HHS Public Access

Author manuscript

Curr Opin Neurol. Author manuscript; available in PMC 2022 August 01.

Published in final edited form as:

Curr Opin Neurol. 2021 August 01; 34(4): 480–487. doi:10.1097/WCO.0000000000000952.

Unveiling the Neuroimaging-Genetic Intersections in the Human Brain

Ibai Diez[1,2], **Jorge Sepulcre**[1,2,&]

[1]Gordon Center for Medical Imaging, Massachusetts General Hospital, Harvard Medical School, Boston MA.

[2]Athinoula A. Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Harvard Medical School, Charlestown, MA

Abstract

Purpose of review: The prevalence of new public datasets of brain-wide and single cell transcriptome data have created new opportunities to link neuroimaging findings with genetic data. The aim of this study is to present the different methodological approaches that have been used to combine this data.

Recent findings: Drawing from various sources of open access data, several studies have been able to correlate neuroimaging maps with spatial distribution of brain expression. These efforts have enabled researchers to identify functional annotations of related genes, identify specific cell types related to brain phenotypes, study the expression of genes across life span and highlight the importance of selected brain genes in disease genetic networks.

Summary: New transcriptome datasets and methodological approaches compliment current neuroimaging work and will be crucial to improve our understanding of the biological mechanism which underlie many neurological conditions.

Keywords

Neuroimaging; Genetics; Gene Expression; Transcriptome; Single-cell transcriptome

INTRODUCTION

In recent years, key developments on the integration of neuroimaging and gene expression data have led to a revolution in how we uncover and investigate the underpinnings of the human brain. This nascent neuroimaging discipline, the so-called neuroimaging-genetics, has enabled us to more precisely characterize the association between idiosyncratic imaging patterns and specific neurobiological hallmarks in multiple neurological diseases [1]. The field has seen marked improvement not only of imaging techniques defining large-scale brain networks breakdown, but also in the creation of stereotactic gene expression datasets.

[&]**Corresponding Author** Jorge Sepulcre, PhD, Massachusetts General Hospital, Gordon Center for Medical Imaging, 149 13th Street, Charlestown, MA 02129, USA, sepulcre@nmr.mgh.harvard.edu; Tel: 1-617-726-2899.

Conflict of interest: The authors declare no competing interest.

These publicly available data have provided unprecedented interest to close the gap between *in vivo* neuroimaging and its underlying molecular basis [2–7].

Investigations of how genes contribute to alterations in the structure and function of the human brain are not new. In the past, researchers focused on the selection of participant samples or cohorts with specific genetic variations to identify genetically associated neuroimaging patterns. These approaches require strong hypotheses on the genetic traits of interest. With the growing accessibility to perform Genome Wide Association Studies (GWAS) in large populations, this technique has become popular to link neuroimaging findings with genetic variants derived from this approach. GWAS is particularly effective to associate genetic SNP variations with neurodegenerative and neuropsychiatric disorders. While this strategy allows for a more data-driven identification of brain–genetic relationships, it requires large numbers of individuals. Moreover, the combination of neuroimaging and GWAS data is still insufficient to interpret how the molecular mechanistic of genetic variants impact the actual brain tissue at the *in situ* level (as the genetic information usually comes from blood or saliva samples).

The advent of brain-wide atlases, in addition to measurements of the expression of thousands of genes from different brain locations (using microarray and/or RNA-seq techniques from post-mortem brain tissue) has led to a recent scientific breakthrough in the neuroimaging-genetic field [1]. While previous methods focus on how genetics alter brain phenotypes indirectly, mainly by comparing genetically affected groups, the integration of Image Derived Phenotypes (IDPs) with gene expression atlases allow us to study their fine spatial intersections throughout the brain. In this review, we provide specific hints, orientations and resources about the recent methodological approaches used to study the transcriptomic delineations with neuroimaging patterns.

Considerations on Available Brain Transcriptomic Datasets, and Gene/Sample Filtering

To successfully design and perform the spatial integration of neuroimaging and genetic data, one must become familiar with the various publicly available brain atlases and with their spatial quantification of gene expression across multiple brain regions [2]. Table 1 shows the most significant human transcriptome datasets currently available [3–6], among which the Allen Human Brain Atlas (AHBA) [7] is the most popular, due to its high spatial resolution covering all cortical and subcortical regions (Fig. 1A). AHBA was built by collecting brain samples from healthy donors (Table 1). Other atlases covering several neurodegenerative disorders are included -although at lesser spatial resolution- among the resources of the Allen Institute for Brain Science. Additional brain datasets provide complementary information about the human brain transcriptome, such as temporal information throughout life span and increases in sample size obtained from the harmonization of combined datasets in 534 brain donors (0–102 yo) [8]. These datasets allow us to understand how the expression trajectories of particular genes are modeled in development throughout the lifespan[8–11]. More recently, Single-Cell RNA sequencing approaches have led to the creation of genome-wide expression profiles from brain cells [12–16], expanding the neuroimaging-genetic assessments toward specific cellular typologies rather than diluted brain data [17–20]. Using cell-type-specific gene markers and the HomoloGene database,

Shin et al. [21] have been able to differentiate the contributions of 9 human-like neural cell-types. The 9 major cell type classes include: S1 pyramidal neurons, CA1 pyramidal neurons, interneurons, astrocytes, microglia, oligodendrocytes, ependymal cells, endothelial cells, and mural cells (Fig. 1B). Now, AHBA transcriptomic data can be used to search for cell-type profiles associated with neuroimaging data. For researchers interested in evolutionary analysis, the PsychENCODE dataset stands out to provide gene expression data of 11 comparable cortical regions in the human, chimpanzee and macaque brains [22,23].

