

Sekvenace DNA

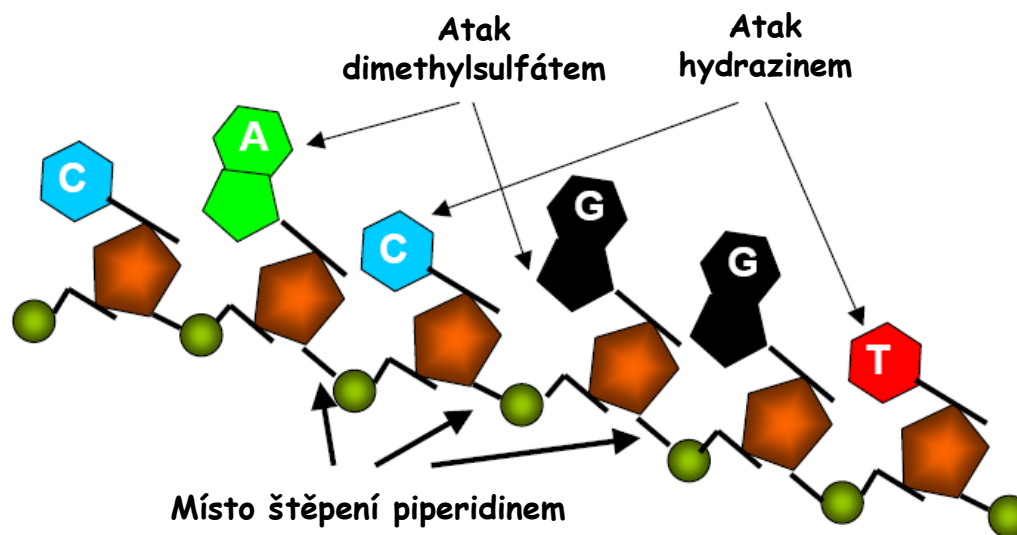
Vývoj sekvenačních
technik za posledních
40 let



Maxam-Gilbertova sekvenační metoda (1977)

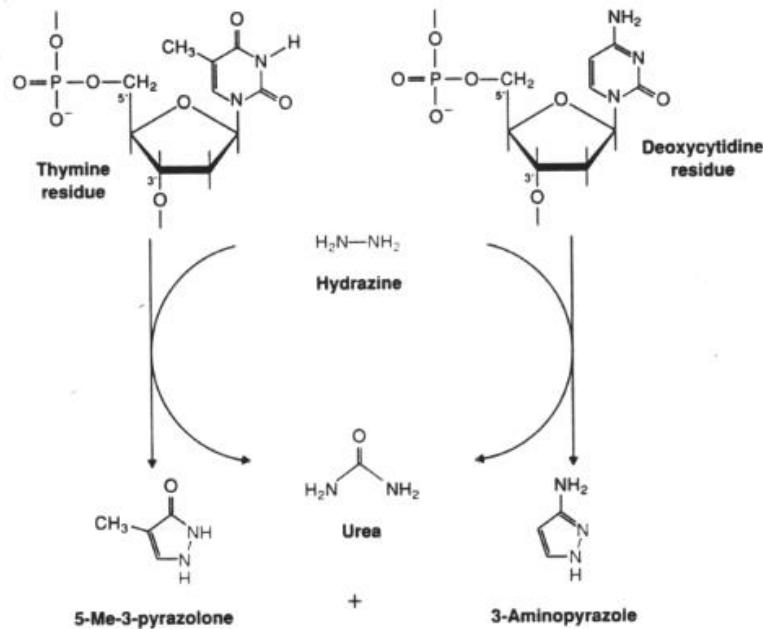
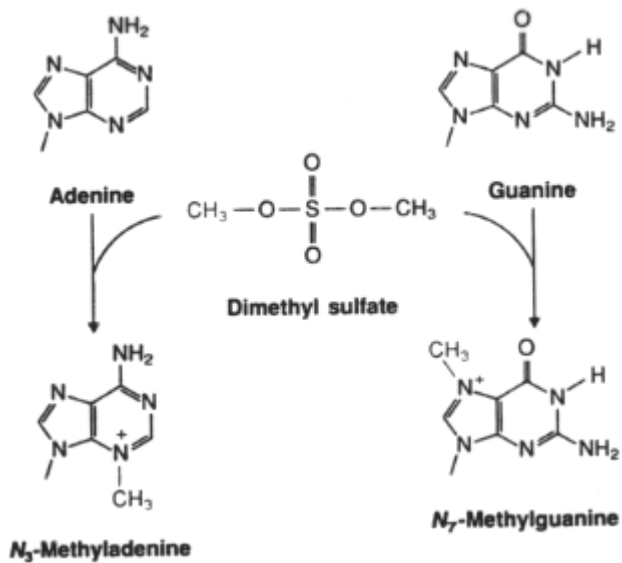
Princip metody - specifická chemická degradace purinových a pyrimidinových bazí

- puriny jsou modifikovány pomocí dimethylsulfátu
- pyrimidiny pomocí hydrazinu
- následně je pomocí 1M piperidinu při 90°C štěpena cukr-fosfátová kostra v místě příslušné modifikované báze



