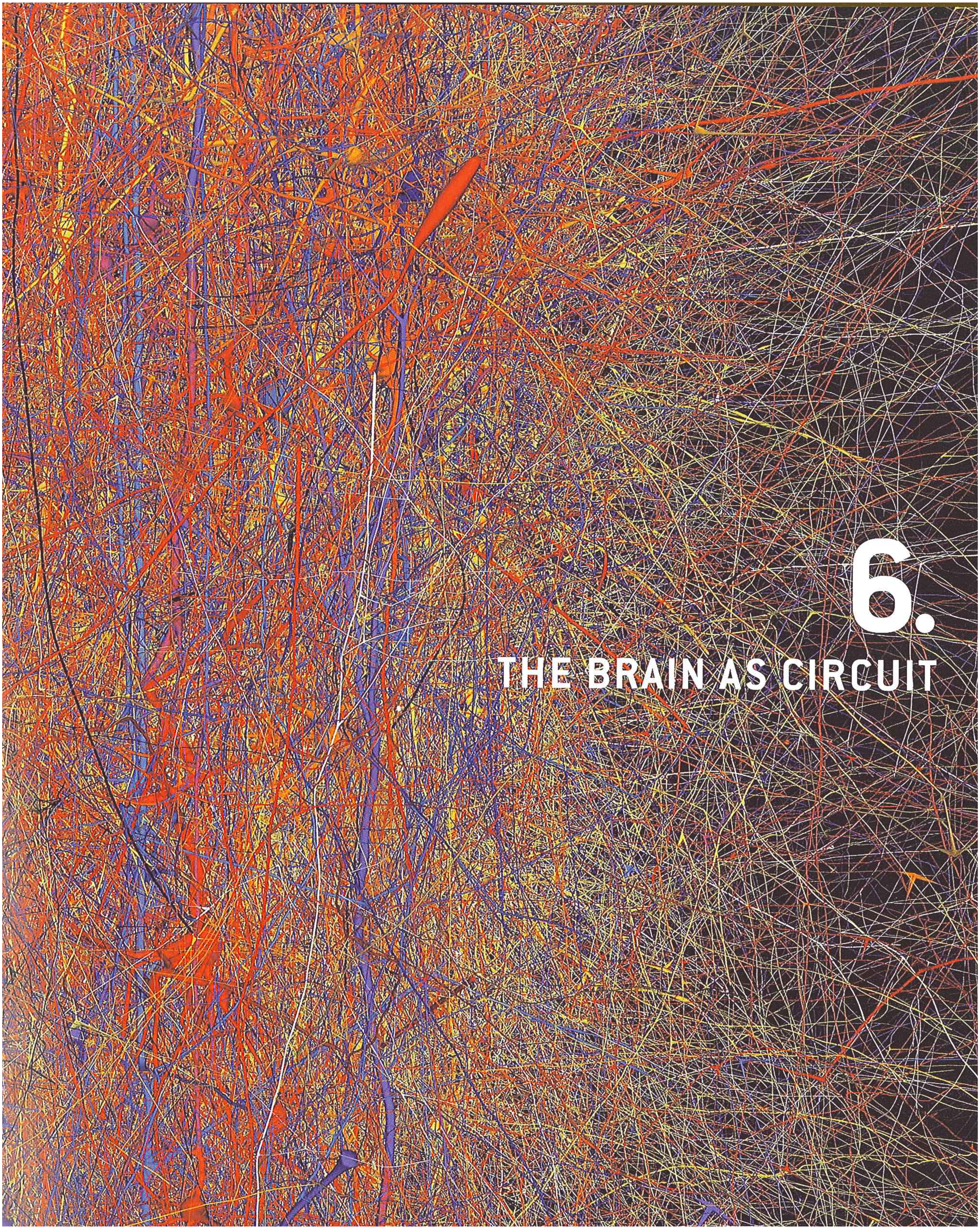


6. THE BRAIN AS CIRCUIT

THE BRAIN AS CIRCUIT

by Terrence Sejnowski
Salk Institute for Biological Studies

Although Santiago Ramón y Cajal is best known for his magnificent drawings of neurons from the brains of many species, it is his insight into brain function that may be his greatest scientific achievement. A close look at the images on pages 67–71 reveals a sequence of small arrows that were obviously not present in the biological samples he examined. He added these arrows to his diagrams to indicate the direction in which he had deduced neural signals progressed within each neuron—from dendrite to soma to axon—and between neurons along their long axonal projections. In the process, Cajal established the overarching conceptual framework under which we approach the brain—the circuit. His little arrows were the germ of a theory of how information flows in the nervous system, a concept we are still trying to grapple with one hundred years after his seminal work.

