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\(^1\), and is reduced in schizophrenia\(^2\), major depression\(^3\) and mesial temporal lobe epilepsy\(^4\). Whereas many brain imaging phenotypes are highly heritable\(^5\), 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 HMG2A, 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\(^6\) and stress regulation\(^7\) 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\(^8\–\^10\). Hippocampal, total brain and intracranial volumes are highly heritable in non-human primates\(^11\,14\,15\) 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\(^9\) (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
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 schizophrenia22 and bipolar disorder23 (Supplementary Table 8). The most significant SNPs in each analysis from the discovery sample (P ≤ 5 × 10^{-5}) are listed (Supplementary Tables 9–16). No markers reached genome-wide significance (P < 1.25 × 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 tests25).

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 × 10^{-7}) compared to the primary fixed-effects analysis (P = 2.42 × 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 × 10^{-16}; N = 21,151).

rs7294919 lies between HRK and FBXW8 (12q24.2; 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...
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 \(4.8 \times 10^{-5}\) 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\(^{35}\). 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\(^{+}/\text{H}^{+}\) exchanger (NHE1)\(^{36}\), 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).

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

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

\(\text{Freq.}\), 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.

\(\text{Mean bilateral hippocampal volume association results were corrected for intracranial volume, sex, age, age}\^2, \text{sex}\times\text{age, sex}\times\text{age}\^2\text{and four MDS components, and individuals with disease were included in the analysis. 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. Association results for intracranial volume were corrected for sex, age, age}\^2, \text{sex}\times\text{age, sex}\times\text{age}\^2\text{and four MDS components, and individuals with disease were excluded from this analysis. rs10784502 is at 12q14.3: position 64,630,077. Effect allele, C; non-effect allele, T. Intrinsic 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 TDCNUG cohort. Analysis for total brain volume was corrected for sex, age, age}\^2, \text{sex}\times\text{age, sex}\times\text{age}\^2\text{and four MDS components, 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. Rs10494373 is at 1q23.3: position 160,885,986. Effect allele, C; non-effect allele, A. Genomic positions are based on the NCBI36/hg18 (March 2006) genome assembly. Association results for intracranial volume were corrected for sex, age, age}\^2, \text{sex}\times\text{age, sex}\times\text{age}\^2\text{and four MDS components, 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. Rs10494373 is at 1q23.3: position 160,885,986. Effect allele, C; non-effect allele, A. Calculated as 2pq \times p^2 + q^2 \times q^2, where p and q are the minor and major allele frequencies. \(\beta\) 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.**
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 (\( r_E = -0.49; P = 0.016 \)) in this cohort. These results support \textit{HMG2A} as a positional candidate gene underlying our observed QTL. \textit{HMG2A} encodes the high-mobility group AT-hook 2 protein, which is a chromatin-associated protein that regulates stem cell renewal during development\(^4\). It is implicated in human growth through genetic association studies and the presence of rare mutations\(^5\) and also has known roles in neural precursor cells\(^6\). 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\(^7\) (\(N = 1642\)). 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\(^8\) (effect size ($\beta$) = 1.29, standard error (S.E.) = 0.47; $P = 0.0073$); phenotypic correlations are shown in \textit{Supplementary Table 26}. This effect was driven by performance (PIQ; $\beta = 1.74$, 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: $C; \beta = 0.126$, 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 \textit{DDR2} (1q23.3; $P = 5.81 \times 10^{-7}$) (\textit{Table 1}), which encodes a receptor tyrosine kinase involved in cell growth and differentiation\(^9\).

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 \textit{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\(^10\). Notably, the discovery sample had 99.92% power to detect variants with effect sizes of 1% of the variance for MAF $\geq 0.05$. 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.

