Amygdalar atrophy as the genetically mediated hub of limbic degeneration in Alzheimer's disease

Sourena Soheili-Nezhad[1]

Neda Jahanshad[1]

Afsaneh Tajer[2,3]

Sebastian Guelfi[4]

Reza Khosrowabadi[5]

Hamidreza Pouretemad[5]

Andrew J. Saykin[6,7,8]

Paul M. Thompson[1]

Mojtaba Zarei[2,3]

for the Alzheimer's Disease Neuroimaging Initiative[1]

[1]Imaging Genetics Center, USC Mark and Mary Stevens Neuroimaging and Informatics Institute, Keck School of Medicine of USC, Marina del Rey, CA 90292

[2]School of Cognitive Sciences, Institute for Research in Fundamental Sciences (IPM), Tehran, Iran

[3]Institute of Medical Science and Technology, Shahid Beheshti University, Tehran, Iran.

[4]Reta Lila Weston Research Laboratories, Department of Molecular Neuroscience, University College London (UCL) Institute of Neurology, London, UK

[5]Institute for Cognitive and Brain Sciences, Shahid Beheshti University, Tehran, Iran

[6]Center for Neuroimaging, Department of Radiology and Imaging Sciences, Indiana University School of Medicine, Indianapolis, IN, USA

[7]Indiana Alzheimer Disease Center, Indiana University School of Medicine, Indianapolis, IN, USA

[8]Department of Medical and Molecular Genetics, Indiana University School of Medicine, Indianapolis, IN, USA

Correspondence to:

Mojtaba Zarei

P.O. 19395-5746

Tehran, Iran

E-mail: mzarei@ipm.ir

[1] Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). 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 can be found at:

http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

Abstract

Mapping the spatiotemporal dynamics of brain atrophy in the Alzheimer's disease (AD) aids in discovering vulnerable brain networks. Various MRI-derived measures including hippocampal volume loss are commonly used as imaging biomarkers of AD progression. Multivariate methods such as independent component analysis (ICA) have recently shown promise in revealing more complex patterns of brain atrophy.

Here, we aimed to extract a data-driven signature of brain atrophy in AD and its heralding risk factor, the mild cognitive impairment (MCI) state. Structural brain changes were quantified in 1,892 MRI scans from 1,100 participants of the Alzheimer's Disease Neuroimaging Initiative (ADNI) by Tensor-Based Morphometry (TBM). Brain morphometry maps of this population were fed into an ICA-based feature extraction algorithm which revealed that AD is associated with a characteristic signature of brain degeneration spanning several areas of the limbic system. The highest probability of brain atrophy in AD was localized to amygdalae (pICA z-score>22), followed by hippocampus (pICA z-score>14), and the mammillary bodies (pICA z-score>9) together with their white-matter connections including fimbrial and fornical tracts as a unified MRI endophenotype. This endophenotype performed better than the hippocampal volume measure in discriminating both AD and MCI subjects from the cognitively normal group, and we evaluated the potential for this feature set to serve as a disease-related endophenotype in a whole-genome association study. A protein-altering polymorphism in the synaptic scaffold gene, *SHARPIN*, affected loss of volume in this endophenotype among 8,852,807 whole-genome sequencing-based variants (rs34173062; $p = 2.1 \times 10^{-10}$). Other suggestive but subgenome-wide signals mapped to *EPHA7* and *FRMD4A*, two known AD risk loci.

In conclusion, by unsupervised decomposition of brain morphometry maps, we have provided a unified imaging endophenotype of AD that closely maps to the brain's memory center. Observation of bilateral amygdalae as the most vulnerable limbic structures spotlights impairment of emotional and memory processing integration in early AD.

Keywords: Alzheimer's disease, Brain atrophy, Tensor-based morphometry, Independent component analysis, endophenotype, genome-wide association study, whole genome sequencing

Abbreviations: Alzheimer's disease (AD), Genome-wide association study (GWAS), Independent component analysis (ICA), Mild cognitive impairment (MCI), Medial temporal circuit (MTC), Single-nucleotide polymorphism (SNP), Tensor-based morphometry (TBM)

Introduction

Sporadic AD is a multifactorial disease predisposed by a strong (58-79%) genetic contribution[1]. However, even the most compelling genetic risk factor, *APOE4*, only explains 6% of disease predisposition variance, leaving the majority of its polygenic risk component largely unresolved by known loci[2].

An important caveat for discovering association of genetic risk factors with various human diseases is selection of optimal outcome measures, or phenotypes, for genome-wide association. Case-control genome-wide association studies (GWAS) identify disease risk genes by stratifying the study groups based on clinical diagnoses. Clinical diagnoses are simple tools that have been fundamentally devised to translate heterogeneous pathologies into simple categorical entities for guiding therapeutic approaches, and represent only the terminal outcomes of the complex molecular and cellular derangements of the disease. As a result, when used as the phenotype of interest in genome-wide research, clinical diagnoses may not give us an ultimately deep insight into the biological *pathways* of disease progression, for instance in the early subclinical stages of brain pathology.

Using *endophenotypes* to search for genetic risk factors of the disease process is thought to address this shortcoming. Endophenotypes are heritable measures, which are shown to be different in individuals with a disease than those without, and are also often disrupted in unaffected siblings of those with the disorder, yet to a lesser degree[3]. In particular, quantitative MRI is a useful modality for biomarker discovery and monitoring brain structure, function and connectivity in living humans. It is the primary objective of the emerging field of imaging genetics to identify the impact of genetic factors on such neuroimaging biomarkers that potentially inform on pathways of disease progression[4,5].

Various measures of brain atrophy are commonly used in AD research, including loss of hippocampal volume[6], reduced thickness of entorhinal cortex[7] and expansion of lateral ventricles[8]. MRI-derived features, such as volume of subcortical structures have also been used in imaging genetic studies[9-11]. In contrast to these *a priori* measures, an

unbiased whole-brain search is expected to provide complementary insight into unexpected patterns of genotype-phenotype linkage in various neuropsychiatric disorders. However, this new approach comes at high computational and methodological costs, since iterating statistical tests across the combined and high-dimensional space of the genome and neuroimage poses significant technical challenges[12], but offers critical insight into unseen disease mechanisms.

Reportedly, network-specific degenerative forces underpin selective vulnerability of brain circuits in various neuropsychiatric disorders and dementias[13]. Extracting these complex patterns from 3D brain morphometry maps may provide us with promising endophenotypes for monitoring disease progression and its genetic modifiers[14]. Here, we aimed to extract the most representative spatial fingerprint of brain atrophy in AD through a brain-wide approach. We quantified voxelwise maps of brain volume and atrophy by processing MRI data of the ADNI cohort, and subsequently decomposed brain atrophy sources by an unbiased, whole-brain computational search using spatial ICA. This analysis revealed a characteristic signature of brain atrophy in AD that connected several structures of the limbic system and their white-matter tracts with a network-like topology in 3D. Whole-genome sequencing variants were used in conjunction with this new endophenotype, which showed that a missense variant in the postsynaptic scaffold protein *SHARPIN* significantly ($p=2.1\times10^{10}$) correlated with the derived measure of limbic system integrity across the study population. Our findings show the promise of data-driven endophenotypes as sensitive disease markers and reveal that bilateral amygdalae are probably the most vulnerable structures of the limbic system in early AD and MCI.

Results

Tensor-based morphometry shows spatial patterns of brain atrophy in aging and cognitive decline that map to known anatomical networks

The ADNI participants in this work included 1,100 individuals (age: 74.5±7.1y) with baseline T1-weighted structural brain MRI scans. Cognitive status of participants included normal, MCI and probable AD subjects based on NINCDS/ADRDA criteria (Table 1). Among these individuals, longitudinal MRI data was also available in a

subpopulation of 1,039 subjects who underwent follow-up scans at 1.07±0.08 year (n=1,009 subjects) and/or 2.07±0.11 year intervals (n=883 subjects).

Table 1. Cognitive status of the study population

	NL	*MCI*	*AD*	*Total*
Cross-sectional	383	456	361	1,100 (491 women)
Longitudinal	269	422	348	1,039 (457 women)

NL: Cognitively normal; MCI: Mild cognitive impairment; AD: Alzheimer's disease

We used Tensor-Based Morphometry (TBM) to track localized brain volume changes at MRI voxel resolution[15] which also formed the basis of subsequent multivariate analyses in this work (Figure S1). Jacobian determinant maps were produced by this analysis, which are 3D scalar fields encoding *cross-sectional* brain volume differences in each study subject (i.e. regional brain shrinkage or expansion) in relation to the average study brain template[16]. For those study subjects who underwent serial MRI scans, a longitudinal TBM workflow was implemented to track brain atrophy in course of disease progression using symmetric alignment of baseline and follow-up MRIs[17,18]. Subtle anatomical changes across the inter-scan intervals were calculated, resulting in production of *longitudinal* Jacobian maps which encoded voxelwise dynamics of brain atrophy progression in each study subject (Fig. 1 and supplementary methods).

