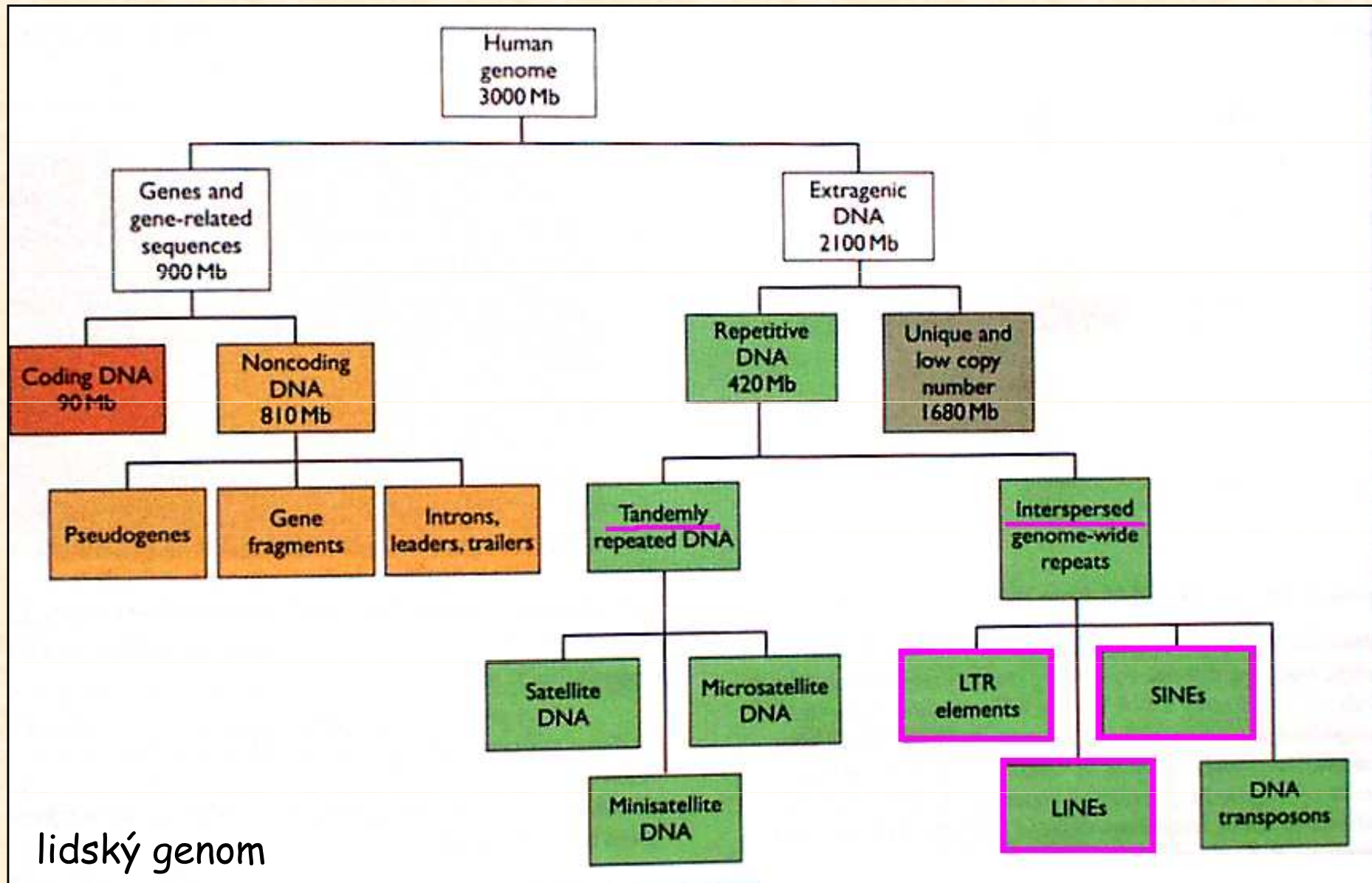








**Kapitola V.**  
**DYNAMIKA GENOMŮ**

# Podstatnou část genomů tvoří opakující se úseky DNA (**repetice**)



# Změna paradigmatu v genetice: Genomy jsou **dynamické**

Classes of interspersed repeat in the human genome

			Length	Copy number	Fraction of genome
LINEs	Autonomous		6–8 kb	850,000	21%
	Non-autonomous		100–300 bp		
Retrovirus-like elements	Autonomous		6–11 kb	450,000	8%
	Non-autonomous		1.5–3 kb		
DNA transposon fossils	Autonomous		2–3 kb	300,000	3%
	Non-autonomous		80–3,000 bp		

lidský genom

- Atributem genomů - inherentní nestabilita
- Barbara McClintock - 1948 (NC 1984)
- Stochastické ale regulované procesy

- Sobecké x altruistické (chování)
- Parazitické x mutualistické (role)
- Zbytečná DNA x geny (podstata)

# Reverzní transkriptáza v moderních genomech

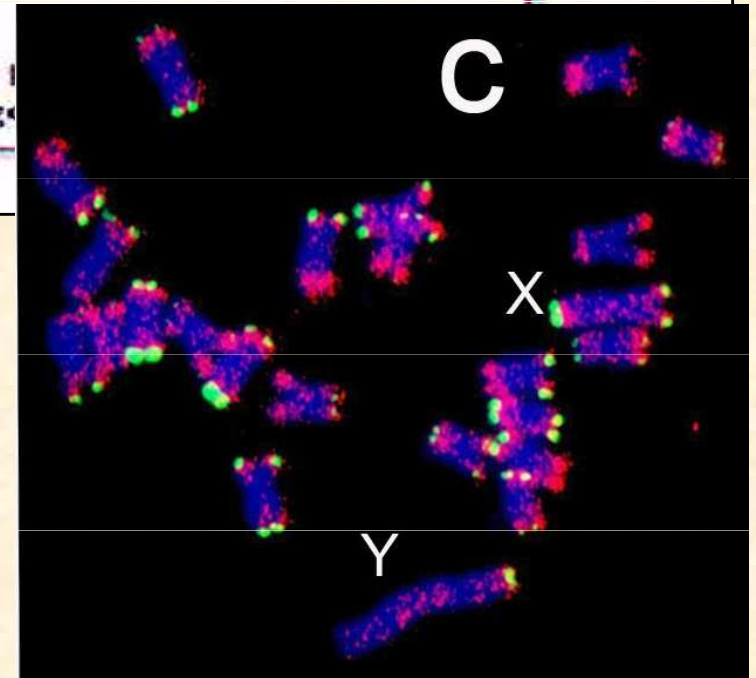
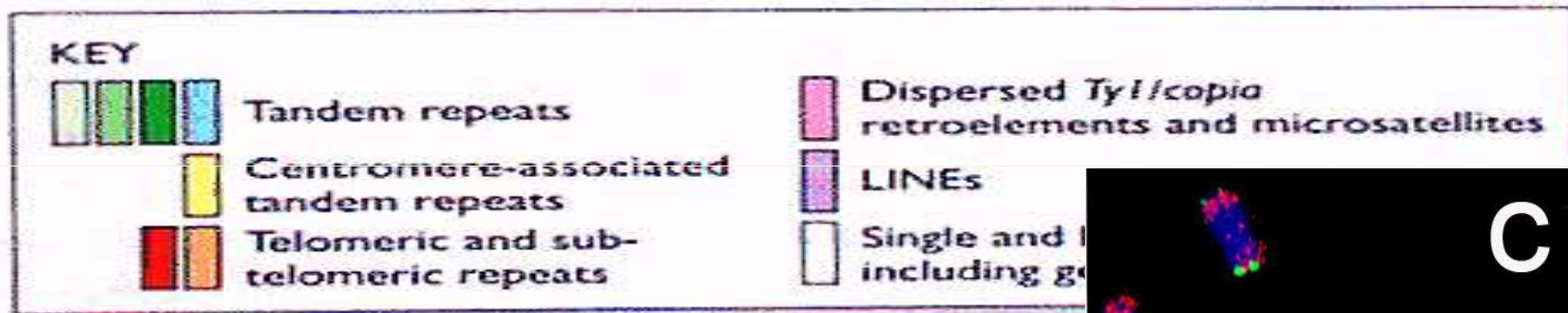
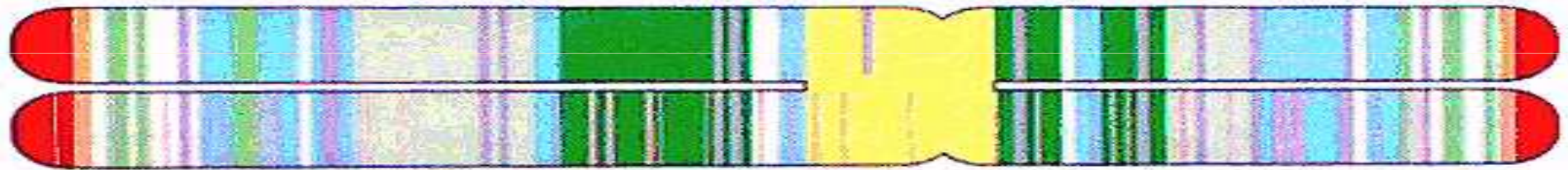
Téměř **polovinu** lidského genomu tvoří mobilní elementy

20 až > 1 000 000 kopií

Eukaryotické genomy: geny plovoucí po moři retrotransposonů (Bushman 2002)

	Number of copies (x 1,000)	Total number of bases in the draft genome sequence (Mb)	Fraction of the draft genome sequence (%)	Number of families (subfamilies)
SINEs	1,558	359.6	13.14	
Alu	1,090	290.1	10.60	1 (~20)
MIR	393	60.1	2.20	1 (1)
MIR3	75	9.3	0.34	1 (1)
LINEs	868	558.8	20.42	
LINE1	516	462.1	16.89	1 (~55)
LINE2	315	88.2	3.22	1 (2)
LINE3	37	8.4	0.31	1 (2)
LTR elements	443	227.0	8.29	
ERV-class I	112	79.2	2.89	72 (132)
ERV(K)-class II	8	8.5	0.31	10 (20)
ERV (L)-class III	83	39.5	1.44	21 (42)
MaLR	240	99.8	3.65	1 (31)
DNA elements	294	77.6	2.84	
hAT group				
MER1-Charlie	182	38.1	1.39	25 (50)
Zaphod	13	4.3	0.16	4 (10)
Tc-1 group				
MER2-Tigger	57	28.0	1.02	12 (28)
Tc2	4	0.9	0.03	1 (5)
Mariner	14	2.6	0.10	4 (5)
PiggyBac-like	2	0.5	0.02	10 (20)
Unclassified	22	3.2	0.12	7 (7)
Unclassified	3	3.8	0.14	3 (4)
Total interspersed repeats		1,226.8	44.83	

# Distribuce repetitivních sekvencí v genomech není rovnoměrná: Genom jako ekosystém elementů



RETROELEMENTY  
A RETROTRANSPOZICE

# Úvod

- RNA svět -----> DNA svět  
význam reverzní transkripce (při vzniku DNA i dnes)
- retroelementy - relikty RNA světa (struktura, replikační mechanismy, všudypřítomnost)
- úloha v evoluci genomů,  
RT - nejhojnější gen kódující protein v lidském genomu

# Co jsou retroelementy?

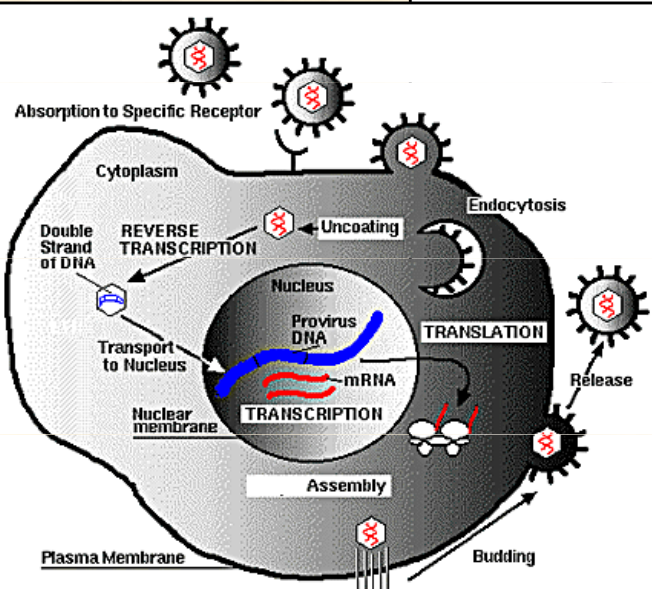
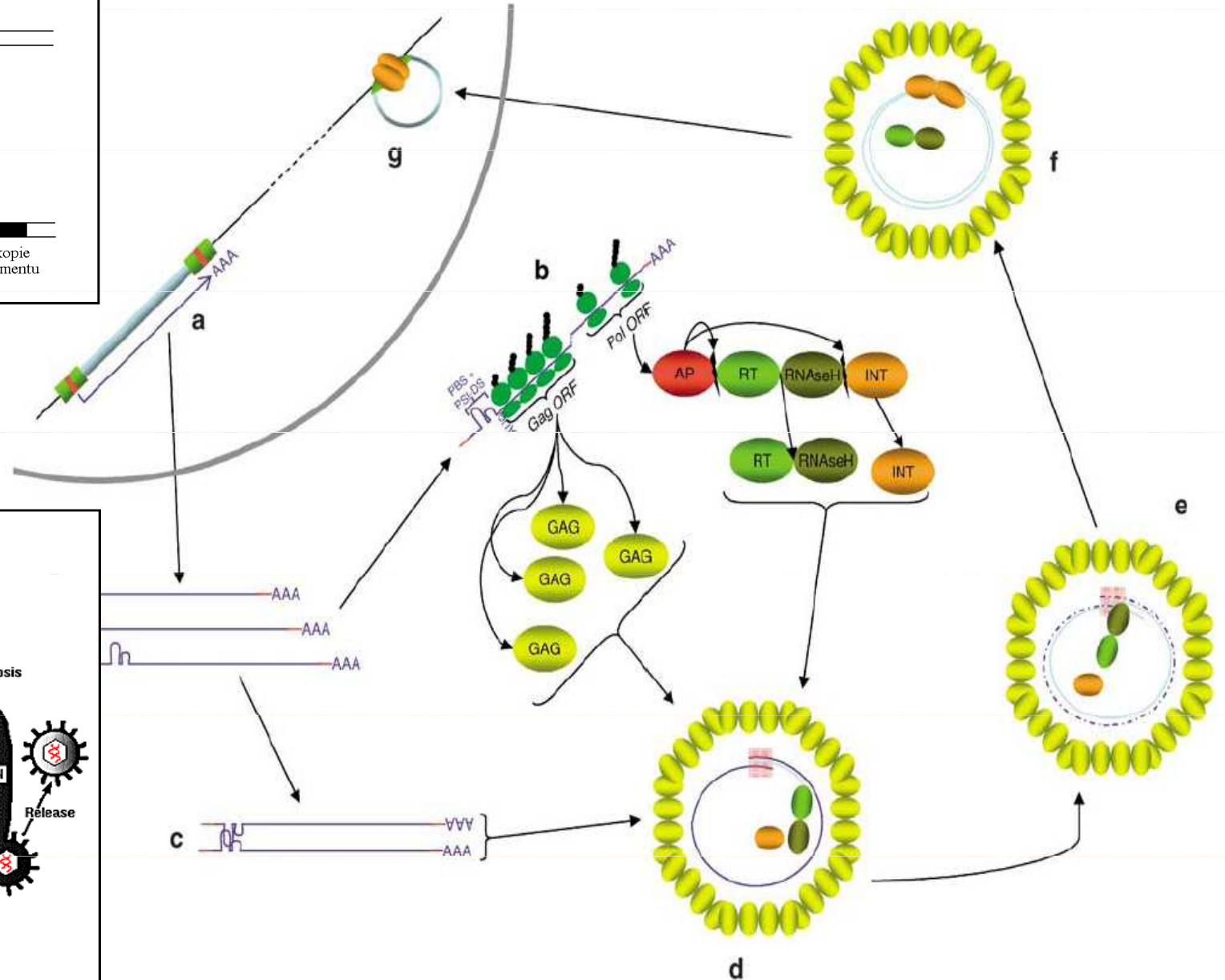
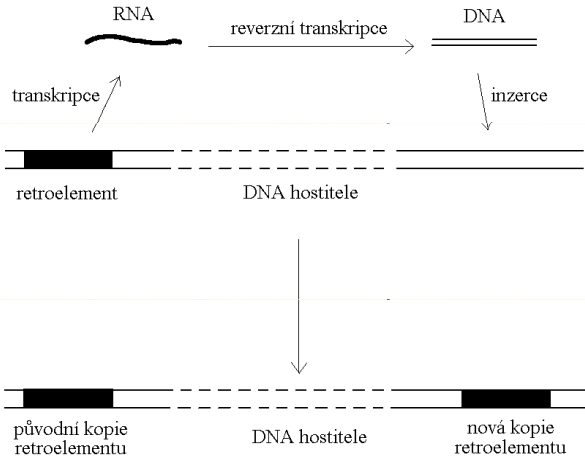
- **Retroelementy** = sekvence DNA nebo RNA obsahující gen pro enzym reverzní transkriptázu (katalyzuje přepis RNA do DNA).  
Širší definice zahrnuje veškeré sekvence vzniklé reverzní transkripcí RNA do DNA.  
Po genomech se šíří procesem retropozice.
- **Retropozice** = přesun genetického materiálu z jednoho místa v genomu do místa jiného prostřednictvím RNA intermediátu. Má vždy duplikativní charakter.



# Životní cyklus retrotransposonů a retrovirů

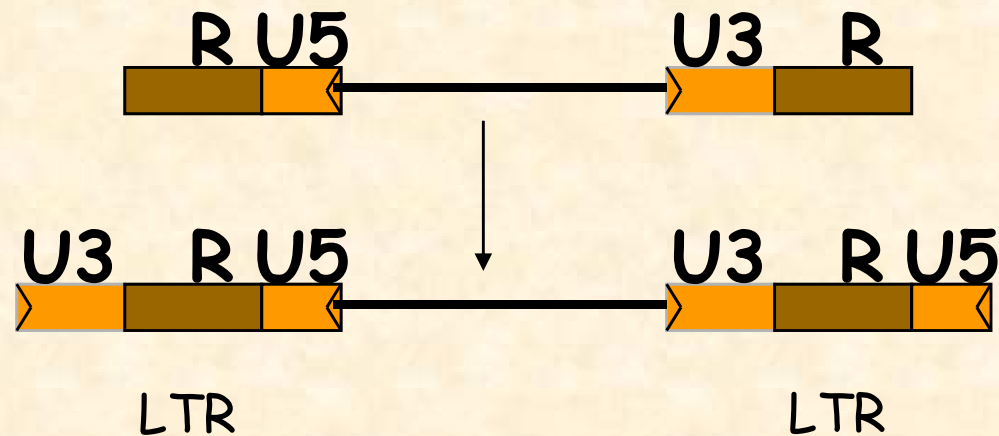
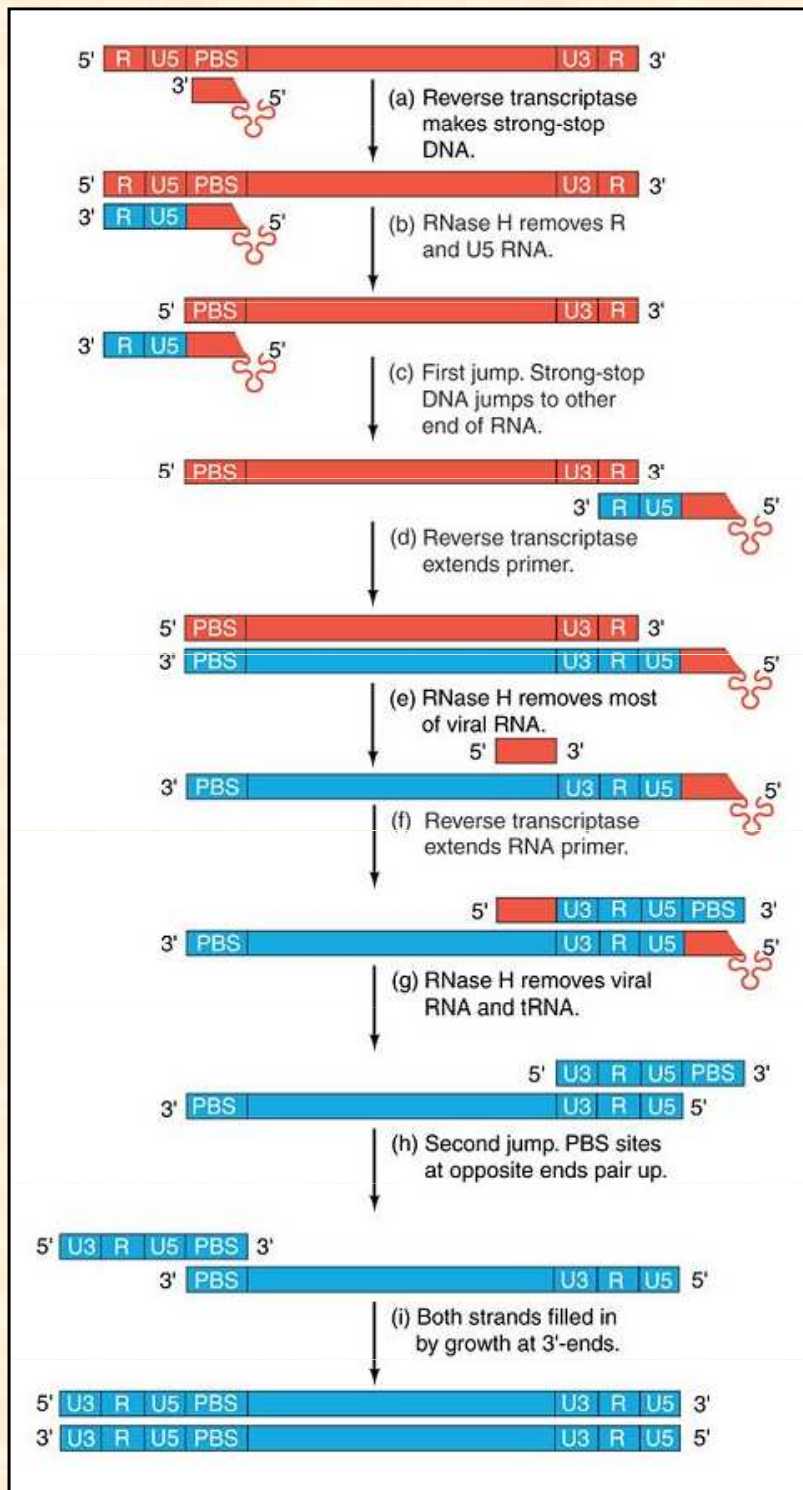
Parasite and parasite of parasites  
F Sabot and AH Schulman

