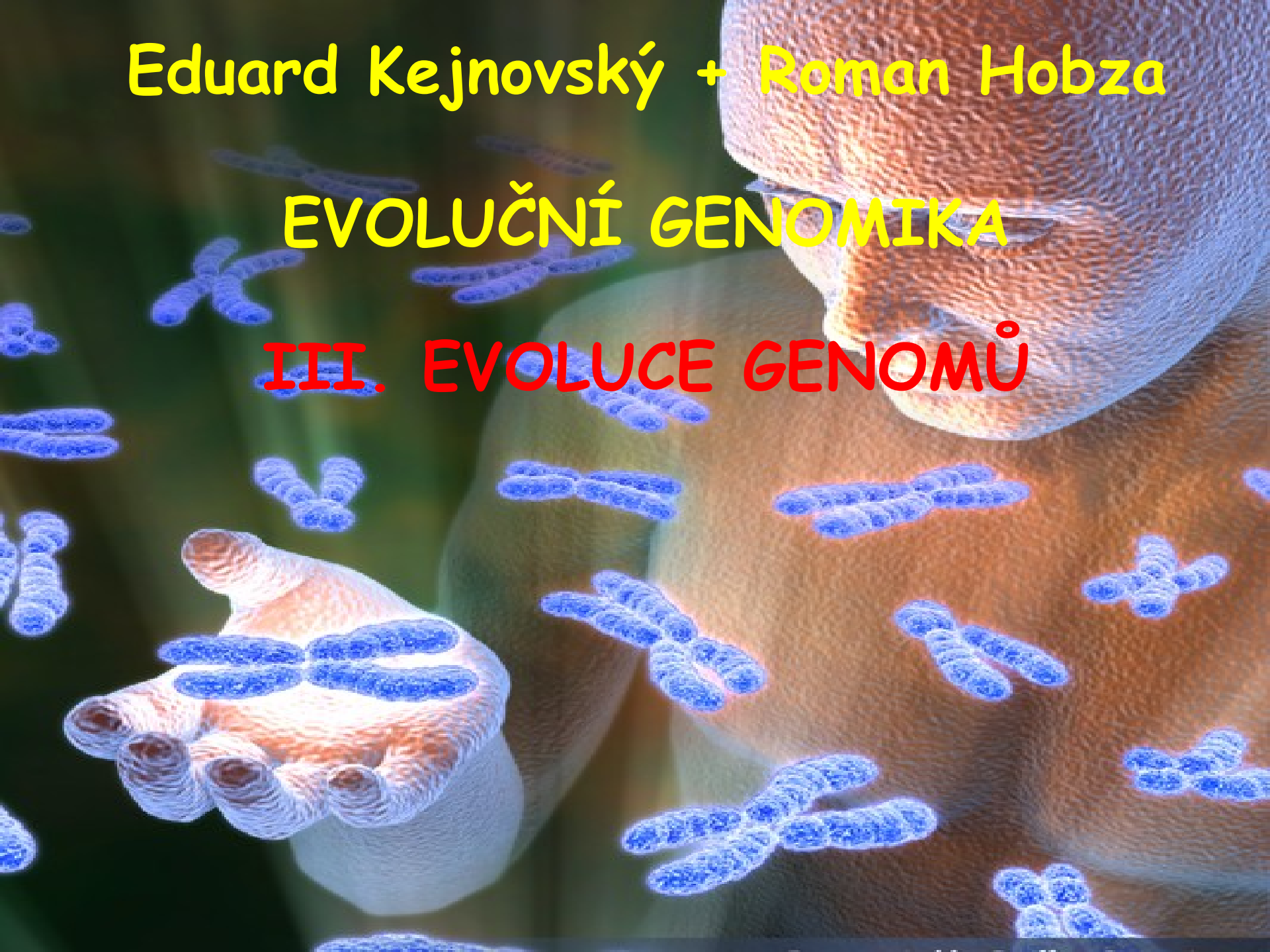


Eduard Kejnovský + Roman Hobza

EVOLUČNÍ GENOMIKA

III. EVOLUCE GENOMŮ



OBSAH

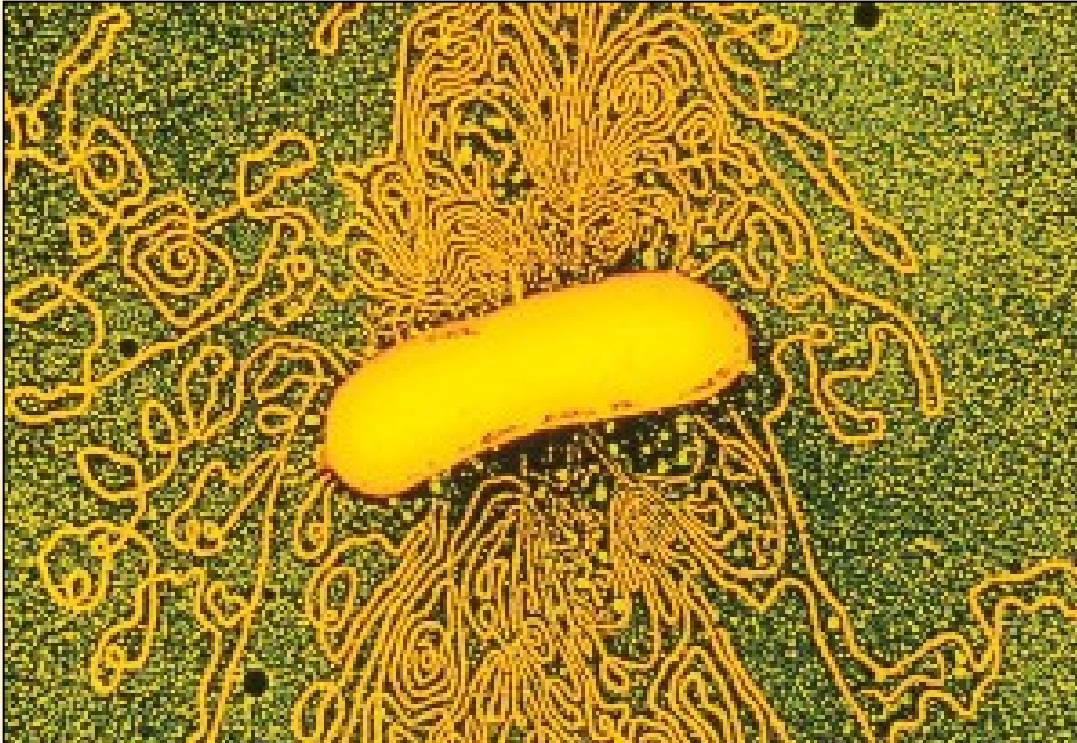
1. Velikost genomu
2. Extrémní genomy
3. Změny ve velikosti genomu
4. Nekódující DNA a velikost genomu
5. Topografie genomu
6. Genomové projekty

VELIKOST GENOMU

Genomy bakterií a virů

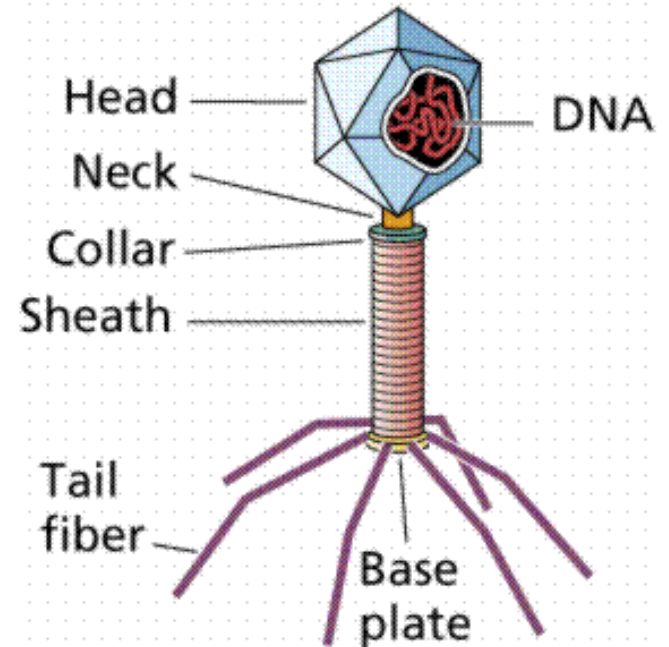
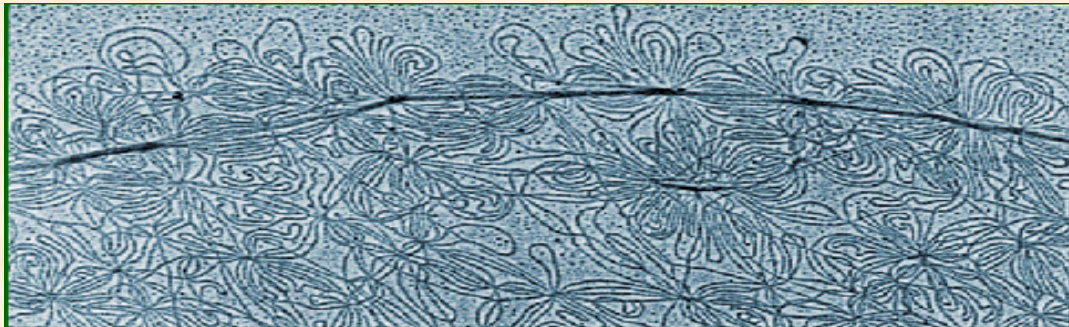
- genomy malé, kompaktní
- efektivní využití

Virus (bakteriofág)
~ 50 kb ~ stovky genů

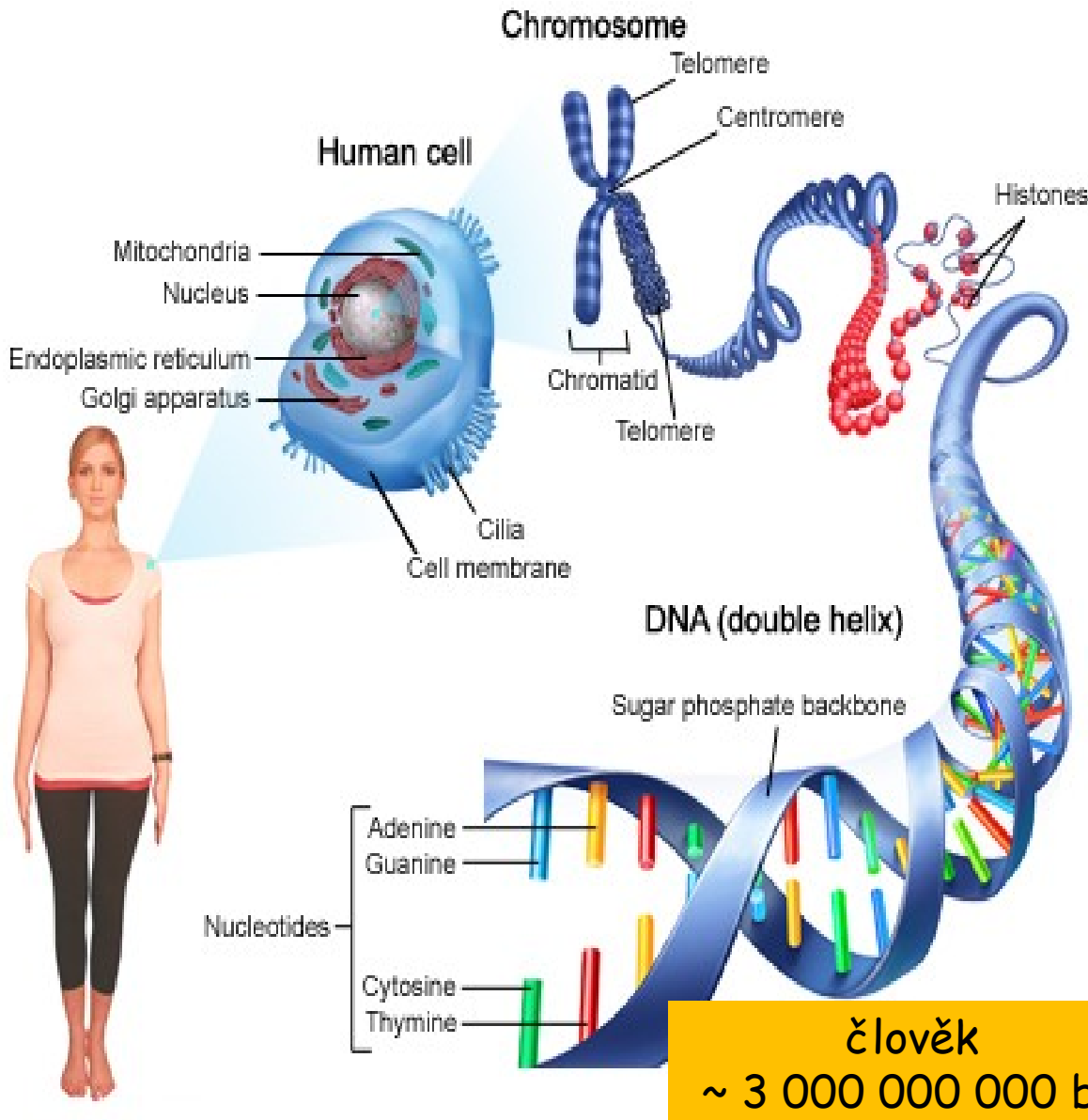


Bakterie:

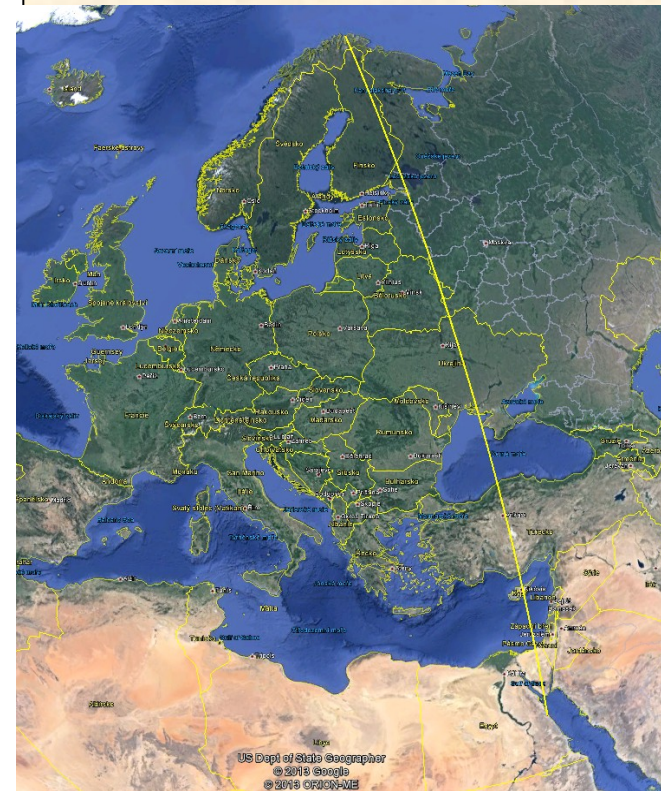
~ 0.5 - 9 mil bp ~ tisíce genů



Vyšší organizmy: kde se nachází DNA?



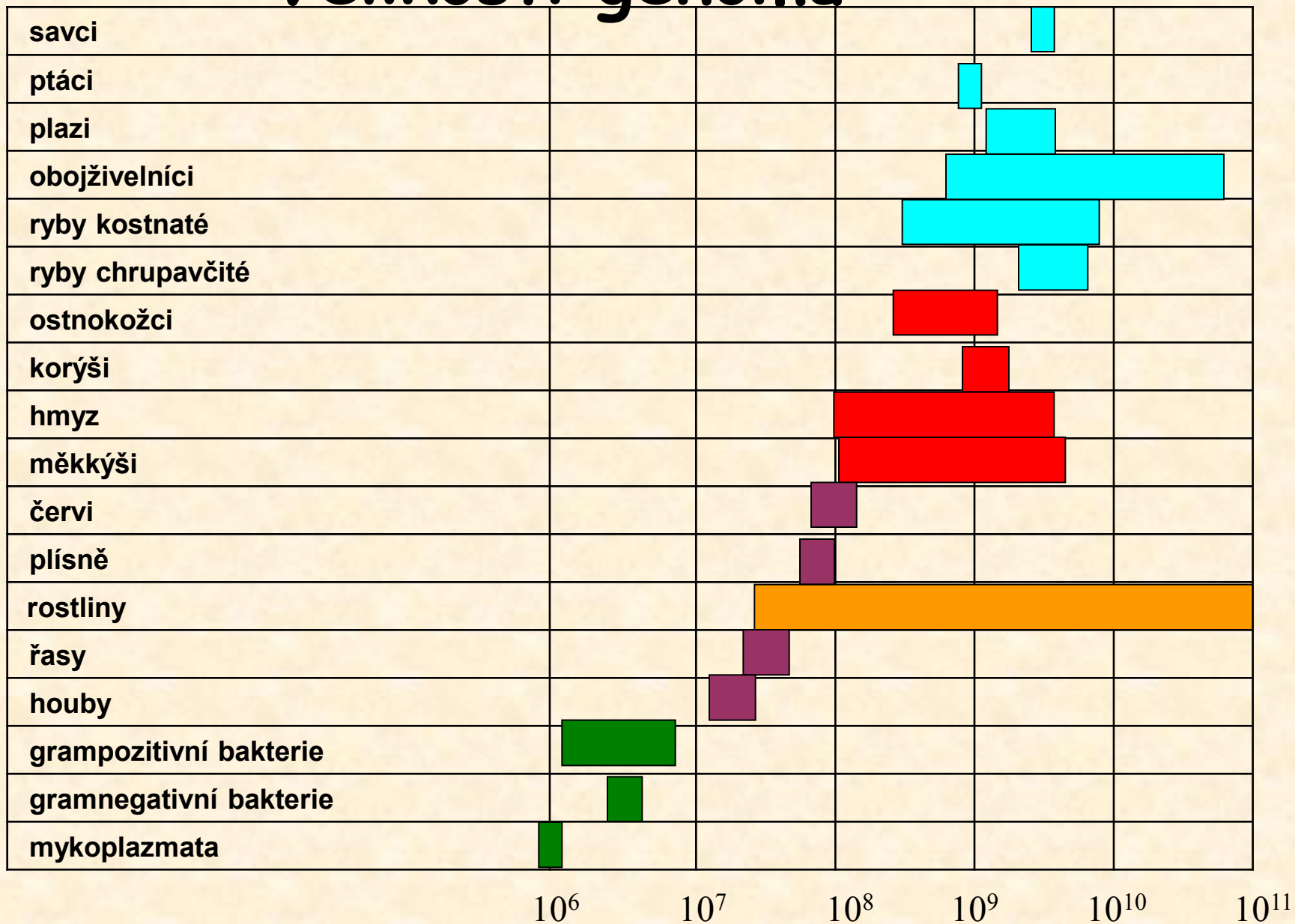
Genom = celková genetická informace buňky nebo viru



člověk
~ 3 000 000 000 bp

2nm x 1m
1cm -> 5000km

Velikosti genomů



Velikosti genomů a paradox hodnoty C



180 Mb

Drosophila melanogaster

100x

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



18,000 Mb

Podisma pedestris

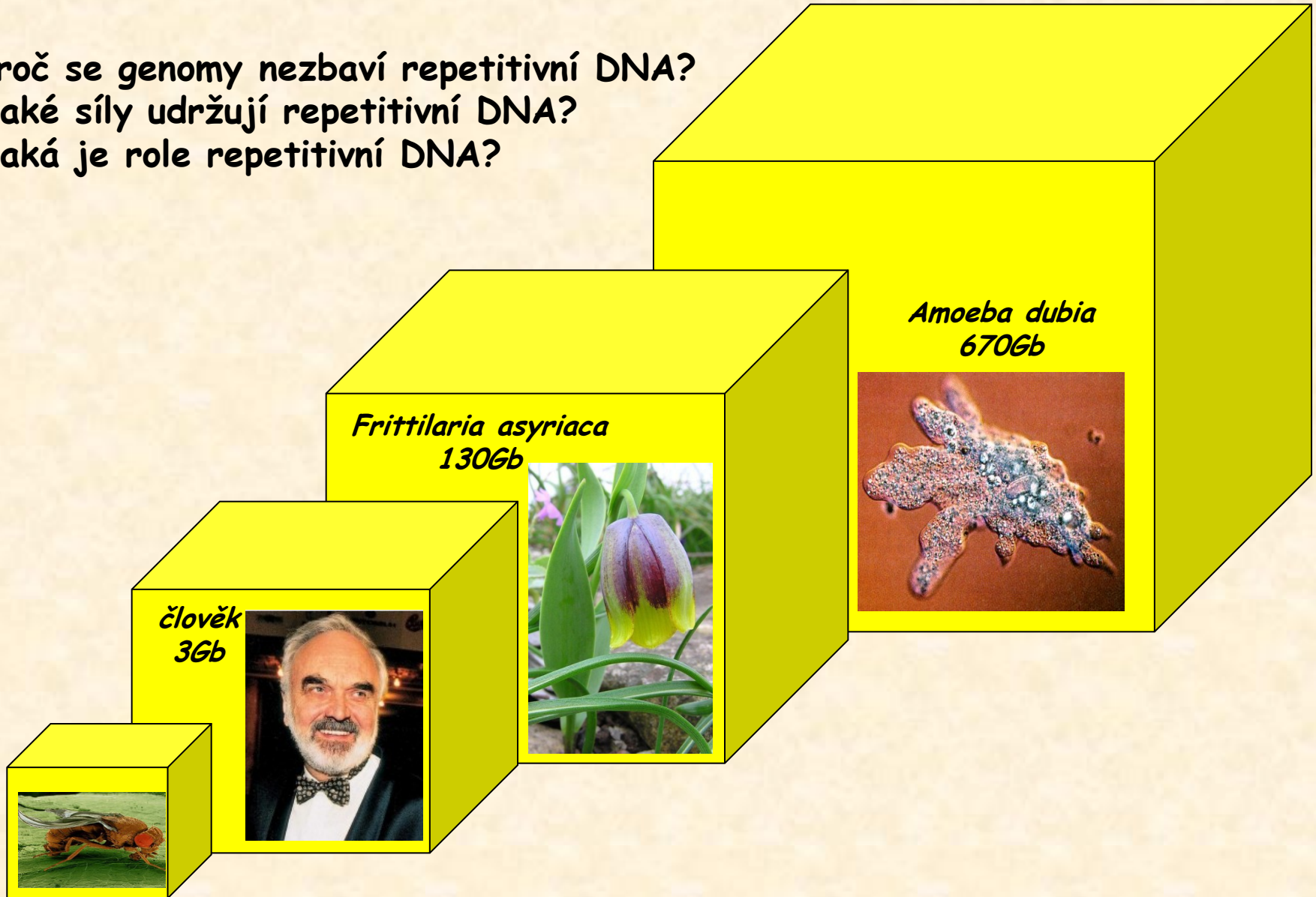
GENOM = celková genetická informace organismu (geny i nengenové oblasti/repetice)

