

**Eduard Kejnovský (Zdeněk Kubát) +
Roman Hobza**

EVOLUČNÍ GENOMIKA

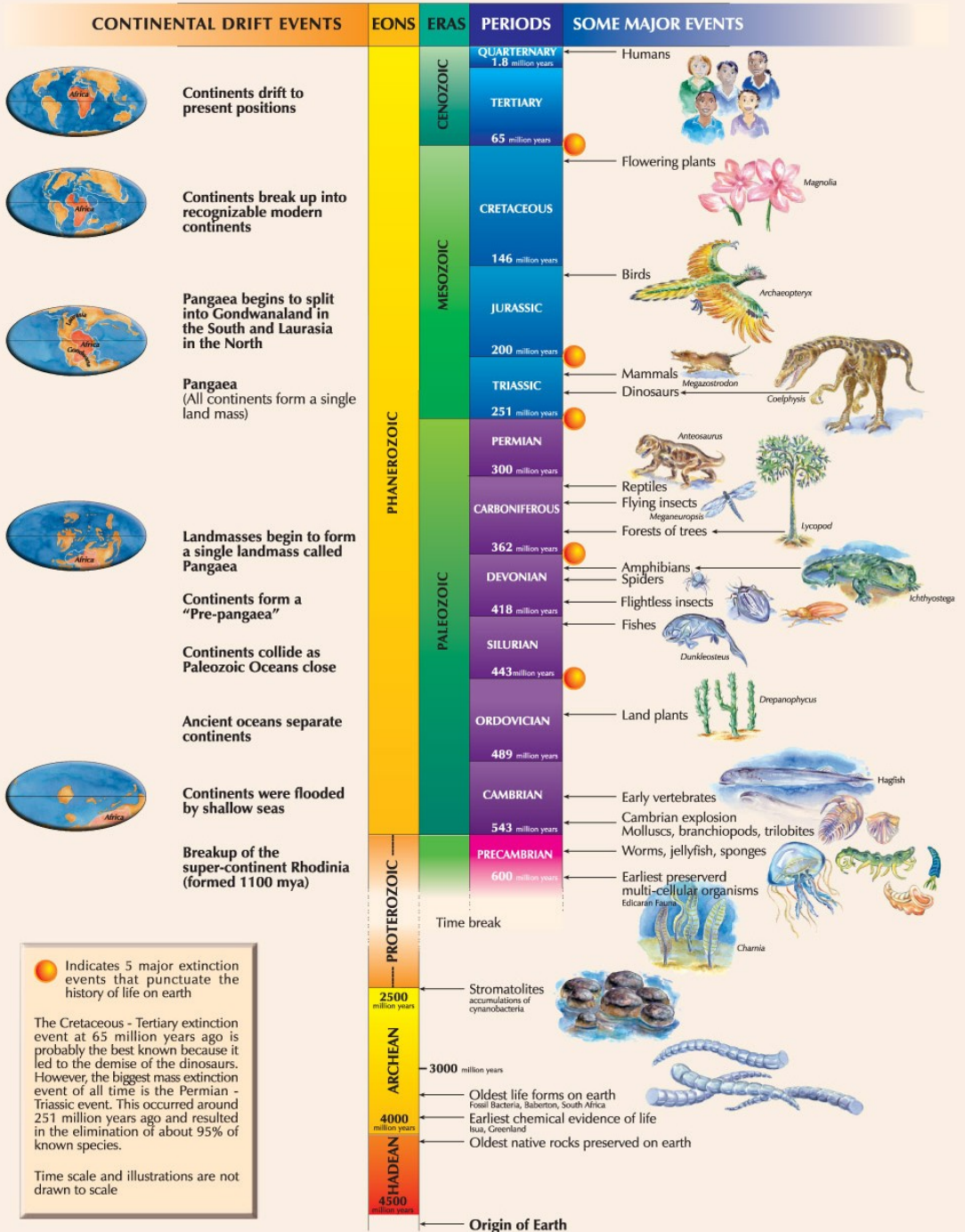
III. EVOLUCE GENOMŮ



OBSAH

1. Velikost genomu
3. Změny ve velikosti genomu
4. Nekódující DNA
- 5 .Topografie genomu
6. Genomové projekty

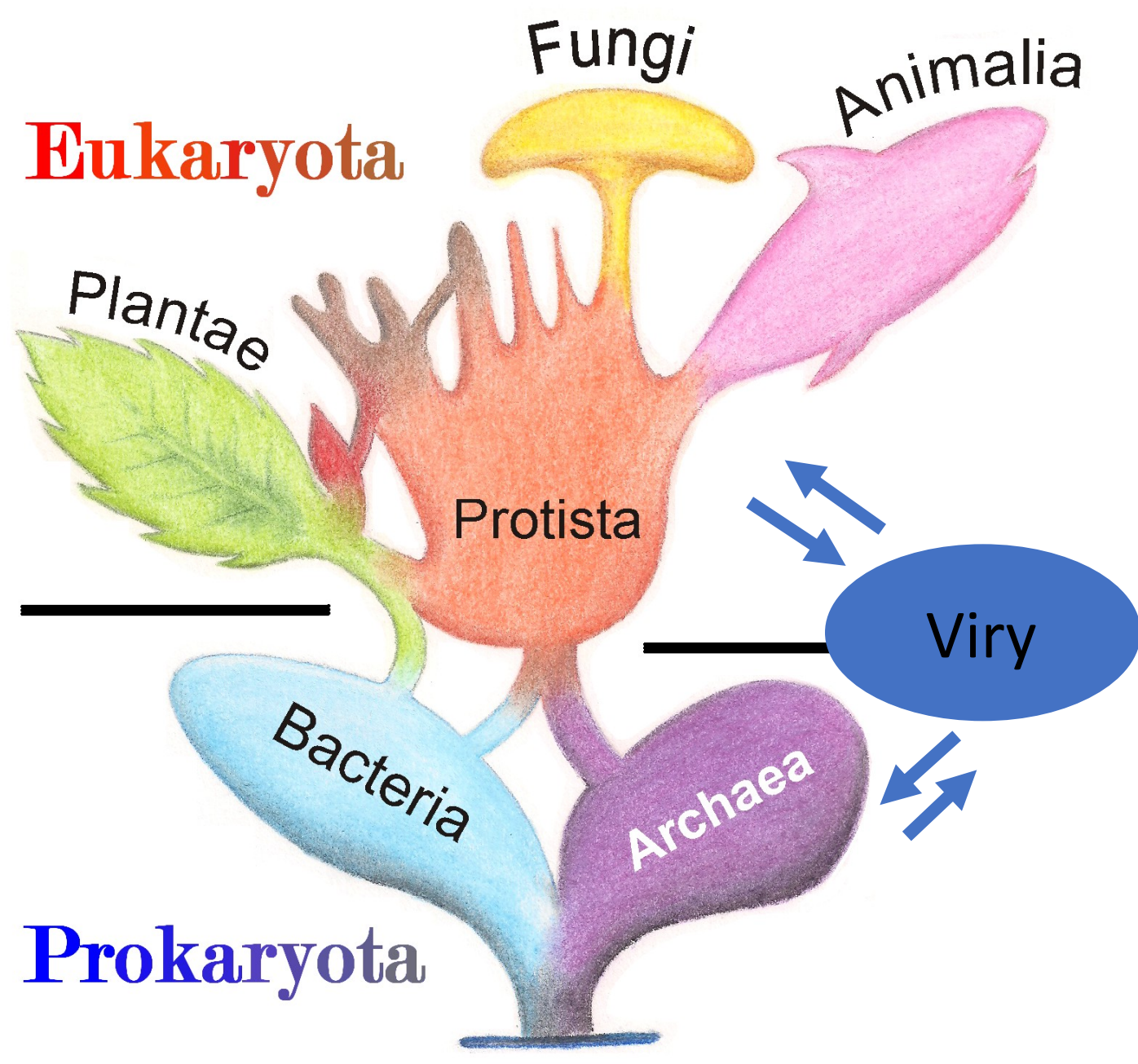
VELIKOST GENOMU



Indicates 5 major extinction events that punctuate the history of life on earth

The Cretaceous - Tertiary extinction event at 65 million years ago is probably the best known because it led to the demise of the dinosaurs. However, the biggest mass extinction event of all time is the Permian - Triassic event. This occurred around 251 million years ago and resulted in the elimination of about 95% of known species.

Time scale and illustrations are not drawn to scale



Velikost genomů

Viry, organely

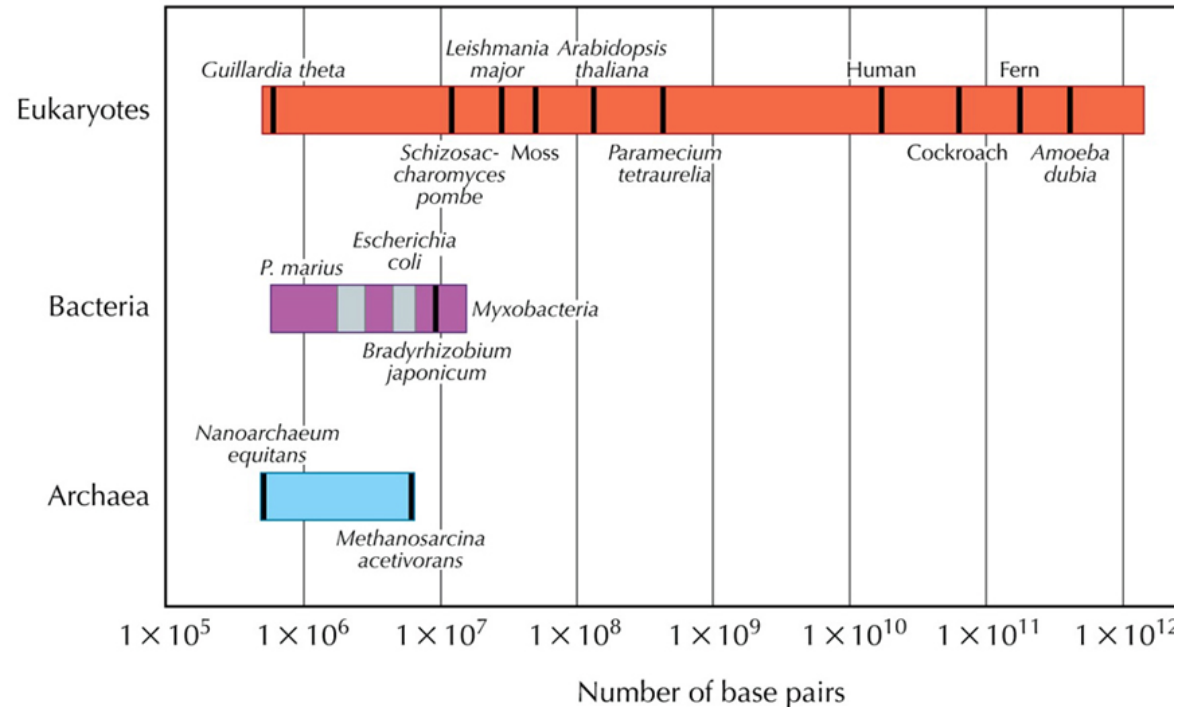
Miniaturní genomy (tisíce bp = kb)
 Vysoká hustota genů
 10-100 genů (bakteriofágové)

Bakterie

Malé genomy (miliony bp = Mb)
 Vysoká hustota genů
 Cca 5000 genů (E.coli)

Eukaryota

Velké genomy (miliardy bp = Gb)
 Nízká hustota genů
 Cca 20000 genů (člověk)

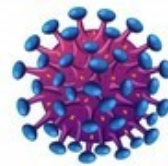


organism	genome size (base pairs)	protein coding genes	number of chromosomes
model organisms			
model bacteria <i>E. coli</i>	4.6 Mbp	4,300	1
budding yeast <i>S. cerevisiae</i>	12 Mbp	6,600	16
fission yeast <i>S. pombe</i>	13 Mbp	4,800	3
amoeba <i>D. discoideum</i>	34 Mbp	13,000	6
nematode <i>C. elegans</i>	100 Mbp	20,000	12 (2n)
fruit fly <i>D. melanogaster</i>	140 Mbp	14,000	8 (2n)
model plant <i>A. thaliana</i>	140 Mbp	27,000	10 (2n)
moss <i>P. patens</i>	510 Mbp	28,000	27
mouse <i>M. musculus</i>	2.8 Gbp	20,000	40 (2n)
human <i>H. sapiens</i>	3.2 Gbp	21,000	46 (2n)
viruses			
hepatitis D virus (smallest known animal RNA virus)	1.7 Kb	1	ssRNA
HIV-1	9.7 kbp	9	2 ssRNA (2n)
influenza A	14 kbp	11	8 ssRNA
bacteriophage λ	49 kbp	66	1 dsDNA
<i>Pandoravirus salinus</i> (largest known viral genome)	2.8 Mbp	2500	1 dsDNA
organelles			
mitochondria - <i>H. sapiens</i>	16.8 kbp	13 (+22 tRNA +2 rRNA)	1
mitochondria - <i>S. cerevisiae</i>	86 kbp	8	1
chloroplast - <i>A. thaliana</i>	150 kbp	100	1
bacteria			
<i>C. ruddii</i> (smallest genome of an endosymbiont bacteria)	160 kbp	182	1
<i>M. genitalium</i> (smallest genome of a free living bacteria)	580 kbp	470	1
<i>H. pylori</i>	1.7 Mbp	1,600	1
<i>Cyanobacteria S. elongatus</i>	2.7 Mbp	3,000	1
methicillin-resistant <i>S. aureus</i> (MRSA)	2.9 Mbp	2,700	1
<i>B. subtilis</i>	4.3 Mbp	4,100	1
<i>S. cellulorum</i> (largest known bacterial genome)	13 Mbp	9,400	1
archaea			
<i>Nanoarchaeum equitans</i> (smallest parasitic archaeal genome)	490 kbp	550	1
<i>Thermoplasma acidophilum</i> (flourishes in pH<1)	1.6 Mbp	1,500	1
<i>Methanocaldococcus (Methanococcus) jannaschii</i> (from ocean bottom hydrothermal vents; pressure >200 atm)	1.7 Mbp	1,700	1
<i>Pyrococcus furiosus</i> (optimal temp 100°C)	1.9 Mbp	2,000	1
eukaryotes - multicellular			
pufferfish <i>Fugu rubripes</i> (smallest known vertebrate genome)	400 Mbp	19,000	22
poplar <i>P. trichocarpa</i> (first tree genome sequenced)	500 Mbp	46,000	19
corn <i>Z. mays</i>	2.3 Gbp	33,000	20 (2n)
dog <i>C. familiaris</i>	2.4 Gbp	19,000	40
chimpanzee <i>P. troglodytes</i>	3.3 Gbp	19,000	48 (2n)
wheat <i>T. aestivum</i> (hexaploid)	16.8 Gbp	95,000	42 (2n=6x)
marbled lungfish <i>P. aethiopicus</i> (largest known animal genome)	130 Gbp	unknown	34 (2n)
herb plant <i>Paris japonica</i> (largest known genome)	150 Gbp	unknown	40 (2n)

Genomy virů

viruses			
hepatitis D virus (smallest known animal RNA virus)	1.7 Kb	1	ssRNA
HIV-1	9.7 kbp	9	2 ssRNA (2n)
influenza A	14 kbp	11	8 ssRNA
bacteriophage λ	49 kbp	66	1 dsDNA
Pandoravirus salinus (largest known viral genome)	2.8 Mbp	2500	1 dsDNA

- genomy malé, kompaktní
- efektivní využití
- ss/ds DNA
- ss RNA
- kombinace DNA a RNA



HIV



Hepatitis B



Ebola Virus



Adenovirus



Influenza

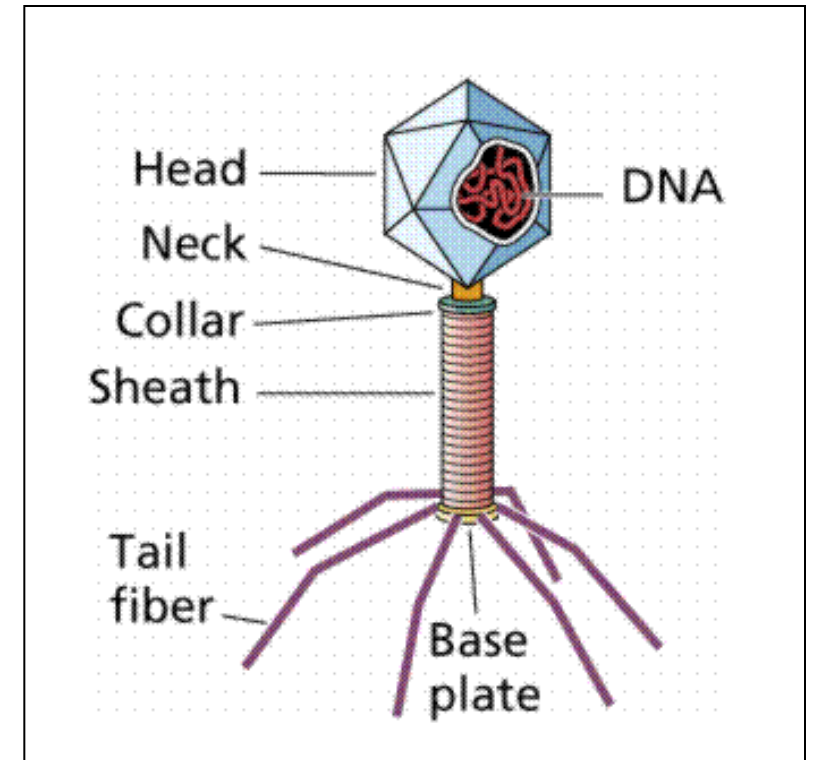


Bacteriophage

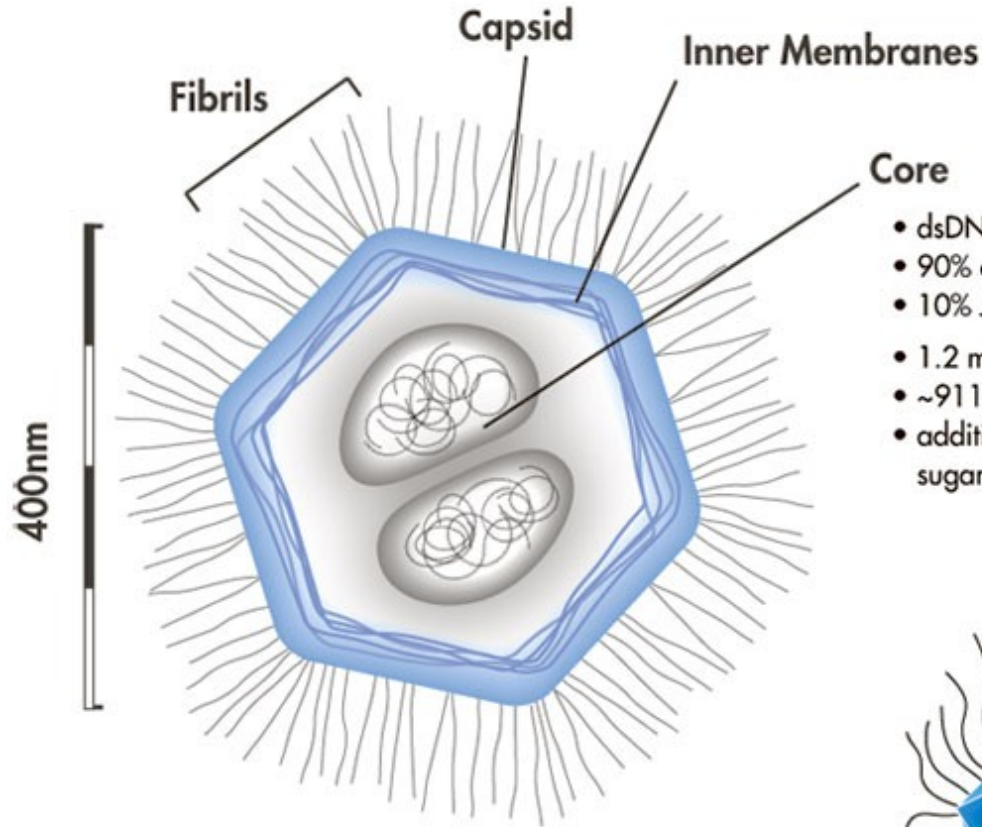
Virus (bakteriofág)
~ 50 kb ~ stovky genů

Viry nejsou monofyletickou skupinou entit

- různé viry mají různý původ, proto i rozdílné genomy z hlediska obsahu genů a kódující molekuly



Extrémní genomy: Mimivirus (APMV)

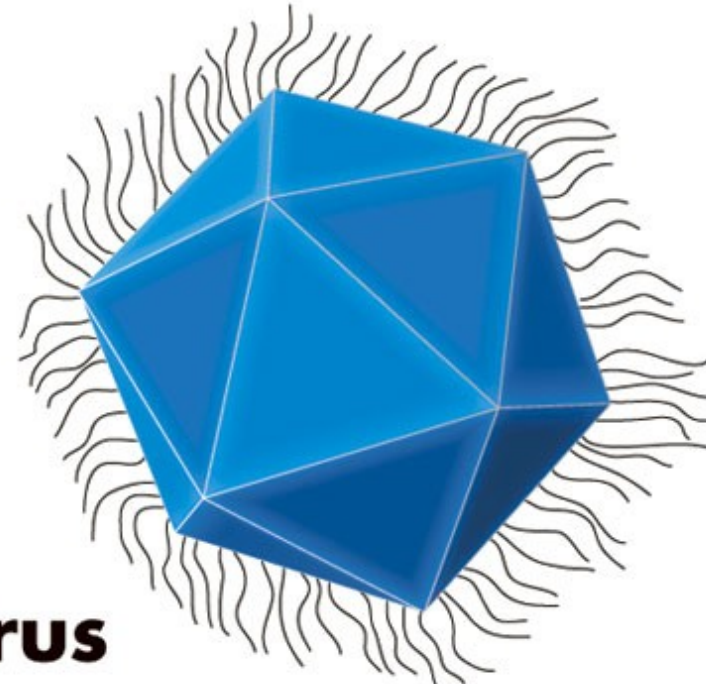


- dsDNA virus
- 90% coding capacity
- 10% Junk DNA
- 1.2 million base pairs
- ~911 protein coding genes
- additional genes (inc. aminoacyl tRNA synthetases; sugar, lipid, and amino acid metabolism)

Megavirus chilensis

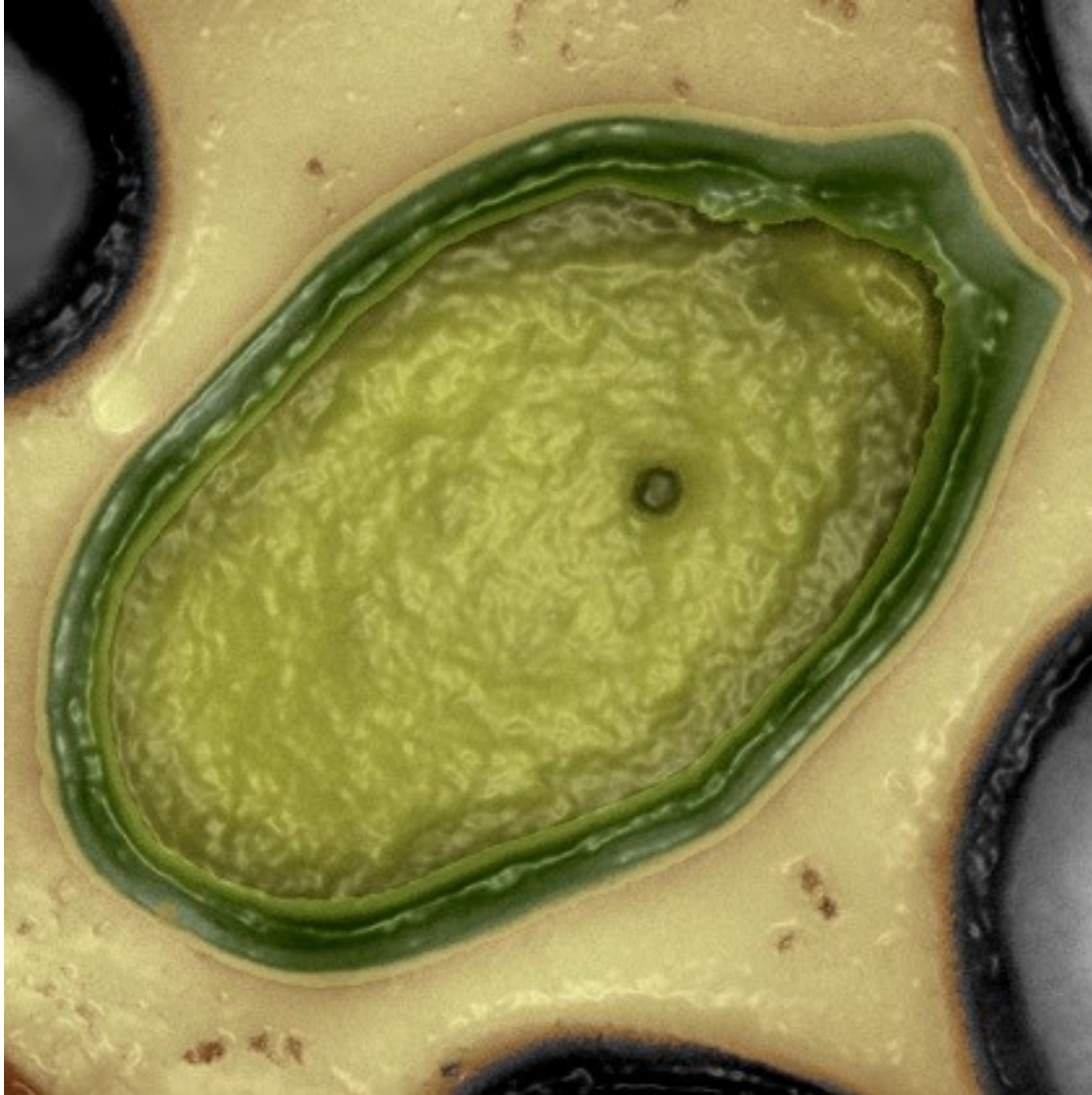
1,26 Mb

1120 genů



**acanthamoeba
polyphaga mimivirus**

Extrémní genomy: Pandoravirus salinus



- 2,47 Mb
- 2556 genů
- „parazité“ améb
- Jen 6 % genů má homologie v genech jiných org. – nová větev života?
- Životní cyklus podobný jako u ostatních virů – závislí na hostiteli v replikaci DNA a syntéze proteinů

Obří viry – megaviry, mimiviry

Velikost **genomu** srovnatelná s bakteriemi - *Pandoravirus salinus* (genom 2.5Mbp)

Velikost **kapsidy** srovnatelná s bakteriemi - *Pithovirus sibericum* (velký 1.5 μ m)

- kódují ribozómy a kapsidu, jen částečná závislost na hostiteli
- metabolické geny (~2500 genů pro proteiny), 10% repetitivní DNA

*Co bylo dříve – buňka
nebo mimivirus?*

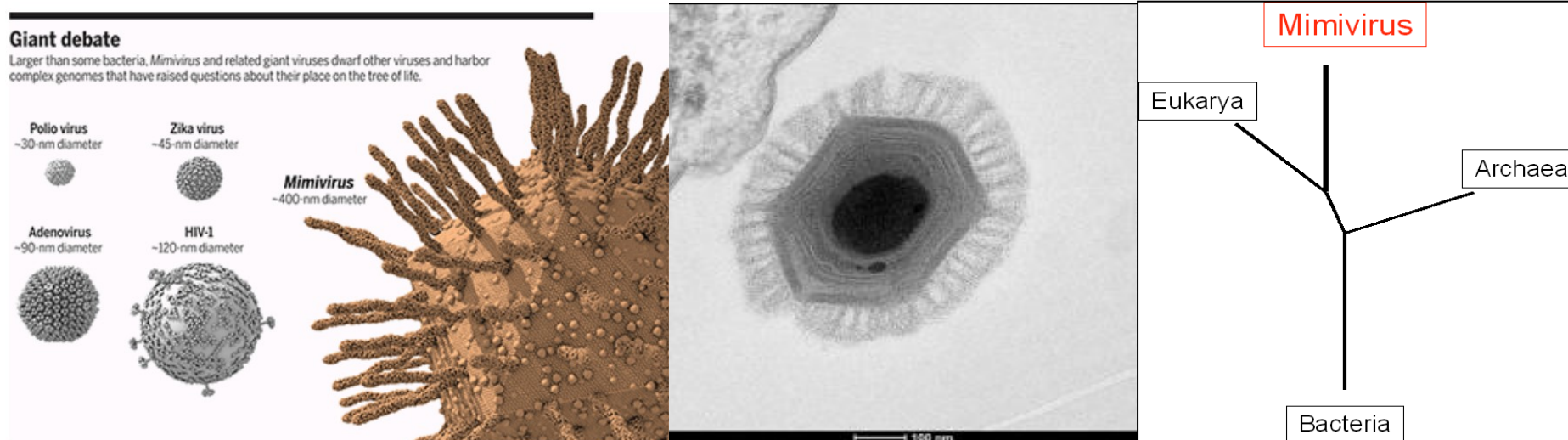
Paraziti megavirů:

Virofágy: např. *Sputnik* - 18kb/21 genů, přenáší geny mezi megaviry (jako fág)

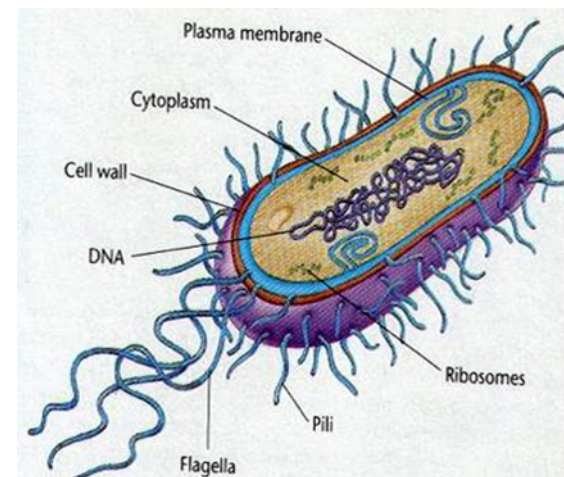
Transpovirony: ~7kb, několik genů, lin. ds. DNA, abundantní

*Eugene Koonin: „Giant viruses crossed the **imaginary boundry** between viruses and cellular organisms.“*

*J.-M. Claverie: „The fact that it can **get sick** makes it more alive.“*



Genomy eubakterií a archeí



bacteria			
<i>C. ruddii</i> (smallest genome of an endosymbiont bacteria)	160 kbp	182	1
<i>M. genitalium</i> (smallest genome of a free living bacteria)	580 kbp	470	1
<i>H. pylori</i>	1.7 Mbp	1,600	1
Cyanobacteria <i>S. elongatus</i>	2.7 Mbp	3,000	1
methicillin-resistant <i>S. aureus</i> (MRSA)	2.9 Mbp	2,700	1
<i>B. subtilis</i>	4.3 Mbp	4,100	1
<i>S. cellulosum</i> (largest known bacterial genome)	13 Mbp	9,400	1
archaea			
<i>Nanoarchaeum equitans</i> (smallest parasitic archaeal genome)	490 kbp	550	1
<i>Thermoplasma acidophilum</i> (flourishes in pH<1)	1.6 Mbp	1,500	1
<i>Methanocaldococcus</i> (<i>Methanococcus</i>) <i>jannaschii</i> (from ocean bottom hydrothermal vents; pressure >200 atm)	1.7 Mbp	1,700	1
<i>Pyrococcus furiosus</i> (optimal temp 100°C)	1.9 Mbp	2,000	1

Procesy ovlivňující velikost bakteriálního genomu:

- genová duplikace, malé delece a inserce, transpozice, horizontální přenos, ztráta genů v parazitických liniích, atd.

