









## TECHNICAL REPORT

# Statistical testing in transcriptomic-neuroimaging studies: A how-to and evaluation of methods assessing spatial and gene specificity

Yongbin Wei<sup>1</sup>  | Siemon C. de Lange<sup>1,2</sup>  | Rory Pijnenburg<sup>1</sup>  |  
Lianne H. Scholtens<sup>1</sup>  | Dirk Jan Ardesch<sup>1</sup>  | Kyoko Watanabe<sup>1</sup>  |  
Danielle Posthuma<sup>1,3</sup>  | Martijn P. van den Heuvel<sup>1,3</sup> 

<sup>1</sup>Department of Complex Trait Genetics, Center for Neurogenomics and Cognitive Research, Amsterdam Neuroscience, Vrije Universiteit Amsterdam, Amsterdam, The Netherlands

<sup>2</sup>Department of Sleep and Cognition, Netherlands Institute for Neuroscience (NIN), Royal Netherlands Academy of Arts and Sciences, Amsterdam, The Netherlands

<sup>3</sup>Department of Child and Adolescent Psychiatry and Psychology, Section Complex Trait Genetics, Amsterdam Neuroscience, Vrije Universiteit Medical Center, Amsterdam UMC, Amsterdam, The Netherlands

## Correspondence

Martijn P. van den Heuvel, De Boelelaan 1085, W&N B-651, 1081 HV Amsterdam, The Netherlands.

Email: martijn.vanden.heuvel@vu.nl

## Funding information

H2020 European Research Council, Grant/Award Numbers: 101001062, ERC-2018-AdG GWAS2FUNC 834057; Nederlandse Organisatie voor Wetenschappelijk Onderzoek, Grant/Award Numbers: ALW open (ALWOP.179), VICI (453-14-005), VIDI (452-16-015), NWO Gravitation (024.004.012); ZonMw, Grant/Award Number: 09120011910032

## Abstract

Multiscale integration of gene transcriptomic and neuroimaging data is becoming a widely used approach for exploring the molecular underpinnings of large-scale brain organization in health and disease. Proper statistical evaluation of determined associations between imaging-based phenotypic and transcriptomic data is key in these explorations, in particular to establish whether observed associations exceed “chance level” of random, nonspecific effects. Recent approaches have shown the importance of statistical models that can correct for spatial autocorrelation effects in the data to avoid inflation of reported statistics. Here, we discuss the need for examination of a second category of statistical models in transcriptomic-neuroimaging analyses, namely those that can provide “gene specificity.” By means of a couple of simple examples of commonly performed transcriptomic-neuroimaging analyses, we illustrate some of the potentials and challenges of transcriptomic-imaging analyses, showing that providing gene specificity on observed transcriptomic-neuroimaging effects is of high importance to avoid reports of nonspecific effects. Through means of simulations we show that the rate of reported nonspecific effects (i.e., effects that cannot be specifically linked to a specific gene or gene-set) can run as high as 60%, with only less than 5% of transcriptomic-neuroimaging associations observed through ordinary linear regression analyses showing both spatial and gene specificity. We provide a discussion, a tutorial, and an easy-to-use toolbox for the different options of null models in transcriptomic-neuroimaging analyses.

## KEYWORDS

coexpression, gene expression, gene specificity, neuroimaging, null model

## 1 | INTRODUCTION

A fast-growing number of imaging-genetic studies have started to investigate associations between spatial patterns of gene transcriptome data and macroscale imaging-derived brain phenotypes

(Anderson et al., 2018; Burt et al., 2018; Fornito, Arnatkeviciute, & Fulcher, 2019; van den Heuvel, Scholtens, & Kahn, 2019; van den Heuvel & Yeo, 2017; Krienen, Yeo, Ge, Buckner, & Sherwood, 2016; Richiardi et al., 2015; Wang, Belgard, et al., 2015). Whole-brain transcriptome data, such as the extensive Allen Human Brain Atlas (AHBA)

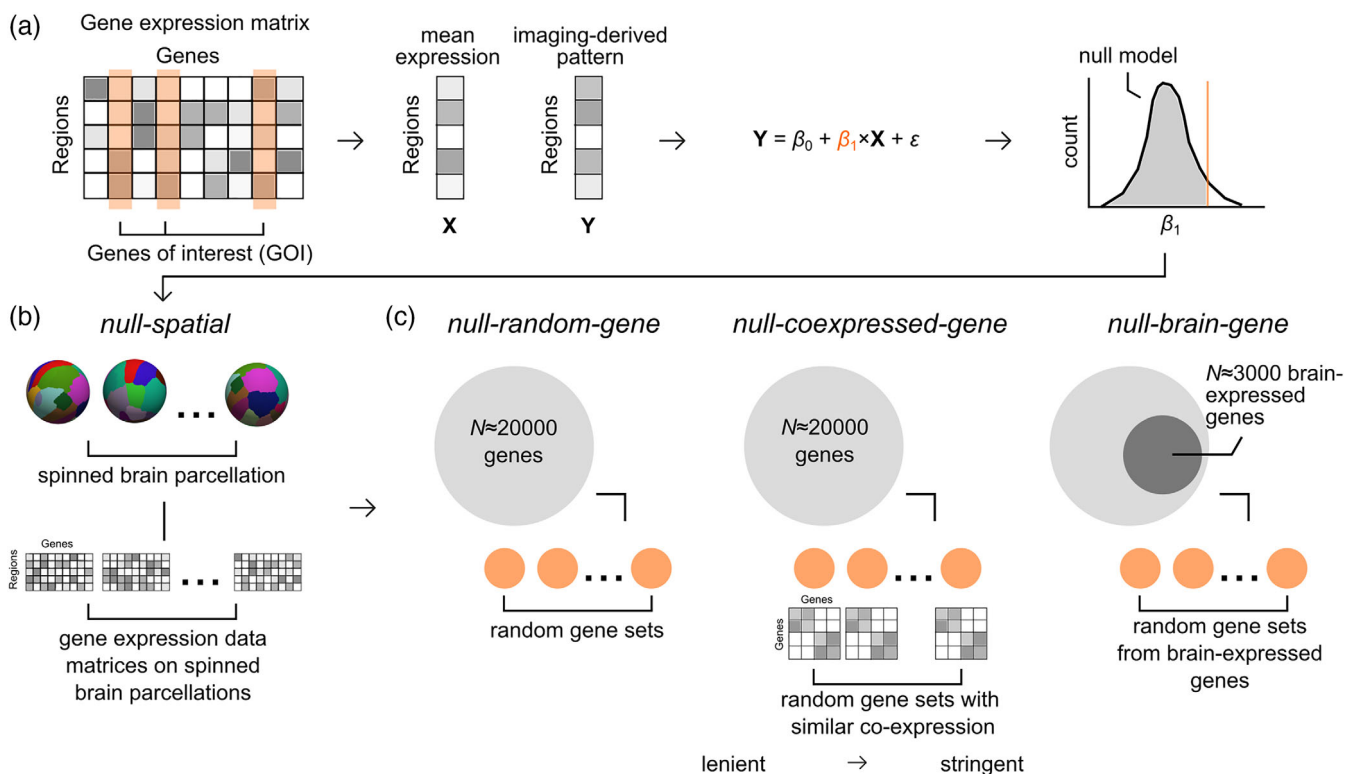
This is an open access article under the terms of the Creative Commons Attribution-NonCommercial License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

© 2021 The Authors. *Human Brain Mapping* published by Wiley Periodicals LLC.

(Hawrylycz et al., 2012), the BrainSpan developing human brain atlas (Kang et al., 2011) and the PsychEncode consortium (Wang et al., 2018), serve as an invaluable quantitative reference to assess transcriptomic-neuroimaging associations. Examples of such studies include reports on genes related to oxidative metabolism to display transcriptional profiles in a similar spatial pattern as the degree of inter-module long-distance connectivity (FC) (Vertes et al., 2016), reports on genes enriched for neuronal and synaptic connectivity to show transcriptional profiles that capture the architecture of brain functional networks (Krienen et al., 2016; Romero-Garcia et al., 2018), and reports of transcriptional profiles of risk genes for brain disorders (e.g., taken from genome-wide association studies (GWAS)) to display overlap with patterns of disorder-specific brain changes, for example for schizophrenia (Romme, de Reus, Ophoff, Kahn, & van den Heuvel, 2017), autism spectrum disorder (Romero-Garcia, Warrier, Bullmore, Baron-Cohen, & Bethlehem, 2019), major depressive disorder (Anderson et al., 2020), among others. This rapidly growing number of powerful transcriptomic-neuroimaging explorations stresses the need for proper statistical evaluation methods that can deal with

the complicating covariance seen between genes and answer the question, for example, “whether observed results are specific to the a-priori selected genes of interest, or reflect a more common effect present across more genes and not related to the genes of interest?”

One commonly examined question (and related hypothesis) in transcriptomic-neuroimaging studies is summarized as follows: “Does the spatial pattern of our gene/gene-set of interest X across brain areas match the spatial pattern that we observe across brain regions for brain phenotype Y”? For example, X here can be gene Apolipoprotein E (APOE), tested against the brain phenotype of cortical atrophy as measured in Alzheimer's disease (AD); or X can be a gene-set enriched in a specific class of cells or cortical layer, such as genes enriched in supragranular layers, and Y the level of regional functional connectivity of cortical areas. After preprocessing and normalization of the transcriptomic data (Arnatkevičiūtė, Fulcher, & Fornito, 2019; Markello et al., 2021), values for gene(s) X are selected from the transcriptomic dataset and the level of cortical expression across areas is correlated with the cortical profile of Y across the same set of regions (Figure 1a). When a significant correlation is observed between X and



**FIGURE 1** Approaches for statistical testing of overlapping patterns of gene transcription and imaging-derived brain phenotypes. (a) Permutation test procedure. The expression profile (X) of a user-defined gene/gene-set of interest (GOI) is computed. Association between the expression profile and the imaging-derived pattern (Y) is assessed using linear regression. Permutation testing is used to examine whether the observed  $\beta_1$  is larger than null distributions of  $\beta_1$  derived from null models. Different statistical null models are possible: (b) Correction for spatial effects. The “null-spatial” model (Alexander-Bloch et al., 2018) is proposed as a method to correct for spatial autocorrelation of the observed associations. Randomized brain parcellations are obtained by spinning the inflated sphere of the real brain parcellation (1,000 randomizations used in the current study). Gene expression data matrices are then rebuilt using these randomized brain parcellations. The rebuilt gene expression data are used to re-evaluate transcriptomic-neuroimaging associations to generate null distributions. (c) Examination of gene specificity. Three null models (from more lenient to more stringent) are used to examine the gene specificity. “null-random-gene” model: random genes are selected from all  $N \approx 20,000$  genes included in AHBA. “null-coexpressed-gene” model: random genes that conserve the mean coexpression level of the original genes are selected from all  $N \approx 20,000$  genes included in AHBA. “null-brain-gene” model: random genes are selected from a subset of genes (2,711 genes) that show up-regulated expression levels in brain tissues in contrast to other body sites

Y (often statistically evaluated by means of simple linear regression), the observed correlation is interpreted as an indication of an association between the gene/gene-set of interest and examined phenotype: for example, “our analysis supports our hypothesis of the level of normative expression of *APOE* across brain regions to be associated with the pattern of cortical atrophy, with regions showing a high level of (normative) expression of *APOE* arguably displaying a higher level of involvement or vulnerability in the disease process.” And this makes sense: *APOE*  $\epsilon 4$  is a well-known risk gene for AD (Cacciaglia et al., 2018), so a high level of normative expression of *APOE* could indeed in some way be related to cortical atrophy in the disease condition.

