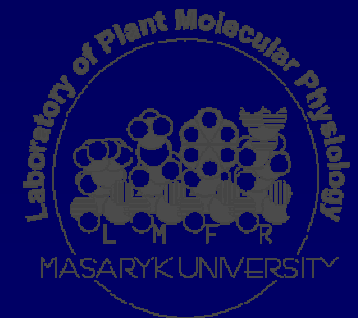


Základy proteomiky 2009

Proč právě proteomika?

Jan Hejátko



Základy proteomiky 2007

▪ Schéma přednášek ze Základů proteomiky 2007

1. blok, 20. 3. 2009

- Proč právě proteomika? (Jan Hejátko)
- Exprese a purifikace rekombinantních proteinů (Radka Fohlerová)

2. blok, 27. 3. 2009

- Funkce proteinů (Lubomír Janda)
- Určování třírozměrných proteinových struktur metodami proteinové krystalografie (Jaromír Marek)

3. blok, 3. 4. 2009

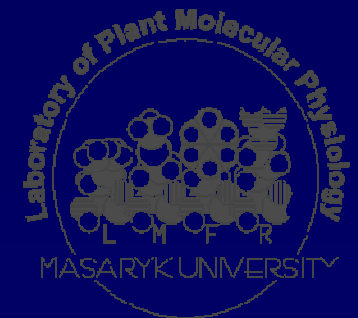
- Metody separace proteinů (Hana Konečná)
- Hmotnostní spektrometrie proteinů (Zbyněk Zdráhal)



Základy proteomiky 2007

Proč právě proteomika?

- Postgenomová éra a co s informacemi, které neumíme číst
- Genotyp vs. fenotyp, aneb co všechno se děje při vyjadřování
- Od genu k proteinu a zpět
- Přístupy současné proteomiky
 - exprese, purifikace a analýza rekombinantních proteinů
 - analýza vztahu mezi strukturou a funkcí proteinů
 - diferenční proteomika
 - analýza posttranslačních modifikací



Základy proteomiky 2007

Proč právě proteomika?

- PROTEOME = PROTEins expressed by genOME (konference 2-D ELFO, Siena, 1994)
 - DNA: GENOME, HAPLOME, EPIGENOME
 - RNA: TRANSCRIPTOME
 - **PROTEIN**: ORFEOME, PROTEOME, LOCALISOME, INTERACTOME, METABOLOME, PHENOME, ...
 - PHENOME: kombinace různých dat, zahrnujících fenotyp, expresní data různých (ideálně všech) genů daného organismu a proteinová data (interakce, jednotlivé vlastnosti proteinů, ...)
- Proč vůbec studovat proteiny, když máme tolik genetických dat? (sekvence genomů, expresní profily genů, fenotypy mutantů,...?)



V koncovém výsledku, tedy *fenotypu*, se vždy projeví regulace na všech úrovních, od genu po protein a jeho modifikaci



Na konci je vždy **BIOLOGICKÝ PROBLÉM !!!!**



Anotace genů a odhad aminokyselinových sekvencí předpokládaných proteinů

Postgenomová éra a co s informacemi, které neumíme číst

- dnes k dispozici v databázích více než 700 000 záznamů (nr databáze) různých organismů
- Získané informace lze zpracovat pomocí **bioinformatiky**



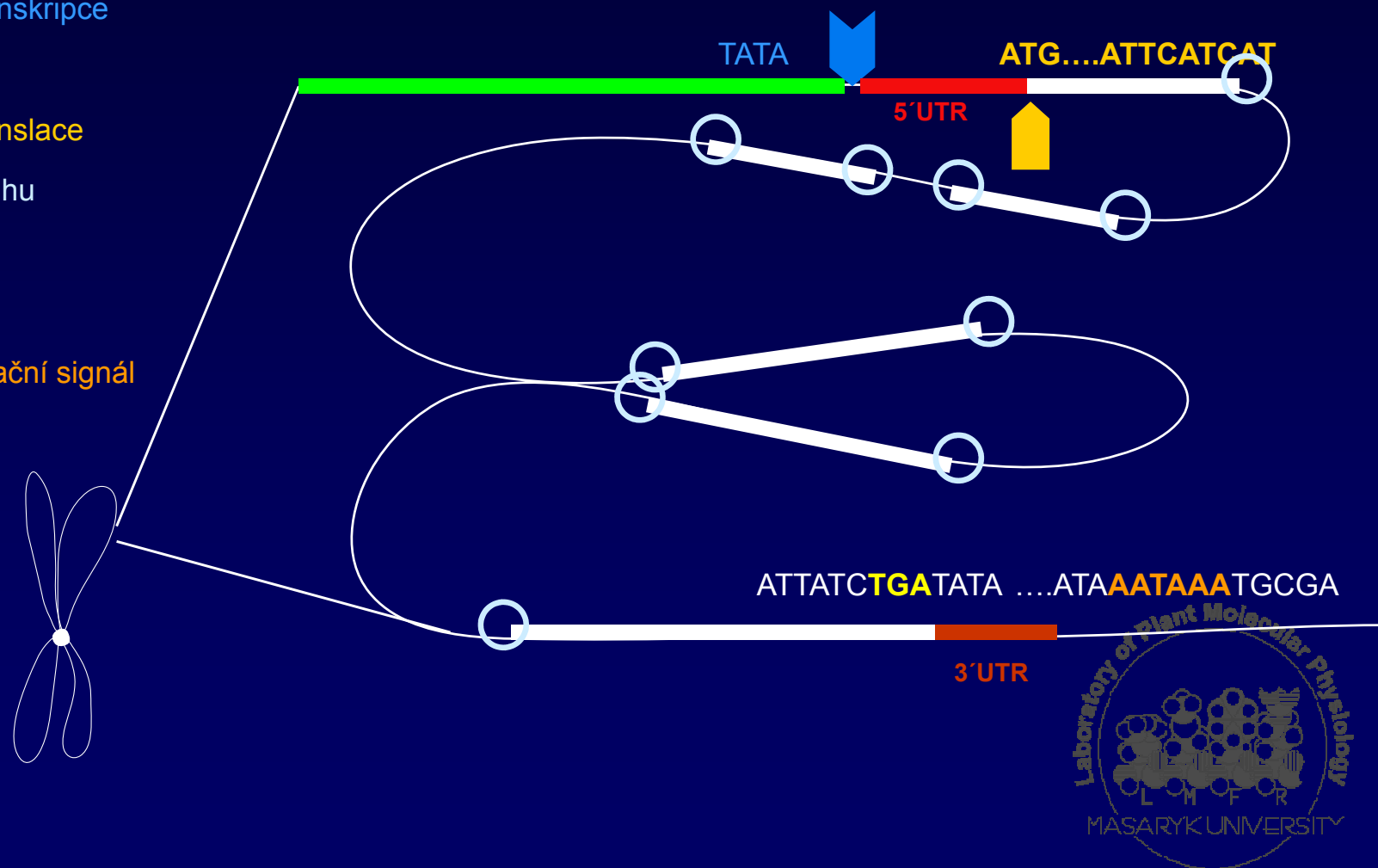
```
19470700 cacacctatagtatagctcaattctagataaaatataagaaatggatcttgagaatcattttttttgtattcttttggtaa
gctecgaggaagaagataaataatgaaaagagcttttttagggttttatcattctctcttgaactttgcaaaaacgtgaaatgtaaggc
19470500 tgttgctttttatagctatcgcttcttacaataagttaacaatgcttctcgtagaatgcaaaaacatttggtagcgcgtgatt
ttcagtggtcttctttgcagcagcttcttcttggaggactaatcaagacagaaatctgttctctaaaaacgatcgcgcttct
19470300 ttgacgagctcttgatctttagaatcaaaatataagggatcacgagatacacgtattaattattttttttttttttgctt
tcactcaaatgatggtgaaagttacaaagcttgggcttcacgtccaattgtggctcttttgcgtcctggtaattctgctttct
19470100 atgattctacatttctactcatctogttcttggtttttcaaatgatataattatttgggtgtatatacaccattctatgtaattt
ttcctgggtggttggtttctgagtgcatttggatctcaaatggcgcaacaacaacggagaacctagtcaagaggctgcttcatt
19469900 caagtctagtttcgggagattgaaaacatcggaaaatttacatatgccaagacaaaacttatctacgatcggtttagcgagaggtt
caacaacgacactggttttacagagattcaaacacaggttggtaaaaactaatcacataaattcaattattcttagttattatc
19469700 tatataacattaactataatttttagttggttggttggttattattggttcttcagatcgcaccattggtggtttagctt
gtctcacaagtttctgtacatcagtagggacggctctcatgtttcttacattgcagaatcaaacacaagtgctcgtgttttgct
19469500 caagtcgtggagactacacttggtagactcaaacctggatcagttaactggctcgttcaacgggaactcaacgaaatctcac
tacagattggttccaagcagcacagagtaataactacactacagcctttgtaggaacgagcttggggaggagaagataacgaga
19469300 gttagcttgtagcaagaaggtcttggtttctttagggtttccggttaagactttaaaccgaagtttgaacagtttgaatct
acatgtggacaaaaggacgggacgggtgcttggctggaaggttcaactgaatgattctttcttcatctccaatggctcogatttgc
19469100 ctccctctggtctcaatgcatccctgaaaattgcagttccagtggtcactcagaggtggagatcaaaagattaagataccaagctt
gtttcgggcttctctggttaaatactgaaacataattcactttgatgtagtaaaaatgcatcgacttgtggttctcagctt
19468900 tttgccagagatacacactcatgtttccaacaaggaggagcaaacgcgatcaagcaccgaagcggaaaaggcaaaatataca
atcttctggcttccggttggcctgtatggtttgtggtttatgatgcaagcaacaaggagagatgcatatgctgcaacgc
19468700 gcgacacaacaagctgagagaaagagcatgaacaagagtcagcatttgcaaatgctagccacgatattagaggtgaccttgc
ttgataatgtggtgatggagttaaaactggctccgaegttagacaccactctcaaccaagtgaaatggttggcgaaggattt
19468500 tctttagctttctcttatgcttttcttcaacttctctcaacagaaaattcttctctatggttggtaaaaattacagctctgct
tgagcaaaatcgaaagcgggaagatgcagtttagtggaagaagatttcaacttgcgaaacttcttgaagacgtcatcgatttt
19468300 gaagaaaggggttgatgtagtttggatccgcacgatggctcggtttcaaaattctcgaatgtacgaggggatagtgccagac
aatcttgttagcaatgctgtcaagttcacctgcagcgggcacattgcggttaagagcttgggctcagagggcaggttccaatac
19468100 catatctaaaggtgtgtccaagtttggtaagagatggttctgcaagaataaagaagagtcatacaacctacgagacagaaatc
caatgcaaacacagatggagtttgggttgaagtggtgatgatactgtaaaaggataacctatggagatgctgtagctgattt
19467900 agagaaacagctcaaggacaccaaggaactggttttagggctcgggatttggcagctctttggtaagctactaaacagaacac
taaattctagatggttttcatttgggtctattattataggttaagattaatgggaggggagataagaatcaccgacaaggccat
19467700 gtttccaattcaatgttttatgacaacattagagtctcctccagtgagtgacatgaaagtgagacaggagatcgaagcagga
gccaacctcgggctgactataaacacttcaacttggaggttagcatgaaatatacgtaacctgagtcctagattcaacaactgct
```

