Discovering Conserved Properties of Brain Organization Through Multimodal Integration and Interspecies Comparison

$SAGE

Ben D Fulcher (iD)

School of Physics, The University of Sydney, Sydney, NSW 2006, Australia.

ABSTRACT: The primate cerebral cortex is broadly organized along hierarchical processing streams underpinned by corresponding variation in the brain's microstructure and interareal connectivity patterns. Fulcher et al. recently demonstrated that a similar organization exists in the mouse cortex by combining independent datasets of cytoarchitecture, gene expression, cell densities, and long-range axonal connectivity. Using the T1w:T2w magnetic resonance imaging map as a common spatial reference for data-driven comparison of cortical gradients between mouse and human, we highlighted a common hierarchical expression pattern of numerous brain-related genes, providing new understanding of how systematic structural variation shapes functional specialization in mammalian brains. Reflecting on these findings, here we discuss how open neuroscience datasets, combined with advanced neuroinformatics approaches, will be crucial in the ongoing search for organization principles of brain structure. We explore the promises and challenges of integrative studies and argue that a tighter collaboration between experimental, statistical, and theoretical neuroscientists is needed to drive progress further.

KEYWORDS: neuroinformatics, interspecies comparison, spatial embedding, data integration

A core question in neuroscience is how the brain's structure enables the rich cognitive function that defines our conscious experience. Clinically, such knowledge would provide a mechanistic basis from which to design treatments to minimize cognitive deficits in disease, or to enhance cognitive function in healthy individuals. Technologically, it may open the door to neuromimicry: brain-inspired design of new algorithmic frameworks for artificial intelligence. Recent unprecedented investments in neuroscience research have yielded large volumes of data ranging from intricate whole-brain maps of the microscopic properties of neural circuits to neural dynamics measured across a wide range of spatial and temporal scales. Despite this deluge of data, we have relatively few organizing principles to explain the patterns, make predictions, and design new experiments to test them.[1–3] Progress in this direction requires closer collaboration between scientists with experimental, statistical, and theoretical modeling expertise. In this commentary, we highlight key results from our recent paper,[4] framing them in the context of the modern neuroscience landscape.

The Prominence of Neuroinformatics on the Landscape of Modern Neuroscience

In Fulcher et al.,[4] we assembled multiple datasets as spatial maps of the mouse cortex: (1) gene expression,[5] (2) cell type densities,[6,7] (3) axonal connectivity,[8] (4) cytoarchitecture,[9] and (5) magnetic resonance imaging (MRI) structural maps.[10] We demonstrated a surprising commonality of spatial variation between these diverse measurements, which together ordered mouse brain areas along a putative functional hierarchy, from primary somatosensory through to transmodal prefrontal areas. Our integrative approach provides clues about the multifaceted structural underpinning of hierarchical functional specialization[11] and allowed us to produce data-driven cortical organization schemes by representing each cortical area as a multimodal set of properties, such as that shown in Figure 1.

