

Technologie sekvenování

- 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)
- ◆ Identifikace mikroorganismů (*de novo* sekvenování a resekvenování)
- ◆ Analýza transkriptomu (RNAseq)
- ◆ Metagenomika (16S rRNA, celé metagenomy)
- ◆ DNA metylační studie (medip-seq, bisulfite treated DNA)
- ◆ Studie transkripčních faktorů, proteinů vázajících se na DNA (ChIPseq)
- ◆ miRNAs, siRNA, piRNA, tRF, aj... (small RNA seq)

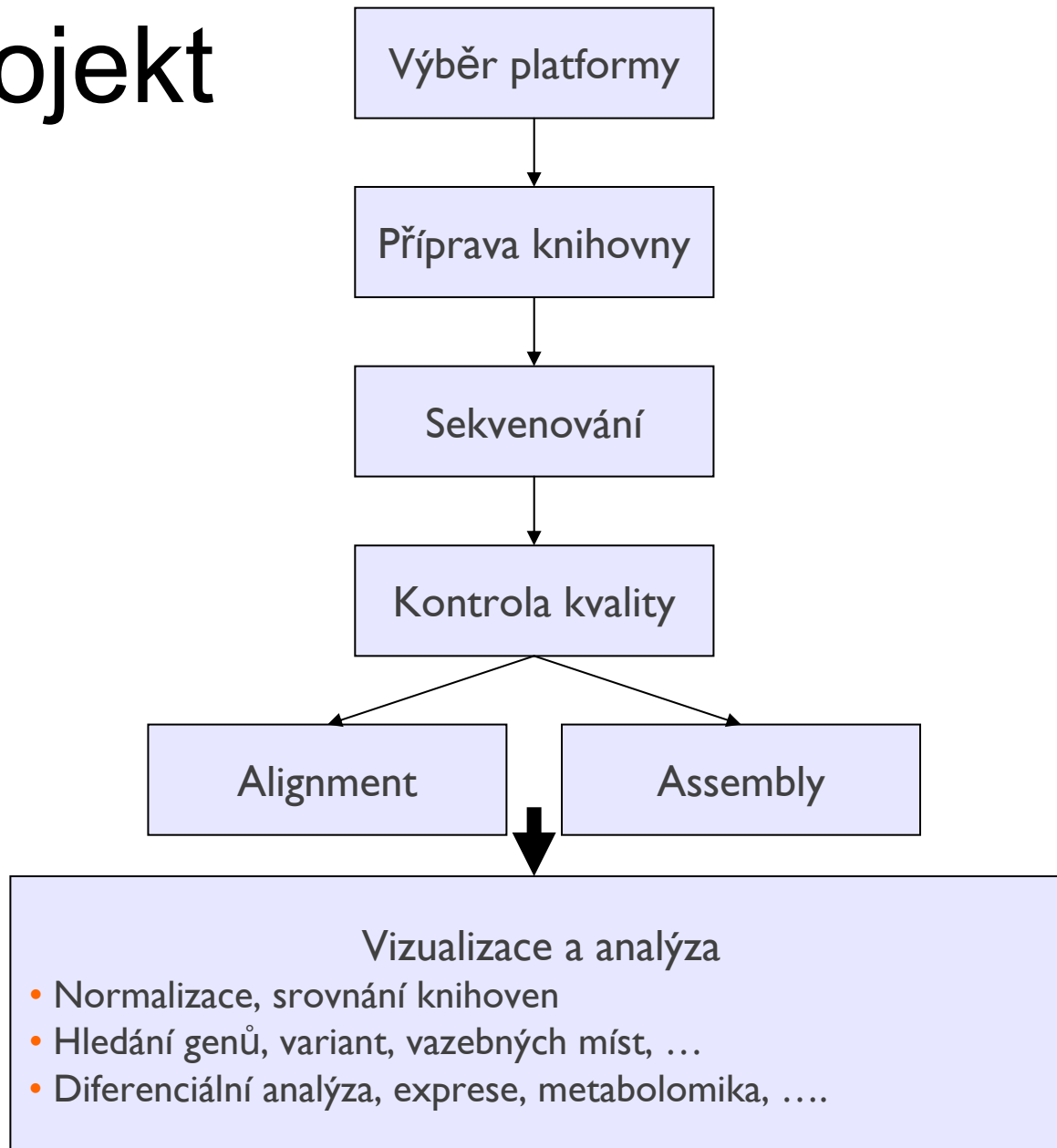
Techniky sekvenování

- 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, IonTorrent, Illumina, SoLiD**
- Metody sekvenování třetí generace
 - **PacBio, NanoPore, Helicos, etc.**

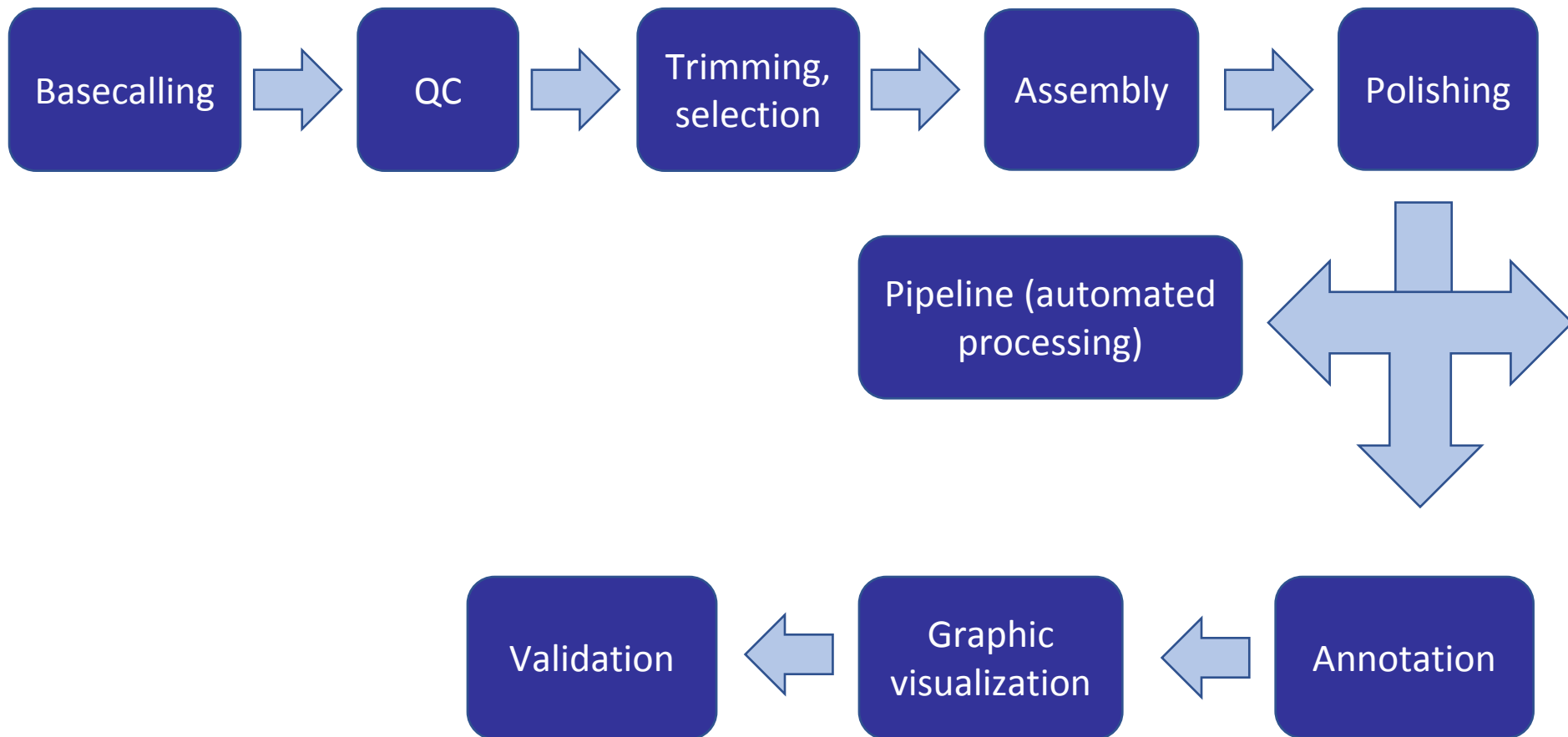
Princip sekvenačních metod

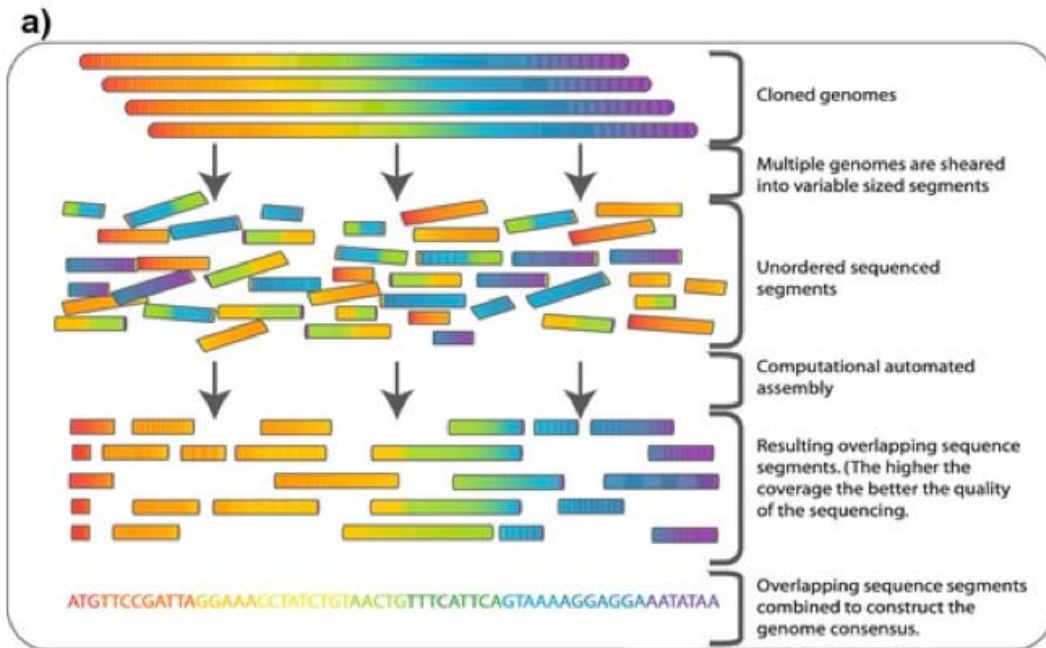
- **Klasické sekvenování kapilární elektroforézou**
 - Vyžaduje velký počet kopií vstupního materiálu DNA jako templátu pro přípravu jednořetězců
- **Sekvenování nové generace**
 - Jako templát slouží jediná molekula, která je amplifikována pro získání dostatečného signálu, produkuje **krátká** čtení,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