While Cajal was developing his ideas about neurons and neural circuits, others were investigating brain function by examining the cognitive deficits associated with injuries to various parts of the brain. For example, focal injury to a region in the brain known as the motor cortex (an area involved in voluntary movement) produces weakness or even paralysis in muscles whose control depends on brain tissue at location of the injury. Patients with injuries to one area of this strip of cerebral cortex might have paralysis in one arm, while those with injuries to another area might have paralysis in a leg. These kinds of observations helped convince scientists of the hypothesis that mental faculties are localized—that is, that certain areas of the brain are responsible for certain functions. But knowing what goes wrong when some part of the brain is damaged is not tantamount to knowing that circuit's function; pinning a broad label like "vision" or "motor" on a brain region is a poor substitute for understanding how it works.

When the late Francis Crick shifted his research program from molecular genetics to neuroscience in the 1970s he channeled his efforts on vision, studying how the connections between neurons in the visual system are organized. Neuroanatomy served as a conceptual framework for his thinking about higher-level processes such as visual awareness and consciousness. The dictum that had served him well in deciphering the function of DNA from its double-helical structure ("from structure follows function"), also offered powerful clues about the function of neural circuits. If a circuit receives input from a visual area, for instance, it is very likely that it in turn exploits or further processes visual information. Crick was particularly intrigued by recurrent loops between brain areas, as he thought that they might have something to do with our ability to pay attention to objects in visual scenes. He conjectured that the feedback connections between neurons in the visual cortex and neurons in parts of the thalamus—a brain structure that sends inputs to the visual cortex—might regulate the incoming information and underlie conscious experience.

But in order to fully account for the anatomical underpinnings of brain connectivity, we need more than just the broad picture of how one major area communicates with another. The Golgi stain reveals the shapes and sizes of neurons in every part of the brain and, to a limited extent, can help researchers deduce the patterns of connectivity between them, but in order to fully understand brain function we will require a complete wiring diagram of both the connections between all the neurons in a single brain area—the local circuit—and the long-range connections between different brain areas. Just as genomics made it possible to compile a complete list of genes and compare them between species, the new field of connectomics seeks to compile complete circuit diagrams of various nervous systems, including the human brain. The wiring diagram of *C. elegans* with its 302 neurons (see page 186) was recently constructed with painstaking human labor, but it will soon become possible to piece together diagrams of much larger systems using powerful computational techniques that automate the process. How will these new maps of connectivity change the way we think about brain function?

A circuit diagram by itself does not reveal a circuit's function, as a diagram is inherently static and brain function is determined by dynamic processes. Today's equivalent of Cajal's little arrows are computer simulations of brain models that take into account the anatomical details and the biophysical realities of neurons and the synapses that connect them. The goal is to follow information as it enters the circuit, such as patterns of light hitting the retina in the eye, and to investigate the resulting signals as they propagate down a chain of neurons, circulating through recurrent loops inside the brain. This is a daunting task since many details about neurons remain unknown, but also because even the fastest supercomputers cannot keep track of all the billions of brain cells that make up a human brain. The system is simply too complex for our computers to fully simulate with today's technology.

