

Evolve bakterialních genomů

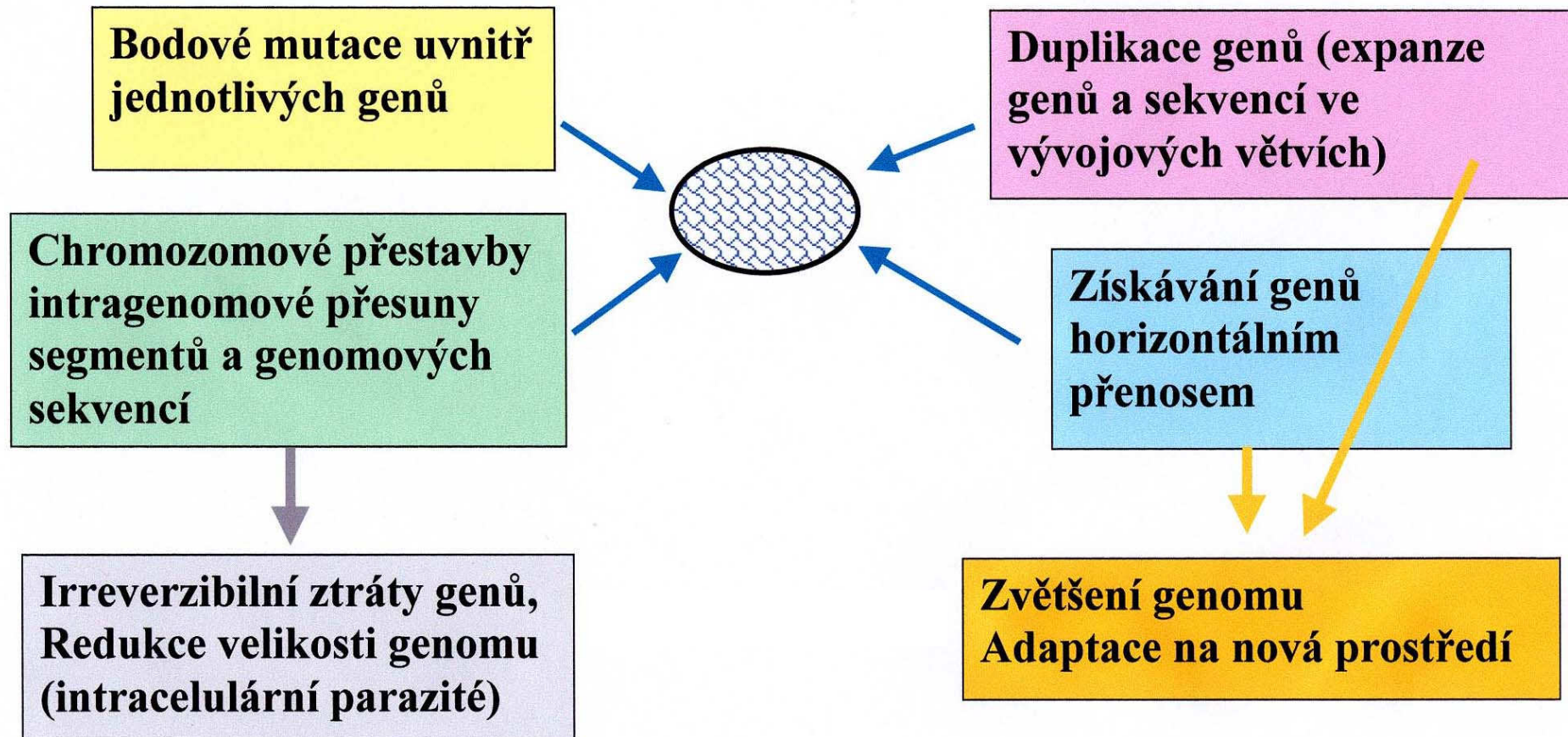
Charakteristické rysy:

Rychlé a rozsáhlé změny ve struktuře a informačním obsahu genomu (plasticita, dynamické změny)

- Vnitřní přestavby
- Získávání a ztráty genů a genetických elementů

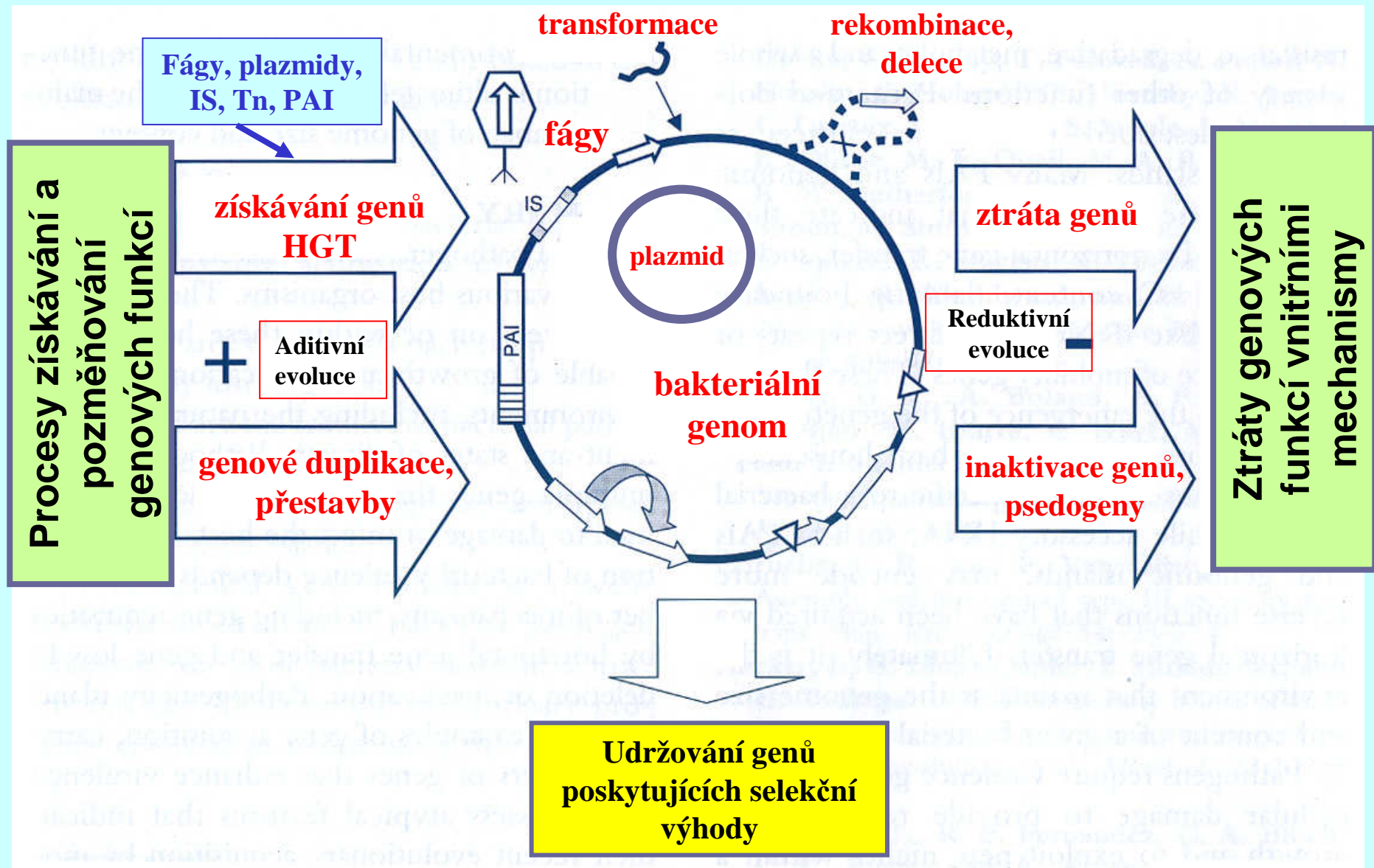
Vývoj kmenů v rámci druhu - adaptace na nové podmínky

Podstata změn v obsahu genomu prokaryot v průběhu evoluce



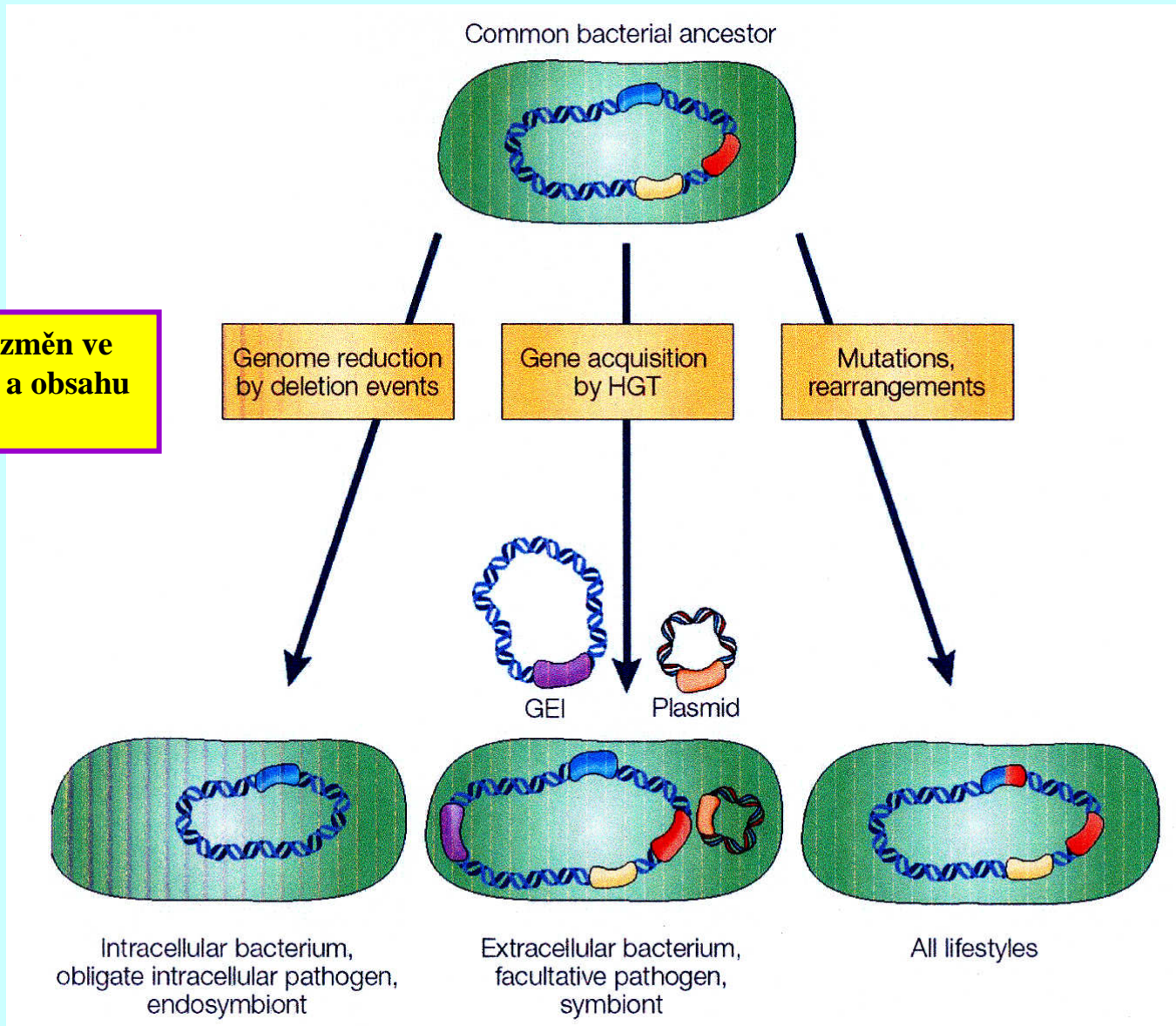
Důsledek: Pořadí genů je zachováno jen u velmi blízké příbuzných druhů

Mechanismy evoluce bakteriálních genomů



Struktura genomu odráží životní styl bakterie

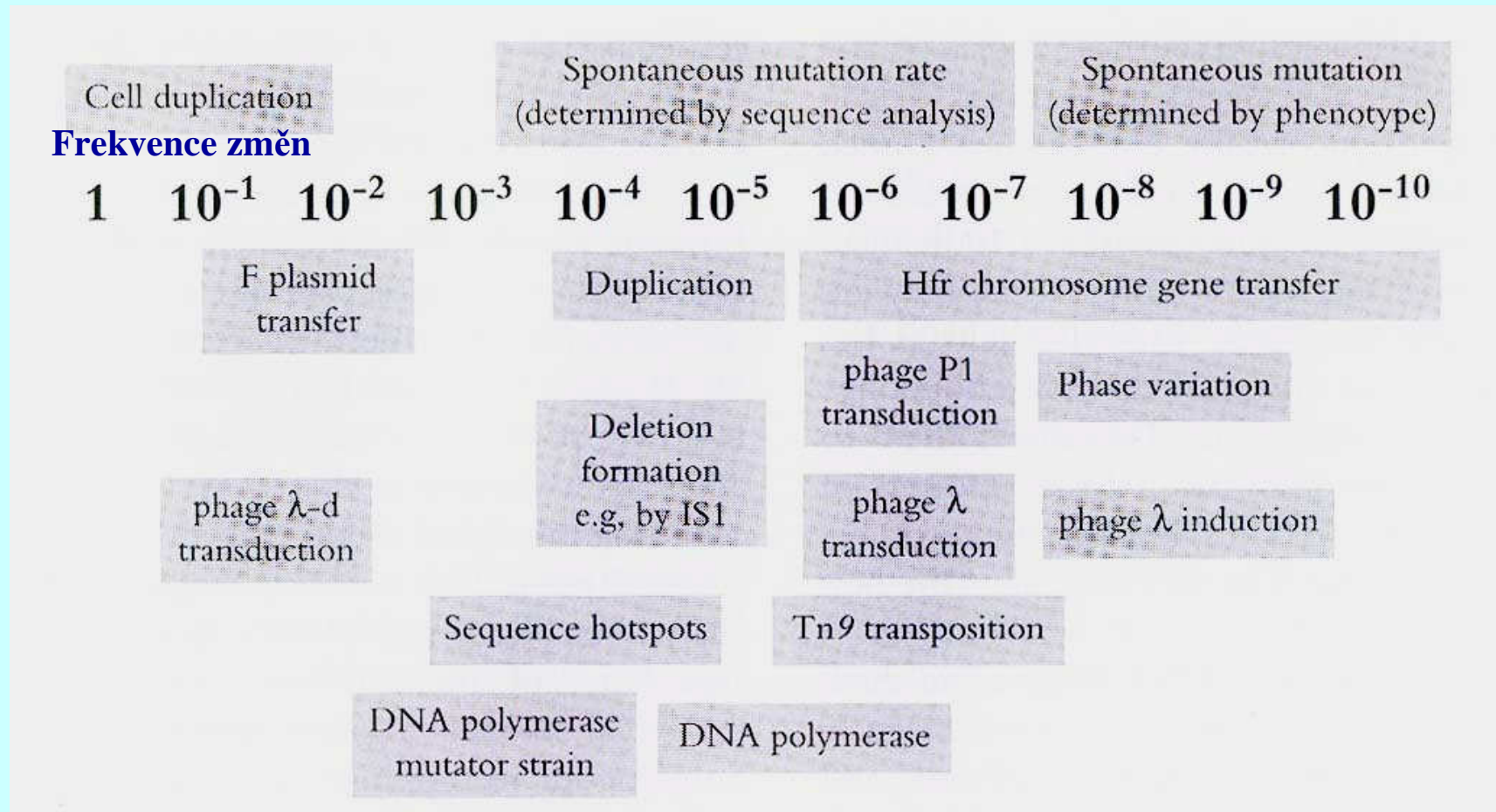
Příčiny změn ve velikosti a obsahu genomu



Mechanismy odpovědné za plasticitu genomu

Genetický element nebo mechanismus	Důsledky
A. Získ vlastností	
Bodové mutace	Změna genové exprese
Homologií rekombinace	Přeskupení DNA, inverze, duplikace, delece DNA Integrace DNA přenesené HGT
Transformace	Získání přídatné genetické informace
IS elementy, transpozony	Inzerce, delece, inverze DNA, změny genové exprese
Integrony	Přenos genů, přeskupení DNA
Konjugativní transpozony, plazmidy	Konjugace, HGT, mobilizace jiných elementů (plazmidů, chromozomu)
Bakteriofágy	HGT, transdukce, fágová konverze
GTA, VTA	HGT, kapsdukce
Genomové ostrovy a ostrůvky	HGT, integrace a delece velkých úseků DNA
B. Ztráta vlastností	
Bodové mutace	Změny v genové expresi, ztráta funkce genů
Homologní rekombinace	Přeskupení DNA, delece DNA, integrace genů získaných HGT
Transpozice	Změny genové exprese, ztráta funkce genů

