Chapter 12

Making and Using Genetically Modified Organisms

After reading this chapter, you should be able to:

- Explain the utility of genetically modified organisms in neuroscience
- Describe and categorize commonly used transgenes, and explain how each transgene is useful for studying the nervous system
- Understand how promoters are used to regulate gene expression and target transgenes to specific cell types
- Differentiate between strategies for introducing DNA sequences into the genome
- Explain the process of creating genetically modified organisms, especially mice, flies, and worms
- Compare methods of disrupting endogenous gene products rather than the genome itself

Techniques covered:

- **Common transgenes in neuroscience:** reporter genes, transgenes used to measure neural activity, transgenes used to manipulate neural activity, transgenes used to silence neural activity, transgenes used to ablate neurons
- **Strategies for genome modification:** nonspecific transgene insertion, classical DNA targeting in mice using homologous recombination, DNA targeting/editing using CRISPR/Cas9
- **Engineering genetically modified organisms:** mice, flies, worms
- **Binary transgenic systems:** Gal4/UAS system, Cre/lox system, TRAP system, Flp/FRT system, Tet-off/Tet-on systems
- **Disrupting gene products:** RNA interference (RNAi), morpholinos

A **genetically modified organism** is any organism whose genome has been modified using genetic engineering techniques. If the genomic modifications are in germ cells (sperm or eggs), then the altered DNA is heritable and passed to all future offspring. Using modern techniques such as CRISPR/Cas9, it is possible to remove genetic sequences, edit and replace genetic sequences, or add foreign genetic sequences to the genomes of virtually any organism.

Guide to Research Techniques in Neuroscience. https://doi.org/10.1016/B978-0-12-818646-6.00003-8

Genetically modified organisms have proven extremely useful for studying the genetic basis of behavior and other neural phenotypes. For example, genetic knockout animals allow scientists to test the hypothesis that a gene is necessary for a certain biological process by removing the gene from the genome; alternatively, an investigator can overexpress a gene and determine the effects of a gain of function of that gene. Genetically modified organisms have also proven extremely useful for targeting specific cell types with genetically encoded tools. By using genetic regulatory elements found only in specific populations of neurons, it is possible to express a transgene useful for studying that population's structure, projection patterns, and activity patterns, as well as to perturb the population's neural activity.

The purpose of this chapter is to survey the rationales for engineering a genetically modified organism and explain the process of removing or inserting genetic material in common model organisms. We will also survey several commonly used transgenes and describe how these transgenes can be targeted to cell types to study discrete populations of neurons. Finally, we describe the process of targeting mRNA transcripts rather than genes themselves.

REASONS FOR ENGINEERING A GENETICALLY MODIFIED ORGANISM

Creating genetically modified organisms, especially animals capable of passing modified genetic material to their offspring, requires much time and skill. Therefore, using these techniques should have a clear experimental goal. There are three general rationales for creating genetically modified organisms: disrupting endogenous gene function, replacing endogenous genetic material, and inserting a functional transgene.

Disrupting Endogenous Gene Function

To test the necessity of a gene for a phenotype or behavior, a scientist can remove the gene from the genome and examine the effects on the mutant animals. Until recently, the technology necessary to deliberately target the coding sequence of a gene was limited to mice. Indeed, over the past several decades, **knockout mice** have proven extremely useful for loss-of-function genetic studies. Loss of gene function studies in other model organisms were typically the result of mutagenesis screens in which scientists caused random mutations in the genome, ultimately discovering a line of animals that lacked a specific gene.

Newer technologies, especially **CRISPR/Cas9** genome editing (described later in this chapter), have made it possible to disrupt gene function in virtually any organism. Additionally, techniques that disrupt the mRNA transcripts of genes (such as RNAi and morpholinos, described at the end of this chapter) can also be used to disrupt endogenous gene function.

Replacing Endogenous Genetic Material

Instead of removing a gene from the genome, a scientist can replace an endogenous DNA sequence with new genetic material that modifies/enhances gene expression for a specific purpose. A **knockin animal** is a genetically modified organism in which a genetic sequence is added at a specific locus in the genome. For example

- To replace a gene of interest with a slightly mutated version to mimic a mutation found in human disease. The gene continues to code for a functional protein with slightly different properties and effects on the cell. For example, a **humanized mouse** is a mouse carrying functional human genetic sequences to better study human gene products for medical research purposes.
- To replace a regulatory region of DNA (such as a **promoter** sequence—described below) to investigate how a gene is regulated.
- To replace a functional gene with a tagged version of the same gene, for example, a gene fused to the coding sequence for green fluorescent protein. The resulting fusion protein can then be used to visualize gene expression patterns in fixed or living tissue.
- To create a **binary transgenic system** (described below) that allows different lines of genetically modified animals to be crossed together to refine the spatial and temporal expression of a gene.

Inserting a Functional Transgene

Instead of deleting or modifying endogenous genes, a scientist can introduce a functional gene that the model organism does not naturally express (Table 12.1). Any gene expressed in an animal that does not normally carry this gene is called a **transgene**, and any animal that carries this foreign DNA is called a **transgenic organism**. Many transgenes are derived from other species (e.g., GFP, channelrhodopsin-2, etc.). The rich diversity of life on Earth provides an extraordinary variety of genes that not only help organisms survive unique niches in their natural habitats, they have also proven very useful for studying the structure and function of the nervous system. Other transgenes are synthetic versions of naturally occurring genes, engineered to serve a useful purpose (e.g., GCaMP, DREADDs).

A transgenic organism can be designed to have a transgene expressed at a specific locus in the genome, in which case it would be considered a knock-in animal. Alternatively, a transgene can be expressed at a random location within a genome, in which case it is usually referred to simply as a transgenic animal.

The modern neuroscientist's toolkit includes many useful transgenes that can be targeted to specific cells in the brain. Below, we survey many commonly used transgenes, then describe how scientists can design and express DNA sequences to create genetically modified organisms.

TABLE 12.1 Examples of Commonly Used Transgenes.

	Type of Transgene	Utility	Examples
Reporter genes	Fluorescent proteins	Marks the location of cells, subcellular structures, and individual proteins	Green fluorescent protein (GFP) and multicolored variants
	Fluorescent proteins to trace neural connections	Addition of tags to fluorescent proteins so they are bound to the membrane or synapse	Farnesylated-GFP (bound to membrane); Synaptophysin-GFP (bound to presynaptic membrane)
	Nonfluorescent proteins	Mark the location of cells, but depend on further histochemical processing for visualization	LacZ; Alkaline phosphatase (AP)
Genes for measuring neural activity	Calcium indicators of cellular activity	Sensors that increase fluorescence in response to increasing intracellular concentrations of calcium ions	GCaMP
	Voltage sensors	Sensors that increase fluorescence in response to increases in membrane potential	Archon
	Indicators of neurotransmitter release	Sensors on the membrane that increase fluorescence in response to binding of neurotransmitters	iGluSnFR (glutamate transmission); iGABASnFR (GABA transmission)
	Indicators of neuromodulator release	Sensors on the membrane that increase fluorescence in response to binding of neuromodulators	dLight and GRAB-DA (dopamine transmission); GRAB-ACh (acetylcholine transmission); GRAB-NE (norepinephrine transmission); GRAB-5HT (serotonin transmission)

Continued

TABLE 12.1 Examples of Commonly Used Transgenes.—cont'd

	Type of Transgene	Utility	Examples
Genes for inducible activation of neural activity	Optogenetic actuators	Membrane-bound channels that depolarize cells when stimulated by light	Channelrhodopsin-2 (ChR2)
	Chemogenetic actuators	Ligand-gated ion channels that depolarize cells in the presence of a ligand	hM3Dq; TRPV1
Genes for inducible inhibition of neural activity	Optogenetic actuators	Membrane-bound channels/pumps that hyperpolarize cells when stimulated by light	Halorhodopsin (NpHR); Archaerhodopsin (Arch)
	Chemogenetic actuators	Ligand-gated ion channels that depolarize cells in the presence of a ligand	hM4Di, allatostatin receptor
Genes for neuronal silencing	Toxins	Toxins that prevent docking of synaptic vesicles, effectively preventing neurotransmission	Tetanus toxin (TeNT); botulinum toxin (BoNT)
	Shabire mutants	Temperature sensitive mutants of the *Drosophila* shibire gene that prevent synaptic transmission at certain temperatures	Shi^{ts}
Genes for neuronal ablation	Toxins	Kills all expressing neurons	Ataxin
	Toxin receptors	Kills all expressing neurons when a toxin is administered	Diphtheria toxin receptor

COMMONLY USED TRANSGENES

There are dozens of commonly used transgenes useful for answering a variety of experimental questions in neuroscience (Table 12.1). Scientists can introduce these transgenes into cells using methods described in Chapter 11, or they can use methods described later in this chapter to create stable lines of genetically modified animals that pass on the transgene to their offspring.

Reporter Genes

A **reporter gene** is any gene whose protein product can visually report the location of a particular protein, cell type, or circuit. The most common reporter genes are **green fluorescent protein (GFP)** and its variants, originally derived from jellyfish. When placed under the control of a cell-specific promoter, GFP can mark the location of specific cells so they can be used in subsequent experiments (for example, recording from specific cells in electrophysiology experiments; identifying cell types in histology experiments, etc.). In addition to GFP, there are dozens of fluorescent proteins with different spectral properties that appear as different colors, such as red fluorescent protein (RFP), cyan fluorescent protein (CFP), and yellow fluorescent protein (YFP).

