

Eduard Kejnovský + Roman Hobza

EVOLUČNÍ GENOMIKA

IV. EVOLUCE GENOMŮ II: VZESTUP



OBSAH

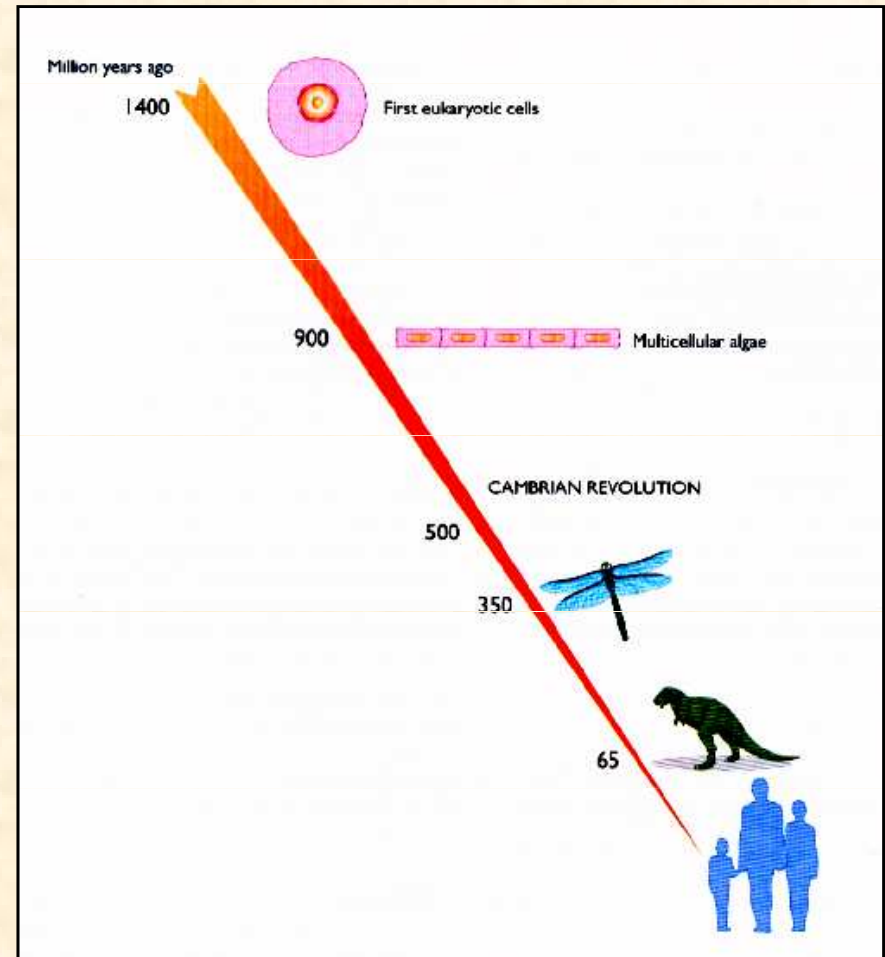
1. Změny ve velikosti genomu
2. Nekódující DNA a velikost genomu
3. Topografie genomu
4. Genomové projekty

ZMĚNY VE VELIKOSTI GENOMU

Zvětšení genomu

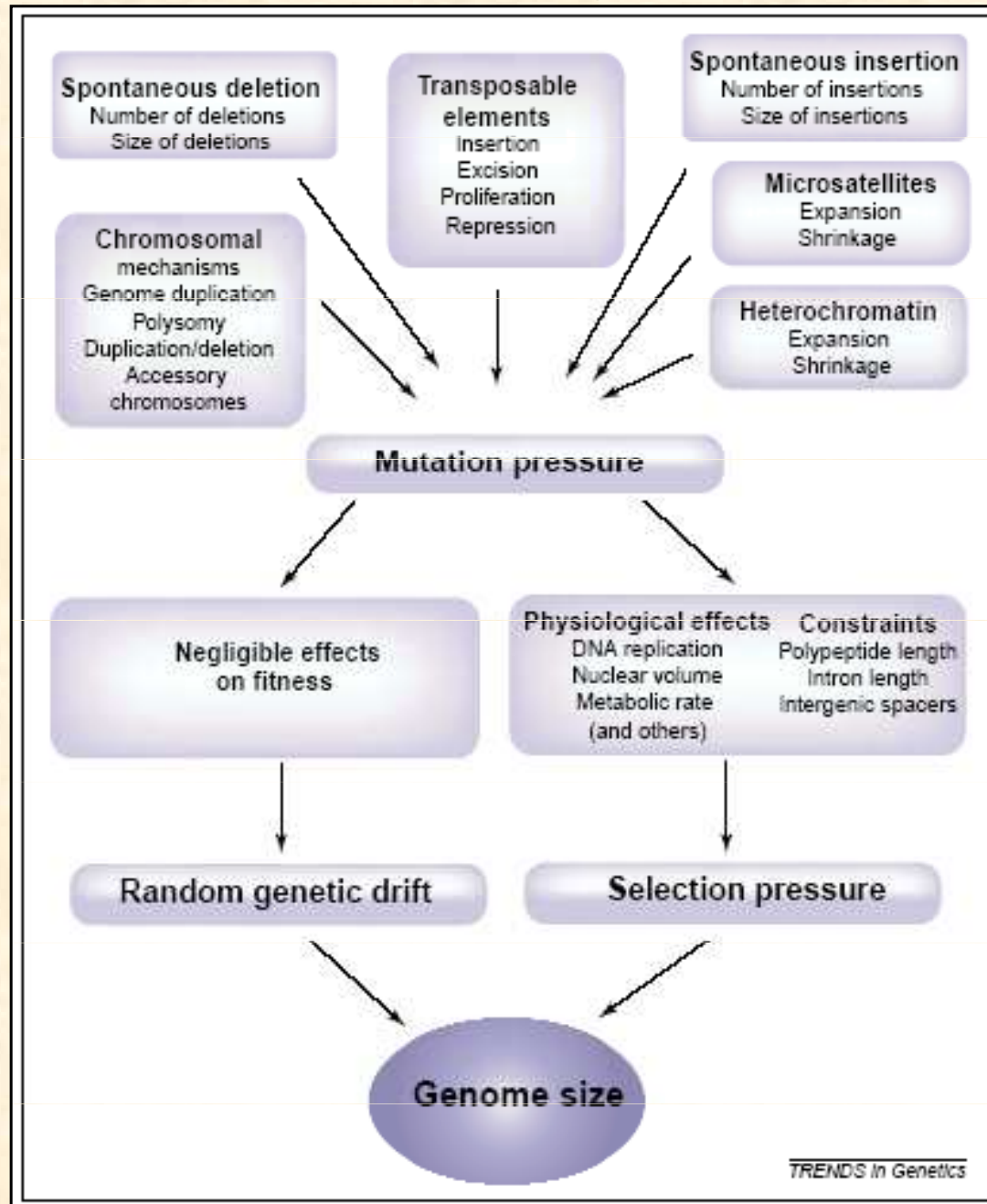
- celkové zvětšení:
polyploidizace (duplikace celého **genomu**)
- duplikace části genomu,
zmnožení počtu **chromosomů**
- duplikace **genů** a skupin genů

- 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?

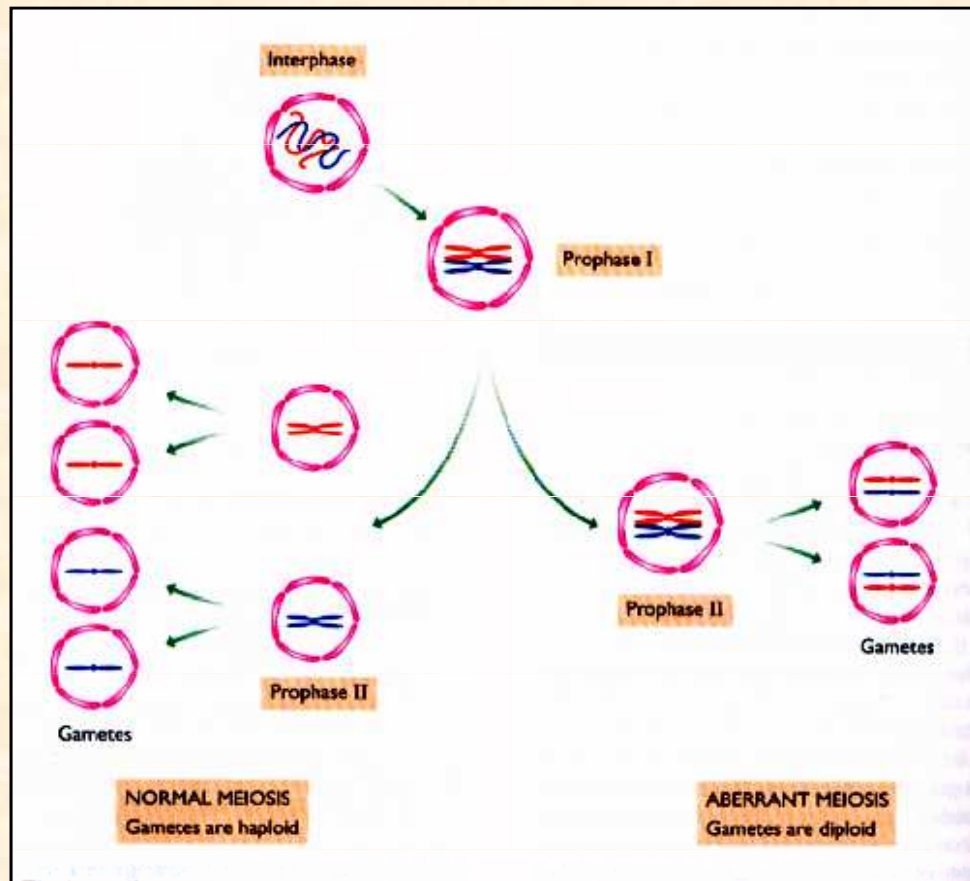
Plynulost:

- 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

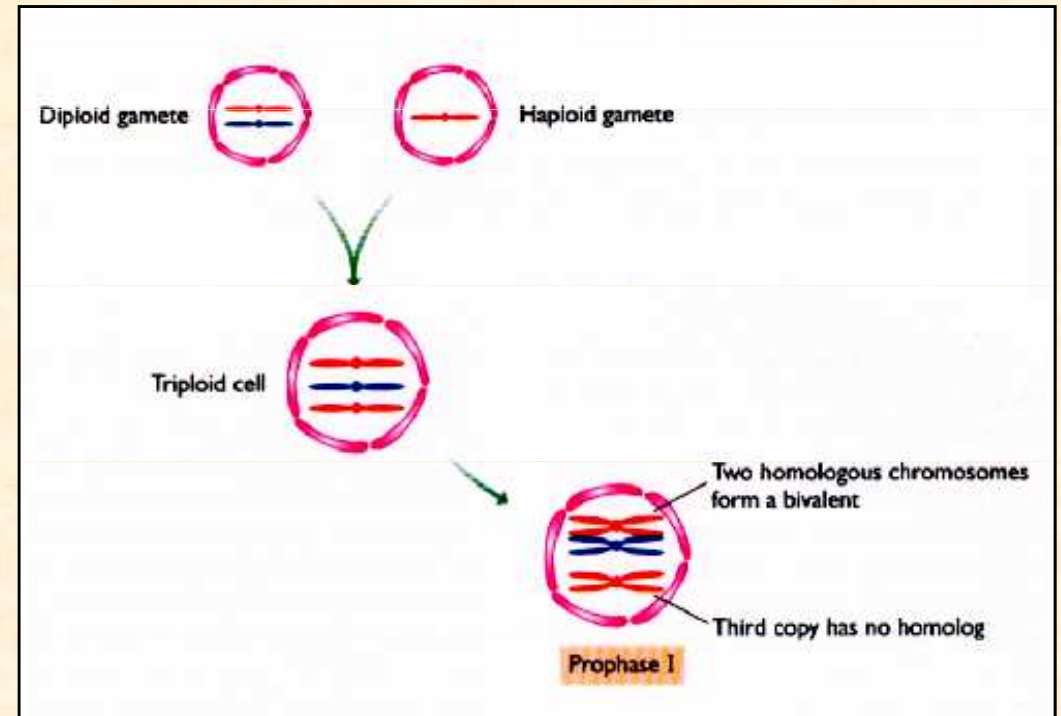
Skoky (kvantové změny):

- bezobratlí, rostliny (50% nahosemenných je polyploidní), výjimečně obratlovci
- polyplodizace, přídavná replikace některého chromosomu, kryptopolyploidizace (koordinovaná replikace všech repetací 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 organizmus, chyba při meioze (růže...)
 - Allopolyploidie:** kombinace geneticky odlišných sad chromosomů (pšenice...)
 - Kryptopolyploidie:** dávná polyploidizace, organizmus 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

