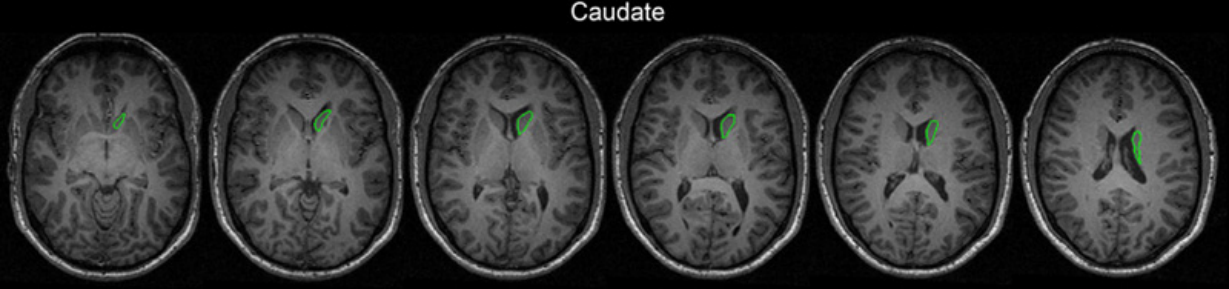
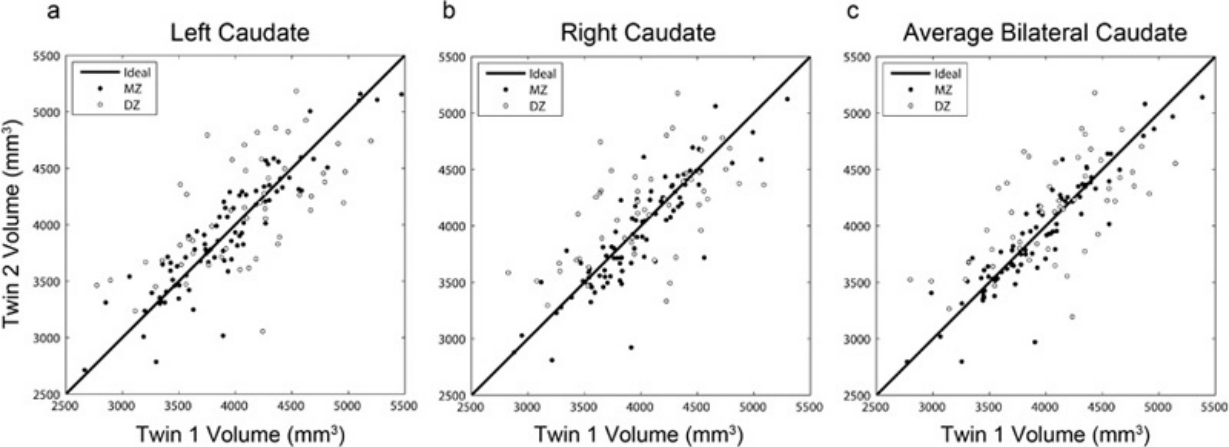
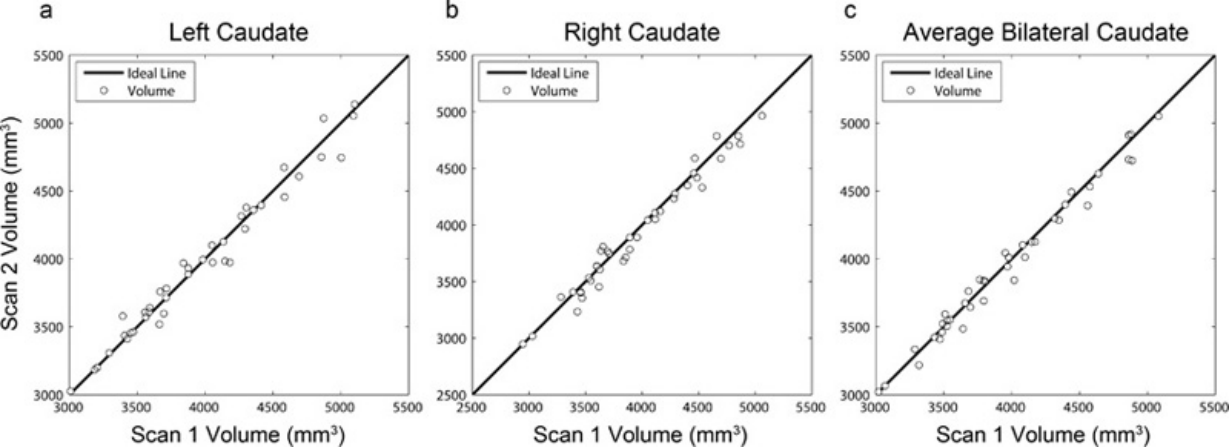
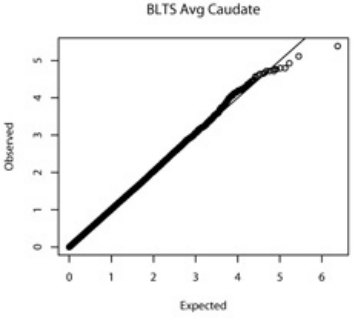
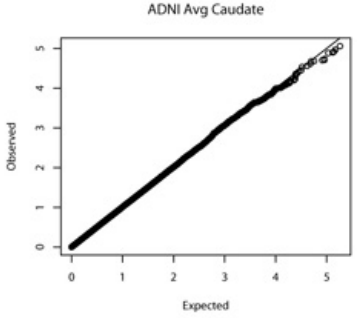
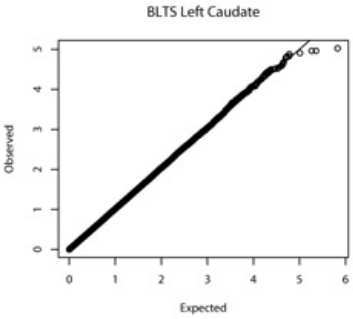

