Bi8940 Developmental biology
Lesson 10

Regulation of Gene Expression during Development

Jan Hejátko
Laboratory of Molecular Plant Physiology,
Department of Functional Genomics and Proteomics,
and
Functional Genomics and Proteomics of Plants
CEITEC
Masaryk University,
Brno, Czech Republic
hejatko@sci.muni.cz, www.ceitec.eu

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ
Tato prezentace je spolufinancována
Evropským sociálním fondem
a státním rozpočtem České republiky

Outline of Lesson 10

Regulation of Gene Expression during Development

- Overview of levels of gene expression regulation
- Transcriptional gene regulation
  - Modification of the chromatin structure and DNA methylation
  - Transcriptional activation
- Post-transcriptional gene regulation
  - Splicing of hnRNA
  - Translation initiation
  - Localization of mRNA
  - Protein localization
- RNA interference
  - Identification and mechanism of gene expression regulation via RNA interference
  - siRNA-mediated silencing
  - miRNA-mediated silencing

EVROPSKÁ UNIE
MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY
OP Vzdělávání pro konkurenceschopnost

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

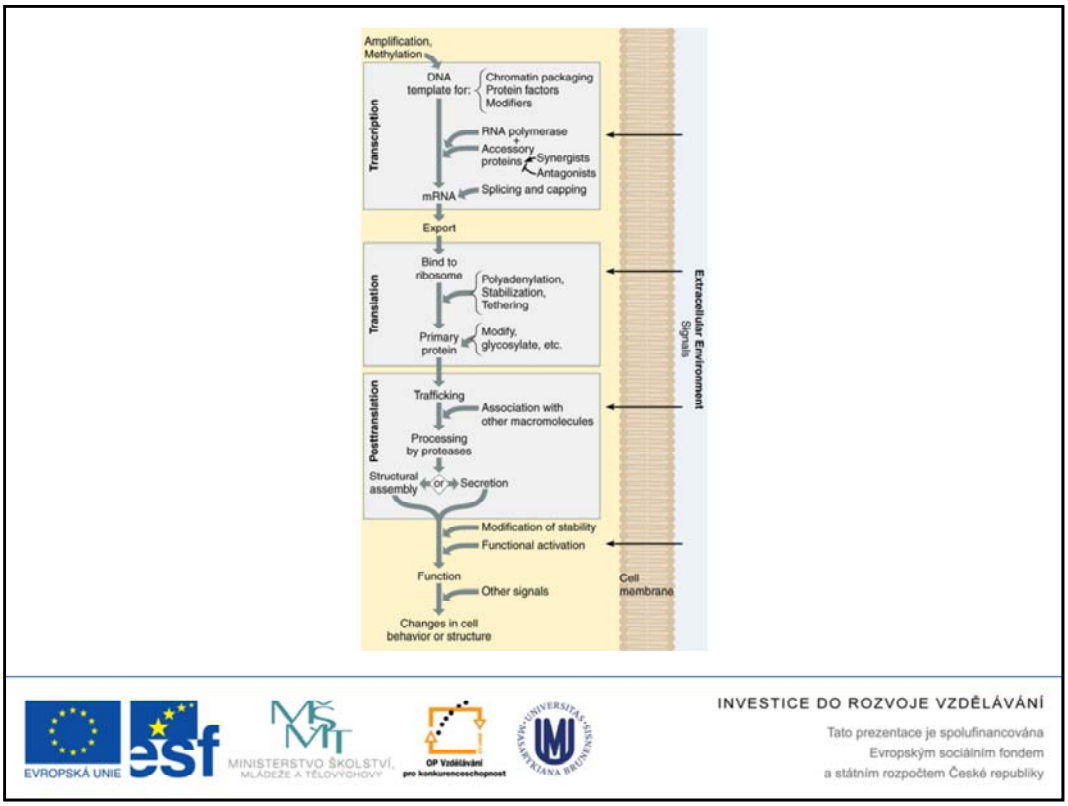
Tato prezentace je spolufinancována Evropským sociálním fondem a státním rozpočtem České republiky

Outline of Lesson 10

Regulation of Gene Expression during Development

- Overview of levels of gene expression regulation

EVROPSKÁ UNIE
MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY
OP Vzdělávání pro konkurenceschopnost
INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ
Tato prezentace je spolufinancována Evropským sociálním fondem a státním rozpočtem České republiky



The regulation of gene expression during development may occur at different levels. Most of the regulations occurs at the transcriptional levels, however, the later ("downstream") type of posttranscriptional regulations e.g. regulation of translation, posttranslational modifications of the proteins, their transport etc. are also important.

Recently, very important type of posttranscriptional regulation was found to be mediated by small RNAs. The molecular mechanisms and their importance for the diverse developmental processes in eukaryotes with emphasis on the development regulation in plants will be discussed in the second part of this lesson.

The regulatory events are diagrammed here as a linear sequence, which, however, is rarely the case. More often, the individual regulatory events occur in a maze of networking interactions.

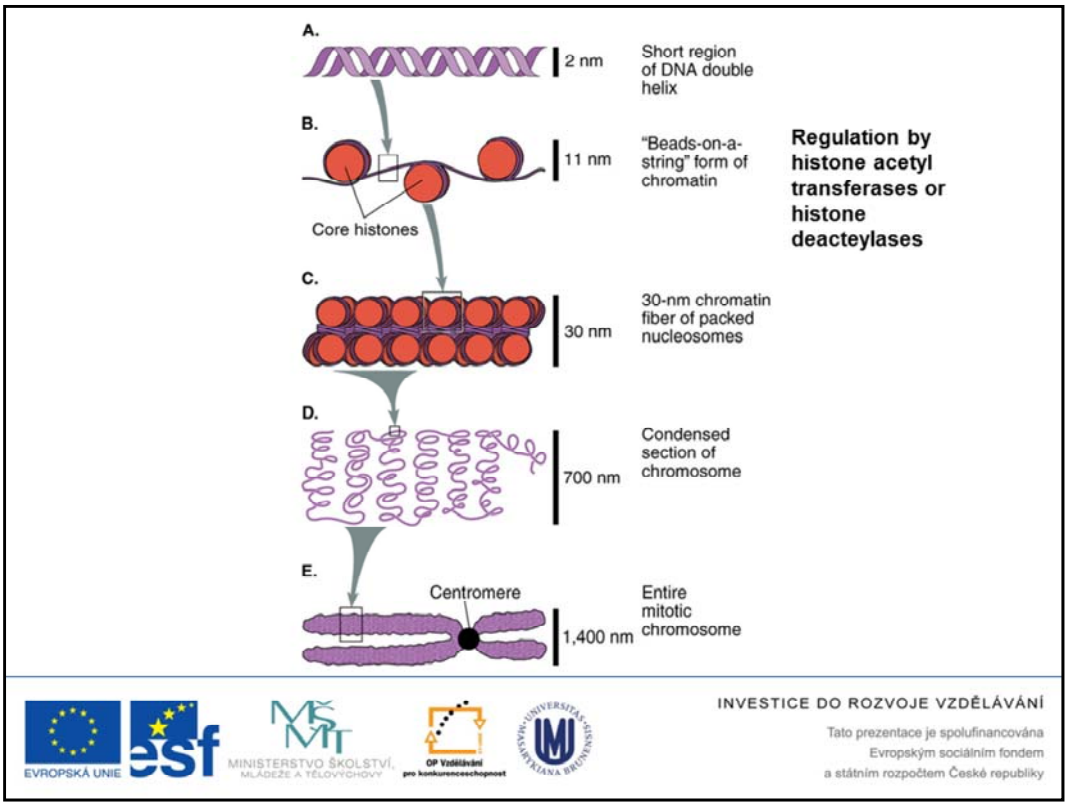
Outline of Lesson 10

Regulation of Gene Expression during Development

- Overview of levels of gene expression regulation
- Transcriptional gene regulation
  - Modification of the chromatin structure and DNA methylation

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

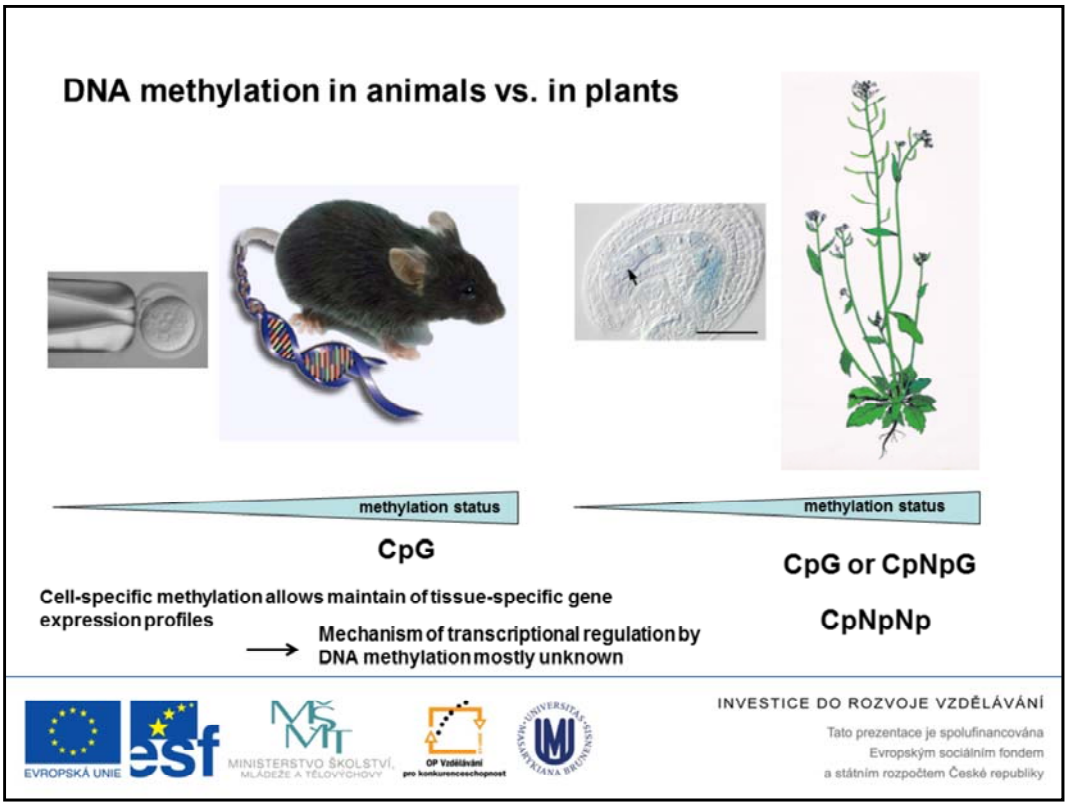
Tato prezentace je spolufinancována
Evropským sociálním fondem
a státním rozpočtem České republiky



Regulation of the chromatin structure represents one of the very basal gene expression regulatory levels. Chromatin is a substrate for DNA-dependent RNA polymerases that transcript the DNA encoded information into the "words and sentences" of RNA.

Regulation of chromatin structure and its accessibility to DNA-dependent RNA polymerases depends on many factors, one of the most important is the regulation of chromatin binding to nucleosomes and chromatin methylation.

Regulation of chromatin interaction with histones, the positively charged proteins forming the core of nucleosomes, is performed via modification of acetylation status of the N-terminal portion of histones, especially histones H3 and H4. This occurs via action of histone acetyl transferases or histone deacteylases.



Modification of the chromatin methylation is performed via DNA methyltransferases.

Interestingly, there is difference in the methylation in animals and in plants.

In animals, the methylation occurs mostly on the cytosine that occurs next to guanosine (the sequence is denoted as CpG). In mammals, 60-90% of all CpGs are methylated.

In plants, cytosines are methylated both symmetrically (CpG or CpNpG) and asymmetrically (CpNpNp), where N is any nucleotide.

Methylation status is usually “reset” in the zygote and is reconstituted during development again. E.g. the methylation is very low in the mouse embryo at the blastula stage, however, DNA derived from later stages when organogenesis is initiated is substantially more modified by methylation.

DNA methylation also stably alters the gene expression pattern in cells such that cells can "remember where they have been"; in other words, cells programmed to be pancreatic islets during embryonic development remain pancreatic islets throughout the life of the organism without continuing signals telling them that they need to remain islets.

DNA methylation is involved in the genomic imprinting, i.e. the genes originating from both parents are often diversely methylated, which results into differential expression of parental genomes (for the importance of the imprinting in the parental conflict and epigenetics, see the lecture “Bi0580 Developmental genetics” by prof. Vyskot).

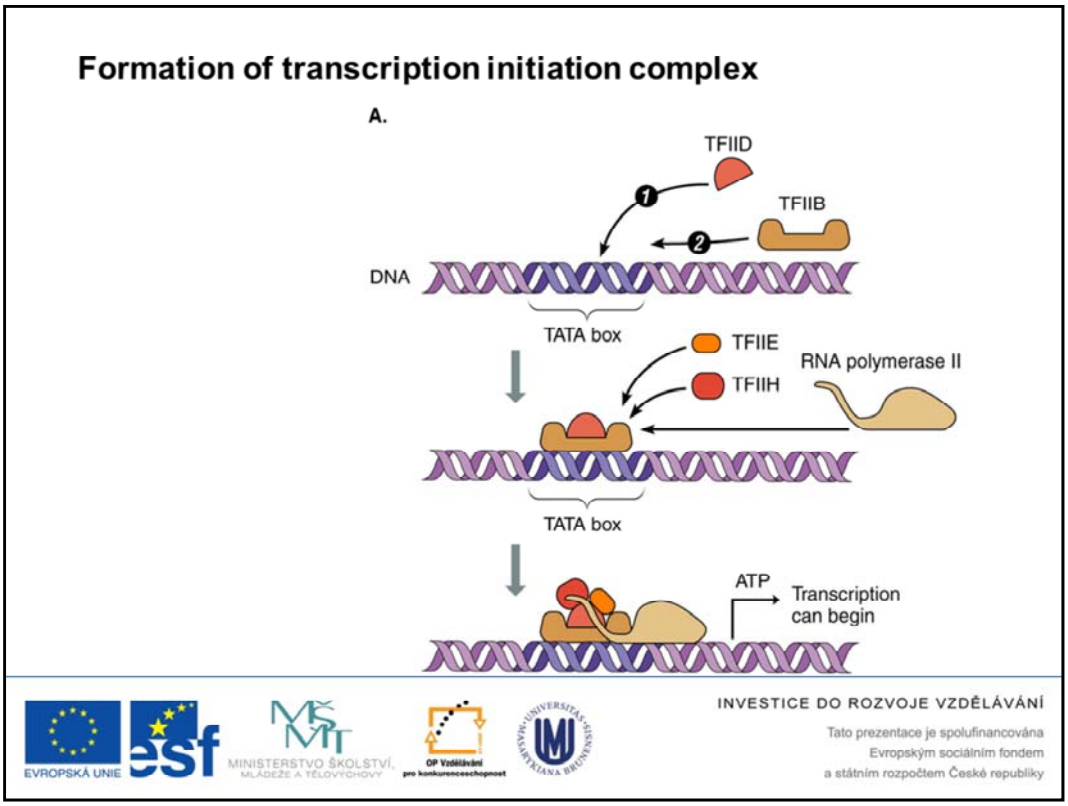
Up to know it is not clear how methylation regulates transcription. Possibly, methylation status affects chromatin configuration or binding general repressor factors.

Outline of Lesson 10

Regulation of Gene Expression during Development

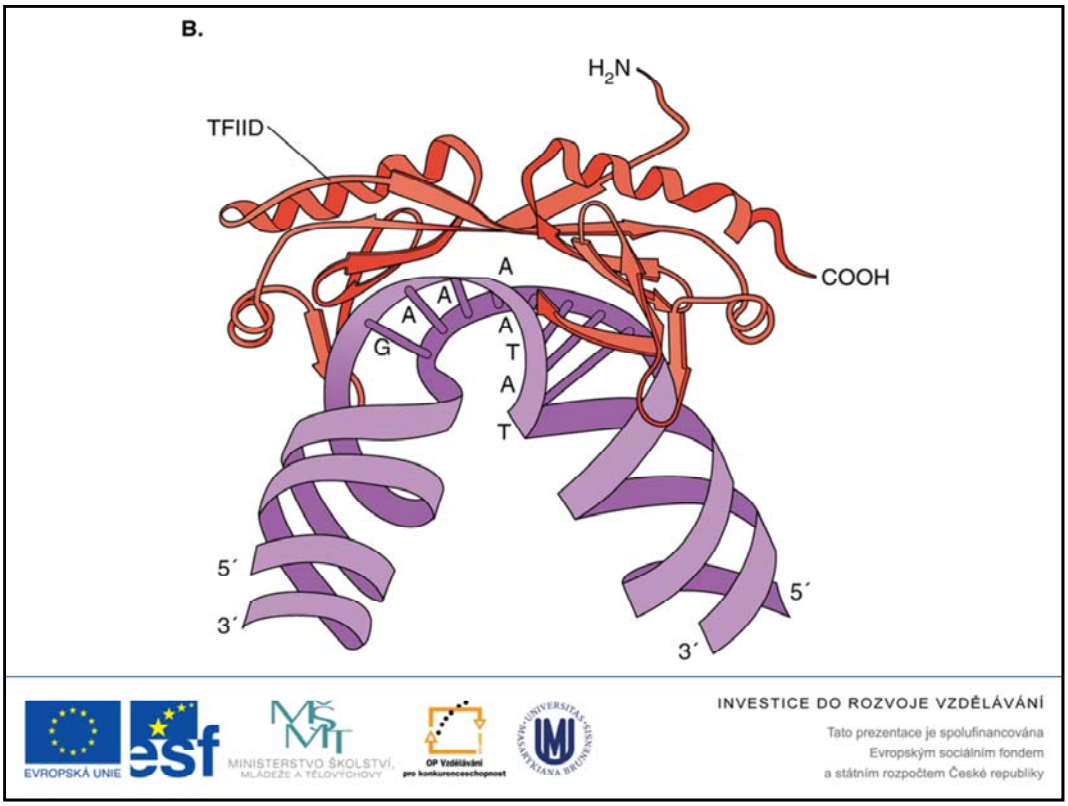
- Overview of levels of gene expression regulation
- Transcriptional gene regulation
  - Modification of the chromatin structure and DNA methylation
  - Transcriptional activation

EVROPSKÁ UNIE
esf
MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY
OP Vzdělávání pro konkurenceschopnost
UNIVERSITAS MASARYKIANA BRUNENSIS
INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ
Tato prezentace je spolufinancována Evropským sociálním fondem a státním rozpočtem České republiky



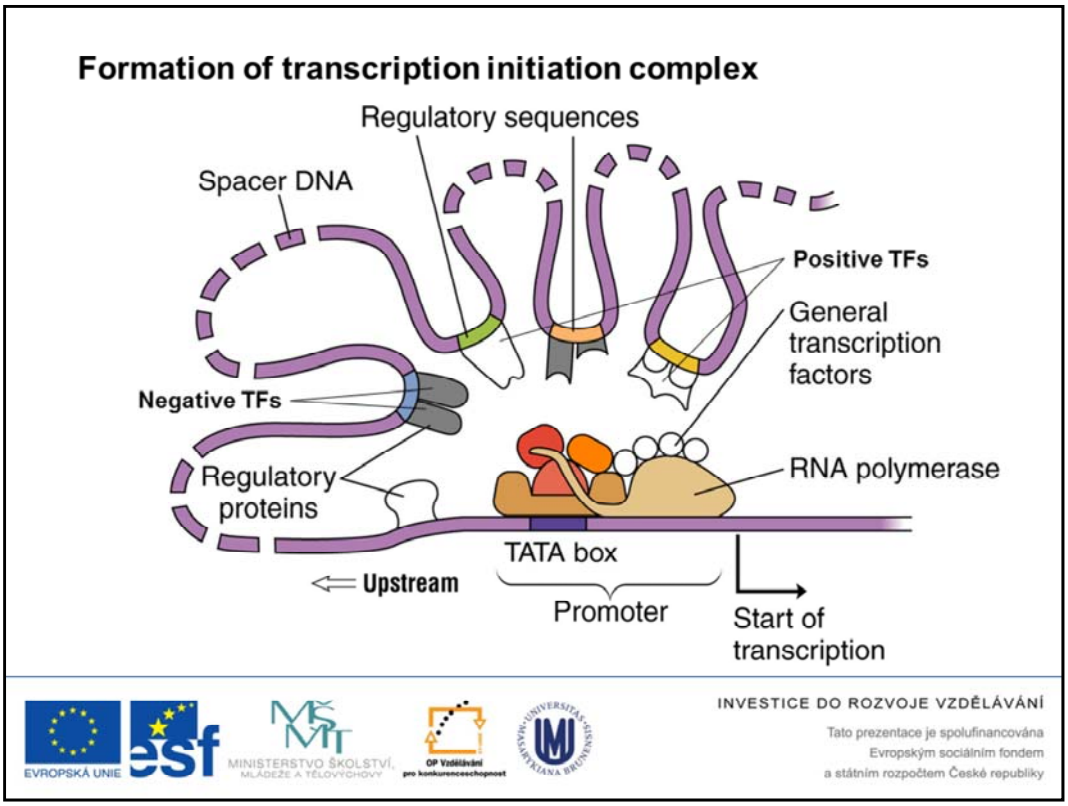
Regulation of transcription occurs via specific interaction of both general and tissue specific transcription factors (TFs) with promoter and/or enhancer sequences.

The scheme above shows simplified subsequent formation of the complex of TFs involved in the regulation of transcription. Interaction of general TFIID with the TATA box induces distortion of the DNA structure (see the next slide).



Induction of structural changes upon interaction of TFIID with DNA. This may be important for the assembly of other TFs involved in the formation of transcription initiation complex.

