[thumb|450px|right|Diagram of the CRISPR prokaryotic antiviral defense mechanism.](/wiki/Image:Crispr.png" \o "Image:Crispr.png)[[1]](#cite_note-1)

**Clustered regularly interspaced short palindromic repeats** (**CRISPR**, pronounced *crisper*[[2]](#cite_note-2)) are segments of [prokaryotic](/wiki/Prokaryotic) [DNA](/wiki/DNA) containing short repetitions of base sequences. Each repetition is followed by short segments of "[spacer DNA](/wiki/Spacer_DNA)" from previous exposures to a bacteriophage virus or [plasmid](/wiki/Plasmid).[[3]](#cite_note-3) The **CRISPR/Cas system** is a prokaryotic [immune system](/wiki/Immune_system) that confers resistance to foreign genetic elements such as those present within [plasmids](/wiki/Plasmids) and [phages](/wiki/Phage),[[4]](#cite_note-4)[[5]](#cite_note-5)[[6]](#cite_note-6) and provides a form of [acquired immunity](/wiki/Acquired_immunity). CRISPR spacers recognize and cut these exogenous genetic elements in a manner analogous to [RNA interference](/wiki/RNA_interference) in [eukaryotic](/wiki/Eukaryote) organisms.[[3]](#cite_note-3) CRISPRs are found in approximately 40% of sequenced [bacterial](/wiki/Bacterial) genomes and 90% of sequenced [archaea](/wiki/Archaea).[[7]](#cite_note-7)[[note 1]](#cite_note-8) By delivering the [Cas9](/wiki/Cas9) nuclease and appropriate guide RNAs into a cell, the cell's [genome](/wiki/Genome) can be cut at a desired location, allowing existing genes to be removed and/or new ones added.[[8]](#cite_note-9)[[9]](#cite_note-10)[[10]](#cite_note-11) CRISPRs have been used in concert with specific [endonuclease](/wiki/Endonuclease) [enzymes](/wiki/Enzymes) for [genome editing](/wiki/Genome_editing) and gene regulation in species throughout the [tree of life](/wiki/Tree_of_life_(biology)).[[11]](#cite_note-12) Cas9 was the first nuclease discovered, followed by Cpf1, which was discovered in the [CRISPR/Cpf1](/wiki/CRISPR/Cpf1) system of [Francisella novicida](/wiki/Francisella_novicida).[[12]](#cite_note-13)[[13]](#cite_note-14) Other such systems are thought to exist.[[14]](#cite_note-15) CRISPR/C2c2 from the bacterium [Leptotrichia shahii](/wiki/Fusobacteria) is RNA-guided CRISPR system that targets RNA rather than DNA, and can either cleave single-stranded RNA targets or [knock them down](/wiki/Gene_knockdown).[[15]](#cite_note-16) The [CRISPR interference](/wiki/CRISPR_interference) technique has many potential applications, including altering the [germline](/wiki/Germline) of humans, animals, and food crops. The use of CRISPR for genome editing[[16]](#cite_note-17)[[17]](#cite_note-18) was the [AAAS's](/wiki/American_Association_for_the_Advancement_of_Science) choice for breakthrough of the year in 2015.[[18]](#cite_note-19) [Bioethical](/wiki/Bioethical) concerns have been expressed about the prospect of using this nascent biotechnology for editing the human germline.[[19]](#cite_note-20) [Template:Toclimit](/wiki/Template:Toclimit)

[Template:Infobox nonhuman protein](/wiki/Template:Infobox_nonhuman_protein)

## Contents

* 1 History[[edit](/index.php?title=(none)&action=edit&section=1)]
  + 1.1 Cas9[[edit](/index.php?title=(none)&action=edit&section=2)]
  + 1.2 Cpf1[[edit](/index.php?title=(none)&action=edit&section=3)]
  + 1.3 Predecessors[[edit](/index.php?title=(none)&action=edit&section=4)]
* 2 Locus structure[[edit](/index.php?title=(none)&action=edit&section=5)]
  + 2.1 Repeats and spacers[[edit](/index.php?title=(none)&action=edit&section=6)]
  + 2.2 Cas genes and CRISPR subtypes[[edit](/index.php?title=(none)&action=edit&section=7)]
* 3 Mechanism[[edit](/index.php?title=(none)&action=edit&section=8)]
  + 3.1 Spacer acquisition[[edit](/index.php?title=(none)&action=edit&section=9)]
    - 3.1.1 Protospacer adjacent motifs[[edit](/index.php?title=(none)&action=edit&section=10)]
    - 3.1.2 Insertion variants[[edit](/index.php?title=(none)&action=edit&section=11)]
  + 3.2 Biogenesis[[edit](/index.php?title=(none)&action=edit&section=12)]
  + 3.3 Interference[[edit](/index.php?title=(none)&action=edit&section=13)]
* 4 Evolution and diversity[[edit](/index.php?title=(none)&action=edit&section=14)]
* 5 Identification[[edit](/index.php?title=(none)&action=edit&section=15)]
* 6 Evolutionary significance[[edit](/index.php?title=(none)&action=edit&section=16)]
* 7 Use by phages[[edit](/index.php?title=(none)&action=edit&section=17)]
* 8 Applications[[edit](/index.php?title=(none)&action=edit&section=18)]
  + 8.1 Genome engineering[[edit](/index.php?title=(none)&action=edit&section=19)]
    - 8.1.1 Major components[[edit](/index.php?title=(none)&action=edit&section=20)]
    - 8.1.2 Structure[[edit](/index.php?title=(none)&action=edit&section=21)]
    - 8.1.3 Delivery[[edit](/index.php?title=(none)&action=edit&section=22)]
  + 8.2 Editing[[edit](/index.php?title=(none)&action=edit&section=23)]
  + 8.3 Knockdown/activation[[edit](/index.php?title=(none)&action=edit&section=24)]
  + 8.4 RNA editing[[edit](/index.php?title=(none)&action=edit&section=25)]
  + 8.5 Disease models[[edit](/index.php?title=(none)&action=edit&section=26)]
  + 8.6 Gene drive[[edit](/index.php?title=(none)&action=edit&section=27)]
  + 8.7 Biomedicine[[edit](/index.php?title=(none)&action=edit&section=28)]
  + 8.8 Gene function[[edit](/index.php?title=(none)&action=edit&section=29)]
  + 8.9 ''In vitro'' genetic depletion[[edit](/index.php?title=(none)&action=edit&section=30)]
* 9 Patents and commercialization[[edit](/index.php?title=(none)&action=edit&section=31)]
* 10 Society and culture[[edit](/index.php?title=(none)&action=edit&section=32)]
  + 10.1 Human germline modification[[edit](/index.php?title=(none)&action=edit&section=33)]
  + 10.2 Recognition[[edit](/index.php?title=(none)&action=edit&section=34)]
* 11 Alternative cutters[[edit](/index.php?title=(none)&action=edit&section=35)]
* 12 See also[[edit](/index.php?title=(none)&action=edit&section=36)]
* 13 Notes[[edit](/index.php?title=(none)&action=edit&section=37)]
* 14 References[[edit](/index.php?title=(none)&action=edit&section=38)]
* 15 Further reading[[edit](/index.php?title=(none)&action=edit&section=39)]

## History[[edit](/index.php?title=(none)&action=edit&section=1)]

Clustered DNA repeats with short unique sequences between the repeat units were first described in the bacterium [*Escherichia coli*](/wiki/Escherichia_coli) by Osaka University researcher [Yoshizumi Ishino](/wiki/Yoshizumi_Ishino) in 1987. The organization of the repeats was unusual because repeated sequences are typically arranged consecutively along the DNA. The function of the interrupted clustered repeats was not known at the time.<ref name=Ishino1987>[Template:Cite journal](/wiki/Template:Cite_journal)[Template:Open access](/wiki/Template:Open_access)</ref><ref name=Hsu2014/> Repeats that were similarly interrupted were observed in an [archaeal](/wiki/Archaea) microbe by Spanish researcher Francisco Mojica in 1993.[[20]](#cite_note-21) He later named them short regularly spaced repeats (SRSR) after he found them in many other microbes.[[21]](#cite_note-22) SRSR was renamed CRISPR in 2002 after a suggestion from Mojica.[[20]](#cite_note-21)[[22]](#cite_note-23)In the same year, a set of genes was found to be associated with CRISPR repeats and was named [*cas*](/wiki/CRISPR-associated_systems), or *CRISPR-associated* genes. The function of CRISPR was assumed to be related to the *cas* genes. Analysis of the protein sequences encoded by the *cas* genes suggested that one was a [helicase](/wiki/Helicase) that unwound DNA and that another was a [nuclease](/wiki/Nuclease) that digested DNA.[[22]](#cite_note-23) [thumb|Simplified diagram of a CRISPR locus. The three major components of a CRISPR locus are shown: *cas* genes, a leader sequence, and a repeat-spacer array. Repeats are shown as gray boxes and spacers are colored bars. The arrangement of the three components is not always as shown.](/wiki/File:SimpleCRISPR.jpg)[[1]](#cite_note-1)[[3]](#cite_note-3)In addition, several CRISPRs with similar sequences can be present in a single genome, only one of which is associated with *cas* genes.[[7]](#cite_note-7)