- genomy malé, kompaktní
- efektivní využití, operony, polycistronní
- kruhová genomová dsDNA + plazmidy
- *Carsonella ruddii* – nejmenší genom endosymbiotické bakterie (160 kb, 182 genů)
- *Mycoplasma genitalium* – nejmenší genom volně žijící (parazitické) bakterie (580 kb, 470 genů)
- *Sorangium cellulosum* – největší genom, půdní bakterie (13 Mb, 9400 genů)



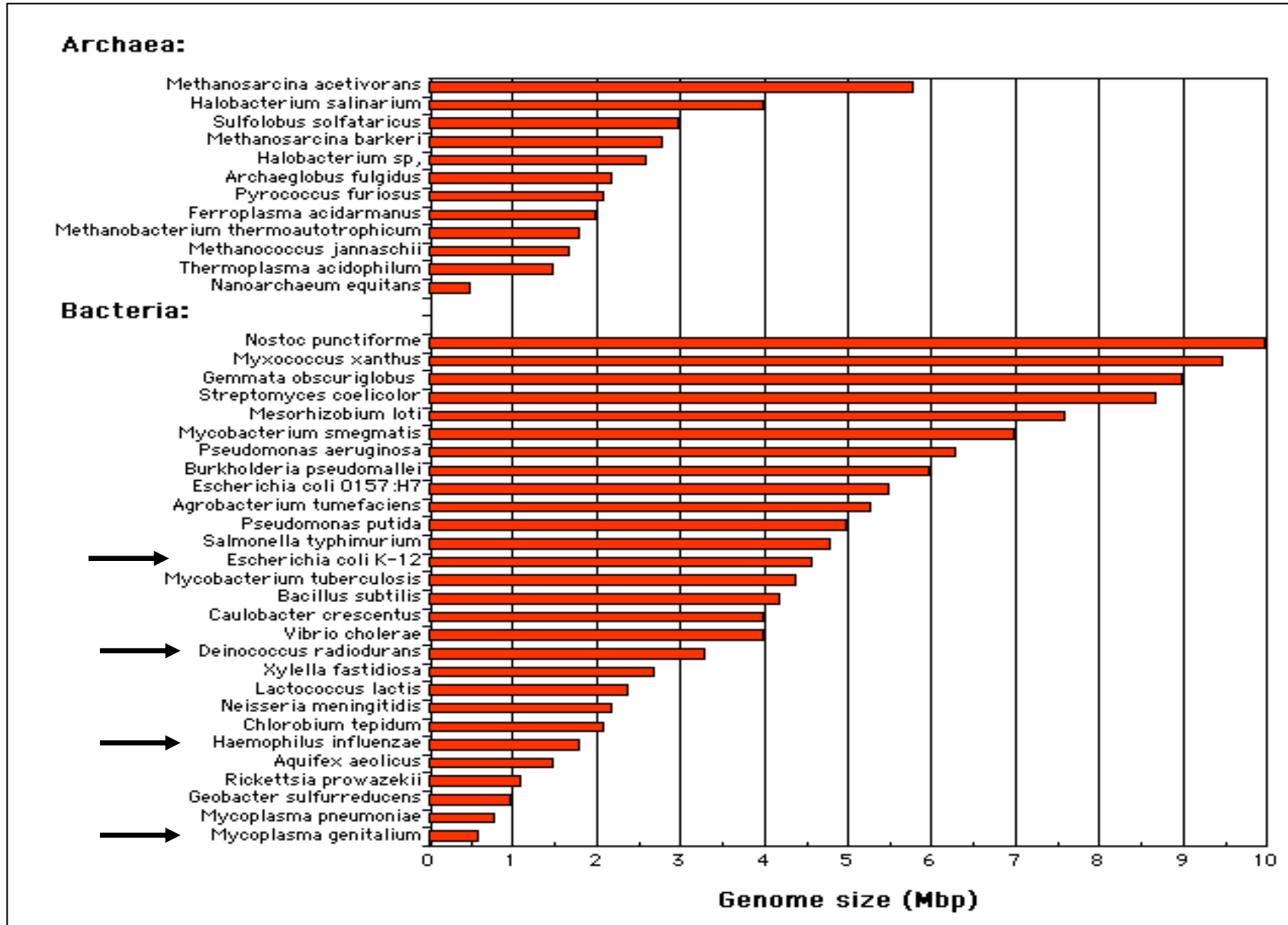
Velikost genomů prokaryot je zhruba úměrná počtu genů

E. coli

~ 4,7 mil bp

~ 4500 genů

Velikosti genomů prokaryot

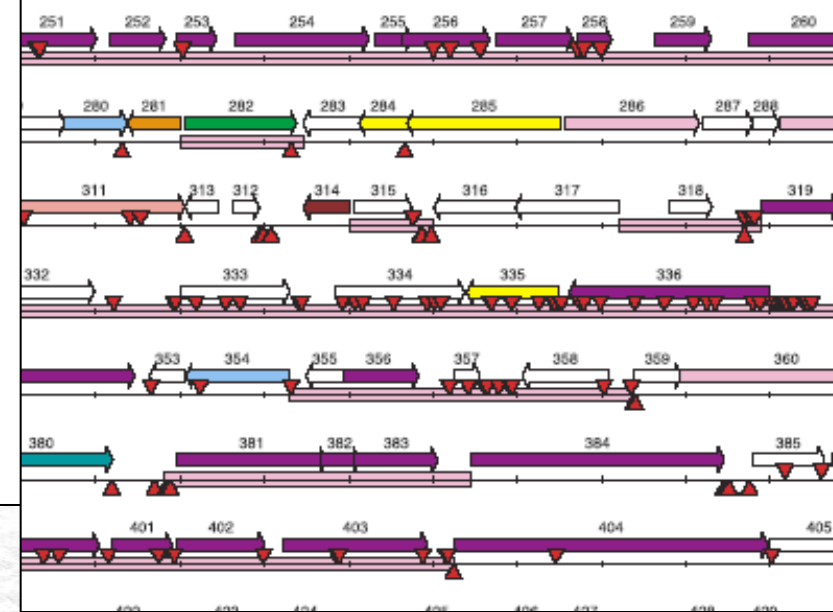


„Životní minimum“ bakterie

Život s pouhými 271 geny

Strategie:

1. odstranění či inaktivace genů (top-down)
2. syntéza minimálního genomu (bottom-up)



TIGR
THE INSTITUTE FOR GENOMIC RESEARCH

Minimal Genome Project

Press Release

Scientists at TIGR Uncover the Minimal Number of Cellular Genes Needed for Life

Publication

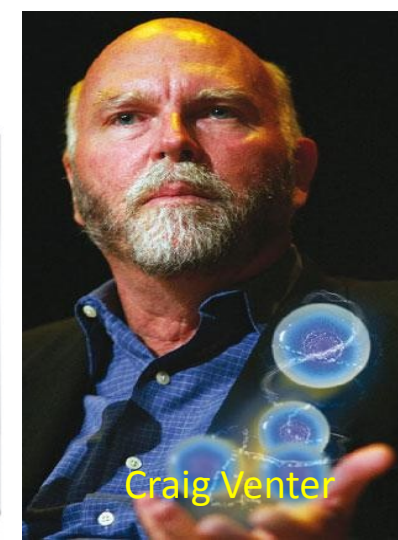
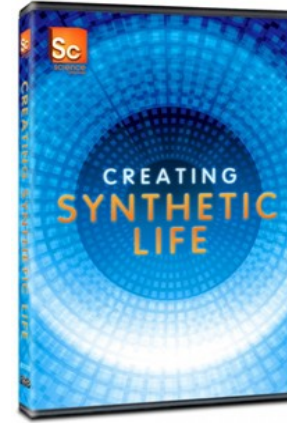
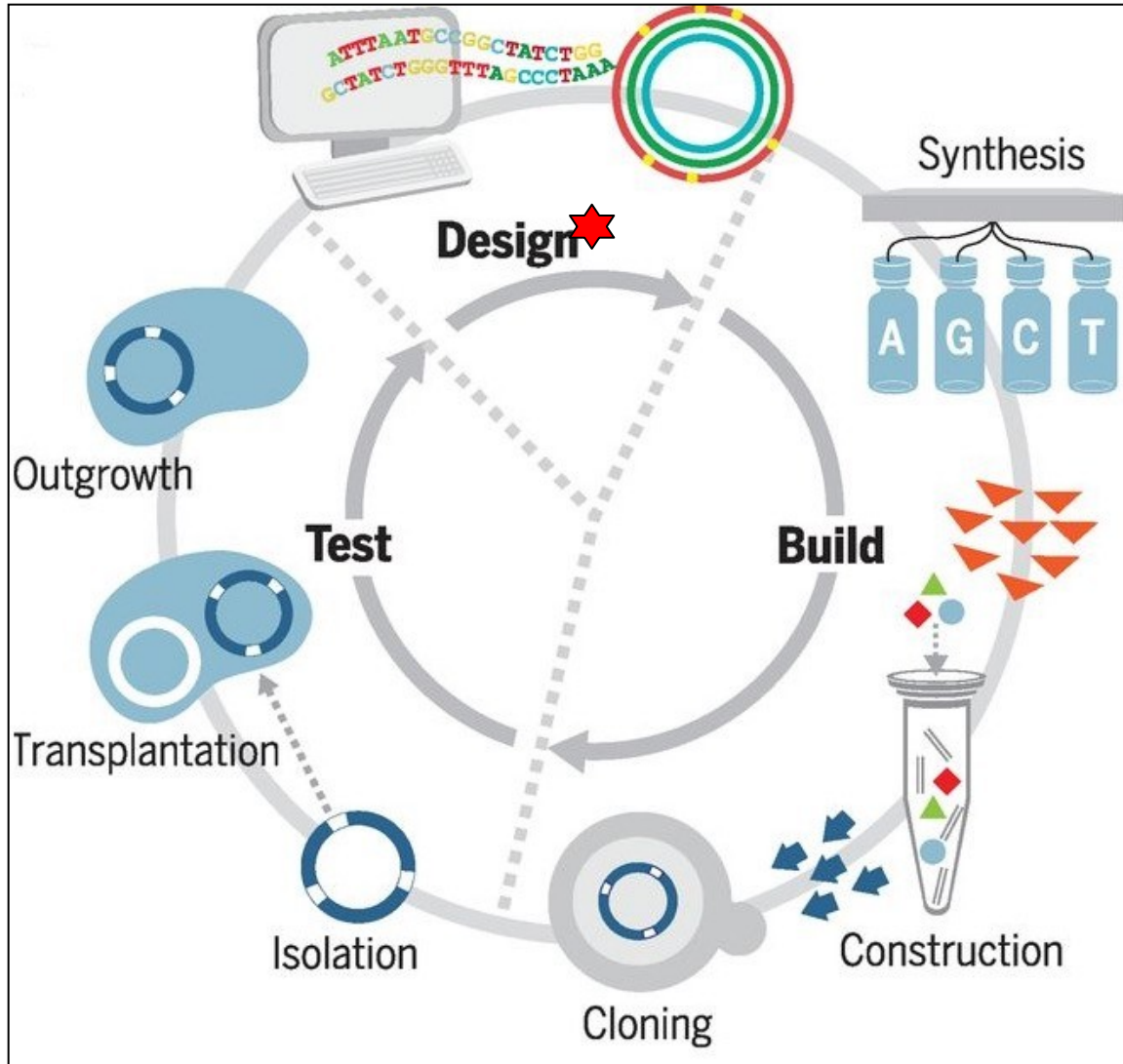
[\(777K PDF Format\)](#) **Global Transposon Mutagenesis and a Minimal Mycoplasma Genome**, Hutchison III, C.A., Peterson, S.N., Gill, S.R., Cline, R.T., White, O., Fraser, C.M., Smith, H.O., Venter, J.C. *Science* 286 (5447): 2165-2169 (1999)

Global Transposon Mutagenesis and a Minimal Mycoplasma Genome

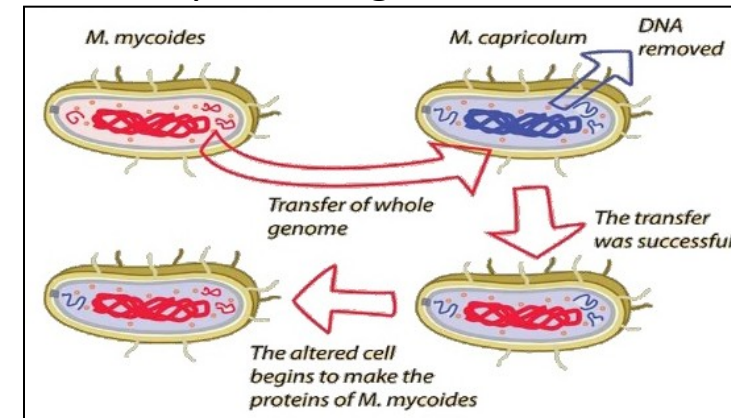
Clyde A. Hutchison III,^{1,2*} Scott N. Peterson,^{1*†} Steven R. Gill,¹
Robin T. Cline,¹ Owen White,¹ Claire M. Fraser,¹
Hamilton O. Smith,^{1†} J. Craig Venter^{1‡§}

Mycoplasma genitalium with 517 genes has the smallest gene complement of any independently replicating cell so far identified. Global transposon mutagenesis was used to identify nonessential genes in an effort to learn whether the naturally occurring gene complement is a true minimal genome under laboratory growth conditions. The positions of 2209 transposon insertions in the completely sequenced genomes of *M. genitalium* and its close relative *M. pneumoniae* were determined by sequencing across the junction of the transposon and the genomic DNA. These junctions defined 1354 distinct sites of insertion that were not lethal. The analysis suggests that 265 to 350 of the 480 protein-coding genes of *M. genitalium* are essential under laboratory growth conditions, including about 100 genes of unknown function.

Syntetický život: Pouhých 271 genů



Transplantace genomu:



*Genomy umíme nejen číst
ale také psát*

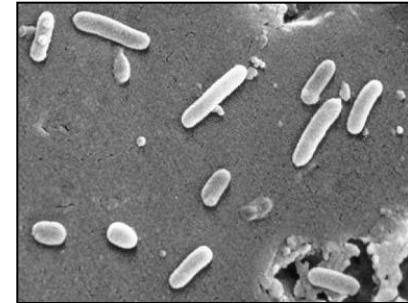
Mycoplasma laboratorium (Synthia)

- nahradili genom *M. capricolum* uměle nasyntetizovanou DNA *M. mycoides*
- nový software přebudoval původní hardware, vodoznaky v umělé DNA

Největší prokaryotické genomy

Pseudomonas aeruginosa (bakterie):

- 5500 genů
- přes 6 Mb
- přirozeně rezistentní k antibiotikům (ochranný obal)
- R-faktor, žije ve společenství jiných bakterií, konjugace
- lidský patogen (kožní n., močové, dýchací a trávicí cesty)



Nostoc punctiforme (sinice):

- 7432 ORF
- 8.9 Mb
- **repetice**, transpozony
- fotoautotrofní, také fakultativně heterotrofní
- možnost symbiózy s rostlinami i houbami

Repetice <i>N.p.</i>	Sites in genome
AATGAC _n (STRR2)	
AATGACA	69
AATGACT	63
AATGACC	39
AATCCC (STRR4)	
AATCCC	41
AATGCC	37
AATTACG (STRR5)	45
AATCCCC (STRR1)	
ATCCCC	39
AATCCCC	19
AGTCCCC	15
AGCAGGGG (STRR6)	29
AAAATTC (STTR7)	13

Genomy eukaryot

Jaderný genom

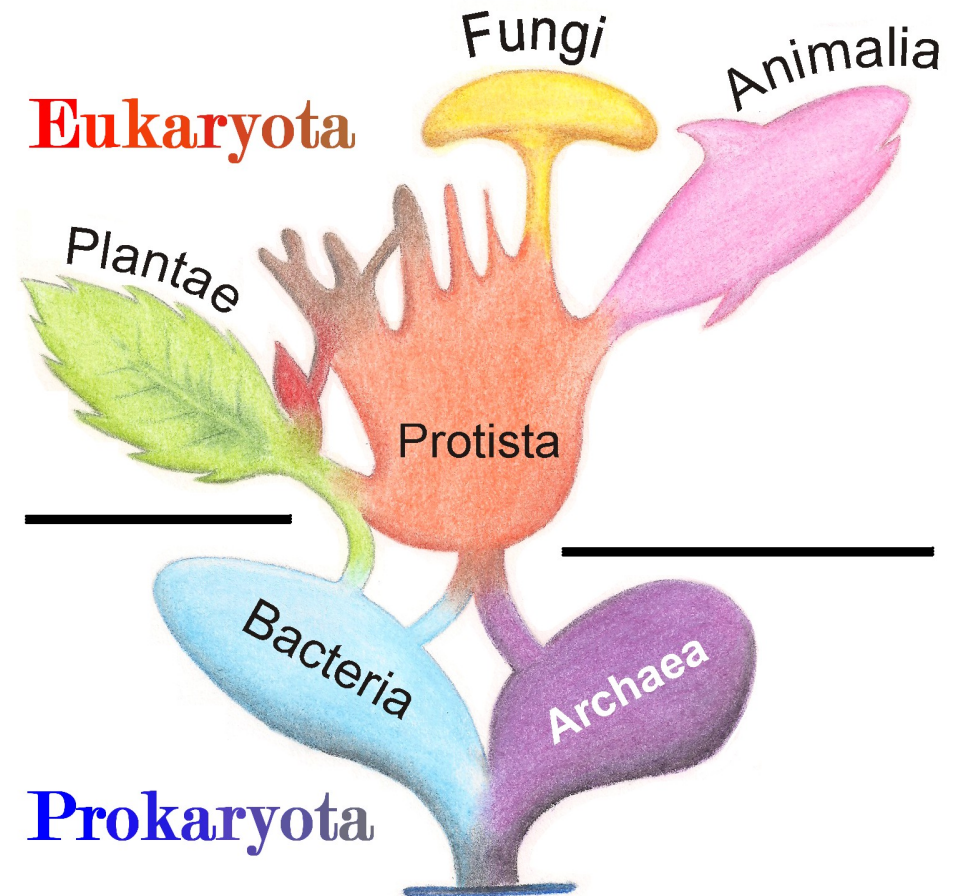
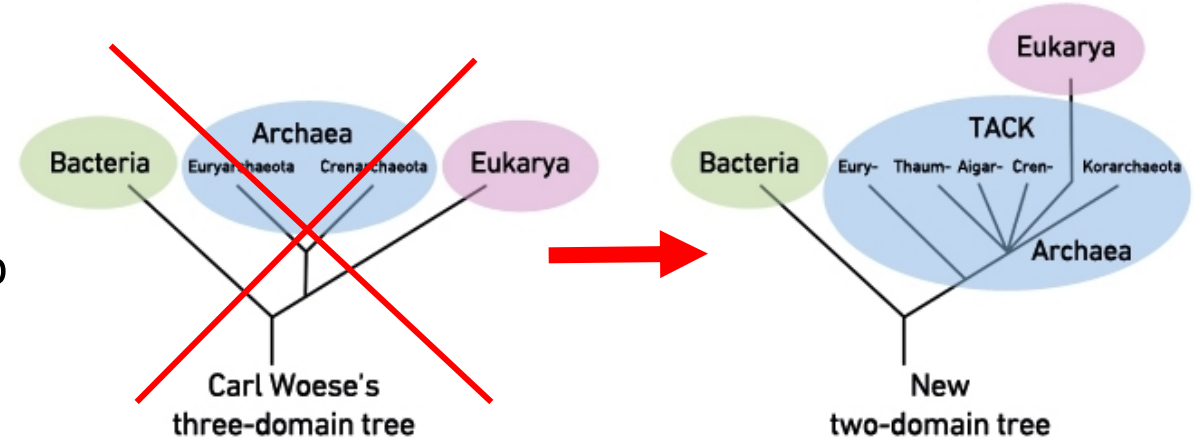
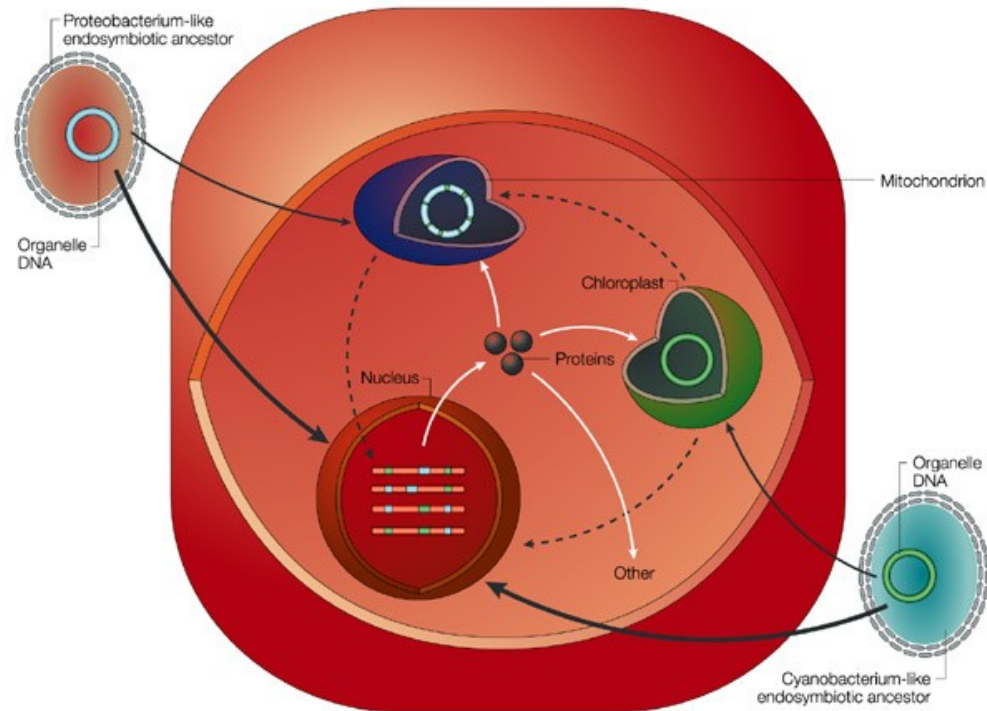
chromozomy z DNA, tisíce genů, monocistronní, Mbp-Gbp

Mitochondriální genom

kruhová DNA, desítky genů, desítky-stovky kbp

Plastidový genom

kruhová DNA, desítky genů, desítky-stovky kbp



Endosymbiotický vznik eukaryot

Endosymbióza cca -2 miliardy let (1,5; 1,2?) → vznik eukaryotické buňky

Archea (= jádro) + alfa-proteobakterie (= mitochondrie)

Reakce na vzrůstající koncentraci kyslíku v atmosféře → silný selekční tlak zvýhodňoval novou eukaryotickou buňku

Primární Endosymbióza cca -1,2 (0,9?) mld let → vznik fotosyntetizujících eukaryot (červené a zelené řasy)

Eukaryota + cyanobakterie (= plastidy)

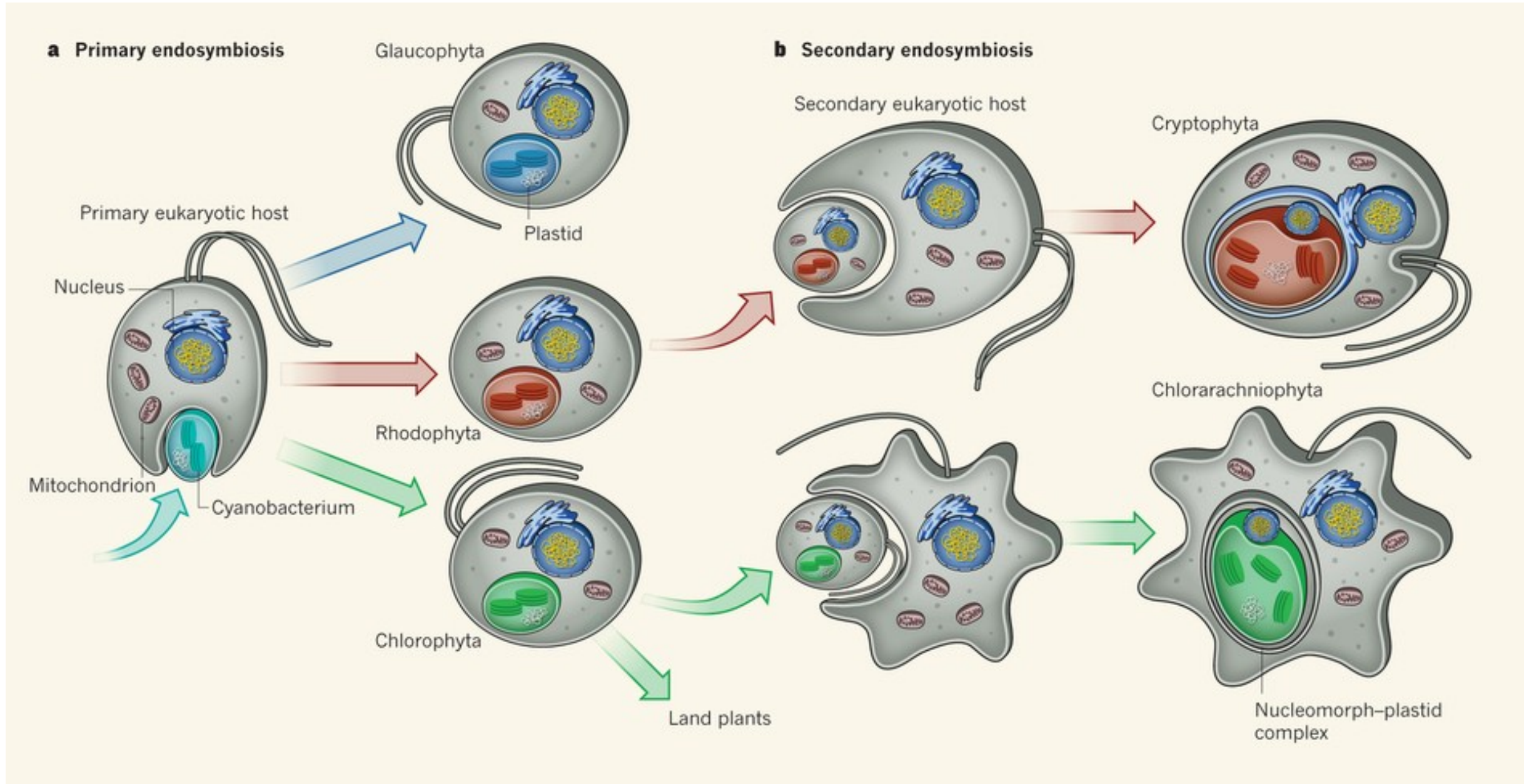
Sekundární Endosymbióza → vznik cryptophyta, chlorarachniophyta

Eukaryota + eukaryota (červená, zelená řasa)

