6 NEUROLOGICAL DISORDERS: TRIALS AND TRIBULATIONS

> In the 50-odd years since the introduction of clinically effective medications for the treatment of behavioral disorders such as depression, anxiety or schizophrenia there has recently been growing unease with a seeming lack of substantive progress in the development of truly innovative and effective drugs for behavioral disorders; an unease indicated by escalating research and development expenditure associated with diminishing returns.
>
> —MCARTHUR AND BORSINI (2008), P. XV

I here define neurological disorders very broadly to include the effects of neuronal injury (e.g., spinal cord injury) as well as neurodegenerative and psychiatric disorders. Other authors may distinguish between neurological and behavioral/mental/psychiatric disorders, but I here prefer a single, comprehensive term (Insel et al., 2010). It is certainly reasonable to include stroke among the neurological disorders because it involves some brain damage, but we already discussed stroke in chapter 5, in the context of cardiovascular disorders (section 5.2.3).

Collectively, neurological disorders cause enormous suffering and inflict huge losses on society in terms of premature death and disability. By some measures, the costs are greater than those caused by cancer or cardiovascular disease. Part of the problem is that neurological disorders often afflict relatively young people who then, in the absence of cures, must live with their problems for many years (figure 6.1). As mentioned in the chapter's opening quote, some effective treatments for neurological disorders do exist, but they are generally not cures, and the rate at which new therapies are being developed is painfully slow. Many promising therapies have emerged from animal and in vitro models, but most of the clinical trials have failed.

In the following sections, I review a representative selection of neurological disorders, focusing on the major models biologists have used to study them and on the

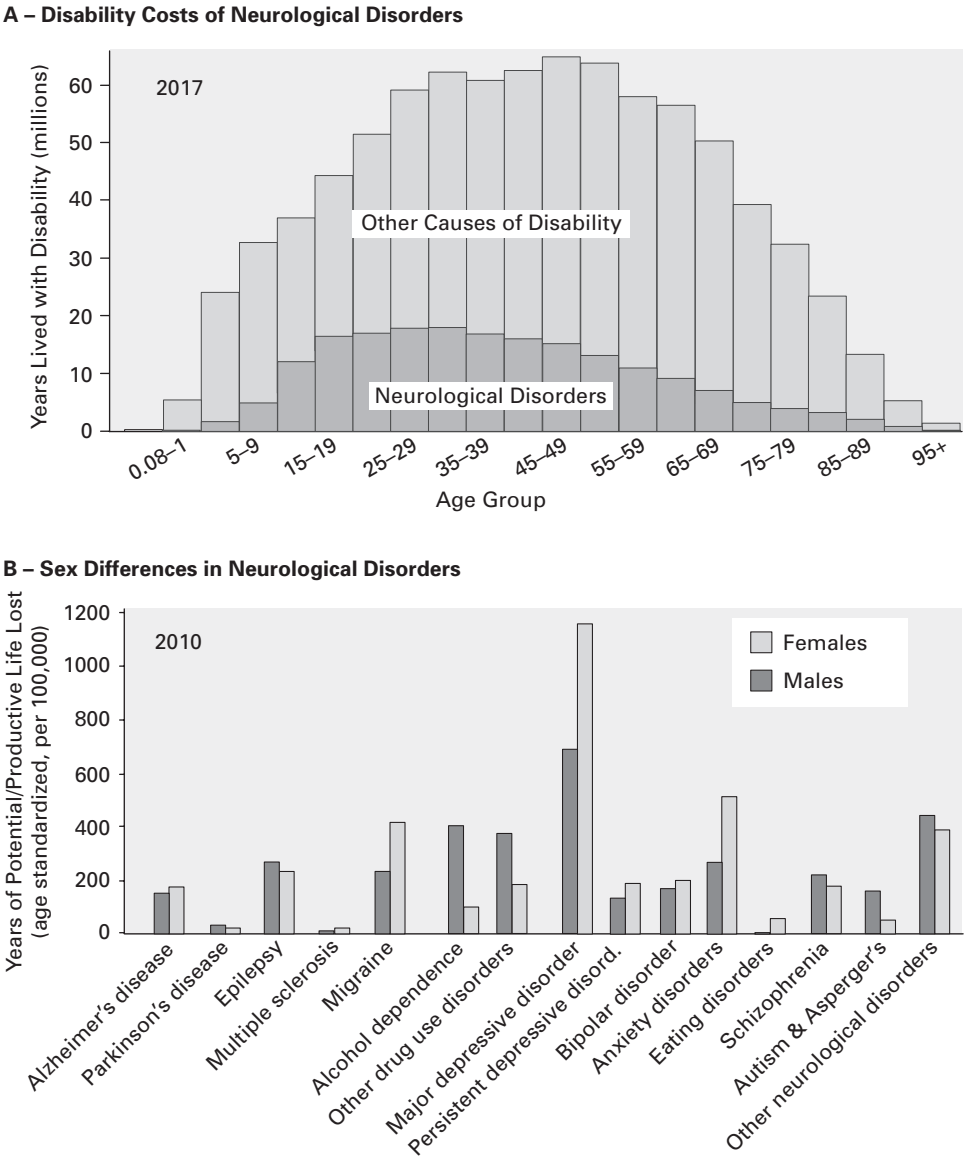


Figure 6.1

The impact of neurological disorders varies across the life span and between the sexes. (*A*) Global burden of neurological disorders (here grouped with "mental disorders" but excluding cases of CNS injury) in terms of total years lived in less-than-ideal health. The burden falls heavily on people of working age. (*B*) Histogram showing that many neurological disorders (again excluding injuries) impact the two sexes differently. The metric here is years lost to disability, either as a result of premature death or in terms of years of productive life lost due to disability (aka disability-adjusted life years, DALY). Neurological and mental disorders are combined. Adapted from (A) Institute for Health Metrics and Evaluation (2018); (B) Whiteford et al. (2015). Data reported in (A) Murray et al. (2012) and (B) James et al. (2018).

breakthrough therapies that have emerged. In addition, we will discuss some instances in which the therapies that worked in animal models have failed in human clinical trials. The chapter ends with some possible explanations for why progress on treating neurological disorders has been so slow.

6.1 BRAIN AND SPINAL CORD INJURIES

Damage to the central nervous system (CNS) typically results from trauma to the head or spine. It often involves the death of neurons at the site of injury as well as damage to the axons that course through the area of injury. In addition, CNS injury typically involves inflammatory responses that help to contain the damage in the short term but may have a variety of long-term negative effects. Because the initial neuronal losses are generally irreversible, potential therapies aim mainly to reduce the deleterious follow-on effects and to restore some of the damaged connections. Not discussed here are mechanisms of compensatory plasticity that contribute to successful rehabilitation after CNS injuries.

6.1.1 Traumatic Brain Injury

Traumatic brain injury (TBI) affects over 25 million people annually (James et al., 2019). It comes in many different forms, ranging from closed-head injury (e.g., when the head hits the windshield during an automobile accident) to penetrating injury (e.g., a bullet wound). The type of damage is also quite varied, ranging from outright neuron death to the disruption of axonal connections and widespread neuroinflammation. It is no surprise, therefore, that biologists have created a wide variety of techniques for modeling such injuries (Ma et al., 2019).

Descriptions of these methods are difficult to read without triggering strong emotional reactions because they emphasize the importance of inflicting serious damage with great precision and then measuring the behavioral sequelae as objectively as possible. Such rigor is essential to the modeling, but it nonetheless elicits ethical concerns, especially when the test animals are sentient. In addition, one must ponder whether the animals should be anesthetized during trauma induction, given that humans in traumatic accidents are usually awake. Some anesthetics are already known to modulate the severity of TBI (Wojnarowicz et al., 2017), but, as the controversy surrounding the University of Pennsylvania monkey experiments revealed (see chapter 3, section 3.6.2), the idea of intentionally damaging the brains of awake monkeys is repugnant to most. To solve these problems, researchers have explored the use of various anesthetics, and some have modeled TBI in flies (Shah et al., 2019) and other presumably "less sentient" species, or even in vitro.

Although numerous preclinical studies on TBI have yielded successful neuroprotection, clinical trials of the putative remedies have frequently failed, just as they did in the case of neuroprotection from stroke (see section 5.2.3). A systematic analysis of 191 clinical TBI trials revealed that only 20% of the interventions resulted in a positive clinical outcome, and even the most promising interventions showed poor consistency across trials (Perel et al., 2007; Bragge et al., 2016). In fact, there currently are no drug treatments approved by the US Food and Drug Administration (FDA) for TBI.

One of the most promising drugs for TBI treatment had been the steroid hormone progesterone. More than 300 studies in various rodent models of CNS injury showed progesterone to be neuroprotective, and the drug passed phase II clinical trials for safety (Stein, 2015). These positive results led to two large phase III trials, but one of them was "stopped for futility," and the other found no difference from placebo after six months. There are many possible explanations for these expensive and frustrating failures, including the possibility of suboptimal drug doses or insufficient treatment duration in the clinical trials (Howard et al., 2017). However, it is also possible that the animal models differ from the human condition in critical ways. Similar uncertainties surround the hormone erythropoietin, which has been reported to be neuroprotective after various forms of TBI in rats (Mammis et al., 2009). A meta-analysis of five clinical trials examining erythropoietin in 2016 indicated some benefits in humans, but the difference from placebo was not statistically significant (Liu et al., 2017). More recent meta-analyses that included additional trials are similarly disappointing (Lee et al., 2019). More promising are efforts to reduce brain damage by cooling the CNS soon after injury, but even those trials have yielded equivocal results (Bragge et al., 2016).

One interesting response to these frustrations has been the formation of a research consortium in which multiple laboratories coordinate their efforts while working with multiple TBI models in rats and pigs (Kochanek et al., 2018). The work with pigs is included mainly because the large brains of these animals have proportionately more white matter and more highly folded cortices than smaller brains, two features that likely influence how brains respond to dangerous impacts (Sorby-Adams et al., 2018). The hope is that this coordinated effort will lead to treatments that are effective in both rodents and pigs and thus are likely to work in humans as well. Sadly, the consortium thus far has found that their tested therapies "have produced less robust benefit than expected from the published literature upon which many of the therapies were chosen and treatment regimens designed" (Kochanek et al., 2020, p. 2357).

6.1.2 Spinal Cord Injury

Traumatic spinal cord injury (SCI) is less common than TBI, but it still affects almost 1 million new patients per year worldwide, leading to a general prevalence of 27 million

people in 2016 (James et al., 2019). It typically kills neurons whose cell body is located in the damaged portion of the spinal cord, and, at least as important, it disrupts the axons passing through that site. The loss of axons that ascend to the brain accounts for the loss of sensations below the site of injury, and the loss of descending axons causes a corresponding loss of voluntary movements. Unfortunately, the vast majority of neurons in the CNS of mammals do not divide in adulthood and, thus, cannot be replaced once they have been lost. Moreover, the axons damaged by SCI in adult mammals also do not regenerate (Huebner & Strittmatter, 2009). Therefore, a major goal for SCI therapy development is to induce axon regrowth experimentally.

In pursuit of this aim, neurobiologists have studied SCI in a variety of species. Some of the earliest research on SCI was done with dogs and cats, but since 1946 approximately 87% of the work has been performed on rodents, especially rats (Sharif-Alhoseini et al., 2017; Filipp et al., 2019). These animal models typically involve surgical exposure of the spinal cord followed by contusion (i.e., interruption of the blood supply), compression, or transection. In most animals, the damaged tissue is quickly removed by phagocytic cells, leaving a fluid-filled cavity that presents an obstacle to axon regrowth. In mice, however, the site of injury fills with a matrix of connective tissue that then contracts longitudinally so that the two ends of the injured spinal cord draw close together (Inman & Steward, 2003); this should facilitate axon regrowth in mice.