Projekt

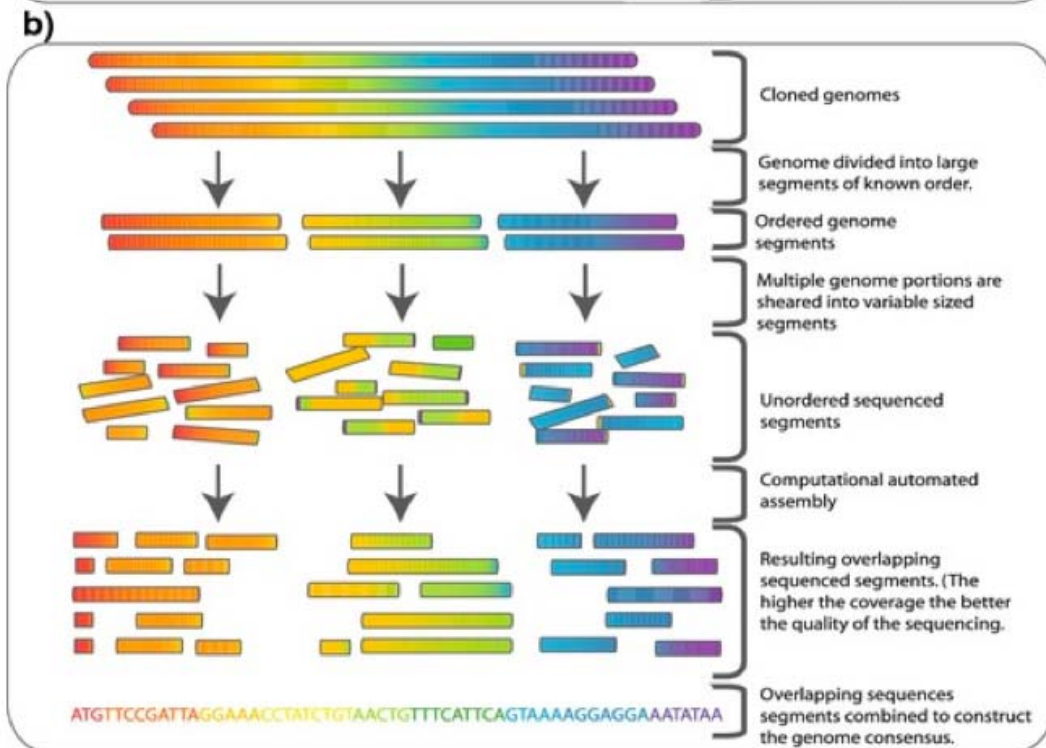


Bioinformatické zpracování





Celogenomové
shotgun (náhodné)
sekvenování



versus

postupné shot-gun
sekvenování

Technické požadavky pro úspěšné stanovení sekvence

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

Postup enzymatického sekvenování DNA

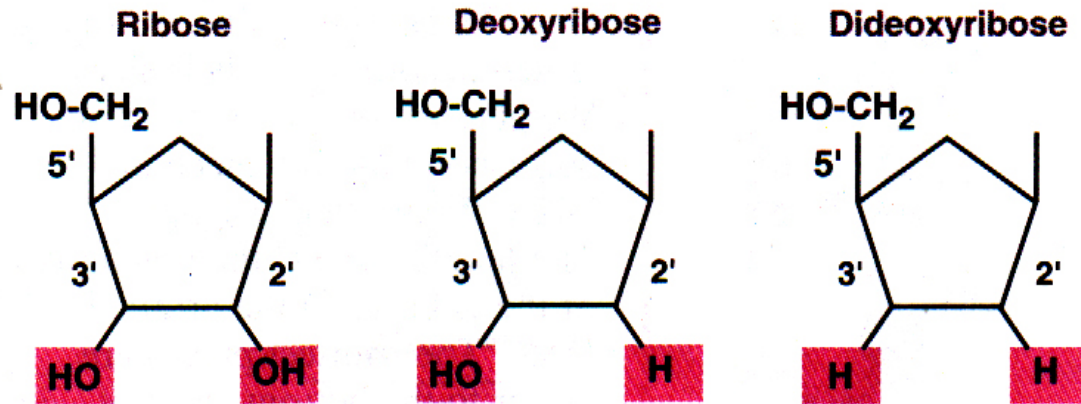
1. Příprava ssDNA jejíž sekvence má být stanovena
2. Rozdělení do 4 vzorků
3. Reakce s DNA polymerázou při níž se začlení do syntetizované DNA místo normálního deoxyribonukleosidtrifosfátu jeho analog, který působí jako koncový inhibitor syntézy DNA - dideoxynukleosidtrifosfát (ddNTP). V každém ze 4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

Reakce obsahuje:

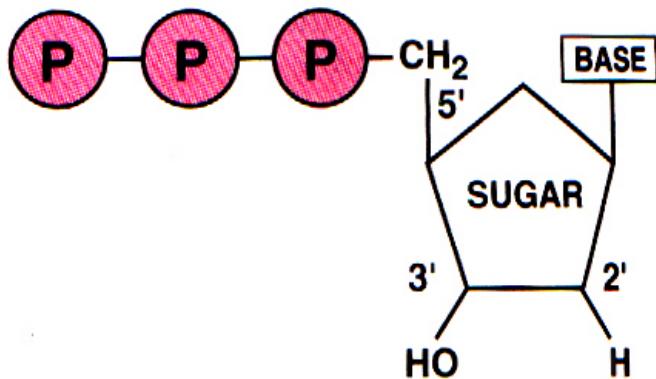
- molekulu DNA
- značený primer při pojující se k části molekuly DNA (místo odkud začínáme stanovovat sekvenci)
- směs obsahující 4 normální nukleotidy
- jeden ddNTP (v každém ze 4 vzorků je jiný)
- DNA polymerázu

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

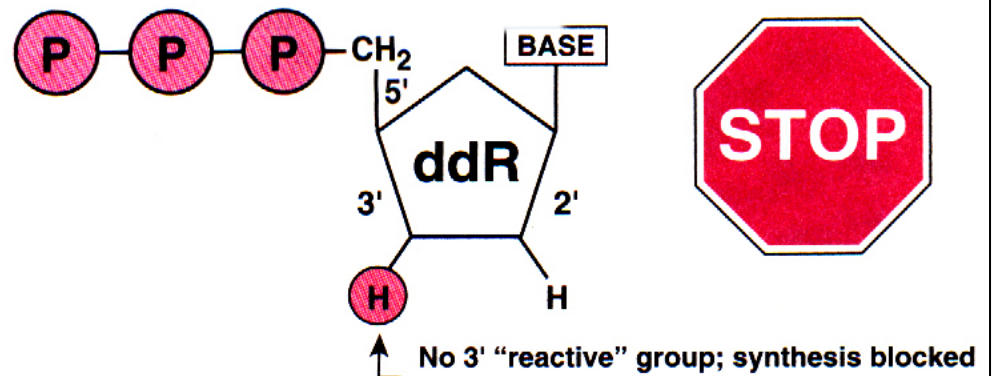
Dideoxynukleotidy nemají hydroxyl na 3'-konci



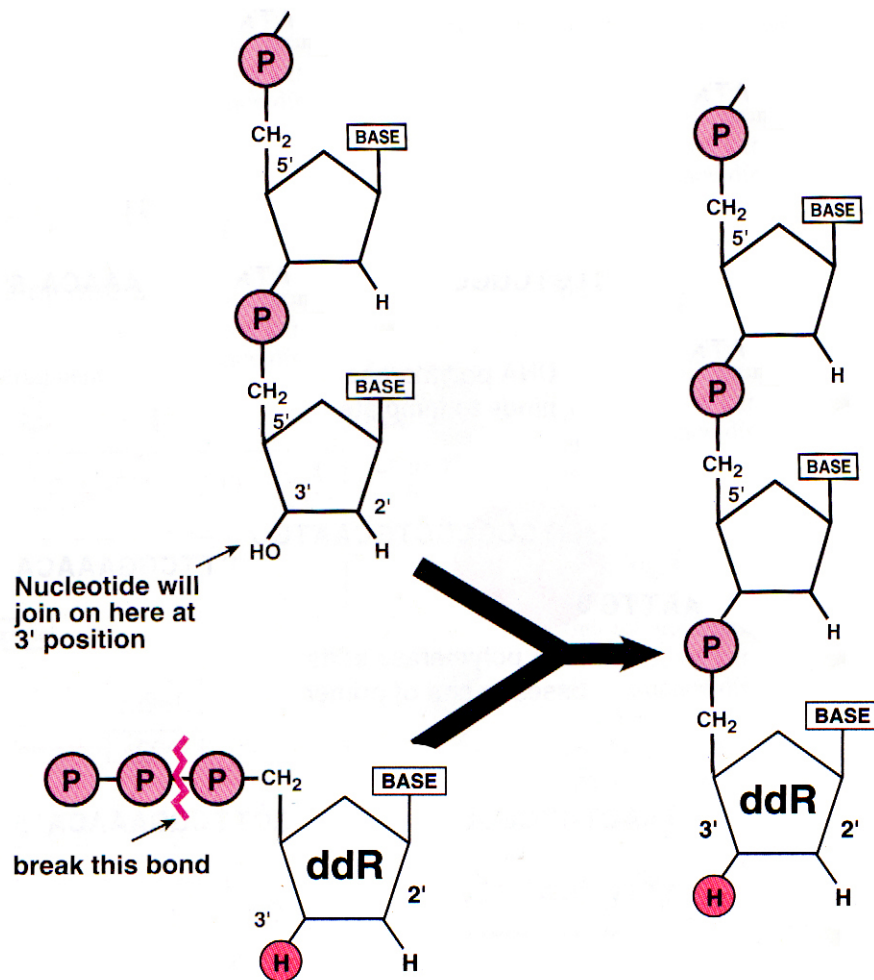
23.4 DEOXYNUCLEOSIDE TRIPHOSPHATE



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

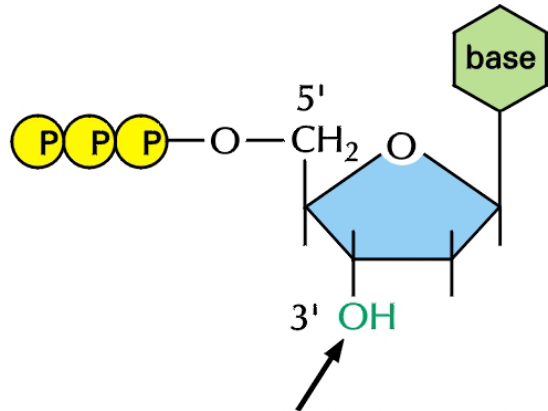
~~A G C C T G G C G A C C A T C G T
T C G G A C C G C T G G T A G C A~~

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

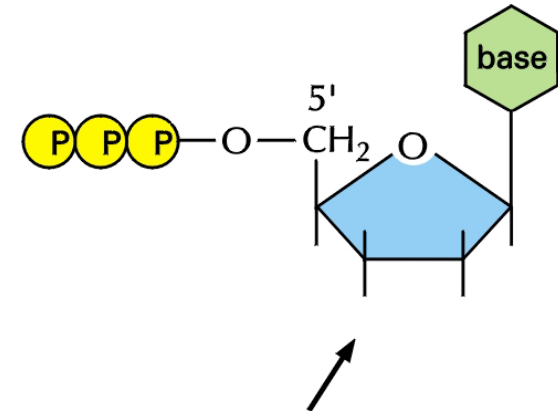
Enzymová metoda sekvenování DNA (Sanger)

(A) **deoxy**ribonukleosidtrifosfát



je možno prodloužit řetězec na 3' -konci

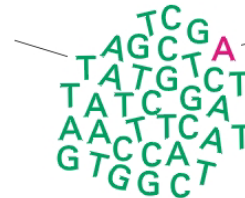
dideoxyribonukleosidtrifosfát



není možno prodloužit řetězec na 3' -konci

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

malé množství jednoho **dideoxy**ribonukleosidtrifosfátu (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

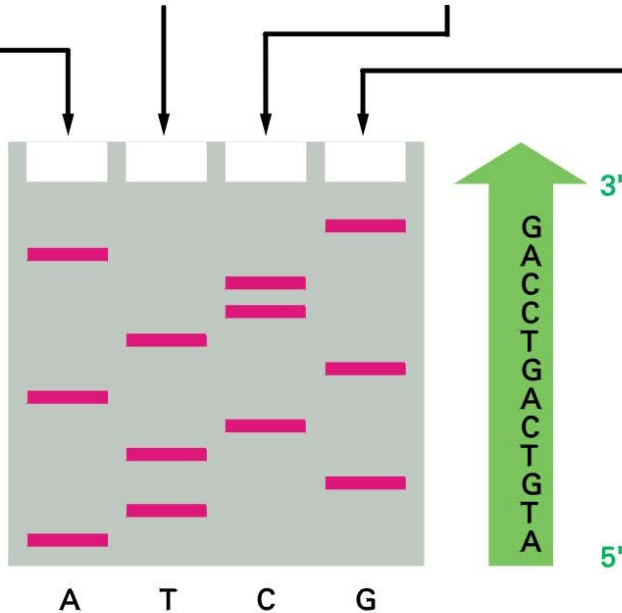
označený primer

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

+ nadbytek dATP
dTTP
dCTP
dGTP

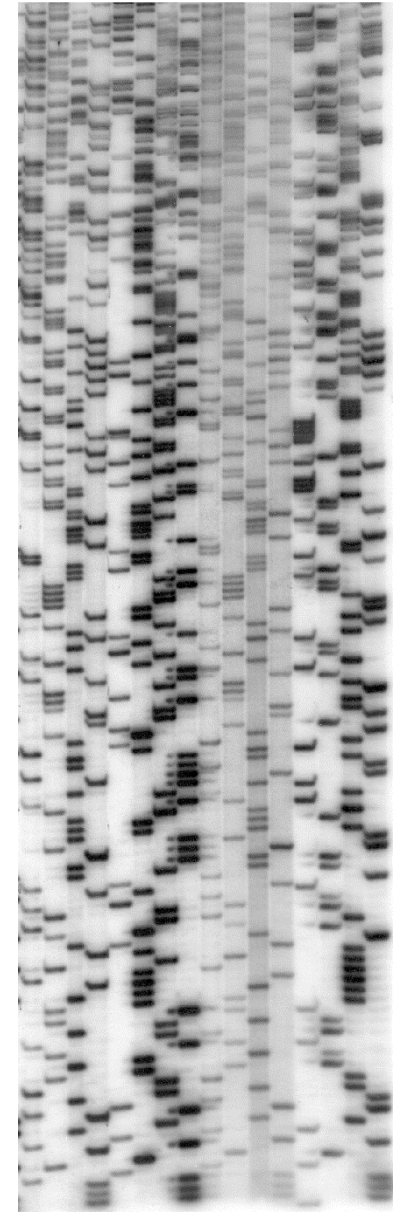
+ ddATP + DNA-polymeráza + ddTTP + DNA-polymeráza + ddCTP + DNA-polymeráza + ddGTP + DNA-polymeráza

