CG020 Genomika

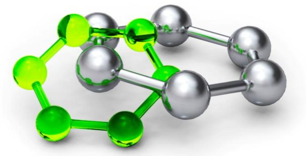
Přednáška 5

RNA interference a editování genomu

Jan Hejátko

Funkční genomika a proteomika rostlin,
Středoevropský technologický institut (CEITEC)
a
Národní centrum pro výzkum biomolekul,
Přírodovědecká fakulta,

Masarykova univerzita, Brno
hejatko@sci.muni.cz, www.ceitec.eu

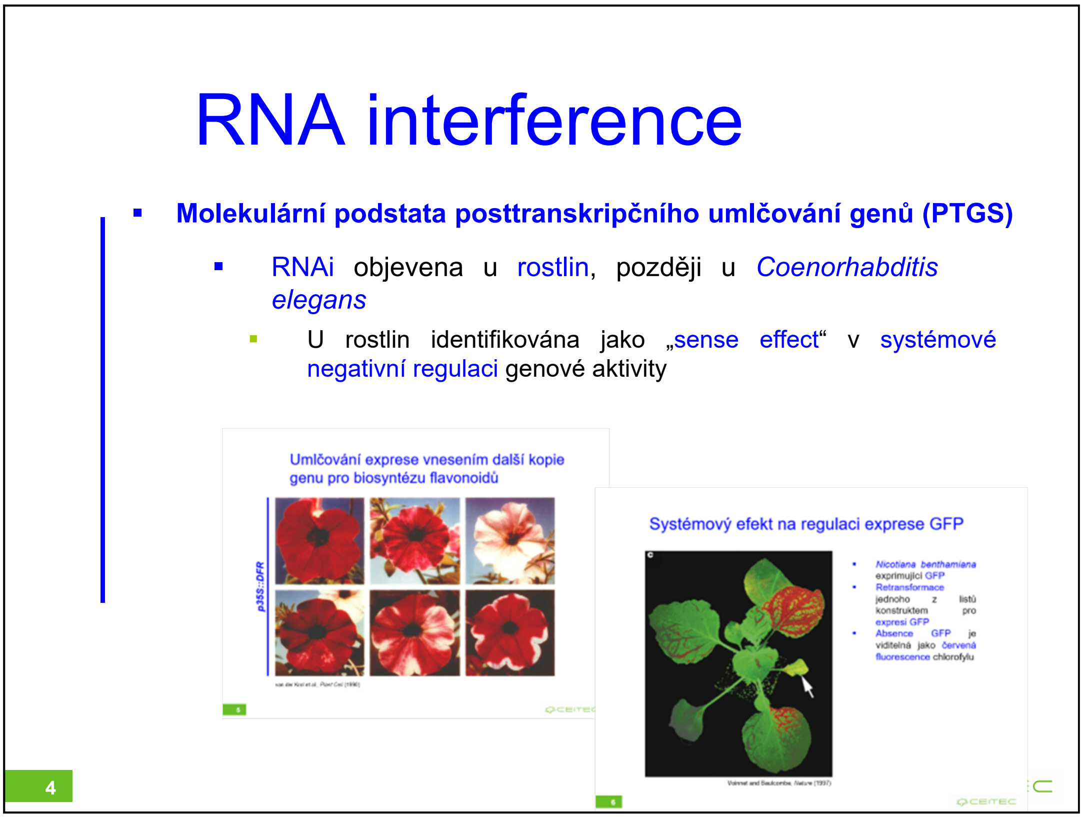


Osnova

- Umlčování genů pomocí RNA interference
  - Mechanismus RNAi
- Editace genomu
  - Princip editace genomu prostřednictvím místně specifických nukleáz (Site Directed Nucleases, SDNs)
  - Zinc-Finger Nucleases (ZFNs)
  - Transcription Activator-Like Effectors (TALENs)
  - Clustered Regularly Interspaced Short Palindromic Repeats/Cas9 (CRISPR/Cas9)

Osnova

- Umlčování genů pomocí RNA interference
    - Mechanismus RNAi



RNA interference

- **Molekulární podstata posttranskripčního umlčování genů (PTGS)**
  - RNAi objevena u rostlin, později u *Coenorhabditis elegans*
    - U rostlin identifikována jako „sense effect“ v systémové negativní regulaci genové aktivity

Umlčování exprese vnesením další kopie genu pro biosyntézu flavonoidů

van der Krol et al., *Plant Cell* (1990)

Systémový efekt na regulaci exprese GFP

- *Nicotiana benthamiana* exprimující GFP
- Retransformace jednoho z listů konstruktem pro expresi GFP
- Absence GFP je viditelná jako červená fluorescence chlorofylu

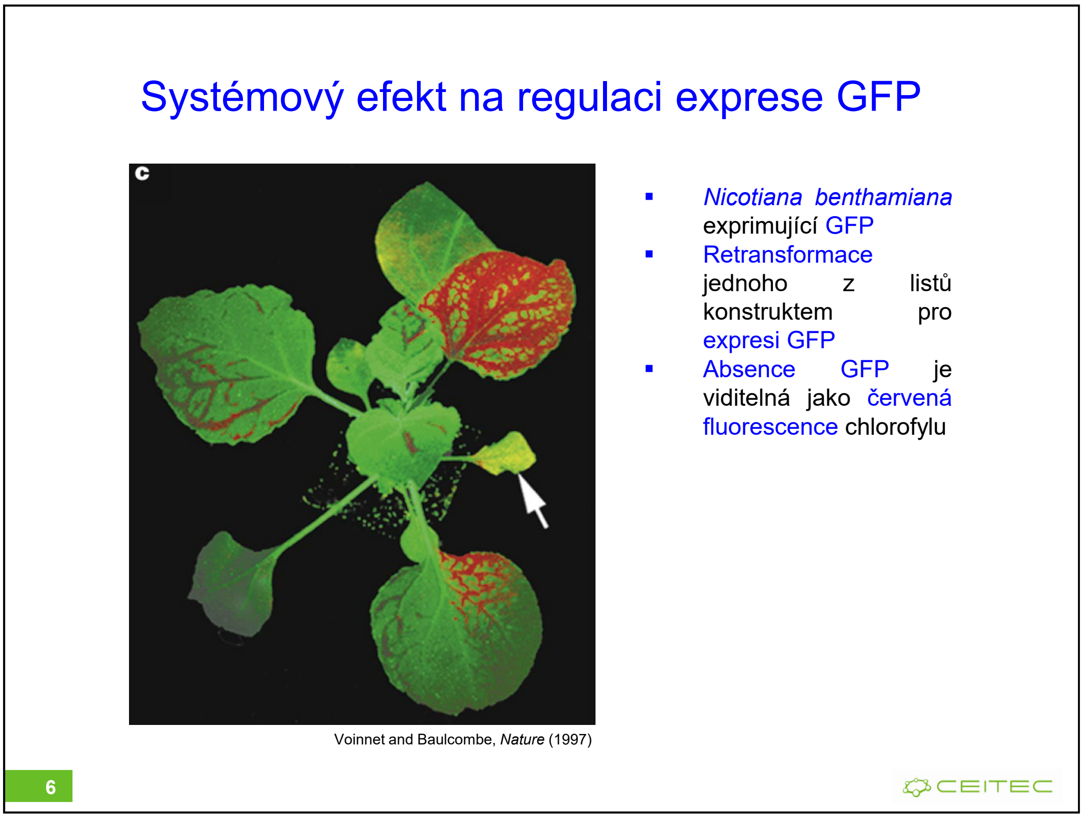
Voinnet and Baulcombe, *Nature* (1997)

Analysis of GUS expression of supertransformed rice callus. Transgenic rice tissue containing a single Gus transgene supertransformed with UbiDGus[s], UbiDGus[ays], UbiDGus[iyr], DGus[iyr].

Umlčování exprese vnesením další kopie genu pro biosyntézu flavonoidů

van der Krol et al., *Plant Cell* (1990)

Flowers on petunia VR plants transformed with the dihydroflavonol-4-reductase (DFR) sense gene construct VIP178 showed either an unaffected flower pigmentation (top left) or a reduction in pigment synthesis. Shown from top left to bottom right: 178-1, 178-14, 178-6, 178-16, 178-10, and 178-15. On transformant 178-16, flower pigmentation varies from fully pigmented to white. Transformant 178-15 shows an ectopic expression pattern, resulting in a white ring at the edge of corolla tissue.