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

Polyploidní plodiny

- tetraploidi: kukuřice, bavlna, brambor, zelí



- hexaploidi: pšenice, chrysan téma

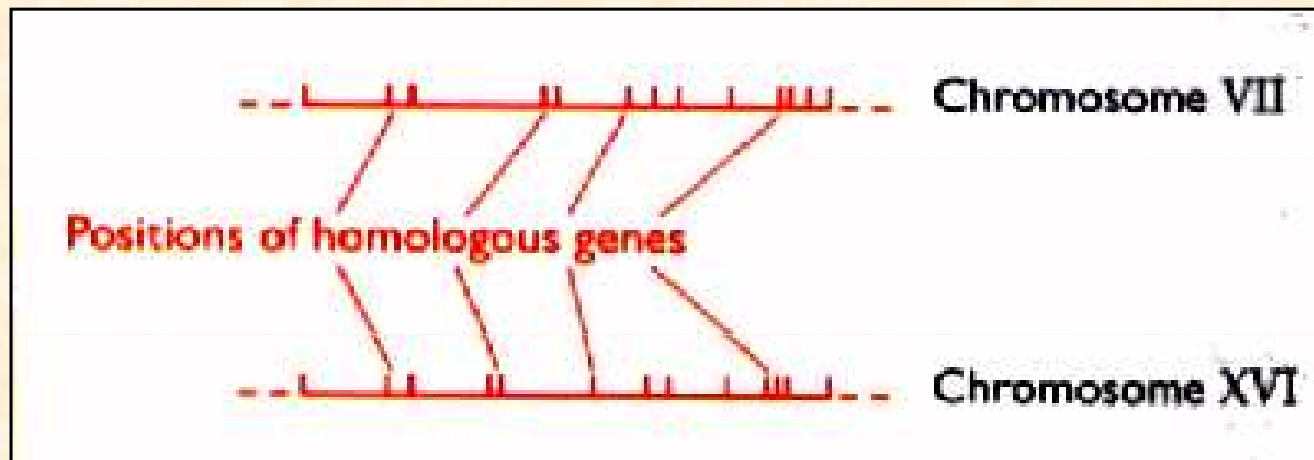


- oktoploidi: jahodník



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*

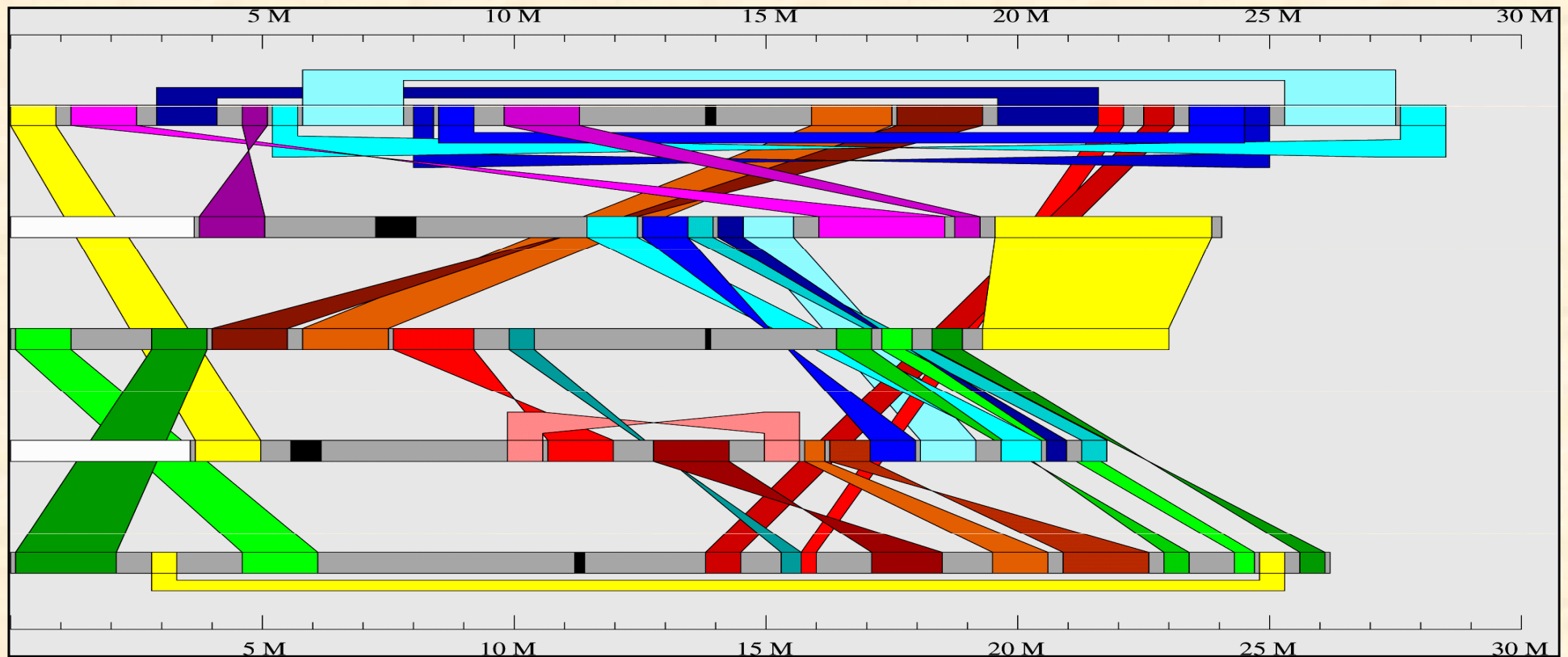
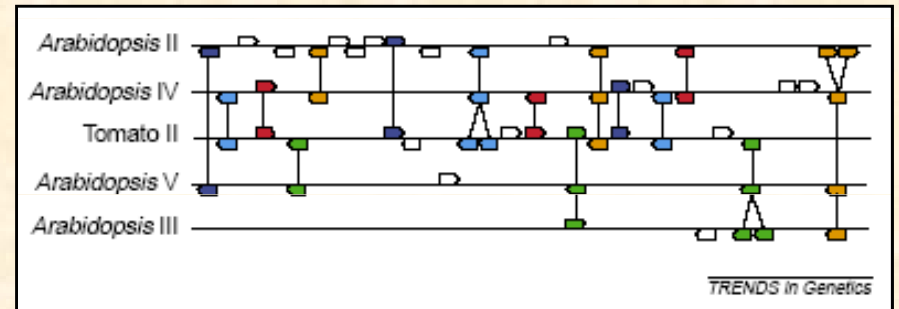


Duplikace jsou
výsledkem
tetraploidie spíše
než postupných
regionálních
duplikací

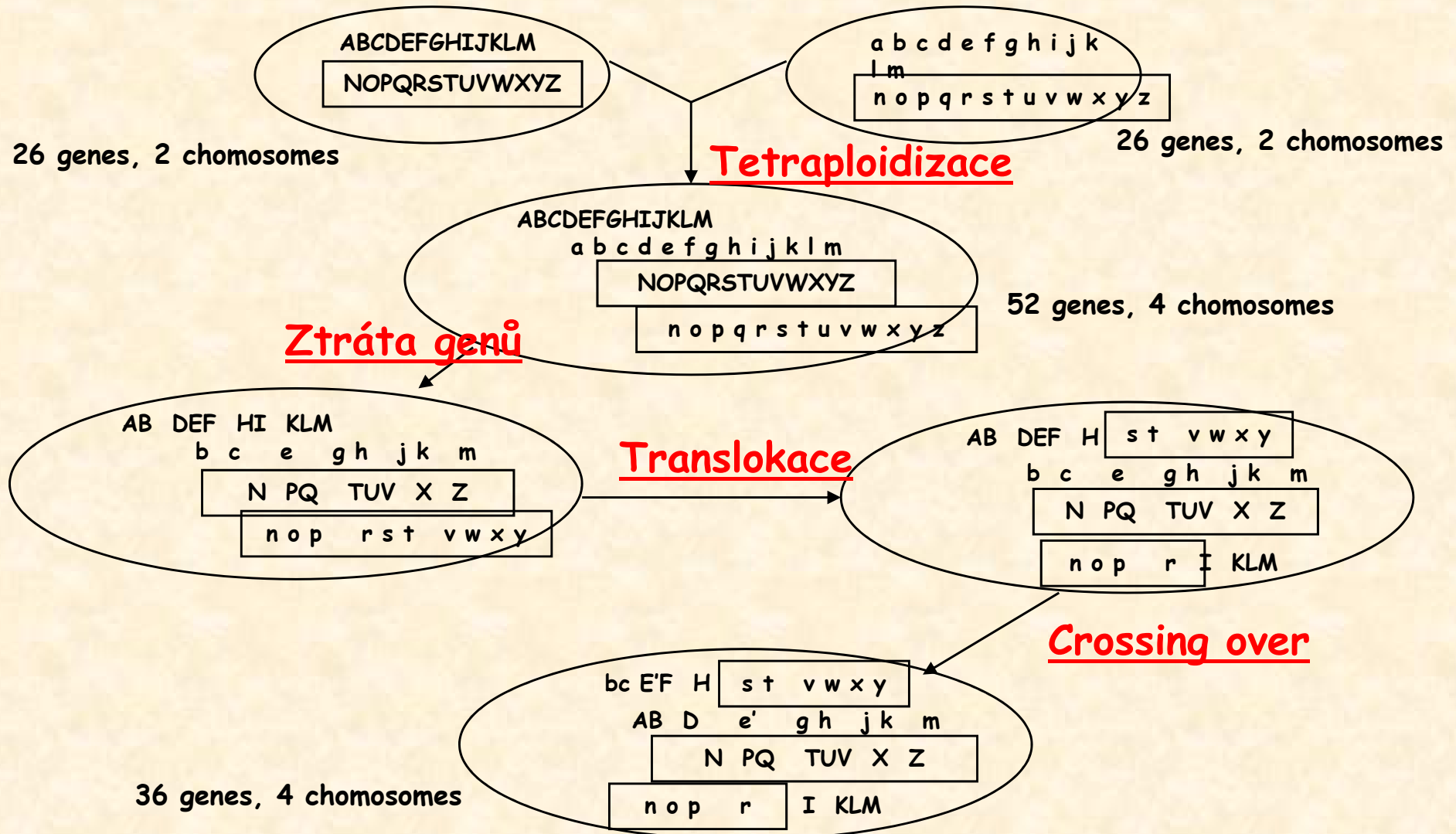
Polyploidizace u rostlin

Arabidopsis:

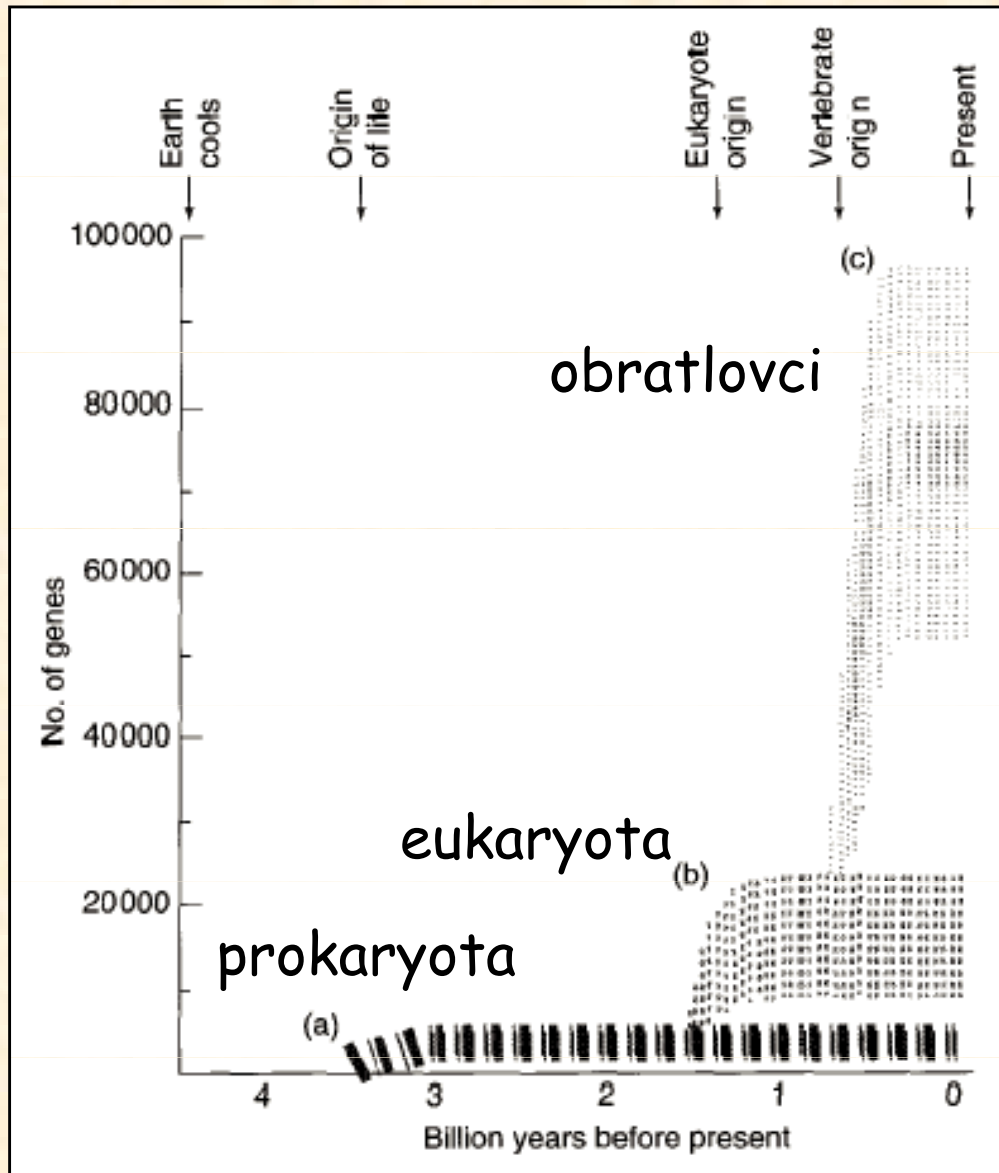
- duplikace před 200 a 80 mil. let
- možná i více než 4 genomové duplikace
- asi 60% genomu leží v duplikovaných segmentech
- 50% genů v nich je konzervováno