Maxam-Gilbertova sekvenační metoda

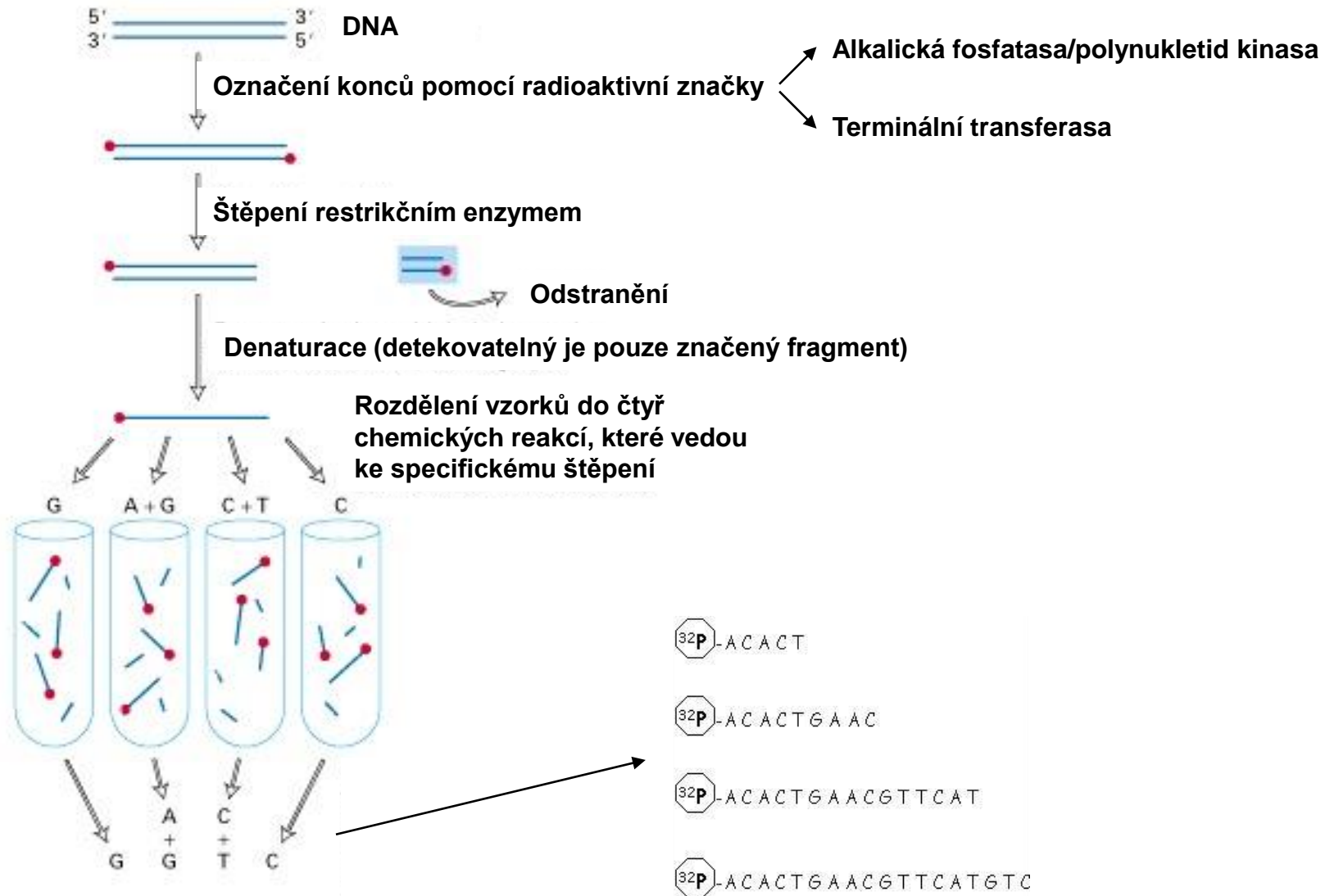
Modifikovaná báze	Specifická modifikace
G	Methylace guaninu pomocí dimethylsulfátu při pH 8.0 - guanin se stává náchylný k odstranění při alkalickém pH.
A + G	Přidání piperidinu v kyselině mravenčí při pH 2.0 vede k odstranění purinovýchází
C + T	Přidání hydrazinu vede k otevření pyrimidinového kruhu a jeho odstranění z DNA
C	Při vysoké iontové síle (1.5 M NaCl) reakce s hydrazinem ionom cytosin





Maxam-Gilbertova sekvenační metoda

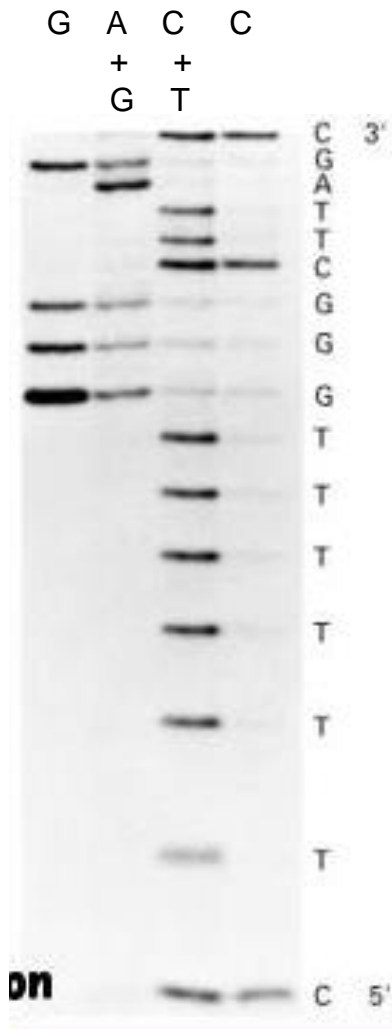
Schéma sekvenace:





Maxam-Gilbertova sekvenační metoda

- Fragmenty jsou separovány na přibližně 6% polyakrylamidovém gelu

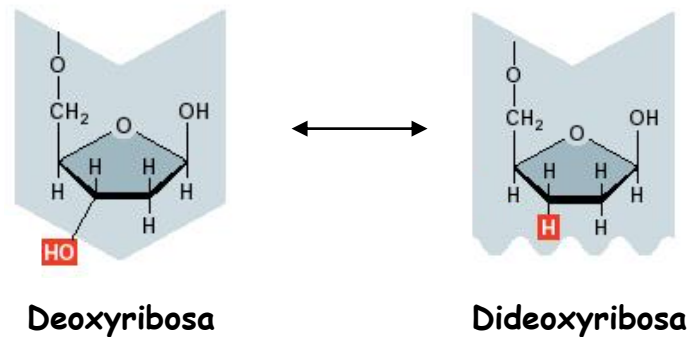


- Touto metodou jsme schopni sekvenovat přibližně 250-300 bp dlouhé fragmenty
- Musíme pracovat s velkým množstvím DNA
- Velká pracnost metody (několik purifikačních kroků), nemožnost plné automatizace, práce s mutagenními chemikáliemi
- Používá se stále pro "footprinting"

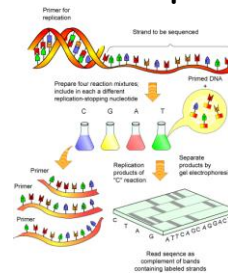


Sekvenační metoda dle Sanger (1980)

- Syntéza DNA in-vitro za použití "terminátorů" - dideoxynukleotidů zabraňujících po svém začlenění do DNA její další elongaci.



