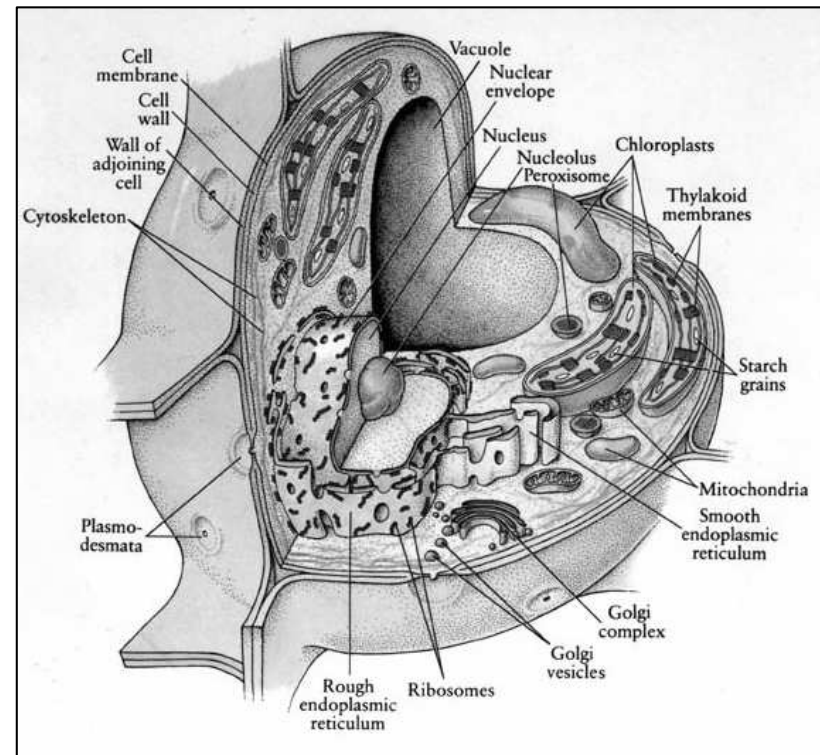
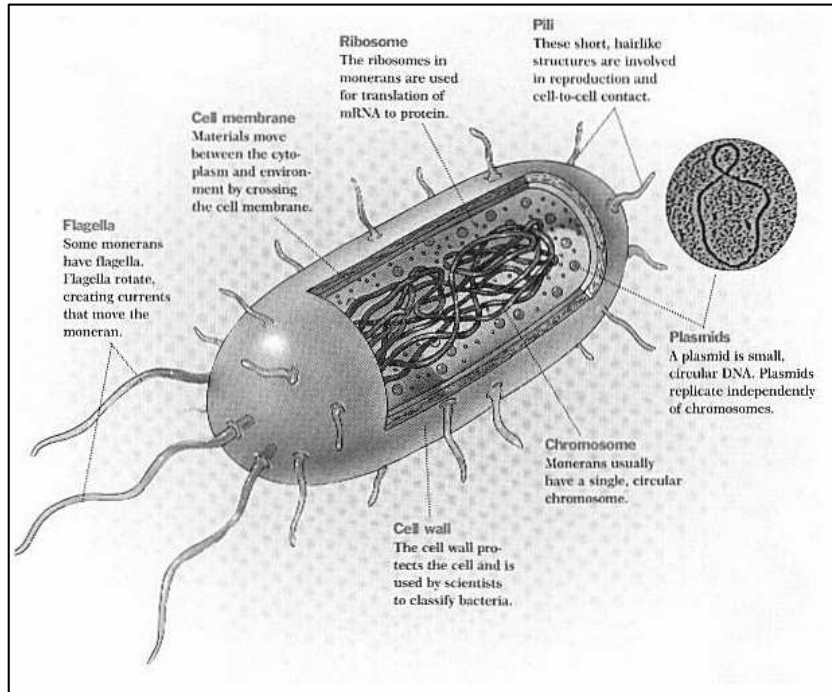


# Obsah

- Velikosti genomů u prokaryot
- Minimální genom a miniaturizace genomu
- Velikost genomů u eukaryot a paradox hodnoty  $C$
- Globální zvýšení velikosti genomu, polyploidizace
- Negenová DNA
- Repetitivní charakter eukaryotického genomu
- Regionální zvětšení genomu
- Distribuce genů v genomu
- Evoluce chromosomů
- Přestavby a pořadí genů
- GC obsah u bakterií a chirochory
- Mozaiková struktura genomu obratlovců, izochory
- Genomové projekty

# Definice genomu

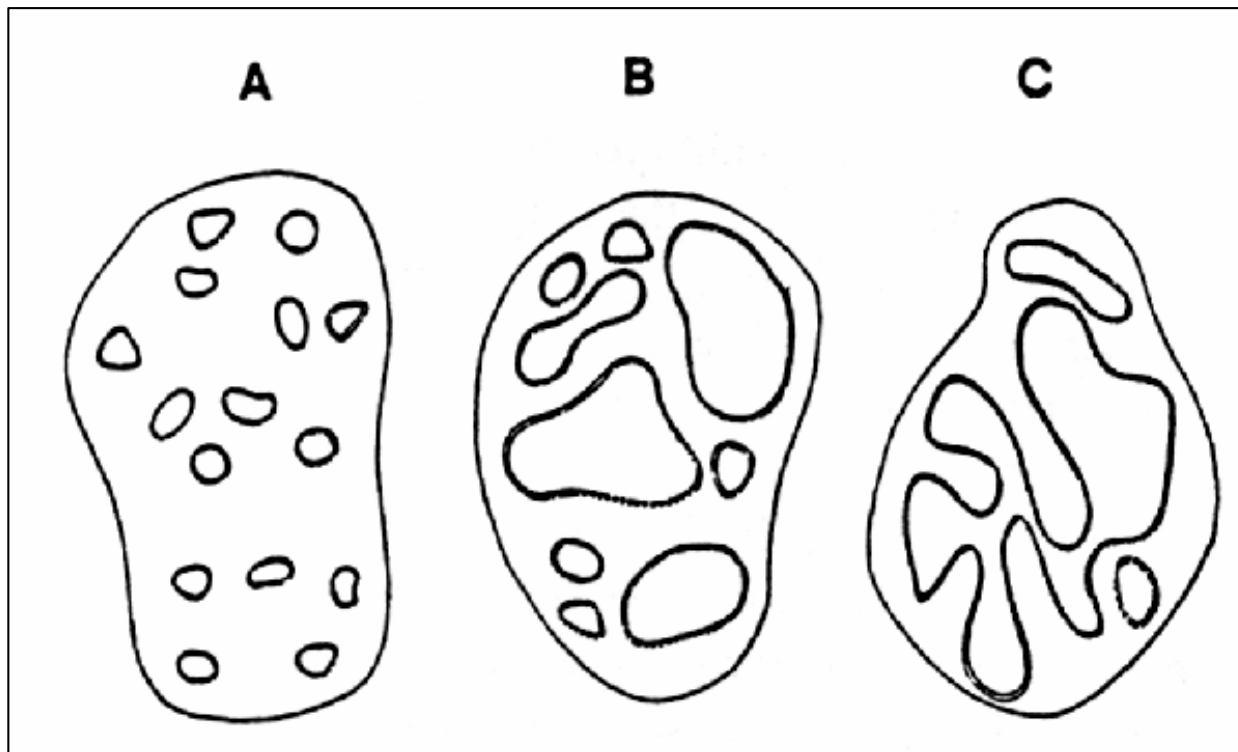
- celková genetická informace organismu
- u autonomních forem života vždy v DNA
- **prokaryota**: cirkulární chromosom + plazmidy
- **eukaryota**: chromosomy v jádře, mitochondrie a chloroplasty



# První genomy - vznik fúzováním malých kružnic DNA

(Trifonov 1995)

- první genomy: lineární nebo cirkulární?
- malé kružnicové DNA genomy, disperzní genom
- fúzování, geny jako autonomní DNA,
- počty kopií statisticky stejné - podobné přenosům plazmidů



Fáze:

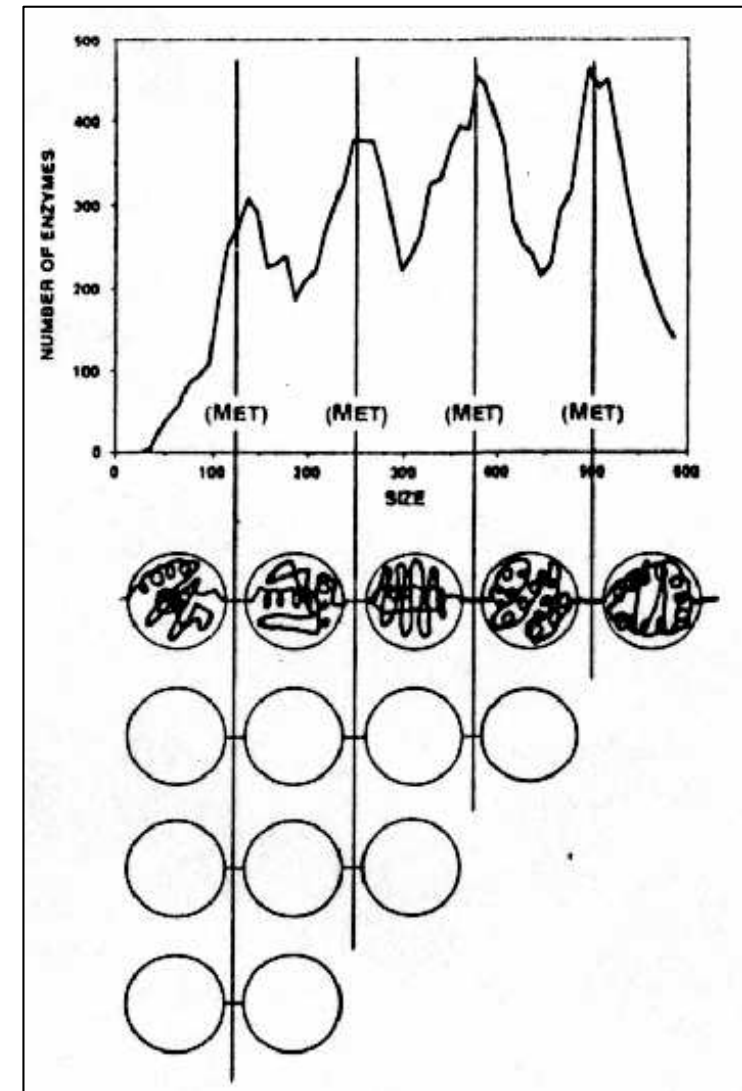
- A. pregenomická
- B. rekombinační
- C. genomická

# První genomy - vznik fúzováním malých kružnic DNA

(Trifonov 1995)

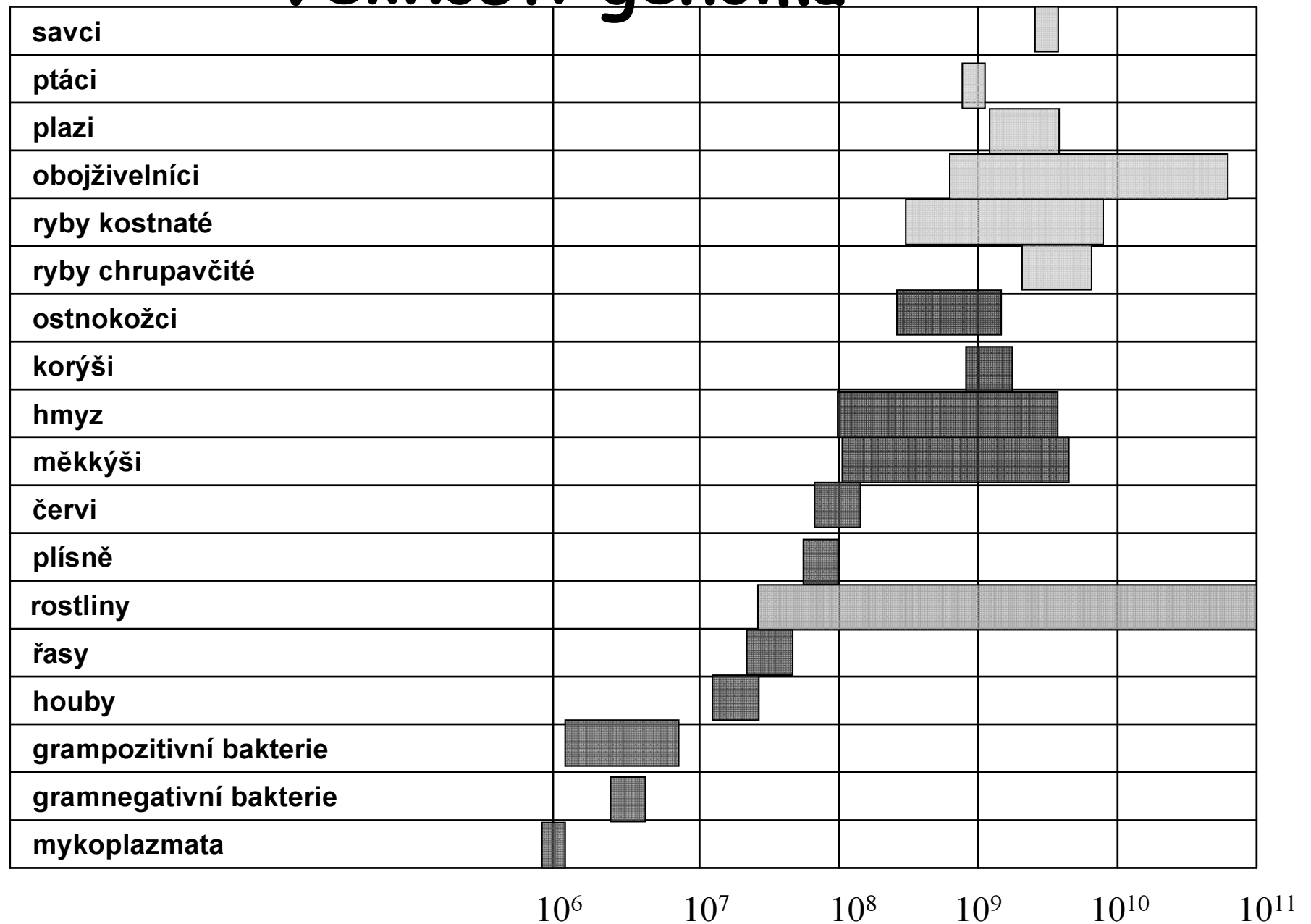
## Důkazy:

- periodičita délek proteinů (123AK u prokaryot, 152 AK u eukaryot)
- nejsnazší cirkularizace 350-450bp
- periodičita výskytu Met
- pozůstatkem extrachromosomální DNA - 10% genomu
- mobilní elementy, lyzogenní fágy, genomy organel, horizontální přenos
- replikace genů a satelitů prostřednictvím eccDNA



VELIKOST GENOMU

# Velikosti genomů



# Paradox hodnoty C

## Definice:

- velikost genomu organismu není v korelaci s komplexitou organismu (*H. sapiens* 200x menší genom než *Amoeba dubia*)
- obsah kódující DNA podobný, příčinou je nekódující DNA, sobecká DNA



180 Mb

*Drosophila melanogaster*

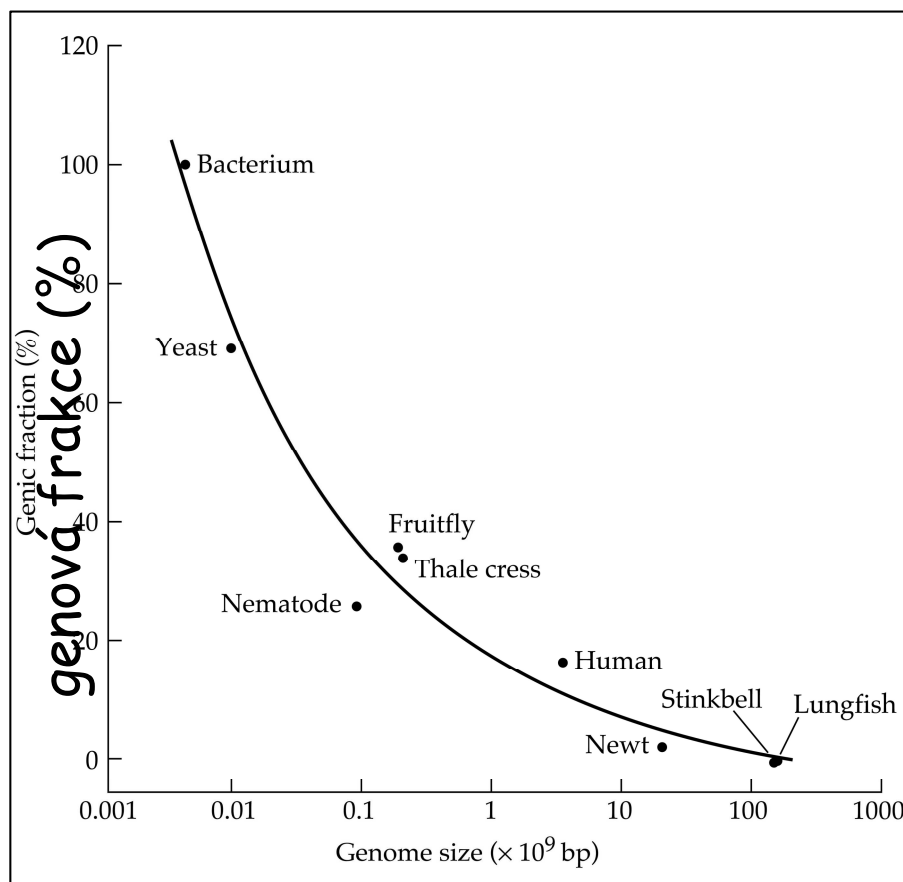


18,000 Mb

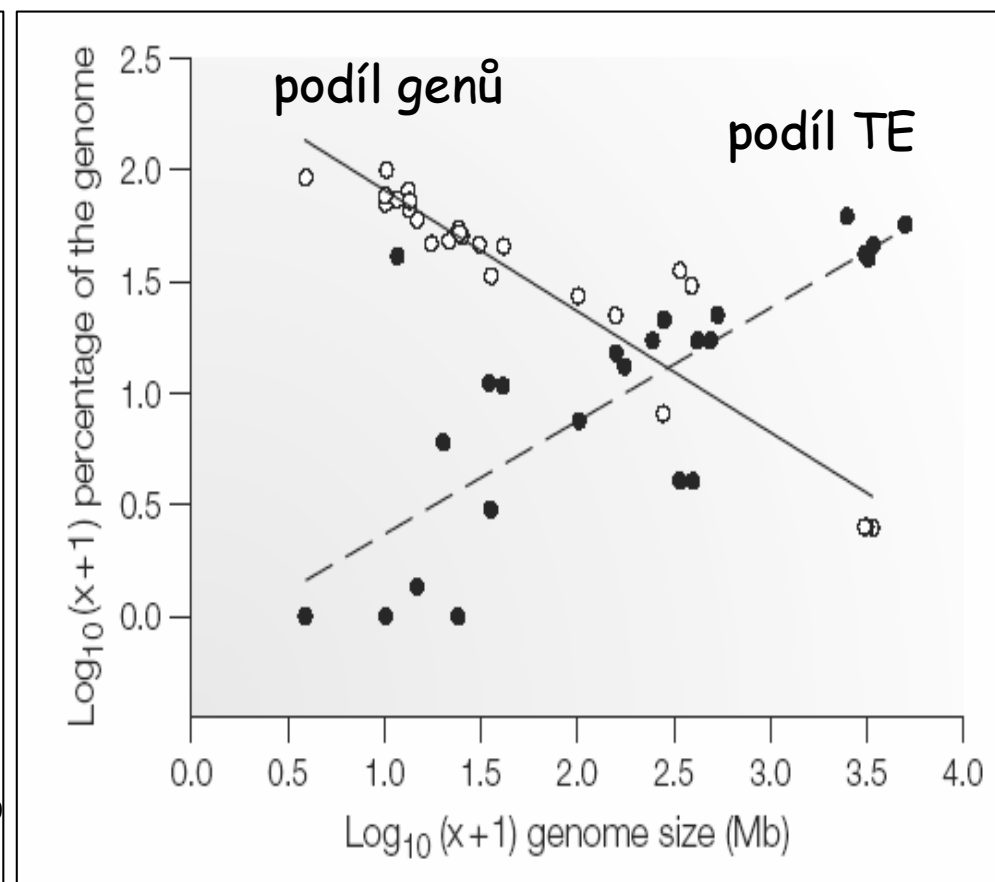
*Podisma pedestris*

Skutečnost, že genomy eukaryot se liší až 80,000-krát -  
snad největší výzva pro genomové holisty