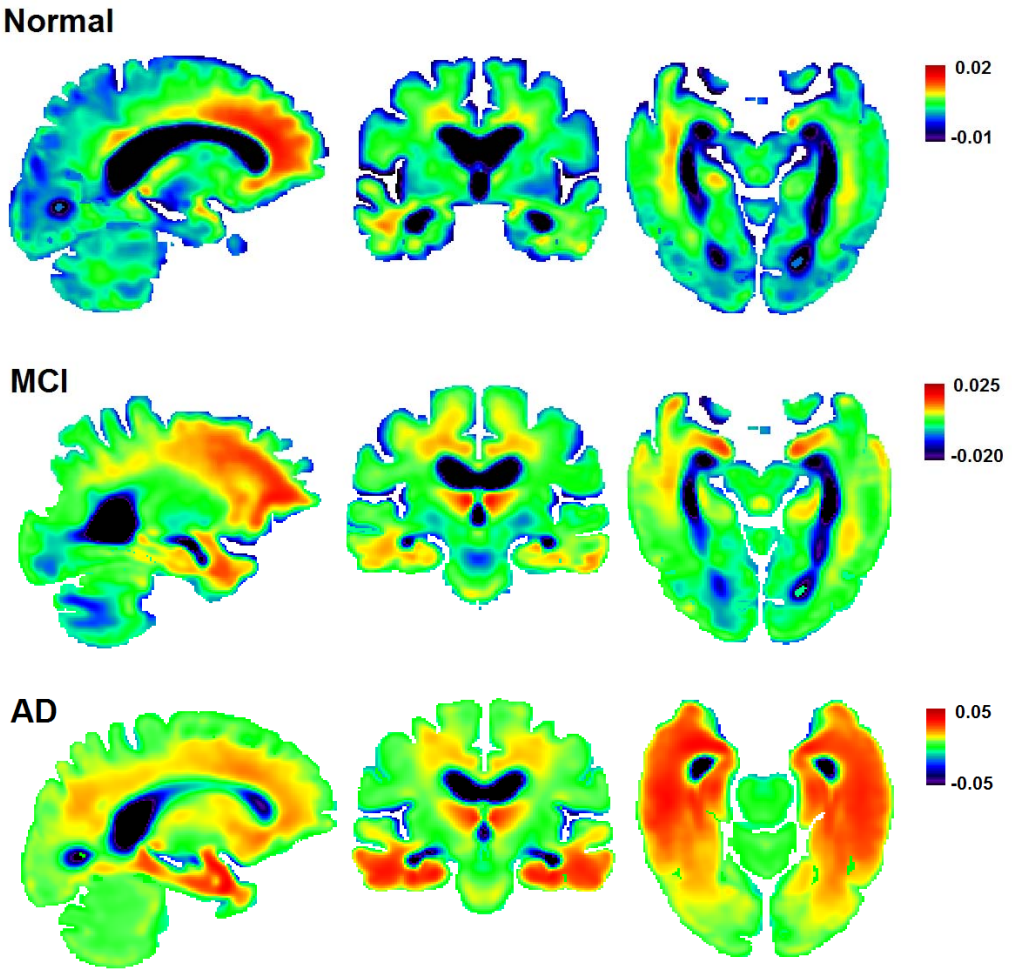

Figure 1. Jacobian maps of brain atrophy. Average longitudinal Jacobian maps of the study diagnosis groups. MCI: Mild cognitive impairment; AD: Alzheimer's disease.

Spatial patterns of regional brain volume differences and atrophy rates in AD

By slightly modifying the ICA-based source-based morphometry approach, we extracted multivariate features in the Jacobian morphometry maps to identify reproducible and independent sources of brain atrophy[19]. To this aim, cross-sectional (1,068,867 voxels × 1,100 subjects) and longitudinal (1,068,867 voxels × 1,039 subjects) Jacobian maps were spatially decomposed into 1,348 independent sources of brain volume change and atrophy by probabilistic

ICA[20]. These data-driven 3D sources inform on brain areas with correlated patterns of volumetric variability and atrophy across the population, and most of them localized to known anatomical structures (Fig. 2).

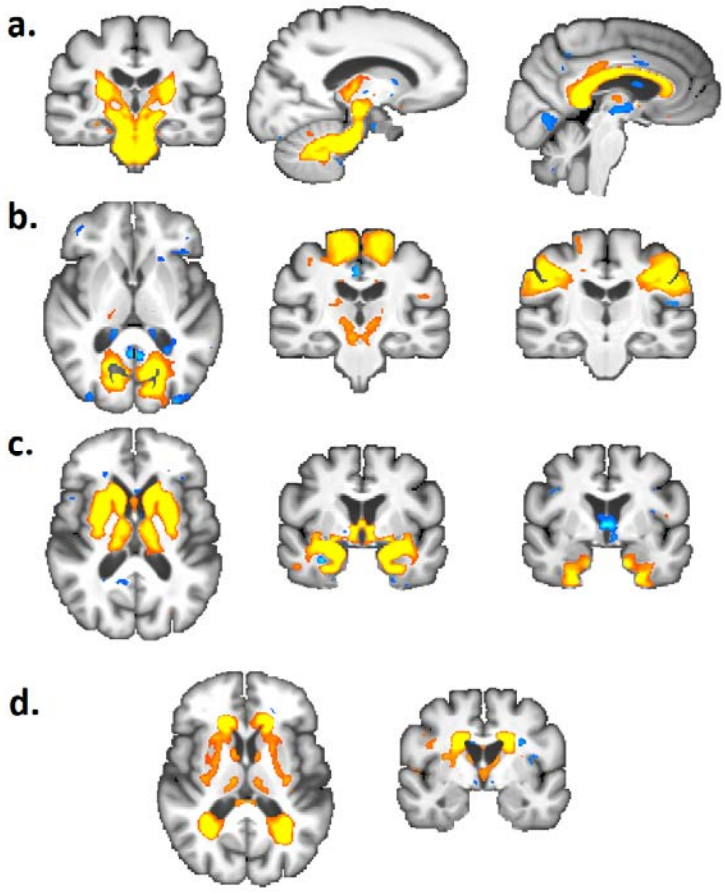

Figure 2: **Spatially-independent components of brain morphometry in the study population.** A number of cross-sectional brain volume sources decomposed by ICA are shown. Components spanned the whole extent of the brain, including white-matter areas (**a**), cortical gyri and sulci (**b**), as well as subcortical structures and the limbic system (**c**). Other features were also captured by ICA, including periventricular white matter lesions (**d**).

We aimed to determine which subset of these imaging features best represented the spatial impact of AD on brain structure. To this aim, all of the 1,348 components were used together as predictor variables in regression models to discriminate AD patients from cognitively normal subjects, after correcting for the confounding effects of age and sex. L1-regularization (LASSO) was used to include as few predictor components as possible and simplify the model[21]. This analysis showed that AD patients could be predicted by reduced volume in 26 ICA components using

cross-sectional MRI data, or 30 ICA components using longitudinal MRI data, with respective classification accuracies of 86% and 87% (Fig 3 and S3).