Systémový efekt na regulaci exprese GFP

- *Nicotiana benthamiana* exprimující GFP
- Retransformace jednoho z listů konstruktem pro expresi GFP
- Absence GFP je viditelná jako červená fluorescence chlorofylu

Voinnet and Baulcombe, *Nature* (1997)

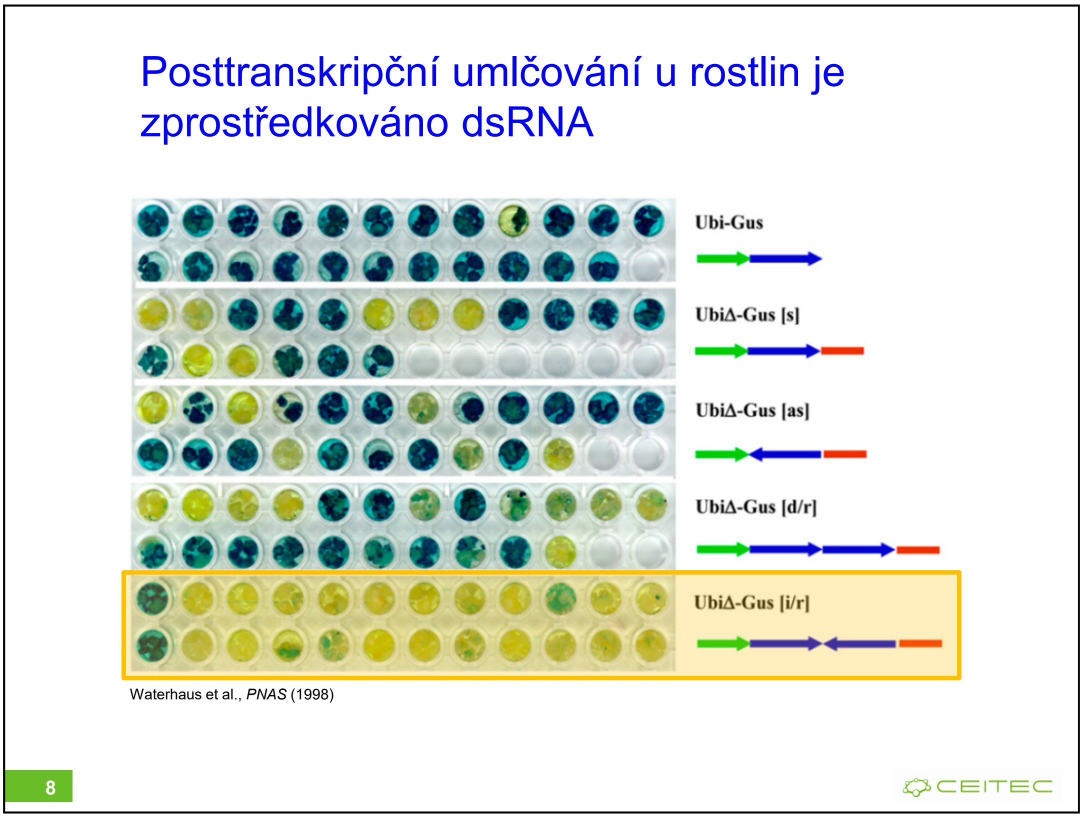
We studied *Nicotiana benthamiana* plants carrying a jellyfish green fluorescent protein (GFP) transgene5. We infiltrated leaves with strains of *Agrobacterium tumefaciens* carrying a GFP reporter gene.Intact GFP transgenic plant infiltrated 18 days previously in a lower leaf (arrow) showing the progression of GFP-silencing.

RNA interference

- **Molekulární podstata posttranskripčního umlčování genů (PTGS)**
  - RNAi objevena u rostlin a později u *Coenorhabditis elegans*
    - U rostlin identifikována jako „sense effect“ v systémové negativní regulaci genové aktivity
    - umlčování bylo indukováno jak sense tak antisense RNA
    - dsRNA indukovala umlčování cca 10-100x účinněji

Posttranskripční umlčování u rostlin je zprostředkováno dsRNA

Analysis of GUS expression of supertransformed rice callus. Transgenic rice tissue containing a single Gus transgene supertransformed with UbiDGus[s], UbiDGus[ays], UbiDGus[iyr], DGus[iyr].



Posttranskripční umlčování u rostlin je zprostředkováno dsRNA

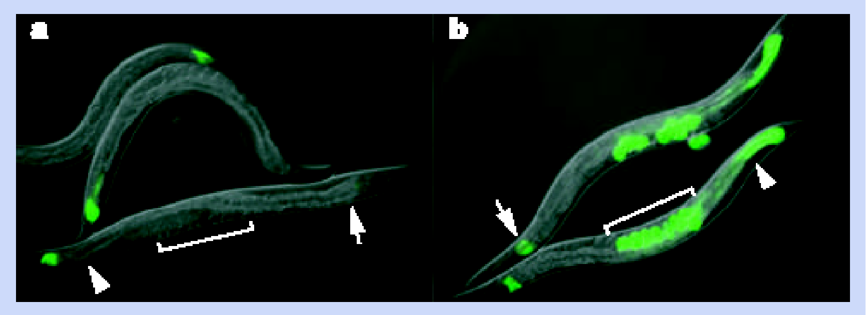
Waterhaus et al., *PNAS* (1998)

Kalusy rýže nesoucí konstrukt pro expresi uidA (GUS), který způsobuje modré zbarvení (1. řádek) byly retransformovány konstrukty pro expresi uidA v sense (2. řádek), anti-sense (3. řádek) a přímé a obrácené repetici (4. resp. 5. řádek). Všimněte si silné represe zbarvení a tedy i exprese uidA genu v případě retransformace konstruktem vedoucím k tvorbě dsRNA (obrácené repetice, 5. řádek).

RNA interference

- **Molekulární podstata posttranskripčního umlčování genů (PTGS)**
  - Umlčování genové exprese prostřednictvím dsRNA je závislé na vlastních genech
    - vyhledávání pomocí přímé genetiky

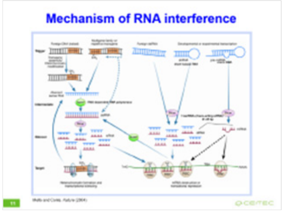
RNAi *rnai*



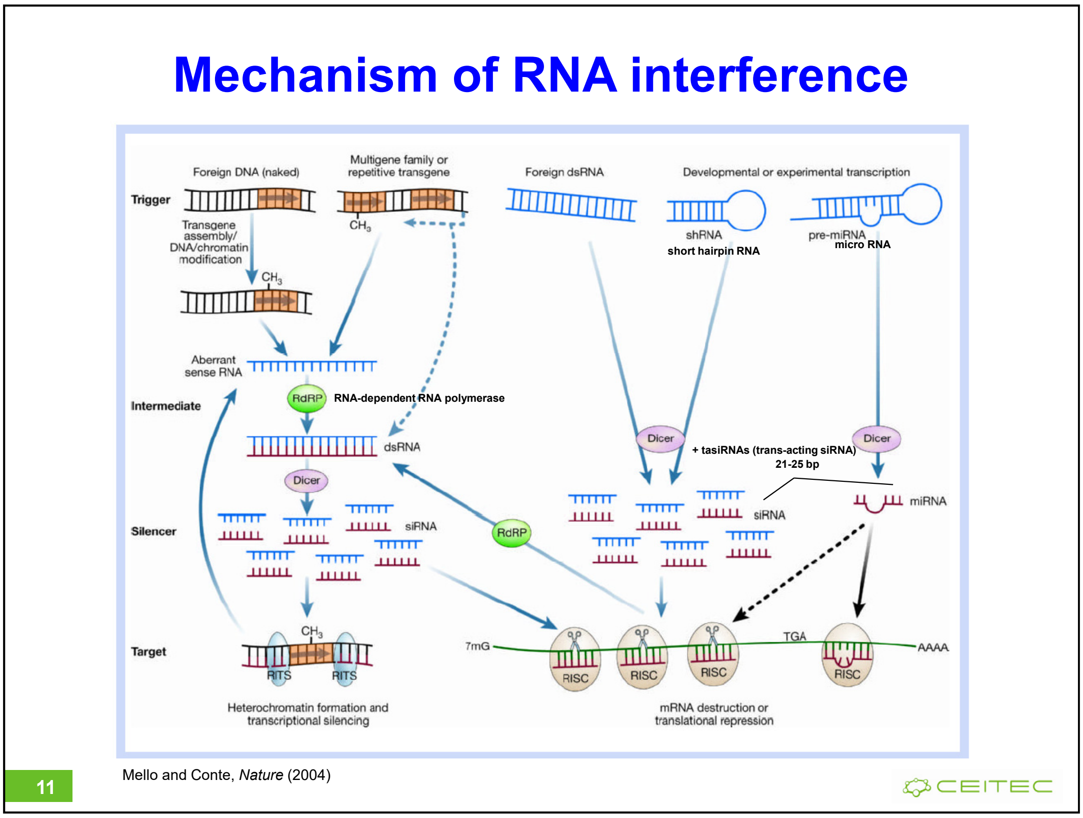
Mello and Conte, *Nature* (2004)

