Title
Genetic Determinants of Brain Structure

Author
Hibar, Derrek Paul

Publication Date
2013

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

Genetic Determinants of Brain Structure

A dissertation submitted in partial satisfaction

of the requirements for the degree Doctor of Philosophy

in Biomedical Engineering

by

Derrek Paul Hibar

2013

ABSTRACT OF THE DISSERTATION

Genetic Determinants of Brain Structure

by

Derrek Paul Hibar

Doctor of Philosophy in Biomedical Engineering

University of California, Los Angeles, 2013

Professor Paul M. Thompson, Chair

Over the past decade, billions of dollars from public and private funding institutions have been invested in the fields of neuroimaging and human genetics. Recently, researchers have realized that quantitative measures from imaging methods are a useful substrate for testing how genes influence brain structure, behavior, and susceptibility to disease. However, properly merging the two well-developed fields requires complex new methods and statistical models. Many studies in imaging genetics are simplistic in that they generally focus on individual association tests of a small set of genetic variants (usually single nucleotide polymorphisms, SNPs) on a single summary measure of the brain. These studies are great at bridging the gap between the two fields, but they often fail to utilize advanced methods from either field. For example, in genetics we know that our genes interact with each other in complex pathways, only

in very rare circumstances is a single mutation or variant responsible for observable differences in phenotype. Modeling genetic associations using simple linear regression to test the effect of a single SNP at a time on an imaging phenotype is a good first step, but methods that more accurately represent the underlying biology will test the joint effect of multiple genetic variants simultaneously. The flipside of this is a similar problem, our brains are under strong genetic control, but our differences are extremely complex; using quantitative summary measures from imaging data will miss fine-grained differences between subjects.

The dissertation of Derrek Paul Hibar is approved.

Arthur W. Toga

Eleazar Eskin

Daniel B. Ennis

Paul M. Thompson, Committee Chair

University of California, Los Angeles

2013

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

To all of my family, friends, colleagues, and mentors thank you for your support and understanding throughout my graduate career; I could not have done this without you.

Funding Agency Acknowledgements

The Australian twins study was supported by the National Institute of Child Health and Hu-man Development (R01 HD050735), and the National Health and Medical Research Council (NHMRC 486682), Australia. Genotyping was supported by NHMRC (389875). Projects were also funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (NIH Grant U01 AG024904). ADNI is funded by the NIA, NIBIB, and through generous contributions from the following: Abbott, AstraZeneca AB, Bayer Schering Pharma AG, Bristol-Myers Squibb, Eisai Global Clinical Development, Elan Corporation, Genentech, GE Healthcare, GlaxoSmithKline, Innogenetics, Johnson and Johnson, Eli Lilly and Co., Medpace, Inc., Merck and Co., Inc., Novartis AG, Pfizer Inc, F. Hoffman-La Roche, Schering-Plough, Synarc, Inc., as well as non-profit partners the Alzheimer's Association and Alzheimer's Drug Discovery Foundation, with participation from the U.S. FDA. Private sector contributions to ADNI are facilitated by the Foundation for the National Institutes of Health (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 UCSD. ADNI data are disseminated by LONI. This research was also supported by NIH grants P30 AG010129, K01 AG030514, and the Dana Foundation. Algorithm development and data collection were also funded by the NIH - NIA, NIBIB, NICHD, NLM, NCRR AG040060, MH097268, MH089722, EB01651, LM05639,

RR019771, EB008432, EB008281, and EB007813, to Paul Thompson. Derrek Paul Hibar was partially supported by NSF grant DGE-0707424.

Co-author Acknowledgements

Chapters 2 through 5 are based on the following articles, I would like to thank all of the co-authors for their contributions:

- **Derrek P. Hibar***, Omid Kohannim*, Jason L. Stein, Ming-Chang Chiang, Paul M. Thompson. Multilocus genetic analysis of brain images. Frontiers in Statistical Genetics. 2011. 2:73. *Denotes equal contribution.
- **Derrek P. Hibar**, Jason L. Stein, Omis Kohannim, Neda Jahanshad, Clifford R. Jack, Jr., Michael W. Weiner, Arthur W. Toga, Paul M. Thompson, and the Alzheimer's Disease Neuroimaging Initiative. Principal components regression: multivariate, gene-based tests in imaging genomics, ISBI, 2011.
- **Derrek P. Hibar**, Jason L. Stein, Omid Kohannim, Neda Jahanshad, Andrew J. Saykin, Li Shen, Sungeun Kim, Nathan Pankratz, Tatiana Foroud, Matthew J. Huentelman, Steven G. Potkin, Clifford R. Jack, Jr., Michael W. Weiner, Arthur W. Toga, Paul M. Thompson, and the Alzheimer's Disease Neuroimaging Initiative. Voxelwise gene-wide association study (vGeneWAS): multivariate gene-based association testing in 731 elderly subjects. Neuroimage. 2011. 56(4):1875-1891.
- **Derrek P. Hibar**, Neda Jahanshad, Jason L. Stein, Omid Kohannim, Arthur W. Toga, Sarah E. Medland, Narelle K. Hansell, Katie L. McMahon, Greig I. de Zubicaray, Grant W. Montgomery, Nicholas G. Martin, Margaret J. Wright, Paul M. Thompson. Alzheimer's Disease Risk Gene, GAB2, is Associated with Regional Brain Volume

Differences in 755 Young Healthy Twins. Twin Research and Human Genetics. 2012. 15(3):286-295.

- **Derrek P. Hibar**, Sarah E. Medland, Jason L. Stein, Sungeun Kim, Li Shen, Andrew J. Saykin, Greig I. de Zubicaray, Katie L. McMahon, Grant W. Montgomery, Nicholas G. Martin, Margaret J. Wright, Srdjan Djurovic, Ingrid Agartz, Ole A. Andreassen, Paul M. Thompson (2013). Genetic clustering on the hippocampal surface for genome-wide association studies, MICCAI 2013, Nagoya, Japan, Sept. 22-26 2013 [8-page paper; peer-reviewed].
- **Derrek P. Hibar**, Jason L. Stein, Neda Jahanshad, Arthur W. Toga, Katie L. McMahon, Greig I. de Zubicaray, Grant W. Montgomery, Nicholas G. Martin, Margaret J. Wright, Michael W. Weiner, Paul M. Thompson (2013). Exhaustive search of the SNP-SNP interactome identifies replicated epistatic effects on brain volume, MICCAI 2013, Nagoya, Japan, Sept. 22-26 2013 [8-page paper; peer-reviewed].
- **Derrek P. Hibar**, Jason L. Stein, April B. Ryles, Omid Kohannim, Neda Jahanshad, Sarah E. Medland, Narelle K. Hansell, Katie L. McMahon, Greig I. de Zubicaray, Grant W. Montgomery, Nicholas G. Martin, Margaret J. Wright, Clifford R. Jack, Jr., Michael W. Weiner, Arthur W. Toga, Paul M. Thompson, and the Alzheimer's Disease Neuroimaging Initiative* (2012). Genome-wide association identifies genetic variants associated with lentiform nucleus volume in N=1345 young and elderly subjects. Brain Imaging and Behavior, accepted June 2012. [Epub Ahead of Print].
- Jason L. Stein, Medland, S.E., Vasquez, A.A., **Derrek P. Hibar**, Senstad, R.E., Winkler, A.M., Toro, R., Appel, K., Bartecek, R., Bergmann, Ø., Bernard, M., Brown, A.A., Cannon, D.M., Chakravarty, M.M., Christoforou, A., Domin, M., Grimm, O.,

Hollinshead, M., Holmes, A.J., Homuth, G., Hottenga, J.-J., Langan, C., Lopez, L.M., Hansell, N.K., Hwang, K.S., Kim, S., Laje, G., Lee, P.H., Liu, X., Loth, E., Lourdusamy, A., Mattingsdal, M., Mohnke, S., Maniega, S.M., Nho, K., Nugent, A.C., O'Brien, C., Papmeyer, M., Pütz, B., Ramasamy, A., Rasmussen, J., Rijpkema, M., Risacher, S.L., Roddey, J.C., Rose, E.J., Ryten, M., Shen, L., Sprooten, E., Strengman, E., Teumer, A., Trabzuni, D., Turner, J., van Eijk, K., van Erp, T.G.M., van Tol, M.-J., Wittfeld, K., Wolf, C., Woudstra, S., Aleman, A., Alhusaini, S., Almasy, L., Binder, E.B., Brohawn, D.G., Cantor, R.M., Carless, M.A., Corvin, A., Czisch, M., Curran, J.E., Davies, G., de Almeida, M.A.A., Delanty, N., Depondt, C., Duggirala, R., Dyer, T.D., Erk, S., Fagerness, J., Fox, P.T., Freimer, N.B., Gill, M., Göring, H.H.H., Hagler, D.J., Hoehn, D., Holsboer, F., Hoogman, M., Hosten, N., Jahanshad, N., Johnson, M.P., Kasperaviciute, D., Kent, J.W., Kochunov, P., Lancaster, J.L., Lawrie, S.M., Liewald, D.C., Mandl, R., Matarin, M., Mattheisen, M., Meisenzahl, E., Melle, I., Moses, E.K., Mühleisen, T.W., Nauck, M., Nöthen, M.M., Olvera, R.L., Pandolfo, M., Pike, G.B., Puls, R., Reinvang, I., Rentería, M.E., Rietschel, M., Roffman, J.L., Royle, N.A., Rujescu, D., Savitz, J., Schnack, H.G., Schnell, K., Seiferth, N., Smith, C., Steen, V.M., Valdés Hernández, M.C., Van den Heuvel, M., van der Wee, N.J., Van Haren, N.E.M., Veltman, J.A., Völzke, H., Walker, R., Westlye, L.T., Whelan, C.D., Agartz, I., Boomsma, D.I., Cavalleri, G.L., Dale, A.M., Djurovic, S., Drevets, W.C., Hagoort, P., Hall, J., Heinz, A., Jack, C.R., Foroud, T.M., Le Hellard, S., Macciardi, F., Montgomery, G.W., Poline, J.B., Porteous, D.J., Sisodiya, S.M., Starr, J.M., Sussmann, J., Toga, A.W., Veltman, D.J., Walter, H., Weiner, M.W., Bis, J.C., Ikram, M.A., Smith, A.V., Gudnason, V., Tzourio, C., Vernooij, M.W., Launer, L.J., DeCarli, C., Seshadri, S.,

Andreassen, O.A., Apostolova, L.G., Bastin, M.E., Blangero, J., Brunner, H.G., Buckner, R.L., Cichon, S., Coppola, G., de Zubicaray, G.I., Deary, I.J., Donohoe, G., de Geus, E.J.C., Espeseth, T., Fernández, G., Glahn, D.C., Grabe, H.J., Hardy, J., Hulshoff Pol, H.E., Jenkinson, M., Kahn, R.S., McDonald, C., McIntosh, A.M., McMahon, F.J., McMahon, K.L., Meyer-Lindenberg, A., Morris, D.W., Müller-Myhsok, B., Nichols, T.E., Ophoff, R.A., Paus, T., Pausova, Z., Penninx, B.W., Potkin, S.G., Sämann, P.G., Saykin, A.J., Schumann, G., Smoller, J.W., Wardlaw, J.M., Weale, M.E., Martin, N.G., Franke, B., Wright, M.J., Thompson, P.M. *Identification of common variants associated with human hippocampal and intracranial volumes.* Nature Genetics. 2012. 44(5):552-61.

VITA

2008-2009	Undergraduate researcher volunteering in the Laboratory of Neuro Imaging with Dr. Arthur Toga, University of California, Los Angeles
2009	BS in Biology at the University of California, Los Angeles
2009-2010	Staff research assistant in the Laboratory of Neuro Imaging with Dr. Paul M. Thompson, University of California, Los Angeles
2010-2013	Grantee of the National Science Foundation Graduate Research Fellowship
2012-2013	Grantee of the Nordic Research Opportunity from the Research Council of Norway
2009-present	Co-author on 36 peer-review publication, 1 book chapter, 8 peer-reviewed conference papers, and 34 conference abstracts.
2009-present	Invited speaker for the Organization of Human Brain Mapping Education Course on Imaging Genetics (Beijing June 2011; Seattle June 2013) and the Nansen Neuroscience Network speaking about ENIGMA (May 2011)
2009-present	Invited peer-reviewer for Biological Psychiatry, NeuroImage, Human Brain Mapping, American Journal of Medical Genetics Part B: Neuropsychiatric Genetics, Neuropsychopharmacology, Translational Psychiatry, Reinvention Journal, Neuroscience Letters

SELECT PUBLICATIONS

1. April J. Ho*, Jason L. Stein*, et al. *Commonly carried allele within FTO, an obesity-associated gene, relates to accelerated brain degeneration in the elderly.* Proceedings of the National Academy of Sciences. 2010. 107:8404-8409.
2. Jason L. Stein, et al. *Discovery and replication of dopamine-related gene effects on*

caudate volume in young and elderly populations (N=1198) using genome-wide search. Molecular Psychiatry. 2011 Sep;16(9):927-37

3. **Derrek P. Hibar**, et al. *Voxelwise gene-wide association study (vGeneWAS): multivariate gene-based association testing in 731 elderly subjects. Neuroimage*. 2011. 56(4):1875-1891.

4. Neda Jahanshad, et al. *Brain structure in healthy adults is related to serum transferrin and the H63D polymorphism in the HFE gene.* Proc Natl Acad Sci. 2012 Apr 3; 109(14):E851-9.

5. Jason L. Stein, et al. *Identification of common variants associated with human hippocampal and intracranial volumes.* Nature Genetics. 2012. 44(5):552-61.

6. Joshua C. Bis, et al. *Common variants at 12q14 and 12q24 are associated with hippocampal volume. Nature Genetics.* 2012. 44(5):545-51.

7. **Derrek P. Hibar**, et al. *Genome-wide association identifies genetic variants associated with lentiform nucleus volume in N=1345 young and elderly subjects*. Brain Imaging and Behavior, accepted June 2012. [Epub Ahead of Print].

8. **Derrek P. Hibar**, et al. *Alzheimer's Disease Risk Gene, GAB2, is Associated with Regional Brain Volume Differences in 755 Young Healthy Twins*. Twin Research and Human Genetics. 2012. 15(3):286-295.

9. Xue Hua, et al. *Unbiased Tensor-Based Morphometry: Improved Robustness and Sample Size Estimates for Alzheimer's Disease Clinical Trials*. NeuroImage, 2012 Nov 12; 66C:648-661

10. Neda Jahanshad, et al. *Genome-wide scan of healthy human connectome discovers SPON1 gene variant influencing dementia severity*. Proceedings of the National Academy of Sciences [EPUB]

11. Ming Li, et al. *Allelic differences between Europeans and Chinese for CREB1 SNPs and their implications in gene expression regulation, hippocampal structure and function and bipolar disorder susceptibility.* Molecular Psychiatry, in press, Accepted March 6, 2013.

12. Priya Rajagopalan, et al. *TREM2 Alzheimer's Risk Gene Carriers Lose Brain Tissue Faster.* New England Journal of Medicine [Accepted May 2013].

CHAPTER 1

Introduction

One of the mainstays of quantitative human genetics is the genome-wide association study, which looks at single base pair changes (single nucleotide polymorphisms or SNPs) across the genome and how they relate to changes in some trait (called the phenotype). The basics of a genome-wide association study (or GWAS) are illustrated in **Figure 1**. The advantage of searching the entire genome is that you are not relying on prior hypotheses about which SNPs are causal; you are performing an unbiased search. Forming a prior hypothesis for a SNP or gene based on function can often lead to false positives and poor replication rates. This is especially true when looking for genetic associations with new or previously untested phenotypes. Similar to genome-wide studies, neuroimaging is a data rich field. 3D scans of the brain contain millions of data points each containing localized and unique information about a specific location in the brain. In the same way that it is difficult to prioritize SNPs and genes for analysis, it is difficult and sometimes biased to select a subset of regions in the brain for analysis. With this in mind, combining the two fields of quantitative genetics and neuroimaging becomes a problem of huge dimensions. Properly analyzing and understanding the complex interactions of our DNA and how changes in our DNA related to the differences in our brain requires complex new statistical models and experience with handling high dimensional data. In this dissertation, we examine just a few small parts of a rapidly evolving field, propose new innovated methods, expand our knowledge of the problems, and raise new questions.

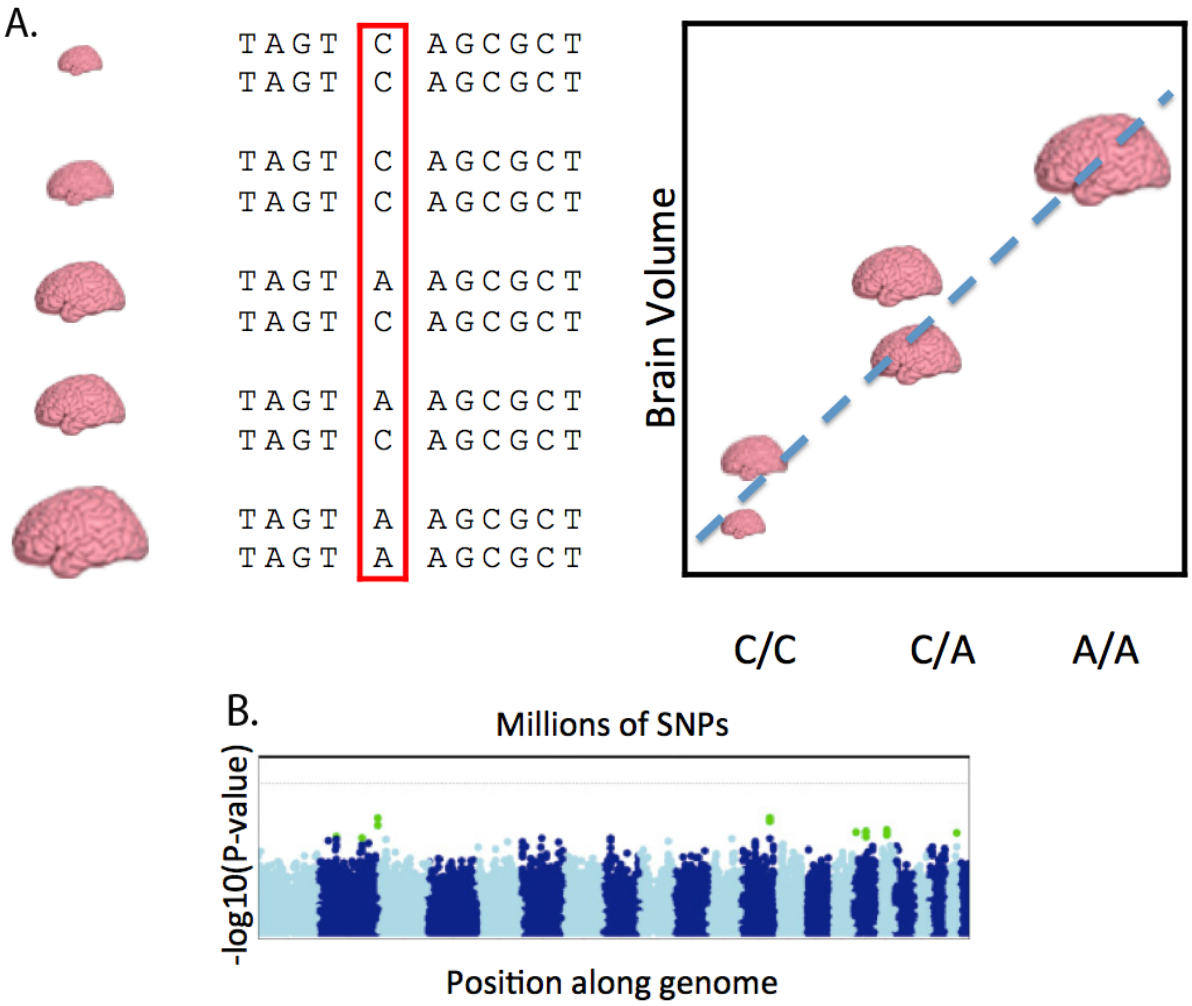


Figure 1. Panel A – a single SNP association test. Collect data on a population of subjects (say brain volume in this example). Next you collect data about the genetic make-up of each subject's DNA. Scan each subject's DNA for locations where some of the subjects are different (single nucleotide polymorphisms or SNPs). Group the subjects by letter pairs at that SNP and fit a regression line. Panel B – there are millions of SNPs at locations across our DNA so we need to repeat the process in Panel A millions of times (once for each SNP) looking for significant relationships between SNPs and our trait. This is a genome-wide association study (or GWAS).

1.1 Merging quantitative human genetics with neuroimaging genetics

Until recently, imaging genetics projects have mainly focused on a specific region of interest for conducting GWA studies. The field has begun to focus on full-brain imaging measures like voxelwise measures of localized volume change [Stein et al., 2010] and fractional anisotrophy [Hibar et al., 2010]. However, both of these studies used traditional multiple linear regression for association testing. No study, to the best of our knowledge, has used multivariate statistical modeling on full-brain, fine-grained imaging phenotypes. Most of these studies could benefit from the rich set of statistical tools developed in the field of quantitative genetics. Specifically, a number of multivariate models like ridge [Hoerl 1962], lasso [Tibshirani 1996], elastic net [Zou and Hastie 2005], partial-least squares [Wold 1984], and principal components regression [Wang and Abbott 2008] have been shown to provide improved power over traditional univariate methods to detect variants associated with phenotypes. In addition, gene-based tests - models that assess the cumulative evidence of association with a phenotype across loci located within a gene transcript - have been shown to increase power over traditional GWA methods [Liu et al., 2010]. To build on this we proposed an imaging genetics study that leverages the complexity of the genome using multivariate gene-based tests with principal components regression (PCReg) to examine the influence of genetic variants on 3-Dimensional maps of regional volume morphometry in the brain. More specifically, we examined whether or not gene-based tests with PCReg will have improved power to detect genetic associations over multiple linear regression. The details of our analyses are discussed in Chapter 2.

1.2 Surface-based morphometry as an endophenotype

One of the most important endeavors in imaging genetics is to identify endophenotypes for genetic analysis. Endophenotypes are quantifiable biological traits that can be reliably and accurately measured and that index the progression or severity of a disease more successfully the standard clinical scores. Many of the imaging traits that we use to study disease are also good examples of endophenotypes for use in genetic analyses of disease. Surface-based morphometric measures have proven to be very powerful for quantifying individual brain differences and indexing disease progression [Gutman et al., 2013]. More specifically, surface-based analyses of subcortical structures like the lateral ventricles, putamen, and hippocampus have been used to detect group differences in Alzheimer's disease [Gutman et al., 2013] and schizophrenia [Styner et al., 2004]. However, no study has looked at the specific genetic variants related to morphometric differences along the surface of subcortical brain structures.

1.2.1 Reducing multiple comparisons correction to boost statistical power

A specific concern with neuroimaging phenotypes including surface-based maps is striking the appropriate balance for finding effects in the whole search space, but also minimizing the correction for multiple comparisons that can reduce power. One way to reduce multiple comparisons is to perform clustering along the image so that analyses can be performed only on a subset (or a local average) of points in the search space. This is possible because many neuroimaging measures contain highly correlated points that are not completely independent. However, the best methods for performing clustering in an imaging measure in order to boost power for genome-wide association studies is not exactly clear. We proposed a study looking at two different types of clustering methods: one based on the Pearson's correlation coefficient and

the other based on the cross-twin cross-trait bivariate genetic correlation coefficient to determine which method is the best for clustering traits to uses as endophenotypes for genetic analysis. The details of our analysis are discussed in Chapter 3.

1.3 Genetic models that more accurately reflect the underlying biology

Continuing with the theme of using more complex statistical models for imaging genetics analysis, we know that traditional univariate methods for testing the association of common genetic variants with complex quantitative traits only consider the marginal effect of a single locus and potentially miss variance explained by synergistic effects of multiple SNPs [Marchinni et al., 2005]. For many complex traits comparisons across families show changes in phenotypic variation exceeding what would be expected as relatedness decreases [Wray et al., 2010]. This implies that there are non-additive (epistatic) interactions involved in the etiology. Previously, there have been a few studies to look at the two-way interactive effects of SNPs on brain structure [Pezawas et al., 2008; Wang et al., 2009; Tan et al., 2007]. However, none of these studies have considered genome-wide genotyping data, instead testing interactive effects on popular candidate genes. This approach potentially misses important interactive effects in regions of the genome where the relevance of the loci in a given SNP pair is not immediately evident. We speculated that genome-wide, SNP-SNP interactions with temporal lobe volumes in the ADNI cohort using the GPU-accelerated regression software would identify novel genetic variants related to brain volume change. We discuss the details of our analysis in Chapter 4.

1.4 Going meta, the future of imaging genetics research?

Imaging genetics projects rely upon the idea that neuroimaging phenotypes are more closely related to disease etiology with a less complex underlying genetic structure than categorical clinical diagnoses. This is the idea of the endophenotype or intermediate phenotype proposed by Gottesman and Gould [2003]. However, a major drawback of measures generated from brain images is that MRI scanning is expensive. Many sites around the world are examining the genetic effects of neuroimaging endophenotypes, however, the cost associated with collecting neuroimaging data means that most sites around the world have sample sizes containing less than 1000 subjects. In the past, many of the imaging genetic studies have found highly significant results in sample sizes of less than 1000, but almost none of the effects found have replicated. The key then to finding new, useful, and exciting genetic variants conferring risk to disease is to boost power by increasing the total number of subjects in your analysis and putting a heavy emphasis on replication and strict significance criteria. However, building a large sample size at a single site, one large enough for believable GWAS analysis, is very difficult. The field of imaging genetics will benefit from joining forces in an international collaborative effort to find genetic variants related to changes in brain structure. There are several limitations to working together, but the benefits far outweigh the negatives. For example, working together allows for the collaborative work to find and replicate believable genetic findings that could not be found alone. However, not every site is willing (or able) to part with their data and send it to a centralized processing pipeline. To avoid this, many sites could keep their data locally, avoiding data sharing issues with an IRB and still contribute to the analysis but just running things locally and contributing test statistics to be used in a meta-analysis framework. This form of collaboration was much more permissive and peeked the interest of many groups around the

world. And thus the Enhancing Neuro Imaging Genetics through Meta Analysis (ENIGMA) Consortium was formed. The details of the pilot project for ENIGMA are discussed in Chapter 5.

1.5 Organization of the dissertation

Each of the studies described in this dissertation are unified by the theme of pushing the field of neuroimaging genetics in new directions. In Chapter 2 we discuss the advantages to multivariate statistics for gene-based tests of association compared to the traditional multiple linear regression. We then apply the analysis in an unbiased search of both the image and gene-wide search space. In Chapter 3 we examine surface-based morphometric values that previously have not been used as endophenotypes for genetic analysis. We also introduce the possibility of using genetic clustering instead of traditional clustering with Pearson's *r* to prioritize and cluster genetically related regions for genetic analysis. In Chapter 4 we build a statistical model of interacting SNPs in the genome reflects the complicated interactions between single base pair changes in the genome. We test for all possible pairwise interactions between SNPs in the SNP-SNP interactome using GPU-accelerated regression software and replicate our findings in an independent dataset. In Chapter 5 we formed an international consortium to facilitate the meta-analysis of GWA studies of neuroimaging phenotypes to improve the power to detect significant variants and increase replication rates compared to individual studies. More specifically, we looked at hippocampal volume, intracranial volume, and total brain volume as endophenotypes in a genetic analysis with 21,151 subjects. Finally, in Chapter 6 we detail our future, planned, and ongoing studies.

CHAPTER 2

Expanding genome-wide search to full brain endophenotypes

2.1 Multilocus genetic analysis of brain images

This section is adapted from:

Derrek P. Hibar*, Omid Kohannim*, Jason L. Stein, Ming-Chang Chiang, Paul M. Thompson. Multilocus genetic analysis of brain images. Frontiers in Statistical Genetics. 2011. 2:73. *Denotes equal contribution.

frontiers in
GENETICS

REVIEW ARTICLE
published: 21 October 2011
doi: 10.3389/fgene.2011.00073

Multilocus genetic analysis of brain images

Derrek P. Hibar[1†], Omid Kohannim[1†], Jason L. Stein[1], Ming-Chang Chiang[1,2] and Paul M. Thompson[1]*

[1] Laboratory of Neuro Imaging, Department of Neurology, University of California Los Angeles School of Medicine, Los Angeles, CA, USA
[2] Department of Biomedical Engineering, National Yang-Ming University, Taipei, Taiwan

Edited by:
Lin S. Chen, The University of Chicago, USA

Reviewed by:
Li Shen, Indiana University School of Medicine, USA
Donghui Yan, Fred Hutchinson Cancer Research Center, USA
Hongyuan Cao, The University of Chicago, USA

***Correspondence:**
Paul M. Thompson, Laboratory of Neuro Imaging, Department of Neurology, University of California Los Angeles School of Medicine, Neuroscience Research Building 225E, 635 Charles Young Drive, Los Angeles, CA 90095-1769, USA.
e-mail: thompson@loni.ucla.edu

[†]Derrek P. Hibar and Omid Kohannim have contributed equally to this work.

The quest to identify genes that influence disease is now being extended to find genes that affect biological markers of disease, or endophenotypes. Brain images, in particular, provide exquisitely detailed measures of anatomy, function, and connectivity in the living brain, and have identified characteristic features for many neurological and psychiatric disorders. The emerging field of *imaging genomics* is discovering important genetic variants associated with brain structure and function, which in turn influence disease risk and fundamental cognitive processes. Statistical approaches for testing genetic associations are not straightforward to apply to brain images because the data in brain images is spatially complex and generally high dimensional. Neuroimaging phenotypes typically include 3D maps across many points in the brain, fiber tracts, shape-based analyses, and connectivity matrices, or networks. These complex data types require new methods for data reduction and joint consideration of the image and the genome. Image-wide, genome-wide searches are now feasible, but they can be greatly empowered by sparse regression or hierarchical clustering methods that isolate promising features, boosting statistical power. Here we review the evolution of statistical approaches to assess genetic influences on the brain. We outline the current state of multivariate statistics in imaging genomics, and future directions, including meta-analysis. We emphasize the power of novel multivariate approaches to discover reliable genetic influences with small effect sizes.

Keywords: GWAS, MRI, brain, penalized regression, sparse regression

INTRODUCTION

Over the past decade, public and private funding institutions have invested billions of dollars in the fields of human neuroimaging and genetics (Akil et al., 2010). Recently, researchers have sought to use quantitative measures from brain images to test how genetic variation influences the brain. Imaging measures are thought to have a simpler genetic architecture than diagnostic measures based on cognitive or clinical assessments (Gottesman and Gould, 2003). In other words, the penetrance of an individual genetic polymorphism is expected to be higher at the imaging level than at the diagnostic level. As such, imaging-derived traits may offer more power to detect how specific genes contribute to brain disease. Genetic analysis of images has been used to discover how susceptibility genes affect brain integrity (Braskie et al., 2011b). Recent studies have revealed gene effects operating within an entire population, in the form of a 3D brain map (Thompson et al., 2001; Stein et al., 2010a; Hibar et al., 2011).

Optimally merging these two well-developed fields requires innovative mathematics and computational methods, guided by genomics and neuroscience. Imaging genetics is still a nascent field, and many studies are relatively simplistic – they generally test how a single genetic variant, or a small set of such variants (usually single nucleotide polymorphisms, or SNPs) are associated with a single summary measure of the brain. These studies begin to bridge the gap between the two fields, but do not take full advantage of advanced methods from either field, which can survey the entire genome or allow an image-wide search. By contrast, multivariate statistical methods such as machine learning and sparse regression can handle high dimensional datasets. Many of these are being adapted to analyze a range of brain processes and biological markers of disease.

In this review, we summarize the recent evolution of imaging genetics, from candidate gene studies to multilocus methods and genome-wide searches to genome-wide, image-wide searches. We explain how images are used in different ways, ranging from single region-of-interest (ROI) methods – that assess the volume or shape of a specific brain region, such as the hippocampus – to voxelwise approaches that survey the whole brain at once in 3D. In these efforts, multivariate, "multilocus" techniques can model how several genetic variants affect the brain at once. Specialized approaches – such as sparse coding methods – can simultaneously handle the high dimensionality and high degree of correlation observed across the genome and in image-derived maps.

CANDIDATE GENE STUDIES

In studies that scan a large number of patients or controls, candidate gene studies have often been used to assess genetic effects on the brain. This approach is appealing as one can test biologically plausible hypotheses and determine how specific, well-studied genetic variations affect brain structure and function. Early studies, for instance, explored how genes related to serotonin transport affected measures extracted from single-photon emission computed tomography (SPECT) and functional magnetic resonance imaging (fMRI; Heinz et al., 2000; Hariri et al., 2002).

Serotonin's role in neurotransmission and neuromodulation – and the well-known anatomy of the monoamine systems – made it possible to frame and confirm testable hypotheses for pertinent regions such as the raphe nuclei and amygdala (Munafo et al., 2008).

Candidate gene studies, such as those above, may assess a single measure derived from a specific ROI in the image. This may be the whole brain, or a subregion such as the gray matter, or the volume or mean activation of a subcortical region. More recently, voxel-by-voxel searches have been conducted to assess candidate gene effects throughout the whole brain in 3D. This unbiased search across the brain makes no prior assumptions on which regions may be affected. Statistical *maps* are also widely used in neuroimaging. Spatial statistics, such as principal components analysis (PCA) or ICA, may also be performed for dimension reduction, and multiple comparisons corrections, such as the false discovery rate (FDR) method, can help to decide if a pattern of gene effects is significant across the voxels searched. For example, Ho et al. (2010) investigated the effects of a proxy SNP in the fat mass and obesity-associated (*FTO*) gene reliably associated with increased risk for obesity (rs3751812; Frayling et al., 2007) on brain structure. They used MRI along with tensor-based morphometry (TBM), to evaluate 206 healthy elderly subjects. *FTO* risk allele carriers had lower frontal and occipital lobe volumes (**Figure 1**). In such studies, maps of statistical associations are created by performing separate association tests at each imaging voxel in the brain. As the number of statistical tests is very large, a standard correction for multiple comparisons can be used, such as the FDR method (Benjamini and Hochberg, 1995) or its more advanced variants such as topological FDR (Chumbley et al., 2010), which consider the geometry of the effects. These corrections assess how likely it is that the overall pattern of associations could be observed by chance. Voxel-based analyses may also be informed by prior hypotheses: ROI may be defined as search regions, such as the temporal lobes, to include prior information on the expected location or patterns of effects (Stein et al., 2010a).

Brain imaging measures used in genetic studies should ideally be highly heritable and be genetically related to a biological process affected by genetic variation, such as a disease process (Gottesman and Gould, 2003; Glahn et al., 2007; Winkler et al., 2010). Some argue that the use of imaging endophenotypes should boost power to detect genetic variants that have reliable but small effects on disease status (Meyer-Lindenberg and Weinberger, 2006). One neuroimaging modality that shows great promise in candidate gene studies is diffusion tensor imaging (DTI), which assesses the fiber integrity of the brain's white matter. DTI is based on the observation that myelination restricts water diffusion, and disease processes typically increase water diffusion across cell membranes (Beaulieu, 2002). Some DTI-derived measures, such as the fractional anisotropy (FA) of diffusion, are widely accepted as measuring brain integrity. FA is highly heritable (Chiang et al., 2009; Kochunov et al., 2010) and is consistently altered in a range of developmental and psychiatric disorders (Thomason and Thompson, 2011). Candidate polymorphisms already associated with brain disorders may be surveyed to discover associations with maps of DTI parameters such as FA. One recent DTI study of young healthy adults (Braskie et al., 2011a), studied the voxelwise effects of the rs11136000 SNP in the recently discovered Alzheimer's disease (AD) risk gene, *CLU*. Significant associations were detected in several anatomical regions that undergo atrophy in AD (**Figure 2**). In similar candidate gene studies using DTI, other genes such as *BDNF* (Chiang et al., 2011a) and *COMT* (Thomason et al., 2010) have been found to influence white matter structure, with carriers of one variant showing consistently higher or lower FA.

GENOME-WIDE ASSOCIATIONS WITH SINGLE IMAGING MEASURES

Candidate gene studies have successfully discovered patterns of brain differences associated with genetic variants whose function is relatively well-known (such as ApoE, for example – a risk gene for late-onset AD; Shaw et al., 2007). The choice of a candidate gene, however, requires a strong prior hypothesis, and most of the genetic determinants of the highly heritable imaging measures

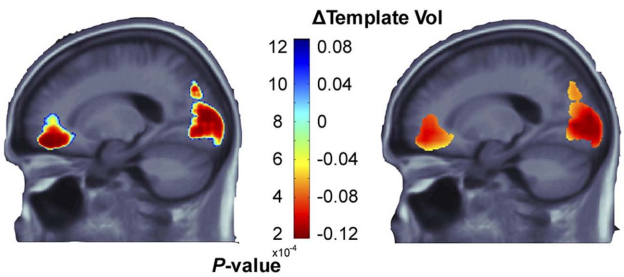


FIGURE 1 | *p*-values (corrected using the false discovery rate method; left panel) and corresponding regression coefficients (right) show the statistical associations between a candidate single nucleotide polymorphism in the *FTO* gene (which is associated with higher risk of obesity) and tensor-based morphometry maps derived from anatomical MRI scans of the brain. Significant associations with regional brain volumes are detected in the occipital and frontal lobes. Clearly, if other regions had been specified in advance as the target of study, association effects may have been missed. Adapted from Ho et al. (2010).

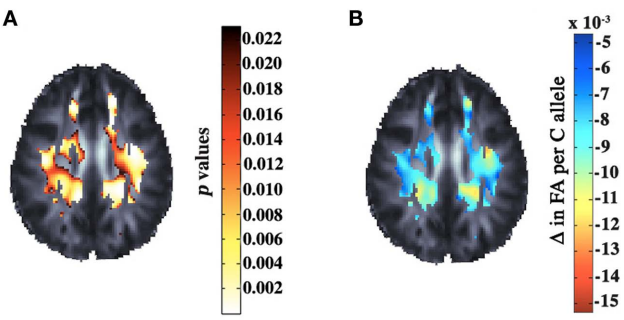


FIGURE 2 | Corrected *p*-values (A) and regression coefficients (B) are shown for the voxelwise effects of a candidate polymorphism in the *CLU* gene – a highly prevalent Alzheimer's susceptibility gene – on fractional anisotropy maps derived from DTI scans of 398 young adults. The axial slice shows the extensive influence of the genetic variant on white matter structure. Adapted from Braskie et al. (2011a).

(connectivity or cortical thickness, for example) are unknown. In most candidate gene studies in imaging, there is a correction for multiple comparisons to control the rate of false discoveries across the image, but this does not take into account the genetic variant tested, or the fact that it could have been selected from a wide list of possibly associated genes. In genetics, and by extension imaging genetics, there is a high risk of false-positive findings unless appropriate corrections are made. Moving beyond candidate gene studies to an unbiased search of the whole genome clearly requires an appropriate genome-wide significance criterion. Otherwise, many false-positive associations will be reported that would not be replicated in the future (Ioannidis, 2005).

Genome-wide association (GWA) studies typically assess associations between hundreds of thousands of SNPs and a phenotype of interest (such as a disease, or a specific image-derived measure). GWA studies have discovered hundreds of common risk loci for diseases and traits in recent years (Hindorff et al., 2009). GWA studies are frequently conducted for discrete, case–control phenotypes, such as the diagnosis of a specific disease (such as AD or schizophrenia vs. healthy control). These studies, however, are limited as participants do not always fall clearly into unique diagnostic categories, and may vary in dimensions not relevant to disease (Pearson and Manolio, 2008). For neuropsychiatric disorders in particular, symptoms expressed by members of specific diagnostic groups may be highly heterogeneous – and there may also be substantial co-morbidity and overlap in symptom profiles across disorders (Psychiatric GWAS Consortium Coordinating Committee et al., 2009; Hall and Smoller, 2010).

Measures derived from brain images in principle are closer to the underlying biology of gene action, offering an alternative target for genome-wide searches, by serving as intermediate phenotypes or endophenotypes for GWA studies (Gottesman and Gould, 2003; Hall and Smoller, 2010). Several imaging GWA scans have been published: Potkin et al. (2009b) identified SNPs in two genes (*RSRC1* and *ARHGAP18*) that showed associations with a blood-oxygen-level dependent (BOLD) contrast measure from a brain region implicated in schizophrenia. Similarly, Stein et al. (2010a) discovered a SNP in the *GRIN2B* gene (rs10845840) and an intergenic SNP (rs2456930) associated with an MRI-derived TBM) measure of temporal lobe volume in 740 elderly subjects from the AD Neuroimaging Initiative. In these and other studies, linear regressions are used to assess the additive or dominant allelic effect of each SNP, after adjusting for covariates such as age and sex, and the confounding effects of population stratification (e.g., Potkin et al., 2009a). This yields *p*-values assessing the evidence for the association of each SNP with the imaging summary chosen. The overall significance of any one SNP effect is then assessed through a form of genome-wide correction for multiple comparisons. Commonly, a nominal *p*-value less than 5×10^{-8} is used.

The GWA study design has been extended to analyze whole images, but one of the shortcomings of all GWAS studies is their limited power (or alternatively, the large sample sizes needed) to detect relevant gene variants. Most SNPs affecting the brain have modest effect sizes (often explaining <1% of the variance in a quantitative phenotype). Meta-analysis can provide added statistical power to discover variants with small effects. Replication, and meta-analysis in particular, have been widely embraced as a way to aggregate evidence from multiple genetic studies, including studies of disease risk, and normally varying traits such as height (de Bakker et al., 2008; McCarthy et al., 2008; Zeggini and Ioannidis, 2009; Yang et al., 2010).

Even so, most imaging GWA studies consider under a thousand subjects, so are limited in detection power. This led many researchers in the field to band together to search for relevant genetic associations with imaging traits meta-analytically, in many large samples. One promising initiative is called Enhancing Neuro Imaging Genetics through Meta-Analysis (ENIGMA) and is currently accepting research groups who want to become involved in meta-analytic imaging genomics projects (http://enigma.loni.ucla.edu/). The ENIGMA pilot project is a large meta-analysis to discover genes associated with hippocampal volume on brain MRI in over 9,000 subjects scanned by 21 research centers (The ENIGMA Consortium, 2011). Future imaging genetics studies may rely on large meta-analyses and international collaborations to overcome the low power and relatively small effect sizes. However, some genetic associations can be found and replicated without vast meta-analytic approaches like ENIGMA. For example, Stein et al. (2011) discovered and replicated an association between caudate volume and the SNP rs163030 located in and around two genes, *WDR41* and *PDE8B*. These genes are involved in dopamine signaling and development; a Mendelian mutation in one leads to severe caudate atrophy. Similarly, Joyner et al. (2009) replicated an association with cortical surface area in a common variant (rs2239464) of the *MECP2* gene, which is linked to microencephaly and other morphological brain disorders.

GENETIC ANALYSIS OF MASS UNIVARIATE IMAGING PHENOTYPES

Studying a single imaging measure with a genome-wide search is as limited as picking a single candidate gene from the entire genome – it may not fully reflect how a given genetic variant influences the brain, or it may miss an important effect by being too restrictive. Important links may be overlooked if a gene variant influences a brain feature present but not measured in the images. To broaden the range of measures surveyed in each image, Shen et al. (2010) studied patients with AD and mild cognitive impairment (MCI) using whole-brain voxel-based morphometry (VBM; Good et al., 2001) and split the brain into 142 cortical and subcortical ROIs using the segmentation software package FreeSurfer (Fischl et al., 2002). The VBM measure within each ROI was averaged for each subject and those values were used as traits for GWA scans. One SNP, rs6463843, from the *NXPH1* gene, was significantly associated with gray matter density in the hippocampus, and had broad morphometric effects in a *post hoc* exploratory analysis. While this study found plausible results, the computation of summaries from ROI may miss patterns of effects that lie only partially within the chosen ROI. As such, a combination of map-based and ROI-based methods seems ideal.

Some researchers have combined unbiased tests of association across the genome with unbiased searches of the entire brain, instead of relying on summary measures derived from ROI. Combining GWA scans with an image-wide search is computationally intensive, requiring new methods to handle the

high dimensionality and multiple statistical comparisons. Three dimensional brain images may contain over 100,000 voxels, and a completely unbiased search may test up to one million SNPs for association at each voxel. This is extremely computationally intensive, but can be completed in a feasible time frame if the process is parallelized. Stein et al. (2010b) performed a full GWA scan at each voxel in maps of regional brain volume calculated by TBM (Leow et al., 2005). Sixteen billion tests of association were conducted – in a so-called "voxelwise genome-wide association study" (vGWAS). To accommodate the huge number of statistical tests performed, only the most highly associated SNP at each voxel was stored. The *p*-value distribution for the top SNP was modeled as a beta distribution, *Beta*(1, *n*), where *n* is an estimate of the effective number of independent tests performed (Ewens and Grant, 2001). The resulting distribution of minimum *p*-values across the genome, assembled from voxels across the image, was transformed into a uniform distribution in the null case for multiple comparisons correction across the image. FDR was used to correct for multiple comparisons across the image, and to assess whether credible effects had been detected (Benjamini and Hochberg, 1995). Several top SNPs were associated with moderate regional brain volume differences; many were in genes that are expressed in the brain (**Figure 3**). However, no SNPs passed the strict correction for multiple comparisons. The Stein et al. study was a *proof of concept*, showing that a completely unbiased search of the genome is feasible with imaging phenotypes. However, the huge correction for multiple comparisons across the image and genome are practically insurmountable unless the effect size or cohort size is very large. In addition, the vGWA study required 27 h when spread across 500 CPUs; this is more computational power than most researchers typically have access to. Clearly, an optimal balance must be made between pure discovery methods, unconstrained by prior hypotheses, and those that invoke prior biological information to boost power and reduce the multiple comparisons correction.

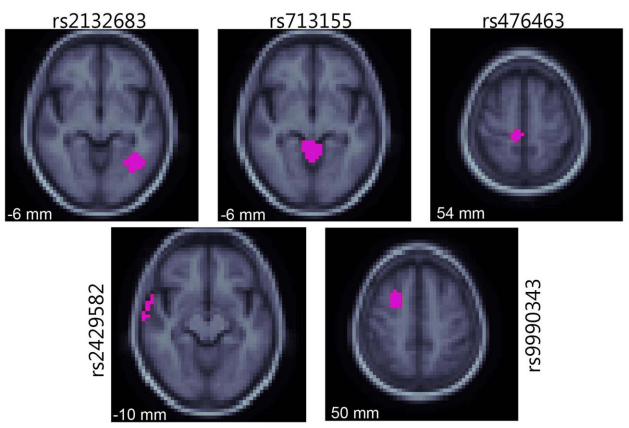


FIGURE 3 | The five most highly associated SNPs identified by vGWAS are shown on slices of an averaged brain MRI template, indicating regions where these SNPs were the most highly associated out of all SNPs (*in purple*). Coordinates refer to the ICBM standard space, and the cohort is the ADNI sample. Adapted from Stein et al. (2010a).

MULTIVARIATE IMAGING GENETICS METHODS

Multivariate methods can be used to assess the joint effect of multiple genetic variants simultaneously, and are widely used in genetics (Phillips and Belknap, 2002; Gianola et al., 2003; Cantor et al., 2010). For example, set-based permutation methods use gene annotation information and linkage disequilibrium values to group univariate *p*-values from traditional GWA studies into gene-based test statistics (Hoh et al., 2001; Purcell et al., 2007). Set-based approaches use prior information on gene structure to incorporate all genotyped SNPs in a given gene into a single test statistic. This can offer, in some cases, greater power than univariate statistical tests to detect SNP effects. Combining univariate *p*-values into a single gene-based test also reduces the total number of tests performed, alleviating the multiple comparisons correction. It can also aggregate the cumulative evidence of association across a gene block to account for allelic heterogeneity (Hoh et al., 2001). Individual SNP *p*-values may not achieve the genome-wide significance level for a traditional GWA study (nominally $p < 5 \times 10^{-8}$), but if several SNPs in the same LD block show moderate association, the combined evidence for association may be enough to beat a gene-wide significance level (nominally $p < 5 \times 10^{-6}$). For example, one study examined SNPs from the SORL1 gene for association with hippocampal volume in healthy elderly controls (Bralten et al., 2011). While they did not find evidence for association of individual SNPs in a discovery and replication dataset, a gene-based test found evidence of association in both datasets. Some set-based statistics may be derived from the separate *p*-values from the individual univariate tests, enabling *post hoc* analysis of published studies. A major issue in applying set-based statistics in imaging genetics is that the permutation procedure applied across SNP groupings would be very computationally intensive. Set-based methods are currently not feasible to apply at >100,000 voxels, as a single gene test takes around 5 min (or 22.8 years to test a single gene at every voxel of the full brain on one CPU). In addition, combining SNPs by *p*-value may miss an important effect where a set of SNPs from the same gene have moderate covariance, but explain different portions of variance in the phenotype. In other words, if they were considered together in the same model, the overall variance explained may be greater than its univariate significance level would imply.

An alternative to set-based methods is to group SNPs into a single statistical model and then test that model for overall association. One classical example of this strategy is multiple linear regression (MLR). However, a problem with applying MLR to genetic data is that SNPs tend to be highly correlated, as they co-segregate in haplotype blocks (Frazer et al., 2007). The MLR is highly sensitive to collinearity among predictors; the inversion step in calculating regression coefficients involves a matrix that is not full rank as the variables are collinear. This leads to wildly inaccurate *Beta* value estimates and SE (Kleinbaum, 2007). To avoid collinearity in multivariate analysis of genetic data, dimensionality is often reduced using sparse regression methods, such as penalized or principal components regression (PCReg).

Some data reduction methods compute a new set of statistically orthogonal variables, for inclusion in a classical MLR model. A data reduction method such as PCA transforms a matrix of SNP predictors into a new orthogonal set of predictors, ranked in

descending order based on the amount of the variance in the data that each component explains (Jolliffe, 2002). The output of PCA is typically a matrix that explains the same amount of the overall variance as the original predictors, but without the collinearity. As the individual components are sorted by amount of variance they explain, the resulting statistical models can strike an efficient balance between the total variance explained (the number of components to include) and the number of degrees of freedom used (model complexity increases as more variance components are included).

One method, known as PCReg first performs PCA on a set of predictors. It then builds a multiple partial-*F* regression model where the number of components included is based on the desired proportion of variance to be explained (Massy, 1965). Wang and Abbott (2008) used PCReg to group SNPs into a single multivariate test statistic. Hibar et al. (2011) extended this method to be applicable to images, conducting gene-based tests at each voxel with PCReg. They used an automated method (Altshuler et al., 2005; Hemminger et al., 2006; Hinrichs et al., 2006) to group SNPs based on gene membership, resulting in 18,044 unique genes. Using the set of SNPs in each individual gene as predictors, Hibar et al. used PCReg to assess the degree of association for every gene at every voxel in the full brain. The resulting method was termed a voxelwise "gene-wide" association study vGeneWAS. By compressing the SNPs into gene-based tests, the total number of tests was reduced to around 500 million tests from the 16 billion tests in vGWAS. However, even with this much smaller number of tests, no genes identified passed correction for multiple comparisons. The most highly associated gene, *GAB2*, showed strong credibility as it is consistently associated with neurodegenerative disorders such as AD (Reiman et al., 2007). In addition, Hibar et al. (2011) simulated full brain parametric maps using statistical priors based on their observed data to show that observed clusters of associated genes were larger than would be expected by chance. This provides evidence that vGeneWAS is a valid and powerful multivariate method to detect gene effects in full brain neuroimaging data. A head-to-head comparison of vGWAS and vGeneWAS was also performed on the same datasets. The cumulative distribution function (CDF) plots of *p*-values for each study show that the FDR in the multivariate vGeneWAS was controlled at a lower rate than in the mass univariate vGWAS method (**Figure 4**).

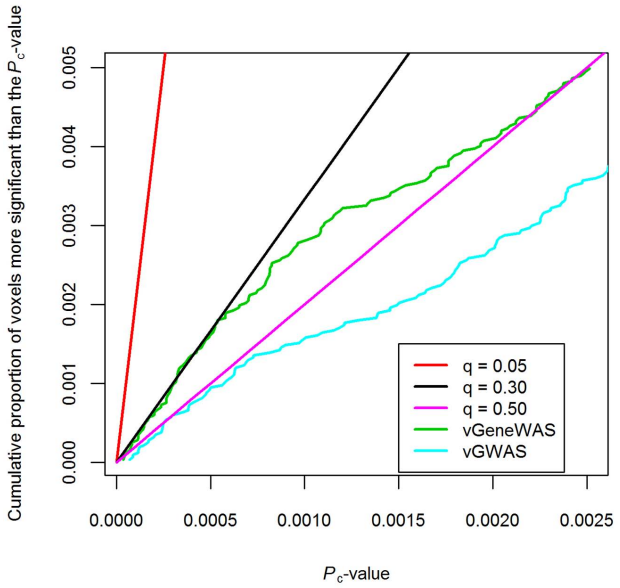


FIGURE 4 | Cumulative distribution function (CDF) plot of corrected *p*-values from vGeneWAS (Hibar et al., 2011) and vGWAS (Stein et al., 2010b) analyses. vGWAS could only be controlled for false-positives at $q = 0.50$ threshold, while vGeneWAS could be controlled for false-positives at $q = 0.30$. The difference in *q*-value thresholds for the CDF of the *p*-values obtained from both studies on the same dataset suggests that the gene-based analysis is more powerful, though neither study controlled the false-positive rate at the nominal $q = 0.05$ threshold.

An extension of PCReg and other data reduction techniques is to perform data reduction on *both* the genome and the 3D brain imaging traits. One approach that appears to be promising is parallel independent components analysis (Parallel ICA or PICA; Liu et al., 2009). Parallel ICA works by first performing PCA on a set of SNPs and also a different PCA on a voxelwise imaging measure. Next, a modified version of ICA is applied to both modalities and independent factors from each modality are chosen simultaneously by a correlation measure (hence "parallel" ICA). Selecting imaging features and SNPs together can be more powerful than mass univariate tests of voxelwise imaging traits as the total number of tests is greatly reduced. For example, Liu et al. (2009) used pre-processed fMRI maps from 43 healthy controls and 20 schizophrenia patients and a pre-selected set of 384 SNPs chosen for their potential associations with schizophrenia. Via a *t*-test, Liu et al. (2009) demonstrated that genetic components ($p = 0.001$) and fMRI BOLD ($p = 0.0006$) response loadings from parallel ICA were able to distinguish healthy subjects from patients with schizophrenia, with reasonable accuracy. Similar approaches have been applied to structural MRI (Jagannathan et al., 2010). The PICA method is quite promising, but several challenges remain. As Parallel ICA requires an initial round of PCA, it is difficult to recover which SNP sets are contributing to a given component and similarly it is difficult to localize the 3D spatial effect contributing to each component from the image. This may make it difficult to interpret and replicate specific findings. In addition, it is not clear how data reduction methods will perform with whole genome and full brain data. Liu et al. (2009) and Jagannathan et al. (2010) both performed considerable downsampling of the images, reducing the total number of voxels included in the Parallel ICA model. In addition, both studies tested only small sets of pre-selected SNPs instead of data from the full genome, or a standard 500,000 SNP genome-wide scan. The power of Parallel ICA to find common components may be greatly reduced if there is additional noise from genome-wide data. Liu et al. (2009) found that as the amount of random noise increased, so did the number of independent components. As the number of independent components increases, the power to detect associations decreases. Also querying full brain phenotypes for effects of genetic variants, another recently proposed multivariate method by Chiang et al. (2011b), identified patterns of voxels in a DTI image with a common genetic determination, and aggregated them to boost power in GWA (**Figure 5**). Approximately 5,000 brain regions were

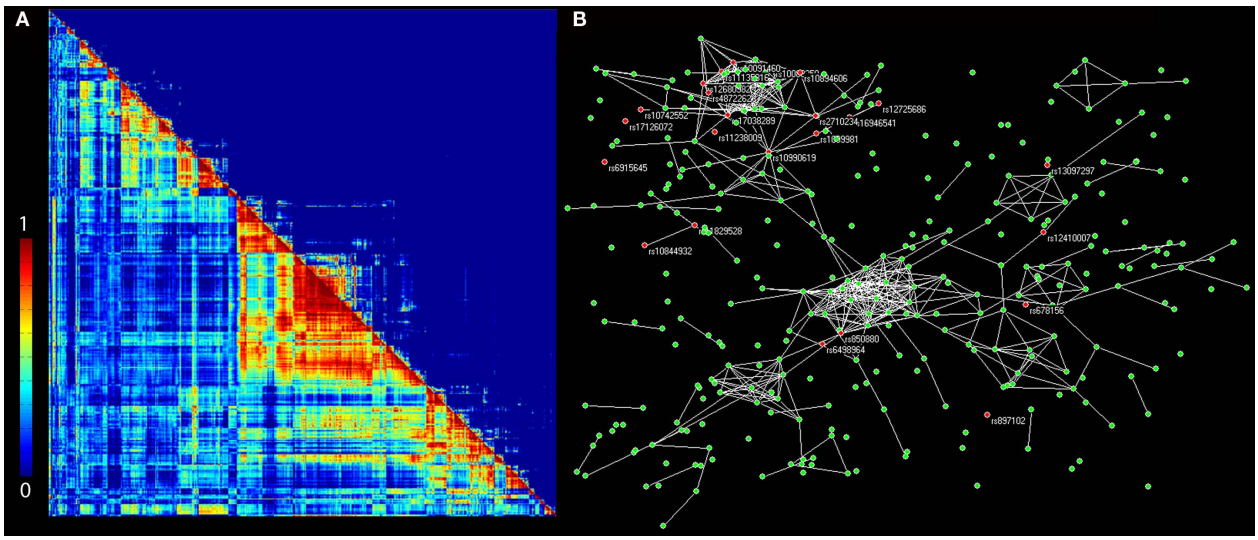


FIGURE 5 | Clustering regions of a brain image that have common genetic determination. In a DTI study of twins, the known kinship structure made it possible to estimate the genetic correlation matrix and a "topological overlap (TO)" index matrix. This was used to gage the similarity of genetic influences on all pairs of brain regions **(A)**. The 18 largest clusters – parts of the image with common genetic influences – were selected as regions of interest (ROIs) for GWAS. By associating the mean white matter integrity of these regions with genetic variants, a genetic interconnection network was obtained **(B)**, where each network node represents a single SNP (colored circles). The figure shows only those SNPs associated with white matter integrity in at least one ROI with a significance p-value $< 10^{-5}$. SNPs whose associations reach genome-wide significance are colored in red, with their names labeled. White lines indicate that SNPs are "connected," i.e., their effects on white matter integrity are strongly correlated. Adapted from Chiang et al. (2011b).

selected, where genetic influences accounted for >60% of the total variation of white matter integrity. From these, a 5,000 × 5,000 correlation matrix was obtained. Hierarchical clustering was used to select the largest clusters, and these voxels were defined to be ROIs. The mean FAs for these ROIs were then tested for evidence of association with all SNPs genotyped across the genome. By identifying a genetic network that influences white matter integrity over multiple brain regions, Chiang et al. (2011b) were able to boost power to detect associations between FA in these brain areas and SNPs from the whole genome. In all, they identified 24 SNPs with genome-wide significance, which is unusual for a study with fewer than 1,000 subjects. To ensure the findings are not false-positives, however, simulations of imaging and genomic data may be necessary (as carried out by Vounou et al., 2010 see below).

Variants near each other on the genome can be highly correlated due to linkage disequilibrium. This leads to problems if all variants are included in a standard multiple regression model to predict the values of a trait. To address this, many new mathematical methods have been used to handle the high dimensionality in the genome (a $p \gg n$ problem) and interactions between genetic variants. These include penalized and sparse regression techniques, such as ridge regression (Hoerl, 1962), the least absolute shrinkage and selection operator (LASSO; Tibshirani, 1996), the elastic net (Zou and Hastie, 2005), and penalized orthogonal-components regression (Malo et al., 2008; Cho et al., 2009; Lin et al., 2009; Zhang et al., 2009; Chen et al., 2010). The various penalty terms (e.g., L^1 in LASSO and L^2 in ridge) in the regularized regression methods can incorporate large numbers of correlated variants with possible interaction terms, in single models. These methods show high statistical power in analyses with both real and simulated data. Although these studies are almost invariably applied to case–control GWA studies, similar approaches may be applied to imaging phenotypes. Kohannim et al. (2011a), for instance, implemented ridge regression to study the association of genomic scanning windows with MRI-derived temporal lobe and hippocampal volume. They reported boosting of power in detecting effects of several SNPs, when compared to univariate imaging GWA. One statistical challenge of such sliding-window approaches is finding optimal window sizes, which can capture the correlation structure in the genomic data without adding excessive degrees of freedom to the model. Kohannim et al. considered several fixed, scanning window sizes (50, 100, 500, and 1000 kbp) in their study, and found boosting of power in detecting SNPs with different window sizes for different genomic regions. A more flexible approach may incorporate information such as the sample size and variant-specific LD structure into the selection of optimal window sizes for each genomic region (e.g., Li et al., 2007). This could ensure that SNPs are not missed due to inappropriate window sizes. In addition, L^1-driven methods, such as LASSO, may provide greater detection power by selecting sparse sets of genomic variants in association with imaging measures (Kohannim et al., submitted). As discussed above, however, multivariate methods can be applied not only to the genome, but also to the images, which are also high dimensional and show high spatial correlations. Sparse and penalized models can be useful in these situations as well. Vounou et al. (2010) applied a sparse reduced-rank regression (sRRR) method to detect whole genome-whole image associations. They computed a matrix of regression coefficients, C, whose rank was p (number of SNP genotypes) times q (the number of imaging phenotypes, or pre-defined anatomical ROIs in their case). They reduced the rank

of this large matrix to r, by factorizing the matrix into the product of a $p \times r$ matrix, B, and an $r \times q$ matrix, A, and constraining A and B to be sparse (**Figure 6**). To evaluate the power of their method and compare it to that of mass univariate modeling, Vounou et al. generated realistic, simulated imaging and genetic data. Using the FoRward Evolution of GENomic rEgions (FREGENE) software, and the ADNI baseline T1-weighted MRI dataset, they obtained a simulated dataset, to which they introduced genetic effects in a number of ROIs. It was not feasible for the investigators to consider all possible genetic effect sizes and sample sizes, but they were able to show boosted power for all parameter settings they explored. Setting the desired, reduced-rank r equal to 2 or 3, they obtained higher sensitivities with sRRR at any given specificity for a sample size of 500. When they increased the sample size from 500 to 1,000, they noted gains in sensitivities with sRRR, which were more considerable than the merely linear gains obtained with univariate modeling. They also demonstrated that boosted sensitivities obtained with sRRR increase with higher numbers of SNPs; sensitivity ratios (sRRR/mass univariate modeling) could be boosted even further to ratios far exceeding 5 (observed with 40,000 SNPs) with numbers of SNPs considered in a typical GWAS (e.g., 500,000 SNPs). Direct power comparisons between association methods on DNA microarray data show that models that incorporate linear combinations of variables perform better than those that perform simple data reduction (Bovelstad et al., 2007). Bovelstad et al. found that the penalized method, ridge regression, was more powerful than LASSO, PCReg, supervised PCReg, and partial least squares regression (PLS), when it comes to predicting survival rates in cancer patients from DNA microarray data. In the future, direct comparisons of methods on imaging genetics data could inform the direction of new methods development.

Comprehensive modeling of whole-brain voxelwise and genome-wide data remains challenging, due to the high dimensionality of the data. This causes both statistical and computational problems. Recently, there have been new developments applying sparse regression methods to genome-wide data; one such method is *iterative sure independence screening* (ISIS; Fan and Lv, 2008; Fan and Song, 2010; He and Lin, 2011). ISIS is an iterative selection procedure that builds a marginal model using the cyclic coordinate descent (CCD; Friedman et al., 2010) algorithm with the LASSO and combines it with a conditional model of interactions based on pairwise correlations. The combined model has lower dimensionality, but effects of individual SNPs are still identifiable, as are SNP–SNP interactions. This method appears to be promising for discovery-based searches of the genome. ISIS has not yet been applied to brain images, but it should be feasible. Methods such as ISIS could also be modified to jointly select imaging phenotypes and genomic data as done by Vounou et al. (2010) but without first having to select ROI or only a small subset of SNPs from the genome.

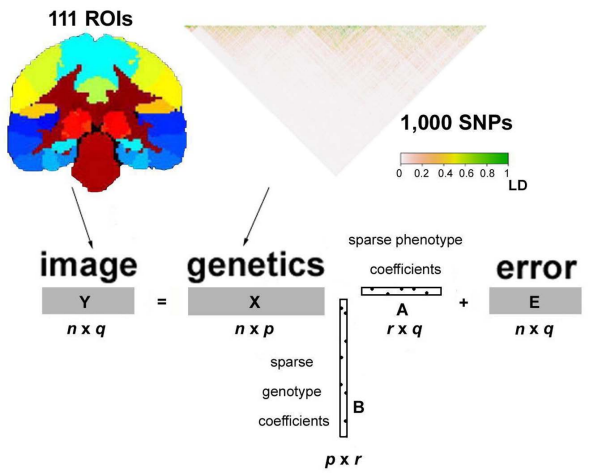


FIGURE 6 | Imaging and genomic data are incorporated into a sparse, reduced-rank regression model, where regions of interest (ROIs), and single nucleotide polymorphisms (SNPs) attain sparse coefficients, simultaneously. This approach can select pertinent SNPs and ROIs. In simulations, it demonstrates higher power than mass univariate models for detecting effects of genetic variants. Adapted with authors' and publishers' permission from Vounou et al. (2010).

CONCLUSION

The field of imaging genetics started with candidate gene studies, where hypotheses about gene action on brain structure and function could be tested in a novel way. More recently, candidate gene studies have been extended to investigate voxelwise associations between genetic variants and images of the brain, to map 3D profiles of genetic effects without requiring *a priori* selection of ROI.

To consider the entirety of the genome and discover potentially new variants, however, GWA studies have been introduced to the field of imaging genetics. In these studies, quantitative measures derived from images are considered as intermediate phenotypes, which are in some respects closer to the underlying biology of brain disorders and processes of interest. Despite their unbiased consideration of the whole genome, the standard, univariate GWA approach considers only one SNP at a time and has several limitations. From a genetic perspective, it does not take into account the interdependence between genetic variants due to linkage disequilibrium; and in regard to imaging, such studies typically rely on single summary measures from images, which only weakly represent the wealth of information in a full 3D scan.

Among the most promising applications of imaging genetics are those that use sparse methods to reduce the data dimensionality. Sparse methods create efficient models, and boost power to identify patterns of association. A major advantage of penalized or sparse regression methods is that they accommodate collinearity inherent in the genome and in the images, but they still offer a familiar regression framework to accommodate covariates and confounding variables. Penalized regression models may include a large number of genetic predictors. This may discover genetic effects undetected by other data reduction methods, such as PICA and PCReg. For studies of large 3D statistical maps of imaging phenotypes, methods to penalize the selection of both voxels from the image and associated genetic variants from the genome seem to have higher power than related discovery-based methods. Even so, this is largely an empirical question that depends on the structure of the true signal. Indeed, Vounou et al. (2010) demonstrated the increased power of the sRRR method, which favors the selection of an efficient set of ROI and a reduced number of SNPs has increased power. A major limitation of penalized methods is that they may fail to converge on a solution when the data dimensions

are very high. Even methods designed for $p \gg n$ problems such as least angle regression (Efron et al., 2004) tend to fail when given a full 3D imaging phenotype. This illustrates why current implementations of penalized regression in imaging genetics often rely on prior "groupings" of voxels or sliding windows in the genome. These prior groupings do not appear to be motivated by strong prior hypotheses, but by limitations in the statistical modeling. Methods similar to ISIS (Fan and Lv, 2008; Fan and Song, 2010; He and Lin, 2011) designed for ultra-high dimensional datasets will likely be useful for future imaging genetics projects.

