

# Mapping the Genetic Variation of Regional Brain Volumes as Explained by All Common SNPs from the ADNI Study

Christopher Bryant<sup>1</sup>, Kelly S. Giovanello<sup>2,4</sup>, Joseph G. Ibrahim<sup>1</sup>, Jing Chang<sup>1</sup>, Dinggang Shen<sup>3,4</sup>, Bradley S. Peterson<sup>5</sup>, Hongtu Zhu<sup>1,4\*</sup>, for The Alzheimer's Disease Neuroimaging Initiative<sup>†</sup>

**1** Department of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, United States of America, **2** Department of Psychology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, United States of America, **3** Department of Radiology, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, United States of America, **4** Biomedical Research Imaging Center, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina, United States of America, **5** The Division of Child and Adolescent Psychiatry, The New York State Psychiatric Institute, New York, New York, United States of America

## Abstract

Typically twin studies are used to investigate the aggregate effects of genetic and environmental influences on brain phenotypic measures. Although some phenotypic measures are highly heritable in twin studies, SNPs (single nucleotide polymorphisms) identified by genome-wide association studies (GWAS) account for only a small fraction of the heritability of these measures. We mapped the genetic variation (the proportion of phenotypic variance explained by variation among SNPs) of volumes of pre-defined regions across the whole brain, as explained by 512,905 SNPs genotyped on 747 adult participants from the Alzheimer's Disease Neuroimaging Initiative (ADNI). We found that 85% of the variance of intracranial volume (ICV) ( $p = 0.04$ ) was explained by considering all SNPs simultaneously, and after adjusting for ICV, total grey matter (GM) and white matter (WM) volumes had genetic variation estimates near zero ( $p = 0.5$ ). We found varying estimates of genetic variation across 93 non-overlapping regions, with asymmetry in estimates between the left and right cerebral hemispheres. Several regions reported in previous studies to be related to Alzheimer's disease progression were estimated to have a large proportion of volumetric variance explained by the SNPs.

**Citation:** Bryant C, Giovanello KS, Ibrahim JG, Chang J, Shen D, et al. (2013) Mapping the Genetic Variation of Regional Brain Volumes as Explained by All Common SNPs from the ADNI Study. PLoS ONE 8(8): e71723. doi:10.1371/journal.pone.0071723

**Editor:** Stefano L. Sensi, University G. D'Annunzio, Italy

**Received:** May 17, 2013; **Accepted:** July 10, 2013; **Published:** August 28, 2013

**Copyright:** © 2013 Bryant et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Funding:** Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health (NIH) grant U01 AG024904). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: Abbott; Alzheimer's Association; Alzheimer's Drug Discovery Foundation; Amofix Life Sciences Ltd.; AstraZeneca; Bayer HealthCare; BioClinica, Inc.; Biogen Idec Inc.; Bristol-Myers Squibb Company; Eisai Inc.; Elan Pharmaceuticals Inc.; Eli Lilly and Company; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; GE Healthcare; Innogenetics, N.V.; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Medpace, Inc.; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Servier; Synarc Inc.; and Takeda Pharmaceutical Company. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health ([www.fnih.org](http://www.fnih.org)). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Disease Cooperative Study at the University of California, San Diego. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of California, Los Angeles. This research was also supported by NIH grants P30 AG010129 and K01 AG030514. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing Interests:** Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI), which includes funds from commercial sources (Abbott, Amofix Life Sciences Ltd.; AstraZeneca; Bayer HealthCare; BioClinica, Inc.; Biogen Idec Inc.; Bristol-Myers Squibb Company; Eisai Inc.; Elan Pharmaceuticals Inc.; Eli Lilly and Company; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; GE Healthcare; Innogenetics, N.V.; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Medpace, Inc.; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Servier; Synarc Inc.; and Takeda Pharmaceutical Company). This does not alter the authors' adherence to all the PLOS ONE policies on sharing data and materials.

\* E-mail: [hzhu@bios.unc.edu](mailto:hzhu@bios.unc.edu)

† Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database ([adni.loni.ucla.edu](http://adni.loni.ucla.edu)). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: [http://adni.loni.ucla.edu/wp-content/uploads/how\\_to\\_apply/ADNI\\_Acknowledgement\\_List.pdf](http://adni.loni.ucla.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf)

## Introduction

The heritability of brain structure and function has received extensive attention for many years, and it continues to be an area of active research [1–4]. Many studies have used twin designs to estimate the heritability of brain structure and function. Twin studies typically compare the similarity of monozygotic twins (MZ), who share the same genetic materials, to that of dizygotic twins (DZ), who share 50% of their genes. Assuming equal

environmental exposure across zygosity, the known differences in genetic similarity allow us to distinguish the effects of genes and environment on a phenotype, including a phenotype characterized using brain imaging measures.

Genome-wide association studies (GWAS) have discovered hundreds of single-nucleotide polymorphisms (SNPs) that are significantly associated with variation in brain structure and function [5–11]. A study using the ADNI database found two SNPs associated with temporal lobe volume [12] while two other

ADNI studies found no SNPs associated with voxelwise volume differences, after multiple comparison adjustment [13], and no significant genes in a multivariate analysis combining the SNPs into genes [14]; five SNPs were found to be significant in the ADNI population for 142 various imaging phenotypes [15]. In cases where significant SNPs are found in GWAS, however, those SNPs typically explain only a small portion of overall phenotypic variation. A question that has not been addressed fully for phenotypic brain measures is where in the genome the missing heritability resides. Explanations could include that causal variants each may explain such a small amount of variation that their effects do not reach stringent significance thresholds and/or that the causal variants may not be in complete linkage disequilibrium (LD) with the SNPs that have been genotyped. Heritability is defined as the proportion of observed phenotypic variation that is due to inherited genetic factors; GCTA (genome-wide complex trait analysis) software allows us to approximate heritability with the proportion of phenotypic variance explained by variation among over 500,000 SNPs, after adjusting for environmental factors. GCTA has been utilized previously to partition the genetic variation of complex traits such as height and BMI into contributions from each chromosome [16] and to estimate the missing heritability of disease from GWAS [17].

