The Brain and Behavior

Eric R. Kandel

The last frontier of the biological sciences—their ultimate challenge—is to understand the biological basis of consciousness and the mental processes by which we perceive, act, learn, and remember. In the last two decades a remarkable unity has emerged within biology. The ability to sequence genes and infer the amino acid sequences for the proteins they encode has revealed unanticipated similarities between proteins in the nervous system and those encountered elsewhere in the body. As a result, it has become possible to establish a general plan for the function of cells, a plan that provides a common conceptual framework for all of cell biology, including cellular neurobiology. The next and even more challenging step in this unifying process within biology, which we outline in this book, will be the unification of the study of behavior—the science of the mind—and neural science, the science of the brain. This last step will allow us to achieve a unified scientific approach to the study of behavior.

Such a comprehensive approach depends on the view that all behavior is the result of brain function. What we commonly call the mind is a set of operations carried out by the brain. The actions of the brain underlie not only relatively simple motor behaviors such as walking or eating, but all the complex cognitive actions that we believe are quintessentially human, such as thinking, speaking, and creating works of art. As a corollary, all the behavioral disorders that characterize psychiatric illness—disorders of affect (feeling) and cognition (thought)—are disturbances of brain function.

The task of neural science is to explain behavior in terms of the activities of the brain. How does the brain marshal its millions of individual nerve cells to produce behavior, and how are these cells influenced by the environment, which includes the actions of other people? The progress of neural science in explaining human behavior is a major theme of this book.

Like all science, neural science must continually confront certain fundamental questions. Are particular mental processes localized to specific regions of the brain, or does the mind represent a collective and emergent property of the whole brain? If specific mental processes can be localized to discrete brain regions, what is the relationship between the anatomy and physiology of one region and its specific function in perception, thought, or movement? Are such relationships more likely to be revealed by examining the region as a whole or by studying its individual nerve cells? In this chapter we consider to what degree mental functions are located in specific regions of the brain and to what degree such local mental processes can be understood in terms of the properties of specific nerve cells and their interconnections.

To answer these questions, we look at how modern neural science approaches one of the most elaborate cognitive behaviors—language. In doing so we necessarily focus on the cerebral cortex, the part of the brain concerned with the most evolved human behaviors. Here we see how the brain is organized into regions or brain compartments, each made up of large groups of neurons, and how highly complex behaviors can be traced to specific regions of the brain and understood in terms of the functioning of groups of neurons. In the next chapter we consider how these neural circuits function at the cellular level, using a simple reflex behavior to examine the way sensory signals are transformed into motor acts.

Two Opposing Views Have Been Advanced on the Relationship Between Brain and Behavior

Our current views about nerve cells, the brain, and behavior have emerged over the last century from a convergence of five experimental traditions: anatomy, embryology, physiology, pharmacology, and psychology.

Before the invention of the compound microscope in the eighteenth century, nervous tissue was thought to function like a gland—an idea that goes back to the Greek physician Galen, who proposed that nerves convey fluid secreted by the brain and spinal cord to the body's periphery. The microscope revealed the true structure of the cells of nervous tissue. Even so, nervous tissue did not become the subject of a special science until the late 1800s, when the first detailed descriptions of nerve cells were undertaken by Camillo Golgi and Santiago Ramón y Cajal.

Golgi developed a way of staining neurons with silver salts that revealed their entire structure under the microscope. He could see clearly that neurons had cell bodies and two major types of projections or processes: branching dendrites at one end and a long cable-like axon at the other. Using Golgi's technique, Ramón y Cajal was able to stain individual cells, thus showing that nervous tissue is not one continuous web but a network of discrete cells. In the course of this work, Ramón y Cajal developed some of the key concepts of modern neuroanatomy—including the idea that individual cells are the elementary signaling elements of the nervous system.

Additional experimental support for the neuron doctrine was provided in the 1920s by the American embryologist Ross Harrison, who demonstrated that the two major projections of the nerve cell—the dendrites and the axon—grow out from the cell body and that they do so even in tissue culture in which each neuron is isolated from other neurons. Harrison also confirmed Ramón y Cajal's suggestion that the tip of the axon gives rise to an expanded call structure called the growth cone, which leads the developing axon to its target (whether to other nerve cells or to muscles).

Physiological investigation of the nervous system began in the late 1700s when the Italian physician and physicist Luigi Galvani discovered that living excitability and nerve cells produce electricity. Modern electrophysiology grew out of work in the nineteenth century by three German physiologists—Emil DuBois-Reymond, Johannes Müller, and Hermann von Helmholtz—who were able to show that the electrical activity of one nerve cell affects the activity of an adjacent cell in predictable ways.

Pharmacology made its first impact on our understanding of the nervous system and behavior at the end of the nineteenth century, when Claude Bernard in France, Paul Ehrlich in Germany, and John Langley in England demonstrated that drugs do not interact with cells arbitrarily, but rather bind to specific receptors typically located in the membrane on the cell surface. This discovery became the basis of the all-important study of the chemical basis of communication between nerve cells.

The psychological investigation of behavior dates back to the beginnings of Western science, to classical Greek philosophy. Many issues central to the modern investigation of behavior, particularly in the area of perception, were subsequently reformulated in the seventeenth century first by René Descartes and then by John Locke, of whom we shall learn more later. In the midnineteenth century Charles Darwin set the stage for the study of animals as models of human actions and behavior by publishing his observations on the continuity of species in evolution. This new approach gave rise to ethology, the study of animal behavior in the natural environment, and later to experimental psychology, the study of human and animal behavior under controlled conditions.

In fact, by as early as the end of the eighteenth century the first attempts had been made to bring together biological and psychological concepts in the study of behavior. Franz Joseph Gall, a German physician and neuroanatomist, proposed three radical new ideas. First, he advocated that all behavior emanated from the brain. Second, he argued that particular regions of the cerebral cortex controlled specific functions. Gall asserted that the cerebral cortex did not act as a single organ but was divided into at least 35 organs (other ideas were added later), each corresponding to a specific mental faculty. Even the most abstract of human behaviors, such as generosity, secretiveness, and religiosity, were assigned their spot in the brain. Third, Gall proposed that the center for each mental function grew with use, much as a muscle bulks up with exercise. As each center
grew, it purportedly caused the overlying skull to bulge, creating a pattern of bumps and ridges on the skull that indicated which brain regions were most developed (Figure 1-1). Rather than looking within the brain, Gall sought to establish an anatomical basis for describing character traits by correlating the personality of individuals with the bumps on their skulls. His psychology, based on the distribution of bumps on the outside of the head, became known as phrenology.

In the late 1820s Gall’s ideas were subjected to experimental analysis by the French physiologist Pierre Flourens. By systematically removing Gall’s functional centers from the brains of experimental animals, Flourens attempted to isolate the contributions of each “cerebral organ” to behavior. From these experiments he concluded that specific brain regions were not responsible for specific behaviors, but that all brain regions, especially the cerebral hemispheres of the forebrain, participated in every mental operation. Any part of the cerebral hemisphere, he proposed, was able to perform all the functions of the hemisphere. Injury to a specific area of the cerebral hemisphere would therefore affect all higher functions equally.

In 1823 Flourens wrote: “All perceptions, all volitions occupy the same seat in these cerebral organs; the faculty of perceiving, of conceiving, of willing merely constitutes therefore a faculty which is essentially one.” The rapid acceptance of this belief (later called the aggregate-field view of the brain) was based only partly on Flourens’s experimental work. It also represented a cultural reaction against the reductionist view that the human mind has a biological basis, the notion that there was no soul, that all mental processes could be reduced to actions within different regions in the brain!

The aggregate-field view was first seriously challenged in the mid-nineteenth century by the British neurologist J. Hughlings Jackson. In his studies of focal epilepsy, a disease characterized by convulsions that begin in a particular part of the body, Jackson showed that different motor and sensory functions can be traced to different parts of the cerebral cortex. These studies were later refined by the German neurologist Karl Wernicke, the English physiologist Charles Sherrington, and Ramón y Cajal into a view of brain function called cellular connectionism. According to this view, individual neurons are the signaling units of the brain; they are generally arranged in functional groups and connect to one another in a precise fashion. Wernicke’s work in particular showed that different behaviors are produced by different brain regions interconnected by specific neural pathways.

The differences between the aggregate-field theory and cellular-connectionism can best be illustrated by an analysis of how the brain produces language. Before we consider the relevant clinical and anatomical studies concerned with the localization of language, let us briefly look at the overall structure of the brain. (The anatomical organization of the nervous system is described in detail in Chapter 17.)

The Brain Has Distinct Functional Regions

The central nervous system is a bilateral and essentially symmetrical structure with seven main parts: the spinal cord, medulla oblongata, pons, cerebellum, midbrain, diencephalon, and the cerebral hemispheres (Box 1-1 and Figures 1-2A, 1-2B, and 1-3). Radiographic imaging techniques have made it possible to visualize these structures in living subjects. Through a variety of experimental methods, such images of the brain can be made while subjects are engaged in specific tasks, which then can be related to the activities of discrete regions of the brain. As a result, Gall’s original idea that different regions are specialized for different functions is now accepted as one of the cornerstones of modern brain science.

Box 1-1 The Central Nervous System

The central nervous system has seven main parts (Figure 1-2A).

- The spinal cord, the most caudal part of the central nervous system, receives and processes sensory information from the skin, joints, and muscles of the limbs and trunk and controls movement of the limbs and the trunk. It is subdivided into cervical, thoracic, lumbar, and sacral regions. The spinal cord continues rostrally as the brain stem, which consists of the medulla, pons, and midbrain (see below). The brain stem receives sensory information from the skin and muscles of the head and provides the motor control for the muscles of the head. It also conveys information from the spinal cord to the brain and from the brain to the spinal cord, and regulates levels of arousal and awareness, through the reticular formation. The brain stem contains several collections of cell bodies, the cranial nerve nuclei. Some of these nuclei receive information from the skin and muscles of
the head; others control motor output to muscles of the face, neck, and eyes. Still others are specialized for information from the special senses: hearing, balance, and taste.

- The **medulla oblongata**, which lies directly above the spinal cord, includes several centers responsible for vital autonomic functions, such as digestion, breathing, and the control of heart rate.

- The **pons**, which lies above the medulla, conveys information about movement from the cerebral hemisphere to the cerebellum.

- The **cerebellum** lies behind the pons and is connected to the brain stem by several major fiber tracts called peduncles. The cerebellum modulates the force and range of movement and is involved in the learning of motor skills.

**Figure 1-2A** The central nervous system can be divided into seven main parts.

- The **midbrain**, which lies rostral to the pons, controls many sensory and motor functions, including eye movement and the coordination of visual and auditory reflexes.

- The **diencephalon** lies rostral to the midbrain and contains two structures. One, the **thalamus**, processes most of the information reaching the cerebral cortex from the rest of the central nervous system. The other, the **hypothalamus**, regulates autonomic, endocrine, and visceral function.

- The **cerebral hemispheres** consist of a heavily wrinkled outer layer—the cerebral cortex—and three deep-lying structures: the basal ganglia, the hippocampus, and the amygdaloid nuclei. The basal ganglia participate in regulating motor performance; the hippocampus is involved with aspects of memory storage; and the amygdaloid nuclei coordinate the autonomic and endocrine responses of emotional states. The cerebral cortex is divided into four lobes: frontal, parietal, temporal, and occipital (**Figure 1-2B**).

The brain is also commonly divided into three broader regions: the **hindbrain** (the medulla, pons, and cerebellum), **midbrain**, and **forebrain** (diencephalon and cerebral hemispheres). The hindbrain (excluding the cerebellum) and midbrain comprise the brain stem.

**Figure 1-2B** The four lobes of the cerebral cortex.
Thus, the right hand, the one most humans use for writing and skilled movements, is controlled by the left hemisphere, the same hemisphere that controls electrically stimulating a localized region of the precentral gyrus of the brain. These discrete regions were invariably located in the contralateral motor cortex. 

Galvanized the scientific community by showing that characteristic and discrete limb movements in dogs, such as extending a paw, can be produced by Broca's work stimulated a search for the cortical sites of other specific behavioral functions—a search soon rewarded. In 1870 Gustav Fritsch and Eduard Hitzig function, from the processes on the function:

One reason this conclusion eluded investigators for so many years lies in another organizational principle of the nervous system known as parallel distributed processing. As we shall see below, many sensory, motor, and cognitive functions are served by more than one neural pathway. When one functional region or pathway is damaged, others may be able to compensate partially for the loss, thereby obscuring the behavioral evidence for localization. Nevertheless, the neural pathways for certain higher functions have been precisely mapped in the brain.

Cognitive Functions Are Localized Within the Cerebral Cortex

The brain operations responsible for our cognitive abilities occur primarily in the cerebral cortex—the furrowed gray matter covering the cerebral hemispheres. In each of the brain's two hemispheres the overlying cortex is divided into four anatomically distinct lobes: frontal, parietal, temporal, and occipital (see Figure 1-2B), originally named for the skull bones that encase them. These lobes have specialized functions. The frontal lobe is largely concerned with planning future action and with the control of movement; the parietal lobe with somatic sensation, with forming a body image, and with relating one's body image with extrapersonal space; the occipital lobe with vision; the temporal lobe with hearing; and through its deep structures—the hippocampus and the amygdaloid nuclei—with aspects of learning, memory, and emotion. Each lobe has several characteristic deep infoldings (a favored evolutionary strategy for packing in more cells in a limited space). The crests of these convolutions are called gyri, while the intervening grooves are called sulci or fissures. The more prominent gyri and sulci are quite similar in everyone and have specific names. For example, the central sulcus separates the precentral gyrus, which is concerned with motor function, from the postcentral gyrus, which is concerned with sensory function (Figure 1-4A).

The organization of the cerebral cortex is characterized by two important features. First, each hemisphere is concerned primarily with sensory and motor processes on the contralateral (opposite) side of the body. Thus sensory information that arrives at the spinal cord from the left side of the body—from the left hand, say—crosses over to the right side of the nervous system (either within the spinal cord or in the brain stem) on its way to the cerebral cortex. Similarly, the motor areas in the right hemisphere exert control over the movements of the left half of the body. Second, although the hemispheres are similar in appearance, they are not completely symmetrical in structure nor equivalent in function.

To illustrate the role of the cerebral cortex in cognition, we will trace the development of our understanding of the neural basis of language, using it as an example of how we have progressed in localizing mental functions in the brain. The neural basis of language is discussed more fully in Chapter 59.

Much of what we know about the localization of language comes from studies of aphasia, a language disorder found most often in patients who have suffered a stroke (the occlusion or rupture of a blood vessel supplying blood to a portion of the cerebral hemisphere). Many of the important discoveries in the study of aphasia occurred in rapid succession during the last half of the nineteenth century. Taken together, these advances form one of the most exciting chapters in the study of human behavior, because they offered the first insight into the biological basis of a complex mental function.

The French neurologist Pierre Paul Broca was much influenced by Gall and by the idea that functions could be localized. But he extended Gall's thinking in an important way. He argued that phrenology, the attempt to localize the functions of the mind, should be based on examining damage to the brain produced by clinical lesions rather than by examining the distribution of bumps on the outside of the head. Thus he wrote in 1861: "I had thought that if there were ever a phrenological science, it would be the phrenology of convolutions (in the cortex), and not the phrenology of bumps (on the head)." Based on this insight Broca founded neuropsychology, a new science of mental processes that he was to distinguish from the phrenology of Gall.

In 1861 Broca described a patient named Leborgne, who could understand language but could not speak. The patient had none of the conventional motor deficits (of the tongue, mouth, or vocal cords) that would affect speech. In fact, he could utter isolated words, whistle, and sing a melody without difficulty. But he could not speak grammatically or create complete sentences, nor could he express ideas in writing. Postmortem examination of this patient's brain showed a lesion in the posterior region of the frontal lobe (now called Broca's area; Figure 1-4B). Broca studied eight similar patients, all with lesions in this region, and in each case found that the lesion was located in the left cerebral hemisphere. This discovery led Broca to announce in 1864 one of the most famous principles of brain function: "Nous parlons avec l'hémisphère gauche!" ("We speak with the left hemisphere!")

Broca's work stimulated a search for the cortical sites of other specific behavioral functions—a search soon rewarded. In 1870 Gustav Fritsch and Eduard Hitzig galvanized the scientific community by showing that characteristic and discrete limb movements in dogs, such as extending a paw, can be produced by electrically stimulating a localized region of the precentral gyrus of the brain. These discrete regions were invariably located in the contralateral motor cortex. Thus, the right hand, the one most humans use for writing and skilled movements, is controlled by the left hemisphere, the same hemisphere that controls speech. In most people, therefore, the left hemisphere is regarded as dominant.
The next step was taken in 1876 by Karl Wernicke. At age 26 Wernicke published a now classic paper, "The

Symptom-Complex of Aphasia: A Psychological Study on an Anatomical Basis." In it he described another type of aphasia, one involving a failure to comprehend language rather than to speak (a receptive as opposed to an expressive malfunction). Whereas Broca's patients could understand language but not speak, Wernicke's patient could speak but could not understand language. Moreover, the locus of this new type of aphasia was different from that described by Broca: the critical cortical lesion was located in the posterior part of the temporal lobe where it joins the parietal and occipital lobes (Figure 1-4B).

On the basis of this discovery, and the work of Broca, Fritsch, and Hitzig, Wernicke formulated a theory of language that attempted to reconcile and extend the two theories of brain function holding sway at that time. Phrenologists argued that the cortex was a mosaic of functionally specific areas, whereas the aggregate-field school argued that mental functions were distributed homogeneously throughout the cerebral cortex. Wernicke proposed that only the most basic mental functions, those concerned with simple perceptual and motor activities, are localized to single areas of the cortex. More complex cognitive functions, he argued, result from interconnections between several functional sites. In placing the principle of localized function within a connectionist framework, Wernicke appreciated that different components of a single behavior are processed in different regions of the brain. He was thus the first to advance the idea of distributed processing, now central to our understanding of brain function.

Wernicke postulated that language involves separate motor and sensory programs, each governed by separate cortical regions. He proposed that the motor program, which governs the mouth movements for speech, is located in Broca's area, suitably situated in front of the motor area that controls the mouth, tongue, palate, and vocal cords (Figure 1-4B). And he assigned the sensory program, which governs word perception, to the temporal lobe area he discovered (now called Wernicke's area). This area is conveniently surrounded by the auditory cortex as well as by areas collectively known as association cortex, areas that integrate auditory, visual, and somatic sensation into complex perceptions.

Thus Wernicke formulated the first coherent model for language organization that (with modifications and elaborations we shall soon learn about) is still of some use today. According to this model, the initial steps in the processing of spoken or written words by the brain occur in separate sensory areas of the cortex specialized for auditory or visual information. This information is then conveyed to a cortical association area specialized for both visual and auditory information, the angular gyrus. Here, according to Wernicke, spoken or written words are transformed into a common neural representation shared by both speech and writing. From the angular gyrus this representation is conveyed to Wernicke's area, where it is recognized as language and associated with meaning. Without that association, the ability to comprehend language is lost. The common neural representation is then relayed from Wernicke's to Broca's area, where it is transformed from a sensory (auditory or visual) representation into a motor representation that can potentially lead to spoken or written language. When the last-stage transformation from sensory to motor representation cannot take place, the ability to express language (either as spoken words or in writing) is lost.

Based on this premise, Wernicke correctly predicted the existence of a third type of aphasia, one that results from disconnection. Here the receptive and motor speech zones themselves are spared but the neuronal fiber pathways that connect them are destroyed. This conduction aphasia, as it is now called, is characterized by an incorrect use of words (paraphasia). Patients with conduction aphasia understand words that they hear and read and have no motor difficulties when they speak. Yet they cannot speak coherently; they omit parts of words or substitute incorrect sounds. Painfully aware of their own errors, they are unable to put them right.

Inspired in part by Wernicke, a new school of cortical localization arose in Germany at the beginning of the twentieth century led by the anatomist Korbinian Brodmann. This school sought to distinguish different functional areas of the cortex based on variations in the structure of cells and in the characteristic arrangement of these cells into layers. Using this cytoarchitectonic method, Brodmann distinguished S2 anatomically and functionally distinct areas in the human cerebral cortex (Figure 1-5).

Thus, by the beginning of the twentieth century there was compelling biological evidence for many discrete areas in the cortex, some with specialized roles in
behavior. Yet during the first half of this century the aggregate-field view of the brain, not cellular connectionism, continued to dominate experimental thinking and clinical practice. This surprising state of affairs owed much to the arguments of several prominent neural scientists, among them the British neurologist Henry Head, the German neuropsychologist Kurt Goldstein, the Russian behavioral physiologist Ivan Pavlov, and the American psychologist Karl Lashley, all advocates of the aggregate-field view.

The most influential of this group was Lashley, who was deeply skeptical of the cytoarchitectonic approach to functional delineation of the cortex. "The 'ideal' architectonic map is nearly worthless," Lashley wrote.

P.12

"The area subdivisions are in large part anatomically meaningless, and misleading as to the presumptive functional divisions of the cortex." Lashley's skepticism was reinforced by his attempts, in the tradition of Flourens's work, to find a specific seat of learning by studying the effects of various brain lesions on the ability of rats to learn to run a maze. But Lashley found that the severity of the learning defect seemed to depend on the size of the lesions, not on their precise site. Disillusioned, Lashley—and, after him, many other psychologists—concluded that learning and other mental functions have no special locus in the brain and consequently cannot be pinned down to specific collections of neurons.

On the basis of his observations, Lashley reformulated the aggregate-field view into a theory of brain function called mass action, which further belittled the importance of individual neurons, specific neuronal connections, and brain regions dedicated to particular tasks. According to this view, it was brain mass, not its neuronal components, that was crucial to its function. Applying this logic to aphasia, Head and Goldstein asserted that language disorders could result from injury to almost any cortical area. Cortical damage, regardless of site, caused patients to regress from a rich, abstract language to the impoverished utterances of aphasia.

Lashley's experiments with rats, and Head's observations on human patients, have gradually been reinterpreted. A variety of studies have demonstrated that the maze-learning task used by Lashley is unsuited to the study of local cortical function because the task involves so many motor and sensory capabilities. Deprived of one sensory capability (such as vision), a rat can still learn to run a maze using another (by following tactile or olfactory cues). Besides, as we shall see, many mental functions are handled by more than one region or neuronal pathway, and a single lesion may not eliminate them all.

In addition, the evidence for the localization of function soon became overwhelming. Beginning in the late 1930s, Edgar Adrian in England and Wade Marshall and Philip Bard in the United States discovered that applying a tactile stimulus to different parts of a cat's body elicits electrical activity in distinctively different subregions of the cortex, allowing for the establishment of a precise map of the body surface in specific areas of the cerebral cortex described by Brodmann. These studies established that cytoarchitectonic areas of cortex can be defined unambiguously according to several independent criteria, such as cell type and cell layering, connections, and—most important—physiological function. As we shall see in later chapters, local functional specialization has emerged as a key principle of cortical organization, extending even to individual columns of cells within a functional area. Indeed, the brain is divided into many more functional regions than even Brodmann envisaged!
incoming sensory information that leads to language production and understanding is processed in more than one pathway.

Recall that Wernicke believed that both written and spoken words are transformed into a representation of language by both auditory and visual inputs. This information, he thought, is then conveyed to Wernicke's area, where it becomes associated with meaning before being transformed in Broca's area into output as spoken language. Posner and his colleagues asked: Must the neural code for a word that is read be translated into an auditory representation before it can be associated with a meaning? Or can visual information be sent directly to Broca's area with no involvement of the auditory system? Using PET, they determined how individual words are coded in the brain of normal subjects when the words are read on a screen or heard through earphones. Thus, when words are heard Wernicke's area becomes active, but when words are seen but not heard or spoken Wernicke's area is not activated. The visual information from the occipital cortex appears to be conveyed directly to Broca's area without first being transformed into an auditory representation in the posterior temporal cortex. Posner and his colleagues concluded that the brain pathways and sensory codes used to see words are different from those used to hear words. They proposed, therefore, that these pathways have independent access to higher-order regions of the cortex concerned with the meaning of words and with the ability to express language (Figure 1-6).

