

Evolve bakteriálních genomů

Charakteristické rysy:

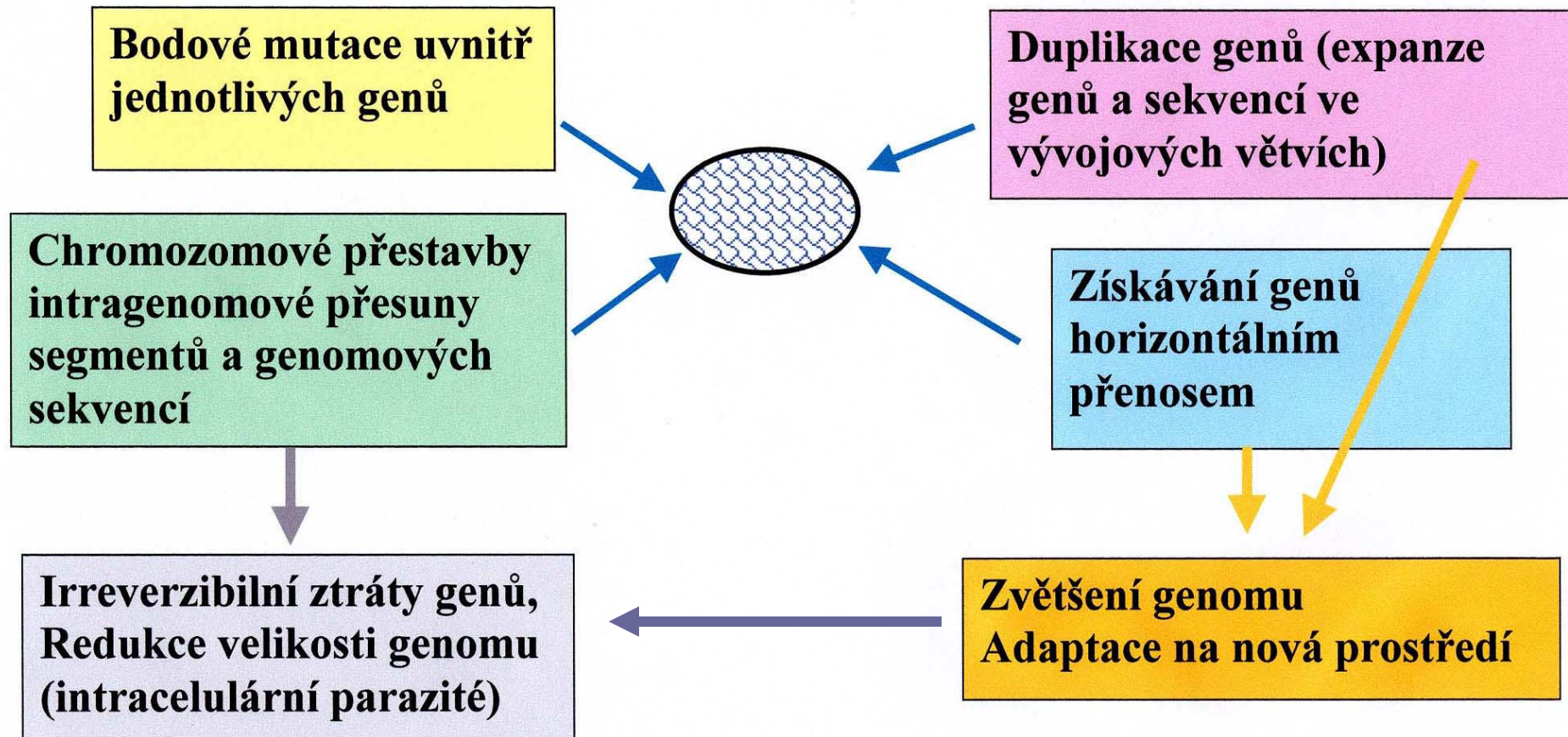
Rychlé a rozsáhlé změny ve struktuře a informačním obsahu genomu

- Vnitřní přestavby
- Získávání a ztráty genů a genetických elementů

Vývoj kmenů v rámci druhu

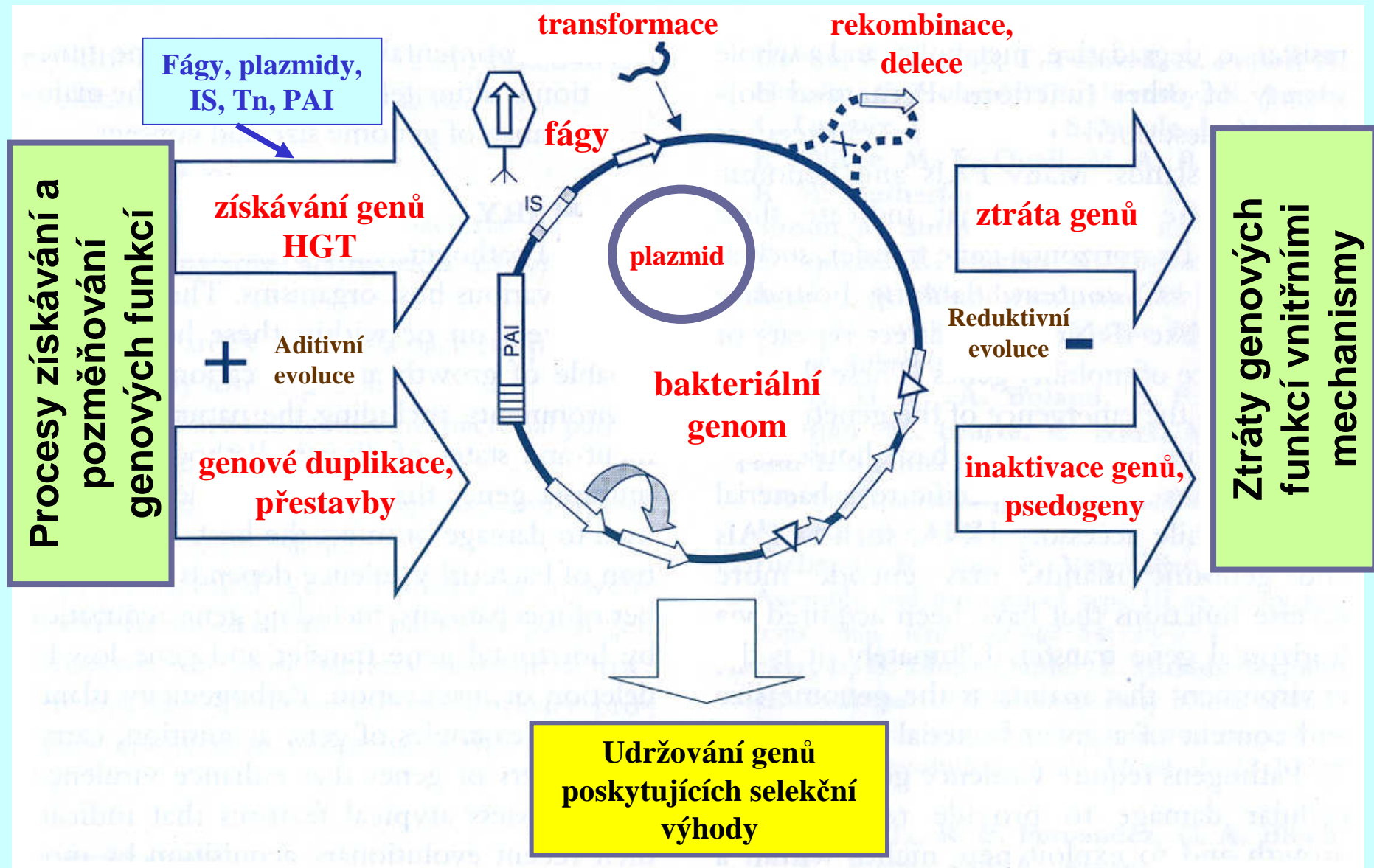
Adaptace na nové podmínky

Podstata změn v obsahu genomu prokaryot v průběhu evoluce



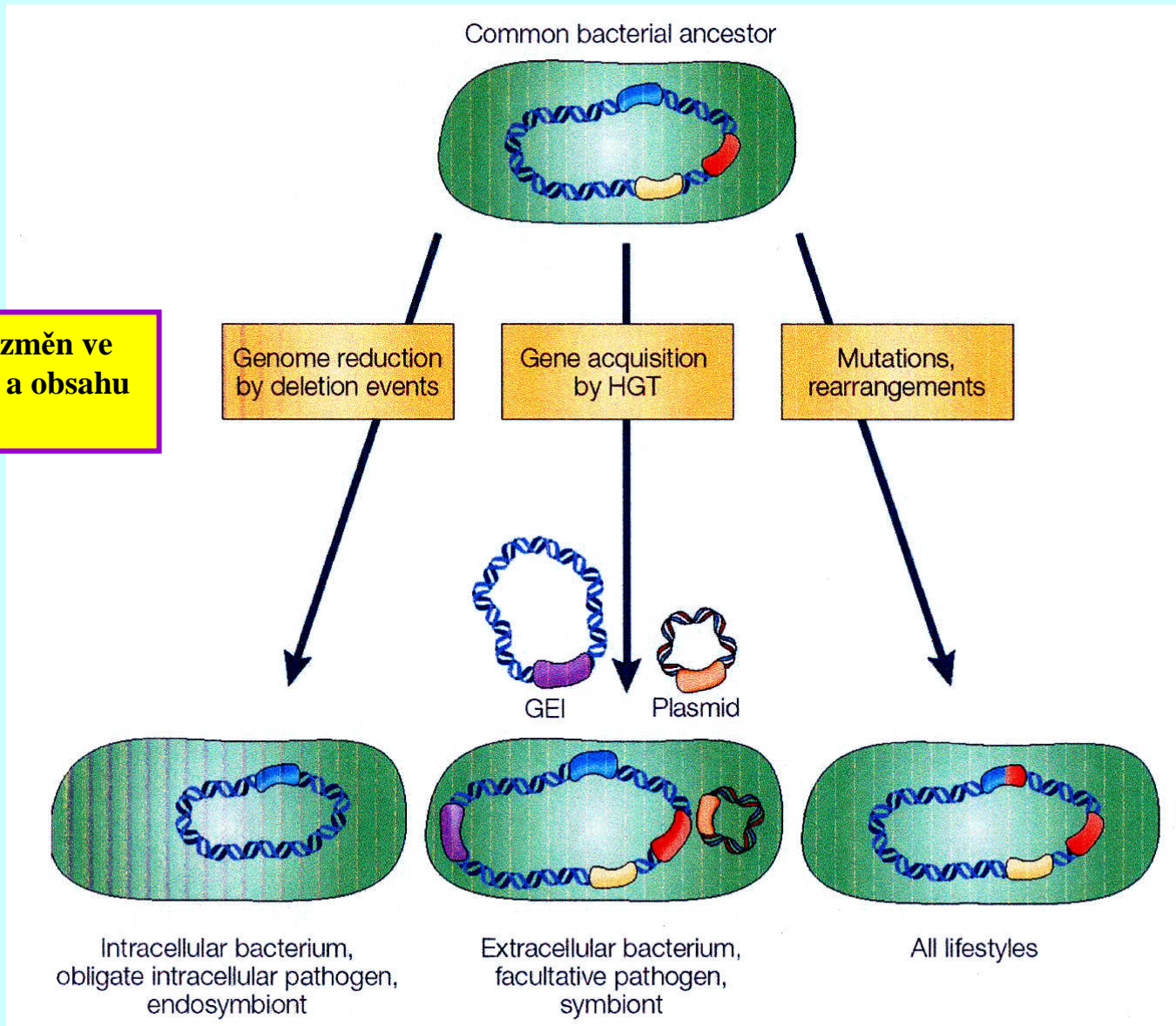
Důsledek: Pořadí genů je zachováno jen u velmi blízké příbuzných druhů