## retrovirů



# 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



# ZÁKLADNÍ TYPY RETROELEMENTŮ

# Typy retroelementů

<i>TYP RETROELEMENTU</i>	Gen pro reverzní transkriptázu	Retropozice (integráza)	Dlouhé koncové repetice (LTR)	Tvorba virionu
<b>Retroviry</b>	+	+	+	+
<b>Pararetroviry</b>	+	-	-	+
<b>Retrotranspozony</b>	+	+	+	-
<b>Retropozony</b>	+	+	-	-
<b>Retrony</b>	+	-	-	-
<b>Mitochondriální plazmidy</b>	+	-	-	-
<b>Introny II. skupiny</b>	+	-	-	-
<b>Telomery</b>	-	-	-	-
<b>Retrogeny</b>	-	-	-	-
<b>Retropseudogeny</b>	-	-	-	-
<b>Retrosekvence</b>	-	-	-	-

# Retroviry - obecné schéma



gag, pol, env ..... geny  
 LTR ... dlouhé koncové repetice  
 PBS .... místo vazby primeru  
 PR ..... proteáza  
 INT .... integráza  
 RT ..... reverzní transkriptáza  
 PPT .... polypurinový úsek  
 > ..... přímé repetice

LTR:

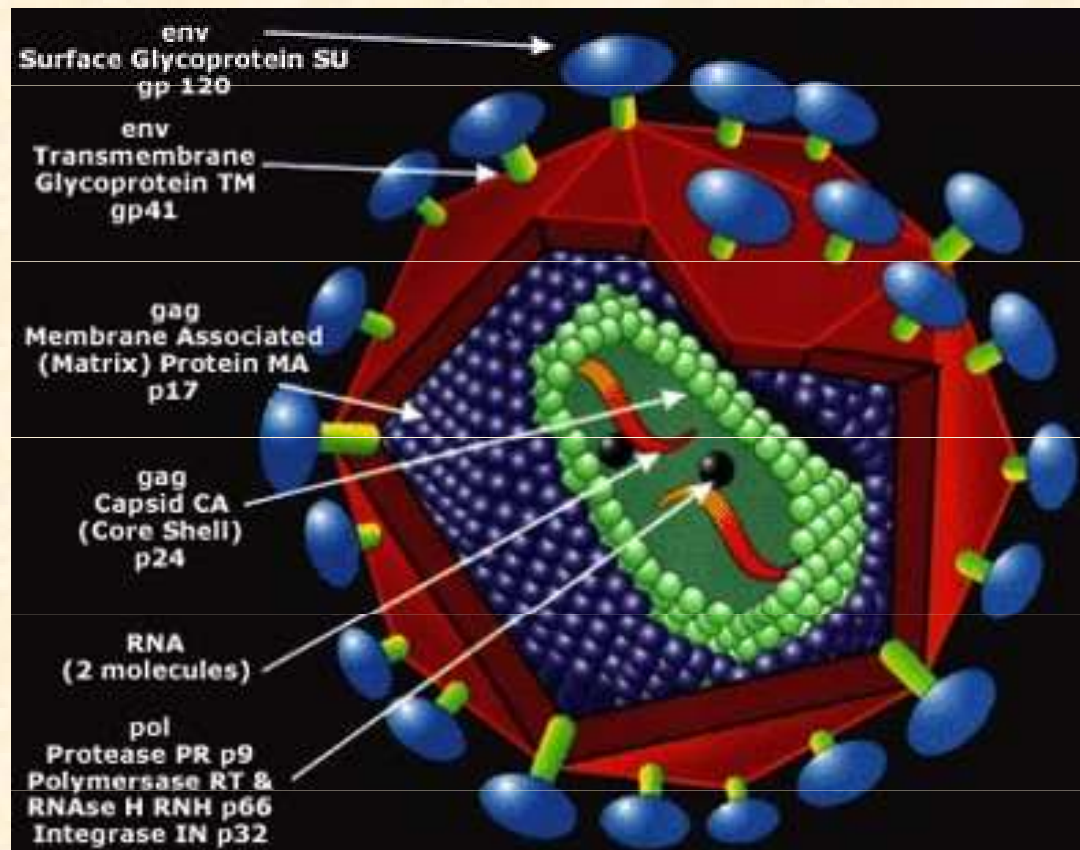


krátká obrácená repetice  
 zesilovače transkripce

promotor

polyadenylační signál

krátká obrácená repetice



U3 ... jedinečná sekvence na 3' konci  
 R ..... repetitivní sekvence  
 U5 ... jedinečná sekvence na 5' konci

# Retrotranspozony - obecné schéma

(a) 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

(b) Typ Ty3-gypsy:



# Retropozony

(polyA, nonLTR retroelementy)

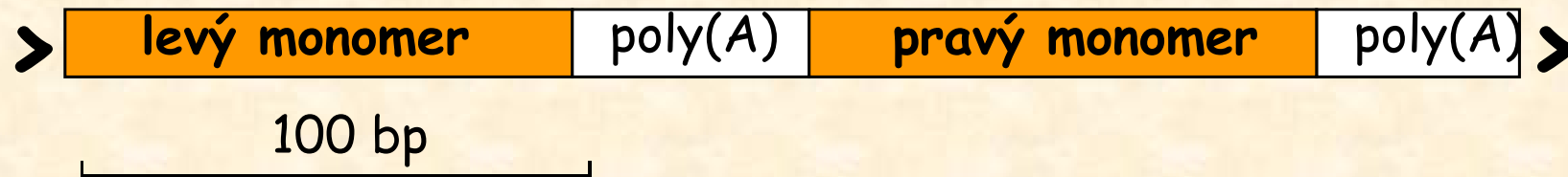
## (a) LINE (L1):



1 kb

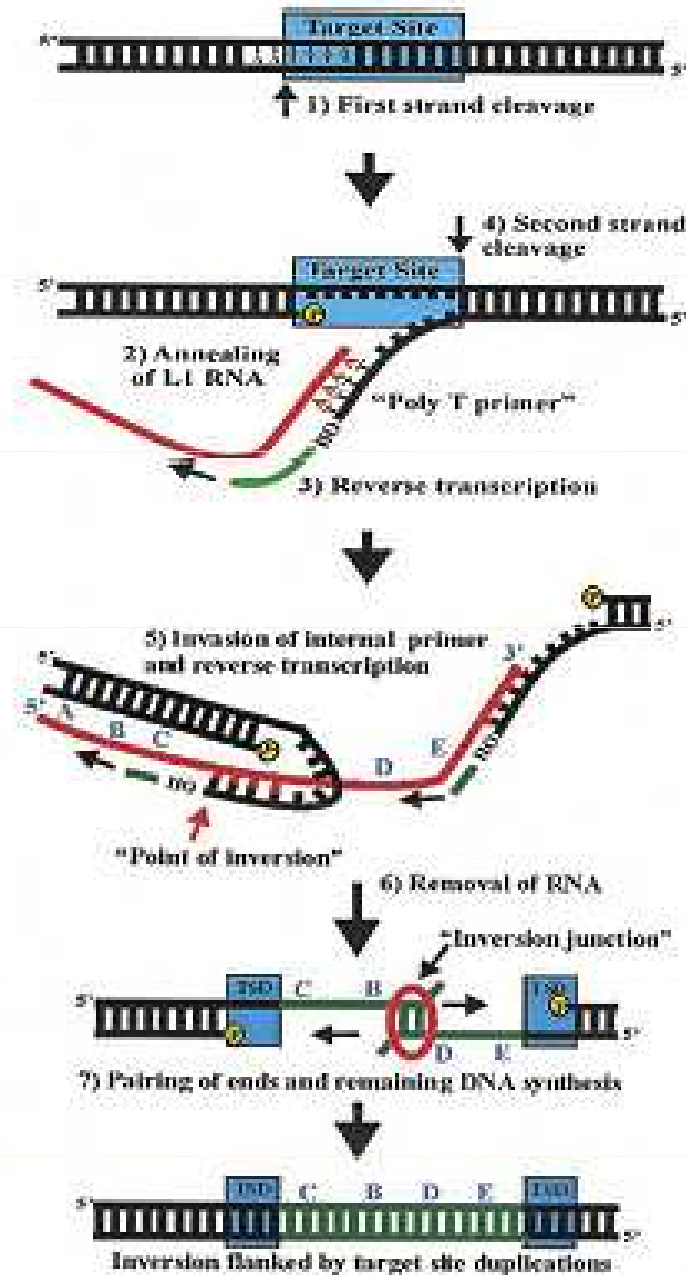
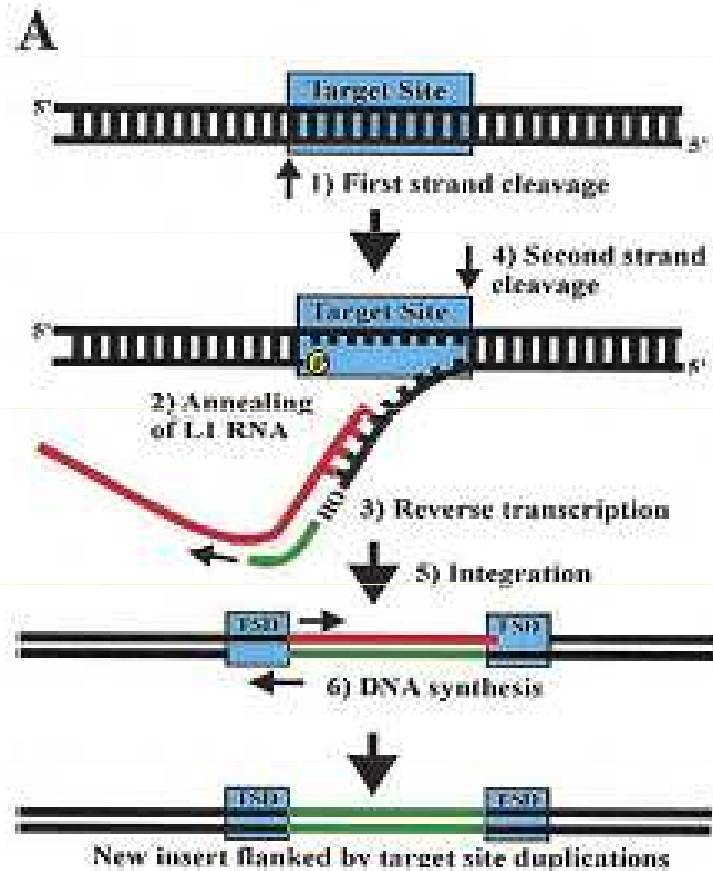
- ORF1 ..... protein vážící RNA
- EN ..... doména kódující endonukleázu
- RT ..... doména kódující reverzní transkriptázu
- 5'UTR ..... netranslatovaná oblast na 5' konci
- 3'UTR ..... netranslatovaná oblast na 3' konci
- PPT ..... polypurinový úsek
- poly(A) ..... polyadeninový úsek
- > ..... přímé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

## (b) SINE (sekvence Alu):



# Včleňování LINE elementů

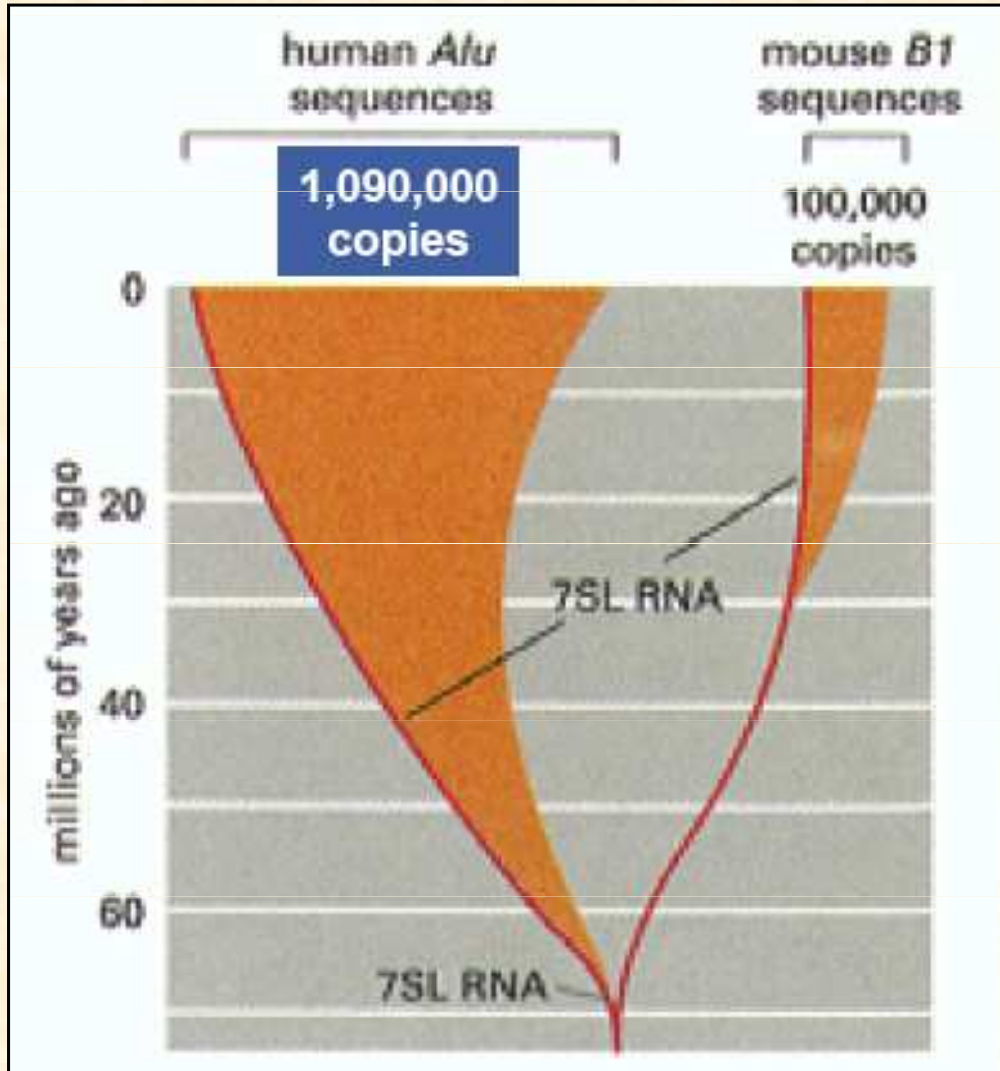
Twin priming:  
mechanismus tvorby inverzí



TPRT (target-primed RT)



# Alu elementy - nejhojnější retroelementy lidského genomu



- odvozeny z 7SL RNA genu kódujícího podjednotku signální rozpoznávací částice (přenos proteinů přes membrány a začleňování do membrán)

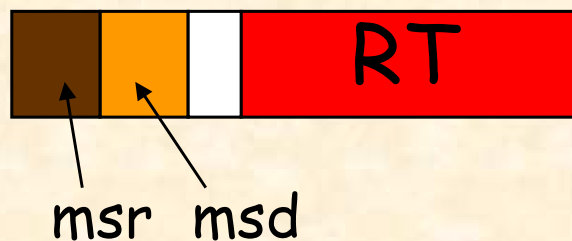
- Alu inserce - u každého 200 narozeného jedince

- transkripce z promotorů RNAPolIII (uvnitř oblasti)

- reverzní transkripce zajišťována RT poskytovanou LINE elementy

# Retrony - primitivní retroelementy bakterií

(a) retron:

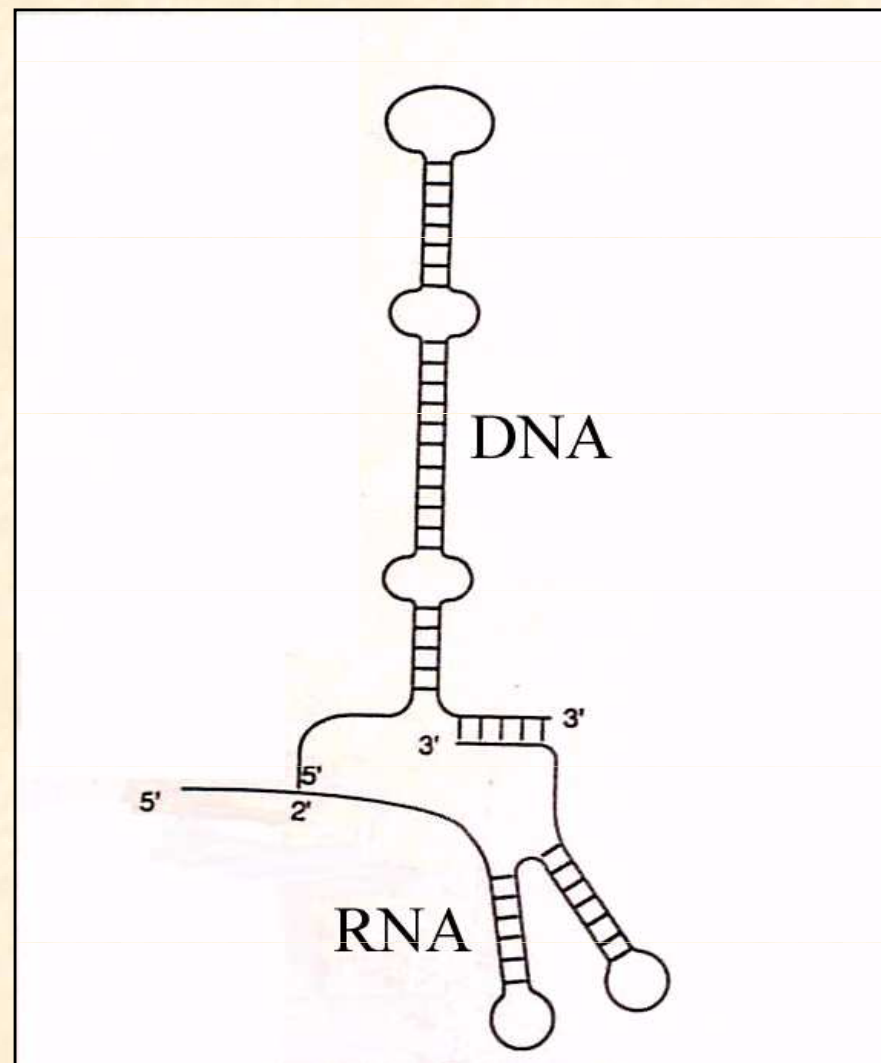


msr ... gen kódující RNA složku  
msd ... gen kódující DNA složku  
RT ..... gen pro reverzní transkriptázu

**msDNA (=multicopy single-stranded)**

- intermediáty extrachromosomálního replikačního cyklu nebo
- abortivní produkty intermediátů

(b) msDNA:



# Retroelementy prokaryot jsou starobylé

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
4. **Telomeráza** (starobylý eukaryotický enzym) je podobná prokaryotickým RT a RT retroposonů

## Typy prokaryotických retroelementů:

1. Retrony
2. Retroplasmidy
3. Retrointrony (introny II. skupiny)

# Další primitivní retroelementy

## 1. Mitochondriální (retro)plazmidy:

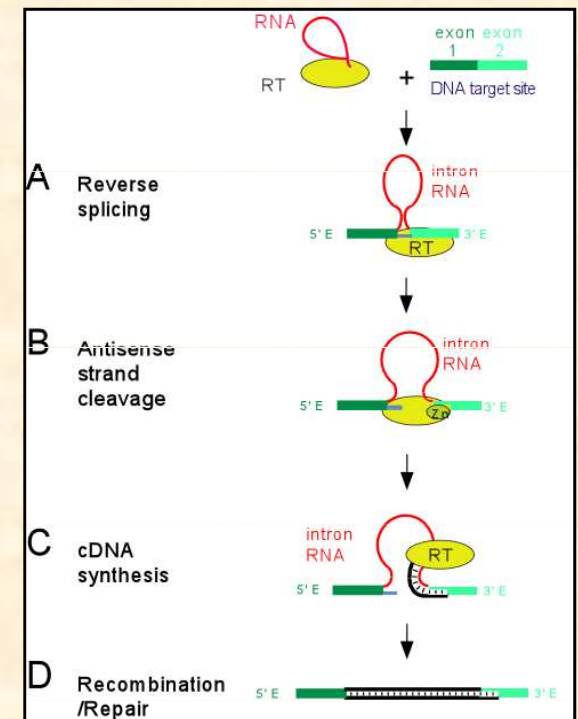
dsDNA plazmidy, v mitochondriích hub, kódují RT, replikace procesem transkripce-reverzní transkripce, homologie s tRNA, sekundární struktury na 3'konci, vlásenky

## 2. Introny II. skupiny:

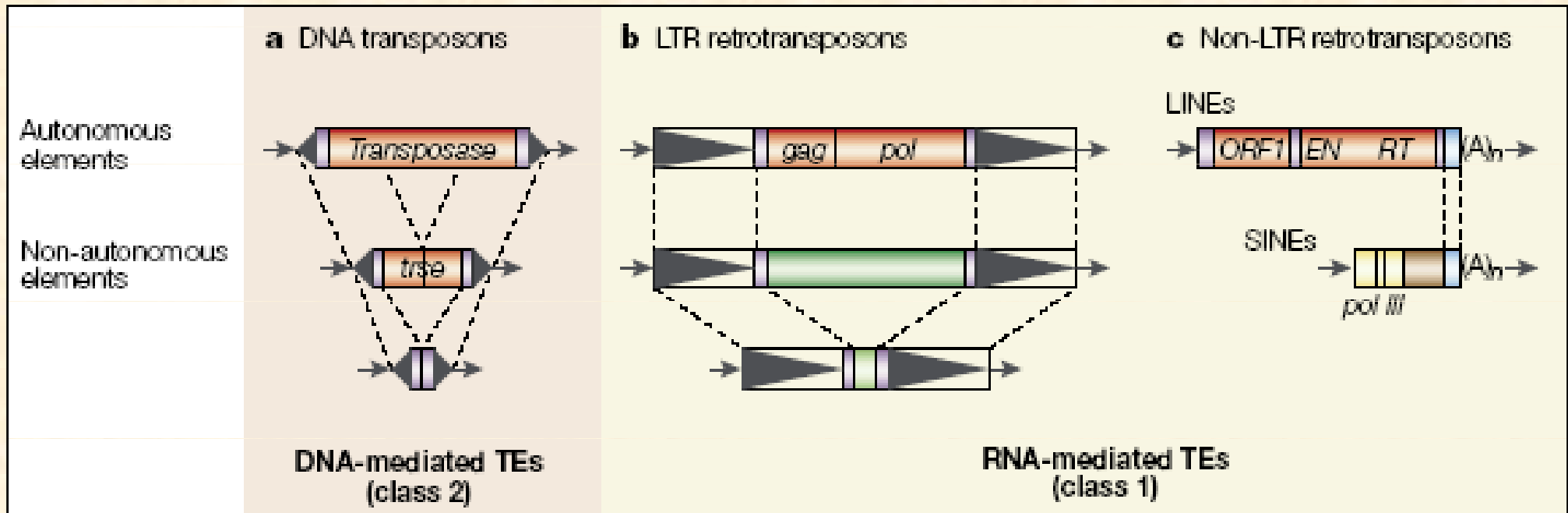
u bakterií a v organelách hub, řas a rostlin, kódují RT, samosestřih, „reverzní sestřih“ do RNA bez intronu a reverzní transkripce

