Genome editing

From Wikipedia, the free encyclopedia

Genome editing, or **genome editing with engineered nucleases** (**GEEN**) is a type of genetic engineering in which DNA is inserted, replaced, or removed from a genome using artificially engineered nucleases, or "molecular scissors." The nucleases create specific double-stranded break (DSBs) at desired locations in the genome, and harness the cell's endogenous mechanisms to repair the induced break by natural processes of homologous recombination (HR) and nonhomologous end-joining (NHEJ). There are currently four families of engineered nucleases being used: Zinc finger nucleases (ZFNs), Transcription Activator-Like Effector Nucleases (TALENs), the CRISPR/Cas system, and engineered meganuclease re-engineered homing endonucleases.[1][2][3]

It is commonly practiced in genetic analysis that in order to understand the function of a gene or a protein function one interferes with it in a sequence-specific way and monitors its effects on the organism. However, in some organisms it is difficult or impossible to perform site-specific mutagenesis, and therefore more indirect methods have to be used, such as silencing the gene of interest by short RNA interference (siRNA) .[4] Yet gene disruption by siRNA can be variable and incomplete. Genome editing with nucleases such as ZFN is different from siRNA in that the engineered nuclease is able to modify DNA-binding specificity and therefore can in principle cut any targeted position in the genome, and introduce modification of the endogenous sequences for genes that are impossible to specifically target by conventional RNAi. Furthermore, the specificity of ZFNs and TALENs are enhanced as two ZFNs are required in the recognition of their portion of the target and subsequently direct to the neighboring sequences.

It was chosen by Nature Methods as the 2011 Method of the Year.[5]

Contents

Concept

A common approach in the modern biological research is to manipulate the genetic sequence (genotype) of an organism (or a single cell) and observe the impact of this change on the organism (phenotype). Such approach is called reverse genetics and its significance for the modern biology lies in its relative simplicity. In contrast, in forward genetics a new phenotype is first observed and then its genetic basis is studied. This course is more complex, since phenotypic changes are often a result of multiple genetic interactions.

Among the key aspects of reverse genetic analysis is the ability to modify the genetic code. This can be achieved by:

- site-directed mutagenesis which employs polymerase chain reaction (PCR) with primers containing the desired mutation. *(In which organisms is it used? Just bacteria?)*
- recombination based methods that utilize the natural ability of cells to exchange DNA between its own genetic information and an exogenous DNA. These methods have been made possible in yeast and mice.

Both approaches have several drawbacks:

- They are less successful in other organisms.
- They also require stringent selection steps and thus addition of selection specific sequences, along with those incorporated into the DNA.
- They can be quite inefficient - e.g. in mouse embryonic stem cells treated with donor DNA, in only 1 of a million the DNA got incorporated at the desired position.[6]

Use of other techniques such as P-element transgenesis in Drosophila also have their limitations, the major one being the randomness of incorporation and the possibility of affecting other genes and expression patterns.

Hence, genomic editing with engineered nucleases, a rapidly growing technology is a promising new approach. It overcomes these shortcomings and uses relatively simple concepts.

Double stranded breaks and their repair

First and foremost in understanding the use of nucleases in genome editing is the understanding of DNA double stranded break (DSB) repair mechanisms. Two of the known DSB repair pathways that are essentially functional in all organisms are the non-homologous end joining (NHEJ) and homology directed repair (HDR).

NHEJ uses a variety of enzymes to directly join the DNA ends in a double-strand break. In contrast, in HDR, a homologous sequence is utilized as a template for regeneration of missing DNA sequence at the break point. The natural properties of these pathways form the very basis of nucleases based genome editing.

NHEJ is error prone such that it was shown to cause mutations at the repair site in approximately 50% of DSB in mycobacteria [7] and also its low fidelity has been linked to mutational accumulation in leukemias.[8] Thus if one is able to create a DSB at a desired gene in multiple samples, it is very likely that mutations will be generated at that site in some of the treatments because of errors created by the NHEJ infidelity.

On the other hand, the dependency of HDR on a homologous sequence to repair DSBs can be exploited by inserting a desired sequence within a sequence that is homologous to the flanking sequences of a DSB which, when used as a template by HDR system, would lead to the creation of the desired change within the genomic region of interest.

