Genetic Engineering

The CRISPR tool kit for genome editing and beyond

Mazhar Adli
Nature Communications volume 9, Article number: 1911 (2018)

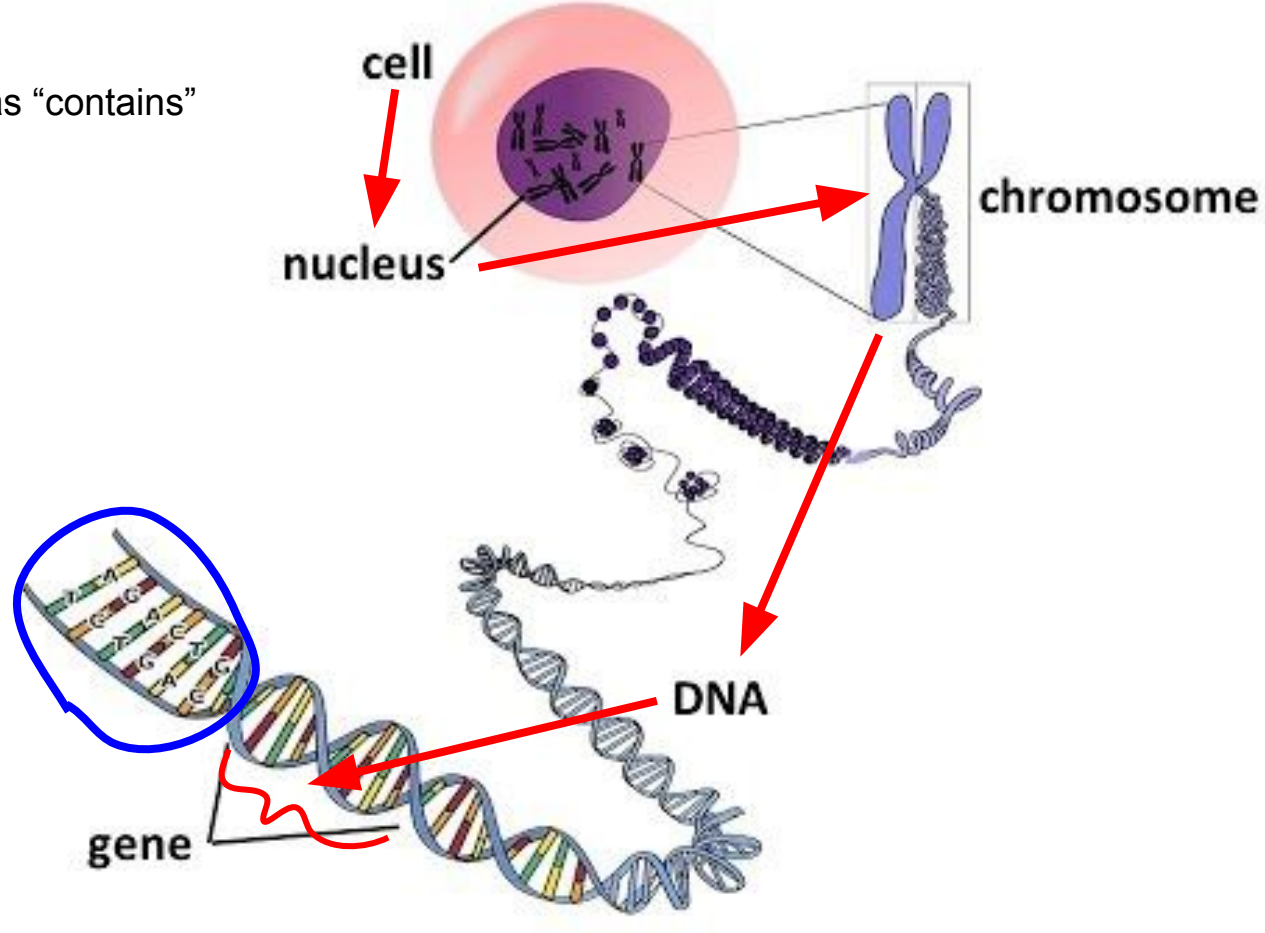
Reviewed by: Eric Wang
University of Virginia
https://qdata.github.io/deep2Read/

