What is the value of a scalable, accessible, molecularly annotated connectomics technology?

Adam H Marblestone, Desiree D Dudley, Edward S Boyden

What we hope to achieve

Today, obtaining detailed structural and molecular maps of neural circuitry is extremely difficult. Briefly, we are proposing to develop a set of robust and accessible tools -- based on next-generation >20x expansion microscopy, multicolor combinatorial neuron labeling, and unique molecular barcoding of neuron identity -- which, we anticipate, will ultimately allow the cheap, fast readout of neural connectivity using conventional optical microscopes.

Value proposition for low-cost, accessible, dense, molecularly annotated optical connectomics

This document aims to detail the prospective value proposition for neuroscience of such a potential future technology, assuming that this technology can be achieved technically and distributed widely. There are at least four key aspects of the value proposition:

1) **Democratizing connectomics**

While a single map of a single brain might be a powerful dataset, neuroscience would benefit most from a fast, cheap, accessible connectomics method that can be customized and applied for a wide variety of questions -- including questions about disease mechanisms, many forms of neural circuit computation, behavior, development and other areas across a wide variety of labs. The proposed optical approach, which combines multicolor morphology tracing and spatially resolved molecular readouts with barcoded error correction inside intact transparent specimens, on conventional and ubiquitously used optical microscopes, should be more accessible than existing approaches based on ultra-thin sectioning, electron microscopy, and large-scale machine learning of image segmentation. The fully co-registered molecular information will also make connectomics relevant to a much wider set of neuroscientific questions, such as molecular and developmental underpinnings of the connectome (i.e., how the genome encodes a “program” for “learning” the connectome based on a complex interaction between nature and nurture) or about disease (see the next item).

2) **Bridging molecular and circuit-level understandings of disease**

Molecular information is important from a pathology/disease standpoint. Many brain disorders are regarded as genetic/molecular disorders. At the same time, we understand that these conditions are circuit-specific, and that they are emergent in the development of neural circuit

anatomy, molecular architecture, homeostasis, and dynamics. Indeed, conditions such as autism and schizophrenia are widely thought to be, at least in part, “connectopathies,” i.e., abnormalities of the connectome.

Maps that show not only the circuitry, but the makeup of that circuitry - with nano-scale spatial resolution, architecture, and *molecular changes* - are needed to understand the complex relationships that give rise to healthy and pathological mechanisms of interaction. These include transcriptomic and proteomic changes in cells, changes in receptor distribution at synapses, changes in neurotransmitter synthesis or neuromodulatory receptor expression, but also including specific pathological molecular variables (like the protein compositions of amyloid plaques in Alzheimer’s). The ability to do this rapidly at both nanoscale resolution, and within a greater map of entire intact circuits, may point to new hypotheses which would connect genomic insights to emergent mechanisms. In short, scalable, accessible, molecularly resolved maps of intact neural circuits could allow us to better interpret genomic variation in the context of actual neural circuit pathology. Here, note that the role of the brain maps is not in and of themselves to provide a complete understanding of pathology, but to provide an unbiased and rich jumping off point to develop new ideas about the mechanisms of disease, which can be tested through hypothesis-driven perturbation, dynamic observation, modeling, and other approaches.

3) Constraining the hypothesis space for neural circuit mechanisms of computation

Neuroscience currently has powerful tools for perturbing small numbers of specific “nodes” in a neural circuit, for observing small numbers of other “nodes,” and for providing stimuli to an organism and observing its behavior. These tools are powerful for *testing* hypotheses, but often don’t constrain the hypotheses themselves as much as researchers would wish. Perhaps this is because the relevant hypotheses, in this context, are statements about the architecture of information flow in a network of circuitry that is far larger and more complex than the set of observable/perturbable nodes. Therefore, seeing the connectivity of the circuitry itself can suggest a more targeted hypothesis space, which can then be tested using more traditional approaches.

We would argue that this idea of using connectomes -- or more generally, molecularly annotated circuit diagrams -- for hypothesis-reduction, hypothesis-generation, and hypothesis-inspiration applies quite broadly. Understanding how the brain works will require a mixture of “big data” analysis on these and other types of maps -- e.g., finding circuit motifs, analyzing variations between cortical columns, or clustering cell types -- and hypothesis-driven experimentation, e.g., validating the anatomical predictions of the thalamic higher-order relay hypothesis. Connectomic maps are likely to be helpful for some of these studies, and essential for others. Finally, theories at multiple levels of abstraction will need to be developed, many of which would derive from higher-level ideas in psychology or AI, rather than being “inferred” from the connectome. Crucially, these analyses and experiments should be carried out by a broad community of scientists. We cannot know in advance with any certainty which approaches will yield the greatest insight. Only after this integrative science has moved forward do we expect

that some of the most transformative potential applications of connectomics could potentially come to light.