RNA interference

- **Molekulární podstata posttranskripčního umlčování genů (PTGS)**
  - je to **přirozený mechanismus** regulace genové exprese u všech eukaryot
  - podstatou je **tvorba dsRNA**, která může být spuštěna několika způsoby:
    - přítomnost **cizí „aberantní“ DNA**
    - **specifické transgeny** obsahující **obrácené repetice** částí cDNA
    - transkripce vlastních genů pro **shRNA** (short hairpin RNA) nebo **miRNA** (micro RNA, endogenní „vlásenková“ RNA)
  - **dsRNA** je procesována enzymovým komplexem (**DICER**), což vede k tvorbě **siRNA** (short interference RNA), která se pak váže buď na enzymový komplex **RITS** (**R**NA-**I**nduced **T**ranscriptional **S**ilencing complex) nebo **RISC** (**R**NA-**I**nduced **S**ilencing **C**omplex)
  - **RISC** zprostředkovává buď **degradaci mRNA** (v případě úplné similarity siRNA a cílové mRNA) nebo vede pouze k **zastavení translace** (v případě neúplné homologie jako je tomu např. v případě miRNA
  - **RITS** zprostředkovává **reorganizaci genomové DNA** (tvorba heterochromatinu a inhibice transkripce)



Mechanism of RNA interference



Mechanism of RNA interference

Mello and Conte, *Nature* (2004)

It has been found that dsRNA might be either an intermediate or a trigger in PTGS.

In the first case, dsRNA is formed by the action of RNA-dependent RNA polymerases (RdRPs), which use specific transcripts as a template. It is still not clear, how these transcripts are recognized, but it might be e.g. abundant RNA that is a result of viral amplification or transcription of foreign DNA.

It is not clear, how the foreign DNA might be recognized, possibly, lack of bound proteins on the foreign “naked” DNA and its subsequent “signature” (e.g. by specific methylation pattern) during packing of the foreign DNA into the chromatin structure might be involved.

The highly abundant transcripts might be recruited to the RdRPs by the defects in the RNA processing, e.g. lack of polyadenylation.

In the case when dsRNA is a direct trigger, there are two major RNA molecules involved in the process: Short interference RNA (siRNA) and micro RNA (miRNA), both encoded by the endogenous DNA.

These two functionally similar molecules differ in their origin:

siRNAs are dominantly product of the cleavage of the long dsRNA that are produced by the action of cellular or viral RdRPs. However, there are also endogenous genes, e.g. short hairpin RNAs (shRNAs) allowing production of the siRNA (see the figure).

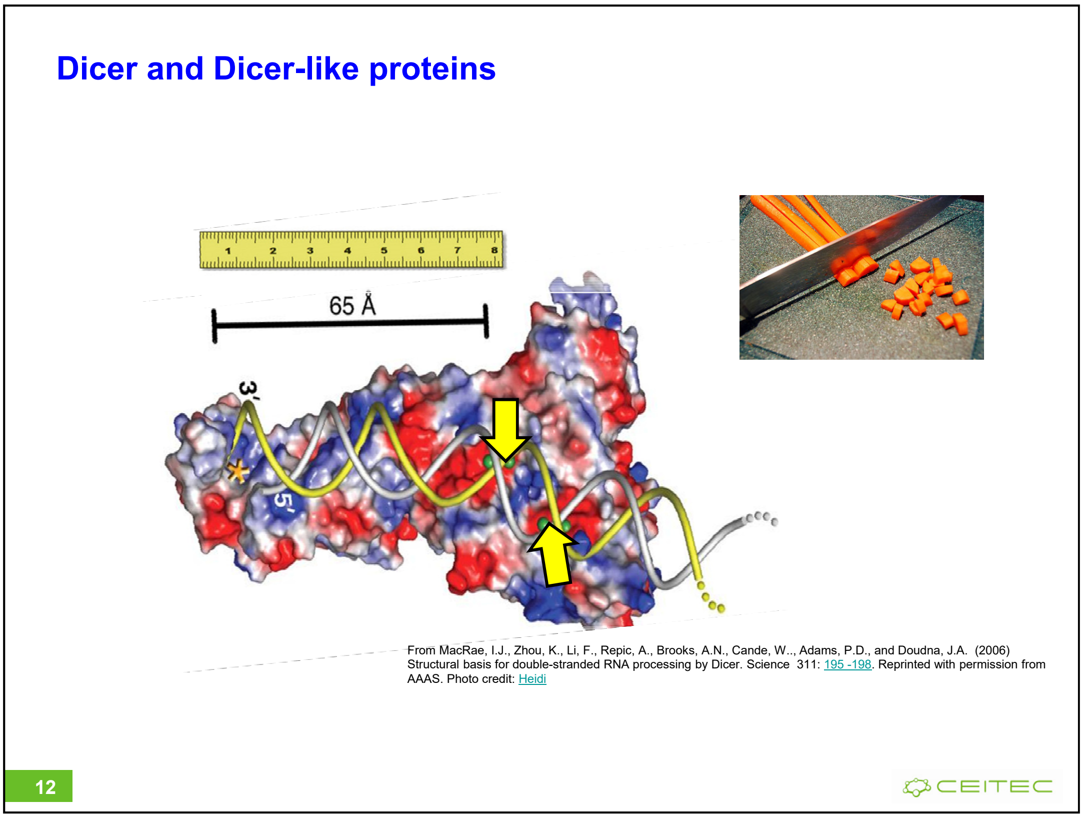
miRNAs are involved in the developmental-specific regulations and are product of transcription of endogenous genes encoding for small dsRNAs with specific structure (see the figure).

In addition to siRNAs, there are trans-acting siRNAs (tasiRNAs) that are a special class of siRNAs that appear to function in development (much like miRNAs) but have a unique mode of origin involving components of both miRNA and siRNA pathways.

Developmental regulations via miRNAs are more often used in animals then in plants.

The dsRNAs of all origins and pre miRNAs are cleaved by DICER or DICER-like (DCL) enzyme complexes with RNAse activity, leading to production of siRNAs and miRNA, respectively.

These small RNAs are of 21-24 bp long and bind either to RNA-induced transcriptional silencing complex (RITS) or RNA-induced silencing komplex (RISC).



Dicer and Dicer-like proteins

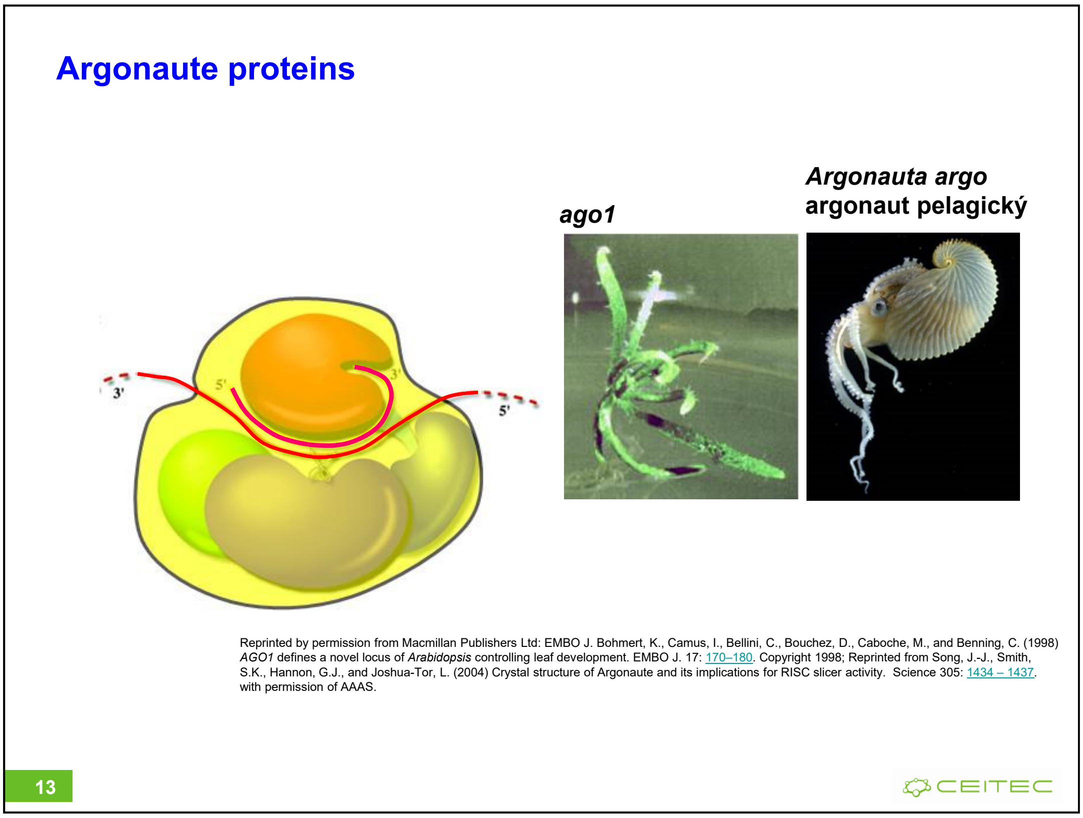
From MacRae, I.J., Zhou, K., Li, F., Repic, A., Brooks, A.N., Cande, W.., Adams, P.D., and Doudna, J.A. (2006) Structural basis for double-stranded RNA processing by Dicer. Science 311: 195 -198. Reprinted with permission from AAAS. Photo credit: Heidi

12

In siRNA and miRNA biogenesis, DICER or DICER-like (DCL) proteins cleave long dsRNA or foldback (hairpin) RNA into ~ 21 – 25 nt fragments.