Spektrum faktorů podílejících se na změnách genomu u kmenů E. coli



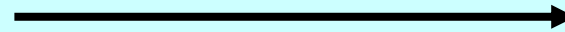
Vnitřní přestavby replikonů navozené přítomností repeticí

Typy repeticí

- geny rRNA a tRNA
- Inzerční sekvence
- transpozony
- krátké repetice
- rhs a Chi-sekvence

Typ přestavby

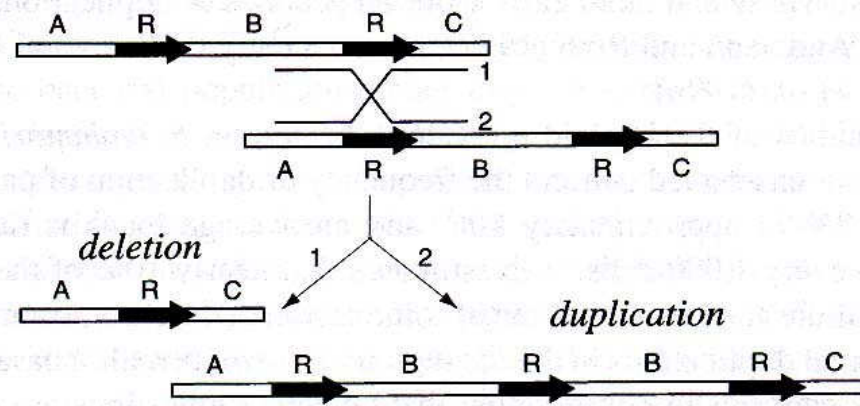
- Duplikace (amplifikace)
- Delece
- Inverze



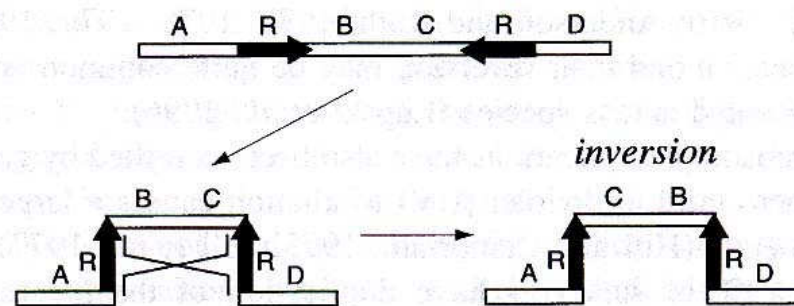
Mechanismus

Homologní rekombinace
Transpozice
Místně-specifická rekombinace
Nerovnoměrný crossing-over

přímá opakování (repetic)



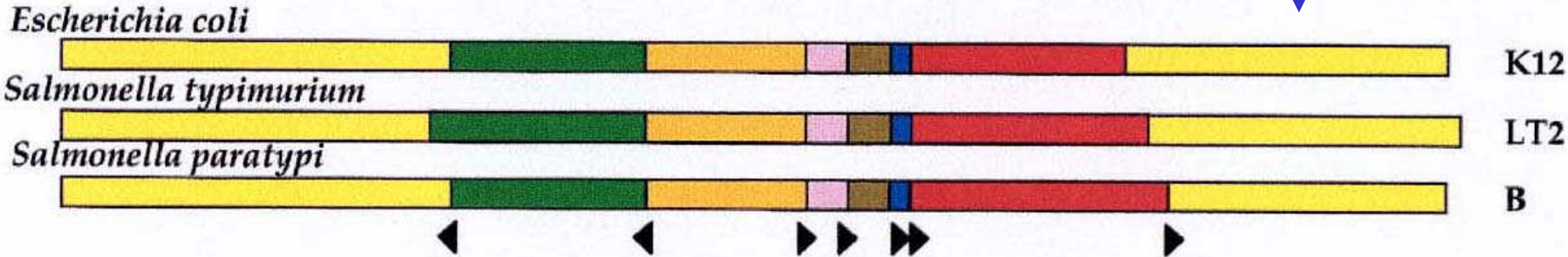
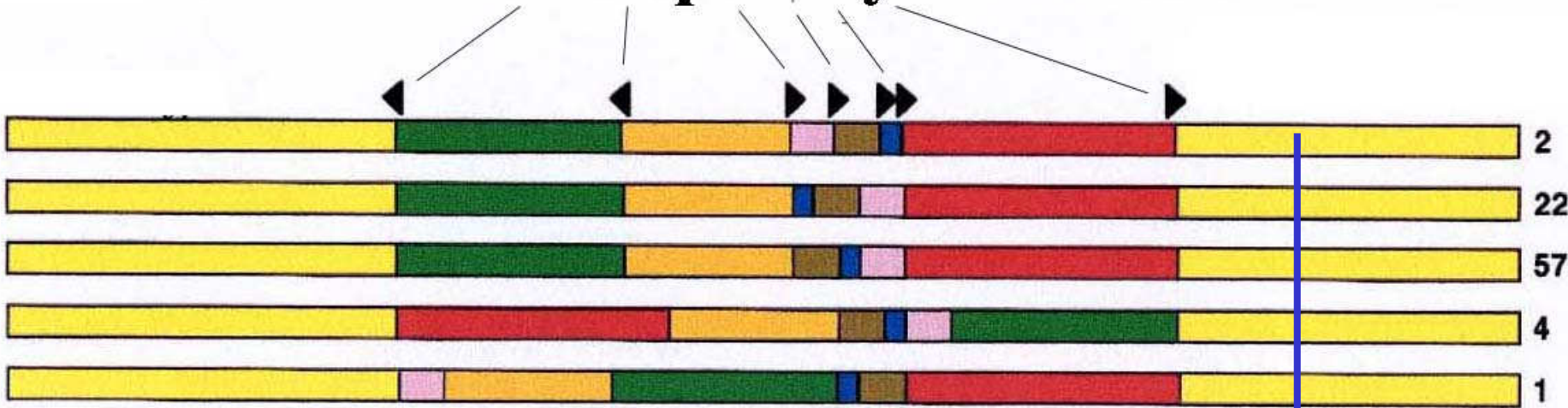
obrácená opakování (repetic)



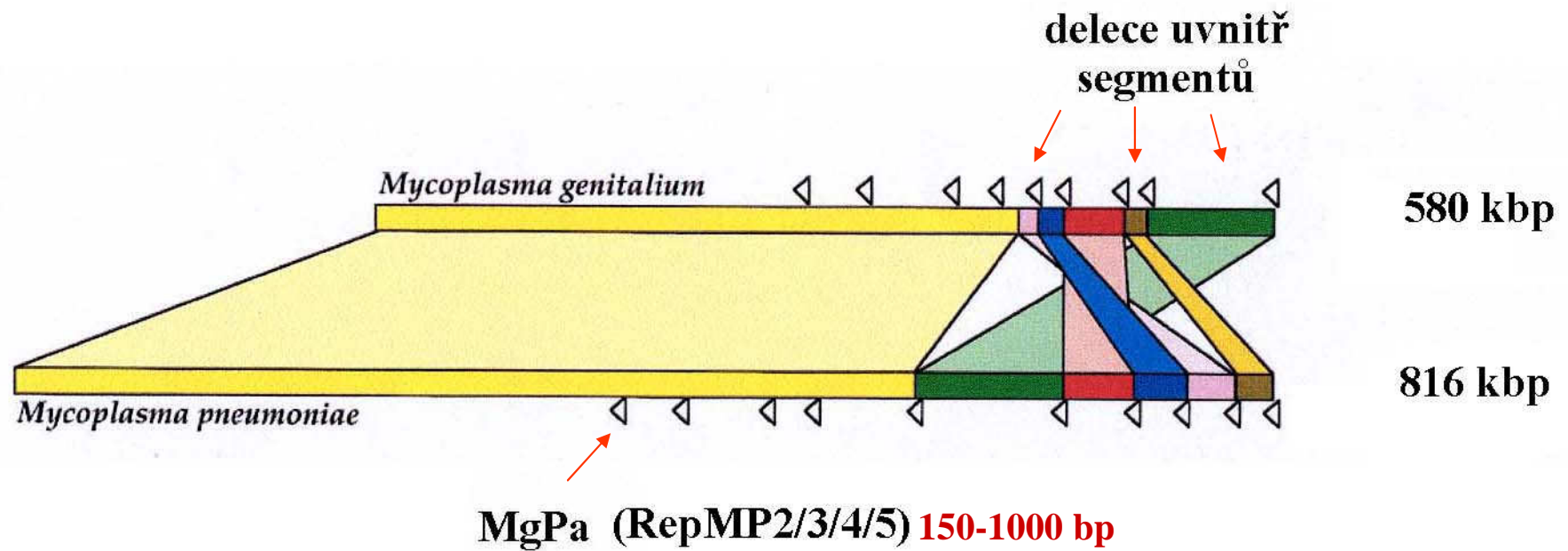
**Přestavby navozené interakcí repetic
(homologní rekombinace)**

Genomová přeskupení navozená rrn operony u kmenů Salmonella typhi

rrn operony



Rozdíly ve struktuře genomů příbuzných druhů



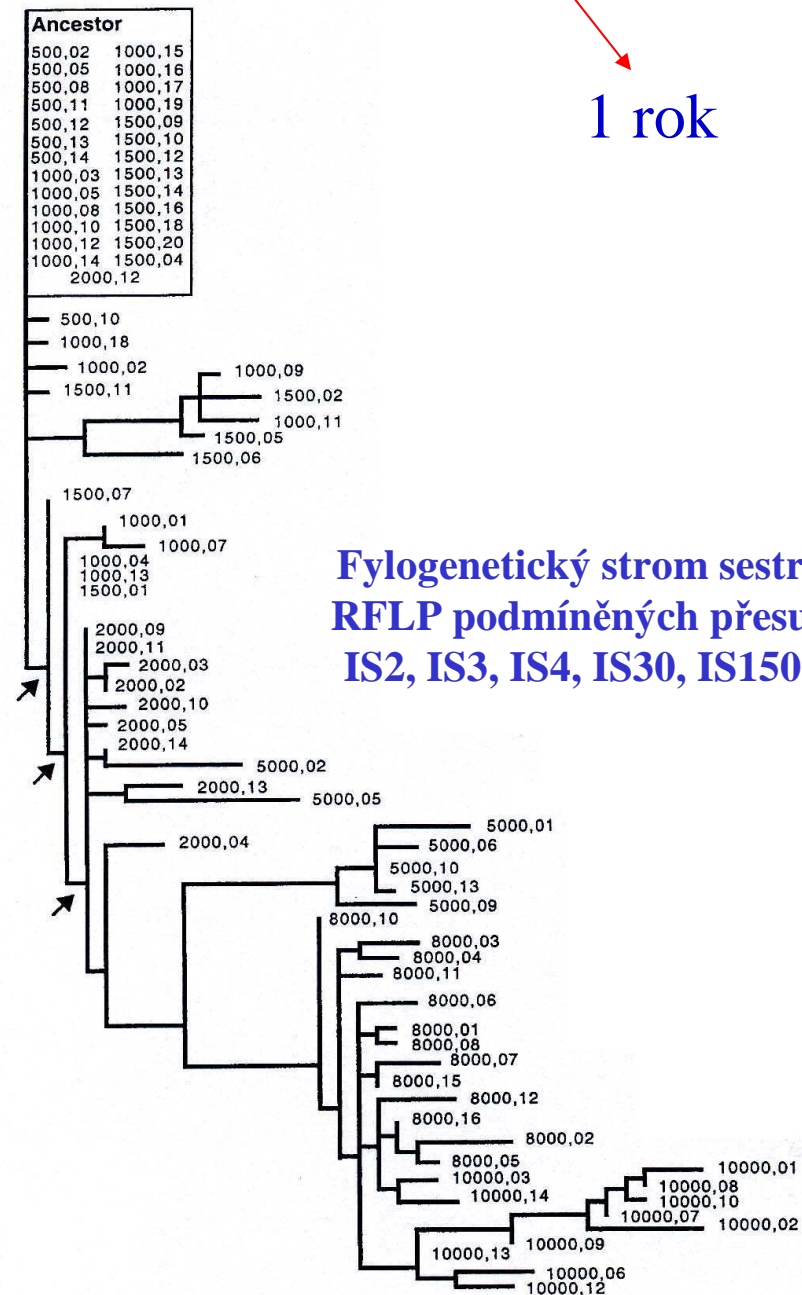
Změny v genomu po
dlouhodobém uchování
kultur *E. coli* a *S. typhimurium*



- změny lokalizace IS
- změny velikosti genomu
- změny fenotypu

Pohyby inzerčních sekvencí v populacích buněk *E. coli*
sledované v průběhu 10 000 generací.

1 rok



Evoluční historie chromozomu *E. coli*

(srovnání *E. coli* K12-MG1655 a kmenů se známou genealogií)

67 událostí: 37 inzercí a 30 delecí

☛ **90% ORF je pro všechny kmeny společné**

☛ **kb až Mb jedinečné DNA:**

*** geny přenesené horizontálně**

➤ **plazmidy**

➤ **bakteriofágy**

➤ **transpozony**

➤ **genové kazety**

**Rozdíly v genomech *E. coli* a *S. typhimurium*
(divergence obou druhů před 120-150 miliony let)**

- rozdíly způsobené rozsáhlými genomovými přestavbami:

- **velké inverze zahrnující až 10% genomu**
- **četné oblasti jedinečné každému druhu**
tzv. „smyčky“ –inzerce nebo delece až 15% délky chromozomu
s náhodnou distribucí

- druhově-specifické geny získané horizontálním přenosem**
- **geny *lac* u *E. coli*, geny pro invazivitu u *S. typhimurium***

Závěry z analýz přestaveb genomu *E. coli* a *S. typhimurium* (u neselektovaných kultur)

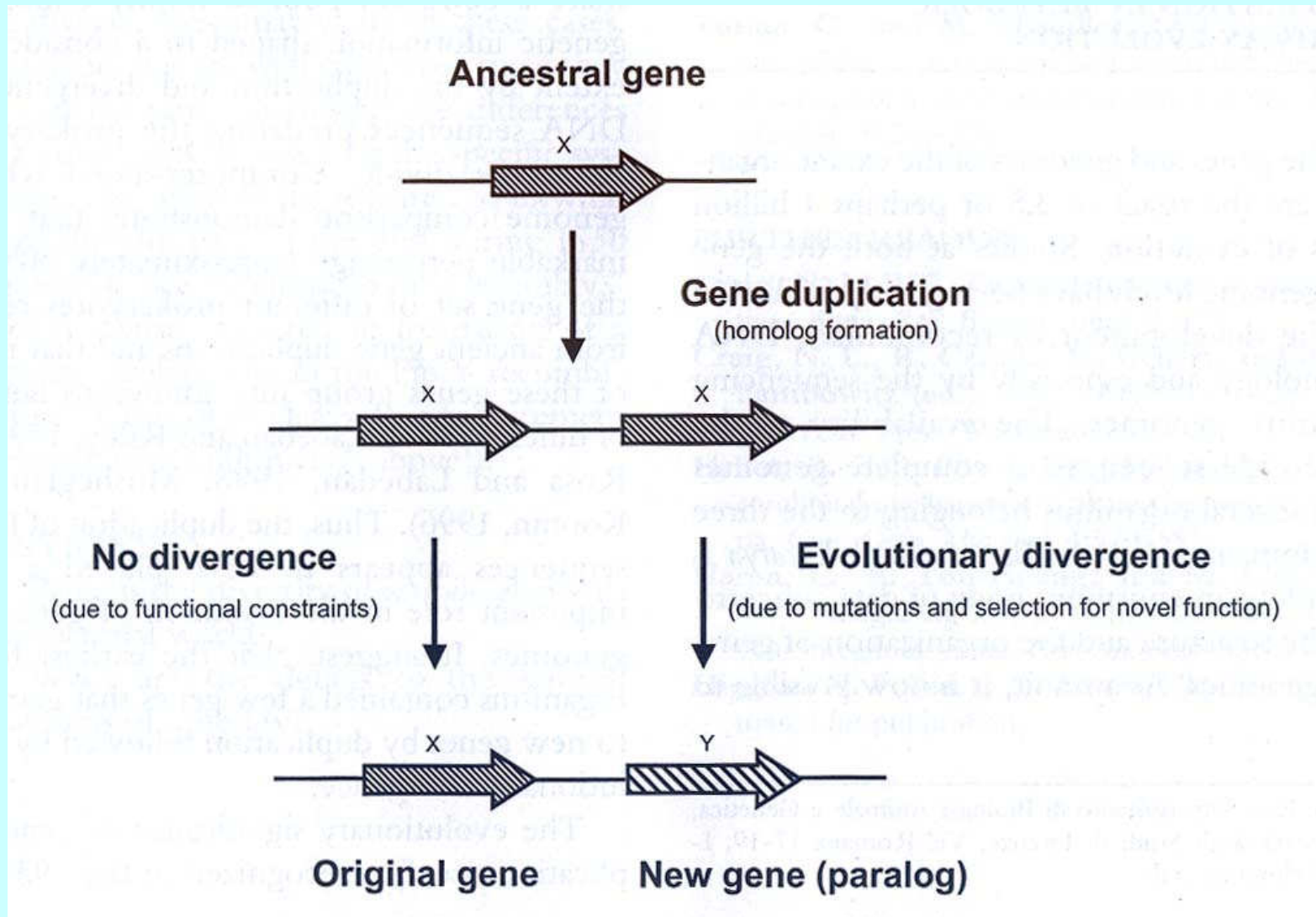
- Průměrný lokus je duplikován v každé z 1000 buněk
 - 10% buněk v kultuře nese duplikaci některé oblasti chromozomu
 - Velikost duplikací: 140 kb – 2100 kb
- Distribuce duplikací není náhodná
- Duplikace jsou ohraničeny dlouhými přímými repeticemi různého typu