Another potentially important species difference is that the projections from the neocortex to the spinal cord are less extensive in rodents than in large-brained primates, where they extend far down the spinal cord and deep into its ventral horn (Bortoff & Strick, 1993; see Striedter, 2005). This probably explains why rodents with damage to the corticospinal tract are generally less impaired than primates with similar injuries (Nardone et al., 2017), and why mice, in stark contrast to primates, can stand within a week or two after SCI, even in the absence of axonal regeneration (Khan et al., 1999; Filipp et al., 2019). Based on these species differences, recovery from SCI should a priori be easier to accomplish in mice than in primates, but this hypothesis has not been tested directly. Be that as it may, biologists have developed several effective therapies for SCI in rodents, and some of them show promise in humans.

One form of SCI therapy is based on the discovery of molecules that inhibit axon growth and, thus, regeneration in the adult CNS. Most of the research has focused on Nogo-A, a protein that is produced by the glial cells that myelinate axons in the CNS (i.e., oligodendrocytes). Antibodies against this protein promote the sprouting and regeneration of axons after SCI in adult rats and macaque monkeys (Schnell & Schwab 1990; Freund et al., 2007). Although the degree of axonal regeneration is limited, especially in primates, the treated animals do recover at least some of their lost

motor abilities (Bregman et al., 1995; Freund et al., 2006). Based on these findings, clinical trials of anti-Nogo therapy have begun and are now in phase II.

One significant concern is that the treatments inhibit Nogo globally and may, therefore, lead to undesirable axon sprouting elsewhere in the brain (Mohammed et al., 2020). It is also becoming clear that Nogo is just one of several molecules that inhibits CNS axon growth and that its inhibition may have modest or even insignificant effects (Zheng et al., 2003). Deletion of the molecule PTEN (phosphatase and tensin homolog) in cortical neurons, for example, enhances axon sprouting after SCI much more than does deletion of Nogo (Liu et al., 2010; Geoffroy et al., 2015). Additional uncertainty is created by the observation that genetic deletion of the Nogo receptor in mice has very little effect on axon outgrowth (Zheng et al., 2005).

A second type of SCI therapy is the implantation of stem cells into or near the site of injury. The implanted cells generally do not replace the damaged cells but instead facilitate recovery by secreting various growth factors, reducing inflammation, or providing a scaffold for axon growth (Drago et al., 2013). Numerous studies supporting the general safety and efficacy of this approach have been conducted in animals, primarily rodents (Cummings et al., 2005; Keirstead et al., 2005; Iwanami et al., 2005; Ahuja et al., 2020). More than a dozen clinical trials have also been performed, although most were relatively small and none have completed phase III (Donovan & Kirshblum 2018; Zhao et al., 2019; Platt et al., 2020).

Overall, stem cell therapy does show significant promise for the treatment of SCI, but questions remain about the optimal timing of the implantation and the optimal number of implanted cells. Arguably the biggest challenge is that the various studies have used a variety of different stem cells, derived from various cellular lineages and treated with different protocols. In light of this variation, it remains unclear which types of cells are optimal, especially since even superficially similar cell types have sometimes yielded very different results in these studies (Anderson et al., 2017). Immune rejection of the implanted cells also remains a concern. Patient-derived stem cells are a potential solution to this problem, but they vary across patients (by definition). Collectively, all this variation makes one wonder whether regulatory agencies should approve induced pluripotent stem cells (iPSCs) for each patient separately, or perhaps approve a limited set of iPSC stock lines that are likely to be "safe" for different patient groups (Tsuji et al., 2019). A third possibility is to approve specific procedures for obtaining high-quality patient-derived stem cells rather than the cells themselves, but this strategy could end up being undermined by lapses in quality control (even if they are relatively rare).

Other treatments for SCI continue to be explored as well, but progress has been unsteady. For example, the drug methylprednisolone (a synthetic corticosteroid) was

approved as an anti-inflammatory drug in the 1950s, and a large clinical trial in 1990 showed that a high dose of methylprednisolone injected intravenously within 8 hours of SCI had some beneficial effects (Bracken, 1992). Based on this finding, methylprednisolone became widely accepted as the standard of care for SCI. However, several additional trials (Donovan & Kirshblum, 2018), as well as analogous studies in rats (Rabchevsky et al., 2002), have raised serious questions about this therapy's safety and efficacy (Hugenholtz, 2003; Hall & Springer, 2004).

Some of the neuroprotective treatments we discussed in the context of stroke (chapter 5) are also being tried on patients with SCI (Ulndreaj et al., 2017). A very recent but also very promising approach to treating SCI is a form of gene therapy that induces expression of a "designer cytokine" in neurons of the motor cortex, which then promotes regeneration of the corticospinal tract and behavioral recovery, at least in mice (Leibinger et al., 2021).

6.2 NEURODEGENERATIVE DISORDERS

The neurodegenerative diseases are those in which various types of neurons die progressively, unrelentingly; they include Alzheimer's, Parkinson's, and Huntington's diseases, as well as amyotrophic lateral sclerosis (ALS). Collectively, neurodegenerative diseases place an enormous burden on individuals and on society because the patients often live for many years while their condition deteriorates. Individual cases are typically characterized as being familial or sporadic, with the former being "inherited" and the latter having no obvious genetic basis, as inferred from family history or genetic screening. The familial cases are sometimes referred to as early onset, because they tend to emerge at younger ages than the sporadic cases, but logically this need not be the case. It is possible, for example, for someone to inherit the mutation that causes Huntington's disease yet develop the disease late in life.

A common feature of neurodegenerative diseases is that a subset of neurons develops intracellular aggregates of mutated and misfolded proteins, with the identity or mix of those proteins varying across the disorders. These aggregates are often considered to be harmful to the affected neurons, although some of the aggregates may be protective, and soluble forms of the mutated proteins can also be neurotoxic (Saudou et al., 1998; Krstic & Knuesel, 2013). Another common element of neurodegenerative diseases is extensive inflammation of the brain, which involves the activation of glial cells, especially microglia. These cells scan the brain for cellular damage and then respond by releasing inflammatory molecules and engulfing cellular debris. These activities are usually beneficial, but microglial cells may destroy reasonably healthy neurons if the inflammation is prolonged. Studying these complex processes in diseased

human brains is difficult because human brain tissue usually becomes available only at autopsy, after many neurons have already died. Therefore, biologists have studied neurodegeneration in diverse animal and in vitro models.

In the following sections, I review this research for the four most intensively studied neurodegenerative diseases, focusing on the models that have been used and the translational efforts. I begin with Huntington's disease, because scientists understand it better than any other neurodegenerative disorder (especially at the genetic level). Huntington's disease therefore provides a good backdrop against which to discuss the other, more mysterious forms of neurodegeneration.

6.2.1 Huntington's Disease

Huntington's disease (HD) is relatively rare, affecting 2 to 3 of every 100,000 people worldwide (Pringsheim et al., 2012). It is characterized by erratic involuntary movements, called chorea, as well as cognitive decline. Sadly, the disease is relentlessly progressive and fatal. Despite its rarity, HD is worth discussing in some depth because it is arguably "the most curable incurable brain disorder" (Wild, 2016). The reason for this optimism is that HD results from a well-described mutation in a gene called *huntingtin*, which encodes a protein called Huntingtin.

Discovered in 1993, the *huntingtin* gene contains a sequence of the nucleotides C, A, and G that is normally repeated 14 to 29 times (MacDonald et al., 1993; Lee et al., 2012). Unfortunately, this CAG repeat has a tendency to expand across generations, and when it is repeated more than 35 times, the person is very likely to develop HD. The longer the repeat, the earlier symptoms tend to appear, although other factors clearly play a role as well (figure 6.2). Importantly, the CAG expansion is inherited in an autosomal dominant fashion, meaning that a single copy of the mutant gene suffices to cause HD in the offspring.

In the brain, HD is associated with the death of neurons in the striatum, which plays a major role in action selection and movement control, but other types of neurons also die as the disease progresses (Han et al., 2010; Rüb et al., 2014). Many of the remaining neurons exhibit proteinaceous aggregates that consist mainly of mutant Huntingtin (DiFiglia et al., 1997), but the extent to which these aggregates are neurotoxic or protective remains a matter of debate (Arrasate et al., 2004; Hoffner & Djian, 2015).

Most of what we know about mutant and wild-type *huntingtin* and the Huntingtin protein comes from animal and in vitro research. By now biologists have mutated or otherwise manipulated *huntingtin* in a wide variety of cell types from many different species, including yeast, fruit flies, pigs, marmosets, and macaques (Li & Li, 2012; Marsh et al., 2012; Hofer et al., 2018). However, the vast majority of studies on mutant *huntingtin* have involved genetically modified mice (Menalled et al., 2014).

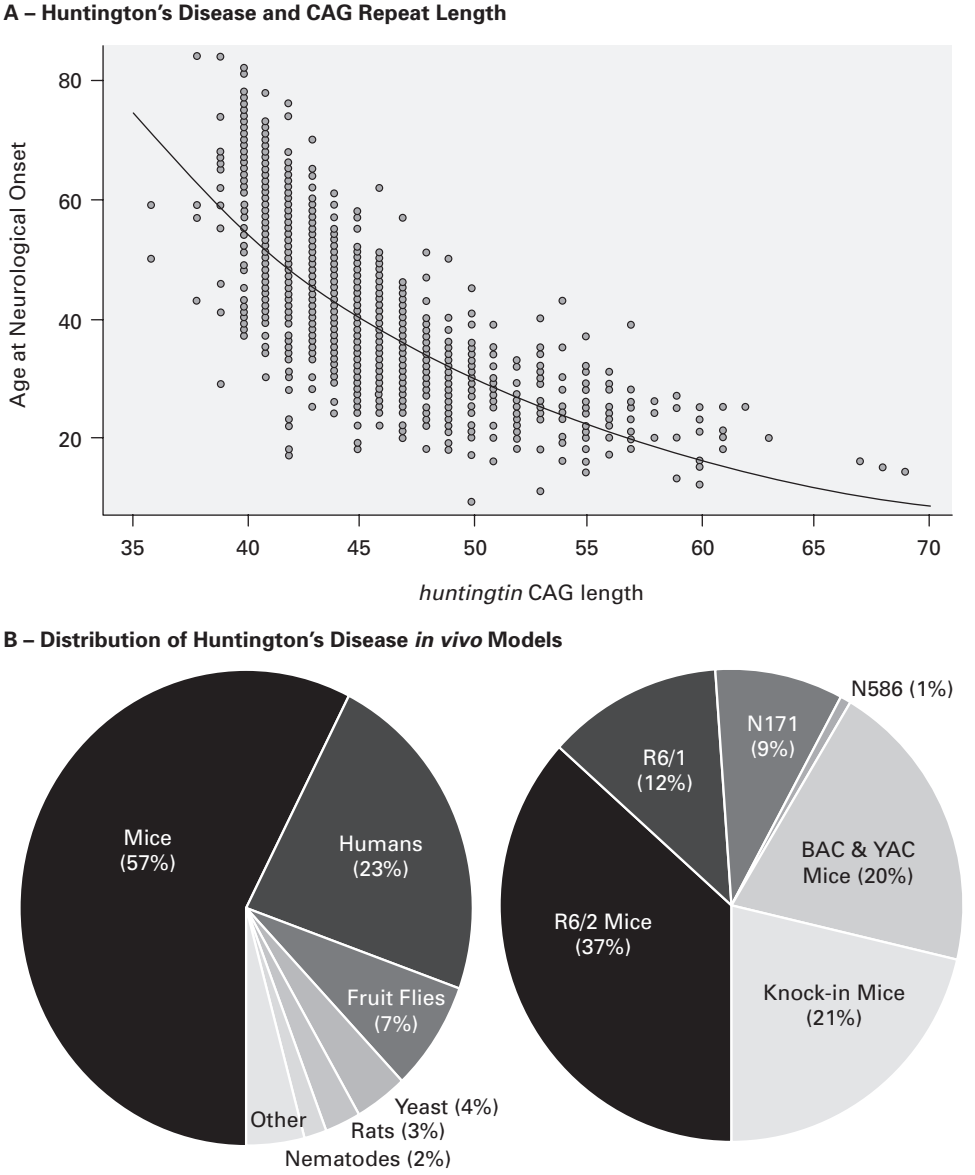


Figure 6.2

Huntington's disease (HD) in humans and in vivo models. (*A*) HD and CAG repeat length. The greater the length of the CAG repeat in the *huntingtin* gene (*htt*), the earlier humans develop the neurological symptoms of HD. Overall, CAG repeat length accounts for 67% of the variation in age at neurological onset; the remaining variation is poorly understood. (*B*) The distribution of species used to model HD in vivo (*left*) and the major genetically modified mouse models of HD (*right*), based on the author's analysis of roughly 1,000 data papers published between 1993 and 2017, retrieved through a PubMed search using the keywords "Huntington's disease gene" or "huntingtin protein" (conducted by the author in October 2018 and searching the full text for species or model mentions). Virtually all the species listed as "other" are vertebrates. The R6/1 and R6/2 mice are transgenic for a mutant form of human *huntingtin* (*mHtt*) exon-1. The BAC and YAC mouse lines express a full-length human *mHtt*, in addition to their endogenous wild-type *htt*. In contrast, the knock-in mice have one or both of their endogenous *htt* alleles modified to include an expanded CAG repeat or, in some cases, a mutant human *htt* exon-1. (A) Adapted from Gusella & MacDonald (2009).