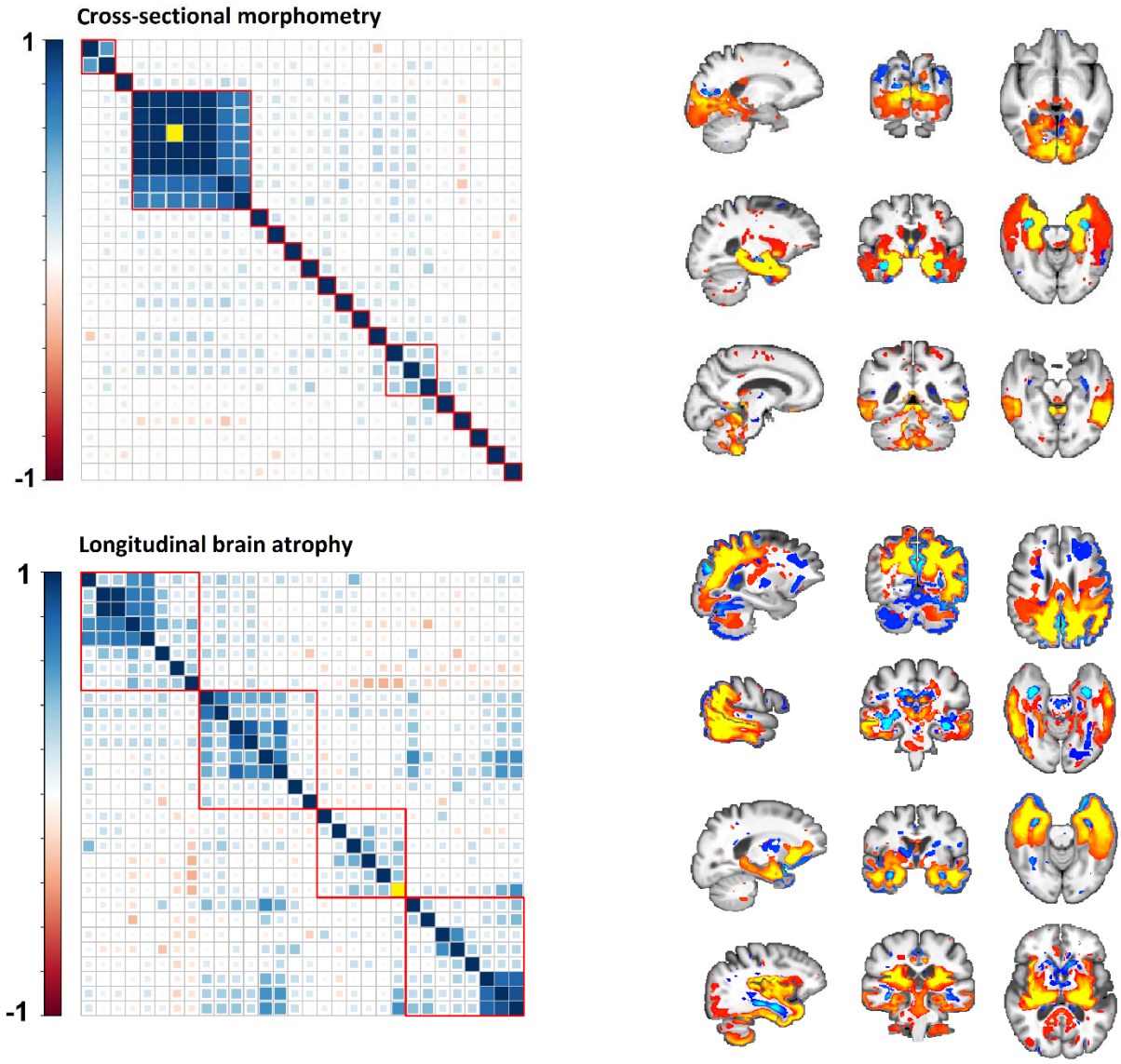

Figure 3. AD-predicting brain volume and atrophy sources: For visualization of the AD-specific brain components, their cross-correlation matrices were constructed and reordered to bring spatially-correlated elements close to the diagonal. Hierarchical clustering was then used to group the components into similar clusters. Cluster-wise ICA z-score maps are depicted in right (red-yellow: atrophy, blue: expansion). The strongest AD discriminator

in both phases was the *medial temporal circuit* (MTC), which is plotted as the yellow diagonal element in both matrices.

The AD-predicting components in cross-sectional phase of the study clustered to brain regions spanning: precuneus, occipital lobes and pulvinar (Fig. 3A); hippocampus, amygdala, parahippocampal gyrus, fornix, mammillary bodies, and uncinate fasciculus (Fig. 3B); as well as lateral temporal lobes and the cerebellum (Fig. 3C). The AD-predicting components in longitudinal phase of the study clustered to: precuneus, cuneus and occipitoparietal lobes (Fig. 3D); thalamus, temporal lobes and its association areas (Figure 3E); medial temporal lobes, uncinate fasciculus and connections to the orbitofrontal cortex (Fig. 3F); as well as insula, basal ganglia and cerebellum (Fig. 3G). Overall, both cross-sectional and longitudinal neuroimaging of AD revealed a strongly reproducible focus of brain atrophy in the medial temporal lobe and its white-matter connections to the limbic system, with a secondary impact observed in parietal lobes, precuneus, cuneus, subcortical grey, and thalamus.

The memory limbic system is the most vulnerable focus of brain degeneration in AD and MCI: a unified endophenotype of early disease impact

Of the 1,348 spatially-independent components of brain volume and atrophy decomposed by 29 rounds of ICA, the most prominent imaging predictor of AD was a unique spatial source which obtained the top odds-ratio rank in both cross-sectional and longitudinal phases in discriminating AD from cognitively-normal subjects. This atrophy source spanned voxels of bilateral amygdalae, hippocampi, entorhinal cortex, insula, mammillary bodies, and fornical tracts with a connected and network-like topology in 3D (Fig. 4 and supplementary video). The maximal focus of atrophy impact within this component mapped to amygdalae, and further extended to other nodes of the memory limbic system in a bilaterally-symmetric probability gradient. This atrophy source, which will be referred to as the *medial temporal circuit* (MTC) in this manuscript, almost identically (r^2>0.94) replicated across five independent ICA runs, reflecting the strong share of variance between its voxels. MTC was also the most consistent predictor of subjects with MCI, as it obtained the top odds-ratio rank in discriminating MCI subjects from the cognitively normal group in cross-sectional MRI, and second in longitudinal MRI (Fig. 4). Of note, MTC was significantly, but only moderately correlated with the hippocampal volume parameter (r^2=0.65, $p<10^{-15}$).

Evaluating the statistical power of this new endophenotype for AD discrimination is subject to circularity and may be inflated, but this is not the case for discriminating the independent MCI group. Compared to the widely-used hippocampal volume parameter, we observed higher power for atrophy in the MTC endophenotype in group-wise discrimination of MCI individuals from the cognitively normal group (MTC t-statistic=4.4, $p=1.2\times10^{-5}$ vs. hippocampal t-statistic=3.2, p=0.002). For completeness, AD discrimination suggested a higher discriminatory power compared to hippocampal volume, as expected (MTC t-statistic=16.8, $p<10^{-15}$ vs. hippocampal t-statistic=4.4, $p=1.1\times10^{-5}$).

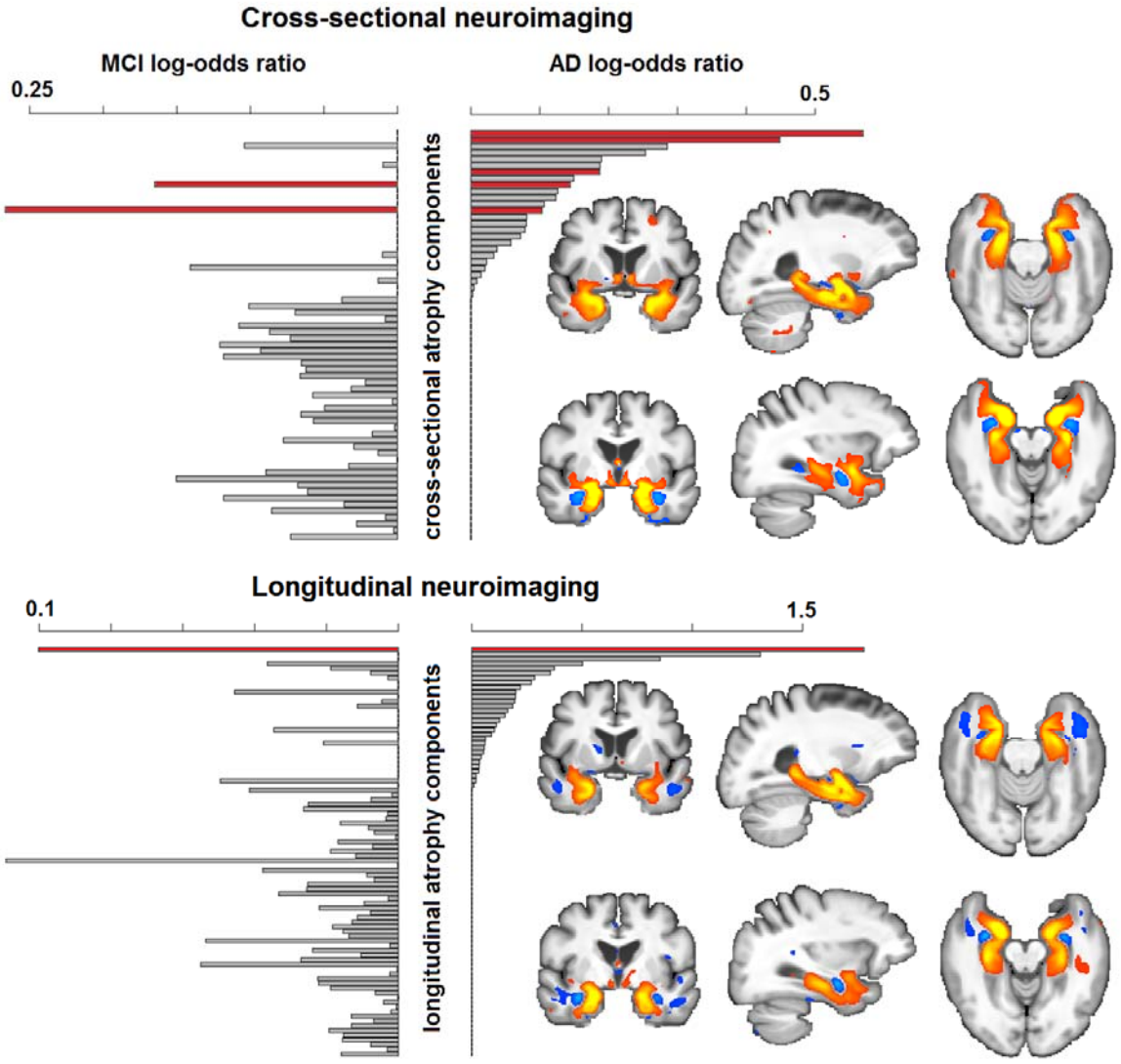

Figure 4. Contribution of brain atrophy components in predicting AD or MCI diagnoses. Each bar represents a brain morphometry component which either predicted MCI subjects (left plots) and/or AD patients (right plots) from cognitively normal subjects in group discriminative regression models. Bar lengths encode regression betas (log-odds ratios) reflecting importance of each component in diagnosis prediction. Logistic regressions in cross-sectional (top) and longitudinal (bottom) neuroimaging both replicated the MTC atrophy endophenotype as the most consistent predictor of AD and MCI, which is plotted as red bars. Spatial map of this component is previewed in right (ICA z-score; red-yellow: atrophy, blue: expansion) and 3D rendered in supplementary video.

