



Biotechnologie léčiv – Základy genového inženýrství I.

Doc. RNDr. Jan Hošek, Ph.D.
hosek@mail.muni.cz

Ústav molekulární farmacie
FaF MU

Studijní materiály

- Prezentace PŘF:Bi8090 Genové inženýrství

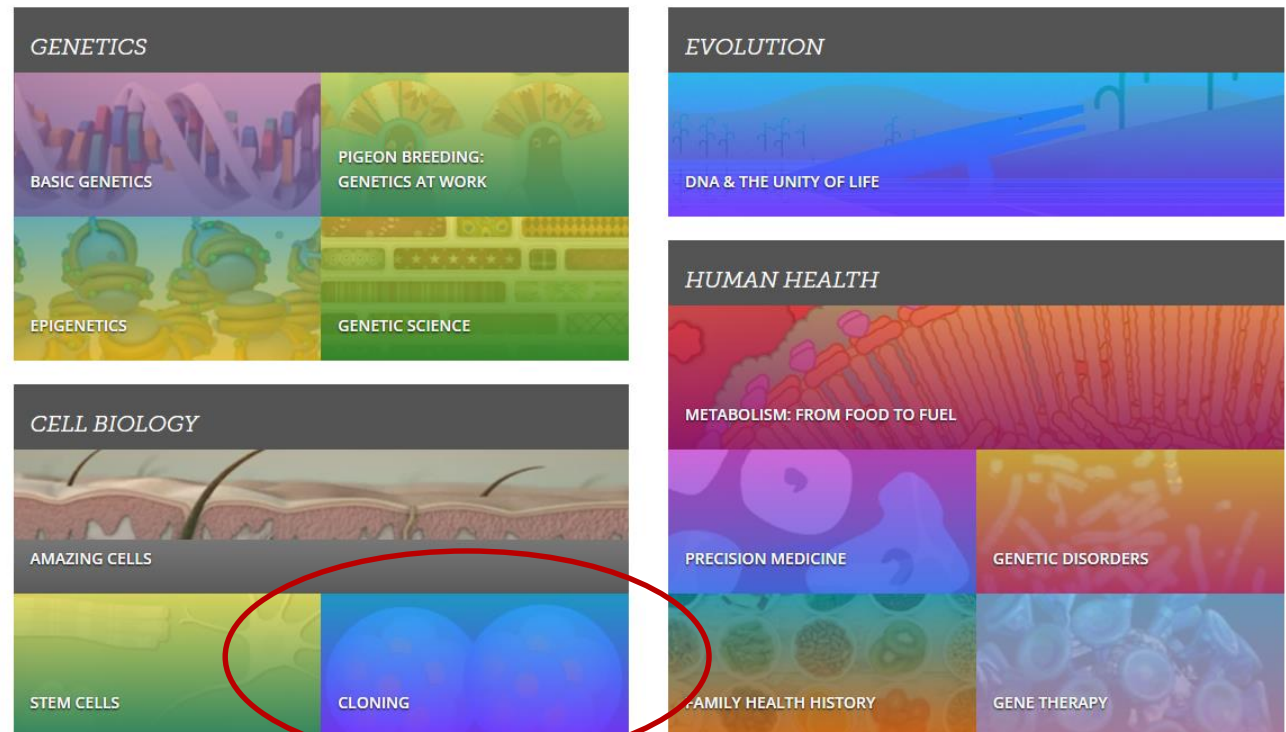




Zajímavá stránka



<https://learn.genetics.utah.edu/>



Biotechnologický produkt může vzniknout:

1. Klasickým biotechnologickým postupem

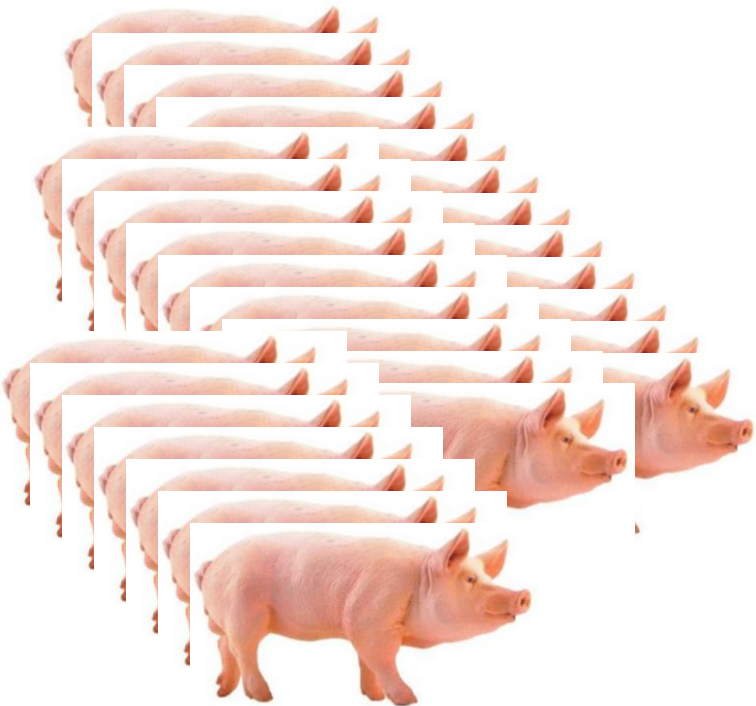
- křížení
- mutageneze populace buněk
- selekce buněk/organismů s vhodnými vlastnostmi
- nelze jimi přinutit organismus, aby produkoval protein, který mu není vlastní

2. Genovým inženýrstvím - rekombinantní DNA technologií

- klonování genů → technologie rekombinantní DNA
- genetické manipulace *in vitro*
- boří bariéry mezi druhy → heterologní systémy

Cíl rekombinantních technologií

Snaha geneticky zakódovat syntézu komerčně výhodného produktu do organismu, který jej bude produkovat levně a s vysokým výtěžkem.



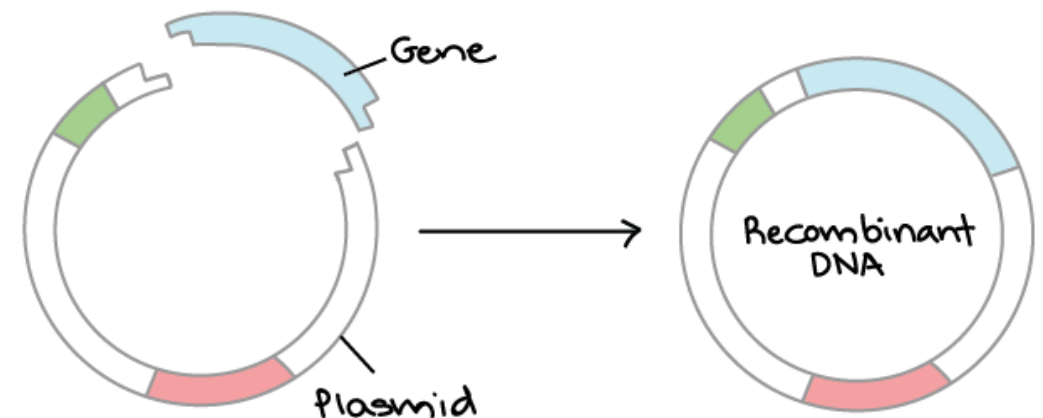
200 mL inzulinu

2000 kg slinivky

Základní principy rekombinantních technologií

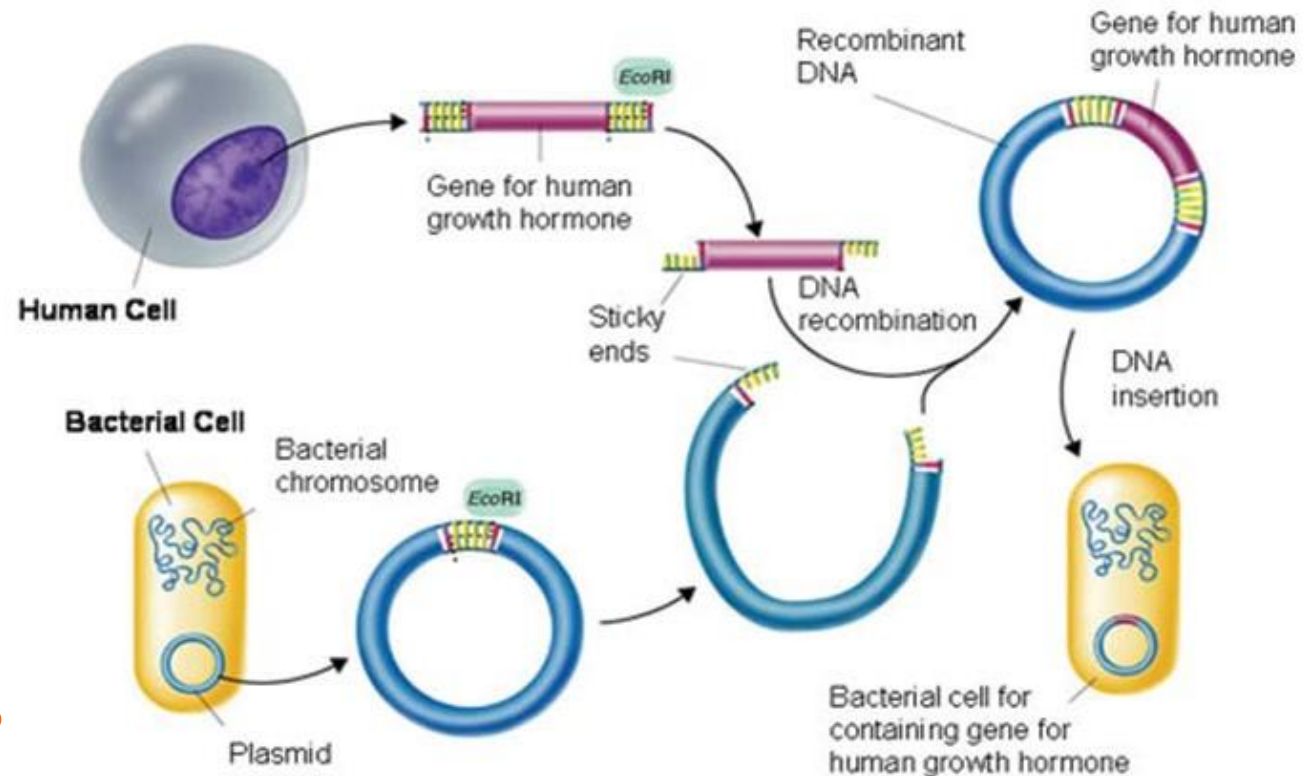
Rekombinantní technologie - klonování genů - potřebuje začlenění cizorodé DNA do hostitelské buňky

- **vnesení DNA (transformace) do buňky**
- **zabezpečení přežití cizorodé DNA**
- **schopnost replikace**
- **zajištění exprese**



Co je potřeba pro klonování genů

- vlastní gen – fragment DNA
- vektor – plasmid, fág, kosmid
- hostitel – příjemce rekombinantní DNA
- **inzert = gen začleněný do vektoru**
- **rekombinantní DNA = vektor s inzertem**



Důsledek klonování genů = transformace

Permanentní dědičná změna v genetickém materiálu buňky způsobená přijetím a začleněním cizí genetické informace

- **schopnost replikace**
- **schopnost exprese**

Zdroje cizí genetické informace

živočišný

rostlinný

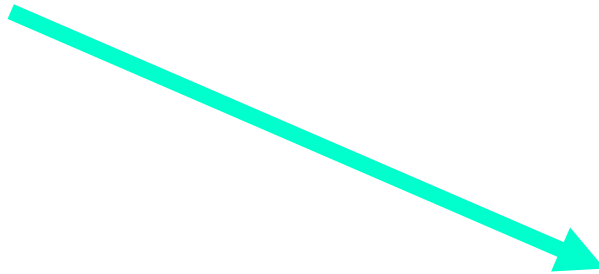
mikrobiální

syntetický

- Rekombinantní protein může být:
 - **Finální produkt** (např. hormon, protilátka)
 - **Nástroj pro další syntézu** (např. různé enzymy)

Jak získat gen nebo jeho fragment?

Izolovat DNA v nativním stavu
(viz Metody mol. biol.)



Opracovat enzymy

Provést PCR



Přímá syntéza



Pouze klonovat



A jak postupovat dál?

1) Získal jsem nukleovou kyselinu?

ANO



NE

2) Mám jí v dostatečném množství a kvalitní?

ANO



NE

a) Opracovat DNA enzymy (viz Metody mol. biol.)

b) Provést PCR (viz Metody mol. biol.)



