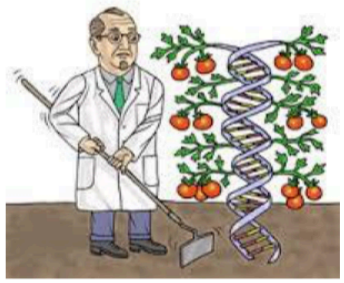
CRISPR/Cas technology

Introduction to CRISPR/Cas - Genetic Engineering

"The deliberate modification of the characteristics of an organism by manipulating its genetic material."

- Research
- Medicine (Protein/Enzyme production)
- Agriculture (Crops)
- Industrial Biotechnology (Biofuel production)
- Entertainment



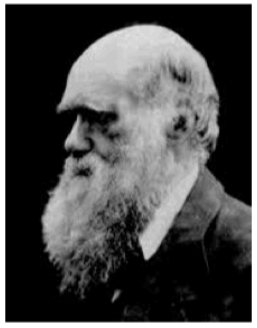
The way towars genetic Engineering

Rules

1859 Darwin
"*Origin of Species*"

1856-66 Mendel
"Mendelian inheritance"

1871 Mieska
Nucleic acids



Information

1944 Avery–MacLeod–McCarty
DNA as the genetic material

1953 Watson, Crick and Franklin
DNA structure

1961-1967 Genetic code

Basics

1970 Restriction Enzymes

1977 Sanger Sequencing

1983 PCR

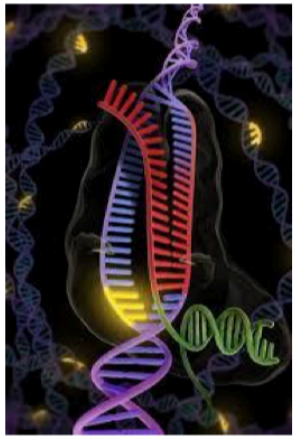
2003 Human Genome Project

Genome editing

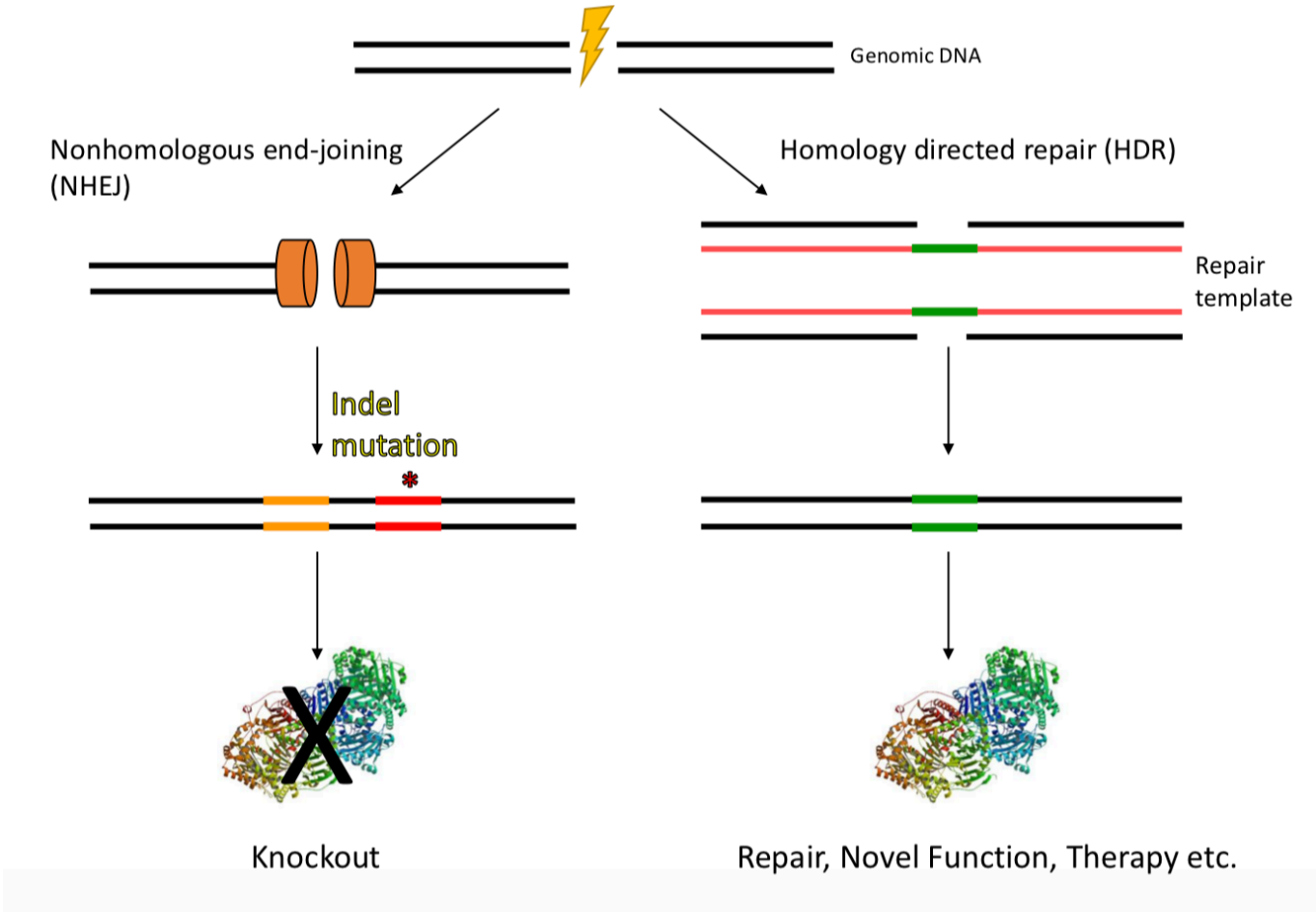
Zink Fingers

TALENs

CRISPRs



Genetic editing uses DNA repair pathways

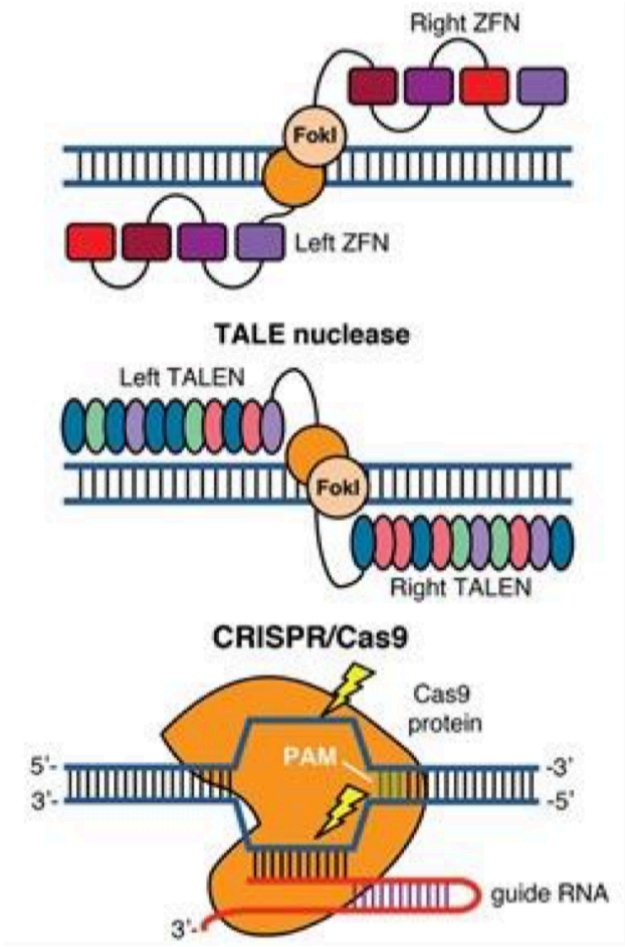


Genome Editing using Site Specific Nucleases

Zinc Fingers

TALEN

CRISPR

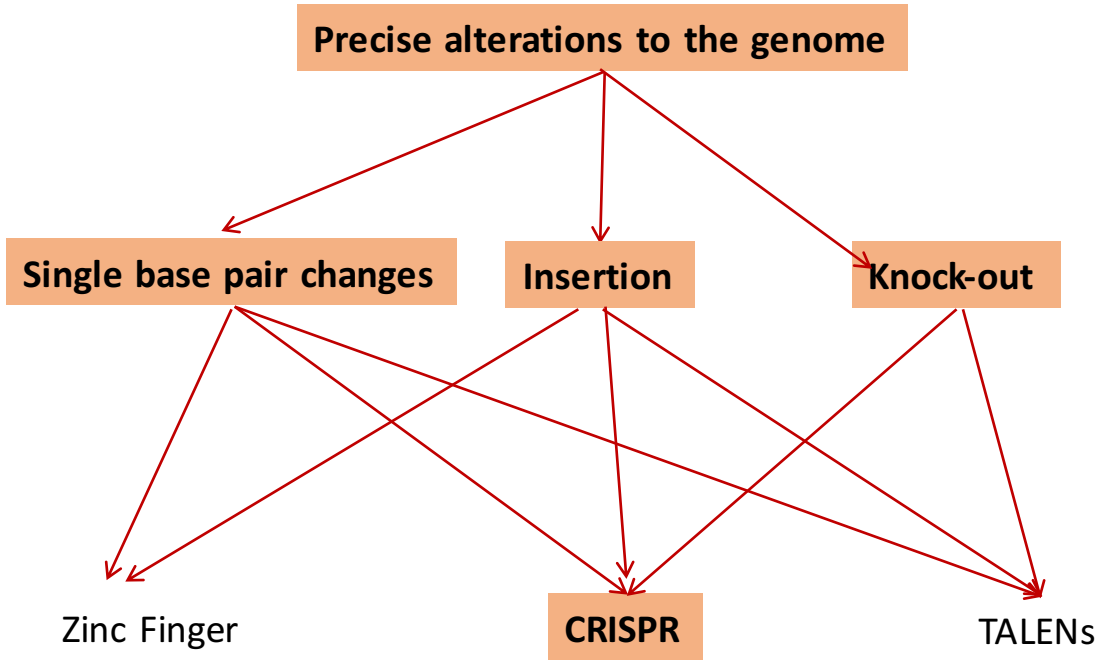


Protein based

RNA based

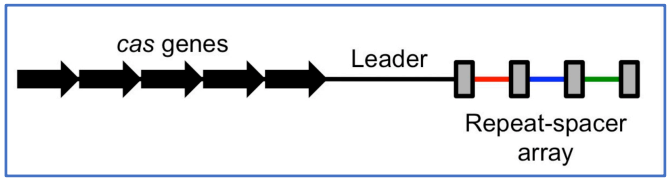
Adapted from Duke University

Genome Editing



CRISPR-Cas – Adaptive immune system in bacteria

A: CRISPR (**clustered regularly interspaced short palindromic repeats**) is a family of DNA sequences found within the genomes of prokaryotic organisms such as bacteria and archaea. These sequences are derived from DNA fragments from viruses that have previously infected the prokaryote and are used to detect and destroy DNA from similar viruses during subsequent infections. Hence these sequences play a key role in the antiviral defense system of prokaryotes.



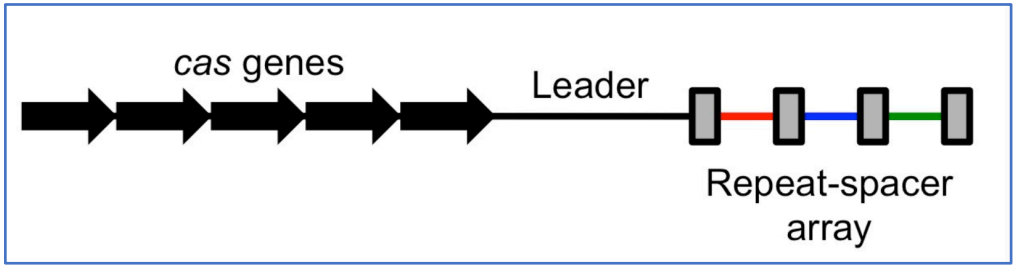
+

B. Cas9 (or "CRISPR-associated 9") is an enzyme that uses CRISPR sequences as a guide to recognize and cleave specific strands of DNA that are complementary to the CRISPR sequence.

