

Sekvencování DNA

- stanovení pořadí nukleotidů v molekule DNA (primární struktury)

Sekven**c**ování / Sekvenování
??

Sequenc**ing** / -

die Sequenz**z**ierung / -

Techniky sekvencování

- 2 klasické metody:
 - **Chemická (Maxamova-Gilbertova)**
 - Již se nepoužívá
 - **Enzymová (Sangerova)**
 - V současnosti se používá automatická varianta
 - společný rys obou metod: příprava a separace fragmentů DNA, jejichž velikost se liší o 1 nukleotid
- Několik nových vysoce účinných metod nové generace (**Pyrosekvencování, SOLiD, Solexa, aj.**)

Základní požadavky pro úspěšné stanovení sekvence:

- ❖ Příprava knihovny pro sekvencování - DNA s přesně definovanými konci
 - s místem pro vazbu primeru
 - s koncovou značkou
- ❖ Příprava souboru všech fragmentů ssDNA lišících se svou délkou o jeden nukleotid
- ❖ Přesné elektroforetické rozdělení těchto fragmentů na základě jejich délky

Postup enzymatického sekvencování DNA

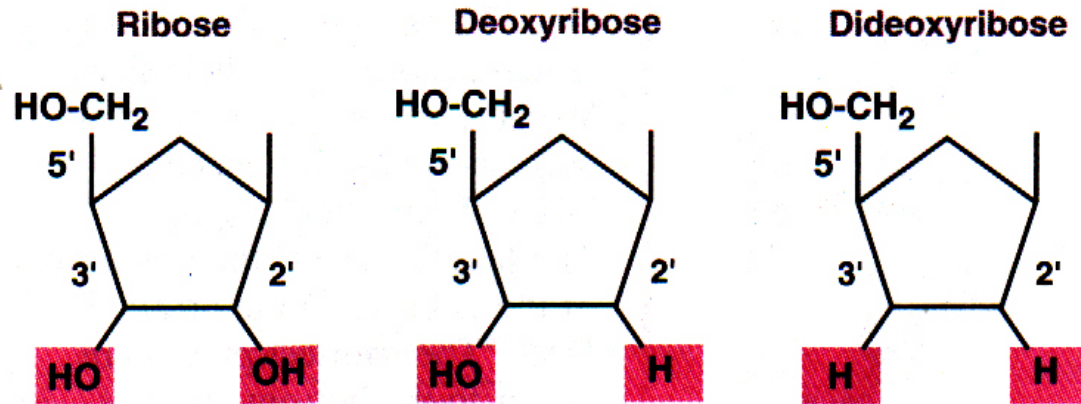
1. Příprava ssDNA jejíž sekvence má být stanovena
2. Rozdělení do 4 vzorků
3. Reakce s DNA polymerázou při níž se začlení do syntetizované DNA místo normálního deoxyribonukleosidtrifosfátu jeho analog, který působí jako koncový inhibitor syntézy DNA - dideoxynukleosidtrifosfát (ddNTP).
V každém ze 4 vzorků vzniknou fragmenty, které jsou zakončeny příslušným ddNTP (ddCTP, ddGTP, ddATP a ddTTP).

Reakce obsahuje:

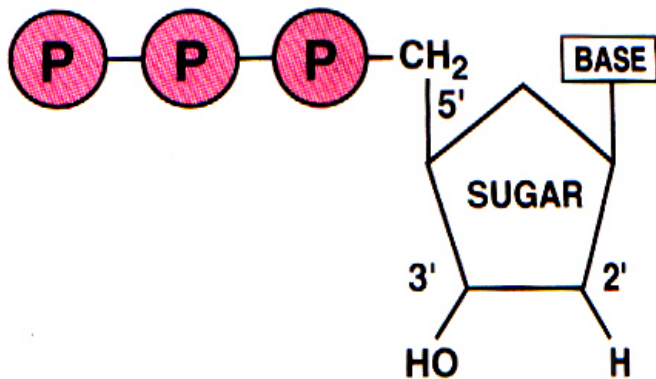
- molekulu DNA
- značený primer při pojující se k části molekuly DNA (místo odkud začínáme stanovovat sekvenci)
- směs obsahující 4 normální nukleotidy
- jeden ddNTP (v každém ze 4 vzorků je jiný)
- DNA polymerázu

4. Denaturace produktů
5. Elektroforetická separace umožňující oddělení fragmentů lišících se délkou o jednu bázi
6. Detekce fragmentů, které nesou označený konec

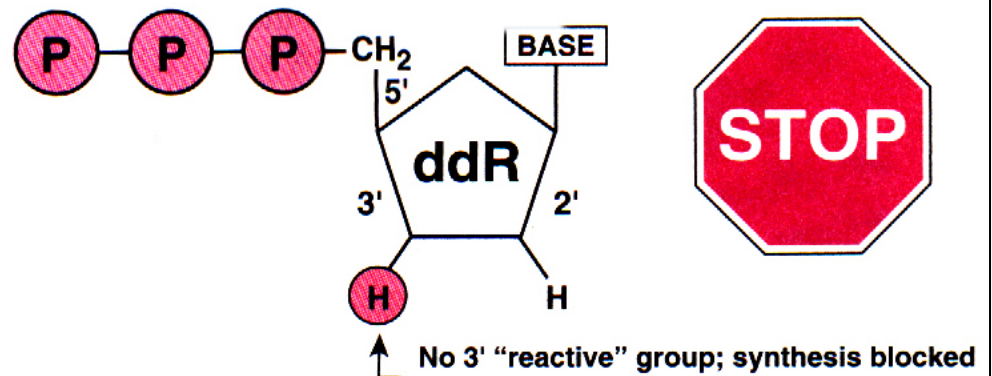
Dideoxynukleotidy nemají hydroxyl na 3'-konci



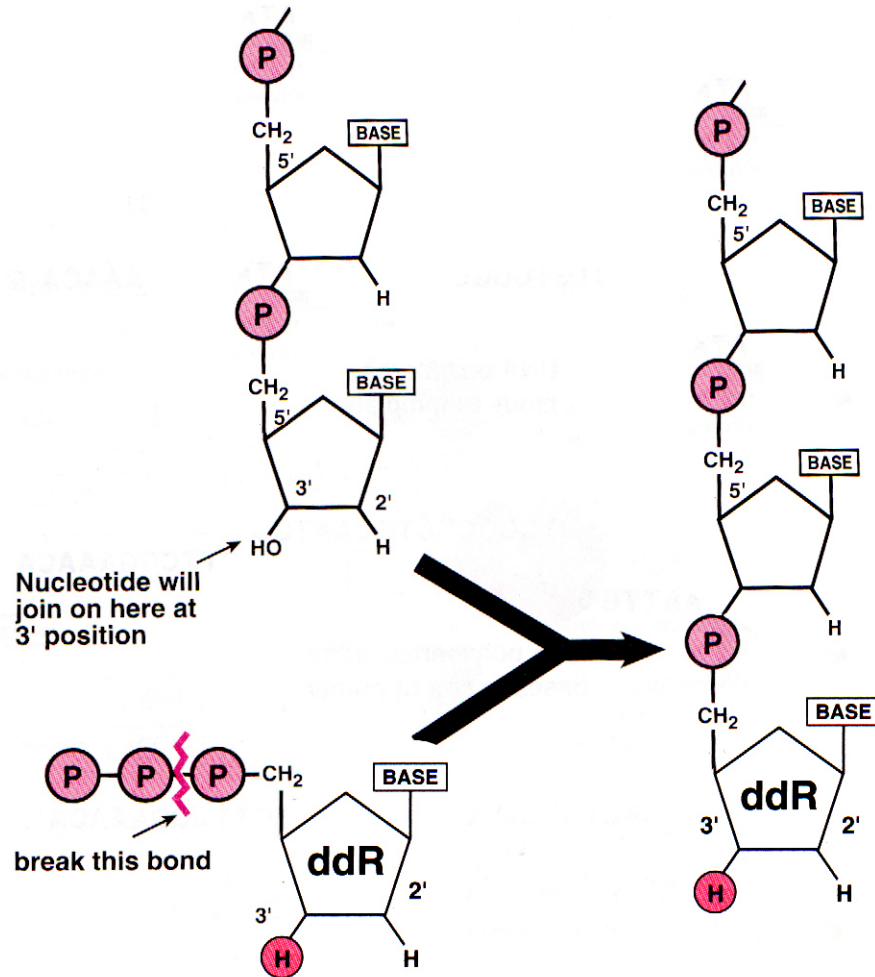
23.4 DEOXYNUCLEOSIDE TRIPHOSPHATE



23.7 DIDEOXYRIBOSE BLOCKS ELONGATION



Pokud je dideoxynukleotid inkorporován do syntetizujícího se řetězce, působí jako terminátor reakce



Replikace DNA s normálními deoxynukleotidy

Původní sekvence

A G C C T G G C G A C C A T C G T
T C G G A C C G C T G G T A G C A

Denaturovaný templátový řetězec

3' A G C C T G G C G A C C A T C G T 5'

Prfektní kopie

A G C C T G G C G A C C A T C G T
T C G G A C C G C T G G T A G C A

Co se stane pokud při reakci použijete směs dGTP a ddGTP?