Not only are reading and listening processed separately, but the act of thinking about a word's meaning (in the absence of sensory inputs) activates a still different area in the left frontal cortex. Thus language processing is parallel as well as serial; as we shall learn in Chapter 59, it is considerably more complex than initially envisaged by Wernicke. Indeed, similar conclusions have been reached from studies of behavior other than language. These studies demonstrate that information processing requires many individual cortical areas that are appropriately interconnected—each of them responding to, and therefore coding for, only some aspects of specific sensory stimuli or motor movement, and not for others.

Studies of aphasia afford unusual insight into how the brain is organized for language. One of the most impressive insights comes from a study of deaf people who lost their ability to speak American Sign Language after suffering cerebral damage. Unlike spoken language, American signing is accomplished with hand gestures rather than by sound and is perceived by visual rather than auditory pathways. Nonetheless, signing, which has the same structural complexities characteristic of spoken languages, is also localized to the left hemisphere. Thus, deaf people can become aphasic for sign language as a result of lesions in the left hemisphere. Lesions in the right hemisphere do not produce these defects. Moreover, damage to the left hemisphere can have quite specific consequences, affecting either sign comprehension (following damage in Wernicke's area) or grammar (following damage in Broca's area) or signing fluency.

These observations illustrate three points. First, the cognitive processing for language occurs in the left hemisphere and is independent of pathways that process the sensory or motor modalities used in language. Second, speech and hearing are not necessary conditions for the emergence of language capabilities in the left hemisphere. Third, spoken language represents only one of a family of cognitive skills mediated by the left hemisphere.

Figure 1-6 Specific regions of the cortex involved in the recognition of a spoken or written word can be identified with PET scanning. Each of the four images of the human brain shown here (from the left side of the cortex) actually represents the averaged brain activity of several normal subjects. (In these PET images white represents the areas of highest activity, red and yellow quite high activity, and blue and gray the areas of minimal activity.) The "input" component of language (reading or hearing a word) activates the regions of the brain shown in A and B. The motor "output" component of language (speech or thought) activates the regions shown in C and D. (Courtesy of Cathy Price.)

A. The reading of a single word produces a response both in the primary visual cortex and in the visual association cortex (see Figure 1-5).

B. Hearing a word activates an entirely different set of areas in the temporal cortex and at the junction of the temporal/parietal cortex. (To control for irrelevant differences, the same list of words was used in both the reading and listening tests.) A and B show that the brain uses several discrete pathways for processing language and does not transform visual signals for processing in the auditory pathway.

C. Subjects were asked to repeat a word presented either through earphones or on a screen. Speaking a word activates the supplementary motor area of the medial frontal cortex. Broca's area is activated whether the word is presented orally or visually. Thus both visual and auditory pathways converge on Broca's area, the common site for the motor articulation of speech.

D. Subjects were asked to respond to the word "brain" with an appropriate verb (for example, "to think"). This type of thinking activates the frontal cortex as well as Broca's and Wernicke's areas. These areas play a role in all cognition and abstract representation.
Affective Traits and Aspects of Personality Are Also Anatomically Localized

Despite the persuasive evidence for localized languagerelated functions in the cortex, the idea nevertheless persisted that affective (emotional) functions are not localized. Emotion, it was believed, must be an expression of whole-brain activity. Only recently has this view been modified. Although the emotional aspects of behavior have not been as precisely mapped as sensory, motor, and cognitive functions, distinct emotions can be elicited by stimulating specific parts of the brain in humans or experimental animals. The localization of affect has been dramatically demonstrated in patients with certain language disorders and those with a particular type of epilepsy.

Aphasia patients not only manifest cognitive defects in language, but also have trouble with the affective aspects of language, such as intonation (or prosody).

These affective aspects are represented in the right hemisphere and, rather strikingly, the neural organization of the affective elements of language mirrors the organization of the logical content of language in the left hemisphere. Damage to the right temporal area corresponding to Wernicke's area in the left temporal region leads to disturbances in comprehending the emotional quality of language, for example, appreciating from a person's tone of voice whether he is describing a sad or happy event. In contrast, damage to the right frontal area corresponding to Broca's area leads to difficulty in expressing emotional aspects of language.

Thus, it is not useful to represent mental processes as a series of links in a chain, for in such an arrangement the entire process breaks down when a single link is disrupted. The better, more realistic metaphor is to think of mental processes as several railroad lines that all feed into the same terminal. The malfunction of a single link on one pathway affects the information carried by that pathway, but need not interfere permanently with the system as a whole. The remaining parts of the system can modify their performance to accommodate extra traffic after the breakdown of a line.

Perception, movement, language, thought, and memory are all made possible by the serial and parallel interlinking of several brain regions, each with specific distributed in many different regions of the brain. Specific brain regions are not concerned with the task of an independent mental organ dedicated to a complete and distinct mental function (much as the pancreas and the liver are independent digestive organs).

In the aftermath of Wernicke's discovery that there is a modular organization for language in the brain consisting of a complex of serial and parallel processing centers with more or less independent functions, we now appreciate that all cognitive abilities result from the interaction of many simple processing mechanisms distributed in many different regions of the brain. Specific brain regions are not concerned with faculties of the mind, but with elementary processing operations. Perception, movement, language, thought, and memory are all made possible by the serial and parallel interlinking of several brain regions, each with specific functions. As a result, damage to a single area need not result in the loss of an entire faculty as many earlier neurologists predicted. Even if a behavior initially disappears, it may partially return as undamaged parts of the brain reorganize their linkages.

Thus, it is not useful to represent mental processes as a series of links in a chain, for in such an arrangement the entire process breaks down when a single link is disrupted. The better, more realistic metaphor is to think of mental processes as several railroad lines that all feed into the same terminal. The malfunction of a single link on one pathway affects the information carried by that pathway, but need not interfere permanently with the system as a whole. The remaining parts of the system can modify their performance to accommodate extra traffic after the breakdown of a line.

Models of localized function were slow to be accepted because it is enormously difficult to demonstrate which components of a mental operation are represented by a particular pathway or brain region. Nor has it been easy to analyze mental operations and come up with testable components. Only during the last decade, with the convergence of modern cognitive psychology and the brain sciences, have we begun to appreciate that all mental functions are divisible into subfunctions. One difficulty with breaking down mental processes into analytical categories or steps is that our cognitive experience consists of instantaneous, smooth operations. Actually, these processes are composed of numerous independent information-processing components, and even the simplest task requires coordination of several distinct brain areas.

To illustrate this point, consider how we learn about, store, and recall the knowledge that we have in our mind about objects, people, and events in our world. Our common sense tells us that we store each piece of our knowledge of the world as a single representation that can be recalled by memory-jogging stimuli or even by the imagination alone. Everything we know about our grandmother, for example, seems to be stored in one complete representation of "grandmother" that is equally accessible to us whether we see her in person, hear her voice, or simply think about her. Our experience, however, is not a faithful guide to the knowledge we have stored in memory. Knowledge is not stored as complete representations but rather is subdivided into distinct categories and stored separately. For example, the brain stores separately information about animate and inanimate objects. Thus selected lesions in the left temporal lobe's association areas can obliterate a patient's knowledge of living things, especially people, while leaving the patient's knowledge of inanimate objects quite intact.

Representational categories such as "living people" can be subdivided even further. A small lesion in the left temporal lobe can destroy a patient's ability to recognize people by name without affecting the ability to recognize them by sight.

The most astonishing example of the modular nature of representational mental processes is the finding that our very sense of ourselves as a self-conscious coherent being—the sum of what we mean when we say "I"—is achieved through the connection of independent circuits, each with its own sense of awareness, that carry out separate operations in our two cerebral hemispheres. The remarkable discovery that even consciousness is not a unitary process was made by Roger Sperry and Michael Gazzaniga in the course of studying epileptic patients in whom the corpus callosum—the major tract connecting the two hemispheres—was severed as a treatment for epilepsy. Sperry and Gazzaniga found that each hemisphere had a consciousness that was able to function independently of the other. The right hemisphere, which cannot speak, also cannot understand language that is well-understood by the isolated left hemisphere.

As a result, opposing commands can be issued by each hemisphere—each hemisphere has a mind of its own! While one patient was holding a favorite book in his left hand, the right hemisphere, which controls the left hand but cannot read, found that simply looking at the book was boring. The right hemisphere commanded the left hand to put the book down! Another patient would put on his clothes with the left hand, while taking them off with the other. Thus in some
commissurotomized patients the two hemispheres can even interfere with each other's function. In addition, the dominant hemisphere sometimes comments on the performance of the nondominant hemisphere, frequently exhibiting a false sense of confidence regarding problems in which it cannot know the solution, since the information was projected exclusively to the nondominant hemisphere.

Thus the main reason it has taken so long to appreciate which mental activities are localized within which regions of the brain is that we are dealing here with biology's deepest riddle: the neural representation of consciousness and self-awareness. After all, to study the relationship between a mental process and specific brain regions, we must be able to identify the components of the mental process that we are attempting to explain. Yet, of all behaviors, higher mental processes are the most difficult to describe, to measure objectively, and to dissect into their elementary components and operations. In addition, the brain's anatomy is immensely complex, and the structure and interconnections of its many parts are still not fully understood. To analyze how a specific mental activity is represented in the brain, we need not only to determine which aspects of the activity are represented in which regions of the brain, but also how they are represented and how such representations interact.

Only in the last decade has that become possible. By combining the conceptual tools of cognitive psychology with new physiological techniques and brain imaging methods, we are beginning to visualize the regions of the brain involved in particular behaviors. And we are just beginning to discern how these behaviors can be broken down into simpler mental operations and mapped to specific interconnected modules of the brain. Indeed, the excitement evident in neural science today is based on the conviction that at last we have in hand the proper tools to explore the extraordinary organ of the mind, so that we can eventually fathom the biological principles that underlie human cognition.

Selected Readings


References

Adrian ED. 1941. Afferent discharges to the cerebral cortex from peripheral sense organs. J Physiol (Lond) 100: 159–191.


2

Nerve Cells and Behavior

Eric R. Kandel

Humans are vastly superior to other animals in their ability to exploit their physical environment. The remarkable range of human behavior—indeed, the complexity of the environment humans have been able to create for themselves—depends on a sophisticated array of sensory receptors connected to a highly flexible neural machine—a brain—that is able to discriminate an enormous variety of events in the environment. The continuous stream of information from these receptors is organized by the brain into perceptions (some of which are stored in memory for future reference) and then into appropriate behavioral responses. All of this is accomplished by the brain using nerve cells and the connections between them.

Individual nerve cells, the basic units of the brain, are relatively simple in their morphology. Although the human brain contains an extraordinary number of these cells (in the order of $10^{11}$ neurons), which can be classified into at least a thousand different types, all nerve cells share the same basic architecture. The complexity of human behavior depends less on the specialization of individual nerve cells and more on the fact that a great many of these cells form precise anatomical circuits. One of the key organizational principles of the brain, therefore, is that nerve cells with basically similar properties can nevertheless produce quite different actions because of the way they are connected with each other and with sensory receptors and muscle.

Since relatively few principles of organization give rise to considerable complexity, it is possible to learn a great deal about how the nervous system produces behavior by focusing on four basic features of the nervous system:

- The mechanisms by which neurons produce signals.
- The patterns of connections between nerve cells.
- The relationship of different patterns of interconnection to different types of behavior.
- The means by which neurons and their connections are modified by experience.

In this chapter we introduce these four features by first considering the structural and functional properties of neurons and the glial cells that surround and support them. We then examine how individual cells organize and transmit signals and how signaling between a few interconnected nerve cells produces a simple behavior, the knee jerk reflex. Finally, we consider how changes in the signaling ability of specific cells can modify behavior.

The Nervous System Has Two Classes of Cells

There are two main classes of cells in the nervous system: nerve cells (neurons) and glial cells (glia).

Glial Cells Are Support Cells

Glial cells far outnumber neurons—there are between 10 and 50 times more glia than neurons in the central nervous system of vertebrates. The name for these cells derives from the Greek for glue, although in actuality glia do not commonly hold nerve cells together. Rather, they surround the cell bodies, axons, and dendrites of neurons. As far as is known, glia are not directly involved in information processing, but they are thought to have at least seven other vital roles:

- Glial cells support neurons, providing the brain with structure. They also separate and sometimes insulate neuronal groups and synaptic connections from each other.
- Two types of glial cells (oligodendrocytes and Schwann cells) produce the myelin used to insulate nerve cell axons, the cell outgrowths that conduct electrical signals.
- Some glial cells are scavengers, removing debris after injury or neuronal death.
- Glial cells perform important housekeeping chores that promote efficient signaling between neurons [Chapter 14]. For example, some glia also take up chemical transmitters released by neurons during synaptic transmission.
- During the brain's development certain classes of glial cells ("radial glia") guide migrating neurons and direct the outgrowth of axons.
- In some cases, as at the nerve-muscle synapse of vertebrates, glial cells actively regulate the properties of the presynaptic terminal.
- Some glial cells (astrocytes) help form an impermeable lining in the brain's capillaries and venules—the blood-brain barrier—that prevents toxic substances in the blood from entering the brain [Appendix B].
- Other glial cells apparently release growth factors and otherwise help nourish nerve cells, although this role has been difficult to demonstrate conclusively.

Glial cells in the vertebrate nervous system are divided into two major classes: microglia and macroglia.
Microglia are phagocytes that are mobilized after injury, infection, or disease. They arise from macrophages outside the nervous system and are physiologically and embryologically unrelated to the other cell types of the nervous system. Not much is known about what microglia do in the resting state, but they become activated and recruited during infection, injury, and seizure. The activated cell has a process that is stouter and more branched than that of inactivated cells, and it expresses a range of antigens, which suggests that it may serve as the major antigen presenting cell in the central nervous system. Microglia are thought to become activated in a number of diseases including multiple sclerosis and AIDS-related dementia, as well as various chronic neurodegenerative diseases such as Parkinson’s disease and Alzheimer’s disease.

Three types of macroglial cells predominate in the vertebrate nervous system: oligodendrocytes, Schwann cells, and astrocytes.

Oligodendrocytes and Schwann cells are small cells with relatively few processes. Both types carry out the important job of insulating axons, forming a myelin sheath by tightly winding their membranous processes around the axon in a spiral. Oligodendrocytes, which are found in the central nervous system, envelop an average of 15 axonal internodes each (Figure 2-1A). By contrast, Schwann cells, which occur in the peripheral nervous system, each envelop just one internode of only one axon (Figure 2-1B). The types of myelin produced by oligodendrocytes and Schwann cells differ to some degree in chemical makeup.

Astrocytes, the most numerous of glial cells, owe their name to their irregular, roughly star-shaped cell bodies (Figure 2-1C). They tend to have rather long processes, some of which terminate in end-feet. Some astrocytes form end-feet on the surfaces of nerve cells in the brain and spinal cord and may play a role in bringing nutrients to these cells. Other astrocytes place end-feet on the brain’s blood vessels and cause the vessel’s endothelial (lining) cells to form tight junctions, thus creating the protective blood-brain barrier (Figure 2-1C).

Astrocytes also help to maintain the right potassium ion concentration in the extracellular space between neurons. As we shall learn below and in Chapter 7, when a nerve cell fires, potassium ions flow out of the cell. Repetitive firing may create an excess of extracellular potassium that could interfere with signaling between cells in the vicinity. Because astrocytes are highly permeable to potassium, they can take up the excess potassium and so protect those neighboring neurons. In addition, astrocytes take up neurotransmitters from synaptic zones after release and thereby help regulate synaptic activities by removing transmitters. But the role of astrocytes is largely a supporting one.

**Figure 2-1** The principal types of glial cells in the central nervous system are astrocytes and oligodendrocytes and in the peripheral nervous system, Schwann cells.

A. Oligodendrocytes are small cells with relatively few processes. In white matter (left) they provide the myelin, and in gray matter (right) perineurial oligodendrocytes surround and support the cell bodies of neurons. A single oligodendrocyte can wrap its membranous processes around many axons, insulating them with a myelin sheath.

B. Schwann cells furnish the myelin sheaths that insulate axons in the peripheral nervous system. Each of several Schwann cells, positioned along the length of a single axon, forms a segment of myelin sheath about 1 mm long. The sheath assumes its form as the inner tongue of the Schwann cell turns around the axon several times, wrapping it in concentric layers of membrane. The intervals between segments of myelin are known as the nodes of Ranvier. In living cells the layers of myelin are more compact than what is shown here. (Adapted from Alberts et al. 1994.)

C. Astrocytes, the most numerous of glial cells in the central nervous system, are characterized by their star-like shape and the broad end-feet on their processes. Because these endfeet put the astrocyte into contact with both capillaries and neurons, astrocytes are thought to have a nutritive function. Astrocytes also play an important role in forming the blood-brain barrier.

There is no evidence that glia are directly involved in electrical signaling. Signaling is the function of nerve cells.

**Nerve Cells Are the Main Signaling Units of the Nervous System**

A typical neuron has four morphologically defined regions: the cell body, dendrites, the axon, and presynaptic terminals (Figure 2-2). As we shall see later, each of these regions has a distinct role in the generation of signals and the communication of signals between nerve cells.

The cell body (soma) is the metabolic center of the cell. It contains the nucleus, which stores the genes of the cell, as well as the endoplasmic reticulum, an extension of the nucleus where the cell’s proteins are synthesized. The cell body usually gives rise to two kinds of processes: several short dendrites and one, long, tubular axon. Dendrites branch out in tree-like fashion and are the main apparatus for receiving incoming signals from other nerve cells. In contrast, the axon extends away from the cell body and is the main conducting unit for carrying signals to other neurons. An axon can convey electrical signals along distances ranging from 0.1 mm to 3 m. These electrical signals, called action potentials, are rapid, transient, all-or-none nerve impulses, with an amplitude of 100 mV and a duration of about 1 ms (Figure 2-3). Action potentials are initiated at a specialized trigger region at the origin of the axon called the axon hillock (or initial segment of the axon); from there they are conducted down the axon without failure or distortion at rates of 1–100 m per second. The amplitude of an action potential traveling down the axon remains constant because the action potential is an all-or-none impulse that is regenerated at regular intervals along the axon.

Action potentials constitute the signals by which the brain receives, analyzes, and conveys information. These signals are highly stereotyped throughout the
nervous system, even though they are initiated by a great variety of events in the environment that impinge on our bodies—from light to mechanical contact, from odorants to pressure waves. Thus, the signals that convey information about vision are identical to those that carry information about odors. Here we encounter another key principle of brain function. The information conveyed by an action potential is determined not by the form of the signal but by the pathway the signal travels in the brain. The brain analyzes and interprets patterns of incoming electrical signals and in this way creates our everyday sensations of sight, touch, taste, smell, and sound.

To increase the speed by which action potentials are conducted, large axons are wrapped in a fatty, insulating sheath of myelin. The sheath is interrupted at regular intervals by the nodes of Ranvier. It is at these uninsulated spots on the axon that the action potential becomes regenerated. We shall learn more about myelination in Chapter 4 and about action potentials in Chapter 9.

Near its end, the tubular axon divides into fine branches that form communication sites with other neurons. The point at which two neurons communicate is known as a synapse. The nerve cell transmitting a signal is called the presynaptic cell. The cell receiving the signal is the postsynaptic cell. The presynaptic cell transmits signals from the swollen ends of its axon's branches, called presynaptic terminals. However, a presynaptic cell does not actually touch or communicate anatomically with the postsynaptic cell since the two cells are separated by a space, the synaptic cleft. Most presynaptic terminals end on the postsynaptic neuron's dendrites, but the terminals may also end on the cell body or, less often, at the beginning or end of the axon of the receiving cell (Figure 2-2).
As we saw in Chapter 1, Ramón y Cajal provided much of the early evidence for the now basic understanding that neurons are the signaling units of the nervous system and that each neuron is a discrete cell with distinctive processes arising from its cell body (the neuron doctrine). In retrospect, it is hard to appreciate how difficult it was to persuade scientists of this elementary idea. Unlike other tissues, whose cells have simple shapes and fit into a single field of the light microscope, nerve cells have complex shapes; the elaborate patterns of dendrites and the seemingly endless course of some axons made it extremely difficult initially to establish a relationship between these elements. Even after the anatomists Jacob Schleiden and Theodor Schwann put forward the cell theory in the early 1830s—when the idea that cells are the structural units of all living matter became a central dogma of biology—most anatomists would not accept that the cell theory applied to the brain, which they thought of as a continuous web-like reticulum.

The coherent structure of the neuron did not become clear until late in the nineteenth century, when Ramón y Cajal began to use the silver staining method introduced by Golgi. This method, which continues to be used today, has two advantages. First, in a random manner that is still not understood, the silver solution stains only about 1% of the cells in any particular brain region, making it possible to study a single nerve cell in isolation from its neighbors. Second, the neurons that do take up the stain are delineated in their entirety, including the cell body, axon, and full dendritic tree. The stain shows that (with rare exceptions we shall consider later) there is no cytoplasmic continuity between neurons, even at the synapse between two cells. Thus, neurons do not form a syncytium; each neuron is clearly segregated from every other neuron.

Ramón y Cajal applied Golgi’s method to the embryonic nervous systems of many animals, including the human brain. By examining the structure of neurons in almost every region of the nervous system and tracing the contacts they made with one another, Ramón y Cajal was able to describe the differences between classes of nerve cells and to map the precise connections between a good many of them. In this way Ramón y Cajal grasped, in addition to the neuron doctrine, two other principles of neural organization that would prove particularly valuable in studying communication in the nervous system.

The first of these has become known as the principle of dynamic polarization. It states that electrical signals within a nerve cell flow only in one direction: from the receiving sites of the neuron (usually the dendrites and cell body) to the trigger region at the axon. From there, the action potential is propagated unidirectionally along the entire length of the axon to the cell’s presynaptic terminals. Although neurons vary in shape and function, the operation of most follows this rule of information flow. Later in this chapter we shall describe the physiological basis of this principle.

The second principle, the principle of connectional specificity, states that nerve cells do not connect indiscriminately with one another to form random networks; rather each cell makes specific connections—at particular contact points—with certain postsynaptic target cells but not with others. Taken together, the principles of dynamic polarization and connectional specificity form the cellular basis of the modern connectionist approach to the brain discussed in Chapter 1.

Ramón y Cajal was also among the first to realize that the feature that most distinguishes one neuron from another is shape—specifically, the number and form of the processes arising from the cell body. On the basis of shape, neurons are classified into three large groups: unipolar, bipolar, and multipolar.

---

**Figure 2-2 Structure of a neuron.** Most neurons in the vertebrate nervous system have several main features in common. The cell body contains the nucleus, the storehouse of genetic information, and gives rise to two types of cell processes, axons and dendrites. Axons, the transmitting element of neurons, can vary greatly in length; some can extend more than 3 m within the body. Most axons in the central nervous system are very thin (between 0.2 and 20 µm in diameter) compared with the diameter of the cell body (50 µm or more). Many axons are insulated by a fatty sheath of myelin that is interrupted at regular intervals by the nodes of Ranvier. The action potential, the cell’s conducting signal, is initiated either at the axon hillock, the initial segment of the axon, or in some cases slightly farther down the axon at the first node of Ranvier. Branches of the axon of one neuron (the presynaptic neuron) transmit signals to another neuron (the postsynaptic cell) at a site called the synapse. The branches of a single axon may form synapses with as many as 1000 other neurons. Whereas the axon is the output element of the neuron, the dendrites (apical and basal) are input elements of the neuron. Together with the cell body, they receive synaptic contacts from other neurons.