History of Genome-editing/Genetic engineering

- Genomes of eukaryotic organisms are composed of billions of DNA bases.
  - **Eukaryotic organism/Eukaryote**: an organism consisting of a cell or cells in which the genetic material is DNA in the form of chromosomes. Eukaryotes include most living organisms such as human beings, insects, even yeast.
  - **Genome**: the *complete* set of genes or genetic material present in a cell or organism
  - **Chromosome**: a threadlike structure found in cells, carrying genetic information in the form of genes.
  - **Gene**: a part of the DNA, in unit of bases, that codes for a molecule.

Adli, M.The CRISPR tool kit for genome editing and beyond Nature Communications. (2018)9:1911.

Red arrows can be interpreted as “contains”



History of Genome-editing/Genetic engineering cont.

- Changing these DNA bases at precisely predetermined locations and introducing desired changes into the genomes are called **Genome-editing/Genetic engineering**.
- Two levels of Genetic engineering
  1. Gene Level: Preparing desired gene fragments.
  2. Cell Level: Incorporating the engineered gene into cells genome.

Adli, M.The CRISPR tool kit for genome editing and beyond Nature Communications. (2018)9:1911.

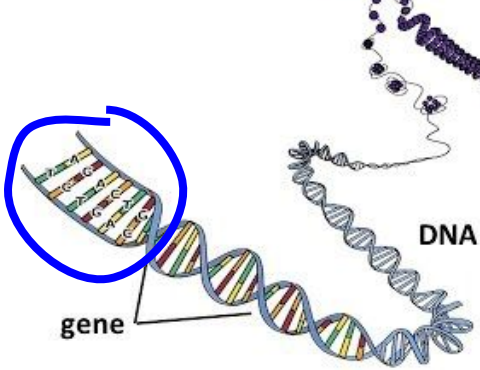
Gene Level: Preparing desired gene fragments

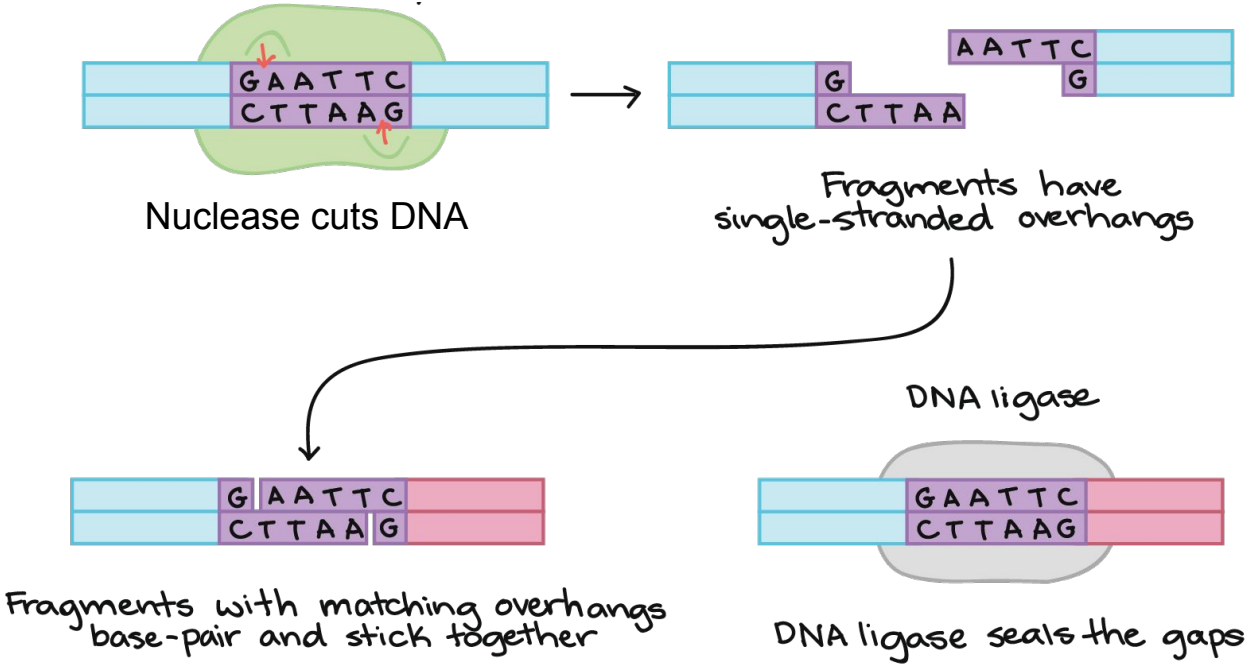
- “Discovery of restriction enzymes in the late 1970s was a turning point that fueled the era of recombinant DNA technology.”
  - **Restriction enzyme**: an enzyme(protein) produced by bacteria, having the property of cleaving DNA molecules at or near a specific sequence of bases. Later in this presentation referred to as **nuclease**.
  - **Recombinant DNA**: DNA that has been formed artificially by combining constituents from different organisms. The *beginning* of genetic engineering and genome-editing.

