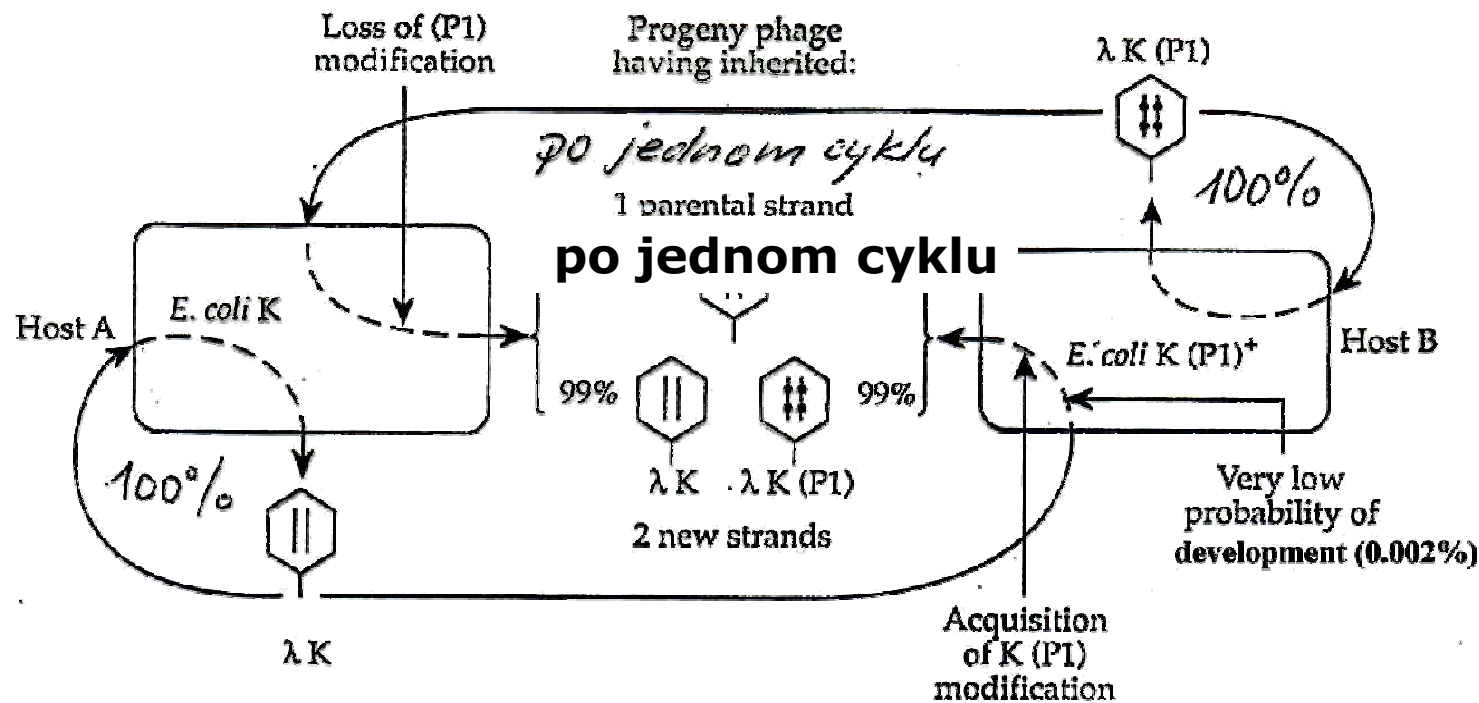


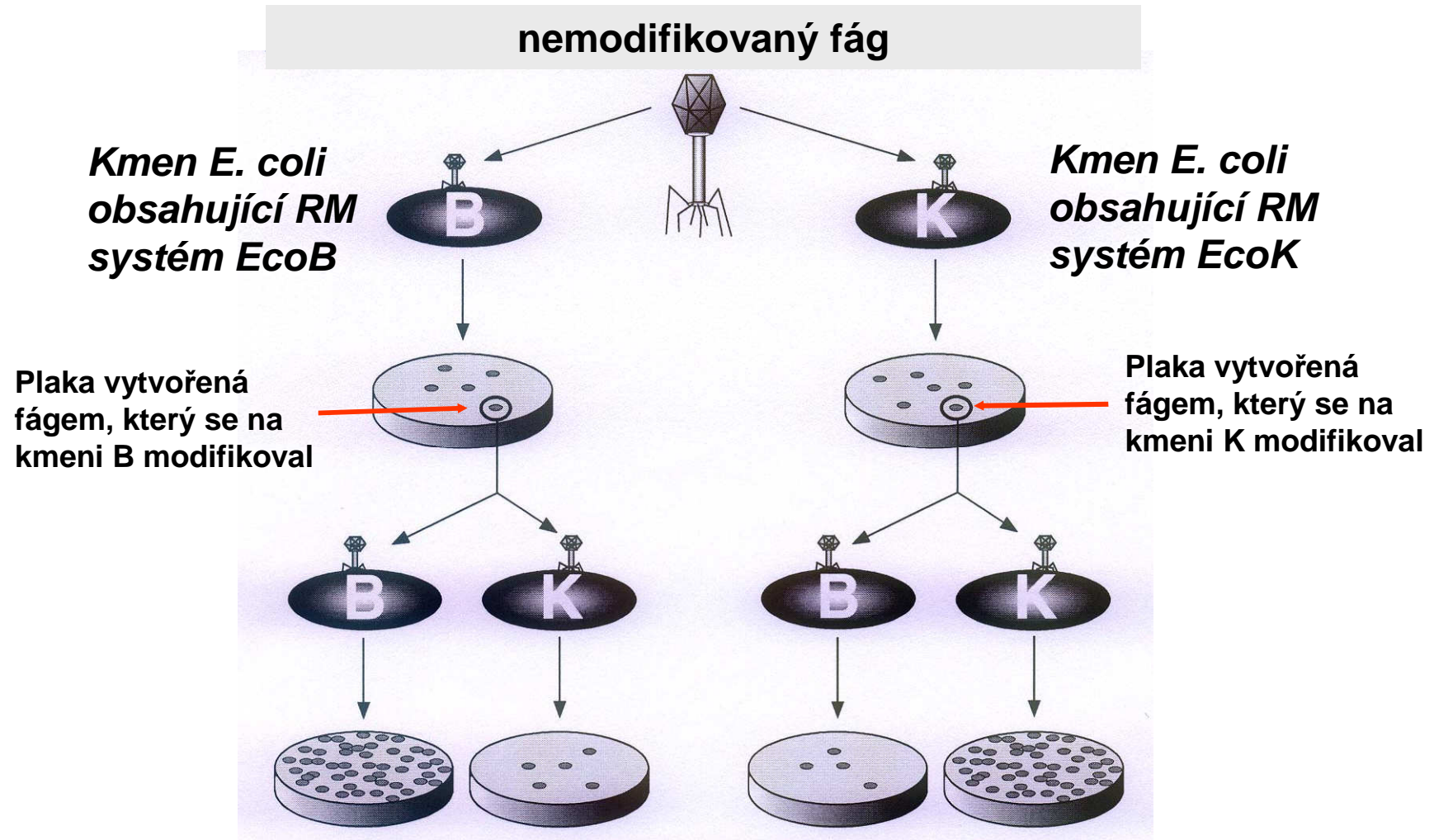
RESTRIKCE A MODIFIKACE FÁGOVÉ DNA



Kmeny E. coli K a K(P1) + mají vzájemně odlišnou hostitelskou specifitu (K a P1) = obsahují odlišné RM-systémy

Fágy po jednom růstovém cyklu na hostiteli získávají nové hostitelské spektrum. Nejedná se o mutaci - změna se týká celé populace a není dědičná – epigenetická záležitost.

Experimentální důkaz přítomnosti a působení RM systémů



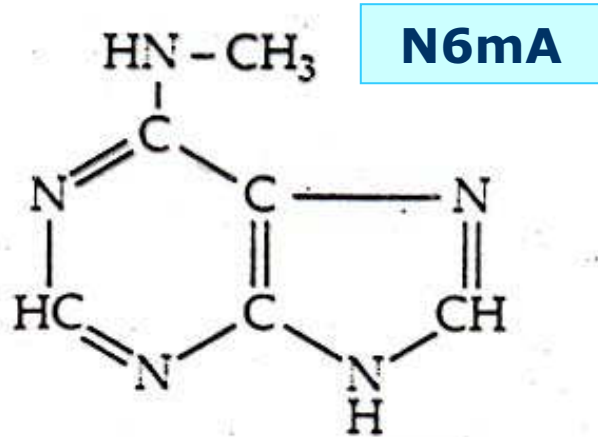
RESTRIKČNĚ-MODIFIKAČNÍ (RM) SYSTÉMY

- Složkami standardního restriktivně-modifikačního (RM) systému jsou sekvencně specifické enzymy:
 - **A. modifikační metyláza (metyltransferáza)**
 - **B. restriktivní endonukleáza**

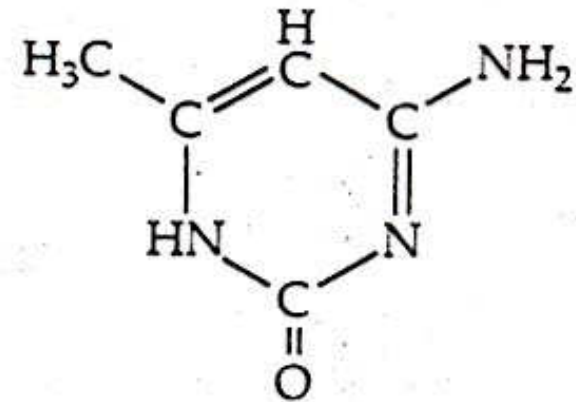
- Restriktivní a modifikační podléhá jen dsDNA:
 1. Při infekci fágem
 2. Při přenosech DNA transdukací a konjugací, omezeně transformací (při umělé tr.)

Monitrování příjmu exogenní DNA – „imunitní systém prokaryot“

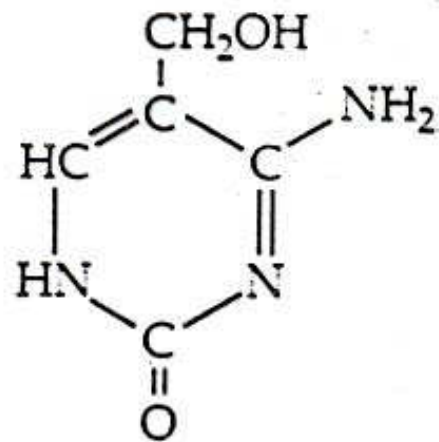
Metylované báze v DNA



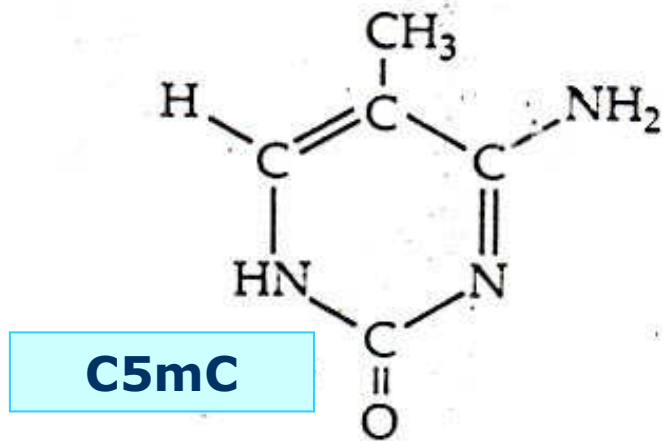
N6-methyl-adenine



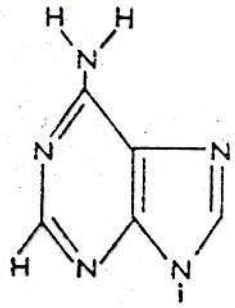
C4-methyl-cytosine



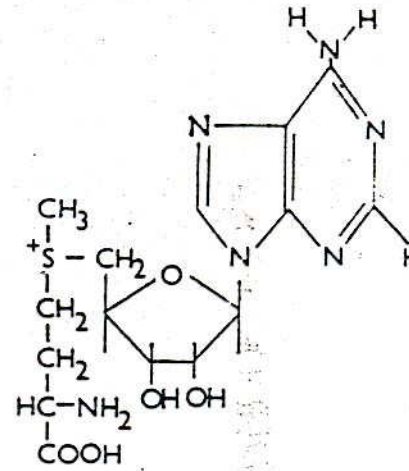
C5-OHmethyl-cytosine



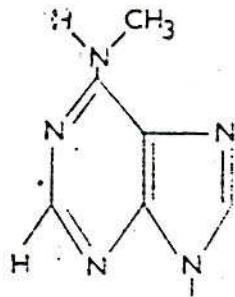
C5-methyl-cytosine



adenin

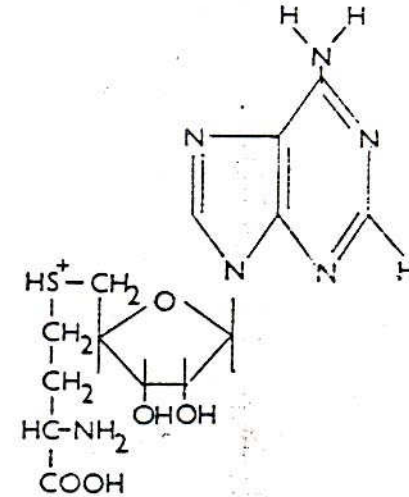


S-adenozylmetionin



6-metylaminoapurin

N6mA



S-adenozylhomocystein

Donor
metylové
skupiny:
SAM
(AdoMet)

SYSTÉMY RESTRIKCE A METYLACE

1. Hsd systémy (host specificity of DNA)

- třídy (typy) I, II, III, IV
restrikční a metylační aktivita

2. Systémy restrikce modifikované DNA

- Mar, Mrr (methylated-adenine), DpnI
- Mcr (methylated-cytosine)

3. Systémy metylující specifické sekvence DNA

- Dam (adenine-methylation)
- Dcm (cytosine-methylation)

Table 6.1 Characteristics of the four classes of Hsd systems.

Properties	Class I	Class II 95%	Class III ^a	Class IV ^a
R and M activities	Trimeric complex	Single enzymes	Trimeric complex	Single protein + 2nd methylase
Genetic organization	Two transcripts: <i>hsdR</i> , <i>hsdSM</i>	Usually two transcripts: <i>hsdR</i> , <i>hsdM</i>	Two transcripts: <i>hsdR</i> , <i>hsdSM</i>	1-2 geny,
Recognition sequence	Non-symmetric, hyphenated	Usually palindromic	Non-symmetric	Palindromic, hyphenated
Requirements for restriction methylation	SAM ^b SAM, Mg ²⁺ , ATP	Mg ²⁺ SAM	SAM, ATP, Mg ⁺⁺ SAM, ATP, Mg ²⁺	Mg ²⁺
Location of cleavage site	Random, at least 10 ³ bases away from recognition site	Usually within the recognition sequence	24–26 bases 3' of recognition site	30 bases 3' of recognition site
Restriction versus methylation	Mutually exclusive	Separate	Simultaneous	protein štěpí jen modifikovanou DNA
Recycling of endonuclease	No	Yes	Yes	
Model systems	<i>EcoB</i> and <i>EcoK</i>	<i>EcoRI</i>	SP1	<i>Eco57I</i> + M- <i>Eco57I</i>

a, Only five Class III and one Class IV cases have been studied. b, SAM: S-adenosylmethionine.

Table 6.2 Recognition sequences of some restriction enzymes.

Enzyme	Name	Organism	Recognition sequence	Observations
Class I	<i>EcoK</i>	<i>E. coli</i> K12	AACN ₆ GTGC	Cleavage $\geq 10^3$ bases away
Class II	<i>EcoRI</i>	<i>E. coli</i> RY13	G/AATTC	Palindromic sequence
	M- <i>EcoRI</i>	<i>E. coli</i> RY13	GAA*TTC	Cognate methylase
	<i>RsrI</i>	<i>Rhodopseudomonas sphaeroides</i>	G/AATTC	Isoschizomer of <i>EcoRI</i>
	<i>AvaI</i>	<i>Anabaena variabilis</i>	C/PyCGPuG	Palindromic degenerate sequence
	<i>DpnI</i>	<i>Diplococcus pneumoniae</i>	GmeA/TC	Necessitates me-DNA
	<i>BamHI</i> <i>MboI</i>	<i>Bacillus amyloliquefaciens</i> <i>Moraxella bovis</i>	G/GATCC GATC	} Compatible enzymes
Class III	SP1	Phage P1	AGACC	Cleavage 24 bases from 3' end
Class IV	<i>Eco57I</i>	<i>E. coli</i> RFL57	CTGGAG	Cleavage 14 bases from 3' end

By convention, only the 5' to 3' strand is shown. N, nucleotide; Py, pyrimidine; Pu, purine; meA, methyl-
Ade; * represents the methylating site; / indicates the cleavage site.

GENY KÓDUJÍCÍ RM-SYSTÉMY JSOU NESENY NA RŮZNÝCH REPLIKONECH

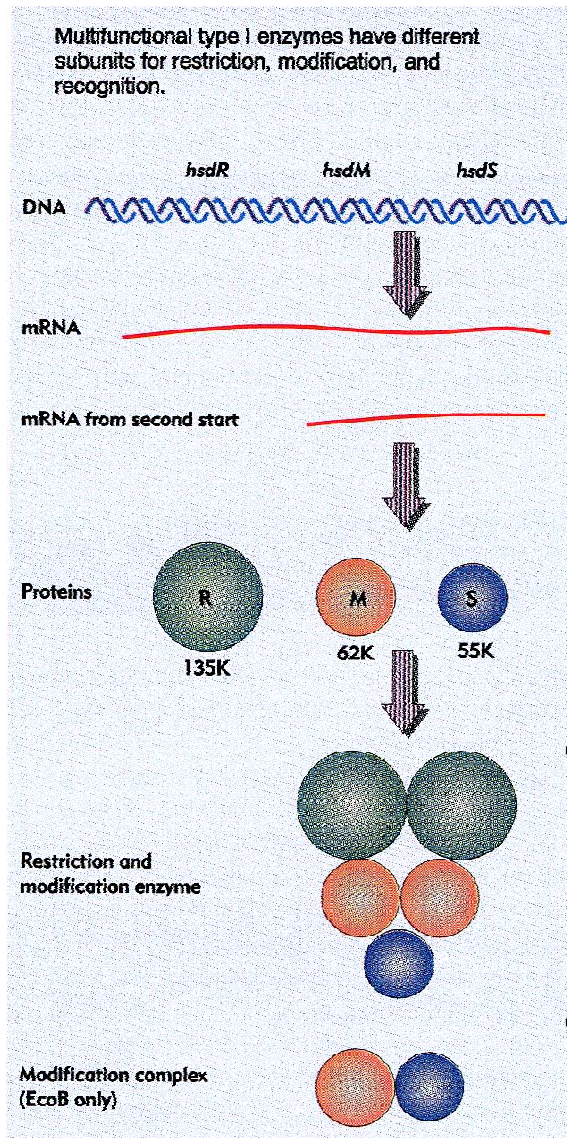
TABLE 2. Organization of Types I and III Restriction–Modification Systems

Group	Members	Recognition sequence	Cellular location of the R–M system
Type IA	<i>EcoK</i>	AAC(N) ₆ GTGC	Chromosomal
	<i>EcoB</i>	TGA(N) ₈ TGCT	Chromosomal
	<i>EcoD</i>	TTA(N) ₇ GTCPy	Chromosomal
	<i>StySB</i>	GAG(N) ₆ PuTAPyG	Chromosomal
	<i>StySP</i>	ACC(N) ₆ GTPuC	Chromosomal
	<i>StySQ</i>	ACC(N) ₆ PuTAPyG	Chromosomal
	<i>EcoDXX1</i>	ATCA(N) ₇ ATTC	Plasmid
Type IB	<i>EcoA</i>	GAG(N) ₇ GTCA	Chromosomal
	<i>EcoE</i>		Chromosomal
Type IC	<i>EcoR124</i>	GAAN ₆ PuTCG	<i>IncFIV</i> plasmid R124
	<i>EcoR124/3</i>	GAAN ₇ PuTCG	<i>IncFIV</i> plasmid R124/3
Type III	<i>EcoP1</i>	AGACC	Prophage P1
	<i>EcoP15</i>	CAGCAG	Plasmid 15B
	<i>HinfIII</i>	CGAAAT	Chromosomal