One of the very first and most frequently studied animal models of HD is the R6/2 line of transgenic mice (figure 6.2B). In addition to their own, endogenous *huntingtin*, these animals express a mutant fragment of human *huntingtin* that was modified to carry 150 CAG repeats (Mangiarini et al., 1996). This degree of CAG expansion is extreme, relative to what is seen in humans, but it is effective for HD modeling insofar as the mutant mice rapidly develop motor symptoms and die prematurely. Indeed, these mice are widely used by researchers precisely because disease progression in these mice is rapid and reliable, which means that sample sizes can be small and publication rates high. The R6/2 mice have also been readily available from commercial sources and, by now, have been characterized in great detail (Hockly et al., 2003; Dragatsis et al., 2009; Menalled et al., 2012). Despite these advantages, R6/2 mice exhibit much less neuron death than human HD patients, develop different types of Huntingtin aggregates (André et al., 2017), and frequently die of epilepsy, which is not a common cause of death in human HD.

Since 1996, when the R6/2 mice were first produced, biologists have created many additional HD model mice that are less divergent from the human condition. These include mice that carry full-length versions of the mutant human gene instead of (rather than in addition to) their endogenous *huntingtin* and various knock-in mice that feature an expanded CAG sequence in their endogenous gene (Menalled et al., 2002; Southwell et al., 2013). An interesting aspect of these later models is that their phenotypes are generally mild and slow to develop. To counteract this problem researchers tend to use homozygous mice with CAG repeat lengths of 150 or more (Woodman et al., 2007; Southwell et al., 2017). In contrast, the vast majority of human HD patients are heterozygous for the mutation and possess CAG lengths below 65 (figure 6.2A). Thus, the efforts to create genetically "more accurate" mouse models of human HD have failed in the sense that the most accurate models are not sick enough to be convenient for laboratory research. In this context it makes sense that many researchers continue to use the R6/2 mice. The data from the more accurate models also suggest that mice are simply less sensitive than humans to the effects of mutant *huntingtin*, although this possibility is rarely discussed.

Some of the discoveries first made in HD models were later found to apply to human patients as well. Most notably, Huntingtin-containing aggregates were described in R6/2 mice before they were identified as such in HD patients (Roizin et al., 1979; DiFiglia et al., 1997). Overall, however, the data from the various HD models present a rather complex picture, replete with species, strain, and cell type differences. For instance, fragments of mutant Huntingtin tend to form insoluble aggregates, but the pattern and extent of aggregate formation vary across species, strains, and cell types (Tallaksen-Greene et al., 2005; Sawada et al., 2007; Malinovska et al.,

2015; Carty et al., 2015; Chongtham et al., 2020). Similarly, the tendency for the CAG repeats to expand somatically (i.e., in nongonadal tissues as the body develops and matures) varies across organs and within the brain across cell types; it also varies with the genetic background of the HD model mice (Telenius et al., 1994; Shelbourne et al., 2007; Gonitel et al., 2008) (figure 6.3). Unfortunately, such variation is rarely studied explicitly (Han et al., 2010; Ooi et al., 2019). In general, scientists have learned a great deal about what *can happen* when *huntingtin* is mutated, but what *actually happens* depends on the genomic and cellular context and is, therefore, not easily predictable (Bard et al., 2014).

If the preceding statement is true, then drugs that ameliorate Huntingtin toxicity in one model system might not work was well in a different model. The extent to which this is the case remains unclear because many therapies that appear promising in one model have not been tested in others. However, in the few cases where such tests have been carried out, they reveal only modest agreement (i.e., concordance) across models. For example, Varma et al. (2007) used a high-throughput screen to identify 29 compounds (out of around 43,000) that selectively reduce cell death in an immortalized striatal cell line expressing a mutant fragment of human *huntingtin*. They then tested these "hit" compounds in a second cellular HD model and in HD model worms (*Caenorhabditis elegans*); only four compounds were effective in multiple models, and only two were effective in all three. Additional testing in a rat model of HD then reduced the number of concordant hits to one, and this remaining drug was not effective in a yeast model of HD. Given this low concordance, it is not surprising that none of the four initially so promising therapies have, as far as I can tell, advanced to clinical trials. In fact, only a handful of drugs have entered clinical trials for HD (Hersch & Ferrante, 2004; Kumar et al., 2015), and virtually none have advanced to regulatory approval.

The principal exception is tetrabenazine, a drug that reversibly inhibits the uptake of dopamine (and, to a lesser extent, serotonin and norepinephrine; Paleacu, 2007). This drug was originally synthesized in the 1950s as an antischizophrenia drug; giving it to patients with HD alleviates their chorea, presumably because it depletes dopamine in neurons that project to the striatum and thus counteracts the effects of Huntington's disease on striatal function (Starr et al., 2008). Following up on this fortuitous observation, a clinical trial with 85 HD patients in 2006 revealed that tetrabenazine is relatively safe and effective (Huntington Study Group, 2006). These positive results led to US regulatory approval in 2008.

An interesting aspect of this process is that tetrabenazine had not been tested in HD models before the clinical trial because it had already been tested for safety in humans as part of the research on its potential to help schizophrenics. However, later work did show that tetrabenazine reduces symptoms in a line of HD model mice

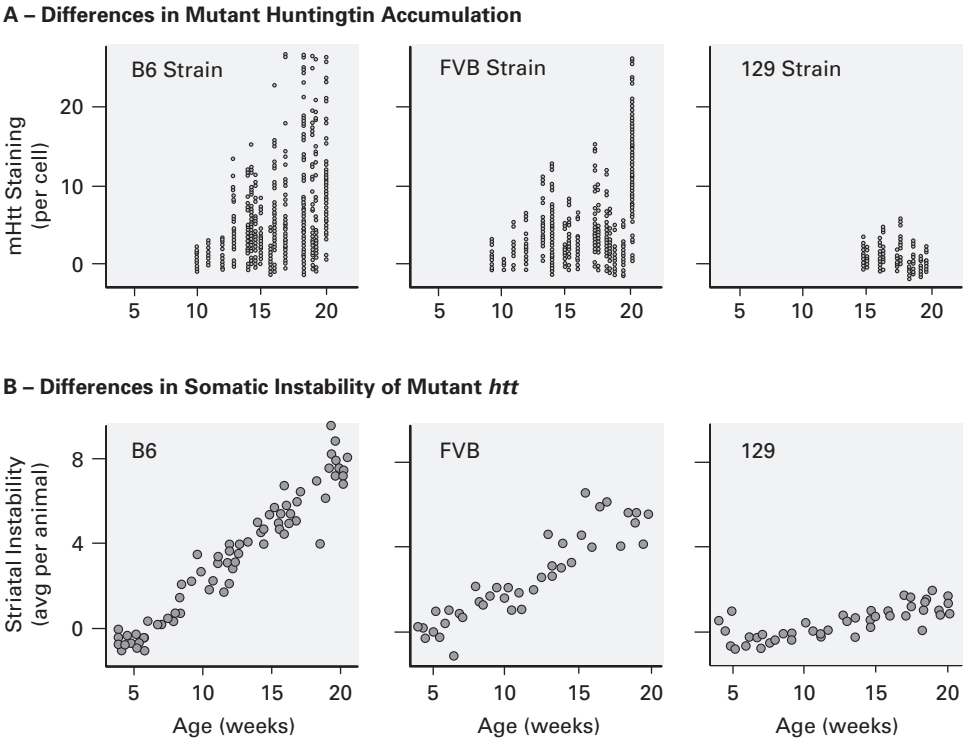


Figure 6.3

Strain differences in Huntington's disease (HD) model mice. (*A*) Differences in mutant *Huntingtin* protein (*mHtt*) accumulation. Ament et al. (2017) found that the intensity of antibody staining for mHtt in individual striatal cells increases with age in knock-in HD model mice (with 111 CAG repeats) but varies across strains. *mHtt* accumulation is fastest in the B6 (aka C57BL/6) mice and slowest in the 129 (aka 129/S2) mice. (*B*) Strain differences in *mHtt* accumulation are paralleled by differences in somatic instability, which is an increase over time in the length of the CAG repeat of the mutant *huntingtin* gene (*mHtt)*. The graphs show how the degree of somatic instability for *mHtt* increases with age in the striatum of these mice (each data point represents an averaged instability index from a single mouse). It is unclear what causes the CAG repeats to expand in these postmitotic cells, but it probably involves malfunctions of the DNA repair machinery. Adapted from Ament et al. (2017).

(Wang et al., 2010), thus retroactively enhancing this model's predictive validity (see chapter 2). Later studies also showed that a related compound called deutetrabenazine works even better because it has a longer half-life and can, therefore, be taken less frequently; it was approved by the FDA in 2017 (Dean & Sung, 2018). No other drugs have been approved for the treatment of HD, but it is instructive to examine some trials that failed even though the preclinical research had been promising.

One failed HD trial examined coenzyme Q10, an antioxidant used by mitochondria to help them generate energy. Coenzyme Q10 had exhibited neuroprotective activity in several mouse HD models (Beal et al., 1994; Ferrante et al., 2002), and

a small clinical trial showed it to be safe for human use (Feigin et al., 1996). Based on these findings, scientists in 2008 began a five-year clinical trial involving 600 HD patients. Unfortunately, the trial was halted early for futility, which means that preliminary analyses indicated a less than 5% chance of success (McGarry et al., 2017). In the meantime, a study on R6/2 mice also failed to demonstrate a positive effect for coenzyme Q10 (Menalled et al., 2010). Similar problems plagued research on creatine, which helps cells store metabolic energy. Creatine had yielded beneficial effects in three different mouse HD models, including two studies on R6/2 mice, but a clinical trial that enrolled 553 HD patients was halted for futility (Hersch et al., 2017).

A third important failed HD trial examined inhibition of sirtuin, which deacetylates various proteins. The effects of manipulating sirtuin expression had been studied in multiple HD models, including nematodes, flies, and mice, where it yielded somewhat contradictory results (Duan, 2013). In 2014, however, a sirtuin inhibitor called Selisistat was shown to be effective in both fly and mouse HD models (Smith et al., 2014). These data led to an exploratory clinical trial, which showed that Selisistat was safe (Süssmuth et al., 2015). Sadly, a larger subsequent trial revealed no statistically significant effects in key parameters, which caused the trial sponsor to halt the drug's development (Haider, 2018).

