Gene Silencing and Genome editing

- RNA interference si RNA silencing
- Zinc Finger Nucleases (ZFNs)
- TALENs Gene Editing
- CRISPR-Cas9 Gene Editing
- Prime editing

Gene Silencing-Small interfering (si) RNA silencing

- An RNA construct complementary to an mRNA sequence leads to inhibition of gene expression and consequently to a decrease of the resultant protein accumulation

- Post-transcriptional gene silencing by RNA interference
- Targets complementary mRNA for degradation

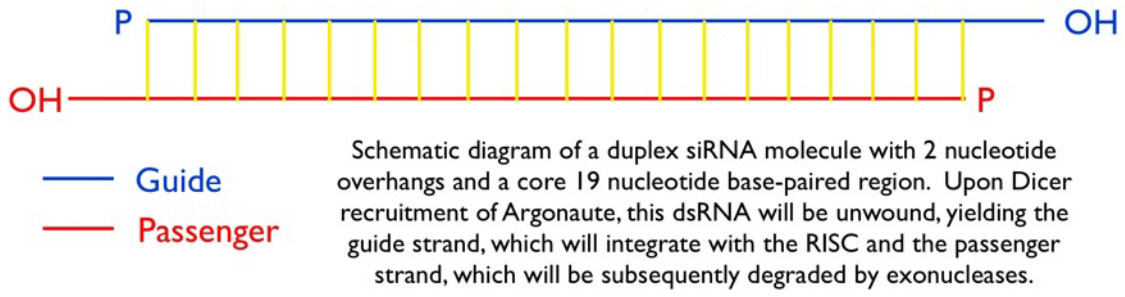
- First developed during experiments with petunia in the ninetieth, later Mello and Fire described RNA interference system for parasite roundworm in 1998 and were awarded by the Nobel price in physiology and medicine in 2006

RNA interference (RNAi)

- RNAi can be in the form of short interfering RNA (siRNA), short hairpin RNA (shRNA) and micro-RNA(miRNA)
- Gene silencing occurs in organisms and more than 6000 micro-RNA (miRNA) sequences have been identified in prokaryotes and eukaryotes with more than 1000 of human origin
- The mechanism of gene silencing can occur either through mRNA degradation or suppression of translation.
- shRNA typically about 80 base pairs in length, that include a region of internal hybridization that creates a hairpin structure.
- shRNAs are recognized by an endogenous enzyme, Dicer, which processes the shRNA into the siRNA duplexes

