

Editace genomů

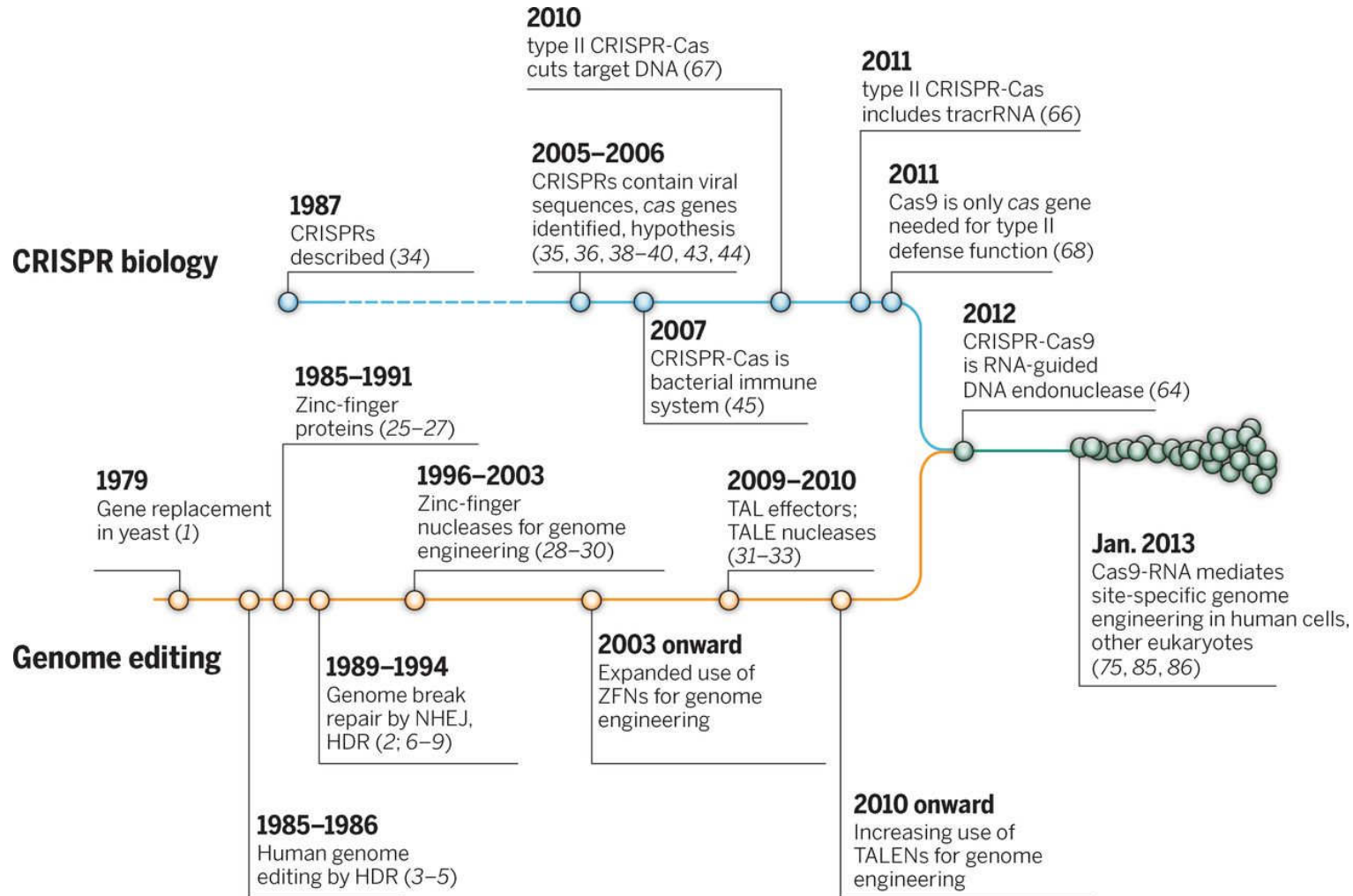
Genome editing, or genome editing with engineered nucleases (GEEN)

Postupy genového inženýrství, při nichž se do vybraného místa v cílové DNA pomocí uměle připravených nukleáz (tzv. molekulárních nůžek) vnáší inserce, delece a nebo se stávající sekvence nahrazuje za jiné (náhrada alel).

Tyto nukleázy vytvářejí na určených místech genomu dvouřetězcové zlomy (DSBs: double-stranded breaks), čímž vyvolávají přirozené endogenní buněčné procesy vedoucí k reparaci zlomů:

- a) Homologní rekombinací (HR)
- b) Nehomologní spojování volných konců (NHEJ: nonhomologous end-joining)

Historické etapy v CRISPR biologii a editování genomů



Genome Editing Glossary

Cas = CRISPR-associated genes

Cas9, Csn1 = a CRISPR-associated protein containing two nuclease domains, that is programmed by small RNAs to cleave DNA

crRNA = CRISPR RNA

dCAS9 = nuclease-deficient Cas9

DSB = Double-Stranded Break

gRNA = guide RNA

HDR = Homology-Directed Repair

HNH = an endonuclease domain named for characteristic histidine and asparagine residues

Indel = insertion and/or deletion

NHEJ = Non-Homologous End Joining

PAM = Protospacer-Adjacent Motif

RuvC = an endonuclease domain named for an *E. coli* protein involved in DNA repair

sgRNA = single guide RNA

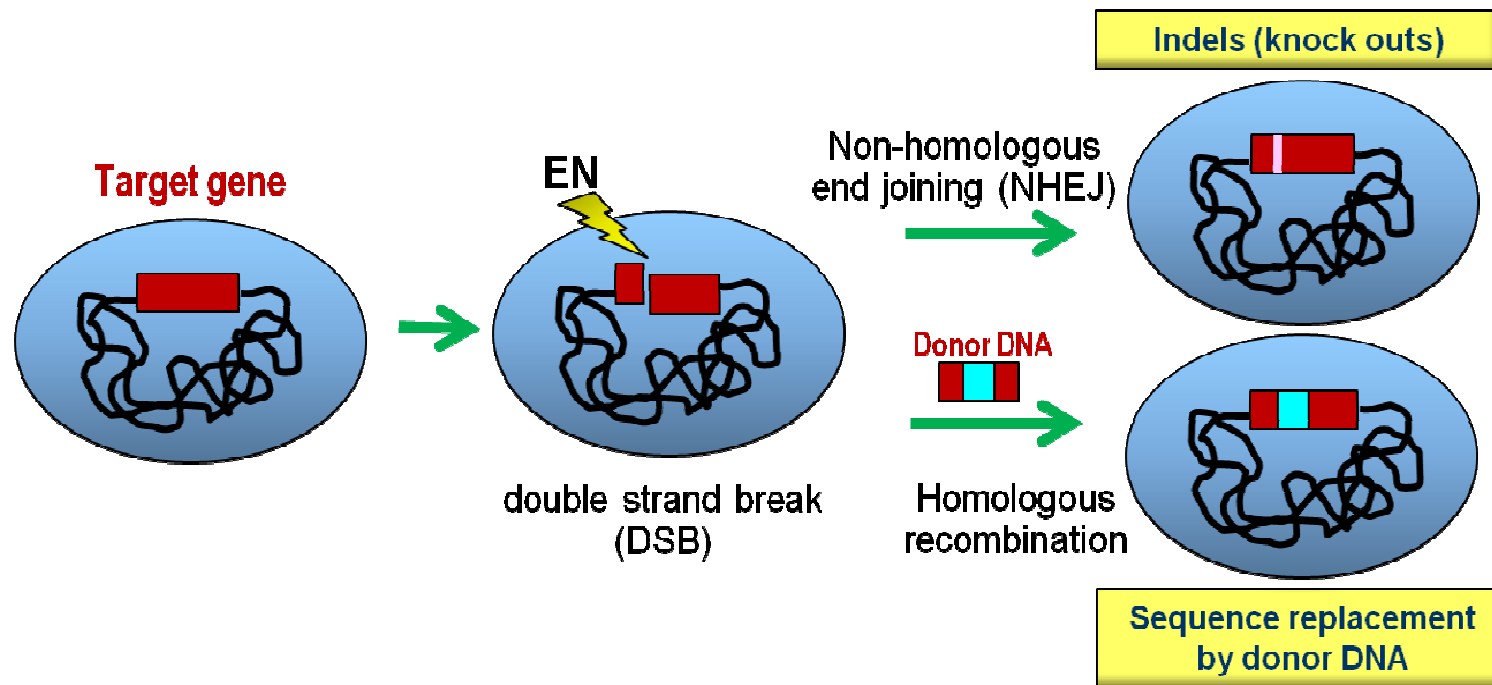
tracrRNA, trRNA = trans-activating crRNA