Dicer’s structure allows it to measure the RNA it is cleaving. Like a cook who “dices” a carrot, DICER chops RNA into uniformly-sized pieces.

Note the two strands of the RNA molecule. The cleavage sites are indicated by yellow arrows.



Argonaute proteins

ago1

Argonauta argo
argonaut pelagický

Reprinted by permission from Macmillan Publishers Ltd: EMBO J. Bohmert, K., Camus, I., Bellini, C., Bouchez, D., Caboche, M., and Benning, C. (1998) *AGO1* defines a novel locus of *Arabidopsis* controlling leaf development. EMBO J. 17: 170–180. Copyright 1998; Reprinted from Song, J.-J., Smith, S.K., Hannon, G.J., and Joshua-Tor, L. (2004) Crystal structure of Argonaute and its implications for RISC slicer activity. Science 305: 1434 – 1437. with permission of AAAS.

ARGONAUTE proteins bind small RNAs and their targets and it is an important part of both RITS and RISC complexes.

ARGONAUTE proteins are named after the *argonaute1* mutant of *Arabidopsis*; *ago1* has thin radial leaves and was named for the octopus *Argonauta* which it resembles (see the figure).

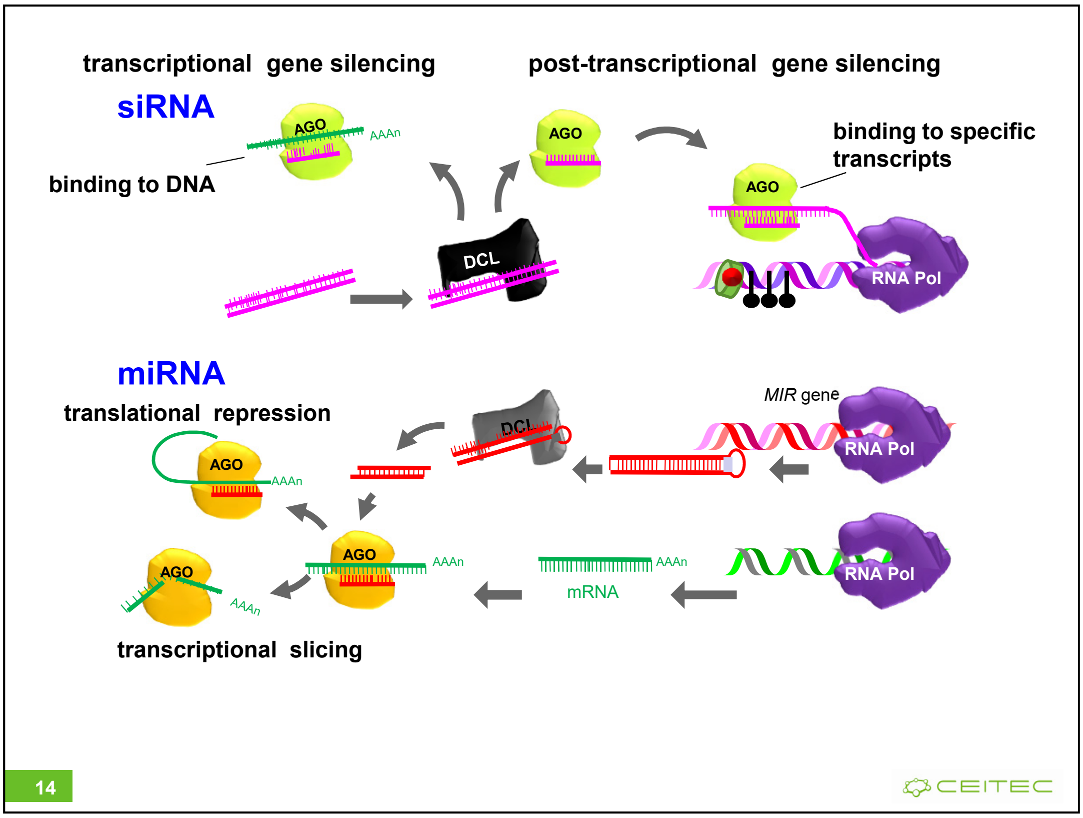
ARGONAUTE proteins were originally described as being important for plant development and for germline stem-cell division in *Drosophila melanogaster.*

ARGONAUTE proteins are classified into three paralogous groups: Argonaute-like proteins, which are similar to *Arabidopsis thaliana AGO1;* Piwi-like proteins, which are closely related to *D. melanogaster PIWI* (P-element induced wimpy testis); and the recently identified *Caenorhabditis elegans*-specific group 3 Argonautes.

Members of a new family of proteins that are involved in RNA silencing mediated by Argonaute-like and Piwi-like proteins are present in bacteria, archaea and eukaryotes, which implies that both groups of proteins have an ancient origin.

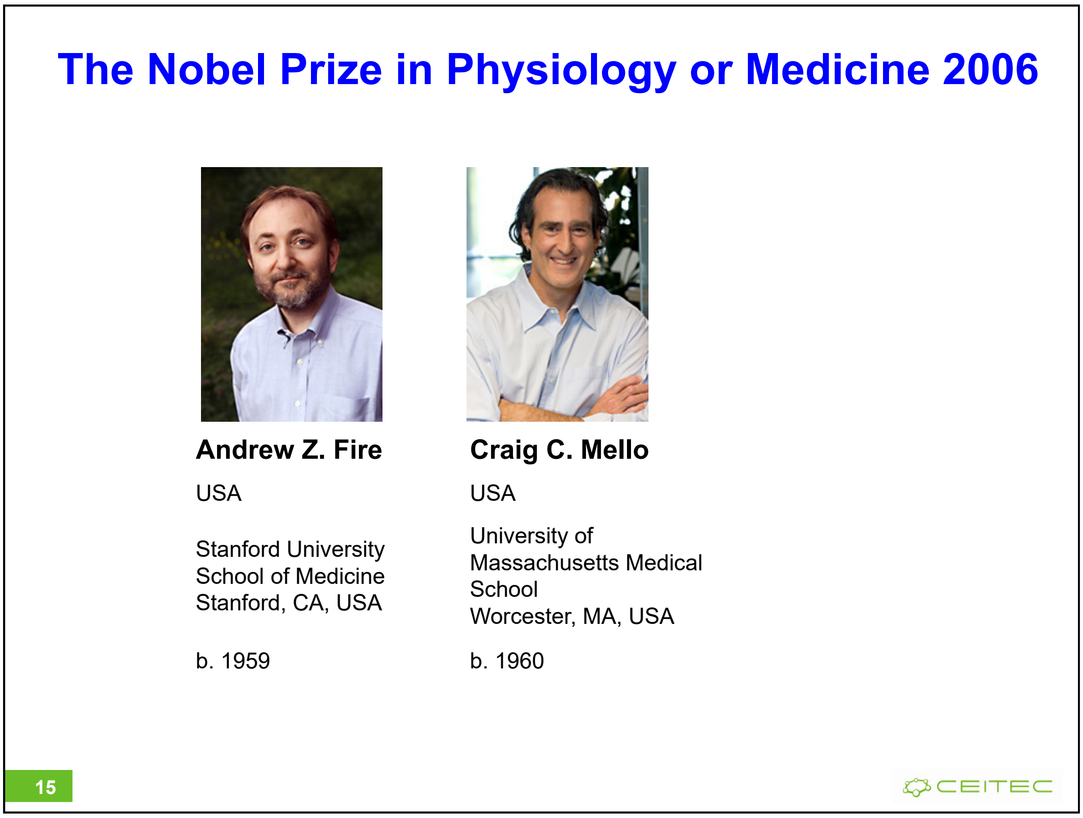
The number of Argonaute genes that are present in different species varies. There are 8 Argonaute genes in humans (4 Argonaute-like and 4 Piwi-like), 5 in the *D. melanogaster genome (2 Argonaute-like and* 3 Piwi-like), 10 Argonaute-like in *A. thaliana, only* 1 Argonaute-like in *Schizosaccharomyces pombe and at* least 26 Argonaute genes in *C. elegans (5 Argonaute-like,* 3 Piwi-like and 18 group 3 Argonautes).

http://youdpreferanargonaute.com/2009/06/



MicroRNAs are encoded by MIR genes, fold into hairpin structures that are recognized and cleaved by DCL (Dicer-like) proteins.