Počet genů a evoluce duplikovaného genomu



Vzrůst počtu genů, potlačení šumu a biologická komplexita



Prokaryota/eukaryota:

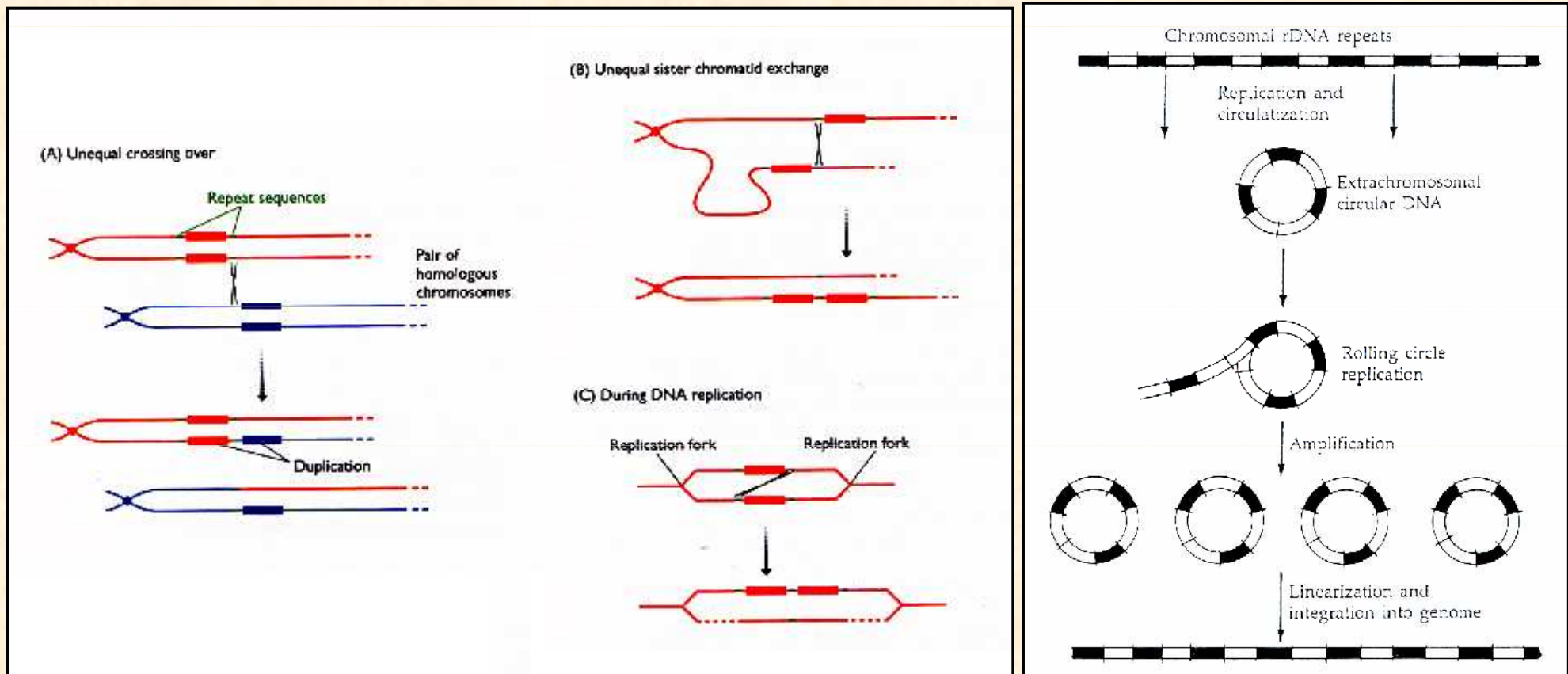
Potlačení šumu - **separace**
transkripce a translace,
jaderná membrána a
histony

Obratlovci:

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

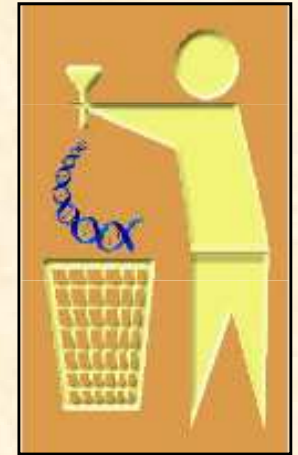
Mechanismy zvětšení genomů

- 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 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

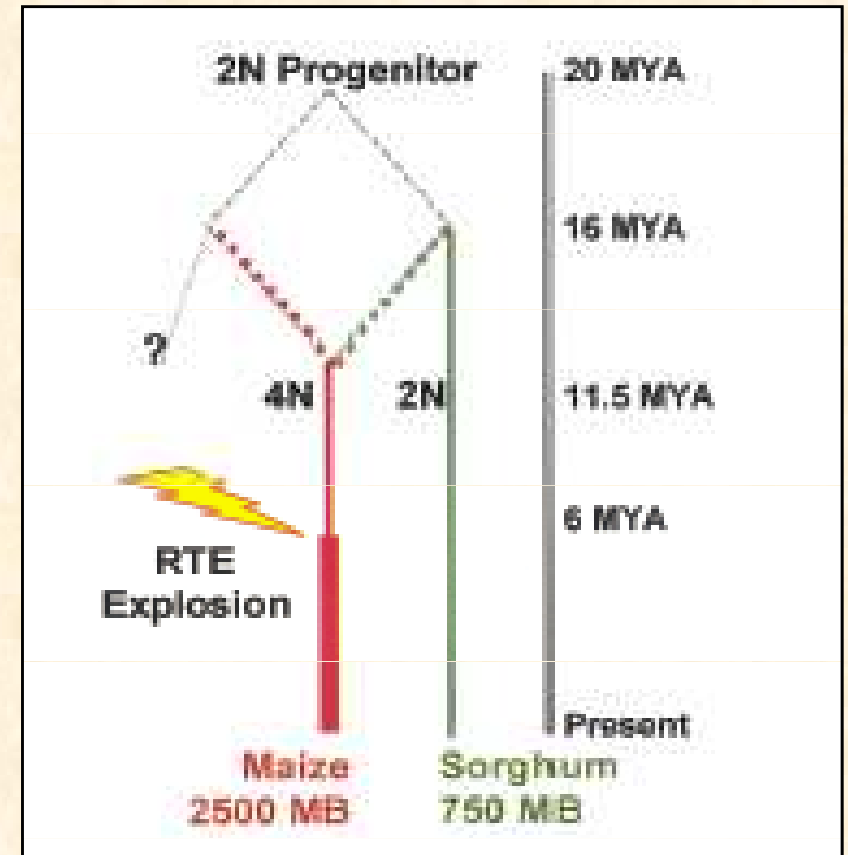
Limity růstu genomů:

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



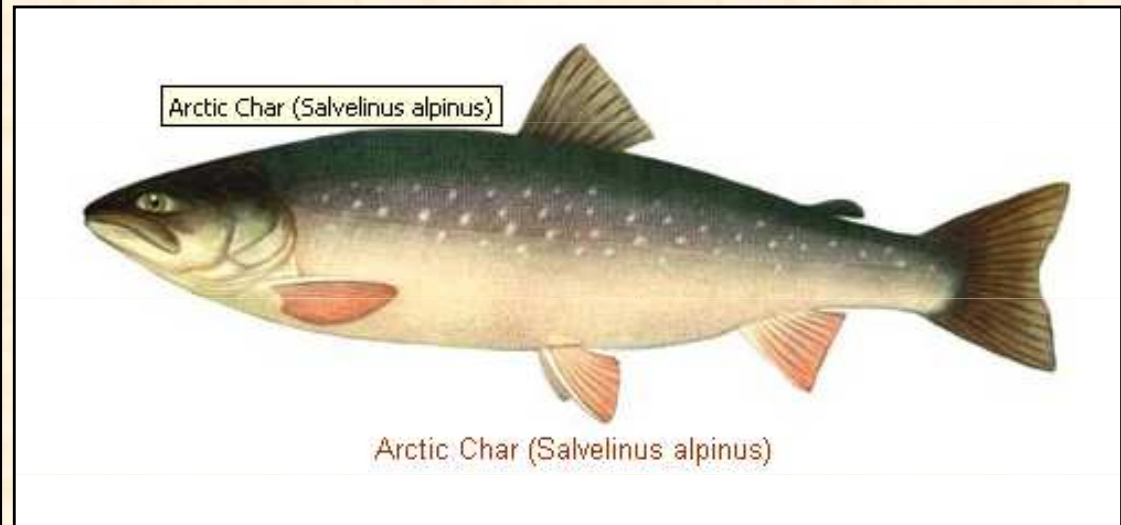
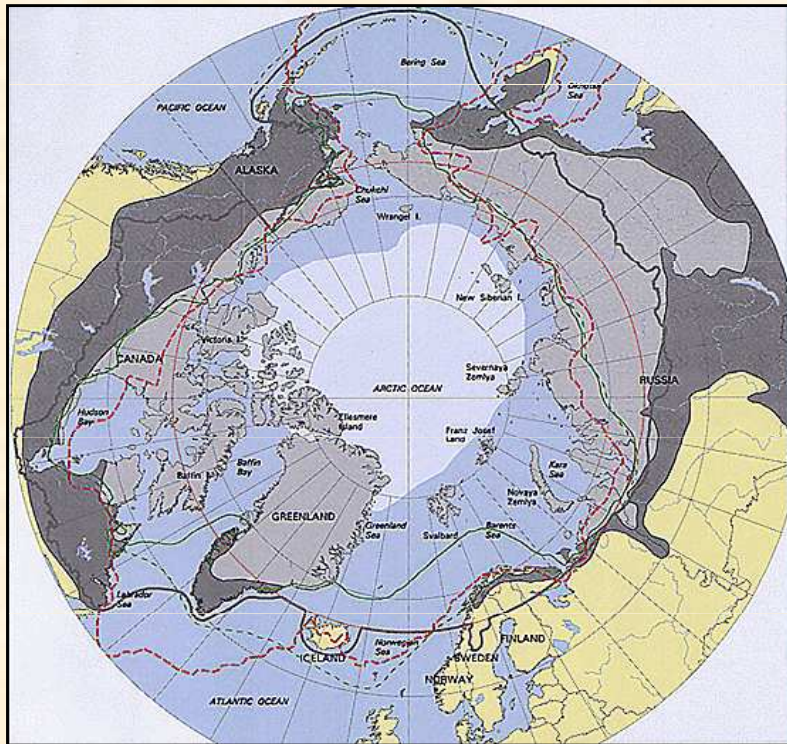
Big Bang v genomu kukuřice

- 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 **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 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:

+

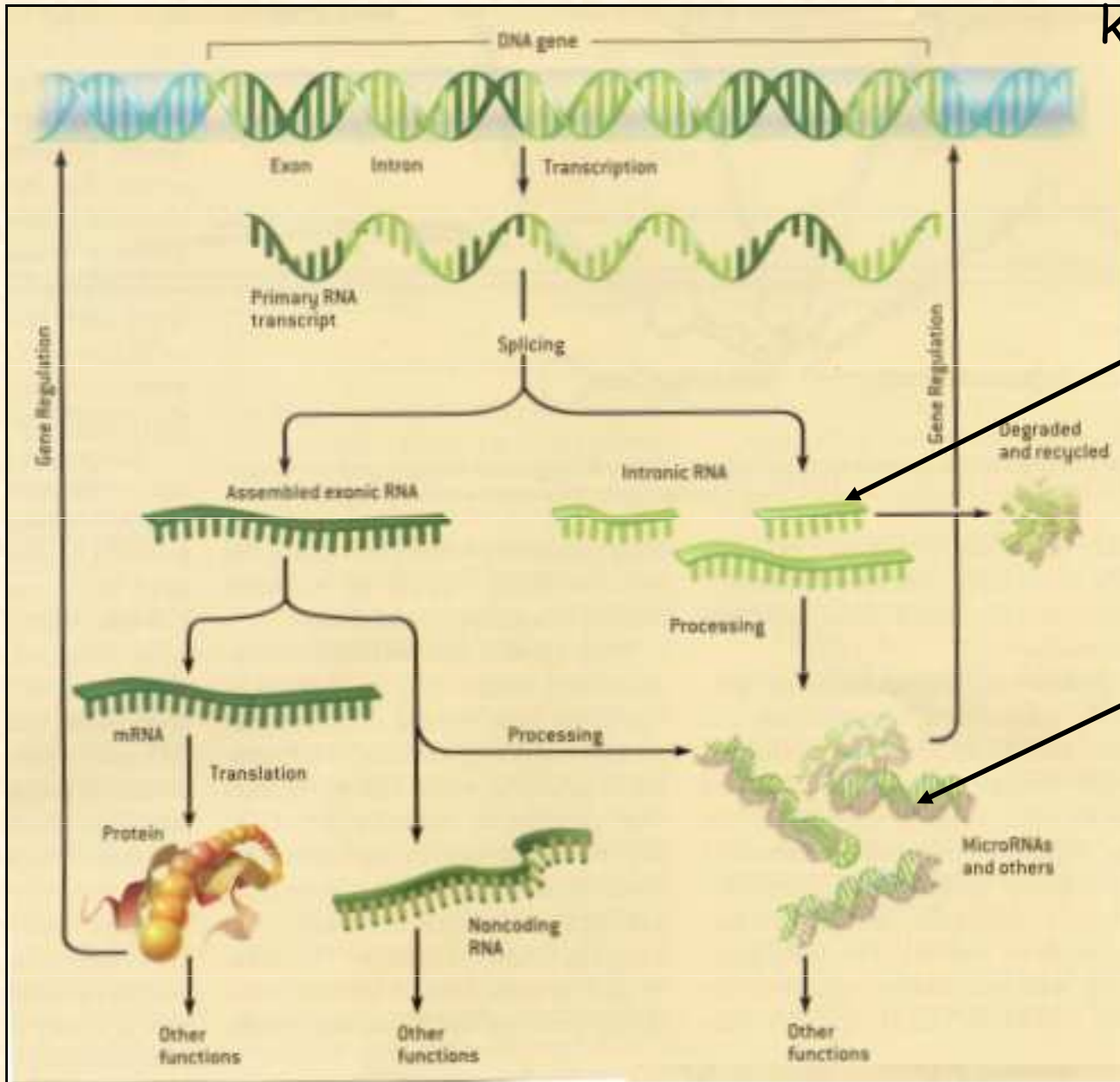
- velikostí jádra
- velikostí buňky (nucleotypic effect)
- dobou mitózy a meiózy
- minimální generační dobou
- velikostí semen
- odpovědí letniček vůči CO_2
- dobou vývoje embrya u mloků