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. Zisk 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
Genomové ostrovy a ostrůvky	HGT, integrace a delece velkých úseků DNA
B. Ztráta vlastností	
Bodové mutace	Změny v genové expresi, ztráta funkce genů
Homologií rekombinace	Přeskupení DNA, delece DNA, integrace genů získaných HGT
Transpozice	Změny genové exprese, ztráta funkce genů

Vnitřní přestavby replikonů navozené přítomností repeticí

Typy repeticí

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

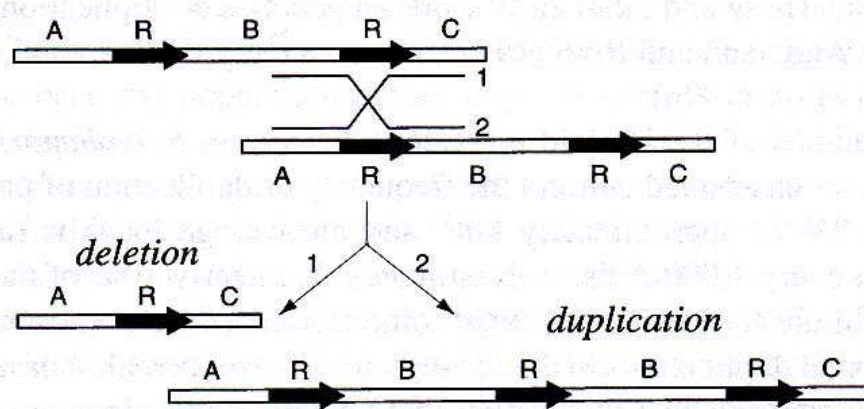
Typ přestavby

- Duplikace (amplifikace)
- Delece
- Inverze

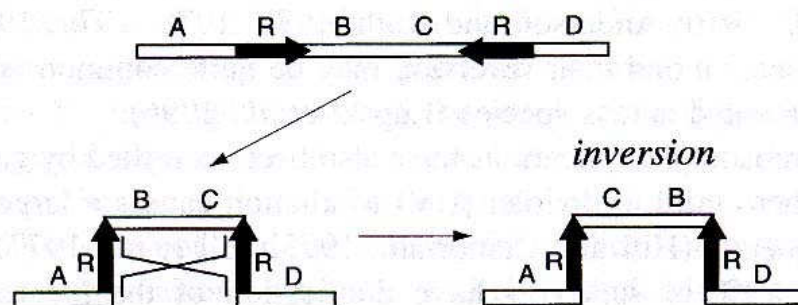
Mechanismus

Homologní rekombinace
Transpozice
Místně-specifická rekombinace
Nerovnoměrný crossing-over

přímá opakování (repetic)



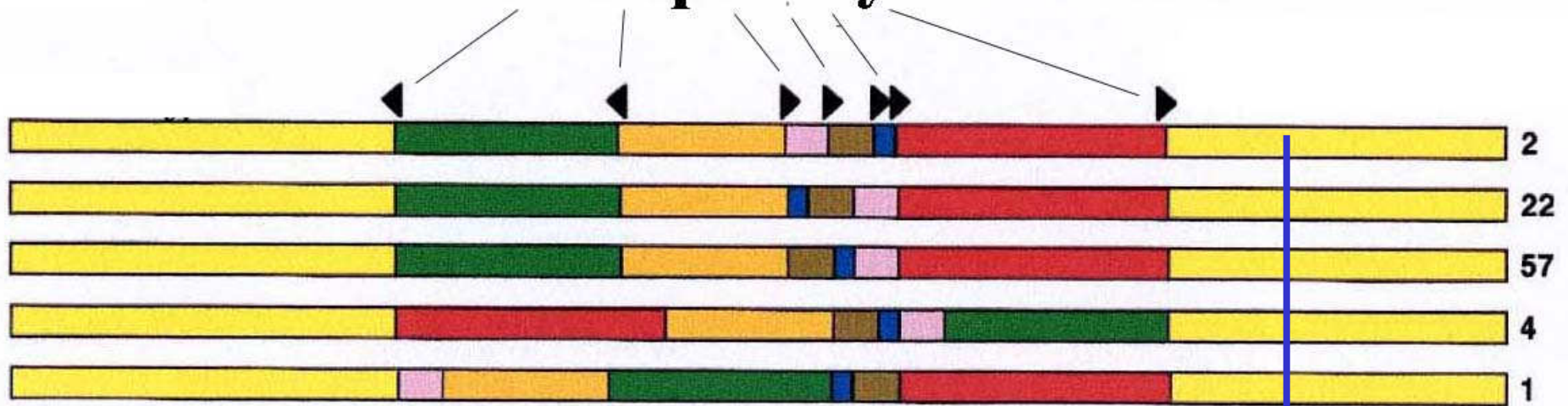
obrácená opakování (repetic)



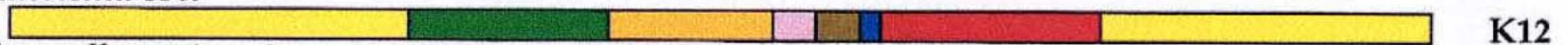
**Přestavby navozené interakcí repetic
(homologní rekombinace)**