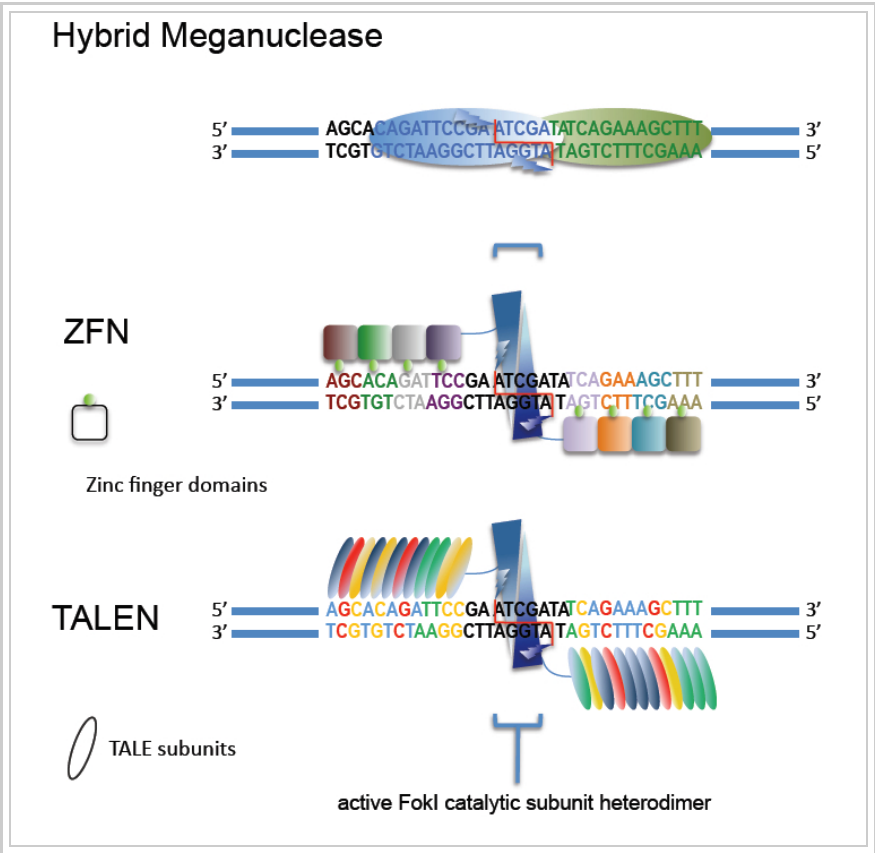
Despite the distinct mechanisms, the concept of the HDR based gene editing is in a way similar to that of homologous recombination based gene targeting. However, the rate of recombination is increased by at least three orders of magnitude when DSBs are created and HDR is at work thus making the HDR based recombination much more efficient and eliminating the need for stringent positive and negative selection steps.[9] So based on these principles if one is able to create a DSB at a specific location within the genome, then the cell's own repair systems will help in creating the desired mutations.

Site-specific double stranded breaks

Creation of a DSB in DNA should not be a challenging task as the commonly used restriction enzymes are capable of doing so. However, if genomic DNA is treated with a particular restriction endonuclease many DSBs will be created. This is a result of the fact that most restriction enzymes recognize a few base pairs on the DNA as their target and very likely that particular base pair combination will be found in many locations across the genome. To overcome this challenge and create site-specific DSB, three distinct classes of nucleases have been discovered and bioengineered to date. These are the Zinc finger nucleases (ZFNs), transcription-activator like effector nucleases (TALENs) and meganucleases. Below is a brief overview and comparison of these enzymes and the concept behind their development.

Current engineered nucleases

Meganucleases, found commonly in microbial species, have the unique property of having very long recognition sequences (>14bp) thus making them naturally very specific.[10][11] This can be exploited to make site-specific DSB in genome editing; however, the challenge is that not enough meganucleases are known, or may ever be known, to cover all possible target sequences. To overcome this challenge, mutagenesis and high throughput screening methods have been used to create meganuclease variants that recognize unique sequences.[11] Others have been able to fuse various meganucleases and create hybrid enzymes that recognize a new sequence.[12] Yet others have attempted to alter the DNA interacting aminoacids of the meganuclease to design sequence specific meganucelases in a method named rationally designed meganuclease (US Patent 8,021,867 B2).



