

Genetické metody v zoologii

Miloš Macholán (macholan@iach.cz)
Josef Bryja (bryja@brno.cas.cz)

Datum	Přednášející	Kde	Téma
21.2.2019	J. Bryja	UKB	Úvod (význam genetických metod v zoologii a evoluční biologii; základní přehled metod, atd.). Analýza DNA I (izolace DNA, genetické markery - jaderná vs. mimojaderná DNA, PCR, real-time PCR, Sangerovo sekvenování)
28.2.2019	J. Bryja	UKB	Analýza DNA II ("single-locus" DNA markery: mikrosatelity, LINE, SINE)
7.3.2019	M. Macholán	UKB	Analýza fenotypu (signální fenotypy, epigenetické znaky, kvantitativní znaky, analýza landmarků)
14.3.2019	J. Bryja	UKB	Analýza DNA III (SNP a jejich analýza: RFLP, DGGE, TGGE, SSCP, klonování, nové techniky SNP genotypizace - SNP chipy atd.)
21.3.2019	M. Macholán	UKB	Cytogenetika (analýza karyotypu, proužkování, FISH, „painting“). Elektroforéza proteinů
28.3.2019	J. Bryja	UKB	Analýza DNA IV ("multi-locus" DNA markery: minisatelitový fingerprinting, RAPD, AFLP)
4.4.2019	J. Bryja	UKB	Analýza DNA V (Úvod do "high-throughput sequencing" = NGS technologií)
11.4.2019	J. Bryja	UKB	Analýza DNA VI (Aplikace technologií NGS, např. metagenomika, hybrid enrichment, RADseq, ddRADseq, atd.)
18.4.2019	J. Bryja	UKB	Analýza genové exprese (microarrays, qPCR, transcriptomika, RNAseq)
25.4.2019	J. Bryja	UKB	Základní manipulace s genetickými daty I (jaderná data založená na frekvencích - základní analýzy genetické variability a struktury populací, HWE, STRUCTURE, atd.)
2.5.2019	M. Macholán	UKB	Základní manipulace s genetickými daty II (analýza sekvencí - datové formáty, alignování sekvencí, základní práce s databázemi - GenBank, NCBI, BLAST, Dryad, TreeBASE aj.)
9.5.2019	J. Bryja + doktorandi	ÚBO AV ČR, Studenec	Analýza DNA v laboratoři (blokové cvičení) - izolace a elektroforéza DNA, PCR, real-time PCR, mikrosatelity, Sangerovo sekvenování, BLAST, ukázka NGS dat
16.5.2019	rezerva		

Doporučená literatura (česká)

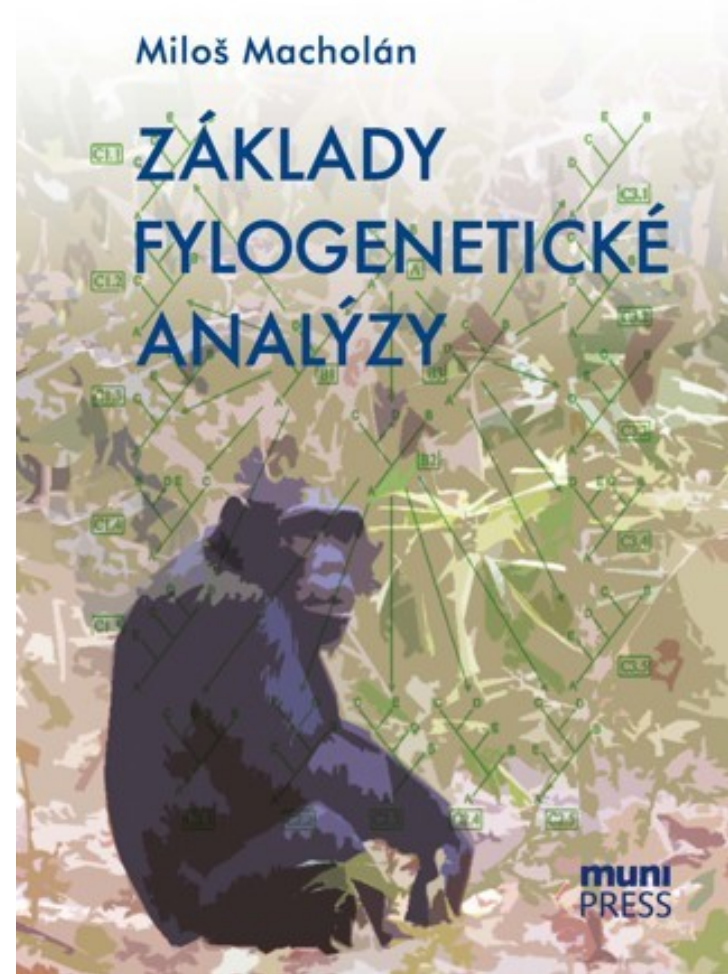
Genetické metody v zoologii

Jan Zima, Miloš Macholán, Pavel Munclinger, Jaroslav Piálek

Nakladatelství Karolinum 2004



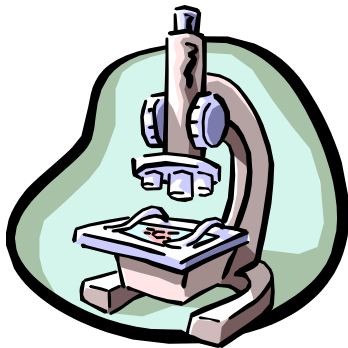
- M. Macholán
- Základy fylogenetické analýzy (2014)



Proč?

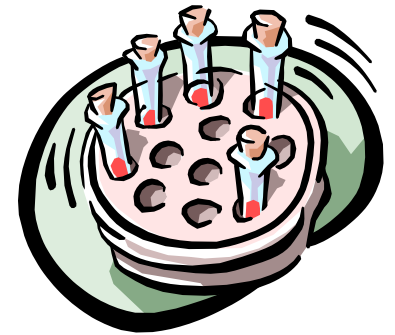
Problém:
zoologie, taxonomie
ekologie, evoluční biologie

Genetické metody:



klasické metody
morfologická,
ekologická,
bionomická
data

genetická data



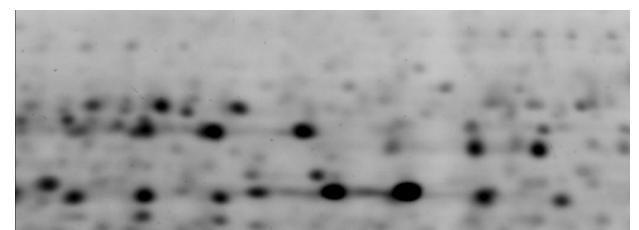
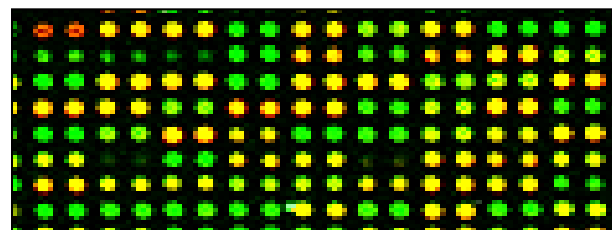
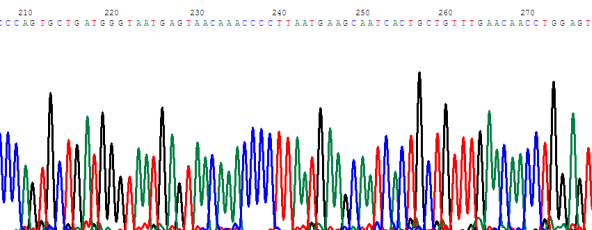
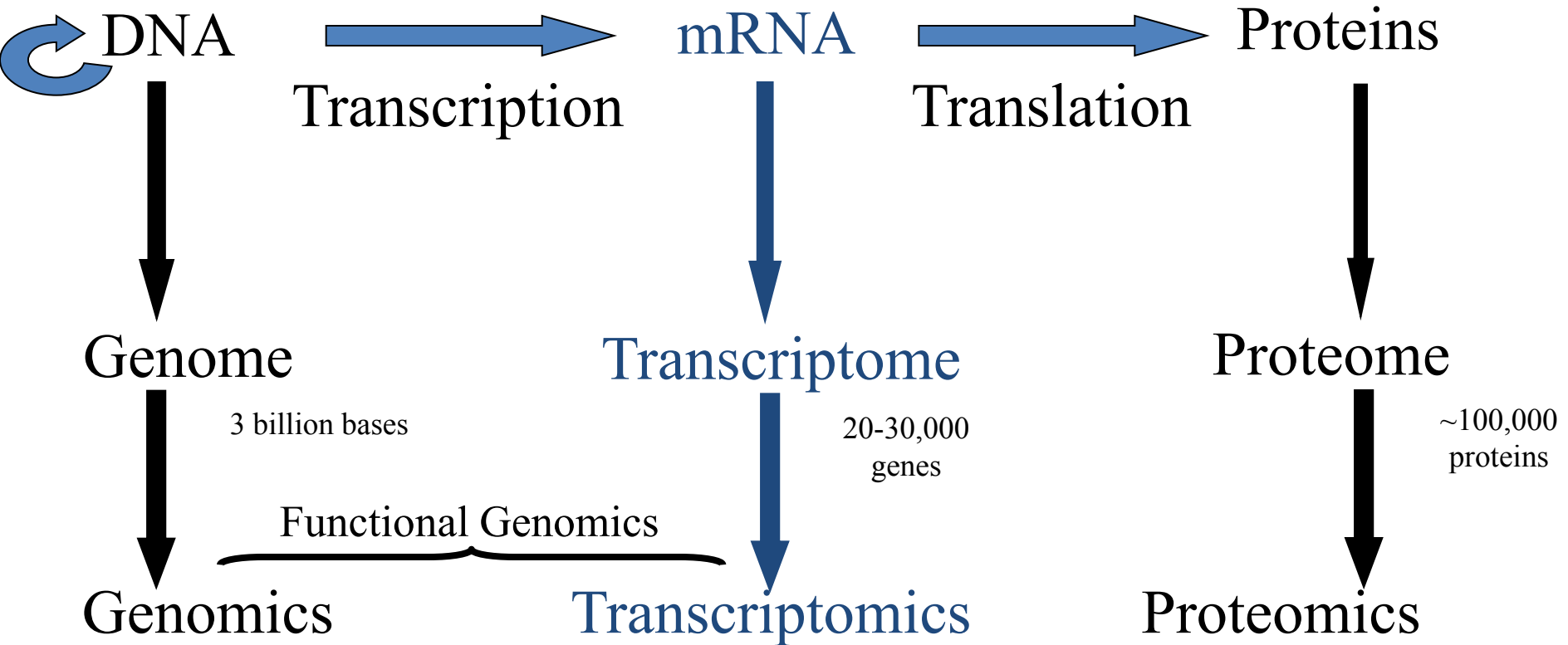
**Další úroveň poznání
Odpovědi na nové otázky**

Na které otázky lze nalézt odpověď jen s využitím genetických metod?

- rekonstrukce fylogenetických vztahů mezi populacemi, druhy či vyššími taxony (konvergence)
- kryptické druhy
- složení společenstev – metabarcoding
- izolace populací (tj. počet migrantů) – nemusí být zřejmá
- identifikace z trusu, chlupů - pohyb jedinců skrytě žijících druhů
- paternita – páření často skryté a nemusí vést k oplození
- a mnoho dalších ...