Fluorescent proteins can be manipulated so that they tag proteins or organelles within a cell. For example, a scientist can use recombinant DNA methods (Chapter 10) to attach the genetic sequence of GFP to a gene that encodes a protein of interest so that the location, expression, and trafficking of the protein can be studied in real time. It is also possible to add specific sequence tags to fluorescent proteins so that they localize to discrete regions of the cell, such as the cell membrane or the synapse. Membrane-bound fluorescent reporters can be useful for tracing axons and dendrites in addition to cell bodies.

Before the widespread use of GFP and other spectral variants, **lacZ** was commonly used in many classic studies as a reporter gene. Originally derived from bacteria, lacZ encodes the enzyme β-galactosidase. This enzyme reacts with a substrate, X-gal, to create a dark blue product. Therefore, lacZ can serve as a faithful reporter of gene expression after histological processing with X-gal. Because fluorescent proteins do not require enzymatic processing for visualization, and because different fluorescent colors can label different structures within the same specimen, fluorescent proteins have almost entirely replaced lacZ/X-gal techniques.

Transgenes Used to Measure Neural Activity

Chapter 7 describes methods of indirectly measuring neural activity by imaging calcium dynamics or membrane voltage. **Genetically encoded calcium indicators (GECIs)**, such as **GCaMP**, or **genetically encoded voltage indicators (GEVIs)**, such as **Archon**, allow targeted observation of neural activity in specific cell types. This specificity is an advantage over calcium sensitive or voltage sensitive dyes that target all neurons at the injection site. Other transgenes can report the binding of neurotransmitters to their receptors on postsynaptic cells. For example, **iGluSnFR** and **iGABASnFR** are modifications of glutamate and GABA receptors, respectively, that increase

fluorescence when these neurotransmitters are released into the synapse. **dLight** and **GRAB-DA** are modifications of the dopamine receptor D1 (Drd1) and dopamine receptor D2 (Drd2), respectively, that increase fluorescence when dopamine is released into a synapse. **GRAB-ACh**, **GRAB-NE**, and **GRAB-5HT** increase fluorescence in the presence of acetylcholine, norepinephrine, and serotonin, respectively.

Transgenes Used to Manipulate Neural Activity

Neural activity can be manipulated through nongenetic methods, such as electrical stimulation or pharmacological inhibition. However, it can be difficult to use these methods to selectively stimulate or inhibit cells of a desired subtype within a local area, especially if the local area is composed of several neural subtypes. Genetically targeting specific cell types with transgenes solve these problems because cells that do not express the transgene are unaffected. These transgenes typically encode optically controlled or ligand-gated transmembrane proteins to selectively stimulate or inhibit neural activity. These techniques are described in more detail in Chapter 8.

Optical methods of neuromodulation allow for high temporal and spatial selectivity. **Channelrhodopsin-2 (ChR2)**, originally derived from single-celled algae, is a cation channel that can depolarize a neuron upon stimulation with blue light. Because ChR2 can be placed under cell-type−specific promoters, it is possible to stimulate specific neurons in a heterogeneous population when delivering blue light to the local area. The chloride pump **Halorhodopsin (NpHR)** and the hydrogen pump **Archaerhodopsin (Arch)** hyperpolarize neurons upon stimulation with yellow light. Using these transgenes, it is possible to silence single action potentials or sustained trains of neuronal spiking.

Another method of manipulating neural activity is to express the gene for a ligand-gated receptor in an animal that normally does not express the receptor. For example, the **capsaicin receptor (TrpV1)** is a ligand-gated channel that depolarizes a neuron upon exposure to the compound capsaicin (the chemical produced by peppers that is perceived as hot by most mammals). Flies and other invertebrates do not normally express this channel, so capsaicin exposure will depolarize only those neurons that express the receptor. This strategy can also be employed in TrpV1 knockout mice that lack functional capsaicin receptors. Other receptors cause hyperpolarization of cells when bound to their ligand. For example, the insect **allatostatin receptor (Alstr)** is not expressed in mammals. In the presence of the insect protein allatostatin, this receptor inhibits neural activity. Thus, if Alstr is expressed in the mammalian brain, addition of allatostatin will reversibly inactivate the genetically targeted neurons until the ligand is washed out from the system.

In addition to naturally occurring ligand-gated receptors, synthetic G-protein—coupled receptors are also useful for modulating neural activity. **Designer Receptors Exclusively Activated by Designer Drugs (DREADDs)** consist of artificial receptors for the normally inert ligand **clozapine-N-oxide (CNO)**. In the presence of CNO, **hM3Dq** increases neural activity, while **hM4Di** causes hyperpolarization.

Transgenes Used to Silence Neural Activity

In addition to the transgenes described above that reversibly inhibit neural activity, there are transgenes that, when expressed, allow a neuron to fire action potentials normally, but not release neurotransmitter. Therefore, these transgenes effectively "silence" neural activity. For example, the neurotoxins **tetanus toxin (TeNT)** and **botulinum toxin (BoNT)** inhibit neurotransmitter release by cleaving proteins necessary for exocytosis of synaptic vesicles in mammals. These transgenes are irreversible, therefore, once expressed the neurons will be silent throughout the life of the animal. In *Drosophila*, the temperature-sensitive **shabire** (dynamin) transgene can reversibly silence neurons at specific temperatures.

Transgenes Used to Ablate Neurons

Lesion studies have been a historical staple of neuroscience research to investigate the necessity of a brain region for a specific behavior or phenotype. Genetic ablation experiments cause the death of specific cell types while leaving the surrounding cells intact. For example, the **ataxin** transgene encodes a protein that kills all expressing neurons. Alternatively, it is possible to use a transgene that encodes a toxin's receptor, such as the **diphtheria toxin receptor (DTR)**. Diphtheria toxin is toxic to neurons only when the receptor is present on the cell's membrane. Neurons that do not express the receptor are not sensitive to the toxin. Therefore, unlike the ataxin transgene, the DTR transgene allows an investigator to ablate neurons at any time point during an animal's lifespan depending on when the investigator administers the toxin.

In addition to the transgenes described above, it is possible to express genes that disrupt endogenous gene function (such as expressing CRISPR/Cas9 constructs or RNAi constructs, described later in the chapter). It is also possible to overexpress endogenous genes that the cells already produce.

The major advantage to using the transgenes described above over nongenetic methods is that they allow for cell-type—specific expression. This specificity is achieved from cell-type—specific promoters and other elements that regulate gene expression. The next sections describe how a scientist uses promoters and gene regulatory elements to expresses foreign DNA in an animal's genome.

USING PROMOTERS TO REGULATE GENE EXPRESSION

For a gene to be transcribed into mRNA (and subsequently translated into a functional protein), several endogenous proteins, called **transcription factors**, interact with neighboring regions of DNA, called **promoter** sequences, to initiate and regulate transcription (Fig. 12.1). Scientists designing a genetically modified organism must be cognizant not only of the genes they wish to introduce or manipulate, but also how those genes are regulated by promoters. For example, if a scientist introduces a functional gene in a nonspecific location in the genome, they must also introduce a functional promoter to drive expression of the gene.

Endogenous promoters regulate the transcription of genes so that they are expressed in specific cells and tissues at specific times. For example, every cell in the human body contains the gene encoding the enzyme tyrosine hydroxylase (TH). However, this gene is only expressed in a subset of neurons in the brain because only those cells have the endogenous transcription factors and other regulatory proteins necessary to interact with the promoter for TH and activate transcription. A transgenic construct containing all the necessary promoter elements for the TH gene will allow a transgene to be expressed specifically in the subset of cells that normally transcribe TH. Therefore, a scientist can express a transgene in a specific neuronal subtype by identifying a gene uniquely expressed in those cells and taking advantage of the promoter elements for that gene.

Promoter sequences are often depicted as existing just upstream of the genes they regulate (as in Fig. 12.1). However, especially in vertebrate animals, endogenous promoter elements necessary to drive expression of a gene of interest may be located thousands or even hundreds of thousands of base-pairs upstream of a gene. They may also be located within the **introns** within the coding sequences of a gene or even downstream of the coding sequence of the gene.

There are some common promoters that scientists can use to drive strong expression of transgenes in specific cell types (Table 12.2). These promoters are located in the 5′ upstream region before the coding sequence of a gene, are relatively short (1−2 kb), and can fit inside viruses (Chapter 11).

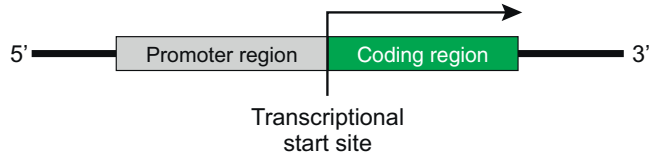


FIGURE 12.1 Promoters drive gene expression. In order for the coding region of a gene to be transcribed into mRNA, transcription factors must bind to a promoter region, often just upstream of the transcriptional start site.

TABLE 12.2 Useful Promoters to Drive Transgene Expression in the Nervous System.

Promoter	Cell Types In Which Expression Occurs
CAG	All cells
EF1a	All cells
Synapsin	All neurons
CamKIIa	All excitatory neurons
GFAP	Astrocytes

Designing a genetically modified organism therefore requires knowledge of not only the endogenous gene to be targeted or the transgene to be inserted, but also the promoter sequences that regulate gene expression. With these considerations in mind, we now consider multiple strategies scientists can employ to create a genetically modified organism.

GENERAL STRATEGIES FOR GENOME MODIFICATION

There are multiple methods of engineering a transgenic organism. All methods require using recombinant DNA techniques (Chapter 10) to produce a desired DNA sequence (the "construct") to introduce into an animal's genome. Below, we describe some general strategies that could be used to introduce this DNA sequence into an animal's genome, and then we describe how a scientist would actually physically deliver these DNA sequences to common model organisms.