# 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:

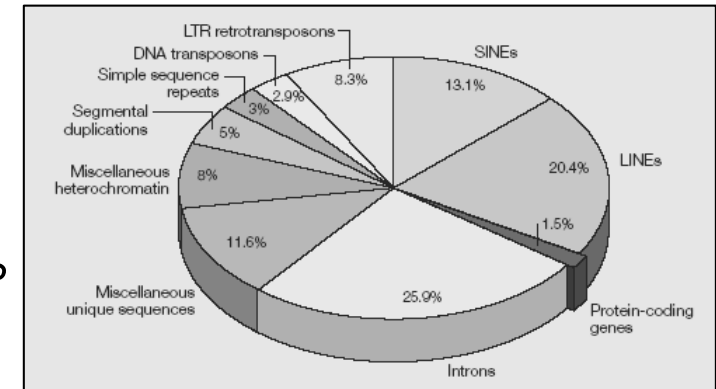
- klasické, retropseudogeny,
- 19000 člověk, 14000 myš, 51 kur, 33 kvasinka, 176 drosophila

## Mobilní elementy:

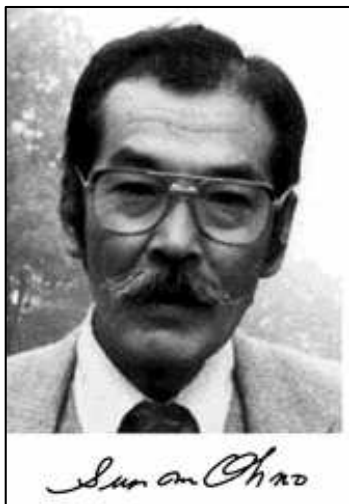
- LTR, nonLTR - SINE, LINE, DNA transposony-MITE

## Numt, Nupt:

- inserce promiskuitní DNA



# Velikosti genomů a zbytečná DNA



„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”

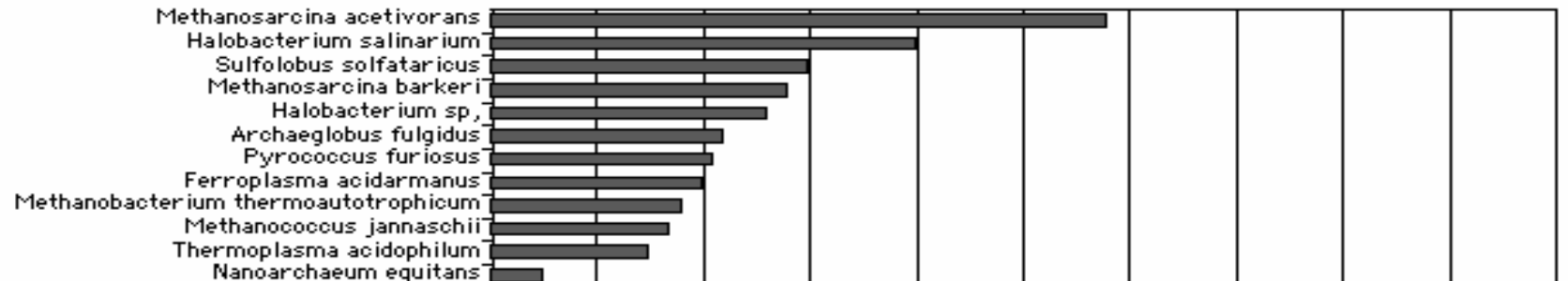
- Maynard Smith

# Velikosti genomů prokaryot

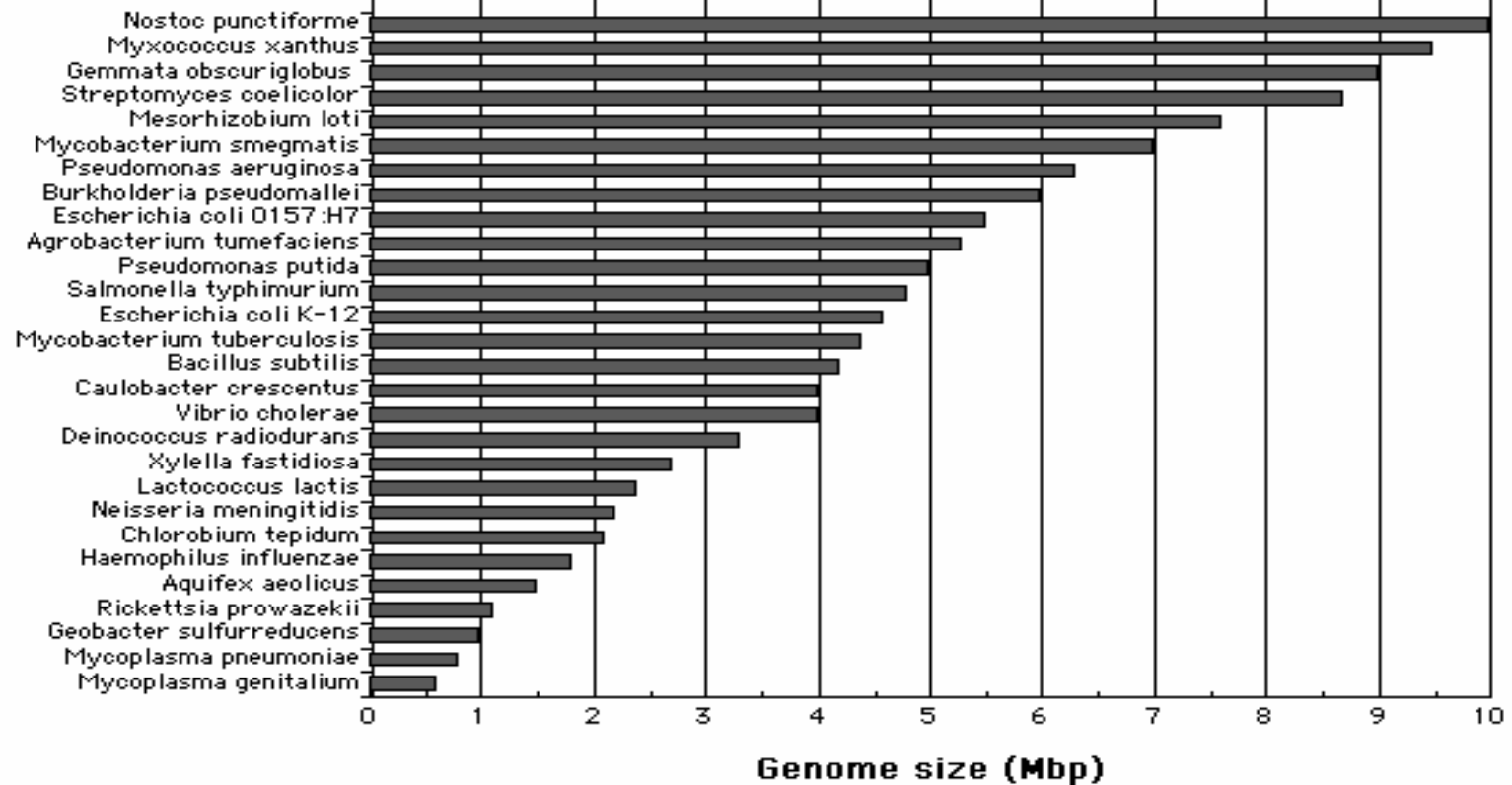
- **Bakteriální genom:  $6 \times 10^5 - > 10^7$  (=0.6-10Mb)**  
nejmenší známý: *Mycoplasma genitalium* (480 genů kódujících proteiny, 3 rRNA geny, 37 tRNA genů)
- **Velikost genomů prokaryot je zhruba úměrná počtu genů**
- **Procesy ovlivňující velikost bakteriálního genomu:**  
Genová duplikace,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

# Velikosti genomů prokaryot

## Archaea:



## Bacteria:



# EXTRÉMNÍ GENOMY

# Minimální velikost genomu



## Minimal Genome Project

Press Release

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

Co je život?

Strategie: 1. odstranění či inaktivace genů  
2. syntéza minimálního genomu

2105-2107 (1999)

TIGR  
Databases

What's  
New

About  
TIGR

TIGR  
Faculty

# Minimální velikost genomu

„One day a scientist will drop gene number 297 into a test tube, then number 298, then 299... and presto: what was not alive a moment ago will be alive now. The creature will be as simple as life can be. But it will still be life. And humans will have made it, in an ordinary glass tube, from off-the-shelf chemicals...”

- Hayden, 1999:

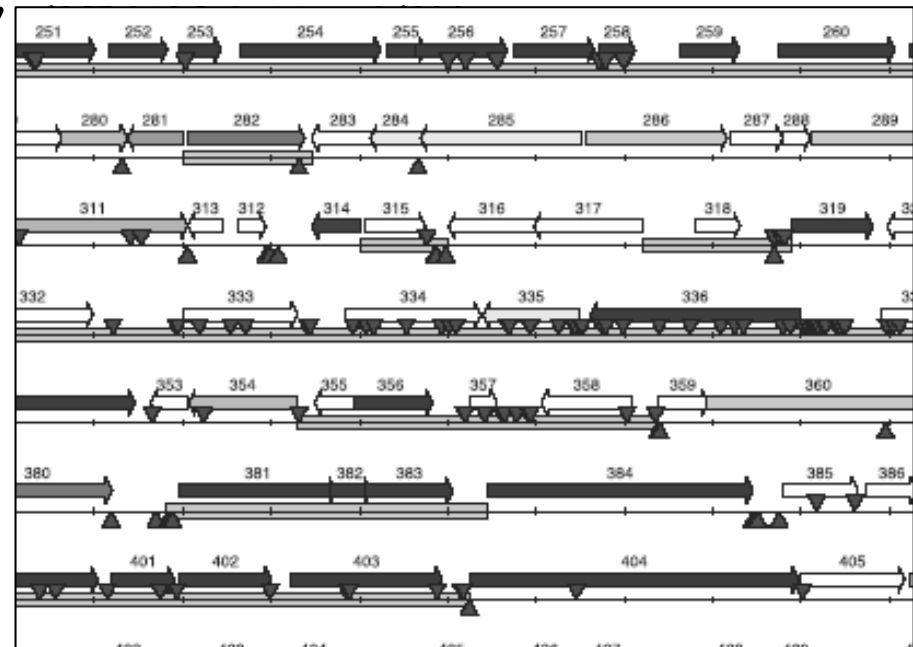
# “Minimální” genom - Mycoplasma

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

## Global Transposon Mutagenesis and a Minimal Mycoplasma Genome

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

*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

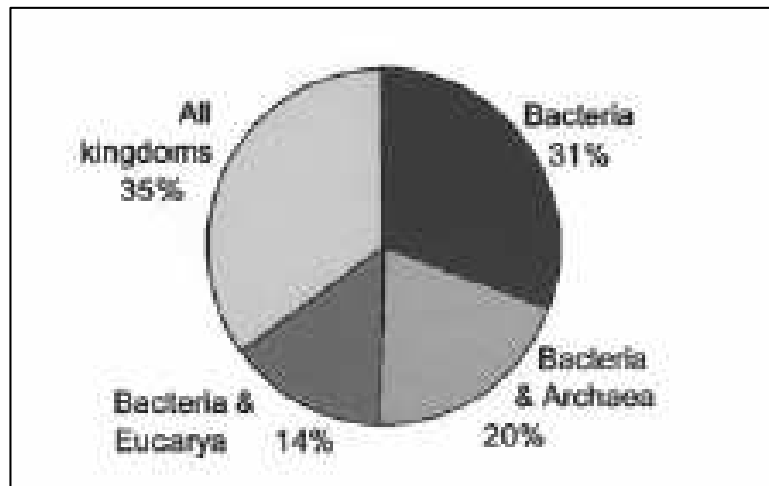


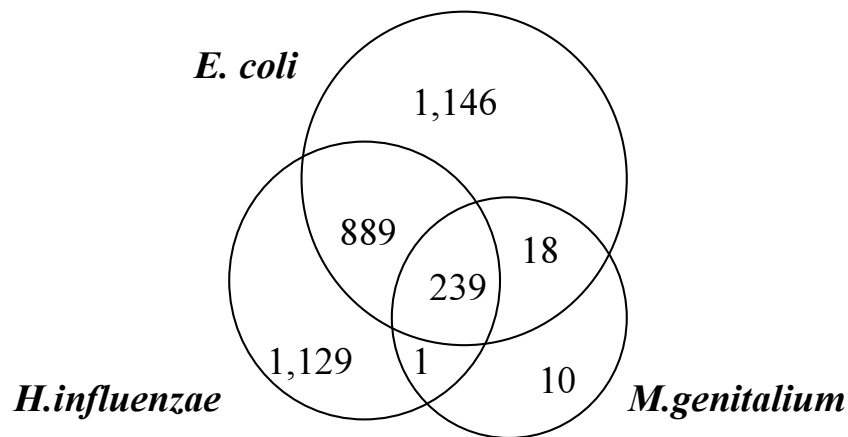
Table 2. *B. subtilis* essential genes

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

# Minimální velikost genomu: (a) Analytický přístup

- (Mushegian a Koonin, 1996):

Srovnání kompletních bakteriálních genomů:



- Translation
- Replication
- Transcription
- Recombination and DNA repair
- Chaperone-like proteins
- Anaerobic metabolism
- Lipid or cofactor biosynthesis
- Transmembrane transporter
- Some unknown function

[Překrývající se ortologní geny (239)] + [přesun neortologních genů] – [geny specifické pro parazitické bakterie nebo funkční redundance] = 256 genes

# Minimální velikost genomu: (b) Experimentální přístup

- (Itaya, 1995):

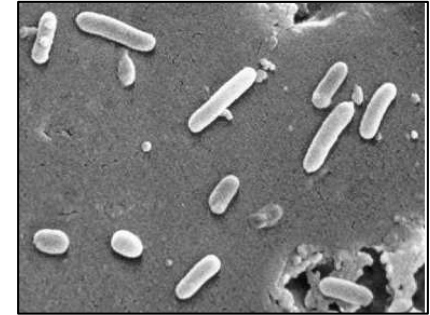
Knock-out 79 náhodně vybraných genů *Bacillus subtilis*:

- pouze 6 letálních, 73 je postradatelných → 7.5% (6/79) genomu je nepostradatelných
- *B.subtilis* genome:  $4.2 \times 10^6 \text{bp} \times 7.5\% = 3.2 \times 10^5 \text{bp}$   
Průměrná velikost genu je 1.25kb, takže minimální velikost genomu  $\approx 254$  genů

# 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 *N.p* Sites in genome

AATGAC <sup>h</sup> (STRR2)	
AATGACA	69
AATGACT	63
AATGACC	39
AATTCCC (STRR4)	
AATTCCC	41
AATGCCC	37
AATTACG (STRR5)	45
A <sup>d</sup> TCCCC (STRR1)	
ATCCCC	39
AATCCCC	19
AGTCCCC	15
AGCAGGGG (STRR6)	29
AAAATTC (STTR7)	13

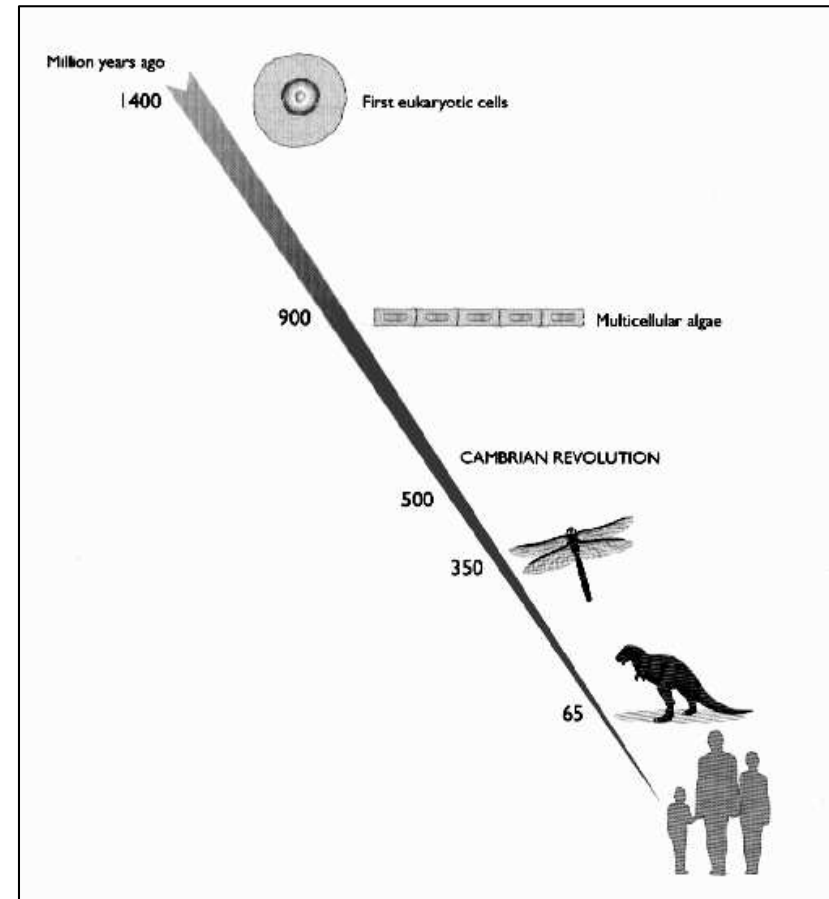
# Velikosti genomů eukaryot

- Eukaryotický genom:  $8.8 \times 10^6 \sim 6.9 \times 10^{11}$   
(8.8Mb - 690Gb)  
Nejmenší známý: *Saccharomyces cerevisiae* a jiné houby
- Velikosti eukaryotických genomů NEjsou úměrné počtu genů nebo komplexitě organismu

# ZMĚNY VE VELIKOSTI GENOMU

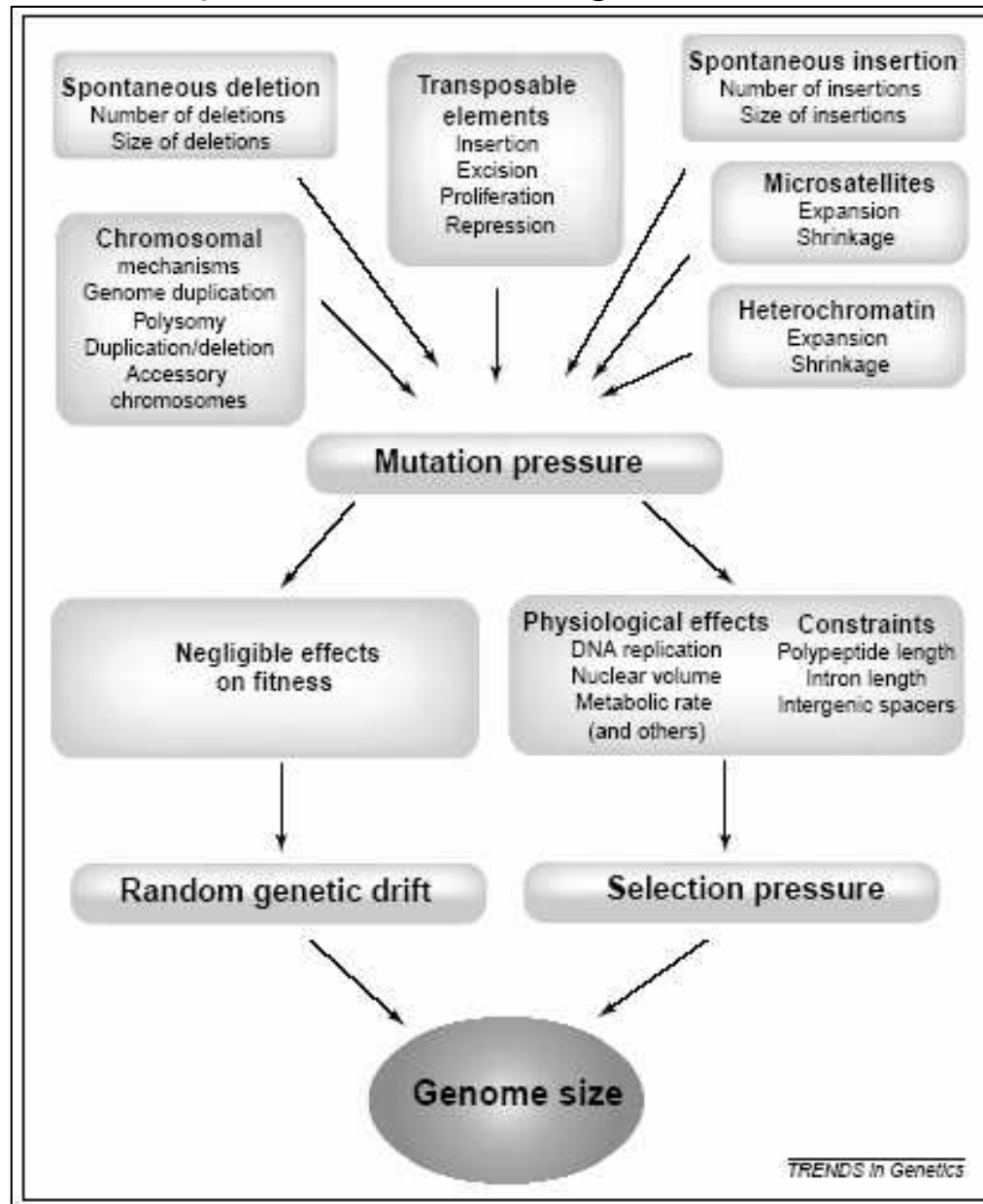
# Zvětšení genomu

- celkové zvětšení:  
polyploidizace (duplikace celého genomu)
- duplikace části genomu, zmnožení počtu chromosomů
- duplikace genů a skupin genů
- expanze heterochromatinu
- amplifikace transpozonů a retroelementů
- inserce virové DNA
- inserce 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?