Not all the genes measured in transcriptome atlases are consistent across donors, and several studies have proposed to reduce the dimensionality of the data by removing highly variable genes[21,24]. For instance, a 2-stage filtering based on spearman correlations between individual data and the median expression of all AHBA donors, as well as the BrainSpan Atlas, has been proposed to filter all inconsistent gene expression patterns, that is genes with high variability in their regional expression profiles between donors [21]. This strategy obtained 2,511 highly reliable genes from the original 20,737 [21]. Other research groups have recommended to only include genes that display brain organ-specific expression relative to other tissues from the Human Protein Atlas (HPA; https://www.proteinatlas.org). A 4-fold gene expression level higher than other organs has been suggested as a selection criterion- obtaining a total of 2,587 brain specific genes[25]. However, alternative approaches based on the inclusion of brain tissue specificity or neuro-related genes from functional annotation systems (e.g. AmiGo) has yielded different criteria (such as 7,971 genes [20] or 3,719 genes [26]). Finally, other options have filtered the genetic data using potential candidate genes from GWAS studies [27].

Not all samples from transcriptome atlases are equally used in neuroimaging-genetic studies. Taking AHBA as a paradigmatic example, there are disparate gene expression levels between cortical and subcortical regions that have forced most researchers to focus on the cortical samples alone. Only a few exceptions have been reported which employ subcortical [28,29], or cortical and subcortical samples [30] with an *ad hoc* normalization strategy. Moreover, it is important to note that only two out of six donors have right hemisphere samples in the AHBA dataset (Table 1). This fact has inclined some researchers to use only the left hemisphere samples, only the samples of the two subjects that cover both hemispheres [31], or the whole dataset w/o applying sample normalization approaches [26]. Finally, due to the potential bias of autocorrelation between nearby samples -locations close to each other can exhibit similar expression values compared to the ones further apart- there have been recent efforts to create specific strategies that account for the spatial relationship between samples. For instance, Altmann et al. have used spatial eigenvectors and linear models to obtain residual data less impacted by spatial autocorrelations[32]. Although there is still needed to reach a consensus and establish a common approach to gene and sample filtering.

Spatial Associations between Image Derived Phenotypes and Genetics

Alzheimer's disease (AD) and Parkinson disease (PD), among other neurological disorders, display characteristic phenotypes on neuroimaging that have been consistently replicated over decades of research. These disorders produce distinctive cortical and subcortical spatial

signatures that transversally affect brain systems; an idea that has been captured in the past with an extended version of the old Hebbian principle ("not only do Neurons that 'fire together, wire together' but also Neurons that 'wire together, die together'"). The ability to produce these highly replicable IDPs have made neurodegenerative disorders an ideal framework to investigate the spatial similarity between neuroimaging and gene expression data. For instance, previous research has effectively identified specific neurobiological traits associated to the vulnerability of neuronal circuits in AD [26,33].

Magnetic Resonance Imaging (MRI) is the prominent method to map *in vivo*, functional and structural properties of the brain associated with a wide variety of neurodegenerative disorders, as well as normal organizational properties of the human brain. Different MRI sequences allow us to identify spatial fingerprints across both locally and distantly connected cerebral areas. Among the most popular IDPs for neuroimaging-genetic studies are: i) T1-weighted-image-related: cortical thickness, or voxel based volumetric information; ii) Functional or BOLD-imaging-related: connectivity properties or activation maps in task functional imaging (in this regard, the Neurosynth brain activation maps dataset -meta-analytic dataset of 14,371 fMRI studies- has inspired several works to identify associations between brain activation states and gene expression maps) [34–36] iii) Diffusion-image-related: diffusion microstructural properties or structural connectivity profiles of brain regions that are connected by white matter tracts. Apart from MRI, other types of IDPs have been successfully used in neuroimaging-genetic studies, such as lesion probability maps [37], and positron emission tomography images [28,38,39].

The integration of IDPs and gene expression data requires the use of spatial association approaches that intersect the two stereotactic maps (Fig. 1A,C,D). First, there are several strategies that have been proposed to match IDPs and gene expression data: i) one can use the phenotype-related imaging values from the exact locations where the brain tissue samples were extracted for genetic assessments; ii) use a brain atlas to define broad regions of anatomical or functional hallmarks, and employ it to compute the average gene expression values and average IDP-related values corresponding to the atlas regions; or iii) perform a voxel-based interpolation within the spatial domain of the genetic data to cover all IDPs space and match them [40]. A detailed pipeline of the steps to process transcriptome and map to an atlas can be found in [24]. Second, a comparison between the two metrics (IDPs and gene expression) overlapping within the same brain regions has to be made (Fig. 1D). While some studies focus on a priori hypotheses and specific regions of interest[41], others compute spatial similarity indexes from the entire cerebral distribution of IDPs and gene expression values using Pearson (Fig. 1C,D) or Spearman correlations [26,42,43], or Euclidean distances [44]. Then, top ranked genes that are highly similar to the imaging phenotype of interest are further studied.