Faktory ovlivňující expresi klonovaných genů

Regulační sekvence pro genovou expresi

1. Transkripce
 - a. Síla promotoru
 - b. Terminátor transkripce
 - c. Stabilita mRNA
2. Translace
 - a. Struktura vazebného místa pro ribozom
 - b. Využití kodonů
3. Transport proteinu
 - a. Charakter signální sekvence

Vlastnosti vektorů

1. Počet kopií vektoru v buňce
2. Stabilita vektoru

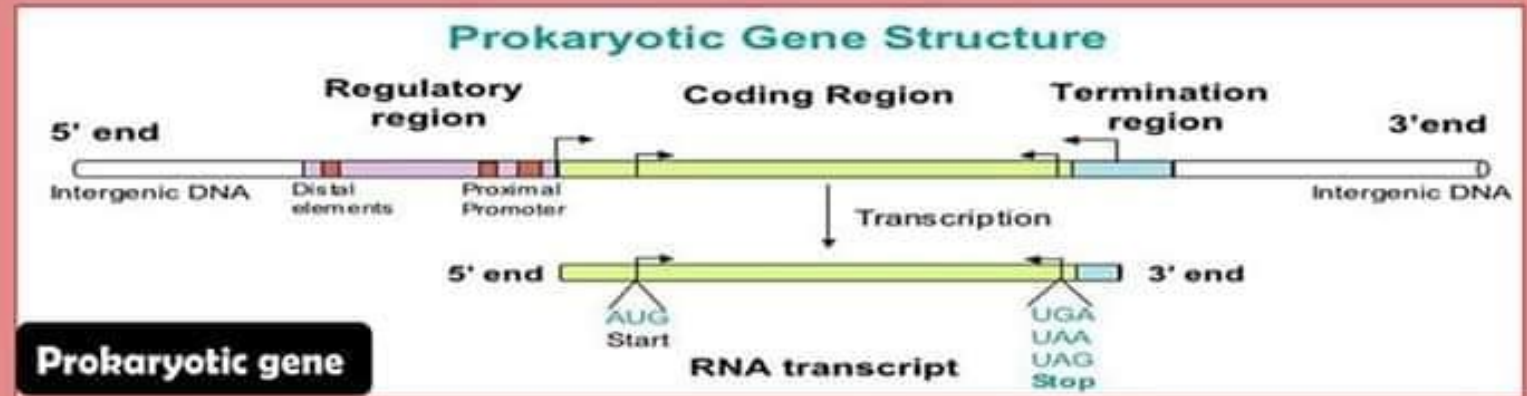
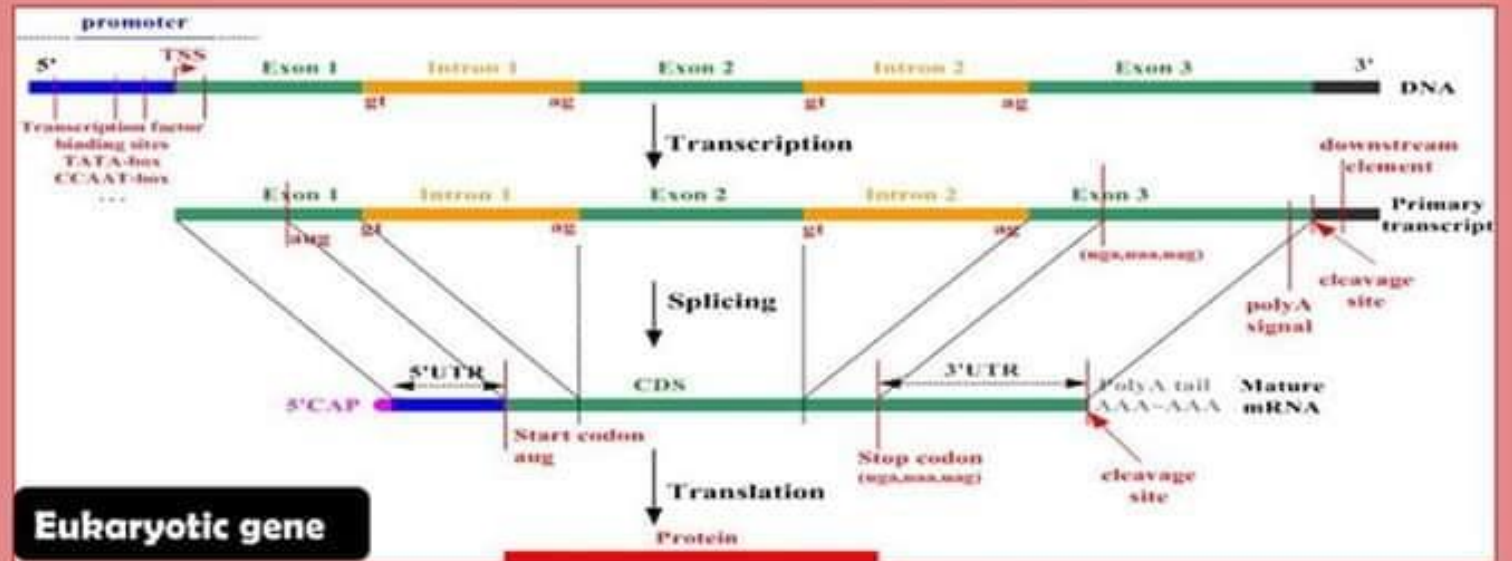
Fyziologie hostitelské buňky

1. Růstové podmínky
2. Enzymový aparát hostitelské buňky

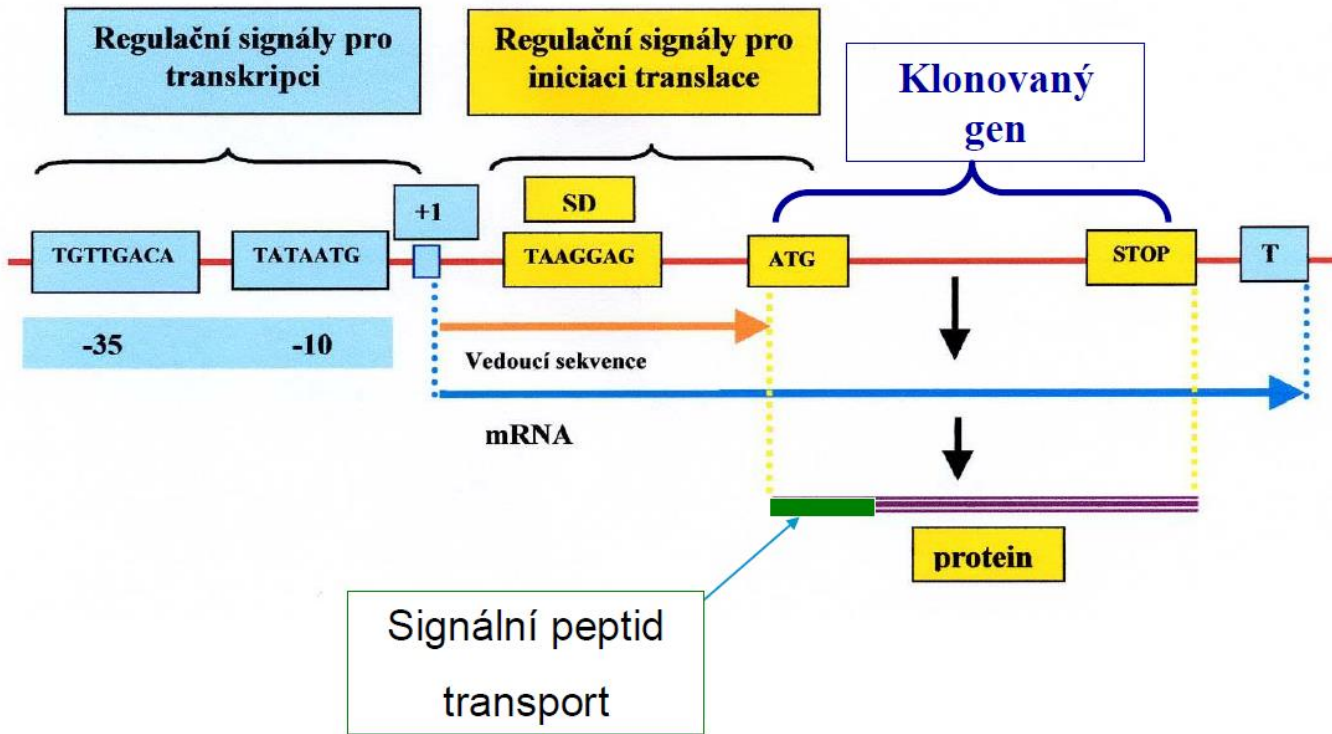
Vlastnosti inzertu

- Obsahuje kódující sekvenci
- Obsahuje začátek translace (start kodon ATG)
- Obsahuje konec translace (stop kodon)
- Nese vhodné kodony pro konkrétní aminokyseliny

