Neuroinform (2010) 8:1–3
DOI 10.1007/s12021-010-9063-0

EDITORIAL

# The Coming of Age of the Hippocampome

**Giorgio A. Ascoli**

Published online: 3 February 2010


The "connectome" or connectivity map of the brain is a long-standing goal in neuroscience. Once a visionary dream of Golgi and Cajal, charting the mammalian connectome has been recently recognized as a pressing, if challenging, priority in biomedical research (e.g. Roysam et al. 2009). Different neuroscience sub-communities are targeting a broad range of scales, from the identification of all individual synapses in model organisms with electron microscopy, to long-range regional connections in the human brain with non-invasive imaging. How complete a connectome could realistically be expected in a finite time frame depends on the analytical level of detail.

Genetic strategies are particularly promising for the potential to visualize the axonal and dendritic arborizations of many or even all individual neurons throughout the nervous system (Lichtman et al. 2008). Still, reconstructing the entirety of the neuropil of a mouse brain as a routine lab technique (i.e. in the span of days rather than years or centuries), would require tools to trace the fine branching processes at a speed greater than that of commercial aircrafts. This corresponds to an improvement of nine orders of magnitudes relative to the current state of the art (e.g. Fan et al. 2009; Yuan et al. 2009; see also Ascoli 2008).

At a coarser scale, the National Institutes of Health have recently released a Request for Applications calling for a human connectome based on diffusion tensor imaging and other existing technology based on magnetic resonance and genotyping (http://grants.nih.gov/grants/guide/rfa-files/RFA-MH-10-020.html). While undoubtedly constituting a most valuable resource in clinical and basic research alike, such a map would only provide a low-resolution blueprint of regional connections, rather than the finer level of neuronal circuitry typically considered in computational neuroscience. A similar consideration applies to recently envisioned programs for systematically tracing fiber tracts in the rodent (Bohland et al. 2009). Although the term "connectome" has been in fact introduced in modern times to describe this level of connectivity mapping (Sporns et al. 2005), others have proposed to refer to projection atlases as the "projectome" (Kasthuri and Lichtman 2007).

There is an important intermediate scale between that of individual synapses and that of regional connections. This is the representation of circuit diagrams that capture the pattern of connectivity among identified *cell types*. In fact, pre- and post-synaptic targets are considered among the most functionally meaningful determinants of neuronal classes (Ascoli et al. 2008). The connection from a given cell type to another can be described in statistical terms, i.e. as a probability that the axon of any particular neuron in the former would contact the dendrite of any particular neuron in the latter.

Current technology and knowledge may be mature to achieve such a probabilistic mapping of connectivity at the level of neuronal classes for specific portions of the mammalian nervous system, such as the rodent hippocampus. An enormous amount of data is known about this particular structure, which has attracted the attention of neuroscientists since the early days of neuroscience (Golgi et al. 1886/2001). A database of statistical synaptic connection between any given pair of individual pre- and post-synaptic neurons would represent an invaluable resource for both computational applications (e.g. Koene et al. 2009) and empirical investigations. In addition, a cell-

G. A. Ascoli (✉)
Center for Neural Informatics, Structures, & Plasticity
and Molecular Neuroscience Department, Krasnow Institute
for Advanced Study, George Mason University,
MS2A1, 4400 University Dr.,
Fairfax, VA 22030, USA
e-mail: ascoli@gmu.edu

level connectivity map could also constitute a natural framework to build a more comprehensive information system by adding physiological and molecular features to the underlying anatomical foundation.

Such a probabilistic atlas of the entire synaptic blueprint of the rat hippocampus may be referred to as the "hippocampome". From the computational point of view this level of representation is particularly useful, because cell assemblies, rather than individual neurons, are commonly considered to implement basic cognitive functions. The hippocampome may be viewed as capturing the general macrostate of hippocampal connectivity in all individual subjects. An individual hippocampus could be instantiated in a neural network simulation by stochastic sampling of a microstate within the statistical constraints provided by the hippocampome probability distributions.

