QUANTITATIVE GENETIC MODELING OF LATERAL VENTRICULAR SHAPE AND VOLUME USING MULTI-ATLAS FLUID IMAGE ALIGNMENT IN TWINS

Yi-Yu Chou[1], *Natasha Leporé*[1], *Marina Barysheva*[1], *Ming-Chang Chiang*[1], *Katie McMahon*[2], *Greig de Zubicaray*[2], *Matthew Meredith*[2], *Margaret Wright*[3], *Arthur W. Toga*[1], *Paul M. Thompson*[1]

[1]Laboratory of Neuro Imaging, UCLA Dept. of Neurology, Los Angeles, CA, USA
[2]Centre for Magnetic Resonance, University of Queensland, Brisbane, Australia
[3]Genetic Epidemiology Laboratory, Queensland Institute of Medical Research, Brisbane, Australia

ABSTRACT

Despite substantial progress in measuring the 3D profile of anatomical variations in the human brain, their genetic and environmental causes remain enigmatic. We developed an automated system to identify and map genetic and environmental effects on brain structure in large brain MRI databases[1]. We applied our multi-template segmentation approach ("Multi-Atlas Fluid Image Alignment") to fluidly propagate hand-labeled parameterized surface meshes into 116 scans of twins (60 identical, 56 fraternal), labeling the lateral ventricles. Mesh surfaces were averaged within subjects to minimize segmentation error. We fitted quantitative genetic models at each of 30,000 surface points to measure the proportion of shape variance attributable to (1) genetic differences among subjects, (2) environmental influences unique to each individual, and (3) shared environmental effects. Surface-based statistical maps revealed 3D heritability patterns, and their significance, with and without adjustments for global brain scale. These maps visualized detailed profiles of environmental versus genetic influences on the brain, extending genetic models to spatially detailed, automatically computed, 3D maps.

1. INTRODUCTION

Major efforts in computational anatomy have analyzed statistical variations in neuroanatomy in large populations. Shape- and intensity-based models have successfully identified systematic differences characteristic of diseases such as Alzheimer's, HIV/AIDS, Williams syndrome, and Fragile X syndrome, revealing factors affecting disease progression and its cognitive correlates. Computational anatomy has been successful in encoding the 3D pattern of normal variations in anatomy, but their genetic or environmental causes remain unknown. A first step in distinguishing these effects is to assemble a large database of MRI scans from healthy individuals with known genetic resemblance (e.g., twins). Genetic models can parse the observed variance in image derived signals, establishing regions that are under strongest genetic control.

Computational morphometry studies of disease routinely use 3D statistical maps to localize regions with greatest effects. Genetic studies of brain images have lagged behind these studies, because (1) tools have been lacking to compute and test genetic models at each 3D location in an image, and (2) to detect effects with reasonable power, large imaging databases must be collected from genetically well-characterized populations, e.g. twins.

Several prior studies found that brain structure volumes are heavily genetically determined; frontal gray matter volume is heritable and correlates with IQ [1]. In healthy elderly male twins [2,3], genetic factors accounted for >70% of the variance in intracranial brain volume, white matter hyperintensity volume, and measures of the corpus callosum and lateral ventricle volume. Twins are of special interest for genetic studies because they provide naturally matched pairs where the confounding effects of a large number of potentially causal factors are removed by comparing twins who share them. Twin studies involving both monozygotic (MZ) or genetically identical twins and dizygotic (DZ) twin pairs that share on average half their genes can disentangle genetic from shared and unique environmental effects, identifying quantitative neurodevelopmental phenotypes that may be heritable in the general population.

To assess the relative influence of genes and environment on brain structure, we developed a algorithm pipeline to analyze lateral ventricular shape and volume in 30 MZ and 28 DZ twin pairs by automatically extracting 3D anatomical surface models. After a fluid segmentation using multiple propagated templates [4], quantitative genetic analysis at each of 30,000 surface vertices computed maps of several standard genetic parameters, such as Falconer's heritability estimate [5] and structural equation models capable of variance partitioning. We compared different surface-based maps of genetic statistics using fields of chi-squared statistics to (1) rank the models, (2) ascribe significance values to the model parameters at each surface vertex, and (3) visualize regions where each model (genes, environment, or both) best explained anatomical variance in the population. The overall goal of this work is to zero in on promising measures to screen for effects of candidate genes that may influence brain morphology.