- Vyžaduje použití iniciálního primeru, DNA polymerasy a směs dNTPs se značenými ddNTPs
- Nasyntetizované řetězce jsou poté separovány pomocí polyakrylamidové gelové elektroforézy nebo kapilární elektroforézy
- Možnost plně automatizované seprace za použití fluorescenčně značených ddNTPs

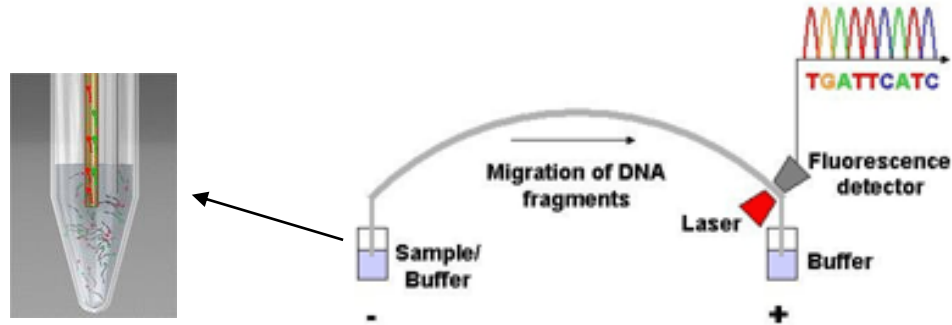




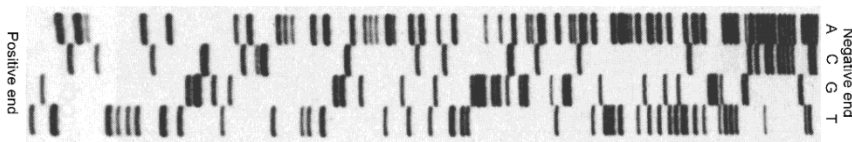
Sekvenační metoda dle Sanger (1980)

Automatická analýza a vyhodnocení

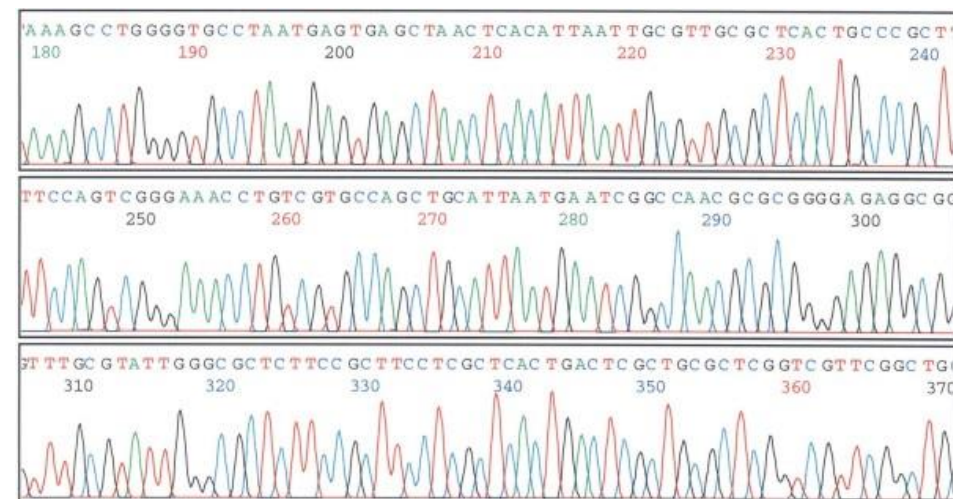
- kapilární elektroforéza spojená s fluorescenční detekcí produktů (BigDye barevný set)
- Genetic analyzer 3000 series (ABI)
- Megabase (GE Healthcare)



Ruční sekvenace



Automatická sekvenace





Sekvenační metoda dle Sangera



Throughput/Performance by Run Module

XLRseq: 768 samples per day (690 Kbases)

LongSeq: 1152 samples/day (980 Kbases)

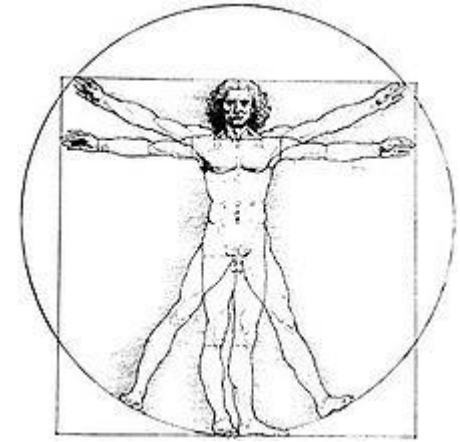
StdSeq: 2304 samples/day (1550 Kbases)

FastSeq: 2304 samples/day (1600 Kbases)

RapidSeq: 3840 samples per day (2100 Kbases)



Human Genome Project (HGP)



- započat v roce 1990 za účasti DOE and [NIH](#)
- sekvenace prováděna pomocí BACs
- prvotní plán počítal s dobou trvání 15 let
- nakonec sekvenace téměř dokončena již v roce 2000
- výsledná sekvenční mapa publikována 14. dubna 2003, 99.99% přesnost
([National Human Genome Research Institute](#))
- celkové náklady projektu 3 miliardy dolarů
- v roce 2000 prezident Bill Clinton ujistil o nepatentovatelnosti lidské DNA

[https://https://www.genome.gov/25019885/online-education-kit-how-to-sequence-a-human-genome/](https://www.genome.gov/25019885/online-education-kit-how-to-sequence-a-human-genome/)



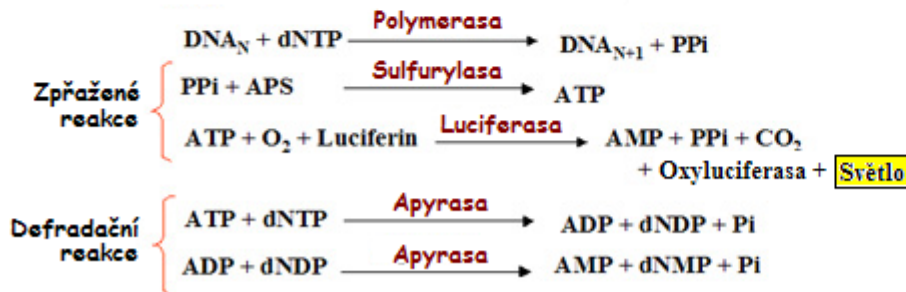
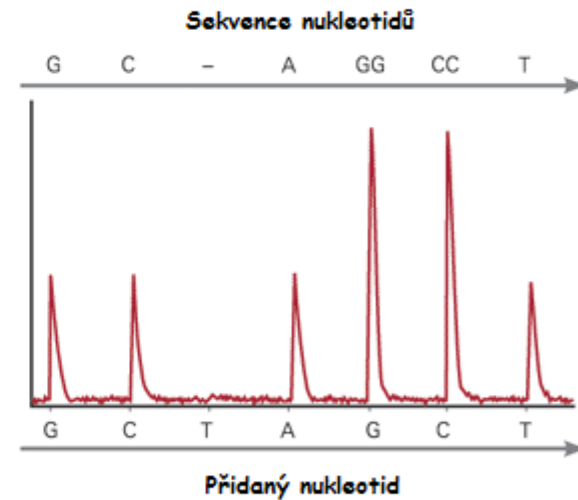
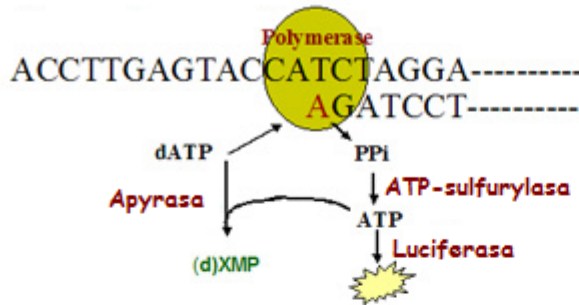
Celera Genomics Project