Původní sekvence

A G C C T G G C G A C C A T C G T
T C G G A C C G C T G G T A G C A

Templátový řetězec

3' A G C C T G G C G A C C A T C G T 5'

Perfektní kopie

~~A G C C T G G C G A C C A T C G T
T C G G A C C G C T G G T A G C A~~

Směs řetězců ukončených GUANINEM
při použití směsi dGTP a ddGTP

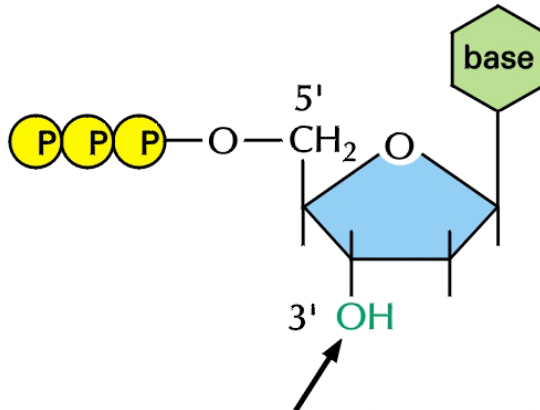
T C G
T C G G
T C G G A C C G
T C G G A C C G C T G
T C G G A C C G C T G G
T C G G A C C G C T G G T A G

Enzymová metoda sekvencování DNA (Sanger)

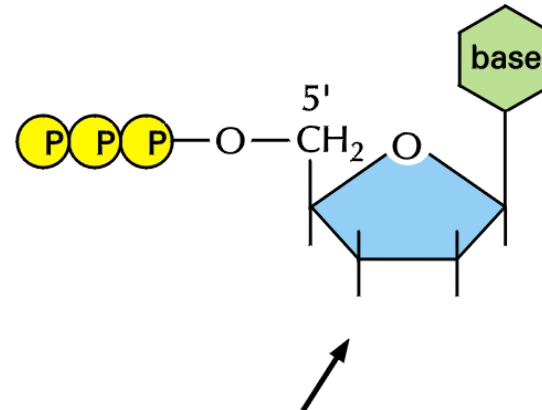
(A)

deoxyribonukleosidtrifosfát

dideoxyribonukleosidtrifosfát



je možno prodloužit řetězec na 3' -konci

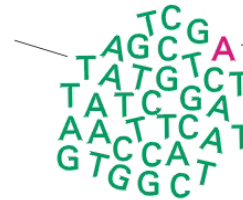


není možno prodloužit řetězec na 3' -konci

(B)

normální prekursorý deoxyribonukleosidtrifosfátů (dATP, dCTP, dGTP, dTTP)

malé množství jednoho dideoxyribonukleosidtrifosfátu (ddATP)



oligonukleotidový primer pro DNA-polymerázu



vzácná inkorporace dideoxyribonukleosidtrifosfátu DNA-polymerázou zastaví další růst molekuly DNA

jednořetězcová DNA, která bude sekvenována

(C)

5' GCATATGTCAGTCCAG 3'
3' CGTATACAGTCAGGTC 5' } dvouřetězcová DNA

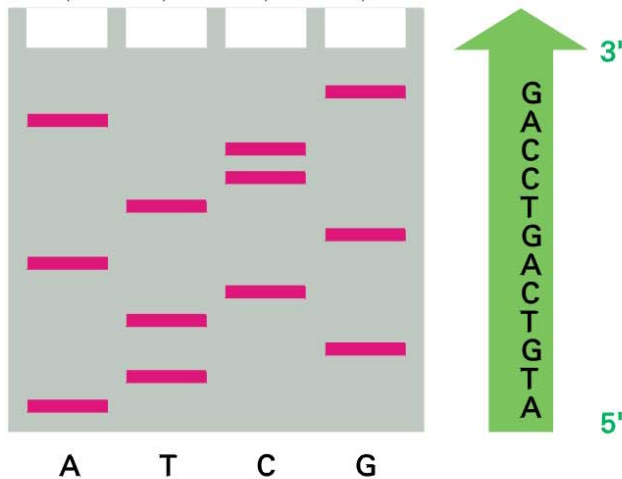
označený primer

5' GCAT 3'
3' CGTATACAGTCAGGTC 5' } jednořetězcová DNA

+ nadbytek dATP
dTTP
dCTP
dGTP

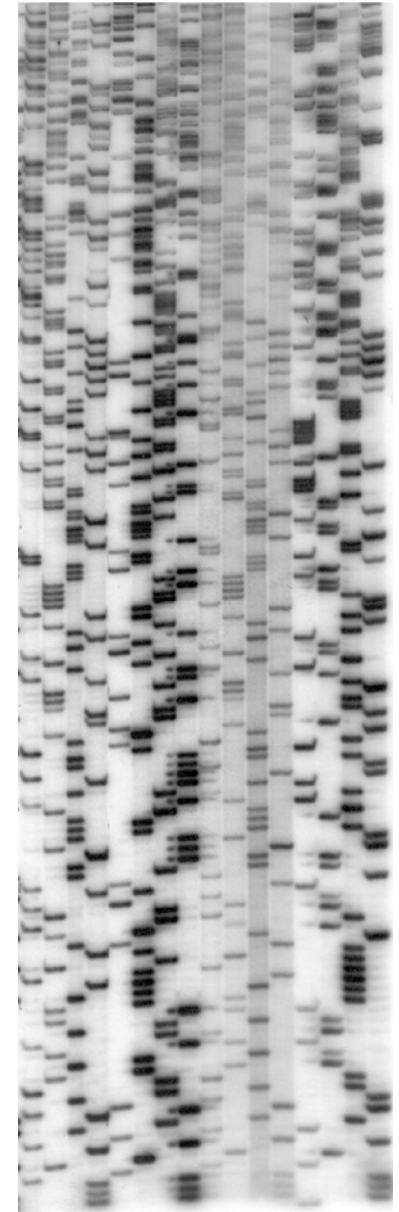
+ ddATP + DNA-polymeráza + ddTTP + DNA-polymeráza + ddCTP + DNA-polymeráza + ddGTP + DNA-polymeráza

GCAT A GCAT AT GCAT ATGTC GCAT ATG
GCAT ATGTCA GCAT ATGT GCAT ATGTCAGTC GCAT ATGTCAG
GCAT ATGTCAGTCCA GCAT ATGTCAGT GCAT ATGTCAGTCC GCAT ATGTCAGTCCAG



[seq4.mov](#)

