Common variants at 12q14 and 12q24 are associated with hippocampal volume

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Published Version</td>
<td>doi:10.1038/ng.2237</td>
</tr>
<tr>
<td>Citable link</td>
<td><a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:10579383">http://nrs.harvard.edu/urn-3:HUL.InstRepos:10579383</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA">http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA</a></td>
</tr>
</tbody>
</table>
Common variants at 12q14 and 12q24 are associated with hippocampal volume


Joshua C. Bis, Charles DeCarli, Albert Vernon Smith, Fedde van der Lijn, Fabrice Crivello, Myriam Fornage and Stephanie Debette contributed equally as first authors.
Cornelia M. van Duijn, Thomas H. Mosley, Reinhold Schmidt, Christophe Tzourio, Lenore J. Launer, M. Arfan Ikram and Sudha Seshadri contributed equally as last authors.
Jason L. Stein, Sarah E. Medland, Alejandro Arias Vasquez, Derrek P. Hibar, Margaret J. Wright, Barbara Franke, Nicholas G. Martin and Paul M. Thompson represent the ENIGMA consortium.

Author Contributions:
Study concept and design:

Acquisition of the data:

Statistical analysis and interpretation of the data:

Drafting the manuscript:

Critical revision of the manuscript:

Obtained funding:

Competing interest statement
The authors declare no competing financial interests related to this manuscript.
Bis et al.

Reinhold Schmidt\textsuperscript{36}, Christophe Tzourio\textsuperscript{1,70,71}, Lenore J. Launer\textsuperscript{51}, M. Arfan Ikram\textsuperscript{6,24,42}, and Sudha Seshadri\textsuperscript{25,26} for the CHARGE Consortium

\textsuperscript{1}Cardiovascular Health Research Unit, Department of Medicine, University of Washington, Seattle, WA, USA \textsuperscript{2}Department of Neurology and Center of Neuroscience, University of California at Davis, Sacramento, CA, USA \textsuperscript{3}Icelandic Heart Association, Kopavogur, Iceland \textsuperscript{4}University of Iceland, Faculty of Medicine, Reykjavik, Iceland \textsuperscript{5}Department of Medical Informatics, Erasmus MC University Medical Center, Rotterdam, The Netherlands \textsuperscript{6}Department of Radiology, Erasmus MC University Medical Center, Rotterdam, The Netherlands \textsuperscript{7}Univ. Bordeaux, Neurofunctional Imaging Group, UMR 5296, F-33000 Bordeaux, France \textsuperscript{8}Centre National de la Recherche Scientifique (CNRS), Neurofunctional Imaging Group, UMR 5296, F-33000 Bordeaux, France \textsuperscript{9}Commissariat à l’Energie Atomique (CEA), Neurofunctional Imaging Group, UMR 5296, F-33000 Bordeaux, France \textsuperscript{10}Brown Foundation Institute of Molecular Medicine, The University of Texas Health Sciences Center at Houston, Houston, TX, USA \textsuperscript{11}Human Genetics Center School of Public Health, The University of Texas Health Sciences Center at Houston, Houston, TX, USA \textsuperscript{12}INSERM, U708, Neuroepidemiology, F-75013, Paris, France \textsuperscript{13}Department of Epidemiology, University of Versailles St-Quentin-en-Yvelines, France \textsuperscript{14}Department of Neurology, Boston University School of Medicine, Boston, MA, USA \textsuperscript{15}Program in Translational NeuroPsychiatric Genomics, Institute for the Neurosciences, Department of Neurology, Brigham and Women’s Hospital, Boston, MA, USA \textsuperscript{16}Program in Medical and Population Genetics, Broad Institute, Cambridge, MA, USA \textsuperscript{17}Institute of Molecular Biology and Biochemistry, Medical University Graz, Austria \textsuperscript{18}Stroke and Ageing Research Centre, Southern Clinical School, Department of Medicine, Monash University, Melbourne, Australia \textsuperscript{19}Menzies Research Institute Tasmania, University of Tasmania, Hobart, Australia \textsuperscript{20}Genetic Epidemiology Unit, Department of Epidemiology, Erasmus MC University Medical Center, Rotterdam, The Netherlands \textsuperscript{21}Department of Neurology, Erasmus MC University Medical Center, Rotterdam, The Netherlands \textsuperscript{22}Rush Alzheimer’s Disease Center, Rush University Medical Center, Chicago, IL, USA \textsuperscript{23}Department of Biostatistics, Boston University School of Public Health, Boston, MA, USA \textsuperscript{24}Department of Epidemiology, Erasmus MC University Medical Center, Rotterdam, The Netherlands \textsuperscript{25}Department of Neurology, Boston University School of Medicine, Boston, MA, USA \textsuperscript{26}The National, Heart, Lung and Blood Institute’s Framingham Heart Study, Framingham, MA, USA \textsuperscript{27}INSERM Unit 744, Lille, France \textsuperscript{28}Institut Pasteur de Lille, Lille, France \textsuperscript{29}Université Lille Nord de France, Lille, France \textsuperscript{30}Department of Radiology, Mayo Clinic, Rochester, MN, USA \textsuperscript{31}Department of Neurological Sciences, Rush University Medical Center, Chicago, IL, USA \textsuperscript{32}Department of Behavioral Sciences, Rush University Medical Center, Chicago, IL, USA \textsuperscript{33}Biospective Inc., Montreal, Canada \textsuperscript{34}Department of Neurology, Sint Franciscus Gasthuis, Rotterdam, the Netherlands \textsuperscript{35}Division of Public Health Sciences, Wake Forest School of Medicine, Winston-Salem, NC, USA \textsuperscript{36}Department of Neurology, Medical University Graz, Austria \textsuperscript{37}University of Queensland, Diamantina Institute, Princess Alexandra Hospital, Brisbane, Australia \textsuperscript{38}Department of Biomedical Engineering, Illinois Institute of Technology, Chicago, IL, USA \textsuperscript{39}Department of Radiology, Leiden University Medical Center, Leiden, the Netherlands \textsuperscript{40}Department of Neurology, University of Mississippi Medical Center, Jackson, MS, USA \textsuperscript{41}Faculty of Applied Sciences, Delft University of Technology, Delft, the Netherlands \textsuperscript{42}Netherlands Consortium for Healthy Ageing, Leiden, the Netherlands \textsuperscript{43}Centre Hospitalier Régional Universitaire de Lille, Lille, France \textsuperscript{44}Department of Psychiatry and Human Behavior, University of Mississippi Medical Center, Jackson, MS, USA \textsuperscript{45}Department of Clinical Genetics, Erasmus MC University Medical Center, Rotterdam, The Netherlands \textsuperscript{46}Laboratory of Neuro Imaging, David Geffen School of Medicine, UCLA, Los Angeles, CA, USA \textsuperscript{47}Genetic Epidemiology Laboratory, Queensland Institute of Medical Research, Brisbane, Australia \textsuperscript{48}Quantitative Genetics, Queensland Institute of Medical Research, Brisbane, Australia \textsuperscript{49}Broad Institute of Harvard and MIT, Cambridge, MA, USA \textsuperscript{50}Department of Human Genetics, Radboud University Nijmegen Medical Centre, Nijmegen, The Netherlands \textsuperscript{51}Department of Psychiatry, Donders Institute for Brain, Cognition and

