INTRODUCTION

Scientists have determined the complete DNA sequences of the genomes for many organisms, including humans. By analyzing patterns in those sequences, they can estimate how many genes an organism has — humans, for example, have about 20,000. But sequence patterns alone don't specifically show what each gene does. How can we figure this out?

In this activity, you will explore a tool that can be used to determine a gene's function. You will then design your own version of this tool to examine genes that affect the colors and patterns on butterfly wings.

PART 1: Using CRISPR-Cas9 to Inactivate Genes

One way to determine a gene's function is to inactivate, or "knock out," that gene and then observe the effect on a cell or organism. Scientists can inactivate genes in cells growing in a lab or in model organisms, such as flies or mice.

1. With a partner, discuss what it means to inactivate a gene and how it could be done. Record your ideas.

2. Can you think of additional ways to find out what genes do? Record one or two ideas.

CRISPR-Cas9, or CRISPR for short, is a biotechnology tool that can edit or inactivate specific genes. To learn more about the CRISPR-Cas9 system, you can explore the "How It Works" section of the *CRISPR-Cas9 Mechanism and Applications* Click & Learn.

As shown in the Click & Learn, the CRISPR-Cas9 tool uses a DNA-cutting enzyme, or nuclease, called **Cas9**. Cas9 was discovered in bacteria as a way for bacteria to fight off viruses. Scientists combine Cas9 with an RNA molecule called a **guide RNA** to form a Cas9-RNA complex. Part of the guide RNA matches a **target DNA sequence** within the gene that the scientists want to inactivate.

You will now go through each step in the process of using CRISPR-Cas9 to inactivate a gene, using a sequence from an actual gene as an example.

Step 1: Targeting

First, the Cas9-RNA complex recognizes and binds to a three-nucleotide sequence called PAM, which stands for "proto-spacer adjacent motif." PAM sequences are abundant throughout the genome and can occur on either strand of DNA. Every PAM sequence has the form 5'-NGG-3', where the "N" can be any DNA nucleotide (A, C, G, or T).

3. The partial gene (DNA) sequence below contains multiple PAM sequences. Highlight **six PAM sequences** in the top (5’ to 3’) strand.

5'-GCACGGCGGAGCGGTTCTTGGCAGCGGCCGCACGATCTCGTTGCCGCCGG-**3'**
3'-CGTGCCGCCTCGCCAAGAACCGTCGCCGGCGTGCTAGAGCAACGGCGGCC-**5'**

Once Cas9 binds to a PAM sequence, it unwinds the DNA. If the guide RNA matches the DNA sequence *next to* the PAM, the guide RNA will bind to the complementary DNA strand. If not, the DNA will zip back together and Cas9 will keep binding to other PAM sequences until it finds the matching target DNA.

4. Below is a partial sequence of a guide RNA. The underlined section of the RNA is designed to match a specific target DNA sequence.

5'-GGCGGAGCGGUUCUUGGCAGGUUUUAGAGCUAGAAAUAGC-**3'**

Review the partial gene sequence reshown below. It contains a target DNA sequence that matches the guide RNA above. Highlight the **one PAM sequence** in the top (5’ to 3’) strand that is *next to* this target DNA sequence. (The sequence upstream, toward the 5’ end, of this PAM should match the underlined sequence in the guide RNA, which makes the opposite DNA strand complementary to the underlined sequence. Remember that U’s in RNA are equivalent to T’s in DNA.)

5'-GCACGGCGGAGCGGTTCTTGGCAGCGGCCGCACGATCTCGTTGCCGCCGG-**3'**
3'-CGTGCCGCCTCGCCAAGAACCGTCGCCGGCGTGCTAGAGCAACGGCGGCC-**5'**

Step 2: Binding

Once Cas9 binds to the correct PAM, the guide RNA binds to the complement of the target DNA sequence.

5. Write down the guide RNA sequence that binds to the DNA, and the DNA sequence that it binds to (the complement of the target DNA). Label the 5' and 3' ends for both the RNA and DNA strands.

Step 3: Cleaving

Once the guide RNA binds to the complement of the target DNA sequence, it activates the nuclease activity (DNA-cutting ability) of the Cas9 enzyme. Cutting DNA is also called “cleaving.” Cas9 always cleaves *both* strands of DNA. It cleaves both the target DNA and its complement three nucleotides upstream (toward the 5’ end) of the PAM sequence.