**Figure 2-3 This historic tracing is the first published intracellular recording of an action potential.** It was obtained in 1939 by Hodgkin and Huxley from the squid giant axon, using glass capillary electrodes filled with sea water. Time marker is 500 Hz. The vertical scale indicates the potential of the internal electrode in millivolts, the sea water outside being taken as zero potential. (From Hodgkin and Huxley 1939.)
Neurons can be classified as unipolar, bipolar, or multipolar according to the number of processes that originate from the cell body.

A. Unipolar cells have a single process, with different segments serving as receptive surfaces or releasing terminals. Unipolar cells are characteristic of the invertebrate nervous system.

B. Bipolar cells have two processes that are functionally specialized: the dendrite carries information to the cell, and the axon transmits information to other cells.

C. Certain neurons that carry sensory information, such as information about touch or stretch, to the spinal cord belong to a subclass of bipolar cells designated as pseudo-unipolar. As such cells develop, the two processes of the embryonic bipolar cell become fused and emerge from the cell body as a single process. This outgrowth then splits into two processes, both of which function as axons, one going to peripheral skin or muscle, the other going to the central spinal cord.

D. Multipolar cells have an axon and many dendrites. They are the most common type of neuron in the mammalian nervous system. Three examples illustrate the large diversity of these cells. Spinal motor neurons (left) innervate skeletal muscle fibers. Pyramidal cells (middle) have a roughly triangular cell body; dendrites emerge from both the apex (the apical dendrite) and the base (the basal dendrites). Pyramidal cells are found in the hippocampus and throughout the cerebral cortex. Purkinje cells of the cerebellum (right) are characterized by the rich and extensive dendritic tree in one plane. Such a structure permits enormous synaptic input. (Adapted from Ramón y Cajal 1933.)

Unipolar neurons are the simplest nerve cells because they have a single primary process, which usually gives rise to many branches. One branch serves as the axon; other branches function as dendritic receiving structures. These cells predominate in the nervous systems of invertebrates; in vertebrates they occur in the autonomic nervous system.

Bipolar neurons have an oval-shaped soma that gives rise to two processes: a dendrite that conveys information from the periphery of the body, and an axon that carries information toward the central nervous system. Many sensory cells are bipolar cells, including those in the retina of the eye and in the olfactory epithelium of the nose. The mechanoreceptors that convey touch, pressure, and pain to the spinal cord are variants of bipolar cells called pseudo-unipolar cells. These cells develop initially as bipolar cells; later the two cell processes fuse to form one axon that emerges from the cell body. The axon then splits into two; one branch runs to the periphery (to sensory receptors in the skin, joints, and muscle), the other to the spinal cord.

Multipolar neurons predominate in the nervous system of vertebrates. They have a single axon and, typically, many dendrites emerging from various points around the cell body. Multipolar cells vary greatly in shape, especially in the length of their axons and in the number, length, and intricacy of dendrite branching. Usually the number and extent of their dendrites correlate with the number of synaptic contacts that other neurons make onto them. A spinal motor cell with a relatively modest number of dendrites receives about 10,000 contacts—2000 on its cell body and 8000 on its dendrites. The dendritic tree of a Purkinje cell in the cerebellum is much larger and bushier, as well it might be—it receives approximately 150,000 contacts!

Neurons are also commonly classified into three major functional groups: sensory, motor, and interneuronal. Sensory neurons carry information from the body's periphery into the nervous system for the purpose of both perception and motor coordination. Motor neurons carry commands from the brain or spinal cord to muscles and glands. Interneurons constitute by far the largest class, consisting of all nerve cells that are not specifically sensory or motor. Interneurons are subdivided into two classes. Relay or projection interneurons have long axons and convey signals over considerable distances, from one brain region to another. Local interneurons have short axons and process information within local circuits.
Nerve Cells Form Specific Signaling Networks That Mediate Specific Behaviors

All the behavioral functions of the brain—the processing of sensory information, the programming of motor and emotional responses, the vital business of storing information (memory)—are carried out by specific sets of interconnected neurons. Here we shall examine in general terms how a behavior is produced by considering a simple stretch reflex, the knee jerk. We shall see how a transient imbalance of the body, which puts a stretch on the extensor muscles of the leg, produces sensory information that is conveyed to motor cells, which in turn convey commands to the extensor muscles to contract so that balance will be restored.

The anatomical components of the knee jerk are shown in Figure 2-5. The tendon of the quadriceps femoris, an extensor muscle that moves the lower leg, is attached to the tibia through the tendon of the kneecap, the patellar tendon. Tapping this tendon just below the patella will pull (stretch) the quadriceps femoris. This initiates a reflex contraction of the quadriceps muscle to produce the familiar knee jerk, an extension of the leg smoothly coordinated with a relaxation of the hamstrings, the opposing flexor muscles. By increasing the tension of a selected group of muscles, the stretch reflex changes the position of the leg, suddenly extending it outward. (The regulation of movement by the nervous system is discussed in Section VI.)

Figure 2-5 The knee jerk is an example of a monosynaptic reflex system, a simple behavior controlled by direct connections between sensory and motor neurons. Tapping the kneecap with a reflex hammer pulls on the tendon of the quadriceps femoris, an extensor muscle that extends the lower leg. When the muscle stretches in response to the pull of the tendon, information regarding this change in the muscle is conveyed by afferent (sensory) neurons to the central nervous system. In the spinal cord the sensory neurons act directly on extensor motor neurons that contract the quadriceps, the muscle that was stretched. In addition, the sensory neurons act indirectly, through interneurons, to inhibit flexor motor neurons that would otherwise contract the opposing muscle, the hamstring. These actions combine to produce the reflex behavior. In this schematic drawing each extensor and flexor motor neuron represents a population of many cells.

Stretch reflexes like the knee jerk are a special type of reflex called spinal reflexes, behaviors mediated by neural circuits that are entirely confined to the spinal cord. As we shall see later in the book, such spinal circuits relieve the major motor systems of the brain of having to micromanage elementary behavioral actions. Stretch reflexes are mediated in good part by monosynaptic circuits, in which the sensory neurons and motor neurons executing the action are directly connected to one another, with no interneuron intervening between them. Most other reflexes, including most spinal reflexes, use polysynaptic circuits that include one or more sets of interneurons. Polysynaptic circuits are more amenable to modification by the brain’s higher processing centers.

The cell bodies of the mechanoreceptor sensory neurons involved in the knee jerk are clustered near the spinal cord in a dorsal root ganglion (Figure 2-5). They are pseudo-unipolar cells; one branch of the cell’s axons goes to the quadriceps muscle at the periphery, while the other runs centrally into the spinal cord. The branch that innervates the quadriceps makes contact with stretch-sensitive receptors called muscle spindles and is excited when the muscle is stretched. The branch in the spinal cord forms excitatory connections with the motor neurons that innervate the quadriceps and control its contraction. In addition, this branch contacts local interneurons that inhibit the motor neurons controlling the opposing flexor muscles. These local interneurons are not involved in the stretch reflex itself, but by coordinating motor action they increase the stability of the reflex response. Thus, the electrical signals that produce the stretch reflex convey four kinds of information:

- Sensory information is conveyed to the central nervous system (the spinal cord) from the body’s surface.
- Motor commands from the central nervous system are issued to the muscles that carry out the knee jerk.
- Complementary, inhibitory commands are issued to motor neurons that innervate opposing muscles, providing coordination of muscle action.
- Information about local neuron activity related to the knee jerk is conveyed to higher centers of the central nervous system, thus permitting the brain to coordinate behavioral commands.
The stretching of just one muscle, the quadriceps, activates several hundred sensory neurons, each of which makes direct contact with 100–150 motor neurons (Figure 2-6A). This pattern of connection, in which one neuron activates many target cells, is called *neuronal divergence*; it is especially common in the input stages of the nervous system. By distributing its signals to many target cells, a single neuron can exert wide and diverse influence. For example, sensory neurons involved in a stretch reflex also contact projection interneurons that transmit information about the local neural activity to higher brain regions concerned with coordinating movements. In contrast, because there are usually five to 10 times more sensory neurons than motor neurons, a single motor cell typically receives input from many sensory cells (Figure 2-6B). This pattern of connection, called *convergence*, is common at the output stages of the nervous system. By receiving signals from numerous neurons, the target motor cell is able to integrate diverse information from many sources.

**Figure 2-6** Diverging and converging neuronal connections are a key organizational feature of the brain.

A. In the sensory systems receptor neurons at the input stage usually branch out and make multiple, divergent connections with neurons that represent the second stage of processing. Subsequent connections diverge even more.

B. By contrast, motor neurons are the targets of progressively converging connections. With convergence, the target cell receives the sum of information from many presynaptic cells.

A stretch reflex such as the knee jerk is a simple behavior produced by two classes of neurons connecting at excitatory synapses. But not all important signals in the brain are excitatory. In fact, half of all neurons produce inhibitory signals. Inhibitory neurons release a transmitter that reduces the likelihood of firing. As we have seen, even in the knee-jerk reflex, the sensory neurons make both excitatory connections and connections through inhibitory interneurons. Excitatory connections with the leg’s extensor muscles cause these muscles to contract, while connections with certain inhibitory interneurons prevent the antagonist flexor muscles from being called to action. This feature of the circuit is an example of *feed-forward inhibition* (Figure 2-7A). Feedforward inhibition in the knee-jerk reflex is reciprocal, ensuring that the flexor and extensor pathways always inhibit each other, so only muscles appropriate for the movement, and not those that oppose it, are recruited.


**Figure 2-7 Inhibitory interneurons can produce either feed-forward or feedback inhibition.**

A. Feed-forward inhibition is common in monosynaptic reflex systems, such as the knee-jerk reflex (see Figure 2-5). Afferent neurons from extensor muscles excite not only the extensor motor neurons, but also inhibitory neurons that prevent the firing of the motor cells in the opposing flexor muscles. Feedforward inhibition enhances the effect of the active pathway by suppressing the activity of other opposing pathways.

B. Negative feedback inhibition is a self-regulating mechanism. The effect is to dampen activity within the stimulated pathway and prevent it from exceeding a certain critical maximum. Here the extensor motor neurons act on inhibitory interneurons, which feed back to the extensor motor neurons themselves and thus reduce the probability of firing by these cells.

Neurons can also have connections that provide feedback inhibition. For example, an active neuron may have excitatory connections with both a target cell and an inhibitory interneuron that has its own feedback connection with the active neuron. In this way signals from the active neuron simultaneously excite the target neuron and the inhibitory interneuron, which thus is able to limit the ability of the active neuron to excite its target (Figure 2-7B). We will encounter many examples of feed-forward and feedback inhibition when we examine more complex behaviors in later chapters.

**Signaling Is Organized in the Same Way in All Nerve Cells**

To produce a behavior, a stretch reflex for example, each participating sensory and motor nerve cell sequentially generates four different signals at different sites within the cell: an input signal, a trigger signal, a conducting signal, and an output signal. Regardless of cell size and shape, transmitter biochemistry, or behavioral function, almost all neurons can be described by a model neuron that has four functional components, or regions, that generate the four types of signals: (1) a local input (receptive) component, (2) a trigger (summing or integrative) component, (3) a long-range conducting (signaling) component, and (4) an output (secretory) component. This model neuron is the physiological representation of Ramón y Cajal’s principle of dynamic polarization.

The different types of signals used by a neuron are determined in part by the electrical properties of the cell membrane. At rest, all cells, including neurons, maintain a difference in the electrical potential on either side of the plasma (external) membrane. This is called the resting membrane potential. In a typical resting neuron the electrical potential difference is about 65 mV. Because the net charge outside of the membrane is arbitrarily defined as zero, we say the resting membrane potential is -65 mV. (In different nerve cells it may range from about -40 to -80 mV; in muscle cells it is greater still, about -90 mV.) As we shall see in Chapter 7, the difference in electrical potential when the cell is at rest results from two factors: (1) the unequal distribution of electrically charged ions, in particular, the positively charged Na\(^+\) ions and the negatively charged amino acids and proteins on either side of the cell membrane, and (2) the selective permeability of the membrane to just one of these ions, K\(^+\).

The unequal distribution of positively charged ions on either side of the cell membrane is maintained by a membrane protein that pumps Na\(^+\) out of the cell and K\(^+\) back into it. This Na\(^+\)-K\(^+\) pump, which we shall learn more about in Chapter 7, keeps the Na\(^+\) ion concentration in the cell low (about 10 times lower than that outside the cell) and the K\(^+\) ion concentration high (about 20 times higher than that outside).

At the same time, the cell membrane is selectively permeable to K\(^+\) because the otherwise impermeable membrane contains ion channels, pore-like structures that span the membrane and are highly permeable to K\(^+\) but considerably less permeable to Na\(^+\). When the cell is at rest, these channels are open and K\(^+\) ions tend to leak out. As K\(^+\) ions leak from the cell, they leave behind a cloud of unneutralized negative charge on the inner surface of the membrane, so that the net charge inside the membrane is more negative than on the outside (Figure 2-9).

**Figure 2-8 Most neurons, regardless of type, have four functional regions in common: an input component, a trigger or integrative component, a conductile component, and an output component.** Thus, the functional organization of most neurons can be schematically represented by a model neuron. Each component produces a characteristic signal: the input, integrative, and conductile signals are all electrical, while the output signal consists of the release of a chemical transmitter into the synaptic cleft. Not all neurons share all these features; for example, local interneurons often lack a conductile component.

Excitable cells, such as nerve and muscle cells, differ from other cells in that their membrane potential can be significantly and quickly altered; this change can serve as a signaling mechanism. Reducing the membrane potential by say 10 mV (from -65 mV to -55 mV) makes the membrane much more permeable to Na\(^+\) than to K\(^+\). This influx of positively charged Na\(^+\) ions tends to neutralize the negative charge inside the cell and results in an even greater reduction in membrane potential—the action potential. The action potential is conducted down the cell’s axon to the axon’s terminals which end on other cells (neurons or muscle), where the action potential initiates communication with the other cells. As noted earlier, the action potential is an all-or-none impulse that is actively propagated along the axon, so that its amplitude is not diminished by the time it reaches the axon terminal. Typically, an action potential lasts about one millisecond, after which the membrane returns to its resting state, with its normal separation of charges and higher permeability to K\(^+\) than to Na\(^+\). We shall learn more about the mechanisms underlying the resting potential and action potential in Chapters 6, 7, 8, 9.
In addition to the long-range signal of the action potential, nerve cells also produce local signals, such as receptor potentials and synaptic potentials, that are not actively propagated and therefore typically decay within just a few millimeters. Both long-range and local signals result from changes in the membrane potential, either a decrease or increase from the resting potential. The resting membrane potential therefore provides the baseline against which all signals are expressed. A reduction in membrane potential (e.g., from -65 mV to -55 mV) is called depolarization. Because depolarization enhances a cell's ability to generate an action potential, it is excitatory. In contrast, an increase in membrane potential (e.g., from about -65 mV to -75 mV) is called hyperpolarization. Hyperpolarization makes a cell less likely to generate an action potential and is therefore inhibitory.

The Input Component Produces Graded Local Signals

In most neurons at rest no current flows from one part of the neuron to another, so the resting potential is the same throughout the cell. In sensory neurons current flow is typically initiated by a sensory stimulus, which activates specialized receptor proteins at the neuron's receptive surface. In our example of the knee jerk, stretch of the quadriceps muscle activates specific proteins that are sensitive to stretch of the sensory neuron. The specialized receptor protein forms ion channels in the membrane, through which Na\(^+\) and K\(^+\) flow. These channels open when the cell is stretched, as we shall learn in Chapters 7 and 8, permitting a rapid influx of ions into the sensory cell. This ionic current disturbs the resting potential of the cell membrane, driving the membrane potential to a new level called the receptor potential. The amplitude and duration of the receptor potential depends on the intensity of the muscle stretch. The larger or longer-lasting the stretch, the larger and longer-lasting the resulting receptor potential (Figure 2-10A). Most receptor potentials are depolarizing (excitatory). However, hyperpolarizing (inhibitory) receptor potentials are found in the retina of the eye, as we shall learn in Chapter 26.

![Figure 2-9 The membrane potential of a cell results from a difference in the net electrical charge on either side of its membrane. When a neuron is at rest there is an excess of positive charge outside the cell and an excess of negative charge inside it.](image)

The receptor potential is the first representation of stretch to be coded in the nervous system. It is, however, a purely local signal. The receptor potential—the electrical activity in the sensory neuron initiated by a stimulus—spreads only passively along the axon. It therefore decreases in amplitude with distance and cannot be conveyed much farther than 1 or 2 mm. In fact, at about 1 mm down the axon the amplitude of the signal is only about one-third what it was at the site of generation. To be carried successfully to the rest of the nervous system, the local signal must be amplified—it must generate an action potential. In the knee jerk the receptor potential in the sensory neuron propagates to the first node of Ranvier in the axon, where, if it is large enough, it generates an action potential. In addition to the long-range signal of the action potential, nerve cells also produce local signals, such as receptor potentials and synaptic potentials, that are not actively propagated and therefore typically decay within just a few millimeters. Both long-range and local signals result from changes in the membrane potential, either a decrease or increase from the resting potential. The resting membrane potential therefore provides the baseline against which all signals are expressed. A reduction in membrane potential (e.g., from -65 mV to -55 mV) is called depolarization. Because depolarization enhances a cell's ability to generate an action potential, it is excitatory. In contrast, an increase in membrane potential (e.g., from about -65 mV to -75 mV) is called hyperpolarization. Hyperpolarization makes a cell less likely to generate an action potential and is therefore inhibitory.

The Trigger Component Makes the Decision to Generate an Action Potential

Charles Sherrington first pointed out that the quintessential action of the nervous system is its ability to weigh the consequences of different types of information and then decide on appropriate responses. This integrative action of the nervous system is clearly seen in the actions of the trigger component of the neuron. Action potentials are generated by a sudden influx of Na\(^+\) ions through voltage-sensitive channels in the cell membrane. When an input signal (a receptor potential or synaptic potential) depolarizes the cell membrane, the change in membrane potential opens the Na\(^+\) ion channels, allowing Na\(^+\) to flow down its concentration gradient, from outside the cell where the Na\(^+\) con-
centration is high to inside the cell where it is low. These voltage-sensitive Na\(^+\) channels are concentrated at the initial segment of the axon, an uninsulated portion of the axon just beyond the neuron’s input region. In sensory neurons the highest density of Na\(^+\) channels occurs at the myelinated axon’s first node of Ranvier; in interneurons and motor neurons the highest density occurs at the axon hillock, where the axon emerges from the cell body.

Because it has the highest density of voltage-sensitive Na\(^+\) channels, the initial segment of the axon has the lowest threshold for generating an action potential. Thus, an input signal spreading passively along the cell membrane is more likely to give rise to an action potential at the initial segment of the axon than at other sites in the cell. This part of the axon is therefore known as the impulse initiation zone, or trigger zone. It is here that the activity of all receptor (or synaptic) potentials is summed and where, if the size of the input signal reaches threshold, the neuron fires an action potential.

### Table 2-1 Comparison of Local (Passive) and Propagated Signals

<table>
<thead>
<tr>
<th>Signal type</th>
<th>Amplitude (mV)</th>
<th>Duration</th>
<th>Summation</th>
<th>Effect of signal</th>
<th>Type of propagation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Local (passive) signals</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Receptor potentials</td>
<td>Small (0.1–10)</td>
<td>Brief (5–100 ms)</td>
<td>Graded</td>
<td>Hyperpolarizing or depolarizing</td>
<td>Passive</td>
</tr>
<tr>
<td>Synaptic potentials</td>
<td>Small (0.1–10)</td>
<td>Brief to long (5 ms to 20 min)</td>
<td>Graded</td>
<td>Hyperpolarizing or depolarizing</td>
<td>Passive</td>
</tr>
<tr>
<td>Propagated (active) signals</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Action potentials</td>
<td>Large (70–110)</td>
<td>Brief (1–10 ms)</td>
<td>All-or-none</td>
<td>Depolarizing</td>
<td>Active</td>
</tr>
</tbody>
</table>

The Conductile Component Propagates an All-or-None Action Potential

The action potential, the conducting signal of the neuron, is all-or-none. This means that while stimuli below the threshold will not produce a signal, all stimuli
above the threshold produce the same signal. However much the stimuli vary in intensity or duration, the amplitude and duration of each action potential are pretty much the same. In addition, unlike receptor and synaptic potentials, which spread passively and decrease in amplitude, the action potential does not decay as it travels along the axon to its target—a distance that can measure 3 m in length—because it is periodically regenerated. This conducting signal can travel at rates as fast as 100 meters per second.

The remarkable feature of action potentials is that they are highly stereotyped, varying only subtly (although in some cases importantly) from one nerve cell to another. This feature was demonstrated in the 1920s by Edgar Adrian, who was one of the first to study the nervous system at the cellular level. Adrian found that all action potentials have a similar shape or wave form on the oscilloscope (see Figure 2-3). Indeed, the voltage signals of action potentials carried into the nervous system by a sensory axon often are indistinguishable from those carried out of the nervous system to the muscles by a motor axon.

Only two features of the conducting signal convey information: the number of action potentials and the time intervals between them (Figure 2-10C). As Adrian put it in 1928, summarizing his work on sensory fibers: “... all impulses are very much alike, whether the message is destined to arouse the sensation of light, of touch, or of pain; if they are crowded together the sensation is intense, if they are separated by long intervals the sensation is correspondingly feeble.” Thus, what determines the intensity of sensation or speed of movement is not the magnitude or duration of individual action potentials, but their frequency. Likewise, the duration of a sensation or movement is determined by the period over which action potentials are generated.

If signals are stereotyped and do not reflect the properties of the stimulus, how do neural signals carry specific behavioral information? How is a message that carries visual information distinguished from one that carries pain information about a bee sting, and how do both of these signals differ from messages that send commands for voluntary movement? As we have seen, and will learn to appreciate even more in later chapters, the message of an action potential is determined by the neural pathway that carries it. The visual pathways activated by receptor cells in the retina that respond to light are completely distinct from the somatic sensory pathways activated by sensory cells in the skin that respond to touch or to pain. The function of the signal—be it visual, tactile, or motor—is determined not by the signal itself but by the pathway along which it travels.

The Output Component Releases Neurotransmitter

When an action potential reaches a neuron’s terminal it stimulates the release of a chemical transmitter from the cell. Transmitters can be small molecules, such as L-glutamate and acetylcholine, or they can be peptides like enkephalin (Chapter 13). The neurotransmitter molecules are held in subcellular organelles called synaptic vesicles, which are loaded into specialized release sites in the presynaptic terminals called active zones. To unload their transmitter, the vesicles move up to and fuse with the neuron’s plasma membrane, a process known as exocytosis. (We shall consider neurotransmitter release in Chapter 14.)

The release of chemical transmitter serves as a neuron’s output signal. Like the input signal, the output signal is graded. The amount of transmitter released is determined by the number and frequency of the action potentials in the presynaptic terminals (see Figure 2-10). After the transmitter is released from the presynaptic neuron, it diffuses across the synaptic cleft to receptors in the membrane of the postsynaptic neuron. The binding of transmitter to receptors causes the postsynaptic cell to generate a synaptic potential. Whether the synaptic potential has an excitatory or inhibitory effect will depend on the type of receptors in the postsynaptic cell, not on the particular neurotransmitter. The same transmitter can have different effects on different types of receptors.

Figure 2-10C

The Transformation of the Neural Signal From Sensory to Motor Is Illustrated by the Stretch Reflex Pathway

We have seen that a signal is transformed as it is conveyed from one component of the neuron to the next and from one neuron to the next. This transformation—from input to output—can be seen in perspective by tracing the relay of signals for the stretch reflex.