Nat Genet. Author manuscript; available in PMC 2012 November 01.
Abstract

Aging is associated with reductions in hippocampal volume (HV) that are accelerated by Alzheimer’s disease and vascular risk factors. Our genome-wide association study of dementia-free persons (n=9,232) identified 46 SNPs at four loci with p-values <4.0×10^{-7}. Two additional samples (n=2,318) replicated associations at 12q24 within MSRB3/WIF1 (discovery + replication, rs17178006; p=5.3×10^{-11}) and at 12q14 near HRK/FBXW8 (rs7294919; p=2.9×10^{-11}). Remaining associations included one 2q24 SNP within DPP4 (rs6741949; p=2.9×10^{-7}) and nine 9p33 SNPs within ASTN2 (rs7852872; p=1.0×10^{-7}) that were also associated with HV (p<0.05) in a third younger, more heterogeneous sample (n=7,794). The ASTN2 SNP was also associated with decline in cognition in a largely independent sample (n=1,563). These associations implicate genes related to apoptosis (HRK), development (WIF1), oxidative stress (MSR3B), ubiquitination (FBXW8), enzymes targeted by new diabetes medications (DPP4), and neuronal migration (ASTN2), indicating novel genetic influences that influence hippocampal size and possibly the risk of cognitive decline and dementia.

Differences in hippocampal volume (HV) that appear with advancing age represent cumulative effects of early life factors, life-course events, and disease. Hippocampal atrophy is a recognized biological marker of Alzheimer’s disease (AD)\textsuperscript{1,2}; however, it is influenced by various vascular and metabolic factors\textsuperscript{3,4}. Because HV is a heritable\textsuperscript{5} widely measurable trait that exhibits meaningful detectable changes throughout the lifespan, it is a suitable endophenotype for aging-related physiological processes and presymptomatic diseases, improving power to detect genetic associations.

We explored genetic influences on HV by conducting a cross-sectional genome-wide association analysis in the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) Consortium\textsuperscript{6} among 9,232 dementia-free persons from eight community-based studies whose mean age ranged from 56 to 84 years (weighted average, 67.1 years). Each study imputed to a common set of Phase II HapMap CEU SNPs using genotype data from...
Illumina or Affymetrix arrays; fit additive genetic models associating total hippocampal volume and genotype dosage with adjustment for age, sex, and familial relationships (if applicable, see Supplementary Note); and applied genomic control. Study-specific results were combined in an inverse-variance weighted meta-analysis.

We then conducted in-silico replication of our genome-wide significant associations and sought additional evidence for suggestive associations in a second-stage targeted meta-analysis of 2,318 subjects from two community-based studies: the Three City study and an independent sample from the third expansion of the Rotterdam study. Characteristics of the discovery and replication samples are shown in Supplementary Table 1.

