Molecular Psychiatry (2011) 16, 927–937

www.nature.com/mp

ORIGINAL ARTICLE

Discovery and replication of dopamine-related gene effects on caudate volume in young and elderly populations ($N = 1198$) using genome-wide search

JL Stein[1], DP Hibar[1], SK Madsen[1], M Khamis[1], KL McMahon[2], GI de Zubicaray[3], NK Hansell[4], GW Montgomery[4], NG Martin[4], MJ Wright[4], AJ Saykin[5], CR Jack Jr[6], MW Weiner[7,8], AW Toga[1], PM Thompson[1] and the Alzheimer's Disease Neuroimaging Initiative[9]

[1]*Department of Neurology, Laboratory of Neuro Imaging, UCLA School of Medicine, Los Angeles, CA, USA;* [2]*Center for Magnetic Resonance, School of Psychology, University of Queensland, Brisbane, Queensland, Australia;* [3]*Functional Magnetic Resonance Imaging Laboratory, School of Psychology, University of Queensland, Brisbane, Queensland, Australia;* [4]*Queensland Institute of Medical Research, Brisbane, Queensland, Australia;* [5]*Department of Radiology and Imaging Science, Center for Neuroimaging, Indiana University School of Medicine, Indianapolis, IN, USA;* [6]*Mayo Clinic, Rochester, MN, USA;* [7]*Departments of Radiology, Medicine, and Psychiatry, UC San Francisco, San Francisco, CA, USA and* [8]*Department of Veterans Affairs Medical Center, San Francisco, CA, USA*

The caudate is a subcortical brain structure implicated in many common neurological and psychiatric disorders. To identify specific genes associated with variations in caudate volume, structural magnetic resonance imaging and genome-wide genotypes were acquired from two large cohorts, the Alzheimer's Disease NeuroImaging Initiative (ADNI; $N = 734$) and the Brisbane Adolescent/Young Adult Longitudinal Twin Study (BLTS; $N = 464$). In a preliminary analysis of heritability, around 90% of the variation in caudate volume was due to genetic factors. We then conducted genome-wide association to find common variants that contribute to this relatively high heritability. Replicated genetic association was found for the right caudate volume at single-nucleotide polymorphism rs163030 in the ADNI discovery sample ($P = 2.36 \times 10^{-6}$) and in the BLTS replication sample ($P = 0.012$). This genetic variation accounted for 2.79 and 1.61% of the trait variance, respectively. The peak of association was found in and around two genes, *WDR41* and *PDE8B*, involved in dopamine signaling and development. In addition, a previously identified mutation in *PDE8B* causes a rare autosomal-dominant type of striatal degeneration. Searching across both samples offers a rigorous way to screen for genes consistently influencing brain structure at different stages of life. Variants identified here may be relevant to common disorders affecting the caudate.
Molecular Psychiatry (2011) **16,** 927–937; doi:10.1038/mp.2011.32; published online 19 April 2011

Keywords: caudate; dopamine; genome-wide association; heritability; PDE8B; WDR41

Introduction

Human brain structure is under strong genetic control,[1,2] but specific genetic variants influencing individual differences are largely unknown. Genes that contribute to structural brain variation are important to identify, as several known examples confer protection or risk for mental illness or brain degeneration. Carriers of the prevalent epsilon 4 allele of the apolipoprotein E gene, for example, have a threefold increased risk for Alzheimer's disease.[3] They also show cortical thinning even in childhood, which may influence their vulnerability to later illness.[4] Searching the genome for associations to biological traits, such as measures of brain structure, may also help to identify genetic variants related to brain disorders.[5] Here, we used a genome-wide search to identify common genetic variants associated with caudate nucleus volume. We used two large independent samples to verify any associations and guard against false-positive findings. As the caudate is implicated in several neurodegenerative, motor, affective and developmental disorders, factors that influence its structure in human populations are of great interest.

Correspondence: Dr P Thompson, Department of Neurology, Laboratory of Neuro Imaging, UCLA School of Medicine, Neuroscience Research Building 225E, 635 Charles Young Drive, Los Angeles, CA 90095-1769, USA.
E-mail: thompson@loni.ucla.edu

[9]Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (www.loni.ucla.edu/ADNI). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators is available at http://www.loni.ucla.edu/ADNI/Collaboration/ADNI_Manuscript_Citations.pdf.

Received 14 July 2010; revised 6 November 2010; accepted 17 February 2011; published online 19 April 2011

The volume of the caudate is a highly heritable feature of brain structure.[2] The strong contrast between the caudate and the surrounding white matter in standard magnetic resonance imaging (MRI) allows it to be accurately and reliably delineated by automated recognition programs. In addition, caudate degeneration is characteristic of several genetic diseases. Caudate degeneration occurs in several rare Mendelian disorders: Huntington's disease,[6,7] pantothenate kinase-associated neurodegeneration,[8] neuroferritinopathy[9] and autosomal-dominant striatal degeneration.[10,11] In these cases, linkage analysis in affected families revealed specific causal genetic variants associated with caudate degeneration and impaired cognition. These disorders are rare, and the genetic variants identified so far are not widely carried in the general population. Caudate volume is altered in several more common disorders such as major depression,[12] Alzheimer's disease,[13] attention deficit hyperactivity disorder[14] and schizophrenia.[15,16] These disorders are highly heritable but their onset and trajectory is thought to be influenced by a large number of genetic polymorphisms, each with a small effect,[17] as well as environmental factors. These features of the caudate make it extremely interesting and tractable for genome-wide association studies.

Previous studies have explored how variants in genes expressed in the brain's monoamine neurotransmitter pathways may also influence caudate volume. The serotonin transporter polymorphism (5-HTTLPR) was associated with reduced caudate volumes in patients with depression ($N=61$),[18] and a DRD2 polymorphism was associated with reduced left caudate volume in memory impaired elderly subjects ($N=49$).[19] A DAT1 polymorphism was also associated with caudate volume in attention deficit hyperactivity disorder patients, their unaffected siblings, and healthy controls ($N=90$).[20] These reports suggest specific candidate genes that may affect caudate volume, but the studies are limited by small sample sizes. To date, no studies have attempted to replicate the findings in new samples, and the findings would be more credible if verified in larger samples.[21,22]

Here, we used an unbiased search across the entire genome, in two separate cohorts scanned with brain MRI, to find common genetic variants associated with caudate volume. We used a large discovery sample of elderly subjects and a replication sample of young adult twins. By design, these samples differ by around 50 years in mean age. As such, they could fail to replicate genetic effects present during only one part of the human lifespan. Even so, the joint use of both samples, young and old, offers a rigorous way to screen for genes that consistently influence brain structure at different stages of life.

Subjects and methods

Subjects

We analyzed two independent samples with neuroimaging and genome-wide genotype data: the Alzheimer's Disease Neuroimaging Initiative (ADNI) and the Brisbane Adolescent/Young Adult Longitudinal Twin Study (BLTS).