CRISPR was proposed to be responsible for [adaptive immunity](/wiki/Adaptive_immune_system) in microbes.[[20]](#cite_note-21) In 2005, three independent research groups showed that some CRISPR spacers are derived from [phage](/wiki/Phage) DNA and [extrachromosomal DNA](/wiki/Extrachromosomal_DNA) such as [plasmids](/wiki/Plasmids).[[23]](#cite_note-24)[[24]](#cite_note-25)[[25]](#cite_note-26) In effect, the spacers are fragments of DNA gathered from viruses that previously tried to attack the cell. The source of the spacers was a sign that the CRISPR/*cas* system could have a role in adaptive immunity in bacteria.[[1]](#cite_note-1)[[26]](#cite_note-27) All three studies proposing this idea were initially rejected by high-profile journals but were eventually published in other journals.[[20]](#cite_note-21) Although researchers came to an understanding that CRISPR-Cas was involved in microbial immunity, the mechanism of how the spacers conferred immunity was a mystery.<ref name=Hsu2014/> Koonin and colleagues proposed that spacers produce small RNA guides to target RNA transcribed from viral DNA, analogous to the [RNA interference](/wiki/RNA_interference) system used by eukaryotic cells.<ref name=Hsu2014/><ref name=Makarova2006>[Template:Cite journal](/wiki/Template:Cite_journal)</ref> Others hypothesized that CRISPR sequences directed Cas enzymes to degrade viral DNA.[[25]](#cite_note-26) Experimental work by several groups revealed the basic mechanisms of CRISPR-Cas immunity. In 2007, Barrangou, Horvath, and other food industry scientists at [Danisco](/wiki/Danisco) provided the first experimental evidence that CRISPR was an adaptive immune system.<ref name=Hsu2014/> A CRISPR region in [*Streptococcus thermophilus*](/wiki/Streptococcus_thermophilus) acquired spacers from DNA of infecting bacteriophage. The researchers manipulated the resistance of *S. thermophilus* to phage by adding and deleting spacers whose sequence matched those found in the phages tested.[[27]](#cite_note-28) In 2008, Brouns and colleagues identified a complex of Cas protein that in *E. coli* cut the CRISPR RNA within the repeats into spacer-containing RNA molecules, which remained bound to the protein complex. In the same year, Marraffini and Sontheimer showed that a CRISPR sequence of [*Staphylococcus epidermidis*](/wiki/Staphylococcus_epidermidis) targeted DNA and not RNA to prevent [conjugation](/wiki/Conjugation_(genetics)). This finding was at odds with the proposed RNA-interference-like mechanism of CRISPR-Cas immunity, although a CRISPR-Cas system that targets foreign RNA was later found in [*Pyrococcus furiosus*](/wiki/Pyrococcus_furiosus).<ref name=Hsu2014/><ref name=Marraffini2015/> A 2010 study provided direct evidence that CRISPR-Cas cuts both strands of phage and plasmid DNA in *S. thermophilus*.<ref name=Garneau2010>[Template:Cite journal](/wiki/Template:Cite_journal)</ref>[[28]](#cite_note-29)

### Cas9[[edit](/index.php?title=(none)&action=edit&section=2)]

[Jennifer Doudna](/wiki/Jennifer_Doudna) and [Emmanuelle Charpentier](/wiki/Emmanuelle_Charpentier) studied a simpler CRISPR system from [*Streptococcus pyogenes*](/wiki/Streptococcus_pyogenes) that relies on the protein [Cas9](/wiki/Cas9). The Cas9 endonuclease is a four-component system that includes two small RNA molecules.<ref name=Barrangou2015>[Template:Cite journal](/wiki/Template:Cite_journal)</ref> In 2012, they re-engineered Cas9 endonuclease into a more manageable two-component system by fusing the two RNA molecules into a "single-guide RNA" that, when mixed with Cas9, could find and cut the DNA target specified by the guide RNA. By manipulating the nucleotide sequence of the guide RNA, the artificial Cas9 system could be programmed to target any sequence in DNA for cleavage.<ref name=pmid22745249/> This technological advance has fueled efforts to edit genomes with the re-engineered CRISPR-Cas9 system.<ref name=Barrangou2016>[Template:Cite journal](/wiki/Template:Cite_journal)</ref>

Feng Zhang's and [George Church's](/wiki/George_M._Church) groups simultaneously described genome editing in human cell cultures using CRISPR-Cas9 systems for the first time.<ref name=Hsu2014>[Template:Cite journal](/wiki/Template:Cite_journal)</ref><ref name=Cong2013>[Template:Cite journal](/wiki/Template:Cite_journal)</ref><ref name=Mali2013>[Template:Cite journal](/wiki/Template:Cite_journal)</ref> It has since been used in a wide range of organisms, including baker's yeast ([*Saccharomyces cerevisiae*](/wiki/Saccharomyces_cerevisiae)),[[29]](#cite_note-30) zebrafish ([*D. rerio*](/wiki/Zebrafish)),[[30]](#cite_note-31) fruit flies ([*Drosophila melanogaster*](/wiki/Drosophila_melanogaster)),[[31]](#cite_note-32) axolotl ([*A. mexicanum*](/wiki/Axolotl)),[[32]](#cite_note-33) nematodes,([*C. elegans*](/wiki/Caenorhabditis_elegans)),[[33]](#cite_note-34) plants,[[34]](#cite_note-35) mice,[[35]](#cite_note-36) monkeys[[36]](#cite_note-37) and human embryos.[[37]](#cite_note-38) CRISPR has been modified to make programmable [transcription factors](/wiki/Transcription_factors) that allow scientists to target and activate or silence specific genes.[[38]](#cite_note-39) Libraries of tens of thousands of [guide RNAs](/wiki/Guide_RNA) are available.<ref name=craze/>

### Cpf1[[edit](/index.php?title=(none)&action=edit&section=3)]

In 2015, the nuclease Cpf1 was discovered in the [CRISPR/Cpf1](/wiki/CRISPR/Cpf1) system of the bacterium [*Francisella novicida*](/wiki/Francisella_novicida).[[12]](#cite_note-13)[[13]](#cite_note-14)Cpf1 showed several key differences to Cas9 including; causing a 'staggered' cut in double stranded DNA as opposed to the 'blunt' cut produced by Cas9, relying on a 'T rich' [Protospacer adjacent motif](/wiki/Protospacer_adjacent_motif) (providing alternate targeting sites to Cas9) and requiring only a CRISPR RNA (crRNA) for successful targeting (with Cas9 requiring both crRNA and a [transactivating crRNA](/wiki/Trans-activating_crRNA) (tracrRNA)).