6. Rewrite the target DNA sequence and its complement below, indicating where Cas9 would cut both strands of DNA with a large space or vertical line (|).

Step 4: DNA Repair

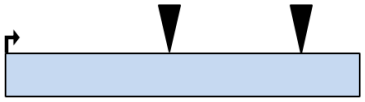
After Cas9 cleaves the DNA, cellular enzymes will attempt to repair the break. Most of the time, these enzymes repair the DNA without errors. However, Cas9 will keep cutting the DNA at the same location until an error is made.

7. DNA repair errors include losing or inserting random nucleotides at the cut site. Explain how these changes might inactivate a gene.

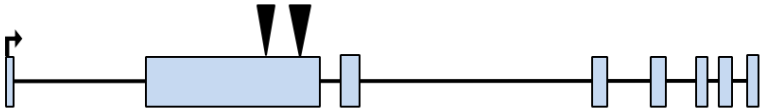
PART 2: Inactivating Genes in Butterflies

Robert Reed, a biologist at Cornell University, wanted to identify genes that are important in butterfly wing patterns. He and his colleagues used CRISPR-Cas9 to inactivate different genes, then observed the effects on butterflies. Models of three genes they inactivated are shown in Figure 1.

***optix* gene:**



***spalt* gene:**



***Distal-less* gene:**

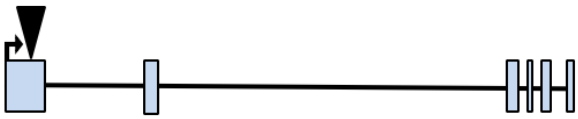


Figure 1. Models of three genes involved in butterfly wing patterns: *optix*, *spalt*, and *Distal-less*. The shaded rectangles represent **exons**, and the horizontal lines between the rectangles represent **introns**. The black triangles above the rectangles point to the **target DNA sequences**, which match the guide RNAs made by the scientists. The arrows at the start of each gene represent the transcription start sites.

1. What are exons and introns, and how do they differ?

2. Figure 1 shows that the target DNA sequences are all in exons. Why might scientists want to target sequences in exons, rather than introns, when inactivating genes?

PART 3: Designing a Guide RNA

You will now design your own guide RNA to inactivate a butterfly gene like Robert Reed did. Your goal is to knock out the *optix* gene in a species of butterfly called the painted lady (*Vanessa cardui*).

Your guide RNA must match a target DNA sequence in the gene that you want to knock out. A partial sequence of an exon from the butterfly's *optix* gene is shown below.

5'-CGACACCGGTTCCAGCGCTCGAGCCACGCGAAGCTTCAGGCGCTGTGGCTGGAAG
3'-GCTGTGGCCAAGGTCGCGAGCTCGGTGCGCTTCGAAGTCCGCGACACCGACCTTC

CGCACTACCAGGAAGCGGAGCGCCTCCGCGGTCGCCCGCTCGGGCCCGTCGACAA
GCGTGATGGTCCTTCGCCTCGCGGAGGCGCCAGCGGGCGAGCCCGGGCAGCTGTT

GTACCGGGTGCGGAAGAAGTTCCCTCTGCCGAGGACTATTTGGGACGGCGAACAG-**3'**
CATGGCCCACGCCTTCTTCAAGGGAGACGGCTCCTGATAAACCCTGCCGCTTGTC-**5'**

1. Underline a 20-nucleotide **target DNA sequence** in the top (5' to 3') strand of the exon above. Remember that this sequence should be directly upstream (on the 5' end) of a PAM sequence (5'-NGG-3').
2. Highlight the **PAM sequence** that is next to your target DNA sequence. This is where Cas9 will bind.
3. Rewrite the target DNA sequence and its complement below, indicating where Cas9 would cut both strands of DNA with a large space or vertical line (|).

4. Record the 20-nucleotide **guide RNA sequence** that matches your target DNA sequence. This sequence should *not* include the PAM.