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

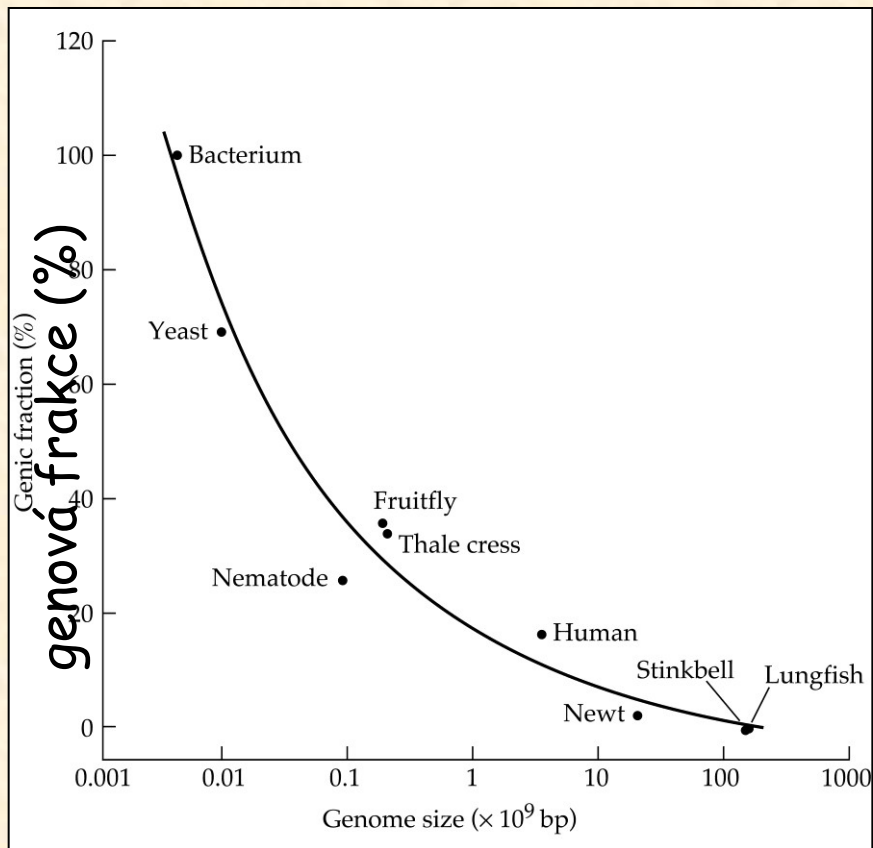
Genomy jsou repetitivní a dynamické

Velikosti genomů a paradox hodnoty C

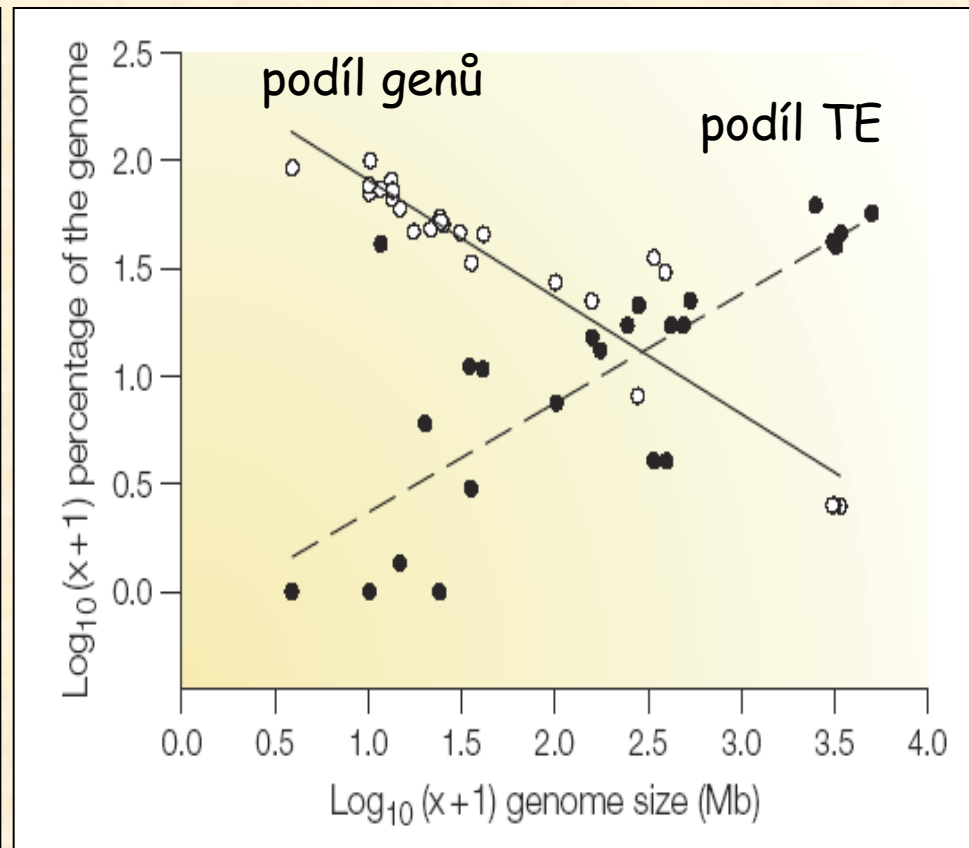
Proč se genomy nezbaví repetitivní DNA?
Jaké síly udržují repetitivní DNA?
Jaká je role repetitivní DNA?



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:

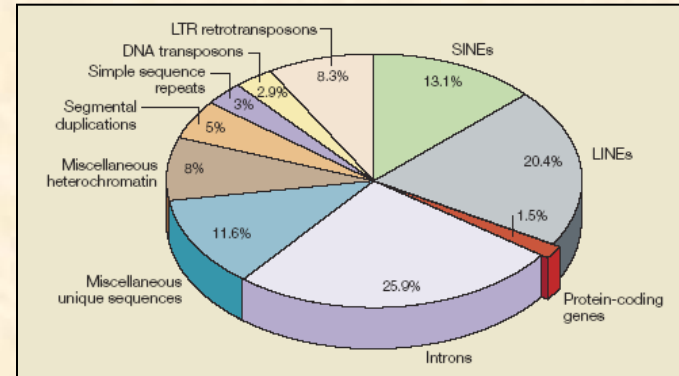
- klasické, retropseudogeny,
- 19000 člověk, 14000 myš, 51 kur, 33 kvasinka, 176 drosophila

Mobilní elementy:

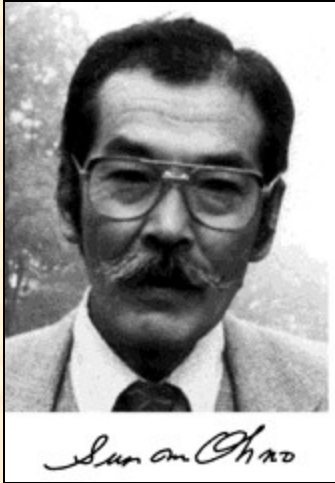
- LTR, nonLTR - SINE, LINE, DNA transposony-MITE

Numt, Nupt:

- inserce promiskuitní DNA



Evolve genomů - citáty



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

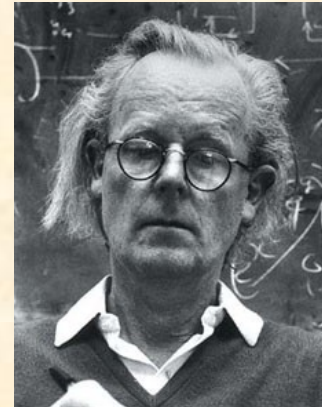
- *Susumu Ohno*

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

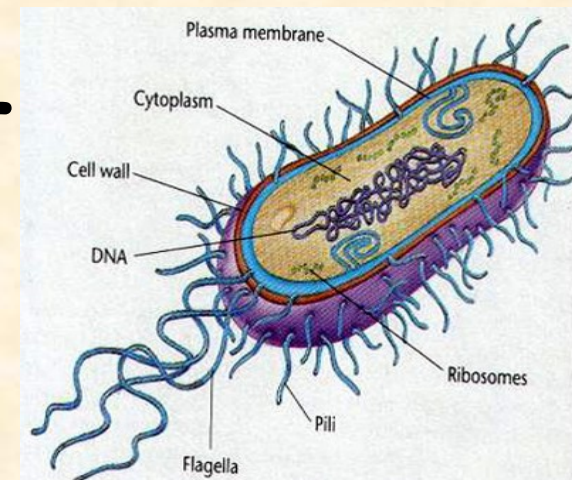
- *Susumu Ohno*

„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*



Velikosti genomů prokaryot

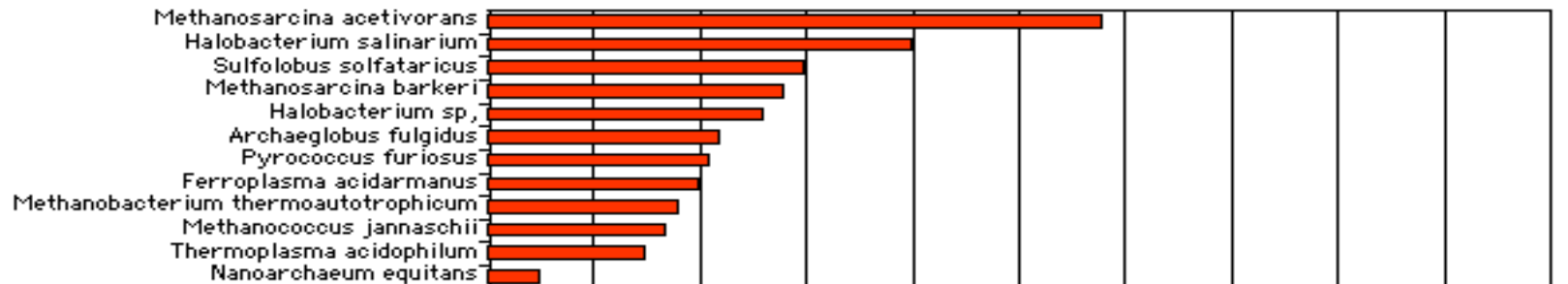


- **Bakteriální genom: $6 \times 10^5 - > 10^7$ (=0.6-10Mb)**
 - nejmenší známý: *Mycoplasma genitalium*
(480 genů kódujících proteiny, 3 rRNA geny, 37 tRNA genů)
- **Procesy ovlivňující velikost bakteriálního genomu:**
 - genová duplikace, malé delece a inserce, transpozice, horizontální přenos, ztráta genů v parazitických liniích, atd.

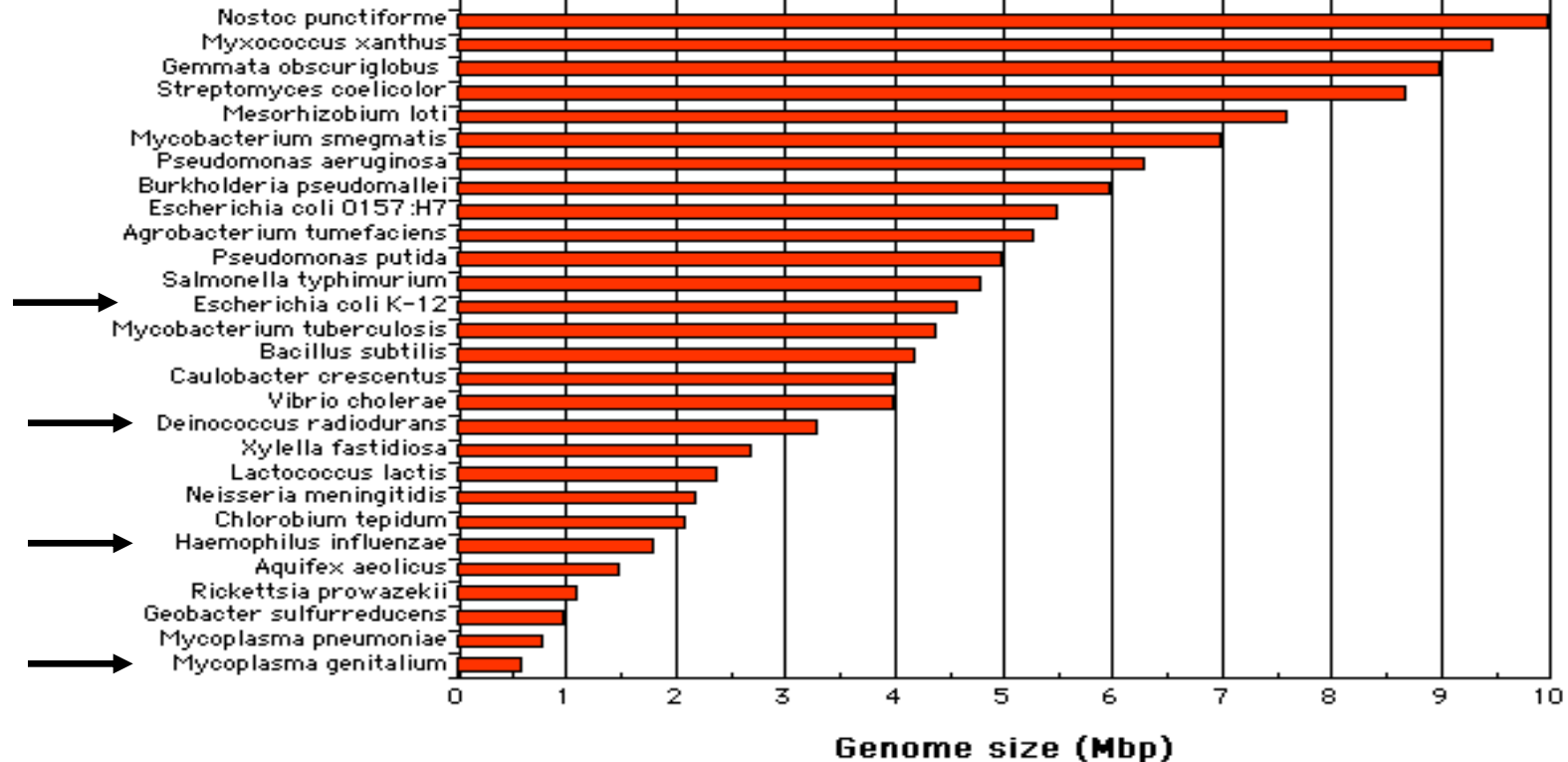
**Velikost genomů
prokaryot je
zhruba úměrná
počtu genů**

Velikosti genomů prokaryot

Archaea:



Bacteria:




EXTRÉMNÍ GENOMY

„Životní minimum“ bakterie

Život s pouhými 271 geny

Strategie:

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



TIGR
THE INSTITUTE FOR GENOMIC RESEARCH

Minimal Genome Project

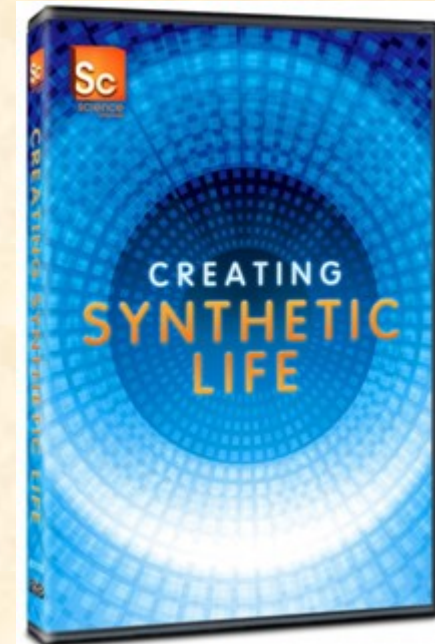
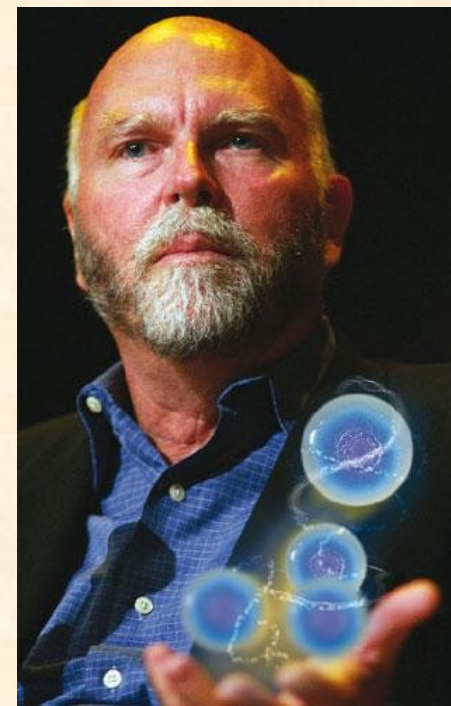
Press Release

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)

TIGR Databases
What's New
About TIGR
TIGR Faculty



- Mycoplasma laboratorium* - Ústav Craiga Ventera (JCVI)
- nahradili genom *M. capricolum* uměle nasyntetizovanou DNA *M. mycoides*
 - nový software přebudoval hardware
 - citáty, adresy a jména vědců v umělé DNA

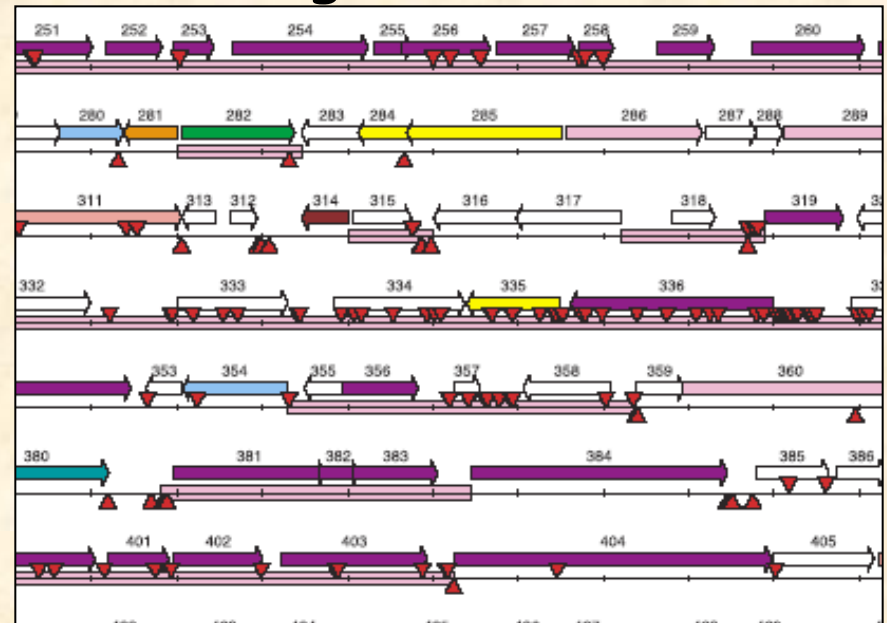
“Minimální” genom - Mycoplasma

580 kb genom/ 480 genů pro proteiny/ 37 genů pro tRNA
2209 inzercí transpozonů/ ve 140 genech
1354 míst, kde inzerce nebyla letální
265-350 genů nepostradatelných (glykolýza)
180-215 genů postradatelných
100 genů má neznámou funkci!!!
různý vliv inzercí podle polohy inzerce v genu

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.