### Predecessors[[edit](/index.php?title=(none)&action=edit&section=4)]

In the early 2000s, researchers developed [zinc finger nucleases](/wiki/Zinc_finger_nuclease), synthetic proteins whose [DNA-binding domains](/wiki/DNA-binding_domains) enable them to create double-stranded breaks in DNA at specific points. In 2010, synthetic nucleases called [TALENs](/wiki/TALEN) provided an easier way to target a double-stranded break to a specific location on the DNA strand. Both zinc finger nucleases and TALENs require the creation of a custom protein for each targeted DNA sequence, which is a more difficult and time-consuming process than that for guide RNAs. CRISPRs are much easier to design because the process requires making only a short RNA sequence.<ref name = MIT>[Template:Cite journal](/wiki/Template:Cite_journal)</ref>

## Locus structure[[edit](/index.php?title=(none)&action=edit&section=5)]

### Repeats and spacers[[edit](/index.php?title=(none)&action=edit&section=6)]

CRISPR repeats range in size from 24 to 48 [base pairs](/wiki/Base_pairs).[[39]](#cite_note-40) They usually show some [dyad symmetry](/wiki/Dyad_symmetry), implying the formation of a [secondary structure](/wiki/Nucleic_acid_secondary_structure) such as a [hairpin](/wiki/Stem-loop), but are not truly [palindromic](/wiki/Palindromic_sequence).[[40]](#cite_note-41) Repeats are separated by spacers of similar length.[[39]](#cite_note-40) Some CRISPR spacer sequences exactly match sequences from [plasmids](/wiki/Plasmids) and [phages](/wiki/Phage),[[23]](#cite_note-24)[[24]](#cite_note-25)[[25]](#cite_note-26) although some spacers match the prokaryote's genome (self-targeting spacers).[[23]](#cite_note-24)[[41]](#cite_note-42) New spacers can be added rapidly as part of the immune response to phage infection.[[42]](#cite_note-43)

### Cas genes and CRISPR subtypes[[edit](/index.php?title=(none)&action=edit&section=7)]

Small clusters of *cas* genes are often located next to CRISPR repeat-spacer arrays. Comparative genomics identified multiple *cas* genes; an initial analysis of 200 bacterial and archaeal genomes suggested as many as 45 *cas* gene families.[[39]](#cite_note-40) The current classification groups CRISPR-Cas systems into two classes. Class 1 systems use a complex of multiple Cas proteins to degrade foreign nucleic acids. Class 2 systems use a single large Cas protein for the same purpose. Class 1 is divided into types I, III, and IV; class 2 is divided into types II and V. The five system types are divided into 16 subtypes. Each type and most subtypes are characterized by a "signature gene" found exclusively in the category. Classification is also based on the complement of *cas* genes that are present. Most CRISPR-Cas systems have a Cas1 protein. The [phylogeny](/wiki/Phylogeny) of Cas1 proteins generally agrees with the classification system.<ref name=Markarova2015>[Template:Cite journal](/wiki/Template:Cite_journal)</ref> Many organisms contain multiple CRISPR-Cas systems suggesting that they are compatible and may share components.[[28]](#cite_note-29)[[43]](#cite_note-44) The sporadic distribution of the CRISPR/Cas subtypes suggests that the CRISPR/Cas system is subject to [horizontal gene transfer](/wiki/Horizontal_gene_transfer) during microbial [evolution](/wiki/Evolution).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Signature genes and their putative functions for the major and minor CRISPR-cas types. | | | | |
| **Class** | **Cas type** | **Signature gene** | **Function** | **Reference** |
| 1 | I | Cas3 | Single-stranded DNA nuclease (HD domain) and ATP-dependent helicase | [[44]](#cite_note-45)[[45]](#cite_note-46) |
|  | IA | Cas8a | Subunit of the interference module. Important in targeting of invading DNA by recognizing the [PAM](/wiki/Protospacer_adjacent_motif) sequence | [[46]](#cite_note-47)[[47]](#cite_note-48) |
|  | IB | Cas8b |
|  | IC | Cas8c |
|  | ID | Cas10d | contains a domain homologous to the palm domain of nucleic acid polymerases and nucleotide cyclases | [[48]](#cite_note-49)[[49]](#cite_note-50) |
|  | IE | Cse1 |
|  | IF | Csy1 | Not determined |  |
|  | IU |  |  |  |
|  | III | Cas10 | Homolog of Cas10d and Cse1 | [[49]](#cite_note-50) |
|  | IIIA | Csm2 | Not Determined |  |
|  | IIIB | Cmr5 | Not Determined |  |
|  | IIIC |  |  |  |
|  | IIID |  |  |  |
|  | IV | Csf1 |  |  |
|  | IVA |  |  |  |
|  | IVB |  |  |  |
| 2 | II | [Cas9](/wiki/Cas9) | [Nucleases](/wiki/Nuclease) RuvC and HNH together produce [DSBs](/wiki/Double-stranded_breaks), and separately can produce single-strand breaks. Ensures the acquisition of functional spacers during adaptation. | [[50]](#cite_note-51)[[51]](#cite_note-52) |
|  | IIA | Csn2 | Ring-shaped DNA-binding protein. Involved in primed adaptation in Type II CRISPR system. | [[52]](#cite_note-53) |
|  | IIB | Cas4 | Not Determined |  |
|  | IIC |  | Characterized by the absence of either Csn2 or Cas4 | [[53]](#cite_note-54) |
|  | V | Cpf1 | Nuclease RuvC. Lacks HNH. |  |

## Mechanism[[edit](/index.php?title=(none)&action=edit&section=8)]

[frame|The stages of CRISPR immunity for each of the three major types of adaptive immunity. (1) Acquisition begins by recognition of invading DNA by](/wiki/File:The_Stages_of_CRISPR_immunity.svg) [Cas1](/wiki/Cas1) and Cas2 and cleavage of a protospacer. (2) The protospacer is ligated to the direct repeat adjacent to the leader sequence and (3) single strand extension repairs the CRISPR and duplicates the direct repeat. The crRNA processing and interference stages occur differently in each of the three major CRISPR systems. (4) The primary CRISPR transcript is cleaved by cas genes to produce crRNAs. (5) In type I systems Cas6e/Cas6f cleave at the junction of ssRNA and dsRNA formed by hairpin loops in the direct repeat. Type II systems use a trans-activating (tracr) RNA to form dsRNA, which is cleaved by [Cas9](/wiki/Cas9) and RNaseIII. Type III systems use a Cas6 homolog that does not require hairpin loops in the direct repeat for cleavage. (6) In type II and type III systems secondary trimming is performed at either the 5’ or 3’ end to produce mature crRNAs. (7) Mature crRNAs associate with Cas proteins to form interference complexes. (8) In type I and type II systems, interactions between the protein and PAM sequence are required for degradation of invading DNA. Type III systems do not require a PAM for successful degradation and in type III-A systems basepairing occurs between the crRNA and mRNA rather than the DNA, targeted by type III-B systems.

CRISPR-Cas immunity is a natural process of bacteria and archaea. CRISPR-Cas prevents bacteriophage infection, [conjugation](/wiki/Conjugation_(genetics)), and [natural transformation](/wiki/Transformation_(genetics)#Natural_transformation) by degrading foreign nucleic acids that enter the cell.<ref name=Marraffini2015>[Template:Cite journal](/wiki/Template:Cite_journal)</ref>

### Spacer acquisition[[edit](/index.php?title=(none)&action=edit&section=9)]

When a [microbe](/wiki/Microbe) is invaded by a [virus](/wiki/Virus), the first stage of the immune response is to capture viral DNA and insert it into a CRISPR locus in the form of a spacer. [Cas1](/wiki/Cas1) and Cas2 are found in all three types of CRISPR-Cas immune systems, which indicates that they are involved in spacer acquisition. Mutation studies confirmed this hypothesis, showing that removal of [cas1](/wiki/Cas1) or cas2 stopped spacer acquisition, without affecting CRISPR immune response.[[46]](#cite_note-47)[[54]](#cite_note-55)[[55]](#cite_note-56)[[56]](#cite_note-57)[[57]](#cite_note-58) Multiple Cas1 proteins have been characterised and their structures resolved.[[58]](#cite_note-59)[[59]](#cite_note-60)[[60]](#cite_note-61) Cas1 proteins have diverse [amino acid](/wiki/Amino_acid) sequences. However, their crystal structures are similar and all purified Cas1 proteins are metal-dependent nucleases/[integrases](/wiki/Integrases) that bind to DNA in a sequence-independent manner.[[28]](#cite_note-29) Representative Cas2 proteins have been characterised and possess either ssRNA-[[61]](#cite_note-62) or dsDNA-[[62]](#cite_note-63)[[63]](#cite_note-64) specific [endoribonuclease](/wiki/Endoribonuclease) activity.