TALEN = Transcription-Activator Like Effector Nuclease

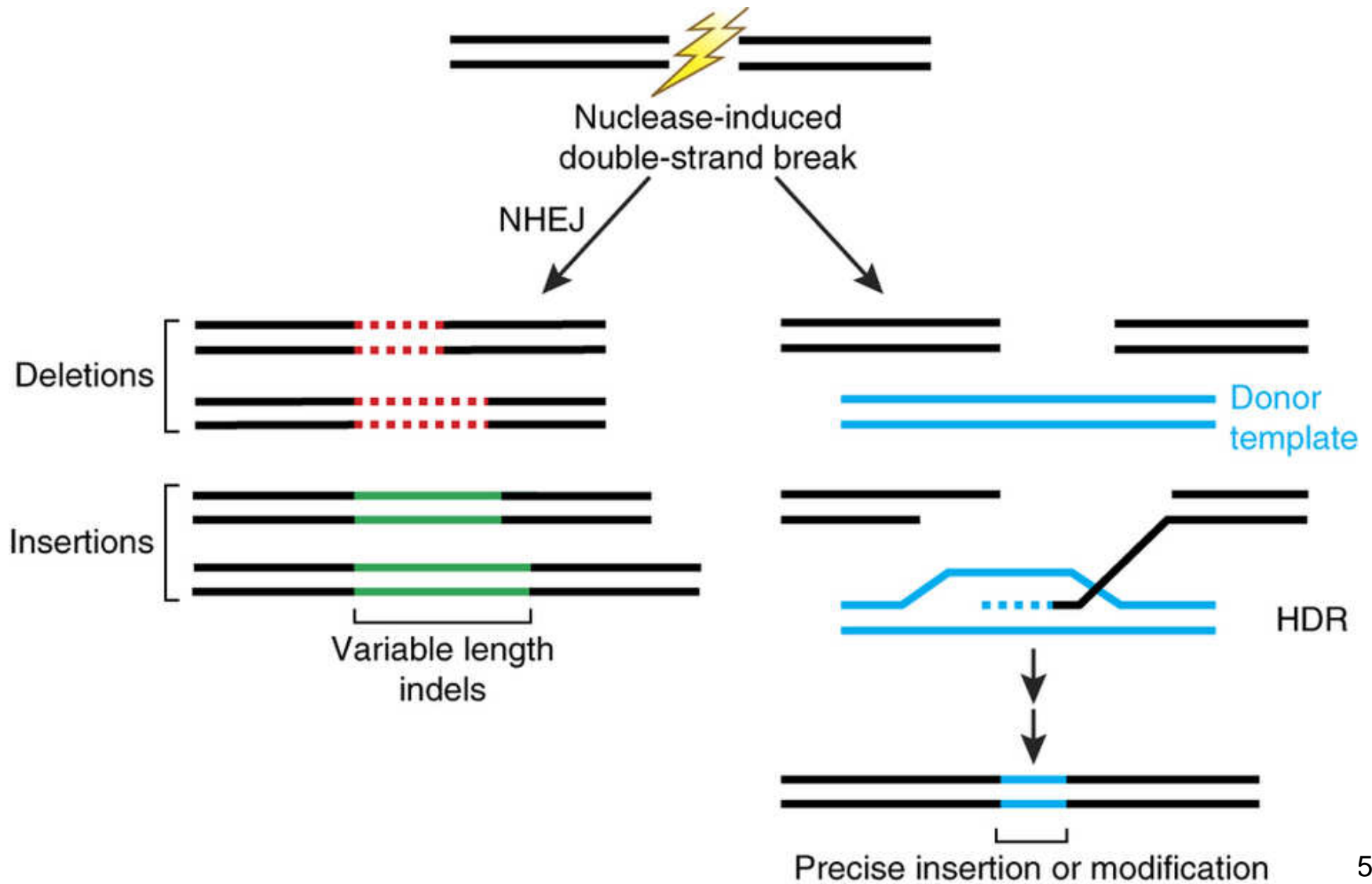
ZFN = Zinc-Finger Nuclease

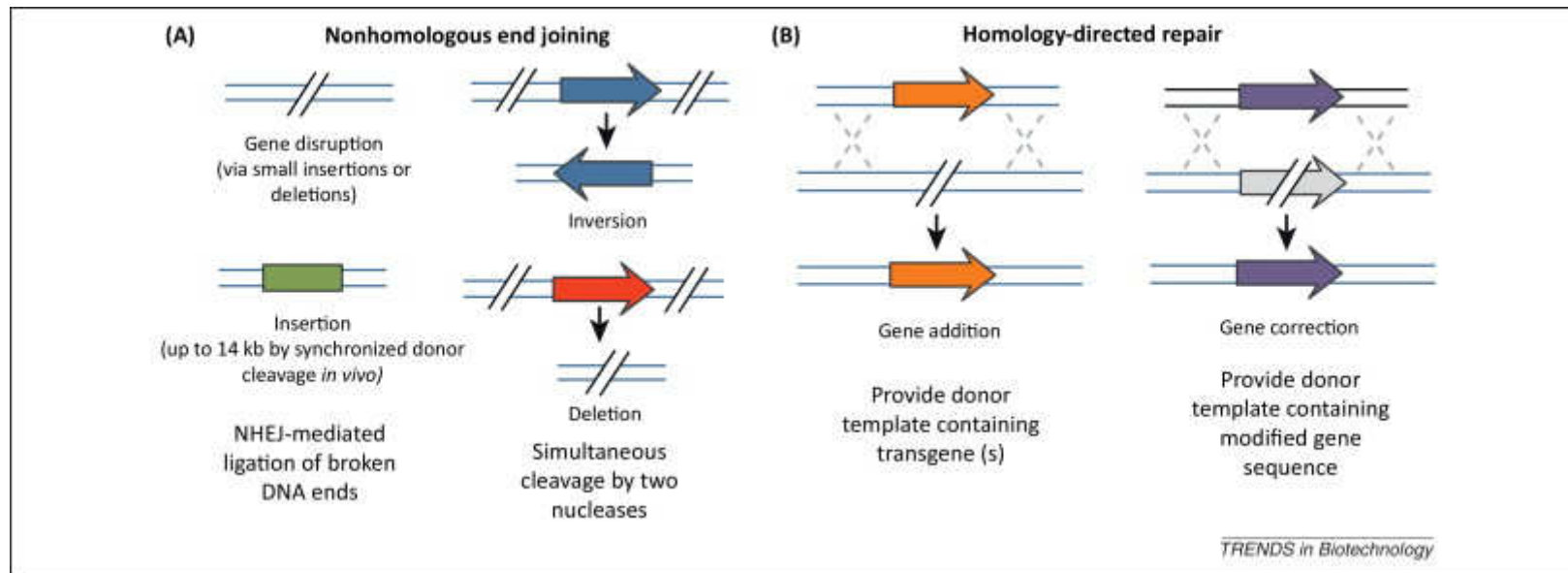
(RGENs) = RNA-guided engineered nucleases

Procesy probíhající po vytvoření DSB umělými nukleázami



Procesy probíhající po vytvoření DSB umělými nukleázami





Overview of possible genome editing outcomes using site-specific nucleases. Nuclease-induced DNA double-strand breaks (DSBs) can be repaired by homology-directed repair (HDR) or error-prone nonhomologous end joining (NHEJ). **(A)** In the presence of donor plasmid with extended homology arms, HDR can lead to the introduction of single or multiple transgenes to correct or replace existing genes. **(B)** In the absence of donor plasmid, NHEJ-mediated repair yields small insertion or deletion mutations at the target that cause gene disruption. In the presence of double-stranded oligonucleotides or *in vivo* linearized donor plasmid, DNA fragments up to 14 kb have been inserted via NHEJ-mediated ligation. Simultaneous induction of two DSBs can lead to deletions, inversions and translocations of the intervening segment.

Typy nukleáz používané pro editaci genomů

Hybrid Meganuclease



ZFN



Zinc finger domains

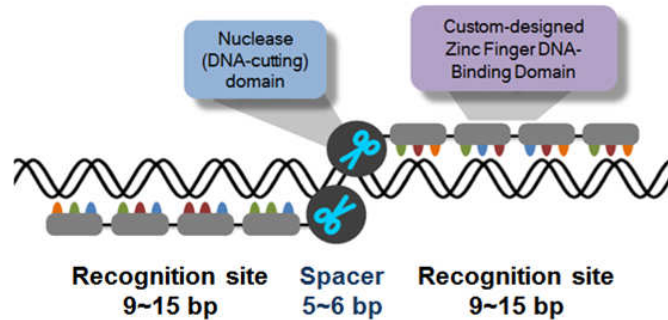
TALEN



TALE subunits

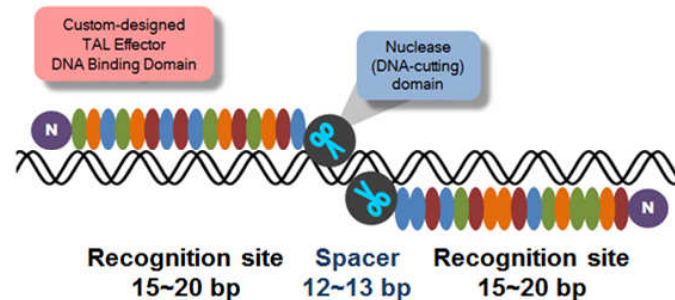
active FokI catalytic subunit heterodimer

Typy nukleáz používané pro editaci genomů



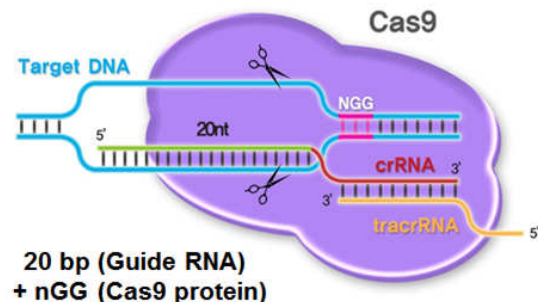
Zinc Finger Nucleases (ZFNs)

- Represent the first generation of engineered nucleases
- DNA binding module: Zinc fingers (each module recognizes 3 bp of target sequence)
- DNA cleavage domain: Fok I restriction enzyme nuclease domain (requires dimerization for cleavage)
- Widely proven in many cells and organisms
- Relatively lower resolution of target sequence programmability
- Relatively lower specificity



TAL Effector Nucleases (TALENs)

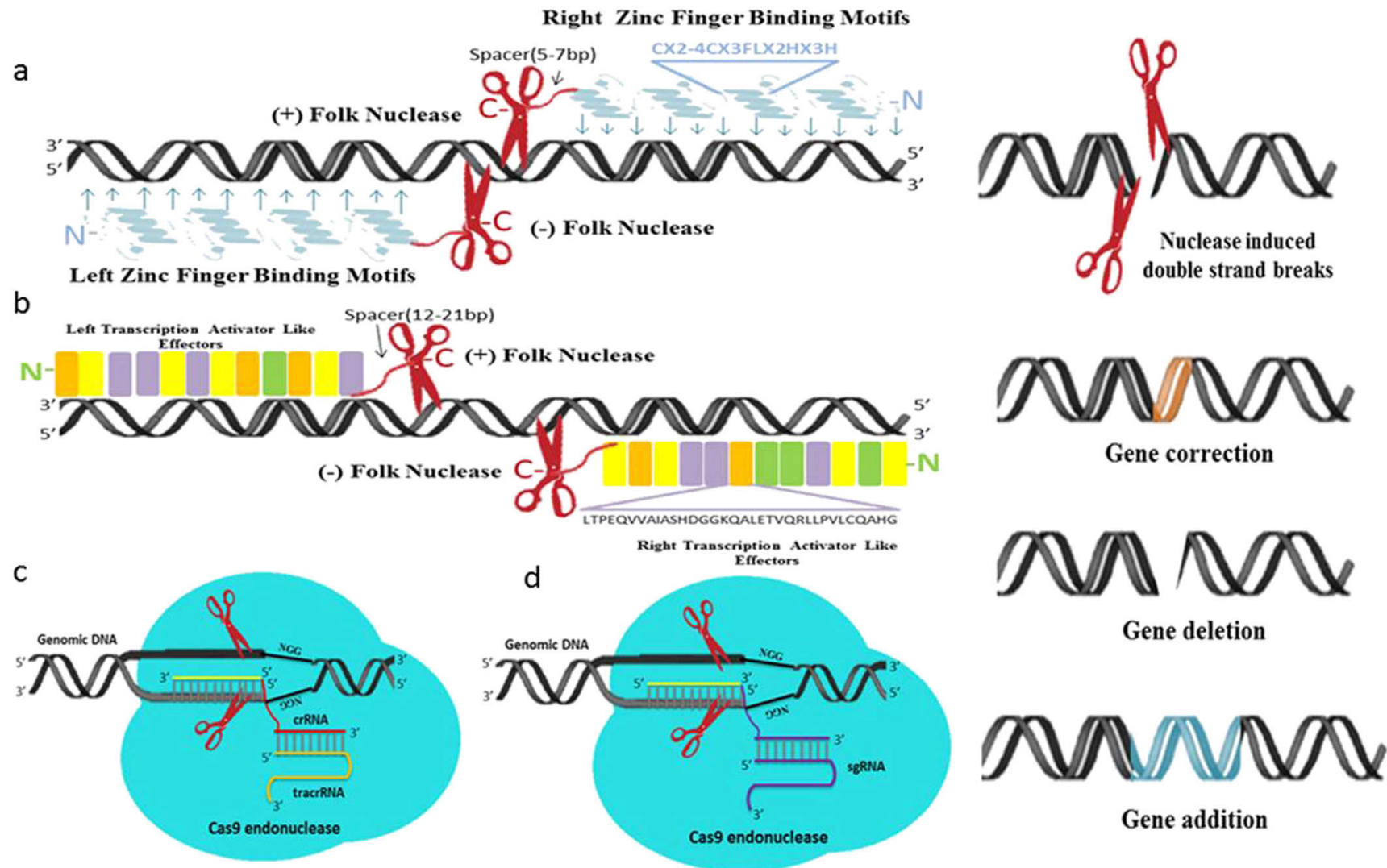
- DNA binding module: TAL effector unit (each module recognizes 1 bp of target sequence)
- DNA cleavage domain: Fok I restriction enzyme nuclease domain (requires dimerization for cleavage)
- High resolution of target sequence programmability
- High specificity and low toxicity



RNA-Guided Endonucleases (RGENs)

- DNA binding module: Guide RNA that hybridizes to the target DNA (1:1 nucleotide base pairing)
- DNA cleavage module: Cas9 protein (contains two nuclease domains)
- High resolution of target sequence programmability
- High specificity and low toxicity

Nukleázy používané pro editaci genomů



MEGANUKLEÁZY

Meganukleázy se vyskytují u různých druhů mikroorganismů, mají velmi dlouhé rozpoznávací sekvence (>14bp) a jsou tak přirozeně sekvenčně velmi specifické.

Nevýhodou je, že je jich známo relativně málo, a tudíž počet cílových sekvencí je omezen.

Mutagenézí byly uměle připraveny varianty meganukleáz, které rozpoznávají další jedinečné sekvence. Byly připraveny rovněž hybridní varianty meganukleáz fúzí dvou s odlišnými cílovými místy.

Byl použit též postup záměn aminokyselin v doménách interagujících s DNA a docíleno vysoké specifiky rozpoznání (*method named rationally designed meganuclease (US Patent 8,021,867 B2)*).

Meganukleázy jsou méně toxické pro buňky než ZNF díky tomu, že mají vyšší specifitu/stringenci rozpoznání cílových sekvencí DNA.

Jejich navrhování je však časově náročné.

LAGLIDADG family of homing endonucleases (kódovány introny nebo inteiny), často v mitochondriích a chloroplastech

I-SceI 18-bp z mitochondrií *Saccharomyces cerevisiae*

I-CreI chloroplasty *Chlamydomonas reinhardtii*

I-Dmol archea: *Desulfurococcus mobilis*.

