



## Bi5000 – Bioinformatika

# Technologie sekvenování

- Sekvenování (sequencing) stanovení pořadí nukleotidů v molekule DNA - primární struktura

# Důvody řešení genomových projektů



- ◆ Genomika modelových organismů (*de novo* sekvenování)
- ◆ Analýza mutací a SNP, výzkum genetických onemocnění (resekvenování genomů, varianty sekvencí, sequence capture)
- ◆ Diagnostika a identifikace (*de novo* sekvenování a resekvenování)
- ◆ Metagenomika (16S rRNA, celé metagenomy)
- ◆ DNA metylační studie (medip-seq, bisulfite treated DNA, ONT)
- ◆ Analýza transkriptomu (RNAseq), diferenciální exprese
- ◆ Analýza malých RNA: miRNAs, siRNA, piRNA, (small RNA seq)
- ◆ CHIP-seq (Chromatin immunoprecipitation sequencing)

# Techniky sekvenování



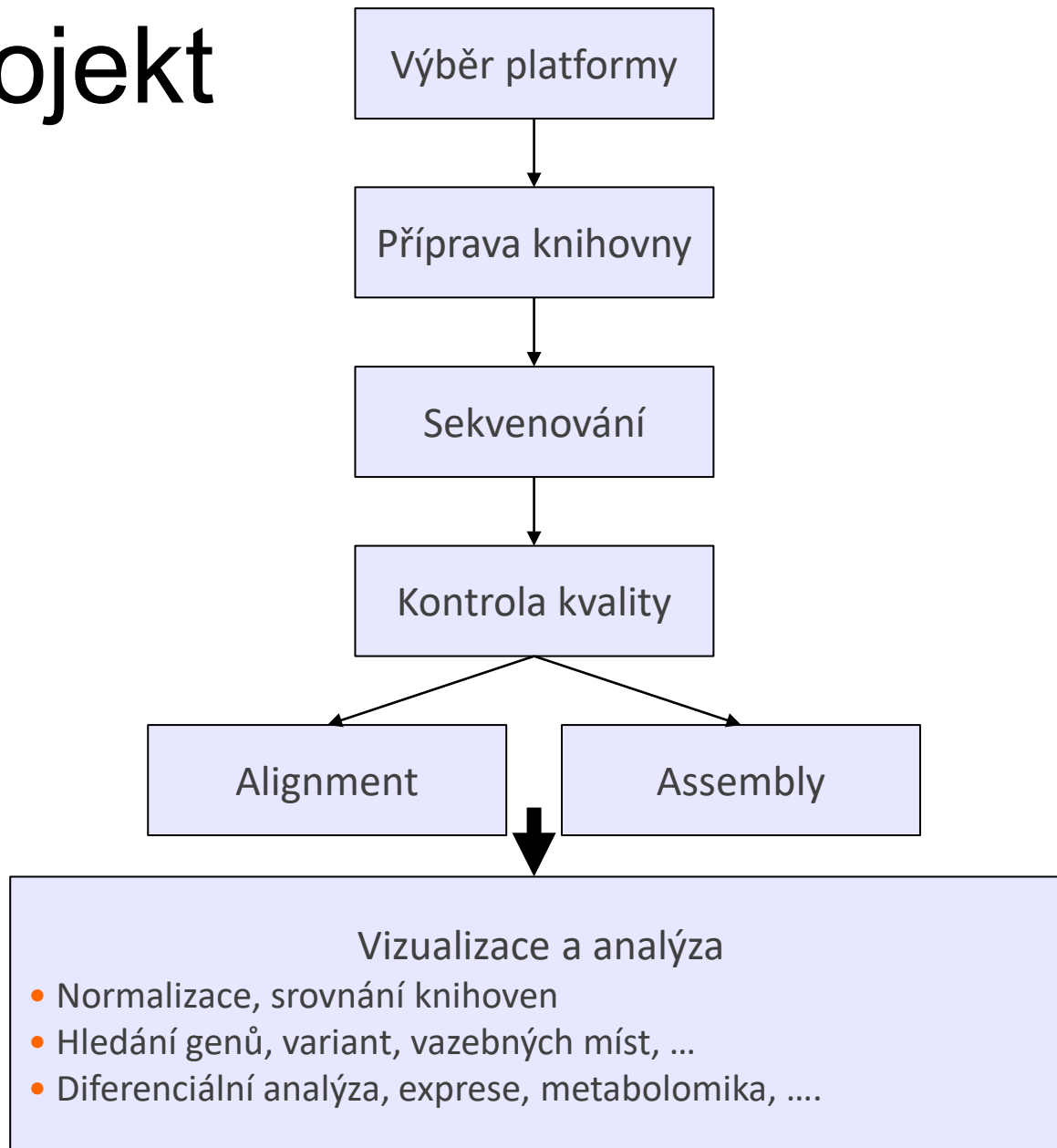
- Klasické techniky:
  - Chemická (Maxamova-Gilbertova)
    - Již se nepoužívá
  - **Enzymová (Sangerova)**
    - V současnosti se používá automatická varianta
- Metody sekvenování nové generace (next generation sequencing, NGS)
  - 454 (pyrosekvenování), **IonTorrent, Illumina, SoLiD**
- Metody sekvenování třetí generace
  - **PacBio, NanoPore**

# Princip sekvenačních metod

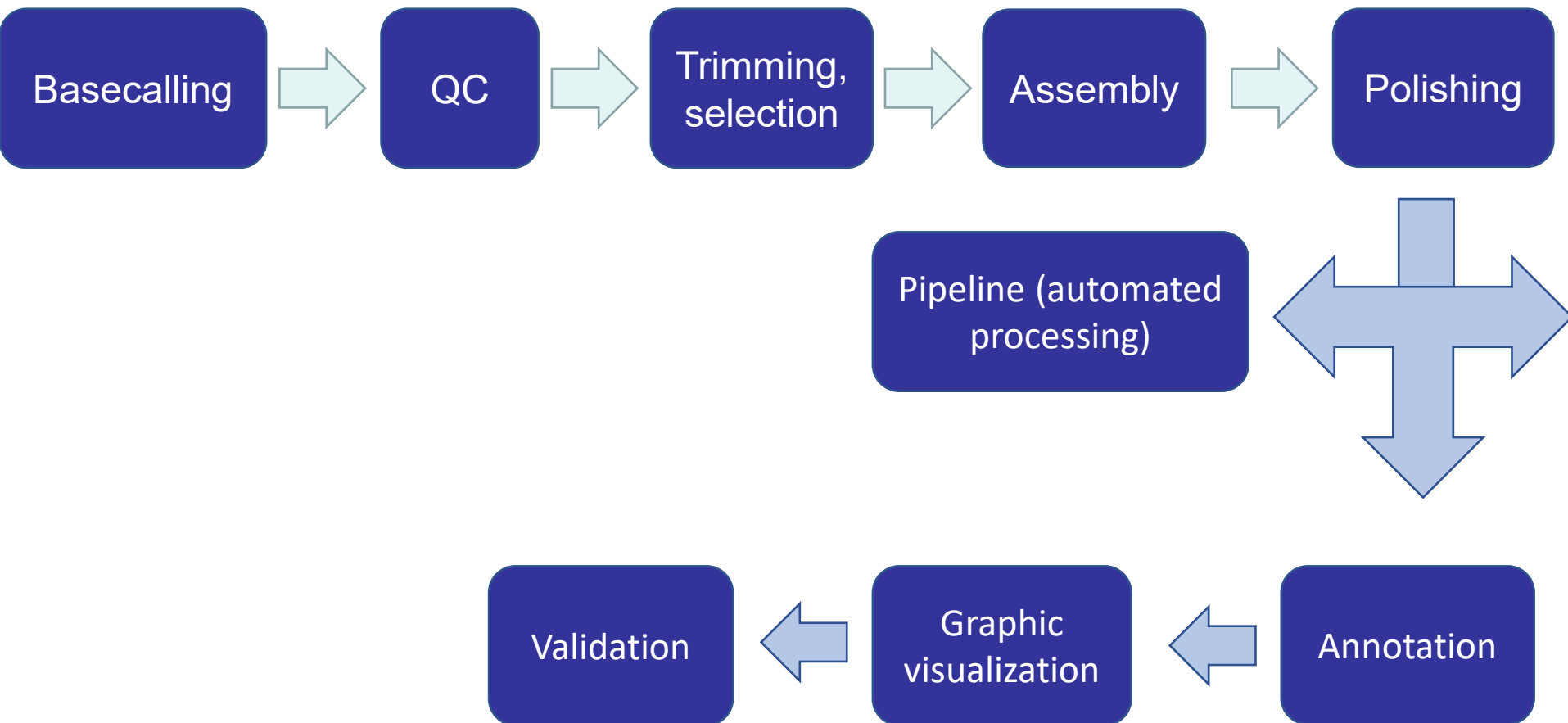


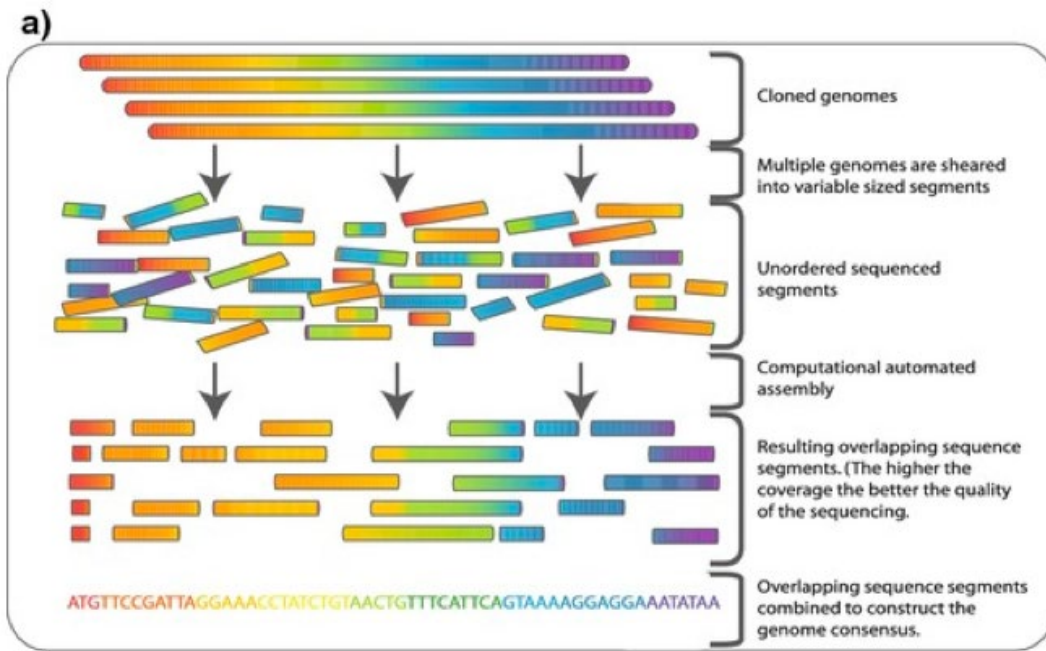
- **Klasické sekvenování kapilární elektroforézou**
  - Vyžaduje **velký počet kopií** vstupního materiálu DNA jako templátu ve formě jednořetězců, abychom zachytili signál
- **Sekvenování nové generace**
  - Jako templát slouží **jediná molekula**, která je **amplifikována** pro získání dostatečného signálu, produkuje **krátká** čtení, možnost kvantitativní analýzy (SNP)
- **Sekvenování třetí generace**
  - Jako templát slouží **jediná molekula**. Nevyužívá amplifikace za účelem zvýšení signálu, produkuje **dlouhá** čtení

# Projekt



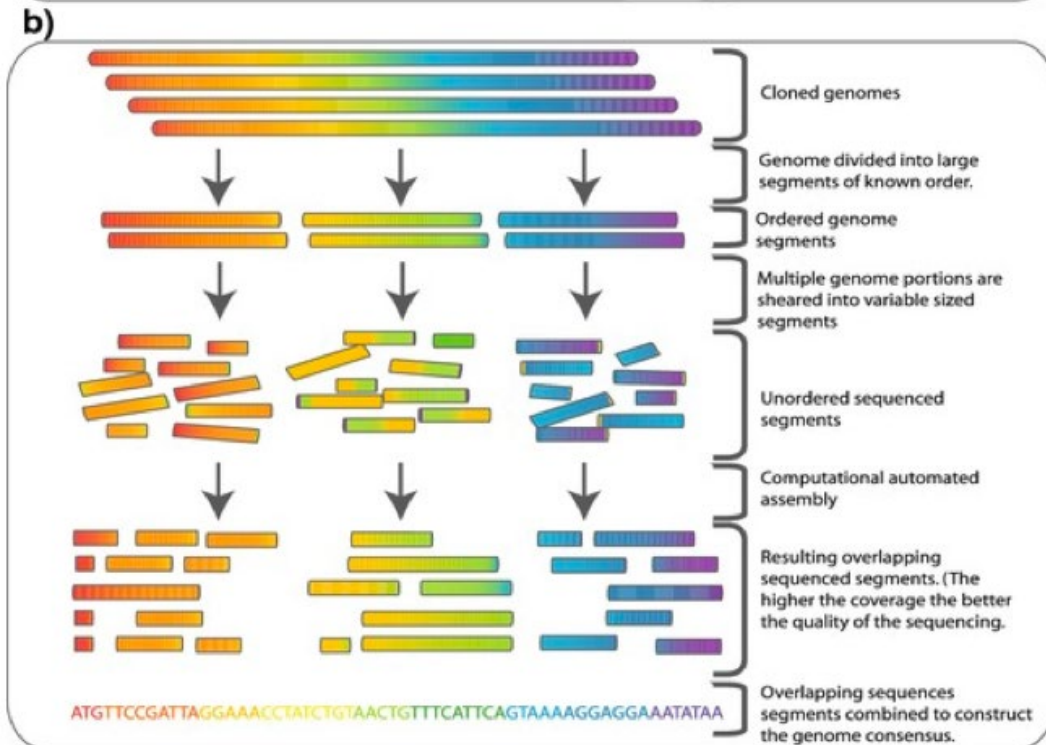
# Bioinformatické zpracování





Celogenomové  
shotgun (náhodné)  
sekvenování

versus



postupné shot-gun  
sekvenování

# Technické požadavky pro úspěšné stanovení sekvence

- ❖ **Fragmentace DNA**
  - zajistit dostatečné pokrytí při čtení genomu
- ❖ **Příprava knihovny pro sekvenování**
  - molekuly s přesně definovanými konci
    - s místem pro vazbu primeru
    - s koncovou značkou
- ❖ **Příprava variant fragmentů (ssDNA) lišících se svou délkou o jeden nukleotid**
  - Syntéza
  - Degradace
- ❖ **Detekce variant fragmentů**
  - Přesná separace na základě jejich délky
  - Detekce koncové báze po jejím připojení
  - Kontinuální čtení



