

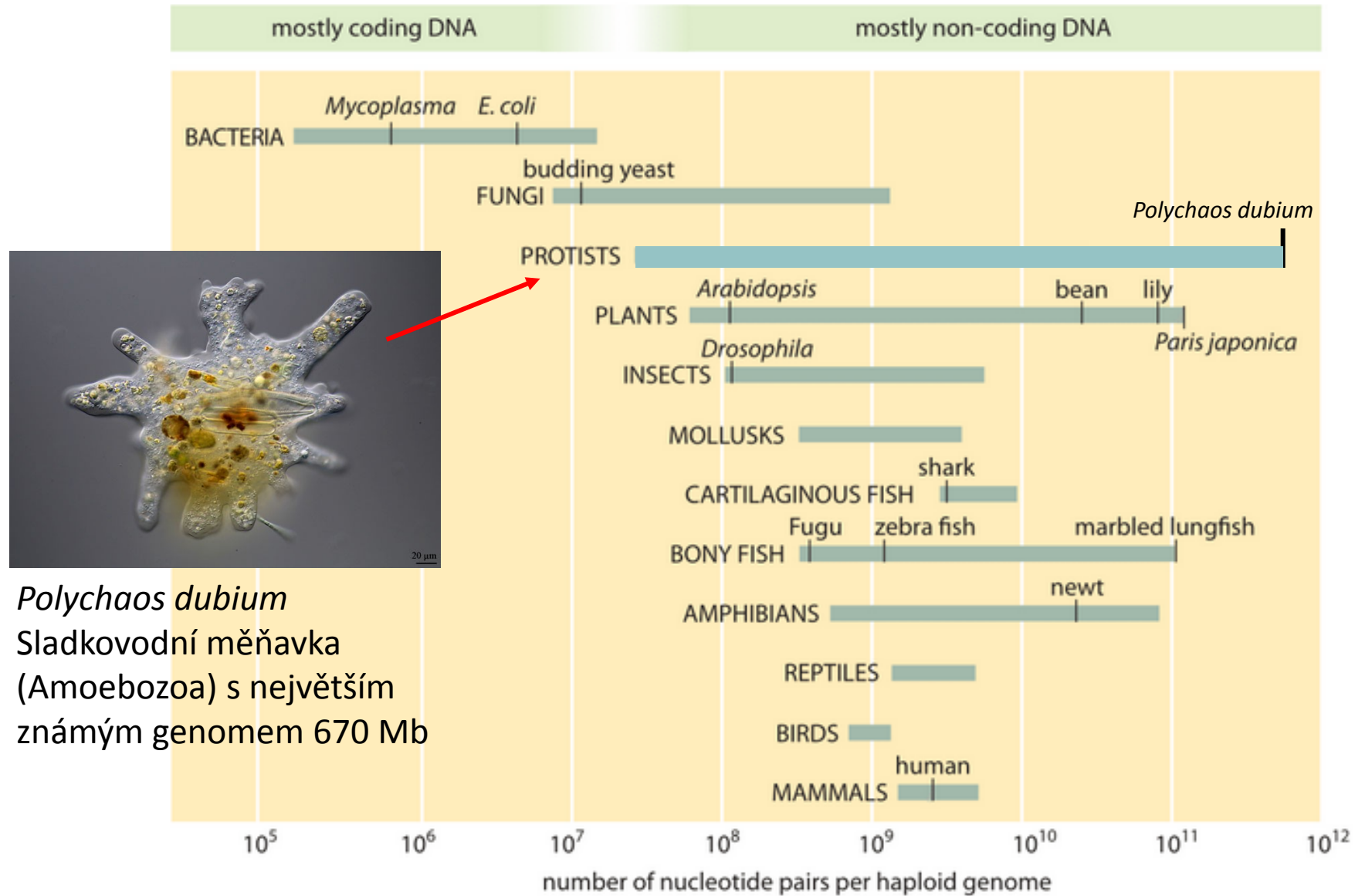


Eduard Kejnovský + Zdeněk Kubát  
+ Roman Hobza

EVOLUČNÍ GENOMIKA:

V. DYNAMIKA GENOMŮ

# Paradox hodnoty C



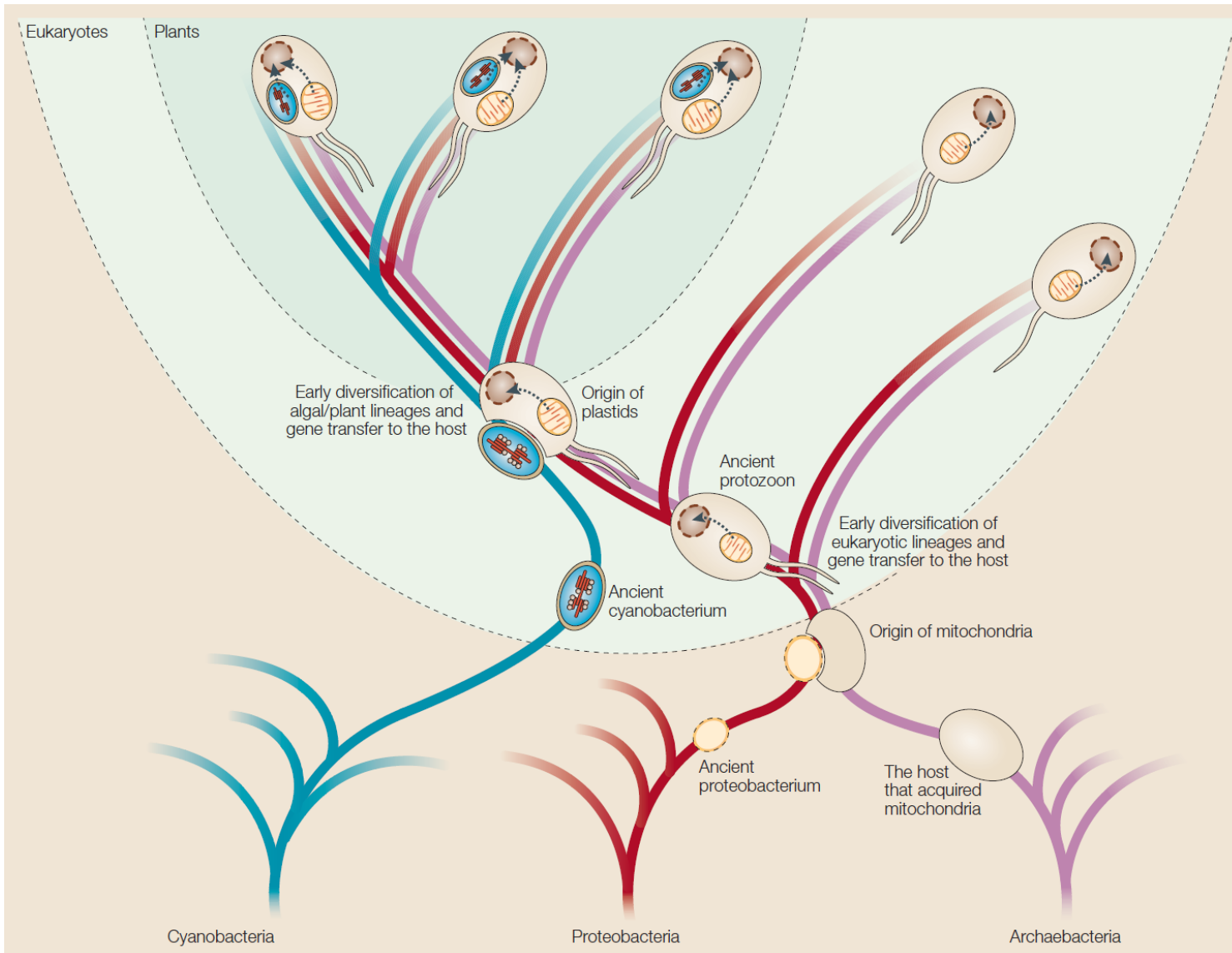
Velikost genomu (hodnota C) není závislá na komplexitě organismu.

# Dynamické složky genomu

1. Promiskuitní DNA
2. Tandemové repetice
3. Transpozony

# **1. PROMISKUITNÍ DNA**

# Endosymbiotická evoluce a strom života



## ENDOSYMBIOTIC GENE TRANSFER: ORGANELLE GENOMES FORGE EUKARYOTIC CHROMOSOMES

*Jeremy N. Timmis<sup>\*</sup>, Michael A. Ayliffe<sup>‡</sup>, Chun Y. Huang<sup>\*</sup> and William Martin<sup>§</sup>*

Genome sequences reveal that a deluge of DNA from organelles has constantly been bombarding the nucleus since the origin of organelles. Recent experiments have shown that DNA is transferred from organelles to the nucleus at frequencies that were previously unimaginable. Endosymbiotic gene transfer is a ubiquitous, continuing and natural process that pervades nuclear DNA dynamics. This relentless influx of organelle DNA has abolished organelle autonomy and increased nuclear complexity.

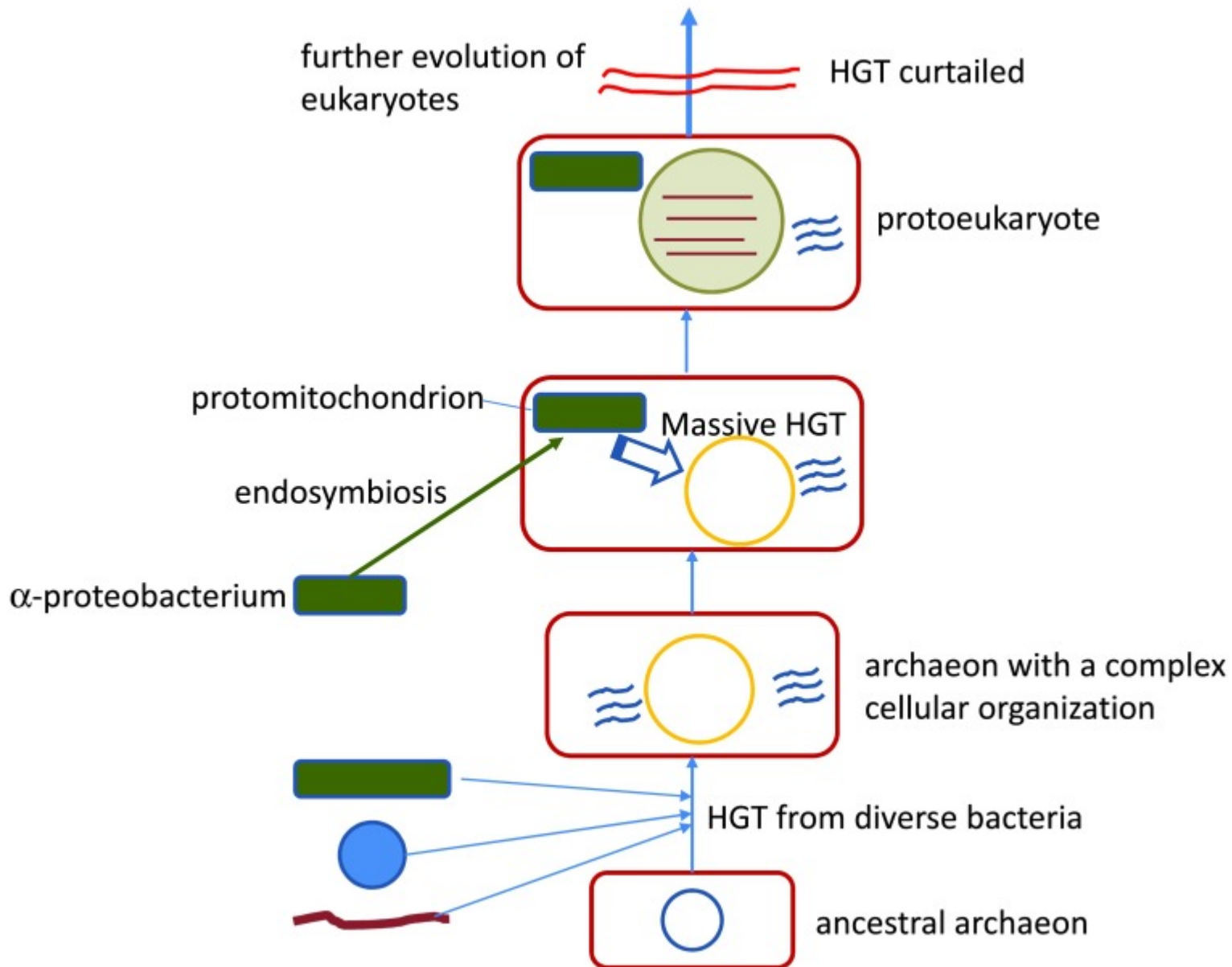
„Endosymbiotic origin of organelles“ (Lynn Sagan (Margulis), 1967)

“Promiscuous DNA” (Ellis, 1982) – cp DNA v mt genomu

“Endosymbiotic gene transfer is ubiquitous...  
... at frequencies that were previously unimaginable”.

(Timmis et al., 2004)

# Vznik Eukaryotické buňky (mitochondrie)



(Koonin, 2016)

**HGT** je velice rozšířený jev u Eubakterií i Archeí, je považován za dominantní proces v evoluci prokaryot, hovoří se o **Pan-genomu** prokaryotických druhů.

-> ancestrální Archaeon obsahoval mnoho bakteriálních genů a patrně měl již znaky Eukaryotické buňky. Podle jiných teorií tyto znaky získal až po pohlcení endosymbionta.

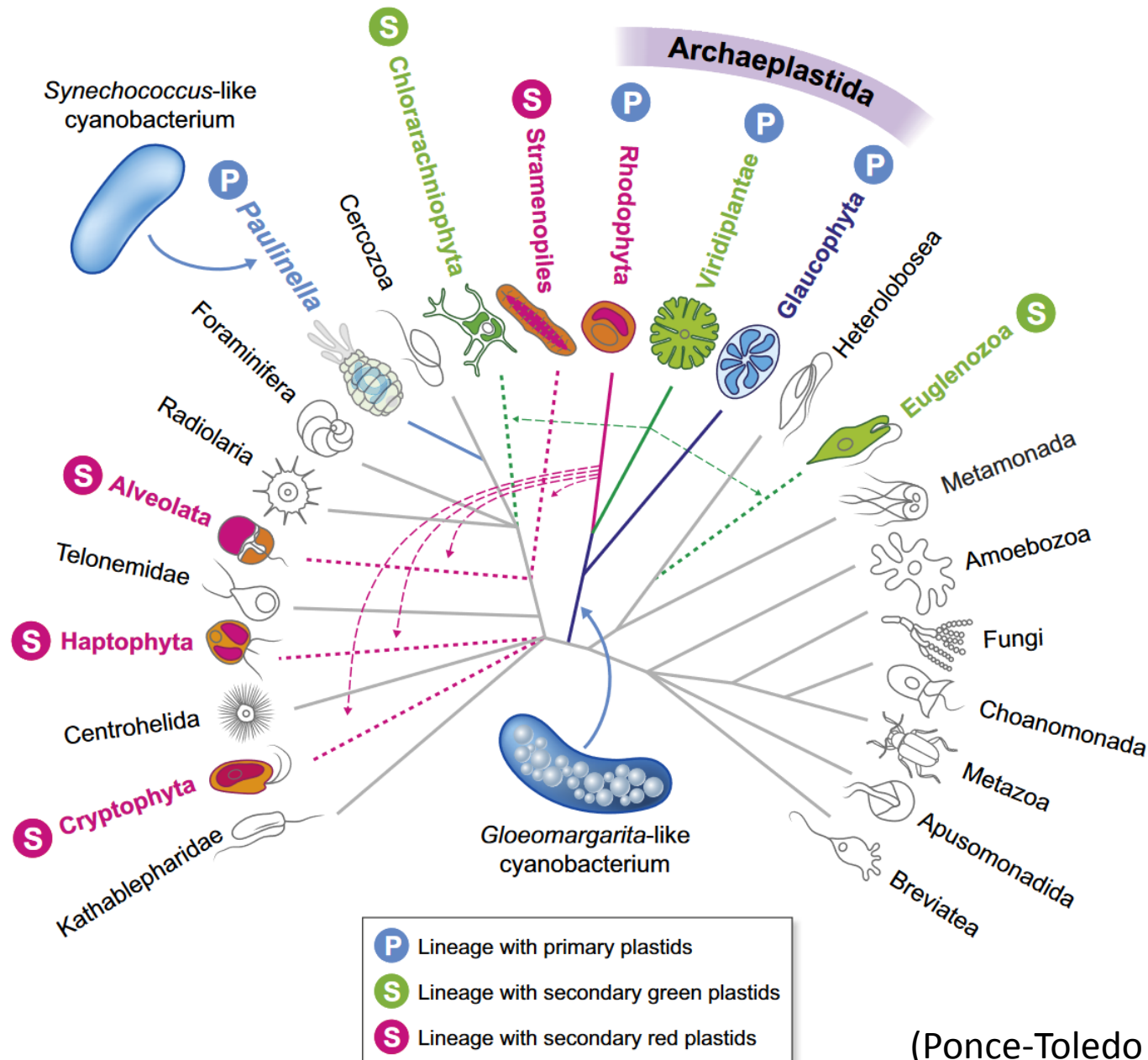
HGT přispěl k evolučnímu přechodu (evolutionary transition) od prokaryotické k eukaryotické buňce.

Masivní tok genů z endosymbionta (protomitochondrie) do jádra.

V současnosti je (EGT) tok genů utlumen.

EGT = Endosymbiotic Gene Transfer  
HGT = Horizontal Gene Transfer  
LGT = Lateral Gene Transfer

# Endosymbiotický vznik plastidů a strom života



## Primární endosymbióza pouze 2x:

- Archeplastida (endosymbiont = *Gloeomargarita*)
- *Paulinella* (endosymbiont = *Synechococcus*)

## Sekundární a terciární endosymbióza mnohokrát:

- Euglenozoa, Chlorarachniophyta (endosymbiont = Viridiplantae)
- Cryptophyta, Haptophyta, Alveolata, Stramenopiles (endosymbiont = Rhodophyta)

## Proč?

- Bariéry proti adopci primárního endosymbionta
- Sekundární endosymbionti vznikají mnohem častěji – odstranění bariér díky toku genů a převzetí mnoha funkcí jádrem

(Ponce-Toledo et al., 2019)

# Primární endosymbióza – charakteristiky plastidů

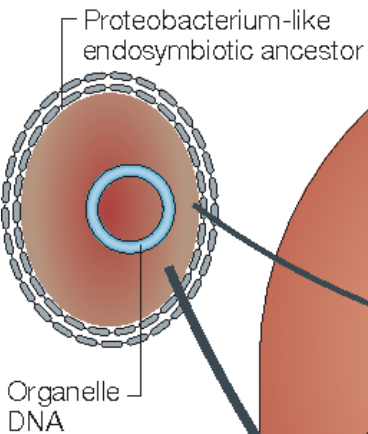
Characteristic	Archaeplastida	<i>Paulinella</i>
Plastid genome size	100–200 kbp	1021 kbp
Number of plastid genes	80–250	911
EGTs in plastid proteome	70–390	> 70
Noncyanobacterial prokaryotic proteins in plastid proteome	40–240	> 170
Host-derived and proteins of uncertain origin in plastid proteome	320–900	> 390
Lineage age estimation	> 1000 Ma	90–140 Ma
Import system of nucleus-encoded proteins into the plastid	Translocons at the outer and inner plastid membranes (TOC/TIC complex)	Vesicles of the host endomembrane system fuse with the outer plastid membrane and proteins cross the inner membrane through a simplified TIC translocon
Phagotrophic capacity	Lost in Rhodophyta, Glaucophyta and most Viridiplantae but preserved in some prasinophytes	Lost
Peptidoglycan wall	Present in Glaucophyta and some Viridiplantae species	Present



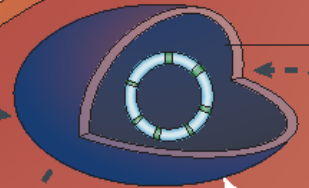
# Promiskuitní DNA

## Endosymbióza a tok genů do jádra a mitochondrií

**Proteobakterie**  
 ~ 4000 kb  
 ~ 4000 genů



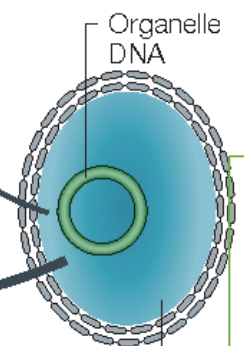
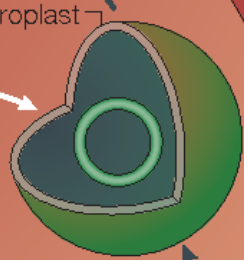
**Mitochondrie**  
 ~ 10-400 kb  
 ~ 60 genů



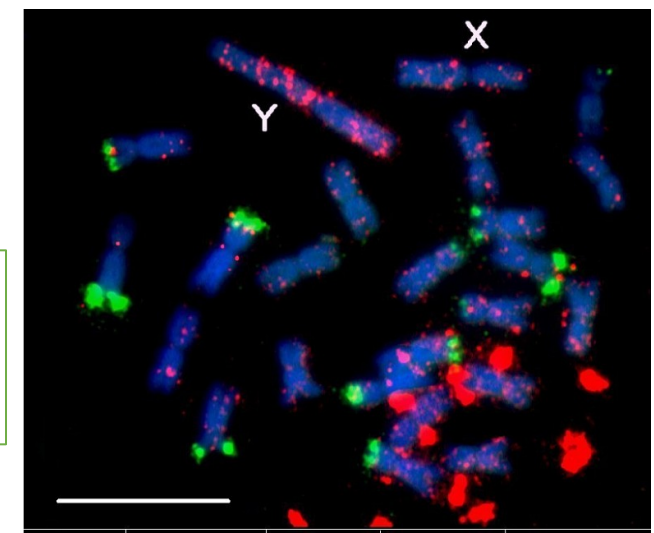
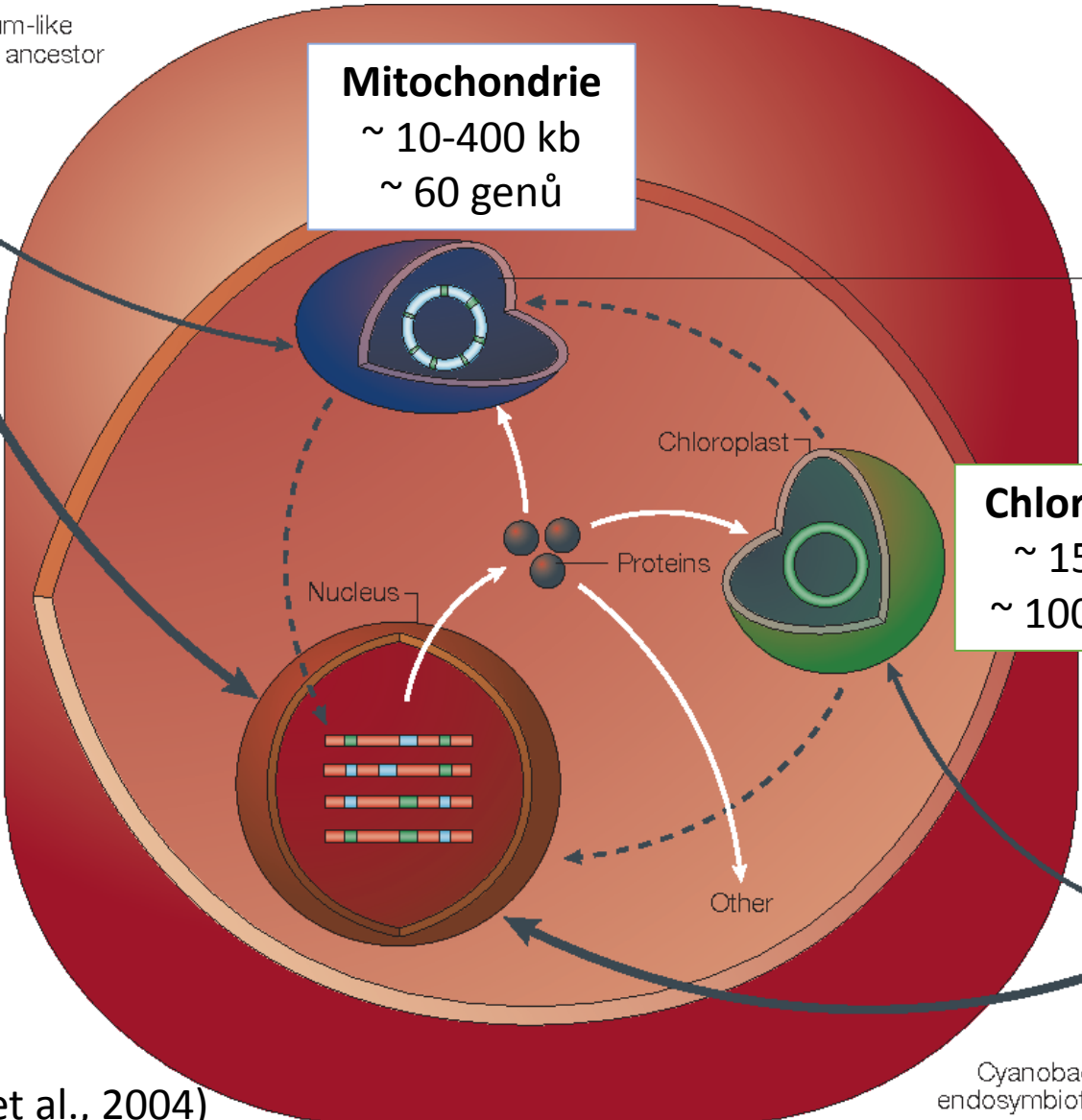
**EGT dnes:**

