

VU Research Portal

Multiscale neuroscience of the healthy and diseased brain

Wei, Y.

2020

document version

Publisher's PDF, also known as Version of record

[Link to publication in VU Research Portal](#)

citation for published version (APA)

Wei, Y. (2020). *Multiscale neuroscience of the healthy and diseased brain*.

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

E-mail address:

vuresearchportal.ub@vu.nl

4

GAMBA: an integrative platform for annotation of gene transcription-neuroimaging associations

Yongbin Wei, Siemon C. de Lange, Rory Pijnenburg, Lianne H. Scholtens, Dirk Jan Ardesch, Kyoko Watanabe, Danielle Posthuma, Martijn P. van den Heuvel
Manuscript in preparation

Abstract

The integration of neuroimaging and genetics enables us to assess the underlying biological mechanisms of human cognition and mental disorders. Here, we present a web-based application, GAMBA (short for Gene Annotation using Macroscale Neuroimaging Association), for scrutinizing gene transcription-neuroimaging associations. GAMBA utilizes gene expression data from the Allen Human Brain Atlas (AHBA) and cross-references the spatial pattern of cortical gene expression to the patterns of a range of neuroimaging-derived phenotypes. We present three examples to demonstrate the usability of GAMBA. In the first example, we show that brain expressions of human supragranular-enriched (HSE) genes are significantly correlated with the nodal strength of the macroscale brain connectome, suggestive of an association between micro- and macro-scale brain connectivity. The second and third examples show the expression profile of *APOE* gene and risk genes for autism spectrum disorder (ASD) respectively to be associated with the pattern of brain structural and functional alterations in Alzheimer's disease and Asperger's syndrome. Together, GAMBA provides a user-friendly, open-source platform for functional annotation of genes concerning neuroimaging-derived phenotypes of the healthy and diseased brain.

4

Introduction

Mapping the biological pathways that link genes with brain functions and dysfunctions enable us to formulate hypotheses on the mechanism of mental disorders and to prioritize genes for functional follow-up research. The recent emergence of neuroimaging-genetic studies has described a growing number of genetic variants associated with multimodal neuroimaging traits by means of large-sample genome-wide association studies (GWAS) [1–3]. These neuroimaging traits have broadly covered anatomical morphology estimated by T1-weighted magnetic resonance imaging (MRI), brain activity measured using functional MRI (fMRI), and brain connectivity derived from diffusion weighted imaging and resting-state fMRI [4]. The single-nucleotide polymorphisms (SNPs) observed in GWAS are commonly assumed to play a role in encoding the amino acid sequences of proteins or regulating the activity of gene expression [5], which in turn influences e.g. synaptic and neuron morphology, neuronal functioning, glial activity, et cetera, and ultimately the macroscale

brain structure and function [6]. Empirical evidence for the associations between gene expression and the macroscale brain traits are however still largely missing.

Brain transcription data from the Allen Human Brain Atlas (AHBA) serves as a valuable quantitative reference to assess the transcription-neuroimaging linkage [7]. Regional variation of gene expressions across the brain have been noted to often run parallel with spatial patterns of features derived from neuroimaging findings (see for reviews, [6, 8]). Examples of these associations include genes related to oxidative metabolism to display transcriptional profiles in a similar pattern as the degree of inter-module long-distance functional connectivity (FC) [9], and genes enriched in supragranular layers of the human cortex to show transcriptional profiles that capture the architecture of both brain structural and functional networks [10, 11]. Transcriptional profiles of genes associated with human accelerated regions of the genome are also correlated to the pattern of cortical expansion in human evolution [12]. Furthermore, brain transcriptional profiles of risk-genes related to brain dysfunction tend to show high similarity with disorder-specific brain changes revealed by neuroimaging techniques. Risk genes of schizophrenia and autism for example show elevated expression levels in brain areas where alterations of brain connectivity and morphology are evident [13, 14] and gene markers for somatostatin interneurons are observed to be more expressed in cortical regions that are more vulnerable to major depressive disorder [15].

Examining transcription-neuroimaging association based on spatial variations thus provides a unique opportunity for the exploration of genes related to specific phenotypes of cognitive abilities or brain disorders. Here, we present an integrative platform, GAMBA (short for Gene Annotation using Macroscale Neuroimaging Association) to scrutinize the potential transcription-neuroimaging associations. GAMBA is an open-access platform and is available online at <http://dutchconnectomelab.nl/GAMBA/> and can be directly used from the gene annotation application FUMA (<https://fuma.ctglab.nl/>) [16]. For any genes of interest, GAMBA displays the cortical expression profile of these genes derived from the AHBA dataset together with the spatial patterns of neuroimaging-related phenotypes, and accompanied statistical testing of the genetic-neuroimaging associations. Neuroimaging phenotypes include the topological layout of resting-state brain functional networks, brain structural connectivity, cognitive components, cortical metabolic properties, functional and structural alterations across a range of brain disorders, etcetera. GAMBA provides a user-friendly open-source platform for functional annotation of genes with respect to macroscale phenotypes of cognitive brain structure and function.

Results

GAMBA overview

GAMBA is built around the notion of cross-referencing human brain transcription data with the macroscale brain imaging data [10, 12, 14, 17]. Transcription data are obtained from the Allen Human Brain Atlas (AHBA; <http://human.brain-map.org>). Data preprocessing of the AHBA data includes gene information reannotation, probe selection, sample assignment, normalization, and averaging among brain donors (see Methods), resulting in the average transcriptional profile of 22,749 genes across 57 cortical areas of the left cortical mantle. The cortical expression profile of each of the 22,749 genes is linked to neuroimaging phenotypes from eight categories using a multivariate linear regression model (Figure 4.1). Neuroimaging phenotypes include the spatial pattern of: a) resting-state functional networks [18], b) brain cognitive components [19], c) regional metrics derived from the brain structural and functional connectome, d) measurements of the level of cortical oxygen and glucose metabolism [20], e) human surface area expansion compared to the chimpanzee [12], f) brain volume alterations across twenty-two brain disorders [21–23], g) brain functional changes in sixteen brain disorders [21–23], h) brain dysconnectivity across nine psychiatric and neurological disorders [24], i) brain functional correlates of 530 terms in relation to cognitive states and brain disorders, as described in the NeuroSynth database [25]. Details about the datasets and statistical analyses are described in Methods.

In the following sections, we present three examples to demonstrate the validity of GAMBA and how GAMBA facilitates our understanding of the transcription-neuroimaging associations.