# Postup enzymatického sekvenování DNA

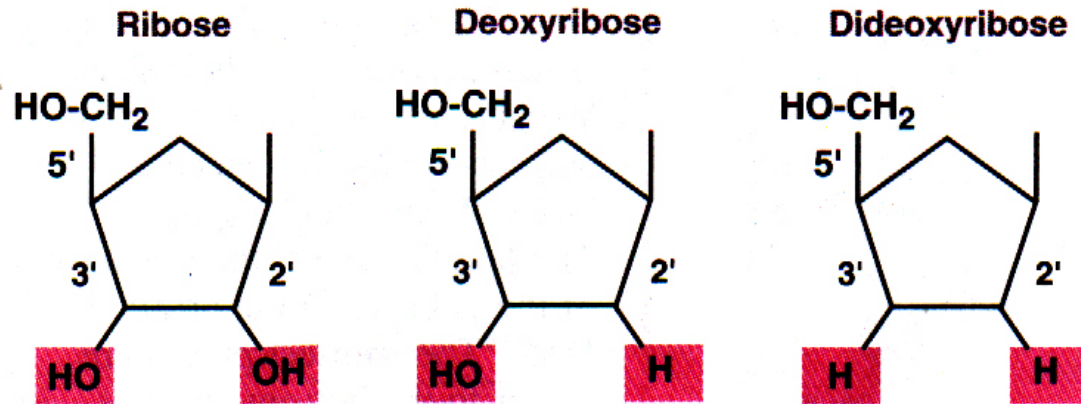
1. Příprava ssDNA jejíž sekvence má být stanovena
2. Rozdělení do 4 vzorků
3. Reakce s DNA polymerázou při níž se začlení do syntetizované DNA místo normálního deoxyribonukleosidtrifosfátu jeho analog, který působí jako koncový inhibitor syntézy DNA - dideoxynukleosidtrifosfát (ddNTP). V každém ze 4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

## Reakce obsahuje:

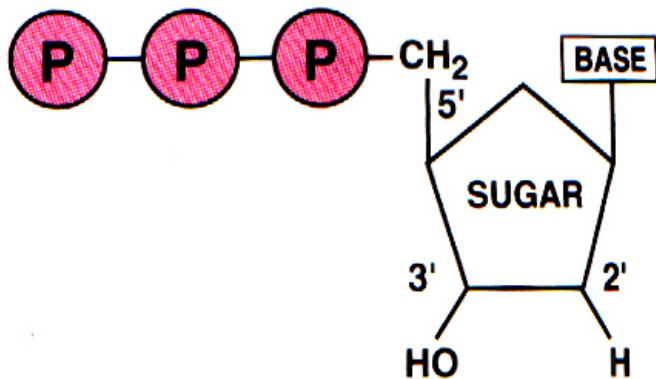
- molekulu DNA
- značený primer při pojující se k části molekuly DNA (místo odkud začínáme stanovovat sekvenci)
- směs obsahující 4 normální nukleotidy
- jeden ddNTP (v každém ze 4 vzorků je jiný)
- DNA polymerázu

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

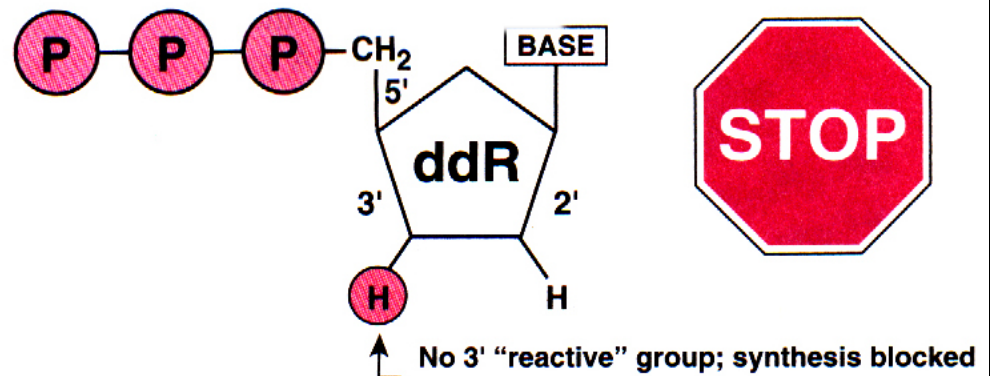
# Dideoxynukleotidy nemají hydroxyl na 3'-konci



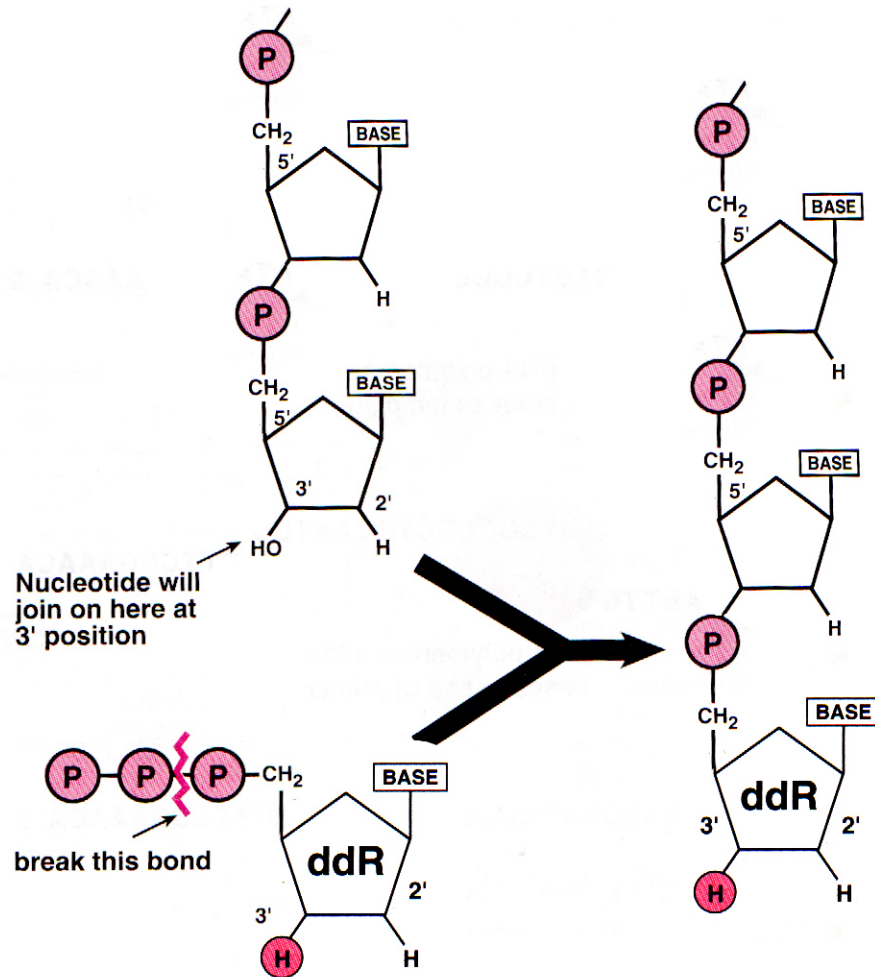
## 23.4 DEOXYNUCLEOSIDE TRIPHOSPHATE



## 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

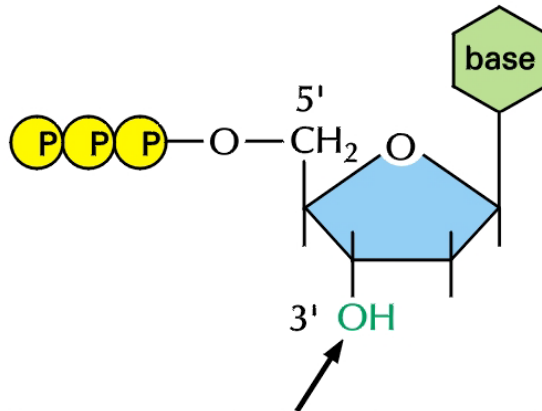
~~A G C C T G G C G A C C A T C G T  
T C G G A C C G C T G G T A G C A~~

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

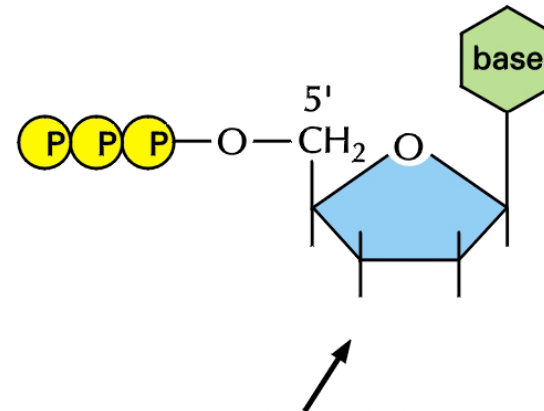
# Enzymová metoda sekvenování DNA (Sanger)

(A) **deoxy**ribonukleosidtrifosfát



je možno prodloužit řetězec na 3' -konci

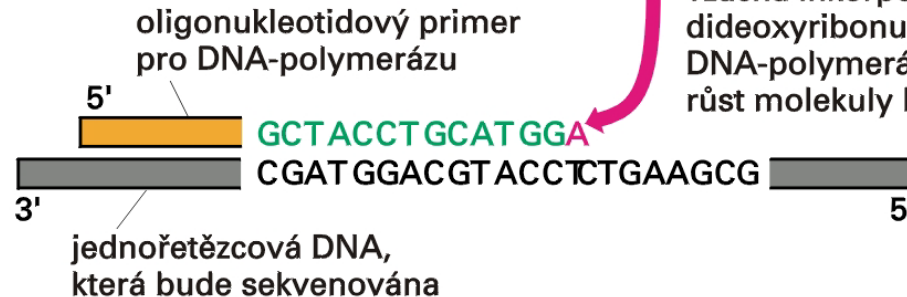
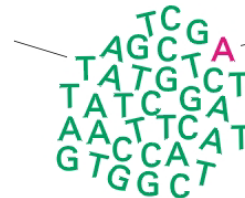
**dideoxy**ribonukleosidtrifosfát



není možno prodloužit řetězec na 3' -konci

(B) normální prekursori **deoxy**ribonukleosidtrifosfátů (dATP, dCTP, dGTP, dTTP)

malé množství jednoho **dideoxy**ribonukleosidtrifosfátu (ddATP)



(C)

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

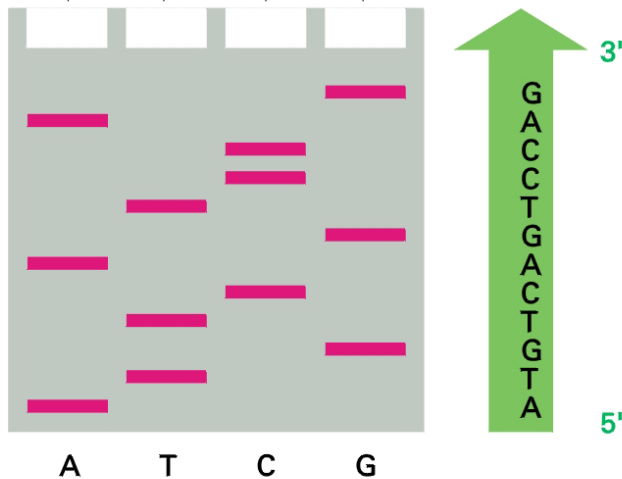
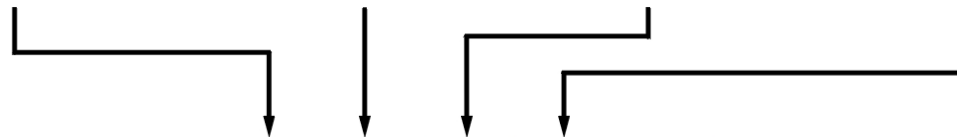
označený primer

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

+ nadbytek dATP  
dTTP  
dCTP  
dGTP

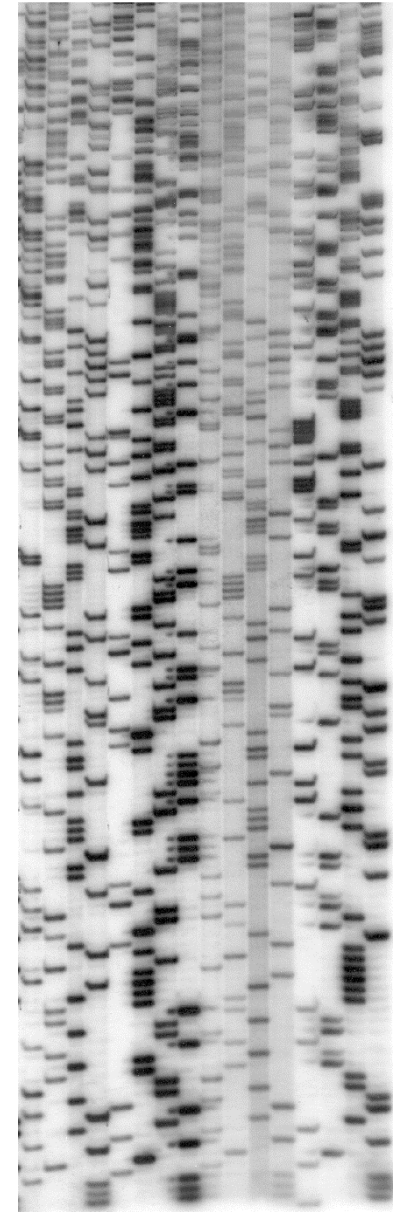
+ ddATP + DNA-polymeráza    + ddTTP + DNA-polymeráza    + ddCTP + DNA-polymeráza    + ddGTP + DNA-polymeráza

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



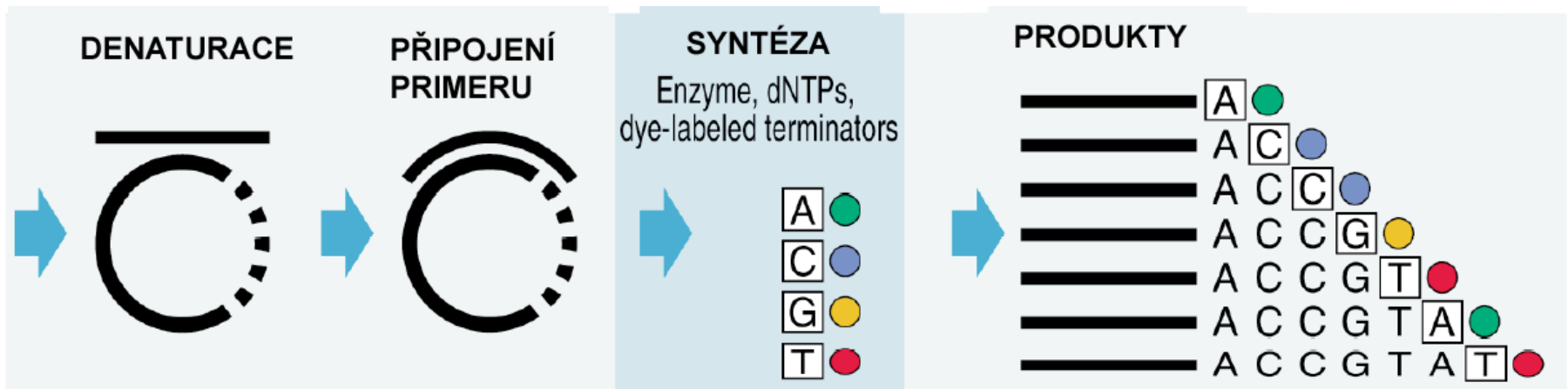
# Automatické sekvenování DNA

- Je variantou Sangerova enzymového sekvenování DNA
- Knihovnu tvoří
  - Klonované fragmenty v plazmidových vektorech
  - Produkty PCR
- Syntéza DNA probíhá v jedné reakci
- Ke značení produktů se používají čtyřmi různými fluorescenčními značkami označené
  - primery
  - dideoxyribonukleotidy