When a muscle is stretched, the features of the stimulus—its amplitude and duration—are reflected in the amplitude and duration of the receptor potential in the sensory neuron. If the receptor potential exceeds the threshold for action potentials in that cell, the graded signal is transformed at the trigger component into an action potential, an all-or-none signal. The more the receptor potential exceeds the threshold, the greater the depolarization and consequently the greater the frequency of action potentials in the axon; likewise, the duration of the input signal determines the number of action potentials. (Several action potentials together are called a train of action potentials.) This information—the frequency and number of action potentials—is then faithfully conveyed along the entire axon’s length to its terminals, where the frequency of action potentials determines how much transmitter is released.

These stages of transformation have their counterparts in the motor neuron. The transmitter released by a sensory neuron interacts with receptors on the motor neuron to initiate a graded synaptic potential, which spreads to the initial segment of the motor axon. If the membrane potential of the motor neuron reaches a
Each of these components is mediated by a single group or several distinct groups of neurons. But, as far as we know, no complex human behavior is initiated by a single neuron. Rather, each behavior is generated by the actions of many cells.

Even cells with similar organization can differ in important molecular details, expressing different combinations of ion channels, for example. As we shall learn in Chapters 6 and 9, different ion channels provide neurons with various thresholds, excitability properties, and firing patterns. Thus, neurons with different ion channels can encode the same class of synaptic potential into different firing patterns and thereby convey different signals.

The model of neuronal signaling we have outlined is a simplification that applies to most neurons, but there are some important variations. For example, some neurons do not generate action potentials. These are typically local interneurons without a conductile component—they have no axon, or such a short one that a conducted signal is not required. In these neurons the input signals are summed and spread passively to the nearby terminal region, where transmitter is released. There are also neurons that lack a steady resting potential and are spontaneously active.

Neurons also differ in the chemical transmitters they use to transmit information to other neurons, and in the receptors they have to receive information from other neurons. Indeed, many drugs that act on the brain do so by modifying the actions of specific chemical transmitters or a particular subtype of receptor for a given transmitter. These differences not only have physiological importance for day-to-day functioning of the brain, but account for the fact that a disease may affect one class of neurons but not others. Certain diseases, such as amyotrophic lateral sclerosis and polionymeltis, strike only motor neurons, while others, such as tubers dorsalis, a late stage of syphils, affect primarily sensory neurons. Parkinson’s disease, a disorder of voluntary movement, damages a small population of interneurons that use dopamine as a chemical transmitter. Some diseases are selective even within the neuron, affecting only the receptive elements, the cell body, or the axon. In Chapter 16 we shall see how research into myasthenia gravis, caused by a faulty transmitter receptor in the muscle membrane, has provided important insights into synaptic transmission. Indeed, because the nervous system has so many cell types and variations at the molecular level, it is susceptible to more diseases (psychiatric as well as neurological) than any other organ of the body.

Despite the differences among nerve cells, the basic mechanisms of electrical signaling are surprisingly similar. This simplicity is fortunate for those who study the brain. By understanding the molecular mechanisms that produce signaling in one kind of nerve cell, we are well on the way to understanding these mechanisms in many other nerve cells.

Nerve Cells Differ Most at the Molecular Level

The stretch reflex illustrates how just a few types of nerve cells can interact to produce a simple behavior. But even the stretch reflex involves populations of neurons—perhaps a few hundred sensory neurons and a hundred motor neurons. Can the individual neurons implicated in a complex behavior be identified with the same precision? In invertebrate animals, and in some lower vertebrates, a single cell (the so-called command cell) can initiate a complex behavioral response. But, as far as we know, no complex human behavior is initiated by a single neuron. Rather, each behavior is generated by the actions of many cells. Broadly speaking, as we have seen, there are three neural components of behavior: sensory input, intermediate (interneuronal) processing, and motor output. Each of these components is mediated by a single group or several distinct groups of neurons.

As discussed in Chapter 1, one of the key strategies of the nervous system is localization of function: specific types of information are processed in particular brain regions. Thus, information for each of our senses is processed in a distinct brain region where the afferent connections typically form a precise map of the pertinent receptor sheet on the body surface—the skin (touch), the retina (sight), the basilar membrane of the cochlea (hearing), or the olfactory epithelium (smell). These maps are the first stage in creating a representation in the brain of the outside world in which we live. Similarly, areas of the brain concerned with movement contain an orderly arrangement of neural connections representing the musculature and specific movements. The brain, therefore, contains at least two types of neural maps: one for sensory perceptions and another for motor commands. The two maps are interconnected in ways we do not yet fully understand.

The neurons that make up these maps—motor, sensory, and interneuronal—do not differ greatly in their electrical properties. They have different functions because of the connections they make. These connections, established as the brain develops, determine the behavioral function of individual cells. Although our understanding of how sensory and motor information is processed and represented in the brain is based on the detailed studies of only a few regions, in those regions in which our understanding is particularly well advanced it is clear that the logical operations of a mental representation can be understood only by defining the flow of information through the connections that make up the various maps.

A single component of behavior sometimes recruits a number of groups of neurons that simultaneously provide the same or similar information. The deployment of several neuron groups or several pathways to convey similar information is called parallel processing. Parallel processing also occurs in a single pathway when different neurons in the pathway perform similar computations simultaneously. Parallel processing makes enormous sense as an evolutionary strategy for building a more powerful brain: it increases both the speed and reliability of function within the central nervous system.

The importance of abundant, highly specific parallel connections is now also being recognized by scientists attempting to construct computer models of the brain. Scientists working in this field, a branch of computer science known as artificial intelligence, first used serial processing to simulate the brain’s higher-level cognitive processes—processes such as pattern recognition, learning, memory, and motor performance. They soon realized that although these serial models solved many problems rather well, including the challenge of playing chess, they performed poorly with other computations that the brain does almost instantaneously, such as recognizing faces or comprehending speech.

As a result, most computational neurobiologists have turned to systems with both serial and parallel (distributed) components, which they call connectionist models. In these models elements distributed throughout the system process related information simultaneously. Preliminary insights from this work are often consistent with physiological studies. Connectionist models show that individual elements of a system do not transmit large amounts of information. Thus, what makes the brain a remarkable information processing machine is not the complexity of its neurons, but rather its many elements and, in particular, the complexity of connections between them. Individual stereotyped neurons are able to convey unique information because they are wired together and organized in different ways.

The Modifiability of Specific Connections Contributes to the Adaptability of Behavior

That neurons make specific connections with one another simple reflexes can undergo modification that lasts minutes, and much learning results in behavioral changes that can endure for years. How can neural activity produce such long-term changes in the function of a set of prewired connections? A number of solutions for these dilemmas have been proposed. The proposal that has proven most farsighted is the plasticity hypothesis, first put forward at the turn of the century by Ramón y Cajal. A modern form of this hypothesis was advanced by the Polish psychologist Jerzy Konorski in 1948:

The application of a stimulus leads to changes of a twofold kind in the nervous system. ... [T]he first property, by virtue of which the nerve cells react to the incoming impulse ... we call excitability, and ... changes arising ... because of this property we shall call changes due to excitability. The second property, by virtue of which certain permanent functional transformations arise in particular systems of neurons as a result of appropriate stimuli or their combination, we shall call plasticity and the corresponding changes plastic changes.
There is now considerable evidence for plasticity at chemical synapses. Chemical synapses often have a remarkable capacity for short-term physiological changes (lasting hours) that increase or decrease the effectiveness of the synapse. Long-term changes (lasting days) can give rise to further physiological changes that lead to anatomical changes, including pruning of preexisting connections, and even growth of new connections. As we shall see in later chapters, chemical synapses can be modified functionally and anatomically during development and regeneration, and, most importantly, through experience and learning.

Functional alterations are typically short term and involve changes in the effectiveness of existing synaptic connections. Anatomical alterations are typically long-term and consist of the growth of new synaptic connections between neurons. It is this potential for plasticity of the relatively stereotyped units of the nervous system that endows each of us with our individuality.

Selected Readings


References


Some primary sensory neurons are also commonly called afferent neurons, and we use these two terms interchangeably in the book. The term afferent (carried toward the nervous system) applies to all information reaching the central nervous system from the periphery, whether or not this information leads to sensation. The term sensory should, strictly speaking, be applied only to afferent input that leads to a perception.
Genes and Behavior

T. Conrad Gilliam
Eric R. Kandel
Thomas M. Jessell

ALL BEHAVIOR IS SHAPED BY the interplay of genes and the environment. Even the most stereotypic behaviors of simple animals can be influenced by the environment, while highly evolved behaviors in humans, such as language, are constrained by hereditary factors. In this chapter we review what is known about the role of genes in organizing behavior. Later in the book we discuss the role of environmental factors.

A striking illustration of how genes and environment interact is evident in phenylketonuria. This disease results in a severe impairment of cognitive function and affects 1 child in 15,000. Children who express this disease have two abnormal copies of the gene that codes for phenylalanine hydroxylase, the enzyme that converts the amino acid phenylalanine, a component of dietary proteins, to another amino acid, tyrosine. Many more children carry only one abnormal copy of the gene and have no symptoms. Children who lack both functional copies of the gene build up high blood levels of phenylalanine. High blood levels of phenylalanine in turn lead to the production of a toxic metabolite that interferes with the normal maturation of the brain. Fortunately, the treatment for this disease is remarkably simple and effective: the mental retardation can be completely prevented by restricting protein intake, thereby reducing phenylalanine in the diet.

Phenylketonuria is a particularly clear example of how an individual's phenotype depends on the interaction between genes and environment (Figure 3-1). In phenylketonuria both heredity and environmental factors in the diet are clearly necessary for the expression of this form of mental retardation. A mere change in diet can rescue the genetic defect and the mental functioning.

In considering genetic factors that control behavior we need first to identify the components of behavior that are heritable. Clearly, behavior itself is not inherited; what is inherited is DNA, which encodes proteins. The genes expressed in neurons encode proteins that are important for development, maintenance, and regulation of the neural circuits that underlie all aspects of behavior. In turn, neural circuits are composed of many nerve cells, each of which expresses a special constellation of genes that direct the production of specific proteins. For the development and function of a single neural circuit, a wide variety of structural and regulatory proteins are required. In simple animals a single gene may control a behavioral trait by encoding a protein that affects the function of individual nerve cells in a specific neural circuit. In more complex animals the circuitry is also more complex and behavioral traits are generally shaped by the actions of many genes. Subtle differences in behavior can be achieved not only by the presence or absence of a given gene product or a set of products, but also by the degree to which different gene products are expressed, or by the specific contribution of gene products.

The interplay of the genes, proteins, and neural circuits underlying behavior has been studied in various organisms ranging in complexity from worms and flies to mice and humans. Molecular genetics provides the techniques to identify the genes involved in a particular behavior and to determine how the proteins they encode control behavior. In worms, flies, and even in vertebrate organisms such as mice and zebrafish, it is possible to examine directly how genes influence behavior because single-gene mutants of these organisms can be bred and isolated.

In this chapter we illustrate how the genetic dissection of behavior in simple animals can provide insight into the mechanisms that regulate human behavioral traits. We then discuss a few important examples of the effects of single-gene defects on human behavior. Finally, we consider complex behavioral traits that typically are determined by the actions of many genes.

Genetic Information Is Stored in Chromosomes

Genes contribute to the neural circuitry of behavior in two fundamental ways. First, through their ability to replicate reliably, each gene provides precise copies of itself to all cells in an organism as well as succeeding generations of organisms. Second, each gene that is expressed in a cell directs the manufacture of specific proteins that determine the structure, function, and other biological characteristics of the cell.

With rare exceptions, each cell in the human body contains precisely the same complement of genes, thought to be about 80,000. The reason cells differ from one another—why one cell becomes a liver cell and another a brain cell—is that a distinct set of genes is expressed (as messenger RNA) in each cell type. Which genes and proteins become activated in a particular cell depends on interactions between the molecules within the cell, between neighboring cells, and between the cell and the organism’s external environment (see Chapter 52). More of the total genetic information encoded in DNA—perhaps 30,000 of the 80,000 genes—is expressed in brain cells than in any other tissue of the body. Genes vary in size from 1 to 200,000 kilobases; the average size is about 10 kilobases. The DNA of each gene that encodes a protein is made up of segments, called exons, which encode parts of the protein and these coding segments are interrupted by noncoding segments called introns.

DNA is not distributed randomly within the nucleus but arranged in an orderly way on structures called chromosomes. The number of chromosomes varies among
different organisms. In addition, different types of organisms contain either one or two copies of each chromosome. With some exceptions, unicellular organisms are haploid; they have only a single copy of each chromosome. By contrast, most complex multicellular organisms (worms, fruit flies, mice, and humans) are diploid; in all their somatic cells they carry two homologous copies of each chromosome and each gene, one from the mother and the other from the father.

The number of chromosomes in the germ, or sex, cells (sperm and egg) is half that found in somatic cells. During the nuclear division that accompanies somatic cell division (the process of mitosis) the chromosomes are partitioned equally—each daughter cell receives one copy of each chromosome in the parent cell. However, during the two successive nuclear divisions that accompany division of the germ cells (meiosis), the number of chromosomes is reduced by half. Fertilization of the egg by the sperm restores the diploid number found in somatic cells, with homologous chromosomes contributed by each parent.

The 80,000 genes in the human genome are arranged in a precise order along the chromosomes. As a result, each gene is uniquely identifiable by its location at a characteristic position (locus) on a specific chromosome. The two copies of a gene at corresponding loci on a pair of homologous chromosomes commonly harbor sequence variations, or polymorphisms, at multiple sites throughout the gene. At any given site, the alternative gene versions are referred to as alleles. Alleles may be identical or, more commonly, differ to some degree because of polymorphisms or mutations, as discussed below.

If two alleles are identical, the organism is said to be homozygous at that locus. If the alleles vary in form (in their nucleotide sequence), the organism is said to be heterozygous at that locus. The recent DNA sequencing of a small number of human genes reveals large variance in the degree of intergenic polymorphism. In general, however, the rate of polymorphic variation between any two individuals is estimated to be 1 per 1000 base pairs in noncoding DNA and 1 per 2000 base pairs in coding DNA. Thus a 10 kilobase gene would harbor, on average, about 10 polymorphisms, including 1 or 2 in the coding sequence DNA. At each of these polymorphic sites, an individual will carry at most two different forms of the same allele, whereas the same allele may exist in many forms within a population. A difference within a population is called allelic polymorphism, or more generally, genetic polymorphism. Prominent examples of allelic polymorphism are the alleles of the genes responsible for hair and eye color.

Humans have 46 chromosomes: 22 pairs of autosomes and two sex chromosomes (two X chromosomes in females, one X and one Y chromosome in males). The parents contribute the sex chromosomes to their offspring differently from the manner they supply the autosomes. A spermatozoon carries either an X (female-determining) or a Y (male-determining) chromosome, whereas an ovum carries only an X chromosome. As a consequence, males inherit their single X chromosome from their mothers.

The 22 autosome pairs and the X and Y sex chromosomes vary in size and cytological banding pattern (Figure 3-2). Chromosome 1 is the largest autosome; it contains 8% of the human genome, or about 6400 genes. Chromosome 22 is the smallest, containing 1% or about 800 genes. Chromosomes also vary in the nucleotide sequence of their DNA, but paired autosomes are usually morphologically (cytogenetically) indistinguishable.

**Gregor Mendel’s Work Led to the Delineation of the Relationship Between Genotype and Phenotype**

The existence of alternative allelic forms of genes were discovered in 1866 by Gregor Mendel, who demonstrated the difference between dominant and recessive alleles using garden peas as an experimental system. Mendel started out with self-breeding experiments on peas. These led to the creation of inbred strains of peas that bred true for given characteristics of the pea such as color or the shape of the pod. He then crossed these inbred strains with each other and observed how the various traits were manifested in the progeny of the pea plant. These crosses allowed Mendel to appreciate that the variability in heredity among the progeny lay in differences in discrete factors that are passed unchanged from one plant generation to another, factors we now call genes. Moreover, Mendel found that each pea had two sets of factors, one from the male parent and the other from the female.

Mendel carried out his studies before it was known how chromosomes behave during cell division. Forty years later it became clear that the segregation pattern of genes noted by Mendel paralleled, almost exactly, the behavior of chromosomes during meiotic cell division, the division that produces the male and female germ cells. These findings were used by Thomas Hunt Morgan to formulate the chromosomal theory of heredity, according to which each chromosome has a linear array of unique genes running from one end to the other, each gene having a definite location on a particular chromosome.

While studying Mendel’s results, Wilhelm Johannsen later distinguished between the genotype of an organism (its genetic makeup) and the phenotype of an organism (its appearance). In the broad sense genotype refers to the entire set of alleles forming the genome of an individual; in the narrow sense it refers to the specific alleles of one gene. Phenotype denotes the functional expression or consequences of a gene or set of genes. The phenotype of an individual may change throughout life, whereas the genotype remains constant except for sporadic mutations.

Most mutations are simply allelic polymorphisms that are silent; that is, they do not have any effect on the phenotype. Some are not silent but are expressed in ways that nevertheless appear neutral and therefore benign.

**Box 3-1.** Benign mutations are allelic polymorphisms that produce differences in body type, such as eye color or hair color, as well as differences in personality characteristics. The consequence of a mutation is often shaped by the environment. A mutation that favored a hunter-gatherer’s survival during periodic food shortages might lead to pathological obesity in a modern-day environment. Many mutations that do not have benign consequences, such as those leading to excessive tallness, dwarfism, or color blindness, do not necessarily impair everyday functions. Some mutations may have significant consequences that are limited to the cell-biological level, without any functional effects. An example would be a mutation that results in the failure of a single type of cell to develop in an animal that can compensate for the loss of that cell type. Only rarely do mutations lead to significant changes in development, cell function, or overt behavior. Some mutations are truly pathogenic, however, and these lead to human disease.
If a mutant phenotype results from one mutant allele in combination with one wild-type (normal) allele, the mutation or phenotypic trait is said to be dominant. Dominant mutations usually lead to the production of an abnormal protein by the mutant allele or to the expression of the wild-type gene product at an inappropriate time or place. Because they give rise to a new, perhaps toxic, variant of the protein or a new pattern of expression in the body, dominant mutations are often referred to as gain of function mutations. Some dominant mutations produce an inactive protein product that can nevertheless interfere with the function of the wild-type protein, thus leading to a complete loss of function of the gene. Such mutations are termed dominant negative mutations.

If a mutant phenotype is expressed only when both alleles of a gene are mutated (that is, only individuals homozygous for the mutant allele will exhibit the phenotype), the mutation or phenotypic trait is said to be recessive. Recessive mutations usually result from the loss or reduction in amount of a functional protein. As a result, recessive mutations are often loss of function mutations. The reason both alleles need to be defective in a recessive mutation in order for a phenotype to become evident is that a 50% reduction of most proteins (such as most enzymes) usually does not cause serious (or even detectable) problems in cell function.

### Box 3-1 The Origins of Genetic Diversity

Although DNA replication generally is carried out with high fidelity, spontaneous errors called mutations do occur. Mutations may result from damage to the purine and pyrimidine bases, mistakes during the DNA replication process, and recombinations that occur between two nonhomologous chromosomes as a result of errors in crossing over during meiosis. It is these mutations that give rise to genetic polymorphisms.

The rate of spontaneously occurring mutations is low. However, the frequency of mutations greatly increases when the organism is exposed to chemical mutagens or ionizing radiation. Chemical mutagens tend to induce point mutations involving changes in a single DNA base pair or the deletion of a few base pairs. By contrast, ionizing radiation can induce large insertions, deletions, or translocations. Both spontaneous and induced mutations can lead to changes in the structure of the protein encoded by the gene (as in a dominant mutation) or to a partial decrease or absence of gene function or expression (as in recessive mutations).

Changes in a single base pair involve one of three types of point mutations: (1) a missense mutation, where the point mutation results in one amino acid in a protein being substituted for another; (2) a nonsense mutation, where a stop codon (triplet) is substituted for a codon within the coding region, thus resulting in a shortened (truncated) protein product; or (3) a frameshift mutation, in which small insertions or deletions change the reading frame, leading to the production of a truncated or abnormal protein.

Large-scale mutations involve changes in chromosome structure that can affect the function of many contiguous genes. Such mutations include rearrangement of genes without the addition or deletion of material (inversion), duplication of genes in a chromosome, or the exchange (crossing over) between segments of DNA. Sometimes large deletions of multiple genes occur. While these mutations are usually fatal if present in both copies of a gene (homozygous lethals), they can result in phenotypes in the heterozygous state (such as the mental retardation associated with the Wilms tumor deletion complex). Chromosomal translocation can also cause fusion between different (nonhomologous) chromosomes.

### The Genotype Is a Significant Determinant of Human Behavior

Independent of Mendel’s work, Francis Galton began to apply genetics to human behavior in 1869. In his book Hereditary Genius, Galton proposed that relatives of individuals with extremely high mental ability were more likely to be endowed with similar abilities than would be predicted by chance: the closer the family relationship, the higher the incidence of such gifted individuals.

Following Galton’s initial insight, genetic studies of human behavior and disease have relied heavily on the analysis of kinship. Relatives share varying degrees of genetic information and are classified as first degree (parents, siblings, and offspring), second degree (grandparents, grandchildren, nephews and nieces, halfsiblings), third degree (first cousins), and so on, depending on the number of steps, more precisely the number of generations (meiotic events), separating the members of the family tree.
Despite the uncontrolled nature of this early study, Galton was among the first to address the interplay of inheritance (nature) and environment (nurture) in the determination of behavior. Galton was well aware that relatives of eminent individuals also share social, educational, and financial advantages, and that these environmental factors might also account for the correlation between eminence and familial relationship. He therefore endeavored to assess more accurately the relative contributions of heritable and environmental factors to behavioral traits. Thus, in 1883 he introduced the idea of the twin study, a method that today remains a primary strategy for evaluating the role of genes and environment in complex behavioral traits.

Identical twins are monozygotic; they develop from a single zygote that splits into two soon after fertilization. As a result, identical twins share all genes; they are as alike genetically as is possible for two individuals. In contrast, fraternal twins are dizygotic; they develop from two different fertilized eggs. Thus, dizygotic twins, like normal siblings, share on average half their genetic information. Systematic comparisons of pairs of identical versus fraternal twins can be used to assess the importance of genes in the development of a particular trait. If identical twins tend to be more similar (concordant) than fraternal twins, the trait is attributable, at least in part, to genes.

The findings from such twin studies are further sup-

ported by studies of identical twins that have been separated early in life and raised in different households. Despite sometimes great differences in their environment, such twins share a remarkable number of behavioral traits that we normally consider to be distinctive features of individuality, such as intellectual, religious, and vocational interests (Figures 3-3 and 3-4). Behavioral similarities between identical twins that have been separated at birth are attributable in part to genes, although environmental factors may also play a role. In general, twin studies reinforce the idea that human conduct is shaped by genetic factors but do not refute the role of environmental influences, which clearly exist.

The environmental contribution to behavioral traits is often divided into shared and nonshared components. Shared environmental influences, such as childrearing practices or income, may underlie observed phenotypic similarities among family members. In contrast, non-

shared influences, such as interactions with peers in school, can create differences among members of the same family. As discussed below, similarities in personality between biological relatives are due primarily to genetic components, with differences arising from genetic factors and nonshared environmental factors.
Although studies of identical twins and kinships provide strong support for the idea that human behavior has a significant hereditary component, they do not tell us how many genes are important, let alone how specific genes affect behavior. These questions can be addressed by genetic studies in experimental animals in which both the gene and the environment are strictly controlled and by studies of human genetic mutations that give rise to diseases.

**Figure 3-4 Variation in personality in studies of twins.** The units express the degree of variance accounted for by various genetic and environmental influences. (Based on Bouchard 1994.)