Nonspecific Transgene Insertion

The simplest method of introducing a DNA sequence to the genome is nonspecific transgene insertion, in which the DNA integrates into a random genomic location. Although there are certain regions of the genome that seem to be "hot spots" of integration activity, scientists generally cannot control where the transgene integrates. Therefore, this method of creating a genetically modified organism can only be used to introduce a transgene but cannot target or edit specific regions of the genome.

Because of the random integration of DNA, position effects can occur that influence the transgene's expression pattern. Depending on where the transgene lands in the genome, neighboring genes and regulatory sequences can alter the levels and pattern of expression. The number of transgene copies that integrate into the genome can also vary the levels and pattern of transgene

expression. Therefore, transgenic organisms must be carefully evaluated to ensure that transgenes are expressed in the desired cell types. Just because a specific promoter is used to drive transgene expression is not a guarantee that the transgene will be expressed in the desired cells. Therefore, although more challenging and time consuming, the alternative strategies described below may allow for more faithful and efficient transgene expression.

Classical DNA Targeting in Mice Using Homologous Recombination

For many years, targeting a DNA sequence to a specific region of the genome was only efficiently used in mice using a process called homologous recombination. This is a naturally occurring process in which regions with strong sequence similarity can exchange genetic material (Fig. 12.2A). In nature, homologous recombination produces new combinations of DNA sequences during chromosomal crossover in meiosis, resulting in enhanced genetic variation in a population. In the lab, homologous recombination can be used to target DNA sequences to specific regions of the genome (Fig. 12.2B).

To use homologous recombination, a scientist designs a DNA sequence with the DNA to be integrated into the genome flanked by long segments of genomic DNA (at least 1 kb in length) that are homologous to sequences in the endogenous mouse genome. These **homology arms** provide targeting specificity: the region between the two homologous segments replaces or inserts into the target gene in the organism. A scientist can easily identify and choose homology arm sequences flanking a target DNA sequence of interest by examining a mouse genome database.

Once a scientist uses recombinant DNA to produce a targeting construct containing a DNA sequence flanked by homology arms, the process of

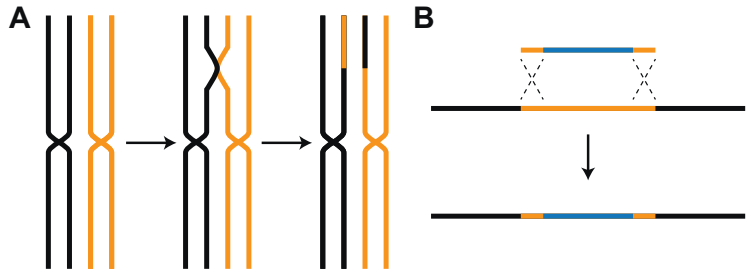


FIGURE 12.2 **Gene targeting using homologous recombination.** (A) Homologous recombination is a natural phenomenon in which genetic sequences with high sequence similarity can exchange genetic material, as occurs with homologous chromosomes during meiosis. (B) In the lab, a scientist can exploit homologous recombination to insert or replace genetic material at a specific locus in the genome. If the foreign DNA is flanked by regions homologous with the locus, homologous recombination will cause the DNA construct to recombine into the genome.

introducing this genetic material is relatively laborious (especially compared to more efficient CRISPR/Cas methods, described below). In brief, the DNA targeting construct is microinjected or electroporated into cultured mouse embryonic stem (ES) cells. Culturing ES cells is challenging in many common model organisms, but highly efficient in mice. Therefore, introducing DNA sequences using homologous recombination has historically been most successful in mice. After introducing the DNA to ES cells, the cells are screened for the correct incorporation of DNA into the genome. Those successfully targeted ES cells are injected into a blastocyst, an early embryonic structure that develops after fertilization of an egg. The blastocyst is then inserted into a female mouse, which produces offspring composed of some cells with the targeted DNA, and some cells with the wild type DNA. If some of the cells with the targeted DNA are germ cells (sperm or egg cells), the mice can be continually bred such that offspring only contain cells with the targeted DNA. This lengthy process can require 1−2 years of work and breeding time before any mice are ready to use for experiments.

This method of targeting is useful for creating knock-out mice that lack a functional gene or knock-in mice in which new DNA is inserted in a specific region of the genome. Transgenes inserted via this method can be targeted to genomic locations that have a specific promoter, making transgene expression very faithful to specific cell types. The disadvantage of using this technique is that it is only widely available in mice, and it is much more time consuming than modern strategies of DNA editing, described below.

DNA Targeting/Editing Using CRISPR/Cas9 Endonucleases

DNA editing with **CRISPR/Cas9** endonucleases has revolutionized the creation of genetically modified organisms. Unlike classical DNA targeting in mice using homologous recombination, these techniques can be used in virtually any organism and are much more efficient in terms of time and success rates.

DNA editing techniques create a double-strand break at a precise location in the genome (Fig. 12.3) using a protein called a **nuclease**, an enzyme that cleaves the bonds between nucleic acids. Nuclease-induced double-strand breaks can be repaired by two different endogenous mechanisms of homologous recombination. In **nonhomologous end-joining (NHEJ)**, naturally occurring DNA break repair processes rejoin the broken ends, often including an insertion or deletion mutation. These mutations can lead to a frameshift mutation of a gene, introducing a premature stop codon and effectively silencing expression. Alternatively, in a process called **homology-directed repair (HDR)**, a scientist introduces a DNA construct with homology arms that have the same sequence as the broken DNA strands. The naturally occurring process of homologous recombination inserts a recombinant DNA construct into the location of the double-stranded break.

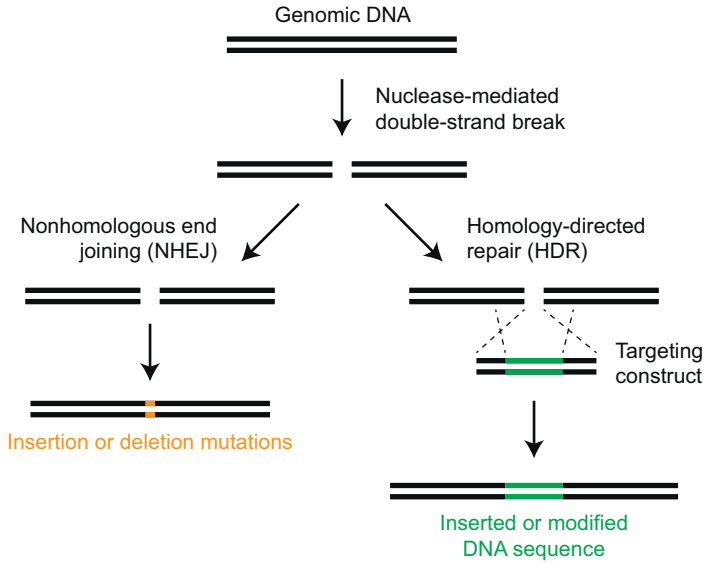


FIGURE 12.3 Conceptual overview of nuclease-mediated genome editing. A nuclease causes a double-strand break at a specific site of the genome. Following the break, a scientist can employ either nonhomologous end-joining (NHEJ) pathways to induce a deletion or insertion, or homology-directed repair (HDR) pathways and homologous recombination to introduce foreign DNA sequences.

CRISPR/Cas9 is currently the most common and most efficient DNA editing technique, used by thousands of laboratories. CRISPR stands for "Clustered Regularly Interspaced Short Palindromic Repeats," and refers to clusters of short, repeated DNA sequences normally used by some strains of bacteria as a method of protection against foreign nucleic acids (usually from invading viruses). Cas is a "CRISPR-associated protein," and Cas9 is originally derived from the bacterium *Streptococcus pyogenes*. In this species, the Cas9 nuclease uses templates provided by the CRISPR sequence to cleave foreign DNA. In the lab, the CRISPR/Cas9 system has been adapted to induce site-specific double-strand breaks in the genome of virtually any organism.

The most common method of using CRISPR/Cas9 technology is to introduce two DNA sequences into a cell or organism (Fig. 12.4). One DNA sequence encodes the Cas9 protein. The other DNA sequence codes for a **single-guide RNA (sgRNA)** that serves as a template for genome targeting. Twenty nucleotides at the 5′ end of the sgRNA direct Cas9 to a specific genomic site using standard DNA−RNA complementary base pairing. Importantly, in the genome to be targeted, the 20 nucleotides must lie immediately 5′ of a sequence called a protospacer adjacent motif (PAM) sequence, a three-nucleotide sequence of NGG (N can be any of the four DNA bases followed by two