Genomová přeskupení navozená rrn operony u kmenů *Salmonella typhi*

rrn operony



Escherichia coli



K12

Salmonella typimurium



LT2

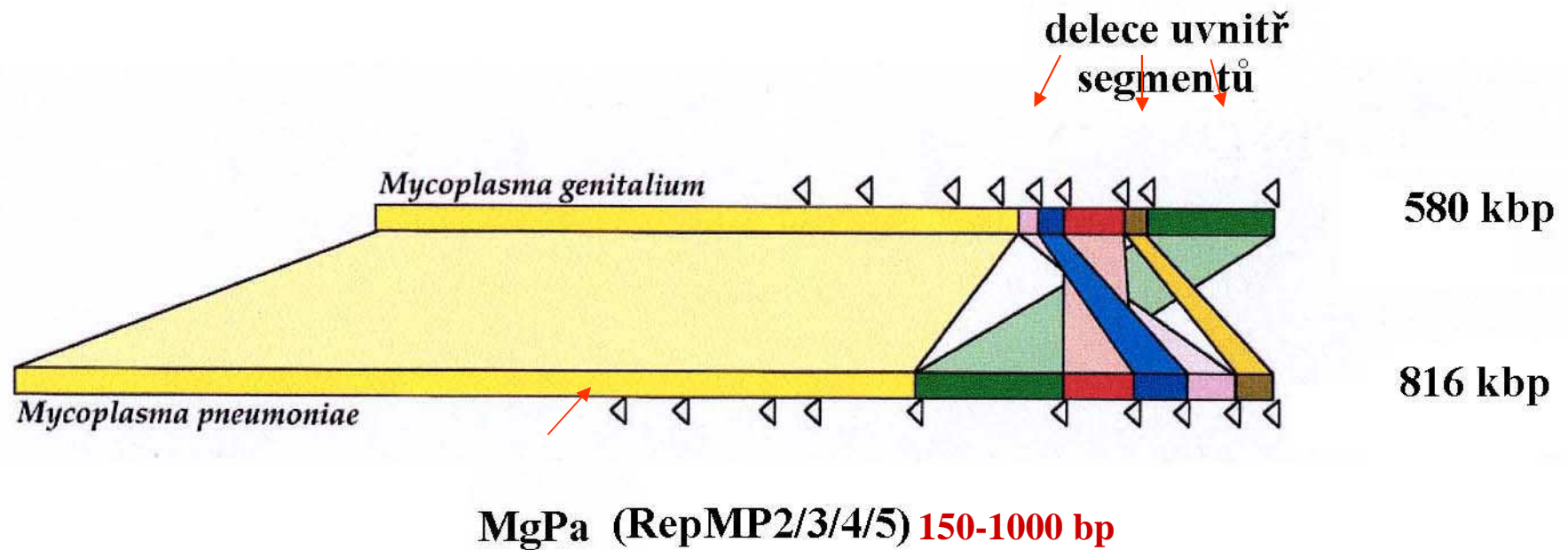
Salmonella paratyphi



B



Rozdíly ve struktuře genomů příbuzných druhů

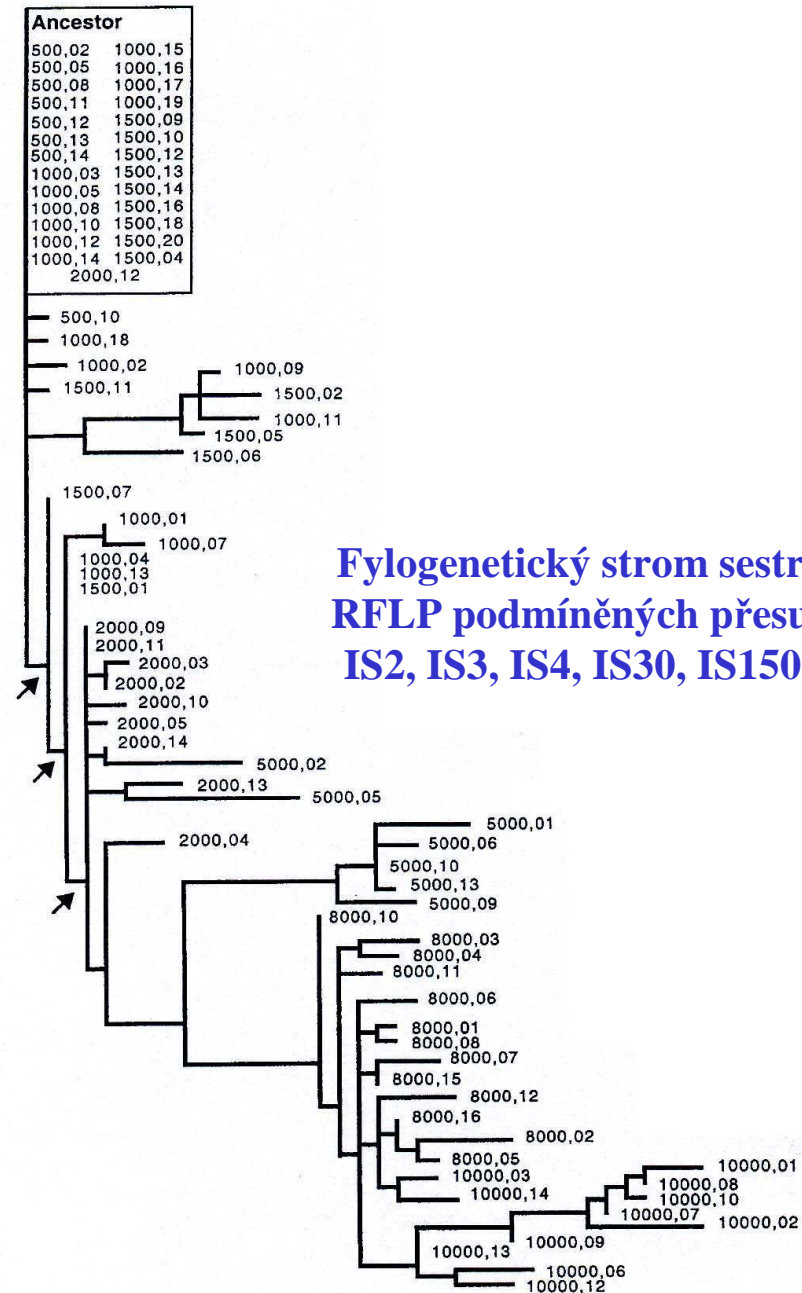


Změny v genomu po
dlouhodobém uchování
kultur *E. coli* a *S. typhimurium*



- změny lokalizace IS
- změny velikosti genomu
- změny fenotypu

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



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 geny 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

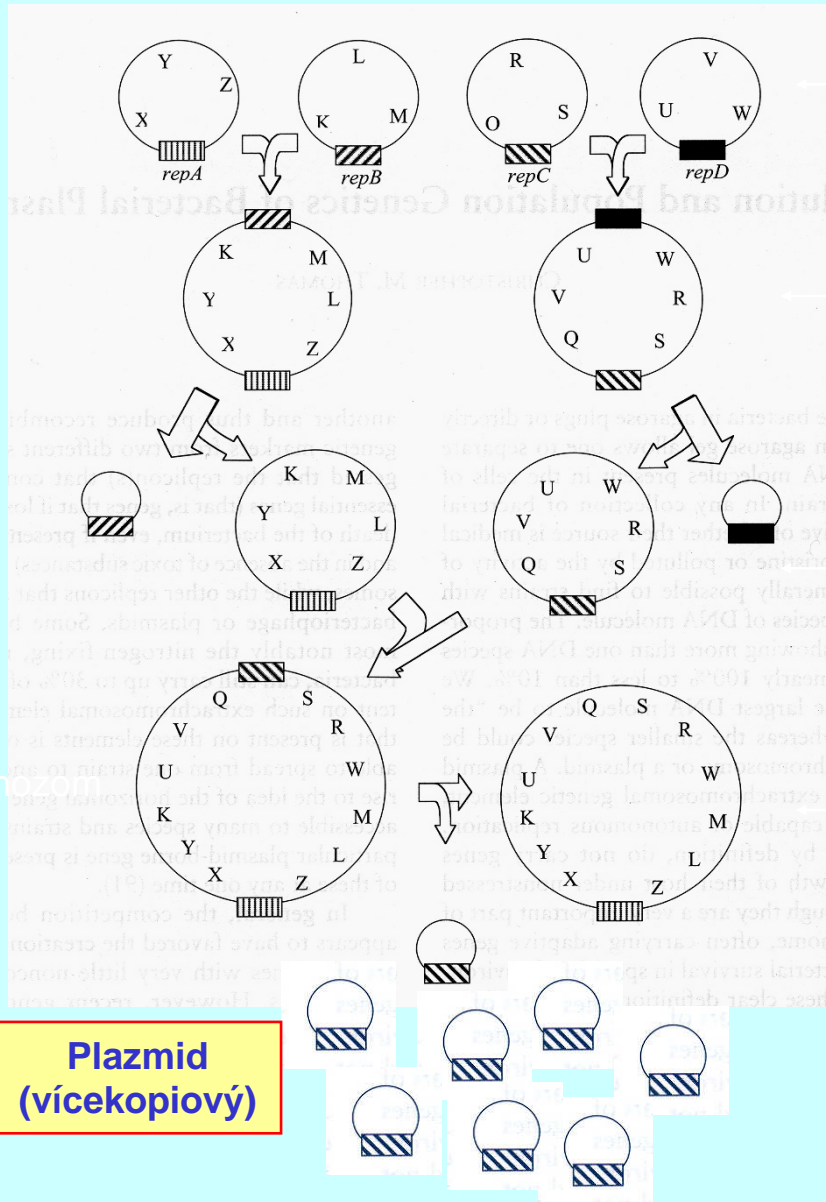
→ paralogní geny – adaptace na nová prostředí

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



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

Plazmid
(vícekopiový)

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)

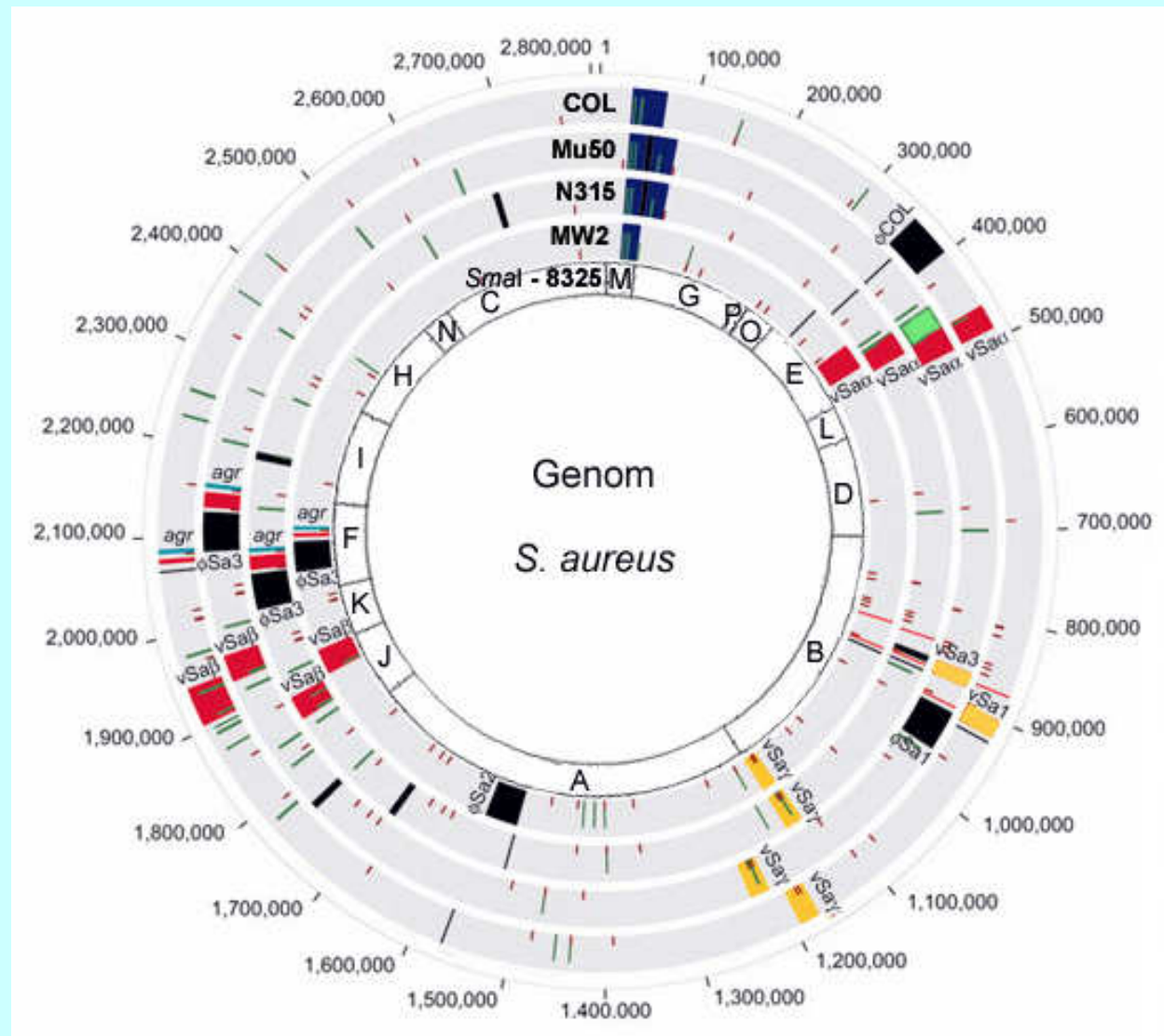
Horizontální přenos genů je spjat s variabilními genetickými elementy