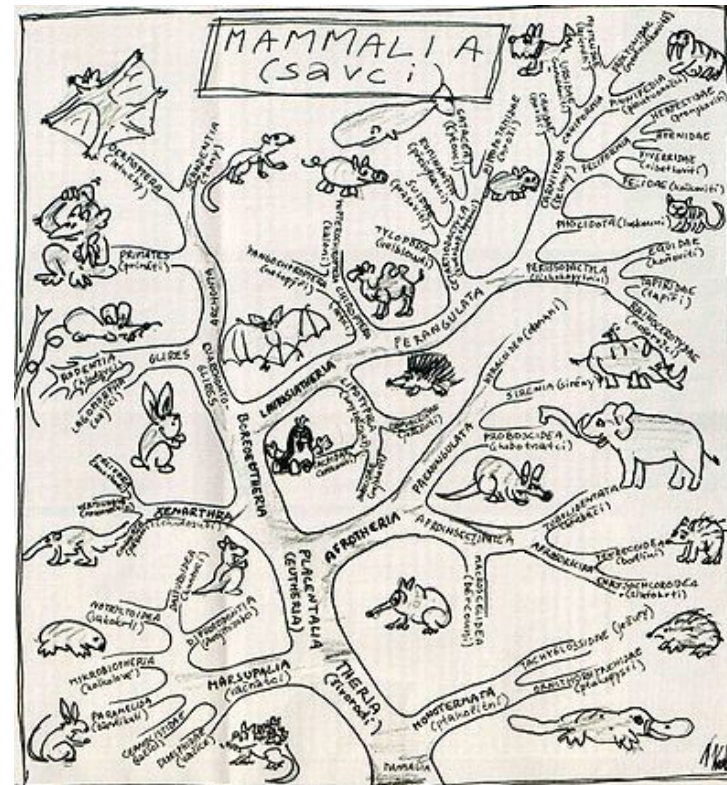
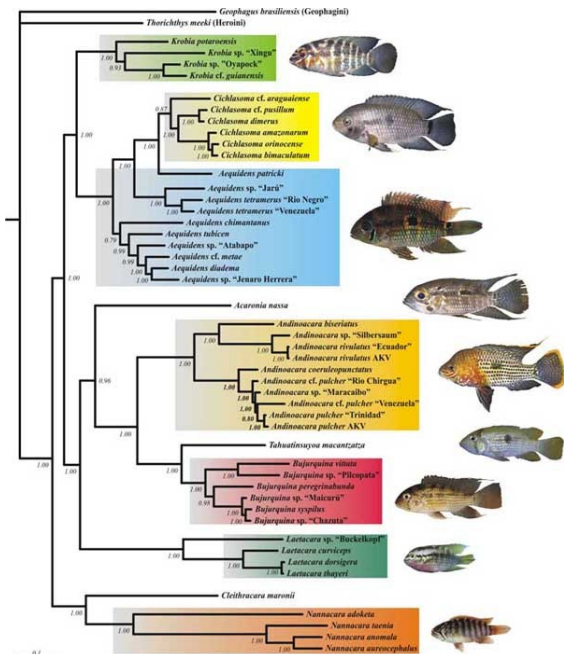
Genetické metody

= studium genetické variability



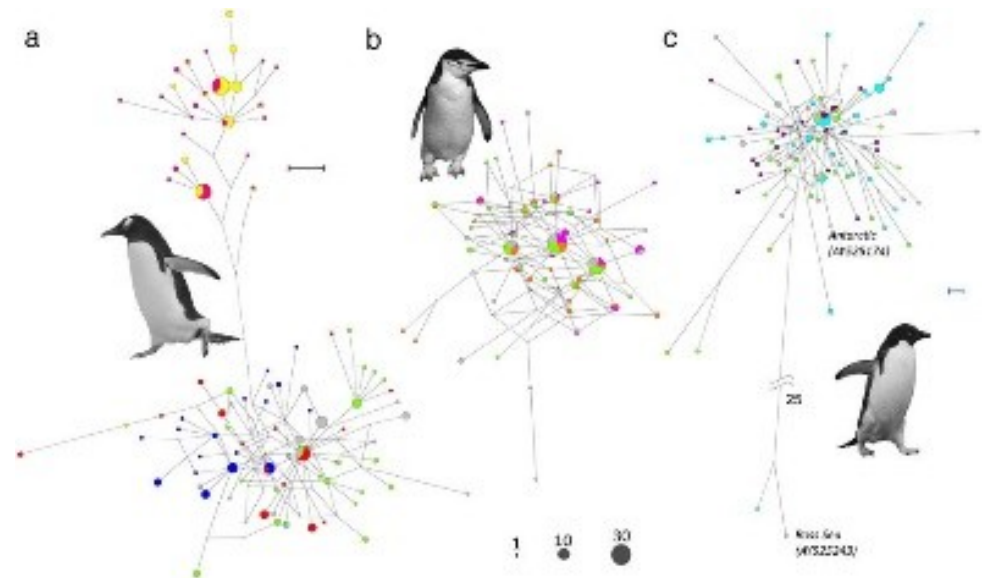
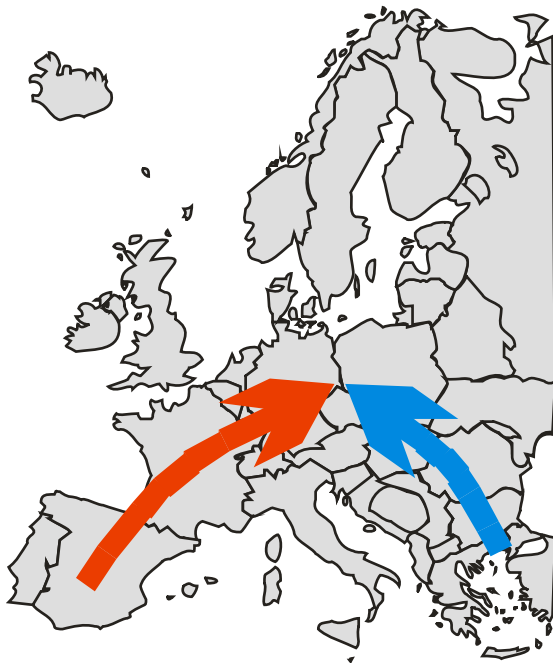
Úrovně genetické variability

- **druhy a vyšší taxony – fylogenetické analýzy** (fylogenetická systematika)



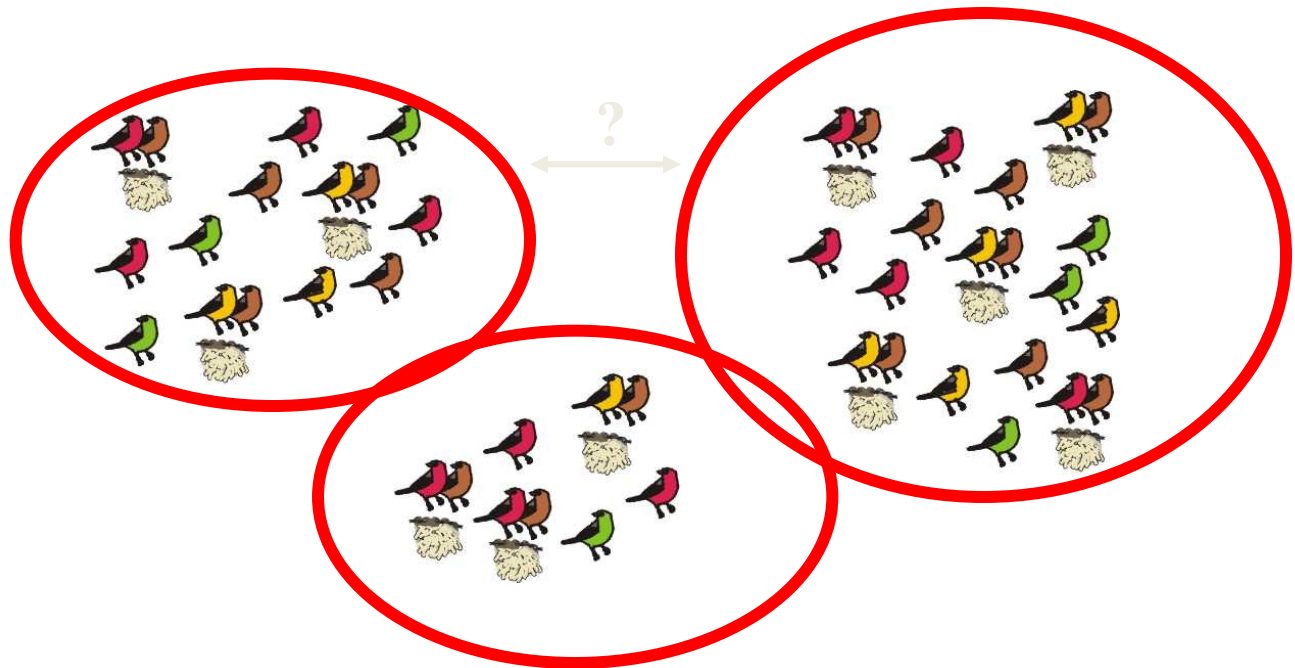
Úrovně genetické variability

- **populace až druh** – studium speciace, fylogeografie, delimitace druhů, hybridizace



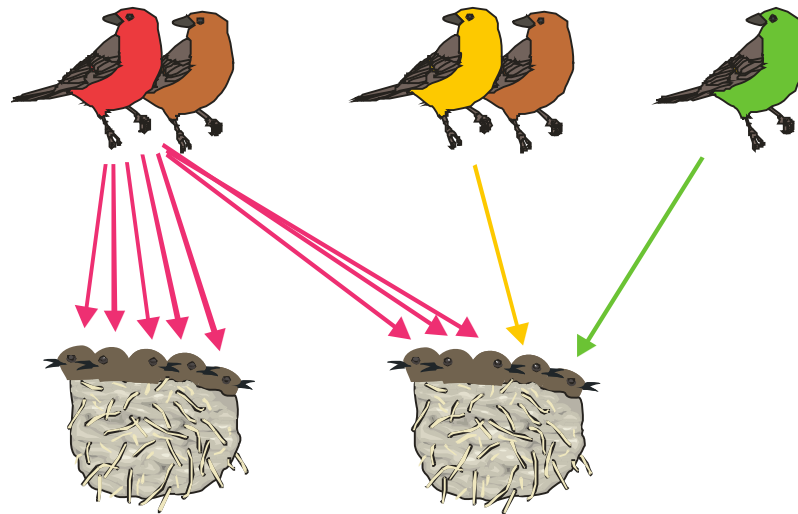
Úrovně genetické variability

- **populace** – populační biologie, ochránářská genetik



Úrovně genetické variability

- **jedinec** – analýzy příbuznosti (behaviorální ekologie, např. analýzy paternity)



Genetické metody v zoologii

- jak genetická data získat, tj. které **techniky** použít
- **Mechanismy mikroevoluce** (jaro)
- základní typy a zpracování (editace) genetických dat
- **Molekulární ekologie** (podzim)

Genetické DNA markery

- **kódující DNA (geny)**
- přepisované sekvence (cca 20-25 tisíc genů u obratlovců)
- genetický kód
- vytvářejí fenotyp
- podléhají přírodnímu výběru
- rostoucí význam v ekologickém výzkumu
- **nekódující DNA**
- nefunkční (neznámá funkce)
- neutrální k přírodnímu výběru
- většina DNA u eukaryot (až 95% u obratlovců)
- pseudogeny
- repetitivní DNA