Application to genes related to neuronal connectivity

As the first example, we apply GAMBA to genes related to the formation of microscopic neuronal connectivity, examining a set of 19 genes known to be enriched in supragranular layers of the human cerebral cortex (referred to as HSE genes) and known to play an important role in long-range cortico-cortical connections [10]. HSE genes showed significantly higher expression than genes on average in the inferior/superior parietal cortex, supramarginal cortex, precuneus, superior/middle frontal cortex, precentral gyrus, and paracentral cortex (Figure 4.2a. , two-sided t -

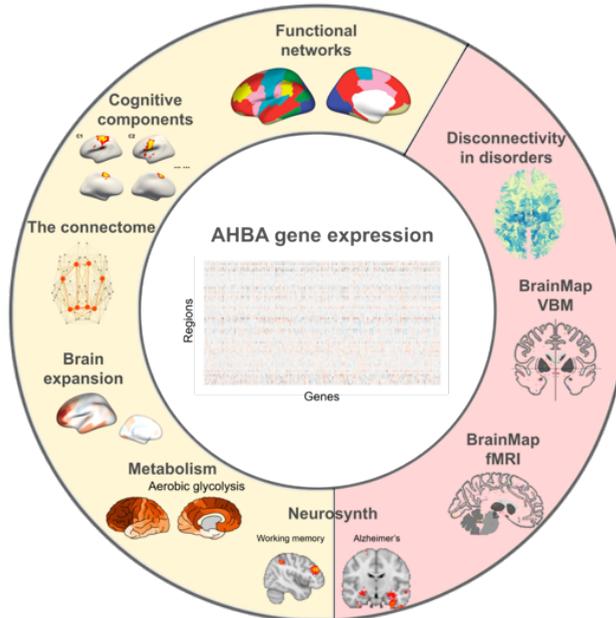


Figure 4.1: An overview of the datasets used in GAMBA. Brain gene expressions are extracted from the Allen Human Brain Atlas (AHBA) and cross-reference to distinct brain imaging phenotypes. Yellow indicates brain imaging phenotypes related to healthy brain organization, red indicates brain imaging phenotypes related to brain disorders.

test; $q < 0.05$, False Discovery Rate [FDR] corrected). Lower HSE gene expression in comparison to genes on average was noted in the anterior cingulate gyrus, entorhinal cortex, parahippocampal gyrus, medial orbital frontal cortex, lingual gyrus, and insular ($q < 0.05$, FDR corrected). Gene expression of HSE genes was significantly associated with the nodal strength of the structural connectome – a measurement describing the extent to which a region is structurally connected to the rest of the brain [standardized beta (β) = 0.520, $p < 0.001$ for the connectivity weighted by the number of streamlines (NOS); $\beta = 0.624$, $p < 0.001$ for the connectivity weighted by streamline density (SD); FDR corrected; Figure 4.2]. A similar correlation was shown between the gene expression profile of HSE genes and the nodal strength of the functional connectome (FC; $\beta = 0.524$, $p < 0.001$; FDR corrected; Figure 4.2). These results suggest that the formation of microscale cortico-cortical connectivity is associated with the organization of macroscale connectome, stressing the usability of GAMBA to uncover the association across different scales of brain organization.

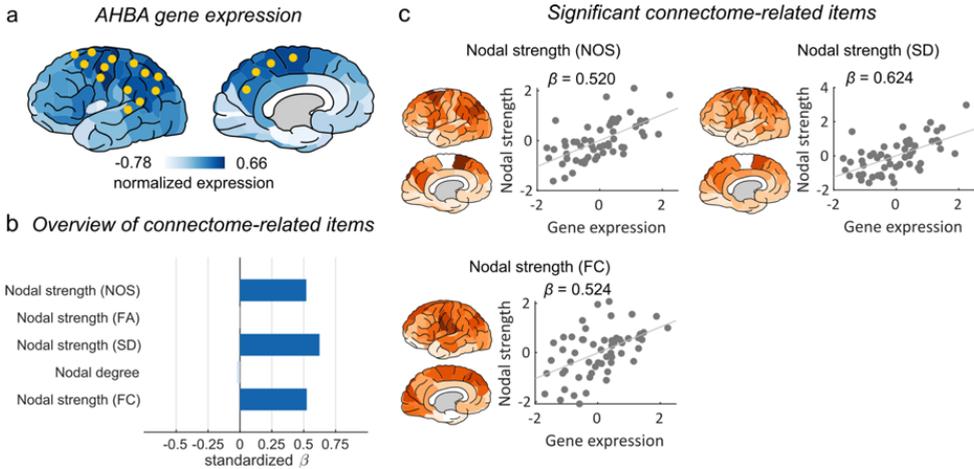


Figure 4.2: Applications of GAMBA on HSE genes. (a) GAMBA figure panels of gene expression levels of HSE genes, showing significantly higher expressions of HSE genes than genes on average in the inferior/superior parietal cortex, supramarginal cortex, precuneus, superior/middle frontal cortex, precentral gyrus, and paracentral cortex. Yellow dots indicate significant regions (two-sided t -test, $q < 0.05$, FDR corrected). (b) GAMBA exhibits significant correlations between HSE gene expression and nodal strength of the structural (NOS-weighted, $\beta = 0.520$ and SD-weighted, $\beta = 0.624$) and functional connectome ($\beta = 0.524$; $q < 0.05$, FDR corrected). (c) Brain maps of the significant brain imaging phenotypes, including the NOS-, SD-, and FC-weighted nodal strength, and scatter plots of the significant associations are displayed.

GAMBA includes null-models to disentangle functional annotations of genes of interest from effects present across all genes. GAMBA tests whether an observed association between gene expression of a set of genes (here HSE) and a macroscale phenotype of interest (here macroscale brain connectome measurements) is a general phenomenon that can be noted for all genes, or whether it is specific to the set of genes of interest. For this purpose, GAMBA uses bootstrap to estimate the mean and standard deviation of the effect size (β) for the same-sized gene-sets randomly selected from all genes (10,000 samplings; referred to as "NULL-1"). Performing z -test showed that the observed effect size for associations between HSE genes and nodal strength weighted by NOS, SD, and FC were significantly larger than the mean effect size of all genes (NOS: $z = 3.874$, $p < 0.001$; SD: $z = 3.492$, $p < 0.001$; FC: $z = 2.408$, $p = 0.016$; $q < 0.05$, FDR corrected).

GAMBA further includes more strict null-models (allowing for a more strict statistical evaluation) to compare the observed effect sizes to the mean effect size of a

set of genes particular expressed in BRAIN tissues (referred to as "NULL-2") [12]. Using this null-model showed significant associations between HSE genes and the connectome metrics (NOS: $z = 4.093$, $p < 0.001$; SD: $z = 3.767$, $p < 0.001$; FC: $z = 2.706$, $p = 0.007$; FDR corrected).

GAMBA includes a third null-model to examine whether the observed association is independent of 'generic anatomical relationships' across brain regions. For this, the gene expression data is rebuilt using randomized brain parcellations that are obtained by spinning the reconstructed sphere of the real brain parcellation (1,000 permutations; further referred to as "NULL-3", see Methods for details) as proposed by [26], and all statistics are reperformed. Using this null model confirmed the significant association between expressions of HSE genes and structural connectivity strength (NOS: $z = 2.448$, $p = 0.007$; SD: $z = 2.377$, $p = 0.009$; FDR corrected), as well as the strength of functional connectivity ($z = 2.853$, $p = 0.002$; FDR corrected). These GAMBA findings implicate a specific role of HSE genes in forming the neurons in supragranular layers and their cortico-cortical connectivity, which are the "building blocks" for the macroscale brain connectivity.