Predikce funkce genů *in silico*

struktura genů

- struktura genů

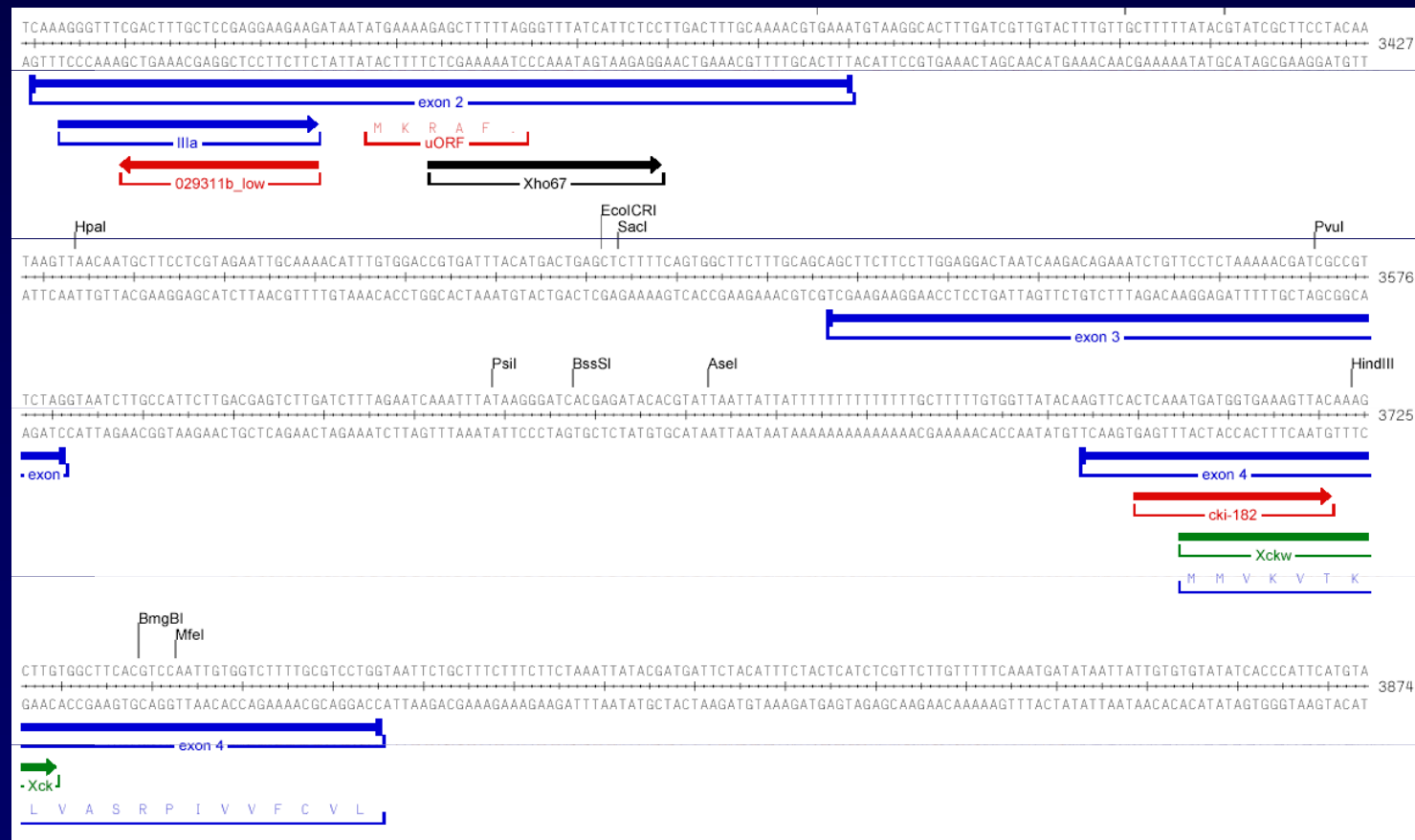
- promotor
- počátek transkripce
- 5'UTR
- počátek translace
- místa sestřihu
- stop kodon
- 3'UTR
- polyadenylační signál



Anotace genů a odhad aminokyselinových sekvencí předpokládaných proteinů

Postgenomová éra a co s informacemi, které neumíme číst

- dnes k dispozici v databázích více než 700 000 záznamů (nr databáze) různých organismů
- Získané informace lze zpracovat pomocí **bioinformatiky**



Základy proteomiky 2007

Proč právě proteomika?

- Postgenomová éra a co s informacemi, které neumíme číst
- Genotyp vs. fenotyp, aneb co všechno se děje při vyjadřování

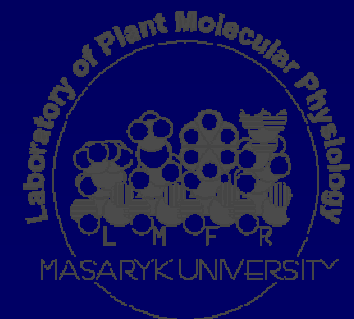


Genotyp vs. fenotyp, aneb co všechno se děje při vyjadřování

- Genom vs. Proteom



Danaus plexippus (monarch)



Genotyp vs. fenotyp, aneb co všechno se děje při vyjadřování

Možná analogie s textem a jeho interpretací

DNA:

Když adoperbtabi jsem dfjfwůcsaknclůsnínjxdalnxckjcnbychcxmasizdciksrdceasnanzxcnlsdlaň.

Když-----jsem-----snídal-----ní-----dal-----bych-----si-----srdce-----na-----dlaň.

RNA:

Když jsem s ní, dal bych si srdce na dlaň.

Když jsem s ní, dalabysissrdce naadtdlaň.

Když jsem snídal srdce.

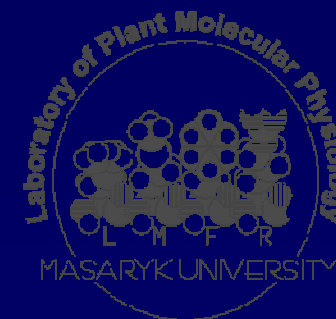
PROTEIN:



Základy proteomiky 2007

Proč právě proteomika?

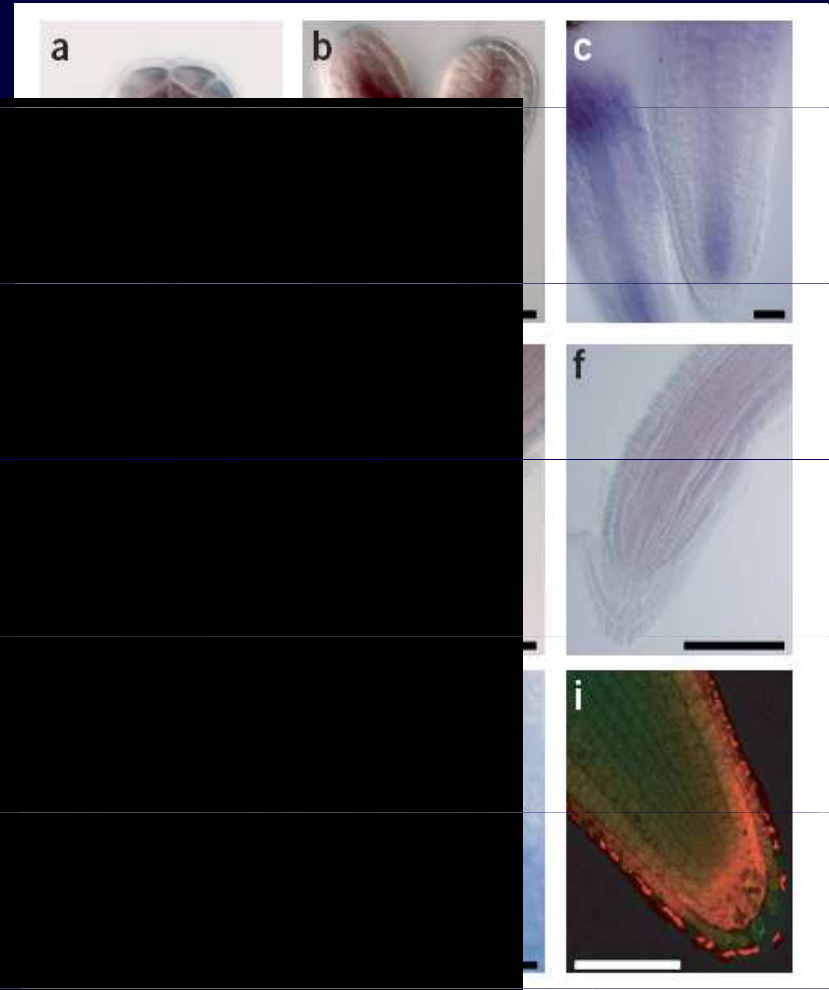
- Postgenomová éra a co s informacemi, které neumíme číst
- Genotyp vs. fenotyp, aneb co všechno se děje při vyjadřování
- Od genu k proteinu a zpět