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



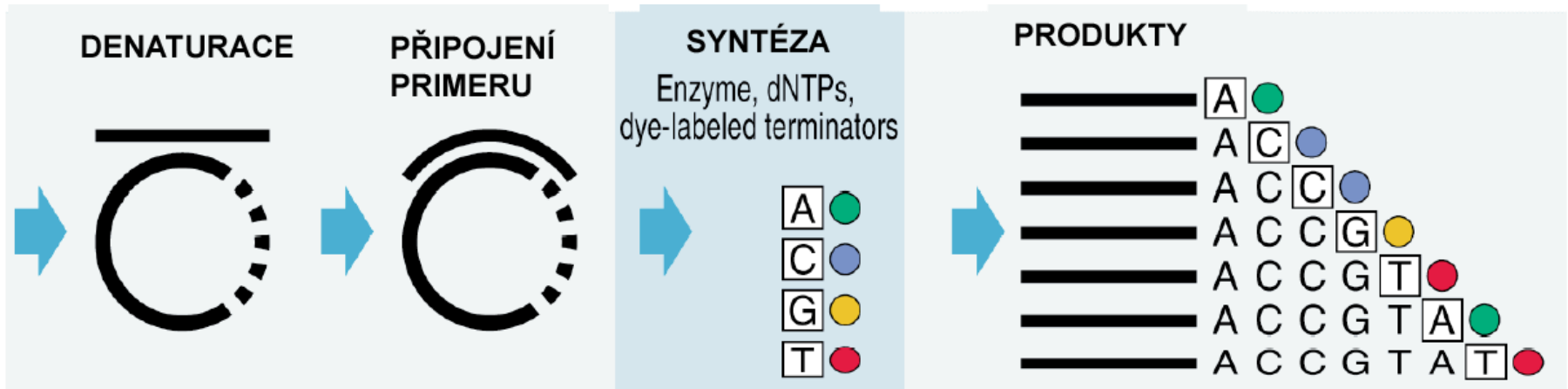
ATGC ATGC ATGC ATGC

Automatické sekvenování DNA

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

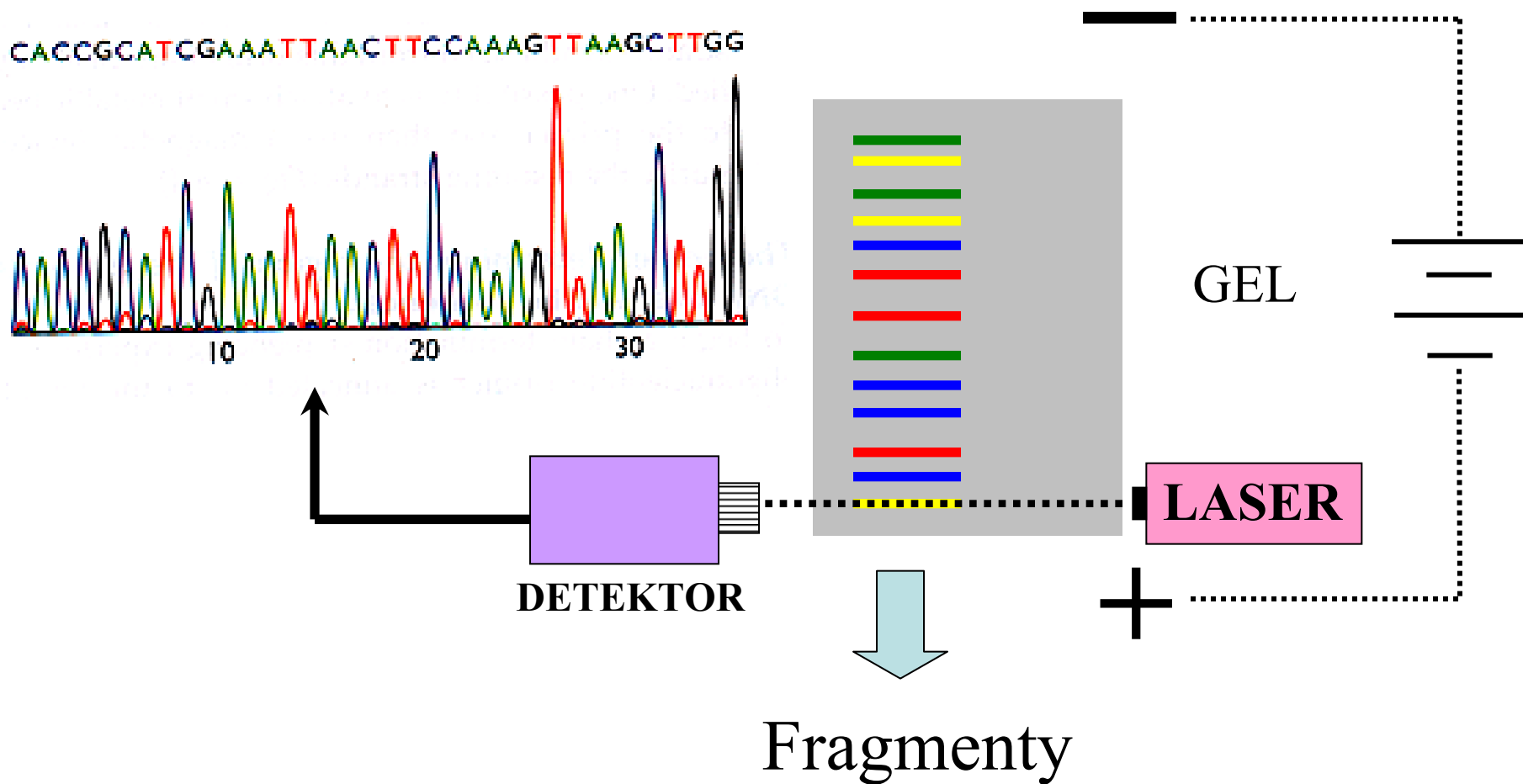
Strategie barevných terminátorů

asymetrická PCR



Princip detekce produktů

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



Genetický analyzátor ABI 3100

Soustava
kapilár

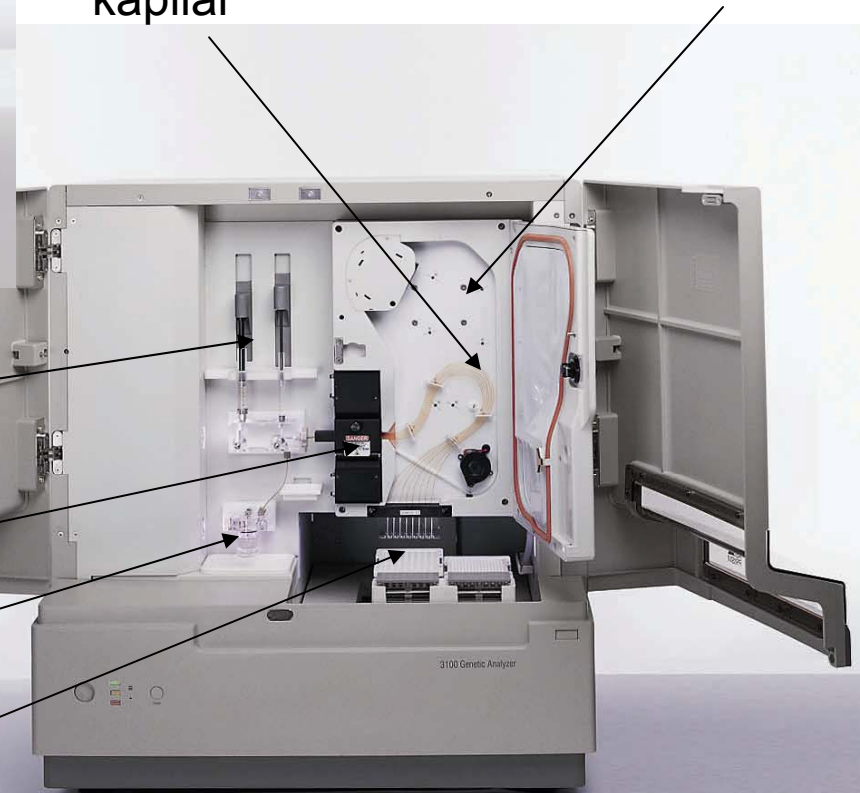
Vyhřívaná deska

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

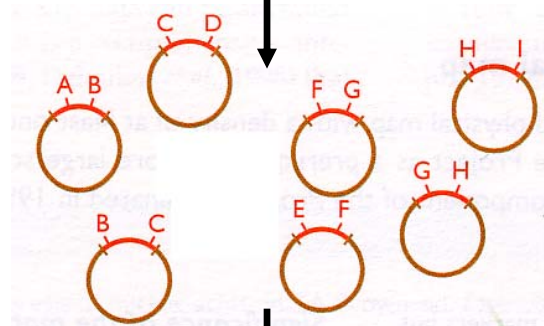


+



Vektor

Úprava konců a klonování

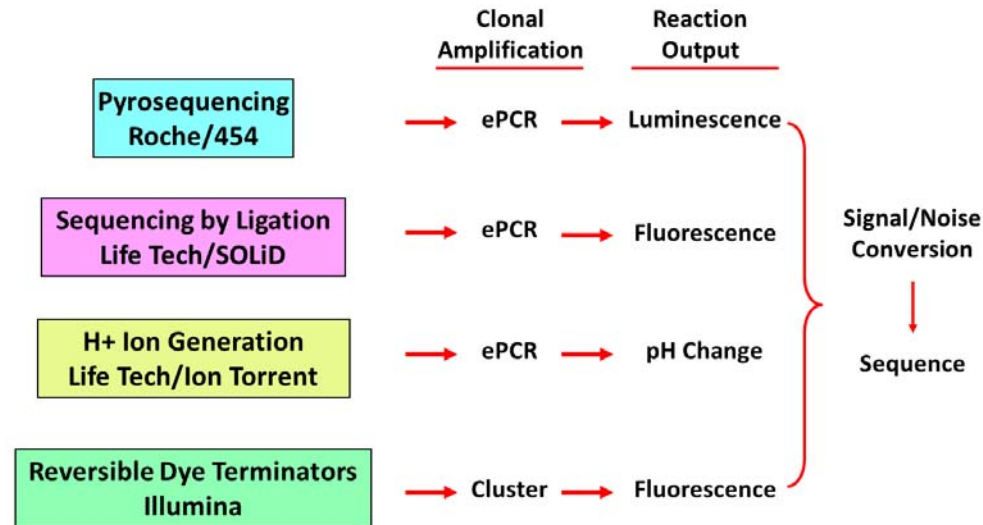


Sestavení překrývajících se klonů



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)