Application to Alzheimer's Disease Risk Gene APOE

In the second example, we show the usability of GAMBA in examining brain patterns of disease-related pathology. We examine the Apolipoprotein E (APOE) gene, widely indicated as a risk gene for Alzheimer's disease (AD) [27]. We hypothesize that the cortical gene expression of *APOE* would be related to the cortical alterations revealed in patients with AD. First, GAMBA analysis showed the highest expression level of *APOE* in the entorhinal cortex ($z = 2.092$, $p = 0.036$; two-sided z -test based on NULL-1; Figure 4.3a), a region well-known to be involved in AD (Bobinski, et al., 1999) and often marked as one of the epi-centers of brain pathology in AD (Van Hoesen, et al., 1991). GAMBA analysis further showed significant associations between the transcriptional profile of *APOE* and the cortical atrophy patterns as reported by the meta-analysis of BrainMap voxel-based morphometry (VBM) studies of i) Alzheimer's disease ($\beta = 0.608$, $p < 0.001$), ii) dementia ($\beta = 0.621$, $p < 0.001$), iii) semantic dementia ($\beta = 0.587$, $p < 0.001$), iv) frontotemporal dementia ($\beta = 0.455$, $p < 0.001$), v) attention deficit hyperactivity disorder (ADHD, which is hypothesized to be a risk factor of dementia pathology (Callahan, et al., 2017), $\beta = 0.384$, $p = 0.002$), and vi) bipolar disorder ($\beta = 0.343$, $p = 0.010$; $q < 0.05$, FDR corrected; Figure 4.3b). These effect sizes exceeded the null distributions of NULL-1 (i.e., mean effect sizes

across all genes) for semantic dementia ($z = 2.019$, $p = 0.044$) and Alzheimer's disease ($z = 2.039$, $p = 0.042$). Examining spatial specificity showed significant associations of *APOE*'s expression with brain volume changes in dementia ($z = 3.234$, $p = 0.001$, NULL-3), Alzheimer's disease ($z = 3.119$, $p = 0.001$ NULL-3), and frontotemporal dementia ($z = 2.785$, $p = 0.005$; FDR corrected, NULL-3). NULL-2 (i.e., comparing the observed effect size to the mean effect size of brain-expressed genes, instead of ALL genes) did not reveal specific significant associations ($p > 0.097$), suggesting that there are other brain-expressed genes showing similar expression patterns to the atrophy pattern of above-mentioned diseases.

Furthermore, cross-referencing transcription data to items in the functional NeuroSynth database again revealed significant associations of *APOE* expression with terms of "Alzheimer" ($\beta = 0.479$, $p < 0.001$) and "dementia" ($\beta = 0.451$, $p < 0.001$; $q < 0.05$, FDR corrected).

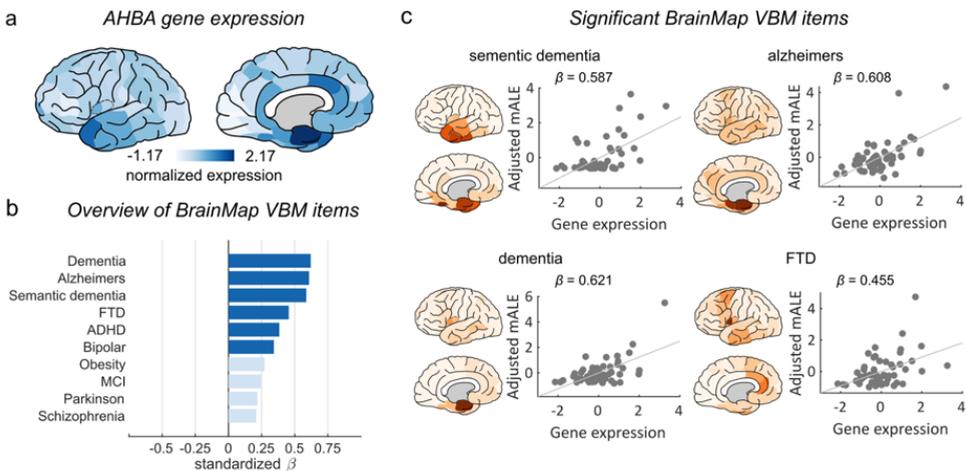


Figure 4.3: Applications of GAMBA on *APOE*. (a) GAMBA figure panels of gene expression levels of *APOE*. (b) GAMBA analysis demonstrates significant correlations between *APOE* gene expression and voxel-based morphometry (VBM) changes in semantic dementia ($\beta = 0.587$), alzheimers ($\beta = 0.608$), dementia ($\beta = 0.621$), frontotemporal dementia ($\beta = 0.455$), ADHD ($\beta = 0.384$), and bipolar disorder ($\beta = 0.343$; $q < 0.05$, FDR corrected). (c) Brain maps of the significant brain imaging phenotypes, including the semantic dementia, Alzheimers, dementia, and frontotemporal dementia, and scatter plots of the significant associations (items showing significance for at least one NULL-model) are displayed.

GAMBA mapping further revealed a significant association of the transcriptional profile of *APOE* with the organization of the limbic network ($\beta = 0.532$, $p < 0.001$;

FDR corrected, $q < 0.05$), exceeding the mean effect size of NULL-1 ($z = 2.260$, $p = 0.024$) and NULL-3 ($z = 2.852$, $p = 0.004$), but not the more strict NULL-2 ($z = 1.840$, $p = 0.066$). We argue that these findings strengthen the importance of including a NULL-2 (i.e., null model of effect sizes of BRAIN genes), on top of the NULL-1 model [13, 14], in examinations of spatial correlations between transcription and imaging features. NULL-2 [12, 15] is a more strict null model to avoid false-positive findings of gene-expression-neuroimaging associations.

Taking together, our findings show that medial temporal and parahippocampal regions with high normative *APOE* expression also display significantly elevated structural and functional brain abnormalities in dementia and AD in disease conditions. GAMBA gene mapping further suggests an important role of *APOE* in regions of the limbic network, a network that is known to subserve brain functions of memory and emotion [28] and to be regularly reported to AD and MCI pathology and clinical outcome [29].