The current groups of engineered nucleases used for GEEN. Matching colors signify DNA recognition patterns

Meganuclease have the benefit of causing less toxicity in cells compared to methods such as ZFNs likely because of more stringent DNA sequence recognition;[11] however, the construction of sequence specific enzymes for all possible sequences is costly and time consuming as one is not benefiting from combinatorial possibilities that methods such as ZFNs and TALENs utilize. So there are both advantages and disadvantages.

As opposed to meganucleases, the concept behind ZFNs and TALENs is more based on a non-specific DNA cutting enzyme which would then be linked to specific DNA sequence recognizing peptides such as zinc fingers and transcription activator-like effectors (TALEs).[13] The key to this was to find an endonuclease whose DNA recognition site and cleaving site were separate from each other, a situation that is not common among restriction enzymes.[13] Once this enzyme was found, its cleaving portion could be separated which would be very non-specific as it would have no recognition ability. This portion could then be linked to sequence recognizing peptides that could lead to very high specificity. A restriction enzyme with such properties is FokI. Additionally FokI has the advantage of requiring dimerization to have nuclease activity and this means the specificity increases dramatically as each nuclease partner would recognize a unique DNA sequence. To enhance this effect, FokI nucleases have been engineered that can only function as heterodimers and have increased catalytic activity.[14] The heterodimer functioning nucleases would avoid the possibility of unwanted homodimer activity and thus increase specificity of the DSB. Although the nuclease portion of both ZFNs and TALENs have similar properties, the difference between these engineered nucleases is in their DNA recognition peptide. ZFNs rely on Cys2-His2 zinc fingers and TALENs on TALEs. Both of these DNA recognizing peptide domains have the characteristic that they are naturally found in combinations in their proteins. Cys2-His2 Zinc fingers typically happen in repeats that are 3 bp apart and are found in diverse combinations in a variety of nucleic acid interacting proteins such as transcription factors. TALEs on the other hand are found in repeats with a one-to-one recognition ratio between the amino acids and the recognized nucleotide pairs. Because both zinc fingers and TALEs happen in repeated patterns, different combinations can be tried to create a wide variety of sequence specificities.[10] Zinc fingers have been more established in these terms and approaches such as modular assembly (where Zinc fingers correlated with a triplet sequence are attached in a row to cover the required sequence), OPEN (low-stringency selection of peptide domains vs. triplet nucleotides followed by high-stringency selections of peptide combination vs. the final target in bacterial systems), and bacterial one-hybrid screening of zinc finger libraries among other methods have been used to make site specific nucleases.

Applications

Over the past decade, efficient genome editing has been developed for a wide range of experimental systems ranging from plants to animals, often beyond clinical interest, and the method holds a promising future in becoming a standard experimental strategy in research labs.[15] The recent generation of rat, zebrafish, maize and tobacco ZFN-mediated mutants testifies to the significance of the methods and the list is expanding rapidly. Genome editing with engineered nucleases will likely contribute to many fields of life sciences from studying gene functions in plants and animals to gene therapy in humans. For instance, the field of synthetic biology which aims to engineer cells and organisms to perform novel functions, is likely to benefit from the ability of engineered nuclease to add or remove genomic elements and therefore create complex systems.[15] In addition, gene functions can be studied using stem cells with engineered nucleases.

Listed below are some specific tasks this method can carry out:

- Targeted gene mutation
- Creating chromosome rearrangement
- Study gene function with stem cells
- Transgenic animals
- Endogenous gene labeling
- Targeted transgene addition

Targeted gene addition in plants

Genome editing using ZFN provides a new strategy for genetic manipulation in plants and is likely to assist engineering desired plant traits by modifying endogenous genes. For instance, site-specific gene addition in major crop species can be used for 'trait stacking' whereby several desired traits are physically linked to ensure their co-segregation during the breeding processes.[14] Progress in such cases have been recently reported in *Arabidopsis thaliana* [16][17][18] and *Zea mays*. In *Arabidopsis thaliana*, using ZFN-assisted gene targeting, two herbicide-resistant genes (tobacco acetolactate synthase SuRA and SuRB) were introduced to SuR loci with as high as 2% transformed cells with mutations.[19] In Zea mays, disruption of the target locus was achieved by ZFN-induced DSBs and the resulting NHEJ. ZFN was also used to drive herbicide-tolerance gene expression cassette (PAT) into the targeted endogenous locus IPK1 in this case.[20] Such genome modification observed in the regenerated plants has been shown to be inheritable and was transmitted to the next generation.[20]