Gene co-expression analysis allows us to detect a group of genes that are jointly expressed in same regions across the brain (Fig. 1A). The most popular gene co-expression approach is the Weighted Gene Co-expression Network Analysis (WGCNA). It uses a similarity value (usually a correlation value) between the expression of all pairs of genes to derive a gene x gene matrix. Using clustering approaches this co-expression matrix provide modules of highly interconnected genes, in which the first principal component of each group of genes

is called eigengene. Similarly, Partial Least Squares (PLS) Regression is a multivariate method that explains a set of observed variables (the IDP) as a linear combination of a set of predictor variables (expression of genes) where the maximum variance of the observed variables is explained. PLS is also well suited for highly collinear data as in the brain transcriptome. Rather than finding a similarity value between a specific gene with the IDPs, PLS finds several linear combinations of genes (components) that explains most of the variance of the IDPs. In neuroimaging-genetic studies, both the eigengenes and PLS scores can be correlated with IDPs to identify genes associated with the regional pattern of the phenotype of interest. In virtual histology, kernel density estimation is used to assess the significance of each cell type with the IDPs [21,45–47]. For each cell type, the distribution with the similarity values between the genes related with the cell type and the IDPs is computed and tested for significance. A linear regression model using the IDPs as dependent variables and the average expression profiles for the statistically significant cell types as independent variables are used to estimate the percent of variation explained by each cell type. Finally, in recent years, researchers have turned their attention toward the combination of transcriptome data and connectivity data of human brain [26,48]. The incorporation of spatial gene expression data to the spreading prediction of neurodegenerative processes in brain networks have opened fascinating opportunities to investigate the intimate network nature of AD or PD pathogenesis [49–52].

Post-Processing of Neuroimaging-Related Genes

One of the most important aspects of studies that investigate the spatial relationship between IDPs and gene expression is the interpretation of the resulting genetic findings. Researchers with a strong hypothesis for a particular gene of interest can search for high IDPs-genetic similarity across the transcriptomic data. They can build null hypothesis distributions in a variety of ways to determine if a targeted gene is among the resulting output of genes. This aprioristic approach has shown positive results in AD, PD and stuttering population studies [44,51,52]. Additionally, data-driven approaches using functional genetic annotations, such as gene-set enrichment analysis, can be performed by using all the ranked list of genes that display high similarity with the IDPs (e.g., red area in histogram of Fig. 1C) [26,53,54]. GeneOntology (GO) enrichment analysis facilitates the biological interpretation of a set of high scoring genes (graph bars in Fig. 1C), as it statistically tests annotation categories related to three main domains: cellular components, molecular function and biological process. Other domains from alternative resources are also useful in many cases. For example, the DisGeNET platform [57], a popular tool to compute disease related gene-set enrichment including associations between 17,549 genes and 24,166 disease traits, can be used to perform sex-chromosome gene-set enrichment[55] and disease related gene-set enrichment [19,56] analyses. Moreover, the most common tools for performing cell-type gene-set enrichment analysis are: i) expression-weighted cell type enrichment (EWCE)[58]; and ii) cell type specific expression analysis (CSEA; http://genetics[45]tl.edu/jdlab/csea-tool-2/)[59]. Importantly, it has been suggested that gene-set enrichment for brain transcriptome data might be affected by statistical bias leading to false positives introduced by gene-to-gene co-expression and spatial autocorrelation. Specific methods to ensure the statistical validity of the enrichment results from transcriptome data are currently being developed [60–62].

A solid neuroimaging-genetic study should incorporate specific approaches that independently confirm the genetic findings. We have found that genetic interactome analysis can provide an additional layer of validity to a discovered gene set. Given that these gene sets come from brain samples and brain IDPs associations, it is optimal if converging results can be obtained from large independent datasets of non-brain genetic interactions, such as co-expression, co-localization, shared pathways or protein domain (Genemania [63]). Today, using graph theory approaches we can test if brain-related gene sets obtained from the IDPs analysis interact in a coordinated and functional genetic manner within a more generalized framework (Fig. 1F). In our experience, an interactome confirmation is essential to understand if a given gene plays a central role in the disease, or condition-related genetic network [43,44,64,65].

Recent Advances on Connectomic-Genetics of Neurodegenerative Diseases

Current views of neurodegenerative disease, such as AD, PD or ALS, characterize their pathological progression as a spreading phenomenon across the cerebral tissue (e.g., accumulation of amyloid and tau in AD). Network organization of neuronal circuits and underlying genetic factors have enabled methodologies to more precisely describe the pathological propagations of neurodegenerative diseases. There is an immediate and intense interest to understand the genetic vulnerability of specific brain networks targeted by neurodegenerative diseases through application of neuroimaging, connectomic and genetic integration techniques. For instance, Sepulcre et al. described the longitudinal tau- and Aβ-related cortical pathways in AD, as well as their spatial associations with transcriptomic profiles [52]. They found that lipid metabolic genes, particularly APOE, played a central role in the overall propagation patterns [52]. Gene set enrichment analysis also classified the tau-imaging-genetic profile as being axon related, while the Aβ-imaging-genetic profile were dendrite related. In PD, the addition of regional expression data from strategic genes, such as GBA, SNCA, and LRRK2 as determinants of α-synuclein synthesis, clearance and/or degradation rate, improved the predictive power of disease spreading models [49,50]. Notably, Basaia et al. used neuroimaging connectivity data to characterize the α-synuclein propagation pathways of PD from the brainstem [51]. They also utilized the whole AHBA transcriptome to search for biological fingerprints of the progression along cerebral spreading pathways. They discovered that SNCA gene expression, as well as α-synuclein histological levels, display high spatial similarities with the cortical distribution of connectivity propagation pathways, in which the transcriptomic profiles associated to the PD cortical spreading showed key linkages to the regulation of dopamine secretion and microtubule neuronal architecture [51].

CONCLUSION

Recent studies linking transcriptome data to image-derived phenotypes have led to a better understanding of the putative biological mechanism underlying normal and neurodegenerative processes, compared to studies just using neuroimaging data. New technical developments are pushing forward the creation of new transcriptome databases that, combined with new methodological approaches, allow a deeper understanding of brain organization. More work is needed to expand the available transcriptome data: more donors,

single cell data covering whole brain, and transcriptome data of the brain with specific diseases. This new data will be translated into new opportunities that, complemented by computational methods, will expand our current knowledge of the human brain and eventually provide important clues to create disease modifying treatments in neurodegenerative disorders.

