CONTENTS

MAPPING THE BRAIN

COMPUTATION

IMPLICATIONS

AFTERWORD

MAPPING THE BRAIN

In the jargon of neuroscience, to map the brain is to understand two things: all of the brain's myriad connections (equivalent to drawing a map of all the roads and buildings in the United States) and all of the "traffic" (neural activity that occurs on those roads). "Connectomes" are like highway maps, "activity maps" record the traffic as the brain is engaged in behavior. Like Google Maps, we ultimately need many "layers" of information, telling us about landmarks (like the folds of the cortex), annotations about particular types of neurons (the brain likely has close to a thousand), and ultimately about the pathways of neurons that are involved in particular kinds of behaviors.

The essays in this part tell a story—from the current, cutting edge to the future—about technological advances that will allow us to map out as much of that territory as possible. Most complex organisms have hundreds of thousands, if not millions or billions, of neurons. For decades, neuroscientists have recorded from just a few at a time, inferring something about a complex system based on incomplete measurements. **Mike Hawrylycz** narrates the history of brain anatomy, from the earliest drawings of neural circuits by Ramón y Cajal to ongoing, cutting-edge efforts to obtain and annotate high-resolution anatomical maps of the entire human brain at cellular resolution. **Misha Ahrens** describes an approach called light-sheet microscopy for monitoring neural activity from the *entire* brain of a transparent organism, the zebrafish, and to do so during behavior in intact animals. **Christof Koch** describes a confluence of emerging methods—anatomical, physiological, and optical—that are making it possible to characterize neural activity across large swaths of the visual cortex of the mouse. Looking further into the future, **Anthony Zador** and **George Church** describe novel approaches to characterizing neural anatomy, specifically neural connectivity, that use genetic techniques to indirectly encode information about connectivity in sequences of DNA. **Church** discusses how these approaches might even be extended to record the firing of neurons over scales much larger than optical or electrophysiological methods currently allow.

BUILDING ATLASES OF THE BRAIN

Mike Hawrylycz

With Chinh Dang, Christof Koch, and Hongkui Zeng

A Very Brief History of Brain Atlases

The earliest known significant works on human anatomy were collected by the Greek physician Claudius Galen around 200 BCE. This ancient corpus remained the dominant viewpoint through the Middle Ages until the classic work *De humani corporis fabrica* (*On the Fabric of the Human Body*) by Andreas Vesalius of Padua (1514–1564), the first modern anatomist. Even today many of Vesalius's drawings are astonishing to study and are largely accurate. For nearly two centuries scholars have recognized that the brain is compartmentalized into distinct regions, and this organization is preserved throughout mammals in general. However, comprehending the structural organization and function of the nervous system remains one of the primary challenges in neuroscience. To analyze and record their findings neuroanatomists develop atlases or maps of the brain similar to those cartographers produce.

The state of our understanding today of an integrated plan of brain function remains incomplete. Rather than indicating a lack of effort, this observation highlights the profound complexity and interconnectivity of all but the simplest neural structures. Laying the foundation of cellular neuroscience, Santiago Ramón y Cajal (1852–1934) drew and classified many types of neurons and speculated that the brain consists of an interconnected network of distinct neurons, as opposed to a more continuous web. While brain tissue is only semitranslucent, obscuring neuronal level resolution, a certain histological stain Franz Nissl (1860–1919) discovered, and known as the *Nissl* stain, can be used to stain negatively charged RNA in the cell nucleus in blue or other visible colors. The development of this stain allowed the German neuroanatomist

Figure 1. a. Cover of the work *De humani Corporis fabrica libri septem*, published by Andreas Vesalius in 1543. The work was the first major advance in human anatomy since the Greek physician Galen. b. A page from the fifth chapter of the book showing the cortex and ventricles of the brain.

Korbinian Brodmann (1868–1918) to identify forty-three distinct regions of the human cerebral cortex based on cytoarchitectural organization using this Nissl stain. These pioneering works of Brodmann, Constantin von Economo, Marthe Vogt, and others mapped cyto- and myeloarchitectural landscape of the human cortex based on painstaking visual inspection and characterization of a few observable cellular properties such as cell shape, density, packing, and such.