Once we have a set of validated genetic variants that affect the brain, multivariate models may be used to combine imaging, genetics, and other physiological biomarkers to predict outcomes in patients with brain disorders. The resulting combination of imaging and genetic data, with other biomarkers, can be used to predict an individual's personalized aggregate risk for specific types of brain disorders. As genomic and proteomic data are added, prognosis and diagnosis may be possible at an earlier stage or more accurate than is possible with current biomarkers. Machine learning algorithms (e.g., decision trees, support vector machines, and neural networks) have shown promise for making disease predictions from genomic and proteomic data (Cruz and Wishart, 2007). Similar approaches may be useful in psychiatry research, and neuroimaging measures such as fiber anisotropy from diffusion imaging may help in making early predictions of brain integrity from genes. In a recent, preliminary study, our group incorporated several candidate polymorphisms in a multi-SNP, machine learning model, to predict personal measures of fiber integrity in the corpus callosum (Kohannim et al., 2011b). Ideally, by incorporating both genomic and proteomic data from larger cohorts, one may be able to obtain personalized "scores" for brain integrity from biomarker profiles. This has considerable implications for prevention and early treatment of brain pathology.

ACKNOWLEDGMENTS

This work is supported in part by NIH grants R01 EB008432, EB008281 HD050735 AG020098, RC2 AG036535, P41 RR013642 (to Paul M. Thompson). Derrek P. Hibar is partially supported by NSF grant DGE-0707424. Omid Kohannim was supported in part by the UCLA MSTP. Jason L. Stein is partially supported by a T32 post-doctoral training grant in Neurobehavioral Genetics.

REFERENCES

Akil, H., Brenner, S., Kandel, E., Kendler, K. S., King, M. C., Scolnick, E., Watson, J. D., and Zoghbi, H. Y. (2010). Medicine. The future of psychiatric research: genomes and neural circuits. *Science* 327, 1580–1581.

Altshuler, D., Brooks, L. D., Chakravarti, A., Collins, F. S., Daly, M. J., Donnelly, P., Gibbs, R. A., Belmont, J. W., Boudreau, A., Leal, S. M., Hardenbol, P., Pasternak, S., Wheeler, D. A., Willis, T. D., Yu, F. L., Yang, H. M., Zeng, C. Q., Gao, Y., Hu, H. R., Hu, W. T., Li, C. H., Lin, W., Liu, S. Q., Pan, H., Tang, X. L., Wang, J., Wang, W., Yu, J., Zhang, B., Zhang, Q. R., Zhao, H. B., Zhao, H., Zhou, J., Gabriel, S. B., Barry, R., Blumenstiel, B., Camargo, A., Defelice, M., Faggart, M., Goyette, M., Gupta, S., Moore, J., Nguyen, H., Onofrio, R. C., Parkin, M., Roy, J., Stahl, E., Winchester, E., Ziaugra, L., Shen, Y., Yao, Z. J., Huang, W., Chu, X., He, Y. G., Jin, L., Liu, Y. F., Shen, Y. Y., Sun, W. W., Wang, H. F., Wang, Y., Wang, Y., Wang, Y., Xiong, X. Y., Xu, L., Waye, M. M. Y., Tsui, S. K. W., Xue, H., Wong, J. T. F., Galver, I. L. M., Fan, J. B., Murray, S. S., Oliphant, A. R., Chee, M. S., Montpetit, A., Chagnon, F., Ferretti, V., Leboeuf, M., Olivier, J. F., Phillips, M. S., Roumy, S., Sallee, C., Verner, A., Hudson, T. J., Frazer, K. A., Ballinger, D. G., Cox, D. R., Hinds, D. A., Stuve, L. L., Kwok, P. Y., Cai, D. M., Koboldt, D. C., Miller, R. D., Pawlikowska, L., Taillon-Miller, P., Xiao, M., Tsui, L. C., Mak, W., Sham, P. C., Song, Y. Q., Tam, P. K. H., Nakamura Y., Kawaguchi, T., Kitamoto, T., Morizono. T., Nagashima, A., Ohnishi, Y., Sekine, A., Tanaka, T., Tsunoda, T., Deloukas, P., Bird, C. P., Delgado, M., Dermitzakis, E. T., Gwilliam, R., Hunt S., Morrison, J., Powell, D., Stranger, B. E., Whittaker P., Bentley, D. R., Daly, M. J., de Bakker, P. I. W., Barrett, J., Fry, B., Maller, J., McCarroll, S., Patterson, N., Pe'er, I., Purcell, S., Richter, D. J., Sabeti, P., Saxena, R., Schaffner, S. F., Varilly, P., Stein, L. D., Krishnan L., Smith, A. V., Thorisson, G. A., Chen, P. E., Cutler, D. J., Kashuk, C. S., Lin, S., Abecasis, G. R., Guan, W. H., Munro, H. M., Qin, Z. H. S., Thomas, D. J., McVean, G., Bottolo, L., Eyheramendy, S., Freeman C., Marchini, J., Myers, S., Spencer, C., Stephens, M., Cardon, L. R., Clarke, G., Evans, D. M., Morris, A. P., Weir, B. S., Tsunoda, T., Mullikin, J. C., Sherry, S. T., Feolo, M., Zhang, H. C., Zeng, C. Q., Zhao, H., Matsuda, I., Fukushima, Y., Macer, D. R., Suda, E., Rotimi, C. N., Adebamowo C. A., Ajayi, I, Aniagwu, T., Marshall, P. A., Nkwodimmah, C., Royal, C. D. M., Leppert M. F., Dixon, M., Peiffer, A., Qiu, R. Z., Kent, A., Kato, K., Niikawa, N., Adewole, I. F., Knoppers, B. M., Foster, M. W., Clayton, E. W., Muzny, D., Nazareth, L., Sodergren, E., Weinstock, G. M., Wheeler, D. A., Yakub, I., Gabriel, S. B., Richter, D. J., Ziaugra, L., Birren, B. W., Wilson, R. K., Fulton, L. L., Rogers, J., Burton, J., Carter, N. P., Clee, C. M., Griffiths, M., Jones, M. C., McLay K., Plumb, R. W., Ross, M. T., Sims, S. K., Willey, D. L., Chen, Z., Han, H., Kang, L., Godbout, M., Wallenburg, J. C., Archeveque, P. L., Bellemare, G., Saeki, K., Wang, H. G., An, D. C., Fu, H. B., Li, Q., Wang, Z., Wang, R. W., Holden, A. L., Brooks, L. D., McEwen, J. E., Bird, C. R., Guyer, M. S., Nailer, P. J., Wang, V. O., Peterson, J. L., Shi, M., Spiegel, J., Sung, L. M., Witonsky, J., Zacharia, L. F., Kennedy, K., Jamieson, R., Stewart, J. and the International HapMap Consortium. (2005). A haplotype map of the human genome. *Nature* 437, 1299–1320.

Beaulieu, C. (2002). The basis of anisotropic water diffusion in the nervous system – a technical review. *NMR Biomed.* 15, 433–455.

Benjamini, Y., and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc. Series B Stat. Methodol.* 57, 289–300.

Bovelstad, H. M., Nygard, S., Storvold, H. L., Aldrin, M., Borgan, O., Frigessi, A., and Lingjaerde, O. C. (2007). Predicting survival from microarray data – a comparative study. *Bioinformatics* 23, 2080–2087.

Bralten, J., Arias-Vasquez, A., Makkinje, R., Veltman, J. A., Brunner, H. G., Fernandez, G., Rijpkema, M., and Franke, B. (2011). Association of the Alzheimer's gene SORL1 with hippocampal volume in young, healthy adults. *Am. J. Psychiatry* 168, 1083–1089.

Braskie, M. N., Jahanshad, N., Stein, J. L., Barysheva, M., McMahon, K. L., de Zubicaray, G. I., Martin, N. G., Wright, M. J., Ringman, J. M., Toga, A. W., and Thompson, P. M. (2011a). Common Alzheimer's disease risk variant within the CLU gene affects white matter microstructure in young adults. *J. Neurosci.* 31, 6764–6770.

Braskie, M. N., Ringman, J. M., and Thompson, P. M. (2011b). Neuroimaging measures as endophenotypes in Alzheimer's disease. *Int. J. Alzheimers Dis.* 31, 490140.

Cantor, R. M., Lange, K., and Sinsheimer, J. S. (2010). Prioritizing GWAS results: a review of statistical methods and recommendations for their application. *Am. J. Hum. Genet.* 86, 6–22.

Chen, L. S., Hutter, C. M., Potter, J. D., Liu, Y., Prentice, R. L., Peters, U., and Hsu, L. (2010). Insights into colon cancer etiology via a regularized approach to gene set analysis of GWAS data. *Am. J. Hum. Genet.* 86, 860–871.

Chiang, M. C., Barysheva, M., Shattuck, D. W., Lee, A. D., Madsen, S. K., Avedissian, C., Klunder, A. D., Toga, A. W., McMahon, K. L., de Zubicaray, G. I., Wright, M. J., Srivastava, A., Balov, N., and Thompson, P. M. (2009). Genetics of brain fiber architecture and intellectual performance. *J. Neurosci.* 29, 2212–2224.

Chiang, M. C., Barysheva, M., Toga, A. W., Medland, S. E., Hansell, N. K., James, M. R., McMahon, K. L., de Zubicaray, G. I., Martin, N. G., Wright, M. J., and Thompson, P. M. (2011a). BDNF gene effects on brain circuitry replicated in 455 twins. *Neuroimage* 55, 448–454.

Chiang, M. C., Barysheva, M., McMahon, K. L., de Zubicaray, G. I., Johnson, K., Martin, N. G., Toga, A. W., Wright, M. J., and Thompson, P. M. (2011b). "Hierarchical clustering of the genetic connectivity matrix reveals the network topology of gene action on brain microstructure: an N=531 twin study," in *International Workshop on Biomedical Imaging (ISBI)*, Chicago, IL, 832–835.

Cho, S., Kim, H., Oh, S., Kim, K., and Park, T. (2009). Elastic-net regularization approaches for genome-wide association studies of rheumatoid arthritis. *BMC Proc.* 15, S7–S25. PMID: 20018015.

Chumbley, J., Worsley, K., Flandin, G., and Friston, K. (2010). Topological FDR for neuroimaging. *Neuroimage* 49, 3057–3064.

Cruz, J. A., and Wishart, D. S. (2007). Applications of machine learning in cancer prediction and prognosis. *Cancer Inform.* 2, 59–77.

de Bakker, P. I., Ferreira, M. A., Jia, X., Neale, B. M., Raychaudhuri, S., and Voight, B. F. (2008). Practical aspects of imputation-driven meta-analysis of genome-wide association studies. *Hum. Mol. Genet.* 17, R122–R128.

Efron, B., Hastie, T., Johnstone, I., and Tibshirani, R. (2004). Least angle regression. *Ann. Stat.* 32, 407–451.

Ewens, W. J., and Grant, G. (2001). *Statistical Methods in Bioinformatics: An Introduction.* New York: Springer.

Fan, J. Q., and Lv, J. C. (2008). Sure independence screening for ultra-high dimensional feature space. *J. R. Stat. Soc. Series B Stat. Methodol.* 70, 849–883.

Fan, J. Q., and Song, R. (2010). Sure independence screening in generalized linear models with Np-dimensionality. *Ann. Stat.* 38, 3567–3604.

Fischl, B., Salat, D. H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., Van Der Kouwe, A., Killiany, R., Kennedy, D., Klaveness, S., Montillo, A., Makris, N., Rosen, B., and Dale, A. M. (2002). Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron* 33, 341–355.

Frayling, T. M., Timpson, N. J., Weedon, M. N., Zeggini, E., Freathy, R. M., Lindgren, C. M., Perry, J. R., Elliott, K. S., Lango, H., Rayner, N. W., Shields, B., Harries, L. W., Barrett, J. C., Ellard, S., Groves, C. J., Knight, B., Patch, A. M., Ness, A. R., Ebrahim, S., Lawlor, D. A., Ring, S. M., Ben-Shlomo, Y., Jarvelin, M. R., Sovio, U., Bennett, A. J., Melzer, D., Ferrucci, L., Loos, R. J., Barroso, I., Wareham, N. J., Karpe, F., Owen, K. R., Cardon, L. R., Walker, M., Hitman, G. A., Palmer, C. N., Doney, A. S., Morris, A. D., Smith, G. D., Hattersley, A. T., and McCarthy, M. I. (2007). A common variant in the FTO gene is associated with body mass index and predisposes to childhood and adult obesity. *Science* 316, 889–894.

Frazer, K. A., Ballinger, D. G., Cox, D. R., Hinds, D. A., Stuve, L. L., Gibbs, R. A., Belmont, J. W., Boudreau, A., Hardenbol, P., Leal, S. M., Pasternak, S., Wheeler, D. A., Willis, T. D., Yu, F., Yang, H., Zeng, C., Gao, Y., Hu, H., Hu, W., Li, C., Lin, W., Liu, S., Pan, H., Tang, X., Wang, J., Wang, W., Yu, J., Zhang, B., Zhang, Q., Zhao, H., Zhou, J., Gabriel, S. B., Barry, R., Blumenstiel, B., Camargo, A., Defelice, M., Faggart, M., Goyette, M., Gupta, S., Moore, J., Nguyen, H., Onofrio, R. C., Parkin, M., Roy, J., Stahl, E., Winchester, E., Ziaugra, L., Altshuler, D., Shen, Y., Yao, Z., Huang, W., Chu, X., He, Y., Jin, L., Liu, Y., Sun, W., Wang, H., Wang, Y., Xiong, X., Xu, L., Waye, M. M., Tsui, S. K., Xue, H., Wong, J. T., Galver, L. M., Fan, J. B., Gunderson, K., Murray, S. S., Oliphant, A. R., Chee, M. S., Montpetit, A., Chagnon, F., Ferretti, V., Leboeuf, M., Olivier, J. F., Phillips, M. S., Roumy, S., Sallee, C., Verner, A., Hudson, T. J., Kwok, P. Y., Cai, D., Koboldt, D. C., Miller, R. D., Pawlikowska, L., Taillon-Miller, P., Xiao, M., Tsui, L. C., Mak, W., Song, Y. Q., Tam, P. K., Nakamura, Y., Kawaguchi, T., Kitamoto, T., Morizono, T., Nagashima, A., Ohnishi, Y., Sekine, A., Tanaka, T., Tsunoda, T., Deloukas, P., Bird, C. P., Delgado, M., Dermitzakis, E. T., Gwilliam, R., Hunt, S., Morrison, J., Powell, D., Stranger, B. E., Whittaker, P., Bentley, D. R., Daly, M. J., de Bakker, P. I., Barrett, J., Chretien, Y. R., Maller, J., McCarroll, S., Patterson, N., Pe'er, I., Price, A., Purcell, S., Richter, D. J., Sabeti, P., Saxena, R., Schaffner, S. F., Sham, P. C., Varilly, P., Altshuler, D., Stein, L. D., Krishnan, L., Smith, A. V., Tello-Ruiz, M. K., Thorisson, G. A., Chakravarti, A., Chen, P. E., Cutler, D. J., Kashuk, C. S., Lin, S., Abecasis, G. R., Guan, W., Li, Y., Munro, H. M., Qin, Z. S., Thomas, D. J., McVean, G., Auton, A., Bottolo, L., Cardin, N., Eyheramendy, S., Freeman, C., Marchini, J., Myers, S., Spencer, C., Stephens, M., Donnelly, P., Cardon, L. R., Clarke, G., Evans, D. M., Morris, A. P., Weir, B. S., Tsunoda, T., Mullikin, J. C., Sherry, S. T., Feolo, M., Skol, A., Zhang, H., Zeng, C., Zhao, H., Matsuda, I., Fukushima, Y., Macer, D. R., Suda, E., Rotimi, C. N., Adebamowo, C. A., Ajayi, I., Aniagwu, T., Marshall, P. A., Nkwodimmah, C., Royal, C. D., Leppert, M. F., Dixon, M., Peiffer, A., Qiu, R., Kent, A., Kato, K., Niikawa, N., Adewole, I. F., Knoppers, B. M., Foster, M. W., Clayton, E. W., Watkin, J., Gibbs, R. A., Belmont, J. W., Muzny, D., Nazareth, L., Sodergren, E., Weinstock, G. M., Wheeler, D. A., Yakub, I., Gabriel, S. B., Onofrio, R. C., Richter, D. J., Ziaugra, L., Birren, B. W., Daly, M. J., Altshuler, D., Wilson, R. K., Fulton, L. L., Rogers, J., Burton, J., Carter, N. P., Clee, C. M., Griffiths, M., Jones, M. C., McLay, K., Plumb, R. W., Ross, M. T., Sims, S. K., Willey, D. L., Chen, Z., Han, H., Kang, L., Godbout, M., Wallenburg, J. C., L'Archevêque, P., Bellemare, G., Saeki, K., Wang, H., An, D., Fu, H., Li, Q., Wang, Z., Wang, R., Holden, A. L., Brooks, L. D., McEwen, J. E., Guyer, M. S., Wang, V. O., Peterson, J. L., Shi, M., Spiegel, J., Sung, L. M., Zacharia, L. F., Collins, F. S., Kennedy, K., Jamieson, R., and Stewart, J. (2007). A second generation human haplotype map of over 3.1 million SNPs. *Nature* 449, 851–861.

Friedman, J., Hastie, T., and Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *J. Stat. Softw.* 33, 1–22.

Gianola, D., Perez-Enciso, M., and Toro, M. A. (2003). On marker-assisted prediction of genetic value: beyond the ridge. *Genetics* 163, 347–365.

Glahn, D. C., Thompson, P. M., and Blangero, J. (2007). Neuroimaging endophenotypes: strategies for finding genes influencing brain structure and function. *Hum. Brain Mapp.* 28, 488–501.

Good, C. D., Johnsrude, I. S., Ashburner, J., Henson, R. N., Friston, K. J., and Frackowiak, R. S. (2001). A voxel-based morphometric study of ageing in 465 normal adult human brains. *Neuroimage* 14, 21–36.

Gottesman, I. I., and Gould, T. D. (2003). The endophenotype concept in psychiatry: etymology and strategic intentions. *Am. J. Psychiatry* 160, 636–645.

Hall, M. H., and Smoller, J. W. (2010). A new role for endophenotypes in the GWAS era: functional characterization of risk variants. *Harv. Rev. Psychiatry* 18, 67–74.

Hariri, A. R., Mattay, V. S., Tessitore, A., Kolachana, B., Fera, F., Goldman, D., Egan, M. F., and Weinberger, D. R. (2002). Serotonin transporter genetic variation and the response of the human amygdala. *Science* 297, 400–403.

He, Q., and Lin, D. Y. (2011). A variable selection method for genome-wide association studies. *Bioinformatics* 27, 1–8.

Heinz, A., Jones, D. W., Mazzanti, C., Goldman, D., Ragan, P., Hommer, D., Linnoila, M., and Weinberger, D. R. (2000). A relationship between serotonin transporter genotype and in vivo protein expression and alcohol neurotoxicity. *Biol. Psychiatry* 47, 643–649.

Hemminger, B. M., Saelim, B., and Sullivan, P. F. (2006). TAMAL: an integrated approach to choosing SNPs for genetic studies of human complex traits. *Bioinformatics* 22, 626–627.

Hibar, D. P., Stein, J. L., Kohannim, O., Jahanshad, N., Saykin, A. J., Shen, L., Kim, S., Pankratz, N., Foroud, T., Huentelman, M. J., Potkin, S. G., Jack, C. R. Jr., Weiner, M. W., Toga, A. W., and Thompson, P. M. (2011). Voxelwise gene-wide association study (vGeneWAS): multivariate gene-based association testing in 731 elderly subjects. *Neuroimage* 56, 1875–1891.

Hindorff, L. A., Sethupathy, P., Junkins, H. A., Ramos, E. M., Mehta, J. P., Collins, F. S., and Manolio, T. A. (2009). Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc. Natl. Acad. Sci. U.S.A.* 106, 9362–9367.

Hinrichs, A. S., Karolchik, D., Baertsch, R., Barber, G. P., Bejerano, G., Clawson, H., Diekhans, M., Furey, T. S., Harte, R. A., Hsu, F., Hillman-Jackson, J., Kuhn, R. M., Pedersen, J. S., Pohl, A., Raney, B. J., Rosenbloom, K. R., Siepel, A., Smith, K. E., Sugnet, C. W., Sultan-Qurraie, A., Thomas, D. J., Trumbower, H., Weber, R. J., Weirauch, M., Zweig, A. S., Haussler, D., and Kent, W. J. (2006). The UCSC genome browser database: update 2006. *Nucleic Acids Res.* 34, D590–D598.

Ho, A. J., Stein, J. L., Hua, X., Lee, S., Hibar, D. P., Leow, A. D., Dinov, I. D., Toga, A. W., Saykin, A. J., Shen, L., Foroud, T., Pankratz, N., Huentelman, M. J., Craig, D. W., Gerber, J. D., Allen, A. N., Corneveaux, J. J., Stephan, D. A., DeCarli, C. S., DeChairo, B. M., Potkin, S. G., Jack, C. R. Jr., Weiner, M. W., Raji, C. A., Lopez, O. L., Becker, J. T., Carmichael, O. T., Thompson, P. M., and Alzheimer's Disease Neuroimaging Initiative. (2010). A commonly carried allele of the obesity-related FTO gene is associated with reduced brain volume in the healthy elderly. *Proc. Natl. Acad. Sci. U.S.A.* 107, 8404–8409.

Hoerl, A. E. (1962). Application of ridge analysis to regression problems. *Chem. Eng. Prog.* 58, 54–59.

Hoh, J., Wille, A., and Ott, J. (2001). Trimming, weighting, and grouping SNPs in human case-control association studies. *Genome Res.* 11, 2115–2119.

Ioannidis, J. P. (2005). Why most published research findings are false. *PLoS Med.* 2, e124. doi:10.1371/journal.pmed.0020124

Jagannathan, K., Calhoun, V. D., Gelernter, J., Stevens, M. C., Liu, J., Bolognani, F., Windemuth, A., Ruaño, G., Assaf, M., and Pearlson, G. D. (2010). Genetic associations of brain structural networks in schizophrenia: a preliminary study. *Biol. Psychiatry* 68, 657–666.

Jolliffe, I. T. (2002). *Principal Component Analysis.* New York: Springer.

Joyner, A. H., Roddey, J. C., Bloss, C. S., Bakken, T. E., Rimol, L. M., Melle, I., Agartz, I., Djurovic, S., Topol, E. J., Schork, N. J., Andreassen, O. A., and Dale, A. M. (2009). A common MECP2 haplotype associates with reduced cortical surface area in humans in two independent populations. *Proc. Natl. Acad. Sci. U.S.A.* 106, 15483–15488.

Kleinbaum, D. G. (2007). *Applied Regression Analysis and Other Multivariable Methods.* Belmont, CA: Brooks/Cole.

Kochunov, P., Glahn, D. C., Lancaster, J. L., Winkler, A. M., Smith, S., Thompson, P. M., Almasy, L., Duggirala, R., Fox, P. T., and Blangero, J. (2010). Genetics of microstructure of cerebral white matter using diffusion tensor imaging. *Neuroimage* 53, 1109–1116.

Kohannim, O., Hibar, D. P., Stein, J. L., Jahanshad, N., Jack, C. R. Jr., Weiner, M. W., Toga, A. W., and Thompson, P. M. (2011a). "Boosting power to detect genetic associations in imaging using multi-locus, genome-wide scans and ridge regression," *International Workshop on Biomedical Imaging (ISBI)*, Chicago, IL, 1855–1859.

Kohannim, O., Jahanshad, N., Braskie, M. N., Stein, J. L., Chiang, M. C., Reese, A. H., Toga, A. W., McMahon, K. L., de Zubicaray, G. I., Medland, S. E., Montgomery, G. M., Whitfield, J. B., Martin, N. G., Wright, M. W., and Thompson, P. M. (2011b). Personalized prediction of brain fiber integrity in 396 young adults based on genotyping of multiple common genetic variants. *Abstr. Soc. Neurosci.*

Leow, A., Huang, S. C., Geng, A., Becker, J., Davis, S., Toga, A., and Thompson, P. (2005). Inverse consistent mapping in 3D deformable image registration: its construction and statistical properties. *Inf. Process. Med. Imaging* 19, 493–503.

Li, Y., Sung, W., and Liu, J. J. (2007). Association mapping via regularized regression analysis of single-nucleotide–polymorphism haplotypes in variable-sized sliding windows. *Am. J. Hum. Genet.* 80, 705–715.

Lin, Y., Zhang, M., Wang, L., Pungpapong, V., Fleet, J. C., and Zhang, D. (2009). Simultaneous genome-wide association studies of anti-cyclic citrullinated peptide in rheumatoid arthritis using penalized orthogonal-components regression. *BMC Proc.* 15, S7–S20. PMID: 20018010.

Liu, J., Pearlson, G., Windemuth, A., Ruano, G., Perrone-Bizzozero, N. I., and Calhoun, V. (2009). Combining fMRI and SNP data to investigate connections between brain function and genetics using parallel ICA. *Hum. Brain Mapp.* 30, 241–255.

Malo, N., Libiger, O., and Schork, N. J. (2008). Accommodating linkage disequilibrium in genetic-association analyses via ridge regression. *Am. J. Hum. Genet.* 82, 375–385.

Massy, W. F. (1965). Principal components regression in exploratory statistical research. *J. Am. Stat. Assoc.* 60, 234–256.

McCarthy, M. I., Abecasis, G. R., Cardon, L. R., Goldstein, D. B., Little, J., Ioannidis, J. P., and Hirschhorn, J. N. (2008). Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nat. Rev. Genet.* 9, 356–369.

Meyer-Lindenberg, A., and Weinberger, D. R. (2006). Intermediate phenotypes and genetic mechanisms of psychiatric disorders. *Nat. Rev. Neurosci.* 7, 818–827.

Munafo, M. R., Brown, S. M., and Hariri, A. R. (2008). Serotonin transporter (5-HTTLPR) genotype and amygdala activation: a meta-analysis. *Biol. Psychiatry* 63, 852–857.

Pearson, T. A., and Manolio, T. A. (2008). How to interpret a genome-wide association study. *JAMA* 299, 1335–1344.

Phillips, T. J., and Belknap, J. K. (2002). Complex-trait genetics: emergence of multivariate strategies. *Nat. Rev. Neurosci.* 3, 478–485.

Potkin, S. G., Guffanti, G., Lakatos, A., Turner, J. A., Kruggel, F., Fallon, J. H., Saykin, A. J., Orro, A., Lupoli, S., Salvi, E., Weiner, M., Macciardi, F., and Alzheimer's Disease Neuroimaging Initiative. (2009a). Hippocampal atrophy as a quantitative trait in a genome-wide association study identifying novel susceptibility genes for Alzheimer's disease. *PLoS ONE* 4, e6501. doi:10.1371/journal.pone.0006501

Potkin, S. G., Turner, J. A., Fallon, J. A., Lakatos, A., Keator, D. B., Guffanti, G., and Macciardi, F. (2009b). Gene discovery through imaging genetics: identification of two novel genes associated with schizophrenia. *Mol. Psychiatry* 14, 416–428.

Psychiatric GWAS Consortium Coordinating Committee, Cichon, S., Craddock, N., Daly, M., Faraone, S. V., Gejman, P. V., Kelsoe, J., Lehner, T., Levinson, D. F., Moran, A., Sklar, P., and Sullivan, P. F. (2009). Genomewide association studies: history, rationale, and prospects for psychiatric disorders. *Am. J. Psychiatry* 166, 540–556.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., Maller, J., Sklar, P., De Bakker, P. I., Daly, M. J., and Sham, P. C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81, 559–575.

Reiman, E. M., Webster, J. A., Myers, A. J., Hardy, J., Dunckley, T., Zismann, V. L., Joshipura, K. D., Pearson, J. V., Hu-Lince, D., Huentelman, M. J., Craig, D. W., Coon, K. D., Liang, W. S., Herbert, R. H., Beach, T., Rohrer, K. C., Zhao, A. S., Leung, D., Bryden, L., Marlowe, L., Kaleem, M., Mastroeni, D., Grover, A., Heward, C. B., Ravid, R., Rogers, J., Hutton, M. L., Melquist, S., Petersen, R. C., Alexander, G. E., Caselli, R. J., Kukull, W., Papassotiropoulos, A., and Stephan, D. A. (2007). GAB2 alleles modify Alzheimer's risk in APOE epsilon4 carriers. *Neuron* 54, 713–720.

Shaw, L. M., Korecka, M., Clark, C. M., Lee, V. M., and Trojanowski, J. Q. (2007). Biomarkers of neurodegeneration for diagnosis and monitoring therapeutics. *Nat. Rev. Drug Discov.* 6, 295–303.

Shen, L., Kim, S., Risacher, S. L., Nho, K., Swaminathan, S., West, J. D., Foroud, T., Pankratz, N., Moore, J. H., Sloan, C. D., Huentelman, M. J., Craig, D. W., Dechairo, B. M., Potkin, S. G., Jack, C. R. Jr., Weiner, M. W., and Saykin, A. J. (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, 1051–1063.

Stein, J. L., Hibar, D. P., Madsen, S. K., Khamis, M., Mcmahon, K. L., De Zubicaray, G. I., Hansell, N. K., Montgomery, G. W., Martin, N. G., Wright, M. J., Saykin, A. J., Jack, C. R. Jr., Weiner, M. W., Toga, A. W., and Thompson, P. M. (2011). Discovery and replication of dopamine-related gene effects on caudate volume in young and elderly populations (N=1198) using genome-wide search. *Mol. Psychiatry* 16, 927–937.

Stein, J. L., Hua, X., Morra, J. H., Lee, S., Hibar, D. P., Ho, A. J., Leow, A. D., Toga, A. W., Sul, J. H., Kang, H. M., Eskin, E., Saykin, A. J., Shen, L., Foroud, T., Pankratz, N., Huentelman, M. J., Craig, D. W., Gerber, J. D., Allen, A. N., Corneveaux, J. J., Stephan, D. A., Webster, J., DeChairo, B. M., Potkin, S. G., Jack, C. R. Jr., Weiner, M. W., Thompson, P. M., and Alzheimer's Disease Neuroimaging Initiative. (2010a). Genome-wide analysis reveals novel genes influencing temporal lobe structure with relevance to neurodegeneration in Alzheimer's disease. *Neuroimage* 51, 542–554.

Stein, J. L., Hua, X., Lee, S., Ho, A. J., Leow, A. D., Toga, A. W., Saykin, A. J., Shen, L., Foroud, T., Pankratz, N., Huentelman, M. J., Craig, D. W., Gerber, J. D., Allen, A. N., Corneveaux, J. J., Dechairo, B. M., Potkin, S. G., Weiner, M. W., and Thompson, P. (2010b). Voxelwise genome-wide association study (vGWAS). *Neuroimage* 53, 1160–1174.

The ENIGMA Consortium. (2011). "Genome-wide association meta-analysis of hippocampal volume: results from the ENIGMA Consortium," in *Organization for Human Brain Mapping Conference*, Quebec City.

Thomason, M. E., Dougherty, R. F., Colich, N. L., Perry, L. M., Rykhlevskaia, E. I., Louro, H. M., Hallmayer, J. F., Waugh, C. E., Bammer, R., Glover, G. H., and Gotlib, I. H. (2010). COMT genotype affects prefrontal white matter pathways in children and adolescents. *Neuroimage* 53, 926–934.

Thomason, M. E., and Thompson, P. M. (2011). Diffusion imaging, white matter, and psychopathology. *Annu. Rev. Clin. Psychol.* 7, 63–85.

Thompson, P. M., Cannon, T. D., Narr, K. L., Van Erp, T., Poutanen, V. P., Huttunen, M., Lonnqvist, J., Standertskjold-Nordenstam, C. G., Kaprio, J., Khaledy, M., Dail, R., Zoumalan, C. I., and Toga, A. W. (2001). Genetic influences on brain structure. *Nat. Neurosci.* 4, 1253–1258.

Tibshirani, R. (1996). Regression Shrinkage and Selection via the Lasso. *J. R. Stat. Soc. Series B Stat. Methodol.* 58, 267–288.

Vounou, M., Nichols, T. E., Montana, G., and Alzheimer's Disease Neuroimaging Initiative. (2010). Discovering genetic associations with high-dimensional neuroimaging phenotypes: a sparse reduced-rank regression approach. *Neuroimage* 53, 1147–1159.

Wang, K., and Abbott, D. (2008). A principal components regression approach to multilocus genetic association studies. *Genet. Epidemiol.* 32, 108–118.

Winkler, A. M., Kochunov, P., Blangero, J., Almasy, L., Zilles, K., Fox, P. T., Duggirala, R., and Glahn, D. C. (2010). Cortical thickness or grey matter volume? The importance of selecting the phenotype for imaging genetics studies. *Neuroimage* 53, 1135–1146.

Yang, J., Benyamin, B., Mcevoy, B. P., Gordon, S., Henders, A. K., Nyholt, D. R., Madden, P. A., Heath, A. C., Martin, N. G., Montgomery, G. W., Goddard, M. E., and Visscher, P. M. (2010). Common SNPs explain a large proportion of the heritability for human height. *Nat. Genet.* 42, 565–569.

Zeggini, E., and Ioannidis, J. P. (2009). Meta-analysis in genome-wide association studies. *Pharmacogenomics* 10, 191–201.

Zhang, D., Lin, Y., and Zhang, M. (2009). Penalized orthogonal-components regression for large p small n Data. *Electron. J. Stat.* 3, 781–796.

Zou, H., and Hastie, T. (2005). Regularization and variable selection via the elastic net. *J. R. Stat. Soc. Series B Stat. Methodol.* 67, 301–320.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 28 July 2011; paper pending published: 15 August 2011; accepted: 03 October 2011; published online: 21 October 2011.

Citation: Hibar DP, Kohannim O, Stein JL, Chiang M-C and Thompson PM (2011) Multilocus genetic analysis of brain images. Front. Gene. ***2****:73. doi: 10.3389/fgene.2011.00073*

This article was submitted to Frontiers in Statistical Genetics and Methodology, a specialty of Frontiers in Genetics.

2.2 Principal components regression: multivariate, gene-based test

This section is adapted from:

Derrek P. Hibar, Jason L. Stein, Omis Kohannim, Neda Jahanshad, Clifford R. Jack, Jr., Michael W. Weiner, Arthur W. Toga, Paul M. Thompson, and the Alzheimer's Disease Neuroimaging Initiative. Principal components regression: multivariate, gene-based tests in imaging genomics, ISBI, 2011.

PRINCIPAL COMPONENTS REGRESSION: MULTIVARIATE, GENE-BASED TESTS IN IMAGING GENOMICS

Derrek P. Hibar[1], Jason L. Stein[1], Omid Kohannim[1], Neda Jahanshad[1], Clifford R. Jack, Jr.[2], Michael W. Weiner[3,4], Arthur W. Toga[1], Paul M. Thompson[1], and the Alzheimer's Disease Neuroimaging Initiative

[1]Laboratory of Neuro Imaging, Dept. of Neurology, UCLA School of Medicine, Los Angeles, CA, [2]Mayo Clinic, Rochester, MN, [3]Depts. of Radiology, Medicine and Psychiatry, UC San Francisco, San Francisco, CA, [4]Dept. of Veterans Affairs Medical Center, San Francisco, CA

ABSTRACT

In imaging genomics, there have been rapid advances in genome-wide, image-wide searches for genes that influence brain structure. Most efforts focus on univariate tests that treat each genetic variation independently, ignoring the joint effects of multiple variants. Instead, we present a *gene-based* method to detect the joint effect of multiple single nucleotide polymorphisms (SNPs) in 18,044 genes across 31,662 voxels of the whole brain in a tensor-based morphometry analysis of baseline MRI scans from 731 subjects from the Alzheimer's Disease Neuroimaging Initiative (ADNI). Our *gene-based* multivariate statistics use principal components regression to test the combined effect of multiple genetic variants on an image, using a single test statistic. In some situations, which we describe, this can boost power by encoding population variations within each gene, reducing the effective number of statistical tests, and reducing the effect dimension of the search space. Multivariate gene-based methods may discover gene effects undetectable with standard, univariate methods, accelerating ongoing imaging genomics efforts worldwide.

Index Terms— principal components regression, multivariate, voxelwise, imaging genomics, GWAS

1. INTRODUCTION

In imaging genomics, the vast amount of information in the images (>100,000 voxels) and across the genome (>12 million known variants) presents computational and statistical challenges when relating genetic variants to the structure and function of the brain. Power issues arise due to the small effect sizes of each genetic variant, and the huge numbers of statistical comparisons. Most techniques use some type of data reduction, limiting the number of genetic variants or imaging features studied, or both. The ultimate goal of these gene-hunting studies is to create a method that discovers which genetic variants affect the brain in a statistically powerful and biologically meaningful way.

In typical GWAS studies, each genetic variant (usually a SNP) is independently tested for its association to the phenotype – a mass univariate method, where no data reduction is used across the genome. For example, one study [1] performed a genome-wide search of around 500,000 SNPs, and found a novel variant in the *GRIN2B* gene that is associated with temporal lobe volume. The gene *GRIN2B* encodes a glutamate receptor that is already the target of drugs (memantine) used to treat Alzheimer's disease. Findings such as these are promising as they have biological relevance, but do not rely on a prior hypothesis about any specific SNP. However, performing mass univariate methods on imaging summary measures (such as temporal lobe volume) or ad hoc regions of interest (ROI), collapses the variation across the brain into a single number.

Several studies now perform genome-wide searches at each voxel across the brain [2]. This approach avoids having to pre-select an *ad hoc* brain region of interest and does not require prior hypotheses about which genetic variants, or which regions of interest, matter. One study [3] performed a genome-wide, brain-wide search, termed a voxelwise genome-wide association study (vGWAS), in 740 subjects from ADNI. However, none of the genetic variants identified was significant after multiple comparisons correction; several variants were promising candidates for further analysis. Future GWAS studies in imaging will likely need to reduce the number of tests and multiple comparisons using Bayesian priors, machine learning, or dimension reduction in the image or the genome. This may prioritize certain regions of the image or the genome, for later meta-analysis across multiple datasets.

Given recent advances in high-throughput genotyping, densely-packed sets of SNPs, or genetic markers, can capture increasing amounts of variation throughout the genome. Methods that consider combinations of SNPs from the same gene should more accurately describe gene effects on images than methods that test the

independent effect of each SNP [4]. By associating the joint effect of multiple SNPs within a gene, in this study we set out to show that gene-based approaches can be more powerful, in some situations, than traditional univariate approaches. For example, if a gene contains multiple causal variants with small individual effects, univariate methods would miss these associations if a very stringent significance threshold is used (as in GWAS).

We assessed whether it would be feasible to extend to a neuroimaging database, a gene-based association method using principal components regression (PCReg). We applied PCReg across all genes, to a large database of voxelwise imaging data. We call our method a voxelwise "gene-wide" association study (vGeneWAS). By performing association tests on whole genes, we greatly reduce the number of tests (from 437,607 SNPs down to 18,044 genes). Using a voxel-based approach, we also avoid known problems associated with focusing on ROIs or summary measures. In addition, we performed direct power comparisons between gene-based tests using PCReg versus traditional univariate regression methods for GWAS.

2. METHODS

2.1. Imaging Measures

Structural MRI data were obtained following the standard ADNI protocol to ensure multisite consistency. Baseline MRI scans for each subject were analyzed using tensor-based morphometry (TBM) as described previously [5]. After quality control selection there were 731 subjects with genotyping data available (172 AD, 356 MCI, and 203 healthy elderly controls; 301 women/430 men; mean age ± sd = 75.56 ± 6.78 years). We did not split the subjects by diagnosis for this analysis, to exploit the broadest phenotypic continuum and maximize statistical power to detect genetic associations [6].

2.2. Genotypes and gene grouping

For details on how genetic data were processed for the ADNI study, please see [7]. We used several quality control measures to filter our SNPs for our analysis as detailed in [1]. Briefly, SNPs were excluded with call rate <95%, significant deviation from Hardy-Weinberg equilibrium P < $5.7\text{x}10^{-7}$, and a minor allele frequency <0.10. After all rounds of quality control and preparation, 437,607 SNPs remained. Remaining SNPs were then grouped by gene, where "gene" is defined by the gene transcript region including both introns and exons. SNPs not located in a gene were excluded. After quality control, SNP annotation, and gene grouping, 18,044 genes were left for analysis.

2.3 Multi-SNP genetic associations

To test the joint effect of all SNPs in a gene on the volume difference (calculated from TBM) at each voxel, we employed a multiple partial-*F* test. This first estimates the fit of a "reduced model" of any number of nuisance variables on a given dependent variable and then estimates the fit of a second "full model" with the nuisance variables and any number of independent variables on the same dependent variable. Each association test results in an *F*-statistic, which represents the joint effect of the independent variables on the dependent variable, controlling for nuisance variables already in the model. The multiple partial-*F* statistic was calculated for each gene at each voxel using equation 1 below. Here *k* is df(full)-df(reduced) and RSS is the residual sum of squares:

$$F_{k,df(full)} = \frac{RSS(reduced) - RSS(full)}{df(reduced) - df(full)} \Big/ \frac{RSS(full)}{df(full)} \qquad (1)$$

Multiple partial-*F* tests are well suited for testing effects of multiple predictors on a given phenotype, but genetic data sometimes complicates testing because SNPs in the same gene are often correlated due to high "linkage disequilibrium" (LD). When the SNP values in a cohort of subjects are treated as a vector (whose components are the SNP value in each subject coded in an additive manner: 0, 1, or 2), then statistical correlations between adjacent SNPs on the genome can make different subjects' vectors highly collinear. The dependence among these almost collinear SNP vectors in the multiple partial-*F* test model can lead to improper signs of beta coefficient estimates, wildly inaccurate magnitudes of beta coefficients, large standard error estimates, and false inferences.

To avoid the complications of collinearity in the statistical model, we first performed principal component analysis (PCA) on the SNPs within each gene, storing all of the orthonormal basis vectors of the SNP matrix that explained the first 95% of the variance in the set of SNPs. Basis vectors with the highest eigenvalues (higher proportions of explained variance) were included until 95% of the variance in the SNPs was explained. The rest were discarded. These new "eigenSNPs" approximate the information in the observed SNPs, but lack the collinearity that disrupts the multiple partial-*F* test models. By first performing PCA followed by a multiple partial-*F* test, our method may be considered a variant of PCReg and produces *F*-statistics equivalent to those proposed previously for non-imaging data [8]. In this study, the independent variables built into the multiple partial-*F* test full model were the column vector output from PCA performed on each gene with age and sex as covariates. In this way, we tested the

joint predictive effect of variation throughout a gene on brain volume variations on a voxel-by-voxel level.

The total number of tests of association for vGeneWAS is very high (18,044 genes x 31,662 voxels). Because of the massive processing requirement, we coded a "threaded" version of the PCA and multiple partial-*F* test steps of PCReg to split processing over multiple cores in a single CPU. Processing was further parallelized over a cluster of 10 high-performance 8-core CPU nodes. As a data reduction step, we only saved data on the gene with the lowest *P*-value at each voxel (the "top gene" at each voxel). The total time required to complete an analysis was approximately 13 days.

2.4 Effective number of test for statistical thresholds

As we noted previously [3], the minimum *P*-value at each voxel, in the null case with *n* independent tests, approximately follows a probability density function (PDF) such that:

$$f_{\min}(x) = n(1-x)^{n-1} \qquad (2)$$

The PDF derived from equation 2 is known as a Beta distribution with parameters $\alpha=1$ and $\beta=n$. At each voxel, selecting the minimum *P*-value for the top gene then follows a Beta(1, *n*) distribution, where *n* is the effective number of independent tests.

However, genetic loci are inherited in contiguous segments, and some genes co-segregate in blocks. The allele frequencies and structure of genes that co-segregate are more similar than would be expected by chance if all variants were assumed to be independent. Because of this, the effective number of independent tests (M_{eff}) is less than the total number of tests performed (M). By determining M_{eff}, we can more accurately estimate the total number of independent tests performed, given the LD structure of our genotype data.

In our sample, we estimated M_{eff} by performing 5000 permutation tests at three randomly selected, uncorrelated voxels in the brain. We regressed each of the 18,044 genes on the permuted residuals of the reduced model after including the age and sex covariates at each run, and stored the minimum *P*-value. As only the minimum *P*-value is retained (for the best fitting gene), one can build up a reference distribution for the minimum *P*-values, to help gauge the level of surprise in seeing associations in the data. Storing the minimum *P*-values of the permutation tests yields the expected null Beta distribution given our data. We used a maximum-likelihood function to estimate the best fit for the null Beta distribution by varying the β parameter of Beta(1,β). The value of β approximates the effective number of independent tests (M_{eff}) performed on our data.

2.5 Estimation of expected values in simulated maps

A certain amount of spatial smoothness is expected among voxels in an image. This is most likely explained by the non-independence of volume difference measures at adjacent voxels. We examined whether the size of voxel clusters associated with the same gene from our vGeneWAS analysis differed from the cluster sizes expected under the null hypothesis of no association at all, given the non-independence of signals at adjacent voxels in our images. In addition, we wanted to determine whether the number of unique, top genes from across the brain significantly differed from the number of top genes expected by chance. We generated 100 3D simulated cluster maps based on a linking algorithm that forms connections between voxels across the brain based on correlation. The probability of any voxel being linked to another voxel was directly related to how correlated they are to each other. By considering the correlation of a given voxel to all other voxels in the image, as opposed to using a single summary measure of smoothness throughout an image, we were able to model the expected 3D clustering among adjacent voxels and non-independent, spatially separated clusters.

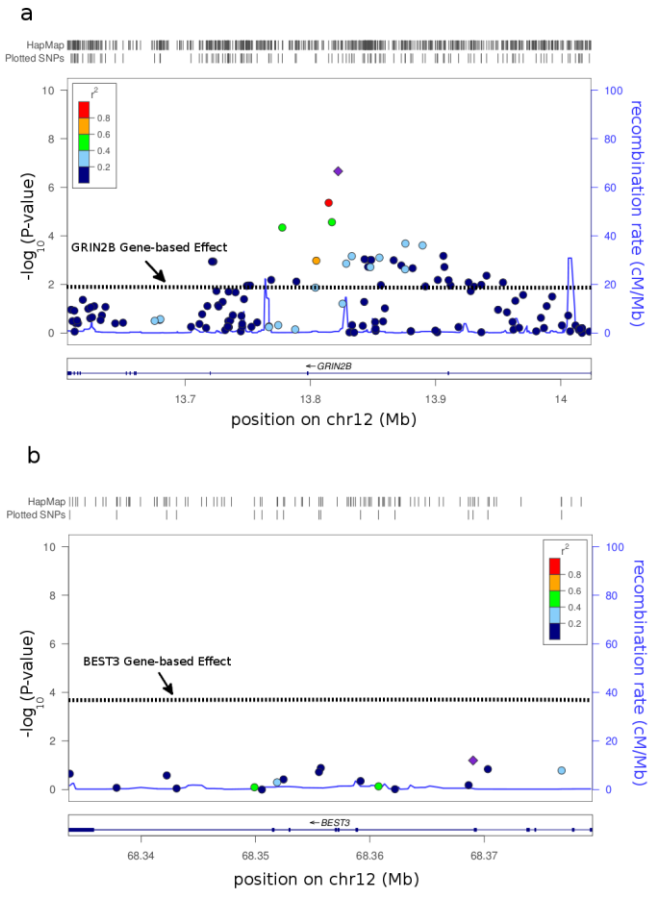
3. RESULTS

3.1 Comparison of methods

To examine differences between gene-based and standard univariate association methods, we compared the results of PCReg to linear regression using the temporal lobe volume (TLV) data from a previous study [1] as the phenotype. We first chose to focus on the top gene or SNP identified by each method, in order to examine performance when the variant chosen is deliberately selected to favor one of the two methods. *GRIN2B* was identified as the gene with the SNP variant that was most significantly associated with TLV using a standard univariate GWAS analysis ($P=4.03\text{x}10^{-7}$). We plotted the $-\log_{10}$(*P*-value) of the univariate test for each of the SNPs in the *GRIN2B* gene, in **Figure 1a.** The PCReg gene test results are overlaid (*black dotted line*). Clearly, the main effect detected with linear regression is much greater in this case, and the p-values are much smaller (i.e., $-\log_{10}$(*P*-value) is higher). Notably, we tested each of the 129 SNPs within the *GRIN2B* gene, which would require any significant *P*-values identified to be corrected for multiple comparisons before further study. In comparison, the gene-based test of *GRIN2B* using PCReg was a single test not requiring correction for multiple comparisons and maintained a nominal significance value ($P=0.012$). Also, we compared *BEST3* - the gene identified to be most significantly associated with TLV via PCReg - with the

linear regression output of each SNP within the gene (**Figure 1b**). The main effect of the gene-based test was much stronger (P=2.9x10^{-4}) than the best linear regression result (P=0.063). This demonstrates a case where variance components from individual markers are not significant via linear regression, but may be combined into a single significant test statistic.

Figure 1. Genetic association plots for univariate linear regression versus multi-locus PCReg. The -log$_{10}$(P-value) of each SNP in *GRIN2B* (a) and *BEST3* (b) is plotted against its position in the gene. Each of the points is color coded by level of LD (compared to the top SNP, *the purple diamond dot*) as measured by r^2.The -log$_{10}$(P-value) of the gene-based PCReg test for each gene is overlaid on the plot for comparison (*dotted black line*). Plots were generated using the LocusZoom software package.



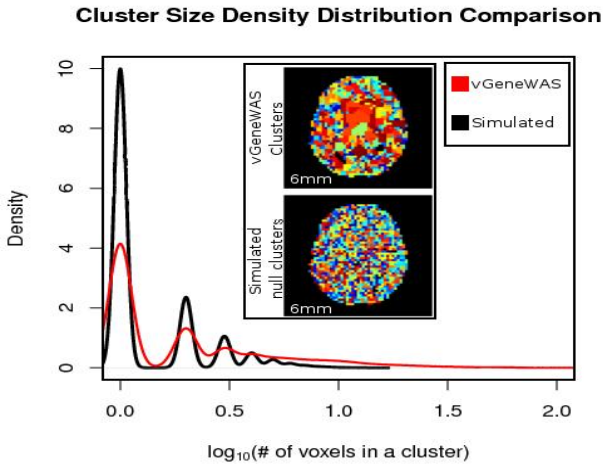
3.2 Voxelwise GeneWAS

By randomly permuting the images, so that they were not assigned to the correct individuals, we compared the distribution of the cluster size values in simulated (null) maps to the cluster sizes obtained from vGeneWAS (**Figure 2**). A large proportion of clusters of voxels associated with the same top gene in vGeneWAS were larger than would be expected based on completely null data. One estimate related to the number of independent voxels is the average number of clusters in simulated maps. This was 11900.8 ± 50.6 (mean ± standard deviation) out of the 31,662 total voxels. We used the number of clusters estimated from the simulation to randomly select (with replacement) from our list of 18,044 genes. We tallied the number of unique genes represented for each simulated cluster map and found the average was 8721.4 ± 44.9 (mean ± standard deviation). We measured the total number of unique genes as 5333 from our run of voxelwise GeneWAS, which is much lower than the number of genes expected based on the null cluster maps. Combined with our cluster size comparisons, this suggests that the top genes identified in our analysis tend to have a much more broadly distributed effect than would be expected if the data were null, even taking into account the intrinsic spatial non-independence of our data.

Among the top genes identified at each voxel across the brain, the *GRB-associated binding protein 2 gene*, *GAB2*, was the most significantly associated gene at any voxel (with P=2.36x10^{-9}) in our analysis and has previously been linked to late-onset Alzheimer's disease (LOAD). One study [9] identified 10 SNPs from the *GAB2* gene that were significantly associated with LOAD and *APOE* allele status in 1411 cases and controls from 20 NIA-sponsored Alzheimer's Disease Centers. *In vivo* testing shows that *GAB2* is over-expressed in certain brain regions such as the hippocampus and posterior cingulate cortex in patients with LOAD [9]. In addition, the *AlzGene* website lists *GAB2* as being in the top 20 genes likely related to AD (October 20, 2010; http://www.alzgene.org/). We identified several other genes highly relevant to brain function; a few are: *LRDD* (P=2.60x10^{-9}), *PRPRB* (P=2.84x10^{-9}), *CHRM5* (P=1.71x10^{-8}), and *S100B* (P=4.75x10^{-8}).

Figure 2. Cluster sizes in vGeneWAS (*red line*) **are compared with a simulated null map** (*black line*). The density of the number of voxels (log$_{10}$ transformed) in a cluster across the brain are plotted. The simulated null map contains a larger proportion of small cluster sizes than vGeneWAS (higher peaks in the *black line* at values close to the origin on the *x*-axis). The vGeneWAS map contains a larger proportion of large cluster sizes than the average simulated null map (the *red line* is higher at larger values and is more extended). A single slice view of the vGeneWAS and average simulated null cluster maps are pictured for comparison (*inset*). Every unique cluster is assigned its own color. There are more unique clusters than

distinct colors making visual inspection difficult, but in general the clusters in the vGeneWAS maps are larger.



3.3 Correction for multiple comparisons

Our Beta-distributed experimental *P*-values (with M_{eff}=15,636) need to be corrected so that their false discovery rate (FDR) can be assessed [10]. Using the analytic β parameter from the null Beta distribution, we fitted a cumulative distribution function (CDF) to our observed data yielding a new distribution of corrected *P*-values that deviate from the uniform distribution only when the data are not null.

We found that the false discovery rate for the second most highly associated gene in our results (*LRDD*) could only be controlled at a threshold of q=0.30 (i.e., allowing a 30% false discovery rate) after applying a statistical threshold of P_c=5.36x10^{-4}. In addition, the pFDR q-value threshold [11] was q=0.23 for the most significantly associated gene at any voxel (*GAB2*). In other words, the vGeneWAS results could not be controlled at the conventional false discovery rate, but show promise.

3.4 Power comparisons

To assess the differences in power afforded by vGeneWAS relative to existing univariate methods, we compared the P_c-values from vGWAS obtained in our previous study [3], with the P_c-values resulting from vGeneWAS (**Figure 3**). The proportion of P_c-values greater than a given FDR threshold for each method is directly related to differences in effect sizes. The FDR of the results from vGWAS could only be controlled at a threshold value of q=0.50, whereas the FDR threshold for vGeneWAS is somewhat lower, although not passing the conventional FDR level (q=0.30; **Figure 3**). This suggests that the vGeneWAS method may have more power, in principle, to detect genetic associations, although neither test controlled the false discovery rate at the conventional level.

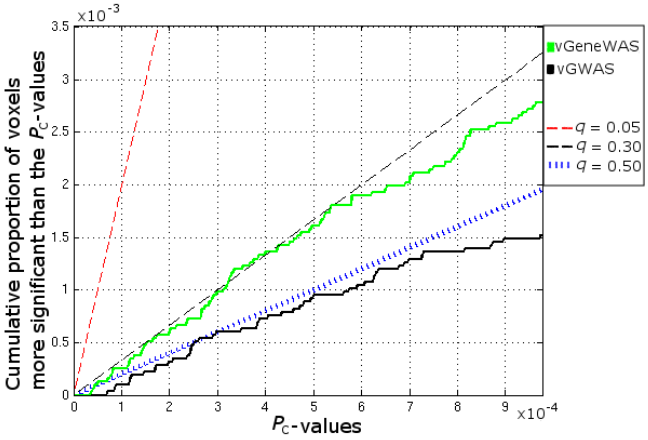


Figure 3. vGeneWAS may control the false discovery rate better than vGWAS. The cumulative distribution function (CDF) of P_c-values from vGeneWAS (*solid green line*) is compared to the CDF of P_c-values from vGWAS [3]. (*solid black line*). Three lines represent different correction thresholds of q=0.05 (*red dashed*), q=0.30 (*black dashed*), and q=0.50 (*blue dotted*).

4. CONCLUSION

We showed that, in certain cases, gene-based methods may offer more power than traditional univariate methods. In addition, our analysis identified a known Alzheimer's risk gene, *GAB2*, lending plausibility to the method. Still, effect sizes may be too small to detect even with multivariate statistics and meta-analytic approaches may prove most useful in the future (e.g., in multi-site efforts such as the ENIGMA consortium [12]).

5. ACKNOWLEDGMENTS

This study was supported by NIH grants EB008281 HD050735, RR013642, AG036535 (to PT).

6. REFERENCES

[1] Stein, J.L., et al., 2010. Genome-wide analysis reveals novel genes influencing temporal lobe structure with relevance to neurodegeneration in Alzheimer's disease. **Neuroimage** 51, 542-554.

[2] Hibar, D., et al., 2010. Voxelwise genome-wide association of Diffusion Tensor Images identifies putative novel variants influencing white matter integrity in 467 related young adults. **Society for Neuroscience**, San Diego, CA.

[3] Stein, J.L., et al., 2010. Voxelwise genome-wide association study (vGWAS). **Neuroimage** 53, 1160-1174.

[4] Neale, B.M., Sham, P.C., 2004. The future of association studies: gene-based analysis and replication. **Am J Hum Genet** 75, 353-362.
[5] Hua, X., et al., 2008. Tensor-based morphometry as a neuroimaging biomarker for Alzheimer's disease: An MRI study of 676 AD, MCI, and normal subjects. **Neuroimage** 43, 458-469.
[6] Cannon, T.D., Keller, M.C., 2006. Endophenotypes in the genetic analyses of mental disorders. **Annu Rev Clin Psychol** 2, 267-290.
[7] Saykin, A.J., et al., 2010. Alzheimer's Disease Neuroimaging Initiative biomarkers as quantitative phenotypes: Genetics core aims, progress, and plans. **Alzheimers Dement** 6, 265-273.
[8] Wang, K., Abbott, D., 2008. A principal components regression approach to multi-locus genetic association studies. **Genet Epidemiol** 32, 108-118.
[9] Reiman, E.M., et al., 2007. GAB2 alleles modify Alzheimer's risk in APOE epsilon4 carriers. **Neuron** 54, 713-720.
[10] Benjamini, Y., Hochberg, Y., 1995. Controlling the False Discovery Rate - a Practical and Powerful Approach to Multiple Testing. **Journal of the Royal Statistical Society Series B-Methodological** 57, 289-300.
[11] Storey, J.D., 2003. The positive false discovery rate: a Bayesian interpretation and the *q*-value. **Annals of Statistics** 31, 2013-2035.
[12] The ENIGMA Consortium, 2011. Genome-wide association meta-analysis of hippocampal volume: results from the ENIGMA consortium. OHBM, June 2011.

2.3 Voxelwise gene-wide association study (vGeneWAS)

This section is adapted from:

Derrek P. Hibar, Jason L. Stein, Omid Kohannim, Neda Jahanshad, Andrew J. Saykin, Li Shen, Sungeun Kim, Nathan Pankratz, Tatiana Foroud, Matthew J. Huentelman, Steven G. Potkin, Clifford R. Jack, Jr., Michael W. Weiner, Arthur W. Toga, Paul M. Thompson, and the Alzheimer's Disease Neuroimaging Initiative. Voxelwise gene-wide association study (vGeneWAS): multivariate gene-based association testing in 731 elderly subjects. Neuroimage. 2011. 56(4):1875-1891.

NeuroImage 56 (2011) 1875–1891

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/ynimg

Voxelwise gene-wide association study (vGeneWAS): Multivariate gene-based association testing in 731 elderly subjects

Derrek P. Hibar [a], Jason L. Stein [a], Omid Kohannim [a], Neda Jahanshad [a], Andrew J. Saykin [b,c], Li Shen [b], Sungeun Kim [b], Nathan Pankratz [c], Tatiana Foroud [c], Matthew J. Huentelman [d], Steven G. Potkin [e], Clifford R. Jack Jr. [f], Michael W. Weiner [g,h], Arthur W. Toga [a], Paul M. Thompson [a,*] and the Alzheimer's Disease Neuroimaging Initiative [1]

[a] Laboratory of Neuro Imaging, Department of Neurology, UCLA School of Medicine, Los Angeles, CA, USA
[b] Center for Neuroimaging, Department of Radiology and Imaging Science, Indiana University School of Medicine, Indianapolis, IN, USA
[c] Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, IN, USA
[d] The Translational Genomics Research Institute, Phoenix, AZ, USA
[e] Department of Psychiatry and Human Behavior, University of California, Irvine, Irvine, CA, USA
[f] Mayo Clinic, Rochester, MN, USA
[g] Departments of Radiology, Medicine, Psychiatry, UC San Francisco, San Francisco, CA, USA
[h] Department of Veterans Affairs Medical Center, San Francisco, CA, USA

ARTICLE INFO

Article history:
Received 20 October 2010
Revised 19 February 2011
Accepted 28 March 2011
Available online 8 April 2011

Keywords:
Principal components regression
Voxelwise
Multivariate
Gene-based
GWAS
GAB2

ABSTRACT

Imaging traits provide a powerful and biologically relevant substrate to examine the influence of genetics on the brain. Interest in genome-wide, brain-wide search for influential genetic variants is growing, but has mainly focused on univariate, SNP-based association tests. Moving to gene-based multivariate statistics, we can test the combined effect of multiple genetic variants in a single test statistic. Multivariate models can reduce the number of statistical tests in gene-wide or genome-wide scans and may discover gene effects undetectable with SNP-based methods. Here we present a gene-based method for associating the joint effect of single nucleotide polymorphisms (SNPs) in 18,044 genes across 31,662 voxels of the whole brain in 731 elderly subjects (mean age: 75.56 ± 6.82SD years; 430 males) from the Alzheimer's Disease Neuroimaging Initiative (ADNI). Structural MRI scans were analyzed using tensor-based morphometry (TBM) to compute 3D maps of regional brain volume differences compared to an average template image based on healthy elderly subjects. Using the voxel-level volume difference values as the phenotype, we selected the most significantly associated gene (out of 18,044) at each voxel across the brain. No genes identified were significant after correction for multiple comparisons, but several known candidates were re-identified, as were other genes highly relevant to brain function. *GAB2,* which has been previously associated with late-onset AD, was identified as the top gene in this study, suggesting the validity of the approach. This multivariate, gene-based voxelwise association study offers a novel framework to detect genetic influences on the brain.

Introduction

Recent efforts in imaging genetics have advanced the field rapidly from identifying heritable features of the brain to genome-wide searches for specific genetic variants that might account for functional and structural variations in large populations (Potkin et al., 2009a, 2009b; Shen et al., 2010; Stein et al., 2010b; Thompson et al., 2010). Variation in the human genome may account for variations in brain integrity, and multi-national consortia have been set up to discover and verify genetic effects on brain images (e.g., the ENIGMA project; http://enigma.loni.ucla.edu). In imaging genomics, the vast amount of information in the images (>100,000 voxels) and across the genome (>12 million known variants) requires powerful methods to relate genetic variants to the structure and function of the brain. Power issues arise due to the small effect sizes, and the huge numbers of statistical comparisons. Most techniques use some type of data reduction, limiting the number of genetic variants studied or the number of imaging features studied, or both. The ultimate goal of

* Corresponding author at: Laboratory of Neuro Imaging, Dept. of Neurology, UCLA School of Medicine, Neuroscience Research Building 225E, 635 Charles Young Drive, Los Angeles, CA 90095-1769, USA. Fax: +1 310 206 5518.
E-mail address: thompson@loni.ucla.edu (P.M. Thompson).

[1] Data used to prepare this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (http://www.loni.ucla.edu/ADNI). 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. For a complete listing of ADNI investigators, please see: http://www.loni.ucla.edu/ADNI/Collaboration/ADNI_Manuscript_Citations.pdf.

1053-8119/$ – see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.neuroimage.2011.03.077

these gene-hunting studies is to create a method that addresses the gene discovery problem in a statistically powerful and biologically meaningful way.

The current mainstay of gene-hunting efforts in imaging genetics is the genome-wide association study (*GWAS*). Most genetic association tests relate individual SNPs to phenotypes, but since there are on average between 20 and 100 SNPS per gene (in our dataset), and alleles at these SNPs are often highly correlated, a method that tests all the SNPs in a gene at once (or most of the variance contributed by SNPS in a gene) would reduce the number of tests required and be more powerful. We will hereafter refer to SNP-based approaches and gene-based approaches. These assess associations between common SNPs and features in an image. In typical GWAS studies, each genetic variant (usually a SNP) is independently tested for its association to the phenotype—a mass univariate method, where no data reduction is used across the genome. For example, Stein et al. (2010b) performed a genome-wide search of around 500,000 SNPs, and found a novel variant in the *GRIN2B* gene that is associated with temporal lobe volume. The gene *GRIN2B* encodes a glutamate receptor that is already the target of drugs (memantine) used to treat Alzheimer's disease (Parsons et al., 2007). Findings such as these are promising as they have biological relevance without relying on a prior hypothesis about any specific SNP. However, performing mass SNP-based tests on imaging summary measures (such as temporal lobe volume, hippocampal volume, etc.) or *ad hoc* regions of interest (ROI), collapses the brain measures into a single number. Studies using an ROI to define the imaging phenotype may miss fine-grained differences throughout the brain, across subjects. In addition, a predefined ROI can lead to false-negative results if a true association signal lies outside or only partially within a chosen ROI.

Several studies now perform genome-wide searches at each voxel across the brain (Hibar et al., 2010). This approach avoids preselecting an *ad hoc* region of interest in the brain and does not require prior hypotheses about which genetic variants, or which regions of interest, matter. Stein et al. (2010a) performed a genome-wide, brain-wide search, termed a voxelwise genome-wide association study (*vGWAS*), in 740 subjects from the ADNI. The experiment was extremely computationally intensive (27 h on 500 nodes), performing around 16 trillion tests of association. However, the correction for multiple comparisons was commensurate with the number of tests performed. None of the variants identified was significant after multiple comparison correction, but several variants were promising candidates for further analysis. In an alternative approach, Vounou et al. (2010) proposed a method that leverages the sparseness of signals to simultaneously select SNP variants and regions of association, reducing the number of SNPs and phenotypes tested. Future GWAS studies in imaging will likely reduce the number of tests and multiple comparisons using Bayesian priors. This can prioritize certain regions of the image or the genome, for later meta-analysis across multiple datasets.

Gene-based association methods complement single-marker GWAS for implicating underlying genetic variants in complex traits and diseases (Neale and Sham, 2004). Given recent advances in high-throughput genotyping, densely packed sets of SNPs, or genetic markers, can capture increasing amounts of variation throughout the genome. Methods that consider combinations of SNPs from the same gene should better describe genetic associations than methods that rely on data from SNPs independently (Neale and Sham, 2004; Schaid, 2004). Whole-gene testing is a biologically plausible approach to the problem, as the ultimate unit of biological activity is the gene (or its protein product; Potkin et al., 2009c). By associating the joint effect of multiple SNPs within a gene, in this study we aimed to show that gene-based approaches can be more powerful than traditional SNP-based approaches (with the relative power depending on how the genetic variants affect the phenotype). For example, if a gene contains multiple causal variants with small individual effects, SNP-based methods will miss these associations if a very stringent significance threshold is used (as in GWAS). In addition, if multiple loci within a gene combine to jointly affect a phenotype, this may also be missed by traditional GWAS. These two scenarios are highly likely, especially if we accept the "common disease, common variant" hypothesis (Reich and Lander, 2001), but they are not accounted for in methods that test each SNP, one at a time.