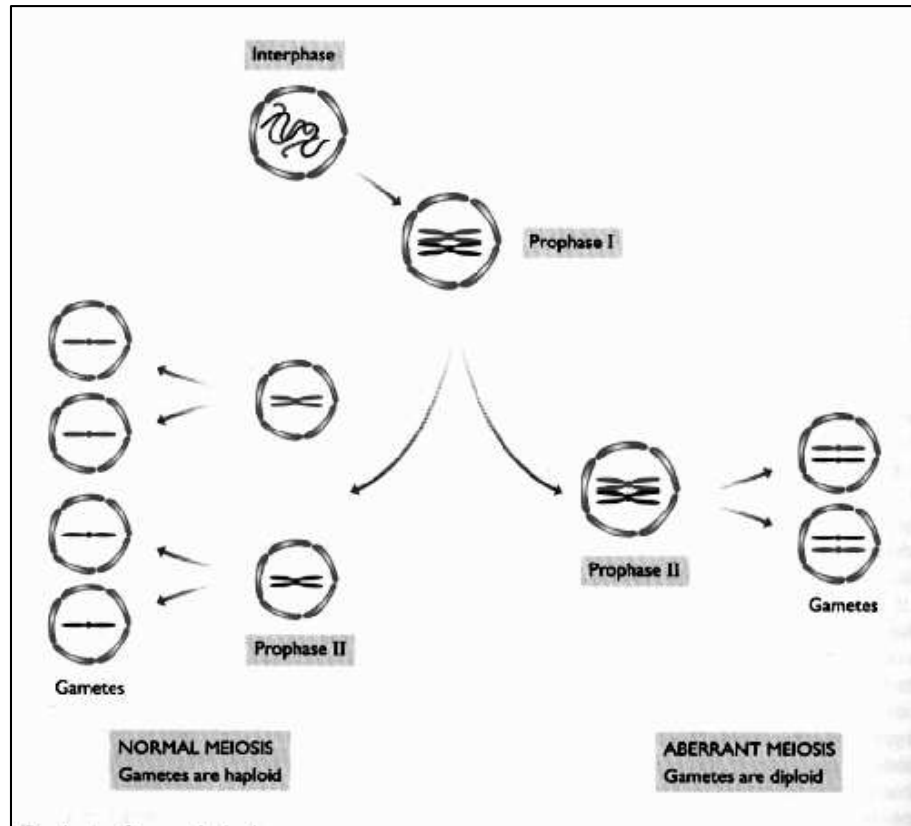
## Plynulost:

- 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

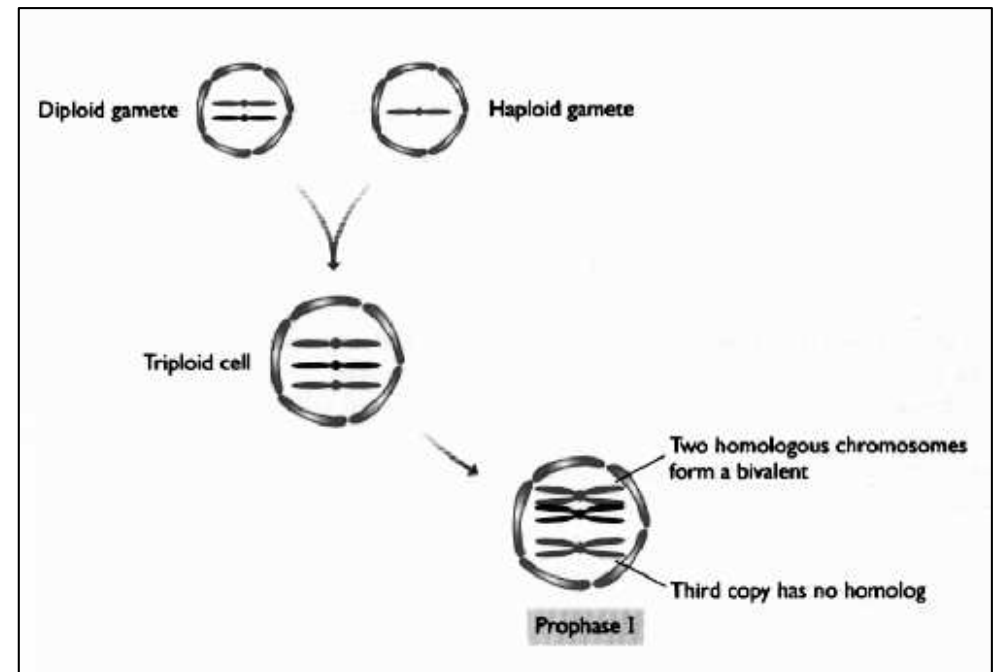
## Skoky (kvantové změny):

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

# Růst velikosti genomu: Polyploidizace



Chyba v meioze vede k  
diploidním gametám



Fúze diploidní a haploidní  
gamety vede k triploidnímu  
jádru, triploidní organizmus je  
však sterilní

# Celkové zvětšení genomu

- **Polyploidizace (duplikace celého genomu):**  
Autopolyploidie: multiplikace jedné základní sady chromosomů, jeden organismus, chyba při meioze (růže...)  
  
Allopolyploidie: kombinace geneticky odlišných sad chromosomů (pšenice...)  
  
Kryptopolyploidie: dávná polyploidizace, organismus se jeví jako diploidní, přestavby - delece, translokace..., kvasinka, obratlovci, obilniny, odhalení až sekvenací
- **Duplikace části genomu:**  
chromosomů - polysomie (trisomie), většinou letální, genová imbalance, evolučně irelevantní genů nebo skupin genů

# Výhody a nevýhody polyploidů

Definice, vznik - chyba nebo indukce, živočichové - partenogeneze, rostliny - fertilní ( $4n$ ,  $8n$ ) vs. sterilní hybridy ( $3n$ , tetra+di)

Skoková (kvantová) změna velikosti genomu - nadějně monstrum musí najít partnera

Schopnost obsazovat nová prostředí, převládá v evoluci rostlin, i u obratlovců

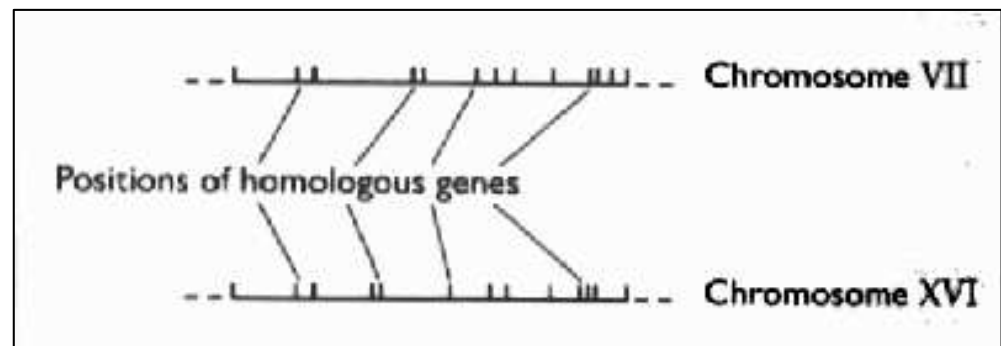
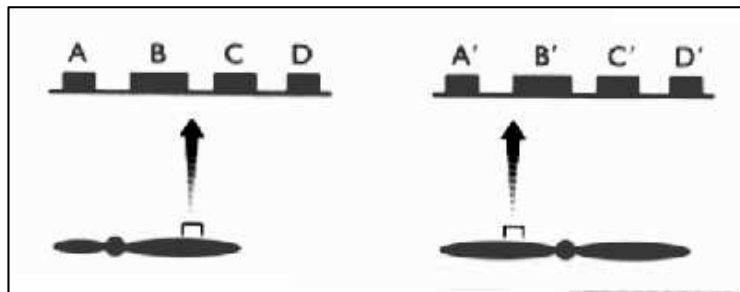
Noví polyploidi - nestabilní - radikální přestavby genů - rychlá diverzifikace - speciace (změněný počet chromosomů vede k reprodukční izolaci)

Multiplikace klíčových genů v regulaci vývoje - multiplikace klastrů homeotických genů vedla k enormní diverzitě druhů ryb

# 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

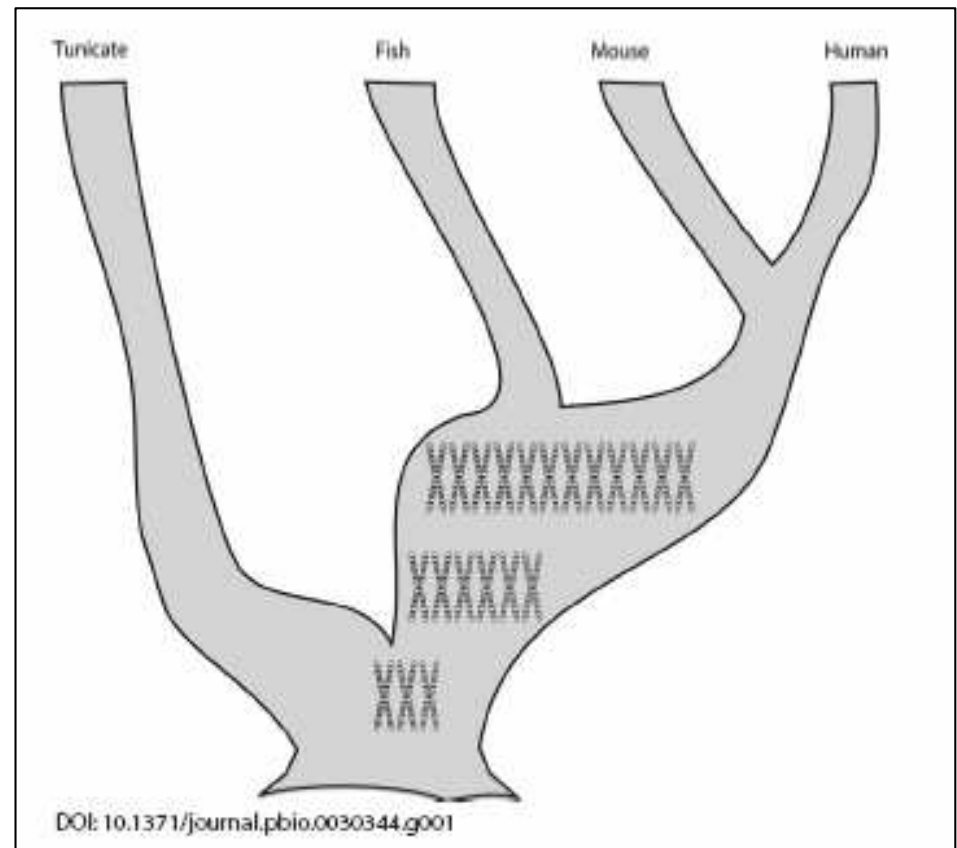


# Polyploidizace v linii obratlovců

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

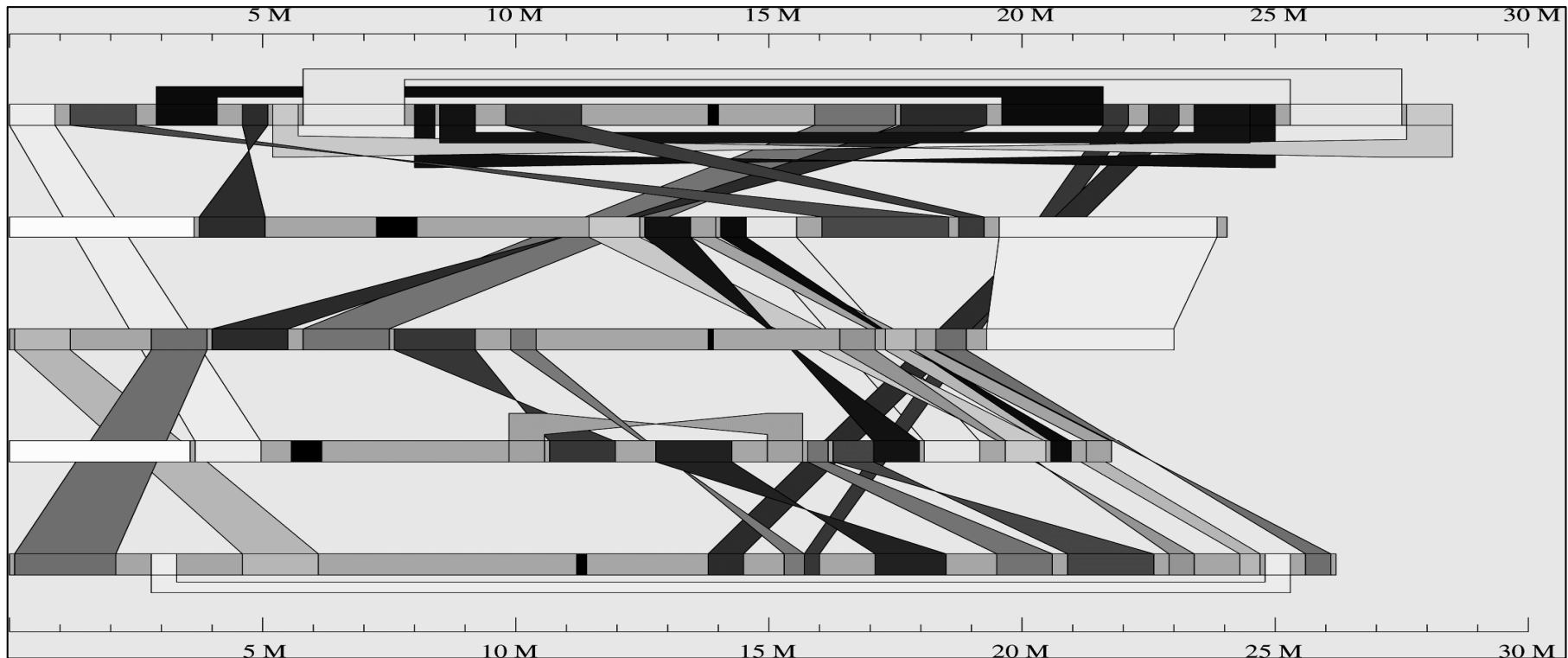
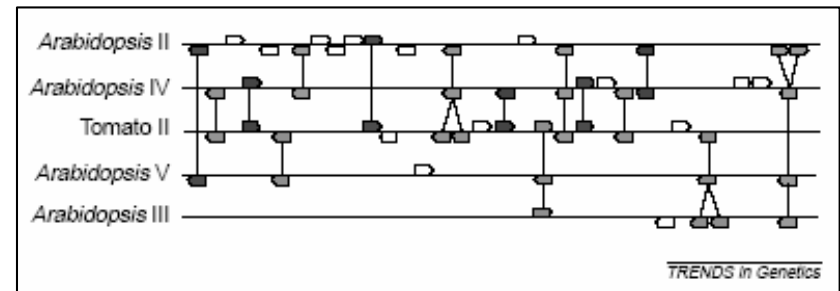
→ Duplikace genomu před 450 mil let napomohla vzrůstu komplexity a diverzifikaci obratlovců



# Polyploidizace u rostlin

## Arabidopsis:

- duplikace před 100-200 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





# Genomová obezita u rostlin - jednosměrný proces?

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

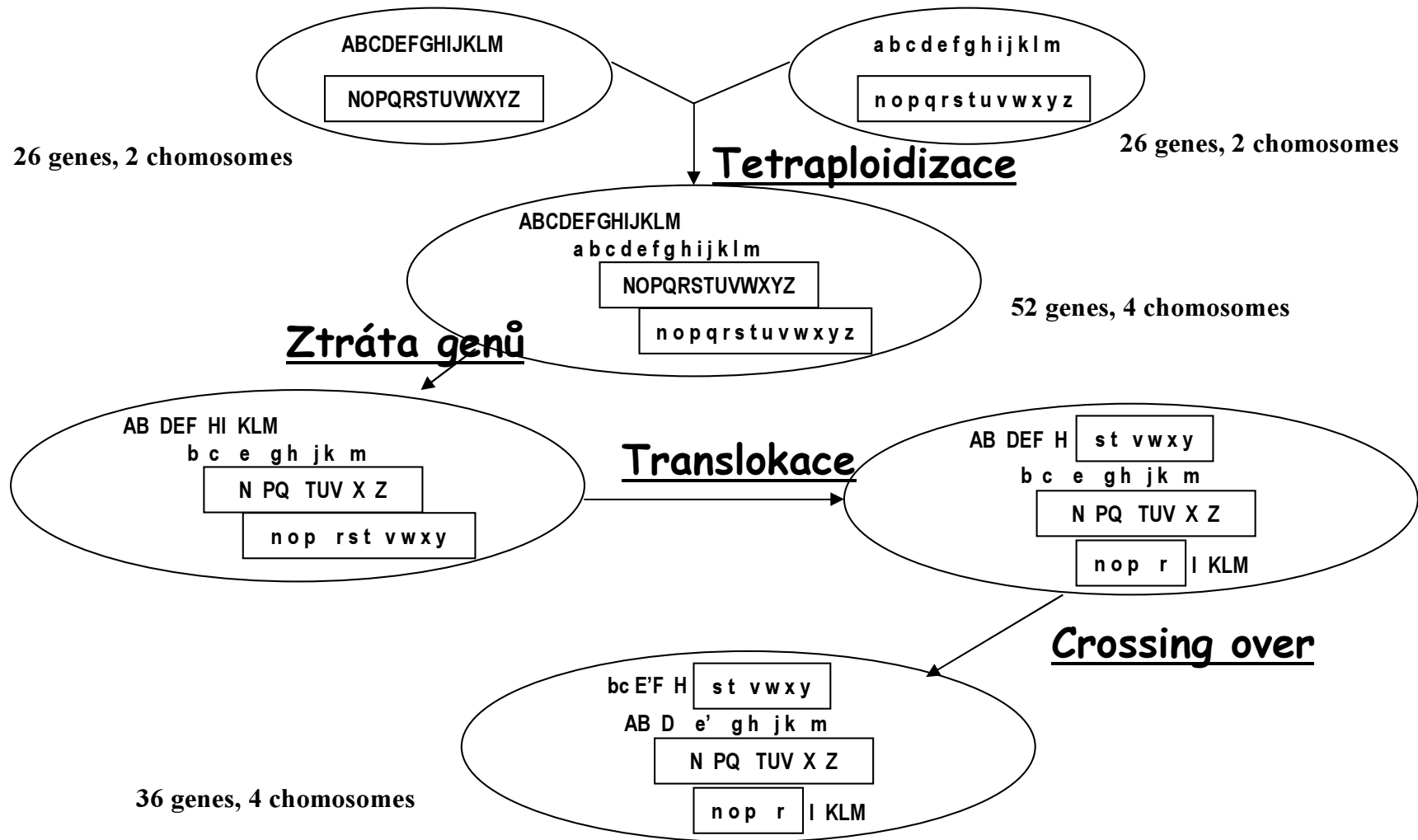
## Dva modely:

- pouze růst genomů
- růst i zmenšování genomů

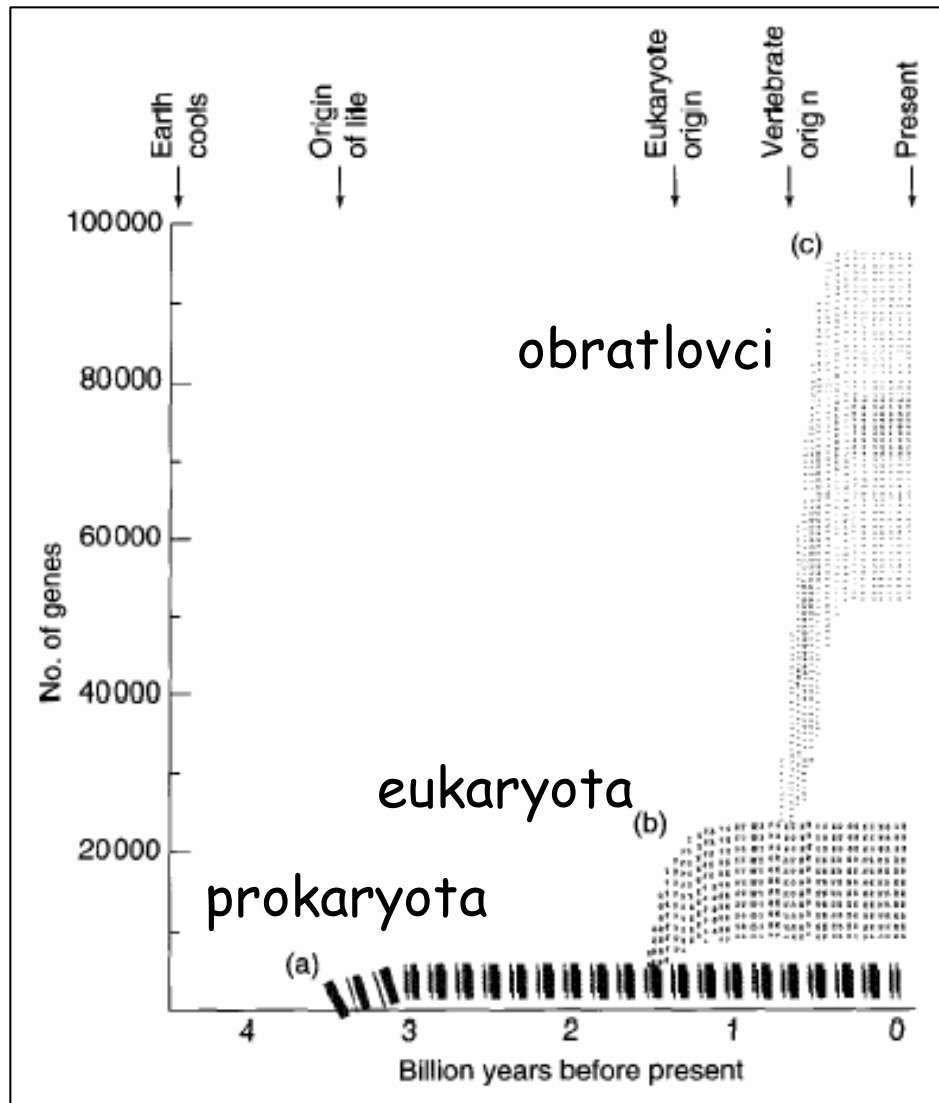
## Limity růstu genomů:

- fitness hostitele
- mechanismy potlačující amplifikaci - metylace
- mechanismy odstraňování repeticí - rekombinace
  - nonLTR u *Laupala* x *Drosophila* - rychlost delecí
  - BARE - 42x více solo LTR než vnitřních částí
  - rekombinace uvnitř i mezi elementy
  - selekce na kratší LTR, solo LTR nejdelší

# Počet genů a evoluce duplikovaného genomu



# Vzrůst počtu genů, potlačení šumu a biologická komplexita



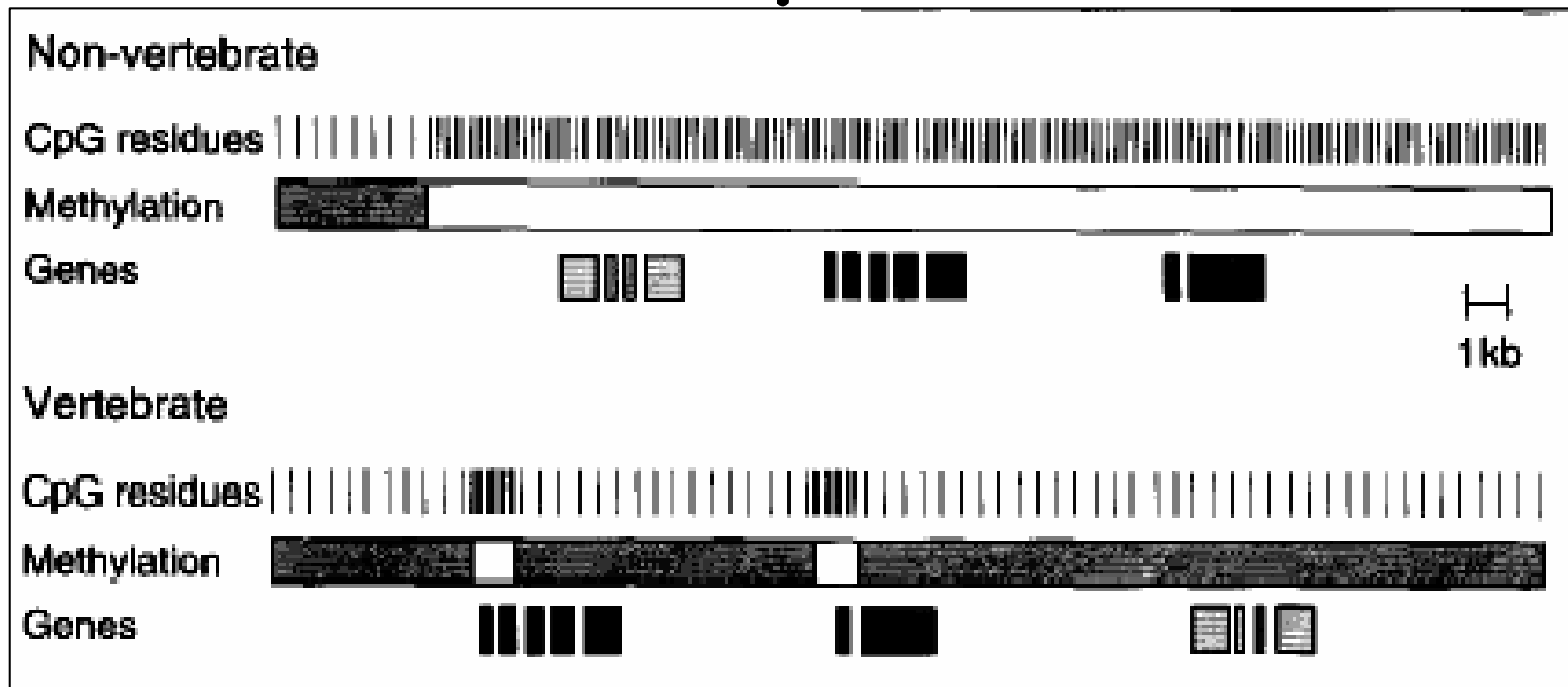
## Prokaryota/eukaryota:

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

## Obratlovci:

Potlačení šumu - metylace genů, mobilních elementů a duplikátů

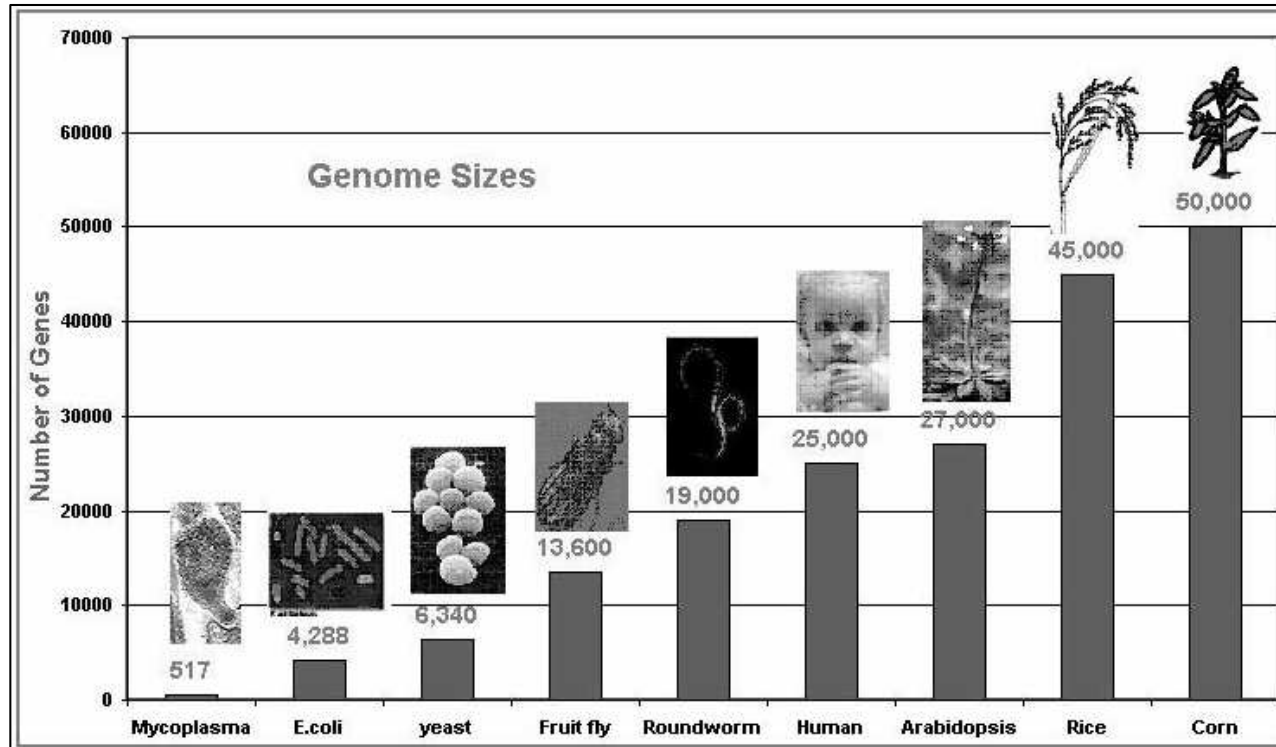
# Metylace u obratlovců - mechanismus potlačení šumu



## Obratlovci:

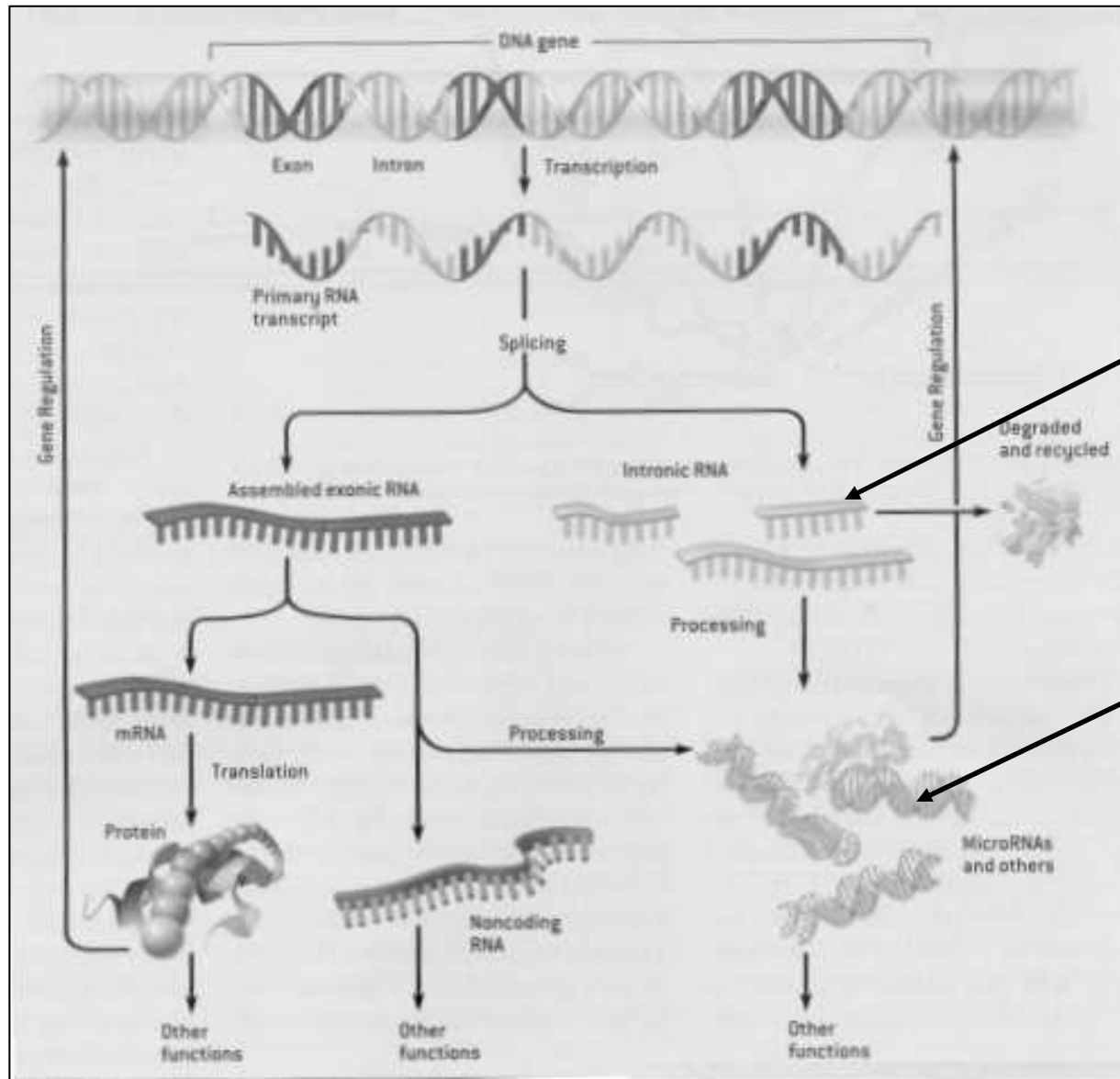
- geny jsou metylovány, CpG-ostrovy nemetylovány
- genové oblasti chudé na dinukleotidy CpG

# Komplexita vzniká snadno, kontroluje se ale obtížně



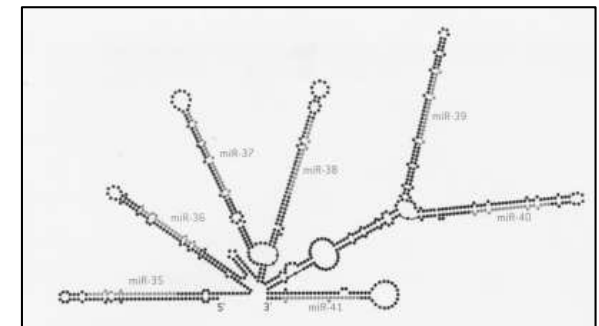
- Komplexita organismu nekoreluje s počtem genů kódujících proteiny (rýže má více genů než člověk)
- „Zbytečná“ DNA může kódovat RNA s různými regulačními funkcemi
- Geny-ostrovy v moři regulačních sekvencí

# Nekódující DNA: Regulační systém na bázi RNA?



**introny**

**MicroRNA**



# Regulační systém na bázi RNA

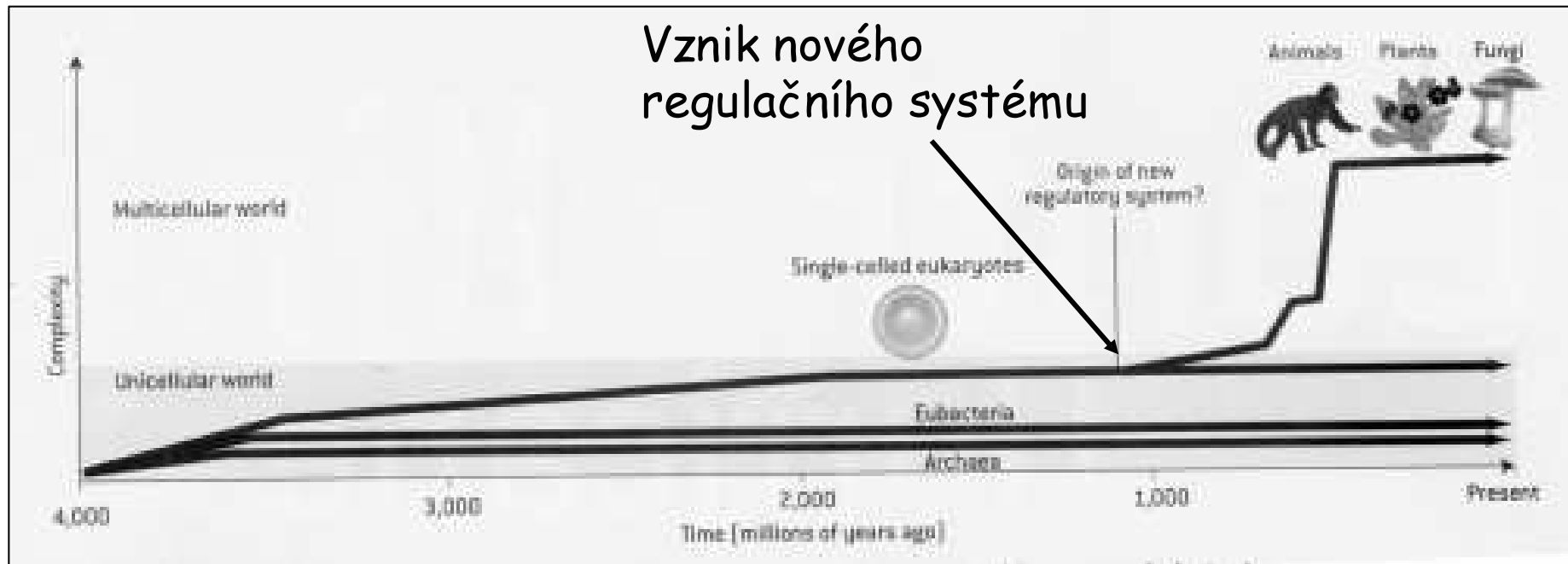
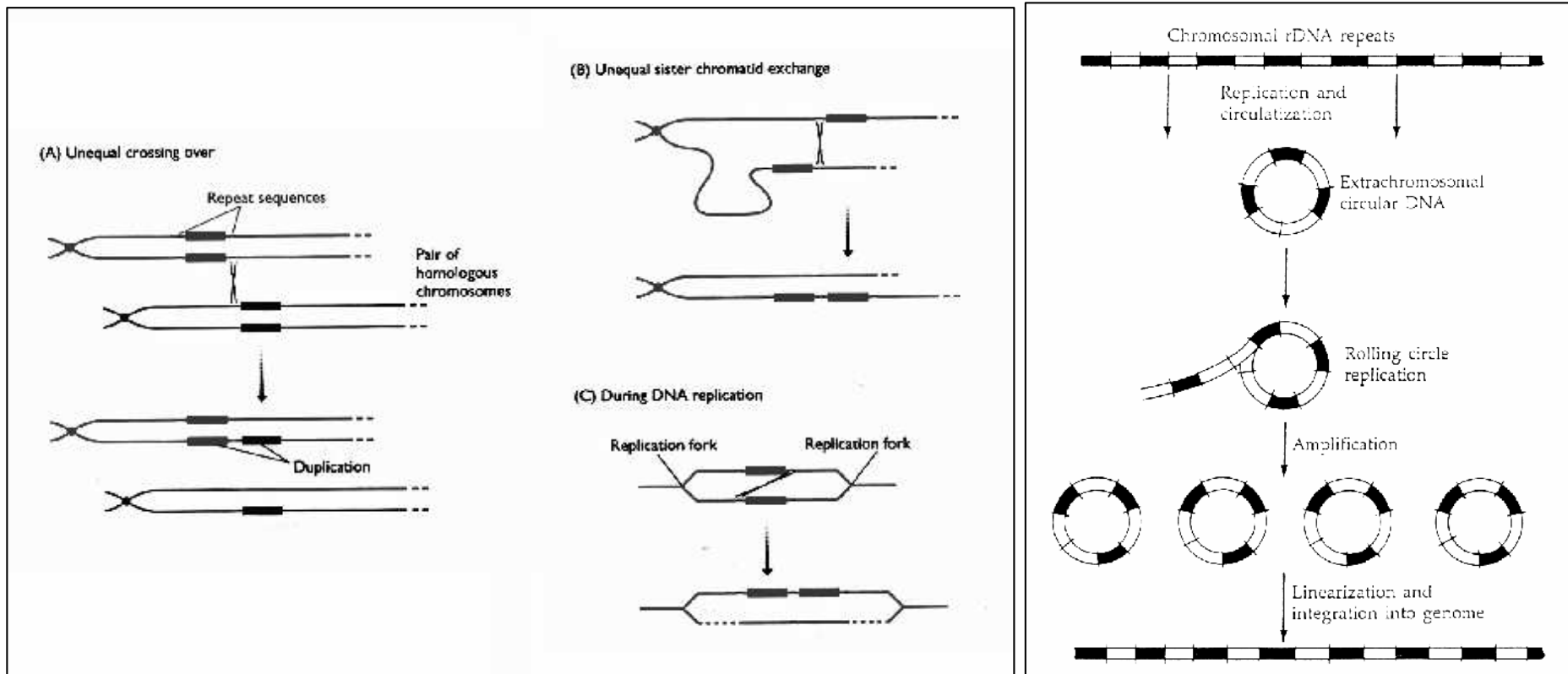


Figure 3 | **A simplified biological history of the Earth.** This graph is intended to present an overview. Some dates are still being debated and the abscissa ('complexity') has an arbitrary scale. Life appeared on Earth approximately 4,200 million years ago (mya), either arising as, or quickly streaming into, three main kingdoms — the eukarya, the bacteria and the archaea. Life remained unicellular, or at best colonial, for at least 3,000 million years. The common ancestor of the animals, plants and fungi is thought to have arisen approximately 1,200 mya, around the time at which the mitochondria entered the lineage through a rickettsial-like endosymbiont, an event that is postulated to have also brought with it type II self-splicing introns<sup>75</sup>. Whether or not these events were coincidental, the incidence of introns (and other non-coding sequences) correlates with the complexity of the organism after that point. In the Cambrian period (~520 mya), complex animal life exploded in an event known as the metazoan radiation, in which recognizable ancestors of all modern phyla appeared only in a single strata of rock<sup>78</sup>. What restrained the appearance of organized multicellular organisms for so long? Was it environmental or biochemical factors (such as oxygen tension and oxidative energy metabolism), or a primitive genetic operating system?

