

# CG920 Genomics

## Lesson 5

### RNA Interference and Genome Editing

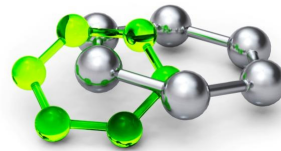
Jan Hejátko

**Functional Genomics and Proteomics of Plants,**  
CEITEC - Central European Institute of Technology  
And

**National Centre for Bimolecular Research,**  
Faculty of Science,

MUNI  
SCI

Masaryk University, Brno  
[hejatko@sci.muni.cz](mailto:hejatko@sci.muni.cz), [www.ceitec.eu](http://www.ceitec.eu)



# Outline

- Knocking-down the genes using RNA interference
  - Mechanism of RNAi
- Genome Editing
  - Principle of genome editing using Site Directed Nucleases, (SDNs)
  - Zinc-Finger Nucleases (ZFNs)
  - Transcription Activator-Like Effectors (TALENs)
  - Clustered Regularly Interspaced Short Palindromic Repeats/Cas9 (CRISPR/Cas9)

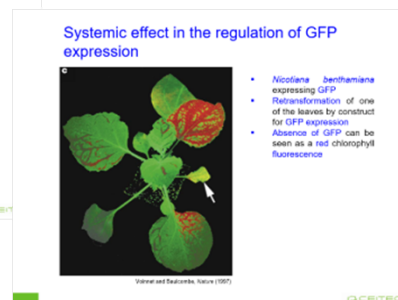
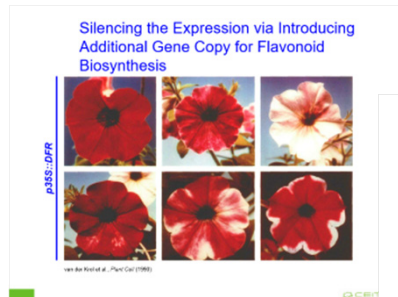
# Outline

- Knocking-down the genes using RNA interference
  - Mechanism of RNAi



# RNA interference

- **Molecular mechanism of post-transcriptional gene silencing (PTGS)**
  - RNAi discovered in plants, later in *Coenorhabditis elegans*
    - In plants identified as „sense effect“ in systemic negative regulation of gene activity



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].

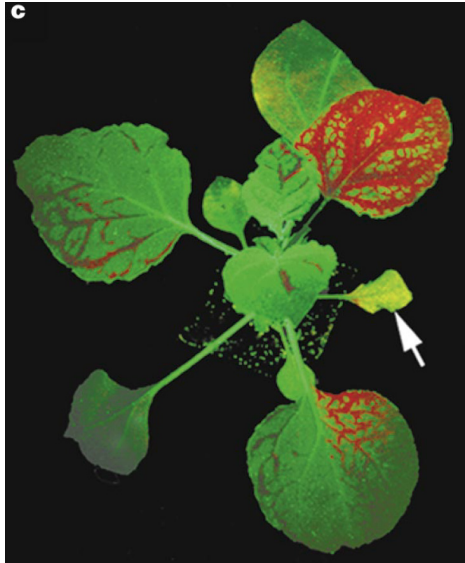
## Silencing the Expression via Introducing Additional Gene Copy for Flavonoid Biosynthesis



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.

## Systemic effect in the regulation of GFP expression



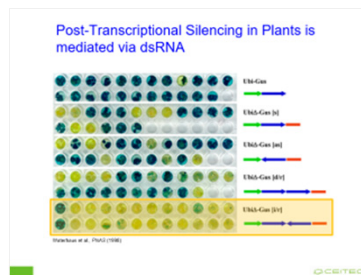
Voinnet and Baulcombe, *Nature* (1997)

- *Nicotiana benthamiana* expressing GFP
- Retransformation of one of the leaves by construct for GFP expression
- Absence of GFP can be seen as a red chlorophyll fluorescence

We studied *Nicotiana benthamiana* plants carrying a jellyfish green fluorescent protein (GFP) transgene<sup>5</sup>. 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