A multi-SNP, gene-based test can consider the combined effect of each variant within the gene, while accounting for linkage disequilibrium (LD) or correlation between markers. As such, at least in theory it may detect associations missed by traditional SNP-based GWAS. Related to this approach is "multi-locus fitting"—a developing field in quantitative genetics, for the analysis of complex traits. Some multi-locus analyses use statistical methods specialized for handling high-dimensional data, including regularized regression methods such as ridge regression (Malo et al., 2008; Sun et al., 2009), the Bayesian lasso (Zou, 2006; Wu et al., 2009), and neural network models (Lucek et al., 1998; Ott, 2001). Another related approach is set-based association testing, implemented in the software Plink (Purcell et al., 2007), which allows for the combination of univariate test statistics into a single univariate test statistic using permutations. Gene-based tests also reduce the effective number of statistical tests by aggregating multiple SNP effects into a single test statistic. However, for gene-based tests to be feasible, the multivariate test statistics need to be computationally efficient to implement. Here we assessed whether it would be feasible to extend to a neuroimaging database, a gene-based association method using principal components regression (PCReg) as proposed by Wang and Abbott (2008) for single-valued traits. We applied PCReg across all genes, to a large database of voxelwise imaging data. We call our method a voxelwise "gene-wide" association study (vGeneWAS). By performing association tests on whole genes, we greatly reduce the number of tests (from 437,607 SNPs down to 18,044 genes) while avoiding the problems associated with focusing on ROIs or summary measures. Our framework shows how to conduct vGeneWAS studies, and identify gene variants that warrant further study.

We hypothesized that vGeneWAS would, in some situations, have greater power to detect associations than existing SNP-based methods. One such situation might be when a gene contains many loci with weak individual effects. In addition, we expected that vGeneWAS would have greater overall power than mass SNP-based methods, like vGWAS, because of the drastic reduction in the effective number of statistical tests performed.

Materials and methods

Study design and subjects assessed

ADNI is a large 5-year study initiated in 2003 as a public–private partnership between the National Institute on Aging (NIA), the National Institute of Biomedical Imaging and Bioengineering (NIBIB), the Food and Drug Administration (FDA), private pharmaceutical companies, and non-profit organizations. The ADNI study aims to identify and investigate biological markers of Alzheimer's disease through a combination of neuroimaging, genetics, neuropsychological tests and other measures in order to develop new treatments, track disease progression, and lessen the time required for clinical trials. The study was conducted according to the Good Clinical Practice guidelines, the Declaration of Helsinki, and U.S. 21 CFR Part 50—Protection of Human Subjects, and Part 56—Institutional Review Boards. Written informed consent was obtained from all participants before protocol-specific procedures were performed.

The study recruited 202 Alzheimer's disease subjects (AD), 413 with mild cognitive impairment (MCI), and 237 normal elderly controls (NC) who were assessed every 6 or 12 months for 3 years. Subjects went through extensive clinical and cognitive tests at the

time of each scan to determine and track diagnosis. Further information on inclusion criteria and the study protocol may be found online (http://www.adni-info.org/). Baseline structural MRI scans and genetic data for 818 subjects were obtained on or before May 5, 2010, from the public ADNI database (http://www.loni.ucla.edu/ADNI/). Scans for 852 subjects were available, but we excluded 121 subjects based on quality control measures (poor registration and image quality) and to avoid a well documented problem in statistical genetics known as population stratification (McCarthy et al., 2008). When performing association tests on latent subpopulations of different ethnicities or relatedness, spurious associations may arise due to differences in allele frequencies between groups, instead of true association with the phenotype (Lander and Schork, 1994). Subjects were removed based on self-reported ethnicity, later verified by multi-dimensional scaling (MDS) analysis (see previous study: Stein et al., 2010b), leaving 172 AD patients (78 women/94 men; mean age ± standard deviation = 75.54 ± 7.62 years), 356 MCI subjects (130 women/226 men; mean age: 75.23 ± 7.22), and 203 healthy elderly controls (93 women/110 men; mean age: 76.15 ± 4.99). We did not split the subjects by diagnosis for this analysis in order to exploit the broadest phenotypic continuum (Petersen, 2000) and maximize statistical power to detect genetic associations (Cannon and Keller, 2006).

Imaging methods

Baseline MRI scans for each subject were analyzed using tensor-based morphometry (TBM) as described previously (Hua et al., 2008). Briefly, high-resolution T1-weighted structural brain MRI scans were acquired at 58 ADNI sites on 1.5 T scanners with a protocol developed for multiple site consistency (Jack et al., 2008; ADNI also collected some data at 3 T, which we did not analyze here; see Ho et al., 2010). Additional image corrections were applied to all images in a processing pipeline including: *GradWarp* correction of geometric distortion (Jovicich et al., 2006), B1-correction to adjust image intensity non-uniformity (Jack et al., 2008), N3 bias correction to adjust intensity inhomogeneity across a scan (Sled et al., 1998), and geometric scaling determined by a phantom scan acquired at each subject's scanning session to adjust for scanner and session-specific calibration errors (Jack et al., 2008). Images were linearly aligned using a 9 parameter algorithm to the International Consortium for Brain Imaging template (ICBM-53; Mazziotta et al., 2001) to align brain positions to a common standard space, adjusting for global scaling.

The TBM analysis was performed following the protocol of our prior study, which showed clinical and cognitive test scores correlated with temporal lobe volumes, in a subset of the ADNI population (Hua et al., 2008). A minimum deformation template (MDT) was created based on a random subset of the healthy elderly controls at baseline. The MDT provides an unbiased representation of MRI scans expected from a group of average healthy elderly persons. We generated maps of localized volume difference for each subject compared to the MDT using an inverse-consistent, symmetric, intensity-based nonlinear warping algorithm (Leow et al., 2005). Maps of localized volume differences (called Jacobian maps) are estimated using the Jacobian determinant of the deformation matrix, which itself is a voxel-level estimate of volume excess or deficit compared to the MDT. Jacobian maps for each individual were then down-sampled using trilinear interpolation to a $4\times4\times4$ mm^3 voxel size to reduce computational burden. The value at each voxel in the Jacobian map represents a percentage volume difference compared to the MDT; we used this voxel-based measure of volume difference as the phenotype for genetic association tests.

SNP filtering and gene grouping

Genome-wide genotype data were collected at 620,901 markers on the Human610-Quad BeadChip (Illumina, Inc., San Diego, CA). For details on how genetic data were processed, please see Saykin et al. (2010) and Stein et al. (2010a). Different types of markers were genotyped (including copy number probes), but only SNPs were used in this analysis. Several SNPs were excluded from the analysis based on standard filtering criteria, measures used in many other GWAS studies (Wellcome Trust Case Control Consortium, 2007): call rate <95% (42,670 SNPs removed), significant deviation from Hardy–Weinberg equilibrium $P<5.7\times10^{-7}$ (871 markers removed), autosomal chromosomes only (10,686 SNPs removed), and an Illumina GenCall quality control score of <0.15 to eliminate "no call" genotypes (variable number of missing genotypes across subjects). We chose to remove SNPs with a minor allele frequency <0.10 (161,354 SNPs removed) based on our sample size. With our sample of 731 images, we are underpowered to detect associations with SNPs where the minor allele frequency is less than 10%, unless effect sizes are large (Wang et al., 2005; Flint et al., 2010). In addition, excluding SNPs with low minor allele frequencies avoids the risk of finding significant associations where only a small subset of subjects have the rare allele type and do not represent an accurate sampling of the phenotype of interest. If a very low minor allele frequency cutoff is used (e.g., MAF<0.01) in samples of fewer than a thousand subjects, this may result in cases where an association is driven by a single subject. Clearly, such a result may be unreliable and is unlikely to replicate, so the higher MAF cut-off guards against this.

Due to the filtering based on Illumina GenCall quality control measures, individual subjects have some residual missing genotypes at random SNPs throughout the dataset. Because PCReg requires data without missing genotypes and to maximize the number of subjects included in the analysis, we performed imputation using the software, Mach (version 1.0), to infer the haplotype phase and automatically impute the missing genotype data (Li et al., 2009). After all rounds of quality control and preparation, 437,607 SNPs remained.

Using the retrieval interface of the software package PLINK (version 1.05; http://pngu.mgh.harvard.edu/~purcell/plink/), SNP annotations were made by continuously soliciting the TAMAL database (Hemminger et al., 2006) based chiefly on UCSC genome browser files (Hinrichs et al., 2006), HapMap (Altshuler et al., 2005), and dbSNP (Wheeler et al., 2008). The newly annotated SNPs were grouped by gene, where "gene" is defined by the gene transcript region including both introns and exons. We chose not to include SNPs upstream/downstream from the gene region. This may miss SNPs in promoter or regulatory regions for a gene, but avoids choosing an arbitrary window that may select regulatory SNPs for some genes, but not for other genes whose regulatory regions lie beyond the window length. SNPs that were not located in a gene were excluded (224,057 SNPs removed). All splice variants were considered as belonging to the same gene. After applying SNP filtering criteria, SNP annotation, and gene grouping, 18,044 genes were left for analysis out of the estimated 20,000–25,000 protein coding genes in the human genome (International Human Genome Sequencing Consortium, 2004).

Gene-based association statistics

Independent tests of statistical association with imaging measures were performed for each gene at 31,622 voxels within a whole-brain mask of the MDT across 731 subjects. To test the joint effect of *all* SNPs in a gene on the volume difference at each voxel, we employed a multiple partial-*F* test. A multiple partial-*F* test works by first estimating the fit of a "reduced model" of any number of nuisance variables on a given dependent variable and then estimating the fit of a second "full model" with the nuisance variables and any number of independent variables on the same dependent variable. Each association test results in an *F* statistic, which represents the joint effect of the independent variables on the dependent variable, controlling for the effects of nuisance variables already in the

model. The multiple partial-*F* statistic was calculated for each gene at each voxel using Eq. (1) below. *k* is df(full) − df(reduced) and RSS is the residual sum of squares:

$$F_{k,df(\text{full})} = \frac{RSS(\text{reduced}) - RSS(\text{full})}{df(\text{reduced}) - df(\text{full})} \Big/ \frac{RSS(\text{full})}{df(\text{full})} \quad (1)$$

Multiple partial-*F* tests are well suited for testing the effects of multiple predictors on a given phenotype, but genetic data sometimes complicate testing because SNPs in the same gene are often correlated due to high LD. When the SNP values in a cohort of subjects are treated as a vector (whose components are the SNP value in each subject coded in an additive manner: 0, 1, or 2), then the adjacent SNPs can make different subjects' vectors collinear. The dependence among these almost collinear SNP vectors in the multiple partial-*F* test model can lead to improper signs of beta coefficient estimates, wildly inaccurate magnitudes of beta coefficients, large standard error estimates (Kleinbaum, 2007), and false inferences. The reason this occurs is that standard regression models require the inversion of a set of "normal" equations, and when predictors (here SNP vectors) are highly correlated, the equations are not of full rank. This leads to unstable or unreliable solutions. One way out of this predicament is to use a type of regularized or penalized regression, such as ridge regression (also known as sparse regression or Tikhonov regularization), which can be used when there are high correlations among the predictors. Alternatively, dimension reduction may be performed (which we do here), to create a set of predictors that explain the variance in the data but that are no longer correlated.

To avoid the complications of collinearity in the statistical model, we first performed principal component analysis (PCA) on the SNPs within each gene, storing all of the orthonormal basis vectors of the SNP matrix that explained the first 95% of the variance in the set of SNPs. Basis vectors with the highest eigenvalues (higher proportions of explained variance) were included until 95% of the SNP variance was explained, and the rest were discarded. These new "eigenSNPs" approximate the information in the observed SNPs, but lack the collinearity that disrupts the multiple partial-*F* test models. By first performing PCA followed by a multiple partial-*F* test, our method may be considered a variant of PCReg and produces *F* statistics equivalent to those proposed in Wang and Abbott (2008). In this study, the independent variables built into the multiple partial-*F* test full model were the column vector output from PCA performed on each gene with age and sex as covariates. In this way, we tested the joint predictive effect of variation throughout a gene on brain volume variations on a voxel-by-voxel level.

The total number of tests of association for vGeneWAS is very high (18,044 genes × 31,662 voxels). Because of the massive processing requirement, we coded the PCA and multiple partial-*F* test steps of PCReg using the R statistical package (version 2.9.2; http://www.cran.r-project.org/) using the doMC "multi-core" package (version 1.2.1; http://www.revolutionanalytics.com/) to split processing over multiple cores in a single CPU. Processing was parallelized over a cluster of 10 high performance 8-core CPU nodes using the Laboratory of Neuro Imaging (LONI) Pipeline (http://pipeline.loni.ucla.edu/). For further data reduction, we only saved data on the gene with the lowest *P*-value at each voxel. This is comparable to our prior work using voxelwise testing of all 500,000 genotyped SNPs, where only the SNP with the lowest *P*-value was retained at each voxel (Stein et al., 2010a). The total time required to complete an analysis was approximately 13 days.

Comparison of SNP-based and gene-based methods

To examine the situations where PCReg exhibits better (or worse) performance than traditional simple linear regression, we compared the two methods directly on real genetic data. Performing tests on real genetic data as opposed to simulated data is important because the power of each method depends upon the underlying LD structure. Generating simulated data that mimics a chosen LD structure can be just as biased as selecting actual genes, though a significant treatment of the issue of power in PCReg is discussed in Wang and Abbott (2008). We used temporal lobe volume (TLV) summary measures obtained by Stein et al. (2010b), as the phenotype for testing associations for both methods. We performed a genome-wide scan of every SNP from our filtered and annotated genotype data (only including SNPs located within genes) using simple linear regression with SNPs coded following an additive model. We took the top SNP from the analysis, and the rest of the SNPs from the same gene, and performed PCReg on all of the SNPs in that gene. In addition, we performed a gene-wide scan of all of the genes in our dataset using PCReg with SNPs coded following the additive model. We selected the top gene from the analysis and then ran individual tests of association using simple linear regression at each SNP within the top gene. In this way, we were able to compare the performance of each method in cases where the underlying genetic structure might favor one method over the other.

Statistical thresholds and correction for multiple comparisons

As we noted in Stein et al. (2010a), the minimum *P*-value at each voxel, in the null case with *n* independent tests, approximately follows a probability density function (PDF) such that (Ewens and Grant, 2001):

$$f_{min}(x) = n(1-x)^{n-1} \quad (2)$$

The PDF derived from Eq. (2) is a Beta distribution with parameters $\alpha = 1$ and $\beta = n$. At each voxel, selecting the minimum *P*-value for the top gene then follows a Beta(1, *n*) distribution, where *n* is the independent number of genes tested.

However, it is well known that the adjacent SNP values within genes are not statistically independent (Frazer et al., 2007). Genetic loci are inherited in contiguous segments, and some genes co-segregate in blocks. The allele frequencies and structure of genes that co-segregate are more similar than would be expected by chance if they were assumed to be independent. Because of this, the effective number of independent tests (M_{eff}) is less than the total number of tests performed (M). By determining M_{eff}, we obtain a more accurate estimate of the total number of independent tests performed with vGeneWAS, given the LD structure of our genotype data.

In our sample, we estimated M_{eff} by performing permutation tests at three randomly selected, uncorrelated voxels in the brain. We regressed each of the 18,044 genes on the permuted residuals of the reduced model after including the age and sex covariates at each run, and stored the minimum *P*-value. Note that, in this case, the phenotype data is null. However, because it is computed from the real data after adjusting for age and sex, the phenotype data (image values) have the same range and statistical distribution as the data tested for genetic associations. By using the genes, one at a time, as regressors on this null data, one can develop a distribution of the resulting *P*-values, under null conditions, that can be used to calibrate the significance values that are ascribed to the observed data. As only the minimum *P*-value is retained (for the best fitting gene), one can build up a reference distribution for the minimum *P*-values, to help gauge the level of surprise in seeing associations in true data. We repeated this process 5000 times at each of the three randomly chosen voxels and merged the data. The distributions of null minimum *P*-values from each voxel were nearly identical (Fig. 1). Storing the minimum *P*-values of the permutation tests yields the expected null Beta distribution given our data. We used a maximum-likelihood function (betafit; Matlab, The Math Works, Inc.) that

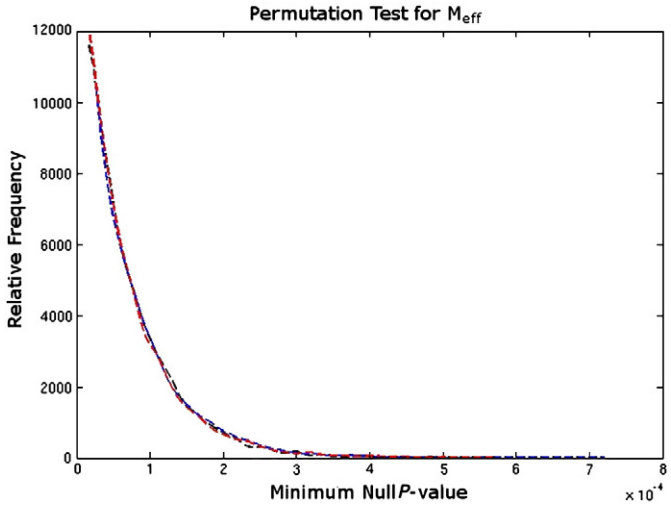


Fig. 1. A histogram shows the minimum null *P*-values obtained from permutation tests. Data from 3 different voxels are shown on the same graph (blue, red, and black lines); they are obtained from 3 randomly chosen, uncorrelated voxels in the brain (5000 permutations each). The distributions are nearly identical, and agree with each other, as well as accurately reflecting the effective number of independent tests (M_{eff}).

estimates the best fit for the null Beta distribution by varying the β parameter of Beta(1, β). The value of β approximates the effective number of independent tests (M_{eff}) performed on our data. We then apply an inverse Beta transform using the approximated β parameter so that the distribution of *P*-values is now a uniform distribution that deviates from the null when there is a signal.

After correcting for the effective number of independent gene-based tests performed at each voxel, we still need to correct for the multiple comparisons across voxels. We used the original false discovery rate method (FDR; Benjamini and Hochberg, 1995), which identifies whether there is any statistical thresholding of the uncorrected *P*-value maps that keeps the rate of false positive results within a predefined threshold (we chose 5%, which is conventionally used). This means that, if the results pass the FDR test, approximately 95% of the voxels declared as significant associations will be true positives (averaged over many experiments). In addition, we tested a less conservative variant of FDR, the positive FDR (pFDR), which operates under the condition that at least one true positive finding exists in the data (i.e., one of the null hypotheses is rejected) and yields *q*-value correction thresholds similar to the original FDR method (Storey, 2003). The pFDR test is implemented in the R statistical package called "qvalue" (Version 1.22.0).

Estimation of expected values in simulated maps

A certain amount of spatial smoothness is expected among voxels in an image. This is most likely explained by the non-independence of volume difference measures at adjacent voxels. Relative volume maps were generated using tensor-based morphometry (TBM), which relies on non-linear registration of each subject's imaging data to a common template. The degree of spatial smoothness in the Jacobian maps derived from the gradient of a deformation field depends on the choice of the regularizer used by the warping algorithm (Laplacian, elastic, fluid, sKL, etc.) and on the resolution of numerical grid chosen to solve the differential equations whose solution is the deformation field. Volume difference maps based on the deformation field vectors are spatially smooth, as are any resulting statistical maps. In addition to image smoothness, certain noncontiguous voxels in distant regions of the brain can have surprising covariance patterns despite their spatial separation (Fillard et al., 2007).

We examined whether the size of voxel clusters associated with the same gene from our vGeneWAS analysis differed from the cluster sizes expected under the null hypothesis of no association at all, given the non-independence of signals at adjacent voxels in our images. In addition, we wanted to determine whether the number of unique, top genes from across the brain significantly differed from the number of top genes expected by chance. We generated simulated cluster maps by first creating a correlation matrix of *r* values representing the Pearson's correlation between any given voxel and all other voxels in the brain. Next, we randomly selected (without replacement) a voxel, V_s, and all corresponding voxels with an r^2 value (proportion of variance explained) greater than 0.8 from the correlation matrix. We chose to only select voxels from the correlation matrix with an r^2 greater than 0.8 because this provided the largest cluster size estimates in the simulated output maps. The r^2 value of each voxel-to-voxel relationship was then used to divide the interval [0,1] of a uniform distribution such that the correlation between a voxel and V_s was directly related to the area under the curve occupied by that voxel's area under the curve or "bucket." The size of each voxel's bucket was recalculated each time we chose a new V_s. We selected a random number on the interval [0,1], and, depending on its value (which "bucket" it fell in), assigned the same categorical variable link (e.g., a, b, and c) to V_s and the voxel whose bucket was selected from the uniform distribution. This linked the two voxels. The probability of a given voxel being chosen from the uniform distribution was directly related to how correlated it was to V_s. We continued the process by randomly choosing a new V_s from the correlation matrix. If a randomly selected voxel V_s did not contain a linking variable, but selected a voxel from the uniform distribution that already contained a link, then V_s was assigned the linking variable of the voxel selected from the uniform distribution. If V_s already contained a linking variable and the voxel chosen from the uniform distribution had not previously been assigned a variable, the two voxels were linked using the existing linking variable of V_s. If both V_s and the voxel selected from the uniform distribution already contained a linking variable, we kept each variable as-is and then continued with the process. Finally, voxels that were not correlated to any other voxels in the image, with an r^2 value greater than 0.8, were assigned a non-linking random variable. After iterating through every voxel in the image, each voxel had a categorical variable that either linked it to other voxels or only to itself. We ran this entire simulation process 100 times, generating a new simulated cluster map each time. By considering the correlation of a given voxel to all other voxels in the image, as opposed to using a single summary measure of smoothness throughout an image, we were able to model the expected clustering among adjacent voxels and non-independent, spatially separated clusters.

Based on the 3D pattern of voxels and the variables linking voxels together, we used a nearest neighbor algorithm to measure cluster sizes of adjacent voxels with the same linking variable value. Using the cluster size estimates from each simulated map, we were able to determine the expected distribution of cluster sizes based directly on our study dataset. In addition, we used the total number of unique linking variables in each simulation as an estimate of the number of independent voxels in our dataset. Because non-independent, correlated voxels may tend to be associated with the same gene, we can use the total number of independent voxels to estimate the number of top associated genes we would expect to find in null cluster maps made from our actual test data. We used the estimated number of independent voxels, V_i, to randomly select (with replacement) a gene from the set of 18,044 genes and repeated the selection V_i times. We found the number of unique genes represented for each simulated output map and then took the average.

Results

Comparison of methods

To examine the differences between gene-based and SNP-based association methods (which are more standard), we compared the results of PCReg to linear regression using temporal lobe volume (TLV) data from a previous study (Stein et al., 2010b) as the phenotype. We chose to focus on the top gene or SNP identified by each method in order to examine performance when the variant chosen is deliberately selected to favor one of the two methods. *GRIN2B* was identified as the gene with the SNP variant that was most significantly associated with TLV ($P = 4.03 \times 10^{-7}$). We identified each of the 129 SNPs within the *GRIN2B* gene, and then performed linear regression at each SNP and PCReg as a single gene test with TLV as the phenotype. The $-\log_{10}$(*P*-value) of each SNP-based test is shown in Fig. 2a with single gene test results overlaid (black dotted line). It is clear that the main effect detected with linear regression is much greater in this case. It is important to note that we tested each of the 129 SNPs within the *GRIN2B* gene, which would require any significant *P*-values identified to be corrected for multiple comparisons before further study. In this example, however, there are several SNPs that beat the Bonferroni-corrected significance threshold ($\alpha = 3.9 \times 10^{-4}$). In comparison, the gene-based test of *GRIN2B* using PCReg was a single test not requiring correction for multiple comparisons and maintained a nominal significance value ($P = 0.012$). Also, we compared *BEST3* – the gene identified to be most significantly associated with TLV via PCReg – with the linear regression output of each SNP within the gene (Fig. 2b). The significance of the main effect of the gene-based test is much stronger ($P = 2.9 \times 10^{-4}$) than the best linear regression result ($P = 0.063$). This demonstrates a case where variance components from individual markers that are not significant via linear regression may be combined into a single significant test statistic.

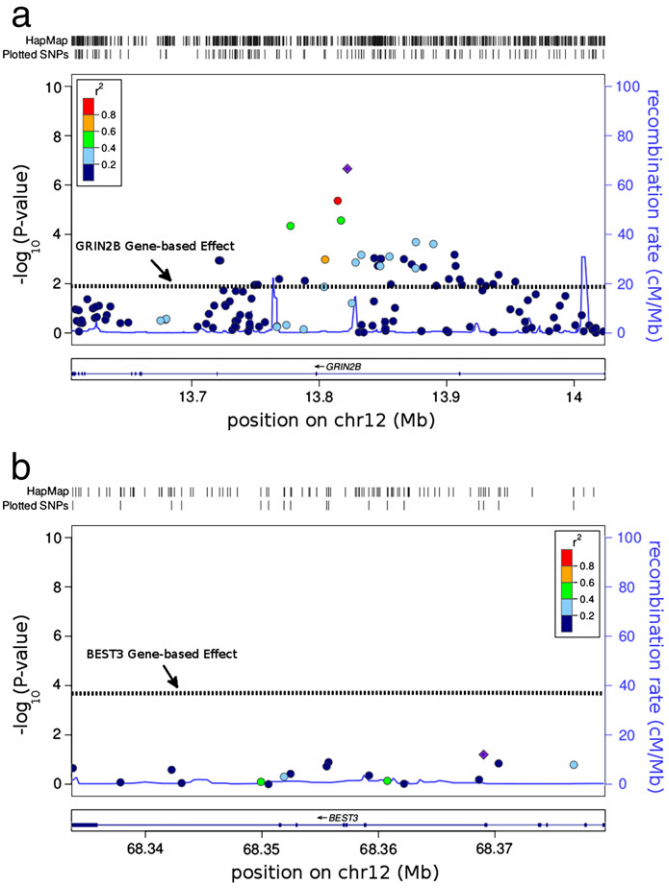


Fig. 2. Genetic association plots for univariate linear regression versus multi-locus PCReg. The $-\log_{10}$(*P*-value) of each SNP in *GRIN2B* (a) and *BEST3* (b) is plotted against its position in the gene. Each of the points is color coded by level of LD (compared to the top SNP, *the purple diamond dot*) as measured by r^2. The $-\log_{10}$(*P*-value) of the gene-based PCReg test for each gene is overlaid on the plot for comparison (*dotted black line*). Plots were generated using the *LocusZoom* software (http://csg.sph.umich.edu/locuszoom/).

Relation of gene significance to number of SNPs

Wang and Abbott examined whether the power to detect associations in genetic data is influenced by the number of eigenSNPs included in a PCReg model. They found that models with greater numbers of eigenSNPs do not have increased power to detect associations (Wang and Abbott, 2008). However, each additional eigenSNP included in the model uses a degree of freedom. It is therefore possible that PCReg and similar regression methods are biased toward selecting effects of smaller versus larger genes (Chapman and Whittaker, 2008). We examined our results for gene-size bias and verified that the number of eigenSNPs in the PCReg model within the top genes from our run of vGeneWAS was not correlated with the observed *P*-value, using a Pearson's product-moment correlation test ($r = 0.0045$; $P = 0.42$). In addition, we verified that the number of eigenSNPs in each of the 18,044 genes at a single voxel was not correlated with its significance level ($r = 0.0066$; $P = 0.29$). We also compared the number of eigenSNPs in each gene (mean and median: 14.3 and 9) with the number of eigenSNPs in the top genes from our analysis (mean and median: 13.6 and 5). It remains possible that we missed effects of very large genes, but this is inevitable in small samples as the number of eigenSNPs needed to adequately encode the majority of variation in large genes tends to approach the sample sizes, reducing the available numbers of degrees of freedom for the whole-gene tests.

Voxelwise GeneWAS

We generated maps of significance where each color-coded voxel in the brain shows the *P*-value of the most highly associated gene at that voxel (Fig. 3). There are several spatially contiguous regions throughout the brain with raw minimum *P*-values lower than 10^{-7}. In addition, some of the top genes identified show symmetric clustering across hemispheres. Brain structures are highly symmetric between hemispheres, at least for most brain regions, so symmetric genetic associations may be biologically plausible because the volumes of symmetric structures co-vary across subjects, so they may share similar genetic determinants. However, evidence of symmetric patterns of association in the brain does not necessarily imply biological plausibility (Fillard et al., 2007).

We used a simulation-based test to build the expected null distribution of cluster sizes given our image data. We compared the distribution of the cluster size values in simulated maps to the cluster sizes obtained from vGeneWAS. The proportion of the null (simulated) maps that contained small clusters is much greater than in vGeneWAS, while the proportion of the vGeneWAS map that contained large clusters was greater than in the null maps (Fig. 4). The minimum and maximum cluster sizes for the simulated maps were 1 and 14 voxels (64 and 896 mm^3), respectively. The minimum and maximum cluster sizes for vGeneWAS were 1 and 429 voxels (64 and 27,456 mm^3). This demonstrates that a large proportion of clusters of voxels associated with the same top gene are larger than would be expected based on completely null data, even taking into account the non-independence of voxels in our dataset.

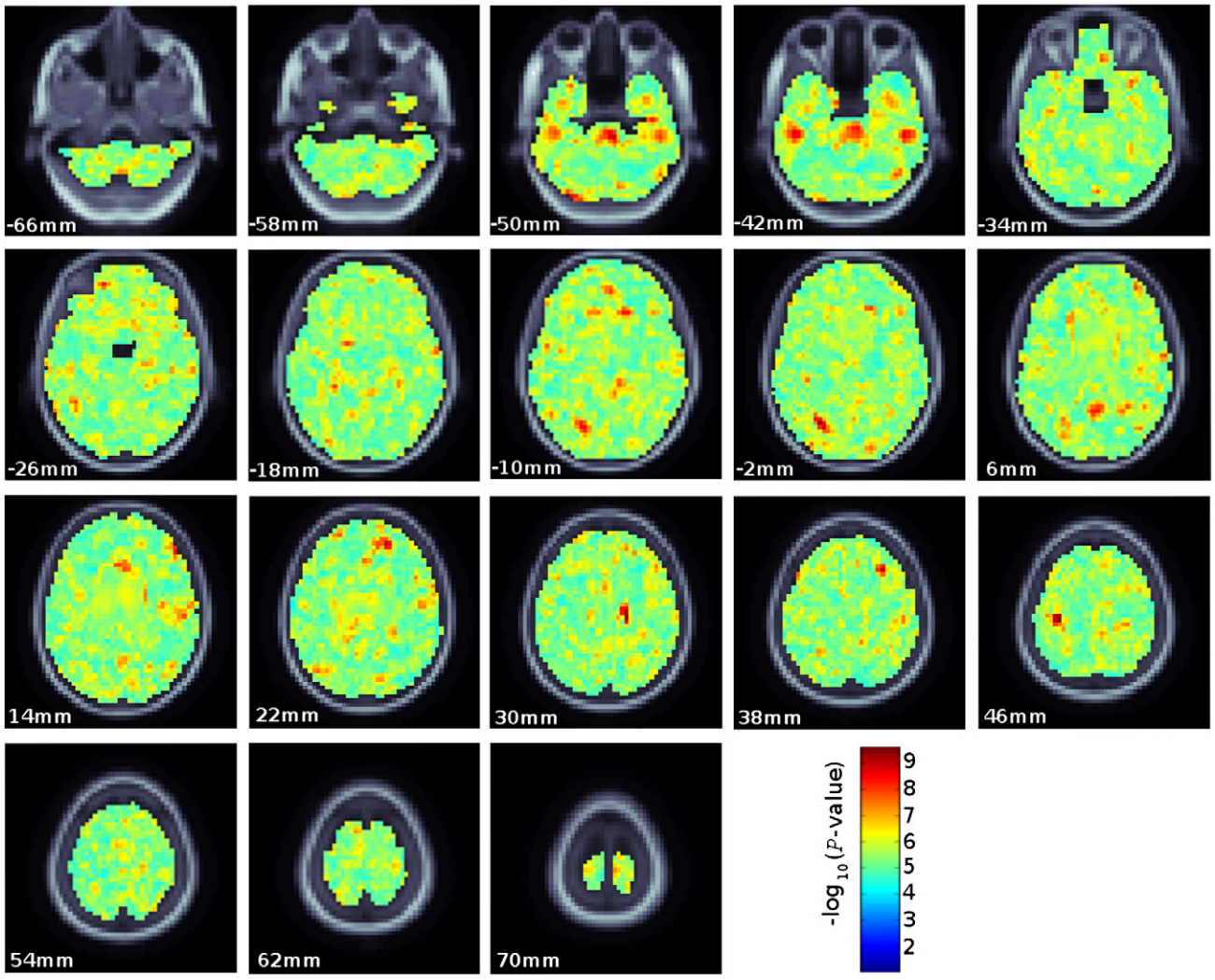


Fig. 3. A color-coded significance map of the top gene at each voxel. Sections are shown at 8 mm intervals throughout the brain. The top of each panel represents the anterior of the brain and bottom the posterior of the brain. The images are in radiological convention (the left side of the image is the patient's right hemisphere). Color coding is based on the $-\log_{10}(P\text{-value})$; warmer colors represent more significant associations.

Based on our simulated cluster maps, we used the number of unique clusters as an estimate of the number of independent voxels. The estimate of the number of independent voxels based on 100 runs of the simulation tests was $11{,}900.8 \pm 50.6$ (mean $\pm$ standard deviation) out of the 31,662 total voxels. We performed tests to estimate the number of genes we should expect to find in our analysis based on the non-independence of voxels in our data. We used the number of independent voxels estimated from the simulation to randomly select (with replacement) from our list of 18,044 genes. We tallied the number of unique genes represented for each simulated cluster map and found the average was 8721.4 ± 44.9 (mean $\pm$ standard deviation). We measured the total number of unique genes to be 5333 from our run of voxelwise GeneWAS. The number of observed genes is significantly lower than the number of genes expected based on the null cluster maps ($P<0.01$). Combined with our cluster size comparisons, this suggests that the top genes identified in our analysis tend to have a much more broadly distributed effect than expected based on null data, even taking into account the intrinsic spatial non-independence of our data. The top 20 genes most significantly associated with any voxel are listed in Table 1.

The GRB-associated binding protein 2 gene, *GAB2*, is the most significantly associated gene in our analysis and has previously been linked to late-onset Alzheimer's disease (LOAD) (Reiman et al., 2007). Reiman et al. (2007) identified 10 SNPs from the *GAB2* gene that were significantly associated with LOAD and *APOE* allele status in 1411 cases and controls from 20 NIA-sponsored Alzheimer's Disease Centers. Replication attempts in independent samples have yielded mixed results (Ramirez-Lorca et al., 2009; Lin et al., 2010; Chapuis et al., 2008), but large meta-analyses of several databases shows that *GAB2* may indeed have a moderate effect on the development of LOAD (Ikram et al., 2009; Schjeide et al., 2009). Specifically, the meta-analysis of the marker rs2373115 in nine studies has an odds ratio of 0.85 and a 95% confidence interval for the odds ratio of [0.76, 0.94]. In addition, the AlzGene website lists *GAB2* as being in the top 20 genes likely related to Alzheimer's disease (September 3, 2010; http://www.alzgene.org/). *In vivo* testing shows that *GAB2* is over-expressed in certain brain regions such as the hippocampus and posterior cingulate cortex in patients with LOAD (Reiman et al., 2007). Experiments with small-interfering RNA (siRNA) and *GAB2* reveal that the normal function of *GAB2* proteins prevents the formation of serine-262 phosphorylated tau tangles (Reiman et al., 2007). No studies, to our knowledge, have considered morphometric effects of *GAB2* variants. The *GAB2* associations show a symmetric signal in the white matter superior to the lateral ventricles (Fig. 5).

The second most highly associated gene, leucine-rich repeat and death domain containing protein (*LRDD*), is expressed in the brain

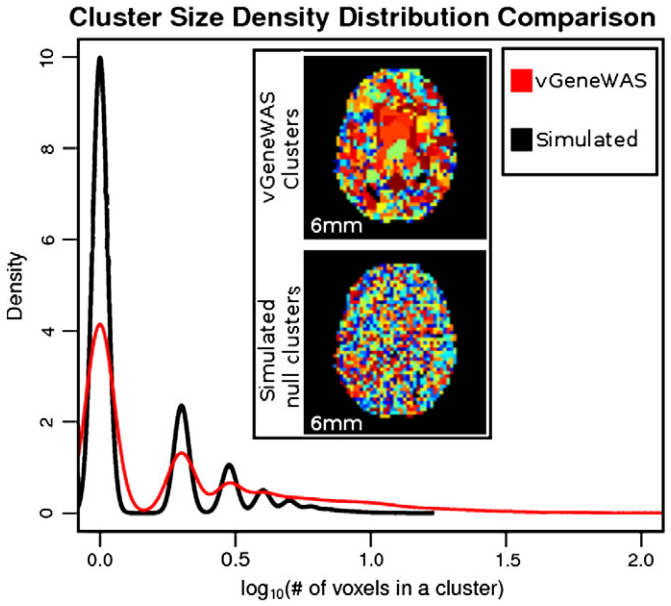


Fig. 4. Cluster sizes in vGeneWAS (*red line*) are compared with an average of simulated null maps (*black line*). We took the $\log_{10}$ of the number of voxels in a cluster (not in mm^3) across the brain in both maps for scaling purposes and for ease of comparison. The $\log_{10}$ cluster sizes are then plotted using a density function such that the total area under each line is equal to 1. The average simulated null map contains a larger proportion of small cluster sizes than vGeneWAS (higher peaks in the black line at values close to the origin on the *x*-axis). The vGeneWAS map contains a larger proportion of large cluster sizes than the average simulated null map (the red line is higher at larger values and is more extended). A single slice view of the vGeneWAS and average simulated null cluster maps are pictured for comparison (*inset*). Every unique cluster is assigned its own color. There are more unique clusters than distinct colors making visual inspection difficult, but in general the clusters in the vGeneWAS maps are larger.

and may mediate cell apoptosis and DNA repair (Telliez et al., 2000). In addition, *LRDD* has been implicated in the p53 tumor-suppression pathway likely by signaling cell apoptosis in response to DNA damage (Brown et al., 2009). *LRDD* was the most significantly associated gene in a cluster of voxels in a white matter tract of the occipital lobe, possibly the optic radiations (Fig. 5).

Associations with protein tyrosine phosphatase receptor type beta, *PTPRB*, are detected in the cerebellum (Fig. 5). *PTPRB* interacts with neural receptors and cell-adhesion molecules and is involved in neurite development and neuronal differentiation (Ishiguro et al., 2008). *PTPRB* has also previously been associated with alcohol and drug abuse via genome-wide search (Ishiguro et al., 2008). In addition, an expression study found that *PTPRB* encoded proteins are present in the gastric mucus and other tissues of gastric cancer patients (Wu et al., 2006).

The fourth and fifth most significantly associated genes are zinc finger protein 462 (*ZNF462*) and immunoglobin superfamily member 5 (*IGSF5*), respectively. *ZNF462* is the most significantly associated gene in a cluster of voxels in the upper-left gray matter of the parietal lobe (Fig. 5). Interestingly, *IGSF5* shows symmetrical clusters of association in the temporal lobe and the surrounding cerebrospinal fluid (CSF) at the base of the brain (Fig. 5). Neither gene is well studied, but *IGSF5* may be involved with junction cell adhesion (Hirabayashi et al., 2003).

Other genes of interest identified in our analysis include *ARALAR*, which encodes a calcium-binding mitochondrial protein that is highly expressed in the brain (del Arco and Satrustegui, 1998). *ARALAR* has previously been associated with autism (Ramoz et al., 2004), but the claims are controversial (Rabionet et al., 2004). *CHRM5*, is a muscarinic acetylcholine receptor M5 coding gene and has previously been associated with schizophrenia (De Luca et al., 2004). *S100B*, encodes a zinc-binding protein over-expressed in patients with Alzheimer's disease and interacts with Tau proteins (Yu and Fraser, 2001).

Correction for multiple comparisons

After permutation testing to determine the effective number of independent gene tests, we need to model the function parameters so that we can transform the data for correction for multiple comparisons. The effective number of independent tests was estimated to be 15,636, which is a moderate reduction from the 18,044 genes measured directly in this experiment. We therefore chose to model the null distribution as Beta(1, 15,636). The probability density function (PDF) of Beta(1, 15,636) on the normalized histogram of observed *P*-values fits the data well with only small deviations from the original Beta(1, 18,044) (Fig. 6a). We note, however, that our

Table 1
The top 20 genes most significantly associated at any voxel, organized by minimum observed *P*-value. The common gene name is listed, with the number of SNPs within that gene (Note: the number of SNPs in a gene will vary depending on the genotyping and quality control methods used). We also list the number of eigenSNPs included in the PCReg models. The mean observed *P*-value is also listed; it is the average of the *P*-values at all voxels where that gene was most significantly associated. Also listed is the volume of voxels (in mm^3) and the percentage of the total volume of voxels in the brain where the gene was the most significantly associated gene. The maximum cluster size observed for each gene is listed as well as the total number of spatially independent clusters.

Chr	Gene	# of SNPs in gene	# of eigenSNPs	Minimum P-value	Mean P-value	Volume (mm^3)	Proportion of brain volume	$Cluster_{max}$ (mm^3)	# of clusters
11	GAB2	20	10	2.36×10^{-9}	1.50×10^{-5}	6336	0.0049	2688	9
11	LRDD	2	2	2.60×10^{-9}	1.32×10^{-5}	8128	0.0063	7872	4
12	PTPRB	17	13	2.84×10^{-9}	1.81×10^{-5}	3200	0.0024	3008	5
9	ZNF462	9	6	3.29×10^{-9}	1.84×10^{-5}	2688	0.0021	2688	1
21	IGSF5	27	14	5.32×10^{-9}	1.62×10^{-5}	16,384	0.013	9344	3
2	SLC25A12	10	5	9.48×10^{-9}	2.66×10^{-5}	1856	0.0014	1792	2
11	MRE11A	9	6	9.86×10^{-9}	8.80×10^{-6}	9344	0.0072	9216	3
19	SLC8A2	11	7	1.06×10^{-8}	3.18×10^{-5}	5632	0.0043	5376	3
15	CHRM5	3	3	1.71×10^{-8}	1.77×10^{-5}	1280	0.00099	1216	2
18	SPIRE1	19	14	2.94×10^{-8}	2.88×10^{-5}	6016	0.0046	3072	12
3	C3orf64	9	8	3.71×10^{-8}	2.43×10^{-5}	4352	0.0034	2112	4
21	S100B	1	1	4.75×10^{-8}	2.81×10^{-5}	9344	0.0072	6656	7
1	CRCT1	1	1	5.54×10^{-8}	2.90×10^{-5}	4096	0.0032	3456	4
19	ZNF626	6	5	5.85×10^{-8}	2.12×10^{-5}	2560	0.0020	2112	3
1	ELK4	1	1	6.05×10^{-8}	3.27×10^{-5}	4032	0.0031	2688	4
11	RSF1	8	6	9.30×10^{-8}	1.31×10^{-5}	768	0.00059	768	1
20	WFDC11	2	2	1.06×10^{-7}	2.49×10^{-5}	1280	0.00099	512	5
6	SCML4	27	18	1.07×10^{-7}	1.67×10^{-5}	3328	0.0026	3328	1
12	ERP27	8	14	1.08×10^{-7}	2.61×10^{-5}	2624	0.0020	2176	2

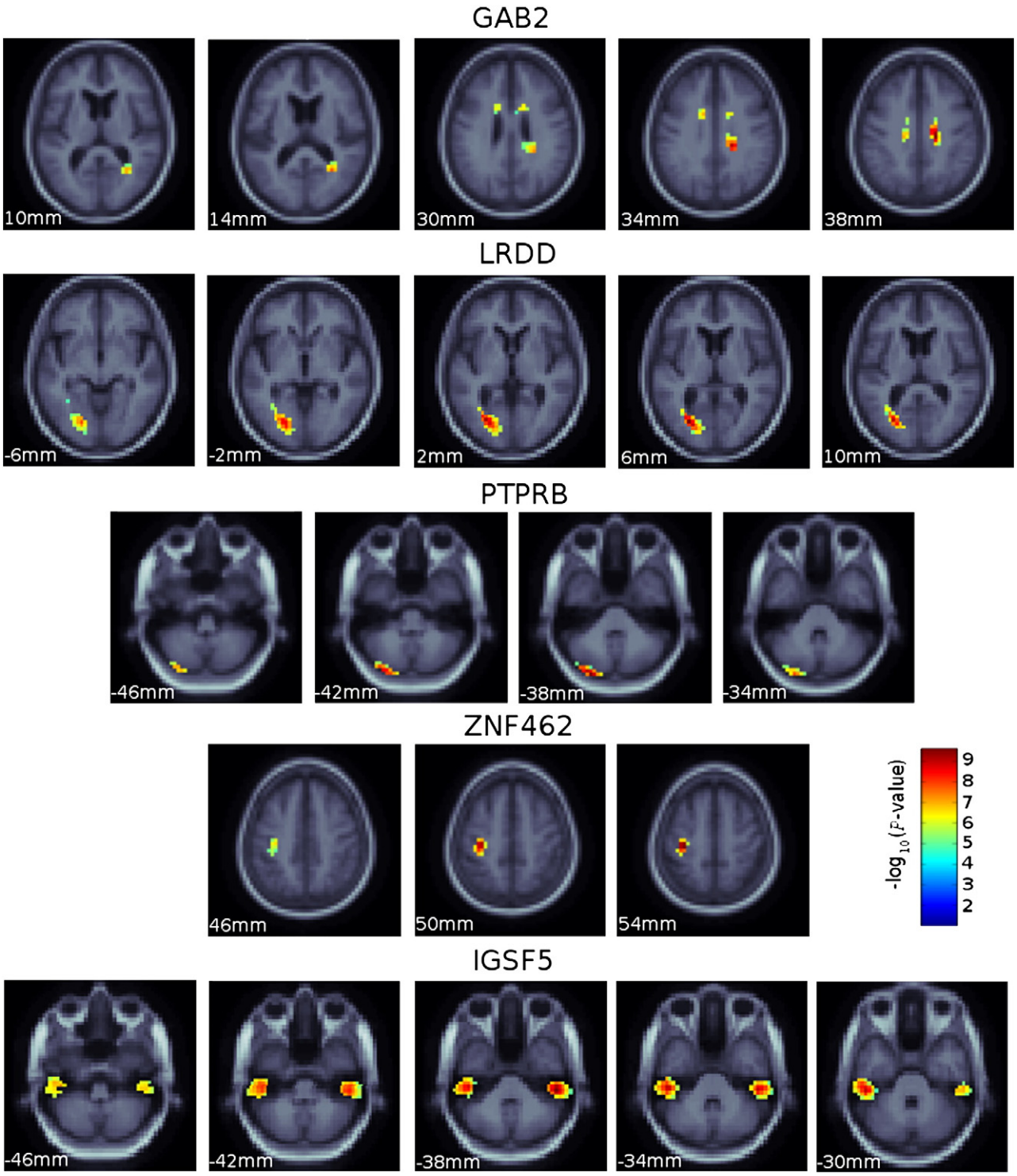


Fig. 5. Regions in the brain associated with the top 5 genes from our vGeneWAS analysis (where the uncorrected *P*-value at a given voxel is overlaid on the minimum deformation template). The slices chosen best represent the regions where each gene was the most significantly associated gene in the brain. Images read from inferior to superior (left to right of the page) following radiological convention and with the top and bottom of each panel representing the anterior and posterior of the brain, respectively.

estimate is determined both by SNP density and degree of coverage in the SNP marking scheme of our study. Experiments that use different SNP-chips, include sex chromosomes, or use different annotation methods may encounter different estimates. To determine how well the expected null distribution compares to the observed PDF, we compare each distribution directly in a Q–Q plot (Fig. 6b). The expected null distribution also fits the observed data well.

P-values suitable for multiple comparison correction via FDR methods should have a probability distribution on the interval [0,1] that is uniform in the null case, i.e., its cumulative distribution is diagonal in the null case (Benjamini and Hochberg, 1995; this is the basis for the false discovery rate method). Our Beta-distributed experimental *P*-values need to be corrected so that they meet the assumptions of the FDR model. Using the analytic β parameter from the null Beta distribution, we fit a cumulative distribution function (CDF) to our observed data yielding a new distribution of corrected *P*-values that deviate from the uniform distribution only when the data are not null. A histogram of the observed corrected

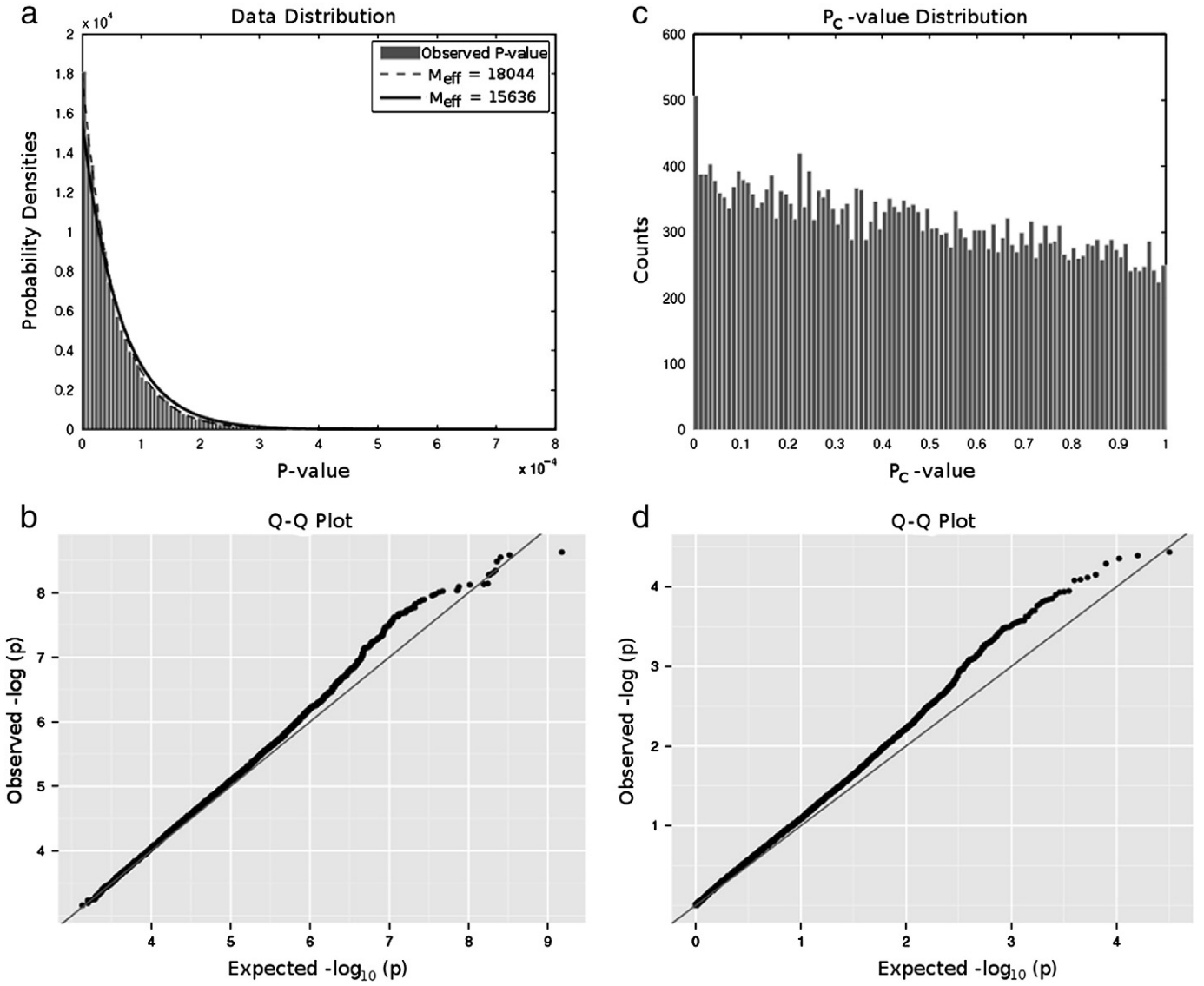


Fig. 6. (a) The normalized histogram of observed *P*-values. The dashed line represents the cumulative distribution function (CDF) of Beta(1, 18,044) where M_{eff} is based on the number of genes tested. The solid line represents the CDF of Beta(1, 15,636) where M_{eff} is an estimate of the number of independent tests from permutation testing. (b) The Q–Q plot shows the observed *P*-values versus those expected from a Beta(1, 15,636) distribution (black dots). The solid gray line represents a purely null distribution of *P*-values. (c) The histogram of corrected *P*-values (P_c) approximately follows a uniform distribution. (d) The Q–Q plot of the observed P_c versus those expected from a null distribution.

P-values (Fig. 6c) shows that the cumulative distribution is approximately equivalent to a uniform distribution. A Q–Q plot of the expected null distribution corrected *P*-values against the observed corrected *P*-values shows that the two distributions differ (Fig. 6d). A Q–Q plot of two identical distributions will lie on the null 45-degree diagonal line ($y = x$). There are two things that can cause a Q–Q plot to deviate from the null: incorrectly modeling a distribution or significant data. The line representing the observed compared to the expected results in the Q–Q plot in Fig. 6d is steeper than and deviates from the null line at lower *P*-values. This suggests that the distribution of P_c values is left skewed and more dispersed than the theoretical uniform distribution (Thode, 2002). It is possible to apply a further correction to our observed P_c distribution by using a Q–Q plot of an analytic null distribution versus the theoretical uniform distribution as a hash table. However, we are only selecting the genes with the lowest *P*-value at each voxel so monotonic *P*-value corrections will not change the distribution of P_c-values.

We used two methods to control the FDR of the corrected *P*-values (P_c). We used the original FDR method (Benjamini and Hochberg, 1995), which appropriately controls for multiple comparisons when the covariance of test statistics shows a positive regression dependency (Benjamini and Yekutieli, 2001). We found that the false discovery rate for the second most highly associated gene in our results (*LRDD*) could only be controlled at a threshold of $q = 0.30$ (i.e., allowing a 30% false discovery rate) after applying a statistical threshold of $P_c = 5.36 \times 10^{-4}$. In addition, the pFDR *q*-value threshold (Storey, 2003) was $q = 0.23$ for the most significantly associated gene at any voxel (*GAB2*). In other words, the vGeneWAS results could not be controlled at the conventional false discovery rate, but show promise.

Post hoc analysis

Voxelwise GeneWAS results in a map that shows only the top gene at each voxel. The top gene at each voxel may be the most significant gene in a certain region, but it may also have a more distributed effect throughout the brain, with effects in additional regions where it is not the top gene. In addition, genes that do not have a large main effect might never be selected in this type of analysis, but still could have a large distributed effect on the brain.

We tested the effect of the top gene in our analysis, *GAB2*, at every voxel across the brain using PCReg. We stored each *P*-value in a map and applied the original FDR method. Voxels surviving the FDR threshold are shown in Fig. 7. These are *post hoc* tests, so are exploratory, and require replication in independent samples, but it is quite clear that *GAB2* has a much greater distributed effect on the brain than could be determined from the vGeneWAS results. Future implementations of vGeneWAS might consider the effects of multiple genes at a voxel to account for the case where a gene is significant in its effect of explaining variations in the image, but is not necessarily the top gene. In addition, vGeneWAS could be further improved by considering the distributed effects of genes. If a gene has an effect over a large region, but is not the top voxel, it will be completely overlooked in the current implementation of vGeneWAS. Adaptation of cluster-level inference to these maps would be of interest, as well as tests that combine cluster extent and height (Hayasaka and Nichols, 2004). Existing adaptations of the original FDR method, such as "searchlight" FDR, could be useful here as it produces region correction thresholds that are sensitive to small clusters of positive signals in imaging data, but appropriately conservative in its correction of false positives (Langers et al., 2007).

To understand the contributions of each individual SNP in the *GAB2* gene, we performed *post hoc* association tests for each SNP with the phenotype value from the top voxel in the brain. It should be noted, however, that choosing the *GAB2* gene to compare the results of SNP-based linear regression with gene-based PCReg provides *P*-values that are biased by the previous gene-wide brain-wide search because *GAB2* was identified using PCReg. There were 20 SNPs from the *GAB2* gene in our genotyped data. Of these, only three passed the nominal significance level in SNP-based association tests ($P=0.05$). The most significant SNP identified, rs7927923, has $P=9.1\times10^{-3}$. The other two significant SNPs, rs1981405 and rs1893447, showed effects with $P=0.027$ and $P=0.049$, respectively. A total of 16 SNPs out of 20 are in high LD with the most significantly associated SNP ($r^2>0.6$). Only one of the SNPs from our analysis overlapped with SNPs used in previous *GAB2* association studies, most likely because we are using different genotyping platforms. Clearly, the gene-based test was more powerful at detecting an association in this case than each SNP tested individually (compare the dotted line with the colored dots in Fig. 8).

Power comparisons

To assess the differences in power afforded by vGeneWAS relative to existing SNP-based methods, we compared the P_c-values from vGWAS obtained in our previous study by Stein et al. (2010a), with the P_c-values resulting from vGeneWAS (Fig. 9). The proportion of P_c-values greater than a given FDR threshold for each method is directly related to differences in effect sizes. The FDR of the results from vGWAS could only be controlled at a threshold value of $q=0.50$, whereas the FDR threshold for vGeneWAS is somewhat lower, although not passing the conventional FDR level ($q=0.30$; Fig. 9). This suggests that the vGeneWAS method may have more power, in principle, to detect genetic associations, although neither test controlled the false discovery rate at the conventional level.

Discussion

Methodological overview

Here we present a method to conduct a voxelwise gene-wide association study (vGeneWAS), testing the aggregate effect of multiple SNPs within each gene. In summary, (1) we implemented a gene-based association test using principal components regression (PCReg); (2) we performed association tests at every voxel within a full brain mask where the value at each voxel was the local volume difference relative to the mean template while controlling for age and sex; (3) we generated a Beta distribution of *P*-values by selecting only the most significant gene at each voxel; (4) with permutation tests, we estimated the effective number of tests performed; and (5) we corrected for multiple comparisons in a two step procedure—we estimated β using the CDF of the analytic Beta distribution and then corrected the new uniform distribution using two different FDR methods.

None of the genes identified passed the standard FDR threshold ($q=0.05$). However, many of the genes identified have previously been associated with brain differences or disorders. The top gene identified is a known Alzheimer's risk gene, *GAB2*, lending plausibility to the method. Many of the genes identified are highly expressed in the brain or differentially expressed, depending on disease status. The

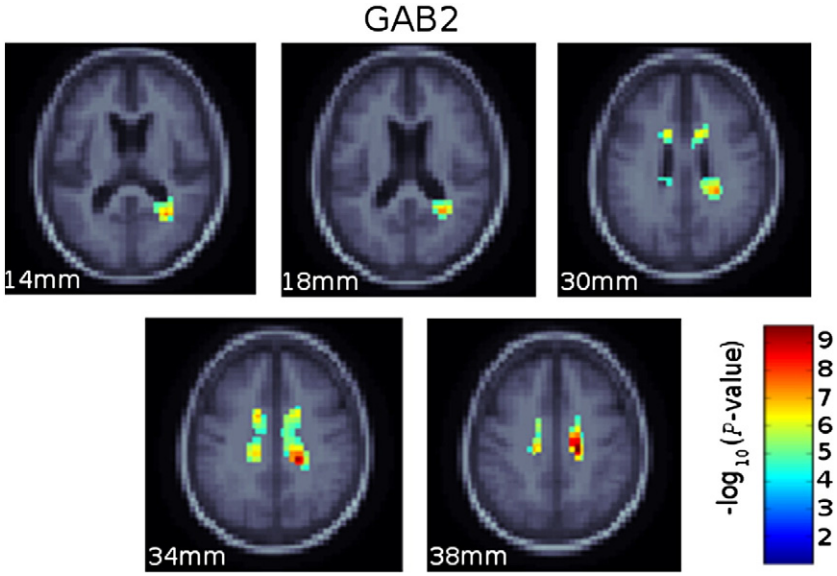


Fig. 7. Map of *P*-values for *GAB2* at every voxel in the brain after correction for multiple comparisons across voxels (but not corrected for search across the genome, as we are only testing one gene) using the original FDR method. *P*-values significant after FDR correction (at $q=3.4\times10^{-4}$) are color-coded. Warmer colors are more significant. *GAB2* has a more distributed effect on the brain than is evident in the vGeneWAS results.

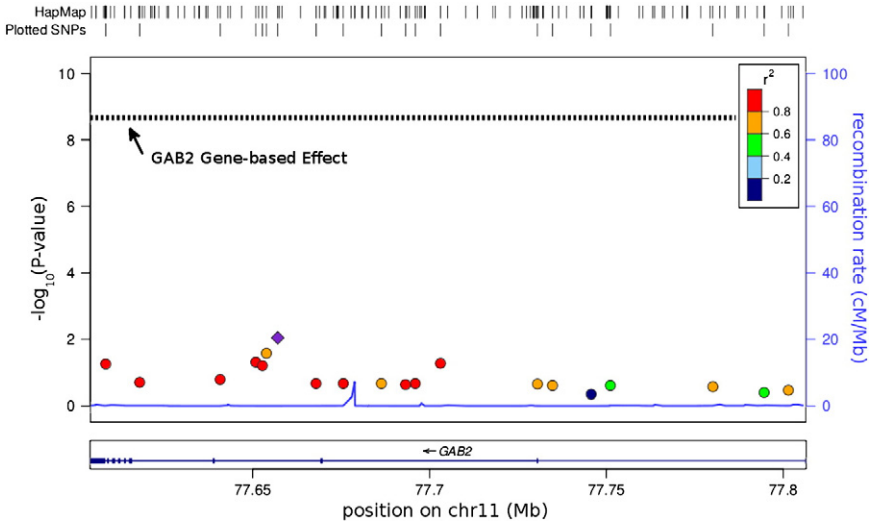


Fig. 8. Genetic association plot, for different SNPs in the *GAB2* gene, at the top voxel from our analysis. The $-\log_{10}$(*P*-value) of each SNP in *GAB2* is plotted against its position in the gene. Each of the points is color coded by level of LD (compared to the top SNP, *the purple diamond dot*) as measured by r^2. The $-\log_{10}$(*P*-value) of the gene-based PCReg test for *GAB2* at this voxel is overlaid on the plot for comparison (dotted black line). In this case, the gene-based test shows a greater effect size than univariate tests on any of the component SNPs treated independently. This shows that the gene-based test can be more powerful than performing separate tests on component SNPs.

findings in this study warrant further examination and replication attempts.

Assessment of the model

Our method selects the top gene at each voxel, to reduce the amount of data. Choosing only the top gene at each voxel, however, can hinder the extensibility of our results. This may miss many genes with distributed effects, if the main effect of the gene is never the largest at any voxel. Future implementations of vGeneWAS could consider the relationship among voxels when performing association tests. Liu et al. (2009) used parallel ICA to relate brain network data from fMRI to SNP data. They selected only a small set of 367 predefined SNPs based on a set of candidate genes for schizophrenia; this does not leverage all of the available data in the genome. Similar approaches have been attempted on voxel-based morphometry (VBM) data from structural MRI (Jagannathan et al., 2010). However, this approach used the same subset of SNPs used by Liu et al. (2009). Vounou et al. (2010) proposed a method called sparse reduced-rank regression (sRRR) which uses the sparseness of signals to simultaneously select phenotypes and genotypes. Power estimates suggest that sRRR is more powerful than using individual tests at each voxel; this may prove to be very useful in the future.

Principal components regression (PCReg) is an efficient method to test the joint association of all SNPs within a gene simultaneously. PCReg detected associations with genes missed by SNP-based regression (Fig. 2b). By leveraging the LD in a gene, PCReg encodes variance throughout a gene to test for associations. We identified situations where SNP-based regression models may have more power (Fig. 2a). If a single SNP has a large main effect, then testing the joint effect of all SNPs within that gene may dilute the association; the cumulative *P*-value from gene-based tests may be lower. However, when one considers the drastic reduction in the number of independent tests when comparing SNP-based linear regression with PCReg, gene-based testing offers advantages.

Another concern with PCReg and related regression models is that each predictor added to the regression model consumes a degree of freedom. There may then be some detection bias in the regression model, where smaller genes are found to be more significantly associated with the phenotype than larger genes, because the regression models of small genes have more degrees of freedom (Chapman and Whittaker, 2008). While we did not observe this effect, it is an important factor to consider when interpreting results. Additionally, SNPs combined into a single test statistic in PCReg could have different directions of effects, disrupting the power to detect an association. However, the situation where a gene contains SNPs with negative correlations with respect to the phenotype may be relatively rare as it requires two nearby loci to be in LD with different causal variants (Chapman and Whittaker, 2008).

Other multivariate regression methods may offer greater power to detect genetic associations than PCReg, which is used here as an example. Wang and Elston (2007) compress genome-wide genotyping data across subjects using a Fourier transform, and assign weights to the low-frequency components in a regression model. This method is similar to PCReg, as it collapses the number of genetic tests performed while capturing much of the variance across markers. Kernel-based methods have been implemented as non-parametric

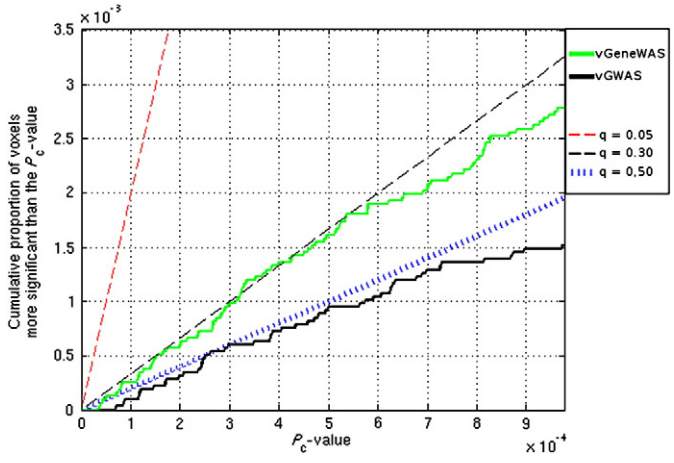


Fig. 9. vGeneWAS may control the false discovery rate better than vGWAS. The cumulative distribution function (CDF) of P_c-values from vGeneWAS (*solid green line*) is compared to the CDF of P_c-values from vGWAS (Stein et al., 2010a; *solid black line*). Three lines represent different correction thresholds of $q = 0.05$ (*red dashed line*), $q = 0.30$ (*black dashed line*), and $q = 0.50$ (*blue dotted line*).

gene-based tests to increase power over SNP-based methods; however, these methods have only been implemented in case–control studies (Mukhopadhyay et al., 2010). Ridge regression (Malo et al., 2008; Sun et al., 2009) and lasso-based penalized regression (Zou, 2006; Wu et al., 2009) can both powerfully detect associations in genetic data. In fact, direct comparisons between ridge regression, lasso-based penalized regression, and PCReg show that the first two methods may be more powerful than PCReg depending on the underlying genetic architecture (Bovelstad et al., 2007; Benner et al., 2010). However, ridge regression and especially lasso-based penalized regression are extremely computationally intensive compared to PCReg. There is a huge computational requirement to complete a vGeneWAS analysis, which searches the whole image in addition to the whole genome. Due to this, we decided to strike a balance between power and computational complexity to complete the analysis in a feasible time frame. Future implementations of vGeneWAS could be improved by using additional multivariate regression methods, although they may need to be modified for speed.