Acknowledgement:

This research was supported by grants from the National Institutes of Health (NIH) (R01AG061811, and R01AG061445 to J.S.).

REFERENCES

1. Fornito A, Arnatkevičiūtė A, Fulcher BD: Bridging the Gap between Connectome and Transcriptome. Trends in Cognitive Sciences 2018, 23:34–50. [PubMed: 30455082]
2. Keil JM, Qalieh A, Kwan KY: Brain Transcriptome Databases: A User's Guide. Journal of Neuroscience 2018, 38:2399–2412. [PubMed: 29437890]
3. Kang HJ, Kawasawa YI, Cheng F, Zhu Y, Xu X, Li M, Sousa AMM, Pletikos M, Meyer KA, Sedmak G, et al.: Spatio-temporal transcriptome of the human brain. Nature 2011, 478:483–489. [PubMed: 22031440]
4. Colantuoni C, Lipska BK, Ye T, Hyde TM, Tao R, Leek JT, Colantuoni EA, Elkahloun AG, Herman MM, Weinberger DR, et al.: Temporal dynamics and genetic control of transcription in the human prefrontal cortex. Nature 2011, 478:519–523. [PubMed: 22031444]
5. Miller JA, Ding S-L, Sunkin SM, Smith KA, Ng L, Szafer A, Ebbert A, Riley ZL, Royall JJ, Aiona K, et al.: Transcriptional landscape of the prenatal human brain. Nature 2014, 508:199–206. [PubMed: 24695229]
6. Aguet F, Brown A, Castel SE, Davis JR, He Y, Jo B, Mohammadi P, Park Y, Parsana P, Segrè AV, et al.: Genetic effects on gene expression across human tissues. Nature 2017, 550:204–213. [PubMed: 29022597]
7. Hawrylycz MJ, Lein ES, Guillozet-Bongaarts AL, Shen EH, Ng L, Miller JA, Lagemaat van de LN, Smith KA, Ebbert A, Riley ZL, et al.: An anatomically comprehensive atlas of the adult human brain transcriptome. Nature 2012, 489:391–399. [PubMed: 22996553]
8. Disorder WC for the A-D, Disorder AS, Disorder B, Disorder MD, Disorder O-C, Groups and SEW: Virtual Histology of Cortical Thickness and Shared Neurobiology in 6 Psychiatric Disorders. Jama Psychiat 2020, doi:10.1001/jamapsychiatry.2020.2694.
9. Xu Q, Liu F, Qin W, Jiang T, Yu C: Multiscale neurobiological correlates of human neuroticism. Hum Brain Mapp 2020, 41:4730–4743. [PubMed: 32839993]
10. Fu J, Liu F, Qin W, Xu Q, Yu C, (ADNI) ADNI: Individual-Level Identification of Gene Expression Associated with Volume Differences among Neocortical Areas. Cereb Cortex 2020, 30:3655–3666. [PubMed: 32186704] *This study uses transcriptome data of spatio-temporal data to perform enrichment in different developmental window and also uses genetic data of different tissues in the body to determine how likely a gene was to express in the brain relative to other tissues.
11. Paquola C, Bethlehem RA, Seidlitz J, Wagstyl K, Romero-Garcia R, Whitaker KJ, Wael RV de, Williams GB, Vértes PE, Margulies DS, et al.: Shifts in myeloarchitecture characterise adolescent development of cortical gradients. Elife 2019, 8:e50482. [PubMed: 31724948] *This study performs developmental enrichment and cell-type specific expression analyses with age related changes in intracortical myelin maps
12. Lake BB, Chen S, Sos BC, Fan J, Kaeser GE, Yung YC, Duong TE, Gao D, Chun J, Kharchenko PV, et al.: Integrative single-cell analysis of transcriptional and epigenetic states in the human adult brain. Nat Biotechnol 2018, 36:70–80. [PubMed: 29227469]

13. Habib N, Avraham-Davidi I, Basu A, Burks T, Shekhar K, Hofree M, Choudhury SR, Aguet F, Gelfand E, Ardlie K, et al.: Massively parallel single-nucleus RNA-seq with DroNc-seq. Nat Methods 2017, 14:955–958. [PubMed: 28846088]

14. Darmanis S, Sloan SA, Zhang Y, Enge M, Caneda C, Shuer LM, Gephart MGH, Barres BA, Quake SR: A survey of human brain transcriptome diversity at the single cell level. Proc National Acad Sci 2015, 112:7285–7290.

15. Li M, Santpere G, Kawasawa YI, Evgrafov OV, Gulden FO, Pochareddy S, Sunkin SM, Li Z, Shin Y, Zhu Y, et al.: Integrative functional genomic analysis of human brain development and neuropsychiatric risks. Science 2018, 362:eaat7615. [PubMed: 30545854]

16. Zhang Y, Sloan SA, Clarke LE, Caneda C, Plaza CA, Blumenthal PD, Vogel H, Steinberg GK, Edwards MSB, Li G, et al.: Purification and Characterization of Progenitor and Mature Human Astrocytes Reveals Transcriptional and Functional Differences with Mouse. Neuron 2016, 89:37–53. [PubMed: 26687838]

17. Zeisel A, Muñoz-Manchado AB, Codeluppi S, Lönnerberg P, Manno GL, Juréus A, Marques S, Munguba H, He L, Betsholtz C, et al.: Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq. Science 2015, 347:1138–1142. [PubMed: 25700174]

18. Anderson KM, Collins MA, Kong R, Fang K, Li J, He T, Chekroud AM, Yeo BTT, Holmes AJ: Convergent molecular, cellular, and cortical neuroimaging signatures of major depressive disorder. Proc National Acad Sci 2020, 117:25138–25149.