Caudate

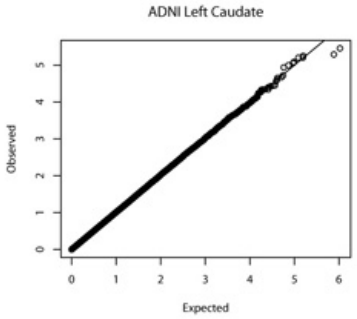

a
Left Caudate
b
Right Caudate
c
Average Bilateral Caudate
Ideal
MZ
DZ
Twin 2 Volume (mm³)
Twin 1 Volume (mm³)
2500
3000
3500
4000
4500
5000
5500

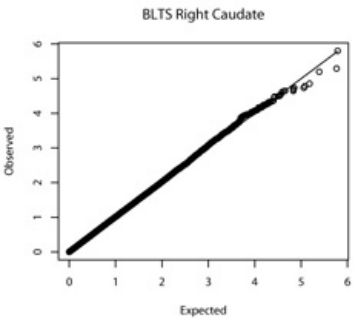

a
Left Caudate
b
Right Caudate
c
Average Bilateral Caudate
Ideal Line
Volume
Scan 2 Volume (mm³)
Scan 1 Volume (mm³)
5500
5000
4500
4000
3500
3000
2500