Cas9 enzymes together with CRISPR sequences form the basis of a technology known as **CRISPR/Cas9** that can be used to edit genes within organisms

CRISPR-Cas – A. CRISPR locus

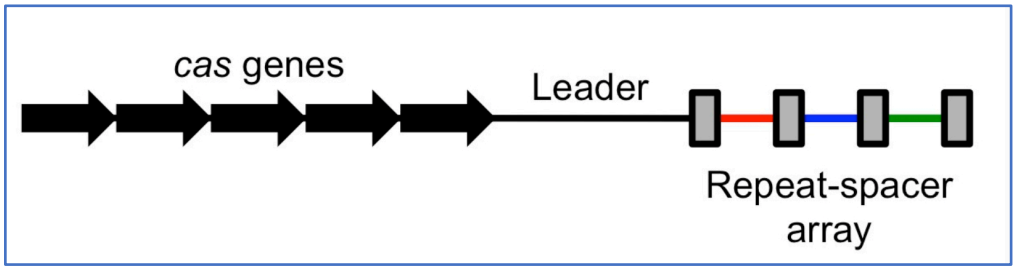
Simplified diagram of a CRISPR locus



The **three major components** of a CRISPR locus are shown: 1. ***cas* genes**, 2. **leader** and 3. **repeat-spacer array**. For the repeat-spacer array, repeats are shown as grey boxesc (typically range in size from 28 to 37 base pairs (bps), though there can be as few as 23 bp and as many as 55 bp), and spacers are colored bars

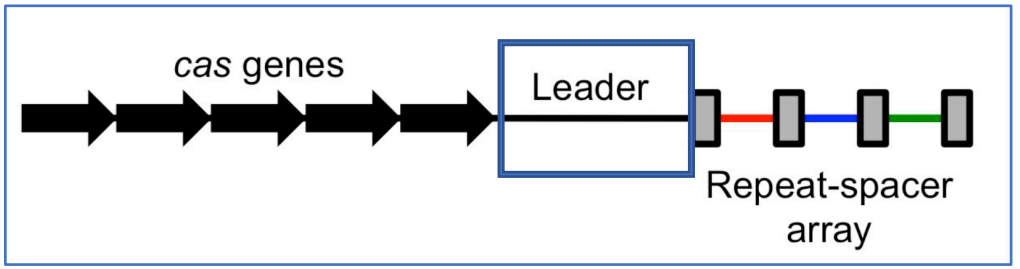
CRISPR-Cas – A. CRISPR locus

Simplified diagram of a CRISPR locus



CRISPR-Cas – A. CRISPR locus

Simplified diagram of a CRISPR locus

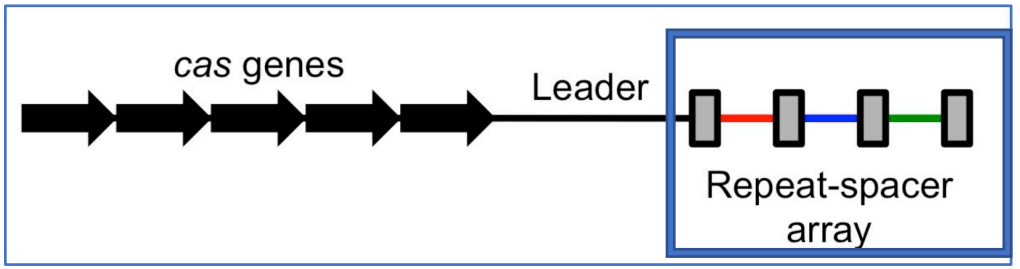


Leader Sequence

This sequence is an A-T reach sequence

CRISPR-Cas – A. CRISPR locus

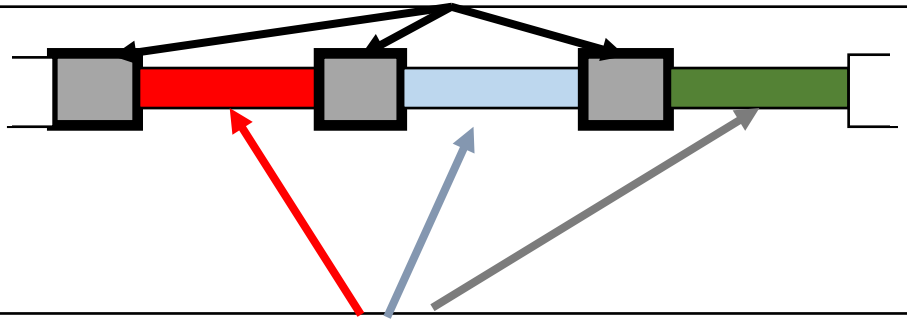
Simplified diagram of a CRISPR locus



REPEAT Spacer ARRAY

Shorts palindromic **repeats** (i.e. this is the same DNA sequence repeated in different places, gray boxes). These are part of the **bacterial genome** and typically range in size from 28 to 37 base pairs

Diagram of CRISPR locus in bacterial genome

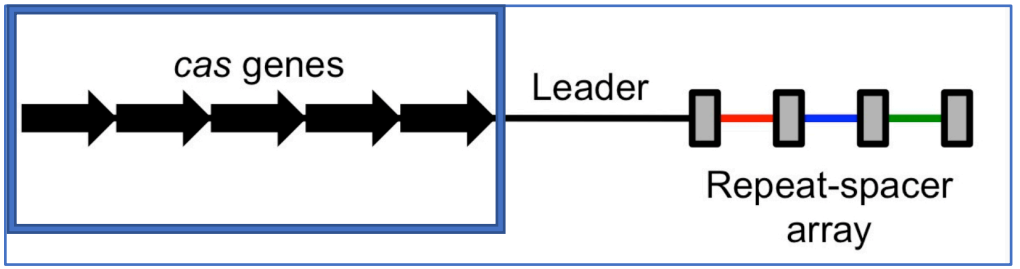


There are usually fewer than 50 units of the repeat-spacer sequence in a CRISPR array.

These bits/**spacers** are derived from **bacteriophage genome** (range 21 to 72 bp) and each one is different and these provide the guidance system for the adaptive immune system

CRISPR-Cas – A. CRISPR locus

Simplified diagram of a CRISPR locus



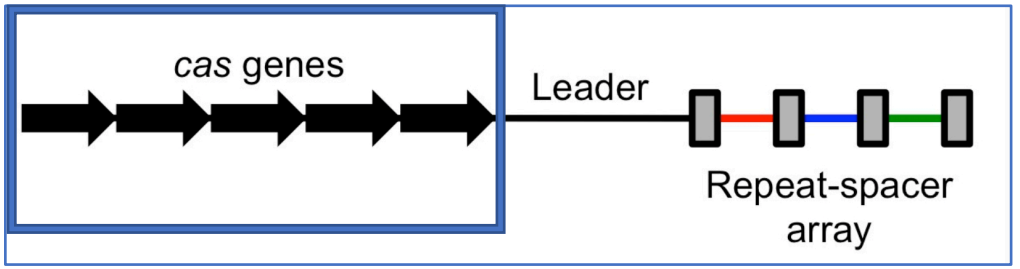
CAS genes

There are several other important regions of the bacterial DNA that are also always associated with the CRISPR locus and these provide the means for the palindromic repeat and the bacteriophage DNA sequences to actually destroy the bacteriophage.

These are called CRISPR Associated Sequences i.e. **Cas** genes

CRISPR-Cas – A. CRISPR locus

Simplified diagram of a CRISPR locus



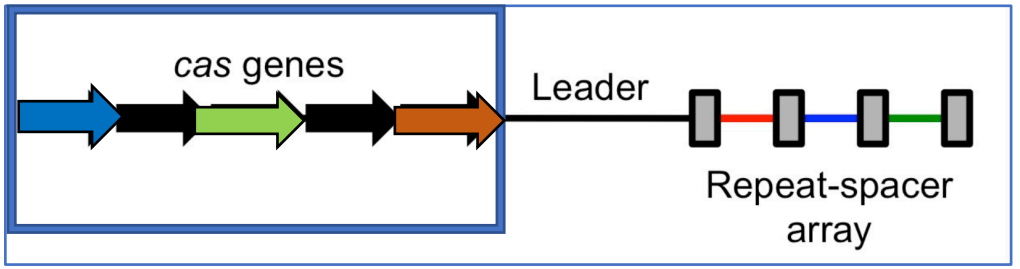
CAS genes

There are several other important regions of the bacterial DNA that are also always associated with the CRISPR locus and these provide the means for the palindromic repeat and the bacteriophage DNA sequences to actually destroy the bacteriophage.

These are called CRISPR Associated Sequences i.e. **Cas** genes

How does this genetic material in CRISPR locus then manage to kill bacteria ?

Simplified diagram of a CRISPR locus



CAS genes

The system can be slighty different in different types of bacteria but the best studies one is ***Streptococcus pyogenes*** so we will focus on that one

For the sake of simplicity lets focus on the 3 Cas genes (now colored arrows) most important for genetic engineering;

Codes for a **trans-activating CRISPR RNA (tracrRNA**) that will help in the process of ensuring the whole process only cuts bacteriophage DNA

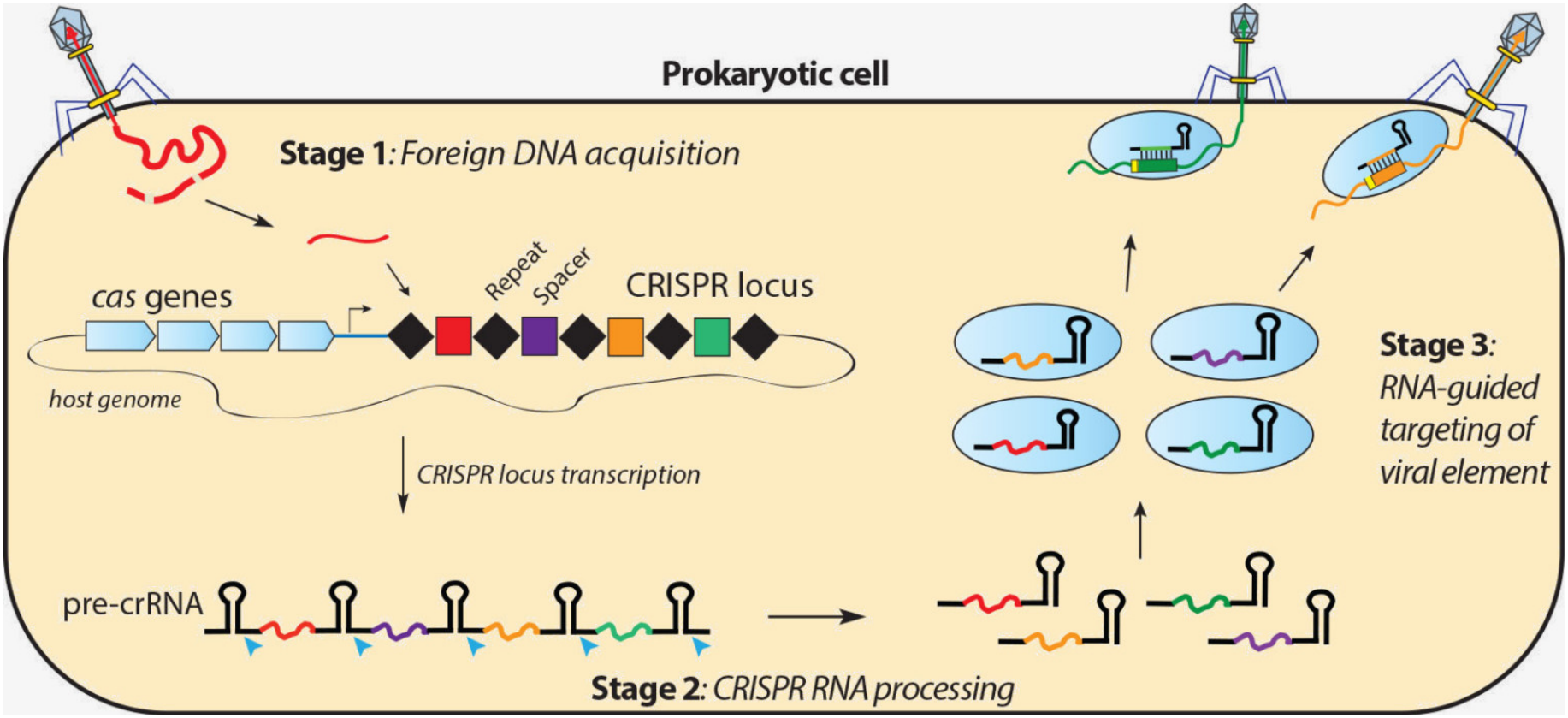
Codes **for a protein** that is a nuclease that cuts DNA but only if it is given a very specific set of signals to do so (otherwise it would potentially damage the bacteria's own DNA). The most common one used in genetic engineering approaches is called Cas9. ; additional Cas1 and Cas2 are responsible for spacer geration

Codes for a very specific piece of RNA (**crRNA or guide RNA**) that will help in the process of ensuring the whole process only cuts bacteriophage DNA