Duplikace funkcí → adaptace na změny prostředí

- zvýšení dávky genů
- vytvoření redundantní DNA pro následnou genetickou divergenci

→ paralogní geny – adaptace na nová prostředí

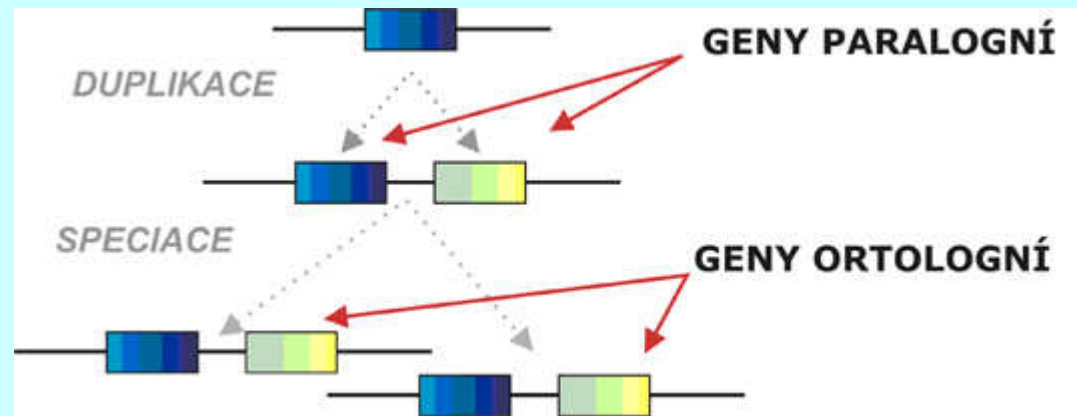
Vytváření paralogních genů duplikací a divergencí



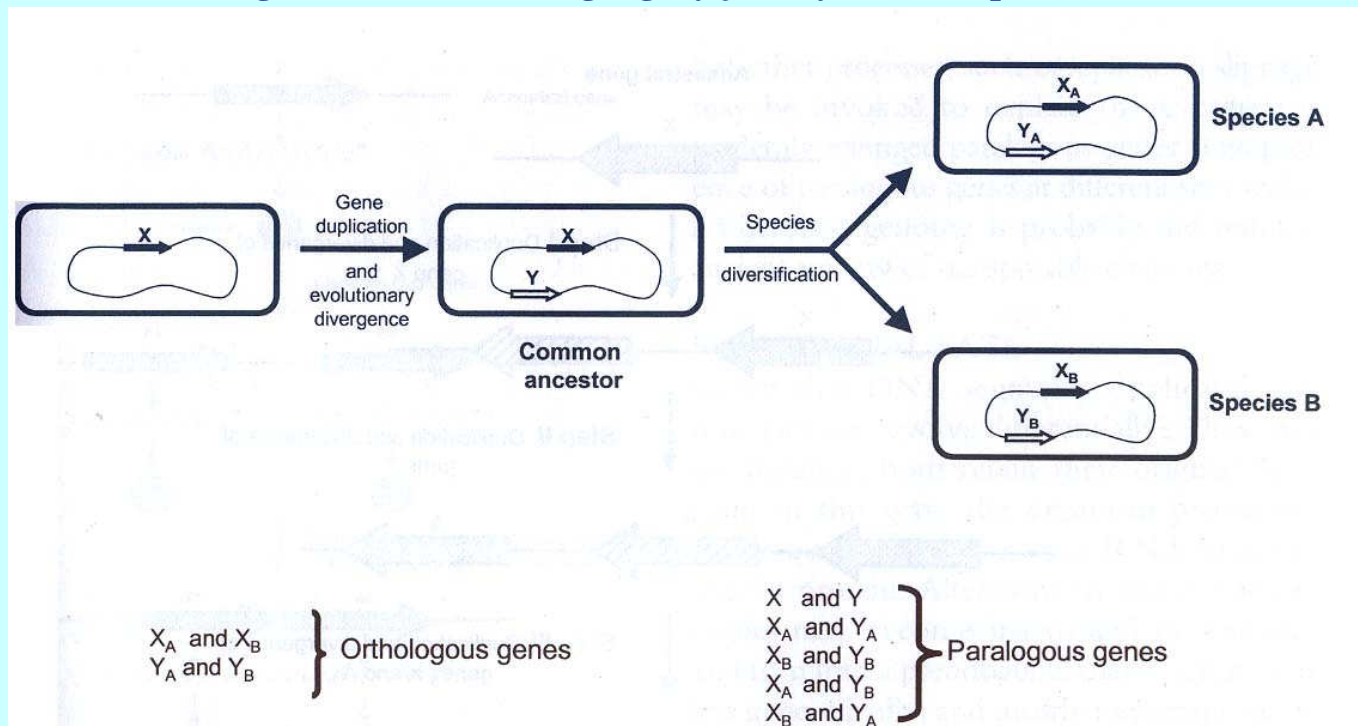
Původní funkce

Nová funkce

Evoluční vztahy mezi ortologními a paralogními geny



Jako homologní jsou označovány geny odvozené z jednoho společného (ancestrálního) genu. Již v roce 1970 bylo navrženo dělení homologických genů na dva typy: geny paralogní a geny ortologní. Paralogní geny jsou výsledkem duplikace ancestrálního genu, zatímco ortologní geny jsou výsledkem speciace.

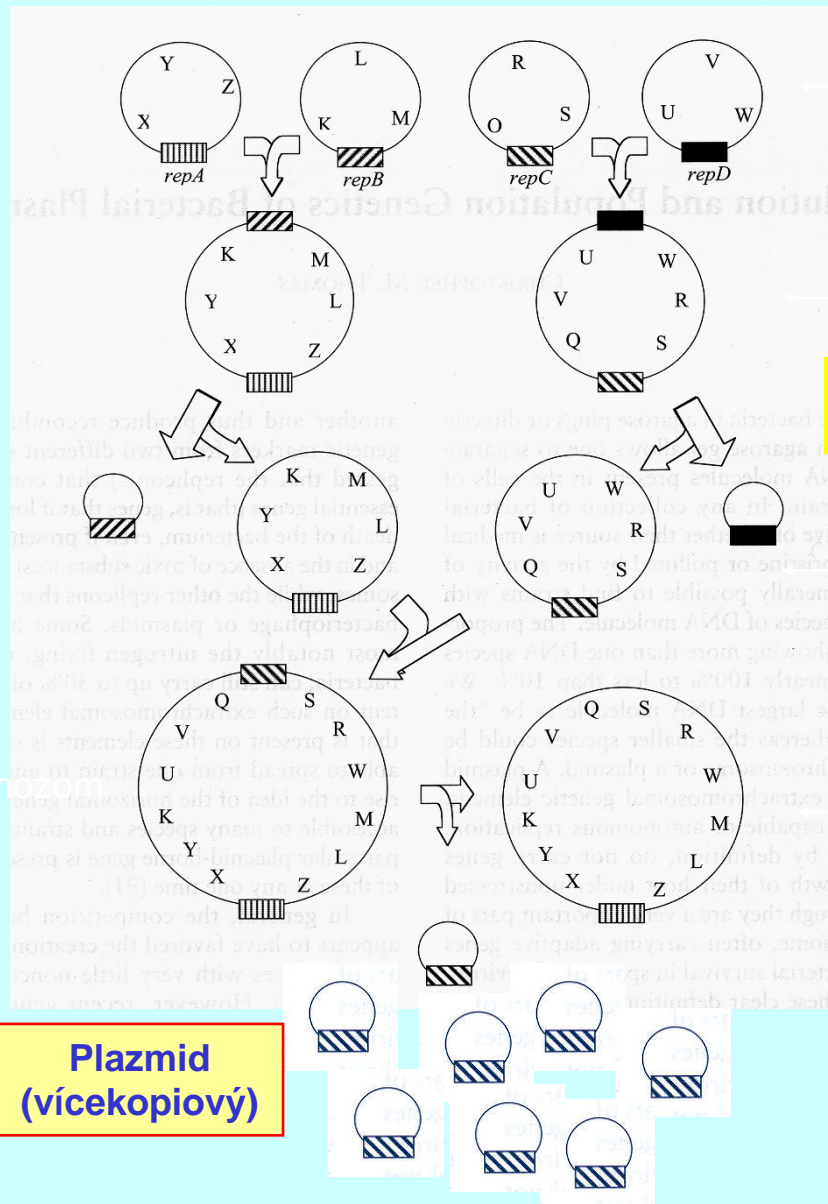


Počty paralogních genů v genomech bakteriálních druhů

Organismus	Velikost genomu (Mbp)	Počet ORF	Počet paralogů
<i>T. pallidum</i>	1.14	1 040	129 (12%)
<i>B. burgdorferi</i>	1.44	1 751	707 (40%)
<i>H. pylori</i>	1.66	1 657	266 (16%)
<i>A. fulgidus</i>	2.18	2 437	719 (30%)
<i>B. subtilis</i>	4.20	4 100	1 947 (47%)
<i>M. tuberculosis</i>	4.41	3 924	2 000 (51%)
<i>E. coli</i>	4.60	4 288	2 272 (53%)

zvýšený adaptivní potenciál

Vznik plazmidů během evoluce bakteriálních replikonů



$F \times F'$

Plazmid
(vícěkopiový)

Původní genom tvořený několika menšími replikony

Vytváření hybridů těchto replikonů vzájemnou integrací

Rozklad hybridů za vzniku větších nízkokopiových stabilních replikonů (chromozomů) nesoucích většinu genů, a malých vysokokopiových replikonů (plazmidů)

Opakování procesu integrace a rozkladu, optimalizace informačního obsahu replikonů

- Výhoda vyššího počtu kopií:
1. vyšší dávka genů,
 2. vyšší šance mutací
 3. přenos mezi buňkami

Horizontální přenos genů

- **Často přenášené: operační geny** (metabolismus a regulace, buněčná struktura)
- **Zřídka přenášené: informační geny** (transkripce, translace)

Horizontální přenos genů je spjat s variabilními genetickými elementy

**profágy,
plazmidy,
IS-elementy,
transpozony,
integrony**

Počet horizontálně přenesených genů u vybraných druhů bakterií a archeí

Druh	Velikost genomu (Mbp)	Počet ORF	Horizontálně přenesené ORF	
			počet	%
Proteobacteria				
<i>Escherichia coli</i>	4,64	4289	381	9,6
<i>Haemophilus influenzae</i>	1,83	96	96	6,2
<i>Helicobacter pylori</i>	1,67	1553	89	6,4
<i>Rickettsia prowazekii</i>	1,11	834	28	3,6
Gram-pozitivní bakterie				
<i>Bacillus subtilis</i>	4,21	4100	537	14,5
<i>Mycoplasma genitalium</i>	0,58	480	67	14,5
<i>Mycoplasma pneumoniae</i>	0,82	677	39	5,9
<i>Mycobacterium tuberculosis</i>	4,41	3918	187	5,0
Spirochaete				
<i>Borrelia burgdorferi</i>	0,91	850	12	1,56
<i>Treponema pallidum</i>	1,14	1031	77	8,3
Chlamydiae				
<i>Chlamydia trachomatis</i>	1,04	894	36	4,3
<i>Deinococcus radiodurans</i>	2,65	2580	95	3,92
<i>Synechocystis</i> sp.	3,57	3169	219	7,5
<i>Thermotoga maritima</i>	1,86	1846	198	11,63
Archaea				
<i>Aeropyrum pernix</i>	1,67	2694	370	14,0
<i>Methanobacterium therm.</i>	1,75	1869	179	10,3
<i>Methanococcus jannaschii</i>	1,66	1715	77	5,0
<i>Pyrococcus abyssi</i>	1,76	1765	124	7,35



1% bakteriálních
genů bylo získáno
HGT z eukaryot

Horizontálně přenesené geny (HGT) u *E. coli* K12 MG1655 (po divergenci *E. coli* a *S. typhimurium*)

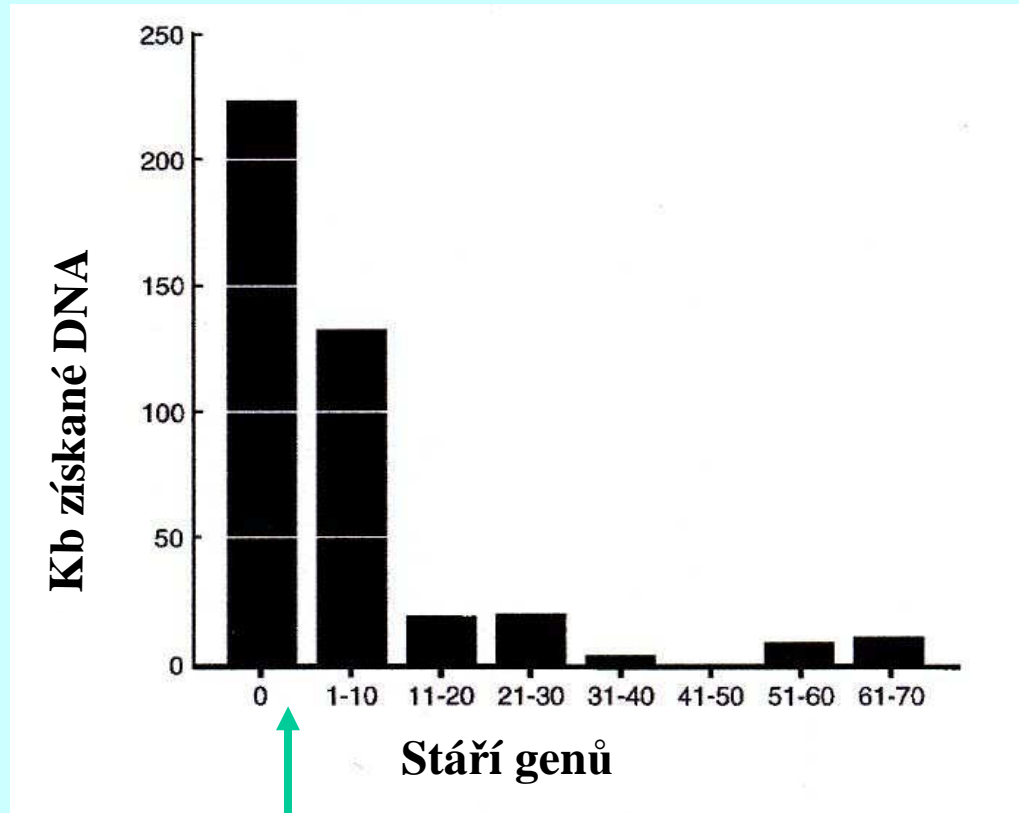
**Genom *E. coli* obsahuje relikty 755 HGT
(18% genomu = 548 kb, 234 přenosových událostí)**

- Vyšší proporce HTG v oblasti terminátoru replikace**
- Lokalizace HTG poblíž genů pro tRNA (přenos pomocí fágů)**
- V blízkosti HGT se nachází 68% všech inzerčních sekvencí**
 - IS jsou přenášeny spolu HTG**
 - IS navozují integraci přenášené DNA**