5. Robert Reed's lab at Cornell University designed guide RNAs to knock out the *optix* gene in three species of butterflies: the painted lady (*V. cardui*), the common buckeye (*Junonia coenia*), and the Gulf fritillary (*Agraulis vanillae*). Figures 2–4 compare the wings of the wild-type (control) butterflies to those of *optix* knockouts (butterflies that had their *optix* gene inactivated).

Figure 2. *V. cardui* butterflies (left: wild-type control; right: *optix* knockout).

Figure 3. Wings of *A. vanillae* butterflies (left: wild-type control; right: *optix* knockout).

Figure 4. Wings of *J. coenia* butterflies (left: wild-type control; right: *optix* knockout).

a. For each species shown, describe how the wings of the *optix* knockout compare to those of the wild-type butterfly.

b. Based on the figures, predict the function of the *optix* gene. Does the gene have the same function in all three butterfly species? If not, explain how it may differ among the species.

c. Justify your prediction with evidence from all three figures. Be specific.

6. Launch the *CRISPR-Cas9: Mechanism & Applications* Click & Learn and select "How It's Used." Scroll through the videos on the right side of the screen and watch the three video interviews with Robert Reed. Revise your answers to Question 5 based on information from the videos.

EXTENSION: Determine the Function of a Different Gene

Use what you've learned from the previous parts of this activity to explore knocking out a different gene: the *spalt* gene in the *V. cardui* butterfly.

1. A partial sequence of an exon from the butterfly's *spalt* gene is shown below. This time, only one of the DNA strands is shown.

5'-
CGATATCTGGACCAATTTCAATAGCAACAGGACTACGTACTTTTCCTTCATATCCACTATTTCCAAATTCCCCA
CCAAGCAGTGTCTCATCTGGATGTCTTACACCTTTCCAAAGTAATCCCAACAGCATAATAGACAGTGACATA
ACTCGTGATCCCATATTTTATAATTCACTTTTACCGCGTCCTGGAAGTAATGACAACTCTTGGGAAAGTTTGA
TTGAAATTACTAAAACTTCAGAAACGTCAAAATTGCAACAGTTAGTAGATAATATTGATAACAAAGTTACTG
ATCCTAACGAGTGTATTGTATGTCATCGCGTCTTATCTTGTAAAAGTGCTTTACAGATGCACTACCGAACTCA
TACCGGGGAAAGACCTTTCAGATGTAAATTGTGCGGTCGTGCTTTTACTACAAAGGGCAATTTAAAAACTCA
TATGGGTGTCCATCG-**3'**

   a. Underline a 20-nucleotide **target DNA sequence** in the exon above. Remember that this sequence should be directly upstream of a PAM sequence.

   b. Highlight the **PAM sequence** that is next to your target DNA sequence.

   c. Rewrite the target DNA sequence, indicating where Cas9 would cut the DNA with a large space or vertical line (|).

2. Record the 20-nucleotide guide RNA sequence that matches your target DNA sequence.

3. The Reed lab used CRISPR-Cas9 to knock out the *spalt* gene in two species of butterflies: *J. coenia* and *V. cardui*. Figures 5–6 compare the wings of the wild-type butterflies to those of *spalt* knockouts.

Figure 5. Wings of *J. coenia* butterflies (left: wild-type control; right: *spalt* knockout).

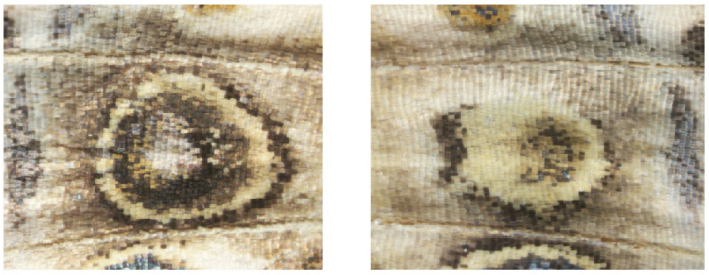


Figure 6. Closeups of markings on the wings of *V. cardui* butterflies

a. For each species shown, describe how the wings of the *spalt* knockout compare to those of the wild-type butterfly.

b. Based on the figures, predict the function of the *spalt* gene. Does the gene have the same function in both butterfly species? If not, explain how it may differ between the species.

c. Justify your prediction with evidence from both figures. Be specific.