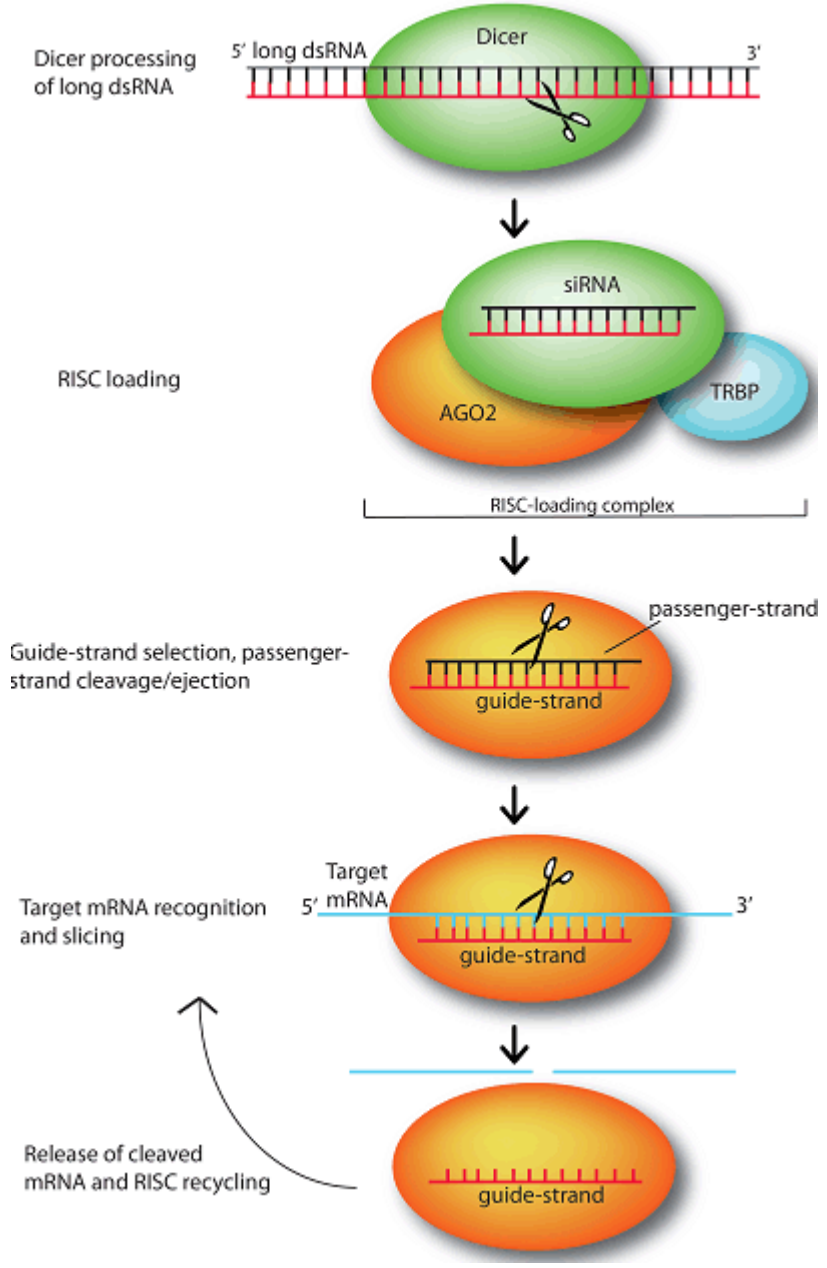
si RNA mechanism

- siRNAs are cleaved off from the long double-stranded RNA molecules by the endoribonuclease Dicer (member of the Rnase III enzyme family) and generates siRNA.
- In the cels, long dsRNAs can arise from the replication of RNA viruses, from the transcription of convergent cellular genes or mobile genetic elements, and from self-annealing cellular transcripts.
- Dicer helps to load small RNA products into large ribonucleoprotein complex called siRISC (RNA induced silencing complex) where it interacts with its Argouate 2 (Ago 2) and functions as a sequence-specific guide targetting the complex to transcripts through base-pairing interactions.
- Ago 2 is a versatile RNA-guided molecular machine that cleaves, or otherwise represses, target RNAs.
- Ago 2 must bind to the guide RNA strand, eject the passenger strand of the siRNA during loading, and subsequently recognize the target RNA.
- The incorporation of siRNAs into the RISC requires the presence of 5′ phosphate groups and 3′ dinucleotide overhangs at the termini



Schematic diagram of a duplex siRNA molecule with 2 nucleotide overhangs and a core 19 nucleotide base-paired region. Upon Dicer recruitment of Argonaute, this dsRNA will be unwound, yielding the guide strand, which will integrate with the RISC and the passenger strand, which will be subsequently degraded by exonucleases.

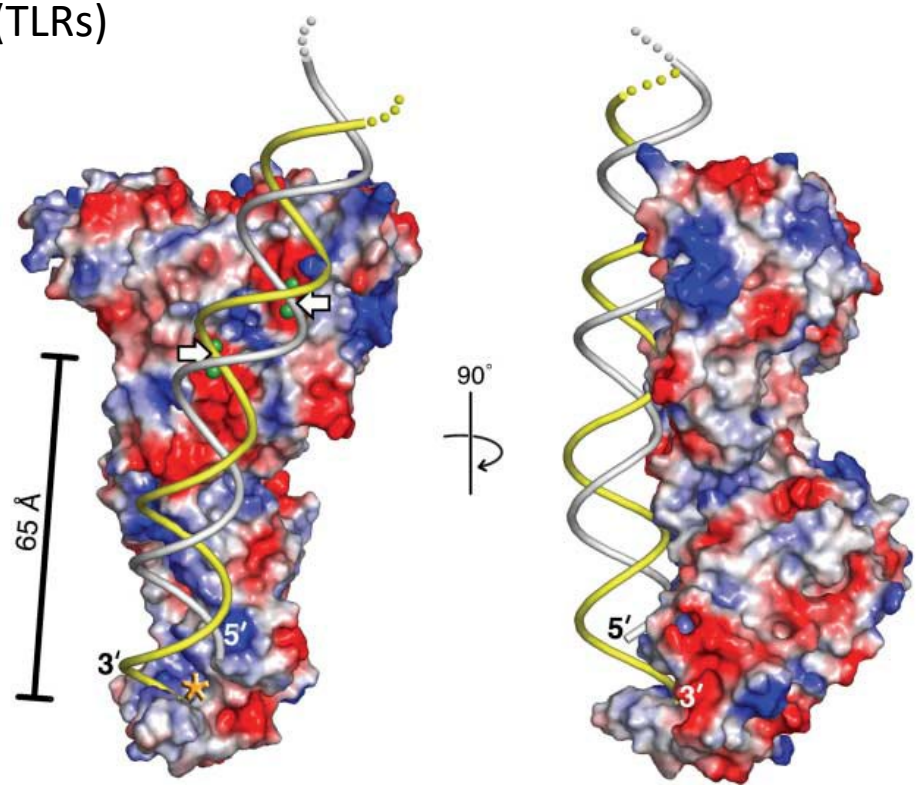
Fmcmanuslab.ucsf.edu



https://images.app.goo.gl/yLpEn4gvvu1Azrx78

Processing by Dicer

Products of RNA processing are short ds RNA (usually 21-25 bp) that are complementary to portions of the transcripts that they regulate.
siRNA longer than 30 nucleotides leads to interferon induction and immune reactions via the activation of toll-like receptors (TLRs)



The active Dicer site is about 65 A long,
The length matches the 25 dsRNAbp segment.
The segments are cleaved from the end of ds RNA.

A model for dsRNA processing by Dicer. Front and side views of a surface representation of Giardia Dicer with modeled dsRNA. Red and blue represent acidic and basic protein surface charge, respectively. Putative catalytic metal ions are shown as green spheres.

MacRae et al., 2006, Science 311, 195 (Doudna lab). The picture is for educational purposes only

Delivery systems

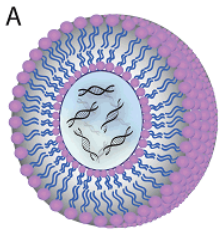
si RNA is a negatively charged macromolecule and is susceptible to degradation by serum endonucleases
Different types of delivery systems for siRNA protection from degradation and enhancing the cellular uptake.