Application to risk genes of autism

We next apply GAMBA to risk genes of autism spectrum disorder (ASD) to further demonstrate the utility of the platform. A set of 25 ASD risk genes was obtained from a recent meta-analysis of genome-wide association study on 18,381 ASD cases and 27,969 controls [30], 24 of which are included in the AHBA data (further referred to as ASD genes, averaged gene expression levels displayed in Figure 4.4a). GAMBA analysis indicated high expression of ASD genes in the superior parietal lobe ($z = 2.654$, $p = 0.008$), precuneus ($z = 2.304$, $p = 0.021$), cuneus ($z = 2.684$, $p = 0.007$), lingual gyrus ($z = 2.591$, $p = 0.010$), lateral occipital lobe ($z = 2.573$, $p = 0.010$), and pericalcarine ($z = 2.909$, $p = 0.004$; uncorrected; two-sided z -test comparing the NULL-1, Figure 4.4a). This transcriptional profile showed a significant association with brain functional alterations as reported by the meta-analysis of BrainMap fMRI studies in "Asperger's syndrome" ($\beta = 0.443$, $p < 0.001$; FDR corrected) (Figure 4.4b), a major diagnosis of ASD with difficulties in social interaction and nonverbal communication as derived from BrainMap data [21–23]. The observed effect size was significantly larger than NULL-2 ($z = 2.097$, $p = 0.036$) and NULL-3 ($z = 3.120$, $p = 0.002$).

GAMBA further subdivided the 24 ASD genes into four communities according to their co-expression network [Modularity index $Q = 0.243$ (see methods), which was significantly different from the null distribution of modularity index derived from

random genes ($p = 0.001$, 10,000 permutations)]. Two communities, including seven genes (*HDAC4*, *KANSL1*, *KIZ*, *PLEKHM1*, *WDPCP*, *WNT3*, *XRN2*) and seven genes (*CRHR1*, *MDH1*, *MFHAS1*, *NTM*, *PINX1*, *SOX7*, *XKR6*), separately, showed significant association with brain functional alterations as observed in "Asperger's syndrome" ($\beta = 0.417$, $p = 0.002$ and $\beta = 0.407$, $p = 0.002$; Figure 4.4c). GAMBA findings thus suggest that cortical areas with high expression of ASD risk genes are more likely to show structural and functional brain alterations related to Asperger's syndrome, a condition strongly related to ASD [31].

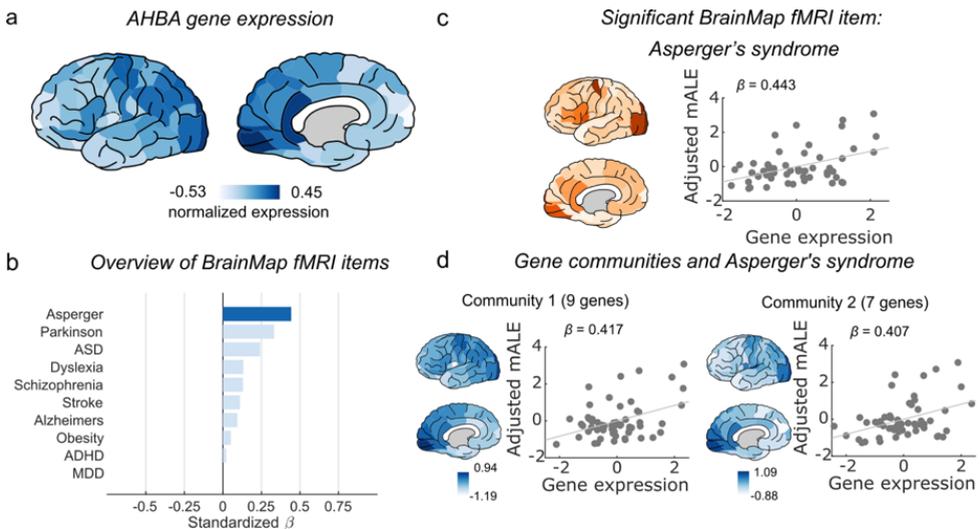


Figure 4.4: Applications of GAMBA on ASD genes. (a) GAMBA figure panels of gene expression levels of ASD genes. (b) GAMBA demonstrates significant correlations between gene expression of ASD genes and brain activity changes in Asperger's syndrome ($\beta = 0.444$; $q < 0.05$, FDR corrected). (c) Maps of brain activity changes in Asperger's syndrome and its association between ASD gene expression. (d) GAMBA detects communities within the set of ASD genes and shows two communities to be significantly correlated to brain activity changes in Asperger's syndrome ($\beta = 0.417$ and 0.407 , separately).

Discussion

The present study introduces a web-based application – GAMBA – that can be used to probe the association between transcription and brain structure/function derived from a wide range of neuroimaging data. Given a single gene or a set of genes, GAMBA tests whether brain regions in which genes of interest are preferentially ex-

pressed show overlapping patterns of a wide range of brain imaging phenotypes.

The application of GAMBA is demonstrated in three examples, HSE genes, *APOE*, and ASD genes, showing GAMBA as a tool for gene function annotation in the context of both health and disease conditions. The applications we show for example indicate that the cortical transcriptional profile of genes involved in microscale neuronal connectivity runs parallel with variation in structural and functional connectivity across brain regions. HSE genes are known to be enriched in supragranular layers of the human cortex and are responsible for long-range corticocortical projections of pyramidal neurons [32]. Cortical transcriptional profiles of these genes capture the architecture of distributed brain functional networks [10] and correlate with the pattern of cortical thickness covariance network [11]. These findings suggest a putative role of HSE genes in shaping large-scale brain connectome, which is compatible with observations of association areas to display the more complicated dendritic structure of layer III pyramidal neurons and to be more connected by white matter tracts at the macroscale level of the connectome organization [6, 33]. GAMBA analysis of HSE genes provides a new way of easily assessing such cross-scale associations of brain connectivity.

The utility of GAMBA is further shown in the application of disease conditions. *APOE* gene expression pattern across cortical areas is shown to correlate with the patterns of structural and functional brain alterations in AD. The *APOE* $\epsilon 4$ variant is known as one of the strongest genetic risk factors for late-onset AD [27, 34, 35]. AD patients are known to have elevated expression levels of the *APOE* gene in the hippocampus and medial temporal regions, areas that are well-known to be involved in the pathology of the disorder [36–38]. Brain imaging studies have further shown that *APOE* $\epsilon 4$ carriers show a larger extent of atrophy in these brain areas compared to healthy aging [39–41]. GAMBA analysis may further extend these findings by stressing that high normative levels of *APOE* expression in for example medial temporal regions may be related to a higher risk for atrophy of these brain regions in AD pathology.

As for a third example we present an association between ASD risk genes and brain functional alterations in Asperger's syndrome that is part of the diagnosis of ASD. ASD is a highly heterogeneous disorder and patients with Asperger's syndrome are the subtype showing the highest heritability [30]. Our example adds to this by showing that genetic variants observed in ASD might be related to the vulnerability of brain functions in Asperger's syndrome. This finding is also in line with the observed association between ASD polygenic risk score and grey matter volume [42], pointing

to the contribution of genetic variants in macroscopic brain alterations reported in the disease. On top of that, this example shows the usage of GAMBA in post-GWAS analysis – exploring brain-related functional annotation for genes identified in GWAS studies.

GAMBA complements other tools that link genetics and brain functions. Tools like Neurosynth [25] and Brain Annotation Toolbox [43] provide platforms to visualize the whole-brain map of gene expressions of a given gene and the maximal correlations between gene expressions and brain maps. GAMBA adds to this by integrating a large gene set or a broader range of brain imaging findings and by performing statistical evaluations of resulting associations. GAMBA is developed to provide a more complete inclusion of brain imaging findings across a wide range of modalities and datasets, and a quick, user-friendly, web-based interface.