Our study is an example of integrative research that bridges the specialized expertise of many individual laboratories through data sharing. Following pioneering large-scale experimental work by the Allen Institute[5] and its associated informatics challenges,[13] the culture of data sharing in neuroscience has become widespread. For computational and theoretical scientists, this provides new opportunities to leverage what are large financial investments in cutting-edge experimental protocols to discover new properties of brain organization, including those that extend across datasets that would not traditionally be compared. This line of research is not without substantial challenges: compared with prototypical collaborations of experimental and theoretical scientists, analyses of open data involve an unusual level of disconnect between the process of data collection and statistical analysis. This requires much greater care on the part of the data analyst to proactively ensure that their analyses respect the subtleties of the dataset and any limitations in its measurement. For example, recent work comparing studies analyzing transcriptional data from the Allen Human Brain Atlas[14] found a high degree of inconsistency in how researchers subsequently processed the data and demonstrated that these choices can have a large impact on the final results and their interpretation.[15]
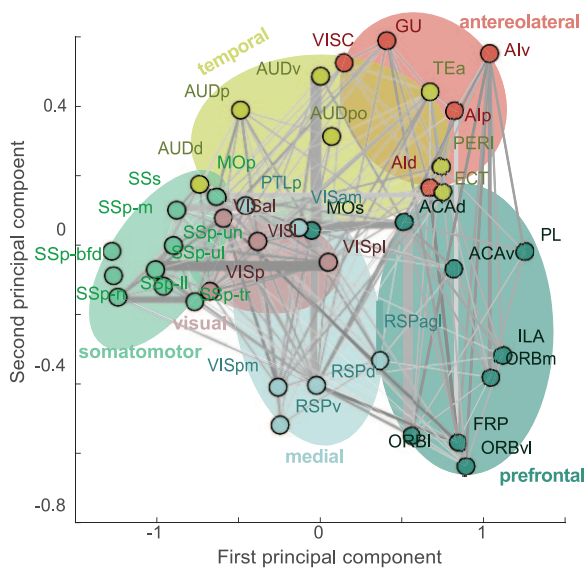
Figure 1. A low-dimensional projection of mouse brain cortical areas, represented as a multimodal feature vector. The feature vector combines properties of gene expression (*Pvalb, Grin3a*, and *Grik2*), cytoarchitecture type, weighted axonal in-degree, T1w:T2w, estimated hierarchical level, and mean density of parvalbumin-containing cells. Brain areas with a similar set of properties are close in this projection space. Shading has been added manually to highlight the different functional families, labeled according to Harris et al.[12]

Data-Driven Interspecies Comparison

Properties of brain organization that are shared across scales and species are strong candidates for providing an evolutionary advantage by enabling cost-efficient information processing. Although some such conserved properties are derived from abstractions of the brain as topological networks,[16] studies of comparative anatomy (most commonly performed between different primate species) require a cortical parcellation mapped between species. Both the choice of parcellation and its mapping require subjective choices on the part of the researcher.[17,18] In Fulcher et al.,[4] we circumvented the need to define a mouse-human homology of cortical areas, instead using the noninvasive MRI contrast map, T1w:T2w, as a common spatial reference for interspecies comparison. As well as confirming the increase of T1w:T2w from dysgranular to eulaminate areas in mouse, as in primate and human,[11] our results revealed a compelling consistency in the transcriptional patterning of brain-related genes relative to T1w:T2w. Mapping ortholog genes between mouse and human, we found significant mouse-human correspondence: Spearman $\rho = 0.44$ ($P = 1 \times 10^{-4}$), a result that was reproduced across a wide variety of gene sets: all genes, brain-expressed genes, astrocyte-enriched genes, neuron-enriched genes, and oligodendrocyte-enriched genes. As T1w:T2w is commonly interpreted as an MRI imaging marker for myelin content,[19] interspecies correspondence may be expected on the basis of common variation of genes coding for myelin content. Whereas myelin-related gene expression did

show interspecies correspondence, e.g., *Mobp* ($\rho_{mouse} = 0.43$, $\rho_{human} = 0.41$) and *Mbp* ($\rho_{mouse} = 0.34$, $\rho_{human} = 0.45$), the overall correlation was driven by functionally diverse genes, including the striking interspecies correspondences of N-methyl-D-aspartate (NMDA) receptor signaling genes such as *Grin3a* ($\rho_{mouse} = -0.63$, $\rho_{human} = -0.65$) and the interneuron marker *Pvalb* ($\rho_{mouse} = 0.57$, $\rho_{human} = 0.70$).

The agreement is striking given (1) measurement noise in both datasets, (2) differences in measurement modalities between mouse (high-throughput in situ hybridization) and human (post-mortem microarray from 6 adults), and (3) vast differences in spatial scale. Although we developed new quality-control methods for the data, we note that the atlas-based expression measurements can be very noisy and that a noisy measurement in one species would manifest as a reduced correlation with T1w:T2w. In this respect, our data are again consistent with a meaningful interspecies correspondence: a high correlation in one species consistently matched the same direction in the other species. This mouse-human conservation of gene expression gradients suggests that common structural features may shape hierarchical functional specialization in both mammalian brains.