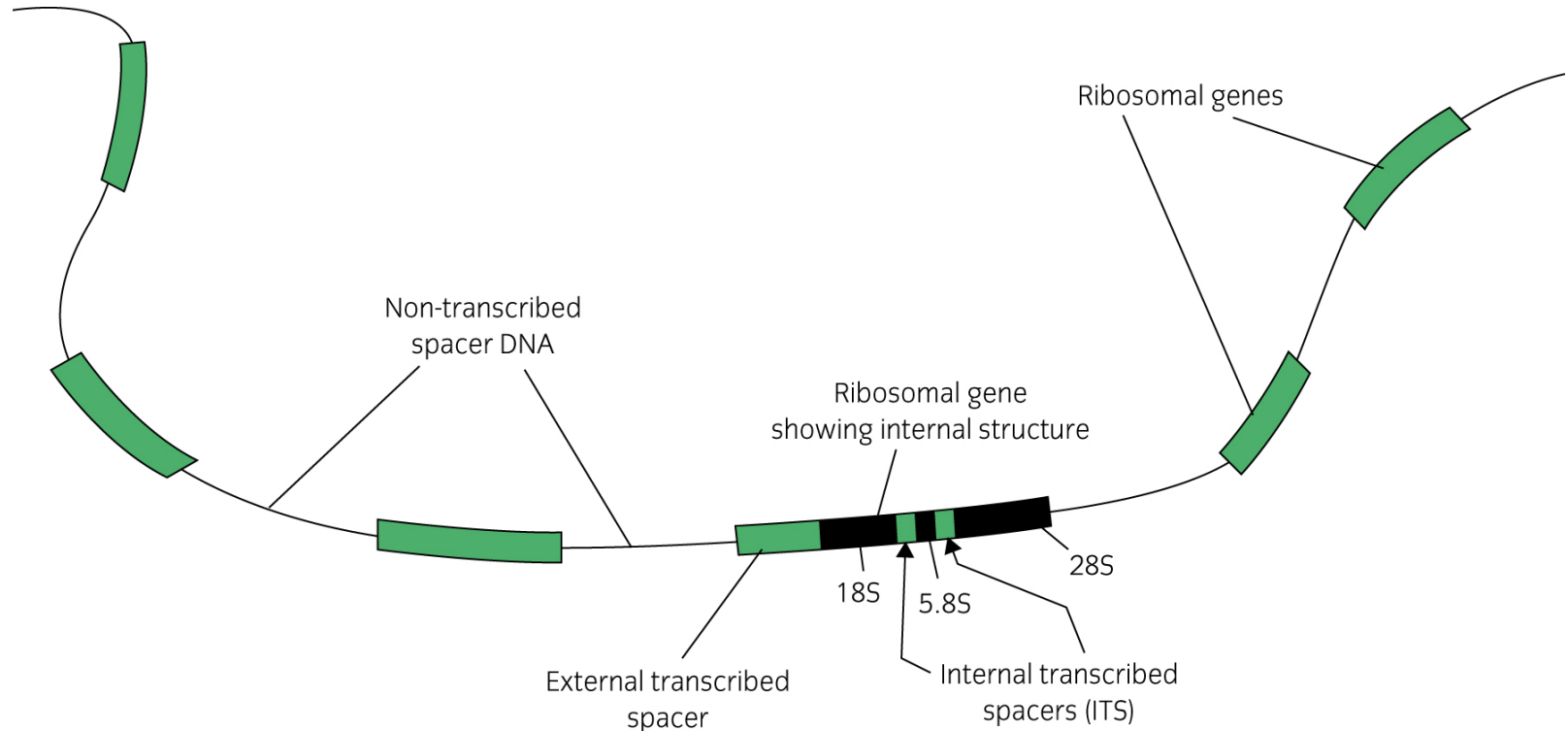
Repetitivní DNA

DNA	Typical sequence length (bp)	Location
Satellites (>10 ⁶ repeats/genome)	5-100	Tandem arrays, scattered throughout the genome
Minisatellites (>10 ³ loci/genome)	20-300	Tandem arrays up to 5 kb in length, scattered throughout the genome
Microsatellites (>10 ⁴ loci/genome)	1-6	Tandem arrays up to a few 100 bp in length, scattered throughout the genome
Telomeres	4-8	Tandem arrays up to 1kb in length, at the ends of each chromosome
SINEs (>10 ⁵ /genome)	50-500 (100-300)	Interspersed throughout the genome
LINEs (>10 ³ /genome)	1-5 k	Interspersed throughout the genome

Kódující („funkční“) DNA

- 1) ribosomální DNA
- 2) jaderné strukturální geny (protein-coding genes)
- 3) mitochondriální DNA

1. Ribosomální DNA

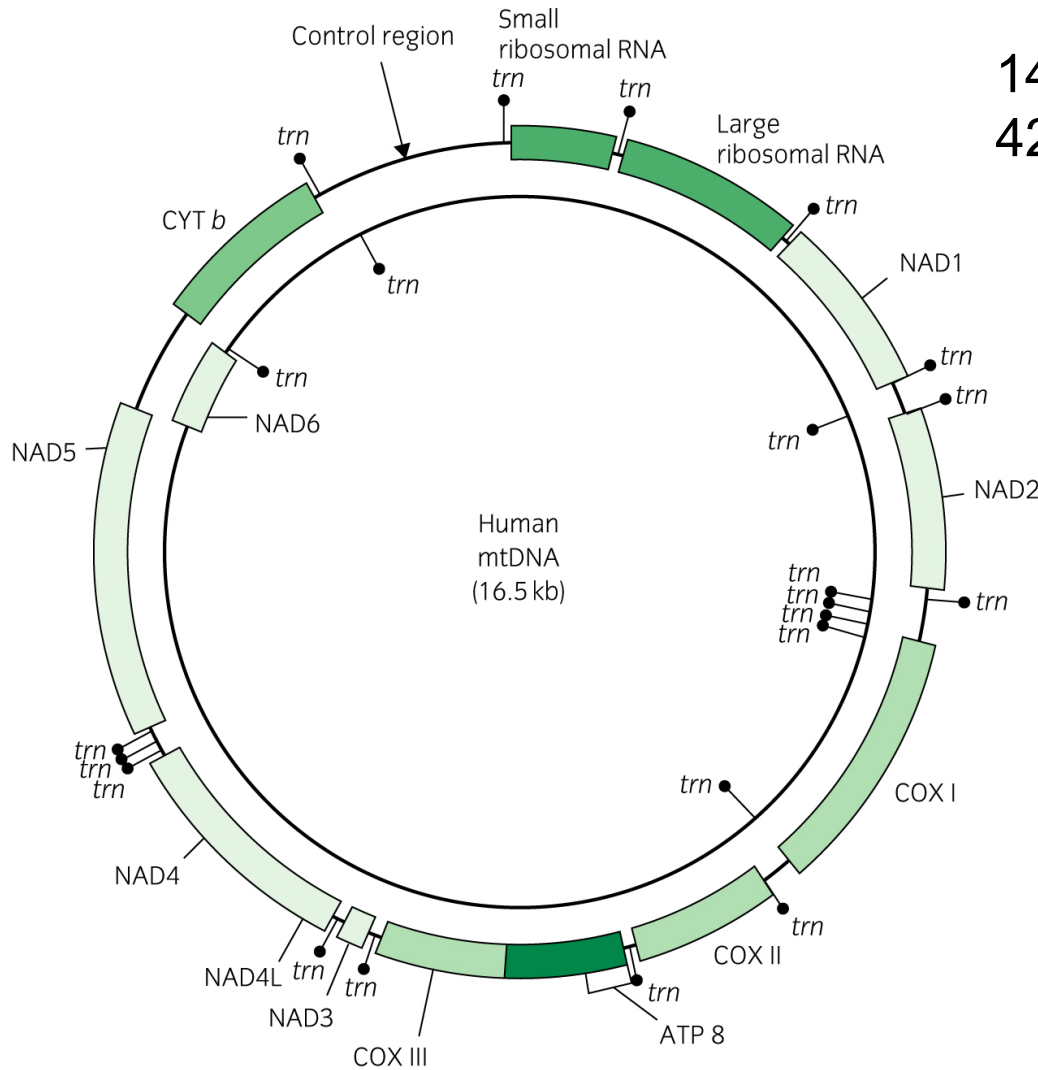


- geny pro ribosomální RNA – mnoho shluků (operonů) u eukaryot
- 16S, 23S, and 5S – málo kopií u prokaryot
- rDNAs – phylogeny, ITS – population structure, barcoding (fungi, helminths)

2. Jaderné geny

- nízká variabilita mezi jedinci – významná funkce, purifikující selekce (nejsou často používány jako genetické markery)
- introny – více variabilní než exony, často ve fylogenetických analýzách
- alozymy
- MHC geny
- SNPs – narůstající význam (jednoduché mutace způsobují významnou funkční změnu)
- studium transkripce - transkriptomika

3. Mitochondriální DNA



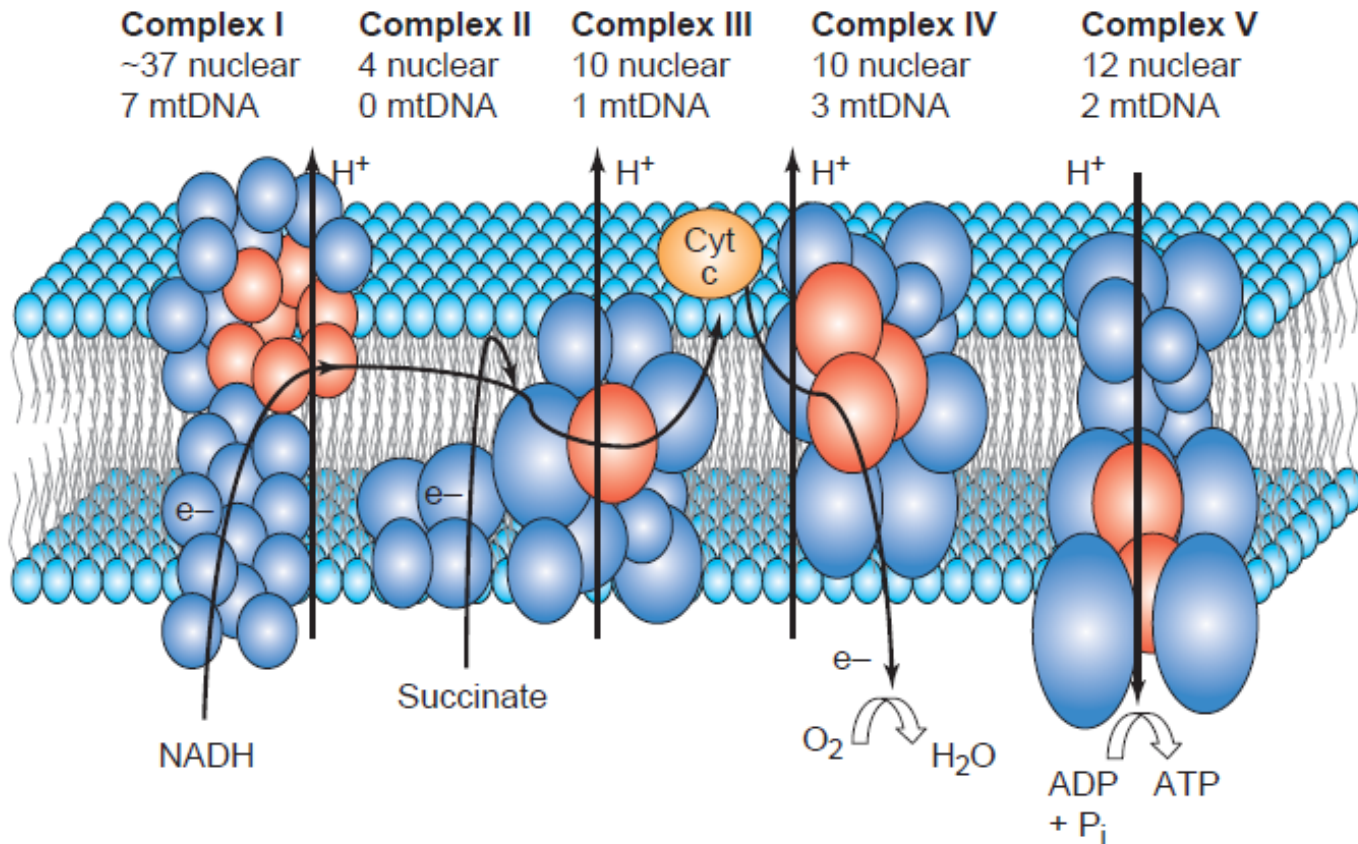
14 kbp (*Caenorhabditis*)

42 kbp (*Placopecten*)

- maternally inherited (?)
- no recombination (?)
- no heterozygotes (?)
- mnoho kopií v každé buňce (ca. 8000 u člověka) – lépe se analyzuje
- « numts » nuclear copies of mtDNA
- vhodná pro fylogenetické a fylogeografické analýzy

Dýchací řetězec v mitochondriích

red = mtDNA blue = nDNA



- koevoluce jaderné a mitochondriální DNA → DNA-barcoding, nejčastější marker pro určování druhů

„Molekulárně-genetické“ metody

- analýza polymorfismu DNA (genetické markery)
- konzervativní vs. variabilní úseky (« loci »)
- **sekvenční polymorfismus:**

CGCATCTCTAGCTTC**GATTCAGGAA**

CGCATCTCTAGCTTT**GATTCAGGAA**

„Molekulárně-genetické“ metody

- analýza polymorfismu DNA (genetické markery)
- konzervativní vs. variabilní úseky (« loci »)
- sekvenční polymorfismus
- **délkový polymorfismus**