A number of methodological considerations have to be kept in mind when interpreting findings from GAMBA. First, GAMBA is based on gene expression data from the AHBA dataset, which were collected from post-mortem brain tissues from six human donors. Inter-individual gene expression variability is inherently not taken yet into account [44, 45]. As an integrative platform, GAMBA will allow for extensions of other datasets of brain gene expressions, such as PsychENCODE [46] and BrainSpan [47]. Second, AHBA gene expression data are extracted from post-mortem human brains with no psychiatric and neurologic conditions. Therefore, GAMBA results regarding brain disorders should be interpreted in the context of gene expression of risk genes in brain regions to be a potential marker for a higher susceptibility of these regions in relevant brain disorders [14, 48]. Third, GAMBA detects spatial associations between cortical gene expression profile and brain imaging findings, making it hard to directly infer causal relationships. GAMBA is best used to examine brain annotations of genes with hypothesized functions. Functional animal follow-up studies aiming to provide causal relations from gene expression and brain function and structure are needed.

In conclusion, GAMBA provides a user-friendly platform to annotate transcriptome-neuroimaging associations, integrating brain gene expression data with a wide range of neuroimaging data. GAMBA can be used to gain new insights into brain-relevant functional annotations for genes of interests in relation to healthy brain structure and function, and in relation to structural and functional brain alterations across various brain disorders. By linking transcriptomic and neuroimaging data, GAMBA may help post-GWAS gene annotation, facilitating the prioritization of genes for further functional follow-up experiments.

Methods

Gene expression

Microarray gene expression data were obtained from the Allen Human Brain Atlas (<http://human.brain-map.org>), with data collected from six postmortem brains of human donors without any neuropathological or neuropsychiatric conditions (demographics tabulated in Chapter 3). Microarray analyses are described in detail in <http://help.brain-map.org/display/humanbrain/Documentation>. An average of 466 brain tissue samples in left hemisphere were obtained from four donors (466 ± 72.6 samples from H0351.1009, H0351.1012, H0351.1015, and H0351.1016), and 946 and 893 samples covering both hemispheres from the remaining two donors (H0351.2001 and H0351.2002). We included tissue samples in cortical and subcortical regions of the left hemisphere and used the expression of 58,692 probes for each brain donor [12, 14].

We performed probe-to-gene re-annotations using the BioMart data mining tool (<https://www.ensembl.org/biomart/>) [49]. Gene symbols were replaced by the previous and alias gene symbols by the approved symbols obtained from the HUGO Gene Nomenclature Committee (HGNC) database (<http://biomart.genenames.org/>), resulting in the inclusion of 22,749 genes. Per brain, and per tissue sample, expressions of probes annotated to the same gene symbol were averaged, followed by log₂ transformation with pseudocount 1. Tissue samples were spatially mapped to the FreeSurfer cortical regions to obtain region-wise gene expression profiles [50]. The cortical parcellation of 114 regions (57 per hemisphere) based on the Desikan-Killiany atlas (DK-114) was obtained for the Montreal Neurological Institute (MNI) ICBM152 template using FreeSurfer [51–55]. Brain tissue samples were annotated based on MNI coordinates to cortical regions in the DK-114 atlas by computing the nearest gray matter voxel in the MNI ICBM152 template in the FreeSurfer space. Tissue samples with a distance shorter than 2mm to the nearest gray matter voxel were included. Gene expression profiles of tissue samples belonging to the same cortical region were averaged, resulting in a $6 \times 57 \times 22,749$ data matrix (i.e., donors \times cortical regions \times genes). Gene expression was normalized within each donor to z-scores across all cortical regions per gene. Normalized gene expression profiles were averaged across the six donors obtaining a group-level gene expression matrix of size of $57 \times 22,749$. The resulting cortical patterns of expression of each gene was linked to the patterns of various neuroimaging findings as described in the

following sections.

Resting-state functional networks

Cortical regions in the DK-114 atlas were assigned to seven well-defined functional networks [18, 56, 57] using the Yeo 7-network atlas [18]. Networks included the visual, somatomotor, dorsal-attention, ventral-attention, limbic, frontal-parietal, and default-mode network. Surface-based annotation of the functional networks of the averaged subject in FreeSurfer (i.e., fsaverage) was translated to a 3D brain image in volumetric space, in which each gray matter voxel was annotated to a network label. Per DK region the ratio of voxels that belonged to each of the seven functional networks was computed, resulting in a mapping matrix of the size of 114×7 to indicate the involvement of each DK region in each of the functional networks.

Brain cognitive components

Brain regions were mapped to functional ontology of twelve distinct brain cognitive components related to various cognition and behavioral processes based on functional neuroimaging data from 10,449 experiments [19]. Brain components on the averaged subject in FreeSurfer were downloaded from https://surfer.nmr.mgh.harvard.edu/fswiki/BrainmapOntology_Yeo2015 and translated to 3D brain images in volumetric space. A mapping matrix of the size of 114×12 was computed indicating the level of involvement of each of the 114 cortical areas in each of the twelve cognitive components.

The connectome

The human structural connectome was reconstructed using T1-weighted and diffusion-weighted MRI (dMRI) data of 487 subjects (age [mean \pm standard deviation]: 29.8 ± 3.4 years old) from the Human Connectome Project [58]. White matter tracts were reconstructed from the dMRI data using deterministic tractography [59, 60]. The 18 sets of $b = 0$ volumes were averaged, and the 270 diffusion images were realigned and corrected for motion and gradient-induced distortions [61]. The diffusion profile within each voxel was reconstructed using generalized q-sampling imaging (GQI) [62] and deterministic tractography was used to reconstruct white matter tracts, performing fiber assignment using the Fiber Assignment by Continuous Tracking (FACT) algorithm [63]. Eight streamline seeds were started per voxel and

tracking was stopped if the streamline reached a voxel of low preferred diffusion direction [fractional anisotropy (FA) < 0.1], exited the gray matter/white matter mask, or made a sharp turn (> 45°). Next, a structural connectome was generated from the set of reconstructed tractography streamlines and the cortical parcellation of each subject. The number of streamlines (NOS), streamline density (SD), mean fractional anisotropy (FA) were extracted and used as the weight of structural connectivity between cortical regions. A group-averaged binary network was formed by averaging values across subjects for connections observed in more than 60% of the subjects [64]. Nodal degree was calculated as the number of connections linked to each region. Nodal strength was computed for each cortical region by summing up the weights of all connections linked to that region.