The ADNI sample has been described previously[23] as detailed in the Supplementary Information. The ADNI cohort included multiple diagnostic groups, and the genome-wide analysis was deliberately not split into diagnostic groups as the goal was to analyze the full phenotypic continuum,[24] maintaining greatest power to detect genetic associations. Our final analysis included 734 individuals (average age $\pm$ s.d. $= 75.5 \pm 6.8$ years; 432 men/302 women) including 172 Alzheimer's disease patients (75.5 ± 7.6 years; 94 men/78 women), 357 mild cognitive impairment (MCI) subjects (75.2 ± 7.3 years; 227 men/130 women), and 205 healthy elderly controls (76.1 ± 5.0 years; 111 men/94 women). Effect sizes for individual genetic variants on brain measures are expected to be small, so the large phenotypic variation in this continuum of subjects should increase power to detect genetic effects.[20,23]

The BLTS sample consists of healthy, young adult Australian monozygotic and dizygotic twins and their singleton siblings (see Supplementary Information). The final analysis included 464 individuals from 239 families (85 monozygotic twin pairs, 97 dizygotic twin pairs, 2 sets of triplets and 94 singletons; 23.7 ± 2.1 years; 188 men/276 women).

Genotyping

Genotyping for both the ADNI and BLTS samples was performed using the Illumina 610-Quad BeadChip (San Diego, CA, USA). After quality control, as previously outlined[23,25] (see Supplementary Information), 546 314 single-nucleotide polymorphisms (SNPs) remained in ADNI, and 542 478 SNPs in the BTLS sample. A total of 520 459 SNPs were jointly analyzed in both samples.

Imaging acquisition and pre-processing

High-resolution structural brain MRIs were acquired in both the ADNI and BLTS samples. Standard pre-processing was applied including registration to a standard template[26] so that images were globally matched in size and mutually aligned, but local differences in shape and size remained intact. Acquisition parameters and pre-processing details are found in the Supplementary Information.

Automatic delineation of caudate volume

We extracted models of the caudate from each structural MRI using an automated segmentation method based on adaptive boosting, which learns the features that best differentiate the caudate based on expert manual delineations of a small subset of the MRI scans.[13,27] Caudate nuclei were traced according to previously published anatomical protocols[28,29] and manual tracing guidelines and algorithm usage details are found in the Supplementary Information and Supplementary Figure 1.

Heritability analyses
The heritability of caudate volume was calculated using structural equation modeling (s.e.m.) implemented using Mx software (version 1.68; http://www.vcu.edu/mx/). This method estimates path coefficients in the widely used 'ACE' model,[30,31] fitted to the observed covariance matrices of monozygotic and dizygotic twin traits (see Supplementary Information).

Genetic analysis
For the ADNI sample, association was conducted using *Plink* software[32] (version 1.05; http://pngu.mgh.harvard.edu/purcell/plink/) to conduct a regression at each SNP with the number of minor alleles, age and sex as the independent variables and the quantitative phenotype (caudate volume) as the dependent variable, assuming an additive genetic model. In the BLTS sample, we performed mixed-model regression to conduct genetic association while adjusting for family relatedness,[33] sex and age. This analysis was performed using Efficient Mixed-Model Association[34] (http://mouse.cs.ucla.edu/emma/) within the R statistical package (see Supplementary Information). Note that if all subjects had been unrelated, the results from the mixed-model regression would be equivalent to the results from the standard regression in *Plink*. Methods for additional genetic analyses including within-group permutation to control for diagnostic status and meta-analysis of genetic results can be found in the Supplementary Information.

Results

Segmented caudate volumes
In the ADNI sample, the average left and right volumes (±s.d.) were: $3521.1 \pm 576.9\,mm^3$ and $3396.9 \pm 624.1\,mm^3$, respectively. Left and right caudate volumes were highly correlated ($r = 0.828$; $P < 2.2 \times 10^{-16}$; $N = 734$). In the BLTS sample, the average left and right volumes (±s.d.) were: $3956.0 \pm 507.0\,mm^3$ and $3986.7 \pm 462.9\,mm^3$, respectively. The correlation between left and right caudate volumes was also high ($r = 0.920$; $P < 2.2 \times 10^{-16}$; $N = 464$). The volumes in the BLTS sample were larger than in ADNI, as expected for younger subjects (right: $t = 17.5$; $P < 2.2 \times 10^{-16}$; Left: $t = 13.3$; $P < 2.2 \times 10^{-16}$; average bilateral: $t = 16.1$; $P < 2.2 \times 10^{-16}$).

Reliability of caudate volume measurement
To assess the reliability of the volume measurements, 40 BLTS subjects were scanned twice (time between scans, ±s.d.: 120 ± 55 days) and the left and right caudate were separately segmented using the algorithm. Measured volumes were highly reproducible for the left (intraclass correlation coefficient (ICC) = 0.986), right (ICC = 0.985) and average bilateral (ICC = 0.990) caudate volumes (Supplementary Figure 3).

Heritability estimates
Heritability analyses use the classical twin design to ascribe proportions of the observed variance (for example, of the volume or shape of a brain structure) in various degrees, to several factors: additive genetic effects (*A*), common environment shared by both twins (*C*) and unique environment/experimental error (*E*).[30,35] Heritability estimates, which were computed for caudate volume using the BLTS sample, were high (around 90%) relative to other brain structures;[2] additive genetic effects significantly contributed to the model fit (Table 1). The heritability is also evident in Supplementary Figure 2, where caudate volumes in monozygotic twins resemble each other (black dots) more closely than those in dizygotic twins (open dots).

Table 1 Heritability of caudate volume in the BLTS cohort

	Model fit statistics					*Estimates (%)*		
	−2LL	*df*	*Δ−2LL*	*Δdf*	P	*A*	*C*	*E*
Left caudate								
ACE	1076.585	451	—	—	—	76.6	14.0	9.4
AE	**1077.836**	**452**	**1.251**	1	**0.26**	**90.8**	**—**	**9.2**
CE	1133.893	452	57.308	1	3.73×10^{-14}	—	67.3	32.7
Right caudate								
ACE	1117.580	451	—	—	—	76.5	10.0	13.4
AE	**1118.184**	**452**	**0.603**	**1**	**0.44**	**86.8**	**—**	**13.2**
CE	1158.719	452	41.139	1	1.42×10^{-10}	—	63.0	37.0
Average caudate								
ACE	1068.968	451	—	—	—	78.0	13.6	8.4
AE	**1070.190**	**452**	**1.222**	**1**	**0.27**	**91.7**	**—**	**8.3**
CE	1132.567	452	63.599	1	1.53×10^{-15}	—	67.8	32.2

Best-fitting models are shown in bold.