“Minimální” genom: *Bacillus subtilis* život s 271 geny

4100 genů

192 genů zcela **nepostradatelných** na bohatém prostředí: zpracování informace, buněčná stěna, energie

79 genů **podstatných**

Jen 4% genů má **neznámou** funkci

Essential *Bacillus subtilis* genes

K. Kobayashi¹, S. D. Ehrlich^{2,5}, A. Albertini¹, G. Amati¹, K. K. Andersen⁶, M. Arnaud⁷, K. Asai⁸, S. Ashikaga⁹, S. Aymerich¹, P. Bessieres¹, F. Boland¹⁰, S. C. Brignell¹¹, S. Bron¹², K. Bunai¹³, J. Chapuis¹⁴, L. C. Christiansen¹⁵, A. Danchin¹⁶, M. Débarbouillé¹⁷, E. Dervyn¹⁸, E. Deuring¹⁹, K. Devine²⁰, S. K. Devine²¹, O. Dreesen²², J. Errington²³, S. Fillinger²⁴, S. J. Foster²⁵, Y. Fujita²⁶, A. Galizzi²⁷, R. Gardan²⁸, C. Eschevins²⁹, T. Fukushima³⁰, K. Haga³¹, C. R. Harwood³², M. Hecker³³, D. Hosoya³⁴, M. F. Hullo³⁵, H. Kakeshita³⁶, D. Karamata³⁷, Y. Kasahara³⁸, F. Kawamura³⁹, K. Koga⁴⁰, P. Koski⁴¹, R. Kuwana⁴², D. Imamura⁴³, M. Ishimaru⁴⁴, S. Ishikawa⁴⁵, I. Ishio⁴⁶, D. Le Coq⁴⁷, A. Masson⁴⁸, C. Mauel⁴⁹, R. Meima⁵⁰, R. P. Mellado⁵¹, A. Moir⁵², S. Moriya⁵³, E. Nagakawa⁵⁴, H. Nanamiya⁵⁵, S. Nakai⁵⁶, P. Nygaard⁵⁷, M. Ogura⁵⁸, T. Ohanan⁵⁹, M. O'Reilly⁶⁰, M. O'Rourke⁶¹, Z. Pragai⁶², H. M. Pooley⁶³, G. Rapoport⁶⁴, J. P. Rawlins⁶⁵, L. A. Rivas⁶⁶, C. Rivolta⁶⁷, A. Sadaie⁶⁸, Y. Sadaie⁶⁹, M. Sarvas⁷⁰, T. Sato⁷¹, H. H. Saxild⁷², E. Scanlan⁷³, W. Schumann⁷⁴, J. F. M. L. Seegers⁷⁵, J. Sekiguchi⁷⁶, A. Sekowska⁷⁷, S. J. Séror⁷⁸, M. Simon⁷⁹, P. Stragier⁸⁰, R. Studer⁸¹, H. Takamatsu⁸², T. Tanaka⁸³, M. Takeuchi⁸⁴, H. B. Thomaides⁸⁵, V. Wagner⁸⁶, J. M. van Dijl⁸⁷, K. Watabe⁸⁸, A. Wipat⁸⁹, H. Yamamoto⁹⁰, M. Yamamoto⁹¹, Y. Yamamoto⁹², K. Yamane⁹³, K. Yata⁹⁴, K. Yoshida⁹⁵, H. Yoshikawa⁹⁶, U. Zuber⁹⁷, and N. Ogasawara⁹⁸

Communicated by Richard M. Losick, Harvard University, Cambridge, MA, January 27, 2003 (received for review November 10, 2002)

To estimate the minimal gene set required to sustain bacterial life in nutritive conditions, we carried out a systematic inactivation of *Bacillus subtilis* genes. Among ~4,100 genes of the organism, only 192 were shown to be indispensable by this or previous work. Another 79 genes were predicted to be essential. The vast majority of essential genes were categorized in relatively few domains of cell metabolism, with about half involved in information processing, one-fifth involved in the synthesis of cell envelope and the determination of cell shape and division, and one-tenth related to cell energetics. Only 4% of essential genes encode unknown functions. Most essential genes are present throughout a wide range of Bacteria, and almost 70% can also be found in Archaea and Eucarya. However, essential genes related to cell envelope, shape, division, and respiration tend to be lost from bacteria with small genomes. Unexpectedly, most genes involved in the Embden-Meyerhof-Parnas pathway are essential. Identification of unknown and unexpected essential genes opens research avenues to better understanding of processes that sustain bacterial life.

M. genitalium and *Haemophilus influenzae*, led to a description of a smaller set of some 260 genes (2). More recently, an experimental approach involving high-density transposon mutagenesis of the *H. influenzae* genome led to a much higher estimate of ~670 putative essential genes (3), whereas transposon mutagenesis of two mycoplasma species led to an estimate of 265–360 essential genes (4). Another experimental approach using antisense RNA to inhibit gene expression led to the identification of some 150 essential genes in *Staphylococcus aureus* (5). However, these approaches have limitations. Computation is likely to underestimate the minimal gene set because it takes into account only those genes that have remained similar enough during the course of evolution to be recognized as true orthologues. Transposon mutagenesis might overestimate the set by misclassification of nonessential genes that slow down the growth without arresting it but can also miss essential genes that tolerate transposon insertions (3, 6). Finally, the use of antisense RNA is limited to the genes for which an adequate expression of the inhibitory RNA can be obtained in the organism under study.

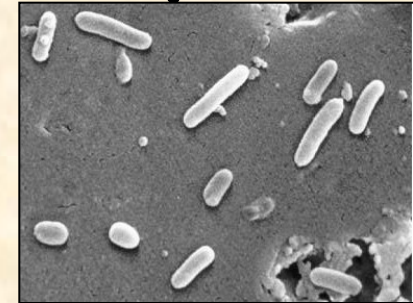
Table 2. *B. subtilis* essential genes

DNA metabolism	27
Basic replication machinery	16
Packaging and segregation	9
Methylation	2
RNA metabolism	14
Basic transcription machinery	4
RNA modification	6
Regulation	4
Protein synthesis	95
Ribosomal proteins	52
Aminoacyl-tRNA synthetases	24
Translation factors	10
Protein folding and modification	3
Protein translocation	6
Cell envelope	44
Membrane lipids	16
Cell wall	28
Cell shape and division	10
Glycolysis	8
Respiratory pathways	22
Isoprenoids	8
Menaquinone	8
Cytochrome biogenesis	3
Thioredoxin	3
Nucleotides	10
Cofactors	15
CoA	1
Folate	3
NAD	4
S-Adenosylmethionine	1
Iron-sulfur cluster	6
Other	15
Unknown	11
Total	271

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
AdTCCCC (STRR1)	
ATTCCCC	39
AATCCCC	19
AGTCCCC	15
AGCAGGGG (STRR6)	29
AAAATTC (STTR7)	13

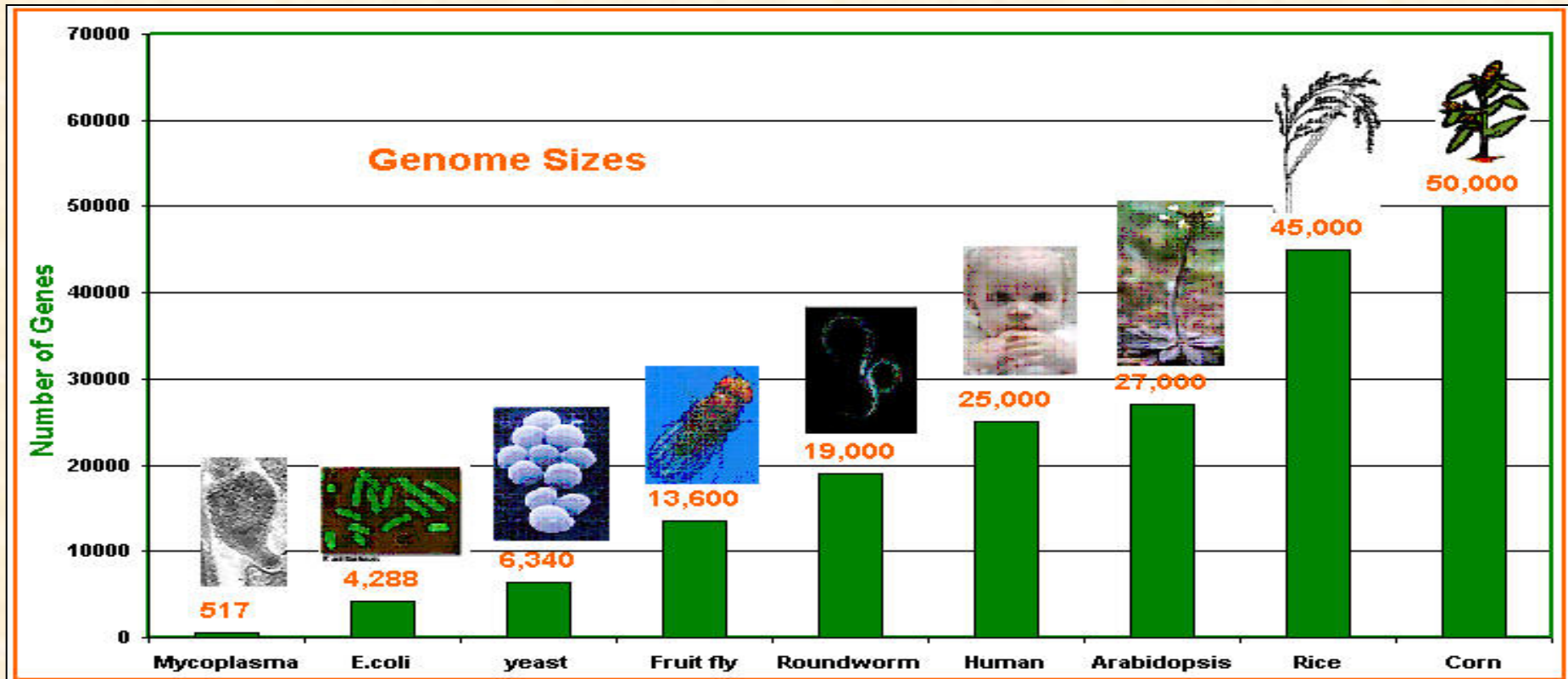
Velikosti genomů eukaryot

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



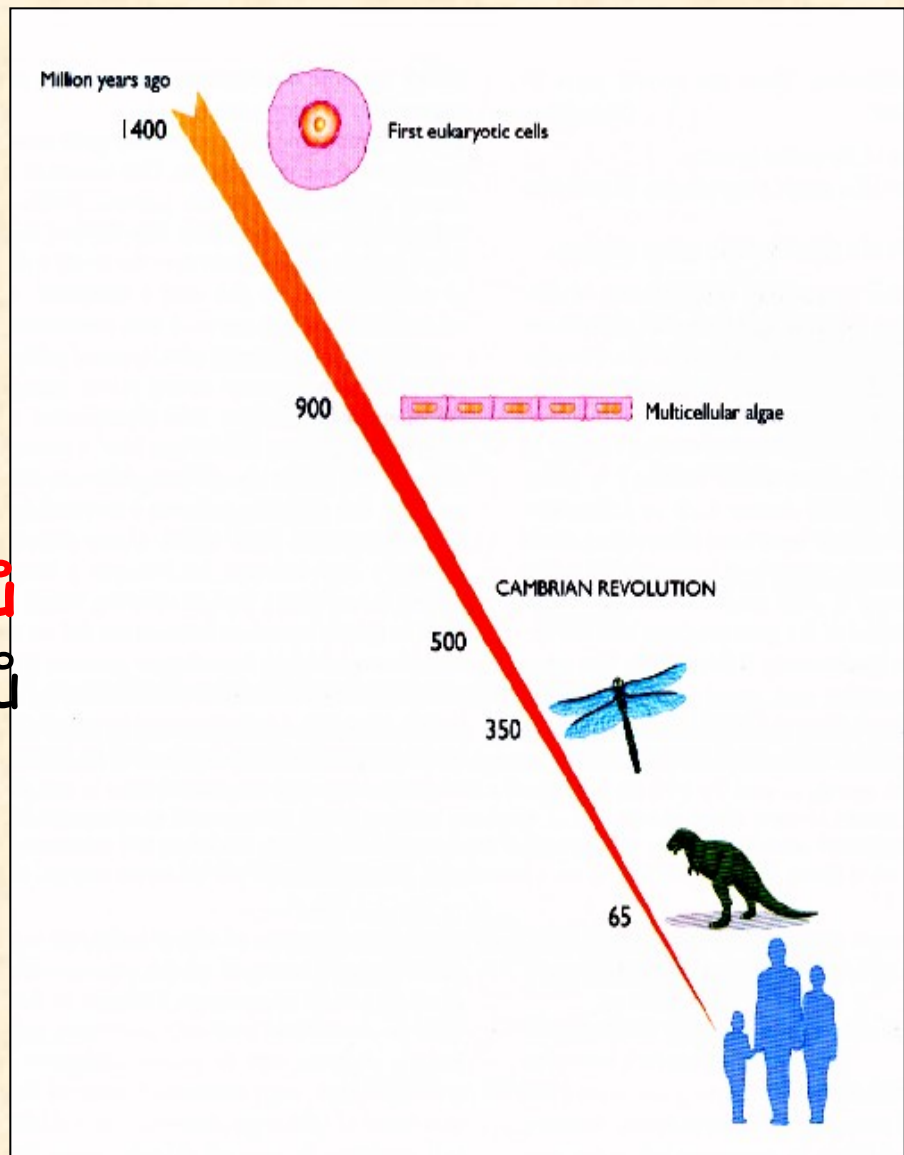
Organism	Genome size (base pairs)	Note
Virus, Bacteriophage MS2	3,569	First sequenced RNA-genome ^[2]
Virus, SV40	5,224	^[3]
Virus, Phage Φ -X174;	5,386	First sequenced DNA-genome ^[4]
Virus, Phage λ	50,000	
Bacterium, <i>Haemophilus influenzae</i>	1,830,000	First genome of living organism, July 1995 ^[5]
Bacterium, <i>Carsonella ruddii</i>	160,000	Smallest non-viral genome. ^[6]
Bacterium, <i>Buchnera aphidicola</i>	600,000	
Bacterium, <i>Wigglesworthia glossinidia</i>	700,000	
Bacterium, <i>Escherichia coli</i>	4,000,000	^[7]
Amoeba, <i>Amoeba dubia</i>	670,000,000,000	Largest known genome. ^[8]
Plant, <i>Arabidopsis thaliana</i>	157,000,000	First plant genome sequenced, Dec 2000. ^[9]
Plant, <i>Genlisea margaretae</i>	63,400,000	Smallest recorded flowering plant genome, 2006. ^[9]
Plant, <i>Fritillaria assyrica</i>	130,000,000,000	
Plant, <i>Populus trichocarpa</i>	480,000,000	First tree genome, Sept 2006
Yeast, <i>Saccharomyces cerevisiae</i>	20,000,000	^[10]
Fungus, <i>Aspergillus nidulans</i>	30,000,000	
Nematode, <i>Caenorhabditis elegans</i>	98,000,000	First multicellular animal genome, December 1998 ^[11]
Insect, <i>Drosophila melanogaster</i> aka Fruit Fly	130,000,000	^[12]
Insect, <i>Bombyx mori</i> aka Silk Moth	530,000,000	
Insect, <i>Apis mellifera</i> aka Honey Bee	1,770,000,000	
Fish, <i>Tetraodon nigroviridis</i> , type of Puffer fish	385,000,000	Smallest vertebrate genome known
Mammal, <i>Homo sapiens</i>	3,200,000,000	
Fish, <i>Protopterus aethiopicus</i> aka Marbled lungfish	130,000,000,000	Largest vertebrate genome known

ZMĚNY VE VELIKOSTI GENOMU

Zvětšení genomu

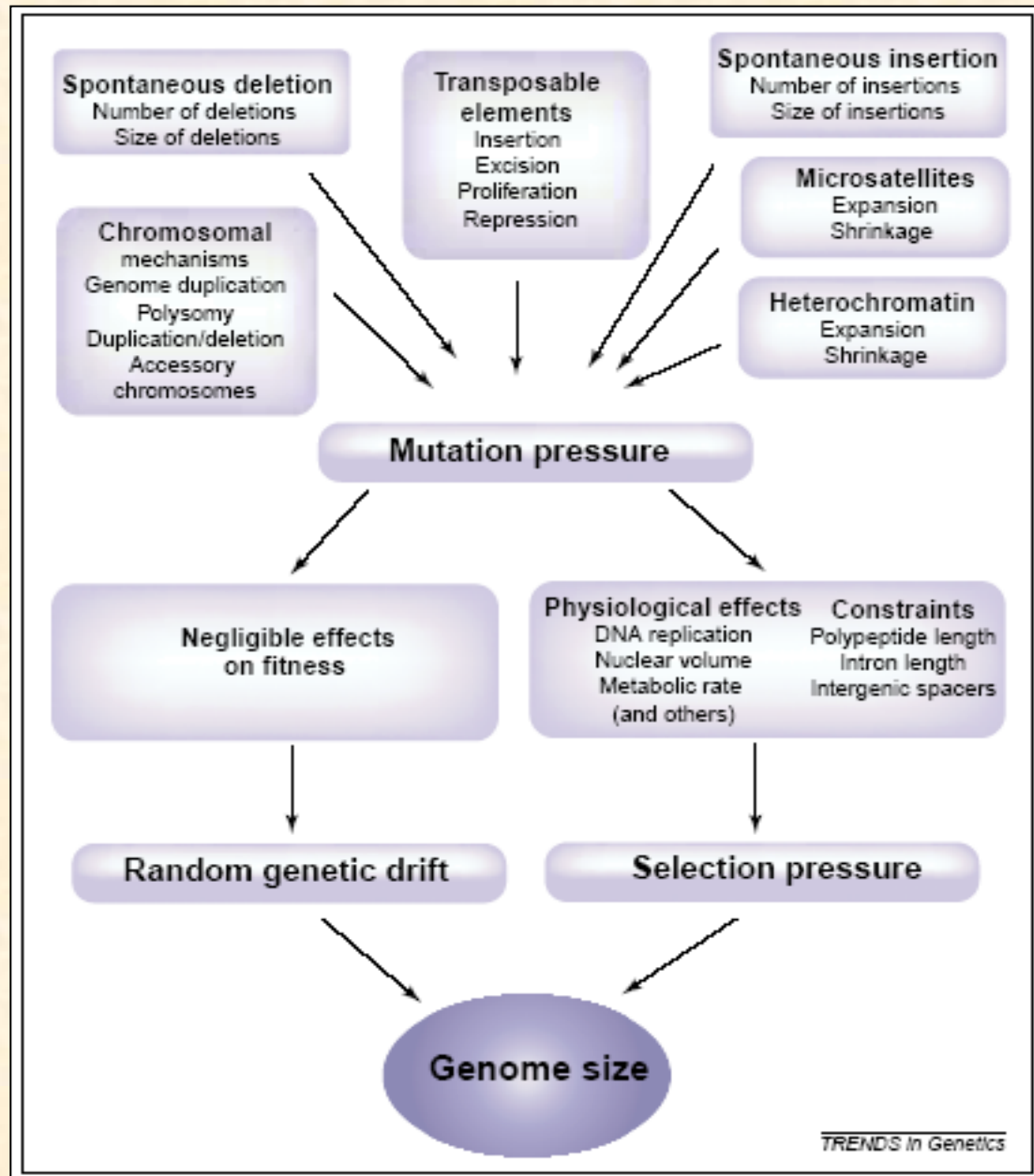
- celkové zvětšení:
polyploidizace (duplikace celého **genomu**)
- duplikace části genomu,
zmnožení počtu **chromosomů**
- duplikace **genů** a skupin genů

- amplifikace **transpozonů**
- inserce **virové DNA**
- inserce **organelové DNA**
- expanze (mikro)**satelitů**



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

Evoluční síly ovlivňující velikost genomu



Změny ve velikosti genomů: plynulé nebo skoky?

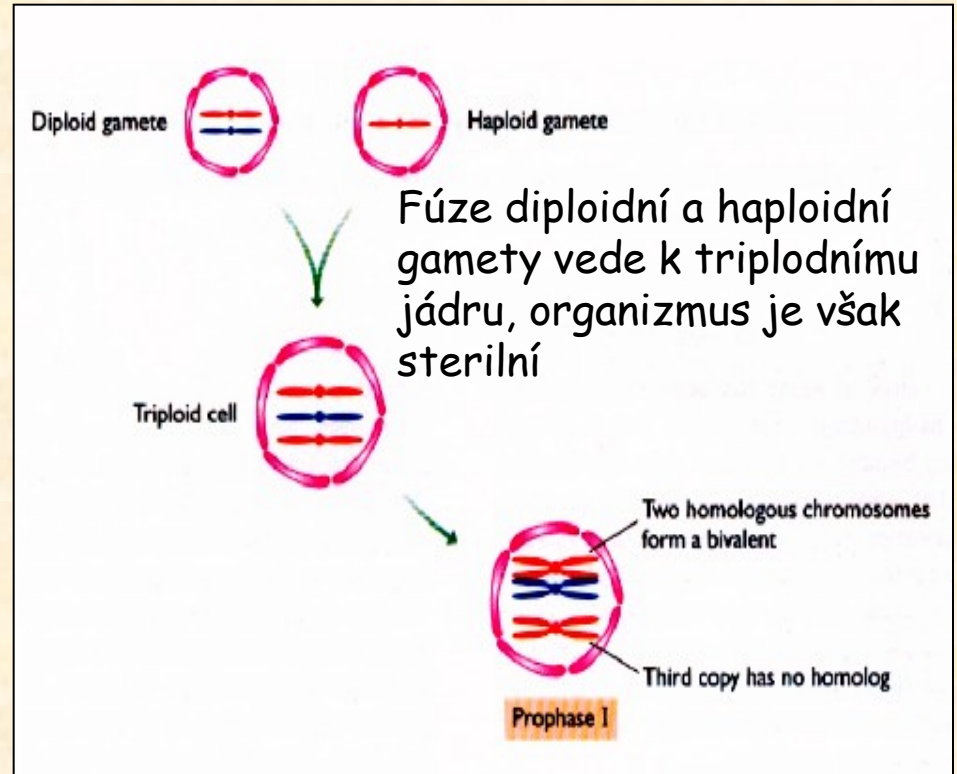
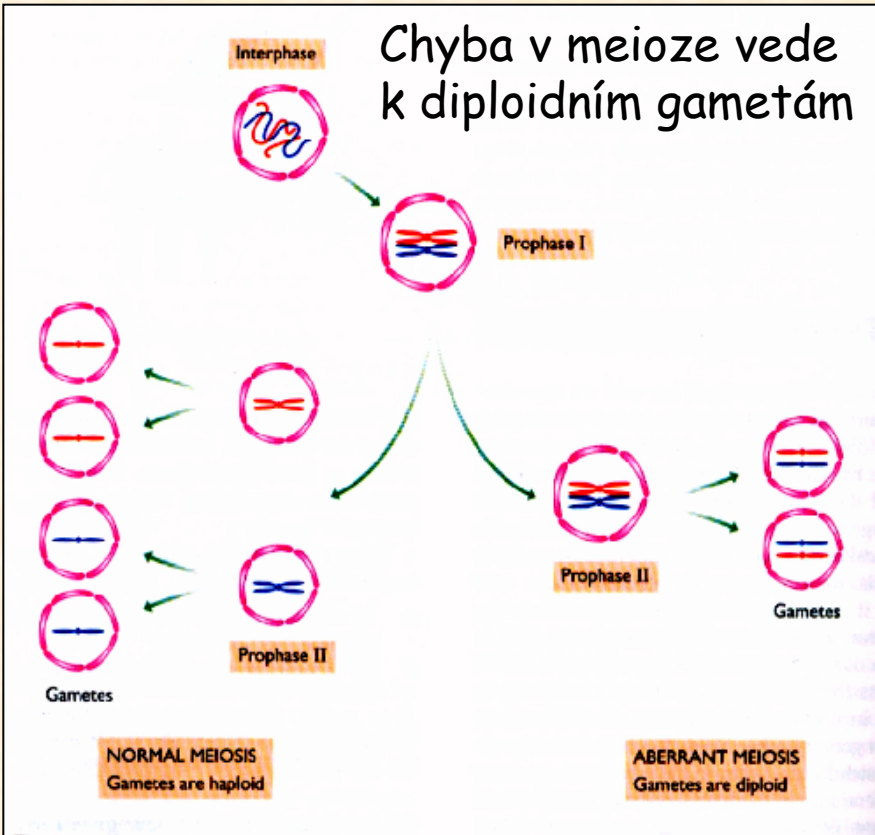
Plynulost:

- 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), vyjíměčně obratlovci
- polyplodizace, přídavná replikace některého chromosomu, kryptopolyploidizace (koordinovaná replikace všech repetitivních sekvencí v genomu - *Chironomus*)
- kvantové rozdíly mezi blízkými příbuznými, velké rozdíly ve velikostech genomů

Růst velikosti genomu: Polyploidizace



Autopolyploidie: multiplikace jedné základní sady chromosomů, jeden organizmus, chyba při meioze (růže...)