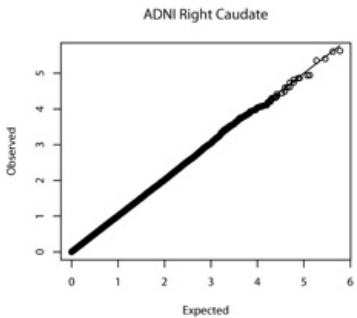

BLTS Avg Caudate
Observed
Expected

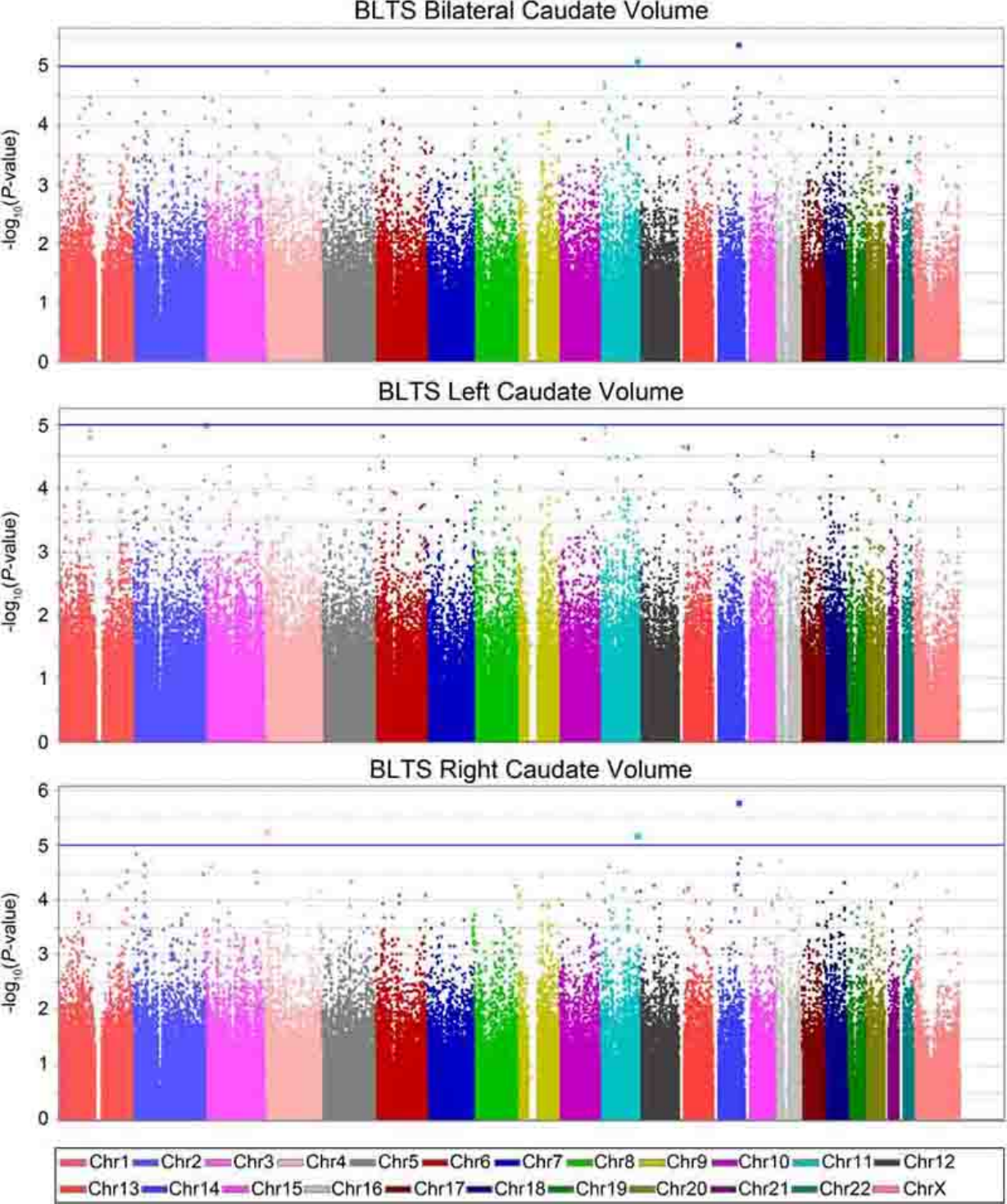

ADNI Avg Caudate
Observed
Expected


BLTS Left Caudate
Observed
Expected


ADNI Left Caudate
Observed
Expected


BLTS Right Caudate
Observed
Expected


ADNI Right Caudate
Observed
Expected


BLTS Bilateral Caudate Volume
BLTS Left Caudate Volume
BLTS Right Caudate Volume
-log10(P-value)
0
1
2
3
4
5
6
Chr1 Chr2 Chr3 Chr4 Chr5 Chr6 Chr7 Chr8 Chr9 Chr10 Chr11 Chr12
Chr13 Chr14 Chr15 Chr16 Chr17 Chr18 Chr19 Chr20 Chr21 Chr22 ChrX

Supplementary Information:

