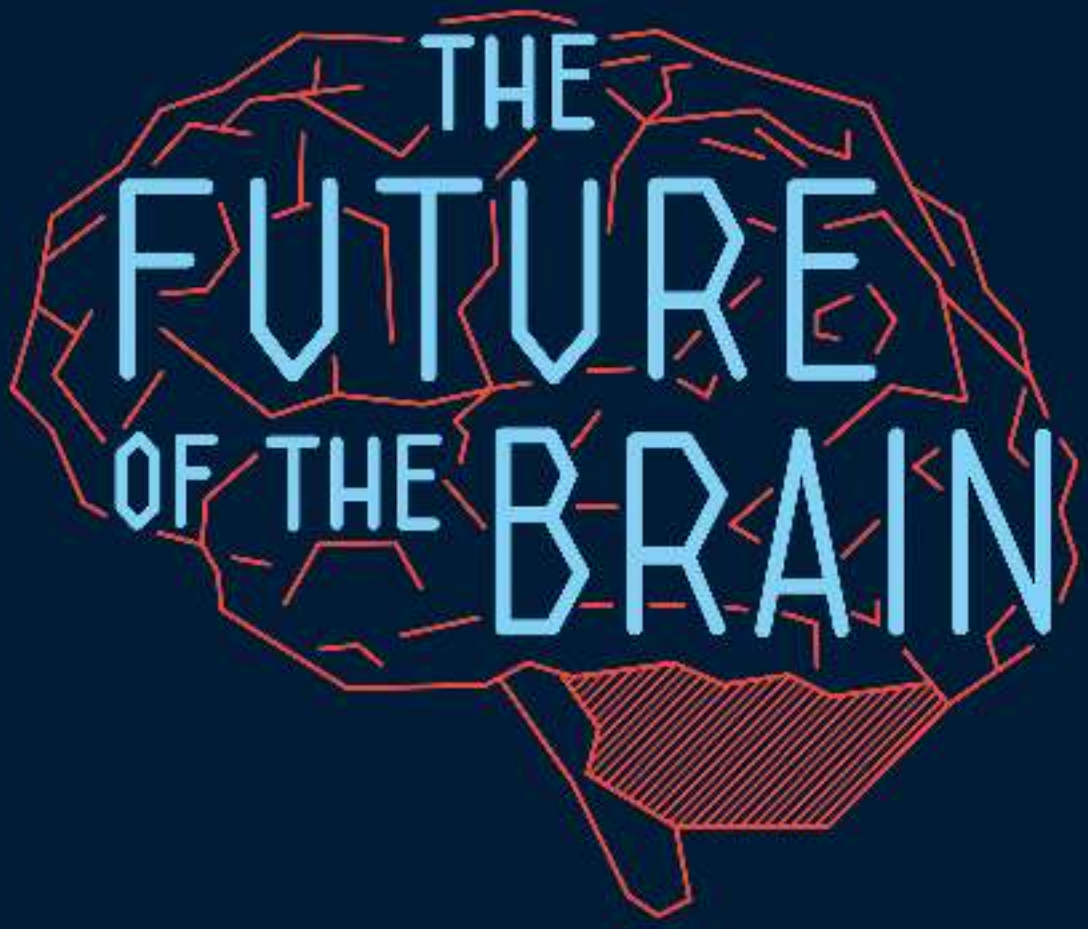
Including a chapter by 2014 Nobel laureates May-Britt and Edvard Moser

ESSAYS BY THE WORLD'S LEADING NEUROSCIENTISTS

THE FUTURE OF THE BRAIN



EDITED BY GARY MARCUS AND JEREMY FREEMAN

THE FUTURE OF THE BRAIN

ESSAYS BY THE WORLD'S LEADING NEUROSCIENTISTS

EDITED BY
GARY MARCUS AND JEREMY FREEMAN

PRINCETON UNIVERSITY PRESS
PRINCETON AND OXFORD

Published by Princeton University Press, 41 William Street, Princeton, New Jersey 08540
In the United Kingdom: Princeton University Press, 6 Oxford Street, Woodstock, Oxfordshire OX20 1TW

press.princeton.edu

Jacket design by Karl Spurzem

ISBN 978-0-691-16276-8

Library of Congress Control Number: 2014938489

British Library Cataloging-in-Publication Data is available

This book has been composed in Minion Pro

Printed on acid-free paper. ∞

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

CONTENTS

SIMULATING THE BRAIN

LANGUAGE

SKEPTICS

IMPLICATIONS

AFTERWORD

CONTRIBUTORS

Misha B. Ahrens
HHMI, Janelia Farm Research Campus, Ashburn, Virginia

Ned Block
New York University

Arthur Caplan
New York University, Langone Medical Center

Matteo Carandini
University College London

George Church
Harvard Medical School

John Donoghue
Brown Institute for Brain Science, Brown University

Chris Eliasmith
University of Waterloo, Ontario, Canada

Simon E. Fisher
Max Planck Institute for Psycholinguistics and Donders Institute for Brain, Cognition & Behaviour, Radboud University, Nijmegen, the Netherlands