A current limitation of our method, as described here, is that in its current form family-based designs (such as pedigree structures) cannot be used. The patterns in allele frequencies in a family cohort depend on kinship, and mixed-effects models would be required to control for kinship structure. Several such methods exist for SNP-based linear regression, such as Efficient Mixed-Model Association (EMMA; Kang et al., 2008). However, to the best of our knowledge, multivariate mixed-model regression has not been attempted in a genetic context. One multivariate gene-based method called versatile gene-based association (VEGAS) avoids multivariate mixed-model regression by converting SNP-based *P*-values into a multivariate gene-based test statistic (Liu et al., 2010). As there are already many methods for SNP-based mixed-model regression, VEGAS is aptly suited to perform gene-based association tests in family-based populations. As the VEGAS test statistic is determined by SNP-based *P*-values, it will not be able to detect associations where the cumulative *P*-values are not significant. In this way, if a series of SNPs contribute a small amount of variance in a gene, VEGAS will miss them, because the SNP-based method will as well.

While this is one of the largest imaging genetics studies to date, our sample size still may be underpowered to detect moderate effects of genetic variants on the brain. Future studies in imaging genetics are likely to benefit from meta-analysis, which aggregates GWAS results from multiple cohorts to determine reproducible genetic associations. This aggregation of large datasets can be used to boost power to detect SNPs with smaller effects. One such effort now underway is the ENIGMA project (ENIGMA Consortium, 2011). In cases where there is not enough data available to perform a true meta-analysis, discovery and replication datasets may be useful. In early tests using brain images, some genetic associations seen in a discovery cohort have been replicated in independent samples (e.g., Rajagopalan et al., 2011). However, our main purpose in this study is to demonstrate a method to conduct voxelwise, gene-based analyses, which will become more powerful as imaging databases continue to grow rapidly worldwide in size and content. In assessing whether our results may generalize to new datasets, we note that we examined only a tight age range in our study (elderly subjects), and this may affect the genes found to have morphometric effects on the brain. If the genes have a varying expression pattern over time some of the top genes detected in our analysis may not be dominant during other parts of the lifespan. Although we controlled for age effects on brain structure in our analysis, it is still unknown whether the identified genetic associations with brain morphology are under some age-related influence; it is also not known if these genes are expressed in a typical/atypical age-related fashion. Datasets drawn from different parts of the lifespan would offer maximal power to detect genetic variants relevant for brain structure (as discussed in Braskie et al., in press; see Rajagopalan et al., 2011, for an example).

One limitation of GWAS analyses is that they overlook rare variants, which are also emerging as key players in the genetic underpinnings of mental disorders (Bansal et al., 2010). Our method does not consider these rare variants, but they may play an important role in explaining variance in complex traits that is not accounted for by common variants. Examination of rare variants is still relatively costly (as it requires deep sequencing of large numbers of subjects). Some types of rare variant can be genotyped on SNP-chip platforms (such as copy number variants) they require separate analytical techniques from those considered in this paper (Bansal et al., 2010). Each of these limitations will be more feasible to address when the cost of deep sequencing drops and sample sizes are large enough to reliably implicate many genes simultaneously.

Our gene-based test may be more powerful than univariate methods in certain cases, but not always. The five top genes identified in the present study do have some biological plausibility; some are known to be expressed in the brain and implicated in brain disorders. However, there are other genes missing in the short list of genes in the present findings that are frequently found using univariate approaches and strongly implicated in complex behavioral pathologies across mammalian species, such as the *BDNF* val66met substitution.

To explain this, we note that the analysis of currently available imaging genetics data is very underpowered. In addition, sample sizes needed to reliably detect a genetic association are even greater when multiple genes are assessed or when genome-wide search is conducted. By contrast, the *BDNF* val66met substitution is often treated as a candidate gene, and if that is done, it is conventionally agreed that its effects must only satisfy a nominal significance level of $P<0.05$, if no other genes or SNPs within it are tested. In our own work on a different cohort scanned with DTI (Chiang et al., 2010), we were able to detect robust associations between the *BDNF* val66met polymorphism and fiber integrity (fractional anisotropy) assessed with DTI. We were also able to replicate these associations in two non-overlapping (independent) samples of subjects. Even so, the significance level used ($P<0.05$) was far more lenient than the very strict thresholds required to control for false positives when the whole genome is searched. As such, false negative results in a GWAS study (e.g., not finding a significant association of the *BDNF* val66met substitution when it is in fact a true association) does not mean that a gene does not affect a phenotype, just that an association was not detected at the very stringent statistical cut-off applied to account for multiple comparisons across the genome.

A review of the literature also suggests that *BDNF* val66met, while a popular target of study, has a mixed history. The BDNFval66met substitution has been inconsistently implicated in mood disorders, Alzheimer's disease, and quantitative measures of memory (Bagnoli et al., 2004; Combarros et al., 2004; Nacmias et al., 2004; Nishimura et al., 2004; Tsai et al., 2004; Desai et al., 2005; Matsushita et al., 2005; Vepsalainen et al., 2005). In a secondary, *post hoc*, analysis, we tested the effect of the BDNFval66met allele in this current sample using standard univariate regression (with a dominant model, controlling for age and sex) and it did not survive correction for multiple spatial comparisons. An association test of the *BDNF* val66met allele with the hippocampal volume of the 731 subjects in this dataset did not survive the nominal significance level of $P=0.05$. As such, we were not able to use *BDNF* as a "gold standard" gene; arguably, there are not yet any such genes with universally replicable associations on brain structure that can be used to gauge the face validity of novel methods.

Another limitation is that any approach that stops after selecting only one gene per voxel is not biologically plausible as a model of phenotypes with complex genetic determination. As such, the development of gene-based tests should be considered as a way station towards a more sophisticated treatment of complex genetic effects, in pathway or gene–gene interaction models (Inkster et al., 2010). However, testing gene–gene interactions in the vGeneWAS framework is computationally intractable as the number of tests

quickly approaches ${}^{n}C_k = n!/((n-k)!k!)$ at each voxel in the whole brain to test for interactions among all sets of k different genes drawn from an overall pool of N genes. Additionally, interaction effect sizes are generally much smaller than the main effects we are searching for in this paper, so our sample size would need to be much larger to accommodate the necessary correction for multiple comparisons and smaller effect sizes (Cordell, 2009). As genome-wide interaction analysis (GWIA; Marchini et al., 2005) is computationally intensive and underpowered with current imaging datasets, we recently developed an alternative method (Chiang et al., 2011a,b) to detect likely gene–gene interactions among SNPs, without having to compute all $N(N-1)/2$ pairwise or all $n!/((n-k)!k!)$ kth-order interactions on the genome. Two advantages of this genetic network analysis, relative to genome-wide interaction analysis (GWIA) are apparent: (1) genetic correlation can tap into the natural latent structure of gene action in a brain image; and (2) voxel clustering by genetic affinity leads to high power to find SNPs with correlated effects in genome-wide scans.

Biological significance of the findings

Gene-based tests of association across the genome and brain have not been attempted before, to our knowledge. Recently, imaging genetics studies have focused on single-locus associations with summary brain volume measures or 3D statistical parametric maps. vGeneWAS advances the burgeoning field of imaging genetics by providing the framework to perform multivariate, gene-based association tests. It does not restrict analyses by requiring prior hypotheses about a specific causal variant or *ad hoc* region of interest. vGeneWAS is the first attempt to apply gene-based tests to morphometric imaging data and opens up more possibilities to discover putative genetic variants that contribute to differences in brain structure. This may help when the main effect of each variant in a gene is too small to detect with traditional SNP-based methods.

Although vGeneWAS is a multivariate, gene-based method, we identified genes previously associated with brain disease using SNP-based tests. Many variants in the *GAB2* gene are implicated in the development of late-onset Alzheimer's disease (LOAD) and are thought to interact with the *APOE* epsilon 4 allele. In the pattern of effects for *GAB2* on the brain (Fig. 7), the highlighted areas are generally periventricular, and ventricular enlargement is a prominent characteristic of AD (de Leon et al., 1989; Chou et al., 2009). As we noted in our prior papers on TBM in Alzheimer's disease (Hua et al., 2008), there is occasionally a ring of voxels around the lateral ventricles that show partial volume effects that mostly like reflect ventricular expansion. Clearly, the ventricular expansion itself indirectly results from the diffuse loss of brain parenchyma, so the changes detected there may also reflect, to some extent, atrophic processes more remote than the voxels singled out in the voxel-based maps. *LRDD* is highly expressed in the brain and is involved with DNA repair including signaling apoptosis in tumor cells. *PTPRB* is associated with addiction to drugs and alcohol and may be involved with tumor regulation (Telliez et al., 2000; Wu et al., 2006; Ishiguro et al., 2008; Brown et al., 2009). Based on gene expression and links to brain diseases, many of the genes identified in our analysis may have differential morphometric effects across the brain. In addition to some of the more well-studied genes, we identified many genes such as *IGSF5* and *ZNF462* that have little research available to infer their plausibility. However, almost all of the genes identified in our analysis are highly expressed in the brain, which at least suggests that the genes may have a role in brain function. Further analysis is required to examine to what extent each gene variant identified in our analysis mediates brain differences.

Many of the associations identified here seem to have a plausible story, but we need to consider that some of the patterns of association, especially clusters of association, may be due to short-range spatial correlations in the images. Adjacent voxels in brain scans tend to covary, as do Jacobian maps used to represent a localized measure of volume difference. These methods rely on non-linear algorithms that generate spatially smooth deformations. In addition to the simulated (null) cluster size maps in this study, Stein et al. (2010a) found that a small amount of spatial clustering is seen even if the genetic data is null. Also, voxels were down-sampled which may introduce partial volume effects. However, performing a vGeneWAS scan on non-down-sampled, original sized images is estimated to take 4372 days (or approximately 12 years) to complete. To this end, the extent to which a gene affects regions of the brain should be interpreted cautiously; however, certain patterns of gene effects that appear in non-adjacent structures or in large clusters may signal gene effects not attributable purely to spatial smoothing or partial volume effects.

To better understand the contribution of genes to global versus local brain structure differences, we conducted association analysis on both (1) the globally normalized brain images and (2) estimates of total intracranial volume (eTIV) that contain information on overall differences in brain scale. We searched for specific gene effects on global brain volume differences using our gene-based method. We computed brain volume measures (eTIV) from our dataset using the automated FreeSurfer package (Fischl et al., 2002). Using the eTIV measure as the phenotype we tested each of the 18,044 genes for association. Looking at the top genes that we found in our analysis of the normalized images, none of the top 20 most highly associated genes was associated with eTIV phenotype. This provides further evidence that the genetic effects we are detecting exert influences on regional brain volumes rather than simply reflecting non-specific effects on the overall volume of the brain.

Additionally, results should be interpreted cautiously when global anatomical normalization is used. By removing, as far as possible, the effects of individual brain size variation from the data, it is possible to discover genes that may have a specific effect on a particular structure, above and beyond any overall genetic effects on brain size (as brain size itself is heritable). Global normalization is commonly performed in all brain mapping studies, as many extraneous factors affect an individual's head size, body size or height that may not be relevant for cognition or for understanding brain function. Global anatomical normalization adjusts for this source of variance in the data, to a large extent, making more localized effects easier to identify. Even so, by applying global anatomical normalization, some genes may be missed that influence the total size of the brain. In fact, if a gene were responsible for influencing brain size, but had a uniform effect on all brain regions, it would be missed in the current analysis, as global effects are discounted. As such, in addition to mapping gene effects, it makes sense to also perform genetic analyses of whole brain size, as has been performed in two recent studies (Paus et al., 2011; ENIGMA Consortium, 2011).

In conclusion, our method may be used to perform gene-based tests on any 3D brain maps, such as data from voxel-based morphometry, diffusion tensor imaging, and cortical surface data. In addition, we found a set of candidate genes that may substantially affect brain morphometry and warrant further study.

Acknowledgments

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904, 3U01AG024904-03S5). 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, AstraZeneca AB, Bayer Schering Pharma AG, Bristol-Myers Squibb, Eisai Global Clinical Development, Elan Corporation, Genentech, GE Healthcare, GlaxoSmithKline, Innogenetics, Johnson and Johnson, Eli Lilly and Co., Medpace, Inc., Merck and Co., Inc., Novartis AG, Pfizer Inc, F. Hoffman-La Roche, Schering-Plough, Synarc, Inc., as

well as non-profit partners the Alzheimer's Association and Alzheimer's Drug Discovery Foundation, with participation from the U.S. Food and Drug Administration. Private sector contributions to ADNI are facilitated by the Foundation for the National Institutes of Health (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, K01 AG030514, and the Dana Foundation. We also thank the many contributors to ADNI-1 genotyping sample curation at NCRAD (Kelley Faber), performing BeadChip assays at TGen (David Craig), and bioinformatics problem solving (Indiana U: Kwangsik Nho; UC Irvine: Anita Lakatos, Guia Guffanti; Pfizer: Bryan DeChairo). Additional support for algorithm development was provided by R01 EB008281, R01 HD050735, RC2 AG036535, and R01 AG020098.

References

Altshuler, D., Brooks, L.D., Chakravarti, A., Collins, F.S., Daly, M.J., Donnelly, P., Gibbs, R.A., Belmont, J.W., Boudreau, A., Leal, S.M., Hardenbol, P., Pasternak, S., Wheeler, D.A., Willis, T.D., Yu, F.L., Yang, H.M., Zeng, C.Q., Gao, Y., Hu, H.R., Hu, W.T., Li, C.H., Lin, W., Liu, S.Q., Pan, H., Tang, X.L., Wang, J., Wang, W., Yu, J., Zhang, B., Zhang, Q.R., Zhao, H.B., Zhao, H., Zhou, J., Gabriel, S.B., Barry, R., Blumenstiel, B., Camargo, A., Defelice, M., Faggart, M., Goyette, M., Gupta, S., Moore, J., Nguyen, H., Onofrio, R.C., Parkin, M., Roy, J., Stahl, E., Winchester, E., Ziaugra, L., Shen, Y., Yao, Z.J., Huang, W., Chu, X., He, Y.G., Jin, L., Liu, Y.F., Shen, Y.Y., Sun, W.W., Wang, H.F., Wang, Y., Wang, Y., Wang, Y., Xiong, X.Y., Xu, L., Waye, M.M.Y., Tsui, S.K.W., Xue, H., Wong, J.T.F., Galver, I.L.M., Fan, J.B., Murray, S.S., Oliphant, A.R., Chee, M.S., Montpetit, A., Chagnon, F., Ferretti, V., Leboeuf, M., Olivier, J.F., Phillips, M.S., Roumy, S., Sallee, C., Verner, A., Hudson, T.J., Frazer, K.A., Ballinger, D.G., Cox, D.R., Hinds, D.A., Stuve, L.L., Kwok, P.Y., Cai, D.M., Koboldt, D.C., Miller, R.D., Pawlikowska, L., Taillon-Miller, P., Xiao, M., Tsui, L.C., Mak, W., Sham, P.C., Song, Y.Q., Tam, P.K.H., Nakamura, Y., Kawaguchi, T., Kitamoto, T., Morizono, T., Nagashima, A., Ohnishi, Y., Sekine, A., Tanaka, T., Tsunoda, T., Deloukas, P., Bird, C.P., Delgado, M., Dermitzakis, E.T., Gwilliam, R., Hunt, S., Morrison, J., Powell, D., Stranger, B.E., Whittaker, P., Bentley, D.R., Daly, M.J., de Bakker, P.I.W., Barrett, J., Fry, B., Maller, J., McCarroll, S., Patterson, N., Pe'er, I., Purcell, S., Richter, D.J., Sabeti, P., Saxena, R., Schaffner, S.F., Varilly, P., Stein, L.D., Krishnan, L., Smith, A.V., Thorisson, G.A., Chen, P.E., Cutler, D.J., Kashuk, C.S., Lin, S., Abecasis, G.R., Guan, W.H., Munro, H.M., Qin, Z.H.S., Thomas, D.J., McVean, G., Bottolo, L., Eyheramendy, S., Freeman, C., Marchini, J., Myers, S., Spencer, C., Stephens, M., Cardon, L.R., Clarke, G., Evans, D.M., Morris, A.P., Weir, B.S., Tsunoda, T., Mullikin, J.C., Sherry, S.T., Feolo, M., Zhang, H.C., Zeng, C.Q., Zhao, H., Matsuda, I., Fukushima, Y., Macer, D.R., Suda, E., Rotimi, C.N., Adebamowo, C.A., Ajayi, I., Aniagwu, T., Marshall, P.A., Nkwodimmah, C., Royal, C.D.M., Leppert, M.F., Dixon, M., Peiffer, A., Qiu, R.Z., Kent, A., Kato, K., Niikawa, N., Adewole, I.F., Knoppers, B.M., Foster, M.W., Clayton, E.W., Muzny, D., Nazareth, L., Sodergren, E., Weinstock, G.M., Wheeler, D.A., Yakub, I., Gabriel, S.B., Richter, D.J., Ziaugra, L., Birren, B.W., Wilson, R.K., Fulton, L.L., Rogers, J., Burton, J., Carter, N.P., Clee, C.M., Griffiths, M., Jones, M.C., McLay, K., Plumb, R.W., Ross, M.T., Sims, S.K., Willey, D.L., Chen, Z., Han, H., Kang, L., Godbout, M., Wallenburg, J.C., Archeveque, P.L., Bellemare, G., Saeki, K., Wang, H.G., An, D.C., Fu, H.B., Li, Q., Wang, Z., Wang, R.W., Holden, A.L., Brooks, L.D., McEwen, J.E., Bird, C.R., Guyer, M.S., Nailer, P.J., Wang, V.O., Peterson, J.L., Shi, M., Spiegel, J., Sung, L.M., Witonsky, J., Zacharia, L.F., Kennedy, K., Jamieson, R., Stewart, J., Consortium, I.H., 2005. A haplotype map of the human genome. Nature 437, 1299–1320.

Bagnoli, S., Nacmias, B., Tedde, A., Guarnieri, B.M., Cellini, E., Petruzzi, C., Bartoli, A., Ortenzi, L., Sorbi, S., 2004. Brain-derived neurotrophic factor genetic variants are not susceptibility factors to Alzheimer's disease in Italy. Ann. Neurol. 55, 447–448.

Bansal, V., Libiger, O., Torkamani, A., Schork, N.J., 2010. Statistical analysis strategies for association studies involving rare variants. Nat. Rev. Genet. 11, 773–785.

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate — a practical and powerful approach to multiple testing. J. R. Stat. Soc. B Methodol. 57, 289–300.

Benjamini, Y., Yekutieli, D., 2001. The control of the false discovery rate in multiple testing under dependency. Ann. Stat. 29, 1165–1188.

Benner, A., Zucknick, M., Hielscher, T., Ittrich, C., Mansmann, U., 2010. High-dimensional Cox models: the choice of penalty as part of the model building process. Biom. J. 52, 50–69.

Bovelstad, H.M., Nygard, S., Storvold, H.L., Aldrin, M., Borgan, O., Frigessi, A., Lingjaerde, O.C., 2007. Predicting survival from microarray data — a comparative study. Bioinformatics 23, 2080–2087.

Braskie, M.N., Ringman, J.M., Thompson, P.M., in press. Neuroimaging measures as endophenotypes in Alzheimer's disease. Int. J. Alzheimers D. Article ID 490140. doi:10.4061/2011/490140.

Brown, C.J., Lain, S., Verma, C.S., Fersht, A.R., Lane, D.P., 2009. Awakening guardian angels: drugging the p53 pathway. Nat. Rev. Cancer 9, 862–873.

Cannon, T.D., Keller, M.C., 2006. Endophenotypes in the genetic analyses of mental disorders. Annu. Rev. Clin. Psychol. 2, 267–290.

Chapman, J., Whittaker, J., 2008. Analysis of multiple SNPs in a candidate gene or region. Genet. Epidemiol. 32, 560–566.

Chapuis, J., Hannequin, D., Pasquier, F., Bentham, P., Brice, A., Leber, I., Frebourg, T., Deleuze, J.F., Cousin, E., Thaker, U., Amouyel, P., Mann, D., Lendon, C., Campion, D., Lambert, J.C., 2008. Association study of the GAB2 gene with the risk of developing Alzheimer's disease. Neurobiol. Dis. 30, 103–106.

Chiang, M.C., Barysheva, M., Toga, A.W., Medland, S.E., Hansell, N.K., James, M.R., McMahon, K.L., de Zubicaray, G.I., Martin, N.G., Wright, M.J., Thompson, P.M., 2010. BDNF gene effects on brain circuitry replicated in 455 twins. Neuroimage 55 (2), 448–454.

Chiang, M.C., Barysheva, M., McMahon, K.L., de Zubicaray, G.I., Johnson, K., Martin, N.G., Toga, A.W., Wright, M.J., Thompson, P.M., 2011a. Understanding the network topology of gene action on brain microstructure: an N=531 twin study. Organization for Human Brain Mapping Conference [June 2011].

Chiang, M.C., McMahon, K.L., de Zubicaray, G.I., Martin, N.G., Toga, A.W., Wright, M.J., Thompson, P.M., 2011b. Hierarchical clustering of the genetic connectivity matrix reveals the network topology of gene action on brain microstructure. Int. Symp. Biomed. Imaging. Chicago, Il.

Chou, Y.Y., Lepore, N., Chiang, M.C., Avedissian, C., Barysheva, M., McMahon, K.L., de Zubicaray, G.I., Meredith, M., Wright, M.J., Toga, A.W., Thompson, P.M., 2009. Mapping genetic influences on ventricular structure in twins. Neuroimage 44, 1312–1323.

Combarros, O., Infante, J., Llorca, J., Berciano, J., 2004. Polymorphism at codon 66 of the brain-derived neurotrophic factor gene is not associated with sporadic Alzheimer's disease. Dement. Geriatr. Cogn. Disord. 18, 55–58.

Cordell, H.J., 2009. Detecting gene–gene interactions that underlie human diseases. Nat. Rev. Genet. 10, 392–404.

de Leon, M.J., George, A.E., Reisberg, B., Ferris, S.H., Kluger, A., Stylopoulos, L.A., Miller, J.D., La Regina, M.E., Chen, C., Cohen, J., 1989. Alzheimer's disease: longitudinal CT studies of ventricular change. AJR Am. J. Roentgenol. 152, 1257–1262.

De Luca, V., Wang, H., Squassina, A., Wong, G.W.H., Yeomans, J., Kennedy, J.L., 2004. Linkage of M5 muscarinic and alpha 7-nicotinic receptor genes on 15q13 to schizophrenia. Neuropsychobiology 50, 124–127.

del Arco, A., Satrustegui, J., 1998. Molecular cloning of Aralar, a new member of the mitochondrial carrier superfamily that binds calcium and is present in human muscle and brain. J. Biol. Chem. 273, 23327–23334.

Desai, P., Nebes, R., DeKosky, S.T., Kamboh, M.I., 2005. Investigation of the effect of brain-derived neurotrophic factor (BDNF) polymorphisms on the risk of late-onset Alzheimer's disease (AD) and quantitative measures of AD progression. Neurosci. Lett. 379, 229–234.

ENIGMA Consortium, 2011. Genome-wide association meta-analysis of hippocampal volume: results from the ENIGMA consortium. Organization for Human Brain Mapping Conference. [June 2011].

Ewens, W.J., Grant, G., 2001. Statistical Methods in Bioinformatics: An Introduction. Springer, New York.

Fillard, P., Arsigny, V., Pennec, X., Hayashi, K.M., Thompson, P.M., Ayache, N., 2007. Measuring brain variability by extrapolating sparse tensor fields measured on sulcal lines. Neuroimage 34, 639–650.

Fischl, B., Salat, D.H., Busa, E., Albert, M., Dieterich, M., Haselgrove, C., van der Kouwe, A., Killiany, R., Kennedy, D., Klaveness, S., Montillo, A., Makris, N., Rosen, B., Dale, A.M., 2002. Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. Neuron 33, 341–355.

Flint, J., Greenspan, R.J., Kendler, K.S., 2010. How Genes Influence Behavior. Oxford University Press, Oxford; New York.

Frazer, K.A., Ballinger, D.G., Cox, D.R., Hinds, D.A., Stuve, L.L., Gibbs, R.A., Belmont, J.W., Boudreau, A., Hardenbol, P., Leal, S.M., Pasternak, S., Wheeler, D.A., Willis, T.D., Yu, F., Yang, H., Zeng, C., Gao, Y., Hu, H., Hu, W., Li, C., Lin, W., Liu, S., Pan, H., Tang, X., Wang, J., Wang, W., Yu, J., Zhang, B., Zhang, Q., Zhao, H., Zhou, J., Gabriel, S.B., Barry, R., Blumenstiel, B., Camargo, A., Defelice, M., Faggart, M., Goyette, M., Gupta, S., Moore, J., Nguyen, H., Onofrio, R.C., Parkin, M., Roy, J., Stahl, E., Winchester, E., Ziaugra, L., Altshuler, D., Shen, Y., Yao, Z., Huang, W., Chu, X., He, Y., Jin, L., Liu, Y., Sun, W., Wang, H., Wang, Y., Xiong, X., Xu, L., Waye, M.M., Tsui, S.K., Xue, H., Wong, J.T., Galver, L.M., Fan, J.B., Gunderson, K., Murray, S.S., Oliphant, A.R., Chee, M.S., Montpetit, A., Chagnon, F., Ferretti, V., Leboeuf, M., Olivier, J.F., Phillips, M.S., Roumy, S., Sallee, C., Verner, A., Hudson, T.J., Kwok, P.Y., Cai, D., Koboldt, D.C., Miller, R.D., Pawlikowska, L., Taillon-Miller, P., Xiao, M., Tsui, L.C., Mak, W., Song, Y.Q., Tam, P.K., Nakamura, Y., Kawaguchi, T., Kitamoto, T., Morizono, T., Nagashima, A., Ohnishi, Y., Sekine, A., Tanaka, T., Tsunoda, T., Deloukas, P., Bird, C.P., Delgado, M., Dermitzakis, E.T., Gwilliam, R., Hunt, S., Morrison, J., Powell, D., Stranger, B.E., Whittaker, P., Bentley, D.R., Daly, M.J., de Bakker, P.I., Barrett, J., Chretien, Y.R., Maller, J., McCarroll, S., Patterson, N., Pe'er, I., Price, A., Purcell, S., Richter, D.J., Sabeti, P., Saxena, R., Schaffner, S.F., Sham, P.C., Varilly, P., Stein, L.D., Krishnan, L., Smith, A.V., Tello-Ruiz, M.K., Thorisson, G.A., Chakravarti, A., Chen, P.E., Cutler, D.J., Kashuk, C.S., Lin, S., Abecasis, G.R., Guan, W., Li, Y., Munro, H.M., Qin, Z.S., Thomas, D.J., McVean, G., Auton, A., Bottolo, L., Cardin, N., Eyheramendy, S., Freeman, C., Marchini, J., Myers, S., Spencer, C., Stephens, M., Donnelly, P., Cardon, L.R., Clarke, G., Evans, D.M., Morris, A.P., Weir, B.S., Mullikin, J.C., Sherry, S.T., Feolo, M., Skol, A., Zhang, H., Matsuda, I., Fukushima, Y., Macer, D.R., Suda, E., Rotimi, C.N., Adebamowo, C.A., Ajayi, I., Aniagwu, T., Marshall, P.A., Nkwodimmah, C., Royal, C.D., Leppert, M.F., Dixon, M., Peiffer, A., Qiu, R., Kent, A., Kato, K., Niikawa, N., Adewole, I.F., Knoppers, B.M., Foster, M.W., Clayton, E.W., Watkin, J., Muzny, D., Nazareth, L., Sodergren, E., Weinstock, G.M., Yakub, I., Birren, B.W., Wilson, R.K., Fulton, L.L., Rogers, J., Burton, J., Carter, N.P., Clee, C.M., Griffiths, M., Jones, M.C., McLay, K., Plumb, R.W., Ross, M.T., Sims, S.K., Willey, D.L., Chen, Z., Han, H., Kang, L., Godbout, M., Wallenburg, J.C., L'Archeveque, P., Bellemare, G., Saeki, K., An, D., Fu, H., Li, Q., Wang, Z., Wang, R., Holden, A.L., Brooks, L.D., McEwen, J.E., Guyer, M.S., Wang, V.O., Peterson, J.L., Shi, M., Spiegel, J., Sung, L.M., Zacharia, L.F., Collins, F.S., Kennedy, K., Jamieson, R., Stewart, J., 2007. A second generation human haplotype map of over 3.1 million SNPs. Nature 449, 851–861.

Hayasaka, S., Nichols, T.E., 2004. Combining voxel intensity and cluster extent with permutation test framework. Neuroimage 23, 54–63.

Hemminger, B.M., Saelim, B., Sullivan, P.F., 2006. TAMAL: an integrated approach to choosing SNPs for genetic studies of human complex traits. Bioinformatics 22, 626–627.
Hibar, D., Stein, J.L., Jahanshad, N., Baresheva, M., Feng, A., Kogachi, S., McMahon, K., De Zubicaray, G., Hansell, N., Martin, N.G., Wright, M.J., Toga, A., Thompson, P., 2010. Voxelwise Genome-Wide Association of Diffusion Tensor Images Identifies Putative Novel Variants Influencing White Matter Integrity in 467 Related Young Adults. Society for Neuroscience, San Diego, CA.
Hinrichs, A.S., Karolchik, D., Baertsch, R., Barber, G.P., Bejerano, G., Clawson, H., Diekhans, M., Furey, T.S., Harte, R.A., Hsu, F., Hillman-Jackson, J., Kuhn, R.M., Pedersen, J.S., Pohl, A., Raney, B.J., Rosenbloom, K.R., Siepel, A., Smith, K.E., Sugnet, C.W., Sultan-Qurraie, A., Thomas, D.J., Trumbower, H., Weber, R.J., Weirauch, M., Zweig, A.S., Haussler, D., Kent, W.J., 2006. The UCSC genome browser database: update 2006. Nucleic Acids Res. 34, D590–D598.
Hirabayashi, S., Tajima, M., Yao, I., Nishimura, W., Mori, H., Hata, Y., 2003. JAM4, a junctional cell adhesion molecule interacting with a tight junction protein, MAGI-1. Mol. Cell. Biol. 23, 4267–4282.
Ho, A.J., Hua, X., Lee, S., Leow, A.D., Yanovsky, I., Gutman, B., Dinov, I.D., Lepore, N., Stein, J.L., Toga, A.W., Jack, C.R., Bernstein, M.A., Reiman, E.M., Harvey, D.J., Kornak, J., Schuff, N., Alexander, G.E., Weiner, M.W., Thompson, P.M., and the Alzheimer's Disease Neuroimaging Initiative, 2010. Comparing 3 T and 1.5 T MRI for tracking Alzheimer's disease progression with tensor-based morphometry. Hum. Brain Mapp. 31, 499–514.
Hua, X., Leow, A.D., Parikshak, N., Lee, S., Chiang, M.C., Toga, A.W., Jack, C.R., Weiner, M.W., Thompson, P.M., and the Alzheimer's Disease Neuroimaging Initiative, 2008. Tensor-based morphometry as a neuroimaging biomarker for Alzheimer's disease: an MRI study of 676 AD, MCI, and normal subjects. Neuroimage 43, 458–469.
Ikram, M.A., Liu, F., Oostra, B.A., Hofman, A., van Duijn, C.M., Breteler, M.M.B., 2009. The GAB2 gene and the risk of Alzheimer's disease: replication and meta-analysis. Biol. Psychiatry 65, 995–999.
Inkster, B., Nichols, T.E., Saemann, P.G., Auer, D.P., Holsboer, F., Muglia, P., Matthews, P.M., 2010. Pathway-based approaches to imaging genetics association studies: Wnt signaling, GSK3beta substrates and major depression. Neuroimage 53, 908–917.
International Human Genome Sequencing Consortium, 2004. Finishing the euchromatic sequence of the human genome. Nature 431, 931–945.
Ishiguro, H., Gong, J.P., Hall, F.S., Arinami, T., Uhl, G.R., 2008. Association of PTPRB gene polymorphism with drug addiction. Am. J. Med. Genet. B Neuropsychiatr. Genet. 147B, 1167–1172.
Jack, C.R., Bernstein, M.A., Fox, N.C., Thompson, P., Alexander, G., Harvey, D., Borowski, B., Britson, P.J., Whitwell, J.L., Ward, C., Dale, A.M., Felmlee, J.P., Gunter, J.L., Hill, D.L. G., Killiany, R., Schuff, N., Fox-Bosetti, S., Lin, C., Studholme, C., DeCarli, C.S., Krueger, G., Ward, H.A., Metzger, G.J., Scott, K.T., Mallozzi, R., Blezek, D., Levy, J., Debbins, J.P., Fleisher, A.S., Albert, M., Green, R., Bartzokis, G., Glover, G., Mugler, J., Weiner, M.W., and the Alzheimer's Disease Neuroimaging Initiative, 2008. The Alzheimer's Disease Neuroimaging Initiative (ADNI): MRI methods. J. Magn. Reson. Imaging 27, 685–691.
Jagannathan, K., Calhoun, V.D., Gelernter, J., Stevens, M.C., Liu, J., Bolognani, F., Windemuth, A., Ruaño, G., Assaf, M., Pearlson, G.D., 2010. Genetic associations of brain structural networks in schizophrenia: a preliminary study. Biol. Psychiatry 68, 657–666.
Jovicich, J., Czanner, S., Greve, D., Haley, E., van der Kouwe, A., Gollub, R., Kennedy, D., Schmitt, F., Brown, G., MacFall, J., Fischl, B., Dale, A., 2006. Reliability in multi-site structural MRI studies: effects of gradient non-linearity correction on phantom and human data. Neuroimage 30, 436–443.
Kang, H.M., Zaitlen, N.A., Wade, C.M., Kirby, A., Heckerman, D., Daly, M.J., Eskin, E., 2008. Efficient control of population structure in model organism association mapping. Genetics 178, 1709–1723.
Kleinbaum, D.G., 2007. Applied Regression Analysis and Other Multivariable Methods, 4th ed. Brooks/Cole, Australia; Belmont, CA.
Lander, E.S., Schork, N.J., 1994. Genetic dissection of complex traits. Science 265, 2037–2048.
Langers, D.R., Jansen, J.F., Backes, W.H., 2007. Enhanced signal detection in neuroimaging by means of regional control of the global false discovery rate. Neuroimage 38, 43–56.
Leow, A., Huang, S.C., Geng, A., Becker, J., Davis, S., Toga, A., Thompson, P., 2005. Inverse consistent mapping in 3D deformable image registration: its construction and statistical properties. Inf. Process. Med. Imaging 19, 493–503.
Li, Y., Willer, C., Sanna, S., Abecasis, G., 2009. Genotype imputation. Annu. Rev. Genomics Hum. Genet. 10, 387–406.
Lin, K., Tang, M., Han, H., Guo, Y., Lin, Y., Ma, C., 2010. GAB2 is not associated with late-onset Alzheimer's disease in Chinese Han. Neurol. Sci. 31, 277–281.
Liu, J., Pearlson, G., Windemuth, A., Ruano, G., Perrone-Bizzozero, N.I., Calhoun, V., 2009. Combining fMRI and SNP data to investigate connections between brain function and genetics using parallel ICA. Hum. Brain Mapp. 30, 241–255.
Liu, J.Z., McRae, A.F., Nyholt, D.R., Medland, S.E., Wray, N.R., Brown, K.M., Hayward, N.K., Montgomery, G.W., Visscher, P.M., Martin, N.G., Macgregor, S., 2010. A versatile gene-based test for genome-wide association studies. Am. J. Hum. Genet. 87, 139–145.
Lucek, P., Hanke, J., Reich, J., Solla, S.A., Ott, J., 1998. Multi-locus nonparametric linkage analysis of complex trait loci with neural networks. Hum. Hered. 48, 275–284.
Malo, N., Libiger, O., Schork, N.J., 2008. Accommodating linkage disequilibrium in genetic-association analyses via ridge regression. Am. J. Hum. Genet. 82, 375–385.
Marchini, J., Donnelly, P., Cardon, L.R., 2005. Genome-wide strategies for detecting multiple loci that influence complex diseases. Nat. Genet. 37, 413–417.
Matsushita, S., Arai, H., Matsui, T., Yuzuriha, T., Urakami, K., Masaki, T., Higuchi, S., 2005. Brain-derived neurotrophic factor gene polymorphisms and Alzheimer's disease. J. Neural Transm. 112, 703–711.
Mazziotta, J., Toga, A., Evans, A., Fox, P., Lancaster, J., Zilles, K., Woods, R., Paus, T., Simpson, G., Pike, B., Holmes, C., Collins, L., Thompson, P., MacDonald, D., Iacoboni, M., Schormann, T., Amunts, K., Palomero-Gallagher, N., Geyer, S., Parsons, L., Narr, K., Kabani, N., Le Goualher, G., Boomsma, D., Cannon, T., Kawashima, R., Mazoyer, B., 2001. A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM). Philos. Trans. R. Soc. Lond. B Biol. Sci. 356, 1293–1322.
McCarthy, M.I., Abecasis, G.R., Cardon, L.R., Goldstein, D.B., Little, J., Ioannidis, J.P., Hirschhorn, J.N., 2008. Genome-wide association studies for complex traits: consensus, uncertainty and challenges. Nat. Rev. Genet. 9, 356–369.
Mukhopadhyay, I., Feingold, E., Weeks, D.E., Thalamuthu, A., 2010. Association tests using kernel-based measures of multi-locus genotype similarity between individuals. Genet. Epidemiol. 34, 213–221.
Nacmias, B., Piccini, C., Bagnoli, S., Tedde, A., Cellini, E., Bracco, L., Sorbi, S., 2004. Brain-derived neurotrophic factor, apolipoprotein E genetic variants and cognitive performance in Alzheimer's disease. Neurosci. Lett. 367, 379–383.
Neale, B.M., Sham, P.C., 2004. The future of association studies: gene-based analysis and replication. Am. J. Hum. Genet. 75, 353–362.
Nishimura, A.L., Oliveira, J.R., Mitne-Neto, M., Guindalini, C., Nitrini, R., Bahia, V.S., de Brito-Marques, P.R., Otto, P.A., Zatz, M., 2004. Lack of association between the brain-derived neurotrophin factor (C-270T) polymorphism and late-onset Alzheimer's disease (LOAD) in Brazilian patients. J. Mol. Neurosci. 22, 257–260.
Ott, J., 2001. Neural networks and disease association studies. Am. J. Med. Genet. 105, 60–61.
Parsons, C.G., Stoffler, A., Danysz, W., 2007. Memantine: a NMDA receptor antagonist that improves memory by restoration of homeostasis in the glutamatergic system—too little activation is bad, too much is even worse. Neuropharmacology 53, 699–723.
Paus, T., Bernard, M., Chakravarty, M., Lourdusamy, A., Leonard, G., Perron, M., Pike, B., Richer, L., Schumann, G., Veillette, S., Pausova, Z., 2011. Association between KCTD8 and brain volume as revealed in a genome-wide study. Organization for Human Brain Mapping Conference. [June 2011].
Petersen, R.C., 2000. Aging, mild cognitive impairment, and Alzheimer's disease. Neurol. Clin. 18, 789–806.
Potkin, S.G., Guffanti, G., Lakatos, A., Turner, J.A., Kruggel, F., Fallon, J.H., Saykin, A.J., Orro, A., Lupoli, S., Salvi, E., Weiner, M., Macciardi, F., 2009a. Hippocampal atrophy as a quantitative trait in a genome-wide association study identifying novel susceptibility genes for Alzheimer's disease. PLoS One 4, e6501.
Potkin, S.G., Turner, J.A., Guffanti, G., Lakatos, A., Fallon, J.H., Nguyen, D.D., Mathalon, D., Ford, J., Lauriello, J., Macciardi, F., 2009b. A genome-wide association study of schizophrenia using brain activation as a quantitative phenotype. Schizophr. Bull. 35, 96–108.
Potkin, S.G., Turner, J.A., Guffanti, G., Lakatos, A., Torri, F., Keator, D.B., Macciardi, F., 2009c. Genome-wide strategies for discovering genetic influences on cognition and cognitive disorders: methodological considerations. Cogn. Neuropsychiatry 14, 391–418.
Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., de Bakker, P.I., Daly, M.J., Sham, P.C., 2007. PLINK: a tool set for whole-genome association and population-based linkage analyses. Am. J. Hum. Genet. 81, 559–575.
Rabionet, R., Jaworski, J.M., Ashley-Koch, A.E., Martin, E.R., Sutcliffe, J.S., Haines, J.L., Delong, G.R., Abramson, R.K., Wright, H.H., Cuccaro, M.L., Gilbert, J.R., Pericak-Vance, M.A., 2004. Analysis of the autism chromosome 2 linkage region: GAD1 and other candidate genes. Neurosci. Lett. 372, 209–214.
Rajagopalan, P., Jahanshad, N., Chiang, M.C., Stein, J.L., Hibar, D.P., Ryles, A., McMahon, K.L., de Zubicaray, G.I., Martin, N.M., Wright, M.J., Saykin, A.J., Jack Jr., C.R., Weiner, M.W., Toga, A.W., Thompson, P.M., and the Alzheimer's Disease Neuroimaging Initiative, 2011. Folate gene variant is associated with brain volume differences: replication in ADNI (N = 740) and Queensland Twins (N = 577). Organization for Human Brain Mapping Conference. [June 2011].
Ramirez-Lorca, R., Boada, M., Saez, M.E., Hernandez, I., Mauleon, A., Rosende-Roca, M., Martinez-Lage, P., Gutierrez, M., Real, L.M., Lopez-Arrieta, J., Gayan, J., Antunez, C., Gonzalez-Perez, A., Tarraga, L., Ruiz, A., 2009. GAB2 gene does not modify the risk of Alzheimer's disease in Spanish APOE 4 carriers. J. Nutr. Health Aging 13, 214–219.
Ramoz, N., Reichert, J.G., Smith, C.J., Silverman, J.M., Bespalova, I.N., Davis, K.L., Buxbaum, J.D., 2004. Linkage and association of the mitochondrial aspartate/glutamate carrier SLC25A12 gene with autism. Am. J. Psychiatry 161, 662–669.
Reich, D.E., Lander, E.S., 2001. On the allelic spectrum of human disease. Trends Genet. 17, 502–510.
Reiman, E.M., Webster, J.A., Myers, A.J., Hardy, J., Dunckley, T., Zismann, V.L., Joshipura, K.D., Pearson, J.V., Hu-Lince, D., Huentelman, M.J., Craig, D.W., Coon, K.D., Liang, W.S., Herbert, R.H., Beach, T., Rohrer, K.C., Zhao, A.S., Leung, D., Bryden, L., Marlowe, L., Kaleem, M., Mastroeni, D., Grover, A., Heward, C.B., Ravid, R., Rogers, J., Hutton, M.L., Melquist, S., Petersen, R.C., Alexander, G.E., Caselli, R.J., Kukull, W., Papassotiropoulos, A., Stephan, D.A., 2007. GAB2 alleles modify Alzheimer's risk in APOE epsilon4 carriers. Neuron 54, 713–720.
Saykin, A.J., Shen, L., Foroud, T.M., Potkin, S.G., Swaminathan, S., Kim, S., Risacher, S.L., Nho, K., Huentelman, M.J., Craig, D.W., Thompson, P.M., Stein, J.L., Moore, J.H., Farrer, L.A., Green, R.C., Bertram, L., Jack Jr., C.R., Weiner, M.W., 2010. Alzheimer's Disease Neuroimaging Initiative biomarkers as quantitative phenotypes: genetics core aims, progress, and plans. Alzheimers Dement. 6, 265–273.
Schaid, D.J., 2004. Evaluating associations of haplotypes with traits. Genet. Epidemiol. 27, 348–364.
Schjeide, B.M., Hooli, B., Parkinson, M., Hogan, M.F., DiVito, J., Mullin, K., Blacker, D., Tanzi, R.E., Bertram, L., 2009. GAB2 as an Alzheimer disease susceptibility gene: follow-up of genomewide association results. Arch. Neurol. 66, 250–254.

Shen, L., Kim, S., Risacher, S.L., Nho, K., Swaminathan, S., West, J.D., Foroud, T., Pankratz, N., Moore, J.H., Sloan, C.D., Huentelman, M.J., Craig, D.W., Dechairo, B.M., Potkin, S.G., Jack Jr., C.R., Weiner, M.W., Saykin, A.J., 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, 1051–1063.

Sled, J.G., Zijdenbos, A.P., Evans, A.C., 1998. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. IEEE Trans. Med. Imaging 17, 87–97.

Stein, J.L., Hua, X., Lee, S., Ho, A.J., Leow, A.D., Toga, A.W., Saykin, A.J., Shen, L., Foroud, T., Pankratz, N., Huentelman, M.J., Craig, D.W., Gerber, J.D., Allen, A.N., Corneveaux, J.J., Dechairo, B.M., Potkin, S.G., Weiner, M.W., Thompson, P., and the Alzheimer's Disease Neuroimaging Initiative, 2010a. Voxelwise genome-wide association study (vGWAS). Neuroimage 53, 1160–1174.

Stein, J.L., Hua, X., Morra, J.H., Lee, S., Hibar, D.P., Ho, A.J., Leow, A.D., Toga, A.W., Sul, J.H., Kang, H.M., Eskin, E., Saykin, A.J., Shen, L., Foroud, T., Pankratz, N., Huentelman, M.J., Craig, D.W., Gerber, J.D., Allen, A.N., Corneveaux, J.J., Stephan, D.A., Webster, J., DeChairo, B.M., Potkin, S.G., Jack Jr., C.R., Weiner, M.W., Thompson, P.M., and the Alzheimer's Disease Neuroimaging Initiative, 2010b. Genome-wide analysis reveals novel genes influencing temporal lobe structure with relevance to neurodegeneration in Alzheimer's disease. Neuroimage 51, 542–554.

Storey, J.D., 2003. The positive false discovery rate: a Bayesian interpretation and the q-value. Ann. Stat. 31, 2013–2035.

Sun, Y.V., Shedden, K.A., Zhu, J., Choi, N.H., Kardia, S.L., 2009. Identification of correlated genetic variants jointly associated with rheumatoid arthritis using ridge regression. BMC Proc. 3 (Suppl 7), S67.

Telliez, J.B., Bean, K.M., Lin, L.L., 2000. LRDD, a novel leucine rich repeat and death domain containing protein. Biochim. Biophys. Acta 1478, 280–288.

Thode, H.C., 2002. Testing for Normality. Marcel Dekker, New York.

Thompson, P.M., Martin, N.G., Wright, M.J., 2010. Imaging genomics. Curr. Opin. Neurol. 23, 368–373.

Tsai, S.J., Hong, C.J., Liu, H.C., Liu, T.Y., Hsu, L.E., Lin, C.H., 2004. Association analysis of brain-derived neurotrophic factor Val66Met polymorphisms with Alzheimer's disease and age of onset. Neuropsychobiology 49, 10–12.

Vepsalainen, S., Castren, E., Helisalmi, S., Iivonen, S., Mannermaa, A., Lehtovirta, M., Hanninen, T., Soininen, H., Hiltunen, M., 2005. Genetic analysis of BDNF and TrkB gene polymorphisms in Alzheimer's disease. J. Neurol. 252, 423–428.

Vounou, M., Nichols, T.E., Montana, G., 2010. Discovering genetic associations with high-dimensional neuroimaging phenotypes: a sparse reduced-rank regression approach. Neuroimage 53, 1147–1159.

Wang, K., Abbott, D., 2008. A principal components regression approach to multilocus genetic association studies. Genet. Epidemiol. 32, 108–118.

Wang, T., Elston, R.C., 2007. Improved power by use of a weighted score test for linkage disequilibrium mapping. Am. J. Hum. Genet. 80, 353–360.

Wang, W.Y., Barratt, B.J., Clayton, D.G., Todd, J.A., 2005. Genome-wide association studies: theoretical and practical concerns. Nat. Rev. Genet. 6, 109–118.

Wellcome Trust Case Control Consortium, 2007. Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. Nature 447, 661–678.

Wheeler, D.L., Barrett, T., Benson, D.A., Bryant, S.H., Canese, K., Chetvernin, V., Church, D.M., DiCuccio, M., Edgar, R., Federhen, S., Feolo, M., Geer, L.Y., Helmberg, W., Kapustin, Y., Khovayko, O., Landsman, D., Lipman, D.J., Madden, T.L., Maglott, D.R., Miller, V., Ostell, J., Pruitt, K.D., Schuler, G.D., Shumway, M., Sequeira, E., Sherry, S.T., Sirotkin, K., Souvorov, A., Starchenko, G., Tatusov, R.L., Tatusova, T.A., Wagner, L., Yaschenko, E., 2008. Database resources of the national center for biotechnology information. Nucleic Acids Res. 36, D13–D21.

Wu, C.W., Kao, H.L., Li, A.F.Y., Chi, C.W., Lin, W.C., 2006. Protein tyrosine-phosphatase expression profiling in gastric cancer tissues. Cancer Lett. 242, 95–103.

Wu, T.T., Chen, Y.F., Hastie, T., Sobel, E., Lange, K., 2009. Genome-wide association analysis by lasso penalized logistic regression. Bioinformatics 25, 714–721.

Yu, W.H., Fraser, P.E., 2001. S100 beta interaction with tau is promoted by zinc and inhibited by hyperphosphorylation in Alzheimer's disease. J. Neurosci. 21, 2240–2246.

Zou, H., 2006. The adaptive lasso and its oracle properties. J. Am. Stat. Assoc. 101, 1418–1429.

2.4 Alzheimer's disease risk gene in young healthy twins

This section is adapted from:

Derrek P. Hibar, Neda Jahanshad, Jason L. Stein, Omid Kohannim, Arthur W. Toga, Sarah E. Medland, Narelle K. Hansell, Katie L. McMahon, Greig I. de Zubicaray, Grant W. Montgomery, Nicholas G. Martin, Margaret J. Wright, Paul M. Thompson. Alzheimer's Disease Risk Gene, GAB2, is Associated with Regional Brain Volume Differences in 755 Young Healthy Twins. Twin Research and Human Genetics. 2012. 15(3):286-295.

Twin Research and Human Genetics
Volume 15 ▌ Number 3 ▌ pp. 286–295 ▌ © The Authors 2012 ▌ doi:10.1017/thg.2012.15

Alzheimer's Disease Risk Gene, *GAB2*, is Associated with Regional Brain Volume Differences in 755 Young Healthy Twins

Derrek P. Hibar,[1] Neda Jahanshad,[1] Jason L. Stein,[1] Omid Kohannim,[1] Arthur W. Toga,[1] Sarah E. Medland,[2,3,4] Narelle K. Hansell,[2] Katie L. McMahon,[5] Greig I. de Zubicaray,[6] Grant W. Montgomery,[2] Nicholas G. Martin,[2] Margaret J. Wright,[2] and Paul M. Thompson[1]

[1]Laboratory of Neuro Imaging, Department of Neurology, UCLA School of Medicine, Los Angeles, USA
[2]Genetic Epidemiology Laboratory, Queensland Institute of Medical Research, Brisbane, Australia
[3]Quantitative Genetics Laboratory, Queensland Institute of Medical Research, Brisbane, Australia
[4]Broad Institute of Harvard and MIT, Boston, USA
[5]Centre for Advanced Imaging, University of Queensland, Brisbane, Australia
[6]Functional Magnetic Resonance Imaging Laboratory, School of Psychology, University of Queensland, Brisbane, Australia

The development of late-onset Alzheimer's disease (LOAD) is under strong genetic control and there is great interest in the genetic variants that confer increased risk. The Alzheimer's disease risk gene, *growth factor receptor bound protein 2-associated protein* (*GAB2*), has been shown to provide a 1.27–1.51 increased odds of developing LOAD for rs7101429 major allele carriers, in case-control analysis. *GAB2* is expressed across the brain throughout life, and its role in LOAD pathology is well understood. Recent studies have begun to examine the effect of genetic variation in the *GAB2* gene on differences in the brain. However, the effect of *GAB2* on the young adult brain has yet to be considered. Here we found a significant association between the *GAB2* gene and morphological brain differences in 755 young adult twins (469 females) ($M = 23.1$, $SD = 3.1$ years), using a gene-based test with principal components regression (PCReg). Detectable differences in brain morphology are therefore associated with variation in the *GAB2* gene, even in young adults, long before the typical age of onset of Alzheimer's disease.

■ **Keywords:** *GAB2*, imaging genetics, tensor-based morphometry, Alzheimer's disease

Numerous heritability studies show that brain structure is under moderately strong genetic control (Kremen et al., 2010; Peper et al., 2007; Thompson et al., 2001). However, very few genetic variants have been identified that reliably explain a significant proportion of brain variation in human populations. Several researchers advocate the use of brain measures to empower the search for disease risk genes (Bis et al., 2012; Meyer-Lindenberg & Weinberger, 2006; Stein et al., 2012). If specific genetic variants can be related to brain differences, they may offer new molecular targets for drug development, and a deeper understanding of disease susceptibility and treatment response. For example, young adult carriers of a recently discovered risk allele for Alzheimer's disease (AD), *CLU*-C, have detectable differences in brain integrity approximately 50 years before the typical age of AD onset (Braskie et al., 2011), as do carriers of a common variant in an iron-overload gene, *HFE* (Jahanshad et al., 2012). Brain measures related to genetic liability can be helpful for studying factors that may avert or promote disease (Gogtay et al., in press), and for defining the biological spectrum underlying psychiatric disorders (Akil et al., 2010). For instance, carriers of the common Val66Met polymorphism in the brain-derived neurotrophic factor (*BDNF*) gene are at heightened risk for bipolar disorder (Fernandes et al., 2011) and schizophrenia (Green et al., 2011). They show detectable differences in white matter microstructure (Chiang et al., 2010),

RECEIVED 1 February 2012; ACCEPTED 26 March 2012.

ADDRESS FOR CORRESPONDENCE: Dr Paul Thompson, Professor of Neurology, Laboratory of Neuro Imaging, Dept. of Neurology, UCLA School of Medicine, Neuroscience Research Building 225E, 635 Charles Young Drive, Los Angeles, CA 90095-1769, USA. E-mail: thompson@loni.ucla.edu

hippocampal volume (Hajek et al., 2012), and prefrontal cortex morphology (Pezawas et al., 2004).

The success of these candidate gene studies has led to increased interest in using neuroimaging endophenotypes to study genetic determinants of brain disorders (Hibar et al., 2011a; Thompson et al., 2010). Recently, there has been increased interest in an AD risk gene, *growth factor receptor bound protein 2-associated protein* or *GAB2*. The *GAB2* gene is well characterized as a risk gene for the development of late-onset AD (Chapuis et al., 2008; Lin et al., 2010; Ramirez-Lorca et al., 2009; Reiman et al., 2007). Large meta-analyses confirm that it has a moderate effect on disease risk with an odds ratio of 1.27–1.51, using a 95% confidence interval (Ikram et al., 2009; Schjeide et al., 2009). There is convincing evidence that *GAB2* is expressed throughout the brain and throughout life (Reiman et al., 2007; Trollmann et al., 2010). In our previous study of 731 elderly subjects from the Alzheimer's Disease Neuroimaging Initiative (ADNI), we identified the *GAB2* gene as significantly associated with morphometric brain volume differences using gene-based tests (Hibar et al., 2011b). We showed also that gene-based tests — based on principal component regression (PCReg) encoding the set of single nucleotide polymorphisms (SNPs) in a gene — have more power to detect significant associations than standard univariate statistical methods, in certain cases. The association of *GAB2* with morphometric brain volume differences was undetectable with traditional univariate methods, but showed significant effects in the brain when the cumulative evidence of variation across the gene was incorporated into a gene-based association statistic. Another prior analysis of *GAB2* in a separate, independent dataset found that a protective *GAB2* haplotypic variant is associated with increased glucose metabolism in brain regions typically affected in AD (including the left temporal lobe, right frontal lobe, bilateral parietal lobes, and the precuneus) measured with fluorodeoxyglucose positron emission tomography (FDG-PET) (Liang et al., 2011). In addition, Liang et al. (2011) showed, when studying *APOE*4 allele carriers specifically, that carriers have increased glucose metabolism if they also carry the protective *GAB2* haplotype. Lower levels of glucose metabolism in the brain tissue of AD and mild cognitive impairment (MCI) patients may reflect a decrease in neuronal density, as well as abnormal metabolism (Magistretti & Pellerin, 1996; Mark et al., 1997; Piert et al., 1996). While the Liang et al. (2011) study focused on haplotype blocks, certain gene-based tests, such as PCReg performed by Hibar et al., (2011b), may be considered comparable, as the *GAB2* gene lies in a single linkage-disequilibrium (LD) block (Reiman et al., 2007). This replication of *GAB2* effects in a separate sample and imaging modality lends credibility to *GAB2* as a risk gene that influences the brain. It is then reasonable to suspect that *GAB2* gene variants have significant, observable effects on brain morphology, perhaps even in early adulthood. However, no studies have examined how *GAB2* relates to brain structure in young adults, long before the onset of neurodegeneration. In this study, we hypothesized that gene-based tests in the AD-risk gene, *GAB2*, would reveal significant effects on brain morphology in young adults.

Methods

Subjects

A total of 755 healthy, young adult (mean = 23.1, SD = 3.1 years) twins and their siblings (469 females) from 439 families (294 dizygotic twins, 222 monozygotic twins, 143 singletons, three dizygotic trios, and 87 siblings) were examined in this study. All subjects had standard T_1-weighted MRI brain scans and genome-wide genotyping information available. All subjects were of European ancestry and recruited as part of the Queensland Twin Imaging Study (QTIM) in Australia. The QTIM study is an ongoing five-year longitudinal study of healthy young twins with structural and functional MRI, diffusion tensor imaging (DTI), genetics, and comprehensive cognitive assessments (de Zubicaray et al., 2008). Subjects were excluded if they reported any pathologies known to affect the brain, head injuries, or major illnesses. All subjects were right-handed, as determined by the 12-item Annett's handedness questionnaire (Annett, 1970). Informed consent was obtained from each subject and the study was approved by the institutional review boards of the University of California, Los Angeles, and the Queensland Institute of Medical Research.

Genotyping and Quality Control Filtering

Genome-wide genotype data were collected on the Human610-Quad BeadChip (Illumina, Inc., San Diego, CA). Several SNPs were dropped from the dataset based on standard quality control filtering measures used in other large GWAS analyses (Wellcome Trust Case Control Consortium, 2007). Individual SNPs were removed based on the following criteria: call rate < 95% (8,447 SNPs removed), minor allele frequency < .01 (33,347 SNPs removed), significant deviation from Hardy-Weinberg equilibrium $p < 1 \times 10^{-6}$ (2,841 SNPs removed), autosomal chromosomes only, and a platform-specific score of .07 to eliminate blank genotype calls (results in a variable number of missing genotypes for each SNP).

After filtering, the hard genotype calls were imputed to the HapMap Phase II release 21 reference dataset (Altshuler et al., 2010). Imputation was performed using the robust and freely available software, MaCH (Abecasis et al., 2010). The imputed genotypes were further filtered based on minor allele frequency < .01 (38,481 SNPs removed) and R^2 < .3 (54,337 SNPs removed). After imputation, 2,439,807 SNPs passed all quality control criteria.

In this study, we were only interested in SNPs found in the candidate gene, *GAB2*, as it was strongly implicated in our prior study of a different dataset (Hibar et al., 2011b). We used the gene annotation function in the KGG software

package (Li et al., 2011; Li et al., 2010) to select SNPs from our set of imputed genotypes in the *GAB2* gene group. SNPs within 50 Kb upstream or downstream of the *GAB2* gene border were included in the final group of selected SNPs.

Image Acquisition and Processing

High resolution structural MRI scans were obtained for each subject on a single 4-Tesla scanner (Bruker Medspec). T_1-weighted images were acquired with an inversion recovery rapid gradient echo sequence (TI/TR/TE = 700/1500/3.35 ms; flip angle = 8°; slice thickness = 0.9 mm, with a 256^2 acquisition matrix).

Non-brain regions were removed from the T_1-weighted scans using Robex (Iglesias et al., 2011) specifically trained on manually edited skull-stripped images. Next, scans were corrected for image field non-uniformity using FreeSurfer (http://surfer.nmr.mgh.harvard.edu/), and linearly aligned to a common template using a 9-parameter model (Holmes et al., 1998).

3D Maps of Morphometric Brain Differences

The minimum deformation template (MDT) represents a nonlinear average of anatomical differences throughout the brain, and is used as a reference template to help compare the brain structure of subjects in the study. Using the MDT as a target, 3D 'Jacobian' maps of regional brain volume differences were generated for each subject, with a nonlinear, inverse consistent registration algorithm (Leow et al., 2005), and then downsampled to a $2 \times 2 \times 2$ mm isometric voxel size. Each voxel value in the 3D maps measures the difference in volume between the subject's image and the template, based on the gradient of the deformation field required to deform the subject's scan to match the common template. As subjects are all registered to the same MDT, we can examine regional morphometric differences in brain volume by analyzing the determinant of the Jacobian matrix at the same voxel in each subject. In this study, we performed a whole-brain analysis of the volume differences from these 3D morphometric maps for each voxel in the brain and across all subjects.

Gene-Based Tests

Tests of association of the *GAB2* gene with whole-brain, voxel-wise regional brain volume differences were conducted inside the whole-brain mask of the MDT, using PCReg (Wang & Abbott, 2008). PCReg test statistics are generated by first performing principal components analysis (PCA) on a set of n SNPs across our sample of subjects. The PCA outputs a set of orthogonal eigenvectors that represent the variance components from the set of SNPs. This transformation is defined so that the first principal component has the largest possible variance (accounts for as much as possible of the variability in the data), and each succeeding component in turn accounts for the highest variance possible under the constraint that it be orthogonal to (linearly uncorrelated with) the preceding components. Eigenvectors are then selected (in descending order of the amount of variance explained) until they explain at least 95% of the total variance in the original SNP set. These PCA steps provide an efficient means for dimension reduction but, more importantly, they provide a means to test the cumulative evidence for association across a full gene without being vulnerable to problems due to multicollinearity. Association tests with PCReg use a multiple partial-F test framework, which works by fitting two separate regression models ('full' and 'reduced') such that the full model contains a set of eigenvectors with age and sex covariates as regressors, and morphometric volume values as the dependent variable. The variance explained by the full model is then compared to the variance explained by a reduced model that fits the age and sex covariates as regressors, with the same morphometric volume values as the dependent variable. An F statistic and p-value are assigned to each voxel, based on the amount of additional variance explained by the full model compared to the reduced model. Our PCReg tests were implemented using the Efficient Mixed-Model Association (EMMA) software package (Kang et al., 2008). EMMA uses a kinship matrix to control for relatedness in family-based and twin samples, such as this one, using a mixed-effects model.

Multiple Comparisons Correction

Performing a large number of statistical tests at multiple voxels across the brain increases the potential of identifying false-positive findings (i.e., inflating Type I errors). To correct for the total of number of tests performed, and appropriately control Type I errors, we employed the searchlight false discovery rate procedure (searchlight-FDR; Langers et al., 2007). Searchlight-FDR controls for the regional significance of test statistics under the null hypothesis, and incorporates information on the spatial extent of the statistical effects and the smoothness of the underlying image. The searchlight-FDR procedure implemented in this study thresholds statistical maps at the standard $q = .05$ false-positive rate. Correction for the number of 'eigen-SNPs' included in the model is not required, as gene-based tests with PCReg evaluate association with a multiple partial-F test that includes the full set of eigen-SNPs in a single statistical test.

Results

The *GAB2* Gene and PCReg

After all quality control filtering, imputation, and annotation, there were 51 SNPs in the *GAB2* gene grouping. Positional and functional annotation data for each SNP in the analysis are shown in Table 1. After applying PCA to the set of 51 SNPs, we found that the first 10 principal components were sufficient to explain 95% of the total variance in

TABLE 1

Gene Annotation Results for the *GAB2* Gene using KGG Software[a]

SNP	Position	Gene feature	Conservation score
rs10899500	78125246	intron	0.071
rs10899496	78123831	intron	0.009
rs2063724	78133077	-	0.008
rs2511175	77975081	intron	0.075
rs1981405	77976208	intron	0.002
rs948662	77979829	intron	0.202
rs2511170	77980582	intron	0.11
rs10793302	78040961	intron	0.069
rs7115850	78045071	intron	0.001
rs10899488	78088754	intron	0
rs1893447	77973182	intron	0.023
rs2450130	77943457	intron	0.551
rs1017908	77948095	intron	0.128
rs1893448	77969180	intron	0.102
rs2510038	77966034	intron	0.014
rs2510054	77959659	intron	0.005
rs731600	77963133	intron	0.006
rs1007837	77941076	intron	0.001
rs2248407	77937800	coding-synonymous	0.005
rs1385600	77936166	coding-synonymous	0.813
rs1385601	77936064	intron	0.004
rs1318241	77930792	intron	0.063
rs901104	77930499	intron	0.283
rs6592772	78015563	intron	0.008
rs7939646	78111613	intron	0.014
rs4944196	78008731	intron	0.026
rs4945261	77990260	intron	0.005
rs7101429	77992967	intron	0.021
rs10899456	78000155	intron	0.006
rs4291702	78001248	intron	0.001
rs4944195	78003499	intron	0.728
rs11602622	78010830	intron	0.047
rs2450129	77940385	intron	0.459
rs10899469	78018313	intron	0.286
rs4945265	78027458	intron	0.782
rs10899485	78072383	intron	0.003
rs10501426	78057122	intron	0.296
rs2292573	78053139	intron	0.017
rs2292572	78052864	utr-5	0.001
rs2373115	78091150	intron	0.002
rs10899489	78095373	intron	0.016
rs7112234	78102470	intron	0.017
rs1046780	77926769	utr-3	0
rs866901	77926309	near-gene-3	0.003
rs2912	77926292	near-gene-3	0
rs7927923	77979414	intron	0.001
rs3740677	77928036	utr-3	0.025
rs10793294	77996403	intron	0.879
rs11237451	78025459	intron	0.233
rs2450135	77927995	utr-3	0.049
rs11601726	78068039	intron	0.002

Note: SNP = single nucleotide polymorphism. The starting position of *GAB2* (Entrez Gene ID: 9846) on chromosome 11 is 77916336 (base pairs) with a total length of 222532 (base pairs). For each of the 51 SNPs in our analysis, we give the positional information in base pairs, annotation of possible SNP function, and a conservation score from the UCSC Genome Browser (http://hgdownload.cse.ucsc.edu/).

[a] see Li et al. (2011) & Li et al. (2010)

the full SNP set. We used this set of 10 eigen-SNPs in the PCReg association tests at each voxel across the full brain.

Statistical Parametric Maps

The significant regions of the searchlight-FDR adjusted *p*-map for the whole-brain *GAB2* association tests is shown in Figure 1. In total, seven clusters survived the searchlight-FDR correction (here we use cluster to mean sets of contiguous voxels, even though we used a voxel-wise correction method). The associated regions after correction for multiple comparisons are summarized in Table 2.

TABLE 2

A Summary of Clusters Significantly Associated with *GAB2* after Correction for Multiple Comparisons Using Searchlight FDR

Cluster #	$MNI_{(x)}$	$MNI_{(y)}$	$MNI_{(z)}$	Size (mm^3)	Minimum corrected *p*-value
1	24	−73	42	10696	.0387 (.0019)
2	48	−58	48	2489	.0387 (.0017)
3	−48	−34	20	1884	.0387 (.0016)
4	25	41	44	1164	.0408 (.004)
5	−21	−75	−49	694	.0387 (.0014)
6	67	−24	−3	690	.0387 (.0018)
7	18	−24	−21	117	.0470 (.0045)

Note: MNI = Montreal Neurological Institute. Here we define 'cluster' to mean any set of significant contiguous voxels after correction for multiple comparisons. The location of the most highly associated voxel in each cluster is given in MNI coordinates: $MNI_{(x,y,z)}$. The size or extent of each cluster is given in mm^3. The minimum corrected *p*-value is the most highly associated voxel in a given cluster in the searchlight-FDR corrected map; uncorrected *p*-values for that voxel are given in parentheses.