CG**CACA**TCTCTAGCTTCGATTCAGGAA

CG**CA**TCTCTAGCTTTGATTCAGGAA

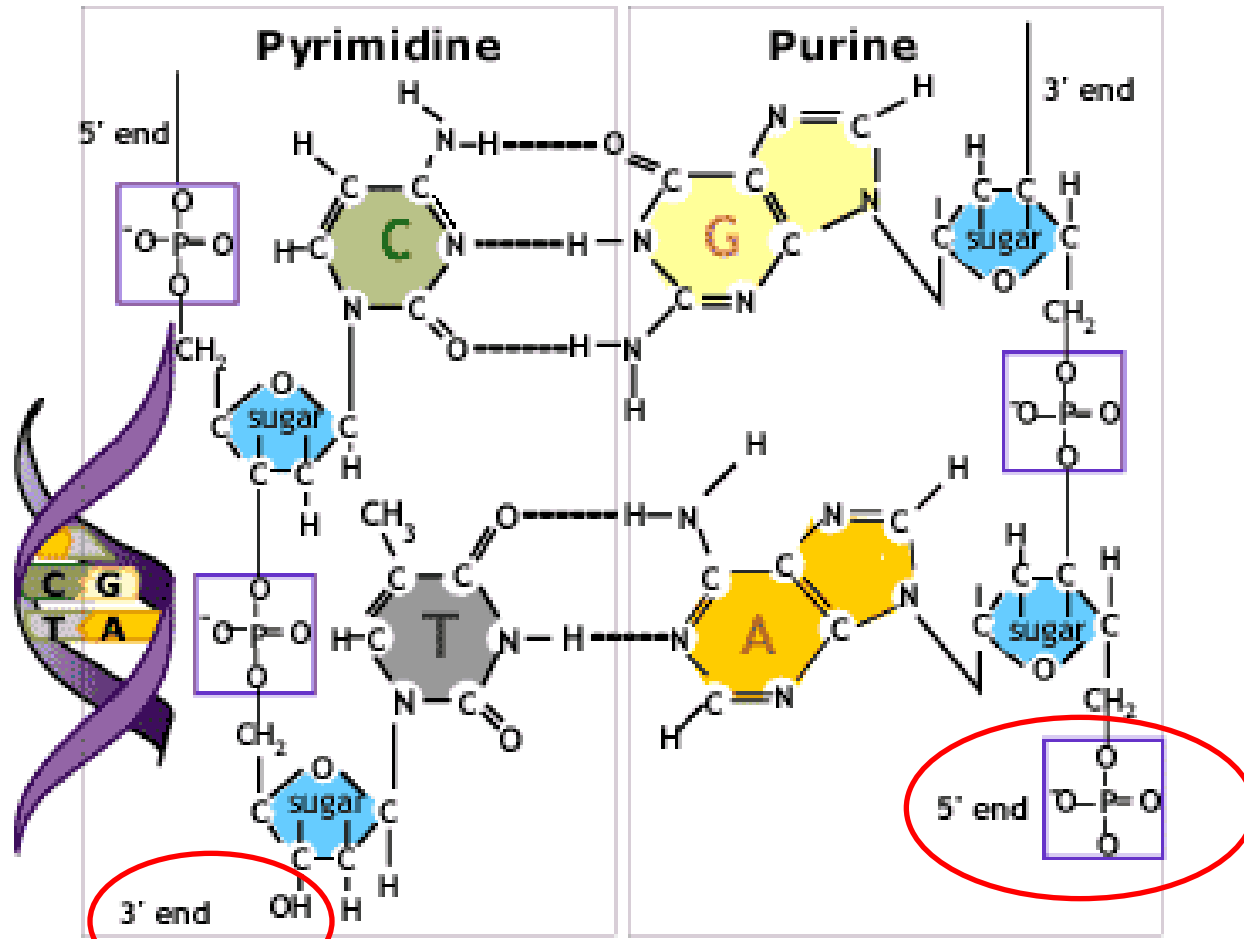
Vznik DNA polymorfismu

- mutace (transice, transverze, inzerce, delece)
- rekombinace (kombinace změn vzniklých mutacemi, duplikace a delece při rekombinačních chybách)
- transpozice
- ⇒ obecná molekulární genetika

Genotypizace – stanovení genotypu

- stanovení formy určitého úseku DNA
(alely = $2n$, haplotypu = $1n$)
 - 1) izolace celkové DNA z tkání
 - 2) amplifikace požadovaného úseku DNA
(PCR-based methods)
 - 3) studium variability daného úseku
(lokus)

Základní struktura molekuly DNA



3' - OH konec

(nutný k navázání dalšího nukleotidu při syntéze DNA)

5' - fosfátový konec

(ve vodném roztoku způsobuje záporný náboj)

Enzymy používané při molekulárně-genetických manipulacích

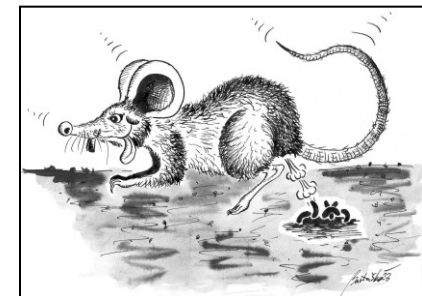
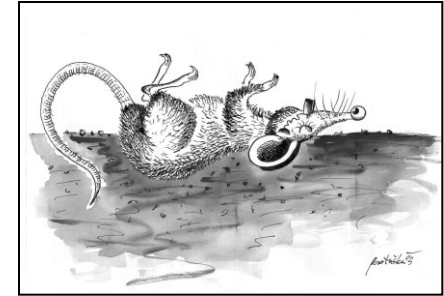
- DNA-polymeráza
- DNA-exonukleáza, DNA-endonukleáza
- DNA-ligáza
- DNA-transkriptáza
- RNA-reverzní transkriptáza

Izolace DNA

- rozmanitý biologický materiál – musí obsahovat buněčná **jádra nebo mitochondrie** s nedegradovanou DNA
- dnes většinou komerční kity (cca 50-100 Kč/vzorek, ale pro některé aplikace i doslova „za pár korun“)
- Izolace RNA (exprese specifických genů) – dříve problém, dnes RNAlater

Způsoby získání DNA z volně žijících živočichů:

- 1. destrukční** – živočich je usmrcen kvůli získání tkání potřebných na genetické analýzy
- 2. nedestrukční (invazivní)** – živočich je odchycen a je mu odebrán vzorek tkáně nebo krve
- 3. neinvazivní** – zdroj DNA je „zanechán za živočichem“ a je získán bez potřeby odchyty, manipulace či dokonce pozorování



Fixace materiálu

+

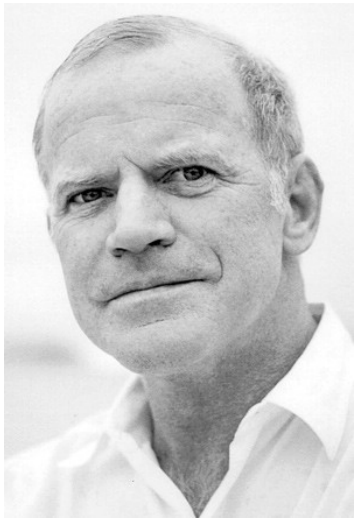
- čerstvá tkáň
- čistý EtOH
- rychlé vysušení
- speciální extrakční pufry
- zamražení

-

- formaldehyd
 - opakované zamrazování
 - rozvlhčování sušeného materiálu
 - další fixační média
-
- speciální metody pro izolaci ze subrecentního materiálu (mamuti, hmyz v jantaru, neandrtálci apod.)

PCR

Polymerase chain reaction (jak z málo DNA udělat hodně)



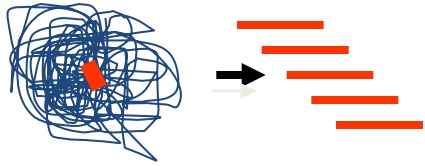
Kary Mullis (1983 – na dálnici ze San Francisca do Mendocina)

- odměna 10 000 USD
- patent pak prodán Roche za 300 000 000 USD)

1993 – Nobelova cena

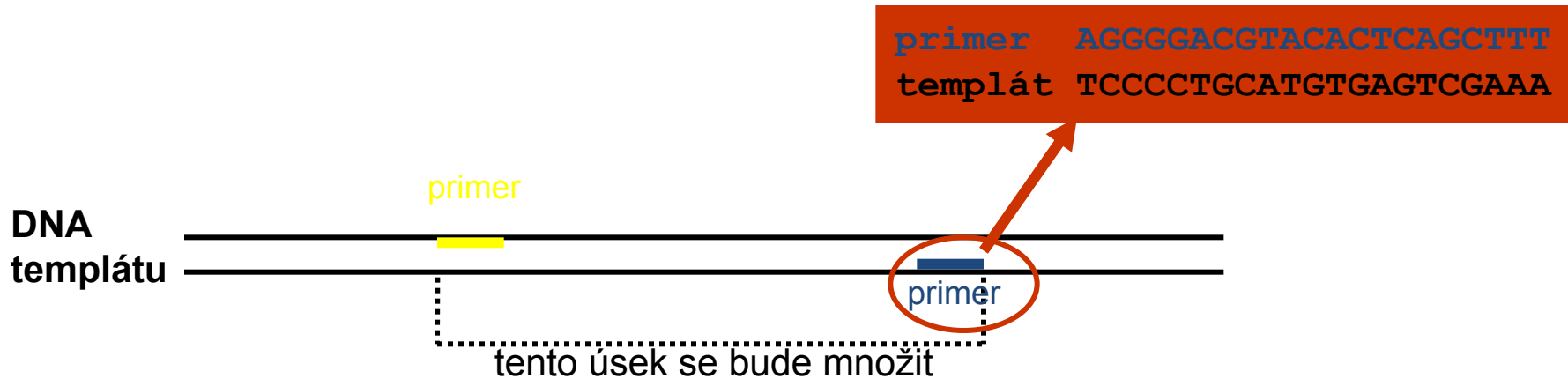
Amplifikace DNA – PCR

Druh	Velikost genomu (bp)	Počet chromozómů (1n)
<i>Caenorhabditis elegans</i>	$8,0 \times 10^7$	4
<i>Drosophila melanogaster</i>	$1,65 \times 10^8$	4
<i>Xenopus laevis</i>	$3,0 \times 10^9$	18
<i>Mus musculus</i>	$3,0 \times 10^9$	20
<i>Homo sapiens</i>	$3,0 \times 10^9$	23



PCR

- Z celkové DNA si namnožíme jen úsek, který nás zajímá.
- Co se bude množit? To určí **primery**.
- **Primery** – krátké oligonukleotidy komplementární k úsekům ohraničujícím místo našeho zájmu.



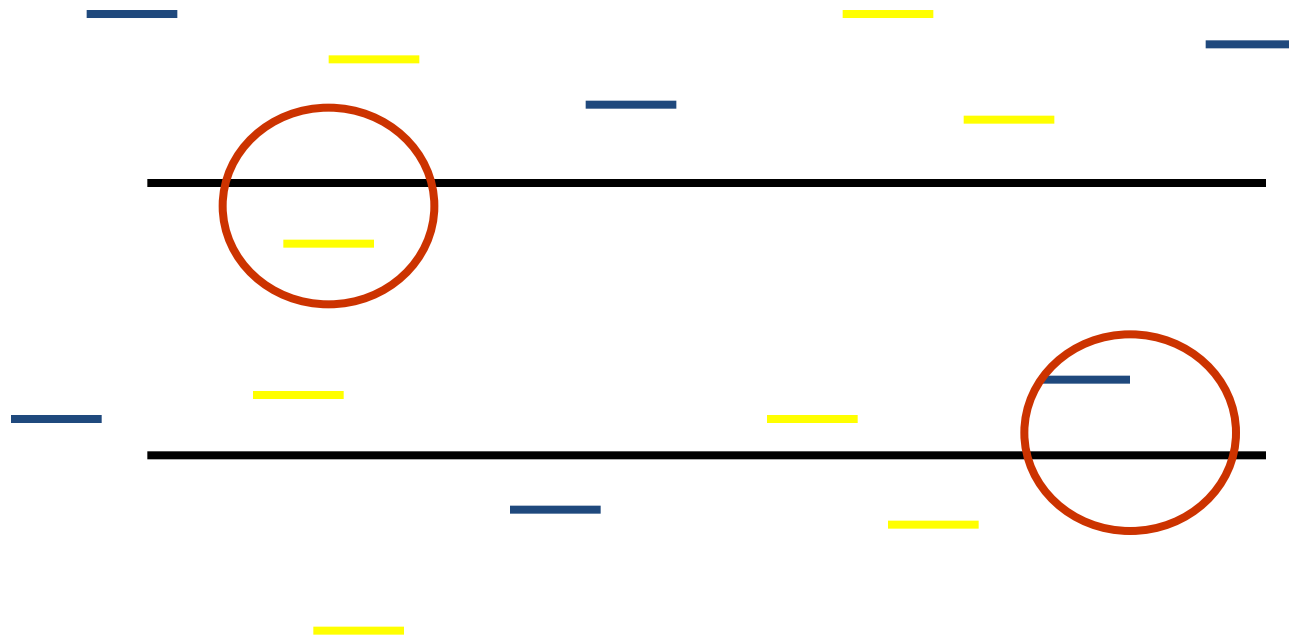
Denaturace (obvykle 95 C)

při **zvýšení teploty** se oddělí komplementární vlákna DNA



Při ochlazení dojde k reasociaci

Primery přidané v nadbytku kmitají díky Brownově pohybu



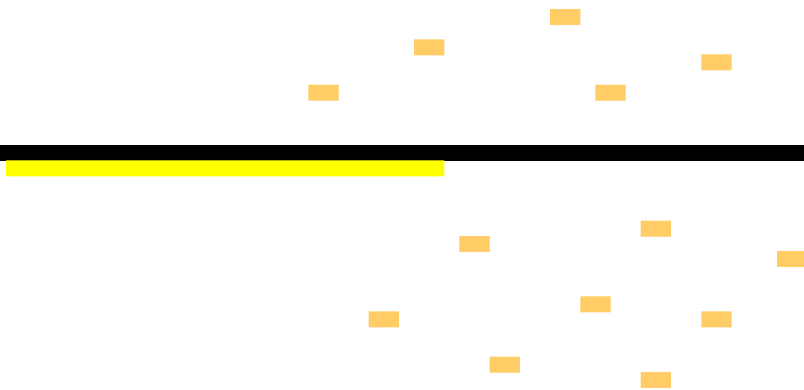
Některé se dostanou do blízkosti komplementárních míst