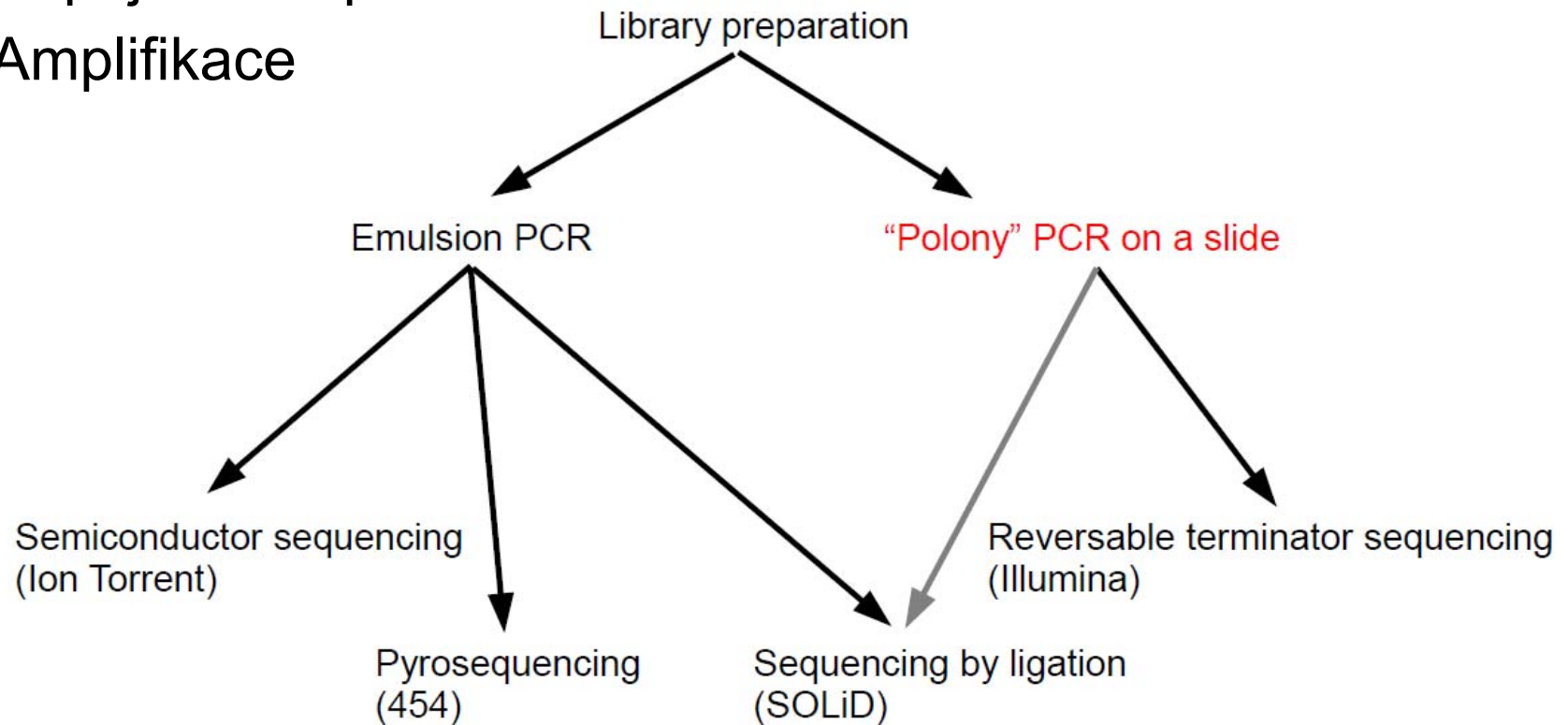


• Sekvenování třetí generace

- Pacific Biosciences
- Helicos Biosciences
- NABsys
- VisiGen Biotechnologies
- Oxford Nanophore Technologies

Metody přípravy knihovny

- Fragmentace
 - Enzymatická
 - Fyzikální (sonikace, nebulizace, Hydroshear)
- Připojení adaptorů
- Amplifikace



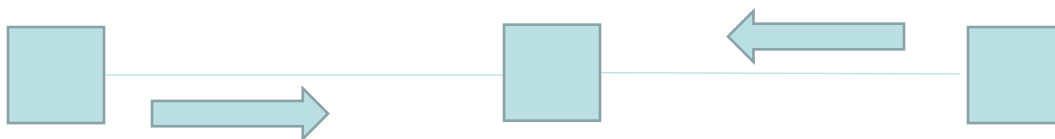
Typy uspořádání NGS sekvenování



Single Read



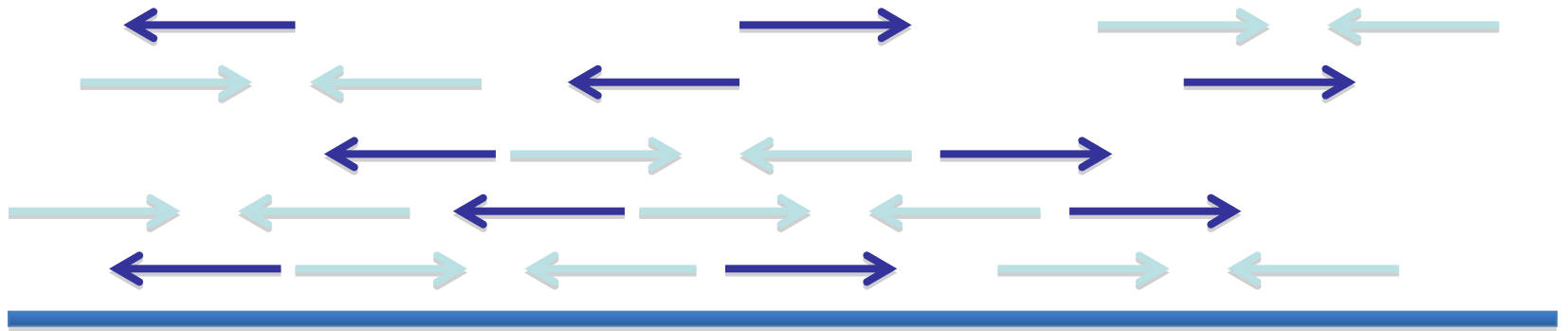
Paired-end read (PE)



Mate-pair read (MP)

výsledkem jsou miliony krátkých čtení sekvencí z náhodných částí genomu

Orientace párových čtení



  Paired end (PE) reads

  Mate pair (MP) reads

Hloubka sekvenování (pokrytí)

- Coverage (pokrytí) = (počet získaných čtení × průměrná délka čtení)/velikost genomu
- Příklad 1: Získali jsme 10^6 čtení s technologií IonTorrent, velikost genomu je 3 Mb. Jaké je pokrytí?
- Příklad 2: Technologie Illumina poskytuje 200 milionů čtení. Kolik genomů o velikosti 5 Mb mohu sekvenovat, abych získal pokrytí 50×?

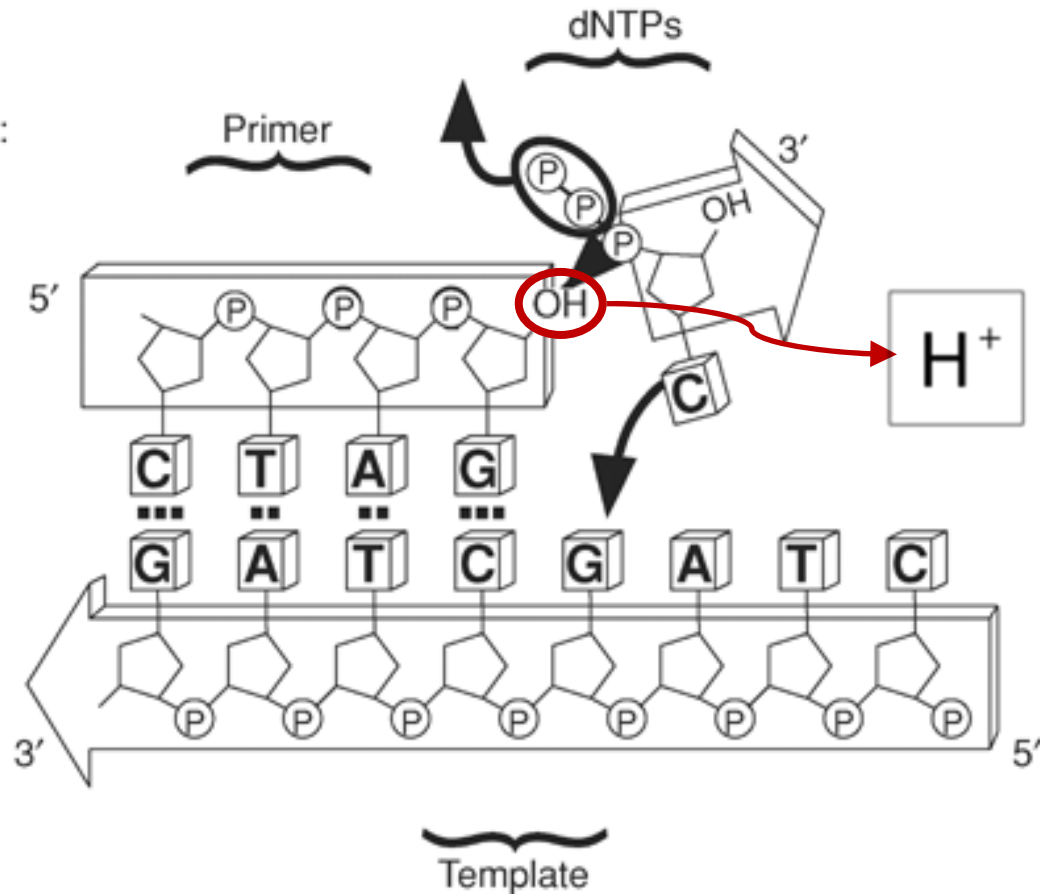
Ion Torrent polovodičové sekvenování

- prvním sekvenátor DNA, který **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

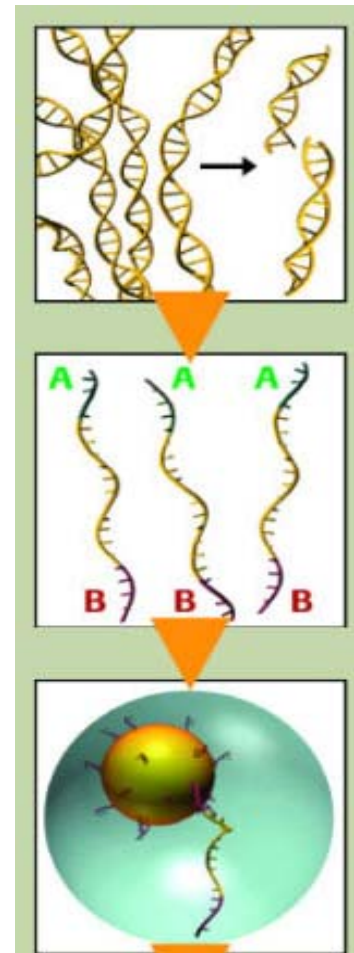


Example:



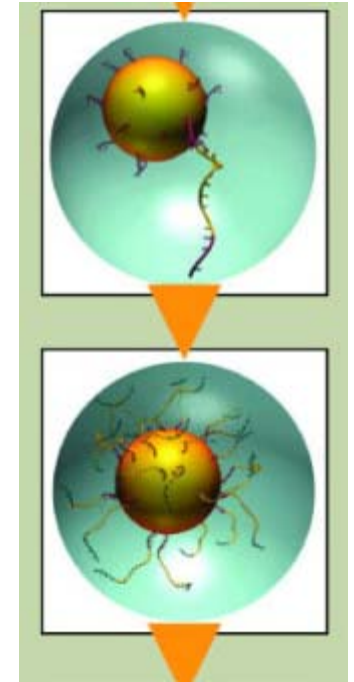
Knihovna

- Příprava jednořetězcové DNA knihovny
 - Umožňuje zpracovat různé typy vzorků
 - genomové DNA
 - produkty PCR
 - BACs
 - cDNA
 - Frakcionace dlouhých úseků na 200- až 400-bp dlouhé fragmenty
 - Vazba dvou adaptorů specifických pro 3'- a 5'-konce na každý fragment
 - Adaptor A obsahuje vazebné místo pro primer
 - Adaptor B B obsahuje 5'-biotinovou značku, která slouží k imobilizaci DNA knihovny na mikrosféry pokryté streptavidinem