Post Hoc Analysis

The multiple partial *F*-test yields *F*-test statistics that are non-directional (in other words, they do not tell you the 'direction' of an effect). So, for our primary analysis, we report only *p*-value statistics for each voxel that passes regional FDR correction, using searchlight-FDR. These *p*-values only indicate that there was a significant association between *GAB2* and volume differences in the cohort at a given voxel, not its direction. To get a better idea of the direction of an effect, we performed a *post hoc* analysis using the first principal component from the PCA of *GAB2* SNPs as a regressor, along with age and sex covariates, in a mixed-effect regression (MER) model (as implemented in EMMA), at each point across the brain. In this way, we were able to obtain approximate directional effects of the *GAB2* gene (i.e., those implied by the first principal component) on brain volume differences using the *Beta* coefficients from the MER model (Figure 2).

Discussion

Here we found a significant association between morphological differences in young-adult brains and variation in the *GAB2* gene. These results augment our previous work with the elderly ADNI sample that demonstrated for the first time that variation in the *GAB2* gene was associated with observable morphological differences in the brain. Our study of the QTIM sample expands on the literature that has established *GAB2* as an AD risk gene. Prior studies have not considered how this gene might affect the brain long before the typical age of onset of the disease. We found significant differences in regional brain volume in a large region of the right parietal lobe, and additional smaller significant clusters in the left parietal, and along the tissue/cerebral spinal fluid (CSF) boundary, in the temporal

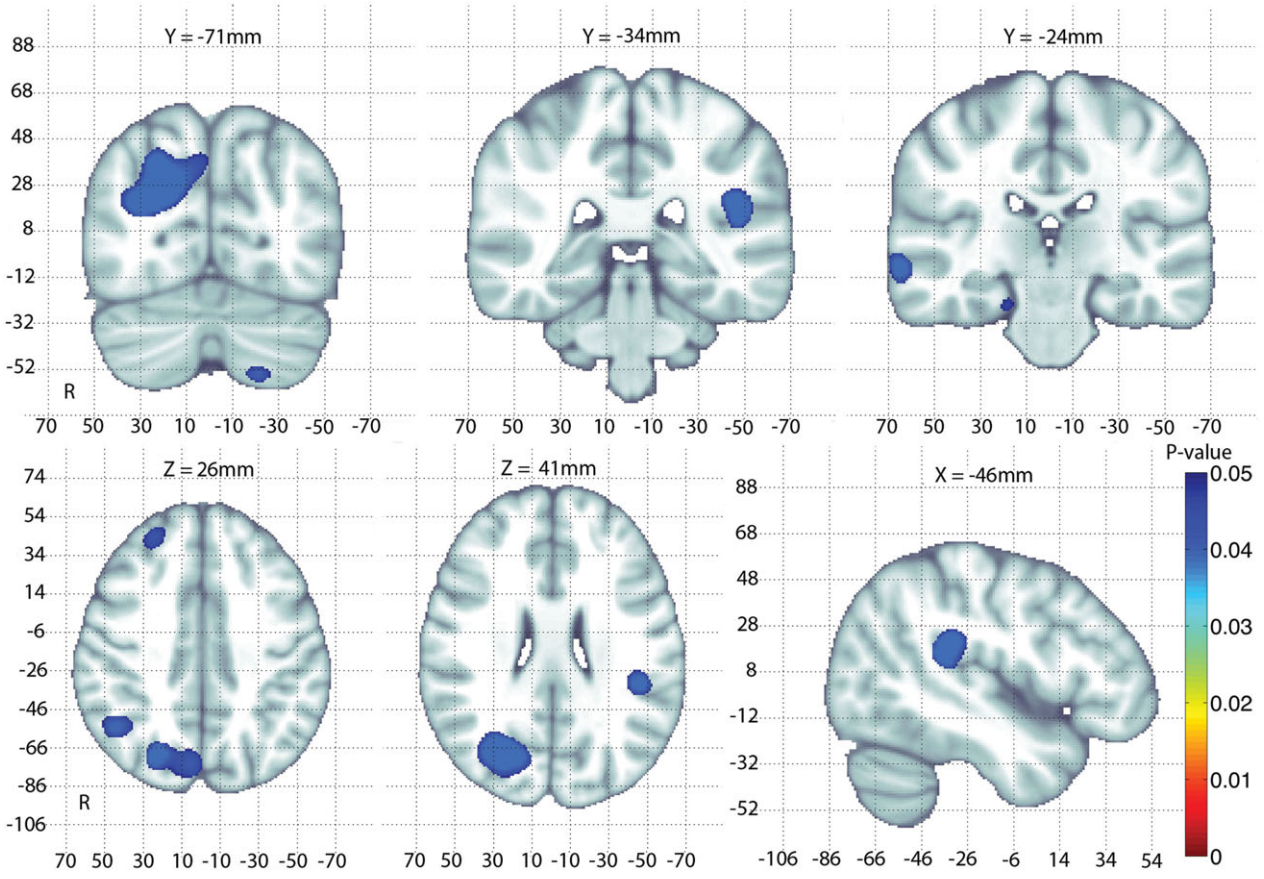


FIGURE 1

An adjusted *p*-map of *GAB2* association tests after correcting for multiple comparisons using searchlight FDR. Adjusted *p*-values < .05 are considered significant and are overlaid on the MNI-152 T1 template for anatomical reference. The largest clusters of significant associations of *GAB2* with morphological differences occur in the right parietal lobe.

lobes of the young-adult twins. A *post hoc* analysis was also performed to determine the direction of the *GAB2* gene effects on brain morphology. There was a negative correlation between overall tissue volume and *GAB2* loading in the right parietal lobe region. In addition, we found a positive correlation between *GAB2* loading and CSF volume along the sulci of the temporal lobes (partial volume effects along the tissue/CSF boundary means that voxels along the border usually represent the effect of CSF volume change), as shown in Figure 2.

Function of *GAB2*

GAB2 encodes a human adapter protein that acts in a number of cell proliferation pathways, especially in endothelial cells (Zhang & Broxmeyer, 2000) and is expressed throughout the human brain (Zhao et al., 1999). The mechanism by which the *GAB2* gene leads to increased neurodegeneration is well understood. Reiman et al. (2007) showed that normally functioning *GAB2* protein is responsible for suppressing the phosphorylation of *tau* tangles associated with the development of Alzheimer's disease. Using small-interfering RNA (siRNA) to knock-down *GAB2* function, Reiman et al. also demonstrated a significant increase in *tau* tangle formation in neuronal tissue with the *GAB2* protein function diminished. While the role of *GAB2* in the pathway that leads to the formation of *tau* tangles is well understood, there is scant evidence in the literature about functional mechanisms and pathways through which changes in the efficiency of the *GAB2* protein might lead to morphological differences in young-adult brains.

Knowing that an AD risk gene such as *GAB2* may have a detectable effect in the young adult brain is useful, as it may indicate a developmental vulnerability to AD, without directly promoting AD pathology (e.g., amyloid plaque and neurofibrillary tangle formation). This is the most likely scenario as the young-adult brain is fairly low in amyloid deposits (Bartzokis, 2011; Braak & Braak, 1997). Even so, recent studies of some other AD risk genes have also revealed detectable effects in young adults. Shaw et al. (2007) showed that adolescent *APOE4* allele carriers have a thinner cortex and slower cortical thickening than their *APOE2/3* counterparts. A genetic variant (rs11136000) in the AD risk gene,

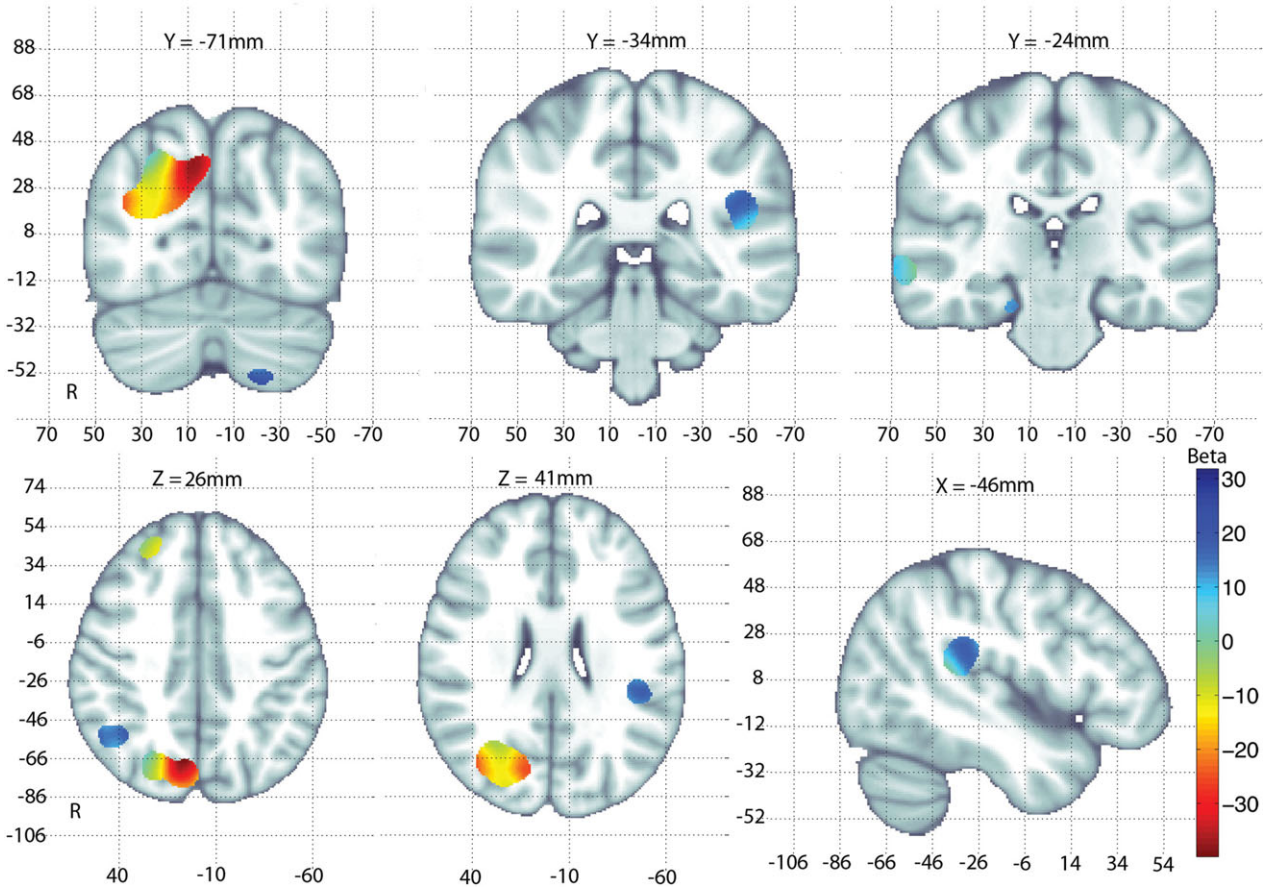


FIGURE 2

Beta coefficient values are shown for the first principal component of the *GAB2* SNPs used in a multiple linear regression with age and sex as covariates. Only regions that were significant after correction for multiple comparisons using searchlight FDR are shown. The coefficient maps are overlaid on the MNI-152 T1 template for anatomical reference. Negative Beta values (warmer colors) in the tissue of the parietal lobe indicate a negative relationship of tissue volume and *GAB2* loading. Positive Beta values (cooler colors) in and around the cerebrospinal fluid of the lateral sulcus and temporal lobes indicate a positive relationship of CSF expansion with *GAB2* loading.

CLU, found in 88% of Caucasian subjects, confers nearly a 20% lifetime increase in the risk for AD and is associated with white matter differences in DTI scans of young adults (Braskie et al., 2011). Similarly, the H63D variant in *HFE*, an iron overload gene, is associated with white matter microstructure in young adults and is thought to be associated with AD risk (Jahanshad et al., 2012).

Conclusions and Interpretations

An essential part of endophenotype theory suggests that complex brain disorders may be better described by traits related to a disease, but with simpler genetic determinants than case-control phenotypes (Gottesman & Gould, 2003). Neuroimaging endophenotypes are now widely used to assess genetic contributions to complex neurodegenerative disorders such as late-onset Alzheimer's disease (LOAD). Given the cost and difficulty of collecting large neuroimaging datasets, it can be challenging to collect enough data for unbiased gene discovery methods such as GWAS (de Geus, 2010). Very large neuroimaging genetics consortia are only just beginning to aggregate samples large enough to discover and replicate GWAS findings from imaging studies (ENIGMA; http://enigma.loni.ucla.edu; Stein et al., 2012). Even so, hypothesis-driven candidate gene studies may be performed in smaller datasets, without such a heavy correction for the number of SNPs assessed. Several such studies of disease risk genes (e.g., *CLU*, Braskie et al., 2011; and *HFE*, Jahanshad et al., 2012) have revealed how variants in these genes affect the brain, offering a plausible mechanistic explanation of how they may promote risk for neurological disease. Endophenotypes can help to localize gene effects, revealing functional and anatomical differences between carriers of different variants; they may also help to classify disease subtypes that may not be apparent in case-control studies (de Geus, 2010).

The current study shows detectable associations between the AD risk gene *GAB2* and morphological brain differences in young adults. Our findings build on our previous work

in a separate cohort of elderly subjects that shows significant morphological differences associated with variation in the *GAB2* gene (Hibar et al., 2011b). While we report significant associations in the right parietal lobe, lateral sulcus, and bilateral temporal lobes, none of the voxels overlap with the voxels that survived statistical thresholding in the original ADNI sample. In this analysis, we performed an unbiased search across each voxel of the full brain without incorporating information from our prior tests in the ADNI dataset. The reason for this is that voxels that are significant after correction for multiple comparisons do not necessarily represent the only voxels where *GAB2* may have an effect. We wanted to perform an analysis that allowed for the possibility to observe an effect anywhere in the brain. As such, the most appropriate test is a brain-wide search, as it asserts only that the gene has some effect in the brain, rather than making a stronger assertion about its specific localization. In addition, the QTIM dataset differs substantially from the ADNI dataset (subjects are, on average, around 50 years younger in the QTIM study, and are scanned at a different field strength on a different continent). Given this, the regions where *GAB2* is associated with brain morphology may be very different from those observed in the elderly ADNI sample. From a biological point of view, it may also be that a gene has regional effects on the brain that either spread out or become more regionally specific over the human lifespan. Or, perhaps more likely, it could be that a very weak effect is spread over the entire brain in both samples, but due to noise and anatomical variability in the two cohorts, different locations in the brain contribute to the statistical results in each cohort. As an alternative, more stringent, approach, a conjunction test (Nichols et al., 2005) could be performed to directly identify voxels where a gene effect is statistically significant in several cohorts at once. Such tests have relatively low power, as they require that the gene effect be found consistently in the same voxels in all cohorts, rather than across the brain in aggregate in each cohort. Different formulations of the null hypothesis may therefore greatly affect the power and scope of the inferences. We found a large set of significant voxels in the parietal and temporal lobes that were associated with tissue volume differences. Regions of both the parietal and temporal lobes have been associated with susceptibility for and progression of Alzheimer's disease (Leow et al., 2009). Parietal lobe atrophy (Scahill et al., 2002), impaired white matter microstructure (Bozzali et al., 2002), and glucose hypometabolism (Langbaum et al., 2009), are well-studied effects of Alzheimer's disease progression on the brain. Additionally, Liang et al. (2011) showed reduced glucose metabolism bilaterally in the parietal and temporal lobes of carriers of a *GAB2* haplotypic risk variant. Many of the significant regions from that study appear to overlap with the regions found in this study, especially in the parietal lobes. Volume changes in the temporal lobes, specifically, have been commonly used as a biomarker for tracking the progression of Alzheimer's disease, and also for image-based diagnostic classification (Jack et al., 1998; Jack et al., 1997). Here we found that *GAB2* is associated with differences in regional brain volumes in young adults, in regions strongly implicated in the progression of Alzheimer's disease. This adds support to the notion that *GAB2* may be used to help identify individuals at heightened risk for AD long before the onset of disease.

Further replication of observed effects of the *GAB2* gene in young adults is still required in order to aggregate evidence about true morphological effects. While searching for genetic determinants of morphological brain differences, the interpretation of results should consider not only individual genetic contributions, but potentially correlated alleles as well. In our study, we detected a very mild positive correlation of the 51 SNPs from the *GAB2* gene and the *CLU* Alzheimer's disease risk SNP rs11136000 (Pearson's $r = .07–.08$), though none of the test statistics passed a Bonferroni correction for the number of SNPs tested ($p < 9.8 \times 10^{-4}$). Tests of other major candidate gene SNPs, such as the *BDNF* Val66Met polymorphism (Chiang et al., 2010), and the H63D polymorphism of the *HFE* gene (Jahanshad et al., 2012), showed no evidence of positive correlation. There could, in principle, be correlations between risk alleles, especially in an elderly population, due to a survivor effect. This could arise if adverse variants in a gene were associated with early mortality. Protective variants might then be present together in the survivors more frequently than their random co-occurrence in randomly selected people. Our analysis provides new evidence for an association between morphological differences and variation in the *GAB2* in young adults, long before the onset of AD pathology. We hope that our findings will inform future research on the functional relevance of *GAB2* in neural development.

Acknowledgments

We thank the twins for their participation, Kori Johnson and the radiographers for MRI scanning and preprocessing the images, Marlene Grace and Ann Eldridge for twin recruitment, Daniel Park for database support, Anjali Henders for DNA processing and preparation, Scott Gordon for quality control and management of the genotypes. The QTIM study was supported by the National Institute of Child Health and Human Development (R01 HD050735), and the National Health and Medical Research Council (NHMRC 486682,1009064), Australia. Genotyping was supported by NHMRC (389875). DPH is partially supported by NSF GRFP grant DGE-0707424. OK was supported in part by the UCLA MSTP. JLS is partially supported by a T32 postdoctoral training grant in Neurobehavioral Genetics. NJ was additionally supported by NIH NLM Grant T15 LM07356. GM was supported by an NHMRC Fellowship 613667; GZ was supported by an ARC Future Fellowship.

References

Abecasis, G. R., Li, Y., Willer, C. J., Ding, J., & Scheet, P. (2010). MaCH: Using sequence and genotype data to estimate haplotypes and unobserved genotypes. *Genetic Epidemiology, 34*, 816–834.

Akil, H., Brenner, S., Kandel, E., Kendler, K. S., King, M. C., Scolnick, E., Watson, J. D., & Zoghbi, H. Y. (2010). Medicine. The future of psychiatric research: Genomes and neural circuits. *Science, 327*, 1580–1581.

Altshuler, D. M., Gibbs, R. A., Peltonen, L., Dermitzakis, E., Schaffner, S. F., Yu, F., et al. (2010). Integrating common and rare genetic variation in diverse human populations. *Nature, 467*, 52–58.

Annett, M. (1970). A classification of hand preference by association analysis. *British Journal of Psychology, 61*, 303–321.

Bartzokis, G. (2011). Alzheimer's disease as homeostatic responses to age-related myelin breakdown. *Neurobiology of Aging, 32*, 1341–1371.

Bis, J. C., DeCarli, C. S., Smith, A. V., van der Lijn, F., Crivello, F., Fornage, M., . . . Seshadri, S., for the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium. (2012). Common variants at 12q14 and 12q24 are associated with hippocampal volume. *Nature Genetics, 44*, 545–551.

Bozzali, M., Falini, A., Franceschi, M., Cercignani, M., Zuffi, M., Scotti, G., Comi, G., & Filippi, M. (2002). White matter damage in Alzheimer's disease assessed in vivo using diffusion tensor magnetic resonance imaging. *Journal of Neurology, Neurosurgery, and Psychiatry, 72*, 742–746.

Braak, H., & Braak, E. (1997). Frequency of stages of Alzheimer-related lesions in different age categories. *Neurobiology of Aging, 18*, 351–357.

Braskie, M. N., Jahanshad, N., Stein, J. L., Barysheva, M., McMahon, K. L., de Zubicaray, G. I., Martin, N. G., Wright, M. J., Ringman, J. M., Toga, A. W., & Thompson, P. M. (2011). Common Alzheimer's disease risk variant within the *CLU* gene affects white matter microstructure in young adults. *Journal of Neuroscience, 31*, 6764–6770.

Chapuis, J., Hannequin, D., Pasquier, F., Bentham, P., Brice, A., Leber, I., Frebourg, T., Deleuze, J. F., Cousin, E., Thaker, U., Amouyel, P., Mann, D., Lendon, C., Campion, D., & Lambert, J. C. (2008). Association study of the *GAB2* gene with the risk of developing Alzheimer's disease. *Neurobiology of Disease, 30*, 103–106.

Chiang, M. C., Barysheva, M., Toga, A. W., Medland, S. E., Hansell, N. K., James, M. R., McMahon, K. L., de Zubicaray, G. I., Martin, N. G., Wright, M. J., & Thompson, P. M. (2010). *BDNF* gene effects on brain circuitry replicated in 455 twins. *Neuroimage, 15*, 448–454.

de Geus, E. J. (2010). From genotype to EEG endophenotype: A route for post-genomic understanding of complex psychiatric disease? *Genome Medicine, 2*, 63.

de Zubicaray, G. I., Chiang, M. C., McMahon, K. L., Shattuck, D. W., Toga, A. W., Martin, N. G., Wright, M. J., & Thompson, P. M. (2008). Meeting the challenges of neuroimaging genetics. *Brain Imaging and Behavior, 2*, 258–263.

Fernandes, B. S., Gama, C. S., Cereser, K. M., Yatham, L. N., Fries, G. R., Colpo, G., de Lucena, D., Kunz, M., Gomes, F. A., & Kapczinski, F. (2011). Brain-derived neurotrophic factor as a state-marker of mood episodes in bipolar disorders: A systematic review and meta-regression analysis. *Journal of Psychiatric Research, 45*, 995–1004.

Gogtay, N., Hua, X., Stidd, R., Boyle, C. P., Lee, S., Weisinger, B., Chavez, A., Giedd, J. N., Clasen, L., Toga, A. W., Rapoport, J. L., & Thompson, P. M. (in press). Delayed white matter growth trajectory in young non-psychotic siblings of childhood-onset schizophrenia patients. *Archives of General Psychiatry*.

Gottesman, I. I., & Gould, T. D. (2003). The endophenotype concept in psychiatry: Etymology and strategic intentions. *The American Journal of Psychiatry, 160*, 636–645.

Green, M. J., Matheson, S. L., Shepherd, A., Weickert, C. S., & Carr, V. J. (2011). Brain-derived neurotrophic factor levels in schizophrenia: A systematic review with meta-analysis. *Molecular Psychiatry, 16*, 960–972.

Hajek, T., Kopecek, M., & Hoschl, C. (2012). Reduced hippocampal volumes in healthy carriers of brain-derived neurotrophic factor Val66Met polymorphism: Meta-analysis. *World Journal of Biological Psychiatry, 13*, 178–187.

Hibar, D. P., Kohannim, O., Stein, J. L., Chiang, M.-C., & Thompson, P. M. (2011a). Multilocus genetic analysis of brain images. [Review]. *Frontiers in Genetics, 2*, 73.

Hibar, D. P., Stein, J. L., Kohannim, O., Jahanshad, N., Saykin, A. J., Shen, L., Kim, S., Pankratz, N., Foroud, T., Huentelman, M. J., Potkin, S. G., Jack, C. R. Jr., Weiner, M. W., Toga, A. W., Thompson, P. M., & Alzheimer's Disease Neuroimaging Initiative. (2011b). Voxelwise gene-wide association study (vGeneWAS): Multivariate gene-based association testing in 731 elderly subjects. *NeuroImage, 56*, 1875–1891.

Holmes, C. J., Hoge, R., Collins, L., Woods, R., Toga, A. W., & Evans, A. C. (1998). Enhancement of MR images using registration for signal averaging. *Journal of Computer Assisted Tomography, 22*, 324–333.

Iglesias, J. E., Liu, C. Y., Thompson, P. M., & Tu, Z. (2011). Robust brain extraction across datasets and comparison with publicly available methods. *IEEE Transactions on Medical Imaging, 30*, 1617–1634.

Ikram, M. A., Liu, F., Oostra, B. A., Hofman, A., van Duijn, C. M., & Breteler, M. M. B. (2009). The *GAB2* Gene and the Risk of Alzheimer's Disease: Replication and Meta-Analysis. *Biological Psychiatry, 65*, 995–999.

Jack, C. R., Petersen, R. C., Xu, Y., O'Brien, P. C., Smith, G. E., Ivnik, R. J., Tangalos, E. G., & Kokmen, E. (1998). Rate of medial temporal lobe atrophy in typical aging and Alzheimer's disease. *Neurology, 51*, 993–999.

Jack, C. R., Petersen, R. C., Xu, Y. C., Waring, S. C., O'Brien, P. C., Tangalos, E. G., Smith, G. E., Ivnik, R. J., & Kokmen, E. (1997). Medial temporal atrophy on MRI in normal aging and very mild Alzheimer's disease. *Neurology, 49*, 786–794.

Jahanshad, N., Kohannim, O., Hibar, D. P., Stein, J. L., McMahon, K. L., de Zubicaray, G. I., Medland, S. E., Montgomery, G. W., Whitfield, J. B., Martin, N. G., Wright, M. J., Toga, A. W., & Thompson, P. M. (2012). Brain structure in healthy adults is related to serum transferrin and the H63D polymorphism in the *HFE* gene. *Proceedings of the National Academy of Sciences USA, 109*, E851–859.

Kang, H. M., Zaitlen, N. A., Wade, C. M., Kirby, A., Heckerman, D., Daly, M. J., & Eskin, E. (2008). Efficient control of population structure in model organism association mapping. *Genetics, 178*, 1709–1723.

Kremen, W. S., Prom-Wormley, E., Panizzon, M. S., Eyler, L. T., Fischl, B., Neale, M. C., Franz, C. E., Lyons, M. J., Pacheco, J., Perry, M. E., Stevens, A., Schmitt, J. E., Grant, M. D., Seidman, L. J., Thermenos, H. W., Tsuang, M. T., Eisen, S. A., Dale, A. M., & Fennema-Notestine, C. (2010). Genetic and environmental influences on the size of specific brain regions in midlife: The VETSA MRI study. *NeuroImage, 49*, 1213–1223.

Langbaum, J. B. S., Chen, K., Lee, W., Reschke, C., Bandy, D., Fleisher, A. S., Alexander, G. E., Foster, N. L., Weiner, M. W., Koeppe, R. A., Jagust, W. J., Reiman, E. M., & Alzheimer's Disease Neuroimaging Initiative. (2009). Categorical and correlational analyses of baseline fluorodeoxyglucose positron emission tomography images from the Alzheimer's Disease Neuroimaging Initiative (ADNI). *NeuroImage, 45*, 1107–1116.

Langers, D. R., Jansen, J. F., & Backes, W. H. (2007). Enhanced signal detection in neuroimaging by means of regional control of the global false discovery rate. *NeuroImage, 38*, 43–56.

Leow, A., Huang, S. C., Geng, A., Becker, J., Davis, S., Toga, A., & Thompson, P. (2005). Inverse consistent mapping in 3D deformable image registration: its construction and statistical properties. *Information Processing in Medical Imaging, 19*, 493–503.

Leow, A. D., Yanovsky, I., Parikshak, N., Hua, X., Lee, S., Toga, A. W., Jack, C. R., Bernstein, M. A., Britson, P. J., Gunter, J. L., Ward, C. P., Borowski, B., Shaw, L. M., Trojanowski, J. Q., Fleisher, A. S., Harvey, D., Kornak, J., Schuff, N., Alexander, G. E., Weiner, M. W., Thompson, P. M., & Alzheimer's Disease Neuroimaging Initiative. (2009). Alzheimer's Disease Neuroimaging Initiative: A one-year follow up study using tensor-based morphometry correlating degenerative rates, biomarkers and cognition. *NeuroImage, 45*, 645–655.

Li, M. X., Gui, H. S., Kwan, J. S. H., & Sham, P. C. (2011). GATES: A rapid and powerful Gene-Based Association Test using Extended Simes procedure. *American Journal of Human Genetics, 88*, 283–293.

Li, M. X., Sham, P. C., Cherny, S. S., & Song, Y. Q. (2010). A knowledge-based weighting framework to boost the power of genome-wide association studies. *PLoS One, 5*, e14480.

Liang, W. S., Chen, K. W., Lee, W., Sidhar, K., Corneveaux, J. J., Allen, A. N., Myers, A., Villa, S., Meechoovet, B., Pruzin, J., Bandy, D., Fleisher, A. S., Langbaum, J. B., Huentelman, M. J., Jensen, K., Dunckley, T., Caselli, R. J., Kaib, S., & Reiman, E. M. (2011). Association between *GAB2* haplotype and higher glucose metabolism in Alzheimer's disease-affected brain regions in cognitively normal *APOE* epsilon 4 carriers. *NeuroImage, 58*, 974–974.

Lin, K., Tang, M., Han, H., Guo, Y., Lin, Y., & Ma, C. (2010). *GAB2* is not associated with late-onset Alzheimer's disease in Chinese Han. *Neurological Science, 31*, 277–281.

Magistretti, P. J., & Pellerin, L. (1996). Cellular bases of brain energy metabolism and their relevance to functional brain imaging: Evidence for a prominent role of astrocytes. *Cerebral Cortex, 6*, 50–61.

Mark, R. J., Pang, Z., Geddes, J. W., Uchida, K., & Mattson, M. P. (1997). Amyloid beta-peptide impairs glucose transport in hippocampal and cortical neurons: Involvement of membrane lipid peroxidation. *The Journal of Neuroscience, 17*, 1046–1054.

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

Nichols, T., Brett, M., Andersson, J., Wager, T., & Poline, J.-B. (2005). Valid conjunction inference with the minimum statistic. *NeuroImage, 25*, 653–660.

Peper, J. S., Brouwer, R. M., Boomsma, D. I., Kahn, R. S., & Hulshoff Pol, H. E. (2007). Genetic influences on human brain structure: A review of brain imaging studies in twins. *Human Brain Mapping, 28*, 464–473.

Pezawas, L., Verchinski, B. A., Mattay, V. S., Callicott, J. H., Kolachana, B. S., Straub, R. E., Egan, M. F., Meyer-Lindenberg, A., & Weinberger, D. R. (2004). The brain-derived neurotrophic factor val66met polymorphism and variation in human cortical morphology. *Journal of Neuroscience, 24*, 10099–10102.

Piert, M., Koeppe, R. A., Giordani, B., Berent, S., & Kuhl, D. E. (1996). Diminished glucose transport and phosphorylation in Alzheimer's disease determined by dynamic FDG-PET. *Journal of Nuclear Medicine, 37*, 201–208.

Ramirez-Lorca, R., Boada, M., Saez, M. E., Hernandez, I., Mauleon, A., Rosende-Roca, M., Martinez-Lage, P., Gutierrez, M., Real, L. M., Lopez-Arrieta, J., Gayan, J., Antunez, C., Gonzalez-Perez, A., Tarraga, L., & Ruiz, A. (2009). *GAB2* gene does not modify the risk of Alzheimer's disease in Spanish *APOE* 4 carriers. *Journal of Nutrition Health and Aging, 13*, 214–219.

Reiman, E. M., Webster, J. A., Myers, A. J., Hardy, J., Dunckley, T., Zismann, V. L., Joshipura, K. D., Pearson, J. V., Hu-Lince, D., Huentelman, M. J., Craig, D. W., Coon, K. D., Liang, W. S., Herbert, R. H., Beach, T., Rohrer, K. C., Zhao, A. S., Leung, D., Bryden, L., Marlowe, L., Kaleem, M., Mastroeni, D., Grover, A., Heward, C. B., Ravid, R., Rogers, J., Hutton, M. L., Melquist, S., Petersen, R. C., Alexander, G. E., Caselli, R. J., Kukull, W., Papassotiropoulos, A., & Stephan, D. A. (2007). *GAB2* alleles modify Alzheimer's risk in *APOE* epsilon4 carriers. *Neuron, 54*, 713–720.

Scahill, R. I., Schott, J. M., Stevens, J. M., Rossor, M. N., & Fox, N. C. (2002). Mapping the evolution of regional atrophy in Alzheimer's disease: Unbiased analysis of fluid-registered serial MRI. *Proceedings of the National Academy of Sciences of the United States of America, 99*, 4703–4707.

Schjeide, B. M., Hooli, B., Parkinson, M., Hogan, M. F., DiVito, J., Mullin, K., Blacker, D., Tanzi, R. E., & Bertram, L. (2009). *GAB2* as an Alzheimer disease susceptibility gene: Follow-up of genome-wide association results. *Archives of Neurology, 66*, 250–254.

Shaw, P., Lerch, J. P., Pruessner, J. C., Taylor, K. N., Rose, A. B., Greenstein, D., Clasen, L., Evans, A., Rapoport, J. L., & Giedd, J. N. (2007). Cortical morphology in children and adolescents with different apolipoprotein E gene polymorphisms: An observational study. *Lancet Neurology, 6*, 494–500.

Stein, J. L., Medland, S. E., Arias Vasquez, A., Hibar, D. P., Senstad, R. E., Winkler, A. M., ... Thompson, P. M., for the Enhancing Neuro Imaging Genetics through Meta-Analysis (ENIGMA) Consortium. (2012). Identification of common variants associated with human hippocampal and intracranial volumes. *Nature Genetics*, *44*, 552–561.

Thompson, P. M., Cannon, T. D., Narr, K. L., van Erp, T., Poutanen, V. P., Huttunen, M., Lönnqvist, J., Standertskjöld-Nordenstam, C. G., Kaprio, J., Khaledy, M., Dail, R., Zoumalan, C. I., & Toga, A. W. (2001). Genetic influences on brain structure. *Nature Neuroscience, 4*, 1253–1258.

Thompson, P. M., Martin, N. G., & Wright, M. J. (2010). Imaging genomics. *Current Opinion in Neurology, 23*, 368–373.

Trollmann, R., Rehrauer, H., Schneider, C., Krischke, G., Huemmler, N., Keller, S., Rascher, W., & Gassmann, M. (2010). Late-gestational systemic hypoxia leads to a similar early gene response in mouse placenta and developing brain. *American Journal of Physiology: Regulatory, Integrative and Comparative Physiology, 299*, R1489–1499.

Wang, K., & Abbott, D. (2008). A principal components regression approach to multilocus genetic association studies. *Genetic Epidemiology, 32*, 108–118.

Wellcome Trust Case Control Consortium. (2007). Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature, 447*, 661–678.

Zhang, S., & Broxmeyer, H. E. (2000). Flt3 ligand induces tyrosine phosphorylation of *gab1* and *gab2* and their association with *shp-2*, grb2, and PI3 kinase. *Biochemical and Biophysical Research Communications, 277*, 195–199.

Zhao, C., Yu, D. H., Shen, R., & Feng, G. S. (1999). *GAB2*, a new pleckstrin homology domain-containing adapter protein, acts to uncouple signaling from ERK kinase to Elk-1. *The Journal of Biological Chemistry, 274*, 19649–19654.

CHAPTER 3

Radial distances as a surface-based endophenotype for genetic association

3.1 Genetic clustering for increased power in genetic studies

This section is adapted from:

Derrek P. Hibar, Sarah E. Medland, Jason L. Stein, Sungeun Kim, Li Shen, Andrew J. Saykin, Greig I. de Zubicaray, Katie L. McMahon, Grant W. Montgomery, Nicholas G. Martin, Margaret J. Wright, Srdjan Djurovic, Ingrid Agartz, Ole A. Andreassen, Paul M. Thompson (2013). Genetic clustering on the hippocampal surface for genome-wide association studies, MICCAI 2013, Nagoya, Japan, Sept. 22-26 2013 [8-page paper; peer-reviewed].

Genetic clustering on the hippocampal surface for genome-wide association studies

Derrek P. Hibar[1], Sarah E. Medland[2], Jason L. Stein[1], Sungeun Kim[3], Li Shen[3], Andrew J. Saykin[3], Greig I. de Zubicaray[4],
Katie L. McMahon[5], Grant W. Montgomery[2], Nicholas G. Martin[2],
Margaret J. Wright[2], Srdjan Djurovic[6], Ingrid Agartz[6,7],
Ole A. Andreassen[6], Paul M. Thompson[1]

[1]Imaging Genetics Center, Laboratory of Neuro Imaging,
UCLA School of Medicine, Los Angeles, CA, USA
[2]Queensland Institute of Medical Research, Brisbane, Australia
[3]Center for Neuroimaging, Department of Radiology and Imaging Sciences,
Indiana University School of Medicine, Indianapolis, IN, USA
[4]Functional Magnetic Resonance Imaging Laboratory, School of Psychology,
University of Queensland, Brisbane, Australia
[5]Centre for Advanced Imaging, University of Queensland, Brisbane, Queensland,
Australia
[6]KG Jebsen Centre for Psychosis Research, Institute of Clinical Medicine,
University of Oslo, Oslo, Norway
[7]Department of Psychiatry, Diakonhjemmet Hospital, Oslo, Norway

Abstract. Imaging genetics aims to discover how variants in the human genome influence brain measures derived from images. Genome-wide association scans (GWAS) can screen the genome for common differences in our DNA that relate to brain measures. In small samples, GWAS has low power as individual gene effects are weak and one must also correct for multiple comparisons across the genome and the image. Here we extend recent work on genetic clustering of images, to analyze surface-based models of anatomy using GWAS. We performed spherical harmonic analysis of hippocampal surfaces, automatically extracted from brain MRI scans of 1254 subjects. We clustered hippocampal surface regions with common genetic influences by examining genetic correlations (r_g) between the normalized deformation values at all pairs of surface points. Using genetic correlations to cluster surface measures, we were able to boost effect sizes for genetic associations, compared to clustering with traditional phenotypic correlations using Pearson's *r*.

Keywords: heritability, GWAS, clustering, hippocampus, 3D surfaces, imaging genetics

1 Introduction

An important focus of biomedical research is the analysis of biomarkers – easily attainable and reproducible measurements that relate to disease severity or predict clinical decline. In neuroimaging, methods that quantify brain morphometry (e.g., anatom-

ical volumes or shapes, expansions, contractions, etc.) offer promising biomarkers for a variety of brain diseases and disorders. Surface-based morphometry of cortical and subcortical structures has been greatly advanced by ideas in computational geometry – many groups have applied surface meshes, "*M*-reps", spectral analysis, differential forms, or partial differential equations – to map disease effects and dynamic changes in the brain [1]. Surface models of subcortical structures such as the hippocampus can reveal 3D shape differences between healthy controls and patients with neurological or psychiatric disorders such as schizophrenia [2] and Alzheimer's disease [3].

More recently, researchers in *imaging genetics* have adapted computational anatomy methods to analyze genetic effects on the brain. Many brain diseases are genetically influenced, and there is an urgent need to find specific variants in our DNA – both common and rare – that contribute to variations in disease and brain measures. It is now feasible to test how variants along the human genome relate to disease biomarkers or imaging measures using genome-wide association scans (GWAS). One study recently applied GWAS to brain MRI data from over 21,000 people, discovering new genetic variants affecting hippocampal volumes [4]. However, GWA studies have low power if they test a large number of individual phenotypes – if GWAS is run at each voxel in an image, an astronomical correction must be made for the multiple statistical tests across the image and genome [5,6]. Here we build on recent work [7,8] using genetic clustering to increase power and prioritize regions for GWAS. We develop a framework to perform GWAS on 3D anatomical surface models. We demonstrate our method on hippocampal surfaces from a large cohort of 1254 subjects, scanned in independent studies on 3 continents.

2 Methods

2.1 Imaging data

3D T1-weighted structural brain MRI and genotyping data were obtained from three independent cohorts: the Alzheimer's Disease Neuroimaging Initiative (ADNI), Queensland Twins Imaging Study (QTIM), and Thematically Organized Psychosis Study (TOP). We focused on healthy controls from each study, but we also included people with mild cognitive impairment (MCI) in the ADNI. In total, there were 511 ADNI subjects (299 males; age mean±sd: 75.5±6.5 years; 323 MCI patients), 571 QTIM subjects (218 males; age mean±sd: 23.9±2.3 years; monozygotic and dizygotic twins and siblings from 335 families), and 172 TOP subjects (90 males; age mean±sd: 35.8±9.8). Genotyping data was filtered to remove SNPs with minor allele frequency <0.01, call rate <95%, violations of Hardy-Weinberg Equilibrium $p<1\text{x}10^{-6}$. The filtered genotype data was imputed to a custom '1000 Genomes' reference set (phase 1, release 3) which excludes non-European samples and singleton SNPs [9].

2.2 Hippocampal surface generation

Hippocampal (HP) segmentations were obtained using the freely-available and automated FSL FIRST segmentation algorithm [10]. Segmentation quality for the left and right hippocampus across all three cohorts was individually inspected by the first

author. Subjects with segmentations not covering the entire HP, or including regions outside the HP (defined by [11]) were removed. The SPHARM-MAT Toolbox for Matlab (V1.0) [12,13] was used to generate hippocampal surface models. First, we ensured that each binary segmentation label had a spherical topology. Binary segmentations were parameterized using triangular mesh surfaces, with a bijective mapping of each point p on the surface to a unit sphere with (θ, φ) coordinates, such that: $p(\theta,\varphi) = (x(\theta,\varphi), y(\theta,\varphi), z(\theta,\varphi))^T$, using the Control of Area and Length Distortions (CALD) algorithm [12]. The object surface was then expanded in terms of a set of spherical harmonic basis functions of order m and degree l [13]. This expansion has the form $p(\theta,\varphi) = \sum_{m=0}^{\infty}\sum_{l=-m}^{m} C_l^m\, Y_l^m(\theta,\varphi)$, where p is defined as above and C_l^m is a set of Fourier coefficient weights for the basis functions: $C_l^m = (c_{xl}^m, c_{yl}^m, c_{zl}^m)^T$. The spherical harmonic models of the surfaces were then aligned using 12 degrees of freedom to a common template model comprised of an average of 40 healthy controls from the QTIM sample using the SHREC algorithm [14]. A translation and rotation matrix for a given mesh to the common template using SHREC matches landmarks on the surface of an object to similar points on the template (a solution is found by minimizing the root mean squared distance) [14]. In this way, we mapped the points along the surface to a common space across subjects and studies, while preserving individual morphometric differences of interest.

2.3 Quantifying morphometric differences on surfaces

We determined the distance a given point on the hippocampal surface had to be deformed to match the equivalent point on the common template surface by first calculating the simple deformation matrix, $M = (x_i, y_i, z_i)$, where i is the index of vertices of length n, from a coordinates matrix V of vertices compared to the vertices in the average template A: $M = V - A$. Next we calculated the vertex normals of each individual's 3D mesh in MATLAB using the *patch* function, which returns an n-by-3 normalization matrix, **N**. We project the deformation onto the vertex normals and obtain a vector of deformation scalars for each vertex, s, such that: $s_i = \sum_{j=1}^{3}(M \cdot N)_{i,j}$. The deformation value preserves in-out differences along the surface normal (a contraction or expansion to match the template). Each value in the normalized deformation vector, s, represents the expansion or contraction required to match a given vertex on the surface of an individual subject's hippocampal surface to the equivalent point on the average template surface.

2.4 Optimizing parameters using test-retest data

To examine the ideal parameters required to maximize the reliability of the hippocampus surface reconstruction while minimizing data smoothing and the density of the reconstructed 3D mesh, we obtained test-retest data from 40 healthy young adults in the QTIM study scanned twice on the same scanner with a mean interval of four months. We examined how the reliability of surface reconstruction within the same subject changes, as a function of the surface sampling density and the extent of heat kernel smoothing [14]. We calculated the intraclass correlation coefficient (ICC) at

each point along the surface to quantify the reproducibility of hippocampal surface models across test-retest data.

2.5 Genetic versus phenotypic clustering and GWAS

We wanted to compare the GWAS performance of clustered regions of interest on the hippocampal surfaces chosen by *genetic correlation* (r_g) relative to those chosen by traditional phenotypic correlations using Pearson's r_p. We calculated the genetic and phenotypic correlations between the normalized deformation values at each point on the surface with all other deformation values on the surface bilaterally, yielding a genetic correlation matrix and a separate phenotypic correlation matrix of the same size. We calculated r_g using the cross-twin, cross-trait method in 142 dizygotic and 120 monozygotic twin pairs, controlling for age and sex [15]. The phenotypic correlation r_p was the partial correlation between traits, controlling for age and sex. The genetic correlation determines areas on the surface of the hippocampus with common genetic determinants by using the known genetic relationships between monozygotic and dizygotic twins. This is not the same as phenotypic correlations, where measures from different regions can covary due to a combination of genetic *and environmental* effects. The genetic correlation is calculated from the covariance between two traits: $Cov(G_x,G_y)/\sqrt{(Var(G_x) * Var(G_y))}$, where G_x and G_y are the genetic effects that influence the two traits *x* and *y*. When the two traits are controlled by overlapping genetic factors they will covary, leading to a high genetic correlation value.We applied *x*-means clustering to the genetic and phenotypic correlation matrices, separately. The *x*-means algorithm is an iterative form of the *k*-means clustering algorithm that chooses the best number of clusters, *k*, using the Bayesian Information Criterion (BIC) [16]. Cluster membership was mapped back onto the 3D surface. Deformation values in the clustered regions were averaged across the cluster. Values in each cluster, for each subject, were used as phenotype values in a GWAS.

Genome-wide association tests were conducted separately within each sample and combined meta-analytically (described below) for the final results. In ADNI and TOP samples, we performed association tests using multiple linear regression, implemented in the *mach2qtl* program [18]. Association tests in the family-based QTIM study employed mixed-effects models to account for twin and family relationships, as implemented in *merlin-offline* [19]. All association tests controlled for sex, age, and intracranial volume (ICV). Each subject's ICV was estimated as the determinant of the affine transformation matrix to the standard FSL template. GWAS results from within each cluster were combined using an inverse variance-weighted meta-analysis, implemented in *metal* [20].

3 Results

Test-retest data show that reproducibility of our hippocampus surface models was moderate but in line with the reproducibility of volume segmentations achieved by others [4] (ICC=0.66 for the left hippocampus and ICC=0.73 for the right) using a low-density icosahedral sampling mesh (called 'icosa2' in SPHARM-MAT) and without smoothing the data (see **Table 1**). We used the most parsimonious model for

our analysis; we examined the surface morphology of the 'icosa2' sample surface at 162 vertices (so 324 vertices left and right) with no heat kernel smoothing [17].

Table 1. Intraclass correlation coefficient values for left and right hippocampal surfaces. 'icosaX' is the name of the sampling mesh provided in SPHARM-MAT; larger values in the name represent a finer sampling mesh (more vertices). Heat kernel smoothing was performed at three different standard deviation values (a parameter of the heat kernel smoothing algorithm) for 100 iterations. The most parsimonious model bilaterally uses the 'icosa2' mesh, with no smoothing.

Left Hippo.	**No Smoothing**	**1mm**	**2mm**	**3mm**
'icosa2'	0.67	0.53	0.53	0.51
'icosa4'	0.67	0.67	0.67	0.67
'icosa6'	0.67	0.67	0.67	0.67
Right Hippo.				
'icosa2'	0.73	0.63	0.62	0.62
'icosa4'	0.73	0.73	0.73	0.73
'icosa6'	0.73	0.73	0.73	0.74

We estimated the number of clusters sufficient to group related vertices based on their phenotypic correlations and separately their genetic correlations with all other points on the hippocampal surface using *x*-means clustering. The most parsimonious models for both the phenotypic and genotypic correlation matrices determined by BIC were *k*-means clustering with 2 groups. To visualize the clusters, we mapped the cluster membership back onto the 3D average template surfaces. The cluster memberships determined by the phenotypic correlation are shown in **Fig. 1** and the genotypic correlation in **Fig. 2**. The cluster regions of interest selected by phenotypic correlation are highly similar to those chosen by genotypic correlation. There does seem to be noticeable differences in the cluster membership along the bottom left hippocampal surface. In addition, there was a clear bilateral symmetry, with cluster 1 (*in green*) occupying the outer curves of the structure and cluster 2 (*in red*) the inner curve.

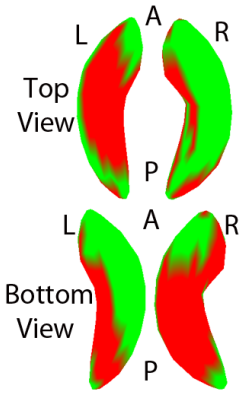


Fig. 1. A 3D projection of the cluster membership determined by phenotypic clustering onto the average template images (*A and P denote anterior and posterior*).

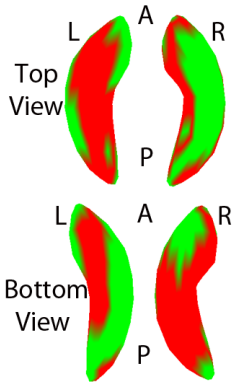


Fig. 2. A 3D projection of the cluster membership from genetic clustering onto the average template images (*A and P denote anterior and posterior*). These are regions where coherent genetic influences are detected, so they are clustered together to provide a coherent signal for GWA.

We conducted a genome-wide association study on the average deformation values in each of the clusters across subjects. Our criterion for significance is the standard genome-wide cut-off ($p<5x10^{-8}$), but after applying a further Bonferroni correction, for testing two separate phenotypes in each condition, our new significance criterion is $p<2.5x10^{-8}$.

After meta-analysis, only cluster 1 from the genetic correlation clustering yielded a region of genome-wide significance. The most strongly associated SNP in the *FBLN2* gene was rs145212527 after meta-analysis: $P_{MA}= 1.25x10^{-8}$; Effect Allele = T; Freq = 0.956; $Beta_{MA}= 0.354$; $SE_{MA}=0.0621$ (**Fig. 3**). Each individual study provided support for this SNP and the same direction of effect (ADNI: $p=0.0073$, $\beta=0.389$, SE=0.145; QTIM: $p=0.00059$, $\beta=0.300$, SE=0.087; TOP: $p=0.00017$, $\beta=0.421$, SE=0.112). Neither of the GWAS analyses of the clusters determined by phenotypic clustering yielded significant results. The top SNP in cluster 1 (rs145212527) was the same SNP found in the genetic clustering analysis of cluster 1. However, the *p*-value was less strong than for the genetic clustering GWAS and did not pass significance ($p=4.6x10^{-7}$).

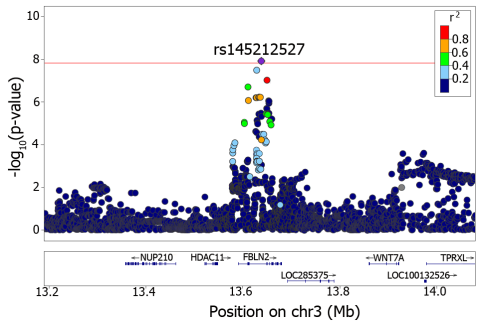


Fig. 3. LocusZoom plot [21] of the most highly associated SNP from the GWAS of cluster 1 from genetic clustering, after meta-analysis. Each point is a SNP; points above the *red horizontal line* are genome-wide significant. Each point's color gives the linkage disequilibrium (r^2) of that point to rs145212527.

4 Discussion

This paper's major contributions are to: 1) perform the first-ever genetic clustering analysis on the hippocampal surface, 2) use genetic correlation values to prioritize and group related regions based on genetic similarity in an image to reduce the multiple comparisons correction, and 3) to demonstrate a case where the added information about common genetic determinants from genetic correlations can boost power for genomic association analyses compared to traditional phenotypic correlation.

In addition, we identified a genome-wide significant SNP affecting hippocampal structure in the *FBLN2* gene. The Allen Human Brain Atlas shows that this gene is differentially expressed in the hippocampus. *FBLN2* is involved with tissue organization, and in differentiation of neurons and other cells [22]. In some ways, clustering the data before performing GWAS is related to performing a GWAS at each point and performing cluster-wise correction for multiple comparisons. The cluster-wise correction methods of Hayasaka and Nichols may be useful for this purpose [23]. However, in this current paper we show that using cluster-wise methods that incorporate genetic correlation methods are more powerful and the methods of [23] do not incorporate genetic correlation into the model. Another paper [24] used sparse models to simultaneously select SNPs from a subset of candidate SNPs and correlated features along the surface. However, the model in [24] has limited utility in high dimensional applications, such as searching the full genome as we did in this study. Additionally, further work is still necessary to confirm that clustering methods are more powerful that voxel-wise analyses. However, this was discussed previously [7]. These are promising findings; further studies will attempt to replicate the genetic results and study the biological pathways they may affect.

References

1. Wang, Y., et al. "Surface-based TBM boosts power to detect disease effects on the brain: An N=804 ADNI study." *NeuroImage* 56.4 (2011): 1993-2010.
2. Styner, M., et al. "Boundary and medial shape analysis of the hippocampus in schizophrenia." *Medical Image Analysis* 8.3 (2004): 197-203.
3. Frisoni, G.B., et al. "Mapping local hippocampal changes in Alzheimer's disease and normal ageing with MRI at 3 Tesla." *Brain* 131.12 (2008): 3266-3276.
4. Stein, J.L., et al. "Identification of common variants associated with human hippocampal and intracranial volumes." *Nature Genetics* 44.5 (2012): 552-561.
5. Stein, J.L., et al. "Voxelwise genome-wide association study (vGWAS)." *NeuroImage* 53.3 (2010): 1160.
6. Hibar, D.P., et al. "Voxelwise gene-wide association study (vGeneWAS): multivariate gene-based association testing in 731 elderly subjects." *NeuroImage* 56.4 (2011): 1875-1891.

7. Chiang, M.-C., et al. "Gene network effects on brain microstructure and intellectual performance identified in 472 twins." *Journal of Neuroscience* 32.25 (2012): 8732-8745.
8. Chen, C.-H., et al. "Hierarchical genetic organization of human cortical surface area." *Science* 335.6076 (2012): 1634-1636.
9. ENIGMA2 Genetics Support Team. ENIGMA2 1KGP Cookbook (v3) [Online]. The Enhancing Neuroimaging Genetics through Meta-Analysis (ENIGMA) consortium. (27 July 2012).
10. Patenaude, B., et al. "A Bayesian model of shape and appearance for subcortical brain segmentation." *NeuroImage* 56.3 (2011): 907-922.
11. Pantel, J., et al. "A new method for the in vivo volumetric measurement of the human hippocampus with high neuroanatomical accuracy." *Hippocampus* 10.6 (2000): 752-758.
12. Shen, L., Makedon, F. "Spherical mapping for processing of 3D closed surfaces." *Image and Vision Computing* 24.7 (2006): 743-761.
13. Brechbuhler, C., Gerig, G., Kubler, O. Parameterization of closed surfaces for 3D shape description. *Comp. Vis. Image Understanding.* 1995;61:154–170.
14. Shen, L., Farid, H., and McPeek, M.A. "Modeling 3-Dimensional Morphological Structures Using Spherical Harmonics." *Evolution* 63.4 (2009): 1003-1016.
15. Neale, M.C., et al. *Methodology for genetic studies of twins and families*. No. 67. Springer, 1992.
16. Pelleg, D., et al. "X-means: Extending k-means with efficient estimation of the number of clusters." *Proceedings of the Seventeenth International Conference on Machine Learning*. Vol. 1. 2000.
17. Chung, M.K. "Heat kernel smoothing on unit sphere." *Biomedical Imaging: Nano to Macro, 2006. 3rd IEEE International Symposium on*. IEEE, 2006.
18. Li, Y., et al. "MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes." *Genetic Epidemiology* 34.8 (2010): 816-834.
19. Chen, W.-M., Abecasis, G.R. "Family-based association tests for genomewide association scans." *The American Journal of Human Genetics* 81.5 (2007): 913-926.
20. Willer, C.J., Li Y., Abecasis, G.R. "METAL: fast and efficient meta-analysis of genomewide association scans." *Bioinformatics* 26.17 (2010): 2190-2191.
21. Pruim, R.J., et al. "LocusZoom: regional visualization of genome-wide association scan results." *Bioinformatics* 26.18 (2010): 2336-2337.
22. Miosge, N., et al. "The extracellular matrix proteins fibulin-1 and fibulin-2 in the early human embryo." *The Histochemical Journal* 28.2 (1996): 109-116.
23. Hayasaka, S, Nichols, T. E. "Combining voxel intensity and cluster extent with permutation test framework." *Neuroimage* 23.1 (2004): 54-63.
24. Wan, Jing, et al. "Hippocampal surface mapping of genetic risk factors in AD via sparse learning models." *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2011*. Springer Berlin Heidelberg, 2011. 376-383.

CHAPTER 4

Examining the effects of epistatic interactions in full-brain phenotypes

4.1 GPU-accelerated interaction testing in the full SNP-SNP interactome

This section is adapted from:

Derrek P. Hibar, Jason L. Stein, Neda Jahanshad, Arthur W. Toga, Katie L. McMahon, Greig I. de Zubicaray, Grant W. Montgomery, Nicholas G. Martin, Margaret J. Wright, Michael W. Weiner, Paul M. Thompson (2013). Exhaustive search of the SNP-SNP interactome identifies replicated epistatic effects on brain volume, MICCAI 2013, Nagoya, Japan, Sept. 22-26 2013 [8-page paper; peer-reviewed].

Exhaustive search of the SNP-SNP interactome identifies epistatic effects on brain volume in two cohorts

Derrek P. Hibar[1], Jason L. Stein[1], Neda Jahanshad[1], Omid Kohannim[1], Arthur W. Toga[1], Katie L. McMahon[2], Greig I. de Zubicaray[3], Grant W. Montgomery[4], Nicholas G. Martin[4], Margaret J. Wright[4], Michael W. Weiner[5,6], Paul M. Thompson[1]

[1]Imaging Genetics Center, Laboratory of Neuro Imaging, UCLA School of Medicine, Los Angeles, USA
[2]Center for Magnetic Resonance, School of Psychology, University of Queensland, Brisbane, Australia
[3]Functional Magnetic Resonance Imaging Laboratory, School of Psychology, University of Queensland, Brisbane, Australia
[4]Genetic Epidemiology Laboratory, Queensland Institute of Medical Research, Brisbane, Australia
[5]Departments of Radiology, Medicine, Psychiatry, UC San Francisco, San Francisco, USA
[6]Department of Veterans Affairs Medical Center, San Francisco, USA

Abstract. The SNP-SNP interactome has rarely been explored in the context of neuroimaging genetics mainly due to the complexity of conducting ~10^{11} pairwise statistical tests. However, recent advances in machine learning, specifically the iterative sure independence screening (SIS) method, have enabled the analysis of datasets where the number of predictors is much larger than the number of observations. Using an implementation of the SIS algorithm (called EPISIS), we used exhaustive search of the genome-wide, SNP-SNP interactome to identify and prioritize SNPs for interaction analysis. We identified a significant SNP pair, rs1345203 and rs1213205, associated with temporal lobe volume. We further examined the full-brain, voxelwise effects of the interaction in the ADNI dataset and separately in an independent dataset of healthy twins (QTIM). We found that each additional loading in the epistatic effect was associated with ~5% greater brain regional brain volume (a protective effect) in both the ADNI and QTIM samples.

Keywords: epistasis, interaction, genome, sure independence, tensor-based morphometry

1 Introduction

Traditional univariate methods can test the association of common genetic variants with complex quantitative traits, but they only consider the marginal effect of a single

[1]**SNP** (=single nucleotide polymorphism): a single-letter variant in the genome; these variations are common, even in healthy human populations, and their effects on brain measures can be assessed using association testing, at one SNP or up to a million genotyped SNPs.
[2]**Interactome**: The study of interactions between genetic variants or sets of variants in terms of their effects on traits such as brain measures.

locus and potentially miss variance explained by synergistic or interacting effects of pairs or sets of SNPs[1] [1]. For many complex traits, the similarity of family members drops faster than would be expected as relatedness decreases [2]. This implies that there are non-additive (epistatic) interactions involved in the etiology of many complex traits. Statistical interactions have been demonstrated to be plausible representations of the complex interactions of genes in biological pathways [3-4].

Some prior studies have examined second-order interactive effects of SNPs on brain structure [5-7]. However, none of these studies has considered genome-wide genotype data; the closest conceptually related study tested for SNP effects on diffusion imaging measures, and aggregated all SNPs with correlated effects into a network [8]. The concept here is different, and aims to assess gene pairs that influence each other's effects on the brain. Prior studies tested interaction effects only for a limited number of popular candidate genes. Any approach based on pre-selecting a pair of genes will overlook a vast search space of potential interactions among SNPs in the genome that have no obvious prior connection. Also, a large main effect is not necessary to be able to detect significant second-order interactions [9]. Given this, prior hypotheses focusing on SNPs with large individual effects may also overlook large second-order effects. Importantly, power estimates for detecting interactive effects are comparable to those for single SNP tests [1]. In simulation studies, the inclusion of interaction terms can boost the power to detect main effects, at least for certain genetic tests [10]. Here we examined the genome-wide, SNP-SNP interactome[2] to test genetic associations with a quantitative biomarker of Alzheimer's disease (temporal lobe volume) in the public Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset. We further examine the whole-brain effects of interaction pairs in statistical parametric maps generated with tensor-based morphometry (TBM); we also replicate our tests in an independent, non-overlapping dataset of young healthy twins from the Queensland Twin Imaging (QTIM) study [11].

2 Methods

2.1 Imaging Parameters and Study Information

We downloaded the full baseline set of 818 high-resolution, T1-weighted structural MRI brain scans from the Alzheimer's Disease Neuroimaging Initiative (ADNI). ADNI is a multi-site, longitudinal study of patients with Alzheimer's disease (AD), mild cognitive impairment (MCI) and healthy elderly controls (HC). Subjects were scanned with a standardized protocol to maximize consistency across sites. We used the baseline 1.5 Tesla MRI scans, i.e., the T1-weighted 3D MP-RAGE scans, with TR/TE = 2400/1000 ms, flip angle = 8°, slice thickness = 1.2 mm, and a final voxel resolution = 0.9375 x 0.9375 x 1.2 mm^3. Raw MRI scans were pre-processed to remove signal inhomogeneity, non-brain tissue, and affine registered to the MNI template (using 9 parameters).

Additionally, we obtained 753 high-resolution, T1-weighted structural MRI brain scans from the Queensland Twin Imaging (QTIM) study. QTIM is a longitudinal neuroimaging and genetic study of young, healthy twins and their family members. All structural MRI scans were acquired on a single 4-Tesla scanner (Bruker Medspec): T1-weighted images, inversion recovery rapid gradient echo sequence,

TR/TE = 1500/3.35 ms, flip angle = 8˚, slice thickness = 0.9 mm, 256 x 256 acquisition matrix, with a final voxel resolution = $0.9375 \times 0.9375 \times 0.9$ mm^3. Raw MRI scans were pre-processed to remove signal inhomogeneity, non-brain tissue, and affine registered to the ICBM template (using 9 parameters).

2.2 Genotype Pre-processing and Study Demographics

Genome-wide genotyping data were available for the full set of ADNI subjects. We performed standard quality control procedures to ascertain the largest homogenous genetic sub-population in the dataset, using multi-dimensional scaling (MDS) compared to a dataset of subjects of known genetic identity (HapMap III; http://hapmap.ncbi.nlm.nih.gov/). The largest subset contained 737 subjects from the CEU population (Caucasians). We therefore removed the remaining 81 subjects from our analysis to limit the effects of genetic stratification on our statistical analyses [12]. Additionally, we applied filter rules to the genotype data to remove rare SNPs (minor allele frequency < 0.01), violations of Hardy-Weinberg Equilibrium (HWE $p < 5.7x10^{-7}$), and poor call rate (<95%). Data were further "phased" to impute any missing individual genotypes after filtering using the MaCH program [13] following the ENIGMA imputation protocol [14]. After filtering and phasing, 534,033 SNPs remained.

All QTIM subjects were ascertained for genetic similarity, so no subjects were removed before analysis. All 753 subjects in the QTIM dataset clustered with the CEU population, in the MDS analysis. The same genotype filter rules from the ADNI dataset were applied to the QTIM sample's genetic data. After filtering and phasing, 521,232 SNPs remained.

After all rounds of genotype pre-processing, the ADNI sample contained 737 subjects (mean age±sd: 75.5±6.8 yrs; 436 males) comprised of 173 patients diagnosed with Alzheimer's disease, 358 subjects with mild cognitive impairment, and 206 healthy elderly controls. The QTIM sample contained 753 subjects (mean age±sd: 23.1±3.0 yrs; 286 males) and consisted of 110 monozygotic twin pairs, 147 dizygotic twin pairs, 3 dizygotic twin trios, 143 singletons, and 87 siblings from 438 families.

2.3 Tensor-Based Morphometric Differences in the Full Brain

We calculated information on regional brain morphometry using an elastic, nonlinear registration algorithm (3DMI) [15] applied to the entire brain. Voxelwise volumetric differences were stored, using the Jacobian value of the deformation matrix obtained by nonlinearly registering a subject's scan to a study-specific minimum deformation template (MDT). Scans from the ADNI and QTIM datasets were processed and analyzed separately (using separate study templates). The MDT for the ADNI sample is a nonlinear average of 40 age-and-sex matched healthy elderly controls [16]. The MDT for the QTIM is a nonlinear average of 32 age- and sex- matched, unrelated subjects [17]. Nonlinear registration with 3DMI yields a 110 x 110 x 110 voxel statistical parametric map, where the Jacobian value at each voxel represents the expansion required to match the same voxel in the study-specific MDT.

2.4 Genome-Wide, Gene-Gene Interaction Testing

The EPISIS software is an implementation of the machine-learning algorithm called *sure independence screening* (SIS) developed by Fan and Lv [18]. The SIS algorithm is a correlation learning method that can be applied to ultra-high dimensional datasets where the number of predictors p is much greater than the number of observations n. Despite the development of robust methods for cases where $p>n$ (e.g., the Dantzig selector of Candes and Tao [19]) the properties of the selector fail when $p>>n$. Fan and Lv [18] developed the SIS algorithm to reduce the ultra-high dimension of p to a moderately-sized subset, while guaranteeing that the subset still explains the maximum amount of variance explained by the full set of predictors.

We conducted an exhaustive search of association tests of genome-wide SNP-SNP interactions with temporal lobe volume (computed by integrating the Jacobian over an temporal lobe ROI on the MDT) [20] in the ADNI dataset using the EPISIS software. EPISIS utilizes the massively parallel processing available in GPGPU (General-purpose computing on graphics processing units) framework to test $p(p\text{-}1)/2$ SNP-SNP interactions in the ADNI dataset in a feasible timeframe. We used the SIS algorithm with cell-wise dummy coding (CDC) [21] to reduce the full predictor space into a subset d of $n/\log(n)$ interaction terms [18]. After screening the full set of possible two-way SNP-SNP interactions, we applied ridge regression [22-23] to the subset of interaction terms (the multiplicative loading of each SNP-SNP pair) and selected significant SNP-SNP interaction terms using the extended Bayesian Information Criterion (EBIC) [24] with $\gamma = 0.5$. The choice of the parameter γ was chosen based on simulations [21]. The EPISIS software is implemented in CUDA and optimized for parallel processing across multiple NVIDIA GPU cards as detailed elsewhere [21]. A single exhaustive search of the genome-wide, SNP-SNP interactome with EPISIS was completed in 7 hours (using one NVIDIA Tesla C2050 GPU card).