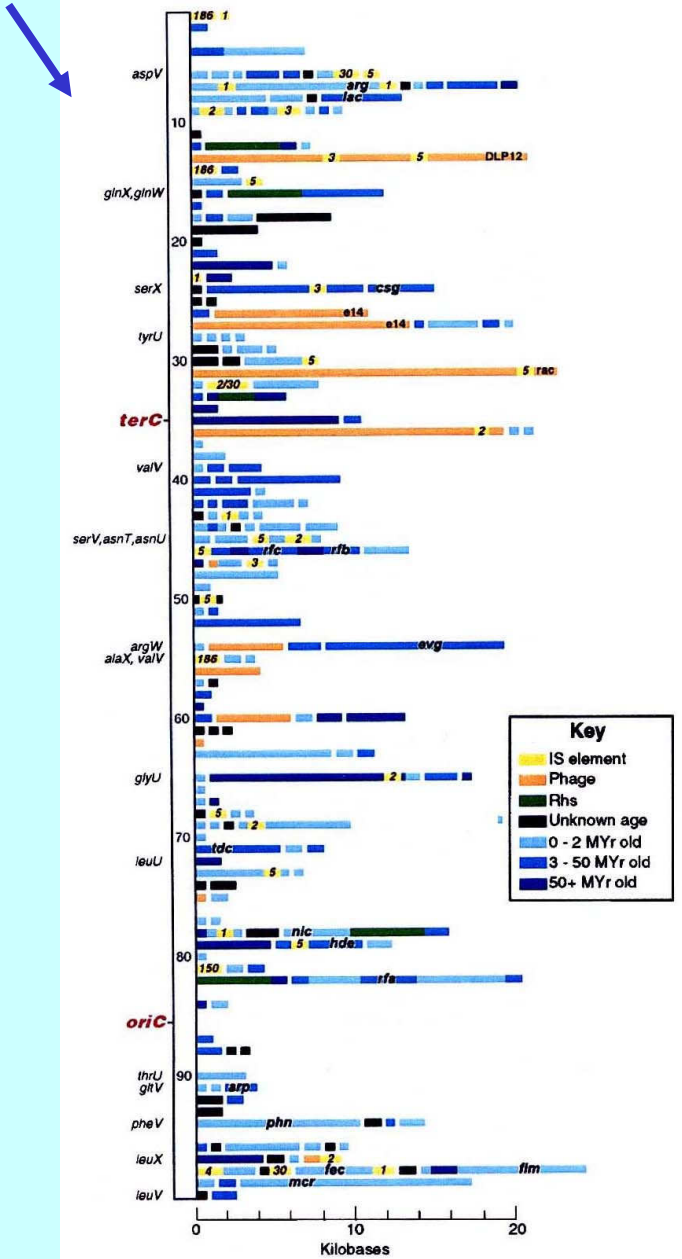
Příklady genetických elementů přenášených horizontálně

Genetický element	Označení	Faktory virulence nebo jiné funkce
Ostrovy patogenity		
Enteropatogenní <i>E. coli</i>	PAI	Adhesiny, hemolyziny, cytotoxiny
Enterohemorhagické <i>E. coli</i>	LEE (esp-LEE)	Adhesiny, enterotoxiny
<i>Vibrio cholera</i> O1, 0139	VPI (vibrio path. island)	Pilusy, regulace
<i>Staphylococcus aureus</i>	TSST-1-PAI (SaPII aj) Exotoxinový PAI Enterotoxinový PAI	Toxin toxického šoku Exotoxin Enterotoxin
Ostrůvky patogenity		
<i>E. coli</i> , <i>Shigella dysenteriae</i>	Lokus chuA a shuA	Příjem hemu
<i>Salmonella enterica</i> sv. <i>Typhimurium</i>	Lokus msgA/pagC	Protein vnější membrány, přežívání v makrofágách
<i>Streptococcus pyogenes</i>	Oblast vir	proteázy
Plazmidy		
<i>E. coli</i> (mimo střevo)	pHly, Vir plazmidy	Hemolyzin, cytotoxický nektrotizující faktor
intestinální <i>E. coli</i>	pO157, Vir plazmidy	Adheziny, enterotoxiny, kataláza, hemolyzin
<i>Shigella flexneri</i>	pWR100, pWR501	Invasiny, enterotoxin
<i>Clostridium tetani</i>	pCL1	Tetanový toxin
Bakteriofágy		
<i>Clostridium botulinum</i>	cI	Botulotoxin A, B
<i>Corynebacterium diphtheriae</i>	β	Difterický toxin A, B
<i>E. coli</i> (enterohemorhagické)	H19, 933	Shiga toxin A, B
<i>S. aureus</i>	φ42	Enterotoxin A, B
<i>V. cholerae</i>	CTXφ	Cholerový toxin A, B

Odhadované stáří genů horizontálně přenesených do chromozomu *E. coli* MG1655 a jejich lokalizace v genomu



IS sekvence,
profágy, Rhs



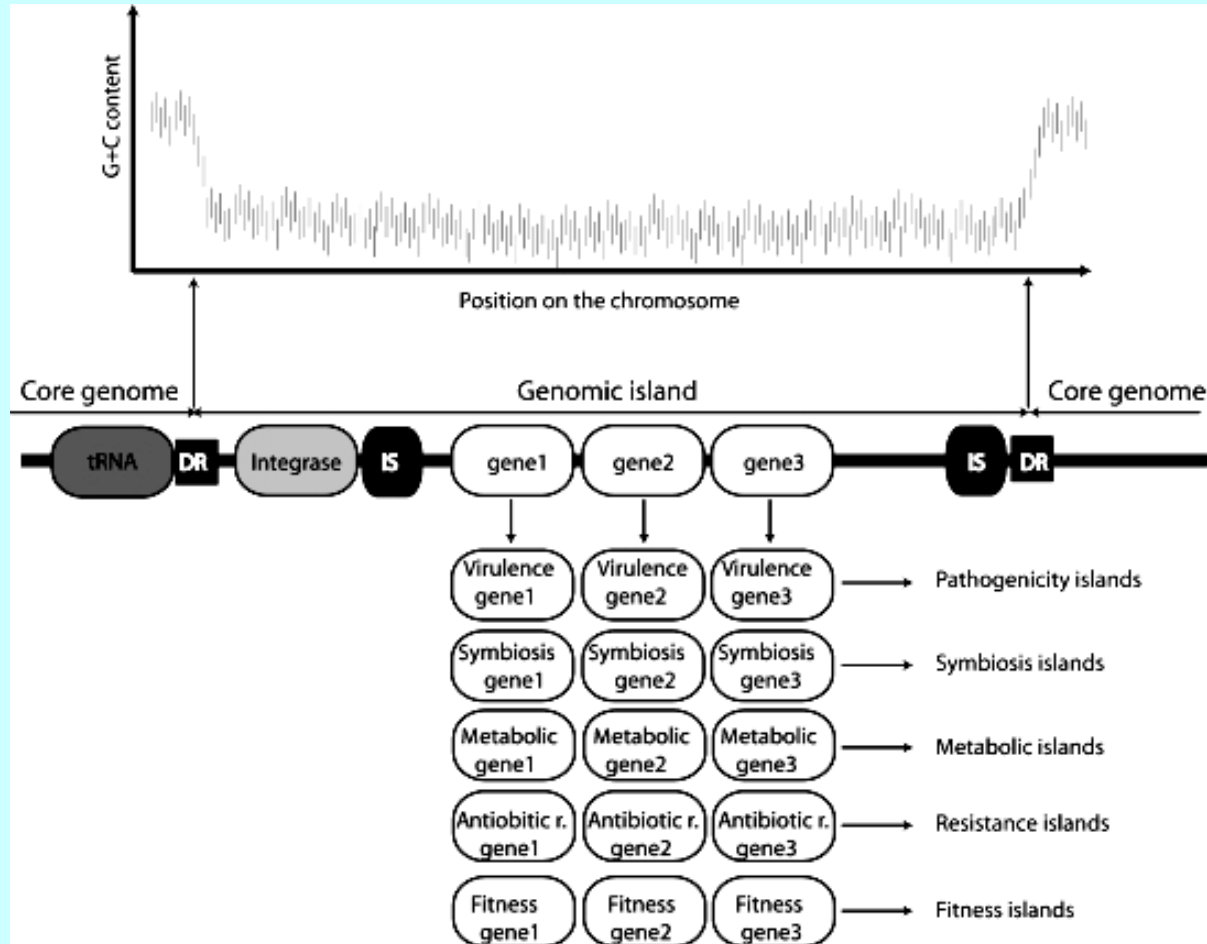
Genomické ostrovy („fitness“ ostrovy)

části genomů se znaky mobilních genetických elementů s odlišným obsahem GC, ohraničené repeticemi a geny pro mobilitu

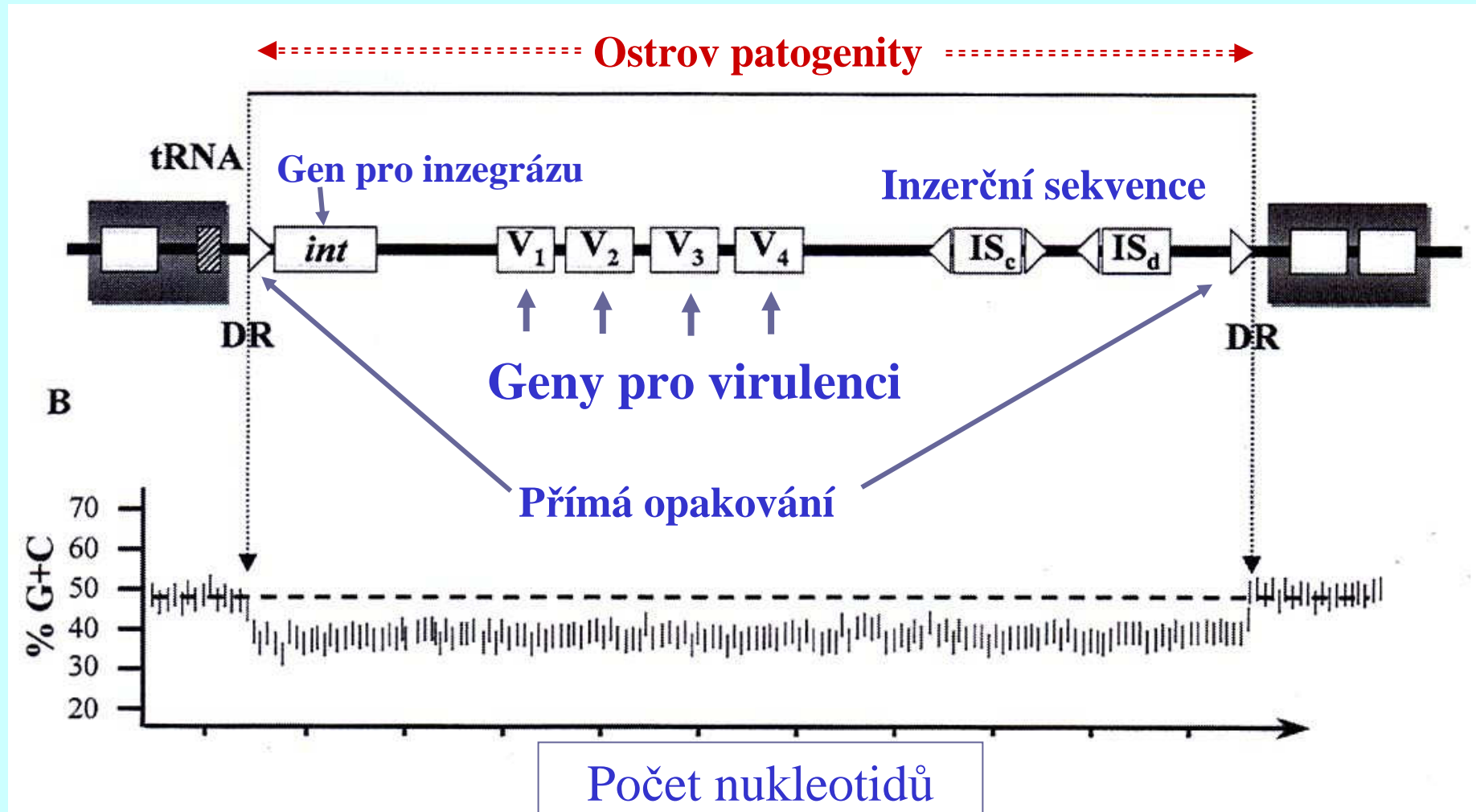
- ostrovy patogenity**
- ekologické ostrovy**
- saprofytické ostrovy**
- symbiosové ostrovy**

Charakteristické pro jednotlivé kmeny v rámci druhu

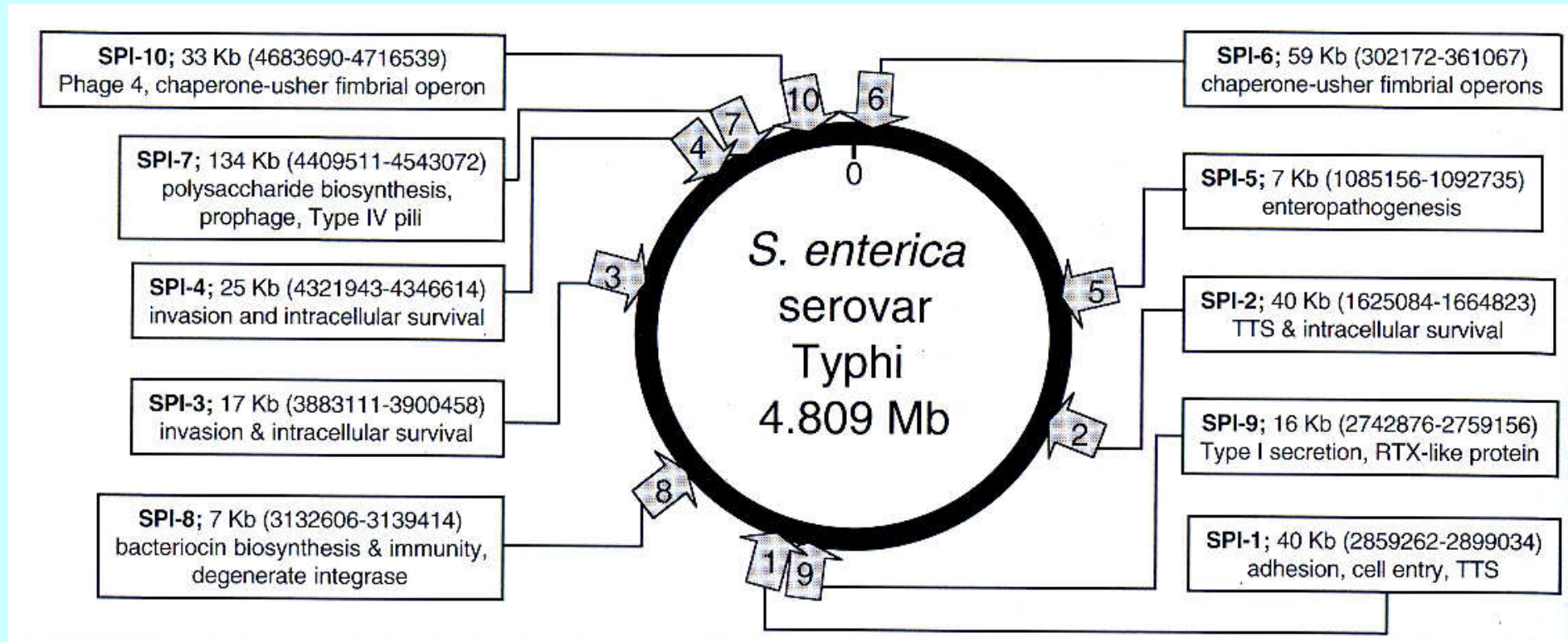
Obecné rysy genomických ostrovů (GEIs)