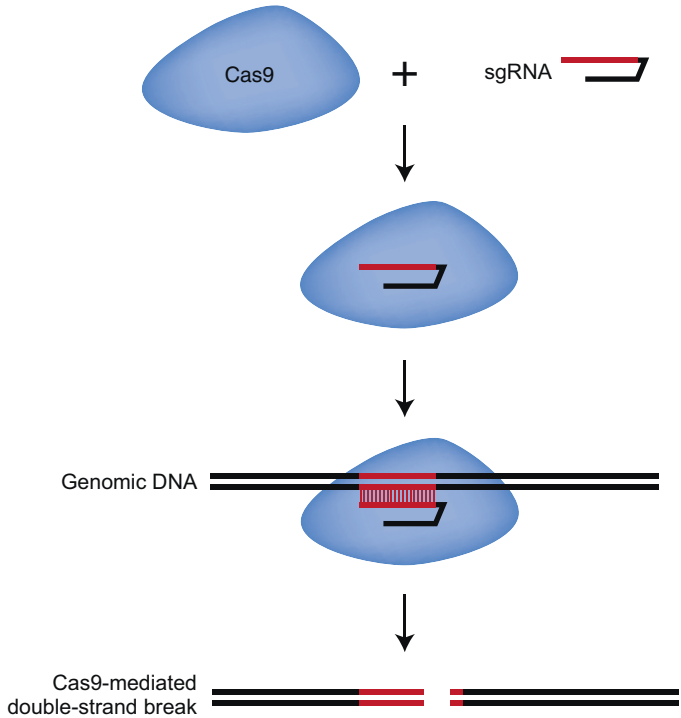


FIGURE 12.4 CRISPR/Cas9 genome targeting. A scientist delivers DNA to a cell encoding the Cas9 protein and a single-guide RNA (sgRNA). The sgRNA contains a 20 bp sequence that allows Cas9 to recognize a specific region of the genome and cause a double-strand break.

guanine bases). Therefore, a subtle limitation to using CRISPR/Cas9 gene targeting is that the break in the genome must occur next to this PAM sequence; however, as the three-nucleotide NGG sequence is incredibly common, it is possible to find a PAM sequence in almost any DNA sequence to be targeted. Thus, a scientist can design and deliver a sgRNA to induce Cas9-mediated genome breaks for any genomic sequence of 20 base-pairs plus the PAM. To use CRISPR/Cas9 to introduce a foreign DNA sequence (e.g., a transgene), a scientist must also include a DNA template containing the novel sequence flanked by homology arms, which will be used for the HDR.

Taken together, the strategies described above can be used to disrupt/remove genetic material at a precise location in a genome, modify/edit genetic material at a precise location in a genome, or insert a functional transgene either in a random or precise location in a genome. We now survey methods of delivering transgenic constructs to germ cells for the process of creating stable lines of genetically modified organisms.

ENGINEERING GENETICALLY MODIFIED ORGANISMS

Many protocols have been developed and refined over several decades to inject DNA constructs into recently fertilized eggs so that they can integrate into the genomes of the offspring. DNA is typically delivered using microinjection techniques (Chapter 11). This section describes the process of physically delivering DNA constructs to eggs to create genetically modified mice, flies, worms, and other animals.

Making Genetically Engineered Mice

To make a genetically modified mouse, a piece of foreign DNA is injected into the nucleus of a recently fertilized mouse egg. After the egg develops into a fetus and eventually a postnatal offspring, the new genetic material can become stably expressed and transferred to future offspring. In most institutions, individual investigators do not microinject the DNA construct into eggs themselves; instead, they pay a service fee to a core facility of trained technicians that specialize in microinjection procedures. As opposed to creating new lines of flies or worms, the harvesting and injection of recently fertilized mouse eggs is a technically challenging process that most laboratories choose to outsource.

The DNA construct is microinjected as a linear piece of DNA into the larger male **pronucleus** of a recently fertilized mouse egg (Fig. 12.5). A pronucleus is the haploid nucleus of either the egg or sperm, both present in the fertilized egg prior to their fusion. This stage is when insertion of foreign DNA into the genome can occur. Because the embryo is at the one-cell stage, the incorporated DNA is replicated and can end up in nearly all of the animal's cells in adulthood. For the DNA to be heritable, it must be incorporated into the germline (eggs or sperm).

After several embryos have been injected with the DNA construct, they are subsequently injected into **pseudopregnant** female mice (Fig. 12.6). The term pseudopregnant refers to the fact that the female is not actually pregnant but has been physiologically "tricked" into thinking she is pregnant by mating

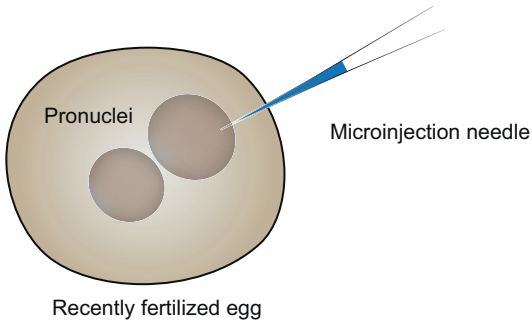


FIGURE 12.5 Microinjection of a DNA construct into mouse embryos. To make a genetically modified mouse, a scientist microinjects a DNA construct into one of two pronuclei in a recently fertilized egg.

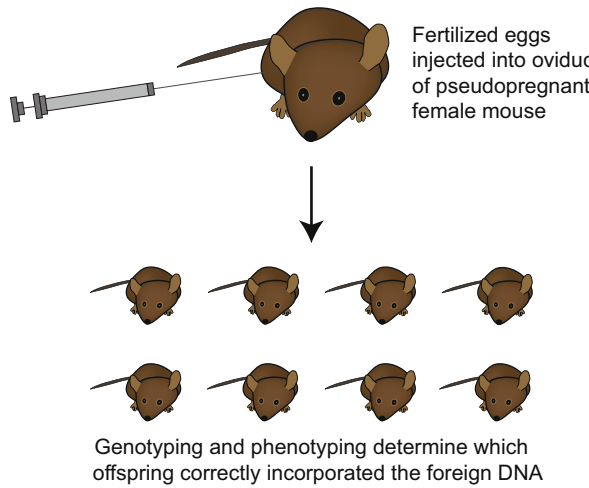


FIGURE 12.6 Injection of embryos to produce genetically modified mice. After microinjecting DNA constructs into embryos, the embryos are injected into the oviducts of pseudopregnant female mice. Each female typically produces 6–10 offspring; however, further genotyping and characterization is required to determine which offspring correctly incorporated the foreign DNA.

with a sterile male mouse. This process causes the females to produce the hormones and physiological changes necessary to develop the embryo into a viable fetus. The mouse eventually carries the embryos to term and genetically modified mice are born. Each genetic engineering core facility uses slightly different procedures, but often 2–4 pseudopregnant females are used to produce 20–30 offspring, called **founders**.

Founders represent the first generation of a new line of mice. However, not all offspring will have successfully incorporated the DNA construct. Therefore, once founders are born, an investigator must detect which mice have incorporated the foreign DNA by genotyping the animals using PCR or sequencing techniques (Chapter 10). Mice that seem to have incorporated the genetic material are mated with wild-type mice so their offspring can be further investigated. These mice are also tested to determine if the foreign DNA construct was transferred to offspring, and rigorously tested to ensure that the DNA is expressed in the appropriate tissues. For example, if a transgene is placed under the control of a specific promoter, the investigator ensures that the transgene is only expressed in the appropriate cell types and locations. The total time necessary to produce and analyze the validity of a genetically modified mouse line is usually around 6–9 months, including the time necessary for breeding, gestation, development, and analysis of genotype and phenotype in the lab.

Scientists can combine the above procedures with CRISPR/Cas9 techniques to create genetically targeted knockouts and knockins. The DNA sequences for Cas9 and the sgRNA, as well as any DNA sequences necessary for

HDR, are microinjected into the pronucleus of a recently fertilized egg. Once expressed, the functional Cas9 protein and sgRNA complex target the specific genomic location specified in the sgRNA. This strategy is now commonly used to produce genetically engineered mice.

Making Genetically Engineered Flies

The classical process of making a transgenic fly takes advantage of specialized, endogenous segments of DNA called **transposons** (also called **transposable elements**) that can move from one position to another within the genome. These mobile genetic sequences have only modest target site selectivity within the genome and can therefore insert themselves into random genomic locations. In **transposition**, an enzyme called **transposase** acts on specific DNA sequences at the end of a transposon, first disconnecting it from the flanking DNA and then inserting it into a new target DNA site.

The most commonly used transposons in *Drosophila* are called **P elements**. To make a transgenic fly using P elements, a DNA construct is first inserted between the two terminal sequences of the *Drosophila* P element (Fig. 12.7). The terminal sequences enable the P element to integrate into *Drosophila* chromosomes when the P element transposase enzyme is also present. The investigator injects the transgenic construct into a very young fruit fly embryo along with a separate plasmid containing the gene encoding

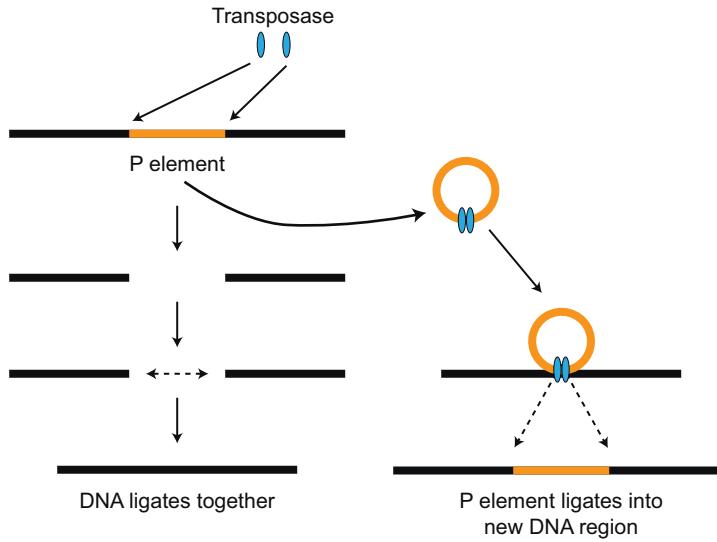


FIGURE 12.7 P elements in flies. The transposase enzyme splices a P element out of its genetic environment and inserts it into a new location in the genome. To make a transgenic fly, a scientist can create a DNA construct containing a P element with a transgene of interest and microinject the construct into fly embryos. The transposase enzyme will then ligate the P element with the foreign DNA into the genome.

the transposase (Fig. 12.8) If this technique is performed correctly, the injected gene enters the germ line as the result of a transposition event.