# 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:

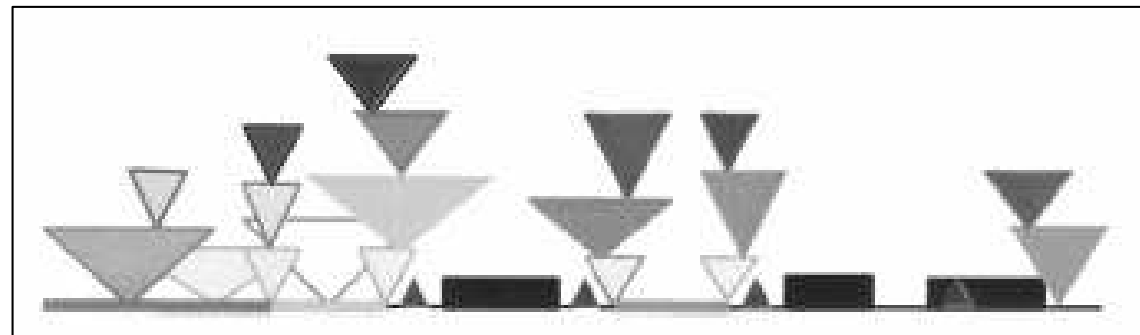
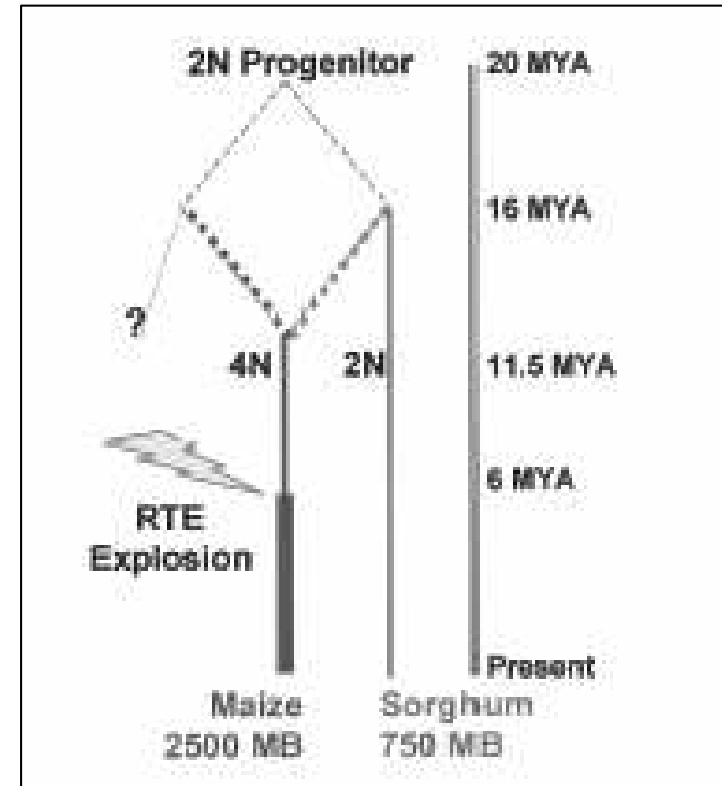
vliv nadbytečné DNA na fenotyp, vliv na velikost jádra a buňky, ochrana kódujících sekvencí před mutacemi, pufrování koncentrace regulačních proteinů

## Sobecká DNA (junk DNA):

Parazitické sekvence, mobilní elementy, fixace genetickým driftem, velikost genomu je tolerovatelné maximum závisající na ekologických a vývojových potřebách organismu

# Big Bang v genomu kukuřice

- polyploidizace
- amplifikace transpozonů a retroelementů
- Exploze před 6 mil let
- Více inzercí do mezigenových oblastí
- Exponenciální růst mezigenových vzdáleností



# Závislost velikosti genomu na zeměpisné šířce a nadmořské výšce

- korelace mezi velikostí genomu a teplotním režimem
- větší genomy nebo polyploidie:
  - 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 velikost vajíčka, vliv na morfologii hlavy indukovanou predátorem
- *Bombyx mori* - 1 000 000 ploidní buňky žláz produkujících 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, jak repetitivní DNA může tak významně ovlivnit genetické mechanismy svého hostitele
- změna obsahu DNA během života jako odpověď na stimuly - rozpor s Darwinismem

# Vliv velikosti genomu na fenotyp

Velikost genomu koreluje s:

+

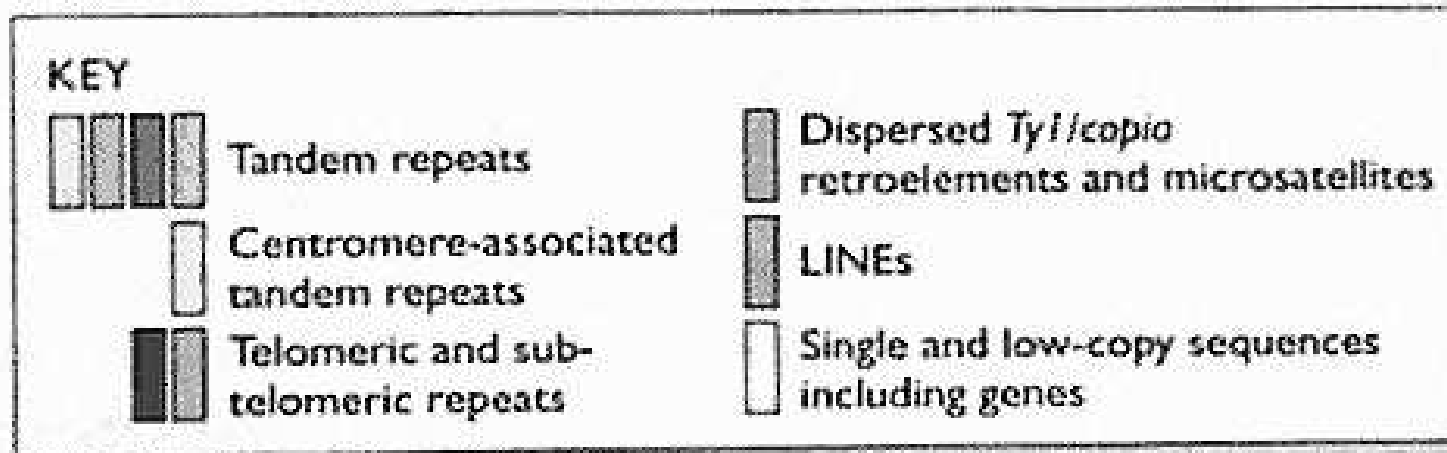
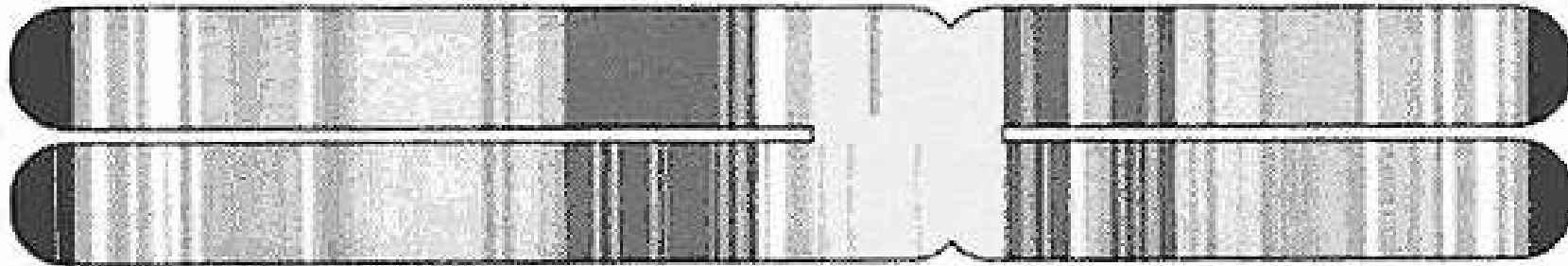
- velikostí jádra
- velikostí buňky (nucleotypic effect)
- dobou mitózy a meiózy
- minimální generační dobou
- velikostí semen
- odpovědí letniček vůči  $CO_2$
- dobou vývoje embrya u mloků

-

- 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)

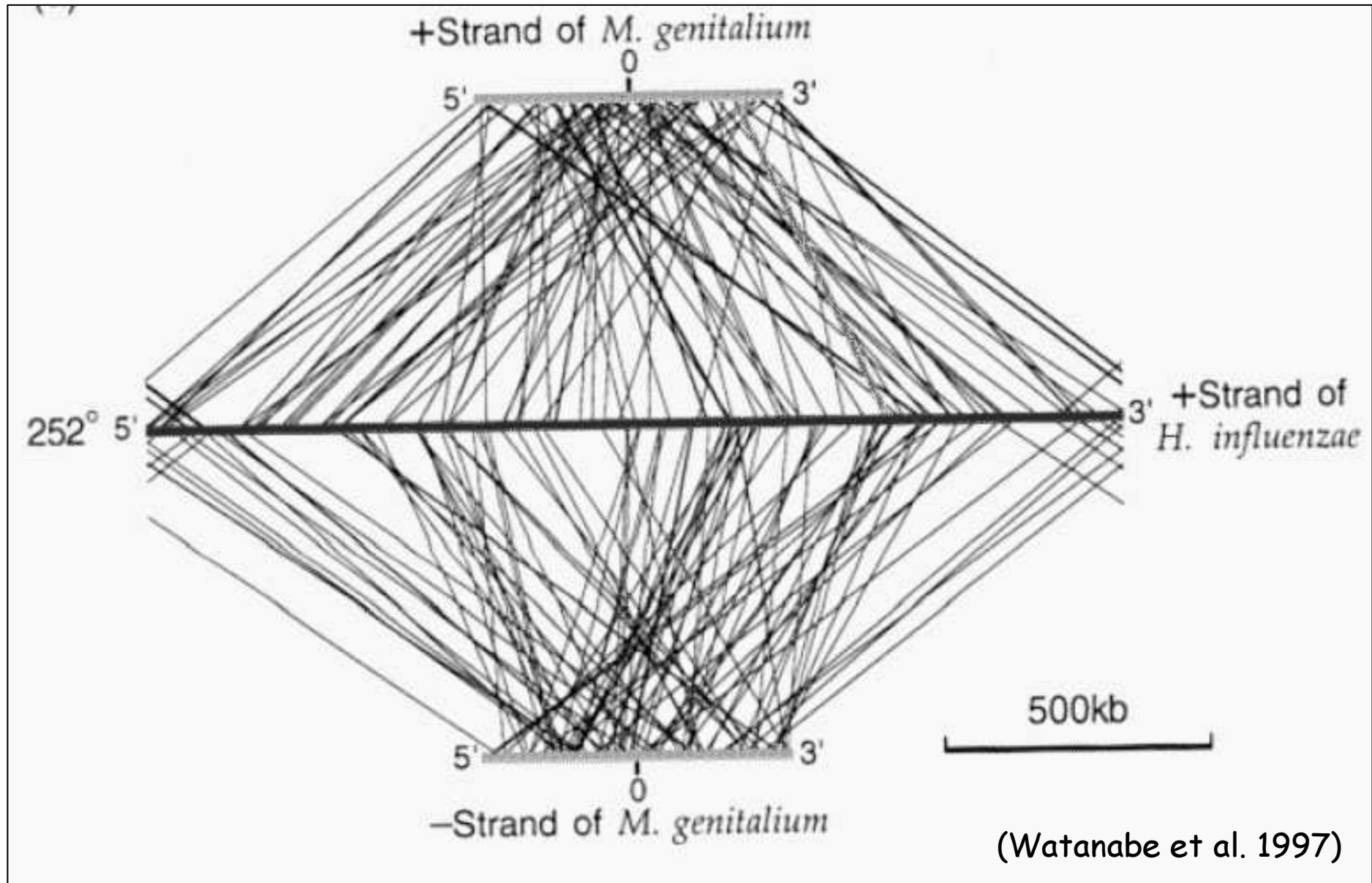
# GEOGRAFIE GENOMU

# Uspořádání genů v genomech

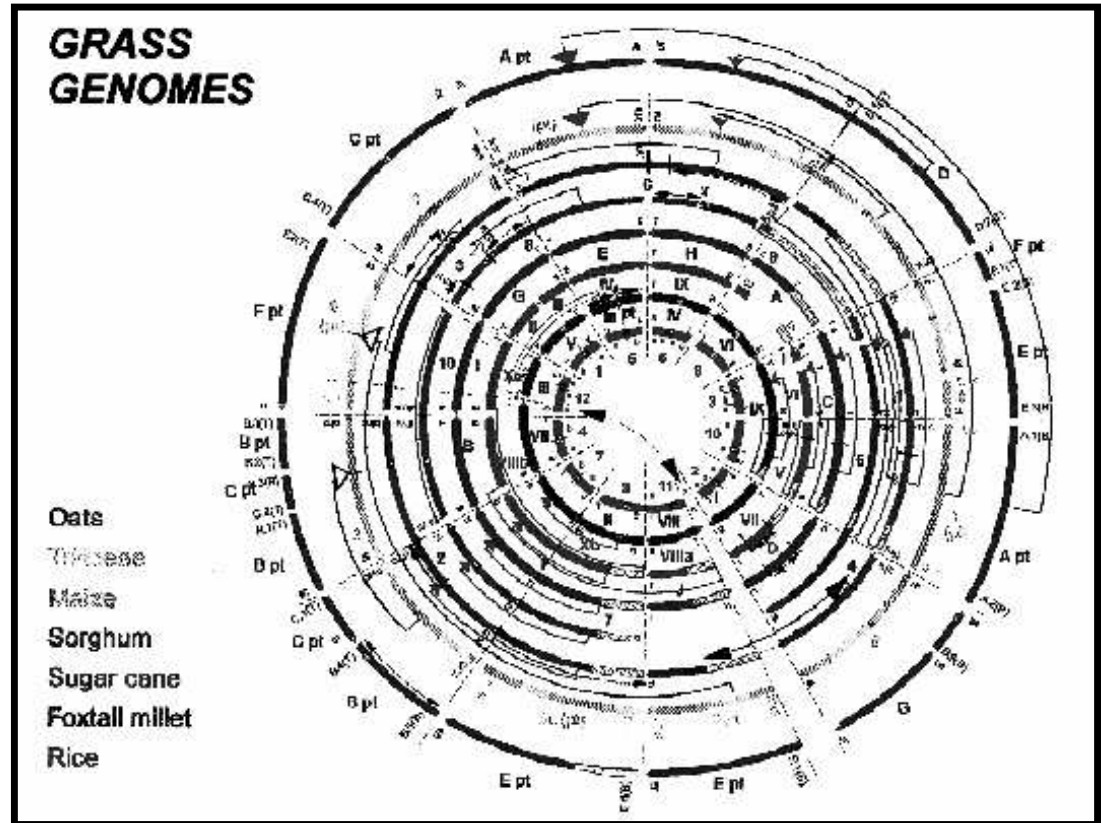
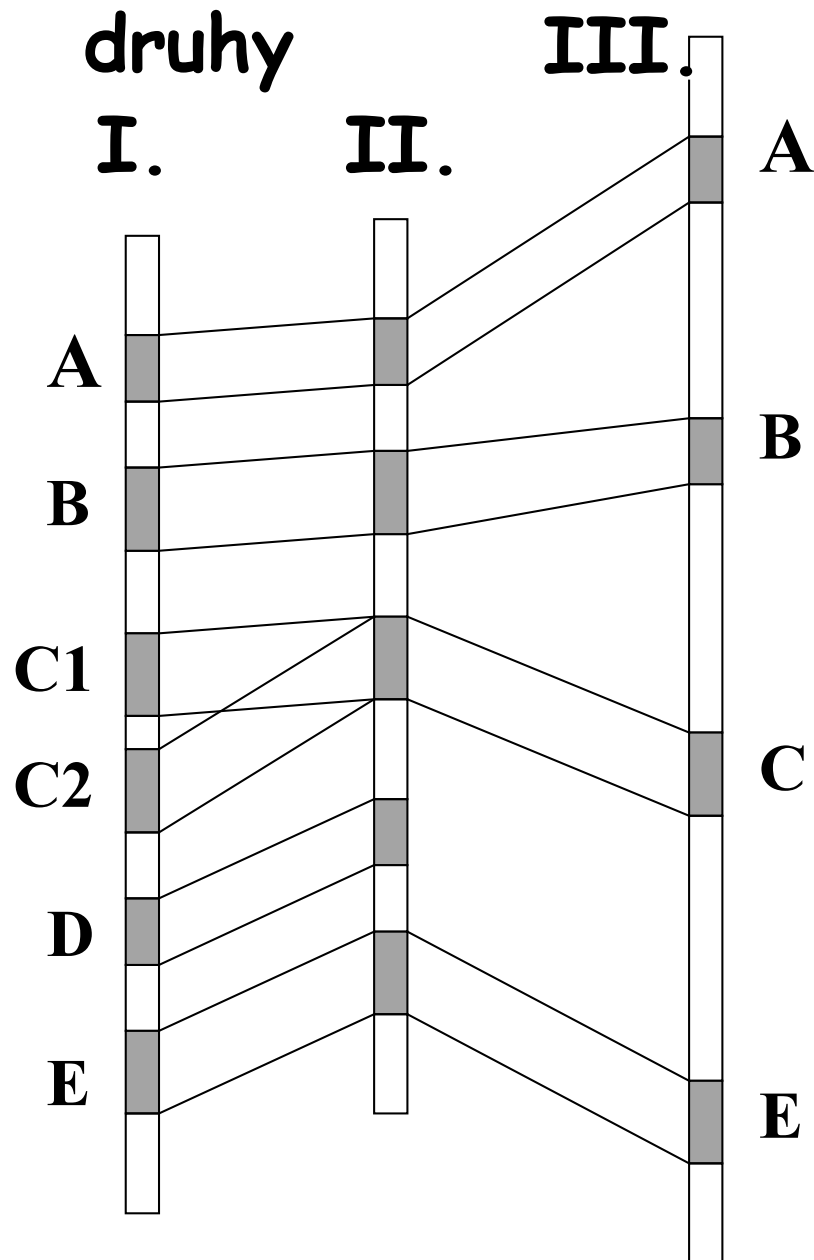