When we find and discuss such an observed association, we are however often quickly tempted to interpret this finding as that it is specifically *APOE*—and not some other gene—that shows this pattern. This is however not directly tested when we test gene X against phenotypic pattern Y. What if another gene would also show this pattern? This would be informative if this gene is also linked to AD (for example, other risk genes identified in AD's GWAS (Jansen et al., 2019)). But what if a gene that is not related to AD would also show this effect? Or when the majority or a considerable part of all genes for which data are available in the transcriptomic dataset would show a similar association? A thorough further investigation into the level of specificity of our shown association to gene or gene-set X would make our claim much stronger and show that it is indeed X that shows this effect, and not just any subset of genes. In contrast, if many other genes would also show the same pattern, one could argue that this makes our observed correlation between X and Y less powerful, and perhaps of less interest for further functional follow-up.

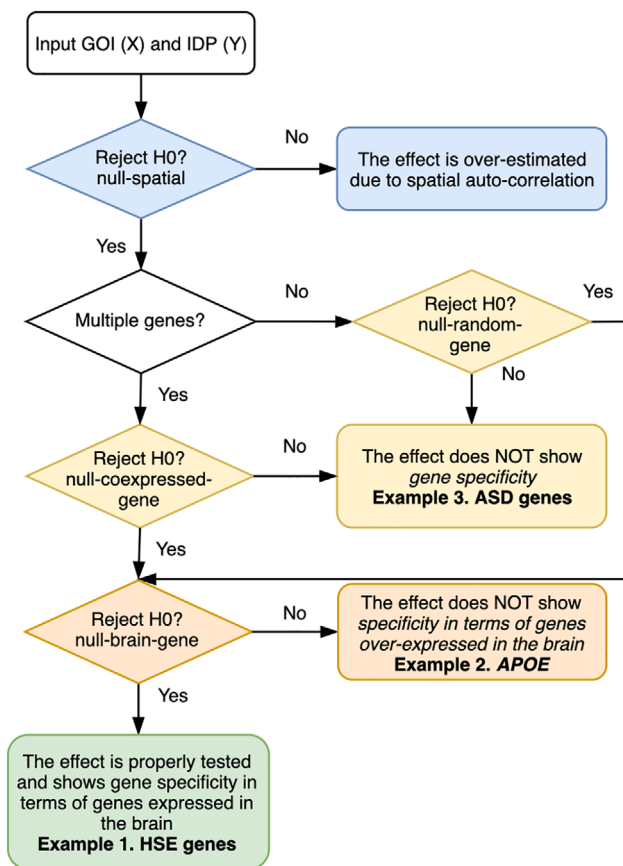
As recently pointed out (Alexander-Bloch et al., 2018; Arnatkevičiūtė et al., 2019; Burt, Helmer, Shinn, Anticevic, & Murray, 2020; Fulcher, Arnatkevičiūtė, & Fornito, 2021; Markello & Misić, 2021) a first important point in this context is that the commonly used linear regression model assumes independent observations. This is however not always the case for brain gene expression data, as expression levels of neighboring regions tend to be often strongly correlated. It has thus been proposed to take these spatial effects (referred to as spatial autocorrelation (Fulcher et al., 2021; Markello & Misić, 2021)) into account by means of a spatially constrained null model (Markello & Misić, 2021), which generates surrogate brain maps based on the parameterized variogram model (Burt et al., 2020) or based on spatial permutations (Alexander-Bloch et al., 2018) (Figure 1b). Implementations of such null models strongly reduce false-positive findings (Fulcher et al., 2021; Markello & Misić, 2021), showing the importance of correcting for spatial autocorrelation effects in transcriptomic-neuroimaging analyses.

What is however not yet answered by means of spatially constrained null models is the importance of evaluating whether an observed correlating transcriptomic-neuroimaging pattern goes beyond effects that one can expect from taking any other gene or gene-set from the data, that is, to what extent is our effect really

unique to our gene(s)-of-interest X or are we looking at a much more general effect? To address the notion of “gene specificity,” statistical comparisons of our effect for gene X to effects derived from other genes are needed. Several options are proposed, the most simple one being to compare the effect of gene X against a null distribution of effects based on selecting random genes from the pool of all  $N \approx 20,000$  genes (Romme et al., 2017) (Figure 1c). Such a “random gene” model may indeed provide an important first indication of whether or not our observed association with Y is relatively unique for our gene(s)-of-interest X or whether it can easily be found for many other genes (resulting in a pseudo  $p > .05$ ). The random-gene null model can serve as a first simple evaluation of whether or not our observed effect has some level of specificity or not.

As mentioned (Fulcher et al., 2021; Wei et al., 2019) the random-gene null model may however not always give us enough information about the specificity of our effect to our gene(s) of interest. It has been noted that the simple random-gene null model ignores common levels of coexpression present among genes (Fulcher et al., 2021). The genes in a gene-set X (taken from a GWAS, or a Gene Ontology (GO) pathway that we want to study) typically show a high level of coexpression among them. If we just compare the gene-expression pattern from the genes in X to a random set X' of randomly selected genes across all  $N \approx 20,000$  genes, we ignore the internal structure of X. Not incorporating information on coexpression into our null condition may lead to an over-evaluation of our effect, that is, we claim that gene-set X has a relatively unique association to Y, but we over-estimate its importance as we compare it to a (too) liberal null condition where we mix signals from nonrelated genes. This nonspecificity strikingly increases when we want to investigate a set of biologically relevant genes (i.e., when we test GO pathways (Fulcher et al., 2021)), which are generally much higher coexpressed than a set of random genes. If our hypothesis is that gene-set X shows a unique pattern of expression that significantly correlates to brain pattern Y, testing with more fine-tuned null models that take into account coexpression and biologically relevant relationships may be warranted.

Here, we evaluate and discuss proposed and used options for statistical evaluation of transcriptomic-neuroimaging associations, including the (most) commonly applied linear regression model, previously proposed null models that maintain spatial relationships, and null models that aim to provide insight into the level of gene specificity on the basis of comparing our effect-of-interest with effects that can occur among random genes. Through means of three simple examples, we point out that controlling for spatial effects is important to reduce false-positive findings (Arnatkevičiūtė et al., 2019; Burt et al., 2018; Fulcher et al., 2021), but we also show that this is not enough if we want to test a meaningful specific hypothesis: We suggest that further examinations of gene specificity using null models that account for (pseudo)random gene effects is equally important in minimizing reports of nonspecific results in transcriptomic-neuroimaging studies. We provide an overview of the several steps to consider (Figure 2), together with a toolbox to easily perform various null-model evaluations for transcriptomic-neuroimaging studies.



**FIGURE 2** Flow chart of the statistical framework that tests associations between the spatial pattern of expressions of the gene(s) of interest (GOI) and imaging-derived phenotypes (IDP)

## 2 | RESULTS

### 2.1 | Overview of results

We demonstrate the use and importance of different null models by means of discussing the analysis strategy and results of three common examples, followed by simulations of effects. We assess how different null models (i.e., null models that give spatial or gene specificity) can serve as a tool to identify possibly inflated statistical effects and separate potential spatial and gene-specific effects from more common, global effects present in the transcriptomic and neuroimaging data. In our examples, we show the usage of the different null-models with respect to gene/gene-set of interest (GOI), including genes related to cortico-cortical connectivity, AD-risk gene *APOE*, and genes associated with autism spectrum disorder (ASD). We hypothesize that the transcriptional profile of a GOI *X* relates to a brain phenotype *Y*, like the spatial pattern of brain atrophy (McColgan et al., 2018) or brain disconnectivity (Romme et al., 2017). We first test our hypothesized association between GOI *X* and the brain phenotype *Y* by means of correlation analysis and, when found significant, then further test the statistical relevance of this association using (1) spatial null models to correct for spatial autocorrelation in our data and (2) random-gene null models to assess gene specificity. We show that statistical

evaluations based on null models that maintain spatial relationships do not necessarily provide gene specificity, nor vice versa.

### 2.2 | Example 1: Genes related to cortico-cortical connectivity

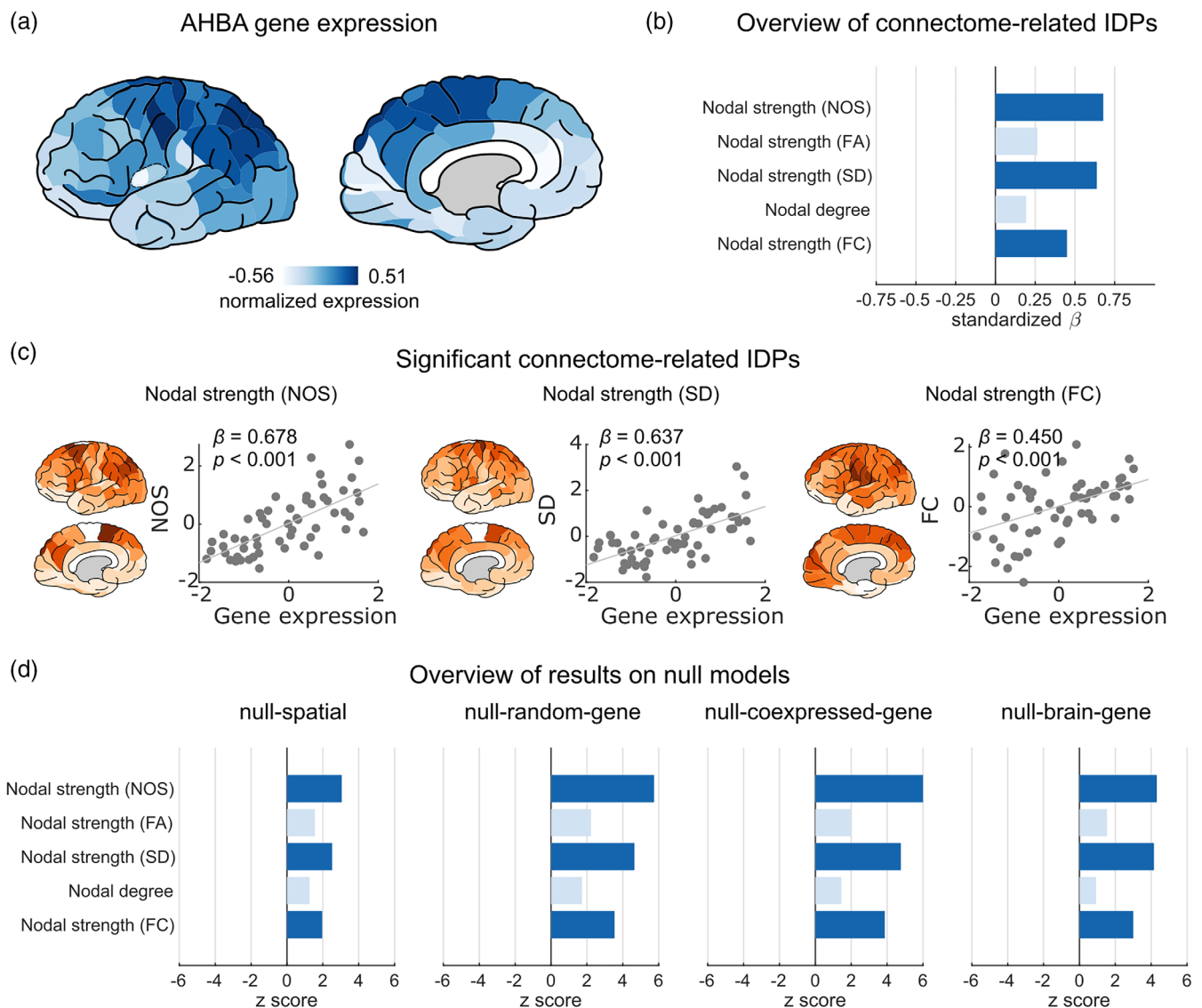
We start by (re-)assessing a commonly examined and powerful relationship between expressions of genes important for neuronal connectivity and macroscale connectome organization in the healthy brain (e.g., Krienen et al., 2016; Romero-Garcia et al., 2018). We test this by focusing on a set of 19 genes known to be enriched in supragranular layers of the human cerebral cortex (referred to as human supragranular enriched [HSE] genes), revealed by a previous study that compared gene expression profiles in different cortical layers between mice and humans (Zeng et al., 2012). These HSE genes are suggested to play a role in shaping long-range cortico-cortical connections of layer III pyramidal neurons and as such to show a spatial expression pattern that runs parallel to the organization of brain structural and functional networks (Krienen et al., 2016; Romero-Garcia et al., 2018). The transcriptional profile of HSE genes is shown in Figure 3a.