- založena vědcem Craig Venterem a v roce 1998 započala sekvenační projekt
- celkové náklady 300 mil. dolarů byly hrazeny plně s privátních zdrojů
- poprvé použita metoda „whole genome shotgun sequencing“
- k analýze sekvenačních dat použit přístup vyvinutý Gene Myersem
- tento přístup však vyžadoval extrémní výpočtové požadavky
- finální výpočet prováděn na 7000 procesorech k získání 1000 násobné rychlosti oproti Pentium počítačům
- tento inovativní přístup dovolil dokončit sekvenaci již za 9 měsíců



Pyrosekvenování (1990)

- umožňuje rychlou sekvenaci krátkých úseků DNA - sekvenace 30 až 50 bazí trvá přibližně 30 až 45 minut.
- Jedná se o bio-luminometrické sekvenování DNA založené na detekci anorganického pyrofosfátu (PPi) uvolněného během inkorporace nukleotidů.





Pyrosekvenování

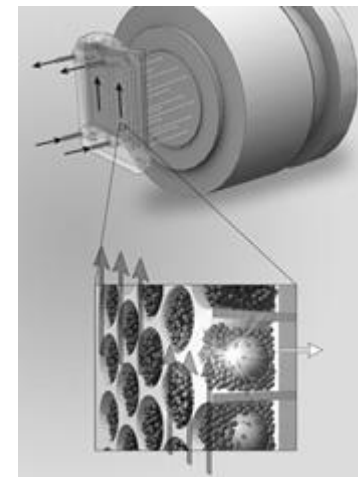
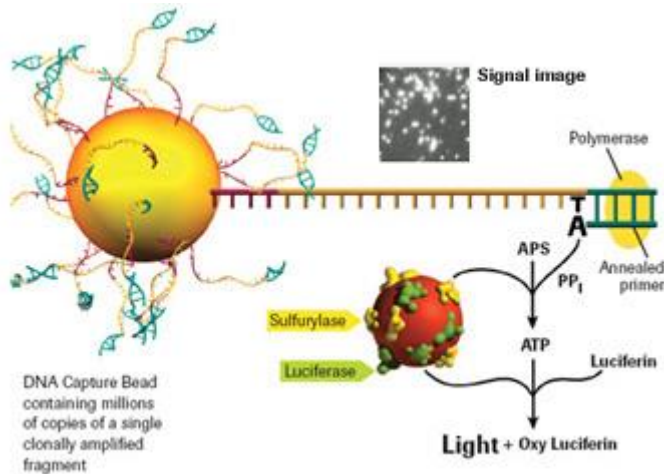
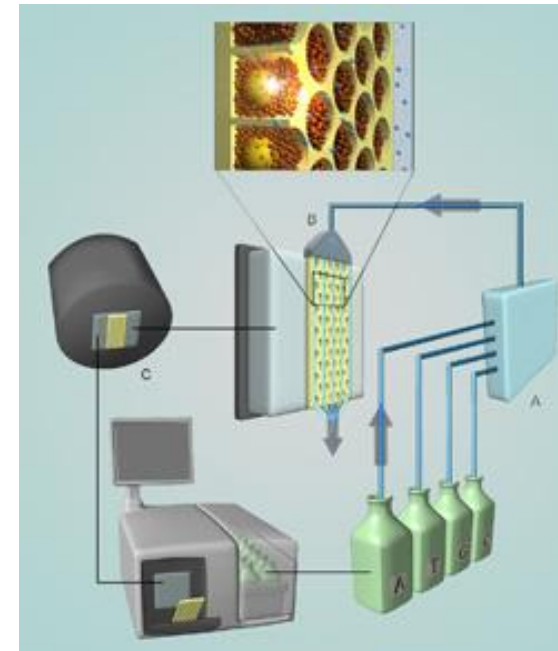
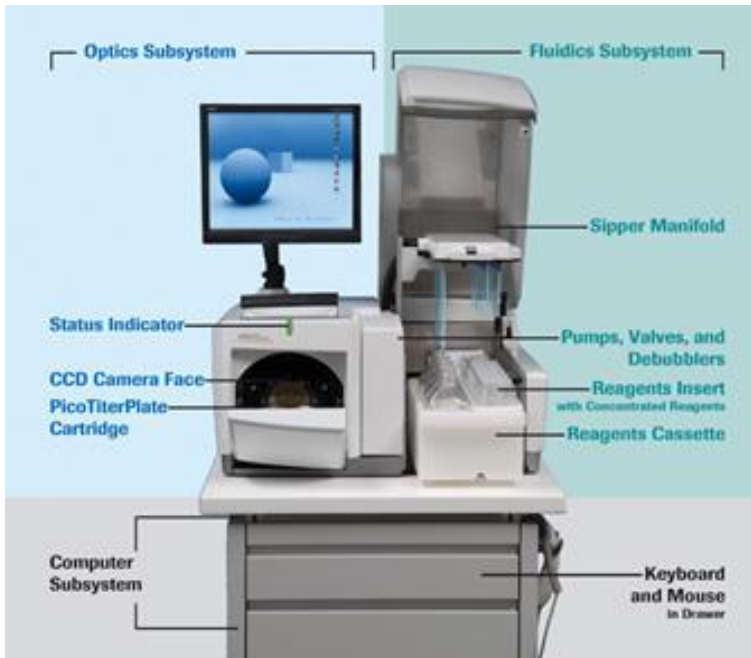
Metoda pyrosekvenování je použita v některých sekvenátorech "druhé generace" (Roche)