For now lets not worry about the other genes in the Cas locus

Acquisition of immunity

Adaptive immunity

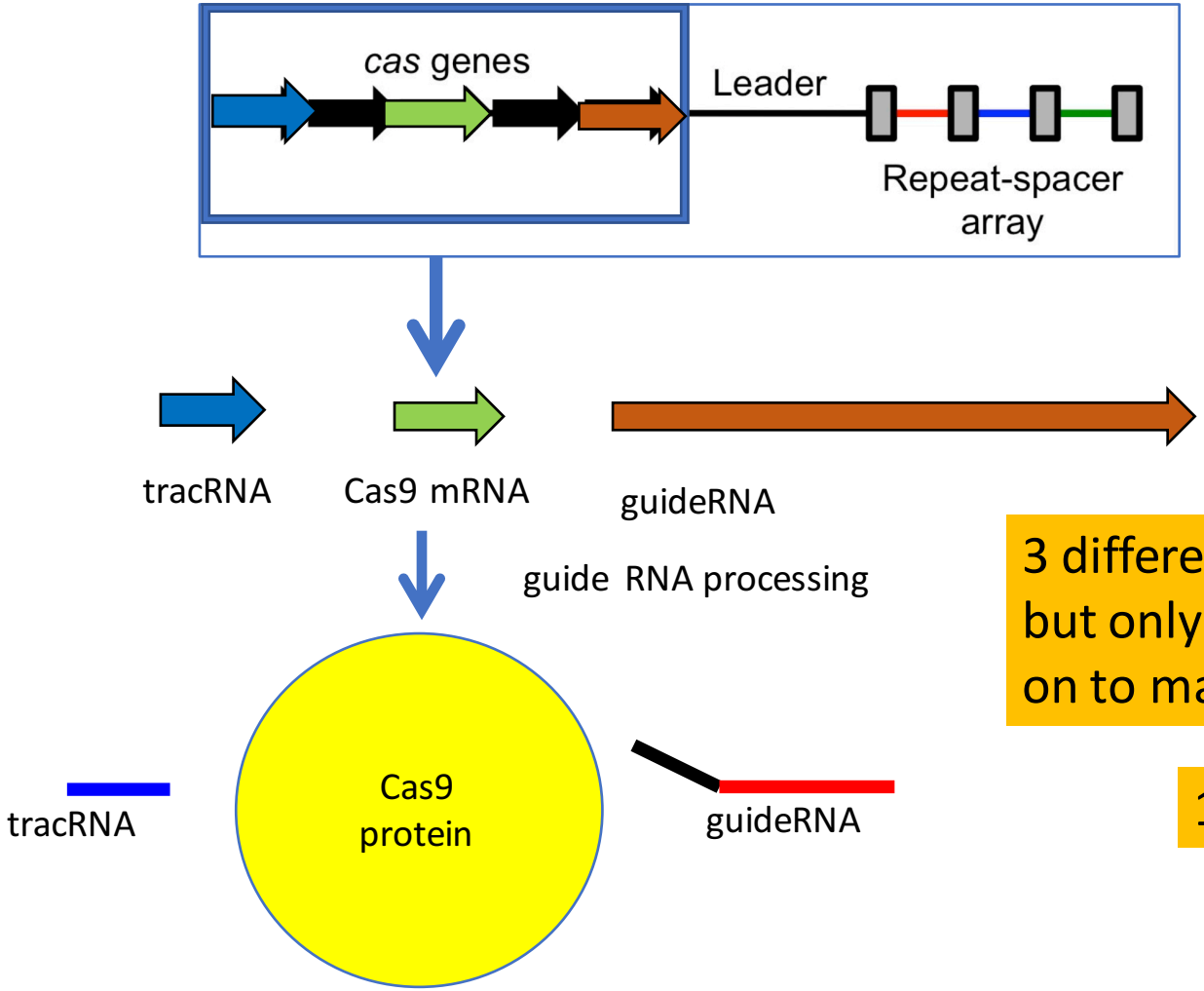


ACQUISITION OF IMMUNITY: Cas1 and Cas2 are responsible for spacer generation: Bioinformatic analysis of regions of phage genomes that were excised as spacers (termed protospacers) revealed that they were not randomly selected but instead were found adjacent to short (3 – 5 bp) DNA sequences termed protospacer adjacent motifs (PAM).

Foreign DNA is inserted into CRISPR locus. IMPORTANT: immunity is passed on to the next generation of cells

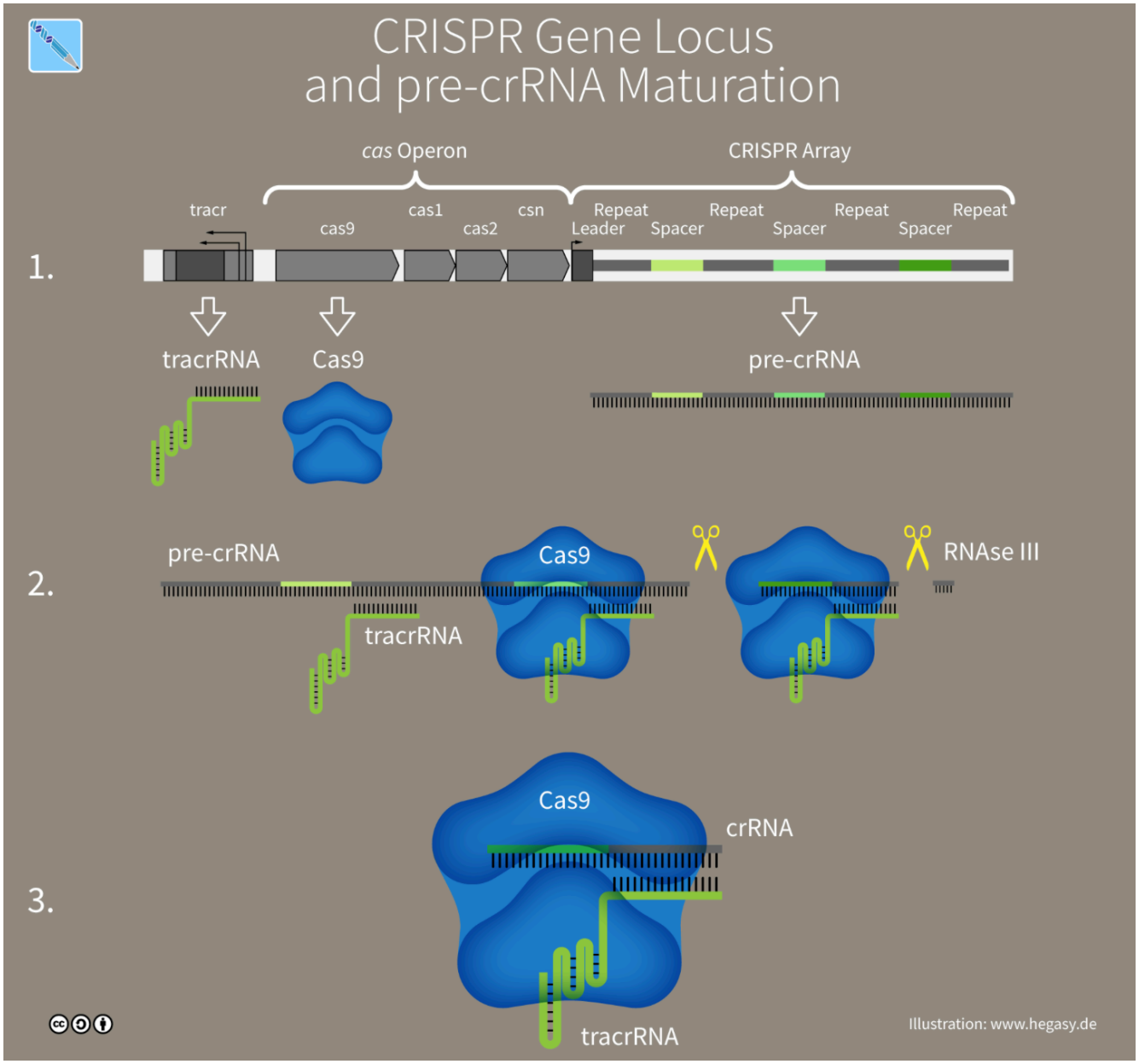
ADAPTIVE IMMUNITY: "daughter cells" are already immune to a bacteriophage that had infected the "mother/father" cell → inherited information

How does this genetic material in CRISPR locus then manage to kill bacteria ?



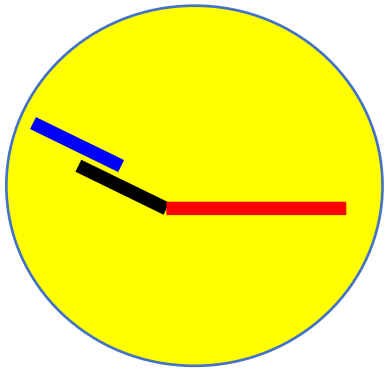
3 different RNAs generated but only one of these goes on to make a protein.

1 protein generated


CRISPR Gene Locus
and pre-crRNA Maturation
cas Operon
CRISPR Array
tracr
cas9
cas1
cas2
csn
Repeat
Leader
Spacer
Repeat
Spacer
Repeat
Spacer
Repeat
1.
tracrRNA
Cas9
pre-crRNA
pre-crRNA
Cas9
RNAse III
tracrRNA
2.
Cas9
crRNA
3.
tracrRNA
Illustration: www.hegasy.de

What is Cas9?

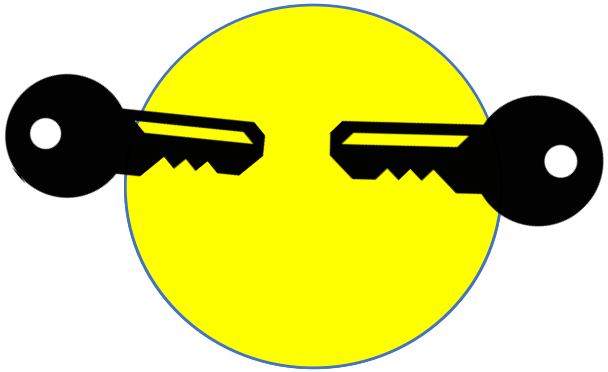
- Cas9 is an endonuclease that can cut double stranded DNA
- Cas 9 is only activated when the tracRNA and the guide RNA are associated with it (i.e it is a nucleoprotein). Imagine this a bit like the fail safe mechanism they use to prevent accidental launch of nuclear missiles where 2 people have to insert keys at exactly the same times
- In fact the tracRNA and the guide RNA have a short overlapping sequence that means they actually have to bind to each other in this complex for this to work properly



Active Cas9

What is Cas9?

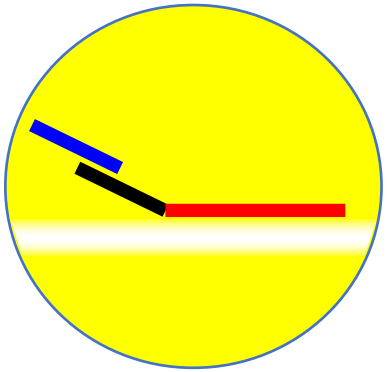
- Cas9 is an endonuclease that can cut double stranded DNA
- Cas 9 is only activated when the tracRNA and the guide RNA are associated with it (i.e it is a nucleoprotein). Imagine this a bit like the fail safe mechanism they use to prevent accidental launch of nuclear missiles where 2 people have to insert keys at exactly the same times
- In fact the tracRNA and the guide RNA have a short overlapping sequence that means they actually have to bind to each other in this complex for this to work properly



Active Cas9

How does Cas9 work?

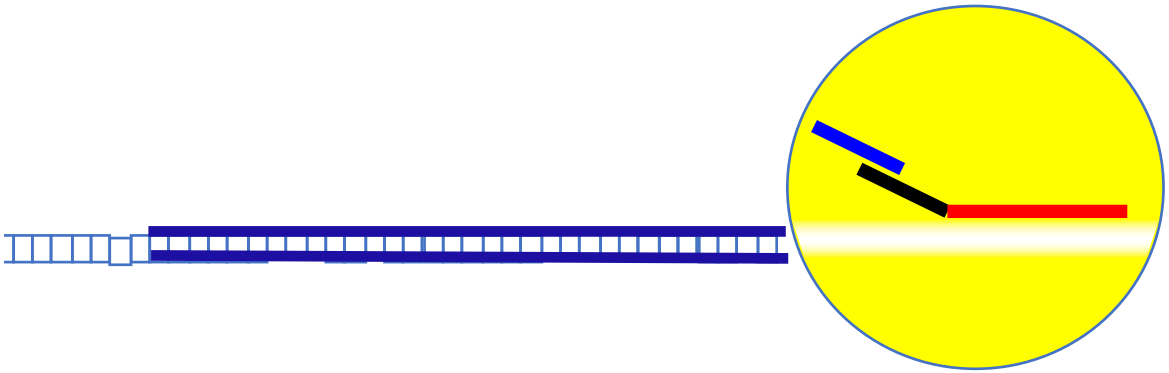
- Cas9 has a channel that DNA can fit into.
- It scans the DNA looking for sequence that match the guide sequence



Active Cas9

How Cas9 works?