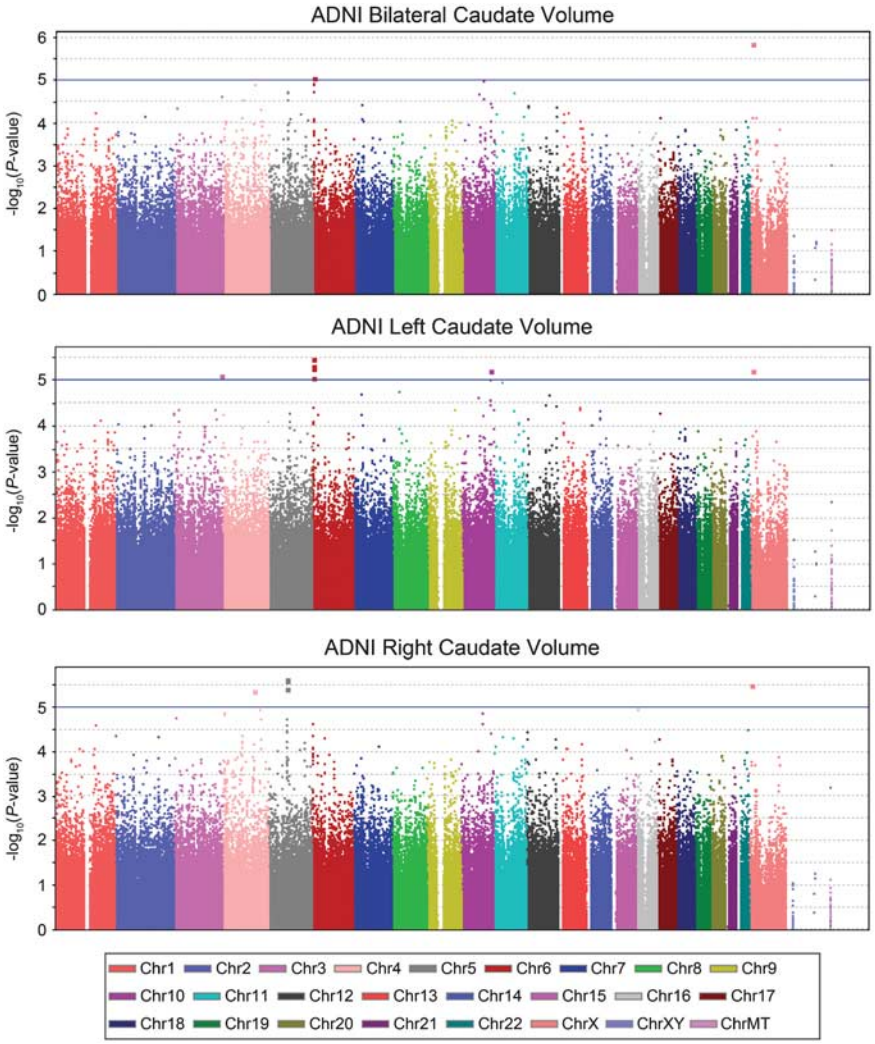

Figure 1 Manhattan plots show the significance of association of each SNP with caudate volume, from genome-wide association analysis conducted in the ADNI cohort. Each marker is represented as a dot and the $-\log_{10}$(*P*-value) is displayed on the *y* axis. Association was conducted separately for average bilateral (top), left (middle), and right (bottom) caudate volumes. Markers above the blue line represent a *P*-value $<1\times10^{-5}$. ChrXY represents the pseudo-autosomal region of the X chromosome, and ChrMT represents mitochondrial SNPs. BLTS Manhattan plots are shown in Supplementary Figure 5.

Genome-wide association

Given the high heritability of caudate volume, we conducted a genome-wide association analysis to search for common genetic variants that might explain some modest proportion of the substantial genetic influence on caudate volume. We analyzed the 734 ADNI subjects as a discovery sample and the 464 BLTS subjects as a replication sample (1198 subjects, in total). For ADNI, we conducted a standard regression of phenotype on the additive allelic effect at each SNP, after statistically controlling for age and sex. Genome-wide association results for the ADNI sample are shown in Figure 1 and the most significant SNPs are presented in Table 2, at a threshold of $P<1\times10^{-5}$. Subsequent replication of the findings was conducted using the BLTS sample. As noted in the Subjects and methods section, a mixed effects model was used to regress the phenotype on the additive allelic effect at each SNP, after statistically controlling for age, sex and for genetic relatedness, through the kinship matrix. Q-Q plots and λ inflation factors[36] show no inflation of statistical significance (Supplementary Figure 4).

Meta-analytic methods were used to combine the two groups rather than a combined 'mega-analysis' (that is, pooling all volume measures from both studies), because of differences in subject demographics and image acquisition. Note that in Table 2, if the reference allele is the same in both samples and the sign of the β-coefficient is same in both samples, then the effect is in the same direction in both samples. Similarly, if the reference allele is *different* in both samples and the sign of the β-coefficient is *opposite* in both samples, then the effect is in the same direction in both samples.

A large peak of replicated association is found in and around two genes: *WDR41* and *PDE8B* (Figure 2).

Table 2 Most associated SNPs from the ADNI discovery cohort ($P < 1 \times 10^{-5}$) and BLTS replication cohort

				ADNI							*BLTS*						*Pooled*		
Chr	*SNP*	*Position*	*Gene*	*A1*	*Freq*	N	*β*	*SE*	*P*	*P | diag*	*A1*	*Freq*	N	*β*	*SE*	*P*	*β*	*SE*	*P*
Average bilateral caudate																			
X	rs5979778	12893224	*TMSB4X*	G	0.429	733	−159.3	32.8	1.42×10^{-6}	4.00×10^{-6}	A	0.607	464	6.55	36.9	0.859	−91.9	24.5	1.78×10^{-4}
6	rs9378688	2180632	*GMDS*	A	0.118	734	−198.1	44.2	8.69×10^{-6}	2.83×10^{-5}	G	0.872	464	60.6	49.3	0.220	−136.8	32.9	3.22×10^{-5}
Left caudate																			
6	rs9378688	2180632	*GMDS*	A	0.118	734	−208.4	44.6	3.53×10^{-6}	7.00×10^{-6}	G	0.872	464	79.1	52.5	0.133	−154.2	34.0	5.72×10^{-6}
6	rs9392374	2168001	*GMDS*	G	0.118	727	−206.3	44.9	5.10×10^{-6}	1.00×10^{-5}	T	0.873	464	80.7	52.8	0.127	−153.6	34.2	7.11×10^{-6}
6	rs9405550	2175063	*GMDS*	T	0.114	732	−208.4	45.5	5.57×10^{-6}	1.20×10^{-5}	T	0.130	463	−83.9	52.3	0.109	−154.8	34.3	6.53×10^{-6}
10	rs4752194	120485154	*C10orf46*	C	0.183	734	−166.2	36.5	6.28×10^{-6}	1.00×10^{-5}	C	0.178	464	14.7	47.9	0.760	−99.7	29.0	5.91×10^{-4}
10	rs4752195	120496674	*C10orf46*	C	0.183	734	−166.2	36.5	6.28×10^{-6}	1.00×10^{-5}	T	0.822	464	−14.7	47.9	0.760	−99.7	29.0	5.91×10^{-4}
X	rs5979778	12893224	*TMSB4X*	G	0.429	733	−150.7	33.2	6.43×10^{-6}	9.00×10^{-6}	A	0.607	464	−1.39	39.3	0.972	−87.4	25.4	5.71×10^{-4}
3	rs1602517	193287854		A	0.311	731	−138.8	30.9	8.20×10^{-6}	6.00×10^{-6}	T	0.303	464	−48.4	38.2	0.205	−64.8	24.0	7.03×10^{-3}
6	rs234937	2175251	*GMDS*	C	0.267	731	−147.9	33.1	8.88×10^{-6}	1.40×10^{-5}	T	0.732	464	34.5	38.4	0.370	−99.6	25.1	7.15×10^{-5}
Right caudate																			
5	rs163030	76817227	*WDR41*	A	0.499	731	147.4	31.0	2.36×10^{-6}	6.00×10^{-6}	C	0.524	462	−83.2	33.2	0.012	117.5	22.7	2.15×10^{-7}
5	rs163035	76815798	*WDR41*	T	0.499	734	147.2	31.0	2.47×10^{-6}	6.00×10^{-6}	G	0.524	461	−83.2	33.2	0.013	117.4	22.7	2.21×10^{-7}
X	rs5979778	12893224	*TMSB4X*	G	0.429	733	−168.0	35.7	3.09×10^{-6}	7.00×10^{-6}	A	0.607	464	13.4	36.3	0.712	−92.0	25.5	3.01×10^{-4}
5	rs335636	76760355	*PDE8B*, *WDR41*	A	0.500	733	143.8	30.9	3.90×10^{-6}	8.00×10^{-6}	G	0.525	464	−81.1	33.0	0.014	114.5	22.6	3.84×10^{-7}
4	rs1299288	132606407		G	0.230	734	−169.4	36.6	4.43×10^{-6}	5.00×10^{-6}	T	0.770	464	−42.0	39.3	0.286	−71.2	26.8	7.84×10^{-3}