## 3. Telomery:

u eukaryot, kompenzace zkracování chromozomů, telomeráza - homologie s RT, vlastní RNA templát



# 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

# PŮVOD A EVOLUCE RETROELEMENTŮ

# Původ a evoluce retroelementů

## 1. Původní genomy byly tvořeny molekulami RNA

- myšlenka **RNA** jako předchůdce DNA (Crick, 1968)
- **všudypřítomnost RNA** (viz Relikty světa RNA)
- objev **ribozymu** (Cech, 1982) - uložení i realizace informace
- **reverzní transkriptáza** (Temin, 1970), homologie s RNA replikázami a telomerázami
- **všudypřítomnost retroelementů**, různá struktura a mechanismus replikace, tRNA a sekundární struktury při převedení informace z RNA do proteinů i do DNA

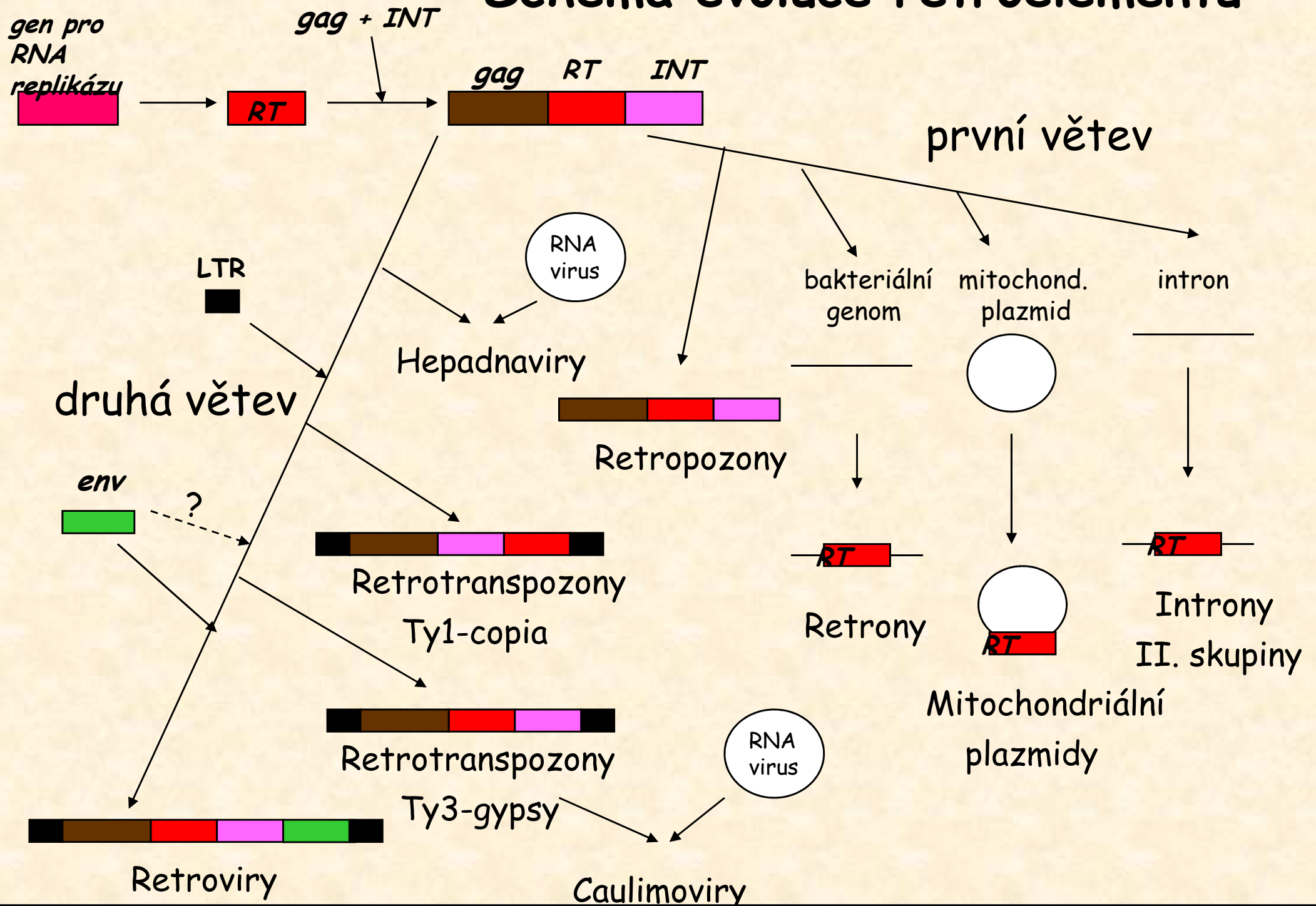
## 2. Hledání společného předka všech retroelementů

- **příbuznost** retroelementů, evoluční strom
- **progenitor** - jednoduchá struktura a všudypřítomnost
- dvě alternativní **hypotézy**:
  - (a) **retrony** (nejjednodušší, codon usage - včleněny dávno)
  - (b) **retropozony** (největší diverzita sekvencí RT, všudypřítomné, jednoduchý mechanismus)

## 3. Dvě hlavní evoluční větve retroelementů (OBR)

- **první větev**: retroelementy bez LTR, jednodušší struktura
- **druhá větev**: retroelementy s LTR, složitější struktura,

# Schéma evoluce retroelementů

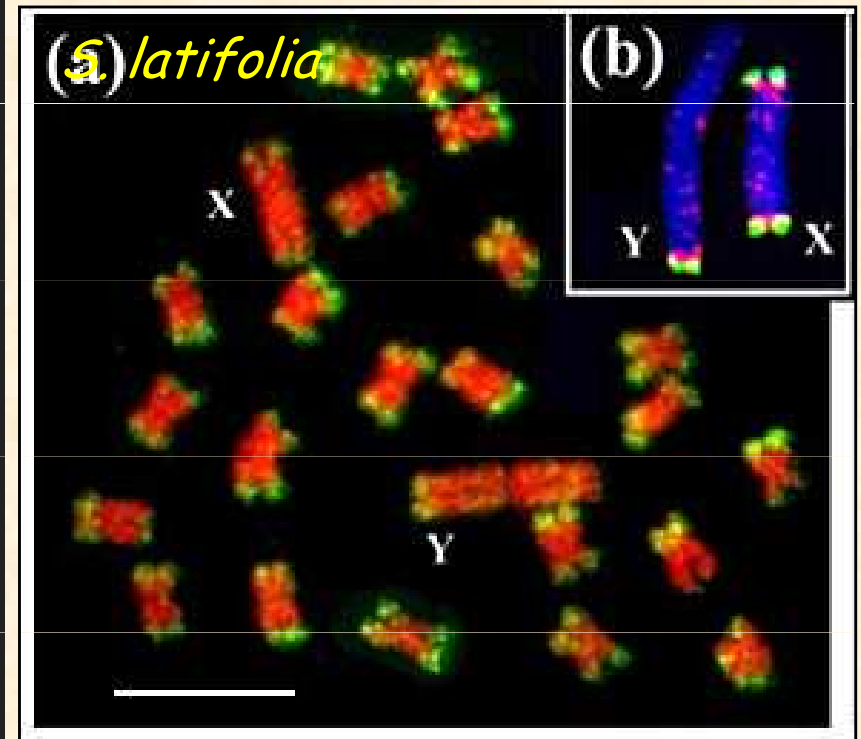
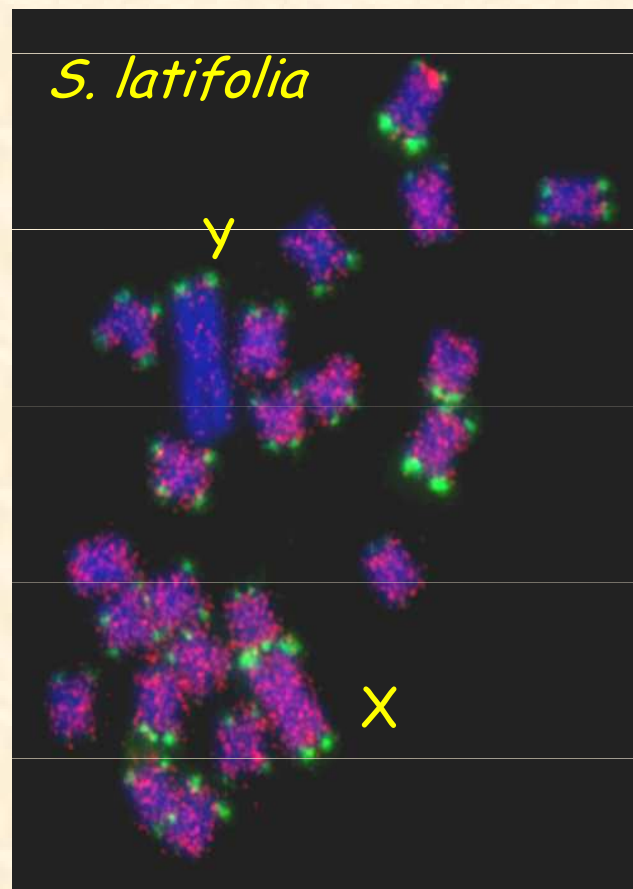
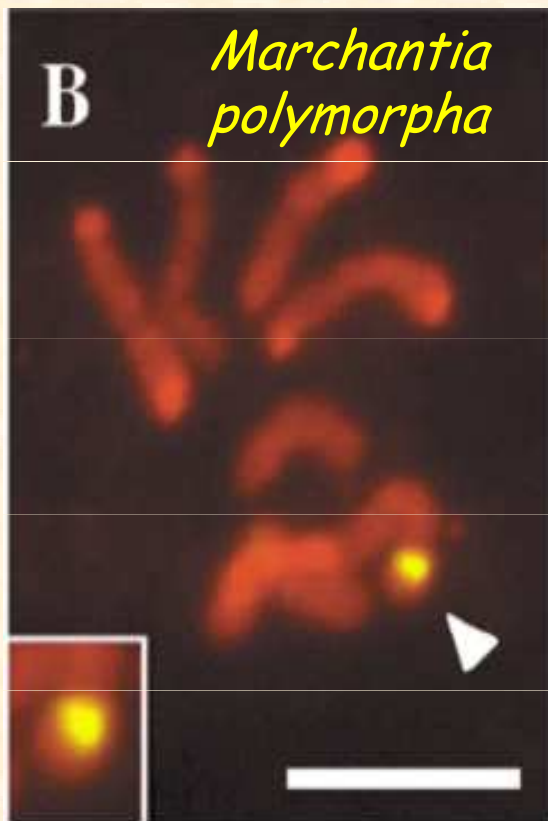




CHROMOSOMÁLNÍ  
DISTRIBUCE  
RETROELEMENTŮ

# Chromozomální distribuce retroelementů

- výskyt na všech chromozomech (hybridizace *in situ*)
- místa s vyšší koncentrací retroelementů i bez retroelementů
- retroelementy v heterochromatinu i euchromatinu, hřbitovy RE (targeting nebo selekce?)
- sex chromozomy - akumulace na chromozomu Y:  
u *D. miranda*, *Cannabis sativa*, *Marchantia polymorpha*, *Silene latifolia*



# Inzerční specificita a inserce retroelementů do heterochromatinu

- inserce není náhodná (retroviry)
- role IN - targeted domain (TD)
- interakce s proteiny chromatinu
- inženýrství (nové specifity), genová terapie
- „local hopping“

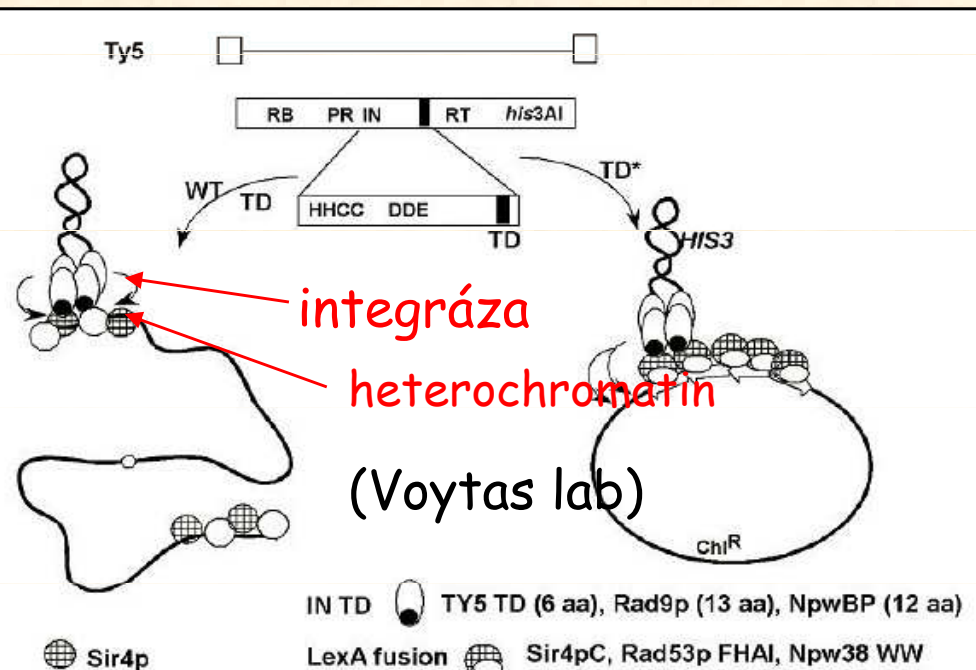
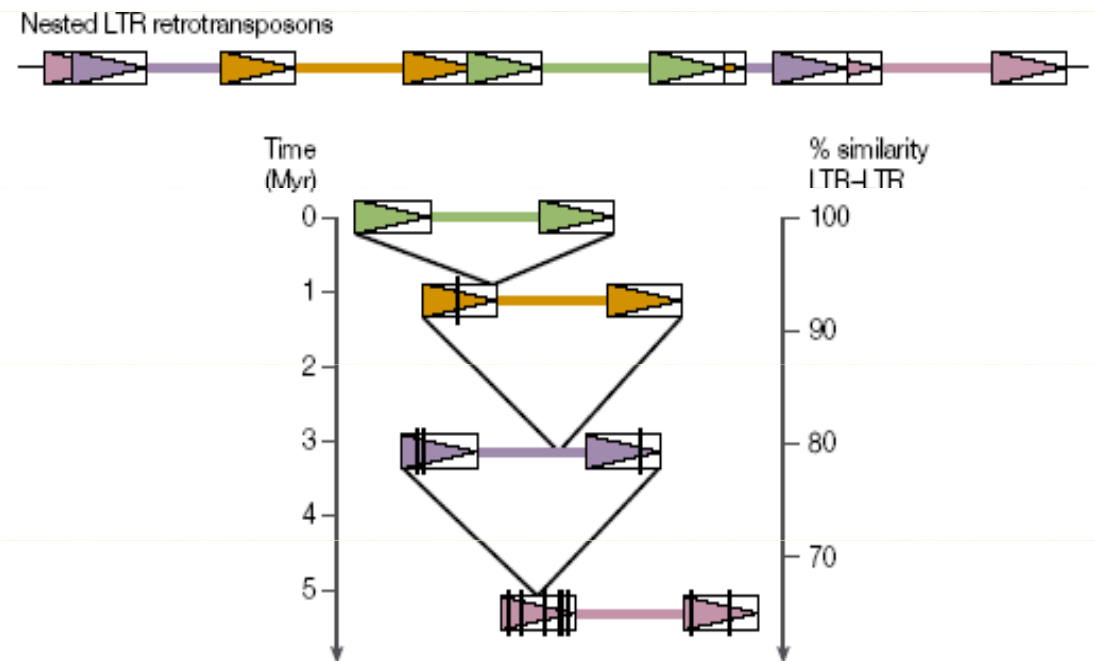
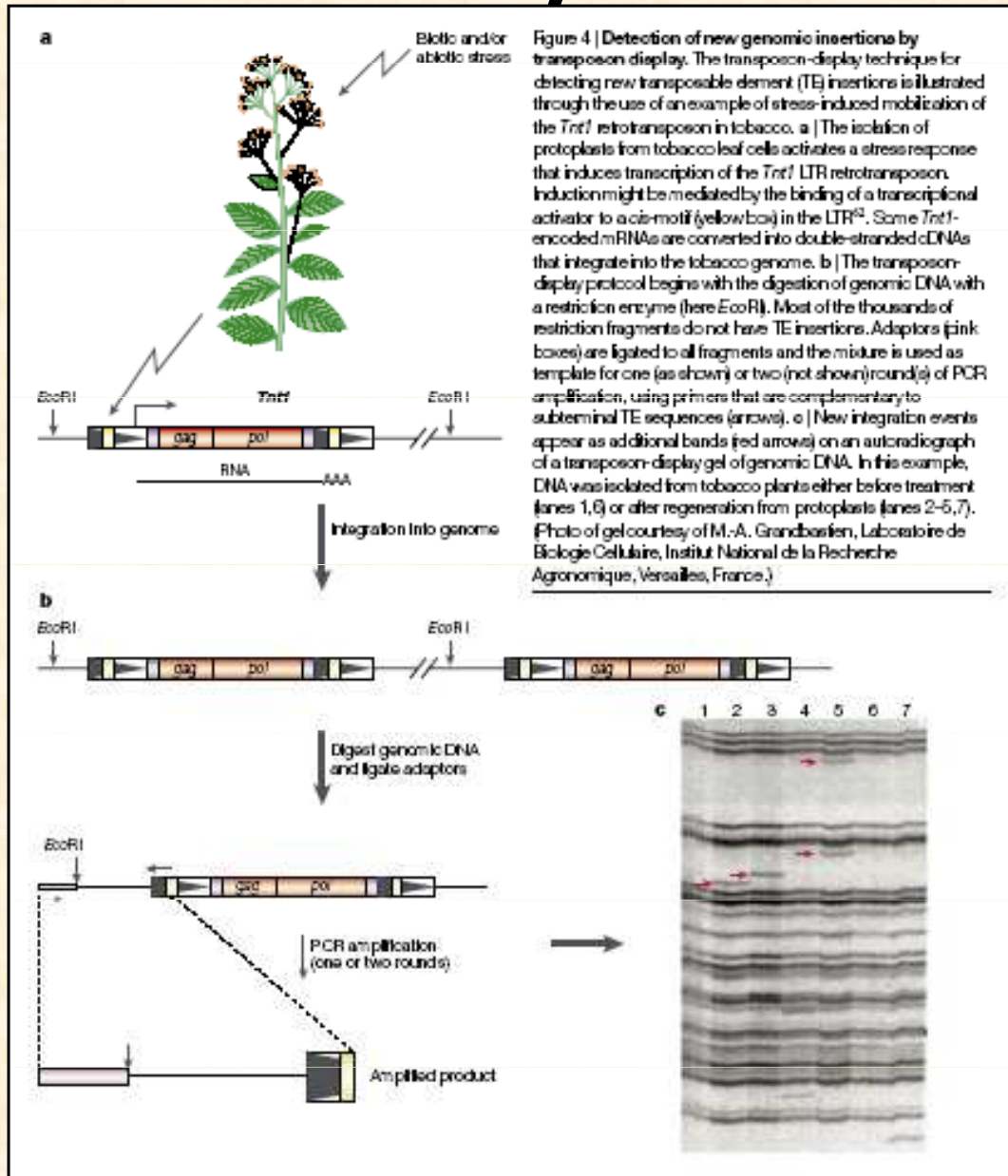


Fig. 1. Strategy for retargeting Ty5 integration. Top, schematic of Ty5 single ORF encoding RNA binding (RB), protease (PR), integrase (IN), reverse transcriptase (RT), and marker gene (*his3AI*) (open box). View of IN is expanded to show conserved residues and targeting domain (TD) (solid). Lower left, preintegration complex showing wild-type IN bound to ends of Ty5 DNA (thick line) and integrating into telomeric heterochromatin, mediated by Sir4p (hatched). Lower right, same as left except that the natural TD is replaced with heterologous domains (TD\*) from Rad9p and NpwBP (solid). LexA DNA-binding domain (open) is expressed fused to Sir4pC, Rad53p FHA1, and Npw38 WW domains (hatched). Integration occurs proximal to LexA-binding sites (open arrows) in plasmid target (closed circle).