[1] This work was funded in part by NIH grant R01 HD050735.

2. MATERIALS AND METHODS

2.1. Subjects

We analyzed MRI data from 30 pairs (60 scans) of identical (MZ) twins (age: 22-27 years; mean 24.6) and 28 pairs (56 scans) of same-sex fraternal (DZ) twins (age: 21-26 years; mean: 22.9 years), scanned as part of a 5-year study of 1150 healthy twins. The male/female ratio was 30/30 for MZ and 20/36 for DZ. 3D T1-weighted images were acquired with an inversion recovery rapid gradient echo (MP-RAGE) sequence, on a 4 Tesla Bruker Medspec whole body scanner at CMR. Acquisition parameters were: TI/ TR/TE = 1500/2500/3.83 msec; flip angle=15°; slice thickness = 0.9mm; 256^3 acquisition matrix. All images were spatially normalized to the ICBM-53 standard template [7] with a 9-parameter (3 translations, 3 rotations, 3 scales) transformation and, for a separate 'unscaled' analysis, with a 6-parameter rigid-body transformation.

2.2. Automated Lateral Ventricle Segmentation and Shape Modeling

Lateral ventricular volumes were automatically estimated for all scans using a technique we recently validated [4]. A small subgroup of 4 images (2 males (1 MZ, 1 DZ), 2 females (1MZ, 1 DZ)) were randomly chosen and the lateral ventricles were manually traced in contiguous coronal brain sections. Lateral ventricular surface models were created in these images and converted into parametric meshes (we will call these labeled image 'atlases') [8]. We fluidly registered each atlas and the embedded mesh models to all 116 subjects, as in [9], treating the deforming image as a viscous fluid governed by the Navier-Stokes equation, as pioneered by Christensen et al. [10]. Transformations resulting from the fluid registration were applied to the manually traced ventricular boundary using tri-linear interpolation, generating a propagated contour on the unlabeled images. A mesh averaging technique combined the resulting fluidly propagated surface meshes for each image. A medial curve was derived from the line traced out by the centroid of the boundary for each ventricular surface model. The local radial size was defined as the radial distance between a boundary point and its associated medial curve. This allows statistical comparisons of local surface geometry at equivalent 3D surface locations across subjects for subsequent shape and thickness analysis.

2.3. Heritability analyses

Heritability is the proportion of phenotypic variation that is attributable to genetic variation in a population. Variation among individuals may be due to genetic and/or environmental factors. Heritability analyses estimate the relative contributions of differences in genetic and non-genetic factors to the total phenotypic variance in a population. To determine the proportion of variance attributable to genetic influences, heritability analyses were performed for lateral ventricle shape and volume, using two different statistical approaches: (i) classical heritability analysis (Falconer's estimate [5]) and (ii) maximum likelihood estimation (MLE) using path analysis [6]. Both methods are described below. Assuming a polygenic model, a heritability estimate of 0% implies no genetic effects; values close to 100% imply strong genetic influences.

2.3.1. Falconer's Estimate

Intraclass correlation coefficients (ICCs) were computed to determine ventricular shape similarity in both MZ and DZ twin pairs. We applied restricted maximum-likelihood (ReML; [11]) to estimate variance components. This technique outperforms more traditional regression analyses as it accounts for all genetic relationships in the dataset, has less strict assumptions about selection patterns and does not require balanced datasets. ReML is therefore more amenable to data from natural populations. Heritability for human traits is frequently estimated by comparing resemblances between twins. Identical twins are twice as genetically similar as fraternal twins, so heritability is approximately twice the difference in correlation between MZ and DZ twins, $h^2 = 2(r(MZ) - r(DZ))$ - known as Falconer's estimate [5]. The proportion of variance between siblings due to their sharing the environment they are raised in, c^2, is approximated by the MZ correlation minus half the heritability, $c^2 = r(MZ) - h^2$. Unique environmental variance, $e^2 = 1 - r(MZ)$, measures the dissimilarity of identical twins raised together; this variance term also includes measurement errors.