These trials may have failed because the test subjects were so far along in the disease that many neurons were irreversibly lost. It is also possible, however, that the HD model mice and flies are just too different from human HD patients to have strong predictive validity. To address this potential problem, it might be useful to examine monkeys expressing mutant *huntingtin*, as they might share with humans some features that are not shared by nonprimates. Indeed, researchers have used viruses to transfect mutant *huntingtin* fragments into the striatum of macaques (Palfi et al., 2007); a transgenic line of monkeys has also been created (Yang et al., 2008; Chan et al., 2014). These nonhuman primate models of HD have yielded some interesting results (Emborg, 2017), but most of the transgenic monkeys died long before they could reproduce.

An alternative approach is to focus more energy on the study of human cells, especially iPSCs derived from HD patients (HD iPSC Consortium, 2012; An et al., 2012; Csobonyeiova et al., 2020). Unfortunately, a major limitation of such studies is that they generally fail to replicate the in vivo interactions between different cell types (Creus-Muncunill & Ehrlich, 2019). This could be important because it is already known, for example, that mutant Huntingtin can affect the ability of striatal neurons to obtain important growth factors from cortical neurons and cholesterol from astrocytes (Zuccato et al., 2001; Valenza et al., 2015). In addition, the axonal connections between the cortex and the striatum are known to undergo significant changes as HD

progresses (Raymond et al., 2011; Strong et al., 2012). Co-culture systems and organoids (see chapter 4) may be able to mitigate these limitations, but their development as HD models is only just beginning.

Meanwhile, recent efforts to develop HD therapies have focused mainly on suppressing *huntingtin* expression by means of synthetic RNAs and other oligonucleotides. Some of these approaches suppress both the mutant and the wild-type *huntingtin* allele (Carroll et al., 2011), which could be problematic because the wild-type protein performs a large variety of potentially important intracellular functions. Fortunately, studies in mice and nonhuman primates suggest that this will not be a significant problem (Drouet et al., 2009; Kordasiewicz et al., 2012; Stanek et al., 2014), and tests in a few humans have so far revealed no negative effects (Tabrizi et al., 2019); a large phase III trial is under way. A second concern is the possibility of off-target effects, which is to say the unintended suppression of other genes (Monteys et al., 2014). A third open question is how long the suppression will last, although using viruses to transfect the RNA constructs into the target cells should have quite durable effects (Wild & Tabrizi, 2017).

One way to avoid several of these problems is to use genome editing techniques to "correct" the *huntingtin* mutation in the most vulnerable cells (Monteys et al., 2017), but this genome editing approach is still in the preclinical stage. Overall, the history of HD therapy development offers both hope and lessons in humility.

6.2.2 Amyotrophic Lateral Sclerosis

Another intensively studied neurodegenerative disease is amyotrophic lateral sclerosis (ALS), which in 2014 affected about one in every 20,000 people in the United States (Mehta et al., 2018). It is also known as Lou Gehrig's disease, after a widely admired American baseball player who died from the disease in 1941. The principal symptom of ALS is progressive paralysis, caused by the degeneration of motor neurons, first in spinal cord but then also in the cerebral cortex. The disease typically manifests when people reach their 50s, although Lou Gehrig began to show symptoms at the age of 36.

Five percent to 10% of cases are classified as familial. Of those, 12% to 23% can be linked to mutations in a gene called *superoxide dismutase-1* (*SOD1*) (Andersen 2006; Tan et al., 2017; Gill et al., 2019). Another 16% of the familial cases have been linked to mutations in a gene called *C9orf72,* which also account for 6% to 8% of the sporadic ALS cases (presumably through mutations that were not inherited). In addition to *SOD1* and *C9orf72*, ALS has been linked to a handful of other genes, including *TDP43*, but mutations in these genes are even less common. Remarkably, many of the ASL-linked proteins colocalize with one another in the intraneuronal aggregates characteristic of ALS, even if those proteins are wild type (i.e., not mutated). Whether

these aggregates are neuroprotective or toxic remains, as in the case of Huntington's disease, a matter of debate (Strong et al., 2005; Blokhuis et al., 2013).

To study the mechanisms underlying ALS and seek potential therapies, biologists have created a variety of transgenic mice, most of which express mutant forms of human *SOD1*. In these mice the transgene is usually inserted as multiple copies (as many as 25), which is quite different from the human condition. Moreover, transgene expression is abnormally high in most SOD1 mouse lines, and the degree of motor neuron degeneration and paralysis covaries with the level of mutant *SOD1* activity (Gurney et al., 1994; Gurney 1997; Dal Canto & Gurney, 1997). Specifically, the mice with the highest number of transgene copies and highest levels of *SOD1* activity die earliest and exhibit some pathologies not seen in human ALS. A related problem is that copy numbers and gene activity levels also vary between supposedly similar mouse lines, and most studies fail to specify those crucial aspects of their research animals (Zwiegers et al., 2014). A more recently developed mouse ALS model expresses mutant versions of human *TDP43* and exhibits progressive symptoms similar to those of human ALS (Wegorzewska et al., 2009). However, further study of these mice did not replicate some aspects of the earlier reports (Perrin, 2014) and revealed that these mice actually die from acute bowel obstruction, rather than skeletal muscle wasting (Esmaeili et al., 2013). Thus, the principal mouse models of ALS are not as similar to human ALS as one might hope. Some non-mammalian animal models of ALS have been created over the years, but their contributions to ALS research have thus far been minor (Patten et al., 2016).

In vitro models of ALS have historically involved a variety of immortal cell lines, but these tend to be quite different from the motor neurons that selectively die in ALS. Many researchers have, therefore, turned their attention to the study of human iPSCs that can be differentiated into motor neurons. Such studies have revealed, for example, that iPSC-derived motor neurons are more sensitive than some other types of neurons to *SOD1* aggregates (Benkler et al., 2018). Another study found that iPSC-derived motor neurons from a patient with sporadic ALS developed *TDP43*-positive aggregates that were quite similar to those observed in the patient at autopsy (Burkhardt et al., 2013). A key advantage of these iPSC models is that the cells can be subjected to high-throughput screens for drugs that modulate aggregate formation or cell death.

However, iPSC-based models also have limitations (Preza et al., 2016; Volpato & Webber, 2020) (see also section 4.2.5). As already noted in the section on Huntington's disease, patient-derived iPSCs vary genetically between patients. The methods used to reprogram adult cells into stem cells add to this heterogeneity because they alter the cells' genome and epigenome in ways that are not entirely predictable. The protocols used to differentiate iPSCs also vary between laboratories, and any given

protocol is likely to produce cells that differ from the in vivo target cells, at least in maturity. Finally, when iPSCs are cultured for several generations, they tend to accumulate mutations that adapt them to the culture conditions.

The development of therapies for ALS has been a frustrating affair. More than 70 compounds have been reported to have beneficial effects in SOD1 model mice, but a meta-analysis has indicated that many of these studies were "methodologically poor" and subject to publication bias, meaning that negative results tended not to be published (Benatar, 2007). A nonprofit biotech company then tested dozens of promising drugs in thousands of SOD1 mice, but was unable to replicate the previously published results (Scott et al., 2008; Perrin, 2014) (see figure 1.1B). Even riluzole, which had long been the only drug approved for the treatment of ALS, does not appear to be effective in SOD1 mice, even though it has at least a modest beneficial effect in ALS patients. In retrospect, this is perhaps not surprising, given that the SOD1 mice generally express so much mutant *SOD1* that any treatments for them would have to be very robust (van der Worp et al., 2010). Lithium, which is often used to treat bipolar disorder (see section 6.3.2), did show promising effects in SOD1 mice and even in a pilot study of ALS patients (Fornai et al., 2008). These findings prompted three larger clinical trials on the effects of lithium on ALS, but all of them failed (Petrov et al., 2017).

In spite of all the frustration, ALS therapy development is not hopeless. Particularly promising is that the antioxidant edaravone was recently approved for ALS in both Japan and the United States. European regulators did have some concerns, however, and the drug's producer responded by withdrawing its application. Perhaps the costs of obtaining the requested evidence could not be justified, given the relative rarity of ALS. Another promising drug is masitinib, an anti-inflammatory kinase inhibitor that suppresses microglia. It can reduce some toxic interactions between microglia and motor neurons in vitro and is beneficial in SOD1 rats, reducing symptoms even when it is administered after the onset of paralysis (Trias et al., 2016). Masitinib (in combination with riluzole) passed a phase II/III trial in 2017, and large phase III trial is currently underway.

A third intriguing advance is a recent phase II trial of a drug cocktail that reduces cellular stress and seems to slow ALS progression, at least slightly (Paganoni et al., 2020). It is worth noting that this trial is the first to be funded through the popular Ice Bucket Challenge (2021) to support ALS research.

6.2.3 Alzheimer's Disease

Alzheimer's disease (AD) is by far the most common neurodegenerative disorder, with roughly 6 million patients currently in the United States alone and affecting roughly 10% of people older than 65 (Alzheimer's Association, 2019). Its hallmark symptom

is progressive memory loss, but late stages of AD entail a general cognitive decline. Delineating AD from other age-related dementias is difficult, and a definitive diagnosis still requires postmortem examination (McKhann et al., 2011).

Within the brain, AD manifests as a progressive loss of neurons and synaptic connections, especially in the hippocampus and adjacent cortical areas. It also involves the degeneration of forebrain neurons that use acetylcholine as their neurotransmitter and project to the cortex. Another common feature of AD is the formation of two types of insoluble protein aggregates. Some aggregates consist mainly of beta-amyloid; these are called amyloid plaques. The other principal aggregates in AD brains are neurofibrillary tangles, which consist of twisted strands of a protein called tau.

Familial cases of AD have been linked to mutations in amyloid precursor protein (APP) and presenilin, which is involved in APP processing. These findings helped to spawn the amyloid cascade hypothesis, which posits that the primary cause of AD is the formation of beta-amyloid from APP, and that this beta-amyloid then accumulates in the form of amyloid plaques, damages neurons, and somehow triggers the formation of neurofibrillary tangles (Hardy & Higgins, 1992). Although most of the research on AD is based on this hypothesis (Herrup, 2015), some AD patients have few amyloid plaques, while others possess numerous plaques but relatively intact memory (Dickson et al., 1992; Giannakopoulos et al., 2003). Moreover, familial AD accounts for only about 1% of all AD cases; among these, only about 5% are linked to mutations in APP or presenilin (Drummond & Wisniewski, 2017; Mullane & Williams, 2019).

This leaves most cases of AD without a clear genetic basis. Numerous genetic and environmental risk factors for AD have now been identified (including old age, stress, genetic variation in apolipoprotein E, and a lack of mental and physical exercise), but synthesizing all this information into a coherent mechanistic picture has remained a serious challenge (Medina et al., 2017). In spite of this complexity, most of the preclinical AD research has focused on models that express mutant APP and presenilin. The hope has been that these genetic models can provide insights and yield potential therapies that will translate to AD resulting from other, more mysterious causes.

The vast majority of AD models are transgenic mice that express mutant forms of human APP (Jankowsky & Zheng, 2017; Mullane & Williams, 2019). Most of these mice display some AD-like symptoms, including amyloid deposition in the brain and reduced performance on some memory tasks. However, the mice generally express abnormally high levels of the transgene and fail to develop neurofibrillary tangles or extensive neuron loss. To overcome this limitation, scientists have created mice that combine as many as five different mutations in two or even three different AD-linked genes, including *presenilin* and *tau* (Oddo et al., 2003; Oakley et al., 2006). Triple transgenic mice do exhibit neurofibrillary tangles, but neuron loss is still quite limited.