Figure 1 provides a Manhattan plot of -log_{10}(p-values) from the discovery analysis, where p-values for 46 SNPs at four loci (Supplementary Table 2) surpassed our replication threshold of \( p<4.0 \times 10^{-7} \) —corresponding to one expected false positive. Of these, 18 SNPs at 2 loci surpassed a genome-wide significant threshold of \( p<5.0 \times 10^{-8} \): 12q14, which included WiFi, LEMD3, and MSRB3, and 12q24, which included HRK and FBXW8. We found evidence of replication (\( p<0.01 \)) for both associations. The remaining suggestive associations included SNPs on 2q24 within DPP4 and on 9p33 within ASTN2, which had consistent directions of association in the replication phase, but did not attain genome-wide significance in a combined analysis. Estimates for each stage are shown in Table 1; discovery GWAS results for each signal’s surrounding region annotated with recombination rates and known genes are shown in Figure 2; and study-specific findings appear in Figure 3.

Below, we present association estimates for a meta-analysis combining the discovery and replication results for these four loci. To contextualize the magnitude of a SNP’s association with HV, we divided the regression coefficient for the allele by the mean decrease in HV for each year of chronological age (\(-27.4 \text{ mm}^3\) per year, estimated within the Framingham Heart Study).

The strongest association was for rs7294919, located on 12q24 between HRK and FBXW8, where each copy of the T allele (allele frequency \( [AF] = 0.91 \)) was associated with lower HV (\( \beta = -107.8 \text{ mm}^3, p=2.9 \times 10^{-11} \)), equivalent to 3.9 years of aging.

HRK is expressed throughout the brain with highest levels in the amygdala, entorhinal cortex, and hippocampus. It acts as a key regulator of apoptosis, a complex pathway associated with aging, ischemia, and AD through its interaction with death-repressor proteins Bcl-2 and Bcl-X(L). In rat neuronal cell cultures, a homologous protein, DP5, is induced during A\( \beta \)-mediated cytotoxicity, withdrawal of nerve growth factor (NGF), and induced global ischemia. While treatment strategies aimed at modifying the apoptotic pathway have yet to achieve success, our findings suggest this area of therapeutics might remain promising.

FBXW8 encodes the substrate-recognition component of an SCF (skip1, cullin1, F-box) E3 ubiquitin ligase found in the golgi apparatus of neurons. Different E3 ligase complexes target specific substrates for polyubiquitination leading to proteosome degradation, suggesting a role in clearing abnormal and potentially toxic protein aggregates, particularly hyperphosphorylated tau. Its role in presynaptic development, synapse formation, neurotransmitter release, and promotion of dendrite growth in hippocampal neurons makes a genetic association with HV plausible. Whether one or both HRK and FBXW8 are involved in determining HV is unclear since rs7294919 is an eSNP associated with changes in both.

At the 12q14 locus, the G allele for rs17178006, intronic within MSRB3 (AF=0.10), was associated with decreased HV (\( \beta = -123.8 \text{ mm}^3, p=5.3 \times 10^{-11} \)) equivalent to 4.5 years of...
aging. MSRB3 catalyzes the reduction of methionine R-sulfoxide residues in proteins and requires zinc or selenium as a cofactor. Thus, the association of lower selenium levels with elevated plasma homocysteine, which in turn has been associated with an increased risk of AD and hippocampal atrophy \(^{22-24}\), may be mediated by suppression of MsrB in various organs including the brain\(^{25}\). Several SNPs in low linkage disequilibrium (\(r^2=0.2\)) with rs17178006 were also associated with decreased HV, including rs6581612 (AF=0.27, \(\beta=-63.3\ \text{mm}^3, \ p=7.2\times10^{-11}\) between WIF1 and LEMD3. WIF1 inhibits extracellular signaling Wnt proteins, which play a role in embryonic development—along with \(\beta\)-catenin—and hippocampal aging\(^{26}\). Changes in Wnt signaling mimic the effects of environmental enrichment increasing hippocampal synaptic densities\(^{27}\). LEMD3 is a transforming growth factor-beta antagonist expressed in the hippocampus and upregulated protectively during ischemia and epileptogenesis\(^{28-30}\). Further, it interacts with progerin, the abnormal form of laminin A responsible for premature aging in progeria (Hutchinson-Gilford syndrome)\(^{31}\).

When testing for independent effects of these two SNPs in conditional models, both associations were attenuated, but only rs17178006 remained significant (\(p<0.05\), Supplementary Figure 1), suggesting that the SNPs mark a single locus. Whereas MSRB3 may be most influential, it remains possible that more than one gene in this region is associated with HV. For example, 8 eSNPs in the vicinity of this locus (Supplementary Table 3) were associated with HV at a p-value \(\leq5.3\times10^{-4}\) and have been reported to modify LEMD3 expression.