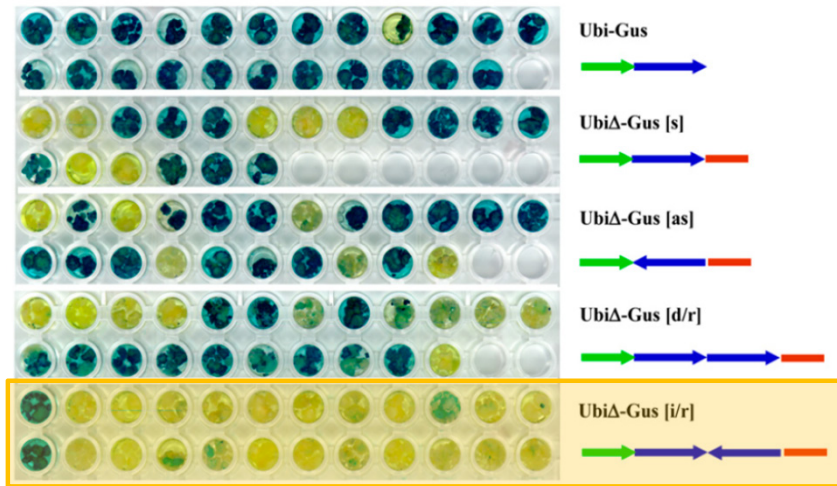
- **Molecular mechanism of post-transcriptional gene silencing (PTGS)**
  - RNAi discovered in plants, later in *Coenorhabditis elegans*
    - In plants identified as „sense effect“ in systemic negative regulation of gene activity
    - Gene silencing induced via both sense and anti-sense RNA
    - dsRNA induced gene silencing approx. 100x more efficiently



CEITEC

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].

## Post-Transcriptional Silencing in Plants is mediated via dsRNA



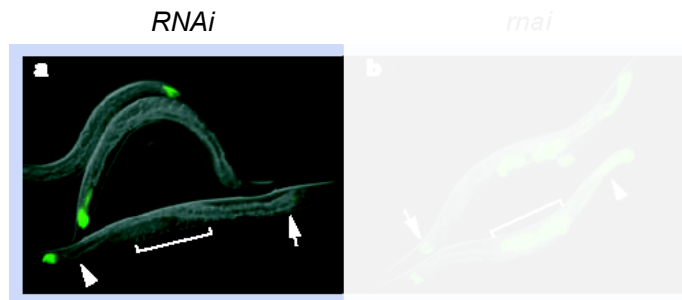
Waterhaus et al., *PNAS* (1998)

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

- **Molecular basis of posttranscriptional gene silencing (PTGS)**
  - dsRNA induction is dependent on own genes
  - gene identification using **forward genetics**

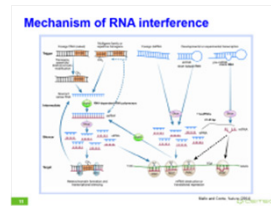


Mello and Conte, *Nature* (2004)