A functional connectome map was constructed using resting-state fMRI data. Data were realigned and co-registered with the T1-weighted image to overlap with the cortical parcellation maps. Blood oxygenation level-dependent (BOLD) time series were corrected for linear trends, as well as global nuisance covariance, including six head motion parameters and mean signals of white matter and the ventricles. We performed band-pass filtering (0.01-0.1 Hz) together with motion scrubbing to minimize the influence of head-motion [65]. Regional time series were computed from the preprocessed fMRI data by averaging the time-series of all voxels within a cortical region. Interregional functional connectivity was calculated as Pearson's correlation coefficient between the time series of cortical regions. Negative correlations were excluded, and correlation coefficients were transformed to z-scores using Fisher transformation. A group-level functional network was constructed by averaging z-scores across all subjects. Nodal strength of functional connectivity was subsequently computed.

Resting cortical metabolism

We used data of five cortical metabolic measurements obtained from positron emission tomography (PET) imaging (tabulated and described in detail in [20]), including the vertebral metabolic rate for oxygen (CMRO₂), cerebral metabolic rate for glucose (CMRGlu), glycolytic index (GI), oxygen-glucose index (OGI), and cerebral blood flow (CBF). Data of metabolic measurements were given based on the Brodmann atlas (BA) and were transformed to the DK-114 atlas using the following steps. We extracted the MNI coordinates of all gray matter voxels of the fsaverage subject in FreeSurfer and mapped MNI coordinates to the Talairach atlas to obtain the an-

notation of BA area, using the BioImage Suite (<https://bioimagesuiteweb.github.io/webapp/mni2tal.html>) [66]. With the resulting annotations, we assigned the value of metabolic measurements to each voxel in the gray matter of the fsaverage subject. For each region of the DK atlas mean values were calculated per metabolic measurement, among all gray matter voxels within that region, yielding a 114×5 data matrix of metabolic measurements.

Cortical expansion

A brain map of evolutionary cortical expansion was obtained from [12], where T1-weighted MRI data from 29 chimpanzees and 50 humans were utilized to compute the surface area expansion of each homologous cortical region in the DK-114 atlas. The expansion was computed as $(\text{AREA}_{\text{human}} - \text{AREA}_{\text{chimpanzee}}) / \text{AREA}_{\text{chimpanzee}}$.

Anatomical and functional involvement in brain diseases

The BrainMap database (<http://www.brainmap.org/>) containing published functional and structural neuroimaging experiments with coordinate-based results was used to extract brain mappings of a wide range of psychiatric and neurologic disorders. MNI coordinates related to 22 brain disorders from the voxel-based morphometry (VBM) database (Number of experiments = 46 ± 48 , ranges from 4 to 206) and of 16 brain disorders from the functional database (Number of experiments = 69 ± 69 , ranges from 13 to 267) were obtained using Sleuth [21–23, 67]. Meta-analysis was performed for each disorder using activation likelihood estimation (ALE) implemented in GingerALE [68, 69]. ALE maps were registered to the MNI152 template in the FreeSurfer space using FLIRT [70, 71]. Furthermore, ALE scores of gray matter voxels within each DK region were averaged, resulting in two matrices with the size of 114×22 and 114×16 , indicating the cortical involvement of each disorder anatomically or functionally.

Neurosynth data

The Neurosynth database includes brain images generated using text-mining, meta-analysis and machine-learning techniques on MRI studies mapping the involvement of brain regions in a large variety of cognitive tasks and states (<http://www.neurosynth.org>) [25]. We downloaded the brain mappings of 749 terms, with voxel-wise z-scores representing the extent to which a voxel is associated with the corresponding term

(FDR corrected, $q < 0.01$). Images of 457 terms with abstract meaning (e.g., "association"), or with names of brain anatomy (e.g., "putamen") were excluded. Resulting 292 images were registered to the MNI152 template in FreeSurfer space using FLIRT [70, 71]. Average z-score within each cortical region of the DK-114 atlas was computed, resulting in a 114×292 matrix to indicate the cortical involvement in each term.

Statistics

We assessed the association between the cortical gene expression profile and the pattern of neuroimaging phenotypes using linear regression with cortical volume of each region (measured on the basis of the fsaverage subject in FreeSurfer) taken as a covariate:

$$Y_i = \beta_0 + \beta_1 X_j + \beta_2 \text{cov} + \varepsilon$$

where Y_i indicates the normalized gene expression profile of gene i or the averaged profile of a *gene set* i , X_j the standardized cortical profile of neuroimaging phenotype j , and cov the standardized covariate. Standardizing is performed by dividing each value by one standard deviation. The standardized regression coefficient β_1 and the corresponding p -value were obtained.

We tested whether the observed association was specific to the given gene(s). Per gene and per brain imaging phenotype, we performed a z-test to examine whether the observed β_1 (i.e., the effect size) was higher than the average effect size observed for all genes. Bootstrapping was performed to estimate the mean (μ) and standard deviation (σ) of effect sizes, by sampling 10,000 random genes/gene sets with the same size of the input gene set and re-performing linear regression.

$$z = \frac{\beta_1 - \mu}{\sigma}$$

where μ , σ indicate the mean and standard deviation of β_1 over 10,000 random genes/gene sets. A two-sided p -value was computed as follows:

$$p = 2\Phi(|-z|)$$

where Φ is the standard normal cumulative distribution function. Moreover, we performed z-test to examine whether the observed β_1 (i.e., the effect size) was larger than the mean effect size of BRAIN genes, a pool of 2979 genes significantly more

expressed in brain tissues compared to other available body sites [12]. Z -score and p -value were computed in a similar way.

GAMBA also validates whether the observed association is dependent on anatomical relationships between brain regions. To this end, the observed β_1 is compared to the mean β_1 generated by 1,000 permutations, in which gene expression data is rebuilt using randomized brain parcellations. Randomized parcellations were obtained by spinning the reconstructed sphere of the real brain parcellation with random angles [26].

GAMBA includes Bonferroni and FDR correction with adjustable thresholds to correct for multiple comparisons in the analysis of each imaging modality. Given a gene or a gene set, standardized β or z -score of all above neuroimaging phenotypes are displayed. Results reaching significance are shown in darker colors, otherwise in lighter colors. All results and raw data are available for downloading.

Gene modularity detection

When more than one gene is examined GAMBA computes a co-expression network calculating the Pearson's correlation of gene expression profiles between all gene pairs. Negative correlations are set to 0, as the subsets of genes with similar transcriptional regulatory programs are of special interest [72]. Modularity detection is performed in the resultant co-expression network using the Louvain method [73]. The organization of modular organization is quantified by the modularity (Q) [74]. Permutation testing is performed by examining whether the resulted Q is different from null distributions of Q derived from random genes/BRAIN genes. The modular structure is optionally displayed in GAMBA with the provided option to further examine the expression patterns as separate gene subsets.

Acknowledgements

The work of M.P.v.d.H. was supported by an ALW open (ALWOP.179) and VIDI (452-16-015) grant from the Netherlands Organization for Scientific Research (NWO) and a Fellowship of MQ. Y.W. was supported by the China Scholarship Council (201506040039). D.P. was supported by The Netherlands Organization for Scientific Research (NWO VICI 453-14-005). Human neuroimaging data was kindly provided by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Insti-

tutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University.