In addition to the strong findings discussed above, SNPs at two additional loci showed suggestive evidence for association, but did not reach genome-wide significance in our combined meta-analysis. The first was rs6741949 in a DPP4 intron on chromosome 2q24, where the G allele (AF=0.53) was associated with smaller hippocampal volume (\(\beta=-52.8\ \text{mm}^3, \ p=2.9\times10^{-7}\)). Many bioactive peptides whose levels are altered in AD and vascular brain injury are substrates for DPP4\(^{32}\), and DPP4 reduces extracellular \(\beta\)-amyloid deposition in mouse models of AD\(^{33}\). Further, DPP4 is an intrinsic membrane glycoprotein and a widely expressed serine exopeptidase\(^{34}\). It is also an adipokine over expressed in visceral adipose tissue of obese persons and those with diabetes\(^{35}\), conditions associated with smaller HV\(^{3,36}\). A novel class of antidiabetic medications (sitaagliptin, and related incretin compounds) inhibits DPP4 to improve insulin sensitivity and glucose tolerance through increased levels of glucagon like proteins-1 and 2 (GLP-1, -2). Interestingly, endogenous incretin GLP-1 is also heavily expressed in some hippocampal neurons and has neuroprotective properties\(^{37-39}\).

The second suggestive association was for rs7852872, located in an ASTN2 intron on chromosome 9 where the C allele (AF=0.63) was associated with lower hippocampal volume (\(\beta=-47.7\ \text{mm}^3, \ p=1.0\times10^{-7}\)). ASTN2 is a cell adhesion molecule expressed in neurons, including those in the dentate gyrus and hypothesized to function in glial-guided neuronal migration\(^{40, 41}\).

We sought additional replication of our significant and suggestive associations by testing the lead (or proxy) SNP from each locus in the Enhancing Neuro Imaging Genetics through Meta Analysis (ENIGMA) consortium. Briefly, this group (\(n=7,794, \text{mean age} 39.9\ \text{years}\)) combined multiple studies including normal older individuals, a developmental sample, and cases symptomatic for cognitive or affective diseases. Among the ENIGMA sample, we observed a consistent direction of association for all cross-study comparisons. For the loci where the lead SNP was available in ENIGMA, replication was strongest at HRK/FBXW8 (rs7294919, \(p=1.6\times10^{-7}\)) and nominal at DPP4 (rs6741949, \(p=0.04\)). While the lead SNPs were not available at the other loci, one proxy SNP in weak (\(r^2 = 0.3\)) linkage disequilibrium
with rs17178006 (MSRB3) and another in strong LD with rs7852872 (ASTN2) both had p-values < 0.05 in ENIGMA (Table 2).

Because the ENIGMA and CHARGE samples differed in two key aspects—ENIGMA’s inclusion of younger adults (8 of 13 studies had no participants older than 65 years) and of some persons with cognitive impairment and dementia (13% of the sample)—we examined the top loci in subsamples of healthy persons (n=5,775, mean age 34.8 years) and of cognitively intact older persons (n=816, mean age 67.2 years). Association estimates, were generally similar to those of the full sample (Supplementary Table 4).

Given the established relationship between hippocampal atrophy and AD, we investigated whether SNPs from published AD GWAS were associated with HV in our discovery meta-analysis (Supplementary Table 5). We found nominal associations of risk alleles in four AD genes with smaller HV: APOE (p=0.005), BIN1 (p=0.02), MS4A4E (p=0.001) and TOMM40 (p=0.01). However, in aggregate, various known AD SNPs explained less than 1% of the observed variance in HV.

We also examined our five lead SNPs for associations with cognitive decline among 1,593 participants (mean age 78.6 years) in the Religious Orders Study and Rush Memory and Aging Project (described in Supplementary Table 6) and found that rs7852872 (ASTN2 locus) was associated with an accelerated rate of global cognitive decline (p=0.009) and an accelerated rate of memory loss (p=0.01) (Supplementary Figure 2 and Supplementary Table 7). The magnitude of effect was comparable to that noted previously for a CR1 SNP (rs6656401) in the same sample, providing evidence for the potential importance of this region.

The strengths of the current study include the large population-based sample. In the discovery sample, our power to detect genome-wide significant associations on the order of 0.2 standard deviations in HV (~128 mm³, in the largest single sample: the AGES_Reykjavik study) was modest for rare variants (68% for 0.05 MAF) and strong for more common variants (>99% for 0.10 MAF). Additional power estimates are shown in Supplementary Figure 3. The concordance of these associations in ENIGMA provides additional biological validation in a population that included younger persons, suggesting that these genes are developmentally important and may be related to maximal adult HV. ENIGMA also has a substantial proportion of persons with dementia, which indicates that these genes may remain important in regulating response to injury. The ability to explore the association of our lead SNPs with cognitive decline provided further context to our findings. Finally, we demonstrated modest associations between previously described Alzheimer’s disease risk SNPs and smaller HV in our samples.

The study also has limitations. A single cross-sectional assessment was used in all studies and MRI and reading protocols varied across participating studies: some studies used manually-traced boundaries (the gold standard) whereas others used computerized algorithms. Although correlation between these two methods is good (Pearson’s r=0.7), the heterogeneity of measurement techniques may have compromised our ability to detect small associations. Although our sample size was reasonably large, we may have missed associations with small effect sizes as well as rare variants not covered by commercial genotyping arrays.