Adli, M.The CRISPR tool kit for genome editing and beyond Nature Communications. (2018)9:1911.

History of Genetic Engineering -> Gene Level cont.

Nuclease Mechanism: cut DNA at specific sequence

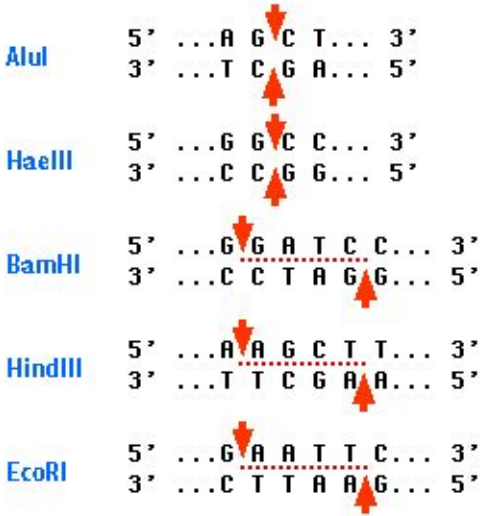




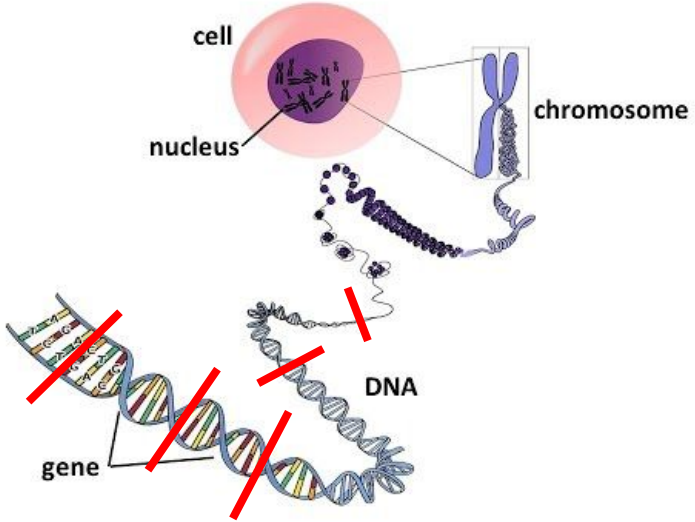
History of Genetic Engineering -> Gene Level -> Nuclease cont.

Each (restriction enzyme) nuclease recognizes *specific* **short** “cut-site sequences”, which limits the *variety* and **scale** of genetic engineering.