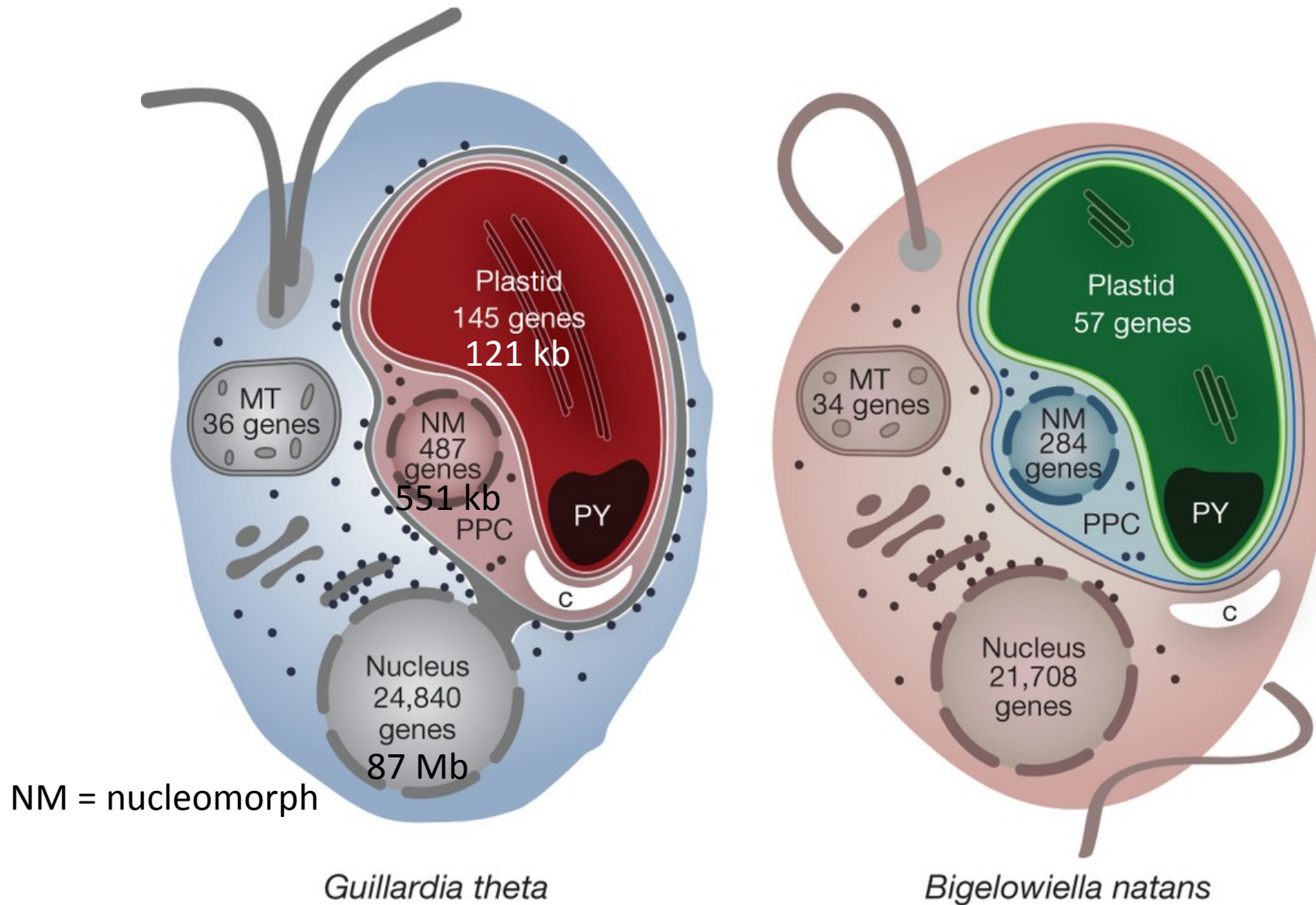
Eukaryota:

- Jaderné procesy jsou podobné archeálním
- Cytoplazmatické a organelové procesy podobné eubakteriím
- Genový transfer z organel do jádra, většinu kódujících funkcí časem převzalo centrální jádro
- Časové zařazení vzniku eukaryotické buňky je velmi orientační kvůli nejasným fosilním záznamům, někdy se uvádí první eukaryota až před 1,5 miliardami let nebo dokonce 800 miliony lety
- První mnohobuněčná eukaryota prokazatelně až před 600 miliony lety (Ediakara)

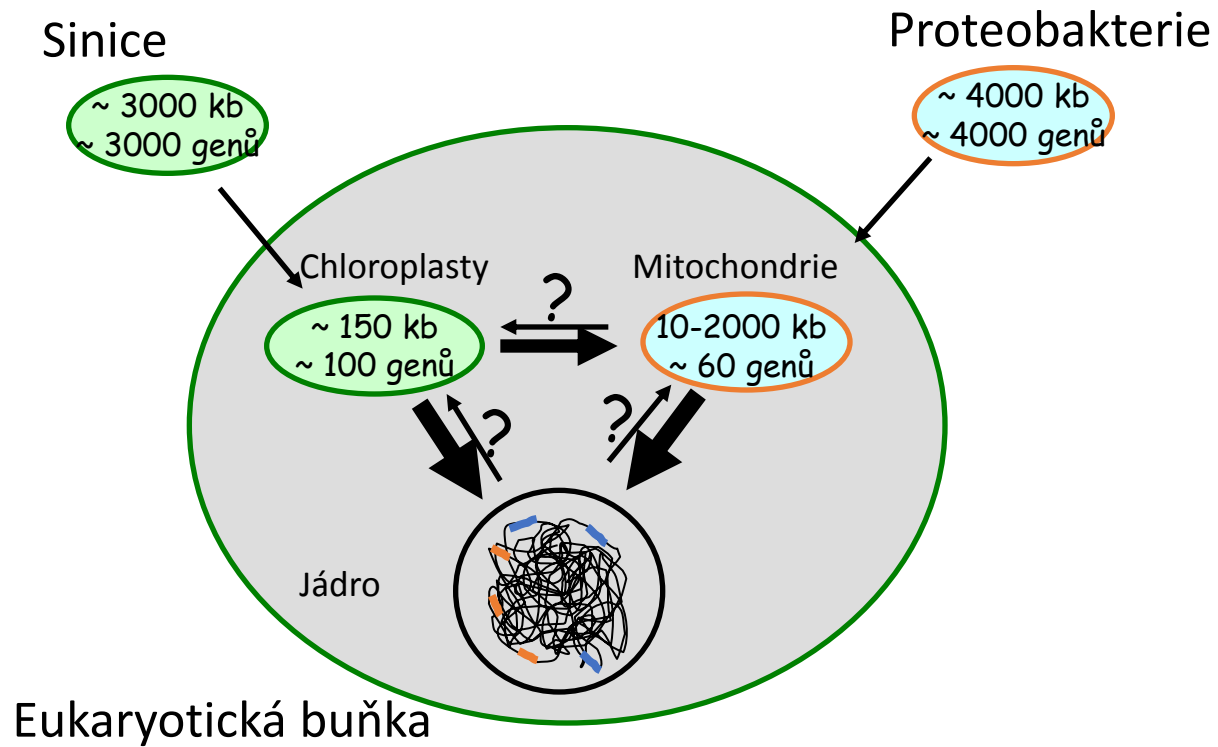
Primární a Sekundární endosymbióza



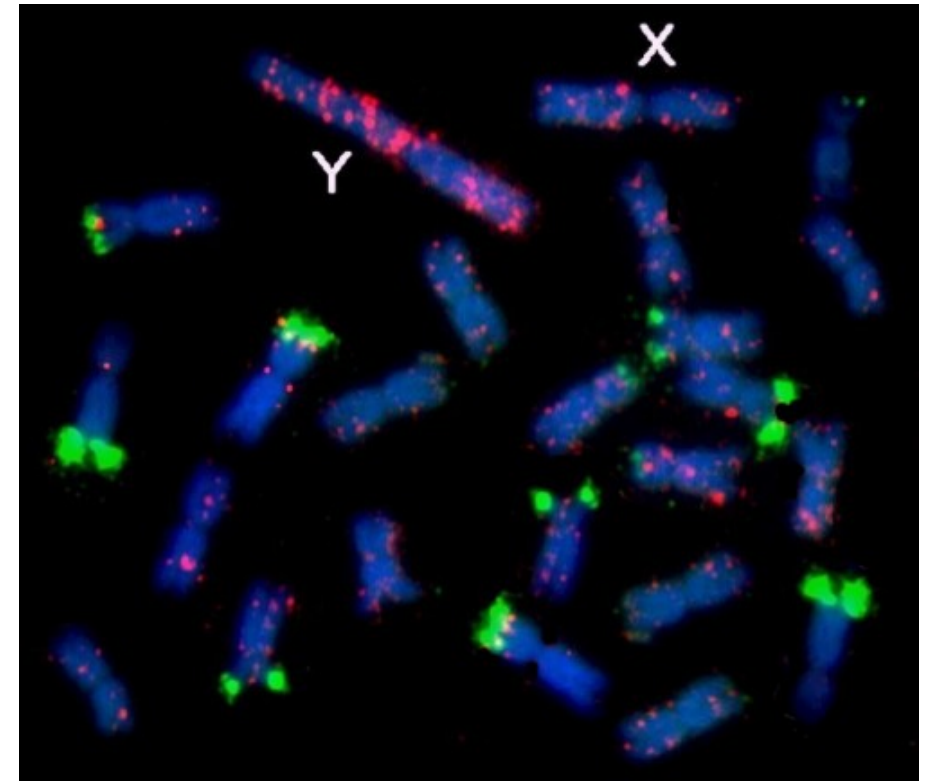
Extrémní genomy: sekundární endosymbióza – *Guillardia theta*



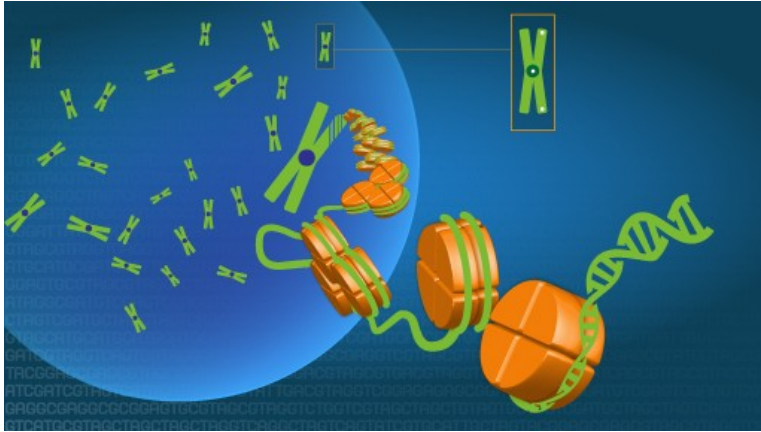
Endosymbióza a promiskuitní DNA



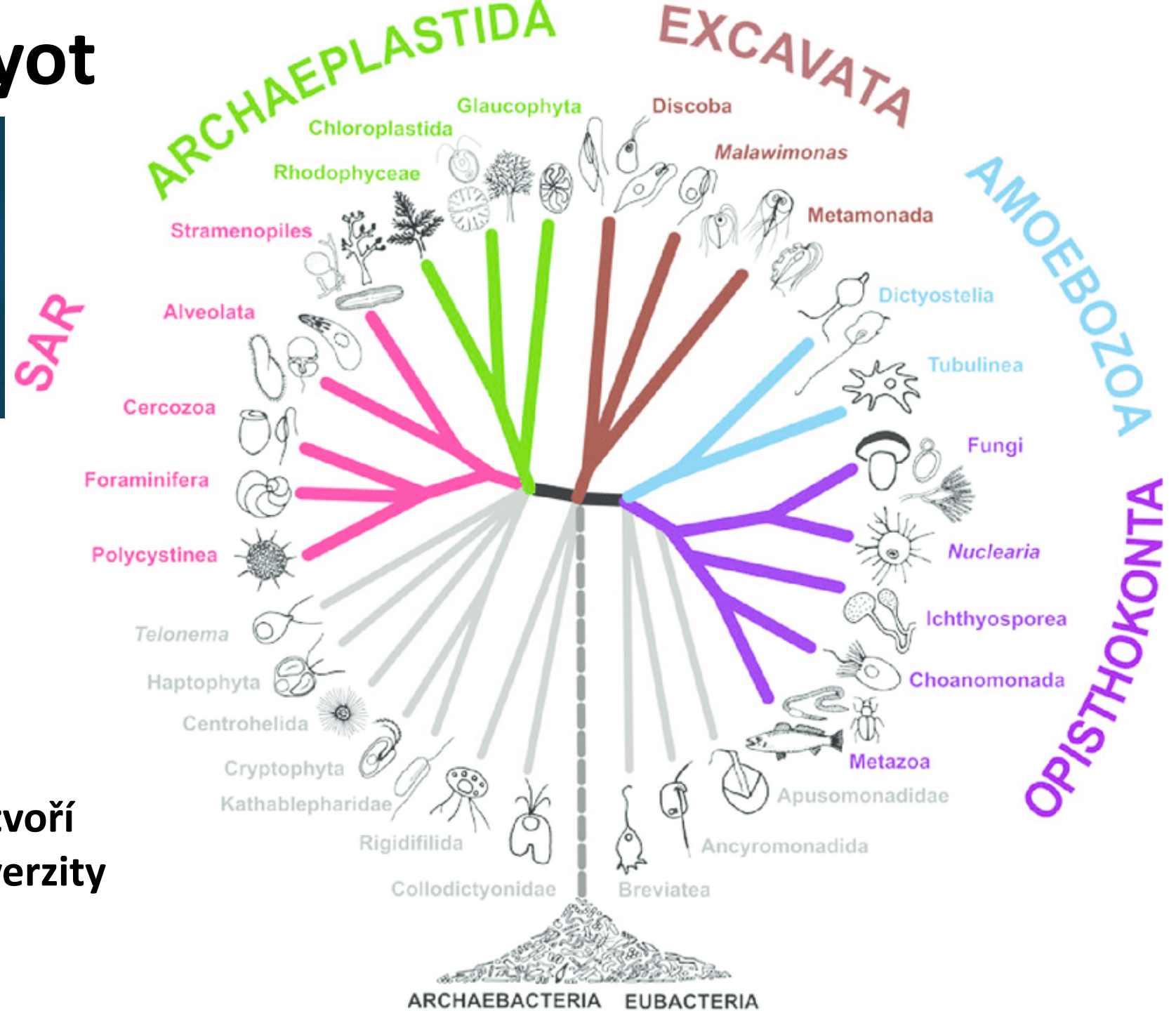
Chloroplastová DNA se hromadí na chromosomu **Y** u *Silene latifolia*



Genomy eukaryot

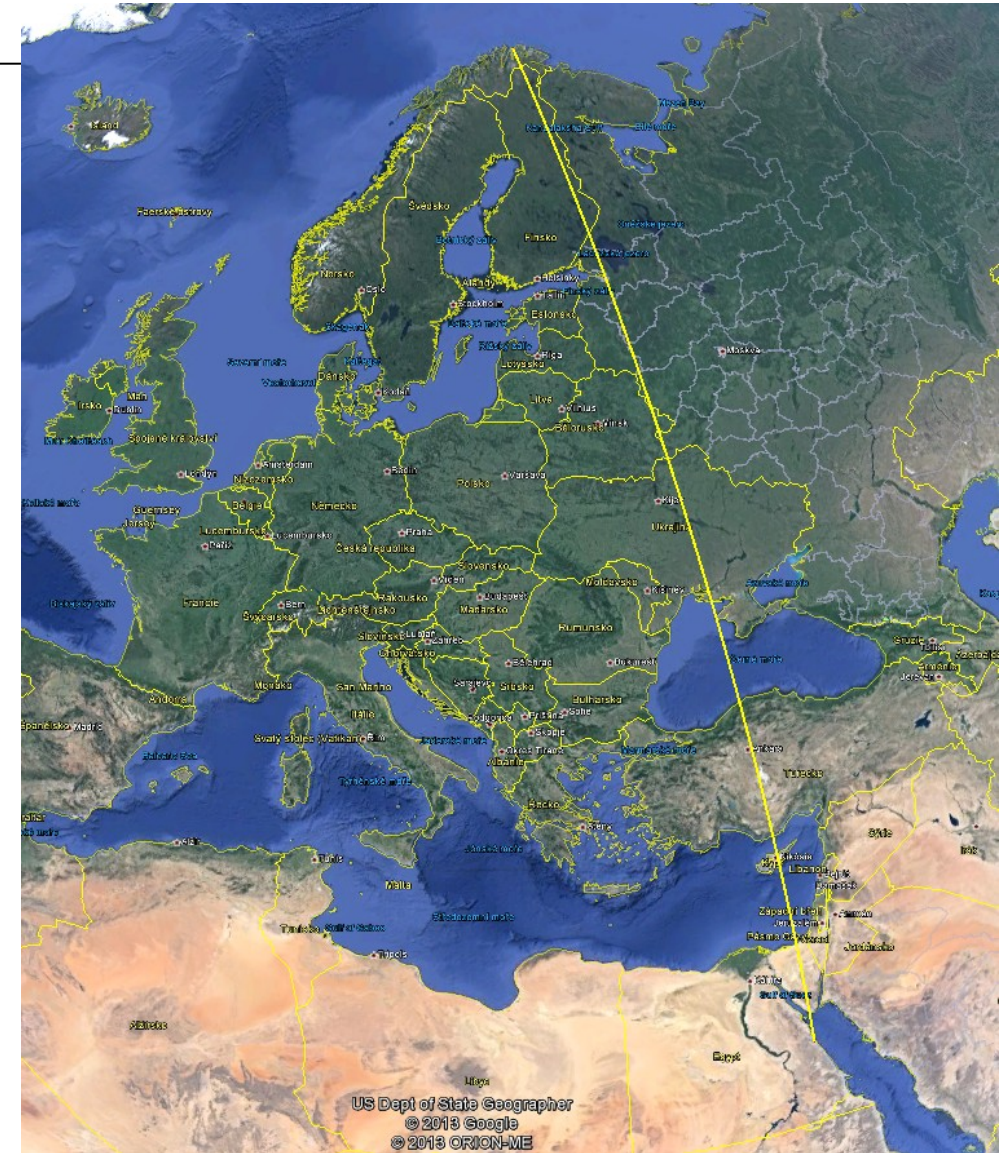
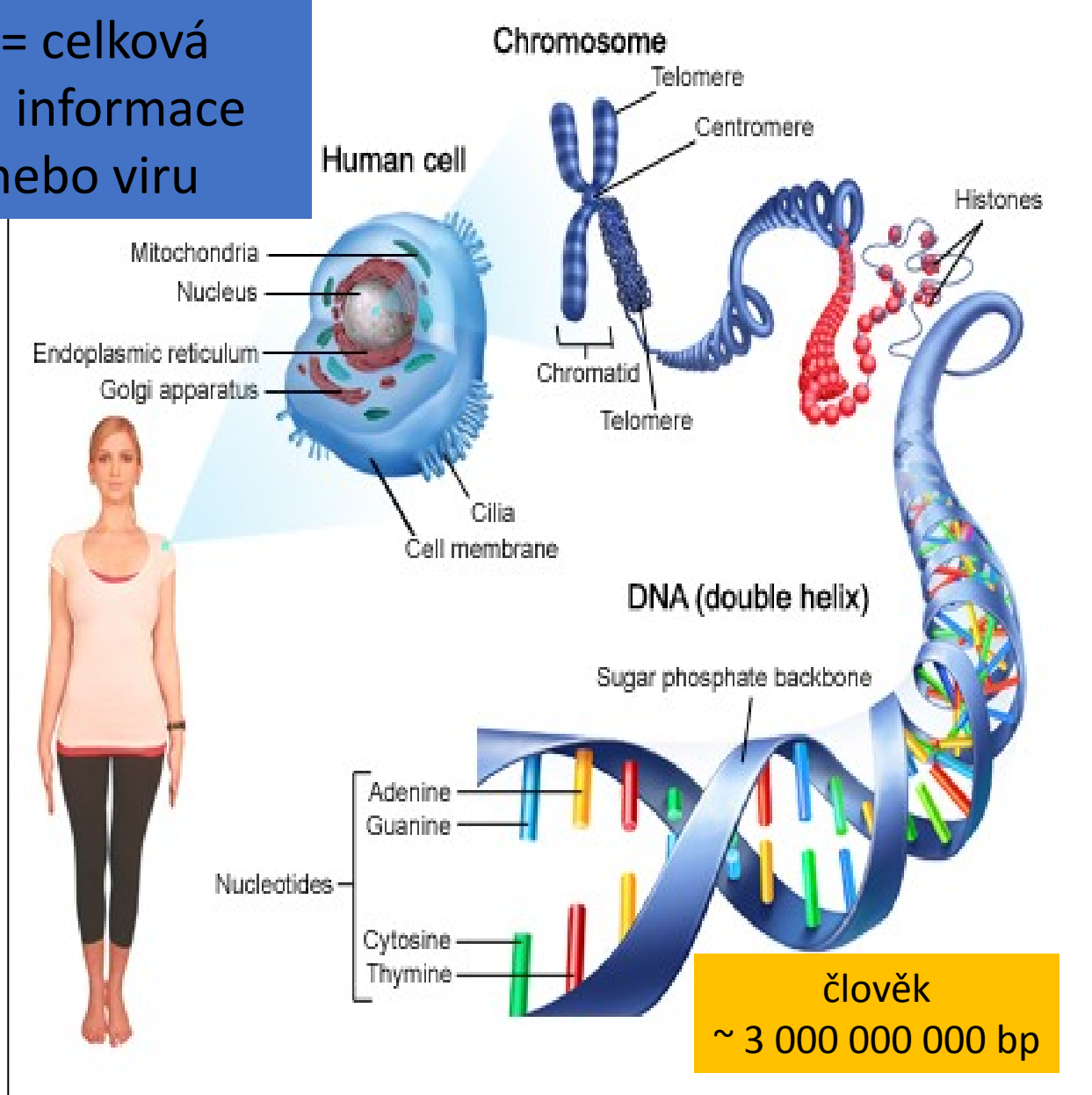


Mnohobuněčná eukaryota tvoří jen malý zlomek celkové diverzity eukaryot!



Vyšší organizmy: kde se nachází DNA?

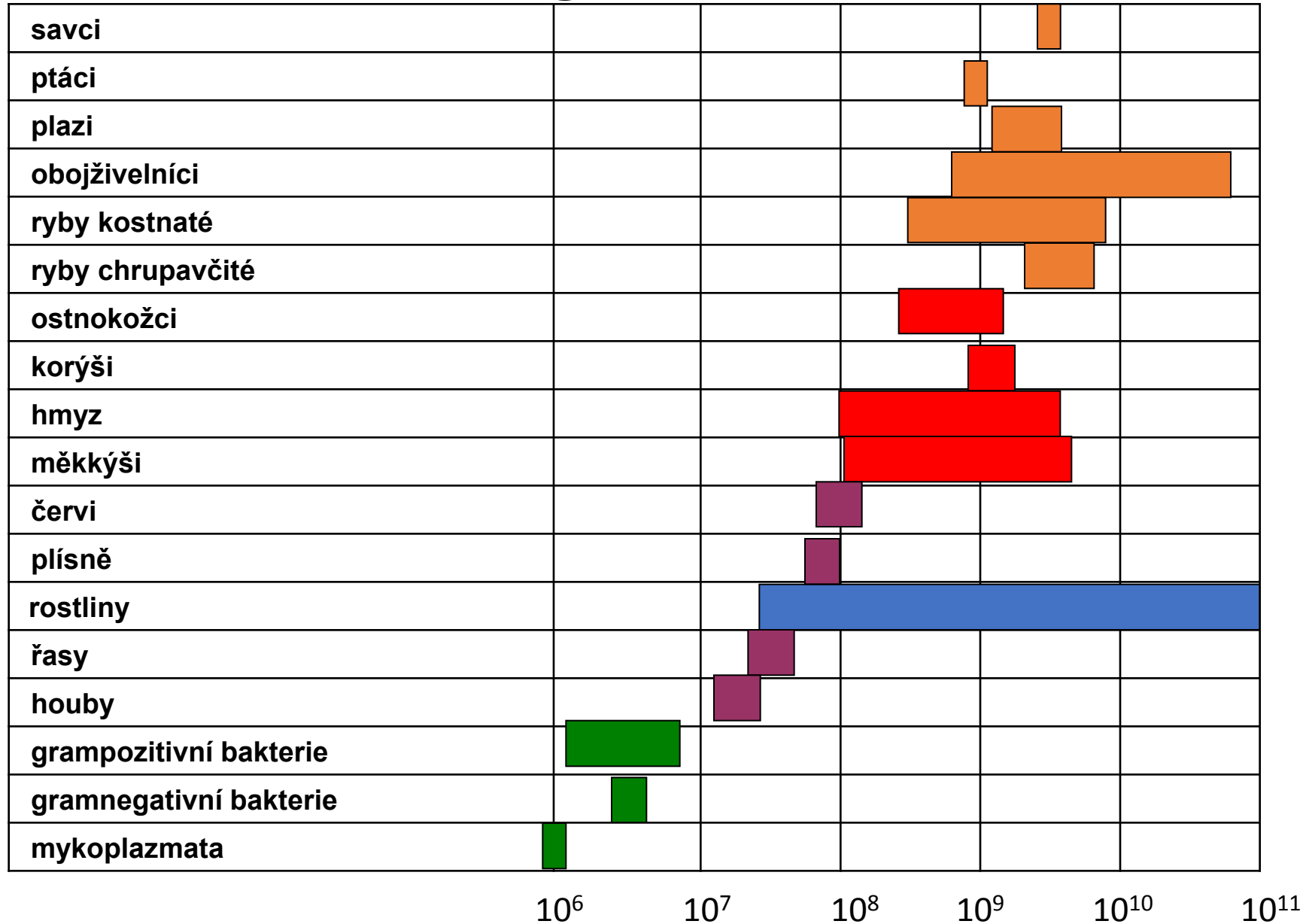
Genom = celková
genetická informace
buňky nebo viru



2nm x 1m
1cm -> 5000km

Variabilita velikosti genomů

Velikosti genomů



Velikosti genomů a paradox hodnoty C



Drosophila melanogaster

100x

Klíčem k řešení paradoxu hodnoty C jsou opakující se úseky DNA (repetice)



Podisma pedestris

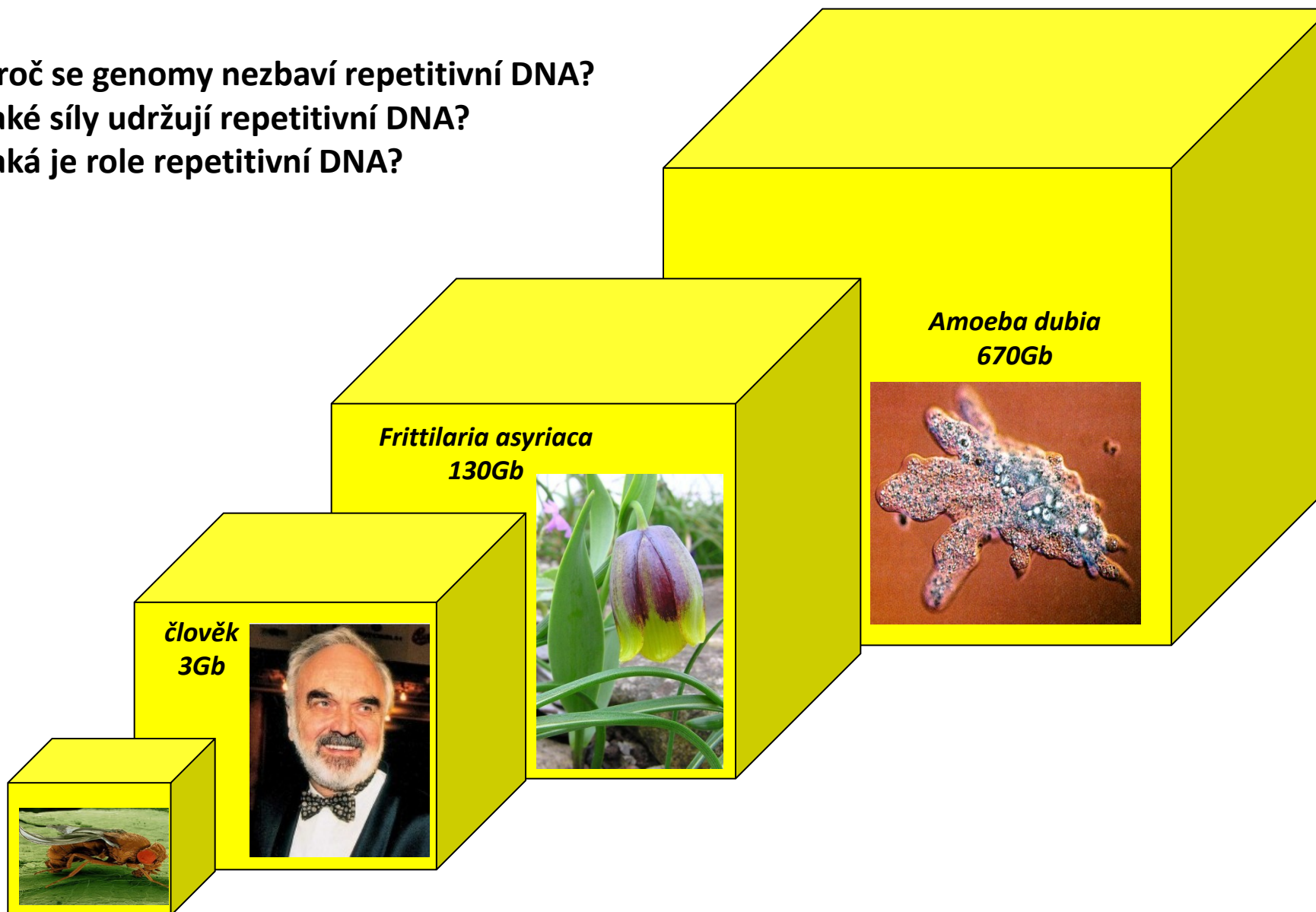
GENOM = celková genetická informace organismu (geny i negenové oblasti/repetice)

Paradox hodnoty C = velikost genomu organismu **není v korelaci** s komplexitou organismu
např. jednobuněčná měňavka (*Amoeba dubia*) má 200x větší genom než člověk

Genomy jsou repetitivní a dynamické

Velikosti genomů a paradox hodnoty C

Proč se genomy nezbaví repetitivní DNA?
Jaké síly udržují repetitivní DNA?
Jaká je role repetitivní DNA?