<u>Donor</u>		<u>Akceptor</u>
Mitochondrie	->	Jádro
Jádro	->>	Mitochondrie
Chloroplast	->>>	Mitochondrie (1-6 %)
Chloroplast	->>>>	Jádro

**Chloroplast**  
 ~ 150 kb  
 ~ 100 genů



**Sinice**  
 ~ 3000 kb  
 ~ 3000 genů



(Timmis et al., 2004)

Table 1 | Sizes and coding content of some organelle and prokaryote genomes

Genome	Length [kbp]	Number of protein-coding genes	GenBank accession number
<b>Algae</b>			
cp <i>Porphyra purpurea</i>	191	200	PPU38804
cp <i>Cyanidium caldarium</i>	165	197	AF022186
cp <i>Guillardia theta</i>	122	148	AF041468
cp <i>Cyanophora paradoxa</i>	136	136	CPU30821
cp <i>Odontella sinensis</i>	120	124	OSCHLPLXX
cp <i>Euglena gracilis</i>	143	58	CLEGCGA
<b>Land plants</b>			
cp <i>Marchantia polymorpha</i>	121	84	CHMPXX
cp <i>Chlorella vulgaris</i>	151	78	AB001684
cp <i>Nicotiana tabacum</i>	156	76	CHNTXX
cp <i>Oryza sativa</i>	134	76	X15901
cp <i>Zea mays</i>	140	76	ZMA86563
cp <i>Pinus thunbergii</i>	120	69	PINCPTRPG
<b>Non-photosynthetic plastids</b>			
cp <i>Toxoplasma gondii</i>	35	26	U87145
cp <i>Eimeria tenella</i>	35	28	AY217738
cp <i>Epi-fagus virginiana</i>	70	21	EPFCPCG
<b>Cyanobacteria</b>			
<i>Synechocystis</i> sp.	3573	3168	AB001339
<i>Prochlorococcus marinus</i>	1660	1884	NC_005071
<i>Nostoc</i> PCC 7120	6413	5368	AP003602
<i>Nostoc punctiforme</i>	~9000	~7400	<a href="http://www.jgi/doi/gov">http://www.jgi/doi/gov</a>
<b>Plants and algae</b>			
mt <i>Pylaiella littoralis</i>	59	52	NC_003055
mt <i>Marchantia polymorpha</i>	187	41	MPOMTCG
mt <i>Laminaria digitata</i>	38	39	AJ344328
mt <i>Cyanidioschyzon merolae</i>	32	34	NC_000887
mt <i>Arabidopsis thaliana</i>	367	31	MIATGENA
mt <i>Chondrus crispus</i>	26	25	MTCCGNME
mt <i>Scenedesmus obliquus</i>	43	20	NC_002254
<b>Various protists and fungi</b>			
mt <i>Reclinomonas americana</i>	69	67	NC_001823
mt <i>Malawimonas jakobiformis</i>	47	49	AF295546
mt <i>Naegleria gruberi</i>	50	46	NC_002573
mt <i>Rhodomonas salina</i>	48	44	NC_002572
mt <i>Dictyostelium discoideum</i>	56	40	NC_000895
mt <i>Phytophthora infestans</i>	38	40	NC_002387
mt <i>Acanthamoeba castellanii</i>	42	36	U12386
mt <i>Cafeteria roenbergensis</i>	43	34	NC_000946
mt <i>Monosiga brevicollis</i>	77	32	AF538053
mt <i>Physarum polycephalum</i>	63	20	AB027295
mt <i>Harpochytrium</i> sp	24	14	AY182006
mt <i>Candida albicans</i>	40	13	NC_002653
mt <i>Cryptococcus neoformans</i>	25	12	NC_004336
mt <i>Plasmodium falciparum</i>	6	3	NC_001677
<b>Anaerobic mitochondria</b>			
mt Hydrogenosomes*	0	0	
<b><math>\alpha</math>-proteobacteria</b>			
<i>Caulobacter crescentus</i>	4017	3767	AE006573
<i>Mesorhizobium loti</i>	7596	7281	BA000012
<i>Bradyrhizobium japonicum</i>	~9100	~8300	BA000040
<b>Yeast</b>			
(nuclear)	13,469	6,327	<a href="http://www.ebi.ac.uk">http://www.ebi.ac.uk</a>

# Velikosti organelových a prokaryotických genomů

(Timmis et al., 2004)

# Některé geny se přenášejí do jádra dříve jiné později

Nejdříve – regulační funkce

(sigma factor of RNA polymerase, gamma subunit of ATPase)

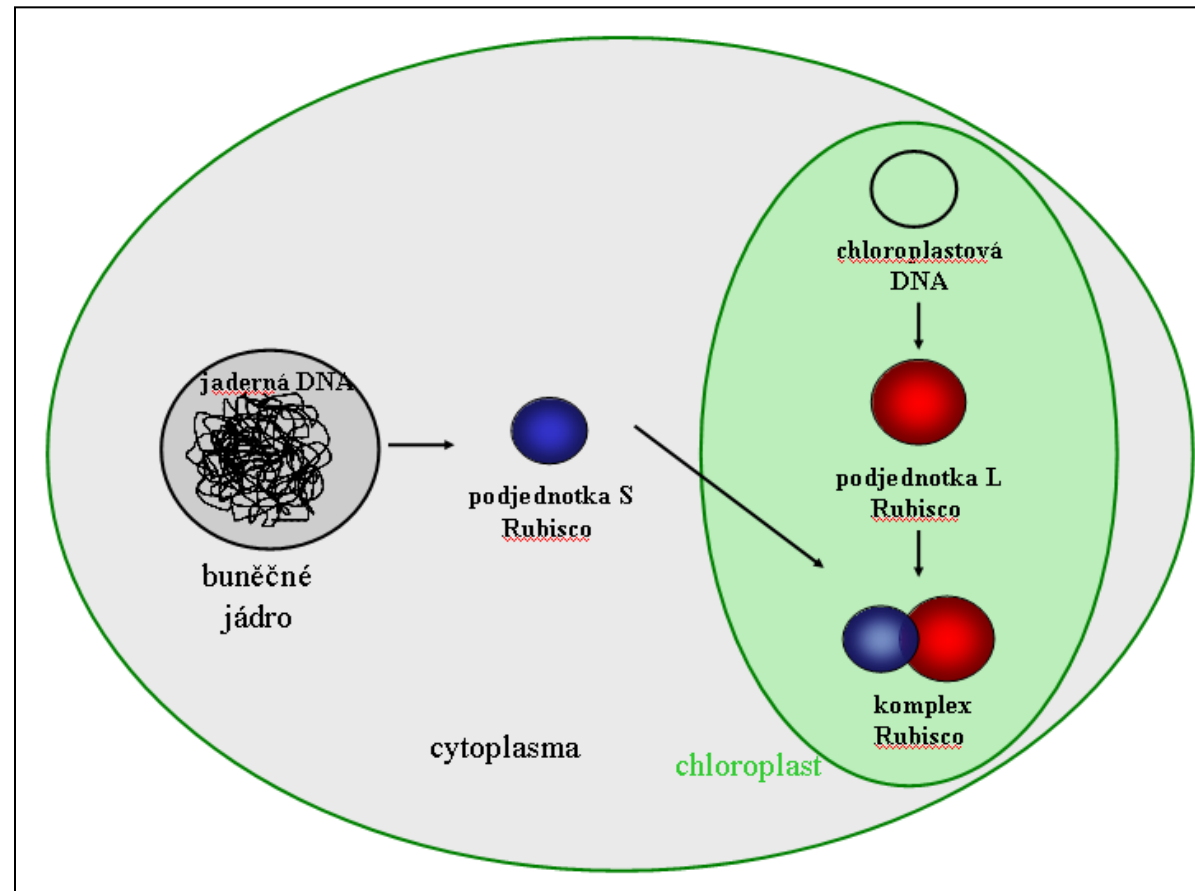
Poslední – translace

- respirace

Rubisco:

katalýza - v plastidu (rbcL)

regulace – v jádře (rbcS)



# Proč některé geny zůstávají v organelách?

1. Hydrophobicita
  - hydrofóbní proteiny jsou těžko importovány do organel
2. Řízení redoxního stavu
  - organely řídí expresi genů, které kódují komponenty jejich elektronového transportu, jejich lokalizace je výhodnější v organelách

Zmenšení genomů u organel a parazitů:

Parazité: - specializace na intracelulární prostředí

- ztráta genů

Organely: - export genů do jádra hostitele

- import produktů

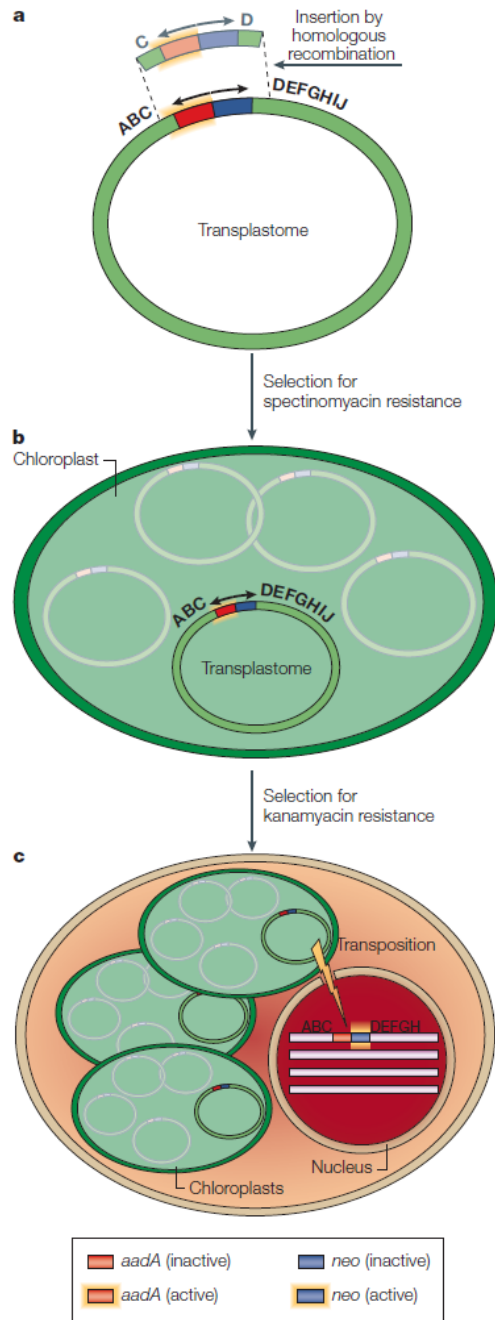
# Mechanizmy genového přenosu

1. Přenos velkých kusů DNA (“bulk DNA” hypothesis)  
intergenové spacery, introns  
experimenty u kvasinek  
>100kb
2. Přenos prostřednictvím cDNA (“cDNA intermediates”)  
přenesená DNA je sestřižena a editována  
rekombinace sestřižené mtDNA s neseestřiženou mtDNA  
heterogenita velikostí mtDNA

A construct that consists of chloroplast sequences (C and D) that flank two selectable marker genes is inserted into the chloroplast genome through homologous recombination, thereby transforming the native plastome into a **TRANSPLASTOME** (a). In the experiments of Huang *et al.*<sup>83</sup>, the flanking chloroplast sequences were in the inverted repeats of the tobacco plastome. One of the selectable genes (*aadA* in the case illustrated) is designed for exclusive expression in the chloroplast and incorporation of this marker confers spectinomycin resistance. The other gene, a neomycin phosphotransferase gene *neoSTLS2* (that encodes NPTII and incorporates a nuclear intron; here *neo*), is tailored, by virtue of a nuclear-specific promoter and the presence of a nuclear intron in the reading frame, for expression only when it is transposed to the nucleus. Continuous selection of growing leaf cells on spectinomycin medium allows transformed plastomes to be selected and eventually the transplastome entirely replaces the native chloroplast genome, such that all copies of the chloroplast genome contain the two selectable marker genes (b).

Selection of cells or progeny seedlings on kanamycin medium allows the detection of the rare cases in which the *neo* gene has changed its location, such that strong expression is promoted from the nuclear environment (c). The progeny of self-fertilized transplastomic plants were not screened directly<sup>83</sup>. Rather, to eliminate low-level expression of *neoSTLS2* from the chloroplast genome, transplastomic plants were crossed with wild-type female plants such that, because of strict maternal inheritance of tobacco plastids (BOX 1), progeny that contained only wild-type chloroplasts were produced. Therefore, chloroplast-to-nucleus transposition must have occurred at some stage during the life cycle of the male parent of the seedlings that were screened on kanamycin plates.

The observation that 1 in 16,000 male tobacco gametes contained a newly integrated segment of chloroplast DNA (REF. 83) was unpredictably high, but it must be an underestimate of the true chloroplast-to-nucleus transposition frequency. In this experiment, the detection strategy enabled the identification only of those events that resulted in an entire, expressed *neoSTLS2* gene in the nucleus. Other regions of the tobacco plastome that integrated in the nucleus without this selectable marker necessarily remained undetected. A similar strategy was used by Stegmann *et al.*<sup>84</sup> and by Lister *et al.*<sup>85</sup>.



# Genový přenos z organel do jádra v reálném čase

## EXPERIMENT:

- exprese/rezistence jen po přenosu do jádra

## Frekvence přenosu:

- v gametách – 1 : 16 000

- v somatických buňkách – 1 : 5 million

## Příčina rozdílu (300x):

- programovaná degenerace plastidů při vývoji pylových zrn zvyšuje frekvenci přenosu

# Kam se přenesená DNA integruje?

- žádné důkazy preferovaných sekvencí či částí chromosomů

Sekvenční **proměnlivost** promiskuitní DNA

>95% identity svědčí o velké obměně organelových sekvencí

Faktory vedoucí k **degeneraci** sekvencí:

- asexualita
- poškozující zplodiny metabolismu
- selekce na malé genomy

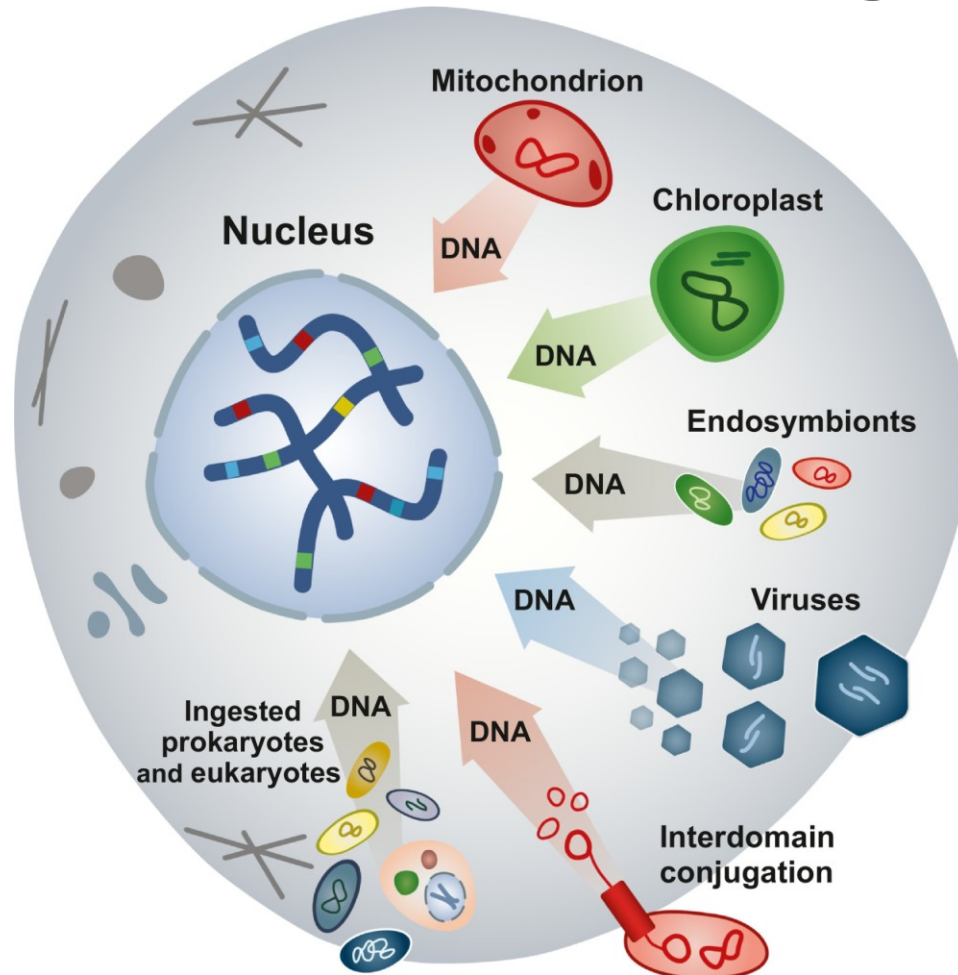
**Kompenzující** faktory (u rostlin)

- polyploidie
- reparace DNA

**Laterální (horizontální) transfer genů  
a „Pan-genom“ Eukaryot?**



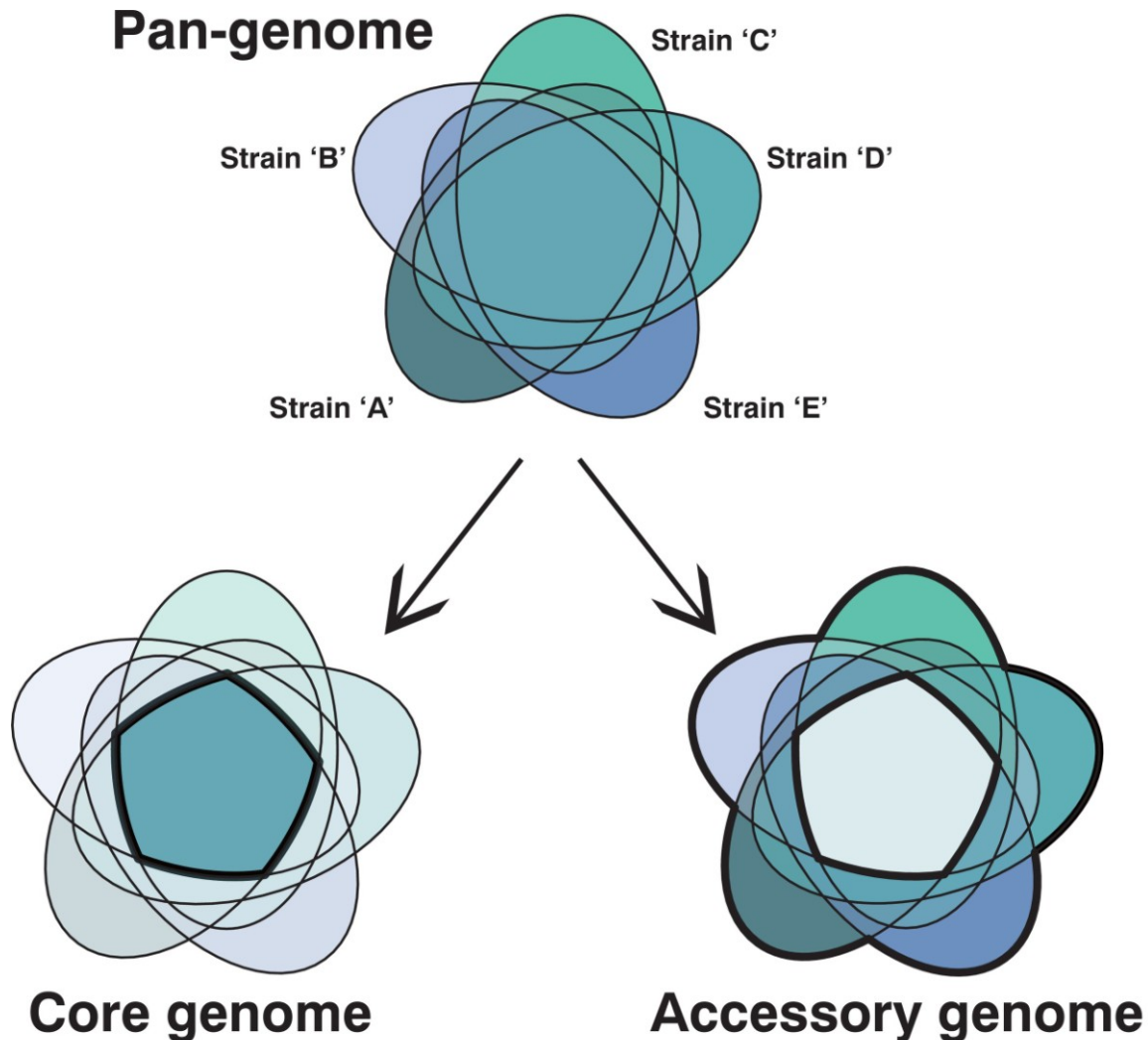
# Laterální (horizontální) transfer genů a „Pan-genom“ u Prvoků?



U jednobuněčných Eukaryot je HGT významným faktorem.

U mnohobuněčných Eukaryot je HGT občasným jevem a souvisí obvykle s parazitismem nebo symbiózou.

# Koncept „Pan-genomu“ u nižších Eukaryot



Studie genomů stejného druhu z různých lokalit odhalily populačně specifické geny odvozené od jiných organismů žijících v místě (bakterií nebo eukaryot) = accessory genome

Pan-genom = Základní (core) genom + Přídavné (accessory) genomy všech jednotlivých populací.

# HGT u mnohobuněčných Eukaryot

Table A1. Selected examples of organisms in which horizontal gene transfer was found.

Gene Donor	Recipient of Gene	References
<b>Bacteria:</b> <i>Agrobacterium rhizogenes</i>	<b>Plant:</b> <i>Nicotiana</i>	[108]
<b>Plant:</b> <i>Oryza</i>	<b>Plant:</b> <i>Zea mays</i>	[109]
<b>Plant:</b> Flowering plant: <i>Asterid</i>	<b>Plant:</b> <i>Gnetum</i>	[110]
<b>Plant:</b> <i>Monocot</i> <i>Eudicot</i> <i>Unknown</i> <i>Ranunculales</i> <i>Monocot</i>	<b>Plant:</b> <i>Aktinidia</i> <i>Amborella</i> <i>Betulaceae</i> <i>Caprifoliaceae</i> <i>Sanguinaria</i>	[58]

Gene Donor	Recipient of Gene	References
<b>Bacteria:</b> <i>Nostoc azollae</i>	<b>Fern:</b> genus <i>Azolla</i>	[111]
<b>Plant:</b> Multiple grass lineages	<b>Plant:</b> <i>Allotroopsis semialata</i>	[112]
<b>Cyanobacteria:</b> <i>Calothrix</i>	<b>Alga:</b> <i>Euglena myxocylindracea</i>	[113]
<b>Alga:</b> <i>Porphyra purpurea</i>	<b>Yeast:</b> <i>Saccharomyces cerevisiae</i>	[114]
<b>Bacteria:</b> <i>Lactococcus</i> sp.	<b>Plant:</b> <i>Amborella trichopoda</i>	[115]
<b>Plant:</b> Flowering plants Mosses	<b>Plant:</b> <i>Phaseolus vulgaris</i>	[82]
<b>Plant:</b> <i>Eudicot</i>	<b>Plant:</b> <i>Apodanthaceae</i>	[116]
<b>Plant:</b> <i>Fabales</i>	<b>Plant:</b> <i>Plantago</i>	[117]
<b>Plant:</b> <i>Orobanchaceae</i>	<b>Yeast:</b> <i>Saccharomyces cerevisiae</i>	[118]
<b>Bacteria:</b> <i>Pseudomonas</i> spp.	<b>Plant:</b> Flowering plant: <i>Santalales</i>	[119]
<b>Plant:</b> <i>Ericaceae</i> <i>Cyrtillaceae</i>	<b>Plant:</b> <i>Ternstroemia</i>	[120]
<b>Plant:</b> Flowering plant: <i>Populus</i>	<b>Plant:</b> Flowering plant: <i>Populus</i>	[121]
<b>Cyanobacteria:</b> <i>Crocospira watsonii</i> <i>Trichodesmium erythraeum</i>	<b>Alga:</b> <i>Heterocapsa triquetra</i> <i>Karlodinium micrum</i> <i>Oxyrrhis marina</i>	[122]
<b>Plant:</b> <i>Poa palustris</i>	<b>Plant:</b> <i>Festuca ovina</i>	[123]
<b>Plant:</b> <i>Orobanche</i>	<b>Plant:</b> <i>Phelipanche</i>	[124]
<b>Fungus:</b> <i>Physcomitrella patens</i> <i>Coprinopsis cinerea</i> <i>Laccaria bicolor</i>	<b>Plant:</b> <i>Oryza sativa</i> <i>Arabidopsis thaliana</i>	[125]
<b>Plant:</b> <i>Cuscuta</i>	<b>Plant:</b> <i>Plantago</i>	[126]
<b>Dinoflagellate:</b> <i>Prorocentrum minimum</i>	<b>Dinoflagellate:</b> <i>Prorocentrum minimum</i> nucleus	[127]
<b>Alga</b>	<b>Fern:</b> <i>Mankyua chejuense</i> <i>Helminthostachys zeylanica</i> <i>Botrychium ternatum</i>	[128]
<b>Bacteria</b>	<b>Alga:</b> Ochrophytes	[129]
<b>Bacteria:</b> <i>Wolbachia</i>	<b>Nematode:</b> <i>Onchocerca volvulus</i>	[130]
<b>Bacteria:</b> <i>Wolbachia</i>	<b>Animal:</b> Insect: <i>Callosobruchus</i>	[131]

Gene Donor	Recipient of Gene	References
<b>Bacteria:</b> <i>Wolbachia</i>	<b>Animal:</b> Insect: <i>Drosophila ananassae</i>	[132]
<b>Alga:</b> <i>Vaucheria litorea</i>	<b>Animal:</b> Snail: <i>Elysia chlorotica</i>	[133]
<b>Bacteria:</b> <i>Wolbachia</i>	<b>Animal:</b> Insect: <i>Aedes aegypti</i>	[134]
<b>Cyanobacteria:</b> <i>Synechococcus</i> sp. <i>Prochlorococcus</i> sp.	<b>Protozoan:</b> Amoeba: <i>Paulinella chromatophora</i>	[135]
<b>Human:</b> <i>Homo sapiens</i>	<b>Protozoan:</b> Malaria parasite: <i>Plasmodium vivax</i>	[136]

Některé studie odhadují, že asi 1-2 % genů eukaryotických genomů je tvořeno geny získanými HGT od virů, bakterií a jiných eukaryot relativně nedávno.