Scientists also routinely engineer genetically modified flies using CRISPR/Cas9, either by microinjecting the DNA sequences for Cas9 and the sgRNA into fly embryos, or by using previously established CRISPR/Cas9 fly lines (Fig. 12.9). The coding sequences for Cas9 and the sgRNA can be driven by

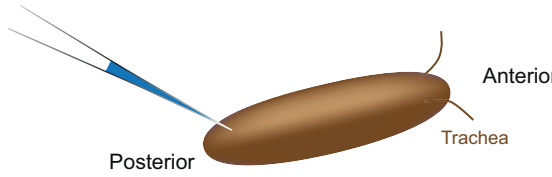


FIGURE 12.8 Microinjection of a DNA construct into fly embryos. To make genetically modified flies, a scientist microinjects a DNA construct into the posterior end (the side opposite the trachea) of fresh *Drosophila* embryos.

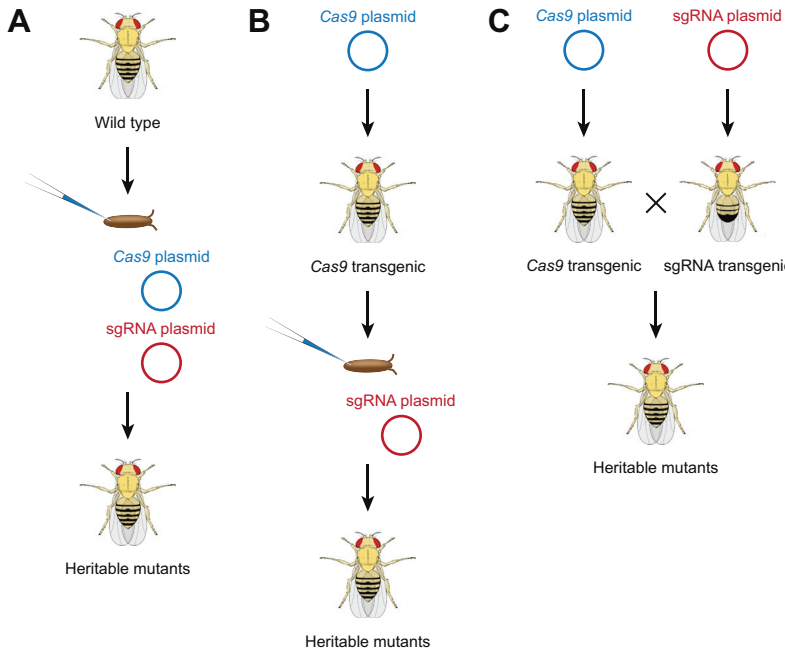


FIGURE 12.9 Methods of CRISPR/Cas9 genome targeting in flies. (A) DNA plasmids containing the sequences for Cas9 and a specific sgRNA are microinjected into wild type fly embryos. (B) A scientist creates or obtains a *Cas9* transgenic fly line, and a DNA plasmid containing a specific sgRNA sequence is microinjected into the embryos. (C) A scientist creates or obtains a *Cas9* transgenic fly line and a fly line expressing a specific sgRNA sequence. Crossing the two fly lines creates heritable CRISPR mutants.

cell-type–specific promoters or combined with the Gal4/UAS system (described below) for cell-type–specific genome editing.

Making Genetically Modified Worms

The process of introducing transgenes in *C. elegans* is conceptually similar to the process of introducing transgenes in flies. An important difference is that worms are hermaphrodites and capable of carrying their own fertilized eggs. DNA sequences are microinjected into the gonad of a young adult hermaphrodite at the time the embryos are dividing. The gonad runs the length of the worm on the dorsal side, and, with practice, an investigator can become skilled at injecting the embryos (Fig. 12.10).

Making Other Genetically Modified Organisms

Although mice, flies, and worms are the most frequently used genetic model organisms, in theory, a scientist can insert genetic material into any animal species by correctly injecting DNA into a recently fertilized egg. To target genetic material to specific locations within the genome, a scientist can microinject DNA sequences encoding the Cas9 protein and a sgRNA to target a specific genomic location. Even just a decade ago, there were only minimal examples of successful genetic targeting in model organisms outside of mice. CRISPR/Cas9 gene editing techniques have now allowed for unprecedented genetic tractability of potentially any organism.

BINARY EXPRESSION SYSTEMS

The most straightforward method of making a genetically modified organism is to insert genetic material into the genome to create a stable line of animals. To further refine the spatial or temporal expression of a gene, a scientist can employ a **binary expression system** (Fig. 12.11 and Table 12.3). To use a binary expression system, a scientist creates two different lines of transgenic

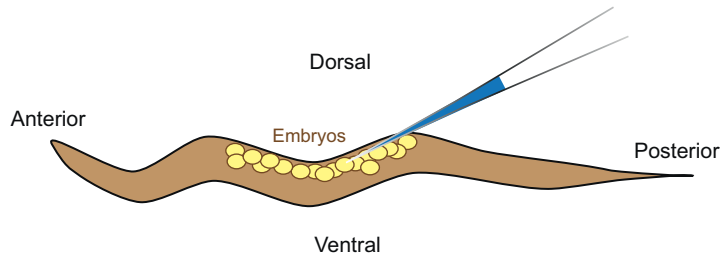


FIGURE 12.10 Microinjection of a DNA construct into worm embryos. To make genetically modified worms, a scientist microinjects a DNA construct into a worm's dorsal surface at the time embryos are dividing.

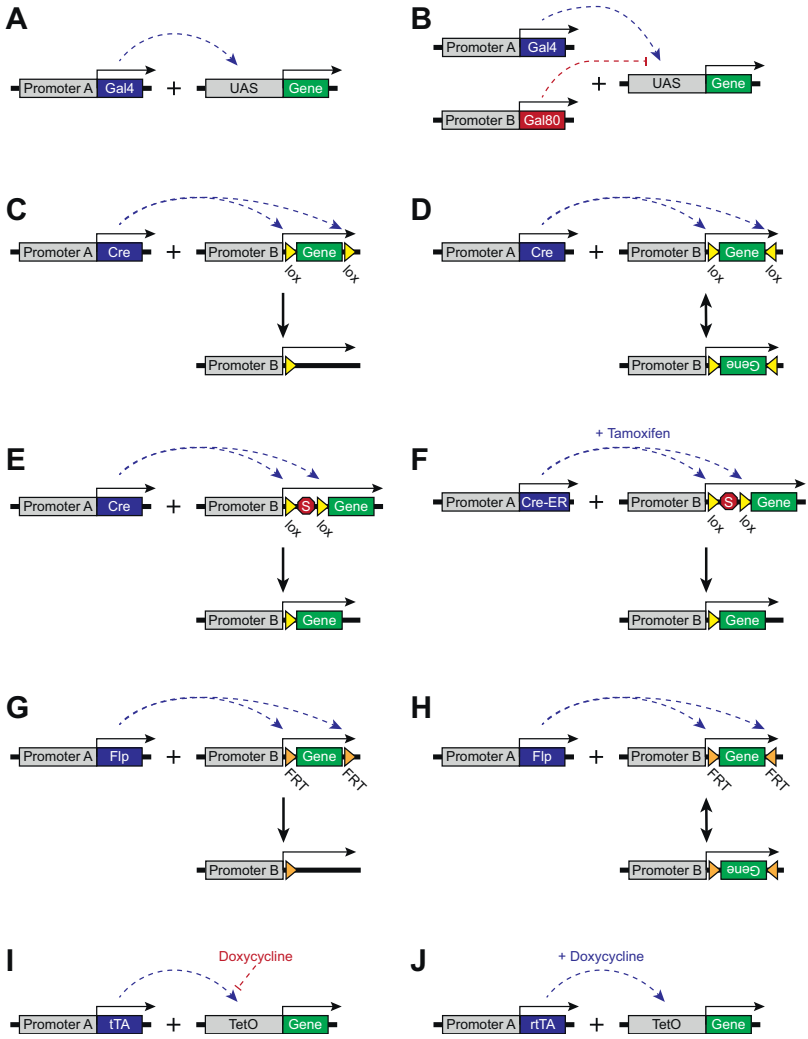


FIGURE 12.11 **Binary expression systems.** These systems use two or more DNA constructs (and sometimes administration of compounds) to spatially and/or temporally restrict gene expression. (A) The Gal4/UAS system. Promoter A drives expression of the transcription factor Gal4. Gal4 binds to the UAS on a second DNA sequence and activates expression of a gene. Therefore, expression of the gene is indirectly controlled by Promoter A. (B) The Gal4/Gal80/UAS system. Promoter A drives expression of Gal4, which binds to the UAS and activates expression of a gene. Promoter B drives expression of Gal80, which prevents binding of Gal4 to the UAS. Therefore, the gene is expressed in cells in which Promoter A is active but not in cells in which Promoter B is active. (C) The Cre/lox system. Promoter A drives the expression of Cre recombinase and Promoter B drives the expression of a gene flanked by lox sites. Cre recognizes lox sites in the same orientation and catalyzes a recombination event that removes DNA within the sites, leaving a single lox site behind. Therefore, the gene will be expressed in cells in which Promoter B is active but not in cells in which Promoter A is active. (D) The Cre/lox system with inverted lox

animals, each line expressing one of the required DNA sequences, and then mates the animals to produce offspring that express both DNA sequences. Alternatively, a scientist can combine a genetically modified organism with another DNA delivery method, such as the use of a virus that contains another genetic sequence (Chapter 11), to express both of the elements of the binary system.