# Strategie barevných terminátorů

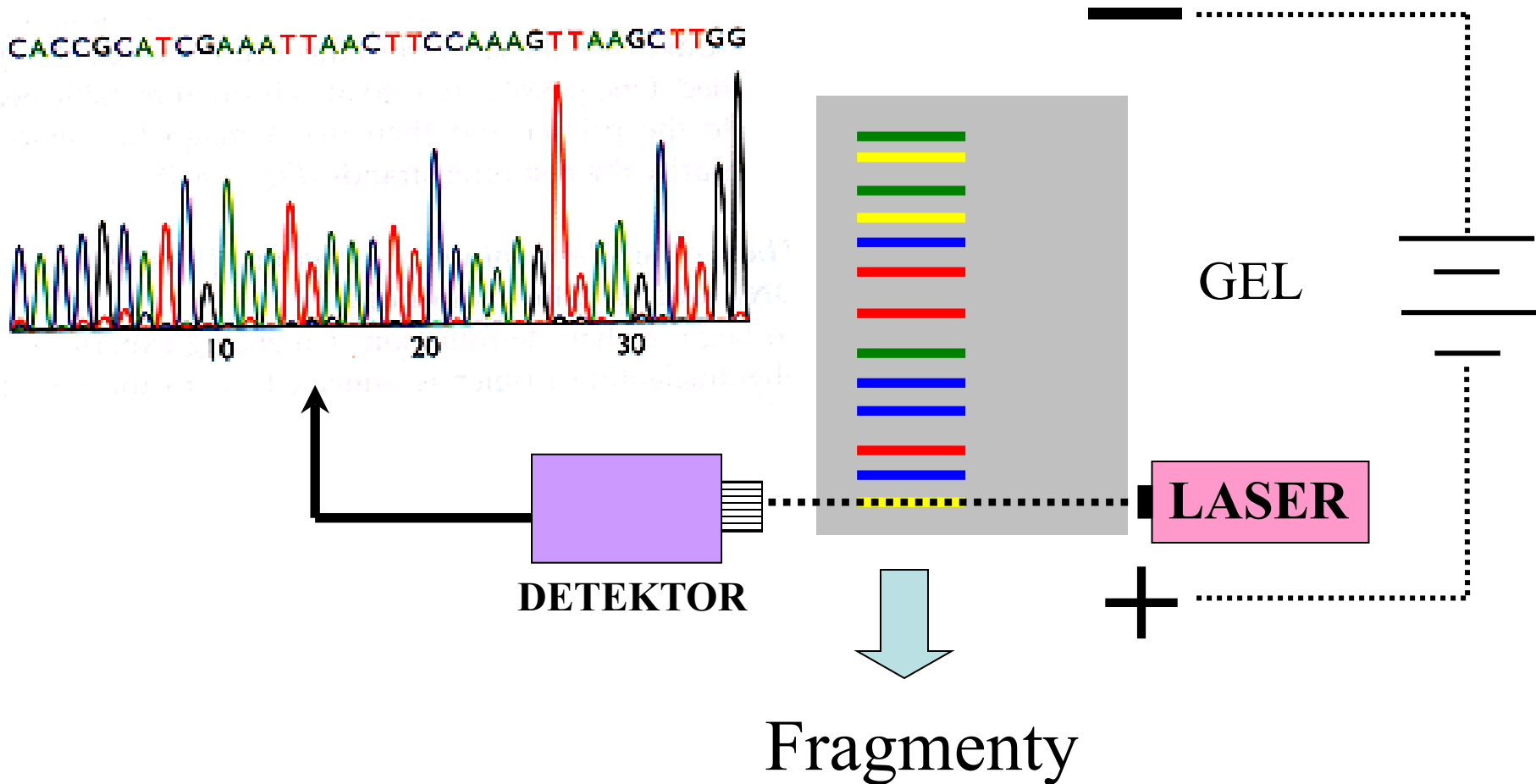


asymetrická PCR

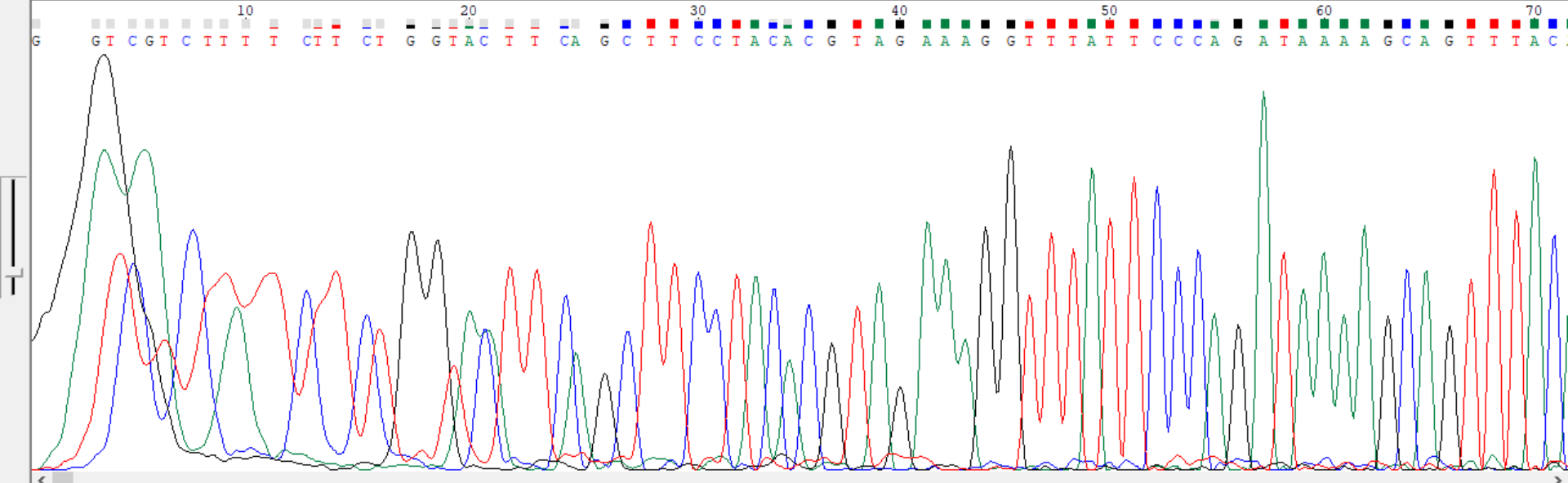


# Princip detekce produktů

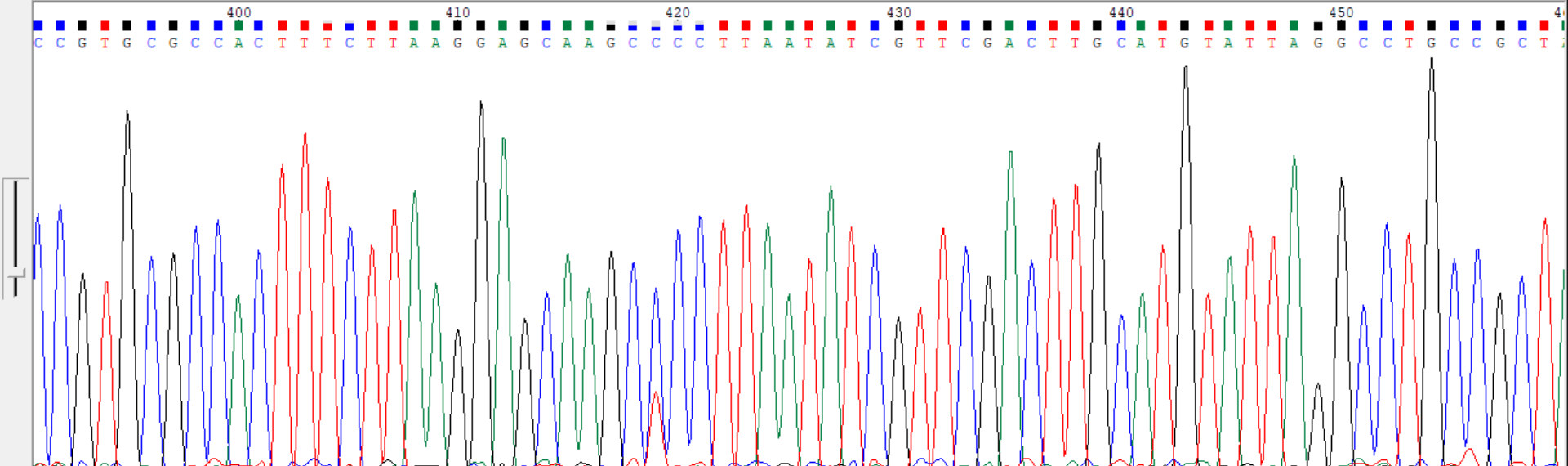
- Detekce produktů probíhá během elektroforézy pomocí laserového detektoru (laserem indukovaná fluorescence, LIF) napojeného na počítač, který vyhodnocuje sekvenci.



Open Save Export Print Next Find Reverse Enhance Sample: 3161A\_R400-ARP0022130



Open Save Export Print Next Find Reverse Enhance Sample: 3161A\_R400-ARP0022130



# Genetický analyzátor ABI 3100

Soustava  
kapilár

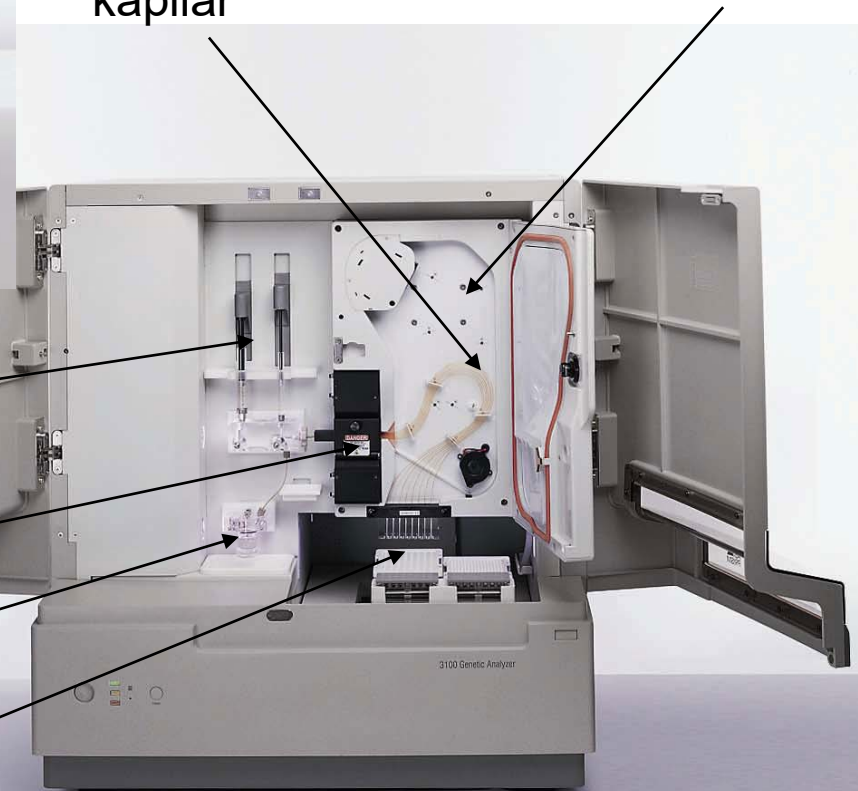
Vyhřívaná deska

Příprava gelu

Laserový detektor

Rezervoár pufry

Automatické nanášení vzorku



# Příprava knihovny pro náhodné sekvenování genomů

- Při náhodném sekvenování 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  $Mn^{2+}$ .

# Příprava knihovny pro Sangerovo sekvenování



Izolace DNA



Fragmentace 1300 – 2000 bp

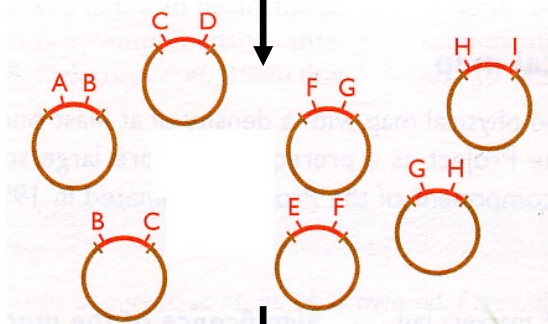


+



Vektor

Úprava konců a klonování



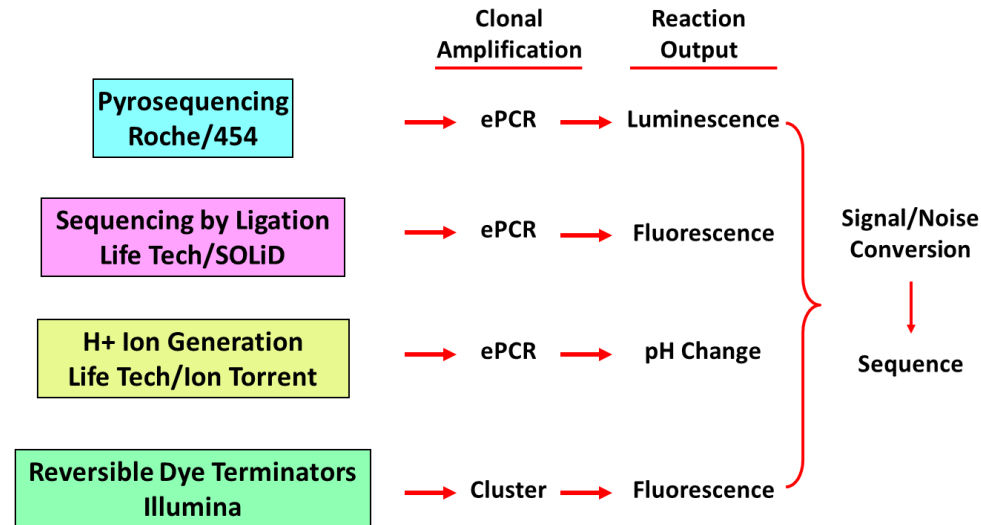
Sestavení překrývajících se klonů





# Metody sekvenování nové generace (next generation sequencing, NGS)

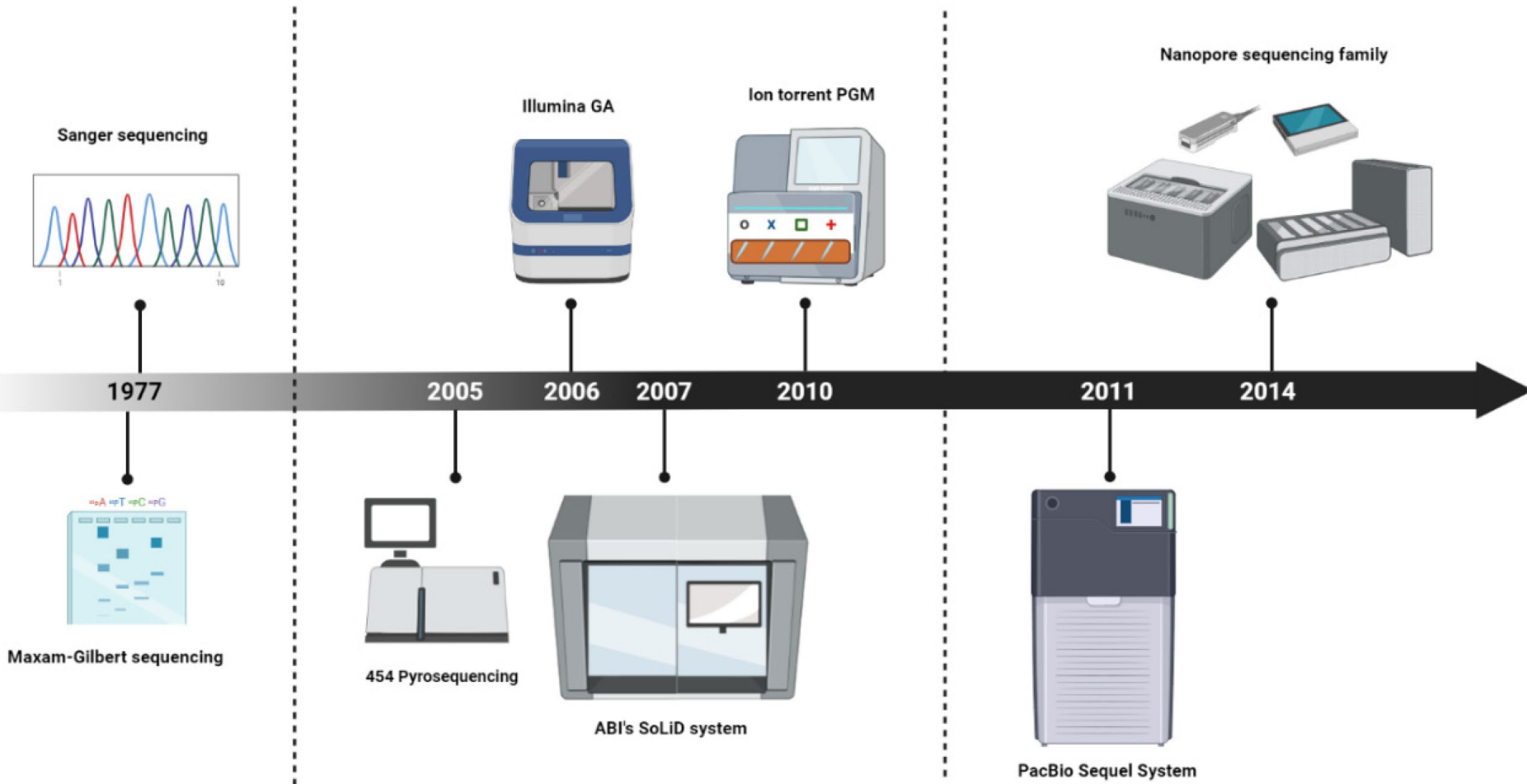
- Liší se v provedení a chemii
- V základu se podobají v sekvenování tisíců až miliónů klonálně amplifikovaných molekul (masivní paralelní sekvenování)
- Probíhá vývoj technik
  - Více čtení
  - Delší čtení
  - Rychlejší sekvenování
  - Nižší cena
- Nové aplikace v klinické oblasti, vývoj kitů cílených na určité části genomů (dědičná onemocnění, nádorová onemocnění, metagenomika, 16S rRNA)



