3 A HISTORY OF ANIMAL MODELS

> For a large number of problems there will be some animal of choice or a few such animals on which it can be most conveniently studied. . . . I have no doubt that there is quite a number of animals which are . . . "created" for special physiological purposes, but I am afraid that most of them are unknown to the men for whom they are "created" and we must apply to the zoologists to find them and to lay our hands on them.
>
> —KROGH (1929), PP. 202–203

The menagerie of animals that biologists have used in their research is a highly nonrandom set. More than a million animal species inhabit the earth, but the vast majority of animal research is concentrated on just a dozen or so species. This taxonomic bias is present in most areas of biological research (Troudet et al., 2017; Rosenthal et al., 2017), but it is especially severe in biomedical research, where the terms "model species" and "model organism" are encountered frequently (Fields & Johnston, 2005; Katz 2016). It is difficult to quantify how frequently researchers employ the various species in their studies, but one can attempt to do so by examining either their publications or regulatory reports (figure 3.1). Both approaches have serious limitations. For one thing, many publications do not mention the names of the examined species in their title (nor, quite frequently, their abstract). For another, the United States does not report animal numbers for the most commonly studied species: mice, rats, birds, and all cold-blooded animals (see chapter 2). Nonetheless, the available data clearly show that a relatively small number of species account for most of the research (see also figure 2.1). They also show that the popularity of the various species in biological research, and the financial support for such research (Farris, 2020), has varied over time. Mice, for example, have become vastly more popular among biologists since the 1960s, while rabbits, hamsters, and guinea pigs have become less frequently studied.

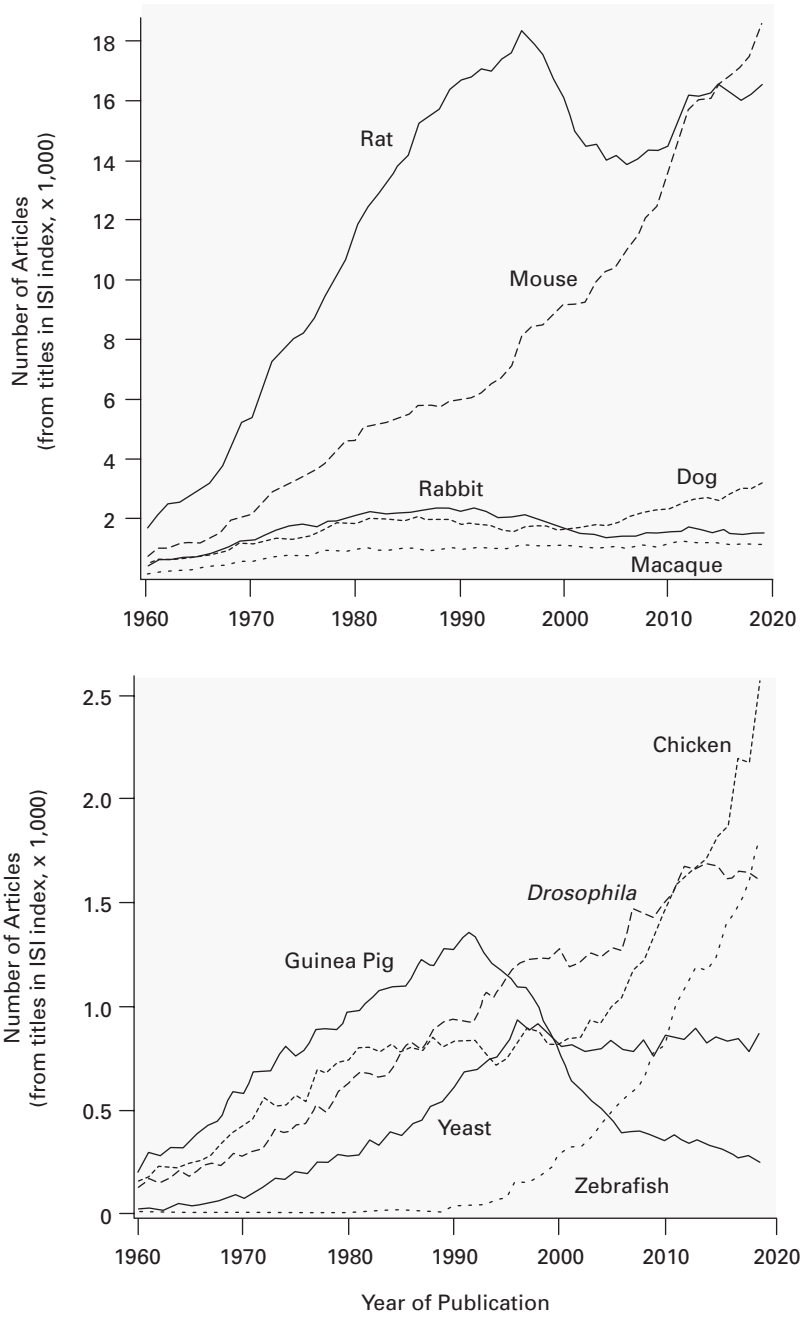


Figure 3.1

Relative distribution of model species featured in journal publications. These graphs indicate the frequency with which several frequently studied species were mentioned in the titles of journal articles (as indexed in the ISI Citation Index), broken down by year of publication. Note the difference in *y*-axis scales. The data were gathered by the author using procedures analogous to those described in Dietrich et al. (2014) but include additional years and species.

All of which raises interesting questions about how and why biologists select their research species. Why did some species become extremely popular among biologists, and why did some of them eventually fade? I here explore these questions in the hope that this exercise may bring to light some general principles of model selection and help inform some model choices going forward. I focus here exclusively on multicellular animal models, but unicellular organisms (notably bacteria and yeast) are covered in chapter 4, together with other in vitro models. The chapter begins with the early history of animal models in biology and then covers the major model species used today, emphasizing their principal contributions to biomedicine as well as their limitations. Due to space constraints, relatively little attention is paid to the biologists driving this research, but this history is well covered in other books (Kohler, 1994; Rader, 2004; Finger, 2004; Endersby, 2007).

3.1 ECLECTIC BEGINNINGS OF BIOLOGY

Early biologists studied a wide variety of species. This was true especially in ancient Greece and Rome, but it remained the case through the rise of experimental physiology in 19th century Europe. Many of those early studies used domesticated animals, including dogs, horses, rabbits, and guinea pigs, but biologists back then usually examined many different species to identify those that were optimally suited to the problem at hand (Logan & Brauckmann, 2015). This approach was championed most effectively by August Krogh, whom we discussed in chapter 2 (section 2.6) and who gave us the opening quotation of the present chapter; but let us start farther back in time.

3.1.1 Ancient Greece and Rome

Aristotle, who lived in Greece during the 4th century B.C., was one of the world's greatest philosophers and the founding father of comparative biology. He examined more than 500 species of animals, including numerous invertebrates (Blits, 1999) and then looked for patterns of similarity and difference. Aristotle was particularly interested in the distinctions between different types of animals and in the various categories of causes that could explain those differences. Thus, Aristotle constructed the first systematic framework for comparative anatomy and a philosophy for explaining nature's diversity.

Roughly 500 years later, Galen made comparative anatomy relevant to human medicine and became Rome's greatest physician. Because of cultural prohibitions against touching human corpses, Galen did not dissect human beings. He did, however, dissect a great number of different animals, including macaques, cats, dogs, weasels, camels, lions, wolves, deer, bears, mice, and even an elephant (Finger, 2004). He also experimented

on diverse living animals, cutting various nerves or parts of the spinal cord to examine their functions, even though surgical anesthetics were not yet available. Collectively, Galen's dissections and vivisections allowed him to make numerous discoveries, which he then extrapolated to humans.

Galen's teachings formed the foundation of medicine for many centuries, but his accounts of the human body were challenged in the first half of the 16th century by Andreas Vesalius, who dissected human cadavers as well as diverse nonhumans. At some point in 1540, Vesalius compared the skeleton of a human and a macaque monkey side by side and suddenly realized that Galen had described the monkey skeleton quite precisely, but had mistakenly assumed that the human skeleton would be the same. Vesalius concluded that Galen's lack of direct experience with human bodies caused him to make numerous unwarranted extrapolations from monkeys to humans. Overall, Vesalius found more than 200 such mistakes in Galen's body of work (Finger, 2004). One famous example was Galen's description in humans of a *rete mirabile*, which is a network of small blood vessels at the base of the brain. Vesalius pointed out that such a *rete mirabile* exists in several mammals that Galen had dissected, such as oxen, but is absent from humans.

3.1.2 The Rise of Experimental Physiology

Galen and Vesalius had done some experiments on live animals, but most of their research was anatomical. Eventually, however, experimental physiology became more prominent. Especially influential was William Harvey's work on the "motion of the heart and blood," which was originally published in 1628 (Harvey, 1628/1928). Harvey carefully studied the anatomy of the circulatory system in more than 100 species (Cole, 1957) but also performed a large number of experiments on living animals (i.e., vivisections), asking, for example, how contractions of specific chambers of the heart correlate with arterial pulsations. By studying a large number of different species, Harvey was able to infer general structure-function principles about how blood is pumped through the body. He fully recognized that these principles are instantiated differently in different species. This is an important point, to which we shall return in chapter 7 (section 7.1.4).

Experimental physiology expanded rapidly toward the end of the 19th century, shortly after the development of ether as a surgical anesthetic (Wood Library Museum of Anesthesiology, 2021) and in tandem with the invention of more sophisticated research equipment. Many of these physiological studies used cold-blooded vertebrates, especially frogs. A major advantage of these species was that they stayed alive longer than warm-blooded animals during invasive procedures. Even isolated nerves and muscles continued to function for hours after the frogs had died. In addition to

cold-blooded vertebrates, experimental physiologists in the late 1800s worked with a wide variety of warm-blooded animals, including pigeons, rabbits, and dogs. As Cheryl Logan pointed out in her thorough review of published studies from this period, the early physiologists "attempted to include as many animals as possible" in their research, and for all of them "generality was the goal" (2002, p. 347).

One of the hot topics during that time was the question of functional localization within the cerebral cortex. Some physiologists had stimulated or destroyed specific brain regions in various animals and found that the behavioral effects varied across cortical areas (Young, 1990). Others argued, instead, that functions were more globally distributed across the cerebral cortex. Ultimately, the proponents of functional localization prevailed, but resolution of the controversy required a large number of animal experiments that surely induced a fair amount of animal suffering. Many of the experiments were performed on rabbits, dogs, or monkeys, and some of them were demonstrated during public lectures to convince a skeptical audience. In retrospect, it is not surprising that these experiments spawned anti-vivisectionist sentiment (especially toward experiments on unanesthetized animals) and the emergence of animal welfare regulation.