#### 2.2.1 | Transcriptomic-neuroimaging overlap: Linear regression

We first test the potential association between the spatial expression pattern of HSE genes and macroscale functional connectivity (Krienen et al., 2016). Using simple linear regression, we indeed find the expression pattern of HSE genes to be significantly associated with the pattern of connectivity strength of cortical areas—a measurement describing the extent to which a region is connected to the rest of the brain—under the assumption of “independent” gene expressions in the brain (standardized beta  $\beta = 0.678$ ,  $p < .001$  for connectivity weighted by the number of streamlines [NOS];  $\beta = 0.637$ ,  $p < .001$  for connectivity weighted by streamline density [SD]; false discovery rate [FDR] corrected  $q < 0.05$  for multiple testing across five connectome-related metrics; Figure 3b,c). A similar association can be found between the pattern of HSE gene expression and the pattern of nodal strength of the functional connectome (FC;  $\beta = 0.450$ ,  $p < .001$ ; FDR corrected; Figure 3b,c), which is in line with the notion of a strong correspondence between structural and functional connectivity (Wang, Dai, Gong, Zhou, & He, 2015).

#### 2.2.2 | Correction for spatial autocorrelation: Null-spatial model

These findings confirm a potential association between HSE gene expression and macroscale connectome organization (Krienen et al., 2016; Romero-Garcia et al., 2018), but it is worth checking whether this association is not accidentally driven by inter-regional



**FIGURE 3** Example 1: HSE gene expression and macroscale connectome properties. (a) Brain plots of normalized gene expression levels of HSE genes. (b) Overview of linear regression results between HSE gene expression profile and five imaging-derived phenotypes (IDPs) that correspond to connectome properties. Dark blue indicates significant ( $q < 0.05$ , FDR corrected across five connectome traits). (c) Scatter plots for significant correlations between HSE gene expression and nodal strength of the structural (NOS-weighted,  $\beta = 0.678$  and SD-weighted,  $\beta = 0.637$ ) and functional connectome ( $\beta = 0.450$ ). (d) Permutation testing results showing whether the observed effect size ( $\beta$  in panel B) is significantly beyond four distinct null distributions of effect sizes for null-spatial, null-random-gene, null-coexpressed-gene, and null-brain-gene models. Dark blue indicates  $q < .05$  (two-tailed z-test, FDR corrected)

auto-correlations in our data (Figure 2). As pointed out recently (Alexander-Bloch et al., 2018; Arnatkevičiūtė et al., 2019; Burt et al., 2020; Fulcher et al., 2021), the commonly used linear regression method assumes independent observations, which means no correlation between values (i.e., expression levels) of the same variable (i.e., gene(s)) across different observations (i.e., regions). However, both the transcriptomic data and neuroimaging data of neighboring regions are likely not independent (Fulcher et al., 2021) and it is thus crucial to examine whether the observed associations are biased by potentially correlated expressions of neighboring brain regions, that is, preserving spatial relationships across regions. To this end, an important null model was introduced (here referred to as “null-spatial”)

where transcriptomic samples are assigned to brain regions from randomized parcellations that are obtained by “spinning” the reconstructed sphere of the real brain parcellation (Alexander-Bloch et al., 2018), importantly preserving spatial relationships across brain regions (1,000 randomizations; see Methods, Figure 1b; we note that other variations are possible but not included here [Burt et al., 2020; Markello & Misic, 2021]). The rebuilt gene expression data matrices are used to generate a null distribution of  $\beta$ , which is used to assess whether the original effect goes beyond the null condition. Using the null-spatial model, the associations between the expression pattern of HSE genes and the pattern of structural/functional connectivity strength are again found to be significant (NOS:  $z = 3.057$ ,  $p = .002$ ;



SD:  $z = 2.516$ ,  $p = .012$ ; FC:  $z = 1.963$ ,  $p = .049$ ; FDR-corrected for multiple testing across three tests, Figure 3d). These effects thus significantly go beyond effects that one can expect using randomized brain regions with the same neighboring relationships, suggesting that the observed associations are specific to the spatial organization of brain regions.

### 2.2.3 | Gene specificity

As laid out, correcting for spatial auto-correlation however does not yet tell us whether the observed effect is specific to this set of genes or whether the same pattern might be present for all genes (including those not related to neuronal processes). We thus further want to examine whether our observed association is also gene-specific (i.e., whether it is relatively unique or can be found for many other genes). This is of particular importance as multiple examinations have pointed out general (global) posterior–anterior gradients of expression levels across brain areas (McColgan et al., 2018; Vogel et al., 2020), effects that may more reflect spatial patterns general to genes expressed in the brain, rather than reflecting patterns unique to the GOI. We argue that a next important step in the evaluation of observed transcriptomic-neuroimaging association(s) is to examine to what extent the computed linear correlations exceed effects that one could also observe when any other gene or set of genes would have been chosen (Figure 2).

#### *Null-random-gene model*

To this end, we can employ permutations to generate null distributions of effect sizes,  $\beta$ , based on gene expression profiles of same-sized gene sets randomly selected from all genes (by default 10,000 permutations are used). We refer to this used null model (Fulcher et al., 2021; Romme et al., 2017) as the “null-random-gene” model (Figure 1c). When the original  $\beta$  significantly exceeds this null model, it indicates the observed association cannot be found for all genes, but is unique for the GOI. For our HSE example, we show that the observed effect sizes for associations between HSE genes and nodal strength of structural and functional connectivity are significantly larger than effect sizes of random genes (NOS:  $z = 5.742$ ,  $p < .001$ ; SD:  $z = 4.646$ ,  $p < .001$ ; FC:  $z = 3.541$ ,  $p < .001$ ; Figure 3d).

#### *Null-coexpressed-gene model*

We (Wei et al., 2019) and others (Fulcher et al., 2021) have earlier noted that when a set of genes (instead of a single gene) is examined, the coexpression level within a set of genes may easily set apart the average expression profile of this set of genes from a set of randomly selected genes. Genes are organized into biological pathways and systems (Ashburner et al., 2000) and our GOI likely contains high biological overlap as the GOI is usually selected as genes associated with the same trait. In contrast, randomly selected sets of genes describe an intersection across less similar biological pathways (see details in Supplementary Results, Appendix S1). The GOI therefore may show high levels of covariance (i.e., coexpression), something that is not

conserved in a random mixture of genes (Fulcher et al., 2021). Therefore, we argue that it would be more informative to further include a null model for stricter statistical evaluation, comparing the observed effect size to the null distribution of effect sizes yielded by random genes conserving the same level of coexpression as the input GOI (referred to as “null-coexpressed-gene”; Figure 1c). For our HSE example, we use simulated annealing to look for random genes with the mean coexpression levels converging to those of HSE genes (see Methods). Application of this null model still shows significant associations between HSE gene expression and connectivity strength (NOS:  $z = 6.002$ ,  $p < .001$ ; SD:  $z = 4.763$ ,  $p < .001$ ; FC:  $z = 3.869$ ,  $p < .001$ ; Figure 3d).

#### *Null-brain-gene model*

We anticipate that the strongest usability of transcriptomic-neuroimaging examinations is to link brain transcriptomic data (like AHBA) to brain phenotypes (like MRI measurements). Most of such examinations will thus be centered on testing a GOI pre-selected based on the findings from previous brain-related studies, for example, GWAS of a disease phenotype such as schizophrenia (Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014) or Alzheimer's disease (Jansen et al., 2019), a specific pathway related to neuronal properties (Kepecs & Fishell, 2014) as in our first HSE example, genetic variants related to brain volume (Jansen et al., 2020), and so on. The result of such pre-selected GOI is that most of them will likely be related to processes related to the brain, and as such reflect genes likely over-expressed in brain tissue. As a consequence, we argue that in these cases a null condition generated by randomly selecting genes from the total set of  $N \approx 20,000$  genes (including genes related to all body processes, certainly not only to the brain) is not fair and too liberal, as the null-condition will include a body of background genes not, or less, expressed in brain tissue. Including such genes into the null-condition will lead to a liberal null-distribution of effects and with that a possible overestimation of the original effect (de Leeuw, Stringer, Dekkers, Heskes, & Posthuma, 2018).

To avoid this and to be able to make more specific claims on our GOI, we advise also testing observed transcriptomic-neuroimaging associations by means of a null model that tests whether the observed effect size is larger than the null distribution of effect sizes based on background genes that are more closely related to the GOI (i.e., here brain tissue). For our neuroimaging analysis, we refer to this null model as “null-brain-gene” (Figure 1c) (Wei et al., 2019), with random genes now selected from the pool of genes that are significantly (more) expressed in brain tissues in contrast to other body sites. This selection can be made by using an external resource like the GTEx database (Consortium, G. TEx, 2015), containing gene expression data of all sorts of body tissues, including the brain (2,711 brain-expressed genes selected by  $q < 0.05$ , FDR correction, one-sided two-sample  $t$ -test comparing brain tissues and other body sites). Alternative less stringent selections can be made when the GOI is not particularly over-expressed in the brain, for example by selecting background genes that are over- or similarly expressed in brain tissue as compared

to other body sites (see Supplemental Methods, Appendix S1 for these alternative selections). In our example, using a stricter null-brain model based on brain-over-expressed genes confirms that HSE genes and connectome metrics are stronger associated than expected for randomly selected brain genes (NOS:  $z = 4.311$ ,  $p < .001$ ; SD:  $z = 4.151$ ,  $p < .001$ ; FC:  $z = 2.997$ ,  $p = .003$ ; Figure 3d).

#### Null-brain-coexpression model

As a further step, one can merge the null-coexpressed-gene model and null-brain-gene model, investigating a more integrated null model that tests whether the observed effect size is larger than the null distribution of effect sizes based on brain-expressed genes with similar coexpression level conserved (referred to as “null-brain-coexpression model”). Using this stringent null model still shows significant associations between HSE gene expression and connectivity strength (NOS:  $z = 4.304$ ,  $p < .001$ ; SD:  $z = 4.184$ ,  $p < .001$ ; FC:  $z = 2.8934$ ,  $p = .004$ ).

In summary, testing the expression pattern of HSE genes against brain connectivity confirms a potentially interesting association between these two modalities, in support of an important multi-scale hypothesis of the expression of genes related to neuronal processes to potentially play an important role in shaping not only local (i.e., synapses), but also more global patterns of brain connectivity and brain network organization (Krienen et al., 2016; Vertes et al., 2016). This association (1) survived correction for spatial effects (i.e., null-spatial) and (2) tends to represent a rather unique and specific pattern compared to general patterns of expression of random genes (i.e., null-random-gene and null-coexpressed-gene), as well as (providing even more specificity) compared with other sets of brain genes (null-brain-gene). Findings on HSE genes and brain connectivity thus would be of more interest to the community and potential follow up examination in for example, disease conditions (Hoftman et al., 2020).