These flexible binary expression systems have numerous advantages. The main advantage is that they provide control over the timing and location of transgene expression. Another substantial advantage is the ability to efficiently mix and match combinations of animals, saving time and effort in making new genetic constructs and genetically modified organisms.

The Gal4/UAS System

The **Gal4/UAS system** is a binary expression system primarily used in *Drosophila*, although it has also been used in zebrafish and other model organisms. In one DNA construct, a cell-specific promoter drives the expression of the gene encoding **Gal4**, a transcription factor normally expressed in yeast. In a second construct, a transgene of interest is regulated by a promoter sequence called an **upstream activation sequence (UAS)**. The Gal4 protein binds to the UAS sequence and causes high expression of the transgene

sites. When Cre recognizes two lox sites in inverted directions (facing each other), Cre rearranges the DNA between the lox sites so that it becomes "flipped" in the reverse orientation. The DNA will continue to flip as long as Cre is present. This system can be useful with different combinations of binary expression systems, as in Fig. 12.13. (E) The Cre/lox system used to remove a stop site. Cre is used to remove a stop codon that prevents the expression of a gene. Therefore, the gene is only expressed in cells in which promoters A and B are both active. (F) The Cre-ER/lox system. ER sequesters Cre in the cytoplasm until a scientist administers tamoxifen, which binds to ER and causes Cre to translocate to the nucleus, where it can interact with lox sites to remove a stop codon preventing expression of a gene. Therefore, the gene is only expressed in cells in which Promoters A and B are active and only when tamoxifen is administered. (G) The Flp/FRT system. Promoter A drives the expression of Flippase recombinase and Promoter B drives the expression of a gene flanked by FRT sites. Flp recognizes FRT sites in the same orientation and catalyzes a recombination event that removes DNA within the sites, leaving a single FRT site behind. Therefore, the gene will be expressed in cells in which Promoter B is active but not in cells in which Promoter A is active. (H) The Flp/FRT system with inverted FRT sites. When Flp recognizes two FRT sites in inverted directions (facing each other), Flp rearranges the DNA between the FRT sites so that it becomes "flipped" in the reverse orientation. The DNA will continue to flip as long as Flp is present. (I) The Tet-off system. Promoter A drives expression of the tetracycline transactivator (rTA) protein. tTA binds to the tetO promoter, activating expression of a gene. Therefore, gene expression is indirectly controlled by Promoter A. If a scientist administers a drug called doxycycline, tTA is reversibly inhibited. (J) The Tet-on system. Promoter A drives the expression of the reverse tetracycline transactivator (rtTA). rtTA is unable to bind to the tetO promoter unless doxycycline is present. Therefore, Promoter A indirectly activates expression of a gene, but only when a scientist administers doxycycline.

TABLE 12.3 Promoters in Binary Transgenic Systems.

Binary transgenic system	Cell and tissue specificity is conferred by the promoter(s) regulating:	The gene of interest is directly regulated by the promoter called:
Gal4/UAS	Gal4	UAS
Gal4/Gal80/UAS	Gal4 and Gal80	UAS
Cre/lox	Cre	Any promoter can be used … lox sites determine the location of DNA recombination
Flp/FRT	Flp	Any promoter can be used … FRT sites determine the location of DNA recombination
tTA/TetO	tTA	tetO
rtTA/tetO	rtTA	tetO

(Fig. 12.11A). Therefore, the transgene will only express in cells defined by the promoter regulating Gal4.

Why not simply place the transgene under the control of the promoter regulating Gal4? There are at least two advantages to using the Gal4/UAS system. First, the UAS is a very strong promoter, expressing transgenes at much higher levels than endogenous promoters. Second, Gal4 lines now exist for thousands of genes in the *Drosophila* genome. There are also many lines of flies in which a UAS drives a specific transgene. Therefore, the *Drosophila* scientific community has created an amazingly efficient system of creating and delivering transgenes to many different cell types (Fig. 12.12).

Gal4/UAS can also be used in combination with another protein, **Gal80**. The Gal80 protein binds to and inhibits the activity of Gal4. Therefore, a second promoter element can drive expression of Gal80 to further restrict the cells expressing the transgene (Fig. 12.11B). Furthermore, Gal80 can be made temperature sensitive: the Gal80ts protein is active at 19°C but not 30°C. Therefore, it is possible to express a transgene in specific cells using the Gal4/UAS system and to regulate the timing of transgene expression by controlling the temperature, and thus the expression of Gal80ts.

The Cre/lox System

The **Cre/lox system** is a site-specific recombination system widely used in mice. The enzyme **Cre recombinase**, originally derived from the P1

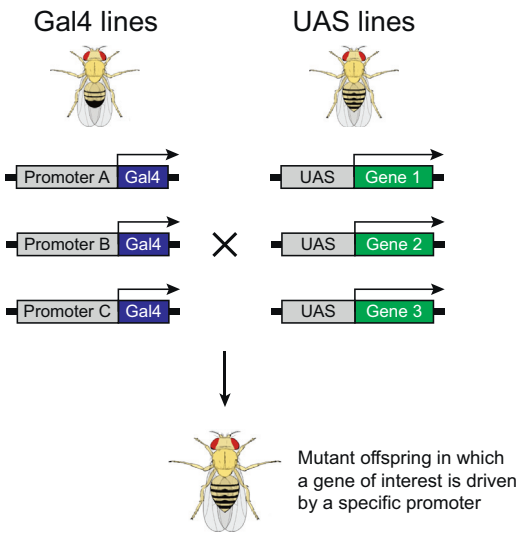


FIGURE 12.12 Utility of the Gal4/UAS system in *Drosophila*. The *Drosophila* community has created thousands of fly stocks in which endogenous promoters drive cell-specific expression of Gal4 transcription factors. In turn, there are thousands of fly stocks in which a transgene of interest is placed under the control of a UAS sequence. Therefore, a scientist can cross any combination of a Gal4 line with a UAS line to drive transgene expression in specific cell types. This system has been used extensively in *Drosophila* as well as other model organisms, including *Xenopus* and zebrafish.

bacteriophage, recognizes specific 34 base-pair DNA sequences called **lox sites**. Pairs of lox sites are placed on either side of a sequence of DNA. The DNA sequence is said to be **floxed**—it is *flanked* by pairs of *lox* sites. Each lox site has an asymmetric sequence in the middle, providing the sequence with directionality. This directionality is important because Cre processes DNA differently depending on the orientation of the lox sites:

- Cre excises any DNA sequences between two lox sites oriented in the same direction (Fig. 12.11C).
- Cre inverts any DNA sequences between two lox sites oriented in inverted directions (Fig. 12.11D)

Therefore, the Cre/lox system allows for rearrangement of DNA in crosses between two different lines of genetically modified mice. Many knockin mouse lines now exist in which Cre recombinase is driven by cell-specific promoters. The ability to engineer various combinations of lox sites and specific gene sequences provides useful experimental possibilities. For example, scientists have produced knockin mice that express a "floxed stop" sequence, a stop codon flanked by lox sites in the same direction, just upstream of an endogenous gene or a transgene. In this mouse line, the gene will not be expressed because of the stop codon; however, in the presence of Cre recombinase, the stop codon will be removed and the gene will be expressed (Fig. 12.11E).

An alternative strategy, usually referred to as a **double-inverse orientation (DIO) construct** or **Flex construct**, is the use of two different pairs of lox sites

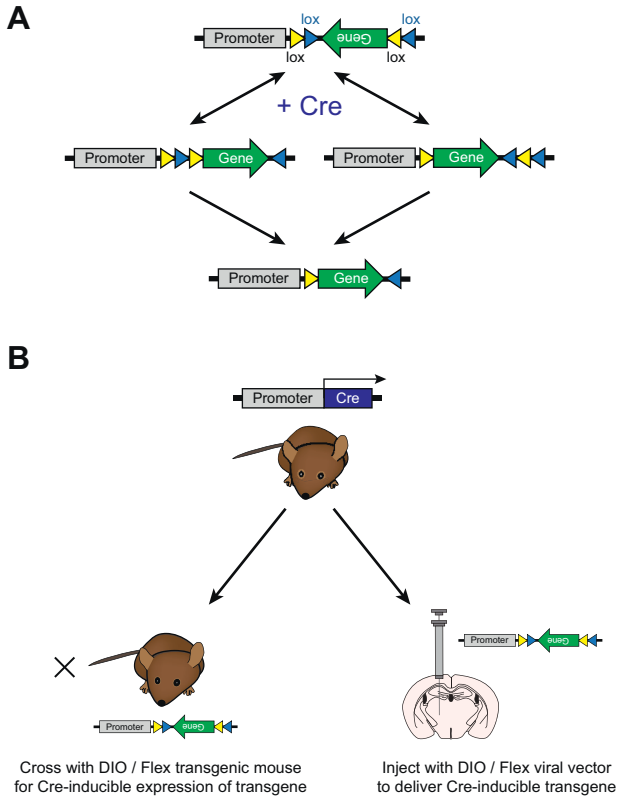


FIGURE 12.13 Double-inverse orientation (DIO) or Flex DNA constructs. (A) A DIO/Flex construct is composed of a promoter driving expression of a gene flanked by two different pairs of inverted lox sites. The gene is oriented in the reverse orientation and cannot code for a functional protein. In the presence of Cre, recombination occurs in one of two ways, depending on which pair of lox sites Cre interacts with first. Eventually, Cre excises one of each of the different lox sites so that the gene is locked in the correct orientation. (B) To use a DIO/Flex construct in practice, a mouse expressing Cre recombinase is either crossed with a transgenic mouse expressing a DIO/Flex construct, or the mouse is stereotaxically injected with a viral vector delivering the DIO/Flex construct. Therefore, the transgene is expressed only in cells expressing Cre.

to cause recombination that leads to Cre-dependent expression of a gene (Fig. 12.13A). A mouse carrying the gene for Cre can be mated with a mouse carrying a DIO/Flex construct for cell-type specific expression of a transgene. Alternatively, a mouse carrying the gene for Cre can be injected with a virus carrying a DIO/Flex viral vector. The virus may infect many cells within a local area, but only the cells expressing Cre will cause expression of the virally delivered transgene.