The MRI endophenotype of medial temporal circuit degeneration is correlated with cognitive decline in aging humans

Cognitive performance of participants, as measured by ADAS-cog, was significantly correlated with the MTC endophenotype at its baseline volume in AD ($p=5.4\times10^{-8}$) and MCI ($p=9.6\times10^{-7}$), adjusting for age and sex confounds. Similarly, longitudinal atrophy rate of MTC showed significant associations with ADAS-cog in AD ($p=0.04$) and MCI subjects ($p=0.002$) adjusted for age and sex confounds. Older age ($p=2.6\times10^{-5}$) and female sex ($p=7.5\times10^{-5}$) were significantly associated with faster atrophy in this ICA endophenotype. Compared to the cognitively-normal group, the annual atrophy rate in the MTC was increased by 2.1-fold in MCI and 5.1-fold in AD patients.

APOE4 is correlated with degeneration of bilateral amygdalae and the MTC endophenotype

Dosage of the *APOE4* allele was significantly correlated with smaller baseline volume ($p=0.006$) and faster atrophy rate ($p=5.0\times10^{-4}$) in the MTC endophenotype, respectively explaining 1.0% and 1.8% of its variance in AD patients. To further elucidate the brain-wide impact of *APOE4*, we conducted a standard non-parametric voxelwise search[22], which showed that longitudinal brain atrophy is modified by *APOE4* dosage in temporal lobes of AD patients. Intriguingly, the *APOE4* influence in cross-sectional morphometry was only significant in bilateral amygdalar voxels, suggesting maximal statistical power for observing a significant effect in these structures.

Common genetic variants modify limbic volume in late life

Observation of the MTC component as the signature of brain atrophy in AD compelled us to study its role as a data-driven endophenotype of limbic integrity in unbiased imaging genetics. Whole-genome sequencing data were available at an average depth of 30-40x coverage in 808 ADNI participants[23], encompassing 40,631,363 single-nucleotide polymorphisms (SNPs). We searched for common variants which were correlated with baseline MTC volume in all study diagnosis groups with mean age (±SD) of 74.0±7.1 years, including 226 normal, 402 MCI and 180 AD subjects, and 363 of whom were women. The top variant observed, rs34173062, changes an amino acid in

the synaptic protein *SHARPIN* and survived the genome-wide Bonferroni correction ($p=2.1\times10^{-10}$, Table 2 and Fig. 5). RegulomeDB suggests that this polymorphism likely affects transcription factor binding by mapping to known binding motifs and matching DNase hypersensitivity footprint[24]. Another top SNP, rs3778470 ($p=1.5\times10^{-7}$), is an intronic variant in *EPHA7*, coding for an Ephrin receptor with roles in axon guidance and dendritic spine development. Upon follow-up analyses, we found this variant regulated hippocampal expression of *EPHA7* in the UKBEC database of normal human brains[25], and the atrophy-increasing allele decreased hippocampal expression of this gene (p=0.03). Another polymorphism, rs28439901, maps to *FRMD4A*, a gene previously shown to modify AD risk[26]. Another hit, rs149101079, maps to an open-reading frame coding for *VEST-1*, a protein-coding transcript of unknown function which is selectively expressed in brain and hippocampus (supplementary figure S5). This polymorphism is also suggestively correlated with amygdalar volume in large-scale ENIGMA-2 meta-analysis of 30,717 humans (p=0.03, supplementary figure S5).

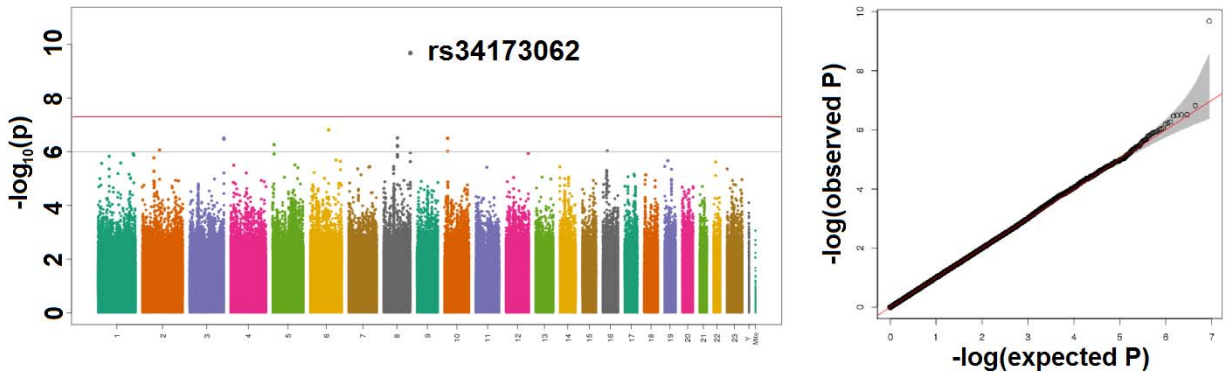

Figure 5. Manhattan and quantile-quantile plot of medial temporal circuit GWAS. A common non-synonymous variant in the *SHARPIN* gene passed the genome-wide threshold.

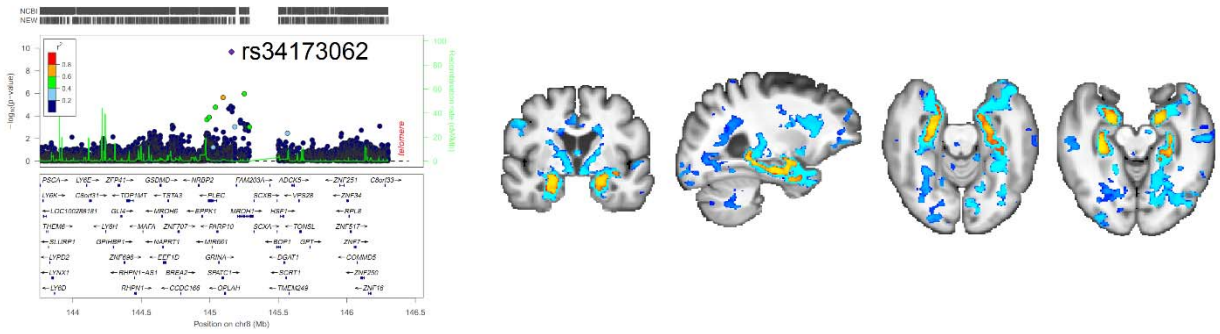

Figure 6. Regional association Plot of the *SHARPIN* locus and correlation of hit SNP with voxel-wise brain volume. The voxel-wise correlation map was calculated by FSL *Randomise* using 50,000 permutations. Blue: uncorrected p<0.05; red-yellow: corrected p<0.05.

Table 2. GWAS results

SNP	Chr	Position[a]	A1	A2[b]	Freq	β[c]	Gene[d]	p-value
rs56112946	3	197077194	C	T	0.03	-0.15	-	3.1×10^{-7}
rs3778470	6	94075684	G	A	0.12	-0.15	EPHA7 (intronic)	1.5×10^{-7}
rs149101079	8	69347181	G	A	0.04	-0.15	C8orf34	3.1×10^{-7}
rs34173062	8	145158607	G	A	0.10	-0.19	SHARPIN (missense)	2.1×10^{-10}
rs28439901	10	14494941	G	A	0.08	-0.15	FRMD4A (intronic)	3.2×10^{-7}

[a] hg19
[b] effect
[c] calculated for variance-normalized phenotype; negative: more atrophy
[d] Ensembl 81 coordinates

Discussions

ICA reveals limbic atrophy as the dominant MRI endophenotype of AD

Our blind ICA approach identified a unified measure of brain degeneration in AD by combining atrophy of several structures engaged in memory mechanisms. In contrast to hippocampal volumetry which is a hypothesis-driven segmentation technique, source-based morphometry can model hidden atrophy patterns with “soft” boundaries and spatial overlap, since contribution of each brain voxel is weighted by taking into account the data covariance

structure within each atrophy source. As a result, the robust share of variance observed in the atrophy of amygdala, hippocampus and the mammillary bodies together with their interconnecting white mater tracts indicates that this unified endophenotype may reveal the higher-order pattern of brain circuit vulnerability in AD. Importantly, although the ICA endophenotype fully embraced bilateral hippocampi, it was only modestly correlated with this commonly used AD endophenotype (r^2=0.65) due to broad inclusion of other vulnerable memory-related structures. The higher discriminatory power of the new endophenotype in predicting MCI, which is an established risk factor for future AD conversion, may show added information of data-driven endophenotypes in monitoring early brain changes in prodromal disease compared to hippocampal volumetry.

The ICA probability map of limbic atrophy clearly demonstrated that bilateral amygdalae are the most vulnerable foci of brain degeneration in AD and MCI. In line, the most consistent effect of the *APOE4* allele was also observed in bilateral amygdalar voxels. Atrophy and functional dysconnectivity of amygdala have been recently shown by ROI studies of early AD[27,28]. In close association with hippocampus, amygdala functionally integrates emotional and memory processing mechanisms. Amygdala is specifically engaged in processing of stimuli with emotional valence[29], and early emotional processing impairment is an emerging topic in AD research[30,31].