NUKLEÁZA FokI

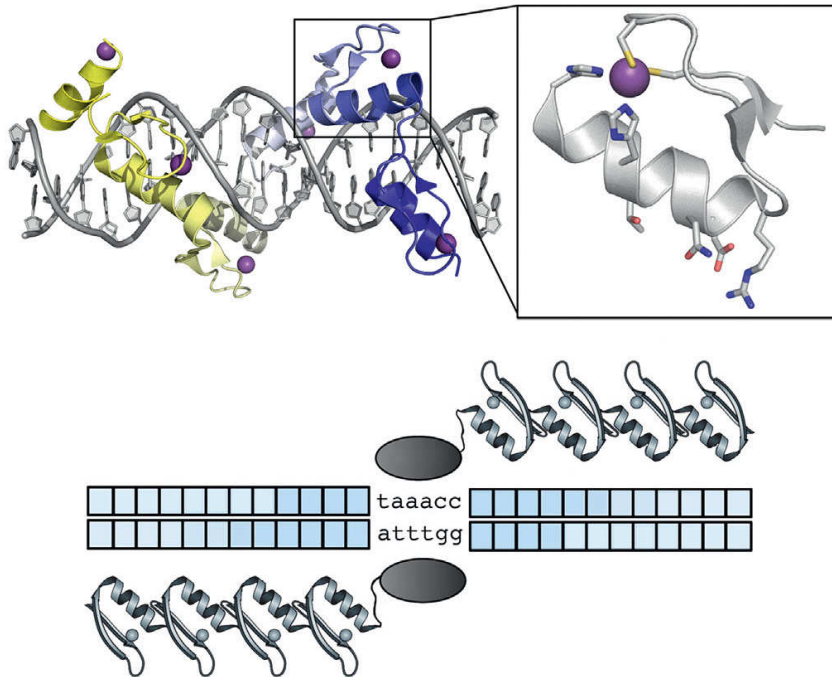
Enzym **FokI** přirozeně se vyskytující u *Flavobacterium okeanoikoites* je restriční endonukleáza typu IIS. Je tvořena N-terminální vazebnou doménou a nespecificky štěpící doménou na C-konci.

Výhody FokI pro její využití v GI:

- Rozpoznávaná sekvence je oddělena od sekvence, která je štěpena – to umožňuje izolovat doménu enzymu, která štěpí sekvenci nespecificky. Tato doména pak může být spojena s doménou zodpovědnou za rozpoznání cílové sekvence.
- FokI vyžaduje pro svou nukleázovou činnost dimerizaci – zvýší se tím specifita rozpoznání cílového místa.
- Byly připraveny modifikované FokI, které fungují jen jako heterodimery, což zvyšuje specifitu rozpoznání cílových sekvencí a eliminuje možnost vytváření nespecifického štěpení v případě homodimerů.

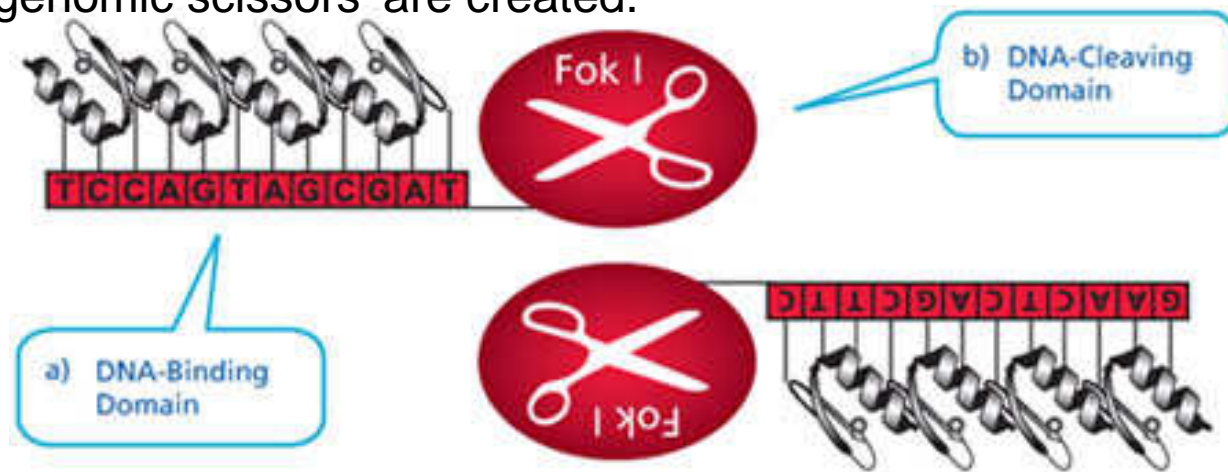
Struktura proteinu obsahujícího zinkové prsty (zinc finger)

Každý zinkový prst sestává asi ze 30 AA v konformaci $\beta\beta\alpha$. Každý prst kontaktuje 3 nebo 4 bp ve velkém žlábků DNA.



Dimer Zinc-finger nukleázy (ZNF) navázaný na DNA. Cílová místa pro vazbu ZNF sestávají ze dvou vazebných míst pro zinkové prsty, která jsou oddělena 5-7 bp dlouhou sekvencí, která je rozpoznávána štěpící doménou FokI.

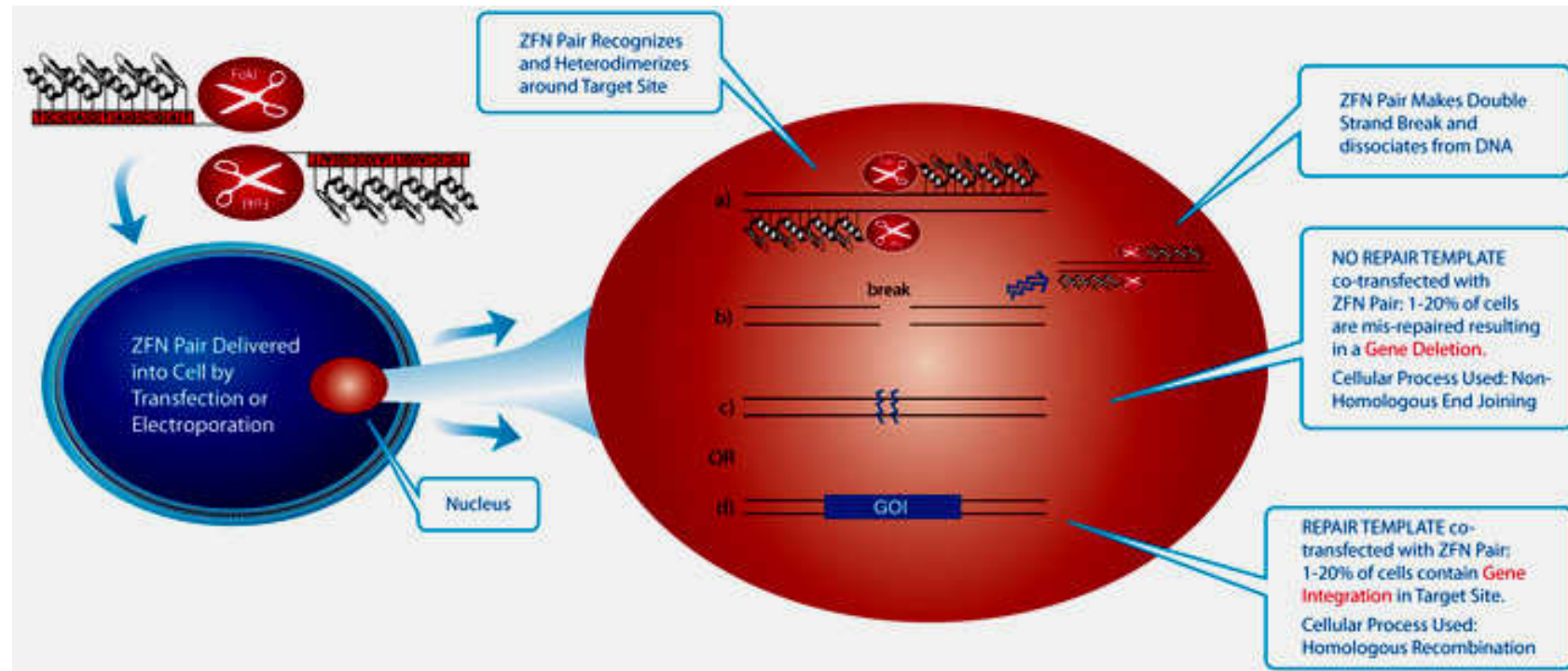
Figure 1: Each Zinc Finger Nuclease (ZFN) consists of two functional domains: a.) A DNA-binding domain comprised of a chain of two-finger modules, each recognizing a unique hexamer (6 bp) sequence of DNA. Two-finger modules are stitched together to form a Zinc Finger Protein, each with specificity of ≥ 24 bp. b.) A DNA-cleaving domain comprised of the nuclease domain of Fok I. When the DNA-binding and DNA-cleaving domains are fused together, a highly-specific pair of 'genomic scissors' are created.

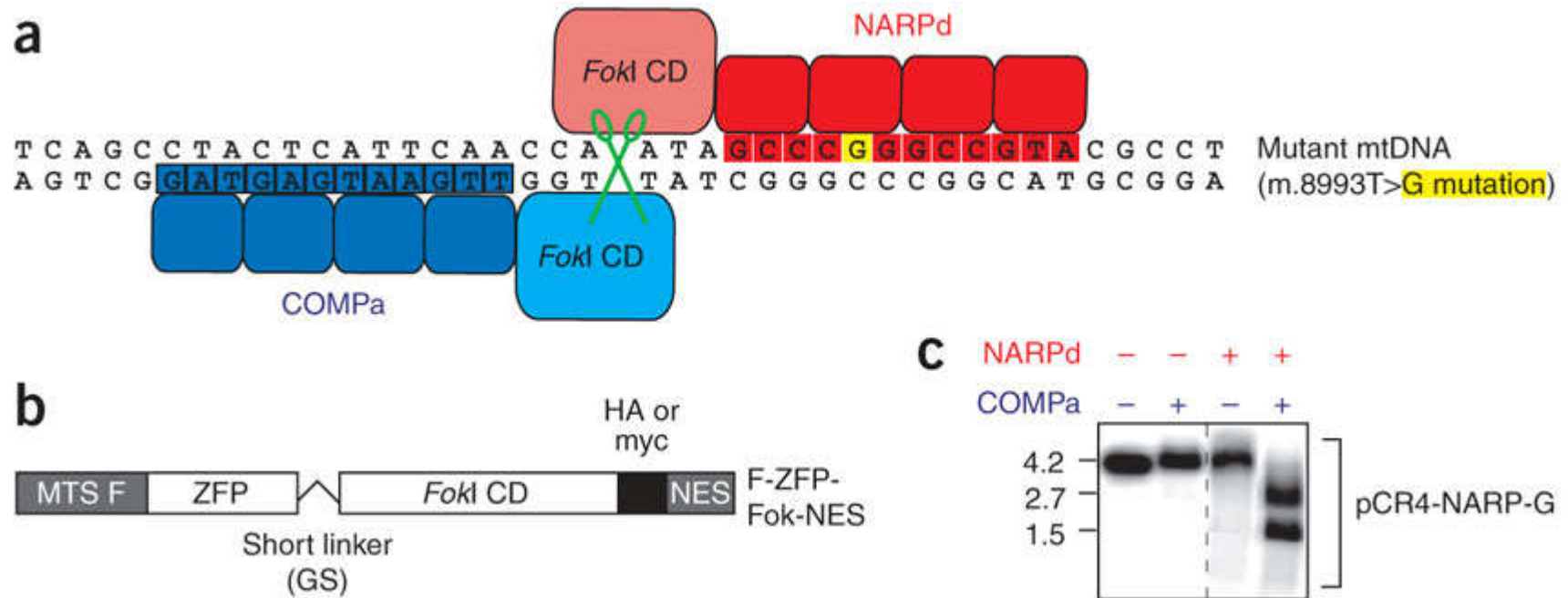


BENEFITS

- Rapid disruption of, or integration into, any genomic loci Mutations made are permanent and heritable
- Works in a variety of mammalian somatic cell types
- Edits induced through a single transfection experiment
- Knockout or knock-in cell lines in as little as two months
- Single or biallelic edits occur in 1–20% of clone population
- No antibiotic selection required for screening

Editace genomů pomocí ZNF





(a) Schematic diagram of a ZFN heterodimer bound to the mutated mtDNA target. Each of the monomeric ZFN consists of the *FokI* nuclease domain (*FokI* CD) linked to a zinc-finger peptide. One of the ZFNs (NARPd, red) was designed to bind to the mutated mtDNA site, whereas its companion ZFN binds a native sequence on the opposite DNA strand (COMPa, blue)¹⁹. (b) Schematic structure of mtZFN used in the cleavage assay in c. 'mitochondrial targeting sequence (MTS) F' denotes the mitochondrial targeting sequence of F1 β -subunit of the human mitochondrial ATP synthase (see ANTICIPATED RESULTS for details). (c) *In vitro* assay testing the specificity of the F–NARPd–Fok–NES and F–COMPa–Fok–NES constructs (see ANTICIPATED RESULTS for details).