2.5 Voxelwise Interaction Analysis and Replication

We tested the significant SNP-SNP interaction pair selected by ridge regression for association with voxelwise, regional volume differences (V) at each point, i, in the full brain. The association test at each voxel in the ADNI dataset followed the multiplicative interaction model in multiple linear regression:

$$V_i \sim \beta_0 + \beta_{age}X_{age} + \beta_{sex}X_{sex} + \beta_{snp1}X_{snp1} + \beta_{snp2}X_{snp2} + \beta_{snp1,2}X_{snp1}*X_{snp2} + \varepsilon \quad (1)$$

Additionally, we used QTIM as an independent replication sample of the top SNP-SNP interaction pair identified by ridge regression after EPISIS. The voxelwise association tests assume the multiplicative interaction model, detailed previously. Due to the family design of the QTIM sample, we tested association using mixed-effects modeling as implemented in the R package *kinship* (version 1.3) in order to account for relatedness.

3 Results

After screening the full set of SNP-SNP interaction pairs for association with temporal lobe volume in the ADNI dataset, we obtained a subset d of SNP-SNP interac-

tion pairs such that $d = n/\log(n)$. The subset is chosen by ranking the marginal correlation coefficients of each interaction pair and selecting the top d SNP-SNP pairs (correlation learning) [18], in this case $d = 111$ pairs. Next, we applied ridge regression to the pruned subset of SNP-SNP interaction pairs. Using the extended BIC ($\gamma = 0.5$) [21] to estimate significance in our ridge regression, we identified a significant interaction between rs1345203 and rs1213205. The distribution of alleles for each SNP and their interaction is given in Table 1.

Table 1. The distribution of alleles for the significant SNPs and the number of subjects with each genotype by study. For rs1345203 the minor allele is G and the major allele is A in both studies. The minor allele is A and the major allele is G for rs1213205. The association testing assumes an additive model (each subject is assigned a value 0,1,2 based on the number of minor alleles they have at a given SNP). The interaction column gives the number of subjects in each category after multiplying together the counts of each of the alleles.

Study	rs1345203	rs1213205	Interaction
ADNI (n=737)	G/G: 27	A/A: 93	0 loadings: 612
	A/G: 223	G/A: 297	1 loadings: 79
	A/A: 487	G/G: 347	2 loadings: 46
QTIM (n=753)	G/G: 5	A/A: 78	0 loadings: 664
	A/G: 193	G/A: 300	1 loadings: 70
	A/A: 555	G/G: 375	2 loadings: 19

We further examined the significant SNP pair, rs1345203 and rs1213205, for whole-brain effects in the statistical parametric maps generated using tensor-based morphometry (TBM). In the ADNI dataset, we found broad effects bilaterally in the temporal and occipital lobes (**Fig. 1**) after correcting for multiple tests at a 5% false discovery rate (FDR) using the searchlight FDR method [25].

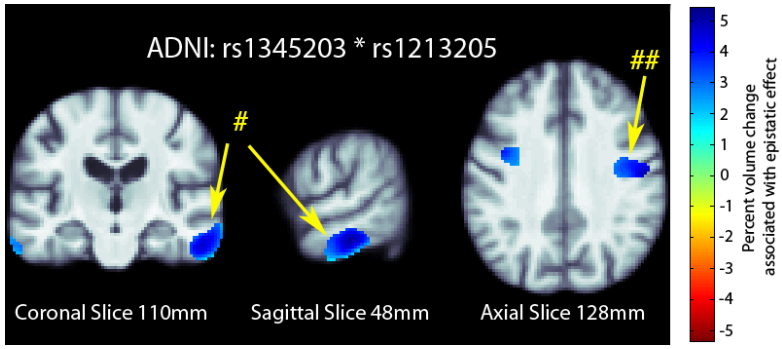


Fig. 1. 3D maps of percent tissue change for each additional genetic variant in the interaction in ADNI. Only significant regions are shown after correcting for multiple comparisons with searchlight FDR [25] at a 5% false discovery rate. Images follow radiological orientation. The origin is placed at the Posterior-Right-Inferior corner. Cooler colors over the tissue represent

tissue expansion (larger regional brain volume) compared to an average template. There is a clear protective effect of the epistatic loadings bilaterally in the temporal (# in the figure) and occipital lobes (## in the figure): as the number of alleles a subject has increases, the amount of local brain tissue they have is also increased on average.

We examined the whole-brain effects of the SNP pair on voxelwise, regional brain volume in the statistical parametric maps in an independent dataset (QTIM). The distribution of alleles for each SNP and their interaction in the QTIM sample is given in Table 1. In the QTIM, we identified significant effects in the left temporal lobe and along the border of the left frontal and occipital lobes (**Fig. 2**) after correction for multiple tests at 5% false discovery rate (FDR) using the searchlight FDR method.

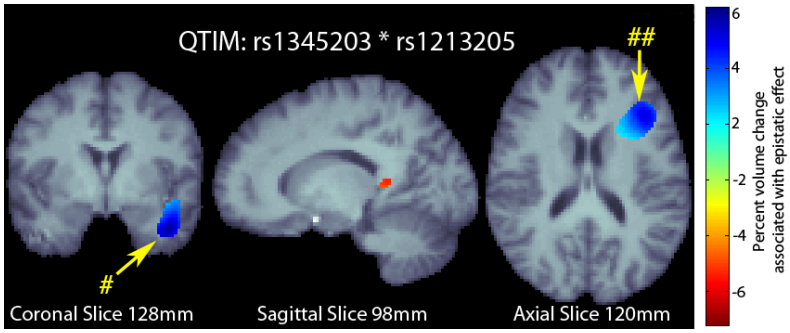


Fig. 2. 3D maps of percent tissue change for each additional genetic variant in the interaction in QTIM. Only significant regions are shown after correcting for multiple comparisons with searchlight FDR [25] at a 5% false discovery rate. Images follow radiological orientation. The origin is placed at the Posterior-Right-Inferior corner. Cooler colors over the tissue represent tissue expansion (larger regional brain volume) compared to an average template. There is a clear protective effect of the epistatic loadings in the left temporal (# in the figure) and along the boundary of the frontal and occipital lobes (## in the figure): as the number of alleles a subject has increases, the amount of local brain tissue they have is also increased on average.

4 Discussion

The genome is incredibly complex and statistical epistasis has been suggested as an appropriate model for the biological interactions among genes and protein products in related pathways [3-4]. Here we examined the multiplicative effect of SNP-SNP pairs on brain volume differences. Significant interaction terms explain additional variance in brain volume beyond what is already explained by the additive SNP terms. In our primary tests of associations with temporal lobe volume in the ADNI dataset, we screened 10^{11} possible SNP-SNP interaction pairs using the GPU acceleration implemented in the EPISIS software. The top 111 interaction pairs were selected after ranking the marginal effect of each SNP-SNP pair on temporal lobe volume, using an implementation of the sure independence screening (SIS) algorithm [18]. We used ridge regression and the extended BIC [24] to identify a significant interaction between rs1345203 and rs1213205. The functional relevance of the two SNPs is as yet

unknown. However, data obtained from the ENCODE dataset (http://genome.ucsc.edu/) show that rs1345203 is located in a transcription factor gene (ELF1/CEBPB) that demonstrates regulatory influence on the DNA structure. The SNP rs1213205 is located in a region of hypersensitivity to cleavage by DNase regulatory elements. It is worth noting that the parameter choices made in the interaction analysis may influence the results, however, parameters were chosen based on the recommended values for EPISIS [21] and SIS [18]. Additional work is still required to identify precisely how these two SNPs might affect brain structure, and to further replicate their interaction. Specifically, we need to identify how changes at a given SNP are related to changes in activity in gene transcription or translation into protein products involved in similar biological pathways.

References

1. Marchini, J., et al. "Genome-wide strategies for detecting multiple loci that influence complex diseases." Nature Genetics 37.4 (2005): 413-417.
2. Wray, N.R., et al. "Multi-locus models of genetic risk of disease." Genome Med 2.10 (2010).
3. Moore, J. H., and Williams, S. M. "Epistasis and its implications for personal genetics." American Journal of Human Genetics 85.3 (2009): 309.
4. Stich, B., et al. "Power to detect higher-order epistatic interactions in a metabolic pathway using a new mapping strategy." Genetics 176.1 (2007): 563-570.
5. Pezawas, L., et al. "Evidence of biologic epistasis between BDNF and SLC6A4 and implications for depression." Molecular Psychiatry 13.7 (2008): 709-716.
6. Tan, H.-Y., et al. "Epistasis between catechol-O-methyltransferase and type II metabotropic glutamate receptor 3 genes on working memory brain function." PNAS 104.30 (2007): 12536-12541.
7. Wang, Y., et al. "Evidence of Epistasis Between the Catechol-O-Methyltransferase and Aldehyde Dehydrogenase 3B1 Genes in Paranoid Schizophrenia." Biological Psychiatry 65.12 (2009): 1048-1054.
8. Chiang, M.-C., et al. "Gene network effects on brain microstructure and intellectual performance identified in 472 twins." Journal of Neuroscience 32.25 (2012): 8732-8745.
9. Cordell, H. J. "Detecting gene–gene interactions that underlie human diseases." Nature Reviews Genetics 10.6 (2009): 392-404.
10. Cordell, H. J., et al. "Statistical modeling of interlocus interactions in a complex disease: rejection of the multiplicative model of epistasis in type 1 diabetes." Genetics 158.1 (2001): 357-367.
11. de Zubicaray G. I., et al. “Meeting the challenges of neuroimaging genetics.” Brain Imaging Behavior 2 (2008): 258–263.
12. Lander, E. S., and Schork, N. J. “Genetic dissection of complex traits.” Science 265.5181 (1994): 2037-2048.
13. Abecasis, G. R., et al. “MaCH: Using Sequence and Genotype Data to Estimate Haplotypes and Unobserved Genotypes.” Genetic Epidemiology 34.8 (2010): 816-834.

14. ENIGMA2 Genetics Support Team. ENIGMA2 1KGP cookbook (v3) [Online]. The Enhancing Neuroimaging Genetics through Meta-Analysis (ENIGMA) consortium. [accessed 27 July 2012]
15. Leow, A., et al. “Inverse consistent mapping in 3D deformable image registration: its construction and statistical properties.” Inf. Process. Med. Imaging 19 (2005): 493–503. 2005.
16. Hua, X., et al. "Unbiased tensor-based morphometry: Improved robustness and sample size estimates for Alzheimer's disease clinical trials." Neuroimage (2012). [EPUB].
17. Jahanshad N., et al. “Brain structure in healthy adults is related to serum transferrin and the H63D polymorphism in the HFE gene.” Proc Natl Acad Sci 109.14 (2012): E851-9.
18. Fan, J., and Lv, J. "Sure independence screening for ultrahigh dimensional feature space." Journal of the Royal Statistical Society: Series B (Statistical Methodology) 70.5 (2008): 849-911.
19. Candes, E., and Tao, T. "The Dantzig selector: Statistical estimation when p is much larger than n." The Annals of Statistics 35.6 (2007): 2313-2351.
20. Stein, J. L., et al. "Genome-wide analysis reveals novel genes influencing temporal lobe structure with relevance to neurodegeneration in Alzheimer's disease." Neuroimage 51.2 (2010): 542-554.
21. Ueki, M., and Tamiya, G. "Ultrahigh-dimensional variable selection method for whole-genome gene-gene interaction analysis." BMC Bioinformatics 13.1 (2012): 72.
22. Hoerl, A. E. “Application of ridge analysis to regression problems.” Chemical Engineering Progress 58 (1962): 54-59.
23. Kohannim, O., et al. "Boosting power to detect genetic associations in imaging using multi-locus, genome-wide scans and ridge regression." Biomedical Imaging: From Nano to Macro IEEE 2011. [EPUB].
24. Chen, J., and Chen, Z. "Extended Bayesian information criteria for model selection with large model spaces." Biometrika 95.3 (2008): 759-771.
25. Langers, D. R., et al. “Enhanced signal detection in neuroimaging by means of regional control of the global false discovery rate.” NeuroImage 38.1 (2007): 43-56.

CHAPTER 5

Analyzing genetic determinants of endophenotypes using meta-analysis

5.1 Genetic analysis of lentiform nucleus volume

This section is adapted from:

Derrek P. Hibar, Jason L. Stein, April B. Ryles, Omid Kohannim, Neda Jahanshad, Sarah E. Medland, Narelle K. Hansell, Katie L. McMahon, Greig I. de Zubicaray, Grant W. Montgomery, Nicholas G. Martin, Margaret J. Wright, Clifford R. Jack, Jr., Michael W. Weiner, Arthur W. Toga, Paul M. Thompson, and the Alzheimer's Disease Neuroimaging Initiative* (2012). Genome-wide association identifies genetic variants associated with lentiform nucleus volume in N=1345 young and elderly subjects. Brain Imaging and Behavior, accepted June 2012. [Epub Ahead of Print].

Brain Imaging and Behavior (2013) 7:102–115
DOI 10.1007/s11682-012-9199-7

ORIGINAL RESEARCH

Genome-wide association identifies genetic variants associated with lentiform nucleus volume in *N*=1345 young and elderly subjects

Derrek P. Hibar · Jason L. Stein · April B. Ryles · Omid Kohannim · Neda Jahanshad Sarah E. Medland · Narelle K. Hansell · Katie L. McMahon · Greig I. de Zubicaray Grant W. Montgomery · Nicholas G. Martin · Margaret J. Wright · Andrew J. Saykin Clifford R. Jack Jr · Michael W. Weiner · Arthur W. Toga · Paul M. Thompson the Alzheimer's Disease Neuroimaging Initiative

Published online: 18 August 2012

Abstract Deficits in lentiform nucleus volume and morphometry are implicated in a number of genetically influenced disorders, including Parkinson's disease, schizophrenia, and ADHD. Here we performed genome-wide searches to discover common genetic variants associated with differences in lentiform nucleus volume in human populations. We assessed structural MRI scans of the brain in two large genotyped samples: the Alzheimer's Disease Neuroimaging Initiative

Some of the data used to prepare this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (www.loni.ucla.edu/ADNI). 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. For a complete listing of ADNI investigators, please see: http://adni.loni.ucla.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

D. P. Hibar · J. L. Stein · A. B. Ryles · O. Kohannim · N. Jahanshad · P. M. Thompson (✉)
Imaging Genetics Center at the Laboratory of Neuro Imaging, Department of Neurology, UCLA School of Medicine, Neuroscience Research Building 225E 635 Charles Young Drive, Los Angeles, CA 90095-1769, USA
e-mail: thompson@loni.ucla.edu

S. E. Medland · N. K. Hansell · G. W. Montgomery · N. G. Martin · M. J. Wright
Genetic Epidemiology Laboratory, Queensland Institute of Medical Research, Brisbane, Australia

S. E. Medland
Neurogenetics Laboratory, Queensland Institute of Medical Research, Brisbane, Australia

S. E. Medland
Broad Institute of Harvard and MIT, Boston, MA, USA

K. L. McMahon
Centre for Advanced Imaging, University of Queensland, Brisbane, Queensland, Australia

G. I. de Zubicaray
Functional Magnetic Resonance Imaging Laboratory, School of Psychology, University of Queensland, Brisbane, Queensland, Australia

A. J. Saykin
Center for Neuroimaging, Department of Radiology and Imaging Science, Indiana University School of Medicine, Indianapolis, IN, USA

C. R. Jack Jr
Mayo Clinic, Rochester, MN, USA

M. W. Weiner
Departments of Radiology, Medicine, Psychiatry, UC San Francisco, San Francisco, CA, USA

M. W. Weiner
Department of Veterans Affairs Medical Center, San Francisco, CA, USA

A. W. Toga
Laboratory of Neuro Imaging, Department of Neurology, UCLA School of Medicine, Neuroscience Research Building 225E 635 Charles Young Drive, Los Angeles, CA 90095-1769, USA

(ADNI; N=706) and the Queensland Twin Imaging Study (QTIM; N=639). Statistics of association from each cohort were combined meta-analytically using a fixed-effects model to boost power and to reduce the prevalence of false positive findings. We identified a number of associations in and around the flavin-containing monooxygenase (*FMO*) gene cluster. The most highly associated SNP, rs1795240, was located in the *FMO3* gene; after meta-analysis, it showed genome-wide significant evidence of association with lentiform nucleus volume ($P_{MA}=4.79\times10^{-8}$). This commonly-carried genetic variant accounted for 2.68 % and 0.84 % of the trait variability in the ADNI and QTIM samples, respectively, even though the QTIM sample was on average 50 years younger. Pathway enrichment analysis revealed significant contributions of this gene to the cytochrome P450 pathway, which is involved in metabolizing numerous therapeutic drugs for pain, seizures, mania, depression, anxiety, and psychosis. The genetic variants we identified provide replicated, genome-wide significant evidence for the *FMO* gene cluster's involvement in lentiform nucleus volume differences in human populations.

Keywords Basal ganglia · Genome-wide association study (GWAS) · MRI · Replication · Morphometry · Drug metabolism

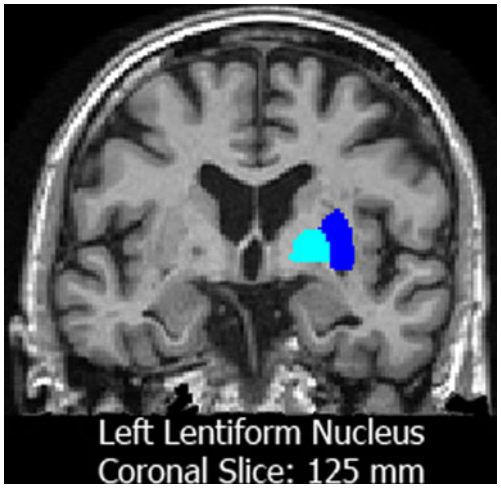


Fig. 1 A coronal slice in a subject from the ADNI sample. The light-blue label represents the left *globus pallidus* and the darker blue label shows the putamen. External and internal portions of the *globus pallidus* are segmented as a single structure in FSL/FIRST (Patenaude et al. 2011)

Introduction

The lentiform nucleus is a lens-shaped, bilateral structure in the basal ganglia bounded by the internal and external capsules. It has three components: the internal and external globus pallidus (Diamond et al. 1985) and the putamen (Fig. 1). The putamen receives dense corticostriate projections from throughout the cortex and funnels information to the external and internal globus pallidus through dense intrabasal ganglionar fibers (Snell 2010). In addition, fibers from the internal globus pallidus project to several thalamic nuclei and continue back to the cortex, primarily to premotor area 6. These projections form the cortico-striato-thalamo-cortical loop, which is involved in initiating and terminating movements (Snell 2010). Both the globus pallidus, and the putamen especially, receive dopamine-rich connections from the *substantia nigra*—the main source of dopamine for the basal ganglia (Snell 2010). Dopamine projections to the basal ganglia are part of the brain's reward circuitry (Schultz 2002).

The lentiform nucleus is implicated in several heritable degenerative and psychiatric disorders. Its role in movement disorders was first identified in studies of *hepatolenticular degeneration* (Wilson 1912)—a disorder that affects both the liver and the lentiform nucleus. Deficits in lentiform nucleus volume have been observed in Parkinson's disease (Dexter et al. 1991; Obeso et al. 2000), Huntington's disease (Marsden et al. 1983; Reiner et al. 1988), and normal aging (Raz et al. 2003). More subtle differences in lentiform nucleus volume are reported in some but not all studies of bipolar disorder (Arnone et al. 2009; Kempton et al. 2008; Strakowski et al. 1999), attention deficit hyperactivity disorder (Castellanos et al. 1996; Ellison-Wright et al. 2008) and schizophrenia (Elkashef et al. 1994; Ellison-Wright et al. 2008). Lesions in the midbrain *tegmentum*—which has reciprocal connections with the lentiform nucleus—are associated with visual and auditory hallucinations (Cascino and Adams 1986). In addition to its many links with known pathology, the lentiform nucleus is a plausible target for genetic analysis, as its volume is highly heritable (Kremen et al. 2010) and can be reliably measured using automated segmentation methods (Morey et al. 2010).

Building on prior studies, here we performed an unbiased genome-wide association study (GWAS) in two large independent cohorts to discover common genetic variants associated with differences in lentiform nucleus volume. The term "unbiased" is often used to describe the type of genetic analysis we performed—a genome-wide association scan—in which we search the whole genome for statistical associations, rather than prioritizing or choosing only a limited subset of variants, such as candidate genes. Arguably, if the genetic loci influencing a given trait are unknown, a broad survey of the genome may avoid missing associations in regions that have no currently known relation to the trait. Association statistics for the genetic variants were combined meta-analytically across two cohorts to boost power and reduce the risk of false positive findings. We assessed both an elderly and a young adult cohort to discover genes with robust associations throughout life.

Methods and materials

Subjects

We analyzed neuroimaging and genome-wide genotype data from two independent samples: the Alzheimer's Disease Neuroimaging Initiative (ADNI) and the Queensland Twin Imaging Study (QTIM). The Alzheimer's Disease Neuroimaging Initiative (ADNI) is a large longitudinal study initiated in 2003 as a public-private partnership between the National Institute on Aging (NIA), the National Institute of Biomedical Imaging and Bioengineering (NIBIB), the Food and Drug Administration (FDA), private pharmaceutical companies, and non-profit organizations. The study aims for ADNI are to identify and investigate biological markers of Alzheimer's disease through a combination of neuroimaging, genetics, neuropsychological tests and other measures in order to develop new treatments, track disease progression, and lessen the time required for clinical trials. The study was conducted according to the Good Clinical Practice guidelines, the Declaration of Helsinki, and U.S. 21 CFR Part 50—Protection of Human Subjects, and Part 56—Institutional Review Boards. Written informed consent was obtained from all participants before protocol-specific procedures were performed. Further information on inclusion criteria and the study protocol may be found online (http://adni-info.org/). Baseline structural MRI scans and genetic data were available for 818 subjects (as of August 1, 2011) from the public ADNI database (http://www.loni.ucla.edu/ADNI/). Here we analyzed all ADNI subjects as a single group, to exploit the broader phenotypic continuum (Petersen 2000) and increase power (Durston et al. 2005; Stein et al. 2010). Some subjects were excluded to eliminate problems caused by population stratification (Lander and Schork 1994; McCarthy et al. 2008) using multi-dimensional scaling as outlined previously for the same dataset (Stein et al. 2010). The remaining sample had 742 Caucasian subjects left for the analysis; this represented the largest homogeneous group attainable from the ADNI cohort. After removing subjects by applying quality control criteria to the lentiform nucleus segmentations (discussed below), the final ADNI dataset consisted of 706 subjects (average age± s.d.: 75.5±6.8 years; 413 men/293 women) including 162 patients with AD (75.6±7.6 years; 88 men/74 women), 346 with mild cognitive impairment or MCI (75.0±7.3 years; 221 men/125 women), and 198 healthy elderly controls (76.1± 4.8 years; 104 men/94 women).

The Queensland Twin Imaging Study (QTIM) is an ongoing, 5-year longitudinal project tasked with identifying genetic influences on brain structure. As of August 1, 2011 there were 672 subjects with both structural MRI scans and genome-wide genotyping data. An additional 40 subjects underwent repeated scans, which we used to verify the reliability of the segmentation algorithm. Each member of every twin pair and their siblings was assessed via extensive diagnostic interviews to exclude anyone with a history of brain related disorders, diseases, or injuries. To avoid problems caused by population stratification in this Caucasian sample, 10 subjects were removed from the QTIM sample, as determined by MDS analysis. All subjects were right-hand dominant as determined by Annett's Handedness Questionnaire (Annett 1970). Written informed consent was obtained from all participants before protocol-specific procedures were performed. After quality control of the lentiform nucleus segmentations (discussed below), the final group we analyzed consisted of 639 subjects from 364 families (98 monozygotic twin pairs; 127 dizygotic twin pairs; 3 dizygotic triplet trios; 117 singletons; 63 siblings; 23.1±3.1 years; 251 men and 388 women).

Genotyping and imputation

Genome-wide genotype data were collected using the Human610-Quad BeadChip (Illumina, Inc., San Diego, CA) in both samples. Several SNPs were excluded from the analysis based on standard filtering criteria, as is standard in many other GWAS studies (Wellcome Trust Case Control Consortium 2007). In the ADNI sample, SNPs were excluded based on: call rate <95 % (42,176 SNPs removed), significant deviation from Hardy-Weinberg equilibrium $P<1\times10^{-6}$ (263 SNPs removed), minor allele frequency <0.01 (60,919 SNPs removed), autosomal chromosomes only, and a platform-specific quality control score of <0.15 to eliminate "no call" genotypes (variable number of missing genotypes across subjects). Similarly, in the QTIM sample SNPs were excluded based on: call rate <95 % (8,447 SNPs removed), significant deviation from Hardy-Weinberg equilibrium $P<1\times10^{-6}$ (2,841 SNPs removed), minor allele frequency <0.01 (33,347 SNPs removed), and a platform-specific quality control score of <0.07 to eliminate "no call" genotypes (variable number of missing genotypes across subjects). Also, we chose to focus only on autosomal SNPs rather than those in mitochondrial DNA and sex chromosomes.

Imputation of hard genotype calls may be used to infer missing values based on the linkage among SNP sets. In addition, imputation may be used to infer SNPs not directly genotyped in a given sample, but genotyped in a reference dataset. The quality of imputed SNPs depends on the strength of the linkage between hard genotyped SNPs and the imputed SNPs. In ADNI, we excluded any SNPs imputed with an R^2 value of <0.3 from the analysis (62,053 SNPs removed); the same steps were taken for the QTIM (54,337 SNPs removed). We also excluded imputed SNPs with a minor allele frequency of <0.01 for both the ADNI (45,818 SNPs removed) and QTIM (38,481 SNPs removed). For each sample, we performed imputation with MaCH, which uses the Markov Chain Monte Carlo method to infer missing genotypes and SNPs robustly and accurately (Abecasis

et al. 2010). After all rounds of quality control filtering, the ADNI dataset had 2,449,382 SNPs and the QTIM dataset had 2,439,807 SNPs. We analyzed the set of overlapping SNPs present in both datasets, totaling 2,380,200 SNPs.

Image acquisition and pre-processing

High-resolution structural brain MR images were collected from both the ADNI and QTIM samples. Structural MRI scans in the ADNI study were obtained using a standardized protocol to maximize consistency across 58 image acquisition sites, using 1.5 Tesla MRI scanners. A T1-weighted 3D MP-RAGE sequence was used (TR/TE=2400/1000 ms; flip angle=8°; FOV=24 cm; with a final voxel resolution= $0.9375 \times 0.9375 \times 1.2$ mm^3).

In the QTIM cohort, structural MRI scans were obtained on a single 4 Tesla scanner (Bruker Medspec). T1-weighted images were acquired with an inversion recovery rapid gradient echo sequence (TI/TR/TE=700/1500/3.35 ms; flip angle=8°; slice thickness=0.9 mm, with a 256×256 acquisition matrix; with a final voxel resolution= $0.9375 \times 0.9375 \times 0.9$ mm^3).

Skull and all other non-brain tissues were removed from each subject's scan using the brain extraction tool (Smith 2002) (BET) prior to analysis. Test-retest data were also available for 40 young normal individuals scanned on two occasions approximately 4 months apart.

Automated delineation of lentiform nucleus volume

We delineated the lentiform nucleus structures using the well-validated, automated FIRST segmentation algorithm (http://www.fmrib.ox.ac.uk/fsl/first/index.html), which is part of the FSL (Smith et al. 2004) image processing package. Morey et al. (2010) showed that the FIRST segmentation algorithm has relatively high reproducibility and accuracy for each of the subcortical structures segmented. Using a Bayesian framework, FIRST provides accurate and validated segmentations of subcortical brain structures (Patenaude et al. 2011).

Quality control of lentiform nucleus segmentation

The quality of segmentations was assessed by examining the left and right globus pallidus and putamen separately; these were checked by hand (by DPH) following established guidelines (Duvernoy and Bourgouin 1999). If any segmentation did not properly delineate any single structure the subject was removed from the analysis. After quality checking, 36 subjects were excluded from the ADNI sample and 23 subjects were excluded from the QTIM sample. As a further measure of segmentation quality, we examined the consistency of independent runs of the FIRST algorithm on repeated scans of 40 subjects, taken a short interval apart. The reliability of lentiform nucleus segmentation was tested by computing intraclass correlation coefficients (ICC) from the 40 repeated scans. All ICC calculations were performed using the *psy* package in the R statistical software (version 2.13.0; http://www.r-project.org/).

Heritability analysis

To evaluate the overall genetic contribution to the variability in volumes, the heritability of the lentiform nucleus volume was estimated using a structural equation model (SEM) as implemented in the software package, Mx (version 1.68; http://www.vcu.edu/mx/). Heritability was estimated using the classic pathway-based 'ACE' model (Chiang et al. 2012; Neale et al. 1992). In families from the full QTIM sample, we used this analysis to compare the observed pattern of covariances in lentiform nucleus volume to what would be expected given different degrees of genetic influence. The heritability of the left, right, and average bilateral lentiform nucleus volumes were analyzed separately. We chose to study the average bilateral lentiform nucleus because it shows higher heritability than the left and right lentiform nucleus separately and because the inevitable segmentation errors should be a smaller proportion of the total volume if both sides are combined. Additionally, we estimated the genetic correlation (r_g) between the putamen and globus pallidus volumes using Mx in the full QTIM sample.

Genetic analysis

In the ADNI sample, we tested each SNP dosage value for association with the lentiform nucleus volume, assuming, by default, an additive model - each SNP dosage value was recorded as the number of minor alleles, with an implicit correction for the accuracy of imputation at that SNP. Tests of association were conducted using linear regression as implemented in the publicly available program, *mach2qtl* (Abecasis et al. 2010). We controlled for age and sex, which both showed significant effects on lentiform nucleus volume. We also covaried for age^2, sex x age, and sex x age^2 to account for any quadratic or interaction effects. In addition we controlled for intracranial volume (ICV), calculated as 1/(determinant) of the transformation matrix from registration to the FSL common template. We chose to correct for head size because we are interested in individual differences in lentiform nucleus volume unrelated to differences in head size (Buckner et al. 2004). In the QTIM sample, association testing was carried out using mixed-model regression, to control for family structure. We also included the same covariates as in the ADNI model. Because of the kinship structure of the twin sample, association tests in QTIM were conducted using the family-based association test implemented in *merlin* (Chen and Abecasis 2007).

Meta-analysis of genetic results

Genome-wide association results from the ADNI and QTIM samples were meta-analyzed using a fixed effects inverse variance-weighted method, as implemented in METASOFT (Han and Eskin 2011). *Beta* coefficients from the regression analysis of each SNP from both studies were pooled based on the inverse of the variance of each *beta* coefficient. In addition to the standard fixed-effects meta-analysis, we performed a random-effects meta-analysis in METASOFT. The random-effects meta-analysis still follows the inverse variance-weighted model, but can more appropriately model the population statistics in cases where the effect size is not the same across cohorts (Han and Eskin 2011).

Gene-based tests and pathway enrichment analysis

The meta-analyzed P_{MA}-values from the full set of SNPs from the GWAS analysis were used for gene-wide, gene-based association testing with the software package KGG (Li et al. 2010). No prioritizing or pre-selection of genes was performed. Gene-based tests in KGG combine univariate association statistics to evaluate the cumulative evidence of association in a gene with a phenotype, using the GATES-Simes test (Li et al. 2011). Similarly, KGG is integrated with biological pathway databases (e.g., KEGG) and combines gene sets to test for significant enrichment of a number of disease and biological pathways (Li et al. 2011). Pathways are considered to be significantly enriched if they contain more significant gene-based test statistics than expected by chance.

Results

Lentiform nucleus segmentations

In the ADNI sample, the volumes of the left (6422.2±723.9 mm^3) and right (6450.9±686.9 mm^3) lentiform nucleus were highly correlated (r=0.83; P<0.0001; df=704). Similarly, in the QTIM sample the volumes of the left (6554.4±744.8 mm^3) and right (6729.5±765.4 mm^3) lentiform nucleus were highly correlated (r=0.84; P<0.0001; df=637). Both samples have a slight asymmetry between left and right lentiform nucleus volume. In the ADNI sample the right lentiform nucleus was 0.4 % larger on average than the left. Similarly, in the QTIM sample the right lentiform nucleus was 2.6 % larger on average than the left. This follows a general trend in the brain where bilateral subcortical structures are slightly larger in the right hemisphere (Toga and Thompson 2003). As expected, because the cohort is younger, the average volumes for the QTIM sample were larger than the ADNI sample (Left: t=3.30; P=0.0010; Δ2.0 %; Right: t=7.00; P<0.0001; Δ4.1 %; Average Bilateral: t=5.37; P<0.0001; Δ3.1 %).

Reliability of lentiform nucleus segmentation

To examine how reliable the automated segmentations were, when measured by FIRST, we obtained repeated scans for 40 subjects from the QTIM sample (time between scans: 120±55 days) and applied the FIRST algorithm to each scan separately. Using the intraclass correlation coefficient (ICC), we found the FIRST segmentations to be very highly reliable for the left (ICC=0.922), right (ICC=0.890), and average bilateral (ICC=0.928) lentiform nucleus volumes (Fig. 2).

Heritability of lentiform nucleus volume

Using twin and family data from the QTIM, we modeled the additive genetic effects (A), effects of the common environment shared by both twins (C), and unique environment effects and experimental error (E). The components of the 'ACE' model are used to estimate the amount of variance in a measure that can be ascribed to purely genetic influences (its heritability). Lentiform nucleus volume is highly heritable (between 70 and 80 %) as has been found for many other structures in the

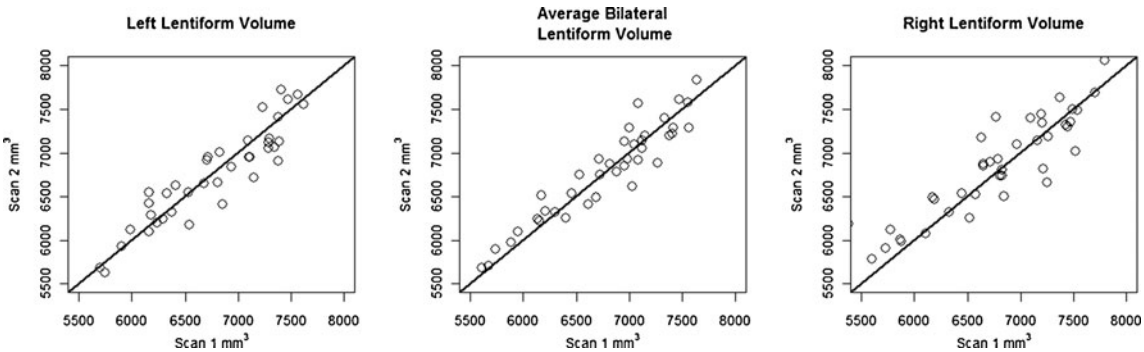


Fig. 2 Scan and re-scan volume values of the lentiform nucleus delineated using the automated FIRST segmentation algorithm. The black line on the diagonal represents the ideal situation where the segmented volumes are identical. In general, there is good agreement between the segmented volumes for both scans

brain (Kremen et al. 2010) (Table 1). The heritability of the lentiform nucleus volume is also evident in a scatterplot showing the similarity among twin pairs, monozygotic twins (*black dots*) have more similar lentiform nucleus volumes in general compared to dizygotic twins (*open dots*; Fig. 3). The genetic correlation (r_g) is the proportion of the observed variance between two traits that can be explained by common genetic influences (Neale et al. 1992). As the structures of the lentiform nucleus are closely related, we expect them to share common genetic determinants. Indeed, the genetic correlation between the putamen and globus pallidus was high: r_g=0.54 (95 % CIs 0.39, 0.82) for the left and r_g=0.56 (95 % CIs 0.40, 0.69) on the right. In addition, the genetic correlation between the left and right lentiform nucleus reveals that the volume of the structure on each side has almost perfect overlap in its genetic determinants: r_g=0.93 (95 % CIs 0.88, 1.00).

Genome-wide association testing

As the lentiform nucleus is involved in a number of brain disorders and its volume is heritable, we conducted genome-wide tests of association on a large set of SNPs from the two independent cohorts to identify genetic variants that help to explain the considerable genetic influence on lentiform nucleus volume. Q-Q plots of the distribution of *P*-values for each individual sample show that the association statistics are approximately Normal (Fig. 4). Genomic inflation factor values (lambda) indicate that the distribution of *P*-values is unbiased and that the results are not likely to be attributable to population stratification.

Meta-analysis

Test statistics from each study were combined meta-analytically to increase the power to detect real effects and to reduce false positives. *Beta* values, and their standard error, for SNPs from each regression model were combined across samples. The signs of *Beta* values were determined based on the reference allele in each study and combined using a fixed-effects, inverse variance-weighted meta-analysis (Han and Eskin 2011). Meta-analysis is preferred in this case, as opposed to combining all subjects into a single *combined analysis*, as the two samples have very different age distributions, image acquisition parameters, and the QTIM is a family-based study that requires complex regression methods (to account for kinship).

In the Manhattan plot of the *P*-values from each meta-analysis, a number of promising genetic variants were associated with lentiform nucleus volume, including one SNP that exceeds the standard, nominal genome-wide significance level $P<5\times10^{-8}$ after meta-analysis (Fig. 5). A list of the top SNPs from each analysis with a meta-analyzed *P*-value (P_{MA}) threshold of $P_{MA}<1\text{x}10^{-6}$ is given in Table 2.

A broad band of SNPs shows high association with lentiform nucleus volume in the flavin-containing monooxygenase gene cluster on chromosome 1 (Fig. 6). The most highly associated SNP, rs1795240, is located approximately 5 Kb outside of the flavin-containing monooxygenase 3 (*FMO3*) gene. It shows genome-wide significant associations with lentiform nucleus volume ($P_{MA}=4.79\times10^{-8}$). Individual association statistics for rs1795240 show significance in both the ADNI ($\beta=-143.48$; SE(β)=28.15; minor allele=A; *P*=

Table 1 Heritability estimates (a^2) for lentiform nucleus volume. These analyses were run in Mx on 637 individuals (i.e., including up to 3 individuals per family so two non-twin siblings who were the 4th family member were not included). Data were winsorised to ±3.3SD. Sex and age were included as covariates

Model	-2LL	df	Δ-2LL	Δdf	AIC	a^2 (95%CI)	c^2 (95%CI)	e^2 (95%CI)
Avg. Lent.								
ACE	1447.604	631	–	–	185.6	0.78 (0.60, 0.84)	0.00 (0.00, 0.18)	0.22 (0.16, 0.29)
AE	1447.604	632	0.00	1	183.6	0.78 (0.71, 0.84)	–	0.22 (0.16, 0.29)
CE	1486.895	632	39.3	1	222.9	–	0.52 (0.43, 0.59)	0.48 (0.41, 0.57)
E	1580.518	633	132.9	2	314.5	–	–	1.00 (1.00, 1.00)
Left Lent.								
ACE	1475.217	631	–	–	213.2	0.76 (0.57, 0.82)	0.00 (0.00, 0.17)	0.24 (0.18, 0.32)
AE	1475.217	632	0.00	1	211.2	0.76 (0.68, 0.82)	–	0.24 (0.18, 0.32)
CE	1508.829	632	33.6	1	244.8	–	0.50 (0.40, 0.58)	0.50 (0.42, 0.60)
E	1595.305	633	120.1	2	329.3	–	–	1.00 (1.00, 1.00)
Right Lent.								
ACE	1498.969	631	–	–	237.0	0.71 (0.52, 0.78)	0.00 (0.00, 0.16)	0.29 (0.22, 0.39)
AE	1498.969	632	0.00	1	235.0	0.71 (0.61, 0.78)	–	0.29 (0.22, 0.39)
CE	1526.486	632	27.5	1	262.5	–	0.46 (0.36, 0.55)	0.54 (0.45, 0.64)
E	1595.929	633	97.0	2	329.9	–	–	1.00 (1.00, 1.00)

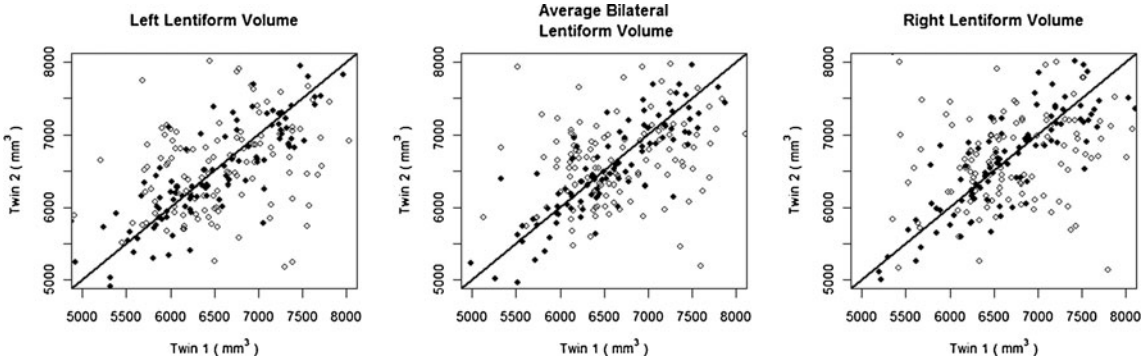


Fig. 3 Scatterplot of lentiform nucleus volume in monozygotic (*black dots*) and dizygotic (*open dots*) twin pairs from the QTIM. Data points closer to the diagonal line represent similar lentiform nucleus volumes across a given twin pair. In general, the lentiform nucleus volumes in the monozygotic twins are closer than their dizygotic counterparts, which is a sign of genetic influence (confirmed by our heritability analysis)

3.46×10^{-7}) and QTIM ($\beta=-76.57$; SE(β)=30.12; minor allele=A; $P=0.011$) samples. The observed effect is likely greater in the ADNI sample due to the greater sample size, but the effect may also increase with age or disease. Additionally, the second most associated SNP, rs1795243 ($P_{MA}=8.76\times10^{-8}$), lies in an untranslated region of the *FMO6P* pseudogene. The variant rs1795243 shows strong evidence for association in both samples (ADNI: $\beta=-141.44$; SE(β)=28.53; minor allele=C; $P=7.12\times10^{-7}$; QTIM: $\beta=-76.805$; SE(β)=30.20; minor allele=C; $P=0.011$). Additionally, a number of the top hits were located in *GATAD2B* and *EPB41L2* among others (detailed in Table 2). After controlling for diagnosis in the ADNI sample, there was little change in observed *P*-values (See the P_{diag} column in Table 2). Two dummy variables were added as covariates to the regression model to account for each of the three different diagnostic categories in the ADNI sample. This was not necessary in the QTIM sample, as they are all healthy young adults. The random effects meta-analysis of these same SNPs gave nearly identical results to the fixed-effects meta-analysis (see Fig. 7).

Gene-based tests and pathway analysis

The genes *FMO3* ($P=1.03\times10^{-6}$) and *FMO6P* ($P=1.32\times10^{-6}$) exceed the nominal gene-wide significance level of $P<5\times10^{-6}$. A number of other genes show promising evidence of association with lentiform nucleus volume: *SLC39A1* ($P=7.56\times10^{-6}$), *DENND4B* ($P=1.53\times10^{-5}$), *GATAD2B* ($P=2.25\times10^{-5}$), and *FOXF2* ($P=9.59\times10^{-5}$). Pathway enrichment analysis in KGG reveals seven pathways that exceed the nominal significance level for pathway enrichment (5×10^{-4}) including the *reactome phase 1 functionalization* pathway ($P=1.34\times10^{-5}$) and the *KEGG drug metabolism pathway of cytochrome P450* ($P=5.66\times10^{-5}$). Additional results of the pathway analysis are given in Table 3.

Discussion

In this study, we identified specific genetic variants associated with differences in lentiform nucleus volume in two large independent samples, including both young and elderly subjects ($N=1345$). We were well powered to find genetic variants that explain some of the heritability of the lentiform nucleus volume, with one SNP exceeding the nominal genome-wide significance threshold. Our two cohorts differed in many ways, but mainly in mean age (50 years). Despite the differences, we identified a number of variants with compelling evidence for association in both samples.

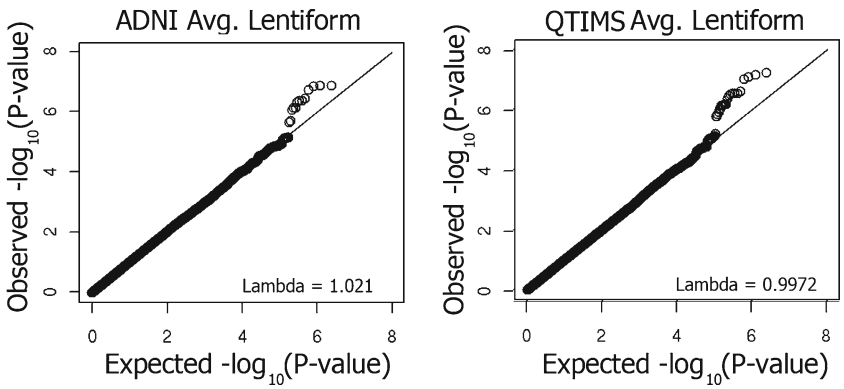


Fig. 4 Q-Q plots for observed association *P*-values of SNPs from both datasets (after removing poorly imputed SNPs and SNPs with a minor allele frequency below 0.01). The genomic inflation factor (lambda) is given for each measure (*inset*). There is no evidence of inflated *P*-values influencing the meta-analysis

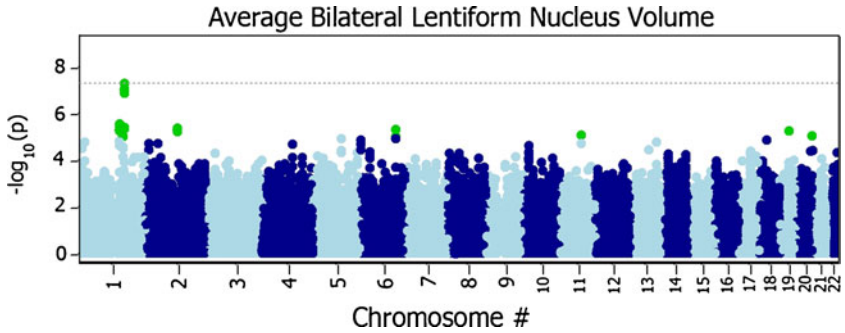


Fig. 5 Manhattan plot of meta-analyzed *P*-values (P_{MA}) from both the ADNI and QTIM samples (N=1345). Each plotted point is the $-\log_{10}(P_{MA})$ of a given SNP sorted by chromosome. The dotted gray line denotes the standard, nominal genome-wide significance level $-\log_{10}(5\times10^{-8})$. Each point plotted above the gray line indicates a genome-wide significant SNP

Associations were detectable despite differences in study protocols; the genes implicated may therefore have a statistical effect on lentiform nucleus volume throughout life. Further replication in independent samples (e.g. as in Stein et al. 2012 and Bis et al. 2012) and examination of functional relevance will still be required to further support a causal role for these variants.

We originally chose to study the lentiform nucleus as it is implicated in several genetically mediated disorders including Parkinsonian syndromes, Huntington's disease, Wilson's disease, Tourette's syndrome, and ADHD. While the putamen is more similar to the caudate histologically, the putamen and globus pallidus are linked by dense intrabasal ganglionar fiber projections. In addition, the genetic correlations between the two structures of the lentiform nucleus were very high (r_g=0.56 and r_g=0.54, for left and right, respectively). This high genetic correlation means that the two structures share many common genetic determinants. This provided empirical support for analyzing the two structures together, in addition to our theoretical reasons for choosing to study the lentiform nucleus. Even so, we note that other natural groupings of structures may be beneficial for future assessment. Although we opted to combine the putamen and globus pallidus, the putamen is functionally more related to the caudate, and together they make up the striatum, which receives afferent projections from large parts of the cortex. In the future, when a broad range of subcortical segmentations are available in large family-based samples, it will be possible to perform genetic clustering to determine logical groupings of subcortical nuclei with coherent genetic determination (C. H. Chen et al. 2012; Chiang et al. 2012). By clustering regions with overlapping genetic determinants, it should be possible to boost the power to detect underlying genetic determinants via GWAS (as shown by (Chiang et al. 2012). In addition, variance component modeling performed in the QTIM sample shows that the left and right lentiform nucleus volume are around 70–80 % heritable (see Table 1). This agrees with published heritability estimates for the putamen and globus pallidus (Kremen et al. 2010; Peper et al. 2007). We examined the reliability of lentiform nucleus segmentations by processing repeated scans in 40 subjects from the QTIM sample. The resulting volumes were highly reliable using the automated FSL FIRST software (Patenaude et al. 2011)(Fig. 2), and the reproducibility also agrees with prior estimates (Morey et al. 2010).

A wide band of SNPs from the flavin-containing monooxygenase (*FMO*) gene cluster on chromosome 1 showed significant evidence of association in both samples and after meta-analysis, with one SNP exceeding the nominal genome-wide significance level. The *FMO* gene cluster consists of five tightly-spaced genes (*FMO1-4* and *FMO6P*) responsible for the metabolism of trimethylamine, methionine, and cysteamine as well as a number of therapeutic medications including *tamoxifen*, *ranitidine*, *sulindac*, and *itopride* (Williams et al. 2004). Additionally, the *FMO* gene cluster is involved in the oxidation of certain environmental toxicants like insecticides and aldicarb (Krueger and Williams 2005). Of the genes in the *FMO* gene cluster, *FMO1* and *FMO3* have been studied extensively. Carriers of a number of common genetic variants have reduced efficacy metabolizing certain drug substrates (Koukouritaki et al. 2002; Overby et al. 1997; Yeung et al. 2000). The role of the *FMO* gene cluster in the metabolism of common environmental toxicants suggests a common underlying mechanism that might yield association results in the young, healthy population of twins that overlaps with association we found in our sample of elderly controls and patients. It is unlikely that the association in these samples were driven by the use of therapeutic medications such as opiates or anti-depressants, as most participants were healthy. Follow up studies are still needed to determine whether commonly used substances, such as alcohol, nicotine, commonly abused drugs, or anti-inflammatory drugs exert detectable and systematic anatomical effects on structures in the reward circuitry, and if they lead to any detectable changes in *FMO* gene expression.

The most highly associated SNP, rs1795240, is located just downstream of the *FMO3* gene. A number of common

Table 2 Top SNPs identified from the meta-analysis ($P_{MA}<1\times10^{-6}$). Individual significance statistics of the average bilateral lentiform nucleus are given for both the ADNI and QTIM samples as well as meta-analyzed statistics (Pooled). The R^2 value gives the estimated Pearson's correlation coefficient of a SNP to the nearest genotyped marker. The minor allele (α) for the reported beta coefficient (β) for each sample is also given. Intergenic SNPs within 20 Kb of an annotated gene region are listed as part of that gene. P_{diag} gives the association *P*-value for a given SNP when controlling for diagnosis in the ADNI sample

				ADNI					QTIM				Pooled	
Chr	SNP	R^2	Gene	Frq	α	β	*P*	P_{diag}	Frq	α	β	*P*	β	P_{MA}
1	rs1795240	Genotyped	FMO3	0.43	A	−143.48	3.46E-07	7.49E-06	0.46	A	−76.57	0.011	−112.27	4.79E-08
1	rs1795243	0.97	FMO6P	0.43	C	−141.44	7.12E-07	1.37E-05	0.46	C	−76.81	0.011	−110.96	8.76E-08
1	rs1795244	0.97	FMO6P	0.46	C	−151.07	1.29E-07	1.57E-06	0.49	T	−66.01	0.031	−111.33	9.77E-08
1	rs1736567	0.97	FMO6P	0.46	C	−151.10	1.29E-07	1.58E-06	0.49	G	−66.02	0.031	−111.34	9.81E-08
1	rs2075992	1.00	FMO3	0.43	C	−144.37	4.08E-07	1.34E-05	0.45	C	−70.34	0.020	−109.50	1.28E-07
1	rs1736556	1.00	FMO3	0.43	C	−144.41	4.04E-07	1.34E-05	0.45	C	−69.92	0.021	−109.34	1.33E-07
1	rs7414227	1.00	GATAD2B	0.49	G	−84.58	0.0042	0.0037	0.49	G	−115.46	0.00016	−99.50	2.78E-06
1	rs7522030	0.98	GATAD2B	0.49	C	−85.67	0.0040	0.0034	0.49	C	−114.64	0.00019	−99.72	3.06E-06
1	rs11264573	1.00	GATAD2B	0.48	A	−85.10	0.0040	0.0036	0.49	A	−113.67	0.00020	−98.93	3.17E-06
1	rs9427232	1.00	GATAD2B	0.49	G	−81.71	0.0057	0.0049	0.49	A	117.41	0.00012	−98.95	3.18E-06
1	rs11264736	0.99	SLC39A1	0.47	T	−79.29	0.0074	0.0051	0.48	T	−119.65	0.00010	−98.69	3.76E-06
1	rs4916390	Genotyped	FMO6P	0.47	T	−117.48	2.72E-05	0.00024	0.48	T	−68.66	0.024	−95.04	3.89E-06
1	rs9426938	1.00	GATAD2B	0.49	C	−79.79	0.0069	0.0059	0.49	C	−117.55	0.00012	−98.03	3.92E-06
2	rs10496584	0.96	Intergenic	0.11	G	165.94	0.00034	0.00099	0.11	G	149.68	0.0037	158.69	4.10E-06
1	rs1920149	1.00	FMO3	0.47	A	−121.46	1.51E-05	0.00018	0.48	A	−63.78	0.036	−94.95	4.18E-06
1	rs12076145	1.00	FMO3	0.47	G	−121.68	1.49E-05	0.00018	0.48	G	−63.33	0.038	−94.94	4.44E-06
1	rs10911192	1.00	FMO3	0.47	C	−121.62	1.50E-05	0.00018	0.48	C	−63.30	0.038	−94.89	4.45E-06
6	rs6941712	Genotyped	EPB41L2	0.13	C	−72.28	0.11	0.17	0.12	C	−235.78	6.83E-07	−149.67	4.61E-06
1	rs1127091	Genotyped	GATAD2B	0.49	A	−78.02	0.0082	0.0069	0.49	A	−117.60	0.00012	−97.12	4.76E-06
2	rs2037892	0.97	Intergenic	0.11	A	165.50	0.00034	0.00095	0.11	A	145.91	0.0043	156.70	4.79E-06
1	rs1139620	Genotyped	GATAD2B	0.49	C	−77.66	0.0085	0.0071	0.49	C	−117.60	0.00012	−96.93	4.95E-06
2	rs1919922	Genotyped	Intergenic	0.11	T	164.79	0.00034	0.00090	0.11	T	144.16	0.0046	155.52	5.17E-06
1	rs9426935	Genotyped	LOC343052	0.49	C	−77.11	0.0089	0.0074	0.49	C	−117.61	0.00012	−96.63	5.25E-06
1	rs4341393	Genotyped	LOC343052	0.49	C	−77.11	0.0089	0.0074	0.49	C	−117.60	0.00012	−96.63	5.25E-06
19	rs11083866	0.92	Intergenic	0.26	A	114.32	0.00074	0.0016	0.26	A	115.32	0.0023	114.77	5.38E-06
1	rs2252508	0.97	DENND4B	0.47	G	−82.13	0.0065	0.0049	0.48	G	−113.85	0.00020	−97.78	5.39E-06
2	rs4848768	0.99	Intergenic	0.11	G	163.94	0.00035	0.00087	0.11	G	143.10	0.0049	154.60	5.71E-06
11	rs17140547	0.75	Intergenic	0.02	T	−409.41	0.0082	0.039	0.01	T	−711.52	0.00014	−532.56	7.96E-06
20	rs6027511	Genotyped	Intergenic	0.17	C	−121.95	0.00076	0.00089	0.17	C	−122.62	0.0038	−122.24	8.96E-06
1	rs10918196	Genotyped	LOC400794	0.41	C	120.25	3.45E-05	0.00061	0.40	C	63.93	0.039	93.94	9.31E-06
1	rs10737511	0.98	LOC400794	0.41	C	120.34	3.47E-05	0.00062	0.40	C	63.95	0.039	93.97	9.39E-06
6	rs12530288	1.00	EPB41L2	0.13	T	−62.45	0.16	0.23	0.12	T	−235.80	6.82E-07	−144.17	9.76E-06

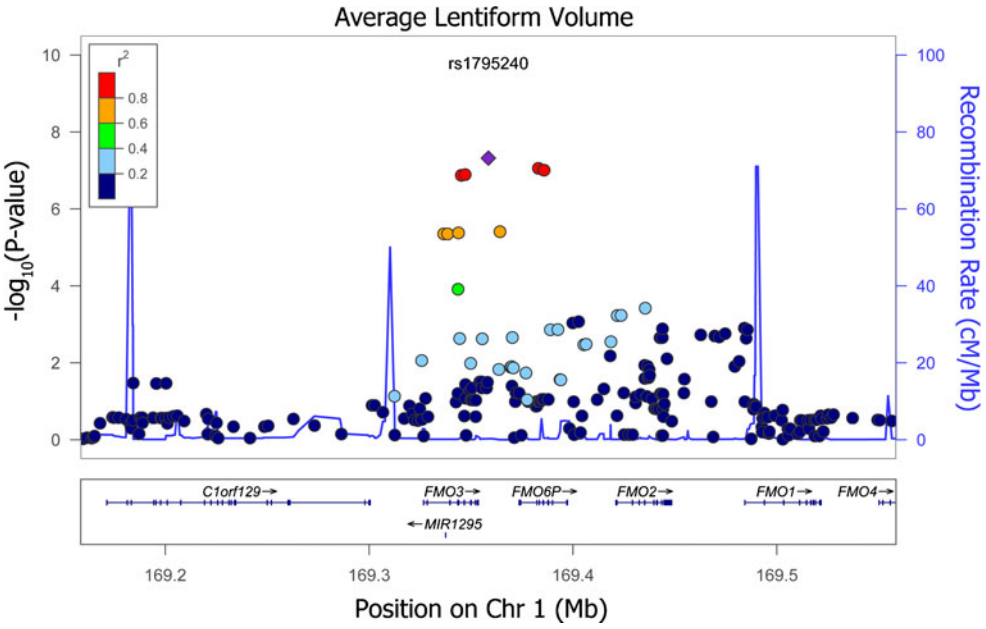


Fig. 6 Detailed view of the flavin-containing monooxygenase gene cluster. Points correspond to the $-\log_{10}(P_{MA}$-value) for the average lentiform nucleus volume. The colors of each point correspond to level of linkage disequilibrium (LD) between a given SNP and the most associated SNP rs1795240. Plots were generated using the LocusZoom software (http://csg.sph.umich.edu/locuszoom/)

genetic variants in the *FMO3* gene have been linked with decreased catalytic activity and the disorder trimethylaminuria (Hines 2006; Koukouritaki et al. 2007). The *FMO3* gene is expressed mainly in the liver but also in the human brain, and may affect how numerous therapeutic drugs are metabolized by the central nervous system (Cashman and Zhang 2002). In addition, an analysis of the Allen Brain Atlas (http://human.brain-map.org/) shows that *FMO3* is differentially expressed in the posterior portion of the lentiform nucleus (Fig. 8).

The second most highly associated SNP, rs1795243, was found in the *FMO6P* pseudogene, which is transcribed into mRNA, but not translated into a protein product (Hines et al. 2002). Pseudogenes are not ultimately converted to proteins, but can act as regulatory elements and are under evolutionary control (Poliseno et al. 2010; Wen et al. 2011). The exact mechanism of action of the *FMO6P* pseudogene is still unknown, but the associations identified in this study may make it an ideal candidate for future genetic studies of neurodegenerative disorders and functional tests of *FMO6P* mechanistic effects. Previously, a large case-control GWAS study found mild evidence of association of the *FMO6P* gene with schizophrenia (Athanasiu et al. 2010).

Gene-based test statistics confirmed the association of *FMO3* and *FMO6P* with lentiform nucleus volume—as found in the univariate study—with both genes exceeding the nominal gene-wide significance level. The gene-based tests also promoted *SLC39A1* to a higher significance level ($P=7.56\times10^{-6}$) than might be expected compared to the other genes in the univariate SNP GWAS. The role of *SLC39A1* is very well studied. It is expressed in the brain and is involved in maintaining an appropriate zinc concentration inside the blood-brain barrier (Bobilya et al. 2008). Pathway enrichment analysis, performed with KGG, combines gene-based test statistics to examine whether known disease and biological pathways are over-represented in the

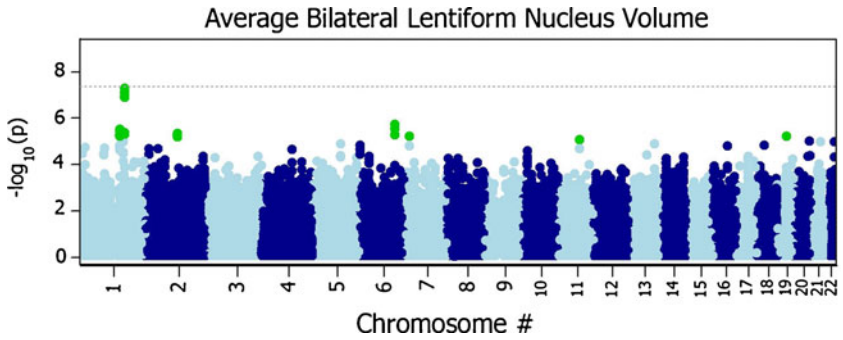


Fig. 7 Manhattan plot of meta-analyzed *P*-values (P_{MA}) from both the ADNI and QTIM samples ($N=1345$) using a *random effects model* (Han and Eskin 2011). Each plotted point is the $-\log_{10}(P_{MA})$ of a given SNP sorted by chromosome; points plotted higher on the *y*-axis are more significant. The dotted grey line denotes the nominal genome-wide significance level $-\log_{10}(5\times10^{-8})$. The results are consistent with those found using the standard fixed effects analysis

Table 3 Significantly enriched pathways determined by pathway analysis with KGG (Li et al. 2010, 2011). Details for the pathways given can be found of the Gene Set Enrichment Analysis website (http://www.broadinstitute.org/gsea/). Pathways that exceed the threshold $P<5\times10^{-4}$ were considered to be significantly enriched

Pathway Name	Pathway P-value	Gene	Gene *P*-value	Chr	Length (bp)	SNP#
REACTOME_PHASE_1_FUNCTIONALIZATION	1.34×10^{-5}	FMO3	1.03×10^{-6}	1	46941	40
KEGG_DRUG_METABOLISM_CYTOCHROME_P450	5.66×10^{-5}	FMO1	0.0073	1	57450	48
–	–	FMO2	0.0077	1	47434	44
–	–	FMO3	1.03×10^{-6}	1	46941	40
REACTOME_PHASE_1_FUNCTIONALIZATION_ OF_COMPOUNDS	6.49×10^{-5}	FMO1	0.0073	1	57450	48
–	–	CYP51A1	0.028	7	42596	5
–	–	FMO2	0.0077	1	47434	44
–	–	FMO3	1.03×10^{-6}	1	46941	40
REACTOME_ZINC_INFLUX_INTO_CELLS_ BY_THE_SLC39_GENES_FAMILY	6.81×10^{-5}	SLC39A1	7.56×10^{-6}	1	28600	2
REACTOME_BIOLOGICAL_OXIDATIONS	1.10×10^{-4}	MAT2A	0.034	2	26115	5
–	–	FMO1	0.0073	1	57450	48
–	–	CYP51A1	0.028	7	42596	5
–	–	FMO2	0.0077	1	47434	44
–	–	FMO3	1.03×10^{-6}	1	46941	40
–	–	NNMT	0.039	11	36703	16
REACTOME_ZINC_TRANSPORTATION	1.21×10^{-4}	SLC39A1	7.56×10^{-6}	1	28600	2
REACTOME_METAL_ION_SLC_TRANSPORTERS	1.66×10^{-4}	SLC39A1	7.56×10^{-6}	1	28600	2

gene sets from our analysis, relative to what might be expected by chance. In all, there were seven biological pathways that reached significance (see Table 3). The most significant pathway, *reactome phase 1 functionalization*, supports the many studies suggesting that the *FMO3* gene is involved in processing environmental toxins (Krueger and Williams 2005; Williams et al. 2004). The next most significant pathway, *KEGG drug metabolism pathway of cytochrome P450*, involves the cytochrome P450 superfamily of enzymes responsible for metabolizing numerous drugs including codeine, morphine, carbamazepine, citalopram, and clozapine (Hines et al. 2008). Cashman and Zhang showed that *FMO3* is expressed in various regions throughout the brain including in the striatum (Cashman and Zhang 2002). Earlier studies using human microsomes showed that numerous brain tissues actively metabolize psychoactive drug substrates including chlorpromazine, imipramine and fluoxetine (Bhagwat et al. 1996; Bhamre et al. 1995). In addition, several positron emission tomography studies have demonstrated significant differences in glucose metabolism

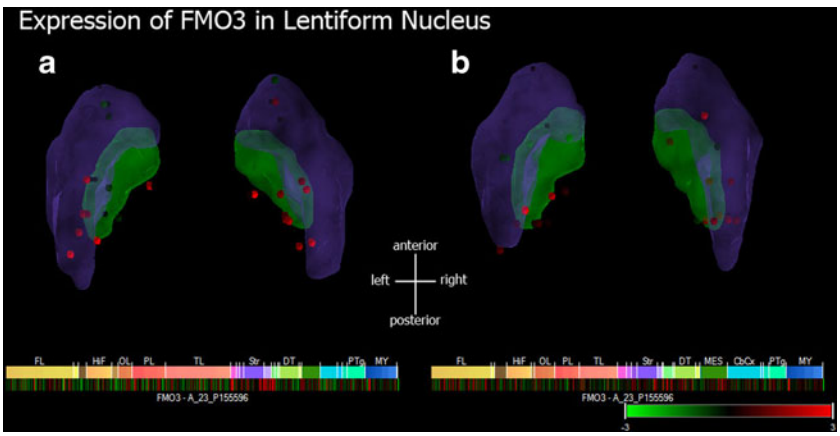


Fig. 8 Expression levels of *FMO3* gene in the lentiform nucleus of two different subjects (**a** and **b**; details can be found at http://human.brain-map.org/). Expression levels were standardized to a mean expression level to eliminate background noise and are presented here as Z-scores (where |Z-score|>2.5 indicate evidence of differential expression of the *FMO3* compared to other regions of the brain). Numerous points in the posterior portion of the lentiform nucleus show evidence of significant differential expression of the *FMO3* gene

in the lentiform nucleus in patient-versus-control comparisons of psychoactive drugs like fluoxetine and chlorpromazine (Chen et al. 2009; Mayberg et al. 2000; Wik et al. 1989). Each of these studies lends credibility to the findings in this study and future endeavors to further understand the mechanisms by which gene variants in the *FMO* gene cluster may influence lentiform nucleus volume.