References

1. Elliott, L. T. *et al.* Genome-wide association studies of brain imaging phenotypes in UK Biobank. *Nature* **562**, 210 (2018).
2. Satizabal, C. L. *et al.* Genetic architecture of subcortical brain structures in 38,851 individuals. *Nature Genetics*, 1–16 (2019).
3. Zhao, B. *et al.* Genome-wide association analysis of 19,629 individuals identifies variants influencing regional brain volumes and refines their genetic co-architecture with cognitive and mental health traits. *Nature Genetics*, 1–8 (2019).
4. Miller, K. L. *et al.* Multimodal population brain imaging in the UK Biobank prospective epidemiological study. *Nature neuroscience* **19**, 1523 (2016).
5. Shastry, B. S. SNPs: impact on gene function and phenotype. *Methods in molecular biology* **578**, 3–22 (2009).
6. Van den Heuvel, M. P., Scholtens, L. H. & Kahn, R. S. Multiscale Neuroscience of Psychiatric Disorders. *Biological Psychiatry* **86**, 512–522 (2019).
7. Hawrylycz, M. J. *et al.* An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature* **489**, 391–399 (2012).
8. Fornito, A., Arnatkeviciute, A. & Fulcher, B. D. Bridging the Gap between Connectome and Transcriptome. *Trends in Cognitive Sciences* **23**, 34–50 (2019).
9. Vértes, P. E. *et al.* Gene transcription profiles associated with inter-modular hubs and connection distance in human functional magnetic resonance imaging networks. *Philosophical Transactions of the Royal Society B: Biological Sciences* **371**, 20150362 (2016).
10. Krienen, F. M., Yeo, B. T., Ge, T., Buckner, R. L. & Sherwood, C. C. Transcriptional profiles of supragranular-enriched genes associate with corticocortical network architecture in the human brain. *Proceedings of the National Academy of Sciences* **113**, E469–E478 (2016).
11. Romero-Garcia, R. *et al.* Structural covariance networks are coupled to expression of genes enriched in supragranular layers of the human cortex. *Neuroimage* **171**, 256–267 (2018).
12. Wei, Y. *et al.* Genetic mapping and evolutionary analysis of human-expanded cognitive networks. *Nature communications* **10**, 1–11 (2019).

13. Romero-Garcia, R., Warrier, V., Bullmore, E. T., Baron-Cohen, S. & Bethlehem, R. A. Synaptic and transcriptionally downregulated genes are associated with cortical thickness differences in autism. *Molecular psychiatry* **24**, 1053–1064 (2019).
14. Romme, I. A., de Reus, M. A., Ophoff, R. A., Kahn, R. S. & van den Heuvel, M. P. Connectome Disconnectivity and Cortical Gene Expression in Patients With Schizophrenia. *Biological Psychiatry* **81**, 495–502 (2017).
15. Anderson, K. M. *et al.* Convergent molecular, cellular, and neural signatures of major depressive disorder. *bioRxiv* (2020).
16. Watanabe, K., Taskesen, E., Van Bochoven, A. & Posthuma, D. Functional mapping and annotation of genetic associations with FUMA. *Nature communications* **8**, 1826 (2017).
17. Anderson, K. M. *et al.* The transcriptional landscape of cortical interneurons underlies in-vivo brain function and schizophrenia risk. *bioRxiv*, 481036 (2018).
18. Thomas Yeo, B. *et al.* The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of neurophysiology* **106**, 1125–1165 (2011).
19. Yeo, R. A. *et al.* Graph Metrics of Structural Brain Networks in Individuals with Schizophrenia and Healthy Controls: Group Differences, Relationships with Intelligence, and Genetics. *Journal of the International Neuropsychological Society : JINS* **22** **2**, 240–9 (2016).
20. Vaishnavi, S. N. *et al.* Regional aerobic glycolysis in the human brain. *Proceedings of the National Academy of Sciences of the United States of America* **107** **41**, 17757–62 (2010).
21. Fox, P. T. & Lancaster, J. L. Mapping context and content: the BrainMap model. *Nature Reviews Neuroscience* **3**, 319–321 (2002).
22. Fox, P. T. *et al.* BrainMap taxonomy of experimental design: description and evaluation. *Human brain mapping* **25** **1**, 185–98 (2005).
23. Laird, A. R., Lancaster, J. L. & Fox, P. T. BrainMap - The social evolution of a human brain mapping database. *Neuroinformatics* **3** **1**, 65–78 (2005).
24. De Lange, S. C. *et al.* Shared vulnerability for connectome alterations across psychiatric and neurological brain disorders. *Nature Human Behaviour*, 1–11 (2019).
25. Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C. & Wager, T. D. Large-scale automated synthesis of human functional neuroimaging data. *Nature methods* **8**, 665 (2011).
26. Alexander-Bloch, A. *et al.* On testing for spatial correspondence between maps of human brain structure and function. *NeuroImage* **178**, 540–551 (2018).
27. Liu, C.-C., Kanekiyo, T., Xu, H. & Bu, G. Apolipoprotein E and Alzheimer disease: risk, mechanisms and therapy. *Nature Reviews Neurology* **9**, 106–118 (2013).

28. Catani, M., Dell'Acqua, F. & de Schotten, M. T. A revised limbic system model for memory, emotion and behaviour. *Neuroscience Biobehavioral Reviews* **37**, 1724–1737 (2013).
29. Nestor, P. J., Fryer, T. D., Smielewski, P. & Hodges, J. R. Limbic hypometabolism in Alzheimer's disease and mild cognitive impairment. *Annals of neurology* **54** **3**, 343–51 (2003).
30. Grove, J. *et al.* Identification of common genetic risk variants for autism spectrum disorder. *Nature Genetics* **51**, 431–444 (2019).
31. Edition, F. *et al.* Diagnostic and statistical manual of mental disorders. *Am Psychiatric Assoc* (2013).
32. Zeng, H. *et al.* Large-Scale Cellular-Resolution Gene Profiling in Human Neocortex Reveals Species-Specific Molecular Signatures. *Cell* **149**, 483–496 (2012).
33. Scholtens, L. H. & van den Heuvel, M. P. Multimodal Connectomics in Psychiatry: Bridging Scales From Micro to Macro. *Biological psychiatry. Cognitive neuroscience and neuroimaging* **3** **9**, 767–776 (2018).
34. Jansen, I. E. *et al.* Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk. *Nature Genetics* **51**, 404–413 (2019).
35. Karch, C. M. & Goate, A. M. Alzheimer's Disease Risk Genes and Mechanisms of Disease Pathogenesis. *Biological Psychiatry* **77**, 43–51 (2015).
36. Akram, A., Schmeidler, J. M., Katsel, P. L., Hof, P. R. & Haroutunian, V. Association of ApoE and LRP mRNA levels with dementia and AD neuropathology. *Neurobiology of Aging* **33**, 628.e1–628.e14 (2012).
37. Linnertz, C., Anderson, L., Gottschalk, W. K., Crenshaw, D. G. & Chiba-Falek, O. The cis-regulatory effect of an Alzheimer's disease-associated poly-T locus on expression of TOMM40 and apolipoprotein E genes. *Alzheimer's & Dementia* **10**, 541–551 (2014).
38. Matsui, T., Ingelsson, M., Fukumoto, H., Ramasamy, K. & Hyman, B. T. Expression of APP pathway mRNAs and proteins in Alzheimer's disease. *Brain Research* **1161**, 116–123 (2007).
39. Agosta, F. *et al.* Apolipoprotein E epsilon4 is associated with disease-specific effects on brain atrophy in Alzheimer's disease and frontotemporal dementia. *Proceedings of the National Academy of Sciences of the United States of America* **106** **6**, 2018–22 (2009).
40. Cacciaglia, R. *et al.* Effects of APOE-ε4 allele load on brain morphology in a cohort of middle-aged healthy individuals with enriched genetic risk for Alzheimer's disease. *Alzheimer's & dementia : the journal of the Alzheimer's Association* **14** **7**, 902–912 (2018).
41. Manning, E. N. *et al.* APOE ε4 is associated with disproportionate progressive hippocampal atrophy in AD. *PloS one* **9** (2014).