Investigators can achieve temporal control using the Cre/lox system by fusing the gene encoding Cre recombinase with the gene encoding the estrogen receptor (ER). This protein is normally sequestered in the cytoplasm, but upon

binding to its ligand, translocates to the nucleus. Scientists have developed a **Cre-ER** fusion transgene in which ER binds to the chemical **tamoxifen** rather than its natural ligand, estrogen. The Cre-ER fusion protein is sequestered in the cytoplasm and unable to access lox sites in the genome. When a scientist injects tamoxifen (or a tamoxifen derivative) into a mouse, it binds to Cre-ER, causing it to translocate into the nucleus where it can recombine DNA (Fig. 12.11F). This strategy provides temporal control over gene expression, as recombination will only occur when a scientist administers tamoxifen.

A useful application of the Cre-ER system, commonly referred to as the **TRAP system** ("Targeted Recombination in Active Populations"), labels transiently active neurons during a fixed time period (Fig. 12.14). In this

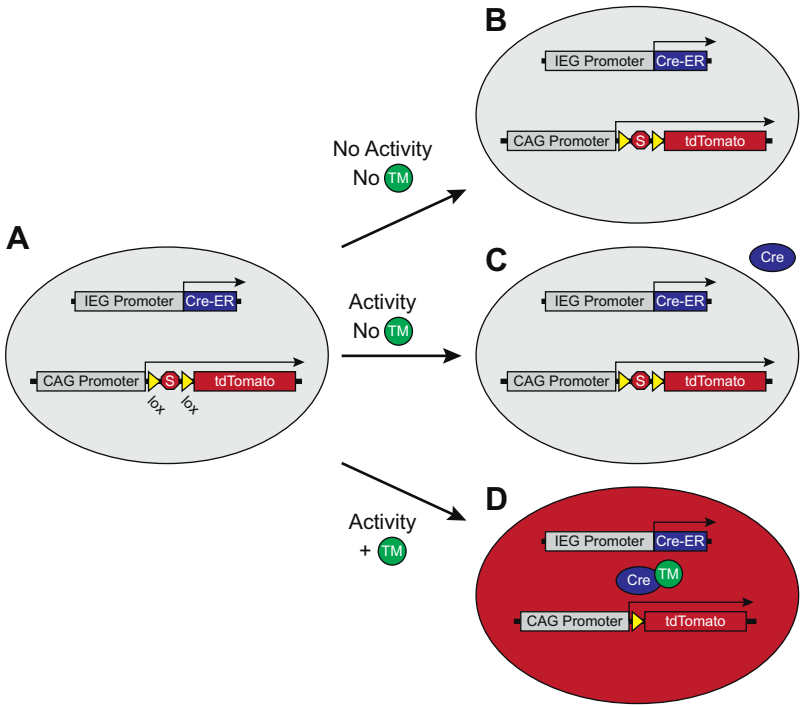


FIGURE 12.14 The TRAP system transiently targets active neurons. Two genetic constructs are expressed in all cells. The first construct uses a promoter for an immediate early gene (IEG) to drive expression of Cre-ER. The second construct uses a universal CAG promoter to drive expression of a reporter protein, such as tdTomato, but a stop codon prevents expression. (A) If the cell is not active and a scientist does not administer tamoxifen, neither Cre-ER nor tdTomato are transcribed. (B) If the cell is active, the IEG promoter will drive expression of Cre-ER. However, without the addition of tamoxifen, Cre-ER will be sequestered in the cytoplasm and cannot interact with lox sites on the tdTomato construct. Therefore, tdTomato will not be transcribed. (C) If the cell is active and the scientist administers tamoxifen, Cre-ER will be expressed and able to translocate to the nucleus where it can interact with lox sites to remove the stop codon. tdTomato will be expressed and cause fluorescence, indicating that the cell was active during the time that tamoxifen was administered (D).

system, Cre-ER is driven by the promoter for an immediate early gene (IEG—Chapter 7), a gene that tends to be expressed only when neurons increase activity. Cre-ER will therefore only be expressed when the neuron is active; however, it will be unable to enter the nucleus unless tamoxifen is administered. A second construct uses a universal promoter to drive expression of a reporter protein; however, a stop codon prevents expression unless Cre is present. Therefore, neurons will only express the reporter protein when the cell is active and when tamoxifen is present. This strategy allows scientists to identify neurons active only when they choose to administer tamoxifen, such as during exposure to a stimulus. In subsequent histological analysis, neurons that were active can be identified by examining expression of the reporter protein. Alternatively, instead of marking active cells by expressing a reporter protein, the TRAP method can be used to drive expression of a neuromodulator transgene, such as ChR2 or a DREADD, or a fluorescent biosensor, such as GCaMP. This method allows scientists to study populations of cells based on their activity patterns as opposed to their gene expression profiles.

The Flp/FRT System

The **Flp/FRT system** is a site-specific recombination system analogous to the Cre/lox system described above. The enzyme **Flippase recombinase (Flp)**, originally derived from yeast, recognizes specific 34 base-pair DNA sequences called **Flippase recognition targets (FRTs)**, excising DNA between identical Frt sites (Fig. 12.11G) and inverting DNA between inversely oriented FRT sites (Fig. 12.11H). This system is used in both flies and mice. The Flp/FRT system and Cre/lox systems of binary expression can be combined for highly specific transgene expression in multiple subsets of cell types in the same organism.

The Tet-off/Tet-on System

The **Tet-off/tet-on system** is a genetic tool used to temporally control transgene expression. In the Tet-off system, a promoter drives the expression of the tetracycline transactivator (tTA) protein. This protein binds to another DNA sequence called the tetO operator, which regulates the expression of a transgene. Therefore, in normal conditions, the transgene expression is regulated by the promoter that regulates tTA. However, a scientist can feed or administer a derivative of the antibiotic tetracycline called **doxycycline** into these mice. Doxycycline binds to tTA and renders it incapable of binding to tetO. Therefore, expression of the transgene can be reversibly inhibited with the addition of doxycycline (Fig. 12.11I). The Tet-on system works in the opposite way. A reversed tTA (rtTA) is only capable of binding to tetO when doxycycline is present. Therefore, the addition of doxycycline causes transcription of the transgene (Fig. 12.11J). Like the Tet-off system, the effect of doxycycline is reversible.

These systems are very useful for expressing transgenes at specific times. It is also possible to control levels of transgene activation. Increasing the concentration of doxycycline using the Tet-off system gradually lowers and then shuts off transgene expression in a dose-dependent manner. Similarly, increasing concentrations of doxycycline using the Tet-on system gradually increases transgene expression.

DISRUPTING GENE PRODUCTS

In addition to disrupting gene expression by directly targeting the genome itself, scientists can perturb endogenous gene function by interfering with a gene's mRNA transcript, effectively inhibiting gene function. These techniques are tractable in a wide variety of vertebrate and invertebrate species.

RNA Interference

RNA interference (RNAi) is a technique for silencing gene expression using endogenous cellular mechanisms that degrade mRNA. Cells naturally use RNAi to regulate normal gene expression, as well as to degrade foreign mRNA transcripts from viruses. Scientists can utilize this natural process to selectively **knock down** gene expression and degrade mRNA for a gene of interest. Though knockdown does not reach 100% efficiency, a good knockdown experiment will eliminate most of the targeted mRNA.

The natural process of RNAi begins when a piece of double-stranded RNA (dsRNA) interacts with an enzyme called **Dicer** (Fig. 12.15). Dicer cleaves the dsRNA into small fragments, about 23 base-pairs long, called **small interfering RNAs (siRNAs).** The siRNA—Dicer complex recruits additional cellular proteins to form an **RNA-induced silencing complex (RISC)**. RISC interacts with and degrades RNA strands that have a complementary nucleotide sequence to one of the siRNA strands. Therefore, if one of the siRNA strands has a sequence complementary to an mRNA strand, the RNAi process will degrade the mRNA before it is translated by a ribosome into a functional protein.