Obraz autoradiogramu ze sekvenačního gelu

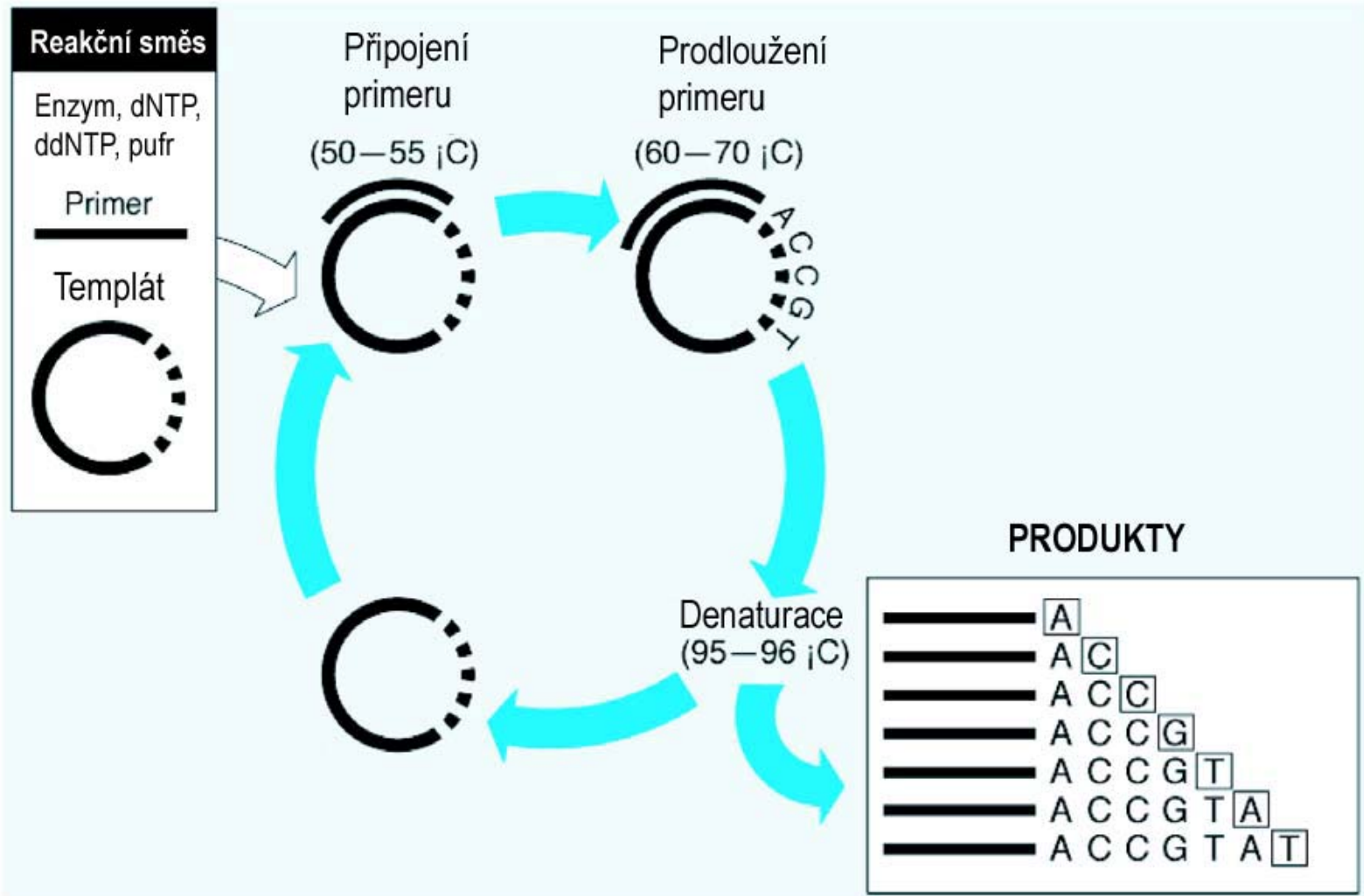


ATGC ATGC ATGC ATGC

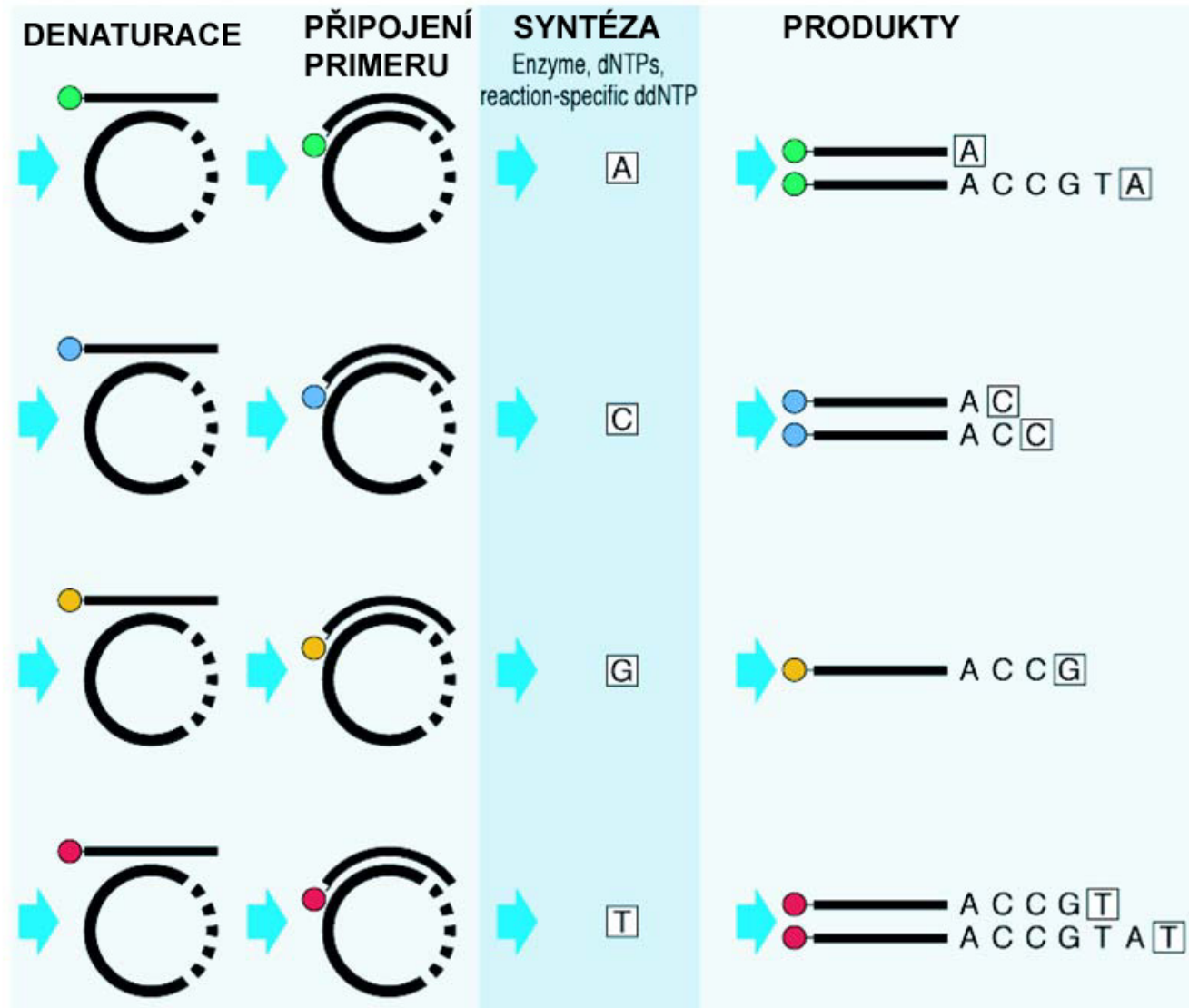
Automatické sekvencování DNA

- Je variantou enzymatického sekvencování DNA.
- Syntéza DNA probíhá v jedné reakci
- Ke značení produktů se používají čtyřmi různými fluorescenčními značkami označené
 - primery
 - dideoxyribonukleotidy

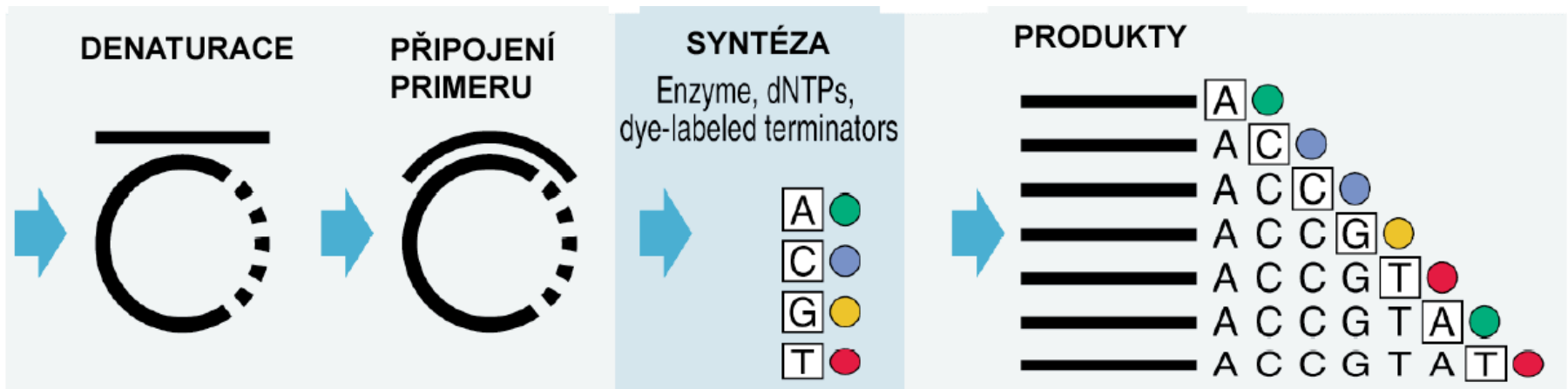
Asymetrická PCR pro sekvencování



Strategie barevných primerů

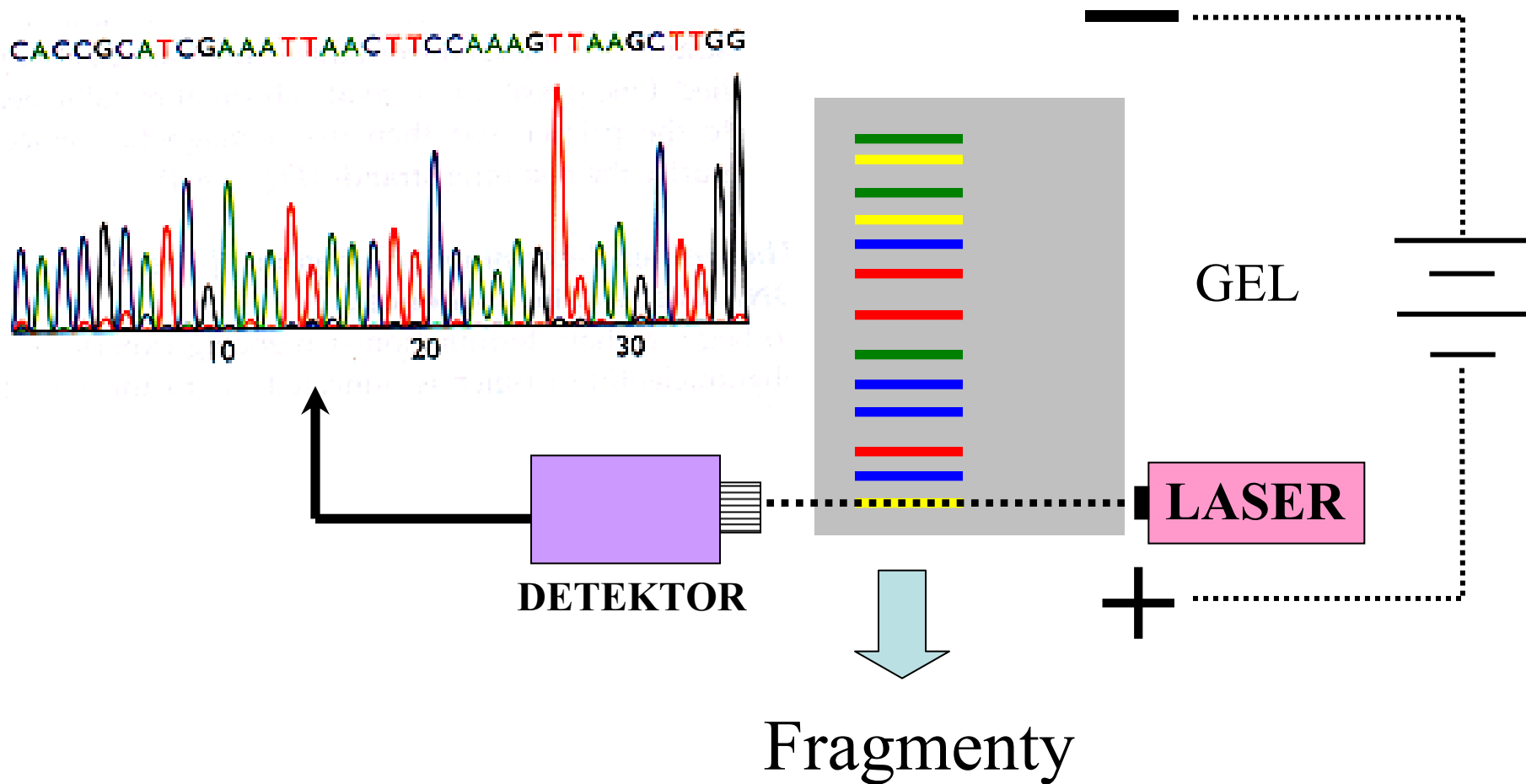


Strategie barevných terminátorů



Princip detekce produktů

- Detekce produktů probíhá během elektroforézy pomocí laserového detektoru (laserem indukovaná fluorescence, LIF) napojeného na počítač, který vyhodnocuje sekvenci.