Klonální amplifikace knihovny emulzní PCR

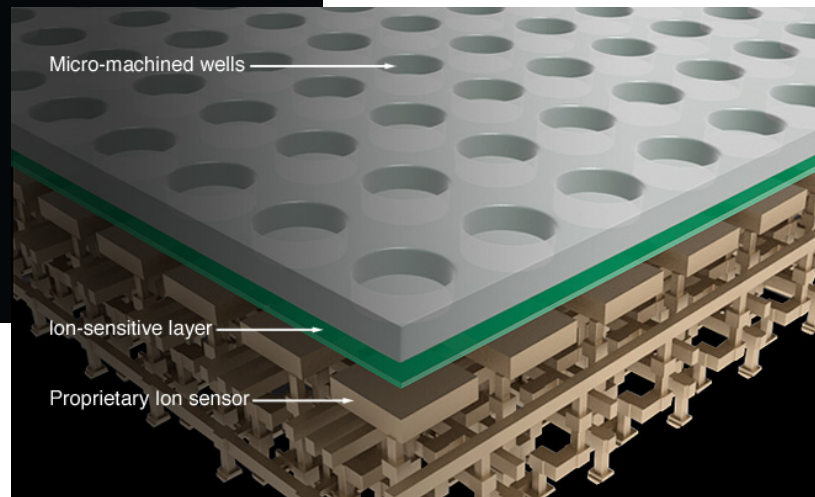
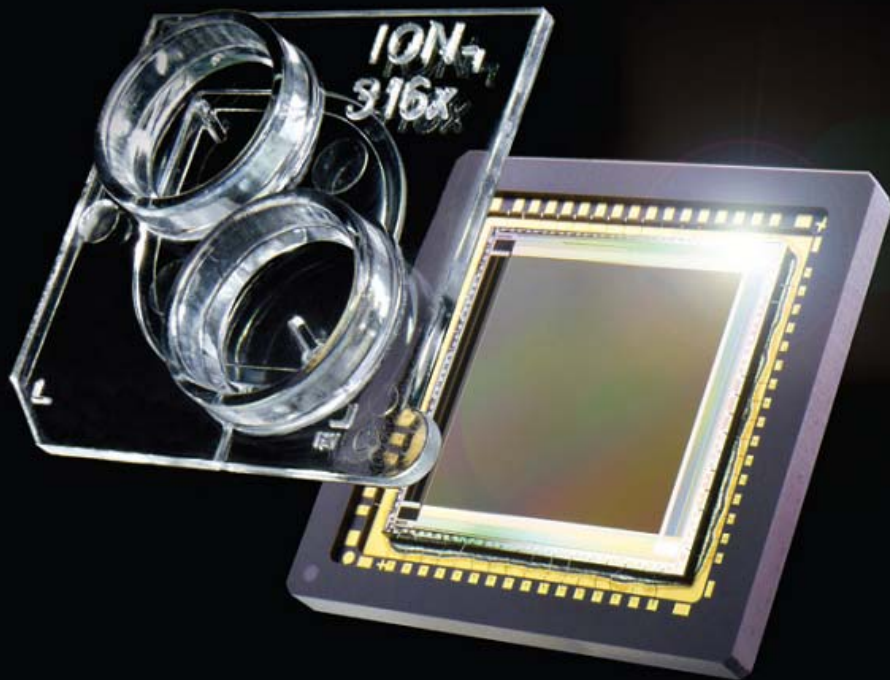
- Klonální amplifikace knihovny
 - **emulzní PCR**, amplifikační reagensie ve směsi voda-olej → mikroreaktory obsahující jedinou kuličku s unikátním fragmentem DNA z knihovny - (10^7) kopií
- Separace kuliček a výběr pouze těch, které obsahují amplifikovanou DNA (enrichment)
- Sekvenační reakce
- Analýza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

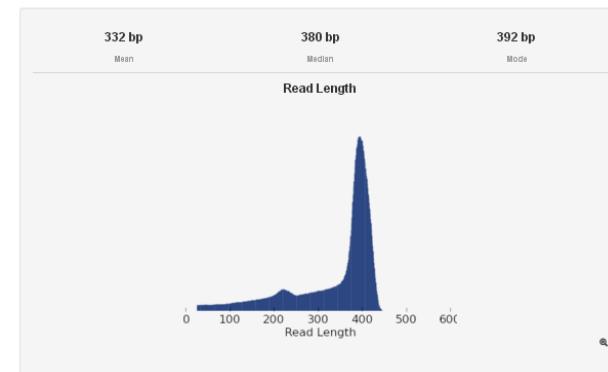
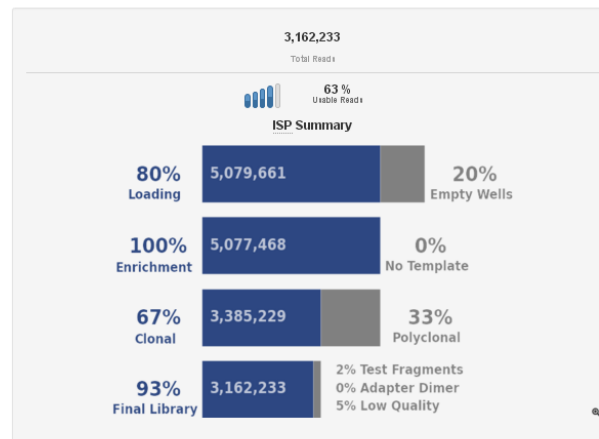
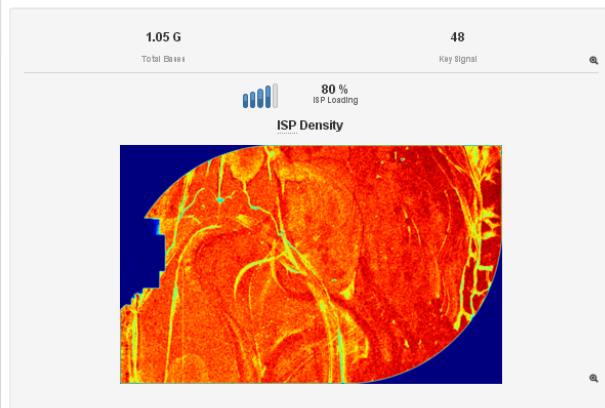


Ion Torrent polovodičové sekvenování

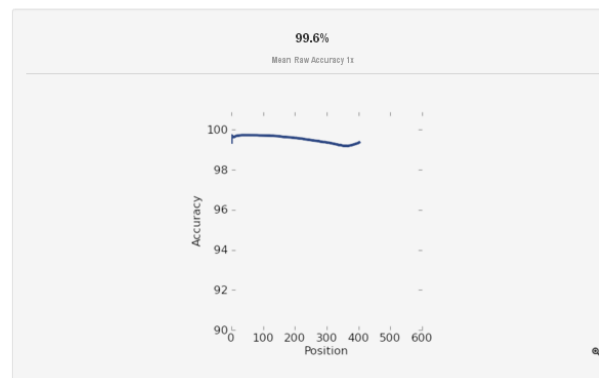
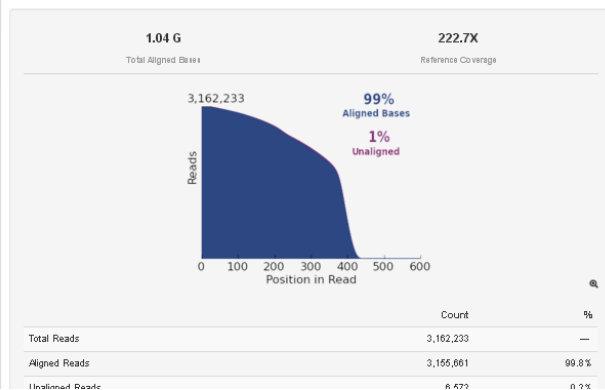
- Technologie Ion Torrentu je otevřená pro sekvenování DNA de-novo i cílené sekvenování vybraných oblastí genomu
- Minimální délka sekvenačního čtení
 - V jedno směru: 400bp
 - Obousměrné (Paired end)
 - Počet vzorků analyzovaných v jednom běhu
 - až 96 vzorků – pomocí označení tzv. barkodem
- Doba sekvenační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

GS De novo Assembler

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

Ready for analysis

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 TCTTCTCCTAAAAACC-TAAAGTATCAAGG-TTTTT-ACAAACCTCTACATT-AGGTAAGT-AACCGAAAGGTTCTATACTAGAATCTGTATATAAACTATTTGT
HK6IEBF01BN89B TCTTCTCCTAAAAACC-TAAAGTATCAAGG-TTTTT-ACAAACCTCTACATT-AGGTAAGT
contig00007 HK6IEBF01A3MAV TCTTCTCCTAAAAACC-TAAAGTATCAAGG-TTTTT-ACAAACCTCTACATT-AGGTAAGT
contig00008 HK6IEBF01BHKX2 TCTTCTCCTAAAAACC-TAAAGTATCAAGG-TTTTT-ACAAACCTCTACATT-AGGTAAGT-AACCGAAAGGTTCTATACTAGAATC
contig00009 HK6IEBF01A5THI TCTTCTCCTAAAAACC-TAAAGTATCAAGG-TTTTT-ACAAACCTCTACATT-AGGTAAGT-AACCGAAAGGTTCTATACTAGAATCTGTATATAAACTATTTGT
contig00010 HK6IEBF01AHUA1 TCTTCTCCTAAAAACC-TAAAGTATCAAGG-TTTTT-ACAAACCTCTACATT-AGGTAAGT-AACCGAAAGGTTCTATACTAGAATCTGTATATAAACTATTTGT
contig00011 HK6IEBF01A9E9P TCTTCTCCTAAAAACC-TAAAGTATCAAGG-TTTTT-ACAAACCTCTACATT-AGGTAAGT-AACCGAAAGGTTCTATACTAGAATCTGTATATAAACTATTTGT
contig00012 HK6IEBF01AV175 TCTTCTCCTAAAAACC-TAAAGTATCAAGG-TTTTT-ACAAACCTCTACATT-AGGTAAGT-AACCGAAAGGTTCTATACTAGAATCTGTATAT
contig00013 HK6IEBF01BJY8L TCTTCTCCTAAAAACC-TAAAGTATCAAGG-TTTTT-ACAAACCTCTACATT-AGGTAAGT-AACCGAAAGGTTCTATACTAGAATCTGTATATAAACT
contig00014 HK6IEBF01AI6AG TCTTCTCCTAAAAACC-TAAAGTATCAAGG-TTTTT-ACAAACCTCTACATT-AGGTAAGT-AACCGAAAGGTTCTATACTAGAATCTGTATATAAACTATTTGT
contig00015 HK6IEBF01AW7YG TCTTCTCCTAAAAACC-TAAAGTATCAAGG-TTTTT-ACAAACCTCTACATT-AGGTAAGT-AACCGAAAGGTTCTATACTAGAATCTGTATATAAACTATTTGT
contig00016 HK6IEBF01B1IE6 TCTTCTCCTAAAAACC-TAAAGTATCAAGG-TTTTT-ACAAACCTCTACATT-AGGTAAGT-AACCGAAAGGTTCTATACTAGAATCTGTATATAAACTATTTGT
contig00017 HK6IEBF01AYV3H TCTTCTCCTAAAAACC-TAAAGTATCAAGG-TTTTT-ACAAACCTCTACATT-AGGTAAGT-AACCGAAAGGTTCTATACTAGAATCTGTATATAAACTATTTGT
contig00018 HK6IEBF01BG5KV TCTTCTCCTAAAAACC-TAAAGTATCAAGG-TTTTT-ACAAACCTCTACATT-AGGTAAGT-AACCGAAAGGTTCTATACTAGAATCTGTATATAAACTATTTGT
contig00019 HK6IEBF01BJFDR TCTTCTCCTAAAAACC-TAAAGTATCAAGG-TTTTT-ACAAACCTCTACATT-AGGTAAGT-AACCGAAAGGTTCTATACTAGAATCTGTATATAAACTATTTGT
contig00020 HK6IEBF01B04U2 TCTTCTCCTAAAAACC-TAAAGTATCAAGG-TTTTT-ACAAACCTCTACATT-AGGTAAGT-AACCGAAAGGTTCTATACTAGAATCTGTATATAAACTATTTGT
contig00021 HK6IEBF01A39J2 TCTTCTCCTAAAAACC-TAAAGTATCAAGG-TTTTT-ACAAACCTCTACATT-AGGTAAGT-AACCGAAAGGTTCTATACTAGAATCTGTATATAAACTATTTGT
contig00022 HK6IEBF01BSDD8 TCTTCTCCTAAAAACC-TAAAGTATCAAGG-TTTTT-ACAAACCTCTACATT-AGGTAAGT-AACCGAAAGGTTCTATACTAGAATCTGTATATAAACTATTTGT
contig00023 HK6IEBF01BPQRL TCTTCTCCTAAAAACC-TAAAGTATCAAGG-TTTTT-ACAAACCTCTACATT-AGGTAAGT-AACCGAAAGGTTCTATACTAGAATCTGTATATAAACTATTTGT
contig00024 HK6IEBF01BBCE9 TCTTCTCCTAAAAACC-TAAAGTATCAAGG-TTTTT-ACAAACCTCTACATT-AGGTAAGT-AACCGAAAGGTTCTATACTAGAATCTGTATATAAACTATTTGT
contig00025 HK6IEBF01B1K98 TCTTCTCCTAAAAACC-TAAAGTATCAAGG-TTTTT-ACAAACCTCTACATT-AGGTAAGT-AACCGAAAGGTTCTATACTAGAATCTGTATATAAACTATTTGT
HK6IEBF01AXRBZ AAAA-CCC-TAA-GTATCAAGG-TTTTT-ACAAACCTCTACATT-AGGTAAGT-AACCGAAAGGTTCTATACTAGAATCTGTATATAAACTATTTGT
HK6IEBF01AG8I1 AAAA-CCC-TAA-GTATCAAGG-TTTTT-ACAAACCTCTACATT-AGGTAAGT-AACCGAAAGGTTCTATACTAGAATCTGTATATAAACTATTTGT
HK6IEBF01BWH8V TTT--ACAAACCTCTACATT-AGGTAAGT-AACCGAAAGGTTCTATACTAGAATCTGTATATAAACTATTTGT
HK6IEBF01AV2R6 CTAGAATCTGTATATAAACTATTTGT