Here we estimate the proportion of variance of 93 regional volumes, as well as ICV, GM, WM, and cerebro-spinal fluid (CSF) volumes, explained by 512,905 SNPs across the 22 autosomes genotyped on 747 adult participants from the ADNI study.

## Materials and Methods

### Sample and Data

The NIH ADNI is an ongoing public-private partnership to test whether genetic data, structural and functional neuroimaging, and clinical data can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimer's disease (AD). The structural brain MRI data and corresponding clinical and genetic data from baseline and follow-up were downloaded from the ADNI publicly available database (<http://adni.loni.ucla.edu/>). This initiative is a collective effort led by Principal Investigator Michael Weiner, M.D., VA Medical Center and University of California – San Francisco, involving many co-investigators and recruitment of participants from over 50 sites in the United States and Canada. More information on the study can be found at [www.adni-info.org](http://www.adni-info.org).

The Human 610-Quad BeadChip (Illumina, Inc., San Diego, CA) was used to genotype 818 participants in the ADNI database, which resulted in a set of 620,901 SNP and copy number variation (CNV) markers. Since the Apolipoprotein E (APOE) SNPs, rs429358 and rs7412, are not on the Human 610-Quad BeadChip, they were genotyped separately. These two SNPs together define a 3 allele haplotype, namely the e2, e3, and e4 variants, and the presence of each of these variants was available in the ADNI database for all the individuals; the APOE gene has been the most significant risk locus in GWAS of AD [18]. The software EIGENSTRAT in the package of EIGENSOFT 3.0 was used to calculate the population stratification coefficients of all participants. To reduce population stratification effects, we initially used 747 Caucasians out of the total 818 participants.

The MRI data were collected across a variety of 1.5 Tesla MRI scanners with protocols individualized for each scanner, including volumetric 3-dimensional sagittal MPRAGE or equivalent protocols with varying resolutions. The typical protocol included: repetition time (TR) = 2400 ms, inversion time (TI) = 1000 ms, flip angle = 8°, field of view (FOV) = 24 cm, with a 256\*256\*170

acquisition matrix in the x-,y-, and z-dimensions yielding a voxel size of 1.25\*1.26\*1.2 mm<sup>3</sup>. All original uncorrected image files are available to the general scientific community, as described at <http://www.loni.ucla.edu/ADNI>.

### Image preprocessing and analysis

The MRI data were preprocessed using standard procedures that included realignment to the anterior commissure and posterior commissure by using MIPAV software, skull-stripping by using Brain Surface Extractor (BSE) and Brain Extraction Tool (BET), cerebellum removal, intensity inhomogeneity correction, segmentation using the FSL-FAST software, and spatial coregistration by using HAMMER [19–21]. Particularly, to establish the longitudinal correspondences in the individual and the inter-subject correspondences between the template and the individual, we combined information across time points for each subject and registered the images to a Jacob template [22] using a fully automatic 4-dimensional atlas warping method called 4D HAMMER [21]. Regional volumetric measurements and analyses are then performed via measurements and analyses of the resulting tissue density maps. Lastly, we carried out automatic regional labeling: first, by labeling the template image and second, by transferring the labels following the deformable registration of subject images. After labeling 93 regions from the Jacob atlas, we were able to compute volumes for each of these regions for each subject.

### Statistical Methods

GCTA (genome-wide complex trait analysis) [23] software running a linear mixed model (LMM) was used to estimate the proportion of variance of an observed phenotype that was explained by a set of SNPs (rather than a single SNP, as in GWAS), which we call genetic variation for simplicity. Here the phenotypes of interest are the 93 regional volumes as well as total intracranial (ICV), grey matter (GM), white matter (WM), and cerebrospinal fluid (CSF) volumes, and the SNPs numbered 512,905 from 747 participants. Genetic relationship matrices (GRMs) captured correlations across participants and were then used to partition the phenotypic variance (see [23] for details). Via GCTA, we fitted LMM separately to each of the 97 volumes as dependent variables to estimate their genetic variation, as explained by the 512,905 SNPs through GRMs. The volumes were standardized to better fit the normal distribution assumed for the LMM, and baseline age, gender, and the interaction of age with gender were included as covariates, to capture variation due to environmental effects. To adjust for population structure, we also included subjects' first ten principal components of the GRM as covariates. ICV was also included as a covariate (for regions other than ICV itself) to remove any potential scaling effects. Within GCTA, we corrected for imperfect LD between the SNPs on the array and causal variants by assuming the same allelic distribution (i.e. minor allele frequency) for those causal variants (which were potentially not in the SNP array) as that observed in the included SNPs. See [24] for details on how the correction is implemented within the GCTA software. This correction allows for better approximation to true heritability within this observational study.

An R package, called APcluster, was used to group standardized volumes into groups of similar regions based on the GCTA measures. Affinity propagation (AP) clustering [25–26] was applied first to the 93 non-overlapping regional (scaled and standardized) volumes separately for each diagnostic group. Clustering was based on the negative squared Euclidean similarity matrix between the regional volumes, finding “exemplars” that

**Table 1.** Participants' gender and baseline age by study group.

	<u>Gender</u>		<u>Total</u>	<u>Age</u>					
	<u>Male</u>	<u>Female</u>		<u>Min</u>	<u>1st quartile</u>	<u>Median</u>	<u>Mean</u>	<u>3rd quartile</u>	<u>Max</u>
<b>Normal</b>	111	95	206	60	73	76	76	79	90
<b>MCI</b>	233	129	362	55	71	75	75	80	90
<b>Alzheimer's</b>	98	81	179	55	71	77	75.5	81	91
<b>Total</b>	442	305	747	55	71	76	75.4	80	91

Gender and baseline age distribution by study group. Chi-squared test of independence between gender and study group yields a p-value of 0.02. ANOVA F-test for differences in mean age between study groups yields a p-value of 0.18.

doi:10.1371/journal.pone.0071723.t001

were the centers of the clusters and iteratively determining the number and membership of the clusters to maximize their "net similarity," an objective function that measures how representative the exemplars are of the data in each cluster. This technique was then applied to cluster the genetic variation estimates for the 93 non-overlapping regions. The vector of the proportions of volumetric variance explained by all SNPs (after adjusting for covariates) for the 93 non-overlapping regions was clustered based on the squared distances of these genetic variation estimates. Fisher's exact tests were used to compare clustering results: pair-wise between study groups and volumetric clustering compared to the clustering of genetic variation estimates.