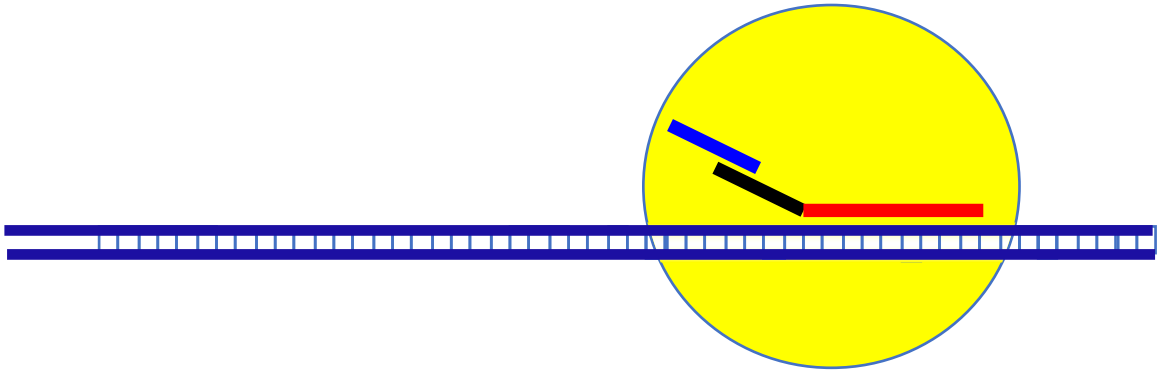
- Cas9 has a channel that DNA can fit into.
- It scans the DNA looking for sequence that match the guide sequence



Active Cas9

How does Cas9 work?

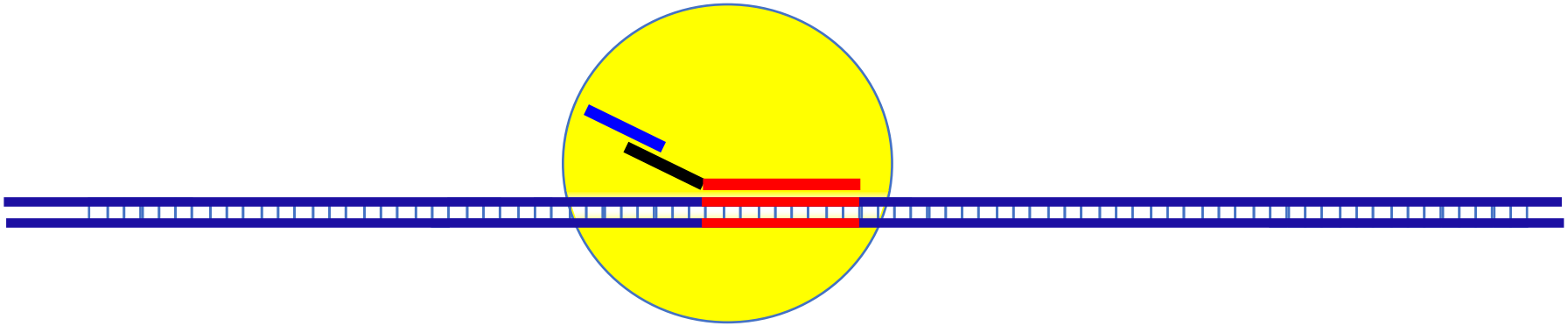
- Cas9 has a channel that DNA can fit into.
- It scans the DNA looking for sequence that match the guide sequence



Active Cas9

How does Cas9 work?

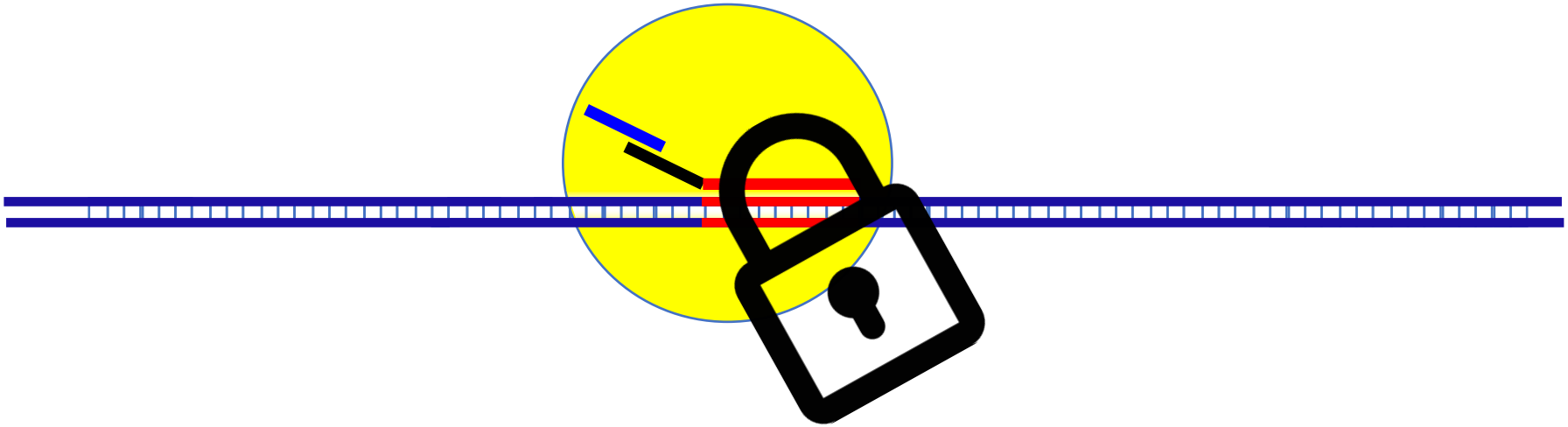
- When a DNA sequence complementary to the guide RNA is found the scanning stops



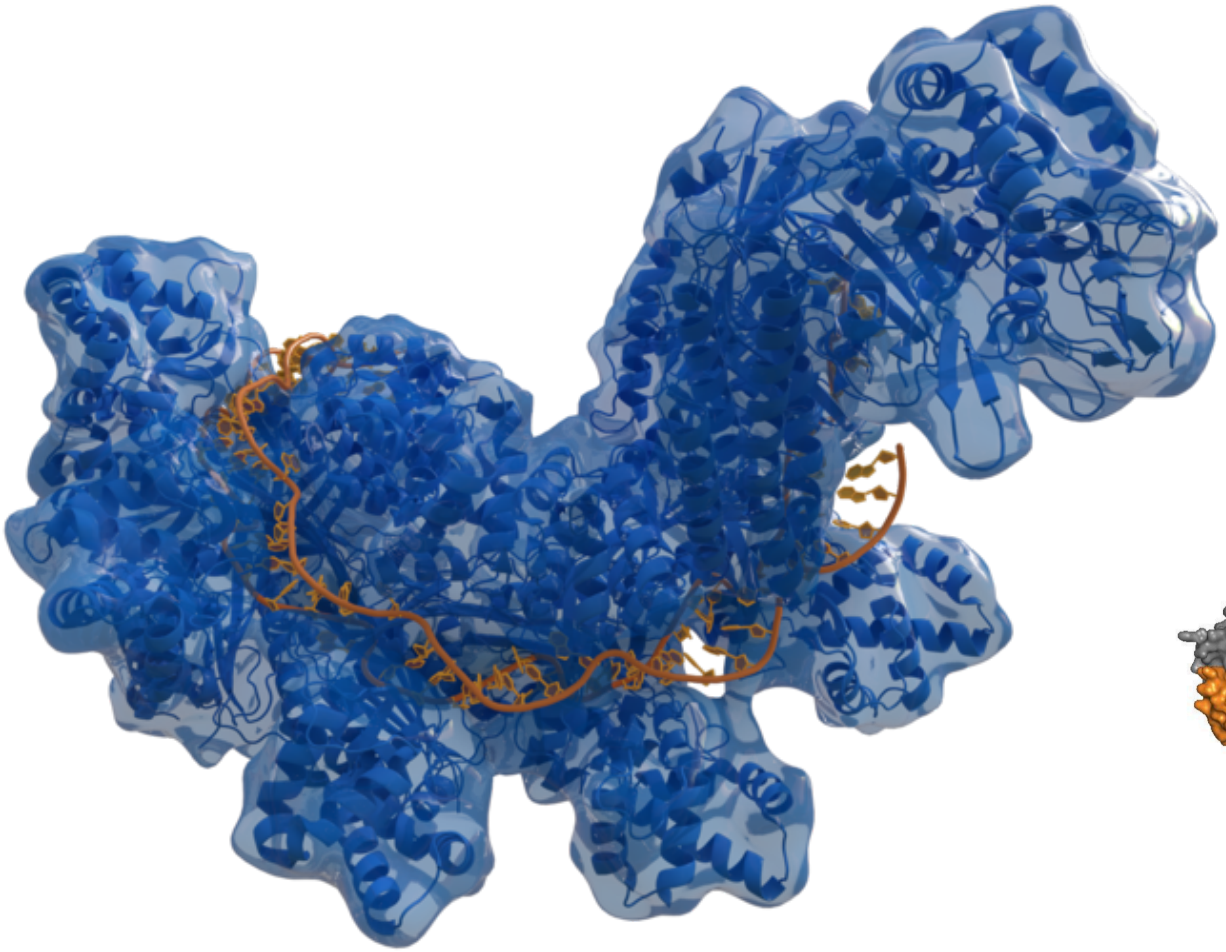
Active Cas9

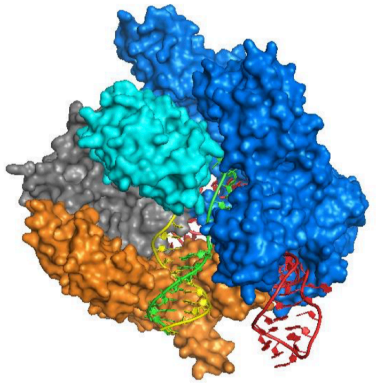
How does Cas9 work?

- When a DNA sequence complementary to the guide RNA is found the scanning stops



Structure of DNA bound to a Cas enzyme





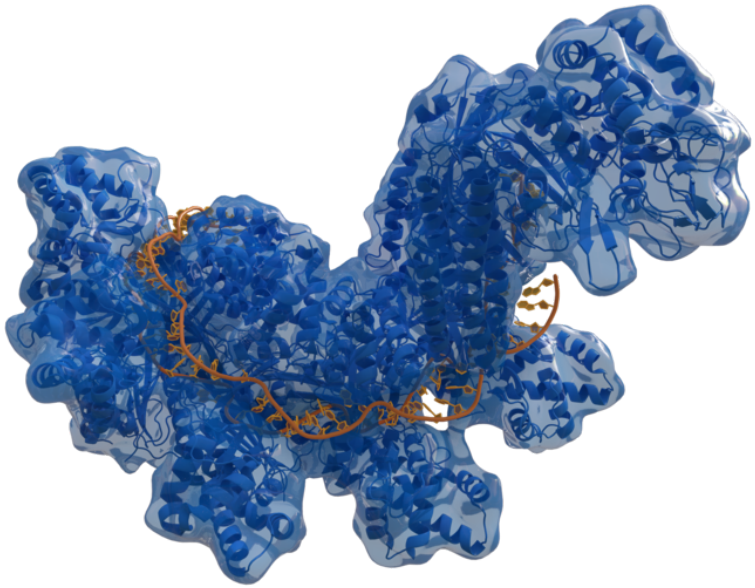
Cas9
(S. pyogenes)

- **Adaptation**
  Recognition of target site
- **Two nuclease domains**
  RuvC (gray) - cleaves non-target DNA strand
  HNH (cyan) - cleaves-target strand of DNA
- **PAM-interacting domain**
  (orange)

Jinek et al. *Science*, 2014

Structure of DNA bound to a Cas enzyme

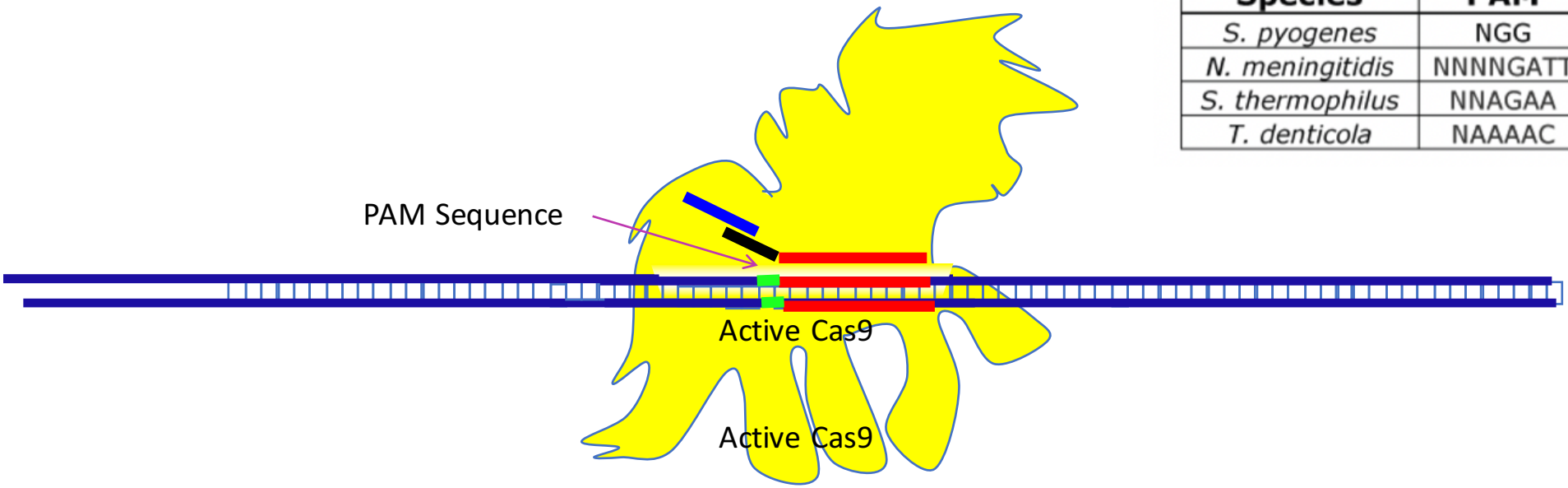
Completely irrelevant aside



How does Cas9 work?

