

lidský genom

jádro:

3.2 miliardy párů bází DNA
(2x)

mitochondrie:

16.5 kb kruhová DNA

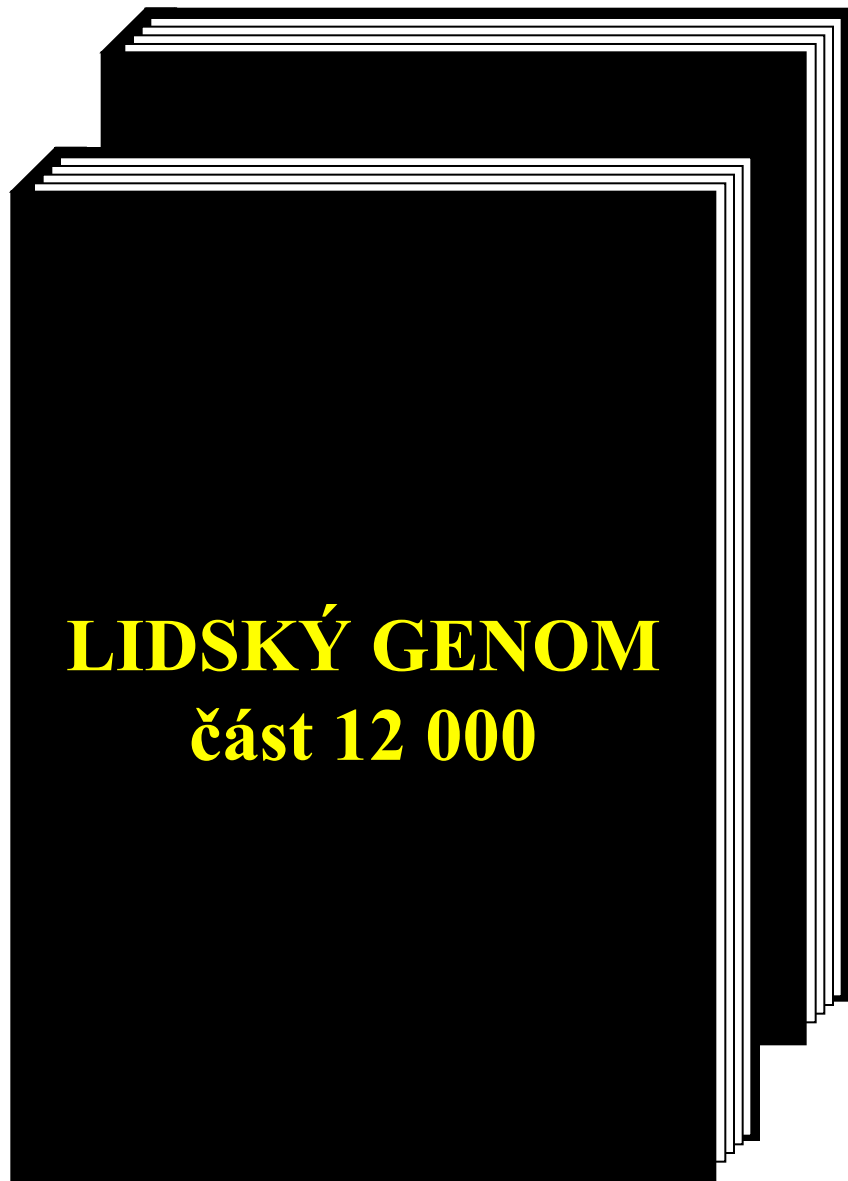
(několik g/mt, stovky mt/b)

geny pro 24 RNA + 13 proteinů

```
TGCTTGGATCGATGGCCCTTAGCTAGCTCCTTTTAGATCTAGTCGGGCTAGCTATGCTTG  
GATCGATGGCCCTTAGCTAGCTCCTTTTAGATCTAGTCGGGCTAGCTATGCTTGGATCGA  
TGGCCCTTAGCTAGCTCCTTTTAGATCTAGTCGGGCTAGCTATGCTTGGATCGATGGCC  
TTAGCTAGCTCCTTTTAGATCTAGTCGGGCTAGCTATGCTTGGATCGATGGCCCTTAGCT  
AGCTCCTTTTAGATCTAGTCGGGCTAGCTATGCTTGGATCGATGGCCCTTAGCTAGCTCC  
TTTTAGATCTAGTCGGGCTAGCTATGCTTGGATCGATGGCCCTTAGCTAGCTCCTTTTAG  
ATCTAGTCGGGCTAGCTATGCTTGGATCGATGGCCCTTAGCTAGCTCCTTTTAGATCTAG  
TCGGGCTAGCTATGCTTGGATCGATGGCCCTTAGCTAGCTCCTTTTAGATCTAGTCGGG  
TAGCTATGCTTGGATCGATGGCCCTTAGCTAGCTCCTTTTAGATCTAGTCGGGCTAGCTA  
TGCTTGGATCGATGGCCCTTAGCTAGCTCCTTTTAGATCTAGTCGGGCTAGCTATGCTTG  
TGCTTGGATCGATGGCCCTTAGCTAGCTCCTTTTAGATCTAGTCGGGCTAGCTATGCTTG  
GATCGATGGCCCTTAGCTAGCTCCTTTTAGATCTAGTCGGGCTAGCTATGCTTGGATCGA  
TGGCCCTTAGCTAGCTCCTTTTAGATCTAGTCGGGCTAGCTATGCTTGGATCGATGGCC  
TTAGCTAGCTCCTTTTAGATCTAGTCGGGCTAGCTATGCTTGGATCGATGGCCCTTAGCT  
AGCTCCTTTTAGATCTAGTCGGGCTAGCTATGCTTGGATCGATGGCCCTTAGCTAGCTCC  
TTTTAGATCTAGTCGGGCTAGCTATGCTTGGATCGATGGCCCTTAGCTAGCTCCTTTTAG  
ATCTAGTCGGGCTAGCTATGCTTGGATCGATGGCCCTTAGCTAGCTCCTTTTAGATCTAG  
TCGGGCTAGCTATGCTTGGATCGATGGCCCTTAGCTAGCTCCTTTTAGATCTAGTCGGG  
TAGCTATGCTTGGATCGATGGCCCTTAGCTAGCTCCTTTTAGATCTAGTCGGGCTAGCTA  
TGCTTGGATCGATGGCCCTTAGCTAGCTCCTTTTAGATCTAGTCGGGCTAGCTATGCTTG  
TGCTTGGATCGATGGCCCTTAGCTAGCTCCTTTTAGATCTAGTCGGGCTAGCTATGCTTG  
GATCGATGGCCCTTAGCTAGCTCCTTTTAGATCTAGTCGGGCTAGCTATGCTTGGATCGA  
TGGCCCTTAGCTAGCTCCTTTTAGATCTAGTCGGGCTAGCTATGCTTGGATCGATGGCC  
TTAGCTAGCTCCTTTTAGATCTAGTCGGGCTAGCTATGCTTGGATCGATGGCCCTTAGCT  
AGCTCCTTTTAGATCTAGTCGGGCTAGCTATGCTTGGATCGATGGCCCTTAGCTAGCTCC  
TTTTAGATCTAGTCGGGCTAGCTATGCTTGGATCGATGGCCCTTAGCTAGCTCCTTTTAG  
ATCTAGTCGGGCTAGCTATGCTTGGATCGATGGCCCTTAGCTAGCTCCTTTTAGATCTAG  
TCGGGCTAGCTATGCTTGGATCGATGGCCCTTAGCTAGCTCCTTTTAGATCTAGTCGGG  
TAGCTATGCTTGGATCGATGGCCCTTAGCTAGCTCCTTTTAGATCTAGTCGGGCTAGCTA  
TGCTTGGATCGATGGCCCTTAGCTAGCTCCTTTTAGATCTAGTCGGGCTAGCTATGCTTG
```

1 strana

1 800 znaků



1 kniha

150 stran

270 000 znaků

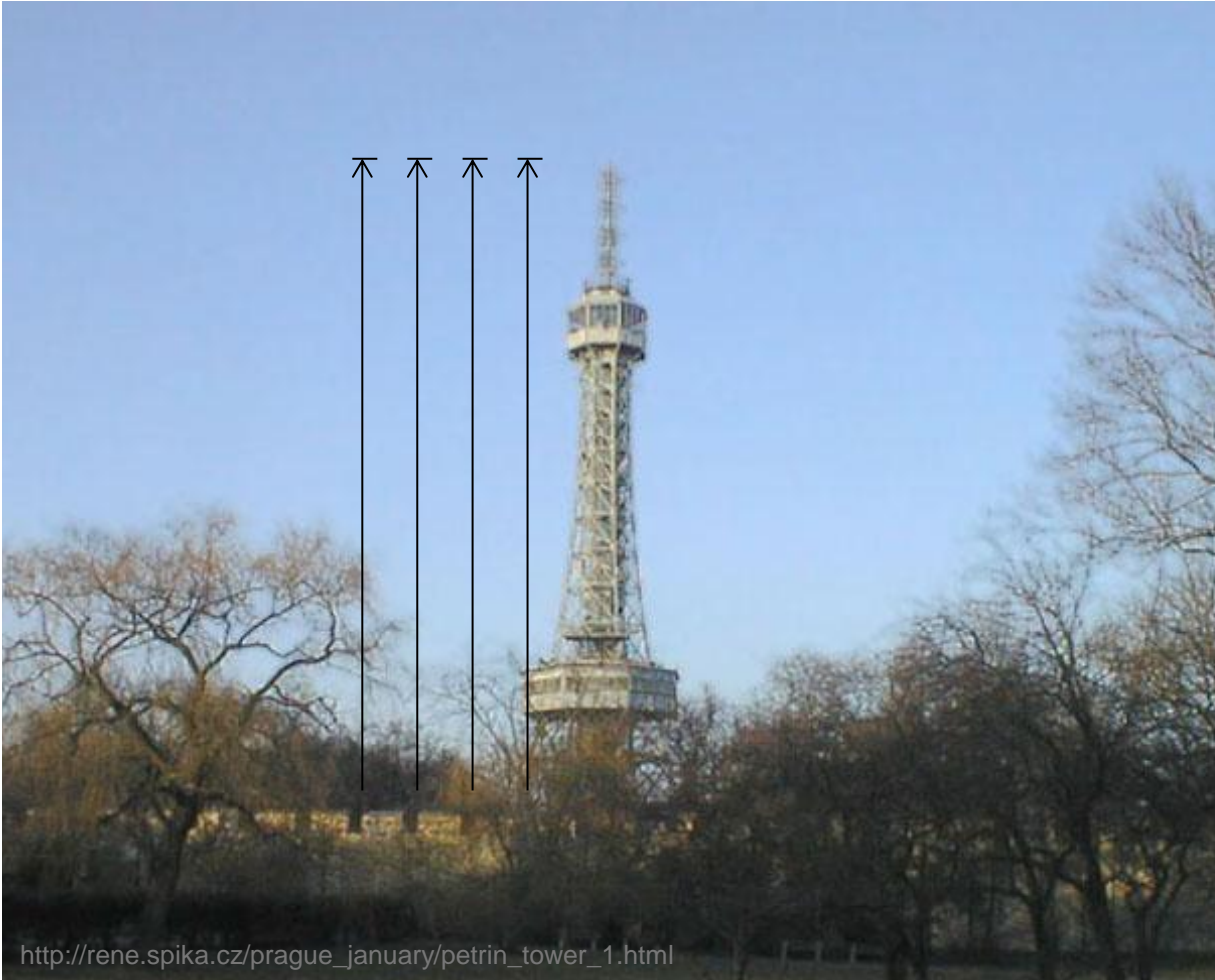
tloušťka 2 cm

celý genom

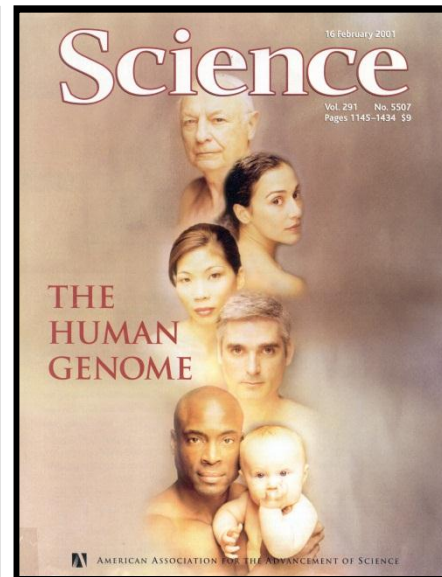
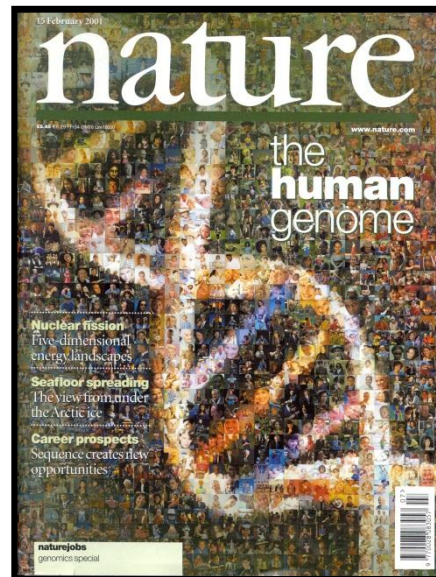
3.2 miliardy znaků