What might come as a surprise is the tremendous complexity that even this level of summary information entails and the robust neuroinformatics framework that the successful completion of such an endeavor will surely require. Considering that we are not attempting to list all connections of every individual neuron, but just the connection probabilities among neuron types, one might expect only a small set of quantities of which to keep track. In neuroanatomy, however, numbers add up quickly. Even after drastic oversimplifications, the sheer number of parameters necessary to fully define this knowledge hovers in the tens of thousands. Why is the number of required parameters so staggering?

A far from comprehensive scheme of the multi-synaptic loop of the rodent hippocampus consists of the dentate gyrus, CA3, CA1, subiculum, and entorhinal cortex. Between five and twenty cell classes have been characterized for each of these areas, for a grand total of at least 70 neuron types. Only roughly 20% of all possible connections among cell classes actually exist: for example, dentate gyrus granule cells contact CA3 pyramidal cells, but not CA1 pyramidal cells. However, noting the absence of a connection is also important. A simple identification of which connections are present, absent, or unknown would require 4900 ternary entries. Then, a probabilistic database of the cellular connectivity of the hippocampus would have to account for the statistics of ~1,000 ($70^2 \times 0.2$) connections, give or take a few.

Connectivity may be expressed as the number of post-synaptic targets a given efferent cell contacts, or the number of pre-synaptic afferents a given neuron receives contacts from. In either case, such number will be represented across all neurons of a class pair (pre-post) as a probability distribution function. Statistical distributions are typically characterized with two variables, such as mean and standard deviation for normal distributions, or minimum and maximum for uniform distributions. Furthermore, each of these descriptors varies considerably with the anatomical locations of both the pre- and the post-synaptic neurons. In the general sense, one should consider connectivity gradients in three spatial dimensions for source and target (i.e., six independent variables). At a minimum, spatial dependence can be (drastically) approximated by the distances between the source and target neurons along just the two principal axes of the hippocampus: transverse (proximal/distal CA3 or CA1, and infra/supra-pyramidal blades of dentate gyrus), and longitudinal (septal/dorsal vs. temporal/ventral).

Assuming for the sake of simplicity (but once again compromising already the realism of description) that each of these two factors affects the probability of connection linearly and independently, every statistical variable can be approximated with a 3-parameter function, bringing the total number of numbers to account for to $2 \times 3 \times 1{,}000 = 6{,}000$. In the laminar organization of the hippocampus, synapses from specific sources are typically located within a characteristic layer in the dendrite of the receiving neuron. The distance from the soma along the dendritic path of the input location has potential implications for synaptic computation, because of cable properties and active signal transformation. Similarly, the distance of the contact from the soma of the pre-synaptic neuron along its axon may affect communication delay and/or failure. Keeping track of the statistics of the synaptic distances along axons and dendrites triples the number of parameters.

This back-of-the-envelope computation invites an important question. For how many of these ~18,000 parameters does the literature offer at least rough estimates? Even if it is likely to be a small proportion, attempting a real answer turns out to constitute an eye-opening exercise. Just determining what information is or is not contained in the massive body of published scientific knowledge is in and by itself a non-trivial problem. Simply listing the parameters for which relevant peer-reviewed references are available, and tentatively identifying the remaining ones as "open problems" could be considerably useful. A related question regards the most suitable informatics framework to represent this information. A combination of spatial visualization tools, ontologies, and relational database appears necessary (e.g. Bezgin et al. 2009). At the same time, the dynamic aspect of information acquisition through ongoing annotation from the continuously growing scientific literature also calls for an integrated wiki system open to curated contributions from active researchers, such as in Neurolex.org (Bug et al. 2008).

Despite such complexity, we submit that this is indeed prime time for this kind of neuroinformatics approach. There are several reasons for this opinion. First, the fraction of parameters for which empirical data is available, while in the minority, already sums up to a respectable amount, at least for structures such as the rodent hippocampus. This

was not necessarily the case even only two decades ago. Second, the continuous rate of knowledge accumulation has already surpassed human capacity. In other words, reading and annotating all scientific papers that are being published on cellular connectivity would take an amount of time that far exceed that typically available to professional researchers. Third, the technology is now being developed to organize, represent and communicate this knowledge: the "informatics infrastructure" of neuroscience that the field of neuroinformatics embodies. Last but not least, there is an opportunity or even necessity to prioritize "what data to acquire next" based on the most pressing open problems.