Non-viral delivery: biocompatible, biodegradable, non-toxic, stable

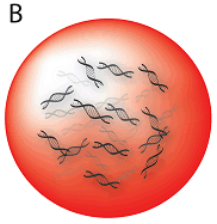
Polymers - cationic polysacharides chitosan, cyclodextran (based nanoparticles),

Peptides - cell penetrating peptides (CPP)

Lipids – liposomes, micelles, emulsions, solid lipid nanoparticles



Liposomes



Polymeric Nanoparticles

Viral delivery:
Adeno-Associated virus (AAV), Lentiviruses (LVs), adenoviruses (AVs) and derived vectors

Hassan, D. et al., 2017 IJBS, 13, 48-56.

Specificity of si RNA silencing

Specificity is not robust

Often the off target exists due to partial sequence homology, especially within 3‘untranslated region

Limitations *in vivo:* si RNA elimination, immune destruction, instability, toxicity, and off-target effects.

Genome Editing

- Hypothesis at the beginning: making double strand breaks at the genome site of interest is toxic and organism evolves mechanism to repair this lesion. The errors may occur during the repair process, which results in a mutation and or exogenous DNA may be incorporated.
- Search for nucleases generating the double strand breaks

Genome Editing Methodologies

Three gene editing methodologies are presently being used:

Zinc Finger Nucleases (ZFNs)

Transcription activator-like effector (TALE) nucleases

CRISPR/Cas system (Clustered regulatory iterspaced short palindromic repeats (RNA-guided Cas 9 nucleases)

DNA is inserted, replaced, removed from a genome using artificially engineered nucleases that create double stranded break and activate cell system to repair the break by homologous recombination (HR) and non-homologous end - joining (NHEJ).

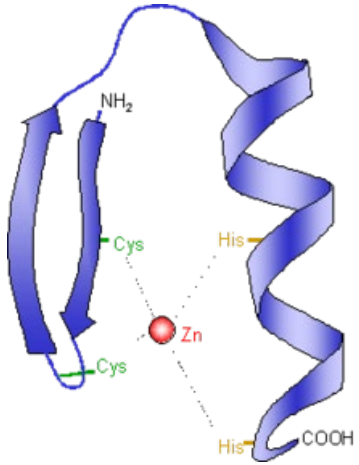
Nucleases can cut any targeted position in the genome

Zinc finger nuclease (ZFN) and TALEN

ZFNs – artificial restriction nucleases consisting of non-specific DNA -cleavage domain from Fok 1 restriction endonuclease fused with recognition - Zn finger domains (1990 technology)

Several Zn fingers (about 30 AA in a conserved ββα configuration) with different specificity for DNA basepairs recognizing.

One Zn finger recognizes 3 bps in DNA .



ZNF

ZNF technology requires design of two unique DNA binding zinc finger domains for sites in the target genome with proper orientation and spacing fused to Fok 1 cleavage domain.

For new target sequence developing new Zinc binding domains.

Dimeriatization of Fok I nuclease induces cleavage of DNA.

DNA cleavage stimulates the cell's natural DNA-repair processes, namely homologous recombination and Non-Homologous End Joining (NHEJ)

TALE and TALENs (Tale nucleases)

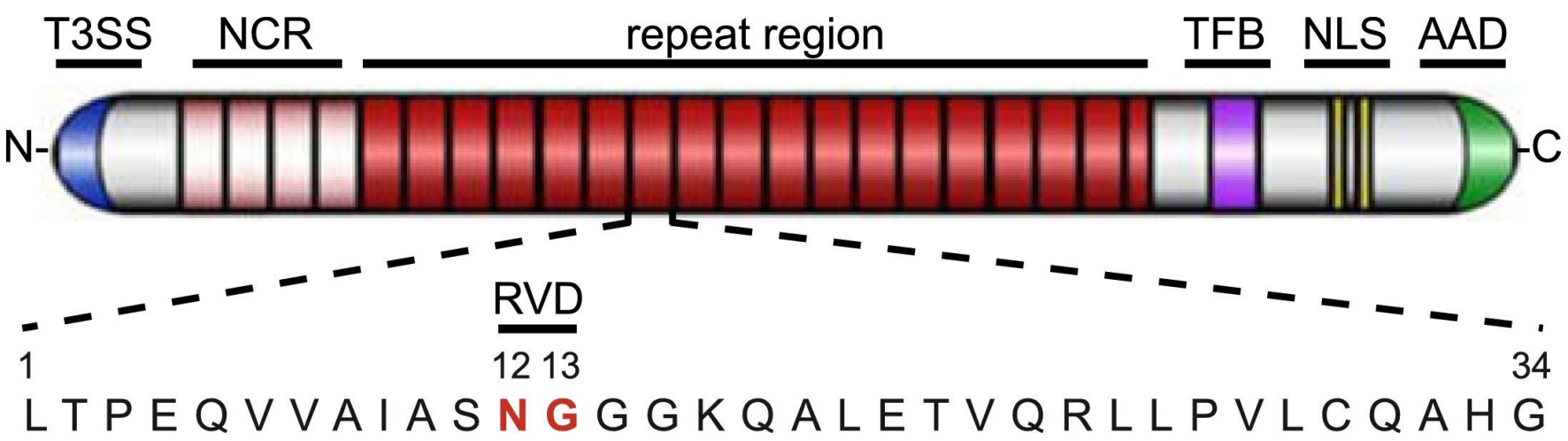
Transcription activator-like effector (TALE): proteins originally produced by plant pathogenic *Xanthomonas* bacteria - activate plant gene expression by binding to promoters (2010)

Talen effector contains repeats of a 33-35 amino acids with divergent 12th and 13th amino acids in the sequence.