### 2.3 | Example 2: Alzheimer's disease risk gene APOE

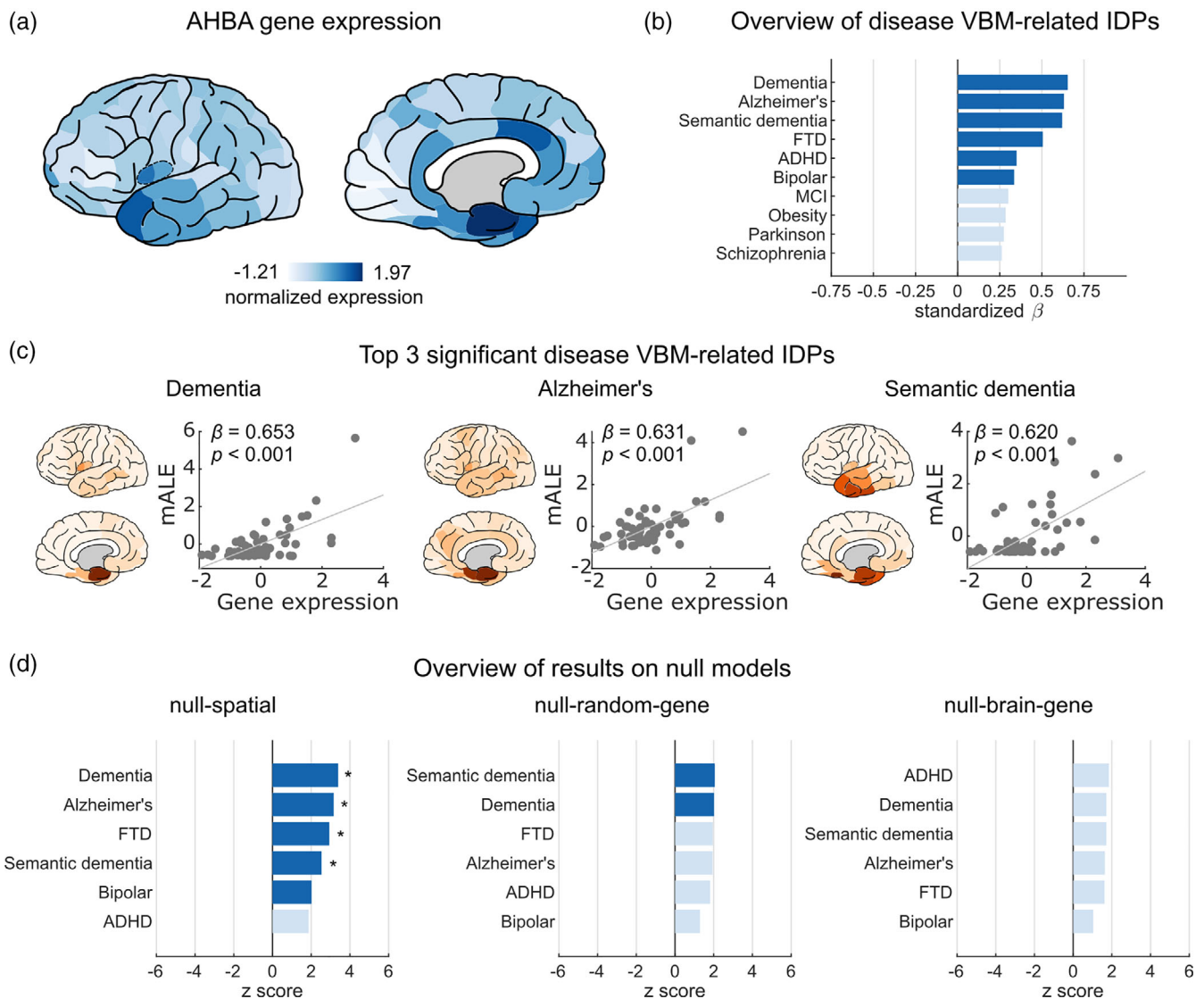
Our HSE example includes a case in which our hypothesized effect survived all different statistical evaluations, from single linear regression to null-spatial (i.e., correcting for spatial bias) to null-coexpressed-gene and null-brain-gene models (i.e., providing gene specificity). In a second example, we show that this is however not always the case, and that the use of a (too) liberal evaluation not testing for gene specificity potentially may lead to the report of a relative nonspecific finding. Here, we zoom in on transcriptomic brain patterns of disease-related pathology, another major topic of combined transcriptomic-neuroimaging studies (Freeze, Pandya, Zeighami, & Raj, 2019; McColgan et al., 2018; Romme et al., 2017). As an example, we examine the expression of the gene APOE, of which variant E4 is widely indicated as a risk variant for Alzheimer's disease (AD) (Liu, Liu, Kanekiyo, Xu, & Bu, 2013). We hypothesize that the cortical gene expression of APOE (as presented in the AHBA dataset) is related to

the cortical alterations revealed in patients with AD and other types of dementia.

We again start by testing for a significant correlation between the transcriptional profile of our GOI, here gene APOE (Figure 4a), and our neuroimaging phenotype of interest, here cortical gray matter atrophy patterns of 22 brain diseases, including AD, dementia, and others as reported by meta-analyses of the BrainMap voxel-based morphometry (VBM) studies (see Methods). Linear regression analysis reveals that the pattern of APOE gene expression is significantly associated with results of VBM studies reporting on atrophy of brain regions in (i) AD ( $\beta = 0.631$ ,  $p < .001$ ), (ii) dementia ( $\beta = 0.653$ ,  $p < .001$ ), (iii) semantic dementia ( $\beta = 0.620$ ,  $p < .001$ ), and (iv) frontotemporal dementia ( $\beta = 0.505$ ,  $p = .001$ ; FDR corrected  $q < 0.05$  for multiple testing across 22 diseases; Figure 4b,c). These findings seem to confirm our hypothesis, bringing promising evidence of a significant association between APOE gene expression and the pattern of cortical atrophy across multiple forms of dementia. Further analysis seems to reveal additional associations between APOE expression and the atrophy pattern of (v) attention deficit hyperactivity disorder (ADHD), which has been mentioned as a risk factor of dementia pathology (Callahan, Bierstone, Stuss, & Black, 2017) ( $\beta = 0.349$ ,  $p = .008$ ), and (vi) bipolar disorder ( $\beta = 0.335$ ,  $p = .010$ ; FDR corrected; Figure 4b).

Motivated by these positive results, we use the null-spatial model to make sure our effects survive when we take into account spatial relationships across brain regions and correct for spatial autocorrelation in the transcriptomic and neuroimaging data. Null-spatial reveals significant associations of APOE's expression with brain atrophy in four out of the six diseases listed above, including dementia ( $z = 3.391$ ,  $p < .001$ ), AD ( $z = 3.161$ ,  $p = .002$ ), frontotemporal dementia ( $z = 2.933$ ,  $p = .003$ ) and semantic dementia ( $z = 2.535$ ,  $p = .011$ ; FDR corrected across six tests; Figure 4d). These findings show encouraging effects that favor a relationship between the transcriptional profile of APOE and patterns of cortical atrophy related to dementia.

However, when we evaluate whether the observed associations between APOE transcriptional profile and disease-related atrophy patterns exceed effects that one could also observe by chance for any other set of genes, these effects diminish: The four significant associations observed in the linear regression analysis that further survived the null-spatial model, show only trend-level effects with respect to the null-random-gene model that constitutes the expression of randomly selected genes as null model, not surviving any correction for multiple testing (semantic dementia:  $z = 2.053$ ,  $p = .040$ ; dementia:  $z = 2.014$ ,  $p = .044$ ; AD:  $z = 1.943$ ,  $p = .052$ ; frontotemporal dementia:  $z = 1.955$ ,  $p = .051$ , ns after FDR correction; Figure 4d). Testing against the more appropriate null-brain-gene model (i.e., zooming in on genes over-expressed in brain tissue) shows that none of the effects are significant (semantic dementia:  $z = 1.715$ ,  $p = .086$ ; dementia:  $z = 1.716$ ,  $p = .086$ ; AD:  $z = 1.637$ ,  $p = .102$ ; frontotemporal dementia:  $z = 1.618$ ,  $p = .106$ , ns; Figure 4d). These statistical evaluations point toward the notion of the originally



**FIGURE 4** Example 2: *APOE* gene expression and atrophy in brain diseases. (a) Brain plots of normalized gene expression levels of *APOE*. (b) Overview of linear regression results, showing the top 10 associations between *APOE* gene expression profile and the imaging-derived phenotypes (IDPs) that correspond to disease atrophy patterns derived from the BrainMap voxel-based morphometry (VBM) studies. Dark blue indicates significant ( $q < 0.05$ , FDR corrected across 22 brain diseases included in GAMBA). (c) Top 3 significant correlations between *APOE* gene expression and VBM changes in dementia ( $\beta = 0.653$ ), Alzheimer's ( $\beta = 0.631$ ), semantic dementia ( $\beta = 0.620$ ). (d) Permutation testing results showing whether the observed effect size ( $\beta$  in panel c) is significantly beyond three distinct null distributions of effect sizes for null-spatial, null-random-gene, and null-brain-gene models. Dark blue indicates uncorrected  $p < .05$  (two-tailed z-test); "\*" indicates FDR corrected  $p < .05$

observed correlation between *APOE* and cortical atrophy to be an effect that holds for many other genes that are expressed in brain tissue, suggesting that the association is not "unique" or "specific"; we would have found the same effect if we would have selected other brain genes as our GOI at the start of our experiment.

While we thus initially found encouraging evidence of an association between our gene of interest *APOE* and the pattern of cortical atrophy in dementia/AD patients, this effect appears to be present among many other genes. In fact, the expression profile of 61 other genes (out of the total set of 2,711 brain-over-expressed genes) shows an equal or stronger correlation with AD cortical atrophy than gene *APOE*. In defense of a hypothesized association of *APOE*, it could

be argued that the 61 other genes are also all associated with the AD cortical atrophy pattern and that testing for gene-specificity is too strict. Posthoc analysis however does not show evidence in that direction: Examining the involvement of these 61 genes in one of the most recent GWAS studies on AD (Wightman et al., 2021) indicates that only one of these genes (only gene *COCH* from the reported 602 AD-associated genes; hypergeometric testing:  $p = .674$ ) has a discovered role in AD. Alternatively examining their potential involvement in differential gene expression in AD (Horesh, Katsel, Haroutunian, & Domany, 2011) reveals that only 2 of the 61 genes (only *GRIN2B*, *OPHN1* from the reported 276 genes; hypergeometric testing:  $p = .071$ ) show differentiated expression levels in AD patients.



## 2.4 | Example 3: Application to autism spectrum disorder risk genes

We further demonstrate the importance of proper null-model selection in avoiding over-interpretation of observed associations in a third example testing a gene-set of interest. We examine gene expressions of 25 risk genes of ASD (24 of which are present in AHBA) obtained from a recent meta-analysis of genome-wide association studies on a total of 18,381 ASD cases and 27,969 controls (Grove et al., 2019). Again, we start with computing the correlation between the expression pattern of ASD genes (Figure 5a) and functional alterations in brain diseases, here derived from the extensive BrainMap database (Fox et al., 2005; Fox & Lancaster, 2002; Laird, Lancaster, & Fox, 2005), which reveals an interesting association between the expression of ASD genes and regions involved in Asperger's syndrome ( $\beta = 0.284$ ,  $p = .032$ , not corrected) (Figure 5b)—a major diagnosis of ASD with difficulties in social interaction and nonverbal communication. In a subsequent analysis we correct for spatial relationships, with the effect positively surviving the null-spatial model ( $z = 2.006$ ,  $p = .045$ ). However, using null models to examine gene specificity does not show any significant result (null-random-gene:  $z = 1.223$ ,  $p = .222$ ; null-coexpressed-gene:  $z = 1.148$ ,  $p = .251$ ; null-brain-gene:  $z = 1.153$ ,