## Inserce retroelementů do již existujících retroelementů



# Metoda „transposon display“ - detekce nových inzercí transposonů



- indukce transpozice stresem
- štěpení genomické DNA
- ligace adaptorů
- PCR okolní oblasti
- nové bandy

# FUNKCE RETROELEMENTŮ

# Funkce retroelementů

## Koadaptace retroelementů a hostitele

- alternativní hypotézy - parazité x významný činitel v evoluci, slučitelnost hypotéz, koadaptace, vliv na hostitele
- **pravidlo 3K**: konflikt-kompromis-kooperace (ekologie genomu)
- **Mutabilita** - stochastické ale regulované „mutageny“
- místo inzerce retroelementu do genomu - exony, introny, regulační sekvence, LTR, RE, sekvenční specificita
- regulace počtu kopií retroelementů v genomu - obranné mechanismy hostitele, metylace, rekombinace, mechanismy RE

## Negativní vliv retroelementů na hostitele

- vyplývá z povahy RE - sobecká a parazitická DNA
- choroby (hemofilie, rakovina), L1, Alu
- příčinou inserce nebo rekombinace (mezi Alu)

## Mechanizmy transposonu minimalizující jejich vliv na hostitele

- omezení transpozice na zárodečou linii
- přednostní včleňování do heterochromatinu
- odstraňování elementů rekombinací mezi homologickými elementy

## Obranné mechanismy hostitele

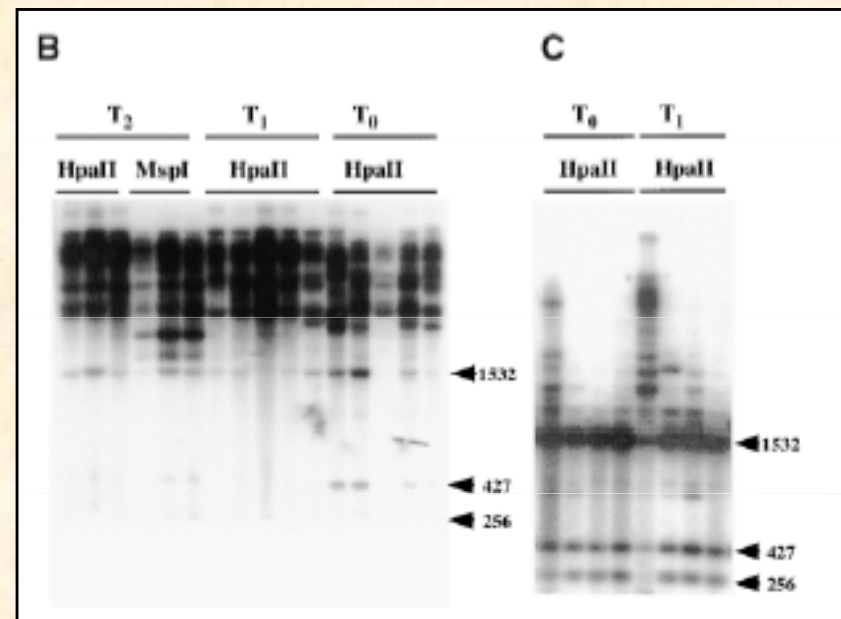
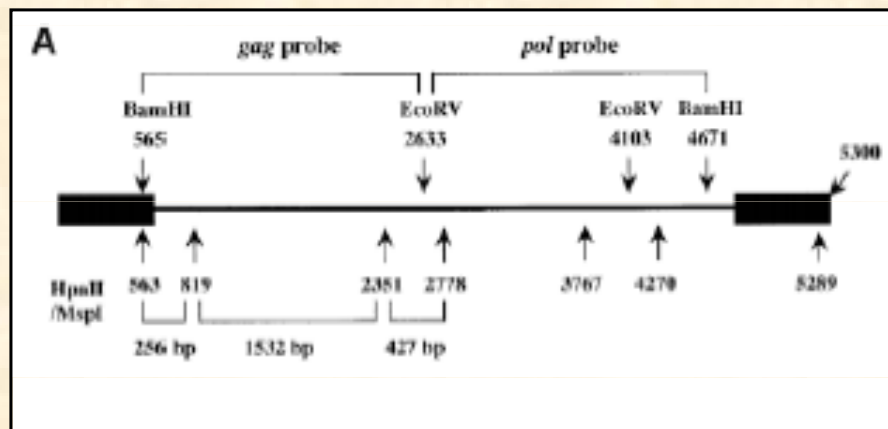
- metylace
- RNA interference

**OBRANNÉ MECHANIZMY  
HOSTITELE:  
UMLČOVÁNÍ TRANSPOSONŮ**

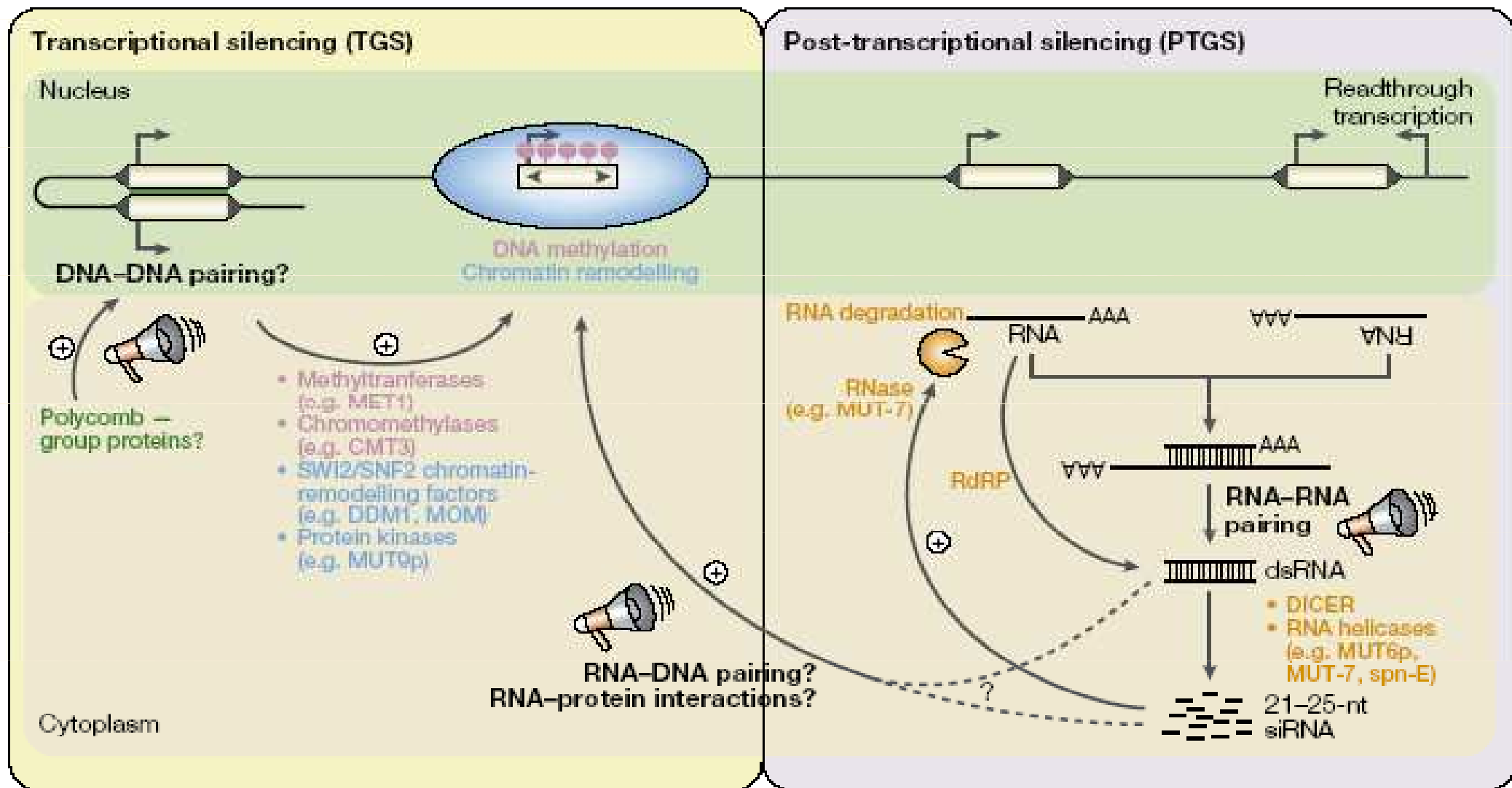


# Umlčování transposonů **metylací** a reaktivace jejich aktivity u mutanta se sníženou metylací DNA:

- Aktivní retrotransposon *Tto1* vnesen z tabáku do *Arabidopsis*
- Ověřena transpozice (obnovení delece v LTR)
- Zvýšení počtu kopií a následná metylace a umlčení (MspI / HpaII)
- Zkřížení s *ddm1* mutantem - vnesení *Tto1* do *ddm1* mutanta
- Snížení metylace *Tto1* a transkripční a transpoziční aktivity



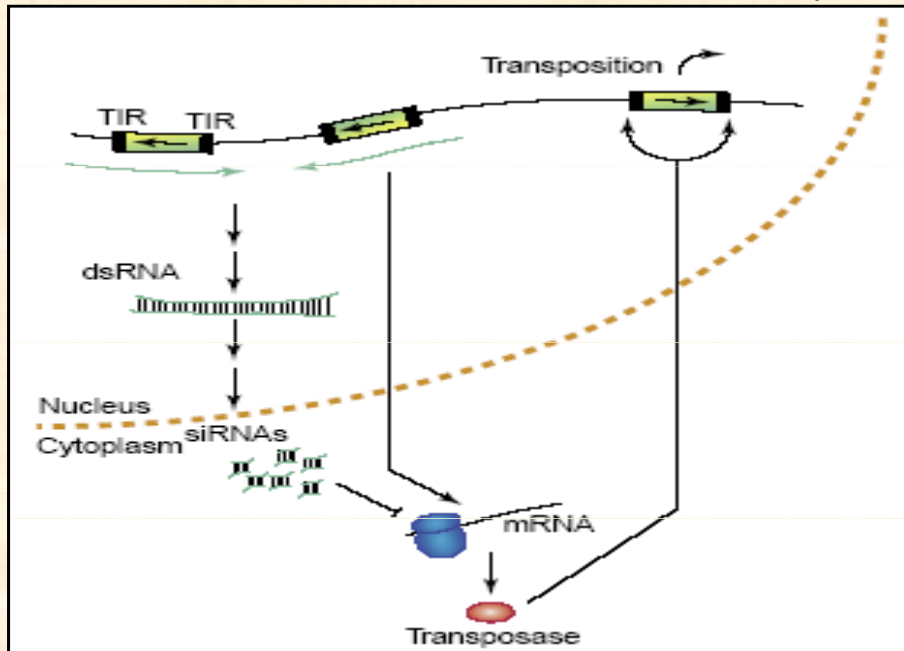
# Umlčování transposonů: transkripční i posttranskripční (metylace a RNAi)



Transkripční umlčování - metylace promotorů TE

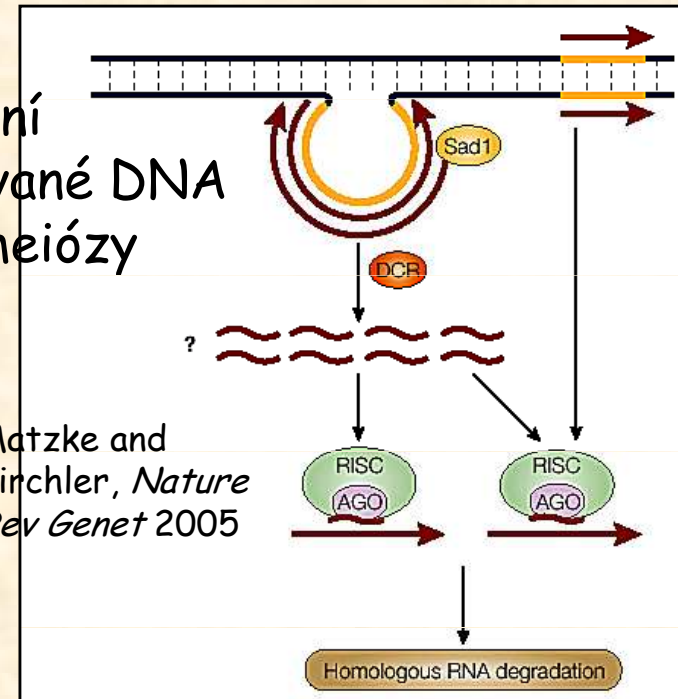
Posttranskripční umlčování - sekvenčně specifická degradace RNA v cytoplazmě

# Umlčování transposonů mechanismem RNAi

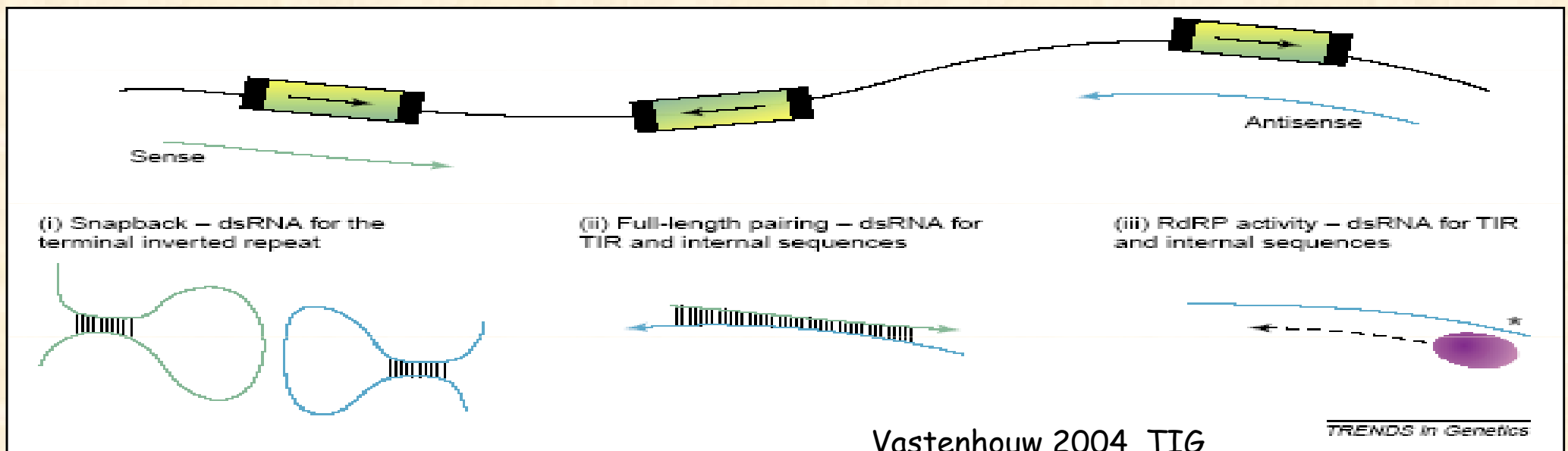


Umlčování nespárované DNA během meiózy

Matzke and Birchler, *Nature Rev Genet* 2005



- *C. elegans* a *Drosophila* nemají metylaci DNA
- Umlčování transposonů je zajišťováno pomocí RNAi



# TRANSPOSONY UŽITEČNÉ PRO HOSTITELE

## DOMESTIKACE TRANSPOSONŮ

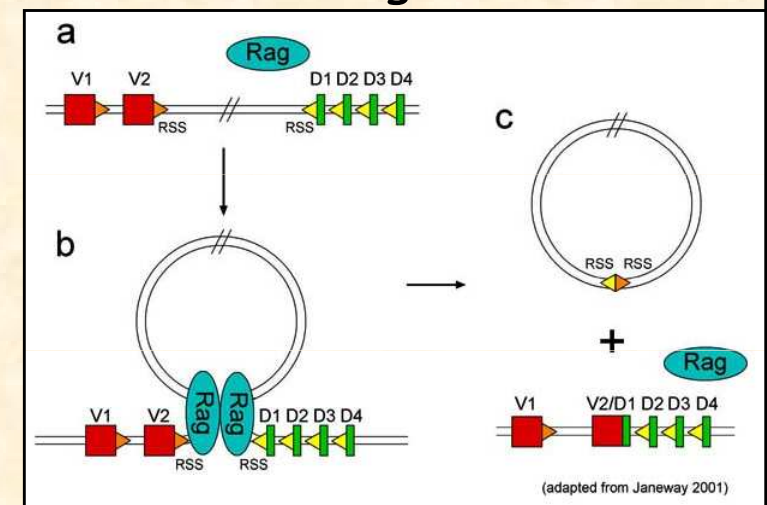
# Pozitivní vliv transposonů na hostitele

## 1. Transposony jako zdroj funkcí užitečných pro hostitele - **strukturní** funkce:

- přestavby genomu, **rekombinace**
- **duplikace**, genové rodiny, vznik nových **genů** (Setmar)
- **telomery** drozofily - Het-A, TART
- centromery - **CENP-B** pochází z transposonů (pogo, Tigger)
- role v **segregaci** chromosomů, reprodukční izolace, speciace

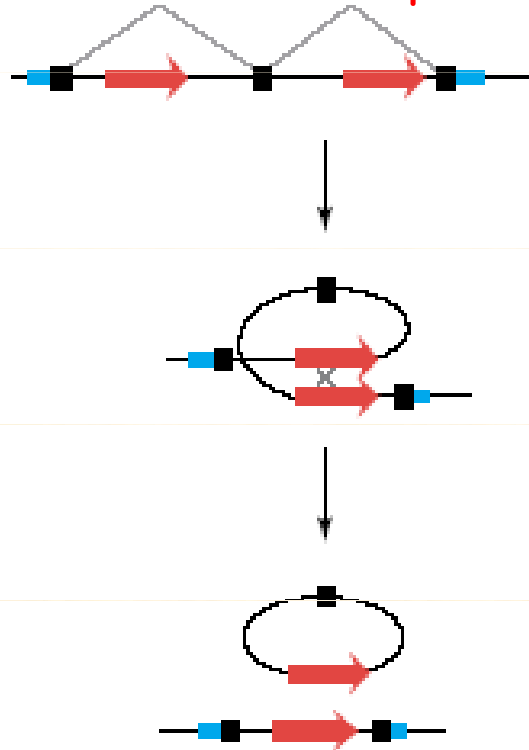
## 2. Transposony jako zdroj funkcí užitečných pro hostitele - **enzymatické** a regulační funkce:

- imunitní systém - **V(D)J** rekombinace
- **rezistence** k virovým chorobám, imunosuprese u živorodých
- **reparace** zlomů v DNA
- **regulační** úloha MITE ležících v intronech a regulačních oblastech genů
- inaktivace chromosomu **X**

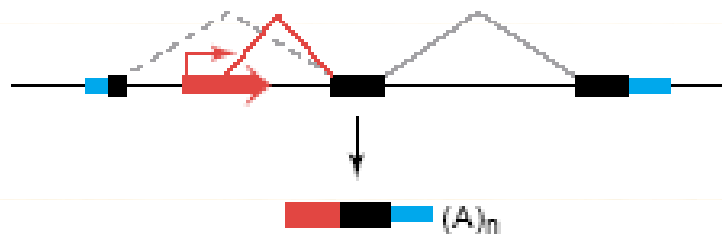


# Přestavby genomu transposony

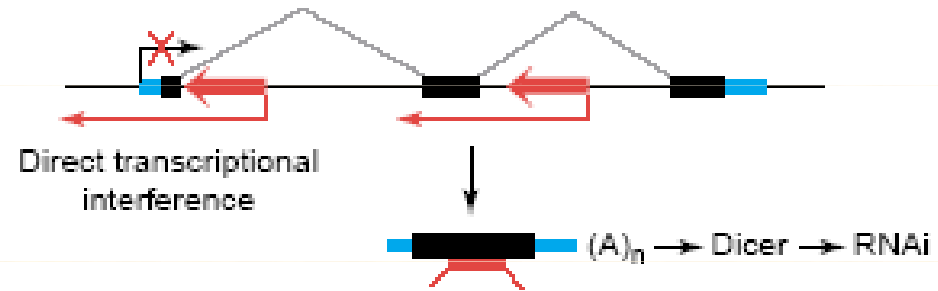
(a) Rekombinační přestavby



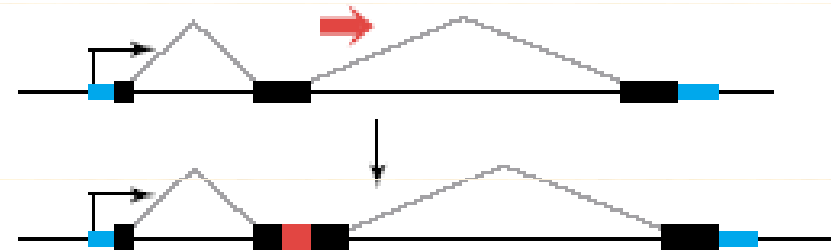
(b) Chimerická mRNA



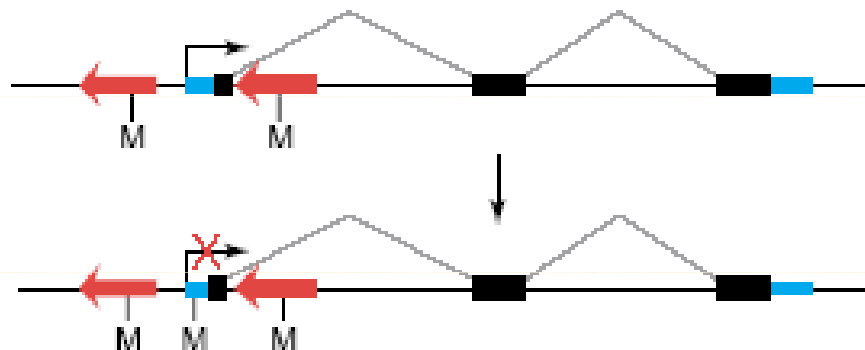
(c) Antisense RNA



(d) Inzerční mutagenéze

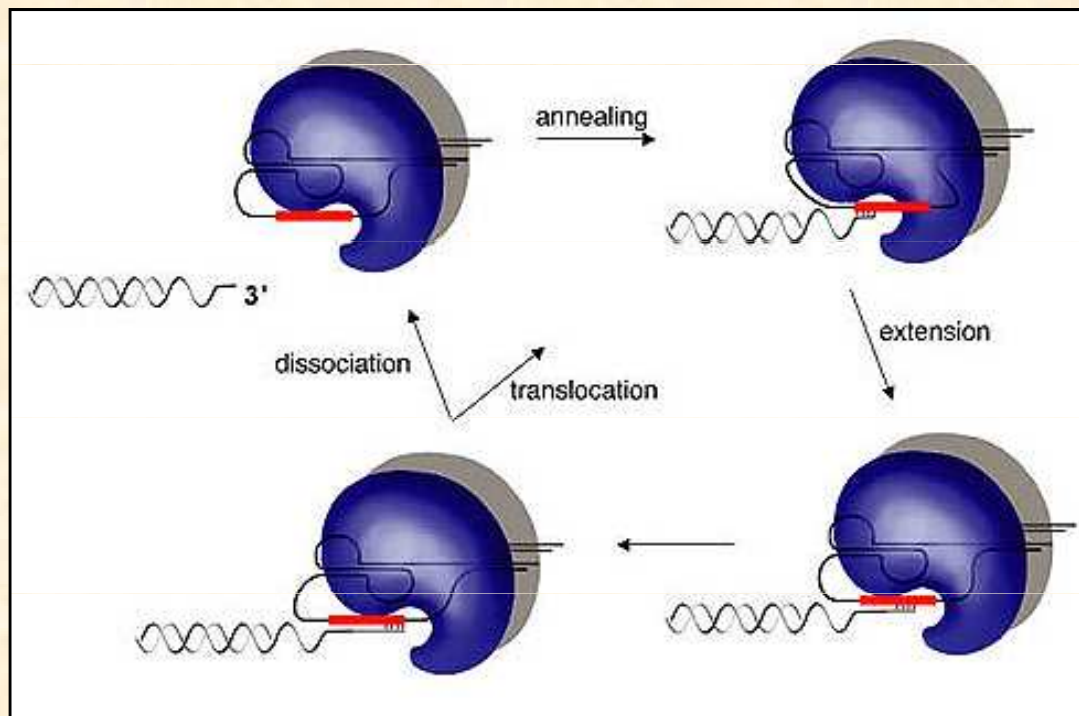


(e) Umlčení způsobené metylací



# Retrotransposony mohou nahradit telomery u drozofily

- **Het-A** a **TART** retrotransposony u **drozofily**
- stále trvající přechod ze světa RNA do DNA,
- TE jako „genome builders“
- L1 RT tvoří **cDNA** sekvenčně nespecificky i z **buněčných genů**
- L1 **zaceluje zlomy** v DNA



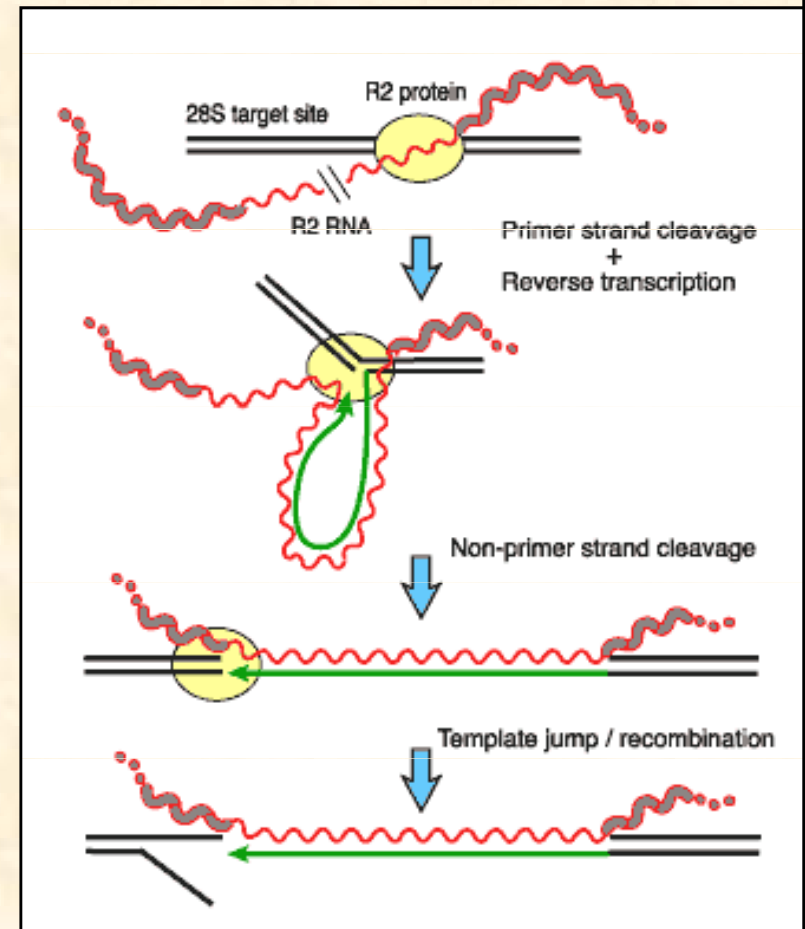
- homologie mezi RT a telomerázou

(Science 276, 561, 1997)

# LINE elements - významný hráč v genomu člověka

- vliv na velikost genomu u primátů (**struktura**), nedávná aktivita u primátů, rozdíly mezi druhy i uvnitř druhů
- místa homologií pro genovou konverzi a rekombinaci (**evoluce**)
- včleňuje se do genu a narušují kódující sekvence, nemoci (**funkce**)
- exon shuffling, zkráceny na 5' konci - nedosyntetizovány
- zacelují místa chromosomových zlomů (**reparace**)

Včleňování mechanismem  
TPRT  
(target primed reverse  
transcription)





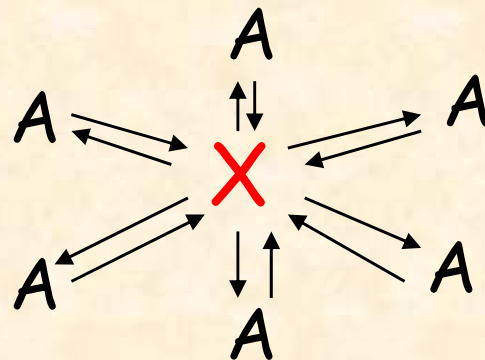
# Chromosom X - křižovatka pohybu genů mechanismem retrotranspozice

## 1. Únik genů z chromosomu X je ovlivněn jeho inaktivací (Xi):

- meiotická inaktivace pohlavních chromosomů (MSCI) - sex chromosomu kondenzovány v samčí meióze, „sex body“, „XY body“, aby nedocházelo k rekombinaci mezi nehomologickými chromosomy
- geny na XY tedy transkripčně suprimovány
- proto backup genů na autosomech, přenos retrotranspozicí (compensatory hypoth.)