# Přestavby pořadí genů u bakterií



# Syntenie - konzervativita pořadí genů

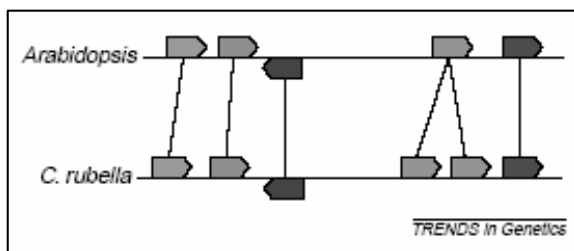


- high conservation of gene sequences and order
- no homologies of intergenic sequences

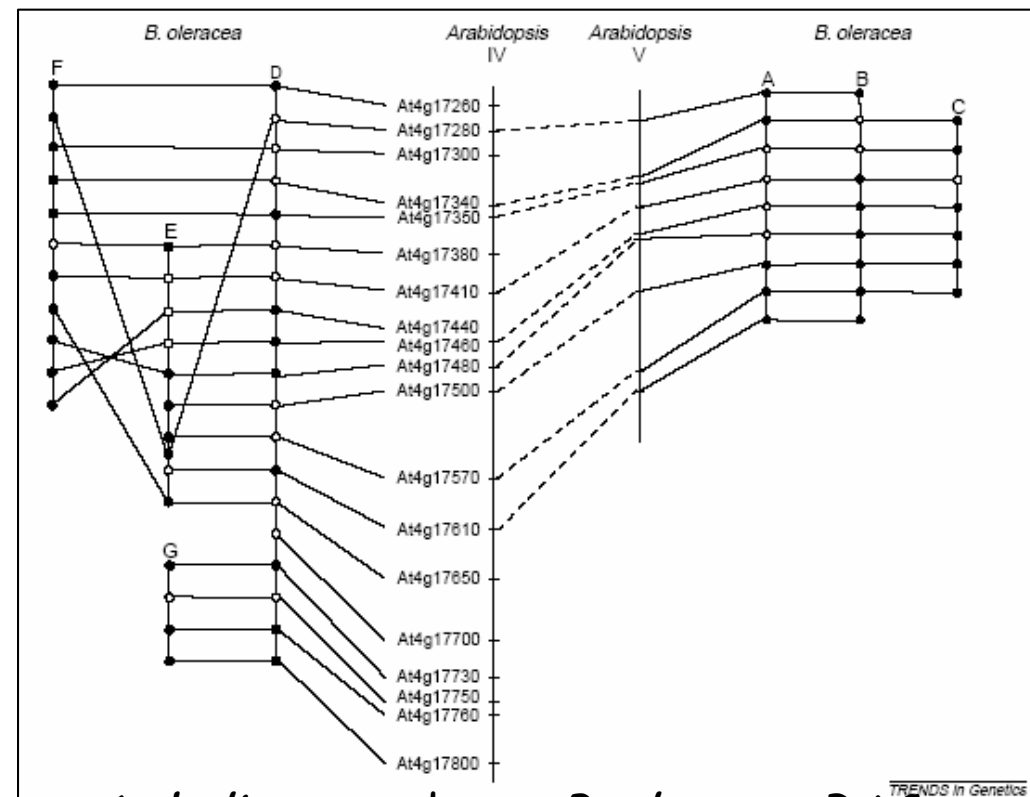
# Makro- vs. mikro- kolinearita

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

(Bancroft, 2001)



*A. thaliana* vs. *Capsella rubella* (7mil)

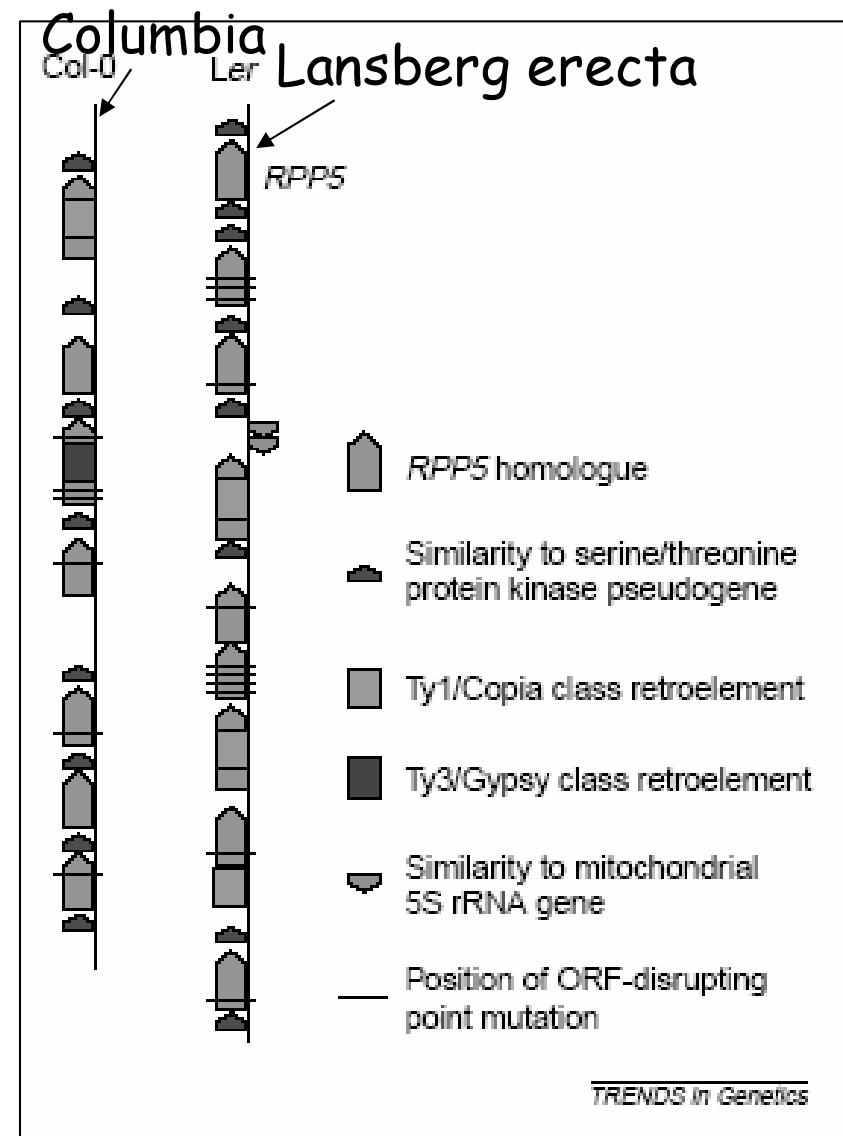


*A. thaliana* sondy vs. *B. oleracea* BACy  
plné kolečko - gen přítomen

# 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>	18	Mold	<i>Penicillium species</i>	4

## Extrémisté:

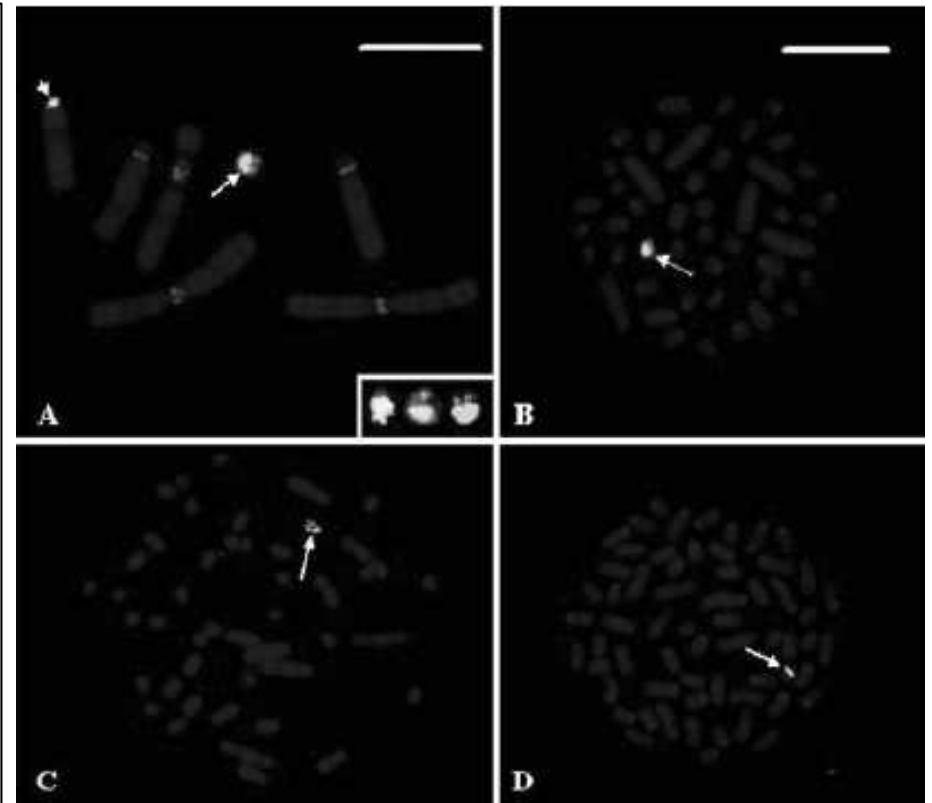
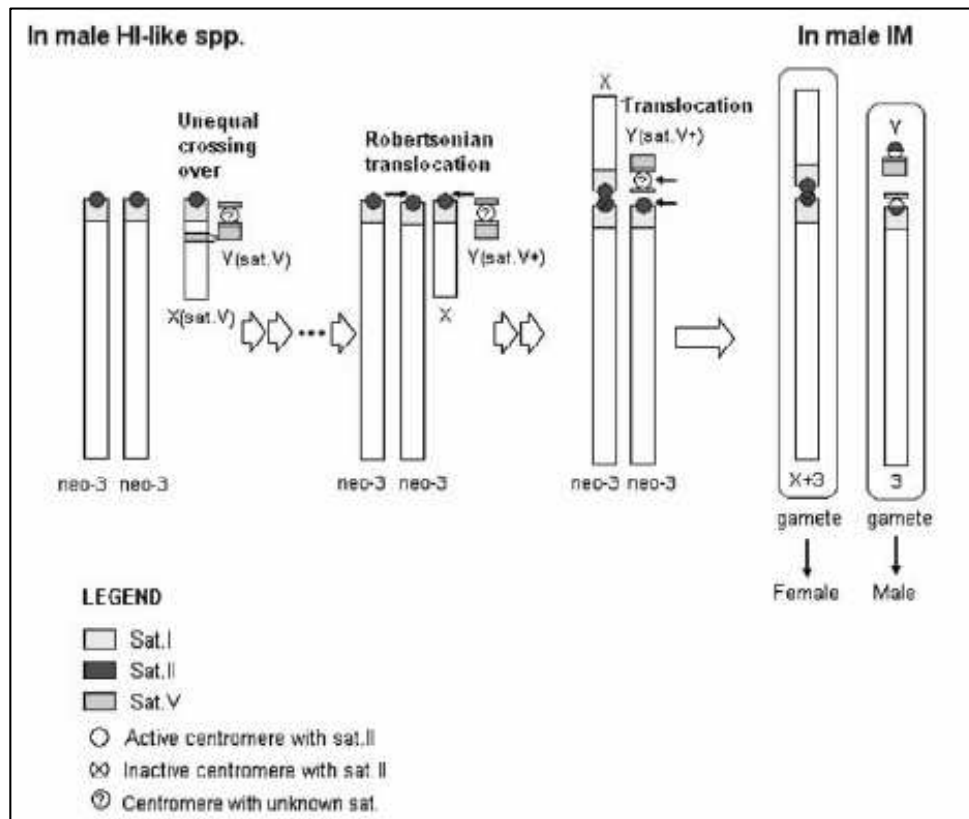
mravenec *Myrmecia pilosula* - 1 pár, samec 1 chromosom

kapradina *Ophioglossum reticulatum* - 630 párů

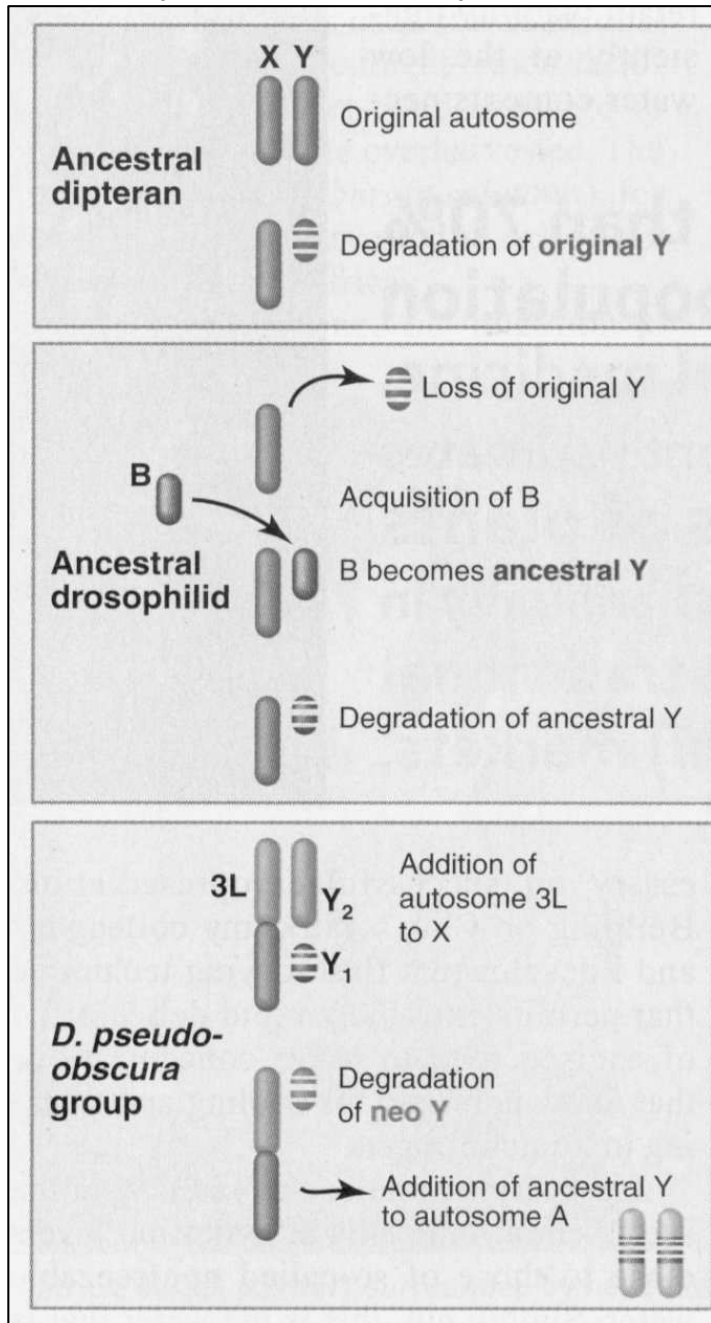
# Různé počty chromosomů u blízce příbuzných druhů jelínek *Muntiacus*

samec a samice

různé druhy



# Recyklace pohlavních chromosomů u drozofily



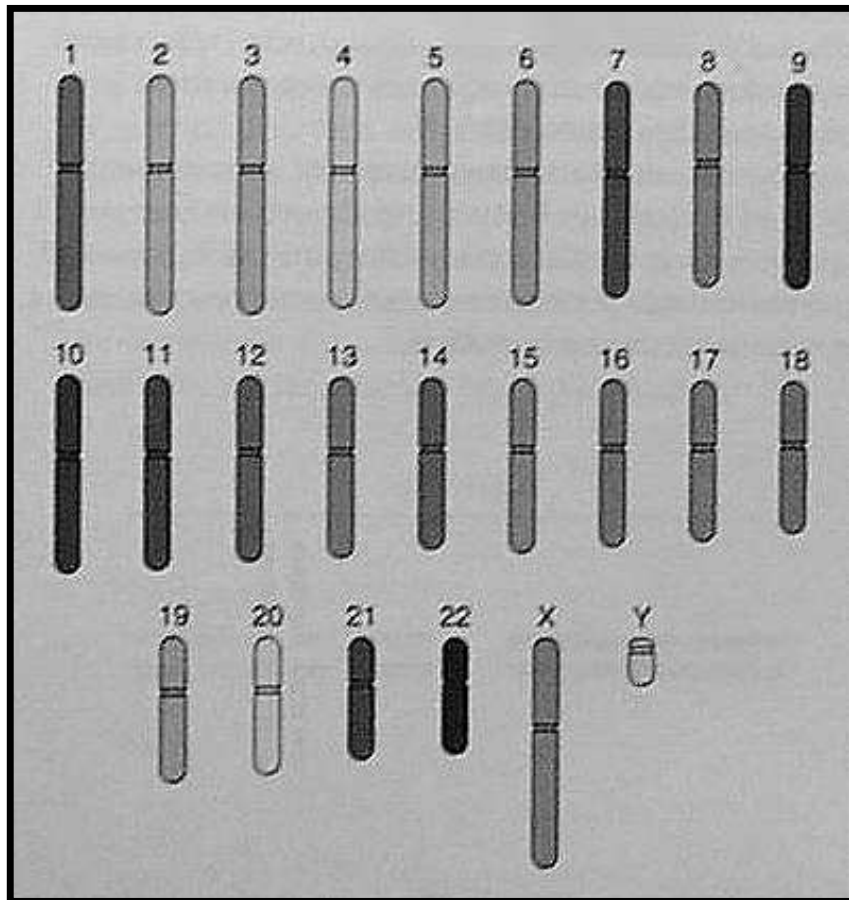
**První Y chromosom: „original Y“:**  
- vymizel před více než 60 mil. lety

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

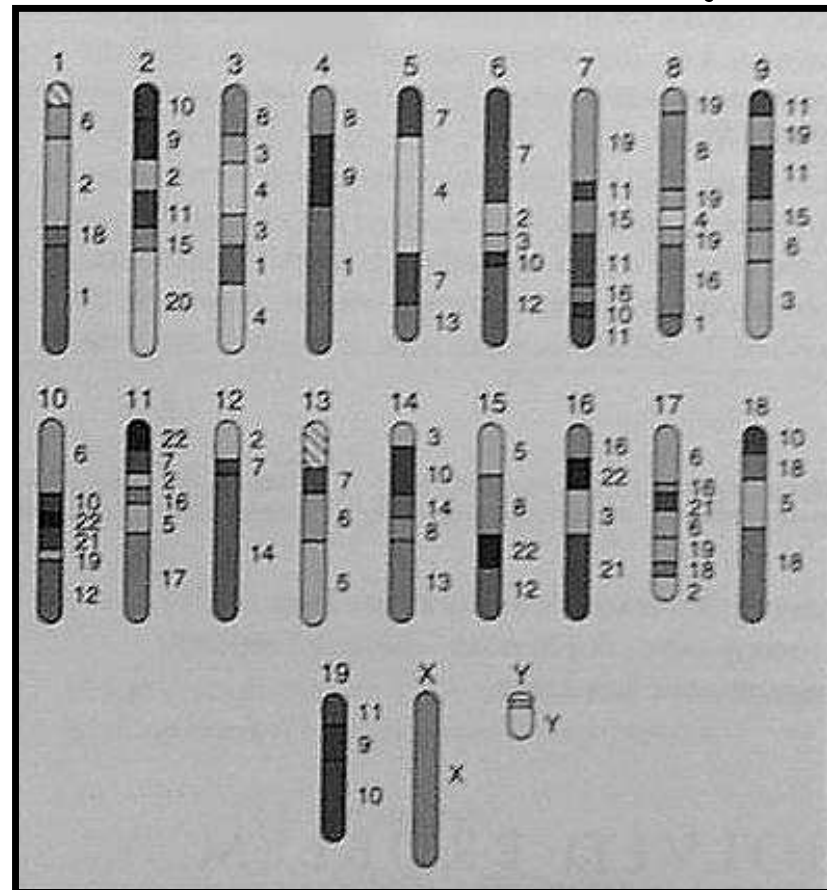
**Třetí Y chromosom: „neo Y“**  
- fúze X s autosomem  
- připojení ancestrálního Y na A  
- opět degenerace neo-Y