Moreover, these mice are poor genetic models of the human disorder insofar as AD patients usually carry just a single mutation, in either APP *or* presenilin, and mutations in *tau* are generally associated with frontotemporal dementia rather than AD. Given these differences, it is not surprising that the changes in neuronal gene expression observed in AD model mice (compared with wild-type mice) tend to be quite different from those observed in humans with AD (Burns et al., 2015; Hargis & Blalock, 2017) (figure 6.4).

A newer, second-generation of mouse AD models modifies the endogenous AD-linked genes of mice rather than inserting additional human genes (Reaume et al., 1996; Baglietto-Vargas et al., 2021). These knock-in mice avoid the overexpression and random transgene insertion problems of the earlier models; but, as in the case of Huntington's disease (see section 6.2.1), their phenotypes tend to be much more subtle. To solve this problem, researchers have created knock-in mice that carry multiple mutations (Siman et al., 2000; Flood et al., 2002; Saito et al., 2014); unfortunately, such combinatorial models sacrifice fidelity at the genetic level for similarity at the neurobiological and behavioral levels.

The observation of milder phenotypes in knock-in AD model mice suggests that mice may be less vulnerable than humans to the effects of AD-linked mutations. This hypothesis is supported by some evidence (Rosen et al., 2016; Espuny-Camacho et al., 2017), but it has not been tested thoroughly. A potentially related observation is that beta-amyloid aggregates in mice tend to be more soluble than their human counterparts, have different binding properties, and generate less inflammation (McGeer, 2003; Jucker, 2010; Drummond & Wisniewski, 2017). Species differences in the *tau* gene may also play a role (Umeda et al., 2014). Alternatively, mice may simply not live long enough to develop AD unless the underlying mechanisms are artificially enhanced. The problem with this proposition is that life span generally scales with body size (Speakman 2005), and mice do age much faster than humans. The transcriptional signature of brain aging, for example, is quite similar between mice and humans despite the differences in chronological age (Burns et al., 2015).

The hypothesis that humans are exceptionally vulnerable to AD is consistent with the observation that few other species naturally develop this disease (Walker & Jucker, 2017). Several nonhuman primate species do accumulate amyloid deposits in their brains as they age, and very old macaques and chimpanzees exhibit neurofibrillary tangles (Edler et al., 2017; Paspalas et al., 2018), but whether these symptoms are associated with neuron loss remains unclear. Some very old mouse lemurs (i.e., strepsirrhine primates) exhibit extensive neurodegeneration, but their pathology is focused on the frontal lobes rather than the temporal lobe (where cell death is most prominent in human AD) (Pifferi et al., 2019).

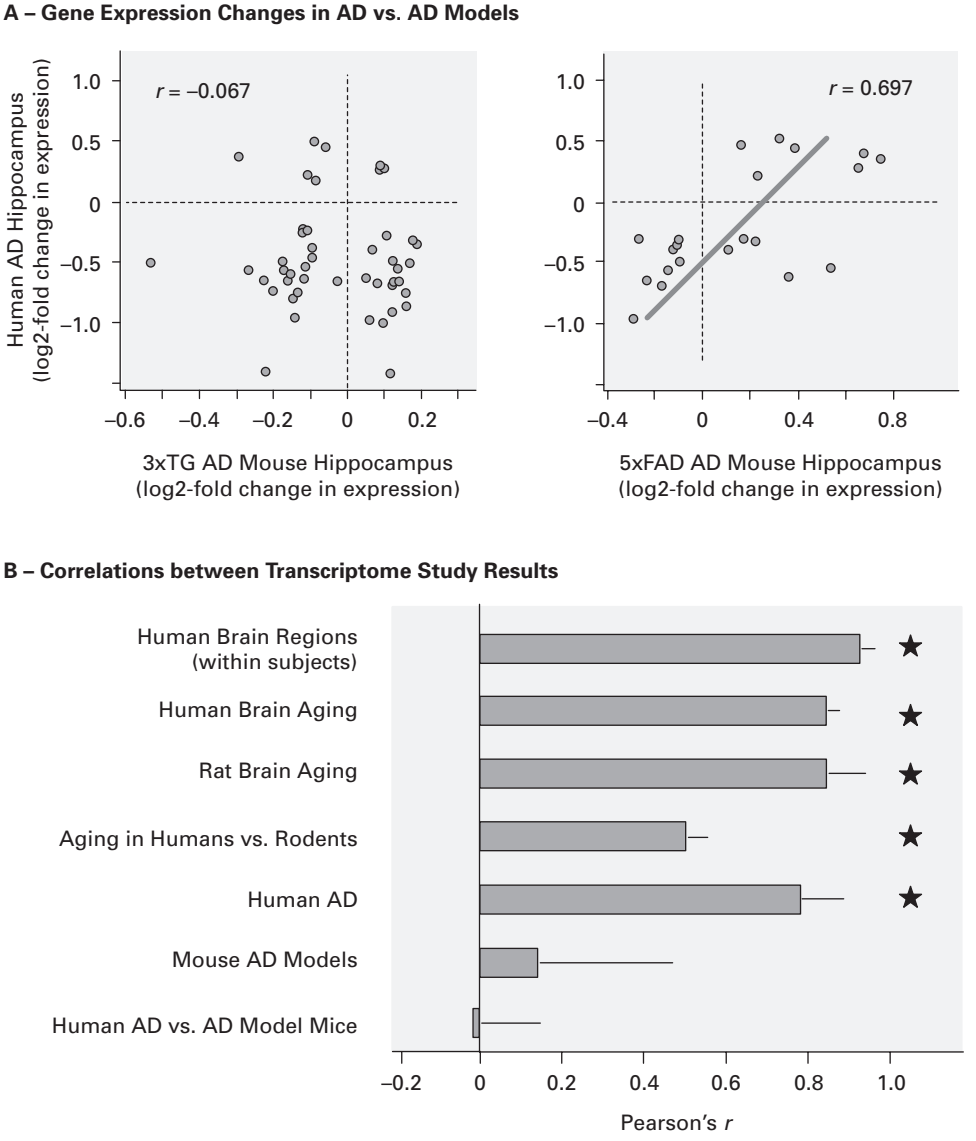


Figure 6.4

Transcriptomic comparisons between human Alzheimer's disease (AD) and mouse AD models. (*A*) Gene expression changes in AD versus AD models. Hargis and Blalock (2017) compared the transcriptional profiles obtained from the hippocampus of human AD patients to those obtained from the hippocampus of two transgenic mouse AD models (3xTG—triple transgenic mice; 5xFAD—mice carrying five mutations linked to familial AD). Each data point represents a single gene that is differentially expressed in both species. Whereas the transcriptional profile of the 5xFAD model correlates positively with that of human AD, the profile of the 3xTG model does not. (*B*) Correlations between transcriptome study results. Based on many such pairwise correlation analyses, Hargis and Blalock determined that the transcriptional profiles for young versus old rodents correlates pretty well with the profiles associated with human aging (stars indicate a statistically significant correlation), but that the mouse AD models correlate poorly with one another and, on average, are a poor reflection of human AD. Adapted from Hargis and Blalock (2017).

Outside of primates, AD-like phenotypes are rare. Some dogs do develop senile dementia and AD-like pathology, but they lack neurofibrillary tangles (Head 2013; Prpar Mihevc & Majdič, 2019). Structures similar to tangles have been reported in the brains of sheep, goats, and leopard cats, but whether these animals develop other AD-like symptoms remains unclear (Braak et al., 1994; Nelson et al., 1994; Chambers et al., 2012). Rats and mice do not normally develop signs of AD, but common degus (which are also rodents) do develop AD-like cognitive deficits, along with amyloid deposits and neurofibrillary tangles, when they get very old (Hurley et al., 2018). Of all these species, the degus may be the ethically most acceptable genetically natural AD model to work with. However, work with this model would be slow, because only a subset of individuals develop the disease and symptoms take years to emerge (much as in humans).

Given the challenges of working with natural AD models, it is not surprising that the development of AD therapies has relied mainly on the genetically modified mouse models. More than 200 different therapies have been reported as being effective in mouse AD models. Unfortunately, of 413 clinical trials examining 244 such compounds between 2002 and 2012, only one passed through to marketing approval (Cummings et al., 2014). The years since then have only produced additional failures, prompting some drug companies to withdraw from neurological research entirely. The one success during this period was memantine, an uncompetitive *N*-methyl-D-aspartic acid (NMDA) receptor blocker that improves some cognitive and behavioral symptoms of AD but has not been shown to delay progression of the underlying disease (Thomas & Grossberg, 2009). Intriguingly, memantine had been used since 1982 to treat a variety of other brain disorders, especially Parkinson's disease, but was found to be effective for AD in the late 1980s (Parsons et al., 1999). Subsequent studies then showed that it is also effective in some mouse AD models (Martinez-Coria et al., 2010).

The only other drugs that have received regular approval for AD are inhibitors of acetylcholinesterase (AChE), an enzyme that is critical for the synthesis of acetylcholine. These AChE inhibitors (e.g., donepezil) had long been used to boost cognition in humans, but their use for the treatment of AD accelerated after scientists discovered that cholinergic neurons (i.e., neurons that use acetylcholine as their transmitter) die relatively early in AD (Summers et al., 1986; Contestabile, 2011). Although these drugs modestly boost cognition in late stages of AD, they (like memantine) do not delay disease progression (Geerts, 2009; Cooper et al., 2013).

Of the clinical AD trials that failed, most involved drugs that were designed to reduce the synthesis of beta-amyloid or increase its clearance. Often this entailed either vaccination with beta-amyloid or the injection of synthetic antibodies against this peptide (Nicoll et al., 2003; McGeer, 2003). Although the tested drugs had been effective

in AD model mice (Schenk et al., 1999; Oddo et al., 2004) and generally reduced levels of beta-amyloid in humans, they ultimately failed; several did so repeatedly (e.g., solanezumab). It is possible that the clinical trials failed because the AD patients were already too far advanced in their disease for the drugs to make much difference. However, several recent trials administered the putative therapies to patients with mild cognitive impairment, which can be a sign of incipient AD; sadly, they failed as well (Lowe, 2020; Bhandari 2020). A lot of hope is currently riding on aducanumab (aka Aduhelm™), another human antibody against beta-amyloid. Although aducanumab apparently did not perform entirely as hoped, it may have limited benefits. (Indeed, aducanumab received "accelerated approval" from the FDA in the summer of 2021, after this book had gone to press. However, this approval was an unusually controversial decision, going against the recommendation of the FDA's own advisory committee, and called for a post-approval, phase 4 confirmatory trial [Cavazzoni, 2021; Steinbrook, 2021].)

In light of these observations, alternatives to the amyloid cascade hypothesis should be considered carefully (Mullane & Williams, 2013). For example, it is interesting to note that mice expressing function-blocking antibodies against nerve growth factor (NGF) develop amyloid plaques, neurofibrillary tangles, and extensive cell loss as they get old (Capsoni et al., 2000). This process probably involves the cholinergic neurons that die in AD, because those neurons require NGF for survival. This unusual AD model is not entirely incompatible with the amyloid cascade hypothesis, but it suggests that modulating growth factors may be more effective in the treatment of AD than manipulating levels of beta-amyloid (Blurton-Jones et al., 2009). Of course, clinical trial failures may also stem from issues with the quality of the preclinical AD research, which does appear to suffer from systemic publication bias and poor, inconsistent methodology (Egan et al., 2016).

6.2.4 Parkinson's Disease

Parkinson's disease (PD) is the second most common neurodegenerative disease, affecting roughly 1% of the population in industrialized countries by age 65 and 4% to 5% by age 85 (Eriksen et al., 2003; de Lau & Breteler, 2006). Its principal symptoms are body rigidity, slow movements, a flexed body posture, tremor of the hands, and sleep disturbances. Within the brain, the principal correlate of PD is the progressive degeneration of dopaminergic neurons in the substantia nigra (SN), which project heavily to the striatum (Bové & Perier, 2012). The damaged SN neurons, as well as some less vulnerable neurons in other brain regions, tend to contain Lewy bodies, which are intracellular aggregates of various proteins, including alpha-synuclein (Spillantini et al., 1997).