Allopolyploidie: kombinace geneticky odlišných sad chromosomů (pšenice...)

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

Polyploidní plodiny

- tetraploidi: kukuřice, bavlna, brambor, zelí



- hexaploidi: pšenice, chrysan téma

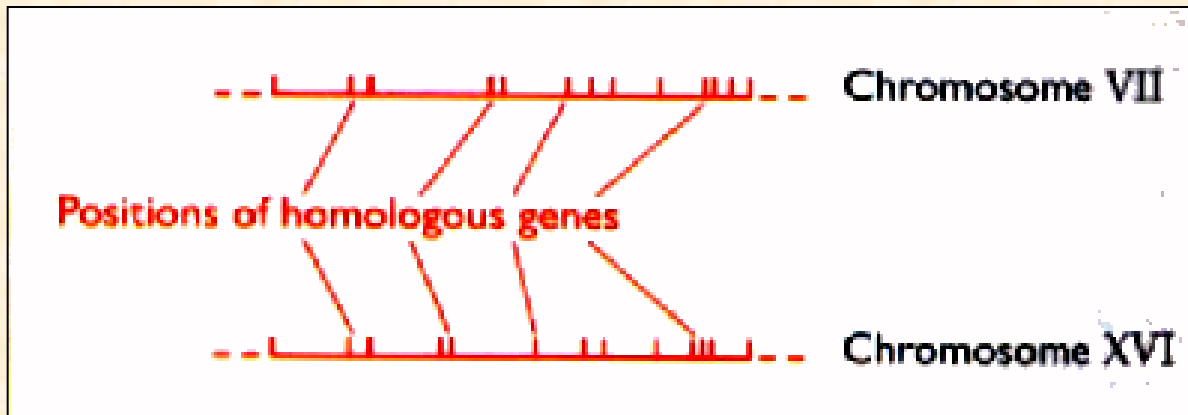


- oktoploidi: jahodník



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*



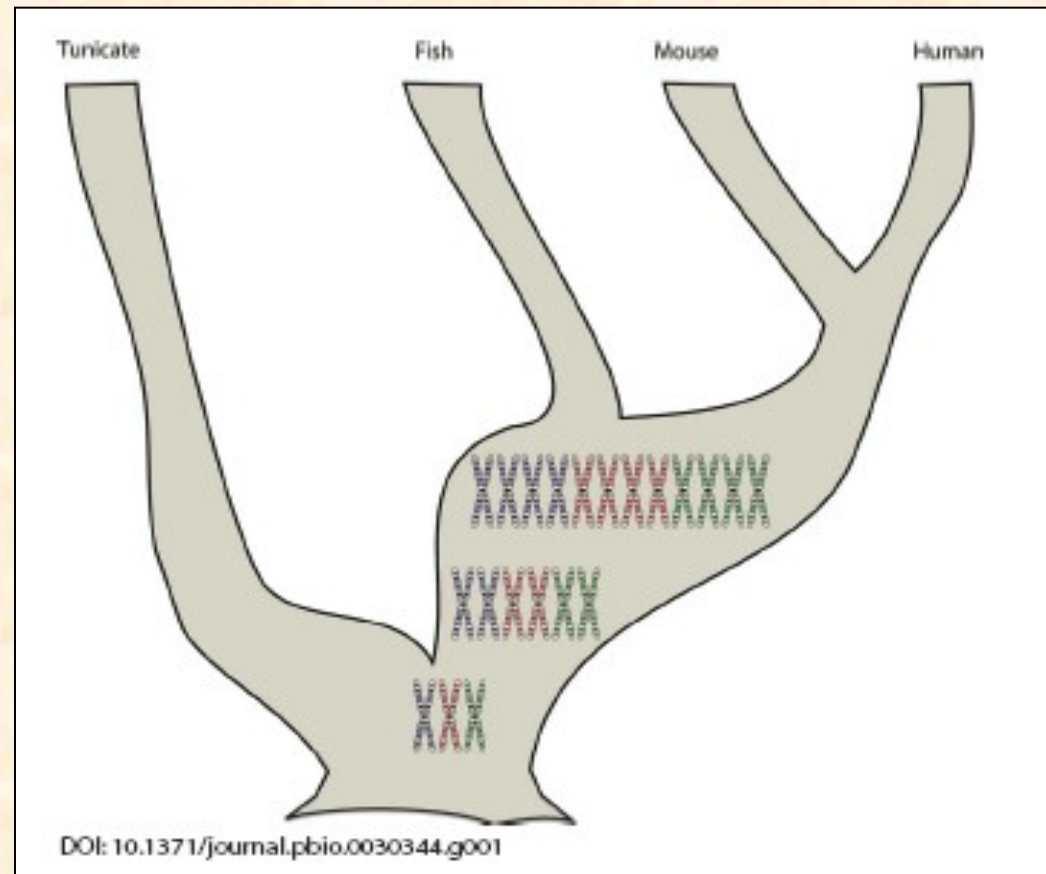
Duplikace jsou výsledkem tetraploidie spíše než postupných regionálních duplikací

Polyploidizace v linii obratlovců

S. Ohno:

- dvě genomové duplikace u obratlovců - hypotéza „2R“ (2 rounds)
- 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

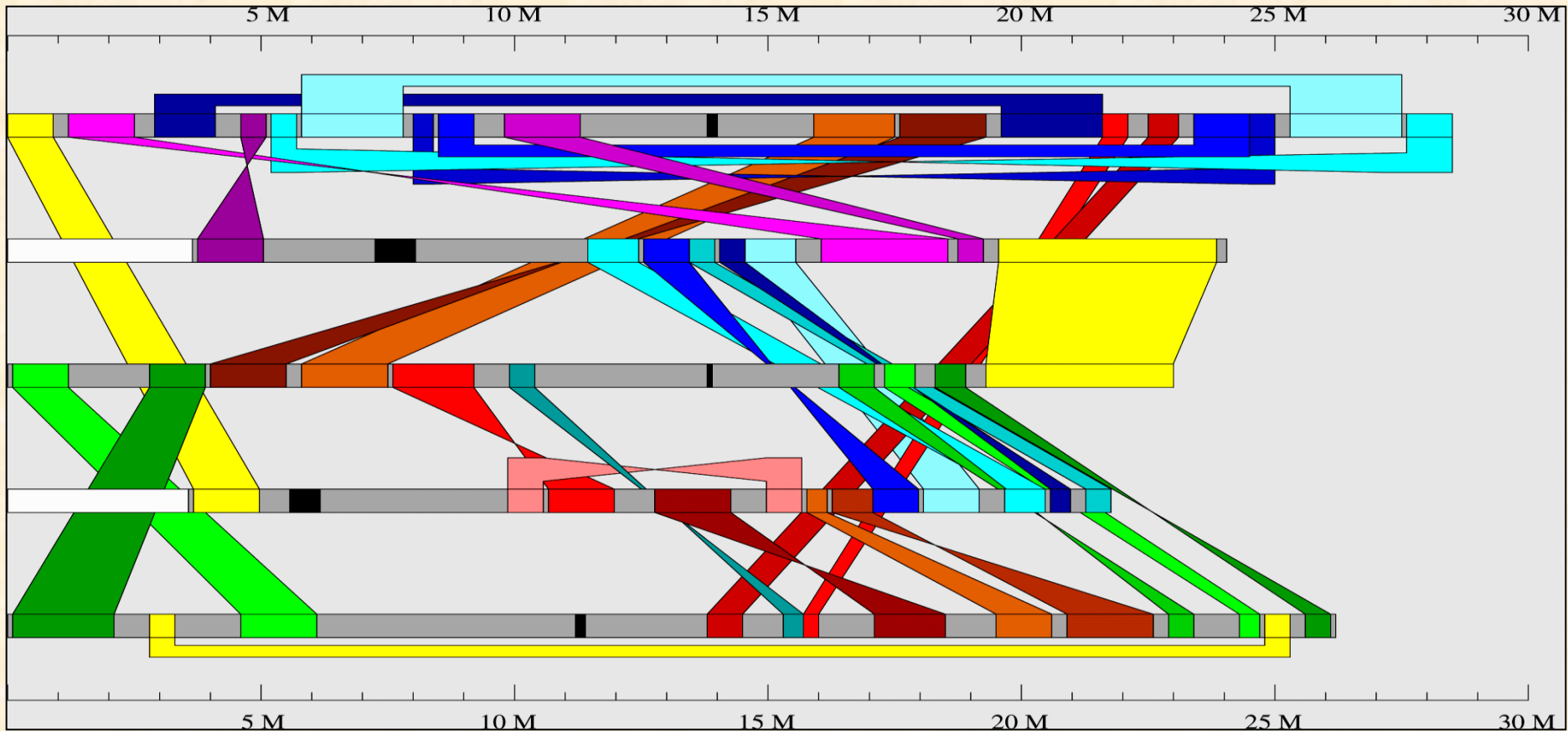
Duplikace genomu před 450 mil let napomohla vzrůstu komplexity a diverzifikaci obratlovců



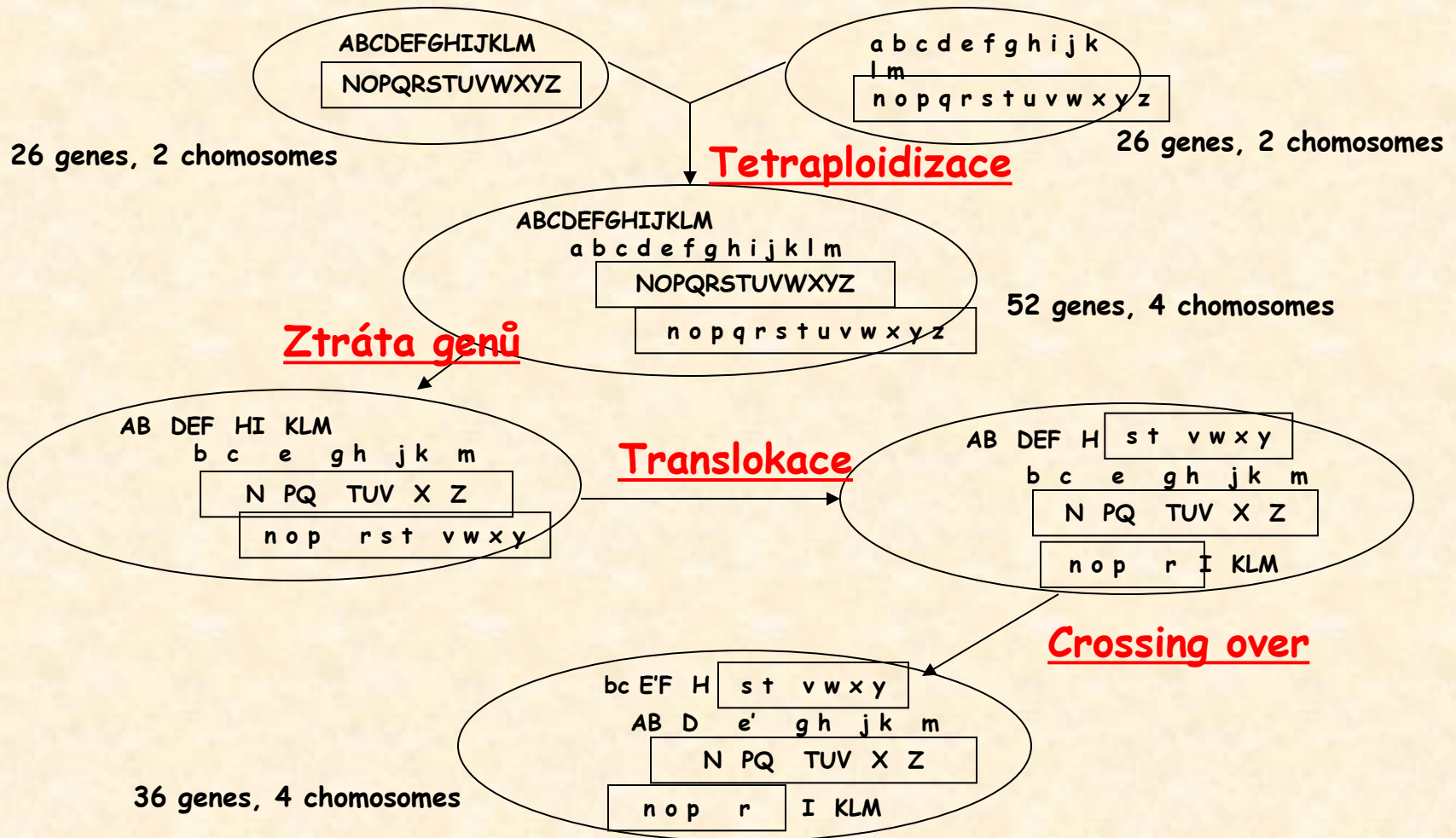
Polyploidizace u rostlin

Arabidopsis:

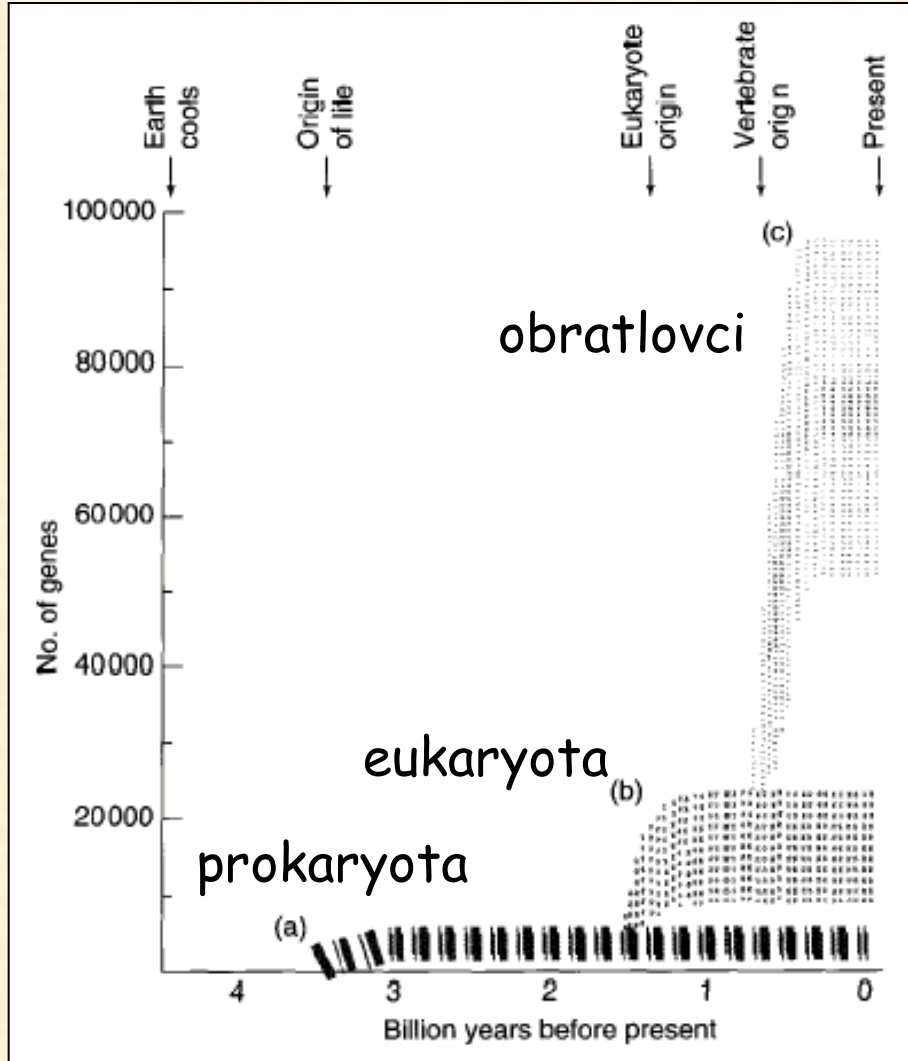
- duplikace před 200 a 80 mil. let
- možná i více než 4 genomové duplikace
- asi 60% genomu leží v duplikovaných segmentech
- 50% genů v nich je konzervováno



Poččet genů a evoluce duplikovaného genomu



Vzrůst počtu genů, potlačení šumu a biologická komplexita



Prokaryota/eukaryota:

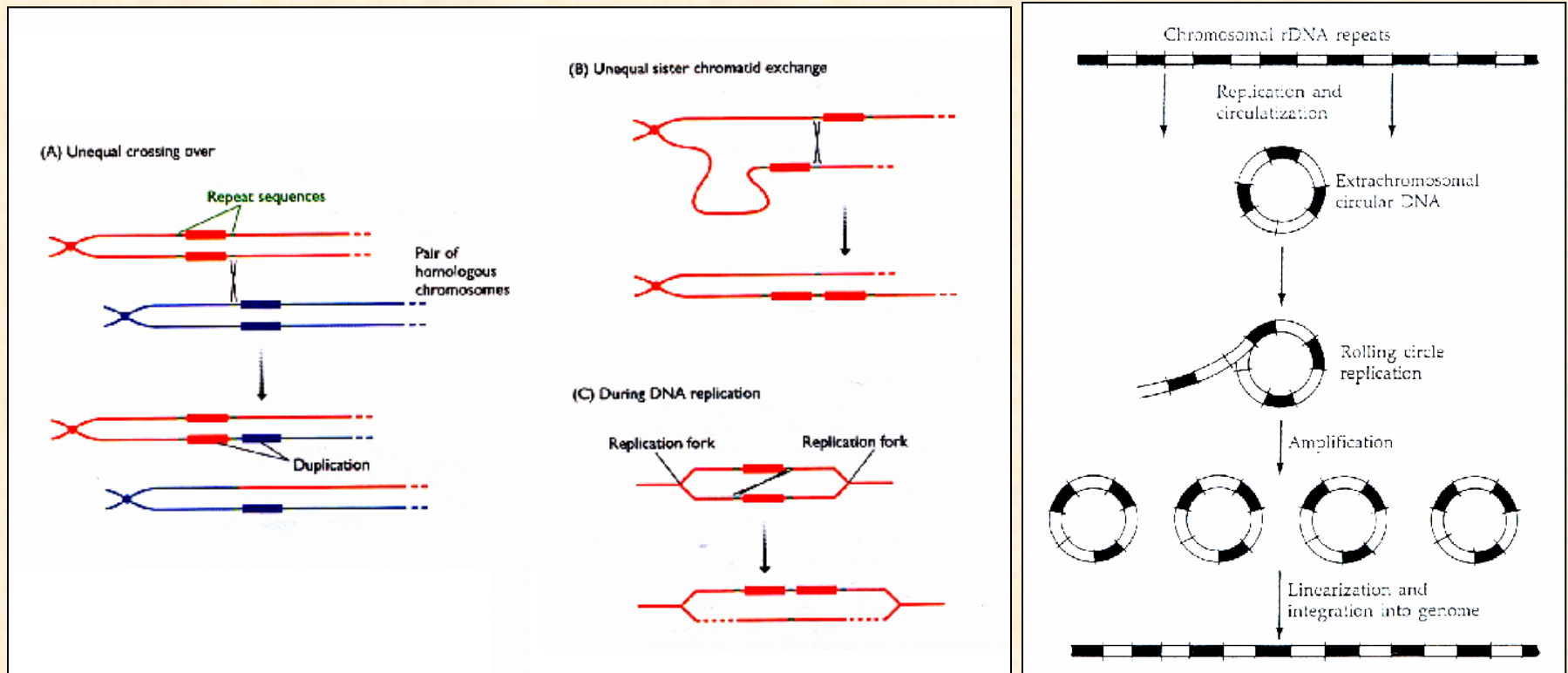
Potlačení šumu - **separace** transkripce a translace, jaderná membrána a histony

Obratlovci:

Potlačení šumu - **metylace** genů, mobilních elementů a duplikátů

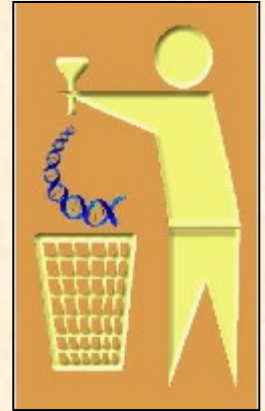
Mechanismy zvětšení genomů

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



NEKÓDUJÍCÍ DNA a VELIKOST GENOMU

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
- pufrování koncentrace regulačních proteinů

Sobeká 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

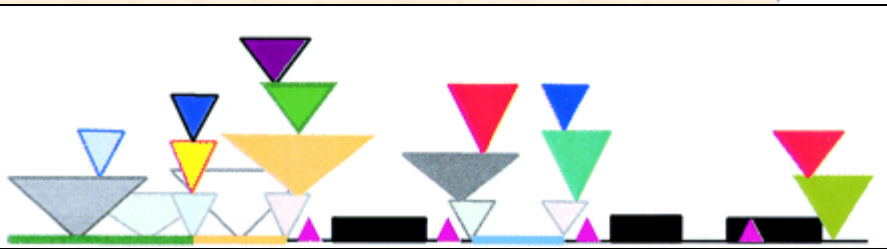
Limity růstu genomů:

- **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

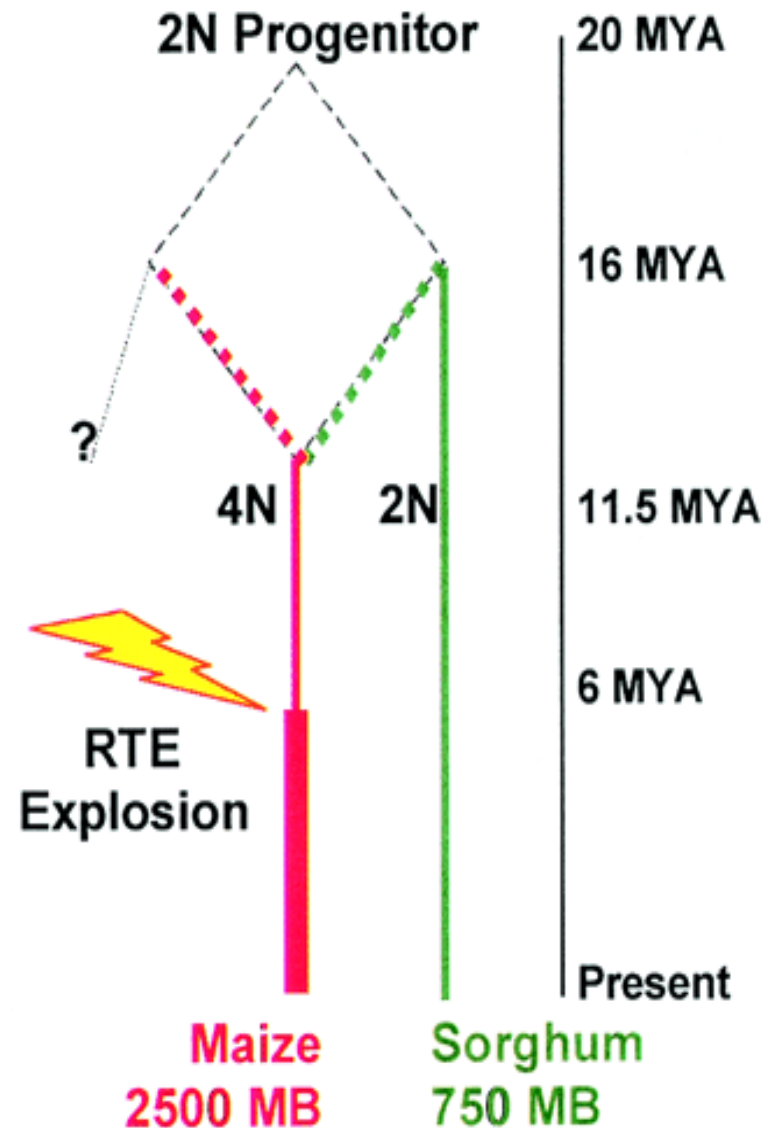


Commentary PNAS 98, 8163-8164 (2001)

Gene galaxies in the maize genome

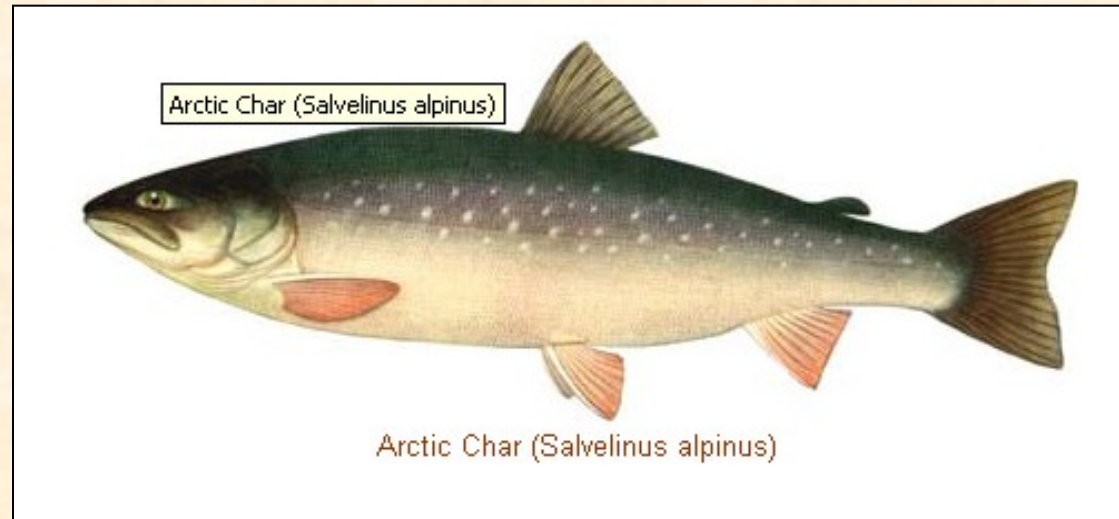
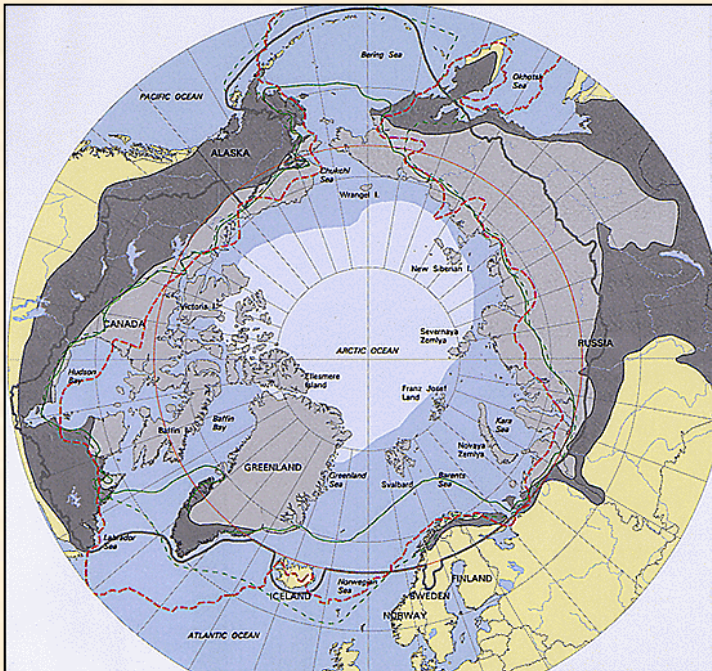
Virginia Walbot* and Dmitri A. Petrov

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



Obsah DNA je proměnlivý i v rámci jedince

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

Vliv velikosti genomu na fenotyp

Velikost genomu koreluje s:

+

- velikostí jádra
- velikostí buňky (nucleotypic effect)
- dobou mitózy a meiózy
- minimální generační dobou
- velikostí semen
- odpovědí letniček vůči CO_2
- dobou vývoje embrya u mloků

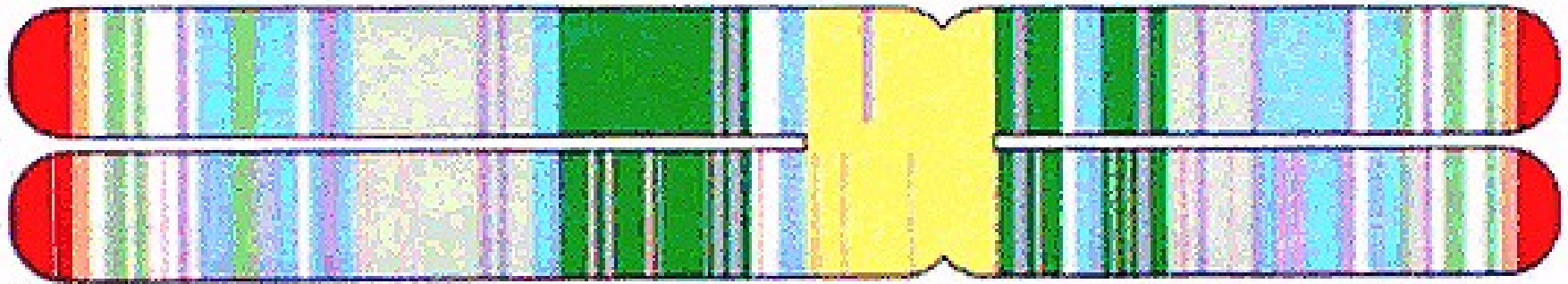


-

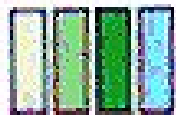
- rychlostí bazálního metabolismu u obratlovců (negativní korelace)
(malý genom ptáků a netopýrů - rychlý metabolismus při letu,
velký genom ryb - estivace za hypoxických podmínek)
- morfologickou komplexitou mozků u žab a mloků (negativní korelace)

TOPOGRAFIE GENOMU

Uspořádání genů v genomech



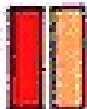
KEY



Tandem repeats



Centromere-associated tandem repeats



Telomeric and sub-telomeric repeats



Dispersed *Ty1/copia* retroelements and microsatellites

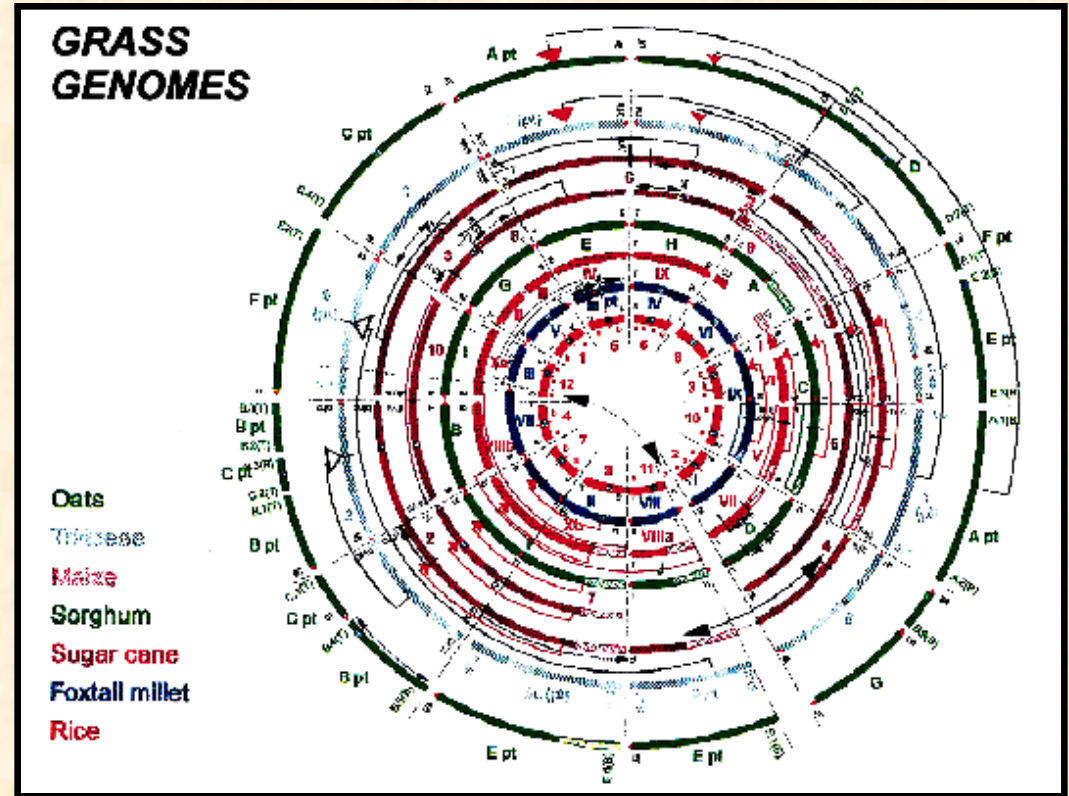
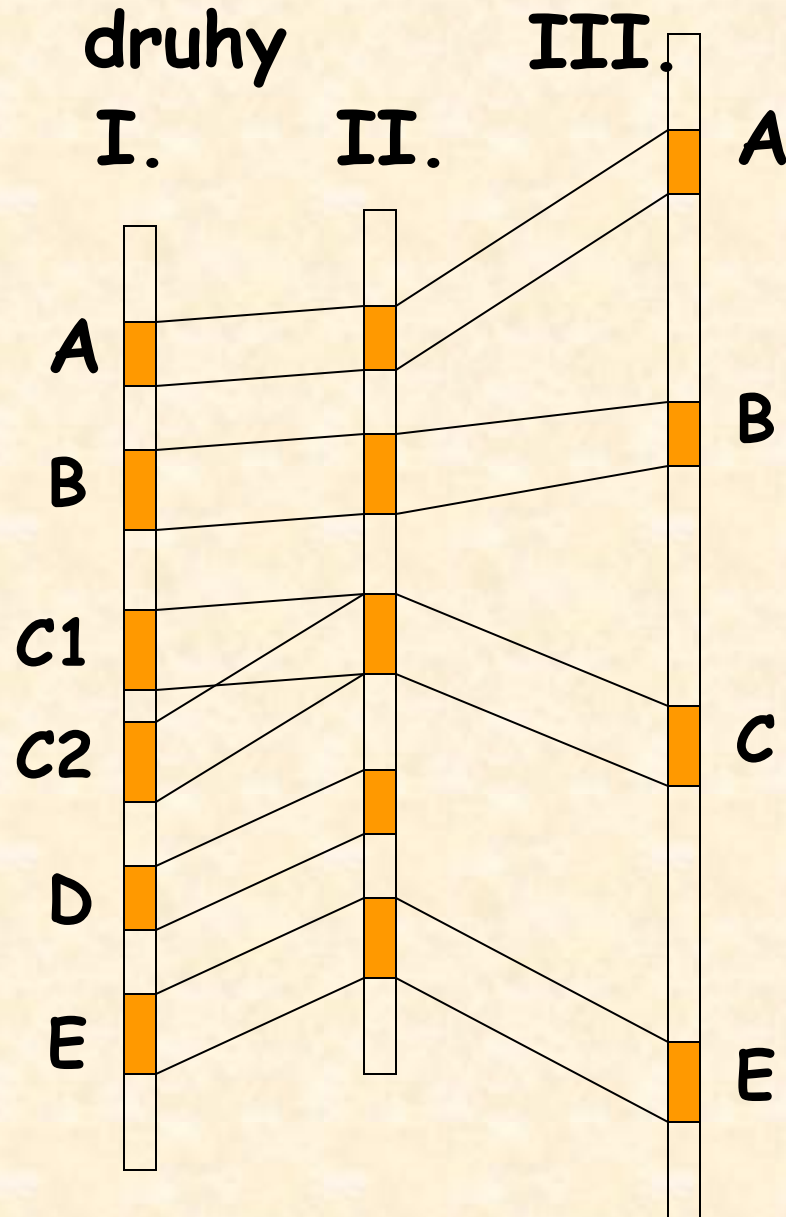


LINEs



Single and low-copy sequences including genes

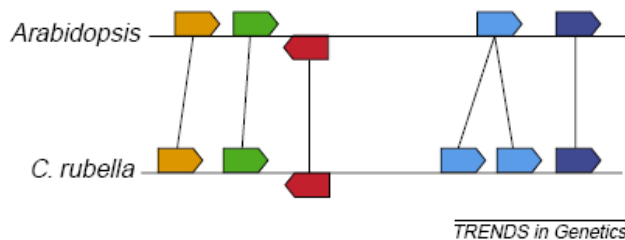
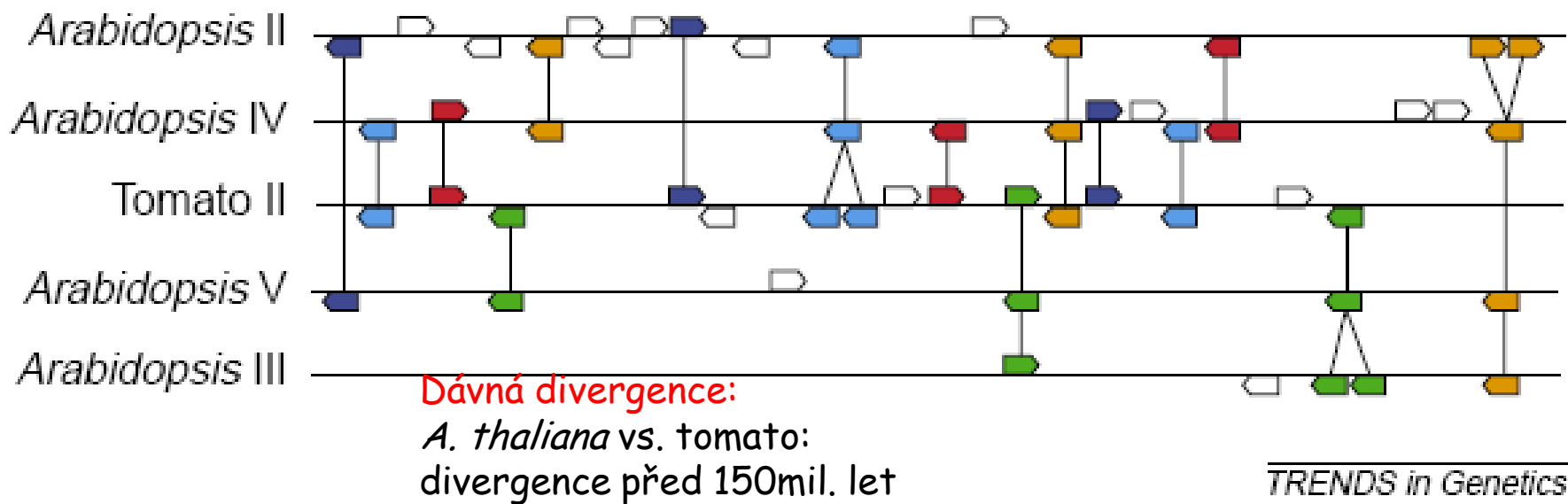
Syntenie - konzervativita pořadí genů



- 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



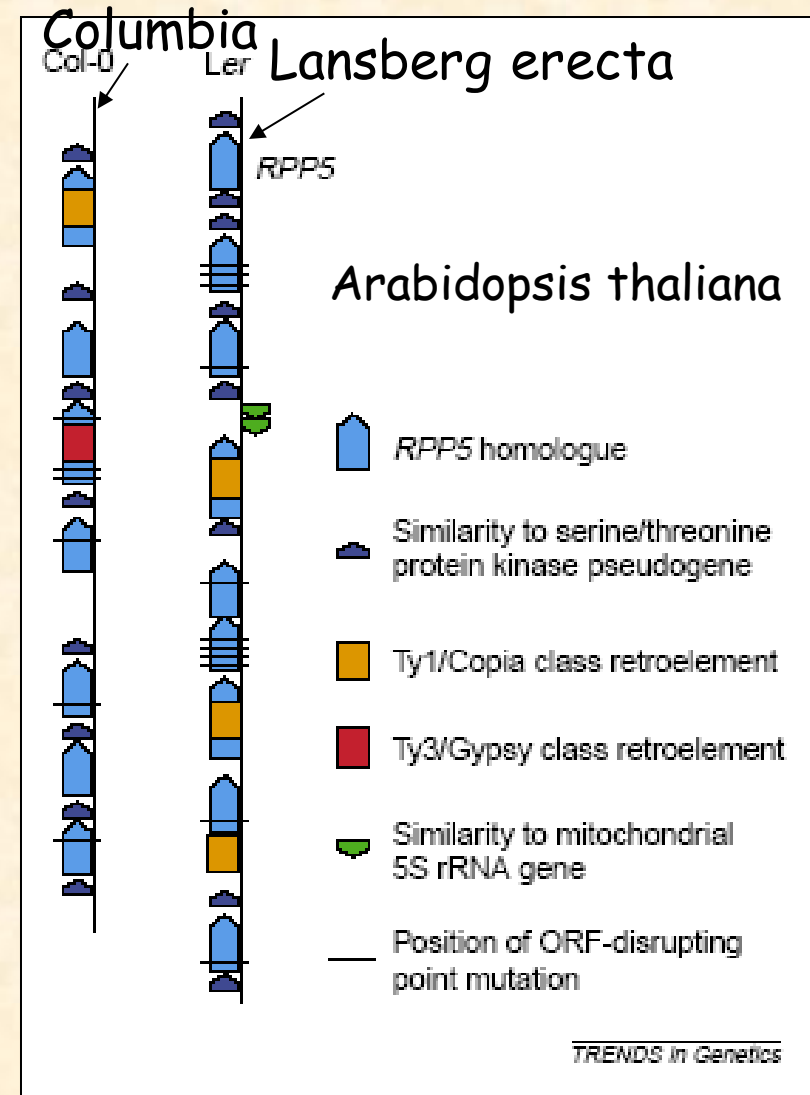
Nedávňá divergence:
A. thaliana vs. *Capsella rubella*:
divergence před 6.2-9.8 mil. let