Prior studies have suggested that cognitive, neuropathological, and MRI endophenotypes of AD might be early and more sensitive markers of genetic risk than clinical dementia. Hence, it could be argued that genes associated with AD risk should also be associated with HV, even in our dementia-free sample. Although four AD genes were associated with HV, several were not; so in this study HV was not a more sensitive measure than clinical AD. It
is clear that genetic analysis of MRI endophenotypes within a healthy older community-based cohort study is not an ideal study design to identify all the genes associated with clinical AD. Our aim, however, was to identify genetic influences on hippocampal development and response to aging and not AD per se.

In summary, we detected four genetic loci associated with HV in a large, population-based, dementia-free sample. Two of these loci replicated in independent community-based samples as well as in ENIGMA, a mixed age sample that included some participants with cognitive impairment indicating that these loci may have broad implications for determining the integrity of the hippocampus across a range of ages and cognitive capacities. Findings from this study identified a series of relevant and potentially important genes associated with HV during development, with aging, and in the presence of Alzheimer’s disease. Exploration of these genomic regions with dense genotyping, expression, and translational studies will be required to understand the role of these genes in determining HV.

**Online Methods**

**Participating studies**

Our analyses were performed among dementia-free participants within the setting of the Cohorts for Heart and Aging Research in Genomic Epidemiology (CHARGE) consortium. The ten discovery samples included the Aging Gene-Environment Susceptibility—Reykjavik Study (AGES), the Atherosclerosis Risk in Communities Study (ARIC), the Austrian Stroke Prevention Study (ASPS), the Erasmus Ruchphen Family study (ERF), the Framingham Heart Study (FHS), the Religious Order Study & Rush Memory and Aging Project (RUSH), three independent phases of the Rotterdam Study (RS I, RS II, RS III), and the Tasmanian Study of Cognition and Gait (TASCOG). The two second stage replication samples included the Three City Study (3C) and another independent sample of the third expansion of the Rotterdam Study (RS R). Details on the discovery samples and second stage studies can be found in the Supplementary Note. Each study has an Institutional Review Board that approved the consent procedures, examination components, data security processes, genotyping protocols and current study design. All participants gave written informed consent for study participation and for use of DNA for genetic research.

**Hippocampal Volume Phenotypes**

Each study evaluated the total hippocampal volume using 1T, 1.5T or 3T MRI and either operator-defined, manually traced boundaries drawn on serial coronal sections or automated methods according to previously described reading protocols. For these analyses, we used data from the baseline examination or the first examination in which an MRI measurement was obtained. Specific details for each study’s MRI protocol are provided in the Supplementary Note.

**Genotyping and imputation**

The studies in these analyses used commercial genotyping platforms available from Illumina or Affymetrix. Each study performed genotyping quality control checks and imputed the approximately 2.5 million polymorphic autosomal SNPs described in the HapMap CEU population for each participant using available imputation methods. Details of per-study genotyping, imputation, and quality control procedures are available in Supplementary Note.

**Statistical analysis within studies**

Each study independently implemented a predefined GWAS analysis plan. For the continuous measure of hippocampal volume, we evaluated cross-sectional associations of hippocampal volume and genetic variation using linear regression models (or linear mixed
effects models, in FHS and ERF to account for family relatedness). For each of the 2.5 million SNPs, each study fit additive genetic models, regressing trait on genotype dosage (0 to 2 copies of the variant allele). In our primary analyses, all studies adjusted for age and sex. Some studies made additional adjustments including study site, familial structure, or for whether the DNA had been whole genome amplified. Additional details of the statistical analyses are available in the Supplementary Note.

Discovery meta-analysis

We conducted a meta-analysis of regression estimates and standard errors using an inverse-variance weighting approach as implemented in METAL\textsuperscript{49}. After verification of strand alignment across studies, QC, filtering, and imputation within each study, we restricted our meta-analysis to autosomal SNPs that were reported in at least 2 studies and that had an average minor allele frequency of at least 1%. Prior to meta-analysis, we calculated a genomic inflation factor ($\lambda_{gc}$) for each study to screen for cryptic population substructure or undiagnosed irregularities that might have inflated the test statistics. Inflation was low, with $\lambda_{gc}$ below 1.05 in all studies. We applied “genomic control” to each study whose genomic inflation factor was greater than 1.00 by multiplying all of the standard errors by the square root of the study-specific $\lambda_{gc}$. We express the association of each SNP and hippocampal volume as the regression slope ($\beta$), its standard error [SE(\beta)] and a corresponding p-value. Standardized gene and SNP annotations were created using a PERL program.\textsuperscript{50} For follow up, we decided a priori on a significance threshold of $p < 4\times10^{-7}$, which corresponds to not more than one expected false positive finding over 2.5 million tests.

Replication meta-analysis

Replication samples were drawn from external studies with available genetic data and measures of hippocampal volume. We provided each collaborating second stage study a list of signal SNPs that attained a p-value of $p < 4\times10^{-7}$ and combined the results from these studies using a fixed-effects meta-analysis as described above.