### Single Gene Alleles Can Encode Normal Behavioral Variations in Worms and Flies

A number of studies of natural populations of flies and worms have found that allelic polymorphisms in single genes can contribute to individual differences in naturally occurring behavior, including social behavior. The first example was provided by Ron Kondoka and his colleagues, who found variants in the circadian rhythm of flies as a result of molecular polymorphisms in the *period* gene. Wild-type flies vary in how well they can maintain their circadian rhythm in the presence of a temperature change, a feature called temperature compensation. As we will discuss below, the protein products of the period and timeless genes are involved in an autoregulatory feedback that is critical for circadian rhythms. The per gene has a repeat region of threonine-glycine that is polymorphic in length. Two of the major variants (with 17 repeats and 20 repeats) are found in Europe along a north-south cline. Flies with long repeats are better able to compensate for temperature shifts than those with short repeats.
A second example of such individual differences was discovered by Marta Sokolowski and her colleagues while examining the natural variation in the foraging behavior of fly larvae. Some larvae are rovers and others are sitters. Rovers follow longer foraging paths, whereas sitters use much smaller paths. The rover larvae also tend to move between patches of food, while the sitters tend to remain feeding within a food pack. This difference between rovers and sitters results from a single gene called forager. The rover allele has complete dominance over the sitter allele. In nature there are 70% rovers and 30% sitters. In fact, sitter larvae can be converted to rover larvae by expressing in them the gene encoding the rover phenotype. The forager gene encodes a cGMP-dependent protein kinase whose activities are higher in rover than in natural sitters, or sitter mutants, which suggests that the protein kinase may be regulated differently in the two natural variants.

Single genes can even account for differences in normal social behavior. In the course of studying 22 natural isolates of the nematode worm Caenorhabditis elegans collected from various locations around the world, Jonathan Hodgkin and Tabitha Doniach had found that, when grown on the surface of agar-filled Petri plates seeded with Escherichia coli, these natural isolates distributed themselves on the agar surface in two ways. Half the strains dispersed evenly across the bacterial patch, but the other strains spontaneously formed large, dense aggregates called clumps. This clumping arises, at least in part, from interaction among the worms in the clump. Mario deBono and Cornelia Bargmann realized that this reflected an example of individual differences in social behavior. They called the dispersing strains solitary and the clumping strains social.

DeBono and Bargmann have identified natural variants in the behavior of worms feeding on E. coli in a Petri dish. Some worms are solitary foragers, moving across the food and feeding alone, while others are social foragers aggregating together on the food while they feed. More than 50 percent of the social foragers are found in groups, whereas less than two percent of the solitary foragers are found in groups. The social worms may aggregate due to the presence of a mutually attractive, as yet unidentified stimulus.

DeBono and Bargmann gathered social strains of worms that arose from mutagenesis screens of solitary strains in several laboratories and found that the mutation encodes for a gene that resembles the neuropeptide Y receptor, a G protein-coupled receptor that is ubiquitous and important in mammals for feeding. Genetic analysis of normal, wild-type strains showed that the difference between social and solitary strains was due to the substitution of a single amino acid in a cytoplasmic loop of the neuropeptide Y receptor gene. Neuropeptides are found in the brain along with conventional small molecules and are often involved in regulating responses over long periods of time. Since neuropeptide Y receptors are associated with feeding and appetite in mammals, it raises the intriguing possibility that closely related peptides might control foraging and eating behaviors in a variety of organisms that are evolutionarily divergent.

**Mutations in Single Genes Can Affect Certain Behaviors in Flies**

The influence of genes on behavior can be explored most rigorously in simple animals, such as the fruit fly Drosophila. Mutations of single genes in Drosophila can produce abnormalities in learned as well as innate behaviors, such as courtship and circadian rhythms. Moreover, mutations that affect specific aspects of behavior can readily be induced in flies (Box 3-2).

**Box 3-2 Introducing Transgenes in Flies and Mice**

Genes can be manipulated in mice by injecting DNA into the nucleus of newly fertilized eggs (Figure 3-5). In some of the injected eggs the new gene, or transgene, is incorporated into a random site on one of the chromosomes and, since the embryo is at the one-cell stage, the incorporated gene is replicated and ends up in all (or nearly all) of the animal’s cells, including the germline.

Gene incorporation is most easily detected by coinjecting the marker gene for pigment production into an egg obtained from an albino strain. Mice with patches of pigmented fur indicate successful expression of DNA. The transgene’s presence is confirmed by testing a sample of DNA from the injected individuals.

A similar approach is used in flies. The DNA need not be injected directly into a nucleus since the vector used, called a P element, is capable of being incorporated into germ cell nuclei at the time the first cells form in the embryo. The development and function of the nervous system of flies can be altered using promoters that are expressed ubiquitously, such as the inducible heat-shock promoter hsp70 in Drosophila. More specific patterns of expression in brain cells can be obtained using promoter and enhancer sequences from genes that are specific to a cell type.

Transgenes may be wild-type genes that rescue a mutant phenotype or novel “designer” genes that drive expression of a gene in new locations or produce a specifically altered gene product.
The genetic analysis of the behavior of flies has its origins in the behavioral screens performed in the 1970s by Seymour Benzer and his colleagues. These screens detected and isolated mutations that affect circadian (daily) rhythms, courtship behavior, movement, visual perception, and memory. The powerful techniques of Drosophila molecular genetics have enabled investigators to identify these genes and characterize how their protein products act. Here we shall focus on one class of genes isolated by Benzer, those that affect circadian rhythms. In Chapter 63 we shall consider genes in Drosophila that influence memory.

Many aspects of animal physiology and behavior fluctuate in rhythmic cycles. Most of these rhythms follow a circadian period; others follow shorter-term (ultradian) periods. Circadian clocks are thought to have a significant adaptive advantage. For example, they provide a means of anticipating dawn and thereby coordinate physiological functions with environmental conditions. Circadian rhythms affect everything from locomotor activity to mood and play a major role in the biology of motivation (see Chapter 51). Because of the ubiquity of these clocks among animals (and even fungi), experimental advances in invertebrates should aid in our understanding of human circadian behaviors.

Clocks have three basic features. First, the core of the clock is an intrinsic oscillator capable of producing a circadian periodicity of approximately 24 hours. Second, this intrinsic oscillator can adapt its rhythm to changes in the duration of the day-night cycle throughout the year. This regulation is primarily achieved through various light-driven signals that are transmitted by the eye to the brain, where the signals in turn act on the oscillator. Third, there are a set of output pathways from the oscillator that control specific behaviors, such as sleep and wakefulness and locomotor activity.

Mutations altering biological rhythms have been isolated in several organisms. The greatest insight into the oscillator has been obtained from studies of two genes in Drosophila, the period (per) gene, identified originally by Benzer and his colleagues, and the timeless (tim) gene identified recently. The period and timeless genes appear to be devoted almost exclusively to the control of rhythms. Even when they are eliminated, the organism has no other major defects.

Mutations in either the per or tim gene affect the circadian rhythms of locomotor activity and eclosion (i.e., the emergence of the adult from the pupa). Arrhythmic per mutants exhibit no discernible rhythms in either of these behaviors. A long-day per allele produces 28-hour cycles for both locomotor activity and eclosion, whereas two short-day per alleles shorten the cycle (to 19 hours in one case and to 16 hours in the other; see Figure 3-6).

How do the per and tim genes keep time? The answer to this question has begun to emerge from genetic and molecular studies of the two genes and their protein products. The protein products of the per and tim genes (PER and TIM) are thought to shuttle between the cytoplasm and nucleus of cells, regulating expression of target genes, including themselves. As a result, the synthesis and accumulation of the messenger RNAs encoding PER and TIM follow a circadian cycle.

For the proteins to function, PER has to bind to TIM (Figure 3-7). Both genes are transcribed in the morning and their mRNAs accumulate during the day, during which the protein products appear not to be functional. A key step in the regulation of this cycle is the light-induced degradation of the TIM protein. During the day tim RNA is transcribed but the level of TIM protein remains low because of a high rate of degradation. In the absence of TIM, PER does not function. As a result, TIM and PER complexes are not formed. After dusk, when the levels of TIM and PER increase, the two proteins bind to one another, thus becoming functional, and enter the nucleus where they inhibit the transcription of their own genes as well as other, unidentified target genes. As a consequence, per and tim mRNA levels decrease and subsequently protein expression decreases. By morning, PER and TIM protein levels have fallen to low enough levels that they no longer repress transcription.

The finding that the per and tim transcripts are regulated by negative feedback raises the question of why the PER and TIM proteins do not immediately repress their own expression. The answer lies in a built-in delay in accumulation and translocation of the proteins to the nucleus. The PER protein cannot accumulate until sufficient TIM protein is present to bind to and stabilize it. TIM protein, on the other hand, cannot enter the nucleus unless it is bound to PER protein. Accurate time-keeping therefore depends on an oscillatory cycle in gene expression and inactivation by negative feedback.

What does this say for mechanisms in normal and short-cycle flies? In the long-day (28-hour) per mutants the binding affinity of PER proteins for TIM appears to be reduced. Binding thus cannot occur until the two proteins reach higher levels, causing a delay in the entry of the PER-TIM complex into the nucleus and thus extending the period of each cycle.

The mechanisms that control circadian rhythms in other organisms are likely to be similar in principle to...
the mechanism that controls the rhythmicity of the per and tim genes in Drosophila. In mammals circadian behavioral rhythms are governed by the suprachiasmatic nucleus in the hypothalamus (see Chapter 47). Because circadian behavior in mice is precise, it is easy to set up quantitative genetic screens for mutations that alter the circadian behavior. Joseph Takahashi took advantage of the regularity of this behavior to carry out a chemical mutagenesis screen. By this means he identified a semidominant autosomal mutation named clock. Mice homozygous for the clock mutation show extremely long circadian periods followed by a complete loss of circadian rhythmicity when transferred to constant darkness (Figure 3-8). The clock gene therefore appears to regulate two fundamental properties of the circadian rhythm in mice: the circadian period itself and the persistence of circadian rhythmicity.

Since no anatomical defects have been observed with the clock mutation, the clock gene appears to encode a protein specific and essential for circadian rhythmicity in the mouse. When the clock gene was cloned it was found to encode a transcription factor, presumably involved in the basic regulation of genes important for the circadian rhythm. Particularly important is the fact that one of the domains of the clock protein (the PAS domain) is also found in PER. This raises the interesting possibility that the clock protein might bind to and interact with a mouse protein homologous to PER. Many mammalian genes related to clock have now been identified and implicated in the control of circadian rhythms.
Defects in Single Genes Can Have Profound Effects on Complex Behaviors in Mice

The use of chemical genetic techniques to identify circadian rhythm mutants in mice underscores the importance of this experimental mammal in behavioral genetic studies. Genetic studies of mouse behavior have begun to provide insight into the genetic bases of some human behavioral disorders. Here we discuss the evidence for a genetic basis for three disorders: obesity, impulsivity, and altered motivational state.

Mutations in the Gene Encoding Leptin Affect Feeding Behavior

Whether an individual is lean, obese, or of intermediate size is determined in large part by the balance between the amount of food consumed and energy expended, a balance governed by both psychological and physiological factors. Genetics studies of obese mice have provided the best insight into the physiological factors that control ingestive behavior.

The physical cloning and characterization of the region around a spontaneous obesity-causing mutation on mouse chromosome 6 led to the identification of the mouse obese (ob) gene and to a highly conserved (homologous) human gene. The mouse ob gene encodes the protein leptin, a small protein of 145 amino acids that is selectively expressed in adipose tissue and released into the bloodstream. Leptin contributes to the homeostatic mechanisms that permit an animal to maintain its weight within 5% of its normal weight for most of its life. Under normal conditions the amount of leptin secreted reflects the total mass of adipose tissue. When adipose tissue decreases, leptin levels decrease and the animal eats more; when adipose tissue increases, leptin levels increase and the animal eats less. Mice with homozygous mutations in the ob gene lack circulating leptin. This lack leads to marked obesity in these mutant animals. When leptin is supplied exogenously, however, food intake and body weight are reduced dramatically.

A receptor for leptin, called OB-R, encodes a protein that is related to a component of certain cytokine receptors that activate specific transcription factors. This leptin receptor is expressed at a high level in the hypothalamus, the part of the brain that controls appetite and feeding (Chapter 32). The gene encoding OB-R is located in the same region of mouse chromosome as the diabetic gene (db). This is interesting because obesity and diabetes are often linked in humans. In fact, db/db mice are also obese and exhibit a phenotype similar to the mice with a mutated ob gene. Moreover, there is good evidence that the db gene encodes the leptin receptor.
Figure 3-8 Locomotor activity records of clock mutant mice. The record shows periods of wheel-running activity by three offspring. All animals were kept on a light-dark cycle (L/D) of 12 hours for the first 7 days, then transferred to constant darkness (D). They later received a 6-hour light pulse (LP) to reset the rhythm. The activity rhythm for the wild-type mouse had a period of 23.1 hours. The period for the heterozygous clock/+ mouse is 24.9 hours. The homozygous clock/clock mice experience a complete loss of circadian rhythmicity upon transfer to constant darkness and transiently express a rhythm of 28.4 hours after the light pulse. (From Takahashi et al. 1994.)

To what extent do these studies of mice provide insight into human disease? Most obese humans are not defective in leptin mRNA or protein levels and indeed produce higher levels than do nonobese individuals. Thus, it is likely that human obesity reflects not a lack of leptin but a failure to respond to normal or even elevated levels of leptin. Failure to respond to leptin could be a result of mutations of the leptin receptor or of molecules that interact with the receptor.

Leptin may affect feeding behavior by regulating neuropeptide and neurotransmitter expression in hypothalamic cells. Lesions of the hypothalamus affect body weight. For example, ablation of the ventromedial hypothalamus or the arcuate nucleus results in obesity. Leptin administration markedly inhibits the biosynthesis and release of neuropeptide Y, a peptide that stimulates food intake when administered to rodents. Remarkably, as we have discussed earlier, the link between neuropeptide Y and food intake appears to have been conserved, in a general sense, between C. elegans and man.

Box 3-3 Generating Mutations in Flies and Mice

Flies

Genetic analysis of behavior in Drosophila relies on behavioral assays of animals in which individual genes have been mutated. Experimental mutations in Drosophila were originally produced through radiation-induced mutagenesis. This method, however, results in large-scale deletions or rearrangements in chromosomes; several genes are often affected, even when small deletions are the target, and molecular characterization of relevant genes is difficult. In contrast, the chemical ethyl methanesulfonate (EMS) induces point mutations and thus facilitates the characterization of mutations at specific loci.

Many spontaneous mutations and chromosomal rearrangements are produced by transposable elements. The most useful class of transposable elements in Drosophila is the P element. P elements encode a transposase enzyme that mediates the mobilization of the element and a repressor product that blocks transposition. P elements have become major tools of the modern Drosophila geneticist.

In one technique, P elements are used to isolate mutations in any Drosophila gene of interest. The investigator screens for mutants of the gene in progeny of crosses between Drosophila strains that carry P elements and those in which they are absent. New mutations result from the transposition of a P element into a gene. A vector is then constructed in which a P element is inserted. This vector is used as a probe to identify and isolate DNA segments that contain P elements; elements inserted into the gene of interest are found within a subset of these segments. The gene can then be cloned and studied.

Mice

Recent advances in molecular manipulation of mammalian genes have permitted in situ replacement of a known, normal gene with a mutant version. The process of generating a strain of mutant mice involves two separate manipulations: the replacement of a gene on a chromosome by homologous recombination in a special cell line known as embryonic stem cells (Figure 3-9), and the subsequent incorporation of this modified cell line into the germ cell
The gene of interest must first be cloned. The gene is mutated, and a selectable marker, usually a drug-resistance gene, is then introduced into the mutated fragment. The altered gene is then transfected into embryonic stem cells, and clones of cells that incorporate the altered gene are isolated. To identify a clone in which the mutated gene has been integrated into the homologous (normal) site, rather than some other random site, DNA samples of each clone are tested.

**Figure 3-9.** Experimentally controlled homologous recombination is the first step in creating mutant mice. Cloned DNA from the mouse gene to be mutated is modified by genetic engineering so that it contains a bacterial gene, neo. Integration of neo into a mouse chromosome makes the mouse cells resistant to drugs that otherwise would be lethal to the cells (drug X). A viral gene, tk, is also added, attached to one end of the mouse DNA. Integration of tk into a mouse chromosome makes the cells sensitive to a different drug (drug Y). (Adapted from Alberts et al. 1994.)

A. Most insertions occur at random sites in the mouse chromosome, and these nearly always include both ends of the engineered DNA fragment. Colonies of cells in which homologous recombination has incorporated the center of the engineered DNA fragment without the ends are obtained by selecting for those rare mouse cells that grow in the presence of both drugs.

B. Most of the cells that grow in the presence of both drugs will carry the targeted gene replacement.

When a suitable clone has been obtained, cells are injected into a mouse embryo at the blastocyst stage (3–4 days after fertilization), when the embryo consists of approximately 100 cells. These embryos are then reintroduced into a female that has been hormonally prepared for implantation and allowed to come to term. Embryonic stem cells in the mouse have the capability of participating in all aspects of development, including the germline. Thus, injected cells can become germ cells and pass on the altered gene.

Since incorporated stem cells generally mix into other tissues besides the germline, their presence can be tested when the injected embryo is born. Initially, this can be done by using a stem cell line from a mouse strain with a fur color different from that of the strain used to obtain the embryo. The mixed (chimeric) offspring appear to have a patchy colored coat. These progeny are then mated to determine if any stem cells have become germ cells. If so, their progeny will carry the altered gene on one of their chromosomes, detectable by analyzing DNA samples from each of the offspring. When the heterozygous individuals are mated together, one-fourth of the progeny will be homozygous mutant. This technique has been used to generate mutations in various genes crucial to development or function in the nervous system.
Figure 3-10. Altered embryonic stem cells derived from mouse blastocysts are used to create transgenic mice. Embryonic stem (ES) cells are transfected with altered DNA. ES cells that have integrated a transgene for a particular trait can be selected by using a donor that carries an additional sequence, such as a drug-resistance gene (see Figure 3-9). An alternative is to assay the transfected ES cells for successful integration of the donor DNA using polymerase chain reaction (PCR) technology. After obtaining a population of ES cells with a high proportion carrying the marker, the cells are then injected into a recipient blastocyst. This blastocyst is implanted into a foster mother to generate a chimeric mouse. Some of the tissues of the chimeric mice will be derived from the cells of the recipient blastocyst; other tissues will be derived from the injected ES cells. To determine whether ES cells have contributed to the germline, the chimeric mouse is crossed with a mouse that lacks the donor trait. Any progeny that have the trait must be derived from germ cells that have descended from the injected ES cells. By this means, an entire mouse is generated from the altered ES cell. (Adapted from Lewin 1994.)

Mutations in the Gene Encoding a Serotonergic Receptor Intensify Impulsive Behavior

Serotonin (5-hydroxytryptamine) is a monoamine that serves as a neurotransmitter in the brain. The level of serotonin is thought to be reduced in depressive illness. As we shall learn later (Chapter 44), neurons that synthesize serotonin are clustered in several nuclei in the brain stem, the most prominent of which are the raphe nuclei. Their axons project to many regions of the brain, notably the cerebral cortex. Neurons that synthesize serotonin modulate the activity of cortical and subcortical neurons in several ways by activating different receptor subtypes: some excitatory, some inhibitory, some both.

Because of its action on different receptors, serotonin has been implicated in the regulation of mood states, including depression, anxiety, food intake, and impulsive violence (see Chapter 61). Several animal studies have shown that aggressive behavior is often associated with decreased activity of serotonergic neurons. These studies are of particular interest because they provide a glimpse of how social and genetic factors interact to modify behavior.

Most animals, including humans, become aggressive when threatened, such as when their territory is invaded, their offspring are attacked, or sexual interactions are prevented. The importance of serotonergic transmission in aggressive behavior is clearly evident in studies of mice in which the gene for the serotonin 1B
receptor has been ablated by targeted deletion (Box 3-3). When mice lacking the serotonin 1B receptor are isolated for four weeks and then exposed to a wild-type mouse, they are much more aggressive than wild-type animals under similar conditions. The mutant mice attack intruders faster than wild-type mice or mice lacking only one copy of the serotonin 1B receptor gene, and the number and intensity of attacks is significantly greater than that of wild-type mice. Thus, the serotonin 1B receptor plays a role in mediating aggressive behavior in mice.

Serotonin activity has been implicated as one of several important biological factors in determining the threshold for violence. People with a history of impulsive aggressive behavior (and of suicide)—and mouse strains that display increased aggressiveness—have low concentrations of serotonin in the brain. Inhibition of serotonin synthesis or destruction of serotonergic neurons increases aggressiveness in mice and monkeys. Finally, certain serotonin agonists that act on the serotonin 1B receptor inhibit aggression.

In humans a variety of social stressors, such as social or sexual abuse during childhood, are thought to lower the biological thresholds for violence, including the level of serotonin in the brain. Indeed, male monkeys raised in isolation have reduced levels of serotonin in their brains, illustrating that both environmental and genetic factors can converge to influence the metabolism of serotonin.

The relationship of serotonin levels to aggression in humans is not simple, however. This complexity is evident in studies of a Dutch family that transmits an X-linked form of mental retardation. Fourteen of the affected males have a history of impulsive behavior that includes arson, rape, and attempted murder. Each of these men carries a point mutation in the gene that encodes the enzyme monoamine oxidase A, one of the two major enzymes that metabolizes monoamines. This class of neurotransmitter includes serotonin, norepinephrine, and dopamine (see Chapters 60 and 61). The mutation apparently leads to increased levels of serotonin, yet the affected people show enhanced impulsiveness. Thus, the relationship between serotonin and aggression is not simply that reduced serotonin causes aggression and enhanced serotonin causes placidity. Both increases and decreases in serotonin levels may enhance aggression. These findings suggest, not surprisingly, that in humans the relationship between serotonin and a complex trait such as aggression is not direct and may be quite subtle. Finally, although monoamines, in particular serotonin, are important in aggressive behavior, other transmitter systems also affect this behavior, as would be expected for a complex behavioral trait.

Deletion of a Gene That Encodes an Enzyme Important for Dopamine Production Disrupts Locomotor Behavior and Motivation

Dopamine, like serotonin, is a major monoaminergic transmitter in the central nervous system. The majority of dopaminergic neurons have their cell bodies in the substantia nigra while their axons project to the corpus striatum. Dopaminergic neurons have been implicated in the regulation of motor behavior—the degeneration of dopaminergic neurons underlies Parkinson's disease, a debilitating disorder of movement. Other dopaminergic pathways are thought to regulate motivated behaviors. Dysfunction of these pathways may contribute to schizophrenia (see Chapter 60).

The role of the dopaminergic system in mammalian behavior has traditionally been studied through pharmacological techniques. Recently, however, gene knockout techniques have been applied to this system. In one set of experiments the ability of neurons to synthesize dopamine was blocked by selectively inactivating the gene that encodes tyrosine hydroxylase, one of the enzymes important in dopamine synthesis. The dopamine-deficient mice were born, began to nurse, and grew normally for about two weeks and then became inactive, failed to eat or drink, and died shortly thereafter. However, daily administration of L-DOPA, the product of tyrosine hydroxylase, restored normal feeding and produced increased activity.

Dopamine is cleared from the synapse by a high-affinity dopamine transporter. In mutant mice with a deficiency in this transporter the amount of extracellular dopamine is 100-fold greater than normal. The mutant mice exhibit spontaneous and excessive locomotion similar to that obtained in normal mice when the dopamine transporter is blocked pharmacologically (as with a psychostimulant such as cocaine).