**lokalizace genů na chromozomu,
plazmidech, nebo profágách**

Přítomnost genů RM systémů na mobilních elementech

Mobilní element	Příklad RM systému
Plazmid	<p><i>paeR71 P. aeruginosa</i> <i>ecoRI E. coli</i> <i>ssoII Shigella sonnei</i> <i>bsp6I Bacillus sp.</i></p>
Bakteriofág/profág	<p><i>hindIII H. influenzae</i> <i>sau42I S. aureus</i> <i>ecoO1091 E. coli</i> <i>bsuMI B. subtilis</i></p>
Integrační komulativní element/genomický ostrov	<p><i>Sth368I Streptococcus thermophilus</i> <i>hsdMS S. aureus</i></p>
Transpozon	<p><i>Rle39BI</i></p>
Integron	<p><i>xbaI Xanthomonas campestris</i> <i>M.Vch0211 Vibrio cholera</i> <i>hphI Vibrio metschnikovii</i> <i>CAA68 Lactococcus lactis</i></p>



PODJEDNOTKOVÉ SLOŽENÍ ENZYMOVÉHO KOMPLEXU TYPU I

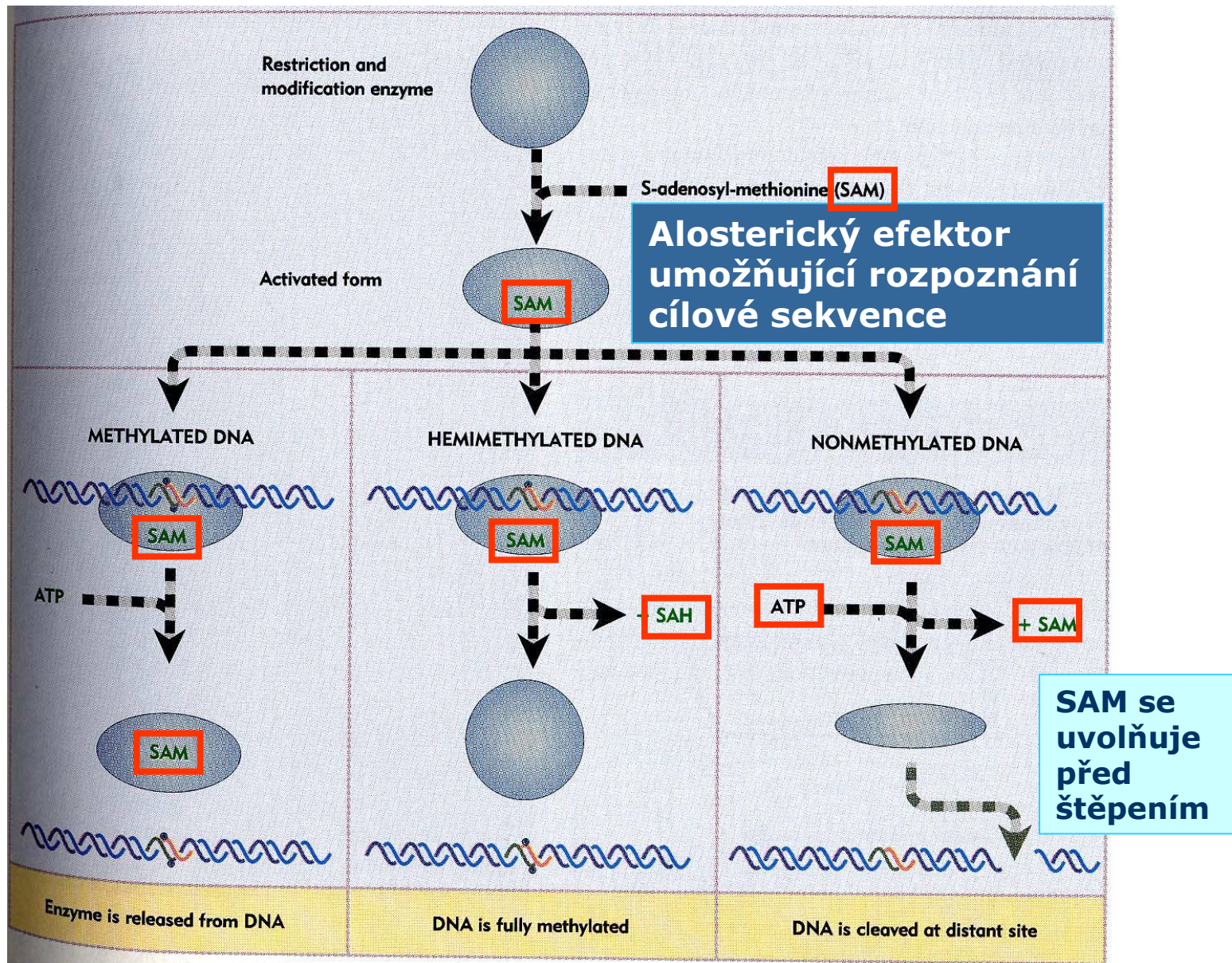
S = rozpoznání cílové sekvence

M = vazebné místo pro SAM a aktivní místo pro metylaci

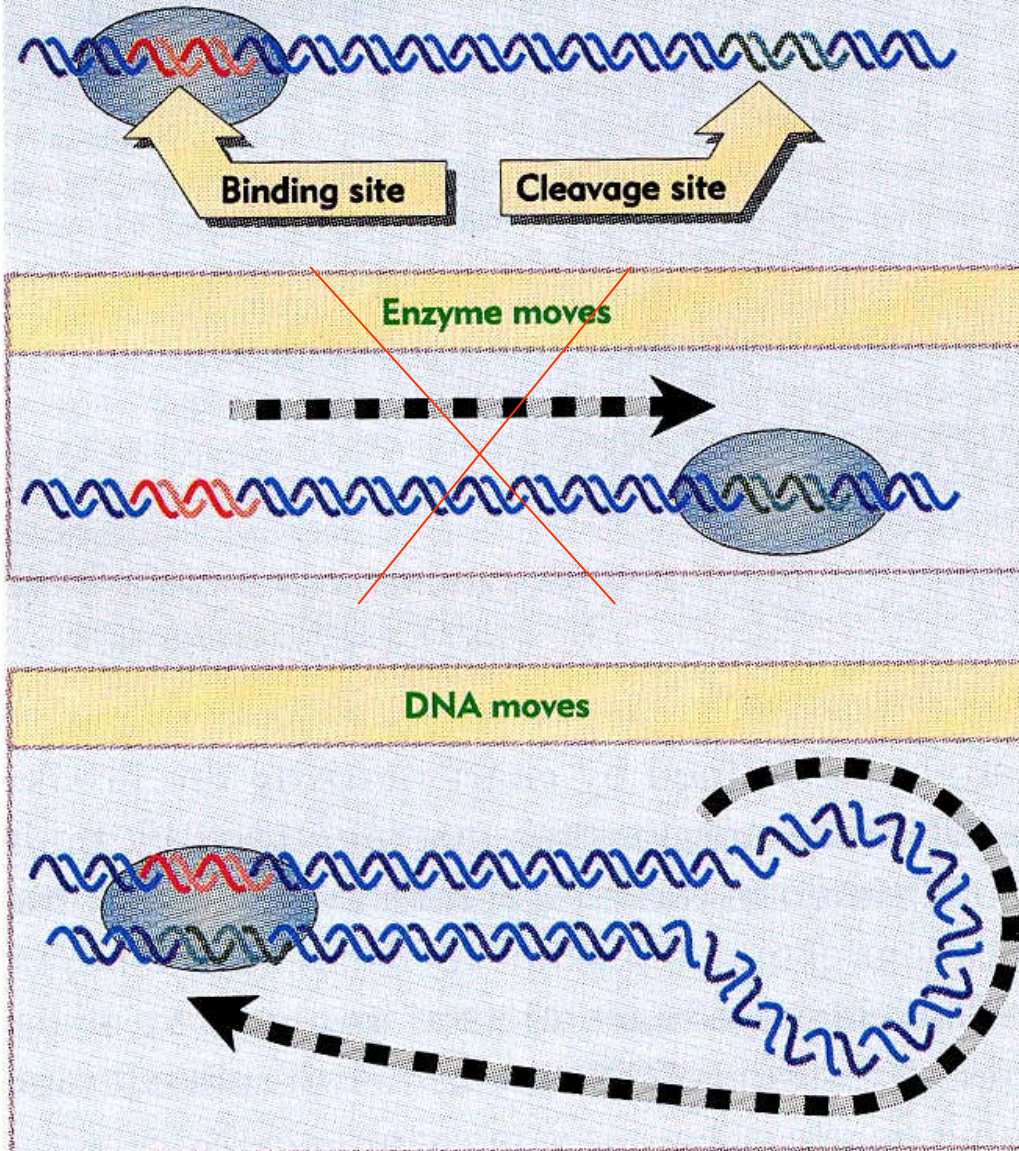
R = aktivní místo pro hydrolýzu ATP, pro translokaci DNA a pro štěpení

Multifunkční komplex

INTERAKCE ENZYMŮ RM SYSTÉMU TYPU I S CÍLOVOU SEKVENCÍ NA DNA



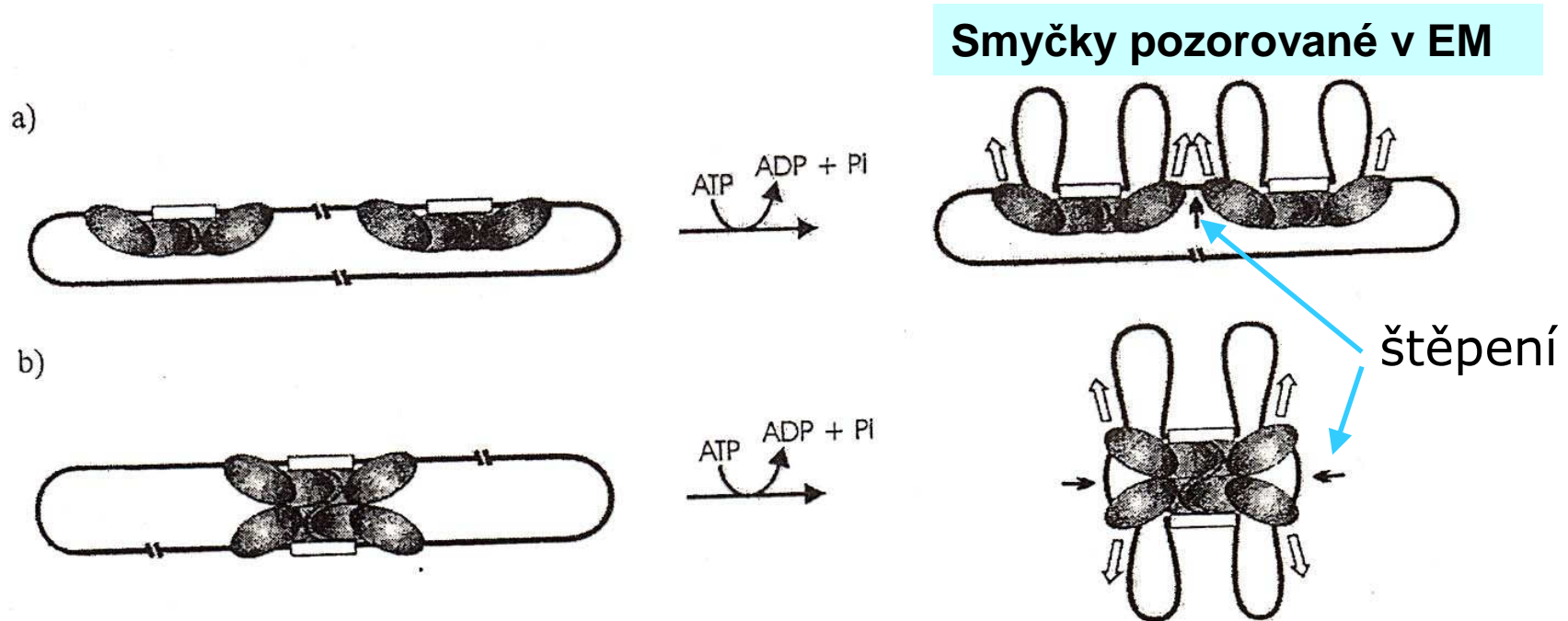
Does a type I enzyme move along DNA or does it remain at its target site, simultaneously pulling the DNA through the protein?



Interakce enzymového komplexu s rozpoznávacím místem

1. Vyhledání rozpoznávací sekvence
2. Podle stavu sekvence (M, NM, H-M) dochází k
 - a) uvolnění enzymu
 - b) restrikci
 - c) metylaci

MODELY TRANSLOKACE DNA ZPROSTŘEDKOVANÉ ENZYMY RM SYSTÉMŮ TYPU I

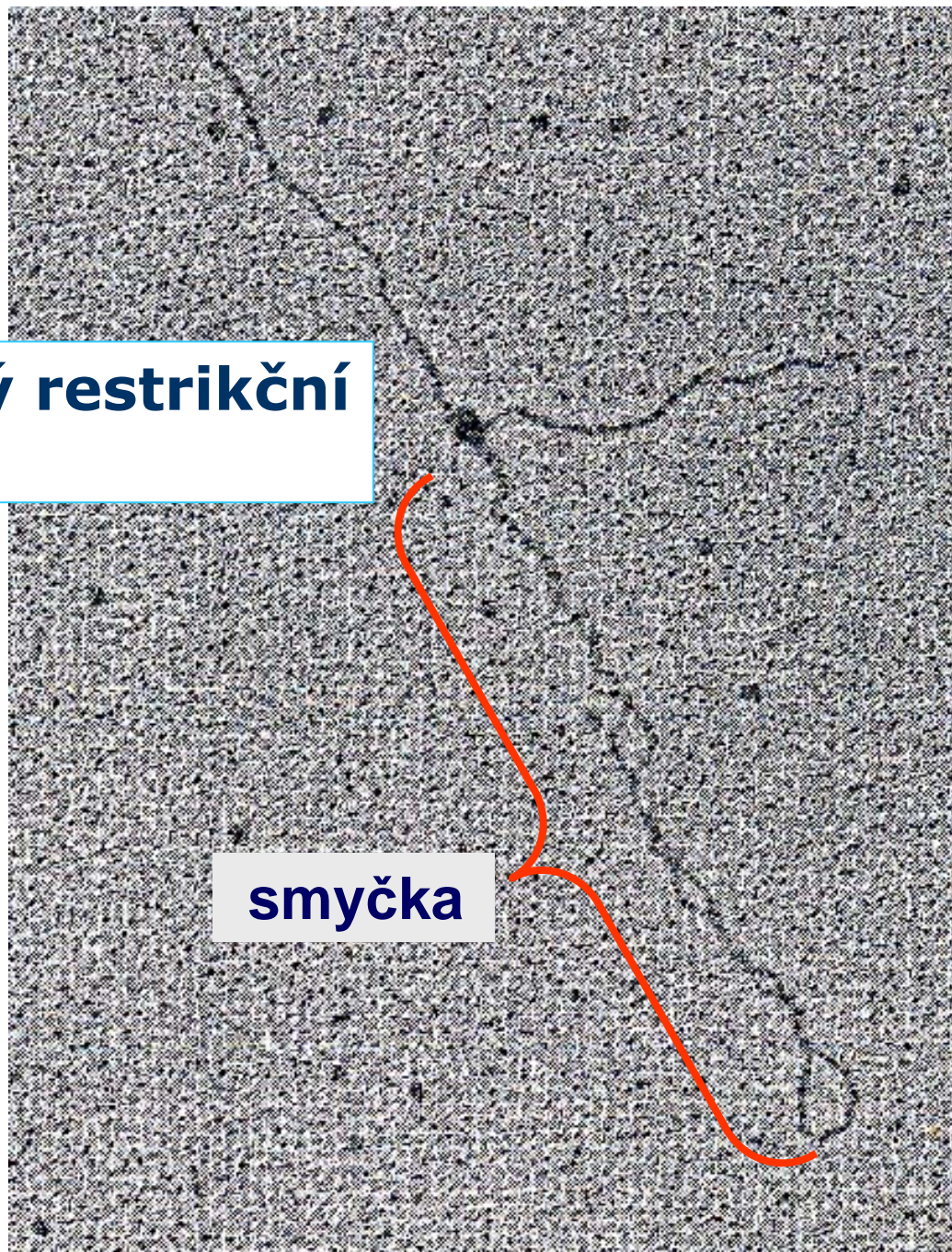


ATP-dependentní translokace DNA - ke štěpení dochází po kolizi DNA řetězců v místě navázaného enzymu

Pokud jsou místa nemodifikovaná, vytváří se rozpoznávací komplex, vázaný na rozpoznávací místo, a jím je DNA protahována (translokována) až k místu vzdálenému zhruba 1000 bp nebo více, kde pak proběhne štěpení.

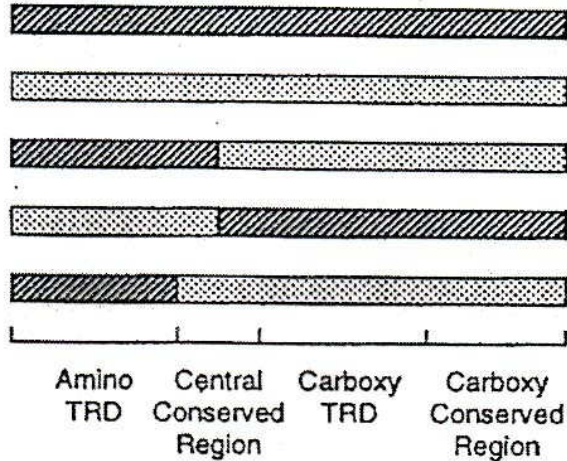
Navázaný restriční enzym

smyčka



Vytváření RM systémů typu I s novými sekvenčními specifitami

a) Struktura rozpoznávací podjednotky



Enzyme	Recognition sequence
<i>StySPI</i>	AAC (N ₆) GTRC
<i>StyLTIII</i>	GAG (N ₆) RTAYG
<i>StySQ</i>	AAC (N ₆) RTAYG
<i>StySJ</i>	GAG (N ₆) GTRC
<i>StySQ'</i>	AAC (N ₆) RTAYG

Přirozený RM systém

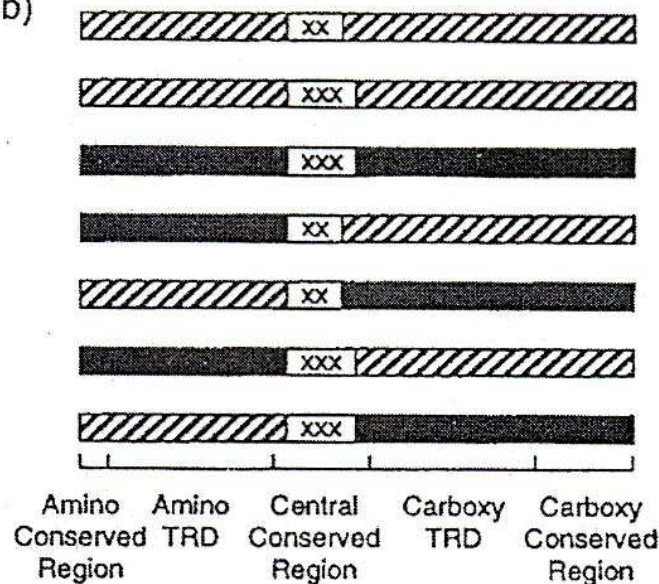
Přirozený RM systém

➤ hybridní hsdS

mutagenizovaná
centrální oblast

Geny hsdS mohou rekombinovat – představují dvě alely

b)



<i>EcoR124I</i>	GAA (N ₆) RTCG
<i>EcoR124II</i>	GAA (N ₇) RTCG
<i>EcoDXXI</i>	TCA (N ₇) RTTC
DR2	TCA (N ₆) RTCG
RD2	GAA (N ₆) RTTC
DR3	TCA (N ₇) RTCG
RD3	GAA (N ₇) RTTC

Změny počtu

aminokyselin

v centrální oblasti

ovlivňují rozpoznání

cílové sekvence

TRD = target recognition domain

MUTACE V GENECH RM SYSTÉMU

Genotype <i>hsdS hsdR hsdM</i>	Phenotype r m	Protein-DNA interaction	Activity R M
+ + +	+ +		+ +
- + +	- -		- -
- + -	- -		- -
+ - +	- +		- +
- - +	- -		- -
+ - -	- -		- -
+ + -	+ -		<u>lethal</u>

$r^{+/-}/m^{-}$

Pro RM systém typu I je nezbytná funkční *hsdS* - její mutace vede k fenotypu r- m-.

