Title
Transcriptional architecture of the human brain

Permalink
https://escholarship.org/uc/item/0k86s826

Journal
Nature Neuroscience, 18(12)

ISSN
1097-6256

Authors
Kelley, KW
Oldham, MC

Publication Date
2015-11-25

DOI
10.1038/nn.4178

Peer reviewed
Transcriptional architecture of the human brain

Kevin W Kelley & Michael C Oldham

The largest survey of gene expression ever performed in the adult human brain reveals highly stereotyped transcriptional patterning across individuals. The most stably patterned genes are enriched for neuronal annotations, disease associations, drug targets, and correspond to resting state functional networks.

An adult human brain consists of approximately 170 billion cells that come in an unknown number of varieties and form interconnectedness that number in the trillions. Is this complexity too much for a human brain itself to understand or is there a path through this paradox? 20 years ago, answers to this question might have tilted toward pessimism. Today, thanks to technological advances that provide new perspectives on the brain’s structural and functional organization from molecules to circuits, there is reason to be hopeful. In this issue of Nature Neuroscience, Hawrylycz et al. mark an important milestone on this journey with publication of the Allen Institute for Brain Science’s definitive survey of gene expression in the neurotypical adult human brain: the Allen Human Brain Atlas.

The study by Hawrylycz et al. is notable for two reasons. First, the data set produced by the authors is the largest of its kind. The authors used custom Agilent microarrays with >62,000 probes to measure gene expression levels in >3,700 meticulously macrodissected or laser-microdissected samples from six adult human brains. This extent of neuroanatomical coverage is unprecedented and provides the neuroscientific community with a massive resource that will be analyzed for many years to come. Second, to make sense of this onslaught of data, the authors introduce the concept of differential stability (DS), a correlation-based metric that quantifies the extent to which a gene shows reproducible expression patterns across neuroanatomical structures in different individuals (Fig. 1). The authors demonstrate that high-DS genes have a number of interesting properties, including significant enrichment with neuronal functions, disease associations and drug targets.

Focusing on high-DS genes, the authors constructed a consensus transcriptional network and used weighted gene coexpression network analysis (WGCNA) to identify 32 distinct groups of coexpressed genes (or modules) that were present in each individual and exhibited diverse neuroanatomical patterns. As shown previously, some of these modules corresponded to broad cell classes, reflecting variation in cellular abundance across samples, others showed marked regional specificity, and a third category consisted of genes associated with ubiquitous cellular organelles and protein complexes. Collectively, these 32 modules explained much of the overall variation in gene expression, despite a reduction from the original microarray probe feature space of more than three orders of magnitude. These findings underscore the power of gene coexpression analysis to distill the critical transcriptional themes that characterize biological systems.

Building on these findings, the authors next sought to determine the extent to which these 32 transcriptional signatures were conserved in the mouse brain using spatially quantified in situ hybridization data from the Allen Mouse Brain Atlas for 55 equivalent neuroanatomical structures. The authors observed strong conservation for a subset of modules with mostly neuronal characteristics; in contrast, modules with non-neuronal characteristics were less well conserved. These results are consistent with previous findings, but are somewhat difficult to interpret because of asymmetries in the distributions of different cell types across brain regions. Nevertheless, the authors uncovered many examples of genes whose neuroanatomical expression patterns appear to differ markedly between humans and mice, and thus warranting further investigation.

The consistency of transcriptional patterning in the human brain suggests that regional or areal differences in gene expression may mirror the organization of functional networks revealed by brain imaging studies. To explore this issue, Hawrylycz et al. calculated correlations between cortical gene expression patterns and resting-state functional activity obtained by functional magnetic resonance imaging (fMRI). These correlations were higher for genes with high DS in cortex, suggesting that conserved transcriptional patterning in the cortex may represent a genetic analog to group-average functional architecture revealed by fMRI. These findings will require further study to establish causality, but such integrative efforts should prove mutually beneficial for molecular and systems neuroscientists.

What next? At present, the Allen Human Brain Atlas represents the pinnacle of human brain gene expression profiling; however, as with most data sets, it was generated from bulk tissue samples comprising different cell types. Although gene coexpression analysis can reveal cell type–specific transcriptional signatures in bulk tissue samples, a conventional wisdom has emerged that single-cell analysis is required to delineate the full extent of cellular heterogeneity in the human brain. For example, a surprising result of bulk tissue profiling is the apparent homogeneity of gene expression in human association cortex, despite its diverse cytoarchitectonic characteristics. Is it possible that transcriptional signatures of rare cell types that distinguish these areas are undetectable at the level of bulk tissue analysis, even with the refined sampling strategy employed by the Allen Institute? In principle, single-cell...
Acquire >3,700 samples from 232 brain structures in 6 individuals
Perform gene expression profiling (>62,000 microarray probes)
Find highly conserved regionally patterned genes by calculating DS

Consistent anatomical patterning across individuals
Example high-DS gene

Use high-DS genes to identify reproducible transcriptional patterns
Assess conservation of transcriptional patterns in mouse

Cellular classes
Brain region enriched
Subcellular processes

Figure 1 Generating and analyzing the Allen Human Brain Atlas. Samples were macrodissected or laser-microdissected from six adult human brains using a standardized workflow. Gene expression was profiled in each sample using custom Agilent microarrays. The reproducibility of gene expression across individuals was quantified with the DS metric, defined as the average Pearson correlation of a gene’s expression levels over a defined set of anatomic regions for all pairs of individuals. Focusing on high-DS genes, the authors constructed a consensus transcriptional network and used WGCNA to identify 32 reproducible transcriptional signatures corresponding to broad cell classes, specific neuroanatomical regions and subcellular processes. Some (mostly neuronal) signatures were highly conserved in mouse brain, whereas others were not. Significant correlations were also observed between transcriptional patterns of genes with high DS in cortex and resting-state functional activation patterns revealed by fMRI. Ctx, cortex; HIP, hippocampus; Str, striatum; Th, thalamus; Med, medulla; Cb, cerebellum.