1.8 milionu stran

12 000 knih



NCBI: <http://www.ncbi.nlm.nih.gov/mapview/>
Sanger/EBI: <http://www.ensembl.org/>
UCSC: <http://genome.ucsc.edu>



Entrez Map View - Microsoft Internet Explorer

Soubor Úpravy Zobrazit Ovládné Nástroje Nápověda

NCBI

PubMed Entrez BLAST OMIM Taxonomy Structure

Search Find in This View Find Advanced Search

Homo sapiens Map View build 31 BLAST the Human Genome

Chromosome: 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 [X] Y

Master Map: Genes On Sequence **Maps & Options**

Total Genes On Chromosome: 1110
Region Displayed: 0-151M bp [Download/View Sequence/Evidence](#)
Genes Labeled: 20 Total Genes in Region: 1110

Contig Uni...	Genes_seq	symbol	orient.
Xp22-33			
Xp22-32			
Xp22-31			
Xp22-22			
Xp22-13			
Xp22-12			
Xp22-11			
Xp21-3			
Xp21-2			
Xp21-1			
Xp11-4			
Xp11-3			
Xp11-23			
Xp11-22			
Xp11-1			
Xp11-2			
Xq12			
Xq13-1			
Xq13-2			
Xq13-3			
Xq21-1			
Xq21-2			
Xq21-31			
Xq21-32			
Xq21-33			
Xq22-1			
Xq22-2			
Xq22-3			
Xq23			
Xq24			
Xq25			

symbol	orient.
ARSD	↑
NX17	↑
FLJ25735	↑
HYPM	↑
RP2	↓
GATA1	↓
VDAC1LP	↑
TRO	↓
ARR3	↓
MKRNP1	↓
FLJ30678	↑
ALEX1	↑
SERPINA7	↑
MDS031	↓
ZNF-kaiso	↓
OR1AA1P	↓

FEATURES

FEATURES	Location/Qualifiers
source	1..1251
	/organism="Mus musculus"
	/mol_type="mRNA"
	/db_xref="taxon:10090"
	/clone="MGC:58456 IMAGE:6770406"
	/tissue_type="Testis, mouse"
	/clone_lib="NIH_MGC_169"
	/lab_host="DH10B"
	/note="Vector: pDNR-LIB"
gene	1..1251
	/gene="DOH6S2654E"
	/note="synonyms: X5L, XAP-5-like"
	/db_xref="LocusID:108161"
CDS	149..1153
	/codon_start=1
	/product="DOH6S2654E protein"
	/protein_id="AAH49659.1"
	/db_xref="GI:29436863"
	/db_xref="LocusID:108161"
	/translation="MAQYKGTMRREAGRAMHL IKKREKQKEQMEVLKQRIAEETIMKSK VDKKFSAHYDAVEAELKSSVTVGLVTLNDMKAKQEAALLREREMQLAKREOLEQRRIQLE MLREKERRRRKRKISNLSFTLDEEEDQEDSRQAESAEVHSAAGAKKNLGRNPDVDTSL FLPDREEREENRRLREELRQEWAKREKVKGEVEITFSYWDGSGHRRTVRMSKGSTV QQFLKRALQGLRRDFRELRAAGVEQLMVVKEDLILPHYHTFYDFIVAKARGKSGPLFS FVDVHDDVRLLSDATMEKDESHAGKVVLRWSWYEKKNKHIFPASRWEPYDPEKKWDRYTI R"

ORIGIN

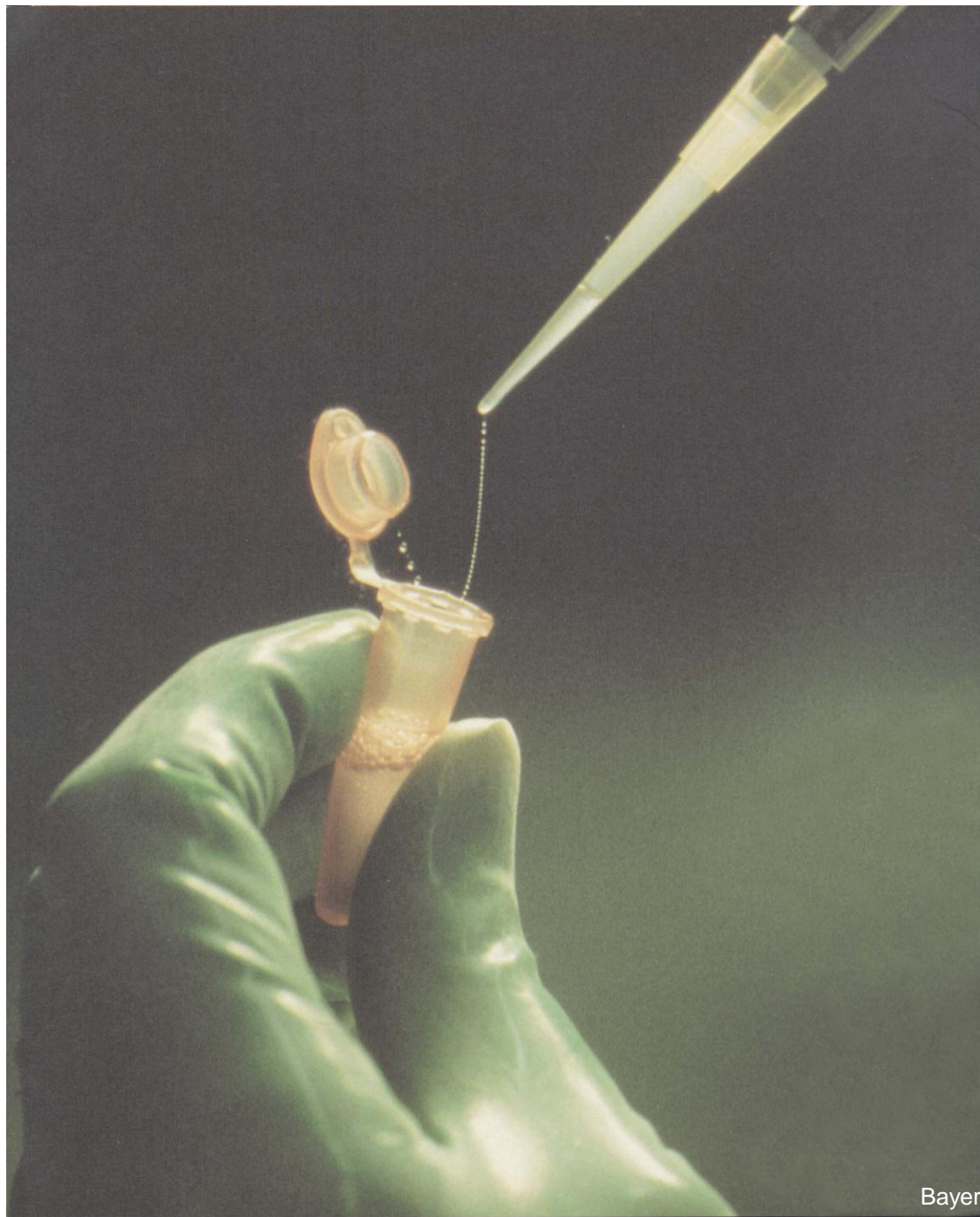
```

1 agaacaaca atattcgcaa aagcaccaga aggaagagtc ttggctcata catcaaaagc
61 tgcagaatct gtgaactgac atcagaccca gaaggctacc agaaacaggg actgggcagg
121 ccaaaaagcc ttgcgctgaa ctgcaggcat ggcgcagtac aaaggcacca tgcgggaagc
181 tggccgggcc atgcacctga tcaagaagcg tgagaagcag aaggagcaga tggaggtgct
241 gaagcagcgc atcgcagagg agaccatcat gaagtcaaaa ttggacaaga agttctcggc
301 aactactgac gccgtggagg ccgagctgaa gtccagtacg ttggcctgg tgaccctgaa
361 tgacatgaag gccaagcagg aggccctgct gagggagcgg gagatgcagc tggccaagag
421 ggagcagctg gagcaacgcc ggatacagct ggagatgctg cgcgagaagg agcgaaggcg
481 agagcgcaag cgcaagatct ccaacctgtc ttccagttg gacgaggaag aaggtgacca
541 agaggacagc cgccaagcgg agagtgccga ggtccacagt gctggagcca agaagaactt
601 gggcaagaat cccgatgtgg acacagagctt cctgcccgac cgcgagcgcg aggaggagga

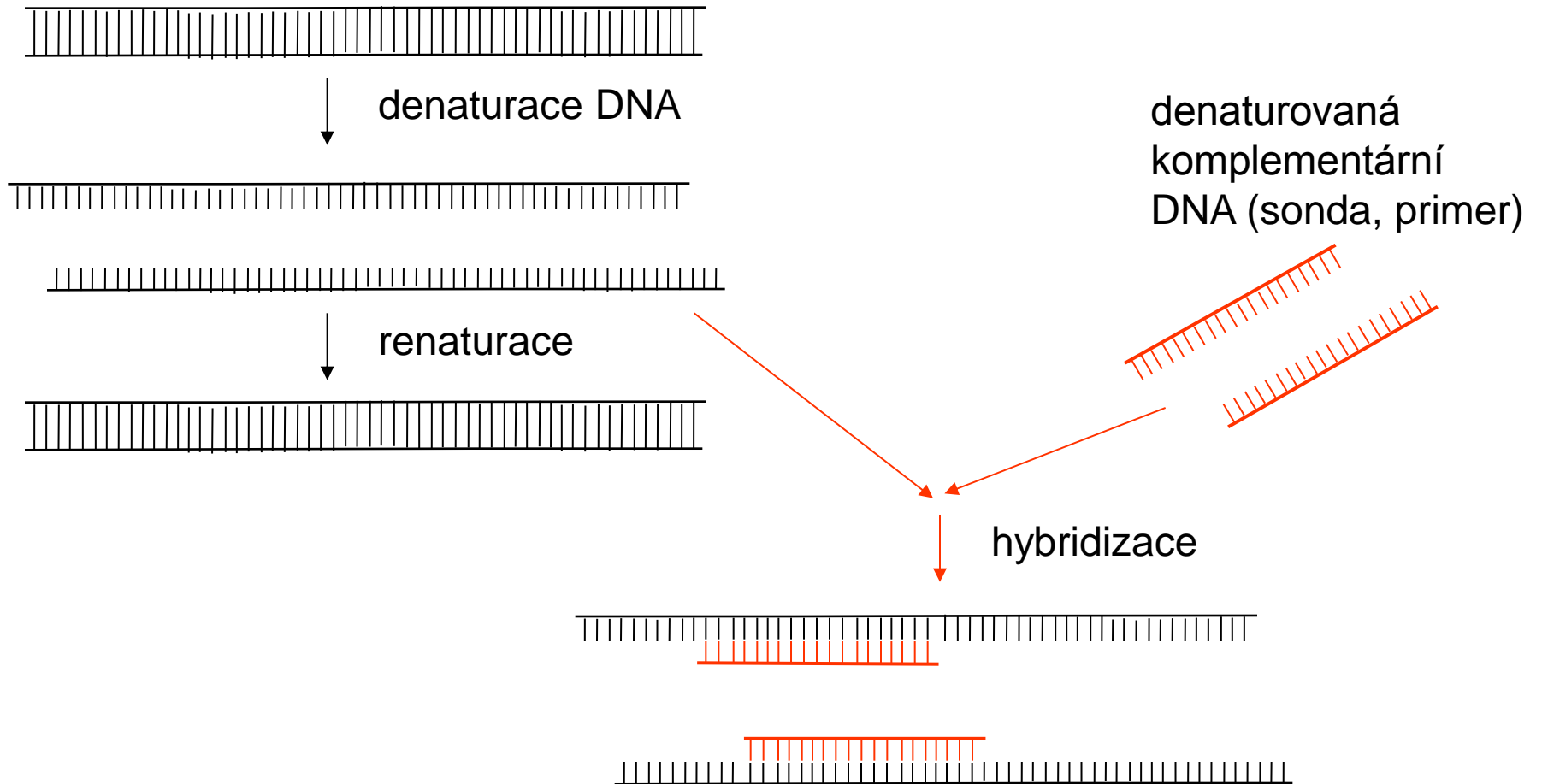
```