\section*{METHODS}

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

\section*{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 fees 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 Daltonics, Abbott, Jurlab 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, Aracdon and Institüt Catala de Neurociencies Apicades, the Gulf War Veterans Illnesses Advisory Committee, VACO, Biogen Idec and Pfizer. M.W. received funding for consulting from AstraZeneca, Aracdon, Medivation/Pfizer, Ipsen, TauRx Therapeutics, Bayer Healthcare, Biogen Idec, Exonhit Therapeutics, SA, Servier, Nanyc, Pfizer and Janssen; for travel from NeuroVigil, CHRU–Hôpital Roger Salengro, Siemens, AstraZeneca, Geneva University Hospitals, Lilly, the University of California, San Diego–ADNI, Paris University, Institüt Catala de Neurociencies Apicades, the University of New Mexico School of Medicine, Ipsen, Clinical Trials on Alzheimer's Disease (CTAD), Pfizer, AD PD Meeting, Paul Sabathier 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 Apicades. 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 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.

\section*{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 AG19771-09) for additional data analysis and NCRRAD (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 Co., 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 NIBIR, 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 neuropyschological 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. Makkijie, M. Hakobyan 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.

BfBRN: BfBRN wishes to acknowledge D.B. Keator for leading BfBRN 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. Vovodic for their leadership contributions to BfBRN scanner and sequence calibration, tool development and data collection efforts. The BfBRN study was supported by the US NIH (U24 RR121992) for phenotypic data collection. Genotyping was performed with the support of the grant R01/045805/WH 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 RR17515). P.K. was also supported by an NIH grant (EB06395).

IMAGEN: IMAGEN is funded by the European Commission Framework Programme 6 (FP-6) Integrated Project IMAGEN (PL07286), the European Commission Framework Programme 7 (FP-7) Programme Alzheimer's Disease, Alcoholism, Memory, Schizophrenia (ADAMS), the FP-7 Innovative Medicines 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 (93538).

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 work was also supported 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 (G01001401 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 (MRC/ES/L002676/1, 961048/1 and 100025401).

IMAGEN is funded by the European Commission Framework Programme 6 (FP-6) Integrated Project IMAGEN (PL07286), the European Commission Framework Programme 7 (FP-7) Programme Alzheimer's Disease, Alcoholism, Memory, Schizophrenia (ADAMS), the FP-7 Innovative Medicines 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 (93538).

MOODS: This work was funded by the German Federal Ministry of Education and Research (BMBF) in the National Genome Research Network (NFGN-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.

MPP: The MPP 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 MIPP and control subjects acquired at the Department of Psychiatry at the Ludwig-Maximilians-University. We wish to acknowledge A. Olynyk 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 GlaasSmithKline 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 (NFGN) (FKZ 01G05081).

NCNG: We would like to thank the personnel involved in recruitment and data collection and, in particular, P. Due-Tonnessen 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 Regionalhelseforetak (RHF) and the Dr Einar Martinsen Fund.

NEFDA-NTR: Funding was obtained from the NWO (MagW/ZonMW 904-61-090; 985-10-902; 904-61-193; 480-04-004; 000-05-717, Addiction -3160008; 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/QLK-T-2001-01254), the European Community's FP6 (HEALTH-F4-2007-20143), the European Research Council (ERC) Genetics of Mental Illness (230374), Rutgers University Cell and DNA Repository (cooperative agreement NIMH U24 MH08457-06), the US NIH (R01D0042157-01A) and the Genetic Association Information Network (a public-private partnership between the NIH and Pfizer, Affymetrix and Abbott Laboratories).

NIHR-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 the RUD 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 (NFGN) (FKZ 01G05081).

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 HD057335) and the Australian National Health and Medical Research Council (NHMRC).
The Big Lottery Fund, the Wolfson Trust and the Epilepsy Society. This work was supported from the Wellcome Trust (grant 084730 to S.M.S.), University College London EPIGEN: and the Science Foundation Ireland (08/IN.1/B1916_Corvin).


COMPETING FINANCIAL INTERESTS
The authors declare no competing financial interests.