When I first began using computers to model neurons and neural circuits in 1980, I used computers that could perform around a million operations per second (at the time this seemed enormously powerful, but today I would be better off using a smartphone). With this technology I was able to simulate a few hundred simplified neurons, each connected to many others. Even using just a small number of these virtual "neurons" I was able to demonstrate how artificial "neural networks" could be configured to perform some amazingly complex tasks, such as pronouncing English words. Rather than hand-wire the network using rules for English pronunciation, which have many exceptions, I created a network trained on examples. Every time the network made a pronunciation mistake, the strengths of "synapses" between "neurons" were changed by a small amount so that the next time the same word or similar words occurred, the pronunciation would improve. This was a slow process. At first the network babbled like a baby as it learned to distinguish consonants from vowels. But eventually it began to sound out simple words and finally even complex words with many syllables. This was astonishing to me and others, who up until then had believed that English pronunciation was beyond the reach of a simple neural-network model. The experiment taught us that problems that can be difficult to program a computer to solve can be quite simple for a network to learn from experience.

Today's computers are thousands of times faster than the ones I used in the 1980s, and we can now simulate neurons in much greater detail, including the fine branching of dendrites and the many synapses on dendritic branches. Theoretical neuroscientists have employed computers to simulate the geometry of the neurons that Cajal first observed, revealing that a single neuron is far more computationally powerful than previously thought; even though the brain derives many of its remarkable abilities from the connections between neurons, some of its prowess is due to the intrinsic dynamical properties of the neurons themselves.

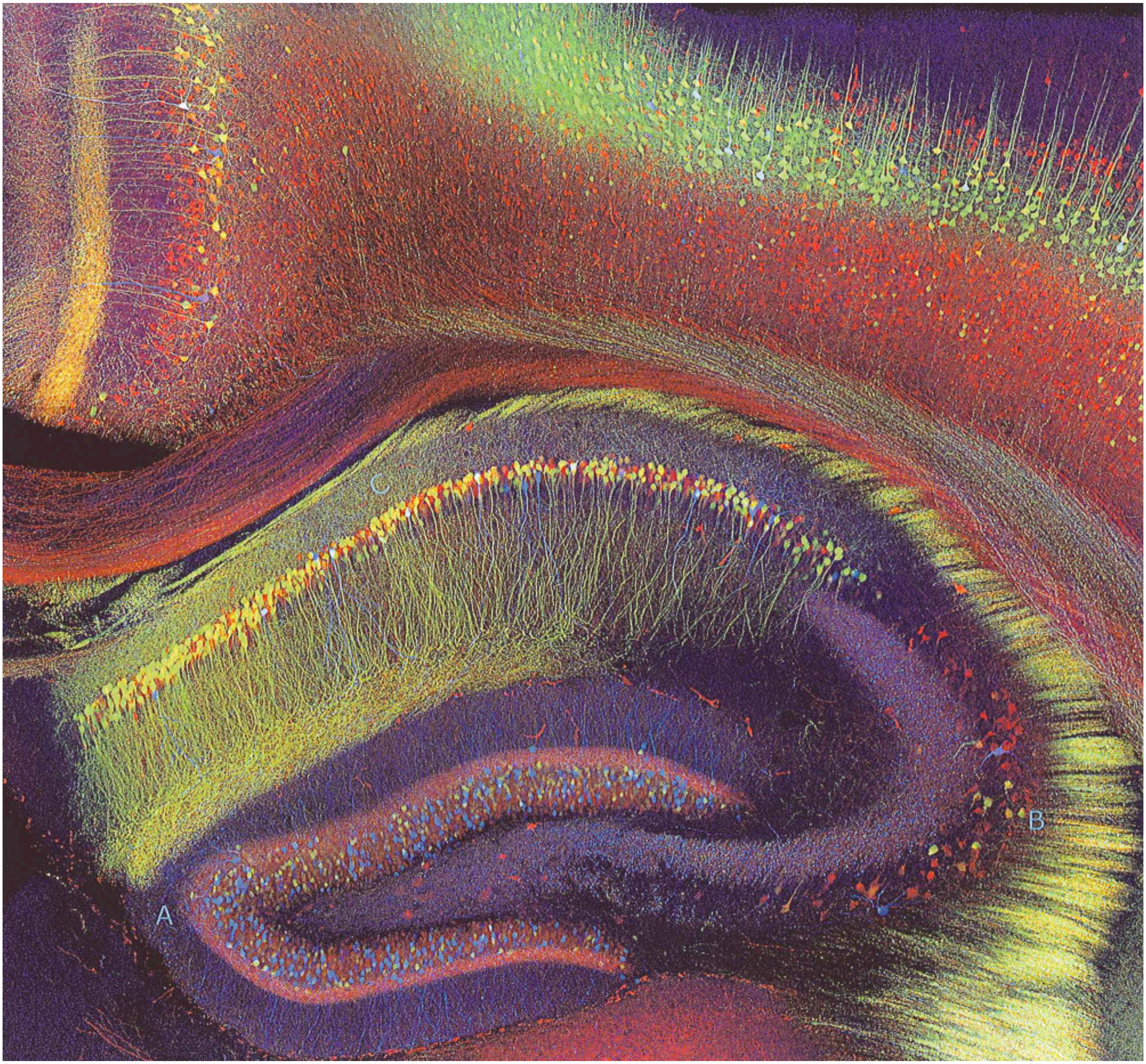
If the study of neural circuits weren't sufficiently complicated, it is now known that circuits are dynamic on many timescales. Every protein in every cell is replaced over a matter of hours or days. Synapses between neurons are plastic and can change their sizes and strengths in response to changes in the patterns of activity within circuits, and new synapses are formed on a daily basis. Massive synapse remodeling occurs during brain development and continues at a reduced rate in adults as a consequence of learning through experience and, sometimes, recovery after brain damage. The loss of neurons due to age or disease leads to a deterioration in memory and agility, another structural shift within the brain that affects neural circuits. In order to fully address the challenge posed by this constant flux, researchers must map many circuits at different stages of development and in many different environments.