Od genu k proteinu a zpět

Základní mechanismy regulace genové exprese

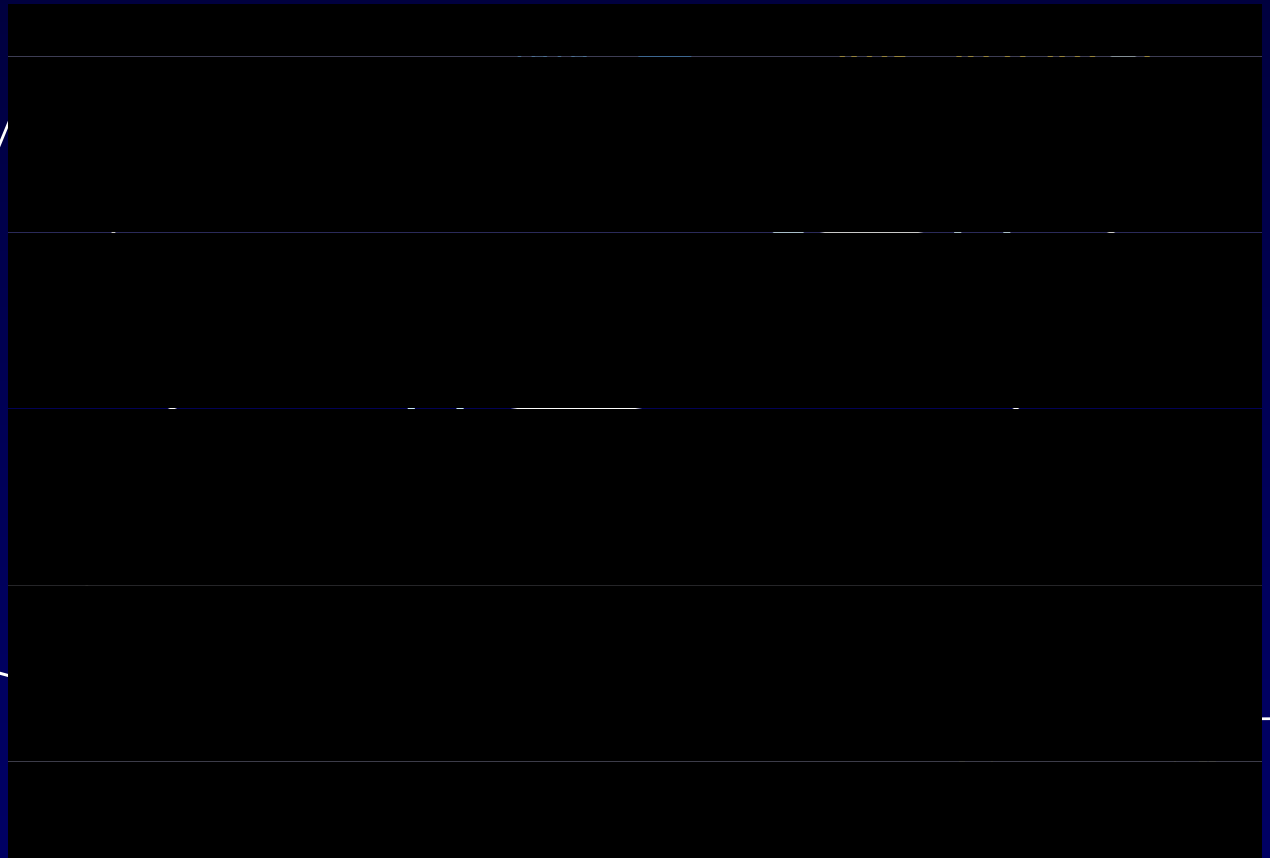
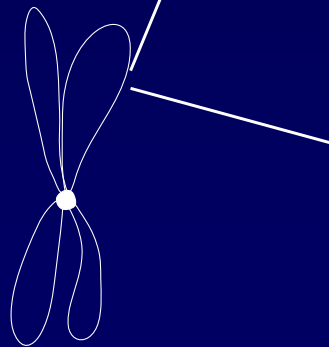
- regulace transkripce



Základní mechanismy regulace genové exprese

struktura genů

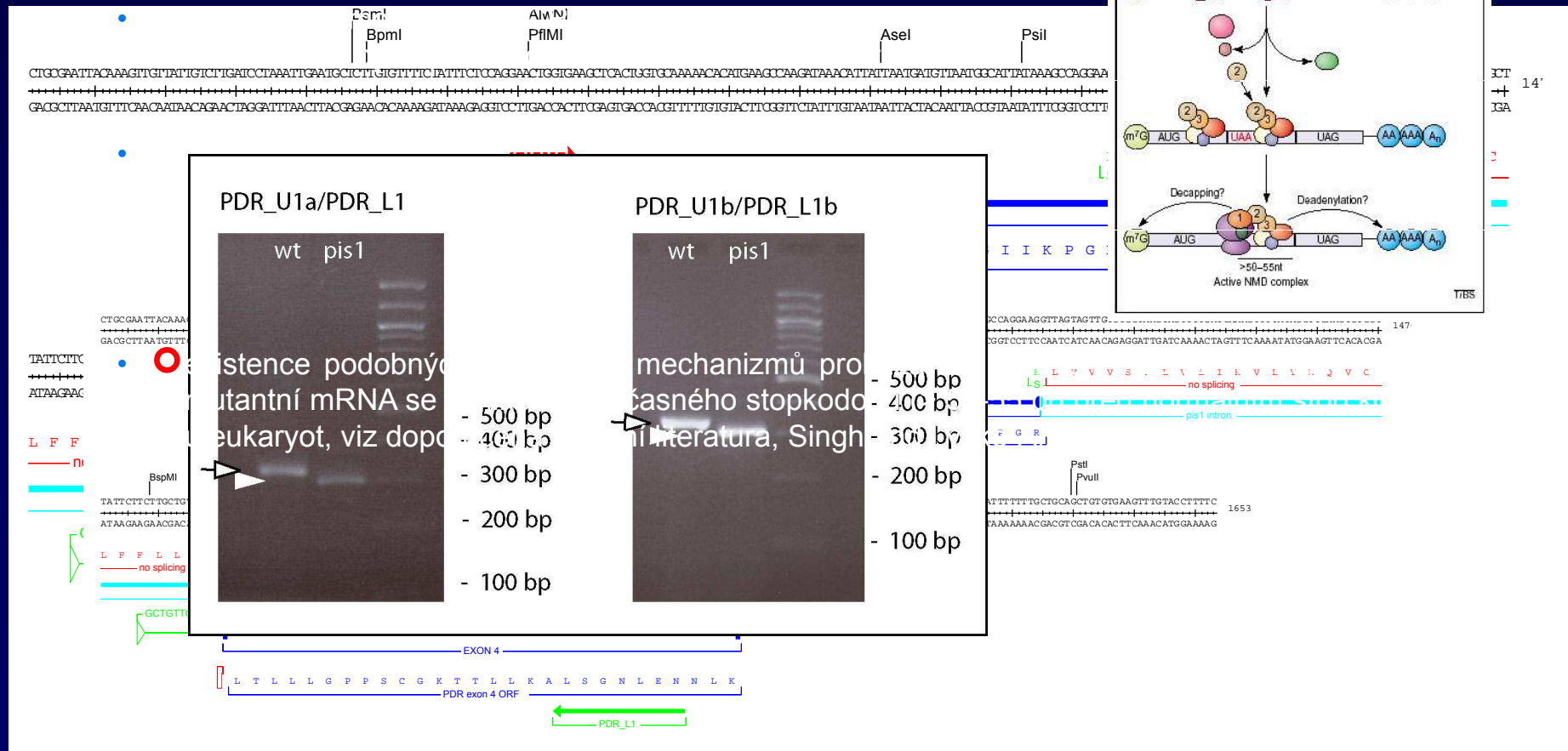
- struktura genů
 - promotor
 - počátek transkripce
 - 5'UTR
 - počátek translace
 - místa sestřihu
 - stop kodon
 - 3'UTR
 - polyadenylační signál



Základní mechanismy regulace genové exprese regulace transkripce

- odchyly rozpoznávání míst sestřihu u rostlin v praxi - příklad vývojové plasticity (nejen) rostlin

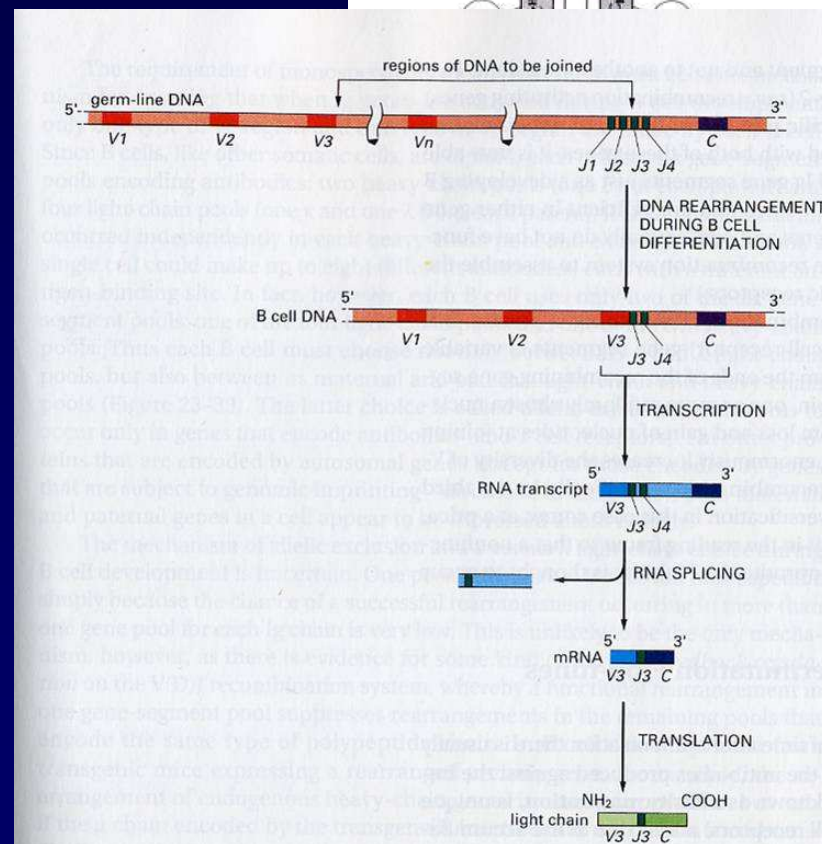
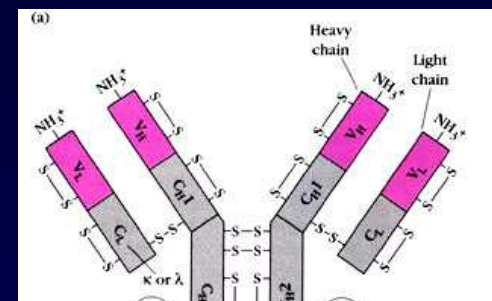
- identifikace mutanta s bodovou mutací (tranzice G→A) přesně v místě sestřihu na 5' konci 4. exonu



Základní mechanismy regulace genové exprese přeskupování subgenů při produkci protilátek