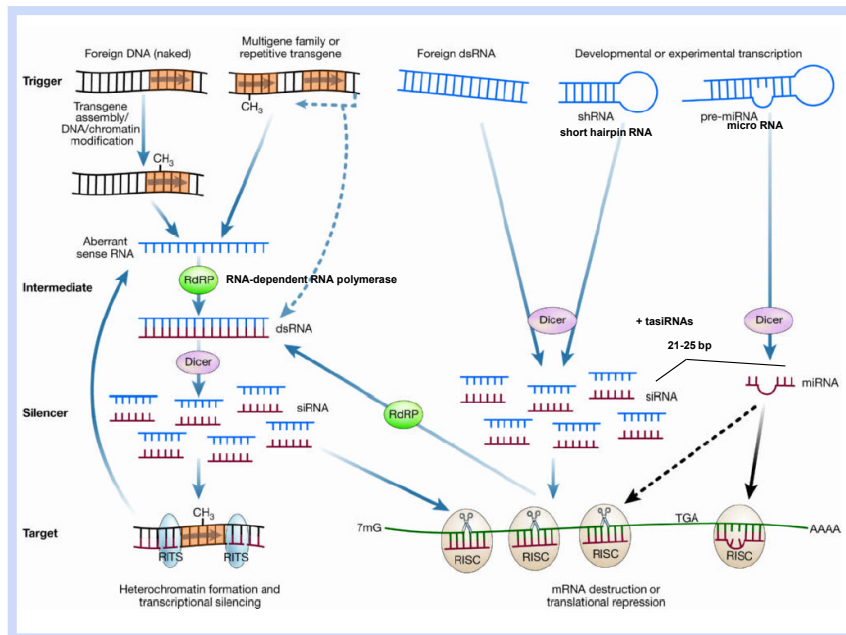
# RNA interference

## ▪ Molecular basis of posttranscriptional gene silencing (PTGS)

- RNAi found in *Coenorhabditis elegans* and in plants
- It is a **natural mechanism** of regulation of gene expression in all eukaryotes
- The principle is **creating dsRNA**, which can be triggered in several ways:
  - By presence of **foreign „aberrant“ DNA**
  - **Specific transgenes** containing **inverted repeats** of the cDNA parts
  - Transcription of own genes for **shRNA** (short hairpin RNA) or **miRNA** (micro RNA, endogenous hairpin RNA)
- dsRNA is processed by enzyme complex (DICER), which leads to the formation of **siRNA** (short interference RNA), which is then bound to enzyme complex **RITS** (RNA-induced transcriptional silencing complex) or **RISC** (RNA-induced silencing complex)
- **RISC** mediates either **degradation of mRNA** (in case of full similarity of siRNA and the target mRNA) or leads only to **termination of translation** (in case of incomplete homology, e.g. as in the case of miRNA)
- **RITS** mediates **reorganization of genomic DNA** (heterochromatin formation and inhibition of transcription)



# Mechanism of RNA interference



11

Mello and Conte, *Nature* (2004)

CEITEC

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).

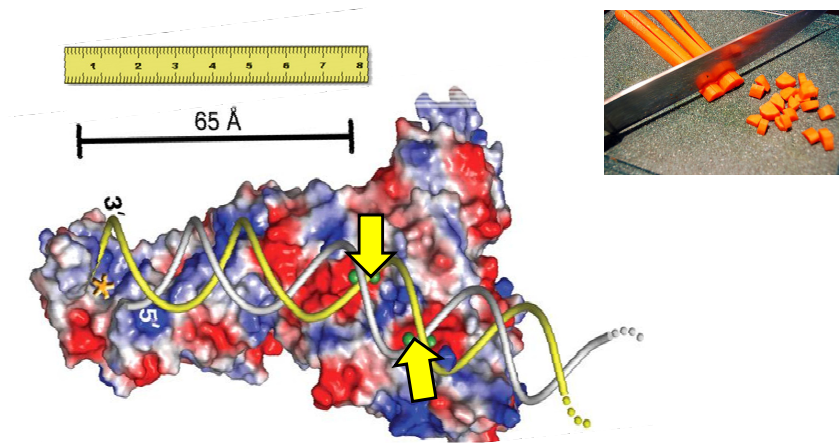
In addition to siRNAs, there are trans-acting siRNAs (tasiRNAs) that are a special class of siRNAs that appear to function in development (much like miRNAs) but have a unique mode of origin involving components of both miRNA and siRNA pathways.

Developmental regulations via miRNAs are more often used in animals than 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 complex (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

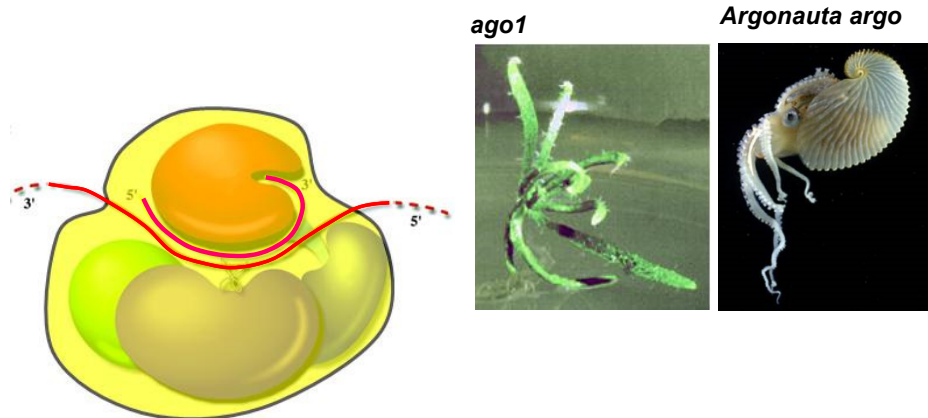
CEITEC

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



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.

13

CEITEC

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).