TAL effector (Transcription Activator-like effectors)

Gram negativní bakterie r. *Xanthomonas* infikují řadu rostlinných druhů, u nichž mohou způsobovat onemocnění. Injikují svými sekrečními systémy do rostlinných buněk řadu efektorových proteinů včetně TAL efektorů. TAL effectors (Transcription Activator-like effectors) mají několik motivů včetně NLS, proto mohou vstupovat do jádra, kde se vážou na promotorové sekvence a aktivují transkripci rostlinných genů, které napomáhají bakteriální infekci.

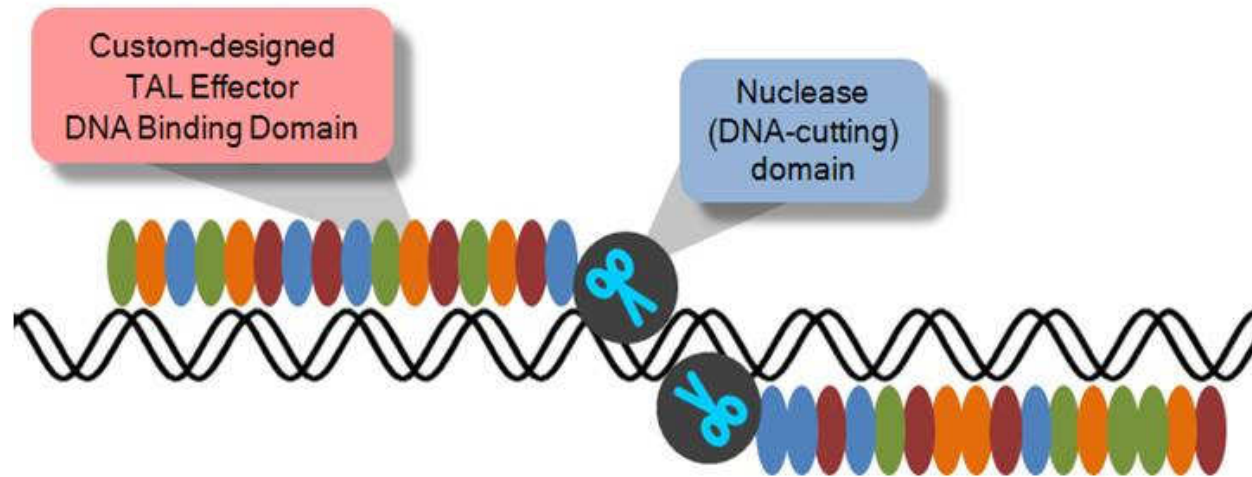
TAL obsahují v centrální části opakující se sekvence aminokyselin v počtu až 33, které jsou dlouhé obvykle 34 AA.

Typická repetice: LTPEQVVAIASHDGGKQALETVQRLLPVLCQAHG, avšak v pozicích 12 a 13 se vyskytují různé aminokyseliny (místo označované jako **repeat variable diresidue n. RVD**)

Byly navrženy TALE (Engineered TAL effectors) schopné se vázat na jakékoliv sekvence DNA díky záměnám aminokyselin. Takové umělé TALE lze využít pro aktivaci nebo represi endogenů u rostlin i živočichů.

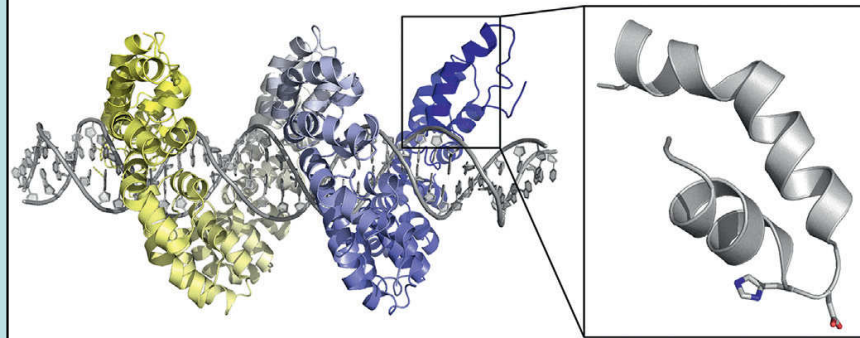
Engineered TAL effectors lze fúzovat se štěpicími doménami nukleáz (FokI) a vytvářet tak TAL Effector Nucleases (**TALENs**).

Struktura TALEN

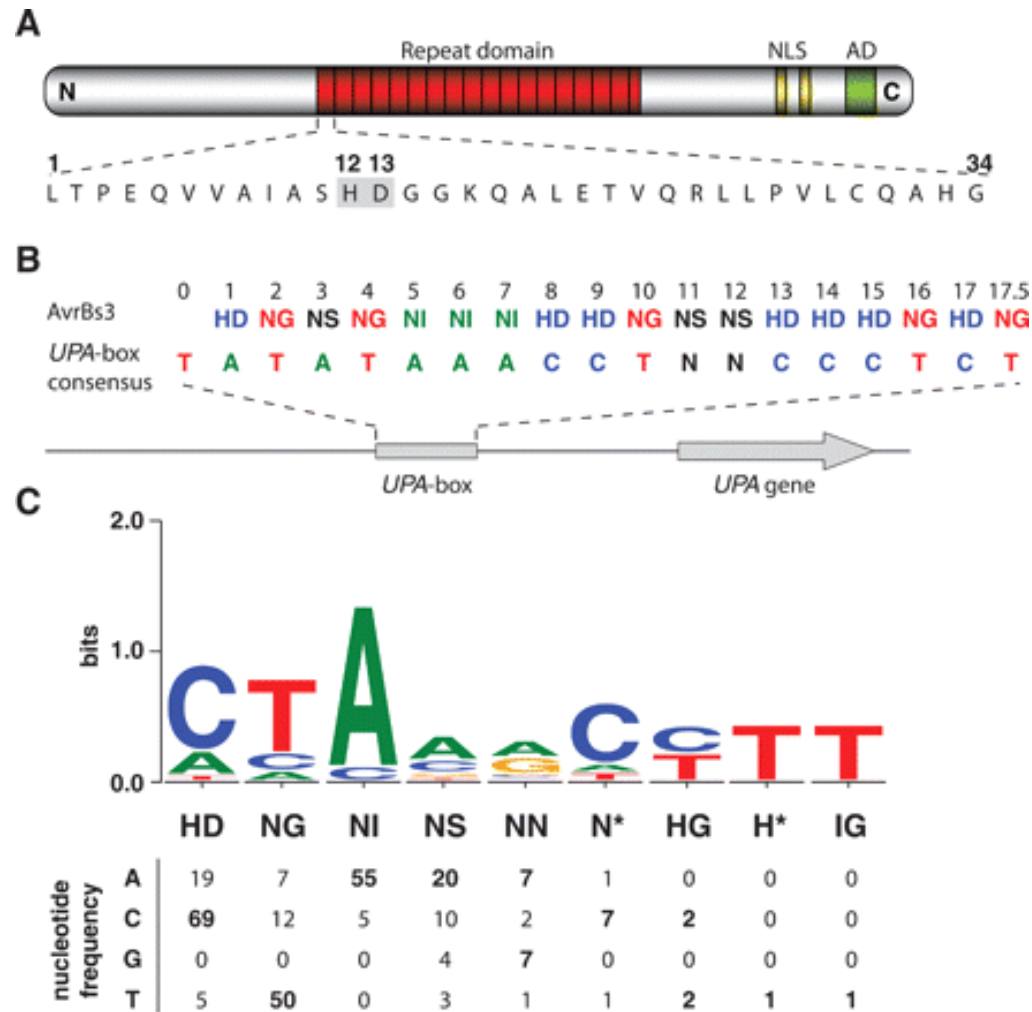


Protein TALE v komplexu s cílovou DNA.

Jednotlivé repetice proteinu TALE obsahují 33-35 AA, které rozpoznávají jednotlivé páry bází prostřednictvím dvou hypervariabilních zbytků aa (repeat-variable diresidues – RVDs)

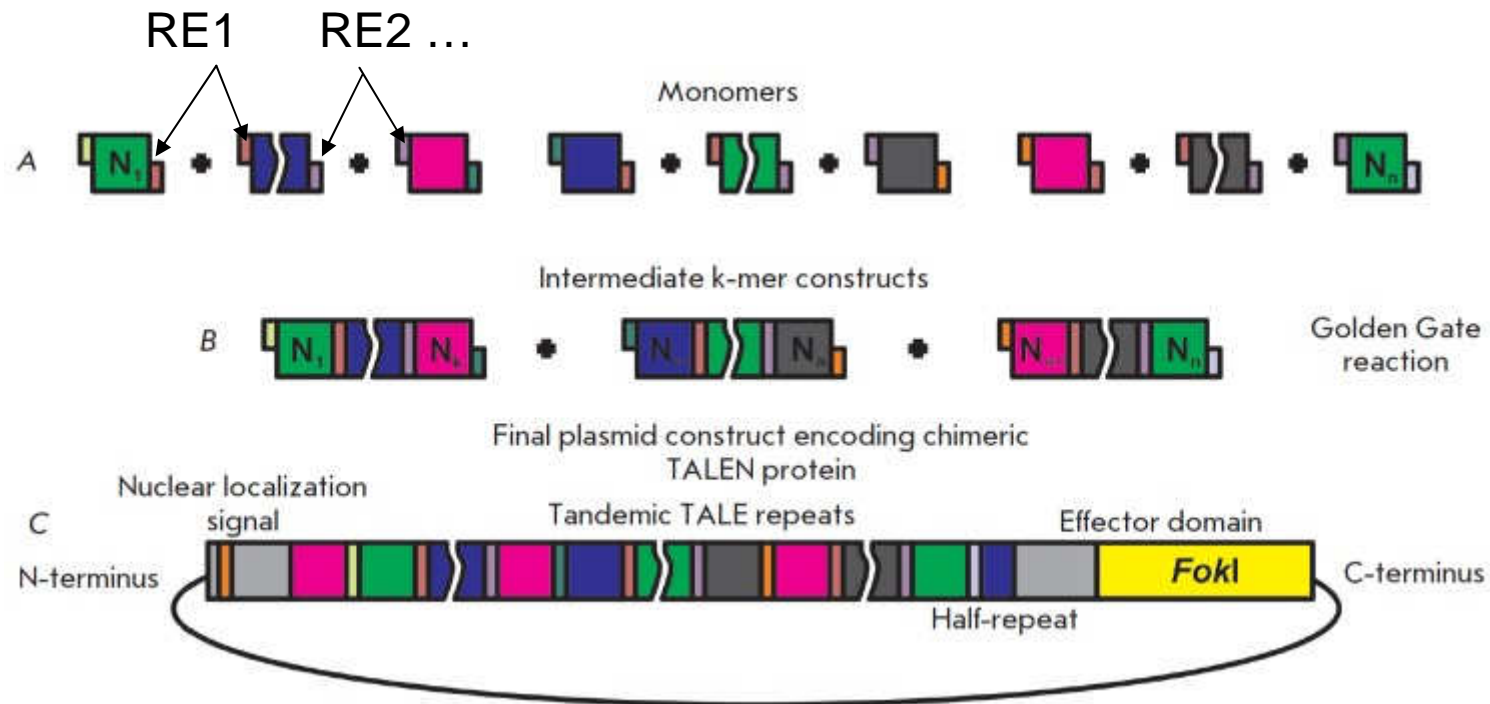


Dimer TALEN navázaný na DNA. Cílová místa pro TALEN jsou tvořena dvěma vazebnými oblastmi pro TALE oddělenými mezerou sekvencí různé délky (12-20 bp). TALE lze upravit tak, aby rozpoznával jedinečné sekvence vlevo a vpravo.