2.3.2. Path Analysis and Structural Equations

Heritability of the lateral ventricular volumes was estimated at each location on the anatomical surfaces by directing the 116-dimensional vector of twin data at each surface vertex into a structural equation model, implemented in the Mx software (version 1.7.03). Briefly, the observed population variance is regarded as arising from the sum of additive genetic variance (A), variance caused by the effects of common environmental factors shared within families (C), and unique environmental variance (E) specific to each individual. The maximum likelihood method was used to estimate the model parameters and expected covariance matrices for both MZ and DZ twins. Then χ^2 values were computed representing the agreement between the observed and expected covariance matrices. These were first used as goodness of fit indices. By convention, a P value *below* 0.05 indicated a *lack* of fit to the data, and led to rejection of the model. The Akaike Information Criterion (AIC), defined as the χ^2 goodness of fit index minus twice the number of degrees of freedom, was used to compare the fit of models including different variance components. **Fig. 1** shows the full twin model as a path diagram.

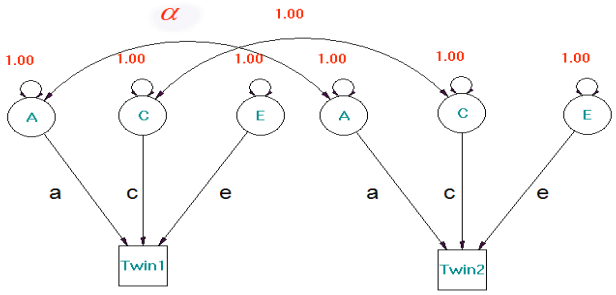

Fig. 1. Structural equation model for the classical ACE design. Three latent variables (*circles*) are denoted A (additive genetic), C (shared environment), and E (unique environment); phenotypes are shown as squares. Single-headed arrows represent causal paths; double-headed arrows represent correlations (a double-headed arrow to the same variable denotes a variance). α represents the genetic correlation between twin pairs (1 for MZ pairs; 0.5 for DZ pairs). Genetic factors increase correlations within MZ twin pairs; common environmental factors increase intrapair correlations for both MZ and DZ twin pairs; unique environmental factors decrease intrapair correlations for both MZ and DZ twin pairs.

3. RESULTS

3.1. Heritability of the Lateral Ventricular Shape

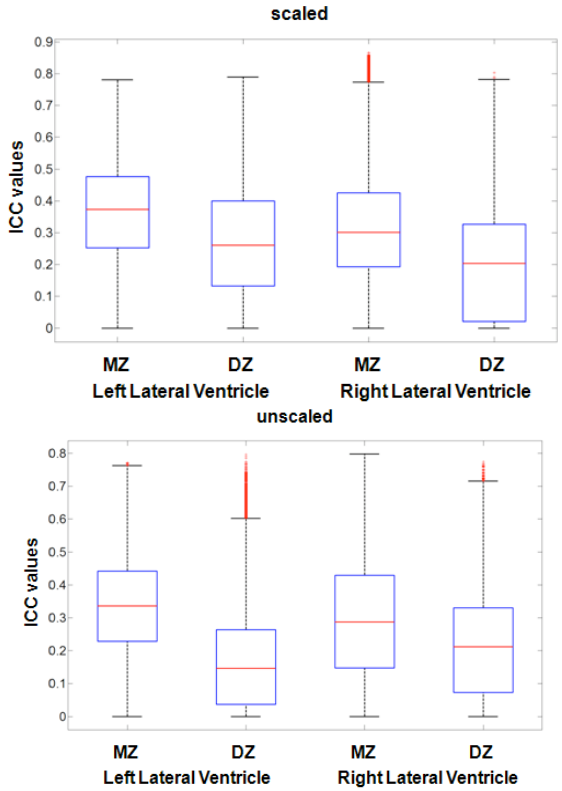

Fig. 2. ICC values of radial ventricular thickness for the left and right lateral ventricles (*top panel*), with and without (*bottom panel*) scaling for brain volume.