BARVY:

- AMIDITOVÉ

HEX (černá)

6-FAM (modrá)

TET (zelená)

filtr

žebříček

C

TAMRA

- ESTEROVÉ

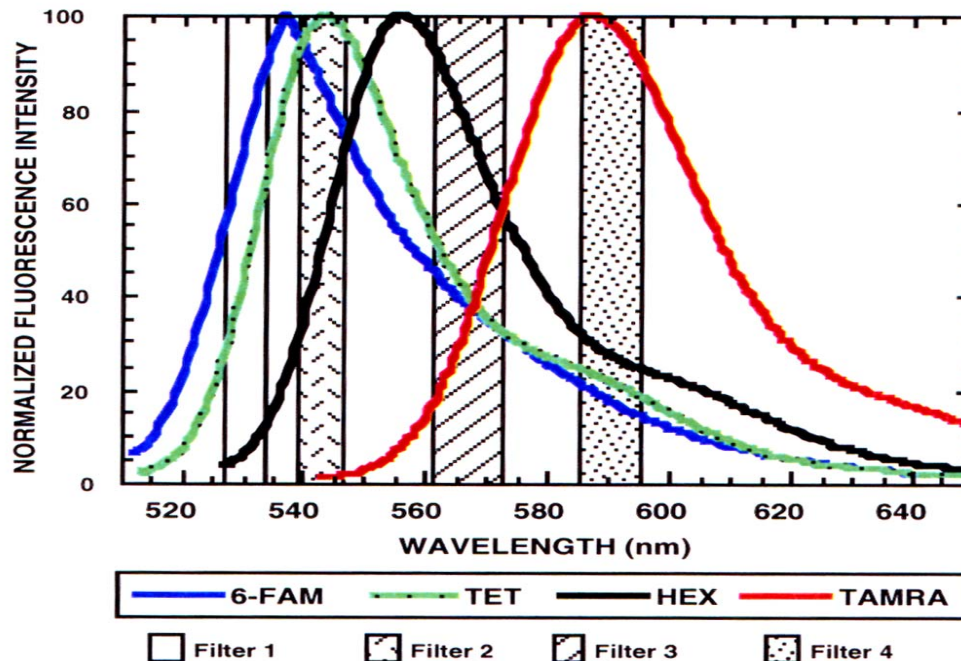
TAMRA (černá)

JOE (zelená)

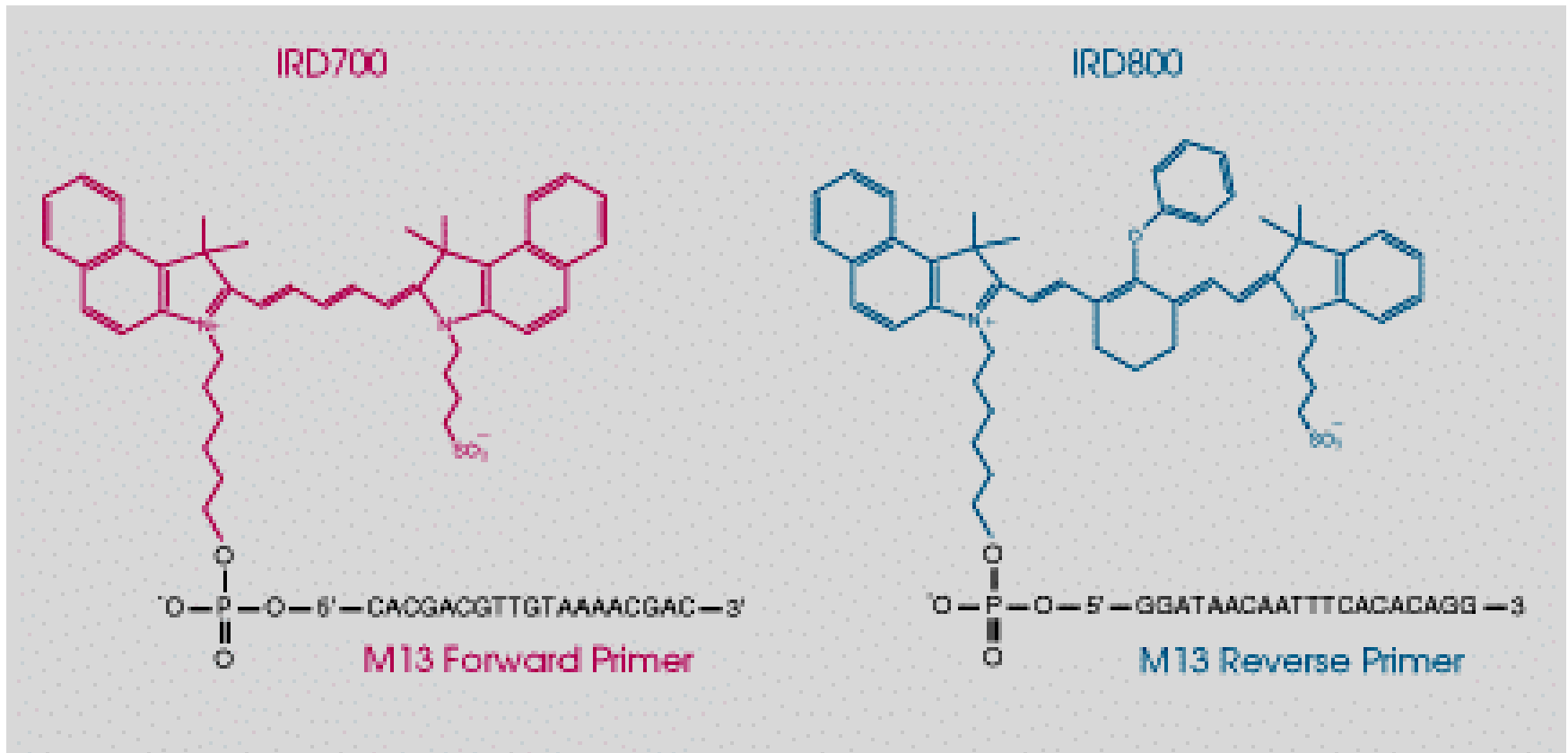
5-FAM (modrá)