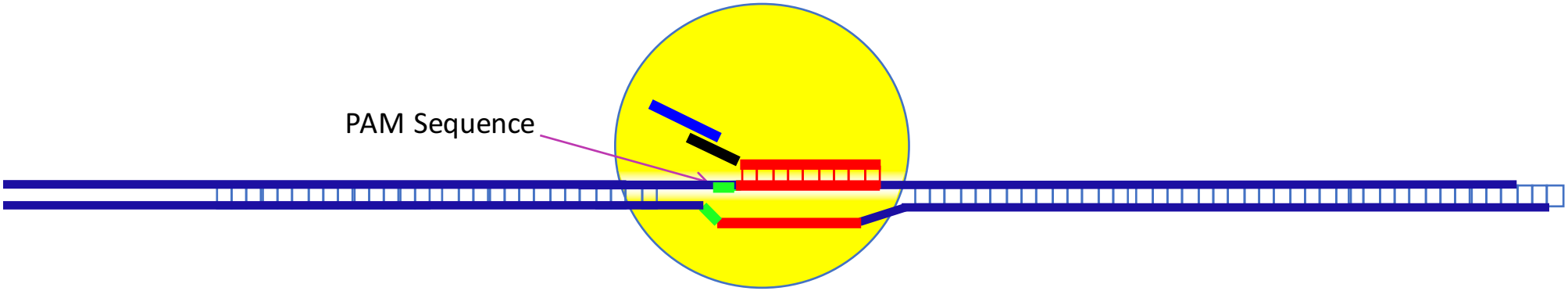
- There is one additional check
- In this control step the target site in the bacteriophage DNA needs to have the PAM sequence (**P**rotospacer **A**djacent **M**otif)
- PAM seqeunces DO NOT APPEAR in the bacterial genome
- PAM seqeunces are requried for Cas9 endonucease activity
- PAM sequences are specific for bacterial strains and protect the Cas locus from beeing cut by Cas9

Species	PAM
S. pyogenes	NGG
N. meningitidis	NNNNGATT
S. thermophilus	NNAGAA
T. denticola	NAAAAC



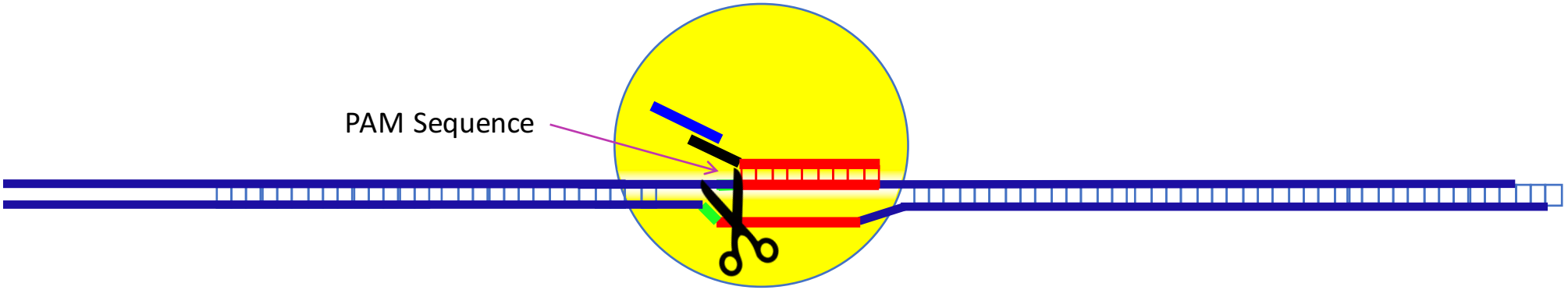
How does Cas9 work?

- Now the RNA binds to the complementary strand of the DNA and opens up the DNA helix



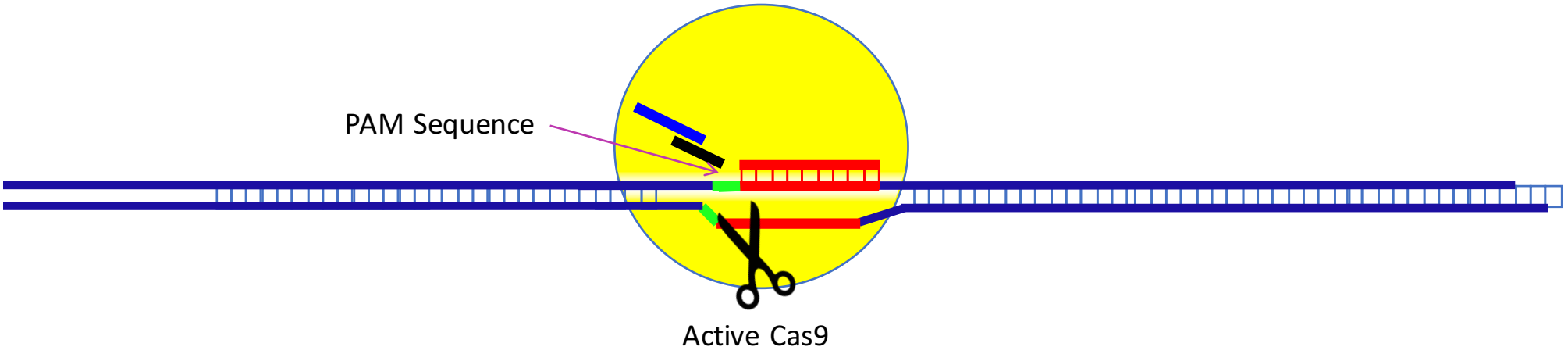
How does Cas9 work?

- Now the bacteriophages DNA gets cut very close to the PAM site



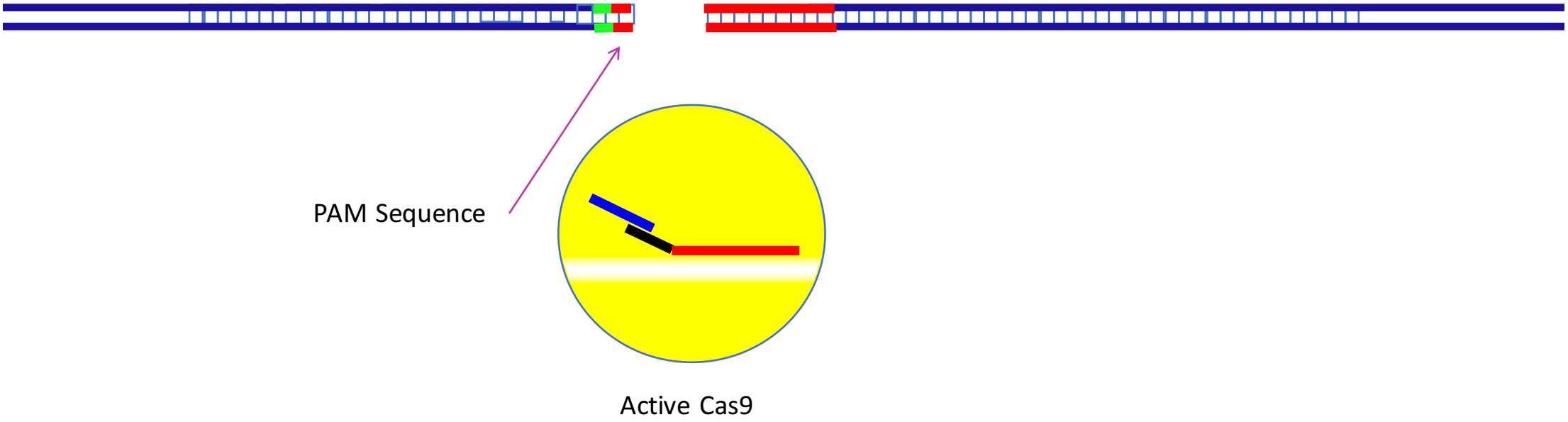
How does Cas9 work?

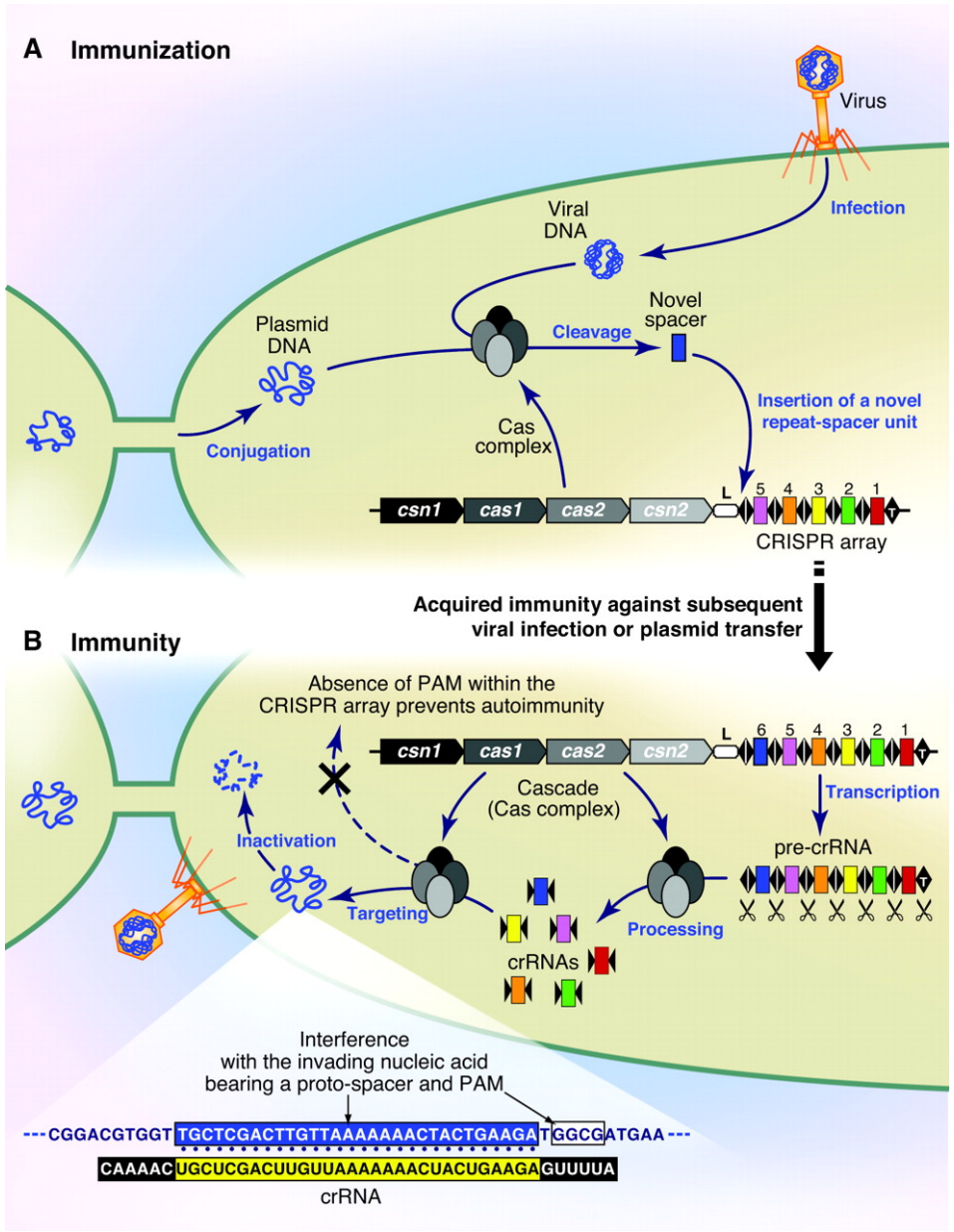
- Now the bacteriophages DNA gets cut very close to the PAM site



How does Cas9 work?