19. Ball G, Seidlitz J, Beare R, Seal ML: Cortical remodelling in childhood is associated with genes enriched for neurodevelopmental disorders. Neuroimage 2020, 215:116803. [PubMed: 32276068] *This study combines postnatal single cell data from different datasets, inhibitory and excitatory gene lists, disease associated gene lists and 10 groups of genes with differential expression across the life span to study developmental cortical changes. Fold enrichment of prenatal and postnatal modules is also performed.

20. Seidlitz J, Nadig A, Liu S, Bethlehem RAI, Vértes PE, Morgan SE, Váša F, Romero-Garcia R, Lalonde FM, Clasen LS, et al.: Transcriptomic and cellular decoding of regional brain vulnerability to neurogenetic disorders. Nat Commun 2020, 11:3358. [PubMed: 32620757]

21. Shin J, French L, Xu T, Leonard G, Perron M, Pike GB, Richer L, Veillette S, Pausova Z, Paus T: Cell-Specific Gene-Expression Profiles and Cortical Thickness in the Human Brain. Cereb Cortex 2017, 28:3267–3277.

22. Sousa AMM, Zhu Y, Raghanti MA, Kitchen RR, Onorati M, Tebbenkamp ATN, Stutz B, Meyer KA, Li M, Kawasawa YI, et al.: Molecular and cellular reorganization of neural circuits in the human lineage. Science 2017, 358:1027–1032. [PubMed: 29170230]

23. Wei Y, Lange SC de, Scholtens LH, Watanabe K, Ardesch DJ, Jansen PR, Savage JE, Li L, Preuss TM, Rilling JK, et al.: Genetic mapping and evolutionary analysis of human-expanded cognitive networks. Nat Commun 2019, 10:4839. [PubMed: 31649260] *This study uses the PsychENCODE dataset to compare genetic expression of human-accelerated genes in 11 comparable cortical regions between species: humans, chimpanzees and macaques

24. Arnatkevičiūtė A, Fulcher BD, Fornito A: A practical guide to linking brain-wide gene expression and neuroimaging data. Neuroimage 2019, 189:353–367. [PubMed: 30648605]

25. Shafiei G, Markello RD, Wael RV de, Bernhardt BC, Fulcher BD, Misic B: Topographic gradients of intrinsic dynamics across neocortex. Elife 2020, 9:e62116. [PubMed: 33331819]

26. Diez I, Sepulcre J: Neurogenetic profiles delineate large-scale connectivity dynamics of the human brain. Nat Commun 2018, 9:3876. [PubMed: 30250030]

27. Hatoum AS, Reineberg AE, Smolker HR, Hewitt JK, Friedman NP: Whole-cortex mapping of common genetic influences on depression and a social deficits dimension. Transl Psychiat 2019, 9:299.

28. Komorowski A, Weidenauer A, Murgaš M, Sauerzopf U, Wadsak W, Mitterhauser M, Bauer M, Hacker M, Praschak-Rieder N, Kasper S, et al.: Association of dopamine D2/3 receptor binding potential measured using PET and [11C]-(+)-PHNO with post-mortem DRD2/3 gene expression in the human brain. Neuroimage 2020, 223:117270. [PubMed: 32818617]

29. Vogel JW, Joie RL, Grothe MJ, Diaz-Papkovich A, Doyle A, Vachon-Presseau E, Lepage C, Wael RV de, Thomas RA, Iturria-Medina Y, et al.: A molecular gradient along the longitudinal axis of

the human hippocampus informs large-scale behavioral systems. Nat Commun 2020, 11:960. [PubMed: 32075960]