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

Genetický element	Označení	Faktory virulence nebo jiné funkce
Ostrovky 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

Variabilní genetické elementy (VGE) (20 % genomu *S. aureus*)

-  Chromozomové kazety rezistence k meticilinu (SCCmec)
-  Jedinečné ostrovy patogenity (SaPI)
-  Společné genomické ostrovy
-  Profágy
-  Lokus přídatného genového regulátoru (*agr*)
-  Integrované plazmidy
-  Repetitivní elementy (STAR)
-  Inzerční sekvence a transpozony



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 sp.</i>	3,57	3169	219	7,5
<i>Thermotoga maritima</i>	1,86	1846	198	11,63
Archaea				
<i>Aeropyrium 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

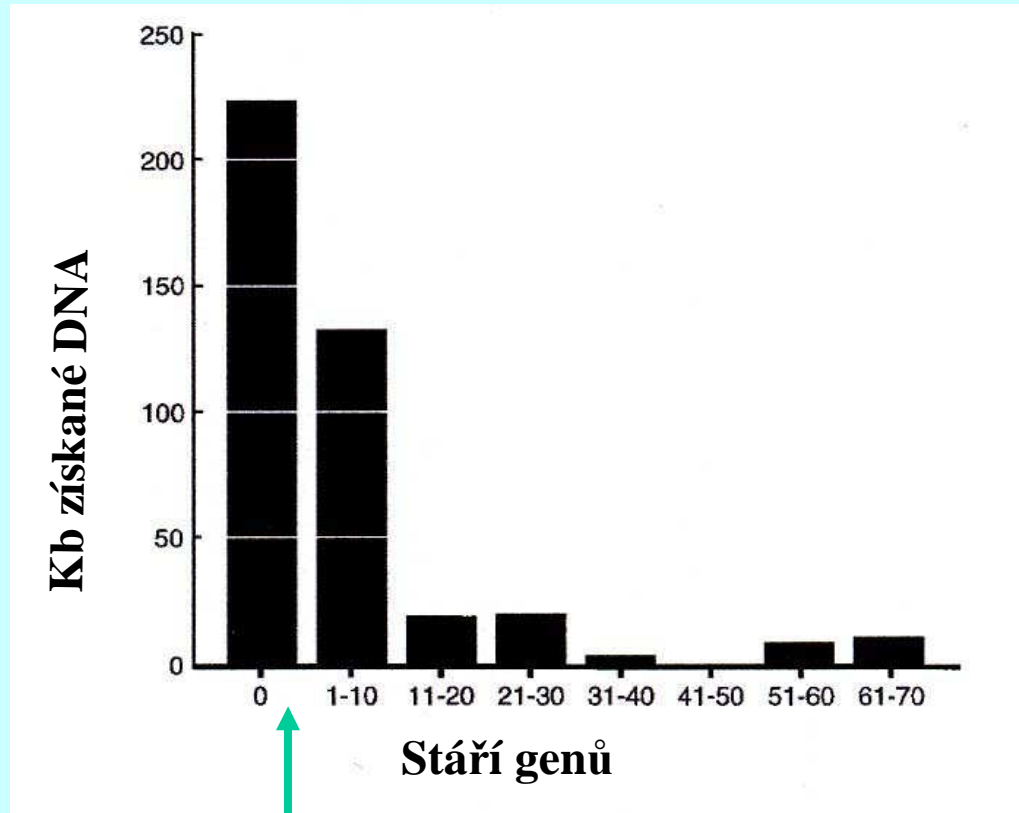


Horizontálně přenesené geny (HGT) u *E. coli* K12 MG1655 (po divergenci *E. coli* a *S. typhimurium*)

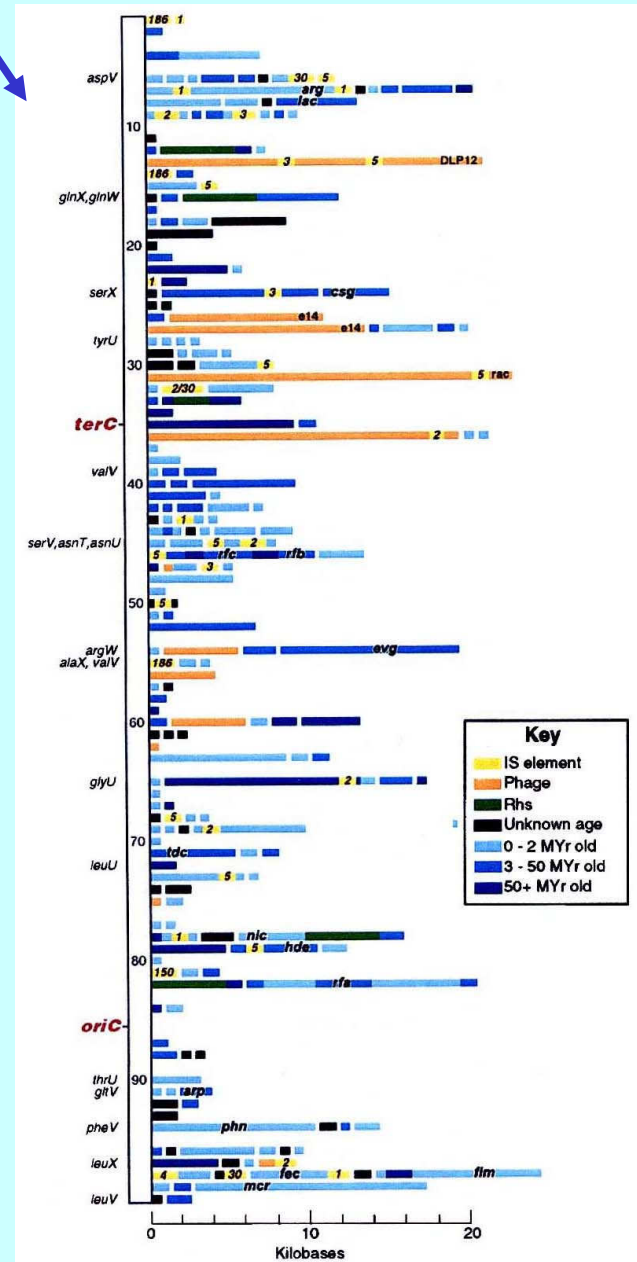
**Genom *E. coli* obsahuje relikty 755 HGT
(18% genomu = 548 kb, 234 přenosových událostí)**

- 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**

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



IS sekvence,
profágy, Rhs



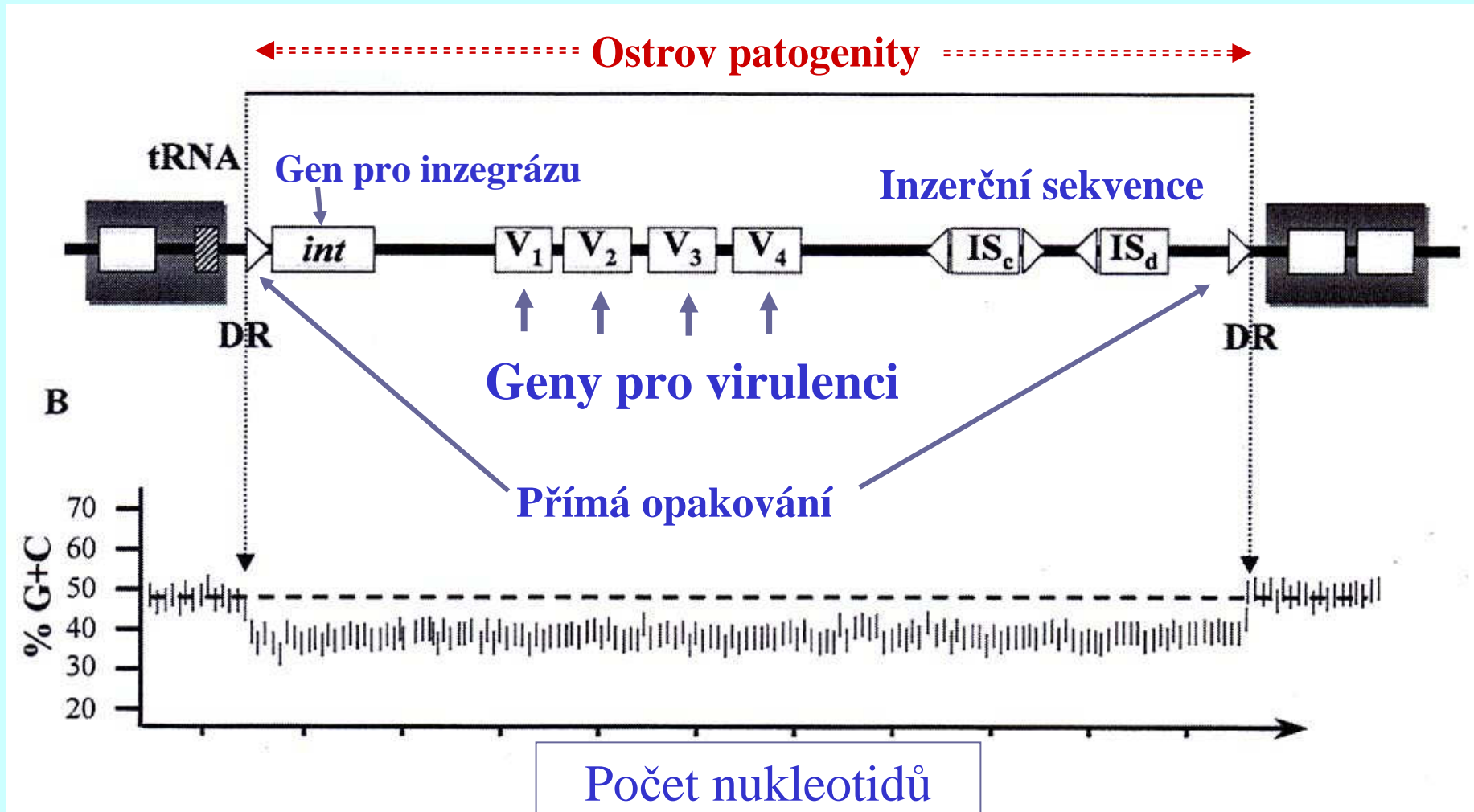
Genomické ostrovy („fitness“ ostrovy)