The principal motivation underlying this research had been to discover general structure-function principles at work in vertebrate brains, rather than to solve a specific human medical problem. However, the principle of functional localization is applicable also to human brains, as first shown by Paul Broca and Carl Wernicke with regard to language (see Finger, 2004) and later extended to other brain functions. Ultimately, this knowledge became clinically relevant because clinicians often use it, for example, to assess brain damage after stroke and guide brain cancer surgery.

3.2 EARLY EXPERIMENTAL MEDICINE: RESEARCH ON PET SPECIES

Because experimental biology in the late 1800s and early 1900s led to several major medical advances, especially with regard to the treatment of infectious diseases (de Kruif, 1926), the number of experiments on living animals increased exponentially during this period. In the United Kingdom, 250 vivisections were performed in 1880, 17,000 in 1900, 90,000 in 1920, and 700,000 in 1940 (French, 1975). It is unclear what fraction of these studies were performed on anesthetized animals, but after the turn of the century the use of anesthetics for invasive procedures on animals became quite widely accepted (Franco, 2013).

Also of interest is that the dramatic rise in animal experiments during this period involved mainly research on mammals, especially guinea pigs, rabbits, and dogs. Not coincidentally, these species were widely kept as pets.

3.2.1 Guinea Pigs

Guinea pigs are neither pigs nor from Guinea. Instead, they are small rodents native to South America. Guinea pigs were domesticated more than 3,000 years ago by the Incas and became popular as pets in Europe during the 16th century. During the 19th century, guinea pigs became popular experimental animals because they were docile and readily available through the pet trade. Moreover, guinea pigs reproduce readily and rapidly in laboratory conditions, becoming sexually mature at three months and able to bear litters every two to three months after that.

Louis Pasteur used guinea pigs (as well as rabbits and dogs) to develop some of the first vaccines in the 1870s (see section 3.2.2 and chapter 5), and Robert Koch a few years later used guinea pigs to isolate the bacteria that cause tuberculosis. Curiously, Koch had initially injected dogs and mice with tuberculosis-causing extracts, but these animals did not get sick. By contrast, guinea pigs were similar to humans in their sensitivity to the tuberculosis bacterium. Also interesting is that Koch claimed to have developed an antitoxin therapy against tuberculosis, which ultimately proved ineffective. His general approach did, however, bear fruit in 1890, when van Behring and Shibasaburo discovered that infected guinea pigs produced antitoxins (later called antibodies; see chapter 5) that were effective at treating and preventing diphtheria when injected into other guinea pigs.

Another line of research that depended heavily on guinea pigs focused on nutritional deficiencies, notably scurvy. It had long been known that sailors on long voyages developed scurvy unless they supplemented their bland diet with citrus fruit, but the specific cause of the disease remained unclear. Researchers had tried to induce scurvy in pigeons and rats by feeding them pure grain diets, but these animals did not get sick. In contrast, impoverished diets rapidly induced scurvy in guinea pigs (Holst & Frölich, 1907), which is why these animals were used in subsequent experiments that led to the discovery of vitamin C (Endersby, 2007). We now know that this species difference in disease susceptibility arises because pigeons and rats can synthesize their own vitamin C, whereas humans and guinea pigs must obtain it from food.

Guinea pigs also played a role in the 1949 discovery of lithium for the treatment of bipolar disorder (see chapter 6) and in a few additional research areas. The use of guinea pigs as "guinea pigs" in biomedical research continued to increase after 1960 but then decreased over the last 30 years (figure 3.1). The most likely explanation for this decline is the increased reliance on mice over this same period. In addition, guinea pigs are difficult to anesthetize (e.g., compared with dogs, rabbits, and rats) (Brodbelt et al., 2008), and efforts to create transgenic guinea pigs are still nascent. Aside from such technical limitations, many biologists may have turned away from guinea pigs

simply because they are so endearing and popular as pets; in contrast, mice and rats elicit weaker ethical concerns.

3.2.2 Rabbits

Rabbits were domesticated thousands of years ago, largely for food, but the details of that process remain unclear. By the 16th century humans had created several distinct breeds of rabbit, and by the 19th century rabbits had become popular also as pets and research animals. Their popularity among experimental biologists probably derived mainly from their notorious proclivity for rapid reproduction (females reach sexual maturity in about six months and can then bear around 12 offspring every month). In addition, rabbits are larger than guinea pigs, which makes it easier to perform surgeries on them and draw substantial amounts of blood.

One important early use of rabbits in biomedical research was the development of a rabies vaccine. Pierre-Victor Galtier in 1879 infected rabbits with ground-up neural tissue from rabid dogs and observed that the rabbits fell ill quite rapidly, roughly twice as fast as similarly injected dogs ("Pierre-Victor Galtier," 2020). Louis Pasteur and his collaborators took this idea further by taking spinal cords from the infected rabbits and exposing them to dry air for several days, which made the virus progressively less virulent (Rappuoli, 2014). When extracts of these dried spinal cords were then injected into dogs, it protected them from later rabies infections and could even stop ongoing infections. Seeing the method work in dogs, Pasteur in 1885 tested his vaccine on a boy who had been bitten by a rabid dog, and the boy lived. Although Pasteur had previously developed vaccines for several other diseases, notably swine flu, cattle anthrax, and chicken cholera (Geison, 1995), the antirabies vaccine was his first vaccine directed at a human disease.

Syphilis is another disease that rabbits helped to treat. In the early 1900s, this devastating sexually transmitted disease affected more than 60 out of every 100,000 persons and accounted for more than 10% of the patients in some insane asylums. In 1905 biologists discovered that syphilis is caused by a bacterium called *Treponema pallidum*. It also became clear around that time that rabbits are more susceptible to this microbe than other nonprimates (Esteves et al., 2018). Paul Ehrlich and his Japanese student Sahachiro Hata then infected countless rabbits with syphilis and tested hundreds of arsenic-related compounds in the hope of finding a treatment (Frith, 2012). They eventually succeeded in 1909 with compound #606 and in 1910 with Salvarsan, which became known as the first "magic bullet" drug for its effectiveness and enormous impact. These drugs required multiple injections over more than a year, but they worked quite well. They became obsolete only in 1943, when it was discovered that penicillin, the world's first antibiotic (see chapter 5), effectively cures syphilis.

Rabbits have continued to be used in diverse research areas, especially in immunology. Many studies have described basic aspects of the immune system in rabbits, revealing a complex pattern of similarities and differences among rabbits, humans, and other mammals (Flajnik, 2002; Haley, 2003; Pinheiro, et al., 2016). Rabbits have also been used extensively in the production of polyclonal antibodies, mainly because their size allows investigators to collect ample amounts of blood. These antibodies were then used in many different research applications, such as immunohistochemistry and Western blot analyses.

In addition, rabbits were heavily used in toxicology, especially in studies using the Draize test, which was developed to predict whether novel cosmetic compounds might cause skin or eye irritation in humans (Wilhelmus, 2001). After both scientists and animal rights advocates pointed out that the Draize test could be replaced with in vitro studies (see chapter 4), the US Food and Drug Administration in 1981 began to accept validated alternatives to the Draize test. These newer models cannot mimic human eyes or skin perfectly, but neither can rabbits (Verstraelen & Van Rompay, 2018). As a consequence of these developments, eye irritation testing in rabbits decreased by 87% between 1982 and 1991. By 2011 rabbits accounted for less than 4% of all animal research procedures in the United Kingdom (van der Staay et al., 2017). In the United States as well, the use of rabbits in research has declined significantly over the last 20 to 30 years (figure 3.1).

3.2.3 Dogs

As noted in the section 3.1, dogs were used quite frequently in the early days of experimental physiology and medicine. Not yet mentioned was that dogs were often used to develop and learn new surgical techniques, which are more difficult to execute in smaller animals. They were also frequently employed in studies of the circulatory system because their heart and major arteries are relatively large and anatomically similar to those of humans. An early example of this work was the use of dogs to develop a hemodialysis machine that filtered nitrogenous waste out of the blood, which later became routine therapy for kidney failure (Abel et al., 1914). Another good example is Vivien Thomas's development of a treatment for blue baby syndrome, which is chronicled in the movie *Something the Lord Made* (Sargent, 2004). Thomas himself reported having used around 200 dogs over several months, first to induce the syndrome in his animals and then to correct the problem (Smith, 2013).

Given that dogs are widely considered "man's best friend," it is not surprising that their use as research animals engendered significant anti-vivisectionist sentiment. Research on dogs was at least partly responsible for passage of the 1876 Cruelty to Animals Act in the United Kingdom. In the United States, passage of the Animal

Welfare Act in 1966 was motivated by a photographic essay in *Life* magazine titled "Concentration Camps for Dogs" (Wayman, 1966). This essay documented deplorable conditions at dubious dog dealers who sold animals to research institutions. The public was most concerned that some of these animals might have been lost or stolen pets. In response, the Animal Welfare Act gave special protections to dogs (as well as cats). Nowadays, the vast majority of dogs used in research are bred specifically for that purpose, either by specialized (class A) dog dealers or in colonies at research institutions. Most of these dogs are beagles, which were selected as the preferred breed primarily because they are relatively small, docile, and social.

Research publications on dogs have become more frequent over the last 20 years (figure 3.1A), but 36% of these publications are veterinary reports that examined dogs for their own sake, rather than as models for other species. By contrast, the number of dogs used at research institutions has remained relatively steady over the last decade or two, both in the United States and the United Kingdom (figure 3.2). Most of these animals were used in mandatory safety tests for "new chemical entities," such as putative new therapeutic drugs. These safety tests generally require that the novel compound be tested in two different types of animals; the first is usually a rodent species, and the second one is often a dog. That said, the ability of studies in dogs to predict human toxicity is imperfect. For example, chocolate is far more toxic to dogs (and cats) than to humans. Overall, the concordance of toxicity results between humans and dogs has been estimated at 63% (versus 43% for rodents), but even this finding remains controversial (Olson et al., 2000; Matthews 2008). Similarly controversial is whether any nonanimal replacements can yield better predictions of human toxicity for novel compounds (see chapter 4, section 4.3).

Due to space constraints, I do not here discuss research on cats except to note that, overall, they have served as research animals less frequently than dogs (Institute of Medicine & National Research Council, 2012), yet contributed significantly to Nobel Prize–winning research on the brain and spinal cord during the 1950s and early 1960s (Eccles et al., 1962; Lienhard, 2017, 2018).