- **Sekvenování třetí generace**
  - Pacific Biosciences
  - Oxford Nanopore Technologies
  - Helicos Biosciences
  - NABsys
  - VisiGen Biotechnologies



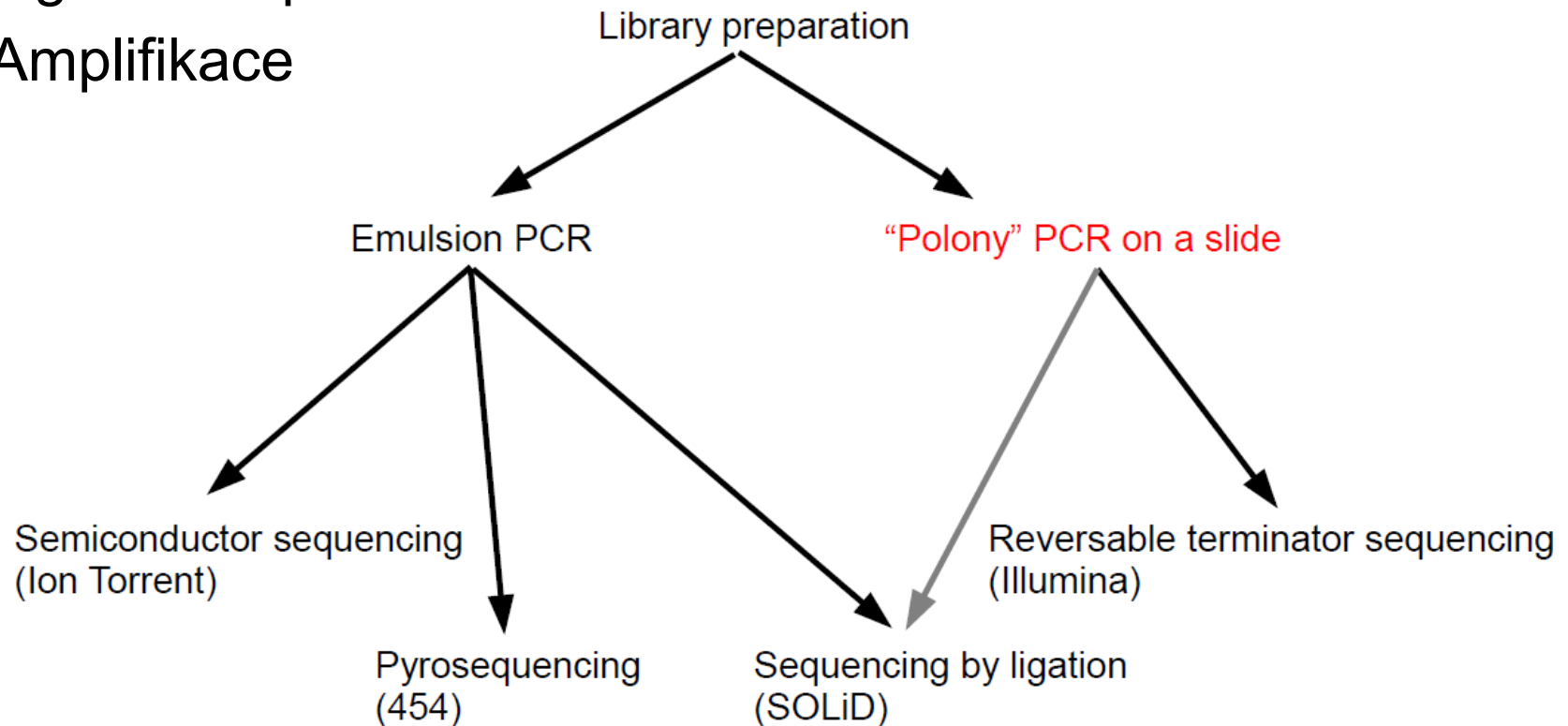
# Vývoj technik sekvenování



# Metody přípravy knihovny

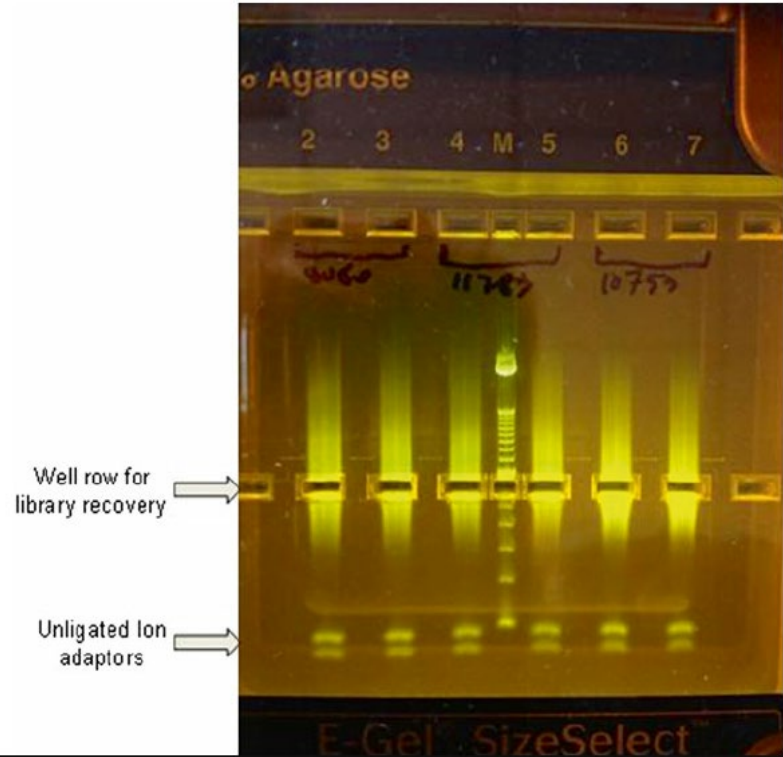


- Fragmentace DNA na uniformní délku
  - Enzymatická
  - Fyzikální (sonikace, nebulizace, Hydroshear)
- Ligace adaptorů
- Amplifikace

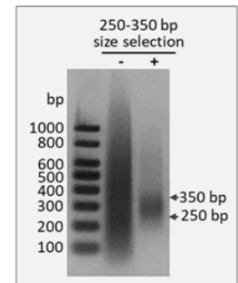
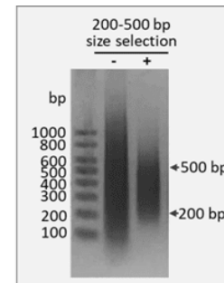
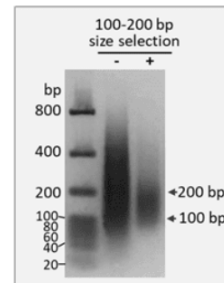
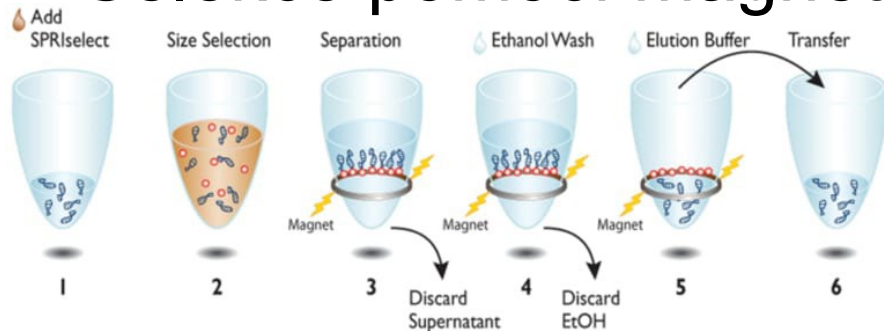


# Selekce fragmentů podle velikosti

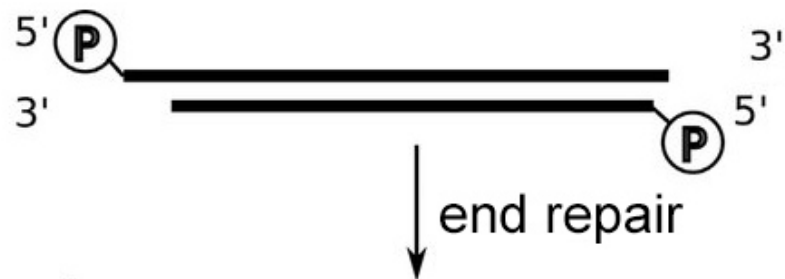
- E-gel Systém



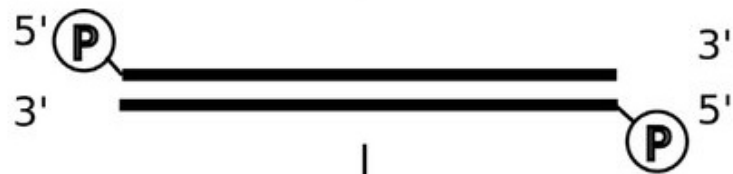
- Selekce pomocí magnetických kuliček



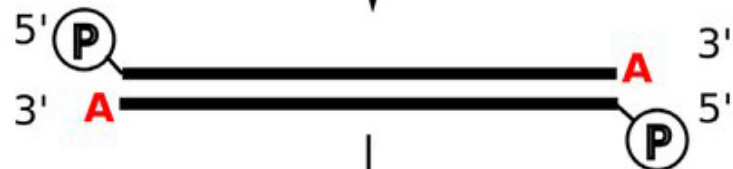
# Ligase adaptor



end repair



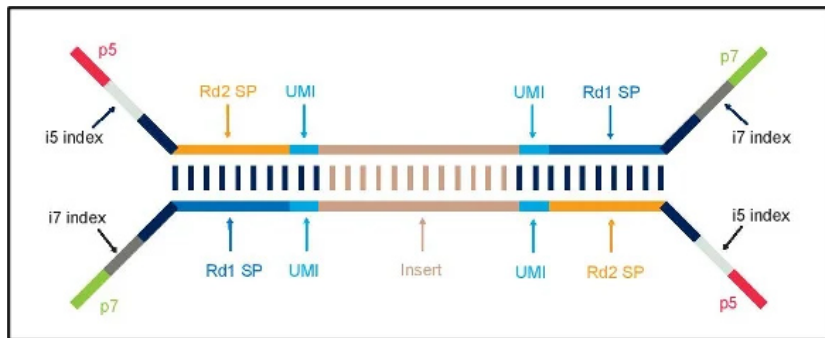
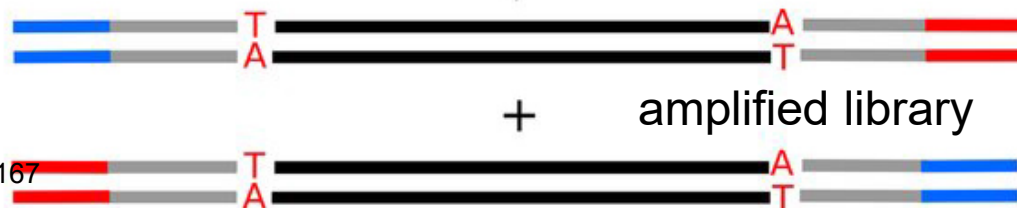
tailing



ligation



PCR (barcode incorporation)





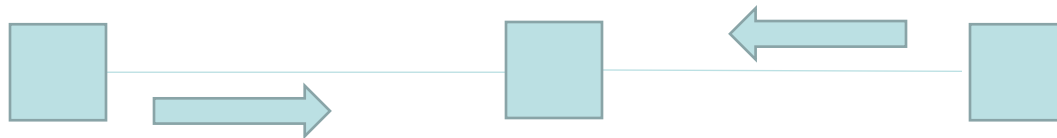
# Typy uspořádání NGS sekvenování



Single Read



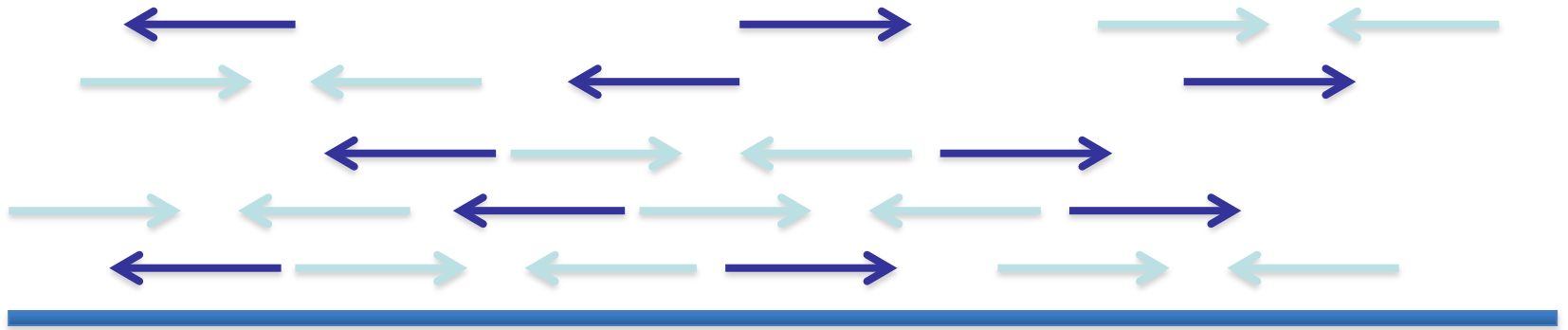
Paired-end read (PE)



Mate-pair read (MP)

výsledkem jsou miliony krátkých čtení sekvencí z náhodných částí genomu

# Orientace párových čtení



Paired end (PE) reads



PE, MP

Mate pair (MP) reads



MP

# Hloubka sekvenování (pokrytí)

- Coverage (pokrytí) = (počet získaných čtení × průměrná délka čtení)/velikost genomu
- Fragment 2000 bází sestavíme z 8 čtení o délce 500 bp ->  $8 \times 500 / 2000 = 2$
- Příklad 1: Získali jsme  $10^6$  čtení s technologií IonTorrent, velikost genomu je 3 Mb. Jaké je pokrytí?
- Příklad 2: Sekvenátor Illumina poskytuje 200 milionů čtení. Kolik genomů o velikosti 5 Mb mohou osekvenovat s použitím chemie poskytující čtení o délce 150 nt, abych získal pokrytí 50×?