Our cluster analysis revealed that AD is associated with a secondary focus of atrophy in precuneus, cuneus and associated parieto-occipital lobes, structures which represent the major nodes of the default mode network. Correlation analysis showed that this posterior focus of brain atrophy was statistically uncorrelated with limbic degeneration, and our preliminary findings indicate that it is more prominent in earlier-onset AD cases and heralds a more aggressive clinical course.

Our study is not without limitation. While the primary objective of this work was identifying a sensitive imaging endophenotype of AD for application in imaging genetics, a whole genome-sequencing-based association of a single cohort of N<1000, cannot, on its own, track the *true* genetic architecture of brain degeneration. Nevertheless, we believe that some of our genetic findings are noteworthy, and could benefit from a replication attempt. The top GWAS variant of our study is a common amino acid-changing SNP in the *SHARPIN* gene coding for a SHANK scaffold protein that interfaces the postsynaptic density and neurotransmitter receptors with the synaptic actin cytoskeleton[32]. This variant is not genotyped or tagged by most of the genotyping chips used in the previous studies

of AD, and its risk allele has a substantially high frequency in the 1000GP cohort (MAF=~0.03-0.15). A sub-genome-wide intronic variant of brain atrophy mapped to *EPHA7*, a novel AD risk gene[33,34] that engages in neurite guidance and synaptic function[35]. *FRMD4A*, another subgenome-wide signal of our study, is a known risk locus for sporadic AD[26]. Taken together, these preliminary findings suggest that genome-wide association of disease-specific imaging endophenotypes may enlighten molecular pathways of disease predisposition.

In conclusion, we provide a data-driven MRI measure for early degeneration of the limbic system in AD that may aid in prognostic stratification of early cognitive decline in aging and prove useful in future clinical trials.

Acknowledgements

Authors are thankful to Professor John Hardy and Professor Charles DeCarli for their useful comments. This work was supported by NIH Big Data to Knowledge (BD2K) grant U54 EB20403. Data collection and sharing for this project was funded by the Alzheimer's Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging, the National Institute of Biomedical Imaging and Bioengineering, and through generous contributions from the following: AbbVie, Alzheimer’s Association; Alzheimer’s Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd. And its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; MesoScale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research is providing funds to support ADNI clinical sites in Canada. Private sector contributions 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. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California. Authors declare no conflicts of interest.

Supplementary methods

ADNI participants

Imaging, whole-genome sequencing and clinical data used in preparation of this manuscript were obtained from the ADNI database (adni.loni.usc.edu). ADNI is a multi-center initiative led by principal investigator Michael W. Weiner, MD, VA Medical Center and University of California, San Francisco, and enrolls subjects with normal cognition, MCI and AD (supplementary section). ADNI protocols are described in detail elsewhere (http://www.adni-info.org/pdfs/adni_protocol_9_19_08.pdf). Normal individuals have MMSE scores between 24-30, CDR of zero, non-depressed, non-demented and with no sign of cognitive impairment. Mild cognitive impairment subjects have MMSE scores between 24-30, memory complaint with objective memory loss measured by education-adjusted scores on Wechsler Memory Scale-Revised (WMS-R) logical memory II, a CDR of 0.5 and absence of significant impairment in other cognitive domains and absence of dementia. Alzheimer's disease cohort had MMSE score between 20 and 26, CDR of 0.5 or 1.0 and fulfilment of NINCDS/ADRDA criteria for probable Alzheimer's disease. ADNI is a longitudinal study and subjects are continuously assessed for cognitive changes, and may possess different diagnosis labels at different study sessions. To elucidate mild and earliest changes in course of AD and MCI, we considered the "worst" diagnosis label of each subject across all clinical sessions to stratify the cohort into three diagnosis groups (AD, MCI, NL). Therefore, AD and MCI subgroups of this study are expected to suffer milder disease than previous ADNI works on average.

MRI preprocessing and construction of common study template

Structural MRI was performed using GE, Philips or Siemens scanners. Structural T1-weigthed MRI volumes of the ADNI subjects were preprocessed for gradient non-uniformity and intensity inhomogeneity by the ADNI imaging core using gradient unwrapping, B1 field and N3 bias correction algorithms[36]. All volumes were recursively registered to an average study space using ANTS (*buildtemplateparallel* script) and a minimum-deformation template was created[37].

Longitudinal atrophy estimation

Brain-skull cropping

Pre-processed T1-weighted volumes of subjects who underwent serial MRI scans were used for longitudinal morphometry. FLIRT[38] with 12 degrees of freedom was used to linearly transform a dilated mask in the common space of the study template into each subject's native MRI space and obtain "cleaned" volumes in all subjects, which included brain and adjacent skull. This common stripping was performed to ensure that the most informative structures would be conserved for driving registration of serial MRI timepoints[39], and anatomical inconsistencies (e.g. soft tissue changes) would be eliminated.

Halfway space registration and longitudinal brain atrophy estimation

Basic steps of our longitudinal TBM workflow have been extensively validated before[17,40]. Cleaned T1-weighted volumes of each subject at $timepoints_{1,2}$ and $timepoints_{1,3}$ were separately registered to halfway spaces ($t_{1,2\text{-}half}$, $t_{1,3\text{-}half}$) using 9 degree-of-freedom FLIRT with normalized-correlation cost function and FSL *midtrans* utility. Halfway space registration prevents bias stemming from arbitrary selection of an MRI timepoint as source or target in registration, and has been implemented in statistical neuroimage analysis software, e.g. SIENA[41]. Nine degree-of-freedom registration removes global brain volume factor and has been shown to be comparable to phantom-based correction in ADNI[42]. In the halfway space, differential MRI bias correction[43] was performed on source and destination imaging volumes using *BtkDifferentialBiasCorrection*, part of Biomechanical toolkit[44]. The calculated bias-correcting fields were transformed to native MRI spaces. This round of halfway space registration was used to calculate an identical skull-striping mask using logical AND for similar cropping of source and destination MRI volumes. After obtaining bias-corrected and identically-masked images in native scan spaces, another iteration of halfway space registration was carried out using the same methods. Final transformation matrices were subsequently converted to ITK format using *c3d_affine_tool* (www.itksnap.org[45]), and used to initialize a non-linear registration of source and destination timepoints by SyN. Parameters of longitudinal SyN registration included histogram matching, a course-to-fine resolution scheme with 20×10 gradient descent iterations, gradient step size of 0.25, an update field variance of 3, a total field variance of 0.5,and a cross-correlation metric with a search radius of 4 voxels.

Independent Component Analysis

In the cross-sectional phase of study, fifteen ICA runs at a broad range of dimensions, spanning 8, 12, 16, 20, 24, 28, 32, 40, 48, 56, 64, 72, 80, 88 and 96 were used to decompose Jacobian maps into independent sources. Each resulting ICA component is represented by a spatial map which is a 3D imaging feature common across the study population, together with a mixing profile vector which linearly reflects the subject-specific brain volume parameter in the corresponding 3D source.

In order to annualize subject-wise atrophy rate estimates as reflected in the columns of the ICA mixing matrix, these values were centered on the *absolute* amount of brain atrophy in each component by spatial regression (*fsl_glm*) using the model:

$$[(1-M)/V]_{(1 \times n\ voxels)} = R_{(1 \times n\ components)} \times C_{(n\ components \times n\ voxels)} + \varepsilon$$

In which M and V denote one-dimensional-collapsed versions of 3D voxelwise Jacobian mean and variance across the study population, C denotes the matrix of spatially-independent components decomposed by ICA, and R is the vector of component-wise centering values to be calculated (i.e. regression betas).

Validating linearity of the atrophy rate estimates

Linear evolution of brain atrophy in temporal domain with a non-significant intercept at timepoint zero is a sign of biological plausibility in longitudinal TBM[40]. For validating our TBM workflow, the ICA-derived atrophy estimates were linearly regressed against the inter-scan interval duration in study diagnosis groups. Regression intercepts at baseline timepoint were not significantly different from zero in the cognitively normal (p=0.9), MCI (p=0.2) or AD subjects (p=0.7). To identify the impact of interscan interval duration on atrophy rate estimates, annualized atrophy rate parameters across $interval_{1\rightarrow 2}$ were regressed against $interval_{1\rightarrow 3}$ to seek their linear association. All regression models in the study diagnosis groups showed non-significant intercepts, as well as slopes acceptably near unity (Normal:1.07, mild cognitive impairment:0.97, Alzheimer's disease:1.07, Fig. S2), indicating that our workflow provides linear atrophy estimates across the two-year timespan of this study.

Genome-wide association methods

Whole-genome sequencing was performed by a non-CLIA laboratory at Illumina[46]. DNA samples were sequenced on Illumina HiSeq2000 systems using paired end read with length of 100 bp at 30-40x read depth. Short read data were mapped to Human Genome build 19 by the proprietary Illumina variant caller CASAVA and single-nucleotide polymorphisms (SNPs) were called, resulting in a total of 40,631,363 variants. We considered SNP calls passing a minimum phred quality score of 30 (>99% accuracy) and read depths between a minimum of 10 and maximum of threefold average chromosomal depth, minor allele frequency of >0.01 in cross-sectional and >0.03 in longitudinal cohorts, Hardy-Weinberg disequilibrium p-value of $>10^{-6}$ and SNP call rate of >0.9. The sequence-called SNPs were validated with the microarray-based genotyping results showing 99.8 ± 0.1% concordance rate.