3.3 FRUIT FLIES: THE FIRST SUPERMODEL

Guinea pigs, rabbits, and dogs were widely used as animal models, but the fruit fly *Drosophila melanogaster* became the world's first "supermodel," which Rowland Davis defined as "an organism that reveals and integrates many and diverse biological findings applying to most living things or to most members of a kingdom" (2003, p. 199). This definition is flexible, but the term is certainly catchy (it also clearly applies to *Escherichia coli* and yeast, which we discuss in the next chapter). *Drosophila* are called

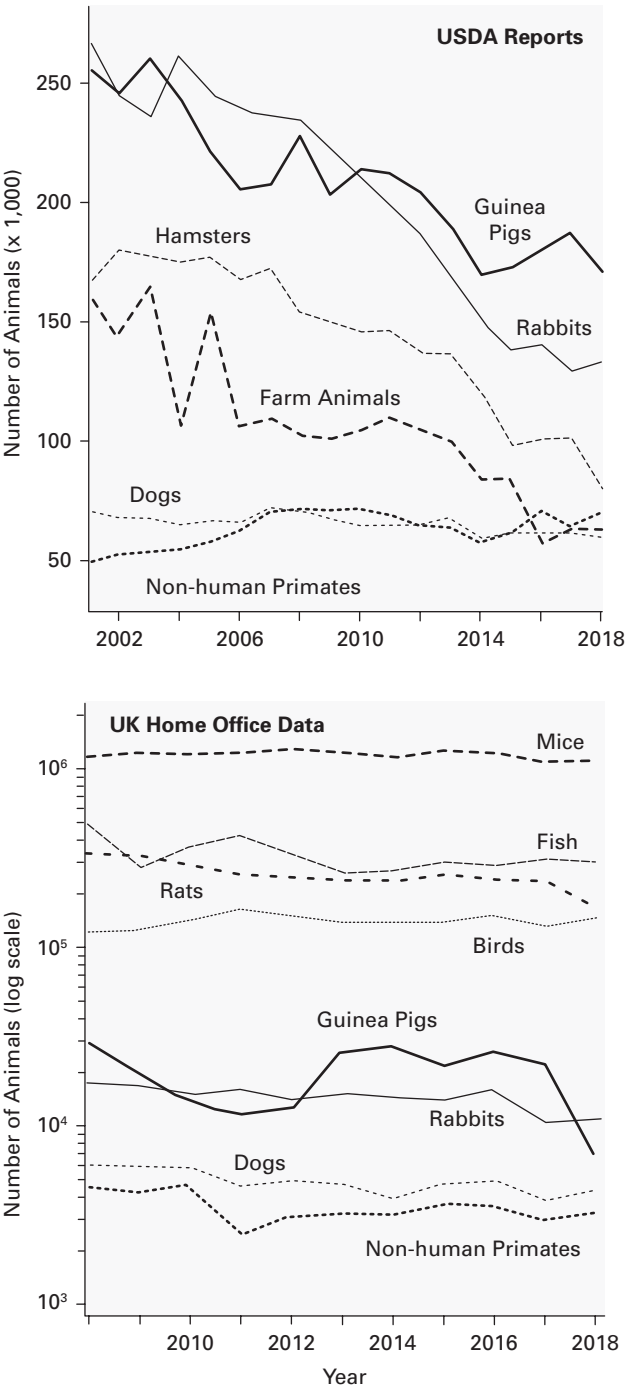


Figure 3.2

Numbers of animals used for research in the United States and the United Kingdom. Shown on the top are the numbers of animals, by species or kind, reported by the US Department of Agriculture as having been used for research between 2001 and 2018. Note that mice, rats, fish, and birds are not counted by the USDA and that pigs and sheep are here lumped with other farm animals. Shown on the bottom are data on equivalent animal numbers reported by the UK Home Office. These data include mice, rats, fish, and birds (which are considered animals in the UK; see chapter 2).

fruit flies because they like to feed on the yeast that grows on fruits. Their ascension to supermodel status began with the work of Thomas H. Morgan and his collaborators in the first few decades of the 20th century (Kohler, 1994; Endersby, 2007).

Morgan had been interested in determining whether new species form by the accrual of many minor mutations, as Darwin had proposed, or by major mutations that cause large changes in the phenotype. To answer this question, Morgan studied a diverse array of species, including pigeons, chickens, mice, and rats, but he eventually settled on *Drosophila* for a number of reasons. Particularly important was that fruit flies breed year-round and rapidly, with a generation time of 10 to 20 days; each female is able to produce around 300 offspring during her lifetime. With this high reproduction rate, Morgan was able to breed tens of millions of flies (Endersby, 2007), which allowed him to identify mutants even when mutation rates were relatively low. Fruit flies were also easy to feed (with bananas in the early days) and maintain as separate breeding colonies in small spaces. In addition, fruit flies tolerate extensive inbreeding better than many other species, which allowed Morgan's team to create very uniform strains in which mutants are more apparent than in heterogeneous populations.

Tracking the inheritance of mutations across multiple generations allowed Morgan and his colleagues to discover that some mutations are sex-linked. Subsequently, they discovered that multiple mutations tend to be inherited together, forming linkage groups. Finally, they discovered that the occasional crossing over (i.e., recombination) of homologous chromosomes causes systematic variations in the strength of those linkages. This last observation ultimately led Morgan and his team to conclude that genes are concrete entities that are arranged linearly along the chromosomes and could be "mapped" by careful analyses of their linkage frequencies (Morgan, 1915). These discoveries were facilitated by the fact that fruit flies have only four pairs of chromosomes and that crossing-over does not occur in the males of this species. The latter trait is relatively rare (John et al., 2016) but made genetic mapping much easier; it is a good example of the kind of experimental "convenience" that August Krogh emphasized in the chapter's opening quotation.

Once Morgan and his fellow Drosophilists had established that genes on chromosomes are the basis of Mendelian inheritance, interest in fruit flies as research animals faded, and geneticists turned their attention to yeast, bacteria, and viruses (see chapter 4). However, fruit flies attracted attention again in the 1960s when Seymour Benzer began to generate fly mutants that were impaired in specific behaviors, such as phototaxis, circadian rhythms, and memory tasks (Weiner, 1999; Greenspan, 2008). In addition, Christiane Nüsslein-Vollhardt and Eric Wieschaus in the 1970s generated more than 130 fruit fly mutants with defects in larval development. Research on these developmental mutants revealed a complex network of genes and proteins that

together control critical aspects of *Drosophila* development. Many of these genes and proteins have vertebrate homologs that are important for development. It is important to note, however, that some of the mechanisms responsible for fruit fly development do not generalize to vertebrates. For example, the mechanisms underlying body segmentation are quite different in fruit flies and vertebrates and involve many nonhomologous genes (Blair, 2008). Still, many of the principles derived from the work on fruit flies, both with regard to behavior and morphological development, have been broadly applicable.

After researchers developed techniques for creating transgenic fruit flies in the 1980s (Rubin & Spradling, 1982), fruit flies came to be used as models of human disease. For example, researchers created a fruit fly model of Huntington's disease by inserting into flies a mutated fragment of the human *huntingtin* gene. These transgenic flies develop abnormal intracellular accumulations of the Huntingtin protein, neuronal degeneration, and behavioral symptoms that are somewhat reminiscent of human Huntington's disease (Warrick et al., 1998). Similarly, researchers in 2000 inserted into flies a mutated version of the human *alpha-synuclein* gene, which has been linked to familial Parkinson's disease (Feany & Bender, 2000). Although the symptoms of these disease-model flies are often only superficially similar to those of human patients (see chapter 6), the transgenic flies allow researchers to test thousands of potential therapies in high-throughput studies. Another rarely discussed but important advantage of working with flies is the lack of animal welfare regulations for insects (Jennings, 2011).

Although the translation of potential therapies from fruit flies to humans remains more of a promise than an obvious success, work on *Drosophila* has provided vast amounts of information on molecular interactions and cellular processes in flies, and much of that basic science has made significant contributions to research on other species. For instance, the molecular mechanisms underlying circadian rhythms were first discovered in *Drosophila*, but this work then facilitated analogous studies in other species, revealing both similarities and differences (Glossop & Hardin, 2002; Rubin et al., 2006; Tomioka & Matsumoto, 2010). Thus, not everything one finds in *Drosophila* extrapolates to other species, but some of the discovered principles are general; in any case, the fruit fly work has often guided subsequent research. It is for those reasons that calling it a supermodel seems appropriate.

3.4 LABORATORY RATS: THE FIRST STANDARDIZED MAMMALS

Rats have long been regarded as thieves of our food and as dirty creatures that live in sewers, spreading disease. However, experimental biologists have frequently used Norway rats (*Rattus norvegicus*) in their research, starting in the 19th century and then

accelerating in the first half of the 20th century. Why did biologists like to experiment with rats? One reason is that rats, as mammals, are more closely related to humans than non-mammalian vertebrates, or invertebrates such as fruit flies. This phylogenetic proximity is generally presumed to correlate with overall similarity and, therefore, to favor translation (see chapters 2 and 7). A second major reason is that rats breed even more rapidly than guinea pigs and are easy to maintain in large numbers. Rats are also a convenient size for many surgical and histological procedures—not too big and not too small. Another factor in favor of rats is that they, in contrast to guinea pigs, are born in a very immature (altricial) state, which makes it easier for biologists to study their early development. For example, biologists can study the effects of early castration without having to castrate embryos in utero. Studies on aging are also relatively easy to perform in rats, because they only live for two to three years.

Another major impetus for the increasing popularity of laboratory rats in the early 20th century was the desire of biologists to minimize the variability in their experiments. During the early days of experimental physiology, novel pieces of equipment had been invented at a rapid clip, but the variability between different machines made it difficult to compare the results obtained in different laboratories. To combat the resulting crisis of replicability, experimentalists began to standardize their equipment and techniques (Logan, 2002). It made sense, therefore, to standardize the research animals as well. Accordingly, researchers created several strains of rats expressly for research. Especially popular were albino rat strains because any cross-contamination from wild brown rats was relatively obvious. Albino rats also lessen the inherent aversion most people have toward wild rats; letting the imagination run, one might even think of these albino rats as wearing white lab coats. The most famous of these albino rat strains was created at the Wistar Institute at the beginning of the 20th century (King, 1918; Clause, 1993). Because these Wistar rats were highly inbred, they exhibited relatively little variation, which meant that their phenotypic traits could be summarized in numerous standard tables and descriptions (Donaldson, 1924). By 1912, the Wistar Institute was shipping around 6,000 Wistar rats per year to scientists around the world (Endersby, 2007).