(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



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

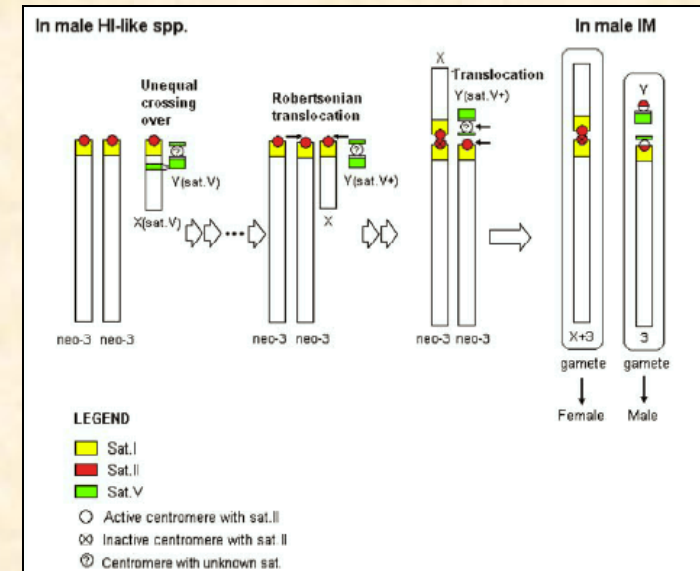
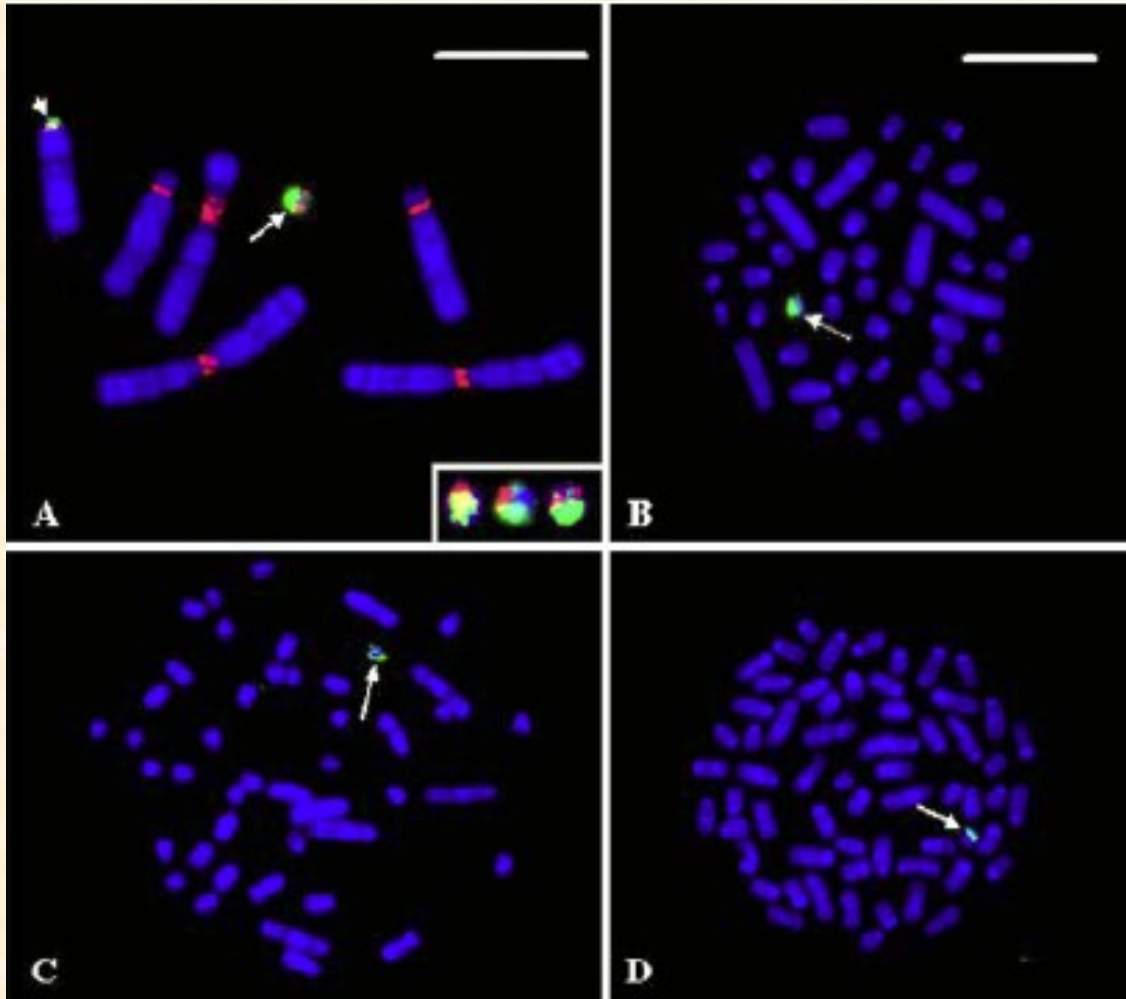


Různé počty chromosomů u blízce příbuzných druhů jelíneků *Muntiacus*

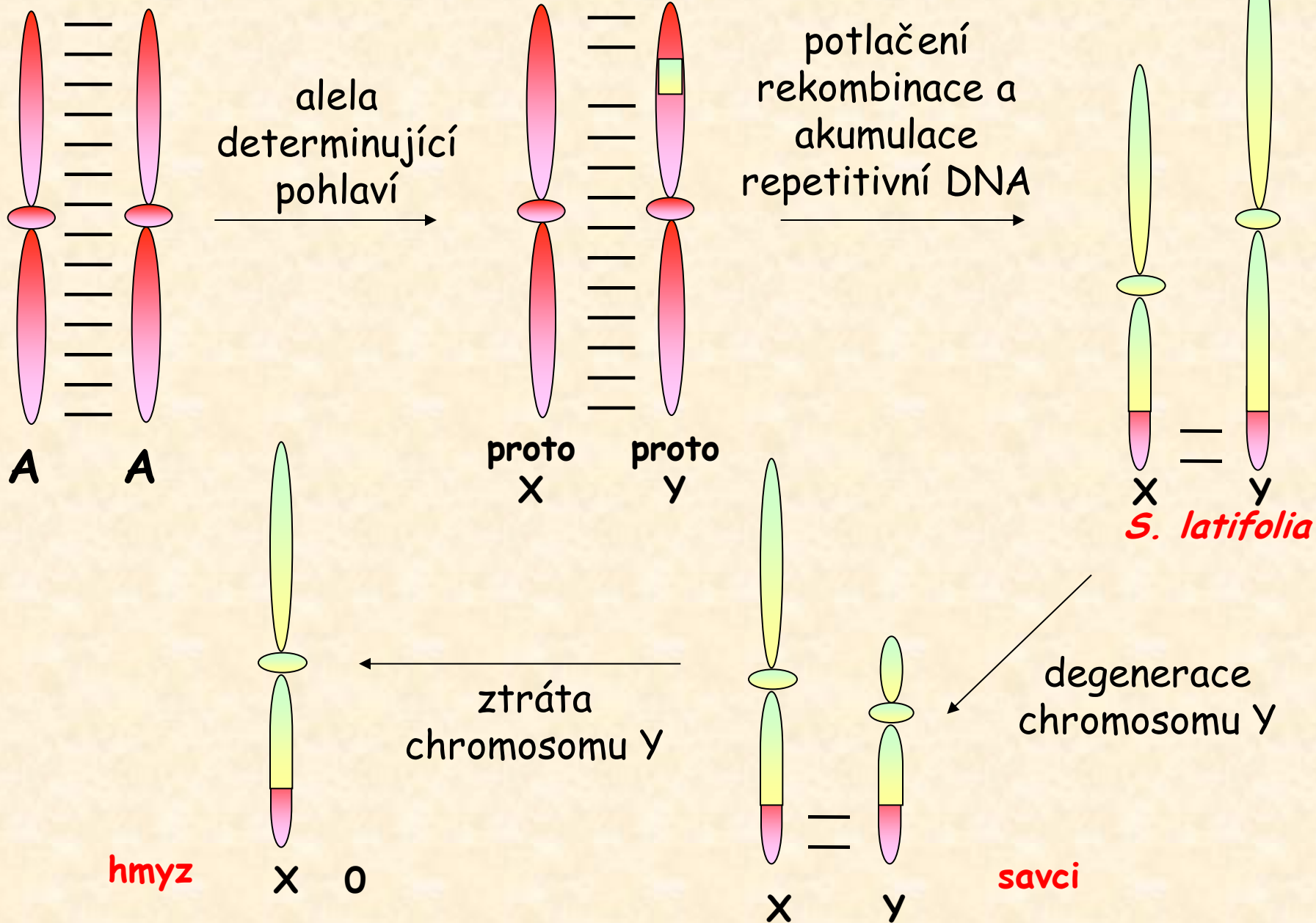
různé druhy:



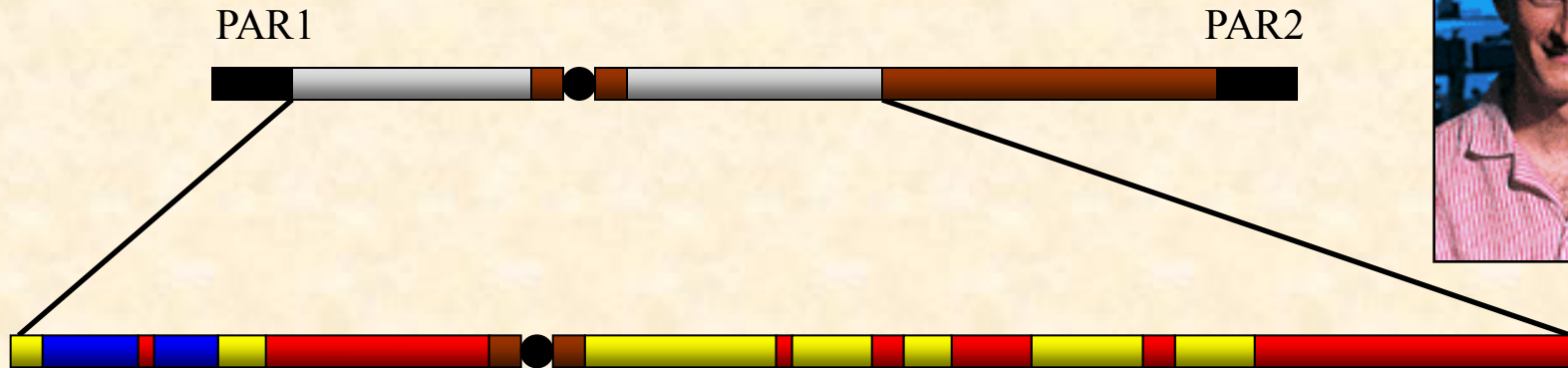
- josef hlasek
www.hlasek.com
antipens-gravesi-6632



Evoluce pohlavních chromosomů






Lidský chromosom Y: „Sál plný zrcadel“

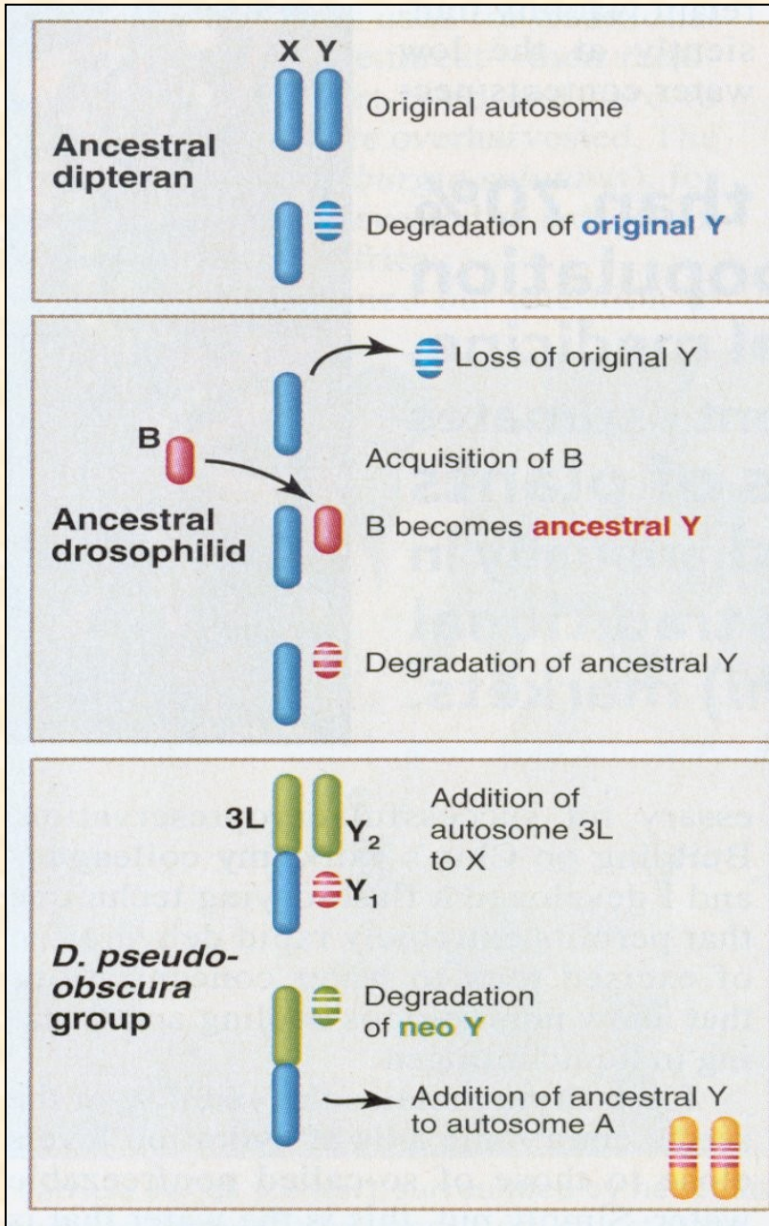


Palindromy:

- 3 000 000 bází dlouhé
- 99.9% identita

-  Homologní oblasti mezi X a Y
-  Duplikativní přenos z X
-  Palindromy

Recyklace pohlavních chromosomů u drozofily



První Y chromosom: „**original Y**“:
- vymizel před více než 60 mil. lety

Druhý Y chromosom: „**ancestral Y**“:
- párování B chromosomu s X chromosomem ~ před 60 mil let
- získání užitečných genů z autosomů
- degenerace

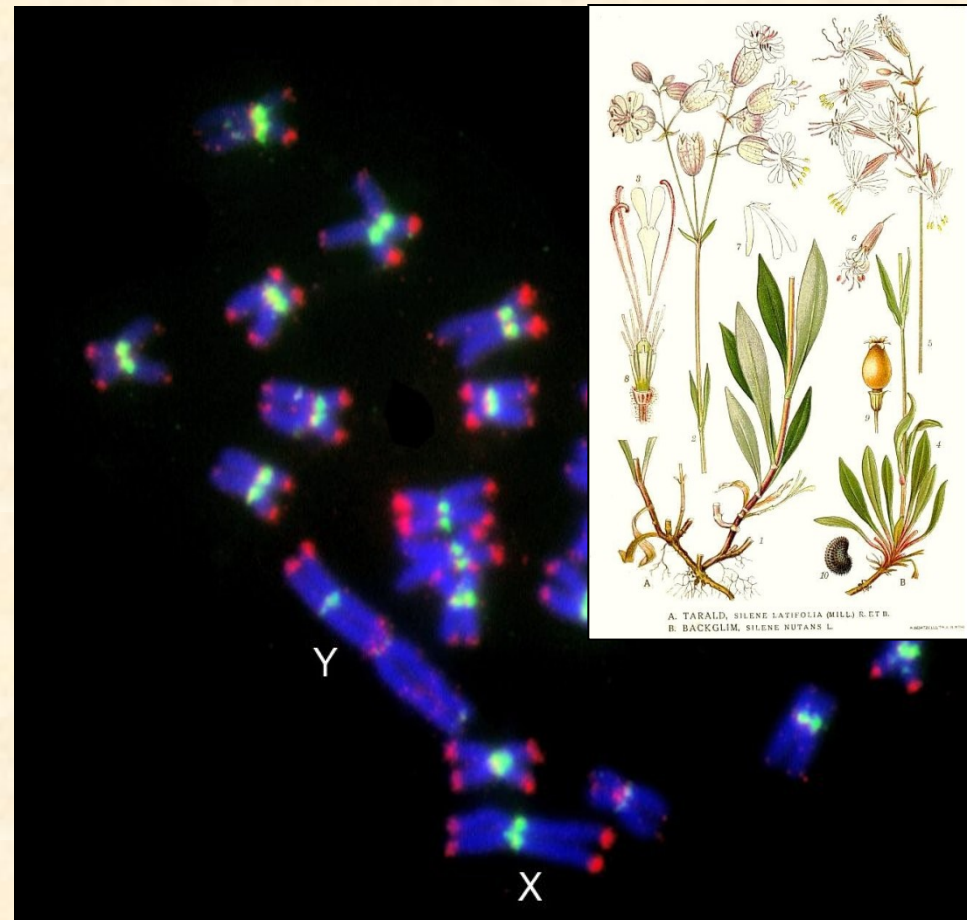
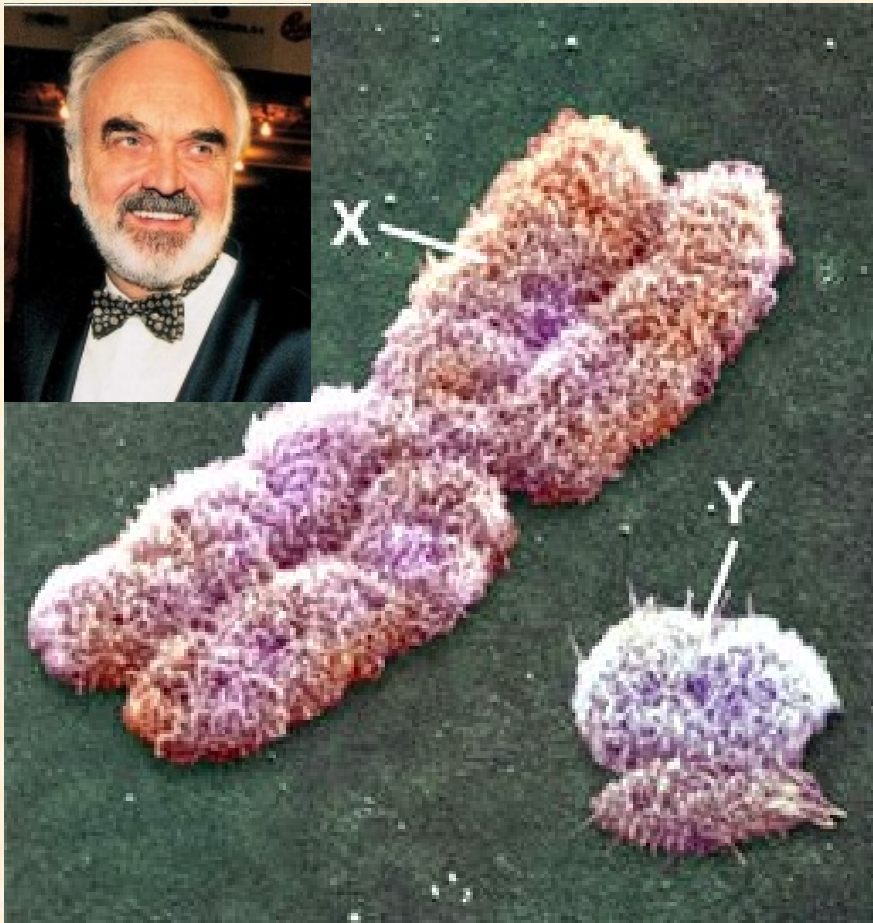
Třetí Y chromosom: „**neo Y**“
- fúze X s autosomem
- připojení ancestrálního Y na A, recovery
- opět degenerace neo-Y