Desítky zdokumentovaných příkladů, ale pozor! Ne všechny jsou spolehlivě prokázané (metodika, kontaminace vzorků nebo veřejných zdrojů dat).

# Jaké geny se nejčastěji přenáší HGT

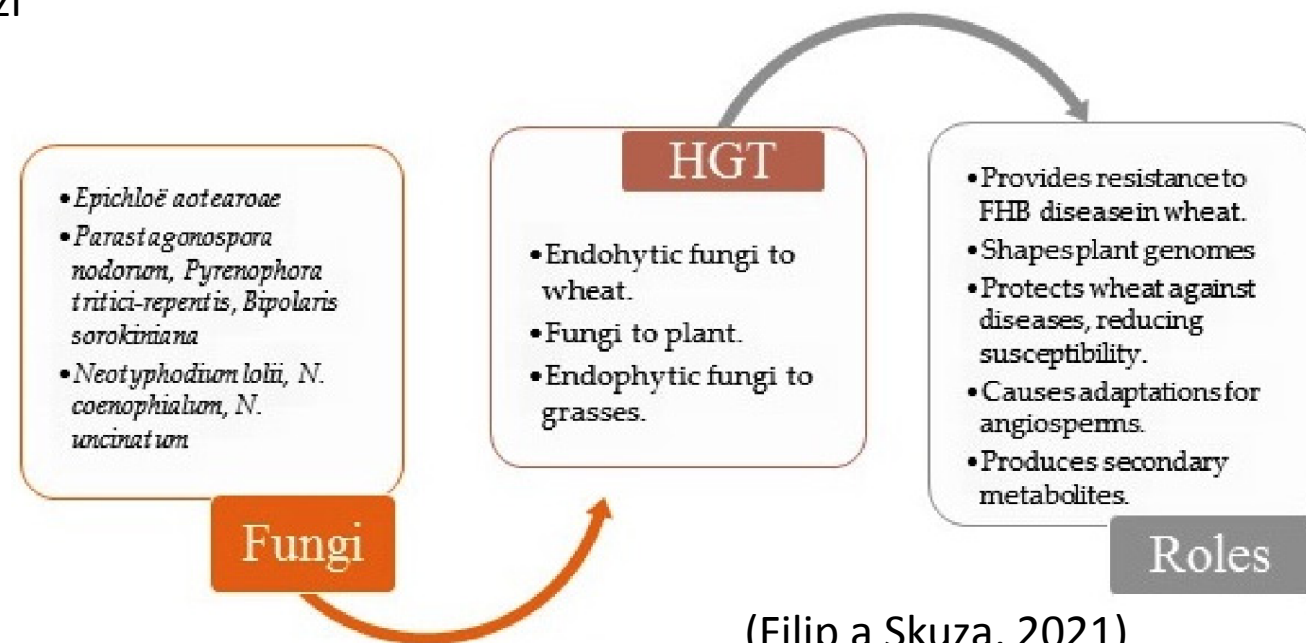
Geny poskytující okamžitou adaptivní výhodu:

Biosyntéza proteinů – Rhodoquinone (akceptor elektronu během syntézy ATP, umožňuje dýchání v anaerobních podmínkách), bakterie -> prvoci (2x) -> další prvoci

Geny pro transportéry látek – transport a využití konkrétních metabolitů, známo mnoho příkladů HGT mezi houbami

Celé bakteriální operony – biosyntéza sideroforu, enterobakterie -> střevní kvasinky hmyzu, lepší využití železa

Rezistence k chorobám a stresu



(Filip a Skuza, 2021)

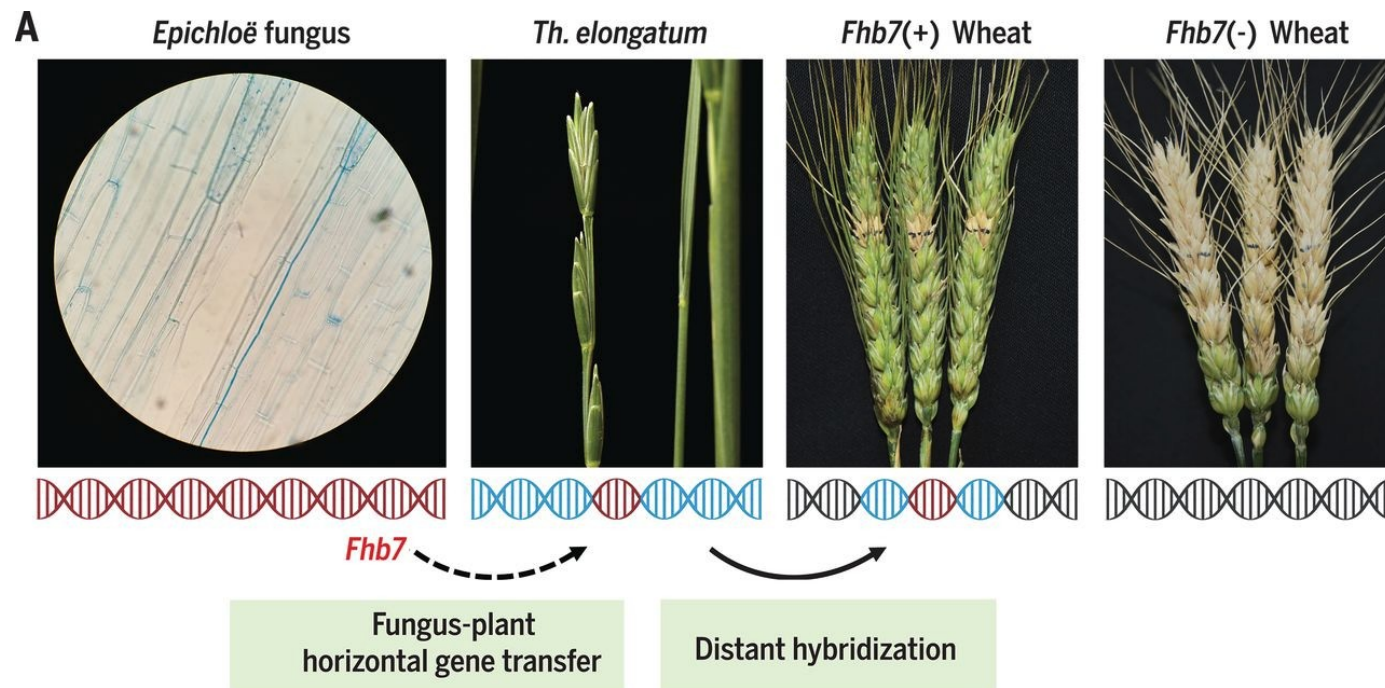
# HGT u mnohobuněčných Eukaryot

Přenos genu *Fhb7* pro rezistenci k fusarióze z endofytické houby do příbuzného pšenice *Thinopyrum elongatum*

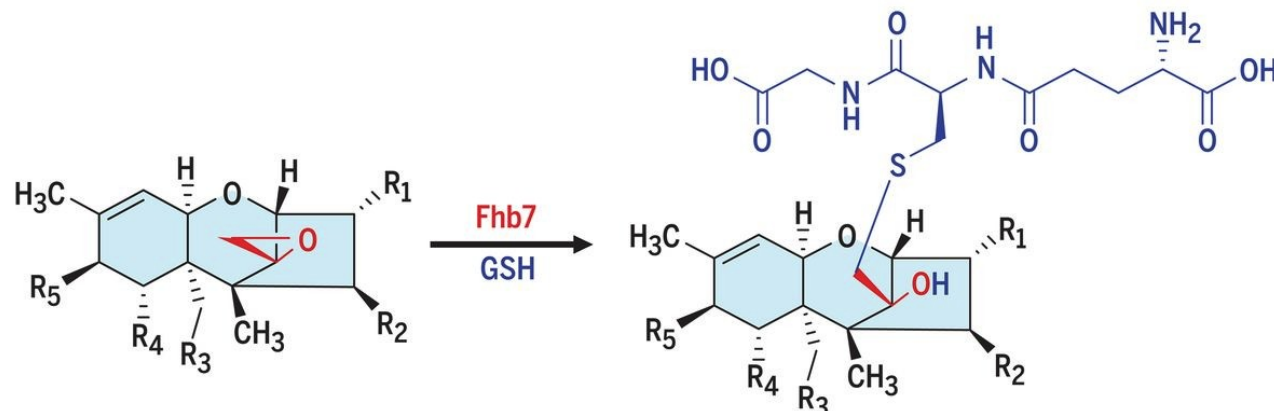
Fusarióza:  
Růžovění pšenice  
Poškození zrn  
Mykotoxiny

*Fhb7* – glutathionsyntetáza, detoxikace trichothecenů produkovaných fusariem prostřednictvím vazby glutathionu.  
97% homologie s genem *E. fungus*  
0% homologie se všemi rostlinnými databázemi

(Wang et al., 2020, Science)



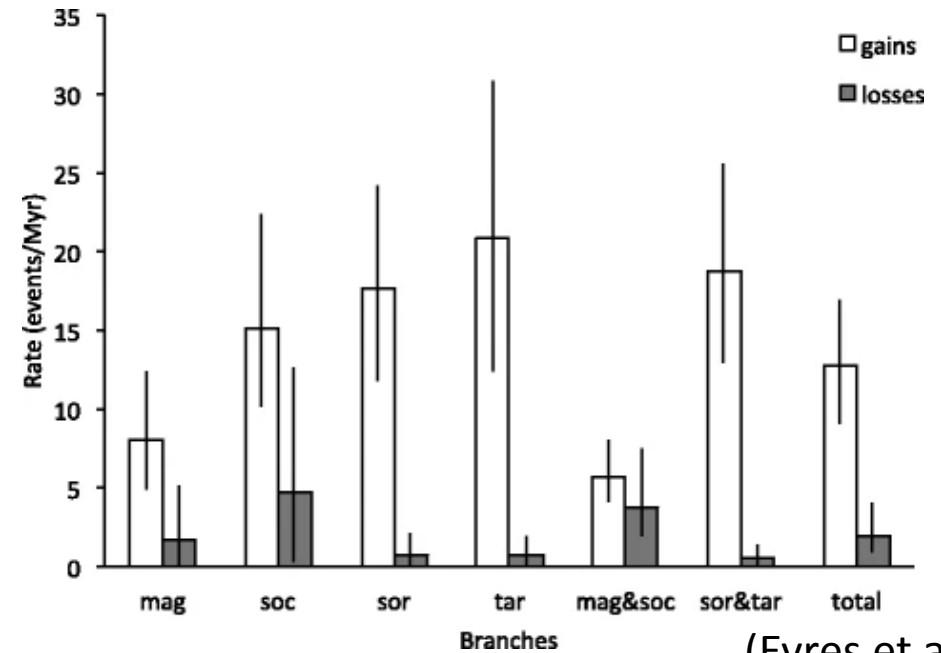
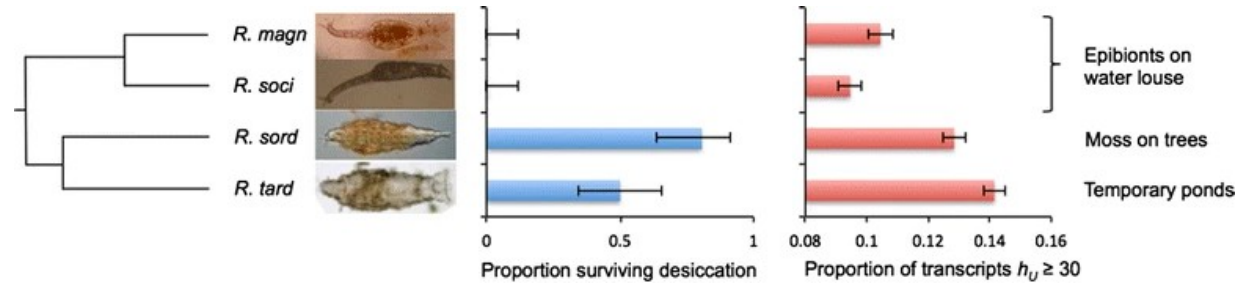
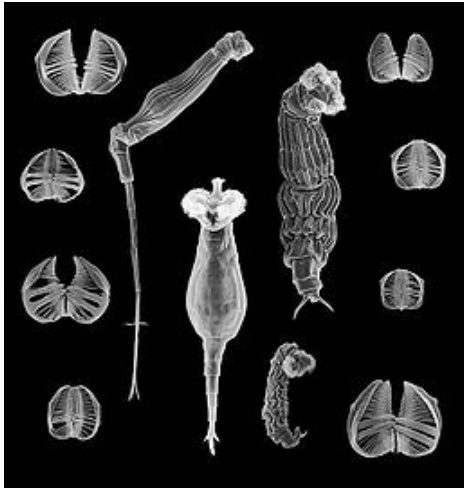
**B**



# Případ Bdelloid Rotifers, vířníci)

Pijavenky (vířníci Bdelloidea) – 40-60 milionů let stará linie, 18 rodů, 360 druhů.

Partenogenetické samičky, žádný sex. Klidová stadia na přežití nepříznivých podmínek - vyschnutí. Unikátní způsob získávání genetické informace od bakterií, hub, prvoků, řas – vypůjčují si jejich geny (horizontální genový transfer) - náhrada sexu?



(Eyres et al., 2015)

# Případ Bdelloid Rotifers, vířníci)

Vířníci si osvojují geny bakterií, hub i řas

Gene ID, name	Contig ID	Introns	AI	% Identity to best hit	Best hit, E-value	Best hit, metazoan	Best hit, taxonomy	Definition
AV10027_XynB	Av212A	0	460	63	0.00E+00	No hits	Bacteria; Bacteroidetes	Xylosidase/arabinosidase
AV10001_NRPS	Av110A	10	460	32	0.00E+00	No hits	Bacteria; (Proteobacteria/ Cyanobacteria)	Nonribosomal peptide synthetase
AV10134_PheA	161F07	0	400	61	1.00E-174	No hits	(Fungi; Bacteria)	Monooxygenase, FAD dependent
AV10002_TrkA	Av110A	0	379	54	1.00E-175	4.00E-11	Bacteria; Proteobacteria	Monooxygenase, NAD dependent
PR10002_MviM	182F10	0	327	67	1.00E-149	2.00E-07	Bacteria; (Acidobacteria/ Chloroflexi)	Oxidoreductase
PR10010_DAP2	182F10	0	310	27	1.00E-140	1.00E-05	Bacteria; (Acidobacteria/ Proteobacteria)	Prolyl oligopeptidase*
AV10104_Dur3	AvTelL.B	1	243	44	1.00E-132	4.00E-27	Eukaryota; Fungi	Urea active transporter*
PR10012_RamA	182J17	0	246	31	1.00E-107	No hits	(Bacteria; Fungi)	$\alpha$ -L-Rhamnosidase
AV10121_NRPS	9907	4	237	30	1.00E-103	No hits	Bacteria; Cyanobacteria	Nonribosomal peptide synthetase
AV10153_XghA	210B3	0	212	50	1.00E-108	2.00E-16	Eukaryota; Fungi	Endo-xylogalacturonan hydrolase
AV10042_HemK	Av240B	1	199	56	2.00E-91	1.00E-04	Bacteria; Proteobacteria	HemK-like methyltransferase
AV10092_ $\beta$ -Gal	AvTelL.A	0	153	33	1.00E-105	4.00E-39	Eukaryota; Viridiplantae	$\beta$ -D-Galactosidase
AV10044_Alr	Av240B	1	152	38	1.00E-67	No hits	Bacteria; Bacteroidetes	Alanine racemase
AV10025_AMH	Av212A	1	150	52	8.00E-77	2.00E-11	Eukaryota; Fungi	Amidohydrolase
AV10045_Ddl	Av240B	1	138	40	1.00E-60	No hits	Bacteria; Bacteroidetes	D-Alanine-D-alanine ligase
AV10140_PLDc	193E18	2	126	31	1.00E-55	No hits	Eukaryota; Fungi	Phospholipase-D active site motif protein*
AV10016_FabG	Av212A	0	98	58	1.00E-74	8.00E-32	Bacteria	Short-chain dehydrogenase/reductase
AV10109_FabG	AvTelL.B	0	92	57	4.00E-73	5.00E-33	Bacteria	Short-chain dehydrogenase/reductase*
AV10011_FabG	Av212A	0	88	54	6.00E-67	2.00E-28	Bacteria	Short-chain dehydrogenase/reductase
AV10071_HAL	AvTelK.A	0	77	48	2.00E-61	1.00E-27	Bacteria	Histidine ammonia-lyase
AV10095_GCNS5	AvTelL.A	0	59	35	2.00E-27	No hits	Bacteria; Proteobacteria	GCNS5-related N-acetyltransferase <sup>+/*</sup>
AV10158_FabG	210B3	2	46	41	2.00E-39	2.00E-19	Bacteria	Short-chain dehydrogenase/reductase

(Gladyshev et al., 2008)

## **2. TANDEMŮVÉ REPETICE**

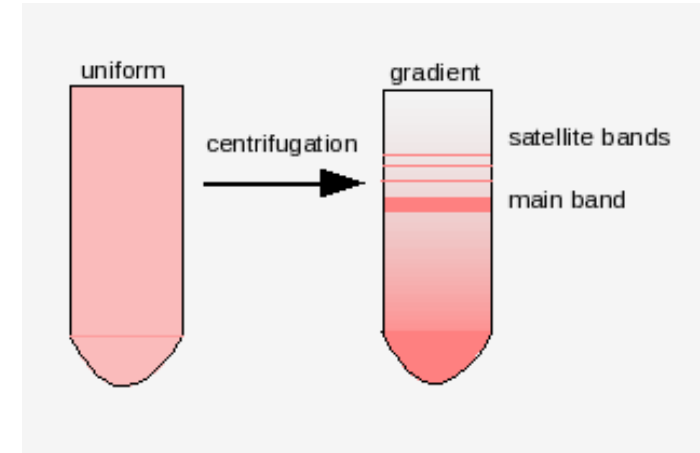


# Tandemové repetice

Jednotka (= monomer) se opakuje n-krát za sebou

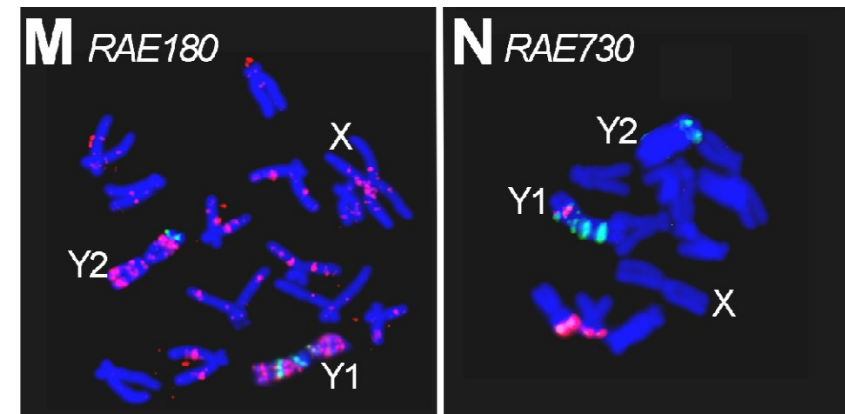


Satelity – objeveny při hustotní gradientové centrifugaci - tvoří satelitní proužek kvůli jinému složení bazí.