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

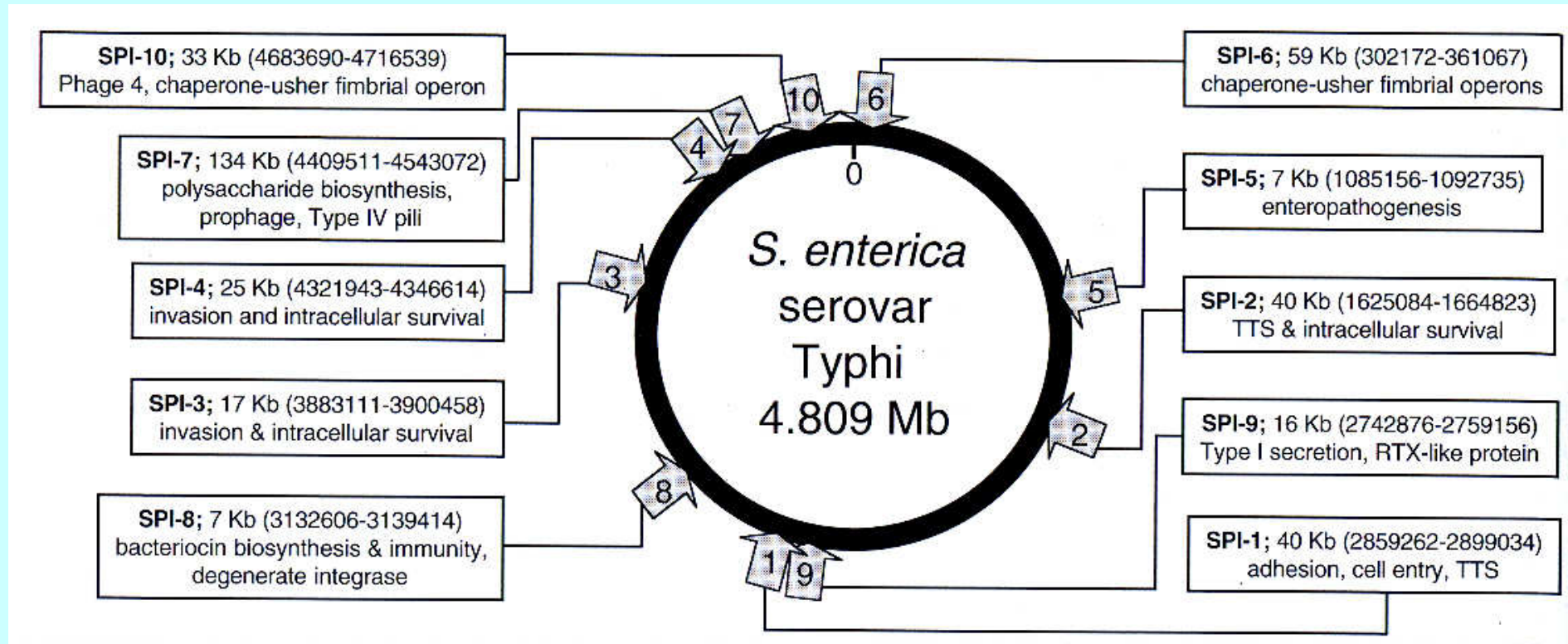
Obecná struktura ostrovů patogenity



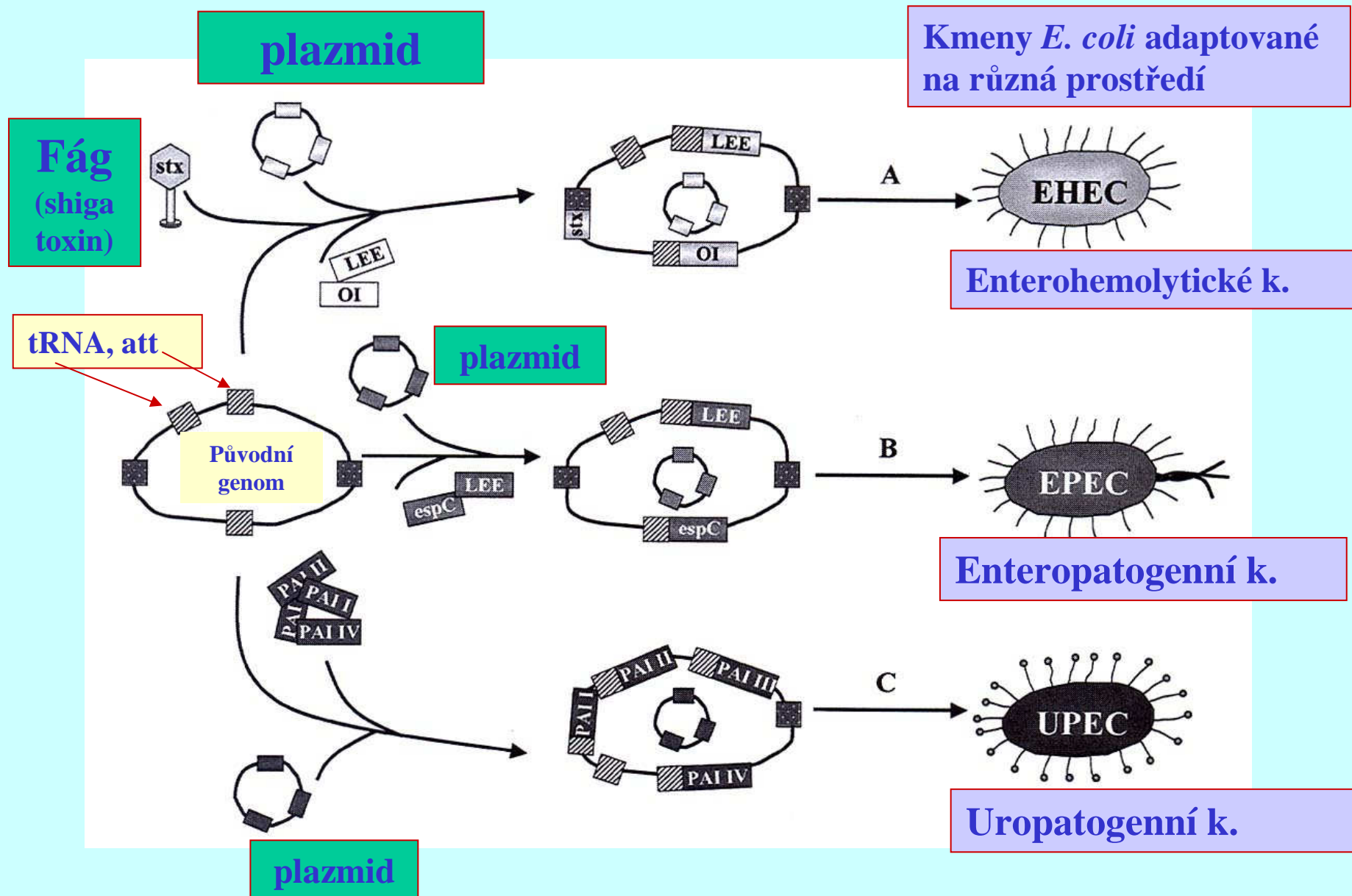
Význačné rysy ostrovů patogenity

- ◆ Nesou jeden nebo několik genů pro virulenci
- ◆ Jsou přítomny jen u patogenních kmenů daného druhu
- ◆ Představují relativně velké úseky genomu (10 – 200 kb)
- ◆ Mají odlišný obsah GC a jiné využívání kodonů
- ◆ Jsou často umístěny poblíž genů pro tRNA (kotvy pro inzerci cizí)
- ◆ Jsou často spojeny s mobilními genetickými elementy.
- ◆ Často jsou ohraničeny DR (16-130 bp) - rozpoznávací místa pro enzymy zajišťující integraci a excizi mobilních elementů (integráza nebo transponáza)
- ◆ Jsou často nestabilní a jsou deletovány s různými frekvencemi.
- ◆ Mají mozaikovitou strukturu – jsou složeny z elementů, které se během evoluce v různé době a z různých zdrojů akumulovaly do určitých míst.