S⁻ → M⁻R⁻

RM SYSTÉMY, U NICHŽ JE ŠTĚPENA MODIFIKOVANÁ DNA (**KMEN NETVOŘÍ METYLÁZU**)

- *E. coli* K12
- Mar = methyladenine restriction (GmeAC, GmeAG)
- Mrr = methyladenine recognition and restriction
- Mcr = methylcytosine restriction - modified cytosine restriction (GmeCG - C5, N4, HM-C5) - štěpení sekvence v několika místech (štěpí též neglukozylovanou fágovou DNA)

- *Diplocooccus (Streptococcus) pneumoniae*:
RM systém DpnI - štěpí **GmACT**
- (*Caulobacter, Neisseria, Acholeplasma, Streptomyces*)

Detekce restričního systému Mar

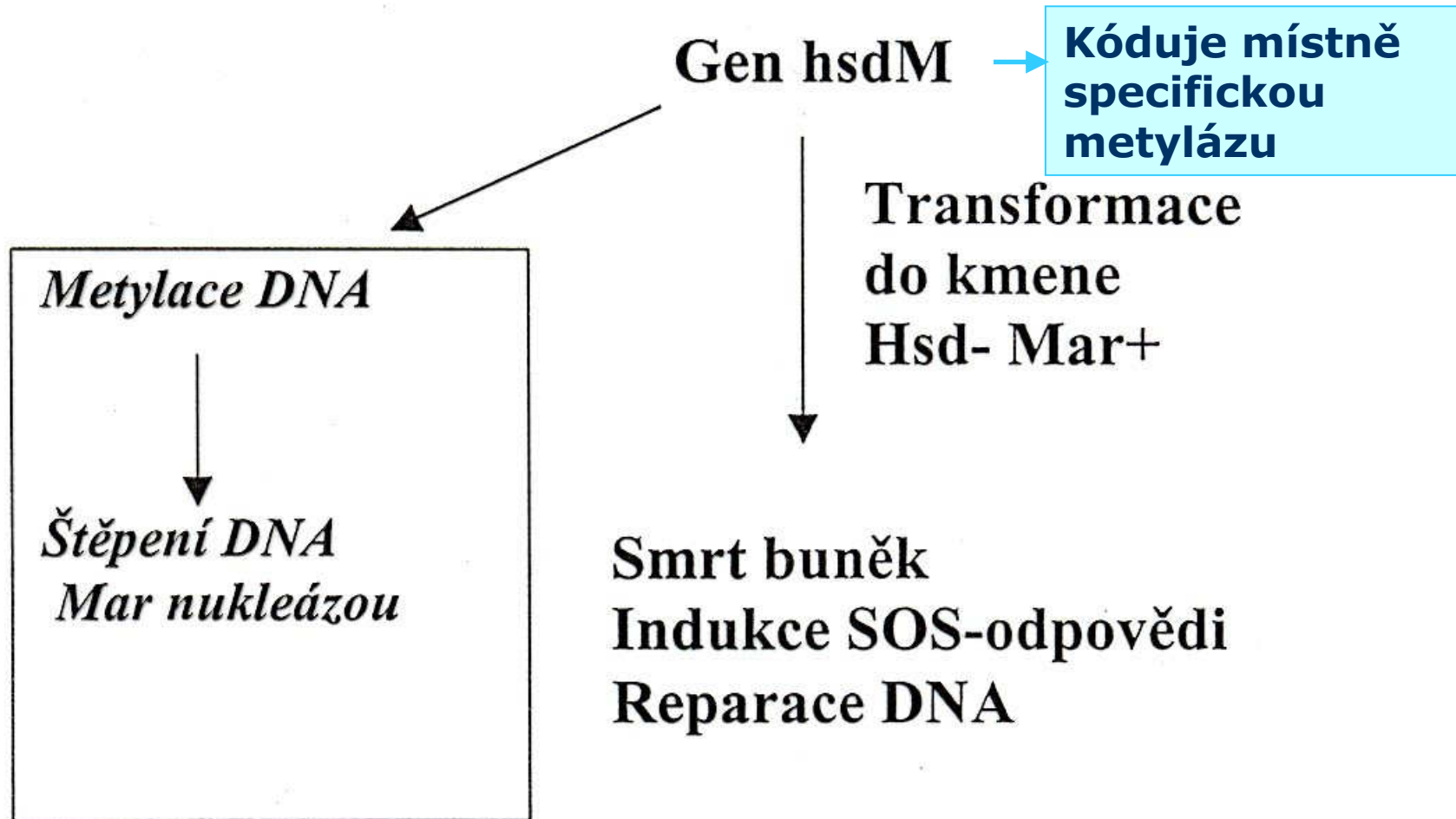


TABLE 3. Antirestriction Strategies

Strategy	Bacteriophage	Antirestriction gene, polypeptide, or base modification	Antirestriction target (R-M type)	Reference
Inhibition of endonuclease	T3, T7	<i>ocr</i>	<i>EcoK</i> , <i>EcoB</i> (type I) <i>EcoP1</i> (type III)	Krüger et al., 1978 Krüger et al., 1982
	φNR2rH, φ1rH	20 kD	<i>BamN_x</i> (type II)	Makino et al., 1979 Makino et al., 1980
Protection of viral DNA by viral-encoded methylase	SPβ, φ3T, SPR	<i>M.BsuRI</i>	M.SPR	Warren, 1980 Noyer-Weidner et al., 1981
	T2, T4	T2,4 <i>dam</i>	Mutant <i>EcoP1</i> (type III)	Bächi et al., 1979 Hattman et al., 1985
Overproduction of host-encoded methylase	λ	<i>ral</i>	<i>EcoK</i> , <i>EcoB</i> (type I)	Zabeau et al., 1980
Modified bases in viral genome	SP01, SP8, SP2G	5-Hydroxymethyluracil (for thymine)	Many type II systems	Hemphill and Whiteley, 1975
	φ25, φe, 2C PBS1, PBS2 T-even phages	Uracil (for thymine) Hydroxymethylcytosine (for cytosine)	<i>EcoB</i> (type I) <i>EcoRI</i> , <i>EcoRV</i> (type II)	Warren, 1980 Li et al., 1975 Revel and Georgopoulos, 1969 Takahashi et al., 1978
Metabolism of S-adenosyl-L-methionine (SAM)	P1	<i>darA</i> , <i>darB</i>	<i>EcoK</i> , <i>EcoB</i> (type I)	Krüger and Bickle, 1983
Underrepresentation of restriction sites in phage genome	φ1, φ29	No GGCC sites	<i>BsuRI</i> (type II)	Kawamura et al., 1981 Ito and Roberts, 1979

Mechanismy, kterými plazmidy a fágy unikají restrikci

1. Neobvyklé modifikace DNA

2. Nízká frekvence cílových míst

3. Tvorba proteinů, které interferují s jednou nebo více aktivitami RM systému

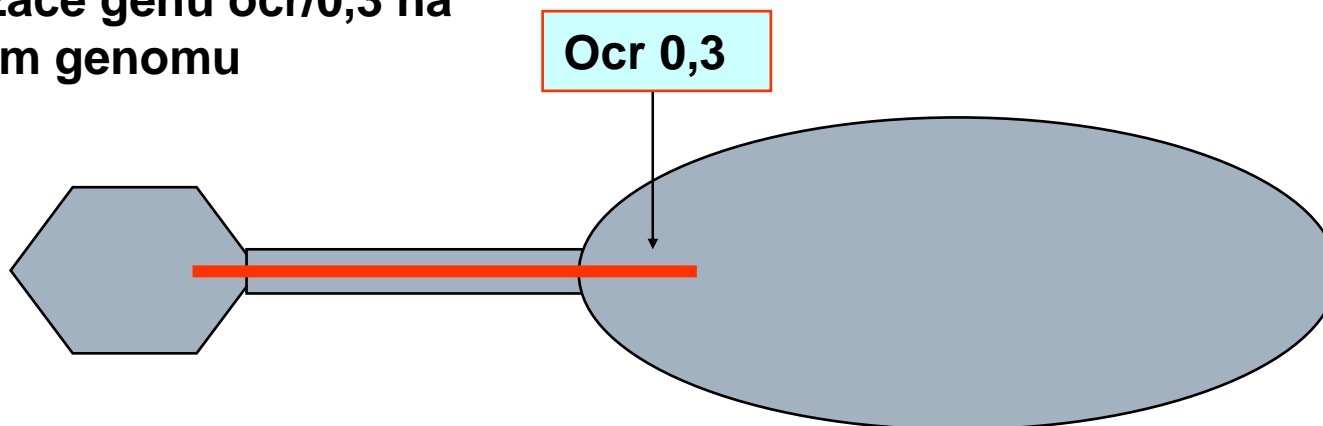
4. DNA vstupující do buňky ve formě ssDNA (konjugativní plazmidy nebo některé fágy), neuniká restrikci, ale stává se k nim citlivá po syntéze druhého vlákna.

Příklady antirestrikčních mechanismů

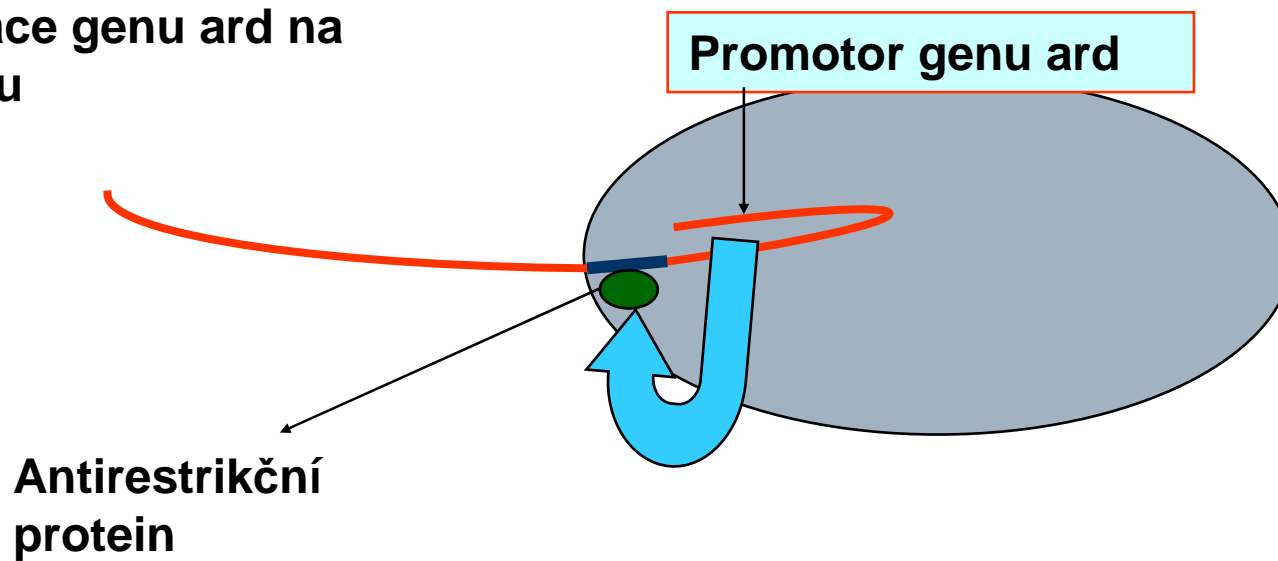
- **Fág MU – funkce MOM: konverze adeninu na A-(1-acetamido)adenin, který je necitlivý k EcoKI a EcoBI.**
- **Plazmidy: Ard (alleviation of restriction) - antirestrikční protein, zmírnění restrickce**
- **Fág lambda: Ral (restriction alleviation) - antirestrikční protein - zvýšení modifikační aktivity enzymů typu AI na NM DNA**
- **Fágy T3 a T7: (proteiny Ocr = overcoming restriction) inaktivace enzymů typu I (zábrana jejich vazby na DNA)**
- **Fág P1: Dar-proteiny (defense against restriction) ? - rozklad SAM**

1. **"Loss" of restriction sites.** Some phage have many fewer recognition sites for certain restriction endonucleases than you would predict by random chance. For example, the phages T3 and T7 lack the 5-bp *EcoRII* recognition site CC(A or T)GG. This is thought to arise by selection for phage with mutations that prevent cleavage by restriction endonucleases commonly encountered in their hosts
2. **Modified bases.** Some phage have modified bases that prevent recognition by restriction endonucleases. For example, phages T2 and T4 have hydroxymethylcytosine instead of cytosine in their DNA.
3. **Self-methylation.** Some phage encode a methylase that can modify their DNA so that it is not cleaved by certain restriction endonucleases. For example, the *E. coli* phages T2, T4, and the *Myxococcus* phage Mx8 encode dAdenine methylases (*dam*). Certain plasmids also encode methylases that may provide protection against restriction endonucleases
4. **Activation of a host methylase.** Some phage encode a protein that stimulates the host's methylase to modify the phage DNA. For example, the Lambda Ral ("restriction alleviation") protein enhances methylation by the HsdM subunit of the *EcoK* and *EcoB* restriction systems]. Ral seems to act by stimulating the expression of the host *hsdS* and *hsdM* genes
5. **Degradation of host S-adenosylmethionine.** Some host restriction systems only recognize modified DNA (e.g. the McrA and McrB systems in *E. coli*). Phage T3 encodes a S-adenosylmethionine hydrolase activity that degrades the substrate required for methylation by host enzymes, thus avoiding modification of the phage DNA and subsequent cleavage by modification-dependent host endonucleases
6. **Inhibition of host restriction endonucleases.** Many conjugal plasmids produce antirestriction proteins (called Ard) that specifically inhibit Type I restriction endonucleases. The Ard proteins have a motif that is very similar to a motif found in the HsdS subunit, thus it is possible that the Ard proteins prevent proper assembly of the restriction endonuclease complex by binding to the other Hsd subunits. Phage T3 produces a protein called Ocr which inhibits host methylases, resulting in resistance to modification-dependent restriction systems
7. **DNA repair systems.** Activity of certain phage genes may allow repair of the double-stranded breaks produced by restriction endonucleases. For example, the lambda Gam and Red proteins may repair the cleaved DNA by recombination with another copy of the phage DNA

Lokalizace genu ocr/0,3 na fágovém genomu

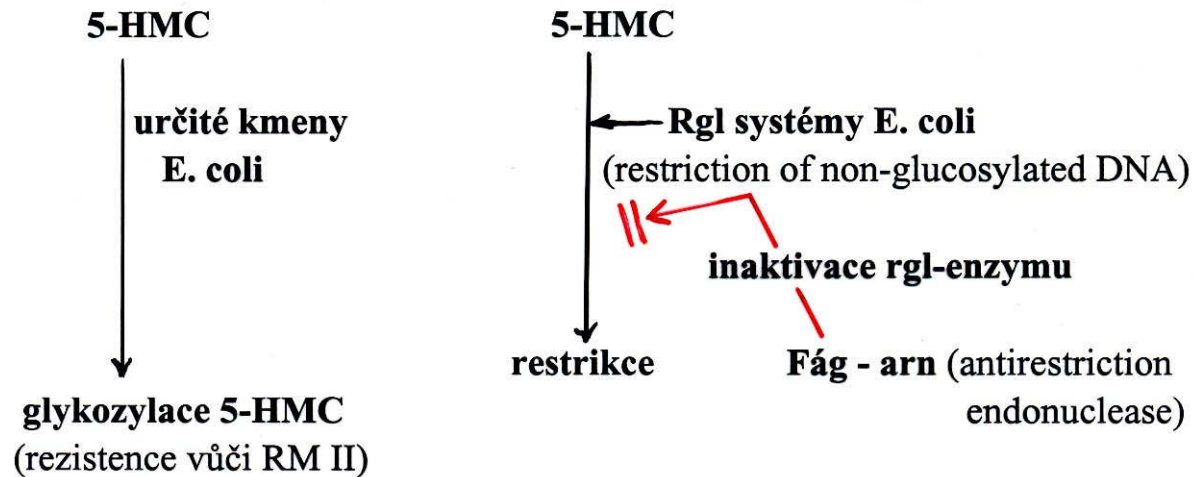


Lokalizace genu ard na plazmidu



Antirestrikční mechanismy T-sudých fágů

Fágová DNA

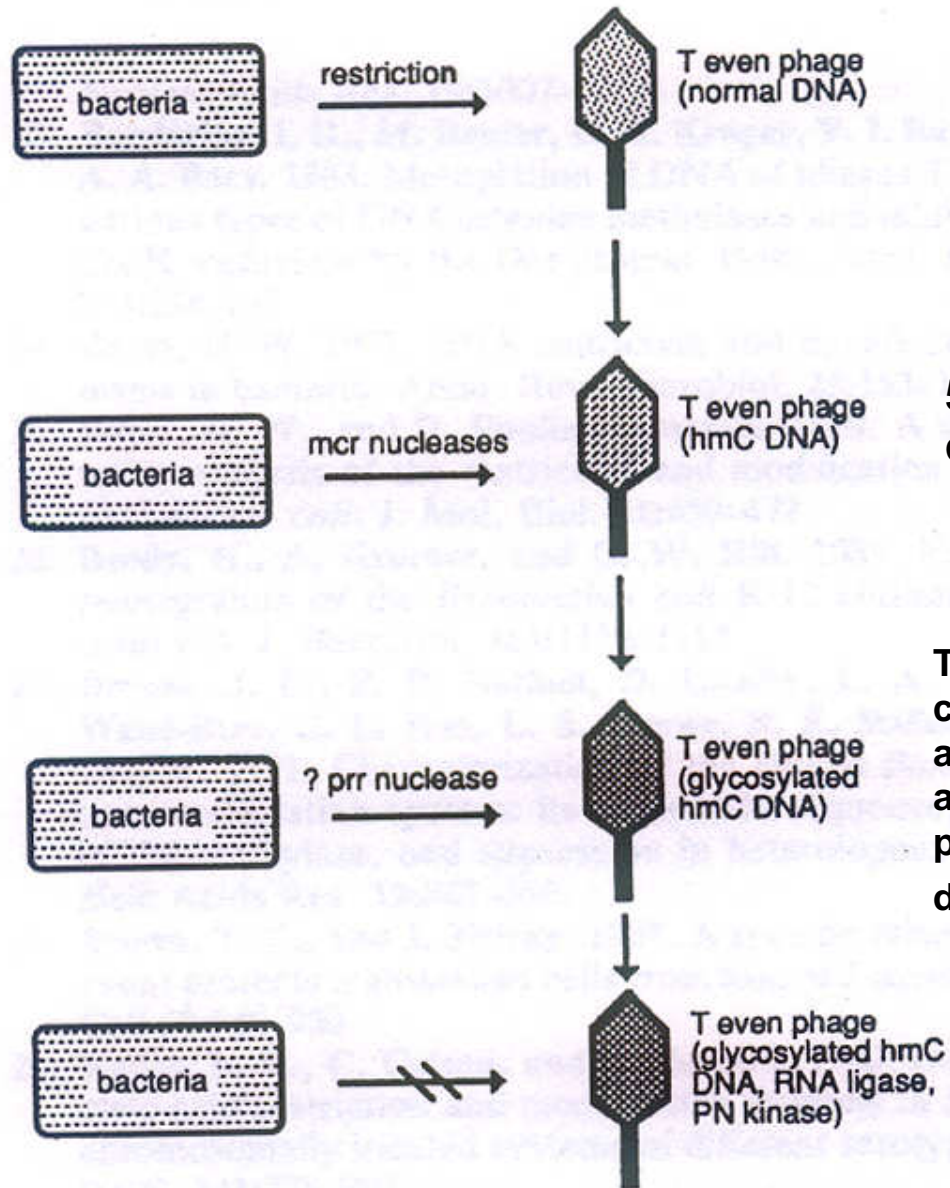


Reakce fága na restrikci

1. T4 ---> HMC
2. E. coli --> rgl
3. T4 ---> arn
4. T4 ---> glukozylace HMC

Glukozylace HMC zbytků DNA T-sudých fágů je účinnou ochranou před většinou RE – a navíc slouží k identifikaci fágové DNA, takže enzymy kódované fágů mohou selektivně degradovat hostitelský chromozom