- Přítomny ve všech genomech (prokaryota i eukaryota)
- Klasifikace:
  - **Satelity** – významná složka genomu ( $i > 50\%$ ), tvoří bloky heterochromatinu, délka monomerů  $i > 2$  kbp)
  - **Minisatelity** – 0,5-30 kbp, délka monomeru 10-100 bp
  - **Mikrosatelity** – krátké úseky, délka monomeru  $\leq 6$  bp

Typické rozložení tandemových repetitiv na chromosomech šťovíku



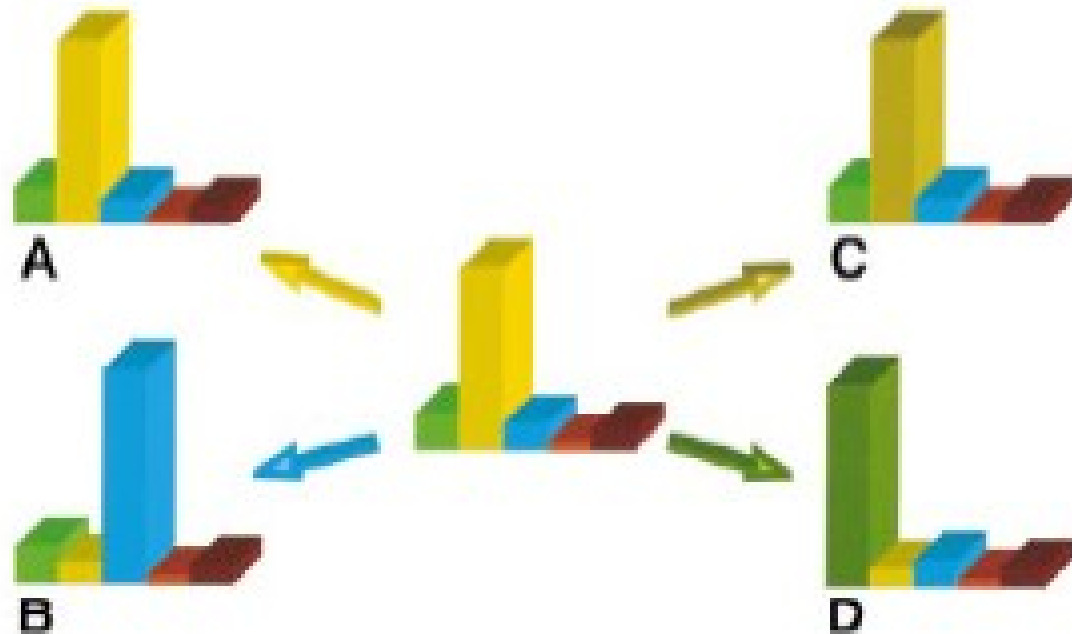
Strukturní a regulační funkce, expanze mikrosatelitů u některých chorob

# Tandemové repetice, „knihovna“ satelitní DNA

Hlavní parametry:

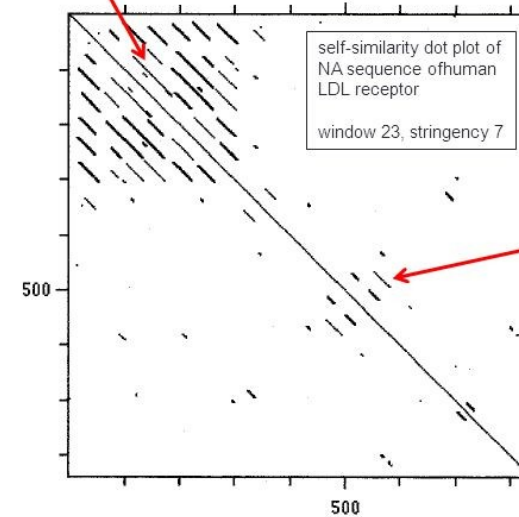
- počet kopií (změna v B a D)
- sekvence DNA (změny C a D)

Đ. Ugarković and M. Plohl



## Repeats in dot plot

minisatellites



direct repeats

inverted repeats



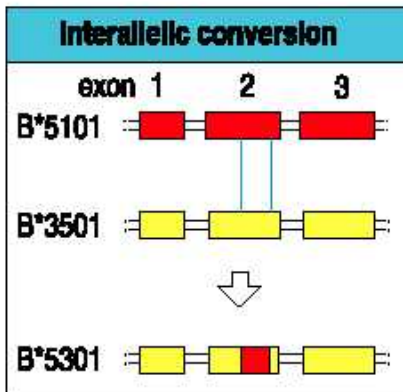
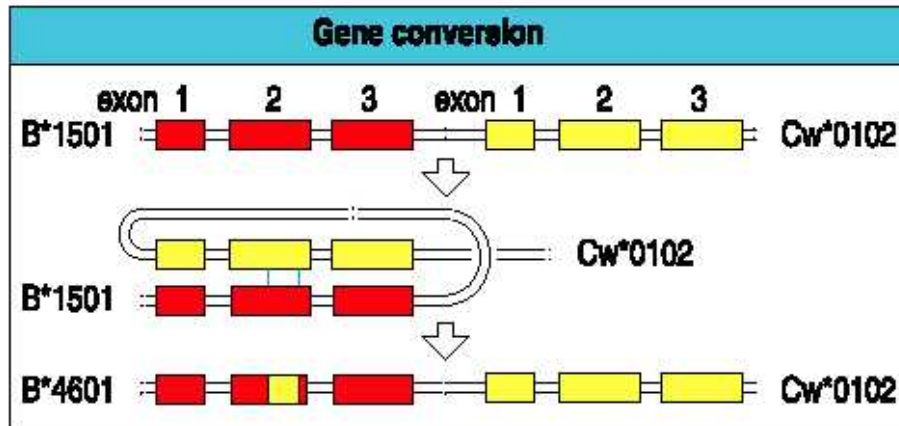
# Evoluce tandemových repeticí

Evoluce v koncertu (concerted evolution)

Genová konverze

Nerovnoměrný crossing-over

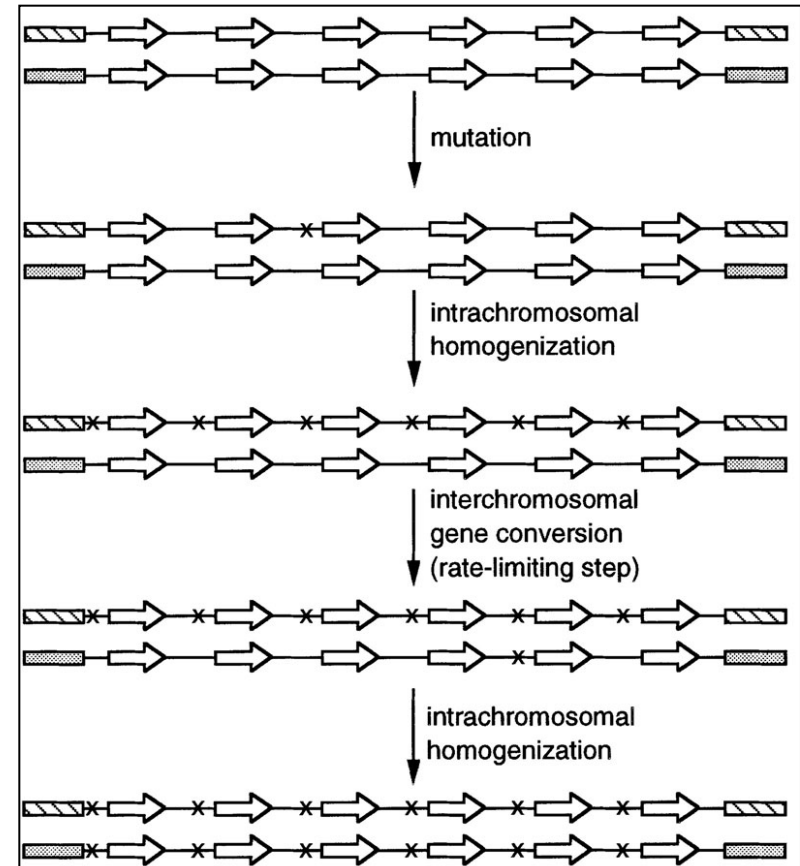
Evoluce satelitních sekvencí – skládání ze segmentů – vznik higher order repeats (HORs)



Genová konverze

- interchromosomální

- intrachromosomální



# Genová konverze transposonů na pohlavních chromosomech

Available online at [www.sciencedirect.com](http://www.sciencedirect.com)

ScienceDirect

Gene 390 (2007) 92–97

ELSEVIER

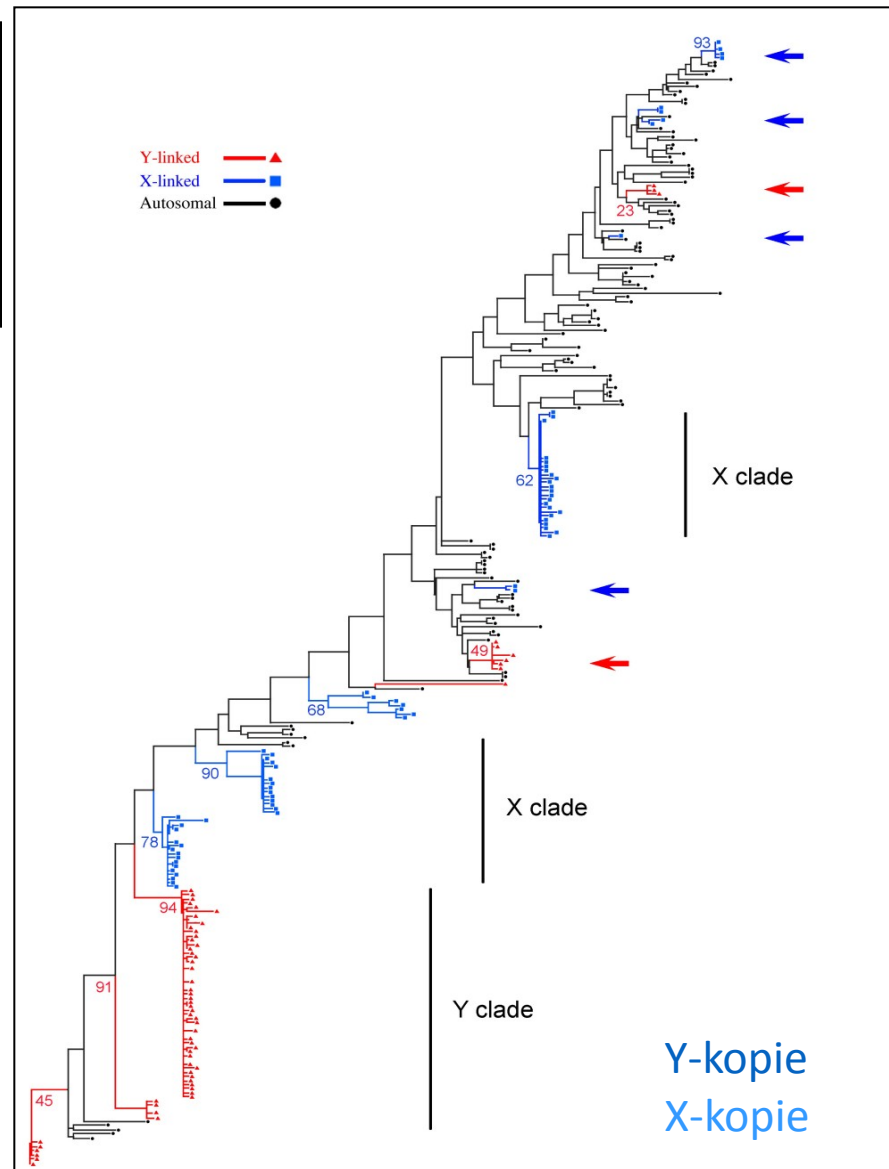
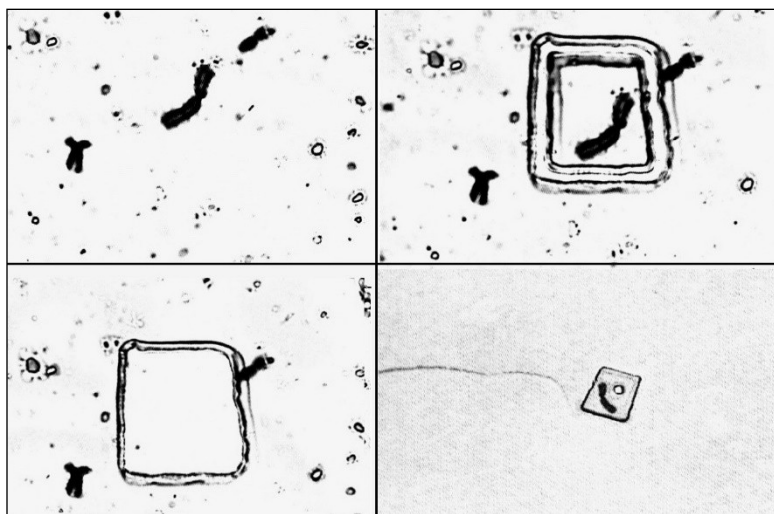
**GENE**  
SECTION  
EVOLUTIONARY GENOMICS  
[www.elsevier.com/locate/gene](http://www.elsevier.com/locate/gene)

High intrachromosomal similarity of retrotransposon long terminal repeats: Evidence for homogenization by gene conversion on plant sex chromosomes?

Eduard Kejnovsky <sup>a,\*</sup>, Roman Hobza <sup>a</sup>, Zdenek Kubat <sup>a</sup>, Alex Widmer <sup>b</sup>, Gabriel A.B. Marais <sup>c</sup>, Boris Vyskot <sup>a</sup>

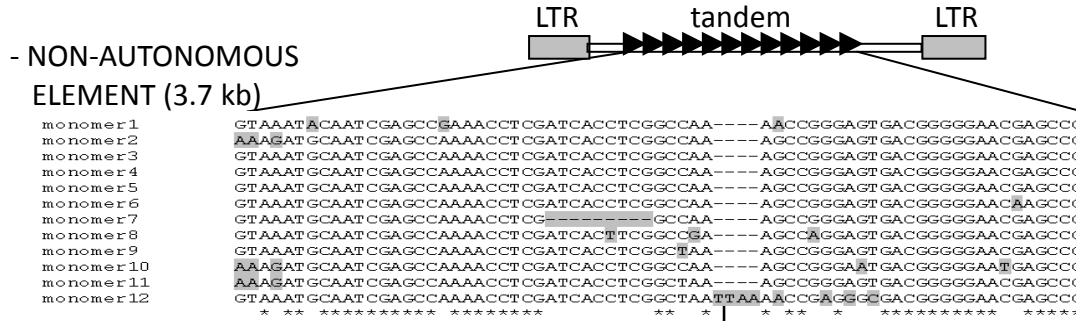
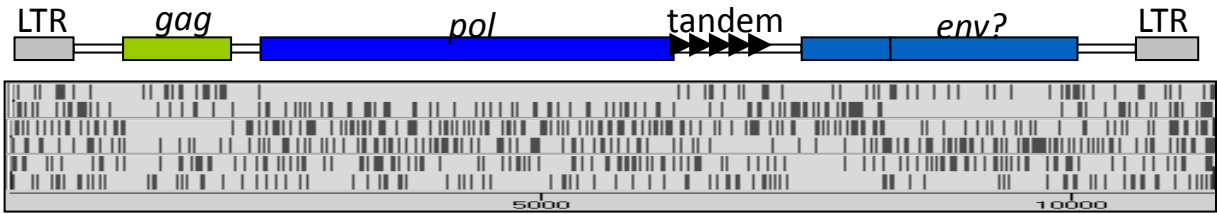
mikrodisekce  
chromosomů Y a X

PCR amplifikace TE na  
mikrodisektovaných  
chromosomech



# Satelitey mohou vznikat i z retroelementů

## Retand element (*Silene latifolia*)



*The Plant Journal* (2020) 101, 484–500

doi: 10.1111/tpj.14546

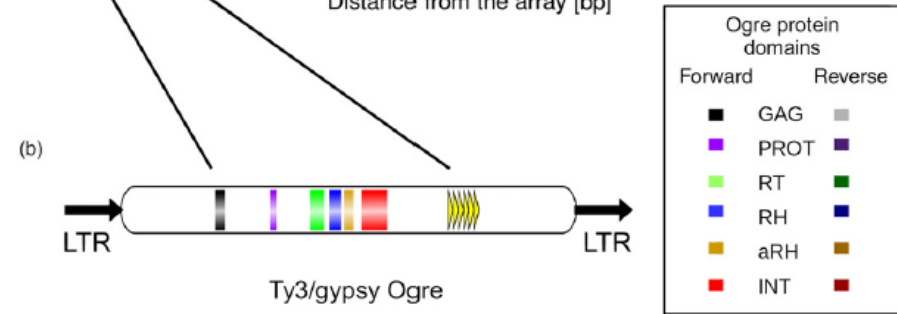
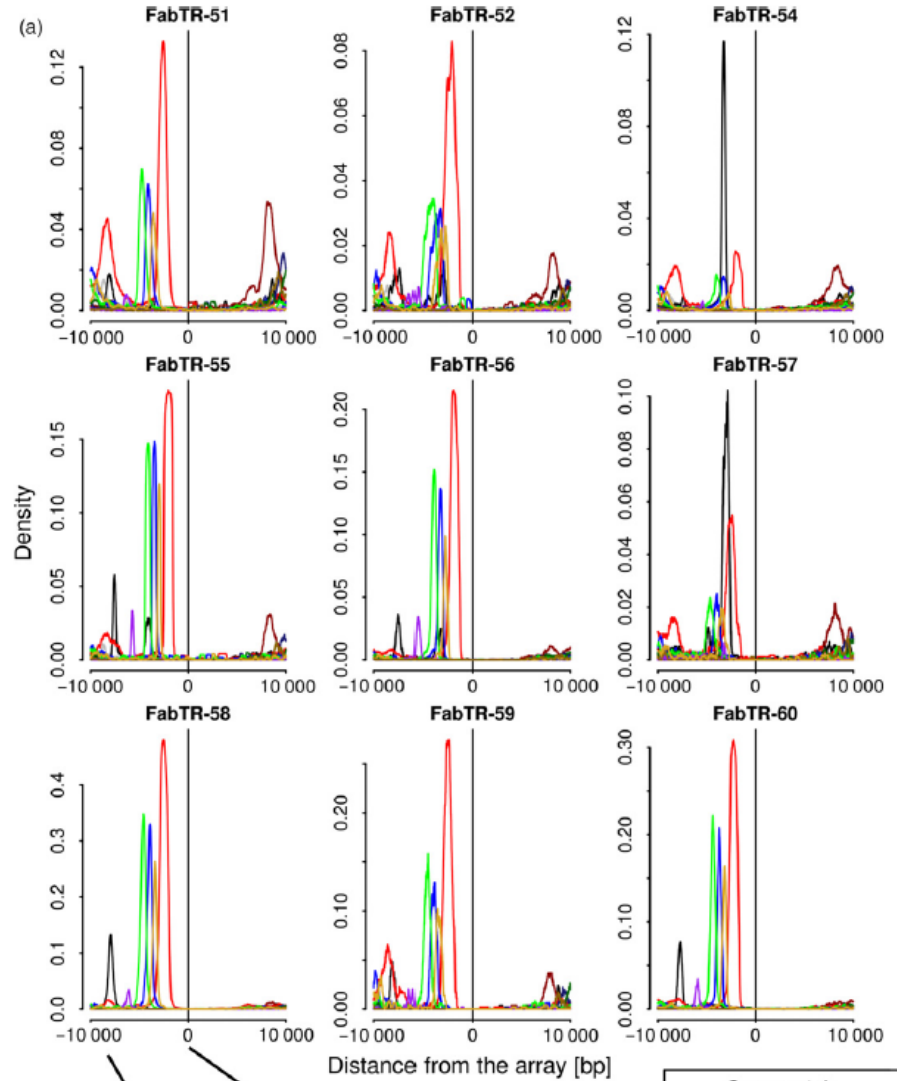
TECHNICAL ADVANCE

### Characterization of repeat arrays in ultra-long nanopore reads reveals frequent origin of satellite DNA from retrotransposon-derived tandem repeats

Tihana Vondrak<sup>1,2</sup>, Laura Ávila Robledillo<sup>1,2</sup>, Petr Novák<sup>1</sup>, Andrea Kobližková<sup>1</sup>, Pavel Neumann<sup>1</sup> and Jirí Macas<sup>1,\*</sup>

<sup>1</sup>Biology Centre, Czech Academy of Sciences, Branišovská 31, České Budějovice CZ-37005, Czech Republic, and

<sup>2</sup>Faculty of Science, University of South Bohemia, České Budějovice, Czech Republic



# **3. TRANSPOZONY**

# Mobilní elementy = transpozony (TE)

**Barbara McClintock (1948)**

Nobelova cena (1983)



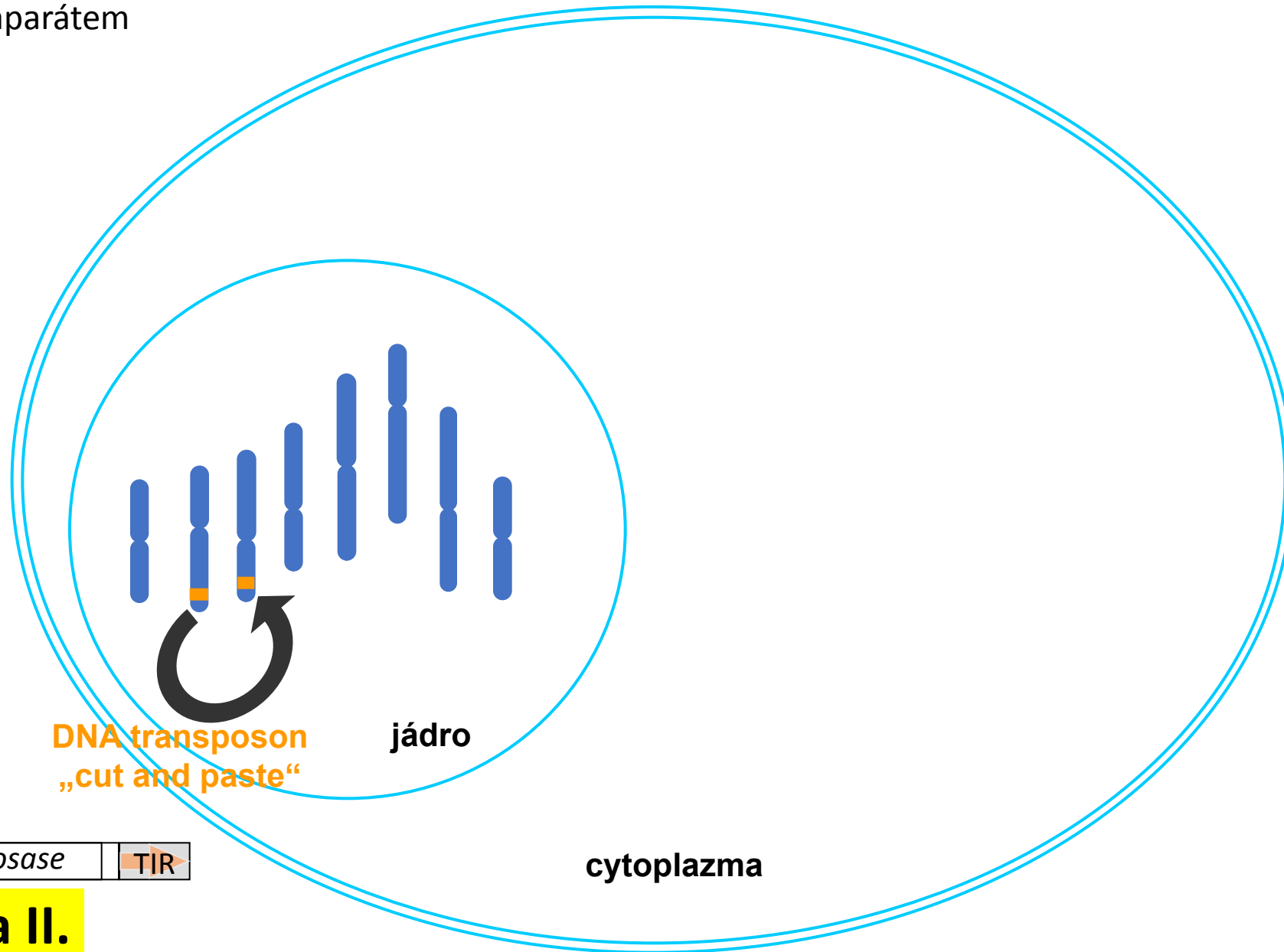
**Disociátor (Ds)  
Aktivátor (Ac)**

Mění polohy na chromozomu a  
ovlivňují přilehlé geny → změna  
barev kukuřičných semen



# Transpozony

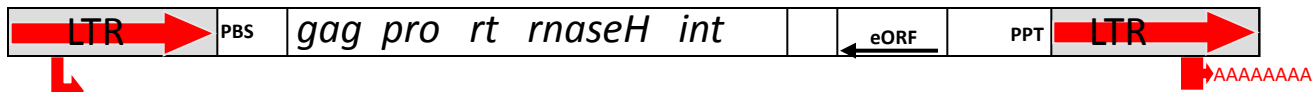
(disponují vlastním aparátem pro transpozici)



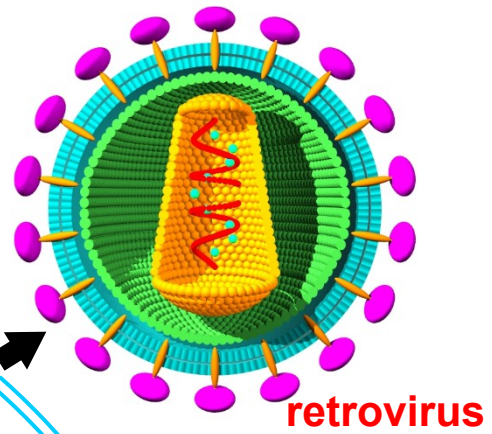
◀ TIR | *transposase* | TIR ▶

**Třída II.**





LTR retrotransposon  
„copy and paste“



retrovirus

Reverzní transkripce

integrace

**Třída I.**

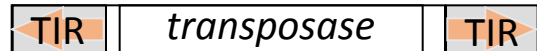
transkripce

translace

DNA transposon  
„cut and paste“

jádro

cytoplazma



**Třída II.**

**Autonomní  
X  
Neautonomní**

# Transpozony

(disponují vlastním aparátem pro transpozici)

# Klasifikace transpozonů

Dle typu  
transpozice

## Třída I. Retrotranspozony

Replikativní transpozice prostřednictvím RNA meziprojektu (copy and paste)

## Třída II. DNA transpozony

Transpozice bez RNA meziprojektu

**Podtřída I** – konzervativní „cut and paste“ mechanismus transpozice

**Podtřída II** – replikativní mechanismus transpozice (např. Helitrony)

(I – štěpí oba řetězce DNA, II – štěpí jen jeden řetězec DNA v místě inzerce i v novém místě)