- Now the bacteriophages DNA gets cut very close to the PAM site, it looks like this and the bacteriophage is essentially inactivated





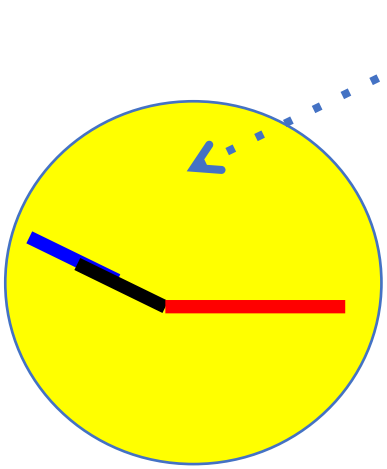
Protospacer adjacent motif (PAM) is a 2-6 base pair DNA sequence immediately following the DNA sequence targeted by the Cas9 nuclease in the CRISPR bacterial adaptive immune system.

PAM is a component of the invading virus or plasmid, but is not a component of the bacterial CRISPR locus. Cas9 will not successfully bind to or cleave the target DNA sequence if it is not followed by the PAM sequence.

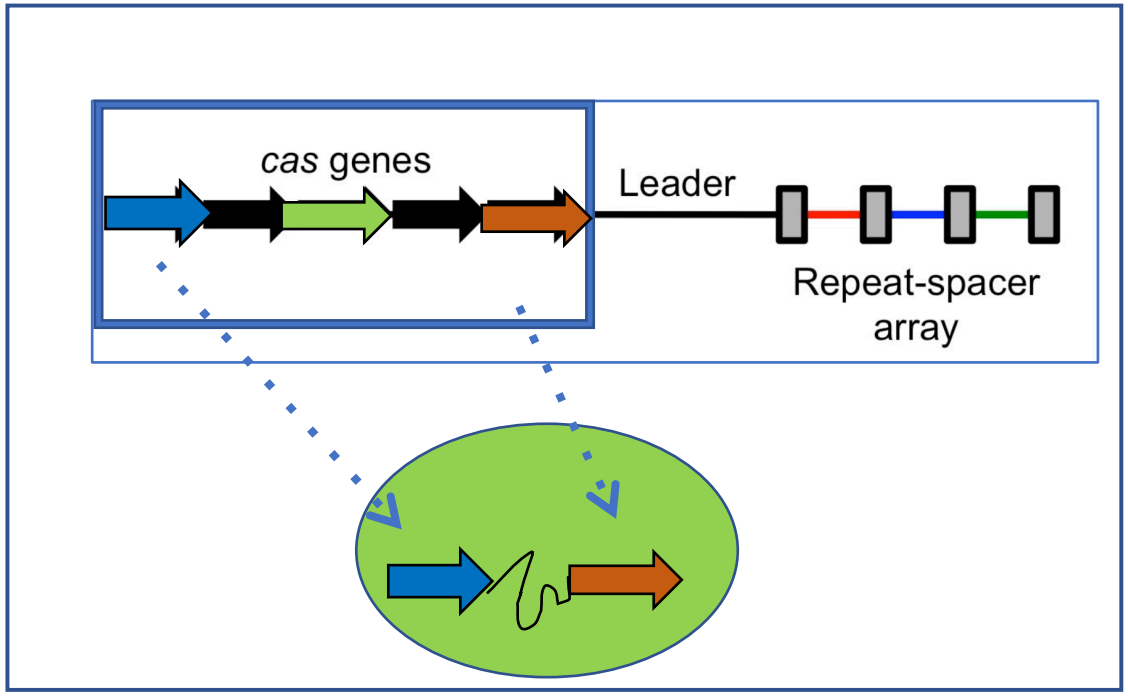
PAM is an essential targeting component (not found in bacterial genome) which distinguishes bacterial self from non-self DNA, thereby preventing the CRISPR locus from being targeted and destroyed by nuclease.

How can we use CRISPR/Cas9 for genetic engineering?

- Jennifer Doudna and Emmanuelle Charpentier re-engineered the Cas9 endonuclease into a more manageable two-component system by fusing the two RNA molecules tracRMA and guide RNA (or crRNA) into a "**SINGLE-GUIDE RNA**" (**sgRNA)** that, when combined with Cas9, could find and cut the DNA target specified by the guide RNA

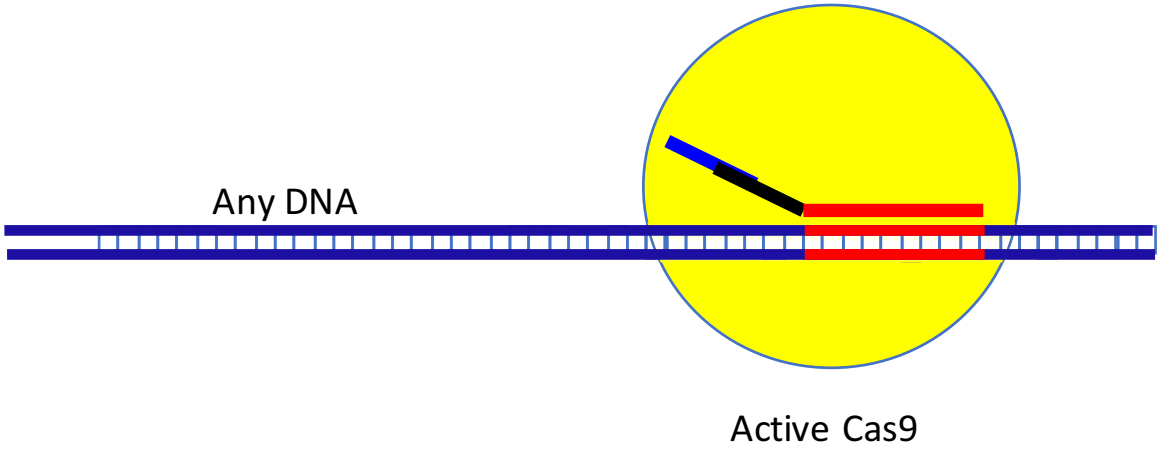


Active Cas9



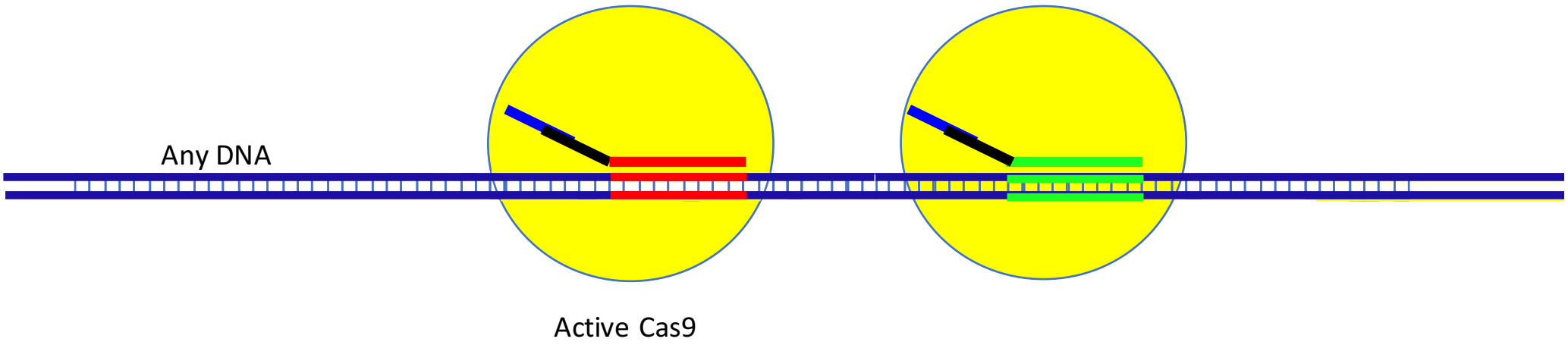
How can we use CRISPR/Cas9 for genetic engineering?

- This means we can artificially make a sgRNA that can be designed to target **any part of the genome** (as long as it has an appropriate PAM sequence nearby)
- **All we have to do is artificially express the Cas9 and the sgRNA together**



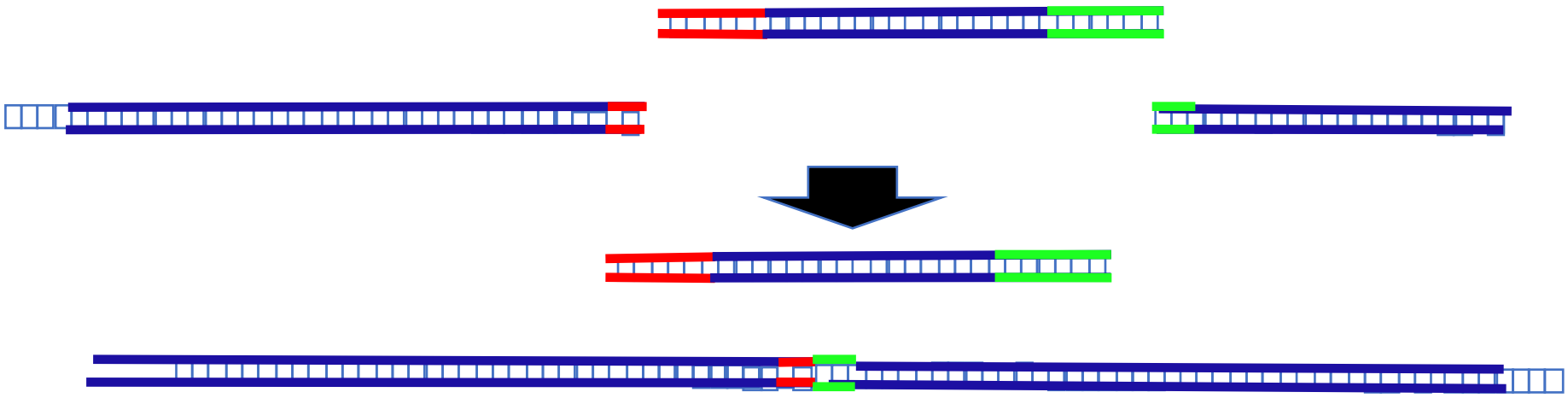
How can we use CRISPR/Cas9 for genetic engineering?

- We can put two different sgRNA into the same protein and cut at 2 places in the genome . → we can cut out large regions of DNA.



How can we use CRISPR/Cas9 for genetic engineering?

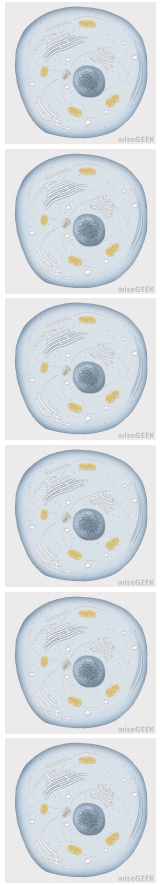
This allows us to selectively “knock out” regions of the genome



Just an example: Knockout of VEGF-a gene

1

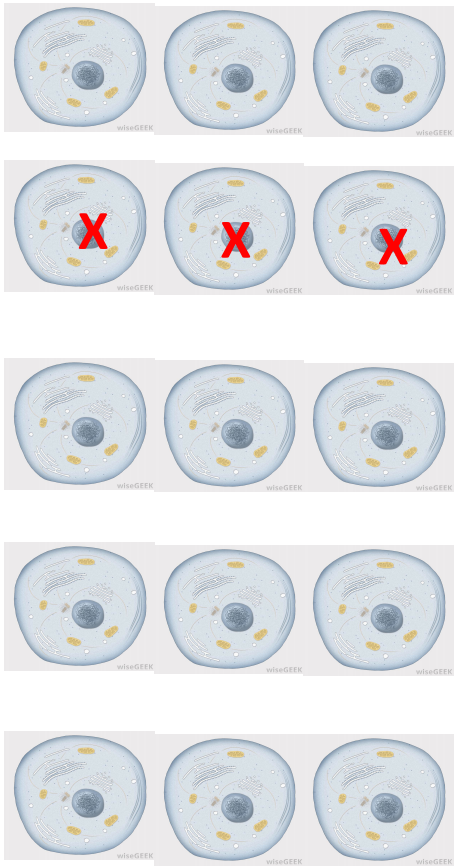
Take lots of cells and add the Cas9 protein plus 2 sgRNA that specifically bind to VEGFA gene



2

Isolate single cells (i.e select clones)