Jeremy Freeman
HHMI, Janelia Farm Research Campus, Ashburn, Virginia

Mike Hawrylycz
Allen Institute for Brain Science, Seattle, Washington

Sean Hill
Brain Mind Institute, École Polytechnique Fédérale de Lausanne, Switzerland

Christof Koch
Allen Institute for Brain Science, Seattle, Washington

Leah Krubitzer
University of California, Davis

Michel M. Maharbiz
University of California, Berkeley

Gary Marcus
New York University and Allen Institute for Brain Science, Seattle, Washington

Kevin J. Mitchell
Smurfit Institute of Genetics and Institute of Neuroscience, Trinity College, Dublin

May-Britt Moser and Edvard I. Moser
Centre for Neural Computation, Kavli Institute for Systems Neuroscience, Norwegian University of Science and Technology, Trondheim, Norway

David Poeppel
New York University

Krishna V. Shenoy
Departments of Electrical Engineering, Bioengineering, and Neurobiology, Stanford University

Olaf Sporns
Indiana University, Bloomington

Anthony Zador
Cold Spring Harbor Laboratory, New York

PREFACE

There's never been a more exciting moment in neuroscience than now. Although the field has existed for two centuries, going back to the days of Phineas Gage and the tamping iron that exploded through his left frontal lobe, progress has in many ways been slow. At present, neuroscience is a collection of facts, still awaiting an overarching theory; if there has been plenty of progress, there is even more that we don't know. But a confluence of new technologies, many described in this book, may soon change that.

To be sure, there is long history of advances, even from the earliest days, often leveraging remarkably crude tools to great effect. In the mid-1800s Paul Broca got the first glimpse into the underpinnings of language by doing autopsies on people who had lost linguistic function because of brain damage to specific cortical areas. Near the end of the nineteenth century, Camillo Golgi discovered that he could visualize neurons under a microscope by staining them with silver nitrate, and Santiago Ramón y Cajal used the technique to develop remarkably prescient characterizations of neuronal structure and function. In 1909 a brilliant ophthalmologist named Tatsuji Inouye launched functional brain mapping, by methodically studying victims of gunshot wounds during the Russo-Japanese war, noting that wounds to the visual cortex impaired his patients' vision, and wounds to particular locations affected vision in *particular* regions of the visual field.

In the latter part of the twentieth century, noninvasive forms of brain imaging, like functional magnetic resonance imaging (fMRI), came on the scene. But as useful as such tools are, current noninvasive techniques are like fuzzy microscopes; they blur the fine detail of neural activity in both space and time. Ultimately, looking at an fMRI scan is like looking at a tiny pixelated version of a detailed, high-resolution photograph.

In nonhuman animals, which can be studied with more invasive techniques, the gold standard until recently was the "single neuron recording," which uses thin electrodes to monitor the electrical activity associated with neural firing. Action potentials are the currency of the

brain, and directly measuring them has led to many fundamental insights, such as Hubel and Weisel's discovery that neurons in the visual cortex are "tuned" or selective for particular visual features. But looking at one neuron at a time tells an incomplete story at best; the neuroscientist Rafael Yuste has likened it to "understanding a television program by looking at a single pixel."

As we write this, it is clear that neuroscience is undergoing a revolution. Optogenetics, introduced in 2005, makes it possible to engineer neurons that literally light up when active, switching them on and off with a laser; multielectrode recordings, which allow recordings from hundreds or even thousands of neurons are finally becoming practical, and new forms of microscopy can record the activity of nearly every neuron in a living, transparent fish. For the first time, it is realistic to think that we might observe the brain at the level of its elementary parts.

• • • • • •

Still, three fundamental truths make the brain more challenging to understand than any other biological system.

First is sheer numbers. Even in the fly or the larval zebrafish brain there are one hundred thousand neurons. In the human brain there are over 85 billion. On top of that, the word *neuron* makes it sound like there is only one kind, whereas in fact there are several hundred kinds, possibly more, each with distinctive physical characteristics, electrical characteristics, and, likely, computational functions. Second, we have yet to discover many of the organizing principles that govern all that complexity. We don't know, for example, if the brain uses anything as systematic as, say, the widespread ASCII encoding scheme that computers use for encoding words. And we are still shaky on fundamentals like how the brain stores memories and sequences events over time. Third, many of the behaviors that seem characteristically human—like language, reasoning, and the acquisition of complex culture—don't have straightforward animal models.