Abbreviations: ADNI, Alzheimer's Disease NeuroImaging Initiative; BLTS, Brisbane Adolescent/Young Adult Longitudinal Twin Study; SNP, single-nucleotide polymorphism.
Genes close to the SNPs are displayed and cells are empty if no gene is within ±50 kb. *A1* is the reference allele, and *Freq* shows the frequency of the reference allele, *β* shows the mean increase in each volume (in mm^3) per added reference allele controlling for age and sex, SE gives the standard error of the β-coefficient, and *P | diag* gives the *P*-value controlling for diagnosis using a permutation-based method. Information was pooled across samples using inverse weighted variance meta-analysis and the *Pooled* β is from the ADNI reference allele. Deviations from Hardy–Weinberg equilibrium are shown in Supplementary Table 2.

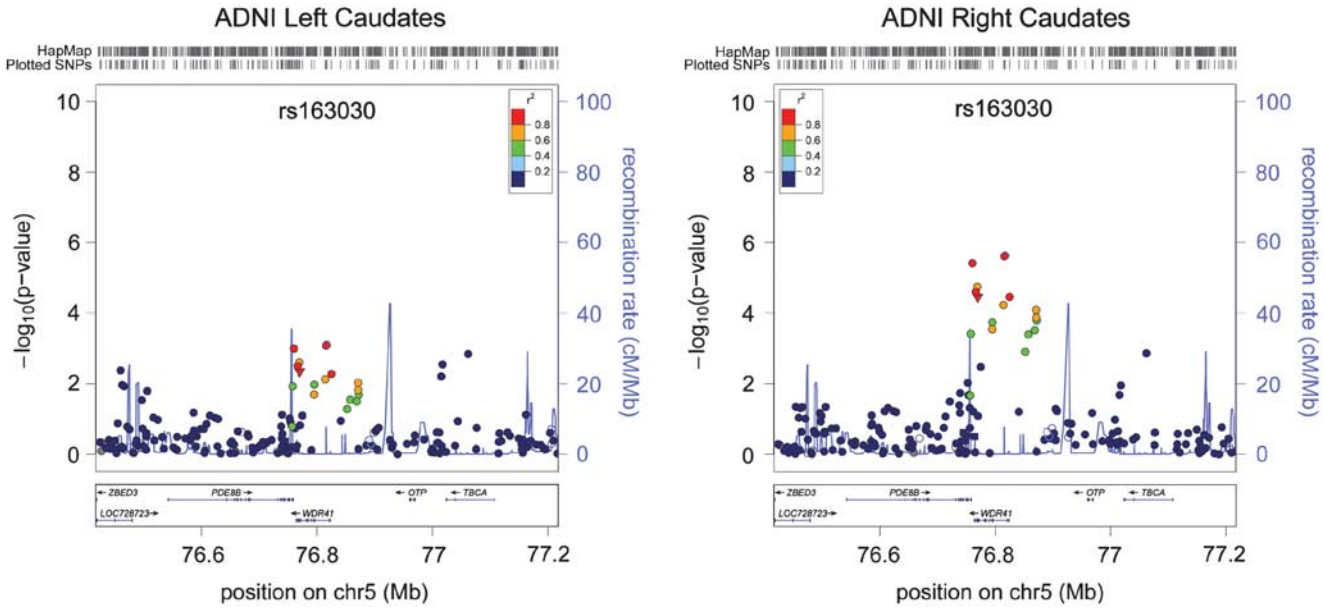

Figure 2 Detailed view of the associated locus. Markers are represented as circles (SNPs with no known function) or downward-pointing triangles (coding non-synonymous mutations). Markers are placed at their position on chromosome 5 (x axis) and graphed based on the $-\log_{10}$(P-values) of their association to the phenotype (y axis). The level of linkage disequilibrium to the most associated SNP (rs163030) is represented in color using the CEU panel from HapMap Phase II. The location of genes is shown below the plots. Images were created using *LocusZoom* (http://csg.sph.umich.edu/locuszoom/).

Association is strongest for the right caudate, but it is also found at a slightly weaker significance level for the left caudate.

A large region encompassing the most highly associated region of the genome contains both the untranslated region of *PDE8B* and several exons of *WDR41*. Many of these variants have functional relevance. rs335636 is found within a 5850-bp deletion region (rs71823322) within the untranslated region of both *WDR41* and *PDE8B* genes (Figure 2). Several other SNPs within *WDR41* are coding non-synonymous base pair changes, meaning that they change the amino acids formed by the *WDR41* gene. A SNP in the same locus, associated at a slightly weaker significance level, rs919224 ($P = 1.82 \times 10^{-5}$) is directly adjacent to two SNPs that code for missense mutations in one of the exons of *WDR41* (rs35774719 and rs17856057). These SNPs are not directly genotyped in the HapMap database so it is not possible to determine the exact linkage disequilibrium pattern between them. However, they do reside within a linkage disequilibrium block so they are likely to be in high linkage disequilibrium.

Other genes of interest to caudate volume identified here—but with little evidence for replication—include *GMDS*, *C10orf46* and TMSB4X. Intergenic SNPs were also identified but not replicated, in chromosomes 3 and 4.

Using a permutation-based approach, there was little change in the P-values of the replicated SNPs when using a null distribution preserving the effect of diagnosis in the ADNI sample (see $P|$diag column in Table 2). Additionally, we tested whether the effect of the top SNP found through this study, rs163030, was present in each of the three diagnostic groups of the ADNI cohort. The effect of the most significant SNP on right caudate volume (controlling for age and sex) was found in the Alzheimer's disease group ($N = 170$; $\beta = 168.8$; $P = 0.0139$), the MCI group ($N = 357$; $\beta = 143.5$; $P = 6.41 \times 10^{-4}$), and as a strong statistical trend in the healthy elderly group ($N = 204$; $\beta = 119.1$; $P = 0.0537$). As expected the significance levels are affected by the number of subjects in each group—the MCI group has a lower P-value than the Alzheimer's disease group, which itself has a lower P-value than the healthy elderly group. Notably though, the effect size is in the same direction and of similar magnitude across all the subsamples split by diagnosis. This shows that the results are driven by all three groups jointly rather than originating as an effect of one group alone.