Combined meta-analysis

Finally, we combined results from the discovery and second stage analyses using inverse variance weighting, as described above, and considered SNPs with a p-value < $5\times10^{-8}$ as genome-wide significant.

External Validation

We sought external replication for our significant and suggestive loci in the ENIGMA consortium, details can be found in the companion paper. The international ENIGMA consortium comprises a wide variety of studies that all have GWAS and hippocampal volume measures (http://enigma.loni.ucla.edu). The sample includes case-control studies of AD and depression, family-based and sib-pair samples as well as population based samples of varying ages and ethnicities (European, African and Hispanic). ENIGMA assesses brain volumes using Freesurfer/FSL-FIRST protocols in most samples but also uses other protocols in a few samples. Hence we chose to compare the results from ENIGMA and CHARGE in a qualitative manner as these two studies vary in the composition of the study sample participants as well as in the methods used to assess HV. We considered replication as a p-value of < 0.05 and consistent direction of association.

Exploration of loci for eQTLs and functional variants

We examined the 4 loci identified as associated with HV for the presence of cis-eQTL associations using the website http://eqtl.uchicago.edu/cgi-bin/gbrowse/eqtl/eQTL. We also
searched for functional SNPs in LD with the 5 index SNPs. We identified over 70 SNPs with an r² > 0.4 that were within 500kb of each index SNP using the SNAP proxy tool (http://www.broadinstitute.org/mpg/snap/) and annotated these SNPs using GeneCruiser (http://genecruiser.broadinstitute.org/genecruiser3/); none of these SNPs were exonic, non-synonymous coding SNPs.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

Aging Gene-Environment Susceptibility-Reykjavik Study: The research has been funded by NIA contract N01-AG-12100 with contributions from NEI, NIDCD and NHLBI, the NIA Intramural Research Program, Hjartavernd (the Icelandic Heart Association), and the Althingi (the Icelandic Parliament).

The Atherosclerosis Risk in Communities Study: The research is carried out as a collaborative study supported by National Heart, Lung, and Blood Institute contracts (HHSN268201100005C, HHSN268201100006C, HHSN268201100007C, HHSN268201100008C, HHSN268201100009C, HHSN268201100010C, HHSN268201100011C, and HHSN268201100012C), R01HL087641, R01HL9367 and R01HL08694 and R01HL7825; National Human Genome Research Institute contract U01HG004402; and National Institutes of Health contract HHSN268200625226C. The authors thank the staff and participants of the ARIC study for their important contributions. Infrastructure was partly supported by Grant Number UL1RR025005, a component of the National Institutes of Health and NIH Roadmap for Medical Research. This project was also supported by grant HL093029 (M.F.).

The Cardiovascular Health Study: co-authors were supported in part by NHLBI grants HL087652 and HL105756 (J.C.B., W.L.) as well as NIA grant AG20098 and AG05133 (J.T.B.).

The Austrian Stroke Prevention Study: The research reported in this article was funded by the Austrian Science Fond (FWF) grant number P20545-P05 and P13180. The Medical University of Graz supports the databank of the ASPS. The authors thank the staff and the participants of the ASPS for their valuable contributions. We thank Birgit Reinhart for her long-term administrative support and Ing Johann Semmler for the technical assistance at creating the DNA-bank.

Erasmus Rucphen Family Study: This study is financially supported by the Netherlands Organization for Scientific Research (NWO), the Internationale Stichting Alzheimer Onderzoek (ISAO), the Hersenstichting Nederland (HSN), and the Centre for Medical Systems Biology (CMSB *1 and 2*) in the framework of the Netherlands Genomics Initiative (NGI). We thank the participants from the Genetic Research in Isolated Populations, Erasmus Rucphen Family, who made this work possible.

Framingham Heart Study: From the Framingham Heart Study of the National Heart Lung and Blood Institute of the National Institutes of Health and Boston University School of Medicine. This work was supported by the National Heart, Lung and Blood Institute’s Framingham Heart Study (Contract No. N01-HC-25195) and its contract with Affymetrix, Inc for genotyping services (Contract No. N02-HL-6-4278). A portion of this research utilized the Linux Cluster for Genetic Analysis (LinGA-II) funded by the Robert Dawson Evans Endowment of the Department of Medicine at Boston University School of Medicine and Boston Medical Center. Analyses reflect intellectual input and resource development from the Framingham Heart Study investigators participating in the SNP Health Association Resource (SHARe) project. This study was also supported by grants from the National Institute of Neurological Disorders and Stroke (NS17950) and the National Institute of Aging (AG08122, AG16495, AG033193, AG031287).

The Religious Order Study (ROS) and Rush Memory and Aging Project (R-MAP): The R-MAP and ROS data used in this article was supported by National Institute on Aging grants P50AG10161, R01AG17917, and R01AG15819, the Illinois Department of Public Health, the Rush Clinical Translational Science Consortium, and a gift from Ms. Marsha Dowd.