The first three axes of population stratification were extracted by *Eigenstrat* (https://github.com/DReichLab/EIG[47]). Genome-wide association study of the MRI endophenotype was conducted by *Plink*[48]. Several confounding variables were used as covariates in the linear regression model, including age, gender, dummy covariates encoding diagnosis groups (AD, MCI, cognitively normal), three principal axes of population structure, MRI pulse sequence (MP-RAGE vs. SPGR), and the *APOE4* allele dosage.

References

1 Gatz, M. *et al.* Role of genes and environments for explaining Alzheimer disease. *Arch Gen Psychiatry* **63**, 168-174, doi:10.1001/archpsyc.63.2.168 (2006).

2 Ridge, P. G., Mukherjee, S., Crane, P. K. & Kauwe, J. S. Alzheimer's disease: analyzing the missing heritability. *PloS one* **8**, e79771 (2013).

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

4 Risacher, S. L. & Saykin, A. J. Neuroimaging biomarkers of neurodegenerative diseases and dementia. *Semin Neurol* **33**, 386-416, doi:10.1055/s-0033-1359312 (2013).

5 Medland, S. E., Jahanshad, N., Neale, B. M. & Thompson, P. M. Whole-genome analyses of whole-brain data: working within an expanded search space. *Nat Neurosci* **17**, 791-800, doi:10.1038/nn.3718 (2014).

6 Schuff, N. *et al.* MRI of hippocampal volume loss in early Alzheimer's disease in relation to ApoE genotype and biomarkers. *Brain* **132**, 1067-1077, doi:10.1093/brain/awp007 (2009).

7 Velayudhan, L. *et al.* Entorhinal cortex thickness predicts cognitive decline in Alzheimer's disease. *J Alzheimers Dis* **33**, 755-766, doi:10.3233/jad-2012-121408 (2013).

8 Nestor, S. M. *et al.* Ventricular enlargement as a possible measure of Alzheimer's disease progression validated using the Alzheimer's disease neuroimaging initiative database. *Brain* **131**, 2443-2454, doi:10.1093/brain/awn146 (2008).

9 Stein, J. L. *et al.* Identification of common variants associated with human hippocampal and intracranial volumes. *Nature genetics* **44**, 552-561 (2012).

10 Hibar, D. P. *et al.* Common genetic variants influence human subcortical brain structures. *Nature* **520**, 224-229 (2015).
11 Hibar, D. P. *et al.* Novel genetic loci associated with hippocampal volume. *Nature communications* **8**, 13624 (2017).
12 Hua, W. Y., Nichols, T. E. & Ghosh, D. Multiple comparison procedures for neuroimaging genomewide association studies. *Biostatistics* **16**, 17-30, doi:10.1093/biostatistics/kxu026 (2015).
13 Seeley, W. W., Crawford, R. K., Zhou, J., Miller, B. L. & Greicius, M. D. Neurodegenerative diseases target large-scale human brain networks. *Neuron* **62**, 42-52, doi:10.1016/j.neuron.2009.03.024 (2009).
14 Gutman, B. A. *et al.* Empowering Imaging Biomarkers of Alzheimer's Disease. *Neurobiology of aging* **36 Suppl 1**, S69-S80, doi:10.1016/j.neurobiolaging.2014.05.038 (2015).
15 Ashburner, J., Good, C. & Friston, K. J. Tensor based morphometry. *NeuroImage* **11**, S465 (2000).
16 Avants, B. B. *et al.* A reproducible evaluation of ANTs similarity metric performance in brain image registration. *Neuroimage* **54**, 2033-2044, doi:10.1016/j.neuroimage.2010.09.025 (2011).
17 Yushkevich, P. A. *et al.* Bias in estimation of hippocampal atrophy using deformation-based morphometry arises from asymmetric global normalization: an illustration in ADNI 3 T MRI data. *Neuroimage* **50**, 434-445 (2010).
18 Hua, X. *et al.* Accurate measurement of brain changes in longitudinal MRI scans using tensor-based morphometry. *Neuroimage* **57**, 5-14 (2011).
19 Xu, L., Groth, K. M., Pearlson, G., Schretlen, D. J. & Calhoun, V. D. Source-based morphometry: the use of independent component analysis to identify gray matter differences with application to schizophrenia. *Hum Brain Mapp* **30**, 711-724, doi:10.1002/hbm.20540 (2009).
20 Beckmann, C. F. & Smith, S. M. Probabilistic independent component analysis for functional magnetic resonance imaging. *IEEE Trans Med Imaging* **23**, 137-152, doi:10.1109/TMI.2003.822821 (2004).
21 Tibshirani, R. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 267-288 (1996).
22 Winkler, A. M., Ridgway, G. R., Webster, M. A., Smith, S. M. & Nichols, T. E. Permutation inference for the general linear model. *Neuroimage* **92**, 381-397 (2014).
23 Saykin, A. J. *et al.* Genetic Studies of Quantitative MCI and AD Phenotypes in ADNI: Progress, Opportunities, and Plans. *Alzheimer's & dementia : the journal of the Alzheimer's Association* **11**, 792-814, doi:10.1016/j.jalz.2015.05.009 (2015).
24 Boyle, A. P. *et al.* Annotation of functional variation in personal genomes using RegulomeDB. *Genome Res* **22**, 1790-1797, doi:10.1101/gr.137323.112 (2012).
25 Trabzuni, D. *et al.* Quality control parameters on a large dataset of regionally dissected human control brains for whole genome expression studies. *Journal of Neurochemistry* **119**, 275-282, doi:10.1111/j.1471-4159.2011.07432.x (2011).
26 Lambert, J. C. *et al.* Genome-wide haplotype association study identifies the FRMD4A gene as a risk locus for Alzheimer's disease. *Mol Psychiatry* **18**, 461-470, doi:10.1038/mp.2012.14 (2013).
27 Poulin, S. P., Dautoff, R., Morris, J. C., Barrett, L. F. & Dickerson, B. C. Amygdala atrophy is prominent in early Alzheimer's disease and relates to symptom severity. *Psychiatry research* **194**, 7-13, doi:10.1016/j.pscychresns.2011.06.014 (2011).
28 Ortner, M. *et al.* Progressively Disrupted Intrinsic Functional Connectivity of Basolateral Amygdala in Very Early Alzheimer's Disease. *Frontiers in Neurology* **7**, 132, doi:10.3389/fneur.2016.00132 (2016).

29 Ball, T. *et al.* Anatomical specificity of functional amygdala imaging of responses to stimuli with positive and negative emotional valence. *Journal of neuroscience methods* **180**, 57-70, doi:10.1016/j.jneumeth.2009.02.022 (2009).

30 Drago, V. *et al.* Emotional indifference in Alzheimer's disease. *J Neuropsychiatry Clin Neurosci* **22**, 236-242, doi:10.1176/appi.neuropsych.22.2.236 (2010).

31 Torres, B. *et al.* Facial expression recognition in Alzheimer's disease: a longitudinal study. *Arq Neuropsiquiatr* **73**, 383-389, doi:10.1590/0004-282x20150009 (2015).

32 Lim, S. *et al.* Sharpin, a novel postsynaptic density protein that directly interacts with the shank family of proteins. *Molecular and cellular neurosciences* **17**, 385-397, doi:10.1006/mcne.2000.0940 (2001).

33 Kunkle B.W. *et al.* Whole-exome sequencing of multiplex families identifies several rare coding variants in known and novel Late-Onset Alzheimer genes. *Presented at the 64th Annual Meeting of The American Society of Human Genetics* (2014).

34 Kohli, M. *et al.* The Dissection of High-Penetrance Variants in Extended Late-Onset Alzheimer Disease Families by Whole-Exome Sequencing (S28. 004). *Neurology* **82**, S28. 004 (2014).

35 Hruska, M. & Dalva, M. B. Ephrin regulation of synapse formation, function and plasticity. *Molecular and cellular neurosciences* **50**, 35-44, doi:10.1016/j.mcn.2012.03.004 (2012).

36 Jack, C. R., Jr. *et al.* The Alzheimer's Disease Neuroimaging Initiative (ADNI): MRI methods. *J Magn Reson Imaging* **27**, 685-691, doi:10.1002/jmri.21049 (2008).

37 Avants, B. B. *et al.* The optimal template effect in hippocampus studies of diseased populations. *Neuroimage* **49**, 2457-2466, doi:10.1016/j.neuroimage.2009.09.062 (2010).

38 Jenkinson, M. & Smith, S. A global optimisation method for robust affine registration of brain images. *Med Image Anal* **5**, 143-156 (2001).

39 Gunter, J. L., Shiung, M. M., Manduca, A. & Jack, C. R. Methodological Considerations for Measuring Rates of Brain Atrophy. *Journal of magnetic resonance imaging : JMRI* **18**, 16-24, doi:10.1002/jmri.10325 (2003).

40 Hua, X. *et al.* Accurate measurement of brain changes in longitudinal MRI scans using tensor-based morphometry. *Neuroimage* **57**, 5-14, doi:10.1016/j.neuroimage.2011.01.079 (2011).