A

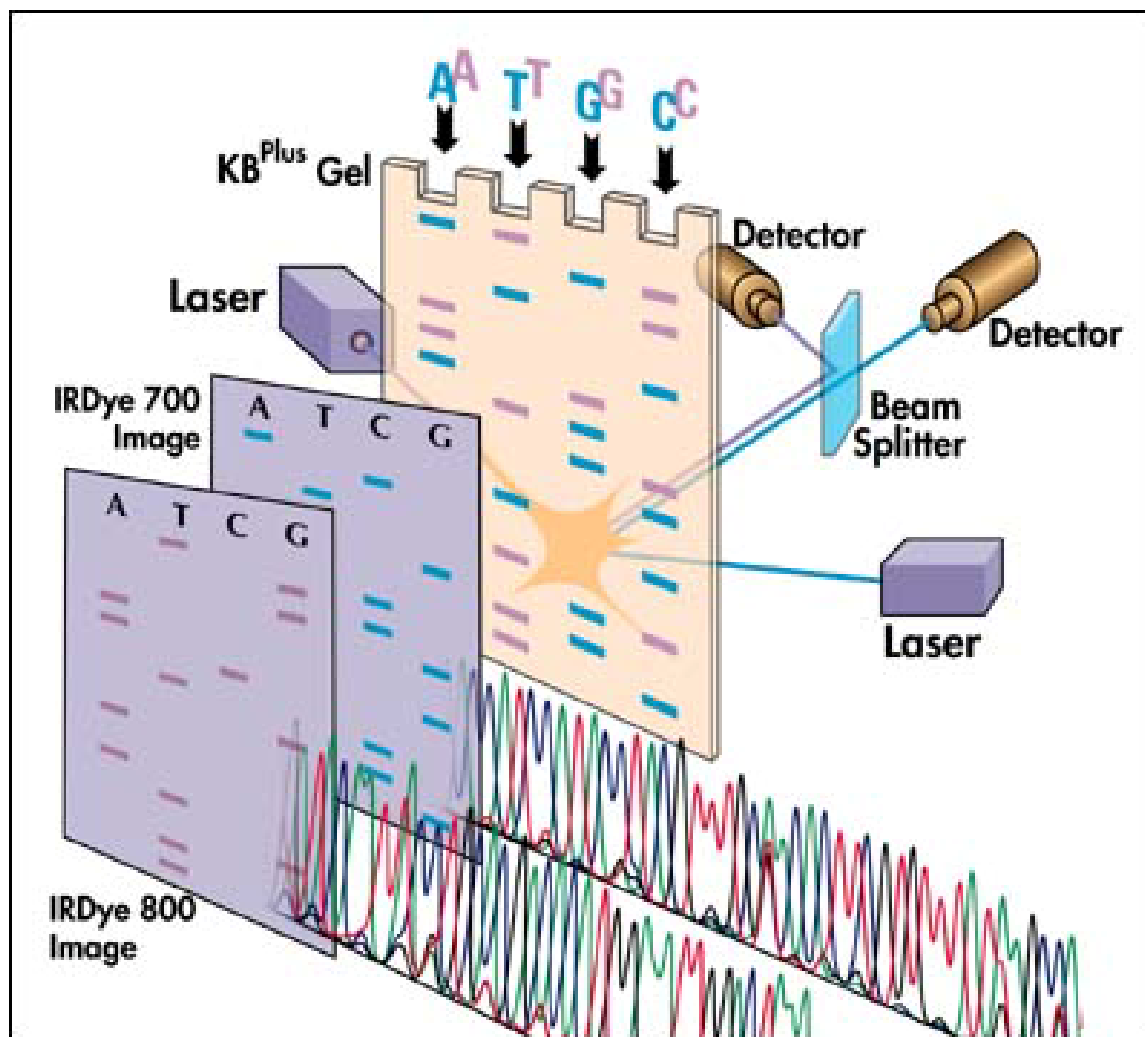
ROX



Příklad fluorescenčního značnického primeru



Vícenásobné detekce u různě značených primerů



Genetický analyzátor ABI 3100

Soustava
kapilár

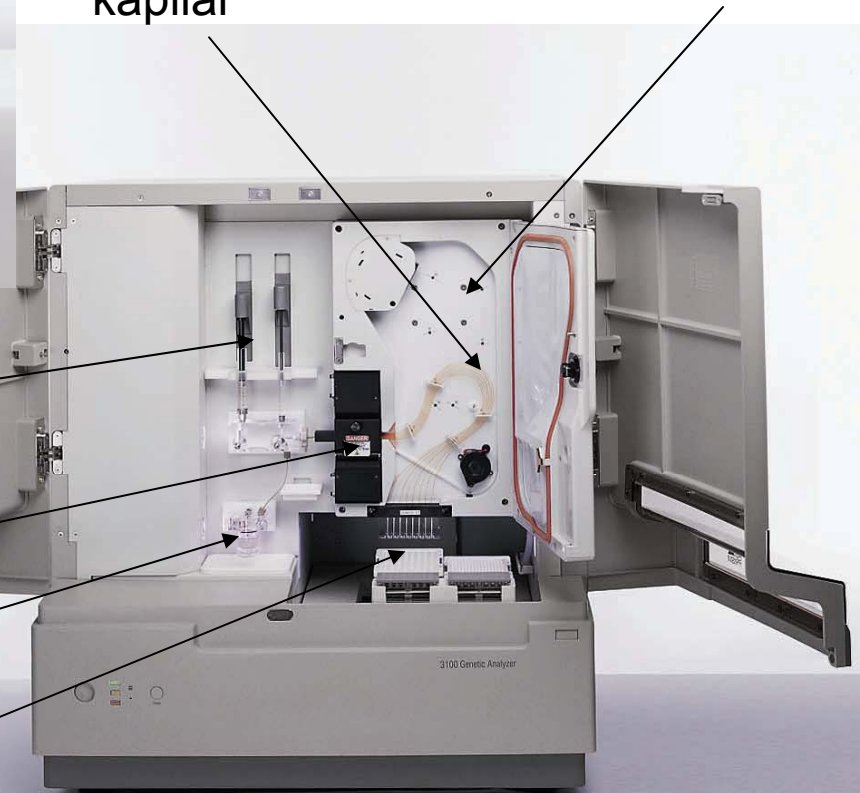
Vyhřívaná deska

Příprava gelu

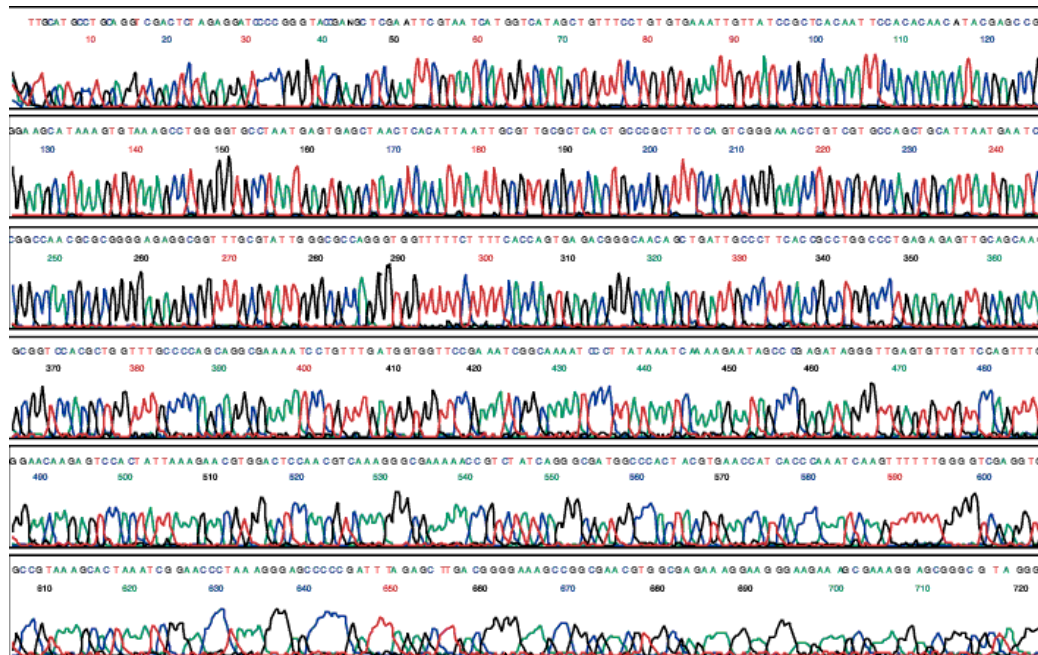
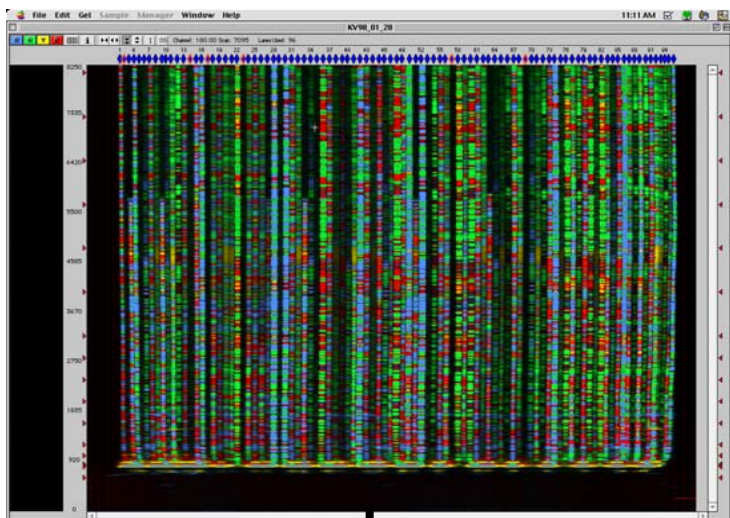
Laserový detektor

Rezervoár pufry

Automatické nanášení vzorku



Příklad výstupu



1 dráha na gelu

Sekvencování genomů

V praxi je velice často potřeba stanovit sekvenci fragmentu DNA, který je delší než průměrná délka 500 – 1 000 bází dosahovaná v jedné reakci.

K tomuto účelu sekvencování genomů mohou být zvoleny dvě zcela odlišné strategie:

- náhodné sekvencování
- uspořádané sekvencování sousedních úseků

Příprava knihovny pro náhodné sekvencování genomů

- Při náhodném sekvencování genomu jsou nejdříve připraveny mechanickým stříháním genomové DNA fragmenty s optimální velikostí (**1 300 – 2 000 bp**) a po úpravě jejich konců jsou nahodile naklonovány do vhodného vektoru za vzniku velkého počtu klonů.
- Ke štěpení DNA se využívá sonikace nebo pankreatická DNáza I za přítomnosti Mn^{2+} .

- Předpokládá se, že celková informace o sekvenci obsažená v připravených malých klonech odpovídá původní DNA a sekvence fragmentů jednotlivých klonů se vzájemně překrývají.
- Pomocí univerzálních sekvenačních primerů připojujících se k vektoru poblíž klonovacího místa jsou stanoveny sekvence krátkých úseků na koncích klonovaných fragmentů (minimálně 500 bází).
- Stanovené sekvence jsou pak použity k uspořádání klonovaných fragmentů z jednotlivých vektorů do souvislých sekvencí genomové DNA - kontigů.