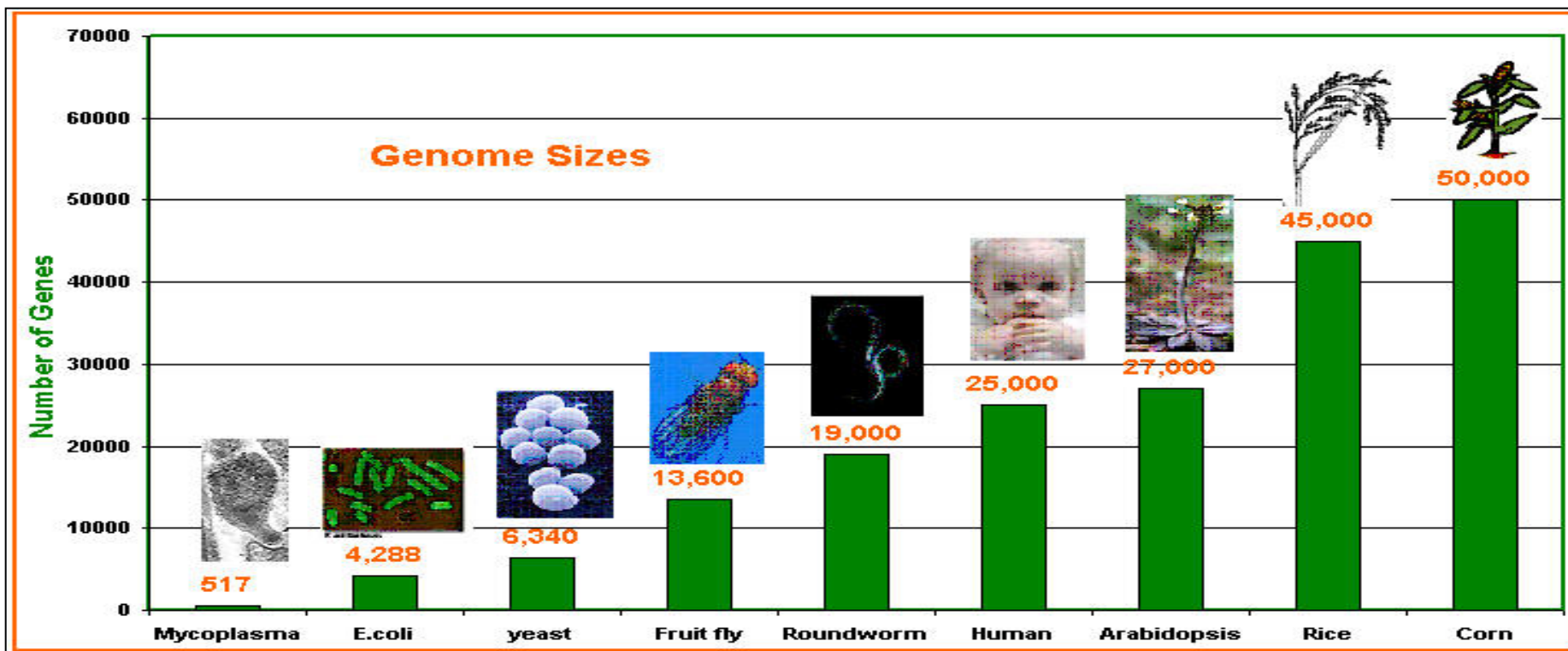
Velikosti genomů eukaryot

Eukaryotický genom: $8.8 \times 10^6 \sim 6.9 \times 10^{11}$
(8.8Mb – 670Gb)

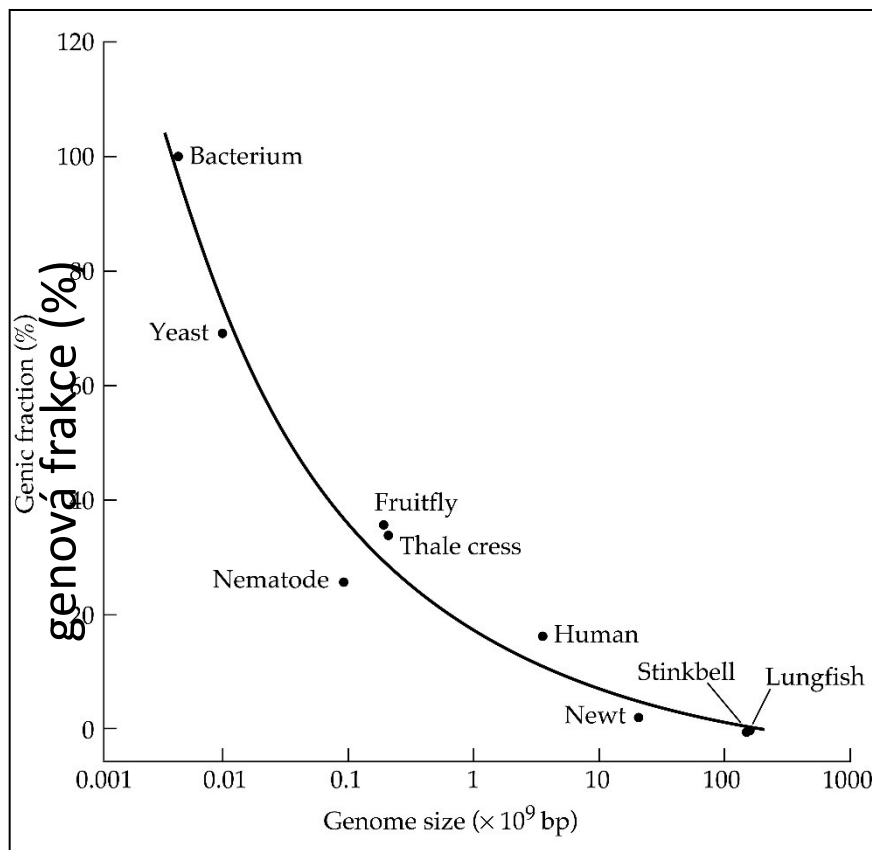
Nejmenší známý: *Saccharomyces cerevisiae*

Největší známý: *Amoeba dubia*

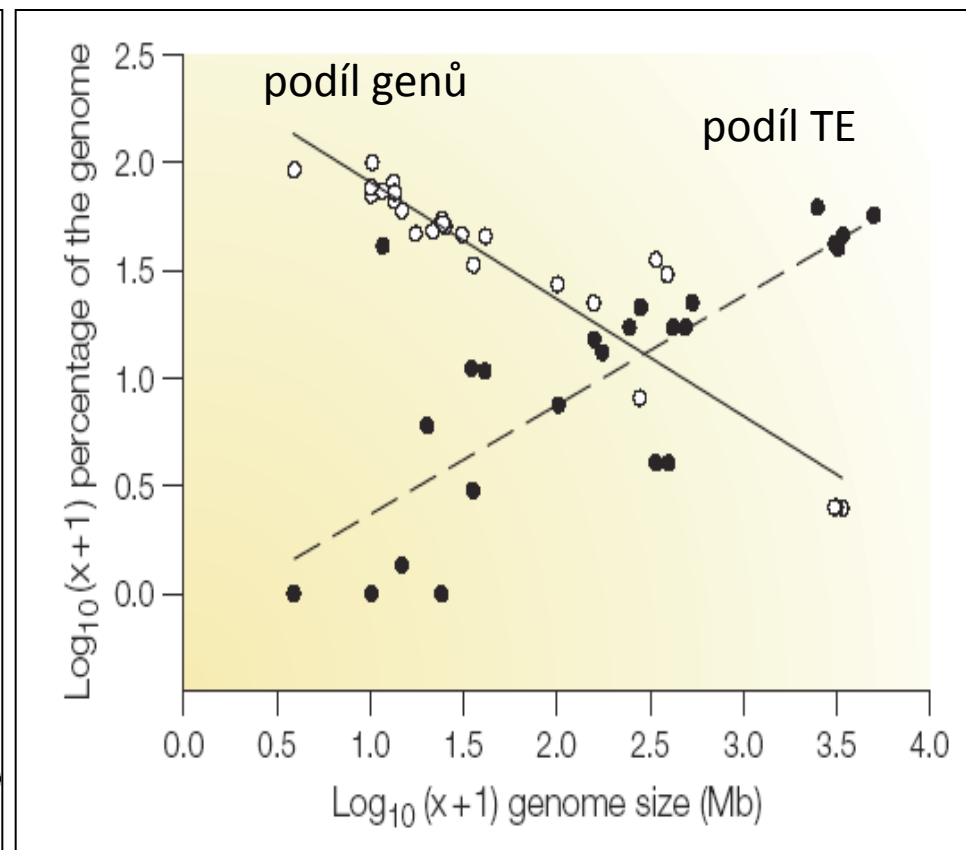
Velikosti eukaryotických genomů NEjsou úměrné počtu genů nebo komplexitě organismu



Klíč k řešení paradoxu hodnoty C se nachází v negenových oblastech



velikost genomu



velikost genomu

Hlavní komponenty eukaryotického genomu

Kódující části genů:

- u prokaryot tvoří většinu genomu
- u eukaryot méně, člověk 24 000 genů – 1.5%

Introny:

- původně považovány za příčinu C-paradoxu,
- tvoří většinu genů

Pseudogeny:

- klasické, retropseudogeny,
- 19000 člověk, 14000 myš, 51 kur, 33 kvasinka, 176 drosophila

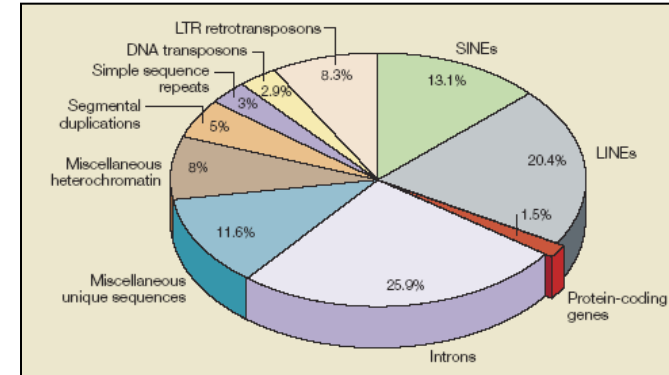
Mobilní elementy:

- LTR, nonLTR – SINE, LINE, DNA transposony- MITE

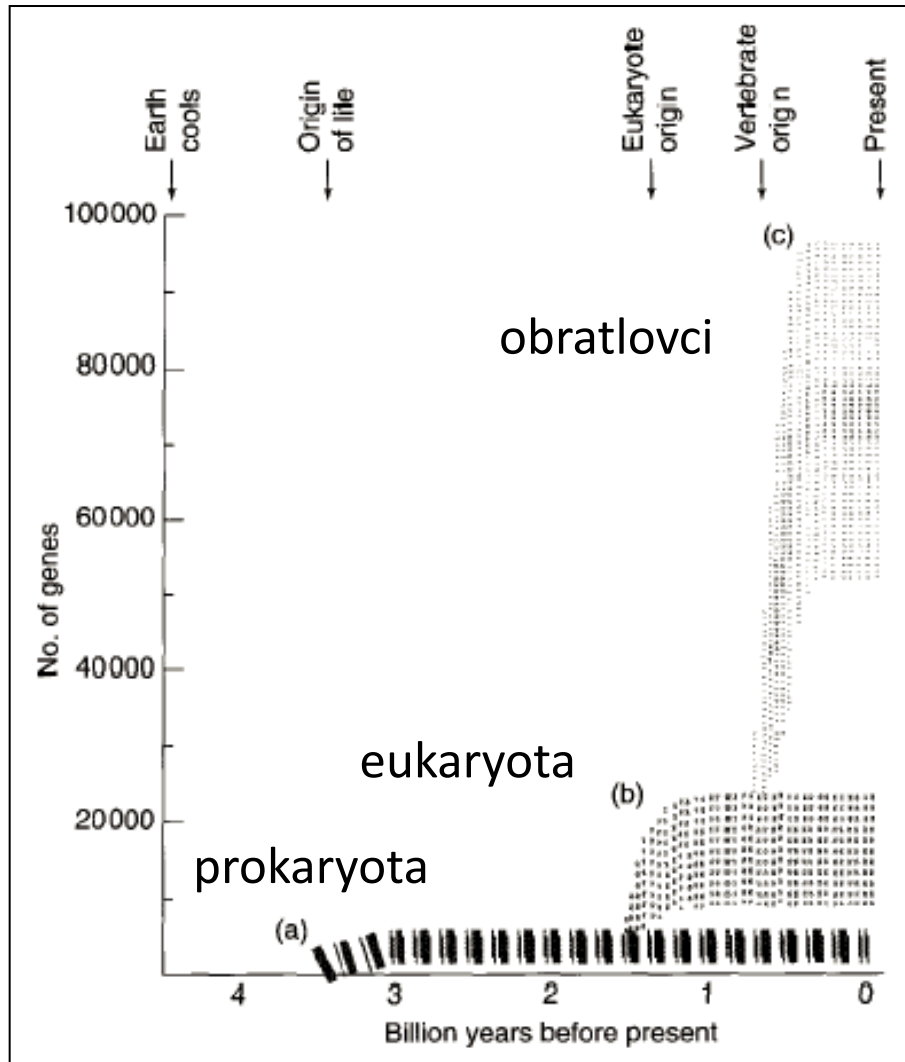
Satelitní DNA:

Numt, Nupt:

- inserce promiskuitní DNA



Vzrůst počtu genů, potlačení šumu a biologická komplexita



Prokaryota/eukaryota:

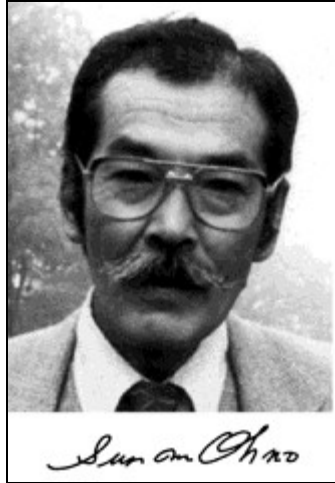
Potlačení šumu - [separace](#)
transkripce a translace, jaderná
membrána a histony

Obratlovci:

Potlačení šumu – [metylace](#)
genů, mobilních elementů a
duplikátů

[RNA interference](#) (smallRNAs)

Evolve genomů - citáty



„The majority of the genome consisted of now-extinct genes“

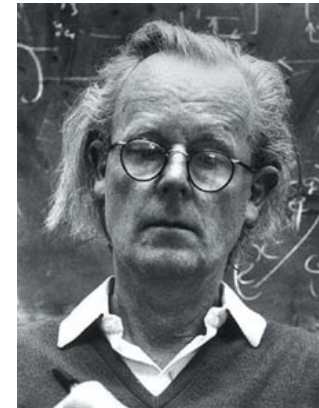
- *Susumu Ohno*

“Triumphs as well as failures of nature's past experiments appear to be contained in our genome”

- *Susumu Ohno*

„The evolution of the large-scale features of the genome is one of the most difficult, perhaps the most difficult, question in evolutionary biology“

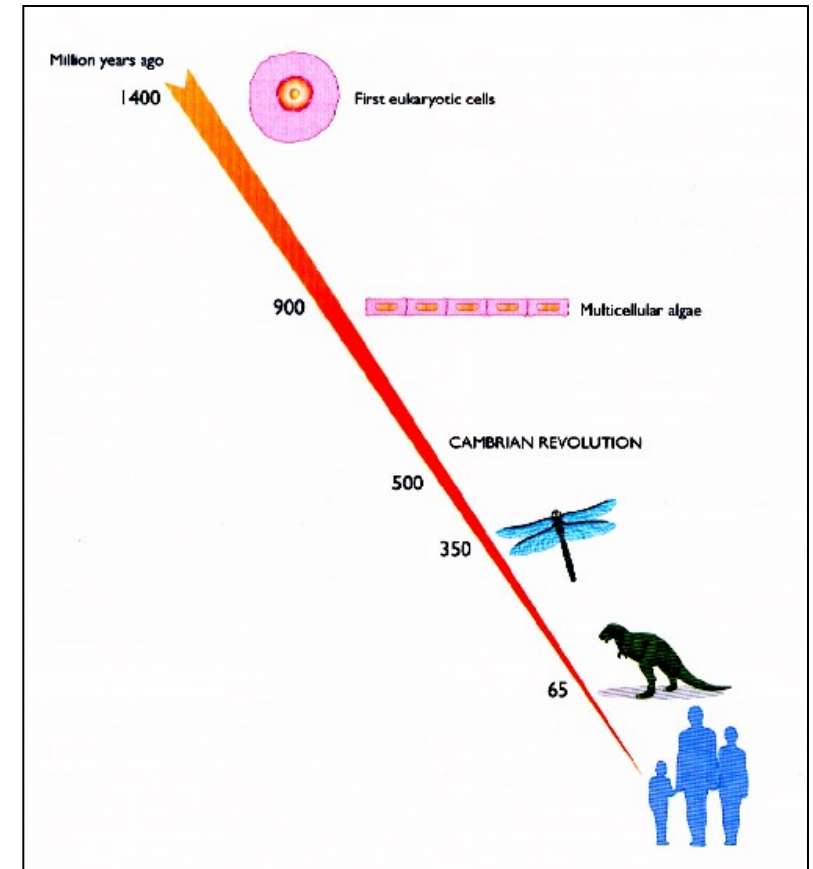
- *Maynard Smith*



MECHANIZMY ZMĚN VE VELIKOSTI GENOMU

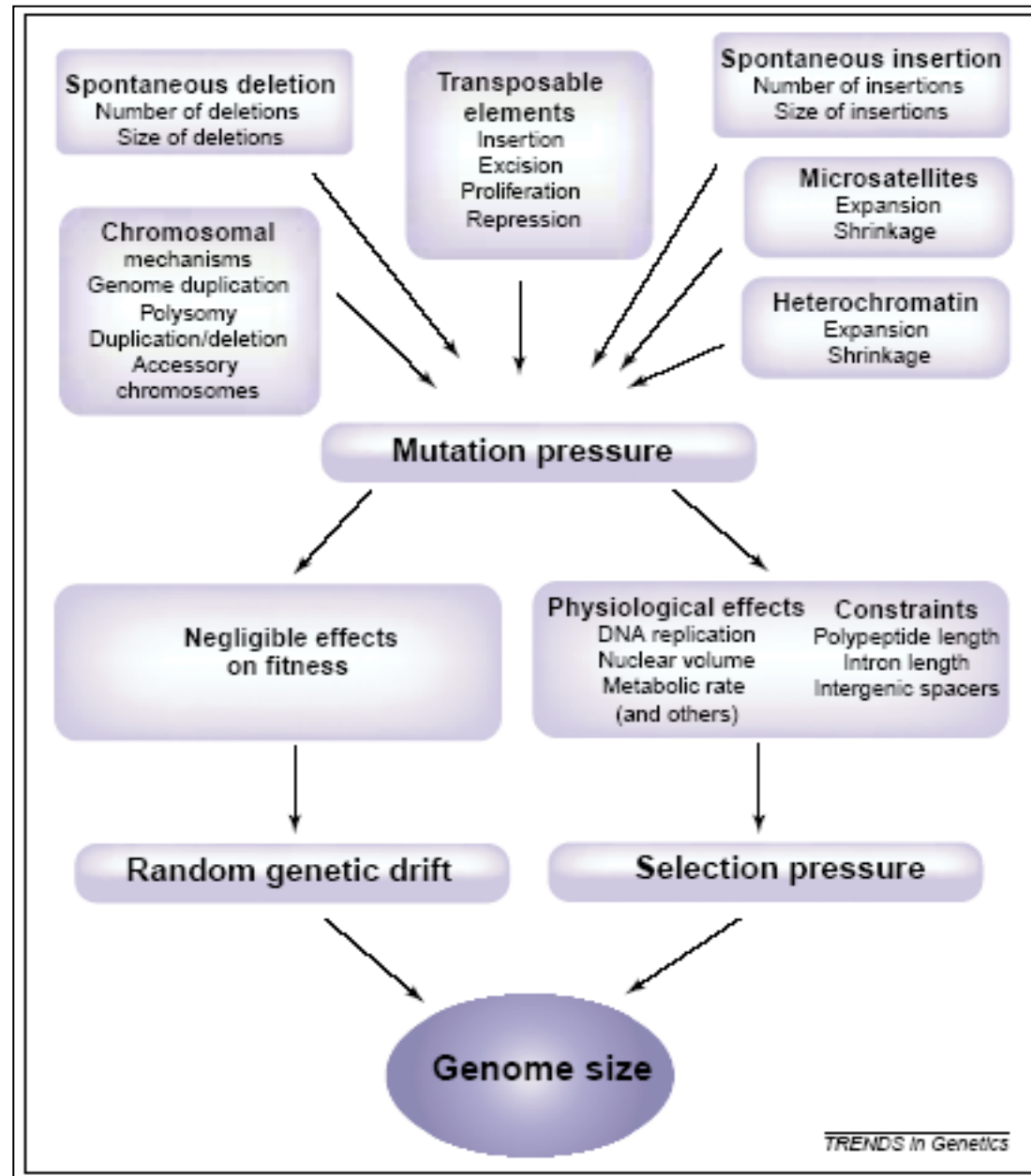
Mechanizmy změn velikosti genomu

- Polyploidizace (duplikace genomu)
- Duplikace a delece chromozomů nebo jejich částí (chromozomové přestavby)
- Inzerce/delece/duplikace genů
- Amplifikace transponovatelných elementů
- Inzerce virové DNA
- Inzerce organelové DNA
- Expanze mikrosatelitů



Vzrůstající komplexita živých forem byla doprovázena vzrůstem velikosti genomů a počtu genů

Evoluční síly ovlivňující velikost genomu



Změny ve velikosti genomů: plynulé nebo skoky?

Plynulost:

- savci, ptáci, kostnaté ryby
- postupná akumulace nebo delece, malé segmenty DNA
- normální distribuce velikostí genomů v těchto skupinách
- podíl jedinečných sekvencí konstantní, mění se spolu se změnou velikosti genomu, je dán velikostí intronů

Skoky (kvantové změny):

- bezobratlí, rostliny (50% nahosemenných je polyploidní), vyjímečně obratlovci
- polyplodizace, přídavná replikace některého chromosomu, kryptopolyploidizace (koordinovaná replikace všech repetací v genomu – *Chironomus*)
- kvantové rozdíly mezi blízkými příbuznými, velké rozdíly ve velikostech genomů

Polyploidizace

Pojmy z hlediska původu:

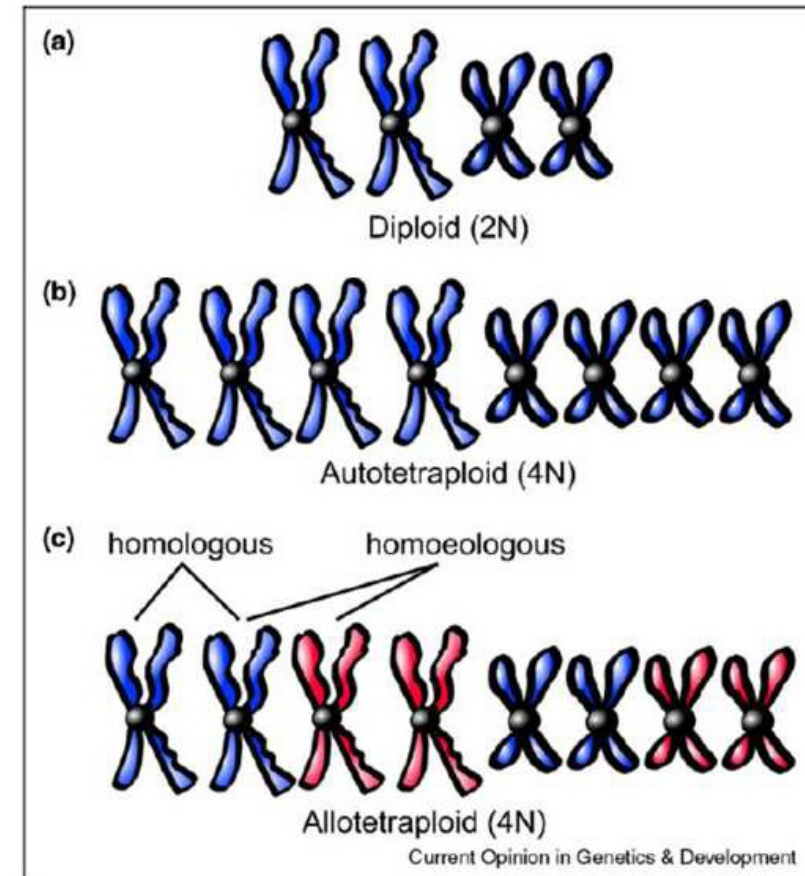
Autopolyploidie (polyploidizace v rámci druhu = jen 1 druh genomu)

Allopolyploidie (polyploidizace mezidruhovou hybridizací = 2 a více druhů genomu)

Allopolyploidie je zřejmě výhodnější:

- Dva různé genomy = více alel a proteinových interakcí
- Dochází k heteróznímu efektu, hybridy mají širší ekologický záběr než rodičovské druhy, mají vyšší rezistenci k parazitům než rodiče

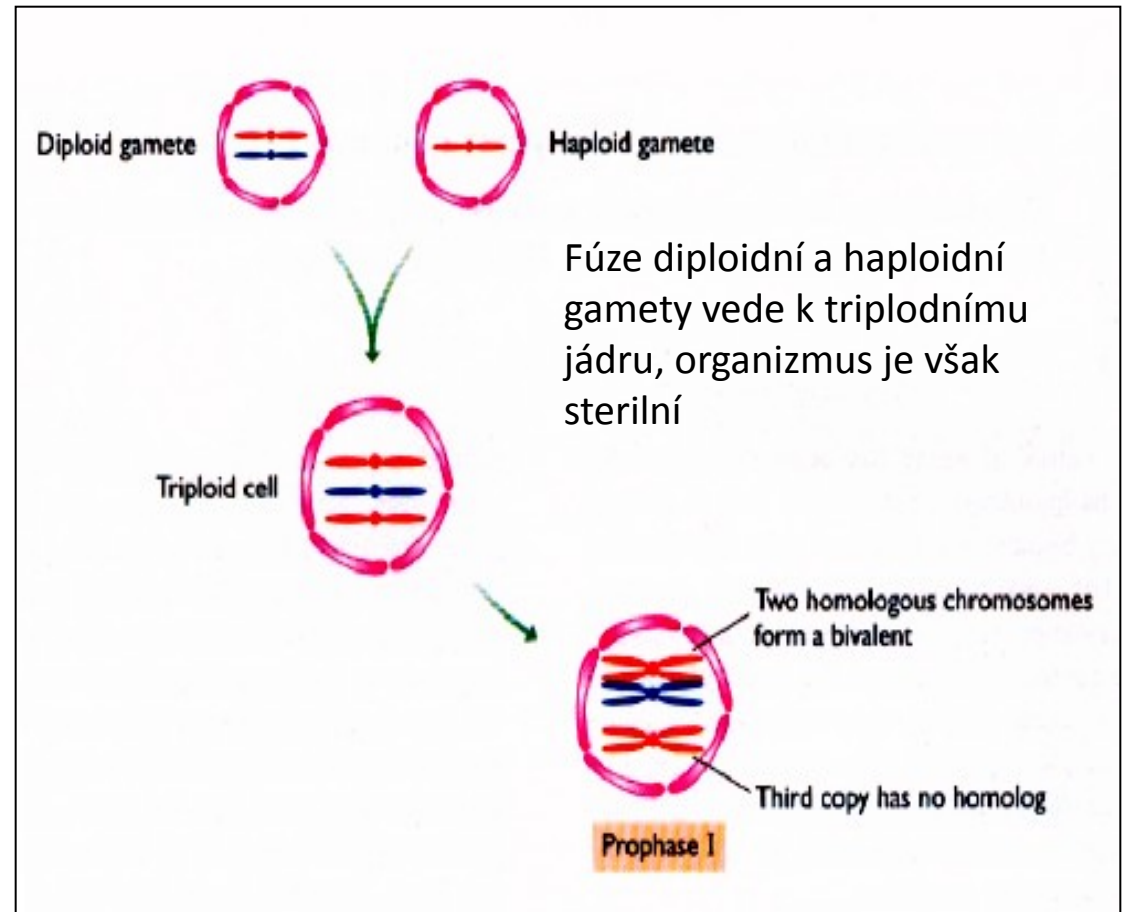
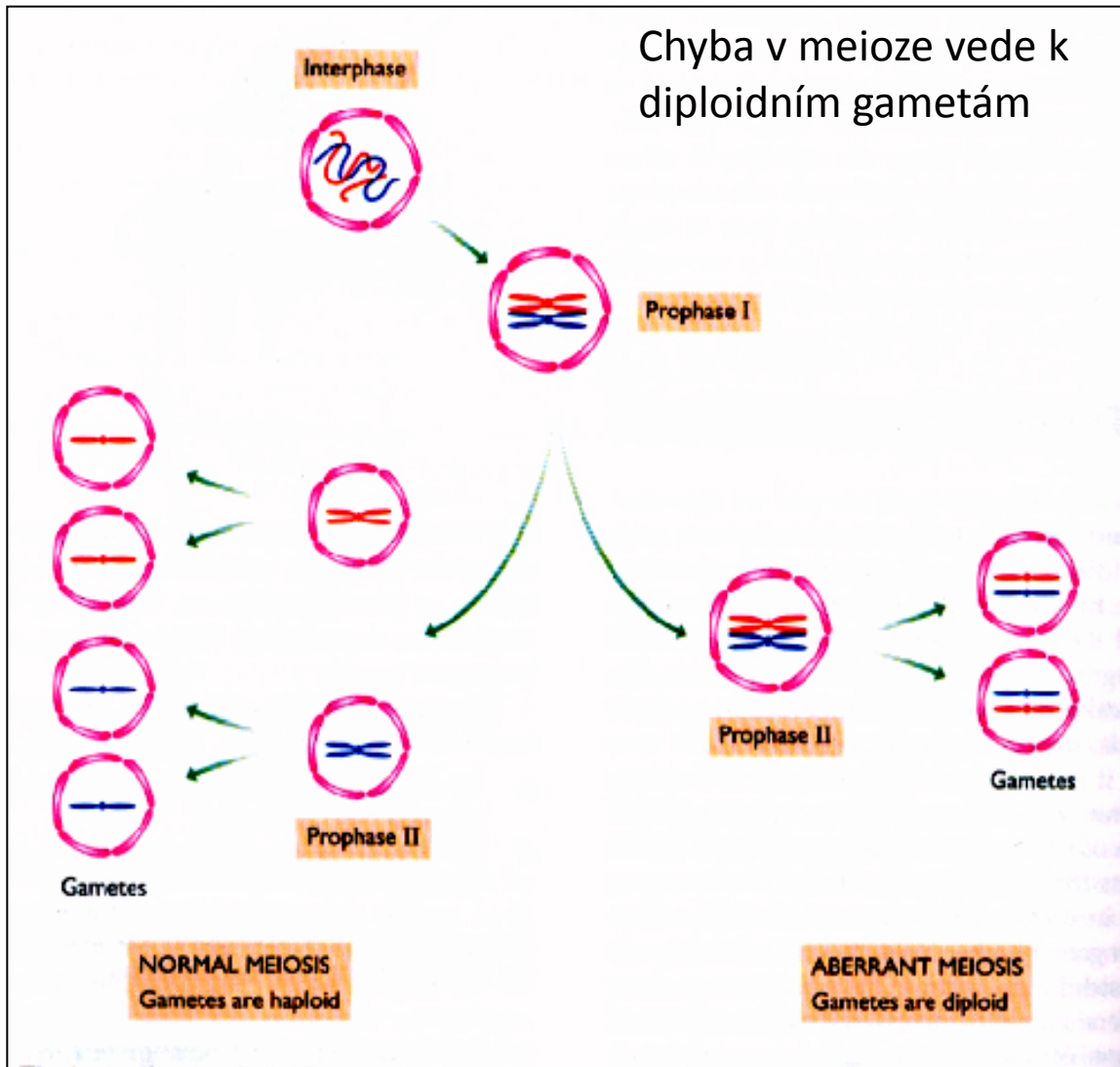
Kryptopolyploidie: dávná polyploidizace, organizmus se jeví jako diploidní, přestavby – delece, translokace..., kvasinka, obratlovci, obilniny, odhalení až sekvenací



Homologní vs. homeologní
chromosomy

Co vede k polyploidizaci?