In a *post hoc* exploratory analysis, we also examined an alternative analysis that used the BLTS cohort as the discovery sample and the ADNI cohort as the replication sample. For this 'switched' analysis, we selected only those SNPs that had strong association with caudate volume ($P < 1 \times 10^{-5}$) within the smaller BLTS cohort, and sought replication in the ADNI cohort. Using the BLTS cohort as a basis to select candidate genes resulted in no replications within the ADNI cohort (see Supplementary Table 1). Selecting the sample with the highest sample size as the discovery sample maximized the likelihood for replication.

Replication attempts of previously tested candidate genes

Neither the 5-HTTLPR polymorphism[18] nor the *DAT1* variable number of tandem repeats polymorphism, rs28363170,[20] were genotyped on the chip used in either sample so replication of the association could not be tested. The *DRD2* Taq1A polymorphism (rs1800497) on chromosome 11 was previously associated with caudate volume.[19] Here, we find little evidence for replication in the ADNI cohort using average bilateral caudate volume ($t = 0.964$; $P = 0.335$; $N = 704$), left caudate volume ($t = 1.263$; $P = 0.207$; $N = 704$) or right caudate volume ($t = 0.600$; $P = 0.549$; $N = 704$). Similarly in the BLTS cohort, little evidence for replication was found for average bilateral caudate volume ($t = 0.400$; $P = 0.691$; $N = 464$), left caudate volume ($t = 0.628$; $P = 0.530$; $N = 464$) or right caudate volume ($t = 0.168$; $P = 0.867$; $N = 464$).

Discussion

This study identified specific genetic variations associated with caudate volume in the human brain, in 1198 subjects. This is one of the largest brain imaging studies ever performed. There was sufficient power to trace heritable variation to specific variations on the genome, although not at a genome-wide significance level. We replicated the same genetic associations in samples from two continents (United States and Australia), separated in mean age by 50 years, and using data collected on scanners with different field strengths (4 Tesla and 1.5 Tesla). Additional replication in still larger samples would be advantageous, but this confirmation in two independent samples suggests that these associations may be robust and may persist throughout life.

The caudate volume is a reasonable starting point for investigating genetic influences on brain structure because it is highly heritable (Table 1), reliably delineated by automated recognition programs[37,38] (Supplementary Figure 3), and has an established link to psychopathology. The estimates of caudate volume heritability from the BLTS cohort (shown in Table 1) are around 76% for the ACE model (one of the standard classical twin models used to assess heritability), and around 90% for the best-fitting AE model. This agrees with a previous study assessing caudate volume in twins,[2] which showed caudate heritability of 70–79% in an ACE model. That study analyzed many other brain structures as well and, although the heritability coefficients of different regions were not directly compared for statistical significance, the caudate showed consistently high heritability relative to other structures.

A relatively large region on chromosome 5 was found to have replicated significance in its association with caudate volume in each of the independent populations, including genes *WDR41* and *PDE8B* (Table 2 and Figure 2). Functionally, the region containing both of these genes is essential to dopaminergic neuron development in zebrafish.[39] *WDR41* was also useful in improving the performance of a diagnostic classification algorithm, which used gene expression patterns to distinguish schizophrenia patients versus healthy controls.[40] *PDE8B* is highly expressed in the rat brain and in neuronal cells.[41] The protein product of the gene, phosphodiesterase, is a key protein in the dopamine signaling casade. Dopamine binding to receptors stimulates or inhibits cAMP production, which is subsequently degraded by phosphodiesterase 8B.[42–44] *PDE8B* is associated with susceptibility to major depression and antidepressant treatment response,[45] and has higher expression in Alzheimer's disease relative to controls.[46] Additionally, autosomal-dominant striatal degeneration is caused by a mutation in *PDE8B*.[10]

The possible relation of these genes to a Mendelian disorder is of great interest. Although specific variants known to cause Mendelian disorders do not necessarily influence normal variability or psychopathology, the same genes may be relevant for normal variability and psychopathology. In genetic studies of obesity, for example, common variants have subtler but similar effects to highly penetrant rare Mendelian mutations.[47] In that study, common SNPs within *ABCG8* and *LCAT* increased risk for dyslipidemia. Mendelian mutations within those same genes are causal for dyslipidemia. Similarly, in our study, common SNPs within the *PDE8B*/*WDR41* region were associated with differences in caudate volume and a Mendelian mutation within the *PDE8B* gene is causal for an autosomal-dominant form of striatal degeneration. This shows that Mendelian mutations may be clues for selectively picking genes to understand the normal variability, although the specific Mendelian mutations themselves may not be involved in the normal variability.

Such replicated genetic hits suggest that our findings are consistent with the literature on dopamine function in the caudate. The caudate receives projections from the dopaminergic neurons of the *substantia nigra* and has high concentration of D_1 and D_2 dopamine receptors.[48] These genes are crucial for the development and function of dopamine neurons. This provides biological plausibility that they may also contribute to variations in caudate anatomy. *WDR41* and *PDE8B*-mediated differences in caudate structure accounted for 2.79 and 1.61% of the trait variance in the ADNI and BLTS samples, respectively at the most associated SNP. These genetic influences on dopamine function and brain structure may also influence behavior, as dopamine is essential for normal cognitive function.[49,50] As such, the genes identified here may become candidates for examination on studies of disorders that affect the caudate, to determine whether they are over-represented in subjects with developmental insufficiencies or deterioration in caudate function.

Three other genes were identified as influencing caudate volume in the ADNI cohort, but were not replicated in the BLTS cohort. *GMDS* encodes an enzyme involved in metabolism pathways, and is also

important for neuronal migration.[51] *C10orf46* (also called *CAC1*) has been characterized as a cell cycle-associated protein.[52] *TMSB4X* is expressed in the brain and involved with corticogenesis[53] and with actin polymerization.[54] Lack of replication in both cohorts may be due to false-positive findings or age-specific gene effects. Additionally, though the *DRD2* Taq1A allele was previously identified as putatively affecting caudate volume[19] as well as availability of striatal dopamine D_2 receptors,[55] we found little evidence for *DRD2* Taq1A association with caudate volume in either cohort.