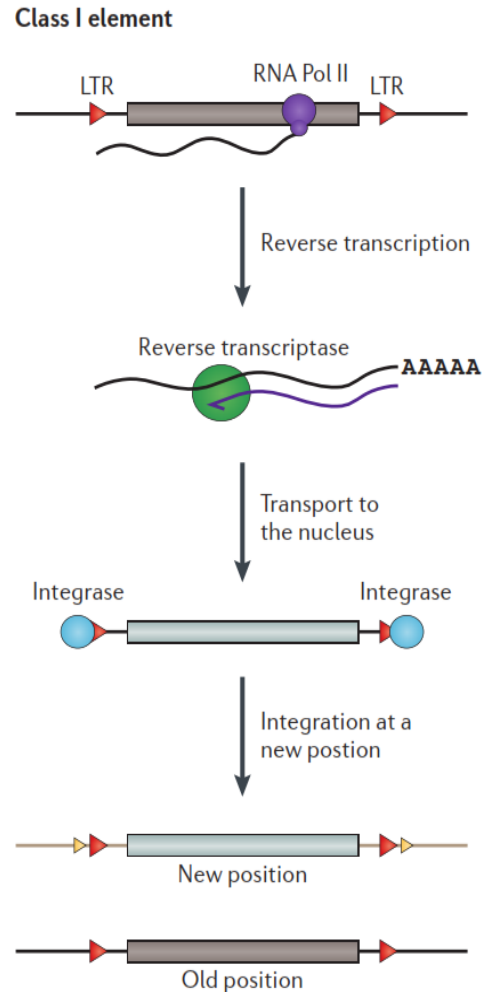
Dle struktury

**Autonomní** - plně funkční, obsahují všechny geny pro transpozici

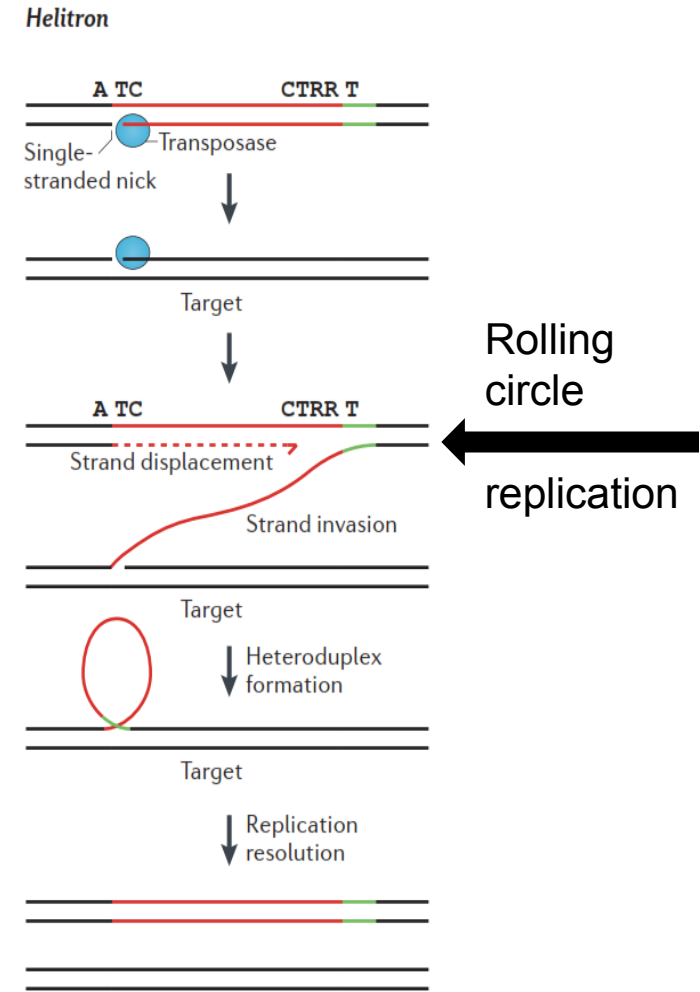
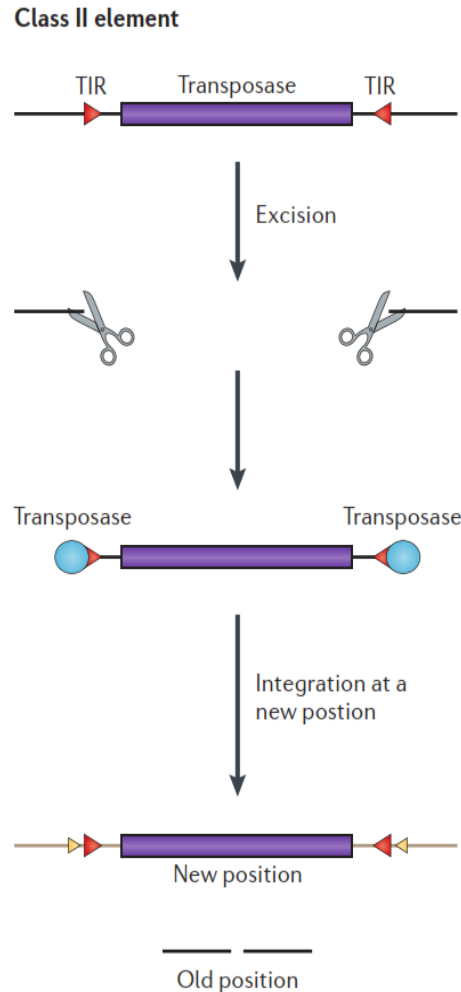
**Neautonomní** – nefunkční, absence genů, transpozice je zajištěna proteiny autonomních elementů

# Základní typy transpozonů dle transpozice

## Třída I Retrotranspozony



## Třída II DNA transpozony



# Třída I.

Classification		Structure	TSD	Code	Occurrence
Order	Superfamily				
<i>Class I (retrotransposons)</i>					
LTR	<i>Copia</i>		4-6	RLC	P, M, F, O
	<i>Gypsy</i>		4-6	RLG	P, M, F, O
	<i>Bel-Pao</i>		4-6	RLB	M
	<i>Retrovirus</i>		4-6	RLR	M
	<i>ERV</i>		4-6	RLE	M
DIRS	<i>DIRS</i>		0	RYD	P, M, F, O
	<i>Ngaro</i>		0	RYN	M, F
	<i>VIPER</i>		0	RYV	O
PLE	<i>Penelope</i>		Variable	RPP	P, M, F, O
LINE	<i>R2</i>		Variable	RIR	M
	<i>RTE</i>		Variable	RIT	M
	<i>Jockey</i>		Variable	RIJ	M
	<i>L1</i>		Variable	RIL	P, M, F, O
	<i>I</i>		Variable	RII	P, M, F
SINE	<i>tRNA</i>		Variable	RST	P, M, F
	<i>7SL</i>		Variable	RSL	P, M, F
	<i>5S</i>		Variable	RSS	M, O

LTR  
retrotranspozony

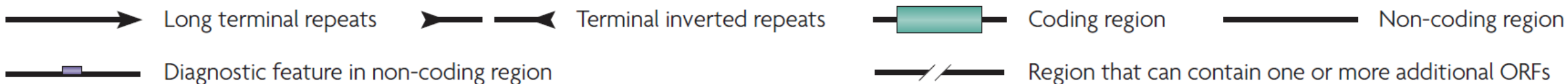
Non-LTR  
retrotranspozony  
(retropozony)

# Třída II.

<i>Class II (DNA transposons) - Subclass 1</i>					
TIR	<i>Tc1-Mariner</i>		TA	DTT	P, M, F, O
	<i>hAT</i>		8	DTA	P, M, F, O
	<i>Mutator</i>		9-11	DTM	P, M, F, O
	<i>Merlin</i>		8-9	DTE	M, O
	<i>Transib</i>		5	DTR	M, F
	<i>P</i>		8	DTP	P, M
	<i>PiggyBac</i>		TTAA	DTB	M, O
	<i>PIF-Harbinger</i>		3	DTH	P, M, F, O
	<i>CACTA</i>		2-3	DTC	P, M, F
	Crypton	<i>Crypton</i>		0	DYC
<i>Class II (DNA transposons) - Subclass 2</i>					
Helitron	<i>Helitron</i>		0	DHH	P, M, F
Maverick	<i>Maverick</i>		6	DMM	M, F, O

## Vysvětlivky k předchozímu obrázku:

### Structural features



### Protein coding domains

AP, Aspartic proteinase	APE, Apurinic endonuclease	ATP, Packaging ATPase	C-INT, C-integrase	CYP, Cysteine protease	EN, Endonuclease
ENV, Envelope protein	GAG, Capsid protein	HEL, Helicase	INT, Integrase	ORF, Open reading frame of unknown function	
POL B, DNA polymerase B	RH, RNase H	RPA, Replication protein A (found only in plants)	RT, Reverse transcriptase		
Tase, Transposase (* with DDE motif)		YR, Tyrosine recombinase	Y2, YR with YY motif		

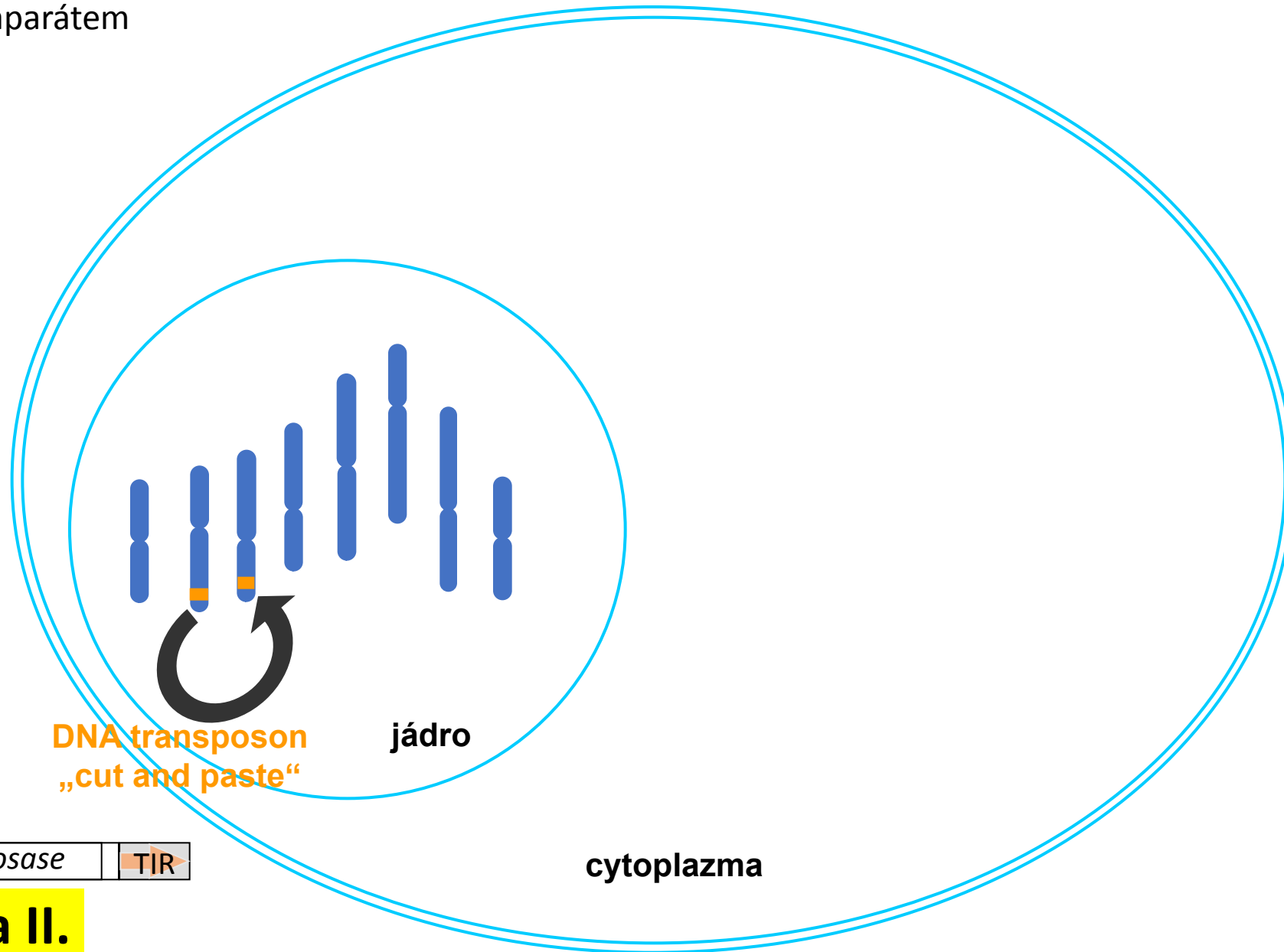
### Species groups

P, Plants      M, Metazoans      F, Fungi      O, Others

# **Třída II. DNA transpozony**

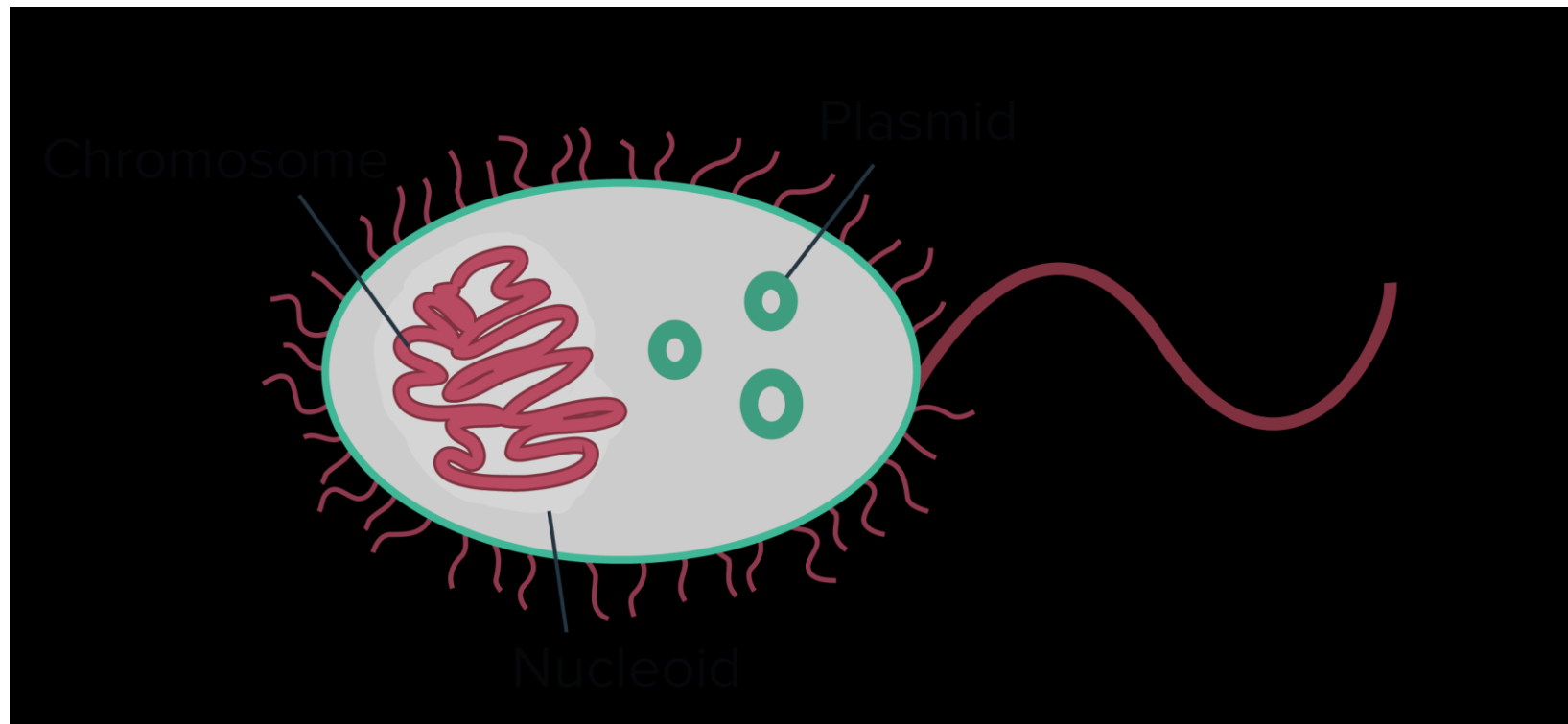
# Transpozony

(disponují vlastním aparátem pro transpozici)



**Třída II.**

# Bakteriální DNA transpozony



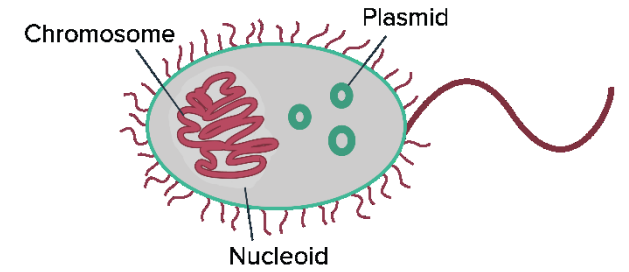


# Transpozony prokaryot

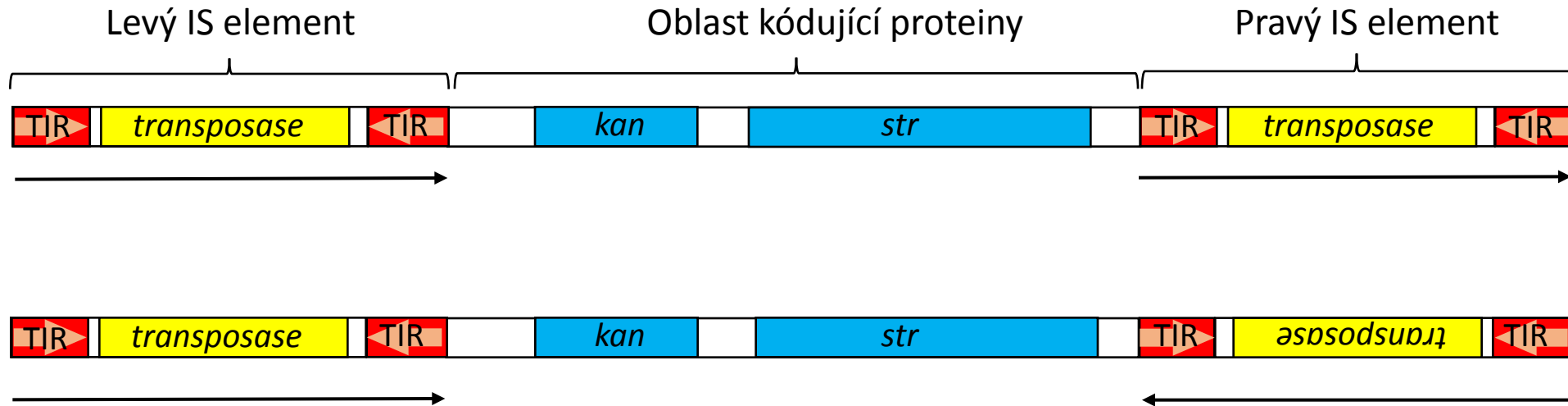
## Inzerční sekvence (IS)



Mobilní v rámci bakteriálního chromozomu nebo mezi chromozomem a plazmidy

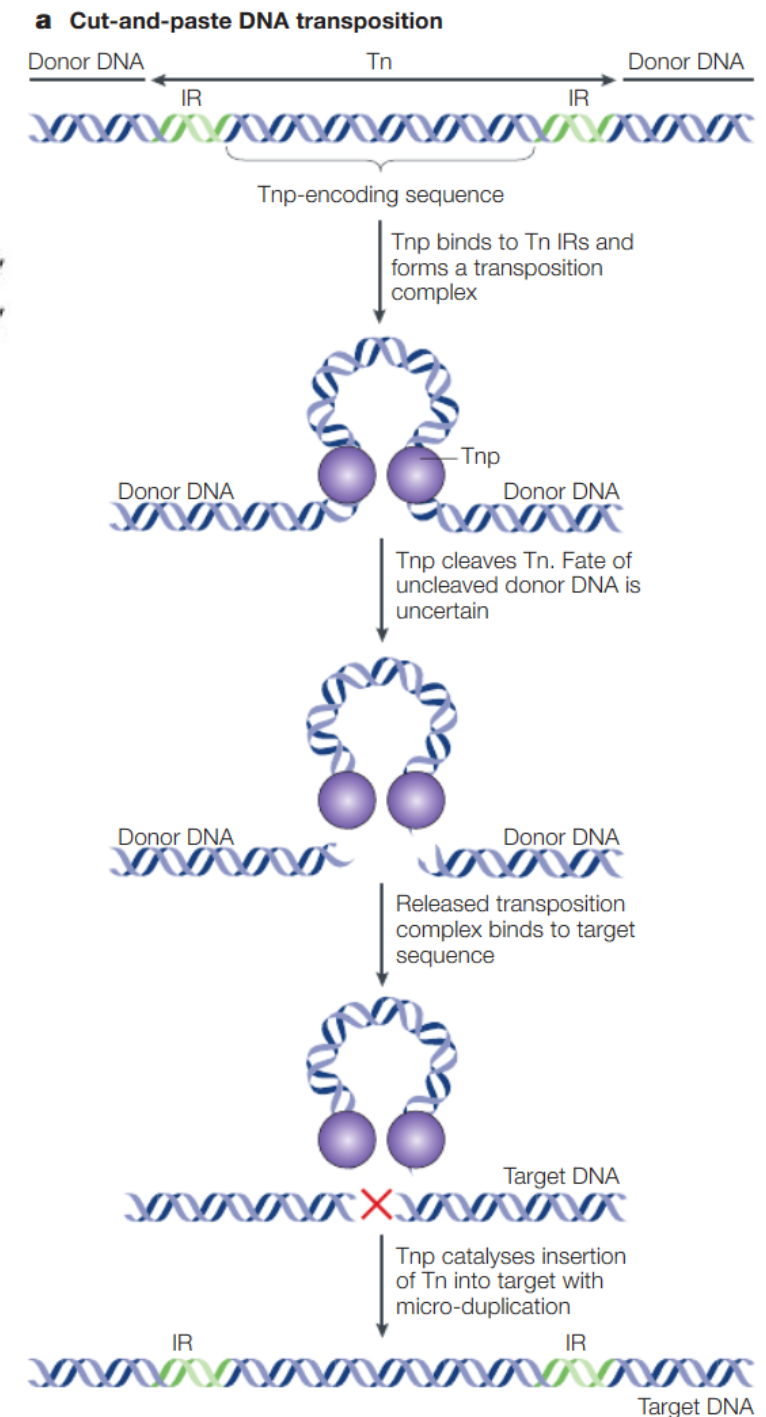
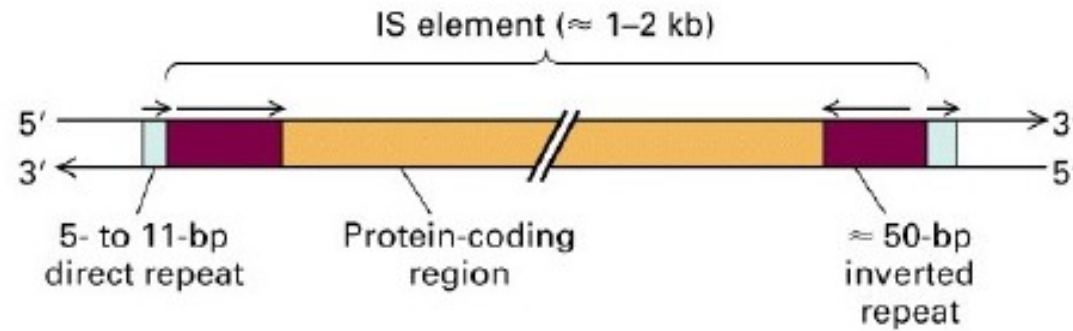