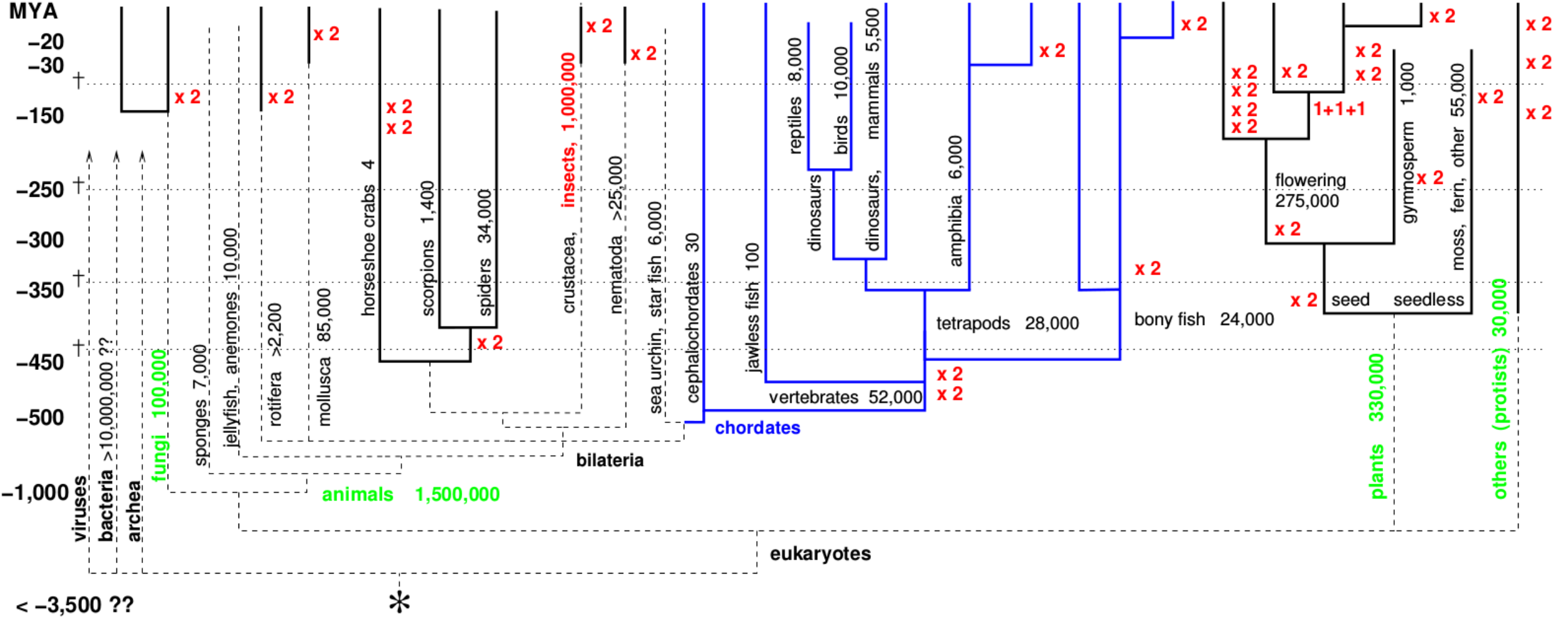
Chyba v meioze vede k diploidním gametám



Duplikace genomu v historii eukaryot



K. lactis *S. cerev.* *Bdelloid* *Bulinus* Horseshoe Scorpion Spider *Nadis* *M. incognita* Lancelet Lamprey *H. sapiens* *X. tropicalis* Gar Tetraodon Banana Poplar *A. thaliana* *B. napus* Paramecium



Duplikace genomu v historii obratlovců

Duplikace genomu obvykle následována masovou speciací (2x na bázi obratlovců, další u Teleostei + další u některých skupin obratlovců (drápatky, lososovití,...))

Susumu Ohno: Hypotéza „2R“ (2 rounds) dvě genomové duplikace u obratlovců před cca 450 mil. lety

- některé genové rodiny mají 4 členy u obratlovců a jednu u bezobratlých
- srovnání ježovky (bezobratlý) s obratlovcem (člověkem, myší a rybou Fugu)
- v genomu člověka dlouhé segmenty ve čtyřech kopiích podél všech chromosomů

Počty druhů obratlovců (celkem 62 tisíc)

Savci – cca 5,5 tis

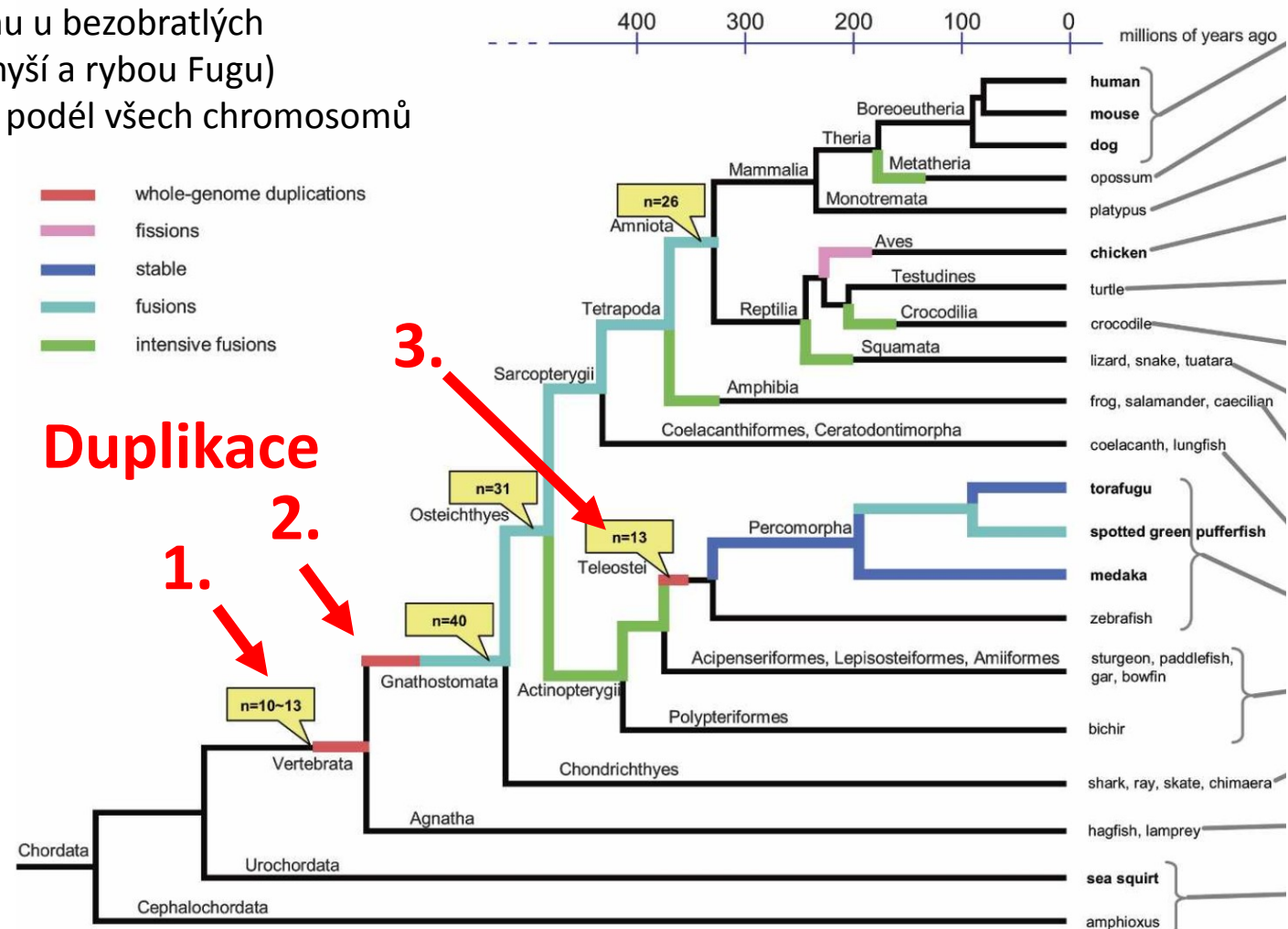
Ptáci – cca 10 tis

Plazi – cca 9 tis

Obojživelníci – cca 6,5 tis

Ryby - > 31 tis

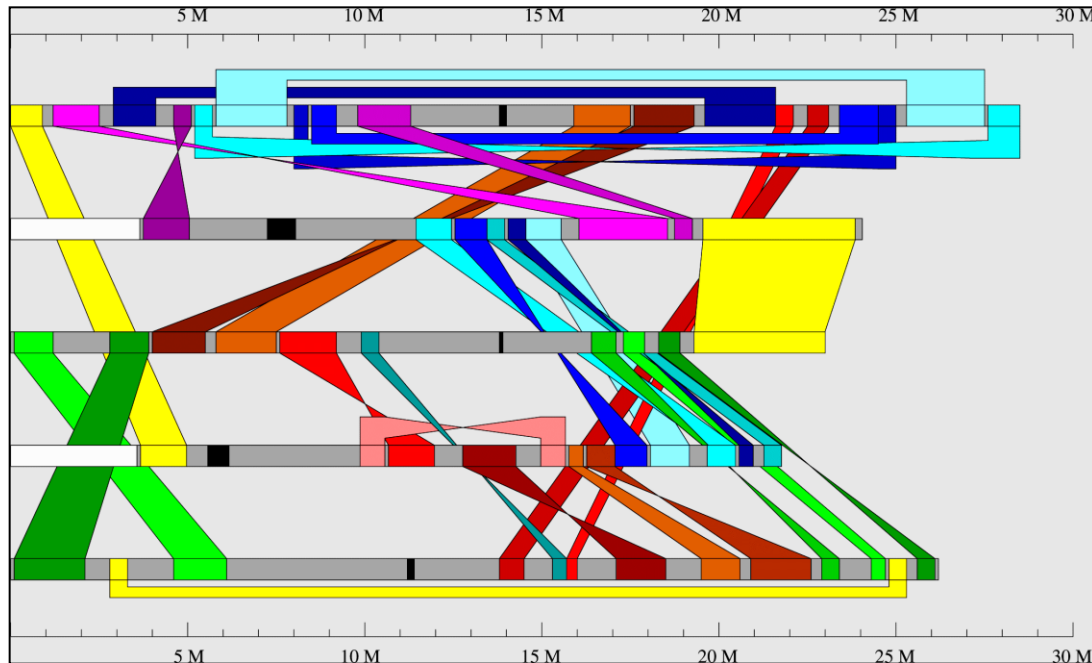
Průměrná doba přežívání genů vzniklých duplikací genomu je mnohem vyšší než u lokálních duplikací genomu (duplikace celých drah)



Polyploidizace u rostlin

Arabidopsis:

- duplikace před 200 a 80 mil. let
- možná i více než 4 genomové duplikace
- asi 60% genomu leží v duplikovaných segmentech
- 50% genů v nich je konzervováno



Polyploidní plodiny

- **tetraploidi:** kukuřice, bavlna, brambor, zelí



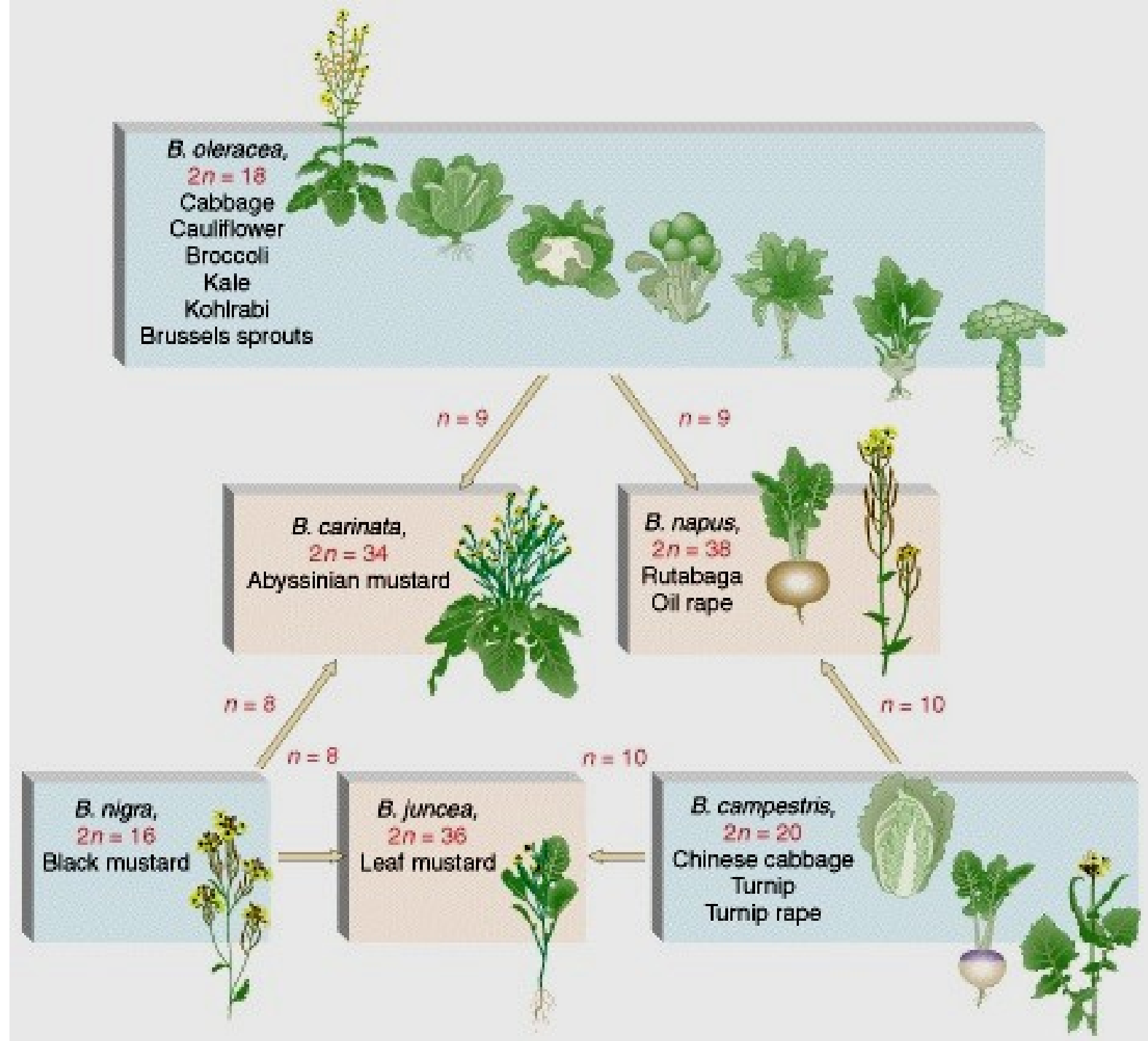
- **hexaploidi:** pšenice, chrysanéma



- **oktoploidi:** jahodník

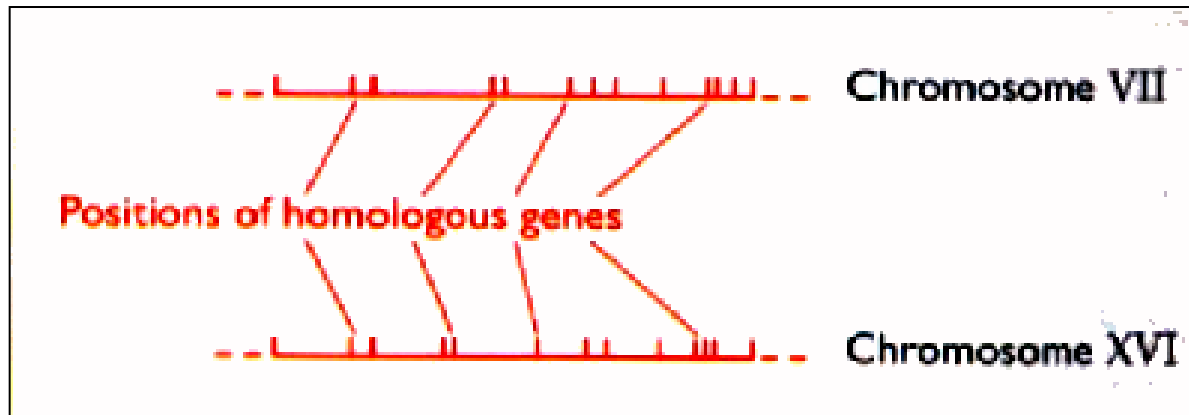


Allotetraploidie brukvovit 



Genom **kvasinky**: Polyploidizace nebo duplikace segmentů?

- 54 nepřekrývajících se duplikovaných oblastí
- 50 z nich si zachovalo stejnou orientaci vzhledem k centroméře
- nebyla nalezena žádná triplikovaná oblast, přičemž 7 oblastí by bylo očekáváno dle Poissonova rozložení
- příbuzné druhy bez duplikace – *Kluveromyces waltii* a *Ashbya gossypii*
- 457 genových párů – 17% se vyvíjí rychleji než jejich homology u *K. waltii*



Duplikace jsou výsledkem tetraploidie spíše než postupných regionálních duplikací

Změna regulace genomu u polyploidů

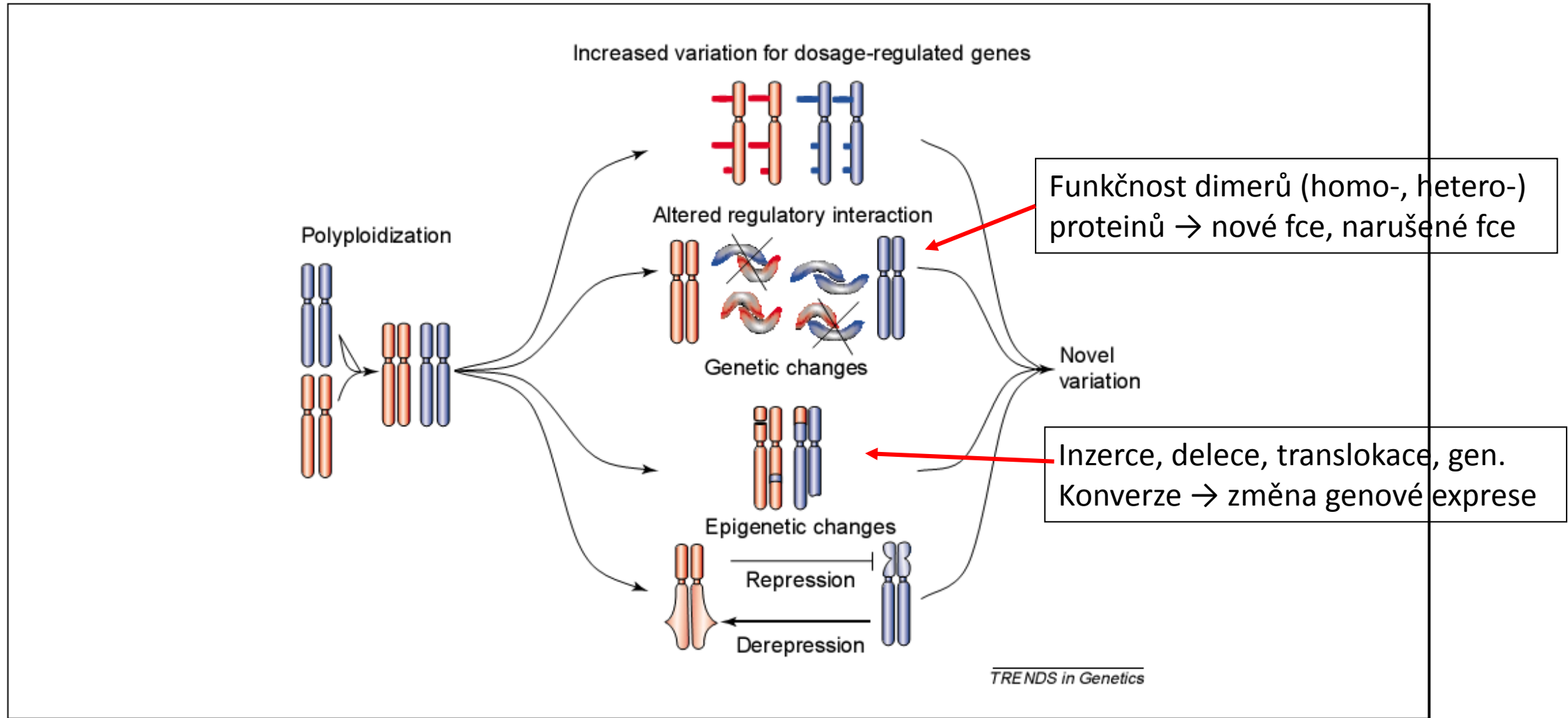


Fig. 4. Potential causes of novel variation in polyploids. The merger of chromosomes from two diploid genomes (red and blue) into a tetraploid genome can cause (1) increased variation of dosage-regulated gene effects and expression (magnitudes of allelic effects and expression shown by size of blocks for three loci); (2) altered regulatory interactions (trans-acting regulatory factors shown as dimeric proteins, with heterodimers not functioning properly); (3) genetic changes affecting gene expression (e.g. insertions, deletions, translocations and gene conversions); and (4) epigenetic changes (repression or derepression of gene expression caused by genome interaction of chromatin modeling factors, which could also trigger movement of transposable elements).

Segmentální duplikace

- Duplikace kratších (stovky bp až desítky kb) unikátních sekvencí s i bez obsahu genů
- Tandemové uspořádání nebo rozptýlené v genomu
- Množství se liší mezi druhy (drosophila málo – člověk mnoho)
- Segmentální duplikace zodpovídají za více než 90 % genetické variability mezi lidskými populacemi (Sudmant et al., Science, 2015)
- V nedávné historii člověka se fixovaly duplikace nesoucí geny se selekční výhodou – imunitní reakce, olfaktorické rozpoznávání, odbourávání xenobiotik, reprodukce, proliferace neuronů (velikost mozku), gen pro amylázu AMY1 (3x více než šimpanz, strava bohatá na škrob).

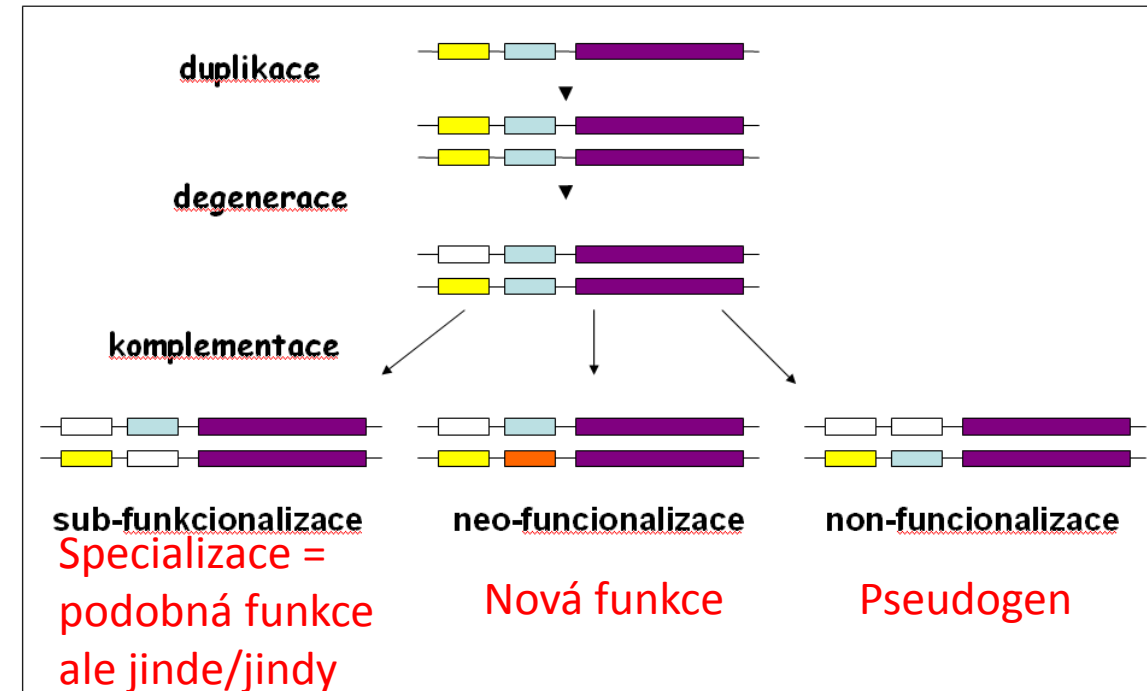
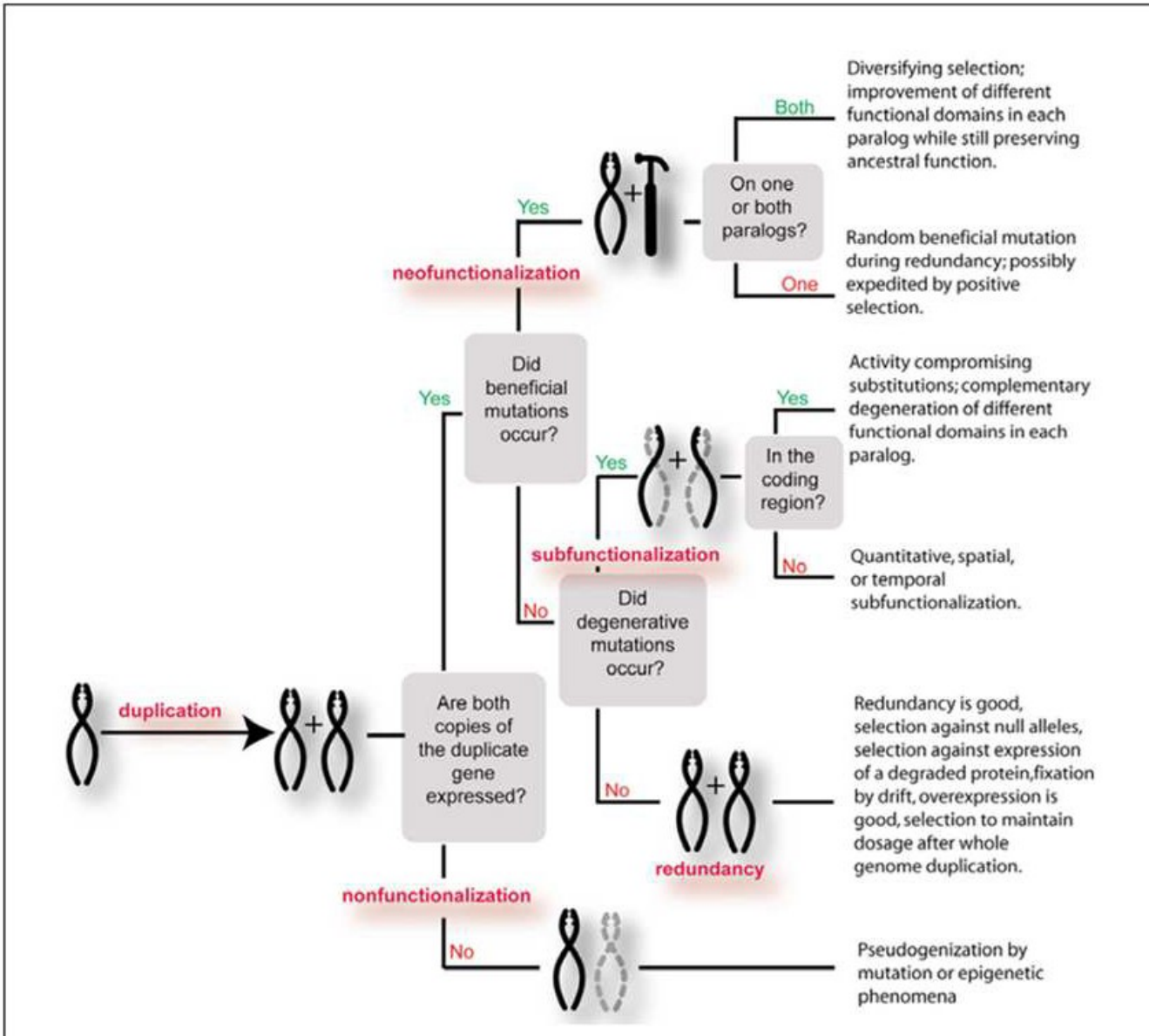
Table 1 | **SD content of sequenced animal genomes**