## Results

Table 1 gives demographic information on the participants used in these analyses. Age is similarly distributed for each of the three study groups and there is a significant association between gender and study group ( $p = 0.02$ ); both age and gender were adjusted for in each analysis. Tables 2 and 3 present genetic variation estimates, their standard errors, the p-values from the associated likelihood ratio tests (LRTs), and clustering results of the estimates for all regions. Figure 1 shows genetic variation estimates (top left) and the corresponding  $-\log_{10}p$ -values from the likelihood ratio tests (top right) for the standardized volumes of each non-overlapping region as explained by the 512,905 SNPs across the entire genome, after adjusting for covariates. Figure 2 shows genetic variation estimates in three dimensions. Genetic variation estimates were found to be uncorrelated with average region size ( $p = 0.2$ ; from GLM with logit link function, Normal distribution).

The AP cluster method identified 9, 10, and 10 clustering groups within the 93 non-overlapping regional volumes for the Control, MCI, and AD groups, respectively. Fisher's exact tests of independence yielded  $p < 0.001$  for each of the pairwise comparisons of clustering membership, indicating similarity between the volumetric clustering results across the three groups. Subsequently all participants were combined into a single analysis, yielding 9 clusters of regional volumes. In addition, AP clustering identified 6 groups of (non-overlapping) regions based on genetic contributions to variation in their standardized, covariate-adjusted volumes (see Table 2). Figure 1 shows the clustering map (bottom) for all of the non-overlapping regions. A Fisher's exact test yielded a p-value of 0.22 when testing for independence between the two sets of clustering results (from the raw volumes and from the genetic variation estimates), leading us to conclude that the groups that are similar in volume are not necessarily similar in the proportion of volumetric variance explained by the SNPs.

## Discussion

GCTA findings showed that about 85% of intra-cranial volume (ICV) variability and 57% of cerebrospinal fluid (CSF) volume variability were explained by genetic variation within these participants, with likelihood ratio tests (LRT) for the significance of the random component yielding p-values of 0.04 and 0.12, respectively. However, the genetic variation estimates for total grey matter (GM) and white matter (WM) volumes (after adjusting for ICV and other covariates) were both near zero. Previous studies of adults reported similar estimates for ICV heritability (73% in [27]; 81% in [28]; 79% in [29]; See [30] for a review). Carmelli et al. [27] estimated heritability of CSF volumes to be 72% in their study of normal elderly twins, similar to our estimates. We found markedly less genetic variation in our sample for GM and WM volumes than did previous studies (WM: 73% heritable in [27]; 62% in [29]; 69%–82% in WM regions and 55–85% in GM regions in [4]), likely because of our adjusting for ICV to remove scaling effects. Analysis of GM and WM volumes, without adjusting for ICV, yielded genetic variation estimates of 41% (LRT p-value = 0.21) and 91% (LRT p-value = 0.02) respectively, which are more consistent with published heritability estimates, indicating that some of the previously published heritability estimates for these volumes may be due to the heritability of overall brain volume rather than genetic contributions that are specific to GM or WM.

In regional analyses, we found estimates that were asymmetric across hemispheres, with no clear systematic difference between left hemisphere and right hemisphere volumes. This is in agreement with one prior report [31] and in a twin study of schizophrenia [4] but at odds with another study that reported symmetric heritability in only 10 monozygotic and 10 dizygotic twin pairs, a sample that likely afforded insufficient statistical

**Table 2.** Genetic Variation estimates for major regional brain volumes.

<u>Region</u>	<u>Estimated genetic variation</u>	<u>Standard error</u>	<u>LRT p-value</u>
ICV	0.845	0.457	0.04
GM Volume	<0.001	0.476	0.5
WM Volume	<0.001	0.483	0.5
CSF Volume	0.574	0.468	0.12

Genetic variation estimates, standard errors, and associated likelihood ratio tests for four aggregated volumes.