Approximately 10% of all PD cases are classified as familial, and 30% of these can be linked to mutations in a specific gene. One of these is the gene that encodes

alpha-synuclein (aSyn), which is a major component of the Lewy bodies mentioned in the previous paragraph. Subsequent research revealed that even overexpression of wild-type *aSyn* can trigger the formation of Lewy bodies and that aSyn accumulations are linked to increased cellular stress, but precisely how aSyn damages neurons remains debatable. A second gene linked to PD is *leucine-rich repeat kinase 2* (*LRRK2*; aka *PARK8*). Mutations in this gene are found in a greater percentage of PD patients than mutations in a*Syn* (Klein & Westenberger, 2012), but the problems caused by these mutations are likewise obscure (Tolosa et al., 2020). The mechanisms underlying sporadic PD apparently include various pesticides (e.g., paraquat and rotenone) and toxins that invade the nervous system via the gut. However, these nongenetic causes remain poorly understood and somewhat controversial (Ascherio & Schwarzschild, 2016; Rietdijk et al., 2017; Parashar & Udayabanu, 2017).

Researchers have created a variety of transgenic mice expressing mutant versions of human *aSyn* or overexpressing wild-type *aSyn* (Chesselet et al., 2012). They have also created mice that express mutant *LRRK2* (Jagmag et al., 2016). These genetically modified PD models do show some PD-like symptoms, including a reduction of striatal dopamine levels, but neuron loss in the SN is minor at best (Fleming et al., 2005; Potashkin et al., 2010).

Greater model fidelity has been achieved with a nongenetic model that was discovered by accident. A rogue graduate student in 1976 accidentally injected himself with a bad batch of a synthetic opioid and rapidly developed symptoms of advanced PD, including the loss of dopaminergic neurons (Davis et al., 1979). A similar incident with several heroin addicts in San Francisco produced analogous results and led to the identification of a putative chemical culprit called MPTP (1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine; Langston et al., 1983). To the initial consternation of scientists, giving MPTP to rats did not cause PD-like symptoms, but subsequent experiments in macaques and several other nonhuman primates did kill the dopamine neurons and cause PD-like behavioral symptoms (Burns et al., 1983; Bové & Perier, 2012; Blesa et al., 2018). Further studies then clarified that rats and a few strains of mice are less susceptible than primates to MPTP poisoning, for reasons that still remain somewhat unclear (Giovanni et al., 1994; Hamre et al., 1999; Smeyne et al., 2005; Kasahara et al., 2013).

Biologists have also created non-mammalian and in vitro models of PD (Delenclos et al., 2019). They have, for example, expressed various PD-linked genes in fruit flies and observed that such flies exhibit protein aggregates resembling Lewy bodies, a progressive degeneration of dopaminergic neurons, and a loss of motor control (Feany & Bender, 2000). These and other non-mammalian PD models (Pienaar et al., 2010) were developed with an eye toward high-throughput drug screening. A yeast PD model, for instance, has been used in a large genetic screen to identify 55 genes

that increase human *aSyn* toxicity and 29 that decrease it (Yeger-Lotem et al., 2009). Because many of the yeast genes have human homologs, these data may reveal useful information about the mechanisms that lead from *aSyn* to cellular toxicity. However, as mentioned previously, molecular homology need not imply conservation of molecular interactions. Moreover, yeast lack an endogenous *aSyn* homolog (Pienaar et al., 2010). One might argue that this actually makes yeast an ideal cellular "clean room" in which to study the functions of human *aSyn* without the confounds caused by an endogenous homolog (Hofer et al., 2018), but that clean room is surely quite different from *aSyn*'s cellular context in humans. We will come back to this in chapter 7.

The most common and oldest PD therapy is the drug L-dopa, which easily traverses the blood-brain barrier and is converted to dopamine within the brain. Once in the brain, L-dopa causes any remaining dopaminergic neurons to synthesize and release more dopamine, which ameliorates the symptoms of PD but does not slow the neurodegeneration. The animal research that revealed L-dopa's effectiveness in the late 1950s (Carlsson et al., 1957) relied on mice that had been given reserpine, which rapidly depletes the brain of dopamine and several other neurotransmitters (notably serotonin and norepinephrine). This reserpine model of PD fell out of favor as the transgenic and MPTP-based models emerged, but it predicted L-dopa's efficacy in humans quite well (Leal et al., 2016). Unfortunately, long-term use of L-dopa generally requires increases in dosage over time and can have negative side effects, including uncontrolled movements and impulsive behavior (Ahlskog & Muenter, 2001).

The principal treatment for L-dopa-resistant PD is deep brain stimulation (DBS). DBS involves implanting an electrode deep in the human brain and applying electrical stimuli to the targeted area. The method was originally developed to address diverse ailments (notably chronic pain), but two important refinements made it highly effective for PD (Gardner, 2013). First, scientists realized that high-frequency stimulation suppresses neuronal activity for some time afterward, rather than increasing it (as one might intuitively expect). Second, researchers focused their stimulation on the subthalamic nucleus, where DBS is most effective for relieving the motor symptoms of PD (Nashold & Slaughter, 1969; Benabid et al., 2009). Some observers have argued that these refinements could have been achieved without the benefit of prior animal research (Bailey, 2014), but the scientists most responsible for developing DBS into a PD therapy have argued forcefully that their approach was inspired by the earlier discovery that MPTP-treated monkeys exhibit a dramatic increase in subthalamic activity (Benabid et al., 2015a, 2015b; Blesa et al., 2018). If the loss of dopaminergic neurons increases activity in the subthalamic nucleus, they argued, then it stands to reason that lesioning or inhibiting this nucleus might reverse or alleviate some of the MPTP-associated symptoms. This was indeed the case (Bergman et al., 1990; Benazzouz et al.,

1993), and human trials soon followed. DBS was FDA approved as a general PD treatment in 2002 and, by 2013, it had been administered in more than 40,000 PD patients (Gardner, 2013). Closely related to DBS is the idea of using gene therapy to suppress subthalamic neurons, but this approach is still under development (Bartus et al., 2014).

Physicians have also tried to treat PD with cell implants. Beginning in the 1970s, scientists grafted cells from the ventral midbrain of very young embryos (where the substantia nigra usually develops) into the striatum of PD model rats, monkeys, and eventually PD patients (Perlow et al., 1979; Sladek et al., 1986; Hagell et al., 1999). They observed good survival and integration of the implanted cells, increased levels of dopamine, and long-term improvements in behavior (W. Li et al., 2016). However, two clinical trials in 2001 and 2003 failed to show the hoped-for clinical benefit, and some subjects actually developed motor problems that probably stemmed from excessive uncontrolled dopamine release (Olanow et al., 2003). These failures dampened enthusiasm for this form of PD therapy, but advances in stem cell biology have now rekindled hope. Specifically, researchers have begun to implant stem cell–derived dopaminergic neurons into the striatum of various PD models (Kriks et al., 2011; Kikuchi et al., 2017; Fan et al., 2020). The results have been promising, and new clinical trials are underway (Barker & TRANSEURO Consortium, 2019).

6.3 NEUROPSYCHIATRIC DISORDERS

Many neurological disorders are not accompanied by neuronal injury or degeneration. People tend to refer to them as behavioral, mental, or psychiatric disorders, but given that the brain is the principal organ of behavior control and thought, they can also be described as "brain disorders" (Insel et al., 2010). Delineating these disorders from one another is difficult because their symptoms tend to overlap and their genetic links remain largely obscure (Nestler & Hyman, 2010; Kas et al., 2007; Cuthbert & Insel, 2013). Definitions aside, neuropsychiatric disorders are real and cause enormous suffering. We here focus on just three of these disorders, namely major depressive disorder, bipolar disorder, and schizophrenia. They are all among the 10 most burdensome psychiatric disorders in terms of years living with disability (see figure 6.1).

6.3.1 Major Depressive Disorder

A key feature of major depression is that the depressive symptoms last for two or more weeks and seriously impair the patient's ability to work or fully function socially. Sadly, it afflicts roughly 14% of people in the United States at some point in their adult life (with women being more susceptible) (Kessler et al., 2012). Scientists used

to think major depression is caused by a depletion of the brain's monoamine transmitters (which include dopamine, norepinephrine, and serotonin), but more recent work has called that relatively simple hypothesis into question without providing a new and widely accepted alternative. Genetic analysis linked 15 different genes to major depression in 2016, and a more recent, larger study identified 200 such genes (Hyde et al., 2016; Cai et al., 2020). However, all these genetic linkages are weak, meaning that mutations in each gene only slightly increase the risk of major depression. Because scientists know so little about the biological mechanisms underlying depression, finding or creating models for this debilitating disease is hampered by uncertainties.

Drugs that deplete one or more of the monoamine transmitters (e.g., reserpine) can induce an inability to express pleasure and other depression-like symptoms in animals (McKinney & Bunney, 1969). So can injections of lipopolysaccharides (LPS), which cause inflammation not just in the body (see chapter 5, section 5.1.4) but also in the brain (Yirmiya, 1996). Another way to trigger depression-like symptoms in rodents is to surgically remove their olfactory bulb, which deprives these animals of their main means for sensing external stimuli (Kelly et al., 1997). Although many of the symptoms emerging in these models can be alleviated by some known anti-depressant drugs (to be discussed shortly), the various drug- and surgery-induced models of depression do not provide a coherent picture of the mechanisms underlying major depression.

More promising are efforts to induce depression-like symptoms though manipulation of an animal's environment. Isolating individuals of highly social species (e.g., parrots or gorillas) can cause them to appear profoundly unhappy for long periods of time and to deteriorate physically (McKinney & Bunney, 1969). Similar effects can be achieved by exposing animals to other stressors, such as a more dominant "bully," repeated electric shocks, physical restraint, or wet bedding (Willner, 1984). These manipulations cause depression-like symptoms especially if they are unpredictable, inescapable, and combined with one another. Although such models of depression are euphemistically called "mild chronic stress" (Willner, 2005), they typically cause the animals to appear as if they had given up and lost interest in things they used to like. To quantify such changes in behavior, scientists have developed various tests that measure, for example, how long an animal will struggle when suspended by its tail or whether it prefers to drink sugar water over the usual fare; it is important, however, not to conflate these tests with the depression-inducing manipulations (Kalueff et al., 2007; Nestler & Hyman, 2010; Demin et al., 2019). The predictive validity of environmentally induced models of depression is demonstrated by the observation that the animals often "get better" when treated with drugs that are effective for depression in humans. Nonetheless, these models have thus far played only a minor role in the development of major new types of antidepressants.

The first major antidepressant drug was iproniazid (aka Marsilid). This compound was originally developed to treat tuberculosis and shown in vitro and in several mammalian species to be an inhibitor of monoamine oxidase, the enzyme that degrades monoamines (Zeller & Barsky, 1952). Observant clinicians in the 1950s discovered that, as a side effect, iproniazid often made patients euphoric and boosted their appetite. These observations prompted tests on depressed subjects in the late 1950s (Loomer et al., 1957), and the positive results of these studies led to the widespread use of iproniazid as an antidepressant (Sandler, 1990). Unfortunately, the drug causes dangerous spikes in blood pressure when it interacts with dietary tyramine, a compound found in cheese. On account of this cheese reaction, iproniazid was withdrawn from the market in 1961.