	<i>Caenorhabditis elegans</i>	<i>Drosophila melanogaster</i>	Human	Mouse	Rat	Chicken	Chimpanzee*
SDs of >1 kb	4.3%	1.2%	5.2%	2.7%	1.6%	2.7%	N.D.
SDs of >10 kb	0.7%	0.1%	4.5%	2.2%	1.5%	0.3%	N.D.
SDs of >20 kb	N.D.	N.D.	4.0%	1.7%	0.9%	0.0%	~4.8%
Genome size	97	123	2,866	2,506	2,566	1,040	2,866

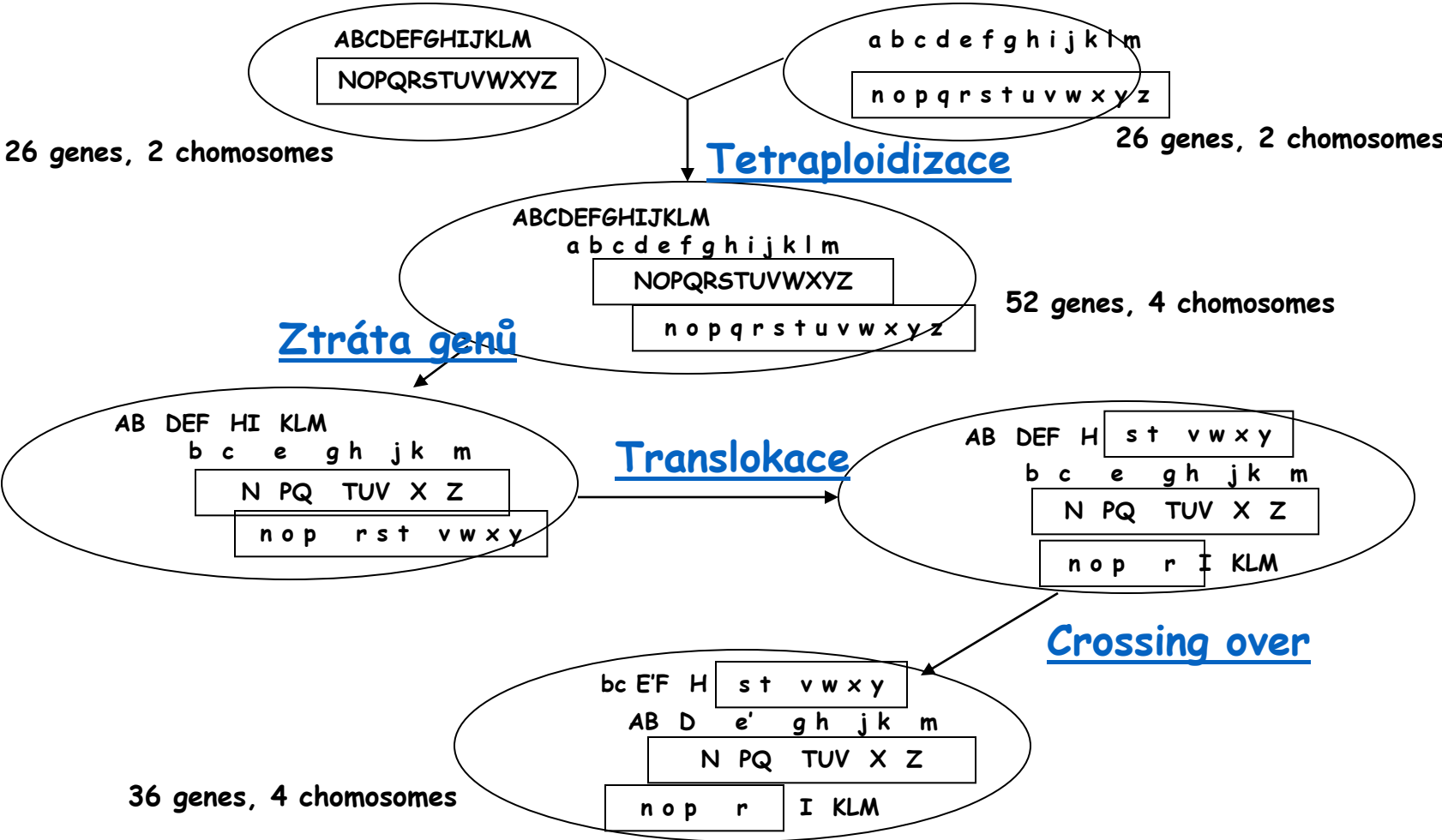
Data taken from REFS 2,7 for pairwise segmental duplications (SDs) with >90% identity. *Given the fragmented nature of SDs in the draft chimpanzee genome, the duplication content can only be estimated indirectly on the basis of human duplication content, adjusting for detected differences in SD compared with chimpanzee whole-genome shotgun sequencing⁶. DNA not assigned to a chromosome was not included in these calculations. Consequently, in other genomes the estimate of recent duplication might rise as the quality of the sequence assembly improves. N.D., not determined.

Bailey & Eichler, NRG, 2006

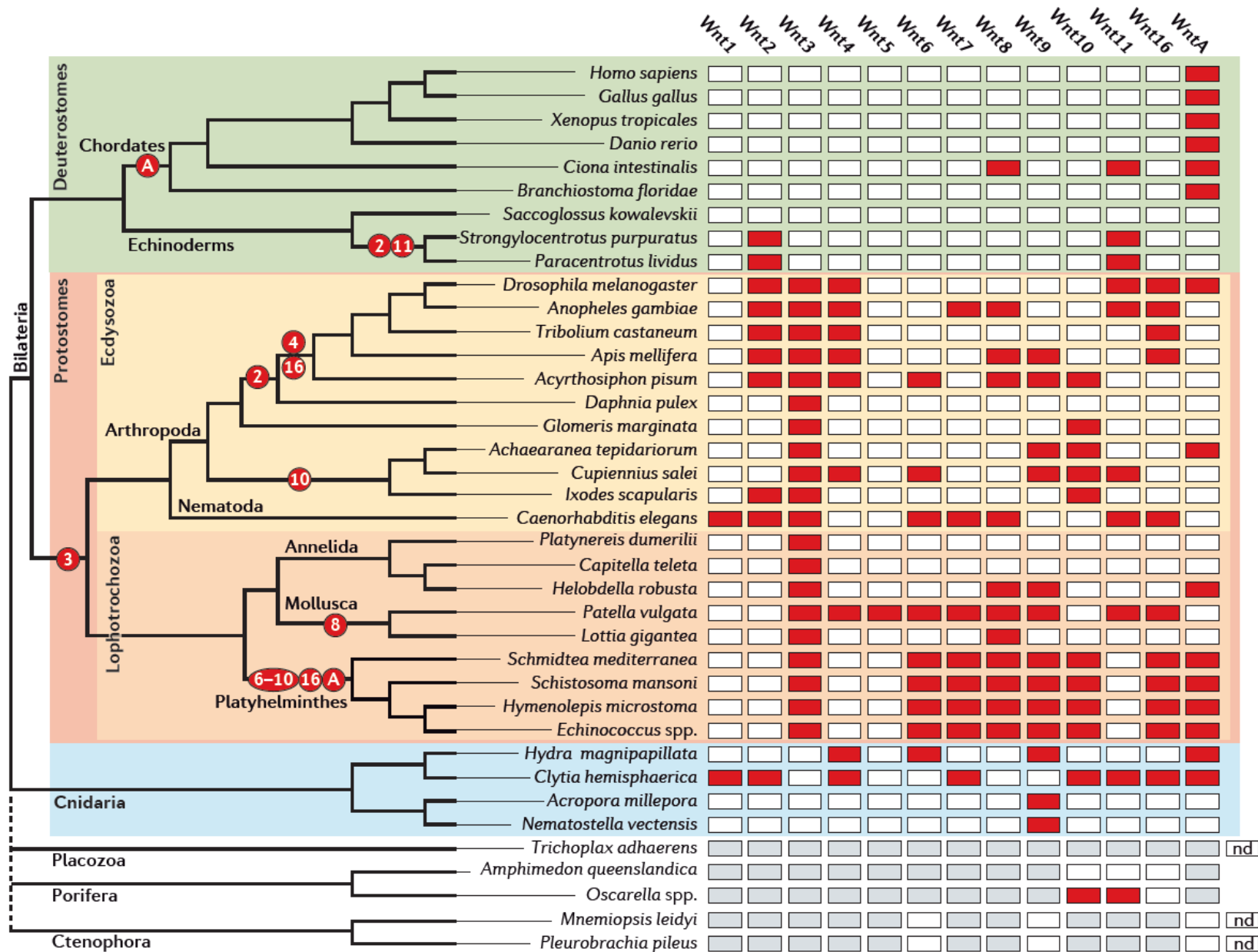
Osudy genů po duplikaci



Diploidizace: postupné snížení počtu genů a delece částí genomu po duplikaci



Evoluce ztrátou genů



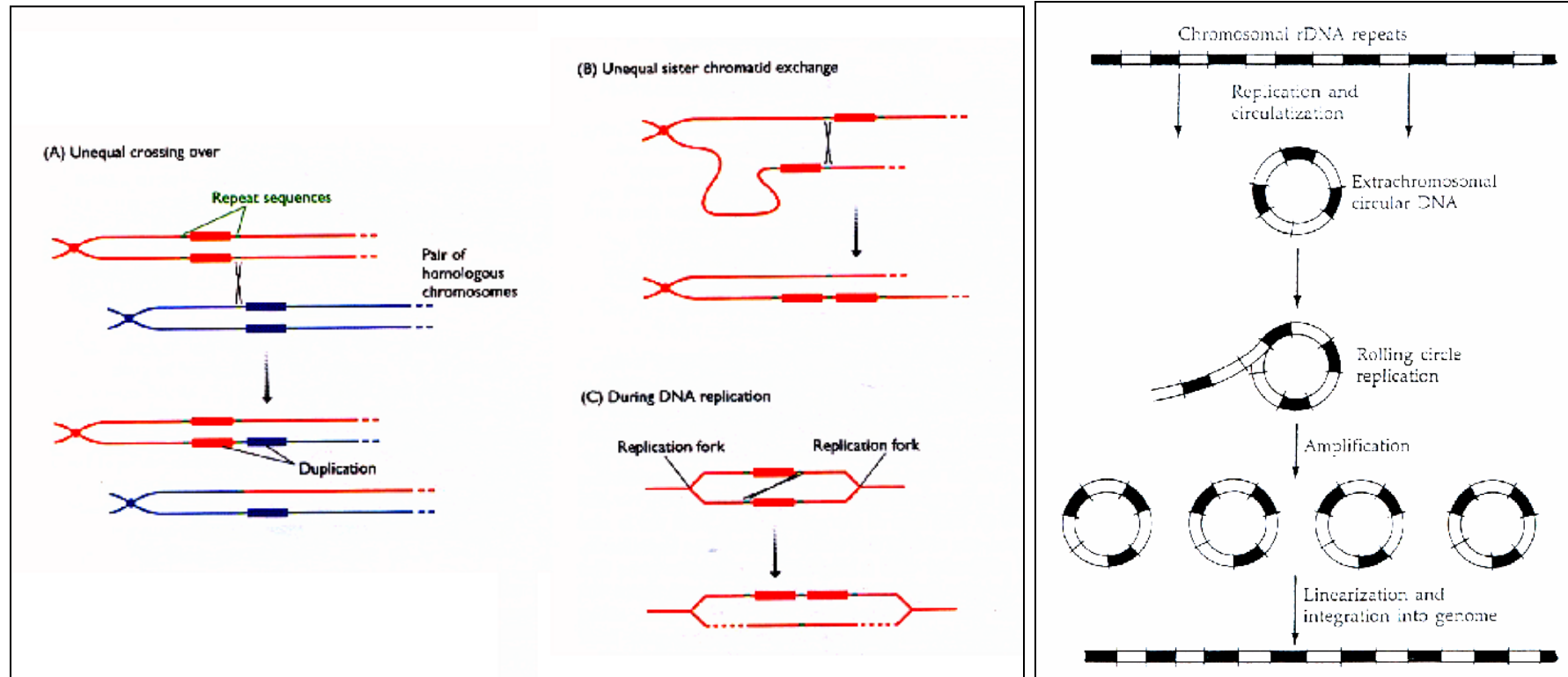
Červená = ztráta genu

Figure 1 | The wingless (Wnt) family: a paradigmatic example of the pervasiveness of gene loss during metazoan evolution. In the past decade, the accumulation of fully sequenced genome data from various species has revealed great

NEKÓDUJÍCÍ DNA a VELIKOST GENOMU **(postupné změny velikosti)**

Mechanismy zvětšení genomů - postupně

- Duplikativní (retro)transpozice
- Nerovnoměrný (unequal) crossing-over
- Replikační klouzání (Replication slippage)
- Genová amplifikace (rolling circle replication)

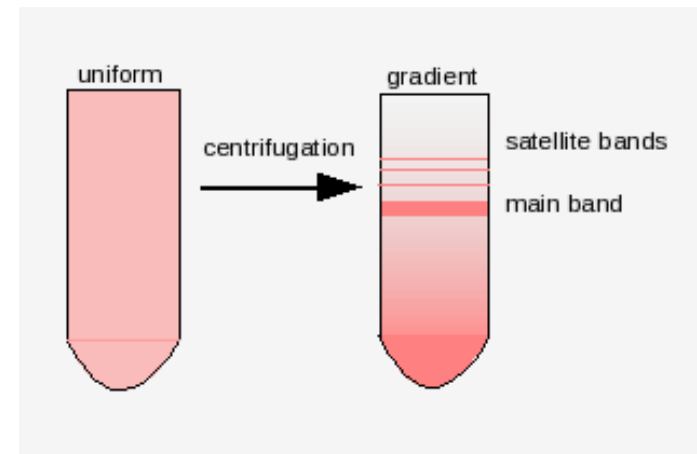


Tandemové repetice = Satelity

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



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

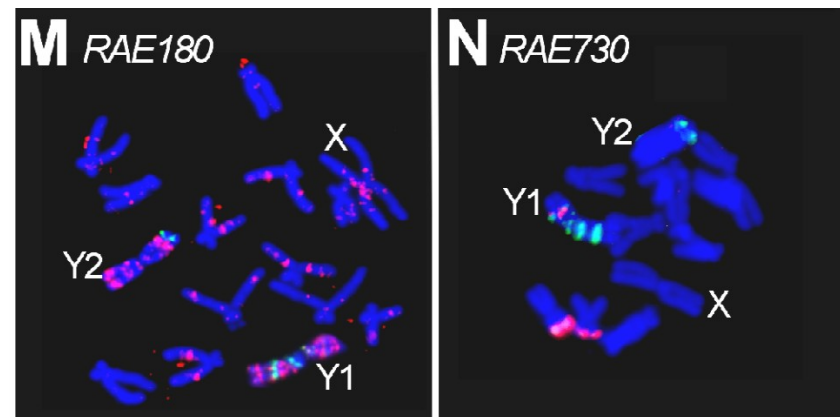


- Přítomny ve všech genomech (prokaryota i eukaryota)

- 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 ≤ 6 bp

Typické rozložení tandemových repetic na chromosomech šťovíku

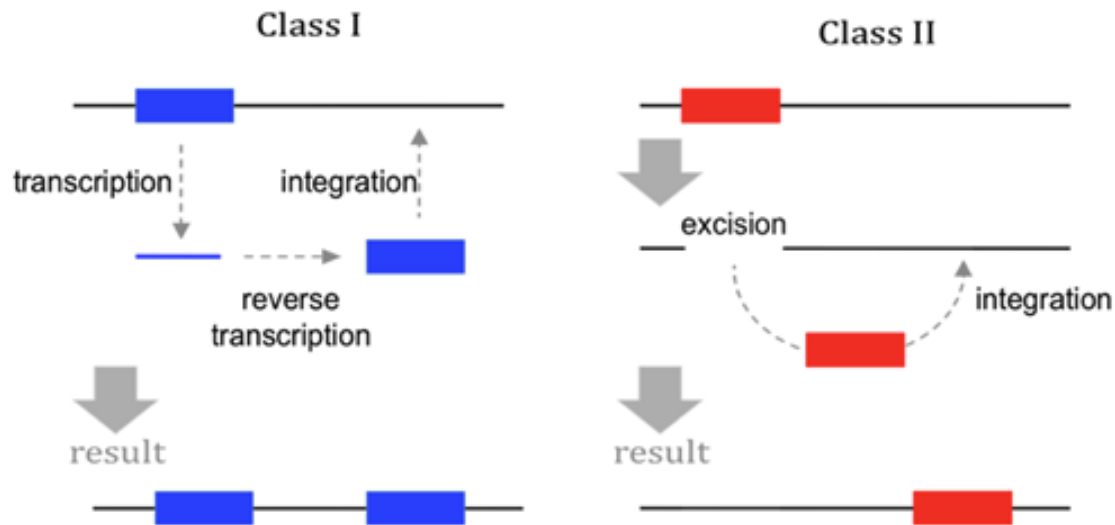


Strukturní a regulační funkce, expanze mikrosatelitů u některých chorob

Mobilní elementy

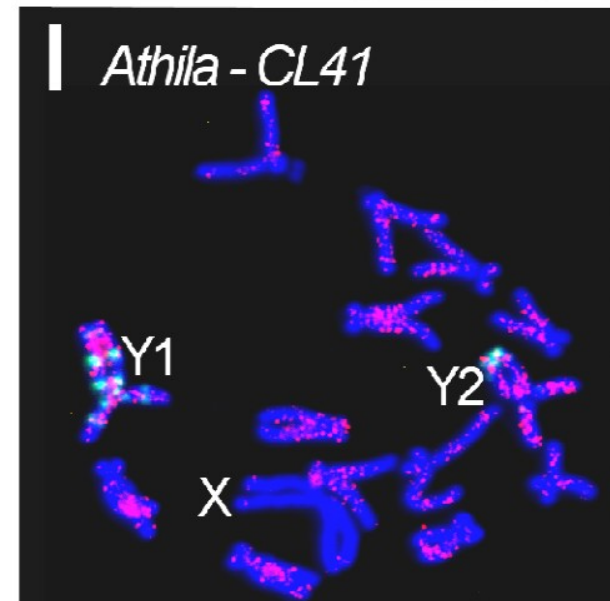
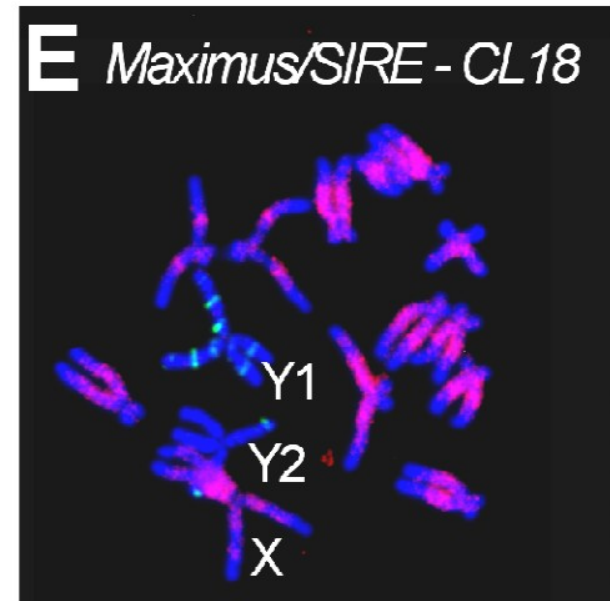
(transposable elements = TEs)

- Schopné samostatného pohybu v genomu
- Typy mobilních elementů:
 - **Třída I – retrotransposony** (copy-and-paste, replikace přes RNA)
 - **Třída II – DNA transposony** (cut-and-paste)



- Pro jedince spíše negativní – inzerce TE může být narušena funkce genů
- Klíčoví hráči v evoluci genomu (vznik nových genů, přestavby chromosomů, ...)
- Často součást heterochromatinu

TEs na chromosomech šťovíku



Teorie úlohy nekódující DNA: adaptivní role vs. sobecká DNA

- Jaké evoluční síly produkují „zbytečnou“ DNA?
- Jaká je role „zbytečné“ DNA?
- Proč selekce toleruje „zbytečnou“ DNA?

Adaptivní role:

- vliv nadbytečné DNA na fenotyp
- vliv na velikost jádra a buňky
- ochrana kódujících sekvencí před mutacemi
- pufování koncentrace regulačních proteinů

Sobecká DNA (junk DNA):

- jsou to jen parazitické sekvence, mobilní elementy, fixace genetickým driftem, velikost genomu je tolerovatelné maximum závisující na ekologických a vývojových potřebách organismu



Genomová obezita u rostlin – jednosměrný proces?

Teorie sobecké DNA (Dawkins, 1976)
- jediným cílem DNA je se replikovat

Dva modely:

- pouze růst genomů
- růst i zmenšování genomů

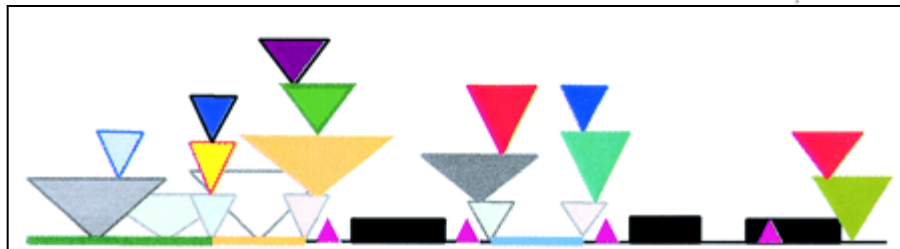
Limity růstu genomů:

- **fitness** hostitele
- mechanismy potlačující amplifikaci – **metylace**
- mechanismy odstraňování repeticí – **rekombinace**
 - BARE – 42x více soloLTR než vnitřních částí (dávná aktivita)
 - rekombinace uvnitř elementů i mezi elementy
 - selekce na kratší LTR, solo LTR nejdelší

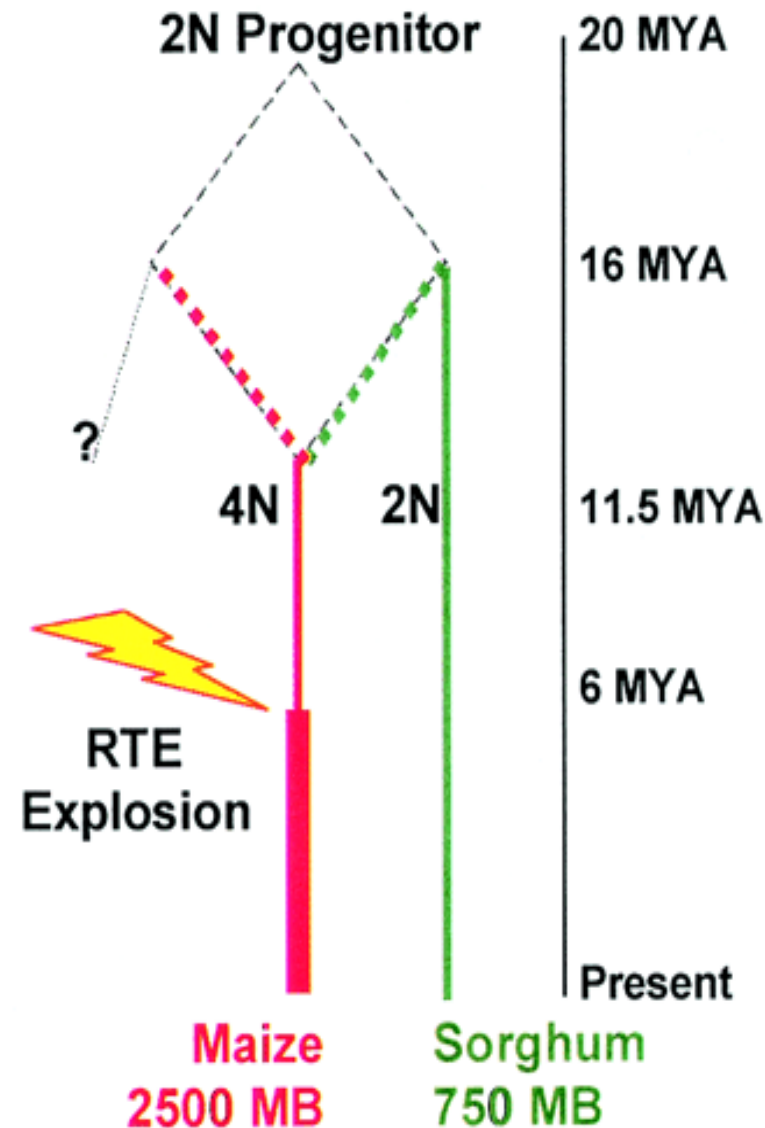


Big Bang v genomu kukuřice

- polyploidizace
- amplifikace TE (6 mil let)
- inserce do mezigenových oblastí

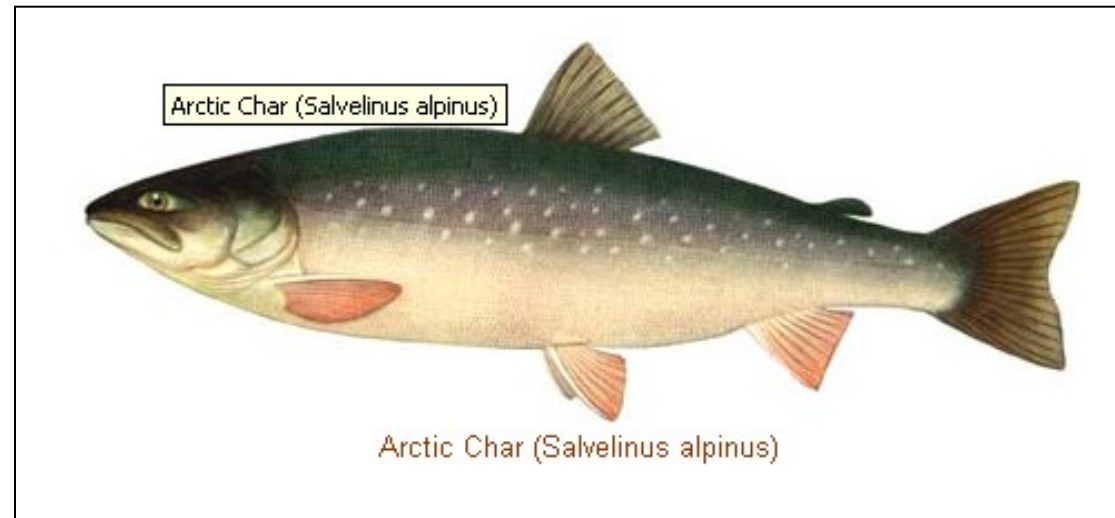
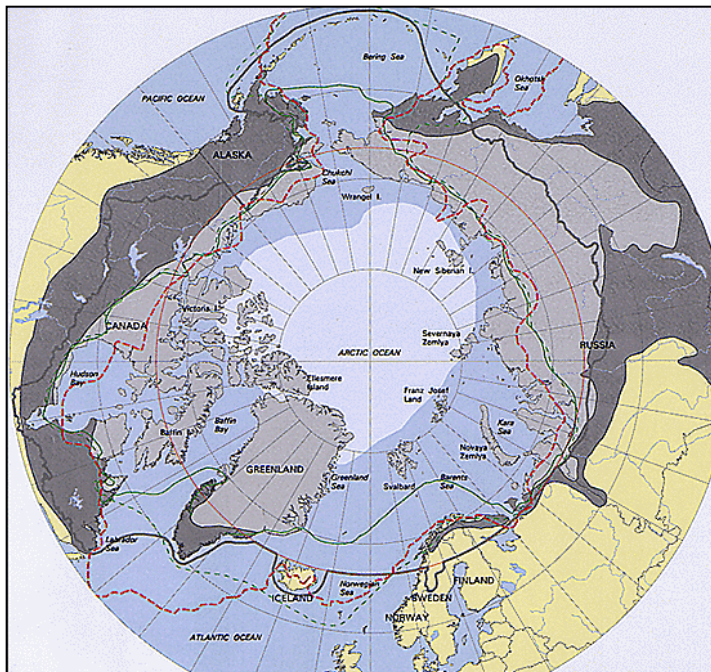


Commentary PNAS 98, 8163-8164 (2001)
Gene galaxies in the maize genome
Virginia Walbot* and Dmitri A. Petrov
Department of Biological Sciences, 385 Serra Mall, Stanford University, Stanford, CA 94305-5020



Závislost velikosti genomu na zeměpisné šířce a nadmořské výšce

- korelace mezi velikostí genomu a **teplotním režimem**
- větší genomy nebo **polyploidii**:
 - **arktické** lososovité ryby
 - zooplankton arktických jezer (*Daphnia*, *Bosmina*)
 - rostliny v polárních oblastech
 - populace v teplých oblastech jsou diploidní



Fenotyp vs. velikost genomu

Velikost genomu ovlivňuje:

- 1) Velikost buňky (pozitivně)**
- 2) Komplexitu organismu (negativně při zachování velikosti těla)**
- 3) Rychlost metabolismu (negativně)**
- 4) Rychlost buněčného dělení (negativně)**
- 5) Rychlost diferenciacce buněk (negativně)**

Velikost genomu ovlivňuje komplexitu mozku

Ocasatí obojživelníci mají jednodušší mozek než většina žab a ostatní obratlovci

- U obojživelníků byla testována korelace mezi velikostí genomu a komplexitou mozku
- Druhy s velkým genomem -> velké buňky - > jednodušší mozek
- Obojživelníci mají ancestrálně malé genomy – genom nezávisle mnohokrát zvětšen -> sekundární zjednodušení mozku



Salamandra salamandra



Thorius narisovalis

Nejjednodušší mozky mají malé druhy mloků



Bombina orientalis
(velké buňky – jednoduchý mozek)

vs.