doi:10.1371/journal.pone.0071723.t002

**Table 3.** Genetic variation estimates and associated clustering results for ROI volumes.

<b>REGION</b>	<b>Genetic Variation</b>	<b>SE</b>	<b>LRT p-value</b>	<b>Cluster</b>
LATERAL VENTRICLE RIGHT	0.999	0.468	0.017	1
GLOBUS PALLADUS RIGHT	0.907	0.440	0.024	1
CAUDATE NUCLEUS RIGHT	0.999	0.459	0.006	1
CUNEUS LEFT	0.999	0.462	0.010	1
NUCLEUS ACCUMBENS LEFT	0.894	0.494	0.053	1
LATERAL VENTRICLE LEFT	0.729	0.476	0.080	2
CAUDATE NUCLEUS LEFT	0.679	0.458	0.073	2
TEMPORAL LOBE WM RIGHT	0.743	0.473	0.068	2
OCCIPITAL LOBE WM LEFT	0.754	0.450	0.055	2
SUPERIOR PARIETAL LOBULE RIGHT	0.822	0.442	0.038	2
LATERAL OCCIPITOTEMPORAL GYRUS RIGHT	0.833	0.472	0.046	2
ENTORHINAL CORTEX RIGHT	0.785	0.442	0.040	2
CUNEUS RIGHT	0.766	0.455	0.050	2
INSULA RIGHT	0.595	0.449	0.094	3
PRECENTRAL GYRUS RIGHT	0.589	0.478	0.117	3
MEDIAL FRONTAL GYRUS LEFT	0.563	0.468	0.120	3
GLOBUS PALLADUS LEFT	0.581	0.483	0.127	3
PUTAMEN RIGHT	0.590	0.470	0.108	3
SUBTHALAMIC NUCLEUS RIGHT	0.586	0.465	0.105	3
OCCIPITAL LOBE WM RIGHT	0.628	0.473	0.102	3
PRECUNEUS LEFT	0.477	0.498	0.191	3
SUPERIOR FRONTAL GYRUS LEFT	0.626	0.462	0.094	3
POSTCENTRAL GYRUS LEFT	0.467	0.476	0.169	3
PERIRHINAL CORTEX RIGHT	0.544	0.464	0.122	3
POSTCENTRAL GYRUS RIGHT	0.569	0.459	0.110	3
LINGUAL GYRUS RIGHT	0.542	0.446	0.108	3
SUPERIOR TEMPORAL GYRUS RIGHT	0.499	0.486	0.165	3
FORNIX RIGHT	0.472	0.451	0.142	3
MIDDLE FRONTAL GYRUS RIGHT	0.263	0.491	0.306	4
INFERIOR FRONTAL GYRUS LEFT	0.266	0.456	0.276	4
ANGULAR GYRUS RIGHT	0.269	0.483	0.296	4
FRONTAL LOBE WM LEFT	0.274	0.475	0.286	4
POSTERIOR LIMB OF INTERNAL CAPSULE INC. CEREBRAL PEDUNCLE LEFT	0.334	0.492	0.261	4
POSTERIOR LIMB OF INTERNAL CAPSULE INC. CEREBRAL PEDUNCLE RIGHT	0.411	0.486	0.212	4
SUPERIOR PARIETAL LOBULE LEFT	0.320	0.478	0.260	4
PARIETAL LOBE WM LEFT	0.276	0.499	0.303	4
PRECENTRAL GYRUS LEFT	0.327	0.490	0.263	4
MEDIAL FRONT-ORBITAL GYRUS LEFT	0.392	0.456	0.195	4
PARIETAL LOBE WM RIGHT	0.291	0.474	0.273	4
PARAHIPPOCAMPAL GYRUS RIGHT	0.230	0.466	0.310	4
OCCIPITAL POLE RIGHT	0.288	0.472	0.275	4
INFERIOR TEMPORAL GYRUS RIGHT	0.293	0.474	0.270	4
MEDIAL FRONT-ORBITAL GYRUS RIGHT	0.176	0.468	0.353	5
SUPERIOR FRONTAL GYRUS RIGHT	0.153	0.497	0.386	5
PUTAMEN LEFT	0.182	0.465	0.346	5
PARAHIPPOCAMPAL GYRUS LEFT	0.112	0.483	0.411	5
FORNIX LEFT	0.205	0.482	0.341	5
PRECUNEUS RIGHT	0.148	0.489	0.386	5
SUPERIOR OCCIPITAL GYRUS RIGHT	0.135	0.453	0.379	5
SUPRAMARGINAL GYRUS LEFT	0.157	0.482	0.376	5