## Složené transpozony

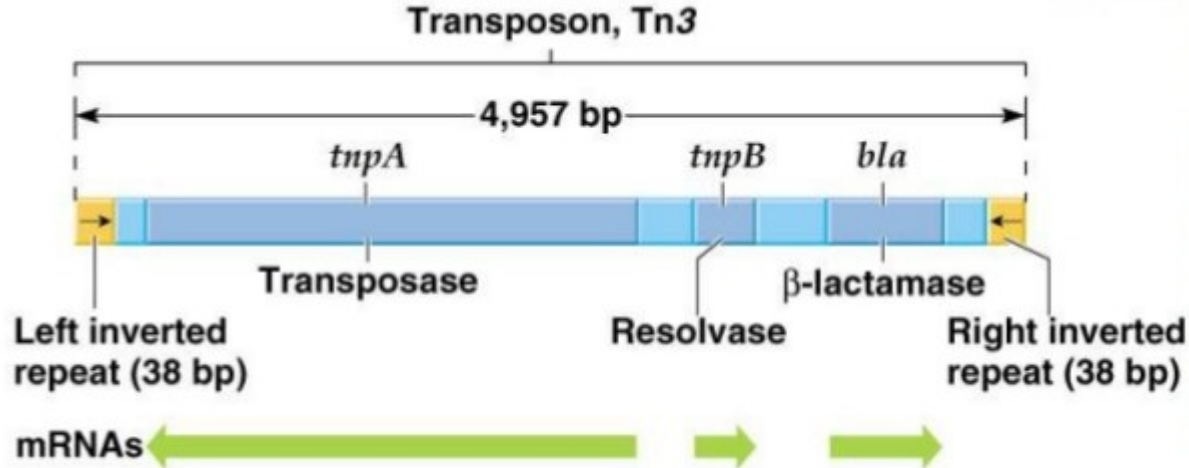


# IS-elementy

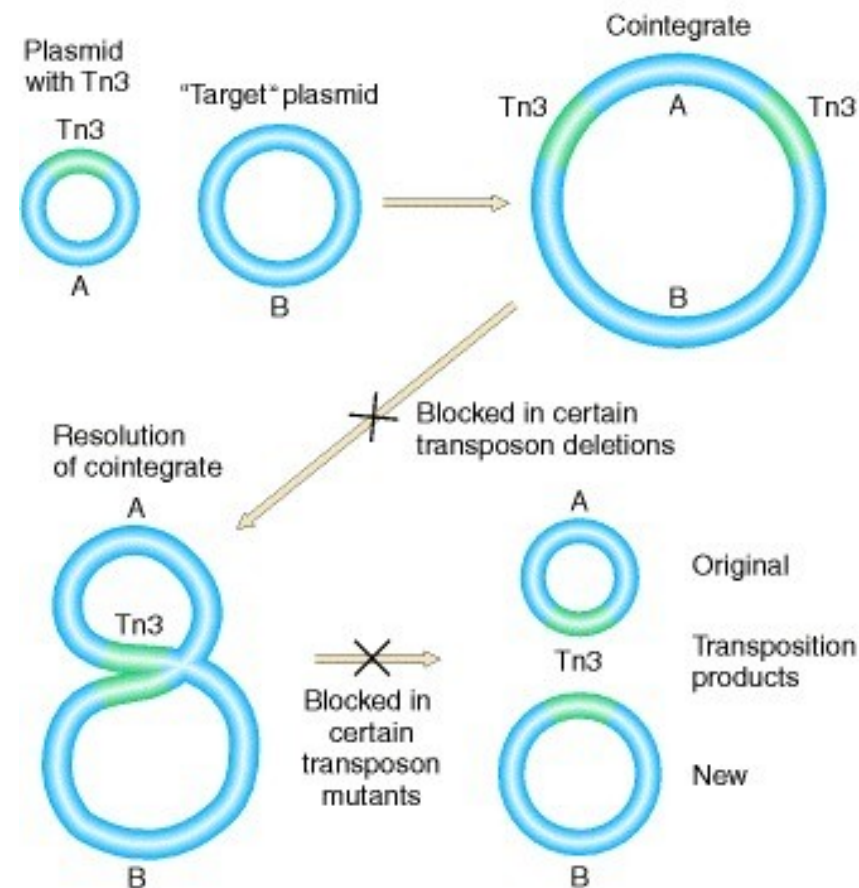
- obvykle méně než 2500 bp
- ohraničeny krátkými identickými sekvencemi – **obrácenými koncovými repeticemi**
- mutace v koncových repeticích eliminuje přemísťovací schopnost
- obsahují pouze geny pro zajištění a regulaci transpozice
- kódují **transpozázu**: váže se ke koncům elementu, kde štěpí oba řetězce DNA – tím je element uvolněn z původního místa



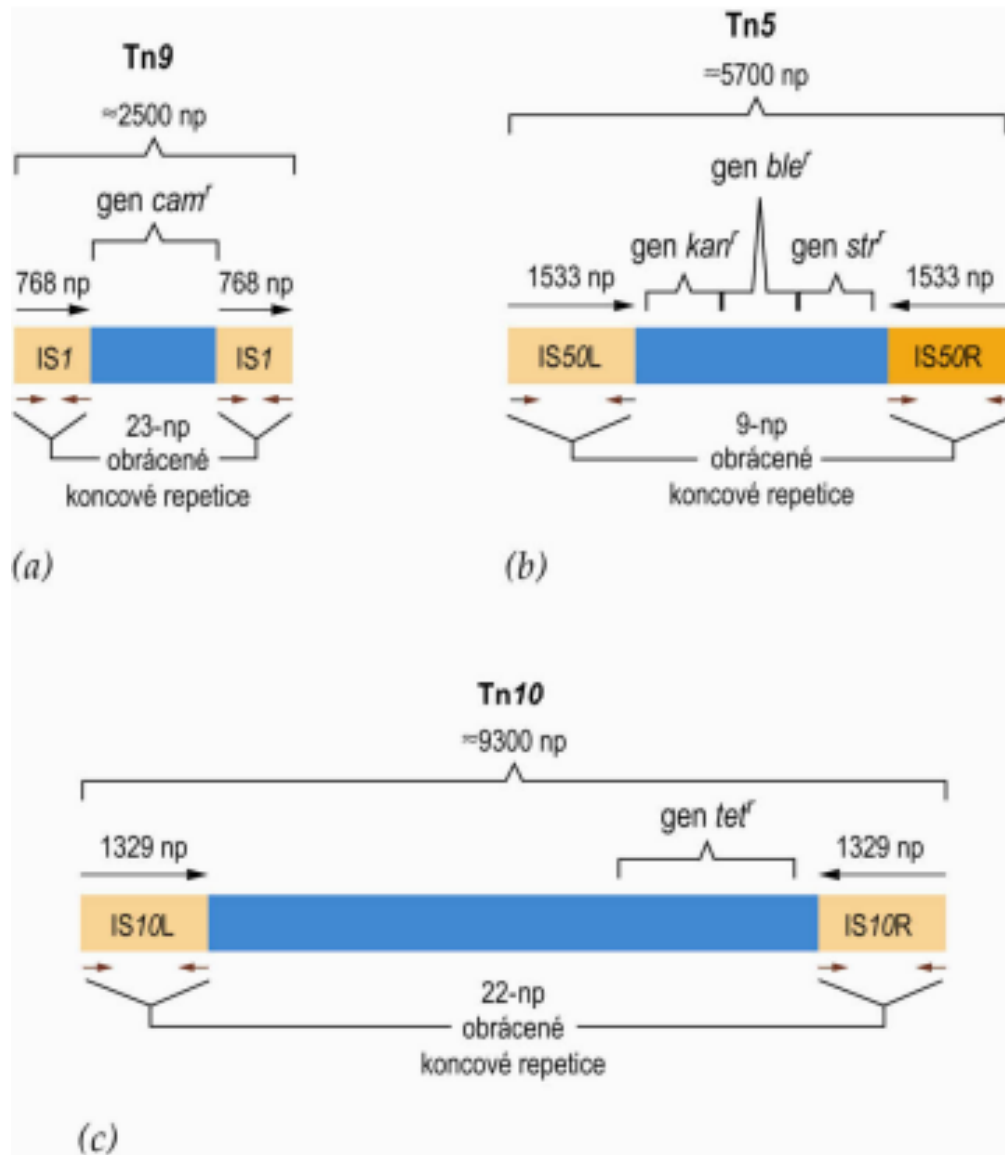
# Transpozon Tn3



- obsahuje geny, které nejsou pro transpozici nezbytné
- na koncích jsou jednoduché obrácené repetice
- replikativní typ transpozonu, dočasně spojuje molekuly za vzniku kointegrátu, po rozložení každá molekula obsahuje 1xTn3
- v cílovém místě nastává duplikace
- Struktura:
  - gen pro transpozázu a resolvázu (reprezor transpozázy)
  - gen pro beta-laktamázu (rezistence k Amp)



# Složené transpozony bakterií



## Rezistence k antibiotikům:

Kanamycin

Gentamycin

Ampicilin

Tetracyklin

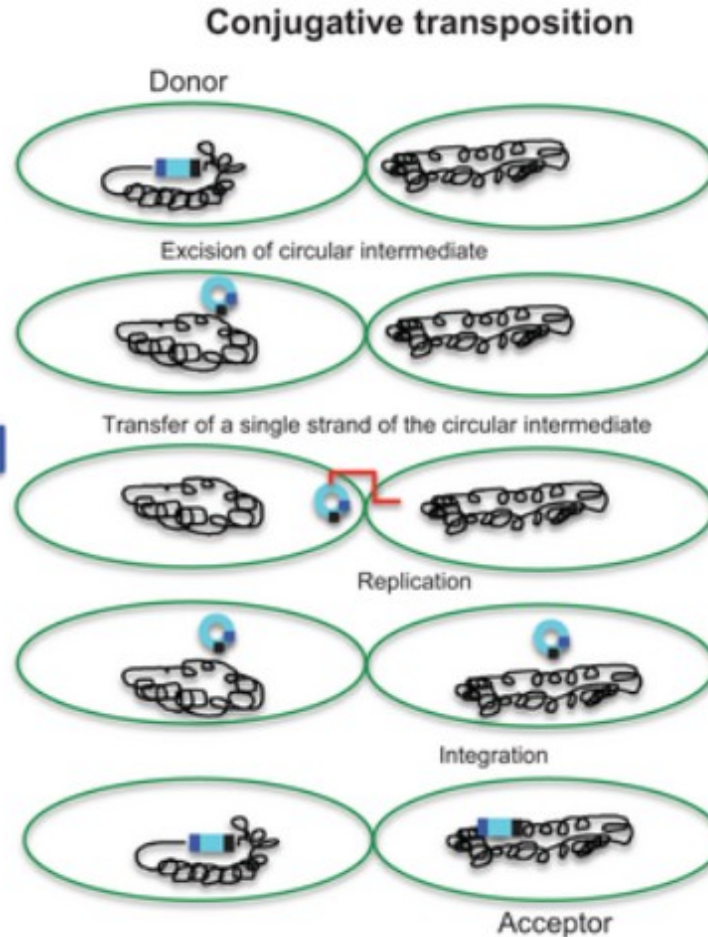
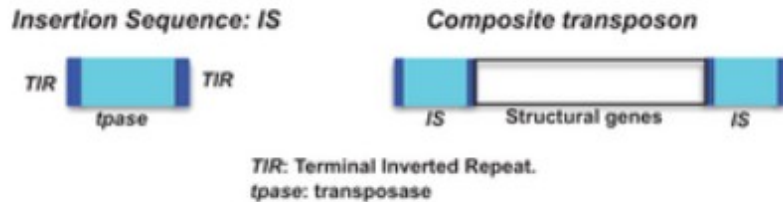
Chloramfenikol

Streptomycin

# Horizontální genový přenos

## Horizontal gene transfer (HGT)

Přenos genů mezi bakteriemi při konjugaci  
Zprostředkován IS elementy



## Význam HGT zprostředkovaného transpozony v lékařství:

- Transpozony často obsahují geny pro rezistenci k antibiotikům
- tím se tyto geny mohou snadno šířit a zvyšovat tak odolnost patogenních bakterií k antibiotikům
- dnes je obtížné léčit řadu infekčních onemocnění (úplavice, kapavka, tuberkulóza, atd.)
- šíření rezistence umožnilo široké používání antibiotik (selekce rezistentů)
- transpozóny (přenos mezi molekulami DNA uvnitř bakteriální buňky) a konjugativní plazmidy (přenos mezi různými bakteriálními kmeny)

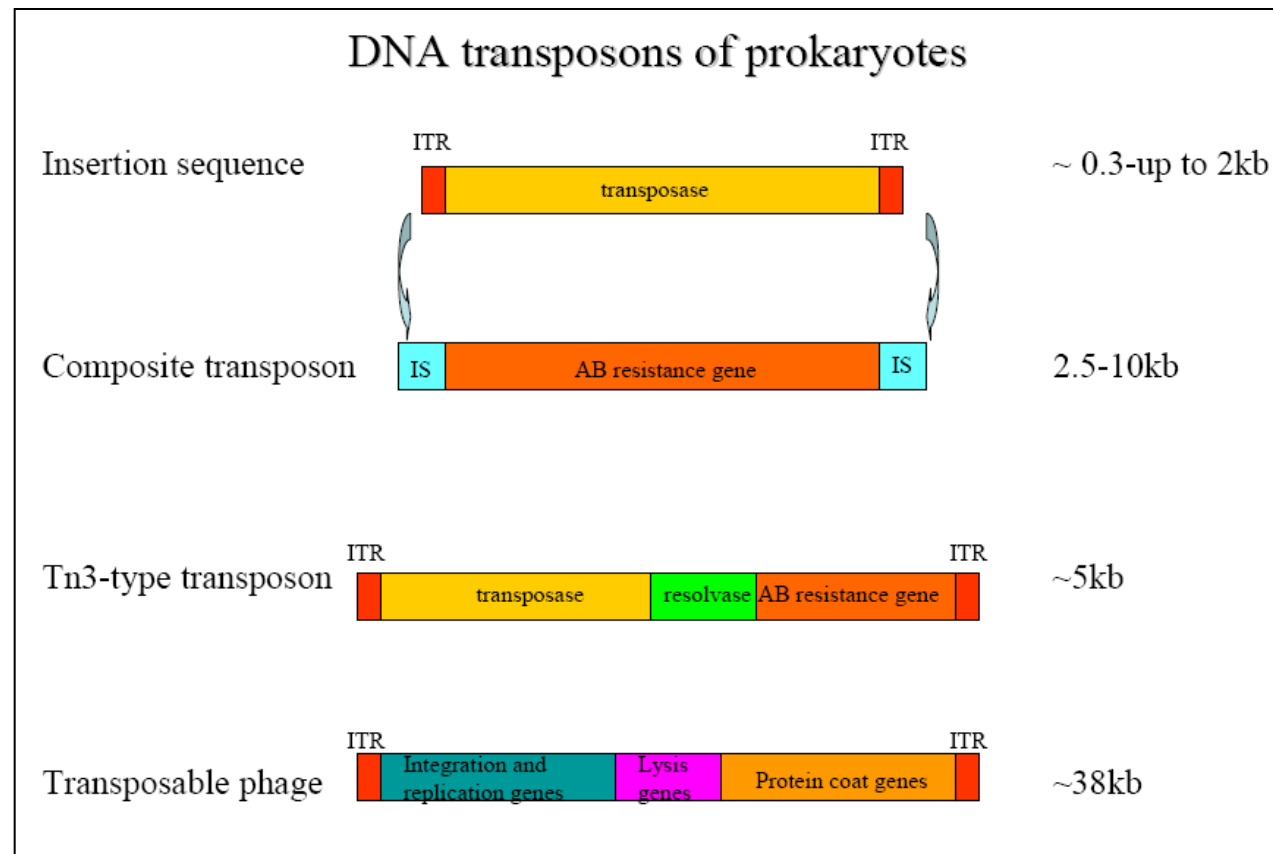
# Eukaryotické DNA transpozony



# DNA transposony

## 1. Prokaryotické:

- IS elementy
- Tn elementy



## 2. Eukaryotické:

- Ac a Ds elementy – autonomní a neautonomní u kukuřice
- P elementy – hybridní dysgeneze u drozofily (samci P+ a samice P-)
- Tc1/mariner – u *C. elegans* (Tc1) a drozofily (Mariner)

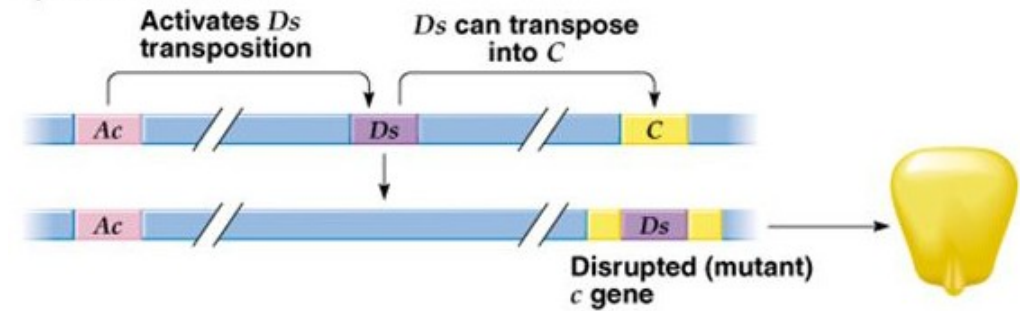
# Ac/Ds transpozony u kukuřice (*hAT* superfamily)



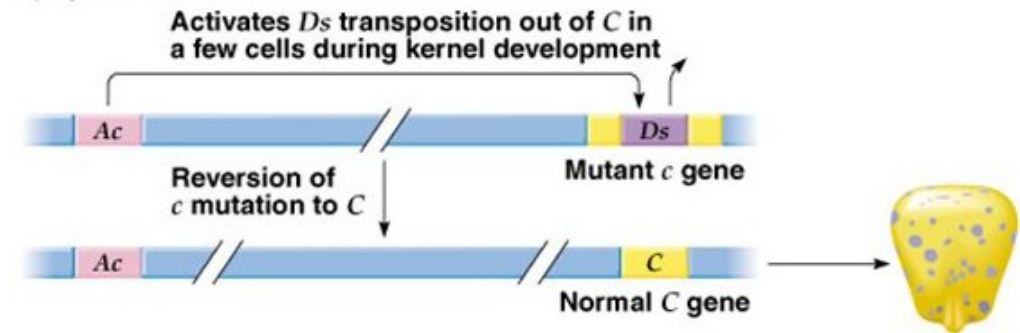
a) Purple kernels



b) Colorless kernels



c) Spotted kernels



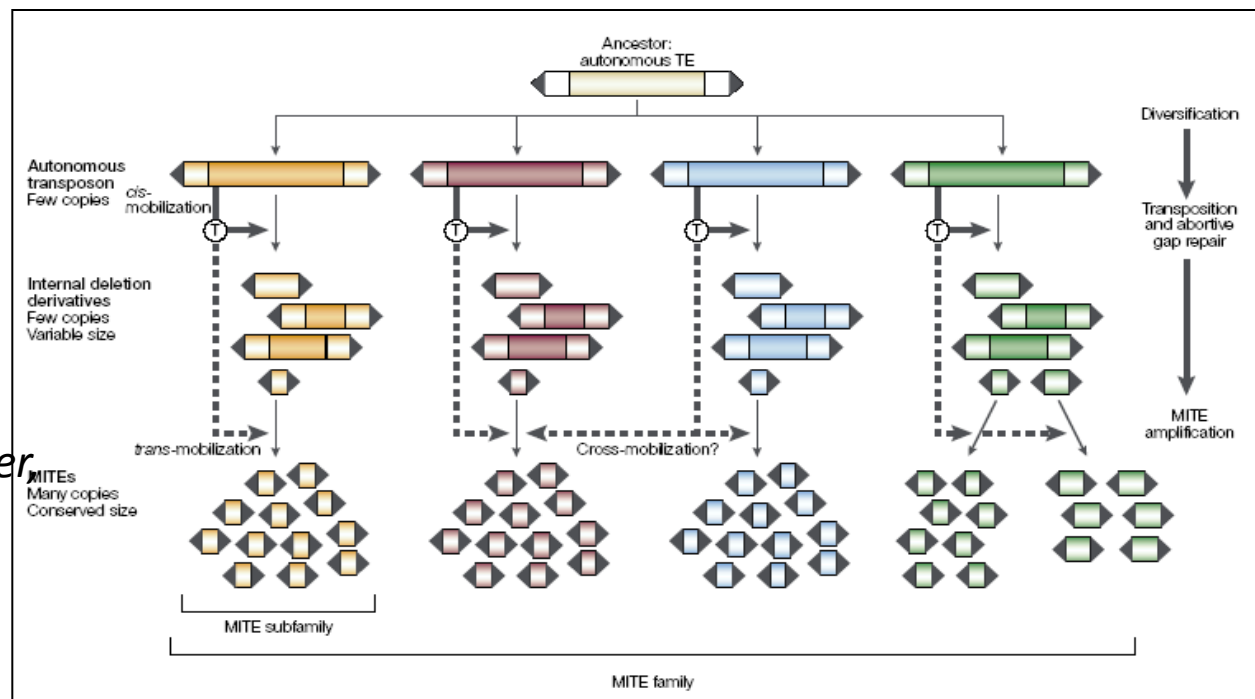


# MITE elementy

- rostlinné ekvivalenty lidských Alu, délka: 125 - 500bp
- neautonomní elementy (master = DNA TE Mariner)
- obrácené koncové repetice (TIR) - konzervativní 10-15 bp
- AT-bohaté (~72% *Stowaway*)
- tvoří sekundární struktury DNA (hairpins)
- preference cílového místa - TA(A)
- asociace s geny – v intronech, poblíž 5' nebo 3' konců genů

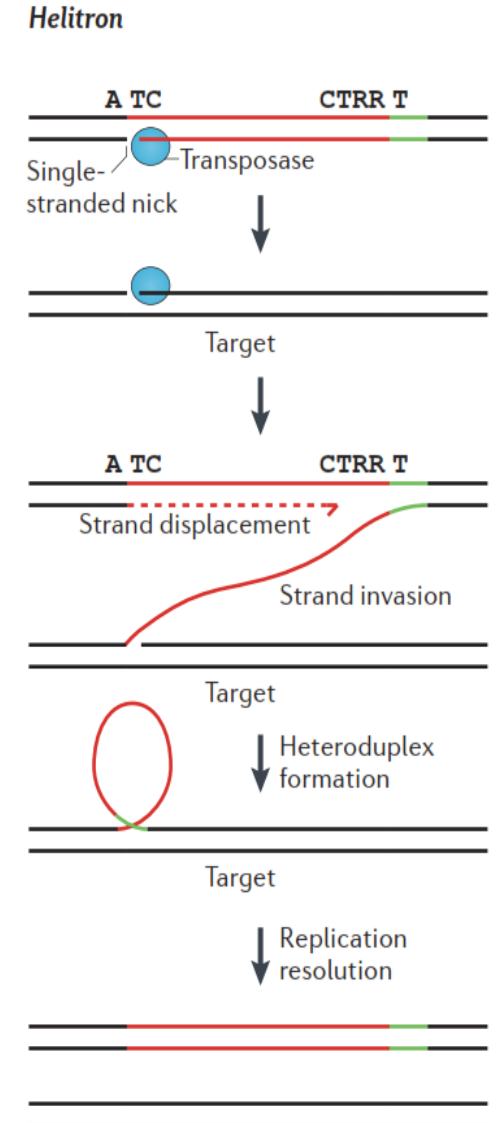
rodiny:

- *Stowaway* (jednoděložné, dvouděložné, živočichové)
- *Tourist* (trávy)
- *Emigrant, Alien, Heartbreaker, Bigfoot, ...* (rostliny)

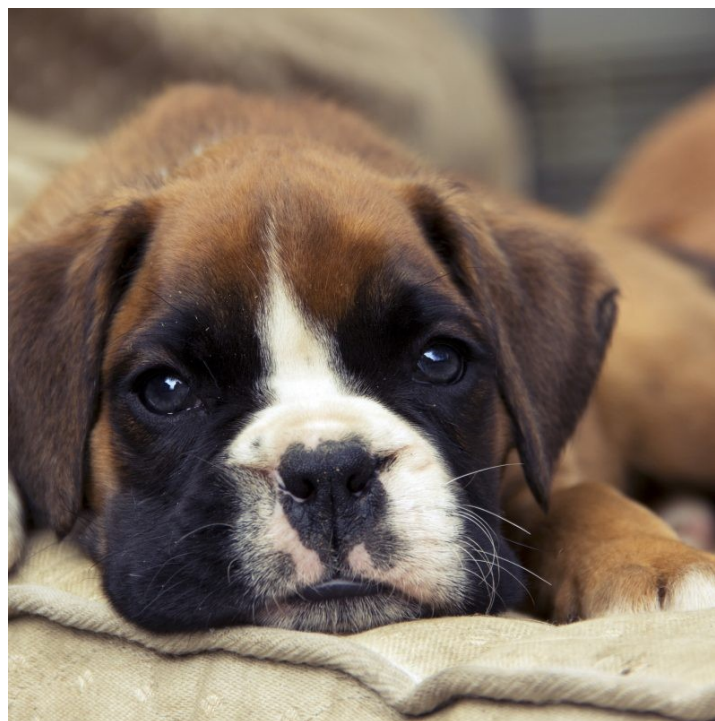


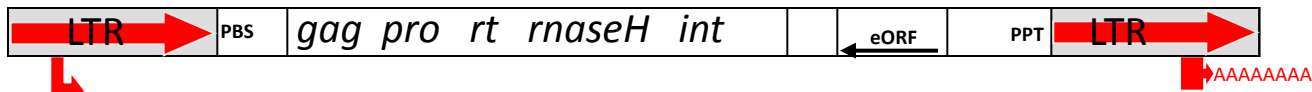
# Helitrony – nedávno objevené transposony využívající mechanismus otáčející se kružnice

- replikace mechanismem otáčivé kružnice (jako plazmidy, ss fágy, geminiviry)
- výskyt u eukaryot - 2% genomu *A.thaliana*, *C. elegans*, také v *Oryza sativa*
- nedělají duplikaci cílové sekvence (TSD)
- cílená inserce do AT dinukleotidu
- začínají 3'-AT a končí CTRR-5', nemají TIR
- konzervace palindromu před 3'CTRR (sekvence není konzervativní)
- většina elementů je neautonomní (0.5-3kb), vzácné dlouhé Helitrony (5.5-15kb)
- kódující proteiny pro RC replikaci: helikázu (HEL), nukleázu/ligázu a proteiny vážící jednořetězcovou DNA (RPA)
- mechanismus tvorby neautonomních elementů nejasný
- Helitrony jako evoluční spojovník mezi prokaryotickými RC elementy a geminiviry (potomci geminivirů integrovaných do genomů časných eukaryot)