The Obama BRAIN Initiative, the European Human Brain Project, and other large-scale programs that may begin in Asia aim to address some of the challenges in understanding the brain. It seems reasonable that we can expect, over the next decade, an enormous amount of new data at an unprecedented level of detail, certainly in animals, and

perhaps in humans as well. But these new data will raise new questions of their own. How can researchers possibly make sense of the expected onslaught of data? How will we be able to derive general principles?

And for that matter, will collecting all these data be enough? How can we scale up data analysis to the terabytes to come, and how can we build a bridge from data to genuine insight? We suggest that one key focus must be on computation. The brain is not a laptop, but presumably it is an information processor of some kind, taking in inputs from the world and transforming them into models of the world and instructions to the motor systems that control our bodies and our voices. Although many neuroscientists might take for granted that the principal process by which the brain does its work is some form of computation, almost all agree that the most foundational properties of neural computation have yet to be discovered. Our hope is that computation can provide a universal language for describing the action of the brain, especially as theorists and experimentalists come closer together in their quest.

Given the complexity of the brain, there is no certainty we will come to fully or even largely understand the brain's dynamics anytime soon; in truth, there is reason for hope, but no guarantees. This book, with chapters by pioneers like Christof Koch and George Church, represents our best guesses—and our esteemed contributors' best guesses—about where we are going, what we are likely to find out, and how we might get there.

But it also admits where we might stumble along the way. If this book is a reader's guide to the future, it's not a foolproof crystal ball; if anything, it's more like a time capsule. Part of the fun will be for scientists, policy makers, and the public to come back to these essays a decade hence, to, as one colleague put it, "reassess its scientific claims, aspirations, and methodological promises, and adjust the aspirations of the next generation of neuroscientific endeavors accordingly." We couldn't agree more.

Gary Marcus and Jeremy Freeman
March 2014

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

a **b**

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

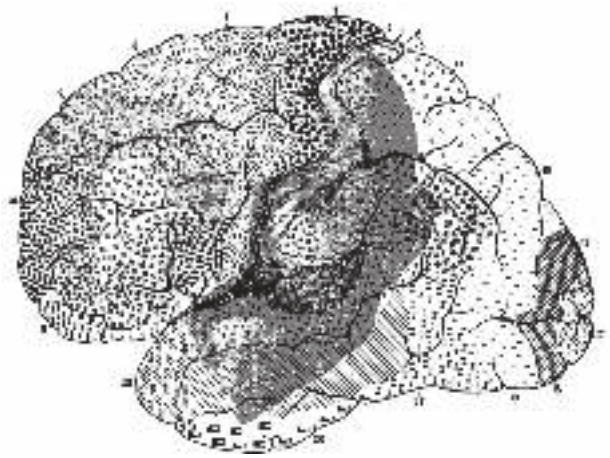


Figure 2. Regions of the human cerebral cortex delineated by Korbinian Brodmann using Nissl stain histology. Brodmann identified forty-three distinct regions that today still serve as a guide for studying distinct functional areas in the human cortex.

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

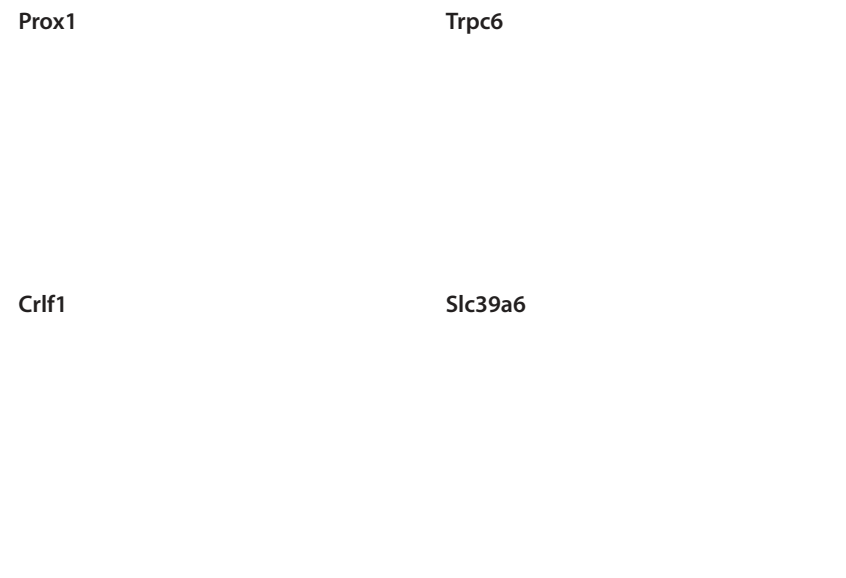


Figure 3. Genes whose expression pattern is highly correlated with *Prox1* (upper left) in dentate gyrus of the hippocampus. These genes were found by starting with the image for gene *Prox1* and searching for patterns whose spatial pattern of gene expression strongly resembled *Prox1*. Combinations 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