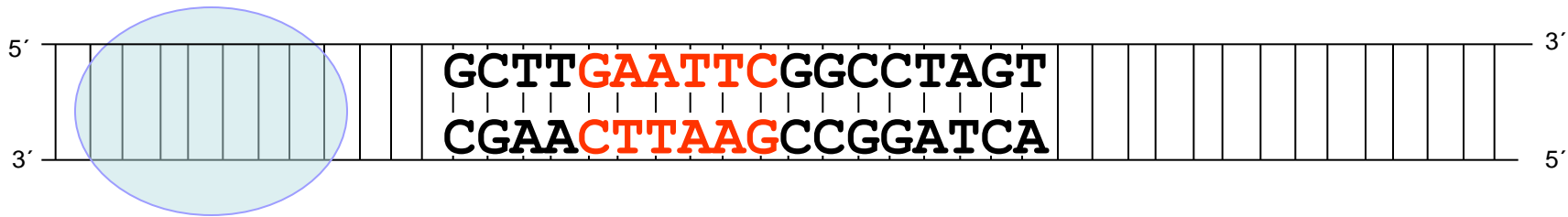
bioinformatika

izolace
DNA

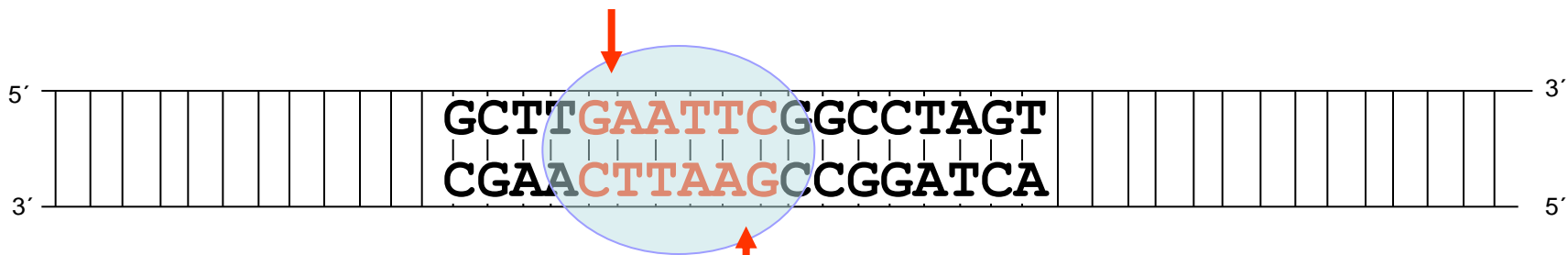


hybridizace DNA

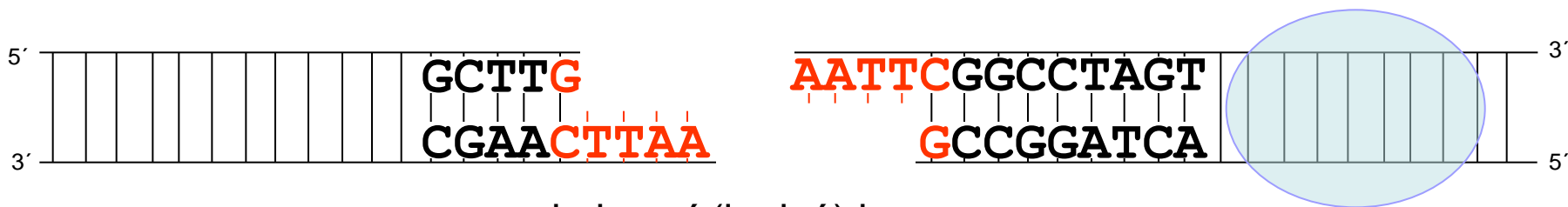




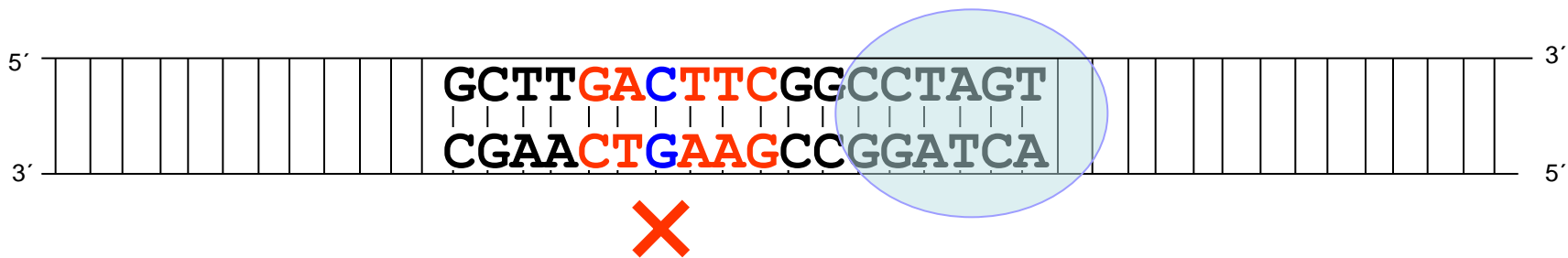
restrikční endonukleáza



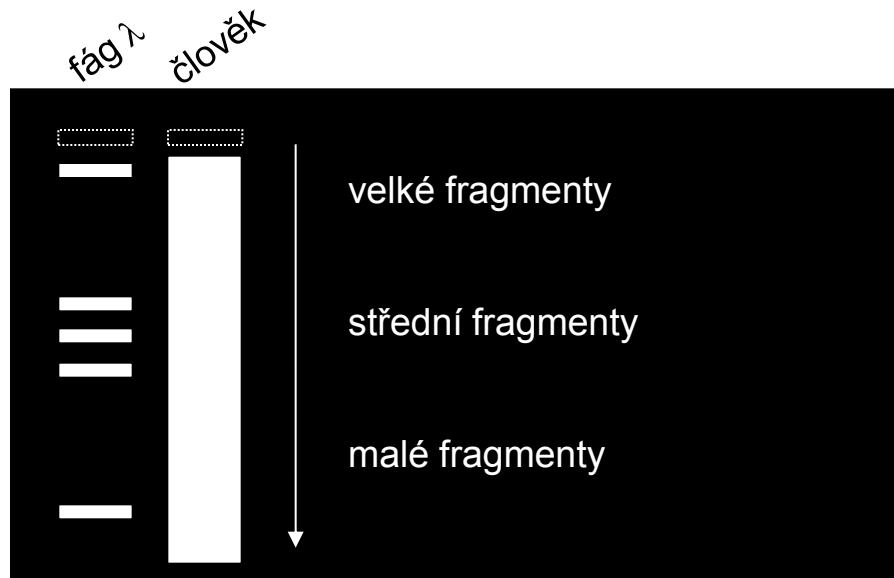
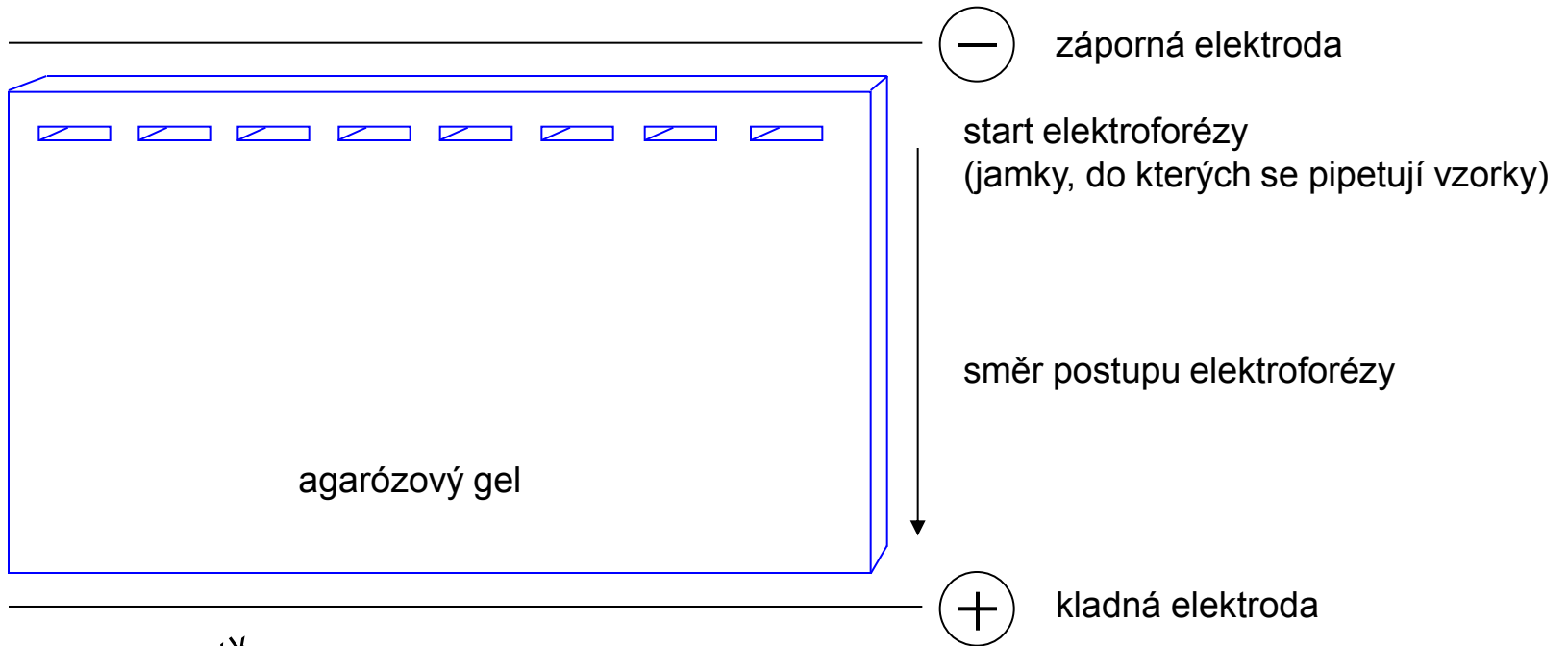
rozpoznávací (štěpné) místo **restrikční štěpení DNA**