Průchodnost	1 miliarda bazí za den
Doba analýza	10.0 hodin
Délka čtení	400
Počet čtení/analýzu	1 000.000
Správnost	>99.0% správnost jednoho čtení na 400 bazích
Potřebné množství DNA	Méně než 100 ng DNA
Multiplexování	Až 192 vzorků/běh





Pyrosekvenování



GS Junior System



Průchodnost	35 bází/běh
Doba analýza	10 hodin sekvenování 2 hodiny zpracování dat
Délka čtení	400 bází
Přesnost	99% přesnost při 400 bází
Počet čtená/běh	100,000 shotgun, 70,000 amplikon
Vzorky	gDNA, amplikony, cDNA



Sekvenátor II generace - Solexa



Tvorba klastrů



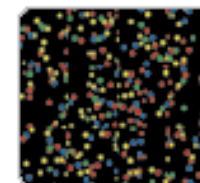
Sekvenace



Párování



Analýza dat



Read Length	Run Time	# of Reads (per flow cell)	High-Quality Output (GB)		Base Calls with Q ≥30	Raw Read Accuracy	Perfect Reads
			Total	Per Day			
1 × 35 bp	~2 days	90–100 M	3–3.5	~ 1.5–1.75	75–90%	≥ 99%	≥ 90%
2 × 35 bp	~4 days	180–200 M	6.5–7	~ 1.6–1.75	75–90%	≥ 99%	≥ 90%
2 × 50 bp	~5 days	180–200 M	9–10	~ 1.8–2.0	75–90%	≥ 99%	≥ 85%
2 × 75 bp	~7.5 days	180–200 M	13–15	~ 1.7–2.0	70–85%	≥ 98.5%	≥ 80%
2 × 100 bp	~9.5 days	180–200 M	18–20	~ 1.9–2.1	≥ 70%	≥ 98%	≥ 70%

Sekvenátor II generace -MiniSeq, MiSeq, NextSeq, HiSeq, NovaSeq

SEQUENCING LIKE NO OTHER

Users can run 1 or 2 flow cells at a time, using any combination of the available read length and flow cell type.

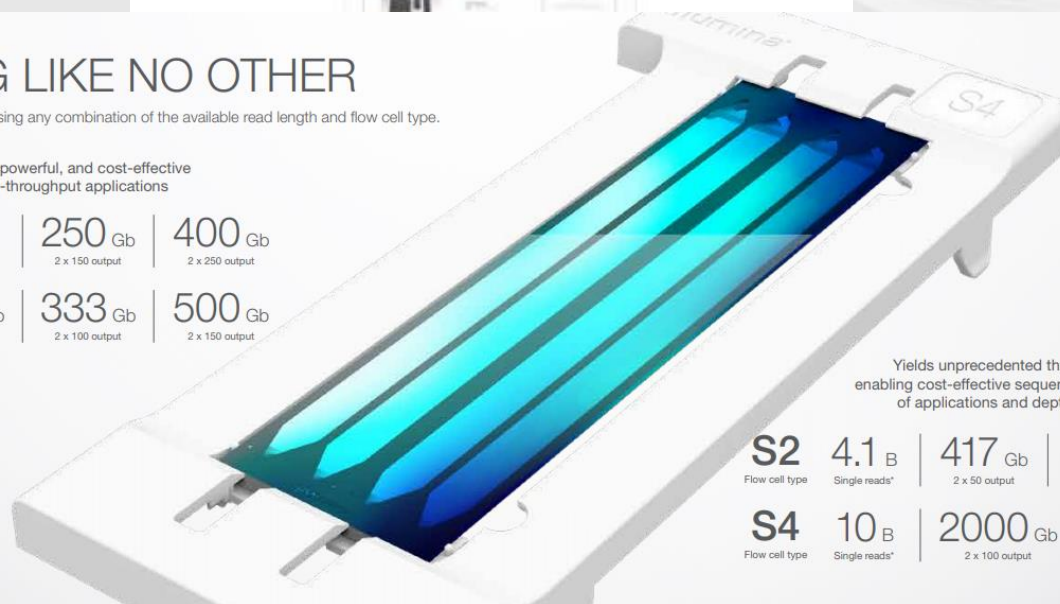
Provides a quick, powerful, and cost-effective option for high-throughput applications

SP Flow cell type	0.8 _B Single reads*	80 Gb 2 x 50 output	250 Gb 2 x 150 output	400 Gb 2 x 250 output
S1 Flow cell type	1.6 _B Single reads*	167 Gb 2 x 50 output	333 Gb 2 x 100 output	500 Gb 2 x 150 output

*Clusters passing filter

Yields unprecedented throughput while enabling cost-effective sequencing across a range of applications and depth of coverage