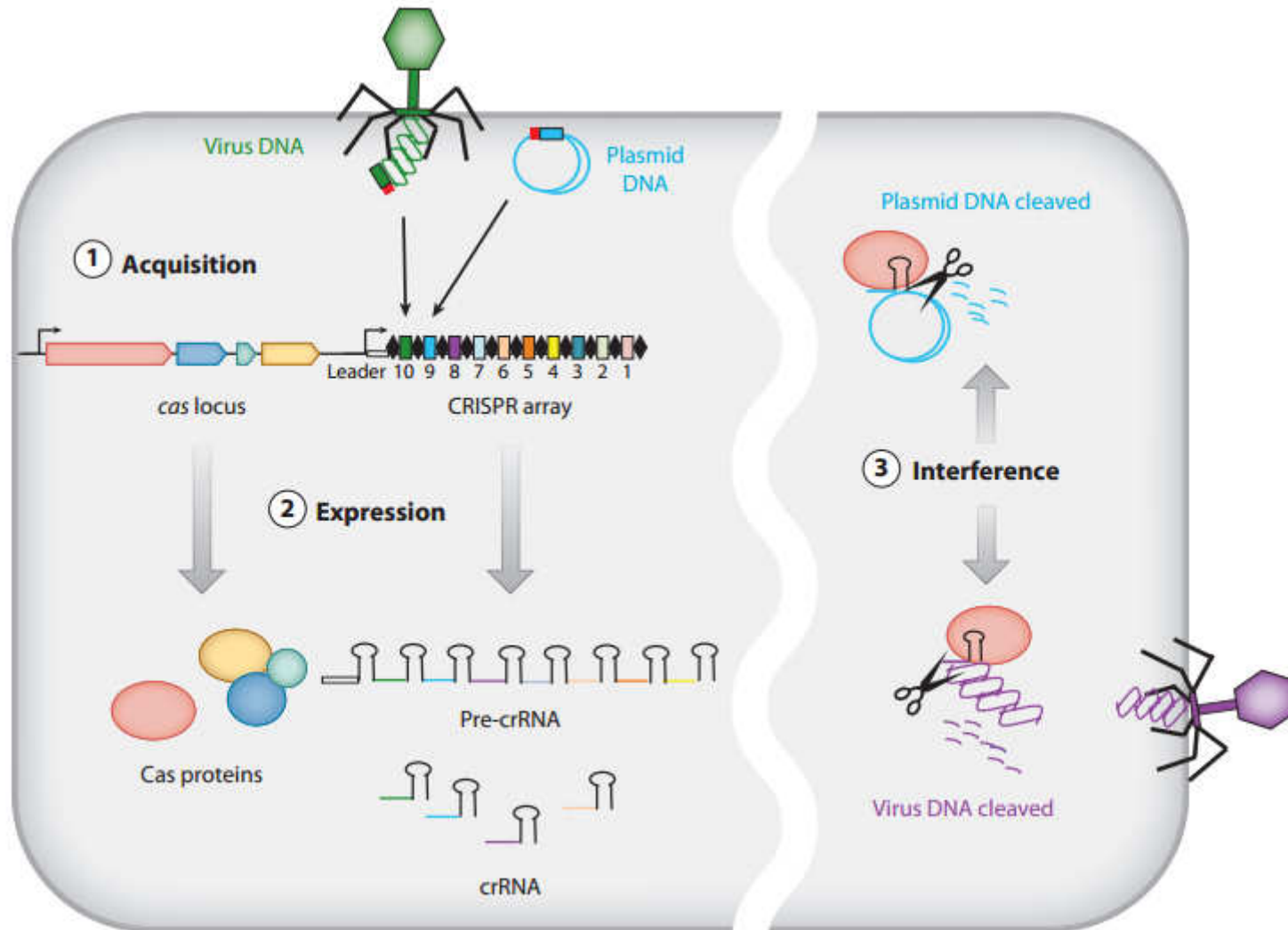
Model for DNA-target specificity of TAL effectors. **(A)** TAL effectors contain central tandem repeats, NLSs, and an AD. Shown is the amino acid sequence of the first repeat of AvrBs3. Hypervariable amino acids 12 and 13 are shaded in gray. **(B)** Hypervariable amino acids at position 12 and 13 of the 17.5 AvrBs3 repeats are aligned to the *UPA* box consensus (14). **(C)** Repeats of TAL effectors and predicted target sequences in promoters of induced genes were aligned manually. Nucleotides in the upper DNA strand that correspond to the hypervariable amino acids in each repeat were counted on the basis of the following combinations of eight effectors and

Schéma strategie pro přípravu genetických konstruktů exprimujících chimerické proteiny TALEN – postupná ligace klonovaných monomerů

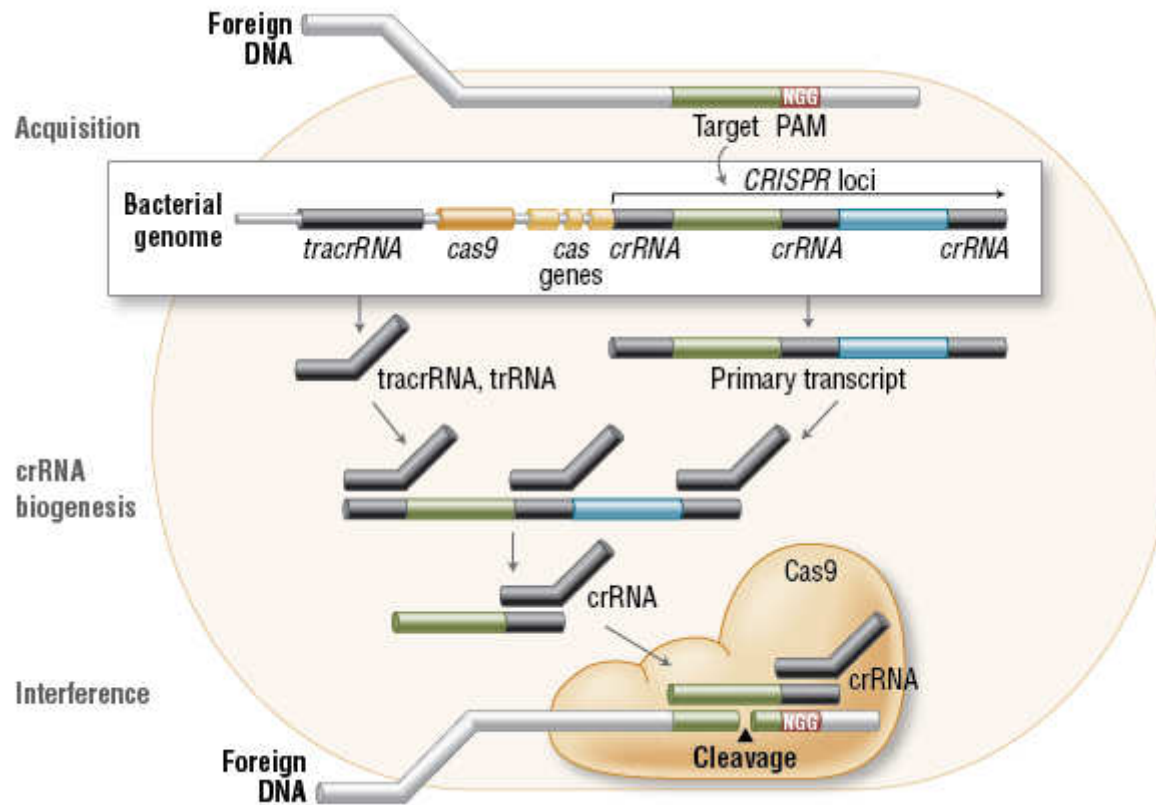


K dispozici jsou knihovny monomerů, dimerů, trimerů a tetramerů

Systém CRISPR/Cas

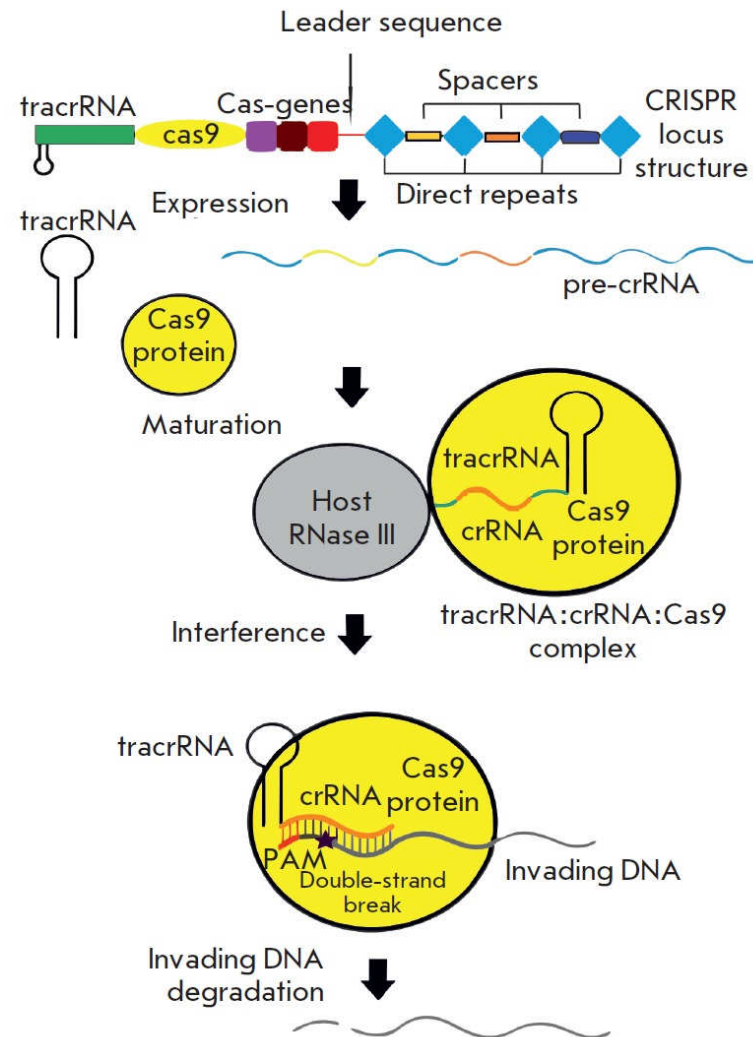


Fungování systému CRISPR-Cas

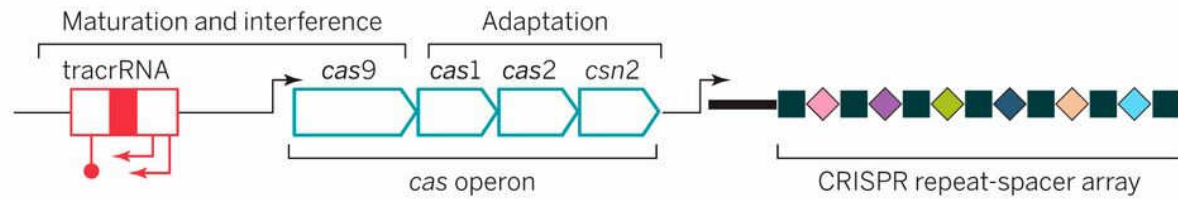


- Úseky cizorodé DNA jsou začleněny do bakteriálního genomu do lokusů CRISPR
- Lokusy CRISPR jsou přepsány a upraveny do *crRNA* (*crRNA* biogenesis)
- Během interference vytváří endonukleáza Cas9 komplex s *crRNA* a *tracrRNA*, který pak štěpí cizorodou DNA obsahující sekvenci 20-ti nukleotidů komplementárních k *crRNA* poblíž sekvence PAM.