Silenka širolistá: model pro studium mladých pohlavních chromosomů

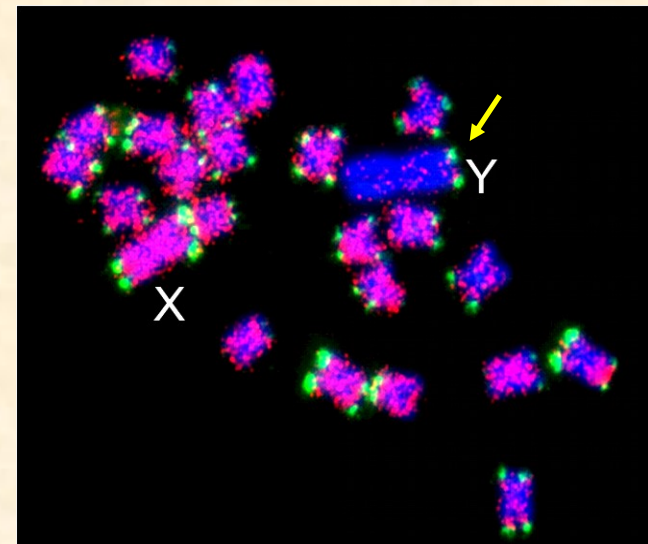
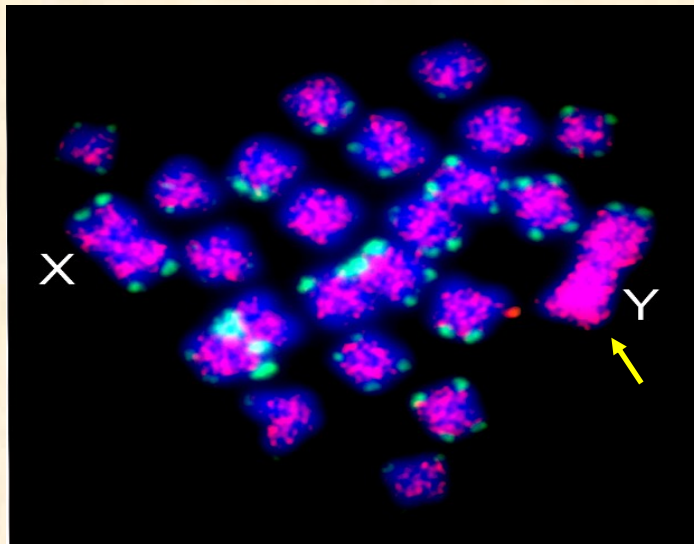
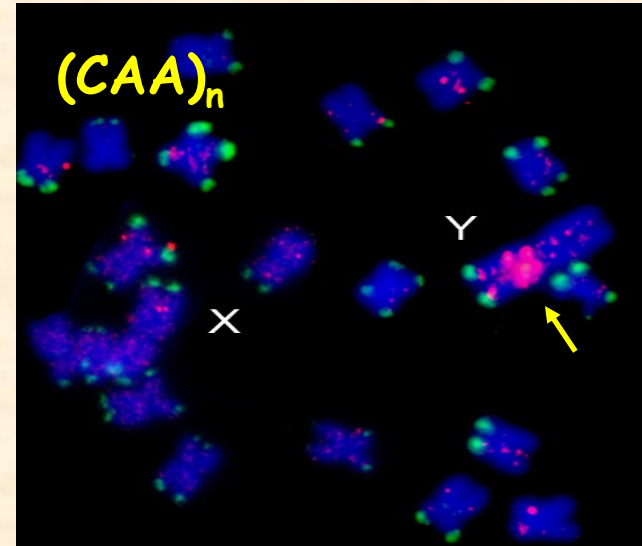
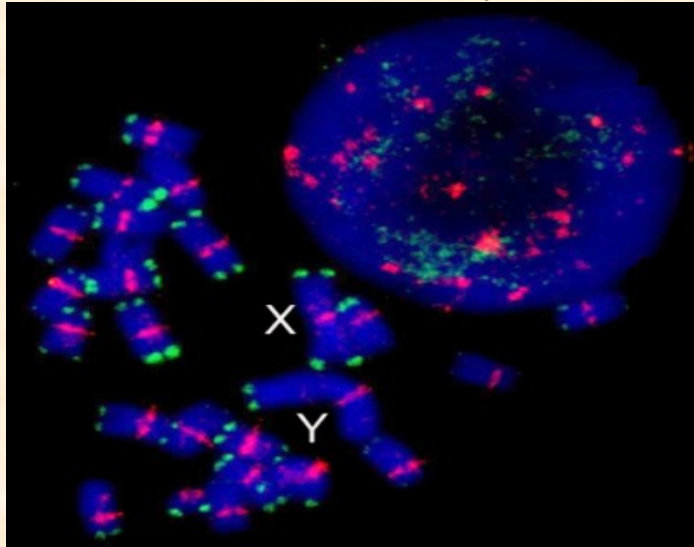
Degenerace chromosomu Y již v časných fázích jeho evoluce

člověk (*Homo sapiens*): 300 mil let

silenka širolistá (*Silene latifolia*): 10 mil let



Akumulace repeticí na mladém chromosomu Y



Chromosome Research (2008) 16:961–976
DOI: 10.1007/s10577-008-1254-2

© Springer 2008

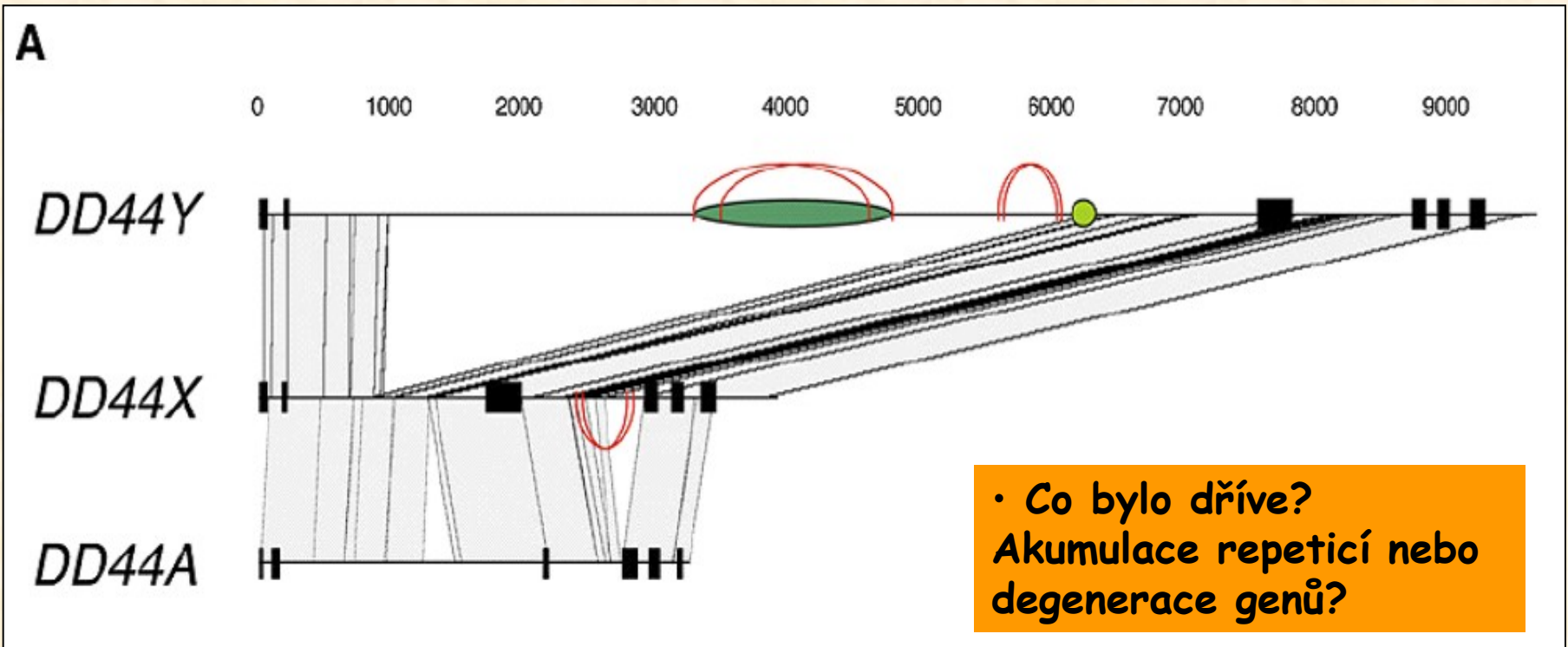
Survey of repetitive sequences in *Silene latifolia* with respect to their distribution on sex chromosomes

Tomas Cermak¹, Zdenek Kubat¹, Roman Hobza¹, Andrea Koblikova², Alex Widmer³, Jiri Macas², Boris Vyskot¹ & Eduard Kejnovsky^{1*}

Microsatellite accumulation on the Y chromosome in *Silene latifolia*

Zdenek Kubat, Roman Hobza, Boris Vyskot, and Eduard Kejnovsky

Degenerace genů na chromosomu Y: Více repetitivní DNA v genech na chromosomu Y



Current Biology 18, 545–549, April 8, 2008 ©2008 Elsevier Ltd All rights reserved DOI 10.1016/j.cub.2008.03.023

Report

Evidence for Degeneration of the Y Chromosome in the Dioecious Plant *Silene latifolia*

Gabriel A.B. Marais,^{1,*} Michael Nicolas,^{2,6} Roberta Bergero,³
Pierre Chambrier,² Eduard Kejnovsky,⁴ Françoise Monéger,²
Roman Hobza,^{4,5} Alex Widmer,⁵ and Deborah Charlesworth³

B chromosomy - parazité

Definice: přídavné postradatelné chromosomy přítomné v některých jedincích některých populací určitých druhů, které zřejmě vznikly z A chromosomů

Vlastnosti B chromosomů:

1. postradatelnost
2. původ z A chromosomů, odlišnost od A chromosomů
3. nerekombinují
4. neúčastní se meiózy, nemendelovská dědičnost

Historie:

Wilson (1906) - „extra“ chromosomy u hmyzu *Metapodius*

Stevens (1908) - u brouka *Diabrotica*

Kuwada (1915) - kukuřice

Rozšíření: 10 druhů hub, 1300 rostlin, 500 živočichů

Převaha u určitých taxonů - Compositae, Graminae, Liliaceae,

Orthoptera, např. u hmyzu Orthoptera ~ u 10-15% druhů

B chromosomy - parazité

Vznik: aneuploidie, ze sex chromosomů (i naopak), z jiných druhů

Počet:

- vyjímečně více než 3-4 B chromosomy, rozdíly mezi populacemi i mezi jedinci
- extrémy: 50 u *Pachyphytum fittkai*, 34 u kukuřice, 26 u *Fritillaria japonica*, 24 u myši *Apodemus peninsulae*
- mitoticky stabilní - stejný počet v buňkách téhož jedince

Velikost a struktura:

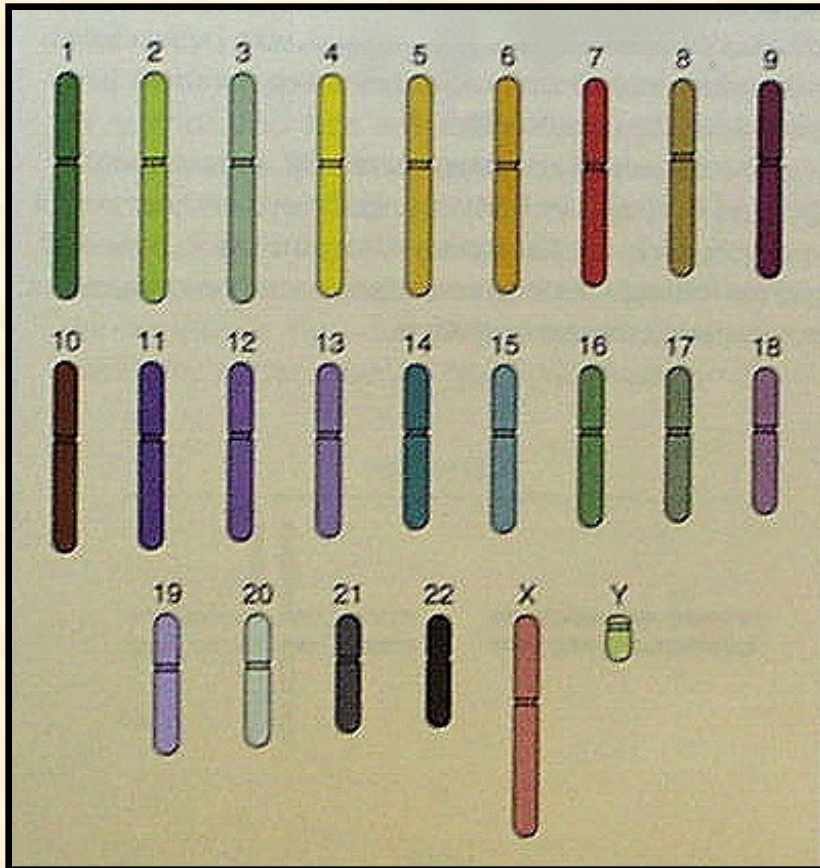
- velikost průměrných chromosomů, extrémy - největší (ryba *Astyanax scabripinnis*) nebo nejmenší (myš *Reithrodontomys megalotis*)
- metacentrické nebo akrocentrické, heterochromatické,
- obsahují repetice, satelitní DNA, rDNA, transpozony (mikrodisekce a PCR), nerekombinují

Chování v meióze:

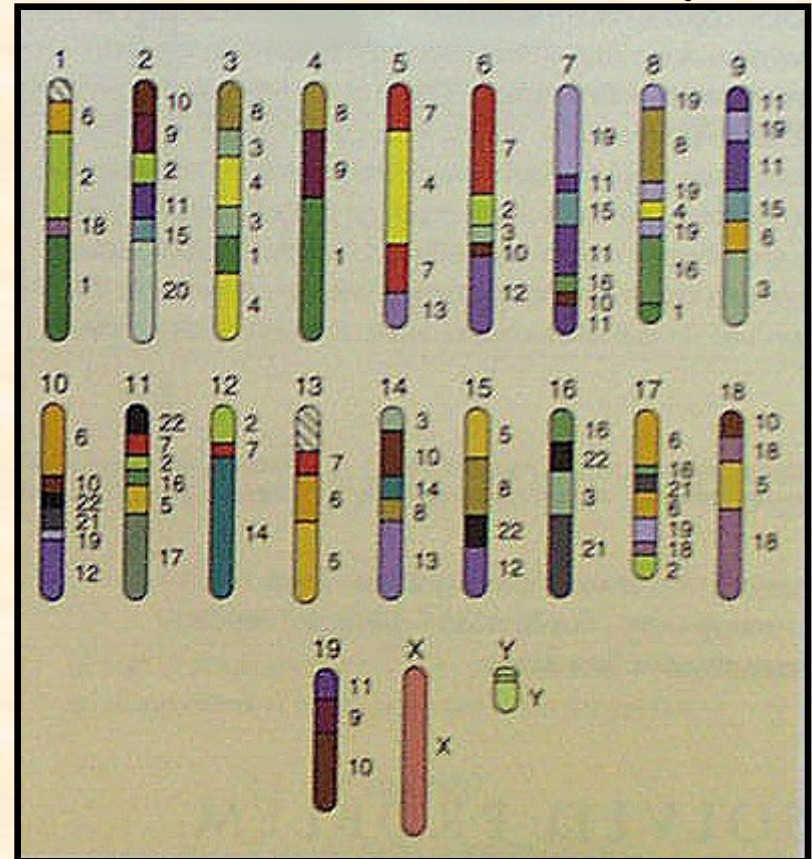
- nepárující se s partnerem, univaletní, segregují přednostně k určitému pólu, podobnost s pohlavními chromosomy, MSCI,

Srovnání lidských a myších chromosomů

člověk

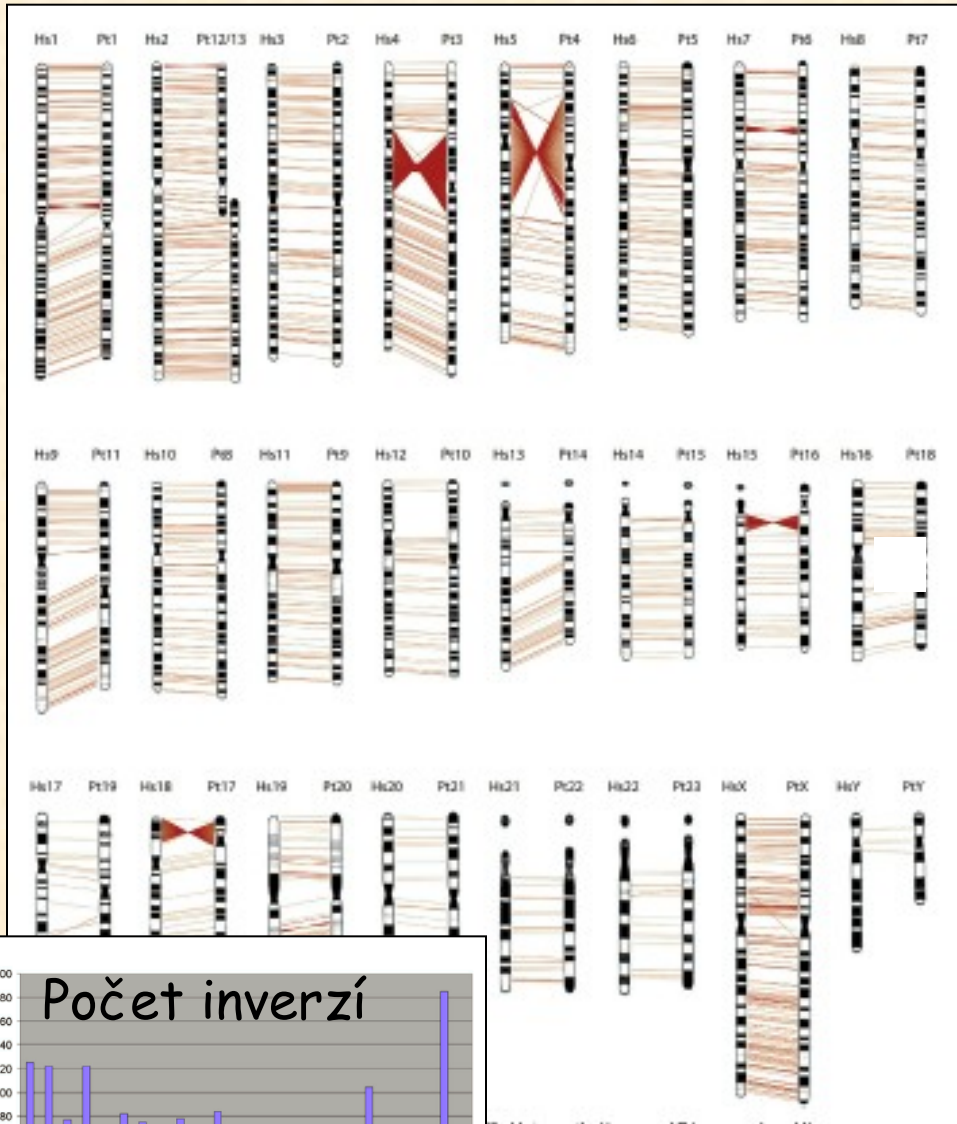


myš

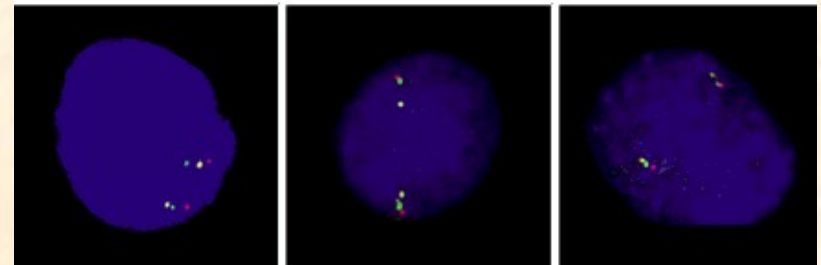


Od evoluční divergence došlo k četným přestavbám, které umístily bloky genů do různých kombinací.

Inverze v genomech člověka a šimpanze

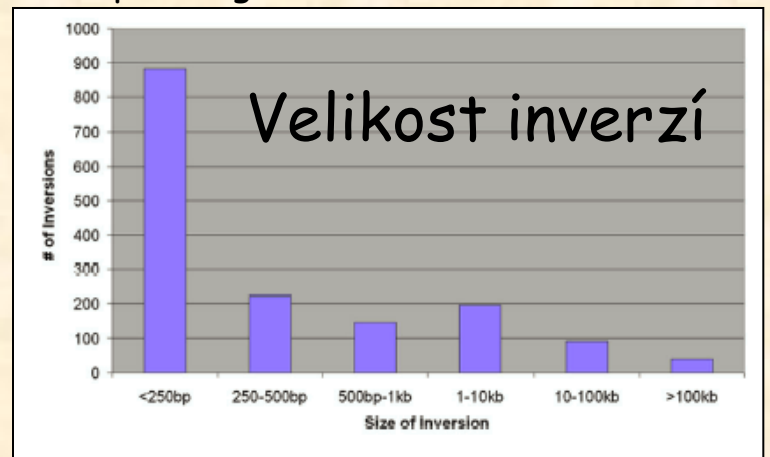


Human Chimpanzee Gorilla
B. 7p22 region



Human Chimpanzee Gorilla

Člověk: červená-žlutá-zelená
Šimpanz a gorila: červená-zelená-žlutá



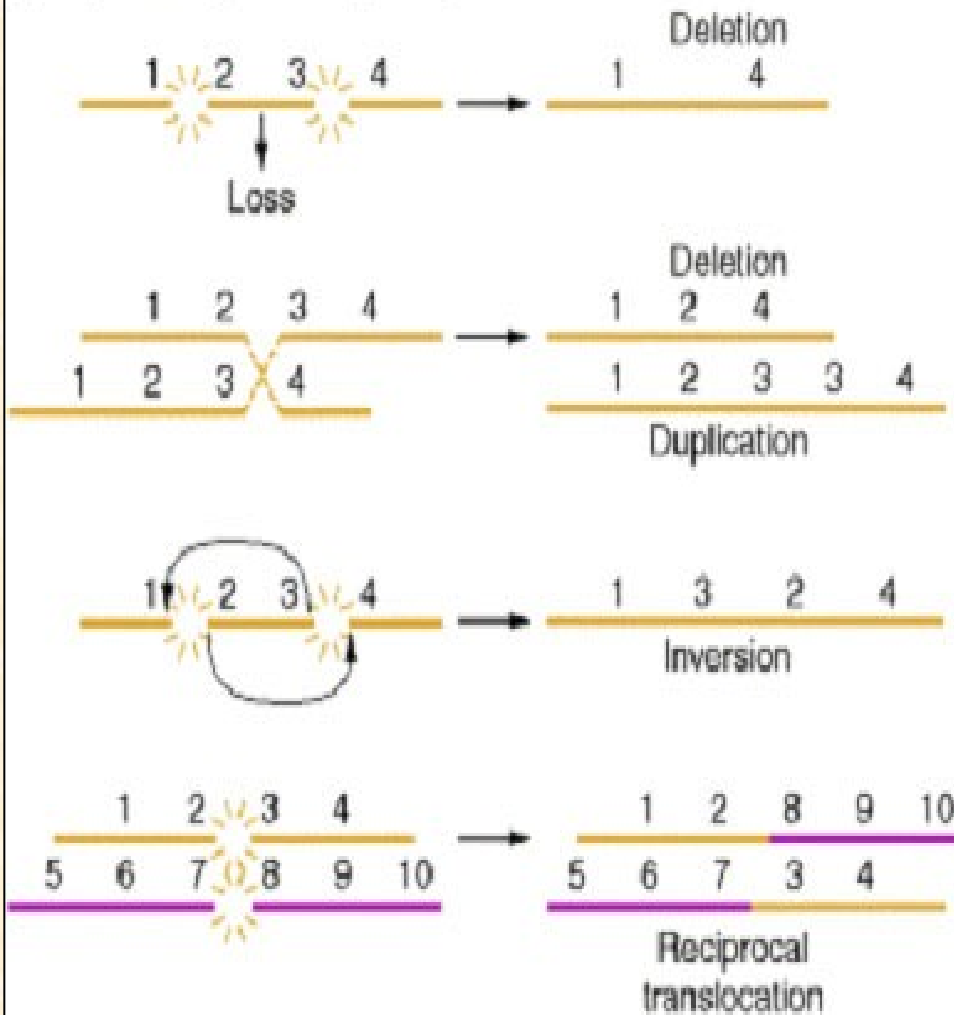
Chromosomové přestavby

zlom a znovuspojení (del., dupl., inv.)