Single Genes Are Critical Factors in Certain Human Behavioral Traits

Mutations in a Dopamine Receptor May Influence Novelty-Seeking Behavior

As we have seen, studies of identical twins suggest that a number of personality characteristics have a significant heritable component, but in no case has this finding been rigorously demonstrated by identifying a specific gene. One fascinating candidate is novelty-seeking behavior, a behavior characterized by exhilaration or excitement in response to stimuli that are novel. People who score high on tests of novelty seeking tend to be impulsive, exploratory, fickle, excitable, quick-tempered, and extravagant. They often do things for thrills, as opposed to thinking things through before coming to a decision.

Twin studies suggest that novelty-seeking behavior has a heritability of about 40%. A significant component (10% of the genetic component) seems to be due to a polymorphism in a single gene, the gene that encodes the D4 dopamine receptor. Dopamine is involved in exploratory and pleasure-seeking behavior. There are at least five known receptors for dopamine, called D1 to D5 (Chapter 60). The D4 receptor is expressed in the hypothalamus and the limbic areas of the brain concerned with emotion.

In general, the coding sequence of the receptors for dopamine are highly conserved (as are the coding sequence for other receptors to chemical transmitters), and polymorphisms are very rare. Nevertheless, an interesting polymorphism has been found in the D4 receptor. One form of the gene, called the short form, has a 48-base pair DNA sequence in one of its cytoplasmic domains. By contrast, the long form of the D4 receptor gene has seven repeats of this domain.

Additionally, the long and short forms of the receptor appear to have slightly different signaling properties in response to dopamine. It appears that these slight differences in the long form of the receptor correlate with novelty seeking.

Mutations in Opsins Genes Influence Color Perception

Color vision is one of the few cases in which variation in normal human perception can be explained at a molecular level. Molecular cloning techniques have been used to identify and clone the genes encoding the proteins for the red, green, and blue pigments that transduce different wavelengths of light (see Chapter 29). Defects in one or more of the genes encoding red and green pigments lead to varying degrees of color blindness.

The genes for red and green pigments are arrayed head-to-tail, close to one another on the X chromosome and differ in only about 1 in 20 of their amino acid residues. Because of this tandem organization and similarity of sequence, crossing over between the red and green pigment genes occurs frequently, leading to gene rearrangement. The resulting abnormality in both genes explains the origin of many cases of red-green color blindness.

Subtle variations in color perception occur even among individuals with normal color vision. This is attributable to polymorphism in the red pigment gene in humans. In 62% of the male population with normal color vision, amino acid 180 is a serine residue while in the remaining 38% it is an alanine residue. The effects of this sequence difference can be revealed in psychophysical tests in which subjects are asked to match the intensity of a mixture of red and green light. The intensity of red light needed to match a standard depends on the amino acid at position 180. Because females have two X chromosomes, they fall into three groups: homozygotes for Ser180, homozygotes for Ala180, and heterozygotes who display an intermediate phenotype. Thus, a major variation in human color perception can be explained by a small change in the coding sequence of a single gene.
Box 3-4 Genetic Polymorphisms

If two genes are located very near one another they are likely to be inherited together. Thus, if an abnormality of one gene produces a disease and a nearby marker gene encodes a readily recognized phenotypic trait (such as hair or eye color) or a readily detectable gene product (such as a protein present in the blood), people who express the marker will likely also express the disease—even though the marker may have nothing to do with the disease. Both the phenotypic trait and the DNA sequence of the gene vary in the normal population. In the past, genetic markers were used to distinguish variations in the protein coding regions of genes, such as blood group antigens, enzymes, and antigens of the histocompatibility complex. However, coding sequences represent only 5–10% of the total human genome; 90 or 95% of the genome contains noncoding regions. It is now possible to saturate the human genome with markers that distinguish variations that occur in otherwise homologous DNA sequences throughout the whole genome (including noncoding as well as coding sequences). This broad coverage has made it much easier to trace the inheritance of a disease to a specific region of a particular chromosome.

![Restriction fragment length polymorphism (RFLP)](image)

**Figure 3-11A.** The presence of a restriction fragment length polymorphism (RFLP) can be detected by analyzing DNA fragments cleaved by restriction endonucleases, enzymes that cut at specific restriction sites in nucleotide sequences. In this example chromosome b is missing a restriction site that is present on chromosome a. As a result, cutting chromosome b produces a larger than normal DNA fragment in this region. After cutting, the DNA from both chromosomes is separated according to size by means of gel electrophoresis and transferred to nylon filters (in a procedure called Southern blotting). Autoradiography is then used to reveal the polymorphism. Because the b fragment is larger, it is distinguishable from the a fragment. (Adapted from Alberts et al. 1994.)

One type of DNA marker, a restriction fragment length polymorphism (RFLP), is created by differences in DNA sequence in paired alleles. At one allele a cutting site for a particular restriction enzyme (an enzyme that cuts DNA only at a specific nucleotide sequence) is eliminated or an extra site added, while the other allele remains normal. As a result, the restriction enzyme produces DNA fragments of different lengths from the two alleles. These so-called restriction fragments can be separated by electrophoresis in agarose gels and distinguished by specific DNA probes (Figure 3-11A, 3-11B, 3-11C).

When such a polymorphic region of the DNA is closely linked to a particular gene, inheritance of the gene can be traced by following the inheritance of a particular pattern of restriction fragments. The method can be used to trace pathogenic genes (Figure 3-11B, C).
Novelty seeking is a natural variation in human behavior. Color blindness is a similar variation in perception. It may be annoying to those who have it, but it interferes only marginally with life’s function and not at all with longevity. These relatively neutral mutations differ importantly from mutations that produce serious disease.

**Mutations in the Huntingtin Gene Result in Huntington Disease**

One of the first complex human behavioral abnormalities to be traced to a single gene is Huntington disease, a degenerative disorder of the nervous system. Huntington disease affects both men and women with a frequency of about 5 per 100,000. It is characterized by four features: heritability, chorea (incessant, rapid, jerky movements), cognitive impairment (dementia), and death 15 to 20 years after the onset of symptoms. In most patients the onset of the disease occurs in the fourth to fifth decade of life. Thus, the disease often strikes after individuals have married and had children.

Huntington disease involves the death of neurons in the caudate nucleus, a part of the basal ganglia involved in regulating voluntary movement. The death of nerve cells in the caudate nucleus is thought to cause the chorea. The basis for the impaired cognitive functions and eventual dementia is less clear and is due either to a loss of cortical neurons or to the disruption of normal activity in the cognitive portion of the basal ganglia (see Chapter 43). The selective loss of neurons in the caudate nucleus can be demonstrated in living patients using imaging techniques.

Huntington disease is inherited as an autosomal dominant disorder and the mutation is highly penetrant. The Huntington disease gene was identified on
chromosome 4 using a technique based on DNA markers to map heritable disease mutations relative to genetic polymorphisms (Box 3-4). This gene encodes a large protein called Huntingtin, the function of which is as yet unknown.

Figure 3-12. The DNA mutation in Huntington disease is an unstable CAG repeat.

A. The nucleotide sequence in the region of the unstable CAG repeat in the Huntington gene.

B. Distribution of CAG repeat lengths on normal and Huntington disease (HD) chromosomes. The percentages of normal and HD chromosomes containing different CAG repeat lengths (from 6 to 125) are compiled from several published studies.

C. A highly significant inverse correlation between age of onset of Huntington disease movements and CAG repeat length occurs across all HD alleles. (Modified from Gusella and MacDonald 1995.)

The mutated form of the Huntingtin protein contains a stretch of glutamine residues that is much longer than in the normal protein. The codon (CAG) that encodes glutamine is repeated 19–22 times in the normal gene but 48 or more times in the mutated gene (Figure 3-12A). This expansion results in abnormally long stretches of polyglutamine in the protein. The number of ways in which the abnormal stretch of glutamines affect protein function is not known.

Diseases that involve trinucleotide expansion have an additional feature: each successive generation of a family that harbors the mutant gene manifests the disease with greater severity at an earlier age (genetic anticipation). Thus, an individual may have a mild case of Huntington disease that was not manifested until age 60, whereas his great grandchild may develop more serious symptoms by age 40 (Figure 3-12B, C). This trend is due to the instability of the expanded trinucleotide repeat. As the repeat passes through the germline, the number of repeats tends to increase, particularly in the paternal line. These repeats are thought to create hairpin-like structures in DNA that interfere with its replication. As the repeats attain a certain length, the hairpin-like structures stabilize, leading to persistent mistakes in replication and consequently further expansion of the trinucleotide repeat. The polyglutamine structures appear to affect the protein in one of two ways: they may make the altered protein destructive to the cell, producing a gain-of-function mutation; or they may bind other proteins required for normal cellular function. Expanded tri-nucleotide repeat diseases are usually genetically dominant.

| Table 3-1 Neurological Diseases Involving Trinucleotide Repeats |
### Table 3-1

<table>
<thead>
<tr>
<th>Disease</th>
<th>Repeat</th>
<th>Repeat length</th>
<th>Gene product</th>
</tr>
</thead>
<tbody>
<tr>
<td>X-linked spinal and bulbar muscular atrophy</td>
<td>CAG</td>
<td>Normal: 11–34</td>
<td>Disease: 40–62</td>
</tr>
<tr>
<td>Fragile X mental retardation†</td>
<td>CGG</td>
<td>Normal: 6 to ~50</td>
<td>Premutation: 52–200; Disease: 200 to &gt;1000</td>
</tr>
<tr>
<td>Ataxin-1</td>
<td>CTG</td>
<td>Normal: 5–30</td>
<td>Premutation: 42–180; Disease: 200 to &gt;1000</td>
</tr>
<tr>
<td>Huntington disease</td>
<td>CAG</td>
<td>Normal: 11–34</td>
<td>Disease: 57–121</td>
</tr>
<tr>
<td>Spinocerebellar ataxia type 1</td>
<td>CAG</td>
<td>Normal: 19–36</td>
<td>Disease: 43–81</td>
</tr>
<tr>
<td>FRA1X mental retardation†</td>
<td>GCC</td>
<td>Normal: 6–25</td>
<td>Disease: &gt;200</td>
</tr>
<tr>
<td>Dentatorubral-pallidoluysian atrophy</td>
<td>CAG</td>
<td>Normal: 7–23</td>
<td>Disease: 49–75</td>
</tr>
</tbody>
</table>

† Eight diseases are now associated with the expansion of a trinucleotide CAG repeat in the coding region of the responsible gene: spinal and bulbar muscular atrophy (SBMA); Huntington disease (HD); dentatorubral-pallidoluysian atrophy (DRPLA); spinocerebellar ataxia type 1 (SCA1); and SCA2, 3, 6, and 7. In addition, three congenital fragile X syndromes, each associated with hypermethylation and unstable trinucleotide repeats, have been identified: FRAXA (CGG); FRAXE (GCC); and FRAXF (GCC). For each of the FRA genes, expression is extinguished by expansion and methylation.

<table>
<thead>
<tr>
<th>Disease</th>
<th>Repeat</th>
<th>Repeat length</th>
<th>Gene product</th>
</tr>
</thead>
<tbody>
<tr>
<td>X-linked spinal and bulbar muscular atrophy</td>
<td>CAG</td>
<td>Normal: 11–34</td>
<td>Disease: 40–62</td>
</tr>
<tr>
<td>Fragile X mental retardation†</td>
<td>CGG</td>
<td>Normal: 6 to ~50</td>
<td>Premutation: 52–200; Disease: 200 to &gt;1000</td>
</tr>
<tr>
<td>Ataxin-1</td>
<td>CTG</td>
<td>Normal: 5–30</td>
<td>Premutation: 42–180; Disease: 200 to &gt;1000</td>
</tr>
<tr>
<td>Huntington disease</td>
<td>CAG</td>
<td>Normal: 11–34</td>
<td>Disease: 57–121</td>
</tr>
<tr>
<td>Spinocerebellar ataxia type 1</td>
<td>CAG</td>
<td>Normal: 19–36</td>
<td>Disease: 43–81</td>
</tr>
<tr>
<td>FRA1X mental retardation†</td>
<td>GCC</td>
<td>Normal: 6–25</td>
<td>Disease: &gt;200</td>
</tr>
<tr>
<td>Dentatorubral-pallidoluysian atrophy</td>
<td>CAG</td>
<td>Normal: 7–23</td>
<td>Disease: 49–75</td>
</tr>
</tbody>
</table>

Although individuals with repeat length in the “premutation” size range are phenotypically normal, the corresponding chromosomes are very likely to expand to the “disease-length” category in the next meiosis. CGG, CTG, and CGG expansions are transcribed into the noncoding region of the mRNAs, whereas the GAG expansions associated with neuro-degenerative disorders are translated into glutamine repeats. (Adapted from Warren 1996.)

---

Strikingly, many other hereditary diseases of the nervous system involve similar expansions in trinucleotide repeats within the coding region of the gene responsible for the disease. These diseases include Friedreich’s ataxia type 1, spinocerebellar ataxia, and certain spinal and bulbar muscular dystrophies (Table 3-1; Figure 3-13). By contrast, fragile X mental retardation is an X-linked recessive disease that involves a trinucleotide repeat in the control region near the coding region of the gene, leading to the inactivation of the FMR (fragile X mental retardation) gene. As in Huntington disease, progressive death of specific subpopulations of neurons or muscle cells occurs in many of these diseases.

### Most Complex Behavioral Traits in Humans Are Multigenic

So far we have considered examples of the effects of single genes on behavior. Classic genetic analysis focuses on Mendelian traits, which, as we have seen, are normally determined by allelic variation within a single gene. However, most behavioral traits as well as most common genetic disorders are multigenic; they are determined by several genes interacting with environmental factors.

In contrast to single-locus Mendelian traits, multigenic traits do not have a simple recognizable pattern of inheritance (autosomal dominant, recessive, or X-linked), and thus the relative contributions of several genes to one trait is difficult to analyze. Nevertheless, determining which genes contribute to complex human traits has profound implications for the care and treatment of human disease.

Most common multigenic diseases, such as diabetes, coronary artery disease, asthma, schizophrenia, and manic-depressive disorder, are thought to represent a variety of disorders both etiologically and genetically. Thus, different mutant alleles and environmental factors are thought to produce indistinguishable phenotypes. In a typical multigenic disease, such as diabetes, there are scores of different alleles (among 10–12 different loci; see below) distributed throughout the human population of the world that are capable of contributing to the disease. In any one family three or four of these mutant alleles are likely to be sufficient to give rise to the disease. In fact, it is possible that each of the alleles that contribute to a multigenic disease function as a normal polymorphism when expressed by itself but gives rise to disease if expressed together with other alleles in a certain genetic background. Moreover, because monozygotic twins with identical genetic endowment are often discordant for multigenic traits, the role of nongenetic factors must be important.

---

**Figure 3-13** This model of a gene containing three exons and two introns (intervening blue line) depicts the location and type of expanded triplets involved in certain neurological diseases.

CGG repeats are found within the 5′ untranslated region of the first exons of the genes for fragile X syndrome, fragile XE mental retardation (MR), and fragile site 11B. CGG repeats are also found at two fragile sites, XF and 16A, which are not known to be in the vicinity of any genes and, like fragile site 11B, are not known to result in any disease phenotype. GAA repeats are found within the first exon of the X25 gene for Friedrich’s ataxia. CAG repeats occur at five loci responsible for neurological diseases. These repeats are coding regions and thus result in the lengthening of a normal polyglutamine tract in their respective gene products. The repeats for Haw River syndrome are at the same locus as those for dentatorubral-pallidoluysian atrophy and similarly involve expansion of the same CAG repeat. A CTG repeat (CAG on the other strand) occurs in the 3′ untranslated region of the final exon of the protein kinase gene for myotonic dystrophy. (Adapted from Warren 1996.)

Several techniques have facilitated the genomewide search for multigenic disorders in humans. The most common genetic mapping strategy is linkage analysis, in which a gene’s locus is determined by comparing the inheritance of the mutant gene with a precisely mapped polymorphic DNA marker in a family afflicted with the particular disease. DNAmarker is useful if it maps to a unique locus within the human genome and it identifies frequent polymorphic variations between individuals at this locus. Coinheritance of a particular DNAmarker with a mutant phenotype (or disease state) suggests that the marker and the mutant gene are physically closer together on the chromosome.

Until 1980 polymorphisms could only be detected by differences in the behavior of the protein, for example, by differences in enzyme activity or electrophoretic mobility. In the early 1980s it was appreciated that the noncoding regions, which make up 90–95% of the DNA, are the sites of frequent DNA polymorphisms. Indeed, single base pair changes that give rise to variants are relatively frequent in the human genome, with rates perhaps as high as 1 in 500 base pairs, and...
most of these changes occur in noncoding regions. The method of restriction fragment length polymorphisms (see Box 3-4) is used to detect polymorphisms throughout the genome.

The coinheritance of a DNA marker and mutant gene can occur by chance, or it can occur because the two loci recombine infrequently during meiosis, a direct result of their physical proximity. The chance that any two unlinked loci—for example, loci from different chromosomes—will be inherited together is 1/2, and the chance that they will be co-inherited in n siblings is (1/2)^n. Thus, if two loci are co-inherited in all eight affected siblings from a single family, the odds against this being a random event would be (1/2)8 = 256:1. In practice this is a more complicated event, one that is better analyzed by computer programs that calculate the ratio of the odds for and against linkage, while considering various statistical issues, and generate a value known as the lod (log of the odds) score. (For practical purposes a lod score equal to or greater than 3 indicates that evidence for linkage between a gene marker is significant. This represents odds of 20:1 in favor of linkage between the two loci.)

A related method of identifying polymorphisms is the characterization of simple sequence repeats by the polymerase chain reaction (PCR). The construction of high-resolution human genetic maps composed of these markers and the application of semi-automated screening technologies have facilitated linkage analysis.

A gene that contributes to a multigenic trait is often called a quantitative trait locus (QTL) to indicate that it contributes to the genetic variance of a particular trait. QTL analysis is currently being used with mice and rats to track the genes that contribute to a number of behaviors (Box 3-5).

**Box 3-5 Analysis of Multigenic Traits**

Quantitative trait locus (QTL) analysis is a method for identifying the multiple genes that condition a single behavioral trait. QTL analysis requires at least two strains of a species, each of which has been inbred until all members of the group are genetically identical and have two uniform sets of chromosomes. In the hypothetical example described here (Figure 3-14), two strains of mice have been selectively bred for aggressiveness (A) and docility (D).

- Aggressive A-type mice are bred with docile D-type mice, producing a first generation (F1) hybrid offspring in which every mouse has one set of chromosomes from each parent. In the F1 generation the chromosomes in the cells that produce eggs and sperm exchange material. Segments of the mother's and father's DNA are recombined on individual chromosomes.
- The F1 generation is bred back to D-type mice, producing offspring with one recombinant set of chromosomes and one set that is pure D. In each offspring the recombinant chromosome will carry a unique mix of genes from both original strains.
- Second-generation mice will show a range of aggressiveness because more than one gene determines aggressiveness and the mix of genes in the recombinant chromosome set varies. In Figure 3-14 the levels of aggressiveness in the second-generation mice are indicated by the different colors.
- Sites in the genome that contain genes that contribute to aggression are identified by searching each mouse's DNA for genetic markers, landmarks scattered throughout the genome that are known to differ between the aggressive and docile strains. Each marker is examined to determine whether a mouse has inherited the A-type or D-type.
- For each marker, the mice are sorted into those that have A-type DNA at that locus and those that have D-type DNA. The aggressiveness scores for the mice in the two groups are then compared. If the A-type group is significantly more aggressive than the D-type group, that marker represents a QTL that may contain a gene contributing to aggressiveness. Since each QTL interval contains many genes, additional methods must be used to find the one conditioning aggression.

(Modified from Barinaga 1994.)
Linkage analysis is very sensitive to the model of transmission—dominant, recessive, X-linked, and others—and loses power when applied to multigenic traits where the mode of transmission is not known a priori. In the study of multigenic traits, therefore, researchers will often analyze the DNA marker data by linkage analysis (where genetic parameters must be specified prior to analysis) and by various nonparametric analyses that are much less dependent upon underlying genetic parameters. An example is sib-pair analysis where one evaluates whether particular alleles (or chromosomal segments) are shared among affected siblings more often than would be predicted by chance alone. When the degree of allele-sharing reaches statistical significance, one concludes that the causal or predisposing mutation is contained within the shared region.

Family, twin, and adoption studies indicate not only that patients who suffer from the major psychiatric disorders have a genetic predisposition to those disorders but also that in the normal population at large components of character and general cognitive abilities have important genetic components. In the past it was generally assumed that these genetic contributions to character and cognitive functioning decline over the course of one's lifetime because of the accumulation over the years of social and environmental experience. However, a study of the cognitive capabilities of 240 pairs of twins in the ninth decade of their life showed that genes continue to account for 50% of the variance in later life, much as they do earlier in life. Thus, while environmental factors are important, genes clearly contribute to a variety of normal higher mental functions.

Similarly, bipolar affective disorder (manic-depressive illness) frequently occurs in both siblings if they are monozygotic twins, but it occurs less frequently in both siblings if they are dizygotic twins. The heritability of
bipolar affective disorder, as well as that of schizophrenia, has been estimated to be about 50-60% (Figure 3-15). Thus, factors other than genes must play a critical role in determining the onset of disease in these multifactorial disorders.

Like other complex traits, schizophrenia and depression are most likely multigenic and multifactorial. It will be important to distinguish between various models of transmission. According to one (monogenic) model, many genes in the population contribute to schizophrenia but each gene is rare and has a strong effect. Genetic linkage studies now indicate that such a monogenic model is likely to account for only a small fraction of schizophrenia patients. Ascend (oligogenic) model assumes that a small number of genes interact together to create a threshold of vulnerability for the disorder. Yet another (polygenic) model assumes that these disorders result from the cumulative effect of many genes, each with a minute effect. Several genetic forms of epilepsy most likely fit the monogenic model, whereas the major psychiatric illnesses are thought to fit the oligogenic model. There may, however, be a subpopulation of people with major mental illness who suffer from the consequence of a powerful gene.

Schizophrenia and bipolar affective disorder were among the first multigenic traits to be analyzed by genetic linkage analysis. In fact, many of the early lessons learned from multigenic gene mapping came from mistakes made in these pioneering studies. Segregation analysis and genetic modeling studies indicate that both schizophrenia and bipolar disorder result from the effect of a small number of mutant genes. Thus, although mutations in a set of 10 or more genes may contribute to schizophrenia on a population basis (due to genetic heterogeneity), the combined effects of even a subset of these mutants would presumably be sufficient to place an individual at high risk for the disorder.

Furthermore, we know from twin studies that environmental and genetic factors together determine the overall likelihood of manifesting these disorders. According to this multifactorial model, a single mutation would produce a relatively small contribution to the overall predisposition to illness in the population and thus would be difficult to detect by genetic-linkage strategies. In any individual, however, one gene could actually be a quite strong contributor. For this reason, current psychiatric genetic studies usually involve international consortia cooperating in the systematic ascertainment and diagnosis of very large clinical samples, which lend sufficient power for the detection of small genetic contributions to illness. We shall see in Chapter 60 that the genotyping of several pedigrees has provided a possible genetic locus for susceptibility to schizophrenia.

An Overall View

Most aspects of behavior are under genetic control. Evidence for this can be seen in the striking biological similarities of human twins and in our ability to select and breed domestic and laboratory animals for particular behavioral traits. Such breeding experiments generally indicate that behavioral traits are multigenic in origin. Only in rare instances has the source of natural variation been traceable to a single predominant genetic factor, as in the development of certain forms of obesity in mice.