Při ochlazení primery přisednou rychleji než dojde k vzájemné reasociaci dlouhých vláken DNA (obvykle 50 - 65 C) – „annealing“



V úseku mezi primery zůstanou vlákna DNA oddělena

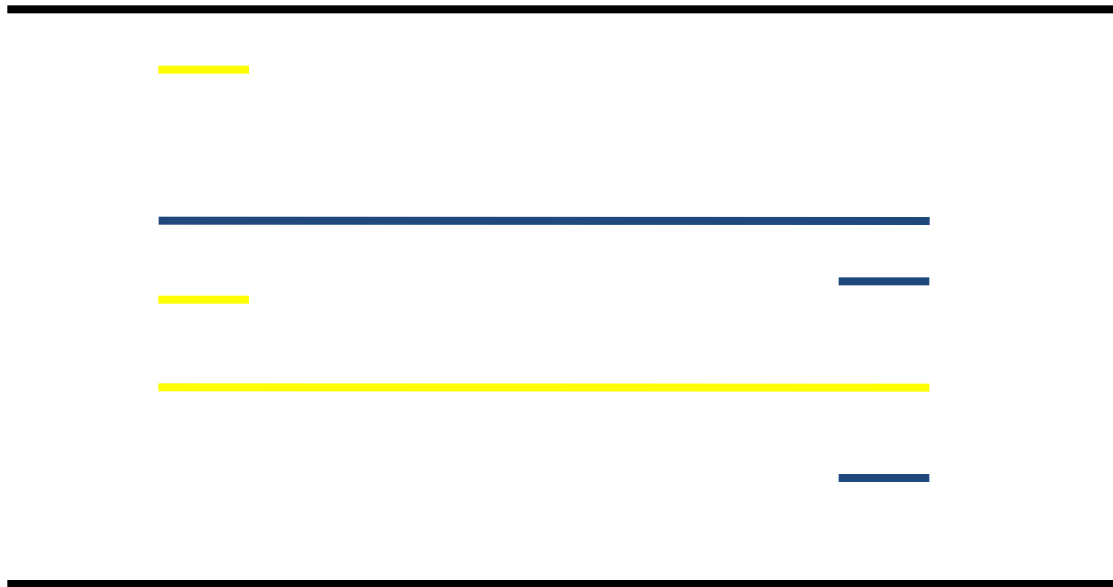
Primery jsou prodlužovány přidáváním nukleotidů podle sekvence templátu (obvykle 72 C – optimum pro *Taq* polymerázu)



Při dalším zahřátí dojde k oddělení templátu a nově vzniklých vláken



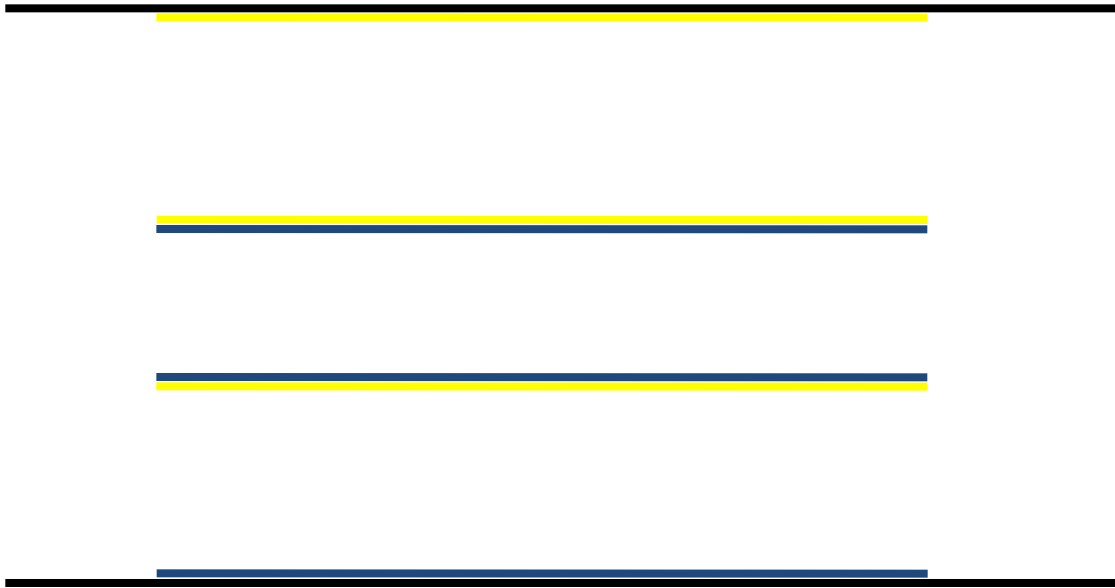
Po ochlazení primery přisednou na templát i nově vzniklé fragmenty („annealing“)



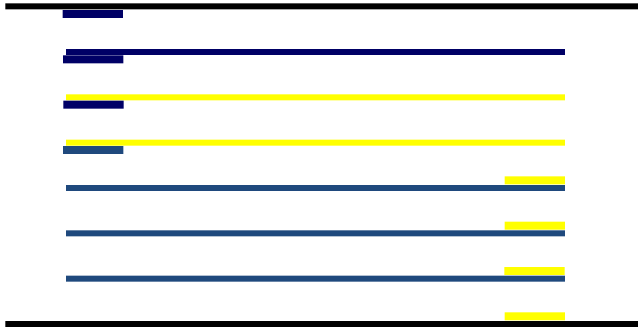
Při 72°C dojde opět k prodlužování primerů a vzniku nových kopií



Při dalším zahřátí...



Ochlazení – nasednutí primerů



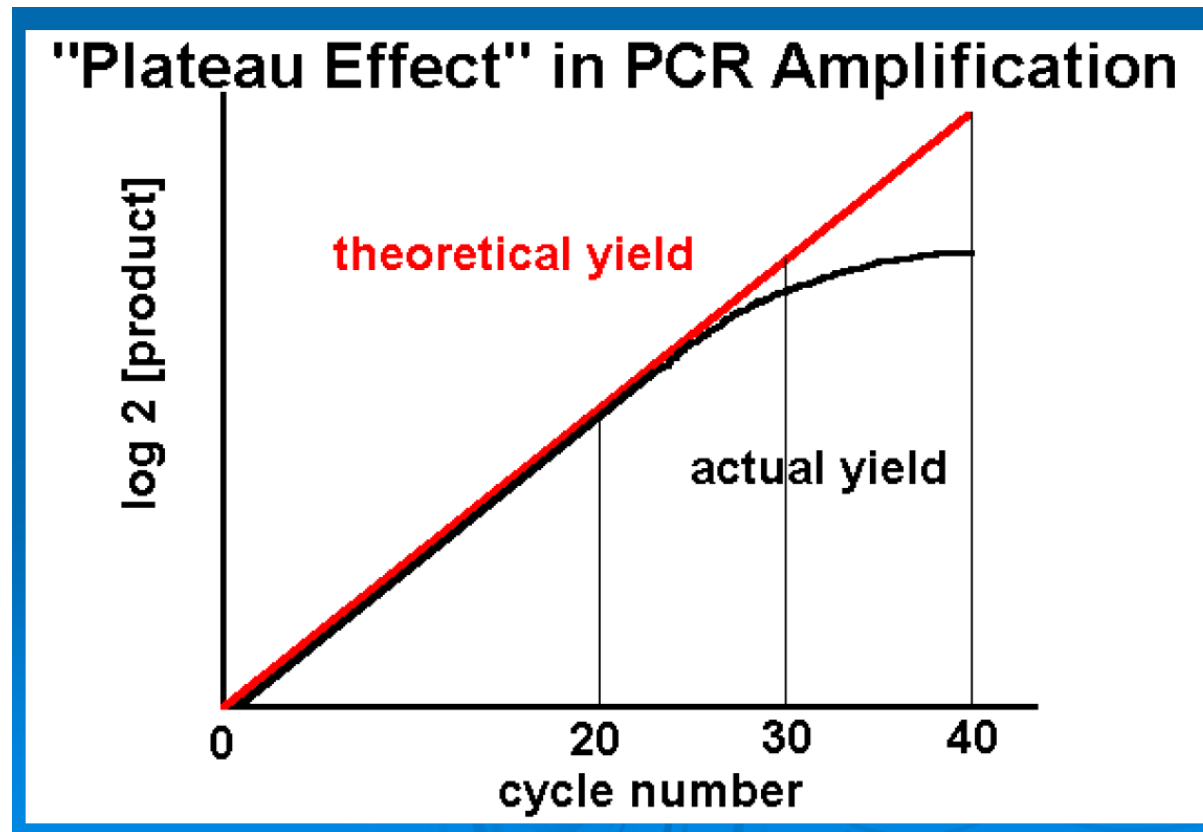
72 C vznik nových fragmentů



95 C denaturace



Inhibice PCR vysokou koncentrací DNA





Cykly (obvykle 20-40):
denaturace (95°C)
nasednutí primerů (50-65°C)
elongace=polymerizace (72°C)

Nejprve však často prodlužená denaturace celkové DNA

Nakonec prodloužená elongace

Příklad
programu

95 C 3 min

95 C 30 s

60 C 30 s

72 C 1 min

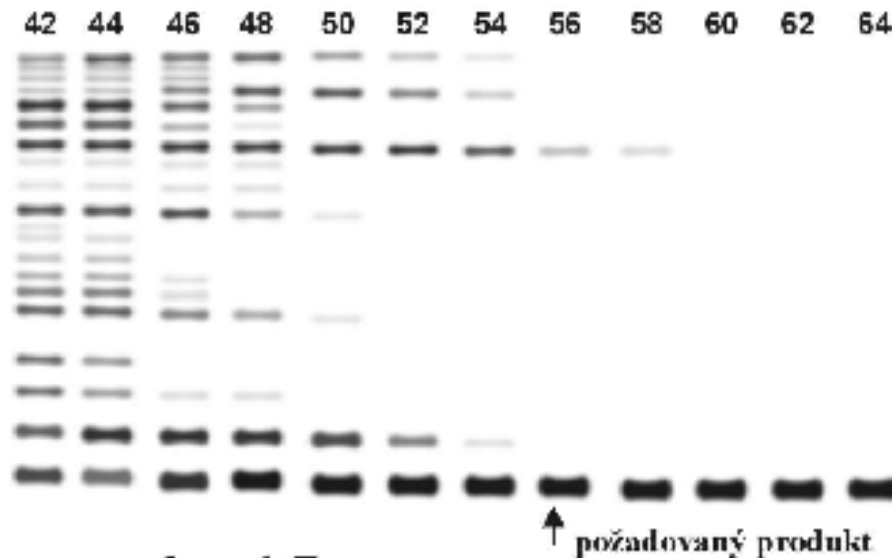
35x zpět

72 C 10 min



Co když PCR nefunguje?