42. Ranlund, S. *et al.* Associations between polygenic risk scores for four psychiatric illnesses and brain structure using multivariate pattern recognition. *NeuroImage: Clinical* **20**, 1026–1036 (2018).
43. Liu, Z. *et al.* Brain annotation toolbox: exploring the functional and genetic associations of neuroimaging results. *Bioinformatics* (2019).
44. Miller, J. A. *et al.* Transcriptional landscape of the prenatal human brain. *Nature* **508**, 199–206 (2014).
45. Tebbenkamp, A. T. N., Willsey, A. J., State, M. W. & Sestan, N. The developmental transcriptome of the human brain: implications for neurodevelopmental disorders. *Current opinion in neurology* **27** **2**, 149–56 (2014).
46. Wang, D. *et al.* Comprehensive functional genomic resource and integrative model for the human brain. *Science* **362**, eaat8464 (2018).
47. Johnson, M. B. *et al.* Functional and evolutionary insights into human brain development through global transcriptome analysis. *Neuron* **62**, 494–509 (2009).
48. McColgan, P. *et al.* Brain regions showing white matter loss in Huntington's disease are enriched for synaptic and metabolic genes. *Biological psychiatry* **83**, 456–465 (2018).
49. Arnatkeviciute, A., Fulcher, B. D. & Fornito, A. A practical guide to linking brain-wide gene expression and neuroimaging data. *NeuroImage* **189**, 353–367 (2019).
50. French, L. & Paus, T. A FreeSurfer view of the cortical transcriptome generated from the Allen Human Brain Atlas. *Front. Neurosci.* **9**, 323 (2015).
51. Cammoun, L. *et al.* Mapping the human connectome at multiple scales with diffusion spectrum MRI. *Journal of Neuroscience Methods* **203**, 386–397 (2012).
52. Desikan, R. S. *et al.* An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage* **31**, 968–980 (2006).
53. Fischl, B. FreeSurfer. *NeuroImage* **62**. 20 YEARS OF fMRI, 774–781 (2012).
54. Fonov, V. S. *et al.* Unbiased average age-appropriate atlases for pediatric studies. *NeuroImage* **54**, 313–327 (2011).
55. Fonov, V., Evans, A., McKinstry, R., Almlí, C. R. & Collins, D. Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. *NeuroImage* **47** (2009).
56. Power, J. D. *et al.* Functional Network Organization of the Human Brain. *Neuron* **72**, 665–678 (2011).
57. Smith, S. M. *et al.* Correspondence of the brain's functional architecture during activation and rest. *Proceedings of the National Academy of Sciences* **106**, 13040–13045 (2009).

58. Essen, D. C. V. *et al.* The WU-Minn Human Connectome Project: An overview. *NeuroImage* **80**, 62–79 (2013).
59. Van den Heuvel, M. P., Scholtens, L. H., de Reus, M. A. & Kahn, R. S. Associated Microscale Spine Density and Macroscale Connectivity Disruptions in Schizophrenia. *Biological Psychiatry* **80**. Molecular Mechanisms of Microcircuit Dysfunction in Schizophrenia, 293–301 (2016).
60. Wei, Y., Scholtens, L. H., Turk, E. & van den Heuvel, M. P. Multiscale examination of cytoarchitectonic similarity and human brain connectivity. *Network Neuroscience* **3**, 124–137 (2019).
61. Andersson, J. L. & Skare, S. A Model-Based Method for Retrospective Correction of Geometric Distortions in Diffusion-Weighted EPI. *NeuroImage* **16**, 177–199 (2002).
62. Yeh, F., Wedeen, V. J. & Tseng, W. I. Generalized q -Sampling Imaging. *IEEE Transactions on Medical Imaging* **29**, 1626–1635 (2010).
63. Mori, S., Crain, B. J., Chacko, V. P. & Van Zijl, P. C. M. Three-dimensional tracking of axonal projections in the brain by magnetic resonance imaging. *Annals of Neurology* **45**, 265–269 (1999).
64. De Reus, M. A. & van den Heuvel, M. P. Estimating false positives and negatives in brain networks. *NeuroImage* **70**, 402–409 (2013).
65. Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L. & Petersen, S. E. Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *NeuroImage* **59**, 2142–2154 (2012).
66. Lacadie, C., Fulbright, R. K., Rajeevan, N., Constable, R. T. & Papademetris, X. More accurate Talairach coordinates for neuroimaging using non-linear registration. *NeuroImage* **42**, 717–725 (2008).
67. Vanasse, T. J. *et al.* BrainMap VBM: An environment for structural meta-analysis. *Human Brain Mapping* **39**, 3308–3325 (2018).
68. Eickhoff, S. B. *et al.* Coordinate-based activation likelihood estimation meta-analysis of neuroimaging data: a random-effects approach based on empirical estimates of spatial uncertainty. *Human brain mapping* **30** **9**, 2907–26 (2009).
69. Turkeltaub, P. E. *et al.* Minimizing within-experiment and within-group effects in Activation Likelihood Estimation meta-analyses. *Human brain mapping* **33** **1**, 1–13 (2012).
70. Jenkinson, M. & Smith, S. A global optimisation method for robust affine registration of brain images. *Medical Image Analysis* **5**, 143–156 (2001).
71. Jenkinson, M., Bannister, P., Brady, M. & Smith, S. Improved Optimization for the Robust and Accurate Linear Registration and Motion Correction of Brain Images. *NeuroImage* **17**, 825–841 (2002).

72. Stuart, J. M., Segal, E., Koller, D. & Kim, S. K. A gene-coexpression network for global discovery of conserved genetic modules. *Science* **302** **5643**, 249–55 (2003).
73. Blondel, V. D., Guillaume, J.-L., Lambiotte, R. & Lefebvre, E. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment* **2008** (2008).
74. Newman, M. E. Finding community structure in networks using the eigenvectors of matrices. *Physical review E* **74**, 036104 (2006).