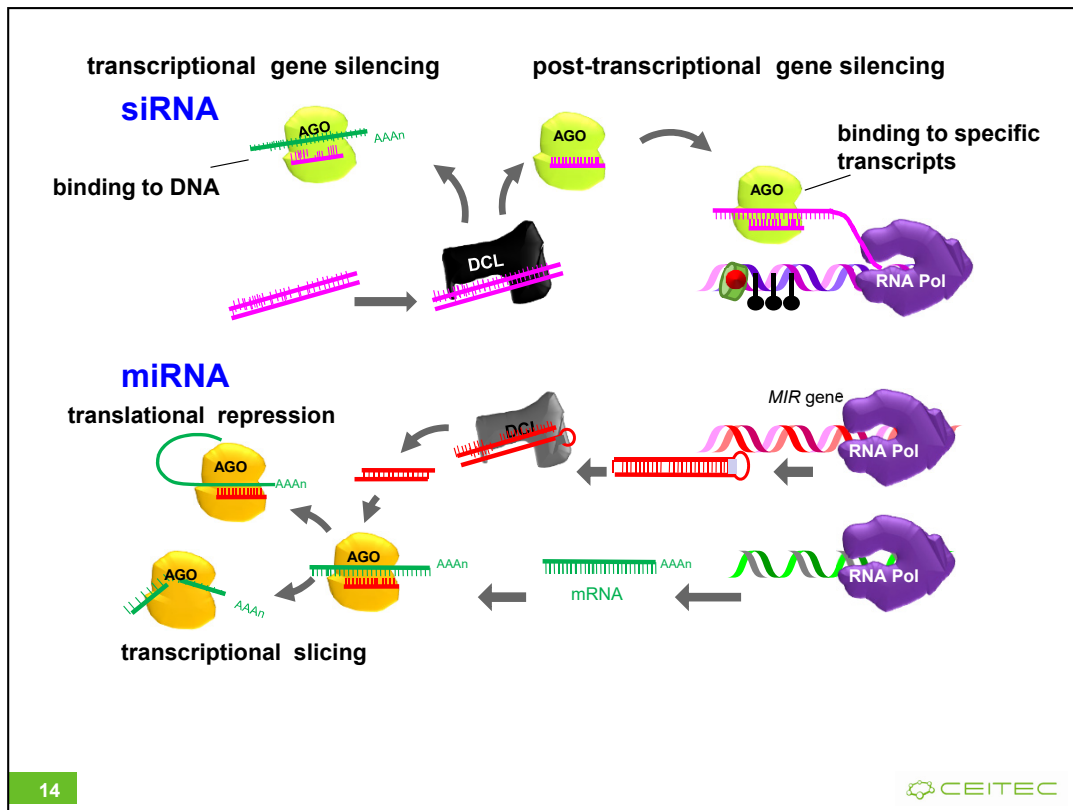
ARGONAUTE proteins were originally described as being important for plant development and for germline stem-cell division in *Drosophila melanogaster*.

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, Andrew 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<sup>1</sup>.

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-9002 Ghent, 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. Valmet, O. & Baulcombe, D. C. *Nature* 389, 553 (1997).
5. Matzke, M., O'Dell, M., Chatter, P. D. & Flavell, R. R. *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:

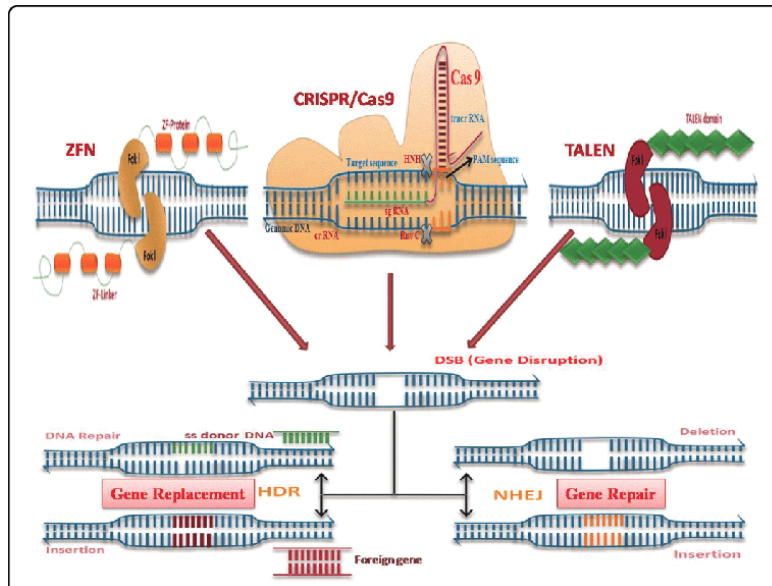
In 2006, Andrew Z. Fire and Craig C. Mello were honored by the Nobel prize "for their discovery of RNA interference - gene silencing by double-stranded RNA".



# Outline

- Knocking-down the genes using RNA interference
  - Mechanism of RNAi
- Genome Editing
  - Principle of genome editing using Site Directed Nucleases, (SDNs)

# Genome Editing via SDNs



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.