Přeskupování subgenů jako specifický mechanismus při produkci protilátek

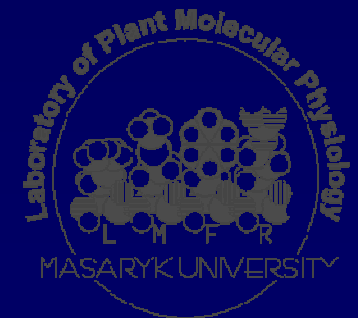
- protilátky variabilní oblast (V) a konstantní oblast (C) a lehký (L) a těžký (H) řetězec
 - každá z V oblastí L řetězce u myši je kódována 2 subgeny (V a J)
 - každá z V oblastí H řetězce u myši je kódována 3 subgeny (V, J a D)
- v zárodečných liniích myších B-lymfocytů dochází k tzv. **kombinatorické diversifikaci** (přeskupování) **subgenů** (místně-specifickou rekombinací)
 - L řetězec (κ): cca 300 V sub-genů a 4 J subgeny (**$300 \times 4 = 1200$** možností)
 - H řetězec: cca 500 V sub-genů, 4 J subgeny a 12 D subgenů (**$500 \times 4 \times 12 = 24000$** možností)
- celkové množství kombinací u myši: cca $1200 \times 24000 = 28$ mil. různých V oblastí (protilátek rozpoznávající různé antigeny)
- antigen indukuje tzv. **afinitní dozrání** mechanismem **somatické hypermutace**
 - po aktivaci B-lymf. pomocnými T-lymf. dochází ke zvýšenému výskytu mutací ve V oblastech (1 mutace/V oblast/generaci, cca 1 mil. X vyšší než je obvyklé (např. u tzv. „house-keeping“ genů) a selekci protilátek se zvýšenou afinitou k antigenu



Od genu k proteinu a zpět

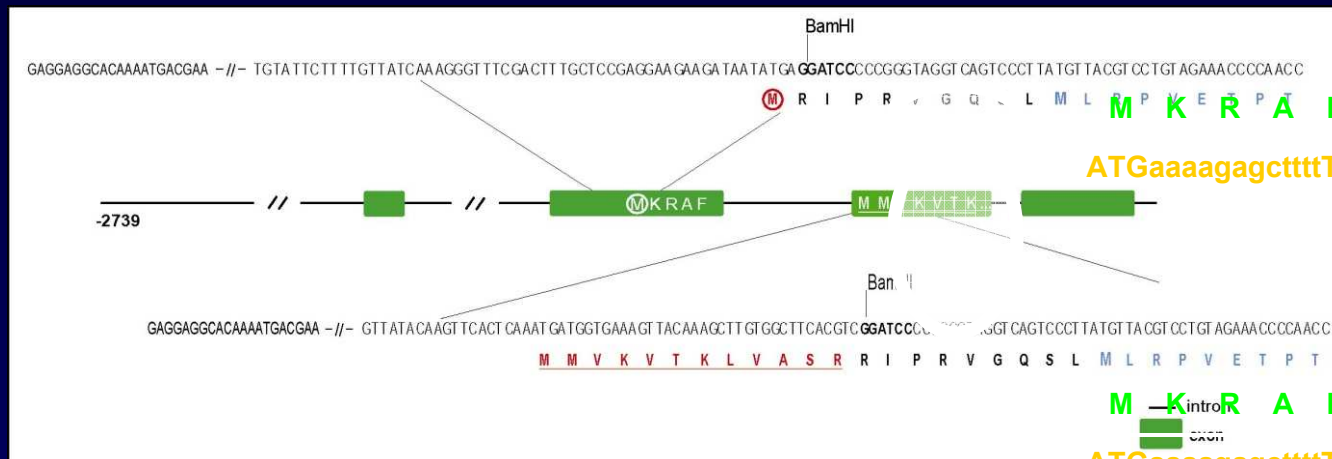
Základní mechanismy regulace genové exprese

- regulace transkripce
- sestřih RNA
- translační represe

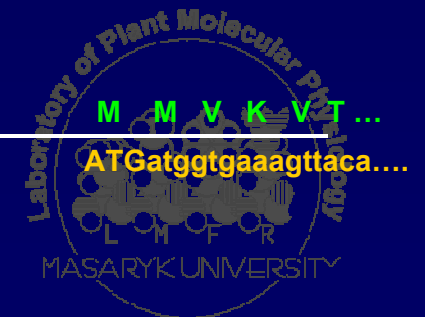
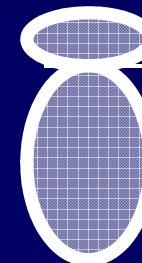
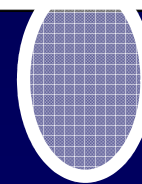


Regulace genové exprese mechanismem translační represe

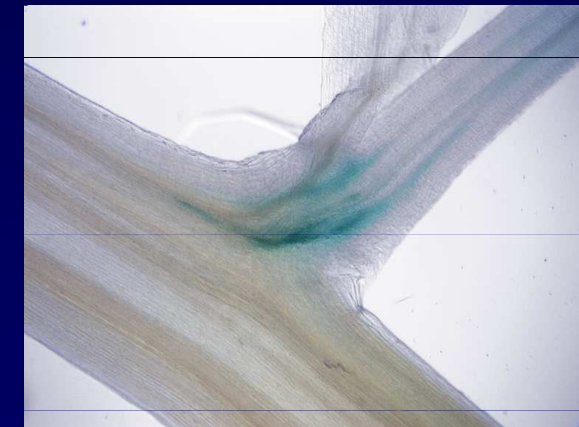
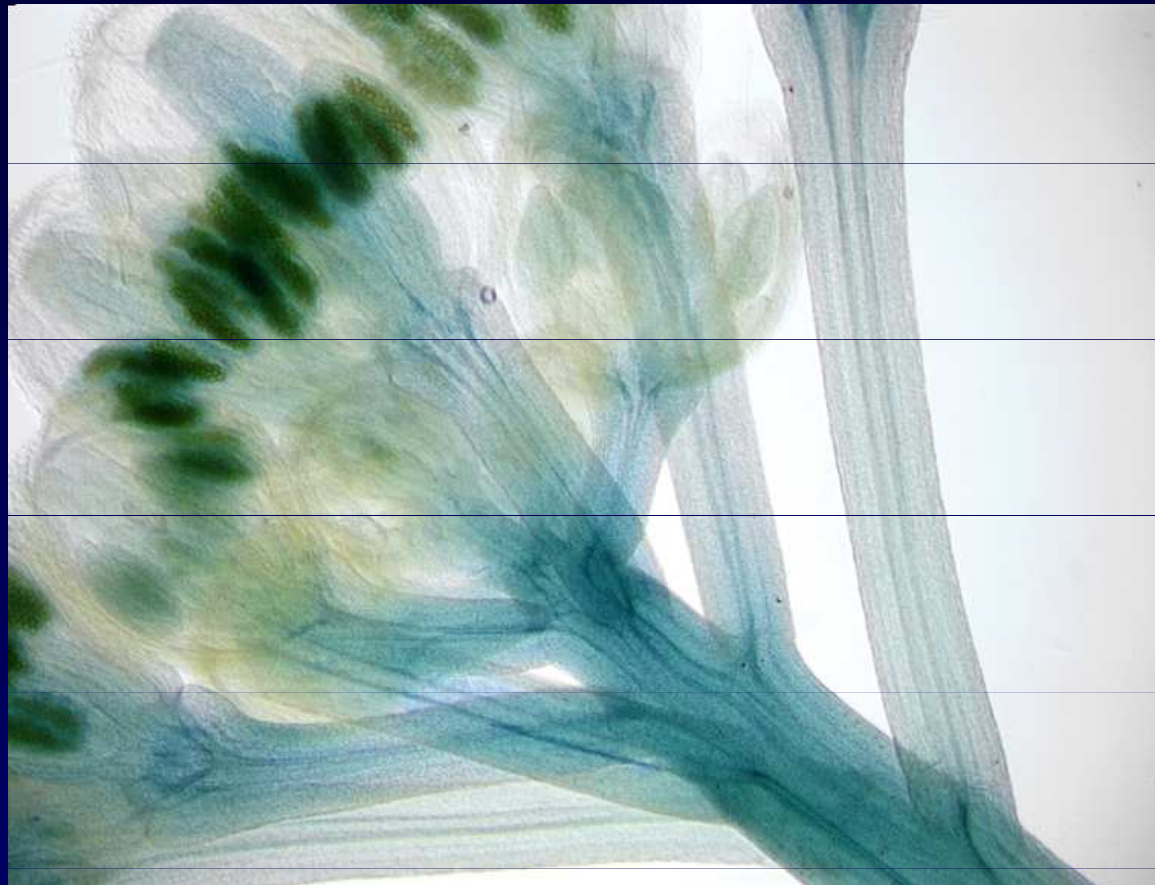
- Funkční význam sestřihu v nepřekládaných oblastech - důležitá regulační součást genů



- V případě CK11 pokus prokázat tento způsob regulace genové exprese pomocí transgenních linií nesoucích *uidA* pod kontrolou dvou verzí promotoru, zatím nepotvrzeno



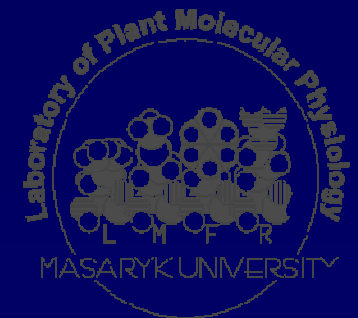
Expression of *CK11* in Diploid Generative Tissue Inflorescence