Exit
New
Open
Start
Stop
About
Help

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

Analysis messages

Variant Calling & Annotation

Mismatches are *potential* variants

```
      ←GTTGAGGCTTGCCTTTTGGTACGCTGGACTTTGT
GTACTCGTCGCTGCGTTGAGGCTTGCCTTTTGGT→
      ATGGTACGCTGGACTTTCTAGGATACCCTCGCTTT→
      TTGCGTTTATGGTACGCTGGACTTTGTAGGATACC→
      CTTGCGTTTATGGTACGCTGGACTTTGTAGGATAC→
      TTGCGTTTATGGTACGCTGGGCTTTGTAGGATACC→
      GCGTTTATGGTACGCTGGACTTTGTAGGATACCCT→
      GAGGCTTGCCTTTATGGTACGCTGGACTTTGTAGG→
←GCGTTGTGGCTTGCCTTTATGGTACGCTGGATTTT
      CGTTTATGGTACGCTGGACTTTGTAGGATACCCTC→
      ATGGTACGCTGGACTTTGTAGGATACCCTCGCTTT→
      ←GTTTTTGGTACGCTGGACTTTGTAGGATACCCTCG
TCTCGTGCTCGTCGCTGCGTTGAGGCTTGCCTTTA→
←TGACTGTCGCTGCGTTGAGGCTTGCCTTTATGGTA
←GCTCGTCGCTGCGTTGAGGCTTGCCTTTATGGTAC
      TATGGTACGCTGGACTTTGTAGGATAGCCTCGCTT→
TCGTGCTCGTCGCTGCGTTGAGGCTTGCCTTTTGG→
←CGTCGCTGCGTTGAGGCTTGCCTTTATGGTACGCT
←GTTGAGGCTTGCCTTTATGGTACGCTGGGCTTTTT
←TTGCGTTTATGGTACGCTGGACTTTGTAGGATACC
```

Ref: CTCTCGTGCTCGTCGCTGCGTTGAGGCTTGCCTTTATGGTACGCTGGACTTTGTAGGATACCCTCGCTTTC

Formát dat

- SAM = Sequence Alignment Map
- BAM = Binary SAM = compressed SAM
- contains information about how sequence reads map to a reference genome
- Supports paired-end reads
- Is produced by bowtie, BWA and other mapping tools (software)

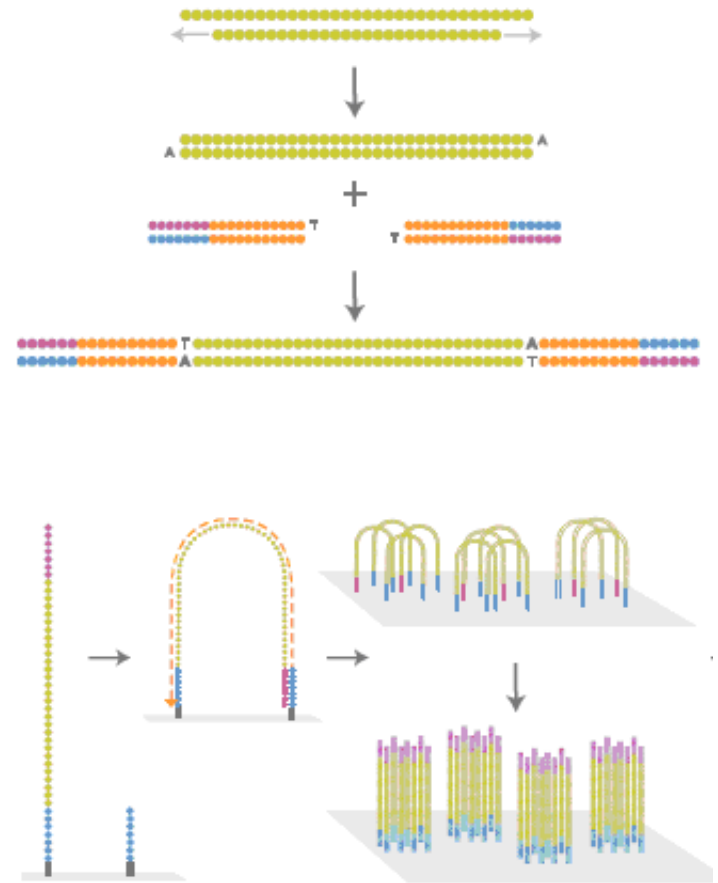
Technologie firmy Illumina/Solexa

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

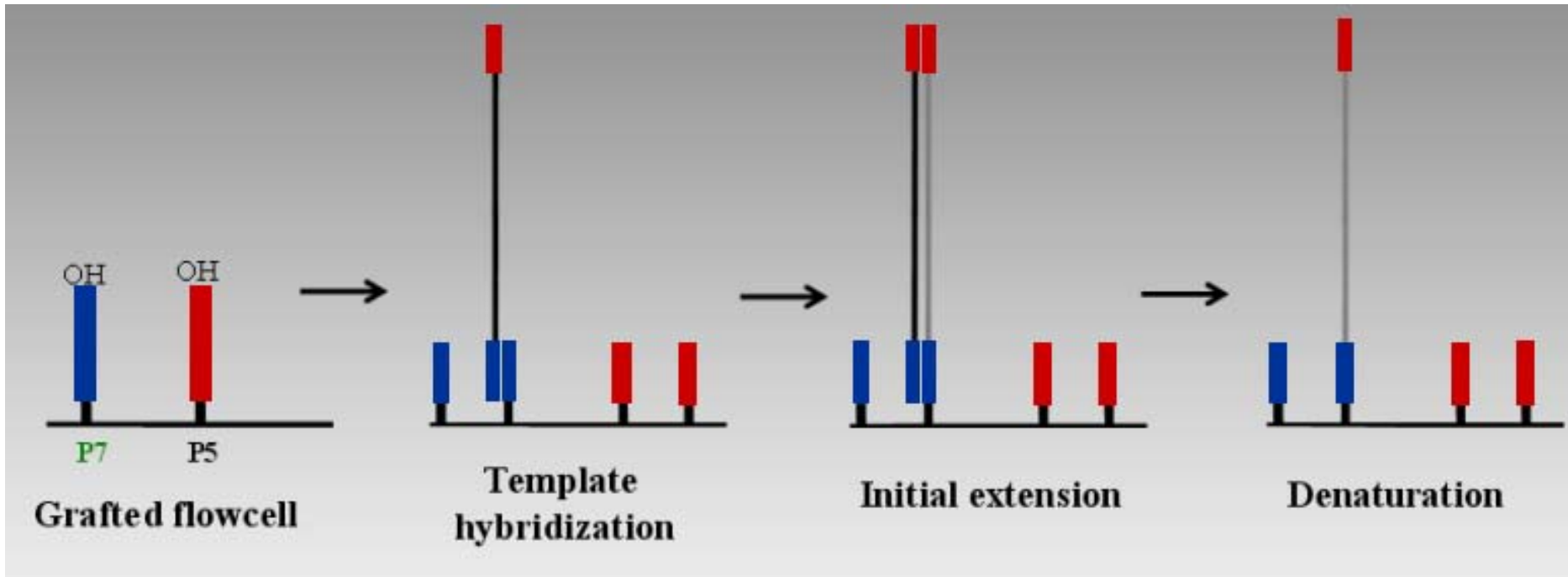
Příprava knihovny – polony

PCR

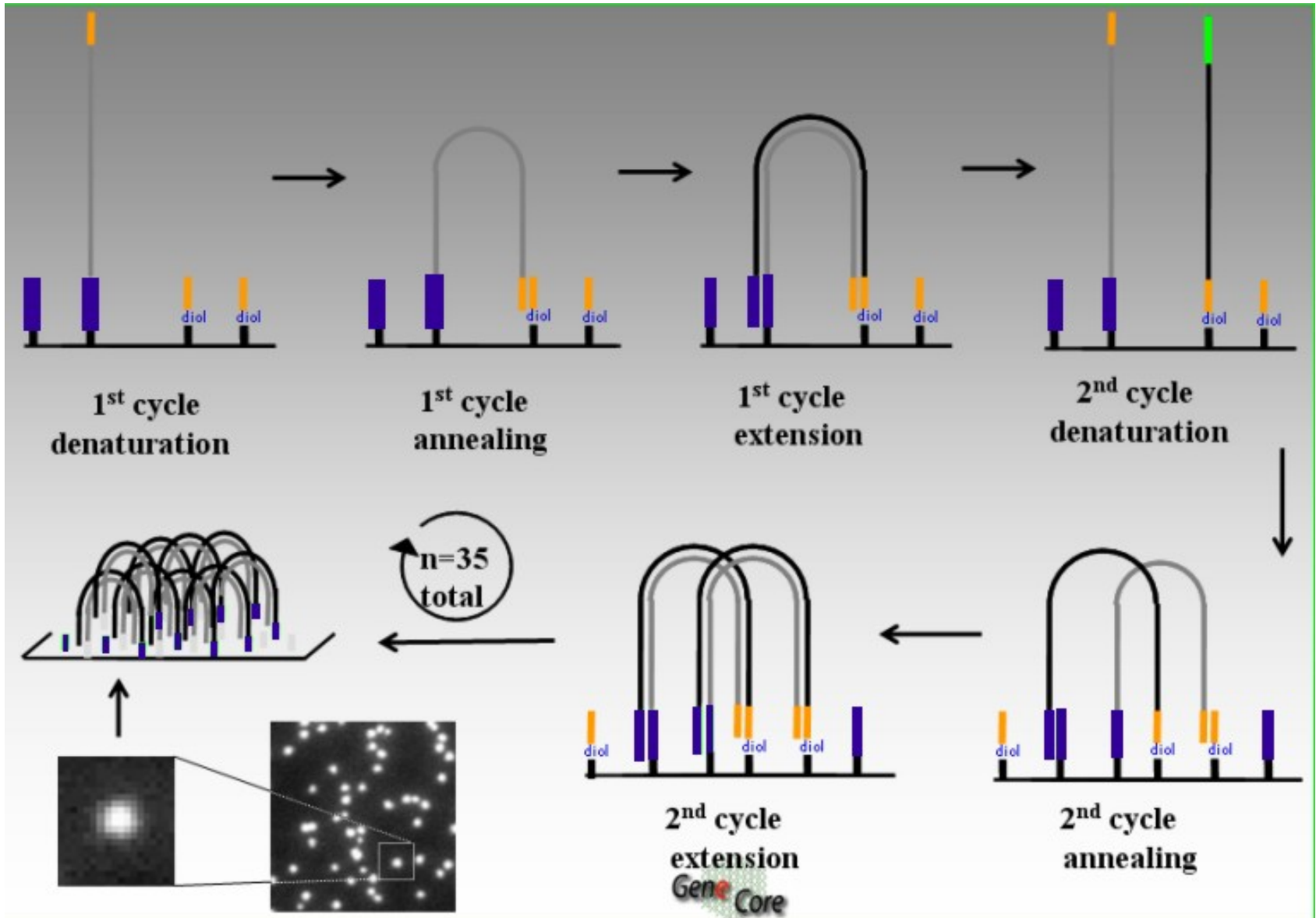
- Náhodná fragmentace DNA na krátké úseky 200 – 500 bp
- Zatupení konců a vytvoření 1 nt A-přesahu na 3'-konci (T4 DNA polymeráza, Klenow a polynukleotid-kináza T4)
- Navázání adaptorů umožňujících kovalentní vazbu na opticky transparentní povrch pro sekvenování
- Každý fragment je na povrchu uchycen 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 μm^2 povrchu
- Na celém povrchu je dosaženo hustoty deseti milionů shluků na cm^2