Xenopus laevis
(malé buňky – komplexní mozek)



Miniaturizace těla a velikost jádra

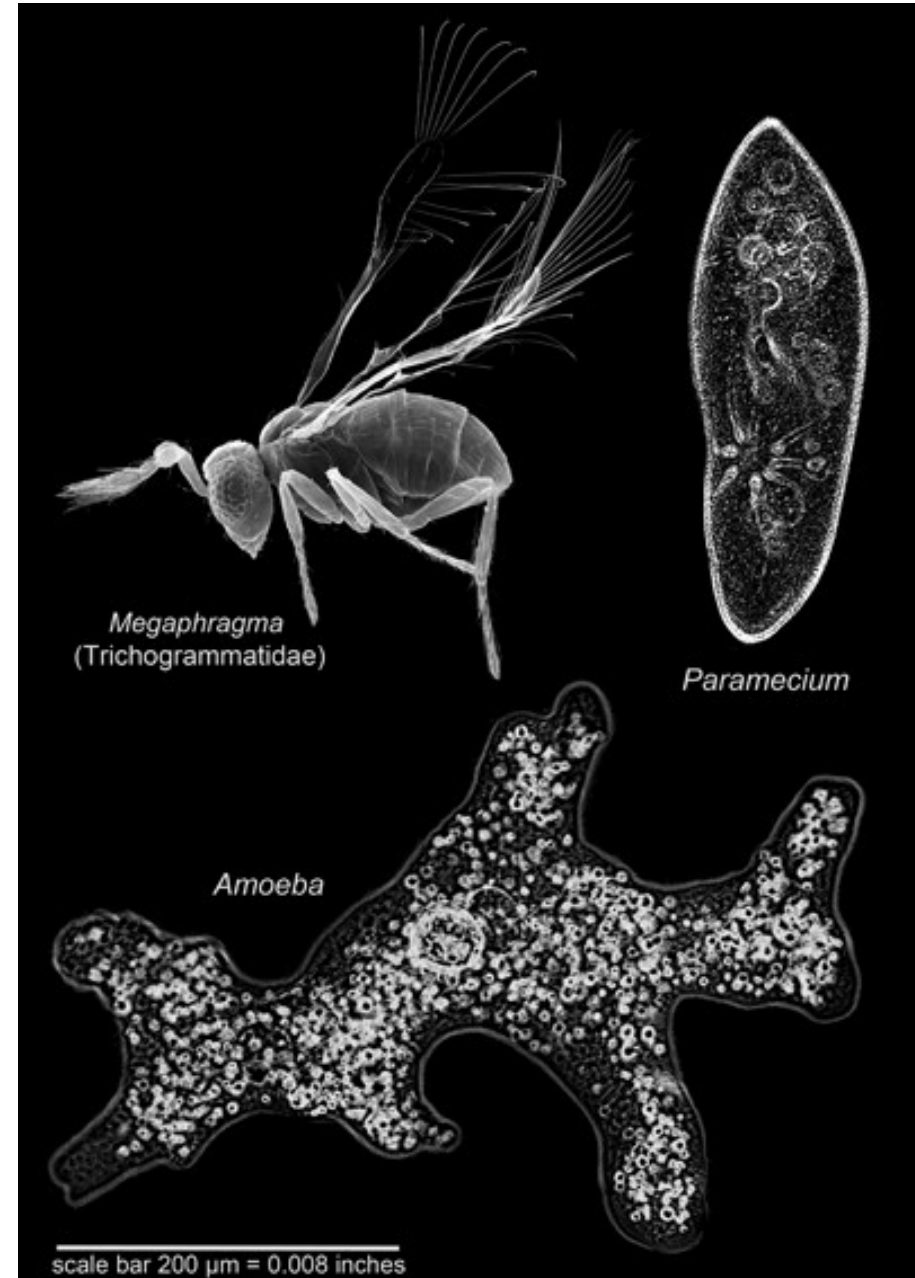
Megaphragma mymaripenne (Blanokřídlí)

Jeden z nejmenších létajících druhů hmyzu (0,2 mm)

-Larva se vyvíjí ve vajíčkách třásněnek
(maličká ->miniaturizace)

-Během metamorfózy zmenšení objemu těla

-asi 100x méně neuronů než včela, 95 % neuronů
bezjaderných



Malý genom -> rychlejší metabolismus

Ptáci mají menší genom než savci – vliv metabolismu?

Ptáci

- > dobří letci -> malý genom
- > špatní letci -> variabilní velikost



Savci

- > netopýři -> malý genom
- > ostatní savci -> větší a variabilní velikost



Pravděpodobné vysvětlení:

- Malý genom -> malá buňka -> velký poměr povrch/objem buňky
- > lepší výměna plynů -> rychlejší metabolismus výhodný při aktivním způsobu života

Mechanismus: redukce délky intronů, množství transpozonů...



Extrémní genomy: miniaturní genom eukaryotických parazitů - Encephalitozoon

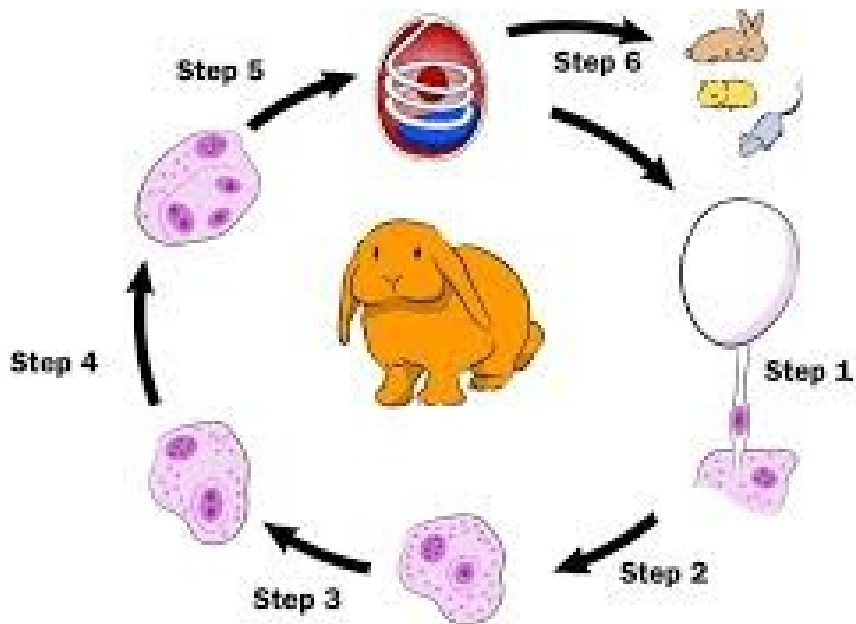
E. cuniculi

2,9Mb

1999 genů

Parazitická spájkivá houba (mikrosporidie)

Encefalitozoonóza králíků



E. intestinalis

2,3Mb

1833 genů

Vznik redukcí genomu *E. cuniculi*



- Některé parazitické organismy jsou jednodušší než jejich neparazitičtí předkové
- Zjednodušování je spojeno se ztrátou genů

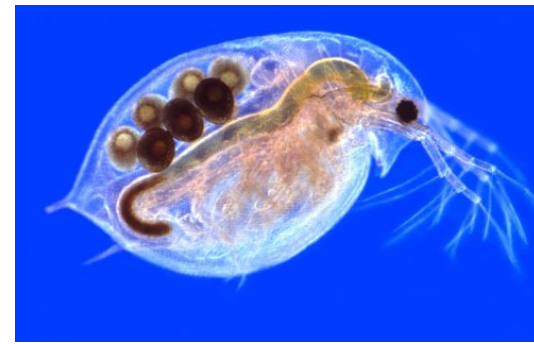
Obsah DNA je proměnlivý i v tkáních jedince (změny během ontogeneze)

Zvýšení obsahu DNA - endopolyploidie a polytenie:

- drosophila - polytenní chromosomy ve slinných žlázách
- *Daphnia* - tkáňově-specifické rozdíly v ploidii, 2-2048C, vliv na morfologii hlavy indukovanou predátorem
- *Bombyx mori* – 1 000 000 ploidní buňky žláz - hedvábí
- korelace ploidie a velikosti buňky

Snížení obsahu DNA v somatických buňkách:

- nematoda, bičíkovci, dvoukřídlí,
- např. *Cyclops strenuus* – somatické b. mají 5% obsahu DNA oproti oplozenému vajíčku
- vysvětlení z hlediska **sobecké DNA** – delece DNA ze slepé somatické linie a ne ze zárodečné linie
- změna obsahu DNA během života jako odpověď na stimuly – Lamarckismus



Daphnia pulex



Bombyx mori



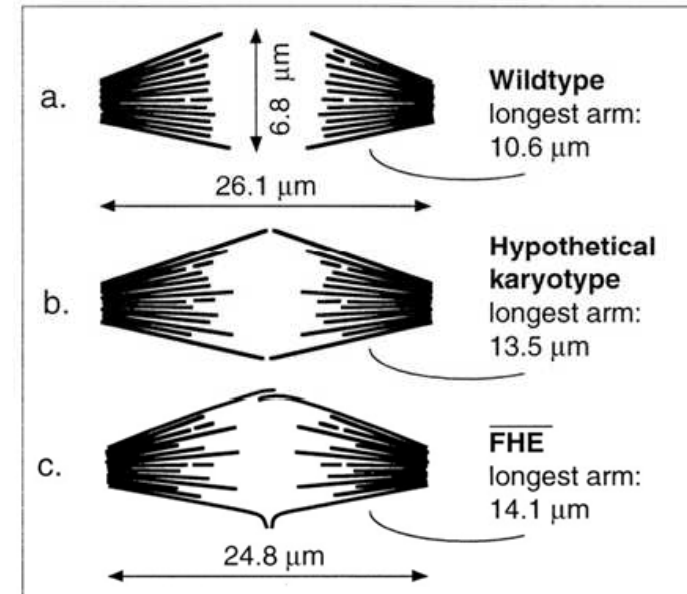
Cyclops strenuus

Evoluce karyotypu

(počet, tvar, velikost chromozomů)

Velikost chromozomů má své limity

- Obvyklá délka mezi 1 a 10 μm
- **Horní limit** – raménko delší než $\frac{1}{2}$ osy dělicího vřeténka – sesterské chromatidy se nemohou rozejít -> přetrženy dělicí se buňkou
- **Spodní limit** – velmi malé špatný rozchod v meióze – nedostatek crossing-overů a chiasmat?



Počet chromozomů - minimální

- Jednobuněčná eukaryota – obvykle 1 chromozom
- Mnohobuněčná eukaryota:



Myrmecia pilosula „Jack jumper ant“, Austrálie: $n=1$



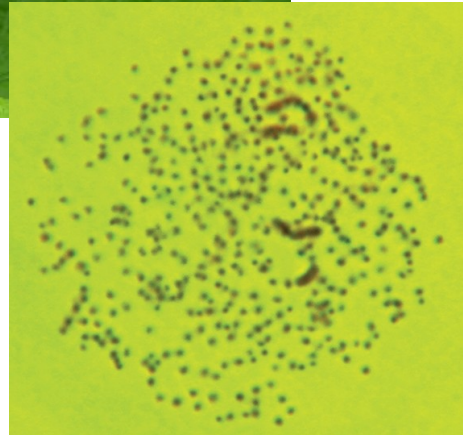
Haplopappus gracilis, Asteraceae, Severní Amerika: $n=2$



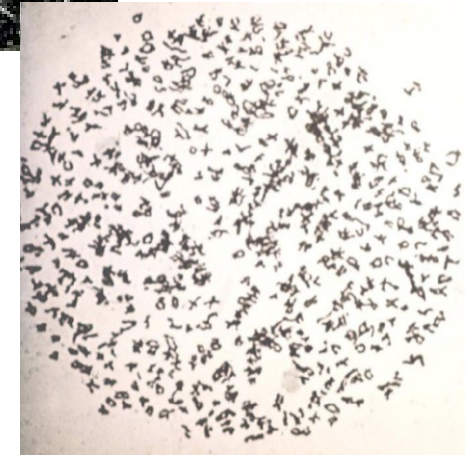
Počet chromozomů - maximální



Polyommatus atlanticus:
n = cca 220



Hadilka neboli hadí jazyk
Ophioglossum reticulatum:
n = cca 530



Horní limit počtu chromozomů je závislý na kapacitě dělicího vřeténka – dostatečný počet mikrotubulů omezen objemem buňky

Příklady počtu chromosomů

Druh		2n	Poznámka
škrkavka	<i>Parascaris univalens</i>	2	nejmenší známý počet
mravenci	<i>Myrmecia pilosula, M. croslandi</i>	2	nejmenší známý počet
modrásek	<i>Polyommatus atlantica</i>	440	největší známý počet u živočichů
kapradina	<i>Ophioglossum reticulatum</i>	1060	největší známý počet u rostlin
drozofila	<i>Drosophila melanogaster</i>	8	modelový organismus
myš	<i>Mus musculus</i>	40	modelový organismus
člověk	<i>Homo sapiens</i>	46	
muntžak červený	<i>Muntiacus muntjak vaginalis</i>	samice 6	nejmenší známý počet u savců
		samec 7	
osmák pouštní	<i>Tympanomyctomys barrerae</i>	102	největší známý počet u savců



Počty chromosomů se mohou lišit i u blízce příbuzných druhů

Počty chromosomů u různých druhů

Common Name	Species	Diploid number	Common Name	Species	Diploid number
Animals (2n)			Plants (2n)		
Human	<i>Homo sapiens</i>	46	Corn	<i>Zea mays</i>	20
Monkey	<i>Macaca mulatta</i>	42	Potato	<i>S. tuberosum</i>	48
Dog	<i>Canis familiaris</i>	78	Green algae	<i>A. mediterranea</i>	20
Cat	<i>Felis domesticus</i>	38			
Mouse	<i>Mus musculus</i>	40	Fungi (2n)		
Frog	<i>Rana pipiens</i>	26	Yeast	<i>S. cerevisiae</i>	32
Fruit fly	<i>Drosophila melanogaster</i>	8	Fungi (1n)		Haploid number
Flatworm	<i>Planaria torva</i>	16	Mold	<i>Penicillium species</i>	4

Extrémisté:

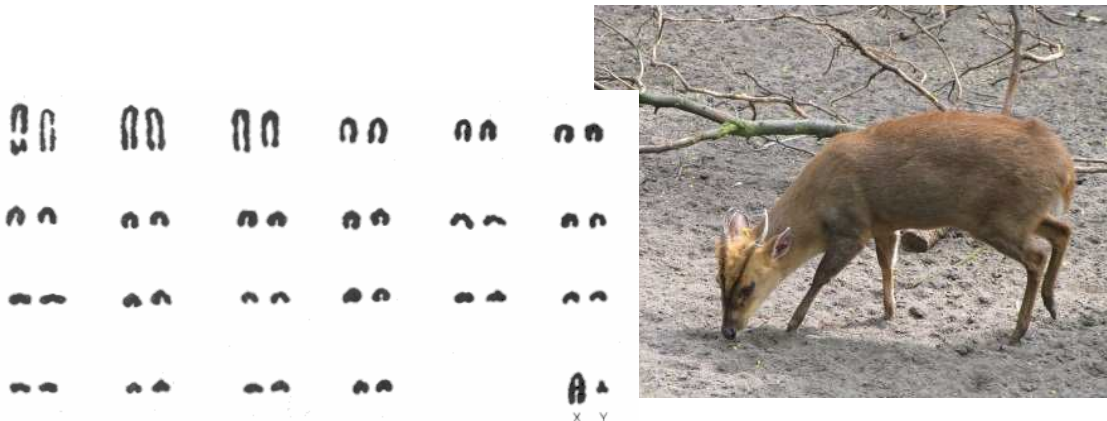
mravenec *Myrmecia pilosula* – 1 pár, samec 1 chromosom

kapradina *Ophioglossum reticulatum* – 630 párů

Muntžak – případ rozdílu v počtu chromosomů u příbuzných druhů

Fenotypově podobné druhy s velkým rozdílem v počtu chromosomů

- *Muntiacus reevesi*: $2n=46$
- *M. feae*: samice $2n=13$, samec $2n=14$
- *M. crinifrons*, *M. gongshanensis*: samice $2n=8$, samec $2n=9$
- *Muntiacus muntjak vaginalis*: samice $2n=6$, samec $2n=7$



Chromosomy muntžaka malého
(*Muntiacus reevesi*, $2n=46$)

Mnohočetné fúze
chromosomů



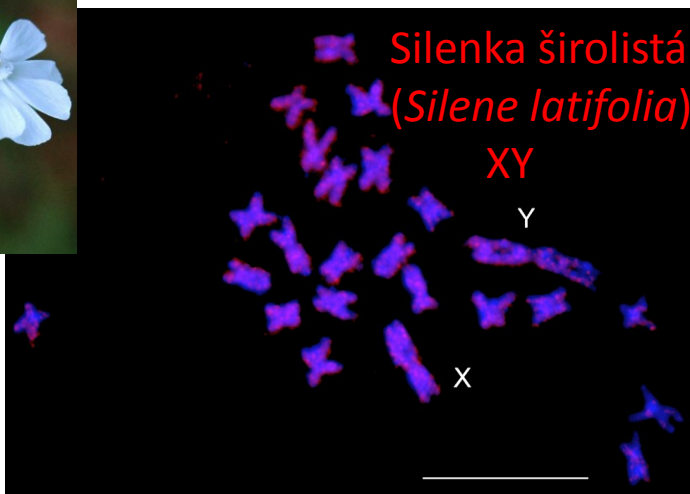
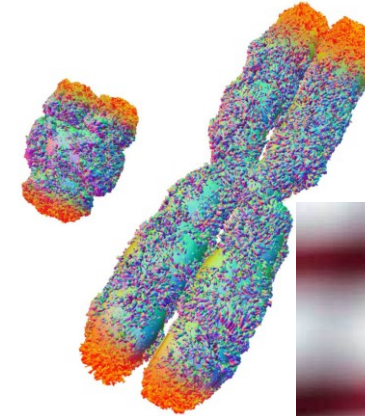
Chromosomy muntžaka červeného
(*Muntiacus muntjak vaginalis*, samice $2n=6$, samec $2n=7$).



Pohlavní chromosomy

Pohlavní chromosomy

- přítomné jen u jednoho pohlaví (XX/XY; ZZ/ZW)
- nebo u obou v různém počtu (X/Z)
- rychlá evoluce (adaptace, speciace)
- běžné u živočichů, vzácné u rostlin



Silenka širolistá
(*Silene latifolia*)

XY

Y

X



Šťovík (*Rumex acetosa*) XY1Y2



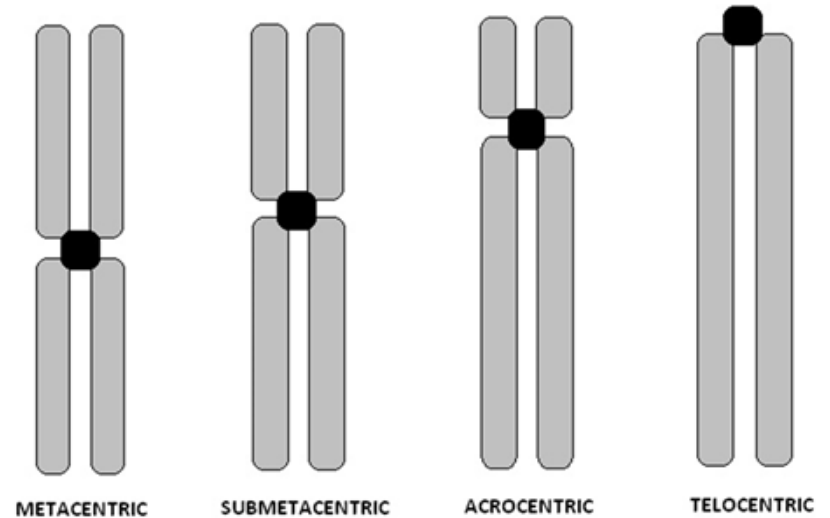
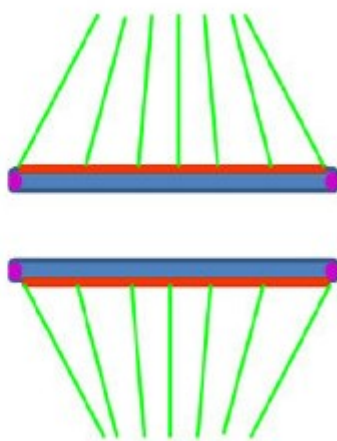
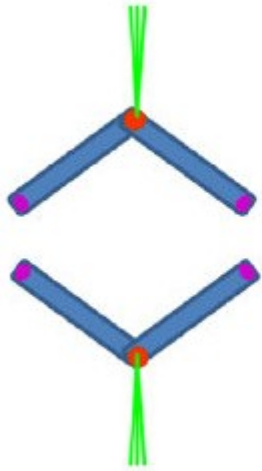
Typ a tvar chromozomů

Monocentrické

x

Holocentrické

Monocentrické



Obratlovci, většina bezobratlých a rostlin

Asi ¼ bezobratlých, 3% rostlinných druhů

Více crossing-overů -> vyšší variabilita potomků

Málo crossing-overů -> pomalejší adaptace

Citlivé na zlomy chromozomů -> zastavení v G2 -> letální

Vysoká odolnost proti UV, gama, kosmickému záření

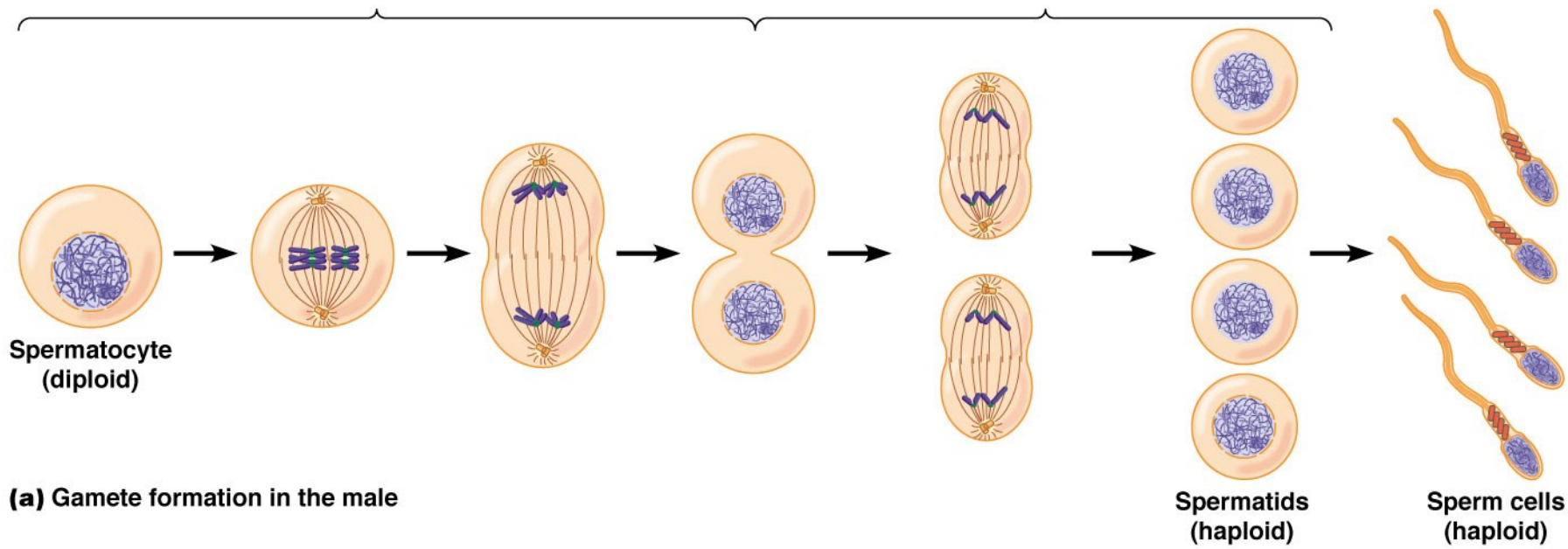
V karyotypu druhu obvykle převažují buď metacentrici nebo akrocentrici – Proč?