-

- 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)

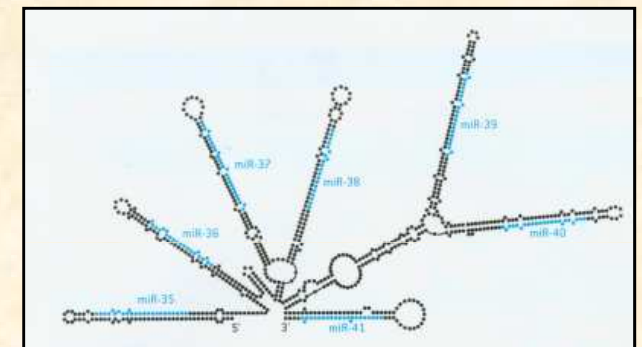
Nekódující DNA: Regulační systém na bázi RNA?

Komplexita vzniká snadno, kontroluje se ale obtížně



introny

MicroRNA



Regulační systém na bázi RNA

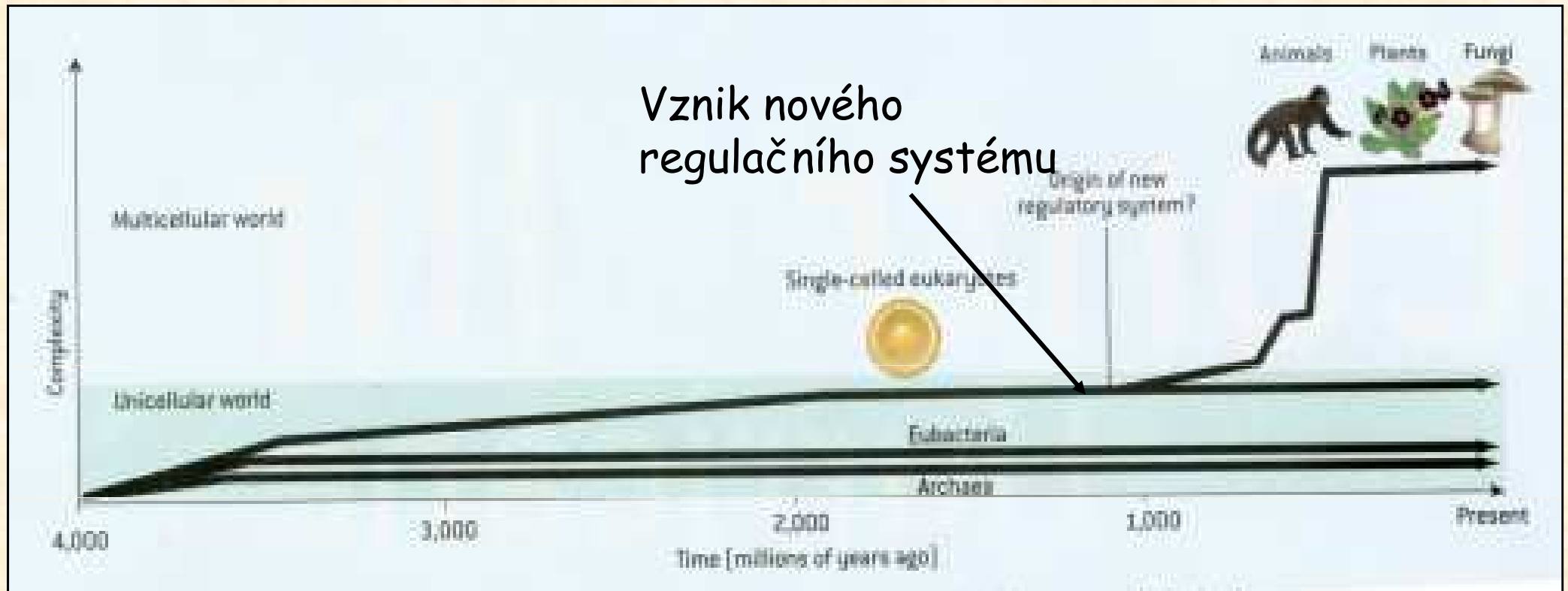


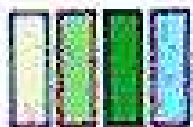
Figure 3 | A simplified biological history of the Earth. This graph is intended to present an overview. Some dates are still being debated and the abscissa ('complexity') has an arbitrary scale. Life appeared on Earth approximately 4,200 million years ago (mya), either arising as, or quickly streaming 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 introns⁷⁵. 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 rock⁷⁸. 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?

TOPOGRAFIE GENOMU

Uspořádání genů v genomech



KEY



Tandem repeats



Centromere-associated tandem repeats



Telomeric and sub-telomeric repeats



Dispersed *Ty1/copia* retroelements and microsatellites

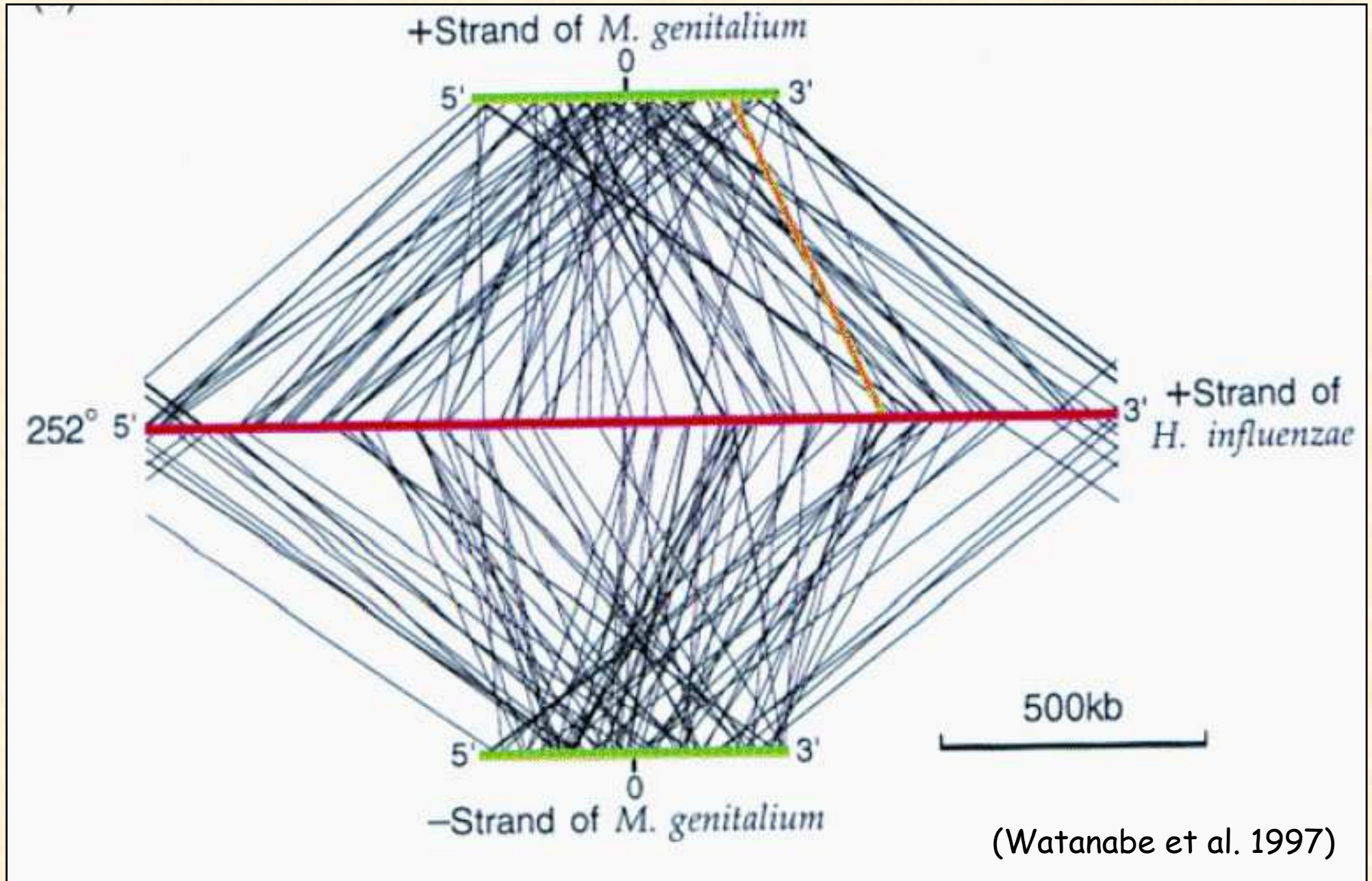


LINEs



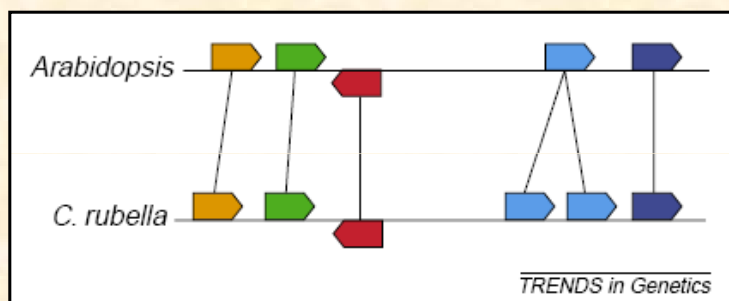
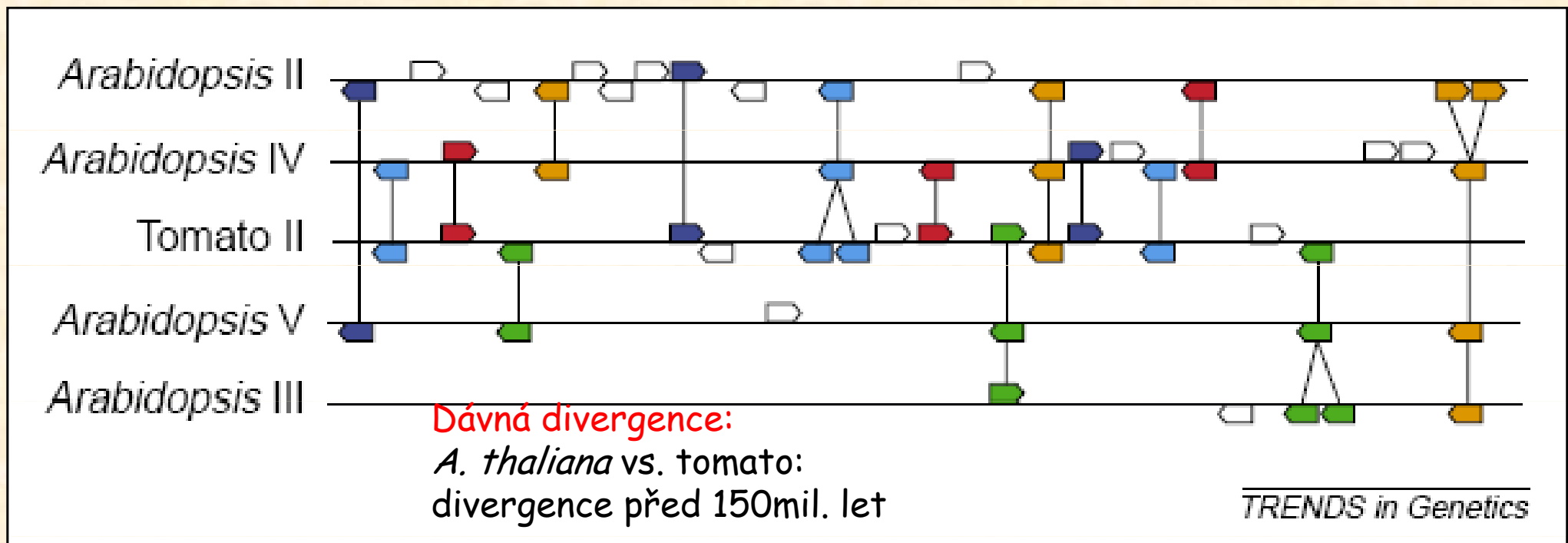
Single and low-copy sequences including genes