Gene structure

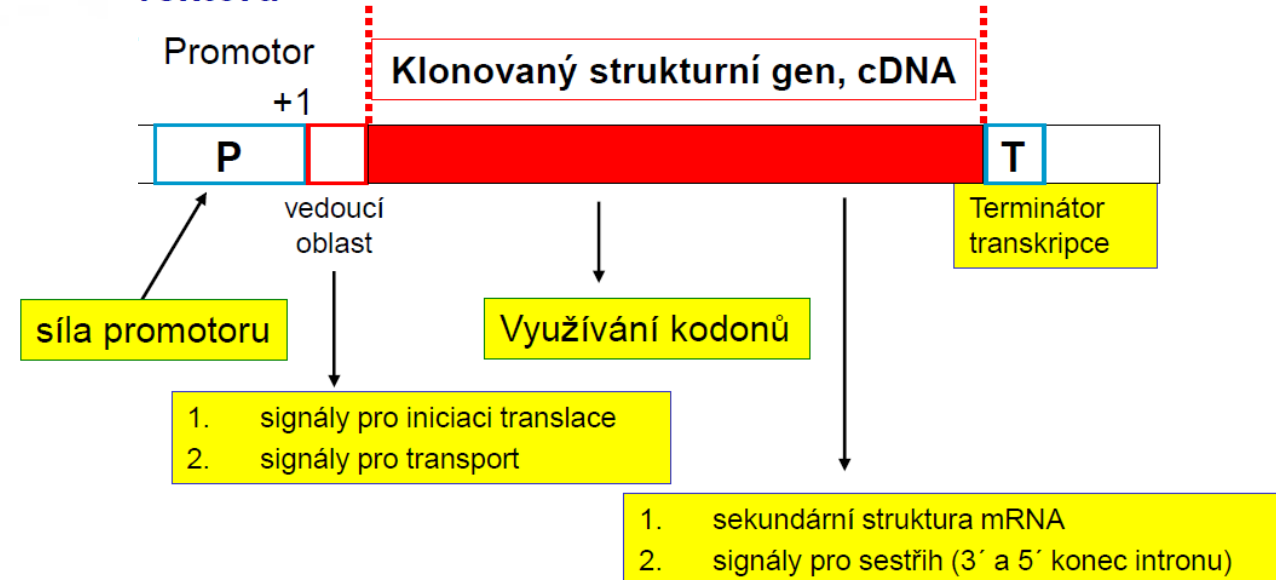


Signály ovlivňující transkripci a translaci strukturního genu prokaryot

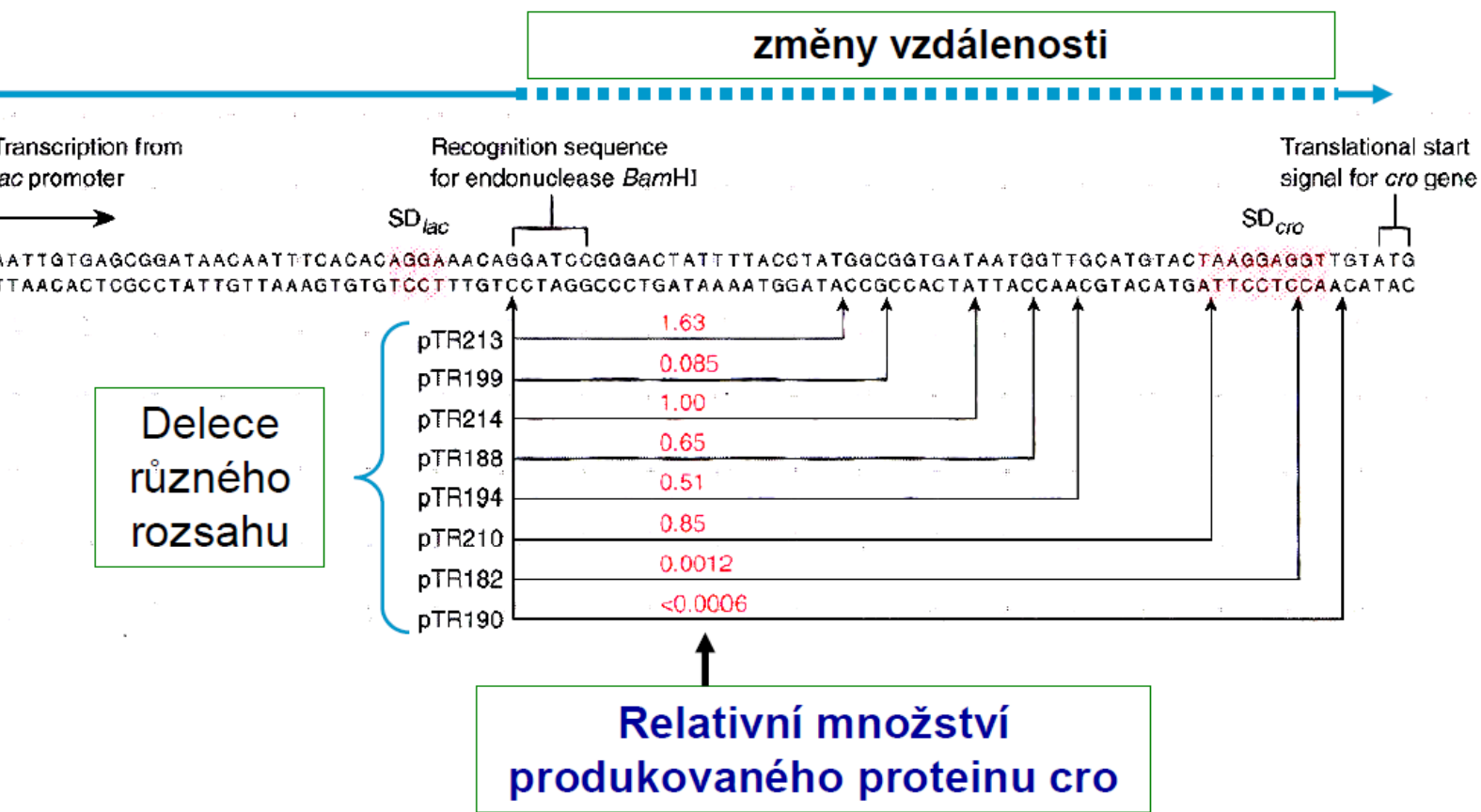


Regulační sekvence vektoru

Vlastnosti inzertu



Vliv vzdálenosti mezi promotorem a startem translace sekundární strukturu mRNA a na množství exprimovaného proteinu



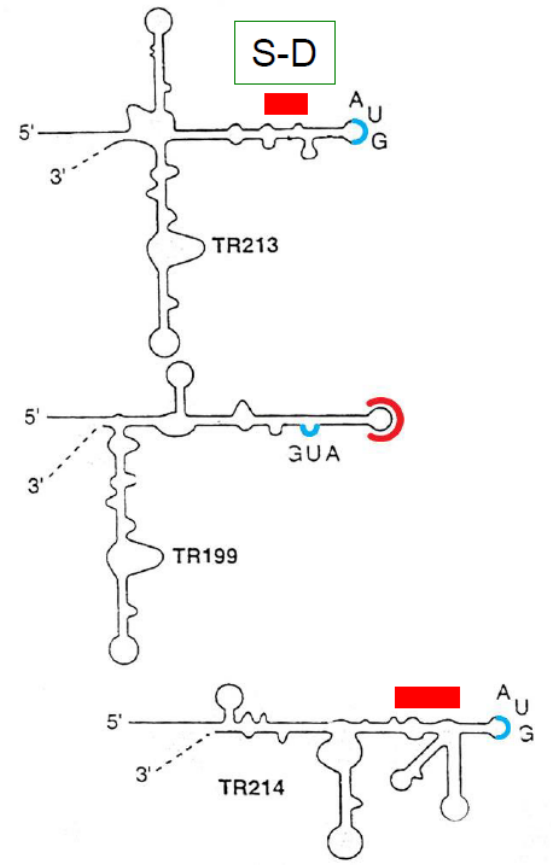
Sekundární struktura *cro*-mRNA

Množství proteinu

163%

8%

100%

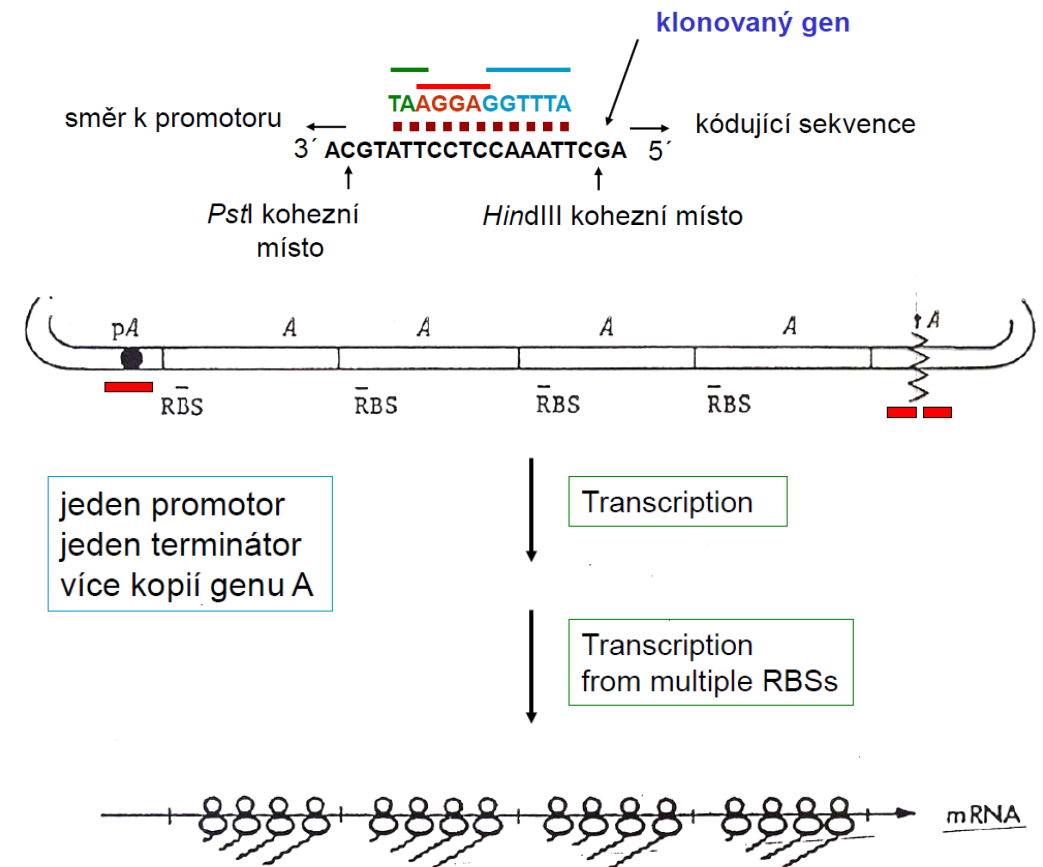


Možnosti zajištění vysoké proteosyntézy

- Rychlost proteosyntézy závisí na množství mRNA v buňce
- Modifikace v 3'-UTR nebo 5'-UTR mohou prodloužit poločas rozpadu mRNA
- Sekvence **PuPuUUUPuPu** poblíž Shine-Dalgarnovy sekvence je nezbytná pro translaci eukaryotických genů v *E. coli*
- Konstrukce homopolycistronických expresních kazet
- Proteiny v periplasmě mají delší poločas než v cytoplasmě
- Cílená inhibice specifických proteáz
- Tvorba fúzních proteinů – protein z hostitelského organismu (stabilizující partner) + protein zájmu

Synteticky připravené ribozomové vazebné místo

1. S-D sekvence
2. RRUUURR (GGTTTAA)
3. Terminační kodon TAA



Vliv využívání kodonů

- Různé druhy využívají kodony s různou frekvencí → souvisí s množstvím tRNA
- Řešení:
 - Klonování vzácných tRNA spolu s genem zájmu
 - Záměna vzácných kodonů za běžné cílenou mutagenézí

CODON USAGE IN *E. COLI* GENES¹

	Codon	Amino acid ²	% ³	Ratio ⁴	Codon	Amino acid	%	Ratio	Codon	Amino acid	%	Ratio	Codon	Amino acid	%	Ratio
U	UUU	Phe (F)	1.9	0.51	UCU	Ser (S)	1.1	0.19	UAU	Tyr (Y)	1.6	0.53	UGU	Cys (C)	0.4	0.43
	UUC	Phe (F)	1.8	0.49	UCC	Ser (S)	1.0	0.17	UAC	Tyr (Y)	1.4	0.47	UGC	Cys (C)	0.6	0.57
	UUA	Leu (L)	1.0	0.11	UCA	Ser (S)	0.7	0.12	UAA	STOP	0.2	0.62	UGA	STOP	0.1	0.30
	UUG	Leu (L)	1.1	0.11	UCG	Ser (S)	0.8	0.13	UAG	STOP	0.03	0.09	UGG	Trp (W)	1.4	1.00
C	CUU	Leu (L)	1.0	0.10	CCU	Pro (P)	0.7	0.16	CAU	His (H)	1.2	0.52	CGU	Arg (R)	2.4	0.42
	CUC	Leu (L)	0.9	0.10	CCC	Pro (P)	0.4	0.10	CAC	His (H)	1.1	0.48	CGC	Arg (R)	2.2	0.37
	CUA	Leu (L)	0.3	0.03	CCA	Pro (P)	0.8	0.20	CAA	Gln (Q)	1.3	0.31	CGA	Arg (R)	0.3	0.05
	CUG	Leu (L)	5.2	0.55	CCG	Pro (P)	2.4	0.55	CAG	Gln (Q)	2.9	0.69	CGG	Arg (R)	0.5	0.08
A	AUU	Ile (I)	2.7	0.47	ACU	Thr (T)	1.2	0.21	AAU	Asn (N)	1.6	0.39	AGU	Ser (S)	0.7	0.13
	AUC	Ile (I)	2.7	0.46	ACC	Thr (T)	2.4	0.43	AAC	Asn (N)	2.6	0.61	AGC	Ser (S)	1.5	0.27
	AUA	Ile (I)	0.4	0.07	ACA	Thr (T)	0.1	0.30	AAA	Lys (K)	3.8	0.76	AGA	Arg (R)	0.2	0.04
	AUG	Met (M)	2.6	1.00	ACG	Thr (T)	1.3	0.23	AAG	Lys (K)	1.2	0.24	AGG	Arg (R)	0.2	0.03
G	GUU	Val (V)	2.0	0.29	GCU	Ala (A)	1.8	0.19	GAU	Asp (D)	3.3	0.59	GGU	Gly (G)	2.8	0.38
	GUC	Val (V)	1.4	0.20	GCC	Ala (A)	2.3	0.25	GAC	Asp (D)	2.3	0.41	GGC	Gly (G)	3.0	0.40
	GUA	Val (V)	1.2	0.17	GCA	Ala (A)	2.1	0.22	GAA	Glu (E)	4.4	0.70	GGA	Gly (G)	0.7	0.09
	GUG	Val (V)	2.4	0.34	GCG	Ala (A)	3.2	0.34	GAG	Glu (E)	1.9	0.30	GGG	Gly (G)	0.9	0.13
	U				C				A				G			

¹ The data shown in this table is from the Arabidopsis Research Companion on the World Wide Web (<http://weeds/mgh.harvard.edu>). Codon frequencies for many other bacteria can be found at <http://morgan.angis.su.oz.au/Angis/Tables.html>.