Since Vesalius, most atlases of the brain have been drawn on paper, with the most recent versions in vivid color delineating hundreds of structures. Such atlases have been drawn for most of the important model organisms studied in the laboratory and provide key bench-side experimental references. As with most aspects of modern biology, however, technology has been a driving factor in improved understanding of brain organization. Neuroimaging techniques evolved over the last twenty years have now allowed neuroscientists to revisit the subject of brain mapping, with the modern brain atlas more akin to a digital database that can capture the spatiotemporal distribution of a multitude of physiological and anatomical data. Modern techniques such as magnetic resonance imaging (MRI), functional magnetic resonance imaging

(fMRI), diffusion MRI, magnetoencephalography (MEG), electroencephalography (EEG), and positron emission tomography (PET) have provided dramatic improvements in brain imaging for research, clinical diagnosis, and surgery. Digital atlases based on these techniques are advantageous since they can be *warped*, mathematically or *in silico*, to fit each individual brain's unique anatomy.

The origin of modern brain mapping for clinical use lies with the seminal work of Jean Talairach, who in 1967 developed a 3D coordinate space to assist deep brain surgical methods. This atlas was generated from two series of sections from a single sixty-year-old female brain, and was later updated by Talairach and P. Tournoux in a printed atlas design for guiding surgery. Today biomedical imaging forms a crucial part of diagnosis and presurgical planning, and much time and resources are invested in the search of imaging biomarkers for diseases. Atlases have been used in image-guided neurosurgery to help plan "stereotaxic," that is, coordinate referenced, neurosurgical procedures. Using this data, surgeons are able to interpret patient-specific image volumes for anatomical, functional, and vascular relevance as well as their relationships.

The field of digital atlasing is extensive and includes high-quality brain atlases of the mouse, rat, rhesus macaque, human, and other model organisms. In addition to atlases based on histology, magnetic resonance imaging, and positron emission tomography, modern digital atlases use gene expression, connectivity, and probabilistic and multimodal techniques, as well as sophisticated visualization software. More recently, with the work of Alan Evans at the Montreal Neurological Institute and colleagues, averaged standards were created such as the Colin27, a multiple scan of a single young man, as well as the highly accessed MNI152 standard. While inherently preserving the 3D geometry of the brain, imaging modalities such as MRI, CT, and PET do not usually allow for detailed analysis of certain structures in the brain because of limitations in spatial resolution. For this reason it is common to use very high-resolution 2D imaging of *in vitro* tissue sections and employ mathematically sophisticated reconstruction algorithms to place these sections back into the 3D context of the brain.

Today digital brain atlases are used in neuroscience to characterize the spatial organization of neuronal structures, for planning and guidance during neurosurgery, and as a reference for interpreting other data
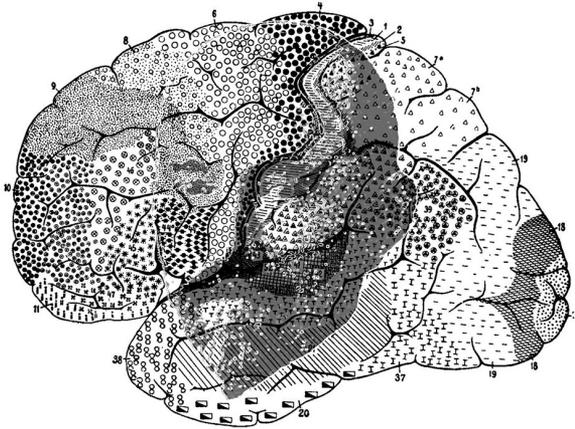
Figure 2. Regions of the human cerebral cortex delineated by Korbinian Brodmann using Nissl stain histology. Brodmann identified forty-three distinct regions that today still serve as a guide for studying distinct functional areas in the human cortex.

modalities such as gene expression or proteomic data. One ultimate aim of neuroscientific inquiry is to gain an understanding of the brain and how its workings relate to activities from behavior to consciousness. Toward this end, brain atlases form a common coordinate framework for summarizing, accessing, and organizing this knowledge and will undoubtedly remain a critical-path technology in the future.

The Genetic Brain

The development of the techniques of modern molecular biology and eventually whole genome sequencing opened the door for understanding the genetics of the brain, and new perspectives on the study of brain anatomy are emerging with the availability of large-scale spatial gene expression data. The brain consists of at least several hundred distinct cell types whose complete classification is still at present elusive. Each cell type is related to its function with its gene expression pattern, for example, on/off, high/low, as a key determinant. Gene expression data can be collected through a variety of techniques, and exploration of these

data promises to deliver new insights into the understanding of relations between genes and brain structure.