Organisms	Genes	Methods of ZFN development
Gene disruption		
Fruitflies	*yellow, rosy, brown*	Modular assembly
Zebrafish	*kdr*	Bacteria one-hybrid
	golden, no tail	Two-finger modules
	trf2, dat, telomerase	OPEN
Human T cells	CCR5	Two-finger modules
Rats	*Rab38, IgM, Il2rg*	Two-finger modules
Gene correction		
Tobacco	*SuRA, SuRB*	OPEN
Arabidopsis thaliana	*AB14, KU80*	Modular assembly
	TT4, ADH1	OPEN
Fruitflies	*yellow, rosy, coilin, pask*	Modular assembly
Human T cells	*IL2RG*	Two-finger modules
Gene addition		
Tobacco	*Chitinase*	Two-finger modules
Zea mays	*Ipk1, Zein protein 15*	Two-finger modules
Human ES cells	*IL2RG, CCR5*	Two-finger modules
	PIGA	OPEN

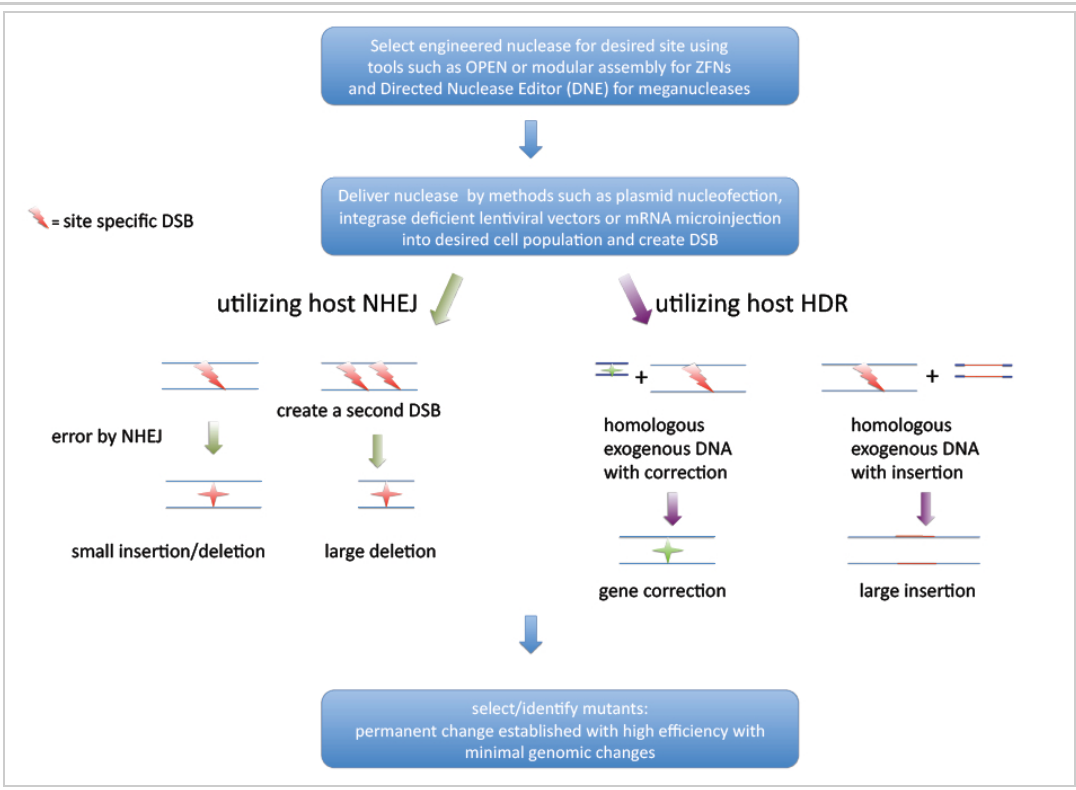
Several optimizations need to be made in order to improve editing plant genomes using ZFN-mediated targeting.[21] These include the reliable design and subsequent test of the nucleases, the absence of toxicity of the nucleases, the appropriate choice of the plant tissue for targeting, the routes of introduction or induction of enzyme activity, the lack of off-target mutagenesis, and a reliable detection of mutated cases.[21]

Gene therapy

The ideal gene therapy practice is that which replaces the defective gene with a normal allele at its natural location. This is advantageous over a virally delivered gene as there is no need to include the full coding sequences and regulatory sequences when only a small proportions of the gene needs to be altered as is often the case.[22] The expression of the partially replaced genes is also more consistent with

normal cell biology than full genes that are carried by viral vectors.