ADNI Cohort Details

Neuroimaging and genetic data were acquired from 818 subjects as part of the ADNI, a large five-year study launched in 2004 by the National Institute on Aging (NIA), the National Institute of Biomedical Imaging and Bioengineering (NIBIB), the Food and Drug Administration (FDA), private pharmaceutical companies, and non-profit organizations. The goal of ADNI is to determine biological markers of Alzheimer's disease through neuroimaging, genetics, neuropsychological tests and other measures to develop new treatments, monitor their effectiveness, and lessen the time of clinical trials. Subjects were recruited from 58 sites in the United States. The study was conducted according to the Good Clinical Practice guidelines, the Declaration of Helsinki, and U.S. 21 CFR Part 50 – Protection of Human Subjects, and Part 56 – Institutional Review Boards. Written informed consent was obtained from all participants before protocol-specific procedures were performed. All ADNI data are publicly available (at http://www.loni.ucla.edu/ADNI/). To avoid the effects of population stratification on genetic analysis (1), we only included Caucasian subjects (non-Hispanic; N=742) identified by self-report and confirmed by multi-dimensional scaling (MDS) analysis (2). This represented the vast majority of the ADNI subjects (91%). Eight subjects were removed due to poor delineation of the caudate (see below).

BLTS cohort details

MRI scanning of the BLTS cohort is ongoing (3) as part of a 5-year research project examining genetic effects on the brain with structural and functional MRI and DTI. 484 subjects had a genome-wide scan available; 475 of these also had MRI scans. 40 subjects underwent repeated imaging to verify the reliability of the structural imaging measures. All twins were screened to exclude pathology known to affect brain structure. No twins reported a history of significant head injury, neurological or psychiatric illness, substance abuse or dependence, or had a first-degree relative with a psychiatric disorder. The cohort is predominately Caucasian; ten individuals were excluded as ancestry outliers, identified by

principal component analysis of genome-wide SNP data. All subjects were right-handed as determined using 12 items from Annett's Handedness Questionnaire (5). Written informed consent was obtained from all subjects before protocol-specific procedures were performed. One subject was removed due to poor delineation of the caudate, when traces were reviewed for accuracy (see below).

Genotype Quality Control Exclusions

Individual markers were excluded from the analysis if they did not satisfy the following quality control criteria. In the ADNI cohort, genotype call rate < 95 % (42 670 SNPs removed), significant deviation from Hardy-Weinberg equilibrium at $P < 5.7\text{x}10^{-7}$ (871 markers removed), minor allele frequency (MAF) < 0.01 (60 867 SNPs removed), and a platform-specific recommended quality control score of < 0.15 (variable number of SNPs removed across subjects). In the BTLS cohort, SNPs were excluded based on a platform specific quality control score < 0.7 (47 148 SNPs removed), genotype call rate < 95% (8447 SNPs removed), significant deviation from Hardy-Weinberg equilibrium at $P < 1\text{x}10^{-6}$ (2841 SNPs removed), SNPs with MAF < 0.01 (33 347 SNPs removed). Filtering SNPs with significant deviation from Hardy-Weinberg equilibrium at $P < 5.7\text{x}10^{-7}$ was less conservative than the threshold of $P < 1\text{x}10^{-6}$ used in the BLTS cohort. We checked if the top SNPs in this study survived the less conservative thresholds from the ADNI cohort (**Supplementary Table 2**). All of the top SNPs identified in this study did not have significant deviations from Hardy-Weinberg equilibrium meeting either of these significance values. The BLTS HWE *P*-values include all subjects regardless of family status.

Image Acquisition and Pre-Processing

In ADNI, high-resolution structural brain MRI scans were acquired at 58 sites using 1.5 Tesla MRI scanners. A sagittal 3D MP-RAGE sequence was used, optimized for consistency across sites (6) (TR/TE = 2400/1000 ms; flip angle = 8°; FOV = 24 cm; final reconstructed voxel resolution = 0.9375 x 0.9375 x 1.2 mm^3). To adjust for global differences in brain positioning and scale, all subjects' scans were linearly

registered to the stereotaxic space defined by the International Consortium for Brain Mapping (ICBM-53) (7) using a 9-parameter (9P) transformation (three translations, three rotations, three scales) (8).

In the BLTS cohort, all subjects were imaged on the same scanner using structural whole-brain MRI at 4 Tesla (Bruker Medspec). T_1-weighted images were acquired with an inversion recovery rapid gradient echo sequence (TI/TR/TE = 700/1500/3.35 ms; flip angle = 8˚; slice thickness = 0.9 mm, with a 256^3 acquisition matrix). All images were initially corrected for intensity non-uniformity (9) within an automatically delineated mask of the brain (10). Images were then spatially normalized to the ICBM-152 template (7) using a 9P transformation that rotated and scaled each image to minimize a normalized mutual information cost function (11). Images were then resampled in the space of the template using sinc interpolation to yield 1 mm^3 isotropic voxels. In this way, each brain was globally matched in size and mutually aligned, but local differences in shape and size remained intact.