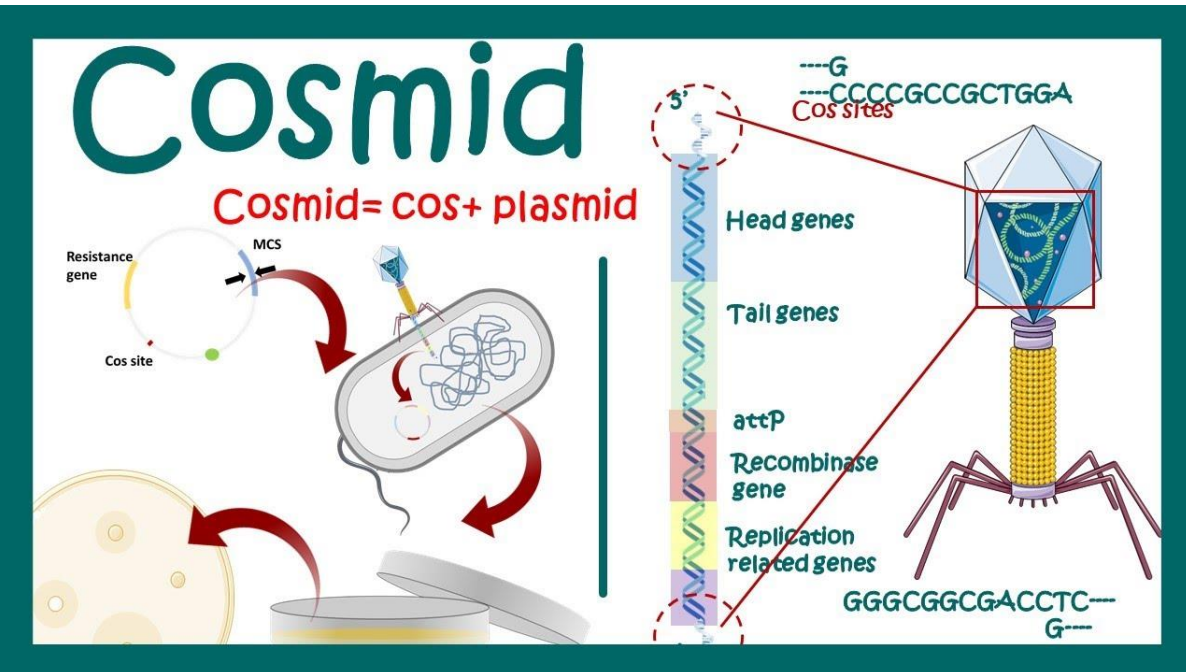
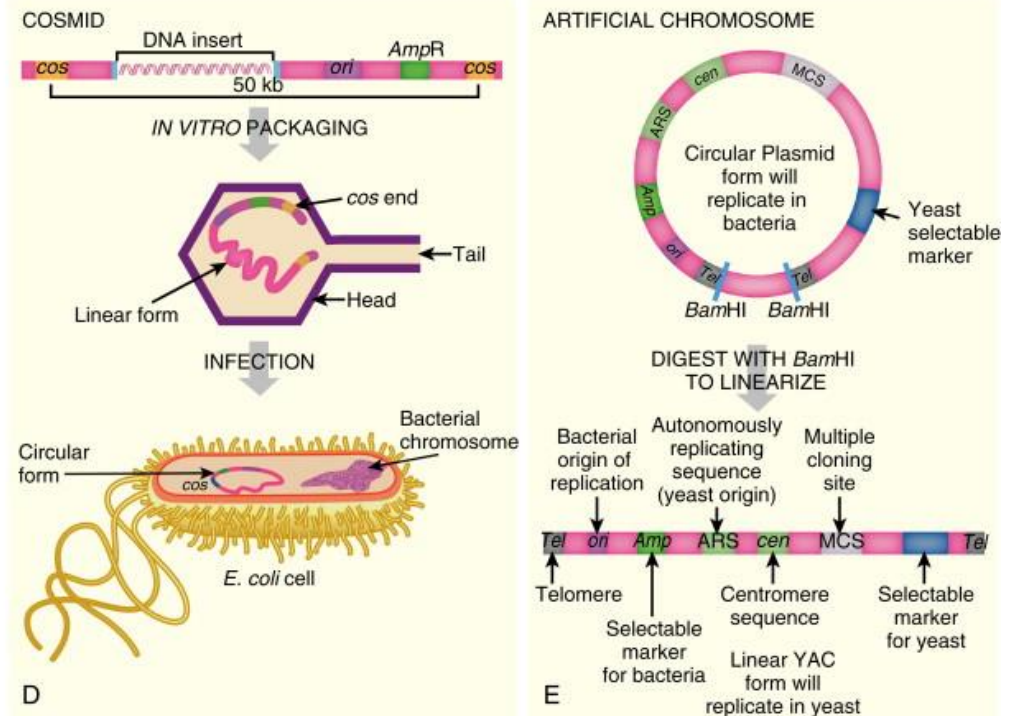
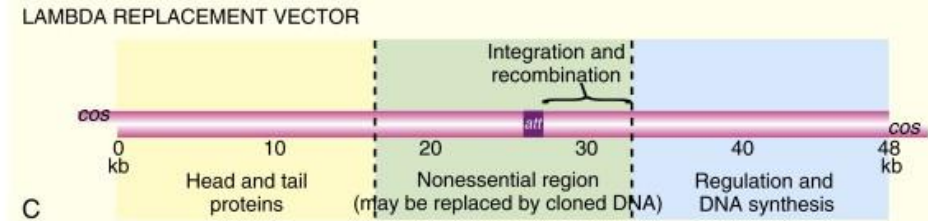
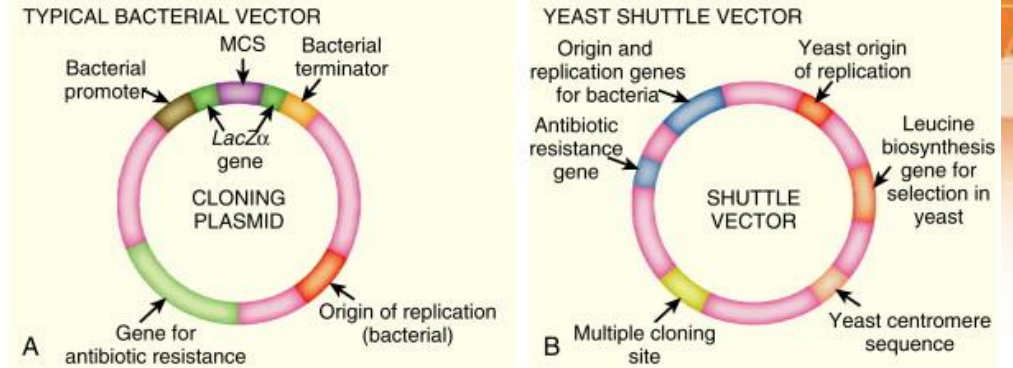
² The letter in parenthesis represents the one-letter code for the amino acid.

³ % represents the average frequency this codon is used per 100 codons.

⁴ Ratio represents the abundance of that codon relative to all of the codons for that particular amino acid.

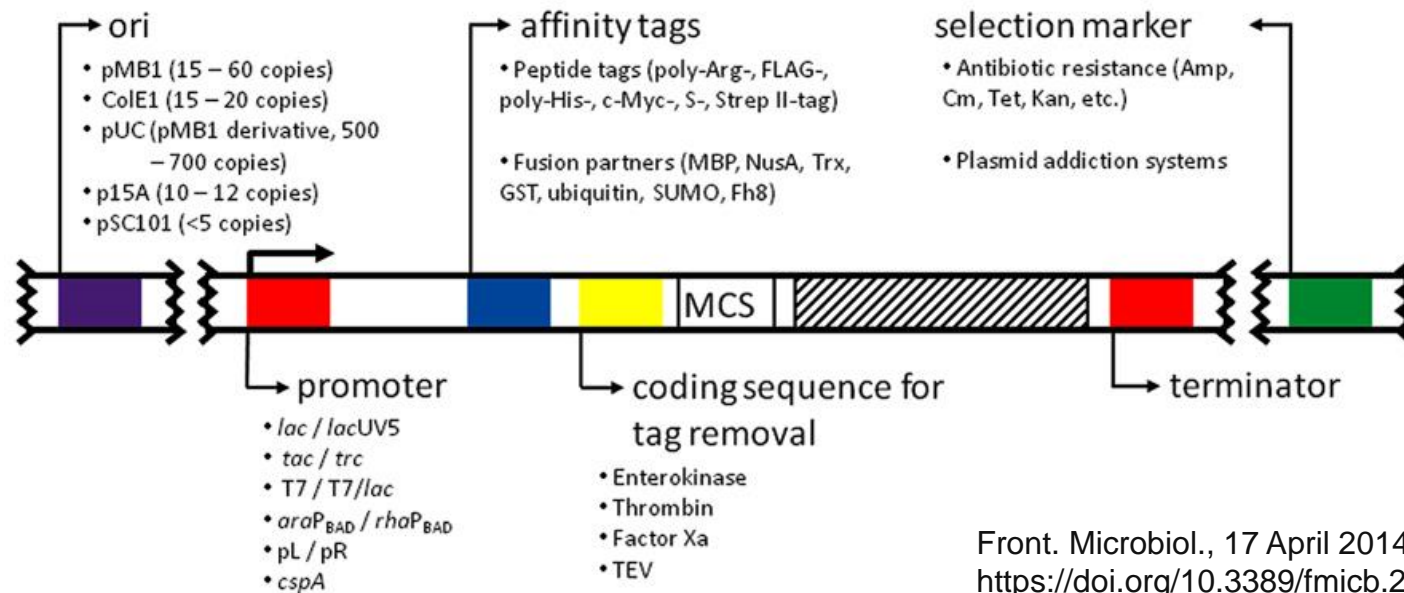
Vektory

- Plasmidy
- Bakteriofágy
- Kosmidy
- Umělé chromozómy
 - BAC (bacterial artificial chromosomes)
 - YAC (yeast artificial chromosomes)



Anatomie expresního vektoru

- 1) počátek replikace (**ori**), který je podmínkou produkce nových kopií
- 2) **inducibilní promotor**, který umožní regulovat expresi požadovaného proteinu
- 3) **selekční znak**, zajišťující zvýhodněný růst transformovaných bakterií
- 4) **klonovací místo (MSC)**, umožňující vložit do plasmidu fragment cizorodé DNA