29. Levay, K. & Slepek, V.Z. Up- or downregulation of tescalcin in HLA-60 cells is associated with their differentiation to either granulocytic or macrophage-like lineage. Exp. Cell Res. 316, 1254–1262 (2010).
LETTERS


© 2012 Nature America, Inc. All rights reserved.
The Netherlands. 62Department of Psychiatry, VU University Medical Center, Amsterdam, The Netherlands. 63Department of Molecular and Cellular Therapeutics, Royal College of Surgeons in Ireland, Dublin, Ireland. 64Department of Genetics, Texas Biomedical Research Institute, San Antonio, Texas, USA. 65Center for Neurobehavioral Genetics, University of California, Los Angeles, USA. 66Division of Neurology, Beaumont Hospital, Dublin, Ireland. 67Department of Neurology, Hospital Erasme, Universite Libre de Bruxelles, Brussels, Belgium. 68Athinaou A. Martins Center for Biomedical Imaging, Massachusetts General Hospital, Boston, Massachusetts, USA. 69Research Imaging Institute, University of Texas Health Science Center at San Antonio, San Antonio, Texas, USA. 70Department of Radiology, University of California, San Diego, La Jolla, California, USA. 71Department of Cognitive Neuroscience, Radboud University Nijmegen Medical Centre, The Netherlands. 72Max Planck Institute for Psycholinguistics, Nijmegen, The Netherlands. 73Department of Clinical and Experimental Epilepsy, University College London, Institute of Neurology, London, UK. 74Maryland Psychiatric Research Center, Department of Psychiatry, University of Maryland School of Medicine, Baltimore, Maryland, USA. 75Department of Genomics, Life and Brain Center, University of Bonn, Bonn, Germany. 76Institute of Human Genetics, University of Bonn, Bonn, Germany. 77Institute for Genomic Mathematics, University of Bonn, Bonn, Germany. 78Department of Psychiatry, Ludwig-Maximilians-University (LMU), Munich, Germany. 79Division of Mental Health and Addiction, Oslo University Hospital, Oslo, Norway. 80Institute of Clinical Chemistry and Laboratory Medicine, University of Greifswald, Greifswald, Germany. 81German Center for Neurodegenerative Disorders (DZNE), Bonn, Germany. 82Department of Psychiatry, University of Texas Health Science Center at San Antonio, San Antonio, Texas, USA. 83Montreal Neurological Institute, McGill University, Montreal, Quebec, Canada. 84Center for the Study of Human Cognition, Department of Psychology, University of Oslo, Oslo, Norway. 85Centre for Advanced Study, Oslo, Norway. 86School of Psychology, University of Queensland, Brisbane, Queensland, Australia. 87Laureate Institute for Brain Research, Tulsa, Oklahoma, USA. 88Department of General Psychiatry, Heidelberg University Hospital, University of Heidelberg, Heidelberg, Germany. 89Department of Psychiatry, Division of Medical Psychology, Bonn, Germany. 90The MRC Sudden Death Tissue Bank in Edinburgh, Department of Pathology, University of Edinburgh, Edinburgh, UK. 91Institute for Community Medicine, University of Greifswald, Greifswald, Germany. 92Department of Research and Development, Diakonhjemmet Hospital, Oslo, Norway. 93Department of Medical Genetics, Oslo University Hospital, Oslo, Norway. 94Aging and Dementia Imaging Research Laboratory, Department of Radiology, Mayo Clinic and Foundation, Rochester, Minnesota, USA. 95Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, Indiana, USA. 96Neurospin, Institut d’Imagerie Biomédicale (I2B), Commissariat à l’Energie Atomique, Gif-sur-Yvette, France. 97Medical Genetics Section, Molecular Medicine Centre, Institute of Genetics and Molecular Medicine, The University of Edinburgh, Western General Hospital, Edinburgh, UK. 98Geriatric Medicine Unit, The University of Edinburgh, Royal Victoria Hospital, Edinburgh, UK. 99Departments of Radiology, University of California, San Francisco, California, USA. 100Department of Medicine, University of California, San Francisco, California, USA. 101Department of Psychiatry, University of California, San Francisco, California, USA. 102Veterans Affairs Medical Center, San Francisco, California, USA. 103A full list of members is provided in the Supplementary Note. 104Cardiovascular Health Research Unit, Department of Medicine, University of Washington, Seattle, Washington, USA. 105Department of Epidemiology, Erasmus Medical Center University Medical Center, Rotterdam, The Netherlands. 106Department of Radiology, Erasmus Medical Center University Medical Center, Rotterdam, The Netherlands. 107Netherlands Consortium for Healthy Aging, Leiden, The Netherlands. 108Icelandic Heart Association, Kopavogur, Iceland. 109Faculty of Medicine, University of Iceland, Reykjavik, Iceland. 110University of Bordeaux, U708, Bordeaux, France. 111Institut National de la Santé et la Recherche Médicale (INSERM), Neuropedriat Orange U708, Bordeaux, France. 112Laboratory of Epidemiology, Demography, and Biometry, NIH, Bethesda, Maryland, USA. 113Department of Neurology, University of California, Davis, Sacramento, California, USA. 114Center of Neuroscience, University of California, Davis, Sacramento, California, USA. 115Department of Neurology, Boston University School of Medicine, Boston, Massachusetts, USA. 116National Heart, Lung, and Blood Institute’s Framingham Heart Study, Framingham, Massachusetts, USA. 117Division of Health Sciences (Medical Physics), The University of Edinburgh, Edinburgh, UK. 118Institute for Neuroscience and Medicine (INM-1), Research Center Juelich, Juelich, Germany. 119Semel Institute for Neuroscience and Human Behavior, David Geffen School of Medicine, University of California, Los Angeles, California, USA. 120Department of Biological and Medical Psychology, Faculty of Psychology, University of Bergen, Bergen, Norway. 121German Center for Neurodegenerative Diseases (DZNE), Rostock/Greifswald, Greifswald, Germany. 122Functional Magnetic Resonance Imaging of the Brain (FMRIB) Centre, Oxford University, Oxford, UK. 123Centre for Advanced Imaging, University of Queensland, Brisbane, Queensland, Australia. 124Department of Statistics, University of Warwick, Coventry, UK. 125Department of Psychiatry, University Medical Center Groningen, Groningen, The Netherlands. 126Information on the consortium is provided in the Supplementary Note. 127These authors contributed equally to this work. 128These 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)\textsuperscript{16} from the FMRIB Software Library (FSL) package of tools\textsuperscript{43} and FreeSurfer\textsuperscript{13}. 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)\textsuperscript{44} package or FreeSurfer. Estimated total intracranial volume was calculated through registration of each MRI scan to a standard brain image template\textsuperscript{35}, using either FSL FLIRT\textsuperscript{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\textsuperscript{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 \(r = 0.66\)\textsuperscript{46} and similar between the two in another (FreeSurfer \(r = 0.73\); FIRST \(r = 0.71\)\textsuperscript{47}). Scan-rescan reliability was also high for both methods (FreeSurfer intraclass correlation (ICC) = 0.98; FIRST ICC = 0.93)\textsuperscript{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 (ICC = 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)\textsuperscript{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)\textsuperscript{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\textsuperscript{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 \(\geq 0.05\).