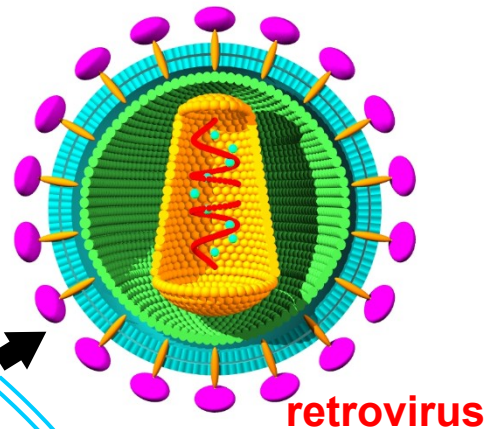


# **Třída I. Retrotranspozony (Retroelementy)**





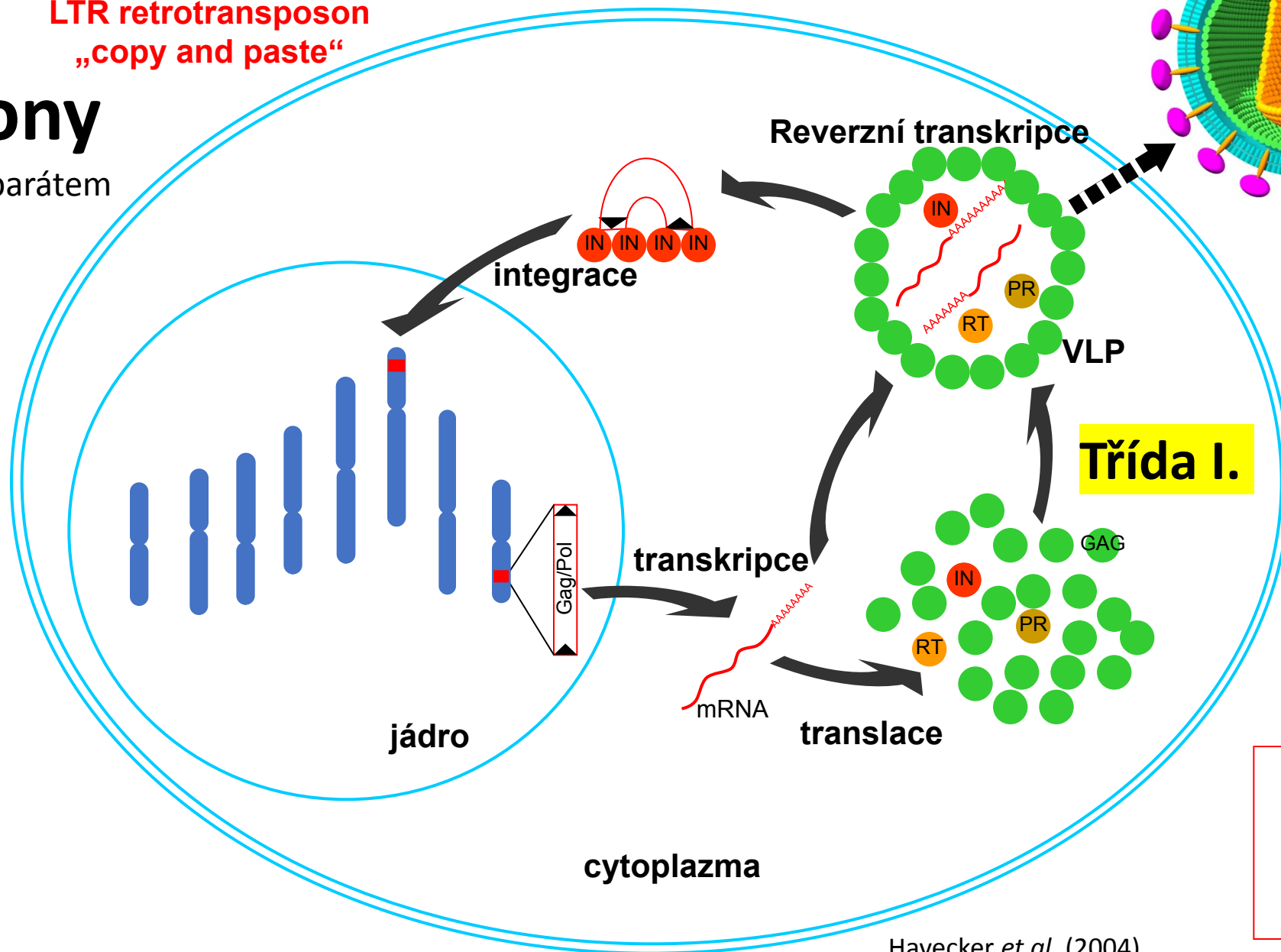
**LTR retrotransposon**  
„copy and paste“



**retrovirus**

# Transpozony

(disponují vlastním aparátem pro transpozici)

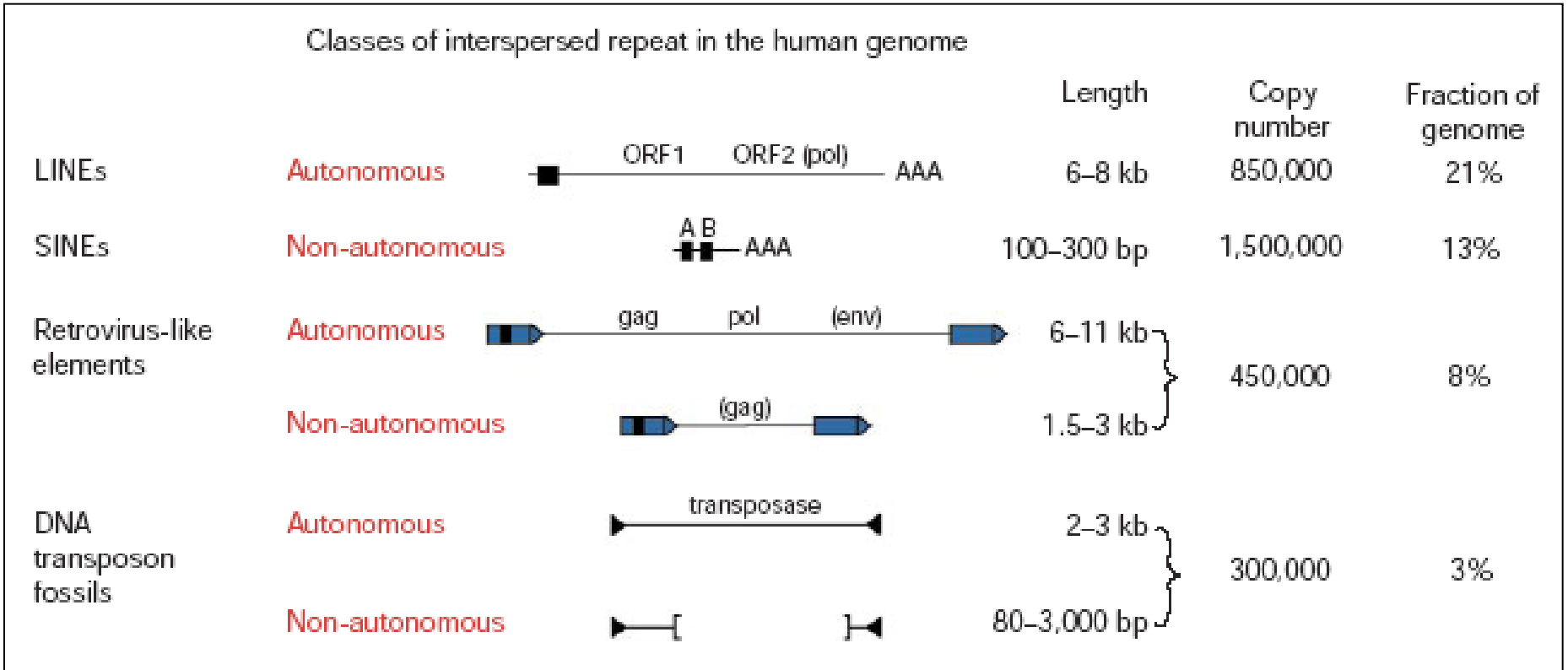


**Třída I.**

**Autonomní**  
**X**  
**Neautonomní**

Havecker *et al.* (2004)

# Téměř polovinu lidského genomu tvoří mobilní elementy



20 až > 1 500 000 kopií

Eukaryotické genomy: geny plovoucí po moři retrotransposonů  
 Endogenní retroviry – otisky dávných infekcí primátů

# Co jsou retroelementy?

- **Retroelementy** = sekvence DNA nebo RNA obsahující gen pro enzym reverzní transkriptázu (katalyzuje přepis RNA do DNA).

Širší definice zahrnuje veškeré sekvence vzniklé reverzní transkripcí RNA do DNA.

Po genomech se šíří procesem retropozice.

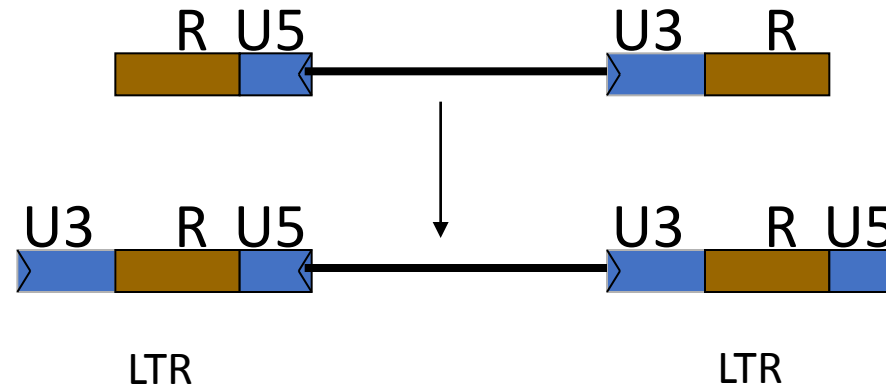
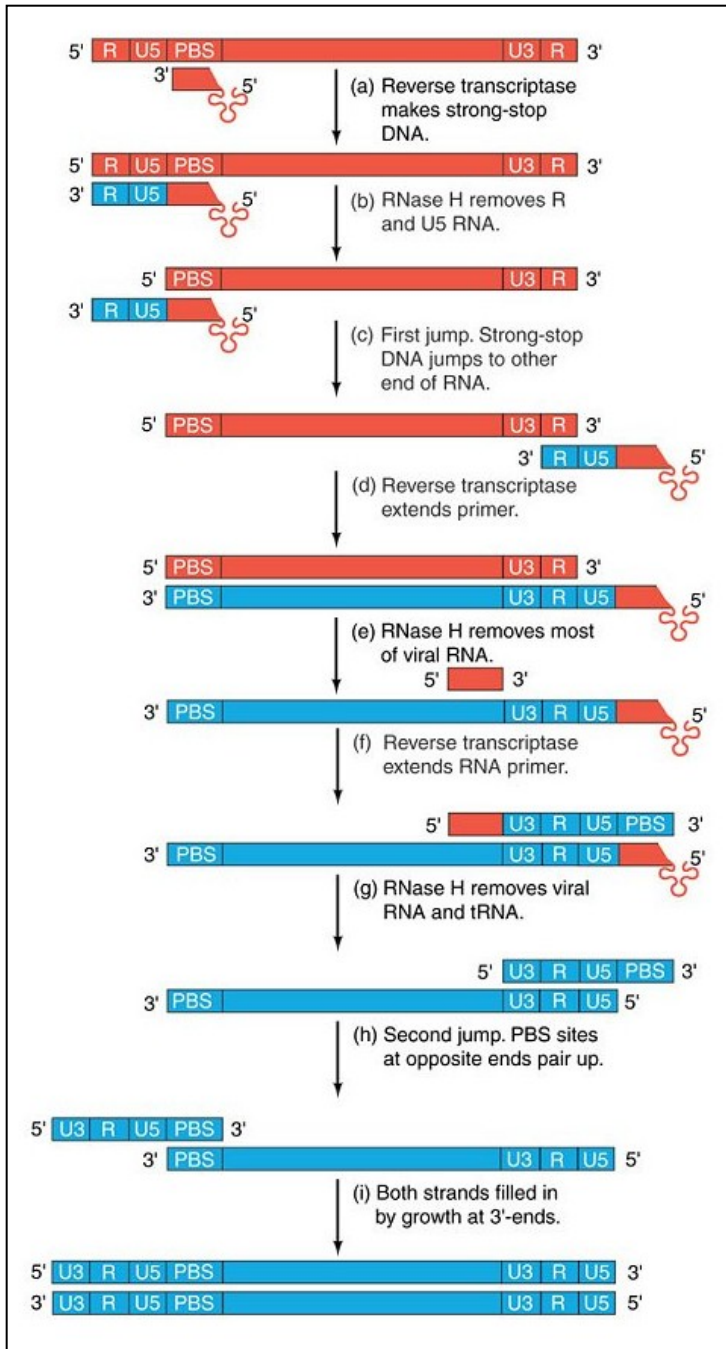
- **Retropozice** = přesun genetického materiálu z jednoho místa v genomu do místa jiného prostřednictvím RNA intermediátu. Má vždy duplikativní charakter.

Jsou retroelementy relikty světa RNA?

Podporuje to jejich struktura, konzervativní mechanismus replikace a všudypřítomnost.

# Reverzní transkripce

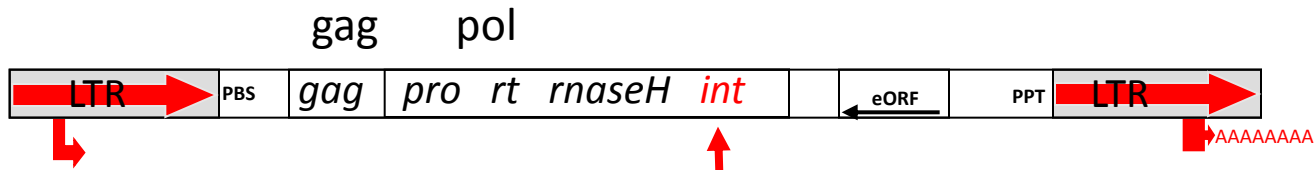
- nasednutí primeru tRNA a extenze
- odstranění RNA oblasti R a U5
- první přeskok a extenze
- odstranění většiny RNA RNázouH
- zbyde primer a extenze
- odstranění zbytku virové RNA a tRNA
- druhý přeskok a dosyntetizování



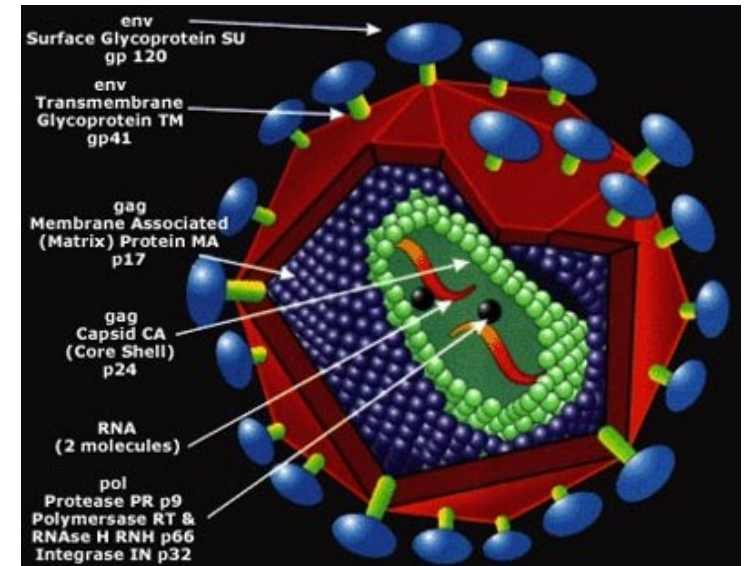
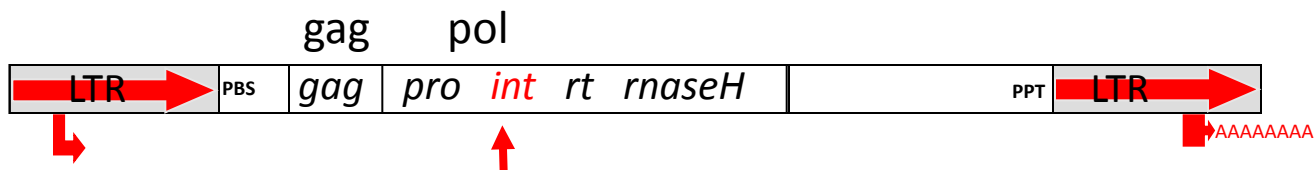
# LTR retrotranspozony

- Dominantní skupina transpozonů u rostlin, u živočichů obvykle méně kopií
- Desítky až stovky podrodin v rámci každé rodiny (jménem Athila, Ogre, Angela, chromovirus...)
- Obvykle jen několik podrodin je u daného druhu extrémně amplifikováno do počtů o tisících až desetitisících kopií a tvoří dohromady až 90 % délky genomu
- Ty3/gypsy LTR retrotranspozony jsou předkové retrovirů

## Typ Ty3/gypsy



## Typ Ty1/copia

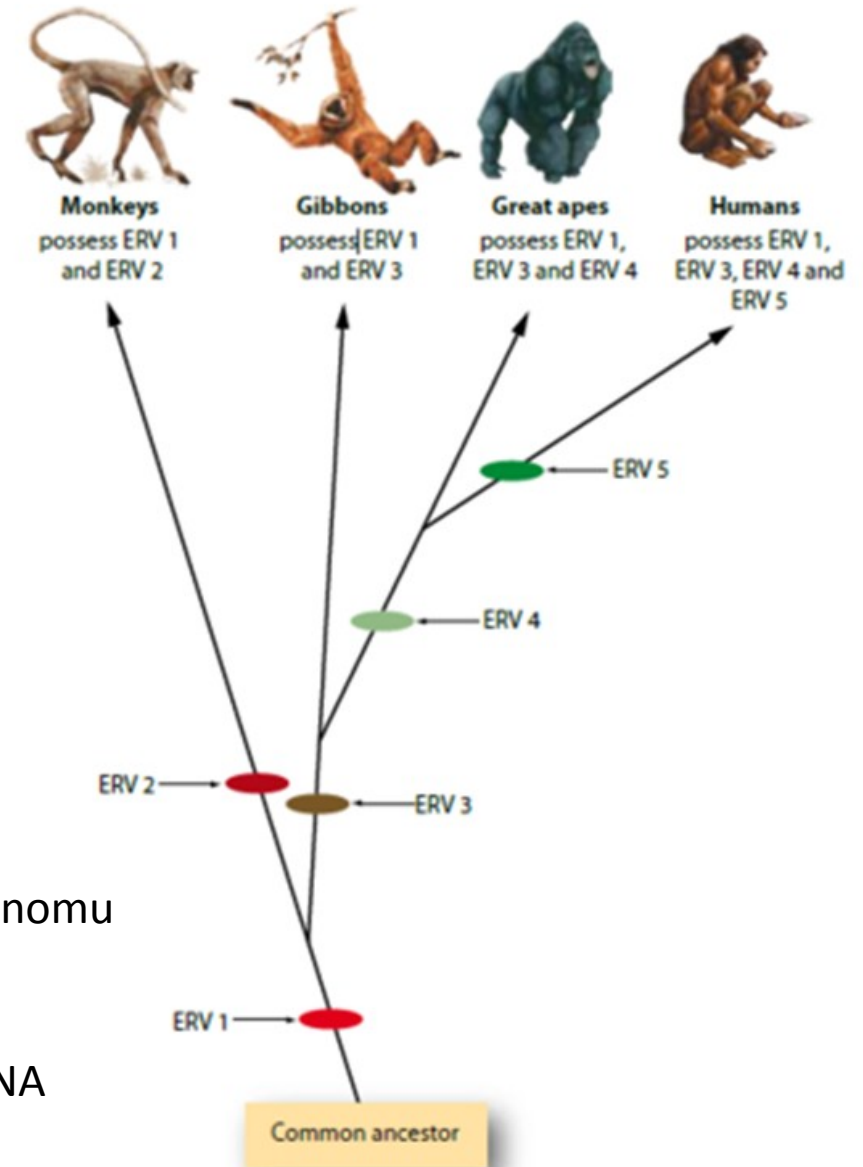
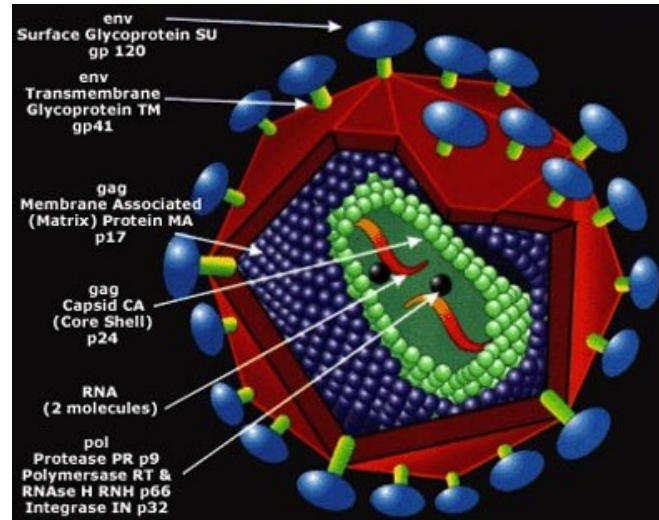
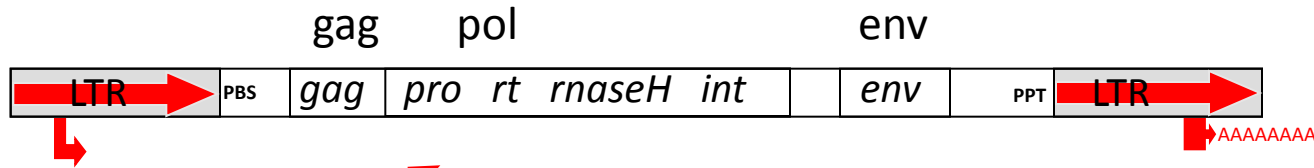


- gag, pol, env ..... geny
- LTR ... dlouhé koncové repetice
- PBS .... místo vazby primeru
- PR ..... doména kódující proteázu
- INT .... doména kódující integrázu
- RT ..... doména kódující reverzní transkriptázu
- PPT ..... polypurinový úsek
- > ..... přímé repetice



# Retroviry

## Endogenní retroviry

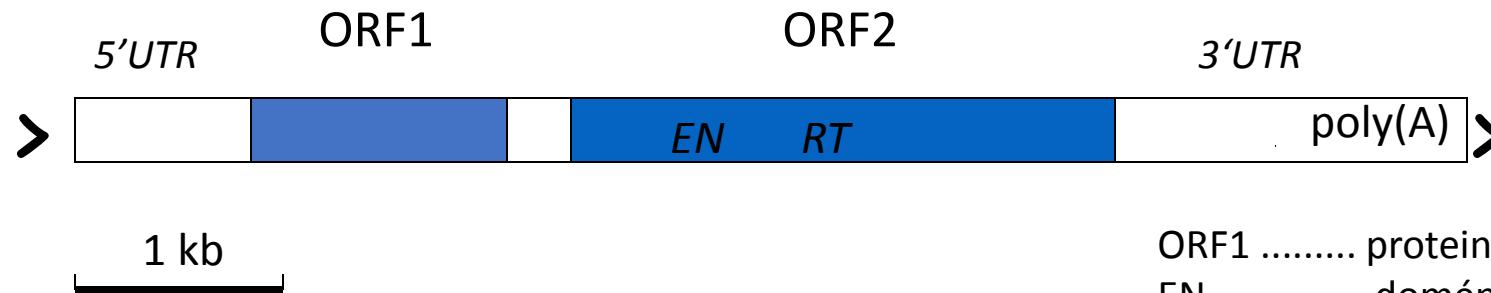


- Původem retroviry, které kolonizovaly genom a přestaly vytvářet infekční částice
- U člověka známé jako Human Endogenous Retroviruses (H-ERVs), tvoří cca 8 % genomu
- Opakované infekce během vývoje savců i během evoluce primátů
- Env – několikrát „domestikován“ – vznik syncitinů v placentě
- Patrně přispívají k imunitě proti retrovirům (například prostřednictvím tvorby siRNA molekul, které v buňce ničí retrovirovou RNA)