Specific sequences

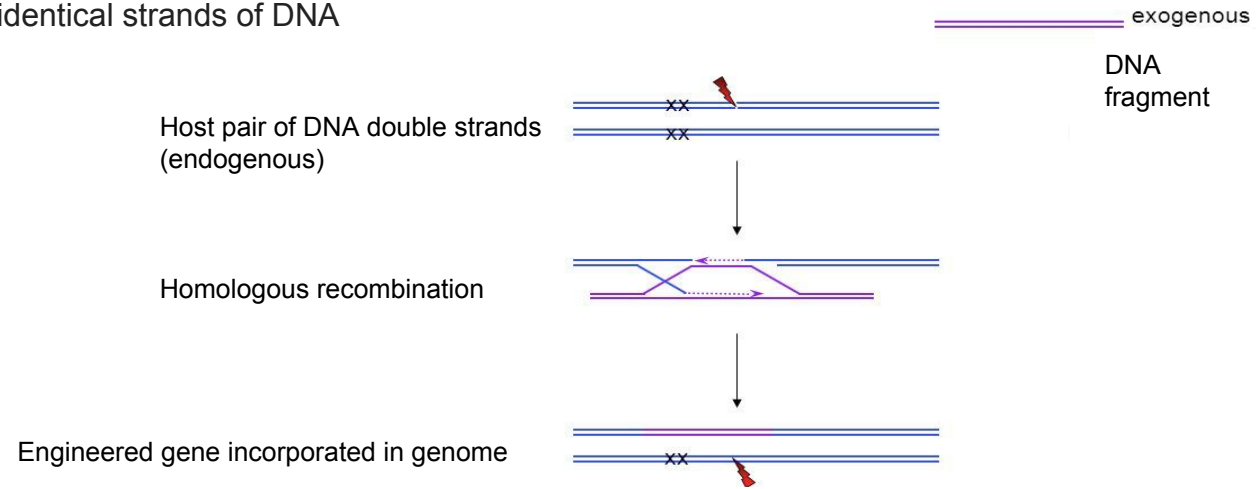


Can’t perform large-scale genome editing-”off target effect”



Cell Level: Incorporating the engineered DNA fragments (gene) into cells

- "Mammalian cells can incorporate an exogenous copy of DNA into their own genome through a process called homologous recombination."
  - Exogenous: foreign DNA not already in the target genome.
  - Homologous recombination: nucleotide sequences are exchanged between two similar or identical strands of DNA



Adli, M.The CRISPR tool kit for genome editing and beyond Nature Communications. (2018)9:1911.

Development of targeted nucleases for genome editing

- “One of the initial breakthroughs came from the realization that the introduction of a **double- strand break (DSB)** at a target site results in a several orders of magnitude increase in the frequency of targeted gene integration.”
- “meganucleases (the nucleases that recognize long stretches of 14–40 bases DNA) increased the genome-editing efficiency”
    - **Double-strand break(DSB)**: the breakage of the double helix of DNA that will be repaired, which gives opportunity for exogenous DNA fragments to be fused at the breakage during repair.

Adli, M.The CRISPR tool kit for genome editing and beyond Nature Communications. (2018)9:1911.

Development of targeted nucleases->Drawbacks

- 1. Hundreds of naturally found meganucleases, each of them has a unique recognition sequence; the *probability* of finding a meganuclease that targets a desired locus was *low*.

- 2. The majority of induced DSBs are repaired through DNA repair mechanism. Thereby, not only may the exogenously introduced DNA template *not incorporate* at the DSBs, but also the repair mechanism may *randomly insert or delete* DNA pieces at the break sites.

Adli, M.The CRISPR tool kit for genome editing and beyond Nature Communications. (2018)9:1911.

Development of targeted nucleases->zinc finger proteins

- “Zinc fingers are zinc ion-regulated small protein motifs that bind to DNA in a sequence-specific manner. Each zinc finger module recognizes a *3-bp* DNA sequence. Therefore, unlike meganucleases, multiple zinc finger modules could be assembled into a larger complex to achieve *higher DNA binding specificity*.”

- “Since each zinc finger recognized a 3-bp DNA code, combinatorial assembly of 6–7 zinc fingers out of the unique 64-finger pool (4^3 combinations) could uniquely target any 18–21 bp genomic sequence.”