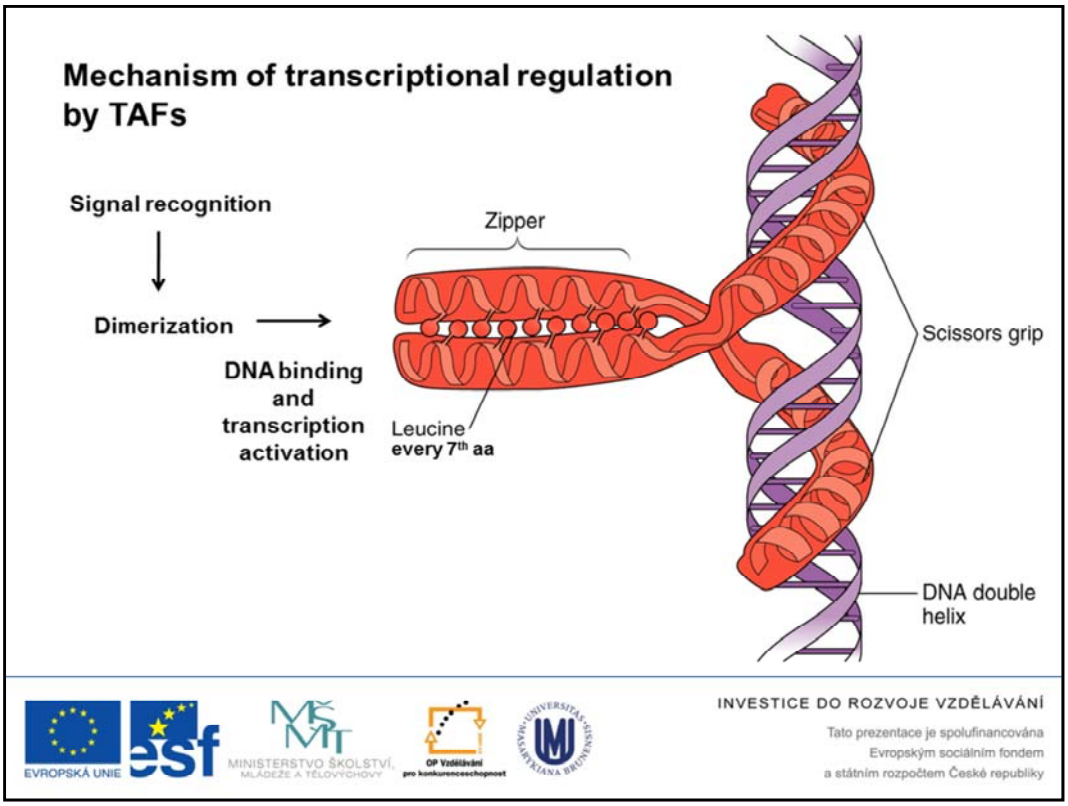
This change of confirmation provides a kind of “signature” that is recognized by other proteins and NA polymerase to recognize the proper binding site. However, there are also TATA box-less promoter, where probably other types of “signatures” occur.



The scheme showing the formation of the transcription initiation complex and the interaction of both positive (open symbols) and negative (solid symbols) factors.

These proteins bind to the regulatory sequences that might be hundreds or even thousands of base pairs away from the promoter. These protein interact with each other and with the RNA polymerase, integrating thus many signals into a “yes” or “no” response of the basal promoter, i.e. the region adjacent to the TATA box and recognized by the RNA polymerase.

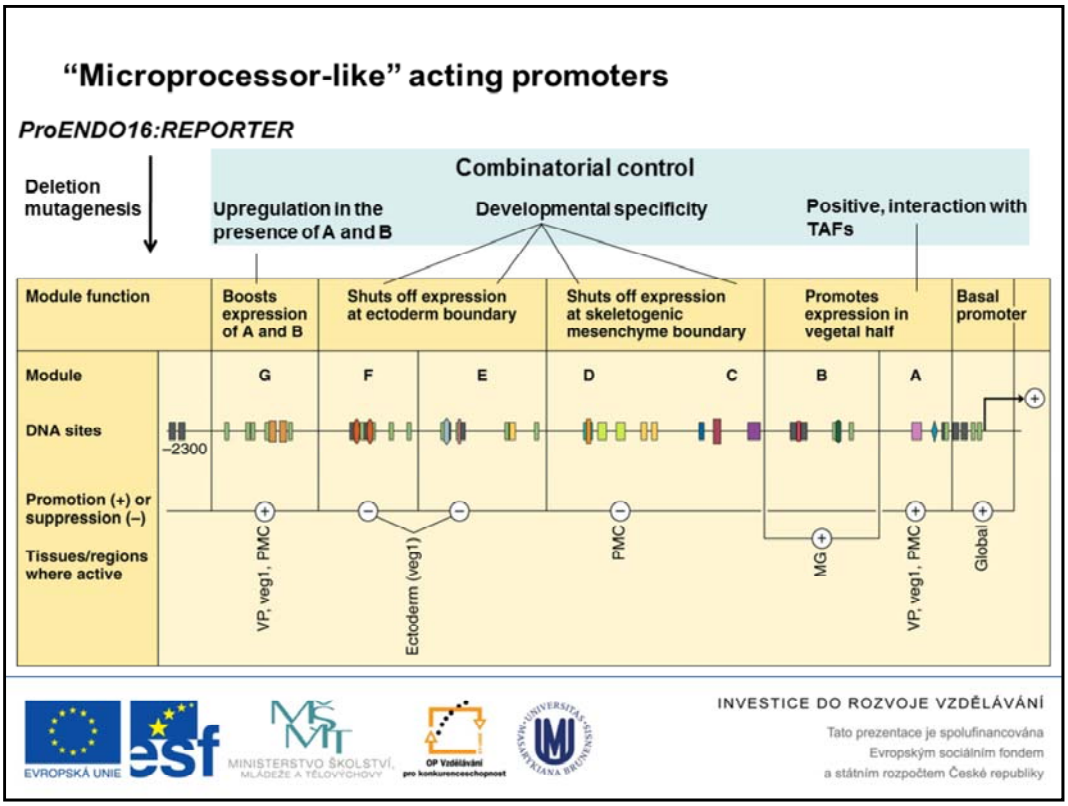
The individual positive or negative factors are complex and their activity might be regulated by their phosphorylation status or via their interaction with other proteins (i.e. monomeric or dimeric) etc..



There is a whole family of transcription activating factors (TAFs) that interact with signalling molecules, e.g. steroid hormones, thyroid hormones or retinoic acid and in a response to the signal translocate to the nucleus and activate transcription.

One of the type of TAF are leucine zipper or bZIP type TAFs. These TAFs are dimeric, with leucine-rich hydrophobic face formed by the Leu that occurs every 7th aa.

That allows the factor to take the proper configuration, which provides the dimer with the ability to bind DNA via charged aa.



An example of the "microprocessor"-like acting promoter is a promoter of the *endo16* gene from the sea urchin.

Endo16 is a single copy gene that codes for the RGD-containing calcium-binding protein of the cells of invaginating archenteron during gastrulation.

As visualized in experiments, the Endo16 protein may be an adhesion molecule, involved in the gastrulation of the embryo (Nocente-McGrath C et al, 1989).

There have been identified several gene regulatory modules in the *endo16* gene that have positive or negative regulatory role. These modules were identified via formation of deletion mutants of the transcriptional fusions with reporter gene.

The analysis has revealed that the module A has a positive function and must interact with its cognate TAFs for transcription to occur.

Module G enhances the expression when the A and B are active.

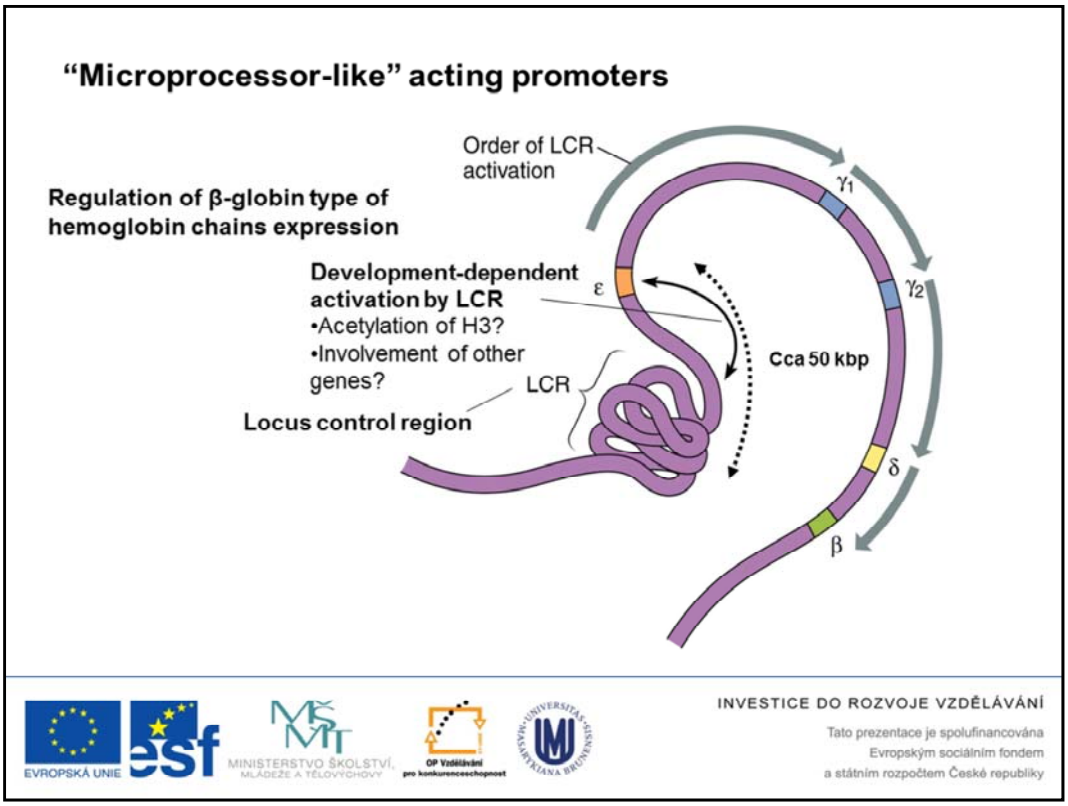
C, D, E and F are responsible for the specificity of the expression of endo16 during sea urchin development.

Each of the modules has several protein interaction sites, some of them general, other unique. Site for the protein SpGCF1 is present in many modules and is probably responsible for looping of chromatin, allowing thus bringing of distal regulatory

modules close to the basal promoter.

This type of regulation, i.e. based on the different activities of diverse regulatory sequences is sometimes called *combinatorial* and is common for development of many living creatures.

In the combinatorial type of regulations, some modules may act synergistically, some of them antagonistically, some may have both positive and negative roles (e.g. the module B, see the figure). This variability allows very precise and responsive regulations towards changing environmental conditions.



An example of the combinatorial gene regulation is the regulation of β-globin type of hemoglobin chains of humans.

As discussed in the Lesson 5, the type of hemoglobin produced by the fetus changes during development. The hemoglobin present in the liver-produced hemoglobin is composed of two α- and two β-type chains. The β-type hemoglobin chains are of several developmental types, produced by ε, γ1, γ2 and β (in this order). In addition, there is minor adult type of β-type hemoglobin, called δ globin.

The genes for the β-type chains are aligned on the chromosome in the order, in which they are expressed during development (se the figure).

For the expression of individual cell types is distinctive an upstream regulatory sequence called locus control region (LCR). LCR is located about 50 kbp away from the most proximal ε gene.

The LCR structure is different in erytrocyte precursor cells in comparison to other cells that could be demonstrated by the changes in the sensitivity to low concentrations of DNase, suggesting low amount of nucleosomes bound.

For the expression of the particular genes, the interaction of their regulatory sequences with LCRs is necessary. Because of LCR can interact only with one regulatory sequence at a time, only one type of genes for the particular β-type chain is activated (the first interaction of LCR with ε gene, which is later in development

replaced by the other one, is shown by the double-headed arrow).

The underlying molecular mechanisms of the specific pattern of the LCR movement from the most proximal towards the most distal gene cannot be satisfactory explained.

Probably, acetylation of H3 histones might play a role and possibly, other genes outside of the β-type chain family are involved in the regulation of LCR activity. That seems to be confirmed by the identification of other human genes with similar structure, suggesting common regulatory mechanisms via LCRs.

Outline of Lesson 10

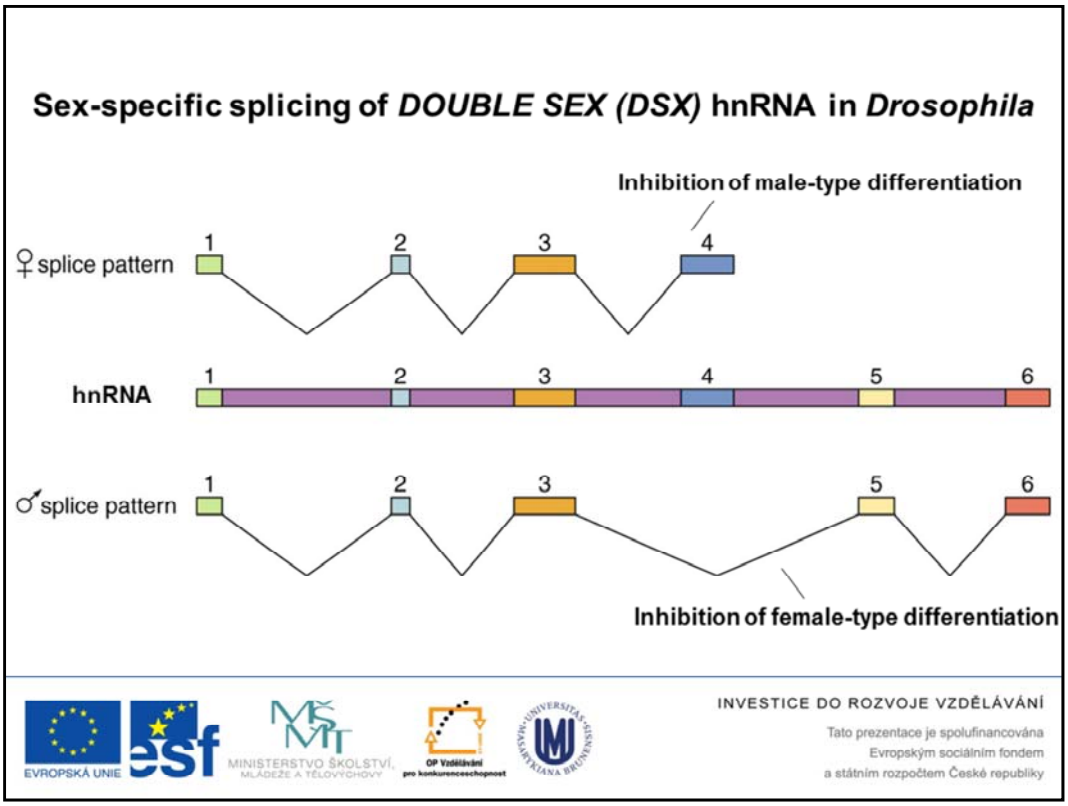
Regulation of Gene Expression during Development

- Overview of levels of gene expression regulation
- Transcriptional gene regulation
  - Modification of the chromatin structure and DNA methylation
  - Transcriptional activation
- Post-transcriptional gene regulation
  - Splicing of hnRNA

EVROPSKÁ UNIE
esf
MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY
OP Vzdělávání pro konkurenceschopnost
UNIVERSITAS MASARYKIANA BRUNENSIS
INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ
Tato prezentace je spolufinancována Evropským sociálním fondem a státním rozpočtem České republiky

Posttranscriptional modifications takes place after transcription and these are a matter of regulation that has developmental consequences, too.

The most important posttranscriptional modification in eukaryotes in splicing of introns. Immediately aftre transcription, the hnRNA is polyadenylated and interacts with proteins and small nuclear RNAs (snRNAs) that are part of small nuclear nucleoproteins. These interact with specific recognition sites and allow splicing of introns and connection of exons into final mRNA.



An example of the developmental importance is an alternative, sex-specific splicing of *DOUBLE SEX (DSX)* gene in *Drosophila.* DSX is a last member of the cascade involved in the sex specification of *Drosophila.*

In females, exons 1 through 4 are connected, while in males, exons 1-3, 5 and 6 are connected into final mRNA that is translated.

Male form of DSX blocks the female-type of differentiation and *vice versa*, the female form of DSX blocks the male-type of differentiation.

Outline of Lesson 10

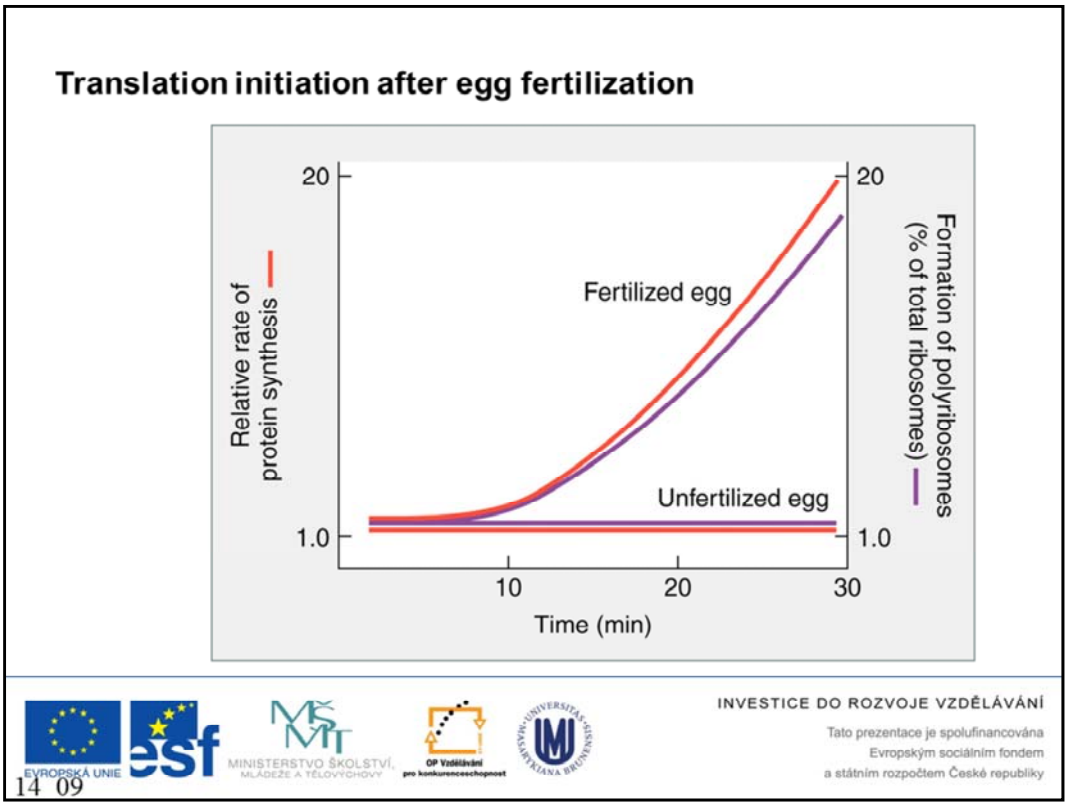
Regulation of Gene Expression during Development

- Overview of levels of gene expression regulation
- Transcriptional gene regulation
  - Modification of the chromatin structure and DNA methylation
  - Transcriptional activation
- Post-transcriptional gene regulation
  - Splicing of hnRNA
  - Translation initiation

EVROPSKÁ UNIE
MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY
OP Vzdělávání pro konkurenceschopnost

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

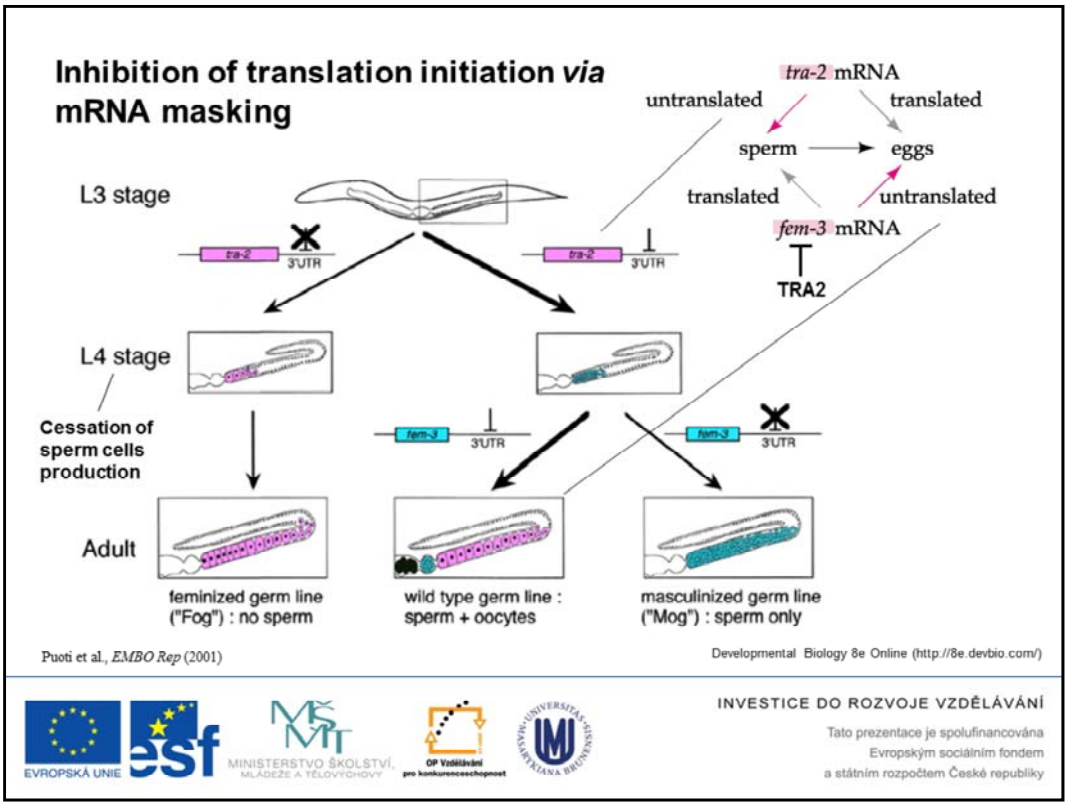
Tato prezentace je spolufinancována Evropským sociálním fondem a státním rozpočtem České republiky



Initiation of translation, elongation of the peptide chain and protein release are all the steps being regulated during development.

An example is translation initiation after fertilization e.g. in the sea urchin eggs. On the figure above, there is shown a dramatic increase of polyribosomes formation (polyribosome is a string of ribosomes on mRNA) and increase in the protein synthesis after urchin egg fertilization.

The polyribosomes form on the preexisting mRNA, originating from the egg.



3’ untranslated region (3’ UTR) of mRNA could interact with specific proteins. This interaction then leads to the inhibition of translation initiation and the process is called *masking* of mRNAs.

This mechanism occurs in many organisms, including *Xenopus*, sea urchin, mouse and *Coenorhabditis*.

The masking regulates also polyA tail length. Once unmasked, the mRNA could be intensely polyadenylated, leading to increase of the polyA tail.

The length of the polyA tail also affects translation initiation and stability of mRNA.

On the figure above, there is shown a mechanism of the sex determination in *Coenorhabditis*, where posttranslational regulation of expression plays important role.

Most *C. elegans* have a female body but are hermaphroditic, producing both sperm and eggs at different times.

The first germ cells to differentiate in the nematode become sperm, which are stored in the uterus for later use. After the fourth molt (from larva to adult), the germ cells cease making sperm and begin to make eggs. These eggs will eventually become fertilized by the stored sperm.

The process determining which path the germ cell follows—to sperm or to egg—depends on the translational repression of different messages.

The initiation of sperm formation is achieved by the repression of the *tra-2* message. The Tra-2 protein is essential for the development of eggs and female body cells, and repression of *tra-2* mRNA translation in germ cells causes them to become sperm.