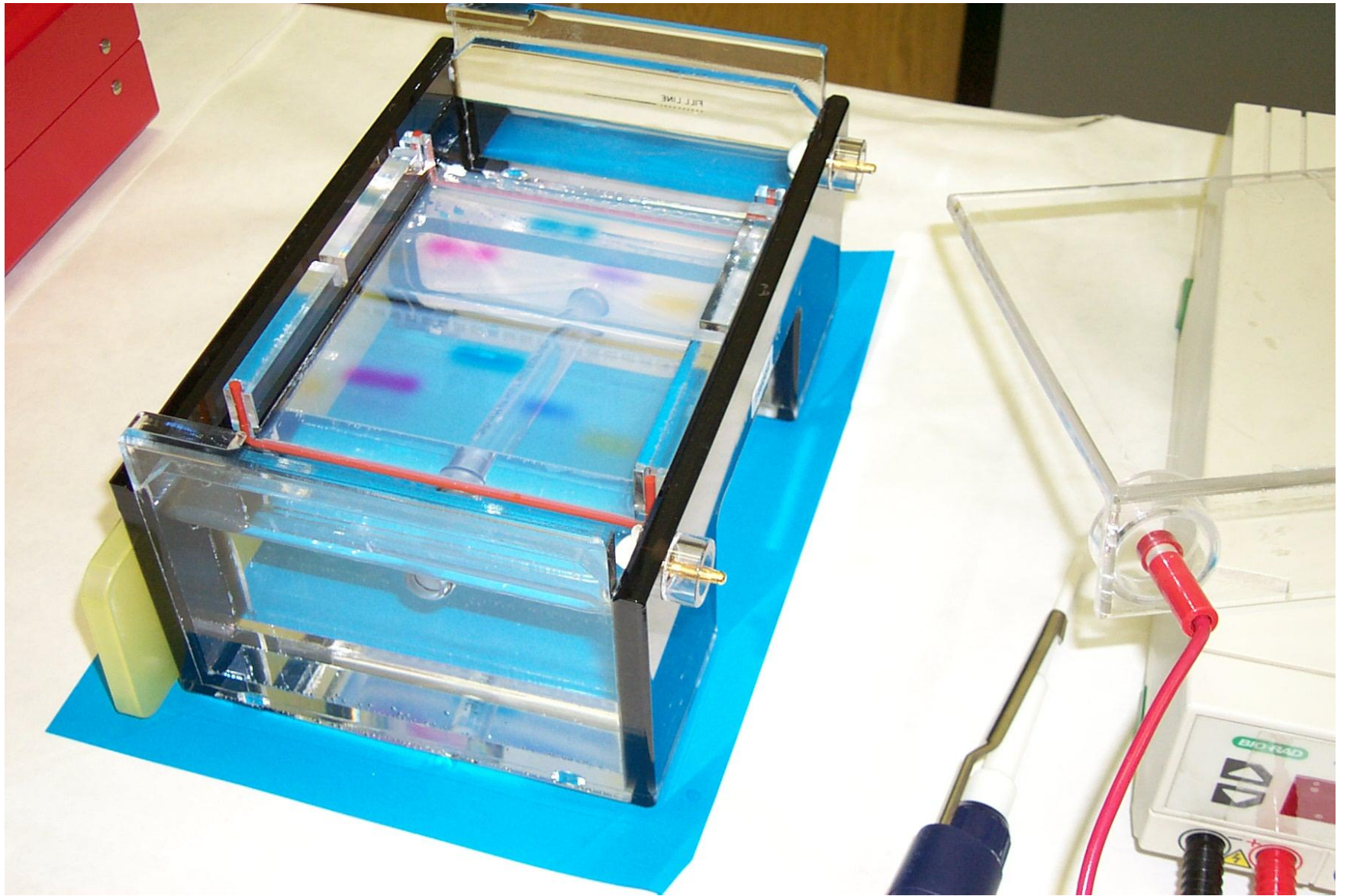


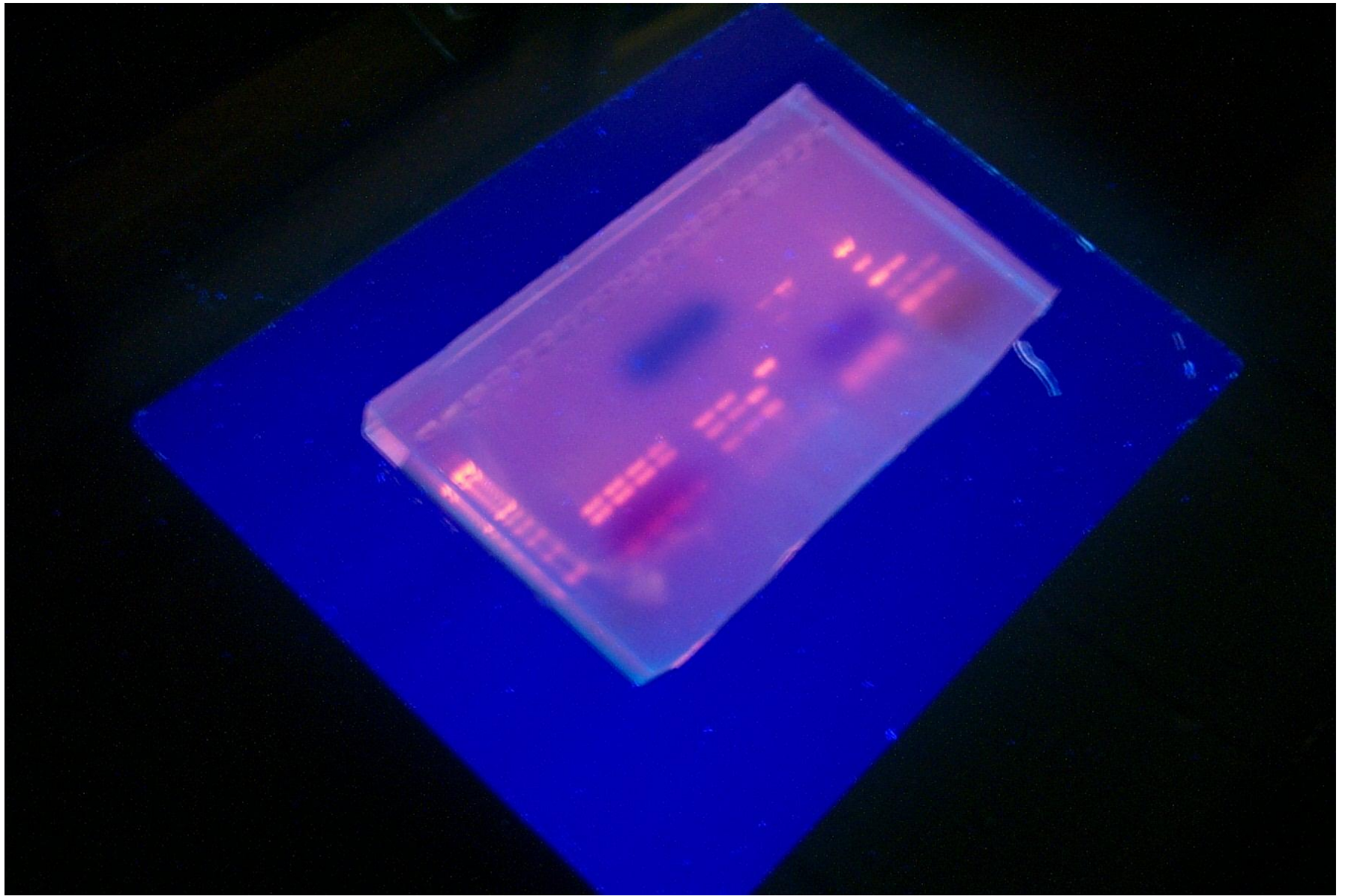
kohezní (lepivé) konce

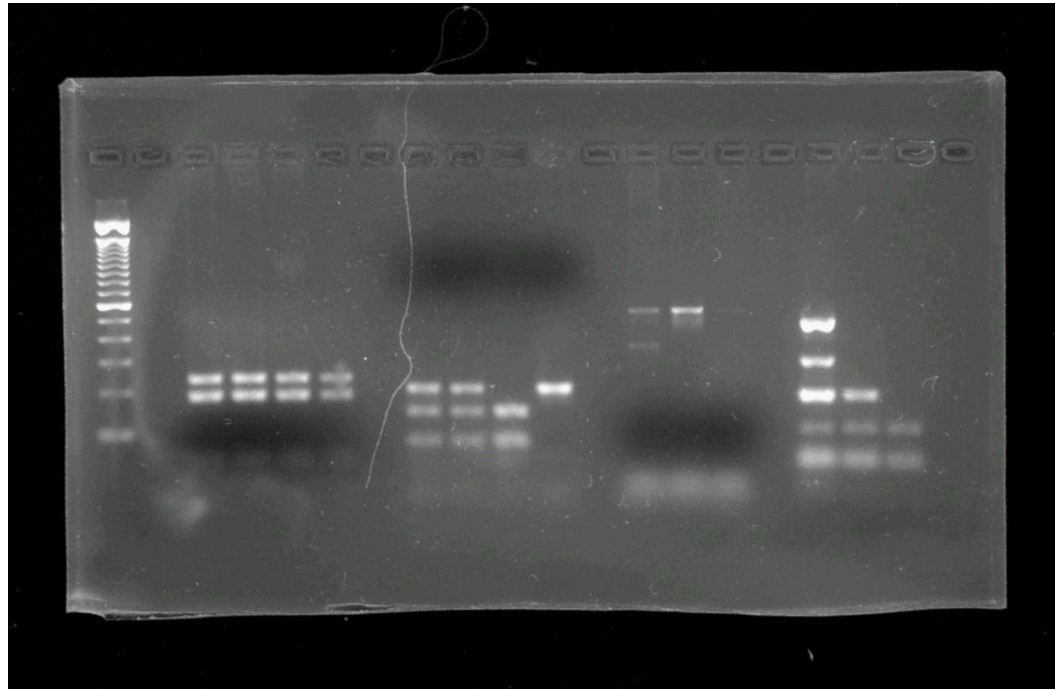


gelová elektroforéza DNA

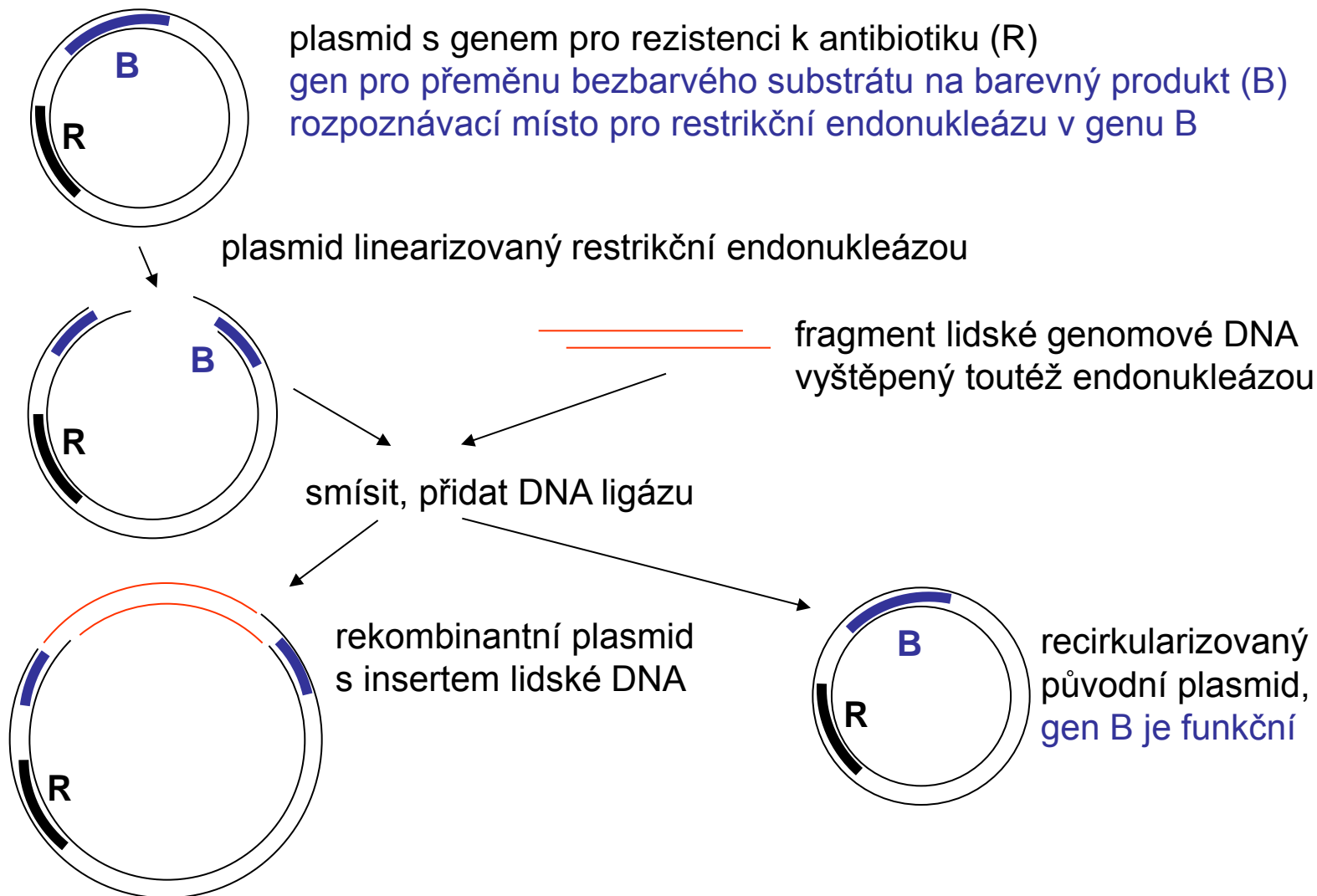