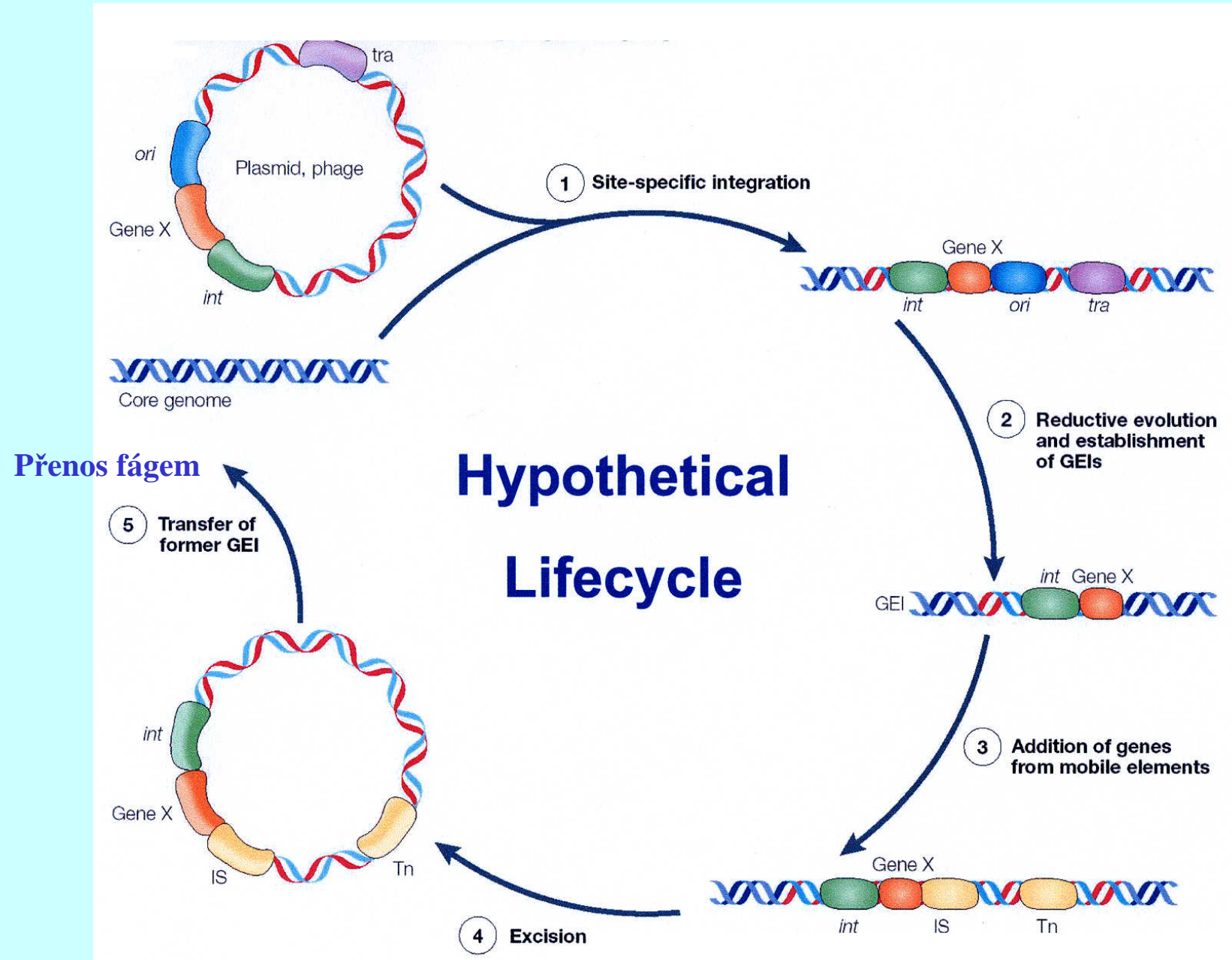
Distribuce ostrovů patogenity u *S. enteritica* serovar Typhi



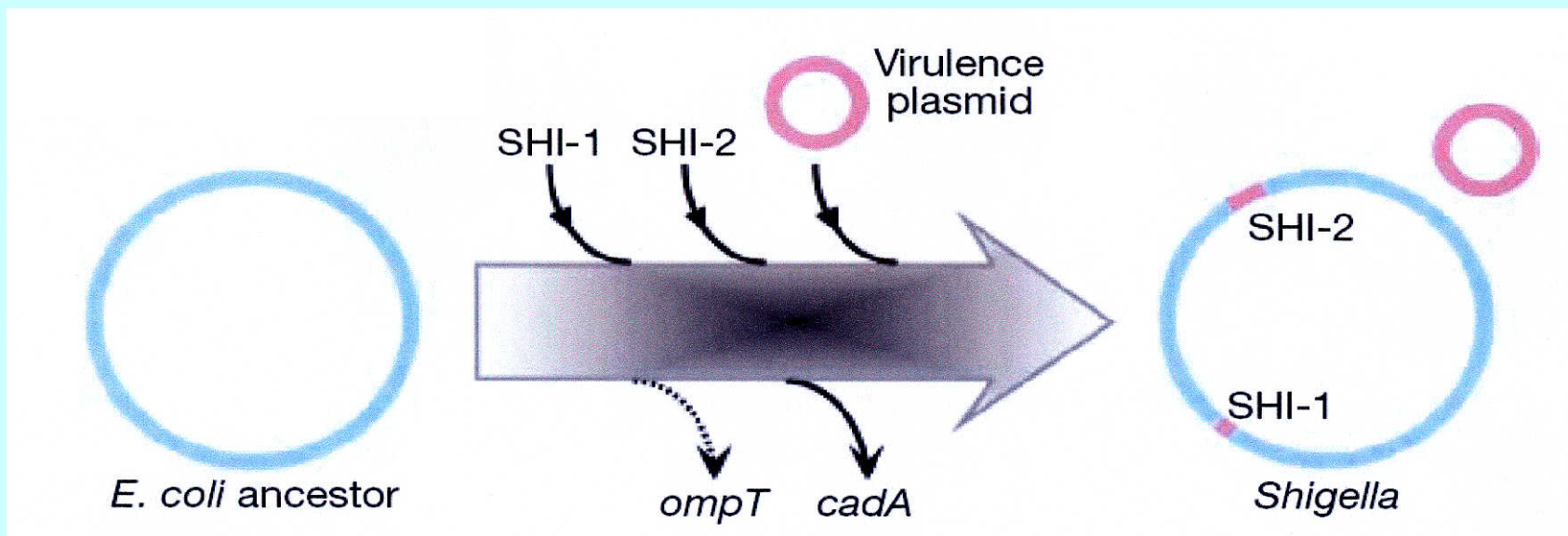
Model vzniku ostrovů patogenity u patogenních *E. coli*



Vznik genomických ostrovů u patogenních a environmentálních mikrobu



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*

r. Shigella x Escherichia coli K-12

90% homologie DNA (!)

Kolinearita genů

Rekombinace po HGT

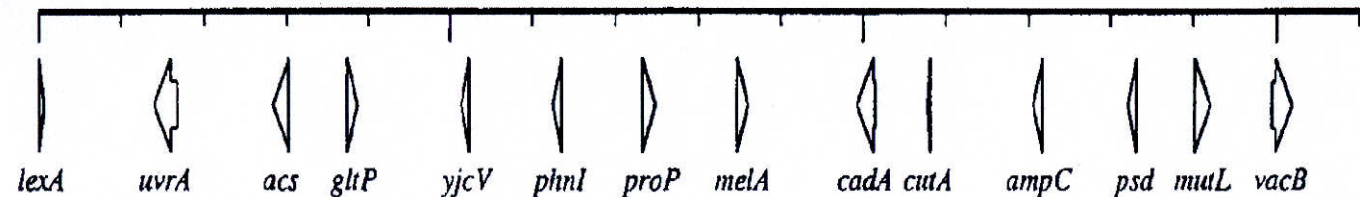
Vliv ztráty genů ztráty genů na patogenitu enterobakterí

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

E. coli K-12

4254428 bp

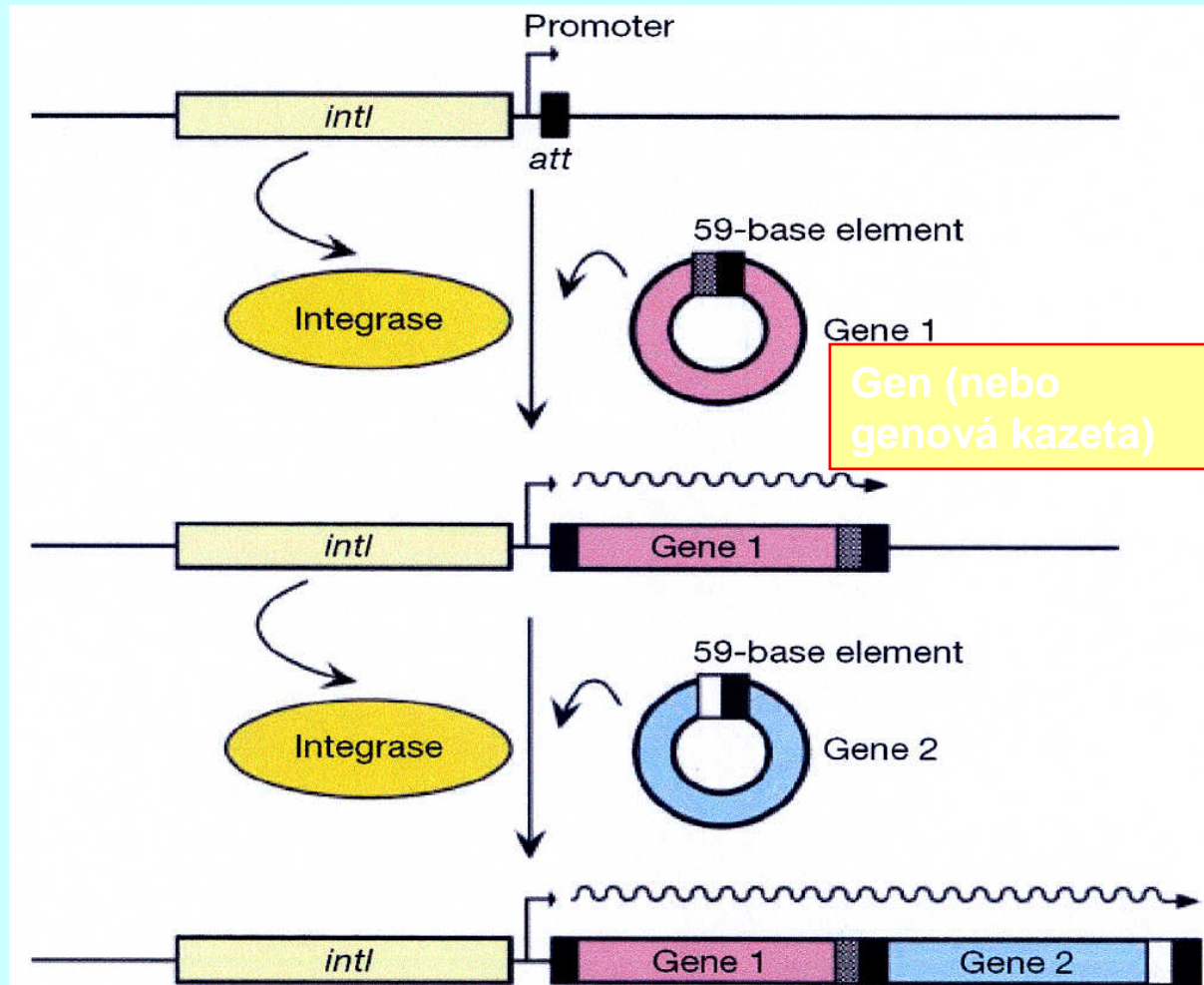
4406306 bp



	<i>lexA</i>	<i>uvrA</i>	<i>acs</i>	<i>gltP</i>	<i>yjcV</i>	<i>phnI</i>	<i>proP</i>	<i>melA</i>	<i>cadA</i>	<i>cutA</i>	<i>ampC</i>	<i>psd</i>	<i>muL</i>	<i>vacB</i>
<i>S. flexneri</i> 2a	+	+	+	+	-	-	+	-	-	-	-	+	+	+
<i>S. dysenteriae</i> 2	+	+	+	+	-	+	+	-	-	-	-	+	+	+
<i>S. boydii</i> 14	+	+	+	+	-	-	+	-	-	+	+	-	+	+
<i>S. sonnei</i>	+	+	+	+	-	-	+	-	-	+	+	+	+	+
EIEC O124:NM	+	+	-	+	-	-	+	-	-	-	+	+	+	+
MC4100	+	+	+	+	+	+	+	+	+	+	+	+	+	+