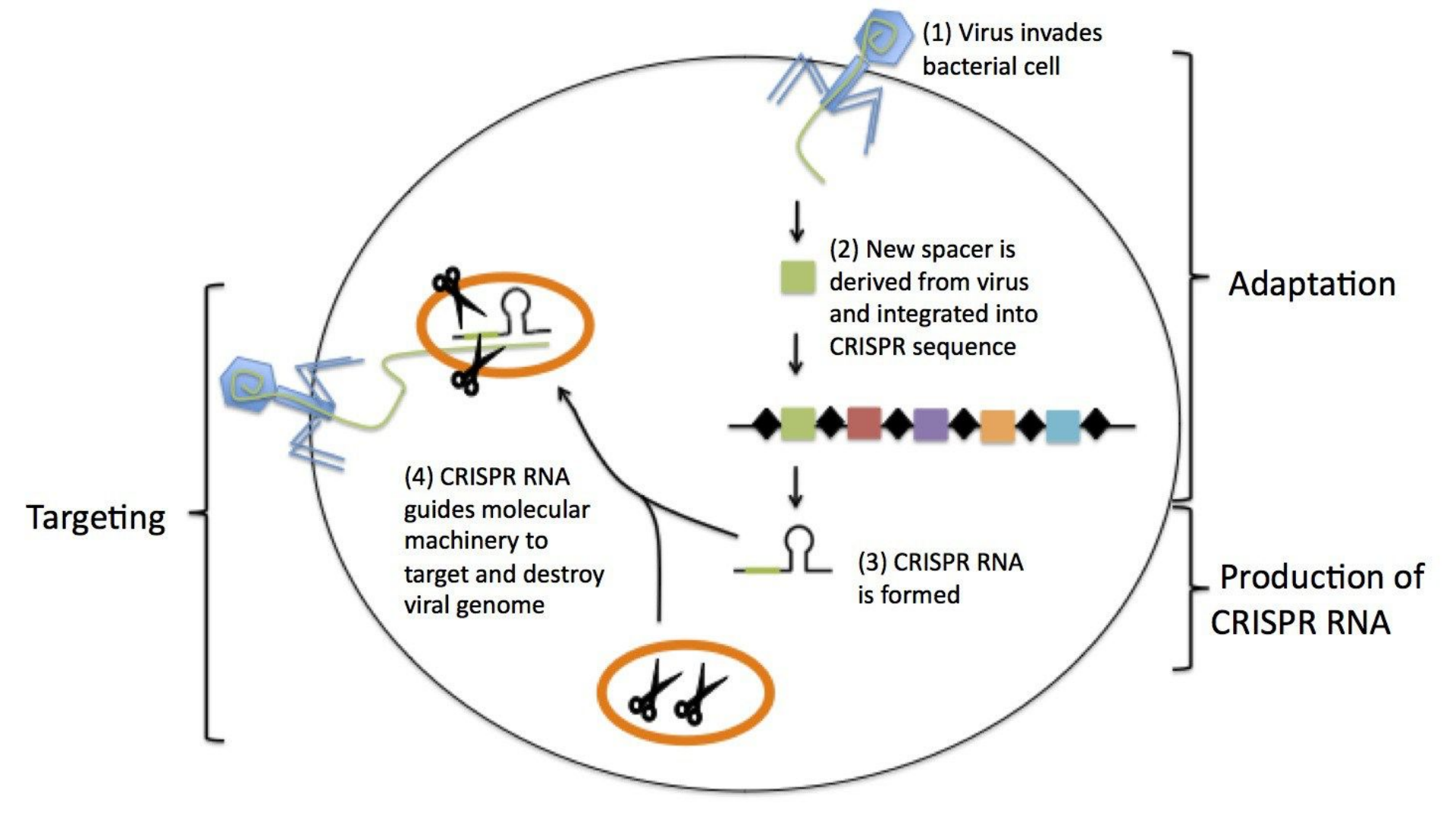
- Transcription activator-like effector (TALE) proteins can specifically recognize one single base instead of three bases.