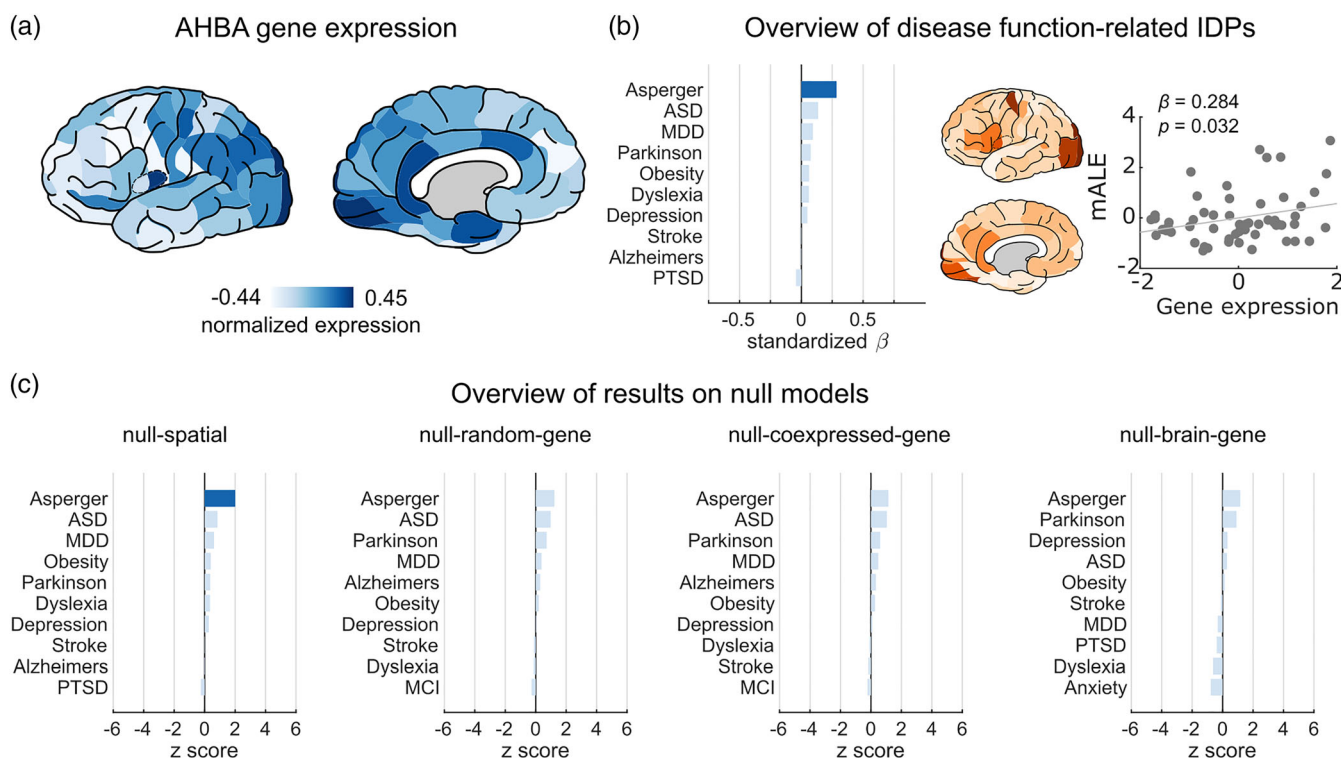
$p = .249$ ; null-coexpressed-brain-gene:  $z = 1.235$ ,  $p = .217$ ) (Figure 5c).

This third simple example shows that even when an expected association is observed that fits with current theories of brain (dis)organization and (dis)connectivity, these effects may be easily over-estimated, with other sets of genes (in this case, random subsets of genes generally expressed in brain tissue) showing highly similar expression patterns. The selection of a proper null model matching one's research question is of importance in providing specificity to an observed relationship between brain expression data and neuroimaging phenotypes (Table 1).

## 2.5 | Quantitative simulations of null models

### 2.5.1 | Brain phenotypes

We continue by simulating the outcome of the discussed statistical evaluation approaches for a wide range of real brain phenotypes and for a number of artificial gradients to get a deeper insight into the comparison of the different null-model strategies and their effect on reporting potential false-positive and/or nonspecific results. We examined the full outcome-space of all associations between single-



**FIGURE 5** Example 3: ASD gene expressions and functional alterations in brain diseases. (a) Brain plots of normalized gene expression levels of 24 ASD genes. (b) Overview of linear regression results, showing the top 10 associations between ASD gene expression profile and the patterns of brain functional alterations derived from the BrainMap fMRI studies, with Asperger's syndrome showing the highest correlation ( $\beta = 0.284$ ,  $p = .032$ , not corrected). (c) Permutation testing results showing whether the observed effect size ( $\beta$  in panel b) is significantly beyond four distinct null distributions of effect sizes for null-spatial, null-random-gene, null-coexpressed-gene, and null-brain-gene models. Dark blue indicates uncorrected  $p < .05$  (two-tailed z-test)

**TABLE 1** Summary of examples statistically testing associations between spatial patterns of gene expression and imaging-derived brain phenotypic patterns

Statistical approach	HSE genes and NOS	APOE and VBM changes in dementia	ASD genes and fMRI changes in Asperger's syndrome	Research question (X = expression of GOI, Y = brain-derived phenotype)	Weakness
Linear regression (standardized $\beta$ )	0.678**	0.653**	0.284*	Is the brain pattern of X and Y correlated?	Overestimation of effect due to spatial auto-correlation in X and Y
Null-spatial (z-score)	3.057**	3.391**	2.006*	Is the observed correlation between X and Y statistically valid?	Does not yet test for gene specificity to our GOI
Null-random-gene (z-score)	5.742**	2.014*	1.223	Is the association between X and Y specific to our a-priori chosen GOI?	Does not correct for spatial autocorrelation; over-estimated if high coexpression among the genes of interest
Null-coexpressed-gene (z-score)	6.002**	NA	1.148	Is the association between X and Y specific to our a-priori chosen GOI?	Does not correct for spatial autocorrelation; does not test for specificity among brain-related processes
Null-brain-gene (z-score)	4.311**	1.716	1.153	Is the association between X and Y relative specific to our GOI beyond what we can commonly expect from any other gene over-expressed in brain tissue?	Does not correct for spatial autocorrelation; coexpression not automatically conserved
Null-brain-coexpression (z-score)	4.304**	NA	1.235	Is the association between X and Y relative specific to our GOI beyond what we can commonly expect from any other set of genes over-expressed in brain tissue taking into account the coexpression structure in X?	Does not correct for spatial autocorrelation;

\*Uncorrected  $p < .05$  (two-sided);

\*\* $q < .05$ , FDR corrected; IDP, imaging-derived phenotype; GOI, gene(s)-of-interest.

gene expression profiles in AHBA (20,949 in total) and 384 imaging-derived brain phenotypes taken from multiple sources such as NeuroSynth (Yarkoni, Poldrack, Nichols, Van Essen, & Wager, 2011), BrainMap (Fox et al., 2005; Fox & Lancaster, 2002; Laird et al., 2005), and Cognitive Ontology (Yeo et al., 2016) (see Methods). Among these gene  $\times$  brain phenotype (20,949  $\times$  384) associations, the linear model indicates 960,963 (11.85%) of these associations as significant ( $p < .05$ , uncorrected), with 208,190 (2.57%) reaching FDR and 56,259 (0.69%) reaching Bonferroni correction with  $\alpha$  level  $< 0.05$  (corrected for 384 tests of all brain phenotypes per gene). We then implemented the null models to examine the spatial/gene specificity for the potential significant associations. Only FDR-corrected results are reported in the following paragraphs for simplicity. Uncorrected and Bonferroni-corrected results are tabulated in Table 2.

Applying the null-spatial model shows that only 26,273 out of the 208,190 associations reported in linear regression remain significant ( $p < .05$ ; Table 2; Figure S1). This suggests that a large proportion of

transcriptomic-neuroimaging associations as revealed by linear regression (87%) are overestimated due to dependencies of expression levels among neighboring brain regions and thus likely involve false-positive findings. This is in line with what was recently reported in studies evaluating null spatial models (see Burt et al., 2020; Fulcher et al., 2021; Markello & Misic, 2021).

When we then further implement the null-random-gene and null-brain-gene models (null-coexpressed-gene is not applicable in the situation of examining the spatial pattern of a single gene) to examine gene specificity of the reported transcriptomic-neuroimaging associations of our gene of interest and all neuroimaging patterns, we can find that only 15,297 out of these 26,273 reported associations that survived the null-spatial model remain further significant using the null-random-gene ( $p < .05$ , null-random-gene and null-spatial combined; Table 2). This suggests that even among spatially specific associations, there is still a considerable proportion of associations (42%) that can be commonly found for a wide range of (random) genes and

**TABLE 2** Summary of all potential associations between spatial patterns of single-gene expression and 384 imaging-derived brain phenotypic patterns

Methods	Number of significant associations ( $\alpha < 0.05$ )		
	Uncorrected	FDR-corrected	Bonferroni-corrected
LR	960,963	208,190	56,259
LR and N-spin	84,149	26,273	8,142
LR and N-spin and N-rand	26,786	15,297	7,610
LR and N-spin and N-brain	19,551	9,802	6,294

Abbreviations: LR, linear regression; N-spin, null-spatial model; N-rand, null-random-gene model; N-brain, null-brain-gene model.

therefore are hard to call an effect that is specific to our gene of interest; instead, they likely reflect a nonspecific finding.

Examining effects under the more stringent null-brain-gene model (providing specificity of findings regarding genes particularly expressed in brain tissue; most often our main topic of investigation) shows that an even smaller proportion of only 9,802 associations out of the 26,273 associations surviving the null-spatial model remain significant (just 37%,  $p < .05$ , null-brain-gene and null-spatial combined; Figure S1).

We note that similar types of results are found when we apply the null models the other way around, namely, first applying the null-brain-gene model and then the null-spatial model. Doing so, we can identify 67,181 significant transcriptomic-neuroimaging associations ( $p < .05$ , null-brain-gene), of which only 15% (9,802) are spatially specific ( $p < .05$ , null-spatial and null-brain-gene combined; Figure S1), indicating that gene specificity and correcting for spatial autocorrelation are not mutually inclusive.

Similar findings can also be observed for gene sets of more than one gene by examining transcriptomic-neuroimaging associations for gene sets curated in the Gene Ontology (GO) database (<http://geneontology.org>) (results for GO terms are presented in the Supplementary Results, Appendix S1).

## 2.5.2 | Simulated phenotypes

In the previous paragraphs, we demonstrate the need of examining gene specificity of observed transcriptomic-neuroimaging associations. We argue that this is particularly important when testing phenotypes with spatial patterns similar to global posterior–anterior or inferior–superior gradients (McColgan et al., 2018; Vogel et al., 2020). To show this point, we simulated spatial maps of seven phenotypes that follow global spatial gradients across the brain, from (i) posterior to anterior, (ii) from inferior to superior, (iii) from medial to lateral, together with the (iv)–(vii) four combinations of these gradients (Figure S2). We then examined all associations between single-gene expression profiles in AHBA and the simulated phenotypes. Among all these possible associations ( $20,949 \times 7$ ), we can find 23,723 (16%) significant associations between single-gene transcriptional profiles and simulated phenotypic profiles (linear regression:  $q < 0.05$ , FDR corrected; corrected for seven tests of all simulated phenotypes per gene; Table 3). These correlations suggest that a large set of genes

have expression profiles that follow general and very global spatial gradients of the brain, and show a nonspecific spatial pattern that is shared with a wide range of other genes (i.e., lack gene specificity). Application of the null-spatial model to these associations shows that only 3,305 out of the observed 23,723 associations (14%) remain significant (Table 3), suggesting that most of the common effects found significant using linear regression are inflated due to the strong spatial dependency of neighboring brain regions and can be filtered out by the null-spatial model. Among those 3,305 spatial-specific associations however, only 974 (29%) further remain significant when we additionally apply the null-random-gene model and examine whether effects are also gene-specific. An even smaller number of significant associations remain (only 626; 19%) when we use a stricter null-brain-gene model (Table 3; Figure S1).