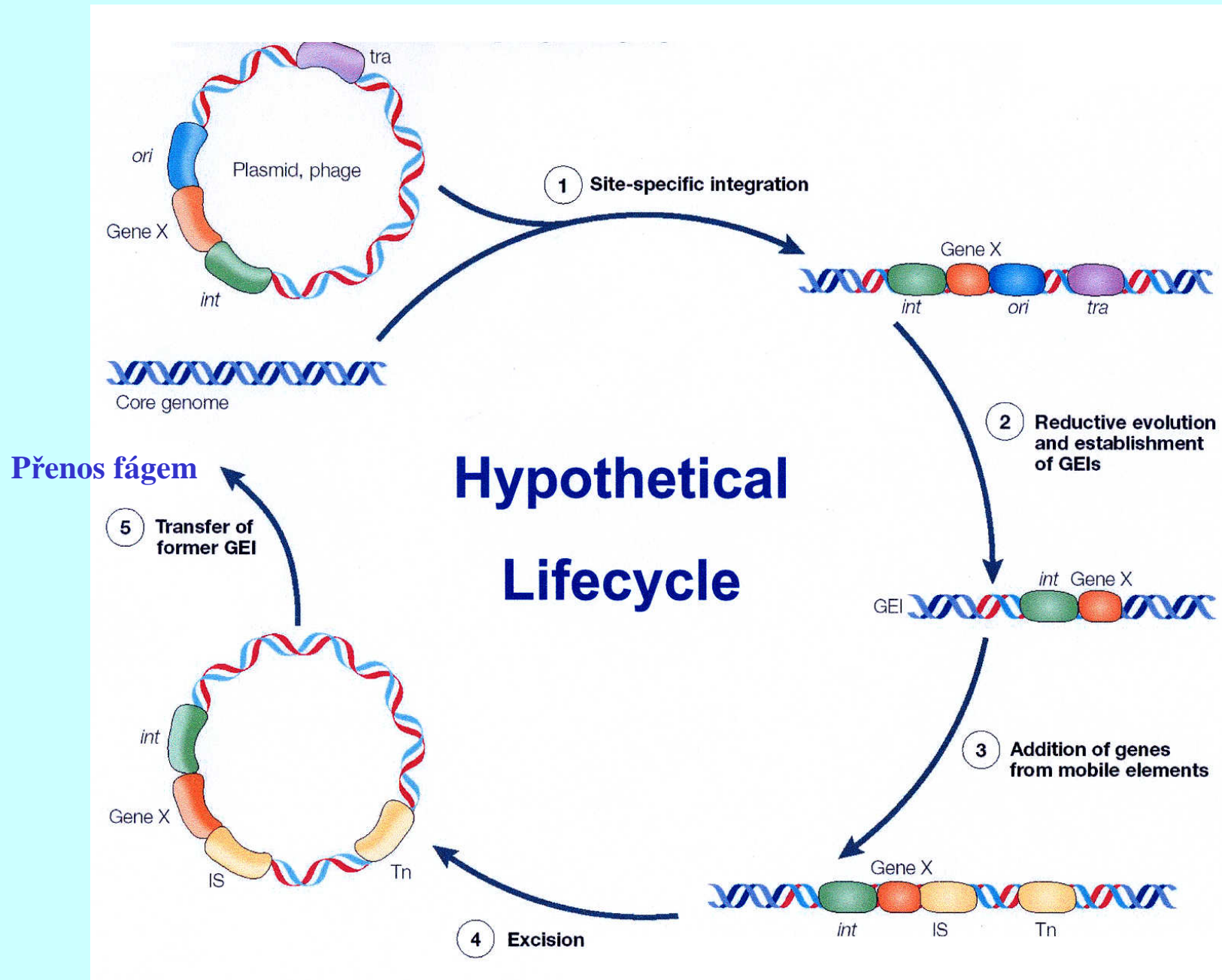
Obecná struktura ostrovů patogenity



Distribuce ostrovů patogenity u *S. enteritica* serovar Typhi

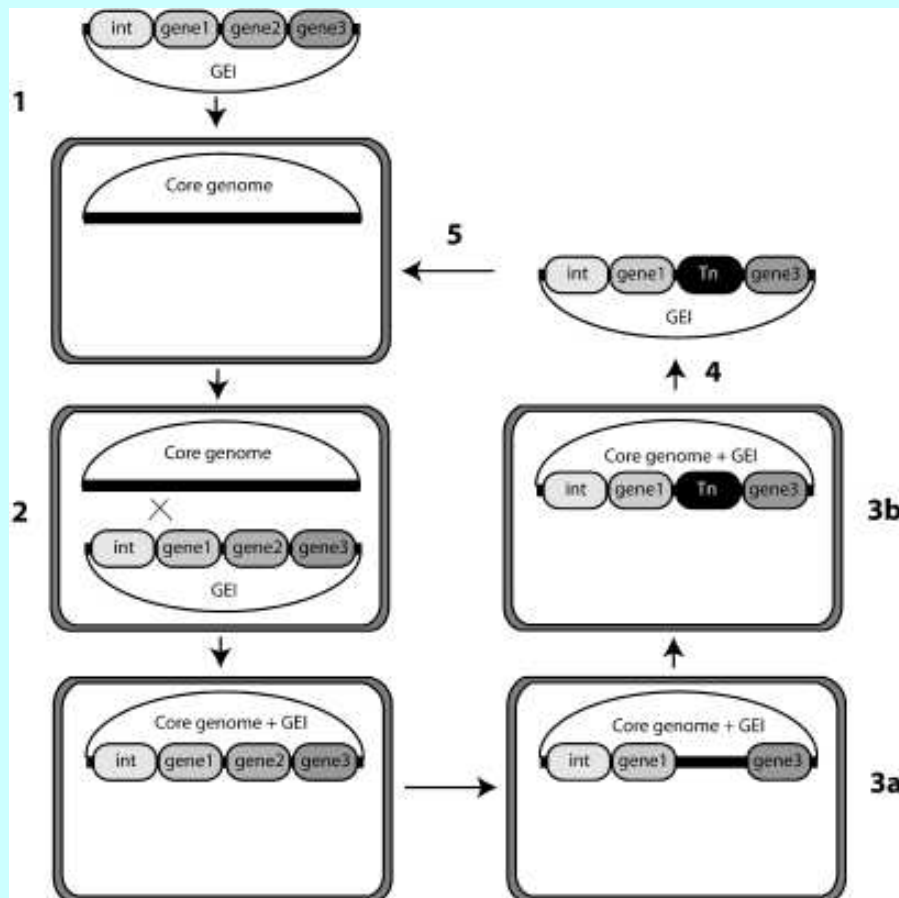


Vznik genomických ostrovů u patogenních a environmentálních mikrobu

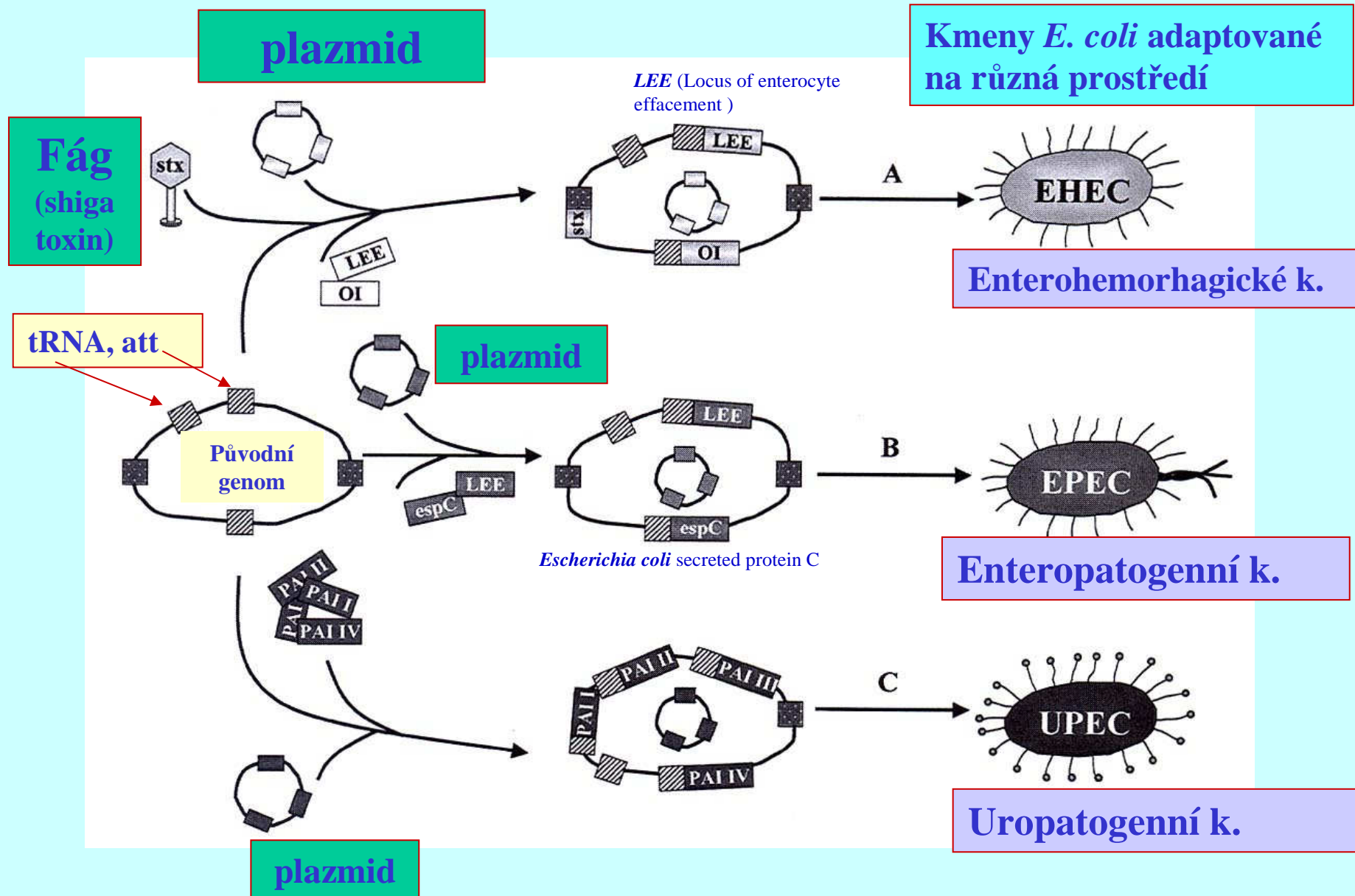


Schematické znázornění životního stylu mobilních genomických ostrovů sestávající z následujících kroků:

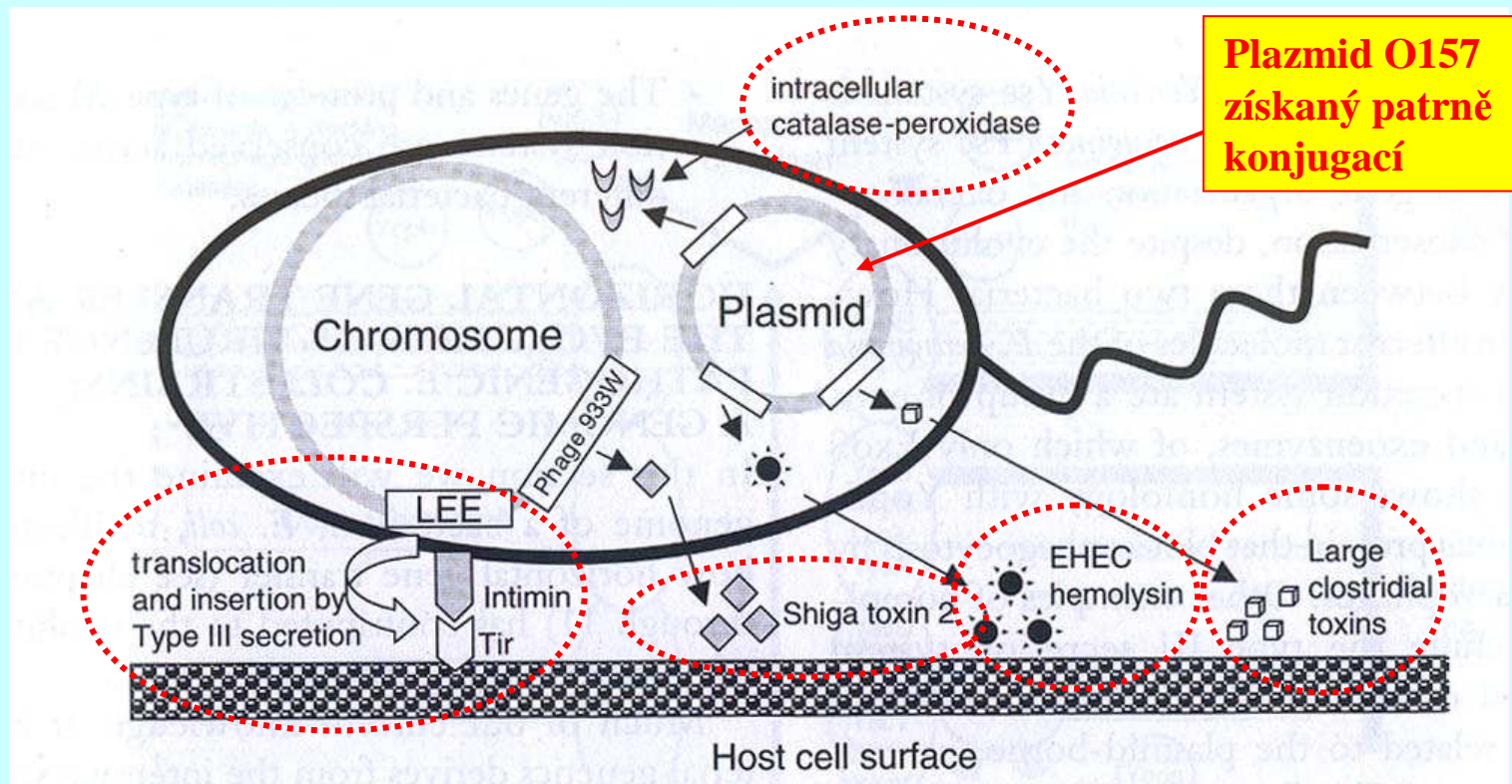
1. Zisk ostrova horizontálním přenosem
2. Začlenění do chromozomu hostitele místně specifickou rekombinací
3. Vývoj ostrova přestavbami n. ztrátami genů (3a) nebo jejich získkem (3b).
4. Vyčlenění ostrova z chromozomu
5. Přenos do dalšího recipienta



Model vzniku ostrovů patogenity u patogenních *E. coli*

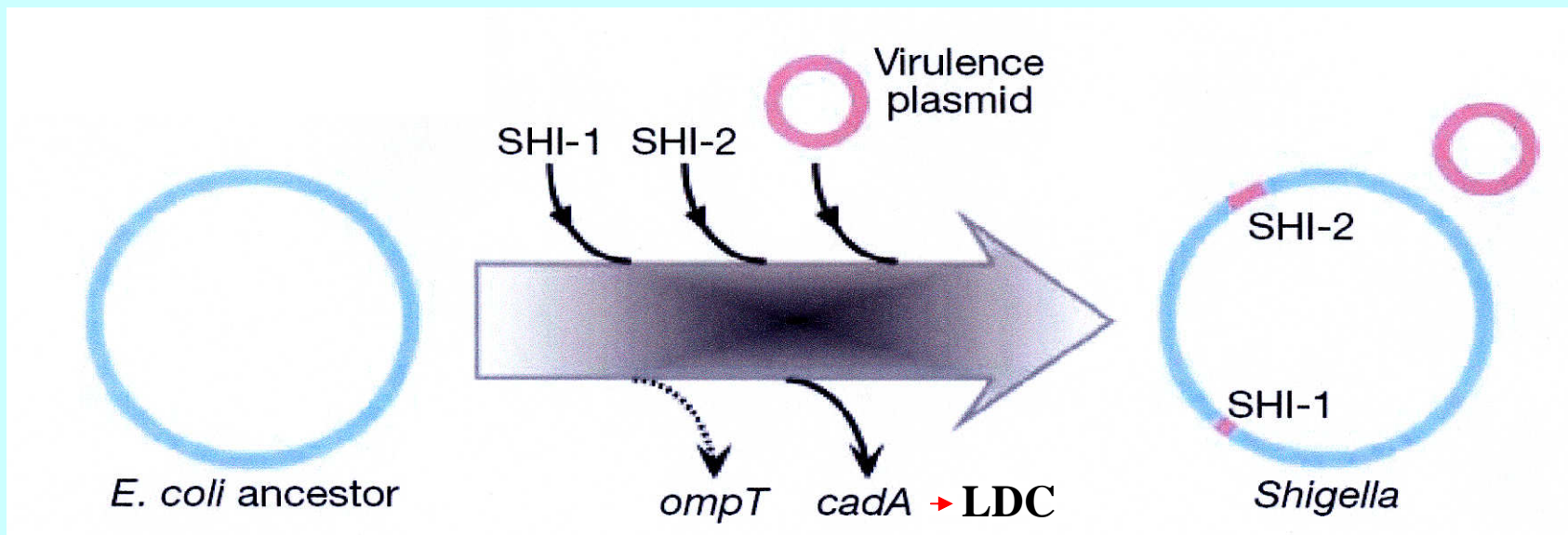


Horizontálně získané virulenční faktory zodpovědné za patogenitu enterohemorhagického kmene E. coli O157:H7



LEE = locus for enterocyte effacement (uchycení na střevní sliznici a její léze)