Whereas our results demonstrate a surprising consistency in how transcriptional gradients vary with T1w:T2w, they also provide insights into key differences in the degree of specialization in the mouse cortex. Previous comparative studies of mouse and primate have focused on quantifying microstructural variation between areas with the most disparate function, e.g., between V1 and frontal cortex in mouse and between V1 and lateral prefrontal cortex in rhesus monkey.[20] This research has painted a consistent picture of the relative uniformity of the mouse brain compared with the highly differentiated primate cortex. In our study, correlations between T1w:T2w and other spatial maps were consistently weaker in mouse relative to macaque and human: for hierarchical level ($\rho_{macaque} = 0.76$, $\rho_{mouse} = 0.29$), cytoarchitectural type ($\tau_{macaque} = 0.87$, $\tau_{mouse} = 0.51$), and the leading principal component of gene transcription ($\rho_{mouse} = 0.53$, $\rho_{human} = 0.81$).[11] In characterizing the most salient mouse-human differences in brain structure, we also flagged candidate genes for further investigation that show the most difference between mouse and human, including NMDA receptor signaling genes *Grin2b/GRIN2B* ($\rho_m = 0.19$, $\rho_h = -0.63$), *Grin2d/GRIN2D* ($\rho_m = -0.41$, $\rho_h = 0.13$), and *Grin3b/GRIN3B* ($\rho_m = -0.34$, $\rho_h = 0.26$).

Validating Gene Markers for Cell Types

Bringing together measurements of gene expression[5] with direct measures of interneuron cell densities[6] allowed us to validate gene markers for interneuron cell types. We demonstrated that using gene expression data as a proxy for cell-type density can be highly accurate, as shown for the strongest correspondence we observed between the density of parvalbumin-containing interneurons[6] and the independently measured expression of
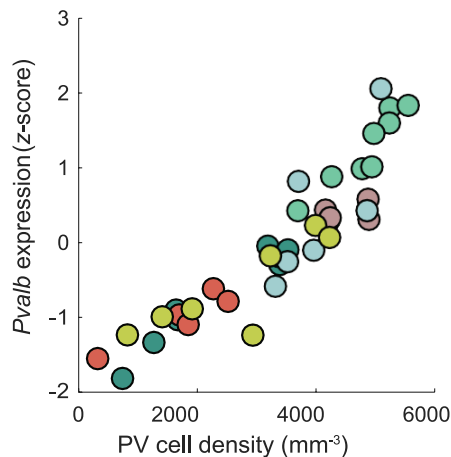
Figure 2. *Pvalb* expression is strongly correlated with an independent measure of parvalbumin containing cell density across the mouse cortex. We plot *Pvalb* expression (*z*-scored expression energy estimated from coronal section data from the Allen Mouse Brain Atlas[3]) as a function of direct quantitative measurement of parvalbumin containing cell density[4] across cortical areas. The 2 measurements are highly correlated, Spearman ρ = 0.95, increasing from prefrontal and anterolateral areas through to somatomotor areas (brain areas are colored as labeled in Figure 1).

Pvalb[5] across cortical areas in Figure 2. The correspondence for vasoactive intestinal polypeptide (VIP) cell density with *Vip* gene expression was also strong, ρ = 0.76 ($P = 3 \times 10^{-7}$), but was much weaker for somatostatin (SST) cell density with *Sst* expression, ρ = 0.24 ($P = 0.1$). The lack of agreement for *Sst* may be due to a breakdown in the relationship between *Sst* expression and SST cell density: *Sst* expression decreases in adults and is regulated by neural activity.[21] Our results suggest that the inference of cell density from gene expression atlases can be accurate, but a direct correspondence can be confounded by the many factors regulating gene transcription, as well as additional factors complicating straightforward inference of protein levels from mRNA.[22] Continuing the search for cell types and their marker genes,[23,24] and validating them against independent cell density measurements,[6,7] will equip us with tools to infer the spatial distributions of cell types, including their contribution to psychiatric phenotypes.