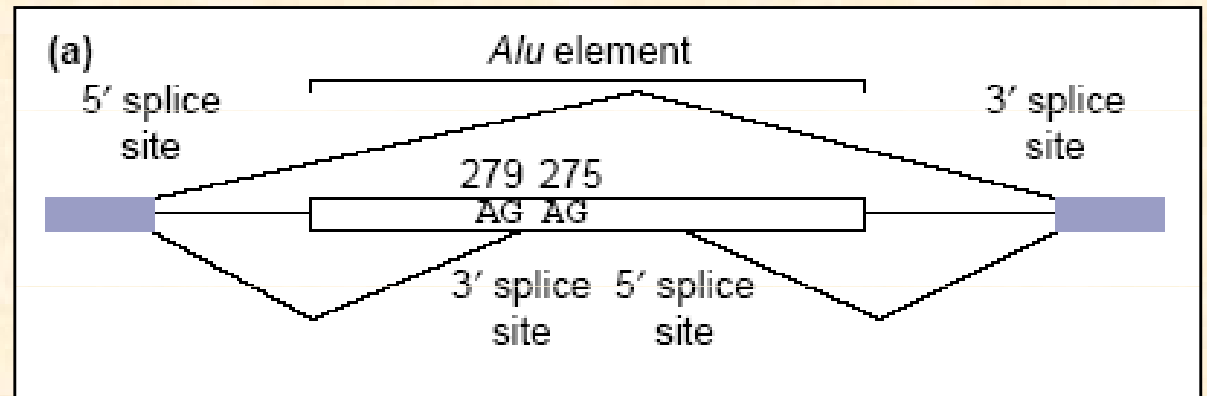
## 2. Přenos genů na chromosom X:

- přednostní včleňování i selekce
- L1 elementy i retropseudogeny, L1 mašinerie
- recesivní geny výhodné pro samce a dominantní geny výhodné pro samice (2/3)



# Alu elementy a alternativní sestřih

- Alu elementy tvoří 10% lidského genomu (>1 milion)
- **5% lidských alternativních exonů je odvozeno z Alu elementů**



## Alternativní sestřih:

- 40-60% lidských genů má alternativní sestřih
- rozdíl mezi člověkem a myší:
- 30 000 genů, fenotypové rozdíly způsobeny druhově-specifickým alternativním sestřihem
- exony - konstitutivní (konzervativní) a alternativní
- alternativní exony - majoritní a minoritní formy

## Vznik alternativních exonů:

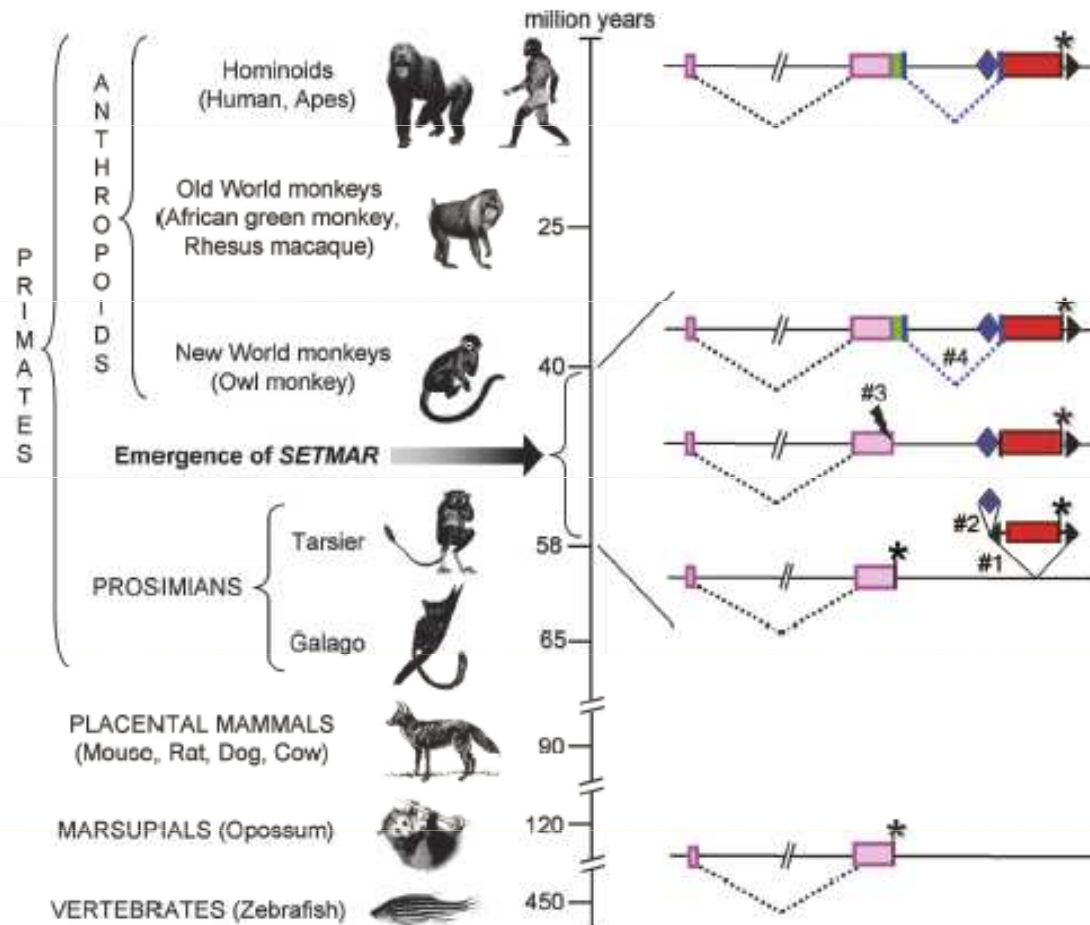
- duplikace exonů,
- inserce transposonu nebo retroviru,
- mutace existující intronové sekvence,
- **z Alu elementů** (5% lidských alternativních exonů)

# Původ genu SETMAR - „recyklace“ transposonu

Birth of a chimeric primate gene by capture of the transposase gene from a mobile element

Richard Cordaux\*, Swalpa Udit†, Mark A. Batzer\*, and Cédric Feschotte††

Histon  
metyltransferáza +  
transpozáza



- zrušení stop
- vznik nového stop
- exonizace
- degenerace TIRu
- vznik intronu
- DNA vazebná doména Tn zachována
- TIR místa v genomu

Fig. 1. Milestones leading to the birth of *SETMAR*. The structure of the *SETMAR* locus (Right) and a simplified chronology of the divergence time of the species examined relative to hominoid primates (Left) are shown. Pink boxes represent the two *SET* exons, which are separated by a single intron (interrupted black line) and form a “*SET*-only” gene whose structure is conserved in all nonanthropoid species examined and terminated with a stop codon (\*) located at a homologous position (except in cow; see Fig. 2a). The *Hsmar1* transposon (event 1) was inserted in the primate lineage, after the split between tarsier and anthropoids, but before the divergence of extant anthropoid lineages. The transposon is shown here with its TIRs (black triangles) and transposase coding sequence (red box). The secondary *Alu5x* insertion within the TIR of *Hsmar1* (event 2) is represented as a blue diamond. The position of the deletion removing the stop codon of the “*SET*-only” gene (event 3) is indicated as a lightning bolt. The *de novo* conversion from noncoding to exonic sequence is shown in green, the creation of the second intron is represented as a dashed blue line (event 4), and the splice sites are shown as thick blue lines.

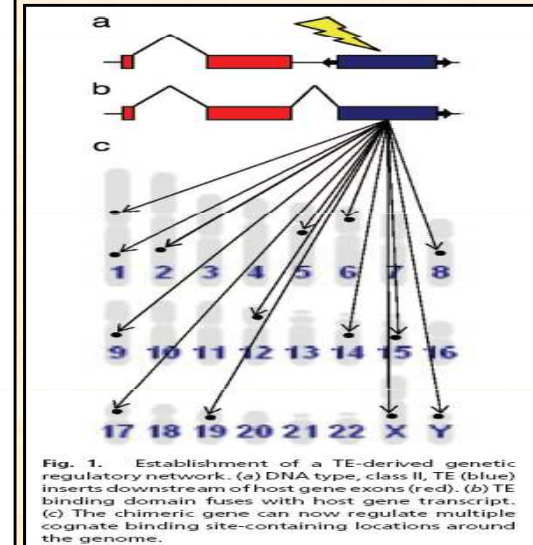


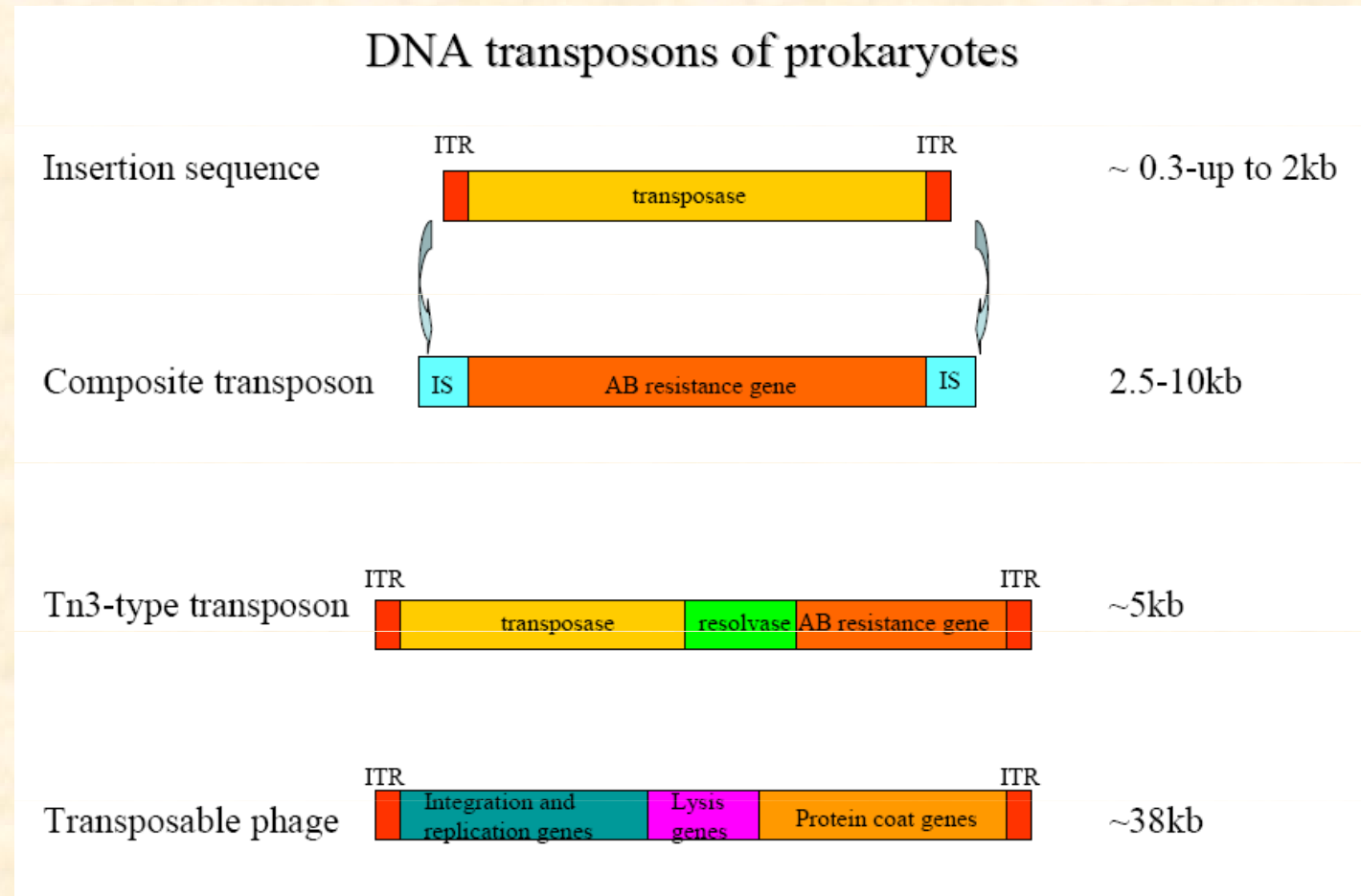
Fig. 1. Establishment of a TE-derived genetic regulatory network. (a) DNA type, class II, TE (blue) inserts downstream of host gene exons (red). (b) TE binding domain fuses with host gene transcript. (c) The chimeric gene can now regulate multiple cognate binding site-containing locations around the genome.

# DNA TRANSPOSONY

# DNA transposony

## 1. Prokaryotické:

- IS elementy
- Tn elementy\_

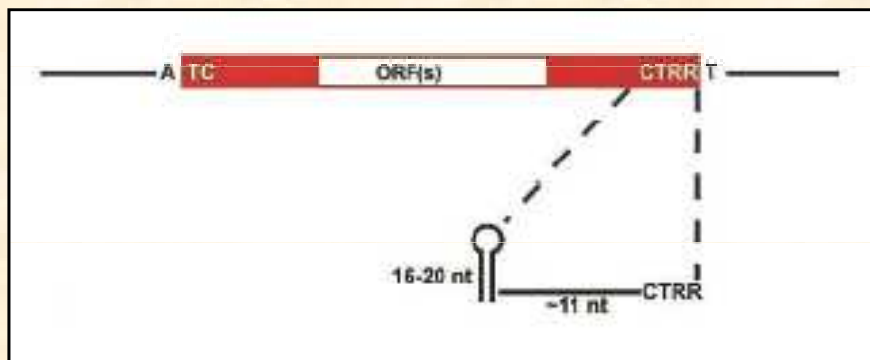


## 2. Eukaryotické:

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

# Helitrony - nedávno objevené transposony využívající mechanismus otáčející se kružnice

- replikace mechanismem otáčivé kružnice (RC, podobně jako plazmidy, jednořetězcové fágy, rostlinné 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)
- méně hojné 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)



# Helitrony putují po genomu a sbírají geny

## NEWS AND COMMENTARY

www.nature.com/hdy

### Plant genomes

## Massive changes of the maize genome are caused by *Helitrons*

SK Lal and LC Hannah

*Heredity* (2005) 95, 421–422. doi:10.1038/sj.hdy.6800764; published online 12 October 2005

## Rolling-circle transposons in eukaryotes

Vladimir V. Kapitonov\* and Jerzy Jurka

Genetic Information Research Institute, 2081 Landings Drive, Mountain View, CA 94043

Communicated by Margaret G. Kidwell, University of Arizona, Tucson, AZ, May 29, 2001 (received for review April 10, 2001)

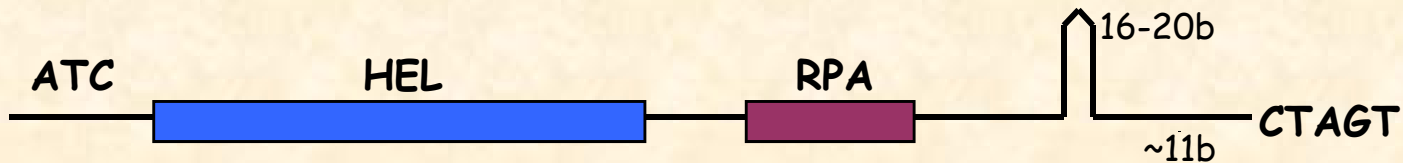
All eukaryotic DNA transposons reported so far belong to a single category of elements transposed by the so-called "cut-and-paste" mechanism. Here, we report a previously unknown category of eukaryotic DNA transposons, *Helitron*, which transpose by rolling-circle replication. Autonomous *Helitrons* encode a CTRR RNA helicase and nuclease/ligase similar to the rolling-circle replicons. *Helitron*-like transposons have 5'-TC and CTRR-3' termini and do not have inverted repeats. They contain 16- to 20-bp hairpin nucleotides from the 3'-end and transpose the 5'-A and T-3', with no modifications. Together with their multiple diverged variants, *Helitrons* constitute ~2% of both the human and *Caenorhabditis elegans* genomes and are found in other eukaryotic genomes. Sequence conservation suggests that they continue to be transposed.

and best illustrated by a recent study of *Sleeping Beauty*, a Tc1-like transposon from fish (13), reconstructed from its inactive copies and demonstrated to be transpositionally active in a test tube. Another much more ancient example is a PiggyBac-like DNA transposon *Loose* discovered in the human genome.

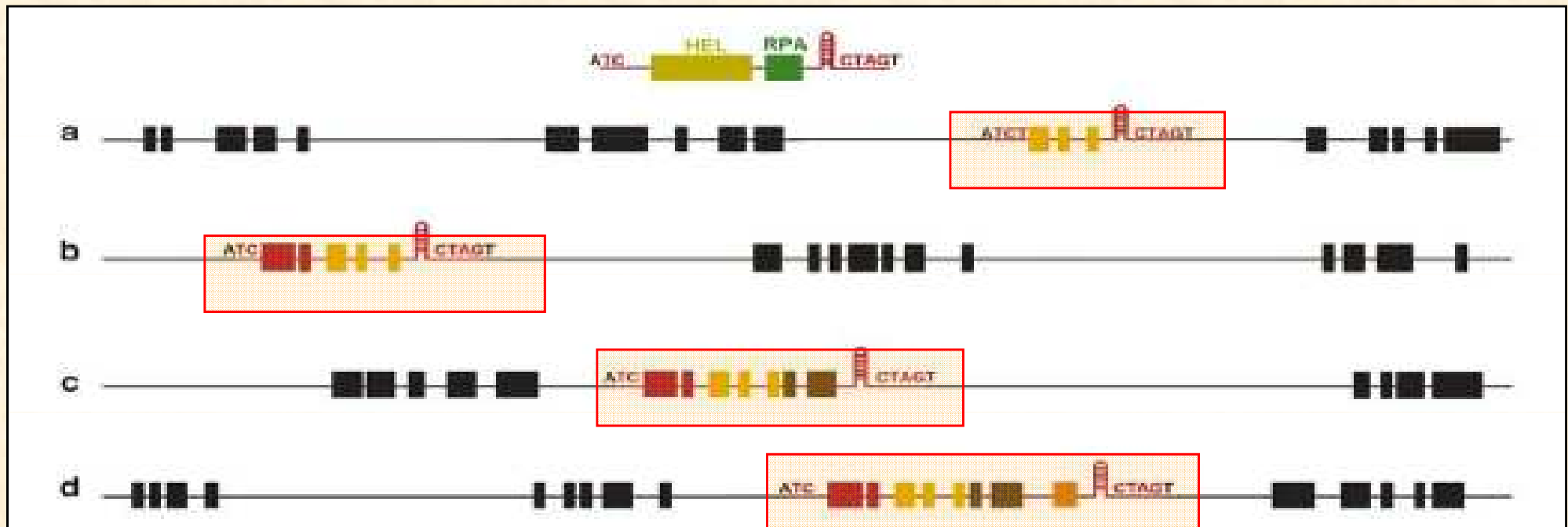
## Treasures in the attic: Rolling circle transposons discovered in eukaryotic genomes