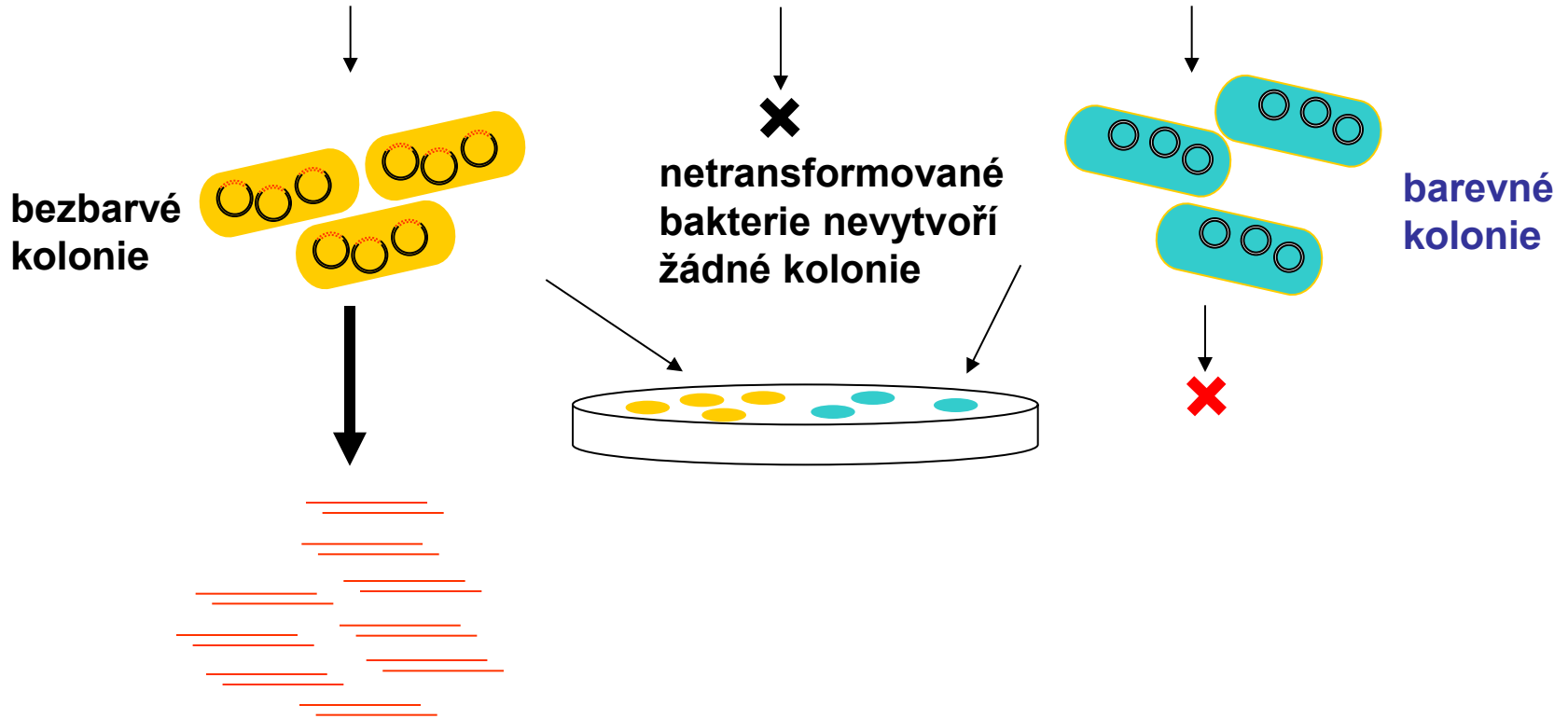
klonování DNA

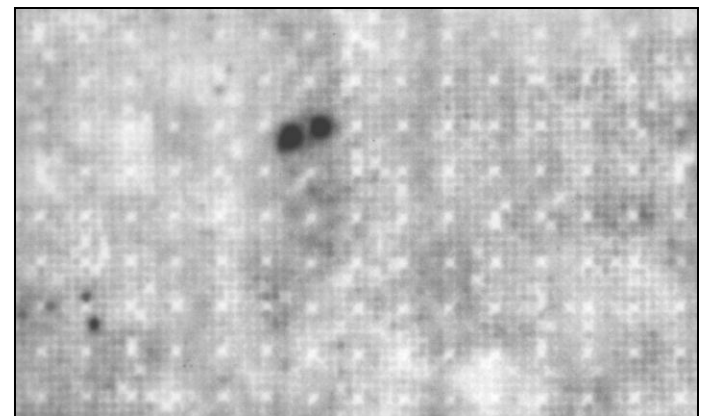
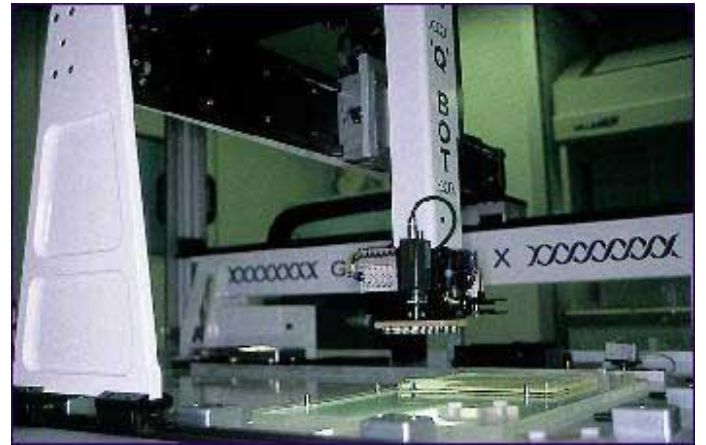
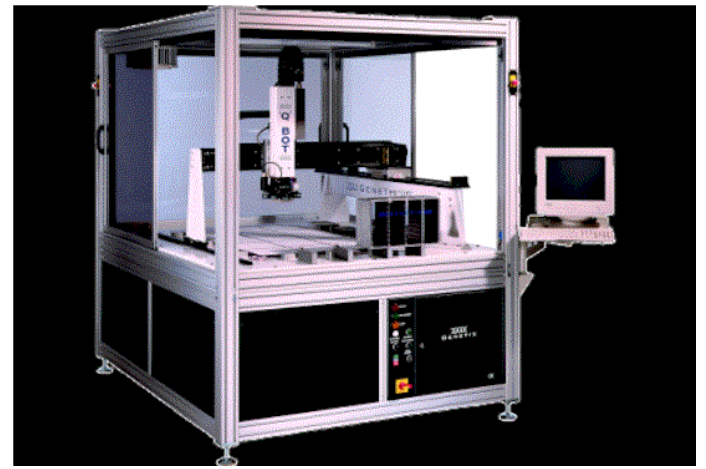
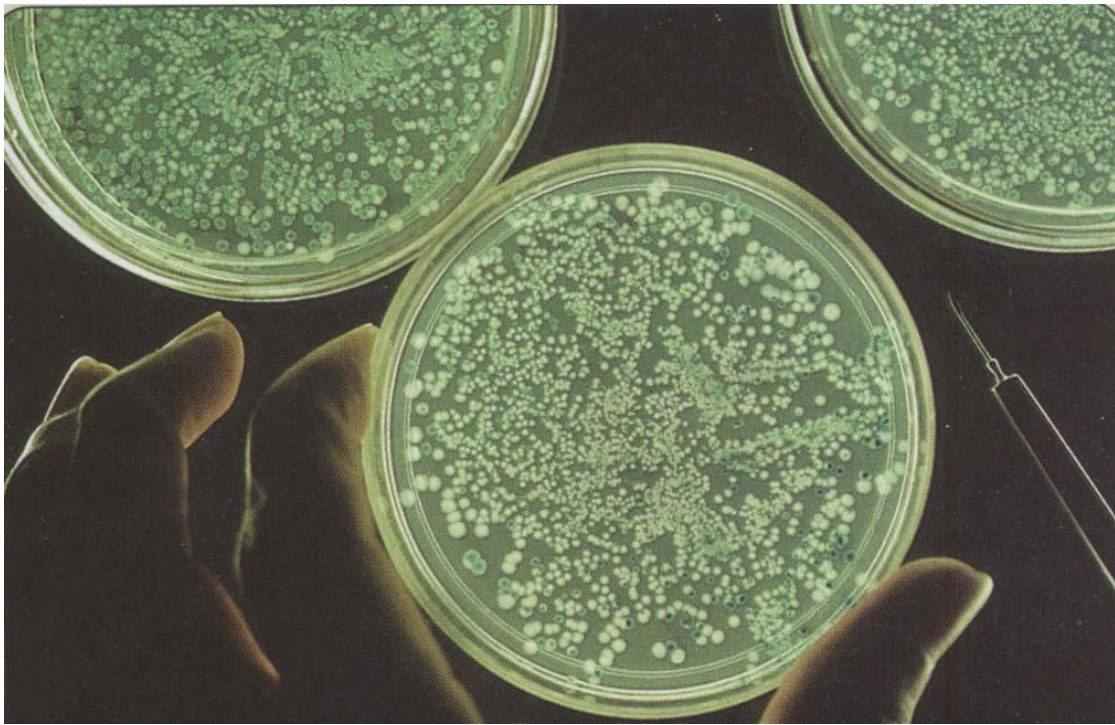