These two positions, referred to as the Repeat Variable Di-residues (RVD),
are highly variable and determine the nucleotide binding

Talen can be mutated to generate sequence specific binding by selecting a combination of repeat segments containing the appropriate RVDs.

TALE Architecture



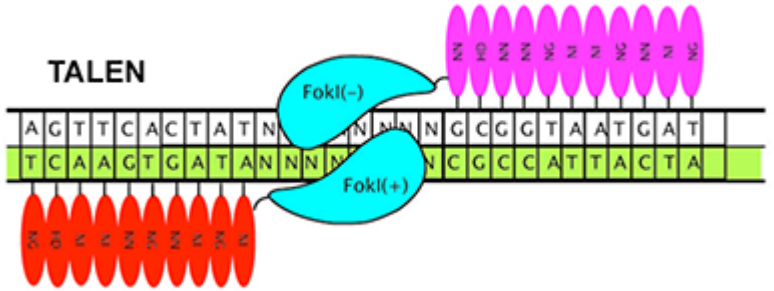
RVD	specificity
NG	T
NI	A
NN	G / A
HD	C
NS	A / G / C / T

Each TALE harbors a repeat region with a varying number of highly conserved repeats of 33–35-AA placed in tandem.
The amino acids responsible for the DNA specificity of a TALE, the repeat variable di-residues (RVD), are in red.

The five most commonly used RVDs and the nucleotides they specify.

Becker and Boch, gene and genome editing, 2021, 2. The picture is for educational purposes only.

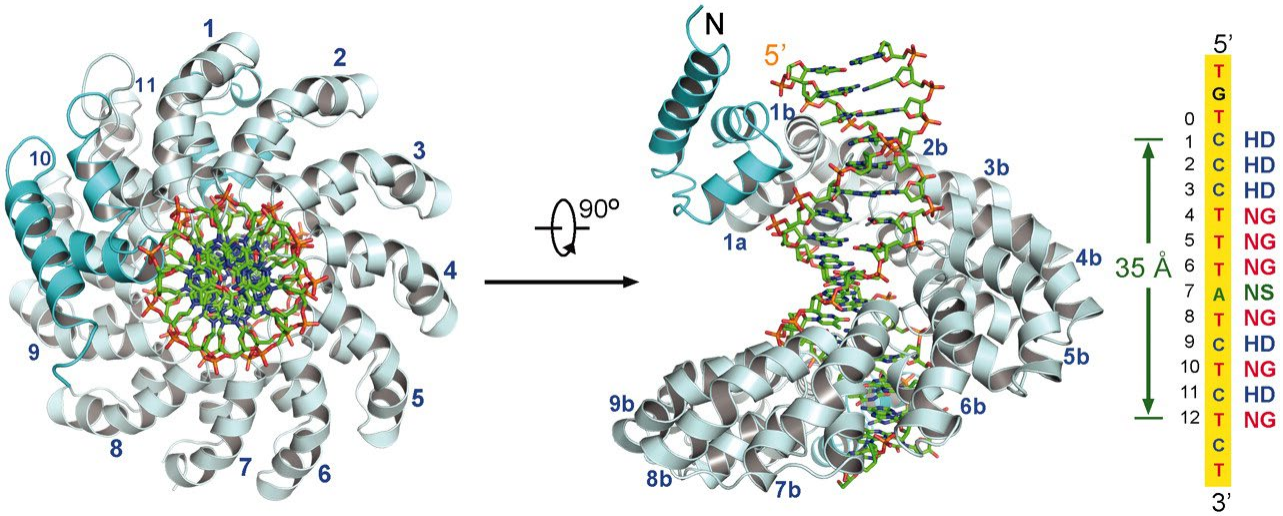
TALENs (Tale nucleases)



TALEN nuclease technology:

Fusion of a TALE DNA-binding domain to a DNA cleavage domain Fok1, which cleaves host DNA between the two distinct TALE array binding sites.

TALE repeats are linked together to recognize DNA sequence.



Overall structure of TALE dHax3, containing 11.5 repeats, bound to DNA.
on the right are the DNA sequence of the sense strand and the corresponding RVDs in dHax3.
The DNA duplex is shown in sticks.

Denk et a.2012, Science 335, 720-723.
The picture is for educational purposes only.

TALEN delivery

Physical methods: microinjection, electroporation

Bacterial-based delivery: directly from *Pseudomonas aeruginosa* via a type III secretion system

Viral-based delivery: viral vectors: adenoviral, lentiviral, baculoviral.

Chemical methods that utilize transfection reagents, liposomes, or PEG.

The delivery of TALEN into plant cells is most often achieved as DNA via *Agrobacterium*-mediated transformation Of embryos, callus, roots or leaf explants.