Because of their strong sex drive, rats were often used to study reproductive behaviors and hormonal mechanisms. In addition, rats were used in many nutritional studies because their dietary requirements are, in some respects, quite similar to those of humans (Richter, 1968). The popularity of laboratory rats extended also to psychology and neuroscience, mainly because rats are relatively fast learners, at least compared to mice. In the words of Ian Whishaw, "At a fork in the evolutionary road . . . the brown rat chose complexity as a survival trait whereas the mouse chose simplicity. The rat became social, intelligent, complex, and skilled, all of which are attributes it shares

with humans" (1999, p. 411). Particularly advantageous is that rats construct and navigate complex tunnel systems in nature (Calhoun, 1963). The laboratory equivalent of these burrows is the various mazes that experimentalists have devised to study learning and memory in rats (Small, 1900; Watson, 1903). These maze experiments were a key building block in the development of behaviorism and other general theories of animal psychology (Watson, 1913; Skinner, 1938; Tolman, 1948). They also anchored many studies on the neural mechanisms of learning and memory (O'Keefe & Nadel, 1978).

As the scope of rat research expanded during the 20th century, biologists increasingly considered the rat to be the ideal animal model for asking many different types of questions and obtaining answers that could be generalized to many other species—a biological Rosetta stone (Gest, 1995). In words of one influential biologist who had research experience with many different mammals, "If someone were to give me the power to create an animal most useful for all types of studies on problems concerned directly or indirectly with human welfare, I could not possibly improve on the Norway rat" (Richter, 1968, p. 403). In parallel with the emergence of the Norway rat as the dominant laboratory animal, the interest in diversity that had characterized earlier generations of biologists gradually waned. Instead of testing whether findings from one species would generalize to others, their generality became widely assumed (Logan, 2001).

Some scientists became alarmed at this heavy reliance on laboratory rats. Frank Beach (1950), for example, became concerned that comparative psychologists were learning a great deal about the behavior and mental processes of laboratory rats but losing sight of the larger goal, which was to learn about animal minds in general. Citing a poem by Lewis Carroll, Beach argued that comparative psychologists had set out to "hunt snarks" but instead found a very dangerous "boojum," which makes the hunters disappear. At the end of his essay, Beach switched stories and compared the laboratory rat to the Pied Piper of Hamelin, who used his flute to lure the city's rats into a river where they drowned. "Now," Beach wrote, "the tables are turned. The rat plays the tune and a large number of human beings follow. . . . Unless they escape the spell that *Rattus norvegicus* is casting over them, experimentalists are in danger of extinction" (1950, p. 117). Other psychologists echoed Beach's concern. Daniel Lehrman, for example, recommended that psychologists should focus on "questions arising from the natural life of a particular species (rather than questions applied to an arbitrarily selected species from a generalized theoretical framework)" (1971, p. 464). To him, generality was something that should be inferred from comparative data, rather than assumed.

That said, many of the findings obtained from rats, both in psychology and other disciplines, have proven to be quite general. The larger question, then, is not whether

some general principles exist, but whether species differences are important to recognize. For example, is it important to recognize that "isolation housing may produce symptoms of psychopathology in rats, but in mice, it reveals normal species-typical behavior" (Miczek & de Wit, 2008, p. 296)? Similarly, is it important that rats, in contrast to humans, mice, and most other mammals, lack gall bladders (McMaster, 1922)? Does it matter that sleep deprivation is more lethal in rats than in humans or mice (Siegel 2008)? The answer to all of these questions is "sometimes yes" or "it depends" (see chapter 7). In short, rats and mice can serve as good animal models for many purposes, yet differ from humans and one another in some important respects (Ellenbroek & Youn 2016).

3.5 LABORATORY MICE: FROM CANCER FIGHTER TO MODEL ORGANISM

Mice played a relatively minor role in the early days of experimental biology, but they attracted interest around the same time that fruit flies became interesting to biologists and for the same reason: the study of genetics. In particular, several biologists in the early 1900s were asking whether inbreeding with strong artificial selection would lead to stable, viable lines. They explored this question in diverse species, including both fruit flies and mice. Most influential was C. C. Little, who founded and directed the Jackson Laboratory, which eventually became one of the world's most prolific producers of laboratory mice. Starting with just a few strains in 1929, the Jackson Laboratory now maintains more than 8,000 genetically defined mouse strains and ships millions of mice to laboratories around the world. The black-6 mouse (C57BL/6) eventually emerged as the world's most common mouse strain, not for any obvious reason but simply because it was already widely used; as one author put it, "use begat ubiquity" (Engber, 2011).

Why did Little and his colleagues focus so heavily on mice as research animals? They liked that mice take up less space per animal than rats and multiply more rapidly (up to 10 litters per year versus around six per year). Little also appreciated that rodents do not attract as much anti-vivisectionist sentiment as other mammals that might be used for research (Little, 1935). However, Little's main selling point for mice as research animals was that he considered them to be ideal for research on cancer, which by the 1930s had emerged as major threat to human health. Little and his colleagues had observed that mice get cancer just as humans do and that some strains of mice develop cancer more frequently than others. Energetically promoting his mice as heroic miniature troops in the war against cancer (Little, 1935), Little thought that breeding such mice could help identify the genes that cause cancer (Rader, 2004).

As we will see in chapter 5, mice have indeed played a major role in understanding and treating cancer, but Little had underestimated the many nongenetic factors (e.g., viruses, tobacco smoke [Brandt, 2012], and other carcinogens) that play a role in tumor formation.

Whereas *Drosophila* research was driven primarily by the discovery of mutant flies, mouse mutants were relatively rare during the early days of mouse model development, presumably because mice breed much more slowly than fruit flies. By 1931, for example, only 25 mutant mouse stocks were recognized (Rader, 2004). This changed after World War II, when scientists exposed enormous numbers of mice to ionizing radiation (Russell, 2013), which greatly increased the animals' mutation rate. The resulting mouse mutants were studied by many different investigators and led to useful insights. However, the project's original mission had not been to find interesting mutants but to inform public policy about what levels and duration of radiation exposure would be harmful to humans. Indeed, the "mouse-house" work did lead to policy revisions. Most interesting for present purposes is that scientists had extensive debates about whether one could extrapolate exposure effects from mice to humans. For example, Alfred Sturtevant (a major figure in early fruit fly research) argued that, despite quantitative species differences, the data from nonhumans were perfectly sufficient to predict that high-energy radiation causes genetic damage in humans (Sturtevant, 1954).

Mouse research again advanced substantially after the invention of genetic engineering in the 1980s. Especially transformative was the ability to create transgenic mice by injecting recombinant DNA into the nuclei of fertilized eggs and then implanting those eggs into female mice (Gordon et al., 1980); for reasons that are still somewhat unclear, this technique had a much higher success rate in mice than rats (Pradhan & Majumdar, 2016). Many of the early transgenic mice contained insertions of human disease genes or parts thereof, but those transgenes varied in copy number and tended to insert at unpredictable locations in the genome. In most of the published cases the transgenes were significantly overexpressed, raising questions about whether the phenotypes were due to mutations in the transgenes or their unusually high expression level (Fisher & Bannerman, 2019). These limitations were overcome by the invention of techniques for modifying the sequence of specific genes by homologous recombination (Thomas et al., 1986). This same technique could also be used disrupt or "knock out" specific genes (Capecchi, 2005), which allowed biologists to investigate the functions of specific genes without having to wait for mutants to arise by chance. In addition, researchers developed mice that express a transgene only under certain conditions. More recently, they started to create "designer mice" using gene editing techniques (such as CRISPR-Cas9; Fischman, 2020).

Because of the ready availability of so many mutant mice and tools to make additional varieties, genetically modified mice came to dominate vast swaths of biological research. Researchers working on bacteria, yeast, and flies also had access to many mutant strains, but mice are far more closely related to humans and, therefore, more similar to humans, at least when averaged across all traits (see section 2.3.1; see also chapter 7). Already in 1952, C. C. Little had referred to mice as "miniature human beings" (see Rader, 2004), and the yeast geneticist Ira Herskowitz reportedly once said, "I don't consider the mouse a model organism. The mouse is just a cuter version of a human, a pocket-size human" (quoted in Rader, 2004, p. 267). These statements may have been tongue in cheek, but by the beginning of the 21st century mice clearly reigned supreme in biomedical research. In the minds of some observers, this dominance created a sort of "group think about the ready translatability" of data from mice to humans, spawning "murine 'model' monotheism" (Libby, 2015).

Of course, biologists readily admit that mice differ from humans in numerous respects (Perlman 2016). For example, the genome's "regulatory landscape" is quite different between humans and mice (Fisher & Bannerman, 2019), as are many aspects of their immune system (Mestas & Hughes, 2004) and nervous system development (Finlay, 2019). Mice also live accelerated lives compared with us, and this affects many different aspects of their biology, ranging from metabolic rate and normal aging to diverse disease processes (Agoston, 2017). Another good example came to light in 2005, when "hydrogen sulfide was reported to be revolutionary for emergency trauma care and treating soldiers on the battlefield based on its extraordinary ability to place a mouse in suspended animation for many hours with apparent complete recovery of function. . . . However, the hydrogen sulfide concept failed to translate into larger animal models because the mouse was the wrong animal model for the translation of that particular question" (Dobson, 2014, p. 480). The reason for this particular translational failure was that mice have a naturally evolved ability to enter torpor, which is a sort of suspended animation similar to hibernation; humans, of course, do not share this ability.

Because mice clearly are not "miniature human beings," researchers sometimes "humanize" their mice: they often insert human genes into mice or edit the endogenous mouse genes so that they possess the human DNA sequence. These modifications are usually limited to single genes or parts of genes, but it is becoming increasingly feasible to humanize multiple mouse genes at the same time. In addition, biologists can humanize mice by injecting them with human cells. For example, they can partially humanize the immune system of mice by injecting human immune cells into mice that are otherwise immunodeficient (Shultz et al., 2007). Similarly, human

gut microbiomes can be transplanted into mice that lack a microbiome of their own (Arrieta et al., 2016).

3.6 NONHUMAN PRIMATES

Nonhuman primates have long been used as models for humans in the belief that their close phylogenetic relationship to us would make them "high fidelity" models. As we discussed in section 2.3.1, we cannot assume such high model fidelity for any given trait; however, on average, traits are more similar between humans and their closest relatives than between, say, humans and fruit flies (see also chapter 7). Unfortunately, this realization simultaneously increases the ethical concerns over performing experiments on nonhuman primates because their neural and mental traits are also likely to be more similar to our own, at least on average. Thus, any discussion of how much benefit has been derived from research on nonhuman primates, or can be gained in the future (Phillips et al., 2014), must be tempered by questions about the costs in terms of animal suffering. These are not simple discussions to have.