Figure 2 shows a boxplot of the intraclass correlations for the radial thicknesses, at all 30000 points on each of the left and right lateral ventricles, with and without adjustment (by registration) for the effects of brain scaling. MZ ICCs were generally higher (0.36 [L], 0.33[R] with scaling; 0.32[L], 0.29 [R] without scaling) than the DZ ones (mean ICC = 0.26 [L], 0.17 [R], with scaling; 0.14[L]; 0.20 [R] without scaling). DZ twins resemble each other less than MZ twins do, which suggests that genetic factors influence lateral ventricular shape differences. Color-coded maps (**Fig. 3**) show local intraclass correlations and associated significance maps for radial thickness at each vertex. Falconer's heritability estimate h^2 is plotted at each surface point (**Fig. 4)**. Genetic control of structure is greatest in the mid-portion of anterior horn and in the posterior horn. **Table 3** shows estimates of a^2, c^2 and e^2 derived from the ACE model of the observed data. The goodness-of-fit statistic is χ^2 distributed. d*f* denotes the number of degrees of freedom. *Large P*-values indicate a good fit between the model and the data. Our maximum-likelihood estimates demonstrate a significant influence of common environmental factors but minimal contribution of genetic factors to lateral ventricular volumes, accounting for 45.63% and 0% of variance in the scaled data, respectively. The ACE model could not be fitted to the unscaled ventricular volumes, due to two MZ outliers (*bottom left*; **Fig. 5**).

4. DISCUSSION

This work complements the pioneering work of Styner et al. [12] who studied genetic influences on ventricular shape using a related modeling approach in 10 MZ and 5 DZ twin pairs. Here we extended it to estimate full structural equation models, plotting genetic and environmental components of variance in 116 twins, based on fluid image segmentation. These high-throughput methods show promise in pinpointing regions for assessing effects of candidate genes on brain morphology.

5. REFERENCES

[1] Narr KL et al., Cerebral Cortex 17(9):2163-2171 (2007).
[2] Carmelli D et al., Stroke 29(6), 1177-1181 (1998).
[3] Pfefferbaum A et al., Neuro. Aging 21(1), 63-74 (2000).
[4] Chou YY et al. NeuroImage (2007).
[5] Falconer DS. Introduction to Quantitative Genetics. Longman, New York, 1981.
[6] Neale MC et al. Mx: Statistical Modeling. Virginia Commonwealth University, Richmond, VA (1999).
[7] Collins DL et al. *JCAT*18(2):192-05 (1994).
[8] Thompson PM et al. *NeuroImage* 3:19-34 (1996).
[9] Gramkow C. *M.Sc. Thesis*, Danish Tech. Univ., (1996)
[10] Christensen GE et al., *IEEE-TIP* 5:1435-1447 (1996).
[11] Lynch M et al., Genetics and Analysis of Quantitative Traits. Sinauer Associates, Sunderland, MA (1998).
[12]. Styner M et al., PNAS 102, 12:4872-4877.