Early gene expression studies used methods such as northern blots, which combine electrophoresis separation of RNA molecules followed by hybridizing probes for detection. At one time this method was the gold standard for confirming gene expression, but it ultimately gave way to more quantitative methods. The microarray revolution dramatically increased our ability to profile genes by hybridizing many gene probes on a single gene chip. Today rapid digital sequencing technology can count individual RNA fragments that can subsequently be mapped back to the genome once it is known for an organism.

In 2001, Paul Allen, cofounder of Microsoft, assembled a group of scientists, including James Watson of Cold Spring Harbor Laboratory and Steven Pinker, then at MIT, to discuss the future of neuroscience and what could be done to accelerate neuroscience research. During these meetings the idea emerged that a complete 3D atlas of gene expression in the mouse brain would be of great use to the neuroscience community. The mouse was chosen due to the wealth of existing genetic studies and for practical reasons. Of the potential possible techniques, the project chose a technique for mapping gene expression called *in situ* hybridization (ISH) (automated by Gregor Eichele of the Max Planck Institute and colleagues), which uses probes that bind to mRNA within sectioned but intact brain tissue and thereby preserves spatial context (see color plate 1).

In 2006, an interdisciplinary scientific team at the Allen Institute for Brain Science, funded by Paul Allen and led by Allan Jones, delivered the first atlas of gene expression in a complete mammalian brain, publically available online at www.brain-map.org. Since then, the Allen Institute has expanded its projects to provide online public resources that integrate extensive gene expression, connectivity data, and neuroanatomical information with powerful search and viewing tools for the adult and developing brain in mouse, human, and nonhuman primate (see figure 3 for an example). In addition to the data there are colorimetric and fluorescent ISH image viewers, graphical displays of ISH, microarray and RNA sequencing data, and an interactive reference atlas viewer ("Brain Explorer") that enables 3D navigation of anatomy and
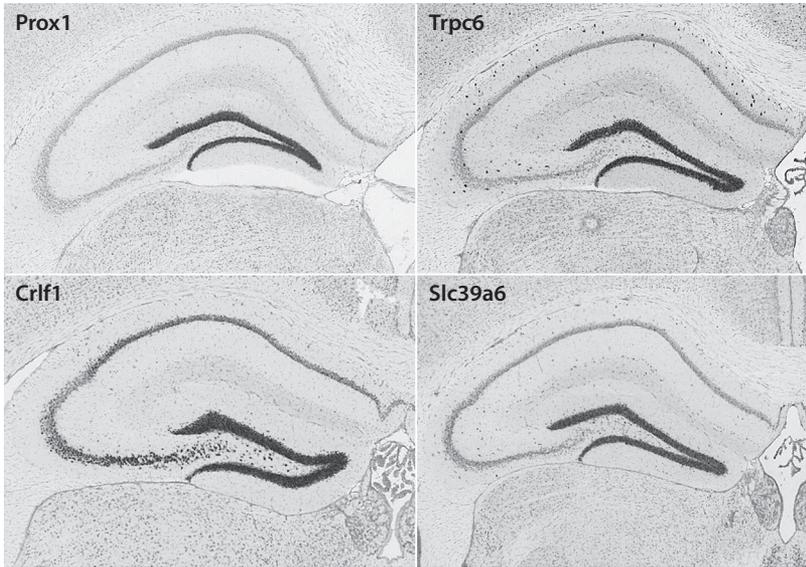
Figure 3. Genes whose expression pattern is highly correlated with *Prox1* (upper left) in dentate gyrus of the hippocampus. These genes were found by starting with the image for gene *Prox1* and searching for patterns whose spatial pattern of gene expression strongly resembled *Prox1*. Combinations of expression patterns such as these may help to refine our present understanding of the function of the hippocampus.

gene expression across these datasets. (Approximately fifty thousand users worldwide access the Allen Brain Atlas resources each month.) Scientists have mined the atlases to search for marker genes in various brain regions associated with diseases, to identify different cell type markers, to delineate brain regions, and to compare gene expression data across species.

Extending this work to humans, the Allen Human Brain Atlas was made public in May 2010 and is the first anatomically comprehensive and genome-wide, three-dimensional map of the human brain. This transcriptional atlas of six adult human brains contains extensive histological analysis and comprehensive microarray profiling of several hundred precise brain subdivisions and has revealed that gene expression varies enormously by anatomical location, with different regions and their constituent cell types displaying robust molecular signatures that are highly conserved between individuals.