Přestavby pořadí genů u bakterií



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

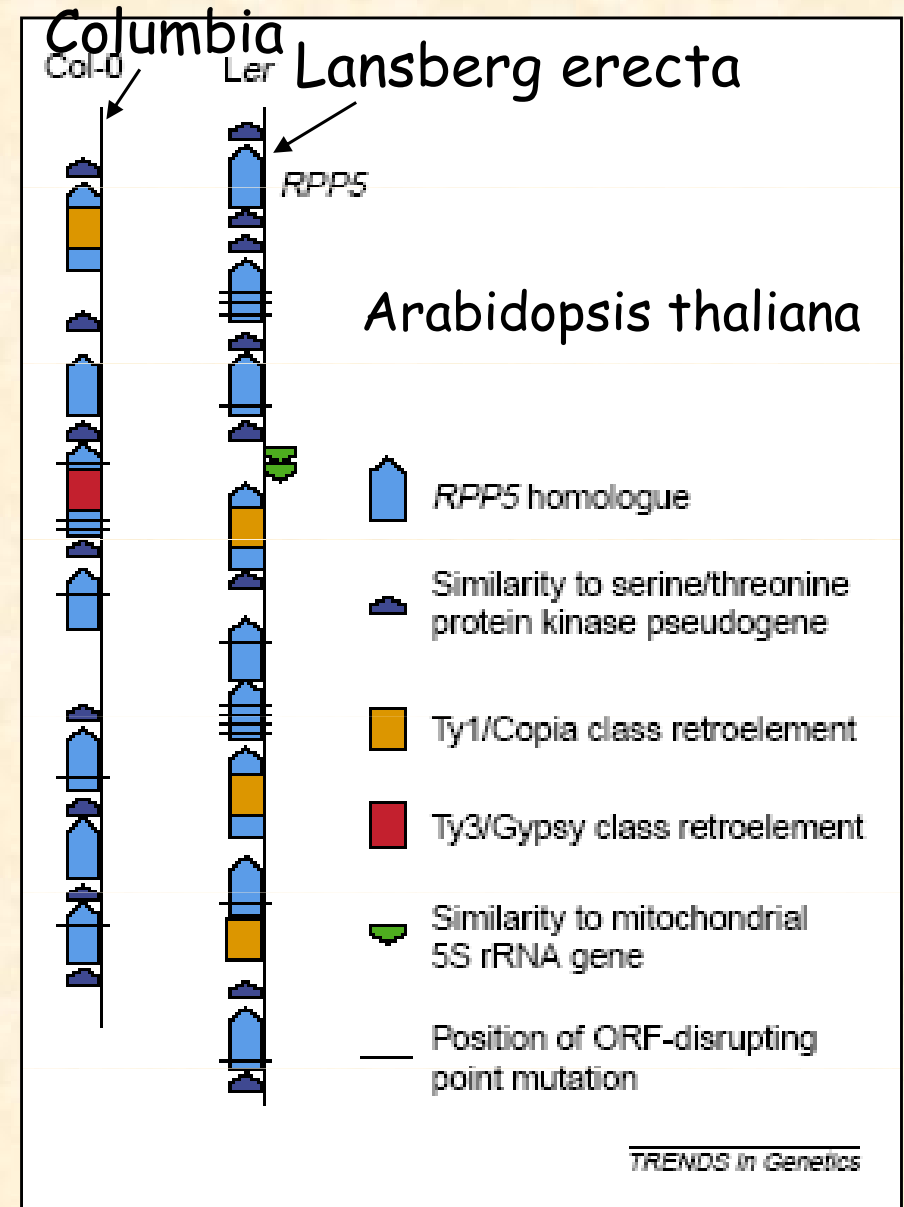


(Bancroft, 2001)

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

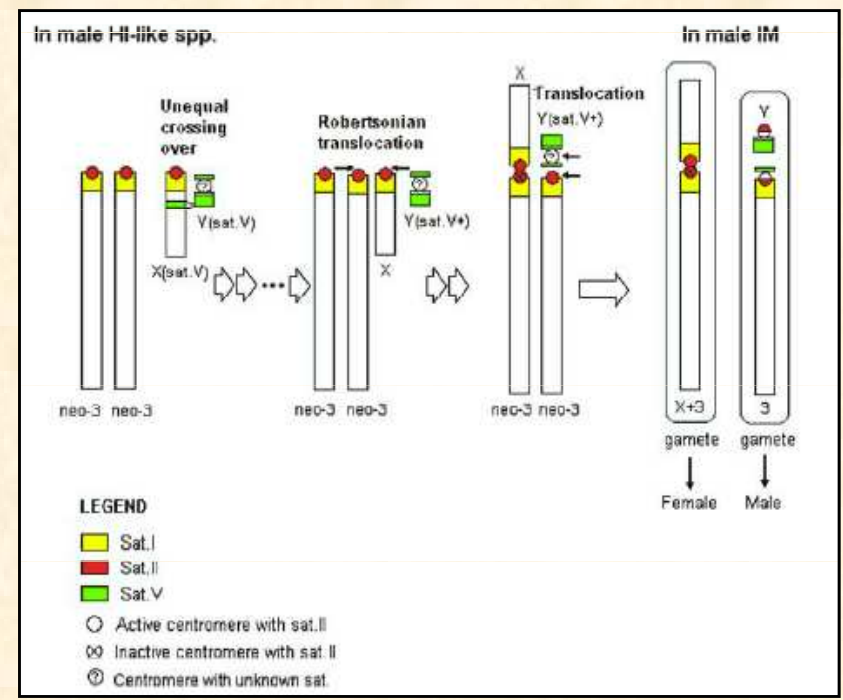
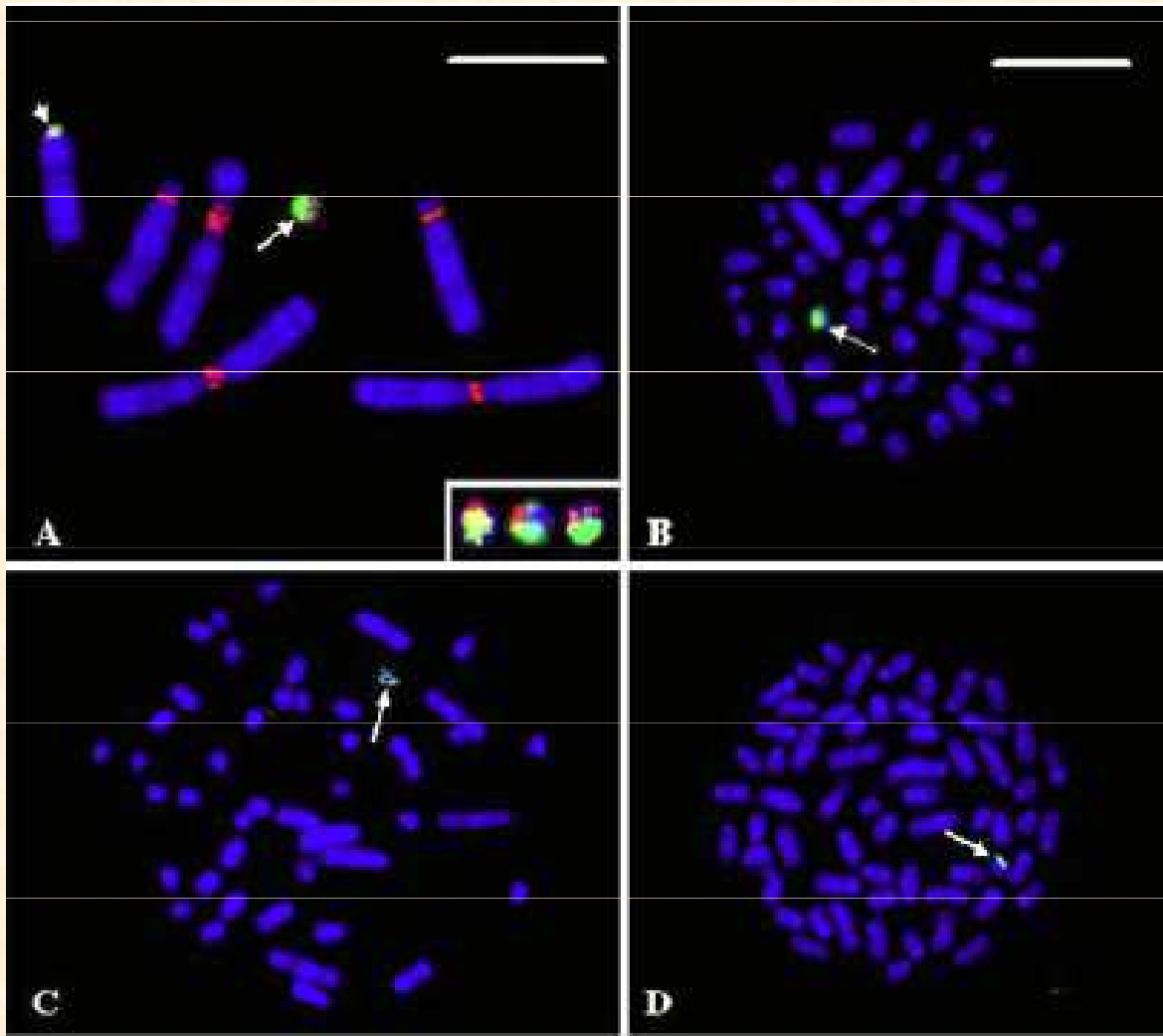
Common Name	Species	Diploid number	Common Name	Species	Diploid number
Animals (2n)			Plants (2n)		
Human	<i>Homo sapiens</i>	46	Corn	<i>Zea mays</i>	20
Monkey	<i>Macaca mulatta</i>	42	Potato	<i>S. tuberosum</i>	48
Dog	<i>Canis familiaris</i>	78	Green algae	<i>A. mediterranea</i>	20
Cat	<i>Felis domesticus</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:

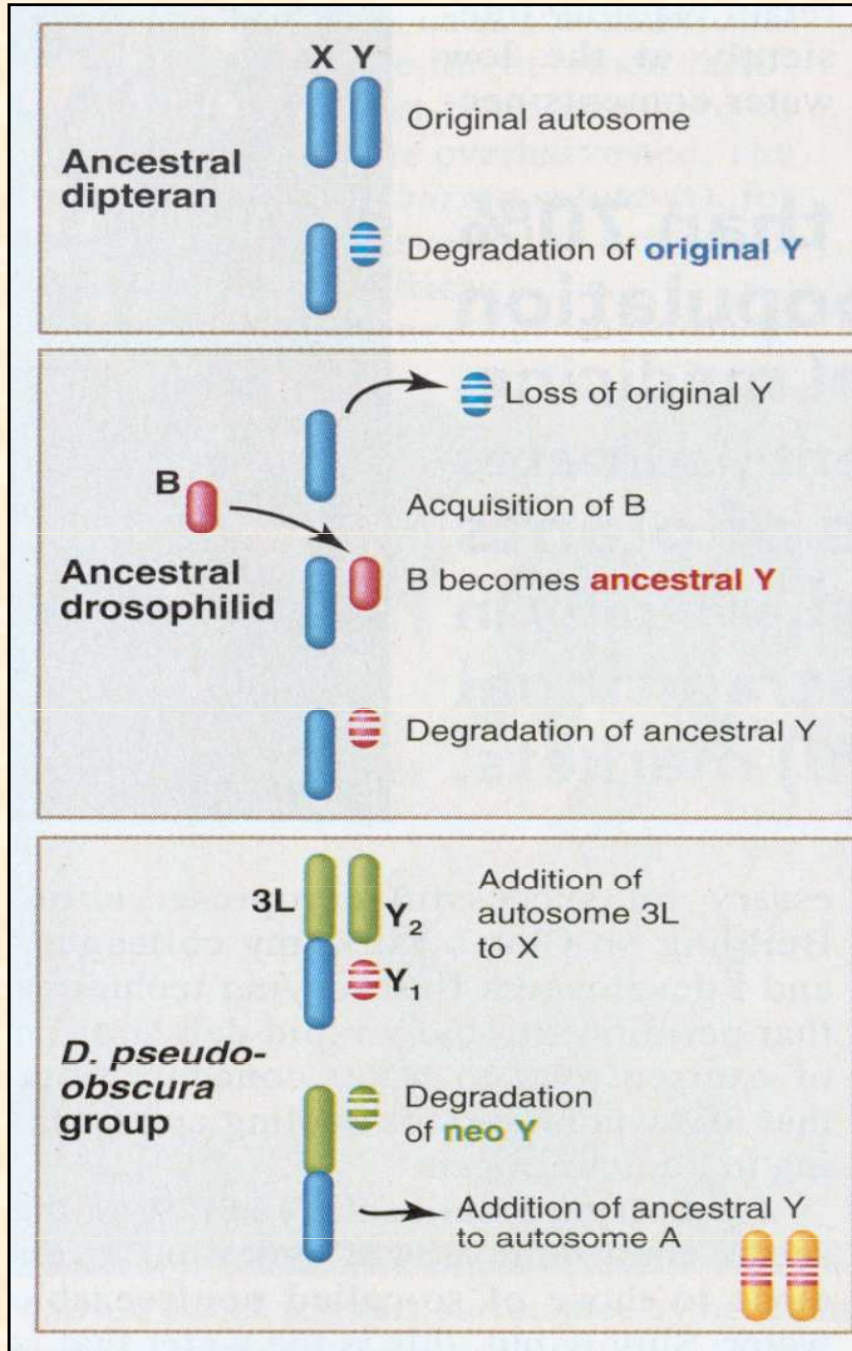
mravenec *Myrmecia pilosula* - 1 pár, samec 1 chromosom
 kapradina *Ophioglossum reticulatum* - 630 párů