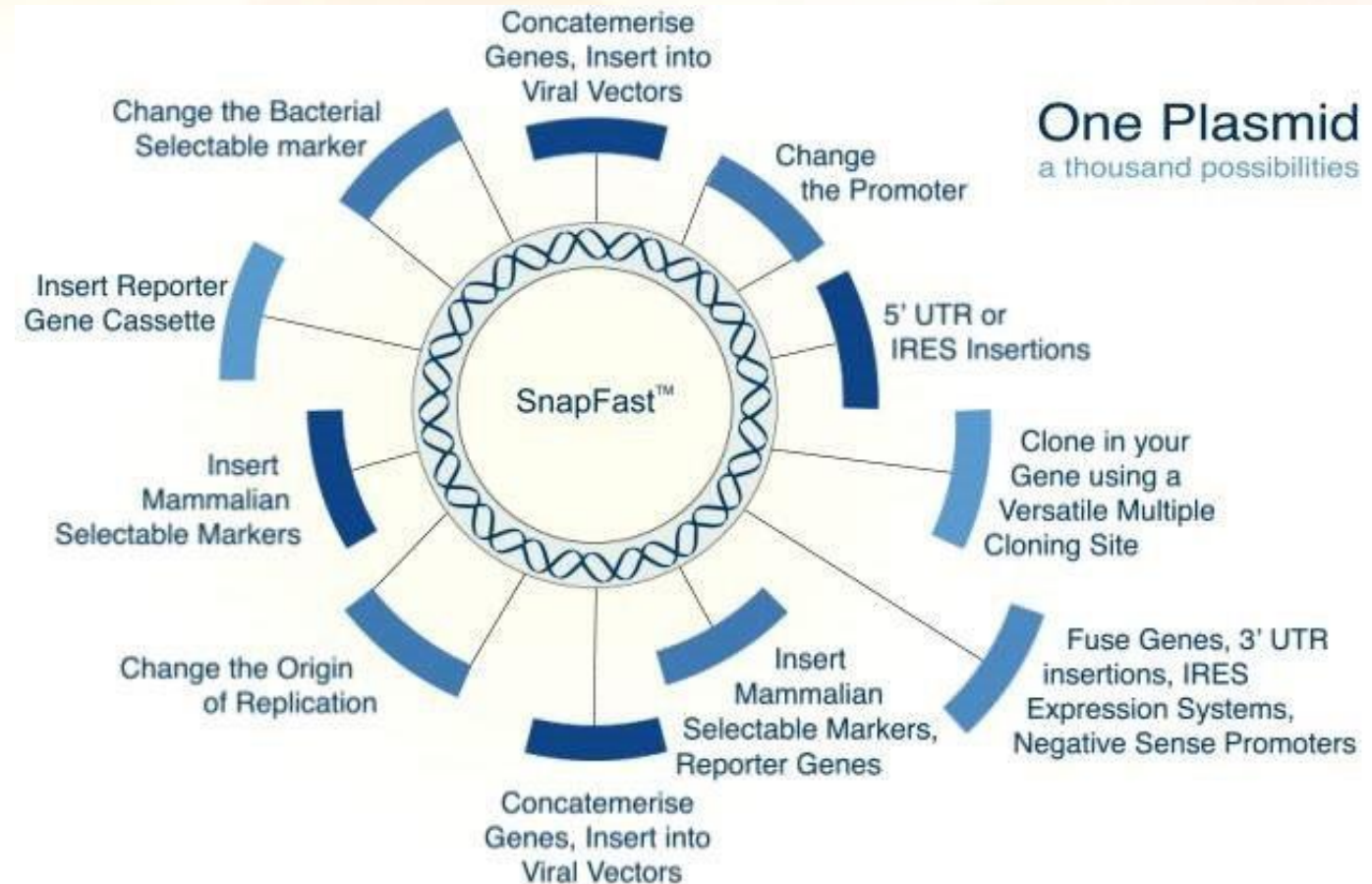


Příklady regulovatelných promotorů v expresních vektorech

		Promotor	Zdroj	Způsob regulace	
				Off	On
<i>E. coli</i>		λ pL, λ pR	Leftward and rightward early promoters of λ	30°C	>37°C (in <i>cl</i> ₈₅₇ host)
<div style="border: 1px solid black; padding: 5px;"> <p>lac-UV5</p> <p>nezávislost na katabolické represi</p> </div>	←	<i>lac</i>	<i>E. coli lac</i> operon	—	IPTG in medium
		<i>trp</i>	<i>E. coli trp</i> operon	Tryptophan in medium	Indoleacetic acid in medium
		<i>tac</i>	<i>trp</i> -35 region <i>lac</i> -10 region hybrid	—	IPTG in medium
		<i>phoA</i>	<i>E. coli</i> alkaline phosphatase operon	Excess phosphate in medium	Phosphate-limited medium
		<i>recA</i>	<i>E. coli recA</i> gene	—	Mitomycin C in medium
<div style="background-color: yellow; border: 1px solid black; padding: 5px;"> <p>Využití u eukaryot</p> </div>	←	<i>tet</i>	Tn10 tetracycline-resistance gene	—	Tetracyclines in medium

Plasmidy jako vektory

- extrachromozomální kružnicové dsDNA
- výskyt u mnoha bakteriálních druhů
- velikost 1 000 až 200 000 bp
- nesou pouze geny kódující druhotné znaky (rezistence k antibiotikům)
- autonomní replikace
- **velikost inzertu = do 25 kbp**



Trend poslední doby → plasmid syntheses on demand

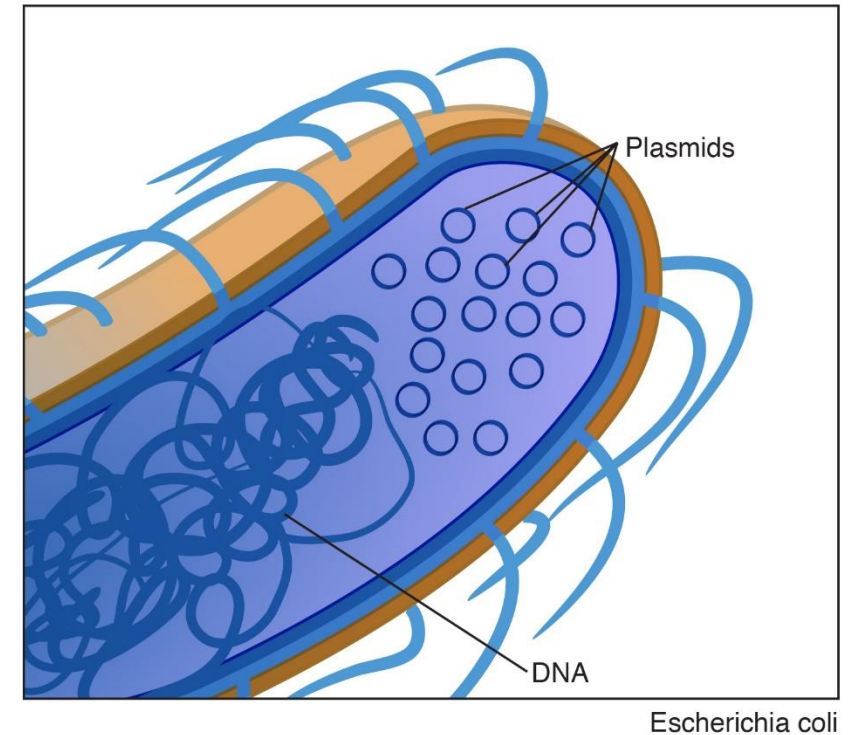
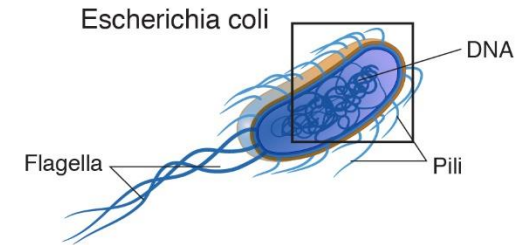
The screenshot shows the VectorBuilder website interface for designing a Mammalian Gene Expression Vector. The page is titled "Vector Design Studio" and includes a navigation menu with options like "Design Vector", "Products & Services", "GMP Manufacturing", "Tools", "Resources", and "Support". The current path is "Home > Choose Vector System > Mammalian Gene Expression Vector".

The main content area is titled "Mammalian Gene Expression Vector" and includes a "Guide" link and a "Design Another Vector" button. A "Select number of ORFs" section has buttons for 1, 2, 3, and 4, with a red callout box stating: "To express multiple ORFs as a polycistron, select the number of ORFs on the left." A "Finish Design" button is also present.

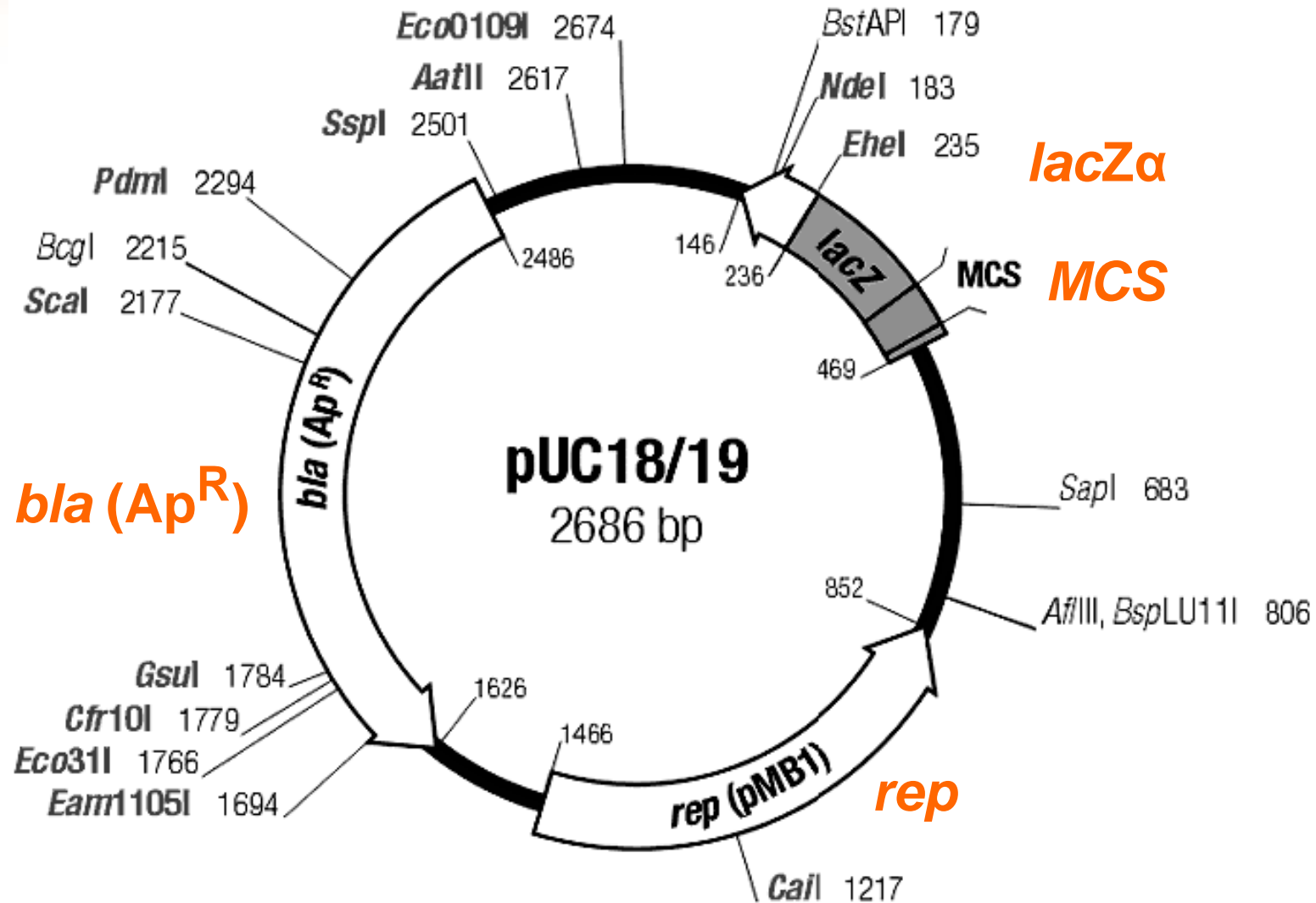
The central diagram is a circular plasmid map with several key features labeled: "Ampicillin" (antibiotic resistance), "pUC ori" (plasmid origin of replication), "SV40 late pA" (polyA signal), "Kozak" (start codon), and three design steps: "Step 1 | Add Promoter", "Step 2 | Add ORF", and "Step 3 | Add Marker".

Kritéria vhodnosti plasmidu

- 1) malá velikost = schopnost transformace
- 2) stabilita plasmidu
- 3) velký počet kopií v buňce = výtěžek
- 4) snadná manipulace
- 5) „shuttle“ vektory = fungují ve více hostitelských druhů (např. *E. coli* + savčí buňky)



Plasmidy pUC18 a pUC19



Polylinker pUC18 a pUC19

MCS = multi-cloning site

pUC19 MCS

EcoRI SacI SmaI XbaI SbfI
 KpnI BamHI SalI PstI SphI HindIII

```
agtgAATTCGAGCTCGGTACCCGGGGATCCTCTAGAGTCGACCTGCAGGCATGCAAGCTTGGcgtaatcatggtcacat
```