The 3' UTR of this message contains two regions of 28 nucleotides, each of which appears to bind a putative repressor protein that is synthesized during the larval stages associated with spermatogenesis. If these regions are mutated, the translation of the *tra-2* mRNA is not repressed, no sperm is made, and the nematode is functionally female instead of hermaphroditic .

The switchover from spermatogenesis to oogenesis also requires suppressing the translation of the *fem-3* gene through its 3' UTR.

The Fem-3 protein is critical for specifying male body cells and sperm production. Transcription of the *fem-3* gene is inhibited by Tra-2 protein, but the repression of existing *fem-3* messages is also needed. This translational repression appears to be effected by the binding of a translational inhibitor by the 3' UTR of the *fem-3* mRNA.

Thus, the initiation of spermatogenesis in hermaphroditic nematodes and the transition from spermatogenesis to oogenesis appears to be regulated by translational repression through the 3' UTR.

Outline of Lesson 10

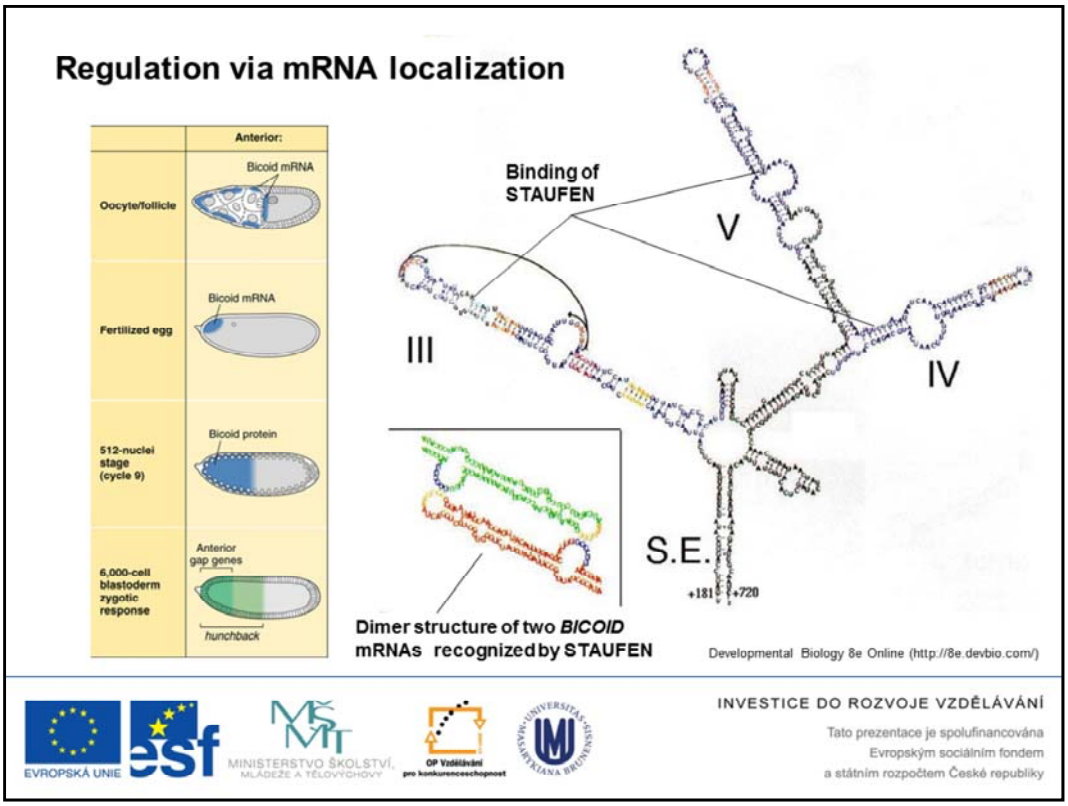
Regulation of Gene Expression during Development

- Overview of levels of gene expression regulation
- Transcriptional gene regulation
  - Modification of the chromatin structure and DNA methylation
  - Transcriptional activation
- Post-transcriptional gene regulation
  - Splicing of hnRNA
  - Translation initiation
  - Localization of mRNA

EVROPSKÁ UNIE
MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY
OP Vzdělávání pro konkurenceschopnost

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

Tato prezentace je spolufinancována Evropským sociálním fondem a státním rozpočtem České republiky



Messenger mRNA may be localized to the specific positions in the cell. Specific mRNA localization was discussed in the Lesson 2, when during syncitial stage of the *Drosophila* embryo formation, BICOID mRNA is located to the dorsoanterior portion of the developing embryo.

Again, as just discussed, 3’ UTR plays an important role in the tethering of BICOID mRNA.

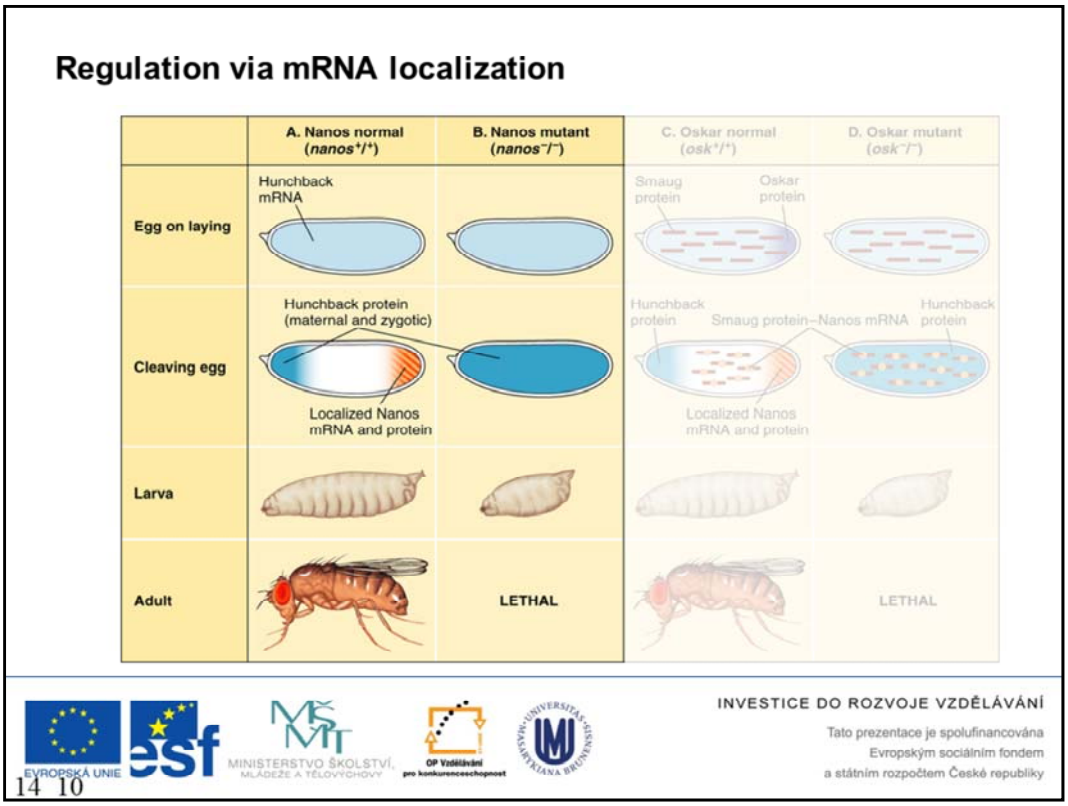
Transcription of the bicoid message occurs in the nurse cells (during stages 4 and 5 of oogenesis) and the mRNA is transported immediately into the oocyte.

BICOID mRNA is localized to the anterior end of the oocyte with the aid of STAUFEN protein. STAUFEN recognizes regions of double stranded mRNA, which are present in arm III of the *BICOID* 3' UTR (see left-hand figure). The STAUFEN-*BICOID* mRNA complex associates with the minus ends of the microtubules, thereby localizing *BICOID* to the anterior.

On the right-hand figure, there is a secondary structure of the *BICOID* mRNA 3'UTR.

STAUFEN protein binds to specific sequences of arms III, IVb and Vb shown here in blue lettering.

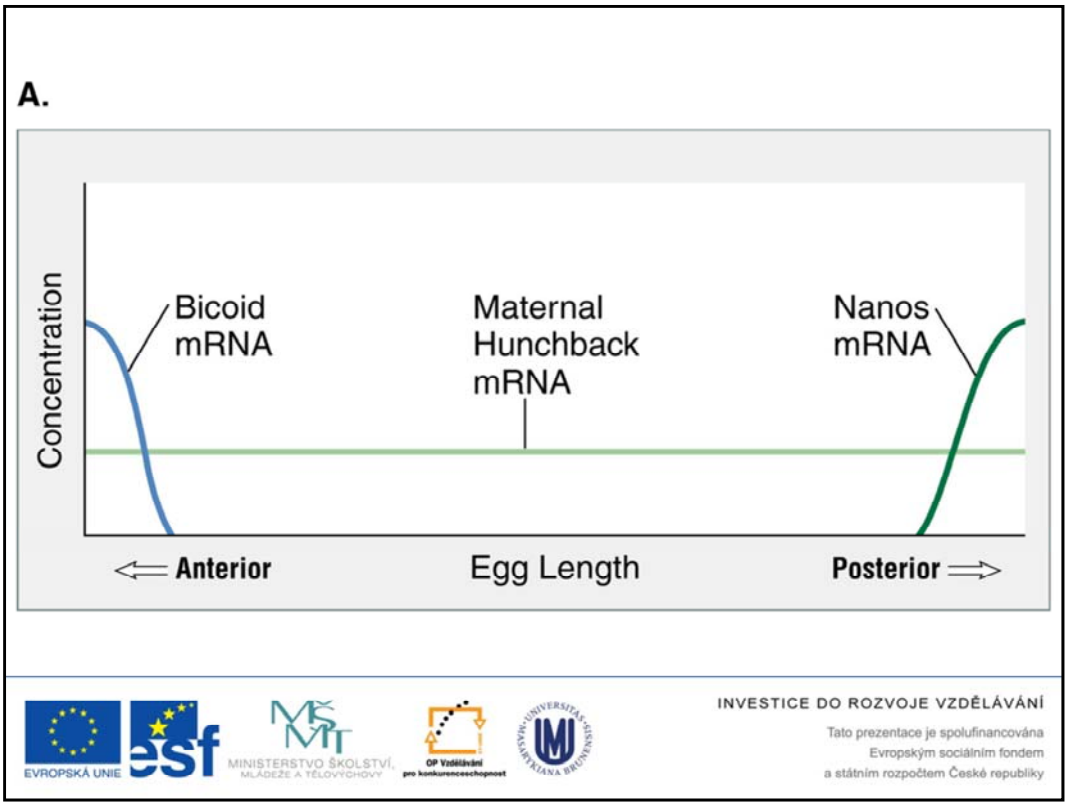
Helix III is essential for the binding to STAUFEN protein and accumulation in the anterior of the oocyte. The insert shows the interactions between helices III of two *BICOID* mRNAs. This structure is recognized by the STAUFEN protein.



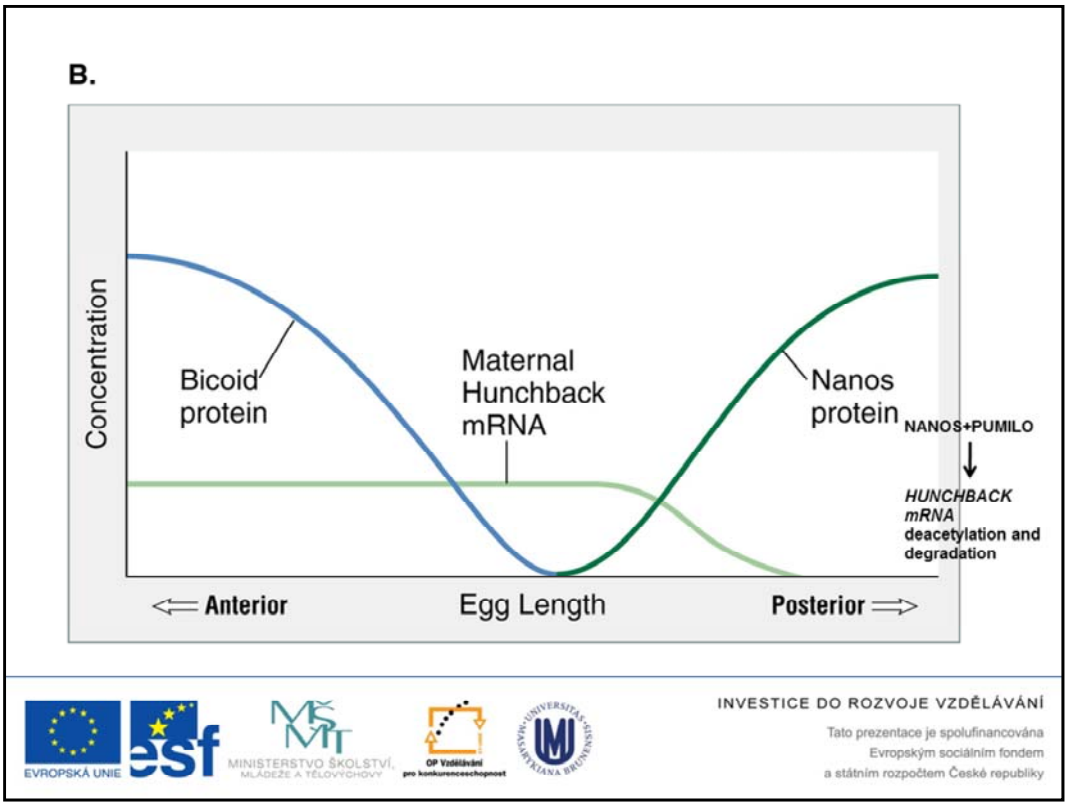
Another example represents localization of *NANOS* mRNA (see the figure above).

NANOS mRNA is localized posteriorly and acts as a translational suppressor of the HUNCHBACK protein (recall the Lesson 2 and next slides).

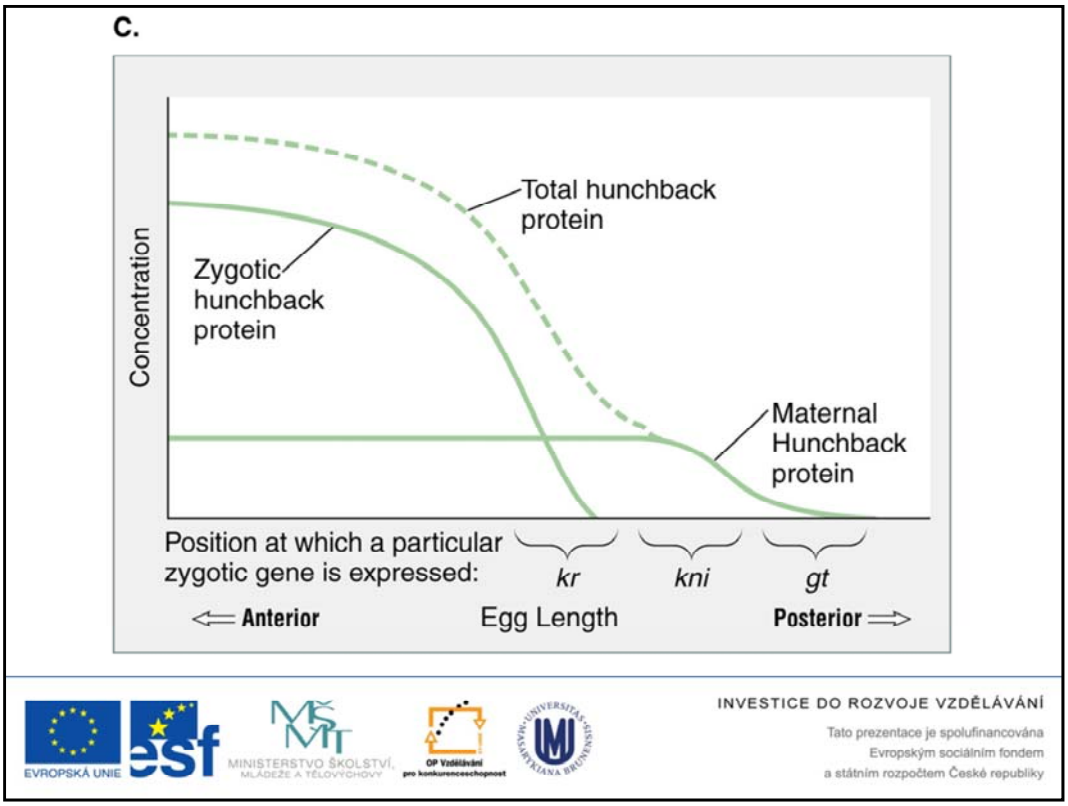
HUNCHBACK is a TF that further acts as a regulator of other donwnstream genes, e.g. *KRUPPEL (KR)*, *KNIRPS (KNI)* and *GIANT (GT)* (see next slides).



The levels of BCD and NANOS at the time of fertilization.



After fertlization, the proteins are translated. NANOS inhibits *HUNCHBACK* translation posteriorly.

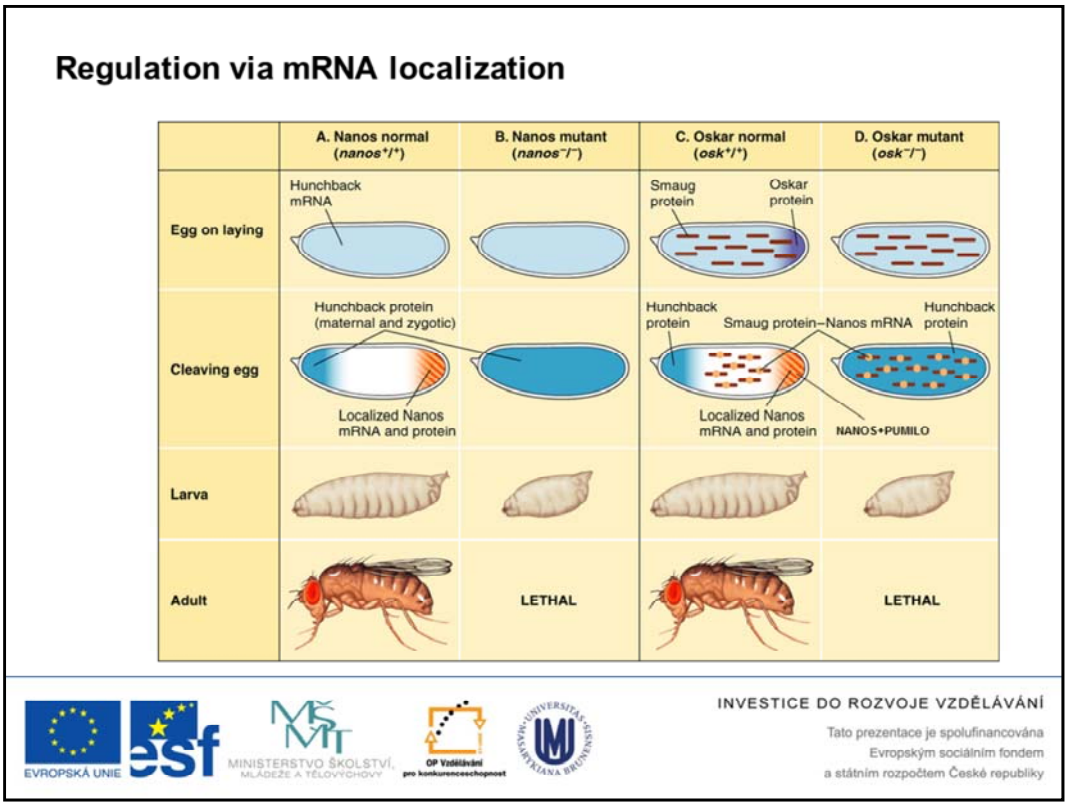


Later in development, *HUNCHBACK* is induced anteriorly by *BCD*.

The gradient of HUNCHBACK regulates the position of expression of *KRUPPEL (KR), KNIRPS (KNI)* and *GIANT (GT).*

HUNCHBACK upregulates *KR* and *KNI*, where *KR* requires higher than maternal level of *HUNCHBACK* .

GT is repressed by *HUNCHBACK*.



NANOS mRNA requires for its function in a translational repression of HUNCHBACK another protein, PUMILO.

PUMILO binds to the 3'UTR of *HUNCHBACK* mRNA and allows thus binding of NANOS and subsequent deadenylation of the *HUNCHBACK* and its degradation.

Localization of NANOS mRNA is under control of several genes (e.g. *OSKAR, STAUFEN, VALOIS, VASA* and *TUDOR*). The proteins of these genes bind top the 3'UTR of *NANOS* mRNA and allow its posterior localization.

Interestingly, the localization of *NANOS* mRNA is predominantly at the posterior pole, but not exclusively. There is still some *NANOS* located in the middle or even anterior portion of the embryo. However, only posteriorly located *NANOS* is translated.

That is ensured by the action of protein SMAUG. SMAUG binds to the specific sequence in the 3'UTR of *NANOS* mRNA and inhibits its translation. The localization of mRNA to the posterior domain relieves this translational repression (see the figure).