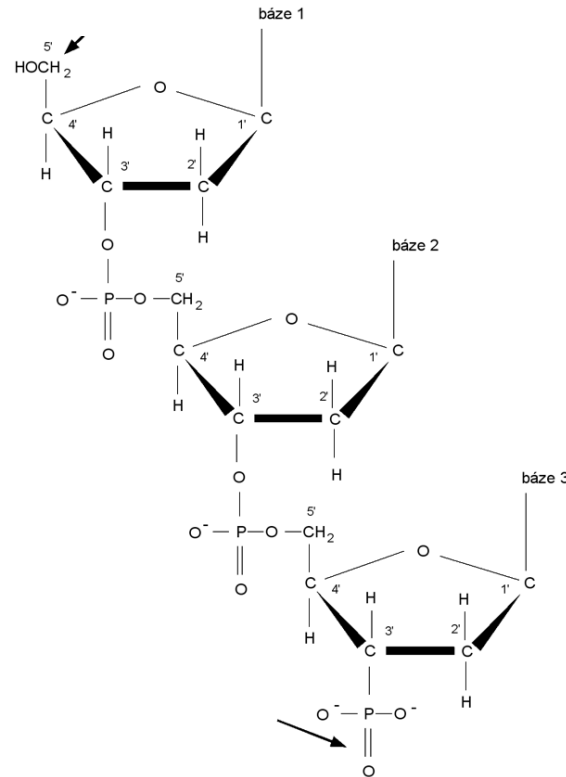
- Měníme teplotu „annealingu“ (nejlépe použijeme gradient teplot, pokud to náš cykler umí)
Vyšší teplota=vyšší specificita
- Měníme koncentraci Mg^{2+} iontů
- Navrhujeme nové primery



Studium variability nasyntetizovaného úseku

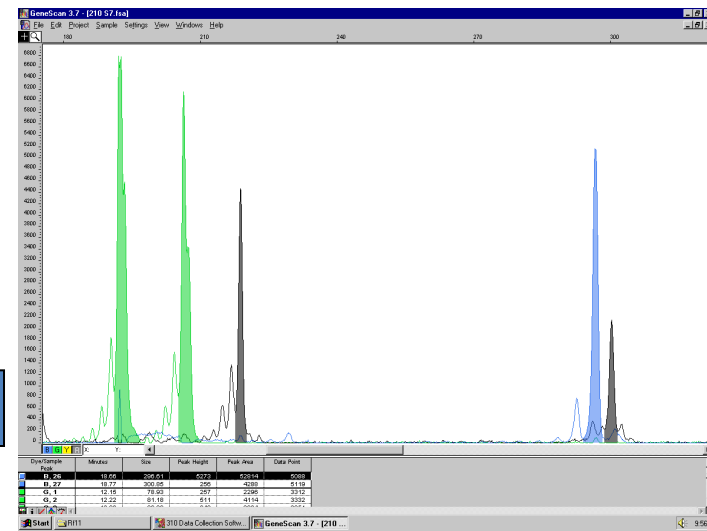
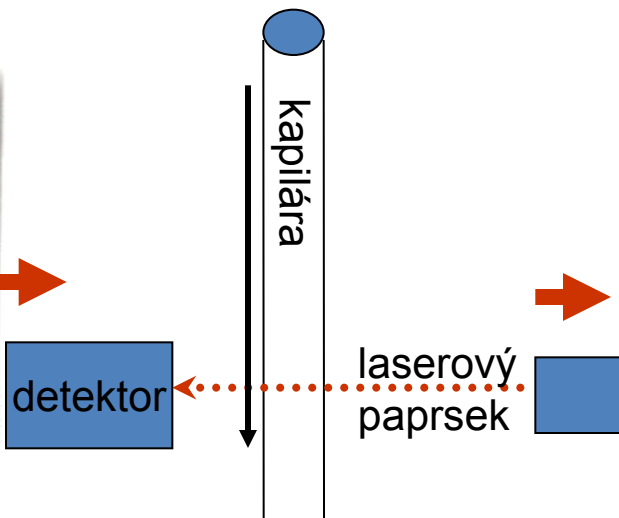
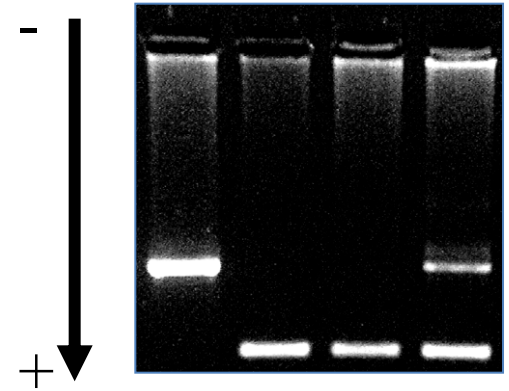
1) délkový polymorfismus

- elektroforéza DNA (DNA = záporný náboj)



Rozdělení fragmentů DNA podle velikosti („molekulové síto“)

- Agarosa - Hrubé rozdělení (do rozdílu 15 bp)
- Polyakrylamid – Přesnější rozdělení (4 bp)
- Sekvenátor, kapilární elektroforéza (fragmentační analýza) – nejpřesnější (fluorescenčně značené PCR fragmenty, např. značené primery)



Studium variability nasyntetizovaného úseku

2) sekvenční polymorfismus

- sekvencování
- SNP („single nucleotide polymorphism“) analýza – mnoho různých metod (viz další přednášky)



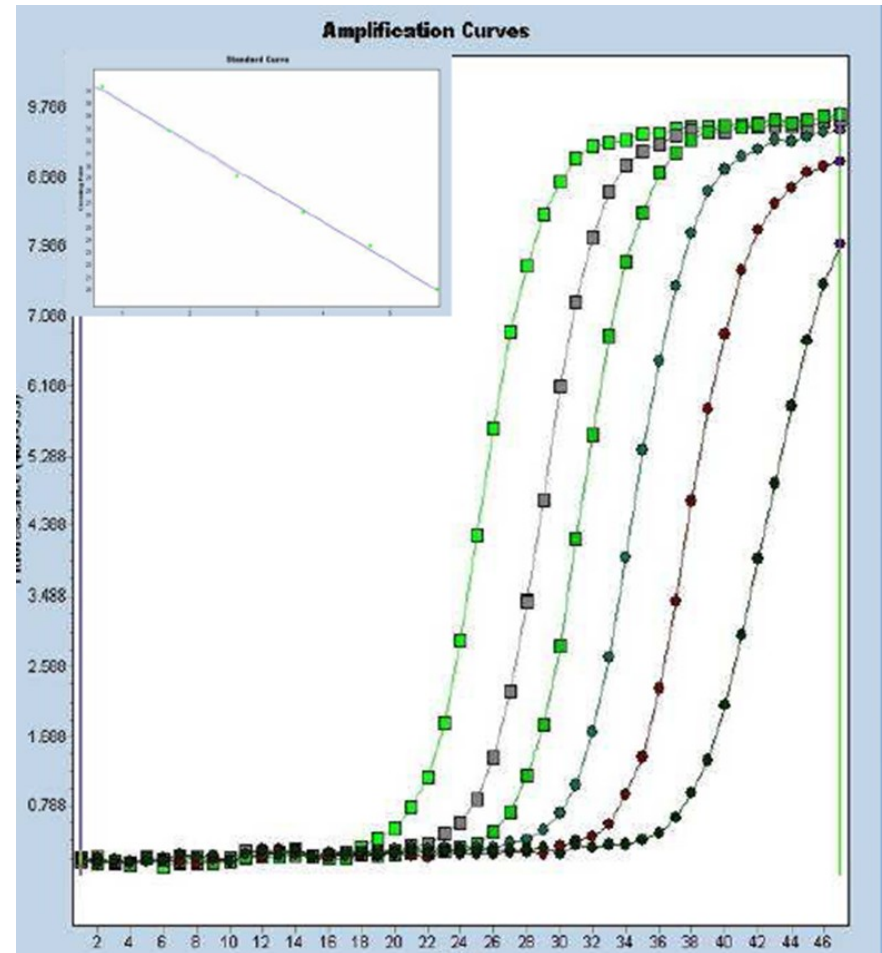
REAL TIME PCR
(= „kvantitativní PCR“)

Problém

- Kvantitativní rozdíly v expresi genů (tj. množství specifické mRNA určitého genu → reverzní transkripce → cDNA)
- Neinvazivní metody – nutnost stanovit, kdy je ještě ve vzorku dostatek DNA cílového druhu pro smysluplnou analýzu
- Genotypizace SNPs atd. (viz další přednášky)

„Real-time PCR“

- fluorescence je měřena v každém cyklu (signál odpovídá množství PCR produktu)
- křivky se zvedají po určitém množství cyklů, které odpovídá **počátečnímu** množství DNA
- srovnání s kalibrační křivkou umožňuje kvantifikaci



Fluorescenční strategie

Nespecifická detekce

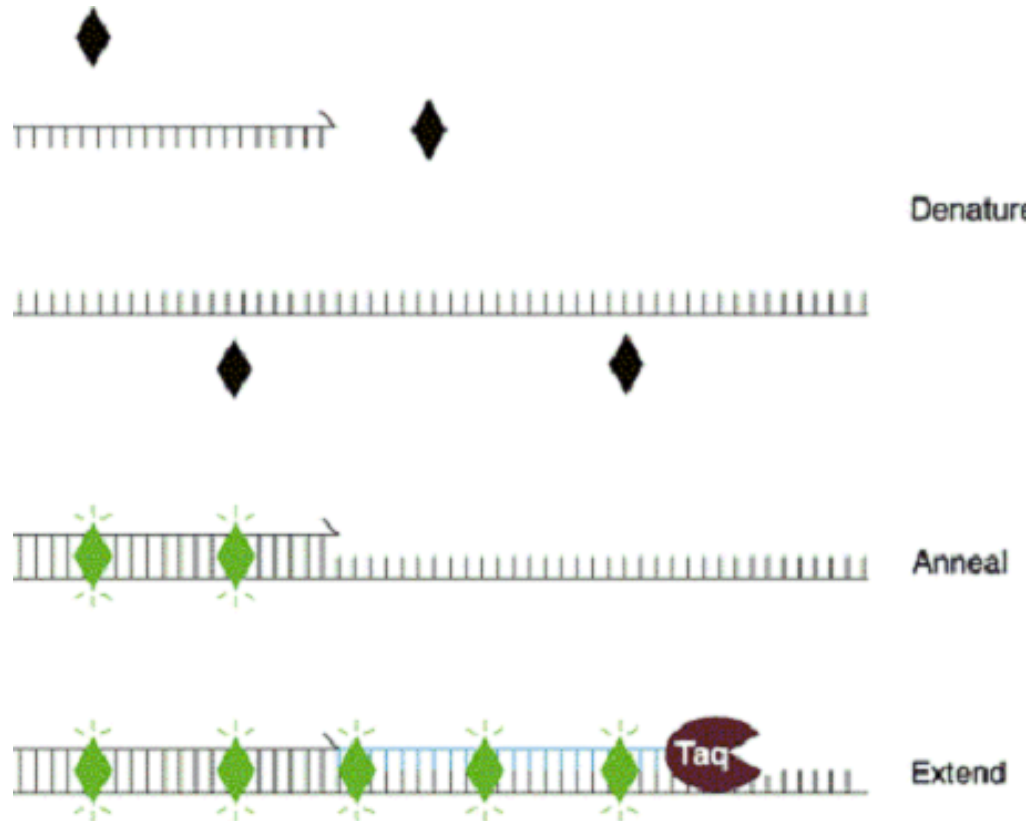
- SYBR Green, EVA Green, ...

Specifická detekce (vyšší přesnost = specificita k amplifikovanému úseku)

- hydrolyzační sondy (TaqMan)
- hybridizační sondy (FRET, Molecular beacon)
- others ...

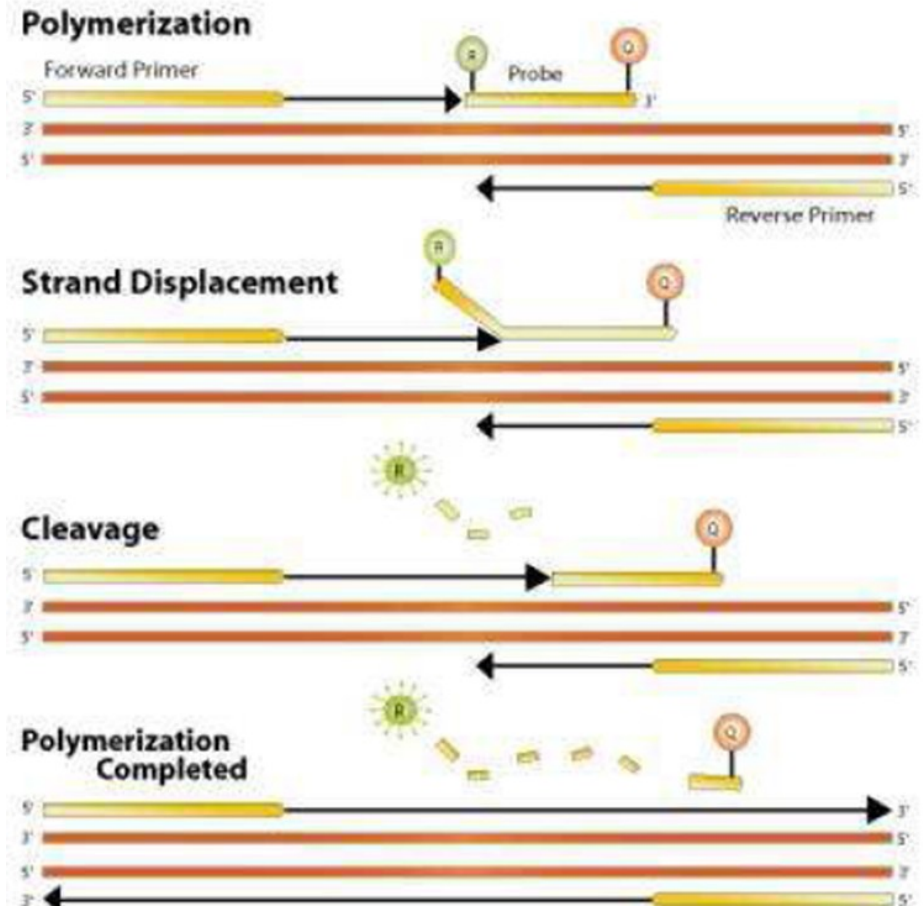
SYBR Green / EVA Green

- „barvička“ po inkorporaci (interkalaci) do dsDNA poskytuje zvýšenou fluorescenci