CRISPR Cas 9 system in bacteria

Inspiration from Bacteria and Archea

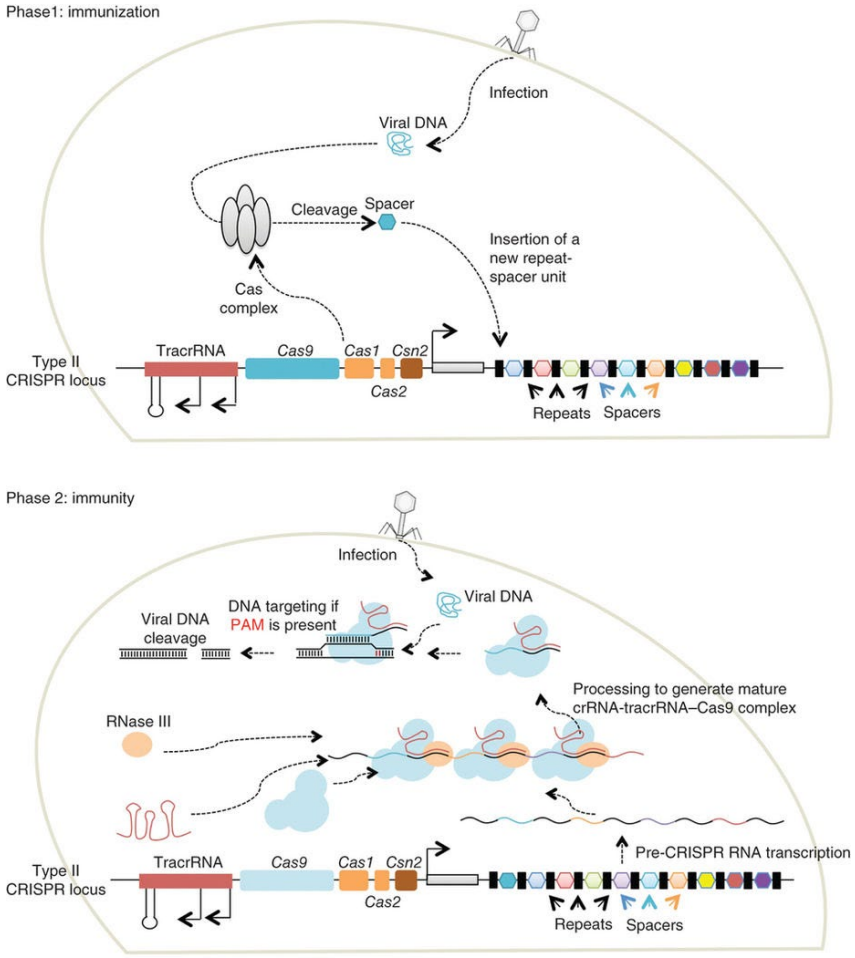
Some prokaryotic microorganisms have defense mechanism against viral invasion.

They integrate DNA sequences identical to past invader into the genome and generate cellular memory- adaptive immune system working as in multicellular organisms.

The integrated sequences allow to recognize the invader and allow to perform degradation of these sequences.

RNA mediated nuclease system based on interaction of RNA with DNA

CRISPR Cas 9 system in bacteria



Mali et al., Nat. Methods 10, 2013, 957.

Incorporation of foreign DNA fragments in highly variable CRISPR loci as a spacers separated by repeat sequences: Repeats (28-37nt), spacers – viral DNA (32-38nt),

CRISPR: Clustered Regulatory Interspaced Short Palindromic Repeats (ssDNA)

Cas 9: RNA guided endonuclease, which unwinds and cleaves the targetted DNA.

Type II CRISPR system from *Streptococcus thermophilus* (a microbibe used in yogurt production)

- Contains highly variable locus in the genome, containing many repeats separated by variable sequences

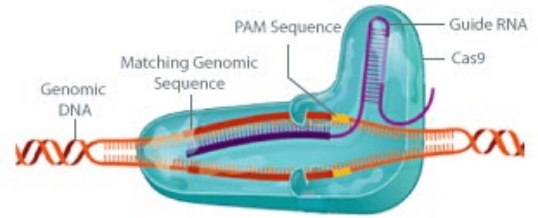
- After acquisition of spacers, crisper RNA (crRNA) with sequences homologous to past invaders is generated and is loaded on Cas, in type II system this RNA is termed pre-crRNA

- Second generated RNA is trans-activating RNA (TracrRNA) from upstream of CRISPR locus, it hybridizes to pre-crRNA and forms dsRNAcrRNA:tracrRNA, which is cleaved by endogenous Rnase III into sequence with one spacer sequence=single guide RNA (sgRNA), which associates with Cas 9.

- This complex is specifically targetted to dsDNA sequence complementary to the spacer sequence and to the Cas binding sequence protospacer-adjacent motif (PAM) in the host genome and Cas9 cuts both strands.