In the I-E system of *E. coli* Cas1 and Cas2 form a complex where a Cas2 dimer bridges two Cas1 dimers.[[64]](#cite_note-65) In this complex Cas2 performs a non-enzymatic scaffolding role,[[64]](#cite_note-65) binding double-stranded fragments of invading DNA, while Cas1 binds the single-stranded flanks of the DNA and catalyses their integration into CRISPR arrays.[[65]](#cite_note-66)[[66]](#cite_note-67)[[67]](#cite_note-68)

#### Protospacer adjacent motifs[[edit](/index.php?title=(none)&action=edit&section=10)]

[Template:Main article](/wiki/Template:Main_article) 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](/wiki/Protospacer_adjacent_motif) (PAM). Analysis of CRISPR-Cas systems from the three major divisions showed PAMs to be important for type I and type II, but not type III systems during acquisition.[[25]](#cite_note-26)[[68]](#cite_note-69)[[69]](#cite_note-70)[[70]](#cite_note-71)[[71]](#cite_note-72)[[72]](#cite_note-73) In type I and type II systems, protospacers are excised at positions adjacent to a PAM sequence, with the other end of the spacer cut using a ruler mechanism, thus maintaining the regularity of the spacer size in the CRISPR array.[[73]](#cite_note-74)[[74]](#cite_note-75) The conservation of the PAM sequence differs between CRISPR-Cas systems and appears to be evolutionarily linked to Cas1 and the [leader sequence](/wiki/Leader_sequence).[[72]](#cite_note-73)[[75]](#cite_note-76) New spacers are added to a CRISPR array in a directional manner,[[23]](#cite_note-24) occurring preferentially,[[42]](#cite_note-43)[[68]](#cite_note-69)[[69]](#cite_note-70)[[76]](#cite_note-77)[[77]](#cite_note-78) but not exclusively, adjacent[[71]](#cite_note-72)[[74]](#cite_note-75) to the leader sequence. Analysis of the type I-E system from *E. coli* demonstrated that the first direct repeat, adjacent to the leader sequence is copied, with the newly acquired spacer inserted between the first and second direct repeats.[[56]](#cite_note-57)[[73]](#cite_note-74) The PAM sequence appears to be important during spacer insertion in type I-E systems. That sequence contains a strongly conserved final nucleotide (nt) adjacent to the first nt of the protospacer. This nt becomes the final base in the first direct repeat.[[57]](#cite_note-58)[[78]](#cite_note-79)[[79]](#cite_note-80) This suggests that the spacer acquisition machinery generates single-stranded overhangs in the second-to-last position of the direct repeat and in the PAM during spacer insertion. However, not all CRISPR-Cas systems appear to share this mechanism as PAMs in other organisms do not show the same level of conservation in the final position.[[75]](#cite_note-76) It is likely that in those systems, a blunt end is generated at the very end of the direct repeat and the protospacer during acquisition.

#### Insertion variants[[edit](/index.php?title=(none)&action=edit&section=11)]

Analysis of [*Sulfolobus solfataricus*](/wiki/Sulfolobus_solfataricus) CRISPRs revealed further complexities to the canonical model of spacer insertion, as one of its six CRISPR loci inserted new spacers randomly throughout its CRISPR array, as opposed to inserting closest to the leader sequence.[[74]](#cite_note-75) Multiple CRISPRs contain many spacers to the same phage. The mechanism that causes this phenomenon was elucidated in the type I-E system of *E. coli*. A significant enhancement in spacer acquisition was detected where spacers already target the phage, even mismatches to the protospacer. This ‘priming’ requires the Cas proteins involved in both acquisition and interference to interact with each other. Newly acquired spacers that result from the priming mechanism are always found on the same strand as the spacer that caused the priming.[[57]](#cite_note-58)[[78]](#cite_note-79)[[79]](#cite_note-80) This observation led to the hypothesis that the acquisition machinery slides along the foreign DNA after priming to find a new protospacer.[[79]](#cite_note-80)

### Biogenesis[[edit](/index.php?title=(none)&action=edit&section=12)]

CRISPR-RNA (crRNA), which later guides the Cas nuclease to the target during the interference step, must be generated from the CRISPR sequence. The crRNA is initially transcribed as part of a single long transcript encompassing much of the CRISPR array.[[3]](#cite_note-3) This transcript is then cleaved by Cas proteins to form crRNAs. The mechanism to produce crRNAs differs among CRISPR-Cas systems. In type I-E and type I-F systems, the proteins Cas6e and Cas6f respectively, recognise stem-loops[[80]](#cite_note-81)[[81]](#cite_note-82)[[82]](#cite_note-83) created by the pairing of identical repeats which flank the crRNA.[[40]](#cite_note-41) These Cas proteins cleave the longer transcript at the edge of the paired region, leaving a single crRNA along with a small remnant of the paired repeat region.

Type III systems also use Cas6, however their repeats do not produce stem-loops. Cleavage instead occurs by the longer transcript wrapping around the Cas6 to allow cleavage just upstream of the repeat sequence.[[83]](#cite_note-84)[[84]](#cite_note-85)[[85]](#cite_note-86) Type II systems lack the Cas6 gene and instead utilize RNaseIII for cleavage. Functional type II systems encode an extra small RNA that is complementary to the repeat sequence, known as a [trans-activating crRNA](/wiki/Trans-activating_crRNA) (tracrRNA).[[54]](#cite_note-55) Transcription of the tracrRNA and the primary CRISPR transcript results in base pairing and the formation of dsRNA at the repeat sequence, which is subsequently targeted by RNaseIII to produce crRNAs. Unlike the other two systems the crRNA does not contain the full spacer but instead is truncated at one end.[[50]](#cite_note-51) CrRNAs associate with Cas proteins to form ribonucleotide complexes that recognize foreign nucleic acids. CrRNAs show no preference between the coding and non-coding strands, which is indicative of an RNA-guided DNA-targeting system.[[6]](#cite_note-6)[[46]](#cite_note-47)[[57]](#cite_note-58)[[86]](#cite_note-87)[[87]](#cite_note-88)[[88]](#cite_note-89) The type I-E complex (commonly referred to as Cascade) requires five Cas proteins bound to a single crRNA.[[89]](#cite_note-90)[[90]](#cite_note-91)

### Interference[[edit](/index.php?title=(none)&action=edit&section=13)]

During the interference stage in type I systems the PAM sequence is recognized on the crRNA-complementary strand and is required along with crRNA annealing. In type I systems correct base pairing between the crRNA and the protospacer signals a conformational change in Cascade that recruits Cas3 for DNA degradation.

Type II systems rely on a single multifunctional protein, [Cas9](/wiki/Cas9), for the interference step.[[50]](#cite_note-51) [Cas9](/wiki/Cas9) requires both the crRNA and the tracrRNA to function and cleaves DNA using its dual HNH and RuvC/RNaseH-like endonuclease domains. Basepairing between the PAM and the phage genome is also required in type II systems, however the PAM is recognized on the same strand as the crRNA (the opposite strand to type I systems).

Type III systems, like type I require six or seven Cas proteins binding to crRNAs.[[91]](#cite_note-92)[[92]](#cite_note-93) The type III systems analysed from *S. solfataricus* and *P. furiosus* both target the mRNA of phages rather than phage DNA genome,[[43]](#cite_note-44)[[92]](#cite_note-93) which may make these systems uniquely capable of targeting RNA-based phage genomes.[[28]](#cite_note-29) The mechanism for distinguishing self from foreign DNA during interference is built into the crRNAs and is therefore likely common to all three systems. Throughout the distinctive maturation process of each major type, all crRNAs contain a spacer sequence and some portion of the repeat at one or both ends. It is the partial repeat sequence that prevents the CRISPR-Cas system from targeting the chromosome as base pairing beyond the spacer sequence signals self and prevents DNA cleavage.[[93]](#cite_note-94) RNA-guided CRISPR enzymes are classified as [type V restriction enzymes](/wiki/Restriction_enzyme#Type_V).