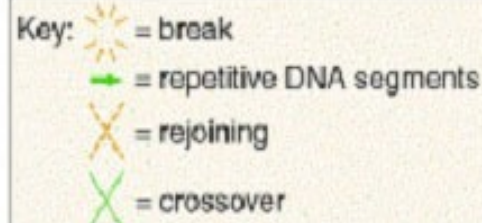
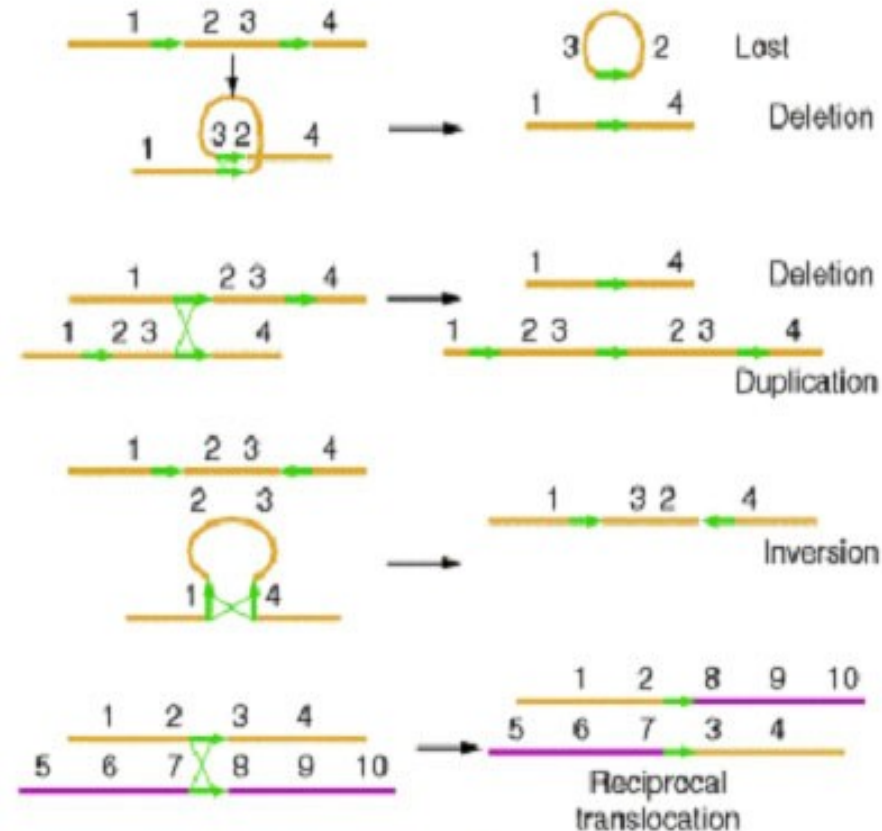
rekombinace mezi repetitivními

Origins of Chromosomal Rearrangements

(a) By breakage and rejoining



(b) By crossing-over between repetitive DNA



Isochorový model organizace genomu

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

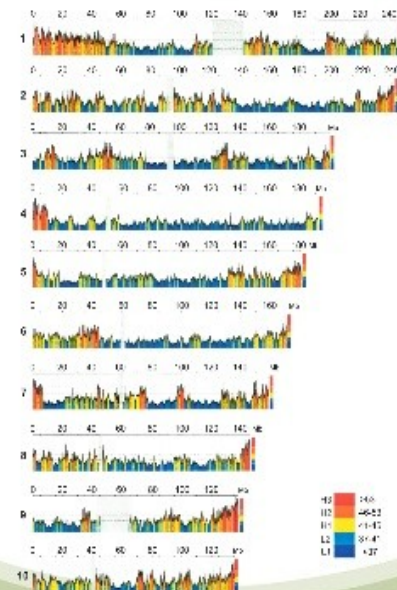
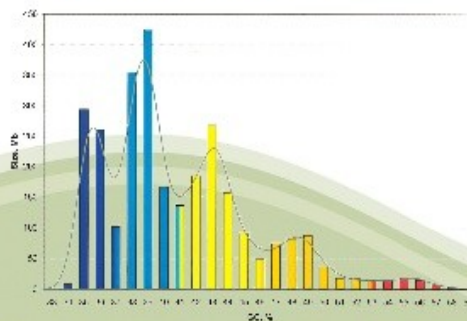


Figure 1 | Large-scale variation in G+C content. a | The classic isochore model. b | G+C content across the three classes of human major histocompatibility (MHC) region of chromosome 6 (data from GenBank); G+C content is plotted as a moving average, and the window size is 100 kb, advanced by 10 kb each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

Finding isochores

- Isochores: homogeneous regions of %GC content

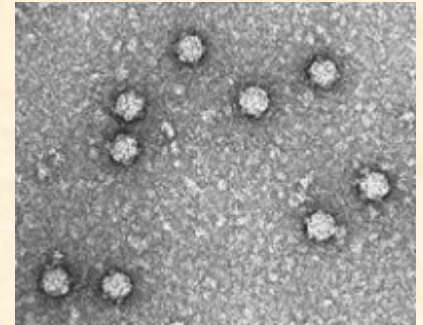
- Easy to find with windowed (100kbp) %GC calculation, from sequenced genomes.
- 3200 isochores characterised in the human genome, consistent with 5 levels (L1, L2, H1, H2, H3) found by staining/hybridisation.



GENOMOVÉ PROJEKTY

První genomové projekty - začátek éry genomiky

1977: první kompletní genom (virus) -
bakteriofág *phiX174* (5 386 bp)
Sanger et al., Nature 265, 687-695.



1995: první kompletní genom autonomního organismu -
bakterie *Haemophilus influenzae* (1.38 Mb)
Fleischmann et al., Science 269, 496-512.



1996: první eukaryotický organismus -
kvasinka *Saccharomyces cerevisiae*
Goffeau et al., Science 274, 563-567.



Genomové projekty



1996: *Saccharomyces cerevisiae* (12 Mb/ 6548 genes)

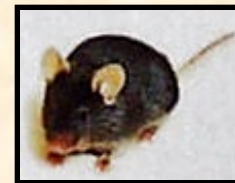


1998: *Caenorhabditis elegans* (97 Mb/19 099 genů)

2000: *Drosophila melanogaster*
(137 Mb/13 500 genů)



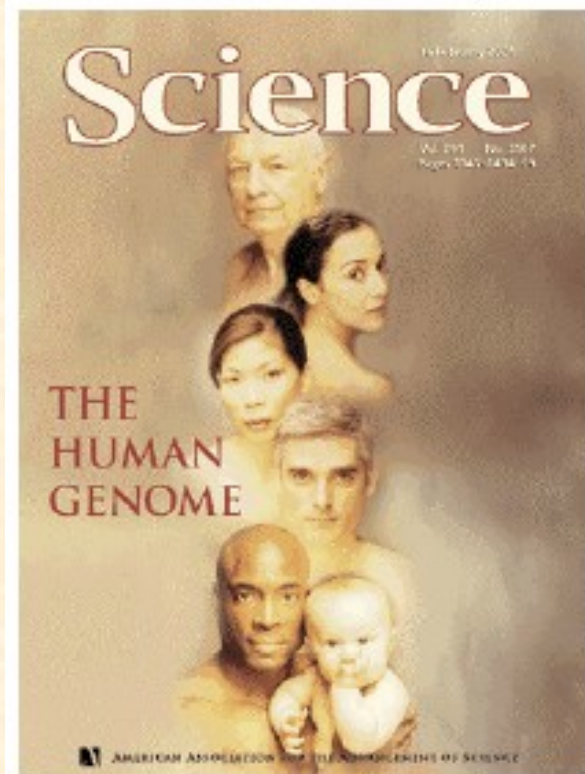
Mus musculus
(3 300 Mb)



2000: *Arabidopsis thaliana*
(125 Mb/25 500 genů)


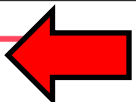




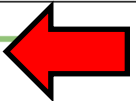


Sekvenování lidského genomu



Jak zjistit stav genomových projektů

Genomes OnLine Database (GOLD): <https://gold.jgi.doe.gov/>

Studies Metagenomic 639 Non-Metagenomic 22 861	Biosamples  Classification Ecosystems Host-associated 13 420 Engineered 2 499 Environmental 9 660	Sequencing Projects   Complete Projects 8 018  Permanent Drafts 33 489  Incomplete Projects 35 609  Targeted Projects 1 553	Analysis Projects Genome Analysis 48 641 Metagenome Analysis 5 632 Combined Assembly 105 Genome from Metagenome 1 501 Metatranscriptome Analysis 1 279 Single Cell (Screened) 1 680 Single Cell (Unscreened) 786 Transcriptome Analysis 0
Organisms  Organisms 73 094 Archaea 1 201 Bacteria 55 303 Eukarya 11 990 Viruses 4 473	Special Projects Type Strain Projects 5 329 GEBA Projects 2 517 HMP Projects 2 920	JGI Projects JGI Studies 1 112 JGI Biosamples 19 736 JGI Sequencing Projects 31 161 JGI Analysis Projects 19 290	Projects with Genbank Data Seq. Projects 42 590 Archaeal Projects 565 Bacterial Projects 35 744

Genbank - sekvence DNA, veřejně přístupná

SRA -sequence read archive - NGS data

SWISS-PROT - sekvence proteinů, popis funkce, struktury domén

dbEST - sekvence cDNA z různých organismů

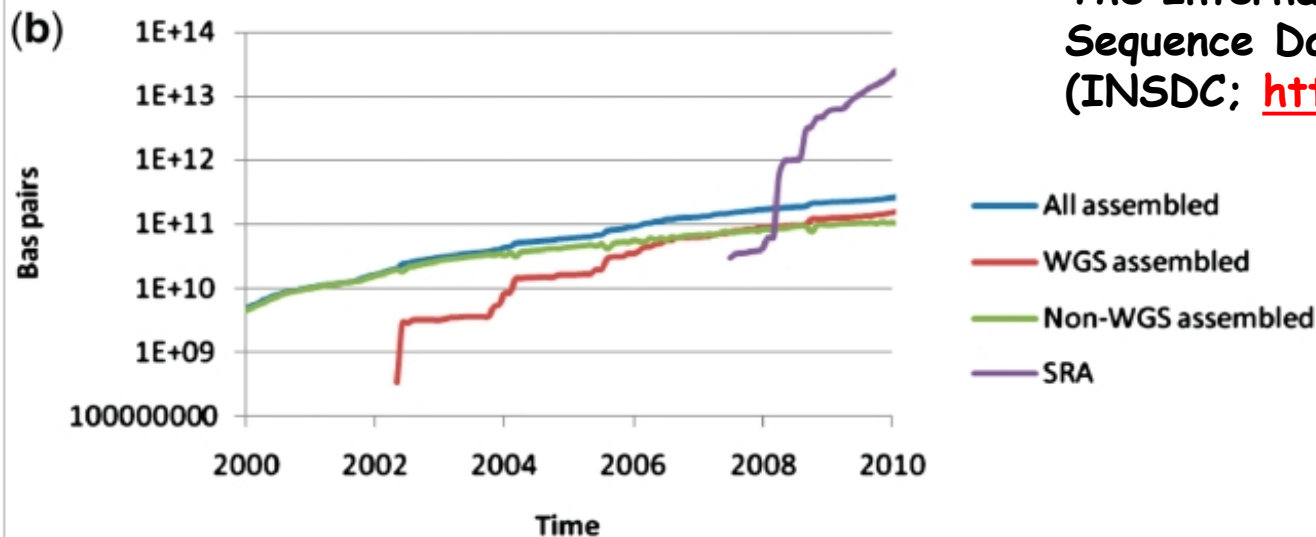
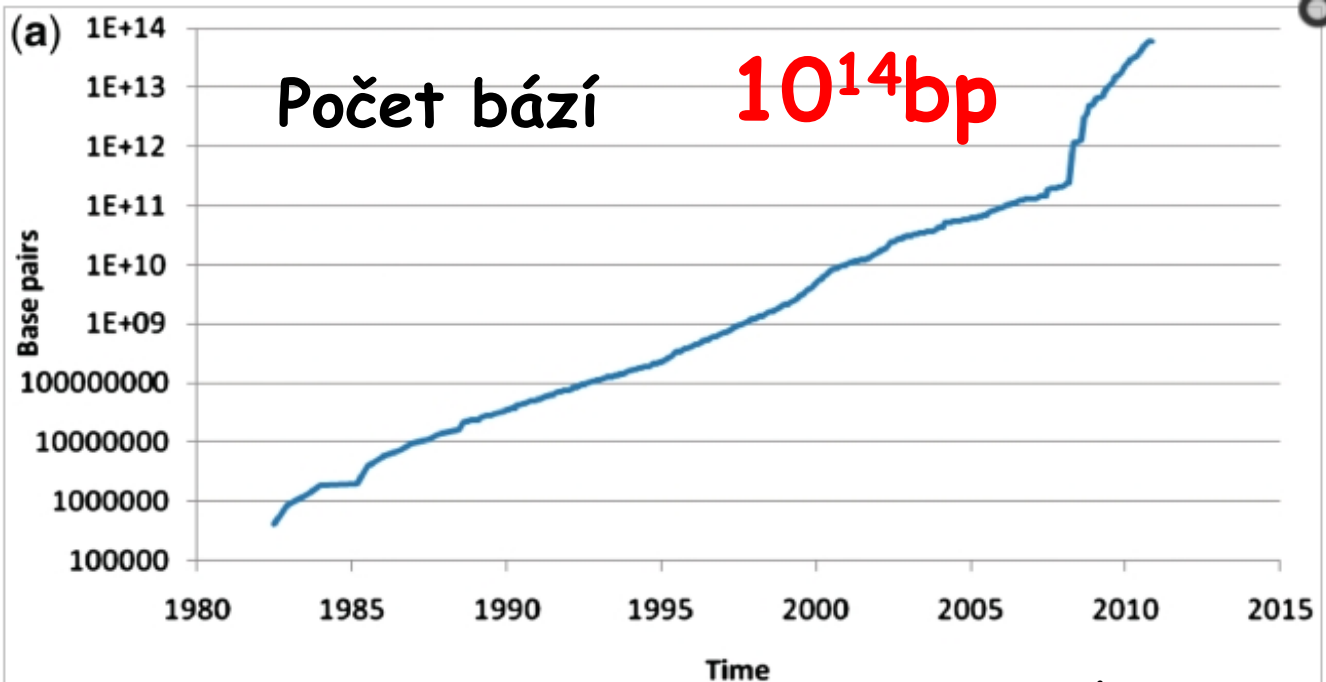
REBASE - restriktázy a metylázy

PEDANT - kompletní nebo parciální sekvence genomů

PDB - 3D koordináty makromolekulárních struktur

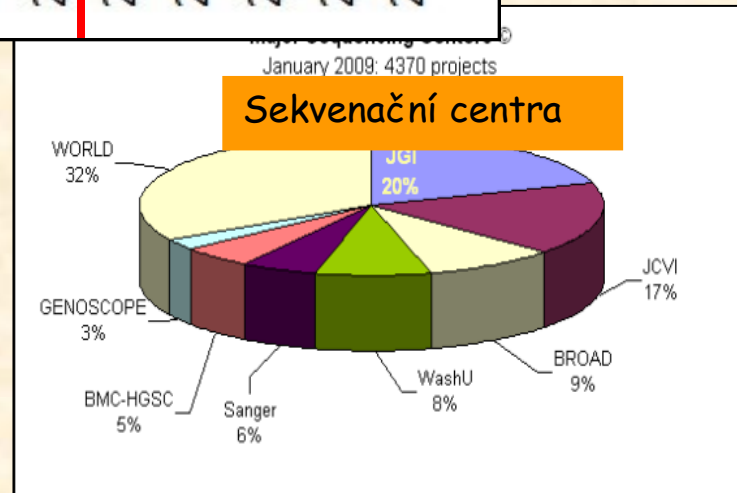
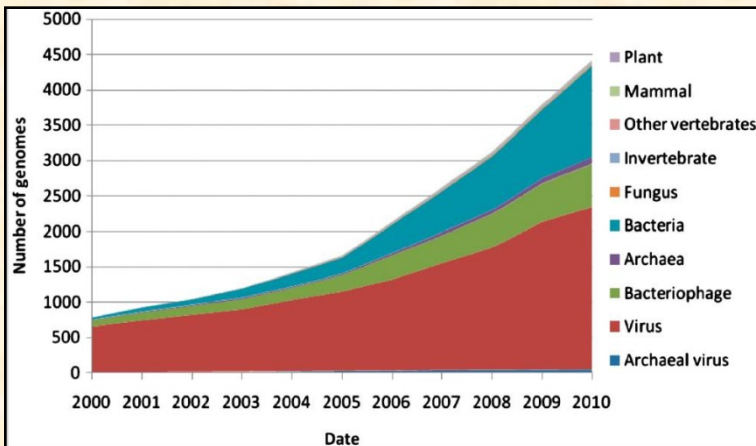
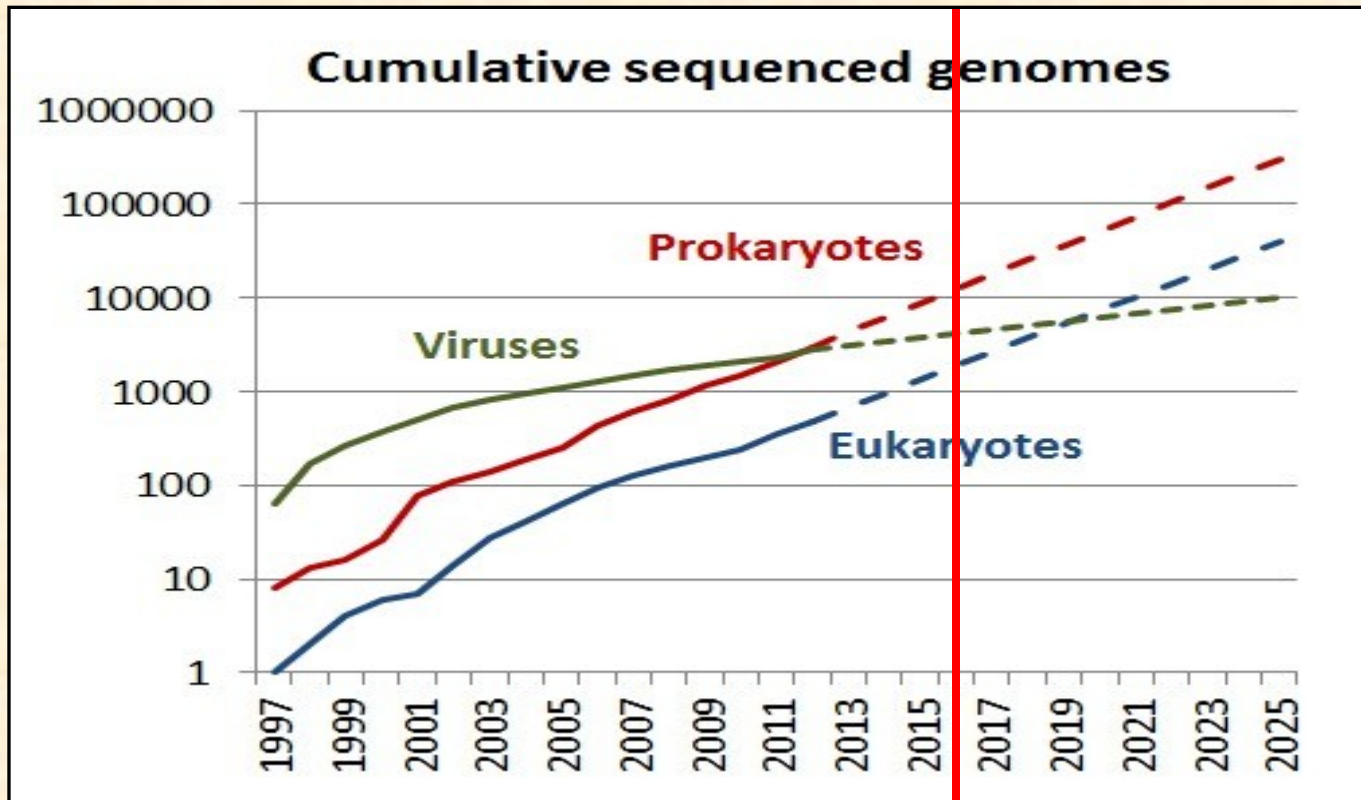
REPBASE - repetitivní sekvence

Objem dat v databázích roste exponenciálně



The International Nucleotide
Sequence Database Collaboration
(INSDC; <http://www.insdc.org>)

Genomové sekvenační projekty (počet genomů)



Prudký pokles ceny sekvenování genomu (člověk)

Cost per Genome

Které jiné
zboží zlevnilo
téměř
1 000 000x ???