The Rotterdam Study: The generation and management of GWAS genotype data for the Rotterdam Study is supported by the Netherlands Organisation of Scientific Research NWO Investments (nr. 175.010.2005.011, 911-03-012). This study is funded by the Research Institute for Diseases in the Elderly (014-93-015; RIDE2), the Netherlands Genomics Initiative (NGI)/Netherlands Organisation for Scientific Research (NWO) project nr. 050-060-810. We thank Pascal Arp, Mila Jhamai, Marijn Verkerk, Lizbeth Herrera and Marjolein Peters for their...
help in creating the GWAS database, and Karol Estrada and Maksim V. Struchalin for their support in creation and analysis of imputed data.

The Rotterdam Study is funded by Erasmus Medical Center and Erasmus University, Rotterdam, Netherlands Organization for the Health Research and Development (ZonMw), the Research Institute for Diseases in the Elderly (RIDE), the Ministry of Education, Culture and Science, the Ministry for Health, Welfare and Sports, the European Commission (DG XII), and the Municipality of Rotterdam. The Rotterdam Scan Study is supported by the Netherlands Organization of Scientific Research (NWO) project nrs. 918-46-615; 904-61-096; 904-61-133; 948-00-010; and Nederlandse Hartstichting 2009B102; and Internationaal Parkinson Fonds. The authors are grateful to the study participants, the staff from the Rotterdam Study and the participating general practitioners and pharmacists.

The Tasmanian Study of Gait and Cognition (TASCOG) is supported by Project Grants from the National Health and Medical Research Council (NHMRC IDs 403000, 491109, 606543), and a grant from the Wicking Dementia Education and Research Centre, Hobart. Velandai Srikanth is supported by an NHMRC/National Heart Foundation Career Development Fellowship (ID 606544). Matthew Brown is supported by an NHMRC Principal Research Fellowship.

Three City Study (3C): We thank the staff and the participants of the 3C Study for their important contributions. The 3C Study is conducted under a partnership agreement between the Institut National de la Santé et de la Recherche Médicale (INSERM), the Victor Segalen–Bordeaux II University, and Sanofi-Aventis. The Fondation pour la Recherche Médicale funded the preparation and initiation of the study. The 3C Study is also supported by the Caisse Nationale Maladie des Travailleurs Salariés, Direction Générale de la Santé, Mutuelle Générale de l’Éducation Nationale (MGEN), Institut de la Longévité, Conseils Régionaux d'Aquitaine et Bourgogne, Fondation de France, and Ministry of Research–INSERM Programme “Cohortes et collections de données biologiques.” Lille Génopôle received an unconditional grant from Eisai. We thank A. Boland (Centre National de Génotypage) for her technical help in preparing the DNA samples for analyses. This work was supported by the National Foundation for Alzheimer’s Disease and Related Disorders, the Institut Pasteur de Lille and the Centre National de Génotypage.

References


Nat Genet. Author manuscript; available in PMC 2012 November 01.
Figure 1. Genome-wide Manhattan plot for hippocampal volume
The plot shows the individual p-values (based on discovery meta-analysis) against their genomic position for hippocampal volume. Within each chromosome, shown on the x-axis, the results are plotted left to right from the p-terminal end. The dashed line indicates the threshold for follow-up, $p<4 \times 10^{-7}$ and the solid line indicates the threshold for genome-wide significance, $p<5 \times 10^{-8}$. The nearest genes are indicated above points that surpassed our significance threshold for follow-up.
Figure 2. Regional plots for hippocampal volume SNPs
Plots are centered on the most significant SNP at a given locus along with the meta-analysis results for SNPs in a region surrounding it (typically ± 100kb). All SNPs are plotted with their discovery meta-analysis p-values against their genomic position, with the most significant SNP in the region indicated as a diamond and other SNPs shaded according to their pairwise correlation ($r^2$) with the signal SNP. The light blue line represents the estimated recombination rates. Gene annotations are shown as dark green lines.
Figure 3. Forest plots for hippocampal volume SNP associations

Plots show the study-specific association estimates ($\beta$) and 95% confidence intervals for the discovery and replication stage studies, presented as rectangles and bars. Arrowheads indicate confidence intervals that span beyond the x-axis. Study specific results are indicated by: AGES, Aging Gene-Environment Susceptibility—Reykjavik Study; ARIC, Atherosclerosis Risk in Communities Study; ASPS, Austrian Stroke Prevention Study; ERF, Erasmus Rucphen Family study; FHS, Framingham Heart Study; RUSH, Religious Order Study & Rush Memory and Aging Project; RS I, RS II, RS III, RS R, independent phases of the Rotterdam Study; TASCOG, Tasmanian Study of Cognition and Gait; 3C, Three City Study. Estimates from the replication phase (3C, RS R) are indicated by open rectangles. The scale is mm$^3$. The association estimate and confidence interval for the meta-analysis combining discovery and second stage results is presented as a diamond. Blank spaces indicate occasions in which a particular study was not able to provide results for a given SNP.
Table 1