- Different Cas recognize different PAMs

CRISPR-Cas9 system for genome editing in eukaryotes

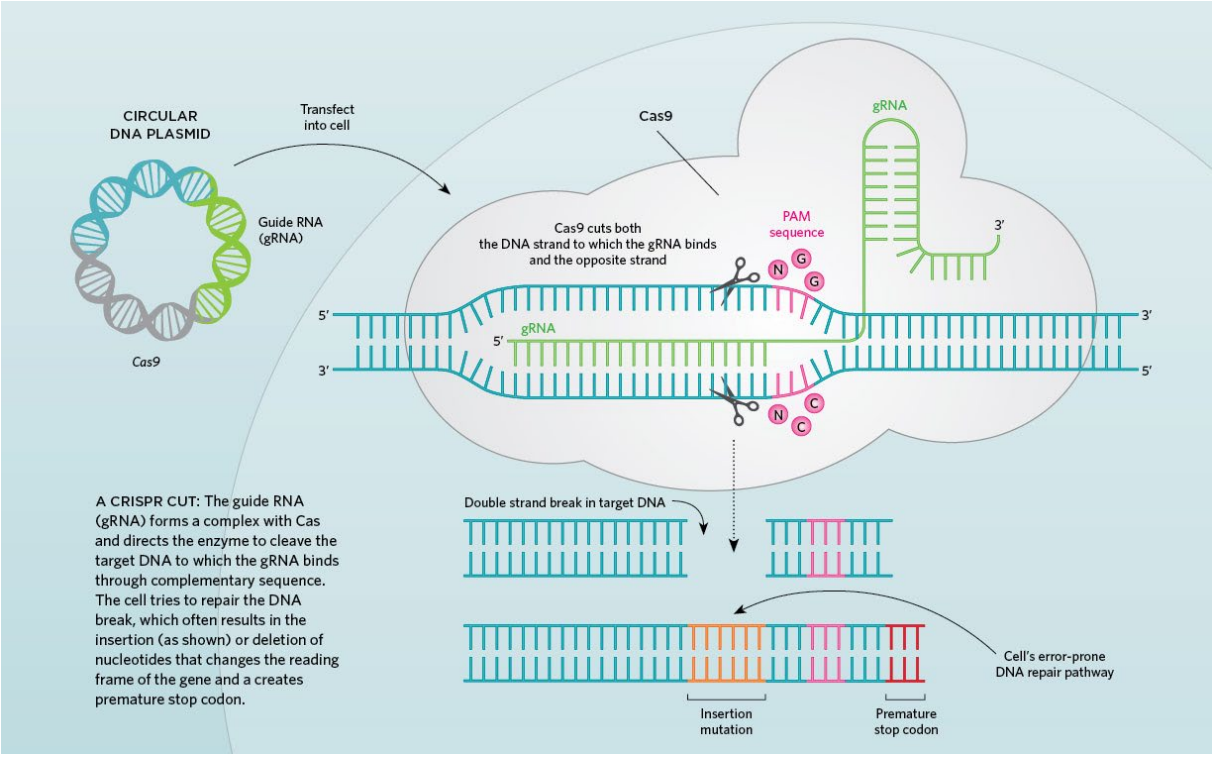


In response to ds break caused by Cas, the cells employ DNA repair pathways to repair the damage using

Non-Homologous End Joining (NHEJ) or Homology-directed repair (HDR).

This is a highly error prone mechanism that causes a few base pair *in*sertions or *del*etions (*indels*) at the cut site. Such an event, in most cases results in frame-shift mutation of the coding sequence, eventually leading to gene disruption (a knock-out).

CRISPR-Cas technology



Chemistry Nobel prize in 2020 awarded to Emmanuelle Charpentier of the Max Planck Institute for Infection Biology and Jennifer Doudna of the University of California, Berkeley.

E. Charpentier discovered tracrRNA in 2011 and together with J. Doudna their described the molecular mechanism of this technology in June,2012 in Science.

They further simplified the tool by showing that crRNA and tracrRNA could be fused to create a single, synthetic guide to direct the cutting mechanism to a site of interest.

In 2013, Doudna's team applied her new technology to human cells.

The guide RNA (gRNA) forms a complex with Cas9 and directs the enzyme to cleave the target DNA to which the gRNA binds through complementary sequence. The cell tries to repair the DNA break, which often results in the insertion (as shown) or deletion of nucleotides that changes the reading frame of the gene and a creates premature stop codon.

CRISPR applications

CRISPR has contributed to discoveries across disciplines, conferring mold and pest resistance to important crops, leading to new cancer therapies and, more recently, repairing damage within mitochondrial DNA.

Discovery of new pathways for developing therapeutics against bacteria, treating genetic diseases, fighting cancer

Find new mechanisms to target genes and their expression.

CRISPR-Cas9 system Application

All components are delivered in one plasmid employed for cell transfection.

For new application we need to: design sgRNA

The NGG PAM motif for Cas 9 ocurs on average once every 8bp in almost any gene of interest to be targeted.

The advantages of the CRISPR/Cas9 genome editing system include the ability of editing multiple genes simultaneously, a simple and fast design process that does not require the reengineering of the nuclease for each target, and an editing efficiency that is similar or greater than ZFNs and TALENs.

CRISPR-Cas9 system Application

Analysis of a specific gene roles
Generating models for diseases studies, e.g. cancer

Targeting viruses – HIV, HBV, papillomaviruses, EBV viruses

Identification of interacting DNA proteins using inactive Cas9

Base editing

Most pathogenic mutations that cause human disease are single nucleotide polymorphisms that only require a single nucleotide change to correct the mutation.

CRISPR, ZFNs, or TALENs can be precise but their repair leads to mutations

Base editors: genome editing approach using components from CRISPR systems together with other enzymes to directly install point mutations into cellular DNA or RNA without making double-stranded DNA breaks (DSBs)

Base editors: two main categories: targeting DNA and RNA.