ZFN-induced targeting can also attack defective genes at their endogenous chromosomal locations. Examples include the treatment of X-linked severe combined immunodeficiency (X-SCID) by ex vivo gene correction with DNA carrying the interleukin-2 receptor common gamma chain (IL-2Rγ) with the correct sequence.[23] Insertional mutagenesis by the retroviral vector genome induced leukemia in some patients, a problem predicted to be avoided by GEEN and ZFNs. However, ZFNs may also cause off-target mutations, in a different way from viral transductions. Currently many measures are taken to improve off-target detection and ensure safety before treatment.



Overview of GEEN workflow and editing possibilities

Recently, Sangamo BioSciences (SGMO) introduced the Delta 32 mutation (a suppressor of CCR5 gene which is a co-receptor for HIV-1 entry into T cells therefore enabling HIV infection) using Zinc Finger Nuclease (ZFN). Their results were presented at the 51st Interscience Conference on Antimicrobial Agents and Chemotherapy (ICAAC) held in Chicago from September 17–20, 2011.[24] Researchers at SGMO mutated CCR5 in CD4+ T cells and subsequently produced an HIV-resistant T-cell population.[25]

Prospect and Current Limitations

In the future, research into genome editing with engineered nucleases needs to focus on improving safety and specificity of the nucleases. For example, improving the ability to detect off-target events can improve our ability to learn about ways of preventing them. In addition, zinc-fingers used in ZFNs are seldom completely specific, and some may cause toxicity. However, the toxicity has been reported to be reduced by modifications done on the cleavage domain of the ZFN.[22]

In addition, the study by Dana Carroll looking at modifying the genome with engineered nucleases shows the requirement of better understanding of the basic recombination and repair machinery of DNA. In the future, a possible method to identify secondary targets would be to capture broken ends from cells expressing the ZFNs and to sequence the flanking DNA using high-throughput sequencing.[22]

Genome editing occurs also as a natural process without artificial genetic engineering. The agents that are competent to edit genetic codes are viruses or subviral RNA-agents.[26]

Lastly, one must also realize that although GEEN has higher efficiency than many other methods in reverse genetics, it is still not highly efficient as in many cases less than half of the treated populations obtain the desired changes.[19] For example, when one is planning to use the cell's NHEJ to create a mutation, the cell's HDR systems will also be at work correcting the DSB with lower mutational rates.

See also

- Genome engineering

Further reading

- E-TALEN.org (http://www.e-talen.org) Design genome editing constructs using TALENs or CRISPR
- Targeted Genome Editing Across Species Using ZFNs and TALENs (http://www.sciencemag.org/content/333/6040/307.full)
- http://investor.sangamo.com/releasedetail.cfm?ReleaseID=606148

References

1. **^** Esvelt, KM.; Wang, HH. (2013). "Genome-scale engineering for systems and synthetic biology." (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3564264). *Mol Syst Biol* **9** (1): 641. doi:10.1038/msb.2012.66 (http://dx.doi.org/10.1038%2Fmsb.2012.66). PMC 3564264 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3564264). PMID 23340847 (https://www.ncbi.nlm.nih.gov/pubmed/23340847).
2. **^** Tan, WS.; Carlson, DF.; Walton, MW.; Fahrenkrug, SC.; Hackett, PB. (2012). "Precision editing of large animal genomes." (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3683964). *Adv Genet* **80**: 37–97. doi:10.1016/B978-0-12-404742-6.00002-8 (http://dx.doi.org/10.1016%2FB978-0-12-404742-6.00002-8). PMC 3683964 (https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3683964). PMID 23084873 (https://www.ncbi.nlm.nih.gov/pubmed/23084873).
3. **^** Puchta, H.; Fauser, F. (2013). "Gene targeting in plants: 25 years later.". *Int. J. Dev. Biol* **57**: 629–637. doi:10.1387/ijdb.130194hp (http://dx.doi.org/10.1387%2Fijdb.130194hp).
4. **^** *Fire, A. et al., Potent and specific genetic interference by double-stranded RNA in* Caenorhabditis elegans" Nature *391 (6669), 806-811 (1998).*
5. **^** *Method of the Year 2011. Nat Meth 9 (1), 1-1.*
6. **^** *Capecchi, M., Altering the genome by homologous recombination*" Science *244 (4910), 1288-1292 (1989).*
7. **^** *Gong, C. et al., Mechanism of nonhomologous end-joining in mycobacteria: a low-fidelity repair system driven by Ku, ligase D and ligase C. Nat Struct Mol Biol 12 (4), 304-312 (2005).*
8. **^** *Feyruz Virgilia, R., DNA double strand breaks (DSB) and non-homologous end joining (NHEJ) pathways in human leukemia. Cancer Letters 193 (1), 1-9 (2003).*
9. **^** *Maria, J., Genetic manipulation of genomes with rare-cutting endonucleases*" Trends in Genetics *12 (6), 224-228 (1996).*