Izolace DNA



Fragmentace

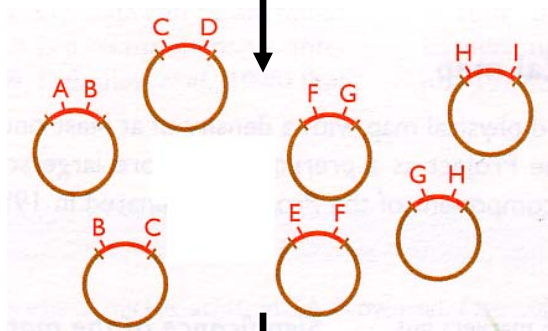


+

Vektor



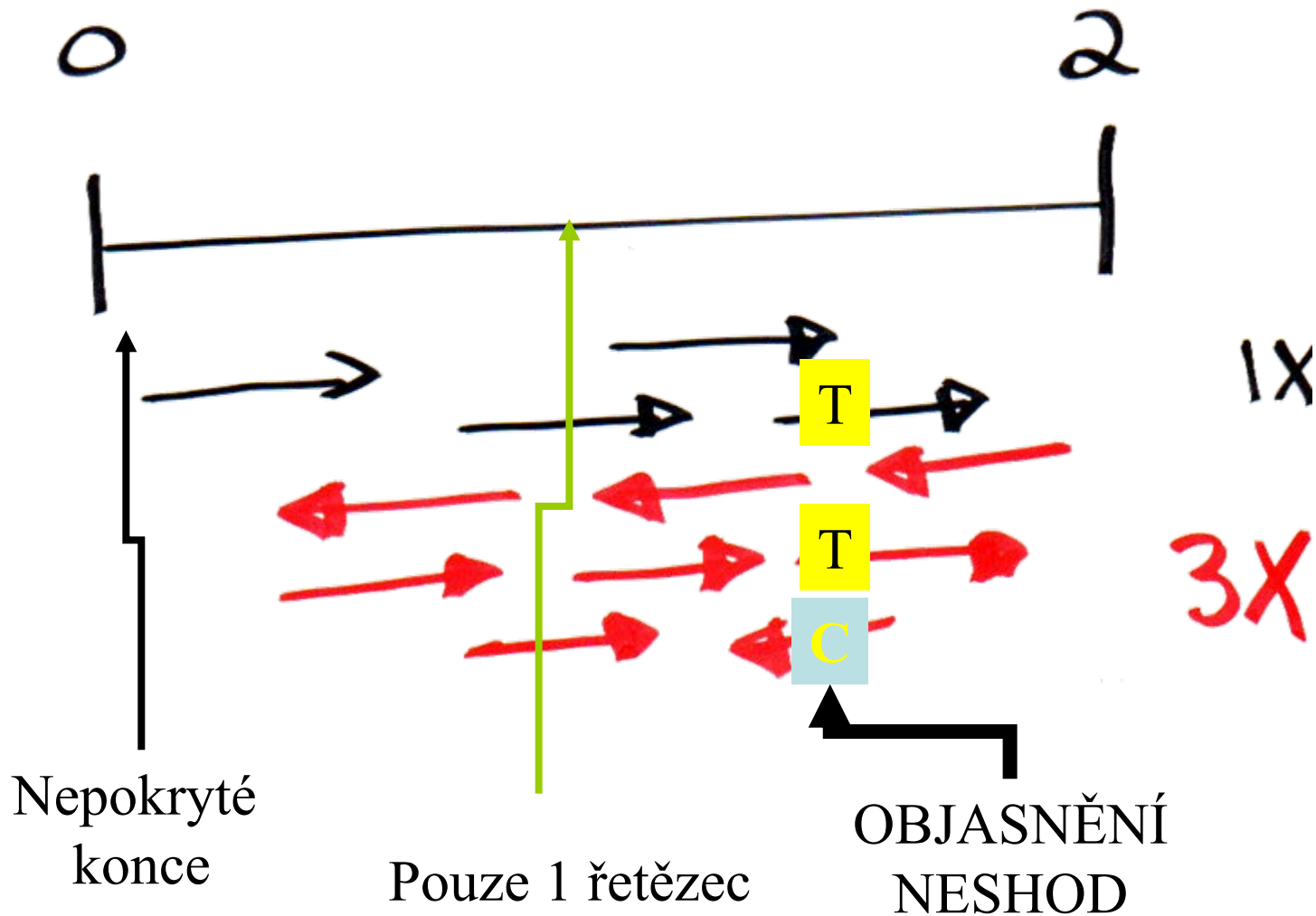
Úprava konců a klonování



Sestavení překrývajících se klonů



Náhodné sekvencování

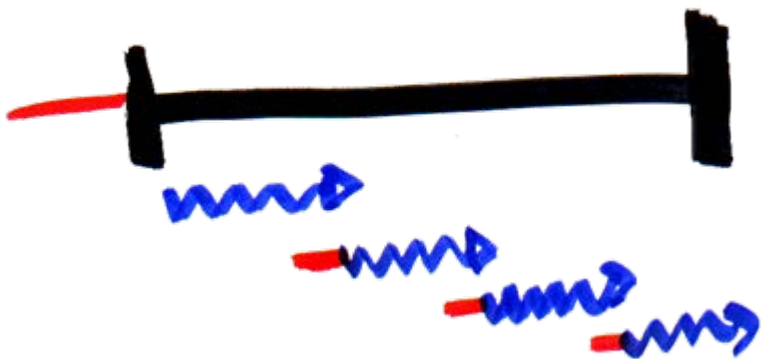


USPOŘÁDANÉ SEKVENCOVÁNÍ

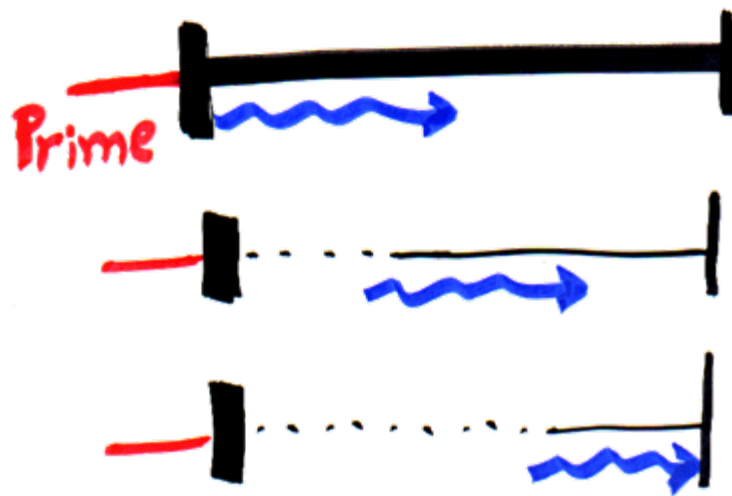
- Pro doplnění sekvence mezer zbývajících po náhodném sekvensování
- Pro sekvensování malých fragmentů DNA
- Dva odlišné přístupy:
 - **procházení primerem**
 - vyžaduje znalost sekvence, ke které se připojuje primer, který umožní prodloužení řetězce DNA-polymerázou.
 - získaná sekvence z první reakce je použita pro návrh primeru pro další reakci a tento krok se opakuje, dokud není dosaženo kompletního stanovení sekvence.
 - Tento proces nevyžaduje další klonování a minimalizuje stanovení nadbytečných sekvencí, ale správnost stanovené sekvence by měla být ověřena sekvensováním **obou řetězců**.
 - **postupné zkracování fragmentu exonukleázou a tvorba sousedících delecí**

Metody uspořádaného sekvencování

Procházení primerem

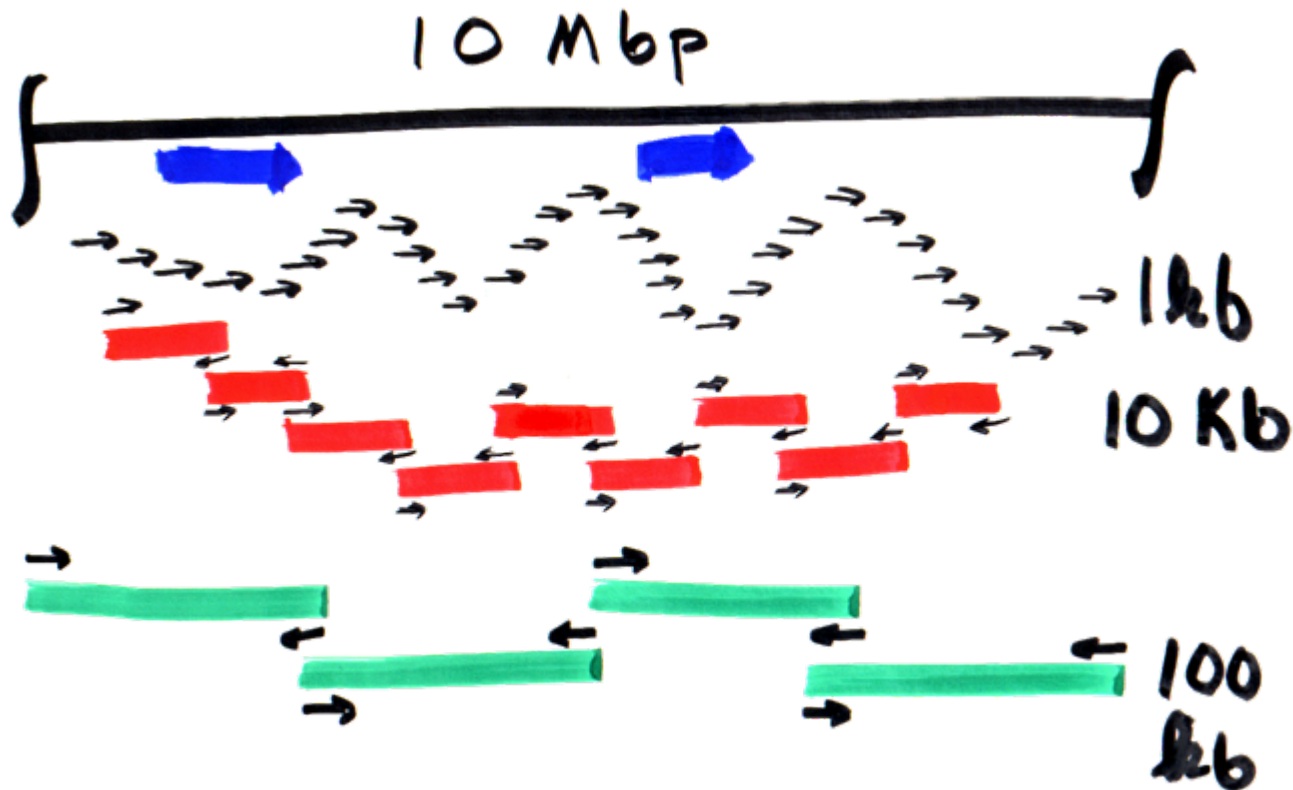


Sousední delece



DOKONČENÍ PROJEKTU

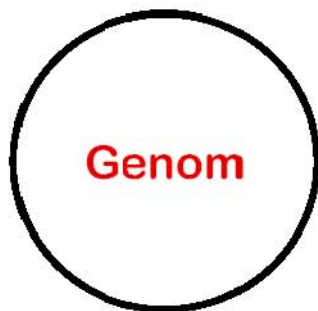
- Sestavení kontigů



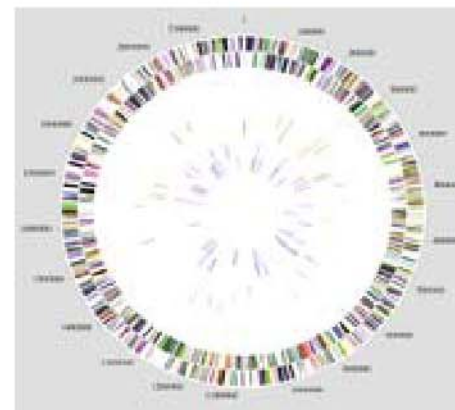
- Anotace (bioinformatika): ORF, repetice, regulační oblasti, → geny, → funkce



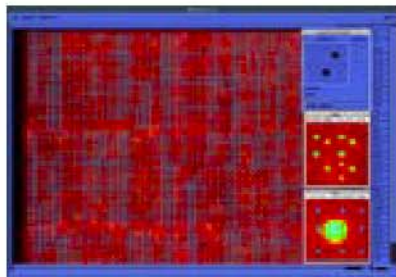
Sekvencování
Kompletace



Hledání ORF
Anotace

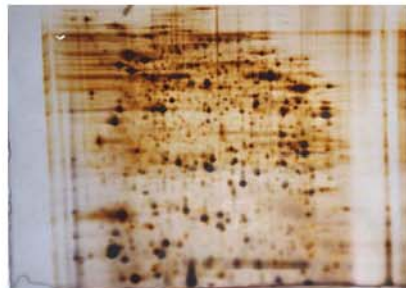


- Transkriptom
- Čipy

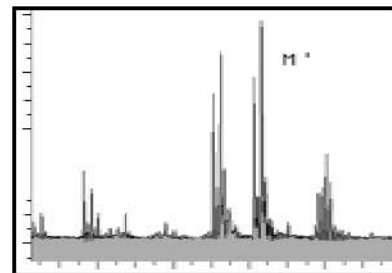


- Proteom

- 2D-Eelektroforéza



- MALDI-TOF-MS



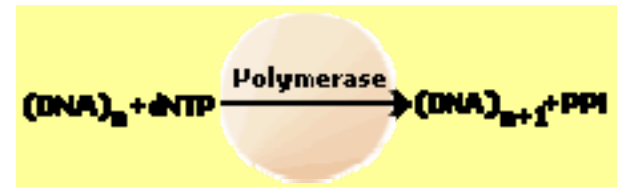
Nové přístupy pro stanovení sekvence DNA

- Pyrosekvencování (454 sekvencování)
- SOLiD
- Solexa
- Komparativní genomové sekvencování prostřednictvím hybridizace

Pyrosekvencování

- Sekvencování se syntézou DNA v reálném čase nevyžadující elektroforézu ani separaci fragmentů
- Je založené na uvolnění pyrofosfátu (PPi) při enzymatické syntéze DNA (Nyren et al., 1987)
- 1. krok: Sekvenační primer hybridizuje k jednořetězcovému templátu a je inkubován s:

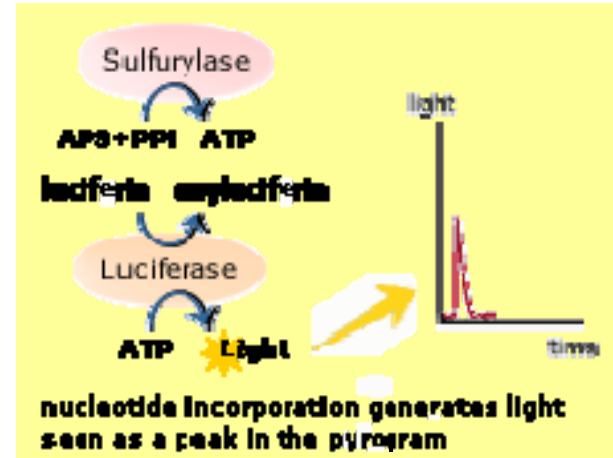
- DNA polymerázou
- ATP sulfurylázou
- Luciferázou
- Apyrázou
- substrátem, adenozin 5'-fosfosulfátem (APS)
- luciferinem



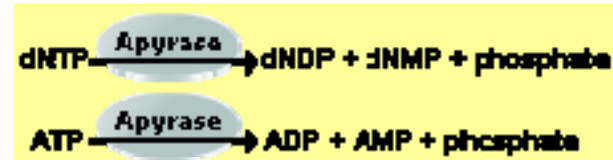
- 2. krok: První ze 4 dNTP – dATP je přidán k reakci
 - Pokud je na matrici komplementární báze, DNA polymeráza katalyzuje připojení nukleotidu k primeru
 - Pokud na matrici není komplementární báze nukleotid bude degradován apyrázou
 - Postupně budou přidány jeden po druhém všechny čtyři dNTP
 - Každé připojení nukleotidu je provázeno uvolněním pyrofosfátu (PPi) v množství ekvimolárním množství přidaného nukleotidu

- 3. krok: ATP sulfuryláza kvantitativně přeměňuje PPI na ATP za přítomnosti APS

- Vzniklý ATP umožní luciferázou zprostředkovanou konverzi luciferinu na oxyluciferin, který vytvoří světelný záblesk zaznamenaný detektorem fotonů a zobrazený jako pík na pyrogramu



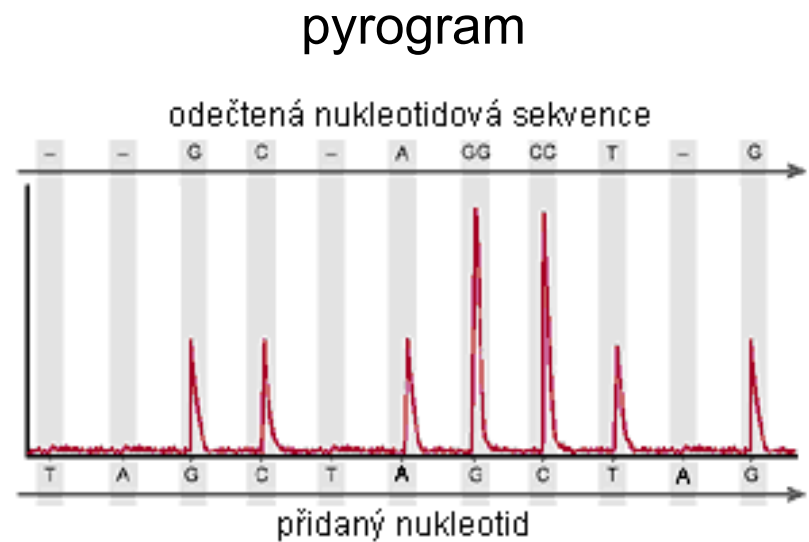
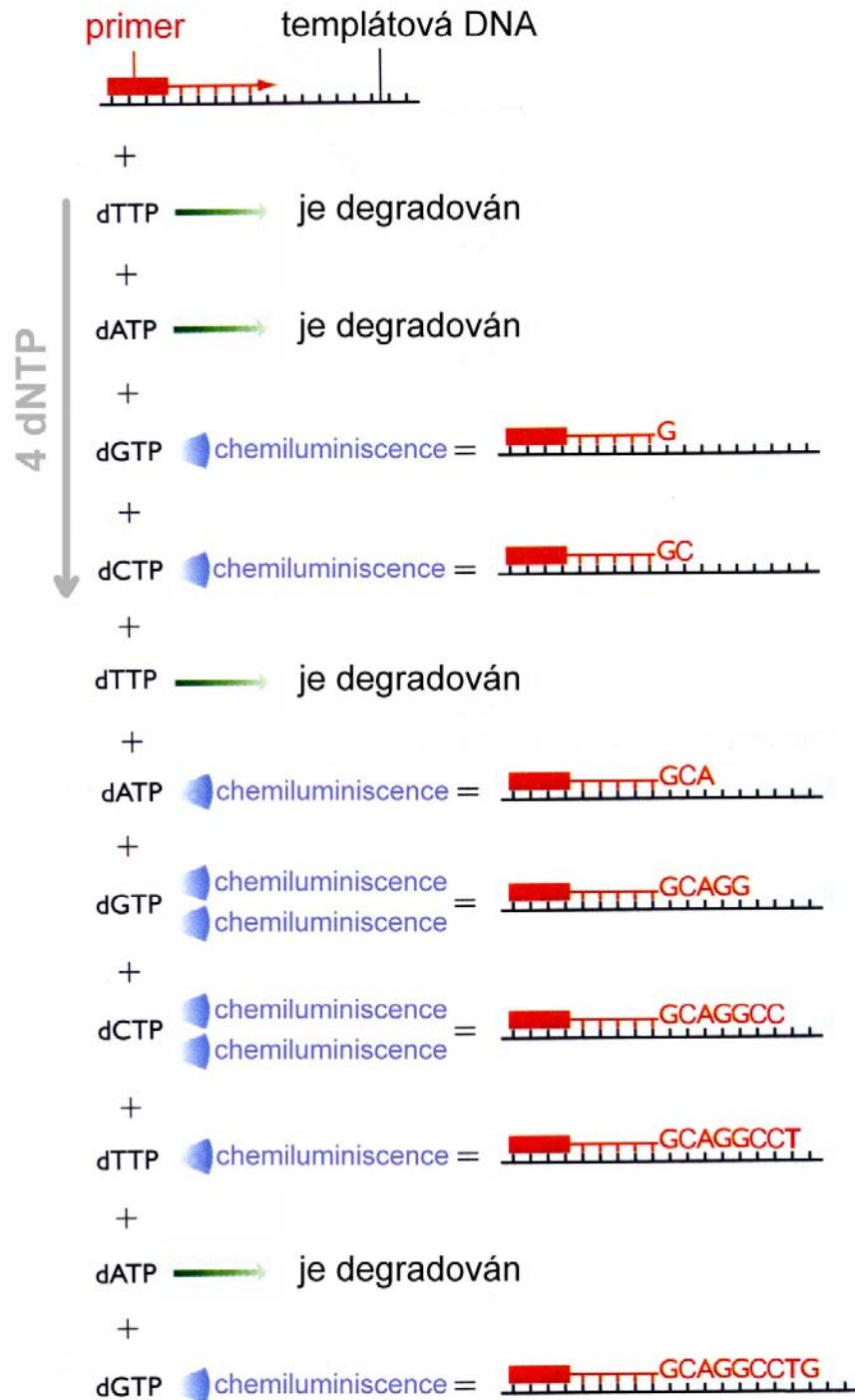
- 4. krok: Apyráza degraduje nepřípojené dNTP a nespotřebované ATP



- Až je degradace kompletní je přidán další dNTP, který je v pořadí

- Proces se znovu opakuje a sekvence je odečítána z pyrogramu
- Namísto standardního dATP je používán α -thiosubstituovaný dATP, který je přijímán DNA polymerázou, ale nikoli luciferázou
- Metoda je ve vývoji a je používána pro identifikaci jednonukleotidových polymorfizmů (SNPs)

Princip pyrosekvencování



Pyrosekvencování

(GS20 firmy 454 Life Sciences, rok 2005)

Výhody

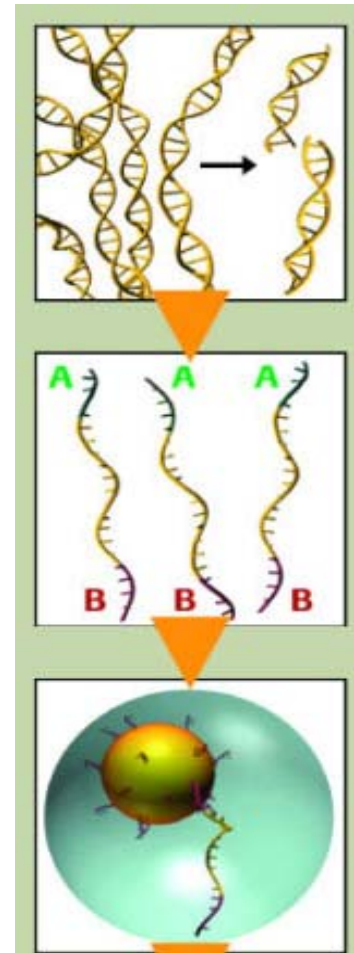
- Vysoká přesnost
- Flexibilita a možnost paralelního zpracování velkého množství vzorků
- Snadná automatizace
- Nevyžaduje elektroforézu a značené primery

Nevýhody

- Rychlost pyrosekvencování je cca 1 odečtená báze/min.
- Maximální délka stanovené sekvence je cca 100 bp
- 0,3 Gbp za týden