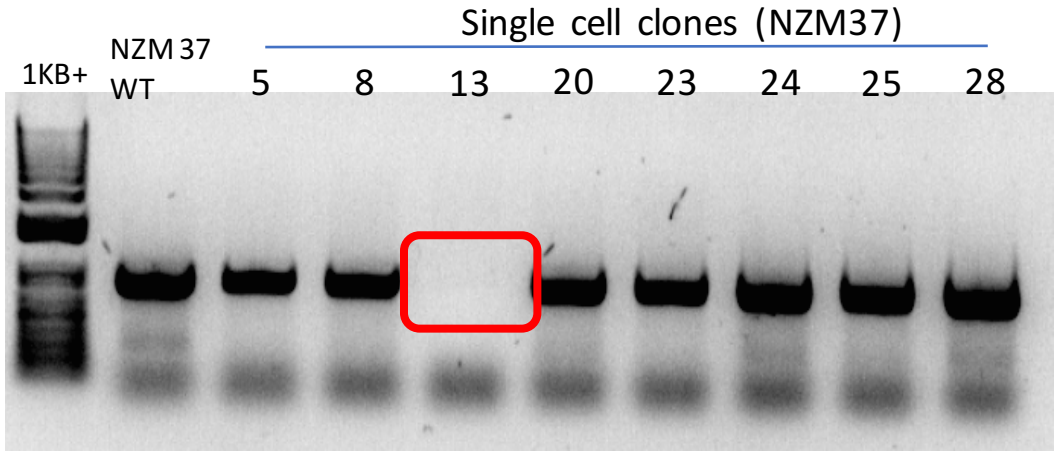
3 Grow cells



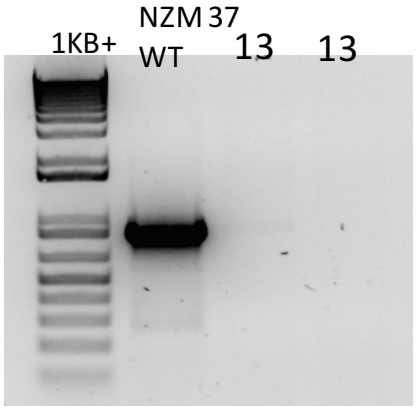
4

Isolate DNA from cells and find cells that have the gene knocked out

Just an example: Knockout of VEGF-a gene



Repeat PCR

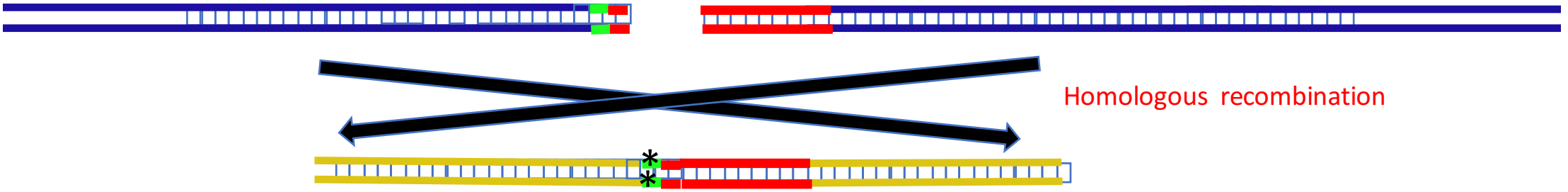


Here is an example of PCR of the VEGFA gene of melanoma cells where we have tried to use CRISPR to "knockout the VEGFA gene (achieved in clone 13)

Using CRISPR/Cas9 to “knockin” bits of DNA

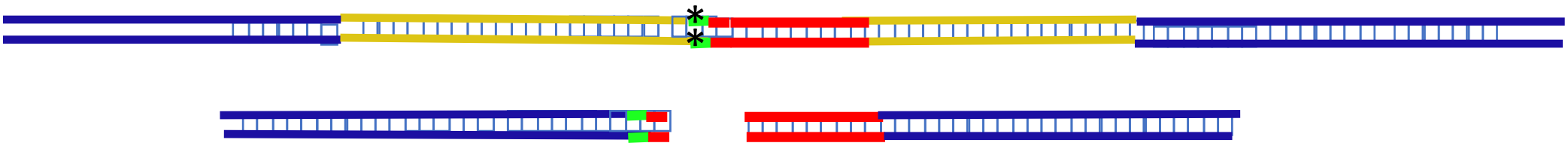
- We can use an artificial piece of DNA that is identical to the cleaved region of DNA (with “corrected” sequence -∗-) ------
- when the cell tries to repair its own chromosomal DNA it will sometimes accidentally incorporate this into its own DNA by homologous recombination!

→ Bring in CRISP-Cas9 components to make a specific cut
→ Bring in a DNA fragement that contains the desired genetic alteration (wt→ mutant; mutant→wt)
→ Strand invasion by cut sequence and HR

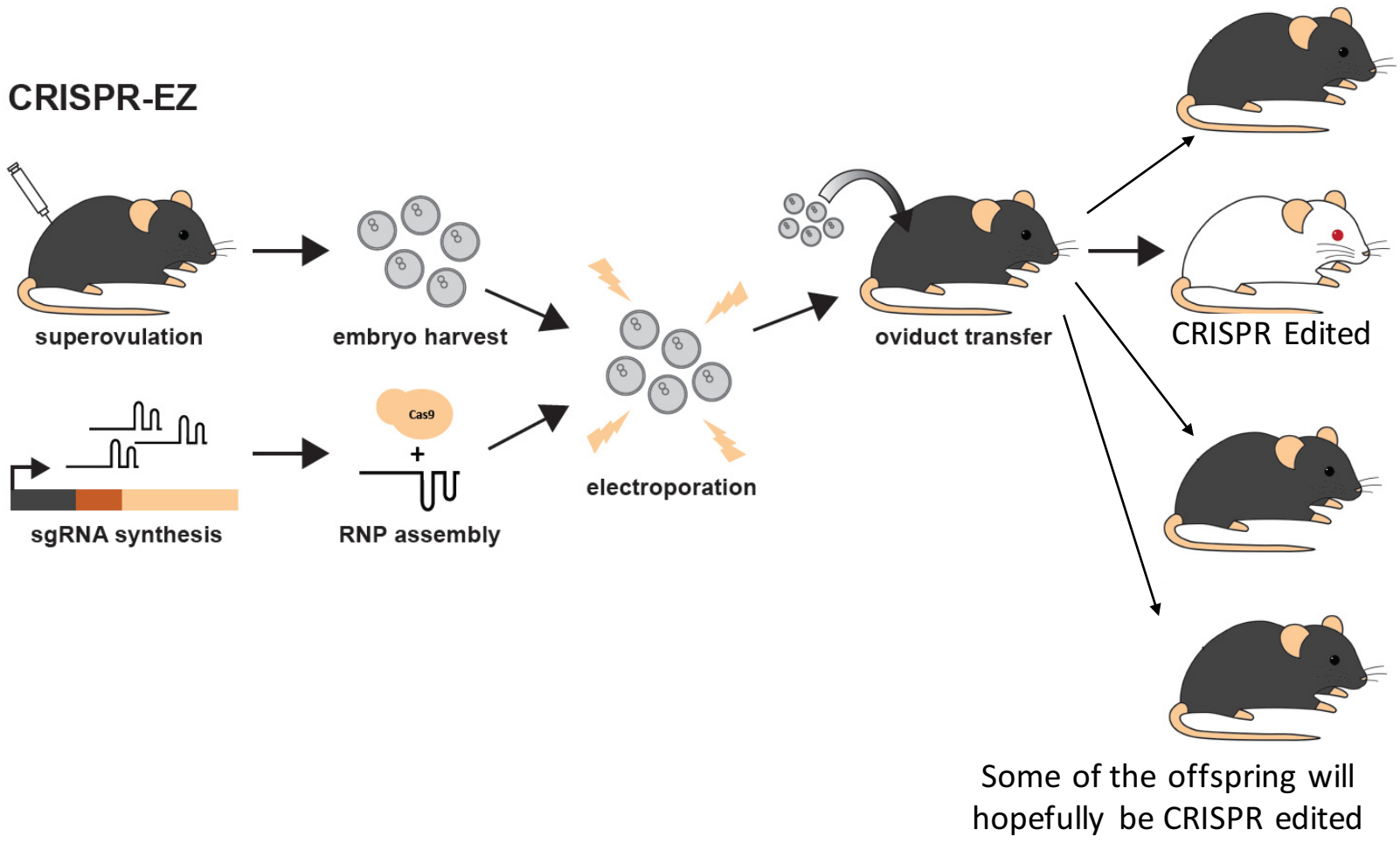


Using CRISPR/Cas9 to “knockin” bits of DNA

- Now the artificially produced piece of DNA is “knocked in” to the genome

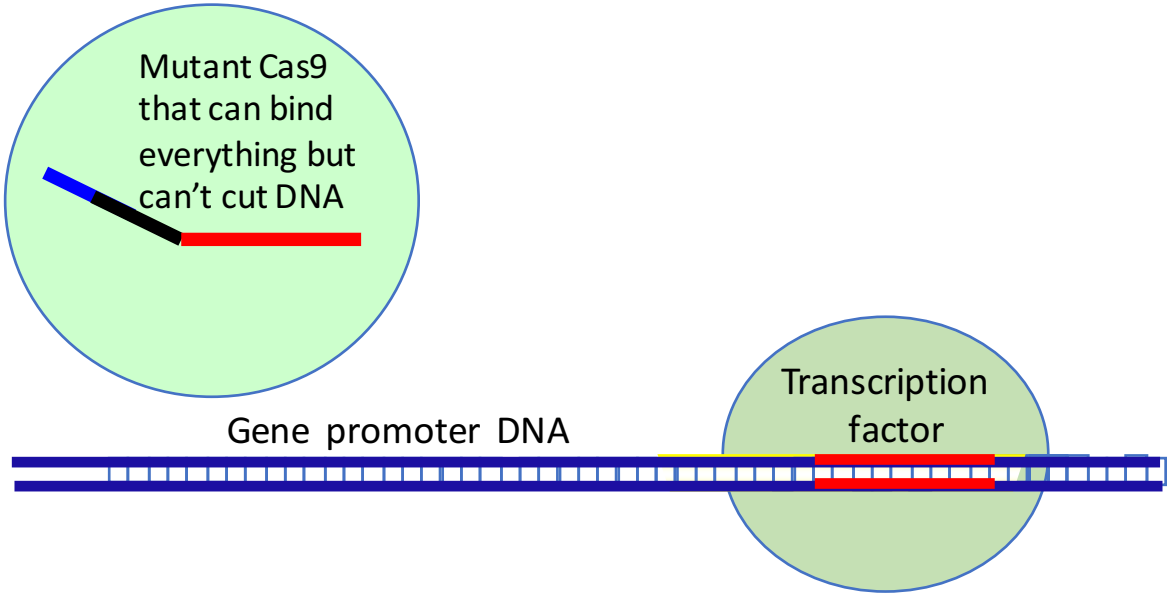


Making mice where genes are knocked out is easier and cheaper



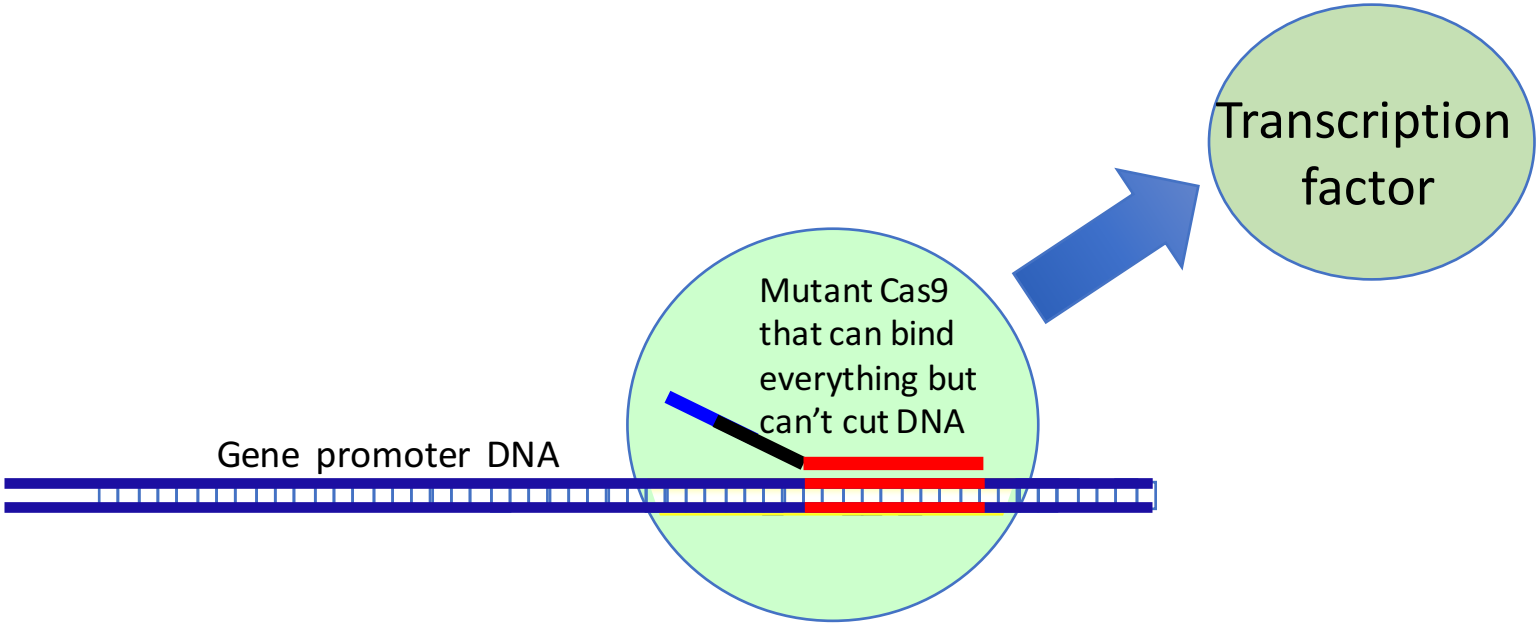
CRISPR/Cas9 to SWITCH ON or OFF GENES

- Uses a mutant **Cas9 that can bind everything but can't cut DNA**
- This means it locks on tightly to the DNA that matches the guide sequence and stops on that region of DNA
- An example of how this can be used is by having a big Cas9 protein sitting at specific transcription factor binding site: we can block the transcription factor from coming into the gene promoter and therefore we switch off the expression of that specific gene controlled by TF in a highly targeted way.
- Negative aspect: keeps DNA in a short RNA:DNA hybrid state