Several weaknesses of our study should be mentioned. First, we provide evidence for association of genetic variants in the flavin-containing monooxygenase (*FMO*) cluster but we do not yet know the mechanistic means by which they may change expression levels or protein structures, or how they might affect lentiform nucleus volumes. Unfortunately, functional and expression data are not yet available for either cohort, but they may be available in future cohorts. Second, the two samples have very different mean ages (over 50 years). Combining data meta-analytically between groups penalizes SNPs that are significant in one sample, but not the other. This analysis of two cohorts is a special case of a meta-analysis, which tends to boost power to detect true positive associations by aggregating information from multiple cohorts. Clearly the power to detect an association with a given effect size depends on the available sample size, so in general the power is increased by increasing the sample size alone. The power of a meta-analysis may be slightly lower than that obtainable in a very large sample all scanned on the same scanner with the same protocol, but practical limitations constrain how many subjects can be scanned and genotyped at any one center, so multi-site efforts can be more efficient than studies at any single site. In that case, meta-analyses may offer high power so long as the chosen phenotypes are measured consistently and reliably across datasets. At the same time, meta-analysis reduces the chance of false positives as it penalizes results that are not consistently detectable across sites; in other words, it finds effects that are known to generalize to other cohorts, and less likely to be spurious associations attributable to the genetic diversity or particular ascertainment or sampling of any one cohort. In cases where the genetic expression has a compact temporal expression pattern, our analysis may lead to false negatives, as an effect could be detected in one sample but not the other. The genes identified in our analysis should be thought of as those associated with lentiform nucleus volume throughout life. Genetic variants that were *not* associated with lentiform nucleus volume could certainly still be involved in cellular or functional differences, so the findings must be interpreted recognizing the power and limitations of the study. Third, the proportion of the sample variance explained by the top SNP, rs1795240, is relatively small (ADNI: 2.68 %; QTIM: 0.84 %). However, a SNP that explains 1–3 % of the overall variability is comparable to the strongest SNP effects observed for other complex traits in even larger studies (Bis et al. 2012; Stein et al. 2012). The small effect sizes and complexity of phenotypic traits mean that individual common SNPs will each probably explain a small portion of the overall observed variability of a given trait. In addition, the proportion of variability explained by the top SNP in each sample is different. Further exploration is needed to determine age related effects of *FMO3* gene variants on lentiform nucleus volume.

The genetic variants identified in our analysis provide replicated, genome-wide significant evidence for the *FMO* gene cluster's involvement in lentiform nucleus volume. In addition, gene-based tests and pathway enrichment analysis provide evidence of probable mechanistic actions through which the variants in our analysis might affect lentiform nucleus volume. Future study is still needed to explain the functional mechanisms of change.

Acknowledgments Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health 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, AstraZeneca AB, Bayer Schering Pharma AG, Bristol-Myers Squibb, Eisai Global Clinical Development, Elan Corporation, Genentech, GE Healthcare, Glaxo-SmithKline, Innogenetics, Johnson and Johnson, Eli Lilly and Co., Medpace, Inc., Merck and Co., Inc., Novartis AG, Pfizer Inc., F Hoffman-La Roche, Schering-Plough, Synarc, Inc., as well as non-profit partners the Alzheimer's Association and Alzheimer's Drug Discovery Foundation, with participation from the US Food and Drug Administration. Private sector contributions to ADNI are facilitated by the Foundation for the National Institutes of Health (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, CA, USA. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of California, Los Angeles, CA, USA. This research was also supported by NIH grants P30 AG010129, K01 AG030514, and the Dana Foundation. Algorithm development for this study was also funded by the NIA, NIBIB, NICHD, the National Library of Medicine, and the National Center for Research Resources (AG016570, EB01651, LM05639, RR019771, EB008432, EB008281 and EB007813, to PMT). For QTIM we thank the twins for their participation, Kori Johnson and the radiographers for MRI scanning and preprocessing the images, Marlene Grace and Ann Eldridge for twin recruitment, Daniel Park for database support, Anjali Henders for DNA processing and preparation, Scott Gordon for quality control and management of the genotypes. The QTIM study was supported by the National Institute of Child Health and Human Development (R01 HD050735), and the National Health and Medical Research Council (NHMRC 486682, 1009064), Australia. Genotyping was supported by NHMRC (389875). DPH is partially supported by NSF grant DGE-0707424. OK was supported by NIH NIA Grant F30AG041681 and the UCLA Graduate Division. JLS is partially supported by a T32 post-doctoral training grant in Neurobehavioral Genetics. NJ was additionally supported by NIH NLM Grant T15 LM07356. GM was supported by an NHMRC Fellowship 613667; GZ was supported by an ARC Future Fellowship. AJS receives support from ADNI as well as NIA R01 AG19771, P30 AG10133 and U01 AG032984.

Conflict of interest The authors declare that they have no conflict of interest.

References

Abecasis, G. R., Li, Y., Willer, C. J., Ding, J., & Scheet, P. (2010). MaCH: using sequence and genotype data to estimate haplotypes and unobserved genotypes. *Genetic Epidemiology, 34*(8), 816–834. doi:10.1002/gepi.20533.

Annett, M. (1970). A classification of hand preference by association analysis. *British Journal of Psychology, 61*(3), 303–321.

Arnone, D., Cavanagh, J., Gerber, D., Lawrie, S. M., Ebmeier, K. P., & McIntosh, A. M. (2009). Magnetic resonance imaging studies in bipolar disorder and schizophrenia: meta-analysis. [Meta-analysis Research Support, Non-U.S. Gov't Review]. *The British Journal of Psychiatry: The Journal of Mental Science, 195*(3), 194–201. doi:10.1192/bjp.bp.108.059717.

Athanasiu, L., Mattingsdal, M., Kahler, A. K., Brown, A., Gustafsson, O., Agartz, I., et al. (2010). Gene variants associated with schizophrenia in a Norwegian genome-wide study are replicated in a large European cohort. *Journal of Psychiatric Research, 44*(12), 748–753. doi:10.1016/j.jpsychires.2010.02.002.

Bhagwat, S. V., Bhamre, S., Boyd, M. R., & Ravindranath, V. (1996). Cerebral metabolism of imipramine and a purified flavin-containing monooxygenase from human brain. [In Vitro Research Support, Non-U.S. Gov't Research Support, U.S. Gov't, Non-P.H.S.]. *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology, 15*(2), 133–142. doi:10.1016/0893-133X(95)00175-D.

Bhamre, S., Bhagwat, S. V., Shankar, S. K., Boyd, M. R., & Ravindranath, V. (1995). Flavin-containing monooxygenase mediated metabolism of psychoactive drugs by human brain microsomes. [Research Support, Non-U.S. Gov't]. *Brain Research, 672*(1-2), 276–280.

Bis, J. C., DeCarli, C. S., Smith, A. V., van der Lijn, F., Crivello, F., Fornage, M., et al. (2012). Common variants at 12q14 and 12q24 are associated with hippocampal volume. *Nature Genetics*.

Bobilya, D. J., Gauthier, N. A., Karki, S., Olley, B. J., & Thomas, W. K. (2008). Longitudinal changes in zinc transport kinetics, metallothionein and zinc transporter expression in a blood-brain barrier model in response to a moderately excessive zinc environment. [Research Support, N.I.H., Extramural]. *The Journal of Nutritional Biochemistry, 19*(2), 129–137. doi:10.1016/j.jnutbio.2007.06.014.

Buckner, R. L., Head, D., Parker, J., Fotenos, A. F., Marcus, D., Morris, J. C., & Snyder, A. Z. (2004). A unified approach for morphometric and functional data analysis in young, old, and demented adults using automated atlas-based head size normalization: reliability and validation against manual measurement of total intracranial volume. *NeuroImage, 23*(2), 724–738. doi:10.1016/j.neuroimage.2004.06.018.

Cascino, G. D., & Adams, R. D. (1986). Brainstem auditory hallucinosis. *Neurology, 36*(8), 1042–1047.

Cashman, J. R., & Zhang, J. (2002). Interindividual differences of human flavin-containing monooxygenase 3: genetic polymorphisms and functional variation. *Drug Metabolism and Disposition, 30*(10), 1043–1052.

Castellanos, F. X., Giedd, J. N., Marsh, W. L., Hamburger, S. D., Vaituzis, A. C., Dickstein, D. P., et al. (1996). Quantitative brain magnetic resonance imaging in attention-deficit hyperactivity disorder. *Archives of General Psychiatry, 53*(7), 607–616.

Chen, W. M., & Abecasis, G. R. (2007). Family-based association tests for genomewide association scans. *American Journal of Human Genetics, 81*(5), 913–926. doi:10.1086/521580.

Chen, C. S., Chiang, I. C., Li, C. W., Lin, W. C., Lu, C. Y., Hsieh, T. J., & Kuo, Y. T. (2009). Proton magnetic resonance spectroscopy of late-life major depressive disorder. [Comparative Study Research Support, Non-U.S. Gov't]. *Psychiatry Research, 172*(3), 210–214. doi:10.1016/j.pscychresns.2009.01.003.

Chen, C. H., Gutierrez, E. D., Thompson, W., Panizzon, M. S., Jernigan, T. L., Eyler, L. T., & Dale, A. M. (2012). Hierarchical genetic organization of human cortical surface area. [Research Support, N.I.H., Extramural Research Support, Non-U.S. Gov't Research Support, U.S. Gov't, Non-P.H.S. Twin Study]. *Science, 335*(6076), 1634–1636. doi:10.1126/science.1215330.

Chiang, M. C., Barysheva, M., McMahon, K. L., de Zubicaray, G. I., Johnson, K., Montgomery, G. W., et al. (2012). Gene network effects on brain microstructure and intellectual performance identified in 472 twins. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 32*(25), 8732–8745. doi:10.1523/JNEUROSCI.5993-11.2012.

Dexter, D. T., Carayon, A., Javoy-Agid, F., Agid, Y., Wells, F. R., Daniel, S. E., et al. (1991). Alterations in the levels of iron, ferritin and other trace metals in Parkinson's disease and other neurodegenerative diseases affecting the basal ganglia. *Brain, 114*(Pt 4), 1953–1975.

Diamond, M. C., Scheibel, A. B., & Elson, L. M. (1985). *The human brain coloring book* (1st ed.). New York: Barnes & Noble Books.

Durston, S., Fossella, J. A., Casey, B. J., Hulshoff Pol, H. E., Galvan, A., Schnack, H. G., et al. (2005). Differential effects of DRD4 and DAT1 genotype on fronto-striatal gray matter volumes in a sample of subjects with attention deficit hyperactivity disorder, their unaffected siblings, and controls. *Molecular Psychiatry, 10*(7), 678–685. doi:10.1038/sj.mp.4001649.

Duvernoy, H. M., & Bourgouin, P. (1999). *The human brain: surface, three-dimensional sectional anatomy with MRI, and blood supply (2nd completely rev. and enl. ed.)*. Wien: Springer.

Elkashef, A. M., Buchanan, R. W., Gellad, F., Munson, R. C., & Breier, A. (1994). Basal ganglia pathology in schizophrenia and tardive dyskinesia: an MRI quantitative study. *The American Journal of Psychiatry, 151*(5), 752–755.

Ellison-Wright, I., Ellison-Wright, Z., & Bullmore, E. (2008). Structural brain change in attention deficit hyperactivity disorder identified by meta-analysis. *BMC Psychiatry, 8*, 51. doi:10.1186/1471-244X-8-51.

Ellison-Wright, I., Glahn, D. C., Laird, A. R., Thelen, S. M., & Bullmore, E. (2008). The anatomy of first-episode and chronic schizophrenia: an anatomical likelihood estimation meta-analysis. *The American Journal of Psychiatry, 165*(8), 1015–1023. doi:10.1176/appi.ajp.2008.07101562.

Han, B., & Eskin, E. (2011). Random-effects model aimed at discovering associations in meta-analysis of genome-wide association studies. *American Journal of Human Genetics, 88*(5), 586–598. doi:10.1016/j.ajhg.2011.04.014.

Hines, R. N. (2006). Developmental and tissue-specific expression of human flavin-containing monooxygenases 1 and 3. *Expert Opinion on Drug Metabolism & Toxicology, 2*(1), 41–49. doi:10.1517/17425255.2.1.41.

Hines, R. N., Hopp, K. A., Franco, J., Saeian, K., & Begun, F. P. (2002). Alternative processing of the human FMO6 gene renders transcripts incapable of encoding a functional flavin-containing monooxygenase. *Molecular Pharmacology, 62*(2), 320–325.

Hines, R. N., Koukouritaki, S. B., Poch, M. T., & Stephens, M. C. (2008). Regulatory polymorphisms and their contribution to interindividual differences in the expression of enzymes influencing drug and toxicant disposition. *Drug Metabolism Reviews, 40*(2), 263–301. doi:10.1080/03602530801952682.

Kempton, M. J., Geddes, J. R., Ettinger, U., Williams, S. C., & Grasby, P. M. (2008). Meta-analysis, database, and meta-regression of 98 structural imaging studies in bipolar disorder. *Archives of General Psychiatry, 65*(9), 1017–1032. doi:10.1001/archpsyc.65.9.1017.

Koukouritaki, S. B., Simpson, P., Yeung, C. K., Rettie, A. E., & Hines, R. N. (2002). Human hepatic flavin-containing monooxygenases 1 (FMO1) and 3 (FMO3) developmental expression. *Pediatric Research, 51*(2), 236–243.

Koukouritaki, S. B., Poch, M. T., Henderson, M. C., Siddens, L. K., Krueger, S. K., VanDyke, J. E., et al. (2007). Identification and functional analysis of common human flavin-containing monooxygenase 3 genetic variants. *Journal of Pharmacology and Experimental Therapeutics, 320*(1), 266–273. doi:10.1124/jpet.106.112268.

Kremen, W. S., Prom-Wormley, E., Panizzon, M. S., Eyler, L. T., Fischl, B., Neale, M. C., et al. (2010). Genetic and environmental influences on the size of specific brain regions in midlife: the VETSA MRI study. *NeuroImage, 49*(2), 1213–1223. doi:10.1016/j.neuroimage.2009.09.043.

Krueger, S. K., & Williams, D. E. (2005). Mammalian flavin-containing monooxygenases: structure/function, genetic polymorphisms and role in drug metabolism. *Pharmacology and Therapeutics, 106*(3), 357–387. doi:10.1016/j.pharmthera.2005.01.001.

Lander, E. S., & Schork, N. J. (1994). Genetic dissection of complex traits. *Science, 265*(5181), 2037–2048.

Li, M. X., Sham, P. C., Cherny, S. S., & Song, Y. Q. (2010). A knowledge-based weighting framework to boost the power of genome-wide association studies. [Research Support, Non-U.S. Gov't]. *PLoS One, 5*(12), e14480. doi:10.1371/journal.pone.0014480.

Li, M. X., Gui, H. S., Kwan, J. S. H., & Sham, P. C. (2011). GATES: a rapid and powerful gene-based association test using extended simes procedure. *American Journal of Human Genetics, 88*(3), 283–293. doi:10.1016/j.ajhg.2011.01.019.

Marsden, C. D., Obeso, J. A., & Rothwell, J. C. (1983). Clinical neurophysiology of muscle jerks: myoclonus, chorea, and tics. *Advances in Neurology, 39*, 865–881.

Mayberg, H. S., Brannan, S. K., Tekell, J. L., Silva, J. A., Mahurin, R. K., McGinnis, S., & Jerabek, P. A. (2000). Regional metabolic effects of fluoxetine in major depression: serial changes and relationship to clinical response. [Research Support, Non-U.S. Gov't Research Support, U.S. Gov't, P.H.S.]. *Biological Psychiatry, 48*(8), 830–843.

McCarthy, M. I., Abecasis, G. R., Cardon, L. R., Goldstein, D. B., Little, J., Ioannidis, J. P., & Hirschhorn, J. N. (2008). Genome-wide association studies for complex traits: consensus, uncertainty and challenges. *Nature Reviews Genetics, 9*(5), 356–369. doi:10.1038/nrg2344.

Morey, R. A., Selgrade, E. S., Wagner, H. R., 2nd, Huettel, S. A., Wang, L., & McCarthy, G. (2010). Scan-rescan reliability of subcortical brain volumes derived from automated segmentation. *Human Brain Mapping, 31*(11), 1751–1762. doi:10.1002/hbm.20973.

Neale, M. C., Cardon, L. R., & North Atlantic Treaty Organization. Scientific Affairs Division. (1992). *Methodology for genetic studies of twins and families*. Dordrecht: Kluwer Academic Publishers.

Obeso, J. A., Rodriguez-Oroz, M. C., Rodriguez, M., Lanciego, J. L., Artieda, J., Gonzalo, N., & Olanow, C. W. (2000). Pathophysiology of the basal ganglia in Parkinson's disease. *Trends in Neurosciences, 23*(10 Suppl), S8–19.

Overby, L. H., Carver, G. C., & Philpot, R. M. (1997). Quantitation and kinetic properties of hepatic microsomal and recombinant flavin-containing monooxygenases 3 and 5 from humans. *Chemico-Biological Interactions, 106*(1), 29–45.

Patenaude, B., Smith, S. M., Kennedy, D. N., & Jenkinson, M. (2011). A Bayesian model of shape and appearance for subcortical brain segmentation. [Research Support, N.I.H., Extramural Research Support, Non-U.S. Gov't]. *NeuroImage, 56*(3), 907–922. doi:10.1016/j.neuroimage.2011.02.046.

Peper, J. S., Brouwer, R. M., Boomsma, D. I., Kahn, R. S., & Hulshoff Pol, H. E. (2007). Genetic influences on human brain structure: a review of brain imaging studies in twins. *Human Brain Mapping, 28*(6), 464–473. doi:10.1002/hbm.20398.

Petersen, R. C. (2000). Aging, mild cognitive impairment, and Alzheimer's disease. *Neurologic Clinics, 18*(4), 789–806.

Poliseno, L., Salmena, L., Zhang, J., Carver, B., Haveman, W. J., & Pandolfi, P. P. (2010). A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. *Nature, 465*(7301), 1033–1038. doi:10.1038/nature09144.

Raz, N., Rodrigue, K. M., Kennedy, K. M., Head, D., Gunning-Dixon, F., & Acker, J. D. (2003). Differential aging of the human striatum: longitudinal evidence. *AJNR. American Journal of Neuroradiology, 24*(9), 1849–1856.

Reiner, A., Albin, R. L., Anderson, K. D., D'Amato, C. J., Penney, J. B., & Young, A. B. (1988). Differential loss of striatal projection neurons in Huntington disease. *Proceedings of the National Academy of Sciences of the United States of America, 85*(15), 5733–5737.

Schultz, W. (2002). Getting formal with dopamine and reward. *Neuron, 36*(2), 241–263.

Smith, S. M. (2002). Fast robust automated brain extraction. *Human Brain Mapping, 17*(3), 143–155. doi:10.1002/hbm.10062.

Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., et al. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage, 23*, S208–S219. doi:10.1016/j.neuroimage.2004.07.051.

Snell, R. S. (2010). *Clinical neuroanatomy* (7th ed.). Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins.

Stein, J. L., Hua, X., Morra, J. H., Lee, S., Hibar, D. P., Ho, A. J., et al. (2010). Genome-wide analysis reveals novel genes influencing temporal lobe structure with relevance to neurodegeneration in Alzheimer's disease. *NeuroImage, 51*(2), 542–554. doi:10.1016/j.neuroimage.2010.02.068.

Stein, J. L., Medland, S. E., Arias Vasquez, A., Hibar, D., Senstad, R. E., Winkler, A. M. et al. (2012). Identification of common variants associated with human hippocampal and intracranial volumes. *Nature Genetics*.

Strakowski, S. M., DelBello, M. P., Sax, K. W., Zimmerman, M. E., Shear, P. K., Hawkins, J. M., & Larson, E. R. (1999). Brain magnetic resonance imaging of structural abnormalities in bipolar disorder. *Archives of General Psychiatry, 56*(3), 254–260.

Toga, A. W., & Thompson, P. M. (2003). Mapping brain asymmetry. *Nature Reviews Neuroscience, 4*(1), 37–48. doi:10.1038/nrn1009.

Wellcome Trust Case Control Consortium. (2007). Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature, 447*(7145), 661–678. doi:10.1038/Nature05911.

Wen, Y. Z., Zheng, L. L., Liao, J. Y., Wang, M. H., Wei, Y., Guo, X. M., et al. (2011). Pseudogene-derived small interference RNAs regulate gene expression in African Trypanosoma brucei. *Proceedings of the National Academy of Sciences of the United States of America, 108*(20), 8345–8350. doi:10.1073/pnas.1103894108.

Wik, G., Wiesel, F. A., Sjogren, I., Blomqvist, G., Greitz, T., & Stone-Elander, S. (1989). Effects of sulpiride and chlorpromazine on regional cerebral glucose metabolism in schizophrenic patients as determined by positron emission tomography. [Research Support, Non-U.S. Gov't]. *Psychopharmacology, 97*(3), 309–318.

Williams, J. A., Hyland, R., Jones, B. C., Smith, D. A., Hurst, S., Goosen, T. C., et al. (2004). Drug-drug interactions for UDP-glucuronosyltransferase substrates: a pharmacokinetic explanation for typically observed low exposure (AUCi/AUC) ratios. *Drug Metabolism and Disposition, 32*(11), 1201–1208. doi:10.1124/dmd.104.000794.

Wilson, S. A. K. (1912). Progressive lenticular degeneration. A familial nervous disease associated with cirrhosis of the liver. *Lancet, 1*, 1115–1119.

Yeung, C. K., Lang, D. H., Thummel, K. E., & Rettie, A. E. (2000). Immunoquantitation of FMO1 in human liver, kidney, and intestine. *Drug Metabolism and Disposition, 28*(9), 1107–1111.

5.2 Genetic analysis of hippocampal volume in 21,151 subjects

This section is adapted from:

Jason L. Stein, Medland, S.E., Vasquez, A.A., **Derrek P. Hibar**, Senstad, R.E., Winkler, A.M., Toro, R., Appel, K., Bartecek, R., Bergmann, Ø., Bernard, M., Brown, A.A., Cannon, D.M., Chakravarty, M.M., Christoforou, A., Domin, M., Grimm, O., Hollinshead, M., Holmes, A.J., Homuth, G., Hottenga, J.-J., Langan, C., Lopez, L.M., Hansell, N.K., Hwang, K.S., Kim, S., Laje, G., Lee, P.H., Liu, X., Loth, E., Lourdusamy, A., Mattingsdal, M., Mohnke, S., Maniega, S.M., Nho, K., Nugent, A.C., O'Brien, C., Papmeyer, M., Pütz, B., Ramasamy, A., Rasmussen, J., Rijpkema, M., Risacher, S.L., Roddey, J.C., Rose, E.J., Ryten, M., Shen, L., Sprooten, E., Strengman, E., Teumer, A., Trabzuni, D., Turner, J., van Eijk, K., van Erp, T.G.M., van Tol, M.-J., Wittfeld, K., Wolf, C., Woudstra, S., Aleman, A., Alhusaini, S., Almasy, L., Binder, E.B., Brohawn, D.G., Cantor, R.M., Carless, M.A., Corvin, A., Czisch, M., Curran, J.E., Davies, G., de Almeida, M.A.A., Delanty, N., Depondt, C., Duggirala, R., Dyer, T.D., Erk, S., Fagerness, J., Fox, P.T., Freimer, N.B., Gill, M., Göring, H.H.H., Hagler, D.J., Hoehn, D., Holsboer, F., Hoogman, M., Hosten, N., Jahanshad, N., Johnson, M.P., Kasperaviciute, D., Kent, J.W., Kochunov, P., Lancaster, J.L., Lawrie, S.M., Liewald, D.C., Mandl, R., Matarin, M., Mattheisen, M., Meisenzahl, E., Melle, I., Moses, E.K., Mühleisen, T.W., Nauck, M., Nöthen, M.M., Olvera, R.L., Pandolfo, M., Pike, G.B., Puls, R., Reinvang, I., Rentería, M.E., Rietschel, M., Roffman, J.L., Royle, N.A., Rujescu, D., Savitz, J., Schnack, H.G., Schnell, K., Seiferth, N., Smith, C., Steen, V.M.,

Valdés Hernández, M.C., Van den Heuvel, M., van der Wee, N.J., Van Haren, N.E.M., Veltman, J.A., Völzke, H., Walker, R., Westlye, L.T., Whelan, C.D., Agartz, I., Boomsma, D.I., Cavalleri, G.L., Dale, A.M., Djurovic, S., Drevets, W.C., Hagoort, P., Hall, J., Heinz, A., Jack, C.R., Foroud, T.M., Le Hellard, S., Macciardi, F., Montgomery, G.W., Poline, J.B., Porteous, D.J., Sisodiya, S.M., Starr, J.M., Sussmann, J., Toga, A.W., Veltman, D.J., Walter, H., Weiner, M.W., Bis, J.C., Ikram, M.A., Smith, A.V., Gudnason, V., Tzourio, C., Vernooij, M.W., Launer, L.J., DeCarli, C., Seshadri, S., Andreassen, O.A., Apostolova, L.G., Bastin, M.E., Blangero, J., Brunner, H.G., Buckner, R.L., Cichon, S., Coppola, G., de Zubicaray, G.I., Deary, I.J., Donohoe, G., de Geus, E.J.C., Espeseth, T., Fernández, G., Glahn, D.C., Grabe, H.J., Hardy, J., Hulshoff Pol, H.E., Jenkinson, M., Kahn, R.S., McDonald, C., McIntosh, A.M., McMahon, F.J., McMahon, K.L., Meyer-Lindenberg, A., Morris, D.W., Müller-Myhsok, B., Nichols, T.E., Ophoff, R.A., Paus, T., Pausova, Z., Penninx, B.W., Potkin, S.G., Sämann, P.G., Saykin, A.J., Schumann, G., Smoller, J.W., Wardlaw, J.M., Weale, M.E., Martin, N.G., Franke, B., Wright, M.J., Thompson, P.M. *Identification of common variants associated with human hippocampal and intracranial volumes.* Nature Genetics. 2012. 44(5):552-61.

Identification of common variants associated with human hippocampal and intracranial volumes

Identifying genetic variants influencing human brain structures may reveal new biological mechanisms underlying cognition and neuropsychiatric illness. The volume of the hippocampus is a biomarker of incipient Alzheimer's disease[1,2] and is reduced in schizophrenia[3], major depression[4] and medial temporal lobe epilepsy[5]. Whereas many brain imaging phenotypes are highly heritable[6,7], identifying and replicating genetic influences has been difficult, as small effects and the high costs of magnetic resonance imaging (MRI) have led to underpowered studies. Here we report genome-wide association meta-analyses and replication for mean bilateral hippocampal, total brain and intracranial volumes from a large multinational consortium. The intergenic variant rs7294919 was associated with hippocampal volume (12q24.22; $N = 21{,}151$; $P = 6.70 \times 10^{-16}$) and the expression levels of the positional candidate gene *TESC* in brain tissue. Additionally, rs10784502, located within *HMGA2*, was associated with intracranial volume (12q14.3; $N = 15{,}782$; $P = 1.12 \times 10^{-12}$). We also identified a suggestive association with total brain volume at rs10494373 within *DDR2* (1q23.3; $N = 6{,}500$; $P = 5.81 \times 10^{-7}$).

The hippocampal formation is a key brain structure for learning, memory[8,9] and stress regulation[10] and is implicated in many neuropsychiatric disorders. Further, overall brain and head sizes are altered in many disorders and are significantly correlated with general cognitive ability[11–13]. Hippocampal, total brain and intracranial volumes are highly heritable in non-human primates[14,15] and in humans[6,7]. Finding loci that influence these measures may lead to the identification of genes underlying susceptibility for neuropsychiatric diseases. Here we sought to identify common genetic polymorphisms influencing hippocampal, total brain and intracranial volumes in a large multinational consortium.

Our discovery sample comprised 17 cohorts of European ancestry from whom genome-wide SNPs and structural MRI data were collected (**Supplementary Tables 1–3**). Unselected population samples and case-control studies were included, with cases ascertained for neuropsychiatric disorders including depression, anxiety, Alzheimer's disease and schizophrenia. To distinguish whether putative effects at these loci varied with disease status, analyses were run in the full sample ($N = 7{,}795$) and in a healthy subsample ($N = 5{,}775$). To help disentangle overall brain size effects from those specific to hippocampal volume, associations were assessed with and without controlling for total brain and intracranial volumes (Online Methods). As the initial goal of the study was to explore associations with hippocampal volume, total brain and intracranial volumes were analyzed in healthy subjects only.

Phenotypes were computed from three-dimensional anatomical T_1-weighted magnetic resonance images, using validated automated segmentation programs[16–18] (**Supplementary Fig. 1** and **Supplementary Tables 4** and **5**). Extensive quality control analysis of segmentation was performed on sample outliers; subjects with poorly delineated brain volume phenotypes were removed (**Supplementary Figs. 2–6**). The mean bilateral hippocampal volume across the discovery cohorts was 3,917.4 mm^3 (s.d. = 441.0 mm^3).

Heritability of structural brain phenotypes was estimated in a sample of Australian monozygotic and dizygotic twins and their siblings (Queensland Twin Imaging (QTIM) study; $N = 646$, including ungenotyped participants; age range = 20–30 years) for hippocampal volume ($h^2 = 0.62$), total brain volume ($h^2 = 0.89$) and intracranial volume ($h^2 = 0.78$). Hippocampal volume was also highly heritable in an extended pedigree cohort of Mexican-Americans from the United States (Genetics of Brain Structure and Function (GOBS); $N = 605$; age range = 18–85; $h^2 = 0.74$), as were total brain volume ($h^2 = 0.77$) and intracranial volume ($h^2 = 0.84$). All heritability estimates were highly significant ($P < 0.001$).

To enable consortium-wide comparison of ancestry and to adjust appropriately for population stratification, each site conducted multidimensional scaling (MDS) analyses comparing their data to the HapMap 3 reference populations (**Supplementary Fig. 7**). All subsequent analyses included the following covariates: sex, linear and quadratic effects of age, interactions of sex with age covariates, MDS components and dummy covariates for different magnetic resonance acquisitions. Analyses were filtered for genotyping and imputation quality (**Supplementary Fig. 8** and **Supplementary Table 6**); distributions of test statistics were examined at the cohort level through Manhattan and quantile-quantile plots (**Supplementary Figs. 9–24**). We conducted fixed-effects meta-analysis with METAL, applying genomic control[19] (**Supplementary Figs. 25–32**). For completeness and to account for heterogeneity across sites, a random-effects meta-analysis was also performed[20] (**Supplementary Figs. 33–40**). We attempted *in silico* replication of the top five loci for each trait within the combined CHARGE Consortium discovery set and 3C replication sample[21] ($N = 10{,}779$), as well as in two cohorts of European ancestry (imputed to the Utah residents of Northern and Western European ancestry (CEU) and/or Toscani in Italy (TSI) HapMap cohorts; $N = 449$) and in two additional cohorts (imputed to combined CEU and Yoruba in Ibadan, Nigeria (YRI), and to

A full list of authors and affiliations appears at the end of the paper.

Received 6 September 2011; accepted 19 March 2012; published online 15 April 2012; doi:10.1038/ng.2250

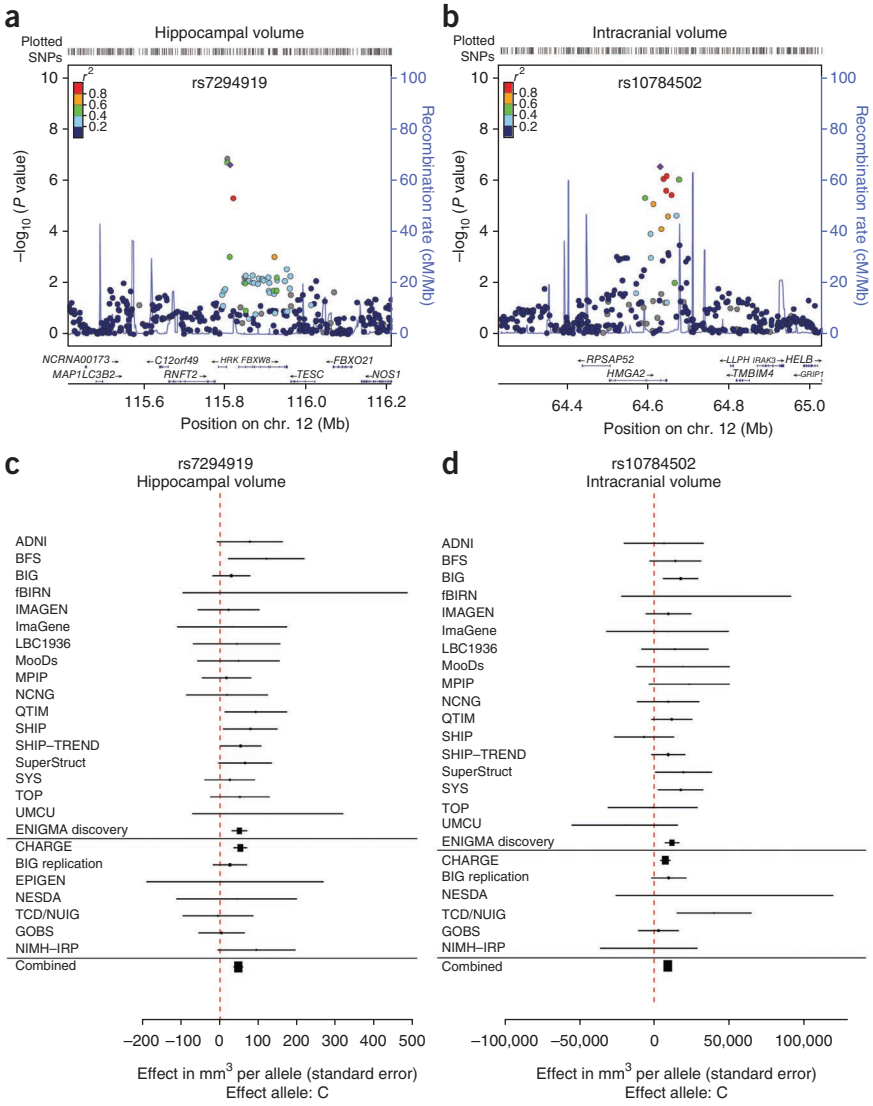


Figure 1 Association results and meta-analysis of effects in individual and combined analyses. (**a**) The strongest association with hippocampal volume was found for rs7294919. Fixed-effects meta-analysis *P* values are shown[41] after controlling for intracranial volume using all subjects in the discovery sample. (**b**) The strongest association with intracranial volume was found for rs10784502. Fixed-effects meta-analysis *P* values are shown in healthy subjects only. (**c**,**d**) The effect within each sample contributing to the meta-analysis is shown in forest plots for hippocampal volume (**c**) and intracranial volume (**d**). Association data using intracranial volume as a phenotype were not available for the EPIGEN sample. Head size was not controlled for in the CHARGE Consortium association analyses.

Mexican ancestry in Los Angeles, California (MEX); $N = 842$). We also undertook custom genotyping of the two most promising SNPs in two additional samples of European ancestry (BIG replication and Trinity College Dublin/National University of Ireland, Galway (TCD/NUIG); $N = 1{,}286$).

In general, previously identified polymorphisms associated with hippocampal volume showed little association in our meta-analysis (*BDNF*, *TOMM40*, *CLU*, *PICALM*, *ZNF804A*, *COMT*, *DISC1*, *NRG1*, *DTNBP1*; **Supplementary Table 7**), nor did SNPs previously associated with schizophrenia[22] and bipolar disorder[23] (**Supplementary Table 8**). The most significant SNPs in each analysis from the discovery sample ($P \leq 5 \times 10^{-5}$) are listed (**Supplementary Tables 9–16**). No markers reached genome-wide significance ($P < 1.25 \times 10^{-8}$; Online Methods) in the discovery sample alone. However, the strongest associations for hippocampal and intracranial volumes were replicated, yielding results at genome-wide significance (**Fig. 1** and **Table 1**; see **Supplementary Tables 17–25** for additional results and gene-based tests[24]).

In our discovery sample, two SNPs in the same linkage disequilibrium (LD) block showed strong associations with hippocampal volume after controlling for intracranial volume (rs7294919 and rs7315280; $r^2 = 0.81$, CEU 1000 Genomes Pilot 1). A random-effects analysis of the discovery sample, conducted to examine heterogeneity between cohorts, reduced significance only slightly for rs7294919 ($P = 4.43 \times 10^{-7}$) compared to the primary fixed-effects analysis ($P = 2.42 \times 10^{-7}$). The association was consistent, although stronger, in the full sample compared to the healthy subset (**Fig. 2**). Notably, the association was robust to the effects of head and brain size (**Fig. 2**), and the locus was not significantly associated with intracranial volume ($P = 0.54$) or total brain volume ($P = 0.41$). This suggests an effect at the level of the hippocampus rather than on brain size in general. The direction of the effect was consistent across samples and ages (**Fig. 1**). Haplotype analysis of directly genotyped variants near rs7294919 in two samples confirmed that the association was present across the haplotype and that the causal variant was well marked by rs7294919 (**Supplementary Note**). rs7294919 was also significantly associated with hippocampal volume in the cohorts from the CHARGE Consortium, which are composed of elderly subjects. Meta-analysis of the Enhancing Neuro Imaging Genetics through Meta-Analysis (ENIGMA) discovery and replication samples with those from the CHARGE Consortium yielded a highly significant association for rs7294919 ($P = 6.70 \times 10^{-16}$; $N = 21{,}151$).

rs7294919 lies between *HRK* and *FBXW8* (12q24.22; **Fig. 1**) and is not in LD with any SNPs within coding sequences, UTRs or splice sites within 500 kb ($r^2 > 0.4$) in the CEU sample from the 1000 Genomes Project Phase 1. To determine whether the observed association is related to a regulatory mechanism, we examined potential *cis* effects of this variant on expression levels of genes within a 1-Mb region. In temporal lobe tissue resected from 71 individuals with mesial temporal lobe epilepsy and hippocampal sclerosis in the University College London (UCL) epilepsy cohort, we examined association between rs4767492 (a proxy for rs7294919, which was not directly genotyped; $r^2 = 0.636$ in 1000 Genomes Project Phase 1) and expression levels. This analysis suggested an association ($P = 0.006$, controlling for age) with expression of the *TESC* gene, which lies 3′ to *FBXW8* (149 kb; **Fig. 3**). To corroborate this finding, we used the publicly available SNPExpress database (see URLs), which includes data on gene expression in post-mortem frontal cortex from 93 subjects. In this independent sample, expression levels of *TESC*

Table 1 Results from the genome-wide association meta-analyses of mean hippocampal, intracranial and total brain volumes

Sample	*N*	Freq. of the effect allele	β (mm^3)	S.E. (mm^3)	*P* value	Heterogeneity *P* value	Variance explained (%)[h]
Mean bilateral hippocampal volume[a]							
rs7294919[b]							
Discovery Fixed-effects model	7,795	0.104	50.27	9.71	2.42×10^{-7}	0.913	0.242
Random-effects model			50.12	9.65	4.43×10^{-7}	0.910	0.241
ENIGMA CEU and TSI replication	1,735	0.101	22.05	19.00	0.246	0.924	0.042
ENIGMA CEU and YRI or MEX replication	842	0.125	27.77	25.96	0.285	0.127	0.095
Discovery and replication	10,372	0.106	42.74	8.22	1.99×10^{-7}	0.347	0.177
CHARGE *in silico* replication	10,779	0.093	52.70	8.45	3.40×10^{-10}	0.442	0.458
ENIGMA and CHARGE	21,151	0.099	47.58	5.89	6.70×10^{-16}	0.419	0.265
Intracranial volume[c]							
rs10784502[d]							
Discovery Fixed-effects model	5,778	0.488	11860.73	2319.00	3.14×10^{-7}	0.783	0.281
Random-effects model			11841.80	2270.07	3.93×10^{-7}	0.771	0.280
ENIGMA CEU and TSI replication[e]	1,130	0.525	15758.59	5244.69	0.003	0.065	0.468
ENIGMA CEU and YRI or MEX replication	699	0.348	1928.43	6215.31	0.756	0.710	0.008
Discovery and replication	7,607	0.479	11395.74	2007.27	1.37×10^{-8}	0.217	0.261
CHARGE *in silico* replication	8,175	0.501	7429.56	1630.92	5.23×10^{-6}	NA	0.110
ENIGMA and CHARGE	15,782	0.491	9006.71	1265.78	1.12×10^{-12}	0.145	0.166
Total brain volume[f]							
rs10494373[g]							
Discovery Fixed-effects model	5,778	0.082	13693.29	3187.51	1.74×10^{-5}	0.688	0.198
Random-effects model			13562.00	3114.17	2.69×10^{-5}	0.728	0.194
ENIGMA CEU and TSI replication	117	0.107	8435.89	20256.09	0.678	NA	0.001
ENIGMA MEX replication	605	0.097	26883.36	8608.20	0.001	NA	0.964
Discovery and replication	6,500	0.085	14778.23	2957.14	5.81×10^{-7}	0.182	0.240

Freq., frequency. CEU, TSI YRI and MEX refer to the HapMap 3 reference panels most representative of the sample and used for imputation; NA, not applicable.
[a]Mean bilateral hippocampal volume association results were corrected for intracranial volume, sex, age, age^2, sex × age and sex × age^2, and individuals with disease were included in the analysis. [b]rs7294919 is located at 12q24.22: position 115,811,975. Effect allele, C; non-effect allele, T. Genomic positions are based on the NCBI36/hg18 (March 2006) genome assembly. [c]Association results for intracranial volume were corrected for sex, age, age^2, sex × age, sex × age^2, and individuals with disease were excluded from this analysis. [d]rs10784502 is at 12q14.3: position 64,630,077. Effect allele, C; non-effect allele, T. [e]Intracranial volume and total brain volume were available for two participants in MPIP and one participant in the BIG cohort who did not have hippocampal volume measures. The proxy SNP rs8756 was genotyped in the TDC/NUIG cohort. [f]Analysis for total brain volume was corrected for sex, age, age^2, sex × age, sex × age^2, and individuals with disease were excluded. Total brain volume was not available for the ENIGMA replication cohorts. Within the CHARGE Consortium, a normalized version of total brain volume was analyzed and defined as total brain volume intracranial volume, and, because of this, the results are not comparable between consortia. [g]rs10494373 is at 1q23.3: position 160,885,986. Effect allele, C; non-effect allele, A. [h]Calculated as $2pq \times \beta^2 / (s.d.)^2$, where p and q are the minor and major allele frequencies, β is the unstandardized regression coefficient and s.d. is from the phenotype in the absence of covariate corrections. Intracranial volume phenotypic variance from the ENIGMA discovery sample was used to calculate percent variance explained in the CHARGE *in silico* replications, as this information was not available from the CHARGE consortium.

again significantly differed by genotype (rs4767492; $P = 0.0021$). Additional replication came from the UK Brain Expression Database, where *TESC* expression in post-mortem brain tissues from 134 individuals free from neurological disorders showed a strong difference by genotype in temporal cortex (rs7294919; $P = 9.7 \times 10^{-4}$ for gene and 4.8×10^{-5} for exon 8). Given the small sample sizes and low minor allele frequency of this SNP (MAF = 0.099), no homozygotes for the minor allele were observed in any brain tissue sample, limiting the inferences we can draw regarding mode of action. Expression of *HRK* showed little evidence of association with the proxy genotype in the UCL epilepsy cohort ($P = 0.11$) or SNPExpress ($P = 0.16$) but was associated with rs7294919 in temporal cortex within the UK Brain Expression Database ($P = 0.0051$). Additional associations were observed in peripheral blood mononuclear cells (PBMCs; **Supplementary Note**).

The expression results in brain tissue suggest that *TESC* is a primary positional candidate for our quantitative trait locus (QTL). Studies of mouse and chicken embryos show that *TESC* is expressed throughout the brain during development, with the strongest expression in the developing telencephalon and mesencephalon and near the developing ventricles[25]. *TESC* also has moderate expression in the human hippocampus during adulthood (Allen Institute Brain Atlas, see URLs; **Fig. 3**). Its protein product, tescalcin, interacts with the Na^+/H^+ exchanger (NHE1)[26], which is involved in the regulation of intracellular pH[21], cell volume and cytoskeletal organization[27]. *TESC* expression is strongly regulated during cell differentiation in a cell lineage–specific fashion[28,29]. Our data suggest that this role in cell proliferation and differentiation is relevant for hippocampal volume and brain development.

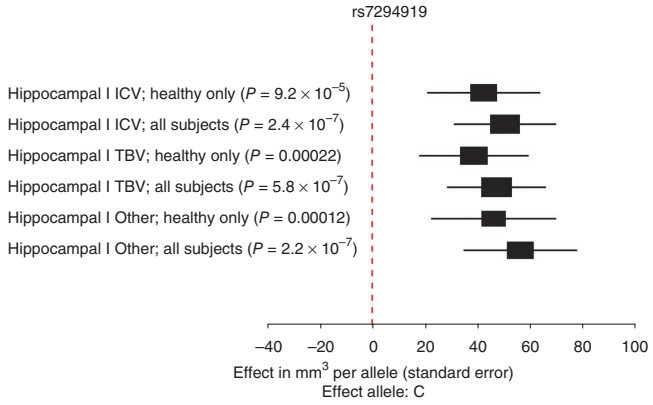


Figure 2 Association of rs7294919 with hippocampal volume stratified by disease and covariates. Effects are consistent in the discovery sample regardless of whether individuals with disease ($N = 7,795$) or only healthy subjects ($N = 5,775$) were included. The effect is also consistent whether accounting for intracranial volume (ICV), total brain volume (TBV) or without a measure of head size (Other).

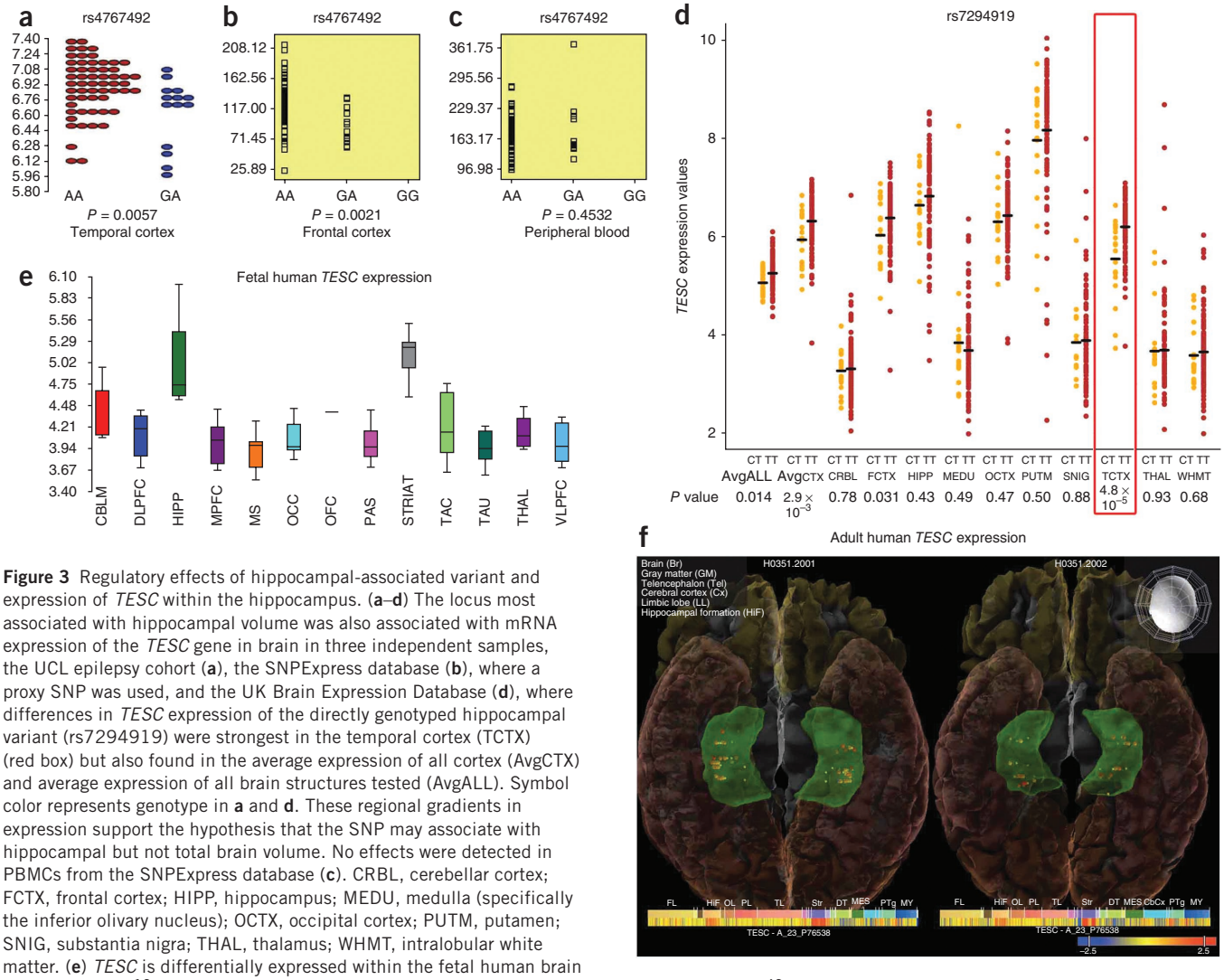


Figure 3 Regulatory effects of hippocampal-associated variant and expression of *TESC* within the hippocampus. (**a**–**d**) The locus most associated with hippocampal volume was also associated with mRNA expression of the *TESC* gene in brain in three independent samples, the UCL epilepsy cohort (**a**), the SNPExpress database (**b**), where a proxy SNP was used, and the UK Brain Expression Database (**d**), where differences in *TESC* expression of the directly genotyped hippocampal variant (rs7294919) were strongest in the temporal cortex (TCTX) (red box) but also found in the average expression of all cortex (AvgCTX) and average expression of all brain structures tested (AvgALL). Symbol color represents genotype in **a** and **d**. These regional gradients in expression support the hypothesis that the SNP may associate with hippocampal but not total brain volume. No effects were detected in PBMCs from the SNPExpress database (**c**). CRBL, cerebellar cortex; FCTX, frontal cortex; HIPP, hippocampus; MEDU, medulla (specifically the inferior olivary nucleus); OCTX, occipital cortex; PUTM, putamen; SNIG, substantia nigra; THAL, thalamus; WHMT, intralobular white matter. (**e**) *TESC* is differentially expressed within the fetal human brain ($P = 1.33 \times 10^{-12}$), with the highest expression in striatum (STRIAT) and hippocampus (HIPP)[42]. Box plots represent median and 25th to 75th percentiles. Upper and lower lines show minimum and maximum values, respectively. CBLM, cerebellum; DLPFC, dorsolateral prefrontal neocortex; MPFC, medial prefrontal neocortex; MS, motor-somatosensory neocortex; OCC, occipital visual neocortex; OFC, orbital prefrontal neocortex; PAS, parietal association neocortex; TAC, temporal association neocortex; TAU, temporal auditory neocortex; THAL, mediodorsal thalamus; VLPFC, ventrolateral prefrontal neocortex. (**f**) *TESC* has moderate to high gene expression throughout the adult human hippocampus (shown in green), as visualized in the Allen Institute Human Brain Atlas using Brain Explorer 2 software. An inferior view of the brain is shown in two subjects; the anterior portion of the brain is at the top. The colors of spheres within the hippocampus indicate the *Z*-scores of *TESC* expression normalized within each subject across brain structures. Heat maps show that expression of *TESC* is higher in the hippocampus (HiF) and striatum (Str) than in other brain structures.

The strongest association with intracranial volume was observed at rs10784502 (**Table 1**), an intronic SNP near the 3′ UTR of the *HMGA2* gene (12q14.3; **Fig. 1**). This locus was associated with intracranial volume across lifespan, as shown by the strong replication in samples from healthy elderly individuals in the CHARGE Consortium. The combined analysis resulted in the identification of a highly significant association ($P = 1.12 \times 10^{-12}$). Of note, rs10784502 has been reliably associated with increased adult height ($P = 3.636 \times 10^{-32}$; effect allele: C)[30]. The genetic correlation between height and intracranial volume within the QTIM sample was significant ($r_g = 0.31$; $P = 1.34 \times 10^{-7}$), as was that observed in the GOBS sample ($r_g = 0.20$; $P = 0.026$), suggesting modest overlap of shared genetic determinants. rs10784502 also had an effect on total brain volume in the discovery sample ($P = 9.49 \times 10^{-5}$). When considering the results from the intracranial volume meta-analysis in SNPs previously associated with height[31–33] ($N_{SNPs} = 175$; **Supplementary Fig. 41**), a clear inflation of the test statistic was observed ($\lambda = 1.44$), indicating that SNPs associated with height are also associated with intracranial volume. This enrichment, which was not observed for hippocampal volume (**Supplementary Figs. 42** and **43**), was due to a systematically higher degree of association throughout the candidate SNP set rather than a small number of large effects. Structural equation modeling showed that the effect of rs10784502 on intracranial volume could not completely be accounted for by the indirect effects of this SNP on height or by the correlation between height and intracranial volume (**Supplementary Fig. 44**).

Examining correlations between rs10784502 and expression levels of genes within a 1-Mb region, we identified a significant effect on the expression of *HMGA2* ($P = 0.0077$) as the single significant result in the GOBS transcriptional profile data. Additionally, *HMGA2* expression levels in PBMCs were significantly negatively genetically

correlated with intracranial volume (r_g = −0.49; P = 0.016) in this cohort. These results support *HMGA2* as a positional candidate gene underlying our observed QTL. *HMGA2* encodes the high-mobility group AT-hook 2 protein, which is a chromatin-associated protein that regulates stem cell renewal during development[34]. It is implicated in human growth through genetic association studies and the presence of rare mutations[35] and also has known roles in neural precursor cells[36]. Whether both functions are due to the same underlying mechanisms warrants further study.

To test for pleiotropic effects of rs7294919 and rs10784502, we examined the influence of these variants on cognition in the Brisbane Adolescent Twin Study[37] (N = 1642). The C allele of rs10784502, which was associated with increased intracranial volume, was also associated with increased full-scale IQ, as measured via the Multidimensional Aptitude Battery[38] (effect size (β) = 1.29, standard error (S.E.) = 0.47; P = 0.0073; phenotypic correlations are shown in **Supplementary Table 26**). This effect was driven by performance (PIQ; β = 1.74, S.E. = 0.61; P = 0.0044) rather than by verbal subtests (VIQ; P = 0.103). rs7294919 was not associated with full-scale IQ (P = 0.139) or PIQ (P = 0.489) but showed nominal association with VIQ (effect allele: C; β = 0.126, S.E. = 0.062; P = 0.043).

No associations at genome-wide significance were detected for total brain volume. Following inclusion of the replication samples, the strongest evidence for association was detected at rs10494373 within *DDR2* (1q23.3; $P = 5.81 \times 10^{-7}$) (**Table 1**), which encodes a receptor tyrosine kinase involved in cell growth and differentiation[39].

The current study identified and replicated two quantitative trait loci for hippocampal and intracranial volumes across lifespan in a large sample including both healthy subjects and those with neuropsychiatric diagnoses. The rs7294919 variant was associated with decreased hippocampal volume of 47.6 mm^3 or 1.2% of the average hippocampal volume per risk allele. Although further work is necessary to confirm the causal variant(s) and functional mechanisms, this QTL influencing hippocampal volume differences may act by regulating expression of *TESC* specifically within the brain. In addition, the C allele of rs10784502 is associated, on average, with 9,006.7 mm^3 larger intracranial volume, or 0.58% of intracranial volume per risk allele and is weakly associated with increased general intelligence by approximately 1.29 IQ points per allele.

It has previously been hypothesized that brain imaging endophenotypes would have large effect sizes; however, this has proven not to be the case for the specific volumetric traits measured here, which had comparable effect sizes to those observed in other genome-wide association studies of complex traits[40]. Notably, the discovery sample had 99.92% power to detect variants with effect sizes of 1% of the variance for MAF ≥ 0.05. It remains to be determined whether specific genetic variations linked to volumetric brain differences are also associated with other neuropsychiatric disorders, brain function and other cognitive traits. If this is the case, neuroimaging genetics may also discover new treatment targets related to the neurobiology of these disorders, in addition to improving phenomenologically based diagnostic criteria.

URLs. Allen Institute Brain Atlas, http://human.brain-map.org/; SNPExpress database, http://compute1.lsrc.duke.edu/softwares/SNPExpress/index.php; ADNI database, http://adni.loni.ucla.edu/; ADNI acknowledgements, http://adni.loni.ucla.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf; the Foundation for the NIH, http://www.fnih.org/; ADNI information, http://www.adni-info.org/; Brain Research Imaging Centre Edinburgh, http://www.bric.ed.ac.uk/; SINAPSE Collaboration, http://www.sinapse.ac.uk/; fBIRN, http://www.birncommunity.org/; SYS, http://www.saguenay-youth-study.org/; SHIP, http://ship.community-medicine.de/; ENIGMA Consortium protocols, http://enigma.loni.ucla.edu/protocols/; Mx, http://www.vcu.edu/mx/; SOLAR, http://solar.txbiomedgenetics.org/; Genetic Power Calculator, http://pngu.mgh.harvard.edu/~purcell/gpc/; HapMap, http://hapmap.ncbi.nlm.nih.gov/; Data upload site for participating studies, http://enigma.loni.ucla.edu/; METAL, http://www.sph.umich.edu/csg/abecasis/Metal/; METASOFT, http://genetics.cs.ucla.edu/meta/; maspD, http://gump.qimr.edu.au/general/daleN/matSpD/.

METHODS

Methods and any associated references are available in the online version of the paper at http://www.nature.com/naturegenetics/.

Note: Supplementary information is available on the Nature Genetics website.

ACKNOWLEDGMENTS

Some authors received commercial funding unrelated to the topic of this paper. N.J.v.d.W. received speaking fees from Eli Lilly & Company and Wyeth and served on advisory panels of Eli Lilly & Company, Pfizer, Wyeth and Servier. A.A. received an investigator-initiated unrestricted research grant from Bristol-Myers Squibb and speaker's bureau honoraria from AstraZeneca, Bristol-Myers Squibb and GlaxoSmithKline. H.J.G. received external research support from the German Research Foundation, the Federal Ministry of Education and Research Germany, speaker's honoraria from Bristol-Myers Squibb, Eli Lilly & Company, Novartis, Eisai, Boehringer Ingelheim and Servier and travel funds from Janssen-Cilag, Eli Lilly & Company, Novartis, AstraZeneca, Lundbeck and the SALUS–Institute for Trend-Research and Therapy Evaluation in Mental Health. M.N. received research grants from the Federal Ministry of Education and Research, Germany, the German Research Foundation, BioRad Laboratories, Siemens AG, Zeitschrift für Laboratoriumsmedizin, Bruker Daltronics, Abbott, Jurilab Kuopio, Roche Diagnostics, Instand and Becton Dickinson. H.V. received external research support via research grants from Hofmann La Roche, the Humboldt Foundation, the Federal Ministry of Education and Research (Germany) and the German Research Foundation. M.W. is on the following scientific advisory boards: Lilly, Araclon and Institut Catala de Neurociencies Aplicades, the Gulf War Veterans Illnesses Advisory Committee, VACO, Biogen Idec and Pfizer. M.W. received funding for consulting from Astra Zeneca, Araclon, Medivation/Pfizer, Ipsen, TauRx Therapeutics, Bayer Healthcare, Biogen Idec, Exonhit Therapeutics, SA, Servier, Synarc, Pfizer and Janssen; for travel from NeuroVigil, CHRU–Hopital Roger Salengro, Siemens, AstraZeneca, Geneva University Hospitals, Lilly, the University of California, San Diego–ADNI, Paris University, Institut Catala de Neurociencies Aplicades, the University of New Mexico School of Medicine, Ipsen, Clinical Trials on Alzheimer's Disease (CTAD), Pfizer, AD PD Meeting, Paul Sabatier University, Novartis and Tohoku University; and research support from: Merck, Avid, DoD, VA. M.W. received honoraria from PMDA/ the Japanese Ministry of Health, Labour, and Welfare, Tohoku University, Neuro Vigil, Insitut Catala de Neurociencies Aplicades. M.W. owns stock options for Synarc, Elan. Organizations contributing to the Foundation for the US NIH and thus to the National Institute on Aging (NIA)-funded Alzheimer's Disease Neuroimaging Initiative included Abbott, the Alzheimer's Association, the Alzheimer's Drug Discovery Foundation, Anonymous Foundation, AstraZeneca, Bayer Healthcare, BioClinica (ADNI 2), Bristol-Myers Squibb, the Cure Alzheimer's Fund, Eisai, Elan, Gene Network Sciences, Genentech, GE Healthcare, GlaxoSmithKline, Innogenetics, Johnson & Johnson, Eli Lilly & Company, Medpace, Merck, Novartis, Pfizer, Roche, Schering Plough, Synarc and Wyeth.

ADNI: The ADNI study was supported by the US NIH (U01 AG024904) and the Foundation for the NIH for genotype and phenotype data collection, the NIH (RC2 AG036535-01) for data analysis, the NIA (R01 AG019771-09) for additional data analysis and NCRAD (U24AG021886) for DNA used in part for the GWAS. Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (see URLs). As such, the investigators within ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. The ADNI sample wishes to acknowledge the investigators who contributed to the design and implementation of ADNI (see URLs). Data collection and sharing for this project were funded by ADNI (NIH grant U01 AG024904). ADNI is funded by the NIA, the National Institute of Biomedical Imaging and Bioengineering (NIBIB) and through generous contributions from Abbott, AstraZeneca AB, Bayer Schering Pharma AG, Bristol-Myers Squibb, Eisai Global Clinical Development, Elan Corporation, Genentech, GE Healthcare, GlaxoSmithKline,

Innogenetics, Johnson & Johnson, Eli Lilly & Company, Medpace, Merck and Cocpany, Novartis AG, Pfizer, F. Hoffman–La Roche, Schering-Plough and Synarc, as well as from nonprofit partners at the Alzheimer's Association and the Alzheimer's Drug Discovery Foundation, with participation from the US Food and Drug Administration (FDA). Private sector contributions to ADNI are facilitated by the Foundation for the NIH (see URLs). 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 of Neuro Imaging at the University of California, Los Angeles. This research was also supported by NIH grants (P30 AG010129 and K01 AG030514) and by the Dana Foundation. ADNI was launched in 2003 by the NIA, the NIBIB, the FDA, private pharmaceutical companies and nonprofit organizations as a 5-year public- private partnership. The primary goal of ADNI has been to test whether serial MRI), positron emission tomography (PET), other biological markers and clinical and neuropsychological assessments can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimer's disease. Determination of sensitive and specific markers of very early Alzheimer's disease progression is intended to aid researchers and clinicians in developing new treatments and monitoring their effectiveness, as well as lessening the time and cost of clinical trials. The Principal Investigator of this initiative is M.W. Weiner. ADNI is the result of efforts of many coinvestigators from a broad range of academic institutions and private corporations, and subjects have been recruited from over 50 sites across the United States and Canada. The initial goal of ADNI was to recruit 800 adults ages 55 to 90 to participate in the research—approximately 200 cognitively normal older individuals to be followed for 3 years, 400 people with MCI to be followed for 3 years and 200 people with early Alzheimer's disease to be followed for 2 years. For up-to-date information, please visit the ADNI website (see URLs).

BIG: The BIG study wishes to acknowledge S. Kooijman for coordination of sample collection and A. Heister, M. Naber, R. Makkinje, M. Hakobjan and M. Steehouwer for genotyping. The BIG study was supported by a Biobanking and Biomolecular Resources Research Infrastructure Netherlands (BBMRI-NL) complementation grant for brain segmentation and the Netherlands Organisation for Scientific Research (NWO) Horizon Breakthrough grant (grant number 93511010 (to A.A.V.).

Bipolar Family Study: The Bipolar Family Study wishes to thank the Scottish Mental Health Research Network for research assistant support, the Brain Research Imaging Centre Edinburgh (see URLs), a center in the Scottish Funding Council Scottish Imaging Network–A Platform for Scientific Excellence (SINAPSE) Collaboration (see URLs), for image acquisition and the Wellcome Trust Clinical Research Facility for genotyping. Genotyping was supported by the National Alliance for Research on Schizophrenia and Depression (NARSAD) Independent Investigator Award (to A.M.M.), and data collection was supported by the Health Foundation Clinician Scientist Fellowship.

fBIRN: fBIRN wishes to acknowledge D.B. Keator for leading fBIRN neuroinformatics development, B.A. Mueller for image calibration and quality assurance and A. Belger, V.D. Calhoun, G.G. Brown, J.M. Ford, G.H. Glover, R. Kikinis, K. Lim, J. Laurriello, J. Bustillo, G. McCarthy, D.S. O'Leary, B. Rosen, A.W.T. and J.T. Voyvodic for their leadership contributions to fBIRN scanner and sequence calibration, tool development and data collection efforts. The fBIRN study was supported by the US NIH (U24 RR21992) for phenotypic data collection. Genotyping was performed with the support of the grant RBIN04SWHR to F.M. from the Italian Ministry of University and Research.

GOBS: The GOBS study was supported by the US NIH (MH0708143 and MH083824 to D.C.G., MH078111 and MH59490 to J.B., C06 RR13556 and C06 RR017515). P.K. was also supported by an NIH grant (EB006395).

IMAGEN: IMAGEN is funded by the European Commission Framework Programme 6 (FP-6) Integrated Project IMAGEN (PL037286), the European Commission Framework Programme 7 (FP-7) Project Alzheimer's Disease, Alcoholism, Memory, Schizophrenia (ADAMS), the FP-7 Innovative Medicine Initiative Project European Autism Interventions (AIMS), the UK Department of Health National Institute of Health Research (NIHR)–Biomedical Research Centre Mental Health program and the MRC programme grant Developmental Pathways into Adolescent Substance Abuse (93558).

ImaGene: ImaGene wishes to acknowledge J. Lee and J. Lane for processing the blood samples, The Easton Consortium for Alzheimer's Disease Drug Discovery and Biomarker Development and the Alzheimer's Disease Research Center (ADRC) funded by the NIA at the University of California, Los Angeles (AG16570).

LBC1936: We thank the participants in LBC1936. We thank C. Murray, A.J. Gow, S.E. Harris, M. Luciano, P. Redmond, E. Sandeman, I. Gerrish, J. Boyd-Ellison, N. Leslie, A. Howden and C. Scott for data collection and preparation. This project is funded by the Age UK's Disconnected Mind programme and also by Research Into Ageing (251 and 285). The entire genome association part of the study was funded by the Biotechnology and Biological Sciences Research Council (BBSRC) (BB/F019394/1). Analysis of brain images was funded by UK MRC grants (G1001401 and 8200). The work was undertaken by The University of Edinburgh Centre for Cognitive Ageing and Cognitive Epidemiology, part of the cross council Lifelong Health and Wellbeing Initiative (G0700704/84698). Imaging was performed at the Brain Research Imaging Centre, Edinburgh, a center in the SINAPSE Collaboration. Funding from BBSRC, the Engineering and Physical Sciences Research Council (EPSRC), the Economic and Social Research Council (ESRC) and the MRC and Scottish Funding Council through the SINAPSE Collaboration is gratefully acknowledged. L.M.L. is the beneficiary of a postdoctoral grant from the AXA Research Fund.

MooDS: This work was funded by the German Federal Ministry of Education and Research (BMBF) in the National Genome Research Network (NGFN-plus) through the MooDs grant Molecular Causes of Major Mood Disorders and Schizophrenia (coordinator M.M.N.). Additional funding for genotyping was provided by a NARSAD Distinguished Investigator award to A.M.-L.

MPIP: The MPIP Munich Morphometry Sample comprises images acquired as part of the Munich Antidepressant Response Signature Study and the Recurrent Unipolar Depression (RUD) Case-Control Study performed at the MPIP and control subjects acquired at the Department of Psychiatry at the Ludwig-Maximilians-University. We wish to acknowledge A. Olynyik and radiographers R. Schirmer, E. Schreiter and R. Borschke for image acquisition and data preparation. We thank D.P. Auer for local study management in the initial phase of the RUD study. We are grateful to GlaxoSmithKline for providing the genotypes of the RUD Case-Control Sample. We thank the staff of the Center of Applied Genotyping (CAGT) for generating the genotypes of the MARS cohort. The study is supported by a grant from the Exzellenz-Stiftung of the Max Planck Society. This work has also been funded by the BMBF in the framework of the National Genome Research Network (NGFN) (FKZ 01GS0481).