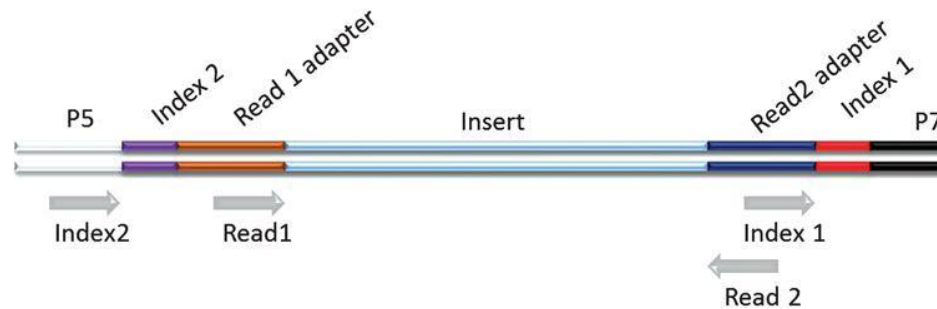
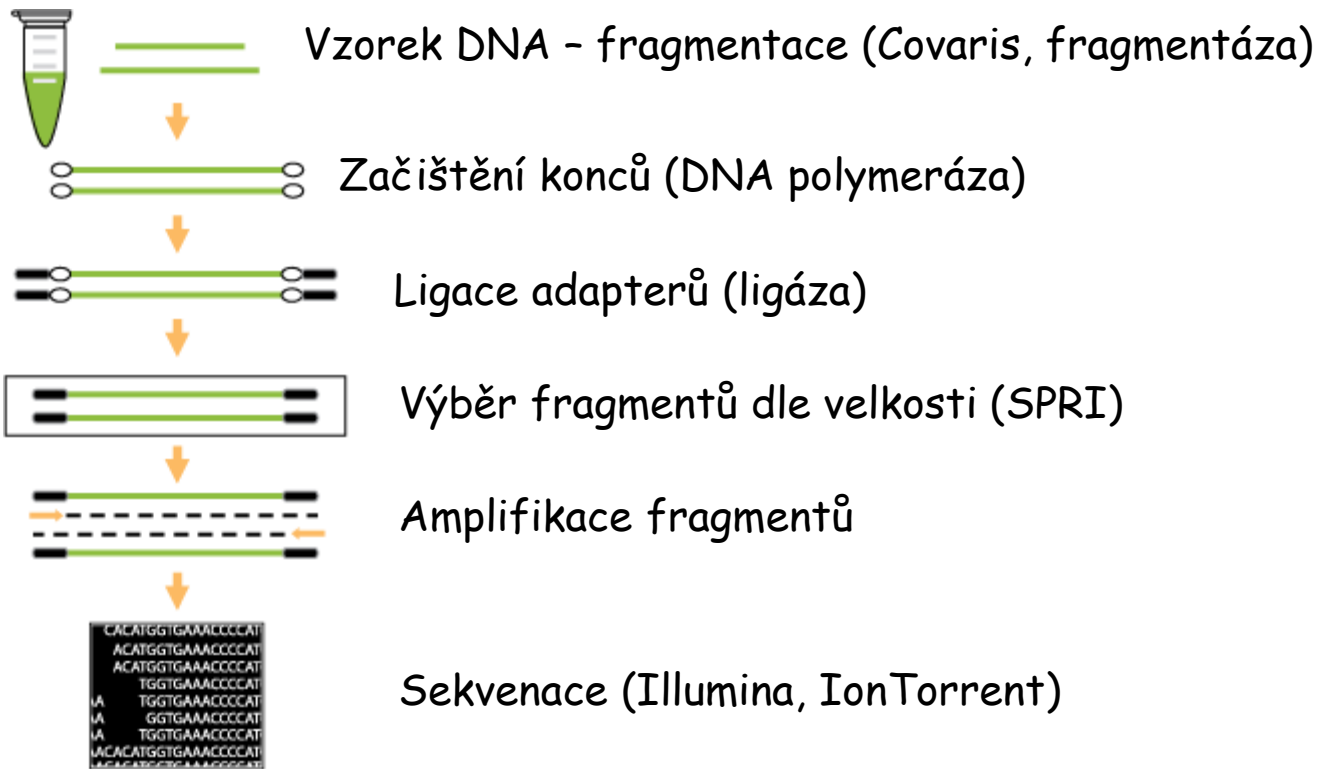
S2 Flow cell type	4.1 _B Single reads*	417 Gb 2 x 50 output	833 Gb 2 x 100 output	1250 Gb 2 x 150 output
S4 Flow cell type	10 _B Single reads*	2000 Gb 2 x 100 output	3000 Gb 2 x 150 output	