The second major antidepressant was imipramine, which is considered a tricyclic antidepressant because of the three rings in its chemical structure. Imipramine is structurally related to chlorpromazine, which is used to treat schizophrenia (which we will discuss shortly), and it was originally synthesized as part of a search for more effective antischizophrenia drugs. When tested in the 1950s on approximately 500 humans with a wide variety of mental disorders, imipramine was ineffective against schizophrenia but effective against major depression (though the improvement over placebo was only approximately 30%; Maxwell & Eckhardt, 2012). Imipramine was approved for treatment of major depression in 1959 and spawned a long series of studies looking for more effective tricyclic antidepressants. Many such drugs have come to market, but they all appear to be similarly effective (Arroll et al., 2005).

The third major class of antidepressant drugs to be discovered are the serotonin reuptake inhibitors (SRIs), which reduce the ability of neurons to recycle previously released serotonin and thus cause the released serotonin to hang around longer. The possibility that serotonin might have a special role in depression dates back to the 1960s, when scientists discovered that depressed suicide victims had low levels of serotonin in the brain. In addition, the earlier antidepressant imipramine had been suggested to function, at least in part, as an SRI (Wong et al., 2005). These observations led to an explicit search for additional serotonin reuptake inhibitors, using in vitro preparations of rat brain and in vivo assays. By 1974, fluoxetine (aka Prozac) emerged as one such drug (Wong et al., 1974). Studies with animal models and early human tests indicated good safety but poor effectiveness. After several additional trials, however, fluoxetine was FDA approved in 1987 for the treatment of major depression. Subsequent research produced a variety of additional SRIs, some of which were even more selective for serotonin (as opposed to other monoamines). However, nonselective SRIs actually tend to work better against depression than highly selective SRIs (Roth et al., 2004). Particularly puzzling is that these drugs, like the tricyclic antidepressants,

take several weeks before the depression lifts (Machado-Vieira et al., 2010), even though the direct cellular effects can be observed almost immediately.

A very recent addition to our armamentarium against major depression is ketamine, which blocks NMDA-type glutamate receptors (Sleigh et al., 2014). Ketamine was synthesized in the 1960s as part of a search for new anesthetics. It was tested on diverse animals and first given to humans in 1964. Deemed relatively safe and effective, it was then used widely as an anesthetic in veterinary medicine. Even in early studies, however, doctors noticed that ketamine causes human subjects to become detached from reality, which explains why it is sometimes abused (e.g., under the name of Special K). Meanwhile, scientists in the 1990s began to suspect, on the basis of research on both humans and mice, that depression might be linked to changes in NMDA receptors. Prompted by this hypothesis, they gave ketamine to a few depressed patients and observed profoundly positive effects (Berman et al., 2000). A larger follow-up study confirmed these findings (Zarate et al., 2006), and ketamine is now recognized as a highly effective treatment for major depression, especially in cases where other treatments have failed. One of its significant advantages over other antidepressants is that its positive effects are felt within a few hours.

Finally, it is interesting that one of the first, and still one of the most effective, treatments for major depression is electroconvulsive therapy (ECT, aka electroshock therapy). It is still unclear exactly how ECT exerts its antidepressant effects, and it certainly has a host of risky side effects. However, it is widely recognized as being effective against major depression and, when administered in conjunction with anesthesia, is recommended as a measure of last resort (Roth et al., 2004; Li et al., 2020).

6.3.2 Bipolar Disorder

Bipolar disorder (BD) affects at least 1% to 2% of the world's population (Yutzy et al., 2012; Brown, 2019) and is characterized by dramatic mood swings that recur at variable intervals, ranging from weeks to years (McCormick et al., 2015). Periods of deep depression tend to alternate with periods of mania, which feature prolonged episodes of great energy and well-being as well as elevated feelings of self-worth, racing thoughts, and often great creativity. Indeed, the list of writers, poets, painters, and musicians who likely struggled with BD is long ("List of People with Bipolar Disorder," 2021). That said, BD is a severely debilitating disease that often ends in suicide. Unfortunately, its underlying mechanisms remain mysterious. Although monozygotic twins are far more likely than dizygotic twins to share a diagnosis of BD, only a few genetic mutations have been linked to this disease, and each of them only slightly elevates the disease risk (Barnett & Smoller, 2009; Orrù & Carta, 2018). It is not surprising, therefore, that animal models of BD are relatively scarce and often of dubious

validity (Malkesman et al., 2009; Beyer & Freund, 2017). For example, some authors have tried to model BD by giving animals cocaine or amphetamines, but this only mimics some aspects of the mania and fails to capture the cyclic nature of BD. Others have manipulated some of the BD-linked genes in mice, and some of these animals respond to lithium, which is the main drug used to treat human BD. However, the extent to which these models can lead to novel therapies remains as yet unknown.

The story of how lithium came to be "psychiatry's most consistently effective medicine" (Draaisma, 2019, p. 584) is both strange and complex. Lithium salts have a long history as treatments for diseases such as gout, but the idea that they could be used to treat depressive disorders originated with the Danish psychiatrist Frederik Lange in 1894 (Shorter, 2009). The idea was then ignored for roughly 50 years, until a psychiatrist named John Cade injected the urine of mental patients (or an 8% solution of urea) into guinea pigs and found that treating those animals with lithium reduced the injection's toxicity (Cade, 1949; Brown, 2019). He further noticed that lithium injected into otherwise untreated animals rendered the animals lethargic, though "fully conscious" (Cade, 1949). Being interested in BD, Cade then made a huge conceptual leap and administered lithium to 10 patients suffering from mania. Although Cade reported that many of his patients became "practically normal" after a few weeks, most readers were skeptical. A few years later, however, Mogens Schou and his collaborators confirmed the essence of Cade's report (Schou et al., 1954). Still, skepticism persisted. It was only in 1970, after Schou and his collaborators showed in a large placebo-controlled clinical trial that lithium prevents the recurrence of manic/depressive episodes (Baastrup et al., 1970) that the medical community took note.

High doses of lithium do cause some serious side effects, but these can be controlled by monitoring drug levels in the blood. Although several alternatives to lithium have been developed over the years (e.g., valproate), lithium remains the drug of choice for many patients with this disorder (Baldessarini & Tondo, 2000).

6.3.3 Schizophrenia

Schizophrenia currently affects roughly 1% of the global population, and its symptoms typically begin in adolescence or early adulthood (with men being more susceptible) (R. Li et al., 2016). Definitions of schizophrenia have changed over the years (Carpenter & Koenig, 2008), but some of the disease's most common symptoms are psychotic episodes (e.g., delusions or hallucinations), reduced emotional expression, low motivation, and asocial behavior. Intriguingly, schizophrenia is among the most heritable psychiatric disorders. If one monozygotic twin is diagnosed with schizophrenia, the other twin's chance of sharing the disease is much higher than the rate for dizygotic twins (48% versus 4%) (Onstad et al., 1991; Sullivan et al., 2012). Nonetheless,

the genetic underpinnings of schizophrenia are very complex, as indicated by the fact that more than 100 (perhaps as many as 1,000) genes have been linked to this disease (Pratt et al., 2012; Schizophrenia Working Group of the Psychiatric Genomics Consortium, 2014). Moreover, each schizophrenia-linked mutation accounts for only a relatively minor increase in disease risk, and diverse environmental factors (e.g., prenatal stress) are almost certainly involved as well (Sullivan, 2005). At the level of the brain, a few differences in brain structure and functional dynamics reportedly correlate with schizophrenia (Dietsche et al., 2017; Vanes et al., 2019), but linking this variation to specific genes or schizophrenia symptoms remains a tall order.

Creating animal models of schizophrenia likewise remains challenging, both because of the complex genetics and because we cannot really know when nonhumans are experiencing hallucinations, delusions, or disordered thought. Scientists have tried to model schizophrenia by injecting animals with amphetamines (yes, the same manipulation also used to model mania) or other drugs that can trigger psychotic episodes in humans (Pratt et al., 2012; Steeds et al., 2015). Others have stressed rats early in their development—for instance, by underfeeding the pregnant mothers, isolating the newborns, or lesioning the ventral hippocampus of juveniles (Powell, 2010). These efforts have allowed scientists to model select aspects of schizophrenia, but the core symptoms have remained intractable.

Hoping to surmount these problems, scientists have created mice with modified schizophrenia-linked genes (Kvajo et al., 2012). For example, they have produced transgenic mice that express function-blocking versions of a gene called *disrupted in schizophrenia-1* (*DISC1*) (Hikida et al., 2007). These mice exhibit some schizophrenia-like symptoms, but mice with humanlike mutations in the endogenous *DISC1* gene exhibit far more subtle symptoms (Koike et al., 2006). Moreover, *DISC1* is associated with multiple psychiatric disorders, mutated in relatively few schizophrenics, and mutated in many people who do not develop schizophrenia (Kvajo et al., 2012). Thus, despite the gene's name, it remains unclear to what extent *DISC1* mice are a good model for schizophrenia. Indeed, thus far none of the genetic or developmental models of schizophrenia have played a major role in the development of novel therapies.

Some of the earliest strategies for treating schizophrenia were seriously misguided. Prefrontal lobotomies, in particular, were widely touted as a way to make schizophrenics (and other agitated patients with mental disorders) more manageable, but the side effects were often severe (Acharya, 2004; Tan & Yip, 2014). This procedure, which surgically destroys the prefrontal cortex or disconnects it from the rest of the brain, was inspired by the results of frontal cortex lesions in chimpanzees (Jacobsen et al., 1935). However, this study was small and was never intended to serve as the basis for clinical translation. Nor were lobotomies tested in rigorous clinical trials (Tierney,

2000). They did help to reduce the population of schizophrenics in overcrowded mental institutions (Swayze, 1995), but in retrospect they were one of medicine's most irresponsible misadventures. Fortunately, lobotomies fell out favor in the late 1950s, in part because pharmacological treatments for schizophrenia became available.

Drug treatments for schizophrenia have focused primarily on controlling the psychotic aspects of the disease, and they generally involve the blocking of dopamine receptors (Carpenter & Koenig, 2008). The first major antipsychotic drug was called chlorpromazine (aka Thorazine). Historically it was preceded by promethazine, which had originally been developed as an antihistamine. The physician scientist Henri Laborit in 1949 recognized that this drug was a mild sedative and could be used as an adjunct to surgical anesthesia. Laborit then encouraged colleagues to develop additional antihistamines with even stronger drowsiness-inducing effects. This search was conducted with rats and soon led to chlorpromazine.

Laborit and his clinical colleagues then gave chlorpromazine to various patients and found that it consistently reduced the agitation of manic patients and quieted the delusions and hallucinations of schizophrenics. In the words of Laborit and colleagues, "in doses of 50–100 mg intravenously, it provokes not any loss in consciousness, not any change in the patient's mentality but a slight tendency to sleep and above all 'disinterest' for all that goes on around him" (1952; trans. Caldwell, 1970, p. 3). By the end of 1952, chlorpromazine had been given to roughly 2 million patients. It dramatically reduced the number of patients confined to mental institutions and remained popular for several years. Unfortunately, chlorpromazine and other, functionally similar "typical anti-psychotics" (Meyer & Simpson, 1997) had serious extrapyramidal side effects, including tremors, erratic movements, slurred speech, anxiety, and paranoia.

The discovery of "atypical antipsychotics" (aka second-generation antipsychotics) began with the development of clozapine. This drug was first synthesized in 1958, quickly tested in animals, and then shown in humans to be an effective antipsychotic that was less likely than chlorpromazine to cause extrapyramidal side effects (Shen, 1999; Crilly, 2007). Unfortunately, clozapine can lead to an alarming loss of white blood cells, a side effect that proved to be a major stumbling block in its development. Nevertheless, a large clinical trial in 1988 demonstrated that clozapine was more effective than chlorpromazine in patients who had not responded to prior treatment with haloperidol, a typical antipsychotic (Kane et al., 1989). Moreover, clozapine's negative side effects could be controlled by carefully monitoring white blood cell counts. Consequently, clozapine was granted US approval for the treatment of schizophrenia in 1990.