Cédric Feschotte and Susan R. Wessler\*

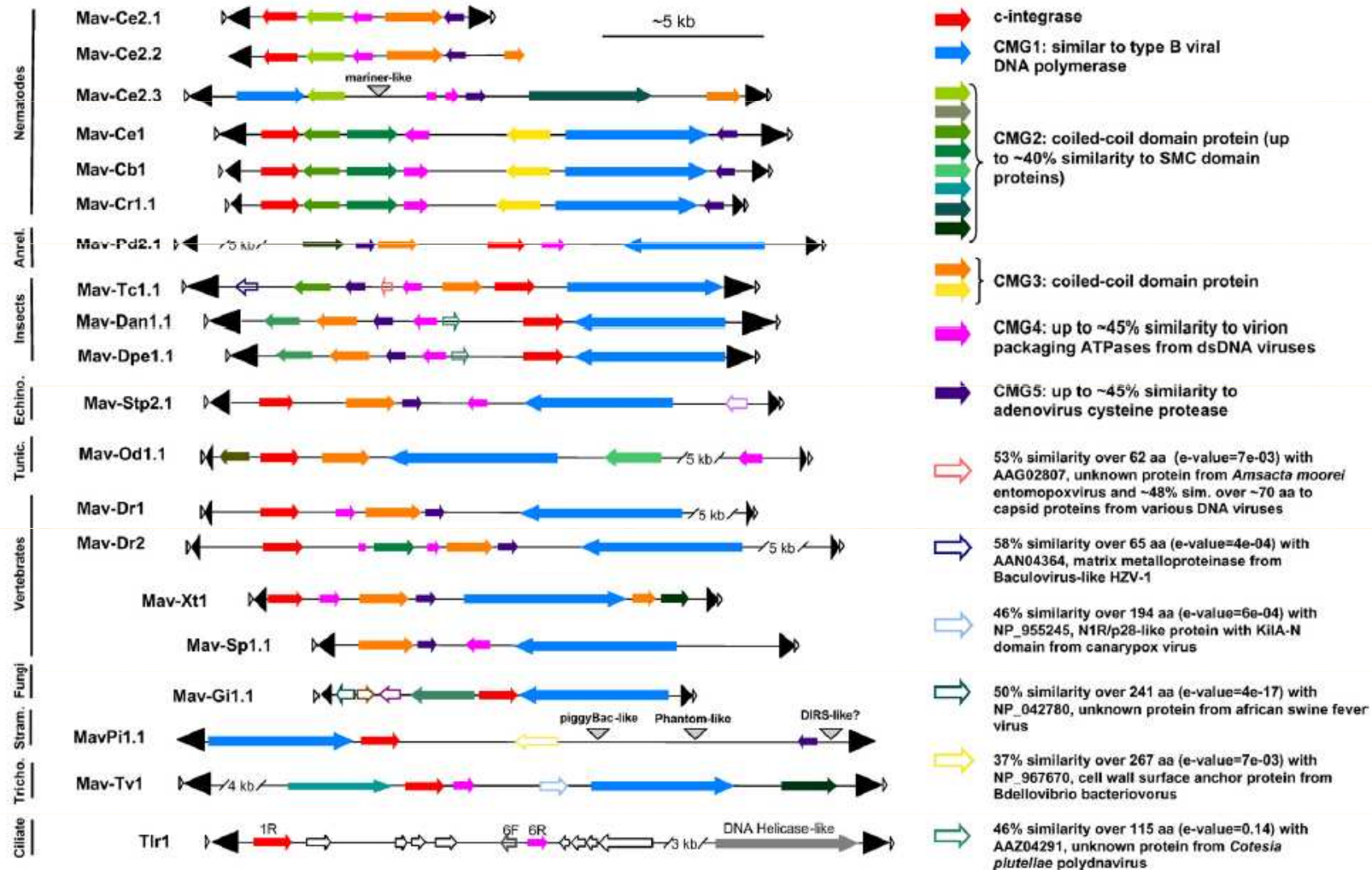
Departments of Botany and Genetics, University of Georgia, Athens, GA 30602



Journey of Helitrons (non-autonomous) through maize genome and capturing genes:



# Maverick - obří transposon příbuzný DNA virům



E.J. Pritham et al. / Gene 390 (2007) 3-17

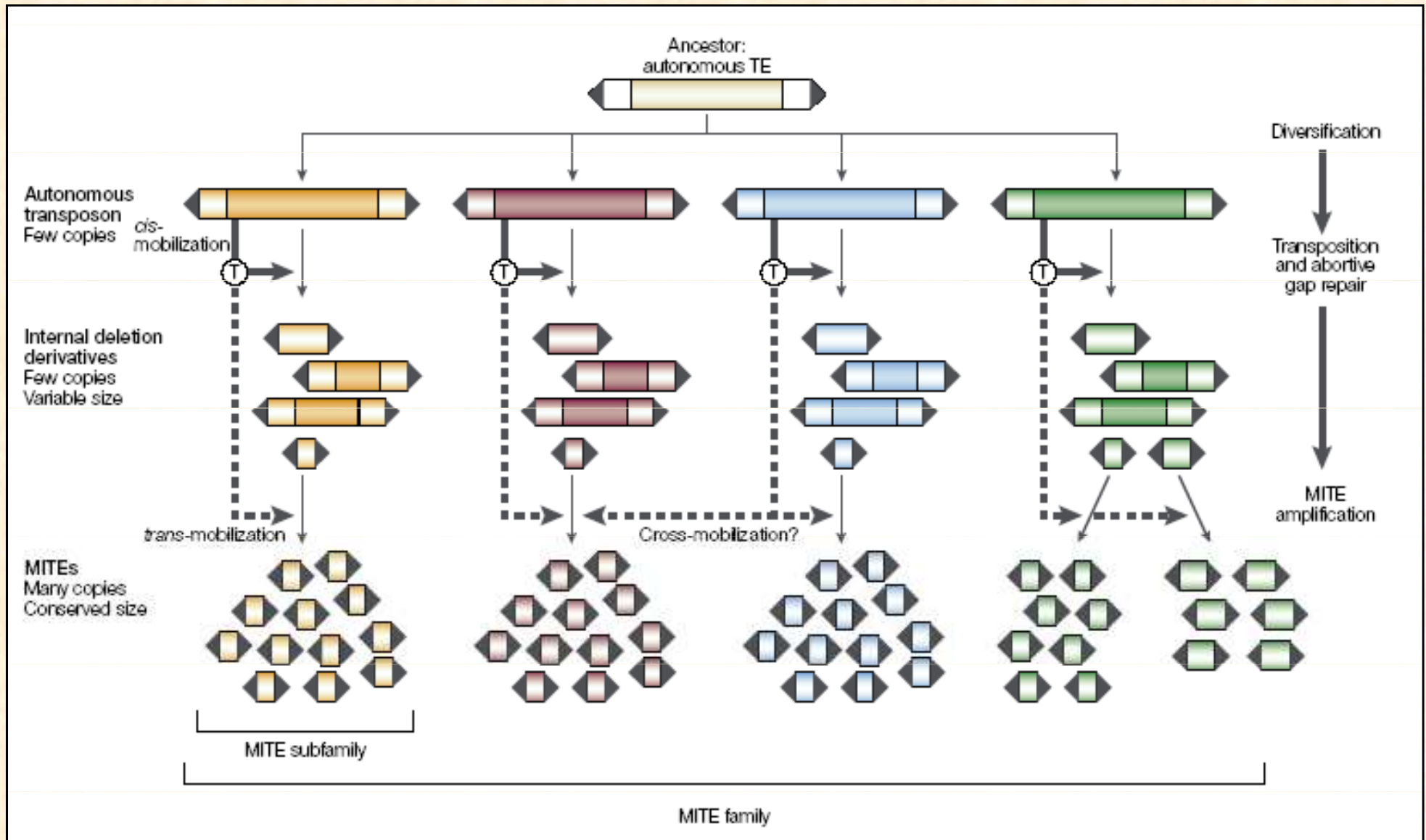
Fig. 1. Coding capacity and genetic organization of *Mavericks*. The individual elements are labeled with the initials of the genus and species, except the previously described element *Tlr1* from *Tetrahymena thermophila*. Ce = *Caenorhabditis elegans*, Cb = *C. briggsae*, Cr = *C. remanei*, Pd = *Platynereis dumerilii*, Tc = *Tribolium castaneum*, Da = *Drosophila ananassae*, Dpe = *D. persimilis*, Stp = *Strongylocentrotus purpuratus*, Od = *Oikopleura dioica*, Dr = *Danio rerio*, Xt = *Xenopus tropicalis*, Sp = *Sphenodon punctatus*, Gi = *Glomus intraradices*, Pi = *Phytophthora infestans* and Tv = *Trichomonas vaginalis*. The black triangles represent the terminal-inverted repeats and the small open triangles the 5- or 6-bp TSD, when identified. Related ORFs are color-coded and a description of each ORF is given in the key. Solid colored arrows represent CMGs, which are defined as ORFs present in *Maverick* elements from multiple species. Arrows not shaded represent ORFs present in *Mavericks* from a single species that are homologous to viral proteins. When secondary insertions of TEs were detected, their position is indicated and the type of nested TEs is given.



# MITE elementy

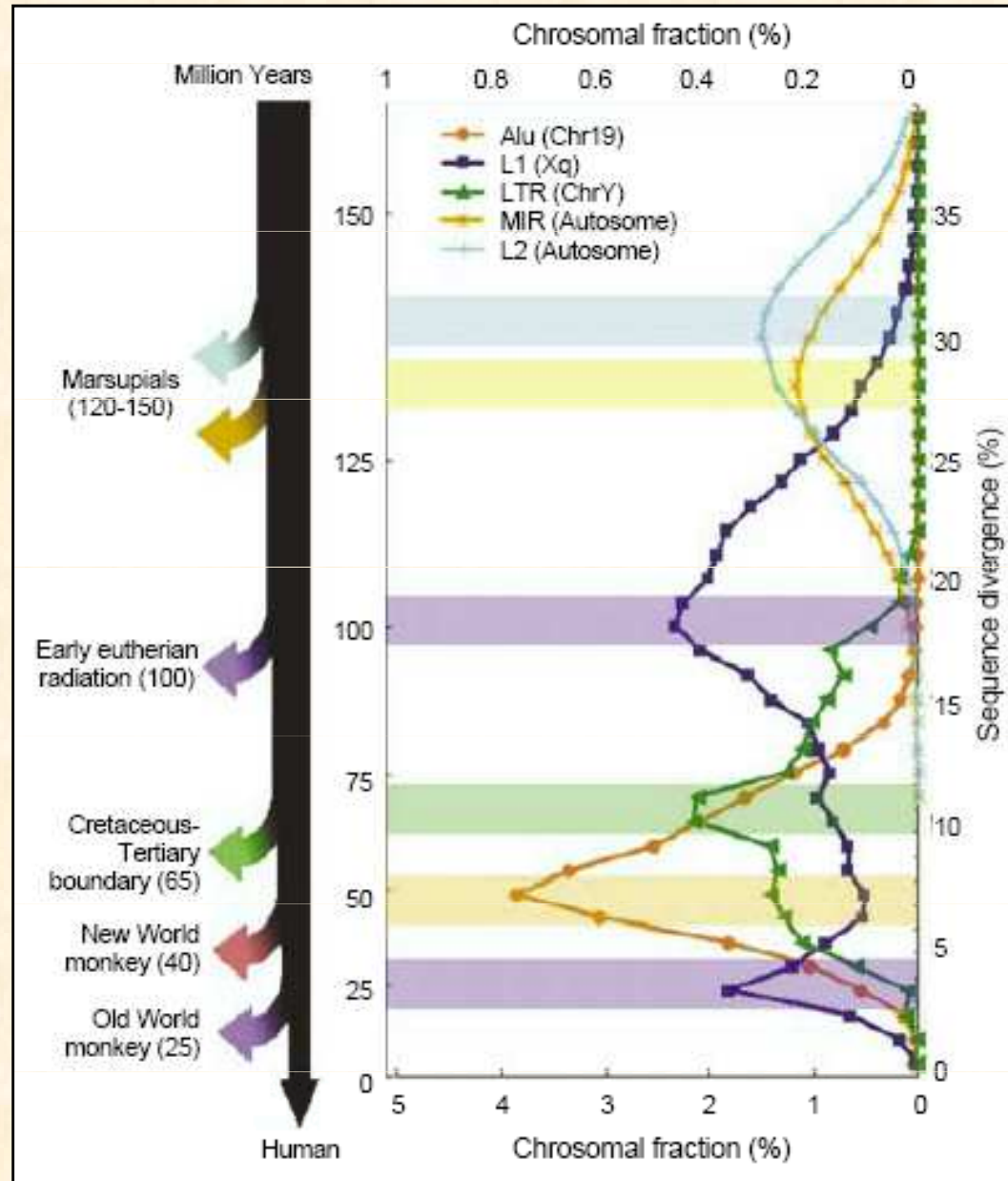
- rostlinné ekvivalenty lidských Alu
- neautonomní elementy (master = DNA TE Mariner)
- malá velikost: 125 - 500bp
- obrácené koncové repetice (TIR) - konzervativní 10-15 bp
- AT-bohaté (~72% *Stowaway*)
- tvoří sekundární struktury DNA (hairpins)
- preference cílového místa - TA(A)
- asociace s geny - v intronech, poblíž 5' nebo 3' konců genů
- rodiny:
  - *Stowaway* (jednoděložné, dvouděložné, živočichové)
  - *Tourist* (trávy)
  - *Emigrant, Alien, Heartbreaker, Bigfoot, ...* (rostliny)

# MITE elementy

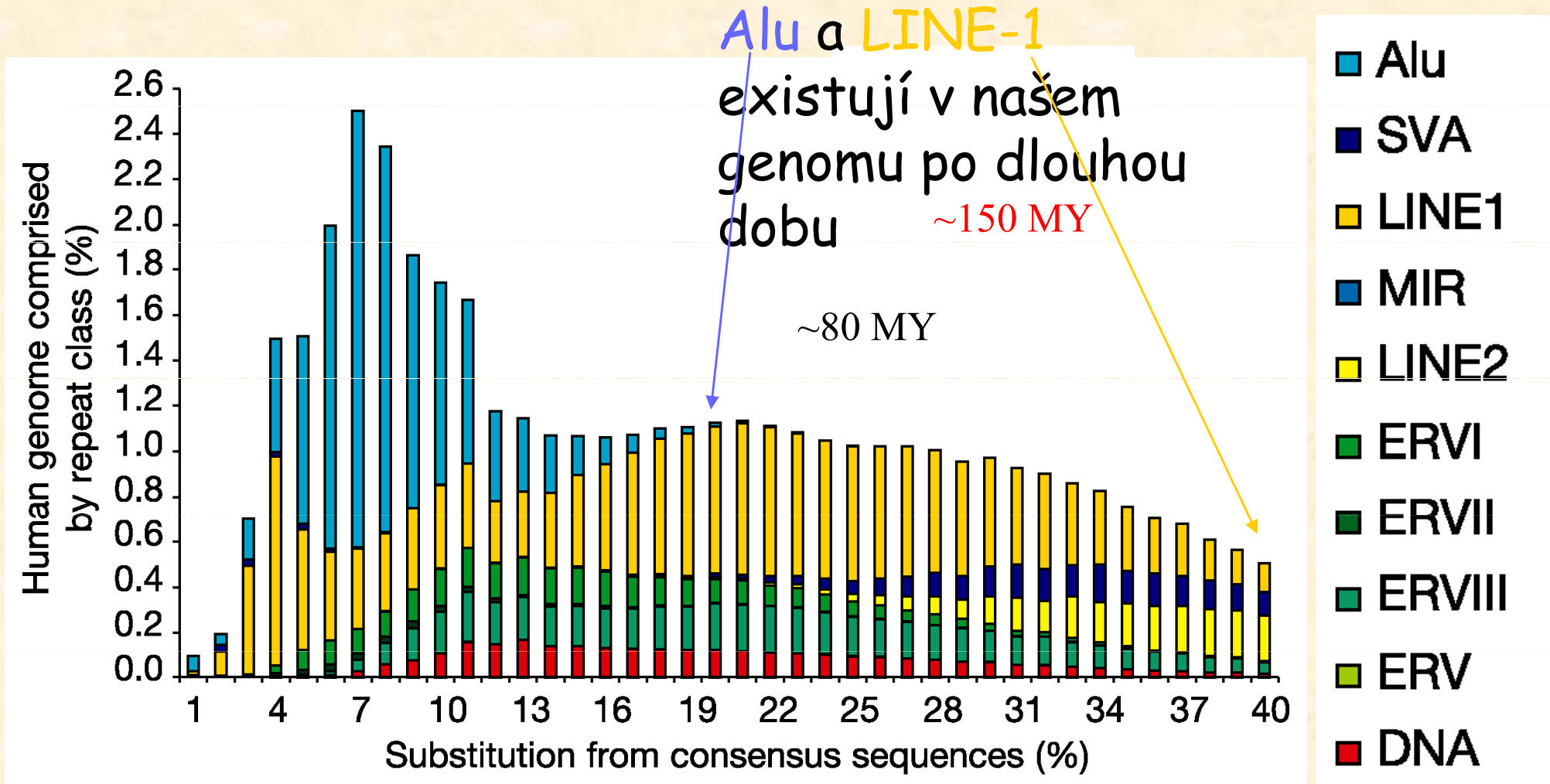


**EXPLOSIVNÍ AMPLIFIKACE  
TRANSPOSONŮ V EVOLUCI  
HOMINIDŮ**

# Korelace mezi expanzemi transposonů a oddělováním větví savců



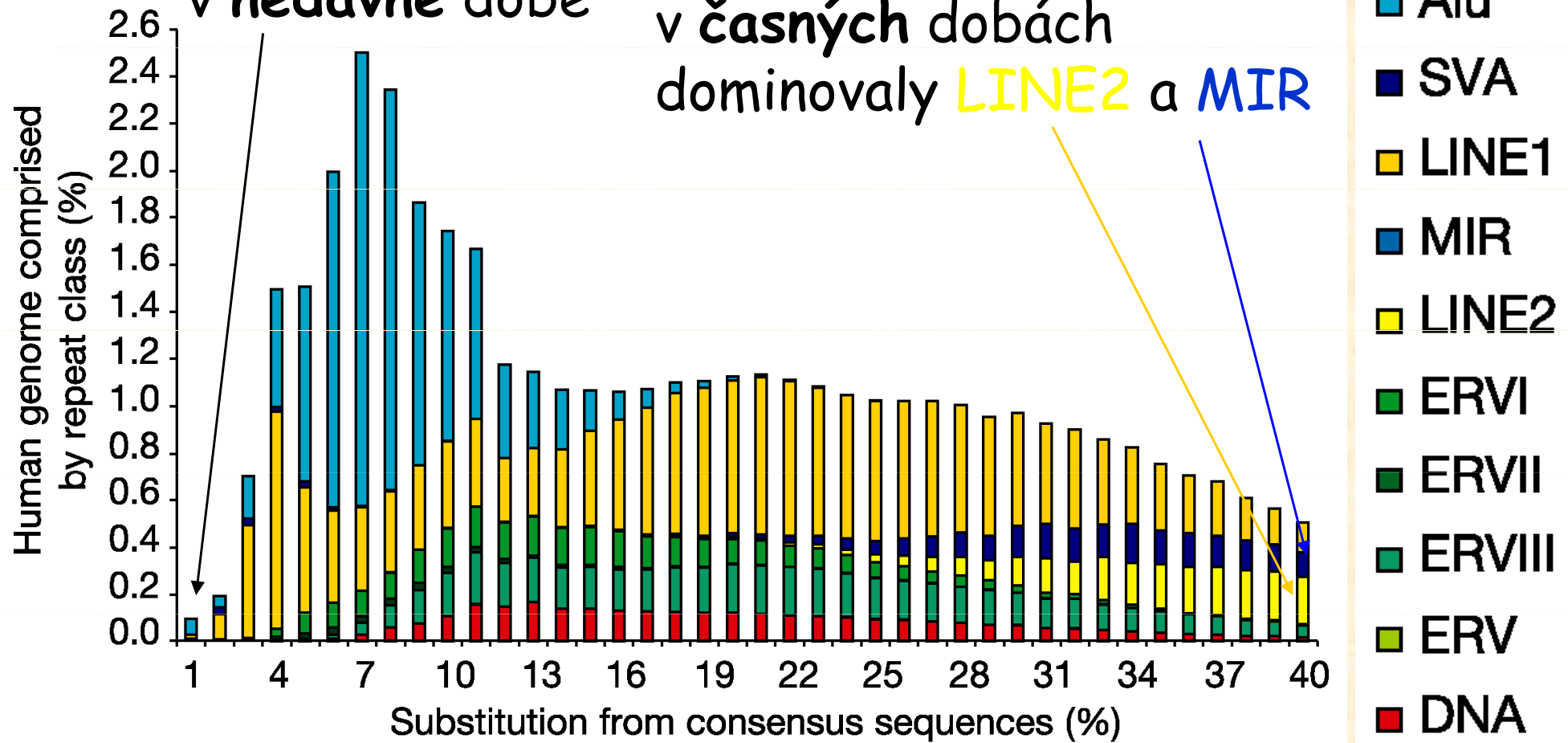
# Korelace mezi expanzemi transposonů a oddělováním větví savců



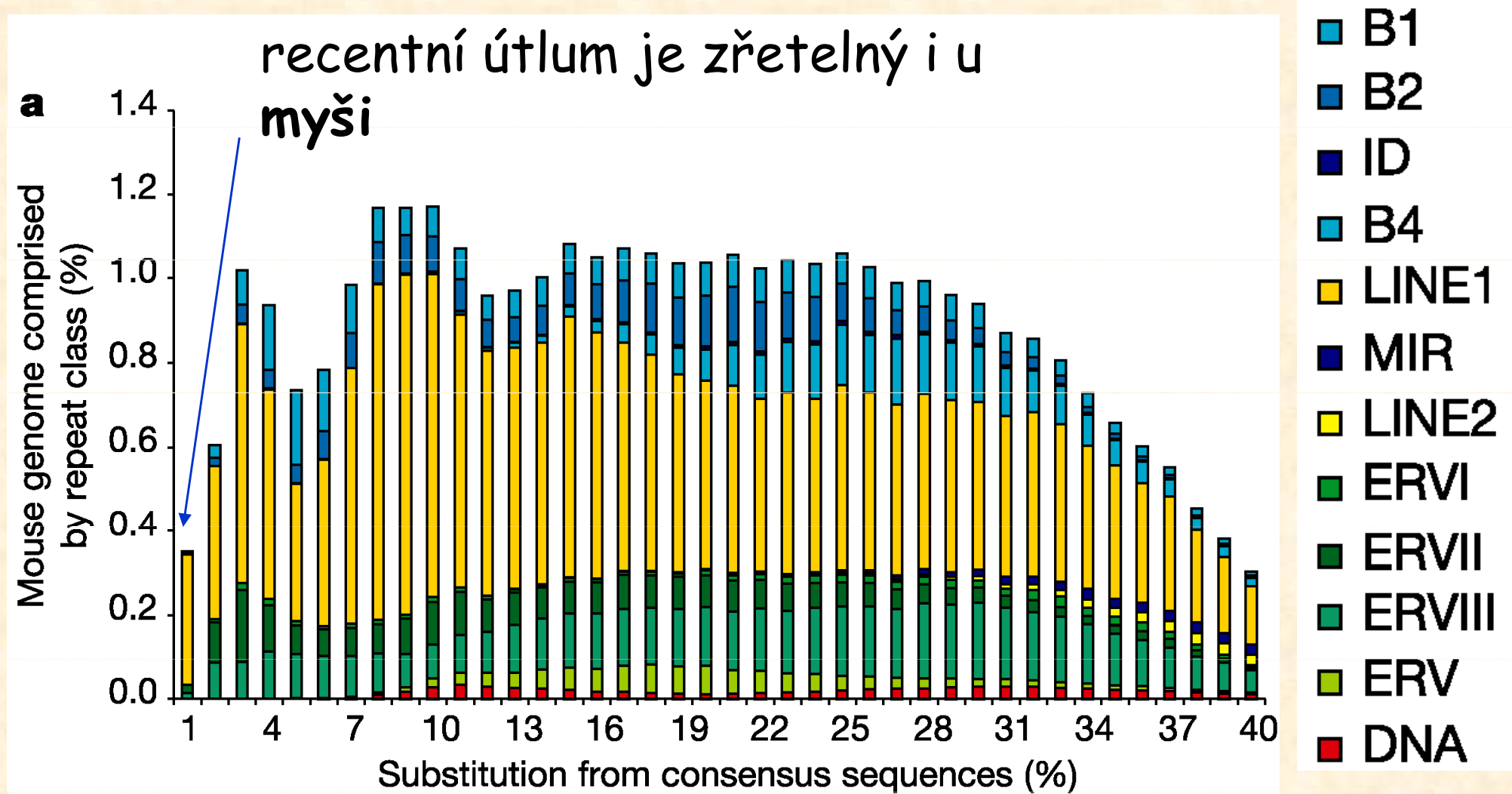
# Korelace mezi expanzemi transposonů a oddělováním větví savců