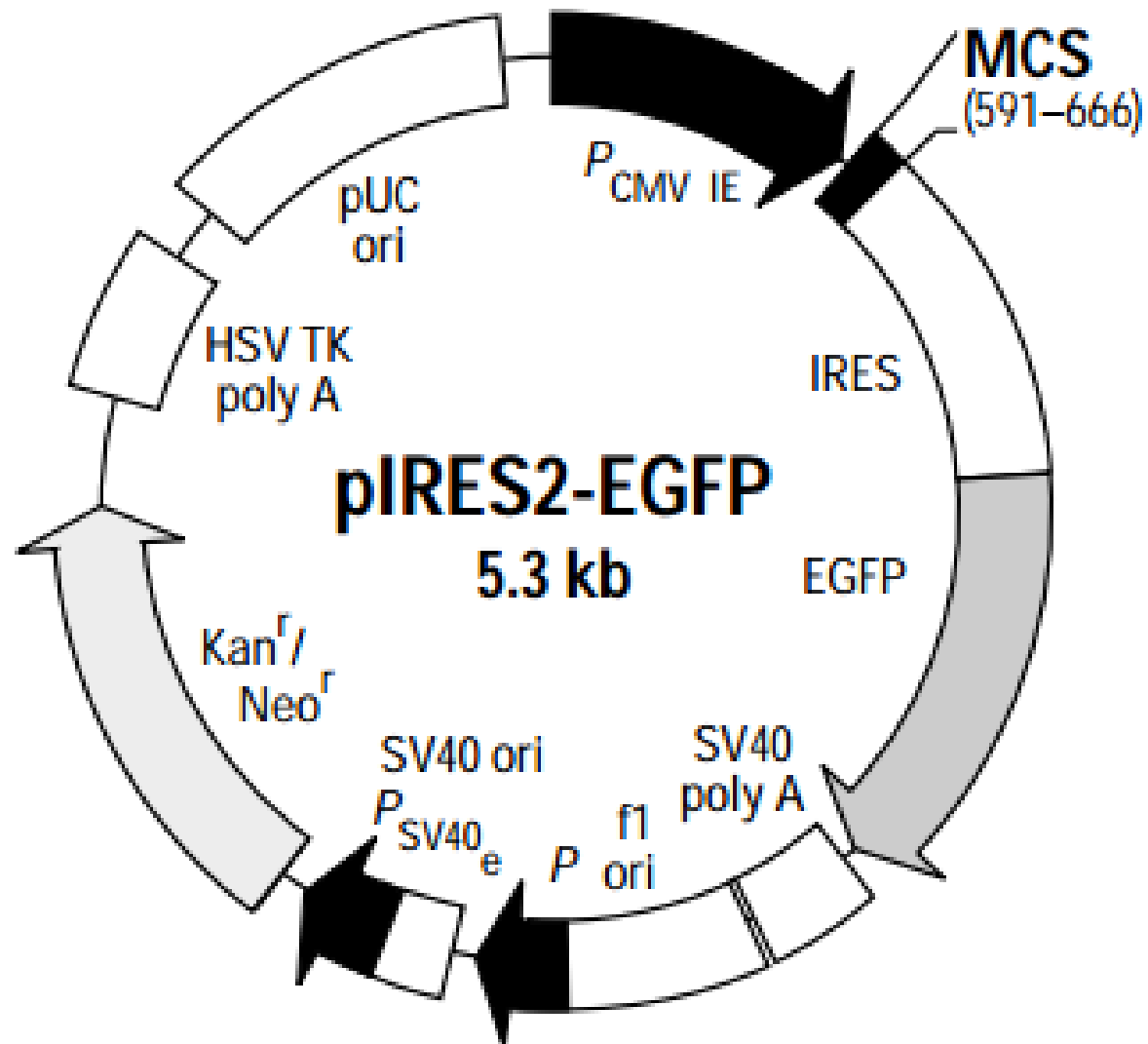
 400 410 420 430 440 450 460

...S N S S P V R P D E L T S R C A H L S P T I M T M

← *lacZα* translational start

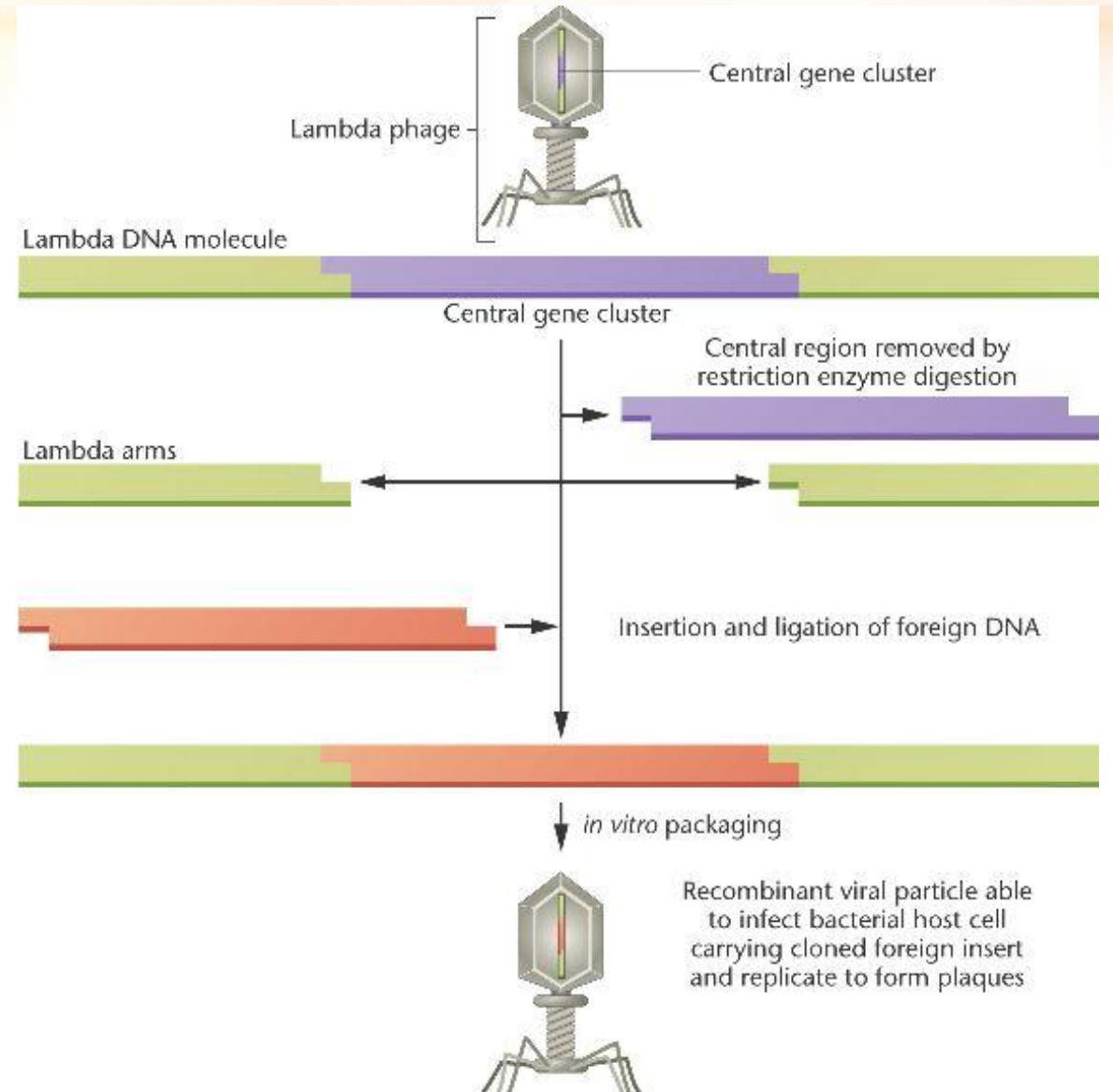
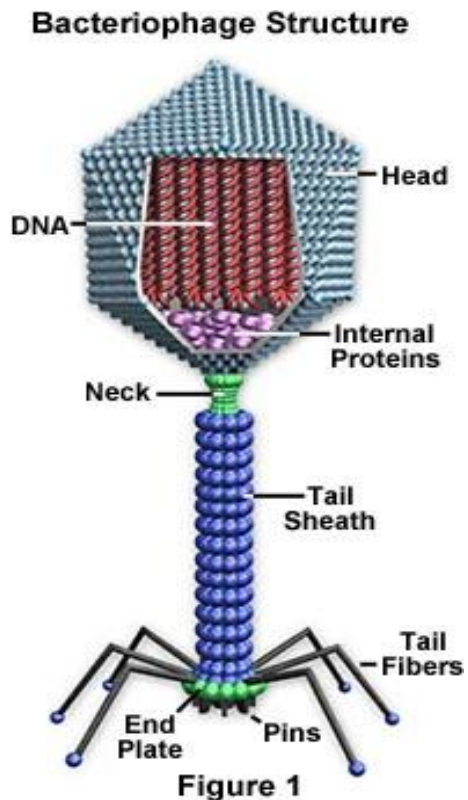
The diagram shows the DNA sequence of the pUC19 Multi-Cloning Site (MCS) region. Above the sequence, restriction enzyme sites are indicated: EcoRI (AATTC), SacI (GAGCTC), KpnI (GTAC), SmaI (ACCGGG), BamHI (GATC), XbaI (CTAGA), SalI (TCGAC), SbfI (CAGGCATG), PstI (GCAAGCTT), SphI (TGG), and HindIII (cgtaatc). The amino acid sequence is shown below the DNA sequence, with the Met (M) residue at the end highlighted in a blue box. An arrow labeled 'lacZα translational start' points to the Met residue, indicating that any gene cloned into this site will be translated from that point.

Expresní plasmid pIRES2-eGFP



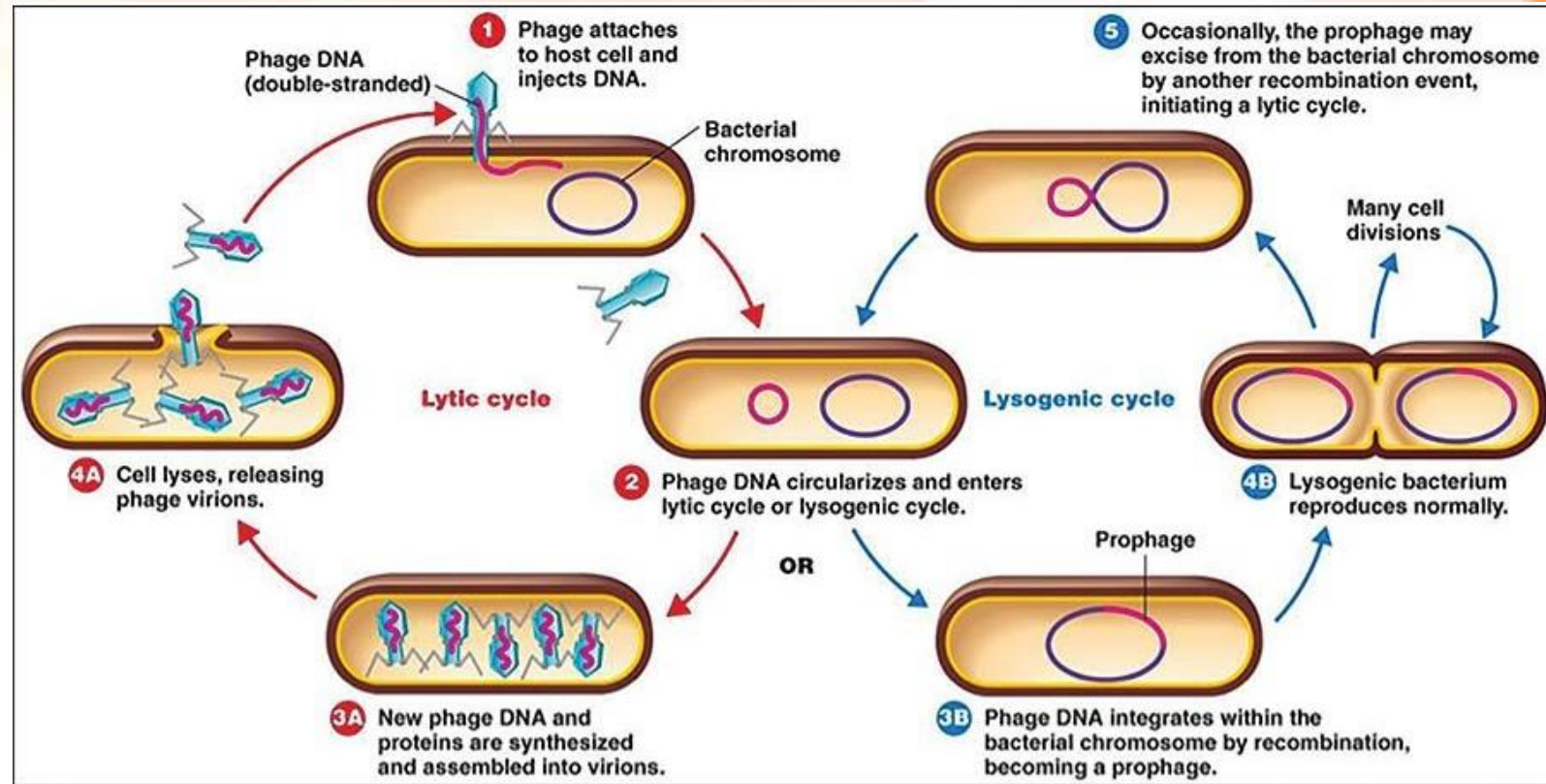
Vektory na bázi λ bakteriofágu

- nahrazují plasmidy je-li třeba klonovat delší fragmenty DNA
- **Inzerční vektory 8 – 10 kbp**
- **Výměnné vektory 8 – 24 kbp**