Now, however, we are entering a new era in which it will be much easier to trace genes that control behavior. The availability of the complete genome for an organism will facilitate our understanding of how genes control genetic pathways important for cellular function, and this advance will allow much more effective and meaningful correlations with behavior. Several genomes are already completed: those of Escherichia coli and several other prokaryotic micro-organisms (5,000 genes, about 5 megabase (Mb) pairs), that of the yeast Saccharomyces cervisiae (6,000 genes, 12 Mb), and that of the worm Caenorhabditis elegans (20,000 genes, 97 Mb). The human genome—all 80,000 genes—is likely to be completed by the year 2003, and work on the genomes of Drosophila and mouse are well underway. From the several genomes that have been completed we have already learned a number of surprising facts.

First, the human genome seems to have undergone two major replications from the primitive genome of single-celled organisms.

Second, fully 40% of the genes in yeast and C. elegans are novel; their function is completely unknown.

Third, from C. elegans we have learned that genes fall into two large classes that perform different functions and have different positions on the chromosomes. One set of 5,000 genes performs the core or housekeeping functions of the cell, the genes encode the proteins for intermediary metabolism for the metabolism of DNA, RNA and protein, for cytoskeletal structures, transport and secretion. The housekeeping genes are highly conserved, in both number and structure, and their ancestors have been found in yeast. Most likely they occur in comparable number in all organisms. In C. elegans these core function genes are clustered together in the central region of their chromosomes where they appear to be protected from evolutionary change.

The second set of about 15,000 genes are more specialized, and newer from an evolutionary perspective; they are not found in yeast. These specialized genes are mostly concerned with intercellular signaling, transcription, and other forms of regulatory control unique to multicellular organisms. These newer genes are positioned at the two ends of the chromosomes, where they appear to be more susceptible to evolutionary pressures. They include genes for 400 protein kinases, 480 zinc finger proteins that appear to be transcription factors, and 790 membrane-spanning receptors. Genes have been identified in C. elegans at the two ends of the chromosomes, where they appear to be more susceptible to evolutionary pressures. They include genes for 400 protein kinases, 480 zinc finger proteins that appear to be transcription factors, and 790 membrane-spanning receptors.

Selected Readings

References


1The specific biochemical processes by which high levels of phenylalanine adversely affect maturation of the brain are still not understood.

2This gene rearrangement is the result of unequal crossing over between the X chromosomes in a female. This unequal crossover appears as a hemizygous condition in male offspring (genes on the male’s X chromosome are called hemizygous because they exist only in one copy).

3Penetrance refers to the frequency with which a heritable trait is manifested phenotypically by individuals carrying the mutant gene(s). Thus the Huntington disease gene is 100% penetrant.

4The term multigenic includes both oligogenic and polygenic traits. An oligogenic trait or disorder is determined by a small number of genes, each contributing to the phenotype in a significant way. In contrast, a polygenic trait is the result of many genes, each with a small effect on the phenotype.
I

The Neurobiology of Behavior
Cave Paintings Contain the First Human Signatures. A paleolithic cave painting from the Chauvet cave in the Ardèche region of France showing a negative image of a right human hand. Cave paintings, found in France and Spain in the regions at the borders of the two countries, primarily show game animals—bison, reindeer, horses, deer, oxen, rhinoceros, and mammoths. Although the purpose of the paintings cannot be known for certain, it is believed that they were used in magical or religious rituals to ensure a good hunt. Images of hands occur either in the negative, as shown here, or in the positive, and always in red. While their meaning is uncertain, it is tempting to think that this hand, which is over 30,000 years old, is early evidence of human cognition. (Reproduced with permission from Chauvet J-M, Deschamps EB, Hillare C. 1996. Dawn of Art: The Chauvet Cave, p. 120. New York: Harry N. Abrams, Incorporated.).

The task of neural science is to understand the mental processes by which we perceive, act, learn, and remember. How does the brain produce the remarkable individuality of human action? Are mental processes localized to specific regions of the brain, or do they represent emergent properties of the brain as an organ? If specific mental processes are represented locally in different brain regions, what rules relate the anatomy and physiology of a region to its specific role in mentation? Can these rules be understood better by examining the region as a whole or by studying its individual nerve cells?

To what extent are mental processes hard-wired into the neural architecture of the brain? What do genes contribute to behavior, and how is gene expression in nerve cells regulated by developmental and learning processes? How does experience alter the way the brain processes subsequent events? This book attempts to address these questions. In so doing it describes how neural science is attempting to link molecules to mind—how proteins responsible for the activities of individual nerve cells are related to the complexity of mental processes.

Today, it is possible to link the molecular dynamics of individual nerve cells to representations of perceptual and motor acts in the brain and to relate these internal mechanisms to observable behavior. New imaging techniques permit us to see the human brain in action—to identify specific regions of the brain associated with particular modes of thinking and feeling.

In the first part of this book we consider the degree to which mental functions can be located in specific regions of the brain. We also examine the extent to which a behavior can be understood in terms of the properties of specific nerve cells and their interconnections in one region of the brain.

The human brain is a network of more than 100 billion individual nerve cells interconnected in systems that construct our perceptions of the external world, fix our attention, and control the machinery of our actions. A first step toward understanding the mind, therefore, is to learn how neurons are organized into signaling pathways and how they communicate by means of synaptic transmission. One of the chief ideas we shall develop in this book is that the specificity of the synaptic connections established during development underlie perception, action, emotion, and learning. We must also understand both the innate (genetic) and environmental determinants of behavior. Specifically, we want to know how genes contribute to behavior. Of course, behavior itself is not inherited—what is inherited is DNA. Genes encode proteins that are important for the
development and regulation of the neural circuits that underlie behavior. The environment, which begins to exert its influence in utero, becomes of prime importance after birth.

Modern neural science represents a merger of molecular biology, neurophysiology, anatomy, embryology, cell biology, and psychology. Along with astute clinical observation, neural science has reinforced the idea first proposed by Hippocrates over two millennia ago that the proper study of mind begins with the study of the brain. Cognitive psychology and psychoanalytic theory have emphasized the diversity and complexity of human mental experience. Both disciplines recognize the importance of genetic as well as learned factors in determining behavior. By emphasizing functional mental structure and internal representation, psychoanalysis served as a source of modern cognitive psychology, a psychology that has stressed the logic of mental operations and of internal representations. Experimental cognitive psychology and clinical psychotherapy can now be strengthened by insights into the cellular neurobiology of behavior. The task for the years ahead is to produce a psychology that—though still concerned with problems of how internal representations are generated, with psychodynamics, and with subjective states of mind—is firmly grounded in empirical neural science.
The Cytology of Neurons

James H. Schwartz
Gary L. Westbrook

The cells of the nervous system vary more than those in any other part of the body. Nevertheless, all neurons have common features that distinguish them from cells in other tissues. For example, they typically are highly polarized. Furthermore, cell functions are compartmentalized, an arrangement that contributes significantly to the processing of electrical signals. The chief functional compartments of neurons—the cell body, dendrites, axons, and terminals—are usually separated by considerable distances, a feature that accounts for the functional polarization discussed in Chapter 2. In most neurons the cell body, which contains the nucleus and the organelles for making RNA and protein, contains less than a tenth of the cell's total volume. The dendrites and axon that originate from the cell body make up the remainder. As discussed in Chapter 2, dendrites are thin processes that branch several times and are specially shaped to receive synaptic input from other neurons. The cell body usually gives off a single axon, another thin process that propagates electrical impulses, often over considerable distances, to the neuron's synaptic terminals on other nerve cells or on target organs.

Neurons also differ from most other cells in being excitable. Rapid shifts in electrical potential are made possible by specialized protein structures (ion channels and pumps) in the cell membrane that control the instantaneous flow of ions into and out of the cells. Polarization and electrical excitability are not unique to neurons, however. Epithelial cells and other nonneuronal secretory cells also are polarized, with basolateral and apical surfaces that differ in structure and function. Some nonneural cells, notably muscle, are excitable, and like nerve cells their excitability depends on special protein molecules that allow ions to pass across the plasma membrane. In neurons, however, polarity and excitability are developed to a higher degree, permitting signals to be received, processed, and conducted over long distances.

Although built on a common plan, neurons are quite diverse—over 50 distinct types have been described. This cytological diversity, which results from developmental differentiation, is also apparent on a molecular level. Each neuron expresses a combination of general and specific molecules. The kinds of proteins a cell synthesizes depends on the genes expressed in the cell; each type of cell synthesizes certain macromolecules (enzymes, structural proteins, membrane constituents, and secretory products) and not others. In essence each cell is the macromolecules that it makes. Many of these molecules are common to all cells in the body; some are characteristic of all neurons, others of large classes of neurons, and still others are restricted to only a few nerve cells.
The Structural and Functional Blueprint of Neurons Is Similar to Epithelial Cells

Neurons develop from epithelial cells and retain fundamental epithelial features. For example, both cell types have distinctive poles: the epithelial cell’s basolateral surface corresponds to the aspect of the neuron’s cell body from which dendrites arise, while the apical surface corresponds to the aspect of the neuron from which the axon arises (Figure 4-1A).

The boundaries of the neuron are defined by the external cell membrane, or plasmalemma. Nerve cell membranes have the general asymmetric bilayer structure of all biological membranes and represent a hydrophobic barrier impermeable to most water-soluble substances. The cytoplasm has two main components: the cytosol (including the cytoskeletal matrix) and the membranous organelles.

The cytosol is the aqueous phase of the cytoplasm. In this phase only a very few proteins are freely soluble, mostly enzymes that catalyze various metabolic reactions. Many cytosolic proteins have general housekeeping functions and are common to all neurons. Others have specific roles in particular types of neurons; for example, the enzymes involved in the synthesis and degradation of the particular substance used as a neurotransmitter. Moreover, some cytosolic proteins are distributed unevenly in the cell because they interact to form aggregates, particles, or matrices. Many cytosolic proteins involved in signaling are concentrated at the cell’s periphery in the cytoskeletal matrix immediately adjacent to the plasmalemma.

Membranous Organelles Are Selectively Distributed Throughout the Neuron

The membranous organelles of the cytoplasm include the mitochondria and peroxisomes as well as a complex system of tubules, vesicles, and cisternae (the vacuolar apparatus) that consists of the rough endoplasmic reticulum, the smooth endoplasmic reticulum, the Golgi complex, secretory vesicles, endosomes, lysosomes, and a multiplicity of transport vesicles that functionally interconnect these various compartments (Figures 4-1B and 4-2).

Membranes of the vacuolar apparatus are thought to be derived from deep invaginations of the cell’s external membrane that become discrete organelles. Their lumen corresponds topologically to the outside of the cell; consequently the inner leaflet of their lipid bilayer corresponds to the outer leaflet of the plasmalemma (Figure 4-1B). Even though the major subcompartments of this system are anatomically discontinuous, membranous and lumenal material are moved from one compartment to another with great efficiency and specificity by means of transport vesicles. For example, proteins and phospholipids synthesized in the rough endoplasmic reticulum are transported to the Golgi complex and then to secretory vesicles destined to fuse with the plasmalemma by exocytosis (the secretory pathway). Conversely, membrane taken into the cell in the form of endocytic vesicles is incorporated into early endosomes, which are sorting compartments concentrated at the cell’s periphery; the membrane is then either shuttled back to the plasmalemma by vesicle recycling or directed to late endosomes and eventually to lysosomes for degradation (the endocytic pathway).
**Figure 4-2 Endoplasmic reticulum in a pyramidal cell.** This micrograph of the basal pole of a pyramidal neuron’s cell body, from which a single dendrite emerges, reveals rough and smooth endoplasmic reticulum (ER) above the nucleus (N). A portion of the Golgi complex (G) appears at the base of the dendrite (Den); some Golgi cisternae have entered the dendrite, as have mitochondria (Mit), lysosomes (Ly), and ribosomes (R). Microtubules (Mt) are the prominent cytoskeletal filaments seen in the cytosol. Axon terminals (AT) are seen synapsing on the neuron. (From Peters et al. 1991.)

**Figure 4-3 Under the light microscope the Golgi complex appears as a network of filaments that extend into dendrites (arrows), but not into the axon.** The arrowheads at the bottom indicate the axon hillock. The Golgi complex in this micrograph is in a large neuron of the brain stem immuno-stained with antibodies specifically directed against this organelle. (From De Camilli et al. 1986.)

A specialized portion of the rough endoplasmic reticulum forms a spherical flattened cisterna called the nuclear envelope, which surrounds the chromosomal DNA and its associated proteins and defines the nucleus (see Figure 4-1). This cisterna is continuous with other portions of the rough endoplasmic reticulum. Because of this continuity, the nuclear envelope is presumed to have evolved to ensheathe the chromosomes by an in-vagination of the plasmalemma. The nuclear envelope is interrupted by the nuclear pores, where fusion of the inner and outer membrane of the nuclear envelope results in the formation of hydrophilic channels through which proteins and RNA are exchanged between the cytoplasm proper and the nuclear cytoplasm. Thus the nucleoplasm and cytoplasm can be considered functionally continuous domains of the cytosol.
Figure 4-4 Neurons develop two distinct types of processes, dendrites and axons, even when grown in isolation. The figure shows a hippocampal neuron grown in isolation in primary culture and stained by double immunofluorescence for the synaptic vesicle protein synaptophysin and the transferrin receptor, a protein involved in iron uptake. When photographed through an appropriate filter, immunofluorescence corresponding to the transferrin receptor is seen only in dendrites (A). When photographed for synapsin, synaptic vesicles are selectively concentrated in the axon (arrow) as revealed by synapsin immunofluorescence (B). (From Cameron et al. 1991.)

Mitochondria and peroxisomes make use of molecular oxygen. Mitochondria generate ATP, the major molecule by which cellular energy is transferred or spent. Peroxisomes engage in detoxification through peroxidation reactions and also prevent the accumulation of the strong oxidizing agent hydrogen peroxide. These two organelles, which are thought to be derived from symbiotic organisms that invaded eukaryotic cells early in evolution, are not functionally continuous with the vacuolar apparatus of the cell.
**The Cytoskeleton Determines the Shape of the Neuron**

The cytoskeleton is the major intrinsic determinant of the shape of a neuron and is responsible for the asymmetric distribution of organelles within the cytoplasm. It contains three main filamentous structures: microtubules, neurofilaments (called intermediate filaments in nonneuronal cells), and actin microfilaments (Figures 4-5 and 4-6). These filaments and their associated proteins account for about 25% of the total protein of the neuron.

Microtubules form long scaffolds that extend the full length of the neuron and play a key role in developing and maintaining the neuron’s processes. A single microtubule can be as long as 0.1 mm. Microtubules are constructed of 13 protofilaments in a tubular array with an outside diameter of 25–28 nm (Figure 4-5A). Each protofilament consists of several pairs of α- and β-tubulin subunits arranged linearly. The polar structure of the tubulin dimer creates a plus and a minus end of the polymer. The tubulins are encoded by a multigene family; at least six genes code for both the α- and β-subunits. More than 20 isoforms of tubulin are present in the brain because of the expression of different genes as well as post-translational modifications.

The cytoplasm of the cell body extends into the dendritic tree without any functional boundary. Generally, all organelles present in the cytoplasm of the cell body are also present in dendrites, although the concentrations of some organelles, such as the rough endoplasmic reticulum, the Golgi complex, and lysosomes, progressively diminish with distance from the cell body. In contrast, a sharp functional boundary exists at the axon hillock, the point of emergence of the axon. For example, ribosomes, the rough endoplasmic reticulum, and the Golgi complex—the organelles that represent the main protein biosynthetic machinery of the neuron—for the most part are excluded from axons (Figure 4-3). Lysosomes and certain proteins, which in epithelial cells are selectively targeted to the basolateral surface of the cell, are also excluded from axons. Axons are, however, rich in synaptic vesicles, synaptic vesicle precursor membranes, and endocytic intermediates involved in synaptic vesicle traffic (Figures 4-1 and 4-4).

Mitochondria and the smooth endoplasmic reticulum are present in all neuronal compartments, including the axon. The smooth endoplasmic reticulum is anatomically continuous with the rough endoplasmic reticulum. One of its functions is to act as a regulated Ca²⁺ store throughout the neuronal cytoplasm. It also performs a variety of enzymatic reactions and is involved in lipid metabolism.
Figure 4-7 The dendritic architecture in the cerebellar cortex is visualized here by immunoperoxidase staining for the microtubule-associated protein MAP2, a dendrite-specific MAP. Dendrites of all classes of neurons are stained. The field is dominated by the dendrites of Purkinje cells. (Courtesy of P. De Camilli.)

Tubulin is a GTPase and microtubules grow by the addition of GTP-bound tubulin dimers at their plus end. Shortly after polymerization GTP is hydrolyzed to GDP. When a microtubule stops growing its plus end becomes capped by GDP-bound tubulin. Given the low affinity of the GDP-bound tubulin for the polymer, this would lead to rapid catastrophic depolymerization unless the microtubule were stabilized by interaction with other proteins. In fact, microtubules undergo rapid cycles of polymerization and depolymerization in dividing cells, but they are much more stable in mature dendrites and axons. This stability is due to microtubule-associated proteins (MAPs), which promote the oriented polymerization and assembly of the microtubules. The MAPs in the axons differ from those in the dendrites. For example, MAP2 is present in dendrites but absent from axons (Figure 4-7), while tau and MAP3 are present in the axon.
A sensory (dorsal root ganglion) cell and a spinal motor neuron form a monosynaptic circuit that controls the knee-jerk stretch reflex.

A. Sensory neuron. **Left:** The axon of the primary sensory neuron is typically quite convoluted before it bifurcates into a central and a peripheral branch. The cell body contains a prominent nucleus. (From Dogiel 1908.) **Right:** Low-power electron micrograph shows the cell body of a large dorsal root ganglion cell. A prominent nucleolus (Nuc) can be seen within the nucleus (N). The cell body of the neuron is surrounded by Schwann cells (Sc), the type of glial cells found in the peripheral nervous system. (Courtesy of R. E. Coggeshall and F. Mandriota.)

B. Motor neuron. **Left:** Many dendrites typically branch from the cell bodies of spinal motor neurons, as shown by five spinal motor neurons in the ventral horn of a kitten. (From Ramón y Cajal 1909.) **Right:** Detail of the cell body of a motor neuron is shown in this photomicrograph. An enormous number of nerve endings from presynaptic neurons (arrows) are visible. These terminals, called synaptic boutons, appear as knob-like enlargements on the cell membrane. The synaptic boutons are prominent in this micrograph because the tissue is specially impregnated with silver. Three dendrites (Den) are also shown. The nucleus and its nucleolus are surrounded by Nissl substance (Nn), clumps of ribosomes associated with the membrane of the endoplasmic reticulum. (Courtesy of G. L. Rasmussen.)

Neurofilaments, 10 nm in diameter, are the bones of the cytoskeleton (see Figure 4-5B). They are the most abundant fibrillar components of the axon. (On average, there are 3-10 times more neurofilaments than microtubules in an axon.) Neurofilaments are related to the intermediate filaments of other cell types, all of which belong to a family of proteins called cytokeratins. (Other cytokeratins include vimentin, glial fibrillary acidic protein, desmin, and keratin.) Unlike microtubules, neurofilaments are very stable and almost totally polymerized in the cell. In Alzheimer’s disease and some other degenerative disorders they become modified and form a characteristic lesion called the neurofibrillary tangle (see Chapter 58).

Microfilaments, 3-5 nm in diameter, are the thinnest of the three main types of fibers that make up the cytoskeleton (see Figure 4-5C). Like the thin filaments of muscle, microfilaments are polar polymers of globular actin monomers (each bearing an ATP or ADP) wound into a double-stranded helix. Actin is a major constituent of all cells, perhaps the most abundant animal protein in nature. Several closely related molecular forms of actin, each encoded by a different gene, have been identified: the actin of skeletal muscle, and at least two other molecular forms, β and γ. Neural actin is a mixture of the β and γ species, which differ from muscle actin at a few amino acid residues. Most of the actin molecule is highly conserved, not only in different cells of an animal but also in organisms as distantly related as humans and protozoa.
Unlike the microtubules and neurofilaments, actin filaments form short polymers: they are concentrated at the cell’s periphery in the cortical cytoplasm lying just underneath the plasmalemma, where, together with a very large number of actin-binding proteins (for example, spectrin-fodrin, ankyrin, talin, and actinin), they form a dense network. This matrix plays a key role in the dynamic function of the cell’s periphery, such as the motility of growth cones during development, generation of specialized microdomains on the cell surface, and the formation of pre- and postsynaptic morphologic specializations.

Like microtubules, microfilaments are in a dynamic state and undergo cycles of polymerization and depolymerization. At any one time about half the total actin in neurons can exist as unpolymerized monomers. The state of actin within the cell is controlled by binding proteins. These proteins facilitate assembly and block changes in polymer length by capping the rapidly growing end of the filament or by severing it. Other binding proteins cross-link or bundle microfilaments. The dynamic state of microtubules and microfilaments permit the mature neuron to retract old processes and extend new ones.
Figure 4-11 The insulating myelin sheath of the axon has regularly spaced gaps called the nodes of Ranvier. Electron micrographs show the region of nodes in axons from the peripheral nervous system, spinal cord, and cerebral cortex. The axon (Ax) runs from the top to the bottom in all three micrographs. The axon is coated with many layers of myelin (M), which is lacking at the nodes (Nd), where the axolemma (Al) is exposed. (In the peripheral nervous system the support cell responsible for myelination is called a Schwann cell (Sc), and in the central nervous system it is an oligodendrocyte.) The elements of the cytoskeleton that can be seen within the axon are microtubules (Mt) and neurofilaments (Nf). Mitochondria (Mit) are also seen. (From Peters et al. 1991.)

In addition to serving as cytoskeleton, microtubules and actin filaments act as tracks along which other organelles and proteins are driven by molecular motors. Since these filamentous polymers are polar, each motor drives its organelle cargo in one direction only. In the axon all microtubules are arranged in parallel, with the plus end pointing away from the cell body and the minus end facing the cell body. This regular orientation permits the orderly movement of distinct classes of organelles along the axon, thus maintaining the special distribution of organelles throughout the cell. In dendrites, however, microtubules with opposite polarities are mixed, and this explains why the organelles of the cell body and dendrites are similar. Actin motors, called myosins, mediate other types of cell motility, including extension of the cell’s processes. Myosin is also thought to translocate membranous organelles within the cortical cytoplasm. Actomyosin in muscle is responsible for contraction (Chapter 34).
The Neurons That Mediate the Stretch Reflex Differ in Morphology and Transmitter Substance

The relationship between neuronal structure and function can be seen by comparing the sensory and motor neurons that mediate the stretch reflex. As described in Chapter 2, the monosynaptic component of the stretch reflex is a simple two-neuron circuit consisting of large sensory neurons that receive information from muscle cells and motor neurons that cause the skeletal muscles of the limb to contract (see Figure 2-5).

The Sensory Neuron Conducts Information From the Periphery to the Central Nervous System

Sensory neurons for the stretch reflex convey information about the state of muscle contraction. Their cell bodies are round with large diameters (60-120 µm) and are located in dorsal root ganglia situated immediately adjacent to the spinal cord. At maturity these neurons possess a single axonal process that bifurcates into two branches a short distance from the cell body (Figure 4-8). The peripheral branch projects to muscle and the central branch to the spinal cord, where it forms synapses on the cell bodies and dendrites of motor neurons (Figure 4-9).

The peripheral branch of the sensory axon coils around a fine, specialized muscle fiber within the muscle spindle, a sensory receptor sensitive to stretch (Figure 4-10). The peripheral branch is 14-18 µm in diameter and is coated with an insulating sheath of myelin 8-10 µm thick. (Myelination is discussed in some detail later in the chapter.) The myelin sheath is regularly interrupted along the length of the axon. At these gaps, called nodes of Ranvier, the plasma membrane of the axon (the axolemma) is exposed to the extracellular space for about 0.5 µm (Figure 4-11). This arrangement greatly increases the speed at which the nerve impulse is conducted along the axon (in humans, 80 m/s) because the signal jumps from one unmyelinated node to the next by saltatory conduction (see Chapters 8 and 9).