Some strengths and limitations of this study deserve comment. First, we identified some variants of interest for caudate volume; however, we are unable to provide mechanistic evidence for how these single base pair differences in the genome affect brain structure. Further mechanistic understanding could be derived by studying both the expression and protein function of the gene products that lie downstream of the SNP variations identified here. Unfortunately, no expression or protein data are currently available in either cohort to directly test these hypotheses. Second, we provide strong support for a particular region in the genome associated to caudate volume, yet it remains to be demonstrated that the genetic factors identified here are of interest for pathophysiology. Third, the two neuroimaging samples are taken from different parts of the lifespan. Replications of SNPs many indicate gene effects that persist, or have different modes of action, throughout life. Lack of replications could either be true negatives, or may reflect age- or cohort-specific effects. In a sense, the use of two very diverse samples on two different continents presents a very high bar for replication. Owing to large differences in the mean age of the samples, it would be logical to assume that some robust genetic events may not be simultaneously found in both of these young and old cohorts. For example, there may be a greater preponderance of aging or apoptotic events in the ADNI sample and more developmental or synaptogenic processes in the BLTS sample. As such, the use of two very different samples is likely to identify genes of enduring relevance across the lifespan. This may miss or fail to replicate effects that are only occurring, or are more dominant, in late or early life. On these grounds, replication should not be taken to imply that the genes found in our study operate on the same biological processes over the lifespan. Nor should it be taken to mean that genes not found in our study are not influential—other genes could impact caudate structure only during one phase of life. Fourth, like other multifactorial traits such as height,[56] individual common variants have small effect sizes and account for only a small proportion of the overall heritability so can only be detected with large sample sizes. Missing heritability might be attributed to low power, rare variants, un-genotyped variants, epistatic interactions or epigenetic contributions to heritability.[57] Finally, the ADNI cohort includes subjects across the continuum of healthy aging to mild cognitive impairment to Alzheimer's disease. Any genetic association in ADNI could be mediated by normal atrophy that occurs with healthy aging or by the disease. To account for this, we were able to perform an analysis controlling for diagnosis through permutation. This showed little change in the degree of association, implying that illness category is not driving the association. Furthermore, the broad range of imaging phenotypes in ADNI is sensitive to effects that may be overlooked if the discovery sample were more narrowly defined. As single genes are likely to have small effects on behavior, several studies advocate examining multiple cohorts where the spectrum of observable variation is larger than that in the general population,[20,23,58,59] especially in the discovery phase. Even so, we replicated the association in our young sample (healthy twins) so the gene effects are not restricted to those who are elderly or ill, and are also detectable in young people.

In this study, we assessed caudate volume rather than surface morphology because volume is an easily measured summary phenotype that is known to associate with disease. Additionally, performing simultaneous searches across both surface vertices and the genome requires complex statistical methods[60,61] not yet optimized for surfaces. Volume effects are also more interpretable and can be readily verified by many other groups.

Large genome-wide association study commonly use a genome-wide significance threshold of $P<5\times10^{-8}$ (see Ref. 56) but less conservative thresholds have been established using permutation testing or by estimating the effective number of tests on the genome.[62] Here, we used a search criterion to select SNPs that were highly associated in the larger cohort, at $P<1\times10^{-5}$, and then tested for replication in a separate cohort. This threshold does not represent a genome-wide significance threshold, but rather a two-stage process that identifies interesting SNPs to carry forward to a second stage in which they can be replicated. This threshold value is somewhat arbitrary but has been used previously in the literature to identify interesting SNPs in large association studies.[63]

It is of interest that although we found a replication across samples, the individual associations did not reach genome-wide significance level in each smaller sample. Similarly in a previous genome-wide association study,[64] a top SNP was found in one cohort that was not genome-wide significant but replicated in others with a lower threshold. The meta-analysis in our study of the individual cohorts separately did not reveal genome-wide significance values for any SNP (Table 2). Thus, despite the replication in two samples, even more studies are needed to verify this association.

The marginally greater effect size for genetic association in the right versus left caudate may be due to the known asymmetries in caudate volume. As we found in a recent non-genetic study of a partially overlapping sample (400 ADNI subjects), the right

caudate was 3.9% larger than the left in controls, on average, and 2.1% larger in MCI subjects—an asymmetry not found in Alzheimer's disease.[13] This same asymmetry is reported in most, but not all, large morphometric studies.[65–69] In the ADNI cohort, which focuses on elderly subjects, lower right caudate volume was associated with conversion from MCI to Alzheimer's disease, with baseline ratings of dementia severity, immediate and delayed logical memory scores, future decline over 1 year in mini-mental state examination scores, and tau and p-tau protein levels in the cerebrospinal fluid.[13] Taken together, these observations suggest that a depletion in caudate volume may be associated with deteriorating cognition, but cognitive associations may not be detectable in healthy subjects as other brain systems may compensate functionally for mild atrophy or developmental insufficiency. Future meta-analyses in even larger samples may be sufficiently powered to relate genetic differences in brain structure to observable differences in cognition or risk for the diseases in which the caudate is implicated.

Here, we demonstrate a replicated—thought not genome-wide significant—association in a sample that is much smaller in size than those used in some current genome-wide association studies.[56] This strongly suggests that MRI-based measures of brain structure are powerful, genetically informative tools with which to search the genome and may be used successfully to find genetic variants in multi-site genetic meta-analyses such as through the Enigma project (http://enigma.loni.ucla.edu). Our results highlight a region of the genome that may provide a stronger understanding of caudate neurobiology, brain structure in humans and predisposition for the development of psychiatric and neurological illness.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgments

Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: Abbott, AstraZeneca AB, Bayer Schering Pharma AG, Bristol-Myers Squibb, Eisai Global Clinical Development, Elan Corporation, Genentech, GE Healthcare, GlaxoSmithKline, Innogenetics, Johnson and Johnson, Eli Lilly and Co., Medpace, Inc., Merck and Co., Inc., Novartis AG, Pfizer Inc., F Hoffman-La Roche, Schering-Plough, Synarc, Inc., as well as non-profit partners the Alzheimer's Association and Alzheimer's Drug Discovery Foundation, with participation from the US Food and Drug Administration. Private sector contributions to ADNI are facilitated by the Foundation for the National Institutes of Health (www.fnih.org). The grantee organization is the Northern California Institute for Research and Education, and the study is coordinated by the Alzheimer's Disease Cooperative Study at the University of California, San Diego, CA, USA. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of California, Los Angeles, CA, USA. This research was also supported by NIH grants P30 AG010129, K01 AG030514, and the Dana Foundation. Algorithm development for this study was also funded by the NIA, NIBIB, NICHD, the National Library of Medicine, and the National Center for Research Resources (AG016570, EB01651, LM05639, RR019771, EB008432, EB008281 and EB007813, to PMT). The BTLS study was supported by the National Institute of Child Health and Human Development (R01 HD050735), and the National Health and Medical Research Council (NHMRC 486682), Australia. Genotyping was supported by NHMRC (389875). JLS was also funded by the ARCS foundation and the NIMH (1F31MH087061).