[Template:Infobox protein family](/wiki/Template:Infobox_protein_family) [Template:Infobox protein family](/wiki/Template:Infobox_protein_family)

[Template:Infobox protein family](/wiki/Template:Infobox_protein_family) [Template:Infobox protein family](/wiki/Template:Infobox_protein_family)

## Evolution and diversity[[edit](/index.php?title=(none)&action=edit&section=14)]

The basic model of CRISPR evolution is newly incorporated spacers driving phages to mutate their genomes to avoid the bacterial immune response, creating diversity in both the phage and host populations. To fight off a phage infection, the sequence of the CRISPR spacer must correspond perfectly to the sequence of the target phage gene. Phages can continue to infect their hosts given point mutations in the spacer.[[93]](#cite_note-94) Similar stringency is required in PAM or the bacterial strain remains phage sensitive.[[69]](#cite_note-70)[[93]](#cite_note-94) A study of 124 *S. thermophilus* strains showed that 26% of all spacers were unique and that different CRISPR loci showed different rates of new spacer acquisition.[[68]](#cite_note-69) Some CRISPR loci evolve more rapidly than others, which allowed the strains' phylogenetic relationships to be determined. A [comparative genomic](/wiki/Comparative_genomics) analysis showed that *E. coli* and *S. enterica* evolve much more slowly than *S. thermophilus*. The latter's strains that diverged 250 thousand years ago still contained the same spacer complement.[[94]](#cite_note-95) [Metagenomic](/wiki/Metagenomics) analysis of two acid mine drainage [biofilms](/wiki/Biofilm) showed that one of the analyzed CRISPRs contained extensive deletions and spacer additions versus the other biofilm, suggesting a higher phage activity/prevalence in one community than the other.[[42]](#cite_note-43) In the oral cavity, a temporal study determined that 7-22% of spacers were shared over 17 months within an individual while less than 2% were shared across individuals.[[77]](#cite_note-78) From the same environment a single strain was tracked using [PCR](/wiki/PCR) primers specific to its CRISPR system. Broad-level results of spacer presence/absence showed significant diversity. However, this CRISPR added 3 spacers over 17 months,[[77]](#cite_note-78) suggesting that even in an environment with significant CRISPR diversity some loci evolve slowly.

CRISPRs were analysed from the metagenomes produced for the [human microbiome project](/wiki/Human_microbiome_project).[[95]](#cite_note-96) Although most were body-site specific, some within a body site are widely shared among individuals. One of these loci originated from [streptococcal](/wiki/Streptococcal) species and contained ~15,000 spacers, 50% of which were unique. Similar to the targeted studies of the oral cavity, some showed little evolution over time.[[95]](#cite_note-96) CRISPR evolution has been studied in [chemostats](/wiki/Chemostat) using *S. thermophilus* to directly examine spacer acquisition rates. In one week, *S. thermophilus* strains acquired up to three spacers when challenged with a single phage.[[96]](#cite_note-97) During the same interval the phage developed [single nucleotide polymorphisms](/wiki/Single_nucleotide_polymorphism) that became fixed in the population, suggesting that targeting had prevented phage replication absent these mutations.[[96]](#cite_note-97) Another *S. thermophilus* experiments showed that phages can still infect and replicate in hosts that have only one targeting spacer. Yet another showed that sensitive hosts can exist in environments with high phage titres.[[97]](#cite_note-98) The chemostat and observational studies suggest many nuances to CRISPR and phage (co)evolution.

## Identification[[edit](/index.php?title=(none)&action=edit&section=15)]

CRISPRs are widely distributed among bacteria and archaea[[48]](#cite_note-49) and show some sequence similarities.[[40]](#cite_note-41) Their most notable characteristic is their repeating spacers and direct repeats. This characteristic makes CRISPRs easily identifiable in long sequences of DNA, since the number of repeats decreases the likelihood of a false positive match. Three programs used for CRISPR repeat identification search for regularly interspaced repeats in long sequences: CRT,[[98]](#cite_note-99) PILER-CR[[99]](#cite_note-100) and CRISPRfinder.[[100]](#cite_note-101) Analysis of CRISPRs in metagenomic data is more challenging, as CRISPR loci do not typically assemble, due to their repetitive nature or through strain variation, which confuses assembly algorithms. Where many reference genomes are available, [polymerase chain reaction](/wiki/Polymerase_chain_reaction) (PCR) can be used to amplify CRISPR arrays and analyse spacer content.[[68]](#cite_note-69)[[77]](#cite_note-78)[[101]](#cite_note-102)[[102]](#cite_note-103)[[103]](#cite_note-104) However, this approach yields information only for specifically targeted CRISPRs and for organisms with sufficient representation in public databases to design reliable PCR primers.

The alternative is to extract and reconstruct CRISPR arrays from shotgun metagenomic data. This is computationally more difficult, particularly with second generation sequencing technologies (e.g. 454, Illumina), as the short read lengths prevent more than two or three repeat units appearing in a single read. CRISPR identification in raw reads has been achieved using purely denovo identification[[104]](#cite_note-105) or by using direct repeat sequences in partially assembled CRISPR arrays from [contigs](/wiki/Contig) (overlapping DNA segments that together represent a consensus region of DNA)[[95]](#cite_note-96) and direct repeat sequences from published genomes[[105]](#cite_note-106) as a hook for identifying direct repeats in individual reads.

## Evolutionary significance[[edit](/index.php?title=(none)&action=edit&section=16)]

A [bioinformatic](/wiki/Bioinformatic) study showed that CRISPRs are evolutionarily conserved and cluster into related types. Many show signs of a conserved secondary structure.[[40]](#cite_note-41) Through the CRISPR/Cas mechanism, bacteria can acquire [immunity](/wiki/Immunity_(medical)) to certain phages and thus halt further transmission of targeted phages. For this reason, [Eugene Koonin](/wiki/Eugene_Koonin) has described CRISPR/Cas as a [Lamarckian](/wiki/Lamarckism) inheritance mechanism.[[106]](#cite_note-107) However, this has been disputed by a recent critic noting "We should remember [Lamarck] for the good he contributed to science, not for things that resemble his theory only superficially. Indeed, thinking of CRISPR and other phenomena as Lamarckian only obscures the simple and elegant way evolution really works.[[107]](#cite_note-108) Analysis of CRISPR sequences revealed [coevolution](/wiki/Coevolution) of host and viral genomes.[[108]](#cite_note-109) Cas9 proteins are highly enriched in [pathogenic](/wiki/Pathogen) and [commensal](/wiki/Commensal) bacteria. CRISPR/Cas-mediated gene regulation may contribute to the regulation of endogenous bacterial genes, particularly during interaction with eukaryotic hosts. For example, [*Francisella novicida*](/wiki/Francisella_novicida) uses a unique, small, CRISPR/Cas-associated RNA (scaRNA) to repress an endogenous transcript encoding a bacterial lipoprotein that is critical for *F. novicida* to dampen host response and promote virulence.[[109]](#cite_note-110)

## Use by phages[[edit](/index.php?title=(none)&action=edit&section=17)]

Another way for bacteria to defend against phage infection is by having [chromosomal islands](/wiki/Genomic_island). A subtype of chromosomal islands called phage-inducible chromosomal island (PICI) is excised from a bacterial chromosome upon phage infection and can inhibit phage replication.[[110]](#cite_note-111) The mechanisms that induce PICI excision and how PICI inhibits phage replication are not well understood. One study showed that lytic ICP1 phage, which specifically targets [*Vibrio cholerae*](/wiki/Vibrio_cholerae) [serogroup](/wiki/Serogroup) O1, has acquired a CRISPR/Cas system that targets a *V. cholera* PICI-like element. The system has 2 CRISPR loci and 9 Cas genes. It seems to be [homologous](/wiki/Homology_(biology)) to the 1-F system found in [*Yersinia pestis*](/wiki/Yersinia_pestis). Moreover, like the bacterial CRISPR/Cas system, ICP1 CRISPR/Cas can acquire new sequences, which allows phage and host to co-evolve.[[111]](#cite_note-112)

## Applications[[edit](/index.php?title=(none)&action=edit&section=18)]

By the end of 2014 some 600 research papers had been published that mentioned CRISPR.<ref name=NatNews2015>[Template:Cite journal](/wiki/Template:Cite_journal)</ref> The technology had been used to functionally inactivate genes in human cell lines and cells, to study [*Candida albicans*](/wiki/Candida_albicans), to modify [yeasts](/wiki/Yeasts) used to make [biofuels](/wiki/Biofuels) and to [genetically modify crop](/wiki/Genetically_modified_crops) strains.<ref name=NatNews2015/>

### Genome engineering[[edit](/index.php?title=(none)&action=edit&section=19)]

CRISPR/Cas9 genome editing is carried out with a [Type II](/wiki/CRISPR#Cas_genes_and_CRISPR_subtypes) CRISPR system. When utilized for genome editing, this system includes [Cas9](/wiki/Cas9), CRISPR RNA (crRNA), trans-activating crRNA (tracrRNA) along with an optional section of DNA repair template that is utilized in either Non-Homologous End Joining (NHEJ) or Homology Directed Repair (HDR).