Adli, M.The CRISPR tool kit for genome editing and beyond Nature Communications. (2018)9:1911.

The rise of CRISPR as the genome-editing technology

- CRISPR stands for *clustered regularly interspaced short palindromic repeat* DNA sequences
- Key features of CRISPR
  - 1. The CRISPR sequences are present in many organisms.
  - 2. The CRISPR elements are adjacent to multiple well-conserved genes called CRISPR-associated (Cas) genes.
  - 3. The non-repeating spacer DNA sequences were recognized to belong to viruses and other mobile genetic elements
  - => CRISPR sequences obtained from outside and incorporated in bacterial genome as part of the immune system against viruses. Next time the bacteria face a viral infection, CRISPR will essentially guide the CRISPR-associated(Cas), which functions as a nuclease(restriction enzyme), to cut the viral DNA and prevent viral infection.

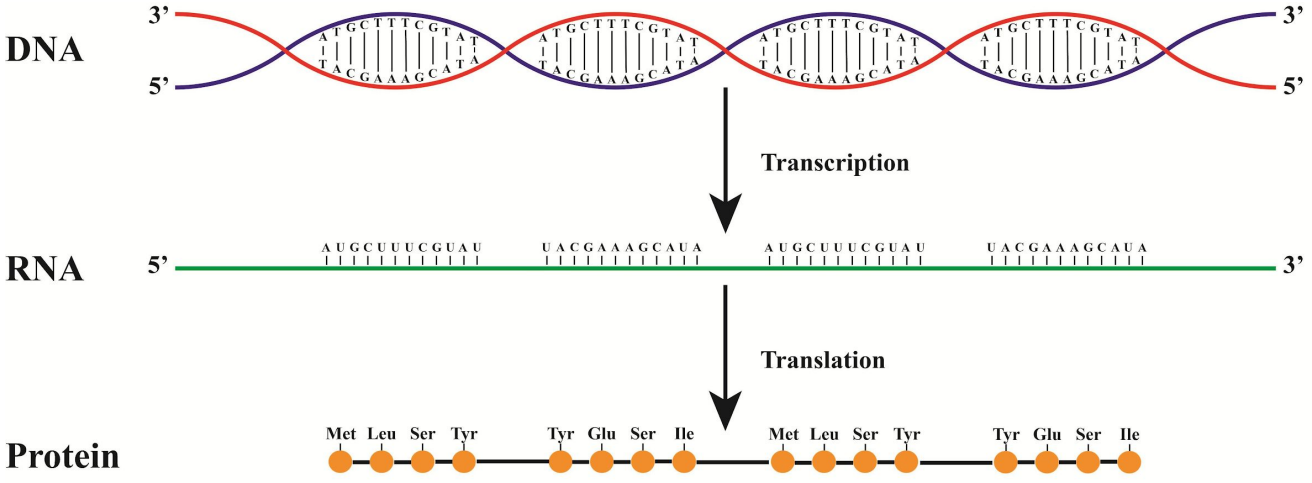
Adli, M.The CRISPR tool kit for genome editing and beyond Nature Communications. (2018)9:1911.


(1) Virus invades bacterial cell
(2) New spacer is derived from virus and integrated into CRISPR sequence
Adaptation
(3) CRISPR RNA is formed
Production of CRISPR RNA
(4) CRISPR RNA guides molecular machinery to target and destroy viral genome
Targeting

Adli, M.The CRISPR tool kit for genome editing and beyond Nature Communications. (2018)9:1911.