In conclusion, our simulations on real-world brain phenotypes (previous section) and on simulated phenotypes (this section) together show that only as much as 3% (626 out of 23,723) of the associations between single-gene transcriptional profiles and 8 simple spatial gradients revealed in ordinary linear regression can be labeled as both surviving correction for spatial effects (null-spatial) and to be gene-specific (i.e., reflect a pattern that is quite unique to the GOI); 97% of the initially found FDR-corrected effects do not survive a statistical evaluation that corrects for spatial effects and represents a pattern that is unique or specific enough to support a main hypothesis of “the expression pattern of gene X to be associated to brain phenotype Y.”

## 2.6 | Toolbox

We made a simple web-based application and a MATLAB toolbox to facilitate quick examinations of transcriptomic-neuroimaging associations and to test observed correlations against different null models, tailored to the hypothesis that is examined and the level of specificity that the researcher desires. Within GAMBA (short for Gene Annotation using Macroscale Brain-imaging Association) website, expression profiles of input GOIs (i.e., a single gene or a set of genes) can be associated with imaging-derived brain traits from nine categories (Figure S3) and tested using the different null models as discussed in this article (including null-spatial, null-random-gene, null-coexpressed-gene, null-brain-gene). Imaging-derived phenotypes included in the tool cover the spatial patterns of (i) resting-state functional networks (Yeo et al., 2011), (ii) brain cognitive components (Yeo et al., 2016),

Methods	Number of significant associations ( $\alpha < 0.05$ )		
	Uncorrected	FDR-corrected	Bonferroni-corrected
LR	33,335	23,723	18,336
LR and N-spin	3,880	3,305	3,175
LR and N-spin and N-rand	1,003	974	967
LR and N-spin and N-brain	647	626	620

Abbreviations: LR, linear regression; N-spin, null-spatial model; N-rand, null-random-gene model; N-brain, null-brain-gene model.

(iii) regional metrics of the brain structural and functional connectome, (iv) measurements of the cortical oxygen and glucose metabolism (Vaishnavi et al., 2010), (v) human surface area expansion compared to the chimpanzee (Wei et al., 2019), (vi) brain volume alterations across twenty-two disorders (Fox et al., 2005; Fox & Lancaster, 2002; Laird et al., 2005), (vii) brain functional changes in sixteen disorders (Fox et al., 2005; Fox & Lancaster, 2002; Laird et al., 2005), (viii) cortical patterns of brain (dis)connectivity across nine psychiatric and neurological disorders (de Lange et al., 2019), and (ix) brain functional correlates of 292 terms in relation to cognitive states and brain disorders, as described in the NeuroSynth database (Yarkoni et al., 2011). Details on the included datasets and statistical analyses are described in the Supplementary Methods, Appendix S1. The web-tool is available online at <http://dutchconnectomelab.nl/GAMBA>. The MATLAB toolbox is available at <https://github.com/dutchconnectomelab/GAMBA-MATLAB>.

### 3 | DISCUSSION

We evaluated the importance of selecting the right null model when examining, testing, and discussing a hypothesized gene-brain association in a transcriptomic-neuroimaging study. We examined the usability of commonly applied statistical methods, such as simple single linear regression to assess the statistical validity of observed transcriptomic-neuroimaging relationships and once again confirm that the use of proper null models that constrain spatial effects (using the null-spatial model) is highly needed. We show that preserving spatial effects and controlling for spatial autocorrelation is not enough to provide information on whether an observed effect of our gene(s) of interest is unique and stands out from effects that can be widely observed across many other random genes in the dataset. Depending on the research question asked (e.g., “does the expression pattern of gene(s)-of-interest X show overlap with brain phenotype-of-interest Y”) and the level of specificity that we aim for in our study (e.g., “We hypothesize this relationship for gene(s) X and we want to rule out that this effect is not also commonly present for many other gene(s)-of-no-interest”), we recommend the use of multiple null models to provide information on the desired level of gene specificity of the observed associations (Figure 2).

The three examples we present do of course not cover all types of examinations one can perform by combining transcriptomic and

**TABLE 3** Summary of all potential associations between spatial patterns of single-gene expression and seven simulated brain phenotypes that represent global spatial gradients

neuroimaging data. They do however illustrate the potential and some of the caveats and limitations of cross-linking large numbers of cortical patterns of transcription and neuroimaging features. Our first example illustrates that transcriptomic-neuroimaging associations can include effects that survive correction for spatial autocorrelation and include a rather unique or specific pattern. We show that the cortical transcriptional profile of HSE genes runs parallel to spatial patterns of macroscale brain traits and that this association goes beyond both general spatial and random-gene-set effects that could be expected based on chance level alone. They thus highlight an interesting relationship between expression patterns of genes related to patterns of macroscale connectivity (Krienen et al., 2016; Romero-Garcia et al., 2018; Zeng et al., 2012).

A different conclusion should however be made when examining APOE and the set of ASD genes, with the two examples underscoring the importance of proper statistical testing. Our starting hypothesis was that “the cortical pattern of normative gene expression of APOE is related to the cortical atrophy pattern of AD,” potentially indicating regions that show higher vulnerability to the disease process. Findings in literature are in support of such associations, arguing in favor of a “true positive” effect. For example, AD patients are known to show elevated APOE expression in the medial temporal regions that are in turn known to be involved in the pathology of the disease (Akram, Schmeidler, Katsel, Hof, & Haroutunian, 2012; Linnertz et al., 2014; Ranlund et al., 2018). However, evaluation using the null models that test gene specificity reveals that the observed association is not particularly specific to APOE, and can be observed for nearly 60 other brain-expressed, AD-irrelevant genes in AHBA (corresponding to  $z = 1.716$  for AD, null-brain-gene model). This indicates that while potentially an interesting effect, it is quite hard to conclude that “the cortical pattern of normative gene expression of APOE is related to the cortical atrophy pattern of AD.” A conclusion that would better fit our results would be “the cortical pattern of normative gene expression of many brain-expressed genes is related to the cortical atrophy pattern of AD.” This is however a somewhat nonspecific conclusion and its implication to advancing our understanding of the disease remains unclear.

Simulations correlating single-gene expression profiles to brain phenotypes show that correcting for spatial autocorrelation (i.e., null-spatial) cannot be regarded as a substitute for testing gene specificity (e.g., null-brain), nor vice versa, pointing to the necessity of testing both the spatial and gene specificity in transcriptomic-neuroimaging

studies. Our simulations show that only 13.6% of all observed linear associations that survive FDR also survive the null-spatial model. Such a small proportion is likely attributable to a high auto-correlation between adjacent brain regions in transcriptomic and phenotypic data, leading to a degree of freedom much smaller than the one applied, such that the effects are inflated (Arnatkevičiūtė et al., 2019; Burt et al., 2020; Fulcher et al., 2021; Markello & Misić, 2021). This stresses the importance of using null models (e.g., the spin-based null-spatial model here (Alexander-Bloch et al., 2018) or other equivalents (Arnatkevičiūtė et al., 2019; Burt et al., 2020; Fulcher et al., 2021; Markello & Misić, 2021)) to reduce false-positive rate introduced by the spatial auto-correlation effects. Extending the important null-spatial model, our simulations further show that implementing the null-spatial model only is, however, not enough, and that 37.3% of the transcriptomic-neuroimaging associations that survived the null-spatial model do not show gene specificity. This proportion appears to be even lower (18.9%) in the simulations that take brain geographic gradients as the phenotypic of interest. This is because the transcriptional profiles of a considerable number of genes (16% of all genes) bear high similarity to the geographic gradients.

The discussed null models are presented in a simple web-tool GAMBA and a MATLAB toolbox, which can be used to probe the association between transcriptomics and common brain structure/function phenotypes derived from a wide range of neuroimaging data. GAMBA complements other tools that link genetics and brain functions, tools such as NeuroSynth (Yarkoni et al., 2011), Brain Annotation Toolbox (Liu et al., 2019), and for example, the ENIGMA toolbox (Larivière et al., 2021) that provide similar platforms to visualize and examine brain maps of gene expression patterns and brain maps, but do not directly provide means to test these effects against multiple null models. We note that the presented null models in the GAMBA MATLAB toolbox could be adapted and applied to other transcriptomic datasets of brain material (e.g., other than the now used default dataset of AHBA, such as PsychENCODE (Wang et al., 2018) and BrainSpan (Johnson et al., 2009)) and/or for the use in other types of datasets. For instance, as a potential application, similar null models could be used to examine gene specificity for associations between expression profiles of genes in single-cell RNAseq data of cell types (Tasic et al., 2018).

Several methodological points have to be considered. The null-brain-gene model takes genes ( $N = 2,711$ ) that are over-expressed in the brain as the background set of genes. We believe that this selection of background genes covers the majority of research contexts that are centered on brain-related processes and genes specifically “active” in the brain, but it does not cover contexts where the GOI plays a role in biological processes throughout the body or specific hypotheses where one wants to study genes overexpressed in non-brain tissues. We propose alternatives for the selection of the set of background genes that are more specifically tailored to alternative research questions in the Supplementary Methods, Appendix S1. Second, we need to consider that AHBA gene expression data is extracted from the postmortem brains of healthy individuals, that is, ones with no psychiatric and neurologic conditions. Therefore, results (when surviving null-spatial and/or null-brain models) should be

interpreted in the context of normative gene expression of risk genes in brain regions to be a potential marker for a higher susceptibility of these regions in relevant disorders (McColgan et al., 2018; Romme et al., 2017). Third, our examples (and the presented null-models) test between static gene expression data and brain phenotypes, data mostly derived from a single point in time. The examined associations and presented null-models thus overlook the important temporal aspects of expression patterns, for example the temporal trajectory of expression of genes during brain development (Kang et al., 2011) or aging (Fraser, Khativich, Plotkin, Pääbo, & Eisen, 2005). Further implementation of the discussed statistical approaches for the integration of neuroimaging data and, for example, datasets on human brain transcriptomics across development such as BrainSpan (Kang et al., 2011) is of great interest. Fourth, our examples and approaches, as the majority of studies examining similar topics, investigate the spatial correspondence between patterns of human brain gene expression and neuroimaging findings obtained at the group-level. Follow-up examinations using animal data where transcriptional data and imaging data are collected from the same specimens are an important next step to confirm and functionally annotate human transcriptomic-neuroimaging findings.

## 4 | CONCLUSION

In summary, we highlight the need of using null models to provide “gene specificity” when examining associations between the spatial patterns of gene transcriptomic profiles and imaging-derived brain traits.

## 5 | METHODS

### 5.1 | AHBA gene expression data

Microarray gene expression data were obtained from the extensive Allen Human Brain Atlas database (<http://human.brain-map.org>), including highly detailed data from six postmortem brains of donors without any neuropathological or neuropsychiatric conditions. Microarray analyses are described in detail in <http://help.brain-map.org/display/humanbrain/documentation>. Brain tissue samples of the left hemisphere were obtained from four donors ( $466 \pm 72.6$  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 of cortical and subcortical regions of the left hemisphere and used the expression of 58,692 probes for each brain donor (Romme et al., 2017; Wei et al., 2019).

We performed probe-to-gene re-annotation using the BioMart data-mining tool (<https://www.ensembl.org/biomart/>) (Arnatkevičiūtė et al., 2019). Outdated gene symbols were updated and alias gene symbols were replaced by symbols obtained from the HUGO Gene Nomenclature Committee (HGNC) database (<http://biomart>).



genenames.org/), resulting in the inclusion of 20,949 genes. Expression levels that were not well-above background were set to *NaN*. Per donor, and per tissue sample, expression levels of probes annotated to the same gene symbol were averaged, followed by log<sub>2</sub>-transformation with pseudocount 1. Tissue samples were spatially mapped to FreeSurfer cortical and subcortical regions to obtain region-wise gene expression profiles (French & Paus, 2015). A cortical parcellation of 114 regions (57 per hemisphere) and subcortical segmentation of 7 regions based on the Desikan-Killiany atlas (DK-114) were obtained for the Montreal Neurological Institute (MNI) 152 template using FreeSurfer (Cammoun et al., 2012; Desikan et al., 2006; Fonov et al., 2011). Brain tissue samples were annotated to cortical regions in the DK-114 atlas based on MNI coordinates, computing the nearest gray matter voxel within the MNI ICBM152 template in the FreeSurfer space. Tissue samples with a distance less than 2 mm 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  $6 \times 64 \times 20,949$  data matrix (i.e., donors  $\times$  brain regions  $\times$  genes). Within each donor, gene expression per gene was normalized to z-scores across all cortical and subcortical regions. Normalized gene expression profiles were averaged across the six donors obtaining a group-level gene expression matrix of size  $64 \times 20,949$ . Considering that most neuroimaging data as described in the following sections only include cortical regions, only the cortical expression pattern of each gene (i.e.,  $57 \times 20,949$  gene expression matrix) was correlated to the patterns of various neuroimaging findings.