TaqMan hydrolyzační sondy

- intaktní sonda = žádná fluorescence („Quencher“ blokuje „Reporter“)
- 5'-exonukleázová aktivita DNA-polymerázy degraduje sondu → uvolnění fluorescence



„Molecular beacon“ hybridizační sondy

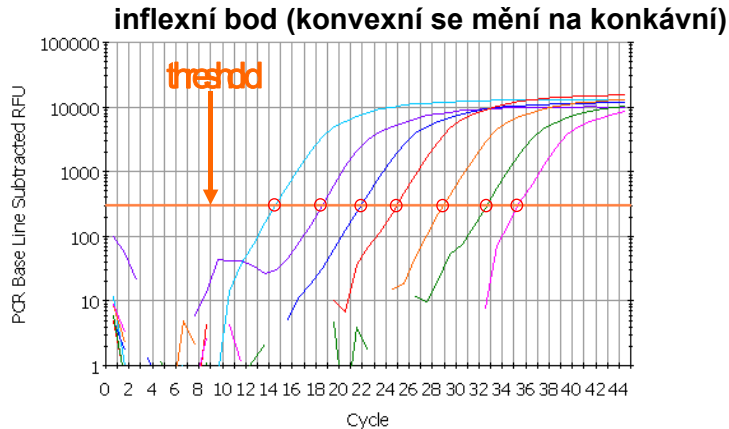
- specifická 3D struktura intaktní sondy („vlásenka“) – nevyzařuje žádnou fluorescenci
- „Quencher“ uvolní „Reporter“ po dosednutí na amplifikovaný úsek (v annealing fázi)



Real-time PCR přístroje

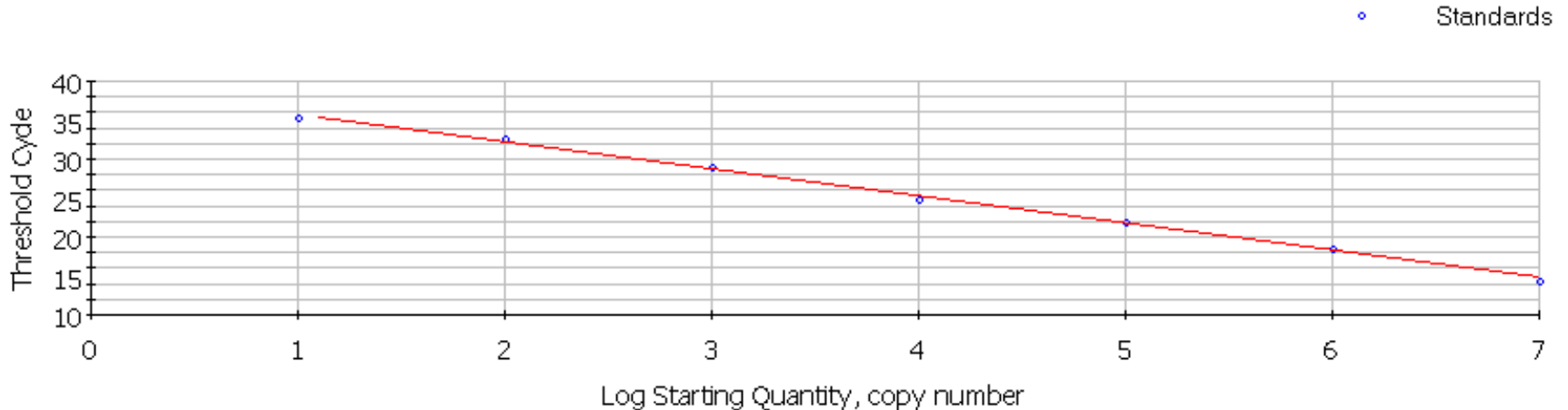


Absolutní kvantifikace



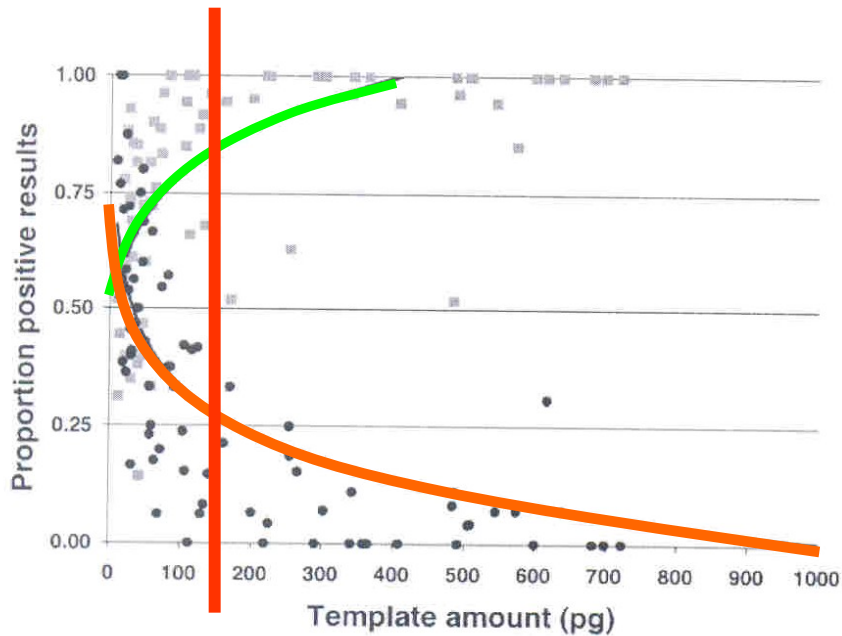
- 1) Vytvoření kalibrační křivky (vzorky se známou koncentrací templátu)
- 2) Real-time PCR se vzorkem s neznámým množstvím DNA, např. z trusu

Correlation Coefficient: 0.999 Slope: -3.488 Intercept: 39.204 $Y = -3.488 X + 39.204$

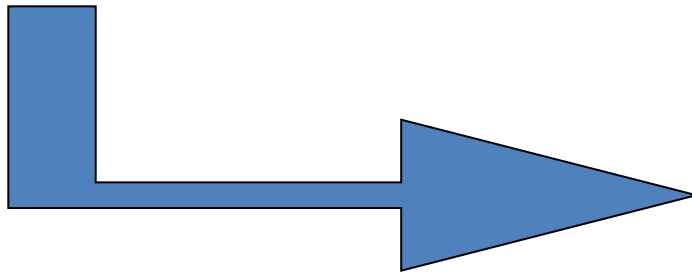


PCR Standard Curve: Data 27-Jan-03 1233ileff.opd

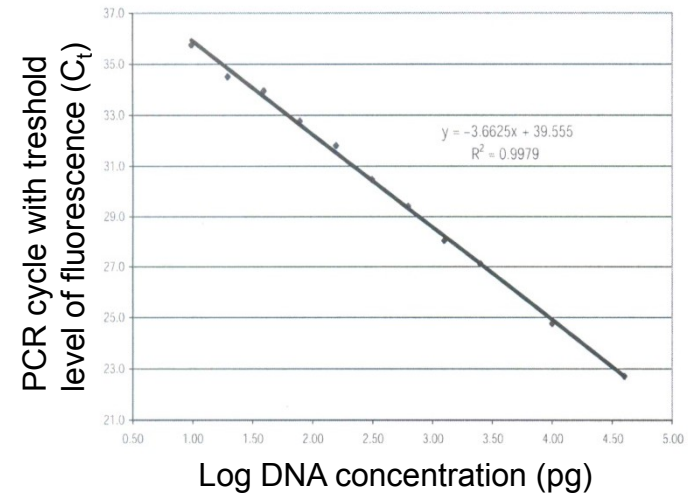
Stanovení koncentrace DNA při neinvazivních analýzách



Positive PCR Allelic dropout



Genotypizace jen „dobrých“ vzorků



Relativní kvantifikace - standardy

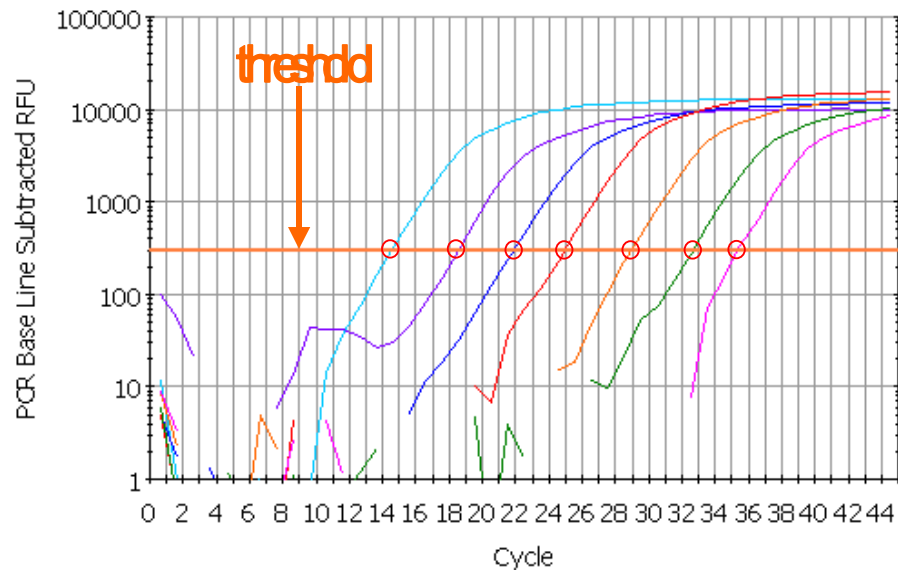
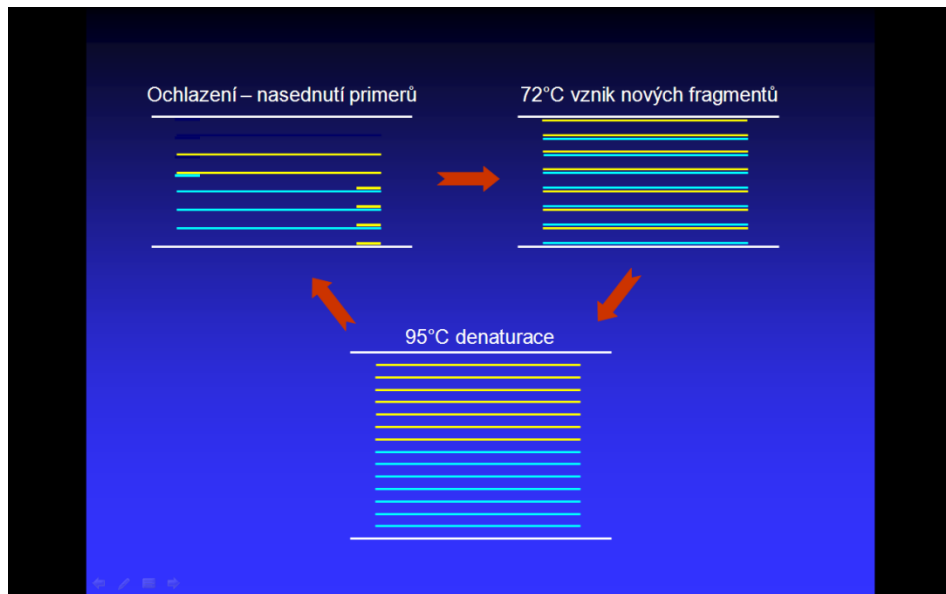
- Měření kvantity určité DNA, např. úroveň exprese (tj. určité RNA=cDNA) nebo množství mitochondrií (tj. mtDNA) (např. v různých typech tkání nebo treatment vs. non-treatment atd.)
- **housekeeping geny** – slouží jako standard pro měření
- stejný počet kopií ve všech buňkách (např. geny kódující cytoskelet)
- exprimované ve všech buňkách, nezávislé na experimentu
- v různých systémech takto fungují různé geny – nutno vždy znovu ověřit a optimalizovat