30. Tan PK, Ananyev E, Hsieh P-J (Brown): Distinct genetic signatures of cortical and subcortical regions associated with human memory. Eneuro 2019, 6:ENEURO.0283–19.2019.
31. Pappas I, Craig MM, Menon DK, Stamatakis EA: Structural optimality and neurogenetic expression mediate functional dynamics in the human brain. Hum Brain Mapp 2020, 41:2229–2243. [PubMed: 32027077]
32. Altmann A, Cash DM, Bocchetta M, Heller C, Reynolds R, Moore K, Convery RS, Thomas DL, Swieten JC van, Moreno F, et al.: Analysis of brain atrophy and local gene expression in genetic frontotemporal dementia. Brain Commun 2020, 2:fcaa122. [PubMed: 33210084]
33. Grothe MJ, Sepulcre J, Gonzalez-Escamilla G, Jelistratova I, Schöll M, Hansson O, Teipel SJ, Initiative ADN: Molecular properties underlying regional vulnerability to Alzheimer's disease pathology. Brain 2018, 141:2755–2771. [PubMed: 30016411]
34. Yarkoni T, Poldrack RA, Nichols TE, Essen DCV, Wager TD: Large-scale automated synthesis of human functional neuroimaging data. Nat Methods 2011, 8:665–670. [PubMed: 21706013]
35. Quintana DS, Rokicki J, Meer D van der, Alnæs D, Kaufmann T, Córdova-Palomera A, Dieset I, Andreassen OA, Westlye LT: Oxytocin pathway gene networks in the human brain. Nat Commun 2019, 10:668. [PubMed: 30737392]
36. Xie Y, Zhang X, Liu F, Qin W, Fu J, Xue K, Yu C: Brain mRNA Expression Associated with Cortical Volume Alterations in Autism Spectrum Disorder. Cell Reports 2020, 32:108137. [PubMed: 32937121] *This study uses the Neurosynth database to link the obtained gene list with implicated neural processes and psychiatric disorders obtained from task-based metanalyses.
37. Mandal AS, Romero-Garcia R, Hart MG, Suckling J: Genetic, cellular, and connectomic characterization of the brain regions commonly plagued by glioma. Brain 2020, 143:3294–3307. [PubMed: 33278823]
38. Kim M-J, Lee J-H, Anaya FJ, Hong J, Miller W, Telu S, Singh P, Cortes MY, Henry K, Tye GL, et al.: First-in-human evaluation of [11C]PS13, a novel PET radioligand, to quantify cyclooxygenase-1 in the brain. Eur J Nucl Med Mol I 2020, 47:3143–3151.
39. Fjell AM, Sederevicius D, Sneve MH, Lange A-MG de, Bråthen AC, Idland A-V, Watne LO, Wang Y, Reinbold C, Dobricic V, et al.: Self-reported Sleep Problems Related to Amyloid Deposition in Cortical Regions with High HOMER1 Gene Expression. Cereb Cortex 2019, 30:2144–2156.
40. Gryglewski G, Seiger R, James GM, Godbersen GM, Komorowski A, Unterholzner J, Michenthaler P, Hahn A, Wadsak W, Mitterhauser M, et al.: Spatial analysis and high resolution mapping of the human whole-brain transcriptome for integrative analysis in neuroimaging. Neuroimage 2018, 176:259–267. [PubMed: 29723639]
41. Manza P, Yuan K, Shokri-Kojori E, Tomasi D, Volkow ND: Brain structural changes in cannabis dependence: association with MAGL. Mol Psychiatr 2020, 25:3256–3266.
42. Ortiz-Terán L, Diez I, Ortiz T, Perez DL, Aragón JI, Costumero V, Pascual-Leone A, Fakhri GE, Sepulcre J: Brain circuit–gene expression relationships and neuroplasticity of multisensory cortices in blind children. Proc National Acad Sci 2017, 114:6830–6835.
43. Bueichekú E, Aznárez-Sanado M, Diez I, Uquillas F d'Oleire, Ortiz-Terán L, Qureshi AY, Suñol M, Basaia S, Ortiz-Terán E, Pastor MA, et al.: Central neurogenetic signatures of the visuomotor integration system. Proc National Acad Sci 2020, 117:6836–6843.*This paper studies the link between visuomotor integration system with genetic data. Interactome analysis is applied to quantify the centrality of candidate genes in the obtained gene to gene network.
44. Benito-Aragón C, Gonzalez-Sarmiento R, Liddell T, Diez I, Uquillas F d'Oleire, Ortiz-Terán L, Bueichekú E, Chow HM, Chang S-E, Sepulcre J: Neurofilament-lysosomal genetic intersections in the cortical network of stuttering. Prog Neurobiol 2020, 184:101718. [PubMed: 31669185]
45. Vidal-Pineiro D, Parker N, Shin J, French L, Grydeland H, Jackowski AP, Mowinckel AM, Patel Y, Pausova Z, Salum G, et al.: Cellular correlates of cortical thinning throughout the lifespan. Sci Rep-uk 2020, 10:21803.
46. Fjell AM, Sørensen Ø, Amlien IK, Bartrés-Faz D, Brandmaier AM, Buchmann N, Demuth I, Drevon CA, Düzel S, Ebmeier KP, et al.: Poor Self-Reported Sleep is Related to Regional Cortical

Thinning in Aging but not Memory Decline—Results From the Lifebrain Consortium. Cereb Cortex 2020, doi:10.1093/cercor/bhaa332.

47. Patel Y, Shin J, Drakesmith M, Evans J, Pausova Z, Paus T: Virtual histology of multi-modal magnetic resonance imaging of cerebral cortex in young men. Neuroimage 2020, 218:116968. [PubMed: 32450248]
48. Liu J, Xia M, Wang X, Liao X, He Y: The spatial organization of the chronnectome associates with cortical hierarchy and transcriptional profiles in the human brain. Neuroimage 2020, 222:117296. [PubMed: 32828922]
49. Zheng Y-Q, Zhang Y, Yau Y, Zeighami Y, Larcher K, Misic B, Dagher A: Local vulnerability and global connectivity jointly shape neurodegenerative disease propagation. Plos Biol 2019, 17:e3000495. [PubMed: 31751329]
50. Yan W, Ye C, Wang T, Sun J, Wu T, Ma T: Misfolded protein propagation in an integrated computational model of structural network and LRRK2 gene expression. 2020 42nd Annu Int Conf Ieee Eng Medicine Biology Soc Embc 2020, 00:2368–2371.
51. Basaia S, Agosta F, Diez I, Bueichekú E, Uquillas F d'Oleire, Delgado-Alvarado M, Caballero-Gaudes C, Rodriguez-Oroz M, Stojkovic T, Kostic VS, et al.: Neurogenetic Traits Conferring Vulnerability to Cortical Progression of Parkinson's Disease. Ssrn Electron J 2020, doi:10.2139/ssrn.3696874.
52. Sepulcre J, Grothe MJ, Uquillas F d'Oleire, Ortiz-Terán L, Diez I, Yang H-S, Jacobs HIL, Hanseeuw BJ, Li Q, El-Fakhri G, et al.: Neurogenetic contributions to amyloid beta and tau spreading in the human cortex. Nat Med 2018, 24:1910–1918. [PubMed: 30374196]
53. Diez I, Larson AG, Nakhate V, Dunn EC, Fricchione GL, Nicholson TR, Sepulcre J, Perez DL: Early-life trauma endophenotypes and brain circuit–gene expression relationships in functional neurological (conversion) disorder. Mol Psychiatr 2020, doi:10.1038/s41380-020-0665-0.
54. Mattsson N, Palmqvist S, Stomrud E, Vogel J, Hansson O: Staging β-Amyloid Pathology With Amyloid Positron Emission Tomography. Jama Neurol 2019, 76:1319–1329. [PubMed: 31314895]
55. Liu S, Seidlitz J, Blumenthal JD, Clasen LS, Raznahan A: Integrative structural, functional, and transcriptomic analyses of sex-biased brain organization in humans. Proc National Acad Sci 2020, 117:18788–18798.
56. Liu W, Peeters N, Fernández G, Kohn N: Common neural and transcriptional correlates of inhibitory control underlie emotion regulation and memory control. Soc Cogn Affect Neur 2020, 15:nsaa073-.
57. Piñero J, Ramírez-Anguita JM, Saüch-Pitarch J, Ronzano F, Centeno E, Sanz F, Furlong LI: The DisGeNET knowledge platform for disease genomics: 2019 update. Nucleic Acids Res 2019, 48:D845–D855.
58. Skene NG, Grant SGN: Identification of Vulnerable Cell Types in Major Brain Disorders Using Single Cell Transcriptomes and Expression Weighted Cell Type Enrichment. Front Neurosci-switz 2016, 10:16.
59. Xu X, Wells AB, O'Brien DR, Nehorai A, Dougherty JD: Cell Type-Specific Expression Analysis to Identify Putative Cellular Mechanisms for Neurogenetic Disorders. J Neurosci 2014, 34:1420–1431. [PubMed: 24453331]
60. Fulcher BD, Arnatkevičiūtė A, Fornito A: Overcoming bias in gene-set enrichment analyses of brain-wide transcriptomic data. Biorxiv 2020, doi:10.1101/2020.04.24.058958.*This study introduces a method and provides a toolbox to correct statistical bias, introduced by within category gene-gene coexpression and spatial autocorrelation, when computing gene-set enrichment analysis of transcriptome data.
61. Burt JB, Helmer M, Shinn M, Anticevic A, Murray JD: Generative modeling of brain maps with spatial autocorrelation. Neuroimage 2020, 220:117038. [PubMed: 32585343] *This paper presents a generative null model to generate surrogate maps with spatial autocorrelation that can be used in gene set enrichment analysis to test hypothesis of interest.
62. Zarkali A, McColgan P, Ryten M, Reynolds RH, Leyland L-A, Lees AJ, Rees G, Weil RS: Dementia risk in Parkinson's disease is associated with interhemispheric connectivity loss and determined by regional gene expression. Neuroimage Clin 2020, 28:102470. [PubMed: 33395965]