The number of nonhuman primates used in biological research is relatively small, compared with the total number of experimental animals, and has been relatively constant over the last 10 years (figures 3.1 and 3.2). Many of these studies have examined the immune system of nonhuman primates (Messaoudi et al., 2011) because it differs substantially from that of rodents and other model animals. Nonhuman primates are also used in neurobiological research, because nonhuman primate brains share many features with our own brains that are not found—or are very different—in other species (Striedter, 2005).

However, the vast majority of nonhuman primates are used for legally mandated safety and efficacy tests of novel drugs and therapies, including vaccines (Weatherall, 2006). For example, macaques and African green monkeys have been used extensively to understand and develop vaccines against the Zika virus, which can cause serious neurological birth defects in humans (Osuna & Whitney, 2017; Haddow et al., 2020). An even better example of how nonhuman primates were used to advance medical research is provided by the development of vaccines against poliomyelitis.

3.6.1 Polio Vaccine Development

Poliomyelitis, or polio, is a viral disease that paralyzes many of its victims (often children) and is frequently fatal. Polio reached epidemic proportions in 1916, when it killed 6,000 people in the United States. Recurring annually during the summer months, polio epidemics peaked in the 1940s and early 1950s, with the 1952 outbreak yielding more than 57,000 cases in the United States, including 3,000 deaths

and 21,000 cases of permanent paralysis. Treatments for the disease were largely ineffective, and thousands of victims with respiratory paralysis could be kept alive only by means of specialized ventilators (the so-called iron lungs). Fortunately, the first vaccines against polio became available in 1955, and extensive vaccination programs then produced an exponential decline in the incidence of polio (Nathanson & Kew, 2010). Although polio has now been eradicated from many countries, outbreaks occasionally recur in some parts of the world.

Nonhuman primates were critical to the discovery of the virus that causes polio and to the early stages of vaccine development. Specifically, researchers in 1908 showed that injecting neural tissue from human polio victims into monkeys caused the monkeys to develop the paralysis and motor neuron degeneration typical of polio. On the heels of this discovery, others used tissue from those monkeys to infect additional monkeys, thus propagating the virus and allowing the scientists to study its behavior. Eventually, some researchers managed to breed a strain of the polio virus that causes the disease in mice (which the human polio virus does not), and this substantially reduced the need for monkeys in polio research (Armstrong, 1939).

Even more important was the 1949 discovery that polio viruses could be grown in cultured human or monkey cells, which made it possible to harvest large quantities of virus with relative ease (Enders et al., 1949). Initially, the cultured cells were derived from human embryos, but immortalized monkey kidney cells later eliminated the need to sacrifice living organisms entirely (Furesz, 2006). Plus, the virulence of the harvested virus could be determined by measuring the extent to which it damaged the cultured cells (Enders et al., 1980), rather than having to inject the virus into whole animals. Thus, the case of polio nicely illustrates both the crucial role of nonhuman primates in biomedical research and the potential impact of replacing animals generally thought to be highly sentient (see chapter 2, section 2.5.1) with other animals or cell culture.

Animal rights advocates sometimes point out that some of the early research on polio in monkeys had been misleading. Indeed, a series of influential early studies on monkeys suggested that the polio virus enters the human brain via the olfactory system, whereas later studies revealed that in humans the virus typically enters through the gut. The discrepancy arose because the researchers working with monkeys had unwittingly evolved a strain of the polio virus that is highly neurotropic, meaning that it preferentially infects neural tissue, which is most directly accessible through the nose. This early misstep caused some delays in vaccine development, but, in the words of Albert Sabin, who ultimately developed one of the successful polio vaccines, "without the use of animals and of human beings, it would have been impossible to acquire the important knowledge needed to prevent much suffering and premature death not

only among humans but also among animals" (quoted by Speaking of Research editor, 2011). Indeed, Sabin reported that, during the four years leading up to his successful vaccine, his laboratory had used "approximately 9,000 monkeys, 150 chimpanzees, and 133 human volunteers" (Sabin, 1956, p. 1589). Importantly, the human volunteers only received vaccines that had been previously tested on monkeys and chimpanzees (Sabin, 1965).

A frequently unappreciated aspect of polio vaccine development is that some early attempts to produce these vaccines ended up giving humans polio, rather than making them immune to it (Horstmann, 1985). These early failures underscored the need for extensive testing of vaccines before they are given to humans. Such tests do not always have to be performed in nonhuman primates, but they should be performed in animals that are susceptible to the disease, which not all species are (Rivera-Hernandez et al., 2014).

Although animal tests are usually mandated by regulatory agencies, those mandates can be waived in the midst of pandemics. During the COVID-19 outbreak of 2020, for example, some vaccines went into clinical trials before animal test results had been obtained (Boodman, 2020). Thus, societies are sometimes willing to risk human health in the hope of reducing overall suffering. In normal times, however, humans are more risk-averse.

In fact, many people do not trust vaccines even when they are widely regarded as safe and effective. It will be interesting to see whether these antivaccine advocates will change their mind now that COVID-19 vaccines have been demonstrated to save numerous lives, or whether they will find their suspicions confirmed if a few individuals get sick after receiving a COVID-19 vaccine. Most likely, both views will persist, as they have for other vaccines (DeStefano et al., 2019).

3.6.2 The Silver Spring and U Penn Monkey Affairs (1980s)

Neurobiological studies on nonhuman primates are less common than toxicological or vaccine-related studies, but they tend to attract more attention from animal rights advocates, in part because their relevance to human health is not as apparent. A prime example is the case of the "Silver Spring monkeys," which in 1981 facilitated a significant expansion of the animal rights movement in the United States and provided the impetus for major changes in US laws and regulations concerning animal welfare.

The case involved the laboratory of Edward Taub at the Institute for Behavioral Research in Silver Spring, Maryland. Taub had been cutting the sensory nerves in one arm of macaques, which led to long-term motor impairments. Taub was doing these experiments to test whether his animals could learn to use their damaged arm if they were prevented from using their "good" arm (Taub, 1980; Taub et al., 1994). In the

midst of this research, Taub's laboratory was infiltrated by an animal activist who was supposed to help care for the animals. This activist secretly documented that some of the animals gnawed on the limbs lacking sensory nerves, had poorly bandaged wounds, and lived in small, filthy cages. While Taub was on vacation, the activist alerted the police and press. As a result, the monkeys were confiscated, and Taub was convicted on several counts of cruelty to animals. The convictions were later overturned, but the case attracted enormous media attention, in part because the confiscated monkeys were briefly kidnapped by individuals who did not want them returned to Taub.

Although this case revealed real problems with monkey research at the time, Taub's original hypothesis has turned out to be largely correct and led to the successful development of "constraint-induced movement therapy" for humans with nerve damage in their arms (Wolf et al., 2008; Fritz et al., 2012). In addition, later studies on Taub's monkeys revealed a dramatic reorganization of their cerebral cortex (Pons et al., 1991). This finding then inspired many later studies on cortical plasticity, which is now recognized as being crucial to rehabilitation after nervous system damage (Kleim & Jones, 2008).

A second good example of the controversies engendered by neurological research on nonhuman primates involves baboons that were used at the University of Pennsylvania to study traumatic brain injury. An activist group broke into this laboratory in 1984 and stole 60 hours' worth of videotapes documenting the experiments. These tapes were then edited down to a 30-minute movie that shocked most viewers (Dusheck, 1985), mainly because many of the monkeys were not fully anesthetized when the trauma was applied, and because their subsequent behaviors were pitiful.

Since the activists' tactics had been blatantly illegal and the need for animal research on traumatic brain injury was widely accepted among scientists, the authorities responded more slowly to this episode than the Silver Spring monkey case. Still, the researchers were ultimately cited for violations of the Animal Welfare Act, the National Institutes of Health (NIH) temporarily halted the relevant grant, and the university stopped supporting nonhuman primate research on head injuries. In conjunction with the Silver Spring monkey case, the episode contributed substantially to the passage of the 1985 amendment of the Animal Welfare Act and to changes in NIH policy.

In the wake of the fracas at U Penn, research on traumatic brain injury became more focused on rats and mice (Xiong et al., 2013; Shultz et al., 2017). These rodent studies have provided extensive information about the cellular and molecular consequences of various types of brain damage, and they produced a number of promising therapies that minimize brain damage in rodents (see chapter 5). Unfortunately, when these therapies were tested in humans, all of them failed (Narayan et al., 2002;

Stein, 2015). One reason for these clinical trial failures is that human brain injuries are multifaceted and variable, whereas the animal models all focus on a specific type of injury and standardize all variables as much as possible. It has also been suggested that rodents, with their small, smooth brains, are poor models for humans when it comes to traumatic brain injury.

To circumvent the latter problem, minipigs (see section 3.7.1) have become an increasingly popular model for this type of research (Kinder et al., 2019). The brains of pigs are 50 to 90 times larger than those of rats and have a highly folded cerebral cortex. Moreover, the public tends not to be as concerned about the welfare of pigs as of monkeys, though the rationale for this differential empathy is rarely spelled out. Whether research on pigs will ultimately lead to treatments for brain trauma in humans remains to be seen.

3.6.3 The Chimpanzee Research Debate

The vast majority of biological research on nonhuman primates has been conducted with macaques, specifically rhesus macaques (*Macaca mulatta*) or crab-eating macaques (*M. fascicularis*, aka cynomolgus monkeys). In addition, biologists have studied some New World monkeys (notably marmosets; see section 3.6.4) and, sometimes, the common chimpanzee *Pan troglodytes.* The latter work has been extremely controversial, because chimpanzees are our closest living relatives. Their DNA sequence is often said to be 99% identical to ours, but that number drops to 96% to 97% when one takes deletions and insertions into account (Chimpanzee Sequencing and Analysis Consortium 2005; Varki & Altheide 2005). Still, chimpanzees and humans share more than 13,400 proteins, and about 29% of them have identical amino acid sequences in the two species; the rest differ by just one or two amino acids.

Consistent with this high level of genetic similarity, chimpanzees and humans are similar in many morphological and physiological respects, as well as mental and emotional capacity (de Waal, 2016, 2019). Some of the cognitive similarities may be "overzealous efforts to dismantle arguments of human uniqueness" (Povinelli & Bering, 2002, p. 115), but it seems fair to say that no other animals are as broadly similar to humans as chimpanzees.