The central branch of the sensory axon enters the spinal cord in the dorsal horn, where it bifurcates into branches that ascend and descend in the spinal cord. Collateral fibers from the axon form synapses on motor neurons in the ventral horn. When excited, the sensory neuron releases the excitatory amino acid neurotransmitter L-glutamate (see Chapter 15) that depolarizes the motor neurons.

The Motor Neuron Conveys Central Motor Commands to the Muscle Fiber

The axon of each sensory neuron directly contacts two classes of motor neurons: those that innervate the muscle within which the sensory ending is located (the homonymous muscle) and those that innervate other muscles that cooperate in stretching the knee joint (synergistic muscles). Both types of motor neurons are located in the ventral horn of the spinal cord. Motor neurons have large cell bodies, and their nucleus is distinctive because of its large and prominent nucleolus (see Figure 4-8B).

Unlike dorsal root ganglion cells, which have no dendrites, motor neurons have several dendritic trees that arise directly from the cell body. Each dendritic tree is complex, generated by extensive branching of primary dendrites (Figure 4-12). The total number of terminal dendritic branches per cell is often more than 100. The average length of a dendrite from the motor neuron's cell body to its end is about 20 cell-body diameters (1 mm), but some branches are twice as long. The branches project radially, so that the entire dendritic structure of a single motor neuron can extend within the spinal cord over an area about 2 to 3 mm in diameter. Such extensive dendritic structures are characteristic of central neurons, whose firing is regulated by input from many neurons. Short specialized dendritic extensions called spines serve to increase the area of the neuron available for synaptic inputs. Dendritic spines provide a biochemical and electrical compartment where incoming signals are initially received and processed; their morphology is discussed later in this chapter.
Figure 4-13 The axon of a spinal motor neuron has branches that make synaptic contact with several interneurons and, rarely, a recurrent (feedback) connection on the motor neuron.

A. An electron micrograph of a cat's spinal motor neuron shows the cell body, axon hillock (AH), initial segment (IS), and the first part of the myelinated portion of the axon. Glial cells surround the initial part of the axon. A cross-section of a capillary (C) is also visible. The inset shows two dendrites emerging from opposite sides of the cell body. (From Conradi 1969.)

B. The axons of motor neurons typically give off from one to five recurrent branches that usually make synaptic contact with inhibitory interneurons. In rare instances an axonal branch (a recurrent collateral) makes direct contact with its own cell body. (Courtesy of R. E. Burke.)

Messenger RNA is transported along dendrites and appears to be concentrated at the base of dendritic spines. Some protein synthesis occurs in dendrites, indicating that the dendrites are functional extensions of the cell body, where most proteins are synthesized. Consistent with this view, the cytoskeleton of dendrites more closely resembles that of the cell body than that of axons. Local protein synthesis at dendrites is thought to play an important role in synaptic plasticity.
A composite illustration of the rat hippocampus and dentate gyrus. A major experimental advantage of the hippocampus for neuroscience research is its highly laminar organization. A Nissl-stained section shows dark bands representing accumulations of neuronal cellbodies in the pyramidal cell layer (stratum pyramidale) of the hippocampus. The hippocampus can be divided into three separate regions—CA1, CA2, CA3—based on the size and connections of the resident pyramidal cells. Typical CA3 and CA1 pyramidal cells are drawn on the Nissl-stained section. Each cell has been traced with an intracellular marker (horseradish peroxidase or Phaseolus vulgaris leukoagglutinin) through adjacent 400 µm slices and reconstructed by computer. The CA3 cell dendrites are shown as thin lines and the axon collaterals as thicker lines. The CA3 axon collaterals innervate other CA3 cells (the associational axon collaterals) and the CA1 pyramidal cells (the Schaffer collaterals). These axons run in the stratum radiatum. Only the dendrites of the CA1 pyramidal cell are illustrated. (Courtesy of D. G. Amaral.)

B. Schematic diagram of the hippocampus showing the connection between the two pyramidal neurons through the Schaffer axon collaterals.

Each motor neuron gives rise to only oneaxon, about 20 µ m in diameter, from a specialized region of the cell body called the axon hillock. The axon hillock and the initial (unmyelinated) segment of the axon extend the length of about one cell-body diameter (Figure 4-13). About half the surface area of the axon hillock and cell body and three-quarters of the dendritic membrane are covered by synaptic boutons, the knob-like terminals of the axons of presynaptic neurons (see Figure 4-8B). The axon hillock and the initial segment of the axon function as a trigger zone, the site at which the many incoming signals from other neurons are integrated and the action potential, the output signal of the neuron, is generated (see Chapter 9).

Close to the cell body the axon gives off several recurrent collateral branches (Figure 4-13). These branches are called recurrent because many of them project back to the motor neuron and modify the activity of the cell. More often, however, recurrent collaterals form synapses on a particular type of interneuron in the spinal cord, the Renshaw cell. These interneurons hyperpolarize the motor neurons, using the neurotransmitter L-glycine, and thus inhibit firing in the motor neurons.

In addition, motor neurons receive recurrent excitatory inputs from other motor neurons, and both excitatory and inhibitory inputs from interneurons driven by descending fibers from the brain that control and coordinate movement. These synaptic inputs, together with the excitatory input from the primary sensory neurons and inhibitory input from Renshaw cells, are integrated by mechanisms that are described in Chapter 12.

A Single Motor Neuron Forms Synapses With Several Muscle Cells

One striking difference between motor and sensory neurons is the location of their synaptic inputs. The sensory neuron has few if any boutons on its cell body or along the peripheral branch of its axon. Its primary input is from sensory receptors at the terminal of the peripheral axon. In contrast, the motor neuron receives primary and modifying inputs throughout its dendrites and cell body. (Almost all presynaptic boutons on motor neurons are located on the dendritic branches; only 5% are located on the cell body.) The synapses on the motor neuron are distributed in a functional pattern. Most inhibitory synapses are on the cell body or close to it, whereas excitatory ones are located farther out along the dendrites. Inhibitory inputs are strategically placed close to the trigger zone to have maximal influence on the final tally of inputs to the neuron (see Chapter 12).
Figure 4-15 Pyramidal cells in the CA3 region of the hippocampus form synapses on the dendrites of CA1 cells in the stratum radiatum.

**Left**: Micrograph of a Golgi-stained CA1 pyramidal cell is shown with dendrites extending downward 350 µm into the stratum radiatum.

**Right**: Three micrographs show synapses formed on this CA1 cell by CA3 cells. **A.** Axons of two CA3 neurons form synapses on a dendrite 50 µm from the CA1 neuron’s cell body. **B.** A single CA3 axon forms synapses on dendrites 259 µm from the cell body. **C.** A single CA3 axon forms synapses on two dendrites 263 µm from the cell body. (From Sorra and Harris 1993.)

The information flow from sensory neurons to motor neurons is both divergent and convergent. Each sensory neuron contacts 500–1000 motor neurons and typically forms two to six synapses on a single motor neuron (divergence of information). At the same time each motor neuron receives input from many sensory neurons (convergence of information); inputs from more than 100 sensory neurons are needed for a motor neuron to reach the threshold for firing.

The axons that mediate the stretch reflex in the leg leave the lumbosacral region of the spinal cord and join the femoral nerve. (The motor axons and sensory fibers travel along the same peripheral path to the muscle.)

When the motor neuron enters the muscle it ramifies into many unmyelinated branches, each with a diameter of only a few micrometers. These terminal fibers run along the surface of a muscle fiber and form many synaptic contacts called neuromuscular junctions. These synapses are the most completely characterized and best understood of all synapses in the nervous system (see Chapter 11).
Figure 4-16 The dendrites of pyramidal cells in the CA1 region of the hippocampus bear a variety of spines.

**Left:** The diversity of dendritic spine shapes is evident along even a short segment of the mature dendrite in this three-dimensional reconstruction from a series of electron micrographs. (From Harris and Stevens 1989.)

**Right:** Three micrographs illustrate the details of different types of dendritic spines. **A.** A thin dendritic spine from the postnatal day-15 rat hippocampus. The postsynaptic density shows as the thickened receptive surface (**open arrow**) located across from the presynaptic axon, which has round clear vesicles. **B.** Stubby spines containing postsynaptic densities (**open arrow**) are both small and rare in the mature hippocampus. Their larger counterparts (not shown) predominate in the immature brain. **C.** Mushroom-shaped spines have a larger head. These spines are present by day 15 as shown here. The immature spines contain flat cisternae of smooth endoplasmic reticulum, some with a beaded appearance (**bd**). Synapse with postsynaptic density is indicated by the **open arrow**. Branched spines did not occur in this dendritic segment. (From Harris et al. 1992.)
The axons of both central and peripheral neurons are insulated by a myelin sheath.

A. An axon in the central nervous system receives its myelin sheath from an oligodendrocyte. (Adapted from Bunge 1968.)

B. An electron micrograph of a transverse section through an axon (Ax) in the sciatic nerve of a mouse. The spiraling lamellae of the myelin sheath (Ml) start at a structure called the inner mesaxon (IM; circled). The spiraling sheath is still developing and is seen arising from the surface membrane (SM) of the Schwann cell, which is continuous with the outer mesaxon (OM; circled). The Schwann cell cytoplasm (Sc Cyt) is still present, next to the axon; eventually it is squeezed out and the sheath becomes compact. (From Dyck et al. 1984.)

C. A peripheral nerve fiber is myelinated by a Schwann cell. (Adapted from Williams et al. 1989.)

Each muscle fiber is contacted by only a single axon, but a single motor axon innervates several muscle fibers. The axon and the muscle fibers it innervates constitute a motor unit. The muscle fibers innervated by any one motor axon are widely spread, overlapping muscle fibers of other motor units. The number of muscle fibers innervated by a single motor axon varies throughout the body, depending on the mass of the body part to be moved. Thus, in the leg a single motor axon innervates more than 1000 muscle fibers, while in the eye an axon contacts fewer than 100 muscle fibers. A lower innervation ratio permits greater precision of movement control.

The sensory and motor neurons that mediate the stretch reflex differ in appearance, location in the nervous system, the distribution of their axons and dendrites, and the inputs they receive. All of these cytological features have important behavioral consequences. In addition, the two types of cells differ biochemically because they use different neurotransmitters (although both transmitters are excitatory). For example, the motor neuron, which uses acetylcholine as a transmitter, requires a set of macromolecules that includes the biosynthetic enzyme choline acetyltransferase and a specific membrane transporter for choline, an essential precursor in the synthesis of acetylcholine (Chapter 15).

**Box 4-1 Defects in Myelin Proteins Disrupt Conduction of Nerve Signals**

Because normal conduction of the nerve impulse depends on the insulating properties of the myelin sheath surrounding the axon, defective myelin can result in severe disturbances of motor and sensory function. Myelin in both the central and peripheral nervous systems contains a major class of proteins, myelin basic proteins (MBP), which have an important role in myelin compaction. At least seven related proteins are produced from a single MBP gene by alternative RNA splicing.

Myelin basic proteins are capable of eliciting a strong immune response. When injected into animals they cause experimental allergic encephalomyelitis, a syndrome characterized by local inflammation and by destruction of the myelin sheaths (demyelination) in the central nervous system. This experimental disease has been used as a model for multiple sclerosis, a common demyelinating disease in humans. Because demyelination slows down conduction of the action potential in the affected neurons' processes, multiple sclerosis and other demyelinating diseases (for example, Guillain-Barré syndrome) can have devastating effects on the function of neuronal circuits in the brain and spinal cord (see Chapter 35).
Many diseases that affect myelin, including some animal models of demyelinating disease, have a genetic basis. The shiverer (or shi) mutant mice have tremors and frequent convulsions and tend to die at young ages. In these mice the myelination of axons in the central nervous system is greatly deficient and the myelination that does occur is abnormal. The mutation that causes this disease is a deletion of five of the six exons of the gene for myelin basic protein, which in the mouse is located on chromosome 18. The mutation is recessive; a mouse will develop the disease only if it has inherited the defective gene from both parents. Shiverer mice that inherit both defective genes have only about 10% of the myelin basic protein found in normal mice.

When the wild-type gene is injected into fertilized eggs of the shiverer mutant with the aim of rescuing the mutant, the resulting transgenic mice express the wild-type gene but produce only 20% of the normal amounts of myelin basic proteins. Nevertheless, myelination of central neurons in the transgenic mice is much improved. Although they still have occasional tremors, the transgenic mice do not have convulsions and live a normal life span (Figure 4-18).

Central and peripheral myelin also contain a distinct protein termed myelin-associated glycoprotein (MAG). MAG belongs to a superfamily that is related to the immunoglobulins and includes several important cell surface proteins thought to be involved in cell-to-cell recognition (for example, the major histocompatibility complex of antigens, T-cell surface antigens, and the neural cell adhesion molecule or NCAM). MAG is expressed by Schwann cells early during peripheral myelination and eventually becomes a component of mature (compact) myelin. It is situated primarily at the margin of the mature myelin sheath just adjacent to the axon. Its early expression, subcellular location, and structural similarity to other surface recognition proteins suggest that it is an adhesion molecule important for the initiation of the myelination process. Two isoforms of MAG are produced from a single gene through alternative RNA splicing.

![Figure 4-18](image)

**Figure 4-18** A genetic disorder of myelination in mice (shiverer mutant) can be partially cured by transfection of the normal gene that encodes myelin basic protein.

A. Electron micrographs show the state of myelination in the optic nerve of a normal mouse, a shiverer mutant, and a mutant transfected with the gene for myelin basic protein. (From Readhead et al. 1987.)

B. Myelination is incomplete in the shiverer mutant. As a result, the shiverer mutant exhibits poor posture and weakness. Injection of the wild-type gene into the fertilized egg of the mutant improves myelination. A normal mouse and a transfected shiverer mutant look perky.

More than half of the total protein in central myelin is a characteristic proteolipid, PLP, which has five membrane-spanning domains. Proteolipids differ from lipoproteins in that they are insoluble in water. Proteolipids are soluble only in organic solvents because they contain long chains of fatty acids that are covalently linked to amino acid residues throughout the proteolipid molecule. In contrast, lipoproteins are noncovalent complexes of proteins with lipids so structured that many serve as soluble carriers of the lipid moiety in the blood.
Many mutations of the proteolipid PLP are known, in humans as well as in other mammals (for example, the jimpy mouse). Pelizaeus-Merzbacher disease, a heterogeneous X-linked disease in humans, results from a PLP mutation. Almost all of these mutations occur in a membrane-spanning domain of the molecule. All of these mutant animals have reduced amounts of the mutated protein and show hypomyelination and degeneration and death of oligodendrocytes. These observations suggest that the proteolipid is involved in the compaction of myelin.

The major protein in mature peripheral myelin, myelin protein zero (MPZ or P0), spans the plasmalemma of the Schwann cell. It has a basic intracellular domain and, like myelin-associated glycoprotein, is a member of the immunoglobulin superfamily. The glycosylated extracellular part of the protein, which contains the immunoglobulin domain, functions as a homophilic adhesion protein during myelin spiraling and compaction by interacting with identical domains on the surface of the opposed membrane. Genetically engineered P0 mice in which the function of myelin protein P0 has been eliminated have poor motor coordination, tremors, and occasional convulsions.

Observation of trembler mouse mutants led to the identification of peripheral myelin protein 22 (PMP22). This Schwann cell protein spans the membrane four times and is normally present in compact myelin. PMP22 is altered by a single amino acid. A similar protein is found in humans, encoded by a gene on chromosome 17.

Although several hereditary peripheral neuropathies result from mutations of the PMP22 gene on chromosome 17, one form of Charcot-Marie-Tooth disease is caused by the DNA duplication of this gene (Figure 4-19). Charcot-Marie-Tooth disease, the most common inherited peripheral neuropathy, is characterized by progressive muscle weakness, greatly decreased conduction in peripheral nerves, and cycles of demyelination and remyelination. Since both duplicated genes are active, the disease results from increased production of PMP22 (a two- to three-fold increase in gene dosage) rather than from a reduction in a mutant protein.

![Figure 4-19 Charcot-Marie-Tooth disease (type 1A) results from gene dosage effects.](image)

**A.** A patient with Charcot-Marie-Tooth shows impaired gait and deformities (from Charcot’s original description of the disease, 1886).

**B.** Sural nerve biopsies from a normal individual (from AP Hays, Columbia University) and from a patient with Charcot-Marie-Tooth (from Lupski and Garcia 1993).

**C.** The disordered myelination in Charcot-Marie-Tooth disease results from the increased production of the peripheral myelin protein PMP22. The increase is caused by a duplication of a normal 1.5 megabase region of the DNA on the short arm of chromosome 17 at 17p11.2-p12. The PMP22 gene is flanked by two similar repeat sequences, as shown in the representation of a normal chromosome 17. Normal individuals have two normal chromosomes. In patients with the disease the duplication results in two functioning PMP22 genes, each flanked by the repeat sequence. The normal and duplicated regions are shown in the expanded diagrams indicated by the dashed lines. (The repeats are thought to have given rise to the original duplication, which was then inherited. The presence of two similar flanking sequences with homology to a transposable element is believed to increase the frequency of unequal crossing-over in this region of chromosome 17 because the repeats enhance the probability of mispairing of the two parental chromosomes in a fertilized egg.)

**D-E.** Although a large duplication, 3 megabases cannot be detected in routine examination of chromosomes in the light microscope, but microscopic evidence for the duplication can be obtained using fluorescence in situ hybridization. With this technique, the PMP22 gene is detected with an oligonucleotide probe tagged with the dye Texas Red. An oligonucleotid probe tagged with fluorescein, a green fluorescent dye that hybridizes with DNA from region 11.2 (indicated in green closer to the centromere), is used for in situ hybridization on the same sample. A nucleus from a normal individual (D)
Pyramidal Neurons in the Cerebral Cortex Have More Extensive Dendritic Trees Than Spinal Motor Neurons

Whereas motor neurons are the major excitatory projection neurons of the spinal cord, pyramidal cells are the excitatory projection neurons in the cerebral cortex. Pyramidal cells in different cortical regions are morphologically similar and use L-glutamate as a transmitter. We shall focus here on the pyramidal cells of the hippocampus, a structure important for memory storage.

The hippocampus is divided into two major regions, CA3 and CA1. In both regions the cell bodies of pyramidal cells are situated in a single continuous layer, the stratum pyramidale (Figure 4-14). In contrast to the motor neurons of the spinal cord, pyramidal cells have not one but two dendritic trees, and these emerge from opposite sides of the cell body: the basal dendrites arise from the side that gives rise to the axon, and the apical dendrites arise from the opposite side of the cell body.

Excitatory input to CA1 pyramidal neurons is extensive. About 5000 CA3 pyramidal cell axons—comprising the Schaffer collateral pathway—converge on a single CA1 cell. These Schaffer collaterals form synapses at all levels of the CA1 cell’s dendritic tree close to the cell body and at more distant levels (Figure 4-15). The connections formed by the Schaffer collaterals are called en passant synapses because CA3 axons continue to pass through the stratum radiatum, making contact with the dendrites of many other CA1 pyramidal cells.

Most of the synapses are made on dendritic spines. In many parts of the brain, spines have two inputs, one excitatory and the other inhibitory. In area CA1, however, each pyramidal cell spine has only one synapse, which is excitatory. These spines have four principal shapes: thin, mushroom, branched, and stubby (Figure 4-16). The neck of the spine restricts diffusion between the head of the spine and the rest of the dendrite. Thus, each spine may function as a separate biochemical region. As we shall see later, this compartmentalization may be important for selectively altering the strength of synaptic connections during learning and memory.

Glia Cells Produce the Insulating Myelin Sheath Around Signal-Conducting Axons

The signal-conducting axons of both sensory and motor neurons are ensheathed in myelin along most of their length (see Figure 4-11). Acting as insulation, myelin speeds transmission along axons and thus is critical for quick reflex movements like the knee jerk. The myelin sheath is arranged in concentric bimolecular layers of lipids interspersed between protein layers (Figure 4-17). Biochemical analysis shows that myelin has a composition similar to that of plasma membranes, consisting of 70% lipid and 30% protein, with a high concentration of cholesterol and phospholipid.

Both the regular lamellar structure and biochemical composition of the myelin sheath are consequences of how myelin is formed from plasma membrane. In the development of the peripheral nervous system, before myelination takes place, the sensory cell axon lies along a peripheral nerve in a trough formed by a class of glia called Schwann cells. Schwann cells line up along the axon at intervals that will eventually become the nodes of Ranvier. The external cell membrane of each Schwann cell surrounds a single axon and forms a double-membrane structure called the mesaxon, which elongates and spirals around the axon in concentric layers (Figure 4-17C). The cytoplasm of the Schwann cell appears to be squeezed out during the ensheathing process when the Schwann cell’s processes condense into the compact lamellae of the mature myelin sheath.

In the femoral nerve, which carries the sensory and motor axons that mediate the stretch reflex, the primary sensory axon is about 0.5 m long and the internodal distance is 1-1.5 mm; thus approximately 300-500 nodes of Ranvier occur along a primary afferent fiber between the thigh muscle and the dorsal root ganglion, where the cell body lies. Since each internodal segment is formed by a single Schwann cell, as many as 500 Schwann cells participate in the myelination of a single peripheral sensory axon.

In the central nervous system myelination of the central branch of dorsal root ganglion cell axons and the axons of motor neurons differs somewhat from myelination in the peripheral system. The glial cell responsible for elaborating central myelin is the oligodendrocyte, which typically ensheathes several axon processes. Schwann cells and oligodendrocytes differ developmentally and biochemically. The expression of myelin genes by Schwann cells in the peripheral nervous system is regulated by the contact between the axon and the myelinating Schwann cell. In contrast, the expression of myelin genes by oligodendrocytes in the central nervous system appears to depend on the presence of astrocytes, the other major type of glial cell in the central nervous system.

Specific diseases can arise from dysfunction of the specialized properties of neurons. In particular, defective myelination of the axon produces severe disturbances of motor and sensory function. Thus, understanding the biochemistry of myelin formation provides important insight into the basis of certain neurological diseases (Box 4-1).

An Overall View

Nerve cells have four distinctive compartments: dendrites, for receiving signals from other neurons; the cell body, which contains the DNA encoding neuronal proteins and the complex apparatus for synthesizing them; the axon, which projects over long distances to target cells (for example, other neurons or muscle); and nerve terminals, for release of neurotransmitters at synapses with targets.

In this chapter we have illustrated this basic cellular plan by describing three types of neurons. Although all of these cells conform to a basic plan, each type differs considerably, most obviously by location in the nervous system—peripheral or central, spinal cord, or brain. They also differ in the location of synaptic inputs on the cell and in the types of target cells to which they project. Furthermore, they differ in cell body size and shape, distribution of their dendritic trees and number of axon branches, and in their degree of myelination. Biochemically, they differ most obviously in transmitter type, and, as we shall see throughout this book, in many other constituents (for example, in the enzymes that synthesize neurotransmitters, the pumps that exchange ions or recapture neurotransmitter substances, and the receptors that transduce physical or biochemical inputs).

The functional significance of many morphological differences is plainly evident. For example, the dorsal root sensory neuron must extend a process in the peripheral nervous system, as must the spinal motor neuron. It also is clear why the motor neuron has a more complex dendritic tree than the sensory neuron: Even simple reflex activity requires coordination of inputs, both excitatory and inhibitory, to regulate specific motor units, and purposeful movements need still more integration because of inputs from the brain.

The functional significance of some other cytological differences is not so obvious, but can be understood in the context of the electrophysiological activities of the particular neurons. Thus the large number of dendrites and axonal branches in cortical pyramidal neurons must contribute to the complexity of information processing in the brain.

Selected Readings


References