transformace hostitelských bakterií



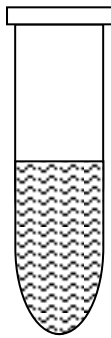
růst bakteriálních klonů na půdě s antibiotikem a substrátem





knihovny DNA

sekvenování DNA

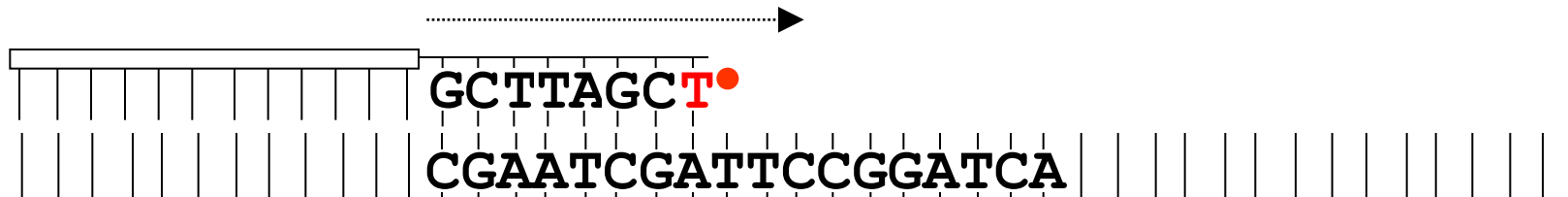


denaturovaná DNA

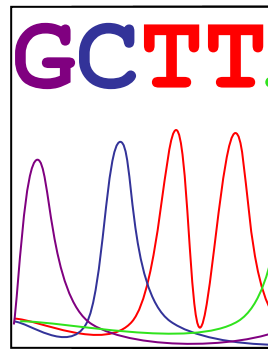
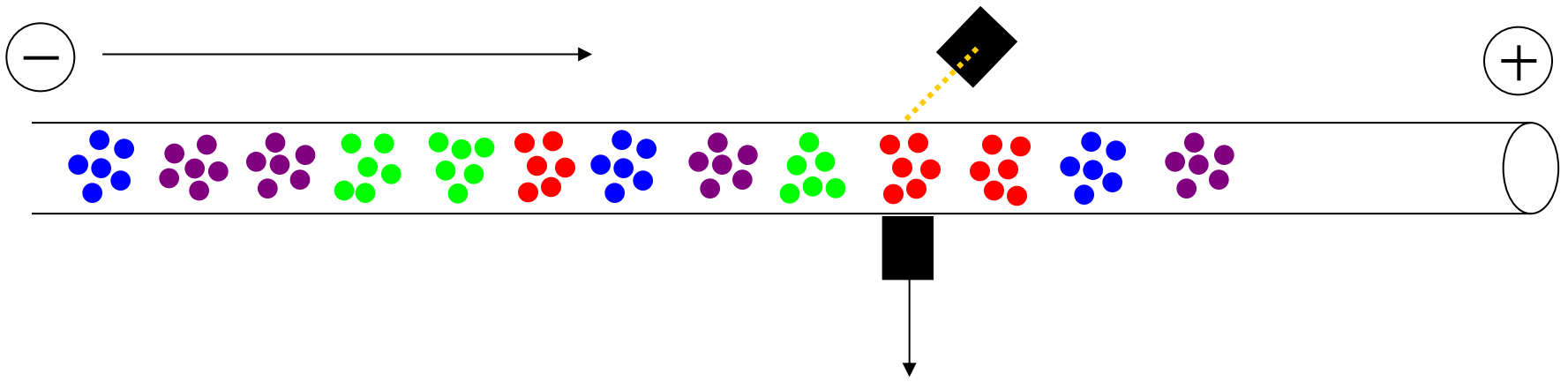
+ primer + polymeráza

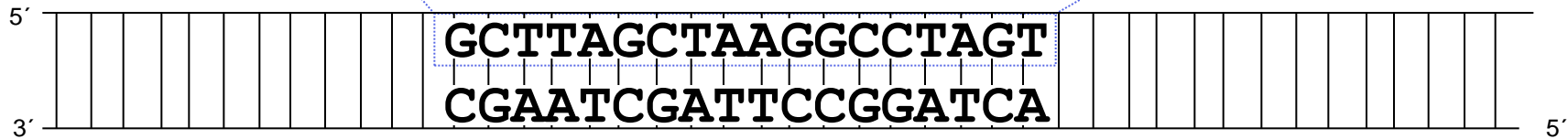
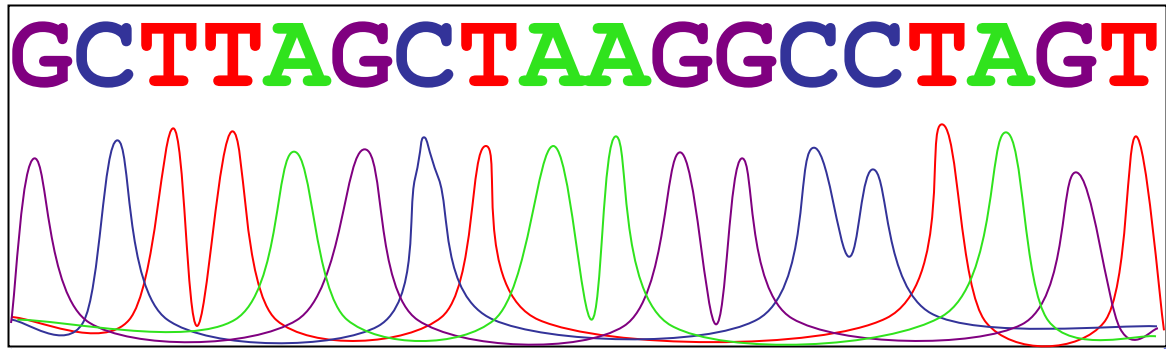
+ deoxynukleotidy G, A, T, C

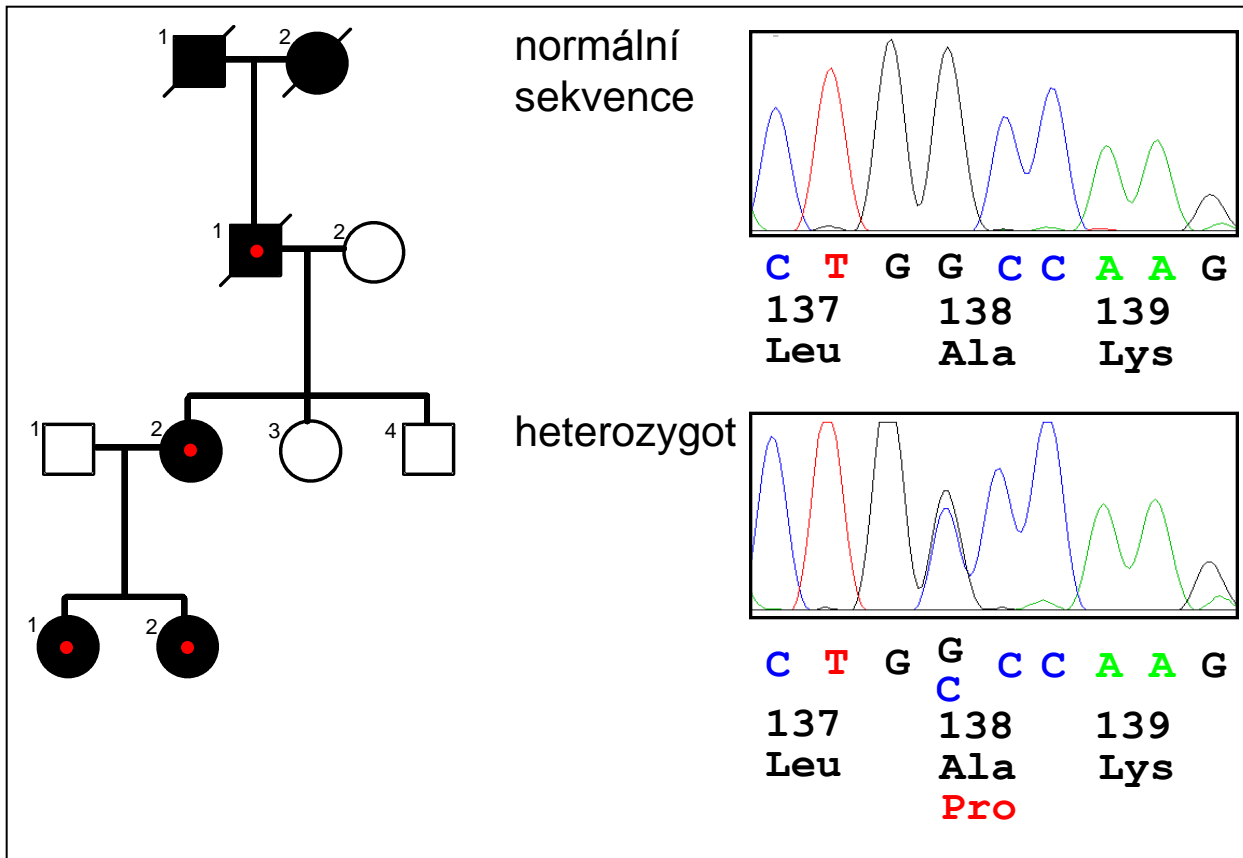
+ dideoxynukleotidy G[•], A[•], T[•], C[•]



GCTTAGCTAAGGCCTAG^T●
GCTTAGCTAAGGCCTAG^G●
GCTTAGCTAAGGCCTA^A●
GCTTAGCTAAGGCC^T●
GCTTAGCTAAGGCC^C●
GCTTAGCTAAGGC^C●
GCTTAGCTAAGG^G●
GCTTAGCTAAG^G●
GCTTAGCTAA^A●
GCTTAGCTA^A●
GCTTAGCT^T●
GCTTAGC^C●
GCTTAG^G●
GCTTA^A●
GCT^T●
GC^T●
GC^C●
G^G●