References

1 Peper JS, Brouwer RM, Boomsma DI, Kahn RS, Hulshoff Pol HE. Genetic influences on human brain structure: a review of brain imaging studies in twins. *Hum Brain Mapp* 2007; **28**: 464–473.
2 Kremen WS, Prom-Wormley E, Panizzon MS, Eyler LT, Fischl B, Neale MC *et al.* Genetic and environmental influences on the size of specific brain regions in midlife: the VETSA MRI study. *Neuroimage* 2010; **49**: 1213–1223.
3 Bertram L, McQueen MB, Mullin K, Blacker D, Tanzi RE. Systematic meta-analyses of Alzheimer disease genetic association studies: the AlzGene database. *Nat Genet* 2007; **39**: 17–23.
4 Shaw P, Lerch JP, Pruessner JC, Taylor KN, Rose AB, Greenstein D *et al.* Cortical morphology in children and adolescents with different apolipoprotein E gene polymorphisms: an observational study. *Lancet Neurol* 2007; **6**: 494–500.
5 Gottesman II, Gould TD. The endophenotype concept in psychiatry: etymology and strategic intentions. *Am J Psychiatry* 2003; **160**: 636–645.
6 Harris GJ, Pearlson GD, Peyser CE, Aylward EH, Roberts J, Barta PE *et al.* Putamen volume reduction on magnetic resonance imaging exceeds caudate changes in mild Huntington's disease. *Ann Neurol* 1992; **31**: 69–75.
7 Rosas HD, Goodman J, Chen YI, Jenkins BG, Kennedy DN, Makris N *et al.* Striatal volume loss in HD as measured by MRI and the influence of CAG repeat. *Neurology* 2001; **57**: 1025–1028.
8 Thomas M, Jankovic J. Neurodegenerative disease and iron storage in the brain. *Curr Opin Neurol* 2004; **17**: 437–442.
9 Curtis AR, Fey C, Morris CM, Bindoff LA, Ince PG, Chinnery PF *et al.* Mutation in the gene encoding ferritin light polypeptide causes dominant adult-onset basal ganglia disease. *Nat Genet* 2001; **28**: 350–354.
10 Appenzeller S, Schirmacher A, Halfter H, Baumer S, Pendziwiat M, Timmerman V *et al.* Autosomal-dominant striatal degeneration is caused by a mutation in the phosphodiesterase 8B gene. *Am J Hum Genet* 2010; **86**: 83–87.
11 Kuhlenbaumer G, Ludemann P, Schirmacher A, De Vriendt E, Hunermund G, Young P *et al.* Autosomal dominant striatal degeneration (ADSD): clinical description and mapping to 5q13-5q14. *Neurology* 2004; **62**: 2203–2208.
12 Sheline YI. Neuroimaging studies of mood disorder effects on the brain. *Biol Psychiatry* 2003; **54**: 338–352.
13 Madsen SK, Ho AJ, Hua X, Saharan PS, Toga AW, Jack Jr CR *et al.* 3D maps localize caudate nucleus atrophy in 400 AD, MCI, and healthy elderly subjects. *Neurobiol Aging* 2010; **31**: 1312–1325.

14 Castellanos FX, Giedd JN, Eckburg P, Marsh WL, Vaituzis AC, Kaysen D *et al.* Quantitative morphology of the caudate nucleus in attention deficit hyperactivity disorder. *Am J Psychiatry* 1994; **151**: 1791–1796.
15 Goldman AL, Pezawas L, Mattay VS, Fischl B, Verchinski BA, Zoltick B *et al.* Heritability of brain morphology related to schizophrenia: a large-scale automated magnetic resonance imaging segmentation study. *Biol Psychiatry* 2008; **63**: 475–483.
16 Wright IC, Rabe-Hesketh S, Woodruff PW, David AS, Murray RM, Bullmore ET. Meta-analysis of regional brain volumes in schizophrenia. *Am J Psychiatry* 2000; **157**: 16–25.
17 Lander ES. The new genomics: global views of biology. *Science* 1996; **274**: 536–539.
18 Hickie IB, Naismith SL, Ward PB, Scott EM, Mitchell PB, Schofield PR *et al.* Serotonin transporter gene status predicts caudate nucleus but not amygdala or hippocampal volumes in older persons with major depression. *J Affect Disord* 2007; **98**: 137–142.
19 Bartres-Faz D, Junque C, Serra-Grabulosa JM, Lopez-Alomar A, Moya A, Bargallo N *et al.* Dopamine DRD2 Taq I polymorphism associates with caudate nucleus volume and cognitive performance in memory impaired subjects. *Neuroreport* 2002; **13**: 1121–1125.
20 Durston S, Fossella JA, Casey BJ, Hulshoff Pol HE, Galvan A, Schnack HG *et al.* Differential effects of DRD4 and DAT1 genotype on fronto-striatal gray matter volumes in a sample of subjects with attention deficit hyperactivity disorder, their unaffected siblings, and controls. *Mol Psychiatry* 2005; **10**: 678–685.
21 Glatt CE, Freimer NB. Association analysis of candidate genes for neuropsychiatric disease: the perpetual campaign. *Trends Genet* 2002; **18**: 307–312.
22 Freimer N, Sabatti C. The use of pedigree, sib-pair and association studies of common diseases for genetic mapping and epidemiology. *Nat Genet* 2004; **36**: 1045–1051.
23 Stein JL, Hua X, Morra JH, Lee S, Hibar DP, Ho AJ *et al.* Genome-wide analysis reveals novel genes influencing temporal lobe structure with relevance to neurodegeneration in Alzheimer's disease. *Neuroimage* 2010; **51**: 542–554.
24 Petersen RC. Aging, mild cognitive impairment, and Alzheimer's disease. *Neurol Clin* 2000; **18**: 789–806.
25 Medland SE, Nyholt DR, Painter JN, McEvoy BP, McRae AF, Zhu G *et al.* Common variants in the trichohyalin gene are associated with straight hair in Europeans. *Am J Hum Genet* 2009; **85**: 750–755.
26 Mazziotta J, Toga A, Evans A, Fox P, Lancaster J, Zilles K *et al.* A probabilistic atlas and reference system for the human brain: International Consortium for Brain Mapping (ICBM). *Philos Trans R Soc Lond B Biol Sci* 2001; **356**: 1293–1322.
27 Morra JH, Tu Z, Apostolova LG, Green AE, Avedissian C, Madsen SK *et al.* Validation of a fully automated 3D hippocampal segmentation method using subjects with Alzheimer's disease mild cognitive impairment, and elderly controls. *Neuroimage* 2008; **43**: 59–68.
28 Looi JC, Lindberg O, Liberg B, Tatham V, Kumar R, Maller J *et al.* Volumetrics of the caudate nucleus: reliability and validity of a new manual tracing protocol. *Psychiatry Res* 2008; **163**: 279–288.
29 Duvernoy HM. *The Human Brain: Surface, Three-Dimensional Sectional Anatomy with MRI, and Blood Supply*. 2nd edn. Springer-Verlag Wien: New York, 1999.
30 Neale MC, Cardon LR. *Methodology for Genetic Studies of Twins and Families*. Kluwer Academic: Dordrecht, The Netherlands, 1992.
31 Chiang MC, Barysheva M, Shattuck DW, Lee AD, Madsen SK, Avedissian C *et al.* Genetics of brain fiber architecture and intellectual performance. *J Neurosci* 2009; **29**: 2212–2224.
32 Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA, Bender D *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007; **81**: 559–575.
33 Yu J, Pressoir G, Briggs WH, Vroh Bi I, Yamasaki M, Doebley JF *et al.* A unified mixed-model method for association mapping that accounts for multiple levels of relatedness. *Nat Genet* 2006; **38**: 203–208.
34 Kang HM, Zaitlen NA, Wade CM, Kirby A, Heckerman D, Daly MJ *et al.* Efficient control of population structure in model organism association mapping. *Genetics* 2008; **178**: 1709–1723.
35 Boomsma D, Busjahn A, Peltonen L. Classical twin studies and beyond. *Nat Rev Genet* 2002; **3**: 872–882.
36 Bacanu SA, Devlin B, Roeder K. The power of genomic control. *Am J Hum Genet* 2000; **66**: 1933–1944.
37 Morey RA, Petty CM, Xu Y, Hayes JP, Wagner II HR, Lewis DV *et al.* A comparison of automated segmentation and manual tracing for quantifying hippocampal and amygdala volumes. *Neuroimage* 2009; **45**: 855–866.
38 Morey RA, Selgrade ES, Wagner II HR, Huettel SA, Wang L, McCarthy G. Scan-rescan reliability of subcortical brain volumes derived from automated segmentation. *Hum Brain Mapp* 2010; **31**: 1751–1762.
39 Ryu S, Mahler J, Acampora D, Holzschuh J, Erhardt S, Omodei D *et al.* Orthopedia homeodomain protein is essential for diencephalic dopaminergic neuron development. *Curr Biol* 2007; **17**: 873–880.
40 Struyf J, Dobrin S, Page D. Combining gene expression, demographic and clinical data in modeling disease: a case study of bipolar disorder and schizophrenia. *BMC Genomics* 2008; **9**: 531.
41 Kobayashi T, Gamanuma M, Sasaki T, Yamashita Y, Yuasa K, Kotera J *et al.* Molecular comparison of rat cyclic nucleotide phosphodiesterase 8 family: unique expression of PDE8B in rat brain. *Gene* 2003; **319**: 21–31.
42 Nicola SM, Surmeier J, Malenka RC. Dopaminergic modulation of neuronal excitability in the striatum and nucleus accumbens. *Annu Rev Neurosci* 2000; **23**: 185–215.
43 Girault JA, Greengard P. The neurobiology of dopamine signaling. *Arch Neurol* 2004; **61**: 641–644.
44 Menniti FS, Faraci WS, Schmidt CJ. Phosphodiesterases in the CNS: targets for drug development. *Nat Rev Drug Discov* 2006; **5**: 660–670.
45 Wong ML, Whelan F, Deloukas P, Whittaker P, Delgado M, Cantor RM *et al.* Phosphodiesterase genes are associated with susceptibility to major depression and antidepressant treatment response. *Proc Natl Acad Sci USA* 2006; **103**: 15124–15129.
46 Perez-Torres S, Cortes R, Tolnay M, Probst A, Palacios JM, Mengod G. Alterations on phosphodiesterase type 7 and 8 isozyme mRNA expression in Alzheimer's disease brains examined by *in situ* hybridization. *Exp Neurol* 2003; **182**: 322–334.
47 Kathiresan S, Willer CJ, Peloso GM, Demissie S, Musunuru K, Schadt EE *et al.* Common variants at 30 loci contribute to polygenic dyslipidemia. *Nat Genet* 2009; **41**: 56–65.
48 Levey AI, Hersch SM, Rye DB, Sunahara RK, Niznik HB, Kitt CA *et al.* Localization of D1 and D2 dopamine receptors in brain with subtype-specific antibodies. *Proc Natl Acad Sci USA* 1993; **90**: 8861–8865.
49 Mozley LH, Gur RC, Mozley PD, Gur RE. Striatal dopamine transporters and cognitive functioning in healthy men and women. *Am J Psychiatry* 2001; **158**: 1492–1499.
50 Nieoullon A. Dopamine and the regulation of cognition and attention. *Prog Neurobiol* 2002; **67**: 53–83.
51 Ohata S, Kinoshita S, Aoki R, Tanaka H, Wada H, Tsuruoka-Kinoshita S *et al.* Neuroepithelial cells require fucosylated glycans to guide the migration of vagus motor neuron progenitors in the developing zebrafish hindbrain. *Development* 2009; **136**: 1653–1663.
52 Kong Y, Nan K, Yin Y. Identification and characterization of CAC1 as a novel CDK2-associated cullin. *Cell Cycle* 2009; **8**: 3544–3553.
53 Ling KH, Hewitt CA, Beissbarth T, Hyde L, Banerjee K, Cheah PS *et al.* Molecular networks involved in mouse cerebral corticogenesis and spatio-temporal regulation of Sox4 and Sox11 novel antisense transcripts revealed by transcriptome profiling. *Genome Biol* 2009; **10**: R104.
54 Zoubek RE, Hannappel E. Influence of the N terminus and the actin-binding motif of thymosin beta4 on its interaction with G-actin. *Ann N Y Acad Sci* 2007; **1112**: 435–441.
55 Pohjalainen T, Rinne JO, Nagren K, Lehikoinen P, Anttila K, Syvalahti EK *et al.* The A1 allele of the human D2 dopamine receptor gene predicts low D2 receptor availability in healthy volunteers. *Mol Psychiatry* 1998; **3**: 256–260.

