits, W.M. van der Flier, H. Vrenken, F. Barkhof, et al. Standardized evaluation of algorithms for computer-aided diagnosis of dementia based on structural MRI: The CADDementia challenge. *NeuroImage*, 111:562–579, 2015.

[45] L.I.K Lin. A concordance correlation coefficient to evaluate reproducibility. *Biometrics*, 45(1):255–268, 1989.

[46] P.M. Visscher, B. McEvoy, and J. Yang. From Galton to GWAS: quantitative genetics of human height. *Genetics Research*, 92:371–379, 2010.

[47] A. Tenesa and C.S. Haley. The heritability of human disease: estimation, uses and abuses. *Nature Reviews Genetics*, 14(2):139–149, 2013.

[48] M. Muñoz, R. Pong-Wong, O. Canela-Xandri, K. Rawlik, Haley C.S., et al. Evaluating the contribution of genetics and familial shared environment to common disease using the UK Biobank. *Nature Genetics*, 48:980–983, 2016.

[49] O. Zuk, E. Hechter, S.R. Sunyaev, and E.S. Lander. The mystery of missing heritability: Genetic interactions create phantom heritability. *Proceedings of the National Academy of Sciences*, 109(4):1193–1198, 2012.

[50] N. Zaitlen, P. Kraft, N. Patterson, B. Pasaniuc, G. Bhatia, et al. Using extended genealogy to estimate components of heritability for 23 quantitative and dichotomous traits. *PLoS Genet*, 9(5):e1003520, 2013.

[51] N. Zaitlen, B. Pasaniuc, S. Sankararaman, G. Bhatia, J. Zhang, et al. Leveraging population admixture to characterize the heritability of complex traits. *Nature genetics*, 46(12):1356–1362, 2014.

[52] D. Mamah, M.P. Harms, L. Wang, D. Barch, P. Thompson, et al. Basal ganglia shape abnormalities in the unaffected siblings of schizophrenia patients. *Biological Psychiatry*, 64(2):111–120, 2008.

[53] M.P. Harms, L. Wang, D. Mamah, D.M. Barch, P.A. Thompson, et al. Thalamic shape abnormalities in individuals with schizophrenia and their nonpsychotic siblings. *The Journal of Neuroscience*, 27(50): 13835–13842, 2007.

[54] D.C Van Essen. A tension-based theory of morphogenesis and compact wiring in the central nervous system. *Nature*, 385:313–318, 1997.

[55] M. Pievani, S. Galluzzi, P.M. Thompson, P.E. Rasser, M. Bonetti, et al. APOE4 is associated with greater atrophy of the hippocampal formation in Alzheimer's disease. *NeuroImage*, 55(3):909–919, 2011.

[56] J. Yang, T.A. Manolio, L.R. Pasquale, E. Boerwinkle, N. Caporaso, et al. Genome partitioning of genetic variation for complex traits using common SNPs. *Nature Genetics*, 43(6):519–525, 2011.

[57] M. Neale and L. Cardon. *Methodology for genetic studies of twins and families*. Springer, 1992.

[58] L. Almasy and J. Blangero. Multipoint quantitative-trait linkage analysis in general pedigrees. *The American Journal of Human Genetics*, 62(5):1198–1211, 1998.

[59] J.K. Haseman and R.C. Elston. The investigation of linkage between a quantitative trait and a marker locus. *Behavior Genetics*, 2(1):3–19, 1972.

[60] R.C. Elston, S. Buxbaum, K.B. Jacobs, and J.M. Olson. Haseman and Elston revisited. *Genetic Epidemiology*, 19(1):1–17, 2000.

[61] P.M. Visscher, G. Hemani, A.A.E. Vinkhuyzen, G.B. Chen, S.H. Lee, et al. Statistical power to detect genetic (co)variance of complex traits using SNP data in unrelated samples. *PLoS Genetics*, 10(4): e1004269, 2014.

[62] G. Molenberghs and G. Verbeke. Likelihood ratio, score, and Wald tests in a constrained parameter space. *The American Statistician*, 61(1):22–27, 2007.

[63] P.H. Westfall and S.S. Young. *Resampling-based multiple testing: Examples and methods for p-value adjustment*, volume 279. John Wiley & Sons, 1993.

[64] B. Efron and R.J. Tibshirani. *An introduction to the bootstrap*. CRC press, 1994.

[65] T. Ge, T.E. Nichols, P.H. Lee, A.J. Holmes, J.L. Roffman, et al. Massively expedited genome-wide heritability analysis (MEGHA). *Proceedings of the National Academy of Sciences*, 112(8):2479–2484, 2015.

[66] M.F. Glasser, S.N. Sotiropoulos, J.A. Wilson, T.S. Coalson, B. Fischl, et al. The minimal preprocessing pipelines for the Human Connectome Project. *NeuroImage*, 80:105–124, 2013.

[67] C.C. Chang, C.C. Chow, L.C.A.M. Tellier, S. Vattikuti, S.M. Purcell, et al. Second-generation PLINK: rising to the challenge of larger and richer datasets. *GigaScience*, 4(7), 2015.

[68] B. Fischl. Freesurfer. *NeuroImage*, 62(2):774–781, 2012.

[69] L.R. Dice. Measures of the amount of ecologic association between species. *Ecology*, 26(3):297–302, 1945.

End Notes

Acknowledgements. This research was carried out in whole at the Athinoula A. Martinos Center for Biomedical Imaging at the Massachusetts General Hospital (MGH), using resources provided by the Center for Functional Neuroimaging Technologies, P41EB015896, a P41 Biotechnology Resource Grant supported by the National Institute of Biomedical Imaging and Bioengineering (NIBIB), National Institutes of Health (NIH). This work also involved the use of instrumentation supported by the NIH Shared Instrumentation Grant Program; specifically, grant numbers S10RR023043 and S10RR023401.

Data were provided in part by the Brain Genomics Superstruct Project (GSP) of Harvard University and MGH, with support from the Center for Brain Science Neuroinformatics Research Group, Athinoula A. Martinos Center for Biomedical Imaging, Center for Human Genetic Research, and Stanley Center for Psychiatric Research. 20 individual investigators at Harvard and MGH generously contributed data to the overall project.

Data were also provided in part by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University.

This research was also funded in part by NIH grants K25CA181632 (to MR); K01MH099232 (to AJH); K99MH101367 (to PHL); R01NS083534, R01NS070963, and 1K25EB013649-01 (to MRS); K24MH094614 and R01MH101486 (to JWS); an MGH ECOR Tosteson Postdoctoral Fellowship Award (to TG); the Brazilian National Research Council (CNPq), grant number 211534/2013-7 (to AMW); and a BrightFocus Foundation grant AHAF-A2012333 (to MRS). JWS is a Tepper Family MGH Research Scholar.

Competing interests. The authors declare no competing financial interests.