Různé počty chromosomů u blízce příbuzných druhů jelíneků *Muntiacus*

různé druhy



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

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:

- vyjíměčně 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:

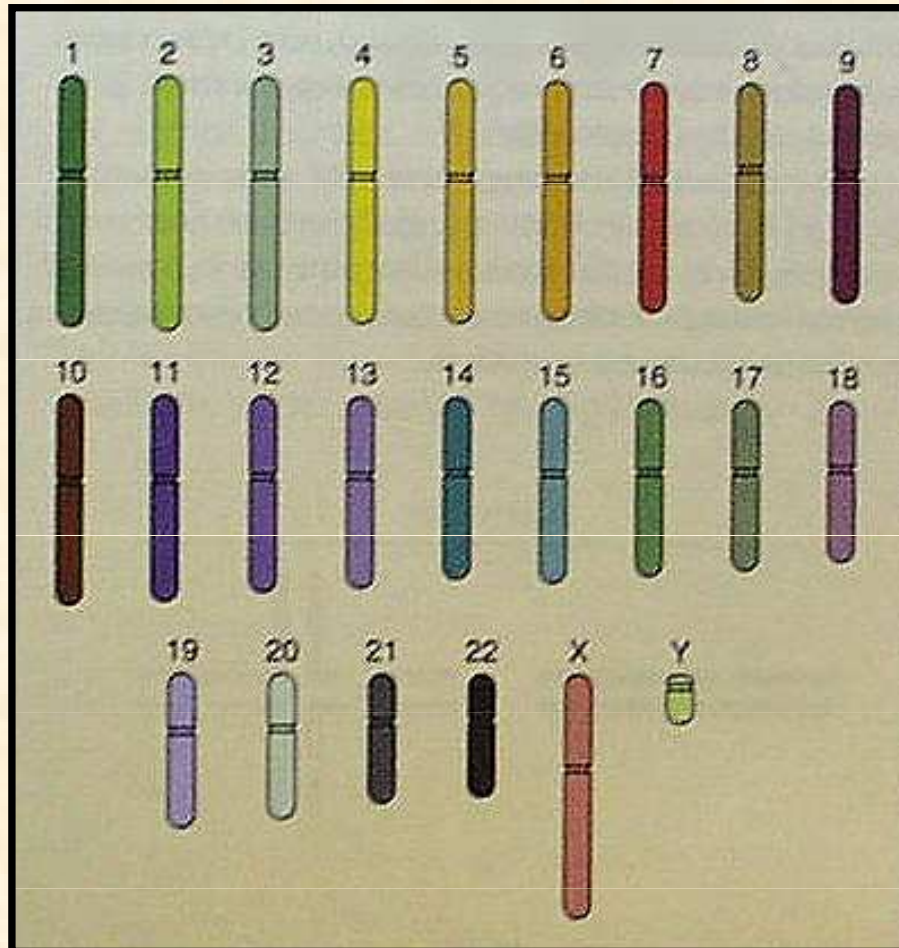
- velikost průměrných chromosomů, extrémy - největší (ryba *Astyanax scabripinnis*) nebo nejmenší (myš *Reithrodontomys megalotis*)
- metacentrické nebo akrocentrické, heterochromatické,
- obsahují repetice, satelitní DNA, rDNA, transpozony (mikrodisekce a PCR), nerekombinují

Chování v meióze:

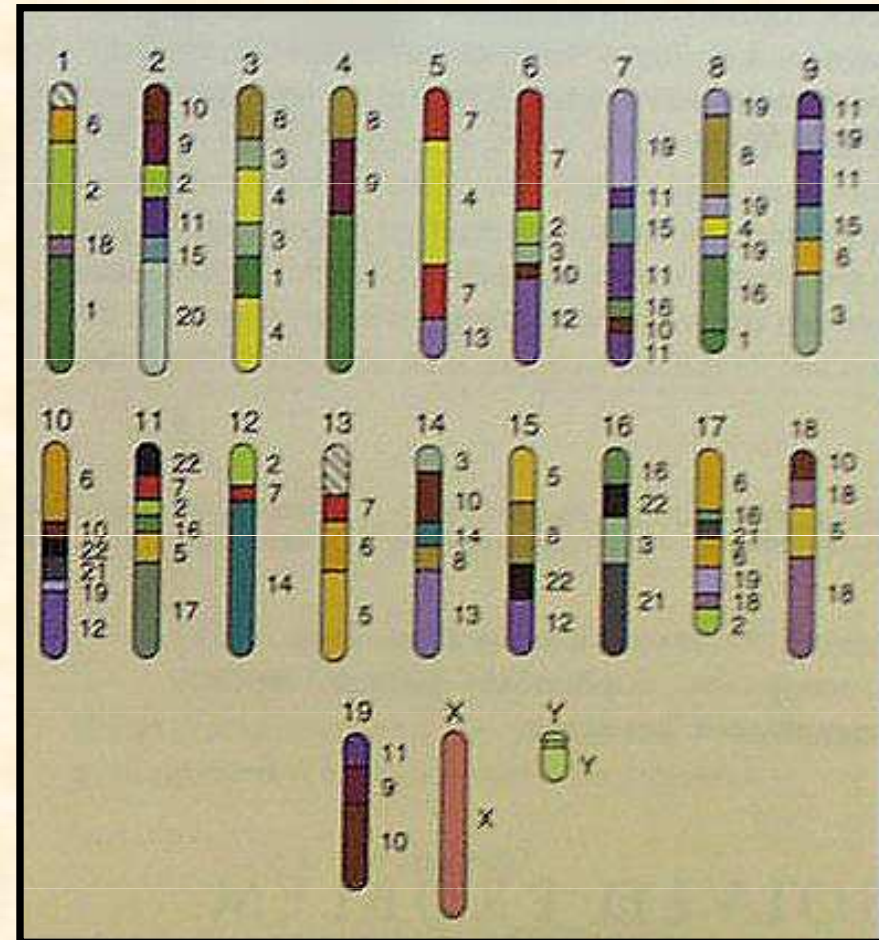
- nepárujíc se s partnerem, univaletní, segregují přednostně k určitému pólu, podobnost s pohlavními chromosomy, MSCI,

Srovnání lidských a myších chromosomů

člověk

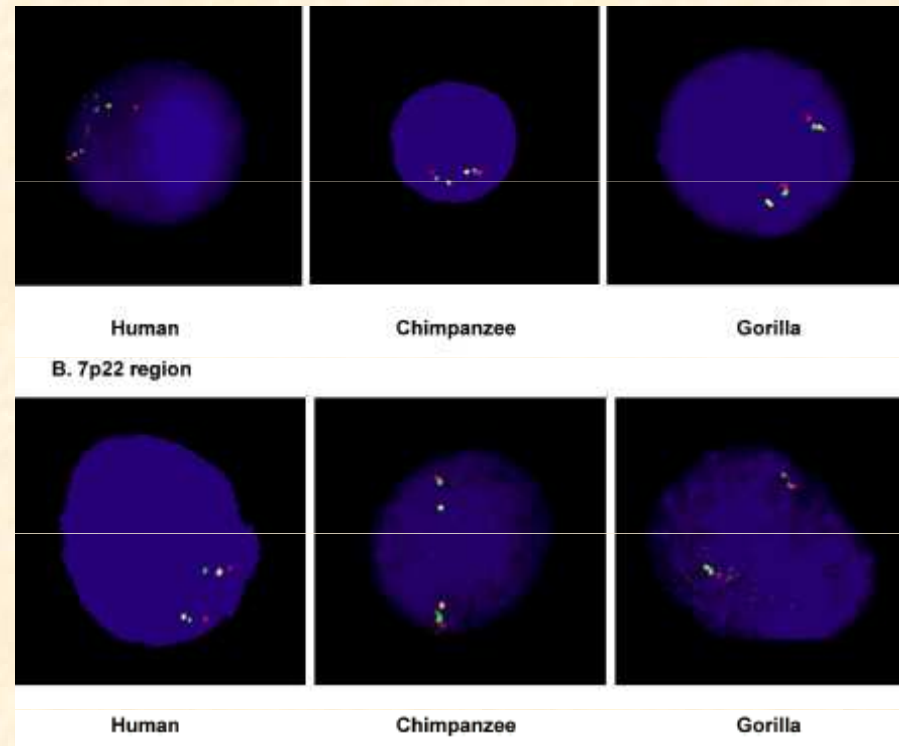
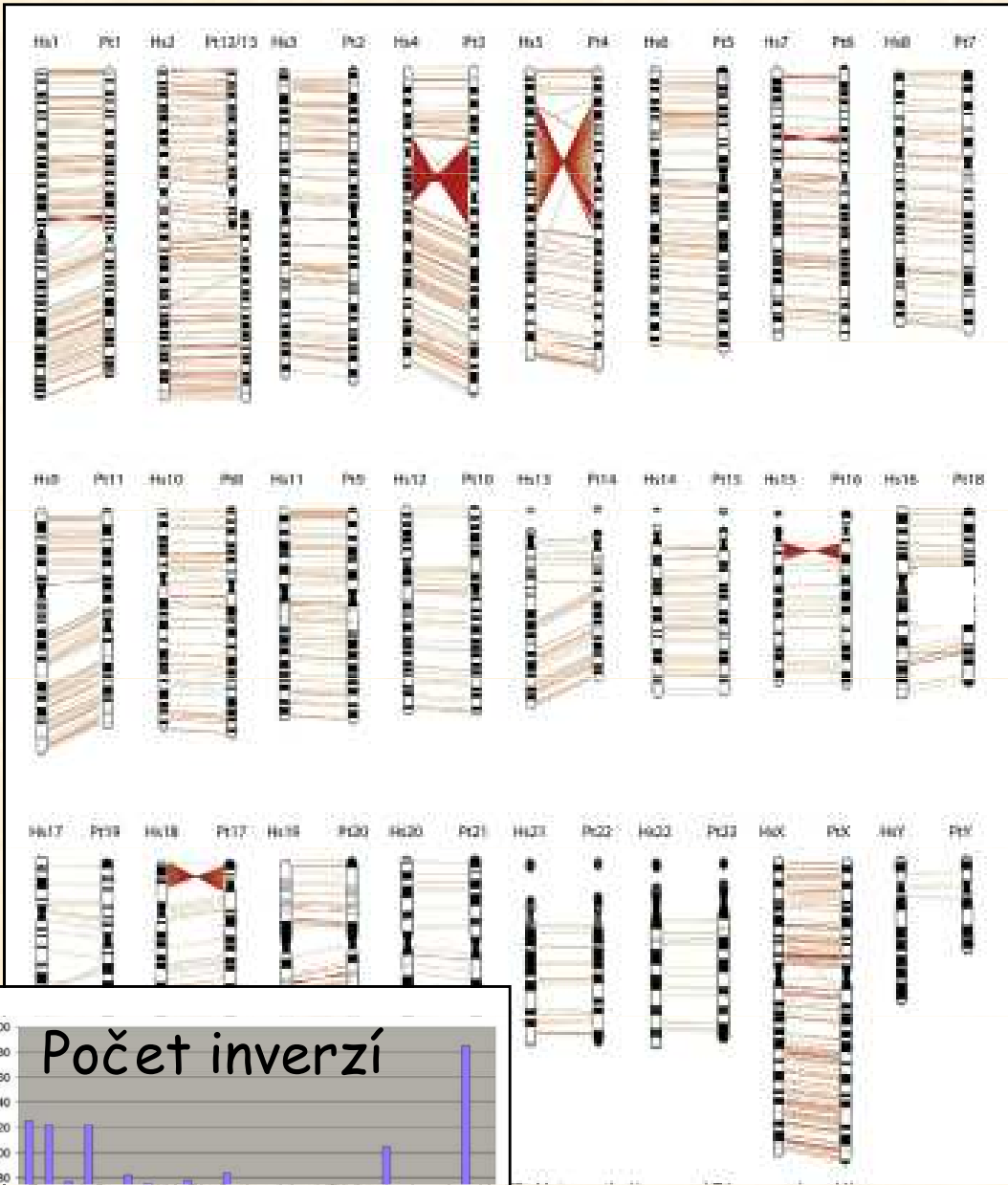


myš

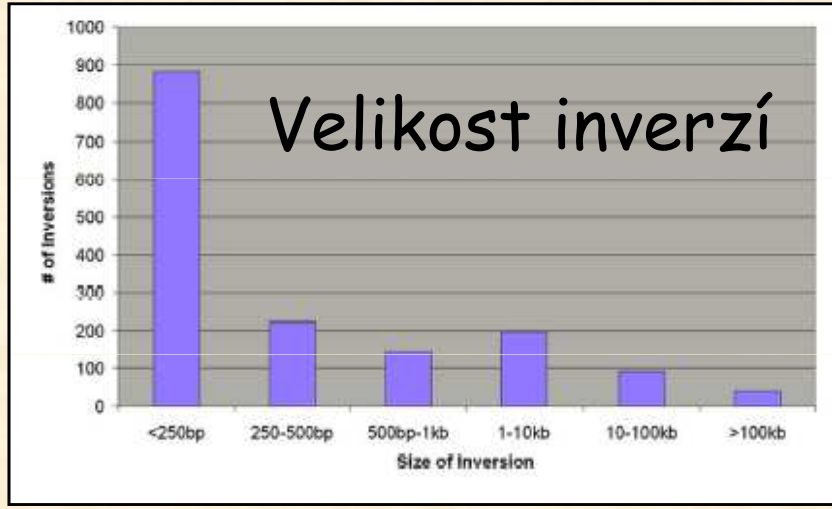


Od evoluční divergence došlo k četným přestavbám, které umístily bloky genů do různých kombinací.