útlum transpozice  
v nedávné době

v časných dobách  
dominovaly **LINE2** a **MIR**



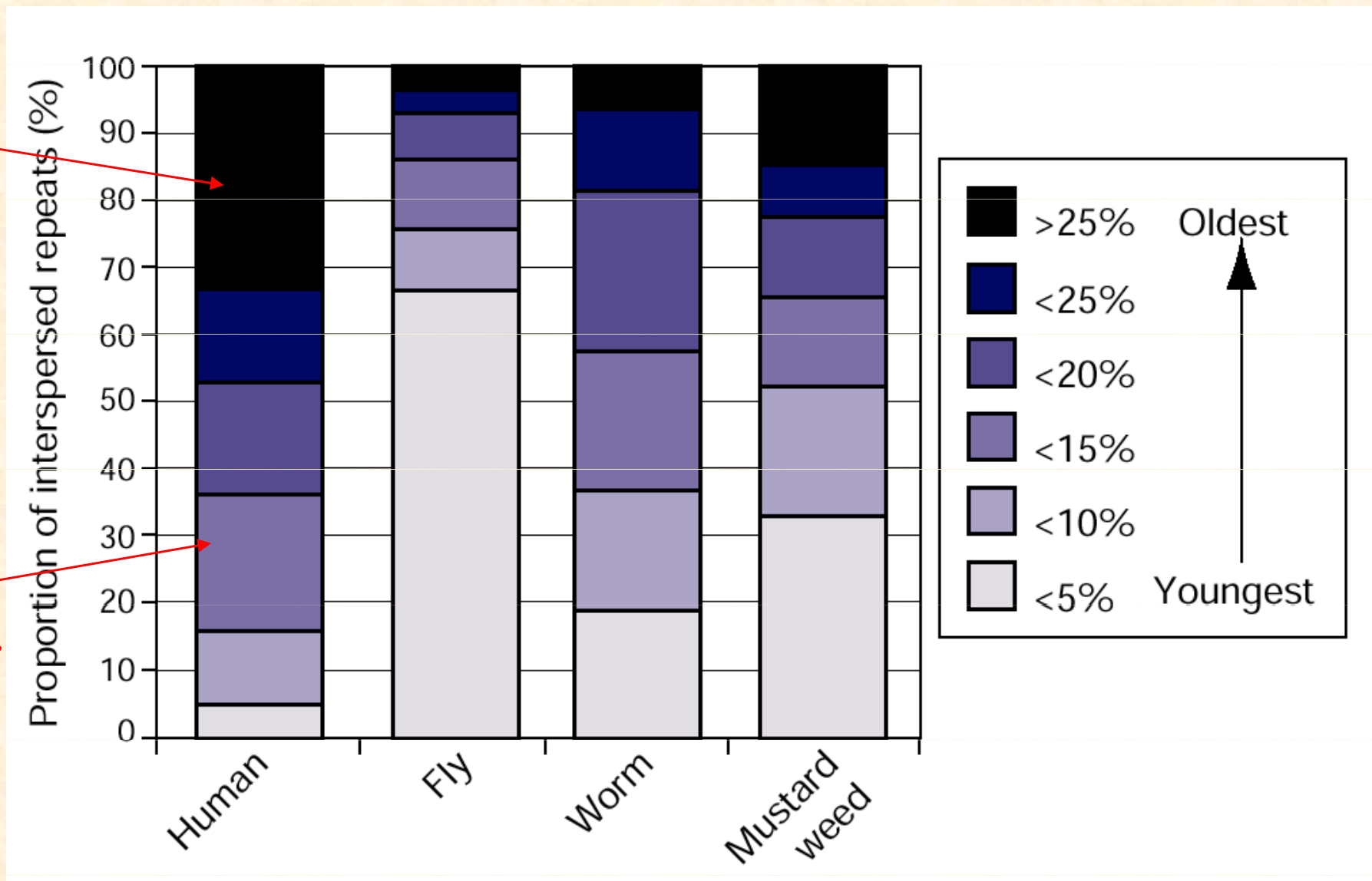
# Dynamika transposonů v evoluci *myši*



# Lidský genom je plný **starých** transposonů zatímco transposony v jiných genomech jsou mladší

staré

mladé



Srovnání stáří transpozonů v eukaryotických genomech



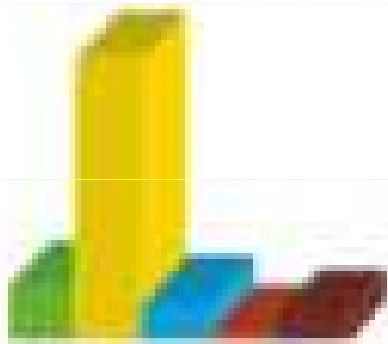
# TANDEMŮVÉ REPETICE

# Tandemové repetice, „knihovna“ satelitní DNA

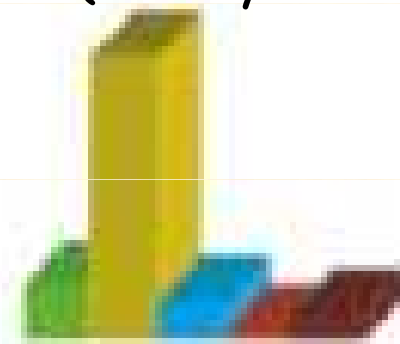
Đ. Ugarković and M. Plohl

Hlavní parametry:

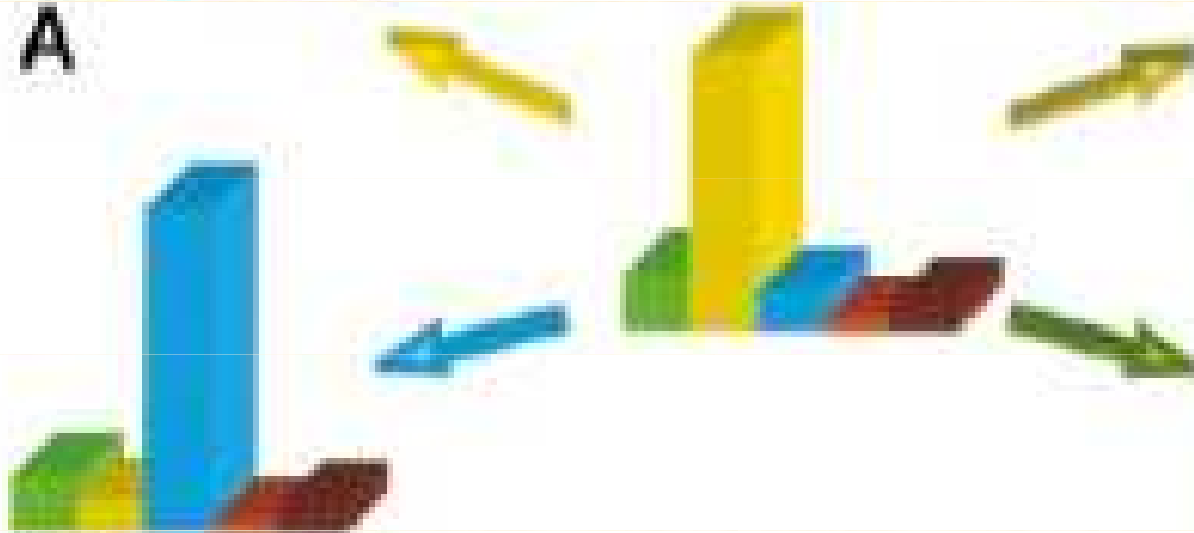
- počet kopií (změna v B a D)
- sekvence DNA (změny C a D)



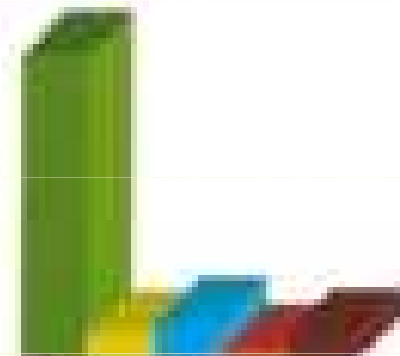
A



C



B



D

# Evolve tandemových repeticií

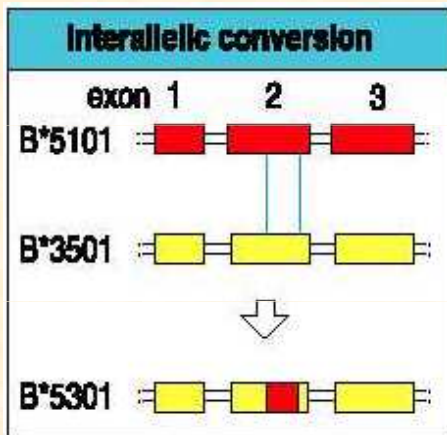
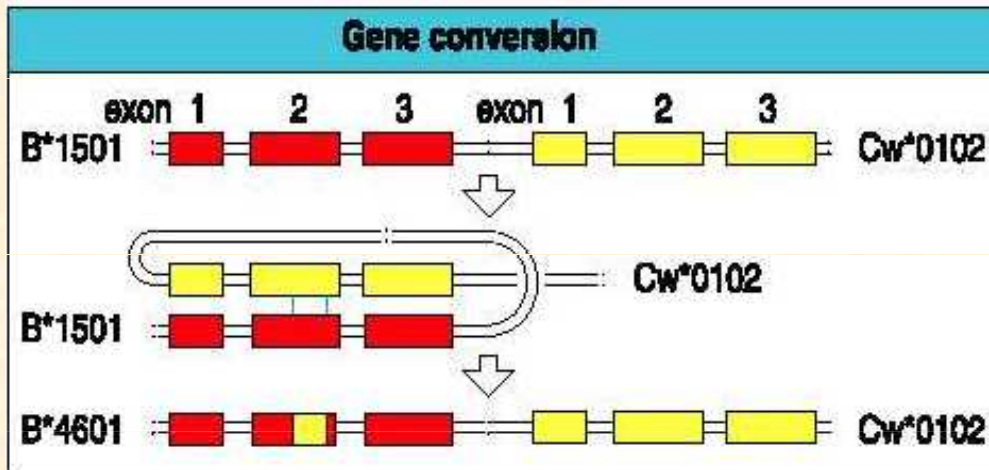
Evolve v koncertu (concerted evolution)

**Genová konverze**

Molekulární tah (molecular drive)

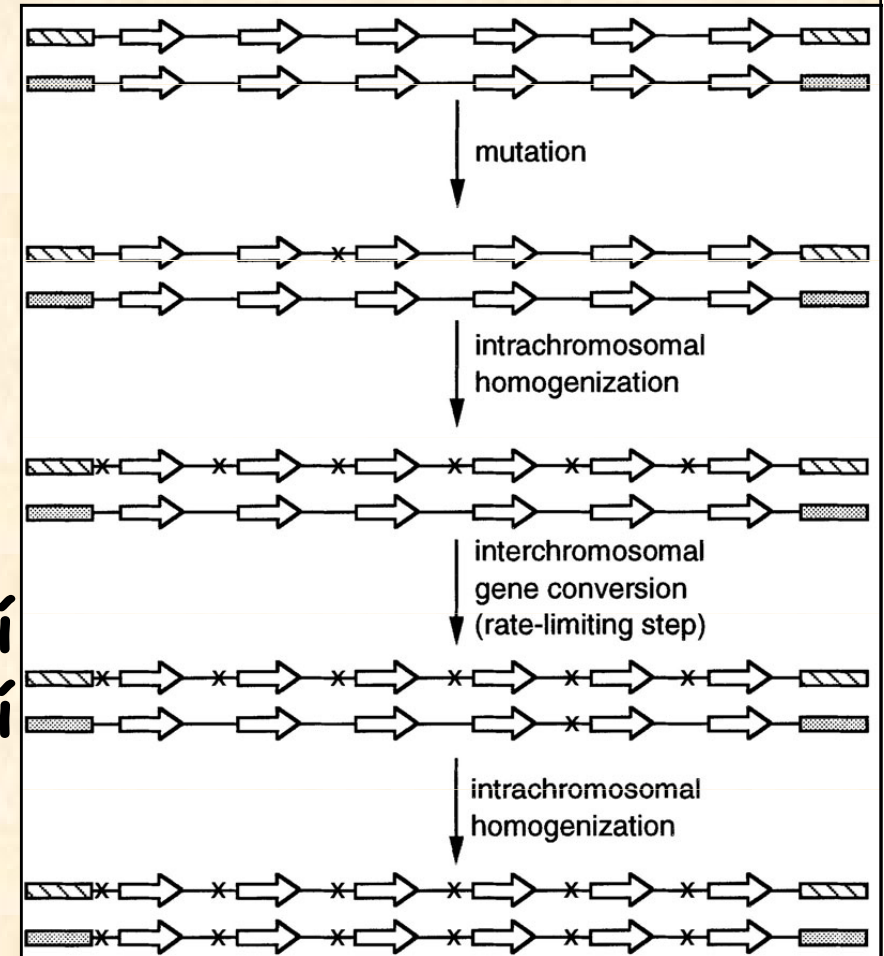
Nerovnoměrný crossing-over

Evolve satelitních sekvencí - skládání ze segmentů

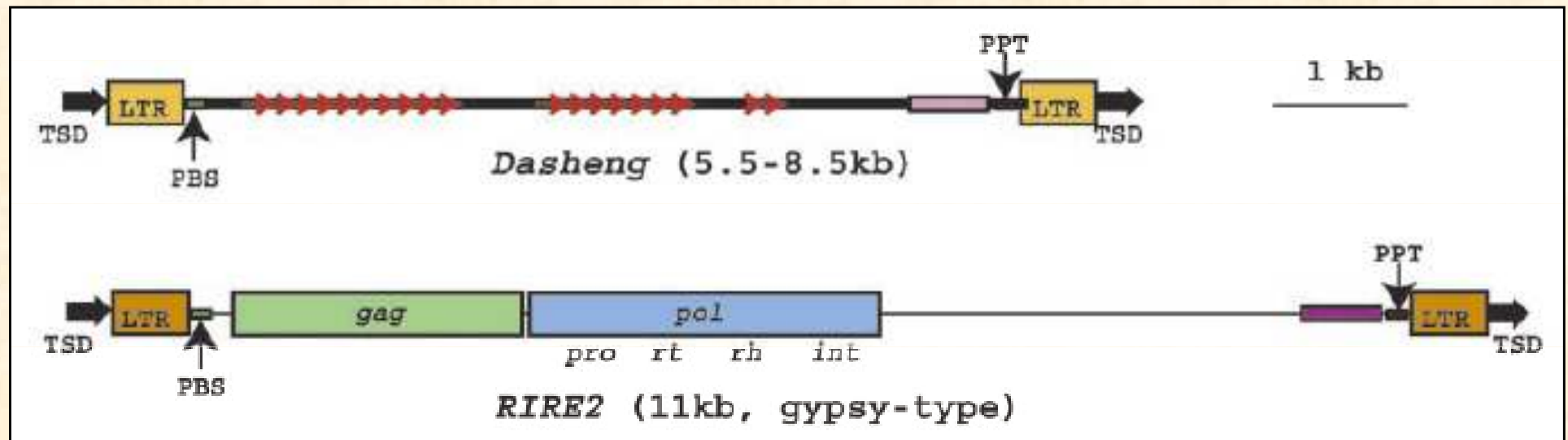
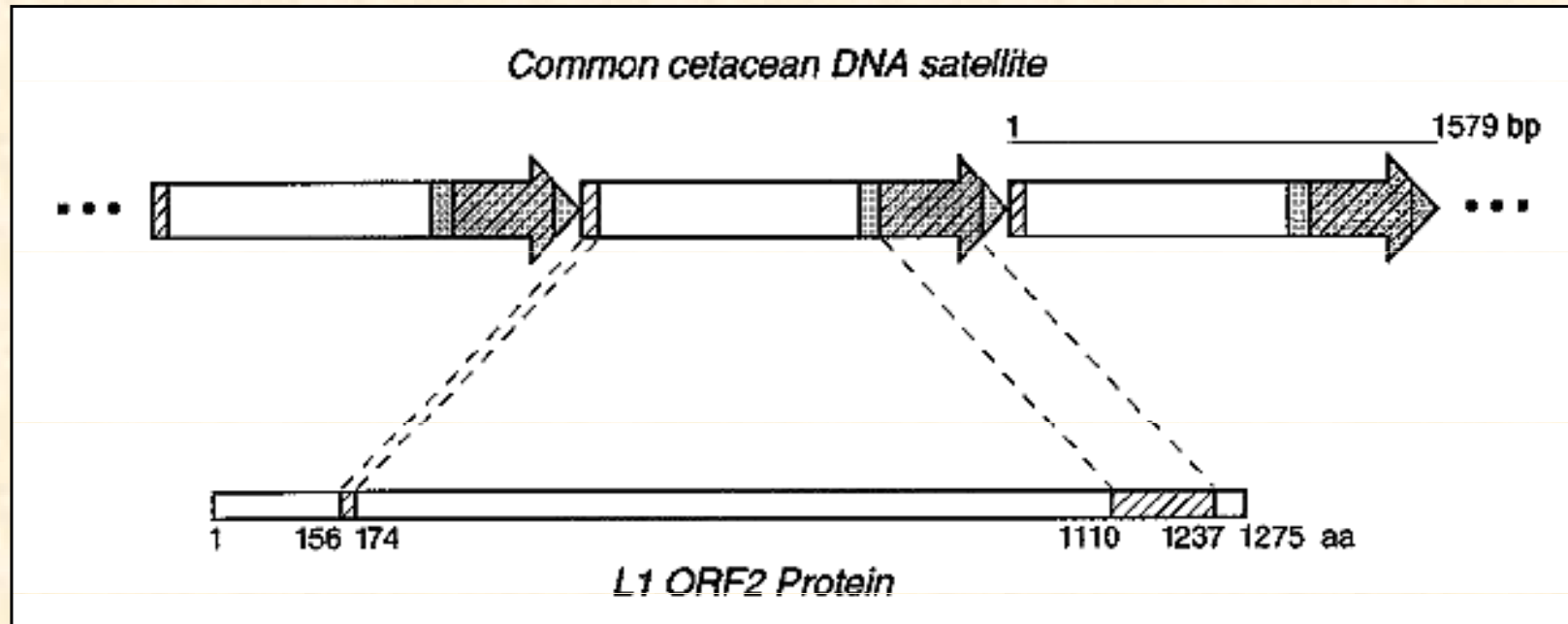


**Genová konverze**

- interchromosomální
- intrachromosomální



# Satelity mohou vznikat i z retroelementů



# PROMISKUITNÍ DNA

# 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.

**“Promiscuous DNA” (Ellis, 1982)**

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

Nature Reviews Genetics, 2004

# Organelové genomy - pozůstatky prokaryot

## (a) chloroplast

20-200 kb

20-200 proteinů

progenitor - **cyanobacteria** (*Synechocystis*)

3.6 Mb

3000 proteinů

## (b) mitochondrie

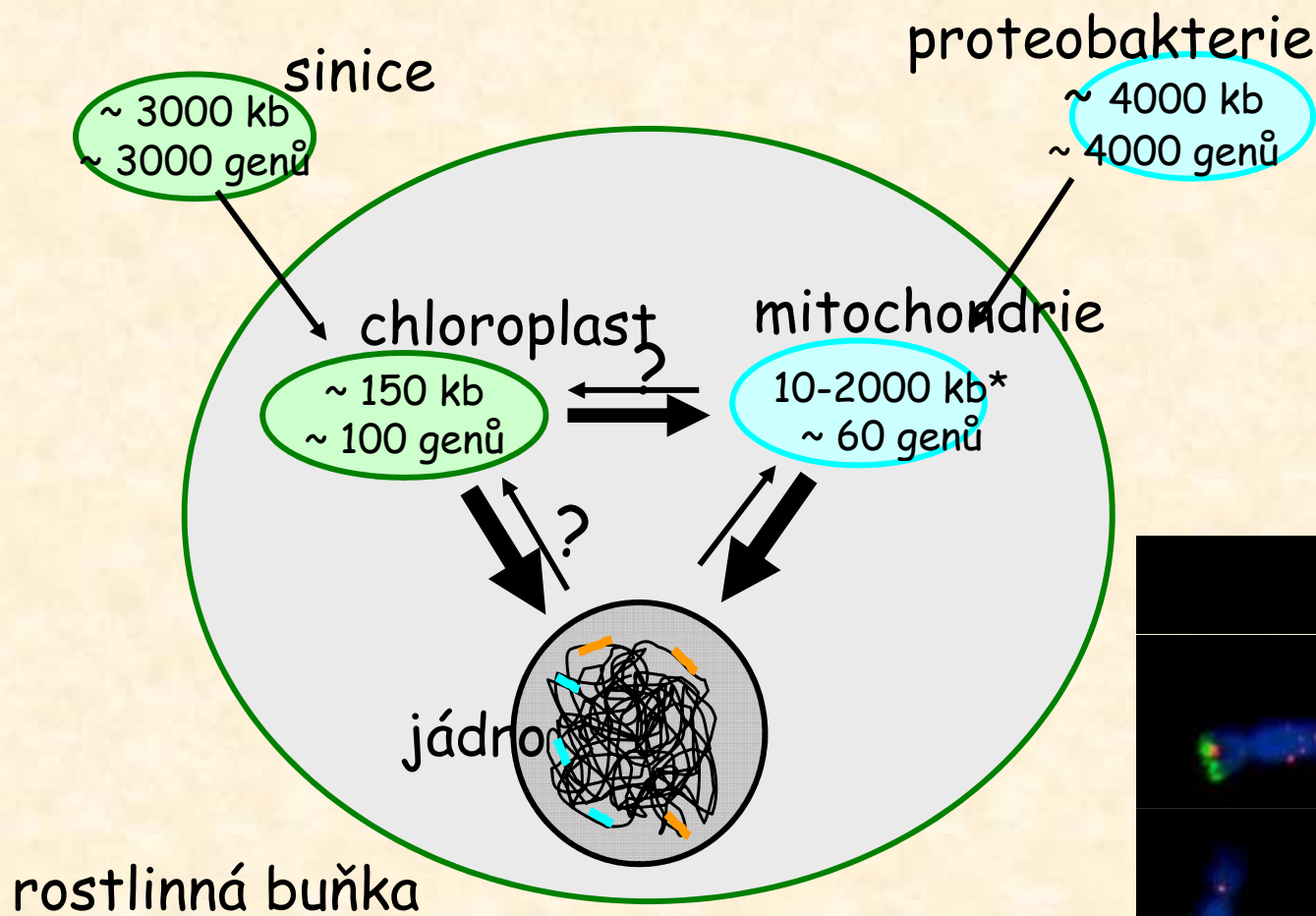
6-400 kb

3-67 proteinů

progenitor - **alpha-proteobacteria** (*Mesorhizobium loti*)