# Ion Torrent polovodičové sekvenování

- Nástupce pyrosekvenování (454-sekvenování)
- prvním platforma, která **nepoužívá pro detekci sekvenovaných bází DNA světelný signál, který by byl uvolňovaný při enzymatických reakcích s fluorescenčními nebo chemiluminescenčními substráty**
- Sekvenování Ion Torrent je založené na elektrochemické detekci vodíkových iontů (pH) které jsou uvolněny při replikaci sekvenovaného templátu na speciálním polovodičovém čipu.

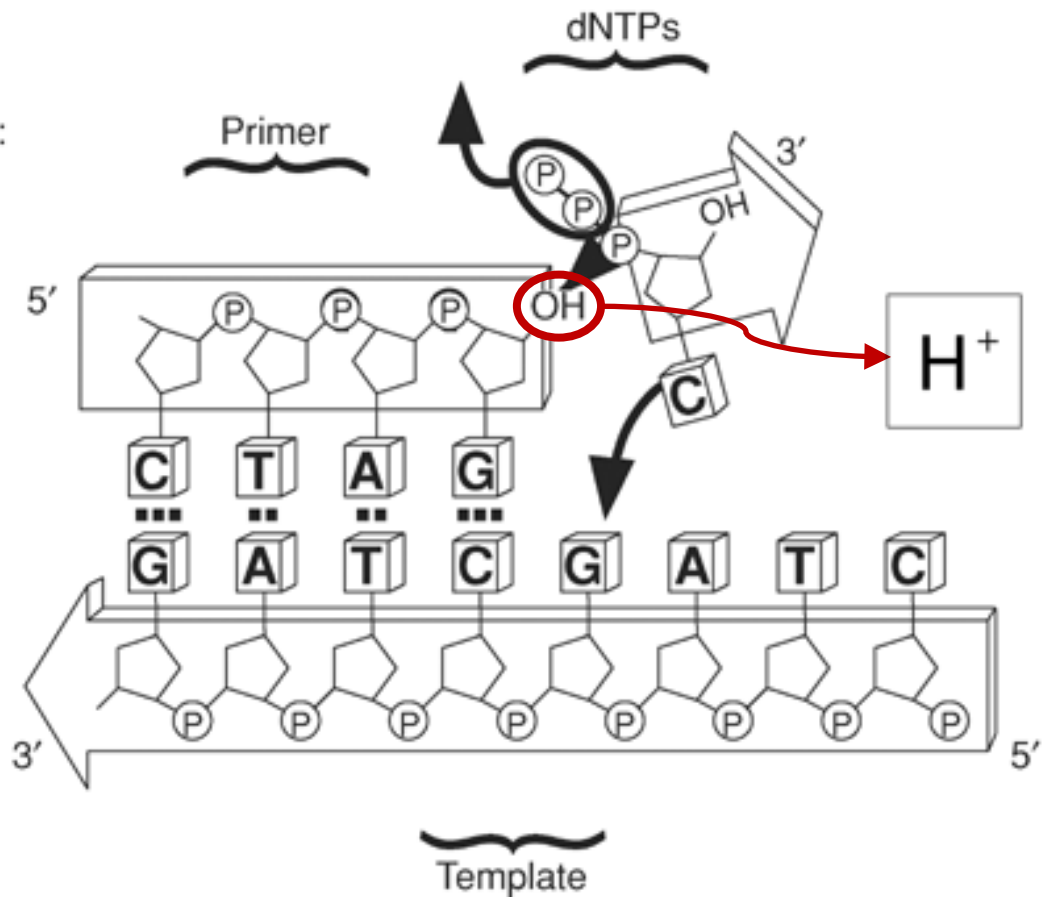




# Uvolnění vodíku v nepufřujících podmínkách

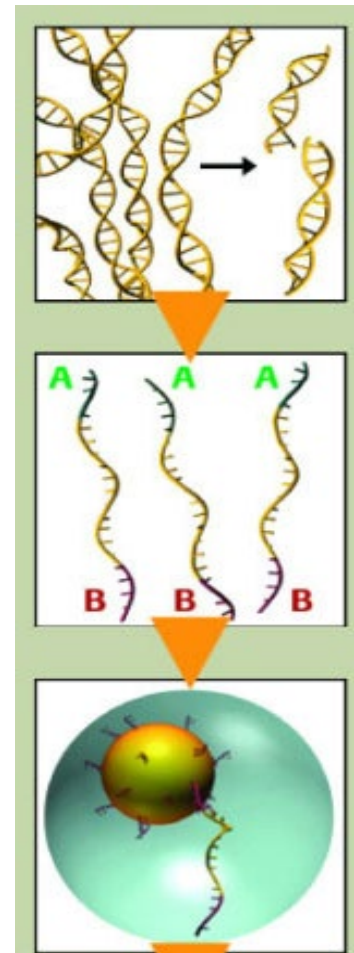


Example:

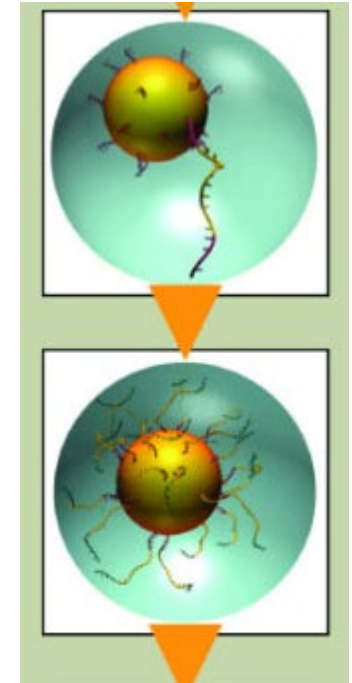


# Příprava knihovny při emulzní PCR

- Příprava jednořetězcové DNA knihovny
  - Umožňuje zpracovat různé typy vzorků
    - genomové DNA
    - produkty PCR
    - klonované DNA
  - Frakcionace dlouhých úseků na 300- až 500-bp dlouhé fragmenty
  - Vazba dvou adaptorů 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 pokryté streptavidinem



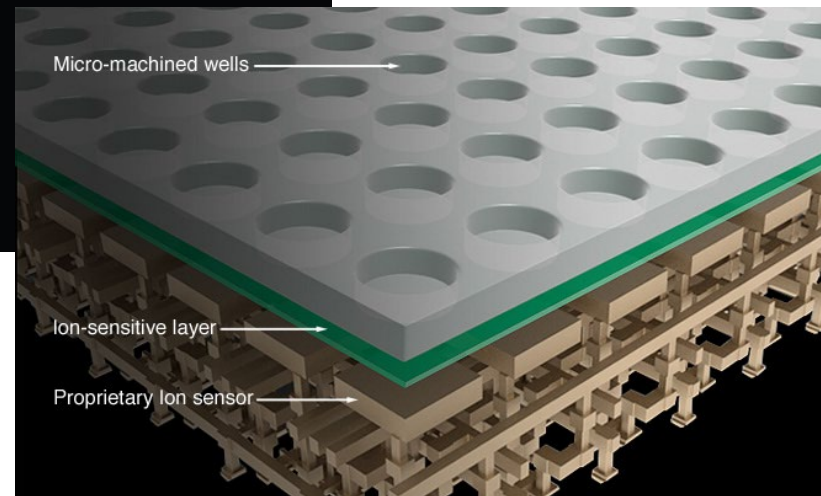
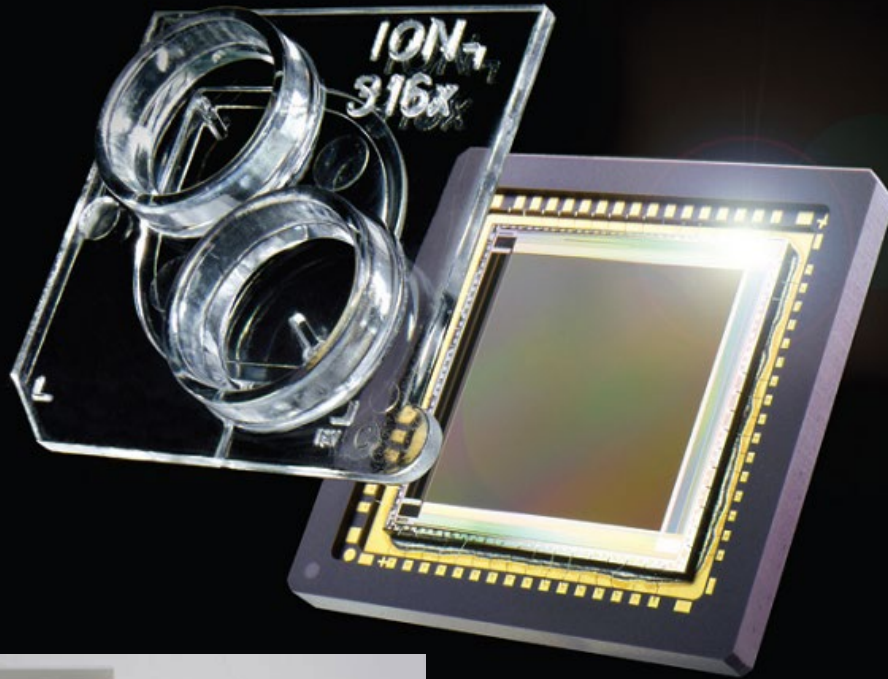
- Klonální amplifikace knihovny
  - **emulzní PCR**, amplifikační reagensie ve směsi voda-olej → mikroreaktory obsahující jedinou kuličku s unikátním fragmentem DNA z knihovny - ( $10^7$ ) kopií
- Separace kuliček a výběr pouze těch, které obsahují amplifikovanou DNA (enrichment)
- Sekvenační reakce
- Analýza dat a assembly



# IonTorrent polovodičové sekvenování

- Na čipu se nachází milióny jednotlivých reakčních cell (reaktorů), ve kterých probíhá sekvenační analýza individuálních kopií DNA na základě kontinuálního střídání reakčních složek (jednotlivých typů nukleotidů) v mikrofluidním systému.
- DNA imobilizovaná na mikrosférách (po emulzní PCR)
- Kapacita závisí na typu použitého polovodičového chipu
  - 10 – 65 miliónů
  - 100 – 600 miliónů
  - > 1000 miliónů

# IonTorrent čip

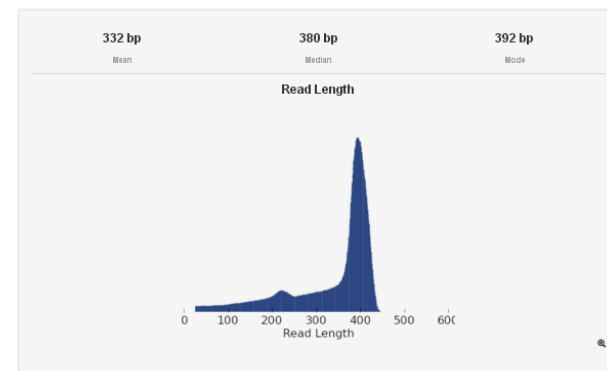
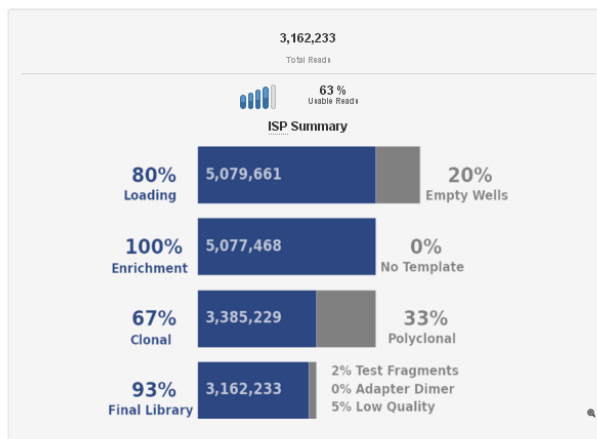
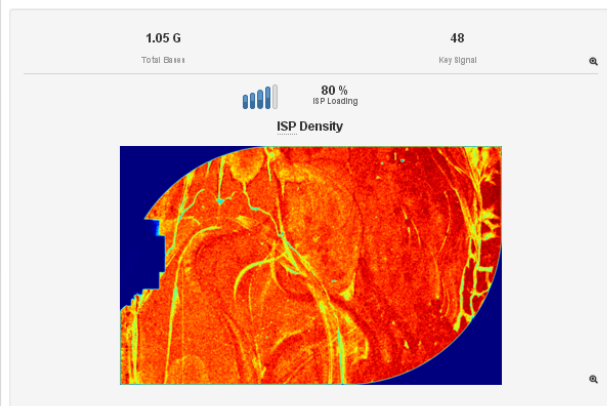


# Ion Torrent polovodičové sekvenování

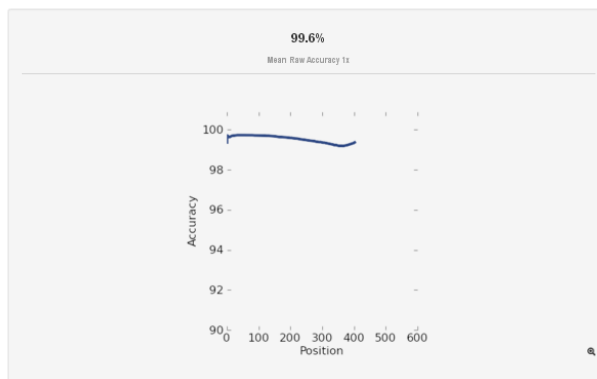
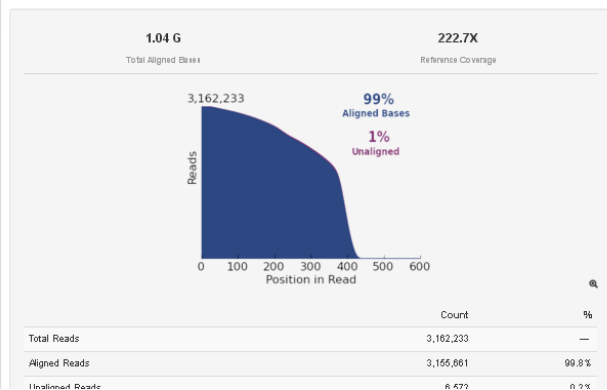
- Technologie Ion Torrentu je otevřená pro sekvenování DNA *de-novo* i cílené sekvenování vybraných oblastí genomu
- Minimální délka sekvenačního čtení
  - V jedno směru: 400bp
  - Obousměrné (Paired end)
  - Počet vzorků analyzovaných v jednom běhu
  - až 96 vzorků – pomocí označení tzv. barkodem
- Doba sekvenační analýzy
  - 2 - 3,5 hodiny – podle použitého čipu

# Výstup z Torrent server

## Read Summary: Unaligned



## Aligned to E. coli DH10B



1.03 G AQ17 Total Bases

Alignment Quality

	AQ17	AQ20	Perfect
Total Number of Bases [bp]	1.03 G	1 G	780 M
Mean Length [bp]	332	325	259
Longest Alignment [bp]	499	499	472
Mean Coverage Depth [x]	220.3	213.8	166.4

# Alignment & Variant Calling

Aplikace Místa Systém Ct, 24. květen, 15:47 LMDM

GS De novo Assembler

Project: Stafal [Genomic] Ready for analysis  
Location: /home/LMDM/Plocha/analyza/Stafal

Overview Project Parameters Result files Alignment results Flowgrams

Contig	Bases
contig00001	20 149
contig00002	18 049
contig00003	14 160
contig00004	11 926
contig00005	8 637
contig00006	7 994
contig00007	6 816
contig00008	6 697
contig00009	4 476
contig00010	3 326
contig00011	3 095
contig00012	3 014
contig00013	2 722
contig00014	2 690
contig00015	2 383
contig00016	2 329
contig00017	2 191
contig00018	2 036
contig00019	1 499
contig00020	1 352
contig00021	1 166
contig00022	1 050
contig00023	1 016
contig00024	948
contig00025	770

Contig: contig00001 - 20 149 bp.