# Srovnání lidských a myších chromosomů

člověk



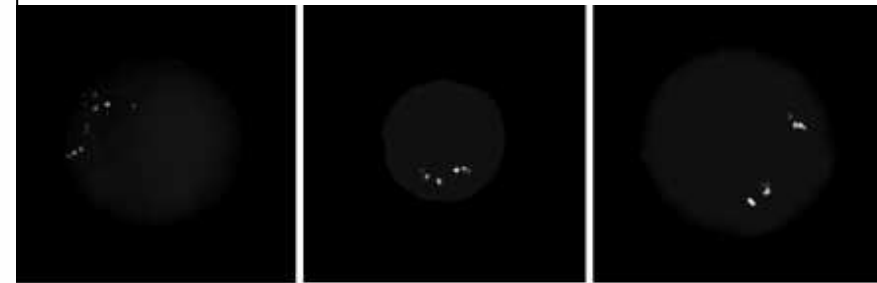
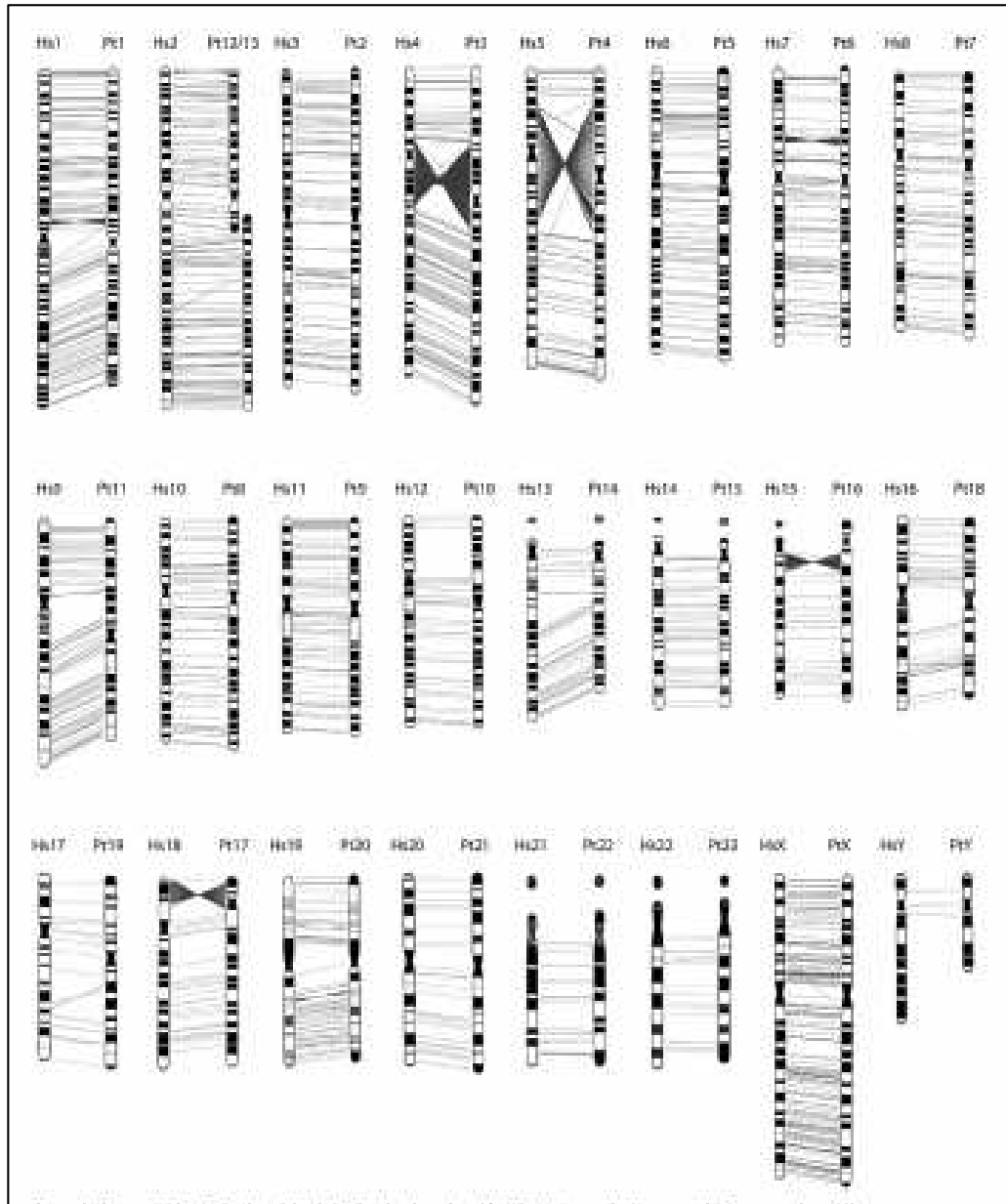
myš



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

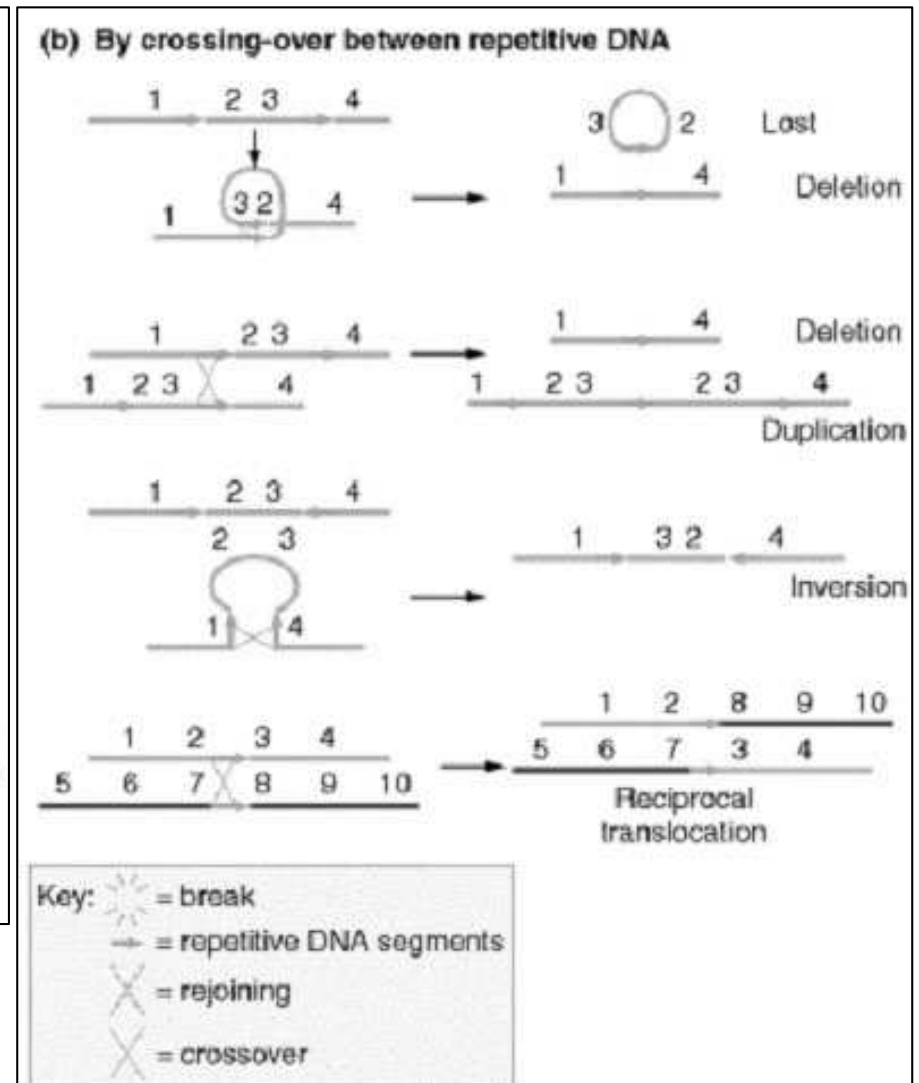
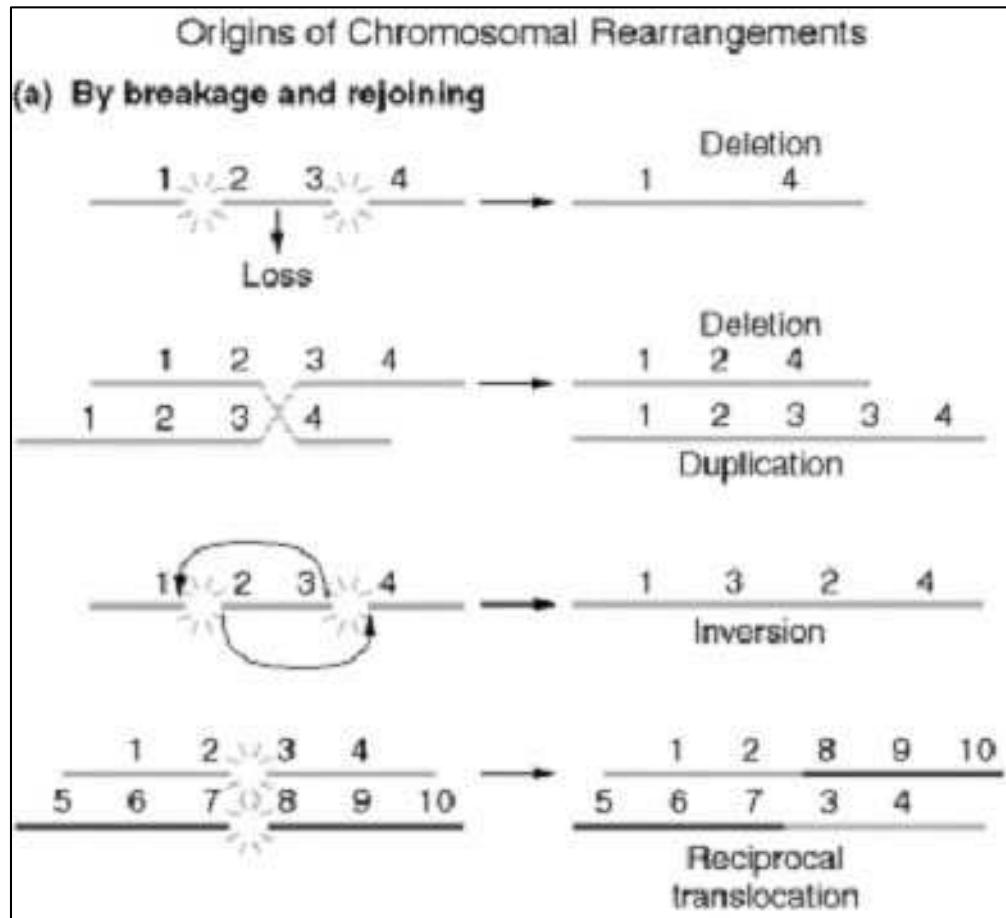
Chimpanzee

Gorilla

# Chromosomové přestavby

zlom a znovuspojení

rekombinace mezi repetitivními



(Introduction to Genetic Analysis. 7th ed)  
 Griffiths, Anthony J.F.; Gelbart, William M.; Miller, Jeffrey  
 H.; Lewontin, Richard C.

# Izochorový model organizace genomu

**Isochory:** bloky genomové DNA (stovky kb-Mb) s charakteristickým GC-obsahem

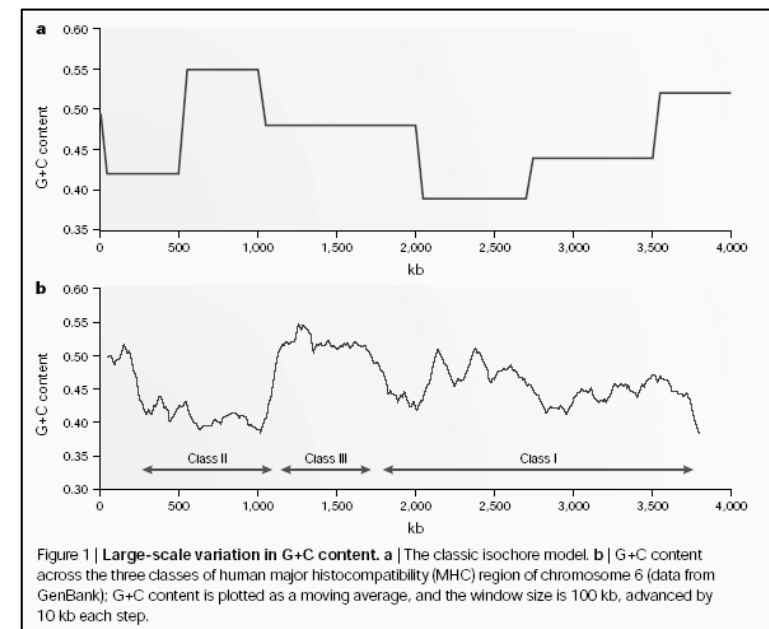
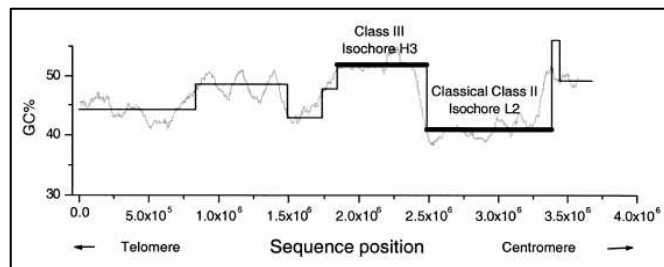
- fragmentace genomové DNA na fragmenty 30-100kb
- separace fragmentů podle obsahu bází
- shlukování do skupin s diskrétními GC-obsahy, u člověka typy L1, L2, H1, H2 a H3 (GC nejbohatší, tvoří jen 3% ale obsahuje 25% genů)
- Mozaiková struktura genomu obratovců a rostlin

## Původ izochor:

Výsledek selekce?

Výsledek mutací?

Genová konverze?



# Procesy vedoucí ke tvorbě izochor

## **Mutační bias:**

- včleňování nukleotidů při replikaci je ovlivněno koncentrací volných nukleotidů
- koncentrace nukleotidů závisí na poloze v jádře
- různé části genomu replikovány v různou dobu
- příčinou izochor je různá účinnost reparace DNA
- izochory jsou důsledkem deaminace cytosinu (C-U), častější v AT-bohatých oblastech, jejich růst

## **Selekce:**

- izochory jsou výsledkem selekce
- Selekce na úrovni teplotní stability DNA a zastoupení AK

## **Genová konverze:**

- korelace mezi rekombinací a obsahem GC

# Mozaiková struktura lidského genomu

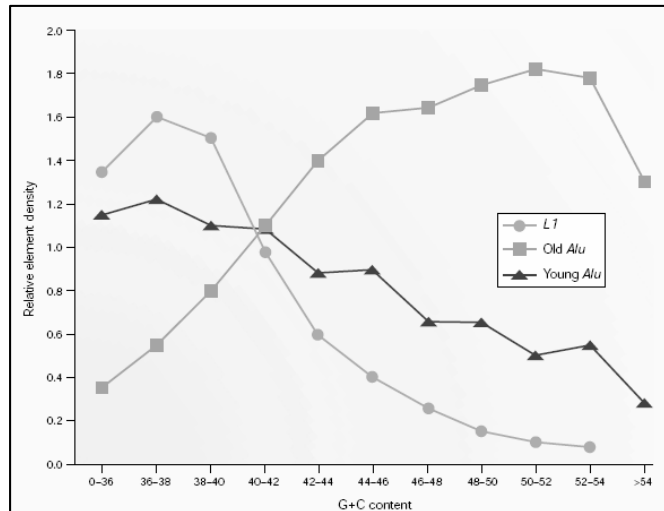
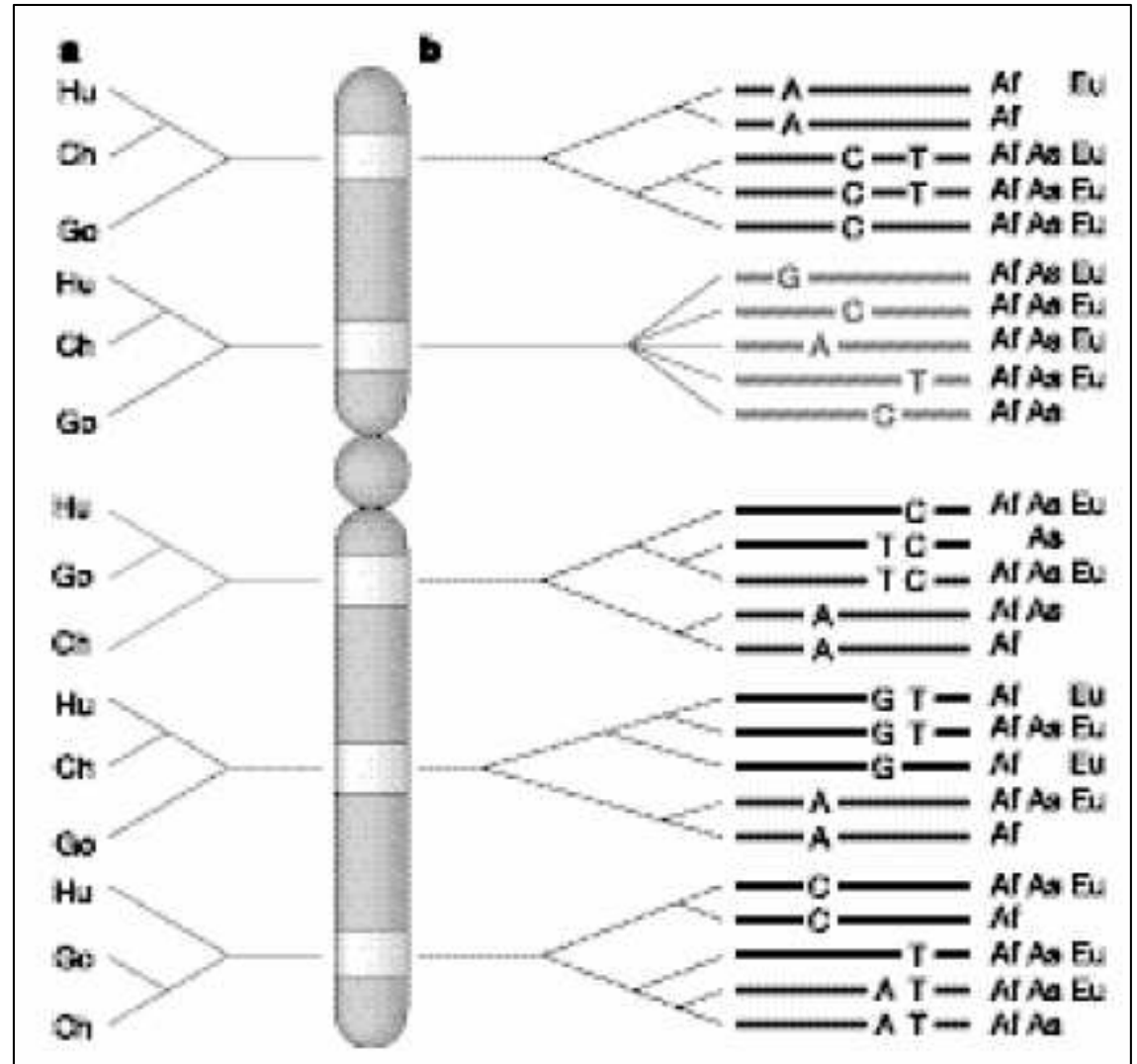
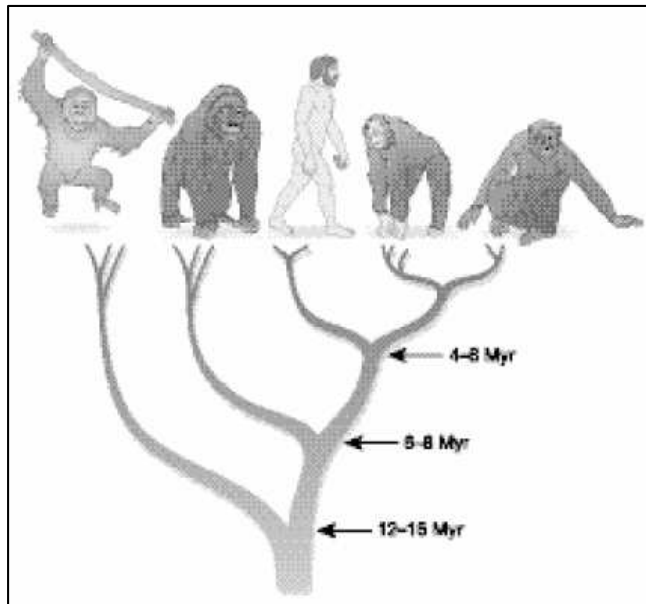