Manual Caudate Delineation

In the ADNI cohort, a representative training set of 21 MRI scans (7 AD, 7 MCI, and 7 controls) was randomly chosen for manual delineation. Likewise, in the BLTS cohort, an additional 20 unrelated subjects (all from DZ twin pairs; 9 male/11 female; mean age ± s.d. 22.85 ± 1.84) were randomly selected for manual delineation of subcortical structures. Manual delineation was completed on the images, after registration to the template, using BrainSuite 2.0 (28) by one rater in the ADNI cohort (SK) and by two independent raters in the BLTS cohort (SK and MK). Left and right manually traced binary caudate images were separately used to train an automated recognition algorithm to delineate the structure in all the other scans. The caudate was defined to include both the caudate head and body, excluding the caudate tail, which tends to be too small to be reliably delineated on conventional MRI. Tracing was started on the axial plane moving from inferior to superior. As the border between the nucleus accumbens and caudate does not have a strong intensity difference on standard T_1-weighted MRI, we defined the first slice containing the caudate as the one where the putamen and caudate were separated by the internal

capsule and were no longer connected by the nucleus accumbens. Tracing continued in superior planes excluding surrounding white matter and ventricles, but including the full gray matter of the caudate. The structure was re-edited in the coronal plane going from anterior to posterior, and then re-edited again in the axial plane to ensure smoothness. **Supplementary Figure 1** shows an example delineation in the BLTS cohort.

Automatic Caudate Delineation program

We used our method based on the AdaBoost and the auto-context algorithms (12). This method takes, as input, the binary training scans of one expert rater and the registered structural images of the manually delineated subjects. Caudates were automatically traced separately for each cohort, using only the training data from that cohort. The algorithm then extracts features from the MRI and tissue segmented images based on image intensity, position, image curvatures, image gradients, and Haar filters. All features are combined to estimate the probability that each voxel belongs to the caudate in each of the other scans. Voxels with greater than 0.5 probability are classified as being within the structure. To ensure validity of the algorithm, each automated structural delineation was visually checked by an experienced rater. Eight ADNI subjects one BLTS subject were removed due to poor segmentations of the caudate. Small islands of voxels that were disconnected from other voxels of the structure were automatically removed. Reliability of segmentations was calculated using intraclass correlation coefficients (ICC) of the volumes in repeated scans of the same subjects. ICC values were calculated using the *psy* package in the R statistical package (version 2.9.2; http://www.r-project.org/).

Heritability Methods

The theoretical weights of the contributing factors are computed that would be the most likely to give rise to the observed pattern of covariances. Normalized path coefficients ascribe the percentage of variance in the structure volume to each of the latent variables A, C, and E. To determine the most parsimonious model, nested sub-models containing only AE or CE sources of influence were compared to the fully

saturated ACE model by examining the difference in a statistic defined as minus two multiplied by the log likelihood, which is distributed under the null hypothesis as a chi-squared variable with one degree of freedom. Left, right, and averaged caudate volumes were analyzed independently.

Association Controlling for Genetic Relatedness

A symmetric $N \times N$ kinship matrix was constructed to describe how each subject was related to all the others. Kinship matrix coefficients were defined as 1 for the relationship of each subject to themselves; 1 for MZ twins in the same family; 0.5 for DZ twins and siblings in the same family; and 0 for a pair of unrelated subjects. Ancestry outliers were removed, so no additional modeling was used in the kinship matrix to adjust for population genetic structure between families. We then used a linear mixed effects model to estimate each SNP's association to caudate volume, after controlling for age, sex, and familial relatedness, according to the formula:

$$\mathbf{y} = \mathbf{X\beta} + \mathbf{Zu} + \mathbf{e}.$$

Here **y** is a vector whose components represent the caudate volume in each subject; **X** is a design matrix of fixed effects containing the additive genetic effect of the SNP, as well as sex and age, for each subject, and a constant term; **β** is a vector representing the fixed-effect regression coefficients; **Z** is the identity matrix; **u** is the random effect with $\mathrm{Var}(\mathbf{u}) = \sigma^2_g\mathbf{K}$, where **K** is the kinship matrix; and **e** is a matrix of residual effects with $\mathrm{Var}(\mathbf{e}) = \sigma^2_e\mathbf{I}$. The resulting *P* value, reported here, is based on the β parameter of the additive genetic effect. This analysis was performed using the Efficient Mixed-Model Association software (13) (EMMA; http://mouse.cs.ucla.edu/emma/), within the R statistical package. SNPs on the X chromosome were coded as 0,1,2 for females and 0,1 for males.