Bridge amplification:
initiation

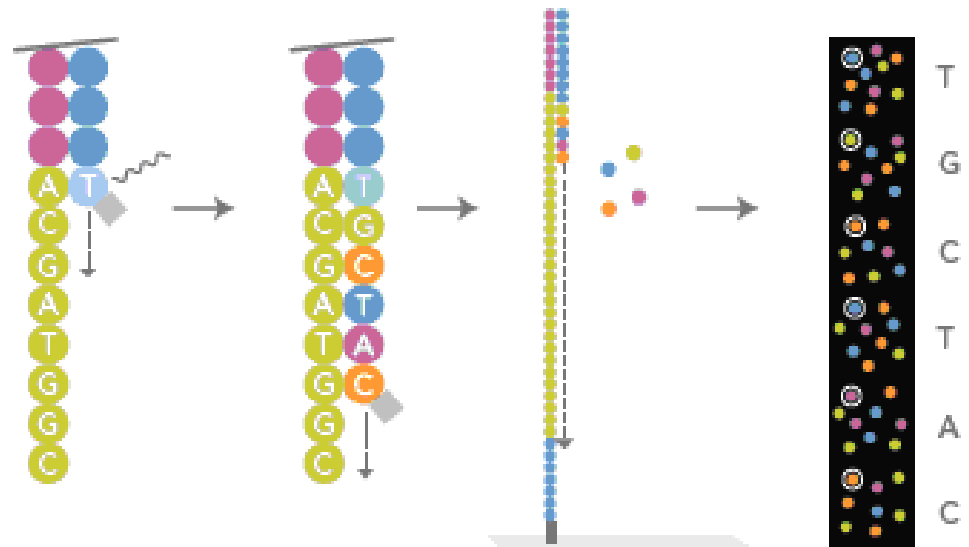


On the surface: complementary oligos



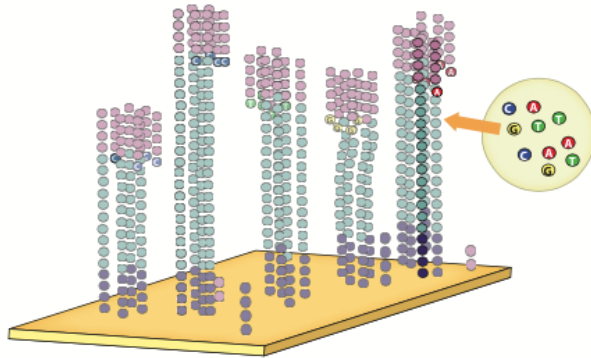
Illumina

- **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



Sequencing (Illumina technology)

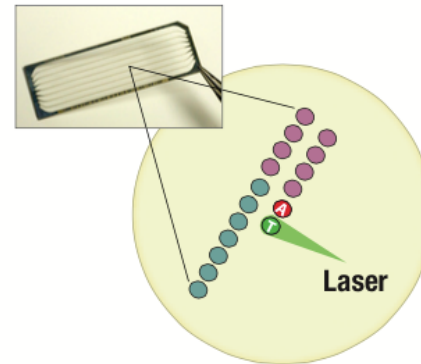
9 DNA SEQUENCING



To initiate the first sequencing cycle and determine the first base, all four labeled reversible terminators and DNA polymerase enzyme are first added. Only one base can incorporate at a time.

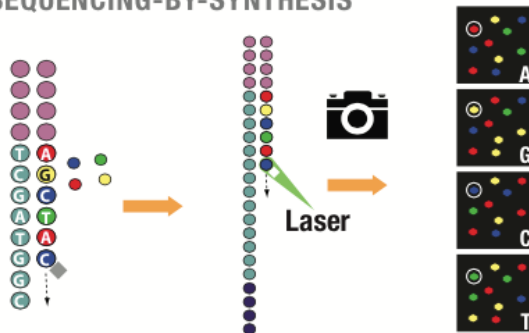


10 BASE CALLING



Lasers excite the fluorescent tags and the images are captured via CCD camera. The identity of the first base in each cluster is recorded, and then the fluorescent tag is removed.

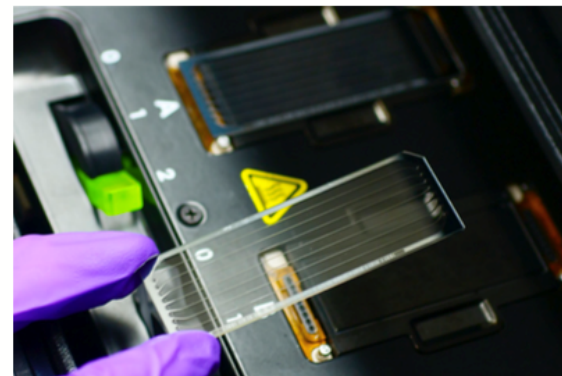
11 SEQUENCING-BY-SYNTHESIS



In the first cycle, the first base is incorporated. Its identity is determined by the signal given off and then recorded. In subsequent cycles, the process of adding sequencing reagents, removing unincorporated bases and capturing the signal of the next base to identify is repeated.



12 DUAL FLOW CELLS



Once the top surface of the flow cell channel has been scanned, the imaging step is repeated on the bottom surface.

Illumina

- Délka čtených úseků:
 - HiSeq : 100 nt or 150 nt
 - MiSeq : 250 nt
- Mnohonásobné pokrytí sekvence (30 x – 100 x)
- Možnost mnohonásobného sekvenování
 - Povrch sklíčka rozdělen do částí
 - Vzorky označeny pomocí **dvanácti** různých oligonukleotidů
 - Současně lze analyzovat až 96 různých vzorků.
- Produkuje desítky až stovky Gb za 1 běh (2 - 8 dnů)
- Aplikace
 - Resekvenování
 - Sekvenování RNA
 - Stanovení metylací

MiSeq



Sekvenování třetí generace

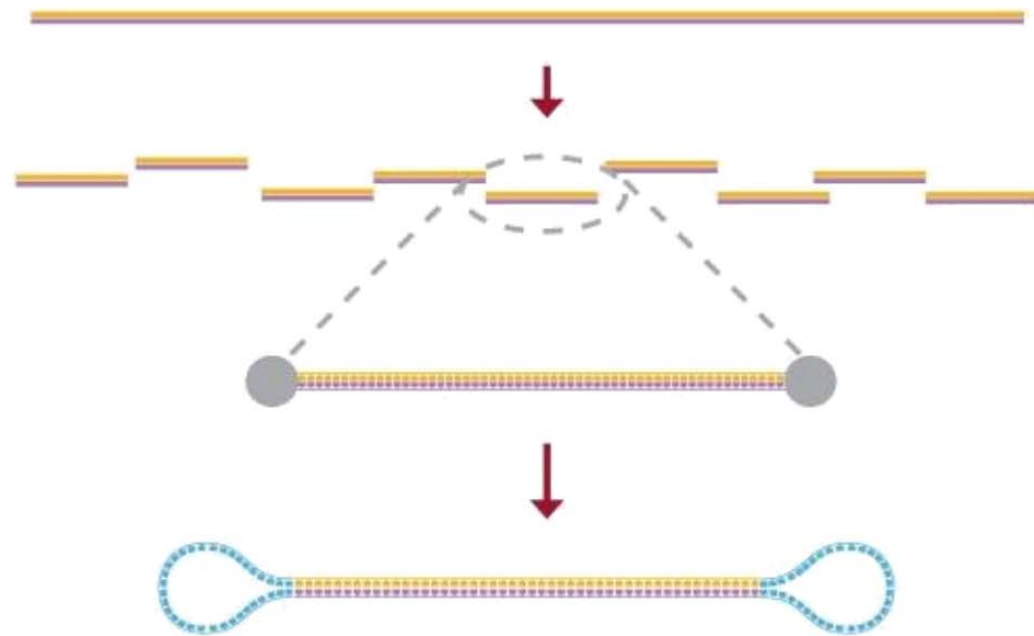
- Analýza jedné molekuly
- Potřeba malého množství vzorku
- Bez rizika kontaminace
- Přesné čtení homopolymerních úseků
- Analýza jakékoli NK (DNA/RNA)
- Analýza poškozených NK (archaické, muzejní, forenzní)
- Analýza DNA z nekultivovatelných organismů
- Absolutní kvantifikace

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

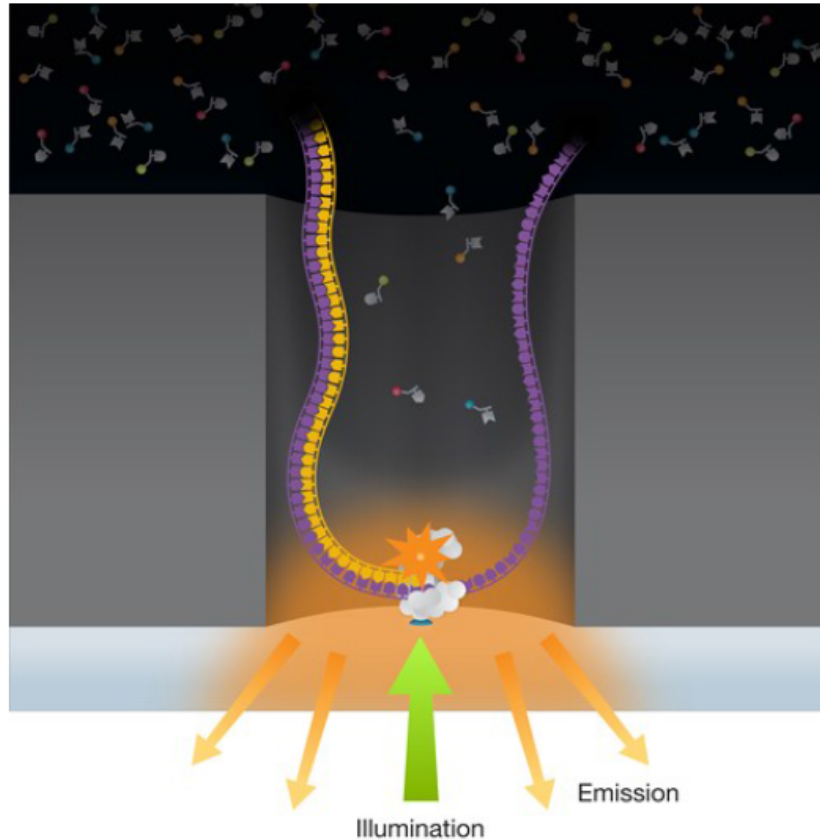
Template Preparation



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

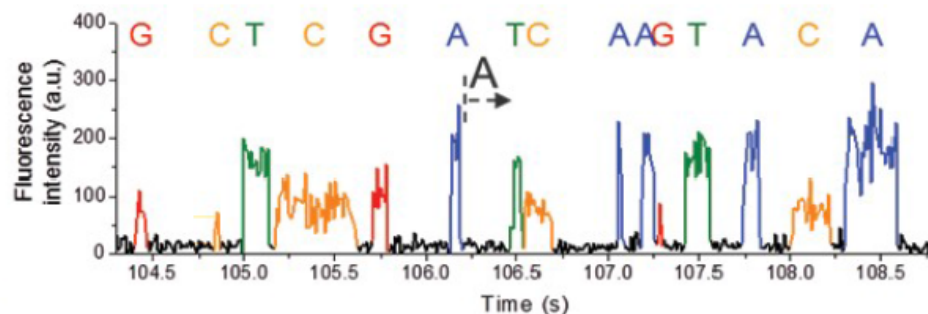
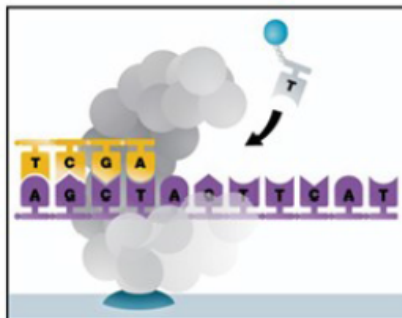
Pacific Biosciences



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



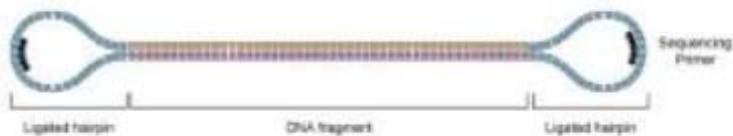
PacBio

High-throughput sequencing

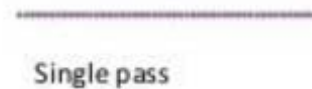


Library preparation

SMRTBell 'template'

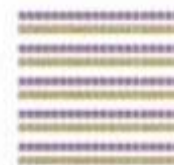


Standard 'Sequencing'



&

Circular 'Consensus' Sequencing'