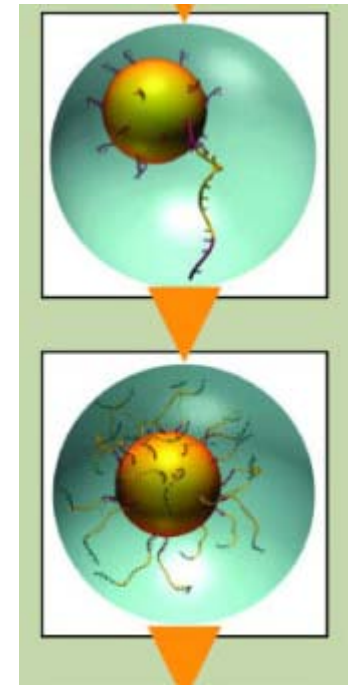
Současná technologie pyrosekvencování

- Příprava jednořetězcové DNA knihovny
 - Umožňuje zpracovat různé typy vzorků
 - genomové DNA
 - produkty PCR
 - BACs
 - cDNA
 - Frakcionace dlouhých úseků na 300- to 800-bp fragmenty
 - Vazba dvou adaptorů A a B – specifických pro 3'- a 5'-konce na každý fragment
 - Adaptor A obsahuje vazebné místo pro primer
 - Adaptor B obsahuje 5'-biotinovou značku, která slouží k imobilizaci DNA knihovny na mikrosféry (DNA Capture Beads) pokryté streptavidinem



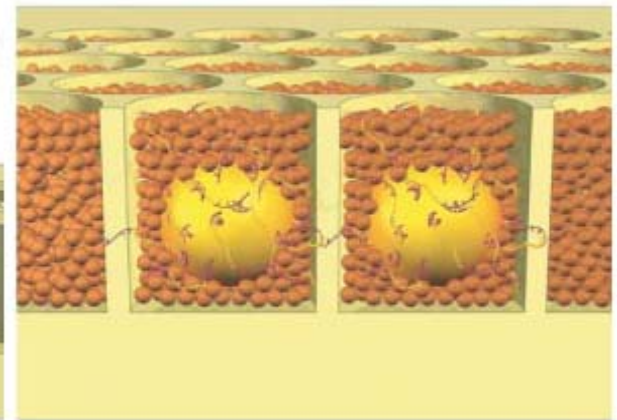
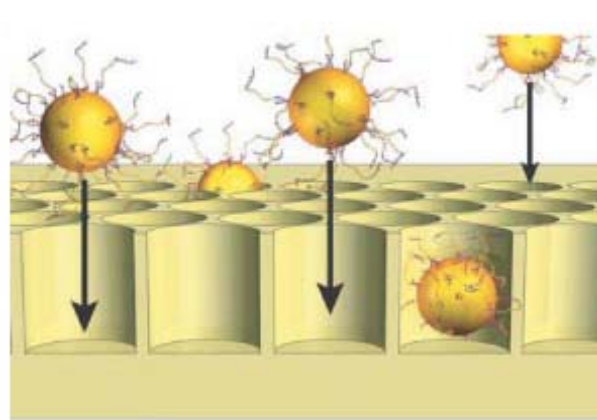
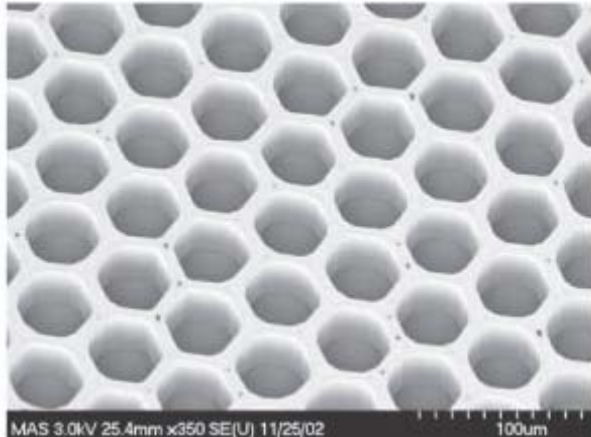
Současná technologie pyrosekvencování

- Separace kuliček a výběr pouze těch, které obsahují navázaný jediný fragment ssDNA
- Klonální amplifikace knihovny v emulzi
 - **emulzní PCR**, amplifikační reagentie ve směsi voda-olej → mikroreaktory obsahující jedinou kuličku s unikátním fragmentem DNA z knihovny - (10^7) kopií
- Pyrosekvencování
- Analýza dat s využitím bioinformatických nástrojů

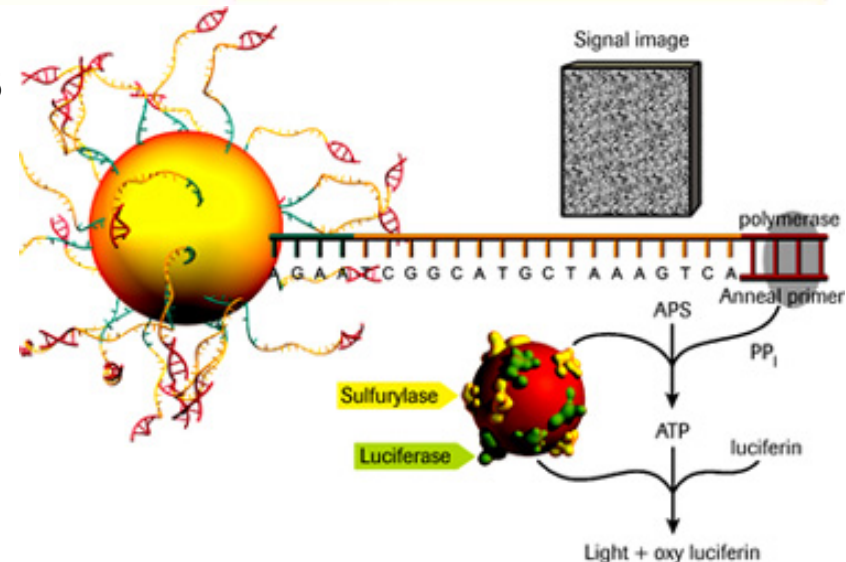


Sekvenátor FLX firmy Roche

- Pyrosekvencování (**One Bead = One Read**)
Kuličky se z emulze extrahují a jsou umístěny pomocí centrifugace do PicoTiterPlate, čipu z optických vláken o rozměrech 70 × 75 mm

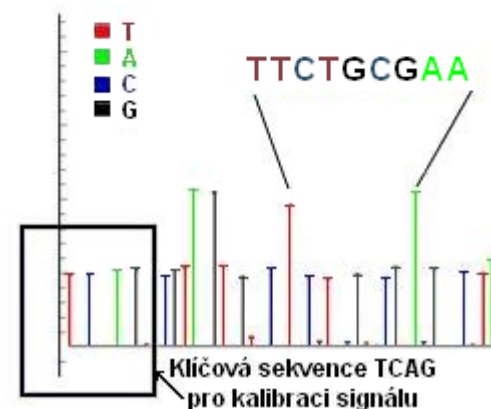
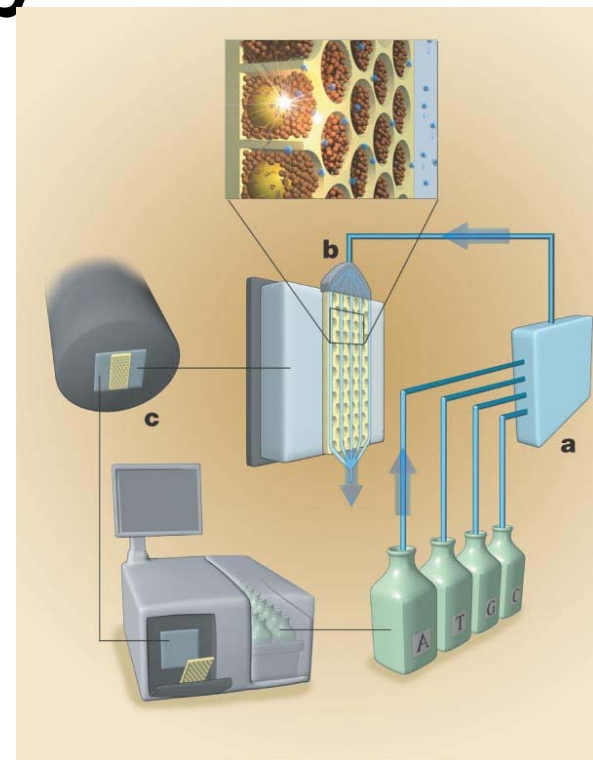


- Průměr jamek PicoTiterPlate dovoluje umístění pouze jedné kuličky průměru 28 μm
 - Na jednom čipu je 1 600 000 jamek
- Převrstvení směsí obsahující reagenty pro pyrosekvencování, enzymy jsou rovněž imobilizovány na malých mikrosférách
 - Spodní část čipu je v optickém kontaktu s dalšími optickými vlákny napojenými na CCD senzor - zachycení až 10000 emitovaných fotonů na 1 nukleotid



Sekvenátor FLX firmy Roche

- Kombinace vysoké intenzity signálu a pozičně specifické informace na mikrodestičce umožňuje současně analyzovat více než 1 milion sekvenačních běhů za 10 hod.
 - Délka stanovené sekvence 400 bp
 - Stanovení až 400 Mbp *de novo*
 - Resekvenace genomů jakékoli velikosti
- Limitující faktory
 - Negativní posuny čtecího rámce
 - 3'→5' exonukleázová aktivita DNA polymerázy
 - homopolymerní úseky
 - Pozitivní posuny čtecího rámce
 - Nedostatečná enzymatická aktivita apyrázy
 - Přesnost 99,5 % na prvních 250 bp



Současné možnosti pyrosekvencování

- Paralelní analýza více vzorků
 - Adaptory obsahující sekvence MID (Multiplex Identifiers)
 - Až 12 různých unikátních sekvencí
 - Přidány během přípravy knihovny
 - Koamplifikovány s fragmenty DNA
 - Rozdělení mikrodestičky do 2, 4, 8 nebo 16 menších oblastí
 - Zabránění vzájemné kontaminace vzorků
 - Možnost sekvenovat současně až 192 různých vzorků DNA

Práce se sekvenčními daty

- **Sekvence** – zápis posloupnosti jednoznačných znaků odpovídajících jednotlivým zbytkům (monomerům), které se nacházejí v odpovídající posloupnosti v dané makromolekule
 - DNA nebo RNA od 5'-konce k 3'-konci
 - protein od N-konce k C-konci
- používají se jednopísmenové kódy dle pravidel IUPAC (tabulka)

Příklady formátů sekvencí

- surová data – elektroforetogramy ve formátu *.abi, *.ab1 nebo *.scf
 - prohlížeče Chromas, ABIView
- formát FASTA *.nt, *.aa nebo *.fa
- formát GenBank
- software pro převod formátů sekvencí, formátování a manipulaci s daty