Spatial Continuity of the Macroscopic Brain

At the macroscale, the brain is a continuous physical system[25] that is commonly approximated as a set of discrete brain areas partitioned by hard boundaries. Our analysis brought together many open datasets of spatial maps reported at the level of the Allen Reference Atlas areas, but neuroscience datasets are now being measured and shared at cellular resolution.[26] These data open the door to future research overcoming the limitations of discrete parcellations, and moving toward a continuous characterization of the fine-scale spatial patterning of the brain's structural and functional properties. The next generation of brain atlases will likely contain a far more subtle representation of the brain's spatial organization that takes into

account multimodal information measured at each voxel,[27] likely yielding a combination of continuous gradient-like variation[28] as well as sudden changes consistent with interareal boundaries.

The Need for Theoretical Frameworks

Linking neuroscience big data to scientific understanding requires statistical analysis to uncover and characterize interesting patterns and theories that explain diverse observations in a coherent conceptual framework.[1] Seen through this lens, the multimodal statistical relationships we present in Fulcher et al.[4] will be important for constraining new mechanistic models of brain dynamics that bridge spatial scales of brain organization by capturing whole-brain activity dynamics in terms of putative physiological mechanisms. Whole-brain models, whether formulated physically (e.g., neural fields) or phenomenologically (e.g., as coupled dynamical systems), have overwhelmingly made the simplifying assumption of spatially uniform parameters when simulating brain dynamics. New models are beginning to incorporate inter-regional heterogeneity, allowing model parameters to vary with an empirical measurement such as spine count[29] or T1w:T2w.[30] Models including this simple 1-dimensional heterogeneity yield better out-of-sample fits to experimental functional-connectivity measurements and the hierarchical organization of intrinsic timescales,[30] providing evidence that hierarchical variation in local circuits plays a key role in shaping cortical dynamics. Physiologically based models, as opposed to phenomenological models, stand to gain the most from the rich new neuroscience datasets, as their formulation and parameters can be directly constrained and verified against data.[25]

The modern landscape of neuroscience has been shaped by advances in neuroinformatics, data-sharing standards, and greater contributions from statistical and computational modeling in generating guiding hypotheses for new experiments. Integrative studies like Fulcher et al.[4] that assemble many diverse datasets across multiple species will play a key role in uncovering the principles of how the brain's organization enables a diverse functional repertoire.

Author Contributions

BDF performed analysis and wrote the paper.

ORCID iD

Ben D Fulcher (iD) https://orcid.org/0000-0002-3003-4055

REFERENCES

1. Wang XJ, Kennedy H. Brain structure and dynamics across scales: in search of rules. *Curr Opin Neurobiol*. 2016;37:92–98.
2. Goulas A, Majka P, Rosa MGP, Hilgetag CC. A blueprint of mammalian cortical connectomes. PLOS Biology, 2019;17(3): e2005346.
3. García-Cabezas MÁ, Zikopoulos B, Barbas H. The Structural Model: a theory linking connections, plasticity, pathology, development and evolution of the cerebral cortex. Brain Structure and Function, 2019;72(12):1–24.
4. Fulcher BD, Murray JD, Zerbi V, Wang XJ. Multimodal gradients across mouse cortex. *Proc Natl Acad Sci U S A*. 2019;116:4689–4695.

5. Lein E, Hawrylycz MJ, Ao N, et al. Genome-wide atlas of gene expression in the adult mouse brain. *Nature*. 2007;445:168–176.

6. Kim Y, Yang GR, Pradhan K, et al. Brain-wide maps reveal stereotyped cell-type-based cortical architecture and subcortical sexual dimorphism. *Cell*. 2017;171: 456–469.