63. Mostafavi S, Ray D, Warde-Farley D, Grouios C, Morris Q: GeneMANIA: a real-time multiple association network integration algorithm for predicting gene function. Genome Biol 2008, 9:S4.

64. Ji Y, Zhang X, Wang Z, Qin W, Liu H, Xue K, Tang J, Xu Q, Zhu D, Liu F, et al.: Genes associated with gray matter volume alterations in schizophrenia. Neuroimage 2021, 225:117526. [PubMed: 33147509]

65. Romero-Garcia R, Seidlitz J, Whitaker KJ, Morgan SE, Fonagy P, Dolan RJ, Jones PB, Goodyer IM, Suckling J, Consortium N, et al.: Schizotypy-Related Magnetization of Cortex in Healthy Adolescence Is Colocated With Expression of Schizophrenia-Related Genes. Biol Psychiat 2020, 88:248–259. [PubMed: 32029217]

KEY POINTS

- Brain-wide transcriptome atlases are a powerful tool to link neuroimage phenotypes with genetic information
- Gene set enrichment analysis allow a better interpretation of biological mechanism behind
- Changes of expression in different age ranges can be computed using datasets including samples from individual across lifespan
- Single cell transcriptome dataset provides new insight into specific cells types linked to our phenotype
- Gene to gene relationship and central genes can be computed with Interactome tools