DNA base editors: **cytosine base editors (CBEs) or adenine base editors (ABEs).**

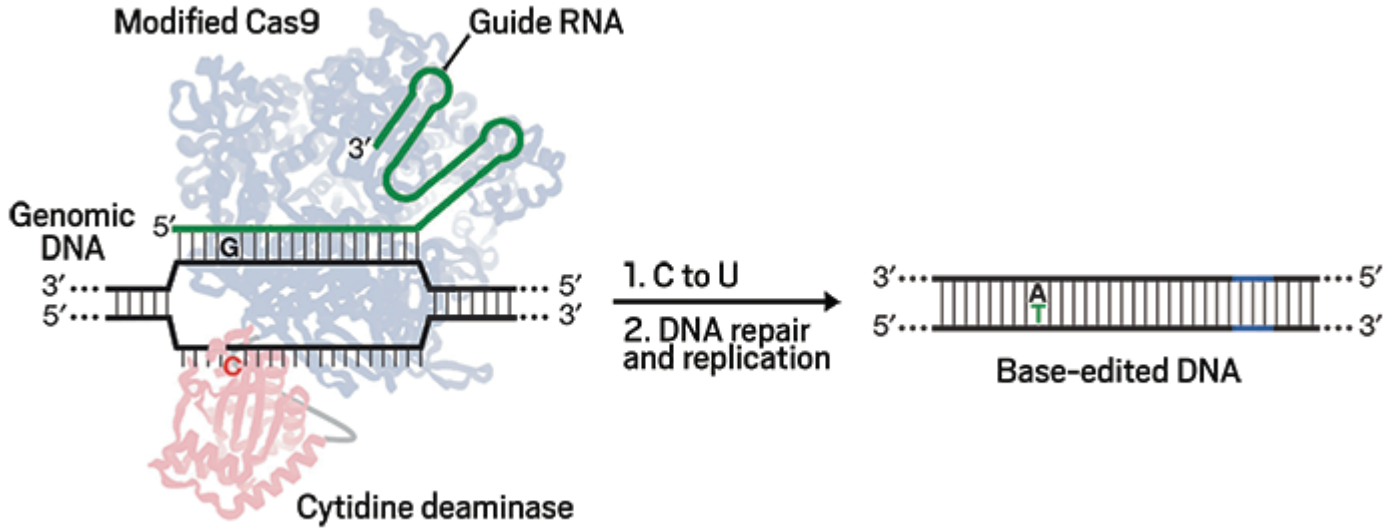
Cytosine base editors (CBEs) convert a C•G base pair into a T•A base pair by deaminating the amine of cytosine to generate uracil

Adenine base editors (ABEs) convert an A•T base pair to a G•C base pair.
Both CBEs and ABEs are powerful tools for the permanent introduction of point mutations in DNA in living cells.

CBEs and ABEs can mediate all four possible transition mutations C to T, A to G, T to C, and G to A

Cytosine base editors (CBE)

Cytosine base editor 1 (BE1, Komor et al.,) fused APOBEC1 cytidine deaminase, accepting only ssDNA as a substrate with dead Cas9 from *Streptococcus pyrogenes* (dCas9, a dead Cas9 mutant containing D10A and H840A)
Fusion of APOBEC1 to dead Cas9 from *Streptococcus pyrogenes* (dCas9, a mutant of Cas9 containing D10A and H840A) resulted in base editor 1 (BE1)



This picture is for educational purpose only

Komor AC, Kim YB, Packer MS, Zuris JA & Liu DR Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. Nature 533, 420–424, 10.1038/nature17946 (2016)
Borman, Biol. Chem. 2016, 94 (17)

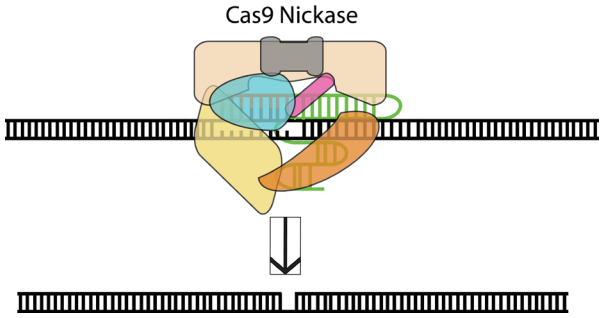
Next generations of CBE

BE2 editor: **BE1 fused with** uracil DNA glycosylase inhibitor (**UGI**), which inhibits uracil N-glycosylase (UNG) and base excision repair of U-G intermediate back to C-G.

BE2 mediates efficient base editing in bacterial cells and moderately efficient editing in mammalian cells, enabling conversion of a C•G base pair to a T•A base pair through a U•G intermediate.
BE2 can edit only one strand of DNA.

BE3: **APOBEC1– Cas9 nickase–UGI:**

Cas9 nuclease may be mutated to create DNA "nickases" capable of introducing a single-strand cut with the same specificity as a regular CRISPR/Cas9 nuclease. The use of Cas9 nickases is essentially the same as the use of the fully functional enzyme.



Adenine base editors (ABE)

ABE converts A to G conversion. ABEs also induce cytosine deamination at the target site

Like cytosine, adenine contains an exocyclic amine that can be deaminated to alter its base pairing preferences. Deamination of adenosine yields inosine.

Use of tRNA adenosine deaminase TadA32 deoxyadenosine deaminase accepting ssDNA tRNA adenosine deaminase enzyme, TadA32

To reduce the cytosine editing activity, adenosine deaminase, TadA7.10 was mutated (D108Q), and resulting ABE7.10 D108Q mutant exhibited tenfold reduced cytosine deamination activity.

Various ABE vectors based on different engineered adenosine deaminase (TadA) proteins fused to Cas9 variants (dCas9, nCas9), enabling efficient A to G editing up to 64% efficiency on-target sites

Base editing

Base editing can efficiently install the four transition mutations (C→T, G→A, A→G, and T→C) without requiring DSBs in many cell types and organisms, including mammals.