Base left selected right

2 032 --- 2 132

contig00001 TCTTCTCCTAAAAACCC-TAAAGTATCAAGG-TTTTT-ACAAACCTCTACATT-AGGTAAGT-AACCGAAAGGTTCTATACTAGAACTGTATATAAACTATTTGT  
HK6IEBF01BN89B TCTTCTCCTAAAAACCC-TAAAGTATCAAGG-TTTTT-ACAAACCTCTACATT-AGGTAAGT  
contig00007 HK6IEBF01A3HAV TCTTCTCCTAAAAACCC-TAAAGTATCAAGG-TTTTT-ACAAACCTCTACATT-AGGTAAGT  
HK6IEBF01BHKX2 TCTTCTCCTAAAAACCC-TAAAGTATCAAGG-TTTTT-ACAAACCTCTACATT-AGGTAAGT-AACCGAAAGGTTCTATACTAGAACTGTATATAAACTATTTGT  
contig00009 HK6IEBF01A5THI TCTTCTCCTAAAAACCC-TAAAGTATCAAGG-TTTTT-ACAAACCTCTACATT-AGGTAAGT-AACCGAAAGGTTCTATACTAGAACTGTATATAAACTATTTGT  
contig00010 HK6IEBF01A9E9P TCTTCTCCTAAAAACCC-TAAAGTATCAAGG-TTTTT-ACAAACCTCTACATT-AGGTAAGT-AACCGAAAGGTTCTATACTAGAACTGTATATAAACTATTTGT  
contig00011 HK6IEBF01BKM7X TCTTCTCCTAAAAACCC-TAAAGTATCAAGG-TTTTT-ACAAACCTCTACATT-AGGTAAGT-AACCGAAAGGTTCTATACTAGAACTGTATATAAACTATTTGT  
contig00012 HK6IEBF01AV175 TCTTCTCCTAAAAACCC-TAAAGTATCAAGG-TTTTT-ACAAACCTCTACATT-AGGTAAGT-AACCGAAAGGTTCTATACTAGAACTGTATATAAACTATTTGT  
contig00013 HK6IEBF01BJY8L TCTTCTCCTAAAAACCC-TAAAGTATCAAGG-TTTTT-ACAAACCTCTACATT-AGGTAAGT-AACCGAAAGGTTCTATACTAGAACTGTATATAAACTATTTGT  
contig00014 HK6IEBF01A1I6AG TCTTCTCCTAAAAACCC-TAAAGTATCAAGG-TTTTT-ACAAACCTCTACATT-AGGTAAGT-AACCGAAAGGTTCTATACTAGAACTGTATATAAACTATTTGT  
contig00015 HK6IEBF01AW7YG TCTTCTCCTAAAAACCC-TAAAGTATCAAGG-TTTTT-ACAAACCTCTACATT-AGGTAAGT-AACCGAAAGGTTCTATACTAGAACTGTATATAAACTATTTGT  
contig00016 HK6IEBF01B1IE6 TCTTCTCCTAAAAACCC-TAAAGTATCAAGG-TTTTT-ACAAACCTCTACATT-AGGTAAGT-AACCGAAAGGTTCTATACTAGAACTGTATATAAACTATTTGT  
contig00017 HK6IEBF01AYV3H TCTTCTCCTAAAAACCC-TAAAGTATCAAGG-TTTTT-ACAAACCTCTACATT-AGGTAAGT-AACCGAAAGGTTCTATACTAGAACTGTATATAAACTATTTGT  
contig00018 HK6IEBF01BGS5K TCTTCTCCTAAAAACCC-TAAAGTATCAAGG-TTTTT-ACAAACCTCTACATT-AGGTAAGT-AACCGAAAGGTTCTATACTAGAACTGTATATAAACTATTTGT  
contig00019 HK6IEBF01BJFDR TCTTCTCCTAAAAACCC-TAAAGTATCAAGG-TTTTT-ACAAACCTCTACATT-AGGTAAGT-AACCGAAAGGTTCTATACTAGAACTGTATATAAACTATTTGT  
contig00020 HK6IEBF01B04U2 TCTTCTCCTAAAAACCC-TAAAGTATCAAGG-TTTTT-ACAAACCTCTACATT-AGGTAAGT-AACCGAAAGGTTCTATACTAGAACTGTATATAAACTATTTGT  
contig00021 HK6IEBF01BSDD8 TCTTCTCCTAAAAACCC-TAAAGTATCAAGG-TTTTT-ACAAACCTCTACATT-AGGTAAGT-AACCGAAAGGTTCTATACTAGAACTGTATATAAACTATTTGT  
contig00022 HK6IEBF01BPQRL TCTTCTCCTAAAAACCC-TAAAGTATCAAGG-TTTTT-ACAAACCTCTACATT-AGGTAAGT-AACCGAAAGGTTCTATACTAGAACTGTATATAAACTATTTGT  
contig00023 HK6IEBF01BBGE9 TCTTCTCCTAAAAACCC-TAAAGTATCAAGG-TTTTT-ACAAACCTCTACATT-AGGTAAGT-AACCGAAAGGTTCTATACTAGAACTGTATATAAACTATTTGT  
contig00024 HK6IEBF01B1K98 TCTTCTCCTAAAAACCC-TAAAGTATCAAGG-TTTTT-ACAAACCTCTACATT-AGGTAAGT-AACCGAAAGGTTCTATACTAGAACTGTATATAAACTATTTGT  
HK6IEBF01AXRBZ AAAAA-CCC-TAA-GTATCAAGG-TTTTT-ACAAACCTCTACATT-AGGTAAGT-AACCGAAAGGTTCTATACTAGAACTGTATATAAACTATTTGT  
HK6IEBF01AG811 AAAAA-CCC-TAA-GTATCAAGG-TTTTT-ACAAACCTCTACATT-AGGTAAGT-AACCGAAAGGTTCTATACTAGAACTGTATATAAACTATTTGT  
contig00025 HK6IEBF01BWH8V TTT---ACAAACCTCTACATT-AGGTAAGT-AACCGAAAGGTTCTATACTAGAACTGTATATAAACTATTTGT  
HK6IEBF01AV2R6 C TAGAACTGTATATAAACTATTTGT

base:  
read:  
val:

Info Stafal: Opened Assembler project: /home/LMDM/Plocha/analyza/Stafal

Analysis messages

[GS De novo Assembl... LMDM [Untitled 1 - OpenOffi... GS De novo Assembler

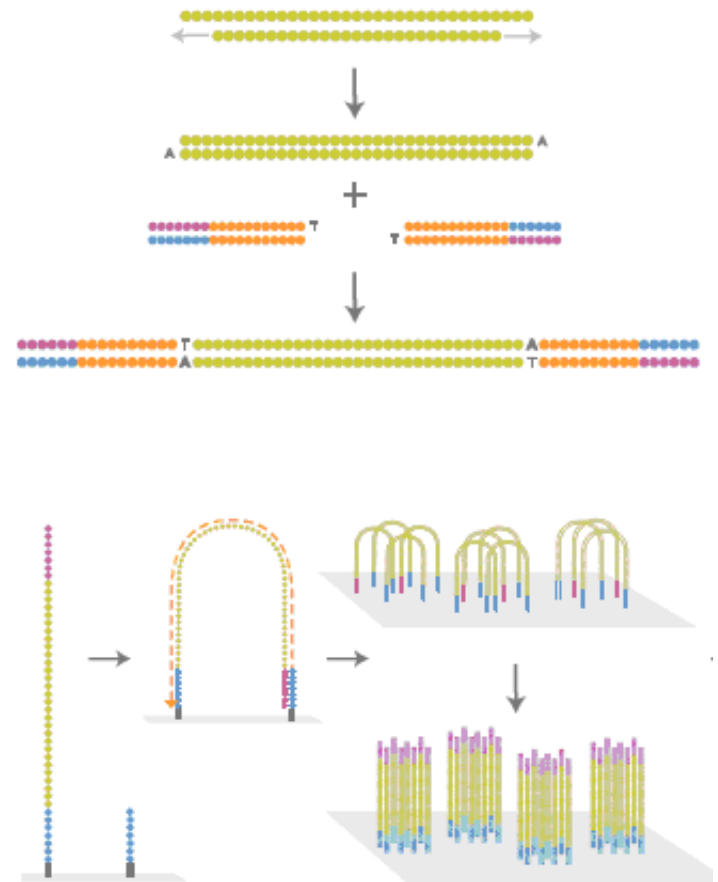


# Technologie firmy Illumina/Solexa

- Uvedena na trh 2006
- Prošla řadou inovací, zejména délky čtení
- Masivní paralelní sekvenování („polony“ sequencing)
- Sekvenování syntézou na čipu (flow cell), které využívá strategii reverzibilních terminátorů
- Nevyžaduje klonování ani přípravu knihovny na mikrosférách
- Dosahuje 99,9 % přesnosti

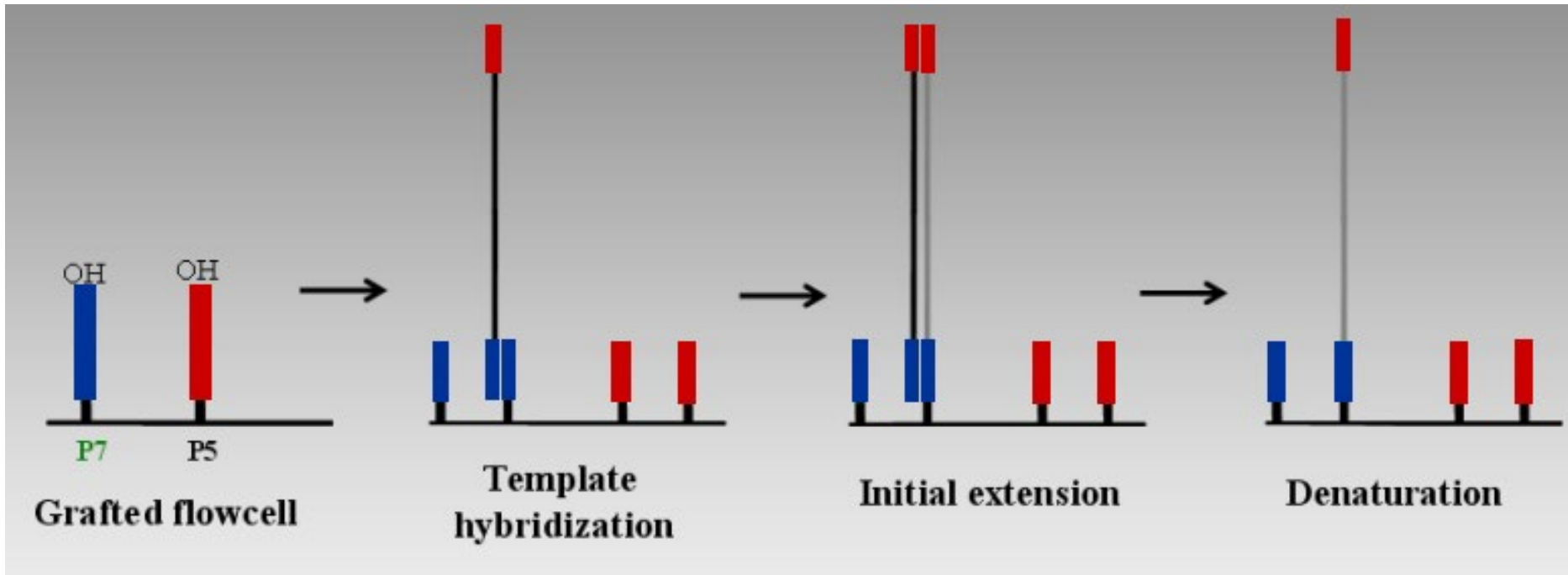
# Příprava knihovny – polony PCR

- Náhodná fragmentace DNA na krátké úseky 200 – 500 bp
- Zatupení konců a vytvoření 1 nt A-přesahu na 3'-konci (T4 DNA polymeráza, Klenow a polynukleotid-kináza T4)
- Navázání adaptorů umožňujících kovalentní vazbu na opticky transparentní povrch pro sekvenování
- Každý fragment je na povrchu uchycen pouze jedním koncem, po přidání enzymů pro amplifikaci dochází k ohnutí fragmentu do „mostu“.
- Výsledkem amplifikace jsou dva řetězce, každý s jedním volným a jedním pevným koncem
- Po denaturaci jsou fragmenty narovnány a uspořádány do shluků, ve kterých je dosažena značná hustota, až 1000 kopií fragmentu na  $\mu\text{m}^2$  povrchu
- Na celém povrchu je dosaženo hustoty deseti milionů shluků na  $\text{cm}^2$