Within-Group Permutation to Control for Effects of Diagnostic Status

To ensure that diagnosis did not confound the analyses in the ADNI cohort, empirical *P*-values were generated through a permutation algorithm which randomly swaps phenotypes between individuals, but only within specified subgroups – in this case diagnosis group (AD, MCI, controls) (2). Any effect of group is then built into the null (reference) distribution formed by permutation tests. The residuals of each phenotype after regression on age and sex were input into the permutation test as the dependent variable. The linear regression analysis was repeatedly performed for each marker and the *t*-statistic for each marker was saved. This process was adaptively repeated to ensure stable *P*-values (up to 100 million replicates). SNPs that proved unlikely to become significant after a small number of permutations were no longer swapped for computational efficiency as high precision for non-associated (high) *P*-values is not desirable here. The number of times a test statistic met or exceeded the observed test statistic was divided by the total number of permutations performed for that SNP. The effect of diagnosis is therefore preserved in each permuted dataset. In this way, we were able to control for diagnosis in the analyses without losing power and without incurring multiple testing issues related to analyzing each group separately.

Meta-Analysis of Genetic Results

Meta-analysis of pooled results was conducted using an inverse variance-weighted method. Pooled β values were calculated using an average of the β coefficients from regression analyses, weighted by the inverse of the variance of the β coefficients. Pooled standard error was computed as the square root of the inverse of the sum of the weights. Combining the pooled β value and the pooled standard error yielded a test statistic, whose distribution, under the null hypothesis, approximates that of a χ^2 statistic with 1 degree of freedom.

References:

1. Lander ES, Schork NJ. Genetic dissection of complex traits. *Science* 1994 Sep 30; **265**(5181)**:** 2037-2048.

2. 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 Jun; **51**(2)**:** 542-554.

3. de Zubicaray GI, Chiang MC, McMahon KL, Shattuck DW, Toga AW, Martin NG *et al.* Meeting the Challenges of Neuroimaging Genetics. *Brain Imaging Behav* 2008 Dec 1; **2**(4)**:** 258-263.

4. Zhu G, Duffy DL, Eldridge A, Grace M, Mayne C, O'Gorman L *et al.* A major quantitative-trait locus for mole density is linked to the familial melanoma gene CDKN2A: a maximum-likelihood combined linkage and association analysis in twins and their sibs. *Am J Hum Genet* 1999 Aug; **65**(2)**:** 483-492.

5. Annett M. A classification of hand preference by association analysis. *Br J Psychol* 1970 Aug; **61**(3)**:** 303-321.

6. Jack CR, Jr., Bernstein MA, Fox NC, Thompson P, Alexander G, Harvey D *et al.* The Alzheimer's Disease Neuroimaging Initiative (ADNI): MRI methods. *J Magn Reson Imaging* 2008 Apr; **27**(4)**:** 685-691.

7. 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 Aug 29; **356**(1412)**:** 1293-1322.

8. Collins DL, Neelin P, Peters TM, Evans AC. Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space. *J Comput Assist Tomogr* 1994 Mar-Apr; **18**(2)**:** 192-205.

9. Sled JG, Zijdenbos AP, Evans AC. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. *IEEE Trans Med Imaging* 1998 Feb; **17**(1)**:** 87-97.

10. Smith SM. Fast robust automated brain extraction. *Hum Brain Mapp* 2002 Nov; **17**(3)**:** 143-155.

11. Jenkinson M, Bannister P, Brady M, Smith S. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage* 2002 Oct; **17**(2)**:** 825-841.

12. 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 Oct 15; **43**(1)**:** 59-68.

13. 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 Mar; **178**(3)**:** 1709-1723.

Figure Legends:

Supplementary Figure 1: Manual delineation of the caudate nucleus in a representative subject's MRI scan from the BTLS cohort. The caudate is outlined in green. Images run from inferior to superior.