4) **Potentially enabling biophysically detailed models of neural circuitry, which can be compared directly with dynamic activity measurements on the exact same tissue**

So far, nobody has been able to create an accurate, fully reductionistic, physio-chemical simulation of an entire neural circuit. An ideal model would take advantage of knowledge about 1) internal correlations inside a single brain, i.e., what is lost by attempted "averaging" over many tiny parts of many different brains under different conditions, 2) detail, i.e., the measurements used to guide the construction of most existing models are extremely sparse. As a notable example of internal correlations, as well as of the links between molecules and connectivity, the Marder lab has found even in the simple crab nervous system that connected cells have highly correlated expression of ion channel genes, apparently to ensure correct circuit-level homeostasis of neural dynamics. We detail why a deeper level of measurement is essential to underpin understanding of neural circuits in our "Assumption-Proof Brain Mapping" essay here:

http://syntheticneurobiology.org/PDFs/14.09.marblestone.short_and_extended.pdf

We are proposing to develop technologies that could enable the first full-circuit connectomes with internal correlations and molecular annotations. Such connectomes could ultimately be "plugged in" rather readily to existing modeling infrastructure. Such dense, fully internally correlated static maps could inform much-improved models, as detailed in the next sub-section.

How can static maps inform dynamical models?

One question that can be asked is: "How can static maps inform dynamic models?" This is a good question, and one that we have thought about in detail. For one thing, there are arguably major advantages of a static representation. It simply may not be possible, in a living brain, to see all the relevant molecules and connections -- achieving sufficient resolution and sufficient diversity of molecular tags would perturb the living brain too much. Thus, to get to the level of biophysical detail needed to specify a circuit's biophysics in a fully reductionistic manner, researchers may have no choice but to make detailed static anatomical and molecular maps. What is new in what we are proposing is a set of technologies prospectively leading to the ability to do this cheaply and accurately for entire circuits, including the molecular information.

Importantly, once we make such a dynamical model, we will have to validate the model through dynamic measurements and perturbations. Conveniently, we and many others are making great progress on recording dynamic activity from large populations of neurons, e.g., through the technique of "calcium imaging." Recent advances report activity mapping across the entire brains of worms and larval zebrafish. Thus, with a scalable connectomic technology, one could

do activity mapping for many seconds or minutes, and then fix the brain and obtain its molecularly annotated connectome. If this were done, there are at least two (and potentially many more) modeling studies that could be tried:

a. Using the molecularly annotated connectome, construct a model, and use it to predict circuit dynamics. Validate the prediction with the actual measured circuit dynamics, in the context of targeted perturbations (e.g., using optogenetics).
b. Combine the molecularly annotated connectome, and the activity maps, to create a predictive model of the future activity of the network, given its past activity. Validate with actual measurements and perturbations of activity.

Needless to say, one cannot expect to be able to make a working model right out of the box. We anticipate that such models would recapitulate some aspects of neural activity, but not others, and thus that they could fail in interesting ways. By observing how they fail, researchers could learn what else needs to be mapped to constrain free parameters, or how a model needs to be tuned, or perhaps might uncover flaws in the entire modeling paradigm (such as a failure to account for molecular changes that occur on longer timescales). What we do anticipate is that detailed static maps -- once they reach a certain level of detail, accuracy and completeness -- will be extremely valuable in informing the construction and refinement of such biophysical models. And such biophysical models can then inform models at higher levels of abstraction.

Responses to common questions

Finally, a few notes dealing with common questions from the neuroscience community regarding "connectomic" efforts in general, in the context of our specific proposition:

1) Scope and diversity of data needed: Connectomes, alone, are insufficient as reductionist descriptions of a neural circuit -- they must be combined with molecular and dynamic information, and perhaps of course behavioral data: Some neuroscientists think of the "connectome" as describing the shapes of cells and the topology of how they are connected, but we argue that knowing the key molecules (ion channels, neurotransmitter receptors, etc.) is critical. Without the ion channels and neurotransmitter receptors, for example, it will not be clear whether a connection between two neurons is fast or slow, strong or weak, excitatory or inhibitory. In contrast, knowing the shapes of cells in the brain and how they are connected, as well as the ion channels, neurotransmitter receptors, and other key molecules that give neurons their physiological properties, will enable a fuller picture of the dynamics of the brain. Many have pointed out that the connectome is "necessary but not sufficient" data for understanding the brain (citing complex neural dynamics as well as complex extra-synaptic neuromodulatory pathways). Molecular annotation of the connectome reduces the risk that the resulting data will be incomplete as a structural characterization of the system. Next steps, for combining connectomes with dynamic information, would be to obtain connectomes on organisms for which large-scale brain activity mapping has already been carried out.

2) Importance of cost reduction and democratic deployment of tools to diverse investigators: Connectomes must be cheap and widely accessible. We will need not just one connectome but many -- studying the effects *on* the connectome of varying genetic programs, disease states, and stimuli, and the effects *of* the connectome on brain activity and behavior. In order to enable this, we must dramatically decrease the cost of connectomic mapping to allow the methods to be widely applied: the technology development necessary to radically reduce costs in this way is perhaps the most important projected impact of our proposal.
3) Theories, models and hypothesis-testing: To reiterate, we view theories, hypotheses, and models as crucial aspects of neuroscience. We wish to enable the creation of improved neural circuit maps to boost these scientific efforts, not to replace them. We are not proposing a monolithic, atheoretical mapping effort to the exclusion of other neuroscience efforts, but rather the development of a set of tools that can be used widely to democratize and enrich neural mapping -- enabling detailed brain maps to be used in a wide variety of contexts to guide and validate diverse theories and experimental approaches.