[thumb|500px|Overview of CRISPR Cas9 plasmid construction](/wiki/File:CRISPR_overview.jpg)[[112]](#cite_note-113)[[113]](#cite_note-114)

#### Major components[[edit](/index.php?title=(none)&action=edit&section=20)]

|  |  |
| --- | --- |
| **Component** | **Function** |
| [crRNA](/wiki/List_of_RNAs) | Contains the guide RNA that locates the correct section of host DNA along with a region that binds to [tracrRNA](/wiki/Trans-activating_crRNA) (generally in a [hairpin loop](/wiki/Stem-loop) form) forming an active complex. |
| [tracrRNA](/wiki/Trans-activating_crRNA) | Binds to [crRNA](/wiki/List_of_RNAs) and forms an active complex. |
| sgRNA | Single guide RNAs are a combined RNA consisting of a [tracrRNA](/wiki/Trans-activating_crRNA) and at least one [crRNA](/wiki/List_of_RNAs) |
| Cas9 | Protein whose active form is able to modify DNA. Many variants exist with differing functions (i.e. single strand nicking, double strand break, DNA binding) due to Cas9's DNA site recognition function. |
| Repair template | DNA that guides the cellular repair process allowing insertion of a specific DNA sequence |

CRISPR/Cas9 often employs a [plasmid](/wiki/Plasmid) to [transfect](/wiki/Transfection) the target cells. The main components of this plasmid are displayed in the image and listed in the table. The crRNA needs to be designed for each application as this is the sequence that Cas9 uses to identify and directly bind to the cell's DNA. The crRNA must bind only where editing is desired. The repair template is designed for each application, as it must overlap with the sequences on either side of the cut and code for the insertion sequence.

Multiple crRNA's and the tracrRNA can be packaged together to form a single-guide RNA (sgRNA). This sgRNA can be joined together with the Cas9 gene and made into a plasmid in order to be transfected into cells (see image for overview).

[thumb|500px|overview of the transfection and DNA cleaving by CRISPR Cas9 (crRNA and tracrRNA are often joined as one strand of RNA when designing a plasmid)](/wiki/File:CRISPR_transfection.png)[[114]](#cite_note-115)

#### Structure[[edit](/index.php?title=(none)&action=edit&section=21)]

CRISPR/Cas9 offers a high degree of fidelity and relatively simple construction. It depends on two factors for its specificity – the target sequence and the PAM. The target sequence is 20 bases long as part of each CRISPR locus in the crRNA array.[[114]](#cite_note-115) A typical crRNA array has multiple unique target sequences. Cas9 proteins select the correct location on the host's genome by utilizing the sequence to bond with base pairs on the host DNA. The sequence is not part of the Cas9 protein and as a result is customizable and can be independently [synthesized](/wiki/Oligonucleotide_synthesis).[[115]](#cite_note-116)[[116]](#cite_note-117) The PAM sequence on the host genome is recognized by the protein structure of Cas9 and generally cannot be easily modified to recognize a different sequence. However this is not too limiting as it is a short sequence and nonspecific (e.g. the SpCas9 PAM sequence is 5'-NGG-3' and in the human genome occurs roughly every 8 to 12 base pairs).[[114]](#cite_note-115) Once these have been assembled into a plasmid and transfected into cells the Cas9 protein with help of the crRNA finds the correct sequence in the host cell's DNA and – depending on the Cas9 variant – creates a single or double strand break in the DNA.

Properly spaced single strand breaks in the host DNA can trigger [homology directed repair](/wiki/Homology_directed_repair), which is less error prone than non-homologous end joining that typically follows a double strand break. Providing a DNA repair template allows for the insertion of a specific DNA sequence at an exact location within the genome. The repair template should extend 40 to 90 base pairs beyond the Cas9 induced DNA break.[[114]](#cite_note-115) The goal is for the cell's HDR process to utilize the provided repair template and thereby incorporate the new sequence into the genome. Once incorporated, this new sequence is now part of the cell's genetic material and passes into its daughter cells.

Many online tools are available to aid in designing effective sgRNA sequences.[[117]](#cite_note-118)

#### Delivery[[edit](/index.php?title=(none)&action=edit&section=22)]

Scientists can use viral or non-viral system for delivery of the Cas9 and gRNA into target cells. [Electroporation](/wiki/Electroporation) of DNA, RNA or ribonucleocomplexes is the most common and cheapest system. This technique was used to edit [CXCR4](/wiki/CXCR4) and [PD-1](/wiki/Programmed_cell_death_protein_1), knocking in new sequences to replace specific genetic “letters” in these proteins. The group was then able to sort the cells, using cell surface markers, to help identify successfully edited cells.[[118]](#cite_note-119) Deep sequencing of a target site confirmed that knock-in genome modifications had occurred with up to ∼20% efficiency, which accounted for up to approximately one-third of total editing events.[[119]](#cite_note-120) However, hard-to-transfect cells (stem cells, neurons, hematopoietic cells, etc.) require more efficient delivery systems such as those based on [lentivirus](/wiki/Lentivirus) (LVs), [adenovirus](/wiki/Adenoviridae) (AdV) and [adeno-associated virus](/wiki/Adeno-associated_virus) (AAV).

### Editing[[edit](/index.php?title=(none)&action=edit&section=23)]

CRISPRs have been used to cut five<ref name=craze/> to 62 genes at once: pig cells have been engineered to inactivate all 62 [Porcine Endogenous Retrovirus](/wiki/Porcine_endogenous_retroviruses) in the pig genome, which eliminated infection from the pig to human cells in culture.[[120]](#cite_note-121) CRISPR's low cost compared to alternatives is widely seen as revolutionary.[[8]](#cite_note-9)[[9]](#cite_note-10) Selective engineered redirection of the CRISPR/Cas system was first demonstrated in 2012 in:<ref name=halemajumder2012>[Template:Cite journal](/wiki/Template:Cite_journal)</ref>[[121]](#cite_note-122)

* Immunization of industrially important bacteria, including some used in food production and large-scale fermentation
* Cellular or organism RNA-guided [genome engineering](/wiki/Genome_engineering). Proof of concept studies demonstrated examples both [*in vitro*](/wiki/In_vitro)[[122]](#cite_note-123)[[50]](#cite_note-51) and [*in vivo*](/wiki/In_vivo)[[35]](#cite_note-36)[[123]](#cite_note-124)[[124]](#cite_note-125)\* Bacterial strain discrimination by comparison of spacer sequences

### Knockdown/activation[[edit](/index.php?title=(none)&action=edit&section=24)]

[Template:Main article](/wiki/Template:Main_article)

Using “dead” versions of Cas9 ([dCas9](/wiki/Cas9#Interference_of_transcription_by_dCas9)) eliminates CRISPR’s DNA-cutting ability, while preserving its ability to target desirable sequences. Multiple groups added various regulatory factors to dCas9s, enabling them to turn almost any gene on or off or subtly adjust its level of activity.[[125]](#cite_note-126) Like RNAi, CRISPR interference (CRISPRi) turns off genes in a reversible fashion by targeting, but not cutting a site. The targeted site is methylated, [epigenetically](/wiki/Epigenetics) modifying the gene. This modification inhibits transcription. Cas9 is an effective way of targeting and silencing specific genes at the DNA level.[[126]](#cite_note-127) In bacteria, the presence of Cas9 alone is enough to block transcription. For mammalian applications, a section of protein is added. Its guide RNA targets regulatory DNA sequences called [promoters](/wiki/Promoter_(genetics)) that immediately precede the target gene.<ref name=craze/>

Cas9 was used to carry synthetic [transcription factors](/wiki/Transcription_factor) that activated specific human genes. The technique achieved a strong effect by targeting multiple CRISPR constructs to slightly different locations on the gene's promoter.<ref name=craze/>