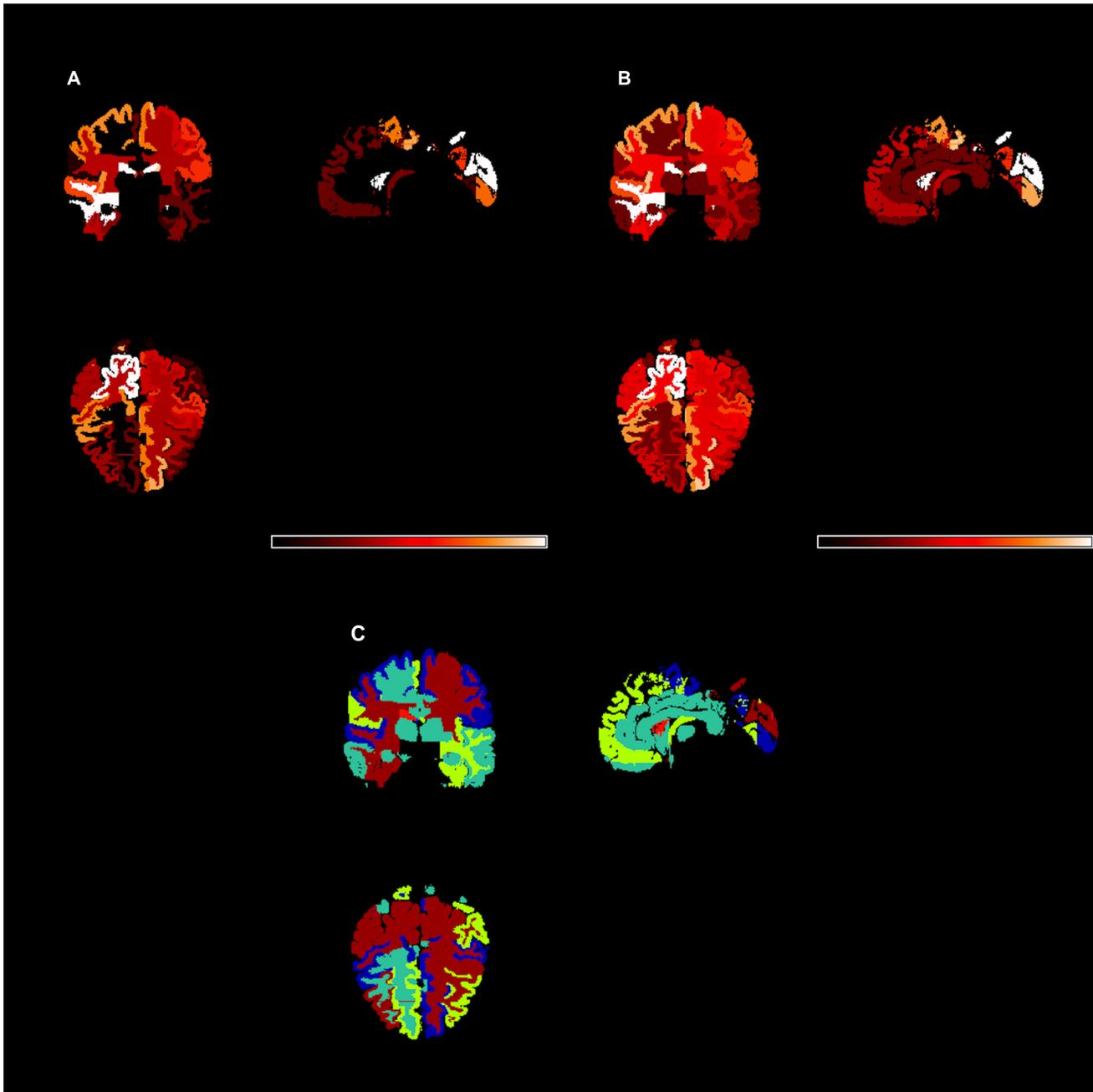
**Table 3. Cont.**

<b>REGION</b>	<b>Genetic Variation</b>	<b>SE</b>	<b>LRT p-value</b>	<b>Cluster</b>
MIDDLE FRONTAL GYRUS LEFT	0.160	0.481	0.373	5
SUPRAMARGINAL GYRUS RIGHT	0.085	0.489	0.434	5
INFERIOR FRONTAL GYRUS RIGHT	0.179	0.476	0.356	5
TEMPORAL LOBE WM LEFT	0.183	0.475	0.352	5
LATERAL FRONT-ORBITAL GYRUS LEFT	0.188	0.474	0.348	5
INSULA LEFT	0.156	0.458	0.365	5
MEDIAL FRONTAL GYRUS RIGHT	0.190	0.467	0.343	5
ANGULAR GYRUS LEFT	0.122	0.462	0.395	5
MEDIAL OCCIPITOTEMPORAL GYRUS RIGHT	0.131	0.473	0.392	5
LATERAL OCCIPITOTEMPORAL GYRUS LEFT	0.217	0.470	0.323	5
OCCIPITAL POLE LEFT	0.104	0.486	0.419	5
LATERAL FRONT-ORBITAL GYRUS RIGHT	<0.001	0.458	0.500	6
CINGULATE REGION RIGHT	<0.001	0.478	0.500	6
FRONTAL LOBE WM RIGHT	<0.001	0.471	0.500	6
TEMPORAL POLE RIGHT	<0.001	0.484	0.500	6
NUCLEUS ACCUMBENS RIGHT	0.059	0.488	0.454	6
UNCUS RIGHT	<0.001	0.490	0.500	6
CINGULATE REGION LEFT	<0.001	0.472	0.500	6
SUBTHALAMIC NUCLEUS LEFT	<0.001	0.466	0.500	6
HIPPOCAMPAL FORMATION RIGHT	<0.001	0.498	0.500	6
INFERIOR OCCIPITAL GYRUS LEFT	<0.001	0.457	0.500	6
ANTERIOR LIMB OF INTERNAL CAPSULE LEFT	<0.001	0.486	0.500	6
SUPERIOR TEMPORAL GYRUS LEFT	<0.001	0.472	0.500	6
UNCUS LEFT	<0.001	0.473	0.500	6
MIDDLE OCCIPITAL GYRUS RIGHT	<0.001	0.471	0.500	6
MIDDLE TEMPORAL GYRUS LEFT	<0.001	0.476	0.500	6
LINGUAL GYRUS LEFT	<0.001	0.425	0.500	6
PERIRHINAL CORTEX LEFT	<0.001	0.478	0.500	6
INFERIOR TEMPORAL GYRUS LEFT	<0.001	0.495	0.500	6
TEMPORAL POLE LEFT	<0.001	0.479	0.500	6
ENTORHINAL CORTEX LEFT	<0.001	0.491	0.500	6
INFERIOR OCCIPITAL GYRUS RIGHT	<0.001	0.463	0.500	6
SUPERIOR OCCIPITAL GYRUS LEFT	0.060	0.473	0.450	6
HIPPOCAMPAL FORMATION LEFT	<0.001	0.479	0.500	6
THALAMUS LEFT	<0.001	0.474	0.500	6
AMYGDALA LEFT	<0.001	0.484	0.500	6
MEDIAL OCCIPITOTEMPORAL GYRUS LEFT	<0.001	0.478	0.500	6
ANTERIOR LIMB OF INTERNAL CAPSULE RIGHT	<0.001	0.486	0.500	6
MIDDLE TEMPORAL GYRUS RIGHT	<0.001	0.480	0.500	6
CORPUS CALLOSUM	<0.001	0.488	0.500	6
AMYGDALA RIGHT	<0.001	0.505	0.500	6
MIDDLE OCCIPITAL GYRUS LEFT	<0.001	0.460	0.500	6
THALAMUS RIGHT	<0.001	0.473	0.500	6

Genetic variation estimates, standard errors, associated LRT p-values, and clustering results for 93 non-overlapping ROIs.  
doi:10.1371/journal.pone.0071723.t003

power to detect differences in heritability across hemispheres [32]. Chen et al. [1] found bilateral symmetry across hemispheres in their mapping of the brain's cortical surface, but their study involved 406 healthy twins and did not consider regional volumes. We found the left and right cuneus to have significant genetic

variation (estimates of 99%,  $p=0.01$  and 77%,  $p=0.05$ , respectively); Niskanen et al. [33] found the cuneus to be involved in the progression of AD, which could explain the genetic variation in our study population. We also found the right caudate nucleus to have significant genetic variation (estimate of 99%,