Evoluce interakcí mezi T-sudými fágy a hostiteli

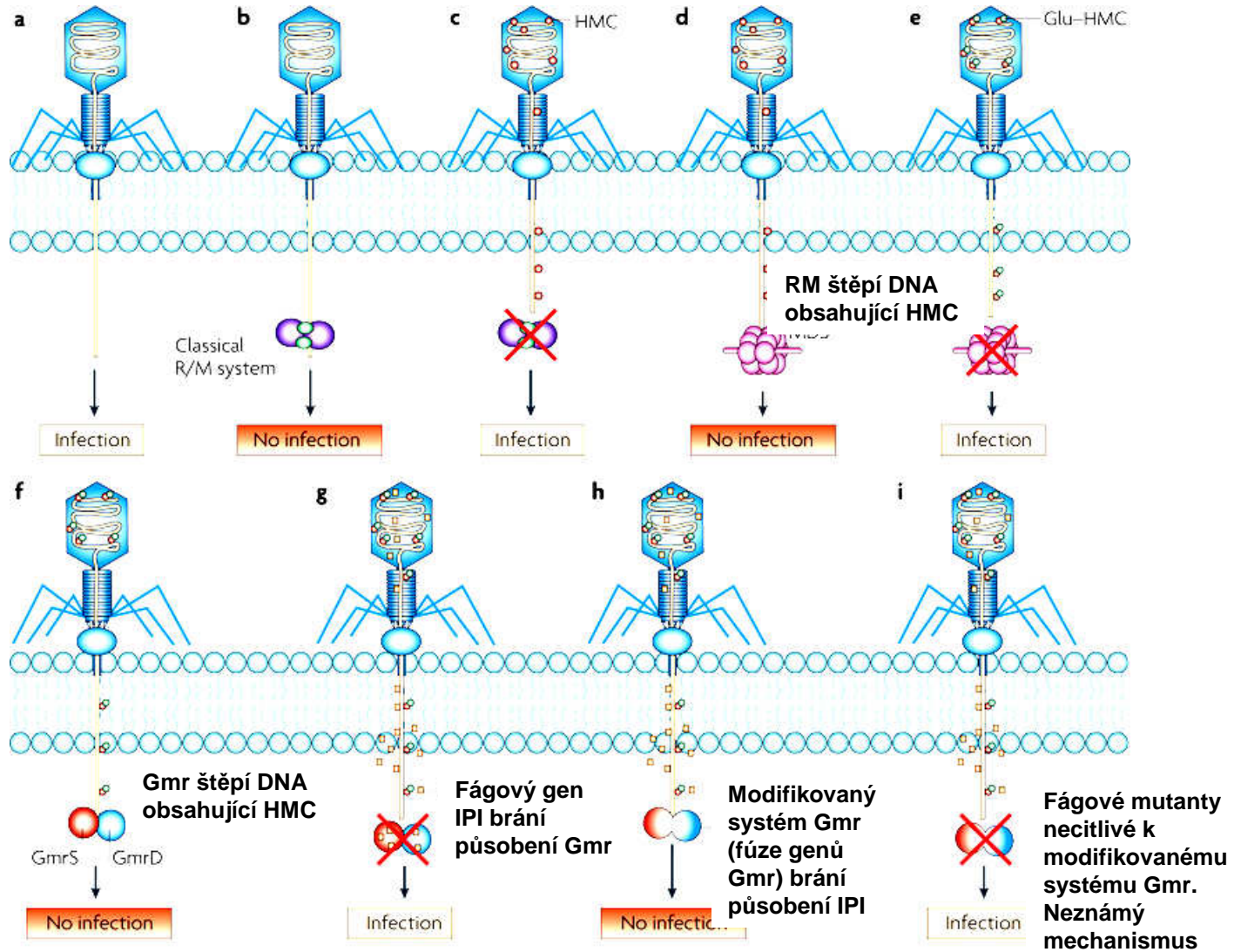


**5-hydroxymethyl-
cytosin**

The prr locus was originally described as coding a ribonuclease that is activated after phage T4 infection to cut within the anticodon of a specific tRNA, inactivating protein synthesis and thus blocking phage development.

glykozylace hmC

Evoluce interakcí mezi T-sudými fagy a hostiteli



SYSTÉMY METYLACE DNA (E. COLI K12)

- V buňkách není přítomna RE rozpoznávající metylovanou sekvenci – nejedná se tudíž o standardní RM systém

- **1. Systém Dam (dam-metyláza)**
 - metylace A v GATC (donor met = SAM)
 - GATC - regulační funkce: **počátek replikace, promotory, reparace, transpozice**
- **sekvence GATC je rozpoznávána 15% RE typu II, je přítomna u 50% všech 4N-RM-systémů, není však cílovou sekvencí pro restriktázy v žádném z druhů enterobaktérií.**

- **2. DNA cytozin metylační - Dcm (E. coli K12)**
 - metylace cytozinu na C5 v sekvenci CC(A/T)GG (?funkce při reparaci na velmi krátké vzdálenosti)

VÝSKYT RM SYSTÉMU U ENTEROBAKTERIÍ (3 DRUHY)

- 30% kmenů (z 1000 studovaných) obsahuje RM systém**
- u 170 RM systémů bylo zjištěno jen 33 různých cílových sekvencí**
- nebyly zjištěny RM systémy rozpoznávající 4 bp-místa (včetně GATC) – účast na regulacích**

VÝZNAM RESTRIKCE

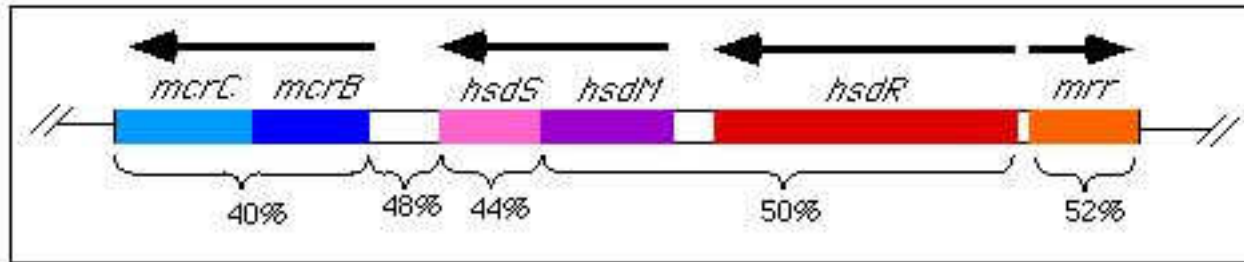
- **Ochrana integrity vlastní DNA**
 - obrana před bakteriofágy (typ II)
...dočasné působení –epigenetické změny, na úrovni populace nutná obměna
- **Systém napomáhající rekombinaci cizorodé DNA (typ I)**
 - štěpení DNA (náhodně) mimo rozpoznávací sekvenci umožní rekombinaci mnoha genů - analýza přenosu genů transdukcí a konjugací podporuje vliv RM systémů při vzniku mozaikových genomů (*E. coli*)
- **? „Selfish DNA“**
 - molekulární paraziti bakterií. RM systémy typu II se udržují prostřednictvím plazmidů, které je kódují (usmrcení bezplazmidových buněk)

RM systémy jako mechanismus postsegregačního zabíjení

Locus (plasmid)	Bacteria	Killer	Target	Anti-killer
<i>B. Restriction-modification systems</i>		restriktáza		metyláza
<i>paeR7I</i> (pMG7)	<i>Pseudomonas aeruginosa</i>	PaeR7I	5' CTCGAG in the chromosome	M.PaeR7I
<i>ecoRI</i> (RTF-1)	<i>Escherichia coli</i>	EcoRI	5' GAATTC in the chromosome	M.EcoRI
<i>ecoRII</i> (Fig. 6; RTF-2; N-3)	<i>Escherichia coli</i>	EcoRII	5' CCWGG in the chromosome	M.EcoRII
<i>ecoRV</i> (pLG13)	<i>Escherichia coli</i>	EcoRV	5' GATATC in the chromosome	M.EcoRV
<i>ssoII</i> (Fig. 6; P4)	<i>Shigella sonnei</i>	SsoII	5' CCNGG in the chromosome	M.SsoII
<i>bsp6I</i> (pXH13)	<i>Bacillus</i> sp. strain RFL6	Bsp6I	5' GCNGC in the chromosome	M.Bsp6I

Odlišný obsah GC – indicie přenosu HGT

"immigration control region"



<u>R-M genes</u>	<u>Recognition sequence</u>
<i>hsdS_K hsdM_K hsdR_K</i>	AACNNNNNNGTGC
<i>hsdS_B hsdM_K hsdR_B</i>	TGANNNNNNNNGTGC
<i>hsdS_K hsdM_B hsdR_B</i>	AACNNNNNNGTGC

Gen *hsdS* může být nahrazen genem *hsdS* z jiného RM systému stejné třídy, nebo geny mohou vzájemně rekombinovat

VYUŽITÍ RM SYSTÉMŮ V EUKARYOTICKÝCH BUŇKÁCH

Transgenní linie myších buněk exprimující geny RM systémů

1. přenos a exprese genu M.PaeR7 (*Pseudomonas aeruginosa*) - **metyláza** CTCGmeAG
2. přenos a exprese genu R.PaeR7 - endonukleáza (**restriktáza**)
3. infekce buněk HSV1 adenoviry - očekávaná rezistence buněk - vytvoření organismu rezistentního k virům (nebo jen určitých tkání)

Další možnost: studium vlivu metylace na genovou expresi

Evolve bakteriálních genomů

Charakteristické rysy:

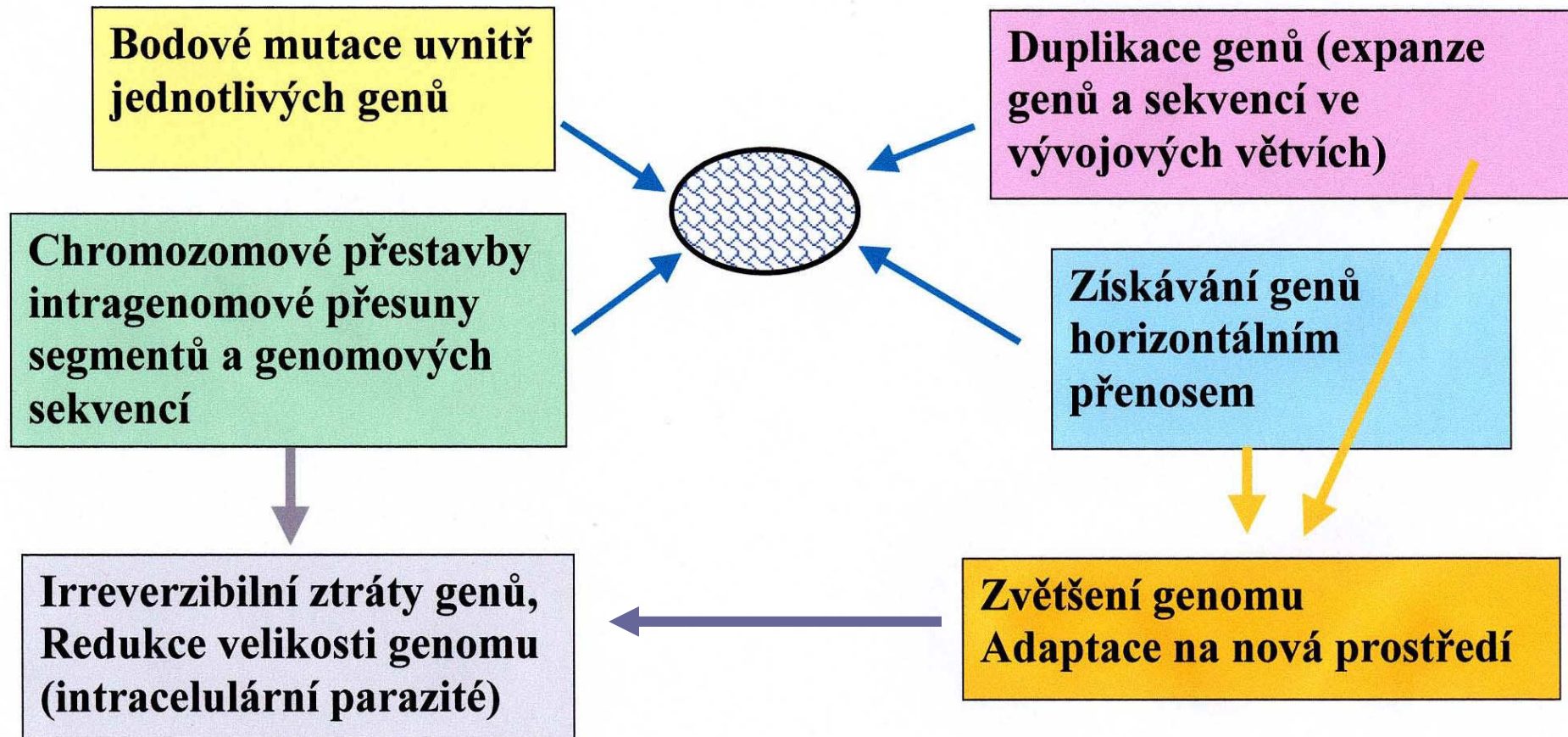
Rychlé a rozsáhlé změny ve struktuře a informačním obsahu genomu (plasticita, dynamické změny)

- Vnitřní přestavby
- Získávání a ztráty genů a genetických elementů

Vývoj kmenů v rámci druhu

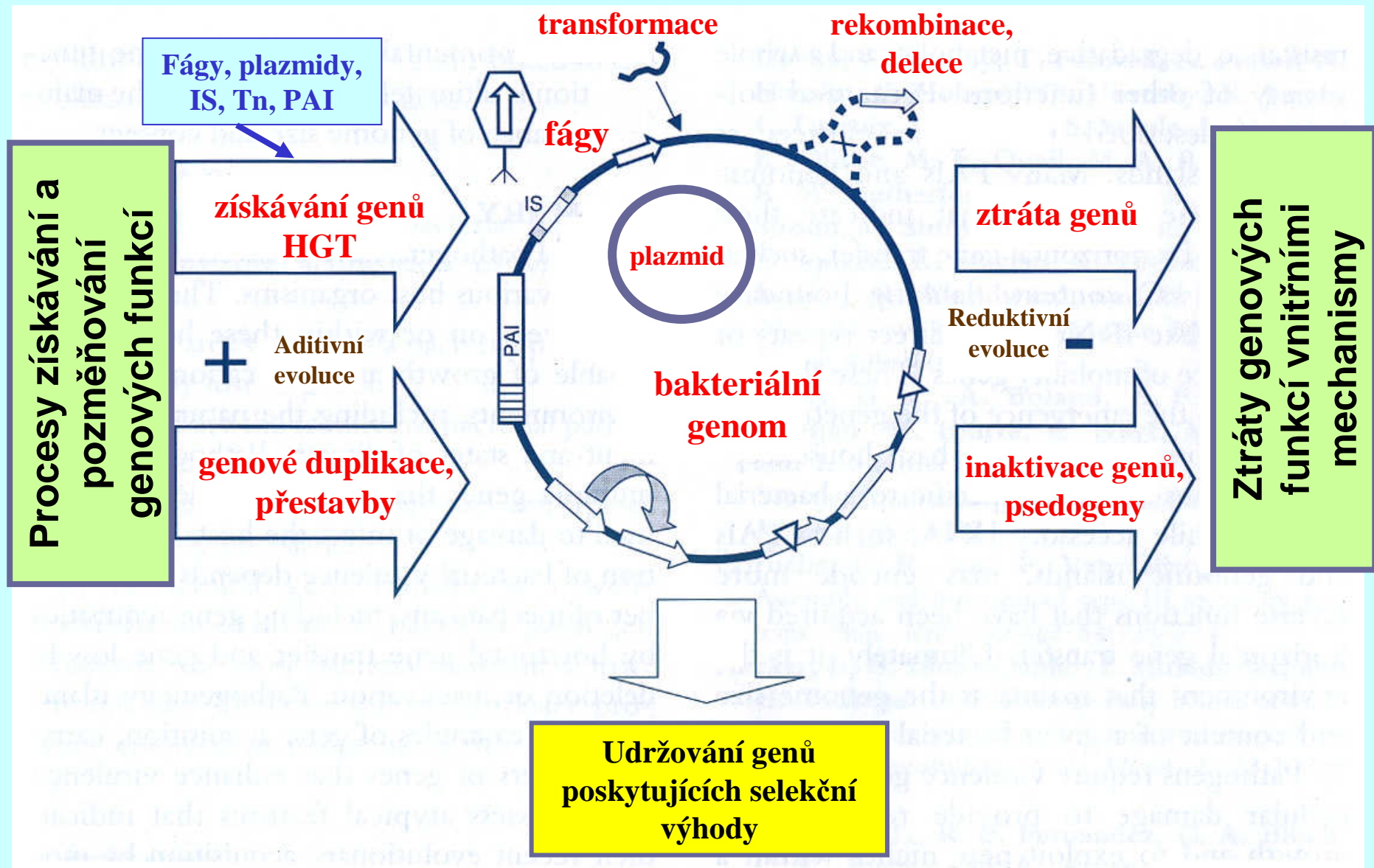
Adaptace na nové podmínky

Podstata změn v obsahu genomu prokaryot v průběhu evoluce



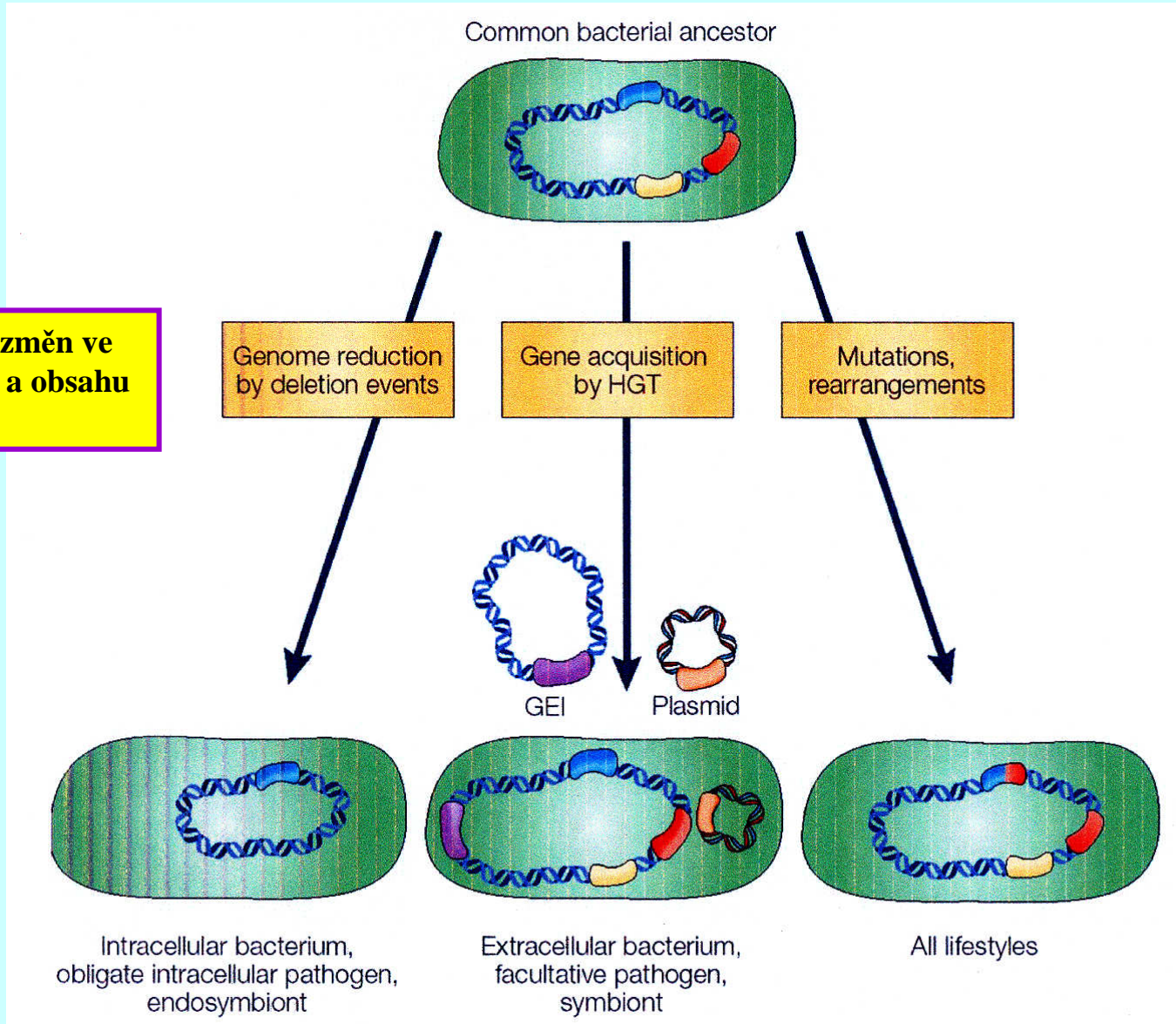
Důsledek: Pořadí genů je zachováno jen u velmi blízké příbuzných druhů