Most important for present purposes is that chimpanzees are very similar to humans in many medically important features (Institute of Medicine & National Research Council, 2011; Phillips et al., 2014). For example, the rates at which several drugs are cleared from the body after oral administration correlate quite well between humans and chimpanzees, but not between humans and monkeys, dogs, or rats (Wong et al., 2004) (figure 3.3). Similarly, several viruses, including human immunodeficiency virus (HIV) and the hepatitis C virus, infect humans and chimpanzees but not most

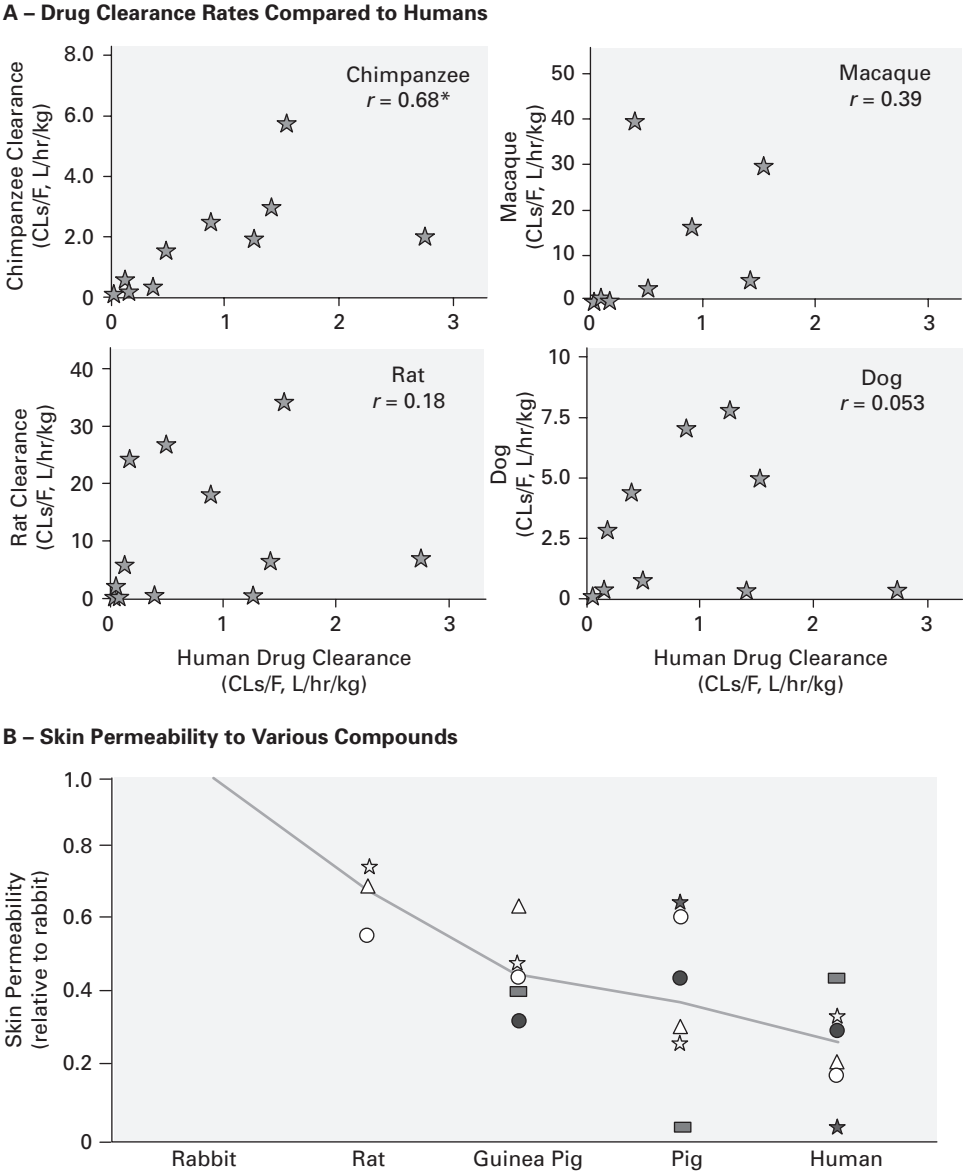


Figure 3.3

Variation across species in drug clearance and absorption. (*A*) Comparison across five species shows the rates at which various (orally administered) drugs are cleared from the animal's system. Clearance rates are significantly correlated between humans and chimpanzees (the asterisk indicates a statistically significant Pearson's *r* coefficient at $p = 0.015$) but not between humans and macaques, beagles, or rats. (*B*) The skin of humans and pigs is much less permeable to various chemicals than the skin of rabbits, rats, or guinea pigs (the data are normalized to each compound's permeability in rabbits; not all compounds were tested in all species). Data such as these are used to justify the use of pig skin as a model for human skin. Adapted from (A) Wong et al. (2004); (B) Calabrese (1991, figure 2-17).

other species (Haigwood & Walker, 2011). Of course, chimpanzees also differ from humans in ways that are medically relevant. For instance, heart attacks tend to have quite different causes in humans versus chimpanzees (Varki et al., 2009). Moreover, chimpanzees tend not to develop acquired immunodeficiency syndrome (AIDS) after being infected with HIV, and they mount a weaker antibody response to hepatitis C infection (Institute of Medicine & National Research Council, 2011).

In general, immune system–related genes have diverged substantially between humans and chimpanzees (Chimpanzee Sequencing and Analysis Consortium, 2005; Bitar et al., 2019). Particularly interesting is that the type of white blood cell that caused the life-threatening cytokine storms in the TGN1412 drug trial (see chapter 1) is less reactive in chimpanzees than in humans, suggesting that these cells were genetically modified in the human lineage after it split from chimpanzees (Chapman et al., 2007; Soto et al., 2010). Thus, chimpanzees exhibit high but still imperfect fidelity as models for human diseases and therapeutics.

In light of these considerations and the inescapable ethical concerns, the US Institute of Medicine concluded in 2011 that research on chimpanzees is no longer necessary for invasive biomedical research (Institute of Medicine & National Research Council, 2011). Instead, it highlighted as alternatives the use of humanized mice and the development of extremely sensitive techniques that make it possible to test humans for adverse reactions to drugs at very low doses (i.e., microdosing). The rapporteurs did leave open the possibility that biomedical research on chimpanzees might become necessary in the future to deal with as yet unknown medical threats. Therefore, rather than recommending a total ban on chimpanzee research, they proposed specific criteria that such research would have to meet. They also recommended a lower bar for comparative genomic and behavioral research with chimpanzees, because such research can be performed without invasive procedures or with animals that have been trained with positive reinforcement to "acquiesce to" anesthesia and other procedures.

Based in part on this report from the Institute of Medicine, the NIH withdrew its support for chimpanzee research in 2015 (Kaiser, 2015). Even before then, the NIH had declared a moratorium on breeding chimpanzees for research and transferred many of its captive chimpanzees to animal sanctuaries. An outright ban on research with great apes, called the Great Ape Protection Act, was introduced in the US Senate repeatedly between 2008 and 2012, but it never passed. Nonetheless, access to chimpanzees for research purposes has become extremely difficult in the United States (VandeBerg & Zola, 2005). Similar barriers to chimpanzee research have been erected in many other countries that carry out significant amounts of biological research. In the United Kingdom, for example, no great apes have been used in research since 1986 (Weatherall, 2006). Some important comparative research is being done with archival

tissue from chimpanzees—especially their brains (see chimpanzeebrain.org)—but even this research has suffered from inconsistent grant support.

3.6.4 Marmosets

Relatively new to the scene of nonhuman primate research is the common marmoset (*Callithrix jacchus*). In contrast to humans, chimpanzees, and macaques, which all belong to the catarrhine lineage of primates, marmosets are platyrrhine primates (aka New World monkeys). Given this phylogenetic position, it is not surprising that marmosets lack some of the features that are unique to catarrhine primates and that their protein coding sequences are, on average, less similar to our own than those of macaques (Burkart & Finkenwirth, 2015; Preuss, 2019). Marmosets also differ from many other primates in that they tend to give birth to twins and are unusually small, weighing an order of magnitude less than a typical macaque. Because of this small body size, marmosets require less vivarium space than other primates, which in turn reduces housing costs. Marmosets also reach sexual maturity after just 1.5 years (versus three to five years for macaques), which facilitates breeding and the creation of transgenic marmosets (Sasaki et al., 2009; Park et al., 2016; Sato & Sasaki, 2017). Yet another important consideration is that marmosets do not carry the macacine herpes virus, which can be transmitted from macaques to humans and cause serious illness or death.

Because of these advantages, marmosets have become increasingly popular as research animals in several countries. Japanese researchers, in particular, have developed a large research program focused on marmosets (Okano et al., 2016). Much of this research is centered on the brain, which is far more similar between humans and marmosets than between humans and nonprimates (Preuss, 2019). Although Japan has clearly taken the lead, the People's Republic of China is also investing in marmoset research. However, China has opted for an even greater focus on macaques, including genetically modified macaques (Hao, 2007; Poo et al., 2016; Park & Silva, 2019). In fact, China is luring prominent primate biologists away from other countries, where nonhuman primate research is more difficult, expensive, and controversial (Abbott, 2014; Cyranoski, 2016; Vogel, 2020).

3.7 THE REST OF THE MENAGERIE

In addition to the previously mentioned species, biologists study a wide array of other species, most of which were selected according to Krogh's principle, which is to say that they are either especially convenient for experimental investigations or highly specialized in other, interesting ways. Many of them are studied to answer basic science questions that have no immediate biomedical relevance. I will return to these

nontranslational research efforts at the end of this section. First, however, let us review four species that have played major roles in translational research, namely pigs, chickens, zebrafish, and the nematode worm *Caenorhabditis elegans*.

3.7.1 Pigs

Roughly 50,000 pigs are used for research purposes in the United States per year, and their prevalence seems to be increasing globally (Brown et al., 2013). Especially popular are several strains of minipigs; weighing 20 to 90 kg, they require less space and food than the large pigs bred in agriculture (Gutierrez et al., 2015). Although pigs reproduce more slowly than mice or rats, they mature relatively fast (four to five months), can have multiple litters per year, and produce five to eight offspring per litter (van der Staay et al., 2017; Luo et al., 2018). Thus, they reproduce significantly faster than nonhuman primates.

The pig genome was sequenced in 2012 and found to be more similar to that of humans than the mouse and rat genomes are, which is surprising given that rodents are more closely related to primates (see figure 2.2). A closer look reveals that it is the mouse genome that has diverged unusually fast, mainly by undergoing large chromosomal deletions (Thomas et al., 2003). Genetic engineering is not as easy in pigs as in mice, but biologists have now developed a variety of ways to modify porcine genomes, and the use of genetically modified pigs has risen substantially (Vidinská et al., 2018; Wolf et al., 2019).