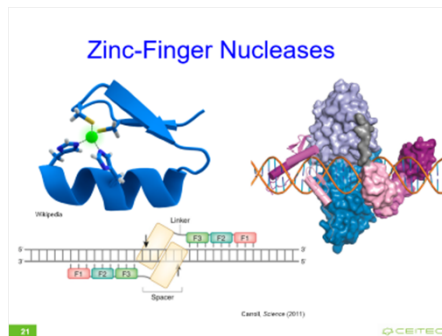
Abbreviations CRISPR: Clustered Regularly Interspaced Short Palindromic Repeats; HDR: Homology Directed Repair; NHEJ: Non- Homologous End Joining; TALENs: Transcription Activator-Like Effectors; ZFN: Zinc Finger Nucleases.

# Outline

- Knocking-down the genes using RNA interference
  - Mechanism of RNAi
- Genome Editing
  - Principle of genome editing using Site Directed Nucleases, (SDNs)
  - Zinc-Finger Nucleases (ZFNs)

# Zinc-Finger Nucleases - ZFNs

- **Sequence-specific endonucleases recognizing the target sequence via set of “zinc fingers”**
  - Each zinc „finger“ is recognizing nucleotide triplet
  - Nuclease domain acts as heterodimer – possibility to enhance the specificity by designing the set of „fingers“ recognizing 9 bp on both sides of the target sequence
- **Shortcomings**
  - Difficult to “program”
  - Delimited specificity



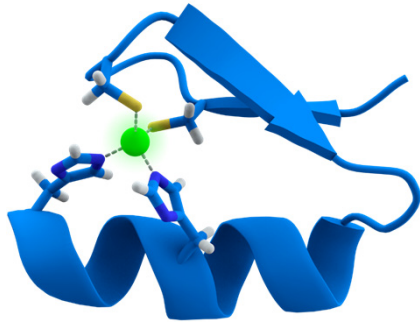
20

CEITEC

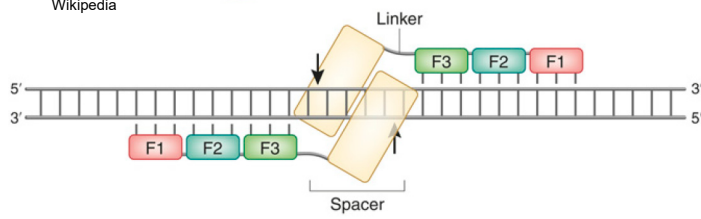
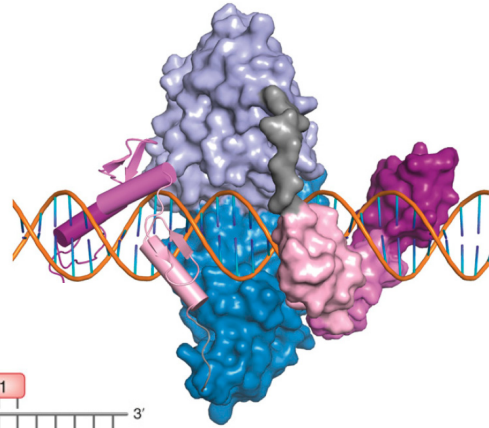
CEITEC

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

# Zinc-Finger Nucleases



Wikipedia



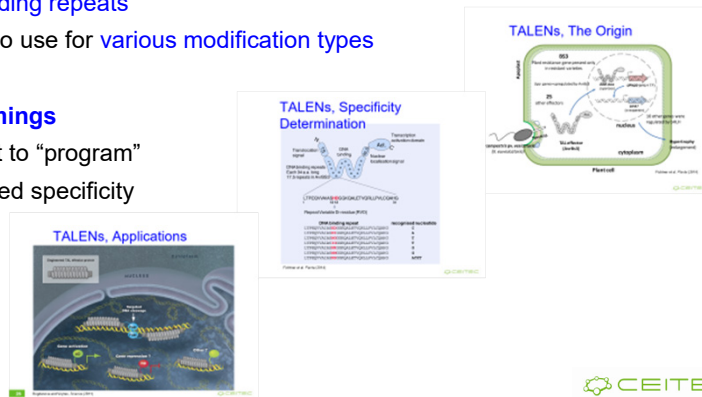
Carroll, *Science* (2011)

# Outline

- Knocking-down the genes using RNA interference
  - Mechanism of RNAi
- Genome Editing
  - Principle of genome editing using Site Directed Nucleases, (SDNs)
  - Zinc-Finger Nucleases (ZFNs)
  - Transcription Activator-Like Effectors (TALENs)