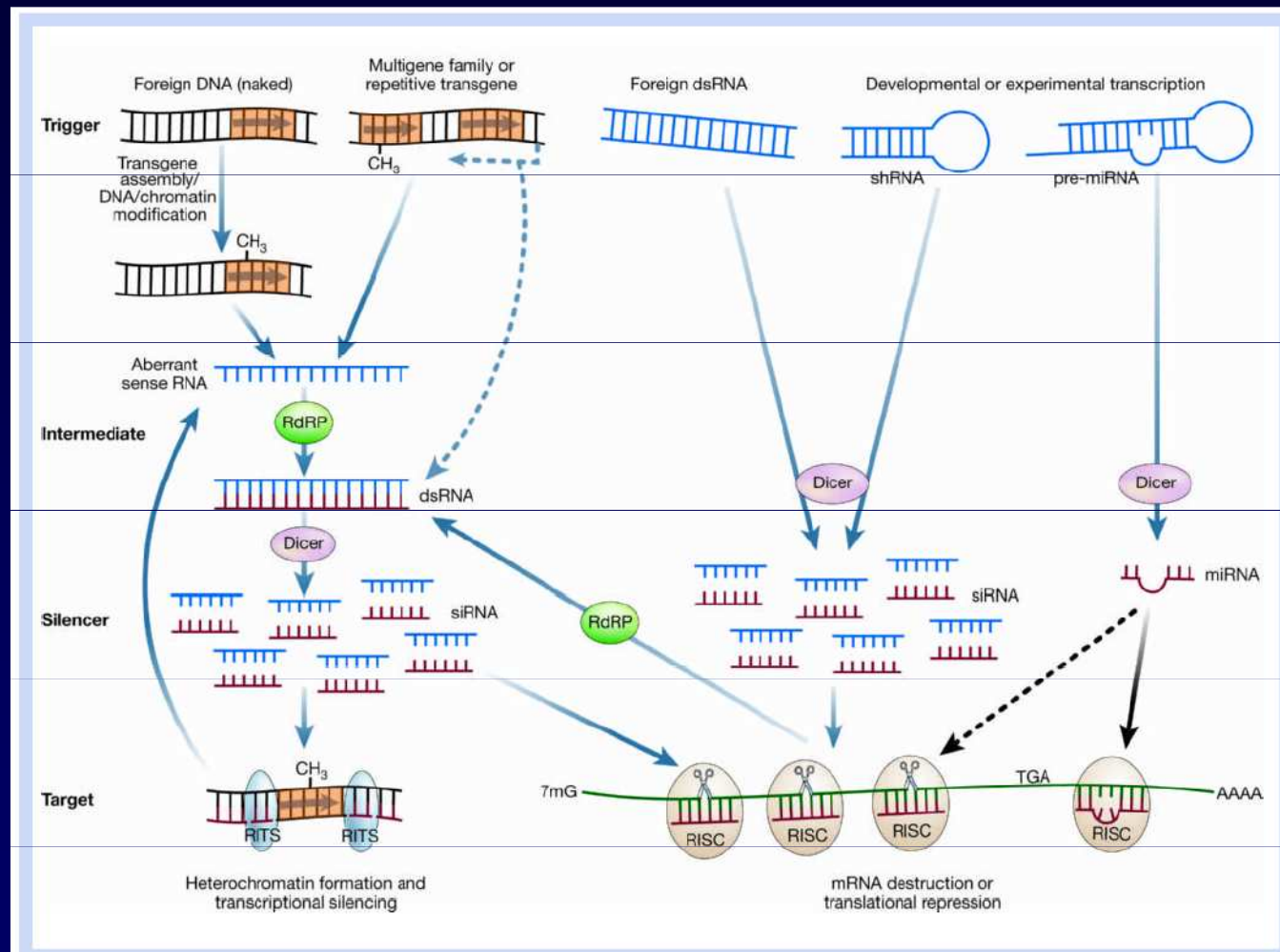
Od genu k proteinu a zpět

Základní mechanismy regulace genové exprese

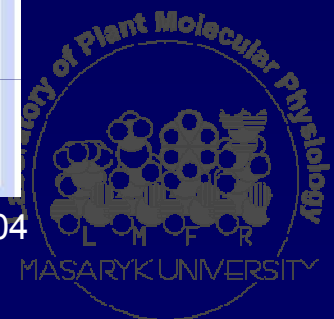
- regulace transkripce
- Sestřih RNA
- translační represe
- posttranskripční umlčování mechanismem siRNA



Od genu k proteinu a zpět umlčování genů mechanismem RNA interference (RNAi)



Mello, 2004



The Nobel Prize in Physiology or Medicine 2006

"for their discovery of RNA interference - gene silencing by double-stranded RNA"



Andrew Z. Fire

USA

Stanford University School of Medicine
Stanford, CA, USA

b. 1959

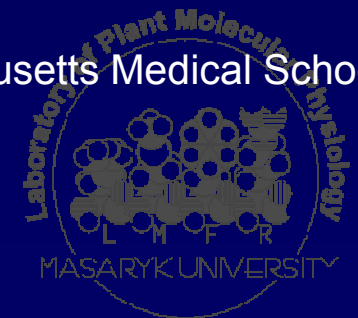


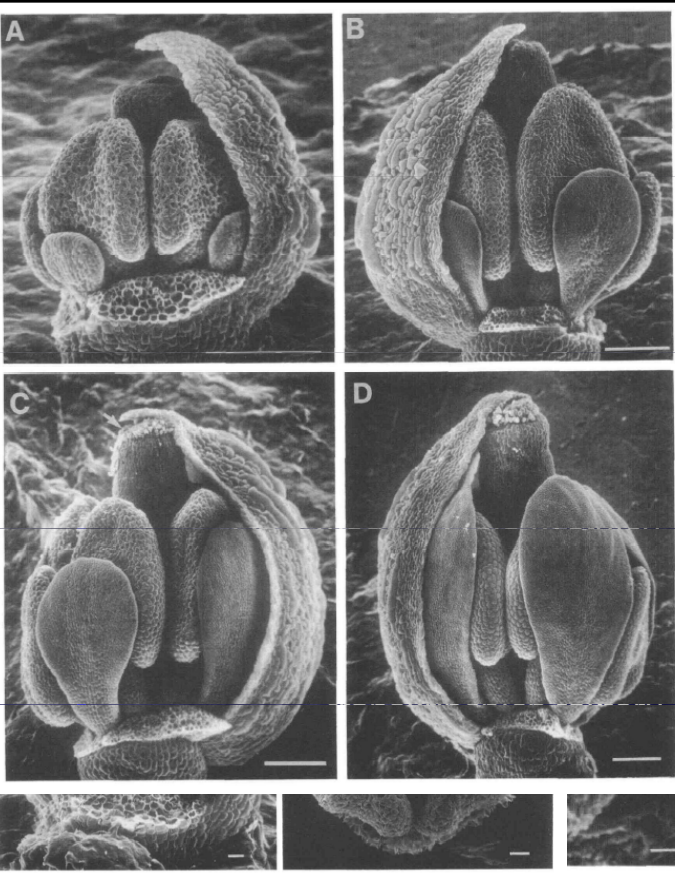
Craig C. Mello

USA

University of Massachusetts Medical School
Worcester, MA, USA

b. 1960





Od genu k proteinu a zpět

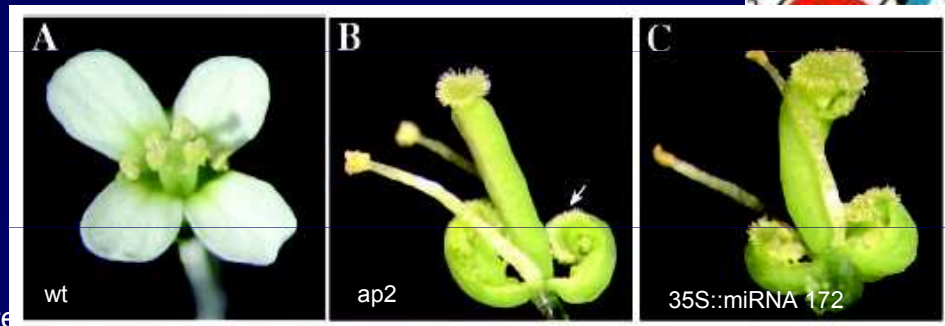
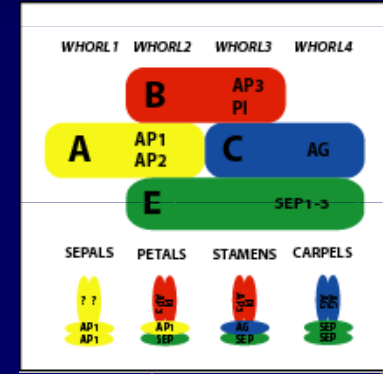
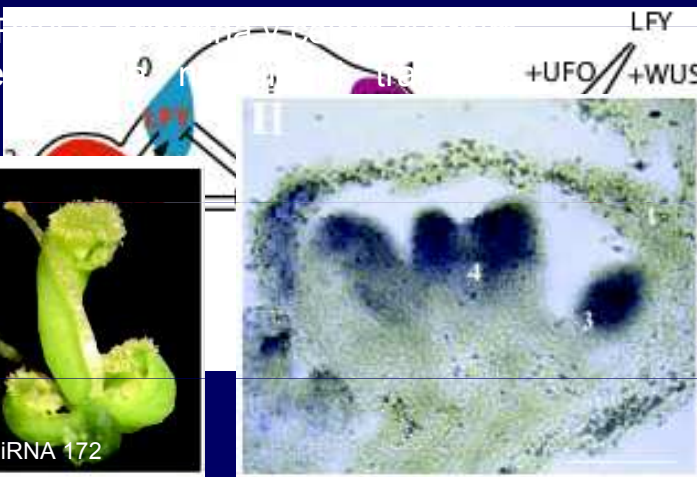
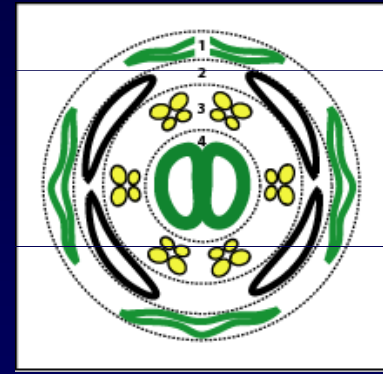
transkripční umlčování mechanismem siRNA

květů u *Arabidopsis* prostřednictvím miRNA

orgánů u rostlin
 orgánů dochází k určování identity jednotlivých květních orgánů
 homeotických genů
 dují většinou rostlinné homolgy MADS-box

kými geny dochází k tzv. katastrálním
 no genu inhibuje expresi dalšího

- např. *AP1* je nejprve aktivní v celém květním meristému, po indukcii exprese *AG* pak *AG* inhibuje expresi *AP1* ve vnitřních dvou kruzích)
- výjimkou je exprese genu *AP2*, jehož mRNA je v celém meristému, ale exprese *AP2* je reprimována prostřednictvím miRNA (gen miRNA 172)

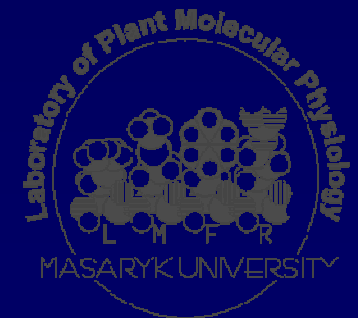


in situ lokalizace miRNA172 v 3. a 4. kruhu

Od genu k proteinu a zpět

Základní mechanismy regulace genové exprese

- regulace transkripce
- Sestřih RNA
- translační represe
- posttranskripční umlčování mechanismem siRNA
- směřování proteinů