Sekvenátor II generace

Příprava DNA knihovny

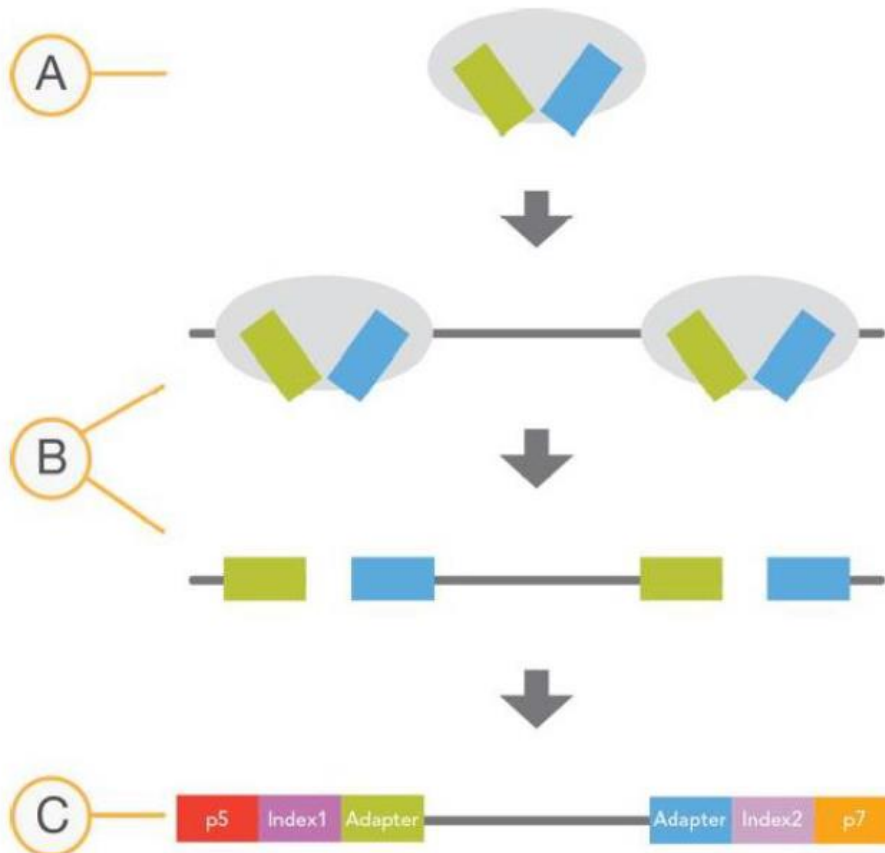




Sekvenátor II generace

Library Preparation

Illumina Nextera DNA Sample Preparation Kit



Separate fragmentation not required

Tag with enzyme mix

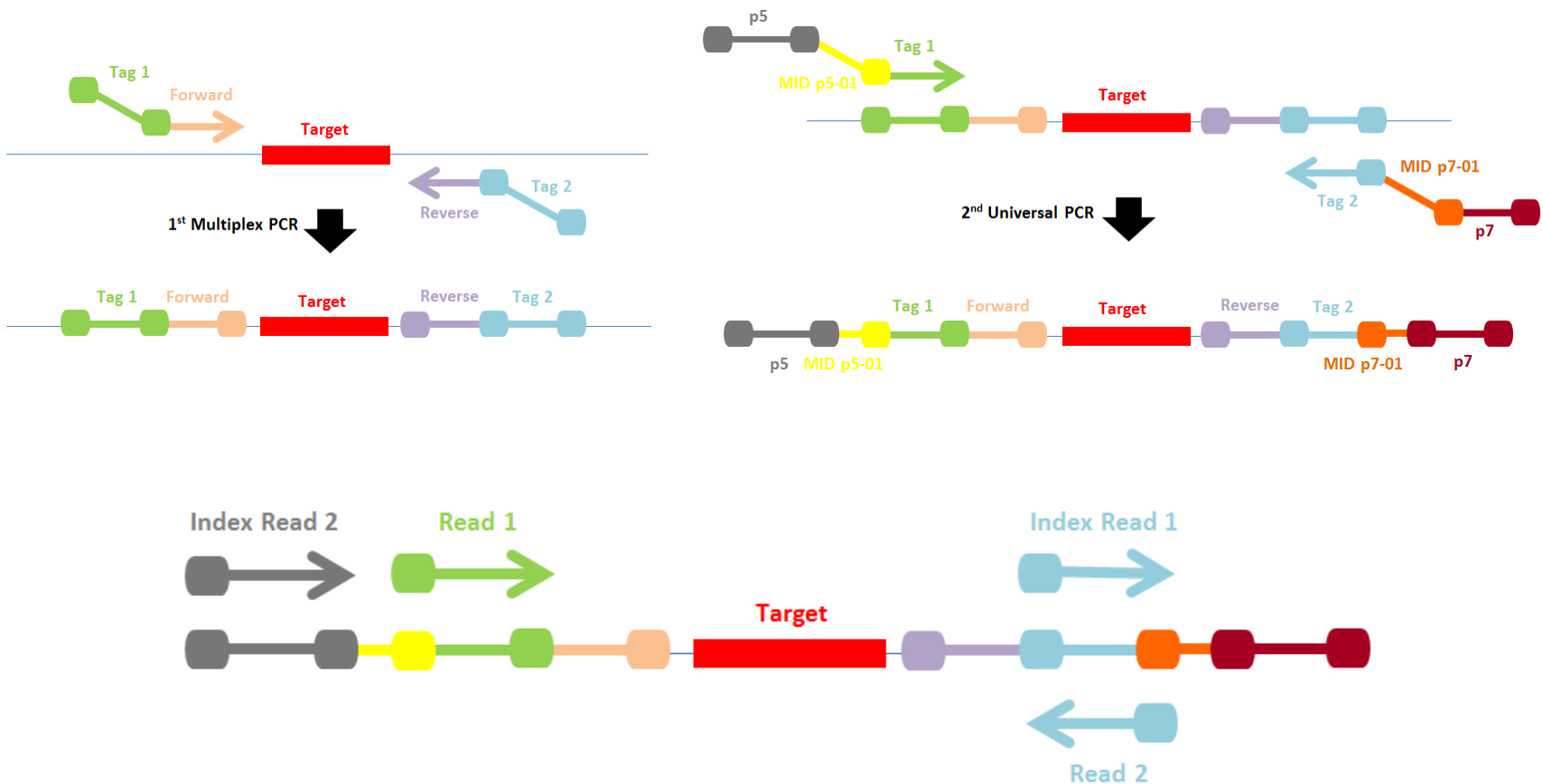
PCR
Polishes fragment ends and incorporates optional indices

- A Nextera Transposome with Adaptors
- B Tagmentation to Fragment and Add Adaptors
- C Limited Cycle PCR to Add Sequencing Primer Sequences and Indices



Sekvenátor II generace

Sekvenace amplikonů



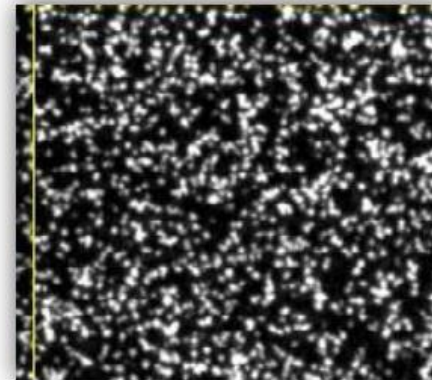


Sekvenátor II generace

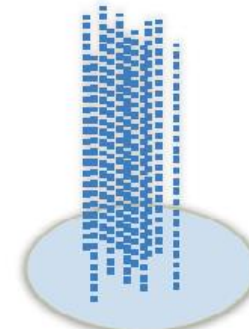
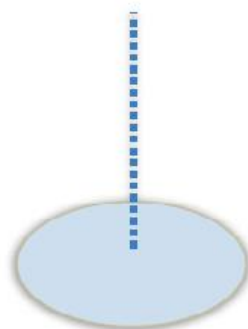
Sekvenace probíhá v klastrech

Clusters are bright spots on an image

Each cluster represents thousands of copies of the same DNA strand in a 1–2 micron spot