NCNG: We would like to thank the personnel involved in recruitment and data collection and, in particular, P. Due-Tønnessen for clinical assessment of the MRI images. The NCNG study was supported by Research Council of Norway grants (154313/V50 and 177458/V50). The NCNG GWAS was financed by grants from the Bergen Research Foundation, the University of Bergen, the Research Council of Norway (FUGE; Psykisk Helse), Helse Vest Regionalt Helseforetak (RHF) and the Dr Einar Martens Fund.

NESDA-NTR: Funding was obtained from the NWO (MagW/ZonMW 904-61-090; 985-10-002; 904-61-193; 480-04-004; 400-05-717, Addiction-31160008; 911-09-032; SPI 56-464-14192 and Geestkracht Program, 10-000-1002), the Center for Medical Systems Biology (CMSB; NWO Genomics), NBIC/BioAssist/RK/2008.024, BBMRI-NL, Biobanking and Biomolecular Resources Research Infrastructure, the VU University, the EMGO Institute for Health and Care Research and Neuroscience Campus Amsterdam, the European Science Foundation (EU/QLRT-2001-01254), the European Community's FP7 (HEALTH-F4-2007-201413), the European Science Council (ERC) Genetics of Mental Illness (230374), Rutgers University Cell and DNA Repository (cooperative agreement NIMH U24 MH068457-06), the US NIH (R01D0042157-01A) and the the Genetic Association Information Network (a public-private partnership between the NIH and Pfizer, Affymetrix and Abbott Laboratories).

NIMH-IRP: This study was supported by funding from the Intramural Research Program of the National Institute of Mental Health (NIMH) from the NIH and the US Department of Health and Human Services (K99 MH085098 to G.L., 1ZIA MH002810 to F.J.M. and 1ZIA MH002790 to W.C.D.). The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products or organizations imply endorsement by the US government.

QTIM: We are extremely grateful to the twins for their participation, the radiographers at the Centre for Advanced Imaging at the University of Queensland for image acquisition and the many research assistants and support staff at the Queensland Institute of Medical Research for twin recruitment and daily management, and we especially thank K. Johnson for MRI scanning and processing, A. Henders for DNA processing and preparation and S. Gordon for quality control and management of the genotypes. Phenotyping was funded by the US National Institute of Child Health and Human Development (R01 HD050735) and the Australian National Health and Medical Research Council (NHMRC)

(project grant 496682). Genotyping was funded by the NHMRC (Medical Bioinformatics Genomics Proteomics Program, 389891). G.M. was supported by an NHMRC Fellowship (613667), and G.Z. was supported by Australian Research Council (ARC) Future Fellowship (FT0991634). S.E.M. is funded by an ARC Future Fellowship (FT110100548). J.L.S. was supported by the Achievement Rewards for College Scientists foundation and the US NIMH (F31 MH087061). D.P.H. is partially supported by a National Science Foundation (NSF) Graduate Research Fellowhip Program (GRFP) grant (DGE-0707424). P.T. was also supported by the NIH (grants U01 AG024904, AG040060, EB008432, P41 RR013642, HD050735, AG036535, AG020098 and EB008281).

SYS: The Saguenay Youth Study Group wishes to thank the following individuals for their contribution in acquiring and analyzing the data: N. Arbour, M.-È. Bouchard, A. Houde, A. Gauthier and H. Simard for the recruitment and assessment of participating families, M. Bérubé, S. Masson, S. Castonguay and M.-J. Morin for MRI acquisition and E. Ding and N. Qiu for MR data management. We thank J. Mathieu for the medical follow up of participants in whom we detected any medically relevant abnormalities. We are grateful to all families for participating in the study. The Saguenay Youth Study Group is supported by the Canadian Institutes of Health Research, the Heart and Stroke Foundation of Quebec and the Canadian Foundation for Innovation. For more information, please see the study website (see URLs).

SHIP: The Study of Health in Pomerania (SHIP) is supported by the German Federal Ministry of Education and Research (grants 01ZZ9603, 01ZZ0103 and 01ZZ0403) and the German Research Foundation (DFG; GR 1912/5-1). Genome-wide data and MRI scans were supported by the Federal Ministry of Education and Research (grant 03ZIK012) and a joint grant from Siemens Healthcare, Erlangen, Germany, and the Federal State of Mecklenburg–West Pomerania. The University of Greifswald is a member of the Center of Knowledge Interchange program of the Siemens AG. We thank all staff members and participants of the SHIP study, as well as all of the genotyping staff for generating the SHIP SNP data set. The genetic data analysis workflow was created using the Software InforSense. Genetic data were stored using the database Caché (InterSystems).

SHIP-TREND: The authors from SHIP are grateful to M. Stanke for the opportunity to use his Server Cluster for SNP Imputation. This cohort is part of the Community Medicine Research net (CMR) of the University of Greifswald, which is funded by the German Federal Ministry of Education and Research and the German Ministry of Cultural Affairs, as well as by the Social Ministry of the Federal State of Mecklenburg–West Pomerania. CMR encompasses several research projects that share data from the population-based Study of Health in Pomerania (SHIP; see URLs). The work is also supported by the German Research Foundation (DFG; GR 1912/5-1) and the Greifswald Approach to Individualized Medicine (GANI_MED) network funded by the Federal Ministry of Education and Research (grant 03IS2061A). Genome-wide data and MRI scans were supported by the Federal Ministry of Education and Research (grant 03ZIK012) and a joint grant from Siemens Healthcare, Erlangen, Germany, and the Federal State of Mecklenburg–West Pomerania. The University of Greifswald is a member of the Center of Knowledge Interchange program of the Siemens AG.

Superstruct: We thank the investigators and participants who contributed to the brain genomics data collection for Superstruct at Massachusetts General Hospital and Harvard University, with funding from the Simons Foundation, the Howard Hughes Medical Institute and the US NIH (grant MH079799).

TOP: We thank the study participants of TOP and the personnel involved in data collection and logistics, especially T.D. Bjella. This work was supported by the Oslo University Hospital–Ullevål, the Eastern Norway Health Authority (2004-123), the Research Council of Norway (167153/V50, 163070/V50 and 183782/V50), and by Eli Lilly & Company (who covered part of the genotyping costs).

TCD: We wish to express our sincere thanks to all participants and to clinical staff who facilitated patients' involvement. In particular, we acknowledge colleagues from the Trinity College Institute of Neuroscience A. Bodke, J. McGrath, F. Newell, H. Garavan, and J. O'Doherty for their support in sample collection. Collection and analysis of these samples were funded by the Wellcome Trust (072894/z/03/z-Gill) and the Science Foundation Ireland (08/IN.1/B1916_Corvin).

EPIGEN: Work from the London Cohort was supported by research grants from the Wellcome Trust (grant 084730 to S.M.S.), University College London (UCL)/University College London Hospitals (UCLH) Comprehensive Biomedical Research Centre/Specialist Biomedical Research Centres (CBRC/SBRC) (grant 114 to S.M.S.), the European Union Marie Curie Reintegration (to M. Matarin and S.M.S.), the UK NIHR (08-08-SCC), the Comprehensive Local Research Network (CLRN) Flexibility and Sustainability Funding (FSF) (grant CEL1300 to S.M.S.), The Big Lottery Fund, the Wolfson Trust and the Epilepsy Society. This work was undertaken at UCLH/UCL, which received a proportion of funding from the UK Department of Health's NIHR Biomedical Research Centres funding scheme. Work from the Royal College of Surgeons in Ireland was supported by research grants from the Science Foundation Ireland (Research Frontiers Programme award 08/RFP/GEN1538) and Brainwave–the Irish Epilepsy Association. The collection of Belgian subjects was supported by the Fonds National de la Recherche Scientifique (grant FC 63574 / 3.4.620.06 F) and the Fonds Erasme pour la Recherche Médicale at the Université Libre de Bruxelles.

UCL Institute of Neurology Control Brain Tissue Collection: Funding was provided by the UK MRC (grant G0901254), the MRC Sudden Death Brain and Tissue Bank and the Sun Health Research Institute Brain Bank.

UMCU: The UMCU study was supported by the Netherlands Organization for Health Research and Development ZonMw (917.46.370 to H.E.H.) and the US NIMH (MH078075 to R.A.O.).

AUTHOR CONTRIBUTIONS

The ENIGMA support group designed the project, established the consortium, determined the analysis and quality control procedures, offered analytical support and performed and coordinated cross-site and replication analyses. This group included J.L.S., S.E.M., A.A.V., D.P.H., M.J.W., B.F., N.G.M. and P.M.T. The imaging protocols group determined and refined protocols for computing brain measures from the MRI scans and helped sites implement them as needed. This group included J.L.S., R.T., A.M.W., T.E.N., M.J. and M. Rijpkema. The genetics protocols group created analysis methods for imputation, quality control and association testing of genome-wide data and helped to ensure that protocols were implemented consistently across all sites. This group included S.E.M., J.L.S., A.A.V. and D.P.H. The meta-analysis was carried out by the meta-analysis group, consisting of S.E.M., R.E.S., J.L.S., D.P.H., A.A.V., M.J.W., N.G.M., B.F. and P.M.T. The first draft was written by J.L.S., S.E.M., A.A.V., D.P.H., M.J.W., B.F., N.G.M. and P.M.T. Local image processing, involving statistical analysis and analysis of the data, was performed by J.L.S., A.M.W., D.P.H., R.B., Ø.B., M.M.C., O.G., M. Hollinshead, A.J.H., S.M.M., A.C.N., M. Rijpkema, N.A.R., M.C.V.H., T.G.M.v.E., S.W., D.G.B., S.L.R., J.L.R., M.-J.v.T., A.A., S.E., P.T.F., P.K., J.L.L., R.M., G.B.P., J. Savitz, H.G.S., K.S., A.W.T., M.V.d.H., N.J.v.d.W., N.E.M.V.H., H.W., A.M.D., C.R.J., D.J.V., E.J.C.d.G., G.I.d.Z., T.E., G.F., P.H., H.E.H.P., K.L.M., A.J.S., L.S., J.B., D.C.G., K.N., E.L., A.M.-L., P.G.S., L.G.A., K.S.H., T.P., M.D., R.P., N.H., K.W., I.A., Ø.B., A.M.D., D.H., M.C., S.A., N.D., C. Depondt, M. Pandolfo, E.J.R., D.M.C., J.C.R., J.R., J.T., R.T., C.L., S.M., A.H., C.D.W., N.J., D.J.H., L.T.W. and M. Hoogman. Local genetics processing, involving statistical analysis and analysis of the data, was performed by J.L.S., S.E.M., A.A.V., A.M.W., D.P.H., M.B., A.A.B., A. Christoforou, G. Davies, J.-J.H., L.M.L., G.L., P.H.L., D.C.L., X.L., M. Mattingsdal, K.N., E. Strengman, K.v.E., T.G.M.v.E., S.W., S.K., L.A., R.M.C., M.A.C., J.E.C., R.D., T.D.D., N.B.F., H.H.H.G., M.P.J., J.W.K., M. Mattheisen, E.K.M., T.W.M., M.M.N., M. Rietschel, V.M.S., A.W.T., J.A.V., S.C., S.D., T.M.F., P.H., S.L.H., G.W.M., O.A.A., H.G.B., R.A.O., B.W.P., A.J.S., L.S., J.B., D.C.G., M.J.W., N.G.M., A.L., E.B.B., C.W., B.P., B.M.-M., G.C., Z.P., G.H., M.N., A.T., D.K., M. Matarin, S.M.S., G.L.C., N.K.H., M.E.R., D.W.M., C.O., A. Corvin, M.G., J.F., J.C.R., A.R., M. Ryten, D.T., N.S., C.S., R.W., J. Hardy, M.E.W. and M.A.A.d.A. Local study oversight and management, involving joint supervision of research, contribution of reagents, materials and/or analysis tools, was carried out by R.L.B., R.D., P.T.F., R.S.K., I.M., R.L.O., I.R., I.A., W.C.D., P.H., F.M., A.M.-L., D.J.P., S.G.P., J.M.S., M.W.W., O.A.A., M.E.B., H.G.B., E.J.C.d.G., I.J.D., G.I.d.Z., T.E., G.F., H.E.H.P., F.J.M., K.L.M., R.A.O., T.P., Z.P., B.W.P., A.J.S., L.S., J.W.S., J.M.W., J.B., D.C.G., M.J.W., B.F., P.M.T., A.M.M., J. Hall, M. Papmeyer, E. Sprooten, J. Sussmann, S.M.L., J.B.P., L.G.A., G.C., D.R., E.M., G.S., K.S.H., P.G.S., E.B.B., D.I.B., H.J.G., H.V., K.A., C.M., G. Donohoe, F.H., A.V.S., V.G., C.T., M.W.V., L.J.L., C. DeCarli, S.S., J.C.B., M.A.I., J. Hardy.

COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Published online at http://www.nature.com/naturegenetics/.
Reprints and permissions information is available online at http://www.nature.com/reprints/index.html.

1. Jack, C.R. Jr. *et al.* Steps to standardization and validation of hippocampal volumetry as a biomarker in clinical trials and diagnostic criterion for Alzheimer's disease. *Alzheimers Dement.* **7**, 474–485 e4 (2011).
2. Simić, G., Kostovic, I., Winblad, B. & Bogdanovic, N. Volume and number of neurons of the human hippocampal formation in normal aging and Alzheimer's disease. *J. Comp. Neurol.* **379**, 482–494 (1997).
3. Wright, I.C. *et al.* Meta-analysis of regional brain volumes in schizophrenia. *Am. J. Psychiatry* **157**, 16–25 (2000).

4. Videbech, P. & Ravnkilde, B. Hippocampal volume and depression: a meta-analysis of MRI studies. *Am. J. Psychiatry* **161**, 1957–1966 (2004).
5. Keller, S.S. & Roberts, N. Voxel-based morphometry of temporal lobe epilepsy: an introduction and review of the literature. *Epilepsia* **49**, 741–757 (2008).
6. Peper, J.S., Brouwer, R.M., Boomsma, D.I., Kahn, R.S. & Hulshoff Pol, H.E. Genetic influences on human brain structure: a review of brain imaging studies in twins. *Hum. Brain Mapp.* **28**, 464–473 (2007).
7. Kremen, W.S. *et al.* Genetic and environmental influences on the size of specific brain regions in midlife: the VETSA MRI study. *Neuroimage* **49**, 1213–1223 (2010).
8. Maguire, E.A. *et al.* Navigation-related structural change in the hippocampi of taxi drivers. *Proc. Natl. Acad. Sci. USA* **97**, 4398–4403 (2000).
9. Burgess, N., Maguire, E.A. & O'Keefe, J. The human hippocampus and spatial and episodic memory. *Neuron* **35**, 625–641 (2002).
10. Snyder, J.S., Soumier, A., Brewer, M., Pickel, J. & Cameron, H.A. Adult hippocampal neurogenesis buffers stress responses and depressive behaviour. *Nature* **476**, 458–461 (2011).
11. Freitag, C.M. *et al.* Total brain volume and corpus callosum size in medication-naive adolescents and young adults with autism spectrum disorder. *Biol. Psychiatry* **66**, 316–319 (2009).
12. Stanfield, A.C. *et al.* Towards a neuroanatomy of autism: a systematic review and meta-analysis of structural magnetic resonance imaging studies. *Eur. Psychiatry* **23**, 289–299 (2008).
13. Posthuma, D. *et al.* The association between brain volume and intelligence is of genetic origin. *Nat. Neurosci.* **5**, 83–84 (2002).
14. Fears, S.C. *et al.* Identifying heritable brain phenotypes in an extended pedigree of vervet monkeys. *J. Neurosci.* **29**, 2867–2875 (2009).
15. Rogers, J. *et al.* On the genetic architecture of cortical folding and brain volume in primates. *Neuroimage* **53**, 1103–1108 (2010).
16. Patenaude, B., Smith, S.M., Kennedy, D.N. & Jenkinson, M. A Bayesian model of shape and appearance for subcortical brain segmentation. *Neuroimage* **56**, 907–922 (2011).
17. Fischl, B. *et al.* Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain. *Neuron* **33**, 341–355 (2002).
18. Buckner, R.L. *et al.* A unified approach for morphometric and functional data analysis in young, old, and demented adults using automated atlas-based head size normalization: reliability and validation against manual measurement of total intracranial volume. *Neuroimage* **23**, 724–738 (2004).
19. Willer, C.J., Li, Y. & Abecasis, G.R. METAL: fast and efficient meta-analysis of genome-wide association scans. *Bioinformatics* **26**, 2190–2191 (2010).
20. Han, B. & Eskin, E. Random-effects model aimed at discovering associations in meta-analysis of genome-wide association studies. *Am. J. Hum. Genet.* **88**, 586–598 (2011).
21. Bis, J.C. *et al.* Common variants at 12q14 and 12q24 are associated with hippocampal volume. *Nat. Genet.* published online (15 April 2012; doi:10.1038/ng.2237).
22. Ripke, S. *et al.* Genome-wide association study identifies five new schizophrenia loci. *Nat. Genet.* **43**, 969–976 (2011).
23. Sklar, P. *et al.* Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near *ODZ4*. *Nat. Genet.* **43**, 977–983 (2011).
24. Li, M.X., Gui, H.S., Kwan, J.S. & Sham, P.C. GATES: a rapid and powerful gene-based association test using extended Simes procedure. *Am. J. Hum. Genet.* **88**, 283–293 (2011).
25. Bao, Y. *et al.* Expression and evolutionary conservation of the tescalcin gene during development. *Gene expression patterns. Gene Exp. Patterns* **9**, 273–281 (2009).
26. Baumgartner, M., Patel, H. & Barber, D.L. Na^+/H^+ exchanger NHE1 as plasma membrane scaffold in the assembly of signaling complexes. *Am. J. Physiol. Cell Physiol.* **287**, C844–C850 (2004).
27. Slepkov, E.R., Rainey, J.K., Sykes, B.D. & Fliegel, L. Structural and functional analysis of the Na^+/H^+ exchanger. *Biochem. J.* **401**, 623–633 (2007).
28. Levay, K. & Slepak, V.Z. Tescalcin is an essential factor in megakaryocytic differentiation associated with Ets family gene expression. *J. Clin. Invest.* **117**, 2672–2683 (2007).
29. Levay, K. & Slepak, V.Z. Up- or downregulation of tescalcin in HL-60 cells is associated with their differentiation to either granulocytic or macrophage-like lineage. *Exp. Cell Res.* **316**, 1254–1262 (2010).
30. Lango Allen, H. *et al.* Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature* **467**, 832–838 (2010).
31. Gudbjartsson, D.F. *et al.* Many sequence variants affecting diversity of adult human height. *Nat. Genet.* **40**, 609–615 (2008).
32. Sanna, S. *et al.* Common variants in the *GDF5-UQCC* region are associated with variation in human height. *Nat. Genet.* **40**, 198–203 (2008).
33. Weedon, M.N. *et al.* Genome-wide association analysis identifies 20 loci that influence adult height. *Nat. Genet.* **40**, 575–583 (2008).
34. Fusco, A. & Fedele, M. Roles of HMGA proteins in cancer. *Nat. Rev. Cancer* **7**, 899–910 (2007).
35. Litterman, N. *et al.* An OBSL1-Cul7Fbxw8 ubiquitin ligase signaling mechanism regulates Golgi morphology and dendrite patterning. *PLoS Biol.* **9**, e1001060 (2011).
36. Hammond, S.M. & Sharpless, N.E. HMGA2, microRNAs, and stem cell aging. *Cell* **135**, 1013–1016 (2008).
37. Wright, M.J. & Martin, N.G. Brisbane adolescent twin study: outline of study methods and research projects. *Aust. J. Psychol.* **56**, 65–78 (2004).
38. Jackson, D.N. *MAB: Multidimensional Aptitude Battery Manual* (Research Psychologists Press, Port Huron, Michigan, 1984).
39. Vogel, W. Discoidin domain receptors: structural relations and functional implications. *FASEB J.* **13** (suppl), S77–S82 (1999).
40. Hindorff, L.A. *et al.* Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc. Natl. Acad. Sci. USA* **106**, 9362–9367 (2009).
41. Pruim, R.J. *et al.* LocusZoom: regional visualization of genome-wide association scan results. *Bioinformatics* **26**, 2336–2337 (2010).
42. Johnson, M.B. *et al.* Functional and evolutionary insights into human brain development through global transcriptome analysis. *Neuron* **62**, 494–509 (2009).

Jason L Stein[1,127], Sarah E Medland[2–4,127], Alejandro Arias Vasquez[5–7,127], Derrek P Hibar[1,127], Rudy E Senstad[1], Anderson M Winkler[8,9], Roberto Toro[10–12], Katja Appel[13,14], Richard Bartecek[15], Ørjan Bergmann[16], Manon Bernard[17], Andrew A Brown[16,18], Dara M Cannon[19], M Mallar Chakravarty[21], Andrea Christoforou[22,23], Martin Domin[24], Oliver Grimm[25], Marisa Hollinshead[26,27], Avram J Holmes[26], Georg Homuth[28], Jouke-Jan Hottenga[29], Camilla Langan[20], Lorna M Lopez[30,31], Narelle K Hansell[2], Kristy S Hwang[1,32], Sungeun Kim[33,34], Gonzalo Laje[35], Phil H Lee[36,37], Xinmin Liu[35,38], Eva Loth[39], Anbarasu Lourdusamy[39], Morten Mattingsdal[16,40], Sebastian Mohnke[41], Susana Muñoz Maniega[30,42,43], Kwangsik Nho[33,44], Allison C Nugent[45], Carol O'Brien[46,47], Martina Papmeyer[48], Benno Pütz[49], Adaikalavan Ramasamy[50], Jerod Rasmussen[51], Mark Rijpkema[7,52], Shannon L Risacher[33], J Cooper Roddey[53], Emma J Rose[46,47], Mina Ryten[54], Li Shen[33,34], Emma Sprooten[48], Eric Strengman[55,56], Alexander Teumer[28], Daniah Trabzuni[54,57], Jessica Turner[58], Kristel van Eijk[55,56], Theo G M van Erp[51], Marie-Jose van Tol[59–61], Katharina Wittfeld[13], Christiane Wolf[49], Saskia Woudstra[62], Andre Aleman[61], Saud Alhusaini[63], Laura Almasy[64], Elisabeth B Binder[49], David G Brohawn[36], Rita M Cantor[65], Melanie A Carless[64], Aiden Corvin[46,47], Michael Czisch[49], Joanne E Curran[64], Gail Davies[31], Marcio A A de Almeida[64], Norman Delanty[63,66], Chantal Depondt[67], Ravi Duggirala[64], Thomas D Dyer[64], Susanne Erk[41], Jesen Fagerness[36], Peter T Fox[69], Nelson B Freimer[65], Michael Gill[46,47], Harald H H Göring[64], Donald J Hagler[70], David Hoehn[49], Florian Holsboer[49], Martine Hoogman[5,7,71,72], Norbert Hosten[24], Neda Jahanshad[1], Matthew P Johnson[64], Dalia Kasperaviciute[73], Jack W Kent Jr[64], Peter Kochunov[69,74], Jack L Lancaster[69], Stephen M Lawrie[48], David C Liewald[30], René Mandl[15], Mar Matarin[73], Manuel Mattheisen[75–77], Eva Meisenzahl[78], Ingrid Melle[16,79], Eric K Moses[64], Thomas W Mühleisen[75,76], Matthias Nauck[80], Markus M Nöthen[75,76,81], Rene L Olvera[82], Massimo Pandolfo[67],

G Bruce Pike[83], Ralf Puls[24], Ivar Reinvang[84,85], Miguel E Rentería[2,86], Marcella Rietschel[25], Joshua L Roffman[37], Natalie A Royle[30,42,43], Dan Rujescu[78], Jonathan Savitz[45,87], Hugo G Schnack[15], Knut Schnell[88,89], Nina Seiferth[41], Colin Smith[90], Vidar M Steen[22,23], Maria C Valdés Hernández[30,42,43], Martijn Van den Heuvel[15], Nic J van der Wee[59,60], Neeltje E M Van Haren[15], Joris A Veltman[5], Henry Völzke[91], Robert Walker[90], Lars T Westlye[84], Christopher D Whelan[63], Ingrid Agartz[16,92], Dorret I Boomsma[29], Gianpiero L Cavalleri[63], Anders M Dale[53,70], Srdjan Djurovic[16,93], Wayne C Drevets[45,87], Peter Hagoort[7,52,72], Jeremy Hall[48], Andreas Heinz[41], Clifford R Jack Jr[94], Tatiana M Foroud[34,95], Stephanie Le Hellard[22,23], Fabio Macciardi[51], Grant W Montgomery[2], Jean Baptiste Poline[96], David J Porteous[30,97], Sanjay M Sisodiya[73], John M Starr[30,98], Jessika Sussmann[48], Arthur W Toga[1], Dick J Veltman[62], Henrik Walter[41,89], Michael W Weiner[99–102], the Alzheimer's Disease Neuroimaging Initiative (ADNI)[103], EPIGEN Consortium[103], IMAGEN Consortium[103], Saguenay Youth Study Group (SYS)[103], Joshua C Bis[104], M Arfan Ikram[105–107], Albert V Smith[108,109], Vilmundur Gudnason[108,109], Christophe Tzourio[110,111], Meike W Vernooij[105–107], Lenore J Launer[112], Charles DeCarli[113,114], Sudha Seshadri[115,116], Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium[103], Ole A Andreassen[16,79], Liana G Apostolova[1,32], Mark E Bastin[30,42,43,117], John Blangero[64], Han G Brunner[5], Randy L Buckner[26,27,37,68], Sven Cichon[75,76,118], Giovanni Coppola[32,119], Greig I de Zubicaray[86], Ian J Deary[30,31], Gary Donohoe[46,47], Eco J C de Geus[29], Thomas Espeseth[84,85,120], Guillén Fernández[7,52,71], David C Glahn[8,9], Hans J Grabe[13,121], John Hardy[54], Hilleke E Hulshoff Pol[15], Mark Jenkinson[122], René S Kahn[15], Colm McDonald[20], Andrew M McIntosh[48], Francis J McMahon[35], Katie L McMahon[123], Andreas Meyer-Lindenberg[25], Derek W Morris[46,47], Bertram Müller-Myhsok[49], Thomas E Nichols[122,124], Roel A Ophoff[15,65], Tomas Paus[21], Zdenka Pausova[17], Brenda W Penninx[59,60,62,125], Steven G Potkin[51], Philipp G Sämann[49], Andrew J Saykin[33,34,96], Gunter Schumann[39], Jordan W Smoller[36,37], Joanna M Wardlaw[30,42,43], Michael E Weale[50], Nicholas G Martin[2,125], Barbara Franke[5–7,128], Margaret J Wright[2,128] & Paul M Thompson[1,128] for the Enhancing Neuro Imaging Genetics through Meta-Analysis (ENIGMA) Consortium[126]

[1]Laboratory of Neuro Imaging, David Geffen School of Medicine, University of California, Los Angeles, California, USA. [2]Genetic Epidemiology Laboratory, Queensland Institute of Medical Research, Brisbane, Queensland, Australia. [3]Quantitative Genetics Laboratory, Queensland Institute of Medical Research, Brisbane, Queensland, Australia. [4]Broad Institute of Harvard University and MIT, Cambridge, Massachusetts, USA. [5]Department of Human Genetics, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands. [6]Department of Psychiatry, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands. [7]Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, Nijmegen, The Netherlands. [8]Olin Neuropsychiatry Research Center, Institute of Living, Hartford Hospital, Hartford, Connecticut, USA. [9]Department of Psychiatry, Yale University School of Medicine, New Haven, Connecticut, USA. [10]Laboratory of Human Genetics and Cognitive Functions, Institut Pasteur, Paris, France. [11]Centre Nationale de Recherche Scientifique (CNRS) Unité de Recherche Associée (URA) 2182 Genes, Synapses and Cognition, Institut Pasteur, Paris, France. [12]Department of Neuroscience, Université Paris Diderot, Sorbonne Paris Cité, Paris, France. [13]Department of Psychiatry and Psychotherapy, University of Greifswald, Greifswald, Germany. [14]Institute of Psychology, Department of Clinical Psychology and Psychotherapy, University of Heidelberg, Heidelberg, Germany. [15]Department of Psychiatry, Rudolf Magnus Institute, University Medical Center Utrecht, Utrecht, The Netherlands. [16]Institute of Clinical Medicine, University of Oslo, Oslo, Norway. [17]The Hospital for Sick Children, University of Toronto, Toronto, Ontario, Canada. [18]Institute of Basic Medical Sciences, Department of Biostatistics, University of Oslo, Oslo, Norway. [19]Clinical Neuroimaging Laboratory, Department of Anatomy, National University of Ireland Galway, Galway, Ireland. [20]Clinical Neuroimaging Laboratory, Department of Psychiatry, National University of Ireland Galway, Galway, Ireland. [21]Rotman Research Institute, University of Toronto, Toronto, Ontario, Canada. [22]Dr Einar Martens Research Group for Biological Psychiatry, Department of Clinical Medicine, University of Bergen, Bergen, Norway. [23]Center for Medical Genetics and Molecular Medicine, Haukeland University Hospital, Bergen, Norway. [24]Department of Diagnostic Radiology and Neuroradiology, University of Greifswald, Greifswald, Germany. [25]Central Institute of Mental Health, University of Heidelberg–Medical Faculty Mannheim, Mannheim, Germany. [26]Department of Psychology, Center for Brain Science, Harvard University, Cambridge, Massachusetts, USA. [27]Howard Hughes Medical Institute, Cambridge, Massachusetts, USA. [28]Interfaculty Institute for Genetics and Functional Genomics, University of Greifswald, Greifswald, Germany. [29]Department of Biological Psychology, Neuroscience Campus Amsterdam, VU University, Amsterdam, The Netherlands. [30]Centre for Cognitive Ageing and Cognitive Epidemiology, The University of Edinburgh, Edinburgh, UK. [31]Department of Psychology, The University of Edinburgh, Edinburgh, UK. [32]Department of Neurology, David Geffen School of Medicine, University of California, Los Angeles, California, USA. [33]Department of Radiology and Imaging Sciences, Center for Neuroimaging, Indiana University School of Medicine, Indianapolis, Indiana, USA. [34]Center for Computational Biology and Bioinformatics, Indiana University School of Medicine, Indianapolis, Indiana, USA. [35]Mood and Anxiety Disorders Section, Human Genetics Branch, Intramural Research Program, National Institute of Mental Health (NIMH), US National Institutes of Health (NIH), US Department of Health and Human Services, Bethesda, Maryland, USA. [36]Psychiatric and Neurodevelopmental Genetics Unit, Center for Human Genetic Research, Massachusetts General Hospital, Boston, Massachusetts, USA. [37]Department of Psychiatry, Massachusetts General Hospital, Boston, Massachusetts, USA. [38]Taub Institute for Research on Alzheimer Disease and the Aging Brain, Columbia University Medical Center, New York, New York, USA. [39]Medical Research Council (MRC)–Social, Genetic and Developmental Psychiatry (SGDP) Centre, Institute of Psychiatry, King's College London, London, UK. [40]Research Unit, Sørlandet Hospital, Kristiansand, Norway. [41]Department of Psychiatry and Psychotherapy, Charité-Universitaetsmedizin Berlin, Campus Mitte, Berlin, Germany. [42]Scottish Imaging Network, A Platform for Scientific Excellence (SINAPSE) Collaboration, UK. [43]Brain Research Imaging Centre, The University of Edinburgh, Edinburgh, UK. [44]Division of Medical Informatics, Regenstrief Institute, Indianapolis, Indiana, USA. [45]Section on Neuroimaging in Mood and Anxiety Disorders, Intramural Research Program, NIMH, NIH, US Department of Health and Human Services, Bethesda, Maryland, USA. [46]Neuropsychiatric Genetics Research Group, Department of Psychiatry, Institute for Molecular Medicine, Trinity College, Dublin, Ireland. [47]Trinity College Institute of Neuroscience, Trinity College, Dublin, Ireland. [48]Division of Psychiatry, University of Edinburgh, Royal Edinburgh Hospital, Edinburgh, UK. [49]Max Planck Institute of Psychiatry, Munich, Germany. [50]Department of Medical & Molecular Genetics, King's College London, London, UK. [51]Department of Psychiatry and Human Behavior, University of California, Irvine, California, USA. [52]Donders Centre for Cognitive Neuroimaging, Radboud University Nijmegen, Nijmegen, The Netherlands. [53]Department of Neurosciences, University of California, San Diego, La Jolla, California, USA. [54]Department of Molecular Neuroscience, University College London, London, UK. [55]Department of Medical Genetics, University Medical Center Utrecht, Utrecht, The Netherlands. [56]Rudolf Magnus Institute, University Medical Center Utrecht, Utrecht, The Netherlands. [57]Department of Genetics, King Faisal Specialist Hospital and Research Centre, Riyadh, Saudi Arabia. [58]Mind Research Network, Albuquerque, New Mexico, USA. [59]Department of Psychiatry, Leiden University Medical Center, Leiden, The Netherlands. [60]Leiden Institute for Brain and Cognition, Leiden University, Leiden, The Netherlands. [61]Behavioural and Cognitive Neuroscience Neuroimaging Center, University Medical Center Groningen, Groningen,

The Netherlands. [62]Department of Psychiatry, VU University Medical Center, Amsterdam, The Netherlands. [63]Department of Molecular and Cellular Therapeutics, Royal College of Surgeons in Ireland, Dublin, Ireland. [64]Department of Genetics, Texas Biomedical Research Institute, San Antonio, Texas, USA. [65]Center for Neurobehavioral Genetics, University of California, Los Angeles, USA. [66]Division of Neurology, Beaumont Hospital, Dublin, Ireland. [67]Department of Neurology, Hopital Erasme, Universite Libre de Bruxelles, Brussels, Belgium. [68]Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Boston, Massachusetts, USA. [69]Research Imaging Institute, University of Texas Health Science Center at San Antonio, San Antonio, Texas, USA. [70]Department of Radiology, University of California, San Diego, La Jolla, California, USA. [71]Department of Cognitive Neuroscience, Radboud University Nijmegen Medical Centre, The Netherlands. [72]Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands. [73]Department of Clinical and Experimental Epilepsy, University College London, Institute of Neurology, London, UK. [74]Maryland Psychiatric Research Center, Department of Psychiatry, University of Maryland School of Medicine, Baltimore, Maryland, USA. [75]Department of Genomics, Life and Brain Center, University of Bonn, Bonn, Germany. [76]Institute of Human Genetics, University of Bonn, Bonn, Germany. [77]Institute for Genomic Mathematics, University of Bonn, Bonn, Germany. [78]Department of Psychiatry, Ludwig-Maximilians-University (LMU), Munich, Germany. [79]Division of Mental Health and Addiction, Oslo University Hospital, Oslo, Norway. [80]Institute of Clinical Chemistry and Laboratory Medicine, University of Greifswald, Greifswald, Germany. [81]German Center for Neurodegenerative Disorders (DZNE), Bonn, Germany. [82]Department of Psychiatry, University of Texas Health Science Center at San Antonio, San Antonio, Texas, USA. [83]Montreal Neurological Institute, McGill University, Montreal, Quebec, Canada. [84]Center for the Study of Human Cognition, Department of Psychology, University of Oslo, Oslo, Norway. [85]Centre for Advanced Study, Oslo, Norway. [86]School of Psychology, University of Queensland, Brisbane, Queensland, Australia. [87]Laureate Institute for Brain Research, Tulsa, Oklahoma, USA. [88]Department of General Psychiatry, Heidelberg University Hospital, University of Heidelberg, Heidelberg, Germany. [89]Department of Psychiatry, Division of Medical Psychology, Bonn, Germany. [90]The MRC Sudden Death Tissue Bank in Edinburgh, Department of Pathology, University of Edinburgh, Edinburgh, UK. [91]Institute for Community Medicine, University of Greifswald, Greifswald, Germany. [92]Department of Research and Development, Diakonhjemmet Hospital, Oslo, Norway. [93]Department of Medical Genetics, Oslo University Hospital, Oslo, Norway. [94]Aging and Dementia Imaging Research Laboratory, Department of Radiology, Mayo Clinic and Foundation, Rochester, Minnesota, USA. [95]Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, Indiana, USA. [96]Neurospin, Institut d'Imagerie Biomédicale (I2BM), Commissariat à l'Energie Atomique, Gif-sur-Yvette, France. [97]Medical Genetics Section, Molecular Medicine Centre, Institute of Genetics and Molecular Medicine, The University of Edinburgh, Western General Hospital, Edinburgh, UK. [98]Geriatric Medicine Unit, The University of Edinburgh, Royal Victoria Hospital, Edinburgh, UK. [99]Departments of Radiology, University of California, San Francisco, California, USA. [100]Department of Medicine, University of California, San Francisco, California, USA. [101]Department of Psychiatry, University of California, San Francisco, California, USA. [102]Veterans Affairs Medical Center, San Francisco, California, USA. [103]A full list of members is provided in the **Supplementary Note**. [104]Cardiovascular Health Research Unit, Department of Medicine, University of Washington, Seattle, Washington, USA. [105]Department of Epidemiology, Erasmus Medical Center University Medical Center, Rotterdam, The Netherlands. [106]Department of Radiology, Erasmus Medical Center University Medical Center, Rotterdam, The Netherlands. [107]Netherlands Consortium for Healthy Aging, Leiden, The Netherlands. [108]Icelandic Heart Association, Kopavogur, Iceland. [109]Faculty of Medicine, University of Iceland, Reykjavik, Iceland. [110]University of Brodeaux, U708, Bordeaux, France. [111]Institut National de la Santé et la Recherche Médicale (INSERM), Neuroepidemiology, U708, Bordeaux, France. [112]Laboratory of Epidemiology, Demography, and Biometry, NIH, Bethesda, Maryland, USA. [113]Department of Neurology, University of California, Davis, Sacramento, California, USA. [114]Center of Neuroscience, University of California, Davis, Sacramento, California, USA. [115]Department of Neurology, Boston University School of Medicine, Boston, Massachusetts, USA. [116]National Heart, Lung, and Blood Institute's Framingham Heart Study, Framingham, Massachusetts, USA. [117]Division of Health Sciences (Medical Physics), The University of Edinburgh, Edinburgh, UK. [118]Institute for Neuroscience and Medicine (INM-1), Research Center Juelich, Juelich, Germany. [119]Semel Institute for Neuroscience and Human Behavior, David Geffen School of Medicine, University of California, Los Angeles, California, USA. [120]Department of Biological and Medical Psychology, Faculty of Psychology, University of Bergen, Bergen, Norway. [121]German Center for Neurodegenerative Diseases (DZNE), Rostock/Greifswald, Greifwald, Germany. [122]Functional Magnetic Resonance Imaging of the Brain (FMRIB) Centre, Oxford University, Oxford, UK. [123]Centre for Advanced Imaging, University of Queensland, Brisbane, Queensland, Australia. [124]Department of Statistics, University of Warwick, Coventry, UK. [125]Department of Psychiatry, University Medical Center Groningen, Groningen, The Netherlands. [126]Information on the consortium is provided in the **Supplementary Note**. [127]These authors contributed equally to this work. [128]These authors jointly directed this work. Correspondence should be addressed to P.M.T. (thompson@loni.ucla.edu).

ONLINE METHODS

All participants provided written informed consent, and studies were approved by the respective Local Research Ethics committees or Institutional Review Boards. MRI scans came from previously collected data. Suggested protocols for imaging analysis are publicly available on the ENIGMA Consortium website (see URLs); however, any validated segmentation software was permitted. Accuracy of segmentation programs is influenced by scanner and head-coil type and scanner sequences and by participant characteristics, such as age. Each site was permitted to use any validated automated segmentation algorithm that worked most accurately in their data set. The two most commonly used hippocampal segmentation packages were the FMRIB's Integrated Registration and Segmentation Tool (FIRST)[16] from the FMRIB Software Library (FSL) package of tools[43] and FreeSurfer[17]. Brain volume, the sum of gray and white matter excluding ventricles and cerebrospinal fluid (CSF), was calculated using the FSL FMRIB's Automated Segmentation Tool (FAST)[44] package or FreeSurfer. Estimated total intracranial volume was calculated through registration of each MRI scan to a standard brain image template[18], using either FSL FLIRT[45] or FreeSurfer (exceptions referenced in **Supplementary Table 2**). To calculate intracranial volume, the inverse of the determinant of the transformation matrix was multiplied by the template volume (1,948,105 mm^3). Extensive quality control analysis on phenotype segmentations included manual examination of phenotype volume histograms (**Supplementary Figs. 2–6**) and box plots of all volumetric measures. Outliers were manually evaluated by overlaying the automated segmentations on the original MRI scan. Subjects were excluded from the analysis if structures were poorly segmented.

As assessed previously, the correlation in volumes between automatic and manually segmented hippocampi was high; the accuracy was reported to be higher with FreeSurfer than with FIRST in one study (FreeSurfer $r = 0.82$; FIRST $r = 0.66$)[46] and similar between the two in another (FreeSurfer $r = 0.73$; FIRST $r = 0.71$)[47]. Scan-rescan reliability was also high for both methods (FreeSurfer intraclass correlation (ICC) = 0.98; FIRST ICC = 0.93)[48]. We undertook a large-scale assessment to determine the correspondence between segmentations from both FSL and FreeSurfer in the same subjects. Correspondence was found to be reasonably high for average bilateral hippocampal segmentation ($r = 0.75$; $N = 6{,}093$; **Supplementary Table 4**). This is close to the agreement between different human raters, as quantified by inter-rater reliability (ICC = 0.73–0.85)[49,50], which may be a reasonable upper bound on the accuracy of automated segmentation. Brain volume and intracranial volume were delineated with high correspondence between the two methods ($r = 0.95$, $r = 0.90$, respectively; $N = 4{,}321$).

Heritability estimates for trait measures were calculated in two family-based samples, QTIM and GOBS. Estimates for the QTIM sample used a twin and sibling analysis within Mx. An extended family analysis in Sequential Oligogenic Linkage Analysis Routines (SOLAR)[51] was used for the GOBS sample.

Given sample size and the heritability of hippocampal volume, power calculations were performed using the Genetic Power Calculator[52]. We had 99.92% power to detect variants with effect sizes of 1% of the variance and 71.16% power to detect variants with effect sizes of 0.5% of the variance for MAF ≥ 0.05.

All cohorts were genotyped using commercially available arrays. Genetics protocols were developed to standardize the filtering, imputation and association of genome-wide genotype data (see ENIGMA protocols in URLs). SNPs were filtered out of samples on the basis of standard quality control criteria, including low MAF (<0.01), poor genotype calling (call rate of <95%) and deviations from Hardy-Weinberg equilibrium indicating possible errors in genotyping ($P < 1 \times 10^{-6}$). Genotyping methods and exceptions to these thresholds are summarized in **Supplementary Table 3**.

Genetic homogeneity within each sample was assessed through MDS plots (**Supplementary Fig. 7**). Ancestry outliers were excluded through visual inspection. A standardized population template from HapMap 3 representing those sampled was selected for imputation. Performance of software for imputation is generally similar between the most used methods[53,54] for common variants (MACH[55], IMPUTE[56] and BEAGLE[57]); the protocols provided included use of the MACH tool. As raw genotype data were not directly transferred to the meta-analysis site, a histogram of allele frequency differences between each contributing group and the HapMap 3 CEU population was generated for each group (**Supplementary Fig. 8**) to further examine genotyping and imputation quality. A simulation to determine the effect of varying quality control thresholds on imputation quality (**Supplementary Table 6**) showed that the minor variation in quality control thresholds and imputation reference panels between sites was unlikely to have influenced imputation accuracy.

Genome-wide association analyses were performed that included and excluded individuals with disease. Including individuals with disease (all subjects) offers advantages of greater sample size and wider phenotype distribution, which may provide greater power to detect genetic effects[58–60]. We reanalyzed phenotypes after we excluded individuals with disease to confirm that the observed associations were not due to confounds relating to disease, medication or the possibly altered environments and experiences of these persons. To aid in the interpretation of results, we reanalyzed hippocampal volume after controlling for intracranial volume and total brain volume in two separate analyses. This helped to determine whether the observed associations were caused by direct effects on hippocampal volume or were attributable to more global associations with head size. In addition, genome-wide association analyses of intracranial volume and brain volume were conducted in the healthy controls to clarify whether observed associations were specific to hippocampal volume or influenced brain size in general. Participating sites were asked to conduct five genome-wide association analyses (three analyses of hippocampal volume, intracranial volume and brain volume). In addition, cohorts with groups of individuals with disease were asked to perform hippocampal analyses including data from these individuals.

Evidence for association was assessed using the allelic dosage of each SNP (accounting for familial relationships in the GOBS, QTIM and SYS samples). SNP-derived covariates were tested as fixed effects, while explicitly modeling the genetic relationships between family members in these pedigree-based studies[51,61,62]. Analyses used multiple linear regression with the phenotype of interest as a dependent variable and the additive dosage of each SNP as an independent variable of interest, controlling for covariates of population stratification (four MDS components), age, age^2, sex and the interactions between age and sex and age^2 and sex. Dummy covariates were used to control for different scanner sequences or equipment within a site when needed. We refer to these covariates as 'other covariates', and these were included in all analyses. The extensive regression model was used to statistically control for factors known to affect hippocampal volume that are not specific genetic influences. Recommended protocols for association were provided to the studies based on those used in mach2qtl software (see ENIGMA protocols).

To combine information from multiple studies, we generated a secure web-accessible upload site for participants to upload their association results. An automated system parsed the uploaded results files (see URLs). This parser was designed to read raw results files from a variety of association software packages (mach2qtl, PLINK, SOLAR, SNPTEST, QUICKTEST, Merlin-offline and ProbABEL), perform a series of tests on the incoming data to ensure quality, correctly assign the effect allele (dependent on both the imputation and association programs used) and correctly scale the β values and standard errors from association into the same units. Quality control was performed on imputed SNPs to filter out SNPs with low frequency (MAF of <0.01) or poor imputation quality (estimated R^2 of <0.3). Result files and summary statistics from each group were pooled for meta-analysis. Meta-analysis was undertaken for each SNP across all groups based on a fixed-effects model using an inverse standard error–weighted meta-analysis protocol implemented in METAL[19]. Genomic control was applied at the level of each study and at the meta-analysis level to adjust for population stratification or cryptic relatedness not accounted for by MDS components[63]. To account for heterogeneity across samples, a random-effects meta-analysis[20] was also conducted via the program METASOFT without using genomic control. Using KGG[64] we performed gene-based tests on the double genome-controlled meta-analysis results, using the extended Simes test[24] to obtain an overall P value for association of the entire gene with a 50-kb boundary on either side. Results from genes with $P \leq 1 \times 10^{-4}$ are presented (**Supplementary Tables 18–25**).

Meta-analysis was performed separately on the discovery sample, the CEU and TSI replication sample and the CEU and YRI or MEX replication sample. These results were then pooled to form the combined meta-analysis statistics for discovery and replication. The *in silico* replication results from the CHARGE Consortium were added to this, and a final meta-analysis was conducted. The location of Manhattan and quantile-quantile plots is specified in **Supplementary Table 27.**

doi:10.1038/ng.2250

To appropriately account for the multiple comparisons conducted, we first sought to determine the effective number of independent phenotypes among the eight highly correlated genome-wide association analyses. This was calculated by creating an 8×8 matrix derived from cross-correlations of meta-analytic *t* statistics of association for each SNP across phenotypes. The resulting correlation matrix provided an estimate of the similarity between phenotypes after adjusting for covariates of interest and appropriately controlling for family structure. The effective number of tests was then calculated by summing eigenvalues of the correlation matrix, weighted by a formula that appropriately controls false positive rates in simulation[65]. The effective number of tests was determined to be 4 and an overall genome-wide significance threshold of $5 \times 10^{-8}/4 = 1.25 \times 10^{-8}$ was used throughout the manuscript.

Regulatory potential of SNPs identified in the genome-wide association analysis was examined in three samples. In the UCL epilepsy cohort, tissue was obtained from resection material from affected individuals who had undergone surgery for drug-resistant mesial temporal lobe epilepsy with hippocampal sclerosis, according to established clinical protocols. Total RNA from the middle temporal cortex (Brodmann areas 20 and 21) from 86 subjects was isolated and randomly hybridized to Affymetrix Human Exon 1.0 ST arrays, and quality control analysis was performed using standard methods. The effects of several methodological (day of expression hybridization, RNA integrity number (RIN)) and biological covariates (sex, age and medication) on exon–gene expression relationships were tested for significance. Of these individuals, 71 had participated in a published epilepsy genome-wide association study, and, therefore, genotyping data were available. Details of sample collection and genotyping quality control steps have been published previously[66]. These samples were assayed with Illumina HumanHap550v3 ($N = 44$) and Illumina Human610-Quadv1 ($N = 27$) arrays.

In the UK Brain Expression database, post-mortem brain tissues from 134 individuals free from neurological disorders were obtained from the MRC Sudden Death Brain Bank in Edinburgh and Sun Health Research Institute[67]. Genotype information was obtained using Illumina HumanOmni 1M arrays and standard quality control methods. Expression profiling was conducted in up to ten separate brain regions for each individual brain using the Affymetrix GeneChip Human Exon 1.0 ST array. Expression levels were normalized using the Robust Multi-array Analysis (RMA) algorithm restricting to probe sets containing more than three probes, unique hybridization target (probes that map to a single position within the genome) and supported by evidence from EntrezGene. The average signals for all neocortex (AvgCTX) and all brain regions (AvgAll) were tested, as were individual cortical and subcortical regions. Any significant association where the probe set contained the SNP or a SNP in high LD ($r^2 > 0.50$) were removed from further analysis.

SNPExpress, a publically available database, was also used for replication of the findings. The SNPExpress database[68] used autopsy-collected frontal cortex brain tissue in 93 samples from human subjects with no neuropsychiatric conditions and PBMCs in 80 samples. In this database, transcript expression levels were measured on Affymetrix Human ST 1.0 exon arrays, and genome-wide genotyping was performed using Illumina HumanHap550K arrays.

Raw gene expression data from human fetal brain were gathered from a published study[42]. Post-mortem specimens from four late mid-fetal human brains (18, 19, 21 and 23 weeks of gestation) were collected from the Human Fetal Tissue Repository at the Albert Einstein College of Medicine. Details of specimens, tissue processing, microdissection and neuropathological assessment have been described elsewhere[42].

43. Smith, S.M. *et al.* Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* **23** (suppl 1), S208–S219 (2004).
44. Zhang, Y., Brady, M. & Smith, S. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Trans. Med. Imaging* **20**, 45–57 (2001).
45. Jenkinson, M., Bannister, P., Brady, M. & Smith, S. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage* **17**, 825–841 (2002).
46. Morey, R.A. *et al.* A comparison of automated segmentation and manual tracing for quantifying hippocampal and amygdala volumes. *Neuroimage* **45**, 855–866 (2009).
47. Pardoe, H.R., Pell, G.S., Abbott, D.F. & Jackson, G.D. Hippocampal volume assessment in temporal lobe epilepsy: how good is automated segmentation? *Epilepsia* **50**, 2586–2592 (2009).
48. Morey, R.A. *et al.* Scan-rescan reliability of subcortical brain volumes derived from automated segmentation. *Hum. Brain Mapp.* **31**, 1751–1762 (2010).
49. Pantel, J. *et al.* A new method for the *in vivo* volumetric measurement of the human hippocampus with high neuroanatomical accuracy. *Hippocampus* **10**, 752–758 (2000).
50. Morra, J.H. *et al.* Validation of a fully automated 3D hippocampal segmentation method using subjects with Alzheimer's disease mild cognitive impairment, and elderly controls. *Neuroimage* **43**, 59–68 (2008).
51. Almasy, L. & Blangero, J. Multipoint quantitative-trait linkage analysis in general pedigrees. *Am. J. Hum. Genet.* **62**, 1198–1211 (1998).
52. Purcell, S., Cherny, S.S. & Sham, P.C. Genetic Power Calculator: design of linkage and association genetic mapping studies of complex traits. *Bioinformatics* **19**, 149–150 (2003).
53. Pei, Y.F., Li, J., Zhang, L., Papasian, C.J. & Deng, H.W. Analyses and comparison of accuracy of different genotype imputation methods. *PLoS ONE* **3**, e3551 (2008).
54. Marchini, J. & Howie, B. Genotype imputation for genome-wide association studies. *Nat. Rev. Genet.* **11**, 499–511 (2010).
55. Allen, A.S., Martin, E.R., Qin, X. & Li, Y.J. Genetic association tests based on ranks (GATOR) for quantitative traits with and without censoring. *Genet. Epidemiol.* **30**, 248–258 (2006).
56. Marchini, J., Howie, B., Myers, S., McVean, G. & Donnelly, P. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat. Genet.* **39**, 906–913 (2007).
57. Browning, B.L. & Browning, S.R. A unified approach to genotype imputation and haplotype-phase inference for large data sets of trios and unrelated individuals. *Am. J. Hum. Genet.* **84**, 210–223 (2009).
58. Durston, S. *et al.* Differential effects of DRD4 and DAT1 genotype on fronto-striatal gray matter volumes in a sample of subjects with attention deficit hyperactivity disorder, their unaffected siblings, and controls. *Mol. Psychiatry* **10**, 678–685 (2005).
59. Dick, D.M. *et al.* Genome-wide association study of conduct disorder symptomatology. *Mol. Psychiatry* **16**, 800–808 (2011).
60. Stein, J.L. *et al.* Discovery and replication of dopamine-related gene effects on caudate volume in young and elderly populations (*N*=1198) using genome-wide search. *Mol. Psychiatry* **16**, 927–937, 881 (2011).
61. Chen, W.M. & Abecasis, G.R. Family-based association tests for genomewide association scans. *Am. J. Hum. Genet.* **81**, 913–926 (2007).
62. Aulchenko, Y.S., Struchalin, M.V. & van Duijn, C.M. ProbABEL package for genome-wide association analysis of imputed data. *BMC Bioinformatics* **11**, 134 (2010).
63. Devlin, B. & Roeder, K. Genomic control for association studies. *Biometrics* **55**, 997–1004 (1999).
64. Li, M.X., Sham, P.C., Cherny, S.S. & Song, Y.Q. A knowledge-based weighting framework to boost the power of genome-wide association studies. *PLoS ONE* **5**, e14480 (2010).
65. Li, J. & Ji, L. Adjusting multiple testing in multilocus analyses using the eigenvalues of a correlation matrix. *Heredity* **95**, 221–227 (2005).
66. Kasperaviciūte, D. *et al.* Common genetic variation and susceptibility to partial epilepsies: a genome-wide association study. *Brain* **133**, 2136–2147 (2010).
67. Trabzuni, D. *et al.* Quality control parameters on a large dataset of regionally-dissected human control brains for whole genome expression studies. *J. Neurochem.* **119**, 275–282 (2011).
68. Heinzen, E.L. *et al.* Tissue-specific genetic control of splicing: implications for the study of complex traits. *PLoS Biol.* **6**, e1 (2008).

doi:10.1038/ng.2250

CHAPTER 6

Future works

From the work presented in this dissertation, there is a clear trend toward collaborative science especially in decentralized consortia that protect the investments of individual investigators by allowing participation without requiring full access to datasets while still allowing for the consortia to tackle problems and hypotheses that would be difficult to convincingly address alone. Further, I think we will see a renewed push to apply collaborative efforts to the study of disease and disease biomarkers.

6.1 GWAS meta-analysis of subcortical brain volumes, ENIGMA2

Neuroimaging genetics has the potential to identify genetic contributions to disease pathology by discovering both common and rare genetic variants that relate to brain structure and function. Many studies have identified significant genetic associations with brain measures, but effect sizes are generally small so vast samples are needed to find and replicate genetic associations (Stein et al., 2012). In addition to small effect sizes, data collection is expensive and small datasets are common. To maximize power to detect genetic effects on brain measures, we formed the ENIGMA Consortium (http://enigma.loni.ucla.edu/) to help coordinate and harmonize neuroimaging genetics efforts at sites around the world.

Our pilot project, studying genetic influences on hippocampal volume, intracranial volume, and total brain volume found several genome-wide significant SNPs

(Stein et al., 2012). Expanding the scope of the initial project, we are now examining genetic influences on the volumes of subcortical structures: caudate, putamen, pallidum, thalamus, nucleus accumbens, amygdala, hippocampus and intracranial volume (ICV).

The ENIGMA Consortium is comprised of 28 groups that span 5 continents, including 16,125 subjects. ENIGMA follows a meta-analysis framework, where analyses are conducted at local sites and group-level, de-identified statistics are contributed for meta-analysis. To harmonize analyses across sites, we developed standardized protocols for image analysis, imputation of genetic data, and genetic association analysis (http://enigma.loni.ucla.edu/protocols/).

Image analysis was conducted using fully-automated and validated neuroimaging segmentation algorithms (either FSL FIRST or FreeSurfer). The heritability of each structure was estimated using structural equation modeling (SEM) from 801 twins and siblings from the QTIM study. Genetic data were imputed to the latest 1000 Genomes reference panel (phase I, version 3) using MaCH and minimac (http://www.sph.umich.edu/csg/abecasis/MACH/download/). The imputation results were cleaned to remove poorly imputed SNPs (Rsq < 0.3) and SNPs with low minor allele frequency (MAF < 0.005). Association testing was conducted using mach2qtl for samples with only unrelated subjects and merlin-offline (http://www.sph.umich.edu/csg/abecasis/Merlin/) for samples with a family design. Association tests conducted at each SNP controlled for age, sex, 4 MDS components, intracranial volume (ICV), site variables (for multi-site studies) and disease status (if applicable). Association was conducted separately for each brain structure phenotype. Data were combined across sites using the inverse variance-weighted meta-analysis.

Inverse variance-weighted meta-analysis combines effect sizes (regression coefficients) across sites in a way that penalizes SNPs with an inconsistent direction of effect. In this way, we minimize the chance of finding false-positive results while boosting power to detect even small genetic effects across sites. In future, the ENIGMA Consortium will examine genetic influences on cortical surface phenotypes, 3D morphometry, white-matter integrity, and many other brain-derived measures. In addition, we are examining disease-specific hypotheses in bipolar disorder, schizophrenia, and major depressive disorder. New groups with neuroimaging data are encouraged to join in or propose new projects; many ongoing projects do not require genotyped cohorts.

6.2 Meta-analysis of structural brain differences in bipolar disorder: the ENIGMA-Bipolar project

Neuroimaging analyses of bipolar disorder (BD) have reported significant structural differences in volumes of subcortical brain structures including the amygdala (DeBello 2004; Strakowski 1999), hippocampus, thalamus, accumbens and lateral ventricles (Rimol 2010). A literature search of subcortical structures indicated that the lateral ventricle was the only structure consistently reported across studies to be significantly different between patients with BD and controls (Kempton 2008). More recently, a large international multisite mega-analysis of a large cohort (321 cases and 442 controls) found that the right lateral ventricle, left temporal lobe, and right putamen differed in volume between BD patients and controls (Hallahan 2011). The inconsistency across sites and

studies, and the relative small number of subjects ascertained per site makes a combined meta- analytic approach advantageous.

Here we developed a standardized image analysis and consensus quality control protocol at a large number of international sites participating in the Enhancing Neuro Imaging Genetics Through Meta- Analysis (ENIGMA) Consortium. We formed the ENIGMA-Bipolar Disorder Working Group to coordinate our efforts across multiple international laboratories interested in bipolar disorder. In our pilot project, presented here, we examine patient v. control effect sizes for subcortical volume differences, in the largest neuroimaging study of BD to date.

Current members of the ENIGMA Bipolar Working Group are listed in **Table 6.1** and we are actively recruiting new members to join future analyses. High-resolution structural MR images were processed with fully automated, validated segmentation software (FSL FIRST or FreeSurfer) to extract a set of subcortical structures: caudate, hippocampus, putamen, pallidum, thalamus, nucleus accumbens, amygdala, lateral ventricles, as well as intracranial volume (ICV). We obtained Cohen's d effect size estimates for the left, right, and average volume of each subcortical structure. Effect size estimates were calculated per site and combined using an inverse variance-weighted meta-analysis framework.

Table 6.1. Demographic breakdown of the number of patients diagnosed with bipolar disorder and healthy controls contributed by each site.

	# of cases	# of controls	total n
Cardiff	60	43	103
Pittsburgh	72	32	104
CLING	35	321	356
HMS	41	55	96
Oslo-Malt	44	44	88
NUIG	60	60	120
TOP	193	305	498
UCLA-Families	153	374*	527
Combined	**658**	**860**	**1892**

The heterogeneity of results across previous BD studies has created the need for an international effort to examine the effects of BD on the brain through meta-analysis, while recognizing that cohorts differ in symptom profiles, duration of illness and medication exposure. The ENIGMA framework allows new sites to participate in analyses without requiring them to share raw scan data, just pre-agreed summary statistics for meta-analysis. In addition, ENIGMA provides standardized protocols to process, quality check, and analyze data. In future, the ENIGMA-Bipolar Disorder

Project will examine cortical phenotypes, 3D surface morphometry, and genetic influence on brain-derived measures, and we welcome additional projects.

6.3 A prospective meta-analysis of subcortical brain volumes in schizophrenia: the ENIGMA-Schizophrenia project

Schizophrenia patients show significant subcortical brain abnormalities but there is considerable heterogeneity of findings across studies. Retrospective meta-analyses of published results are limited because image segmentation and analysis methods differ across the studies reviewed. Here we present a coordinated, large-scale meta-analysis that applies consistent quality assurance (QA) metrics and statistical models across independent datasets. Using the methods developed by the Enhancing Neuro Imaging Genetics Through Meta-Analysis (ENIGMA) Consortium, participating researchers provide their analysis results for prospective aggregation and meta-analysis.

The ENIGMA Schizophrenia project currently comprises datasets from sites listed in **Table 6.2** and encourages other sites to join ongoing collaborative analysis efforts. At each site, total intracranial and subcortical volumes (for pallidum, hippocampus, putamen, lateral ventricle, amygdala, caudate, thalamus, nucleus accumbens) were extracted using FreeSurfer from high-resolution structural brain MRI scans of schizophrenia patients and healthy volunteers of similar mean age and sex distribution. The analyses of each cohort included age, sex, and total intracranial volume as covariates, and dummy variables for site effects in multi-site datasets. For each

subcortical region, we computed Cohen's d effect sizes within each study, and weighted mean effect sizes for patient-control group differences across all studies.

Table 6.2. Demographic breakdown for the number of patients diagnosed with schizophrenia and healthy controls contributed by each site. In addition, mean age and gender for each group is given. For some groups duration of illness (DOI) was also available.

Datasets	N (HC)	N (SZ)	M/F (HV)	M/F (SZ)	Mean Age	Mean DOI
TOP	305	219	160/145	130/89	34	7
AMC	170	206	114/56	180/26	23	2
FBIRN	177	186	127/51	141/46	38	17
UMCU	116	159	72/44	130/29	32	11
MCIC	165	158	103/62	119/39	32	11
NU	93	108	52/41	74/34	33	13
CLING	323	49	132/191	36/13	26	
HMS	55	46	21/34	32/14	34	
Total Analyzed	1401	1131	781/624	842/290	31	

Combining data using harmonized methods from many large cohorts through consortia such as ENIGMA can provide robust effect size estimations. This helps focus on the most consistent neuroimaging measures for genetic analysis, for example; ongoing analyses are exploring potential causes of variations in effect sizes across studies. Such meta-analyses may be particularly useful in research areas where study samples are traditionally small (e.g., high-risk, first-episode, and medication studies of schizophrenia). Future work will explore clinical and cognitive factors that influence disease effects within the consortium infrastructure. These more granular analyses can be addressed in subsets of the overall samples with relevant measures in common. The

ENIGMA- Schizophrenia group actively encourages other participating research groups to contribute their analyses and collaborative efforts, as ongoing meta-analyses are developed and expanded.

BIBLIOGRAPHY

DelBello, Melissa P., et al. "Magnetic resonance imaging analysis of amygdala and other subcortical brain regions in adolescents with bipolar disorder." Bipolar disorders 6.1 (2004): 43-52.

Gottesman, I. I., & Gould, T. D. "The endophenotype concept in psychiatry: etymology and strategic intentions." *The American Journal of Psychiatry* 160.4 (2003): 636–645.

Gutman, B. A., et al. "Maximizing power to track Alzheimer's disease and MCI progression by LDA-based weighting of longitudinal ventricular surface features." *NeuroImage* 70 (2013): 386–401.

Hallahan, Brian, et al. "Structural magnetic resonance imaging in bipolar disorder: an international collaborative mega-analysis of individual adult patient data." Biological psychiatry 69.4 (2011): 326- 335.

Hibar D. P., et al. "Voxelwise genome-wide association of Diffusion Tensor Images identifies putative novel variants influencing white matter integrity in 467 related young adults." *Proc. Society for Neuroscience*, San Diego, CA (Nov. 2010).

Hoerl, A. E. "Application of ridge analysis to regression problems." *Chemical Engineering Progress* 58 (1962): 54-59.

Kempton, Matthew J., et al. "Meta-analysis, database, and meta-regression of 98 structural imaging studies in bipolar disorder." Archives of general psychiatry 65.9 (2008): 1017.

Liu, J. Z., et al. “A versatile gene-based test for genome-wide association studies.” *Am. J. Hum. Genet.* 87 (2010): 139–145.

Marchini, J., Donnelly, P., & Cardon, L. R. "Genome-wide strategies for detecting multiple loci that influence complex diseases." *Nature genetics* 37.4 (2005): 413-417.

Pezawas, L., et al. "Evidence of biologic epistasis between BDNF and SLC6A4 and implications for depression." *Molecular psychiatry* 13.7 (2008): 709-716.

Rimol, Lars M., et al. "Cortical thickness and subcortical volumes in schizophrenia and bipolar disorder." Biological psychiatry 68.1 (2010): 41-50.

Stein, J. L., et al. "Voxelwise genome-wide association study (vGWAS)." *Neuroimage* 53.3 (2010): 1160.

Styner, M., et al. "Boundary and medial shape analysis of the hippocampus in schizophrenia." *Medical Image Analysis* 8.3 (2004): 197-203.

Tan, H-Y., et al. "Epistasis between catechol-O-methyltransferase and type II metabotropic glutamate receptor 3 genes on working memory brain function." *Proceedings of the National Academy of Sciences* 104.30 (2007): 12536-12541.

Strakowski, Stephen M., et al. "Brain magnetic resonance imaging of structural abnormalities in bipolar disorder." Archives of General Psychiatry 56.3 (1999): 254.

Tibshirani, R. “Regression Shrinkage and Selection via the Lasso.” J R Stat Soc B 58 (1996): 267-288.

Ueki, M., & Tamiya, G. "Ultrahigh-dimensional variable selection method for whole-genome gene-gene interaction analysis." *BMC bioinformatics*13.1 (2012): 72.

Wang, K., & Abbott, D. “A principal components regression approach to multilocus genetic association studies.” *Genet Epidemiol* 32 (2008): 108-118.

Wang, Y., et al. "Evidence of Epistasis Between the Catechol-< i> O</i>-Methyltransferase and Aldehyde Dehydrogenase 3B1 Genes in Paranoid Schizophrenia." *Biological psychiatry* 65.12 (2009): 1048-1054.

Wold, S., et al. "The collinearity problem in linear regression. The partial least squares (PLS) approach to generalized inverses." *SIAM Journal on Scientific and Statistical Computing* 5.3 (1984): 735-743.

Wray, N.R., et al. "Multi-locus models of genetic risk of disease." *Genome Med* 2.10 (2010).

Zou, H., & Hastie, T. “Regularization and variable selection via the elastic net.” *J R Stat Soc B* 67 (2005): 301–320.