Aside from being similar to humans in some genetic respects, pigs resemble humans in diverse other ways. In the words of one pig model enthusiast, "pigs are not just pigs but almost human" (Douglas, 1972, p. 226). Indeed, several organs are quite similar between humans and pigs, including the skin, the eyes, the immune system, and the gastrointestinal tract. These similarities have encouraged the use of pigs as models for a variety of human maladies, ranging from sunburn to brain injury and depression (Gieling et al., 2011; Gutierrez et al., 2015; Kinder et al., 2019). The large size of pigs also makes them frequent subjects for surgery technique development and medical device testing (Swindle et al., 2012; Hennessy & Goldstein, 2019).

The most common use of pigs in biological research, however, is as a "second non-rodent species" in toxicology and therapy testing (Bode et al., 2010; Brown et al., 2013). Regulatory agencies usually require such a second species, and experiments on dogs and primates tend to raise more ethical concerns than those on pigs. One should note, however, that this disparity arises mainly because pigs are so widely consumed for food. Numerous studies have suggested that pigs are actually quite intelligent and emotionally complex (Marino & Colvin, 2015).

3.7.2 Chickens

Domestic chickens descended from Red Jungle Fowl more than 8,000 years ago. The number of chickens used for research in the United States is difficult to estimate because those numbers are not reported to the US Department of Agriculture (USDA). In the United Kingdom, however, researchers typically perform experiments on just over 100,000 birds per year (mostly chickens; see figure 3.2). Much of this research deals with agricultural issues, notably selective breeding, nutrition, and animal health. However, chickens also serve as models for the study of more general biological questions.

At least since Aristotle, biologists have realized that chicken embryos are easy to study because they are easy to maintain and breed, develop outside of the mother, can be incubated artificially, and are relatively large. They also float on top of the yolk, which makes them easy to observe (through a window in the egg shell) and manipulate experimentally. Researchers have also learned how to create genetically modified chickens (van de Lavoir et al., 2006; Collarini et al., 2015; Woodcock et al., 2017). Collectively, these studies have revealed many features of vertebrate development that seem to be broadly conserved (Stern 2005).

In addition, chicken embryos and chicken cell lines have been used extensively for vaccine development; in fact, approximately 82% of all flu vaccines in the United States are grown in chicken eggs (this technique does not work for coronavirus vaccines [Yeung, 2020]). One may further note that chickens have contributed significantly to cancer research, notably through the discovery of cancer-causing viruses (see chapter 5) and the study of ovarian cancer, which develops spontaneously in chickens and humans, most likely because both of these species ovulate more frequently than other species (Bahr, 2008). Finally, chickens and their eggs are widely used as models in environmental toxicology, where the aim is to determine the effects of diverse chemicals on humans or wildlife (Giesy et al., 2003; E. G. Xu et al., 2019).

3.7.3 Zebrafish

The zebrafish (*Danio rerio*) is a small teleost fish native to southern Asia. It is popular in the pet trade, hardy, and inexpensive to maintain. Zebrafish become sexually mature by 10 to 12 weeks, lay on the order of 100 eggs at a time, and can mate every couple of weeks (Meyers, 2018). Thus, they reproduce extremely rapidly (for vertebrates) and can be housed cheaply in very large numbers. These features make zebrafish very convenient for genetic studies, especially when mutagens are used to increase the rate of mutant production. Indeed, tens of thousands of mutant zebrafish lines have now been created and can be ordered from centralized facilities (e.g., zebrafish .org). Although inbreeding significantly suppresses fertility in zebrafish (Monson & Sadler, 2010), mutant lines can be maintained by experimentally inducing the females

to reproduce parthenogenetically (i.e., without DNA from the male) or, more recently, by combining frozen sperm, in vitro fertilization, and strategically designed breeding regimes (Geisler et al., 2016).

The zebrafish genome was sequenced in 2013, and 70% of all zebrafish genes were found to have human homologs (Howe et al., 2013). However, more than half of these genes do not have simple one-to-one homology (i.e., orthology) relationships to their mammalian counterparts, mainly because teleosts apparently duplicated their entire genome early in their history and then lost a subset of the duplicated genes (Inoue et al., 2015). Because of these complex genetic differences, the effects of single gene mutations in zebrafish can be difficult to discern if the duplicate gene remains intact and compensates for the mutated gene. However, duplicated genes sometimes diverge in such a way that each of the duplicates assumes a subset of the ancestral gene's functions (Force et al., 1999). In such instances, identifying those functions may actually be easier in zebrafish than in species where the genes are not duplicated (Kleinjan et al., 2008).

Another difference between zebrafish and the more traditional model organisms—notably fruit flies and mice—is that the wild-type lines of zebrafish are not well standardized, both because of inbreeding suppression and because of divergence between strains maintained in different facilities, which are periodically outcrossed to wild zebrafish to increase their vigor. In addition, the various wild-type zebrafish lines currently in use were derived from different subsets of the wild population, which is very heterogenous genetically (Suurväli et al., 2020).

Starting in the 1960s, zebrafish were used extensively for the study of embryonic development. Zebrafish were especially convenient for such research because their embryos develop outside the mother's body and do so very rapidly (most of their organs are formed within four days). Furthermore, zebrafish embryos are transparent, which makes it possible to trace the lineage and movements of individual cells in living embryos, be they mutant or wild type. In addition, researchers have developed a variety of methods for surgical or genetic manipulation of zebrafish embryos. Collectively, these studies have revealed a variety of developmental mechanisms that apply not only to zebrafish but also, at least in principle, to other vertebrates (Grunwald & Eisen, 2002).

A second area where zebrafish are used extensively is toxicology, mainly because it is relatively easy to expose multiple zebrafish larvae or embryos to multiple test compounds "in parallel" (in multiwell cell culture plates) (Horzmann & Freeman, 2018; Cassar et al., 2020). As explained more thoroughly in the next chapter, such high-throughput screens are desirable because regulatory agencies have a large backlog of compounds that have never been tested for toxicity. Indeed, several zebrafish toxicology screens have shown at least moderate concordance with data from mammalian

species. For example, a set of drugs to known to affect mammalian heart function has somewhat similar effects in zebrafish (Milan et al., 2003; Dyballa et al., 2019).

Collectively, these data indicate that the predictive validity of the zebrafish model in toxicology is far from perfect, especially when it comes to toxin sensitivity, but no worse than that for other species—and sometimes better than that of human cell culture systems (Dyballa et al., 2019). At the very least, the results from zebrafish screens can be used to prioritize some chemicals for additional scrutiny. Although this tiered approach is widely endorsed, the risk of false negatives (i.e., missing adverse effects) appears to be relatively high even in mammalian screens (Olson et al., 2000; Monticello et al., 2017). One should also note that terrestrial mammals are typically exposed to toxins via inhalation, ingestion, or across the skin, whereas the zebrafish in typical screens are simply (and very conveniently!) immersed in water-soluble chemicals; such different routes of exposure may well produce divergent results.

Over the last two decades, improved techniques for targeted mutagenesis and genome editing have allowed researchers to model a variety of human diseases in zebrafish (Bradford et al., 2017; Davis & Katsanis, 2017). To create these models, researchers typically insert the human disease genes into the zebrafish genome or, using gene editing, mimic the disease-linked mutations in the endogenous zebrafish homologs. When the resulting fish exhibit symptoms similar to those of the human disease, the model is usually considered successful, and the models are then used for high-throughput screening of potential therapeutics (Wiley et al., 2017).

Some of this work has led to novel therapies for diseases in zebrafish models, which were then tested in humans (Cully, 2019). For example, a wide variety of cancers have been genetically induced in zebrafish and found to be quite similar to cancers in other species (Kirchberger et al., 2017). Some researchers have even implanted zebrafish with cells derived from human tumors and then tested how well the tumors in the fish respond to various potential therapies (Fior et al., 2017); the drugs found to be most effective in the "avatar fish" can then, at least in theory, be administered to the patient (Fazio et al., 2020). These uses of zebrafish to fight cancer are very promising, but they are not yet in clinical use. One anticancer drug that had emerged from zebrafish research (i.e., ProHema) passed phase II clinical trials but then "fell by the wayside" (Cully, 2019) as other forms of cancer therapy emerged.

Similarly, zebrafish are being used to model Dravet syndrome, a rare but severe form of epilepsy that involves mutations in a specific sodium channel gene. Zebrafish with analogous mutations exhibit abnormal patterns of brain activity as well as swimming movements that are "reminiscent" of epileptic seizures in humans (Griffin et al., 2016). These abnormalities abate when the fish are treated with some known antiepileptic drugs (Baraban et al., 2013). Thus, this model scores relatively high in terms

of construct, face, and predictive validity (see chapter 2, section 2.2.1). Moreover, a high-throughput drug screen performed on these fish revealed several effective treatments, the most promising of which is clemizole. A phase II trial testing whether this compound (aka EPX-100) is safe for children with Dravet syndrome is ongoing (no results have been published as of July 2021).

In short, zebrafish show a great deal of promise as models for the study of human disease, but the extent to which such work will deliver new medicines remains an open question. This is not surprising given that the use of zebrafish to model human diseases only began in earnest 20 years ago and clinical trials often take a decade or more to come to fruition.

3.7.4 "The Worm" *Caenorhabditis elegans*

The roundworm *C. elegans* is just over 1 mm long as an adult and can live at high density in petri dishes containing agar and bacteria. Most individuals are females that reproduce as hermaphrodites via self-fertilization, and each female can produce approximately 300 offspring, which reach adulthood in three to five days. *C. elegans* anatomy is quite simple, with each adult female containing 959 cells, including 302 neurons.

Because the animals are transparent, researchers have been able to trace the developmental lineages of all these cells. Remarkably, they found the entire cell lineage map and most of the neuronal connections to be highly stereotyped across genetically identical animals (Sulston, 2002; Ankeny, 2007). Laser ablation of specific cells subsequently revealed that the fate of many cells depends on interactions with other cells and is not, therefore, rigidly preprogrammed (Kenyon, 1988). The cell lineage studies also showed that a significant number of cells die during normal development, which is to say that they undergo programmed cell death. More recent studies have combined the natural transparency of *C. elegans* with the use of calcium indicator dyes to monitor neuronal activity during behavior (Larsch et al., 2013), something that is much more difficult to do in larger animals.

A major aim of *C. elegans* research has long been to study the functions of genes (Brenner, 2002). It was convenient, therefore, that *C. elegans* breeds so rapidly and can be maintained in enormous numbers. The use of chemical mutagens and the ability to rapidly screen thousands of animals further facilitated the identification of mutants, and the worms' asexual mode of reproduction made it easy to create homozygous mutant lines. Having identified interesting mutants, researchers were able to identify the altered genes and compare their sequences to those of genes in other species. Using this general approach, researchers have identified numerous *C. elegans* genes with important developmental functions, including programmed cell death (Horvitz, 2002).