Continued generations of reads

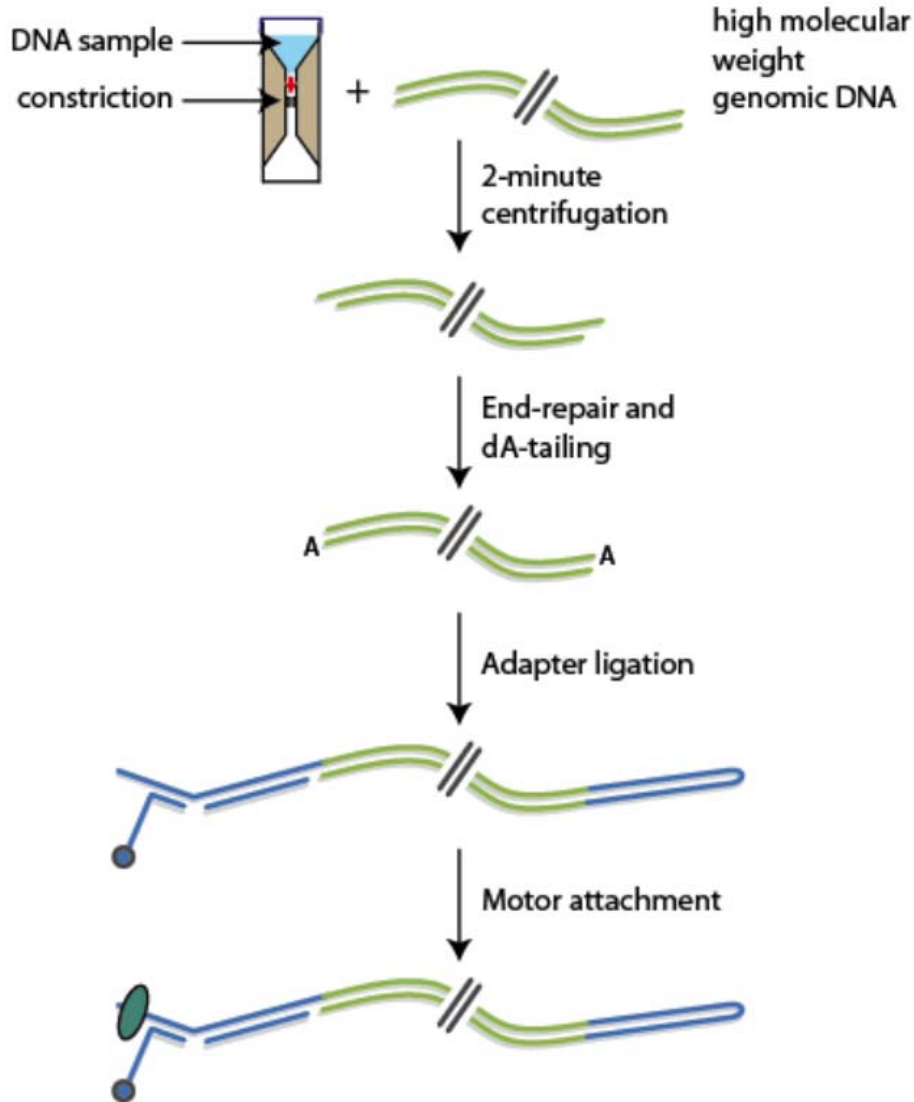
Multiple passes

&

NanoPore (1995)

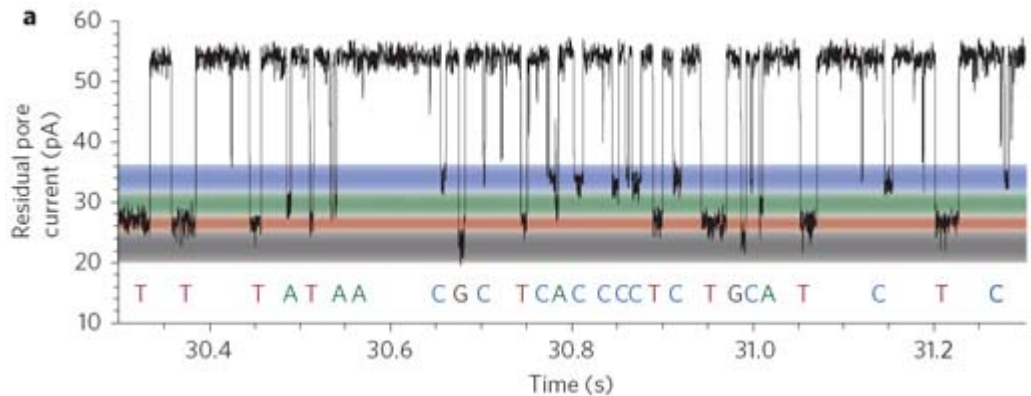
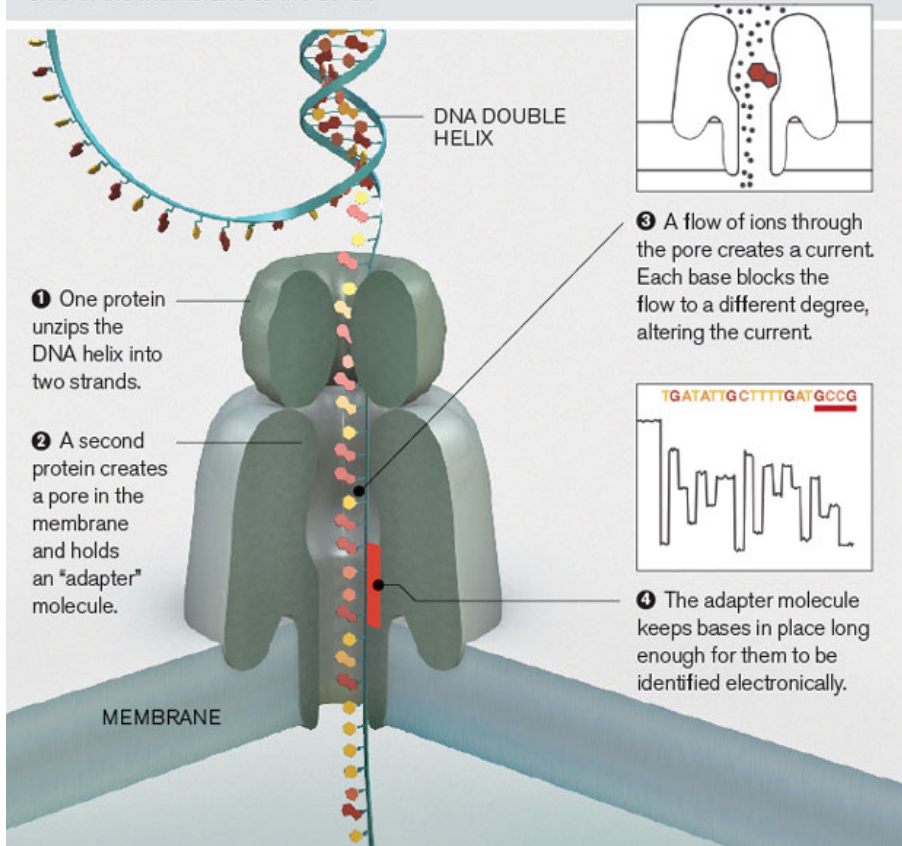
- **Využívá měření elektrického potenciálu přes membránu**
- Elektricky odolná polymerní membrána (synтетická lipidová dvouvrstva) s transmembránovým porózním proteinem
- Modifikovaný α hemolysin (α HL) nebo porin A (MspA) *Mycobacterium smegmatis*
- Nanopore je ponořený ve vodivém roztoku
- Průchod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

G-tube sense/antisense



Příprava knihovny MinION

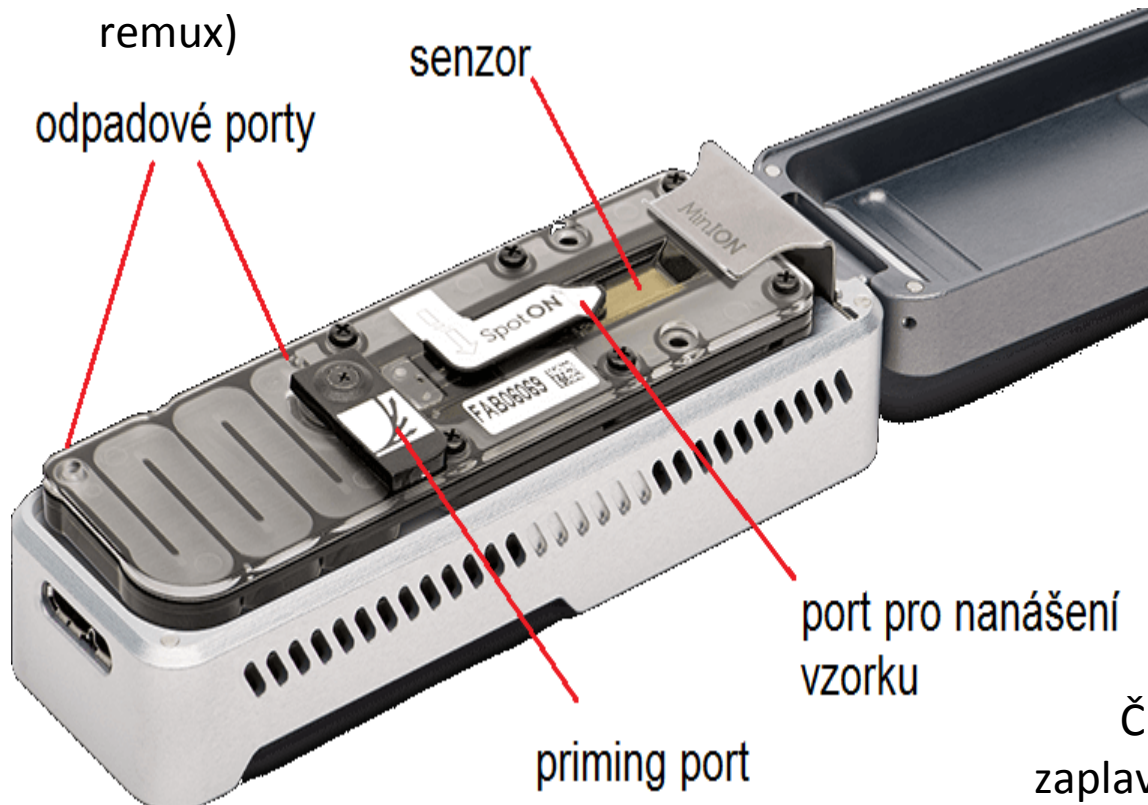
DNA can be sequenced by threading it through a microscopic pore in a membrane. Bases are identified by the way they affect ions flowing through the pore from one side of the membrane to the other.



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 formát dat

The screenshot shows the hdfview application interface. On the left is a tree view of the data structure. The main area is divided into three windows:

- Text View:** Displays FASTQ data for a specific read. The text includes read identifiers, quality scores, and sequence data.
- Table View - Trace:** Shows a table of signal values across different channels (0-7) for each read index (0-10).
- Table View - Signal:** Shows a table of signal values for a specific read (0) across different channels (0-9).
- Lineplot:** A plot of the signal values for the selected read, showing a noisy waveform over time.

Table View - Trace Data:

	0	1	2	3	4	5	6	7
0	98	105	39	12	0	0	0	1
1	99	102	43	9	0	1	0	0
2	100	101	43	9	0	1	0	0
3	101	100	43	9	0	1	0	0
4	102	100	43	9	0	1	0	1
5	103	98	43	9	0	1	0	1
6	107	94	43	9	0	1	0	1
7	170	0	79	2	2	0	0	1
8	171	0	33	1	2	0	47	1
9	183	0	21	1	2	0	47	1
10	184	0	21	1	1	0	47	1

Table View - Signal Data:

	0
0	769
1	462
2	447
3	437
4	446
5	433
6	454
7	458
8	485
9	370

Lineplot Data:

Index	Value
0	769
1	716
2	663
3	610
4	558
5	505
6	452
7	400
8	347
9	294
10	242

Srovnání sekvenačních metod

Srovnání metod pro sekvenování

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

Srovnání metod pro sekvenování

	Metoda	Princip reakce
Klasické	Maxam-Gilbert (chemická)	Degradace chemickými činidly
	Sanger (enzymová)	Syntéza DNA polymerázou s terminátory
NGS	454 (pyrosekvenování)	Syntéza DNA polymerázou, detekce PP
	Polovodičové (IonTorrent)	Syntéza DNA polymerázou, detekce H ⁺
	Illumina	Syntéza DNA polymerázou s reverzibilními terminátory
	Solid	Série ligací sond + posun rámce
Třetí generace	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
Klasické	Maxam-Gilbert (chemická)	150 – 300 nt
	Sanger (enzymová)	600 – 1000 nt
NGS	454 (pyrosekvenování)	500 nt
	Polovodičové (IonTorrent)	400 nt
	Illumina	200 nt
	Solid	30 – 40 nt
Třetí generace	PacBio	10 – 15 kb
	Nanopore	> 10 kb
	Helicos	200 nt

Srovnání metod pro sekvenování

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