## 5.2 | Neuroimaging phenotypic data

Connectome metrics used in *Example 1* were obtained from a human structural connectome map reconstructed using T1-weighted and diffusion-weighted MRI (dMRI) data of 487 subjects (age (mean  $\pm$  SD):  $29.8 \pm 3.4$  years old) from the Human Connectome Project (Van Essen et al., 2013). Disease maps used in *Example 2* and *Example 3* were computed based on coordinate-based results obtained from the extensive BrainMap database (<http://www.brainmap.org/>) that contains published functional and structural neuroimaging experiments of psychiatric and neurological disorders (see Supplementary Methods, Appendix S1 for a detailed description of the used procedures).

## 5.3 | Statistics and null models

*Linear regression.* Linear regression is used to test for an association between the cortical expression profile of a gene (or the average expression pattern across a set of genes) and the pattern of an imaging-derived brain phenotype:

$$Y_i = \beta_0 + \beta_1 X_j + \varepsilon \quad (1)$$

where  $Y_i$  indicates the standardized gene expression profile of gene  $i$  or the standardized, averaged profile of a gene set  $i$ , and  $X_j$  the

standardized cortical profile of neuroimaging phenotype  $j$ . Standardization is performed by dividing each value  $X$  or  $Y$  by the *SD*. The standardized regression coefficient  $\beta_1$  and the corresponding correlation coefficient and  $p$ -value are obtained.

### 5.3.1 | Null-spatial model

An important variant on the linear regression model was recently introduced (Alexander-Bloch et al., 2018), now testing whether the observed association is specific to spatial-anatomical relationships between brain regions. To this end, in (1) the observed  $\beta_1$  is compared to  $\beta_1$ s generated by 1,000 permutations, in which the gene expression data matrix is rebuilt using randomized brain parcellations by spinning the reconstructed sphere of the real brain parcellation with random angles ( $0$ – $360^\circ$ ) conserving the spatial relationship of neighboring regions.

### 5.3.2 | Null-random-gene model

The null-random-gene model tests whether the observed association is specific to the given gene or set of genes of interest, that is, comparing against effects that can be observed when any other set of genes would have been selected. To this end, it is tested whether the observed  $\beta_1$  (i.e., the effect size) is different from null distributions of  $\beta_1$  observed for randomly selected, same-sized, gene sets with the null-random-gene distribution estimated from 10,000 permutations and the mean ( $\mu$ ) and *SD* ( $\sigma$ ) of effect sizes in the null distribution obtained.

### 5.3.3 | Null-coexpressed-gene model

The null-coexpressed-gene model includes a stricter null model where random genes with similar coexpression levels as the given GOI are selected to generate null distributions of  $\beta_1$  (null-coexpressed-gene model). Random genes are selected according to the following steps: first, a set of random genes are initially selected. Then the mean coexpression level of the random gene-set is compared to the mean coexpression level of the given GOI. If the coexpression difference is larger than the maximum difference allowed, the gene with the highest/lowest coexpression level is excluded and a new random gene is included in the set. This step repeats until the coexpression difference is smaller than the maximum difference allowed. A total number of 1,000 sets of random genes were obtained due to the limitation of computational capacity.

### 5.3.4 | Null-brain-gene model

The null-brain-gene model is generated using random genes selected from a pool of 2,711 genes over-expressed in brain tissue, with brain-

expressed genes identified by performing one-tail two-sample *t*-tests on gene expression levels between brain tissues and other body sites ( $q < 0.05$ , FDR corrected), using gene expression data from the GTEx portal (<https://www.gtexportal.org>). Permutation is performed 10,000 times.

For all null models, the mean ( $\mu$ ) and SD ( $\sigma$ ) of effect sizes ( $\beta_1$ ) in the null distributions are estimated. A two-tailed *z*-test is performed to examine whether the observed  $\beta_1$  (i.e., the effect size) is larger than the mean effect size derived from null models:

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

where  $\mu$ ,  $\sigma$  indicate the mean and SD of  $\beta_1$  over random permutations. A two-tailed *p*-value is computed as follows:

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

where  $\Phi$  is the standard normal cumulative distribution function. We use *z*-tests to compute *p*-values in the GAMBA website as we aim to develop a light and quick website for the exploration of transcriptomic-imaging associations. Pseudo *p*-values may provide more accurate estimates, in particular when the null distribution is not normal, and are therefore implemented in the GAMBA MATLAB toolbox. The reported results in the current study are not influenced by the choice of these two *p*-value calculation approaches.

## ACKNOWLEDGMENTS

Martijn P. van den Heuvel was supported by an ALW open (ALWOP.179), a VID1 (452-16-015) grant from the Netherlands Organization for Scientific Research (NWO), and an ERC Consolidator grant (ID 101001062) of the European Research Council. D.P. was supported by The Netherlands Organization for Scientific Research (NWO VICI 453-14-005 and NWO Gravitation: BRAINSCAPES: A Roadmap from Neurogenetics to Neurobiology 024.004.012) and a European Research Council advanced grant (ERC-2018-AdG GWAS2FUNC 834057). Siemon C. de Lange was supported by the Amsterdam Neuroscience alliance grant and ZonMw Open Competition, project REMOVE 09120011910032. Human neuroimaging data were kindly provided by the Human Connectome Project, WU-Minn Consortium (Principal Investigators: David Van Essen and Kamil Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research; and by the McDonnell Center for Systems Neuroscience at Washington University.

## CONFLICT OF INTEREST

The authors declare no potential conflict of interest.

## AUTHOR CONTRIBUTIONS

Yongbin Wei: methodology, software, investigation, writing—original draft, writing—review & editing; Siemon C. de Lange: methodology, software, writing—review & editing; Rory Pijnenburg: validation,

writing—review & editing; Lianne H. Scholtens: visualization, writing—review & editing; Dirk Jan Ardesch: visualization, writing—review & editing; Kyoko Watanabe: software, writing—review & editing; Danielle Posthuma: supervision, writing—review & editing; Martijn P. van den Heuvel: conceptualization, methodology, writing—review & editing, supervision.

## DATA AVAILABILITY STATEMENT

Gene expression data are available in the Allen Brain Atlas (<http://human.brain-map.org>). The GTEx data are available in the GTEx Portal (<https://www.gtexportal.org>). MRI data (used for the connectome reconstruction) that support the findings of this study are available from the Human Connectome Project (Q3 release, <https://www.humanconnectome.org>). BrainMap data are available at <http://www.brainmap.org/>.