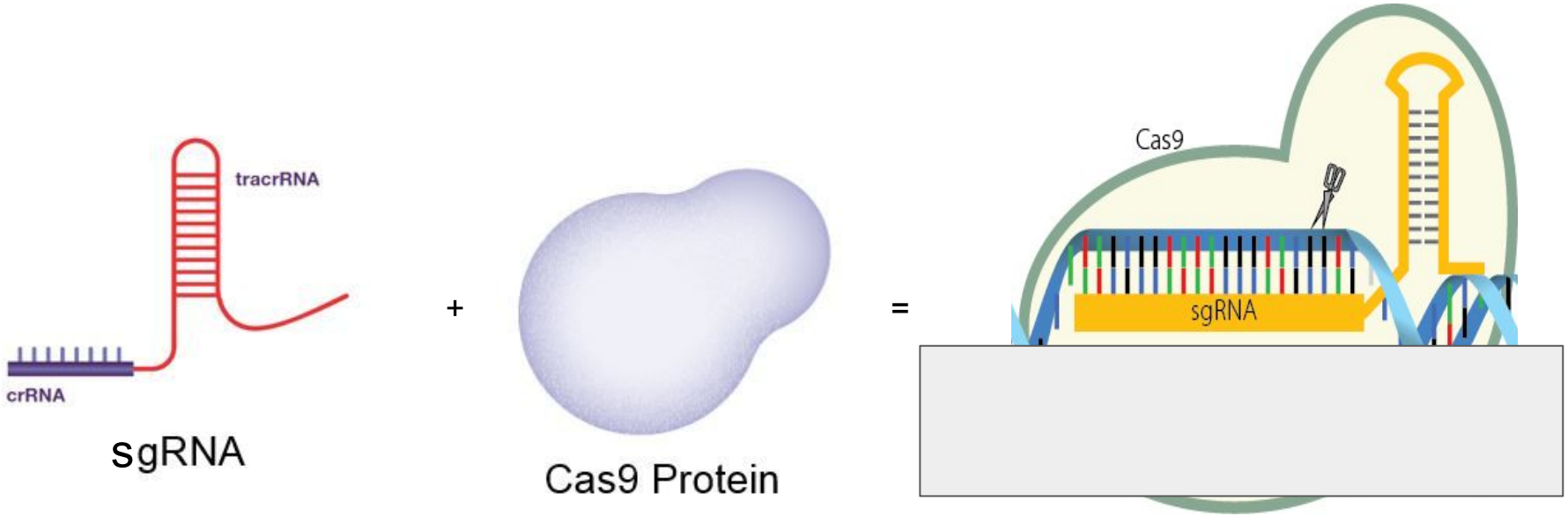
The rise of CRISPR as the genome-editing technology->CRISPR system parts 1

- sgRNA: a fusion RNA artificially created by combining two RNAs coded by the spacer sequences of CRISPR. sgRNA serves as a sequence-specific guide for Cas9 to cut target DNA at sites that complement the sgRNA.
  - RNA: act as a messenger carrying instructions from DNA for controlling the synthesis of proteins. DNA->RNA->Protein.



The rise of CRISPR as the genome-editing technology->CRISPR system parts 2

- Cas/Cas9: enzyme coded by the CRISPR-associated genes, with a function of cutting the target DNA like a restriction enzyme/nuclease.



Adli, M.The CRISPR tool kit for genome editing and beyond Nature Communications. (2018)9:1911.