# Transcription Activator-Like Effectors - TALENs

- **Proteins derived from sequence-specific transcription activators**
  - Identified (so far only) in plant pathogenic bacteria *Xanthomonas sp.* as bacterial effectors, able to control the transcription of target genes in plants
  - Sekvenční specificity determined by aminoacid sequence of DNA-binding repeats
  - Possible to use for various modification types
- **Shortcomings**
  - Difficult to “program”
  - Delimited specificity

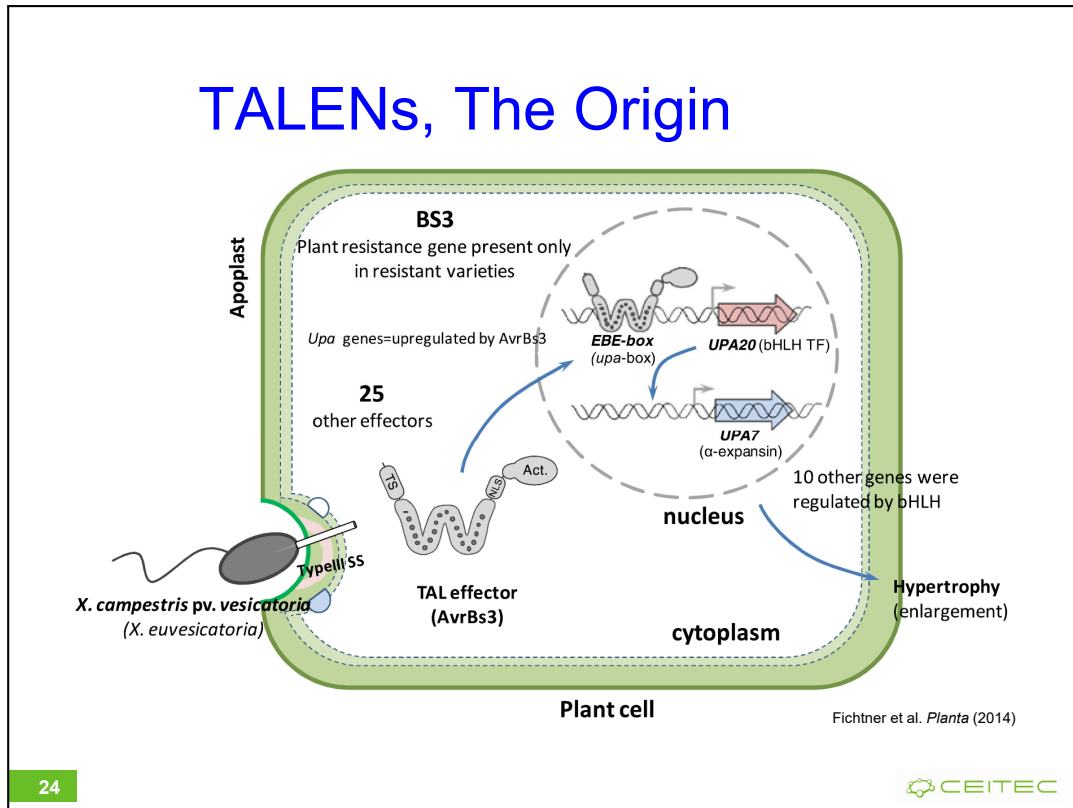


23

CEITEC

Specifita některých RVDs je omezena 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, The Origin

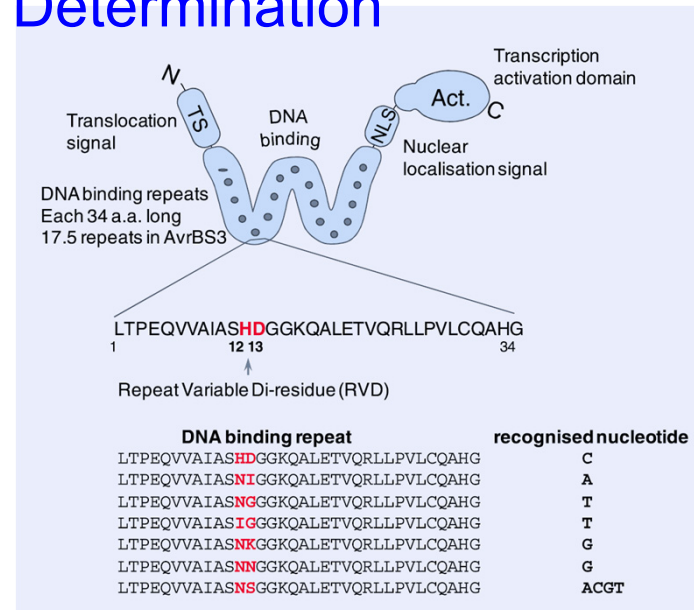


Discovery of *Xanthomonas* TAL effectors and the proposed mode of action of AvrBs3. AvrBs3 TAL effector protein is secreted into the plant cell via a Type III secretion system. The internal natural nuclear localization signal of AvrBs3 leads to import to the nucleus, where this TALE searches for the 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.