### RNA editing[[edit](/index.php?title=(none)&action=edit&section=25)]

In 2016 researchers demonstrated that CRISPR from an ordinary mouth bacterium could be used to edit [RNA](/wiki/RNA). The researchers searched databases containing hundreds of millions of genetic sequences for those that resembled Crispr genes. With Zhang they considered the [fusobacteria](/wiki/Fusobacteria) *Leptotrichia shahii*. It had a group of genes that resembled CRISPR genes, but with important differences. When the researchers equipped other bacteria with these genes, which they called C2c2, they found that the organisms gained a novel defense.[[127]](#cite_note-128) Many viruses encode their genetic information in RNA rather than DNA that they repurpose to make new viruses. HIV and poliovirus are such viruses. Bacteria with C2c2 make molecules that can dismember RNA, destroying the virus. Tailoring these genes opened any RNA molecule to editing.[[127]](#cite_note-128)

### Disease models[[edit](/index.php?title=(none)&action=edit&section=26)]

CRISPR simplifies creation of [animals for research](/wiki/Genetically_modified_organism#Mammals) that mimic disease or show what happens when a gene is [knocked down](/wiki/Gene_knockdown) or mutated. CRISPR may be used at the [germline](/wiki/Germline) level to create animals where the gene is changed everywhere, or it may be targeted at non-germline cells.[[128]](#cite_note-129)[[129]](#cite_note-130)[[130]](#cite_note-131) CRISPR can be utilized to create human cellular models of disease. For instance, CRISPR was applied to human [pluripotent stem cells](/wiki/Cell_potency) to introduce targeted mutations in genes relevant to [polycystic kidney disease](/wiki/Polycystic_kidney_disease) (PKD) and [focal segmental glomerulosclerosis](/wiki/Focal_segmental_glomerulosclerosis) (FSG).[[131]](#cite_note-132) These CRISPR-modified pluripotent stem cells were subsequently grown into human kidney [organoids](/wiki/Organoid) that exhibited disease-specific phenotypes. Kidney organoids from stem cells with PKD populations formed large, translucent cyst structures from kidney tubules. Kidney organoids with mutations in a gene linked to FSG developed junctional defects between [podocytes](/wiki/Podocyte), the filtering cells affected in that disease. Importantly, these disease phenotypes were absent in control organoids of identical genetic background, but lacking the CRISPR modifications.[[131]](#cite_note-132) A similar approach has been taken to model long QT syndrome in [cardiomyocytes](/wiki/Cardiac_muscle_cell) derived from pluripotent stem cells.[[132]](#cite_note-133) These CRISPR-generated cellular models, with isogenic controls, provide a new way to study human diseases and test drugs.

### Gene drive[[edit](/index.php?title=(none)&action=edit&section=27)]

In 2003 evolutionary biologist Austin Burt envisioned attaching a gene that coded for a desired trait to “selfish” DNA elements that could copy themselves from one chromosome position to another. That would bias daughter cells to inherit it, quickly spreading it throughout a population. In 2015 a U.S. team used CRISPR to create a “mutagenic chain reaction” that drove a pigmentation trait in lab-grown [Drosophila](/wiki/Drosophila) to the next generation with 97% efficiency. With another research group they created a gene drive in mosquitoes that spread genes that prevented the insects from harboring [malaria](/wiki/Malaria) [parasites](/wiki/Parasites). Only weeks later, the team reported a second drive with genes that rendered female mosquitoes infertile and could quickly wipe out a population. The work was done in the lab, stimulating debates over the desirability of field testing.[[125]](#cite_note-126)

### Biomedicine[[edit](/index.php?title=(none)&action=edit&section=28)]

CRISPR/Cas-based "RNA-guided nucleases" can be used to target [virulence factors](/wiki/Virulence_factors), genes encoding [antibiotic resistance](/wiki/Antibiotic_resistance), and other medically relevant sequences of interest. This technology thus represents a novel form of antimicrobial therapy and a strategy by which to manipulate bacterial populations.[[133]](#cite_note-134) Some of the affected genes are tied to human diseases, including those involved in muscle differentiation, cancer, inflammation and fetal [hemoglobin](/wiki/Hemoglobin).[[27]](#cite_note-28) Clinical researchers are applying it to develop tissue-based treatments for cancer and other diseases.[[125]](#cite_note-126) CRISPR may revive the concept of [transplanting](/wiki/Organ_transplantation) animal organs into people. [Retroviruses](/wiki/Retroviruses) present in animal genomes could harm transplant recipients. In 2015 a team eliminated 62 copies of a retrovirus’s DNA from the pig genome.[[125]](#cite_note-126) It may also have applications in tissue engineering and regenerative medicine, such as by creating human blood vessels that lack expression of [MHC class II](/wiki/MHC_class_II) proteins, which often cause transplant rejection.[[134]](#cite_note-135)

### Gene function[[edit](/index.php?title=(none)&action=edit&section=29)]

In 2015, multiple studies attempted to systematically disable each individual human gene, in an attempt to identify which genes were essential to human biology. Between 1,600 and 1,800 genes passed this test—of the 20,000 or so known human genes. Such genes are more strongly activated, and unlikely to carry disabling mutations. They are more likely to have indispensable counterparts in other species. They build proteins that unite to form larger collaborative complexes. The studies also catalogued the essential genes in four cancer-cell lines and identified genes that are expendable in healthy cells, but crucial in specific tumor types and drugs that could target these rogue genes.[[135]](#cite_note-136) The specific functions of some 18 percent of the essential genes are unidentified. In one 2015 targeting experiment, disabling individual genes in groups of cells attempted to identify those involved in resistance to a [melanoma](/wiki/Melanoma) drug. Each such gene manipulation is itself a separate "drug", potentially opening the entire genome to CRISPR-based regulation.[[125]](#cite_note-126)

### ''In vitro'' genetic depletion[[edit](/index.php?title=(none)&action=edit&section=30)]

Unenriched sequencing libraries often have abundant undesired sequences. Cas9 can specifically deplete the undesired sequences with double strand breakage with up to 99% efficiency and without significant off-target effects as seen with restriction enzymes. Treatment with Cas9 can deplete abundant rRNA while increasing pathogen sensitivity in RNA-seq libraries.[[136]](#cite_note-137)

## Patents and commercialization[[edit](/index.php?title=(none)&action=edit&section=31)]

As of December 2014, [patent rights](/wiki/Patent_rights) to CRISPR were contested. Several companies had formed to develop related drugs and research tools.[[137]](#cite_note-138) As companies ramp up financing, doubts as to whether or not CRISPR can be quickly monetized were raised.[[138]](#cite_note-139) As of November 2013, SAGE Labs (now part of [Horizon Discovery](/wiki/Horizon_Discovery) group) had [exclusive rights](/wiki/Exclusive_right) from one of those companies to produce and sell genetically engineered rats and non-exclusive rights for mouse and rabbit models.[[139]](#cite_note-140) [Template:As of](/wiki/Template:As_of), [Thermo Fisher Scientific](/wiki/Thermo_Fisher_Scientific) had licensed intellectual properly from ToolGen to develop CRISPR reagent kits.[[140]](#cite_note-141)