7. Erö C, Gewaltig MO, Keller D, Markram H. A cell atlas for the mouse brain. *Front Neuroinf*. 2018;12:e17727.

8. Oh SW, Harris JA, Ng L, et al. A mesoscale connectome of the mouse brain. *Nature*. 2014;508:207–214.

9. Goulas A, Uylings HBM, Hilgetag CC. Principles of ipsilateral and contralateral cortico-cortical connectivity in the mouse. *Brain Struct Funct*. 2016; 252:1.

10. Bakker R, Tiesinga P, Kötter R. The scalable brain atlas: instant web-based access to public brain atlases and related content. *Neuroinf*. 2015;13:353–366.

11. Burt JB, Demirtas M, Eckner WJ, et al. Hierarchy of transcriptomic specialization across human cortex captured by structural neuroimaging topography. *Nat Neurosci*. 2018;21:1251–1259.

12. Harris JA, Mihalas S, Hirokawa KE, et al. The organization of intracortical connections by layer and cell class in the mouse brain. *BioRxiv*. 2018;292961:1–33.

13. Ng L, Pathak SD, Kuan C, et al. Neuroinformatics for genome-wide 3-D gene expression mapping in the mouse brain. *IEEE/ACM Trans Comput Biol Bioinf*. 2007;4:382–393.

14. Hawrylycz MJ, Lein E, Guillozet-Bongaarts AL, et al. An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature*. 2012;489:391–399.

15. Arnatkevičiūtė A, Fulcher BD, Fornito A. A practical guide to linking brain-wide gene expression and neuroimaging data. *Neuroimage*. 2019;189: 353–367.

16. Van den Heuvel MP, Bullmore ET, Sporns O. Comparative connectomics. *Trends Cognit Sci*. 2016;20:345–361.

17. Orban GA, Van Essen DC, Vanduffel W. Comparative mapping of higher visual areas in monkeys and humans. *Trends Cognit Sci*. 2004;8:315–324.

18. Mars RB, Sotiropoulos SN, Passingham RE, Sallet J, Verhagen L, Khrapitchev, AA, et al. Whole brain comparative anatomy using connectivity blueprints. *eLife,* 2018;7:870.

19. Glasser MF, Van Essen DC. Mapping human cortical areas in vivo based on myelin content as revealed by T1- and T2-weighted MRI. *J Neurosci*. 2011;31: 11597–11616.

20. Luebke JI. Pyramidal neurons are not generalizable building blocks of cortical networks. *Front Neuroanat*. 2017;11:13644–13660.

21. Urban-Ciecko J, Barth AL. Somatostatin-expressing neurons in cortical networks. *Nat Rev Neurosci*. 2016;17:401–409.

22. Schwanhäusser B, Busse D, Li N, et al. Global quantification of mammalian gene expression control. *Nature*. 2011;473:337–342.

23. Tasic B, Menon V, Nguyen TN, et al. Adult mouse cortical cell taxonomy revealed by single cell transcriptomics. *Nat Neurosci*. 2016;19: 335–346.

24. Mancarci BO, Toker L, Tripathy S, et al. Cross-laboratory analysis of brain cell type transcriptomes with applications to interpretation of bulk tissue data. *eNeuro*. 2017;4:e0212.

25. Robinson PA. Physical brain connectomics. *Phys Rev E*. 2019;99:012421.

26. Fürth D, Vaissière T, Tzortzi O, et al. An interactive framework for whole-brain maps at cellular resolution. *Nat Neurosci*. 2017;21:139–149.

27. Glasser MF, Coalson TS, Robinson EC, et al. A multi-modal parcellation of human cerebral cortex. *Nature*. 2016;536:171–178.

28. Huntenburg JM, Bazin PL, Margulies DS. Large-scale gradients in human cortical organization. *Trends Cognit Sci*. 2017;22:21.

29. Chaudhuri R, Knoblauch K, Gariel MA, Kennedy H, Wang XJ. A large-scale circuit mechanism for hierarchical dynamical processing in the primate cortex. *Neuron*. 2015;88:419–431.

30. Demirtas M, Burt JB, Helmer M, et al. Hierarchical heterogeneity across human cortex shapes large-scale neural dynamics. *Neuron*. 2019;101:1181–1194.