In particular, these data show that 84 percent of all genes are expressed somewhere in the human brain and in patterns that while complex are substantially similar from one brain to the next. The analysis of differential gene expression and gene coexpression relationships demonstrates that brain-wide variation strongly reflects the distributions of the major cell types such as neurons, oligodendrocytes, astrocytes, and microglia, all of which are essential to brain function. Interestingly, the neocortex displays a relatively homogeneous transcriptional pattern but with distinct features associated selectively with primary sensorimotor cortices and with enriched frontal lobe expression. Interestingly, the spatial topography of the neocortex is strongly reflected in its molecular topography, that is, the closer two cortical regions are, the more similar their gene expression patterns remain.

Several other significant efforts toward understanding the genetic basis of brain organization are underway, including the Edinburgh Mouse Atlas Project (EMAP) (www.emouseatlas.org), which contains substantial spatial and temporal data for mouse embryonic development, and the Rockefeller University–based GENSAT project of Nathanial Heintz and colleagues that seeks to characterize gene expression patterns using Bacterial Artificial Chromosomes (BAC) in genetically modified mice (www.gensat.org), as well as BGEM (www.stjudebgem .org), GenePaint (www.genepaint.org), EurExpress (www.eurexpress .org), and MGI (http://www.informatics.jax.org), all generally user friendly with useful tutorials.

A Standard Brain?

Does a standard or normal brain exist? This is less likely for humans than genetically bred mice, but mapping neuroscientific and clinical data onto a common frame of reference allows scientists and physicians to compare results between individuals. One main reason for standardization is that multiple and diverse brains can be transformed into a standard framework that maximizes our ability to understand their similar features. Another is that it allows us to identify how unique or unusual features in a particular brain may differ from an average population. With modern advanced image processing capabilities, digital

atlases can serve as the framework for building standard atlases and for surveying the information linked to it. In contrast to basic data repositories, which allow for simple access to data through a single interface, sophisticated digital atlases backed by appropriate technology can act as hubs facilitating access to multiple databases, information sources, and related documents and annotations. These may act as a scaffold from which to share, visualize, analyze, and mine data of multiple modalities, scales, and dimensions.

Many of these ideas of standardization grew out of a major initiative of the National Institutes of Health in the 1990s called the "Decade of the Brain," where a number of digital and electronic resources were created to enable the unification and integration of the various subfields of neuroscience. One outcome of this work is the field of *neuroinformatics*, or the application of computer- and mathematical-based technologies to organize and understand brain data. The ultimate goal of neuroinformatics is to bring together brain architecture, gene expression, and 2D and 3D imaging information into common frames of reference. Major organizations have evolved around mapping brain data, such as the International Consortium for Human Brain Mapping (www.loni.ucla.edu/ICBM/About) and the International Neuroinformatics Coordinating Facility (INCF, www.incf.org). These efforts have led to atlases such as the standard Talairach Atlas and the Montreal Neurological Institute (MNI) standard that have been extensively used in neuroscience.

One consideration in standardizing brain atlases is the type of coordinate system used. As Alan Evans of the Montreal Neurological Institute remarks, "The core concept within the field of brain mapping is the use of a standardized or 'stereotaxic' 3D coordinate framework for data analysis and reporting of findings from neuroimaging experiments. This simple construct allows brain researchers to combine data from many subjects such that group-averaged signals, be they structural or functional, can be detected above the background noise." The concept of a coordinate system is fundamental to digital atlases and requires two basic components: the specification of an origin in the stereotaxic space and a mapping function that transforms each 3D brain from its native coordinates to that of the atlas. A major step in addressing these issues, and a standard tool set that allows different types of neuroscience data to be combined and compared, is now in development for one of

the most important subjects in experimental neuroscience: the mouse, *Mus musculus*. This project is an international collaboration in digital atlasing and is sponsored in part by the International Neuroinformatics Coordinating Facility (INCF).

The Connected Brain

Much recent evidence indicates that the vast interconnected network of the human brain is responsible for our advanced cognitive capabilities, rather than a simple expansion of specialized regions of the brain such as the prefrontal cortex. This may apply in particular to diseases associated with potentially aberrant wiring such as schizophrenia, autism, and dyslexia. The importance of *circuit* considerations for differentially characterizing disorders such as major depression, anxiety, and obsessive-compulsive disorder, and substance abuses including nicotine addiction, is now being widely recognized.