Outline of Lesson 10

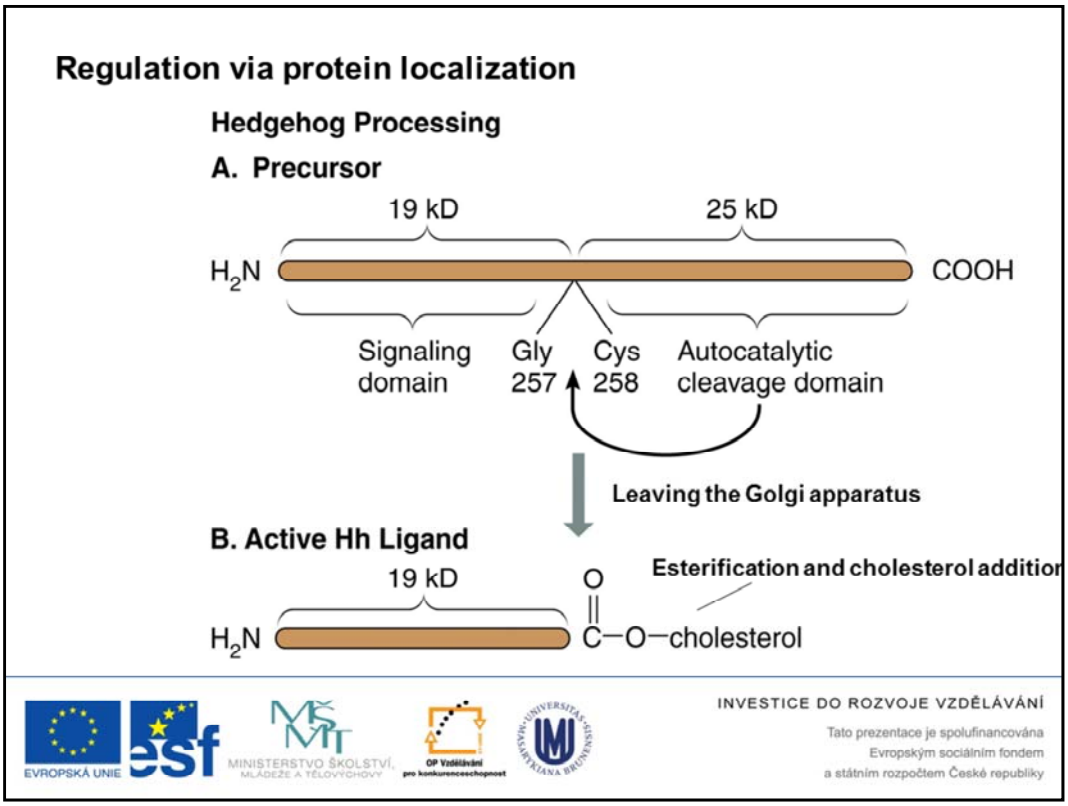
Regulation of Gene Expression during Development

- Overview of levels of gene expression regulation
- Transcriptional gene regulation
  - Modification of the chromatin structure and DNA methylation
  - Transcriptional activation
- Post-transcriptional gene regulation
  - Splicing of hnRNA
  - Translation initiation
  - Localization of mRNA
  - Protein localization

EVROPSKÁ UNIE
MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY
OP Vzdělávání pro konkurenceschopnost

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

Tato prezentace je spolufinancována Evropským sociálním fondem a státním rozpočtem České republiky



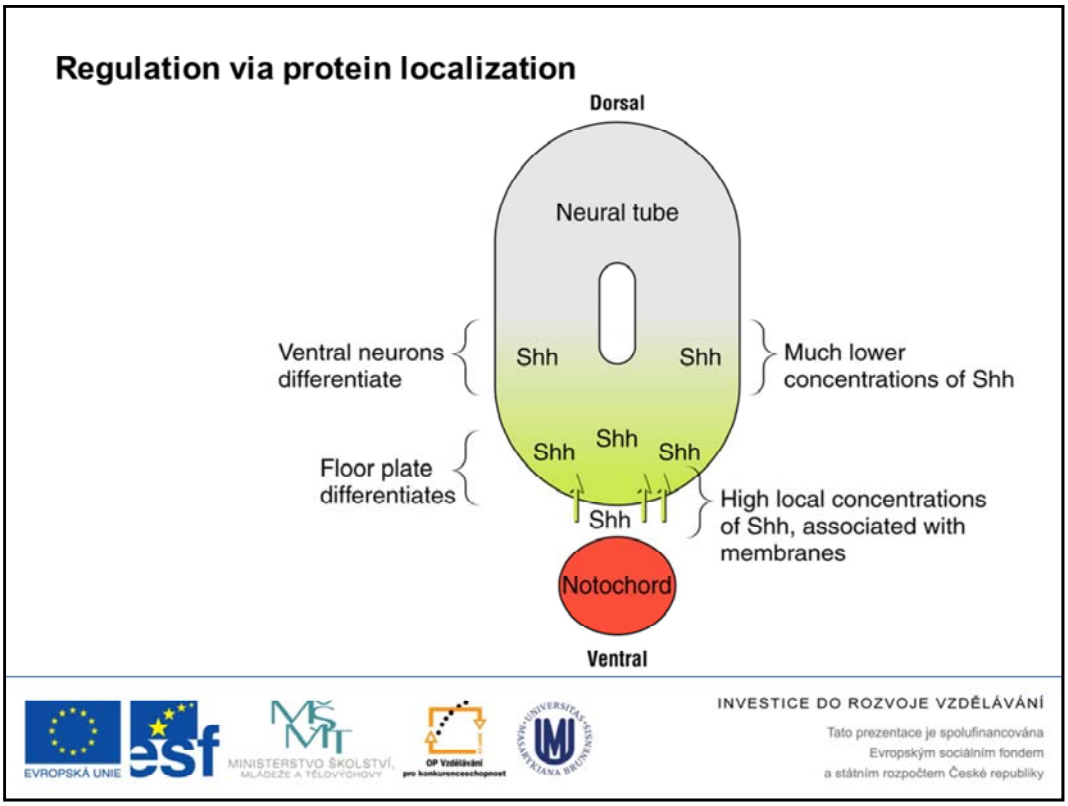
Autocatalytic cleavage and esterification represent another example of posttrransaltional modifications during development.

HEDGEHOG (HD) protein from *Drosophila* is important signalling ligand involved in the formation of boundaries from body segments and other patterned elements during development.

HD contains the signal peptide that directs the HD protein to the ER and later to the Golgi apparatus. After leaving the Golgi apparatus, the protein is cleaved into two parts: The N-terminal (19 kD) and C-terminal portion (25 kD).

The N-terminal portion is later modified via esterification and cholesterol is added to its C-terminus. Only the N-terminal part acts as an active protein that bids to its receptor.

Modification of HD via cholesterol binding has an important role in the regulation of the HD diffusion from the *HD* expressing cells. Genes like *DISPATCHED* or *TOUT VELU* do affect mobility of HD that is important for HD-mediated signalling.



Similar function of cholesterol as in case of HD protein seems to be applied also in the case of its vertebrate homologue, SONIC HEDGEHOG (SHH), discussed previously (see Lesson 4) during the formation of dorsolateral axis in the neural tube development.

Notochord acts as a source of the SHH production that leads to the differentiation of floor plate in the ventral portion of the neural tube while motor neurons differentiate in the dorsolateral poprtion, where SHH concentration is much lower.

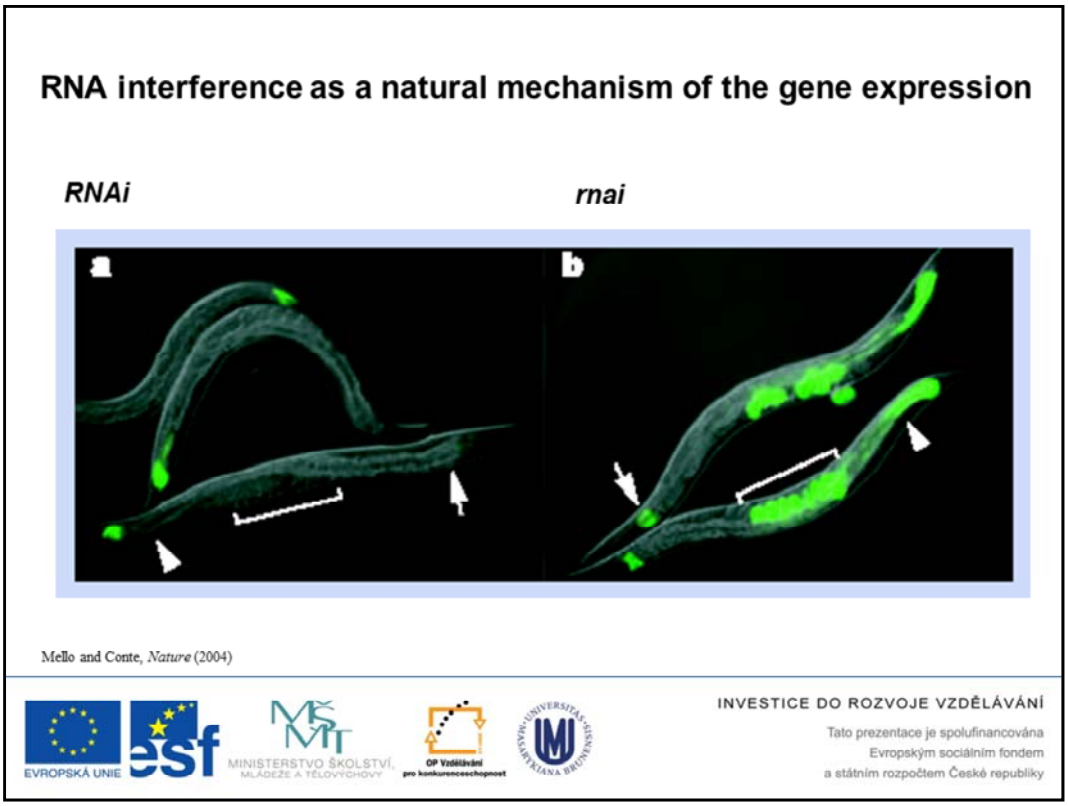
Probably, the cholesterol modification is involved in the regulation of SHH distribution.

Outline of Lesson 10

Regulation of Gene Expression during Development

- Overview of levels of gene expression regulation
- Transcriptional gene regulation
  - Modification of the chromatin structure and DNA methylation
  - Transcriptional activation
- Post-transcriptional gene regulation
  - Splicing of hnRNA
  - Translation initiation
  - Localization of mRNA
  - Protein localization
- RNA interference
  - Identification and mechanism of gene expression regulation via RNA interference

EVROPSKÁ UNIE
esf
MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY
OP Vzdělávání pro konkurenceschopnost
INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ
Tato prezentace je spolufinancována Evropským sociálním fondem a státním rozpočtem České republiky



RNA interference or RNAi, is a mechanism of targeted posttranscriptional regulation of gene expression, what is also called *posttranscriptional gene silencing (PTGS*).

The molecular mechanism o PTGS was discovered in experiments with gene silencing in *Coenorhabditis elegans*. It was found that surprisingly both sense and anti-sense mRNA when injected into Coenorhabditis was able to silence the gene of interest.

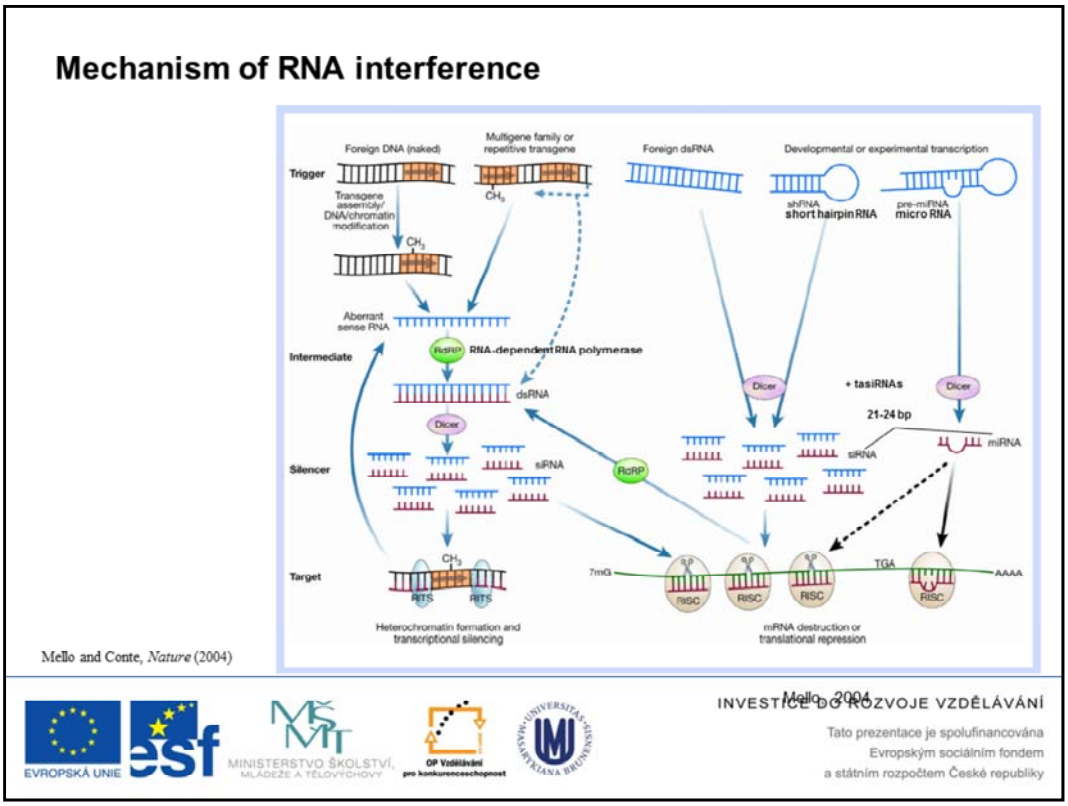
One of possible hypothesis predicted that contamination during in vitro transcription might lead to the formation of dsRNa that would be responsible for the observed phenomenon.

The hypothesis was confirmed by the introducing dsRNA in *Coenorhbditis* and it was found that the dsRNa induced gene silencing by about an order or two better then anti sense mRNA.

Importantly, in then forward genetic screens were identified mutants that were affected in the dsRNA-induced gen silencing. This suggested that the mechanism of dsRNA-mediated gene silencing is an **intrinsic regulatory mechanism for regulation of gene expression**.

That finding triggered an explosion of subsequent discoveries and led to the

identification of many members of the entire pathway.



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 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).

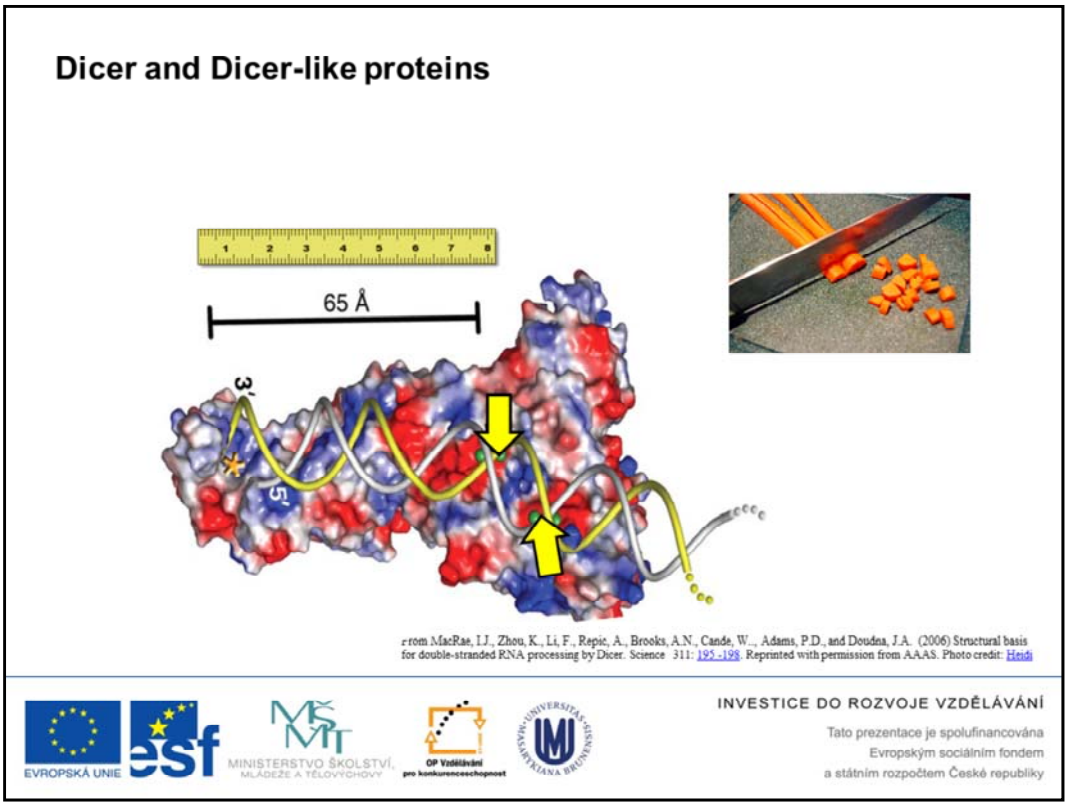
miRNAs are involved in the developmental-specific regulations and are product of transcription of endogenous genes encoding for small dsRNAs with specific structure (see the figure).

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.

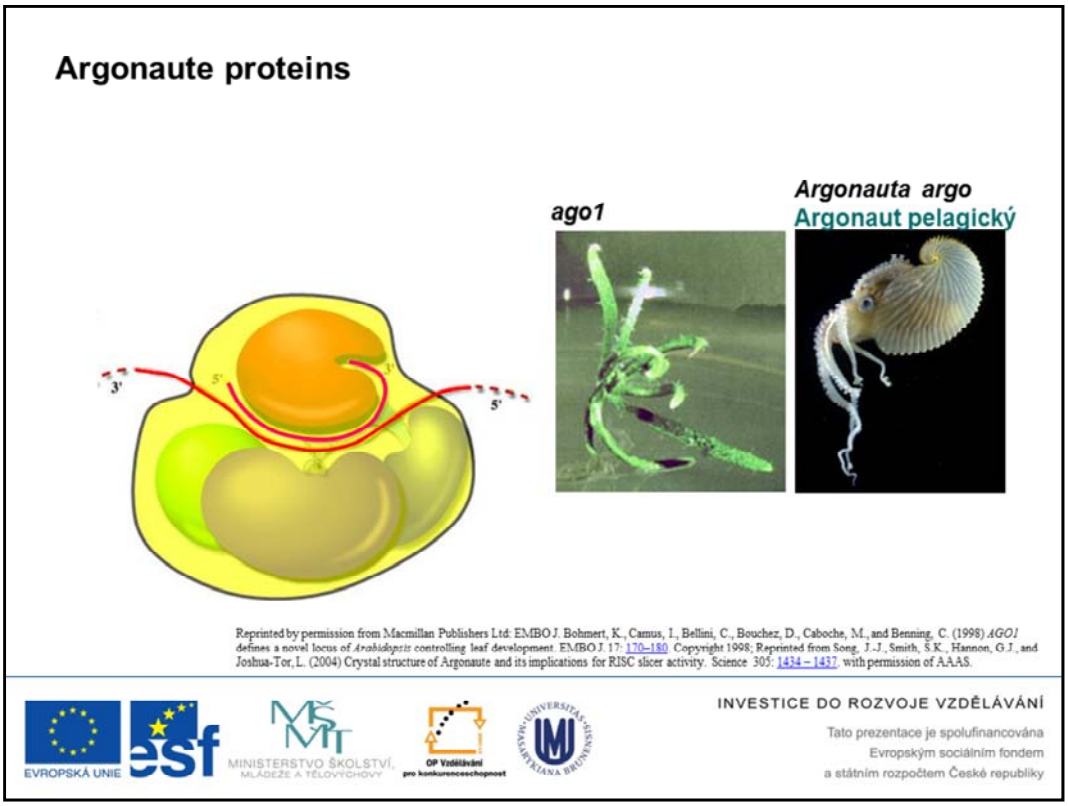
These small RNAs are of 21-24 bp long and bind either to RNA-induced transcriptional silencing complex (RITS) or RNA-induced silencing komplex (RISC).



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 bind small RNAs and their targets.

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.*

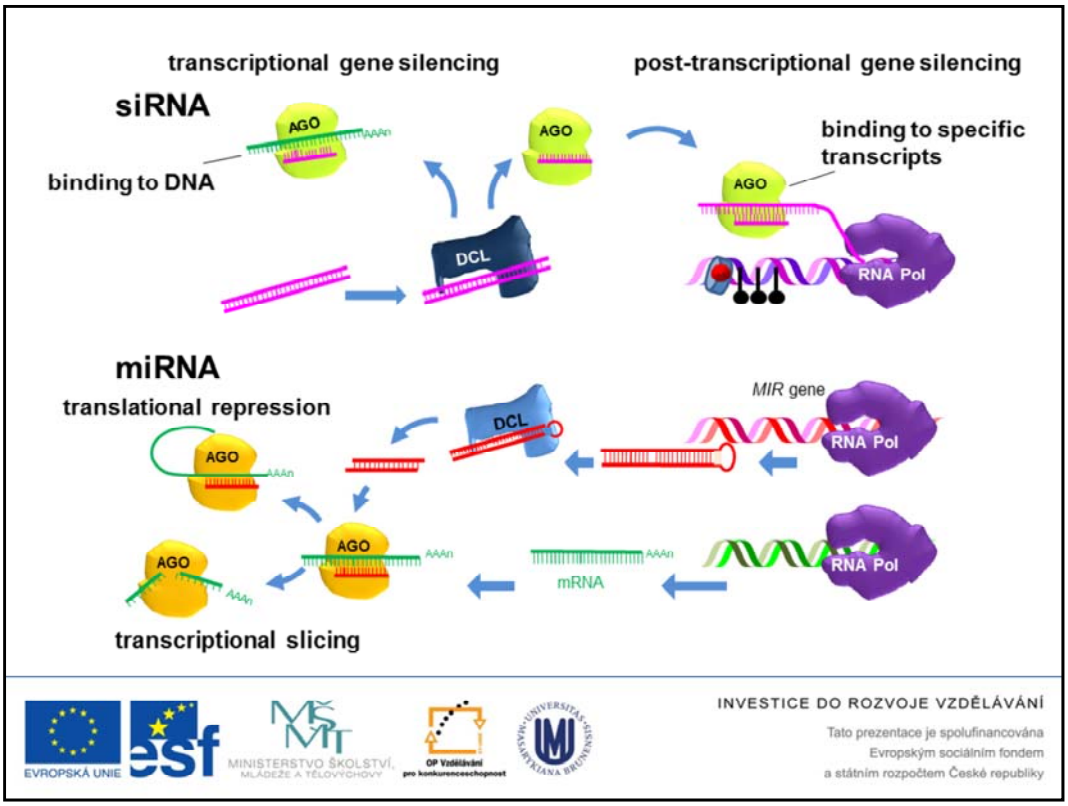
ARGONAUTE proteins are classified into three paralogous groups: Argonaute-like proteins, which are similar to *Arabidopsis thaliana AGO1;* Piwi-like proteins, which are closely related to *D. melanogaster PIWI* (P-element induced wimpy testis); and the recently identified *Caenorhabditis elegans*-specific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

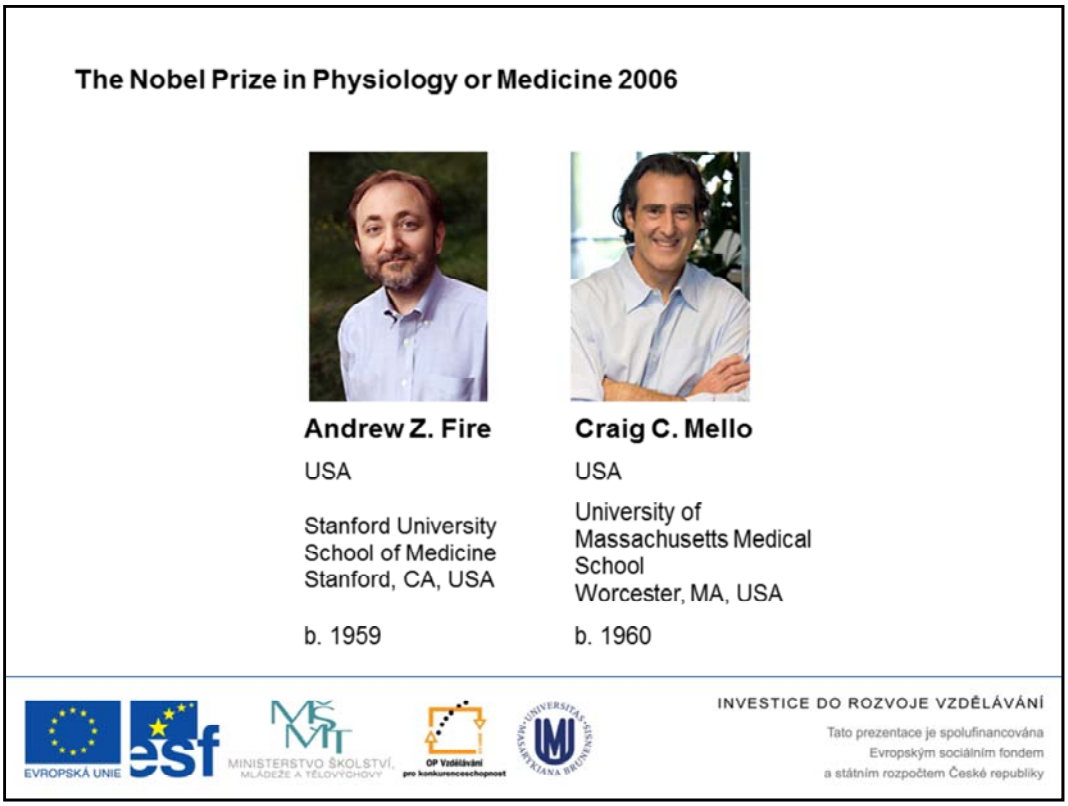
8 Argonaute genes in humans (4 Argonaute-like and 4 Piwi-like), 5 in the *D. melanogaster genome (2 Argonaute-like and* 3 Piwi-like), 10 Argonaute-like in *A. thaliana,* only 1 Argonaute-like in *Schizosaccharomyces pombe and at* least 26 Argonaute genes in *C. elegans (5 Argonaute-like,* 3 Piwi-like 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.