Mechanismy evoluce bakteriálních genomů



Struktura genomu odráží životní styl bakterie

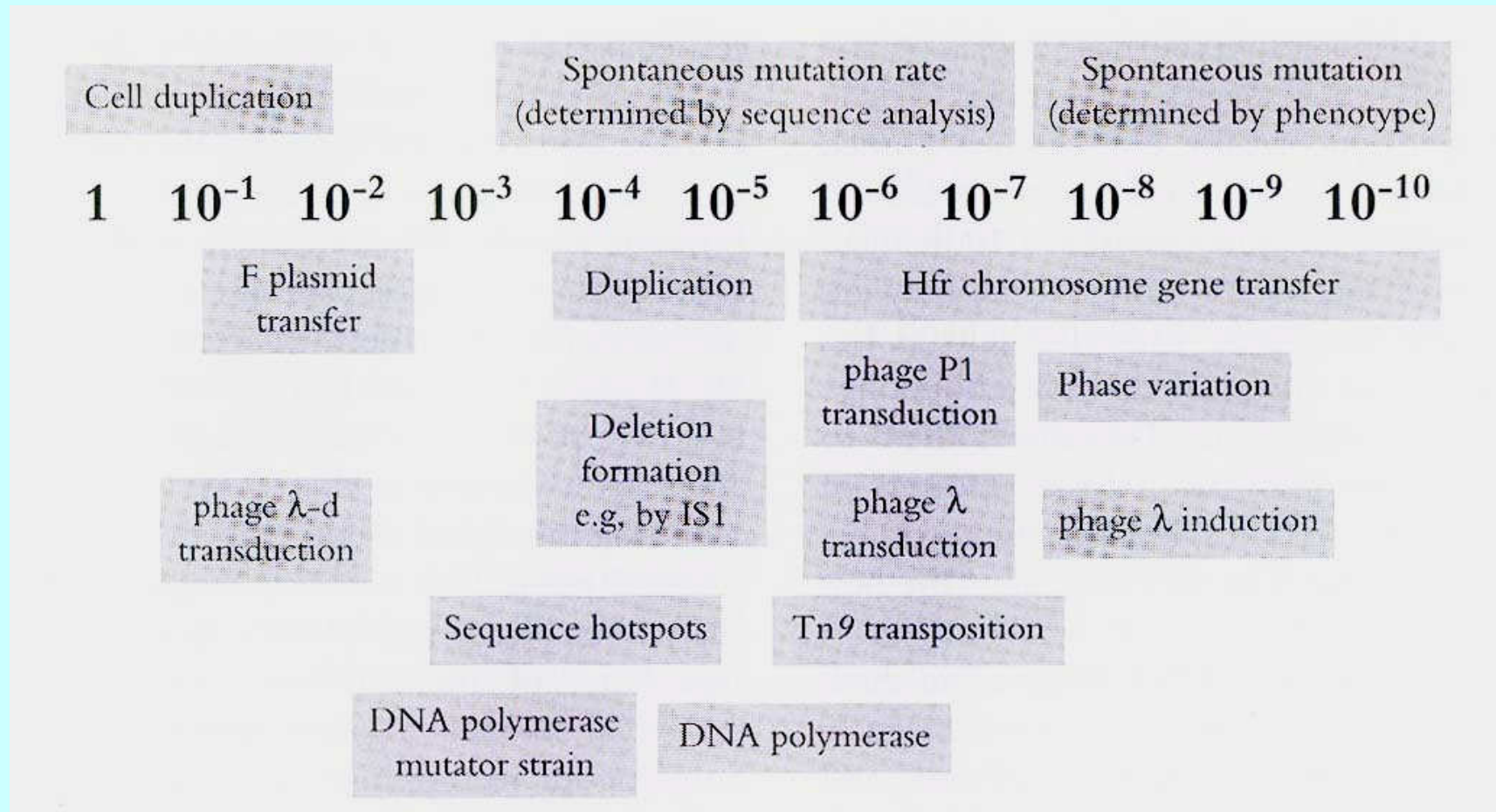
Příčiny změn ve velikosti a obsahu genomu



Mechanismy odpovědné za plasticitu genomu

Genetický element nebo mechanismus	Důsledky
A. 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

Spektrum faktorů podílejících se na změnách genomů



Vnitřní přestavby replikonů navozené přítomností repeticí

Typy repeticí

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

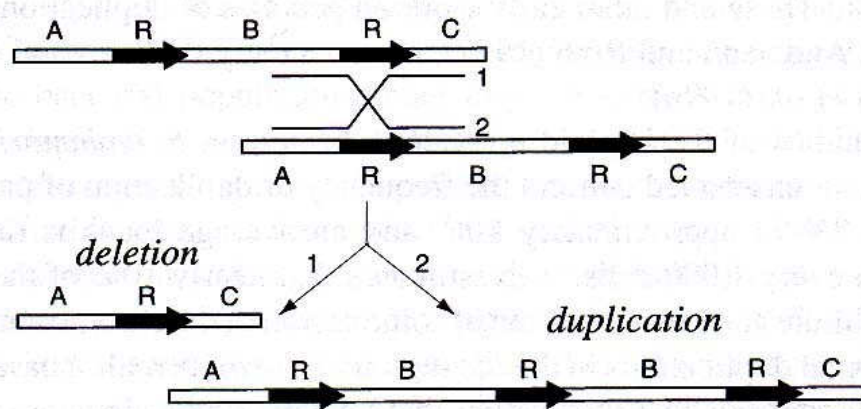
Typ přestavby

- Duplikace (amplifikace)
- Delece
- Inverze

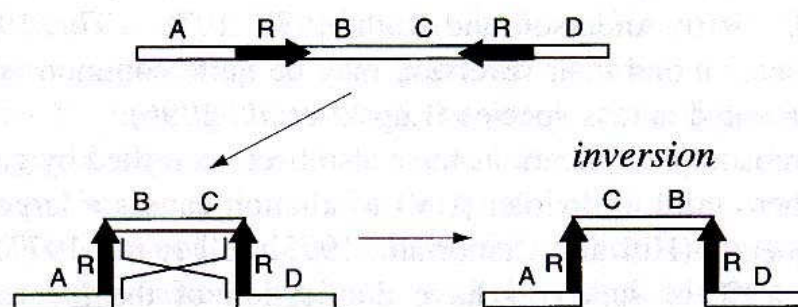
Mechanismus

Homologní rekombinace
Transpozice
Místně-specifická rekombinace
Nerovnoměrný crossing-over

přímá opakování (repetic)



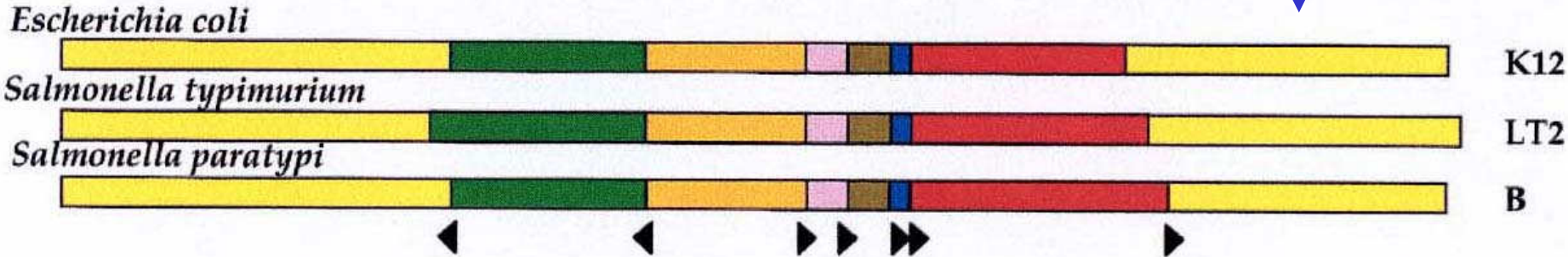
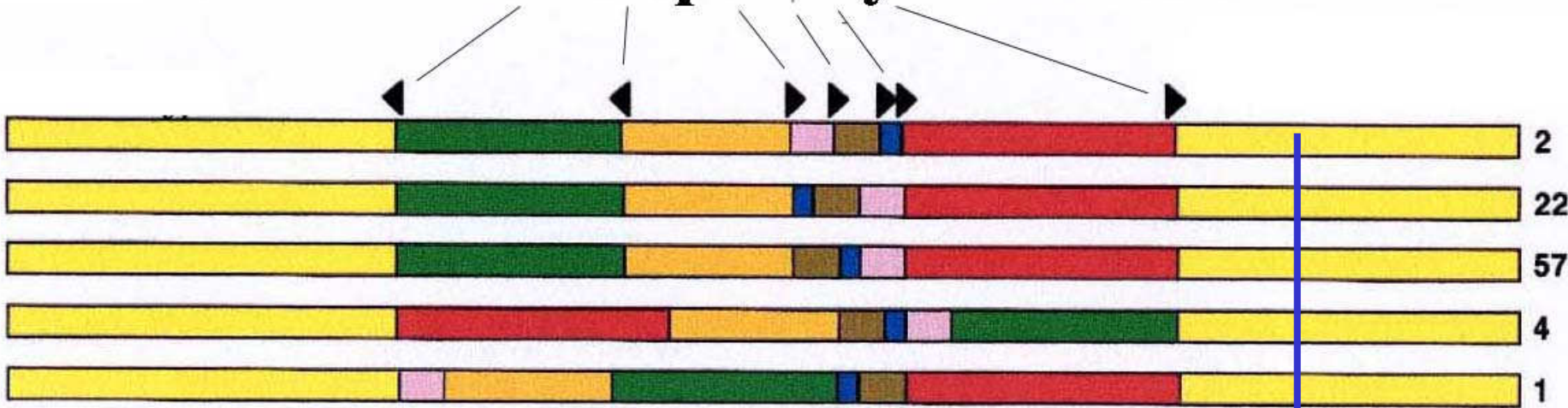
obrácená opakování (repetic)



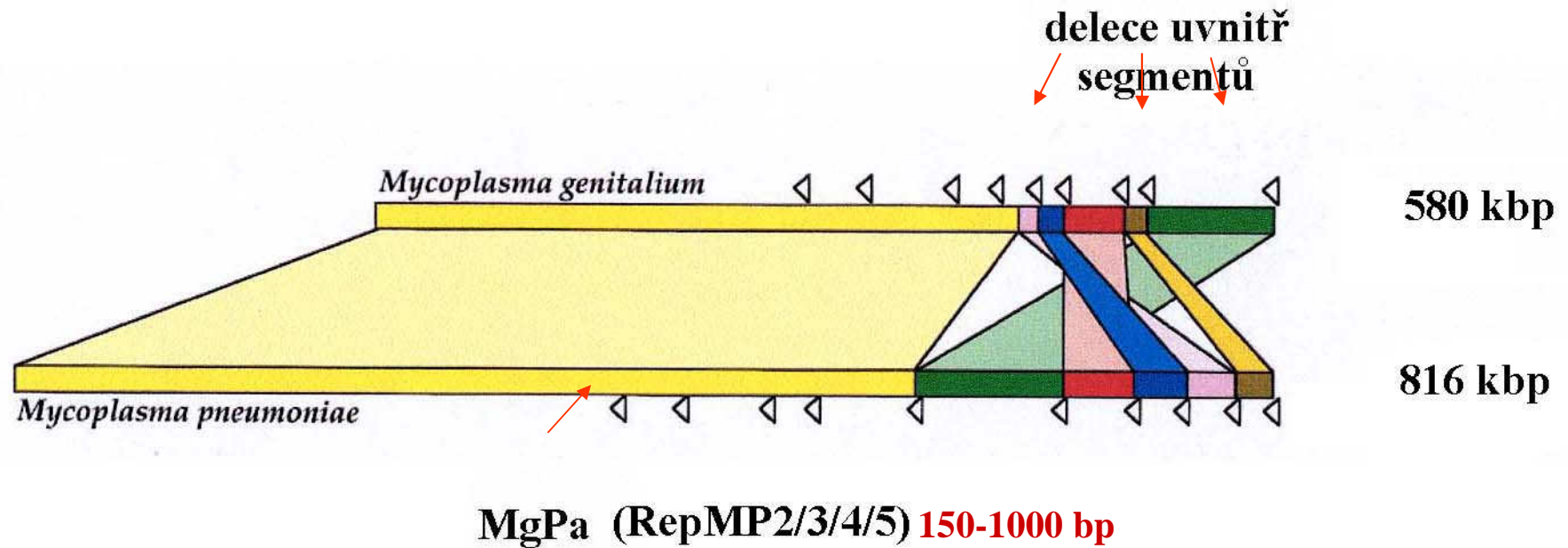
**Přestavby navozené interakcí repetic
(homologní rekombinace)**

Genomová přeskupení navozená rrn operony u kmenů Salmonella typhi

rrn operony



Rozdíly ve struktuře genomů příbuzných druhů

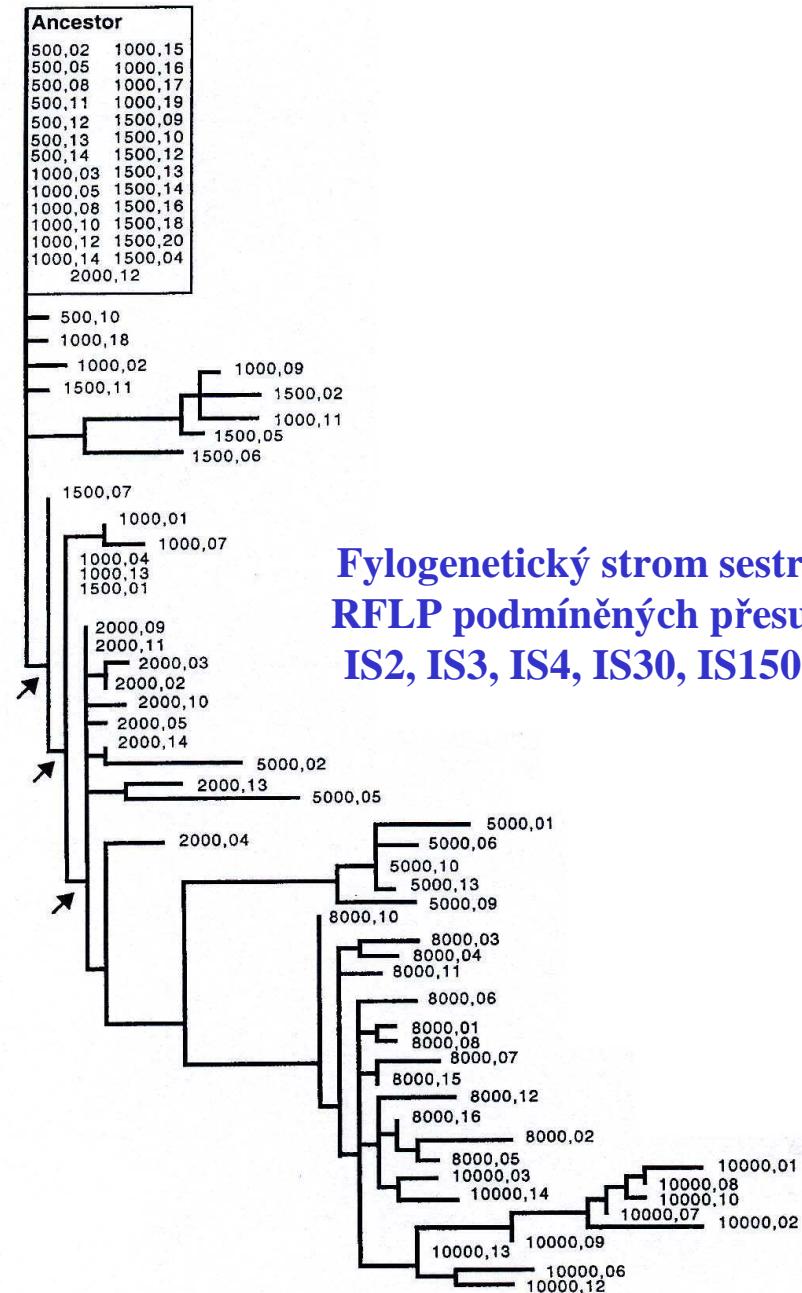


**Změny v genomu po
dlouhodobém uchování
kultur *E. coli* a *S. typhimurium***



- změny lokalizace IS
- změny velikosti genomu
- změny fenotypu

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



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)