Sukcese genetických událostí vedoucích k virulenci druhů *Shigella*



Kmeny *Shigella* jsou odvozeny z *E. coli* po získání virulenčního plazmidu a dvou chromozomových genů (SHI-1, SHI-2) a po ztrátě několika málo genů z genomu *E. coli* (*lyzindekarboxyláza – inhibice toxinů*)

r. Shigella x Escherichia coli K-12

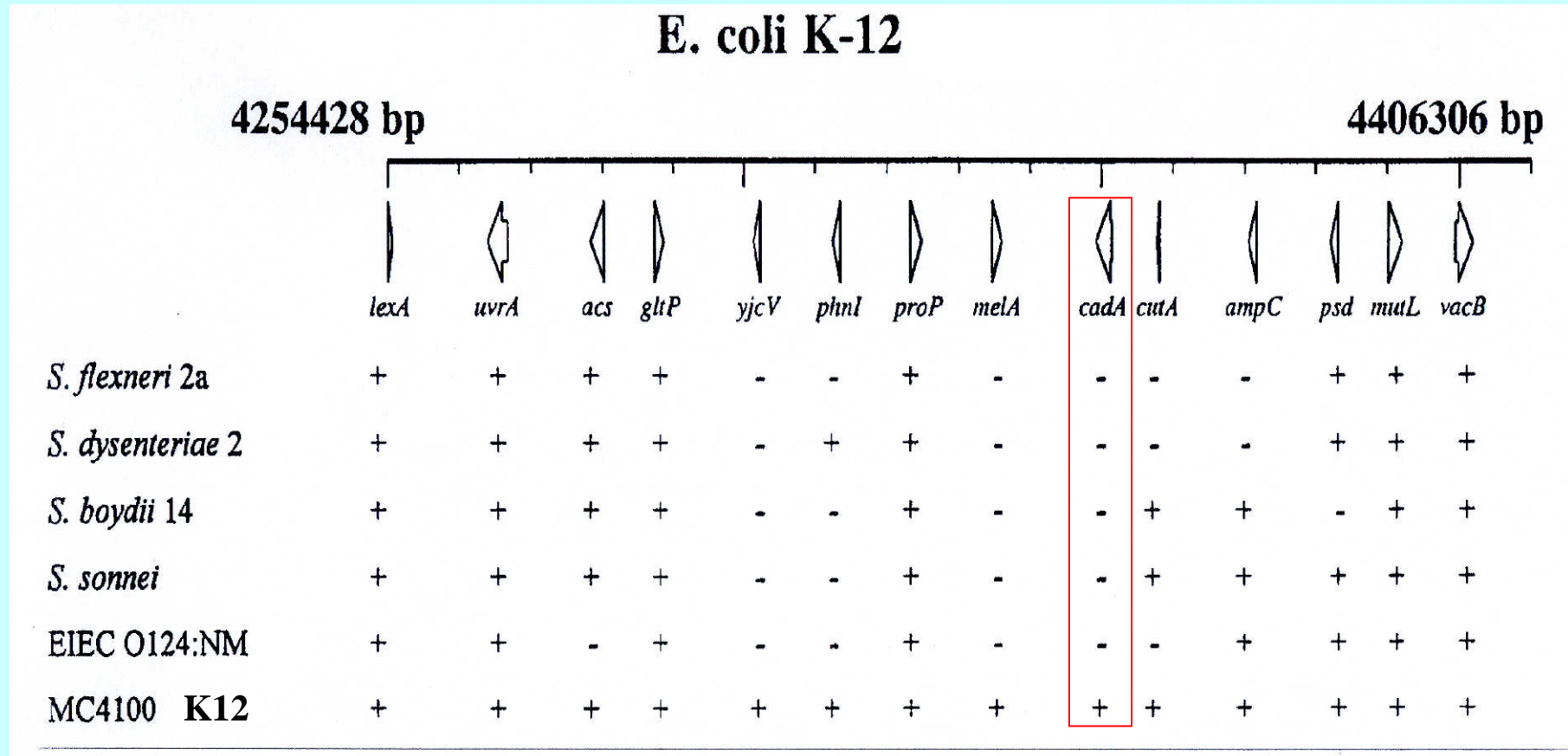
90% homologie DNA (!)

Kolinearita genů

Rekombinace po HGT

Vliv ztráty genů ztráty genů na patogenitu enterobakterí

Genomové delece („černé díry“) zvyšující virulenci u *Shigella* spp. a u enteroinvazivních kmenů *E. coli* (EIEC)



Výsledek hybridizace sond z 14 různých genů *E. coli* K12 z oblasti genomu 4254428-4406306 bp k genomové DNA reprezentativních kmenů *Shigella* a EIEC (+ = pozitivní hybridizace, - = negativní hybridizace)

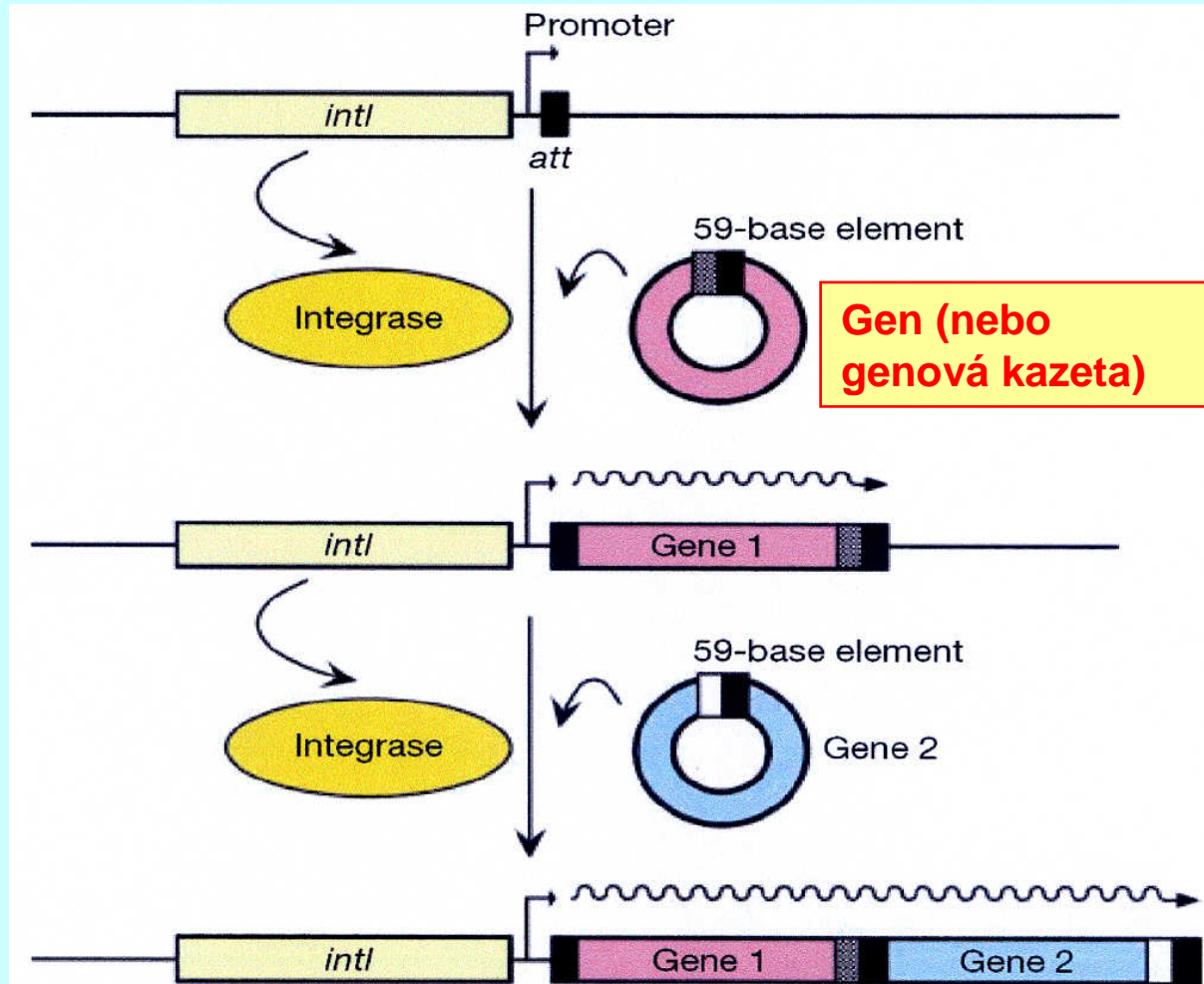
***Shigella* spp., původce bacilární dysentérie, se liší od příbuzných komensálních kmenů *Escherichia coli* přítomností plazmidu, který kóduje virulenční funkce. Patogenní bakterie však vedle toho mohou postrádat vlastnosti, které jsou charakteristické pro nepatogenní druhy.**

Enzym lysindekarboxyláza (LDC) je přítomen u $\approx 90\%$ kmenů *E. coli* kmenů, avšak vždy chybí u kmenů *Shigella*. Pokud je gen *cadA* kódující LDC zaveden do *Shigella flexneri* 2a, její virulence se sníží a je silně inhibována aktivita enterotoxinu.

Inhibitorem enterotoxinu je kadaverin, což je produkt reakce katalyzované LCD. Srovnání genomů *S. flexneri* 2a a laboratorního kmene *E. coli* K-12 ukázalo, že v oblasti, kde se nachází *cadA* je u shigely rozsáhlá delece. Vybrané kmeny *Shigella* spp. a enteroinvazivních kmenů *E. coli* mají podobné delece genu *cadA*.

Tyto výsledky naznačují, že patogenní kmeny *Shigella* spp. se vyvinuly z *E. coli* nejen po získání virulenčních genů nesených na plazmidu, ale současnou ztrátou genů v důsledku delece. Vytvoření těchto “černých děr”, tj. delece genů, které jsou škodlivé pro patogenní způsob života, představují evoluční proces, který umožňuje patogenu zvýšit jeho virulenci. To, že kadaverin snižuje aktivitu enterotoxinu, může představovat obecný návod a model pro “antitoxinovou terapii” – nový způsob léčby infekčních onemocnění.

ZACHYTÁVÁNÍ GENŮ INTEGRONY



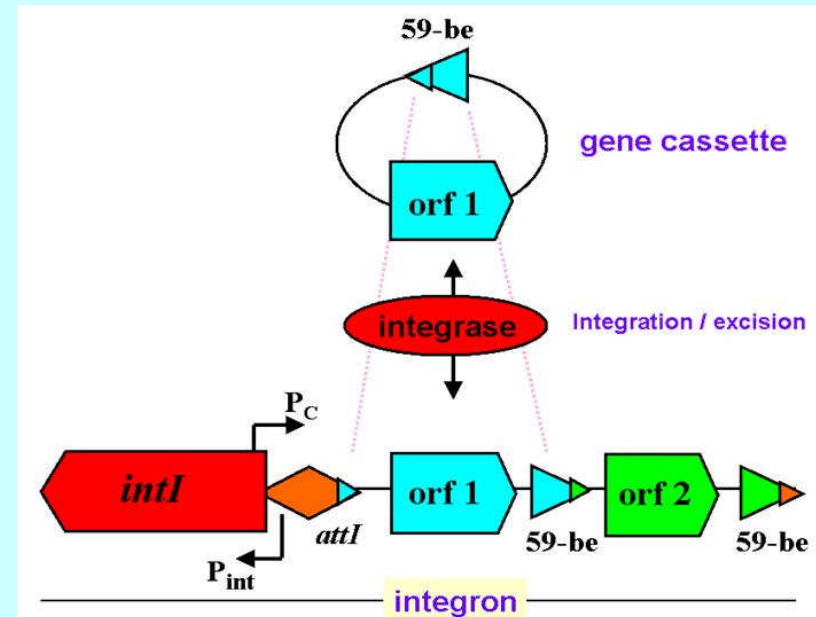
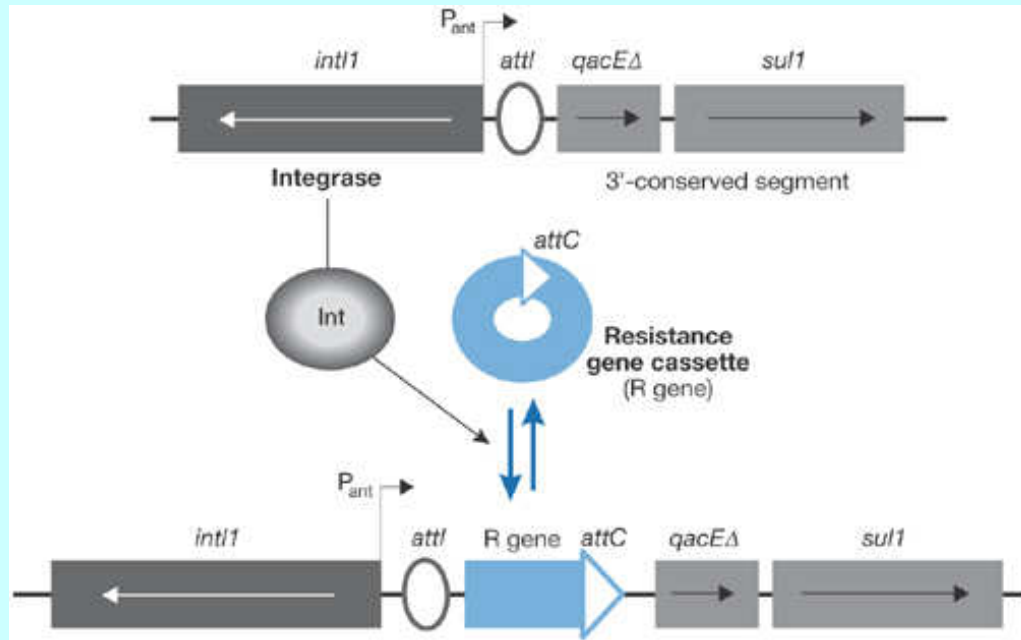
Genová kazeta v CTn (případně v plazmidu)

Integron obsahuje:

1. att místo,
umožňující opakované zachycení genů nebo genových kaset

2. Gen *intl* kódující integrázu,
rozpoznávající různá 59 bp rekombinační místa

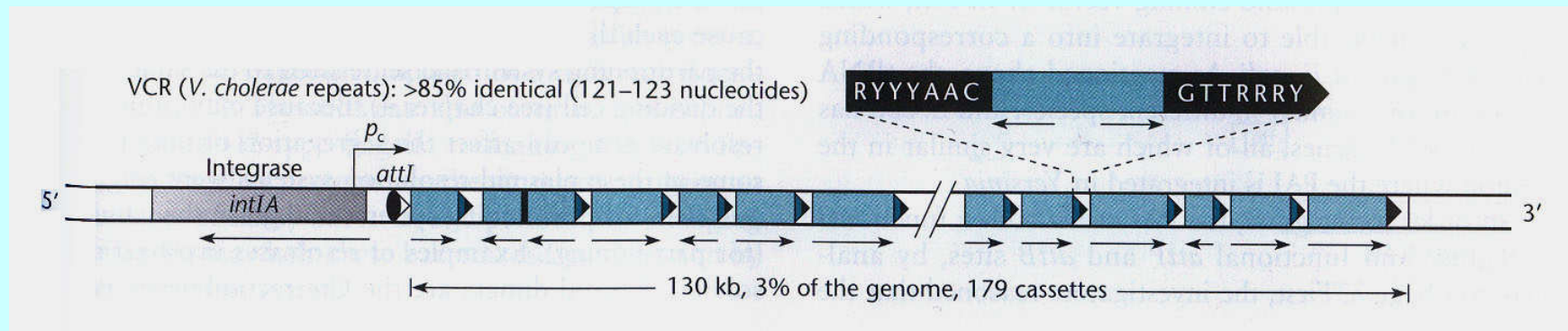
3. Promotor
umožňující expresi vloženého genu



Integron obsahuje místně-specifický rekombinační systém schopný začleňovat a exprimovat geny přítomné ve strukturách nazývaných mobilní genové kazety. Integrony byly původně identifikovány na mobilních elementech patogenních bakterií jako hlavní rezervoáry genů antibiotické rezistence. Patří mezi starobylé vzájemně fylogeneticky odlišné struktury, zjištěny u 10% sekvenovaných bakteriálních genomů. Diverzita kazet je extrémně vysoká – mají tedy významnou úlohu v adaptaci, nejen vzhledem k rezistenci k antibiotikům.