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 samples was selected for imputation. Performance of software for imputation is generally similar between the most used methods\textsuperscript{53,54} for common variants (MACH\textsuperscript{55}, IMPUTE\textsuperscript{56} and BEAGLE\textsuperscript{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\textsuperscript{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\textsuperscript{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, SNPTST, 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 \(\beta\) 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 \(R^2 < 0.3\)). 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\textsuperscript{59}. 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\textsuperscript{63}. To account for heterogeneity across samples, a random-effects meta-analysis\textsuperscript{50} was also conducted via the program METASOFT without using genomic control. Using KGG\textsuperscript{64} we performed gene-based tests on the double genome–controlled meta-analysis results, using the extended Simes test\textsuperscript{65} to obtain an overall \(P\) value for association of the entire gene with a 50-kb boundary on either side. Results from genes with \(P < 1 \times 10^{-4}\) 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 \textit{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.
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 metanalytic 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. The effective number of tests was determined to be 4 and an overall genome-wide significance threshold of 5 × 10⁻⁵ 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 assessed by linear model. 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. 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. Genotype information was obtained using Illumina HumanOmni 1M arrays and standard quality control methods. Expression profiling was conducted in up to 10 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 probesets 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 (r² > 0.50) was removed from further analysis.

SNPExpress, a publicly available database, was also used for replication of the findings. The SNPExpress database used autopsy-collected frontal cortex brain tissue in 93 samples from human subjects with no neuropyschiatric 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. 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.