41 Smith, S. M., De Stefano, N., Jenkinson, M. & Matthews, P. M. Normalized accurate measurement of longitudinal brain change. *Journal of computer assisted tomography* **25**, 466-475 (2001).

42 Clarkson, M. J. *et al.* Comparison of phantom and registration scaling corrections using the ADNI cohort. *Neuroimage* **47**, 1506-1513, doi:10.1016/j.neuroimage.2009.05.045 (2009).

43 Lewis, E. B. & Fox, N. C. Correction of differential intensity inhomogeneity in longitudinal MR images. *Neuroimage* **23**, 75-83, doi:10.1016/j.neuroimage.2004.04.030 (2004).

44 Barre, A. & Armand, S. Biomechanical ToolKit: Open-source framework to visualize and process biomechanical data. *Computer methods and programs in biomedicine* **114**, 80-87, doi:10.1016/j.cmpb.2014.01.012 (2014).

45 Yushkevich, P. A. *et al.* User-guided 3D active contour segmentation of anatomical structures: significantly improved efficiency and reliability. *Neuroimage* **31**, 1116-1128, doi:10.1016/j.neuroimage.2006.01.015 (2006).

46 Saykin, A. J. *et al.* Genetic studies of quantitative MCI and AD phenotypes in ADNI: Progress, opportunities, and plans. *Alzheimer's & Dementia* **11**, 792-814 (2015).

47 Price, A. L. *et al.* Principal components analysis corrects for stratification in genome-wide association studies. *Nature genetics* **38**, 904-909, doi:10.1038/ng1847 (2006).

48 Purcell, S. *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* **81**, 559-575, doi:10.1086/519795 (2007).

Supplementary Figures

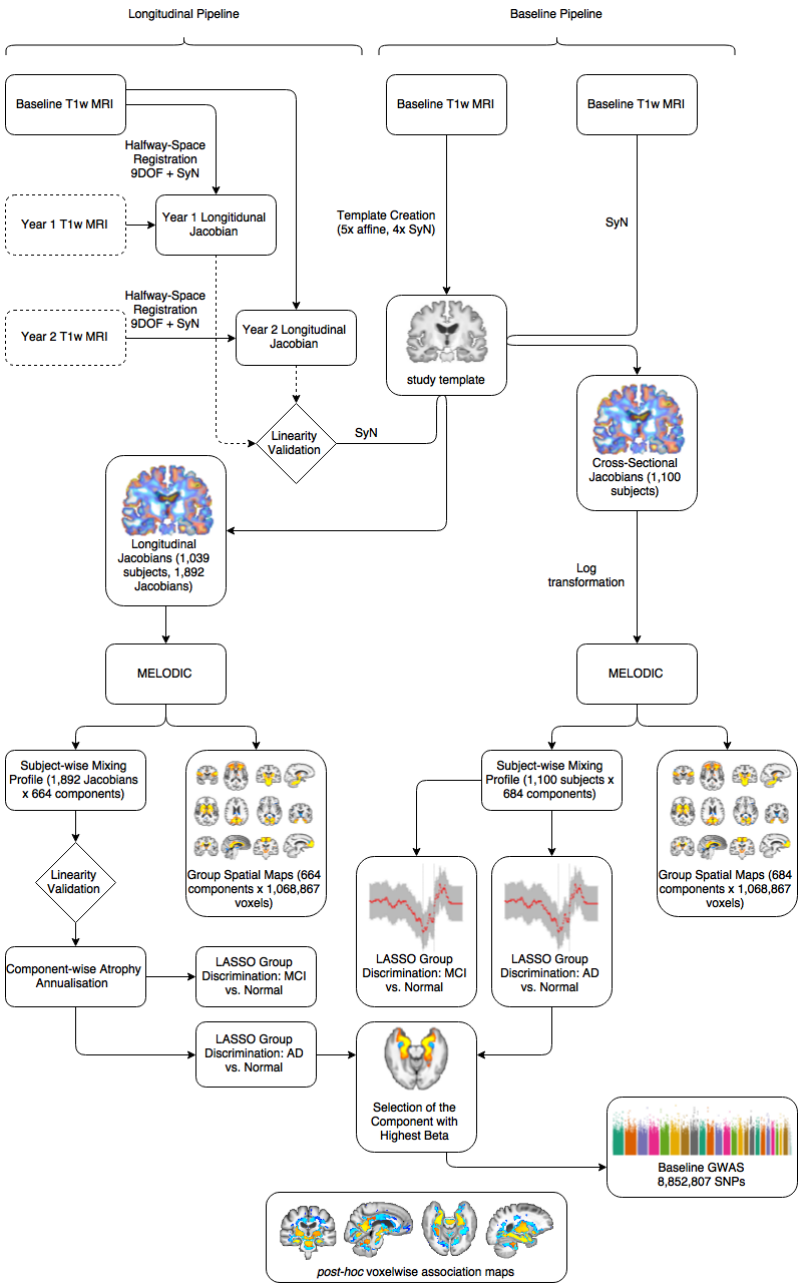

Figure S1: **Outline of Study methods**. T1-weighted MRIs were used to identify structural brain changes in cross-sectional and longitudinal studies. ICA decomposed 1,348 spatial sources of brain atrophy, among which the medial temporal circuit (MTC) was replicated as the top imaging predictor of AD and MCI and subsequently brought to GWAS.

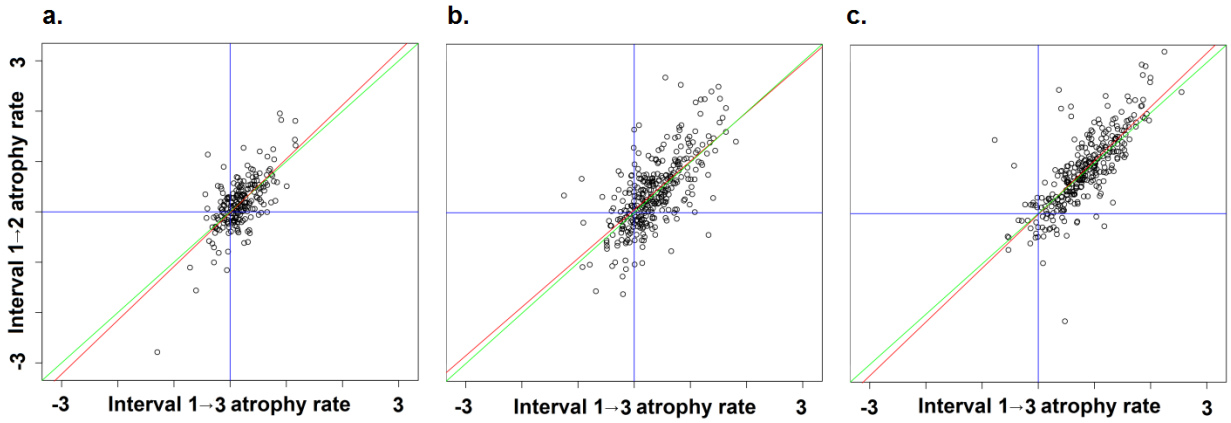

Figure S2: **Linearity of atrophy rates estimated by ICA**. Annualized atrophy rate scatter plots of longitudinal MRI across $timepoint_{1\rightarrow3}$ (baseline and year-2) against $timepoint_{1\rightarrow2}$ (baseline and year-1) in cognitively normal (**plot A**), MCI (**plot B**) and AD (**plot C**) subjects. Regression lines (red) and the ground truth identity line (green) are shown which disclose minimal bias.