Super-integron *Vibrio cholerae*

Obsahuje více než 100 kazet kódujících rezistenci k různým antibiotikům a další funkce

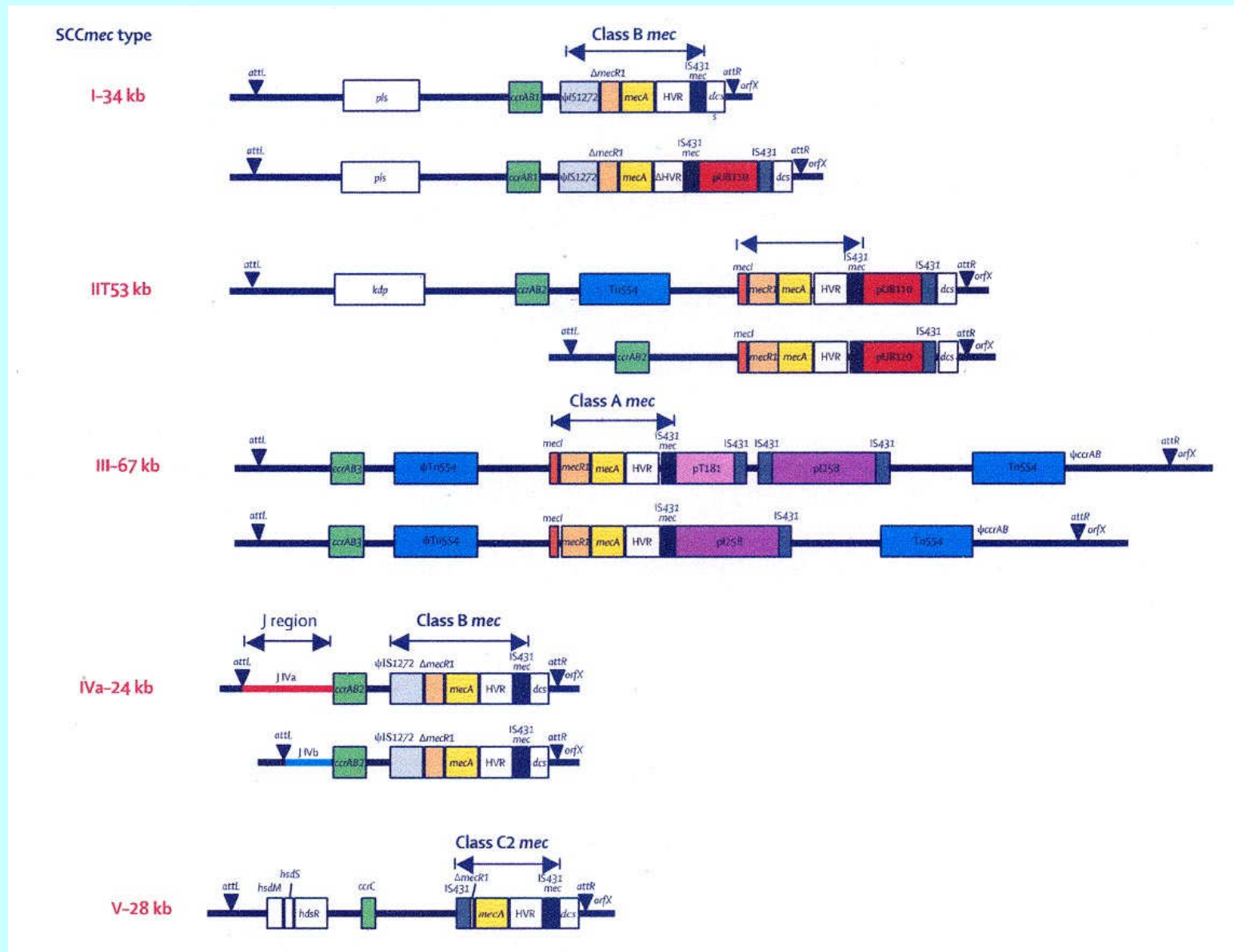


- ▶ konzervativní sekvence s vysokou homologií
- oblasti mezi kazetami odpovídající potenciálním attC místům

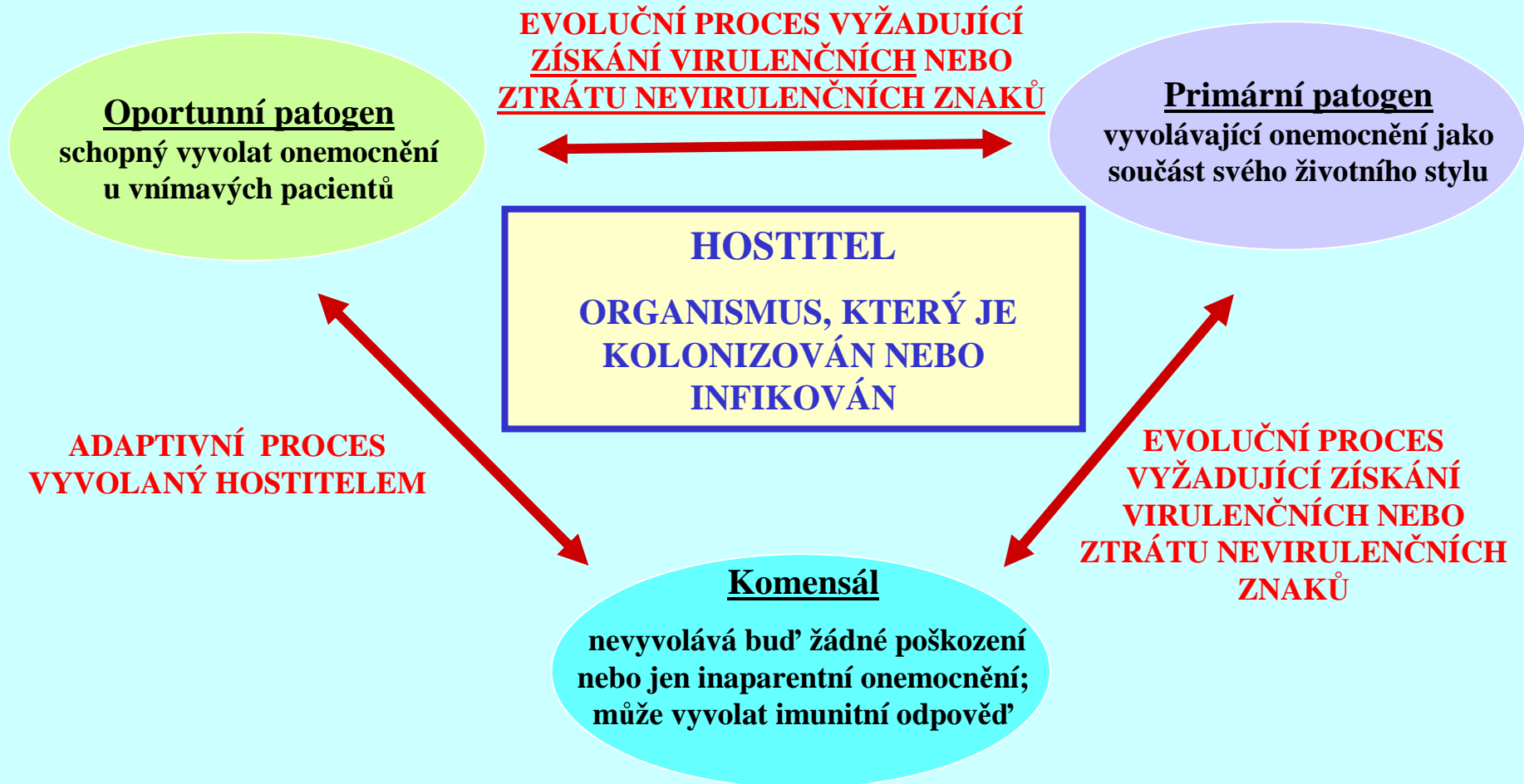
Integrony jsou genetické struktury schopné zachytit nebo vyčlenit genové kazety, které obvykle kódují determinanty antimikrobiální rezistence. I když integrony nejsou schopné se samy mobilizovat (přenášet) jsou obvykle asociovány s transpozony a často jsou lokalizovány na plazmidech, což usnadňuje jejich přenos. Jsou tak ideálně vhodné pro šíření a re-kombinace genů antimikrobiální rezistence.

Integrony bývají často součástí větších transpozonů, často přítomných na plazmidech nebo na integrativních konjugativních elementech.

Typy stafylokokových chromozomových kazet (SCCmec) zodpovědných za rezistenci kmenů *S. aureus* k meticilinu



Interakce patogen-hostitel u bakteriálních infekčních onemocnění



Minimální genom

Definice: Základní sada esenciálních genů, kterou daný organismus potřebuje k udržení života.

Představuje silně redukovanou sestavu genů jeho genomu, která se liší v závislosti na životním stylu a podmínkách prostředí (růstové požadavky, dostupné zdroje živin atp).

Závisí na podmínkách experimentu, při nichž jsou esenciální geny určovány.

Cíle studia minimálního genomu

1. Poznání životně nezbytných enzymových funkcí – pochopení prebiotické existence – předchůdců prvních bakterií
2. Pochopení vývoje bakteriálních druhů
3. Předpoklad pro přípravu bakterií s umělým genomem (*Mycoplasma laboratorium*)
 - produkce látek pro průmyslové využití (paliva, plasty, farmaka)

Přístupy ke stanovení minimálního genomu

A. Teoretické přístupy

- 1. Bioinformatický přístup – srovnání genomů různých organismů a vyhledání těch genů (genových funkcí), které jsou konzervovány u většiny druhů – tyto geny jsou esenciální, udržují se v podobných formách u všech**
- 2. Modelování buněčných procesů, zejména metabolických drah. Vyhledání biochemických modulů a jeho konfrontace s genetickým základem.**

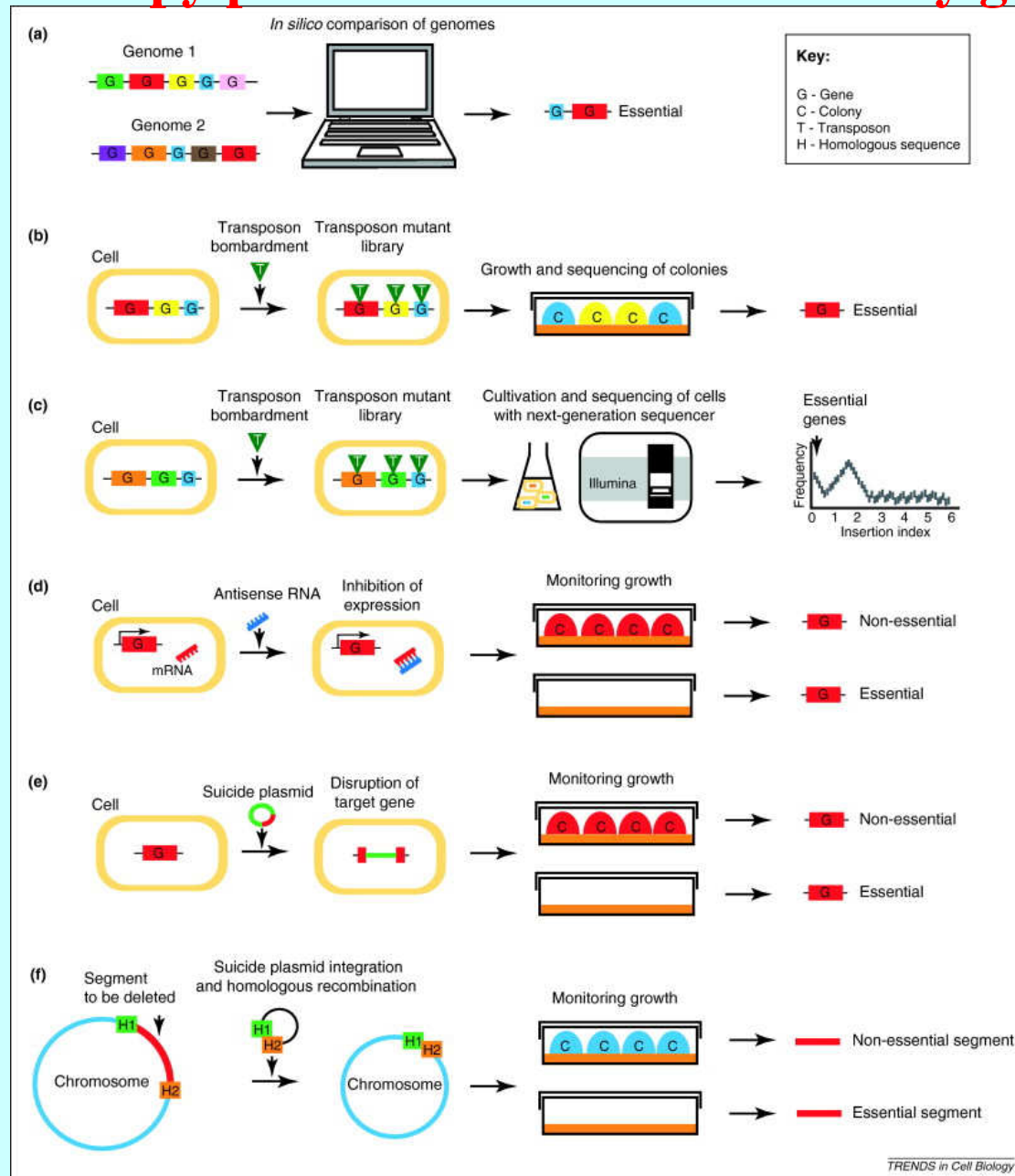
Přístupy ke stanovení minimálního genomu