Inverze v genomech člověka a šimpanze



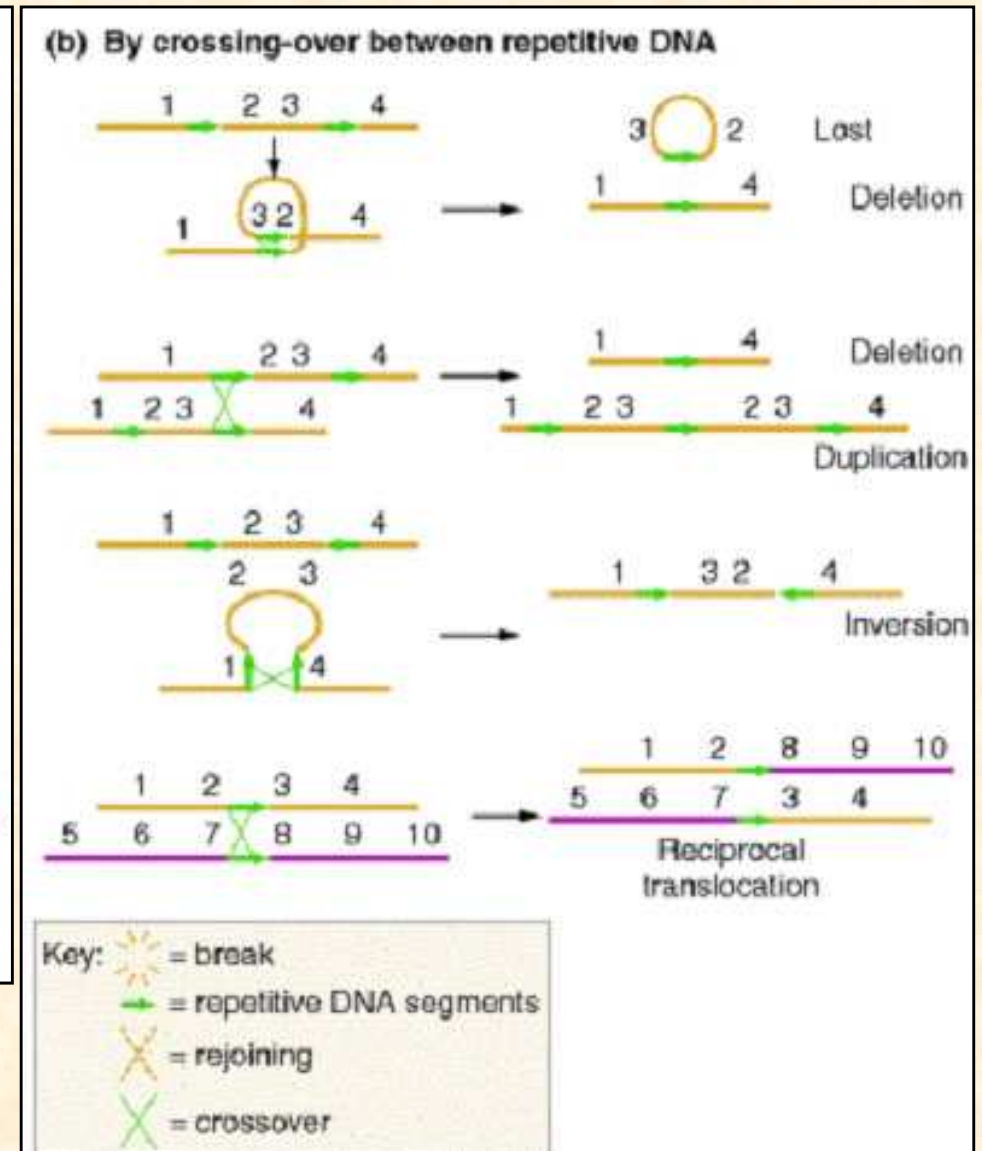
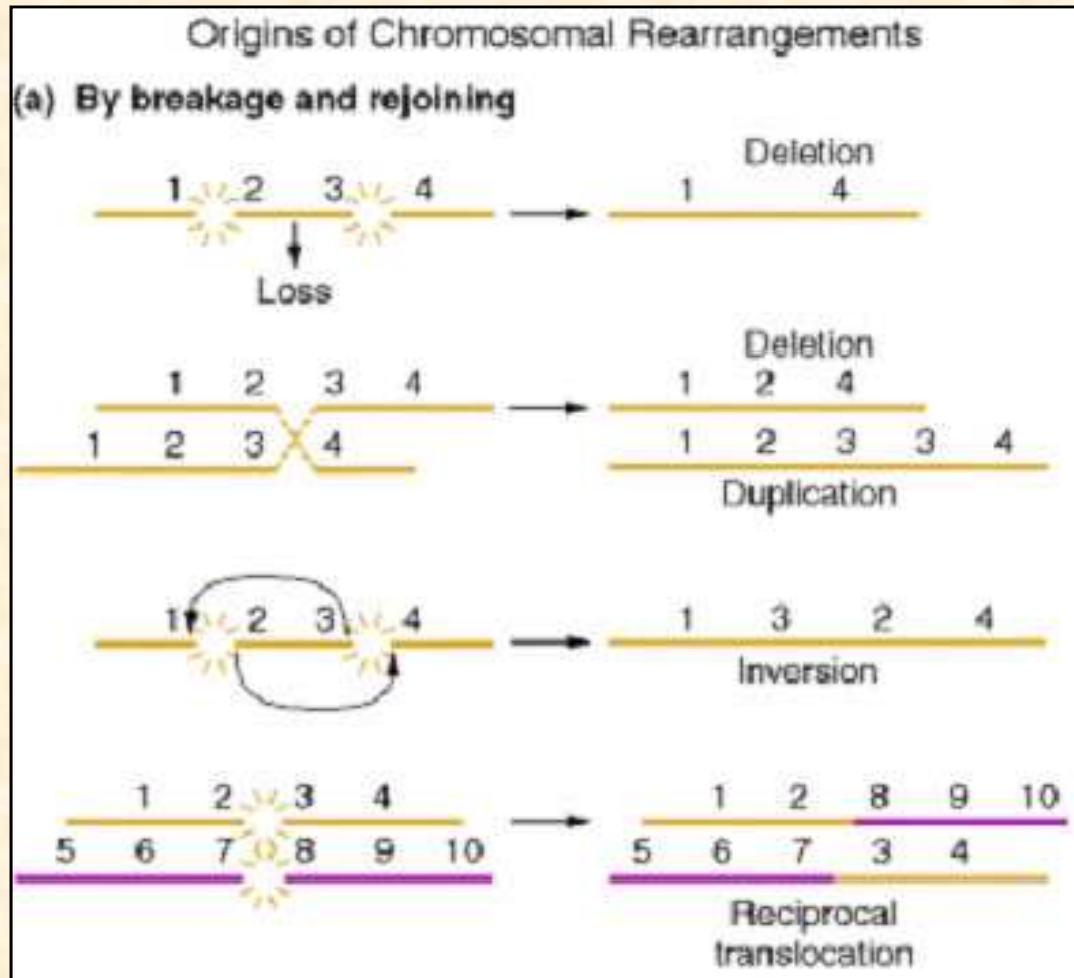
Člověk: červená-žlutá-zelená
 Šimpanz a gorila: červená-zelená-žlutá



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.

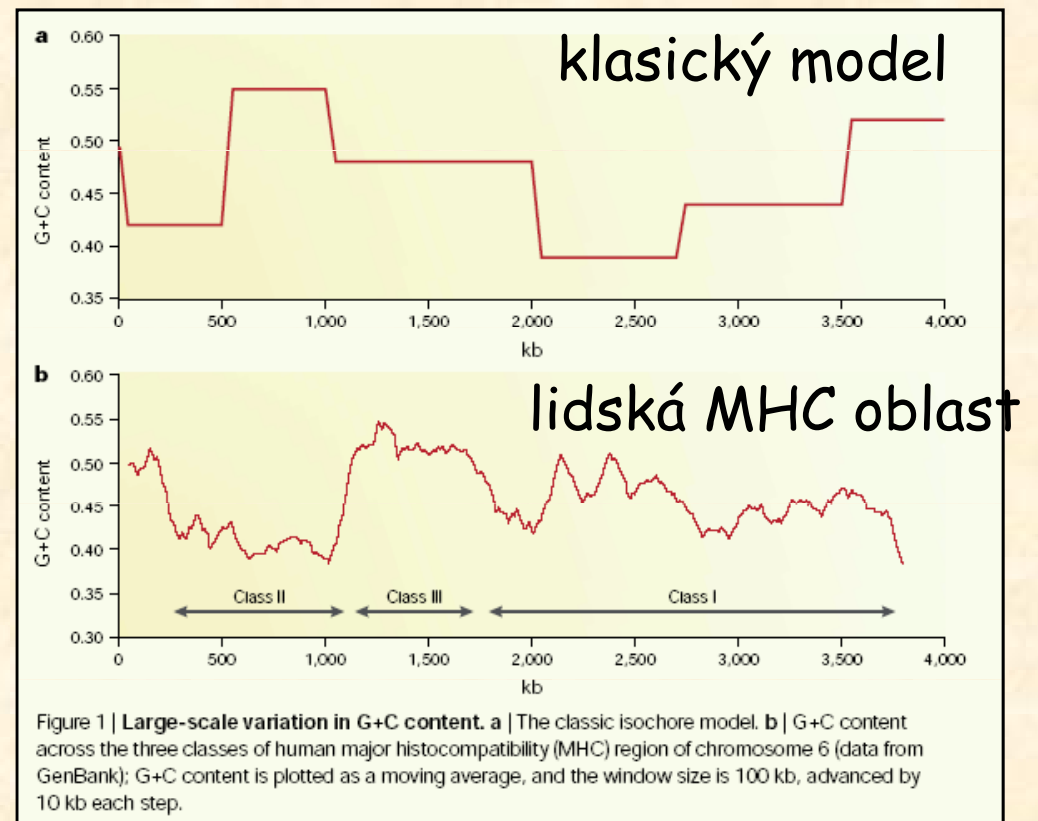
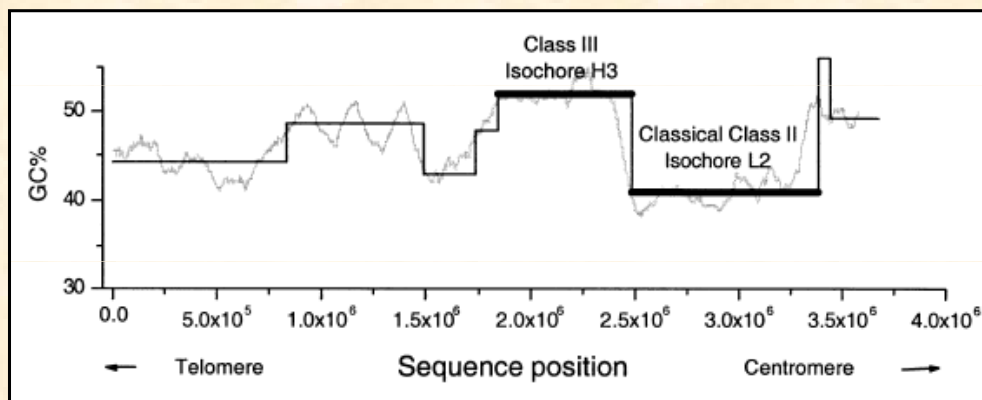
Izochorový model organizace genomu

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

- 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:

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



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 organizmus -
kvasinka *Saccharomyces cerevisiae*
Goffeau et al., Science 274, 563-567.