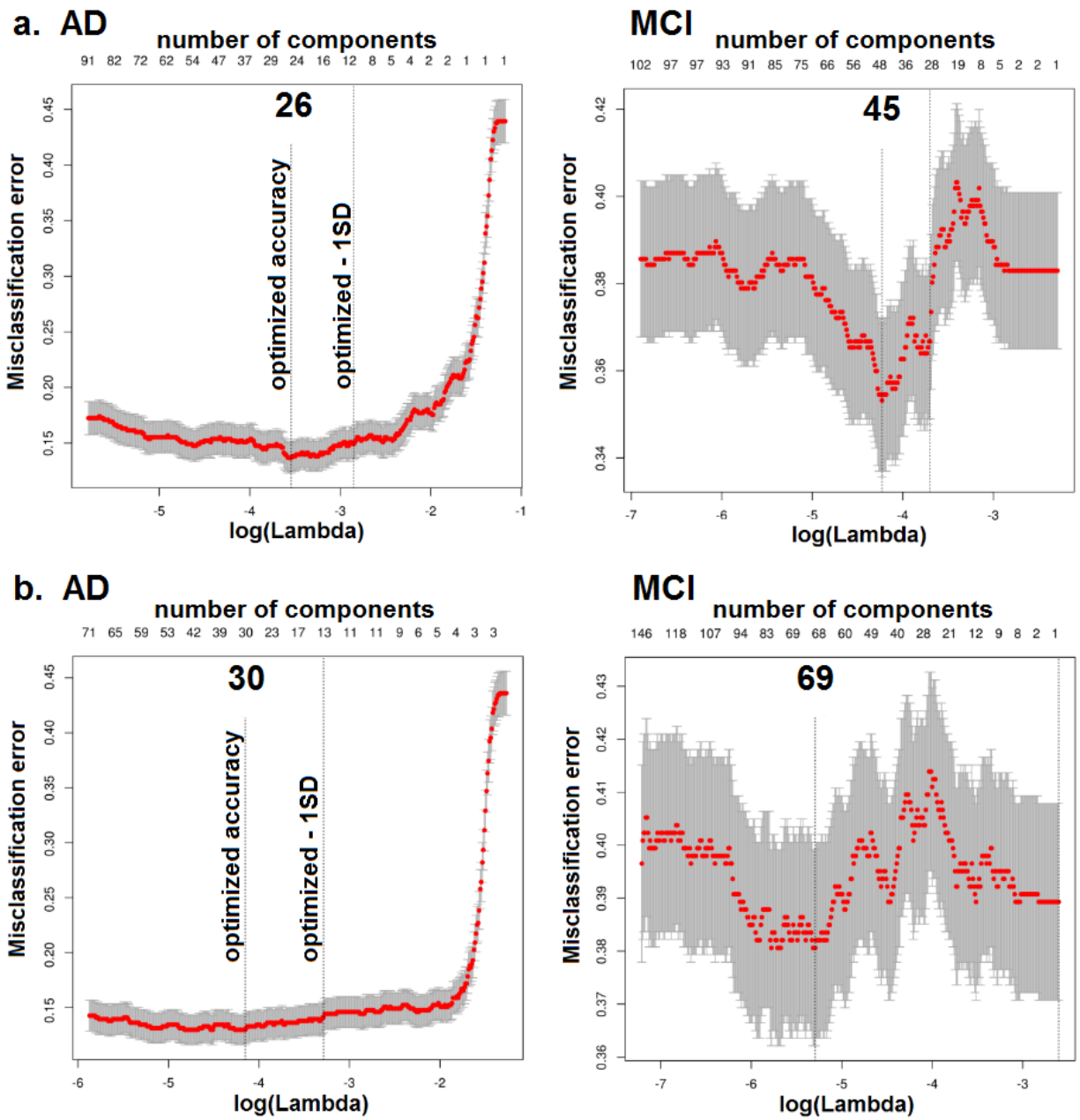

Figure S3. L_1-regularized logistic regression models for diagnosis prediction. Leave-one-out cross-validation results for discriminating AD or MCI subjects from the cognitively normal group using ICA components. **a)** All of the 684 brain atrophy components in the cross-sectional study phase were used as regression predictors for diagnosis discrimination. Increased L_1 regularization parameter (lambda values, x-axis: bottom) shrinks the regression model and results in inclusion of fewer predictor components (x-axis: top). Separate regression models were implemented to discriminate AD (**left**) or MCI subjects (**right**) from the cognitively normal group. **b)** A similar regression model was implemented in longitudinal neuroimaging using 664 components to discriminate AD (**left**) or MCI subjects (**right**) from the normal population. While AD prediction was possible with good accuracy by as few as five ICA components, the MCI discrimination model did not perform at a comparable accuracy.

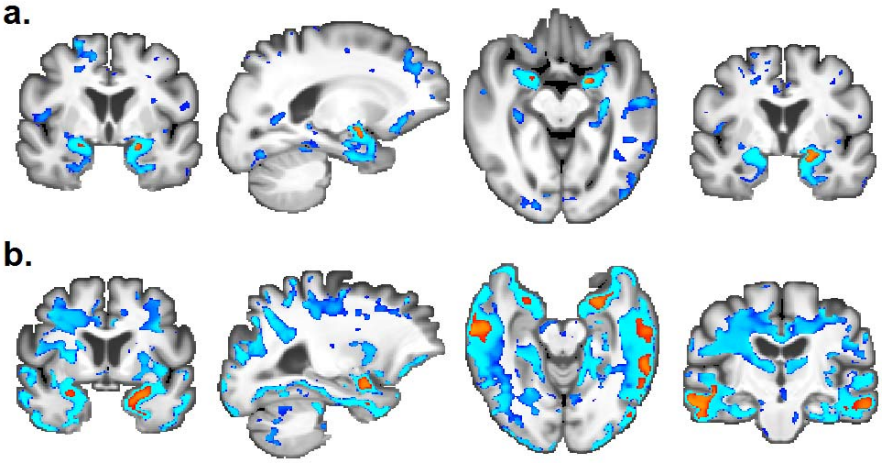

Figure S4. Voxelwise atrophy impacts of the *APOE4* allele. Correlation maps of the *APOE4* dosage in cross-sectional (**a**) and longitudinal (**b**) brain morphometry. The most consistent impact was observed in bilateral amygdalar areas (Blue: uncorrected $p<0.05$, Red-yellow: corrected $p<0.05$).

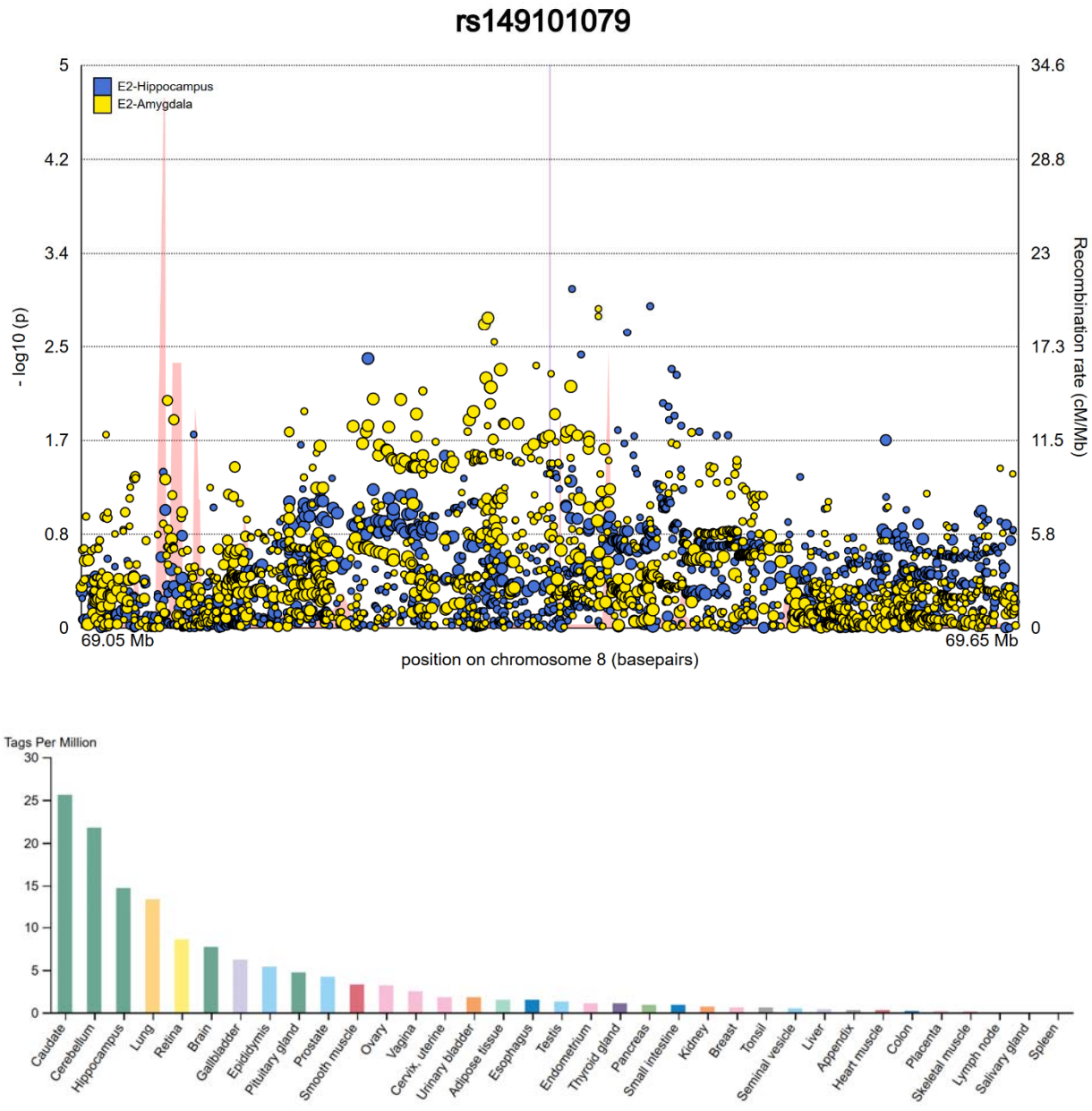

Figure S5. C8orf34 locus. Regional association plot of the *C8orf34* locus in ENIGMA-2 meta-analysis (**top,** *https://www.enigma-brain.org/enigmavis/visualizer/visualizer*) and tissue expression profile of *C8orf34* in FANTOM5 database (**bottom**, shown as short sequence reads per million, *https://www.proteinatlas.org*).

Supplementary Video

3D volume rendering of the medial temporal ICA component. The z-score map of this component is demonstrated at incremental z-thresholds. The highest probability of brain atrophy is observed in bilateral amygdalae, further extending to hippocampus, their white-matter connections to fornix and other structures of the limbic system.