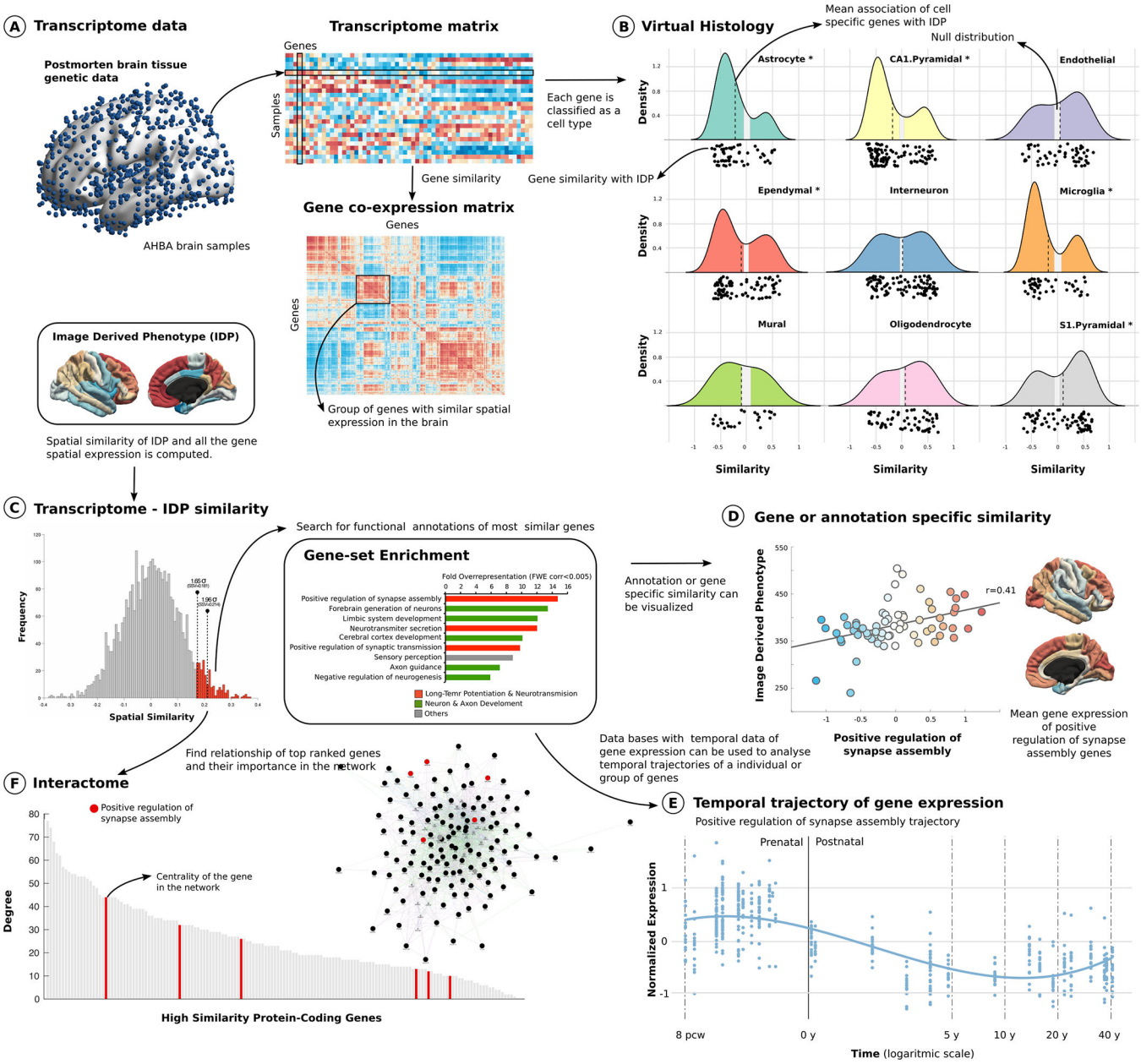

Figure 1. Main methods applied to study the spatial association of Image Derived Phenotypes and genetic information.

Transcriptome data is obtained from postmortem brain tissue from different brain regions across the cortex (A, AHBA example). The transcriptome matrix (brain samples x genes) contains the expression of all the measured genes for each sample of the brain. Computing the similarity of the spatial expression of each pair of genes in the brain the gene-coexpression matrix is obtained, which identifies groups of genes with similar spatial patterns. Single cell transcriptome data allow to classify each gene into a cell type (B). This information allows us to compute the distribution of the spatial similarity of the IDP with the genes pertaining to each cell type. The spatial similarity of the whole transcriptome can be computed with the IDPs to obtain a distribution of similarity values (C has been adapted

from [26], and C-F shows the significant fold enrichment on *Positive Regulation of Synapse Assembly* as a case example). The top ranked genes (most similar to the IDPs) can be used in a gene-set enrichment analysis to find significantly associated functional annotations with the genes (C). The functional annotations or genes of interest can be further explored to obtain the correlation value with IDP, project their values onto the brain surface and compute the temporal trajectory along the lifespan (D-E). This top ranked list can also be used to find relationships with other genes and find the importance of our candidate genes in the obtained interactomic gene network (F). The position of each node in the network is plotted as a function of the hubness of each gene (or degree centrality, the number of connections to other genes of the network) in the bars figure in F.

Table 1.

Key publicly available human brain transcriptome datasets

Human Transcriptome databases						
Database	**Subjects**	**Number of Samples**	**Number of genes**	**Spatial distribution**	**Temporal distribution**	**URL**
Allen Human Brain Atlas (AHBA)[7]	6	3,702	20,737	Whole brain	24–57 years	http://human.brain-map.org
Human Brain Transcriptome[9]	57	1,340	17,565	16 brain regions	5.7wpc–82 years	https://hbatlas.org/
Brain Cloud[10]	269	269	~14,500	Dorsolateral prefrontal cortex (BA46/9)	14wpc->80 years	https://www.ncbi.nlm.nih.gov/projects/gap/cgi-bin/study.cgi?study_id=phs000417.v2.p1
Developmental Transcriptome (BrainSpan)	35	492	17,604	16 brain regions	8wpc–40 years	http://www.brainspan.org/rnaseq/search/index.html
Human Prenatal Brain Development (BrainSpan) [11]	4	1,203	29,180	~25 areas (9 layers per area)	15wpc-21wpc	http://www.brainspan.org/lcm
Genotype-Tissue Expression project (GTEx)[12]	399	3,326	19,670	13 brain regions	20 years -71 years	https://www.gtexportal.org/
Single Cell Transcriptome databases						
Database	**Subjects**	**Number of single nuclei**	**Samples**		**Years**	**URL**
Allen Cell Type Database: M1 – 10x genomics	2	76,533	Primary motor cortex		50,60 years	https://celltypes.brain-map.org/
Allen Cell Type Database: Multiple Cortical Areas	50	49,495	Middle temporal gyrus, anterior cingulate cortex, primary visual cortex, primary motor cortex, primary somatosensory cortex, primary auditory cortex		18–83years	https://celltypes.brain-map.org/
single-cell analysis of transcriptional and epigenetic states in the human adult brain [13]	6	>60,000	Visual cortex, frontal cortex, and cerebellum		20 years – 49 years	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE97942
Genotype-Tissue Expression project – Single Cell data[14]	5	14,963	Hippocampus, frontal cortex		40 years – 65 years	https://gtexportal.org/home/datasets
human brain transcriptome diversity at the single cell level[15]	8 adults 4 fetal	466	Temporal lobe		21 years – 63 years 16–18w	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE67835
PsychENCODE - Human Brain Development [16]	3 adults 9 prenatal	17,093	Dorsolateral frontal cortex		18 years -64 years 5pcw-20pcw	http://development.psychencode.org/
RNA-Seq of human astrocytes [17]	14 adult 6 fetal		Temporal lobe		8 years – 63 years 17–20 pcw	https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE73721