10. ^ **a** **b** *de Souza, N., Primer: genome editing with engineered nucleases. Nat Meth 9 (1), 27-27 (2011).*
11. ^ **a** **b** **c** Smith, J. et al., A combinatorial approach to create artificial homing endonucleases cleaving chosen sequences" *Nucleic Acids Research* 34 (22), e149 (2006).
12. **^** *Chevalier, B.S. et al., Design, Activity, and Structure of a Highly Specific Artificial Endonuclease*" Molecular Cell *10 (4), 895-905 (2002).*
13. ^ **a** **b** Baker, M., Gene-editing nucleases. Nat Meth 9 (1), 23-26 (2012).
14. ^ **a** **b** Urnov, F.D., Rebar, E.J., Holmes, M.C., Zhang, H.S., & Gregory, P.D., Genome editing with engineered zinc finger nucleases" *Nat Rev Genet* 11 (9), 636-646 (2010).
15. ^ **a** **b** McMahon, M.A., Rahdar, M., & Porteus, M., Gene editing: not just for translation anymore. Nat Meth 9 (1), 28-31 (2012).
16. **^** *Townsend, J.A. et al., High-frequency modification of plant genes using engineered zinc-finger nucleases*" Nature *459 (7245), 442-445 (2009).*
17. **^** *Zhang, F. et al., High frequency targeted mutagenesis in Arabidopsis thaliana using zinc finger nucleases*" Proceedings of the National Academy of Sciences *107 (26), 12028-12033 (2009).*
18. **^** *Osakabe, K., Osakabe, Y., & Toki, S., Site-directed mutagenesis in Arabidopsis using custom-designed zinc finger nucleases*" Proceedings of the National Academy of Sciences *107 (26), 12034-12039 (2010).*
19. ^ **a** **b** Townsend, J.A. et al., High-frequency modification of plant genes using engineered zinc-finger nucleases" *Nature* 459 (7245), 442-445 (2009).
20. ^ **a** **b** Shukla, V.K. et al., Precise genome modification in the crop species Zea mays using zinc-finger nucleases" *Nature* 459 (7245), 437-U156 (2009).
21. ^ **a** **b** Puchta, H. & Hohn, B., Breaking news: Plants mutate right on target" *Proceedings of the National Academy of Sciences* 107 (26), 11657-11658 (2010).
22. ^ **a** **b** **c** Carroll, D., Progress and prospects: Zinc-finger nucleases as gene therapy agents. Gene Ther 15 (22), 1463-1468 (2008).
23. **^** *Lombardo, A. et al., Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery. Nat Biotech 25 (11), 1298-1306 (2007).*
24. **^** *Sangamo BioSciences Announces Presentation of Groundbreaking Clinical Data From ZFN Therapeutic for HIV/AIDS at ICAAC 2011 (2011).*
25. **^** *Perez, E.E. et al., Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases. Nat Biotech 26 (7), 808-816 (2008).*
26. **^** Witzany, G. The agents of natural genome editing. J Mol Cell Biol 3: 181-189 (2011). PMID: 21459884 doi:10.1093/jmcb/mjr005 (http://dx.doi.org/10.1093%2Fjmcb%2Fmjr005)

Categories: Genetic engineering