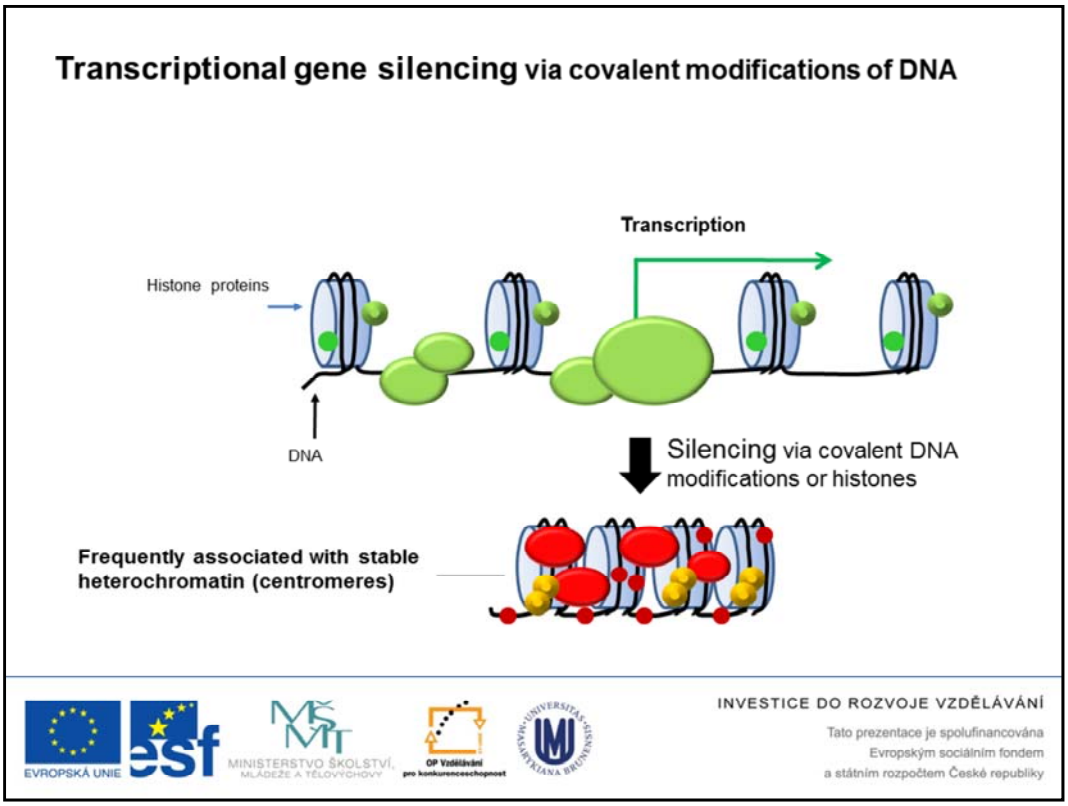
In 2006, Andrwe Z. Fire and Craig C. Mello were honored by the Nobel prize “for their discovery of RNA interference - gene silencing by double-stranded RNA“.

Outline of Lesson 10

Regulation of Gene Expression during Development

- Overview of levels of gene expression regulation
- Transcriptional gene regulation
  - Modification of the chromatin structure and DNA methylation
  - Transcriptional activation
- Post-transcriptional gene regulation
  - Splicing of hnRNA
  - Translation initiation
  - Localization of mRNA
  - Protein localization
- RNA interference
  - Identification and mechanism of gene expression regulation via RNA interference
  - siRNA-mediated silencing

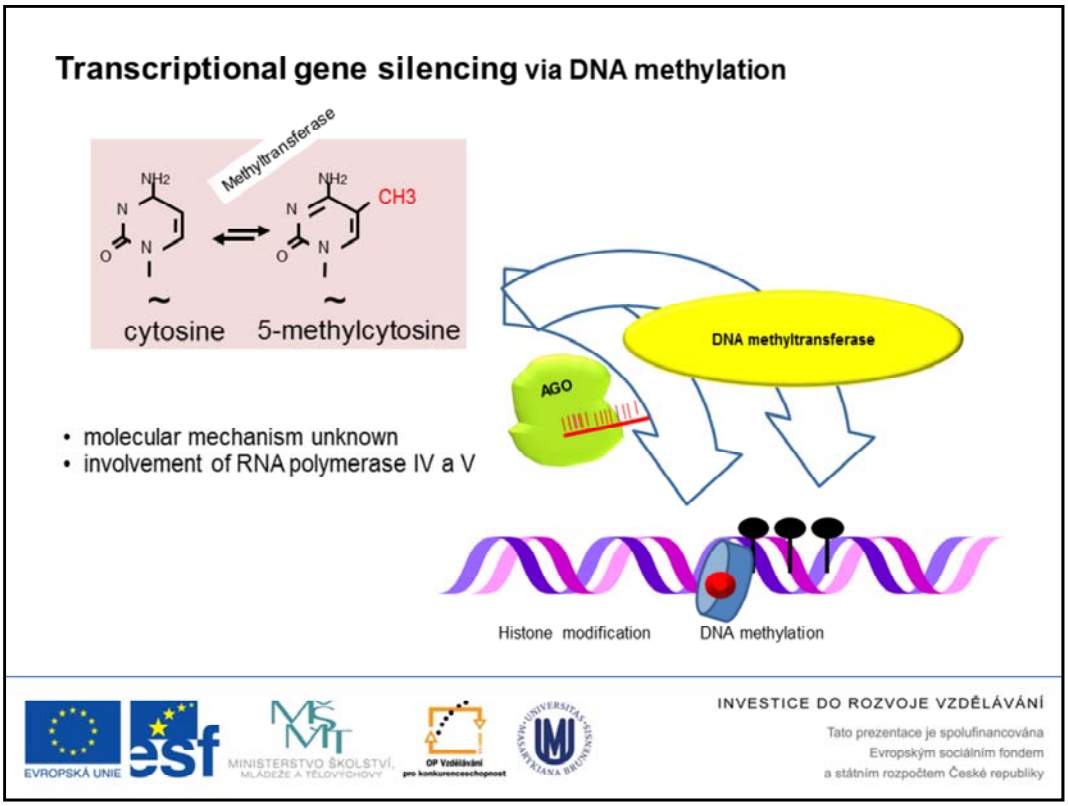
EVROPSKÁ UNIE
esf
MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY
OP Vzdělávání pro konkurenceschopnost
INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ
Tato prezentace je spolufinancována Evropským sociálním fondem a státním rozpočtem České republiky



Small RNAs can initiate gene silencing through covalent modifications of the DNA or its associated histone proteins, interfering with transcription.

This form of silencing is frequently associated with stably silenced DNA including centromeres and transposons, but also occurs at genes.

This figure represents chromatin (DNA wrapped around histones) in an open and closed conformation. Chromatin modifications include covalent modifications to DNA and the histone proteins.



siRNAs can target DNA for silencing by cytosine methylation or histone modifying enzymes.

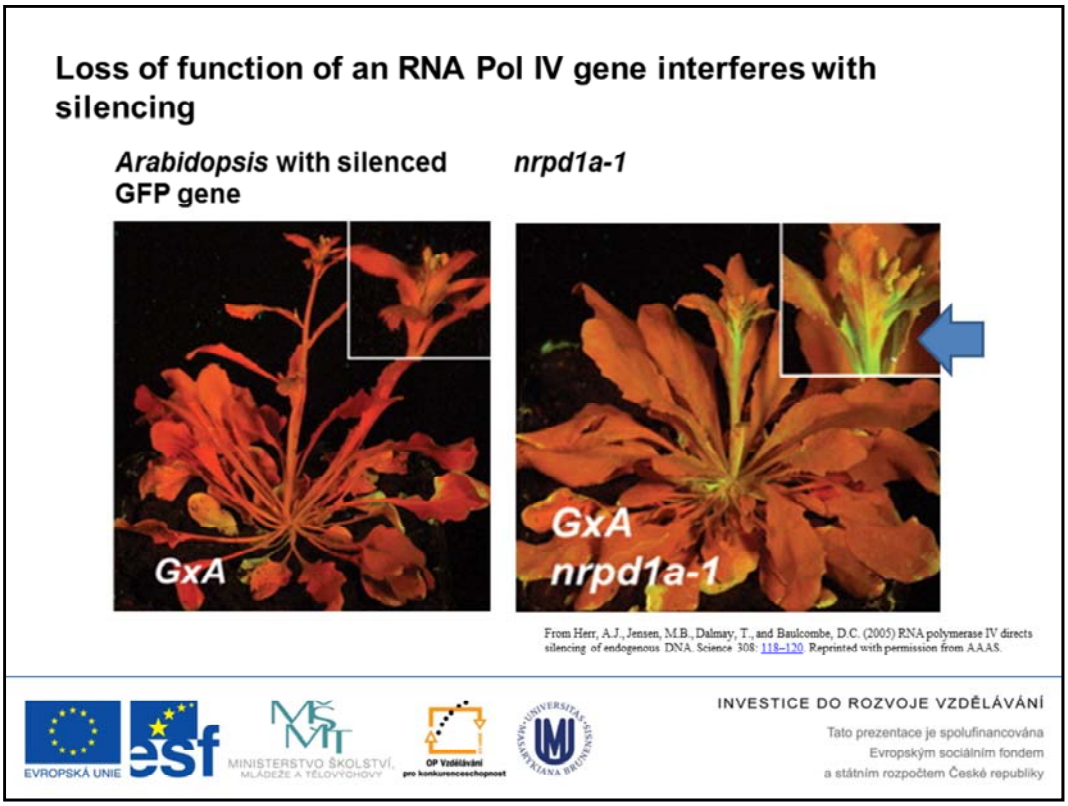
DNA can be covalently modified by cytosine methylation, carried out by DNA methyltransferases.

The precise mechanisms by which siRNAs target DNA for silencing are not known, but involve the action of two plant-specific RNA-polymerase complexes, RNA Polymerase IV (Pol IV) and RNA Polymerase V (Pol V).

Complex	Distribution	Function
RNA Polymerase I	All eukaryotes	Production of rRNA
RNA Polymerase II	All eukaryotes	Production of mRNA, microRNA
RNA Polymerase III	All eukaryotes	Production of tRNA, 5S rRNA
RNA Polymerase IV	Land plants	Production of siRNA
RNA Polymerase V	Angiosperms	Recruitment of AGO to DNA

Plants have additional RNA Polymerase complexes that contribute to silencing.

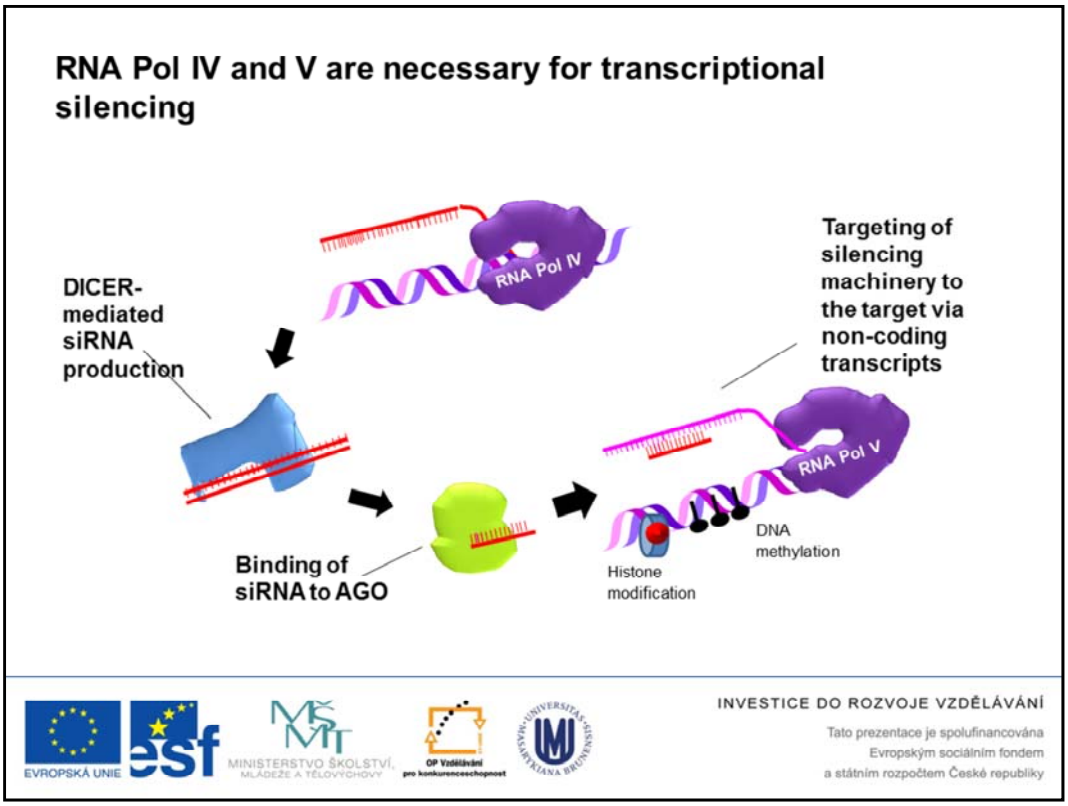
Besides RNA polymerases I-III that occur also in other eucaryotes, plant posses RNA polymerase IV that is involved in the siRNAs and RNA polymerase V that is necessary for the recruitment of AGO to DNA (see the table above).



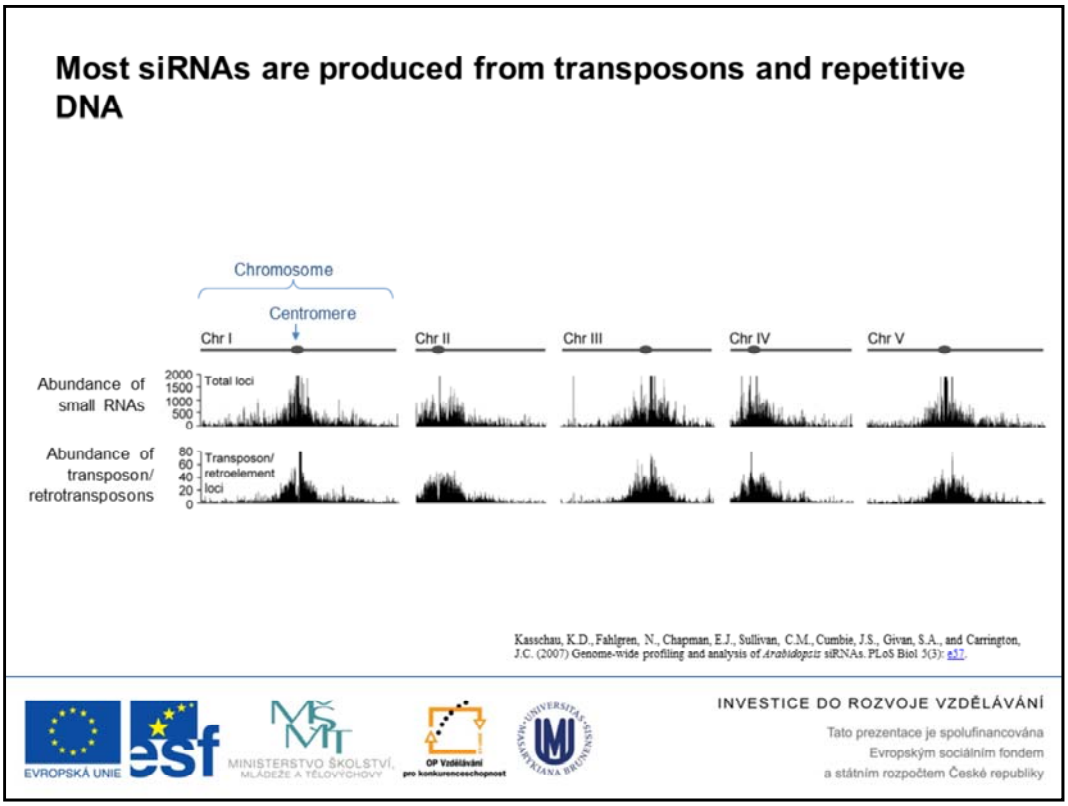
In the wild-type background, the red indicates endogenous chlorophyll fluorescence with no GFP expression; it is silenced.

NRPD1A encodes a subunit of RNA Polymerase IV. In the *nrpd1a* mutant, RNA Polymerase IV isn't produced, interfering thus with silencing.

Green signal (arrow) indicates GFP is expressed, showing that Pol IV is required for gene silencing.



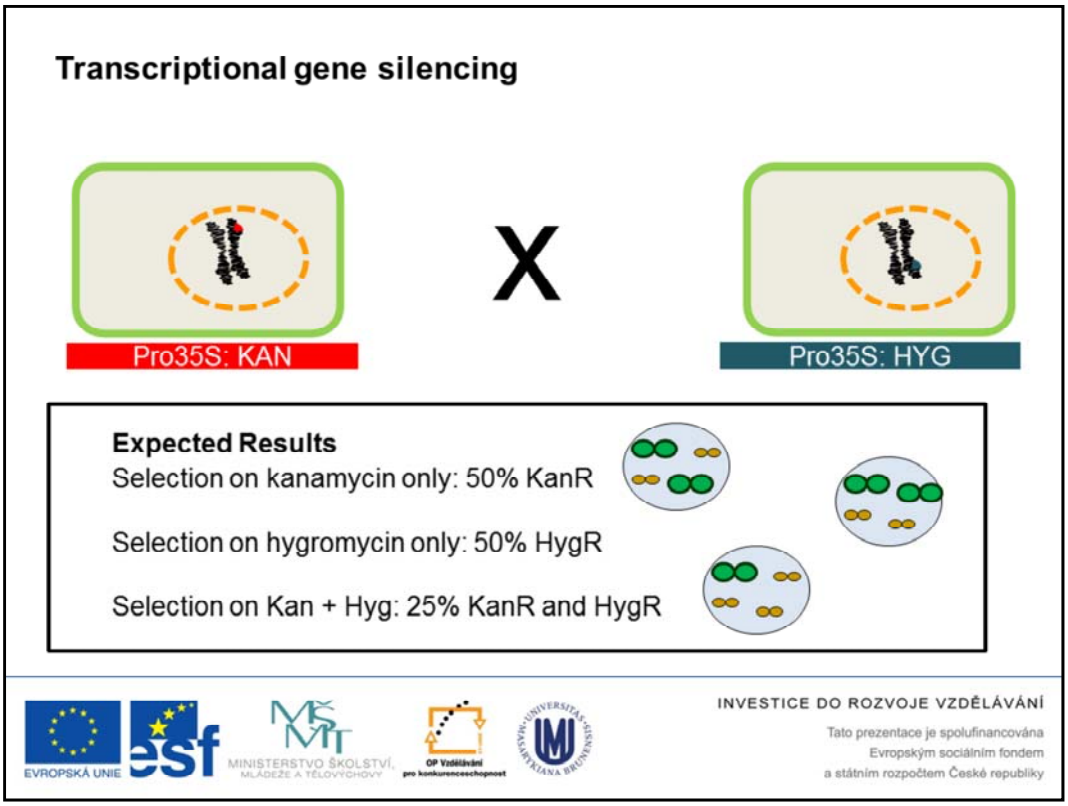
RNA Pol IV contributes to siRNA production. Non-coding RNAs produced by RNA Pol V direct silencing machinery to target sites.



Most of the cellular siRNAs are derived from transposons and other repetitive sequences. In *Arabidopsis*, as shown above, there is a high density of these repeats in the pericentromeric regions of the chromosome.