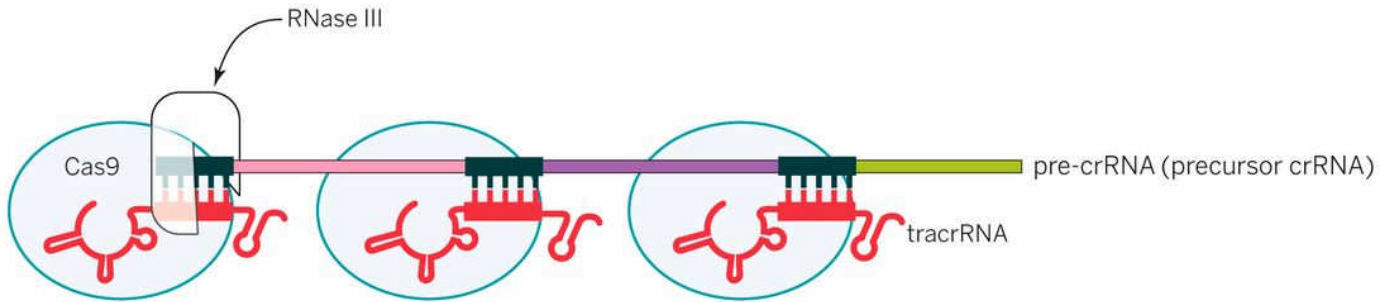
Fungování CRISPR-Cas9 komplexu



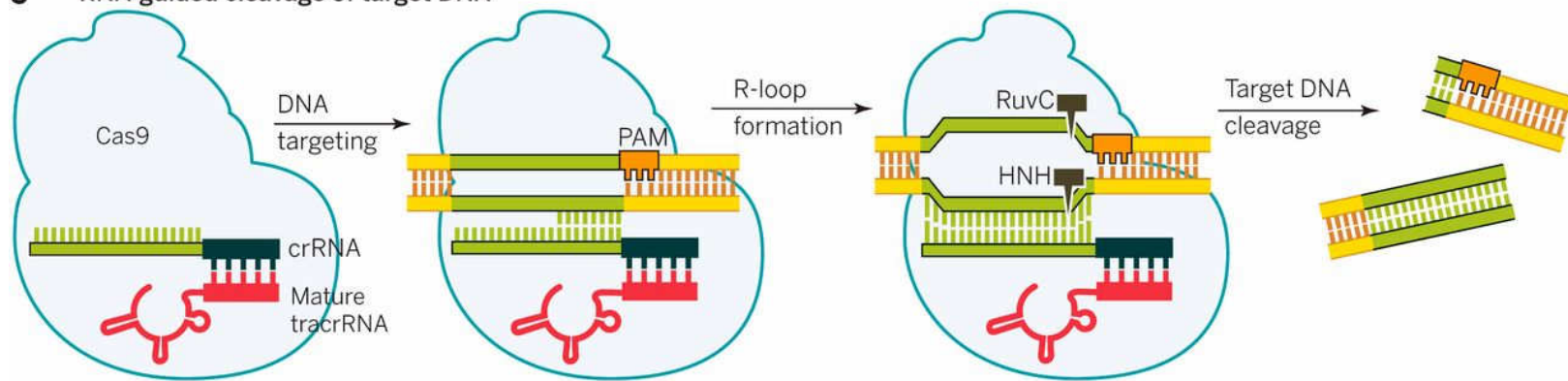
A Genomic CRISPR locus



B *tracrRNA*:*crRNA* co-maturation and Cas9 co-complex formation

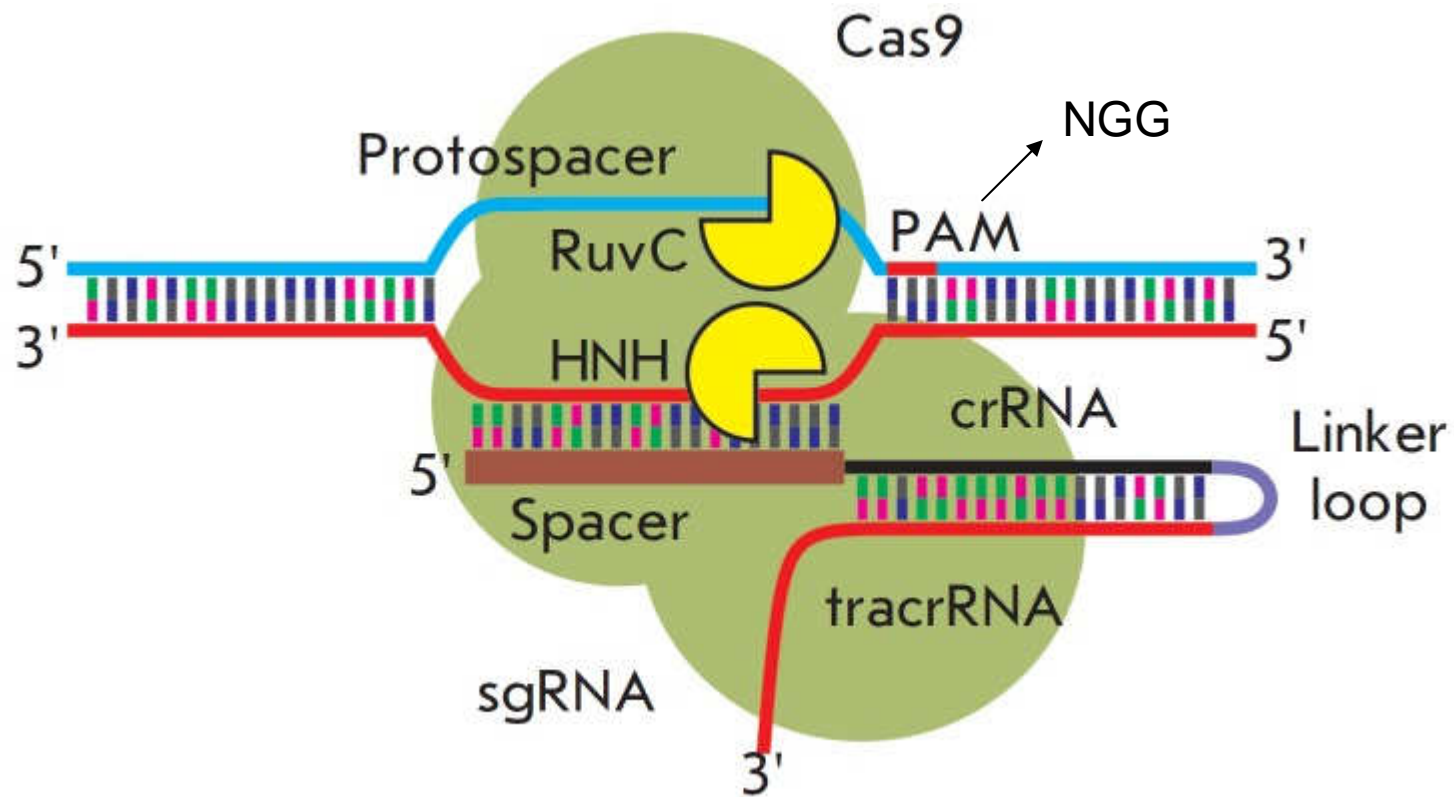


C RNA-guided cleavage of target DNA

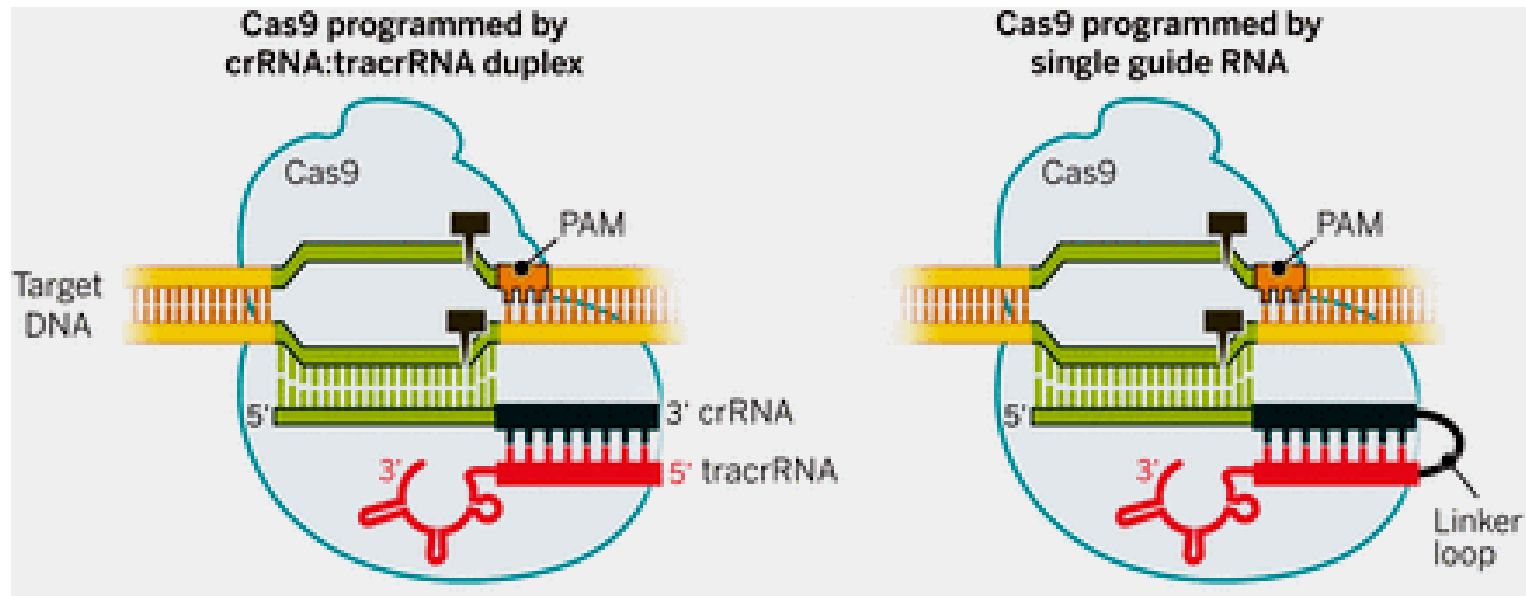


Struktura chimerické sgRNA a komplexu sgRNA s Cas9 pro vytváření DSB v cílových místech.

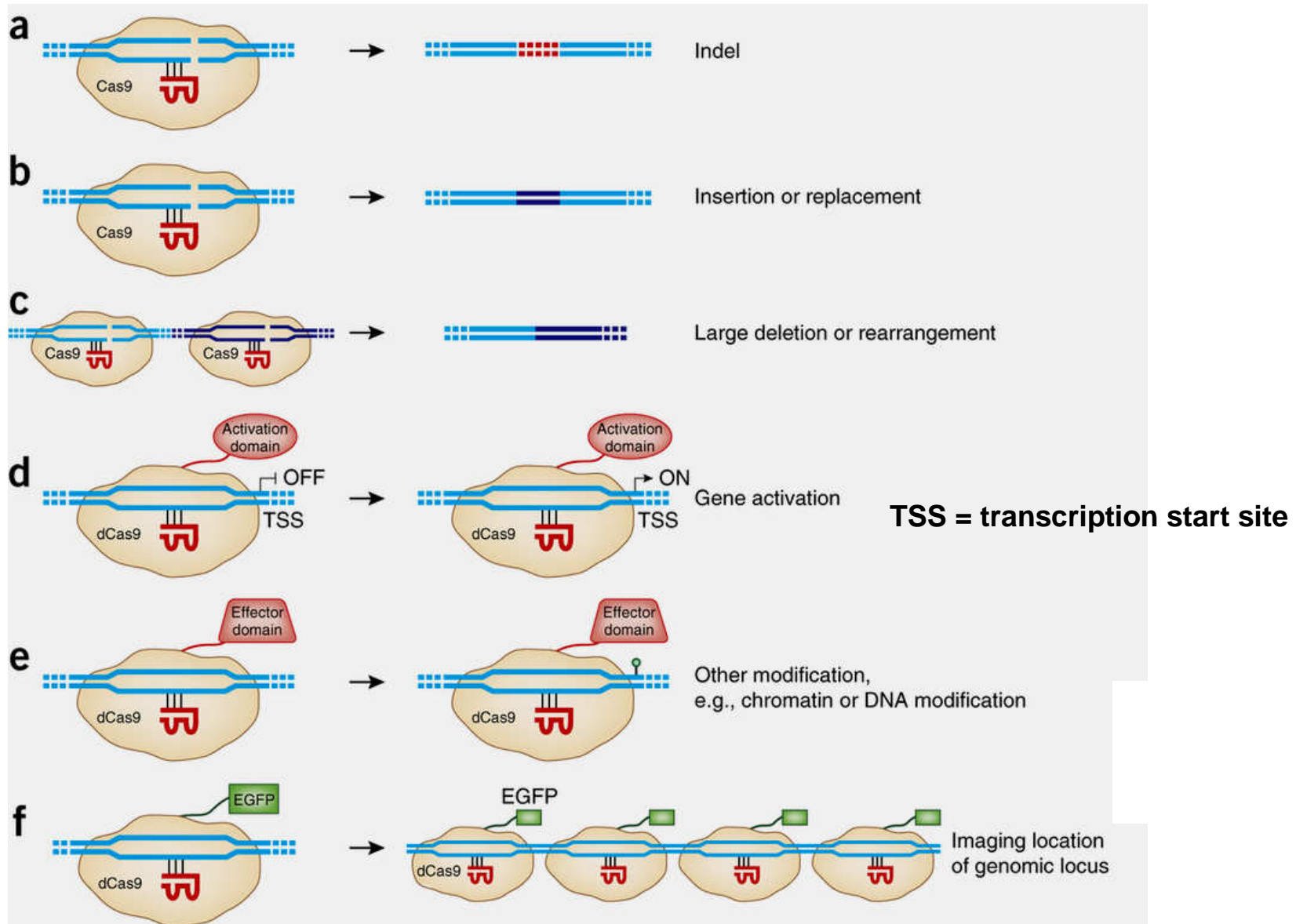
sgRNA je vytvořena spojením crRNA a tracrRNA.