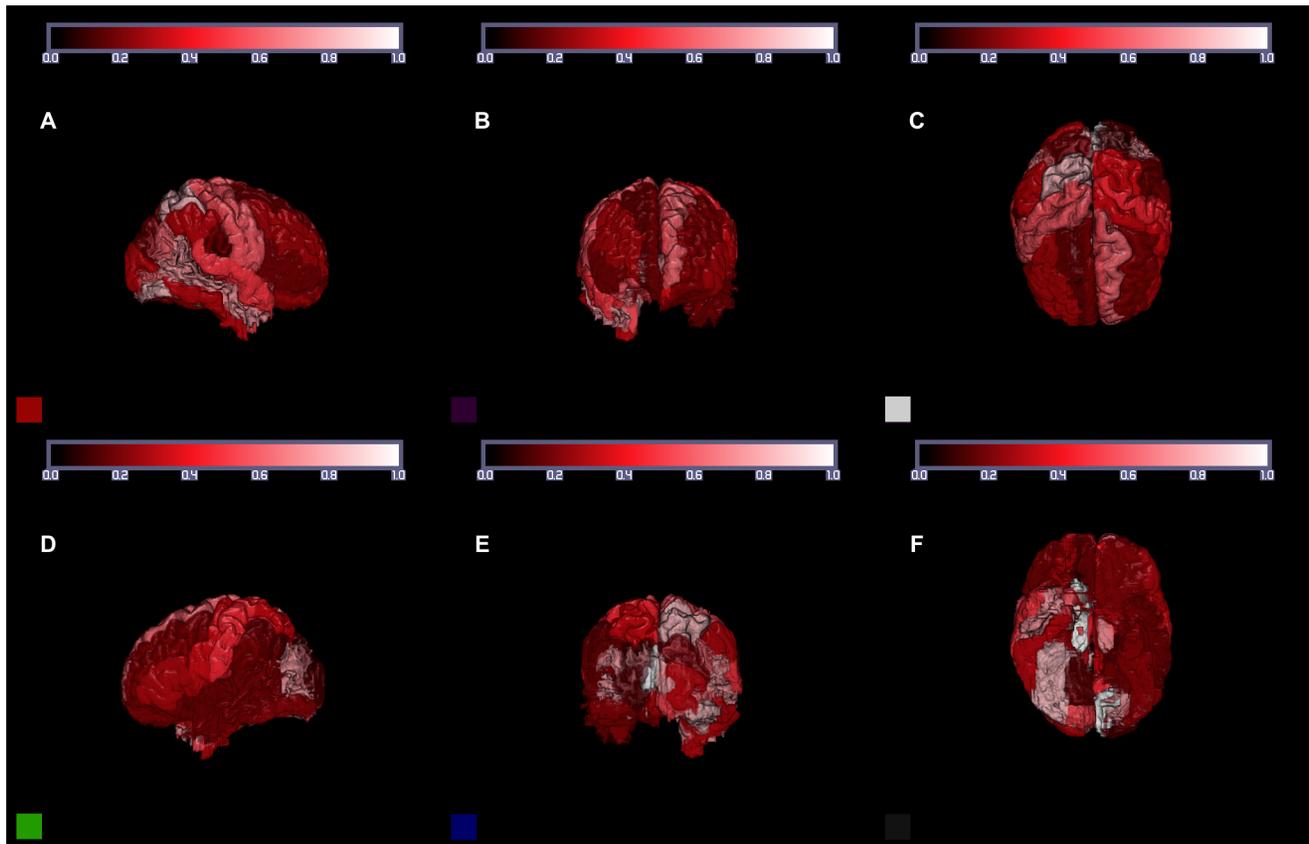
**Figure 1. Genetic variation estimates and additional results for non-overlapping brain regions.** Genetic variation estimates (top left; A) and the associated  $-\log_{10}p$ -values from LRT (top right; B). Hotter colors (black<red<white) indicate larger values. Clusters of genetic variation estimates (below center; C) using AP cluster method. doi:10.1371/journal.pone.0071723.g001

$p = 0.02$ ; left caudate nucleus: estimate of 68%,  $p = 0.07$ ); this is a region found previously to be significantly smaller in AD patients than in normal controls [34], including in the ADNI population [35], suggesting that size differences in the right hemisphere in AD may be heritable.

The AP cluster analyses indicated the presence of 6 groups of brain regions within the genetic variation estimates; the largest group was comprised of those regions with little to no volumetric variability explained by the SNPs. The clusters of regional volumes differed significantly from the clusters of the corresponding genetic variation estimates, but average regional volumes were not

associated significantly with genetic variation estimates. Thus, the regions that were similar in size were not necessarily similar in genetic variation, and there was no apparent relationship between a region's size and the genetic variation of the region.

In the current study, we systematically estimated the variability of 93 non-overlapping regional volumes, ICV, and GM, WM, and CSF volumes explained by 512,905 SNPs across the 22 autosomes genotyped on 747 adult participants from ADNI. These results give some evidence for a genetic component of the variation between individuals in several functional brain systems.



**Figure 2. 3-dimensional map of genetic variation estimates.** Views (clockwise from top left): right lateral (A), anterior (B), superior (C), inferior (D), posterior (E), and left lateral (F). Hotter colors (black<red<white) indicate larger values.  
doi:10.1371/journal.pone.0071723.g002

First, there were several medial/mid-line brain structures with high genetic variation estimates, including the bilateral medial frontal-orbital gyrus, left medial frontal gyrus, bilateral precuneus, right perirhinal cortex and right entorhinal cortex. These structures, along with other medial neural regions, have been characterized as being part of the default mode network (DMN) - defined as a set of functionally connected brain regions that exhibit task-induced deactivation and increase activity at rest [36–37]. DMN changes have been observed in MCI and AD, particularly in the posterior region of medial parietal cortex referred to as the precuneus [38]. Additionally, studies of resting glucose metabolism have demonstrated hypometabolism in the inferior parietal lobule that progresses with AD and correlates with mental status [39–40] and is present in individuals at genetic risk for AD [41].

Second, many subcortical nuclei, as well as white matter tracts, had large genetic variation estimates, including the right insula, bilateral globus pallidus, bilateral putamen, bilateral caudate and left nucleus accumbens. Many of these structures make up the basal ganglia (i.e., globus pallidus, putamen, caudate nucleus, nucleus accumbens), a group of nuclei that act as a cohesive functional unit. The basal ganglia are associated with a variety of functions, including motor control, procedural learning relating to routine behaviors or “habits”, eye movements, and emotional functions [42]. Additionally, several theories implicate the basal ganglia in action selection (i.e., deciding which of several possible behaviors to execute at a given time [43–44]). Beyond the basal ganglia, another subcortical region with high estimated genetic

variation was the insula, a structure believed to play a role in functions usually linked to emotion (i.e., disgust) or the regulation of the body’s homeostasis. Although changes in subcortical regions are not the hallmark of MCI or AD, as AD is considered a cortical dementia, the large estimates observed in the current study indicate that future research in this area is warranted.

Finally, large genetic variation estimates were observed for two prominent perceptual cortical pathways, namely the dorsal “where” pathway and the ventral “what” pathway. First described by Ungerlieder and Mishkin [45], the dorsal pathway leads from striate cortex to the parietal lobe and is responsible for determining an object’s location in space, whereas the ventral pathway leads from striate cortex to the temporal lobe and is responsible for determining an object’s identity. Via an interaction between bottom-up (visual perceptual information) and top-down (knowledge information) processing these pathways are utilized to rapidly to perceive “what” objects are in an environment and “where” those objects are spatially located relative to the observer. Given the clear adaptive nature of such visual processing streams, it is perhaps unsurprising that large genetic variation estimates were observed for several key regions that comprise these pathways, including bilateral occipital white matter tracts, bilateral occipital pole, bilateral temporal lobe white matter tracts, bilateral parietal lobe white matter tracts, bilateral superior parietal lobule and bilateral occipito-temporal gyrus.