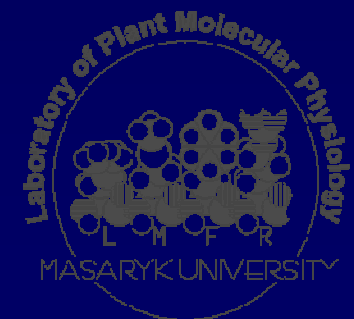
Od genu k proteinu a zpět směrování (cílování) proteinů

■ Intracelulární lokalizace proteinů

- v rostlinných buňkách dochází k velice dynamickým procesům, zprostředkovaným zejména tzv. endomembránovým transportem (viz film, GFP směrované do ER)

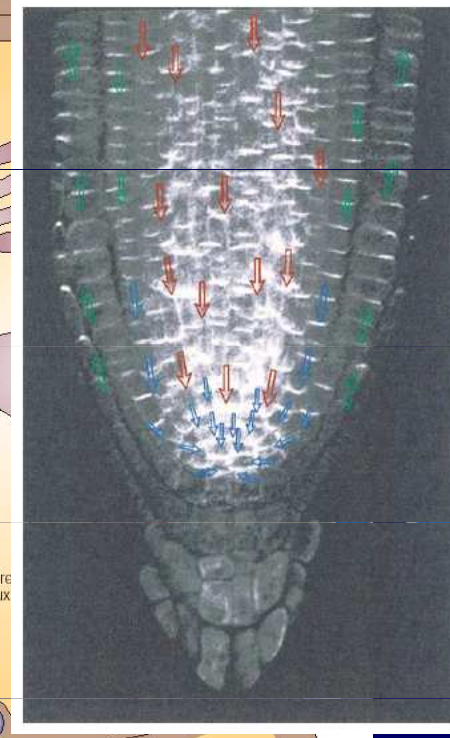
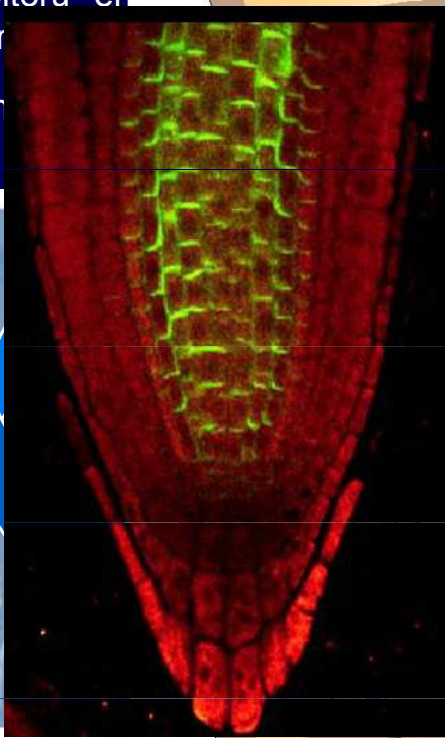


CV, central vacuole; DV, central vacuole; GA, Golgi apparatus; LV, large vacuole; PSV, protein storage vacuole; ER, endoplasmic reticulum; PSV, protein storage vacuole; SV, secretory vacuole



Od genu k proteinu a zpět směřování (cílování) proteinů

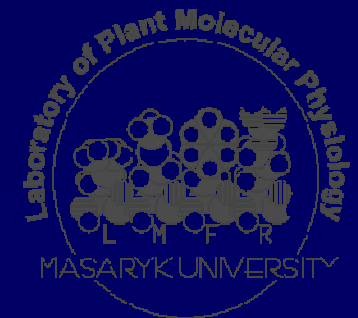
- **Cyklování auxinových přenašečů u *Arabidopsis***
 - auxin je rostlinný hormon se silným morfogenním účinkem
 - proteiny podílející se na transportu proteinů jsou tzv. PIN proteiny, polárně lokalizované v buněkách kořene u *Arabidopsis*
 - PIN proteiny cyklují v endomembránovém systému rostlinné buňky
 - v přítomnosti inhibitorů endocytózy těchto proteinů v intracelulárním prostoru
 - ...čímž je zároveň narušena polarita



Od genu k proteinu a zpět

Základní mechanismy regulace genové exprese

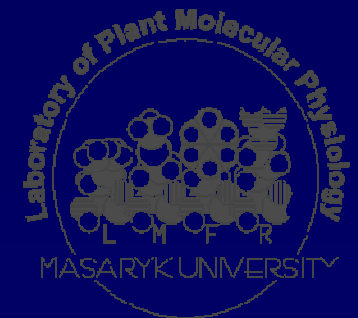
- regulace transkripce
- sestřih RNA
- translační represe
- posttranskripční umlčování mechanismem siRNA
- směřování proteinů
- posttranslační modifikace proteinů



Od genu k proteinu a zpět postranslační modifikace proteinů

Význam postranslačních modifikací proteinů

- regulace enzymové aktivity
- regulace interakcí proteinu s dalšími proteiny nebo jinými biomolekulami
- lokalizace proteinu v buňce
- změna mechanických vlastností proteinu
- přenos signálu

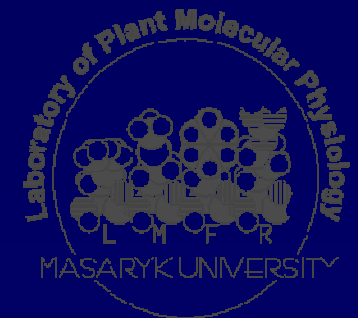


Od genu k proteinu a zpět

postranslační modifikace proteinů

Typy posttranslačních modifikací proteinů

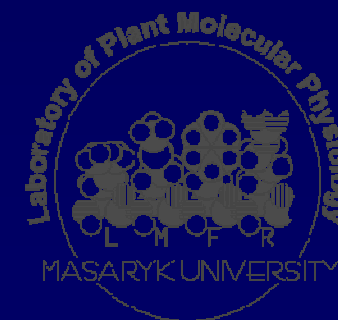
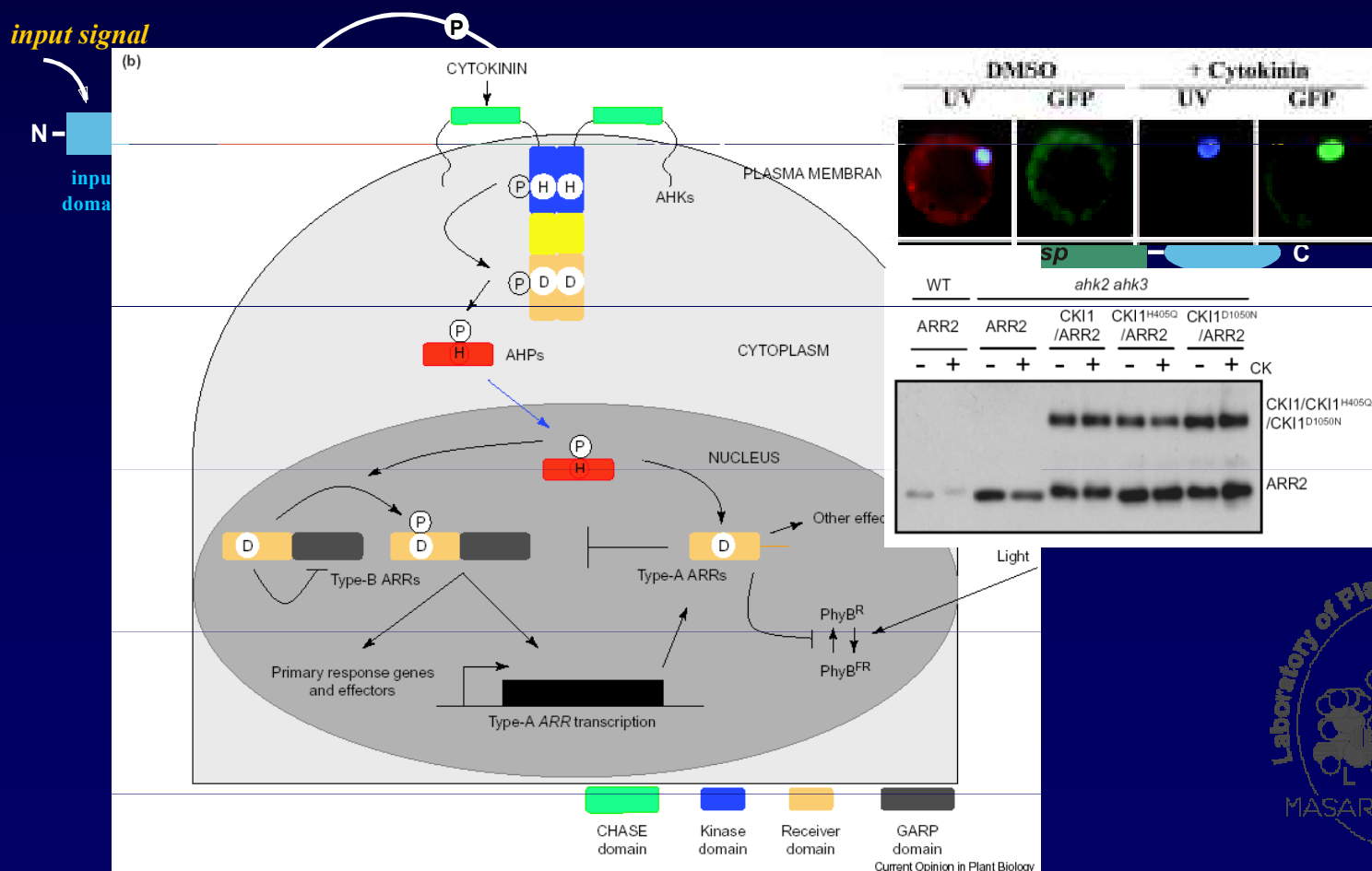
- přidání glykosylfosfatidylinositolové (GPI) kotvy
- fosforylace
- sulfonace
- glykosylace
- N-myristolyace
- N-metylace
- hydroxylace
- karboxylace
- prenylace
-



Od genu k proteinu a zpět postranlační modifikace proteinů

Přenos signálu a regulace genové exprese prostřednictvím fosforylace

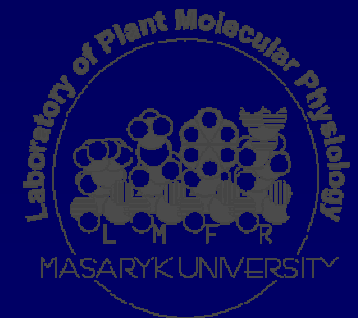
- přenos cytokininového signálu u rostlin



Od genu k proteinu a zpět postranlační modifikace proteinů

Přenos signálu a regulace genové exprese prostřednictvím fosforylace

- přenos signálu prostřednictvím TGF β (Transforming Growth Factor) u živočichů



Základy proteomiky 2007

Proč právě proteomika?