mezinárodní konsorcium

DNA 12 anonymních dárců

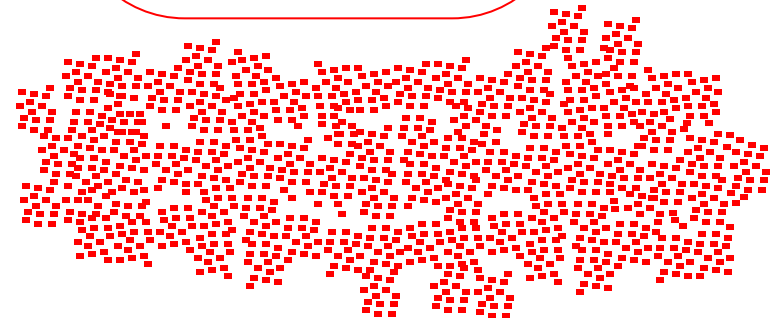


mapování



Celera Genomics

5 dárců různé etnicity



whole-genome shotgun



paired-end sequencing, scaffolds

K 10564



KEVIN
DAVIES

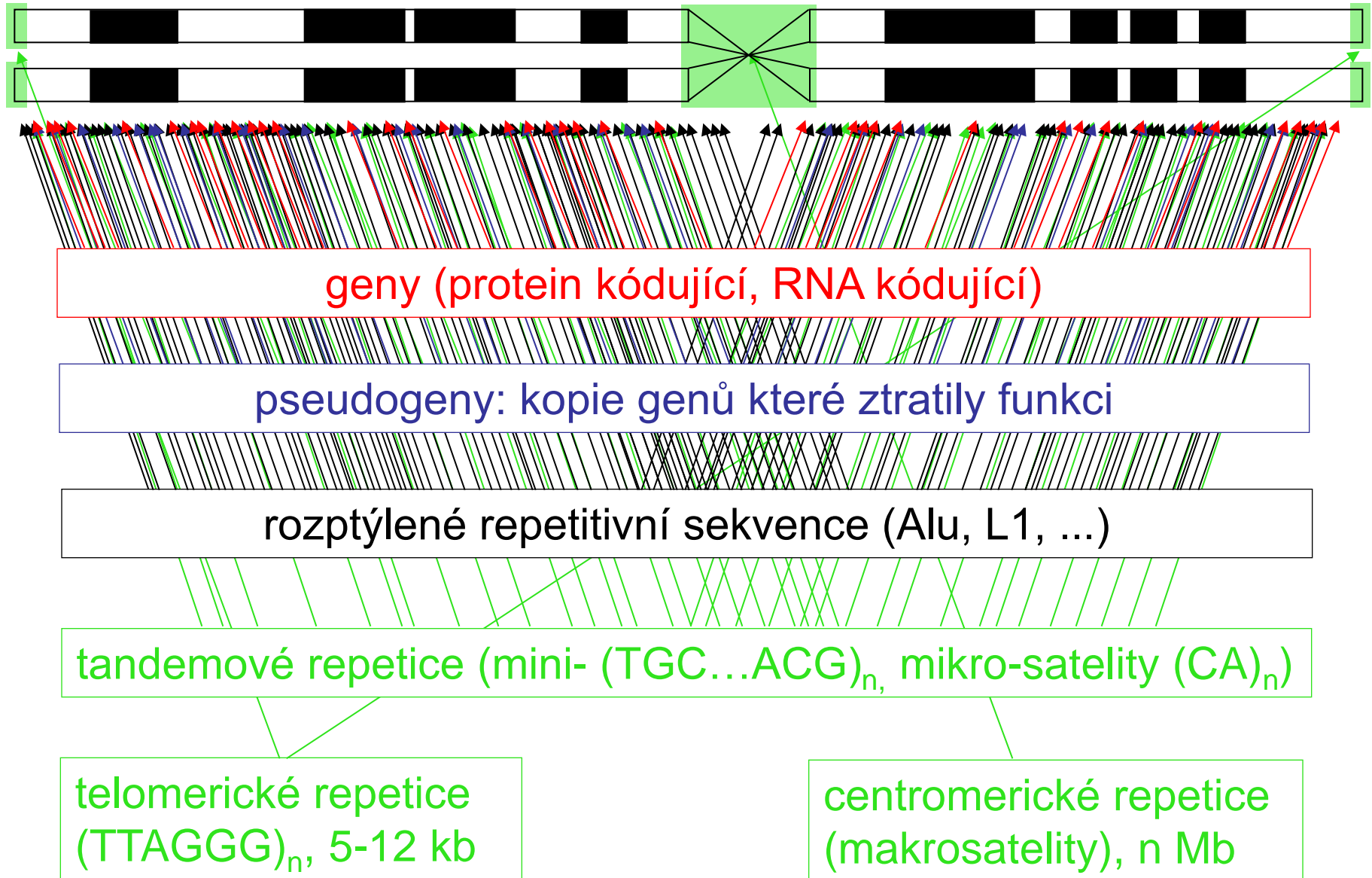
Příběh
největšího
vědeckého
objevu
naší doby

Rozluštěný genom

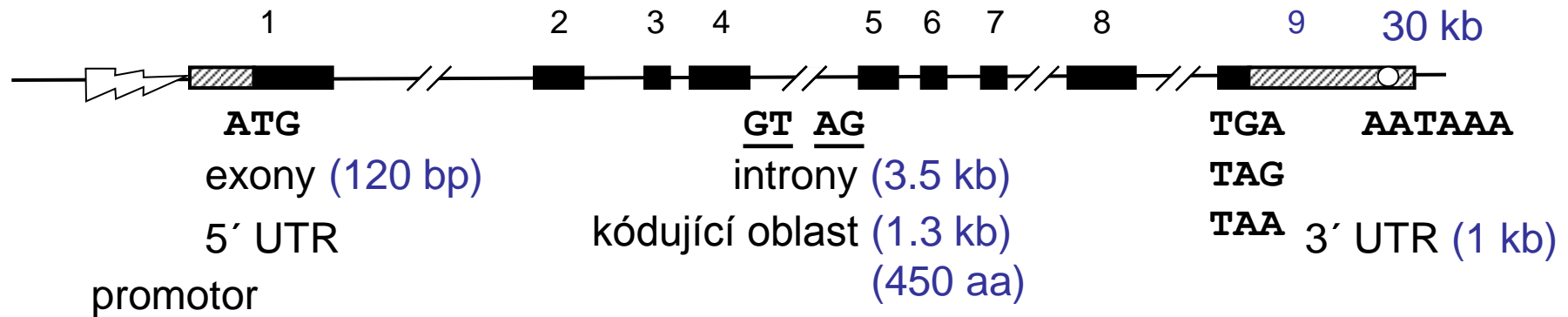
K A

F
E
N
I

typy sekvencí v lidském genomu



geny kódující proteiny



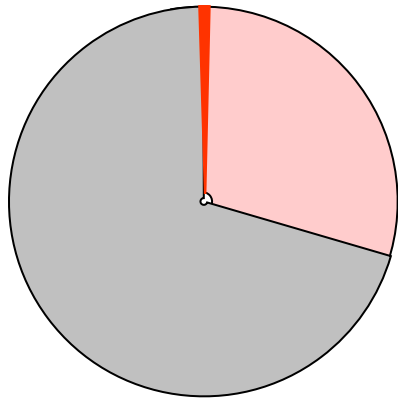
nejdelší gen: dystrofin (2.5 Mb, 89 exonů)
nejvíce exonů: titin (363 exonů)

genové rodiny

počet genů: méně než 25 tisíc

geny kódující nekódující RNA (ncRNA)

pseudogeny



geny (exony + introny) cca. 1/4

exony protein kódujících genů cca. 1.5%

mezigenová DNA



‘junk DNA’ 98.5%

podstata, původ a funkce ‘junk DNA’ ?

2500

nepravidelná distribuce genů v genomu



1



600



13



400



18

1500



19



300



21

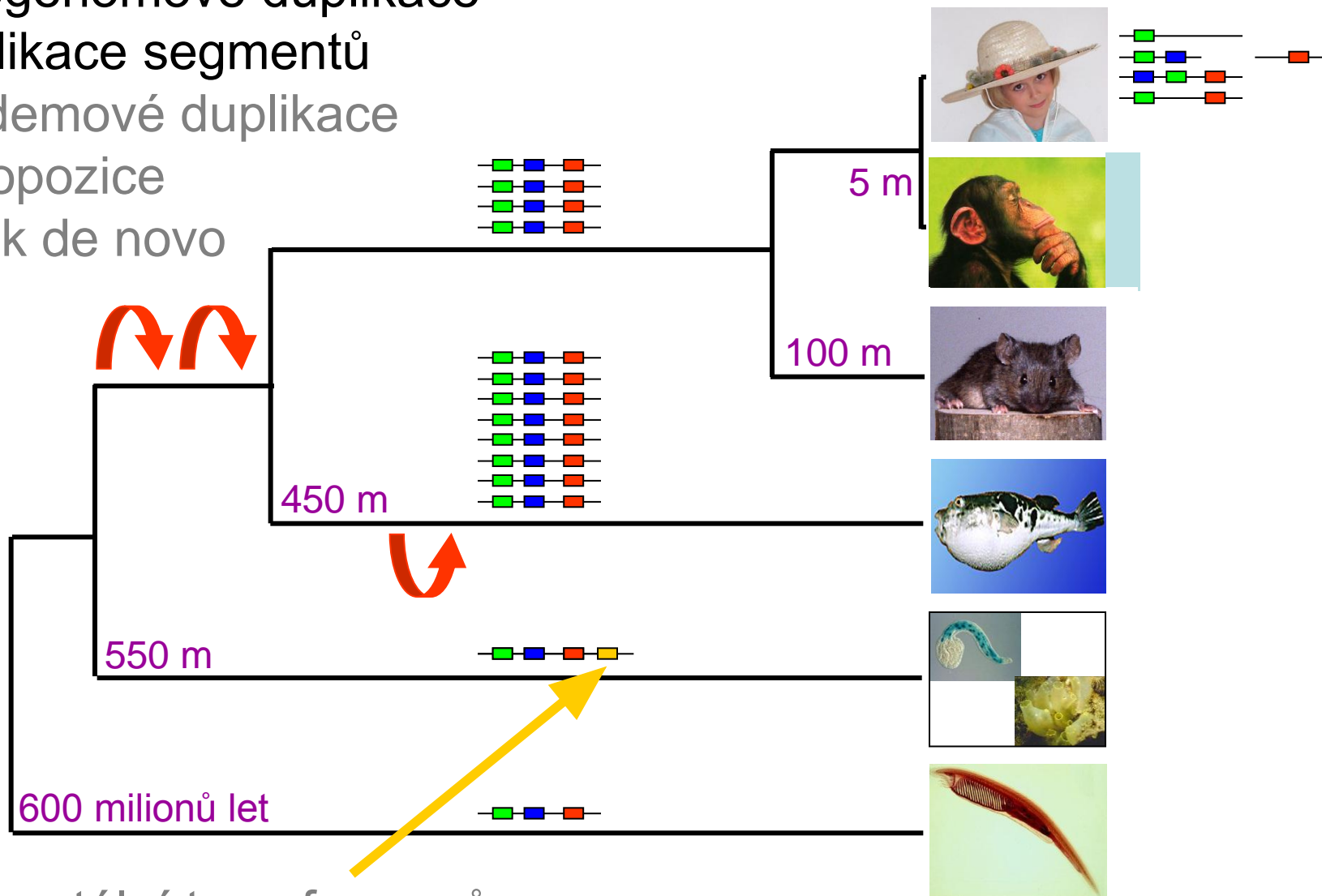
700



22

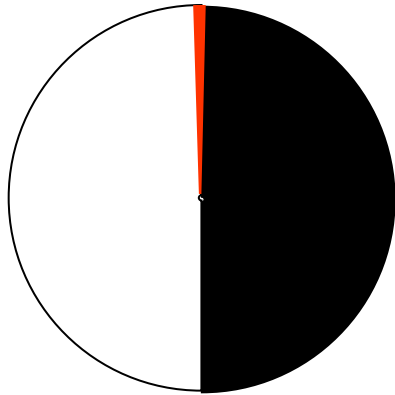


původ genů
celogenomové duplikace
duplikace segmentů
tandemové duplikace
retropozice
vznik de novo



horizontální transfer genů
(také transfer z genomu organel)