Base editing cannot currently perform the eight transversion mutations

(C→A, C→G, G→C, G→T, A→C, A→T, T→A, and T→G)

No DSB-free method cannot perform targeted deletions, such as the removal of the 4-base duplication or targeted insertions

CRISPR Prime Editing (PE)

- Versatile and precise genome editing method that directly writes new genetic information into a specified DNA site using a catalytically impaired Cas9 endonuclease (nickage) fused to an engineered reverse transcriptase, programmed with a prime editing guide RNA (pegRNA) that specifies the target site and encodes the desired edit.
- No making of double-strand DNA breaks (DSBs) by active Cas9
- PE mediates all 12 possible base-to-base conversions, insertions, deletions

Anzalone et al., 2019, Nature 576 (149-157)

CRISPR Prime Editing

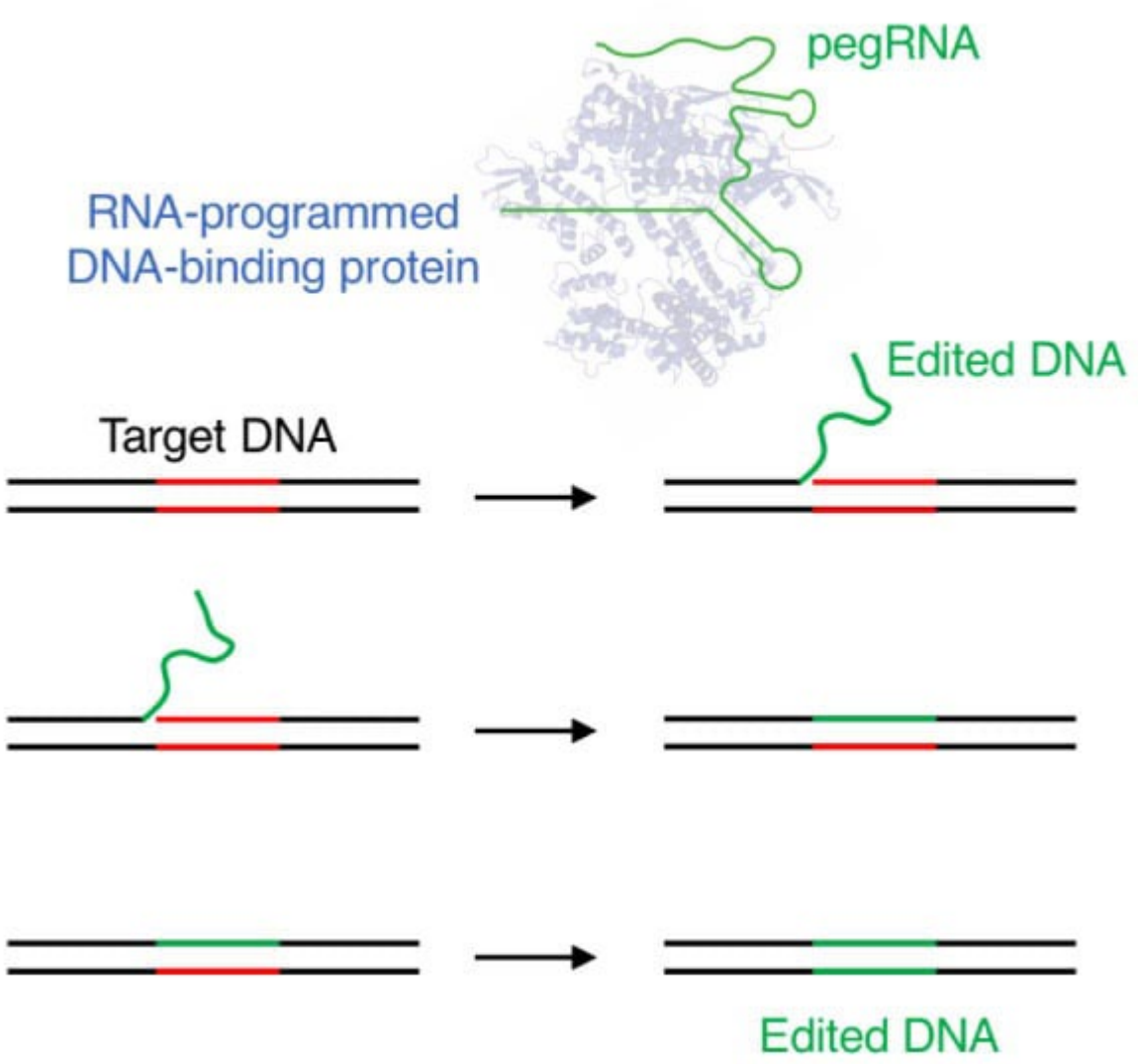
PE shows higher or similar efficiency and fewer byproducts than homology-directed repair

PE induces much lower off-target editing than Cas9 nuclease at known Cas9 off-target sites

PE requires:
Prime editor protein: Nicked Cas 9 fused with domain of reverse transcriptase. (M-MLV reverse transcriptase (RT) with the Cas9 H840A). Nickage is guided to the DNA target site by the pegRNA

pegRNAs:specify both the DNA target and contain new genetic information that replaces target DNA nucleotides. The pegRNA is a guide RNA that also encodes the RT template, which includes the desired edit and homology to the genomic DNA locus

After nicking by Cas9, the reverse transcriptase domain uses the pegRNA to template reverse transcription of the desired edit, directly polymerizing DNA onto the nicked target DNA strand.



Jenifer Tsang, Addgene.org. This picture is for educational purpose only

The edited DNA strand replaces the original DNA strand, creating a heteroduplex containing one edited strand and one unedited strand.

The editor guides resolution of the heteroduplex to favor copying the edit onto the unedited strand, completing the process.