In summary, **siRNAs**-mediates silencing via post-transcriptional and transcriptional gene silencing, while **miRNAs** -mediate slicing of mRNA and translational repression.



The Nobel Prize in Physiology or Medicine 2006

Andrew Z. Fire

USA

Stanford University
School of Medicine
Stanford, CA, USA

b. 1959

Craig C. Mello

USA

University of
Massachusetts Medical
School
Worcester, MA, USA

b. 1960

In 2006, Andrwe Z. Fire and Craig C. Mello were honored by the Nobel prize “for their discovery of RNA interference - gene silencing by double-stranded RNA“.

The Nobel Prize in Physiology or Medicine 2006

?


Andrew Z. Fire

USA


Craig C. Mello

USA


David Baulcombe

UK


CORRESPONDENCE

NATURE|Vol 443|26 October 2006

RNAi Nobel ignores vital groundwork on plants

SIR — The Nobel prize, by recognizing the individuals behind breakthroughs, inspires all scientists to do great science. The discovery of RNA interference (RNAi) changed the face of gene regulation, a feat deservedly recognized with this year's Nobel Prize in Physiology or Medicine[1].

As undergraduates, we witnessed with great excitement the discovery of gene silencing. At that time, almost all research in that area was being conducted by plant

values at the centre of the prize and is sending a discouraging message, especially to young researchers.

Marc Bots*, Spencer Maughan†, Jeroen Nieuwland†

*Flanders Interuniversity Institute for Biotechnology, Technologiepark 927, BE-9052 Gent, Belgium

†Institute of Biotechnology, University of Cambridge, Cambridge CB2 1QT, UK

1. *Nature* 443, 488 (2006).
2. Baulcombe, D. C. *Plant Mol. Biol.* 32, 79–88 (1996).
3. Van der Krol, A. R. *et al. Plant Cell* 2, 291–299 (1990).
4. Voinnet, O. & Baulcombe, D. C. *Nature* 389, 553 (1997).
5. Metzlaff, M., O'Dell, M., Cluster, P. D. & Flavell, R. B. *Cell* 88, 845–854 (1997).

will not do so in the future. We believe that Iranian scientists can and will respond appropriately to the country's needs.

Kamran B. Lankarani

Ministry of Health and Medical Education of I. R. Iran, Tehran, I. R. Iran

Iran: productivity is not simple to evaluate

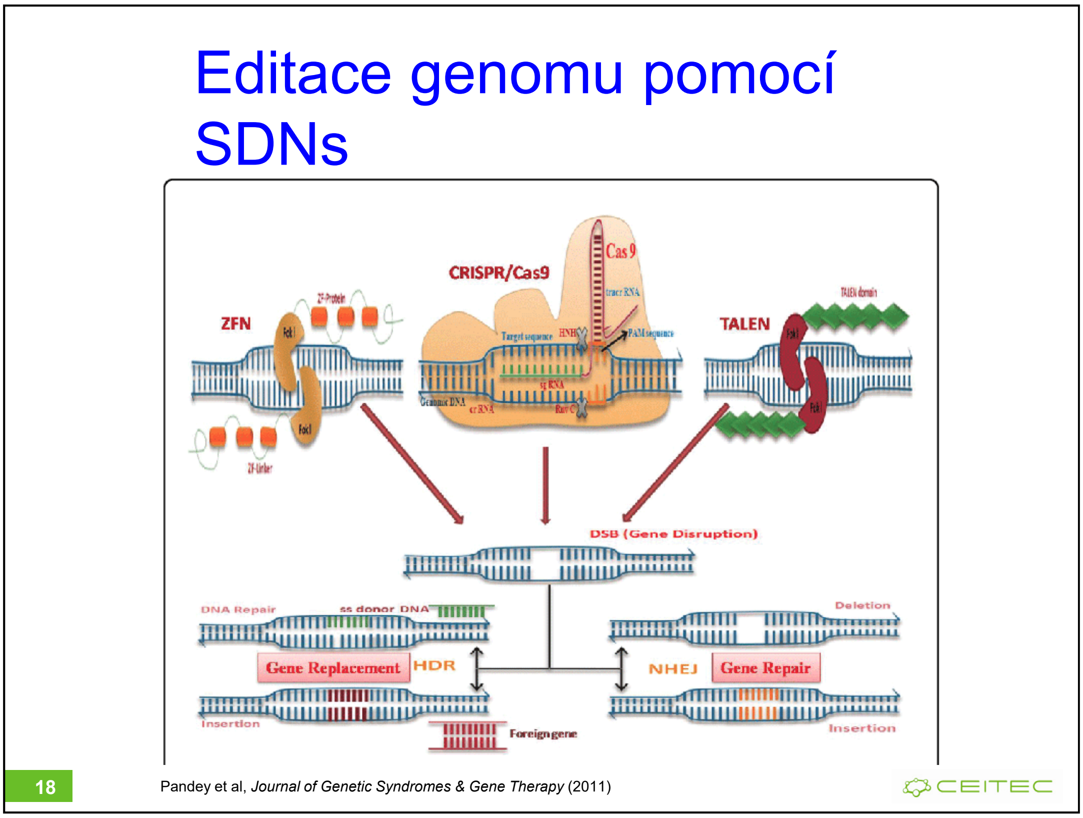
SIR — Eran Meshorer, in Correspondence ("Iran is sixth, not second, in Middle East publication list" *Nature* 443, 271; 2006), states:

CEITEC

In 2006, Andrwe Z. Fire and Craig C. Mello were honored by the Nobel prize "for their discovery of RNA interference - gene silencing by double-stranded RNA".

Osnova

- Umlčování genů pomocí RNA interference
    - Mechanismus RNAi
- Editace genomu
    - Princip editace genomu prostřednictvím místně specifických nukleáz (Site Directed Nucleases, SDNs)



Editace genomu pomocí SDNs

Pandey et al, *Journal of Genetic Syndromes & Gene Therapy* (2011)

CRISPR–Cas9, TALEN and ZFN mediated genome editing. (A) ZFN-fok1, TALEN domain and Cas9·sgRNA-induced DSBs can be repaired by either NHEJ or by HDR pathways. NHEJ mediated repair is highly efficient but error-prone process, which causes small insertions and/or deletions (indels) at the cleave site. HDR requires a donor DNA homologous template to repair the cleavage site and this process can be used to introduce specific point mutations, correction of mutation or to knock-in of corrected DNA sequences at cleavage site.

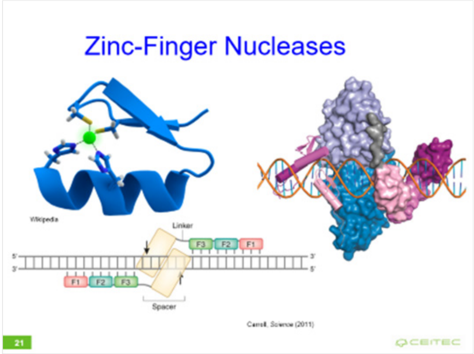
Abbreviations CRISPR: Clustered Regularly Interspaced Short Palindromic Repeats; HDR: Homology Directed Repair; NHEJ: Non- Homologous End Joining; TALENs: Transcription Activator-Like Effectors; ZFN: Zinc Finger Nucleases.