Srovnání struktury crRNA:tracrRNA a sgRNA

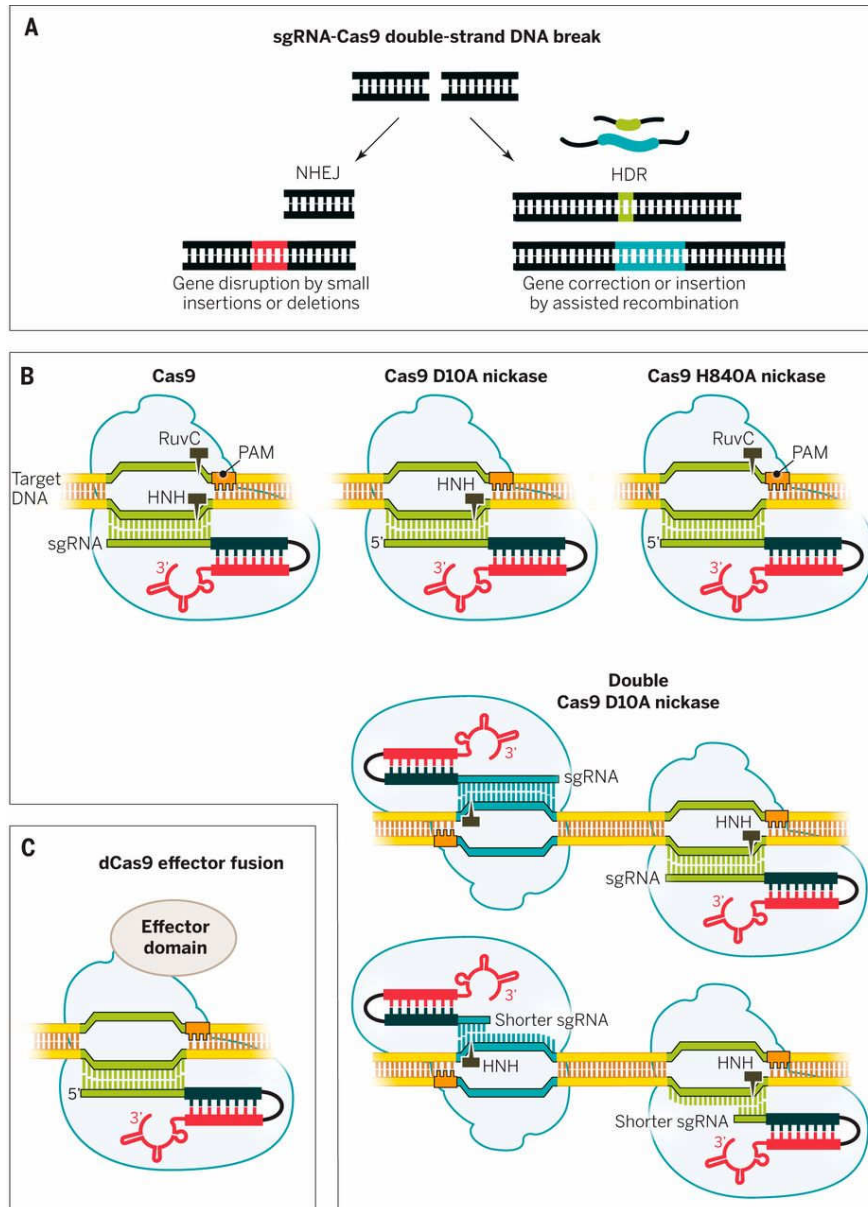


Možnosti editace genomů pomocí systému CRISPR/Cas9

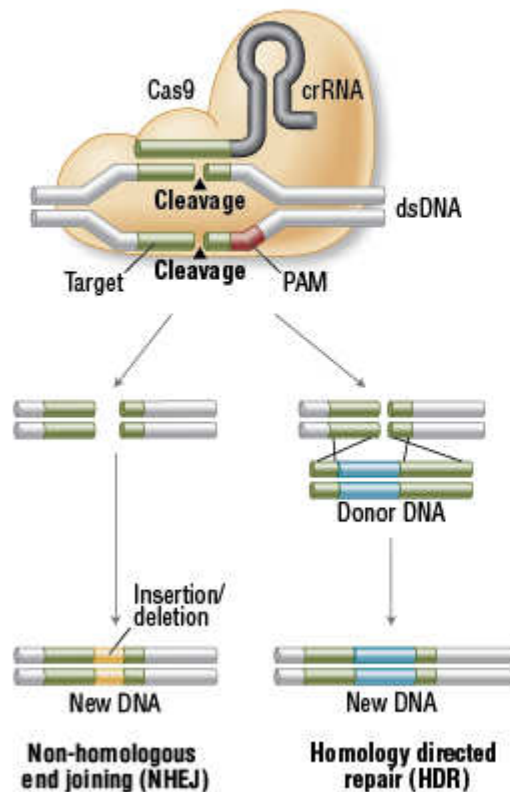


EGFP = fluorescenční proteiny

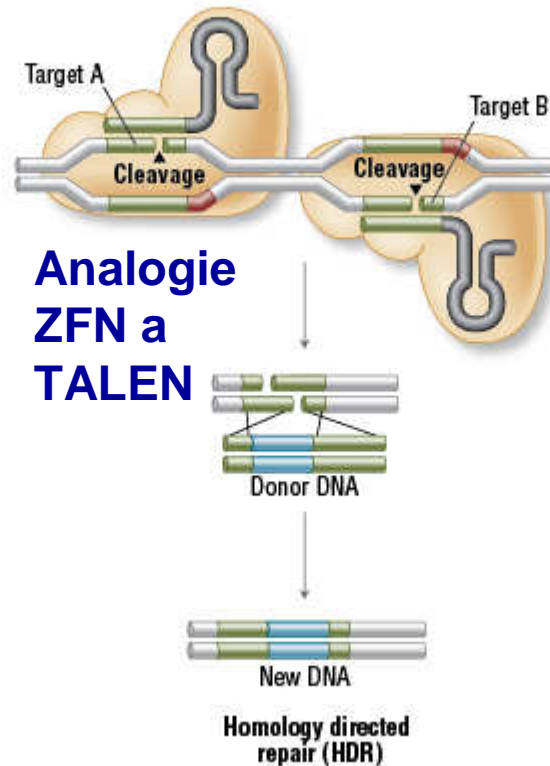
Působení modifikovaného systému CRISPR/Cas9



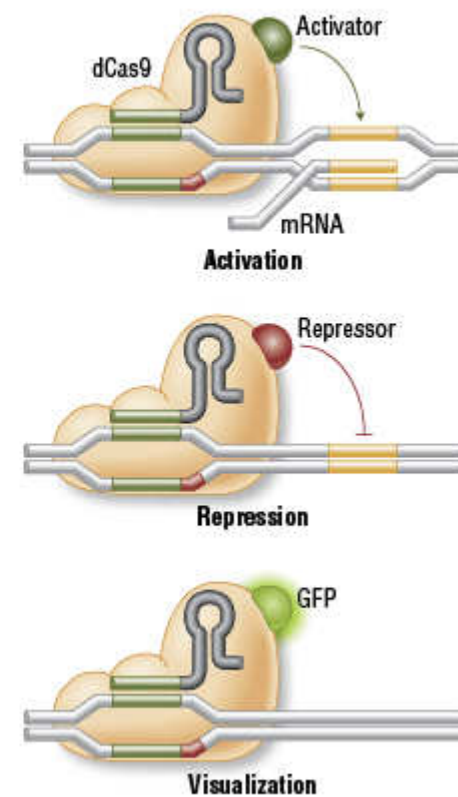
A. Genome Engineering With Cas9 Nuclease