The rise of CRISPR as the genome-editing technology->CRISPR system parts 3

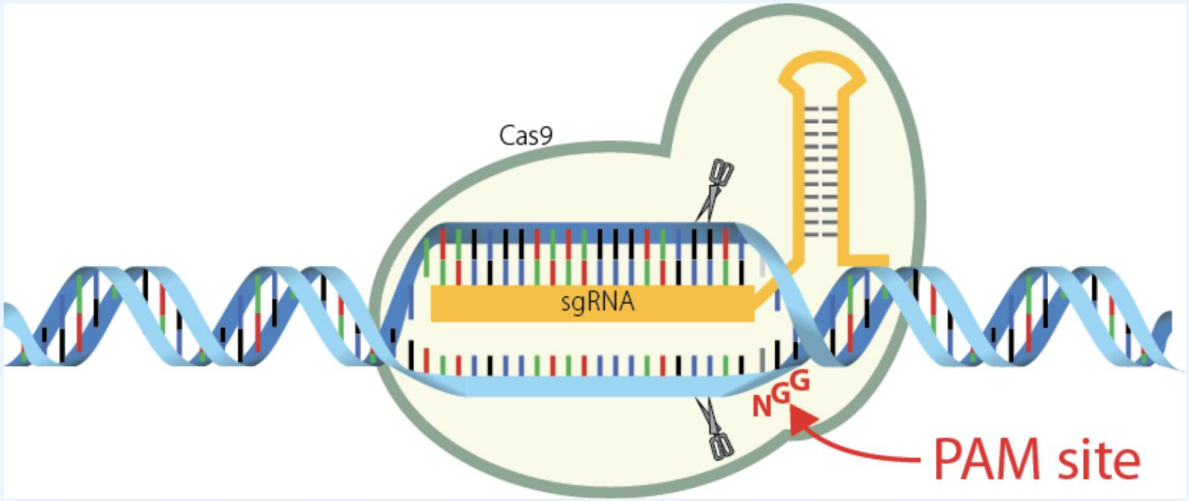
- TAM(protospacer-adjacent motifs): DNA sequence recognized by different Cas enzymes from varying organisms; specific to each Cas. TAM is on the target DNA close to the sequence that's complement to the sgRNA, to signal Cas9 attachment.
  - Example:
  - SaCas9 requires a 5′-NNGRRT-3′ PAM sequence
  - CjCas9 requires a 5′-NNNNACAC-3′ PAM sequence

Adli, M.The CRISPR tool kit for genome editing and beyond Nature Communications. (2018)9:1911.

IDT. What is TAM?

The rise of CRISPR as the genome-editing technology->CRISPR system parts 3

CRISPR-Cas9 mechanisms recognize DNA targets that are complementary to a short CRISPR sgRNA sequence. The part of the sgRNA sequence that is complementary to the target sequence is known as a protospacer. In order for Cas9 to function it also requires a specific protospacer adjacent motif (PAM) that varies depending on the bacterial species of the Cas9 gene. The most commonly used Cas9 nuclease, derived from *S. pyogenes*, recognizes a PAM sequence of NGG that is found directly downstream of the target sequence in the genomic DNA, on the non-target strand.



Recognition of the PAM by the Cas9 nuclease is thought to destabilize the adjacent sequence, allowing interrogation of the sequence by the sgRNA, and resulting in RNA-DNA pairing when a matching sequence is present [1,2]. Cas9 nucleases with alternative PAMs have also been characterized and successfully used for genome editing [3]. It is important to note that the PAM is not present in the sgRNA sequence but needs to be immediately downstream of the target site in the genomic DNA.

Different CRISPR systems and their uses in genome editing

- Naturally found Cas9 variants are large proteins, which adds particular limitation when it comes to their packaging and delivery into different cells.

- For example, the widely used SpCas9 protein is 1,366 aa, which creates a particular therapeutic delivery challenge. Thus, smaller Cas9 variants have greater therapeutic potential. To this end, the discoveries of 1082 aa NmCas9, 1053 aa SaCas9, and 984 aa CjCas9 are major forward steps toward this goal.

- However, the tradeoff is that these smaller Cas9 proteins require more complex PAM sequences. The SaCas9 requires a 5′-NNGRRT-3′ PAM sequence, whereas CjCas9 requires a 5′-NNNNACAC-3′ PAM sequence. Therefore, these smaller Cas9 proteins have relatively limited *targeting scope and flexibility* in genome targeting compared to SpCas9 despite the reduction in *size*.

Adli, M.The CRISPR tool kit for genome editing and beyond Nature Communications. (2018)9:1911.

Re-engineering CRISPR-Cas9 tools

- Three major goals:
  - 1. Reducing the size of Cas9 nucleases
  - 2. Increasing their fidelity (precision, less off-target effect)
  - 3. Expanding the targeting scope of Cas9 variants

** “It remains a **challenge** to truly determine an inspection process that maps all of the CRISPR-Cas9-mediated DNA *cleavage and binding sites*, as these can be dependent on sgRNA guiding sequences, the cell type, and sgRNA/Cas9 delivery methods.”

Adli, M.The CRISPR tool kit for genome editing and beyond Nature Communications. (2018)9:1911.