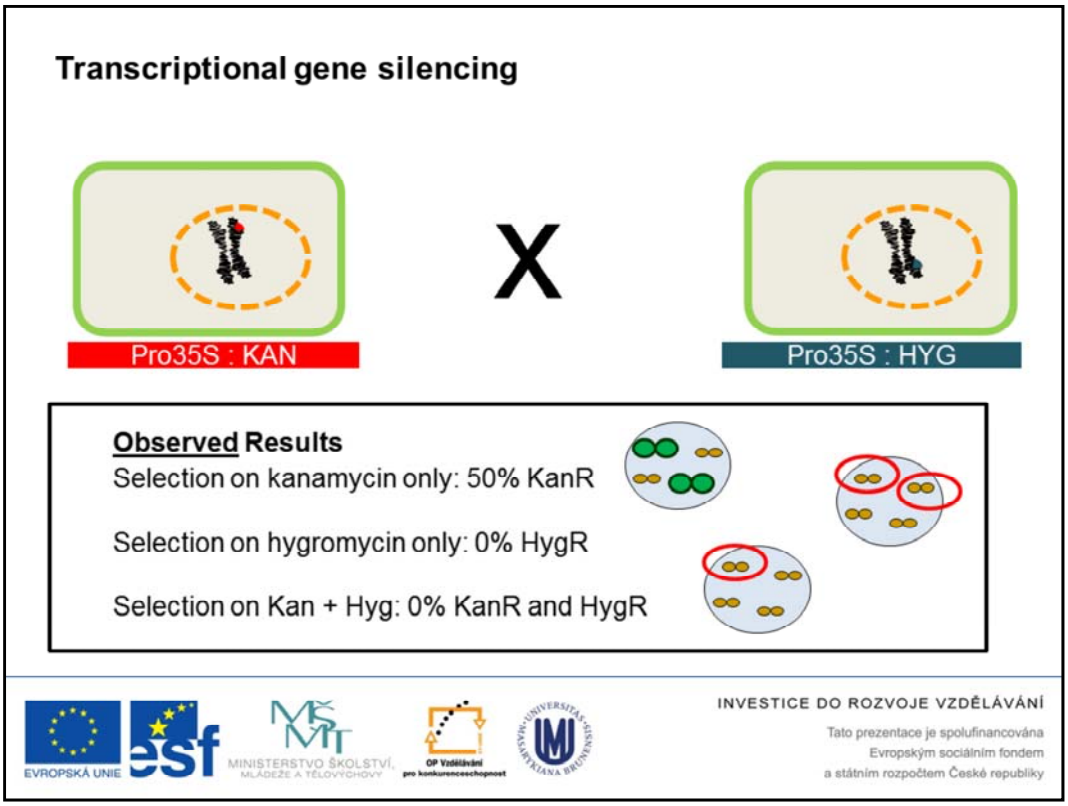
Deep sequencing methods allow the population of siRNAs in a cell to be mapped onto their regions of homology. The abundance of small RNAs and transposons/retrotransposons along each of the five *Arabidopsis* chromosomes is shown relative to the centromere (circle on each line representing the chromosomes). In *Arabidopsis* these repetitive elements occur primarily in pericentromeric regions.

Transposons and repetitive elements are more dispersed in organisms with larger genomes.



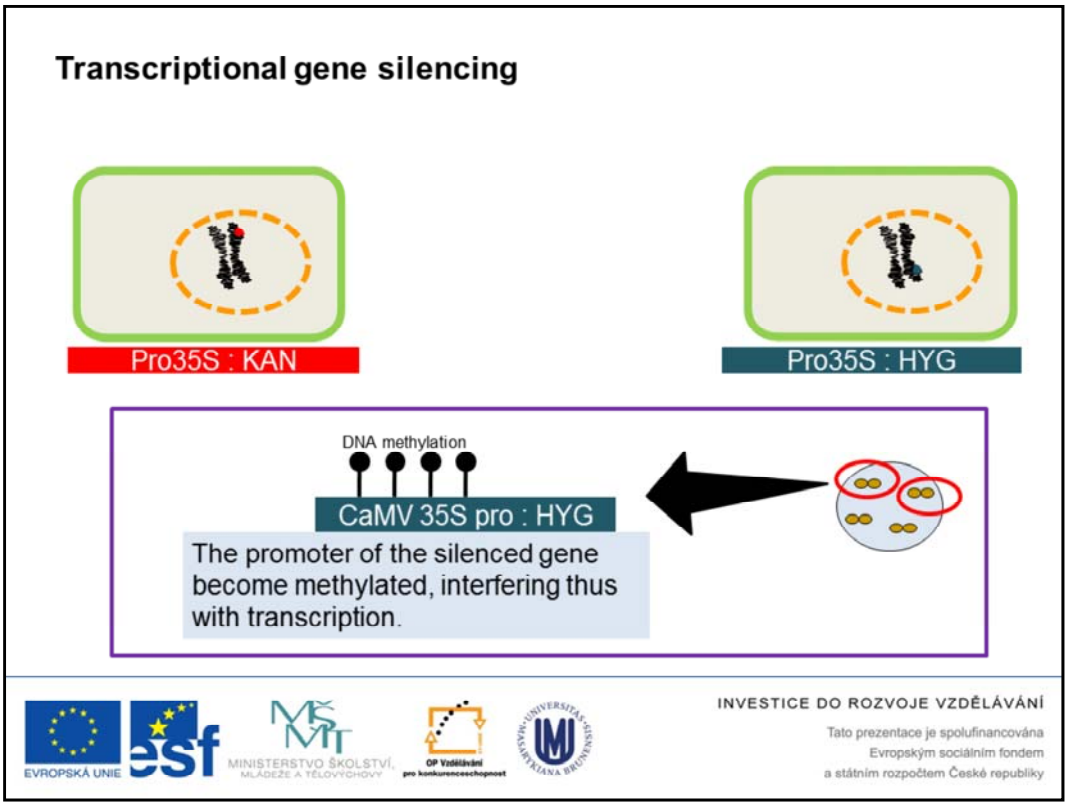
Transcriptional silencing on the example of crossing of two lines of transgenic plants, each of them being heterozygous (hemizygous) for a single copy transgene with different resistance marker.

The green ovals represent cotyledons of healthy, antibiotic resistant seedlings. The parental lines carry one copy of the transgene which acts as a dominant trait. Therefore, 50% of the progeny from the cross should carry any one dominant trait, and 25% should carry both. Note that these transgenes insert at different loci – they are independent. They are not two alleles of a single locus.



Sometimes one of the transgenes was silenced in the progeny carrying both genes.

That is reflected in the observed ratios of observed frequency of plant resistant to both on of the used antibiotics. In this case, the gene for resistance to hygromycin was silenced.



The mechanism of silencing in this case is promoter methylation.

siRNAs - summary

- The siRNA pathway silences foreign DNA, transposons and repetitive elements.
- In plants, siRNAs are produced by the action of Dicer-like proteins dicing dsRNA into 24 nt siRNAs
- The siRNAs associate with AGO proteins and form silencing complexes
- The silencing complexes can act post-transcriptionally on RNA targets, cleaving them or interfering with translation
- **The silencing complexes can also act on chromatin, silencing their targets by DNA methylation or histone modification**

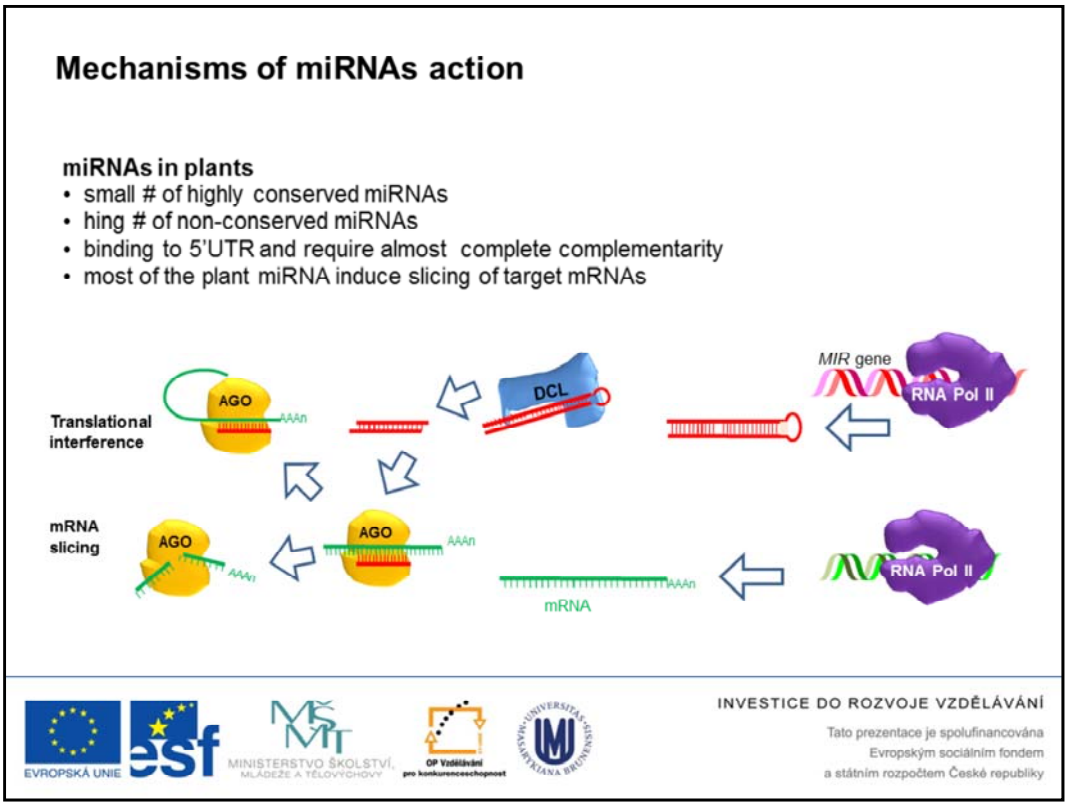
EVROPSKÁ UNIE
MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY
OP Vzdělávání pro konkurenceschopnost
INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ
Tato prezentace je spolufinancována Evropským sociálním fondem a státním rozpočtem České republiky

Outline of Lesson 10

Regulation of Gene Expression during Development

- Overview of levels of gene expression regulation
- Transcriptional gene regulation
  - Modification of the chromatin structure and DNA methylation
  - Transcriptional activation
- Post-transcriptional gene regulation
  - Splicing of hnRNA
  - Translation initiation
  - Localization of mRNA
  - Protein localization
- RNA interference
  - Identification and mechanism of gene expression regulation via RNA interference
  - siRNA-mediated silencing
  - miRNA-mediated silencing

EVROPSKÁ UNIE
esf
MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY
OP Vzdělávání pro konkurenceschopnost
INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ
Tato prezentace je spolufinancována Evropským sociálním fondem a státním rozpočtem České republiky



microRNAs slice mRNAs or interfere with their translation. Thus, in contrast to siRNAs, miRNAs do not induce transcriptional silencing.

miRNAs are thought to have evolved from siRNAs, and are produced and processed somewhat similarly. Plants have a small number of highly conserved miRNAs, and a large number of non-conserved miRNAs.

miRNAs are encoded by specific *MIR* genes but act on other genes – they are trans-acting regulatory factors. miRNAs in plants regulate developmental and physiological events.

Plant miRNA bind preferentially to the 5'UTR or coding sequence and require almost complete complementarity with their targets.

In contrast to that, in animals there is much higher proportion of genes being regulated by miRNAs. The most of animal miRNA bind to the 3'UTR and extensive sequence complementarity is not required.

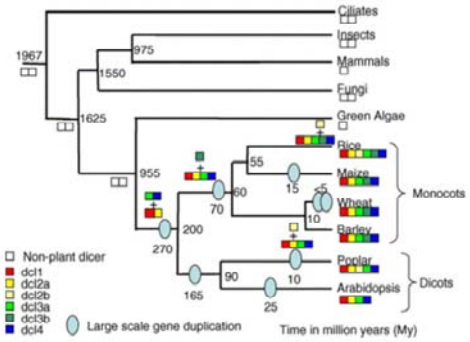
In animals, dominant regulatory mechanism mediated by miRNAs is

translational repression, while in plants, most of the miRNAs initiate slicing of target mRNAs.

miRNAs and siRNAs are processed by related but different DCL proteins

AtDCL1 produces **miRNA**

AtDCL2 - 4 produce **siRNA**



Reprinted from Margis, R., Fusaro, A.F., Smith, N.A., Curtin, S.J., Watson, J.M., Finnegan, E.J., and Waterhouse, P.M. (2006) The evolution and diversification of Dicers in plants FEBS Lett. 580: 2442-2450 with permission from Elsevier.

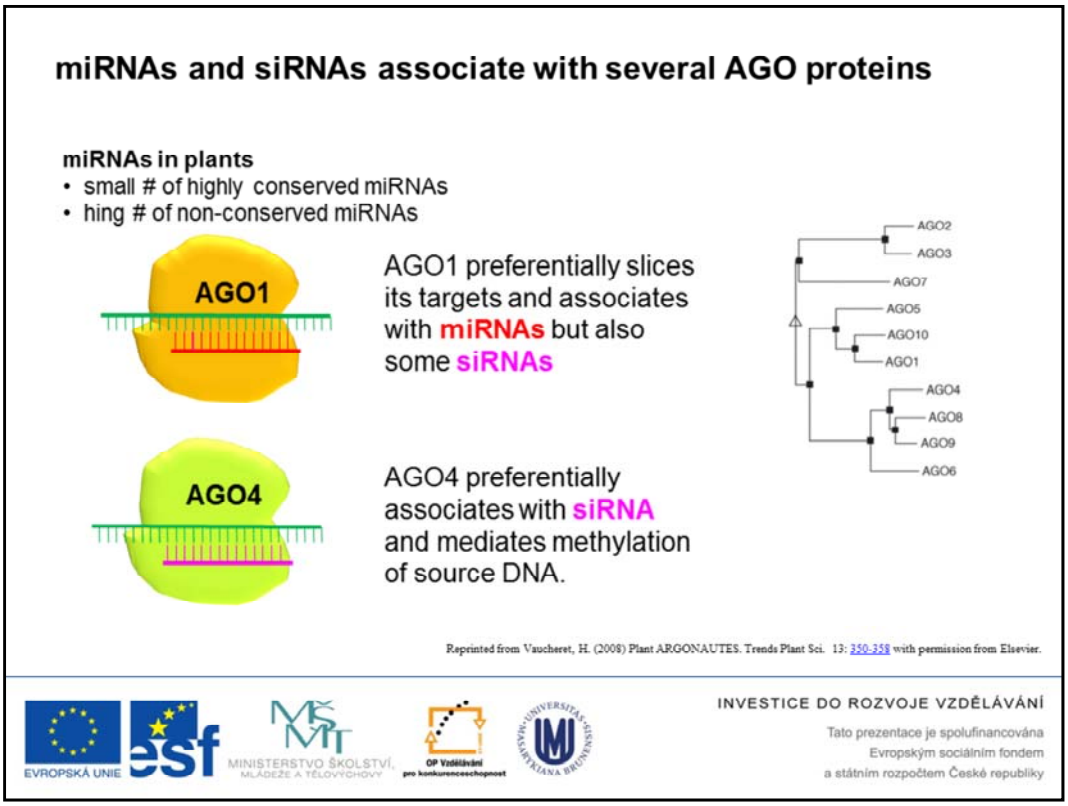
EVROPSKÁ UNIE esf MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY OP Vzdělávání pro konkurenceschopnost

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

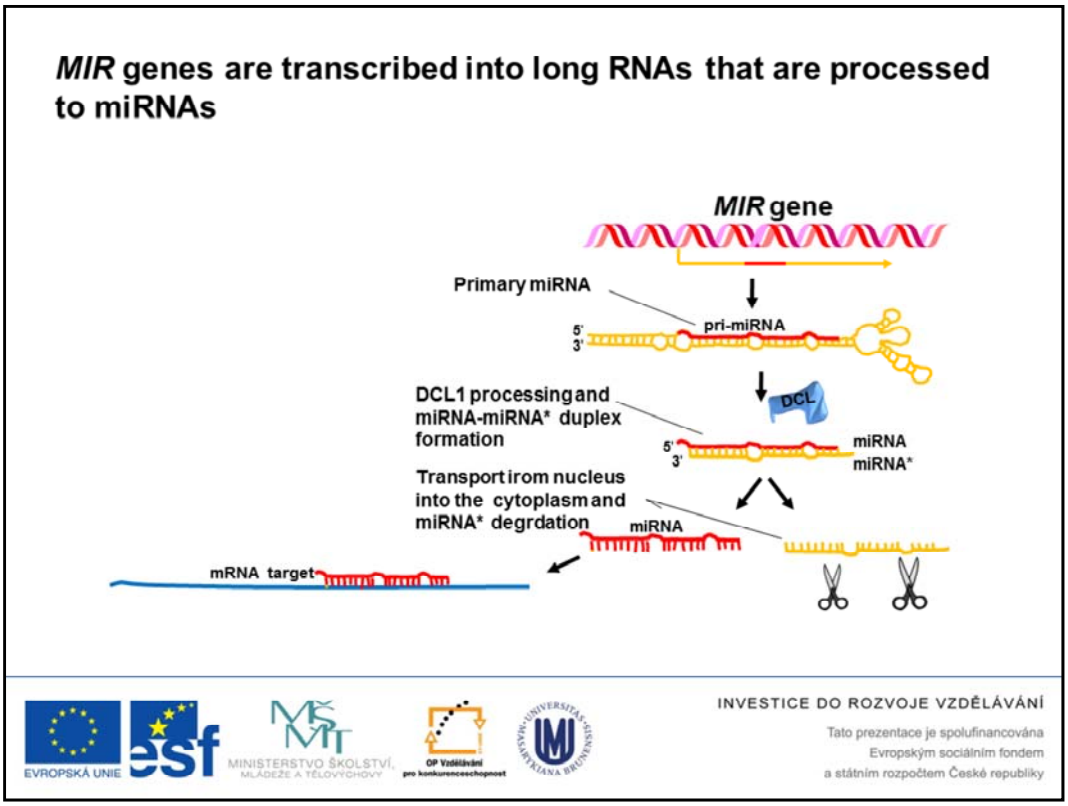
Tato prezentace je spolufinancována Evropským sociálním fondem a státním rozpočtem České republiky

Plants have 4 or more DCL proteins, more than found in other organisms. The amplification of DCL proteins is thought to allow plants great flexibility in pathogen defence responses.

Note that mammals make do with one dicer, and insects and fungi with two. Like most components of the siRNA pathway, dicer-like genes are amplified in plants.



Arabidopsis has 10 AGO proteins. They are not all well characterized and there is some functional overlap.

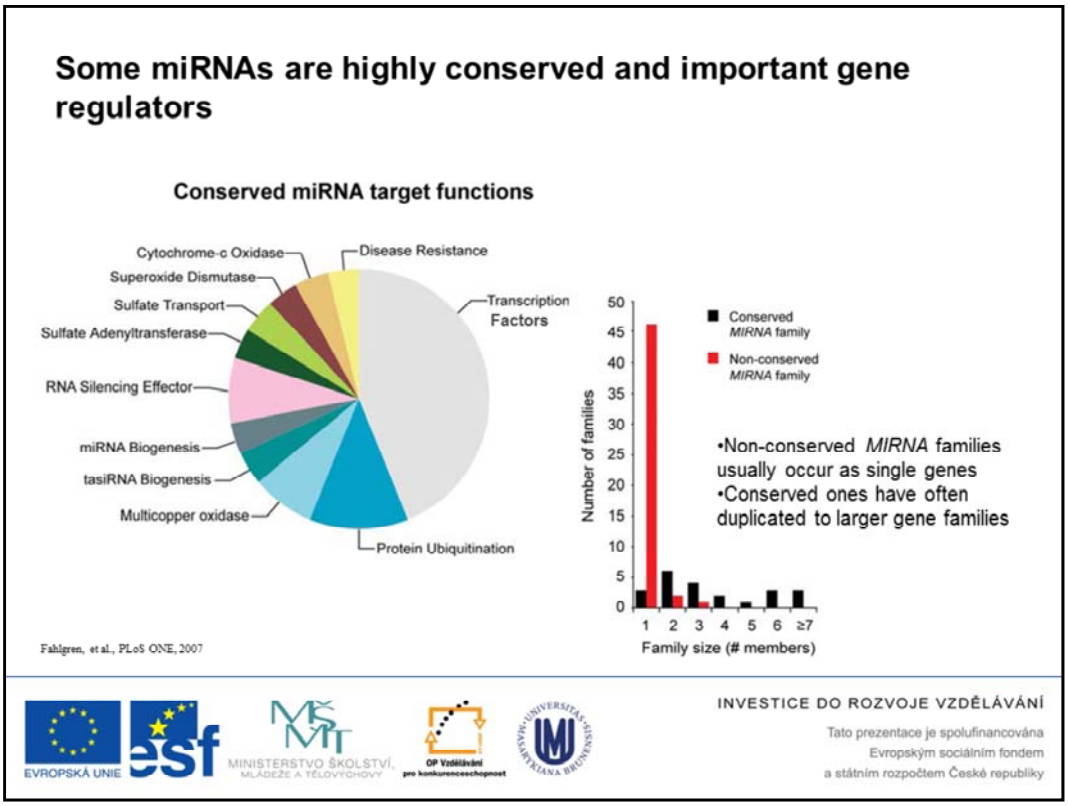


miRNAs are encoded by *MIR* genes. *MIR* genes are transcribed into long RNAs called primary miRNA (pri-miRNA) with partial complementarity.

The pri-miRNA transcript folds back into a partially double-stranded stem-loop or hairpin structure, which is processed by DCL1. The product of DCL cleavage dsRNA with a 2-nucleotide 3´ overhang is called a miRNA-miRNA* duplex.

The miRNA-miRNA* duplex is transported from the nucleus into the cytoplasm by the homologue of exportin 5 HASTY (will be discussed later).

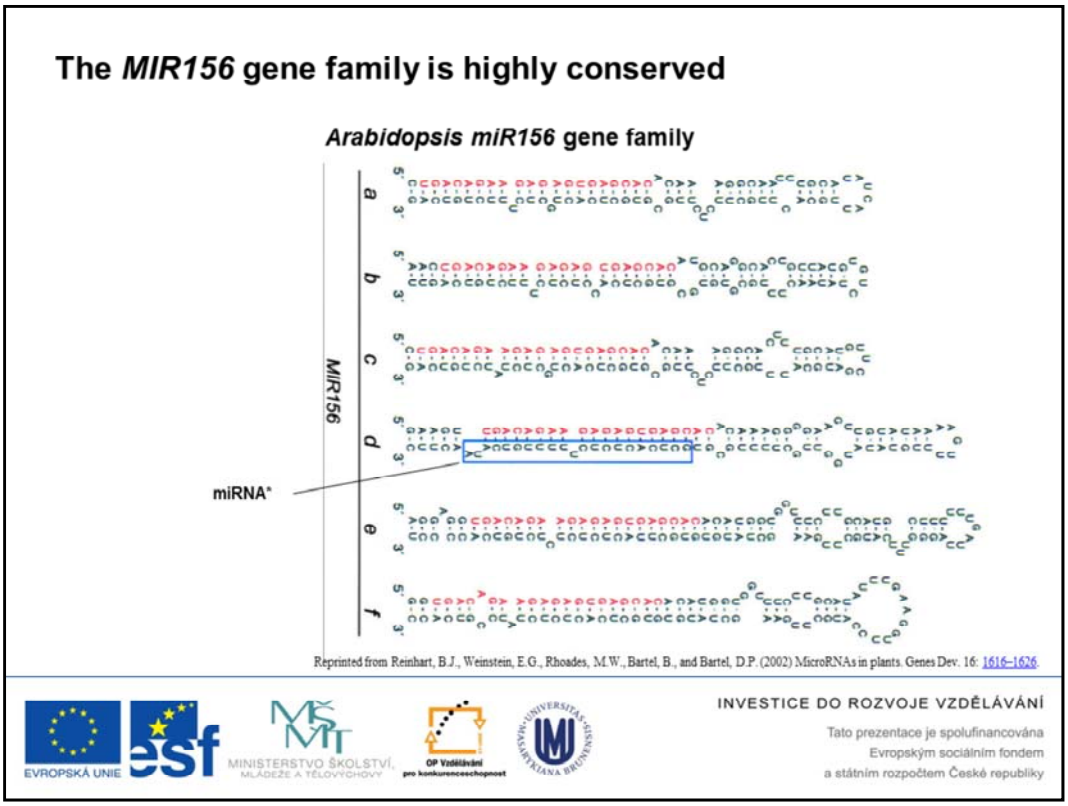
Subsequently, the miRNA* strand is degraded.