- Postgenomová éra a co s informacemi, které neumíme číst
- Genotyp vs. fenotyp, aneb co všechno se děje při vyjadřování
- Od genu k proteinu a zpět
- Přístupy současné proteomiky
 - exprese, purifikace a analýza rekombinantních proteinů

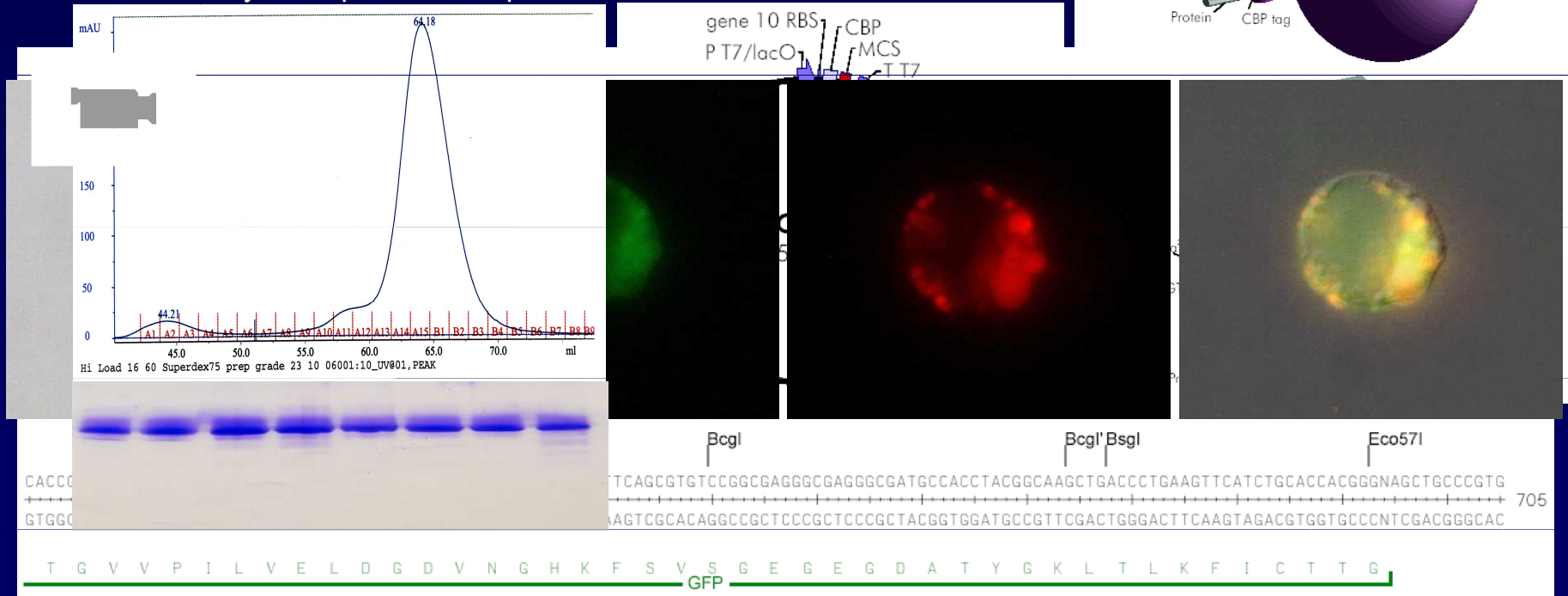
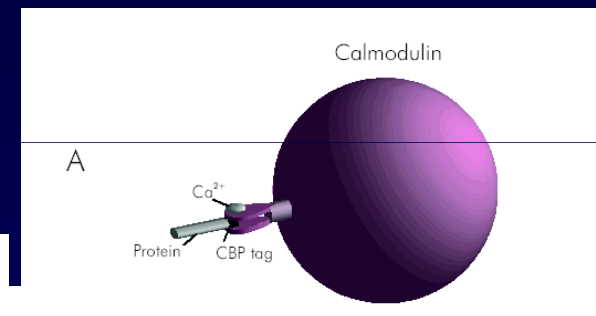


Přístupy současné proteomiky

exprese, purifikace a analýza rekombinantních proteinů

Technologie rekombinantních proteinů

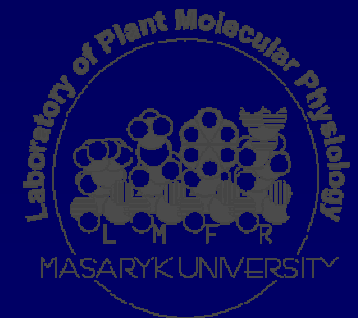
- umožňuje získat velké množství analyzovaného proteinu ve velké čistotě
- využívá technologie rekombinantní DNA
- principem je vložení „přívěsku“ prostřednictvím přípravy rekombinantní DNA, který usnadní purifikaci (afinitní purifikace)
- možnost využití „přívěsku“ i pro lokalizaci a další



Základy proteomiky 2007

Proč právě proteomika?

- Postgenomová éra a co s informacemi, které neumíme číst
- Genotyp vs. fenotyp, aneb co všechno se děje při vyjadřování
- Od genu k proteinu a zpět
- Přístupy současné proteomiky
 - exprese, purifikace a analýza rekombinantních proteinů
 - analýza vztahu mezi strukturou a funkcí proteinů

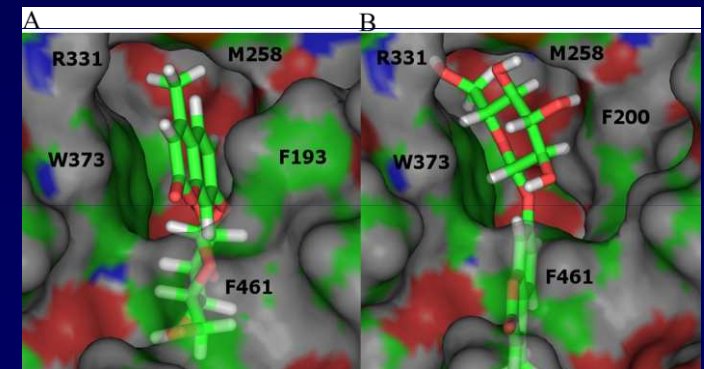
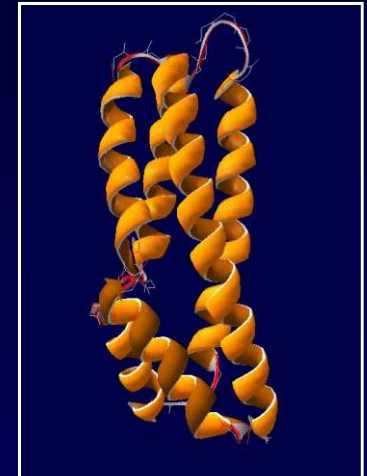


Přístupy současné proteomiky

analýza vztahu mezi strukturou a funkcí proteinů

Analýza vztahu mezi funkcí a strukturou proteinu

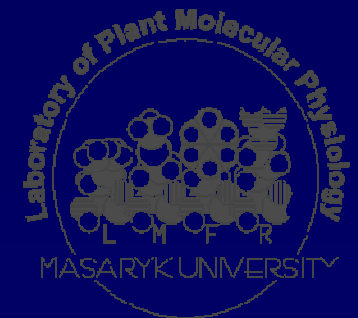
- využívá technologie produkce rekombinantních proteinů a místně řízené mutagenese
- umožňuje analyzovat strukturu rekombinantního proteinu pomocí rentgenové krystalografie nebo NMR
- komparativní analýzou lze pak analyzovat strukturu a funkci jak u standardního typu tak i mutantního proteinu



Základy proteomiky 2007

Proč právě proteomika?

- Postgenomová éra a co s informacemi, které neumíme číst
- Genotyp vs. fenotyp, aneb co všechno se děje při vyjadřování
- Od genu k proteinu a zpět
- Přístupy současné proteomiky
 - exprese, purifikace a analýza rekombinantních proteinů
 - analýza vztahu mezi strukturou a funkcí proteinů
 - diferenční proteomika

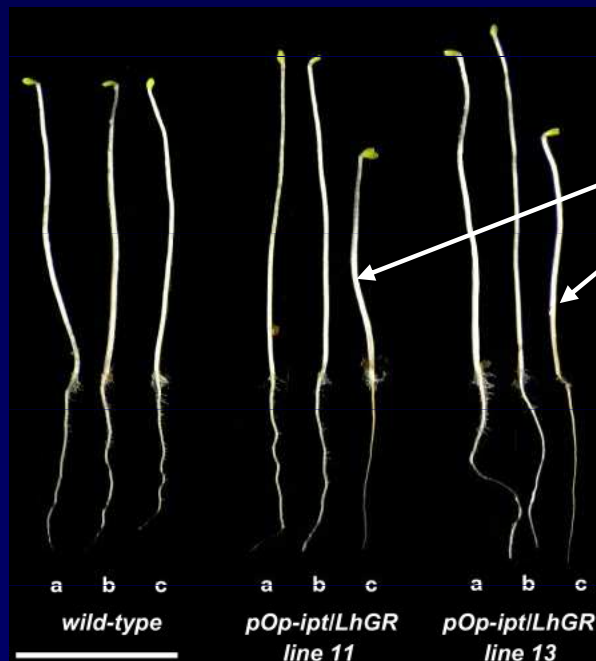


Přístupy současné proteomiky

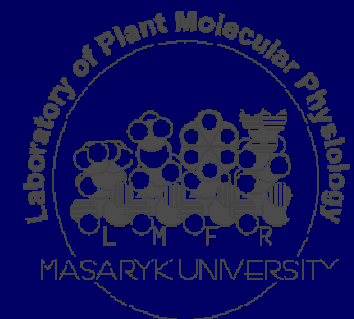
komparativní proteomika

Analýza změn proteomu organismu na různé stimuly

- využívá technologií vícerozměrné separace proteinů (nejčastěji ELFO nebo LC)
- umožňuje analyzovat mnoho proteinů najednou
- pomocí analýzy obrazu lze odhadnout proteiny se změnou vlastností, příp. kvantitativně za daných sledovaných podmínek (mutace, vývojové stadium, fyziologická odpověď)