## ORCID

Yongbin Wei  <https://orcid.org/0000-0002-8044-7667>

Siemon C. de Lange  <https://orcid.org/0000-0002-3981-7996>

Rory Pijnenburg  <https://orcid.org/0000-0003-3817-7138>

Lianne H. Scholtens  <https://orcid.org/0000-0002-1331-3255>

Dirk Jan Ardesch  <https://orcid.org/0000-0003-0341-5554>

Kyoko Watanabe  <https://orcid.org/0000-0002-3303-8860>

Danielle Posthuma  <https://orcid.org/0000-0001-7582-2365>

Martijn P. van den Heuvel  <https://orcid.org/0000-0003-1570-1195>

## REFERENCES

- Akram, A., Schmeidler, J., Katsel, P., Hof, P. R., & Haroutunian, V. (2012). Association of ApoE and LRP mRNA levels with dementia and AD neuropathology. *Neurobiology of Aging*, 33, 628 e1–628 e14.
- Alexander-Bloch, A. F., Shou, H. C., Liu, S. Y., Satterthwaite, T. D., Glahn, D. C., Shinohara, R. T., ... Raznahan, A. (2018). On testing for spatial correspondence between maps of human brain structure and function. *NeuroImage*, 178, 540–551.
- Anderson, K. M., Collins, M. A., Kong, R., Fang, K., Li, J., He, T., ... Holmes, A. J. (2020). Convergent molecular, cellular, and cortical neuroimaging signatures of major depressive disorder. *Proceedings of the National Academy of Sciences of the United States of America*, 117, 25138–25149.
- Anderson, K. M., Krienen, F. M., Choi, E. Y., Reinen, J. M., Yeo, B. T. T., & Holmes, A. J. (2018). Gene expression links functional networks across cortex and striatum. *Nature Communications*, 9, 1428.
- Arnatkevičiūtė, A., Fulcher, B. D., & Fornito, A. (2019). A practical guide to linking brain-wide gene expression and neuroimaging data. *NeuroImage*, 189, 353–367.
- Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., ... Sherlock, G. (2000). Gene ontology: Tool for the unification of biology. The Gene Ontology Consortium. *Nature Genetics*, 25, 25–29.
- Burt, J. B., Demirtaş, M., Eckner, W. J., Navejar, N. M., Ji, J. L., Martin, W. J., ... Murray, J. D. (2018). Hierarchy of transcriptomic specialization across human cortex captured by structural neuroimaging topography. *Nature Neuroscience*, 21, 1251–1259.
- Burt, J. B., Helmer, M., Shinn, M., Anticevic, A., & Murray, J. D. (2020). Generative modeling of brain maps with spatial autocorrelation. *NeuroImage*, 220, 117038.
- Cacciaglia, R., Molinuevo, J. L., Falcón, C., Brugalat-Serrat, A., Sánchez-Benavides, G., Gramunt, N., ... Gispert, J. D. (2018). 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*, 14, 902–912.
- Callahan, B. L., Bierstone, D., Stuss, D. T., & Black, S. E. (2017). Adult ADHD: Risk factor for dementia or phenotypic mimic? *Frontiers in Aging Neuroscience*, 9, 260.
- Cammoun, L., Gigandet, X., Meskaldji, D., Thiran, J. P., Sporns, O., Do, K. Q., ... Hagmann, P. (2012). Mapping the human connectome at multiple scales with diffusion spectrum MRI. *Journal of Neuroscience Methods*, 203, 386–397.
- Consortium, G. TEx. (2015). Human genomics. The genotype-tissue expression (GTEx) pilot analysis: Multitissue gene regulation in humans. *Science*, 348, 648–660.
- Desikan, R. S., Ségonne, F., Fischl, B., Quinn, B. T., Dickerson, B. C., Blacker, D., ... Killiany, R. J. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *NeuroImage*, 31, 968–980.
- Fonov, V., Evans, A. C., Botteron, K., Almlí, C. R., McKinstry, R. C., Collins, D. L., & Brain Development Cooperative, Group. (2011). Unbiased average age-appropriate atlases for pediatric studies. *NeuroImage*, 54, 313–327.
- Fornito, A., Arnatkevičiute, A., & Fulcher, B. (2019). Bridging the gap between connectome and transcriptome. *Trends in Cognitive Sciences*, 23, 34–50.
- Fox, P. T., Laird, A. R., Fox, S. P., Fox, P. M., Uecker, A. M., Crank, M., ... Lancaster, J. L. (2005). BrainMap taxonomy of experimental design: Description and evaluation. *Human Brain Mapping*, 25, 185–198.
- Fox, P. T., & Lancaster, J. L. (2002). Opinion: Mapping context and content: The BrainMap model. *Nature Reviews Neuroscience*, 3, 319–321.
- Fraser, H. B., Khaitovich, P., Plotkin, J. B., Pääbo, S., & Eisen, M. B. (2005). Aging and gene expression in the primate brain. *PLoS Biology*, 3, e274.
- Freeze, B., Pandya, S., Zeighami, Y., & Raj, A. (2019). Regional transcriptional architecture of Parkinson's disease pathogenesis and network spread. *Brain*, 142, 3072–3085.
- French, L., & Paus, T. (2015). A FreeSurfer view of the cortical transcriptome generated from the Allen human brain atlas. *Frontiers in Neuroscience*, 9, 323.
- Fulcher, B. D., Arnatkevičiute, A., & Fornito, A. (2021). Overcoming false-positive gene-category enrichment in the analysis of spatially resolved transcriptomic brain atlas data. *Nature Communications*, 12, 2669.
- Grove, J., Ripke, S., Als, T. D., Mattheisen, M., Walters, R. K., Won, H., ... Børglum, A. D. (2019). Identification of common genetic risk variants for autism spectrum disorder. *Nature Genetics*, 51, 431–444.
- Hawrylycz, M. J., Lein, E. S., Guillozet-Bongaarts, A. L., Shen, E. H., Ng, L., Miller, J. A., ... Jones, A. R. (2012). An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature*, 489, 391–399.
- van den Heuvel, M. P., Scholtens, L. H., & Kahn, R. S. (2019). Multiscale neuroscience of psychiatric disorders. *Biological Psychiatry*, 86(7), 512–522.
- van den Heuvel, M. P., & Yeo, B. T. T. (2017). A spotlight on bridging microscale and macroscale human brain architecture. *Neuron*, 93, 1248–1251.
- Hoftman, G., Forsyth, J., Mennigen, E., Lin, A., Sun, D., Vajdi, A., & Bearden, C. (2020). Association of human supragranular enriched gene expression with cortical thickness patterns in 22q11.2 deletion syndrome. *Biological Psychiatry*, 87(S366), S366.
- Horesh, Y., Katsel, P., Haroutunian, V., & Domany, E. (2011). Gene expression signature is shared by patients with Alzheimer's disease and schizophrenia at the superior temporal gyrus. *European Journal of Neurology*, 18, 410–424.
- Jansen, I. E., Savage, J. E., Watanabe, K., Bryois, J., Williams, D. M., Steinberg, S., ... Posthuma, D. (2019). Genome-wide meta-analysis identifies new loci and functional pathways influencing Alzheimer's disease risk. *Nature Genetics*, 51, 404–413.
- Jansen, P. R., Nagel, M., Watanabe, K., Wei, Y., Savage, J. E., de Leeuw, C. A., ... Posthuma, D. (2020). Genome-wide meta-analysis of brain volume identifies genomic loci and genes shared with intelligence. *Nature Communications*, 11, 5606.
- Johnson, M. B., Kawasawa, Y. I., Mason, C. E., Krsnik, Ž., Coppola, G., Bogdanović, D., ... Šestan, N. (2009). Functional and evolutionary insights into human brain development through global transcriptome analysis. *Neuron*, 62, 494–509.
- Kang, H. J., Kawasawa, Y. I., Cheng, F., Zhu, Y., Xu, X., Li, M., ... Šestan, N. (2011). Spatio-temporal transcriptome of the human brain. *Nature*, 478, 483–489.
- Kepecs, A., & Fishell, G. (2014). Interneuron cell types are fit to function. *Nature*, 505, 318–326.
- Krienen, F. M., Yeo, B. T., Ge, T., Buckner, R. L., & Sherwood, C. C. (2016). Transcriptional profiles of supragranular-enriched genes associate with corticocortical network architecture in the human brain. *Proceedings of the National Academy of Sciences of the United States of America*, 113, E469–E478.
- Laird, A. R., Lancaster, J. L., & Fox, P. T. (2005). BrainMap: The social evolution of a human brain mapping database. *Neuroinformatics*, 3, 65–78.
- Larivière, S., Paquola, C., Park, B.-Y., Royer, J., Wang, Y., Benkarim, O., ... Bernhardt, B. C. (2021). The ENIGMA Toolbox: multiscale neural contextualization of multisite neuroimaging datasets. *Nature Methods*, 18(7), 698–700.
- de Lange, S. C., Scholtens, L. H., van den Berg, L. H., Boks, M. P., Bozzali, M., Cahn, W., ... van den Heuvel, M. P. (2019). Shared vulnerability for connectome alterations across psychiatric and neurological brain disorders. *Nature Human Behaviour*, 3, 988–998.
- de Leeuw, C. A., Stringer, S., Dekkers, I. A., Heskies, T., & Posthuma, D. (2018). Conditional and interaction gene-set analysis reveals novel functional pathways for blood pressure. *Nature Communications*, 9, 3768.
- Linnertz, C., Anderson, L., Gottschalk, W., Crenshaw, D., Lutz, M. W., Allen, J., ... Chiba-Falek, O. (2014). 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.
- Liu, C. C., Liu, C. C., Kanekiyo, T., Xu, H., & Bu, G. (2013). Apolipoprotein E and Alzheimer disease: Risk, mechanisms and therapy. *Nature Reviews Neurology*, 9, 106–118.
- Liu, Z., Rolls, E. T., Liu, Z., Zhang, K., Yang, M., Du, J., ... Wren, J. (2019). Brain annotation toolbox: Exploring the functional and genetic associations of neuroimaging results. *Bioinformatics*, 35, 3771–3778.
- Markello, R. D., Arnatkevičiūtė, A., Poline, J.-B., Fulcher, B. D., Fornito, A., & Misić, B. (2021). Standardizing workflows in imaging transcriptomics with the abagen toolbox. *bioRxiv*. <https://www.biorxiv.org/content/10.1101/2021.07.08.451635v1.abstract>
- Markello, R. D., & Misić, B. (2021). Comparing spatial null models for brain maps. *NeuroImage*, 236, 118052.
- McColgan, P., Gregory, S., Seunarine, K. K., Razi, A., Papoutsis, M., Johnson, E., ... Track-On, H. D. Investigators. (2018). Brain regions showing white matter loss in Huntington's disease are enriched for synaptic and metabolic genes. *Biological Psychiatry*, 83, 456–465.
- Ranlund, S., Rosa, M. J., de Jong, S., Cole, J. H., Kyriakopoulos, M., Fu, C. H. Y., ... Dima, D. (2018). Associations between polygenic risk scores for four psychiatric illnesses and brain structure using multivariate pattern recognition. *NeuroImage: Clinical*, 20, 1026–1036.
- Richiardi, J., Altmann, A., Milazzo, A. C., Chang, C., Chakravarty, M. M., Banaschewski, T., ... Greicius MD, consortium, Imagen. (2015). BRAIN NETWORKS. Correlated gene expression supports synchronous activity in BRAIN networks. *Science*, 348, 1241–1244.
- Romero-Garcia, R., Warrier, V., Bullmore, E. T., Baron-Cohen, S., & Bethlehem, R. A. I. (2019). Synaptic and transcriptionally down-regulated genes are associated with cortical thickness differences in autism. *Molecular Psychiatry*, 24, 1053–1064.
- Romero-Garcia, R., Whitaker, K. J., Vasa, F., Seidlitz, J., Shinn, M., Fonagy, P., ... Vertes, P. E. (2018). Structural covariance networks are

- coupled to expression of genes enriched in supragranular layers of the human cortex. *NeuroImage*, 171, 256–267.
- Romme, I. A., de Reus, M. A., Ophoff, R. A., Kahn, R. S., & van den Heuvel, M. P. (2017). Connectome disconnectivity and cortical gene expression in patients with schizophrenia. *Biological Psychiatry*, 81, 495–502.
- Schizophrenia Working Group of the Psychiatric Genomics Consortium. (2014). Biological insights from 108 schizophrenia-associated genetic loci. *Nature*, 511, 421–427.
- Tasic, B., Yao, Z., Graybiel, L. T., Smith, K. A., Nguyen, T. N., Bertagnoli, D., ... Zeng, H. (2018). Shared and distinct transcriptomic cell types across neocortical areas. *Nature*, 563, 72–78.
- Vaishnavi, S. N., Vlassenko, A. G., Rundle, M. M., Snyder, A. Z., Mintun, M. A., & Raichle, M. E. (2010). Regional aerobic glycolysis in the human brain. *Proceedings of the National Academy of Sciences of the United States of America*, 107, 17757–17762.
- Van Essen, D. C., Smith, S. M., Barch, D. M., Behrens, T. E., Yacoub, E., Ugurbil, K., & Consortium, W. U-Minn HCP. (2013). The WU-Minn human Connectome project: An overview. *NeuroImage*, 80, 62–79.
- Vertes, P. E., Rittman, T., Whitaker, K. J., Romero-Garcia, R., Vasa, F., Kitzbichler, M. G., ... Consortium, Nspn, Bullmore ET. (2016). Gene transcription profiles associated with inter-modular hubs and connection distance in human functional magnetic resonance imaging networks. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 371(1705), 20150362.
- Vogel, J. W., La Joie, R., Grothe, M. J., Diaz-Papkovich, A., Doyle, A., Vachon-Preseau, E., ... Evans, A. C. (2020). A molecular gradient along the longitudinal axis of the human hippocampus informs large-scale behavioral systems. *Nature Communications*, 11, 960.
- Wang, D., Liu, S., Warrell, J., Won, H., Shi, X., Navarro, F. C. P., ... Gerstein, M. B. (2018). Comprehensive functional genomic resource and integrative model for the human brain. *Science*, 362(6420).
- Wang, G. Z., Belgard, T. G., Mao, D., Chen, L., Berto, S., Preuss, T. M., ... Konopka, G. (2015). Correspondence between resting-State activity and brain gene expression. *Neuron*, 88, 659–666.
- Wang, Z., Dai, Z., Gong, G., Zhou, C., & He, Y. (2015). Understanding structural-functional relationships in the human brain: A large-scale network perspective. *The Neuroscientist*, 21, 290–305.
- Wei, Y., de Lange, S. C., Scholtens, L. H., Watanabe, K., Ardesch, D. J., Jansen, P. R., ... van den Heuvel, M. P. (2019). Genetic mapping and evolutionary analysis of human-expanded cognitive networks. *Nature Communications*, 10, 4839.
- Wightman, D. P., Jansen, I. E., Savage, J. E., Shadrin, A. A., Bahrami, S., Holland, D., ... Posthuma, D. (2021). A genome-wide association study with 1,126,563 individuals identifies new risk loci for Alzheimer's disease. *Nature Genetics*, 53, 1276–1282.
- Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C., & Wager, T. D. (2011). Large-scale automated synthesis of human functional neuroimaging data. *Nature Methods*, 8, 665–U95.
- Yeo, B. T., Krienen, F. M., Eickhoff, S. B., Yaakub, S. N., Fox, P. T., Buckner, R. L., ... Chee, M. W. (2016). Functional specialization and flexibility in human association cortex. *Cerebral Cortex*, 26, 465.
- Yeo, B. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., ... Buckner, R. L. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of Neurophysiology*, 106, 1125–1165.
- Zeng, H., Shen, E. H., Hohmann, J. G., Oh, S. W., Bernard, A., Royall, J. J., ... Jones, A. R. (2012). Large-scale cellular-resolution gene profiling in human neocortex reveals species-specific molecular signatures. *Cell*, 149, 483–496.

## SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

**How to cite this article:** Wei, Y., de Lange, S. C., Pijnenburg, R., Scholtens, L. H., Ardesch, D. J., Watanabe, K., Posthuma, D., & van den Heuvel, M. P. (2022). Statistical testing in transcriptomic-neuroimaging studies: A how-to and evaluation of methods assessing spatial and gene specificity. *Human Brain Mapping*, 43(3), 885–901. <https://doi.org/10.1002/hbm.25711>