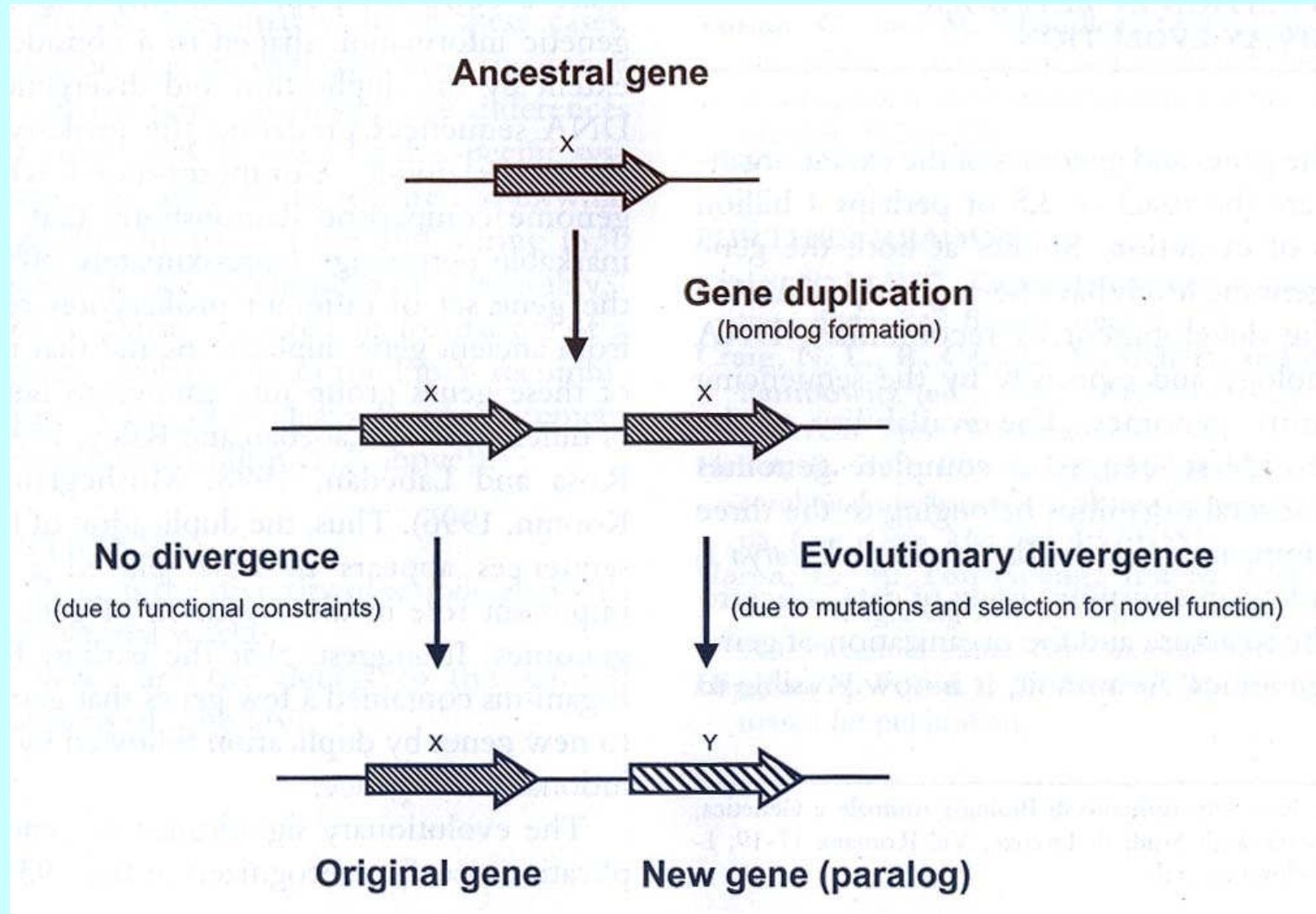
- Průměrný lokus je duplikován v každé z 1000 buněk
 - 10% buněk v kultuře nese duplikaci některé oblasti chromozomu
 - Velikost duplikací: 140 kb – 2100 kb
- Distribuce duplikací není náhodná
- Duplikace jsou ohraničeny dlouhými přímými repeticemi různého typu

Duplikace funkcí → adaptace na změny prostředí

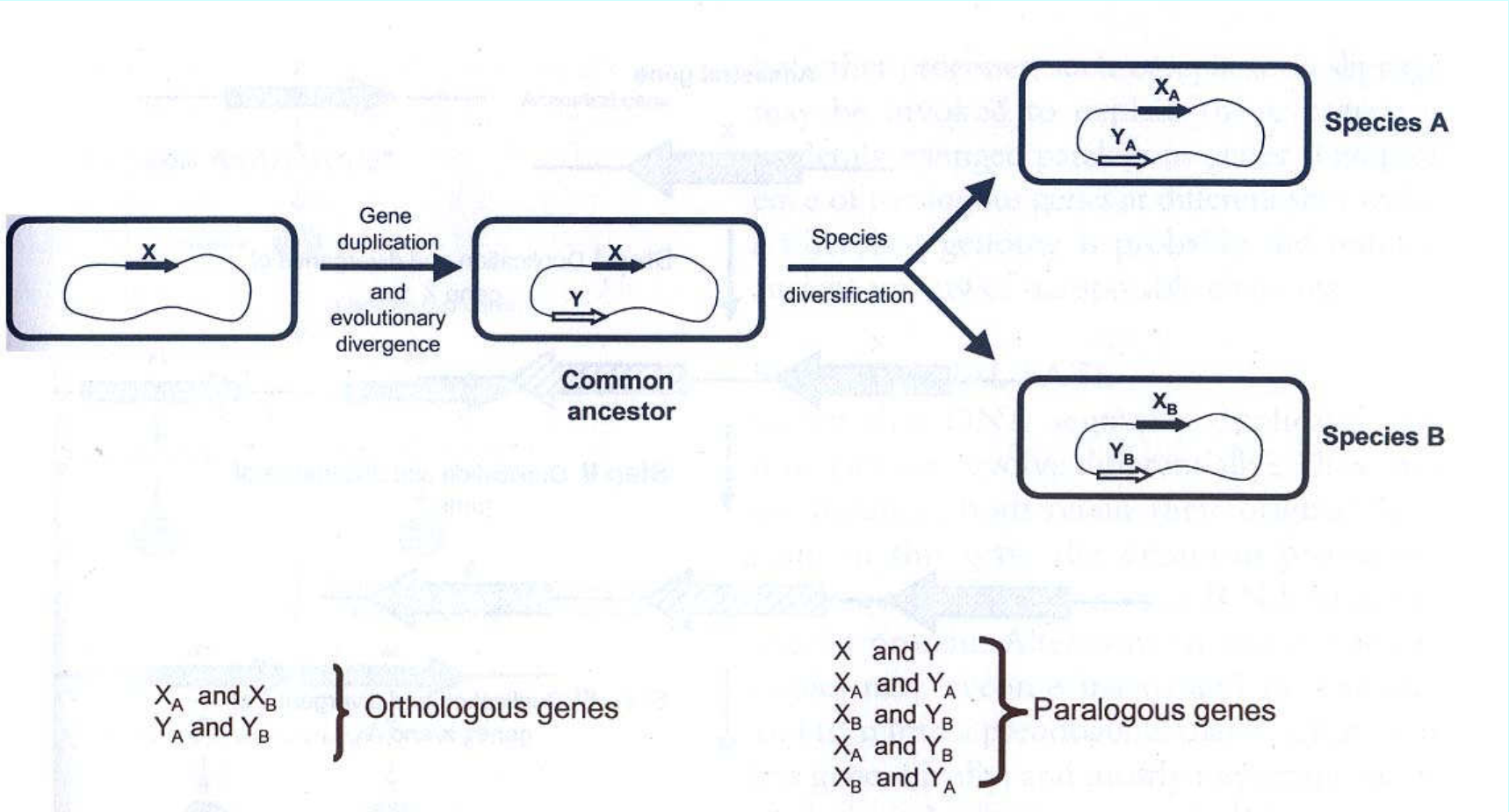
- zvýšení dávky genů
- vytvoření redundantní DNA pro následnou genetickou divergenci

→ paralogní geny – adaptace na nová prostředí

Vytváření paralogních genů duplikací a divergencí



Evoluční vztahy mezi ortologními a paralogními geny

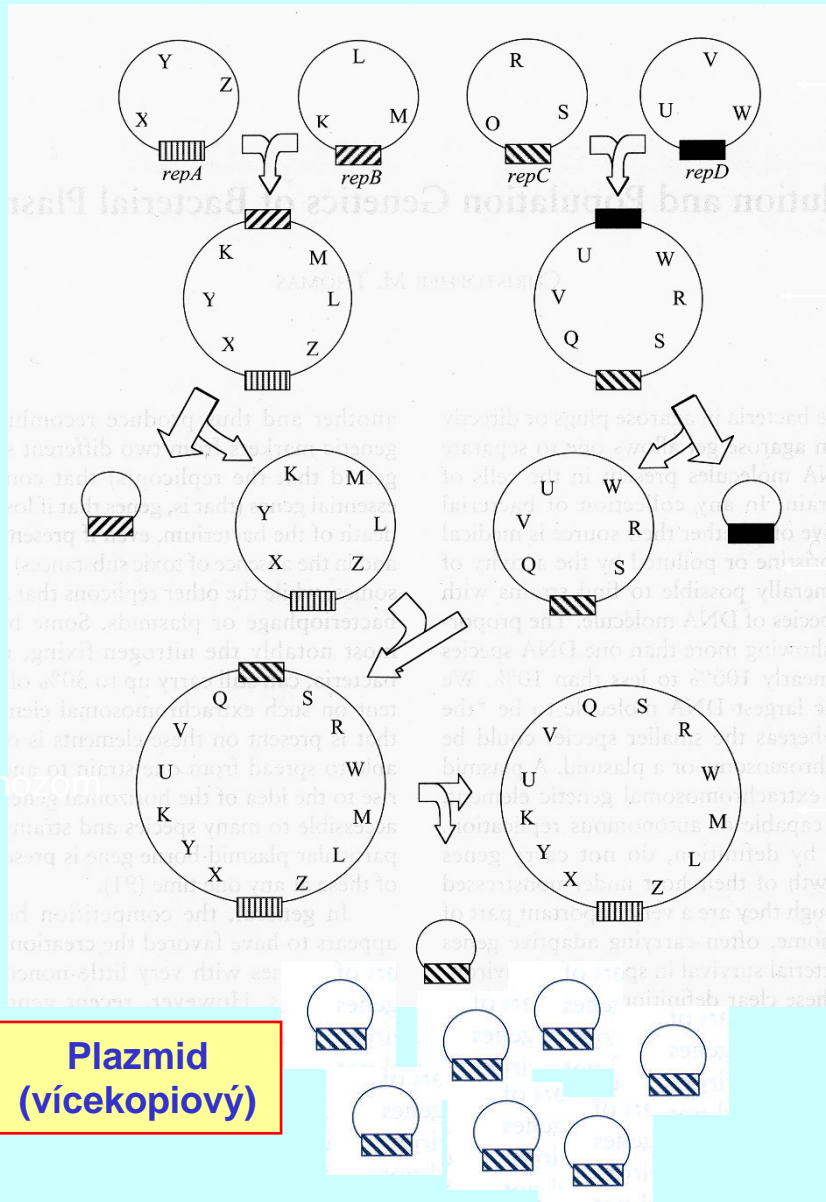


Počty paralogních genů v genomech bakteriálních druhů

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

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 sp.</i>	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

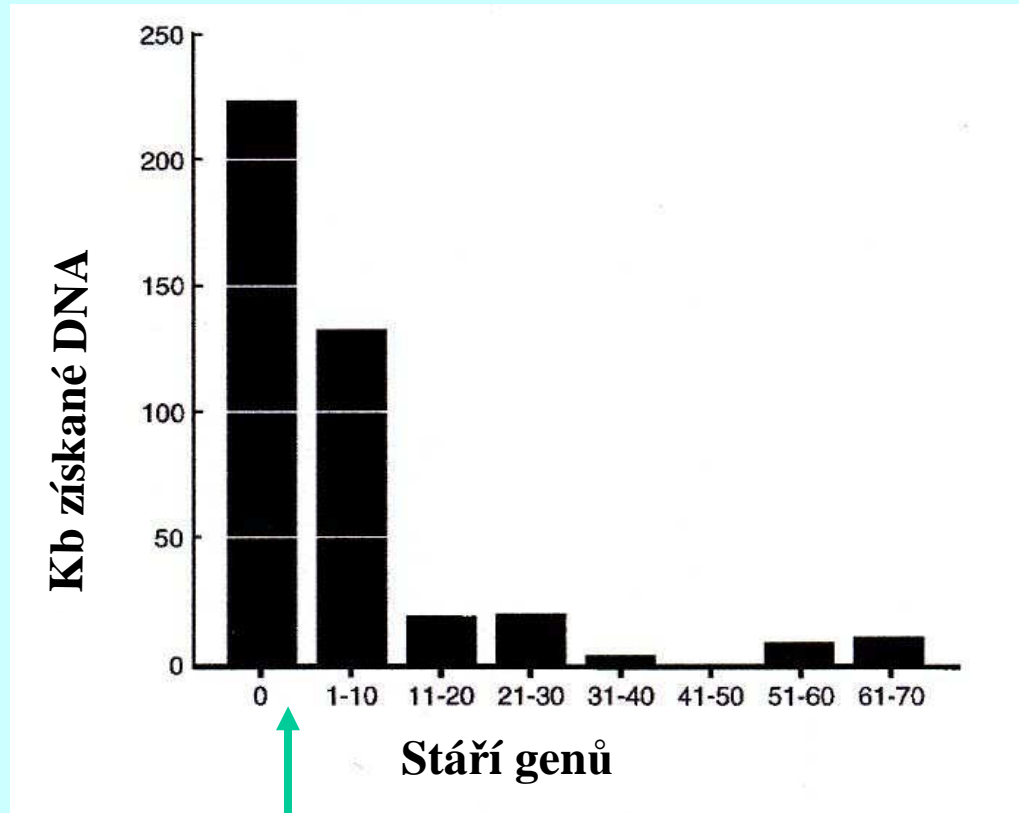
Horizontálně přenesené geny (HGT) u *E. coli* K12 MG1655 (po divergenci *E. coli* a *S. typhimurium*)

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

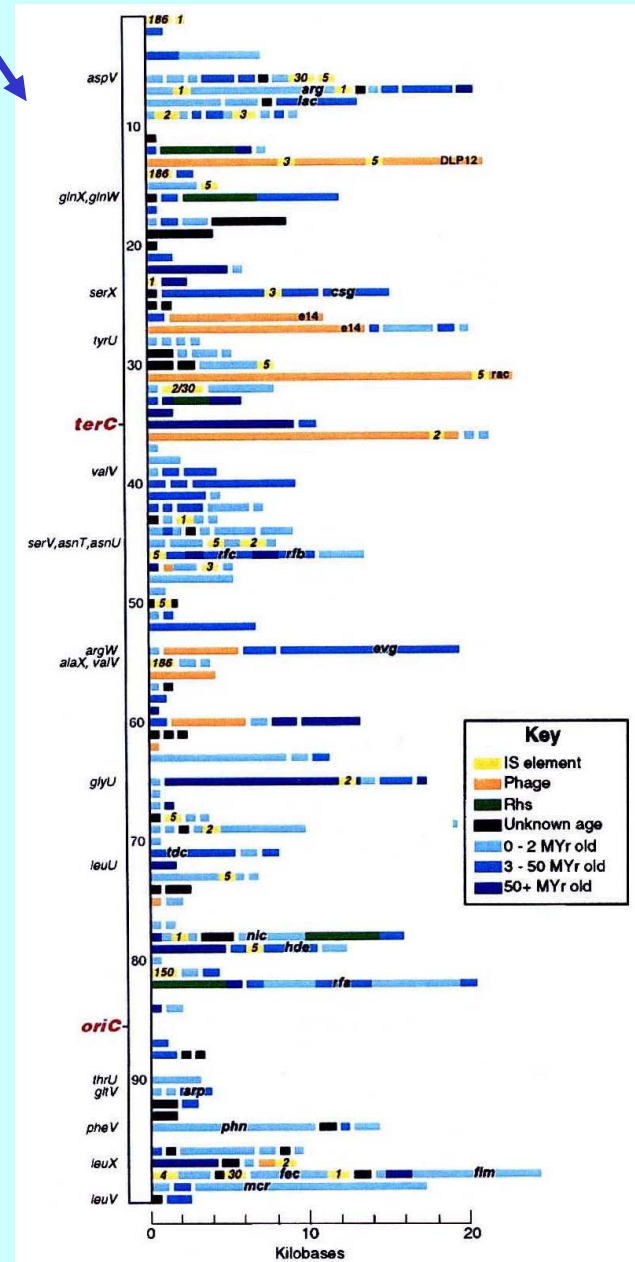
- **Vyšší proporce HTG v oblasti terminátoru replikace**
- **Lokalizace HTG poblíž genů pro tRNA (přenos pomocí fágů)**
- **V blízkosti HGT se nachází 68% všech inzerčních sekvencí**
 - **IS jsou přenášeny spolu HTG**
 - **IS navozují integraci přenášené DNA**

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 (SaPI1 aj) Exotoxinový PAI Enterotoxinový PAI	Toxin toxického šoku Exotoxin Enterotoxin
Ostrůvky patogenity		
<i>E. coli, 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 nekrotizující 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



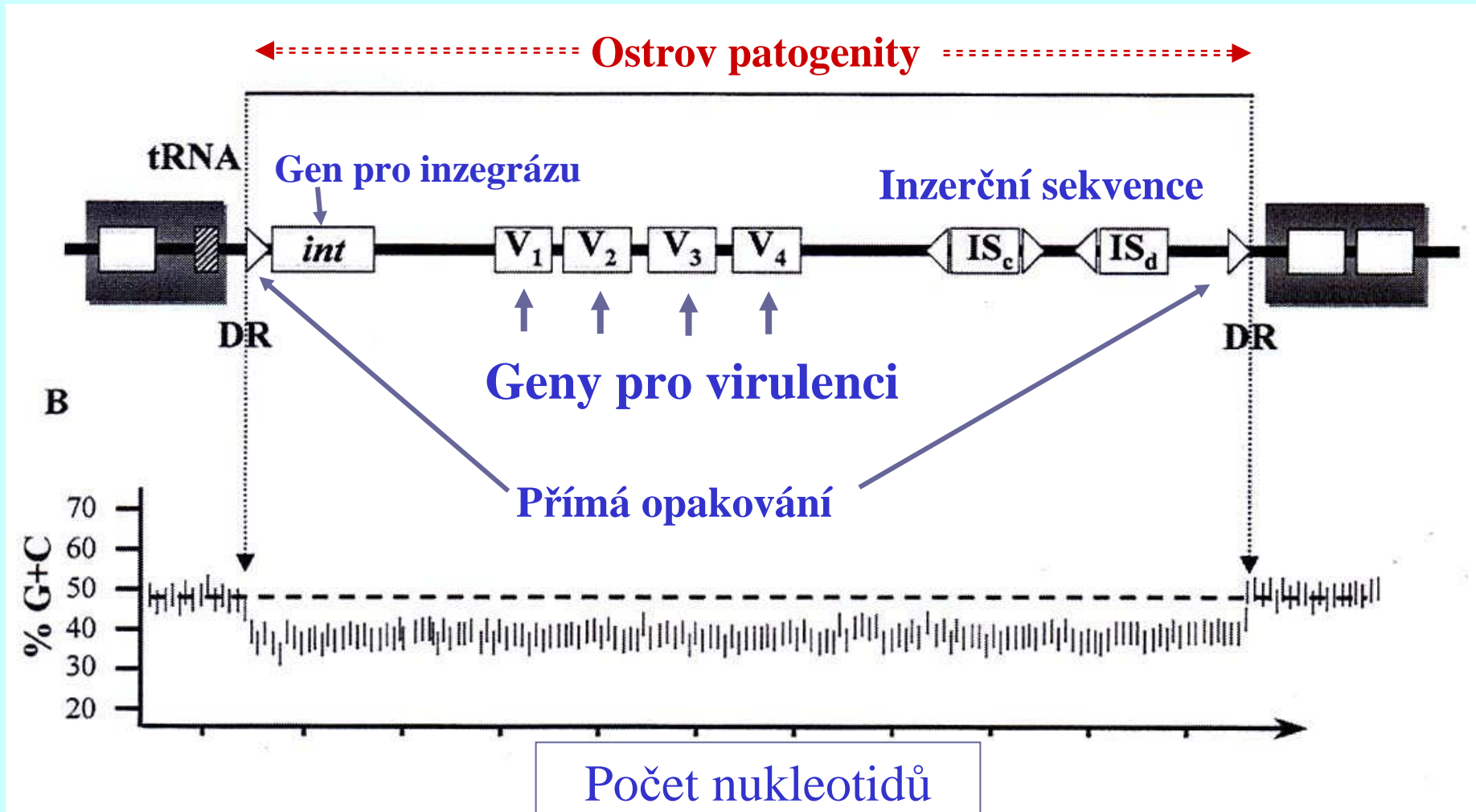
Genomické ostrovy („fitness“ ostrovy)