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)

ZACHYTÁVÁNÍ GENŮ INTEGRONY



Genová kazeta v CTn (v plazmidu)

Integron obsahuje:

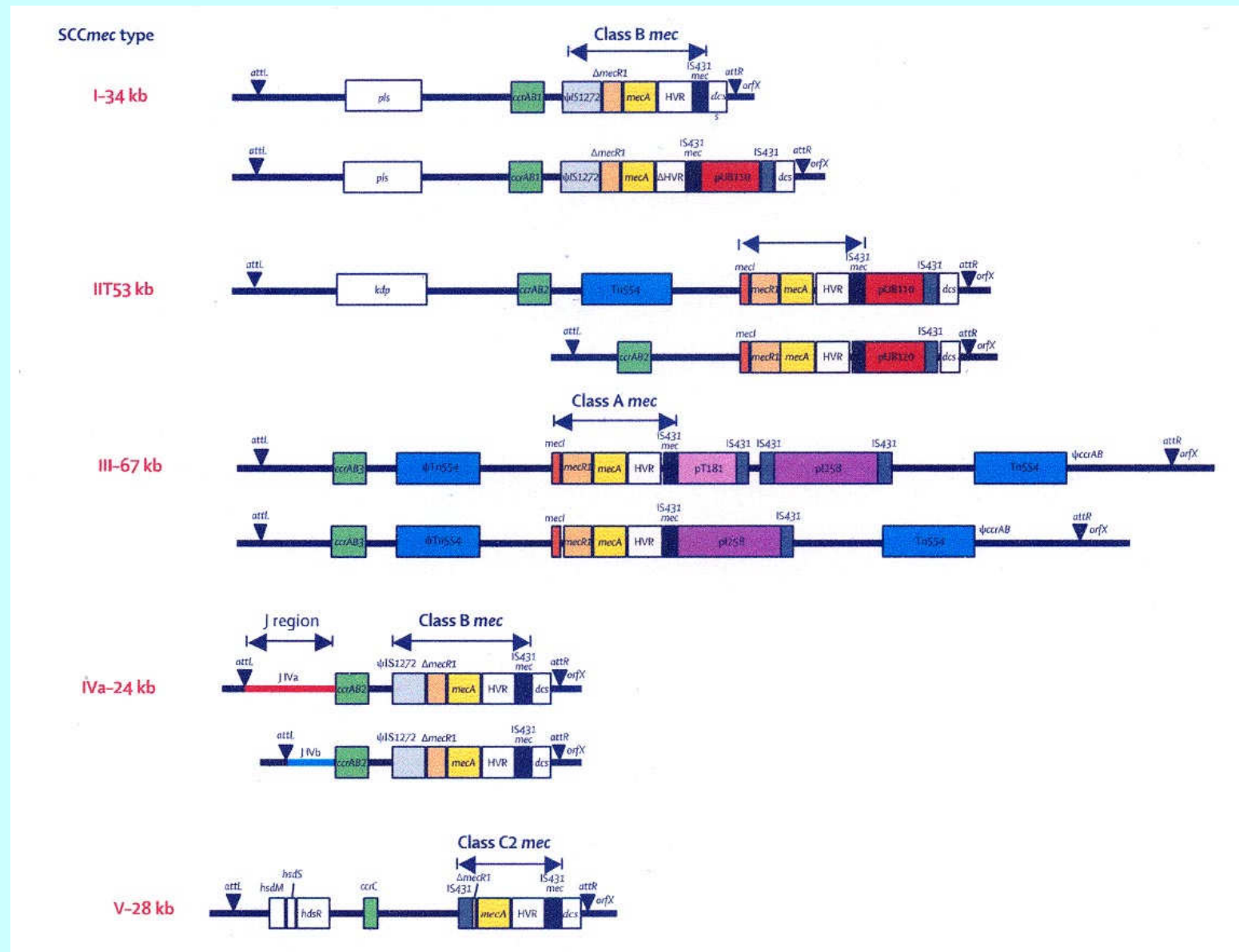
1. att místo, umožňující opakované zachycení genů nebo genových kazet

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

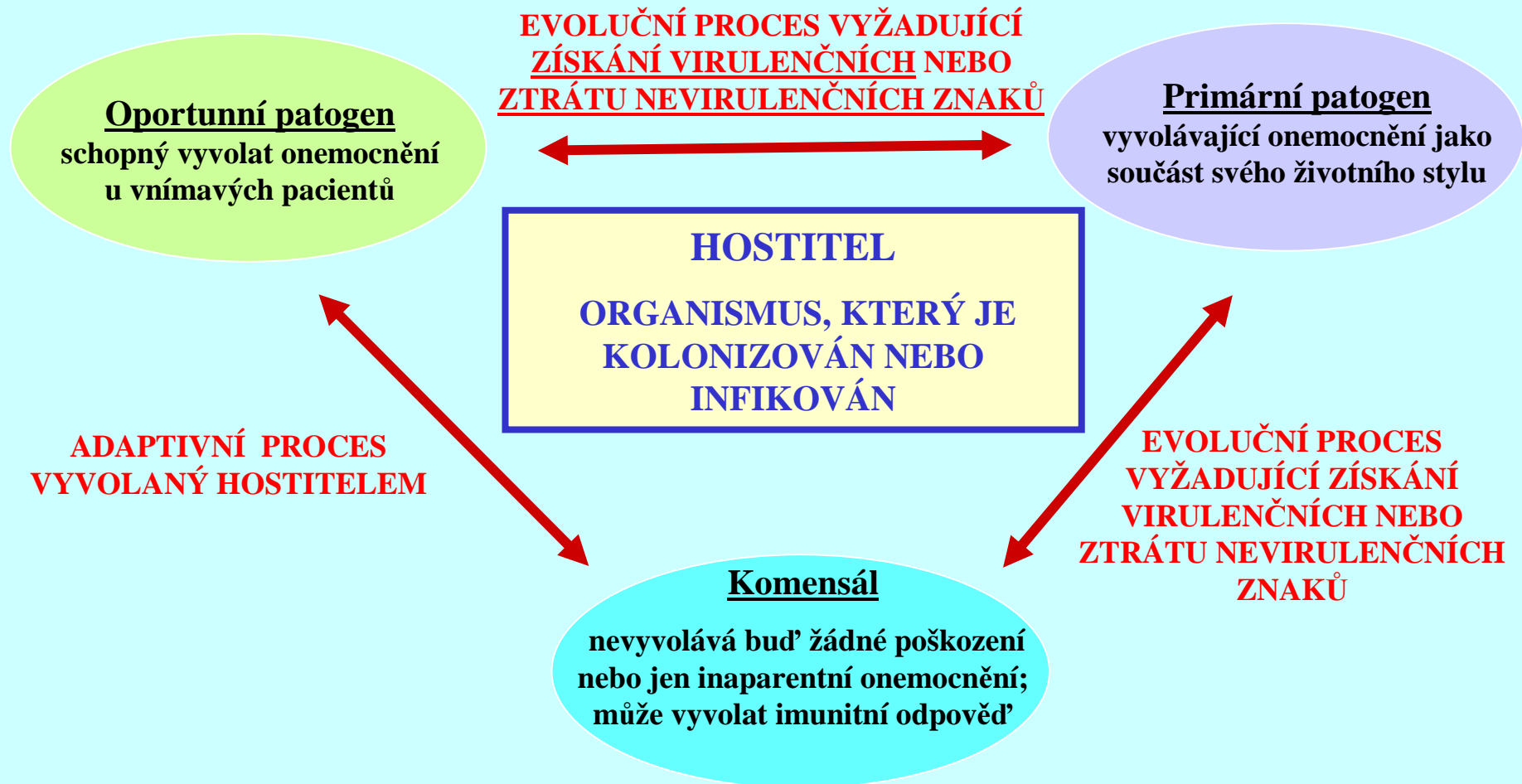
3. Promotor umožňující expresi vloženého genu