In spite of there are quite few genes regulated by miRNAs in plants, there is high number of *MIR* genes.

However, most of them are evolutionary non-conserved and probably provide no competitive advantage. It has been suggested that there is high rate of birth and death and only few of them became stabilized in the genome.

Nearly half of the targets of conserved miRNAs are transcription factors.



Among the evolutionary conserved, i.e. important regulators belong miRNAs that target TF. E.g. miR156 family is highly conserved within the plant kingdom (it is found in angiosperms as well as mosses)

miR156 is encoded by six or more genes in *Arabidopsis* and targets transcription factors that control developmental phase changes (see further slides).

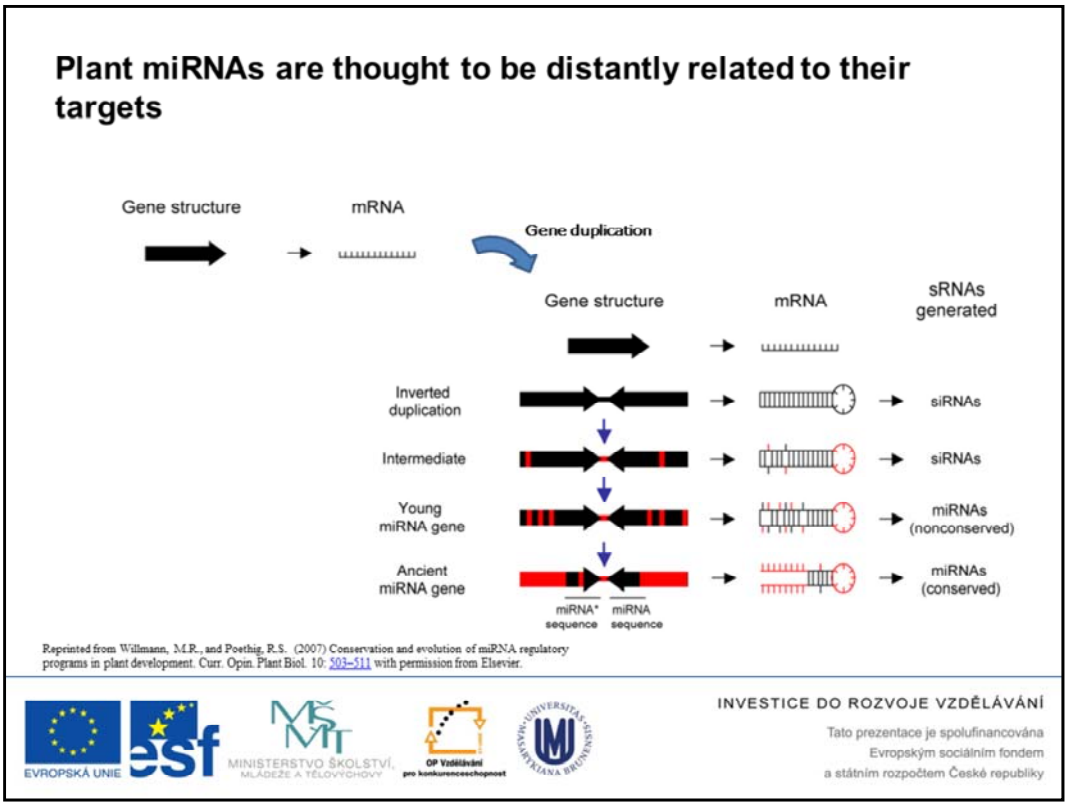
The red sequence indicates the miRNA produced from each of the six *Arabidopsis MIR156* genes (*MIR156A* – *MIR156F*). The boxed sequence shows the miRNA*.

Targets of some conserved miRNAs

miRNA gene family	Target gene family	Function
156	SPL transcription factors	Developmental timing
160	ARF transcription factors	Auxin response, development
165	HD-ZIPIII transcription factors	Development, polarity
172	AP2 transcription factors	Developmental timing, floral organ identity
390	TAS3 (tasiRNA) which acts on ARF transcription factors	Auxin response, development
395	Sulfate transporter	Sulfate uptake
399	Protein ubiquitination	Phosphate uptake

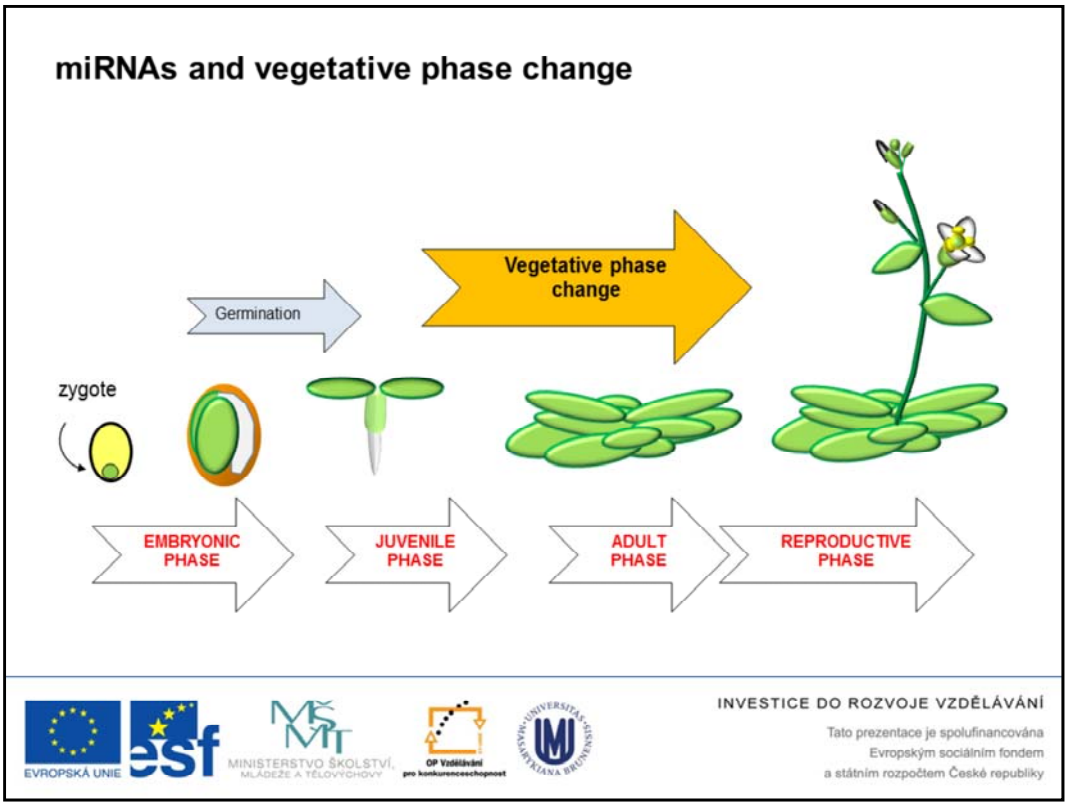
Adapted from Willmann, M.R., and Poethig, R.S. (2007) Conservation and evolution of miRNA regulatory programs in plant development. Curr. Opin. Plant Biol. 10: 503–511.

The table shows targets of some of the evolutionary conserved MIR genes (or rather respective miRNAs). As discussed previously, these are mostly *MIR* genes encoding miRNAs recognizing TFs.



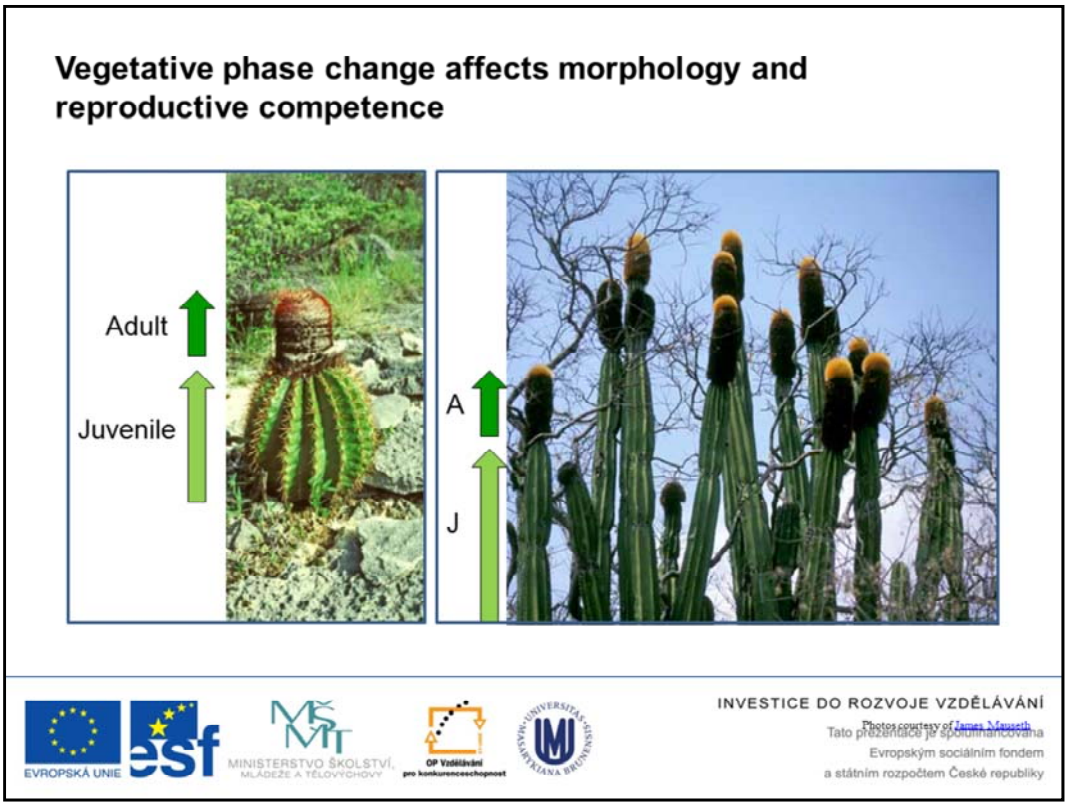
Plant miRNAs are thought to be derived from their target sequences following gene duplication, inverted duplication and divergence.

Only some miRNAs confer selective advantage and are retained and further duplicated.

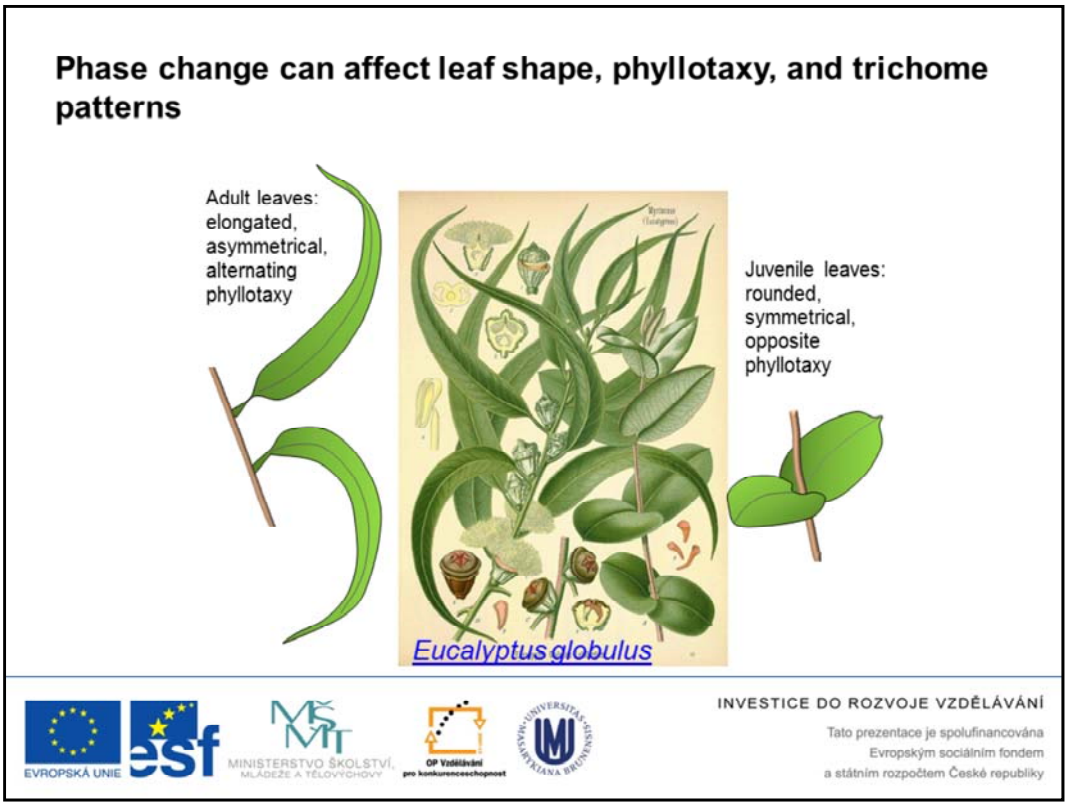


Vegetative phase change is the transition from juvenile to adult growth in plants.

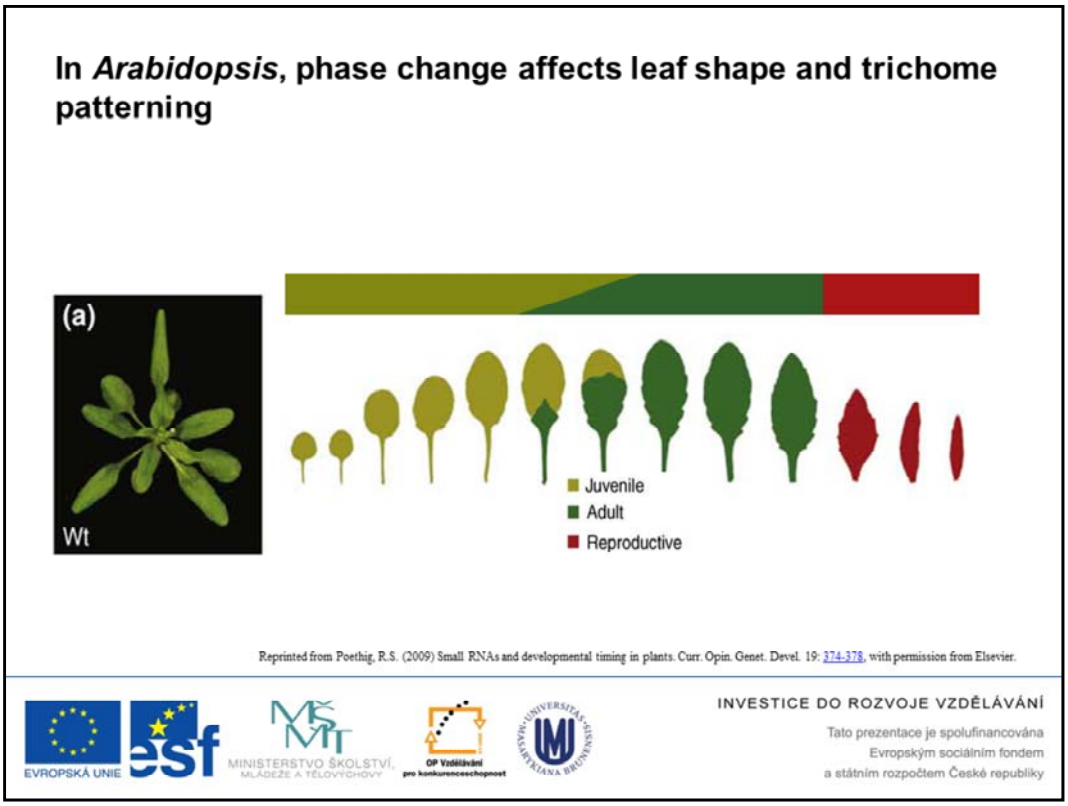
The transition includes changes in the leaf shape and arrangenment, internode length, and the accumulation of epidermal hairs (trichomes) and waxes.



Some cacti have very different juvenile and adult growth patterns.

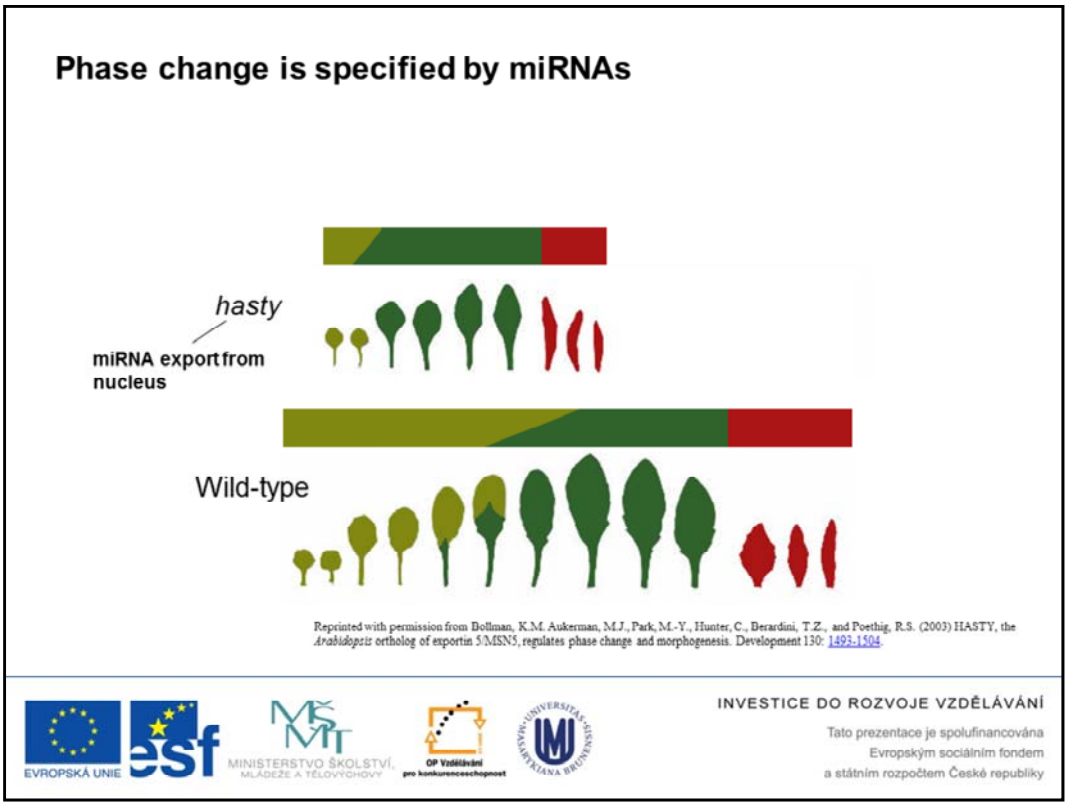


Eucalyptus leaves are strongly dimorphic, as are leaves of holly and ivy. In other plants including *Arabidopsis* and maize the change is more subtle.



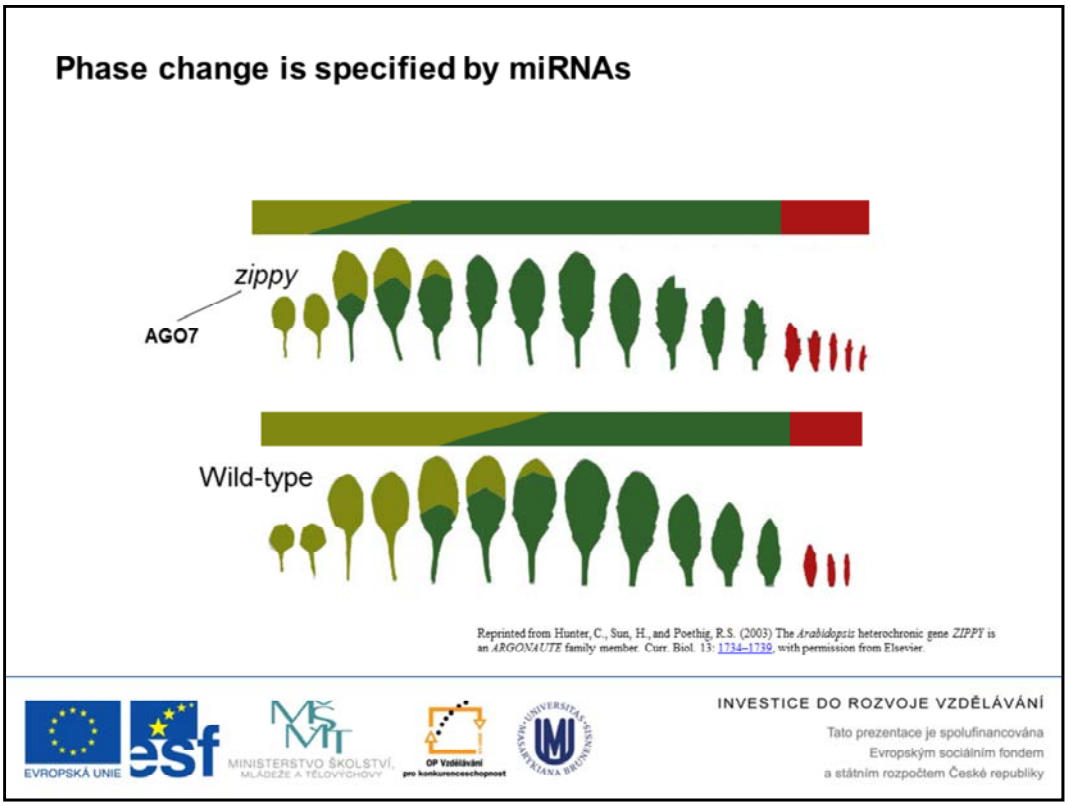
In *Arabidopsis*, juvenile leaves are rounder, less serrated, and have trichomes only on the upper (adaxial) surface; adult leaves also have trichomes on the lower (abaxial) surface.

In the next few slides the leaves are colour coded as juvenile (olive), adult (green) or reproductive, also sometimes called cauline leaves (red) based on their morphology and anatomy. The boxes above the leaves schematically indicate the relative duration of each phase (for comparing wild-type and mutant plants).