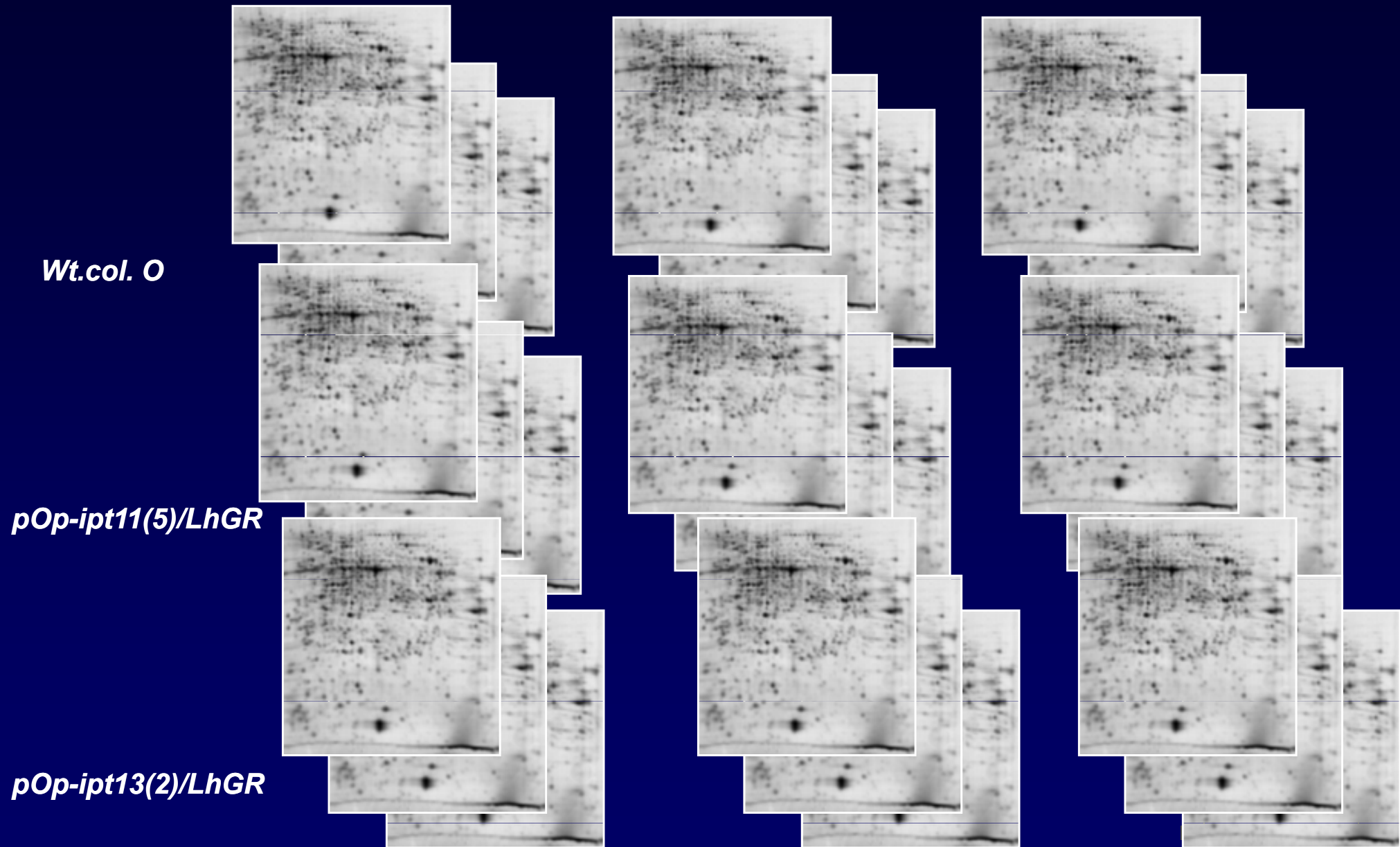


Zvýšená hladina endogenních CK



Přístupy současné proteomiky

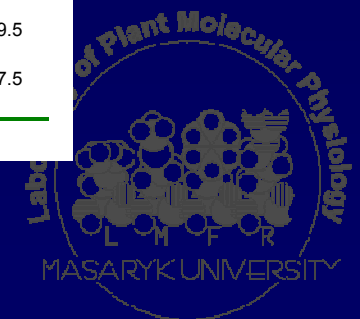
komparativní proteomika



Přístupy současné proteomiky

komparativní proteomika

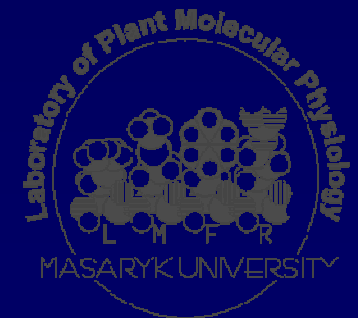
Spot	% Vol	Relative Abundance	Accession No.	Protein Name	MALDI-MS		LC-MSMS		MM [kDa]	pI
					Score	%Cov	Score	%Cov		
1105		3.55	S47970	14-3-3 protein homolog RC12	-	-	123	20	28.1	5.0
2203		< 0.01	BAB01146	AP000373 NID (jasmonate inducible, myrosinase-like)	97	725	-	-	48.5	5.0
2310		< 0.01	BAA84393	AP000423 NID (RuBisCO large subunit)	116	26	-	-	52.9	5.9
2402		3.40	Q8LBJ7	50S ribosomal protein L12-C	-	-	130	16	19.7	6.0
2801		2.50	AAG51430	AC008153 NID (2-cys peroxiredoxin BAS1, chloroplast [Precursor])	87	29	92	11	29.1	6.9
4503		> 100	Q9SUI9	ATP synthase gamma chain, chloroplast	131	29	393	32	33.3	5.1
5204		< 0.01	E71425	Hypothetical protein (putative epoxide hydrolase)	-	-	40	5	33.4	6.2
7009		< 0.01	T05413	cinnamyl-alcohol dehydrogenase (EC 1.1.1.195) F28A23.10	-	-	86	7	38.7	5.4
7401		< 0.01	Q9LJE4	GloEL protein, chaperonin, 60 kDa	-	-	217	10	63.3	5.6
7606		> 100	F86262	F13K23.15 protein (glyceraldehyde-3-phosphate dh)	81	23	388	33	42.8	8.2
7701		0.37	AAM20572	AY099721 NID (Cysteine synthase, mitochondrial [Precursor])	128	26	236	28	45.8	8.4
8202		> 100	AAL36245	AY063889 NID (50S ribosomal protein L4, chloroplast [Precursor])	-	-	59	4	30.5	8.9
8302		2.69	H84808	probable annexin	-	-	90	9	36.2	6.9
8605		> 100	Q941D3	At5g19940/F28I16_90 (Hypothetical protein)	-	-	142	10	26.5	9.5
8803		0.42	T47470	isovaleryl-CoA dehydrogenase	63	11	213	18	44.7	7.5



Základy proteomiky 2007

Proč právě proteomika?

- Postgenomová éra a co s informacemi, které neumíme číst
- Genotyp vs. fenotyp, aneb co všechno se děje při vyjadřování
- Od genu k proteinu a zpět
- Přístupy současné proteomiky
 - exprese, purifikace a analýza rekombinantních proteinů
 - analýza vztahu mezi strukturou a funkcí proteinů
 - diferenční proteomika
 - analýza posttranslačních modifikací

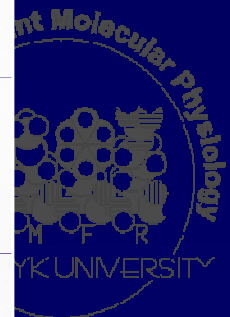


Přístupy současné proteomiky

analýza posttranslačních modifikací

Analýza posttranslačních modifikací

- pomocí specifických metod lze identifikovat kotranslační a posttranslační modifikace, buď v gelu nebo po blotování na membránu (barvení spec. barvičkami)
- identifikace modifikací pomocí MS technik (MALDI TOF, ESI-MS, ...)
- fosforylace
 - přenos signálu
- acetylace
 - regulace chromatinových struktur a transkripční aktivity prostřednictvím regulace vazby histonů
- glykosylace
 - velice heterogenní (aktivátor plasminogenu 3 místa pro glykosylaci na N-konci, až 11. 520 možností izoforem)

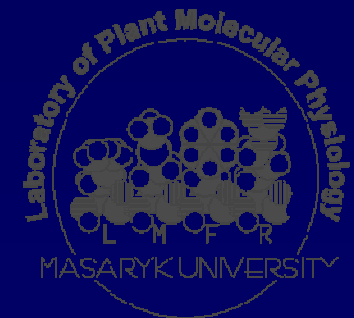


Základy proteomiky 2007

shrnutí

Proč právě proteomika?

- Postgenomová éra a co s informacemi, které neumíme číst
- Genotyp vs. fenotyp, aneb co všechno se děje při vyjadřování
- Od genu k proteinu a zpět
- Přístupy současné proteomiky
 - exprese, purifikace a analýza rekombinantních proteinů
 - analýza vztahu mezi strukturou a funkcí proteinů
 - diferenční proteomika
 - analýza posttranslačních modifikací



Základy proteomiky 2007

zdrojová literatura

■ Zdrojová literatura k první přednášce:

Monografie a učebnice

- Plant Functional Genomics, ed. Erich Grotewold, 2003, Humana Press, Totowa, New Jersey
- Proteome Research: New Frontiers in Functional Genomics, ed. Wilkins, M.R., Williams, K.L., Appel, R.D., Hochstrasser, D.F., 1997, Springer-Verlag, Berlin, Heidelberg
- Dubová J., Hejátko J., Friml J. (2005) Reproduction of Plants, in Encyclopedia of Molecular Cell Biology and Molecular Medicine (ed, R. A. Meyers), pp. 249 – 295. Wiley-VCH, Weinheim, Germany

Publikace v mezinárodních časopisech

- Wang, L. and Wessler, S.R. (1998) Inefficient reinitiation is responsible for upstream open reading frame-mediated translational repression of the maize R gene. *Plant Cell*, **10**, (1733)
- Friml, J. and Palme, K. (2002) Polar auxin transport. Old questions and new concepts?. *Plant Mol. Biol.*, **49**, 273-284
- Mello, C.C. and Conte Jr., D. (2004) Revealing the world of RNA interference. *Nature*, **431**, 338-342
- Surpin, M. and Raikhel, N. (2004) Traffic jams affect plant development and signal transduction. *Nature Reviews/Molecular Cell Biology* **5**, 100-109