Scientists can take advantage of this natural RNAi process to degrade an mRNA molecule and knockdown gene expression. One strategy is to deliver a dsRNA molecule in which one strand is complementary to the mRNA transcript to be silenced. The dsRNA molecule will attract Dicer, and the process of RNAi will begin. This process works well in invertebrate model organisms, but can cause a stress response in mammalian cells. Another strategy that is better for long-term, stable knockdown is to express a DNA sequence that is transcribed into an mRNA product called a **short hairpin RNA (shRNA)**. The shRNA contains complementary sequences at either end that cause the molecule to fold onto itself, producing a hairpin shape (Fig. 12.16) Dicer then recognizes the shRNA as it would a small dsRNA, and RNAi will degrade the

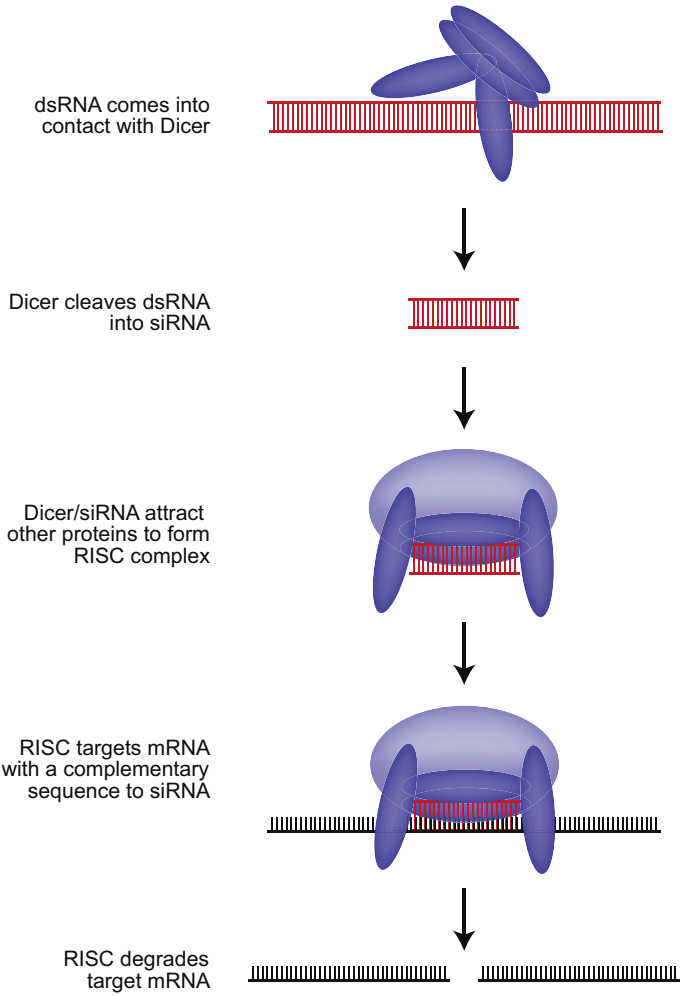
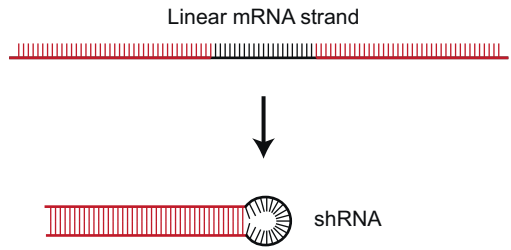


FIGURE 12.15 Conceptual overview of RNAi. When double-stranded RNA (dsRNA) comes into contact with the Dicer enzyme, Dicer cleaves the RNA into small fragments, creating small interfering RNA (siRNA) molecules. The siRNA–Dicer complex recruits other proteins to form an RISC complex. RISC targets mRNA molecules with a complementary sequence to one of the two siRNA strands, degrading the mRNA and knocking down gene expression.

FIGURE 12.16 Short hairpin RNA (shRNA). An shRNA molecule is composed of two complementary sequences separated by a small spacer. The molecule naturally folds into itself, resembling a hairpin shape. The two complementary strands serve as a double-stranded RNA molecule to interact with Dicer.



target RNA. A DNA sequence that codes for a shRNA can be delivered to cells using one of the DNA targeting methods described in Chapter 11, or stably incorporated into the genome using one of the methods described in this chapter.

When designing an RNAi experiment, it is important to consider the efficiency and specificity of the dsRNA or shRNA probes. An effective probe has the ability to knock down the expression of a specific mRNA strand without off-target effects. Usually, a scientist designs multiple RNAi probes against a gene of interest and compares their knockdown efficiency in cell culture experiments before choosing a probe for an actual experiment. The scientist designs these probes to have sequences unique to the gene of interest and no other genes—a process that can be performed using the BLAST application (described in Chapter 9). Other typical RNAi experimental controls include a probe with a scrambled sequence, as well as a probe against a separate, nonessential gene to show the specificity of the RNAi effect. The most conclusive experiments demonstrate that multiple probes again different regions of the same mRNA result in the same knockdown phenotype. The gold standard of RNAi control experiments is a phenotypic rescue, in which a version of the knocked down gene that is not targeted by the RNAi probe is delivered together with the RNAi probe to show that it can restore a normal phenotype. The rescuing gene is often an orthologue from another species with the same function but a slightly different genetic sequence.

Morpholinos

An alternative to RNAi is the use of **morpholinos** to block proper mRNA translation. This approach is useful for the few model organisms for which RNAi techniques do not seem to be effective, such as *Xenopus* frogs. Morpholinos are stable, synthetic 22–25 bp antisense oligonucleotide analogues designed to complement an RNA sequence. A morpholino and its complementary mRNA will hybridize to each other, preventing a ribosome from accessing the mRNA and translating it into a protein (Fig. 12.17). Unlike RNAi, morpholinos do not degrade their target mRNA molecules. Morpholinos have been used effectively in *Xenopus*, zebrafish, and chicks by injecting probes into eggs and producing morphant embryos. Controls used in morpholino experiments are similar to those used for RNAi experiments.

CONCLUSION

This chapter surveyed a wide variety of techniques used to disrupt endogenous gene function, replace endogenous genetic material, and insert functional transgenes into the genomes of a variety of model organisms. Just 10–20 years ago, genome targeting and editing technologies were laborious and only available in a limited number of model organisms. Nowadays, the universal

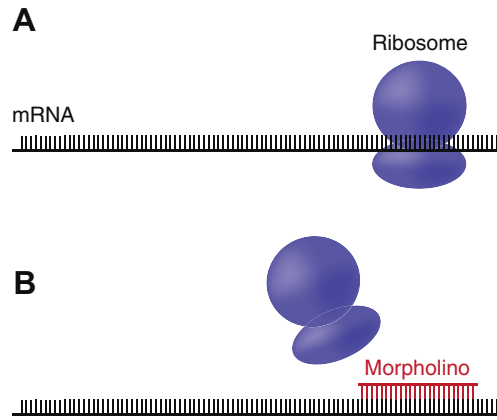


FIGURE 12.17 Morpholinos. (A) During translation, a ribosome binds to an mRNA molecule and translates specific nucleotide sequences into polypeptide chains. (B) A morpholino hybridizes with the mRNA strand, functionally blocking association with a ribosome.

applicability and ease of use of techniques such as CRISPR/Cas9 have made it possible to modify the genome of virtually any animal. Additionally, the wide variety of transgenes useful for studying the structure, activity, gain-of-function, and loss-of-function of neurons have allowed for unprecedented experimental control of specific cell types. The development of newer gene editing tools is an exciting area of research that is sure to allow for even more experimental possibilities in the future.

SUGGESTED READING AND REFERENCES

Books

Appasani, K. (Ed.), 2011. RNA Interference Technology: From Basic Science to Drug Development. Cambridge University Press.

Behringer, R., Gertsenstein, M., Nagy, K., Nagy, A., 2013. Manipulating the Mouse Embryo: A Laboratory Manual, fourth ed. Cold Spring Harbor Laboratory Press.

Review Articles

He, M., Huang, Z.J., 2018. Genetic approaches to access cell types in mammalian neural systems. Curr. Opin. Neurobiol. 50, 109–118.

Hsu, P.D., Lander, E.D., Zhang, F., 2014. Development and applications of CRISPR-Cas9 for genome engineering. Cell 157, 1262–1278.

Luo, L., Callaway, E.M., Svoboda, K., 2018. Genetic dissection of neural circuits: a decade of progress. Neuron 98, 256–281.

Sander, J.D., Joung, J.K., 2014. CRISPR-Cas systems for editing, regulating, and targeting genomes. Nat. Biotechnol. 32, 347–355.

Wiegert, J.S., Mahn, M., Prigge, M., Printz, Y., Yizhar, O., 2017. Silencing neurons: tools, applications, and experimental constraints. Neuron 95, 504–529.

Xu, J., et al., 2015. A toolkit of CRISPR-based genome editing systems in *Drosophila*. J. Genet. Genom. 42, 141–149.

Primary Research Articles—Interesting Examples from the Literature

Cong, L., et al., 2013. Multiplex genome engineering using CRISPR/Cas systems. Science 339, 819–823.

Fire, A., Xu, S., Montgomery, M.K., Kostas, S.A., Driver, S.E., Mellow, C.C., 1998. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. Nature 391, 806–811.

Guenthner, C.J., Miyamichi, J., Yang, H.H., Heller, H.C., Luo, L., 2013. Permanent genetic access to transiently active neurons via TRAP: targeted recombination in active populations. Neuron 78, 773–784.

Livet, J., et al., 2007. Transgenic strategies for combinatorial expression of fluorescent proteins in the nervous system. Nature 450, 56–62.

Panda, S.K., McGrew, M.J., 2021. Genome editing of avian species: implications for animal use and welfare. Lab. Anim.

Root, D.H., et al., 2020. Distinct signaling by ventral tegmental area glutamate, GABA, and combinatorial glutamate-GABA neurons in motivated behavior. Cell Rep. 32, 108094.

Thomas, K.R., Capecchi, M.R., 1987. Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells. Cell 51, 503–512.

Wang, H., et al., 2013. One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. Cell 153, 910–918.

Protocols

Doudna, J., Mali, P. (Eds.), 2016. CRISPR-cas: A Laboratory Manual. Cold Spring Harbor Laboratory Press.

Luo, Y. (Ed.), 2019. CRISPR Gene Editing. Humana Press.

Websites

Addgene CRISPR Resource. https://www.addgene.org/guides/crispr/.