Supplementary Figure 2: Caudate volumes in MZ and DZ twin pairs. MZ twins (*black dots*) are more similar in volumes than DZ twins (*open dots*). Siblings and unpaired twins are not shown here.

Supplementary Figure 3: Caudate volume estimates are highly reproducible in 40 BLTS subjects, scanned twice. Volumes for the later versus the earlier scan are plotted for the (**a**) left, (**b**) right, and (**c**) average bilateral caudate volumes (open circles). A diagonal line indicating perfect correspondence between two scans is also shown.

Supplementary Figure 4: Q-Q plots show the distribution of *P*-values in each sample vs. the *P*-values expected under the null hypothesis of no association (open circles). Observed *P*-values were ordered and $-\log_{10}$ transformed (*y*-axis). Expected *P*-values were random draws from a uniform distribution which were also ordered and $-\log_{10}$ transformed (*x*-axis). The black line shows the expected null distribution. No inflation of the *P*-values is observed (ADNI Avg Caudate: $\lambda = 1.012$; BLTS Avg Caudate: $\lambda = 1.017$; ADNI Left Caudate: $\lambda = 1.012$; BLTS Left Caudate: $\lambda = 1.016$; ADNI Right Caudate: $\lambda = 1.012$; BLTS Right Caudate: $\lambda = 1.018$).

Supplementary Figure 5: Manhattan plots show the significance of association of each SNP with caudate volume, from genome-wide association analysis conducted in the BLTS sample. 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 $< 1 \times 10^{-5}$.

				BLTS						ADNI							Pooled		
Chr	SNP	Position	Gene	A1	Freq	N	β	SE	*P*	A1	Freq	N	β	SE	*P*	*P* | *diag*	β	SE	*P*
Average Bilateral Caudates																			
14	rs1954196	86041768		T	0.813	464	182.7	39.2	4.08×10^{-6}	C	0.1955	734	17.7	36.12	0.624	0.8571	-74.3	26.6	5.14×10^{-3}
11	rs559254	122602143	*ASAM*	G	0.734	464	-164.2	36.1	6.85×10^{-6}	A	0.2159	734	1.82	36.55	0.960	1.00	84.0	25.7	1.07×10^{-3}
Left Caudate																			
3	rs17034704	1187944	*CNTN6*	T	0.920	464	-287.9	64.3	9.42×10^{-6}	C	0.08104	728	72.43	53.97	0.180	0.2063	161.5	41.3	9.37×10^{-5}
Right Caudate																			
14	rs1954196	86041768		T	0.813	464	187.3	38.5	1.61×10^{-6}	C	0.1955	734	22.25	39.32	0.572	0.75	-84.7	27.5	2.07×10^{-3}
4	rs10018002	6848057	*KIAA0232*	G	0.218	464	-170.4	36.9	4.93×10^{-6}	G	0.2265	733	-9.94	38.54	0.7965	1.00	-93.7	26.7	4.42×10^{-4}
11	rs559254	122602143	*ASAM*	G	0.734	464	-163.1	35.5	5.66×10^{-6}	A	0.2159	734	-4.51	39.66	0.9095	0.8571	88.5	26.5	8.16×10^{-4}

Supplementary Table 1: Most associated SNPs using the BLTS as the discovery cohort ($P < 1 \times 10^{-5}$) and ADNI as replication cohort. *Genes* close to the SNPs are displayed and cells empty if no gene is within approximately ±50kb. *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 beta 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.

SNP	ADNI HWE *P*-value	BLTS HWE *P*-value
rs9378688	0.7235	0.3015
rs9392374	0.7204	0.2944
rs9405550	0.8545	0.4065
rs4752194	0.3865	0.5243
rs4752195	0.3865	0.5243
rs1602517	0.73	0.7421
rs234937	0.8507	0.1243
rs163030	1.00	0.5147
rs163035	1.00	0.5753
rs335636	0.9411	0.5154
rs1299288	0.677	0.6939
rs5979778	0.0106	0.07355

Supplementary Table 2: *P*-values representing deviation from Hardy-Weinberg equilibrium at the most associated SNPs from the ADNI discovery cohort and BLTS replication cohort represented in **Table 2**. The BLTS HWE *P*-values include all subjects regardless of family status.