Due to a lack of power in the likelihood-ratio testing procedure and a sample size that was not sufficiently large to make standard

errors small, we did not find estimates to be highly statistically significant. We did not perform multiple comparisons adjustments on the p-values from these analyses, in order to preserve the ability to compare results across different brain regions. Here our goal is to map genetic variation across the entire brain, where the results of particular interest are the comparisons of genetic variation among different regional volumes. The mixed model approach utilized by GCTA allows for an approximation to total heritability based on a large number of SNPs; this methodology is validated by the comparison of our results to published twin studies that estimate “true” heritability and have similar findings. Estimates of 0.999 for three regions do not imply that over 99% of variation in those volumes is explained by genetic variability, but they likely indicate large true parameter values (i.e. volumetric variance explained by the SNPs) and some additional variation due to the (case-control) sampling scheme, as well as noise in the volumetric data.

## References

- Chen CH, Gutierrez ED, Thompson W, Panizzon MS, Jernigan TL, et al. (2012) Hierarchical genetic organization of human cortical surface area. *Science* 335 (6076):1634–6. PubMed PMID:22461613.
- Gilmore JH, Schmitt JE, Knickmeyer RC, Smith JK, Lin W, et al. (2010) Genetic and environmental contributions to neonatal brain structure: A twin study. *Hum Brain Mapp* (8):1174–82. PubMed PMID:20063301.
- Yoon U, Perusse D, Lee JM, Evans AC (2011) Genetic and environmental influences on structural variability of the brain in pediatric twin: deformation based morphometry. *Neurosci Lett* 493 (1–2):8–13. PubMed PMID:21296128.
- Hulshoff Pol HE, Schnack HG, Mandl RC, Brans RG, van Haren NE, et al. (2006) Gray and white matter density changes in monozygotic and same-sex dizygotic twins discordant for schizophrenia using voxel-based morphometry. *Neuroimage* (2):482–8. PubMed PMID:16497519.
- Ikrum MA, Fornage M, Smith AV, Seshadri S, Schmidt R, et al. (2012) Cohorts for Heart and Aging Research in Genomic Epidemiology Consortium. Common variants at 6q22 and 17q21 are associated with intracranial volume. *Nat Genet* 44 (5):539–44. PubMed PMID:22504418.
- Biffi A, Anderson CD, Desikan RS, Sabuncu M, Cortellini L, et al. (2010) Genetic variation and neuroimaging measures in Alzheimer disease. *Arch Neurol* 67 (6):677–85. PubMed PMID:20558387.
- Erten-Lyons D, Jacobson A, Kramer P, Grube A, Kaye J (2010) The FAS gene, brain volume, and disease progression in Alzheimer’s disease. *Alzheimers Dement* (2):118–24. PubMed PMID:19766542.
- Baranzini SE, Wang J, Gibson RA, Galwey N, Naegelin Y, et al. (2009) Genome-wide association analysis of susceptibility and clinical phenotype in multiple sclerosis. *Hum Mol Genet* 18 (4):767–78. PubMed PMID:19010793.
- Inkster B, Strijbis EM, Vounou M, Kappos L, Radue EW, et al. (2013) Histone deacetylase gene variants predict brain volume changes in multiple sclerosis. *Neurobiol Aging* 34 (1):238–47. PubMed PMID:22884548.
- Strijbis EM, Inkster B, Vounou M, Naegelin Y, Kappos L, et al. (2012) Glutamate gene polymorphisms predict brain volumes in multiple sclerosis. *Mult Scler* 31; PubMed PMID:22851457.
- Kremen WS, Panizzon MS, Neale MC, Fennema-Notestine C, Prom-Wormley E, et al. (2012) Heritability of brain ventricle volume: converging evidence from inconsistent results. *Neurobiol Aging* 33 (1):1–8. PubMed PMID:20363053.
- Stein JL, Hua X, Morra JH, Lee S, Hibar DP, et al. (2010) Genome-wide analysis reveals novel genes influencing temporal lobe structure with relevance to neurodegeneration in Alzheimer’s disease. *Neuroimage* 51 (2):542–54. PubMed PMID:20197096.
- Stein JL, Hua X, Lee S, Ho AJ, Leow AD, et al. (2010) Voxelwise genome-wide association study (vGWAS). *Neuroimage* 53 (3):1160–74. PubMed PMID:20171287.
- Hibar DP, Stein JL, Kohannim O, Jahanshad N, Saykin AJ, et al. (2011) Voxelwise gene-wide association study (vGeneWAS): multivariate gene-based association testing in 731 elderly subjects. *Neuroimage* 56 (4):1875–91. PubMed PMID:21497199.
- Shen L, Kim S, Risacher SL, Nho K, Swaminathan S, et al. (2010) 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):1051–63. PubMed PMID:20100581.
- Yang J, Manolio TA, Pasquale LR, Boerwinkle E, Caporaso N, et al. (2011) Genome partitioning of genetic variation for complex traits using common SNPs. *Nat Genet* 43(6):519–25. PubMed PMID:21552263.
- Lee SH, Wray NR, Goddard ME, Visscher PM (2011) Estimating missing heritability for disease from genome-wide association studies. *Am J Hum Genet* 88(3):294–305. PubMed PMID:21376301.
- Schellenberg GD, Montine TJ (2012) The genetics and neuropathology of Alzheimer’s disease. *Acta Neuropathol* 124 (3):305–23. PubMed PMID:22618995.
- Shen D, Davatzikos C (2004) Measuring temporal morphological changes robustly in brain MR images via 4-dimensional template warping. *NeuroImage* 21 (4):1508–1517. PubMed PMID:15050575.
- Davatzikos C, Genc A, Xu D, Resnick SM (2001) Voxel-based morphometry using the RAVENS maps: methods and validation using simulated longitudinal atrophy. *Neuroimage* 14 (6):1361–9. PubMed PMID:11707092.
- Goldszal AF, Davatzikos C, Pham DL, Yan MX, Bryan RN, et al. (1998) An image-processing system for qualitative and quantitative volumetric analysis of brain images. *J Comput Assist Tomogr* 22 (5):827–37. PubMed PMID:9754125.
- Kabani N, MacDonald D, Holmes C, Evans A (1998) A 3D atlas of the human brain. *Neuroimage* 7:5717.
- Yang J, Lee SH, Goddard ME, Visscher PM (2011) GCTA: a tool for genome-wide complex trait analysis. *Am J Hum Genet* 88 (1):76–82. PubMed PMID:21167468.
- Yang J, Benyamin B, McEvoy BP, Gordon S, Henders AK, et al. (2010) Common SNPs explain a large proportion of the heritability for human height. *Nat Genet* 42(7):565–9. PubMed PMID:20562875.
- Frey BJ, Dueck D (2007) Clustering by passing messages between data points. *Science* 315 (5814):972–6. PubMed PMID:17218491.
- Bodenhofer U, Kothmeier A, Hochreiter S (2011) APCluster: an R package for affinity propagation clustering. *Bioinformatics* 27 (17):2463–4. PubMed PMID:21737437.
- Carmelli D, DeCarli C, Swan GE, Jack LM, Reed T, et al. (1998) Evidence for genetic variance in white matter hyperintensity volume in normal elderly male twins. *Stroke* 29 (6):1177–81. PubMed PMID:9626291.
- Pfefferbaum A, Sullivan EV, Swan GE, Carmelli D (2000) Brain structure in men remains highly heritable in the seventh and eighth decades of life. *Neurobiol Aging* 21 (1):63–74. PubMed PMID:10794850.
- Kremen WS, Prom-Wormley E, Panizzon MS, Eyer LT, Fischl B, et al. (2010) Genetic and environmental influences on the size of specific brain regions in midlife: the VETSA MRI study. *Neuroimage* 49 (2):1213–23. PubMed PMID:19786105.
- Peper JS, Brouwer RM, Boomsma DI, Kahn RS, Hulshoff Pol HE (2007) Genetic influences on human brain structure: a review of brain imaging studies in twins. *Hum Brain Mapp* 28 (6):464–73. PubMed PMID:17415783.
- Tramo MJ, Loftus WC, Thomas CE, Green RL, Mott LA, et al. (1995) Surface Area of Human Cerebral Cortex and Its Gross Morphological Subdivisions: In Vivo Measurements in Monozygotic Twins Suggest Differential Hemisphere Effects of Genetic Factors. *J Cogn Neurosci* 7:2, 292–302.
- Wright IC, Sham P, Murray RM, Weinberger DR, Bullmore ET (2002) Genetic contributions to regional variability in human brain structure: methods and preliminary results. *Neuroimage* 17 (1):256–71. PubMed PMID:12482082.
- Niskanen E, Kõnönen M, Määttä S, Hallikainen M, Kivipelto M, et al. (2011) New insights into Alzheimer’s disease progression: a combined TMS and structural MRI study. *PLoS One* 6 (10):e26113. PubMed PMID:22022529.
- Almeida OP, Burton EJ, McKeith I, Gholkar A, Burn D, et al. (2003) MRI study of caudate nucleus volume in Parkinson’s disease with and without dementia with Lewy bodies and Alzheimer’s disease. *Dement Geriatr Cogn Disord* 16 (2):57–63. PubMed PMID:12784028.
- Madsen SK, Ho AJ, Hua X, Saharan PS, Toga AW, et al. (2010) 3D maps localize caudate nucleus atrophy in 400 Alzheimer’s disease, mild cognitive impairment, and healthy elderly subjects. *Neurobiol Aging* 31 (8):1312–25. PubMed PMID:20538376.