Osnova

- Umlčování genů pomocí RNA interference
    - Mechanismus RNAi
- Editace genomu
    - Princip editace genomu prostřednictvím místně specifických nukleáz (Site Directed Nucleases, SDNs)
    - Zinc-Finger Nucleases (ZFNs)

Zinc-Finger Nucleases - ZFNs

- **Místně specifické endonukleázy, schopné rozpoznat cílovou sekvenci prostřednictvím sady „zinkových prstů“**
- Každý zinkový „prst“ je schopen rozpoznat nukleotidový triplet
- Nukleázová doména funguje jako heterodimer – možnost zvýšení specifity navržením sady „prstů“ rozpoznávajících 9 bp z každé strany cílové sekvence
- **Nevýhody**
  - Špatně se „programuje“
  - Omezená specifita

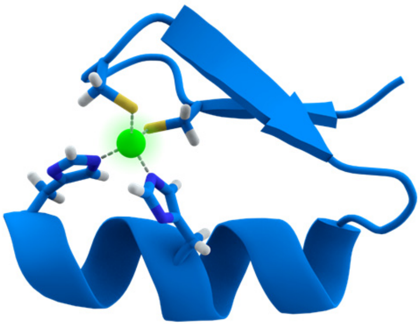


Zinc-Finger Nucleases

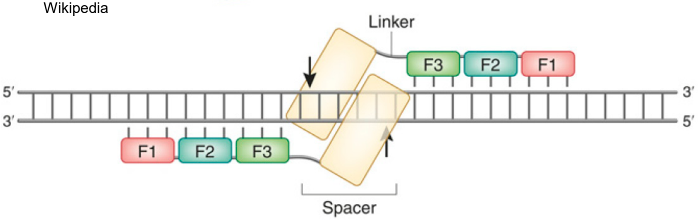
Carroll, Science (2011)

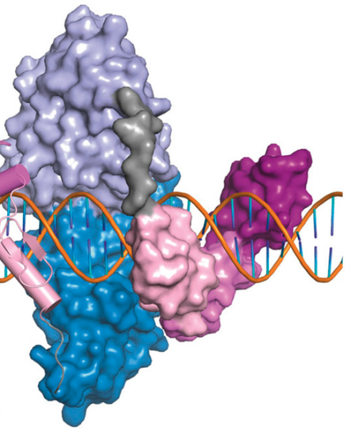
K zásadním nevýhodám ZFNs poatří omezená specifita – některé ‚prsty“ rozpoznávají více tripletů, pro některé trilplety naopak nejsou známy „prsty“ žádné.

Zinc-Finger Nucleases



Wikipedia

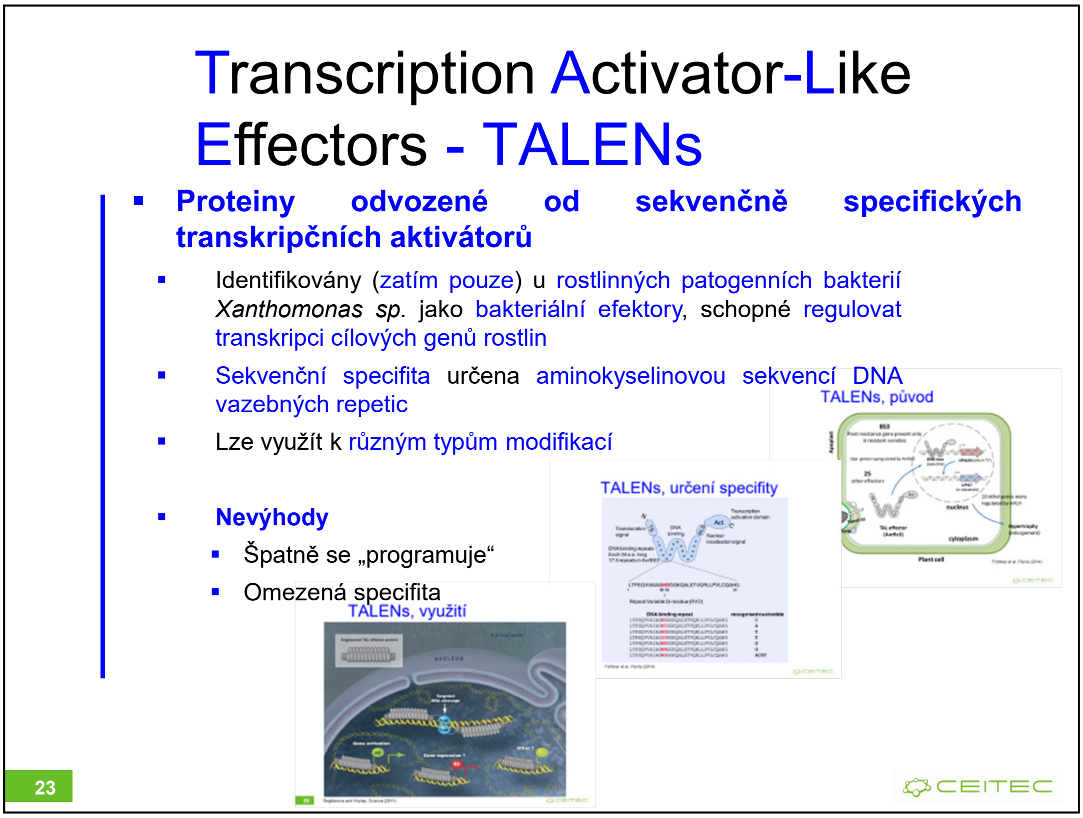




Carroll, *Science* (2011)

Osnova

- Umlčování genů pomocí RNA interference
  - Mechanismus RNAi
- Editace genomu
  - Princip editace genomu prostřednictvím místně specifických nukleáz (Site Directed Nucleases, SDNs)
  - Zinc-Finger Nucleases (ZFNs)
  - Transcription Activator-Like Effectors (TALENs)



Transcription Activator-Like Effectors - TALENs

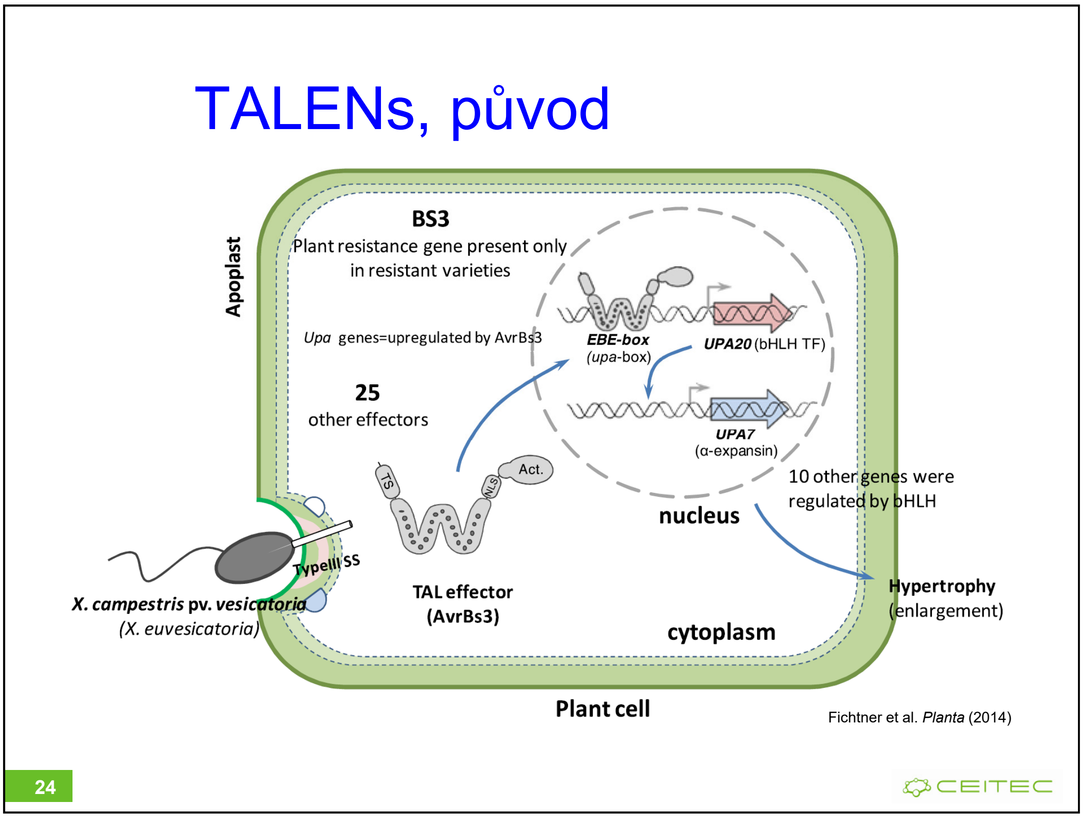
- **Proteiny odvozené od sekvenčně specifických transkripčních aktivátorů**
  - Identifikovány (zatím pouze) u rostlinných patogenních bakterií *Xanthomonas sp.* jako bakteriální efektory, schopné regulovat transkripci cílových genů rostlin
  - Sekvenční specifita určena aminokyselinovou sekvencí DNA vazebných repetic
  - Lze využít k různým typům modifikací
  - **Nevýhody**
    - Špatně se „programuje“
    - Omezená specifita

TALENs, původ

TALENs, určení specifity

TALENs, využití

Specifita některých RVDs je omezena na na in vitro podmínky, RVDs na 5’ konci vazebného motivu přispívají ke specifitě více než ty na 3’ konci (možnost vzniku „mismatches“ na 3’ konci, atd.).