It is now understood that neuropsychiatric disorders likely result from pathologies at the system level, with both complex genetic and environmental factors impacting neural circuitry. As Jason Bohland and colleagues point out in a recent 2009 proposal for a "mesoscale," that is, medium scale, connectome in *PLOS Computational Biology*: "For those [diseases] with heritable susceptibility effects, genetic polymorphism and cellular processes play a greater role, but anatomical circuits remain critical to understanding symptoms and developing therapies." In Parkinson's disease, for example, drug- and stimulation-based therapeutic interventions do not occur at a particular cellular lesion site, but rather are contingent on understanding interactions within the extrapyramidal motor system of neurons.

The first unified approach to defining connectional atlases of the brain was proposed by Olaf Sporns (Indiana University) and Patrick Hagmann (Lausanne). In 2005 they independently suggested the term "connectome" to refer to a complete map of the neural connections within the brain. This term was directly inspired by the concurrently ongoing effort to sequence the human genetic code, and since then the field of *connectomics* (see chapters by Sporns and Zador, this volume) has been concerned with assembling and analyzing connectome datasets. (The

term *connectome* was most recently popularized in Sebastian Seung's book *Connectome*, which discusses the high-level goals of mapping the human connectome, as well as ongoing efforts to build a 3D neural map of brain tissue at the microscale.)

The first complete neural circuit in any organism was found in the common worm *Caenorhabditis elegans*, and research into its molecular and developmental biology was begun in 1974 by Nobel laureate Sydney Brenner. *C. Elegans* has since been used extensively as a model organism in biology. Using high-resolution electron microscopy and manual annotation of hundreds of images, the circuit-mapping project was a major tour de force of neuroanatomy, resulting in a 341-page publication by the Royal Society in 1986 by John White and Brenner titled "The Structure of the Nervous System of the Nematode *Caenorhabditis elegans*." Other landmark studies include a study of the areas and connections of the visual cortex of the macaque published by Daniel Felleman and David Van Essen in 1991 and of the thalamo-cortical system in the feline brain by J. W. Scannel and colleagues in 1999. Since then several neuroinformatics databases of connectivity have emerged, such as the online macaque cortex connectivity tool CoCoMac (www.cocomac.org) and the Brain Architecture Management System (BAMS, http://brancusi .usc.edu).

Several years ago, supported both by public and private funding, a series of independent projects were launched to map the connectome of the laboratory mouse at the mesoscale. Among these projects the Allen Institute embarked on a large-scale effort to develop a regional and cell type specific three-dimensional connectivity map. This Allen Mouse Brain Connectivity Atlas uses a combination of normal and genetically modified mice together with genetic tracing approaches and a high-throughput serial 2-photon tomography system to image the labeled axons throughout the entire brain. High-resolution coronal images are sampled every 100 μm (0.1 mm), resulting in a large 750-GB dataset per brain. At the end of 2013, approximately 1,500 terabytes of data (or 1.5 petabytes) will have been generated, all mapped onto a common 3D reference space of high spatial fidelity that allows for identification of the neural circuitry that governs behavior and brain function.

Mapping the connectome of the human brain is one of the great scientific challenges of the twenty-first century. The Human Connectome

Project (HCP, http://www.humanconnectome.org) is tackling a key aspect of this challenge by elucidating some of the main neural pathways that underlie brain function and behavior. Due to the immense complexity and comparatively large size of the human brain, the HCP (see chapter by Sporns, this volume) is taking a more macro approach to mapping large-scale circuitry, comprehensively mapping human brain circuitry in a target number of 1,200 healthy adults using a combination of noninvasive neuroimaging techniques such as MRI, EEG, and fMRI.

Accurate *parcellation* of fMRI imaging activity into component areas of the brain is an important consideration in deciphering its connectivity, and it takes us back to our original discussion of anatomy. Modern imaging techniques have enabled parcellation of localized areas of cortex and have been accomplished by using diffusion tractography and functional imaging to measure connectivity patterns and define cortical areas based on these different connectivity patterns. Such analyses may best be done on a whole brain scale and by integrating types of noninvasive imaging. It is hoped that more accurate whole brain parcellation may lead to more accurate macroscale connectomes for the normal brain, which can then be compared to disease states. The HCP images and their parcellations are being made available to the public through a public interface called the ConnectomeDB at http://www .humanconnectome.org, mentioned above.