Yet, even if we could reproduce all the anatomical details and signals in a brain, this wealth of knowledge would not in itself explain how a brain functions, or goes awry. What we need is a twenty-first-century Cajal who can understand the function of these circuit diagrams by simulating the signals as they are processed by the circuits themselves. For example, while a person is looking at a particular object, the neurons that represent that object in the brain change their firing patterns, sending out signals to other neurons at faster or slower rates. This is a subtle change that does not involve a shift in the wiring diagram of the brain; we will need to quantify these firing patterns using theoretical tools like information theory from engineering and dynamical systems theory from physics. But even this may not be enough to fully understand brain function. There are many complex systems in the world that have defied our best efforts to understand them both through mathematical analysis and simulation. The weather, to pick a familiar example, remains famously unpredictable despite decades of research, sophisticated mathematical models, and ever more powerful computers.

✳ ✳ ✳

It may be that in order to fully understand brain function we must first understand how the brain develops from an embryo, and how the molecular mechanisms inside cells interact with the information flowing through brain circuits. In particular, we need to resolve how structures within the brain at many different spatial scales interact with each other over a wide range of timescales, a challenge termed "the levels problem." For example, the relatively slow molecular machinery inside cortical synapses, which regulates the size and strength of a synapse, is itself controlled by the relative timing, on a very short timescale, of the action potentials in presynaptic neurons (those that are sending messages) and postsynaptic neurons (those that are receiving them). The patterns of action potentials in a population of neurons can influence the biochemical pathways inside synapses that control the way neurons communicate with each other, which, in turn, influence the subsequent pattern of action potentials. It is with this elegant interplay between scales that the brain is able to solve the countless variety of problems that it is designed to tackle, as well as challenges that nature could not anticipate. And it has proven a source of constant frustration to students of the brain who are attempting to untangle its function.

We are witnessing a period of unprecedented innovation in the techniques for studying the brain. The hope is that just as Golgi's method enabled powerful insight into the structure of the brain, today's dazzling new tools will open up a universe hitherto inaccessible to us. And indeed, they have begun to bear fruit: we can now "watch" the electrical activity of many neurons simultaneously using optical recordings; the advances (described at the end of Chapter 5, pages 164–165) even allow us to manipulate the activities of individual cells and cell types using beams of light. When new techniques are introduced, it is more or less impossible to predict what will come out of them, or whether they will even alter a field's landscape. Golgi's technique lay fallow for more than a decade before Cajal perfected it and applied it to establish the basic facts about neurons. We can only hope that another Cajal will soon emerge to fully exploit this embarrassment of riches and uncover the secrets of our marvelous neural circuits.


C
B
A