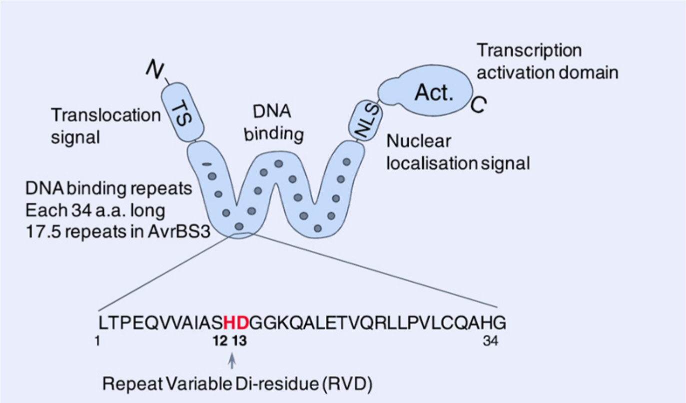
TALENs, původ

Fichtner et al. *Planta* (2014)

Discovery of *Xanthomonas* TAL effectors and the proposed mode of action of AvrBs3. AvrBs3 TAL effector protein is secreted into the plant cell via a Type III secretion system. The internal natural nuclear localization signal of AvrBs3 leads to import to the nucleus, where this TALE searches for the base pair sequence recognised by the internal RVD structure of the DNA binding region. Upon binding of the TAL effector to its recognised EBE-box (Effector Binding Element), also known as upa-box, transcription is initiated, leading to physiological effects in the infected plant cell such as hypertrophy.

Plant resistance to *Xanthomonas* derives from resistance (*R*) genes having a similar EBE-box and mimicking the natural TALE target site. This leads to enhanced expression of *R* genes upon infection.

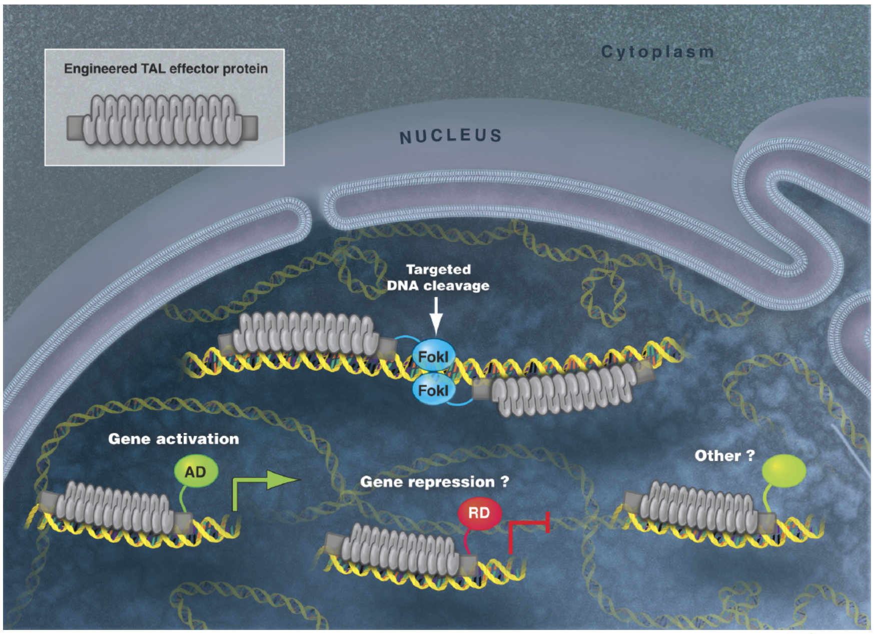
TALENs, určení specifity



DNA binding repeat	recognised nucleotide
LTPEQVVAIAS**HD**GGKQALETVQRLLPVLCQAHG	**C**
LTPEQVVAIAS**NI**GGKQALETVQRLLPVLCQAHG	**A**
LTPEQVVAIAS**NG**GGKQALETVQRLLPVLCQAHG	**T**
LTPEQVVAIAS**IG**GGKQALETVQRLLPVLCQAHG	**T**
LTPEQVVAIAS**NK**GGKQALETVQRLLPVLCQAHG	**G**
LTPEQVVAIAS**NN**GGKQALETVQRLLPVLCQAHG	**G**
LTPEQVVAIAS**NS**GGKQALETVQRLLPVLCQAHG	**ACGT**

Fichtner et al. *Planta* (2014)

TALENs, využití



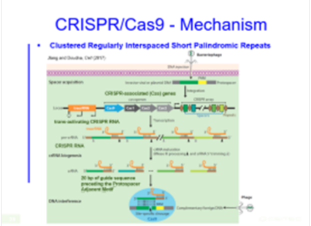
Bogdanove and Voytas, *Science* (2011)

Osnova

- Umlčování genů pomocí RNA interference
  - Mechanismus RNAi
- Editace genomu
  - Princip editace genomu prostřednictvím místně specifických nukleáz (Site Directed Nucleases, SDNs)
  - Zinc-Finger Nucleases (ZFNs)
  - Transcription Activator-Like Effectors (TALENs)
  - Clustered Regularly Interspaced Short Palindromic Repeats/Cas9 (CRISPR/Cas9)

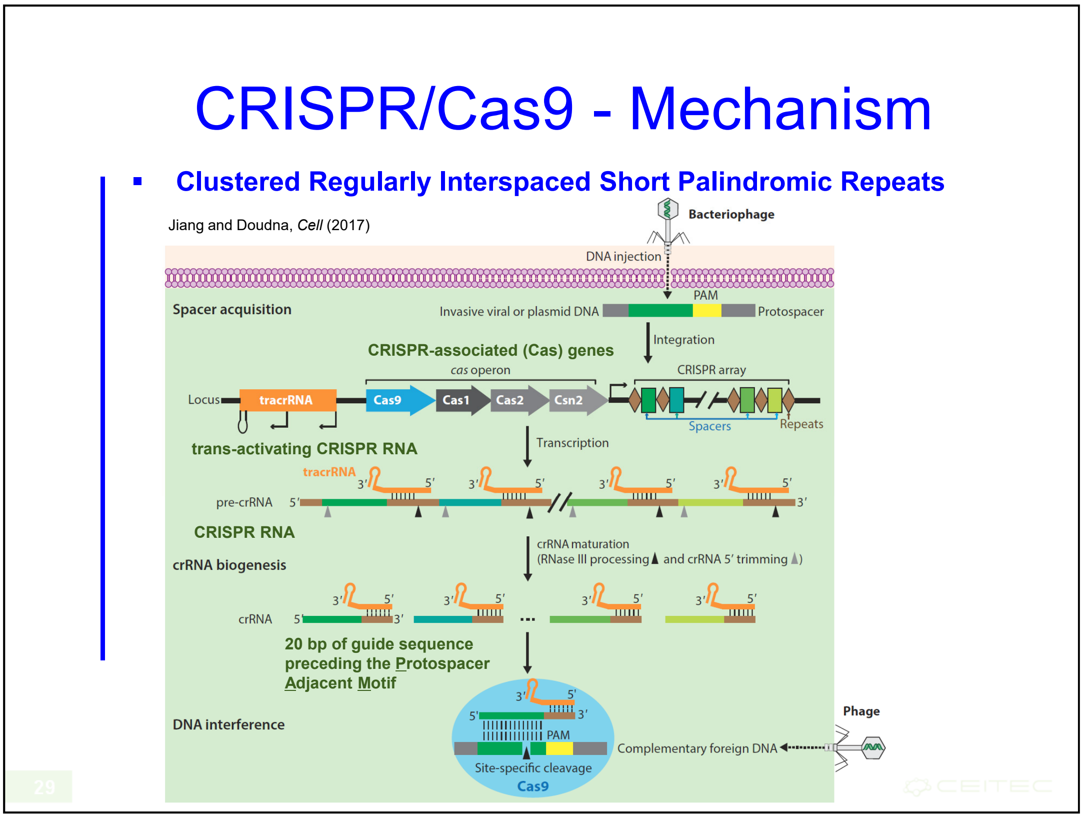
Clustered Regularly Interspaced Short Palindromic Repeats/Cas9 - CRISPR/Cas9

- **Objevena jako součást imunitního systému bakterií**
  - Principem je cílené začlenění cizorodé DNA (typicky fágové DNA) do specifických oblastí genomu bakterií
  - Po přepisu genů pro trans-activating CRISPR RNA (tracrRNA) a oblasti se začleněnými částmi cizorodé DNA a následném procesování vzniklé RNA dojde k tvorbě komplexu crRNA–tracrRNA
  - crRNA–tracrRNA váže Cas9 nukleázu a navádí ji na komplementární (cizorodou/fágovou) DNA, kterou pak Cas9 štěpí
  - crRNA–tracrRNA je v cíleném editování genomu nahrazována single guide RNA (sgRNA nebo také gRNA)
  - **Výhody**
    - Snadno se „programuje“
    - Značná specifita
    - Možná celá řada dalších aplikací


CRISPR/Cas9 - Mechanism

- Clustered Regularly Interspaced Short Palindromic Repeats

Specifita některých RVDs je omezena na na in vitro podmínky, RVDs na 5’ konci vazebného motivu přispívají ke specifitě více než ty na 3’ konci (možnost vzniku „mismatches“ na 3’ konci, atd.).