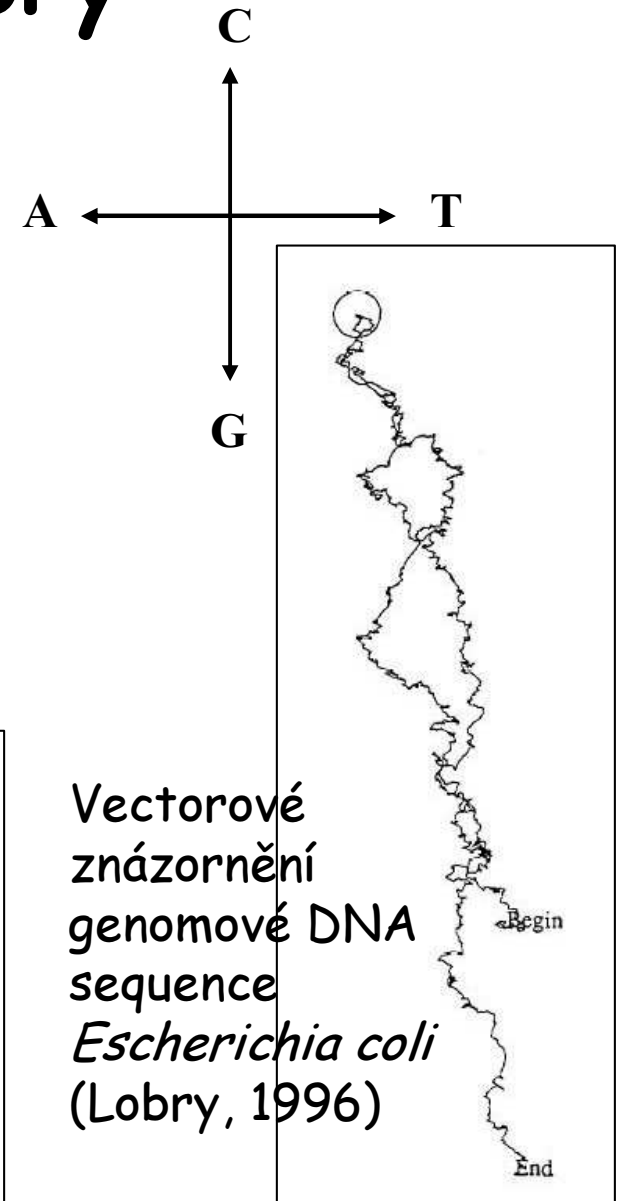
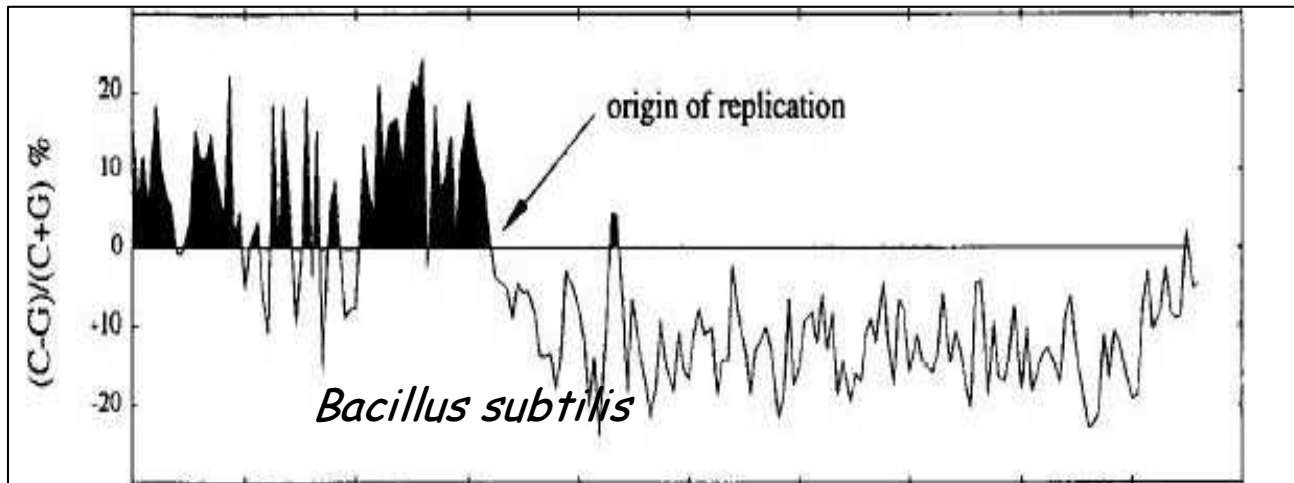
Figure 3 | The density of *Alu* and *L1* elements in the human genome. The density of young *Alu* (<1 million years (Myr) old), old *Alu* (60–100 Myr old) and *L1* (<65 Myr old) elements in the human genome, relative to the average density of *Alu* and *L1* elements, separated according to the G+C content of the 50 kb that surrounds each element. Data derived from REF. 1.



# Další zajímavé struktury v genomu: chirochory

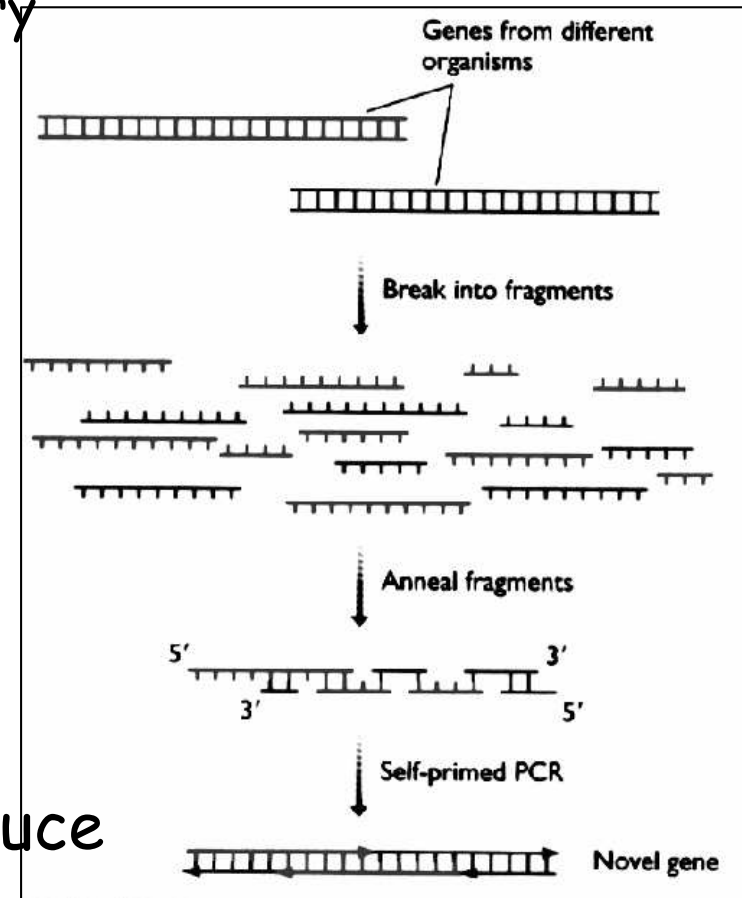
## Chirochory:

- různý obsah bází v komplementárních vlákněch bakteriální DNA
- oblasti s homogenním zastoupením bází v jednom vlákně
- korelace s polohou replikačního počátku



# Jak studovat evoluci genomu?

- Simulace evoluce „*in silico*”
- **Evoluční experimenty „*in laboratory*”**
  - SELEX experimenty - ribozymy, aptamery
  - selekce bakterií
  - řízená evoluce in vitro



řízená evoluce

# GENOMOVÉ PROJEKTY



# První genomové projekty - začátek éry genomiky

1977: první kompletní genom (virus) -  
bakteriofág phiX174

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

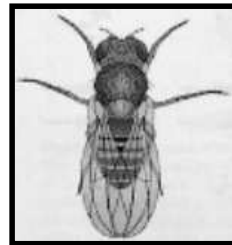


1996: *Saccharomyces cerevisiae* (12 Mb/ 6548 genes)

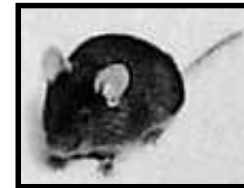


1998: *Caenorhabditis elegans* (97 Mb/19 099 genů)

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



*Mus musculus*  
(3 300 Mb)



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



# Jak zjistit stav genomových projektů



**GOLD<sup>™</sup>**  
**Genomes OnLine Database v 2.0**



<b>Contact:</b> <u><a href="http://www.genomesonline.org">Genomesonline</a></u>	<b>Last Update:</b> <b>January 31, 2006</b>	<b>Location</b> <u><a href="http://www.genomesonline.org">www.genomesonline.org</a></u>
	<u>Search GOLD:</u> <b>1945</b> genome projects	<b>28</b> Metagenomes
<b>340</b> Published Complete Genomes	<b>989</b> Prokaryotic Ongoing Genomes	<b>588</b> Eukaryotic Ongoing Genomes

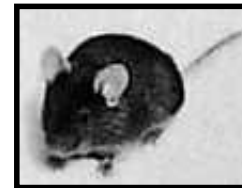
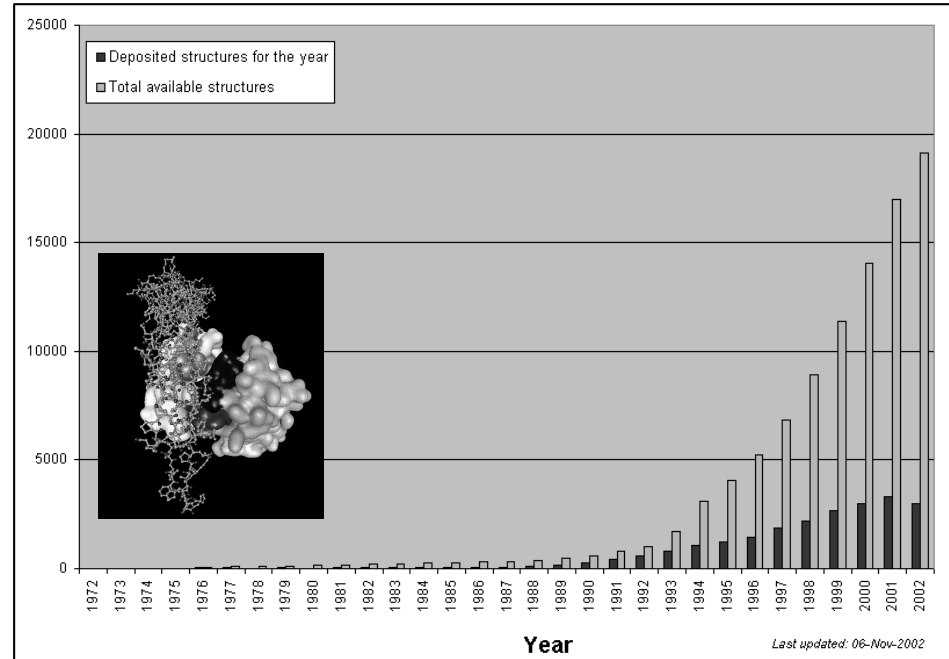
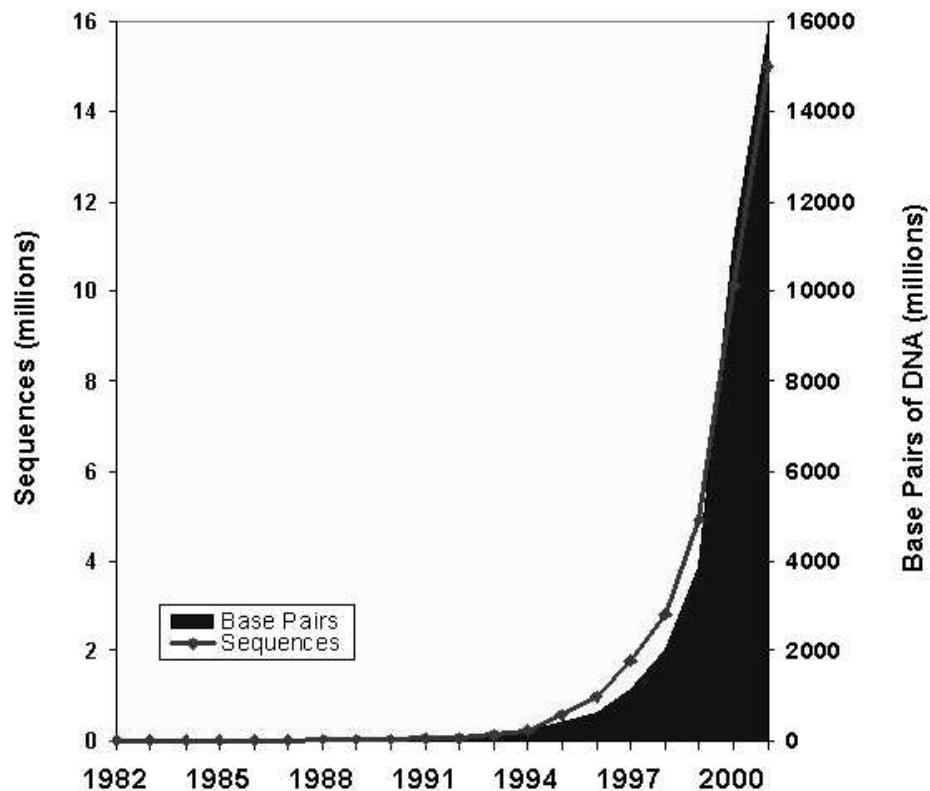
<http://www.genomesonline.org/>

# Objem dat v databázích roste exponenciálně

~ 22 000 000 000 bp

~ 20 000 struktur

## Growth of GenBank



~ 1200 organizmů

# Hlavní databáze

**Genbank** - sekvence DNA, veřejně přístupná

**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

# Genomové sekvenační projekty

únor 2006

## 340 ukončených projektů:

19 archea  
175 bakterie  
28 eukaryota

## 975 probíhajících projektů:

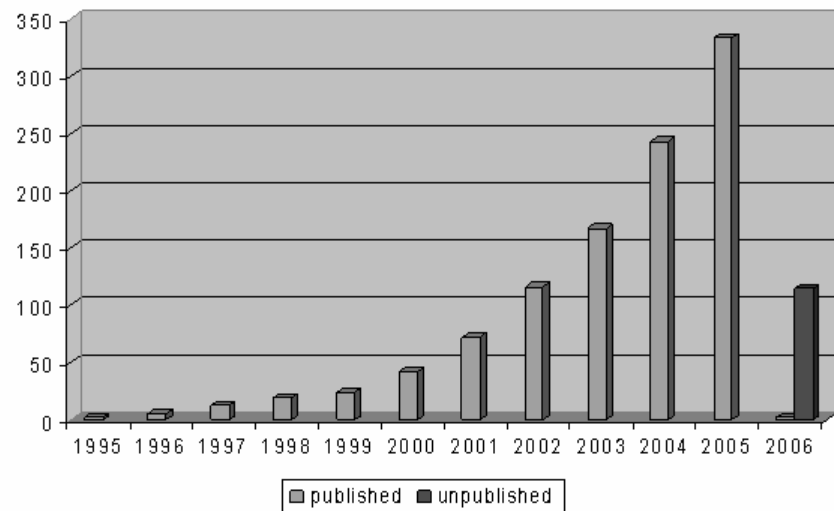
989 prokaryota  
588 eukaryota

## Ukončené eukaryotické genomy:

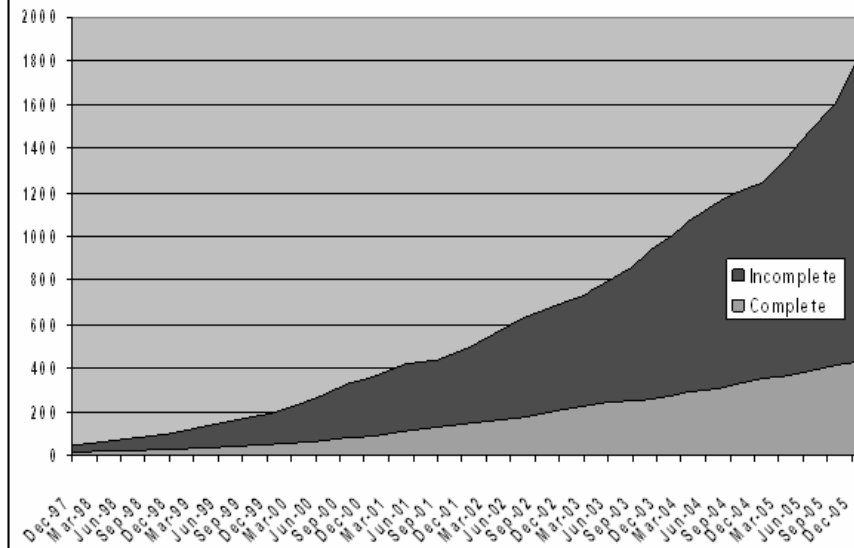
protozoa: *Plasmodium falciparum* (23 Mb)  
fungi: *Saccharomyces cerevisiae* (12 Mb)  
*Neurospora crassa* (43 Mb)  
nematodes: *Caenorhabditis elegans* (97 Mb)  
insects: *Drosophila melanogaster* (137 Mb)  
*Anopheles gambiae* (278 Mb)  
fish: *Ciona intestinalis* (117 Mb)  
mammals: *Homo sapiens* (2900 Mb)  
*Mus musculus* (3300 Mb)  
*Rattus norvegicus* (2750 Mb)  
plants: *Arabidopsis thaliana* (115 Mb)  
*Oryza sativa* (420 Mb)

# Genomové sekvenační projekty

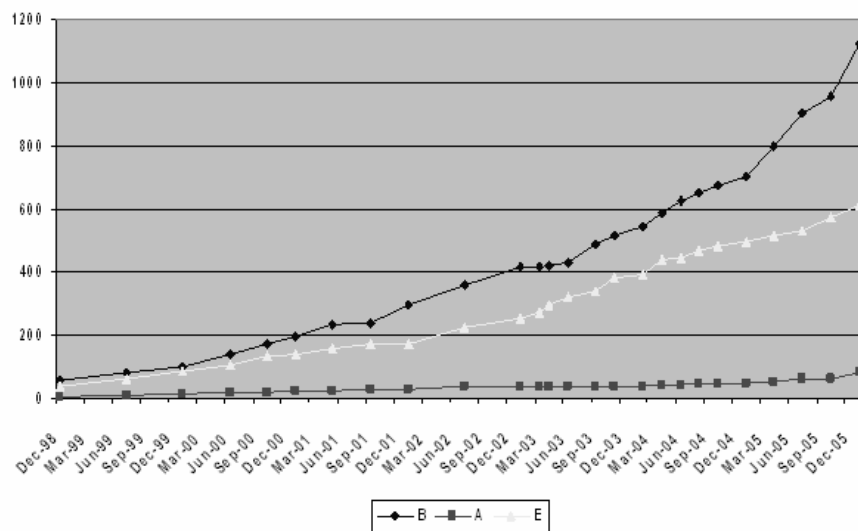
Completely Sequenced Genomes ©  
January 2006



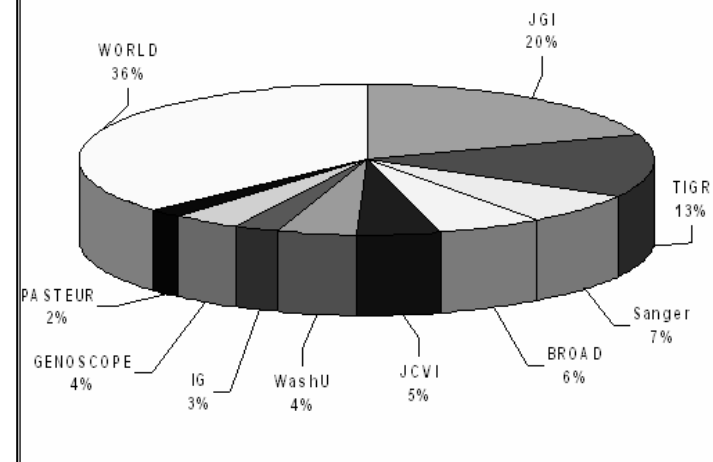
Genome Sequencing Projects on GOLD ©



Genome Projects on GOLD according to Phylogenetic Groups ©



Major Sequencing Centers  
January 06: 1835 projects



# Sekvenování lidského genomu