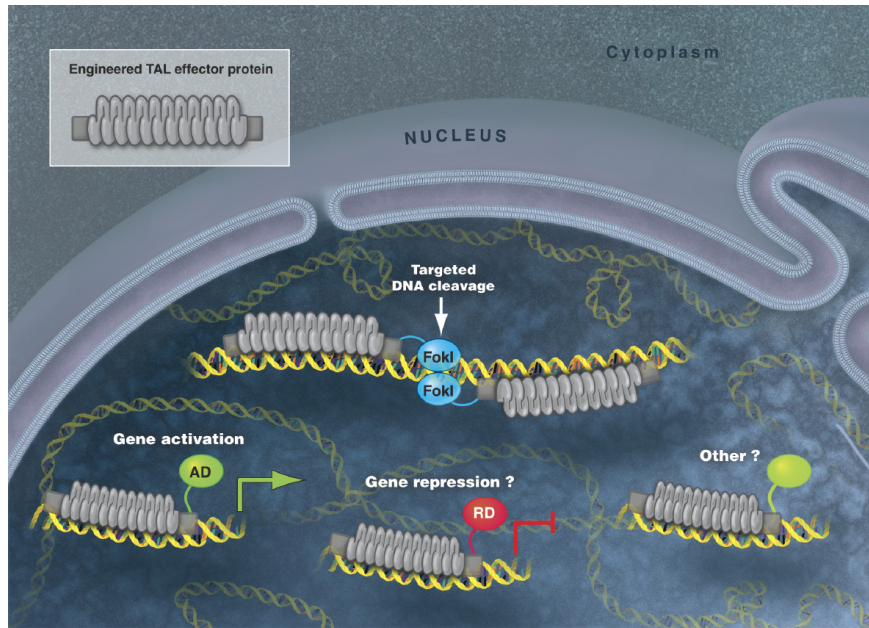


# TALENs, Specificity Determination



Fichtner et al. *Planta* (2014)

# TALENs, Applications

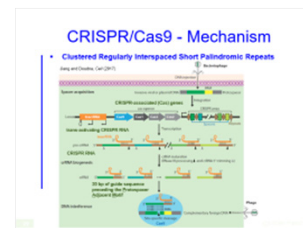


# Outline

- Knocking-down the genes using RNA interference
  - Mechanism of RNAi
- Genome Editing
  - Principle of genome editing using 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

- **Discovered as a mechanism of bacterial immune system**
  - The principle is **targeted insertion of foreign DNA** (typically **phage DNA**) into **specific bacterial genome loci**
  - Transcription of **trans-activating CRISPR RNA (tracrRNA)** and the region with **inserted foreign DNA** followed by **RNA processing** allows formation of **crRNA–tracrRNA complex**
  - **crRNA–tracrRNA** binds **Cas9** nuclease, targeting it to **complementary (foreign/phage) DNA**, that is then digested
  - crRNA–tracrRNA is in the targeted genome editing replaced by a **single guide RNA (sgRNA or gRNA)**
  - **Advantages**
    - Easy to „program“
    - High specificity
    - Number of further applications possible



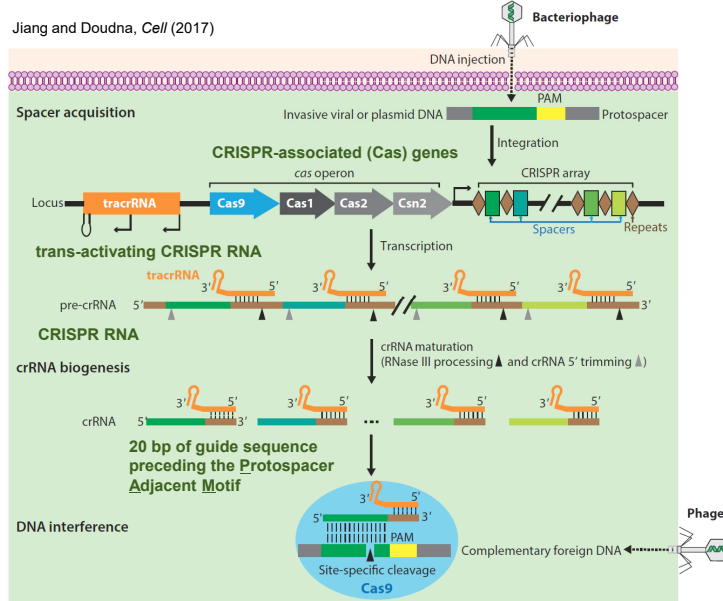
CEITEC

28

Specifita některých RVDs je omezena 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