The stability of gene coexpression modules across individuals also provides a natural biological framework for exploring the functional context of candidate disease genes through the principle of ‘guilt by association’. However, many genetic variants that have been associated with disease do not fall within genes, but rather within genomic regions that are presumed, but in most cases not proven, to control gene expression. To identify genomic regions that control gene expression in the human brain, several consortia, including GTEx, CommonMind (http://www.commonmind.org/) and PsychENCODE (see p. 1707), are analyzing regulatory regions, epigenetic modifications and gene expression levels simultaneously in large collections of human brain samples from different ages, anatomical regions and disease cohorts. These data sets will provide a comprehensive multiomic perspective on transcriptional regulation in the human brain and an opportunity to assess the network organization of candidate control mechanisms in parallel with the network organization of gene expression.

The integration of new data modalities with transcriptomics is an exciting prospect, but one that should not diminish future efforts to understand transcriptome organization in its own right. Gene expression occupies a unique position in the flow of biological information in an organism: it represents the final common output of myriad regulatory mechanisms and the essential input for most biological processes. As such, it is likely that our collective efforts thus far have only scratched the surface of the meaning encoded by the human brain transcriptome. The Allen Human Brain Atlas will provide fertile ground for future efforts to decipher this code, particularly in human brain regions that have been under-represented in previous gene expression studies. At the same time, the information contained in this data set must be combined with other surveys of the human brain transcriptome to increase the representation of individuals and technology platforms.

The Allen Human Brain Atlas is a monumental resource whose scale is unlikely to be replicated anytime soon. However, this does not mean that the study by Hawrylycz et al. should mark the end of human brain transcriptomics; rather, it should mark the end of the beginning. The authors’ efforts have
Does the hippocampus preplay memories?

Howard Eichenbaum

Previous studies have reported ‘preplay’ of hippocampal neural activity patterns associated with events yet to occur. Silva et al. challenge this finding on the basis of large-scale recordings before and after experiences.

The Minority Report, a short story by Philip K. Dick (put to film in 2002 with Tom Cruise starring), described ‘precogs’, human mutants who could ‘preplay’ visions of future crimes in time for police to arrest suspects before the public comes to harm. This was great science fiction action, but it offered no insight into the neural mechanisms of these premonitions. In the real world, however, neural preplay has been observed in hippocampal neural activity patterns that foretell a sequence of place cell activations that will later occur when a mouse runs through a novel environment. In these experiments, hippocampal neurons are recorded as animals sleep, then again while the animals run on a track in a novel part of the environment, and preplay is observed during the sleep period, as sequential activation of neurons in the same order as they will later be observed during the novel run (Fig. 1). A paper in this issue of Nature Neuroscience now challenges these conclusions. Silva et al. report that they did not observe patterns that exhibited preplay during sleep; instead, place cell sequences developed only during experience (Fig. 1). These findings call for a reconsideration of whether the hippocampus employs pre-existing representations or develops representations for new experiences de novo.

Both the earlier reports of preplay and the findings of Silva et al. involve recordings of hippocampal place cells, neurons that fire selectively when a rodent occupies a particular location in its environment. When animals repeatedly traverse a narrow track, place cells fire sequentially, covering the total length of the track, and several studies have shown that these sequential spatial representations are recapitulated as temporal sequences of the same neuronal activations in later ‘offline’ periods when the animal sleeps or is awake but at rest (for example, see ref. 4). The consensual explanation of this ‘replay’ phenomenon was that serial activations of interconnected place cells would strengthen forward connectivity through synaptic plasticity mechanisms (Fig. 1).

**Figure 1** Models of memory-related sequence generation in neural hippocampal activity. In the preplay model, during sleep before the novel run experience, a subset of internal network connections are strong (thick synaptic connections) and specific spatial inputs that will later be made to those neurons are weak (dotted arrows). The sequence of cell activations is observed as a very brief series of spikes from the sequentially preconnected neurons (shown below). Then, during a subsequent run in a novel track, the spatial inputs corresponding to places traversed are strengthened (solid arrows), resulting in longer bursts by sequentially activated cells. By contrast, in the learned sequence model, during sleep before the novel experience, intrinsic connections are all weak (thin synaptic connections), whereas the spatial inputs are assumed to be prewired (solid arrows). In this situation, firing during sleep is observed as randomly ordered spontaneous activity. Then, during the subsequent novel run, the activation of cells by sequentially experienced locations strengthens connections between those neurons (thick synaptic connections) as they fire in sequence, resulting in long bursts of each cell at each sequential location.