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

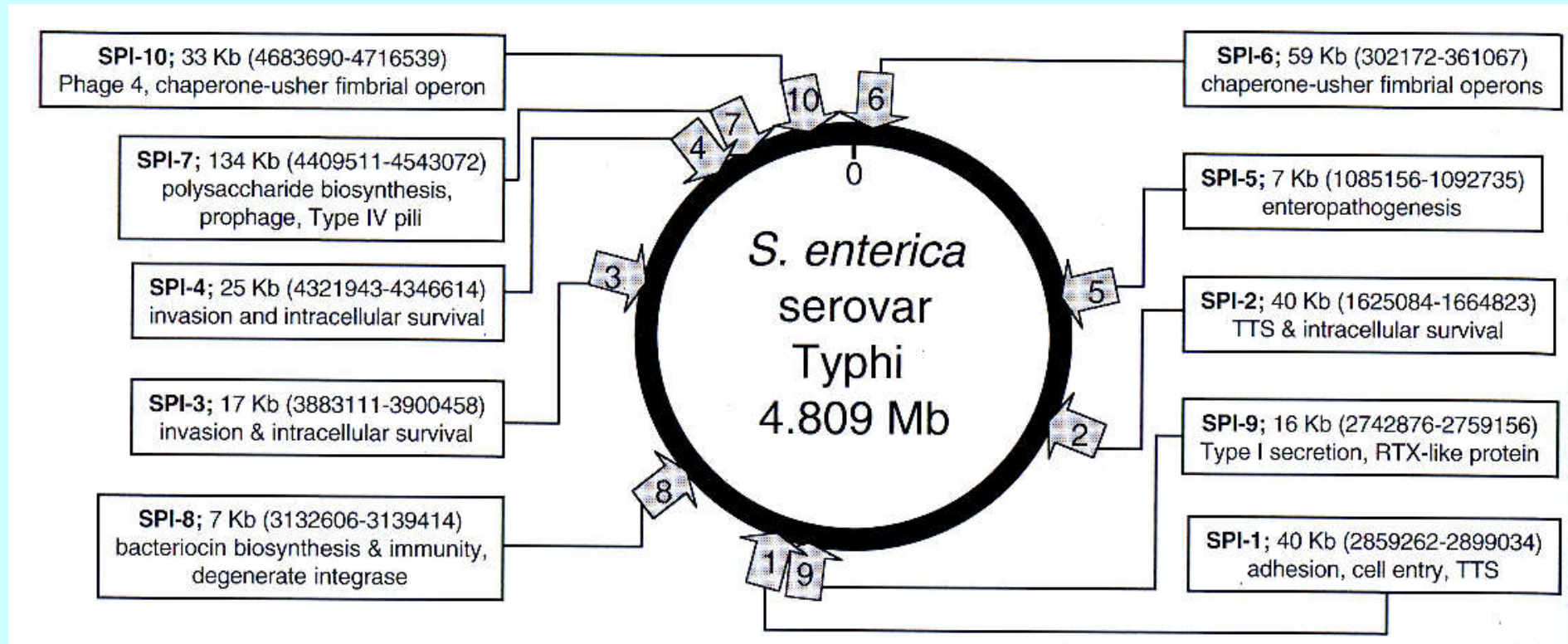
Obecná struktura ostrovů patogenity



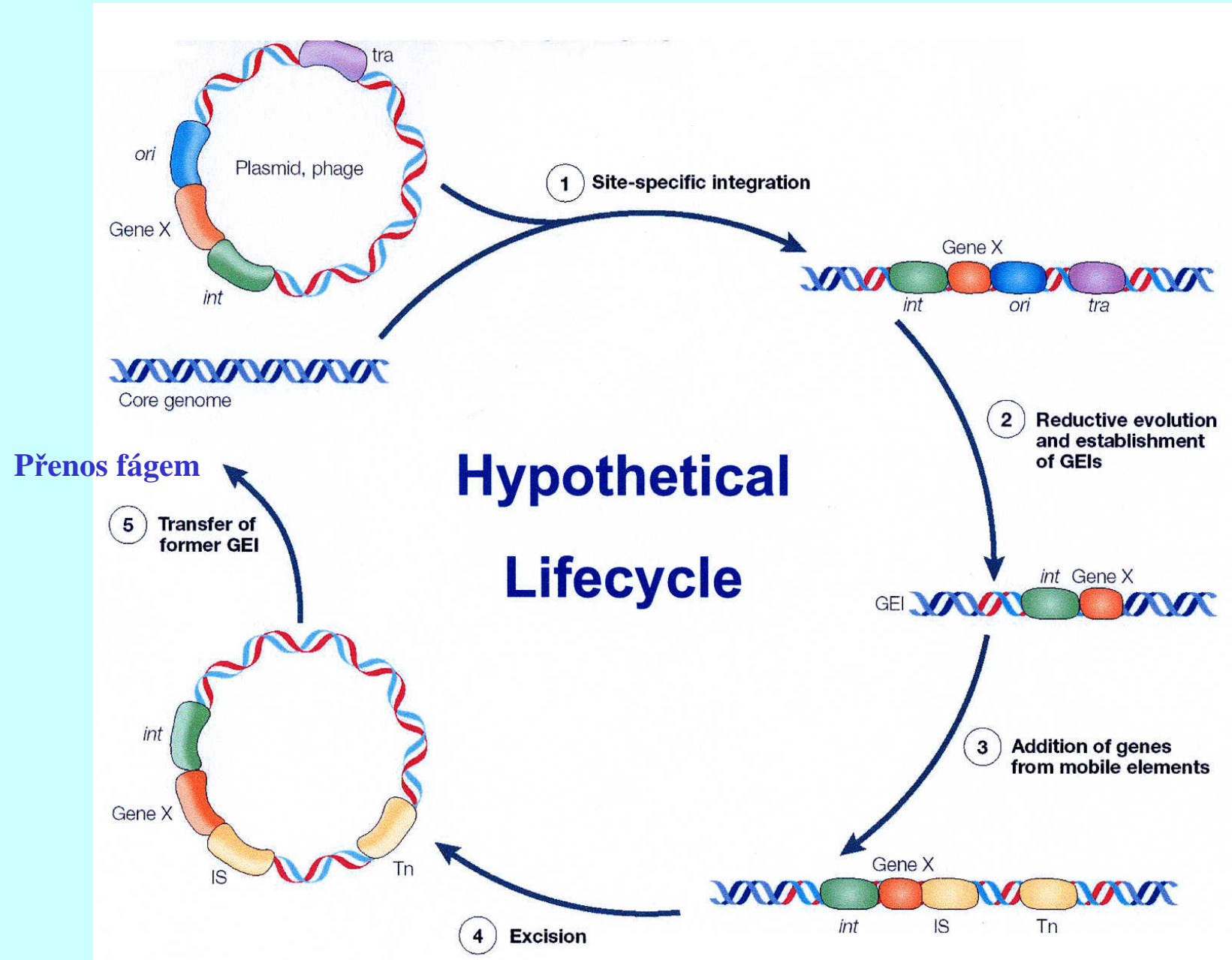
Význačné rysy ostrovů patogenity

- ◆ 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.

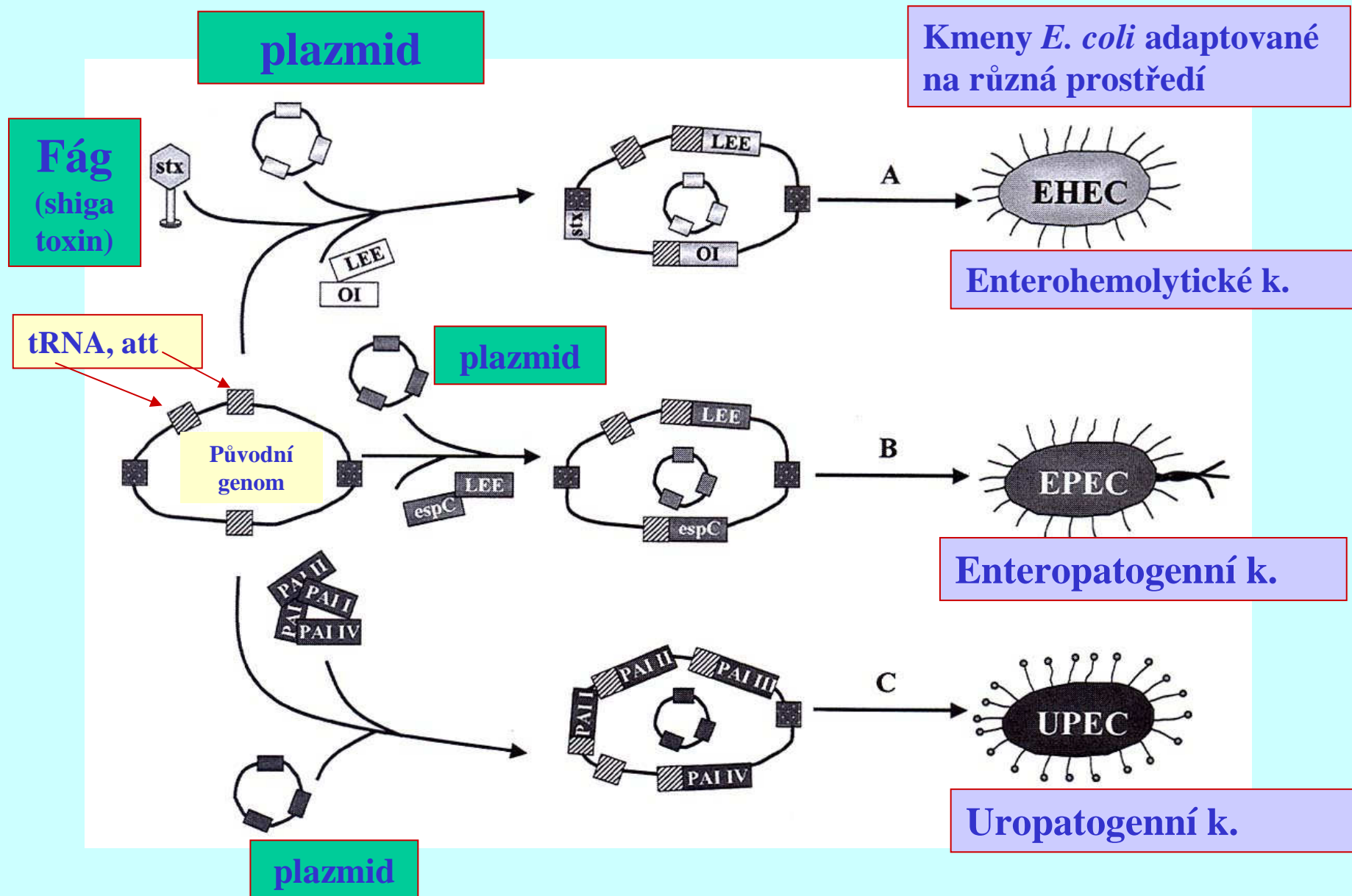
Distribuce ostrovů patogenity u *S. enteritica* serovar Typhi



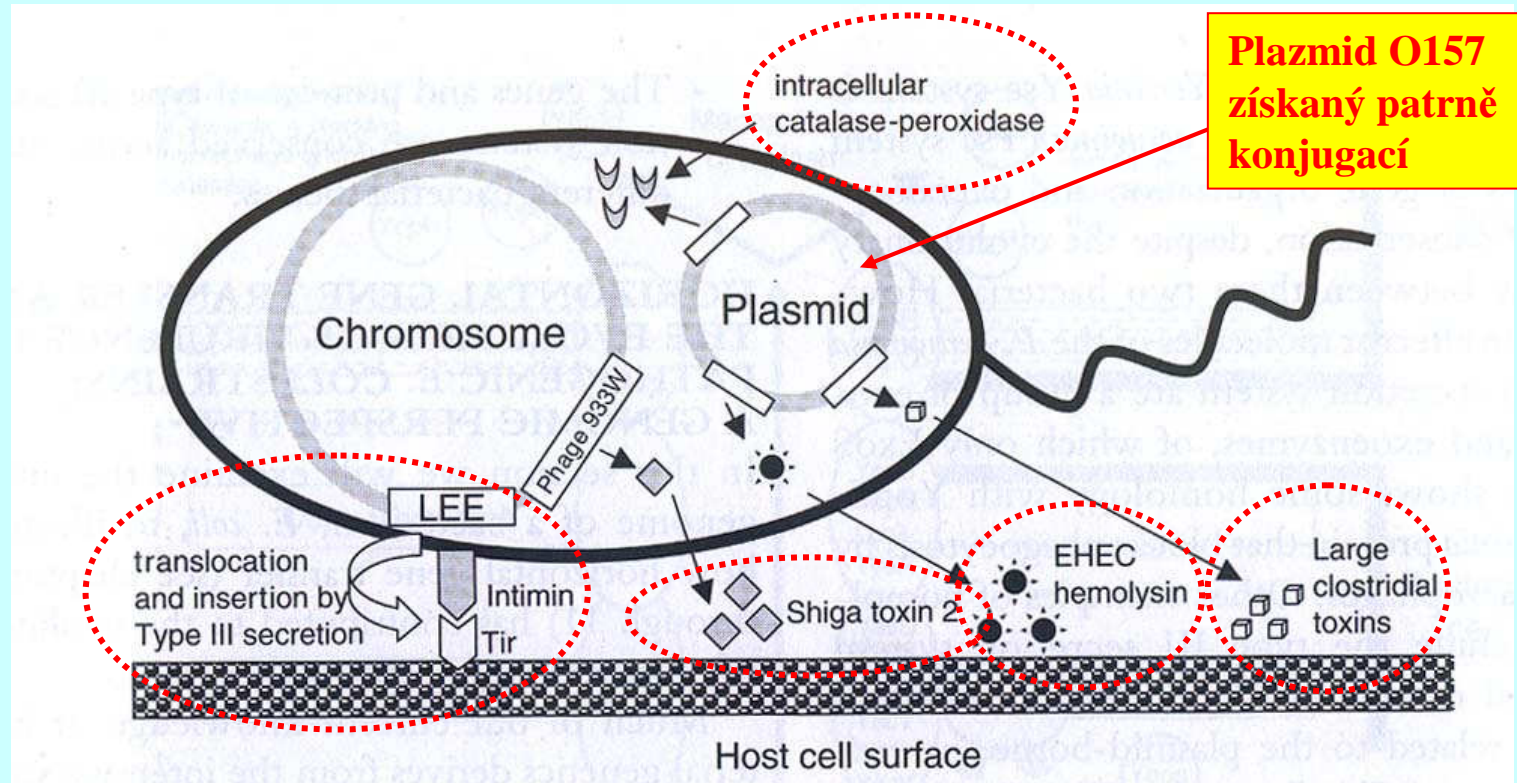
Vznik genomických ostrovů u patogenních a environmentálních mikrobu



Model vzniku ostrovů patogenity u patogenních *E. coli*

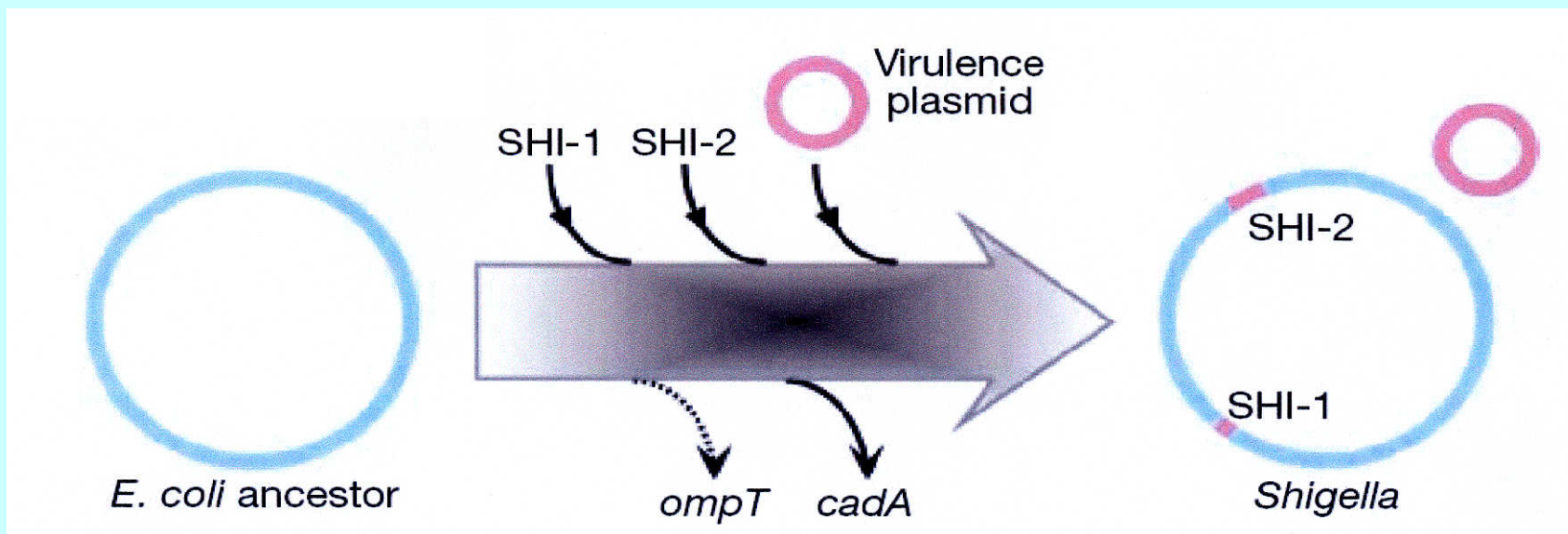


Horizontálně získané virulenční faktory zodpovědné za patogenitu enterohemorhagického kmene E. coli O157:H7



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

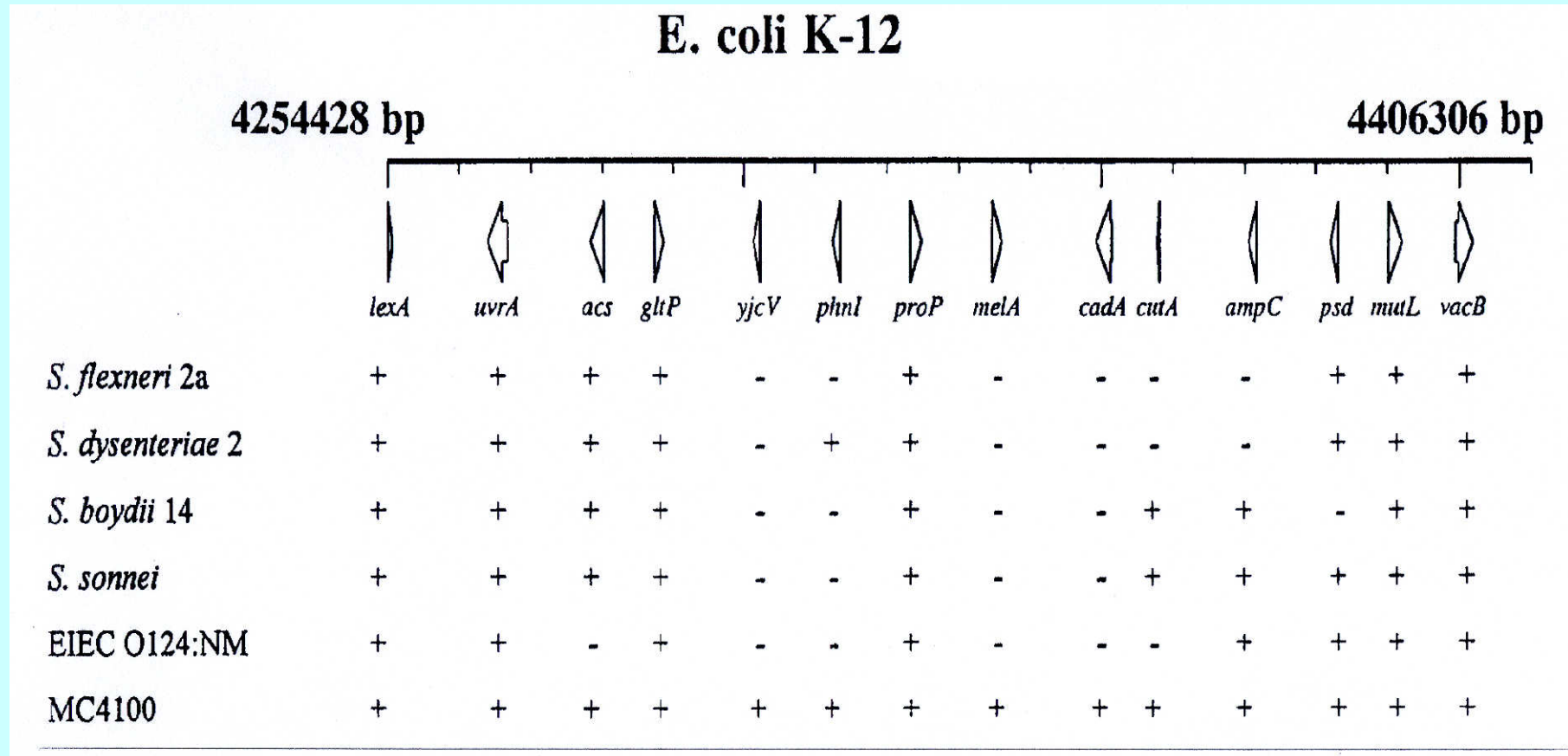
90% homologie DNA (!)