Tvar chromozomů zřejmě nepřináší evoluční výhody/nevýhody

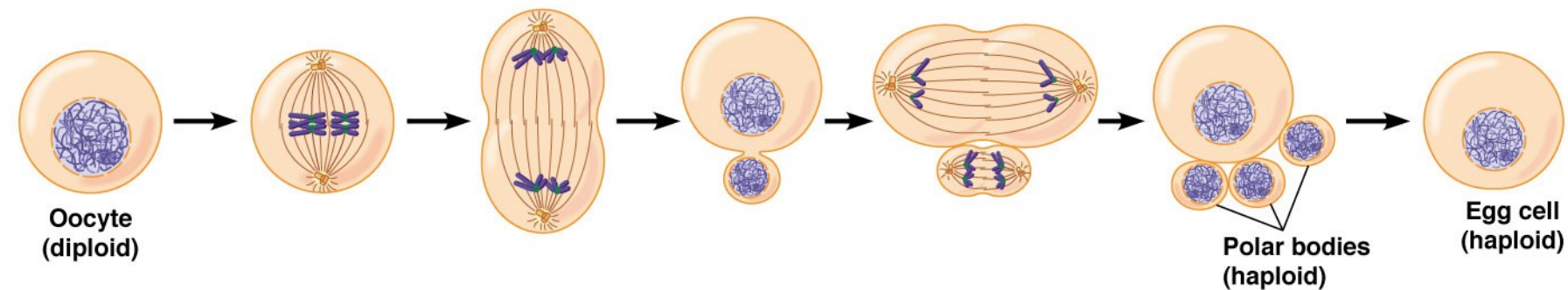
Evolve karyotypu a meiotický tah

Meiosis I

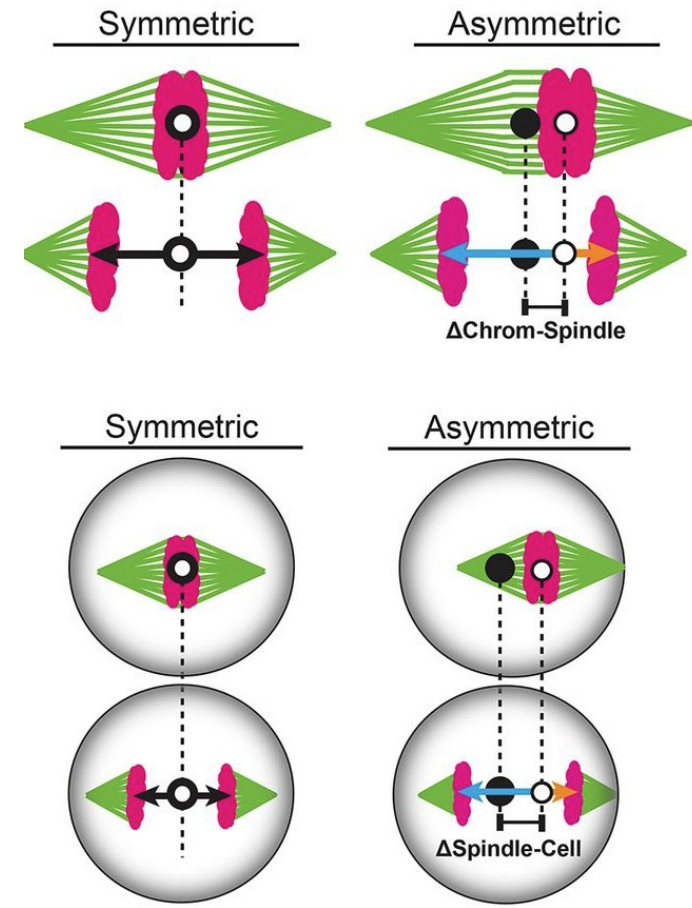
Meiosis II



(a) Gamete formation in the male

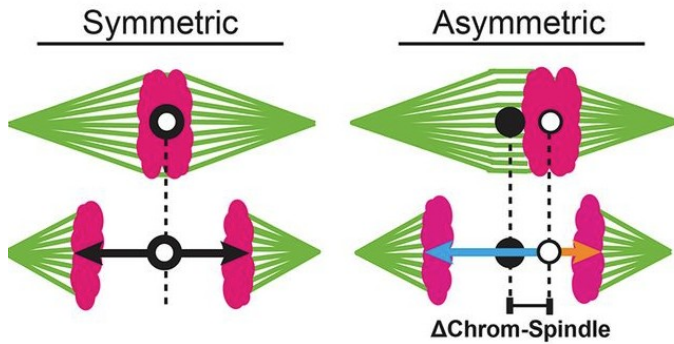


(b) Gamete formation in the female



Evolution karyotypu a meiotický tah

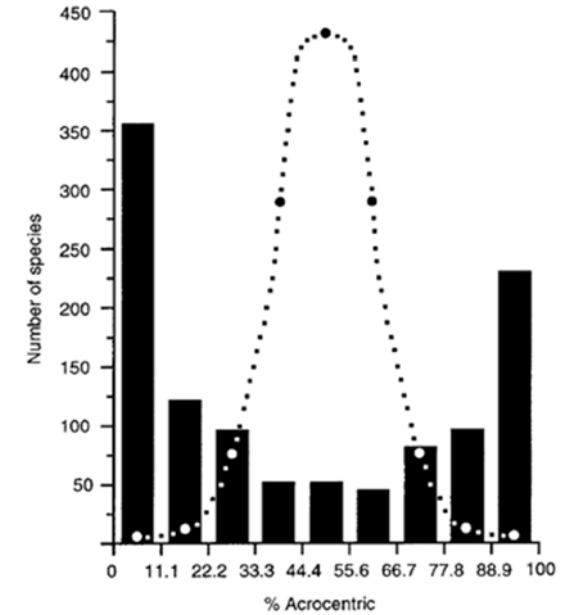
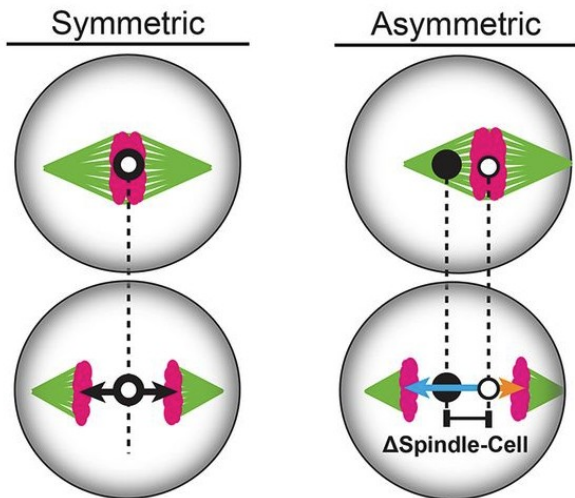
Asymetrické dělicí vřeténko hraje roli jen v samičí meióze



Asymetrické vřeténko -> přednostně kratší/delší chromozomy do vajíčka

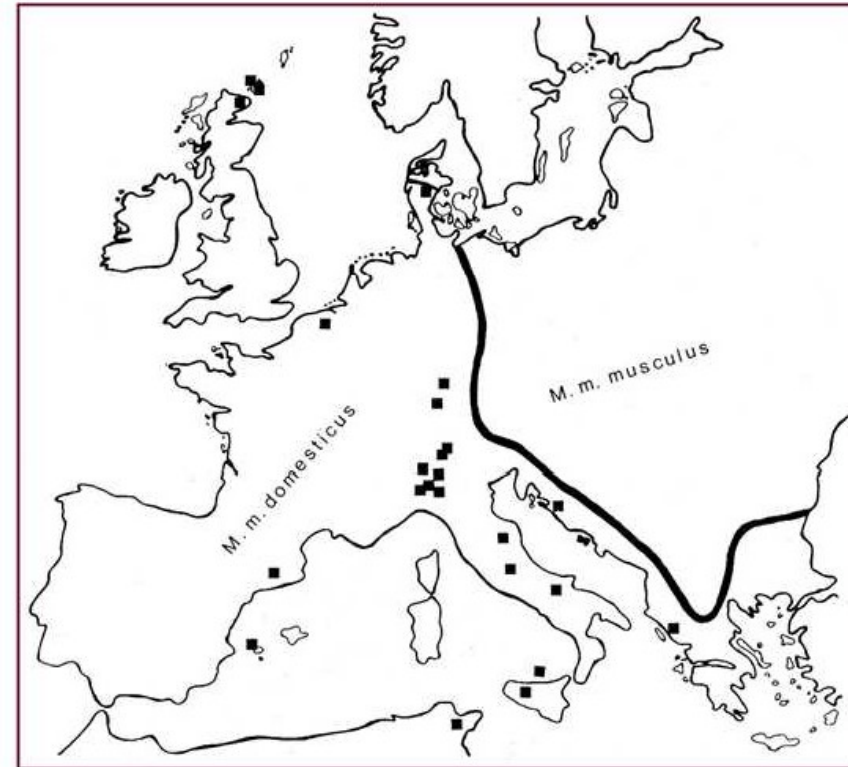
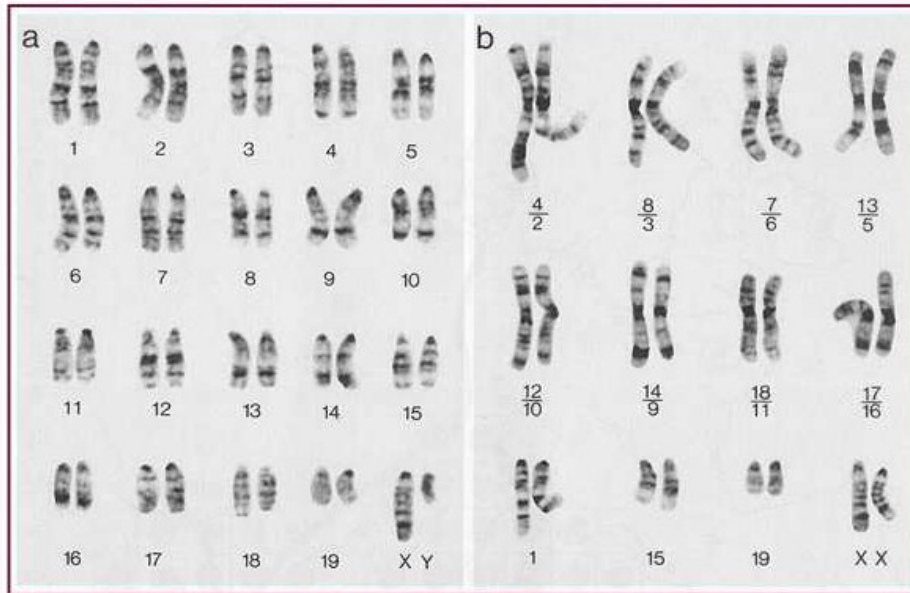
-> Karyotyp savců je převážně tvořen z meta- nebo akrocentriků

Možnost změny orientace vřeténka Prokázána u myši, dvě subpopulace, cca 500 let



	Females		Males	
Mouse	3283 298.10	4839 $P < 10^{-30}$	3494 0.17	3529 not significant
Human	246 12.72	173 $P < 0.0005$	126 0.14	132 not significant
Chicken	264 54.90	119 $P < 10^{-10}$	210 0.30	199 not significant
Centromeres	1	2	1	2

Evolution karyotypu a meiotický tah + Robertsonské translokace



Myš domácí – 2 poddruhy

Mus musculus musculus – původní karyotyp $2n=40$ akrocentrických chromozomů

Mus musculus domesticus – fúze chromozomů $\rightarrow 2n=22-38$

Inverze chromozomů -> speciace

Drosophila pseudoobscura a *D. persimilis*

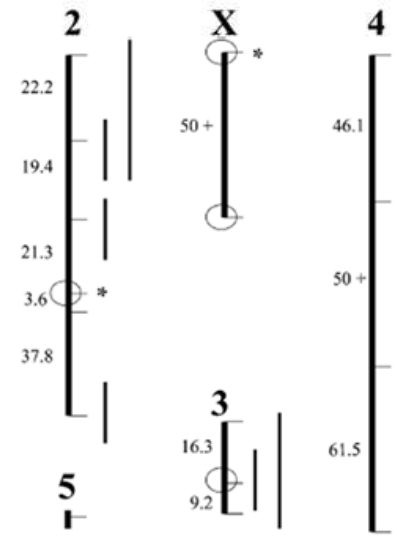
- odděleny před cca 500 tisíci let
- Liší se 4 inverzemi
- V přírodě tvoří hybridy X málo (tvorba prezygotické izolace – rozdílná vibrace křídel při námluvách)
- Samci v F1 jsou sterilní
- Zkoumána genetická podstata inviability/sterility hybridních samců, dysfunkce ve dvoření se samicím, úspěch v páření a samičí preference.
- Všechny znaky zamapovány do inverzí na chromosomech X a 2.



D. pseudoobscura



D. persimilis

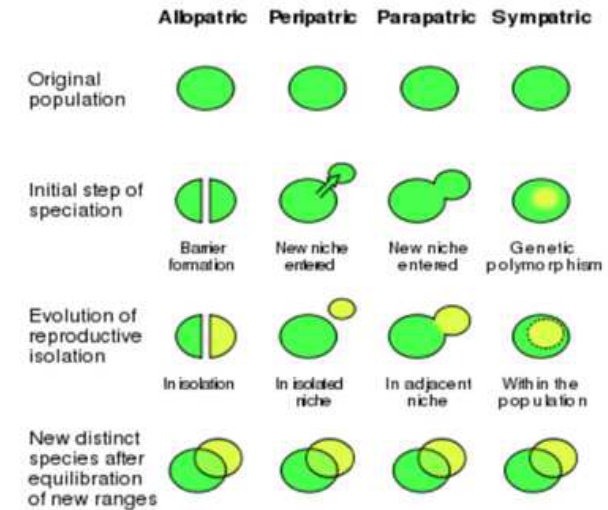


Inverze chromozomů u drosophil

Výskyt inverzí u sympatrických druhů drozofil vs. absence u alopatrických druhů → role v izolaci sympatrických druhů?

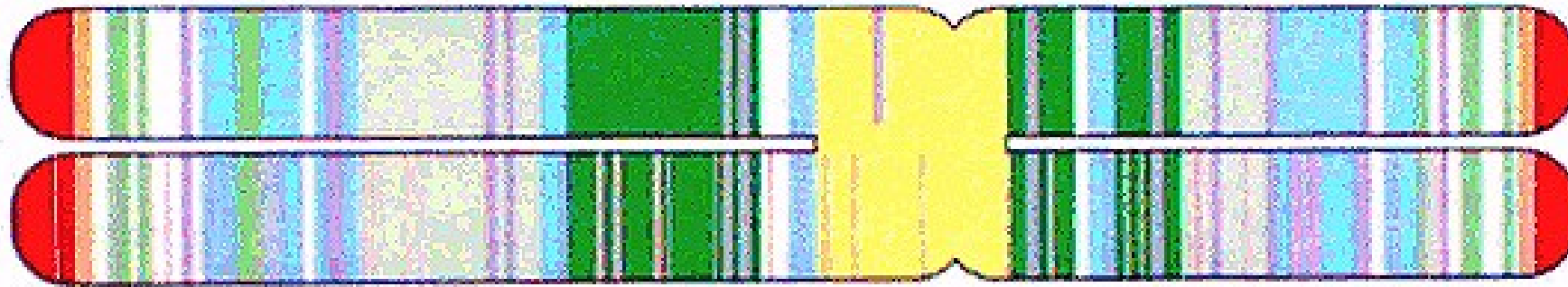
Table 2. Relative abundance of homosequential *Drosophila* taxa vs. those bearing inversion differences compared by sympatry (sym)/allopatry (allo) and genetic distance

		Patry*	D [†]	Sterility [‡]
<i>Drosophila</i> species/subspecies not differing by one or more inversions				
<i>heteroneura</i>	<i>planitibia</i>	allo	0.134	0.500
<i>differens</i>	<i>planitibia</i>	allo	0.138	0.500
<i>planitibia</i>	<i>silvestris</i>	allo	0.191	0.500
<i>pseudoobscura</i> Bogota	<i>pseudoobscura</i> USA	allo	0.194	0.250
<i>sechellia</i>	<i>simulans</i>	allo	0.280	0.500
<i>simulans</i>	<i>sechellia</i>	allo	0.280	0.500
<i>mauritiana</i>	<i>simulans</i>	allo	0.300	0.500
<i>mauritiana</i>	<i>sechellia</i>	allo	0.320	0.500
<i>"aldrichi"[§]</i>	<i>mulleri</i>	sym	1.051	1.000
<i>Drosophila</i> species differing by one or more inversions				
<i>flavomontana</i>	<i>laticola</i>	allo	0.180	0.500
<i>borealis</i>	<i>montana</i>	sym	0.210	0.500
<i>arizonae</i>	<i>mojavensis baja</i>	sym	0.212	0.250
<i>flavomontana</i>	<i>montana</i>	sym	0.290	0.500
<i>borealis</i>	<i>flavomontana</i>	sym	0.380	1.000
<i>persimilis</i>	<i>pseudoobscura</i>	sym	0.410	0.500

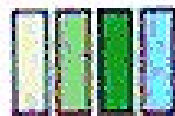


TOPOGRAFIE GENOMU

Uspořádání genů v genomech



KEY



Tandem repeats



Centromere-associated tandem repeats



Telomeric and sub-telomeric repeats



Dispersed *Ty1/copia* retroelements and microsatellites

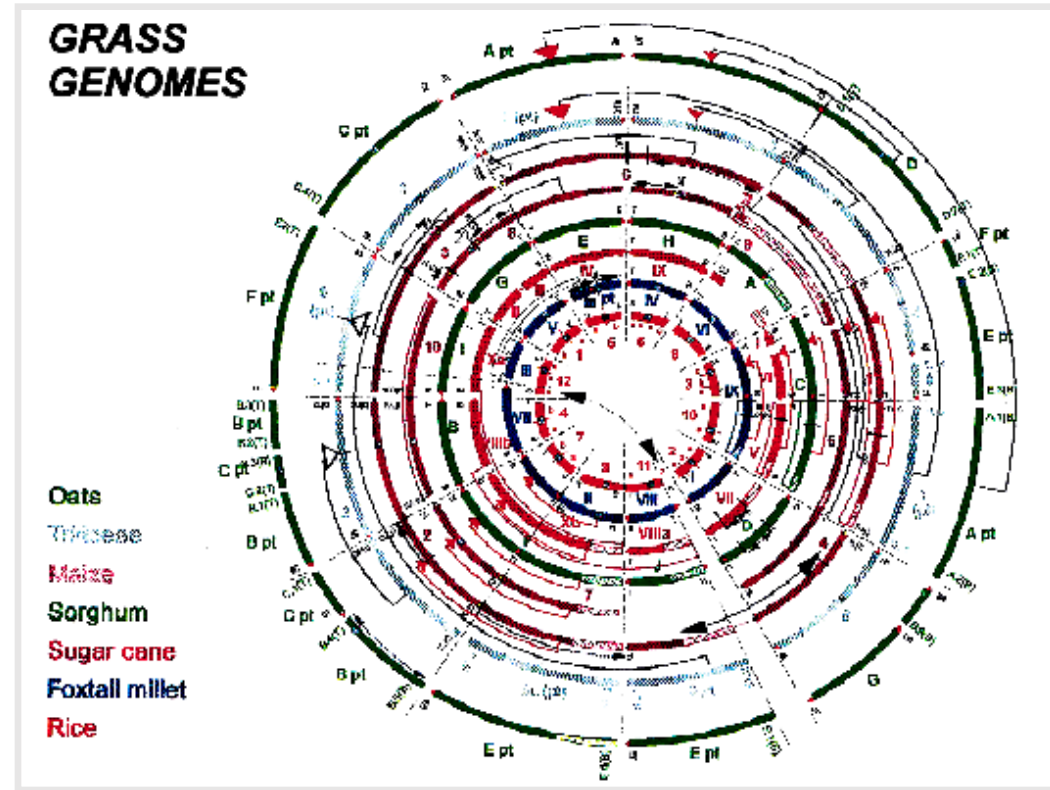
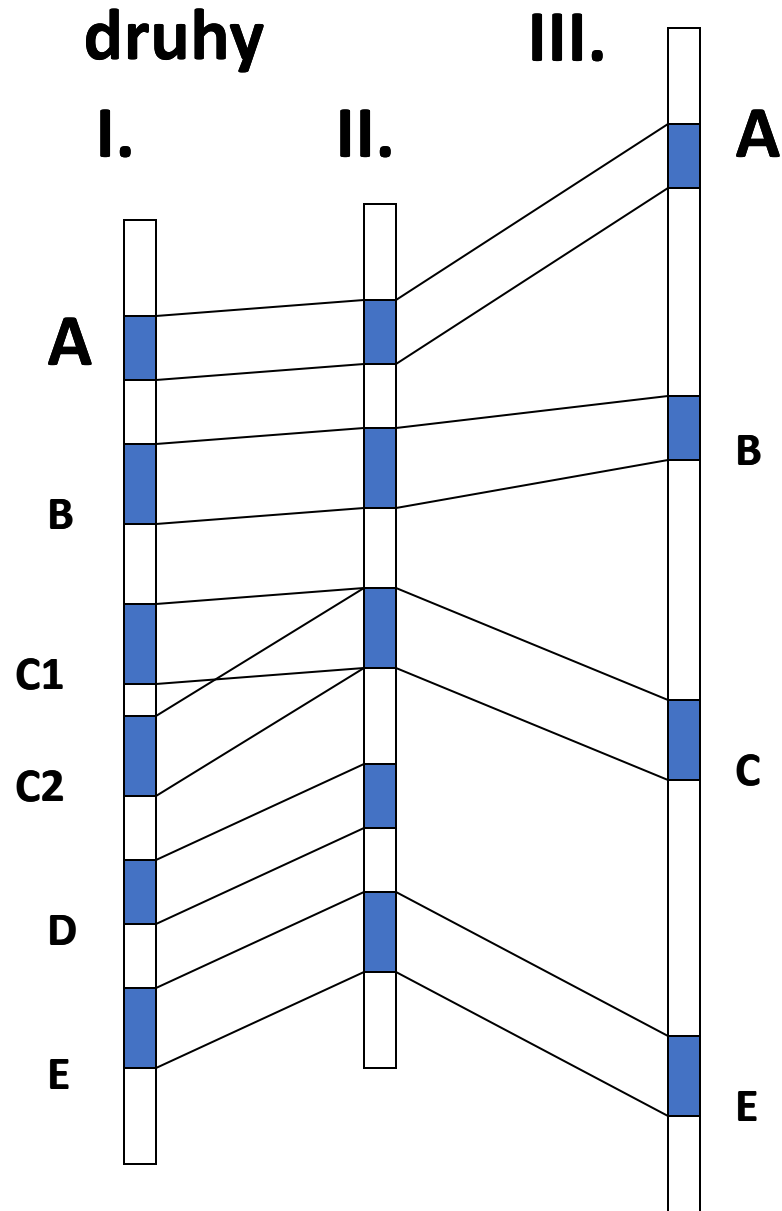


LINEs



Single and low-copy sequences including genes

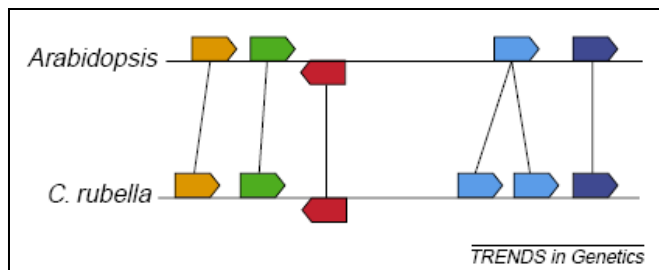
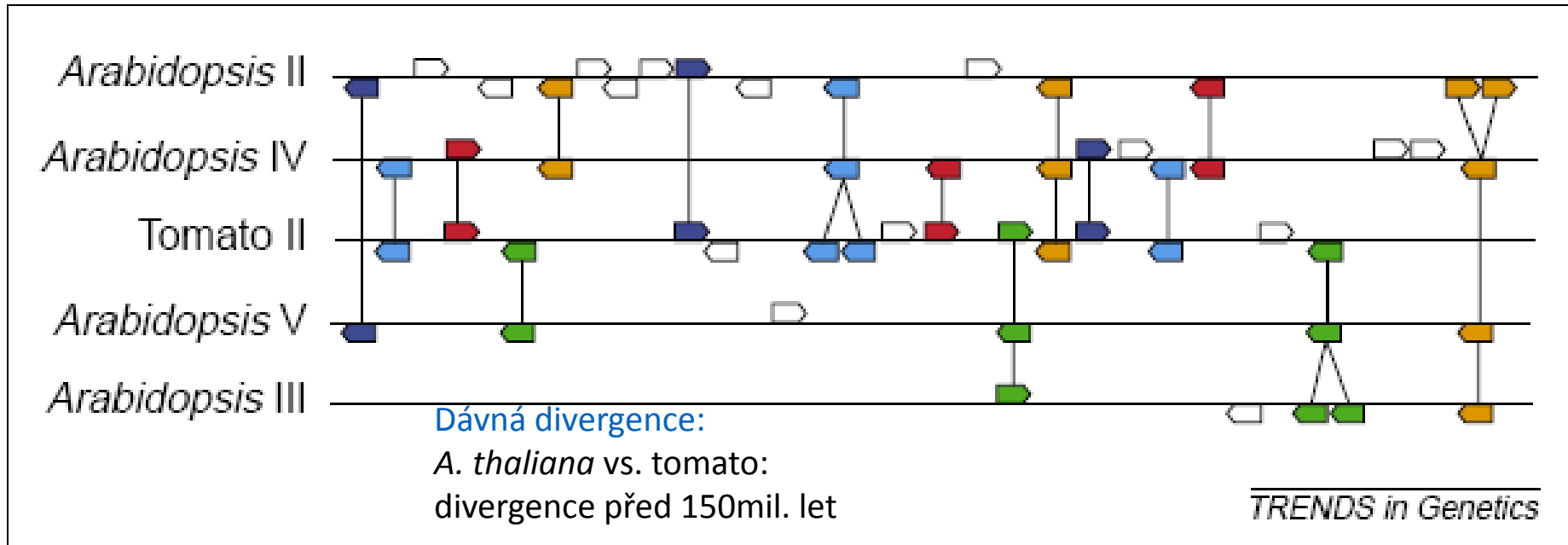
Syntenie - konzervativita pořadí genů



- konzervativita genů
- rozdílné mezigenové sekvence

Makro- vs. mikro- kolinearita

- konzervativita v dlouhých úsecích
- mikrostruktura více dynamická, mix s jinými geny
- polyploidie urychlí evoluci mikrostruktury
- změny již u prvních generací syntetického alotetraploida u Brassica



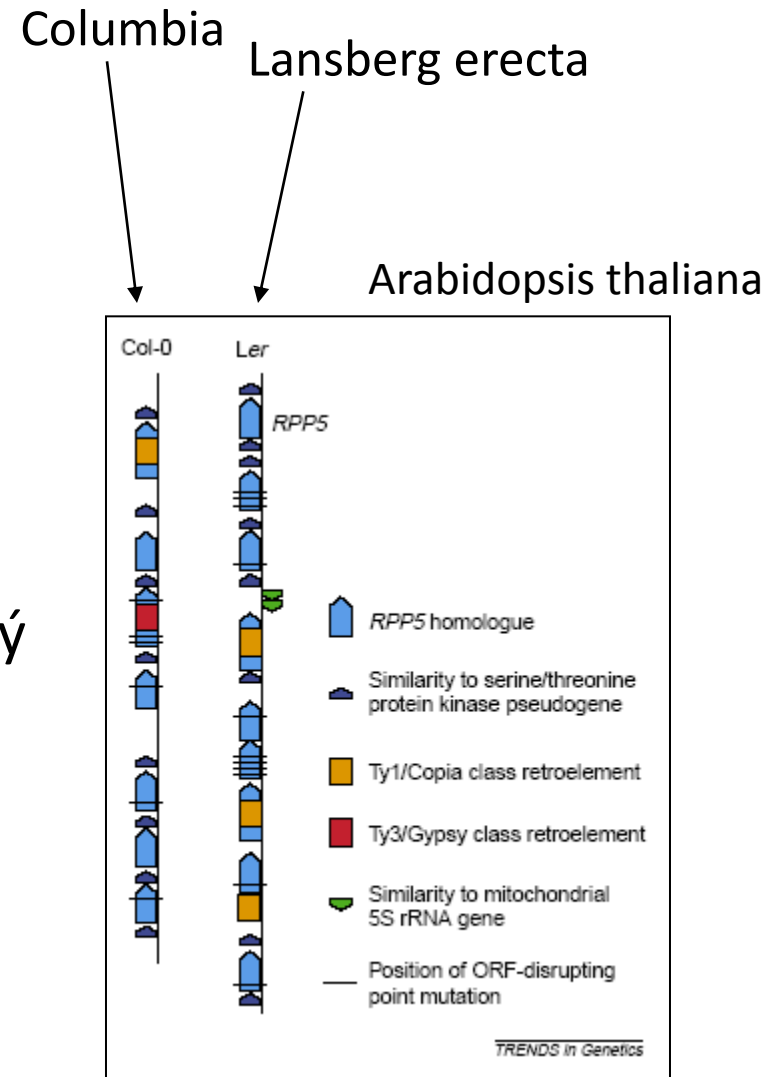
Nedávná divergence:
A. thaliana vs. *Capsella rubella*:
divergence před 6.2-9.8 mil. let

(Bancroft, 2001)

Rychlá evoluce oblasti rezistence k patogenu

(Bancroft, 2001)

- gen „RPP5“ – rezistence k *Perenospora parasitica*
- rozdíly mezi A.t. **ekotypy** Lansberg erecta a Columbia
- většinou pseudogeny
- tandemové duplikace a nerovnoměrný crossing-over je mechanismem rychlé evoluce



Izochorový model organizace genomu

Isochory: bloky genomové DNA (stovky kb-Mb) s charakteristickým GC-obsahem, ostrý přechod (*G. Bernardi*)

- shlukování do skupin s diskretními GC-obsahy, u člověka typy L1, L2, H1, H2 a H3 (GC nejbohatší, tvoří jen 3% ale obsahuje 25% genů)
- Mozaiková struktura genomu obratovců a rostlin



Původ izochor:

1. Výsledek mutací?
2. Výsledek selekce?
3. Genová konverze?

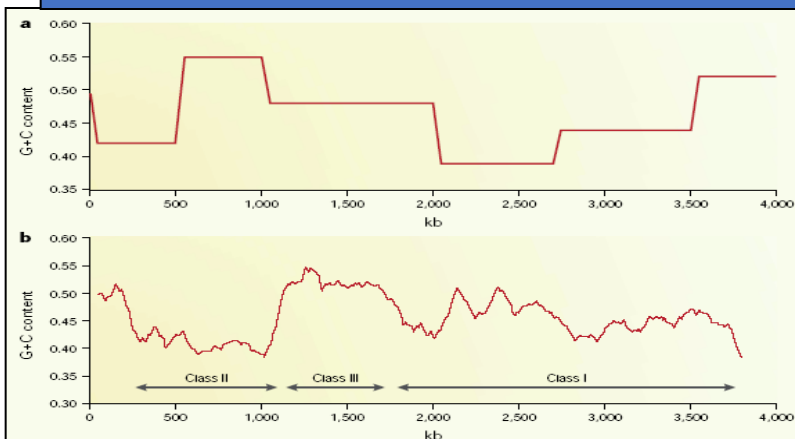
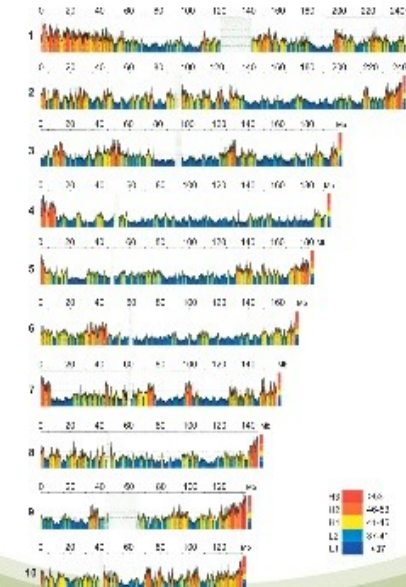
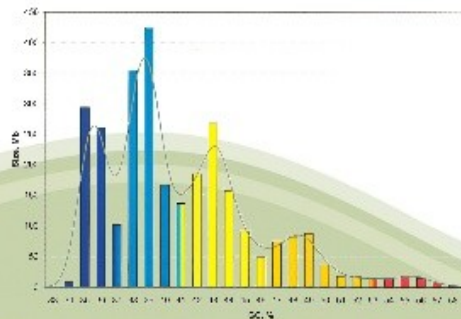


Figure 1 | Large-scale variation in G+C content. a | The classic isochore model. b | G+C content across the three classes of human major histocompatibility (MHC) region of chromosome 6 (data from GenBank); G+C content is plotted as a moving average, and the window size is 100 kb, advanced by 10 kb each step.

Finding isochores

● Isochores: homogeneous regions of %GC content

- Easy to find with windowed (100kbp) %GC calculation, from sequenced genomes.
- 3200 isochores characterised in the human genome, consistent with 5 levels (L1, L2, H1, H2, H3) found by staining/hybridisation.



Shrnutí

- Endosymbiotický původ eukaryotických organismů a organel s vlastními genomy
- Organismy mají různě velké genomy, neodpovídá komplexitě těl → paradox C
- Mechanismy změn velikosti genomu: polyploidizace, indely, expanze repeticí a TE → vše významné v evoluci genomu a speciaci
- Důsledky zvětšeného genomu: velké buňky, zjednodušení orgánů, pomalejší metabolismus, pomalé dělení buněk → selekce na malé genomy u ptáků a letounů
- Evoluce karyotypu – velikost, tvar a počet chromosomů
- Strukturní aberace chromosomů – inverze – role ve speciaci