B. Experimentální přístupy – navození delece genů

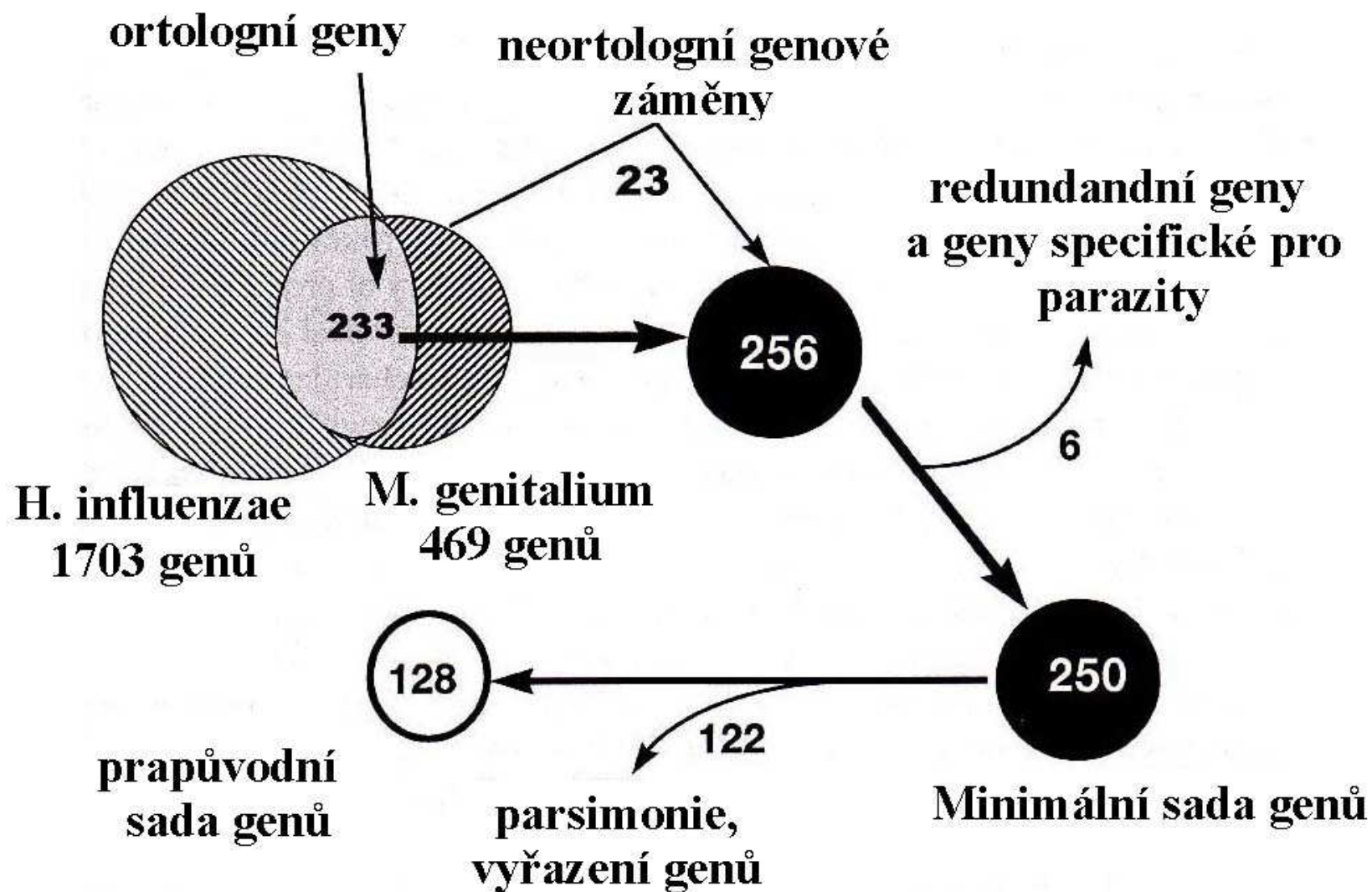
Gen, který je esenciální,

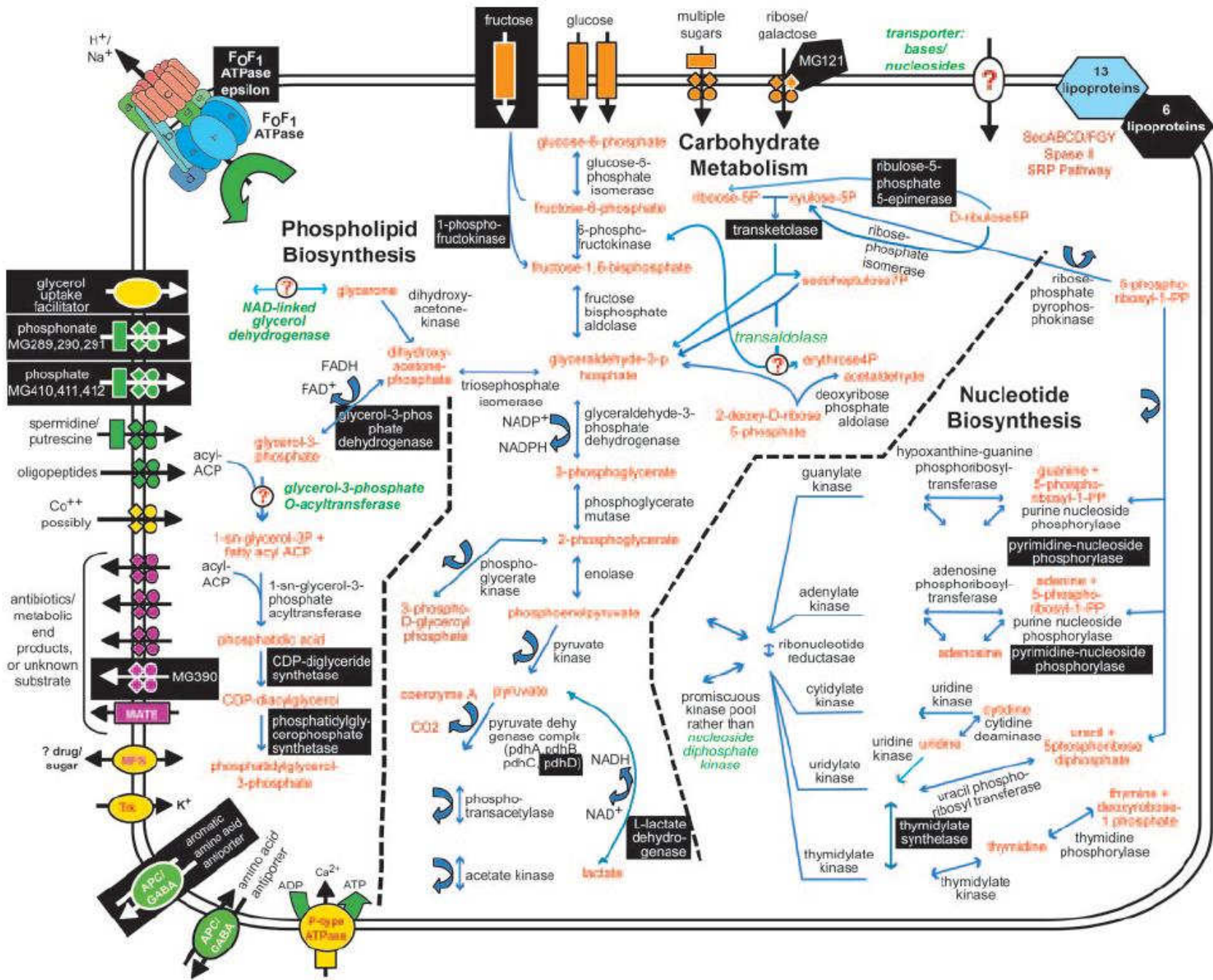
1. **Využití sebevražedných plazmidů.** Plazmid obsahuje sekvence homologní se sekvencemi ohraničujícími úsek genomu určený k deletování. Po začlenění plazmidu do genomu dojde k intramolekulární rekombinaci, která vede buď k odstranění plazmidu nebo žádaného úseku genomu (chromozomu). Lze využít též lineární DNA – princip je stejný jako u plazmidu).
2. **Transpozonová mutageneze.** Náhodné začlenění transpozonu vede k inzerční inaktivaci genů a ztrátě jejich funkcí
3. **Antisense RNA.** Inaktivace transkriptů strukturních genů.

Přístupy pro stanovení minimální sady genů

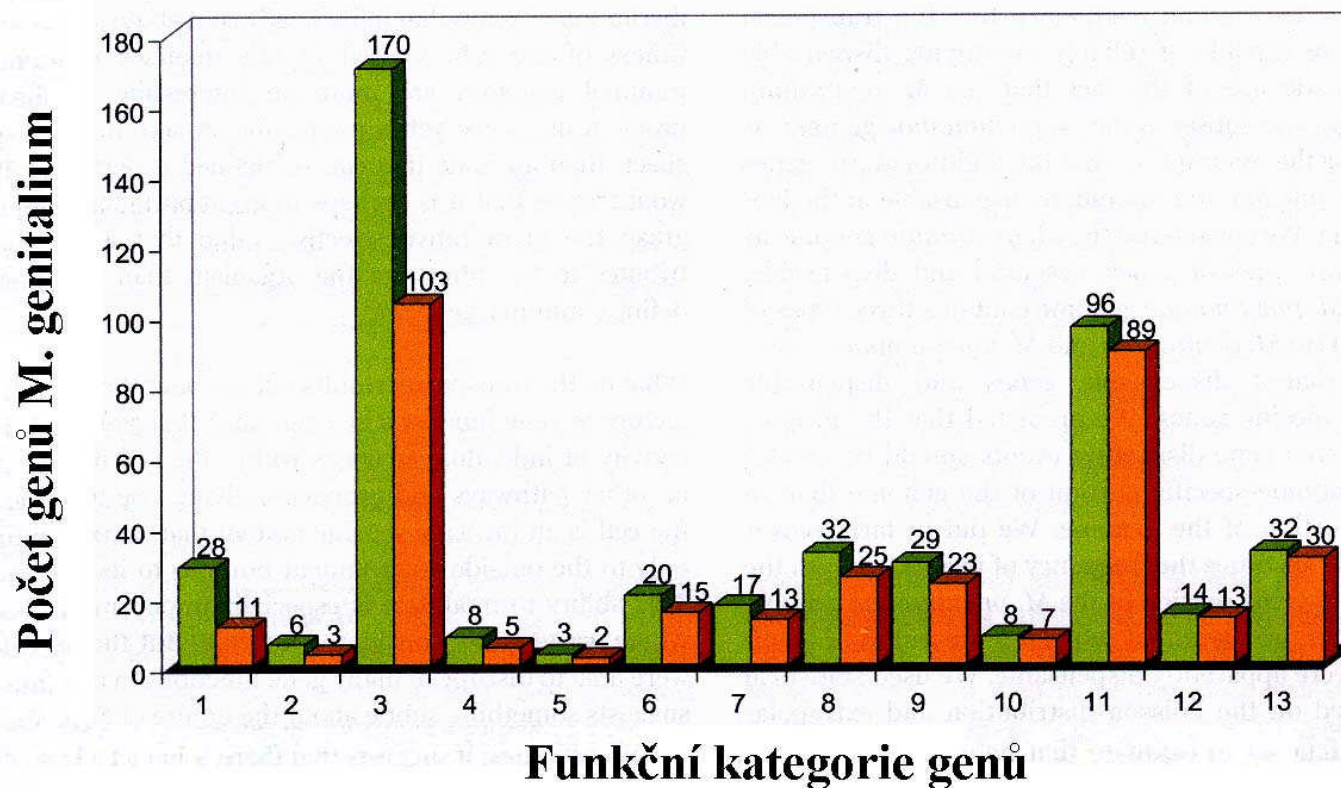


Odhad minimální sady genů pro život buňky ze srovnání genomů *Haemophilus influenzae* a *Mycoplasma genitalium*





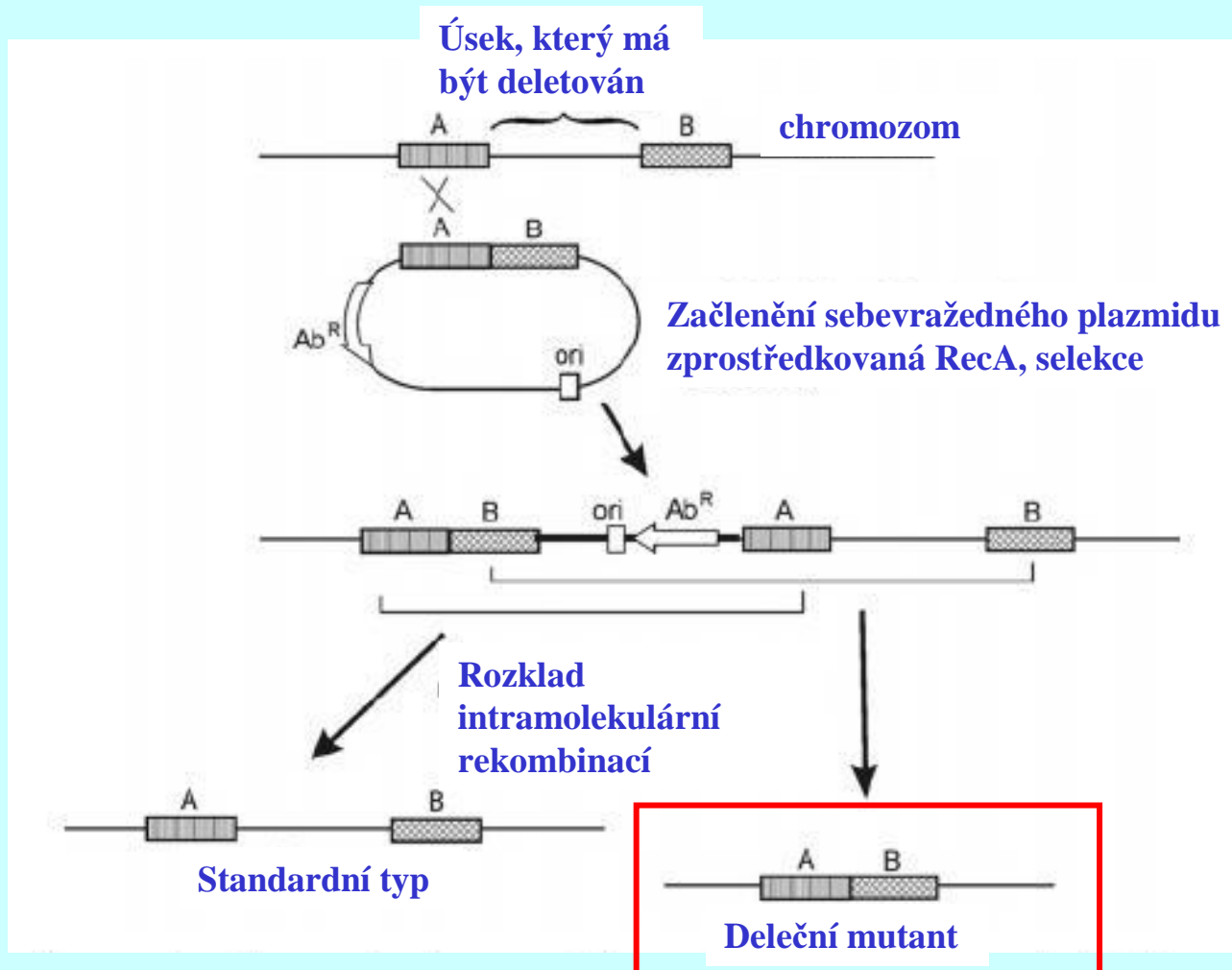
Počty genů *Mycoplasma genitalium* podle jejich funkce a inaktivace transpozonovou mutagenezí



- | | |
|-------------------------|-------------------------------------|
| 1. Buněčný obal | 8. Transport |
| 2. Regulace | 9. Replikace, rekombinace, reparace |
| 3. NEZNÁMÁ FUNKCE | 10. Metabolismus lipidů |
| 4. Metabolismus | 11. Translace |
| 5. Biosyntéza kofaktorů | 12. Buněčné procesy |
| 6. Metabolismus Pu a Py | 13. Energie |
| 7. Transkripce | |

■ Počet genů v jednotlivých funkčních kategoriích u *M. genitalium*
■ Počet genů, které nebyly přerušeny při transpozonové mutagenezi

Navození delece úseku chromozomu pomocí plazmidu



Organismus	Počet esenciálních genů
<i>Escherichia coli</i>	1617
<i>Haemophilus influenzae</i>	642
<i>Streptococcus pneumoniae</i>	244
<i>Mycoplasma genitalium</i>	381
<i>Vibrio cholerae</i>	779
<i>Staphylococcus aureus</i>	653
<i>Saccharomyces cerevisiae</i>	1110

Počet esenciálních genů je odlišný pro různé organismy – každý organismus má odlišný počet esenciálních genů v závislosti na tom, který konkrétní kmen a za jakých podmínek je testován. U bakterií (nebo např. u kvasinek) byly všechny nebo většina genů deletována jednotlivě, aby se zjistilo, které z nich jsou pro přežití nezbytné. Takové testy se obvykle provádějí na bohatých mediích obsahujících všechny živiny – když jsou všechny živiny přítomny, nebudou geny vyžadované pro jejich syntézu „esenciální“, když jsou ale kmeny pěstovány na minimální půdě, řada genů vyžadovaných pro syntézu určitých živin (nebo např. vitamínů) bude esenciálních. V tabulce jsou počty zjištěné na bohatých mediích.

Srovnání informačního obsahu sekvencovaných genomů

- Počet informačních genů je v každém genomu zhruba stejný, i když se jejich velikosti značně liší.
- Počet genů ostatních funkčních kategorií je mnohem variabilnější a má tendenci se zvyšovat.
- Se zvětšováním velikosti genomu přibývá paralogních genů a zvětšuje se též biochemická komplexita organismu.
- Jedna čtvrtina ORF u každého druhu je jedinečná a nemá významnou sekvenční homologii k žádné dostupné proteinové sekvenci.

Zhruba třetina (~100) esenciálních genů nemá žádnou ze známých funkcí

ZÁVĚRY VYVOZENÉ Z ANALÝZY MINIMÁLNÍCH GENOMŮ

- Každý genom obsahuje dva typy genů
 - Esenciální geny zajišťující základní biologické procesy
 - Geny pro dosažení selektivní výhody v daném prostředí (metabolismus – nové substráty, nové faktory virulence)
- Minimální sada genů je společná pro všechny druhy (současný odhad ~ 206 kódujících genů)
- Prostředí určuje, který gen je pro daný druh esenciální a který je postradatelný

Science 2 July 2010: vol. 329 no. 5987 pp. 52-56

Creation of a Bacterial Cell Controlled by a Chemically Synthesized Genome

The assembly of a synthetic *M. mycoides* genome in yeast. A synthetic *M. mycoides* genome was assembled from 1078 overlapping DNA cassettes in three steps. In the first step, 1080-bp cassettes (orange arrows), produced from overlapping synthetic oligonucleotides, were recombined in sets of 10 to produce 109 ~10-kb assemblies (blue arrows). These were then recombined in sets of 10 to produce 11 ~100-kb assemblies (green arrows). In the final stage of assembly, these 11 fragments were recombined into the complete genome (red circle). With the exception of two constructs that were enzymatically pieced together in vitro (white arrows), assemblies were carried out by in vivo homologous recombination in yeast. Major variations from the natural genome are shown as yellow circles. These include four watermarked regions (WM1 to WM4), a 4-kb region that was intentionally deleted (94D), and elements for growth in yeast and genome transplantation. In addition, there are 20 locations with nucleotide polymorphisms (asterisks). Coordinates of the genome are relative to the first nucleotide of the natural *M. mycoides* sequence. The designed sequence is 1,077,947 bp. The locations of the Asc I and BssH II restriction sites are shown. Cassettes 1 and 800-810 were unnecessary and removed from the assembly strategy. Cassette 2 overlaps cassette 1104, and cassette 799 overlaps cassette 811.

- VENTERINSTITVTE
- CRAIGVENTER
- HAMSMITH
- CINDIANDCLYDE
- GLASSANDCLYDE

Prokaryota s nejmenším genomem a počtem genů

Bakteriální druh	počet genů	velikost Mbp
Candidatus Hodgkinia cicadicola Dsem	169	0.14
Candidatus Carsonella ruddii PV	182	0.16
Candidatus Sulcia muelleri GWSS	227	0.25
Candidatus Sulcia muelleri SMDSEM	242	0.28
Buchnera aphidicola str. Cinara cedri	357	0.4261
Mycoplasma genitalium G37	475	0.58
Candidatus Phytoplasma mali	479	0.6
Buchnera aphidicola str. Baizongia pistaciae	504	0.6224
Nanoarchaeum equitans Kin4-M	540	0.49