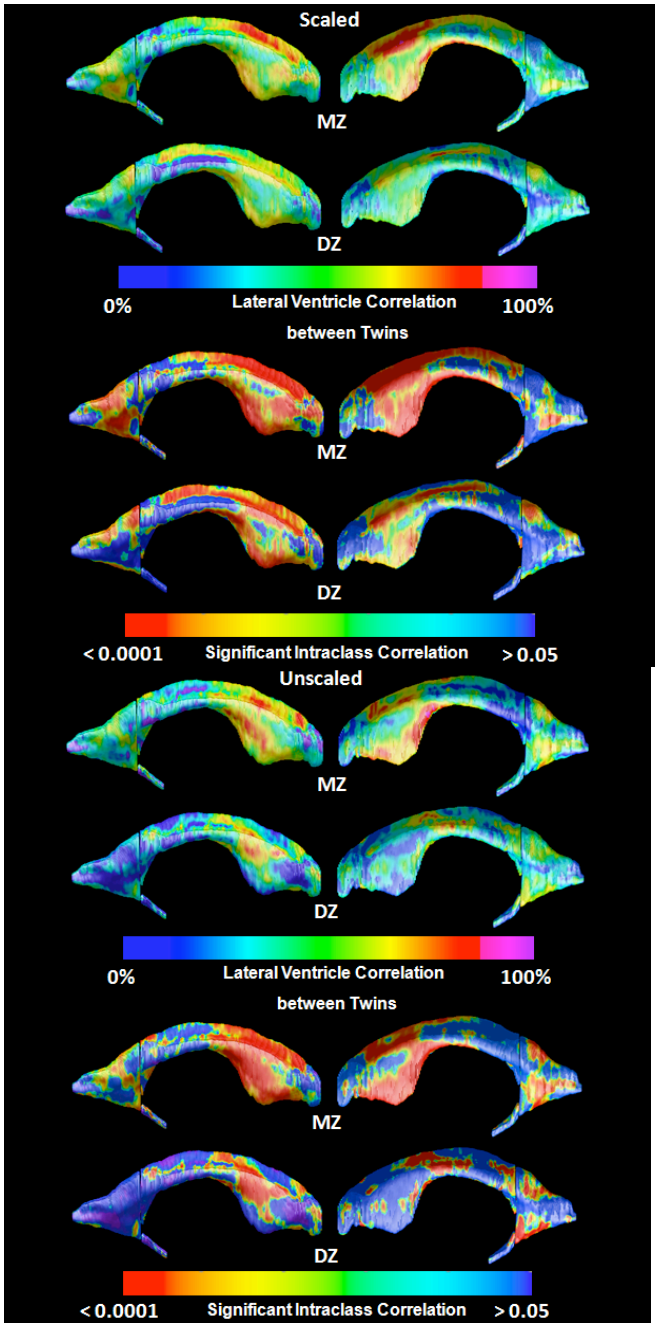

Fig. 3. Maps of twin correlations (and their significance) for radial thickness at each vertex.

	MZ (N = 60)		DZ (N = 56)	
Lateral ventricles	Mean	SD	Mean	SD
Unscaled	18.3	2.5	17.6	1.7
Scaled	23.2	3.7	22.5	4.0

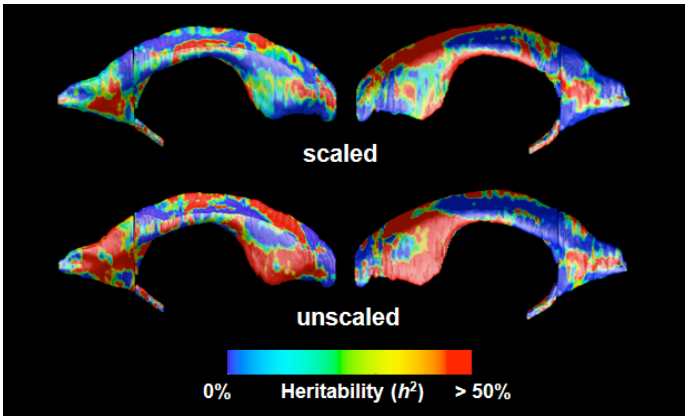

Fig. 4. Color-coded maps for heritability, h^2, of radial thickness at each vertex.

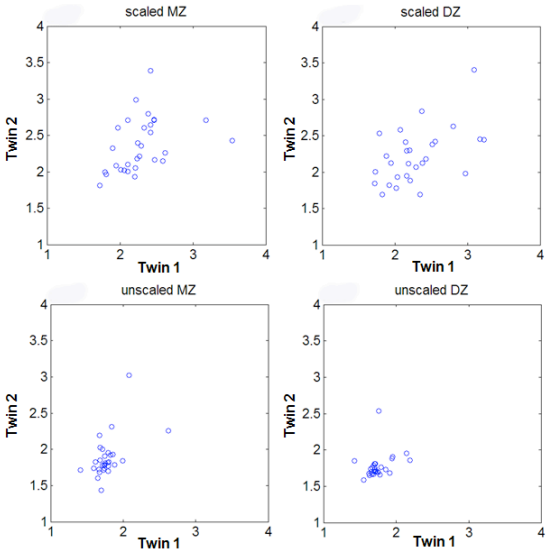

Fig. 5. Scatter plot of ventricular volumes in MZ and DZ twins with and without brain scaling. Each circle represents a twin pair (the order is random). Top left: scaled MZs, top right: scaled DZs, bottom left: unscaled MZs, bottom right: unscaled DZs.

Table 3. ACE Parameter Estimates for Ventricular Volume.

	Heritability Estimates (%)			Model Fit			
	a^2	c^2	e^2	χ^2	p	AIC	df
Unscaled	-	-	-	11.650	0.009	5.650	3
Scaled	0.00	45.63	54.37	1.026	**0.795**	-4.794	3

⇐**Table 2 (left).** Lateral ventricle volumes (in cc); MZ and DZ groups did not differ in average volume (P = 0.36 with scaling; P = 0.09 without scaling).