<table>
<thead>
<tr>
<th>Locus</th>
<th>SNP</th>
<th>Gene</th>
<th>A1/2</th>
<th>AF</th>
<th>B</th>
<th>SE</th>
<th>P</th>
<th>N_Eff</th>
<th>B</th>
<th>SE</th>
<th>P</th>
<th>N_eff</th>
<th>B</th>
<th>SE</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>2q24</td>
<td>rs6741949</td>
<td>DPP4</td>
<td>G/C</td>
<td>0.53</td>
<td>−61.4</td>
<td>11.3</td>
<td>5.2×10^-8</td>
<td>6.673</td>
<td>−10.1</td>
<td>25.2</td>
<td>0.7</td>
<td>1.369</td>
<td>−52.8</td>
<td>10.3</td>
<td>2.9×10^-7</td>
</tr>
<tr>
<td>9q33</td>
<td>rs7852872</td>
<td>ASTN2</td>
<td>C/G</td>
<td>0.62</td>
<td>−53.1</td>
<td>10.0</td>
<td>1.0×10^-7</td>
<td>9.187</td>
<td>−25.0</td>
<td>20.4</td>
<td>0.2</td>
<td>2.318</td>
<td>−47.7</td>
<td>9.0</td>
<td>1.0×10^-7</td>
</tr>
<tr>
<td>12q14</td>
<td>rs17178006</td>
<td>MSRB3</td>
<td>G/T</td>
<td>0.10</td>
<td>−121.0</td>
<td>20.7</td>
<td>5.5×10^-9</td>
<td>5.249</td>
<td>−137.9</td>
<td>45.5</td>
<td>0.002</td>
<td>1.003</td>
<td>−123.8</td>
<td>18.9</td>
<td>5.3×10^-11</td>
</tr>
<tr>
<td></td>
<td>rs6581612</td>
<td>WIF1</td>
<td>C/A</td>
<td>0.27</td>
<td>−60.5</td>
<td>10.8</td>
<td>2.2×10^-8</td>
<td>9.183</td>
<td>−75.2</td>
<td>22.1</td>
<td>0.0007</td>
<td>2.318</td>
<td>−63.3</td>
<td>9.7</td>
<td>7.1×10^-11</td>
</tr>
<tr>
<td>12q24</td>
<td>rs7294919</td>
<td>HRK</td>
<td>T/C</td>
<td>0.91</td>
<td>−97.7</td>
<td>17.9</td>
<td>4.8×10^-8</td>
<td>8.089</td>
<td>−154.0</td>
<td>38.3</td>
<td>5.8×10^-5</td>
<td>1.573</td>
<td>−107.8</td>
<td>16.2</td>
<td>2.9×10^-11</td>
</tr>
</tbody>
</table>

A1/2 indicates coded (risk)/non-coded allele; B=association estimate, in mm^3; SE=standard error, P=p-value, N_eff indicates effective sample size: Σ(imputation quality×N). **Bolded** gene names indicate that a SNP is within the gene.
Replication results in ENIGMA

<table>
<thead>
<tr>
<th>Locus</th>
<th>SNP</th>
<th>Gene</th>
<th>A1/2</th>
<th>AF</th>
<th>B</th>
<th>SE</th>
<th>P</th>
<th>N_Eff</th>
</tr>
</thead>
<tbody>
<tr>
<td>2q24</td>
<td>rs6741949</td>
<td>DPP4</td>
<td>G/C</td>
<td>0.58</td>
<td>-28.2</td>
<td>14.0</td>
<td>0.04</td>
<td>7,794</td>
</tr>
<tr>
<td>9q33</td>
<td>rs7852872</td>
<td>ASTN2</td>
<td>C/G</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>rs7040792</td>
<td></td>
<td></td>
<td></td>
<td>0.64</td>
<td>12.8</td>
<td>0.02</td>
<td>7,794</td>
</tr>
<tr>
<td>12q14</td>
<td>rs17178006</td>
<td>MSRB3</td>
<td>G/T</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>rs1370938</td>
<td></td>
<td></td>
<td></td>
<td>0.24</td>
<td>14.4</td>
<td>0.02</td>
<td>7,794</td>
</tr>
<tr>
<td></td>
<td>rs6581612</td>
<td>WIFI</td>
<td>C/A</td>
<td></td>
<td>0.25</td>
<td>14.6</td>
<td>0.08</td>
<td>7,794</td>
</tr>
<tr>
<td></td>
<td>rs1498792</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12q24</td>
<td>rs7294819</td>
<td>HRK</td>
<td>T/C</td>
<td>0.90</td>
<td>-112.2</td>
<td>21.4</td>
<td>1.6×10^{-7}</td>
<td>7,794</td>
</tr>
</tbody>
</table>

A1/2 indicates coded (risk)/non-coded allele; B=association estimate, in mm^3; SE=standard error, P=p-value, N_eff indicates effective sample size. Bolded gene names indicate that a SNP is within the gene. Proxy SNPs are indicated in (italics); r^2 indicates correlation in between these proxies and lead SNP in Phase II HapMap CEU sample.