The current noninvasive imaging techniques cannot capture the brain's activity at a neuronal level, and mapping the connectome at a cellular level in vertebrates currently requires postmortem microscopic analysis of limited portions of brain tissue. The challenge of doing this on a grand scale is quite major, as the number of neurons comprising the brain easily ranges into the billions in more highly evolved organisms, with the human cerebral cortex alone contains at least 10^{10} neurons and linked by 10^{14} synaptic connections. A few of the main challenges of building a microscale mammalian connectome today include: the data collection would take years given current technology; annotation tools are insufficient to fully delineate and extract information at a neuronal scale; and, not least, the algorithms necessary to map relevant connections and build the connectivity graphs are not yet fully developed. To address these machine-vision and image-processing issues, the Open

Connectome Project (openconnectome.org) is a crowd-sourcing initiative to meet this challenge. Finally, statistical graph theory is an emerging discipline that is developing sophisticated pattern recognition and inference tools to parse these brain graphs.

The Future Brain

Development of large-scale brain atlases is now a major undertaking in neuroscience. While it may not be possible to systematically map each of the one hundred billion neurons any time soon in any given individual brain, modern mapping techniques are providing atlases of remarkable resolution and functionality.

Several recent advances in neuroimaging support the possibility of deep and large-scale mapping, and this goal may be less audacious than seems at first. For example, using a combinatorial color labeling method, Brainbow, which is based on the random expression of several types of fluorescent proteins, Josh Sanes and Jeff Lichtman at Harvard are able to mark individual neurons with one of over one hundred distinct colors. The labeling of individual neurons with a distinguishable hue then allows the tracing and reconstruction of their cellular structure, including long processes within a block of tissue. Labeling techniques such as these allow for classification and visualization of microscopic neurons. Another approach aimed at classifying diversity in the synaptic code, called array tomography, has been developed by Stephen Smith at Stanford, and can also achieve combinatorial labeling of synaptic connections using electron microscopy.

Recently, in a processing tour de force, nearly 7,500 sections of an individual human brain were sliced and scanned and mapped onto a 3D reconstructed brain at 20-micron isotropic resolution, that is, in all three spatial dimensions. This project is the culmination of years of work from the Katrin Amunts and Karl Zilles laboratory at Jülich, Germany, with semiautomated informatics reconstruction by Evans in Montreal. The atlas called BigBrain is a thin-sliced histology project that offers nearly cellular resolution, that is, detail close to the level of the cell. Because of the nearly continuously collected sections and the 3D

image reconstruction, BigBrain is a dataset 125,000 times bigger than a typical MRI! Atlases based on MRIs do not allow for the visualization of information at the level of cortical cells and layers, although this atlas will allow that. However, to make BigBrain into a full-fledged atlas it will need to be annotated, that is, it will need to provide the anatomic structural delineations that outline the fine structure of the brain.

The BigBrain effort indicates that high-resolution 3D microscopy is still not at a level of resolution to map the finest structures in the brain. However, advances are being made in 3D imaging as well. In 2013, in a highly publicized article in *Nature*, a method was developed to subject the brain to a three-dimensional network of hydrophilic polymers and then to remove the lipids from the brain by electrophoresis. The brain remains fully intact but optically transparent and macromolecule permeable. Using mouse brains, the authors show intact-tissue imaging of long-range projections, local circuit wiring, cellular relationships, and subcellular structures. This method, called "CLARITY," uses intact-tissue in situ hybridization and immunohistochemistry with multiple rounds of staining and de-staining in nonsectioned tissue to visualize gene expression or protein binding. The method is still being refined but may be useful for human postmortem imaging as well.

Alternative computational processing techniques will also be necessary to deal with the massive data these new atlases generate. In 2012, a Citizen Science project called EyeWire, launched by Sebastian Seung of MIT, began attempting to crowd source the mapping of the connectome through an interactive game in which contributors try to map the retinal connectome (Zador's chapter herein outlines another possible approach to this problem).

Large-scale atlases of the brain are providing content to the neuroscience community through molecular, cellular, functional, and connectomic data. The transition from print to digital atlases has been revolutionary, as it has allowed navigation, 3D reconstruction, and visualization from the smallest nuclei to macroscale regions. Digital atlases have also transformed clinical neuroscience, and all stages from pre- to postoperation of surgery in some way use digital atlases. It is likely that in the near future we will have annotated 3D microscale atlases of the structure of the human brain. In several years it should be

possible to achieve near complete axon and synaptic connectivity in a substantial segment of the human cortex, thereby elucidating the detailed complexity of its cortical circuitry. Atlases will continue to be more integrated into scientific and clinical workflows, thus aiding in discovery science and providing novel ways of diagnosing, monitoring, and treating disease.

INDEX