B. Genome Engineering By Double Nicking With Paired Cas9 Nickases

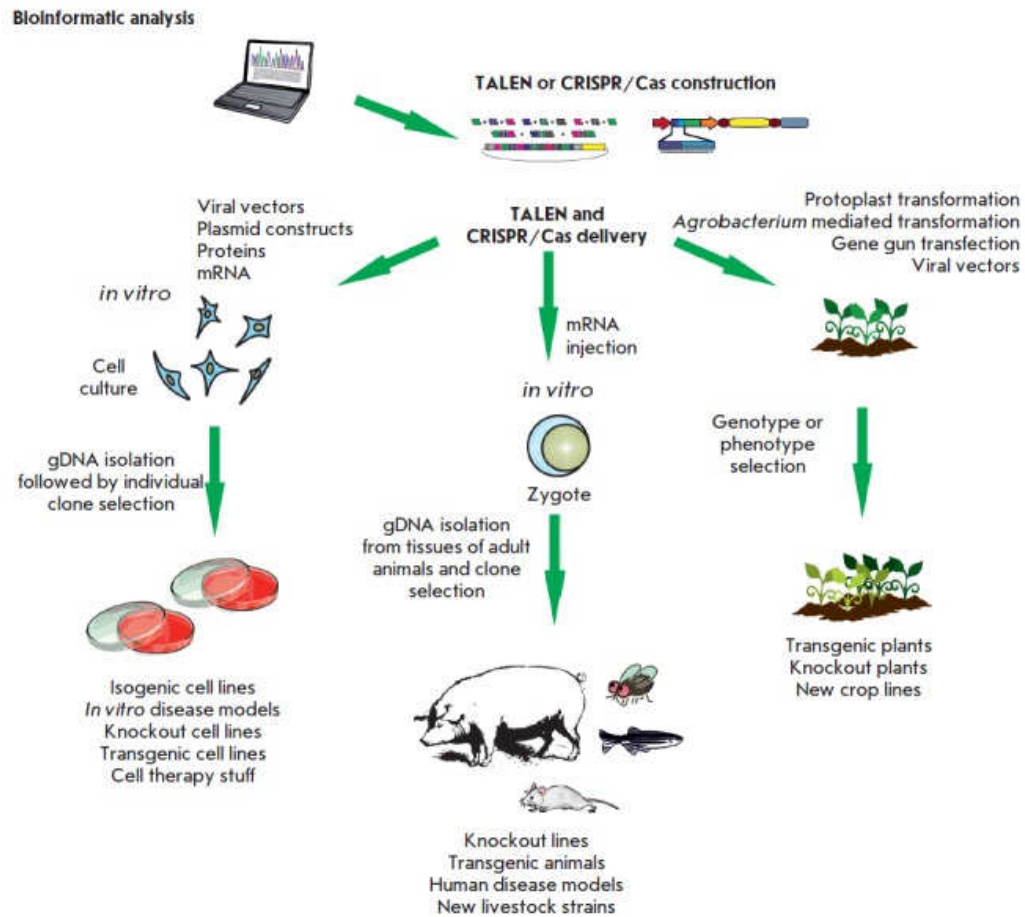


C. Localization With Defective Cas9 Nuclease



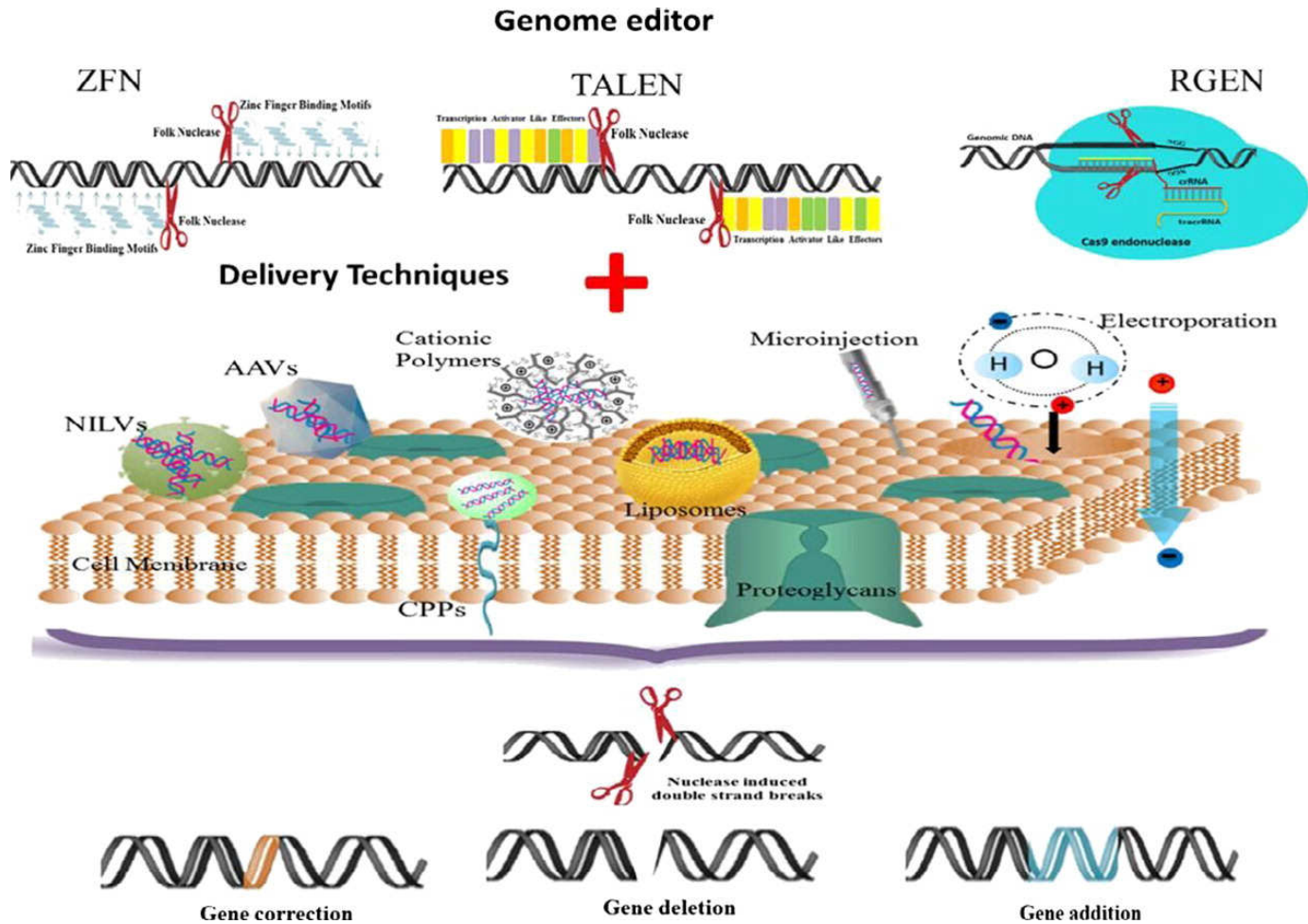
- A. Wt Cas9 nukleáza štěpí specificky dsDNA, vytváří DSB a tím navozuje reparaci. Za nepřítomnosti homologní sekvece může dojít k nehomolognímu spojování konců (NHEJ) a vzniku indelů, přerušujících zasažený gen. Za přítomnosti homologní sekvece může dojít k reparaci s využitím této sekvece a jejími vložením do místa zlomu.
- B. Mutantní Cas9 nukleáza vytváří místně specifické jednořetězcové zlomy. Dvojice takových sgRNA může zavádět posunuté zlomy, které mohou být reparovány s využitím homologní DNA.
- C. Mutantní Cas 9 nukleáza může být připojena k různým efektorovým doménám, které umisťuje na specifické sekvece. Mohou to být TF, represory, fluorescenční značky aj.

Obecná strategie genomového inženýrování



1. Selection of a target nucleotide sequence in the genome;
2. Generation of a nuclease construct directed at the selected target;
3. Delivery of this construct to the cell nucleus; and
4. Analysis of produced mutations.

Current and future delivery systems for engineered nucleases: ZFN, TALEN and RGEN



Genomic engineering using TALEN and CRISPR/Cas

Nuclease	Object	Gene	Objective	Reference
TALEN	Human cells (<i>Homo sapiens</i>)	<i>ccr5, akt2, e17k, angptl3, apob, atgl, c6orf106, celsr2, cftr, ciita, foxo1, foxo3, gli1, glut4, hbb, hdac1, hdac2, hdac6, hmga2, hoxa13, hoxa9, hoxc13, hprt, il2rg, jak2, kras, linc00116, maoa, map2k4, mdm2, met, mlh1, msh2, mutyh, myc, mycl1, mycn, nbn, ncor1, ncor2, nlr5, ntf3, pdgfra, pdgfrb, phf8, plin1, pms2, ppp1r12c (aavs1), ptch1, pten, rara, rbbp5, recq4, ret, runx1, sdhb, sdhc, sdhd, setdb1, sirt6, smad2, sort1, sox2, klf4ss18, suz12, tfe3, tp53, trib1, tsc2, ttn, vhl, xpa, xpc, abl1, alk, apc, atm, axin2, bax, bcl6, bmpr1a, brca1, brca2, cba3, cba8, ccnd1, cdc73, cdk4, cdh4, chd7, ctmb1, cyld, ddb2, ercc2, ewsr1, ext1, ext2, ezh2, fanca, fancd, fancf, fancg, fes, fgfr1, fh, flcn, flt4, mstn, aavs2, oct4, pitx3</i>	Knockout, insertion	[67, 68, 70-72, 74, 92, 176-179, 180]
	Yeast (<i>Saccharomyces cerevisiae</i>)	<i>URA3, ADE2, LYS3</i>	Knockout, insertion	[181]
	Nematode (<i>Caenorhabditis elegans</i>)	<i>ben-1, tex-1, sdc-2</i>	Knockout	[182]
	Drosophila (<i>Drosophila melanogaster</i>)	<i>yellow, crhr1, ponzr1, bmil, cdh5, dip2a, elmo1, epas1b, fh, golden, gria3, hey2, hif1ab, ilczf1, jak3, moesina, myod, phf6, ppp1cab, ryr1a, ryr3, scl6a3, tba6, tnikeb, th, fam46c, smad5</i>	Knockout, insertion	[94, 183-187]
	Silkworm (<i>Bombyx mori</i>)	<i>blos2</i>	Knockout	[188]
	Cricket (<i>Gryllus bimaculatus</i>)	<i>lac2</i>	Knockout	[189]
	Western clawed frog (<i>Xenopus tropicalis</i>)	<i>ets1, foxd3, grp78/bip, hhex, noggin, ptf1a/p48, sox9, vpp1</i>	Knockout	[190]
	Mouse (<i>Mus musculus</i>)	<i>c9orf72, fus, lepr, pak1ip1, gpr55, rprm, fbxo6, smurf1, tmem74, wdr20a, dcaf13, fam73a, mlkl, mstn, pibf1, sepw1, rab38, zic2</i>	Knockout, insertion	[179, 191-196]
	Rat (<i>Rattus norvegicus</i>)	<i>bmpr2, IgM</i>	Knockout	[197, 198]
	Pig (<i>Sus scrofa</i>)	<i>amely, dmd, gdf8, ggta, qhdrhr, il2rg, ldlr, raq2, rela (p65), sry</i>	Knockout	[199]
	Cattle (<i>Bos taurus</i>)	<i>acan, gdf8, ggta, mstn, prnp</i>	Knockout	[179, 199]
	Arabidopsis (<i>Arabidopsis thaliana</i>)	<i>adh1</i>	Knockout	[70]
	Tobacco (<i>Nicotiana benthamiana</i>)	<i>surA, surB, hax3</i>	Knockout, insertion	[156, 157]
	False brome grass (<i>Brachypodium distachyon</i>)	<i>aba1, caxk2, coi1, hta1, rht, sbp, smc6, spl</i>	Knockout	[154]
	Rice (<i>Oryza sativa</i>)	<i>avrxa7, pthxo3, badh2, ckx2, dep1, sd1</i>	Knockout	[154, 155]
CRISPR/Cas	Yeast (<i>Saccharomyces cerevisiae</i>)	<i>CAN1, ADE2</i>	Knockout, insertion	[200]
	Human cells (<i>Homo sapiens</i>)	<i>dnmt3b-tdTomato, pou5f1(oct4), emx1, dyrk1a, grin2b, egfp, ccr5, c4bpb, pvalb, aavs, akt2, celsr2, ciita, glut4, linc00116, sort1, ldlr</i>	Knockout, insertion	[46, 51, 78, 80, 201, 202]
	Nematode (<i>Caenorhabditis elegans</i>)	<i>dpy-11, unc-4, ben-1, unc-36, daf-2, klp-12, lab-1, egfp, dpy-11, lin-5, rol-1, dpy-3, unc-1, dpy-13, unc-119, klp-12</i>	Knockout, insertion	[119-124, 126]
	Drosophila (<i>Drosophila melanogaster</i>)	<i>yellow, white, rosy, cg14251 (k81), cg3708cg17629 (kl-3), light</i>	Knockout, insertion	[130-133]
	Danio rerio (<i>Danio rerio</i>)	<i>etsrp, gata5, etsrp, gsk3b, apoea, fh, fh1, th1, rgs4, tia1l, tph1a, drd3, egfp, tyr, gol, mitfa, ddx19, sema3fb, dre-mir-126a, dre-mir-126b, dre-mir-17a-1-dre-mir-92a-1, dre-mir-17a-2-dre-mir-92a-2, fgd5, ensdarg00000070653, ensdarg00000076787, psmf1, dre-mir-126a, dre-mir-17a-2, dre-mir-92a-2, tardbp, tardbpl, c13h9orf72</i>	Knockout, insertion, chromosomal rearrangements	[81, 82, 203-206]
	Frog (<i>Xenopus tropicalis</i>)	<i>tyr, six3</i>	Knockout	[207]
	Pig (<i>Sus scrofa</i>)	<i>gdf8, p65</i>	Knockout, insertion	[208]
	Mouse (<i>Mus musculus</i>)	<i>tet1, tet2, tet3, sry, uty, rosa26, hprt, egfp, th, rheb, uhrf2</i>	Knockout, insertion	[83, 144, 209, 210]
	Rat (<i>Rattus norvegicus</i>)	<i>dnmt1, dnmt3a, dnmt3b, tet1, tet2, tet3, mc3r, mc4r</i>	Knockout, insertion	[144, 145, 211]
	Arabidopsis (<i>Arabidopsis thaliana</i>)	<i>pds33, fls2, bri1, jaz1, gaj, chl, chl2, 5g13930</i>	Knockout, insertion	[87, 88, 149]
	Tabacco (<i>Nicotiana benthamiana</i>)	<i>Pds</i>	Knockout, insertion	[88, 89]
	Rice (<i>Oryza sativa</i>)	<i>ods, badh2, mrk2, 02g23823, roc5, spp, ysa, myb1, cao1, lazy1, sweet11, sweet14</i>	Knockout, insertion	[86, 150, 152]
	Wheat (<i>Triticum aestivum</i>)	<i>Mlo</i>	Knockout	[86]

Příklady typů buněk a organismů, které byly geneticky modifikovány pomocí Cas9

Biology

Cell lines

HEK293
U2OS
K562

Model organisms

Mice
Rats
Fruit flies
Nematodes
Arabidopsis
Salamanders
Frogs
Monkeys

Biotechnology

Crop plants

Rice
Wheat
Sorghum
Tobacco

Fungi

Kluyveromyces
Chlamydomonas

Biomedicine

Organoids
hESCs
iPSCs