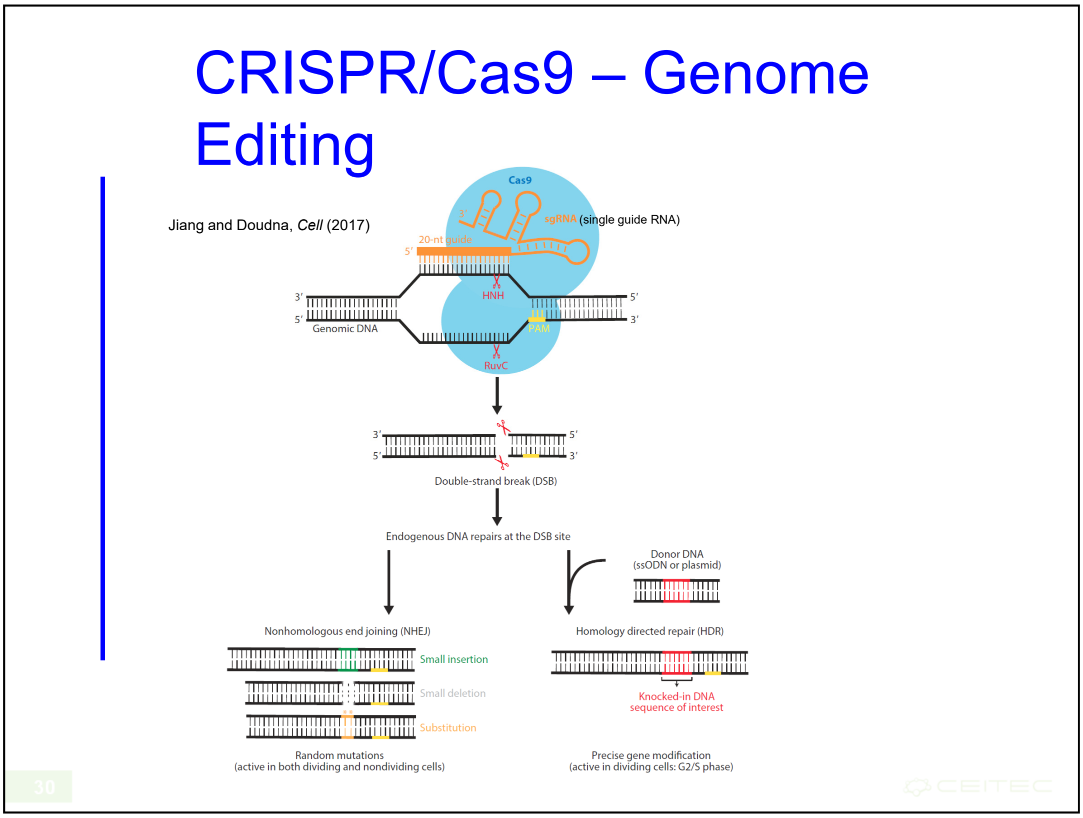


CRISPR/Cas9 - Mechanism

- **Clustered Regularly Interspaced Short Palindromic Repeats**

Jiang and Doudna, *Cell* (2017)

CRISPR–Cas9-mediated DNA interference in bacterial adaptive immunity. A typical CRISPR locus in a type II CRISPR–Cas system comprises an array of repetitive sequences (repeats, brown diamonds) interspaced by short stretches of nonrepetitive sequences (spacers, colored boxes), as well as a set of CRISPR-associated (cas) genes (colored arrows). Preceding the cas operon is the trans-activating CRISPR RNA (tracrRNA) gene, which encodes a unique noncoding RNA with homology to the repeat sequences. Upon phage infection, a new spacer (dark green) derived from the invasive genetic elements is incorporated into the CRISPR array by the acquisition machinery (Cas1, Cas2, and Csn2). Once integrated, the new spacer is cotranscribed with all other spacers into a long precursor CRISPR RNA (pre-crRNA) containing repeats (brown lines) and spacers (dark green, blue, light green, and yellow lines). The tracrRNA is transcribed separately and then anneals to the pre-crRNA repeats for crRNA maturation by RNase III cleavage. Further trimming of the 5’ end of the crRNA (gray arrowheads) by unknown nucleases reduces the length of the guide sequence to 20 nt. During interference, the mature crRNA–tracrRNA structure engages Cas9 endonuclease and further directs it to cleave foreign DNA containing a 20-nt crRNA complementary sequence preceding the PAM sequence. Asterisks denote conserved, key residues for Cas9-mediated DNA cleavage activity. Abbreviations: Arg, arginine-rich bridge helix; crRNA, CRISPR RNA; CTD, C-terminal domain; nt, nucleotide; NUC, nuclease lobe; PAM, protospacer adjacent motif; REC, recognition lobe; tracrRNA, trans-activating CRISPR RNA.



CRISPR/Cas9 – Genome Editing

Jiang and Doudna, *Cell* (2017)

The mechanism of CRISPR–Cas9–mediated genome engineering. The synthetic single guide (sgRNA) or crRNA–tracrRNA structure directs a Cas9 endonuclease to almost arbitrary DNA sequence in the genome through a user-defined 20-nt guide RNA sequence and further guides Cas9 to introduce a double-strand break (DSB) in targeted genomic DNA. The DSB generated by two distinct Cas9 nuclease domains is repaired by host-mediated DNA repair mechanisms. In the absence of a repair template, the prevalent error-prone nonhomologous end joining (NHEJ) pathway is activated and causes random insertions and deletions (indels) or even substitutions at the DSB site, frequently resulting in the disruption of gene function. In the presence of a donor template containing a sequence of interest flanked by homology arms, the error-free homology directed repair (HDR) pathway can be initiated to create desired mutations through homologous recombination, which provides the basis for performing precise gene modification, such as gene knock-in, deletion, correction, or mutagenesis. CRISPR–Cas9 RNA-guided DNA targeting can be uncoupled from cleavage activity by mutating the catalytic residues in the HNH and RuvC nuclease domains, making it a versatile platform for many other applications beyond genome editing. Abbreviations: crRNA, CRISPR RNA; nt, nucleotide; PAM, protospacer adjacent motif; sgRNA, single-guide RNA; tracrRNA, *trans*-activating CRISPR RNA.

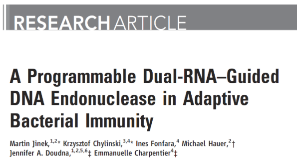
CRISPR/Cas9 – Nobel Prize in 2020!

Francisco Mojica

Emmanuelle Charpentier

Jenifer Doudna

Martin Jinek

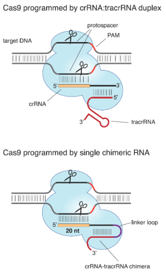


RESEARCH ARTICLE

A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity

Martin Jinek,[1,2*] Krzysztof Chylinski,[3,4*] Ines Fonfara,[4] Michael Hauer,[2†] Jennifer A. Doudna,[1,2,5,6‡] Emmanuelle Charpentier[4‡]

Jinek et al, *Science* (2012)



Základní koncepty

- RNAi
  - Přirozený způsob regulace genové exprese, vyžadující vlastní geny a vysvětlující přítomnost velkého množství DNA nekódující proteiny
  - Možno využít pro cílenou regulaci genové exprese
- Editování genomu
  - Sekvenčně-specifické modifikace genomu s velkou přesností
  - Umožňuje jak vznik náhodných mutací v daném lokusu, tak
  - Cílené vkládání definovaných sekvencí – ideální nástroj pro cílené modifikace genomu včetně genové terapie
  - CRISPR/Cas9 otevřel cestu snadné, rychlé a zejména přesné editaci genomu a dalším odvozeným modifikacím s velkým aplikačním potenciálem

Diskuse