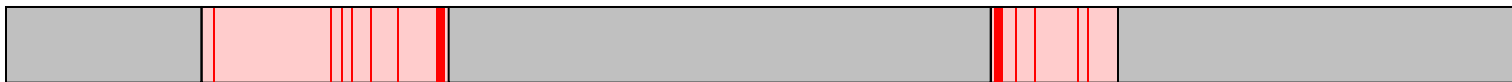
repetitivní DNA



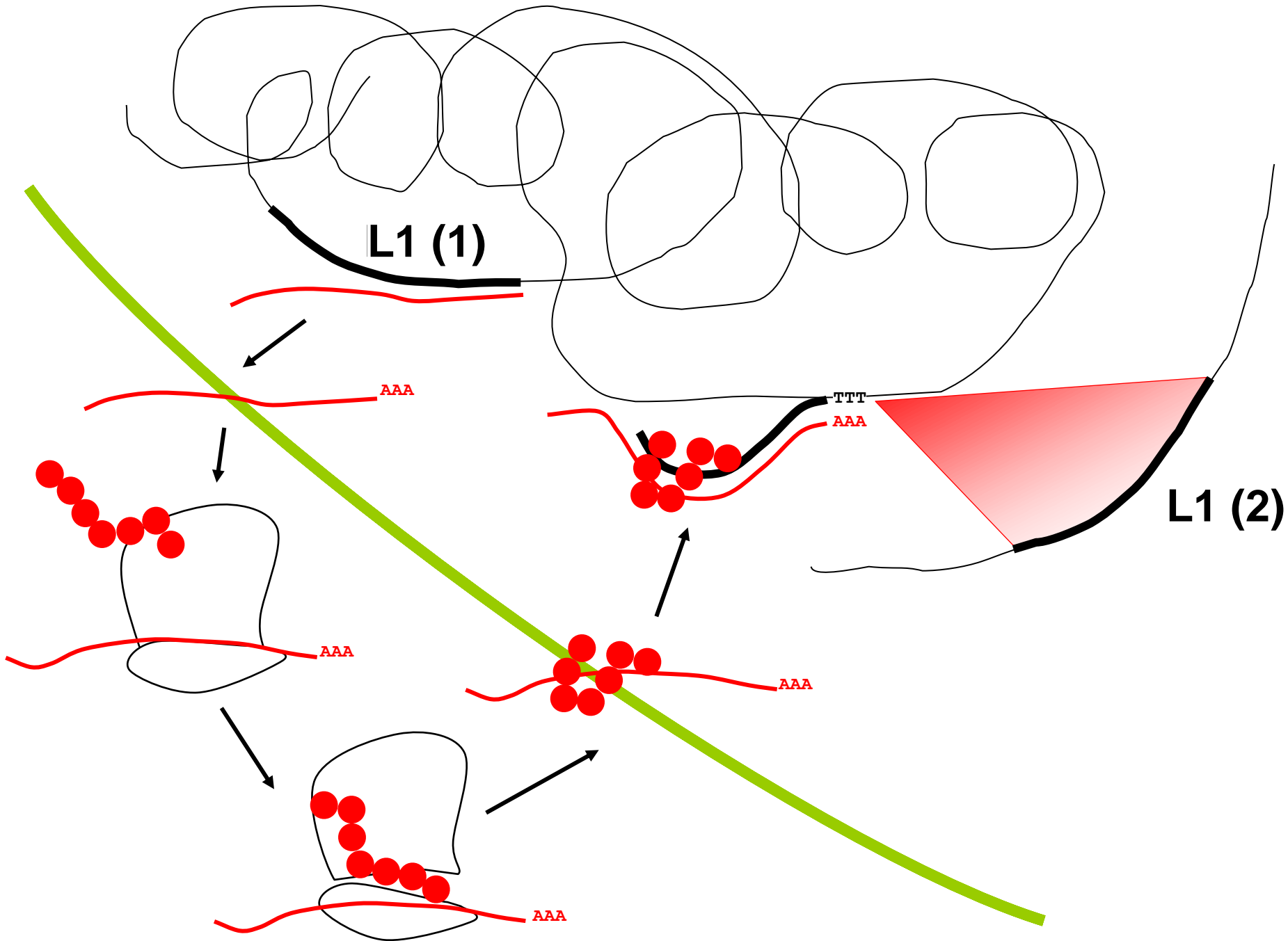
exony cca. 1.5%

repetitivní DNA cca. 50%

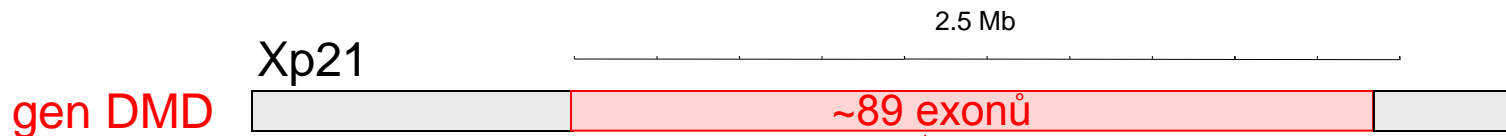
nerepetitivní DNA



mobilní (sobecká) DNA



DMD - Duchennova muskulární dystrofie



DMD intron 44 ← 10 20 30 40 50 60

DMD exon 44 ←

5' aagacttacCTTAAGATACCATTGTATTTAGCATGTTCCCAATTCTCAGGAATTTGTGT
 3' ttctgaatgGAATTTATGGTAAACATAAATCGTACAAGGGTTAAGAGTCCTTAAACACA



TTAAAA CCTTTGTCCTACTAC

CTTTCTGAGAACTGTTCTGTTAGCCACTGATTAATATCTTTATATCA
 GAAAGACTCTTTGACAAGTCGAAGACAATCGGTGACTAATTATAGAAATATAGTCTACT
 K R L F Q E A E T L W Q N F I K I D S S

GTAGGTTGCAAAAATTTCTCCCATGTTGTAGGTTGCTGTTCACTCTGATGGTAGTTTC
 CATCCAACGTTTTTAAAAGAGGGTACAACATCCAACGGACAAGTGAGACTACCATCAAG
 Y T A F I K E W T T P Q R N V R I T T E

TTTTGCTGTGCAGAAGCTCTTTAGTTTAAATAGATCCATTTGTCAATTTTGTCTTTTGT
 AAAACGACACGTCTTCGAGAAATCAAATTAATCTAGGGTAAACAGTTAAAACAGAAAACA
 K A T C F S K L K I L D W K D I K D K T

TGCCATTGCTTTTGGTGTTTTGGACATGAAGTCCTTGCCACGCCTATGTCTGAATGGT
 ACGGTAACGAAAACCAAAAACCTGTACTTCAGGAACGGGTGCGGATACAGGACTTACCA
 A M A K P T K S M F D K G V G I D Q I T

duplikace cílové sekvence

AATGCCTAGGTTTTCTTAGGGTTTTATGGTTTTAGGTTTAACTTTAAATCTTTAAT
 TTACGGATCCAAAAGAAGATCCCAAAAATACAAAATCAAATTCGAAATTTAGAAATTA
 I G L N E E L T K I T K P K V N L D K I



CCATCTTGAATTGATTTTTGTATAAGGTGTAAGGAAGGGATCCAGTTTCAGCTTTCTACA
 GGTAGAACTTAACTAAAACATATCCACATTCCTCCCTAGGTCAAAGTCGAAAGATGT
 W R S N I K T Y P T L F P D L K L K R C

TATGGCTAGCCAGTTTTCCAGCACCATTATTAATAGGGAATCCTTTCCCATTGCTT
 ATACCGATCGGTCAAAGGGTCGTGGTAAATAATTTATCCCTTAGGAAAGGGGTAACGAA
 I A L W N E W C W K N F L S D G W Q K

GCGGCATTAT

GTTTTTCTCAGGTTGTCAAAGATCAGATAGTTGTAGATATGCGGCAAAAAAAAAAAAA
 CAAAAGAGTCCAAACAGTTTCTAGTCTATCAACATCTATACGCGGTTTTTTTTTTTTTTT
 N K E P K D F I L Y N Y I H P F F F F F

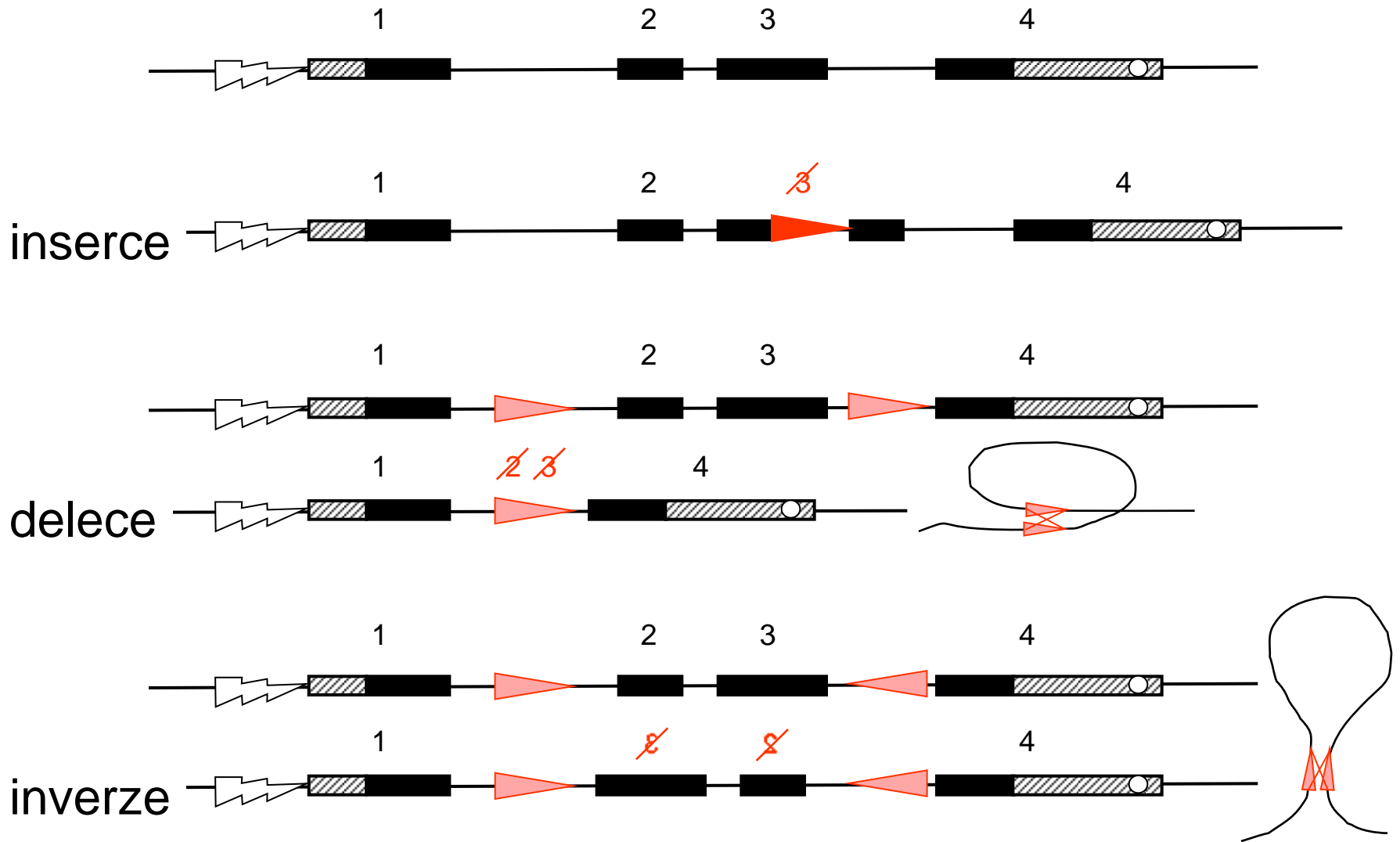
AA
 TT
 F F F F F F F F F F F F F F F F I K I D Y H F R R W

TTTCTCAACAGATCTGTCAAATCGCctgcaggtaaaagcatatggatcaagaaaaataga 3'
 AAAGAGTTGTCTAGACAGTTTAGCGgacgtccattttcgtatacctagttctttttatct 5'
 K E V S R D F R

DMD exon 44 ←

← DMD intron 43

vliv transpozonů na genom



fluidita/plasticita genomu