A number of these genes have homologs in humans, and some of their functions appear to be conserved as well, at least at the cellular and molecular levels. Because this degree of conservation is surprising given the enormous phylogenetic distance between humans and worms, it helped to spawn the widespread belief that all the truly "fundamental" features of animal life are broadly conserved (Horvitz, 2002). However, sequencing of the *C. elegans* genome revealed that only about 35% of its genes have human orthologs (i.e., one-to-one homologs) (Shaye & Greenwald, 2011).

Despite the significant amount of genetic divergence between roundworms and humans (Zdobnov et al., 2005), *C. elegans* make an attractive model for toxicology research because high-throughput screens are so easily performed with this species. Indeed, like zebrafish, roundworms are often viewed as "a bridge between *in vitro assays* and mammalian toxicity testing" (Hunt 2017, p. 56). Unfortunately, the data obtained from the *C. elegans* screens do not closely track the results obtained in zebrafish assays (Boyd et al., 2016). Indeed, the concordance rates between different model systems in toxicology are generally not very high (see chapter 4), which is why regulatory agencies usually request tests to be conducted in at least two different models and why positive test results in non-mammalian systems are mainly used to prioritize *potential* toxins for further testing. In general, toxicological decision making must balance the need to avoid false negatives, which would endanger public health, against the problem of false positives, which can create needless economic distress for chemical or pharmaceutical companies. Achieving this balance is very difficult when concordance rates are low.

Analogous issues arise when *C. elegans*—or, for that matter, other invertebrates—are used to model human diseases. In the words of Titus Kaletta and Michael Hengartner,

> Given that even mammalian models are often not reliably predictive of drug action in humans, it is—from a preclinical model perspective—unrealistic to expect an invertebrate system to give enough confidence to predict drug action and safety in humans. Non-mammalian model organisms will be typically used in early research and should deliver fast answers to a discovery problem, such as the function of a gene, or pioneer medical research to define novel therapeutic entry points. Of the animal models, *C. elegans* is certainly the fastest and most amenable to cost-effective medium/high-throughput technologies. *C. elegans* is a valuable disease model if the disease can be defined on a molecular basis. (2006, p. 387)

Indeed, biologists generally accept that invertebrate or in vitro models of human diseases never fully replicate the human condition (see chapter 6, section 6.4.3). Press releases aside, the principal goal of such studies is to learn how the modified genes function in the model system and then to use this information to guide experiments in humans or, more frequently, in other models that more closely resemble humans.

This is a reasonable, time-honored approach. However, the success of the cross-species extrapolations depends on the degree of sequence similarity between the human disease genes and those of the model, as well as the molecular, cellular, and organismal contexts in which those genes operate.

Given the large genetic differences between *C. elegans* and humans, it is not surprising that attempts to extrapolate between these two species do not always succeed. This moderation of expectations explains why invertebrate disease models are generally considered successful when genetic manipulations produce cellular or behavioral symptoms that bear at least a superficial similarity to the modeled human disease, not when they lead to novel therapies that succeed in clinical trials.

To the best of my knowledge, few if any therapies of human diseases have emerged directly from *C. elegans* research. Perhaps this is why Sydney Brenner, the founding father of *C. elegans* research, reportedly suggested in 2008,

> Throw out the animal models for the moment and focus on man as the primary organism of study. I started on *C. elegans* as a model system because humans were not experimentally accessible. But now we have the human genome. And if man's genes are accessible, it's man's genes that we should be focusing on. (Friedberg 2008, p. 9)

This view is probably not widely shared among biologists, but it is worth considering. We will come back to it in chapters 6 and 7.

3.7.5 Animal Models in Nontranslational Research

Many biologists study species other than those I review in this book, and they frequently do so for reasons that are independent of any translational ambitions. For example, they study how songbirds learn their songs, how some fish use weak electrical signals to communicate and navigate in muddy waters, how hummingbirds or bumblebees fly, how locusts swarm, or bees and ants communicate (Camhi, 1984; Catania, 2020). Such work tends to focus on species that exhibit the traits of interest either uniquely or most robustly than other species, and are well suited for experimental inquiry. Small research communities usually form around the study of these specialist species—sometimes called "Krogh organisms" (Green et al., 2018)—and community members often share research-facilitating resources, such as husbandry techniques and genomic data. Thus, these species qualify as "model species" and may even aim for "model organism" status insofar as the research on them spans diverse levels of analysis (National Research Council, 1985; Ankeny & Leonelli, 2020).

However, most of these nontraditional model species were not initially selected because researchers thought that findings obtained in them would generalize to humans or other distant relatives. On the contrary, they were generally selected for some fascinating

trait that is *not* present in all species. The ensuing research is sometimes said to be curiosity driven, but it can certainly yield unexpected societal benefits (e.g., via bio-inspired engineering or unexpected medical applications). Recalling Pasteur's metaphor of science as a fruit tree (see chapter 1), we can say that this nontranslational, curiosity-driven research is aimed at growing the tree, rather than reaping a harvest.

Unfortunately, a full discussion of nontranslational research and the species it relies on would burst the limits of this book. However, I believe that curiosity-driven research should be valued more highly than is currently the case (Zoghbi, 2013; Lindsley, 2016). It should be viewed in concert with applied/translational science as being part of an overarching search for biological principles that are quite general, even if the details of their implementation are somewhat species-specific (see chapter 7, section 7.1.4).

3.8 MODEL SYSTEM ECOLOGY

As shown in figure 2.1 and in several figures of the present chapter, the popularity of the various species used for research has waxed and waned over the years. Since the various model systems are competing with one another for research funds and attention from researchers, one can think of them as being players in an ecosystem of models. We will discuss this notion of competing model systems more thoroughly in chapter 4, but already some trends and principles are clear.

As the use of guinea pigs, rabbits, dogs, and other mammals in research declined, laboratory rats and mice became the dominant research animals. This shift was largely driven by the fact that rats and mice breed more rapidly than other mammals and are cheaper to house in large numbers. It was also propelled by a shift toward "big science" biology (Logan, 2019) and its attendant view of biomedical research as an industrial enterprise that would proceed much more efficiently with highly standardized and "pure" research materials (Rader, 2004; Kirk, 2012).

Particularly important in this transformation was the push by C. C. Little and others in the late 1930s to view inbred mice as "the biological equivalents of standardized, interchangeable parts in a well-oiled machine of biomedical research production" (Rader, 2004, 155). As funding for the centralized production of laboratory mice (mainly at the Jackson Laboratory) increased, the abundant supply created (or at least supported) more demand by researchers. A postwar effort at one of the US National Laboratories to study the effects of radiation on enormous numbers of mice (Russell, 2013) further entrenched inbred mice as the go-to animals for large-scale biological research. With the rise of experimental techniques for manipulating mouse genomes in the 1980s, the popularity of mice in research laboratories became unparalleled.

Not to be neglected is that rats, mice, and flies are pests, so they evoke relatively little compassion from either scientists or the general public. In general, biologists have moved away from species widely held as pets to species that are pests or bred for food. Worms do not fit neatly into this general rule, but killing them is widely considered acceptable. Zebrafish also occupy an interesting middle ground because, as vertebrates, they have intricate nervous systems, complex behaviors, and, according to some, the capacity for feeling and remembering pain (Sneddon, 2009; Braithwaite, 2010; Key, 2015; Striedter, 2016). However, the small, transparent, and abundant zebrafish larvae used in many studies are unlikely to trigger much empathy in most observers; even adult fish are usually considered low on the (supposed) phylogenetic scale. In the words of two young zebrafish researchers, "if you can study it in fish, all things equal, you'd rather do that than in a mouse or a dog . . . I'm happy working on fish. . . . I'm not sure I'd be happy working on mice. I definitely would not be happy working on primates" (quoted in Endersby, 2007, pp. 407–408). Similar sentiments are widely shared among biologists, though rarely verbalized.

The research questions and aims pursued by biologists have also changed over the years. In the early days, biologists were keenly aware of species differences and sought to discover principles that could explain and organize those differences. This was true even in the early days of genetics, where Mendelian inheritance and chromosome maps emerged as some of the key principles. The pattern began to change with the rise of molecular biology, which unearthed principles so general that variation became largely uninteresting (e.g., the genetic code, which actually does exhibit some interesting variation) (see chapter 4, figure 4.2). The guiding assumption became that all truly important findings would be "fundamentally conserved" (Grunwald & Eisen, 2002). As a result, model selection became almost exclusively a matter of experimental convenience and ethical concerns about animal welfare; species differences became an afterthought, at least for most medically oriented biologists.

Now, however, biologists are reluctantly discovering that many of their findings do not generalize across models as well as they had hoped (Yartsev, 2017). This disappointment is evident both in toxicology, where cross-species concordance is relatively low (see chapter 4), and in disease modeling (see chapters 5 and 6), where therapies that work in animals fail at alarming rates in clinical trials. As I will argue in chapter 7, these observations suggest that biologists should pay more attention to model differences and, once again, take up the search for principles that can accommodate variation.

This is a section of doi:10.7551/mitpress/14366.001.0001

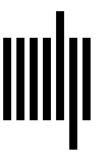
Model Systems in Biology

History, Philosophy, and Practical Concerns

By: Georg Striedter

Citation:
Model Systems in Biology: History, Philosophy, and Practical Concerns
By: Georg Striedter
DOI: 10.7551/mitpress/14366.001.0001
ISBN (electronic): 9780262370028
Publisher: The MIT Press
Published: 2022

The open access edition of this book was made possible by generous funding and support from MIT Press Direct to Open



The MIT Press

The MIT Press would like to thank the anonymous peer reviewers who provided comments on drafts of this book. The generous work of academic experts is essential for establishing the authority and quality of our publications. We acknowledge with gratitude the contributions of these otherwise uncredited readers.

This book was set in Adobe Garamond Pro and Berthold Akzidenz Grotesk by Westchester Publishing Services.

Library of Congress Cataloging-in-Publication Data

Names: Striedter, Georg F., 1962– author.
Title: Model systems in biology : history, philosophy, and practical concerns / Georg Striedter.
Description: Cambridge, Massachusetts : The MIT Press, [2022] | Includes bibliographical references and index.
Identifiers: LCCN 2021033979 | ISBN 9780262046947 (hardcover)
Subjects: LCSH: Animal models in research. | Animal experimentation.
Classification: LCC R853.A53 S77 2022 | DDC 616.02/7—dc23
LC record available at https://lccn.loc.gov/2021033979