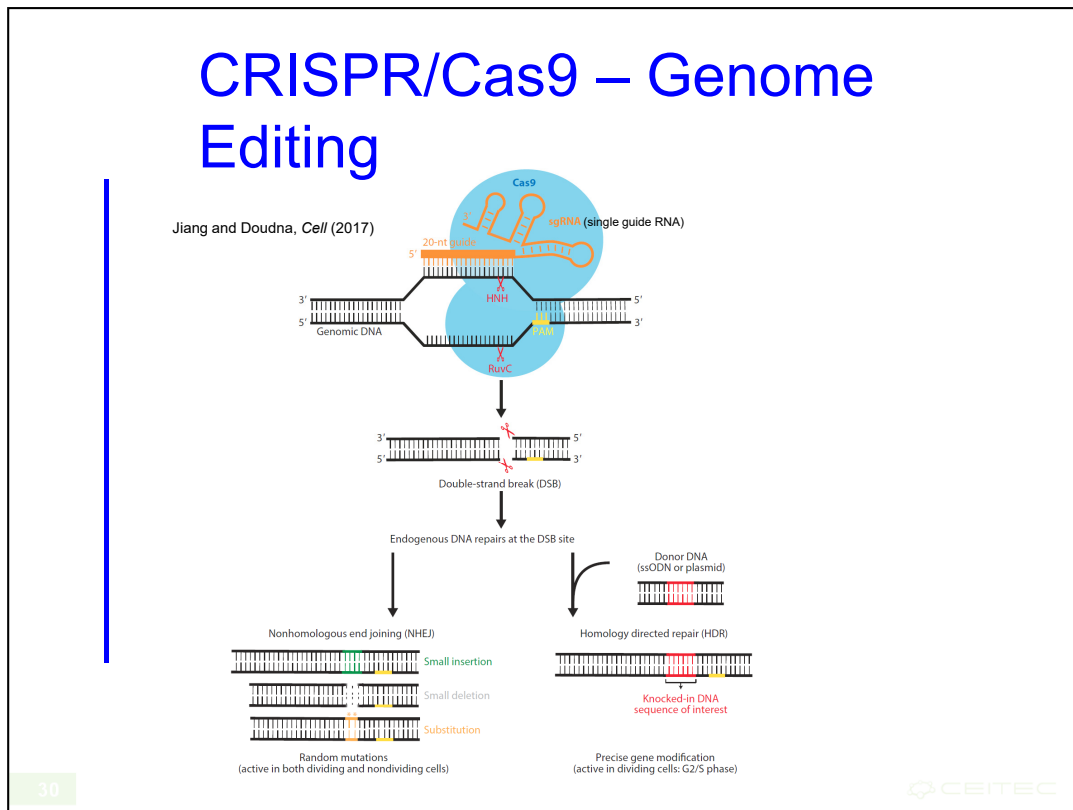
- **Clustered Regularly Interspaced Short Palindromic Repeats**



29

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



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

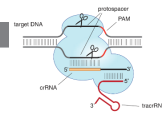
## RESEARCH ARTICLE

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

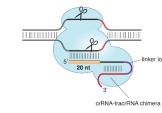
Martin Jinek,<sup>1,2\*</sup> Krzysztof Chylinski,<sup>1,2\*</sup> Ines Fonfara,<sup>3</sup> Michael Haue,<sup>2,4</sup> Jenifer A. Doudna,<sup>2,3,5,6</sup> Emmanuelle Charpentier<sup>1,2</sup>

Jinek et al, *Science* (2012)

Cas9 programmed by crRNA:tracrRNA duplex



Cas9 programmed by single chimeric RNA



# Key concepts

- RNAi
  - Natural mechanism controlling gene expression, partially explaining existence of large amount of non-coding DNA in various genomes
  - Possible use as a tool for specific gene expression control
- Genome editing
  - Sequence-specific high-precision genome modifications
  - Allows generation of both random mutations in a specific locus, as well as
  - introgression/replacement of defined sequence in the target locus, including gene therapy
  - CRISPR/Cas9 paved the way for easy, fast and accurate genome editing and further derived modifications



# Discussion