56 Lango Allen H, Estrada K, Lettre G, Berndt SI, Weedon MN, Rivadeneira F *et al.* Hundreds of variants clustered in genomic loci and biological pathways affect human height. *Nature* 2010; **467**: 832–838.
57 Manolio TA, Collins FS, Cox NJ, Goldstein DB, Hindorff LA, Hunter DJ *et al.* Finding the missing heritability of complex diseases. *Nature* 2009; **461**: 747–753.
58 Hodgkinson CA, Enoch MA, Srivastava V, Cummins-Oman JS, Ferrier C, Iarikova P *et al.* Genome-wide association identifies candidate genes that influence the human electroencephalogram. *Proc Natl Acad Sci USA* 2010; **107**: 8695–8700.
59 Dick DM, Aliev F, Krueger RF, Edwards A, Agrawal A, Lynskey M *et al.* Genome-wide association study of conduct disorder symptomatology. *Mol Psychiatry* 2010 (in press).
60 Stein JL, Hua X, Lee S, Ho AJ, Leow AD, Toga AW *et al.* Voxelwise genome-wide association study (vGWAS). *Neuroimage* 2010; **53**: 1160–1174.
61 Vounou M, Nichols TE, Montana G. Discovering genetic associations with high-dimensional neuroimaging phenotypes: a sparse reduced-rank regression approach. *Neuroimage* 2010; **53**: 1147–1159.
62 Gao X, Becker LC, Becker DM, Starmer JD, Province MA. Avoiding the high Bonferroni penalty in genome-wide association studies. *Genet Epidemiol* 2010; **34**: 100–105.
63 Wellcome Trust Case Control Consortium. Genome-wide association study of 14 000 cases of seven common diseases and 3000 shared controls. *Nature* 2007; **447**: 661–678.
64 Purcell SM, Wray NR, Stone JL, Visscher PM, O'Donovan MC, Sullivan PF *et al.* Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature* 2009; **460**: 748–752.
65 Giedd JN, Snell JW, Lange N, Rajapakse JC, Casey BJ, Kozuch PL *et al.* Quantitative magnetic resonance imaging of human brain development: ages 4-18. *Cereb Cortex* 1996; **6**: 551–560.
66 Ifthikharuddin SF, Shrier DA, Numaguchi Y, Tang X, Ning R, Shibata DK *et al.* MR volumetric analysis of the human basal ganglia: normative data. *Acad Radiol* 2000; **7**: 627–634.
67 Larisch R, Meyer W, Klimke A, Kehren F, Vosberg H, Muller-Gartner HW. Left-right asymmetry of striatal dopamine D2 receptors. *Nucl Med Commun* 1998; **19**: 781–787.
68 Peterson BS, Riddle MA, Cohen DJ, Katz LD, Smith JC, Leckman JF. Human basal ganglia volume asymmetries on magnetic resonance images. *Magn Reson Imag* 1993; **11**: 493–498.
69 Yamashita K, Yoshiura T, Hiwatashi A, Noguchi T, Togao O, Takayama Y *et al.* Volumetric asymmetry and differential aging effect of the human caudate nucleus in normal individuals: a prospective MR imaging study. *J Neuroimag* 2009; **21**: 34–37.

Supplementary Information accompanies the paper on the Molecular Psychiatry website (http://www.nature.com/mp)