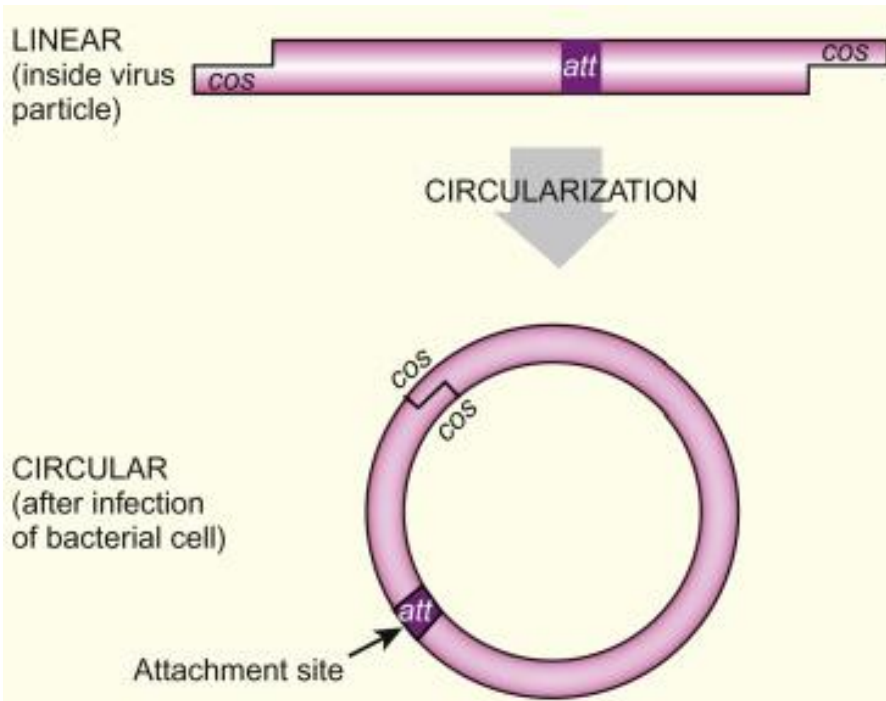


Bakteriofág λ

- 50 kbp dsDNA
- lineární a kružnicová forma
- místa cos



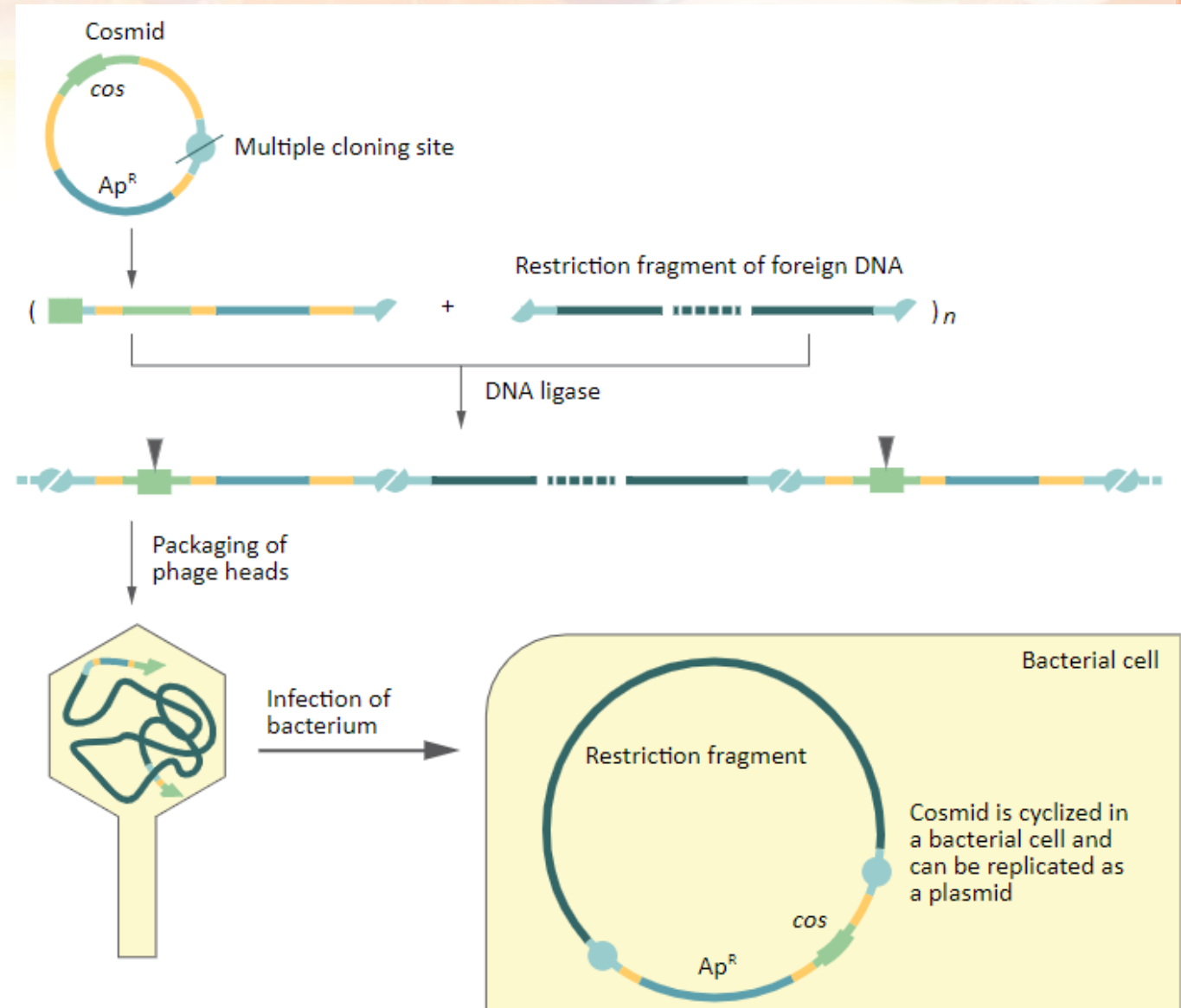
<http://hotcore.info/babki/bacteriophage-lysogenic-cycle.html>



<https://doi.org/10.1016/C2009-0-01986-2>

Kosmidy

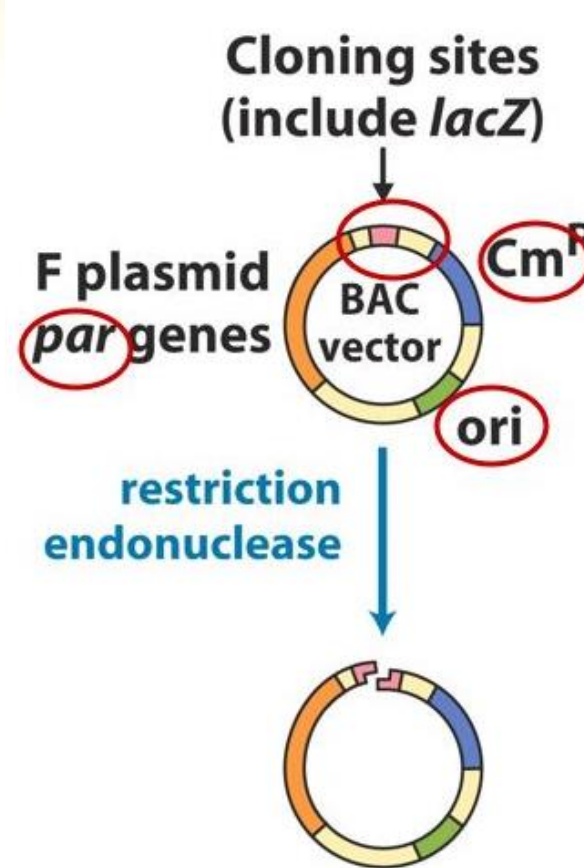
- kombinace plasmidu a fága
- prokaryotický počátek replikace $oriV$
- selekční znak
- klonovací místo
- kapacita 37 - 52 kbp
- místa cos bakteriofága λ
- pro sbalení nutno přidat sbalovací proteiny
- vstupuje jako fág
- v buňce se chová jako plasmid



Bakteriální umělý chromosom

- BAC = bacterial artificial chromosome
- Odvozeny od plasmidu F'
- Určeny ke klonování do bakteriálních buněk
- Vyskytují se v počtu 1-2 kopie na buňku
- Klonovaná DNA je vysoce stabilní
- **Klonovací kapacita až 300 kbp**
(možná i více)

- **Využity v projektu HUGO**
- **Dnes nahrazovány metodami celogenomového sekvenování, next generation sekvenování a sekvenování třetí generace**



Artificial Chromosomes allow for cloning of large pieces of DNA

Bacterial Artificial Chromosome

1. *ori* allows for replication in bacteria,
2. *par* helps segregate BAC evenly between daughter cells,
3. *lac Z* allows for detection of insert,
4. Cm^R allows for selection of transformed cells.

<https://library.uams.edu/assets/COM/BioChem/MolecularTools/MolecularToolsSDL12.html>

Další varianty vektorů

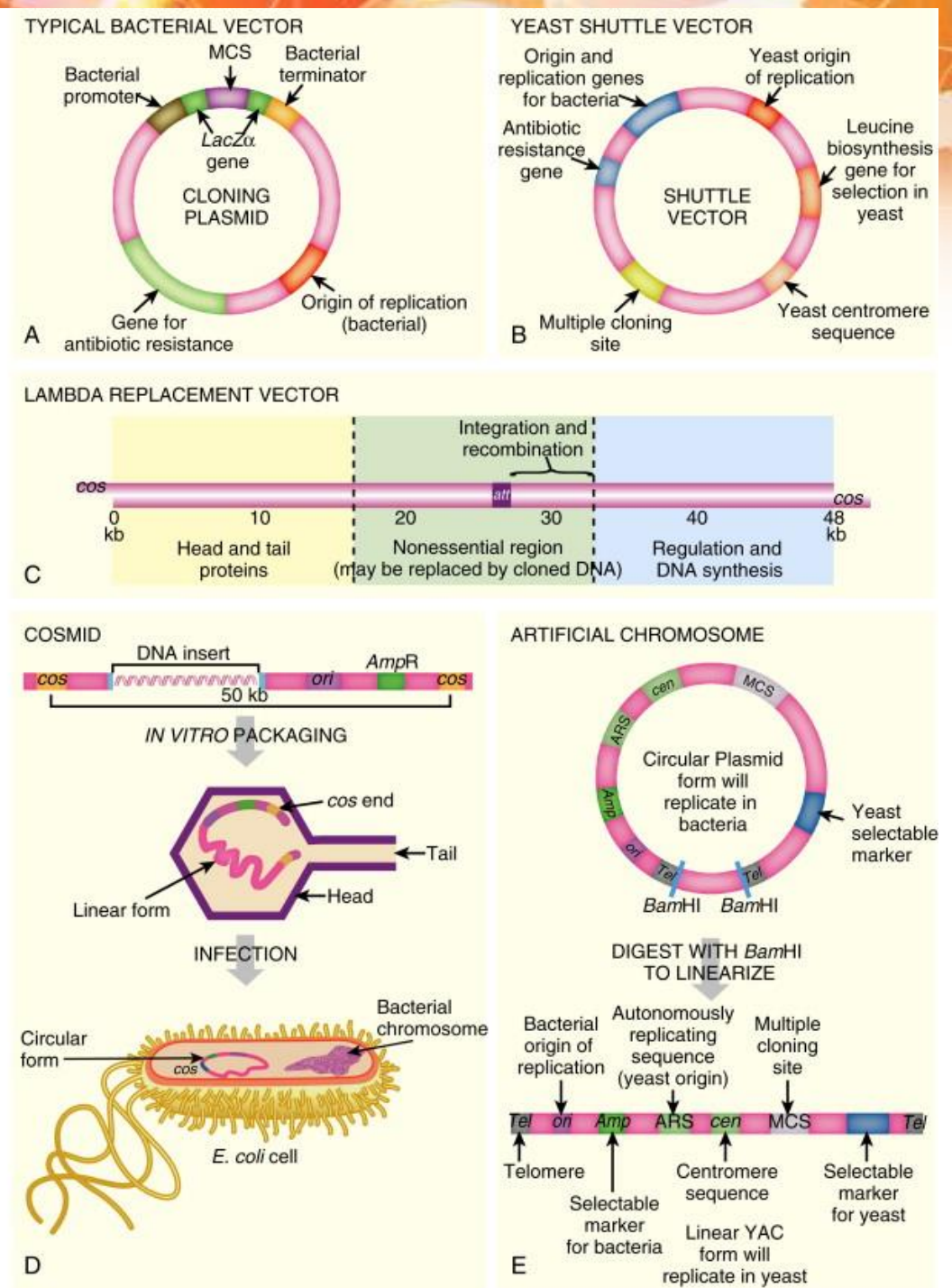
Fosmid (phosmid)

- Podobný kosmidu, ale je založen na bakteriálním F-plasmidu
- Hostitel (*E. coli*) může obsahovat jen jedinou molekulu
- Klonovací kapacita do 40 kbp
- Vhodné ke konstrukci stabilních knihoven z komplexních genomů
- Vysoce stabilní; schopné udržet lidskou DNA po více než 100 generací

Fagemid (phagemid)

- Plasmid obsahující počátek replikace fága M13. Používá se k přípravě ssDNA.
- Nejznámějšími příklady jsou série klonovacích vektorů pBluescript.