PCR a qPCR s hydrolyzační sondou - VIDEO

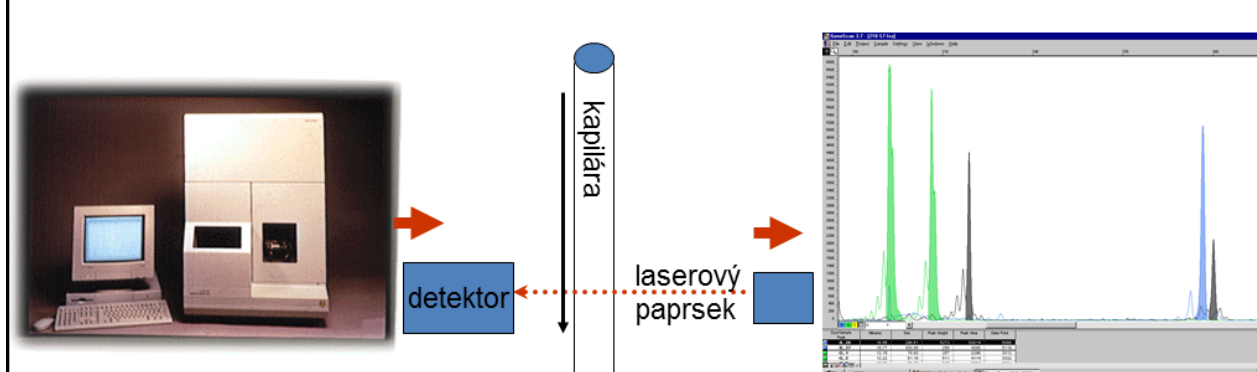
<https://www.youtube.com/watch?v=FIgGKkcLLuo>

PCR

qPCR



- Sekvenátor, fragmentační analýza – nejpřesnější (fluorescenčně značené PCR fragmenty, např. značené primery)



kapilární elektroforéza

Sangerovo sekvenování

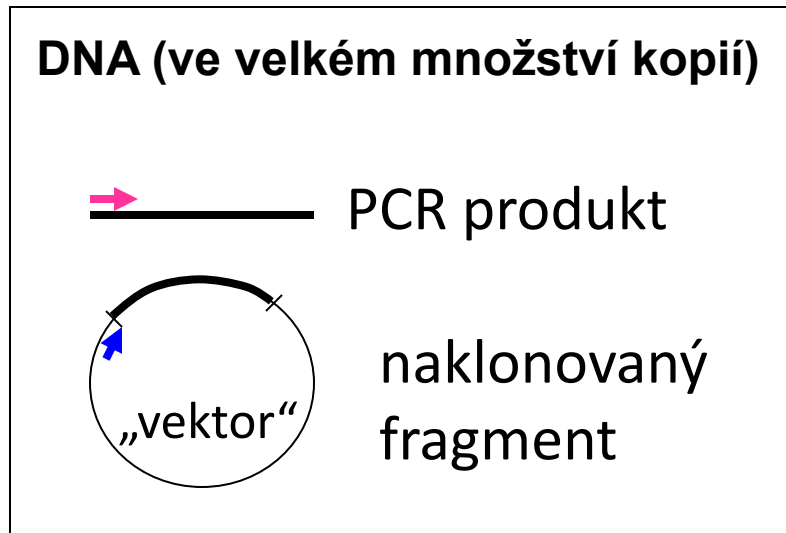
Sekvencování DNA

- ~~Maxam-Gilbertova (chemická) metoda: bázově-specifická chem. modifikace a štěpení fragmentů DNA~~

- Sangerova (enzymatická) metoda: terminace replikace pomocí ddNTP

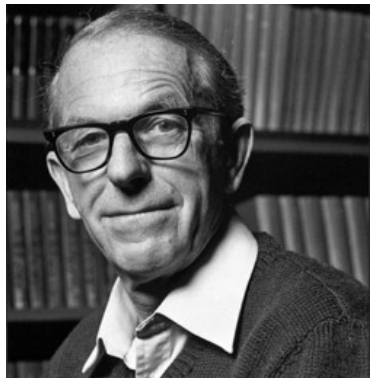
- paralelní „high-throughput“ sekvenování:
= NGS („next-generation sequencing“)

Sangerovo sekvenování DNA



Sekvenační reakce:

- směs standardních nukleotidů a značených dideoxynukleotidů
- **jeden specifický** nebo **universální** primer – poskytuje volný 3 -konec



„dideoxy metoda“

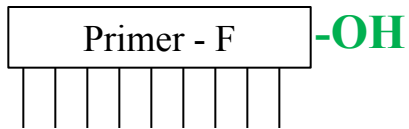
Frederick Sanger (1918-2013)

Nobel prize 1958 (struktura inzulinu) a 1980 (sekvenování bílkovin a nukleových kyselin)

- **jen jeden primer**

- vysoká koncentrace templátu (hodně kopií – buď PCR nebo klony v bakteriích)

- směs deoxynukleotidů a fluorescenčně značených dideoxynukleotidů



1. Denaturace - 96°C

2. Nasednutí primeru - 50-60°C

Sekvenační „PCR“ - jen jeden primer



Přisedání deoxynukleotidů ...



1. Denaturace - 96°C

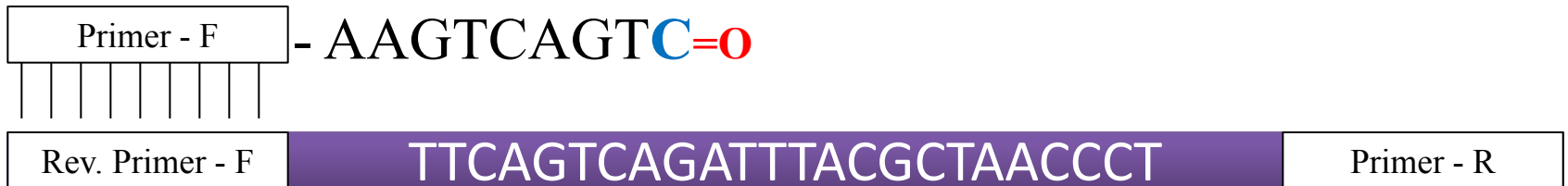
3. Polymerizace - 72°C

2. Nasednutí primeru - 50-60°C

Sekvenační „PCR“ - jen jeden primer



Přisedání deoxynukleotidů ...
... až narazí na dideoxynukleotid



1. Denaturace - 96°C

3. Polymerizace - 72°C

2. Nasednutí primeru - 50-60°C

Sekvenační „PCR“ - jen jeden primer

Primer - F - AAGTCAGTC=O

Primer - F - AAGTCAGTCTAA=O

Rev. Primer - F

TTCAGTCAGATTACGCTAACCT

Primer - R

1. Denaturace - 96°C

2. Nasednutí primeru - 50-60°C

3. Polymerizace - 72°C

35 x

Sekvenační „PCR“ - jen jeden primer

Primer - F - AAGTCAGTC**C=O**

Primer - F - AAGTCAGTCTA**A=O**

Primer - F - AAG**T=O**

Rev. Primer - F

TTCAGTCAGATTACGCTAACCT

Primer - R

1. Denaturace - 96°C

2. Nasednutí primeru - 50-60°C

3. Polymerizace - 72°C

35 x

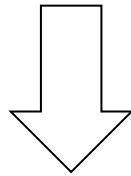
Výsledek sekvenační „PCR“

Primer - F - AAGTCAGTC**C=O**

Primer - F - AAGTCAGTCTAA**A=O**

Primer - F - AAG**T=O**

... a řada dalších fragmentů, každý z nich označený fluorescenčním dideoxynukleotidem



Kapilární elektroforéza

- seřazení fragmentů podle délky
- detekce barvy dideoxynukleotidů

Detekce fragmentů - kapilární elektroforéza

Primer - F - AAGTCAGTCTAA**A**=0

Primer - F - AAGTCAGTCT**A**=0

Primer - F - AAGTCAGTCT**T**=0

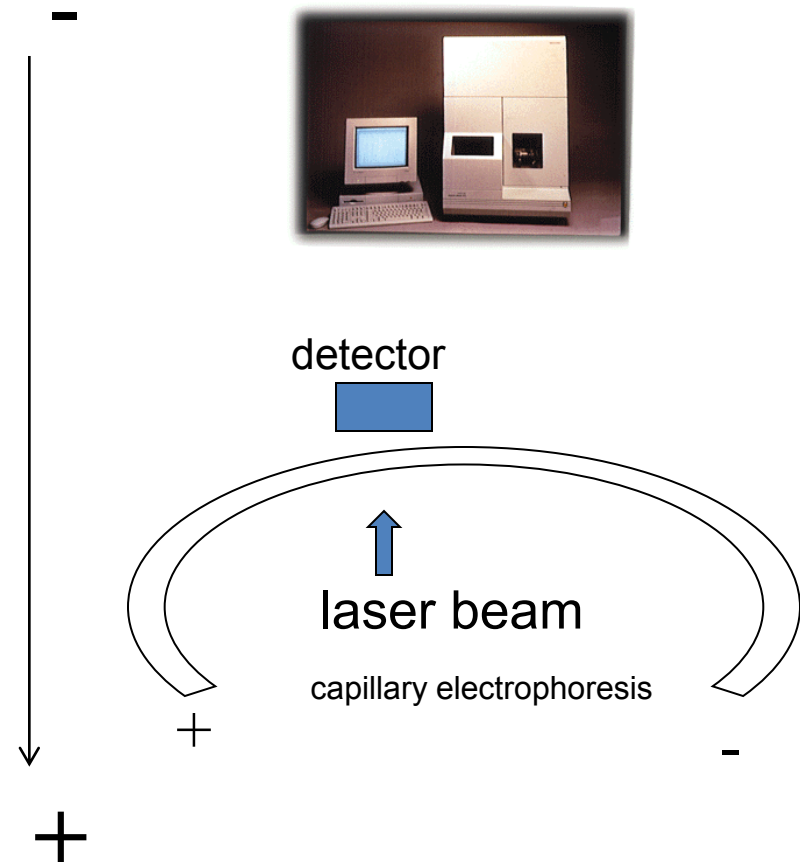
Primer - F - AAGTCAGT**C**=0

Primer - F - AAGTCAG**T**=0

Primer - F - AAGTCAG**G**=0

Primer - F - AAGTC**A**=0

Primer - F - AAGTC**C**=0



detector



laser beam

capillary electrophoresis

+

-

+

Detekce fragmentů - kapilární elektroforéza

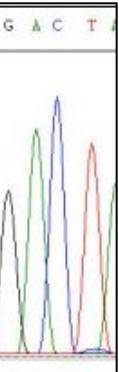
Primer - F

- AAGTCAGTCTAA**A**=0 -

Primer - F

- AAGTCAGTCT**A**=0

- Sekvence délky 500 – 1000 bp (cca 100 Kč za sekvenci, bez PCR a přečištění PCR produktu)
- 4 kapiláry - destička s 96 vzorky za noc
- Jsou i sekvenátory s 96 kapilárami



Primer - F

- AAGTC**C**=0

+

Primer - F

AAGTC**CAGTCTAA**ATGCGATTGGGA

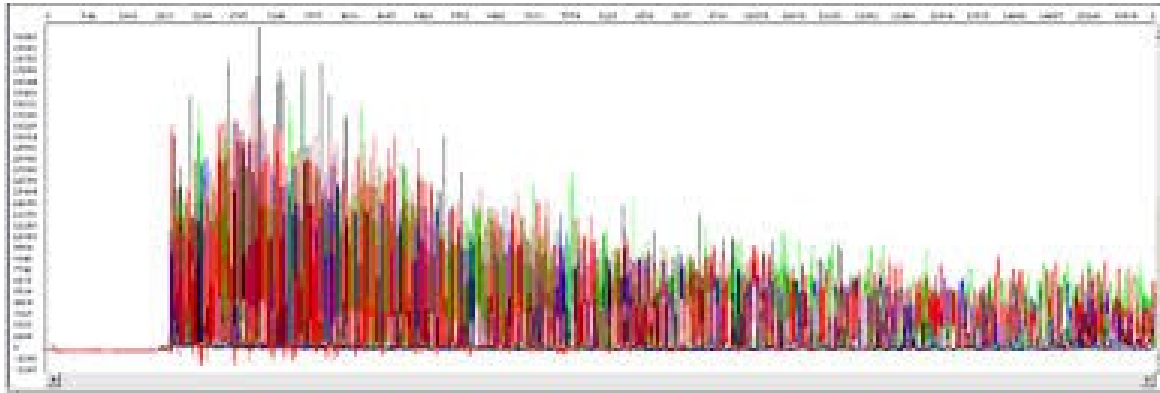
Rev. Primer - R

Rev. Primer - F

TTCAGTCAGATTACGCTAACCT

Primer - R

Editace sekvencí



„raw data“ (.ab1 file)
electrophoretogram



„basecalling“ (specializovaný software)

T T A A C C C A T T T C T C A A T A C T G A T T A T A C T G T G G G G A C G A A A G T C T C T G C T T T T A A C T A G
145 157 169 181 193