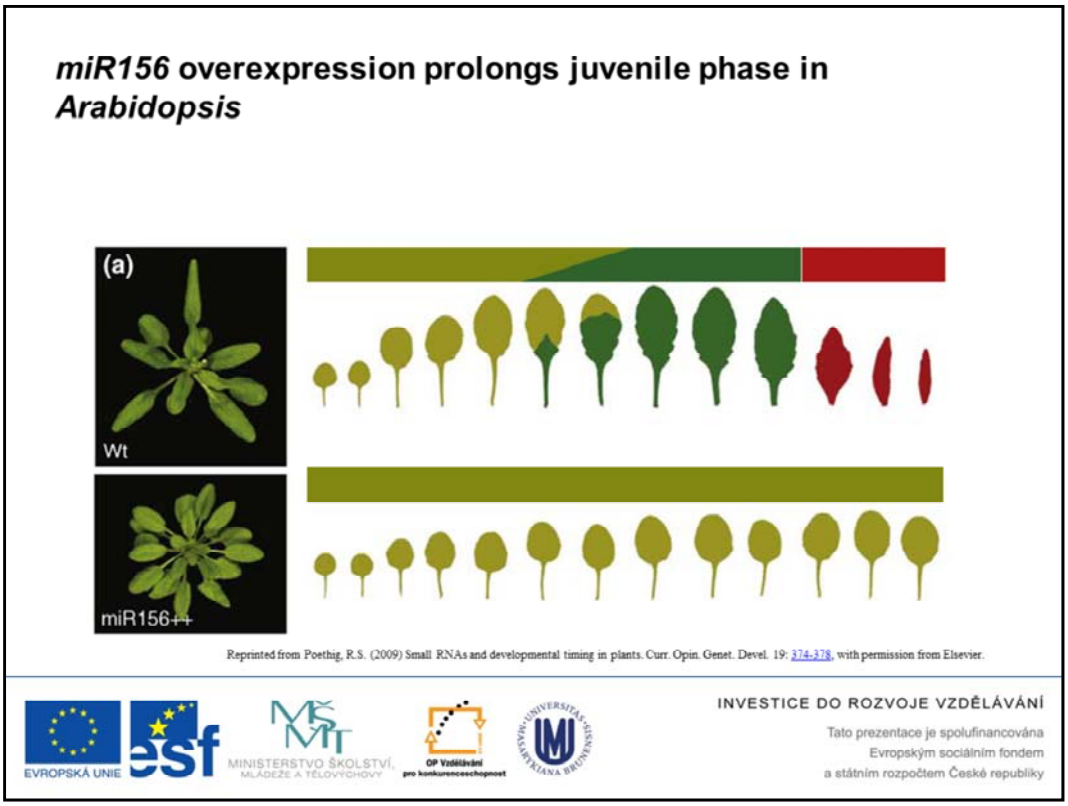
HASTY, with a shortened juvenile phase, encodes a protein needed for miRNA export from nucleus to cytoplasm.

It's easy to see from these figures that *hasty* has a shortened juvenile phase relative to wild-type.



Loss-of-function *zippy* mutants prematurely express adult vegetative traits. *ZIPPY* encodes an ARGONAUTE protein, AGO7.

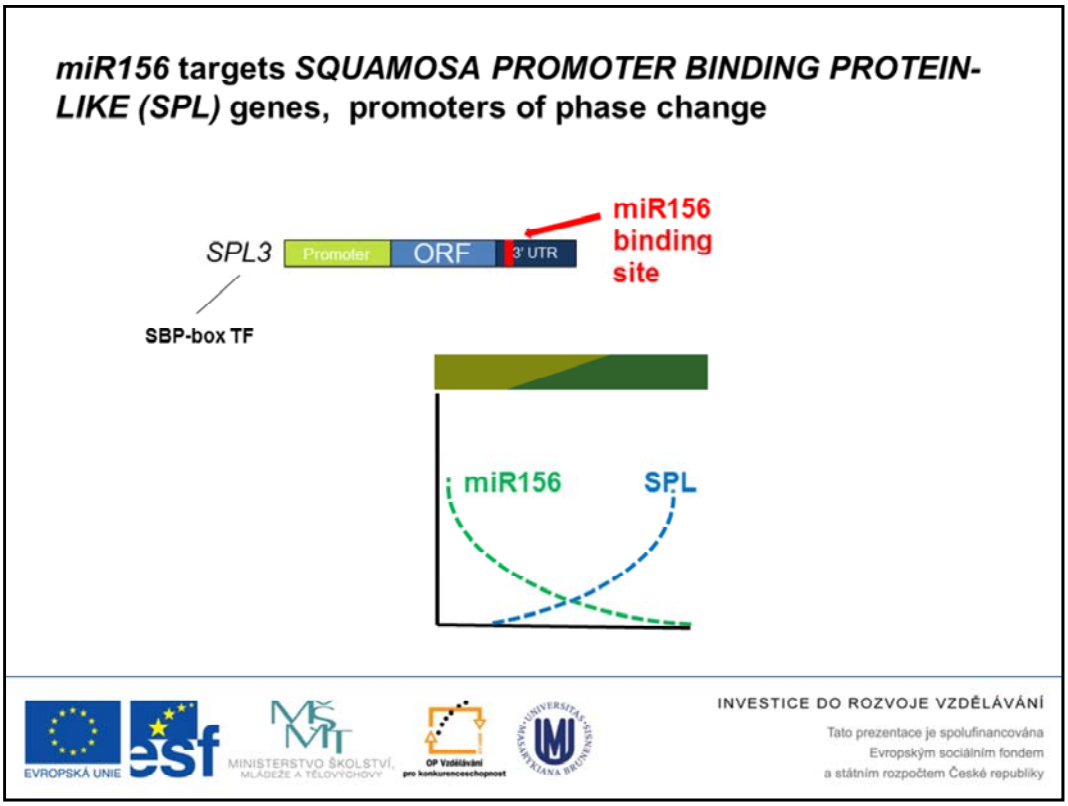
Zippy prematurely expresses adult traits; this is a different effect that hasty, in which the total non-reproductive phase was shortened as well.



As could be clearly seen from the phenotyope of transgenic line with ectopic overexpression of miR156 gene, there is dramatic elongation of the juvenile phase.

Highly conserved miR156 targets SQUAMOSA PROMOTER BINDING PROTEIN-box (SBP-box) genes in maize, called SPL genes in *Arabidopsis.*

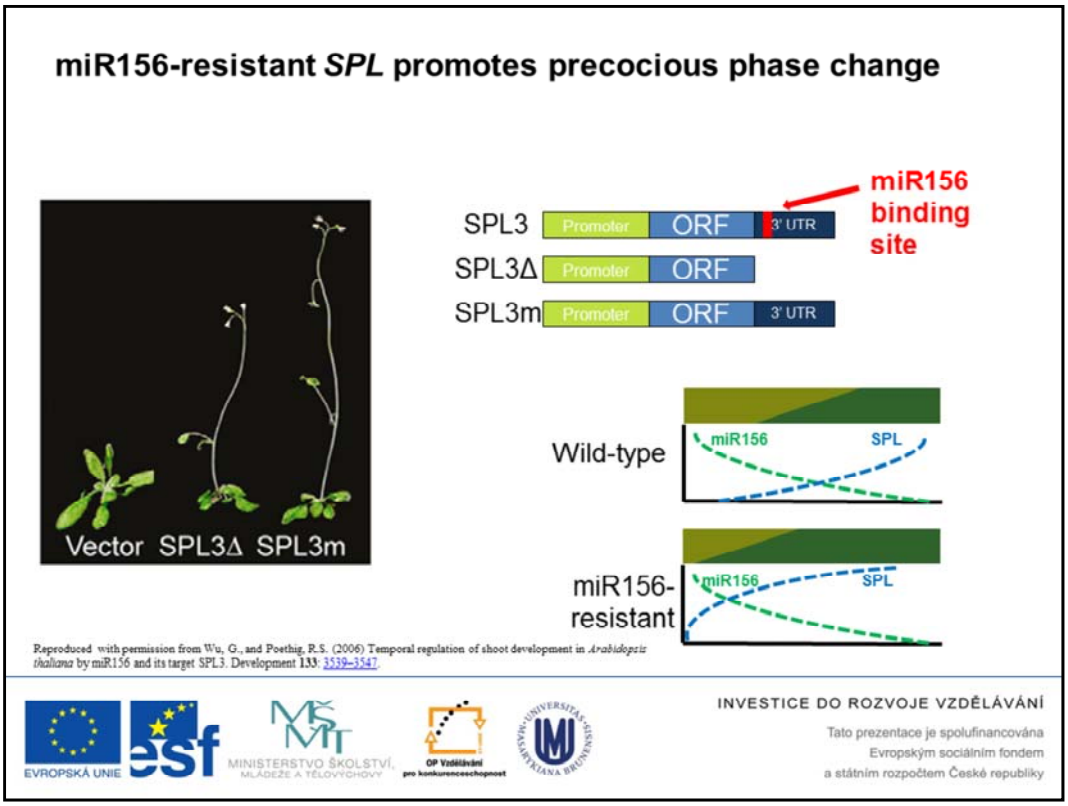
SBP-box genes are TFs that regulate expression of key regulators of flowering TF, SOC1, LFY, and AP1 and thus regulate the transition from juvenile to adult/reproductive growth phase.



The *SQUAMOSA PROMOTER BINDING PROTEIN-LIKE (SPL)* genes are a family of SBP-box transcription factors that are miR156 targets.

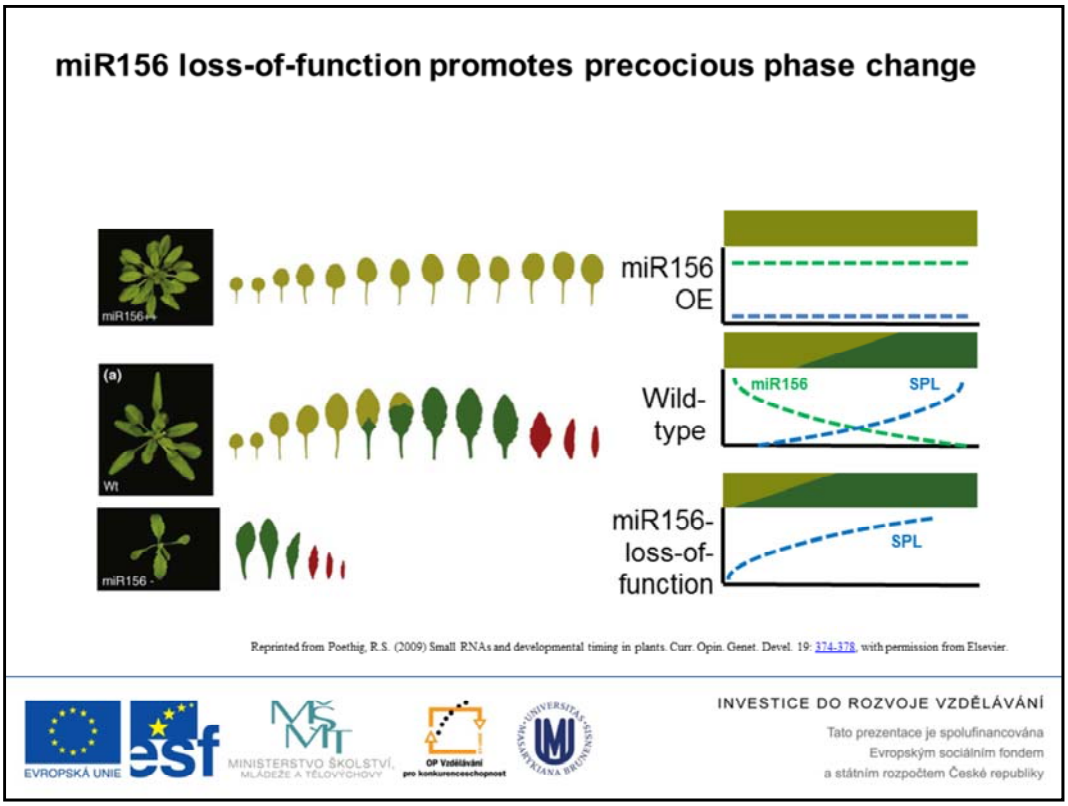
In wild-type plants, *miR156* expression decreases with plant age, allowing SPL to accumulate and promote phase change.

The graph indicates the relative activity of *miR156* and *SPL*, showing that as *miR156* levels decrease *SPL* levels increase. When *miR156* levels are high, *SPL* mRNA is cleaved and the protein can't accumulate.

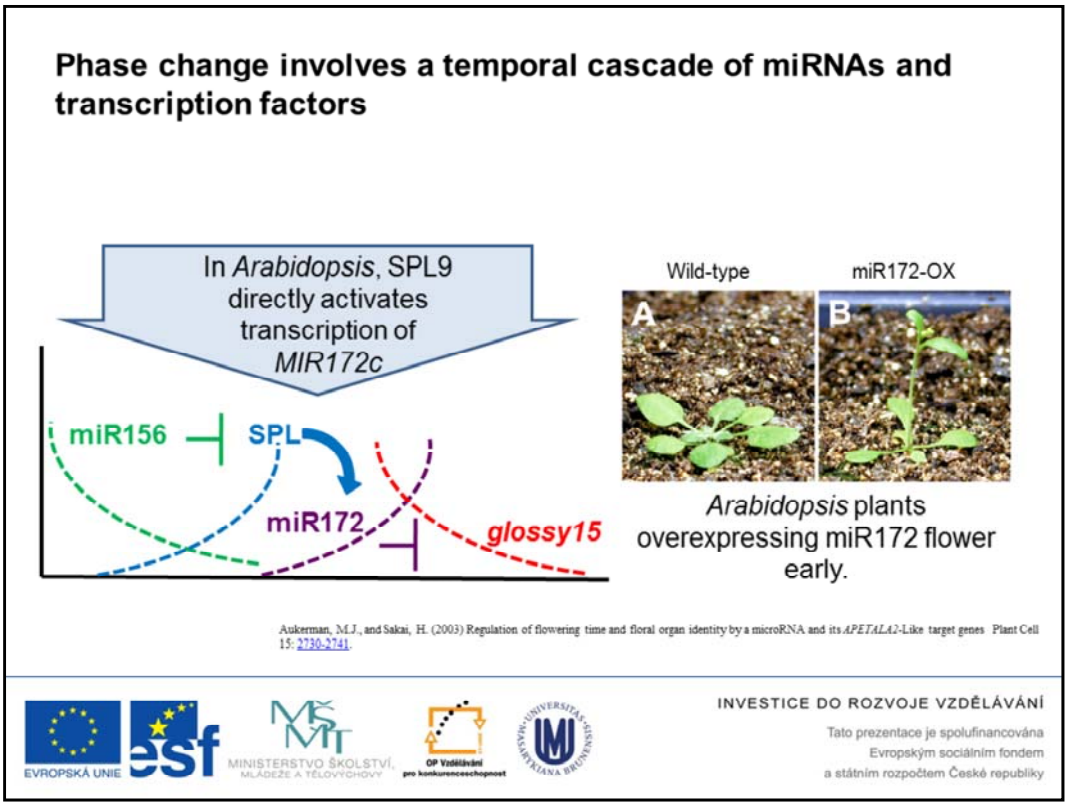


The miR156 effect can be eliminated by deletion of the 3’ UTR or mutations that interfere with miR156 binding, as shown schematically.

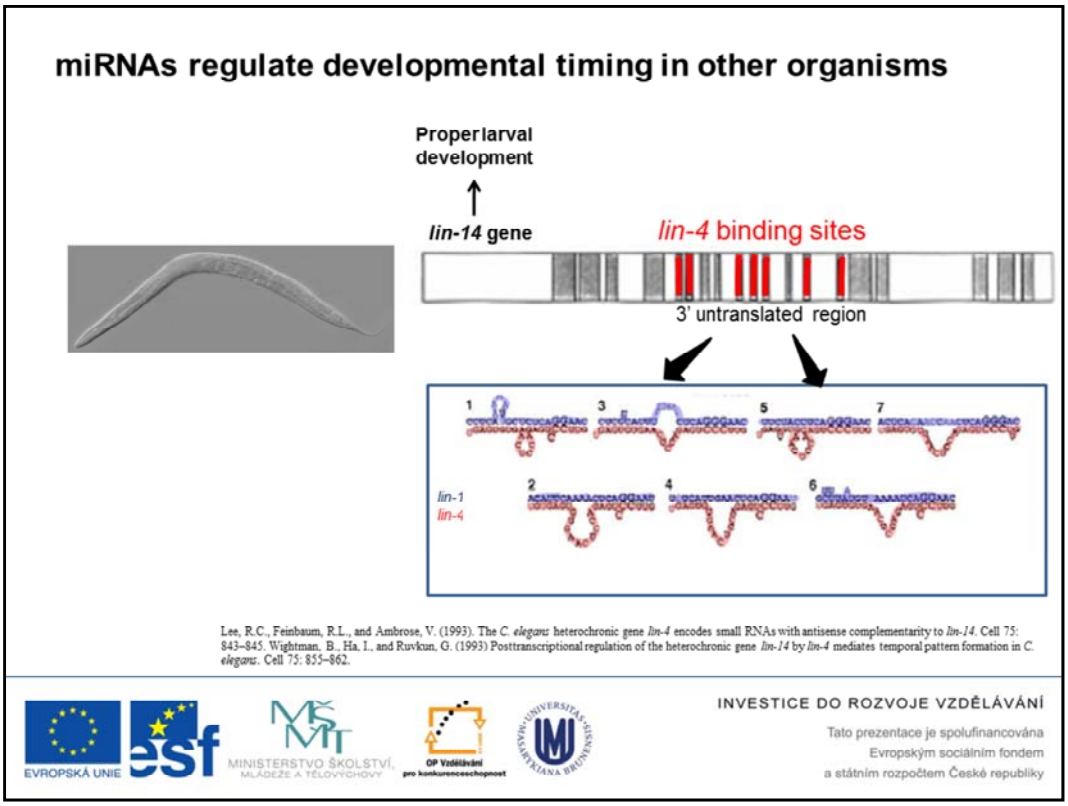
In either case, SPL protein accumulates early (it is indifferent to the presence of mir156), promoting phase change. Note that the miR156 resistant plants are flowering when the control plants are still vegetative.



Similarly, interfering with miR156 production allows SPL to accumulate early and promotes precocious phase change.



Interestingly, in *Arabidopsis*, *MIR172* gene is induced by one of the SPL transcription factors, suggesting that a cascade of miRNA-regulated genes controls phase transition.

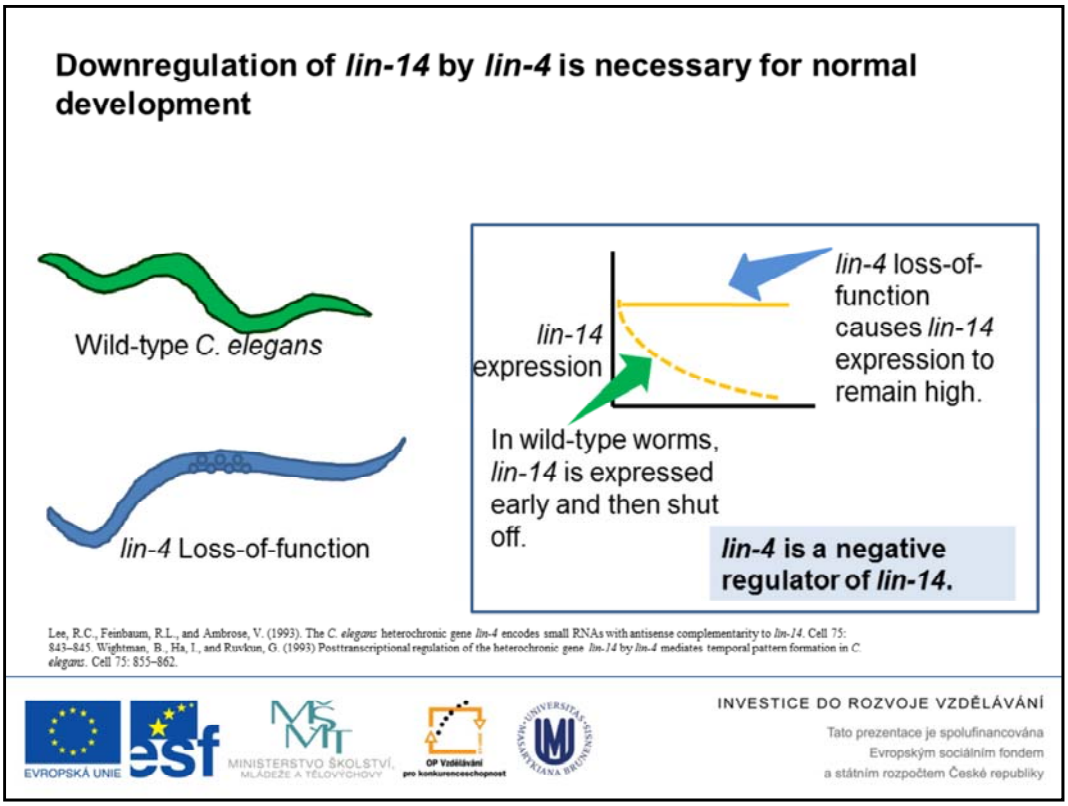


miRNAs were discovered in studies of developmental progressions in the nematode *C. elegans*. miRNA encoded by *lin-4* is required for proper larval development.

Genes *lin-14* and *lin-4* promote or repress developmental progression in the nematode *C. elegans*. Genetic studies showed that *lin-4* acts antagonistically to *lin-14* but that *lin-4* does not encode any protein.

The functional domain of *lin-4* was identified and found to be complementary to several regions in the 3´ UTR of *lin-14* (see the figure, red).

Subsequently, the interaction between these complementary sequences was shown to be necessary for *lin-4's* repression of *lin-14,* revealing the basis for gene regulation by miRNAs.



For proper developmental progression in *C. elegans*, the juvenile-promoting activity of the *lin-14* gene has to be repressed by the action of the miRNA encoded by *lin-4*.

miRNAs and phase change - summary

- Vegetative phase change affects morphology and reproductive competence
- miRNAs contribute to the temporal control of gene expression and phase change
  - *miR156* promotes juvenile phase by preventing *SPL* gene accumulation
  - *SPL* genes promote phase change and flowering
  - In *Arabidopsis*, a SPL protein promotes transcription of *miR172*
  - *mir172* triggers phase change by interfering with *GLOSSY15* expression
- In the nematode *C. elegans*, *lin-4* silencing of *lin-14* is required for developmental progression

Key Concepts

Regulation of Gene Expression during Development

- **Regulation of gene expression** occurs at **different levels**, from **transcriptional** till the **posttranscriptional** and **posttranslational**
- **Basal promoters** are co-regulated in a **combinatorial way** via spectrum of **positive** and **negative factors**
- **mRNA** and **protein localizations** belong to the most important posttranscriptional regulations of gene expression
- **RNA interference** is **natural** and powerful **mechanism** allowing regulation of gene expression at both **transcriptional** and **posttranscriptional** levels
- **dsRNA** is either **trigger** or **intermediate** in the **RNAi-mediated regulation**
- **siRNA** and **miRNA** are **two major effector molecules** regulating different and complementary spectrum of target genes

EVROPSKÁ UNIE
esf
MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY
OP Vzdělávání pro konkurenceschopnost
INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ
Tato prezentace je spolufinancována Evropským sociálním fondem a státním rozpočtem České republiky