(A) Typical bacterial cloning vector. This vector has bacterial sequences to initiate replication and transcription. In addition, it has a multiple cloning site embedded within the *lacZ* α gene so that the insert can be identified by alpha-complementation. The antibiotic resistance gene allows the researcher to identify any *E. coli* cells that have the plasmid. (B) Yeast shuttle vector. This vector can survive in either bacteria or yeast because it has both yeast and bacterial origin of replication, a yeast centromere, and selectable markers for yeast and bacteria. As with most cloning vectors, there is a polylinker. (C) Lambda replacement vectors. Because lambda phage is easy to grow and manipulate, its genome has been modified to accept foreign DNA inserts. The region of the genome shown in green is nonessential for lambda growth and packaging. This region can be replaced with large inserts of foreign DNA (up to about 23 kb). (D) Cosmids. Cosmids are small multicopy plasmids that carry *cos* sites. They are linearized and cut so that each half has a *cos* site (not shown). Next, foreign DNA is inserted to relink the two halves of the cosmid DNA. This construct is packaged into lambda virus heads and used to infect *E. coli*. (E) Artificial chromosomes. Yeast artificial chromosomes have two forms: a circular form for growing in bacteria and a linear form for growing in yeast. The circular form is maintained like any other plasmid in bacteria, but the linear form must have telomere sequences to be maintained in yeast. The linear form can hold up to 2000 kb of cloned DNA and is very useful for genomics research.



Hostitelé = příjemci rekombinantní DNA

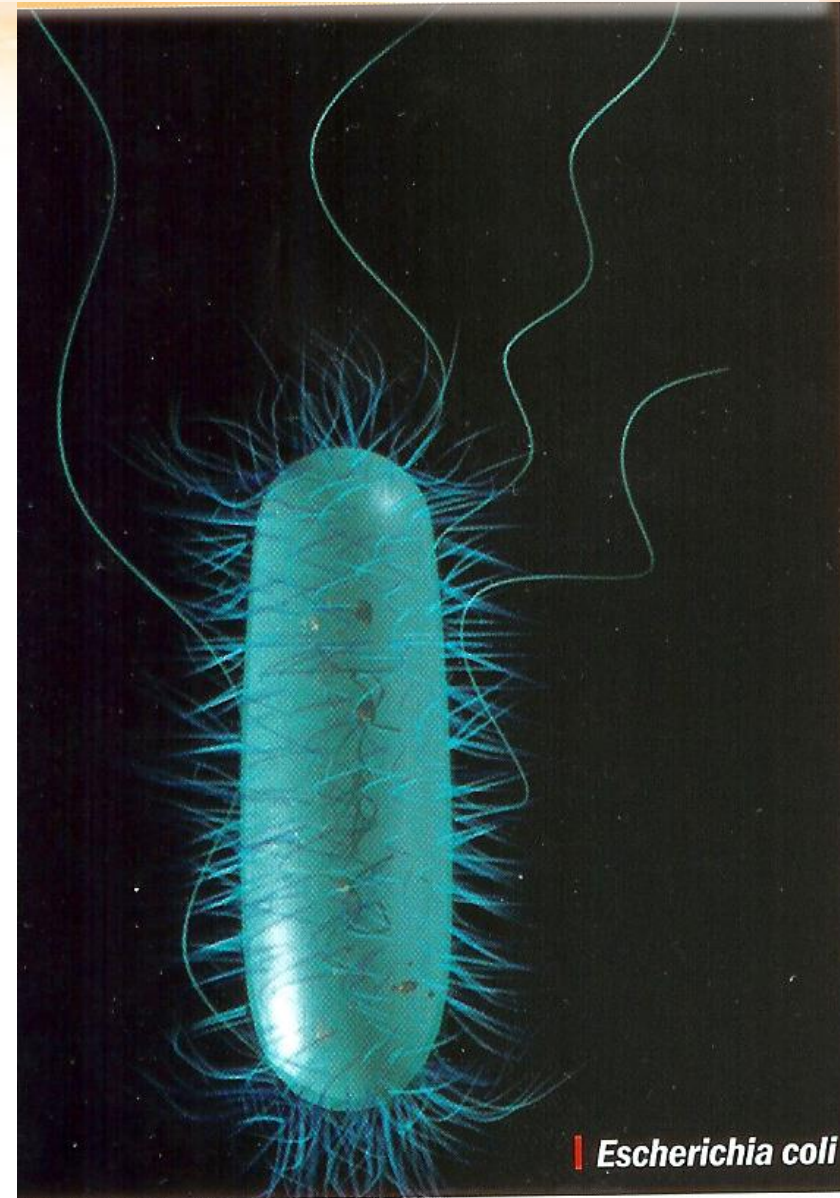
Bez ohledu na typ zdroje

- bakteriální buňky
- kvasinky a plísně
- rostlinné i živočišné buňky
- celá rostlina nebo živočich

Escherichia coli

- G- bakterie, kružnicový chromozóm 3×10^6 bp
- množství použitelných plasmidů
- generační doba 20 min. → rychlá tvorba biomasy
- nenáročná a levná kultivace
- stacionární fáze 2×10^9 b/mL
- řada mutantů (DH5α, HB101, BL21,...)

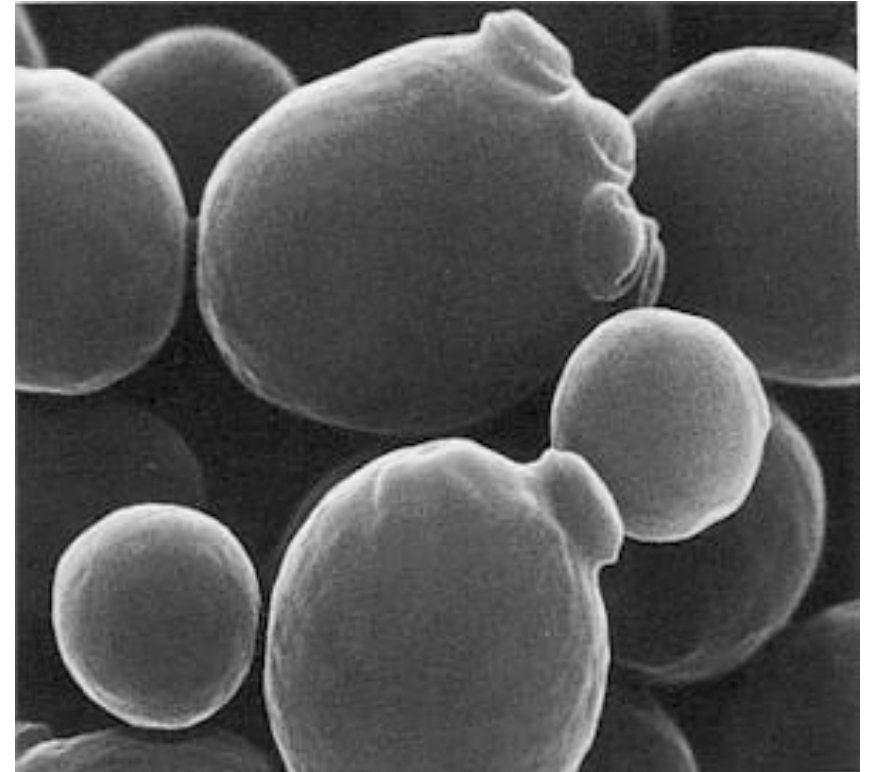
- **nevýhody** – výrazně odlišné posttranslační modifikace oproti eukaryotům



Escherichia coli

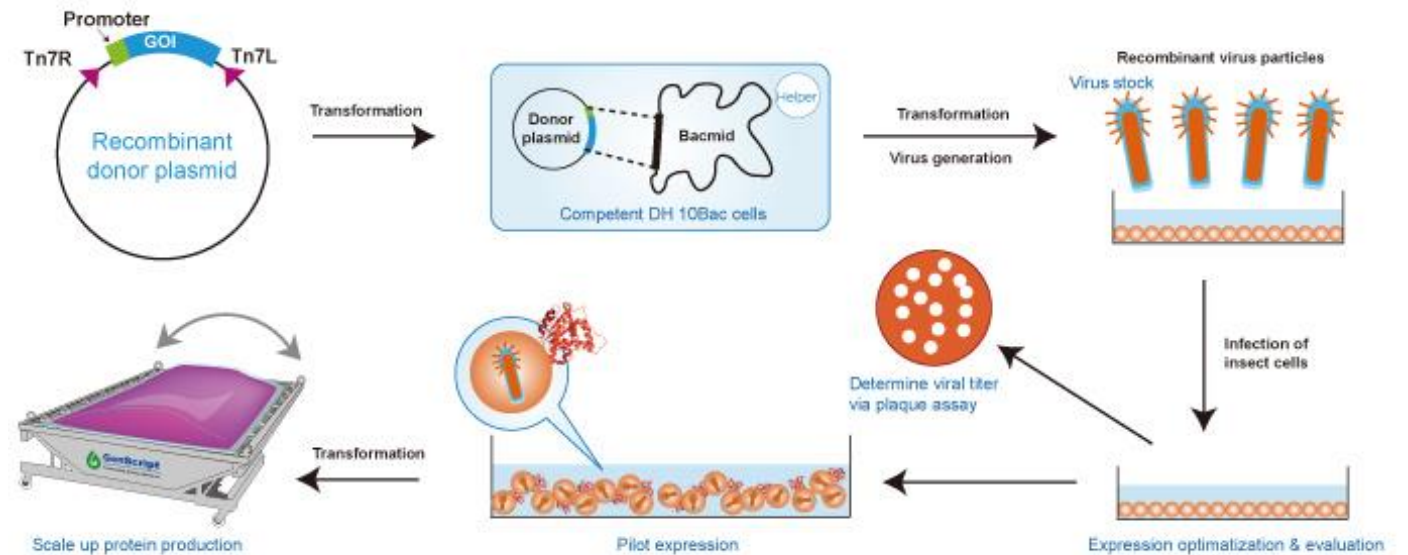
Saccharomyces cerevisiae

- lineární chromozómy
- přibližně 13×10^6 bp
- asi 6 275 genů
- nejjednodušší eukaryotický organismus
- shodný transkripční a translační aparát s dalšími eukaryoty
- rozdíly v posttranslačních procesech, např. **hyperglykosylace manosou**
 - *Schizosaccharomyces pombe*
 - *Pichia pastoris*



Hmyzí buňky a baculoviry

- podobné selekční principy jako u nižších eukaryot či bakterií
- využívání binárních „shuttle“ vektorů (bakterie + hmyz)
- náhrada polyhedrinové sekvence viru rekombinantním genem
- nejčastěji používaná buněčná kultura Sf9 odvozená z *Spodoptera frugiperda*
- **baculoviry napadají pouze hmyzí buňky**



Rostlinné buňky

- vektory jsou typické bakteriální plasmidy, které obsahují rostlinné expresní kazety
- transformace přímá
- transformace pomocí *Agrobacterium*
- transformace virovými vektory
- hostitelské buňky – *Nicotiana tabacum*, *Arabidopsis thaliana*, ...

Relativně bezpečná technologie



Savčí buňky a jejich viry

- člověku nejbližší systémy
- nejběžnějším producentem jsou savčí buňky CHO (Chinese hamster ovary)
- **rozdílná mezidruhová glykosylace**
- jako vektory adenoviry, retroviry, herpesviry



Otázky bezpečnosti technologie !

Tvorba rekombinantních proteinů v různých organismech

Parameter	Bacteria	Yeast	Mammalian cell culture	Transgenic plants
Glycosylation	None	Incorrect	Correct	Generally correct
Assembles multimeric proteins	Limited	Limited	Limited	Yes
Production costs	Medium	Medium	High	Low
Protein-folding accuracy	Low	Medium	High	High
Protein yield	High	High	Medium	Medium
Scale-up costs	High	High	High	Low
Time required	Low	Low	High	Medium
Skill level required for growth	Medium	Medium	High	Low