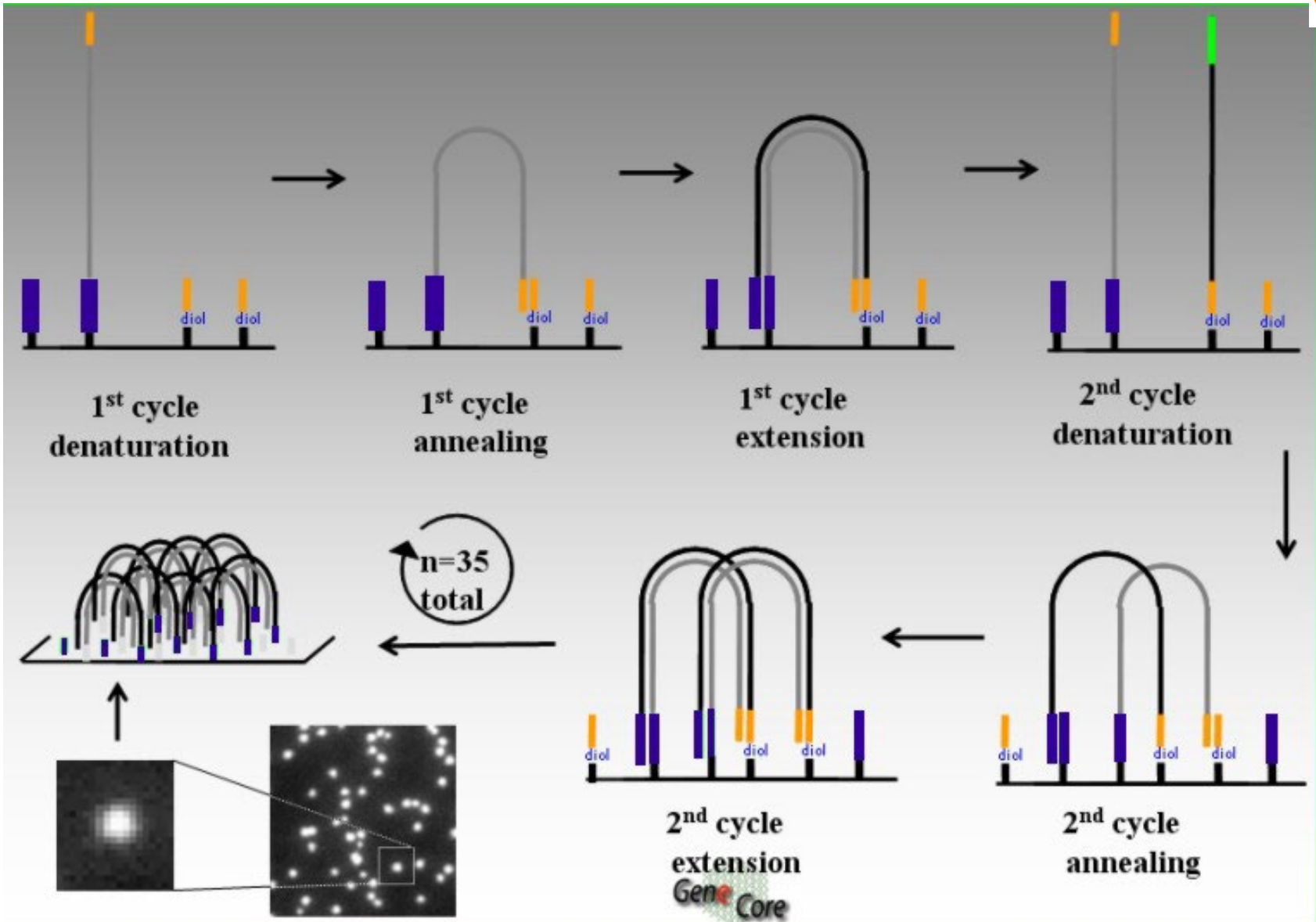




## Bridge amplification: initiation



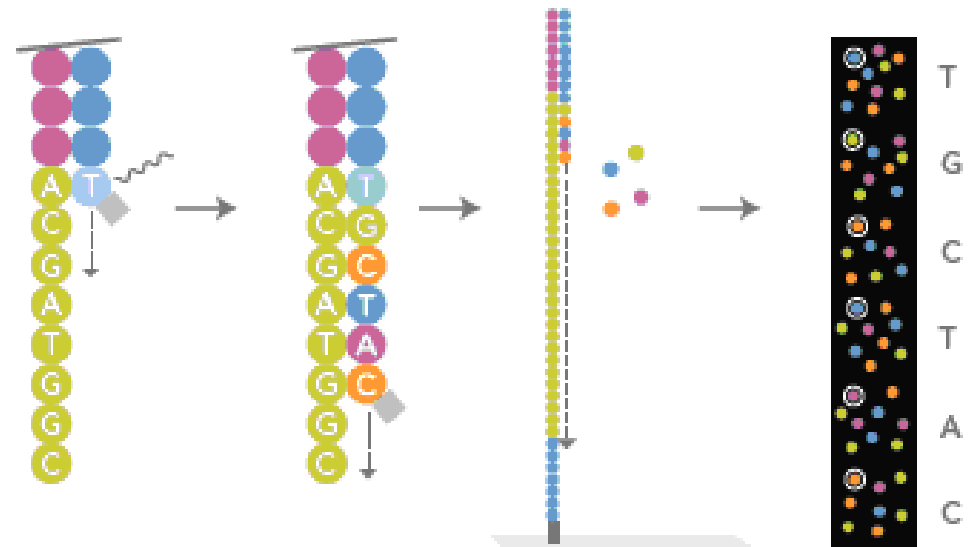
On the surface: complementary oligos





# Illumina

- **Průběh sekvenační reakce**
- Přidání primerů ve směsi s DNA polymerázou a fluorescenčně značenými dNTP
- Nukleotidy jsou na 3'-konci modifikovány tak, že umožňují **reverzibilní ukončení** prodlužujícího se řetězce DNA
- Je tak zajištěno, že se řetězec v každém cyklu prodlouží právě o jednu bázi.
- Po zachycení obrazu připojené báze následuje odblokování 3'-konce



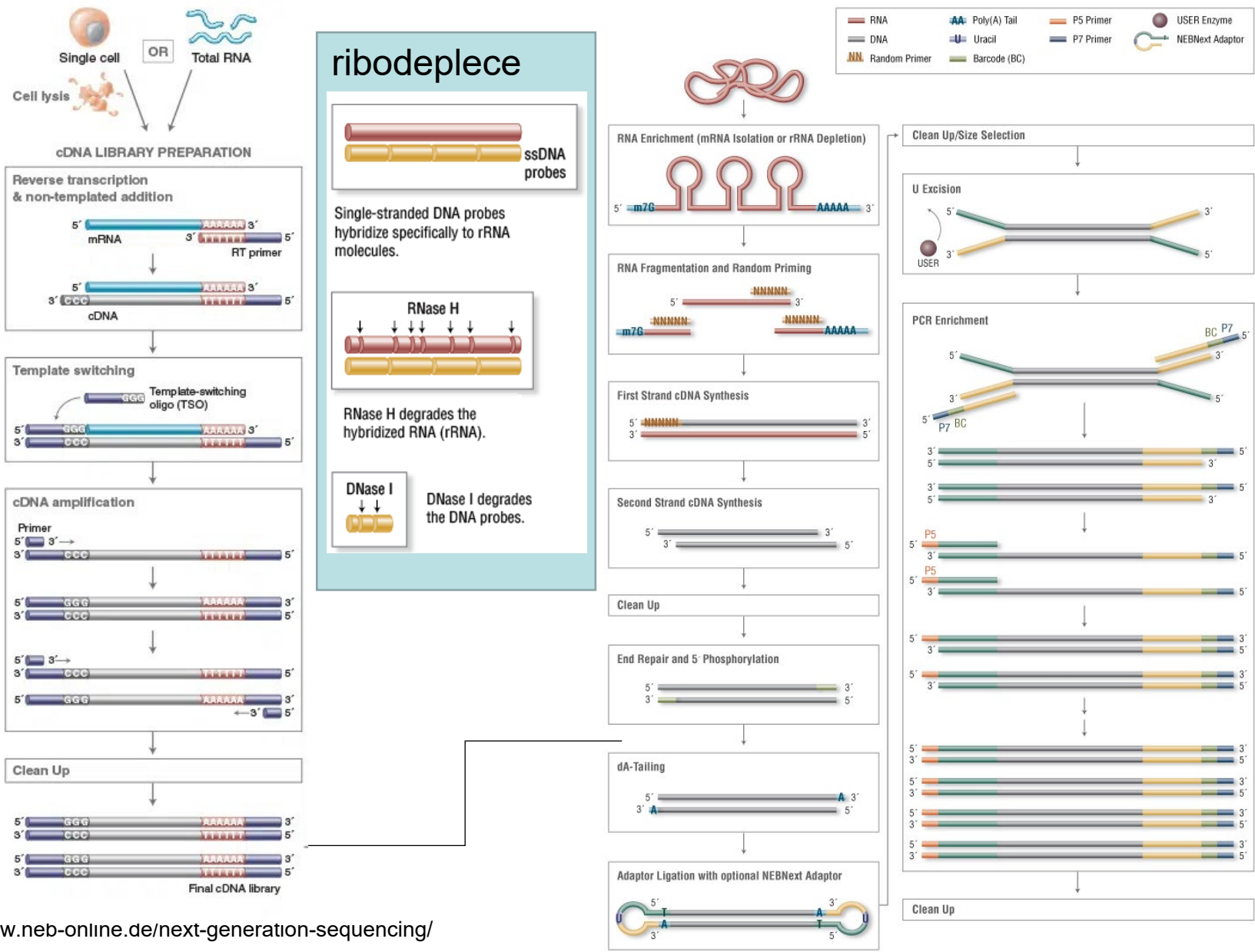
# Illumina

- Délka čtených úseků:
  - HiSeq : 75 nt or 150 nt
  - MiSeq : 300 nt
  - NextSeq : 2 x 150 nt
- Mnohonásobné pokrytí sekvence (30 x – 100 x)
- Možnost mnohonásobného sekvenování
  - Současně lze analyzovat až 96 různých vzorků.
- Produkuje desítky až stovky Gb za 1 běh (2 - 8 dnů)
- Aplikace
  - Resekvenování
  - Sekvenování RNA
  - Stanovení metylací

NextSeq



# Specifika přípravy knihoven pro RNA-seq



# Sekvenování třetí generace

- Analýza jedné molekuly
- Potřeba malého množství vzorku
- Bez rizika kontaminace
- Přesné čtení homopolymerních úseků
- Analýza jakékoli NK (DNA/RNA)
- Analýza poškozených NK (archaické, muzejní, forenzní)
- Analýza DNA z nekultivovatelných organismů
- Absolutní kvantifikace



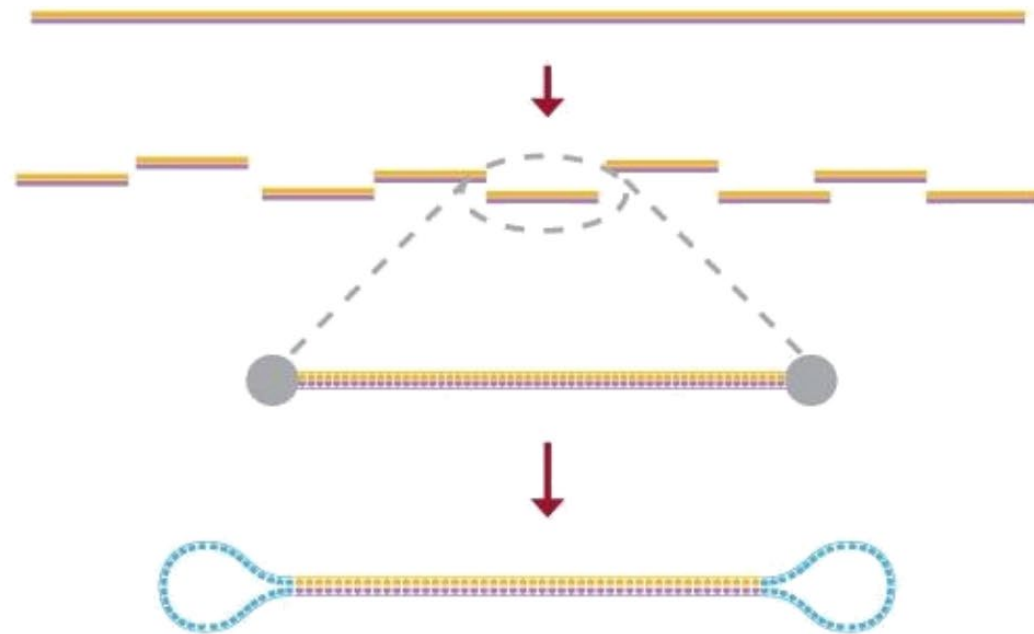
# PacBio

- **Kružnicové kontinuální sekvenování** (Single Molecule Real Time Sequencing – SMRT)
- Templát obsahuje **dvouřetězcovou oblast** (inzert) na obou koncích **uzavřenou jednořetězcovými vlásenkovými smyčkami**
- Vlášenkové smyčky představují jednořetězcovou oblast, na kterou se může vázat sekvenační primer.
- Polymeráza prodlužuje primer z jedné vlásenkové struktury využívá jedno vlákno DNA jako templát a druhé vytěsňuje
- Když se polymeráza dostane zpět k 5'konci primeru, začne vytlačovat již nasyntetizovaný řetězec a pokračuje v syntéze DNA
- Výsledná sekvence je odvozená z obou vláken



# Příprava knihovny pro PacBio

## Template Preparation

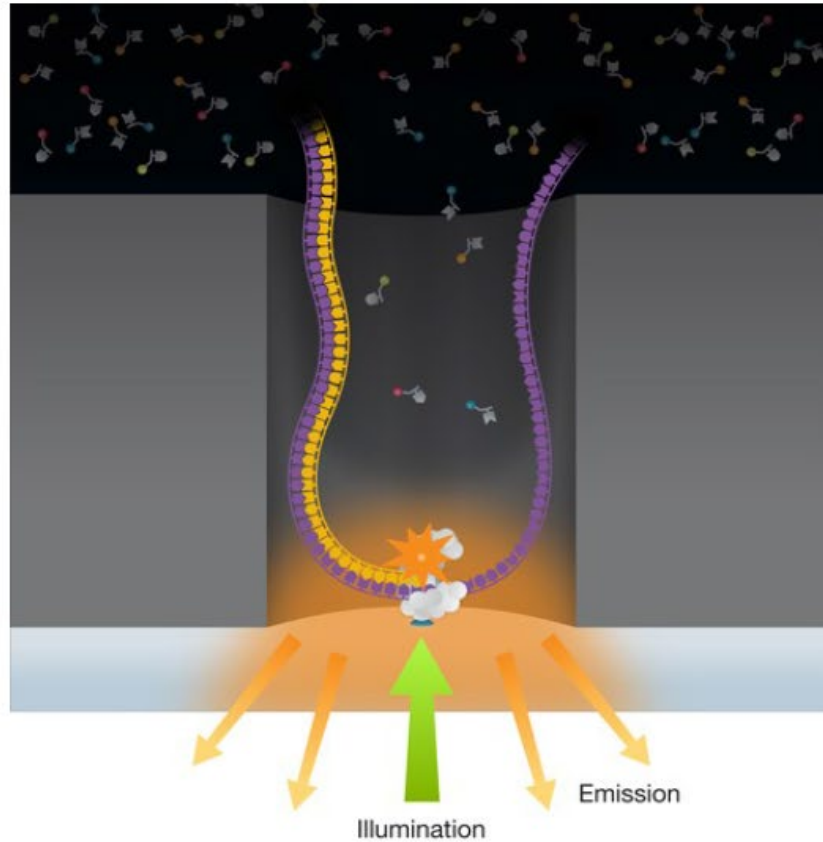


**SMRTbell™** Template preparation can be used to create libraries of various insert sizes from 250 bp to 20,000 bp depending on the needs of the application.