7 Mb

6 700 proteinů



```

BAC7H5 CCTTCGGGAACGCGGACACAGGTGGTGCATGGCTGTCGTCAGCTCGTCCGCTAAGGTGTT 43358
CTY16 -----G-C-T-----C-----T-----T-G-A----- 457

BAC7H5 GGGTTAAGTCCCGCAACGAGCGCAACCCCTCGTGTTTAGTTGCCAACGTTGAGTTGGAAC 43418
CTY16 -----ACC-----T-A-T-----G-C----- 397

BAC7H5 CCTGAACAGACTGCCGGTGATAAGCCGGAGGAGGTGAGGATGACGTCAAGTCATCATGC 43478
CTY16 TT-A-GGTA-C-----G-----G-----C----- 337

BAC7H5 CCCTTATGCCCTGGGCGACACACGTGCTACAATGGCCGGGACAAAGGTCGCGATCCGCG 43538
CTY16 -----C-G-----T-----G-T-----GT--CA--A-G-A-- 277

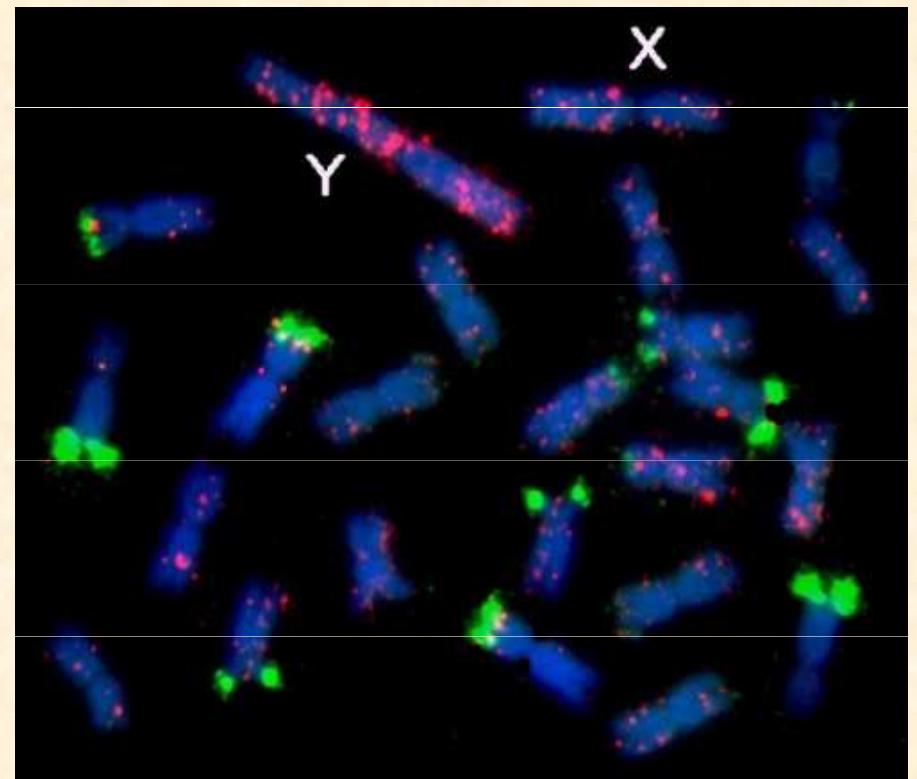
BAC7H5 GAGGGTGAGCTAACCCCAAAAACCCGTCTCAGTTCGGATTGCAGGCTGCAACTCGCCTG 43598
CTY16 --T-----T-T-C-----G-----TTC-T-----AGA- 218

BAC7H5 CATGAAGCCGGAATCGCTAGTAATCGCCGGTCAGCCATACGGCGGTGAATTCGTTCCCGG 43658
CTY16 -----G-----G-A-----G-C-----A----- 159

BAC7H5 GCCTTGTACACCCCGCTCACACTATGGGAGCTGGCCATGCCCGAAGTCGTTACCTTA 43718
CTY16 -----C-----ATTCA-----G--G-C-- 99

BAC7H5 AC-CGCAAG-GAGGGGGATGCCGAAGGCAGGCTAGTGACTGGAGTGAAGTCGTAAACAAG 43776
CTY16 --T-----A---CA-GC-A-C-C--TG--TT--C-----G----- 39

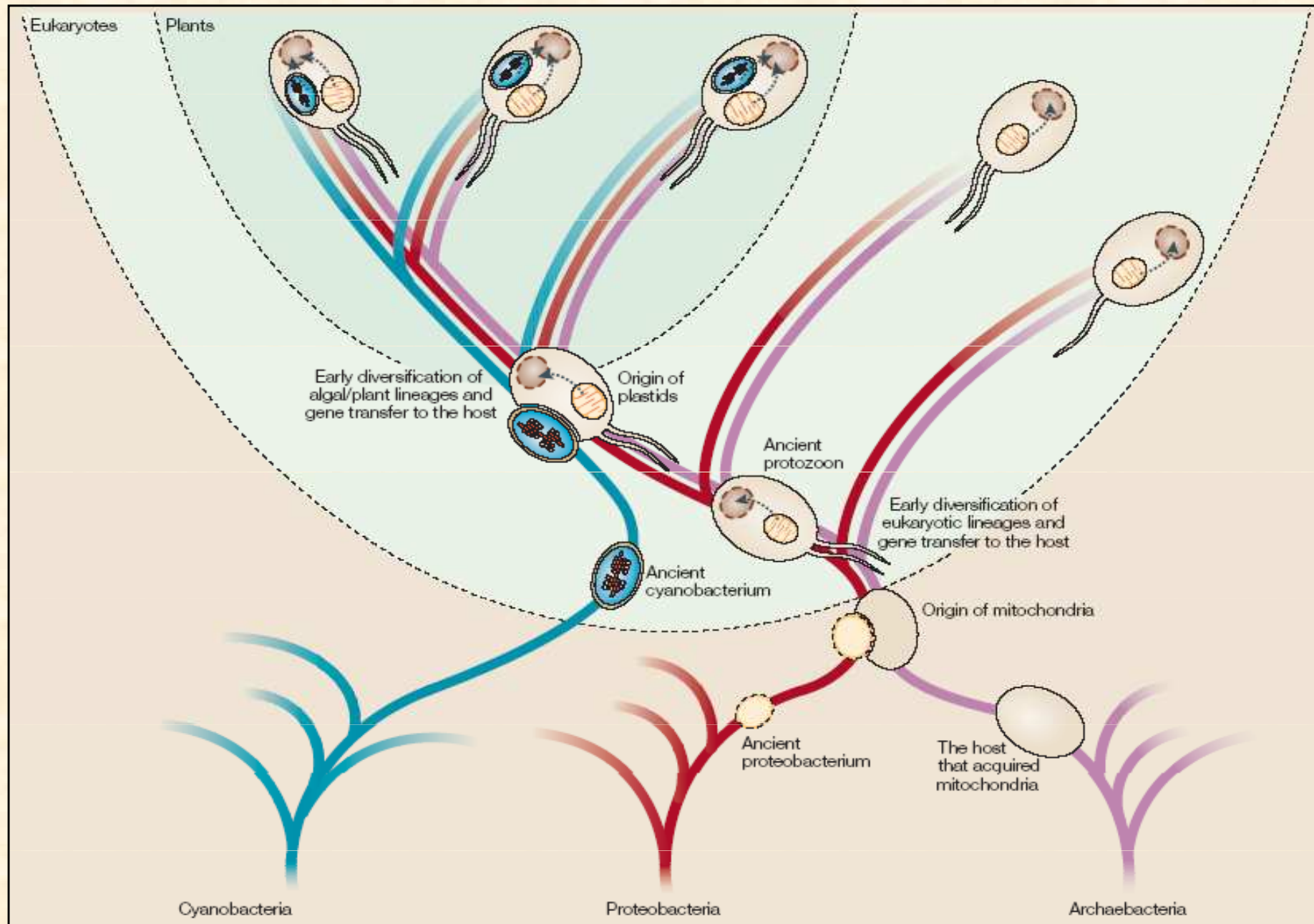
BAC7H5 GTAGCCGCTGGAAGGTGCGGCTGGATCACCTCCTTT 43814
CTY16 -----GG---CC----- 1
  
```



Hromadění cpDNA  
na chromosomu Y u  
*Silene latifolia*



# Endosymbiotická evoluce a strom života



# Velikosti organelových a prokaryotických genomů

Genome	Length [kbp]	Number of protein-coding genes
<b>Algae</b>		
cp <i>Porphyra purpurea</i>	191	200
cp <i>Cyanidium caldarium</i>	165	197
cp <i>Guillardia theta</i>	122	148
cp <i>Cyanophora paradoxa</i>	136	136
cp <i>Odontella sinensis</i>	120	124
cp <i>Euglena gracilis</i>	143	58
<b>Land plants</b>		
cp <i>Marchantia polymorpha</i>	121	84
cp <i>Chlorella vulgaris</i>	151	78
cp <i>Nicotiana tabacum</i>	156	76
cp <i>Oryza sativa</i>	134	76
cp <i>Zea mays</i>	140	76
cp <i>Pinus thunbergii</i>	120	69
<b>Non-phosynthetic plastids</b>		
cp <i>Toxoplasma gondii</i>	35	26
cp <i>Eimeria tenella</i>	35	28
cp <i>Epifagus virginiana</i>	70	21
<b>Cyanobacteria</b>		
<i>Synechocystis</i> sp.	3573	3168
<i>Prochlorococcus marinus</i>	1660	1884
<i>Nostoc</i> PCC 7120	6413	5368
<i>Nostoc punctiforme</i>	~9000	~7400
<b>Plants and algae</b>		
mt <i>Pylaiella littoralis</i>	59	52
mt <i>Marchantia polymorpha</i>	187	41
mt <i>Laminaria digitata</i>	38	39
mt <i>Cyanidioschyzon merolae</i>	32	34
mt <i>Arabidopsis thaliana</i>	367	31
mt <i>Chondrus crispus</i>	26	25
mt <i>Scenedesmus obliquus</i>	43	20

Genome	Length [kbp]	Number of protein-coding genes
<b>Various protists and fungi</b>		
mt <i>Reclinomonas americana</i>	69	67
mt <i>Melawimonas jakobiformis</i>	47	49
mt <i>Naegleria gruberi</i>	50	46
mt <i>Rhodomonas salina</i>	48	44
mt <i>Dictyostelium discoideum</i>	56	40
mt <i>Phytophthora infestans</i>	38	40
mt <i>Acanthamoeba castellanii</i>	42	36
mt <i>Cafeteria roenbergensis</i>	43	34
mt <i>Monosiga brevicollis</i>	77	32
mt <i>Physarum polycephalum</i>	63	20
mt <i>Harpochytrium</i> sp	24	14
mt <i>Candida albicans</i>	40	13
mt <i>Cryptococcus neoformans</i>	25	12
mt <i>Plasmodium falciparum</i>	6	3
<b>Anaerobic mitochondria</b>		
mt Hydrogenosomes*	0	0
<b><math>\alpha</math>-proteobacteria</b>		
<i>Caulobacter crescentus</i>	4017	3767
<i>Mesorhizobium loti</i>	7596	7281
<i>Bradyrhizobium japonicum</i>	~9100	~8300
<b>Yeast</b>		
(nuclear)	13,469	6,327

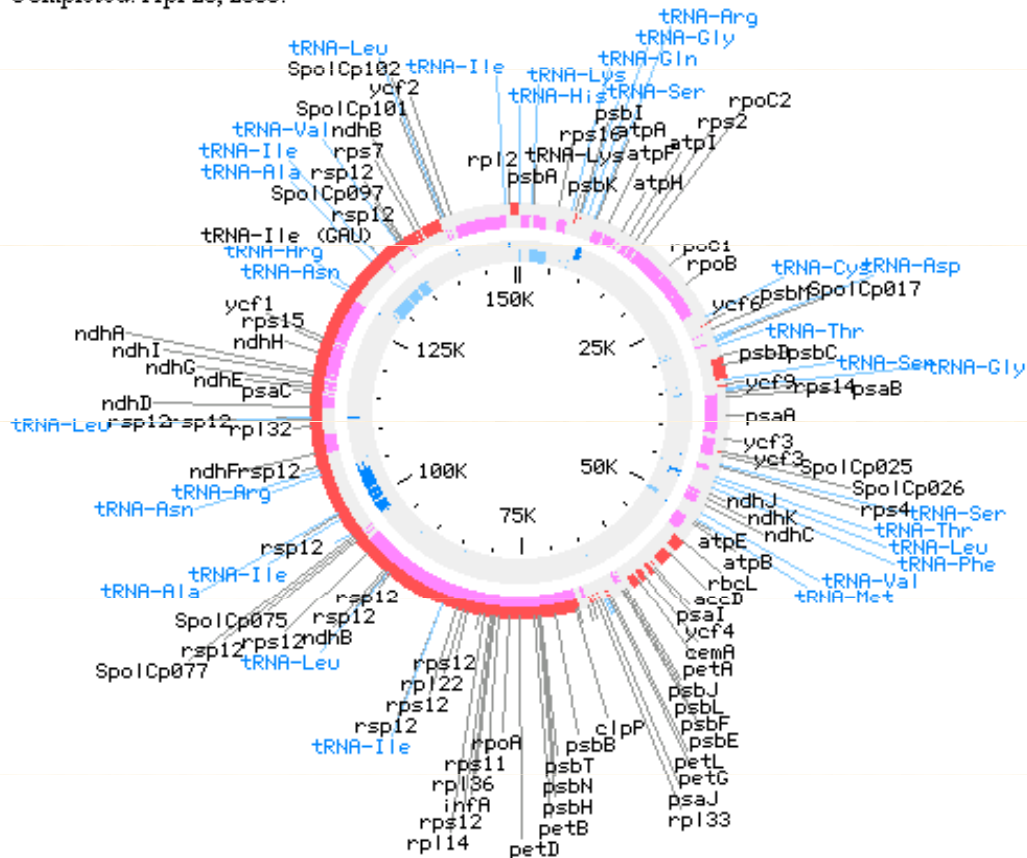
# Typický chloroplastový genom

## Spinacia oleracea plastid, complete genome

Accession: [NC\\_002202](https://www.ncbi.nlm.nih.gov/nuccore/NC_002202)

Total Bases Sequenced: 150725 bp

Completed: Apr 20, 2000.



### Legend:

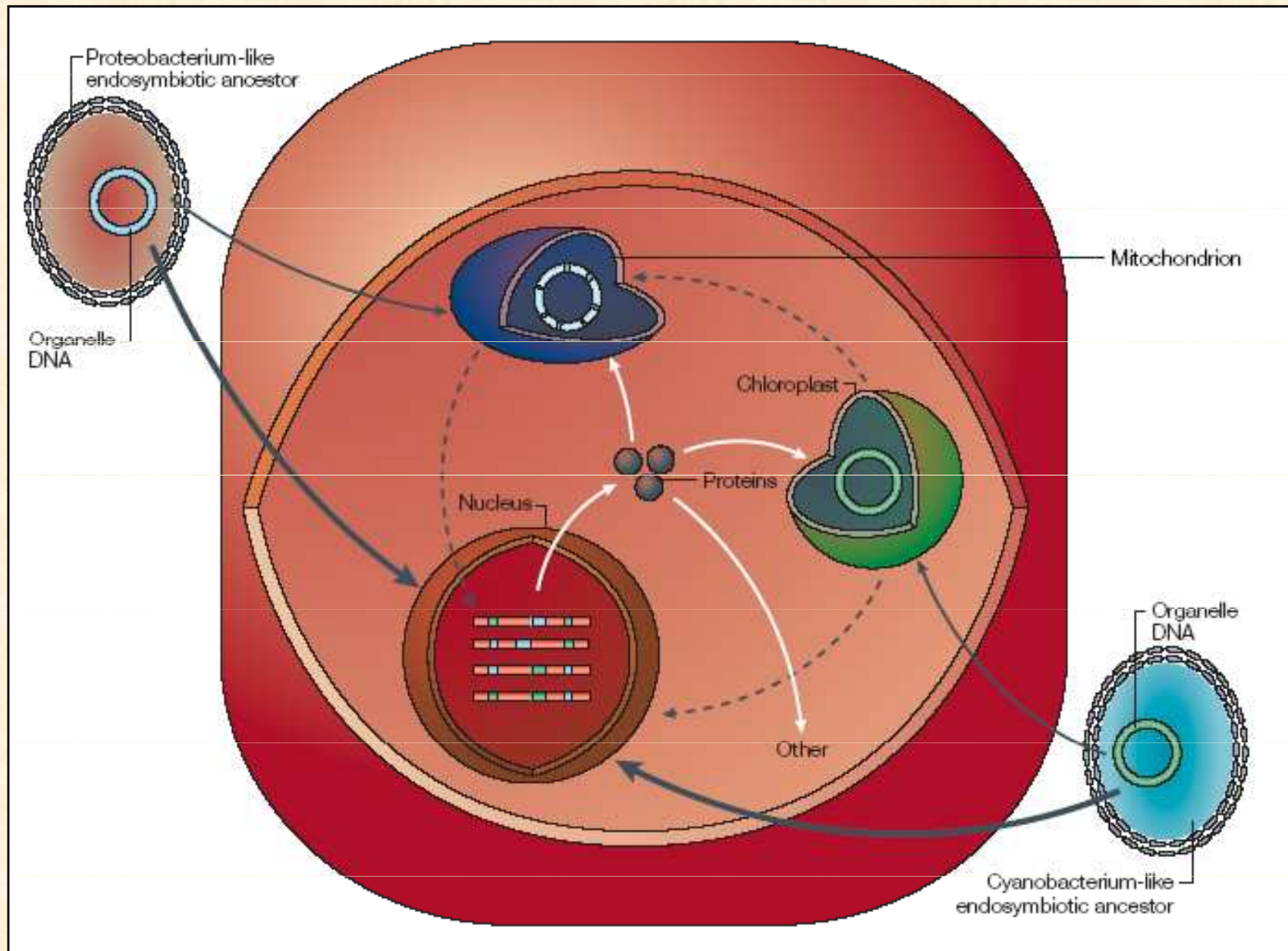
- - CDS +strand
- - RNA +strand
- - CDS -strand
- - RNA -strand

velikost: ~ 150 kb  
 LSC (large single copy) - 80 kb  
 SSC (small single copy) - 20 kb  
 IR (inverted repeats) - 25 kb

118 genů:  
 85 proteiny  
 photosystem I and II  
 cytochrome  
 ATP synthase  
 Rubisco  
 NADH dehydrogenase  
 ribosomal proteins  
 RNA polymerase

29 tRNA  
 4 rRNA

# Endosymbiotický genový přenos: - transport genů, reimport proteinů



# Mechanismy 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

# 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

# Některé geny se přenášejí do jádra dříve jiné později

**Nejdříve** - regulační funkce

(sigma factor of RNAPolymerase, gamma subunit  
of ATPase)

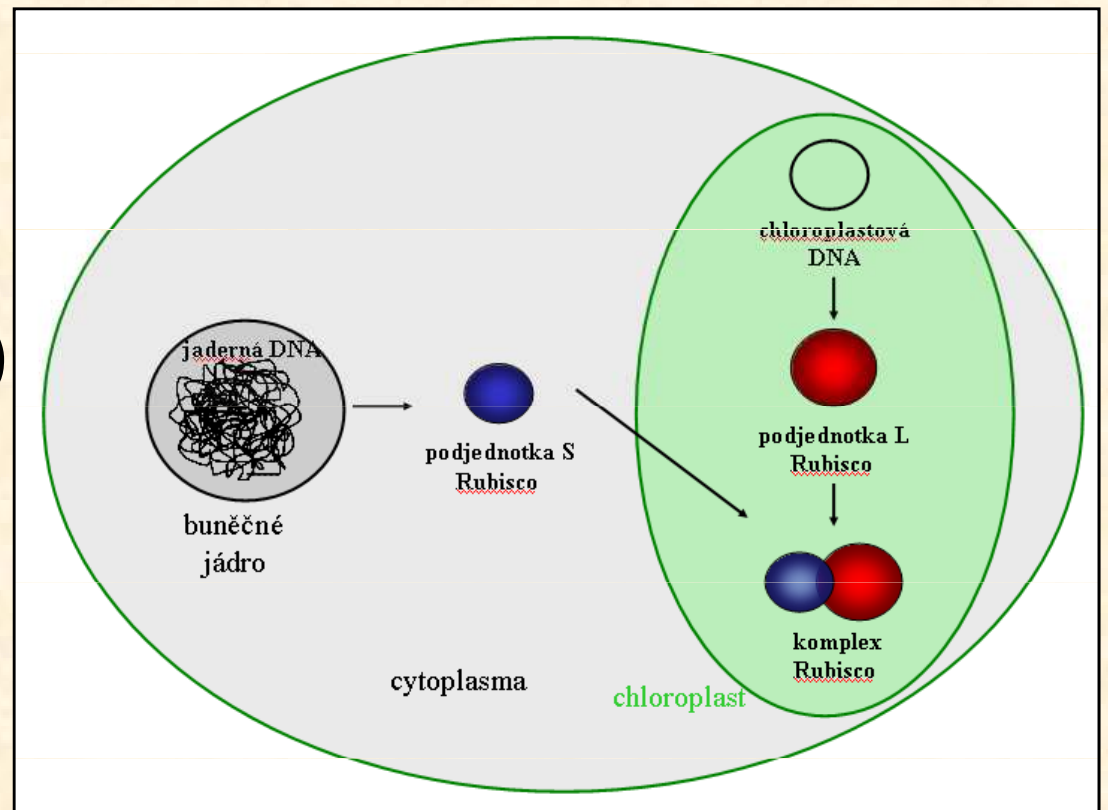
**Poslední** - translace

- respirace

**Rubisco:**

katalýza - v plastidu (rbcl)

regulace - v jádře (rbcS)



## **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



# Genový přenos z organel do jádra v reálném čase

## 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