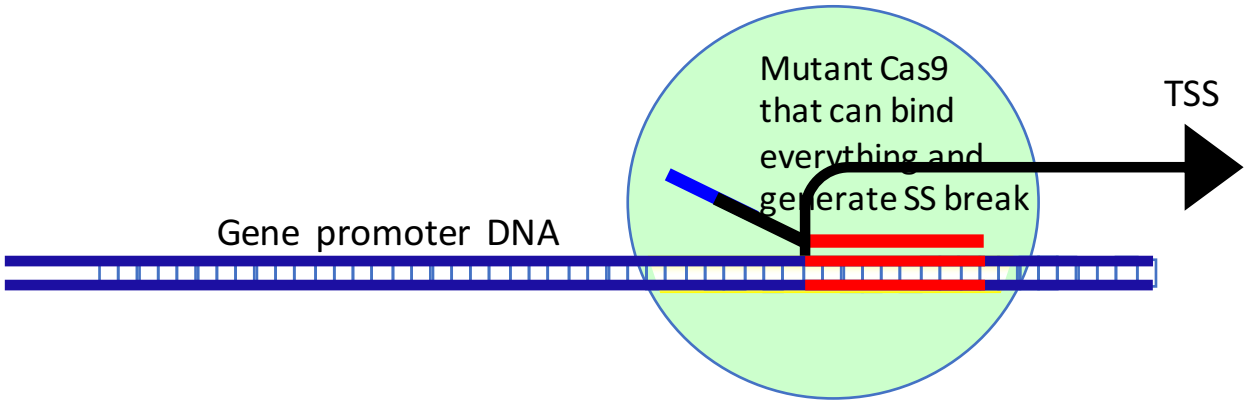
CRISPR/Cas9 to SWITCH ON or OFF GENES

- Uses a mutant **Cas9 that can bind everything but can't cut DNA**
- This means it locks on tightly to the DNA that matches the guide sequence and stops on that region of DNA
- An example of how this can be used is by having a big Cas9 protein sitting at specific transcription factor binding site: we can block the transcription factor from coming into the gene promoter and therefore we switch off the expression of that specific gene controlled by TF in a highly targeted way.



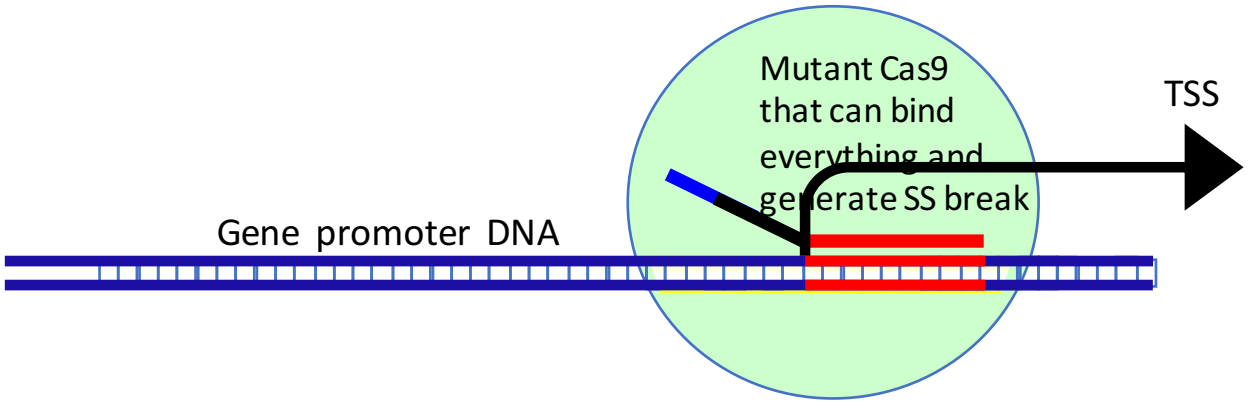
CRISPR/Cas9 to SWITCH ON or OFF GENES

- Alterantively we can use a mutant **Cas9 that can bind everything but generates a single strand cut DNA**
- This means it locks on tightly to the DNA that matches the guide sequence, stops and generate a single strand break that turn-off gene expression



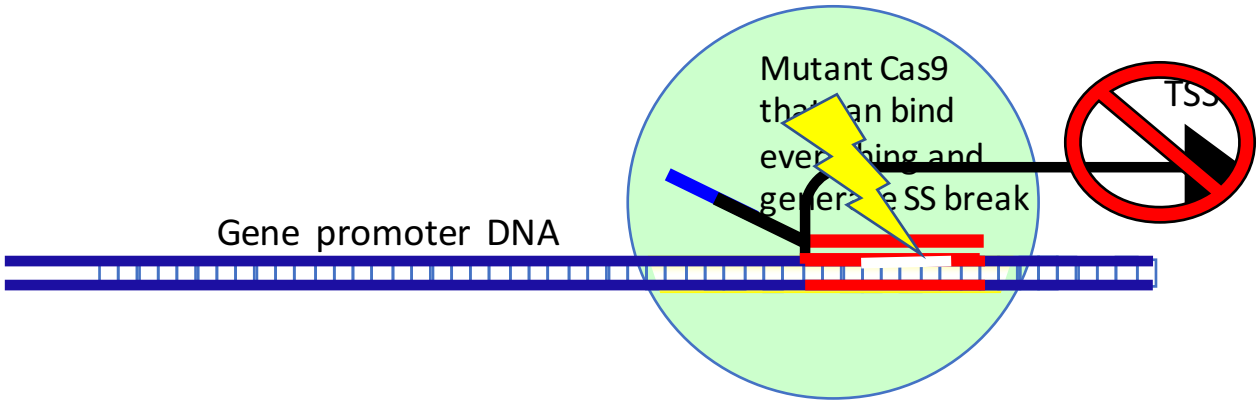
CRISPR/Cas9 to SWITCH ON or OFF GENES

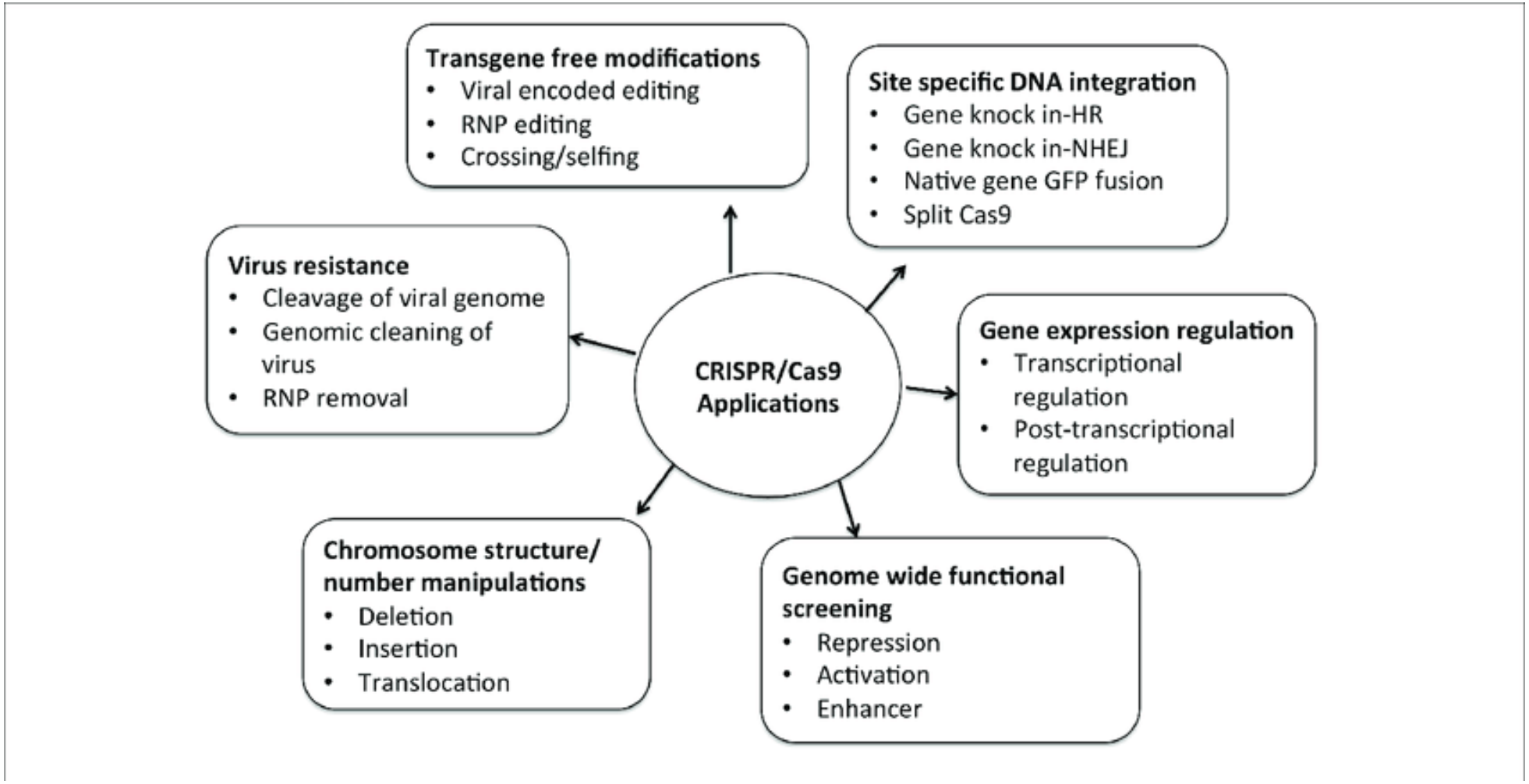
- Alterantively we can use a mutant **Cas9 that can bind everything but generates a single strand cut DNA**
- This means it locks on tightly to the DNA that matches the guide sequence, stops and generate a single strand break that turn-off gene expression



CRISPR/Cas9 to SWITCH ON or OFF GENES

- Alterantively we can use a mutant **Cas9 that can bind everything but generates a single strand cut DNA**
- This means it locks on tightly to the DNA that matches the guide sequence, stops and generate a single strand break that turn-off gene expression




Transgene free modifications
• Viral encoded editing
• RNP editing
• Crossing/selfing
Site specific DNA integration
• Gene knock in-HR
• Gene knock in-NHEJ
• Native gene GFP fusion
• Split Cas9
Virus resistance
• Cleavage of viral genome
• Genomic cleaning of virus
• RNP removal
CRISPR/Cas9 Applications
Gene expression regulation
• Transcriptional regulation
• Post-transcriptional regulation
Chromosome structure/ number manipulations
• Deletion
• Insertion
• Translocation
Genome wide functional screening
• Repression
• Activation
• Enhancer

Using CRISPR a weapon to wipe out mosquitos

A gene drive is a genetic engineering technology that can propagate a particular suite of genes throughout a population. Gene drives can arise through a variety of mechanisms. They have been proposed to provide an effective means of genetically modifying specific populations and entire species.