Genomové projekty

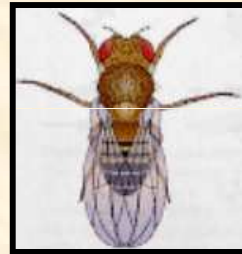


1996: *Saccharomyces cerevisiae* (12 Mb/ 6548 genes)



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

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



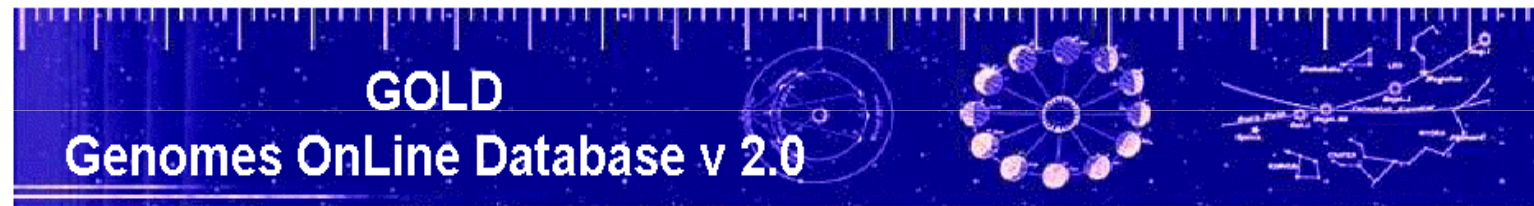
Mus musculus
(3 300 Mb)



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



Jak zjistit stav genomových projektů



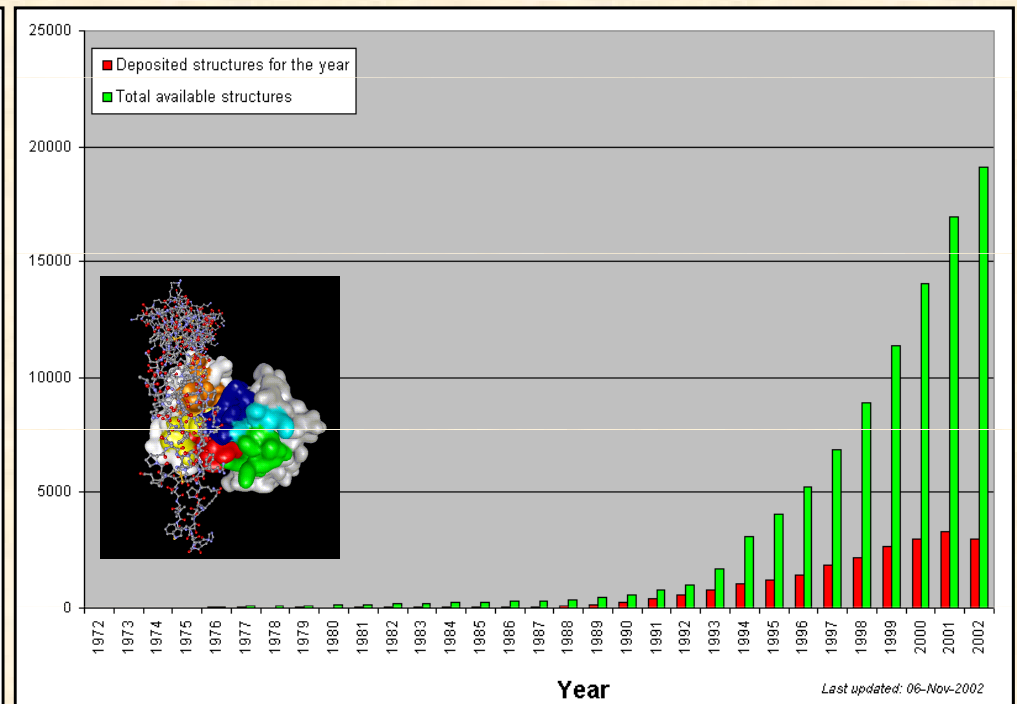
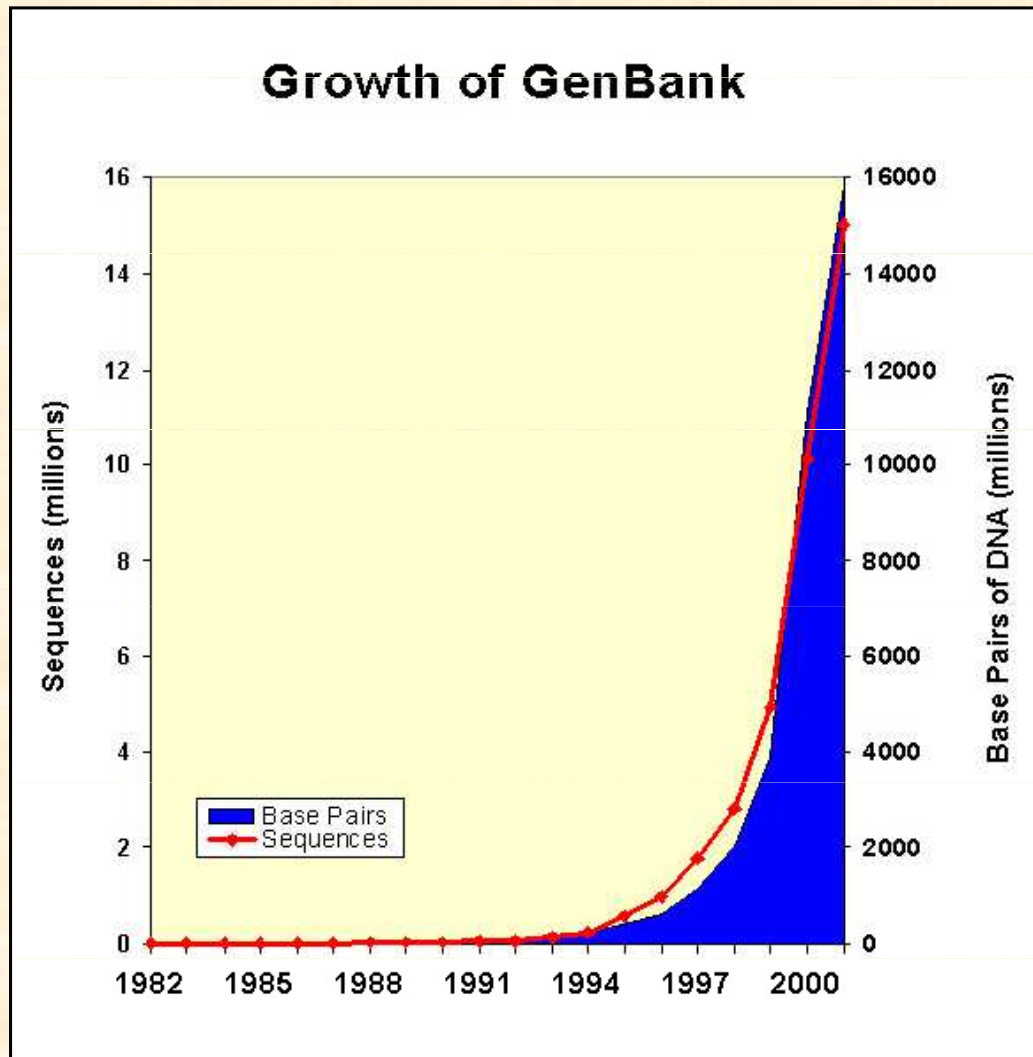
Contact: Genomesonline	Last Update: April 2, 2008	Location www.genomesonline.org
762 Published Complete Genomes	Search GOLD : 3658 genome projects	123 Metagenomes
89 Archaeal Ongoing Genomes	1749 Bacterial Ongoing Genomes	935 Eukaryotic Ongoing Genomes
	Right-click to save all data: DOWNLOAD	

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

Objem dat v databázích roste exponenciálně

>60 000 000 000 bp
>55 000 000 sekvencí

> 20 000 struktur



~ 165 000 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

REPBASE - repetitivní sekvence

Genomové sekvenační projekty

duben 2008

762 ukončených projektů:

53 archea
626 bakterie
83 eukaryota

975 probíhajících projektů:

1749 prokaryota
935 eukaryota
89 archea

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)

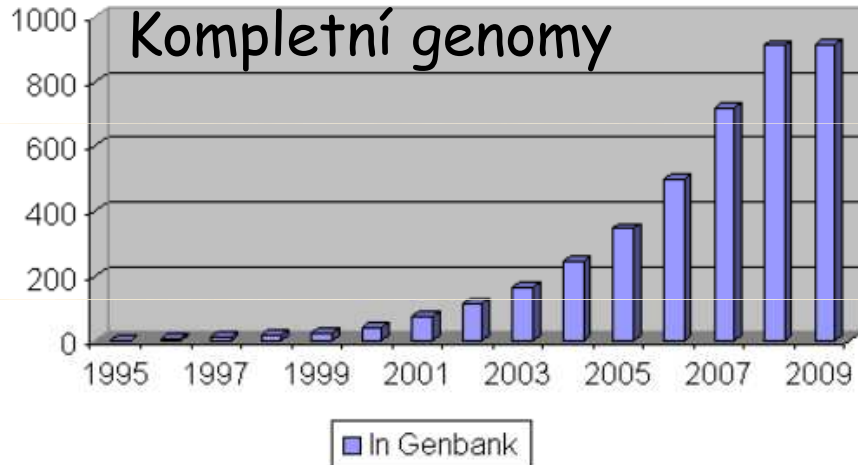
Rattus norvegicus (2750 Mb)

plants: *Arabidopsis thaliana* (115 Mb)

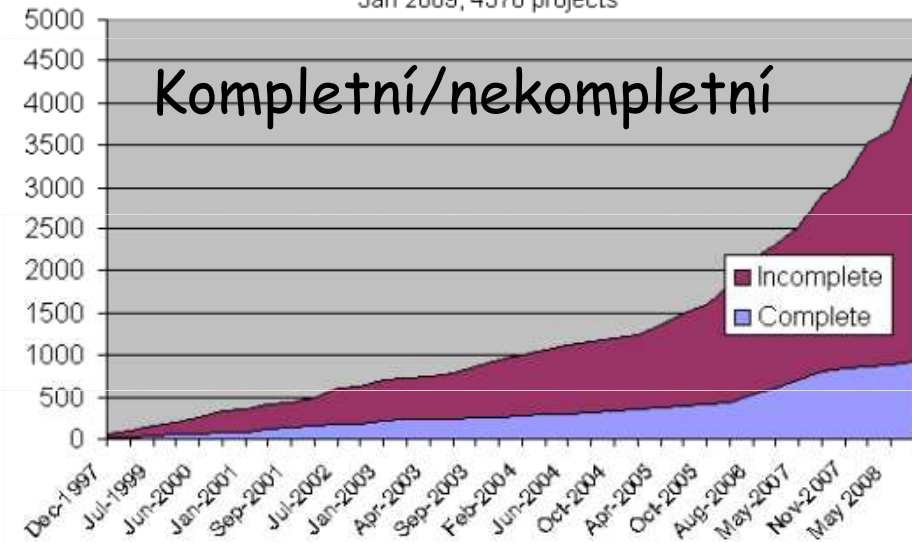
Oryza sativa (420 Mb)

Genomové sekvenační projekty

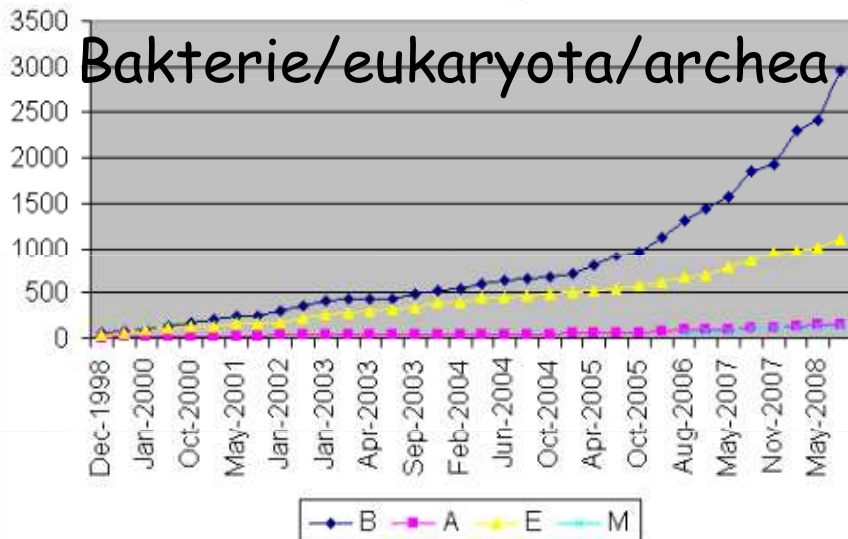
Completely Sequenced Genomes ©
January 2009



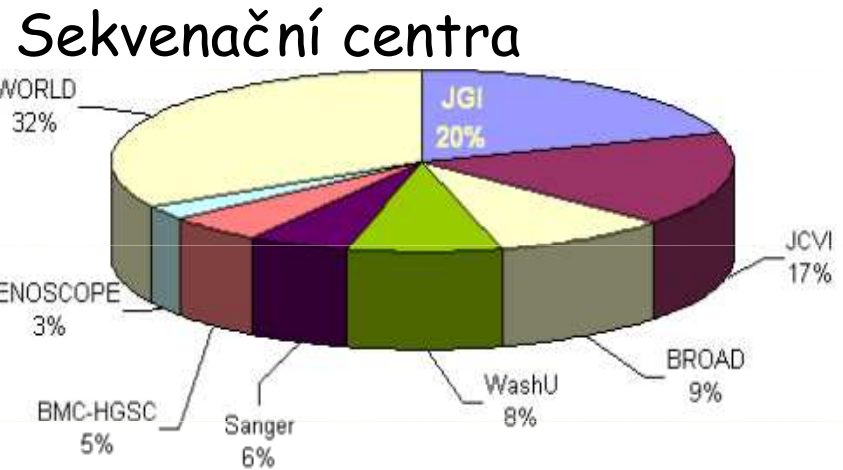
Genome Sequencing Projects on GOLD ©
Jan 2009, 4370 projects



Genome Projects on GOLD according to Phylogenetic Groups ©
Jan 2009 - 4370 Projects



Major Sequencing Centers ©
January 2009: 4370 projects



Sekvenování lidského genomu