Several of the considerations outlined here are valid not just for the hippocampome, but for analogous connectomes of other regions of the mammalian nervous system, including but not limited to the retina, spinal cord, cerebellum, thalamus, basal ganglia, and neocortex. Yet it is fascinating to note the remarkable actuality of the 124-year old words of Golgi (in the 2001 translation by Marina Bentivoglio and Larry Swanson):

> "The *pes Hippocampi major* is a region of the cerebrum that presents a very complex structure. There is a great deal of interest in studying this structure because determining the exact shape of the cells that form it, and the course and organization of its nerve fiber bundles, may shed light on its function […] it seems likely that further work along the same lines will complete the description of these bundles, and that it will be possible to obtain clues about their functions, along with those of the corresponding cell groups."

## References


Ascoli, G. A. (2008). Neuroinformatics grand challenges. *Neuroinformatics, 6*(1), 1–3.

Ascoli, G. A., Alonso-Nanclares, L., Anderson, S. A., Barrionuevo, G., Benavides-Piccione, R., Burkhalter, A., et al. (2008). Petilla terminology: nomenclature of features of GABAergic interneurons of the cerebral cortex. *Nature Reviews Neuroscience, 9*(7), 557–568.

Bezgin, G., Reid, A. T., Schubert, D., & Kötter, R. (2009). Matching spatial with ontological brain regions using Java tools for visualization, database access, and integrated data analysis. *Neuroinformatics, 7*(1), 7–22.

Bohland, J. W., Wu, C., Barbas, H., Bokil, H., Bota, M., Breiter, H. C., et al. (2009). A proposal for a coordinated effort for the determination of brainwide neuroanatomical connectivity in model organisms at a mesoscopic scale. *PLoS Computational Biology, 5*(3), e1000334.

Bug, W. J., Ascoli, G. A., Grethe, J. S., Gupta, A., Fennema-Notestine, C., Laird, A. R., et al. (2008). The NIFSTD and BIRNLex vocabularies: building comprehensive ontologies for neuroscience. *Neuroinformatics, 6*(3), 175–194.

Fan, J., Zhou, X., Dy, J. G., Zhang, Y., & Wong, S. T. (2009). An automated pipeline for dendrite spine detection and tracking of 3D optical microscopy neuron images of in vivo mouse models. *Neuroinformatics, 7*(2), 113–130.

Golgi, C., Bentivoglio, M., & Swanson, L. (2001). On the fine structure of the pes Hippocampi major (with plates XIII–XXIII). 1886. *Brain Research Bulletin, 54*(5), 461–483.

Kasthuri, N., & Lichtman, J. W. (2007). The rise of the 'projectome'. *Nature Methods, 4*(4), 307–308.

Koene, R. A., Tijms, B., van Hees, P., Postma, F., de Ridder, A., Ramakers, G. J., et al. (2009). NETMORPH: a framework for the stochastic generation of large scale neuronal networks with realistic neuron morphologies. *Neuroinformatics, 7*(3), 195–210.

Lichtman, J. W., Livet, J., & Sanes, J. R. (2008). A technicolour approach to the connectome. *Nature Reviews Neuroscience, 9*(6), 417–422.

Roysam, B., Shain, W., & Ascoli, G. A. (2009). The central role of neuroinformatics in the National Academy of Engineering's grandest challenge: reverse engineer the brain. *Neuroinformatics, 7*(1), 1–5.

Sporns, O., Tononi, G., & Kötter, R. (2005). The human connectome: a structural description of the human brain. *PLoS Computational Biology, 1*(4), e42.

Yuan, X., Trachtenberg, J. T., Potter, S. M., & Roysam, B. (2009). MDL constrained 3-D grayscale skeletonization algorithm for automated extraction of dendrites and spines from fluorescence confocal images. *Neuroinformatics, 7*(4), 213–232.