# Sekvenování třetí generace – PacBio RS



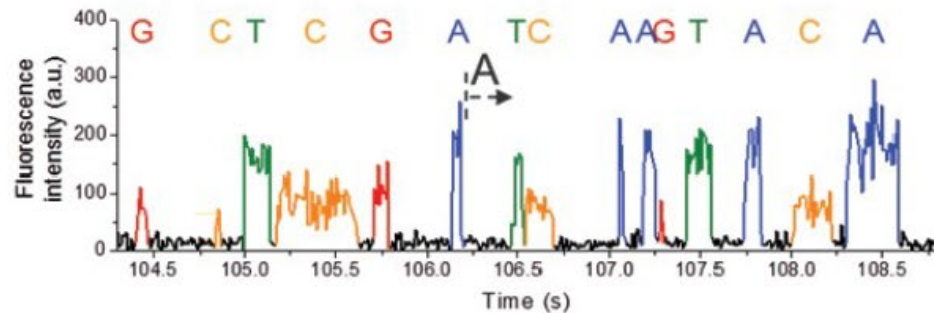
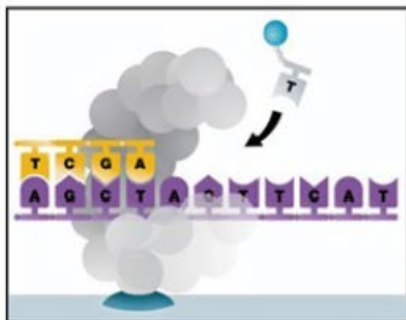
Pacific Biosciences



4 nucleotides with different fluorescent dye simultaneous present

2-3 nucleotides/sec  
2-3 Kb (up to 50) read length  
6 TB data in 30 minutes

laser damages polymerase

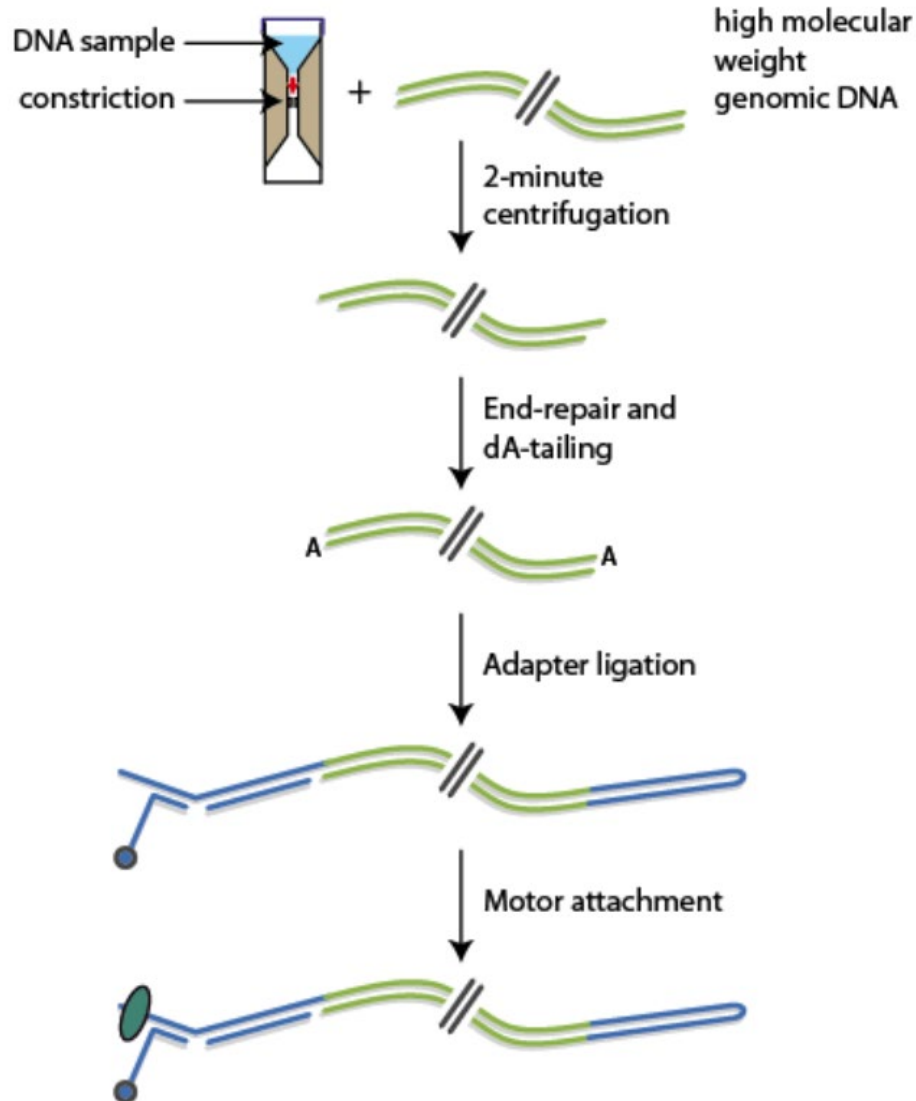


# Technologie NanoPore

- **Využívá měření elektrického potenciálu při průchodu DNA přes membránu (1995)**
- Elektricky odolná polymerní membrána (synтетická lipidová dvouvrstva) s transmembránovým porózním proteinem
- Modifikovaný  $\alpha$  hemolysin ( $\alpha$ HL) nebo porin A (MspA) *Mycobacterium smegmatis*
- Nanopore je ponořený ve vodivém roztoku
- Průchod bází na sekvenovaném řetězci přerušuje proud, který je **specifický podle procházející báze**

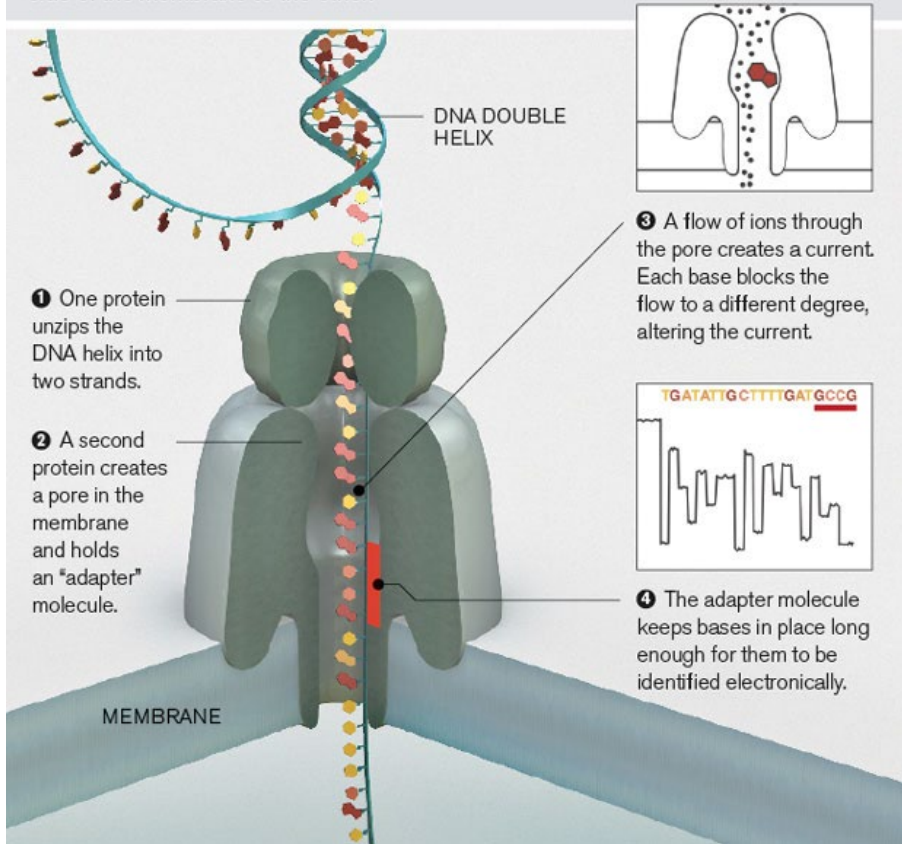


## G-tube sense/antisense

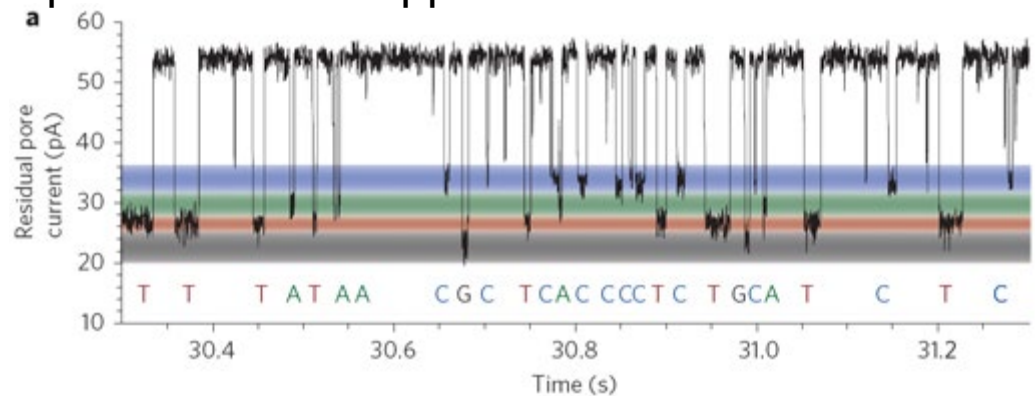


# Příprava knihovny MinION

DNA can be sequenced by threading it through a microscopic pore in a membrane. Bases are identified by the way they affect ions flowing through the pore from one side of the membrane to the other.



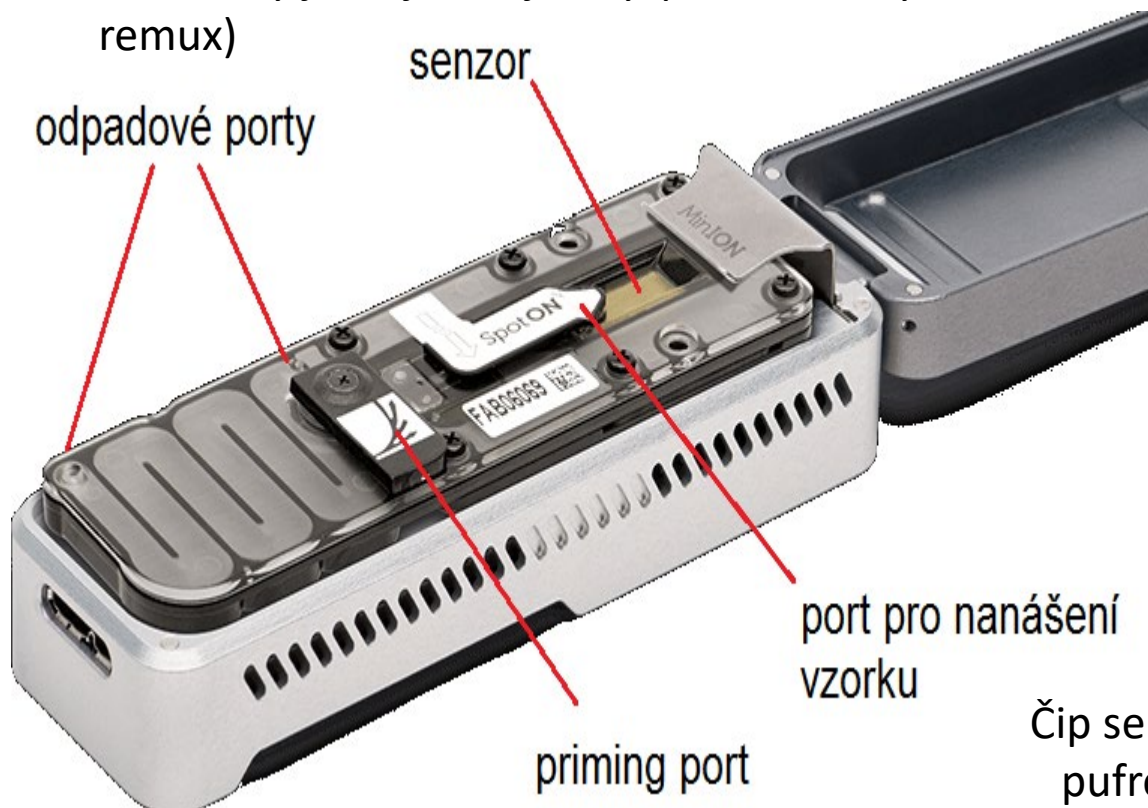
<https://nanoporetech.com/support/how-it-works>



# Flowcell - MinION



- 2048 jamek s póry (záruka na min. 800 funkčních nanopórů)
- 512 kanálů (ASIC obsahuje 1 zesilovač signálu na 4 jamky)
- čtení vždy jen z jedné jamky, po nastavených intervalech se střídají (tzv. remux)



nanášení vzorku

Čip senzoru musí být neustále zaplaven pufrem **vnesení bubliny nebo odsátí pufru vede k nevratnému poškození**

# fastf format dat

The screenshot shows the hdfview application interface. On the left is a file tree for a file named 'FAK84752\_484f7064fad5f...'. The tree includes folders for 'Analyses', 'Raw', and 'Signal'. The 'Analyses' folder is expanded to show 'Basecall\_1D\_000', which contains 'Fastq' and 'Trace'. The 'Raw' folder contains 'Signal'. The 'Signal' folder is selected, and its contents are displayed in the main window.

The main window displays three views:

- Text View:** Shows a text file named 'Fastq'. The content includes a header with read ID, run ID, flow cell ID, protocol group ID, and sample ID. The main body of the text is a FASTQ record for read 0, showing the sequence: `ATCATTGGGGTTTAACCGTTTTTCGATTATCGTAAACGCTTTCGCGTTTTCTGTCGCGCTCAATAGTCTGTAATCCCTCTTTAGTCTAGGAACCCCTCTCTACATCACGTA`.
- Table View - Trace:** Shows a table with 8 columns (0-7) and 11 rows (0-10). The data represents signal intensity for each column across different reads.
- Table View - Signal:** Shows a table with 1 column (0) and 10 rows (0-9). The data represents signal intensity for column 0 across different reads.

The Signal view also includes a **Lineplot** showing the signal intensity for column 0 across the read length (1 to 5872). The plot shows a noisy signal fluctuating around a mean value of approximately 450.



# Srovnání sekvenačních metod



# Srovnání metod pro sekvenování

	Metoda	Příprava knihovny
<b>Klasické</b>	Maxam-Gilbert (chemická)	Radioaktivně značené ssDNA
	Sanger (enzymová)	Inzerty v plazmidových vektorech / produkty PCR
<b>NGS</b>	454 (pyrosekvenování)	Emulzní PCR
	Polovodičové (IonTorrent)	Emulzní PCR
	Illumina	Polony PCR
	Solid	Emulzní PCR + imobilizace
<b>Třetí generace</b>	PacBio	Ligace adaptorů se smyčkou
	Nanopore	Ligace adaptorů + molekulární motor
	Helicos	Ligace adaptorů



# Srovnání metod pro sekvenování

	Metoda	Princip reakce
<b>Klasické</b>	Maxam-Gilbert (chemická)	Degradace chemickými činidly
	Sanger (enzymová)	Syntéza DNA polymerázou s terminátory, elektroforéza
<b>NGS</b>	454 (pyrosekvenování)	Syntéza DNA polymerázou, detekce pyrofosfátu (PP)
	Polovodičové (IonTorrent)	Syntéza DNA polymerázou, detekce H <sup>+</sup>
	Illumina	Syntéza DNA polymerázou s reverzibilními terminátory
	Solid	Série ligací sond + posun rámce
<b>Třetí generace</b>	PacBio	Syntéza DNA polymerázou
	Nanopore	Měření změn el. potenciálu při průchodu inertní membránou
	Helicos	Syntéza DNA polymerázou s reverzibilními terminátory



# Srovnání metod pro sekvenování

Metoda		Délka čtení
<b>Klasické</b>	Maxam-Gilbert (chemická)	150 – 300 nt
	Sanger (enzymová)	600 – 1200 nt
<b>NGS</b>	454 (pyrosekvenování)	500 nt
	Polovodičové (IonTorrent)	400 nt
	Illumina	1× 300 nt, 2× 150 nt, 2× 75 nt
	Solid	30 – 40 nt
<b>Třetí generace</b>	PacBio	10 – 15 kb
	Nanopore	> 10 kb
	Helicos	200 nt

# Srovnání metod pro sekvenování

	Metoda	Uspořádání
<b>Klasické</b>	Maxam-Gilbert (chemická)	Single read
	Sanger (enzymová)	Single read, Pair-end read
<b>NGS</b>	454 (pyrosekvenování)	Single read, Pair-end read
	Polovodičové (IonTorrent)	Single read, Pair-end read
	Illumina	Single read, Mate-pair read
	Solid	Mate-pair read
<b>Třetí generace</b>	PacBio	Single read
	Nanopore	Single read
	Helicos	Single read