***Vibrio cholerae* – obsahuje superintegrony s mnoha genovými kazetami**

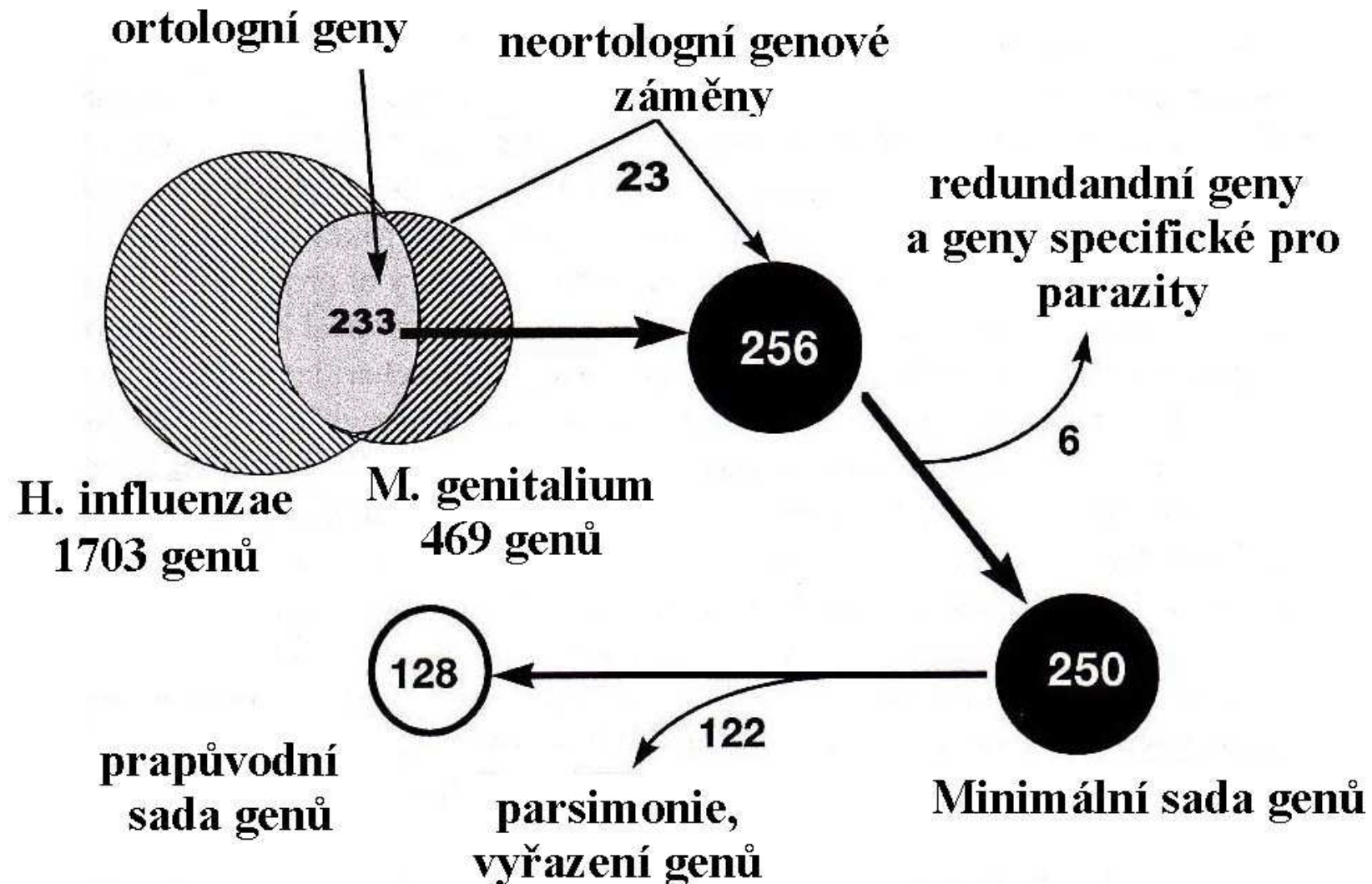
Typy stafylokokových chromozomových kazet (SCCmec) zodpovědných za rezistenci kmenů *S. aureus* k meticilinu



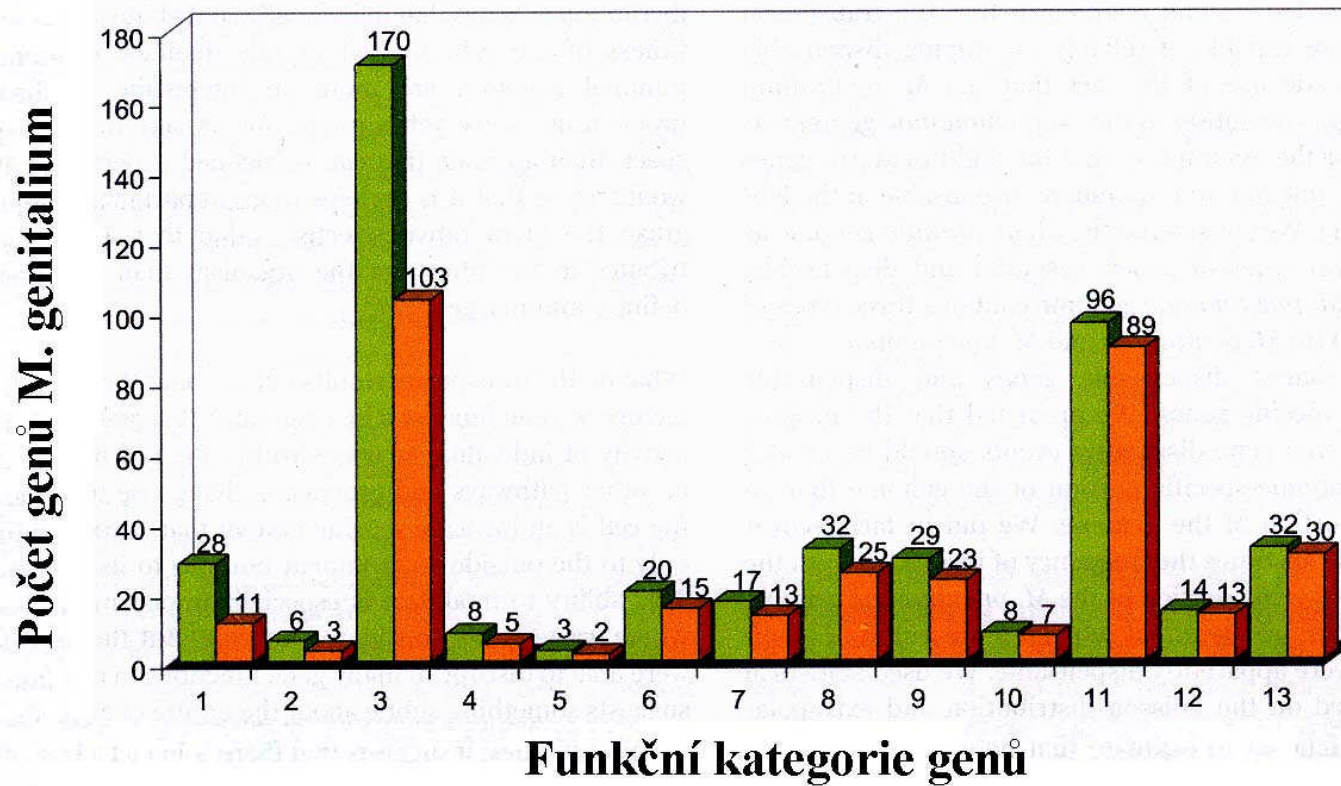
Interakce patogen-hostitel u bakteriálních infekčních onemocnění



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 <i>M. genitalium</i> |
| ■ | Počet genů, které nebyly přerušeny při transpozonové mutagenezi |

Závěry vyvozené z analýzy minimálních genomů

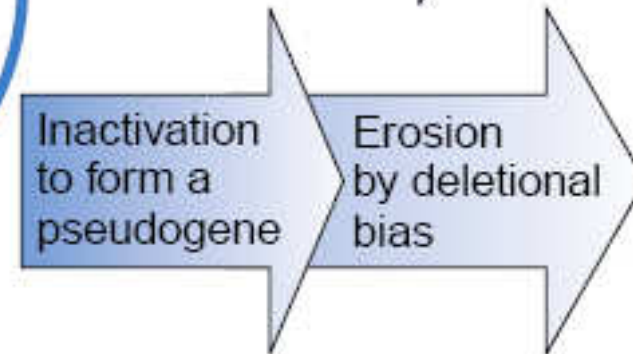
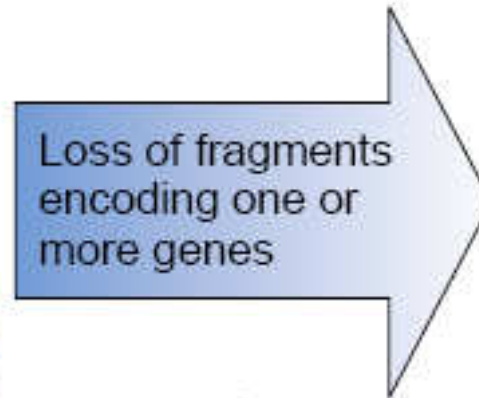
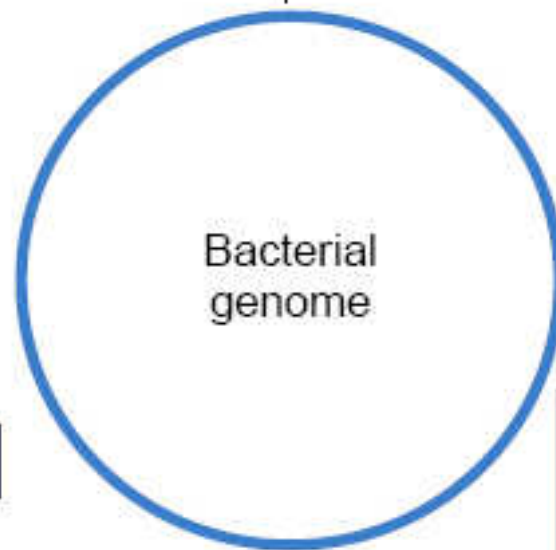
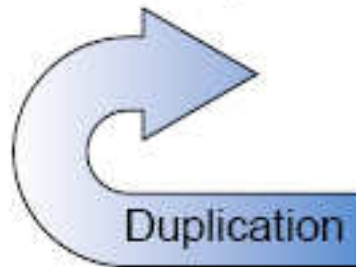
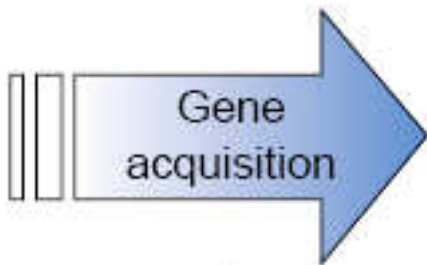
- Každý genom je složen ze dvou typů genů
 - Esenciální geny zajišťující základní biologické procesy
 - Geny pro dosažení selektivní výhody v daném prostředí
- Prostředí určuje, který gen je pro daný druh základní a který postradatelný
- Zhruba třetina (~100) esenciálních genů nemá žádnou ze známých funkcí

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.

Increase in DNA content involving selection for gene functions

Removal of DNA through mutation and genetic drift



TRENDS in Genetics