# Non-LTR retrotranspozony (Retropozony)

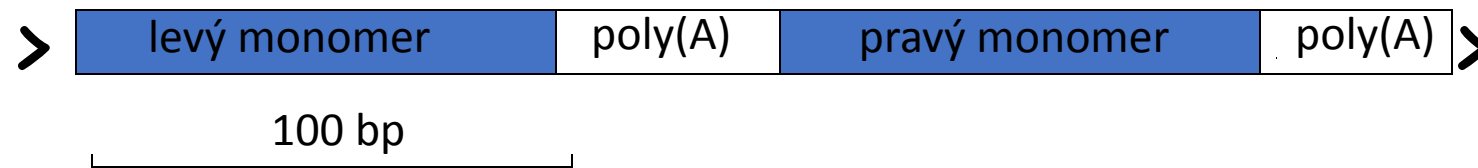
- Velmi rozšířená skupina transpozonů u živočichů, u rostlin obvykle méně kopií
- U člověka dominantní typ transpozonů, L1 – 21 % genomu, Alu – 13 % genomu
- Mnoho různých podrodin, ale vysokého počtu kopií dosahují jen některé z nich
- Hrají roli v kompenzaci dávky genů (Inaktivace X) a vzniku regulačních genových sítí

## LINE (L1):



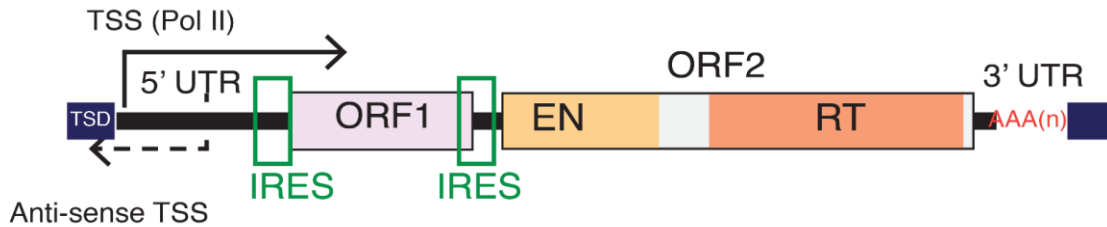
ORF1 ..... protein vážící RNA  
EN ..... doména kódující endonukleázu  
RT ..... doména kódující reverzní transkriptázu  
5'UTR ..... netranslatovaná oblast na 5' konci  
3'UTR ..... netranslatovaná oblast na 3' konci  
poly(A) ..... polyadeninový úsek  
> ..... přímé repetice

## SINE (sekvence Alu):



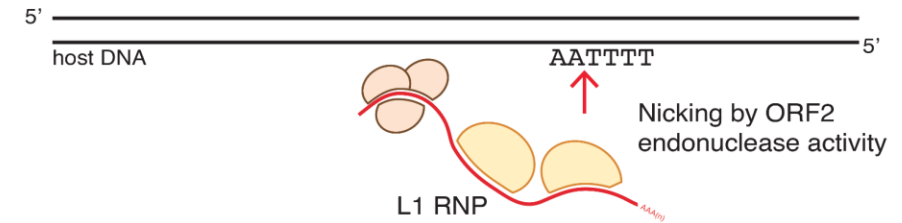
# L1 elementy (LINE)

## mouse LINE1 (~ 6 kb)

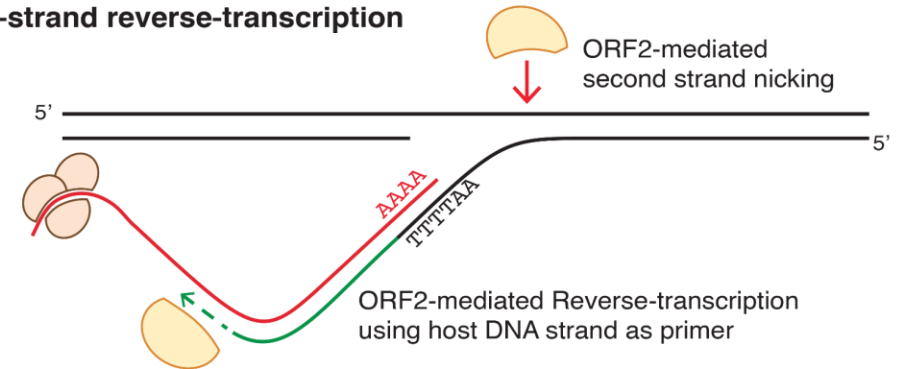


- Autonomní LINE elementy savců, dominují lidskému genomu – 21 %, celkem 850 tisíc kopií
- Člověk: aktivní L1 (516 tis), neaktivní L2 (315 tis) a L3 (37 tis)
- **TPRT** mechanismus mobilizace (Target Primed Reverse Transcription)
- Hrají roli v kompenzaci dávky genů (Inaktivace X) a vzniku regulačních genových sítí
- Průměrný lidský genom obsahuje asi 100 aktivních kopií, vysoká variabilita, někteří jedinci mají daleko vyšší aktivitu L1 – možný vznik chorob inzerční mutagenézí
- Zvýšená aktivita v mozku lidí se schizofrenií (příčina nebo důsledek?)
- Místa homologie pro genovou konverzi a rekombinaci – evoluce
- Inzerce do zlomů chromozomů – reparace DNA

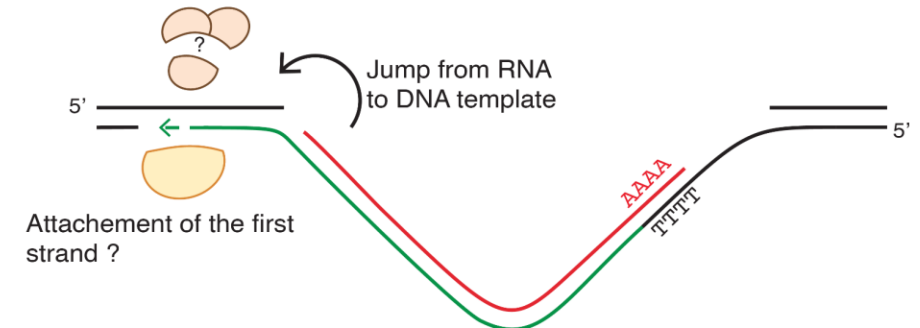
## First-strand cleavage



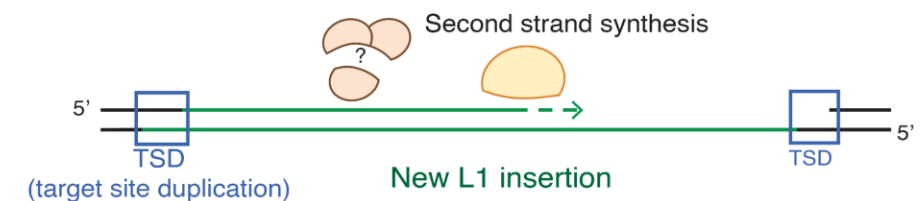
## First-strand reverse-transcription



## Template jump

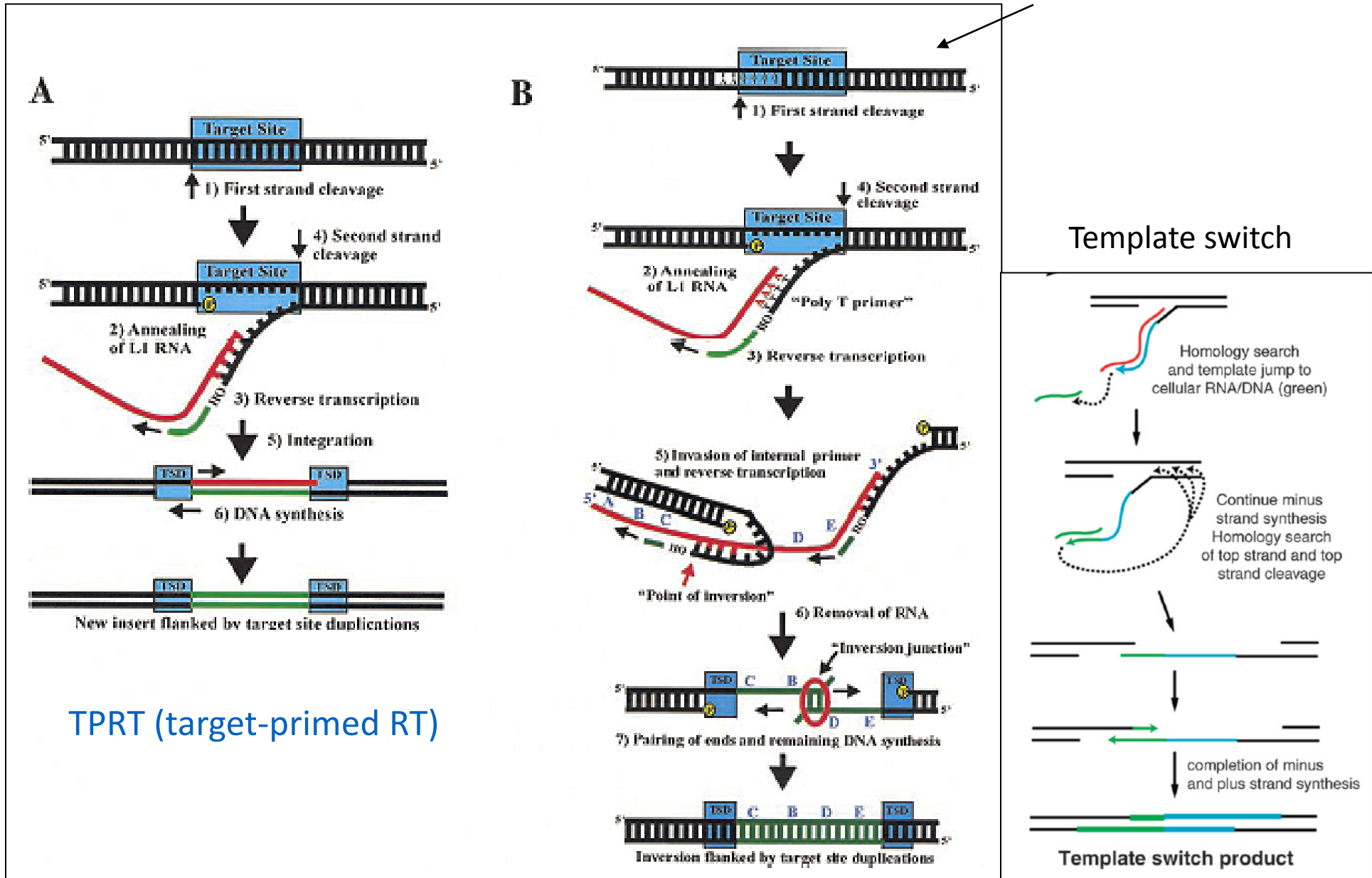


## Second-strand reverse-transcription



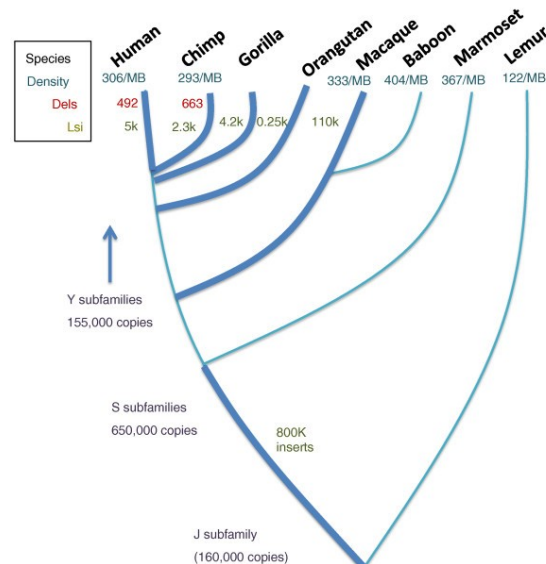
# Včleňování LINE elementů

Twin priming:  
mechanismus tvorby inverzí



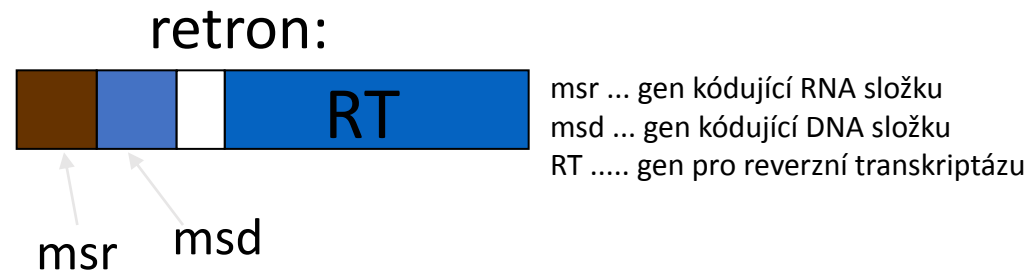
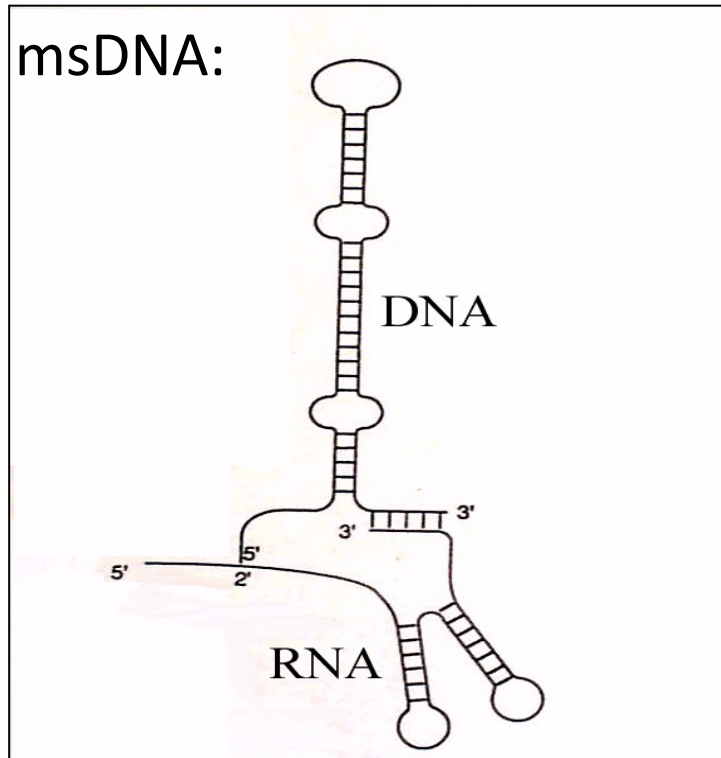
# Alu elementy (SINE)

- Odvozeny od 7SL RNA genu kódujícího podjednotku signální rozpoznávací částice (přenos proteinů přes membrány a začleňování do membrán)
- délka 300 bp, u člověka 1-1,5 milionu kopií
- Vznik před oddělením linie primátů a hlodavců (70 milionů let)
- Závislé na mobilizačním aparátu L1 elementů
- Trvalá kumulace kopií v genomu pravděpodobně kvůli malé délce Alu – nízká pravděpodobnost delece genovou konverzí
- Časté včleňování blízko nebo do genů – ovlivnění exprese přilehlých genů (alu obsahuje vazebná místa pro transkripční faktory) a vznik míst pro alternativní sestřih, předčasné ukončení transkripce...
- Hrají roli při vzniku mnoha genetických chorob a rakoviny



Locus	Chromosome	Subfamily	Disease
3 × <i>HEMB</i> (IX)	X	Ya5, Ya5, Yb8	Hemophilia B
2 × <i>HEMA</i> (VIII)	X	Yb8, Yb9	Hemophilia A
2 × <i>CLCN5</i>	X	Ya6, Ya5	Dent's disease
2 × <i>BTK</i>	X	Y, Y	X-linked agammaglobulinemia
<i>IL2RG</i>	X	Ya5	X-linked severe combined immunodeficiency disease
<i>GK</i>	X	Yc1	Glycerol kinase deficiency
<i>CD40LG</i>	X	Yb8	Hyper IgM syndrome
<i>ATP7A</i>	X	Ya5a2	Menkes disease
<i>CRB1</i>	1	Y	Retinitis pigmentosa
<i>ZFHX1B</i>	2	Ya5	Mowat-Wilson syndrome
<i>BCHE</i>	3	Yb8	Cholinesterase deficiency
<i>OPA1</i>	3	Yb8	Autosomal dominant optic atrophy
<i>CASR</i>	3	Ya4	Hypocalciuric hypercalcemia and hyperparathyroidism
<i>MLVI2</i>	5	Ya5	Associated with leukemia
<i>APC</i>	5	Yb8	Hereditary desmoid disease
<i>P5N1</i>	7	Ya5	Chronic hemolytic anemia
<i>EYA1</i>	8	Y	Branchio-oto-renal syndrome
<i>LPL</i>	8	Yb9	Lipoprotein lipase deficiency
<i>POMT1</i>	9	Ya5	Walker Warburg syndrome
3 × <i>FGFR2</i>	10	Ya5, Yb8, Yc1	Apert's syndrome
<i>TNFRSF6</i>	10	Yb8	Autoimmune lymphoproliferative syndrome
<i>C1NH</i>	11	Yc1	Complement deficiency
<i>AIP</i>	11	Ya5	Acute intermittent porphyria
<i>GNPTAB</i>	12	Y	Mucopolidosis
3 × <i>BRCA2</i>	13	Ya5, Yc1, Y	Breast cancer
<i>PMM2</i>	16	Yb8	Congenital disorder of glycosylation type I
<i>BRCA1</i>	17	Ya5	Breast cancer
15 × <i>NF1</i>	17	Y subfamilies	Neurofibromatosis

# Retrony - primitivní retroelementy bakterií



msDNA (=multicopy single-stranded)

- intermediáty extrachromosomálního replikačního cyklu  
nebo abortivní produkty intermediátů

## Prokaryotické retroelementy:

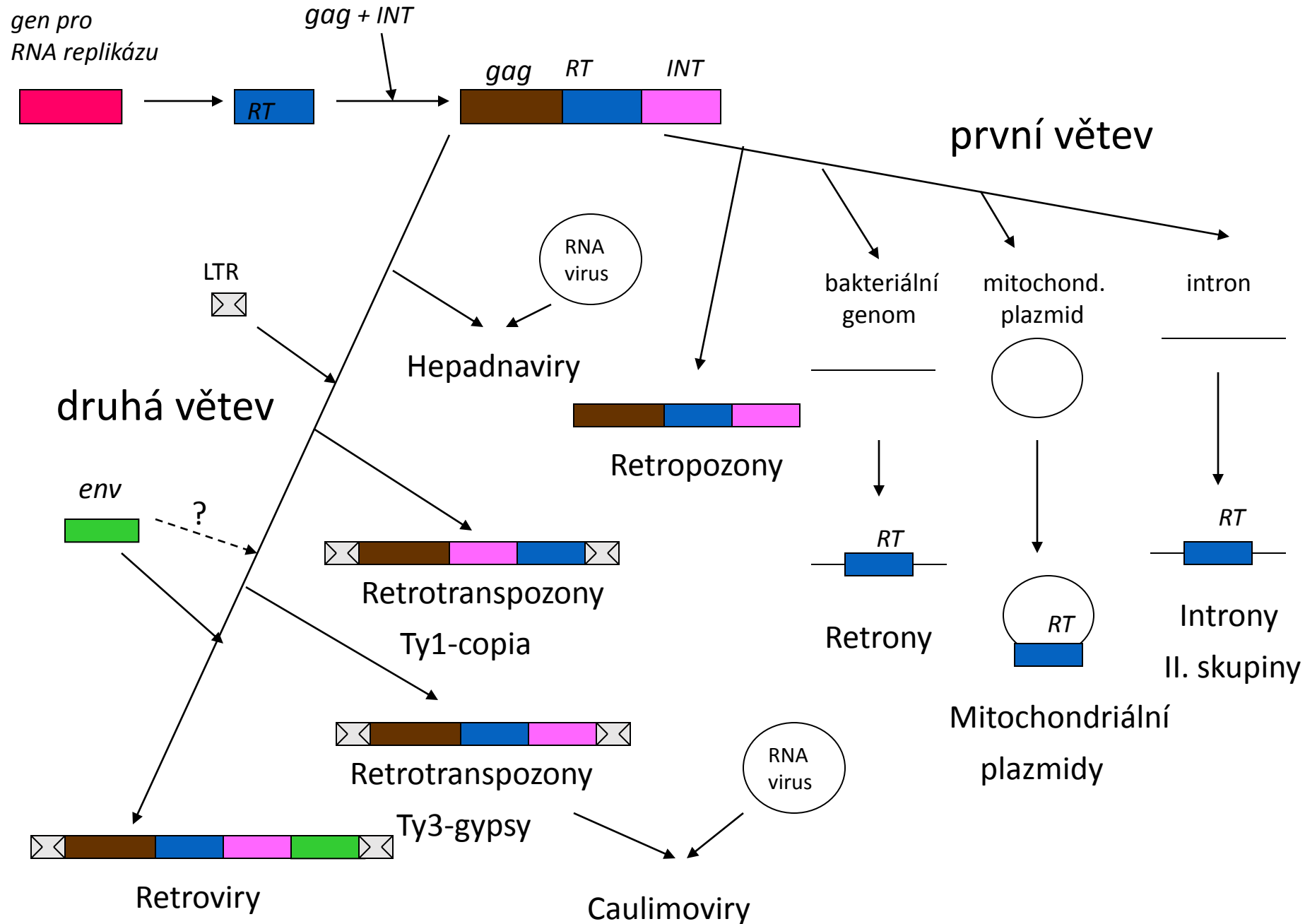
1. Kódují **jediný RT protein** s jedinou enzymatickou aktivitou (doménu).

**Ostatní aktivity** (proteáza, integráza, endonukleáza) retroelementy v průběhu evoluce získaly **od hostitele** (jako onkogeny retrovirů)

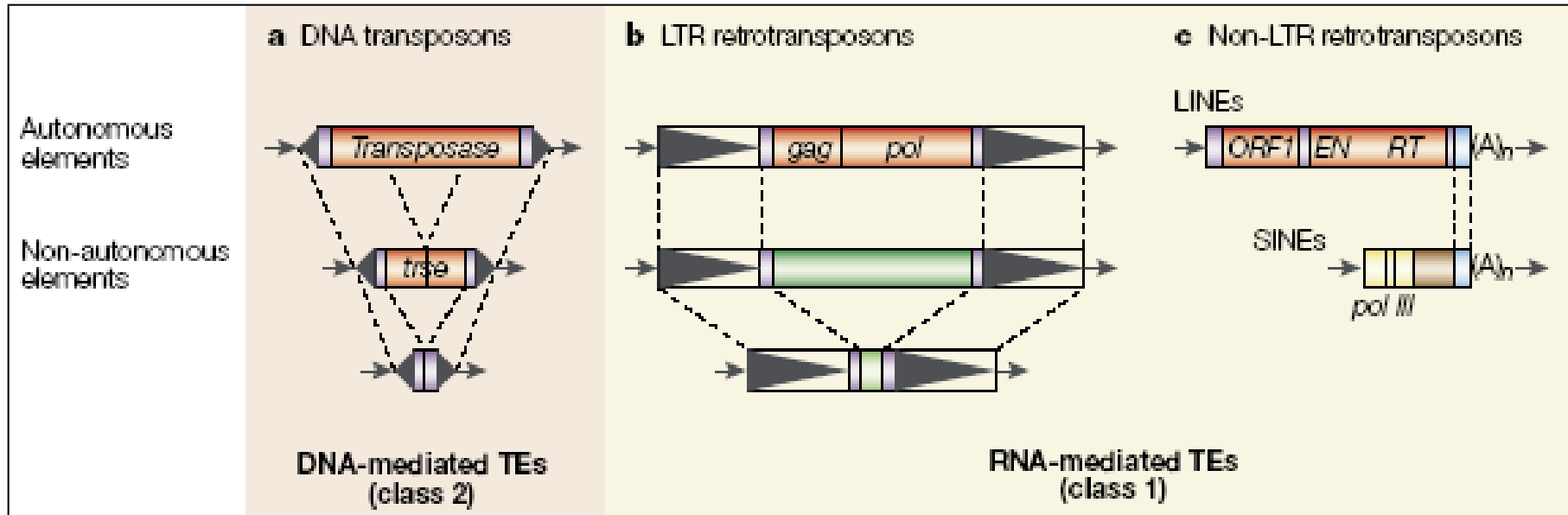
2. RT prokaryot provádí syntézu **nezávislou na primeru** podobně jako RNA polymeráza (předchůdce RT)

3. Prokaryotické RT jsou **podobné RdRP** (RNA-directed RNA polymerase), RT eukaryotických retroelementů jsou méně příbuzné

# Schéma evoluce retroelementů



# Paraziti parazitů: **Neautonomní** elementy tvoří většinu. Balancovaná rovnováha



Vybalancovaná rovnováha mezi autonomními (A) a neautonomními (N) elementy:

- přílišná úspěšnost neautonomních vede k záhubě
- titrace transpozázy neautonomními vede k omezením aktivity
- inhibice nadprodukcí
- **koevoluce A a N** řízena kompeticí o RT
- koevoluce endogenních a exogenních retrovirů – rezistence k infekci
- další mechanismy restrikce: metylace a heterochromatinizace