At a mechanistic level, clozapine differs from the typical antipsychotics mainly in that it blocks dopamine receptors more transiently and, in addition, blocks serotonin receptors (Meltzer, 1994). The latter property explains why clozapine and other

atypical antipsychotics tend to block the psychotic effects of the hallucinogen LSD (lysergic acid diethylamide), which exerts its effects at least in part by activating serotonin receptors (Valeriani et al., 2015).

Considering the various typical and atypical antipsychotic drugs, patients and physicians now have at their disposal a variety of drugs to treat schizophrenia. Some work better in some patients than others, and often a combination of drugs is best. Overall, however, schizophrenia treatments remain inadequate. All the drugs have potentially serious side effects, and it remains unclear how their various mechanisms of action relate to schizophrenia's behavioral and cognitive symptoms. To illustrate some of the mystery that still surrounds this disease, schizophrenics tend to smoke more heavily than control subjects (Kumari & Postma, 2005; Manzella et al., 2015). They seem to be self-medicating with nicotine, but how and why this works remains unknown.

6.4 SUCCESSES, FAILURES, HOPES

As noted in this chapter's opening quote, frustration has been growing over the lack of progress in the development of treatments for neuropsychiatric disorders over the last half century. Progress on treatments for CNS injury and neurodegenerative diseases has likewise been meager. Although advances in the treatment of cancer and some other diseases have also been slower than one would like (see chapter 5), the neurological disorders have been especially refractory to therapy development. Why has progress been so slow?

One very likely explanation is that human brains exhibit virtually no adult neurogenesis (Bergmann et al., 2012; Sorrells et al., 2018), which means that lost neurons cannot be replaced. Therefore, neurological disorders involving neuron death should ideally be treated before a significant number of neurons has been lost. An analogous argument could be made for treating psychiatric disorders before they manifest, especially if they involve irreversible losses of neuronal connections (S. Li et al., 2019). Prophylactic interventions have worked well for the treatment of cancer and cardiovascular disease, and early intervention is quite common in animal research on neurodegenerative disorders (e.g., in mouse models of Parkinson's disease) (Zeiss et al., 2017). However, for neurological disorders in humans, this strategy would require clinical trials that include huge numbers of subjects because it is currently difficult to predict who will develop a neurological disorder later in life. Moreover, treating huge numbers of people for decades before they might develop symptoms would likely be extremely expensive.

In the following section, I discuss three additional potential explanations for the slow progress of neurological treatment development. The first is that drug development has historically involved a great deal of serendipity; perhaps we are currently just

stuck in streak of bad luck. Second, some of the models may be actively misleading us, effectively trapping us in dead ends, while ignoring other research avenues. Third, most of the current models may be mimicking only parts of the targeted disease. All these explanations contain kernels of truth, but each on its own is incomplete.

6.4.1 The Role of Serendipity

Many of the drugs that have ended up being widely used to treat neurological disorders were originally developed for other purposes. Tetrabenazine and imipramine were designed to be antipsychotics but turned out to be at least somewhat effective against Huntington's disease and major depression, respectively. Iproniazid was developed to fight tuberculosis before it was discovered to have antidepressant activity; lithium had long been used a treatment for gout before it was repurposed to fight bipolar disorder; and chlorpromazine was created as part of a search for novel antihistamines. Ketamine, too, was originally developed as an anesthetic, not as a depression fighter. These were all serendipitous discoveries, but hardly dumb luck. They came as a result of insightful clinical observations and minds that were open to unexpected developments. As Pasteur famously opined in 1854, "in the fields of observation chance favors only the prepared mind" (see Pearce, 1912).

A related observation is that many neurological treatments were discovered before 1960, when it was still possible for clinicians to test drugs on themselves (e.g., Cade tested lithium on himself before he gave it to his patients) and on small groups of human volunteers without extensive prior animal testing (Janssen, 2009). The more extensive regulations of the modern era were imposed as justified responses to major medical disasters, but they make chance discoveries less likely. It is still possible to test already approved compounds for off-label purposes, hoping to find new uses for old drugs (Heemskerk et al., 2002), but the odds of making serendipitous discoveries have certainly gone down.

One might think that the problem is much less severe in preclinical research on non-mammalian or in vitro models because they often allow for high-throughput drug screening. However, the odds that those discoveries will translate to humans appear thus far to be quite low. Most likely, these models are just too different from the human condition to allow for consistent extrapolation.

6.4.2 Convenient Models That Mislead

Although it is a given that "all models are wrong" (recall the opening quote of chapter 2), my review of the preclinical research on neurological disorders suggests that many of our current model systems may be too wrong to be useful in the quest for therapies. The models may be convenient insofar as they develop symptoms quickly and reliably, thereby facilitating the rapid publication of results, but most of the models are more

like caricatures of the human condition than facsimiles (Libby, 2015). The R6/2 mice, for example, develop symptoms far more rapidly than humans with Huntington's disease (and develop severe epilepsy as they age). Similarly, mice that combine multiple mutations linked to Alzheimer's disease are quite unnatural. Moreover, most of our models are based on genetic information gleaned from familial versions of the disease, but for most neurological disorders sporadic cases outnumber the familial ones. What if the sporadic diseases differ in ways that matter for therapy development? And, as we just discussed, what if the models that are most suitable for high-throughput drug screening are too divergent to produce more than promising hits?

In short, one may well wonder whether biomedical researchers have been behaving like the proverbial drunk who searches for his keys at night under a streetlight (Schnabel, 2008): when a policeman asks the drunk why he is looking there, the drunk responds that the keys were lost on the other side of the street but the light is better under the lamp. In the words of Hittner et al. (2019), "when it comes to scientific research, the 'streetlight effect' is no longer a joke" (p. 2).

Searching in darkness is difficult, of course. Nor should one underestimate the pressures on researchers to produce publications and tangible results. In the absence of those, grant funding quickly dries up, and ambitious students seek greener pastures. When that happens, it becomes essentially impossible to make any empirical discoveries. Work with model systems that more closely resemble the human condition often takes years before results emerge, especially if those systems involve nonhuman primates or other large animals. Such work requires great patience from administrators and granting agencies as well as the individual investigators. Working with nonhuman primates, dogs, or any other higher species also entails the risk of attacks by animal rights activists, as well as internal ethical conflicts. In vitro studies with human cells can yield results more rapidly and harbor fewer ethical quandaries, but constructing in vitro systems that even remotely approach the complexity of their in vivo target is a much slower, more treacherous path. Given these pressures, why not study whichever models are most easily studied and most likely to yield quick results?

The problem is that the most convenient models may sometimes lead researchers down what Shakespeare called the primrose path (Burian, 1993). One should certainly be wary of turning August Krogh's principle (see chapter 2 and the opening quotation of chapter 3) on its head: there may well be a "most convenient" model in which to study any biological problem (Krebs, 1975), but studying a very convenient model will not necessarily answer the questions that originally motivated the inquiry. Moreover, some of the convenient models may have unsuspected problems. For example, the FVB/N strain of mice is very convenient because the animals breed readily, but this strain (and a few others) also carries a gene causing blindness (Chen et al., 2013; Zeiss, 2015).

In general, it seems preferable to work with natural (aka spontaneous) disease models instead of artificially constructed models. Such models have been very useful, for example, in research on diabetes, hypertension, and atherosclerosis. Unfortunately, spontaneous models of neurological disorders are rare. This is especially frustrating because the artificial models of neurological disorders are, as we discussed, very likely to be misleading in at least some respects. As Felix d'Hérelle, a pioneer in the study of infectious diseases, once complained, "because all illnesses studied by significant authors were 'artificial' illnesses (neither the rabbit nor the guinea pig are affected by cholera or typhus in the natural environment) they have bearing only when talking about the artificial illness and not at all practical for application to real, natural illnesses which occur in humans" (Endersby 2007, p. 290). D'Hérelle turned out to be wrong about the utility of "artificial" models in vaccine development (see chapter 5), but his point may be applicable to more complex illnesses, especially those involving the nervous system.

6.4.3 Partial Models and Endophenotypes

Most biomedical researchers readily admit that their models are imperfect. More interesting is that some, especially those who study neurological disorders, tend to refer to their study systems as partial or incomplete models (Jucker, 2010; Zahs & Ashe, 2010; Drummond & Wisniewski, 2017). Mouse AD models, for example, are considered incomplete if they exhibit amyloid plaques but not neurofibrillary tangles or extensive cell death (Schwab et al., 2004). Similarly, animal models of schizophrenia are said to be partial insofar as they share with their human counterparts only a specific subset of symptoms (Roberts, 2007).

Although the idea of partial models seems innocuous, it does entail the debatable assumption that complex diseases are caused by a set of largely independent processes that can be modeled, and ultimately treated, individually. Moreover, it implies that the various models and treatments can be combined with additive effects (e.g., Oddo et al., 2003). This approach is potentially transformative, but its assumptions should be examined carefully. One might point out, for example, that a single change in a complex system may, over time, lead to a variety of interdependent effects and that a given effect may be produced by diverse causal processes.

A sophisticated and impactful version of the partial modeling approach is the Research Domain Criteria (RDoC) framework developed and promoted by the National Institute of Mental Health (Insel et al., 2010; Kozak & Cuthbert, 2016) (for an analogous European framework, see Schumann et al., 2014). This framework is based on the notion of endophenotypes, which I introduced in chapter 2 (section 2.4.1). The core idea is that between the genotype and the outwardly observable phenotype lies an intermediate level of organization, the endophenotype. Individual endophenotypes (i.e., traits at the endophenotype level) can be observed with special tools and behavioral tests.

They include, for example, the death of specific neurons, changes in neuronal connectivity or function, and deficits in attention. Some authors have argued that endophenotypes must be heritable and have adaptive functions (Gottesman & Gould, 2003; Cuthbert & Insel, 2013); however, in practice these two criteria are rarely treated rigorously. The framework's most important aspect, in my view, is that it formally justifies the modeling of traits that contribute to a disorder without having to claim that one has modeled the entire disorder; that is, it justifies the creation of partial models, including so-called pathway models (Schnabel, 2008). The endophenotypes can then, in theory, be linked to underlying genes and, going in the other direction, to specific mental or behavioral symptoms. Therapy development is envisioned as targeting specific endophenotypes, which implies that the ideal treatment for complex disorders will require multiple simultaneous treatments and that any one treatment may be of use in multiple disorders.

The RDoC framework originated, and is presently employed, mainly in biological psychiatry. This is understandable because psychiatric disorders tend to have a large number of overlapping diagnostic criteria (Lilienfeld, 2014) and, as we have reviewed, remain difficult to treat. The framework may well be applicable to other types of complex diseases, but this remains to be seen. Some authors have hailed the RDoC framework as a new paradigm in which complex diseases like schizophrenia are deconstructed into multiple domains without a "unifying pathophysiology" that ties them together (Carpenter & Koenig, 2008). This is an exciting idea, but the RDoC framework has been criticized for being difficult to falsify, in the sense of providing no clear criteria for determining whether or not it is "mapping well onto the state of nature" (Lilienfeld & Treadway, 2016).

Ultimately, the framework's utility is an empirical question: will it lead to new treatments that the traditional paradigm would not have revealed? This question remains open. What is already clear, however, is that the RDoC framework has allowed research on fundamental aspects of neurobiology and behavior to be considered translational, even if it does not explicitly target a specific psychiatric disorder (Rubio et al., 2010; Anderzhanova et al., 2017). Thus, the RDoC framework facilitates the integration of basic and translational research (McArthur & Borsini 2008) (see chapter 1, section 1.1). Such a rapprochement could well be useful far beyond psychiatry.

Model Systems in Biology

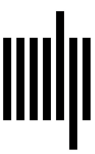
History, Philosophy, and Practical Concerns

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