Kolinearita genů

Rekombinace po HGT

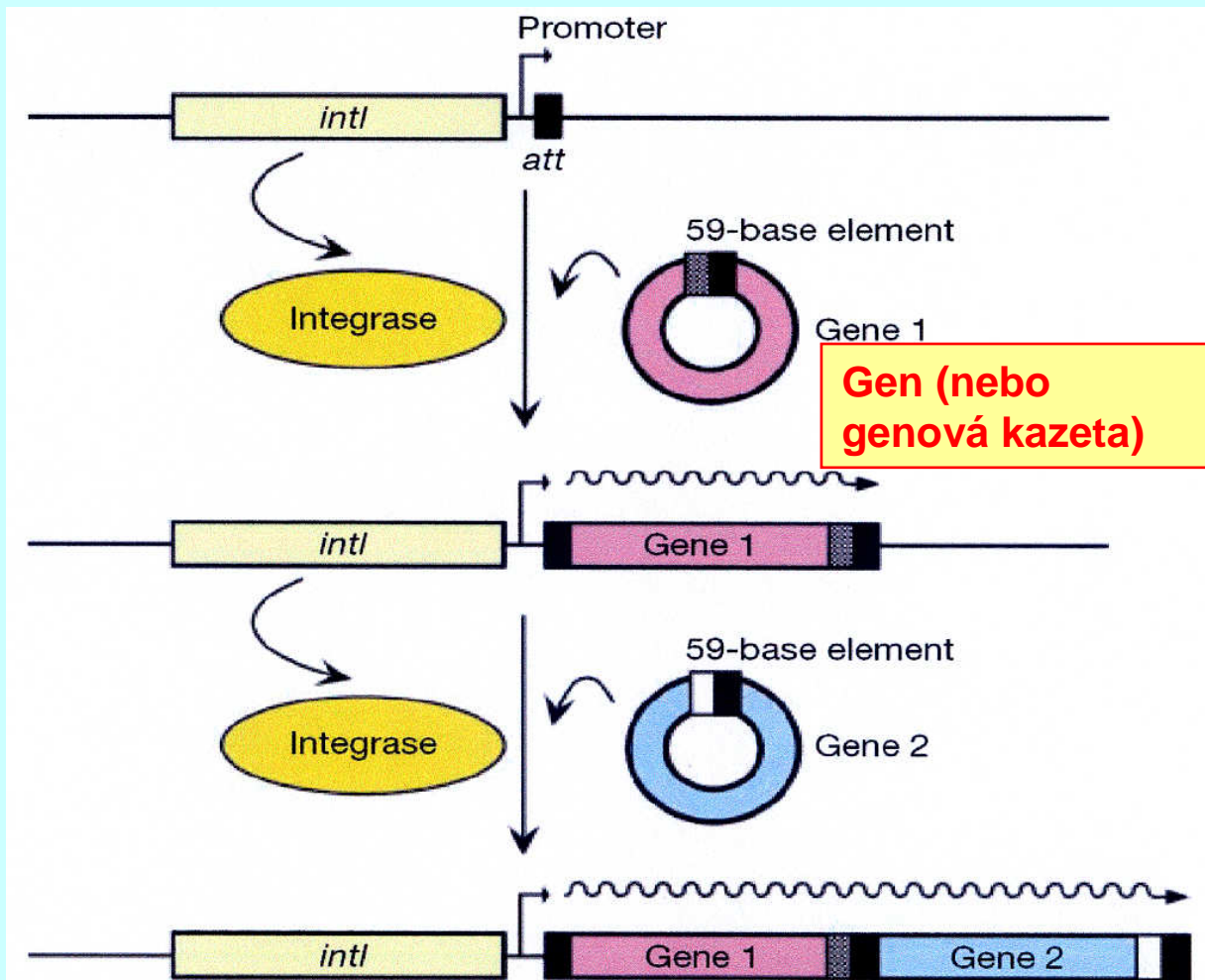
Vliv ztráty genů ztráty genů na patogenitu enterobakterí

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)

ZACHYTÁVÁNÍ GENŮ INTEGRONY



Genová kazeta v CTn (v plazmidu)

Integron obsahuje:

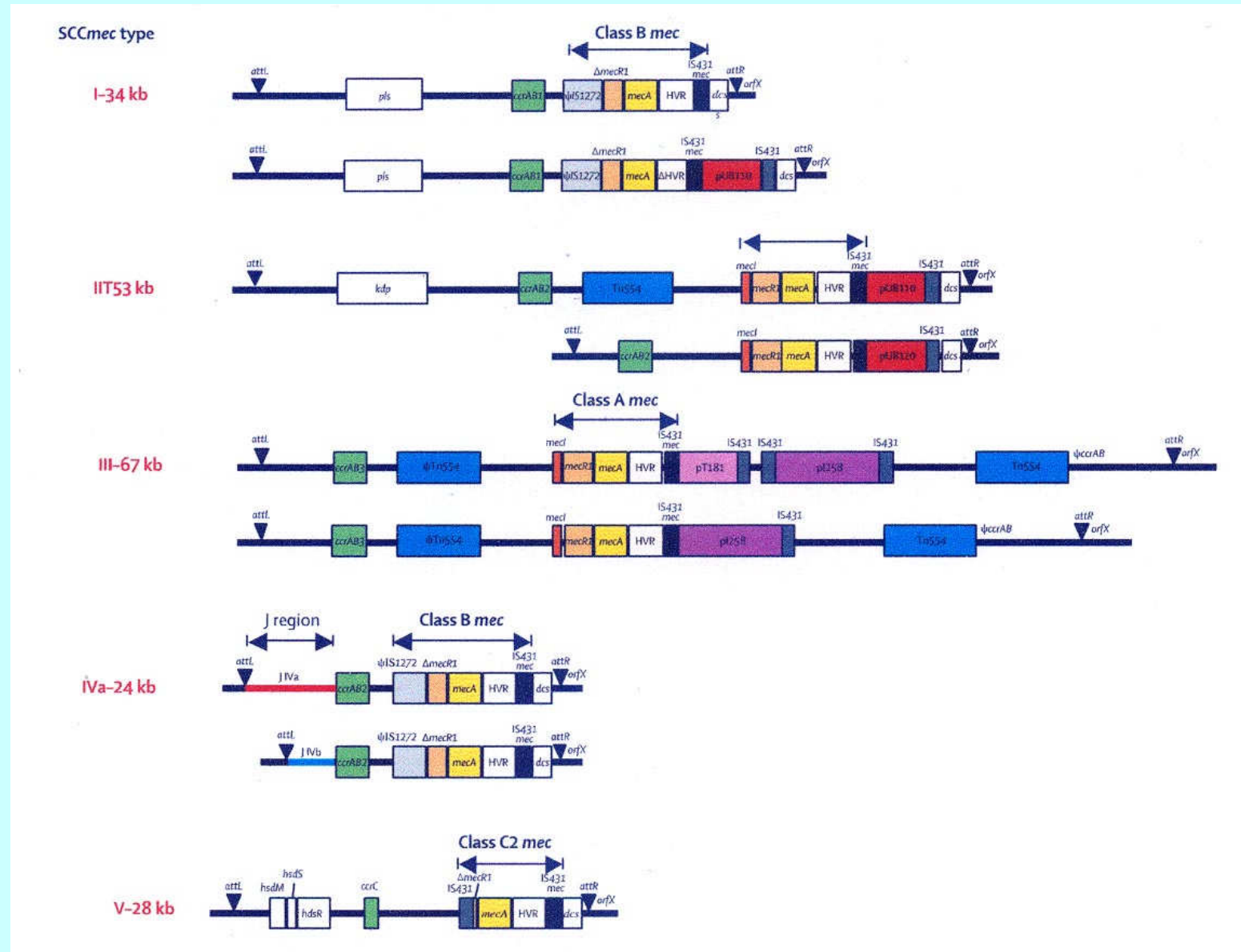
1. att místo, umožňující opakované zachycení genů nebo genových kazet

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

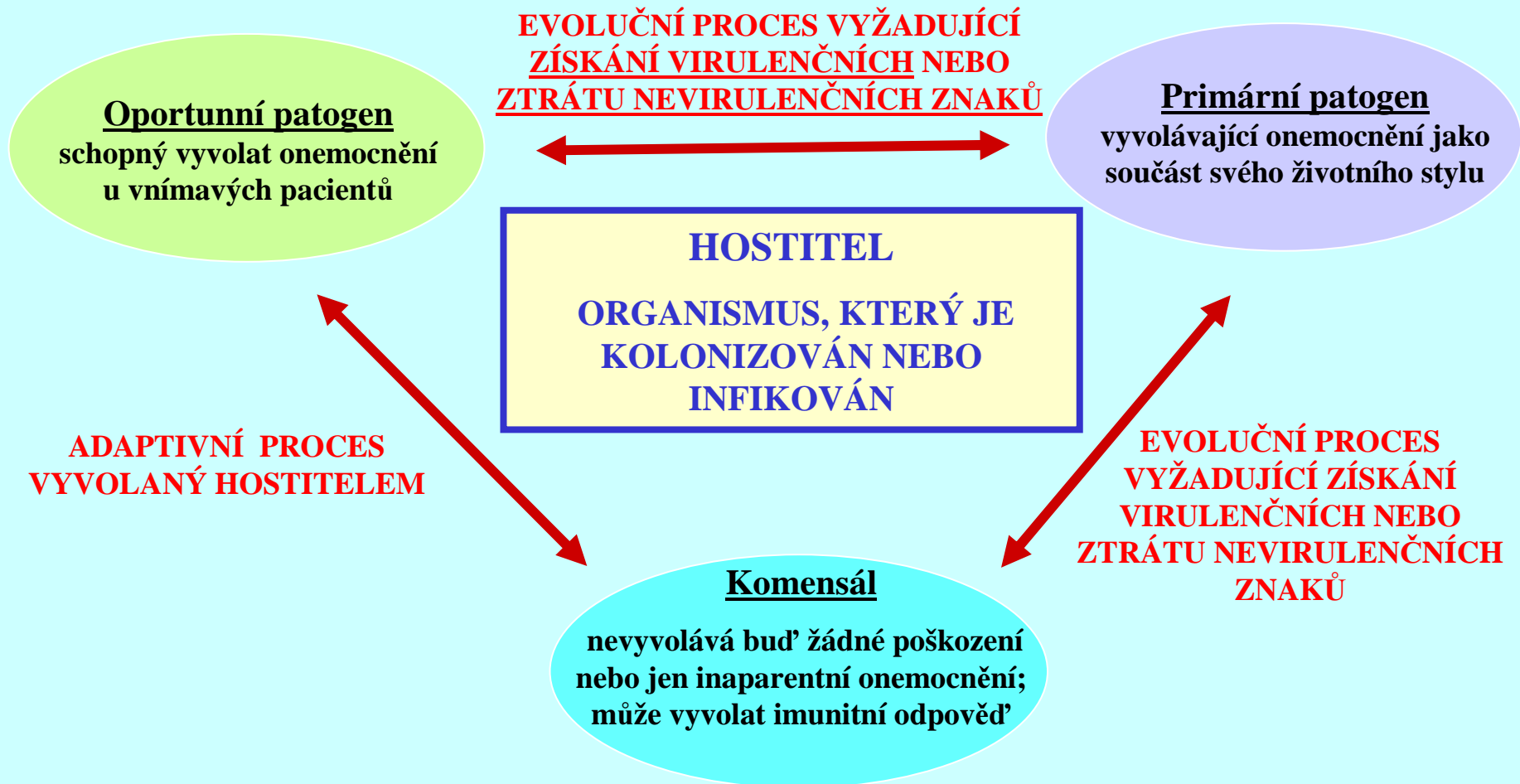
3. Promotor umožňující expresi vloženého genu

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

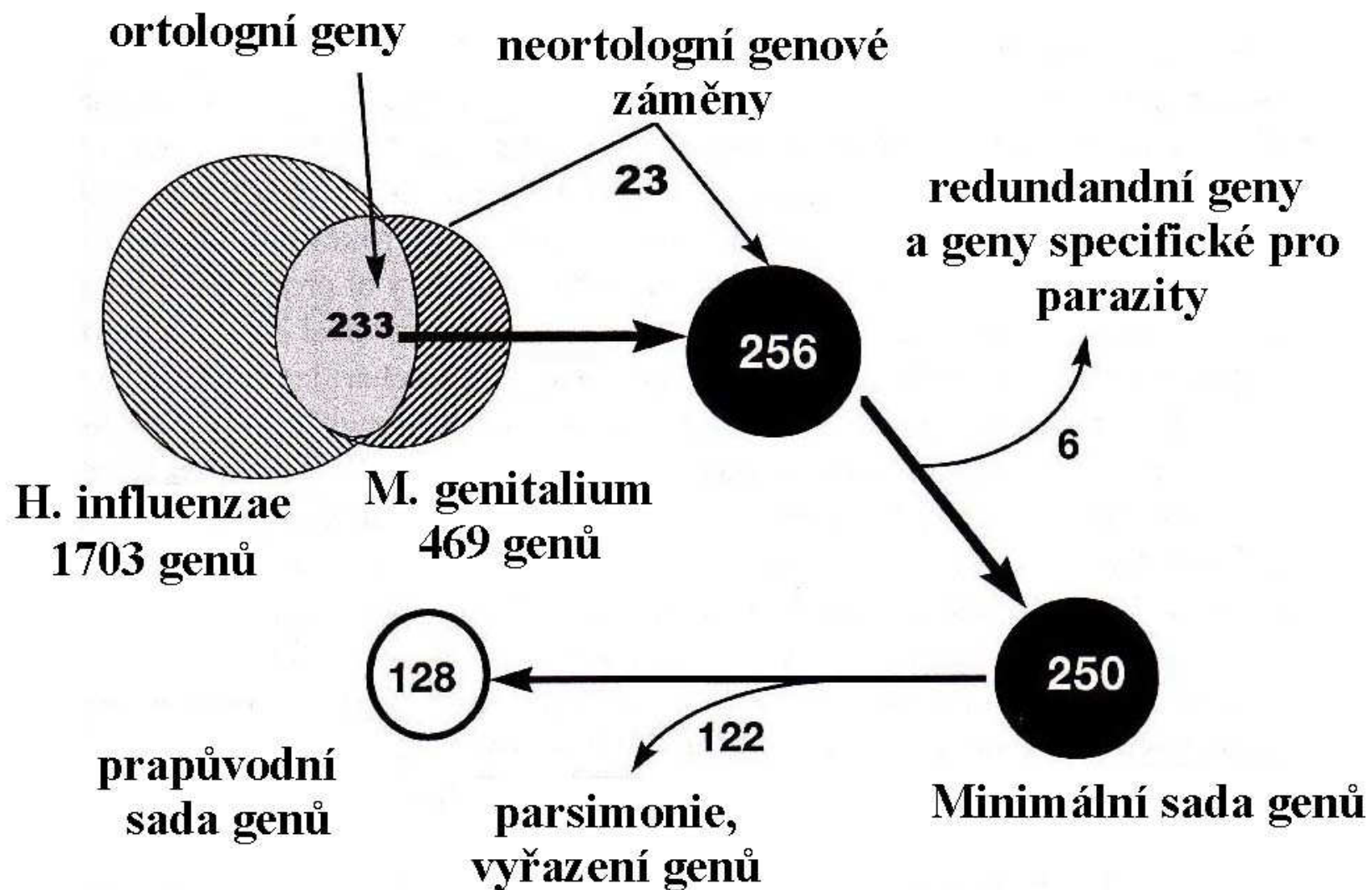
Typy stafylokokových chromozomových kazet (SCCmec) zodpovědných za rezistenci kmenů *S. aureus* k meticilinu



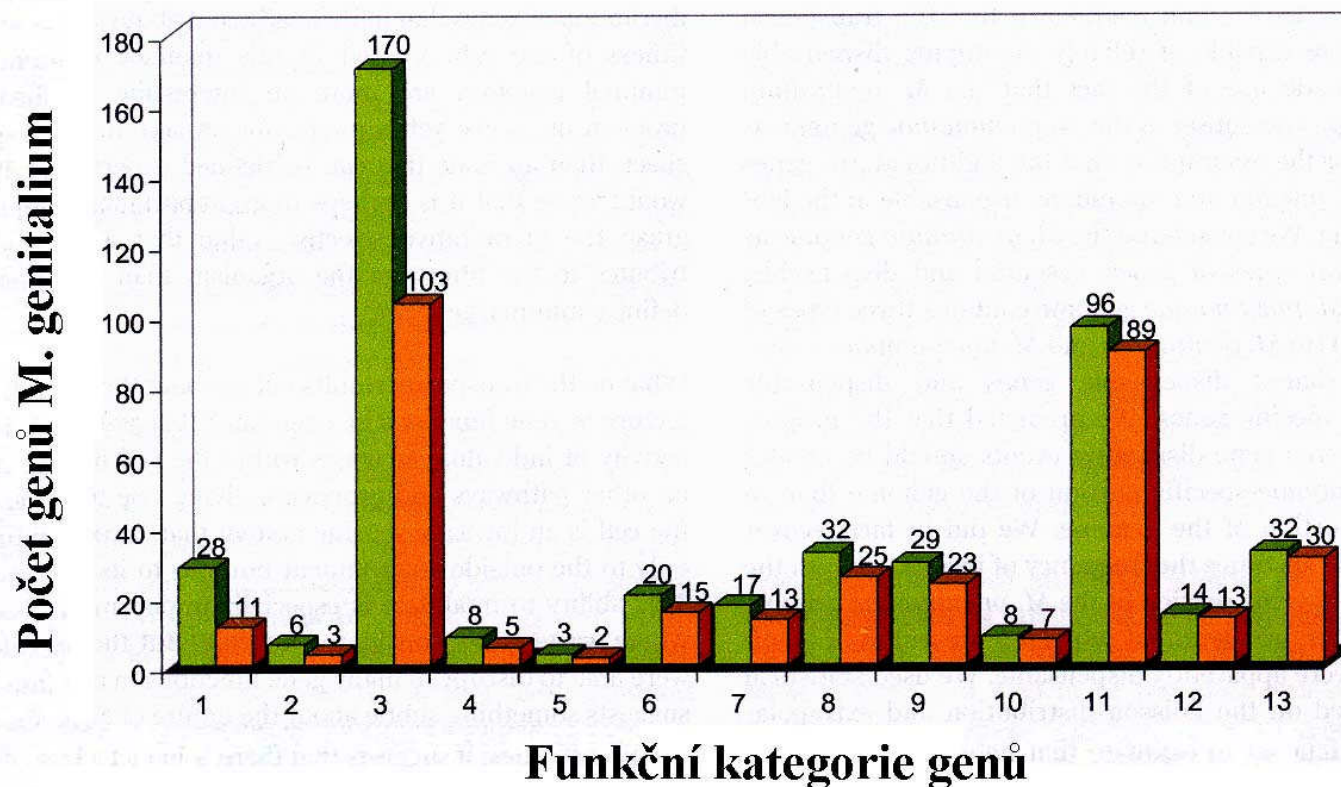
Interakce patogen-hostitel u bakteriálních infekčních onemocnění



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 *M. genitalium*
■ 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

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

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.