## Society and culture[[edit](/index.php?title=(none)&action=edit&section=32)]

### Human germline modification[[edit](/index.php?title=(none)&action=edit&section=33)]

At least four labs in the US, labs in China and the UK, and by a US biotechnology company called Ovascience announced plans or ongoing research to apply CRISPR to human embryos.[[141]](#cite_note-142) Scientists, including a CRISPR co-inventor, urged a worldwide moratorium on applying CRISPR to the human germline, especially for clinical use. They said "scientists should avoid even attempting, in lax jurisdictions, germline genome modification for clinical application in humans" until the full implications "are discussed among scientific and governmental organizations".[[37]](#cite_note-38)[[142]](#cite_note-143) These scientists support basic research on CRISPR and do not see CRISPR as developed enough for any clinical use in making heritable changes to people.[[143]](#cite_note-144) In April 2015, Chinese scientists reported results of an attempt to alter the [DNA](/wiki/DNA) of non-viable [human embryos](/wiki/Human_embryos) using CRISPR to correct a [mutation](/wiki/Mutation) that causes [beta thalassemia](/wiki/Beta_thalassemia), a lethal heritable disorder.[[144]](#cite_note-145)[[145]](#cite_note-146) The study had previously been rejected by both [*Nature*](/wiki/Nature_(journal)) and [*Science*](/wiki/Science_(journal)) in part because of ethical concerns; the journals had no comment.<ref name=NatureNews>[Template:Cite web](/wiki/Template:Cite_web)</ref> The experiments resulted in changing only some genes, and had off-target effects on other genes. The researchers stated that CRISPR is not ready for clinical application in [reproductive medicine](/wiki/Reproductive_medicine)."<ref name=NatureNews/> In April 2016 Chinese scientists were reported to have made a second unsuccessful attempt to alter the DNA of non-viable human embryos using CRISPR - this time to alter the [CCR5](/wiki/CCR5) gene to make the embryo HIV resistant.[[146]](#cite_note-147) In December 2015, the International Summit on Human Gene Editing took place in Washington under the guidance of [David Baltimore](/wiki/David_Baltimore). Members of national scientific academies of America, Britain and China discussed the ethics of germline modification. They agreed to support basic and clinical research under appropriate legal and ethical guidelines. A specific distinction was made between clinical use in [somatic cells](/wiki/Somatic_cells), where the effects of edits are limited to a single individual, versus germline cells, where genome changes could be inherited by future generations. This could have unintended and far-reaching and potentially damaging consequences for human evolution, genetically (e.g. gene/environment interactions) and culturally (e.g. [Social Darwinism](/wiki/Social_Darwinism)). Altering of [gametocytes](/wiki/Gametocytes) and embryos to generate inheritable changes in humans was thus claimed irresponsible. In addition, they agreed to initiate an international forum to address such concerns and harmonize regulations countries.[[147]](#cite_note-148) In February 2016, British scientists were given permission by regulators to genetically modify [human embryos](/wiki/Human_embryos) by using CRISPR-Cas9 and related techniques. The embryos were to be destroyed after seven days.[[148]](#cite_note-149)[[149]](#cite_note-150)

### Recognition[[edit](/index.php?title=(none)&action=edit&section=34)]

In 2012 and 2013, CRISPR was a runner-up in [*Science Magazine*](/wiki/Science_Magazine)'s [Breakthrough of the Year](/wiki/Breakthrough_of_the_Year) award. In 2015, it was the winner of that award.[[125]](#cite_note-126) CRISPR was named as one of [MIT Technology Review's](/wiki/MIT_Technology_Review) 10 breakthroughs technologies in 2014 and 2016.[[150]](#cite_note-151)[[151]](#cite_note-152)

## Alternative cutters[[edit](/index.php?title=(none)&action=edit&section=35)]

<gallery class="center" title = "Gallery of secondary structure images"> Image:RF01315.png|**CRISPR-DR2**: Secondary structure taken from the [Rfam](http://rfam.sanger.ac.uk/) database. Family [RF01315](http://rfam.sanger.ac.uk/family/RF01315). Image:RF01318.png|**CRISPR-DR5**: Secondary structure taken from the [Rfam](http://rfam.sanger.ac.uk/) database. Family [RF011318](http://rfam.sanger.ac.uk/family/RF01318). Image:RF01319.png|**CRISPR-DR6**: Secondary structure taken from the [Rfam](http://rfam.sanger.ac.uk/) database. Family [RF01319](http://rfam.sanger.ac.uk/family/RF01319). Image:RF01321.png|**CRISPR-DR8**: Secondary structure taken from the [Rfam](http://rfam.sanger.ac.uk/) database. Family [RF01321](http://rfam.sanger.ac.uk/family/RF01321). Image:RF01322.png|**CRISPR-DR9**: Secondary structure taken from the [Rfam](http://rfam.sanger.ac.uk/) database. Family [RF01322](http://rfam.sanger.ac.uk/family/RF01322). Image:RF01332.png|**CRISPR-DR19**: Secondary structure taken from the [Rfam](http://rfam.sanger.ac.uk/) database. Family [RF01332](http://rfam.sanger.ac.uk/family/RF01332). Image:RF01350.png|**CRISPR-DR41**: Secondary structure taken from the [Rfam](http://rfam.sanger.ac.uk/) database. Family [RF01350](http://rfam.sanger.ac.uk/family/RF01350). Image:RF01365.png|**CRISPR-DR52**: Secondary structure taken from the [Rfam](http://rfam.sanger.ac.uk/) database. Family [RF01365](http://rfam.sanger.ac.uk/family/RF01365). Image:RF01370.png|**CRISPR-DR57**: Secondary structure taken from the [Rfam](http://rfam.sanger.ac.uk/) database. Family [RF01370](http://rfam.sanger.ac.uk/family/RF01370). Image:RF01378.png|**CRISPR-DR65**: Secondary structure taken from the [Rfam](http://rfam.sanger.ac.uk/) database. Family [RF01378](http://rfam.sanger.ac.uk/family/RF01378). </gallery>

## See also[[edit](/index.php?title=(none)&action=edit&section=36)]

[Template:Div col](/wiki/Template:Div_col)

* [Cas9](/wiki/Cas9)
* [CRISPR/Cas Tools](/wiki/CRISPR/Cas_Tools)
* [CRISPR/Cpf1](/wiki/CRISPR/Cpf1)
* [CRISPR interference](/wiki/CRISPR_interference)
* [Gene drive](/wiki/Gene_drive)
* [Gene knockout](/wiki/Gene_knockout)
* [Gene silencing](/wiki/Gene_silencing)
* [Protospacer adjacent motif](/wiki/Protospacer_adjacent_motif)
* [RNAi](/wiki/RNAi)
* [SiRNA](/wiki/SiRNA)
* [Transcription activator-like effector nuclease](/wiki/Transcription_activator-like_effector_nuclease) (TALEN)
* [Zinc finger nuclease](/wiki/Zinc_finger_nuclease)
* [Bioterrorism](/wiki/Bioterrorism)
* [Biological warfare](/wiki/Biological_warfare)
* [Synthetic Biology](/wiki/Synthetic_Biology)
* [Surveyor Nuclease Assay](/wiki/Surveyor_Nuclease_Assay)

[Template:Div col end](/wiki/Template:Div_col_end)

## Notes[[edit](/index.php?title=(none)&action=edit&section=37)]

[Template:Reflist](/wiki/Template:Reflist)

[[45]](#cite_note-46)

## References[[edit](/index.php?title=(none)&action=edit&section=38)]

[Template:Reflist](/wiki/Template:Reflist)

## Further reading[[edit](/index.php?title=(none)&action=edit&section=39)]

[Template:Refbegin](/wiki/Template:Refbegin)

* [CRISPR-Cas: A Laboratory Manual](http://www.cshlpress.com/default.tpl?cart=14520192576517890&action=full&--eqskudatarq=1073) Edited by Jennifer Doudna, University of California, Berkeley; Prashant Mali, University of California, San Diego
* [Template:Cite journal](/wiki/Template:Cite_journal)
* [Template:Cite journal](/wiki/Template:Cite_journal)
* [Template:Cite journal](/wiki/Template:Cite_journal)
* [Template:Cite journal](/wiki/Template:Cite_journal)
* [Template:Cite journal](/wiki/Template:Cite_journal)
* [Template:Cite journal](/wiki/Template:Cite_journal)
* [Template:Cite journal](/wiki/Template:Cite_journal)
* [Template:Cite journal](/wiki/Template:Cite_journal)
* [Template:Cite journal](/wiki/Template:Cite_journal)
* [Template:Cite journal](/wiki/Template:Cite_journal)
* [Template:Cite journal](/wiki/Template:Cite_journal)
* [Template:Cite journal](/wiki/Template:Cite_journal)
* [Template:Cite journal](/wiki/Template:Cite_journal)
* [Template:Cite news](/wiki/Template:Cite_news)
* [Template:Cite journal](/wiki/Template:Cite_journal) [Template:Closed access](/wiki/Template:Closed_access)

[Template:Refend](/wiki/Template:Refend) [Template:Breakthrough of the Year](/wiki/Template:Breakthrough_of_the_Year)

[Template:DEFAULTSORT:Crispr](/wiki/Template:DEFAULTSORT:Crispr) [Category:Repetitive DNA sequences](/wiki/Category:Repetitive_DNA_sequences) [Category:Non-coding RNA](/wiki/Category:Non-coding_RNA)