36. Buckner RL, Andrews-Hanna JR, Schacter DL (2008) The brain's default network: anatomy, function, and relevance to disease. *Ann N Y Acad Sci* 1124:1–38. PubMed PMID:18400922.
37. Raichle ME, Snyder AZ (2007) A default mode of brain function: a brief history of an evolving idea. *Neuroimage* 37 (4) :1083–90; discussion 1097–9. PubMed PMID:17719799.
38. Buckner RL, Snyder AZ, Shannon BJ, LaRossa G, Sachs R, et al. (2005) Molecular, structural, and functional characterization of Alzheimer's disease: evidence for a relationship between default activity, amyloid, and memory. *J Neurosci* 25 (34) :7709–17. PubMed PMID:16120771.
39. Minoshima S, Giordani B, Berent S, Frey KA, Foster NL, et al. (1997) Metabolic reduction in the posterior cingulate cortex in very early Alzheimer's disease. *Ann Neurol* 42 (1) :85–94. PubMed PMID:9225689.
40. Herholz K, Salmon E, Perani D, Baron JC, Holthoff V, et al. (2002) Discrimination between Alzheimer dementia and controls by automated analysis of multicenter FDG PET. *Neuroimage* 17 (1) :302–16. PubMed PMID:12482085.
41. Reiman EM, Caselli RJ, Yun LS, Chen K, Bandy D, et al. (1996) Preclinical evidence of Alzheimer's disease in persons homozygous for the epsilon 4 allele for apolipoprotein E. *N Engl J Med* 334 (12) :752–8. PubMed PMID:8592548.
42. Weyhenmeyer JA, Gallman EA (2007) *Rapid Review of Neuroscience*. Mosby Elsevier. p.102
43. Chakravarthy VS, Joseph D, Bapi RS (2010) What do the basal ganglia do? A modeling perspective. *Biol Cybern* 103 (3) :237–53. PubMed PMID:20644953.
44. Stocco A, Lebiere C, Anderson JR (2010) Conditional routing of information to the cortex: a model of the basal ganglia's role in cognitive coordination. *Psychol Rev* 117 (2) :541–74. PubMed PMID:20438237.
45. Ungerleider L, Mishkin M (1982) Two cortical visual systems. In D.J. Ingle, M.A. Goodale, & R.J. Mansfield (Eds.), *Analysis of Visual Behavior* (pp. 549–580). Cambridge, MA: MIT Press.