Single
DNA
Library



Amplified
Clonal
Cluster



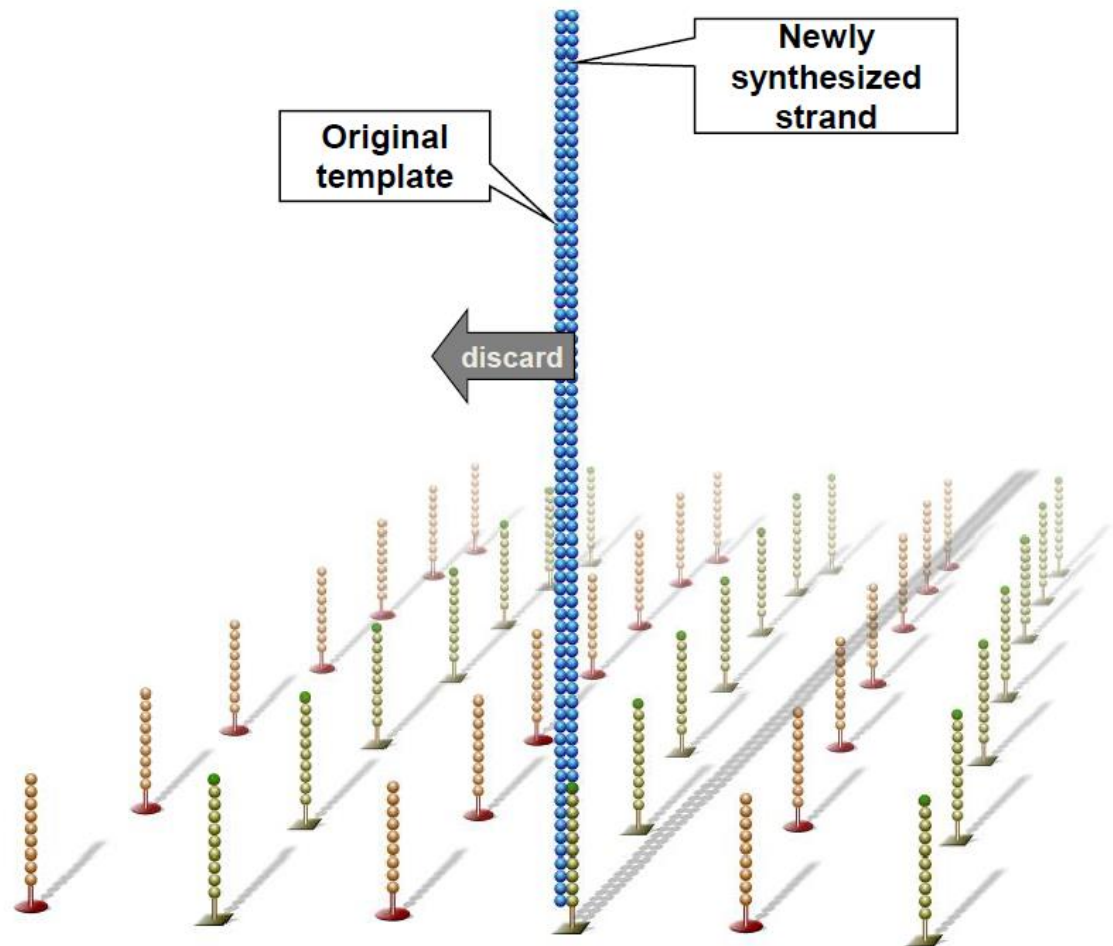
Sekvenátor II generace

1. Krok - hybridizace templátu na průtočnou celu

Double-stranded molecule is denatured

Original template washed away

Newly synthesized strand is covalently attached to flow cell surface



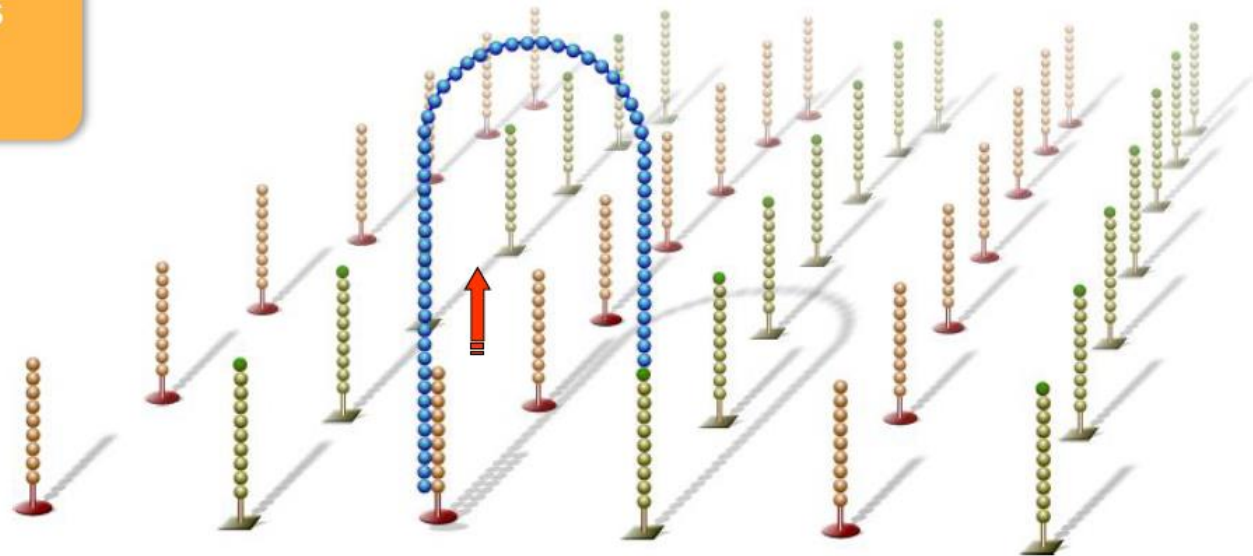


Sekvenátor II generace

2. Krok - Můstková PCR

Single-stranded molecule flips over and forms a bridge by hybridizing to adjacent, complementary primer

Hybridized primer is extended by polymerases

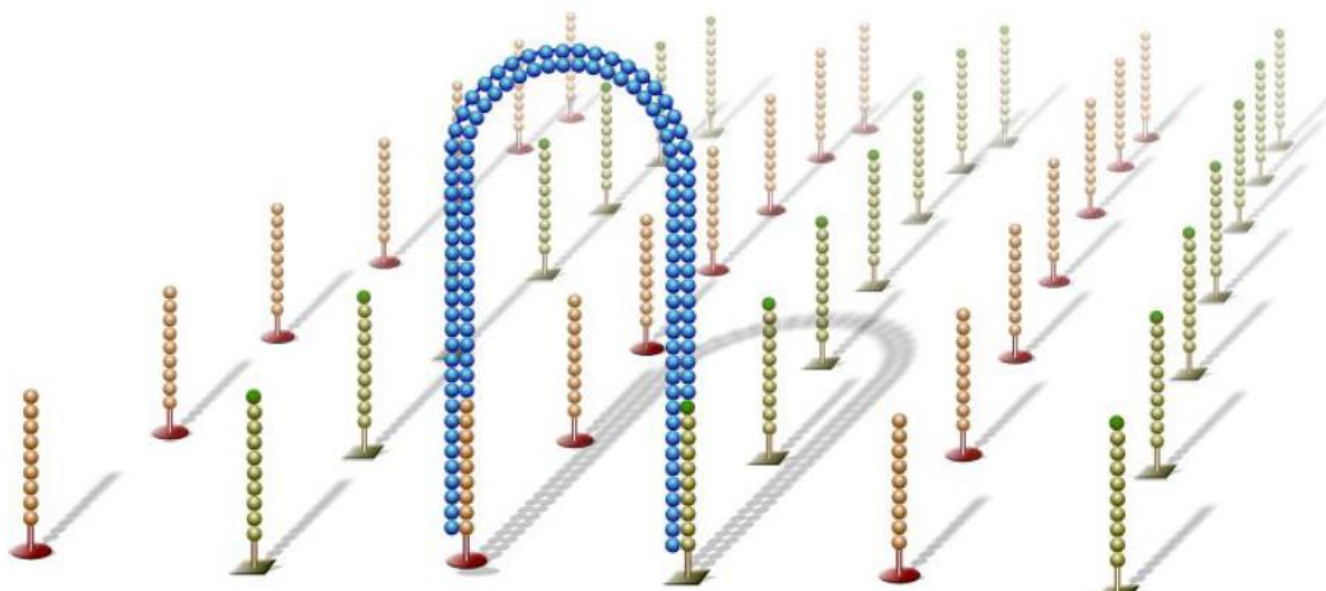




Sekvenátor II generace

2. Krok - Můstková PCR

Double-stranded bridge is formed



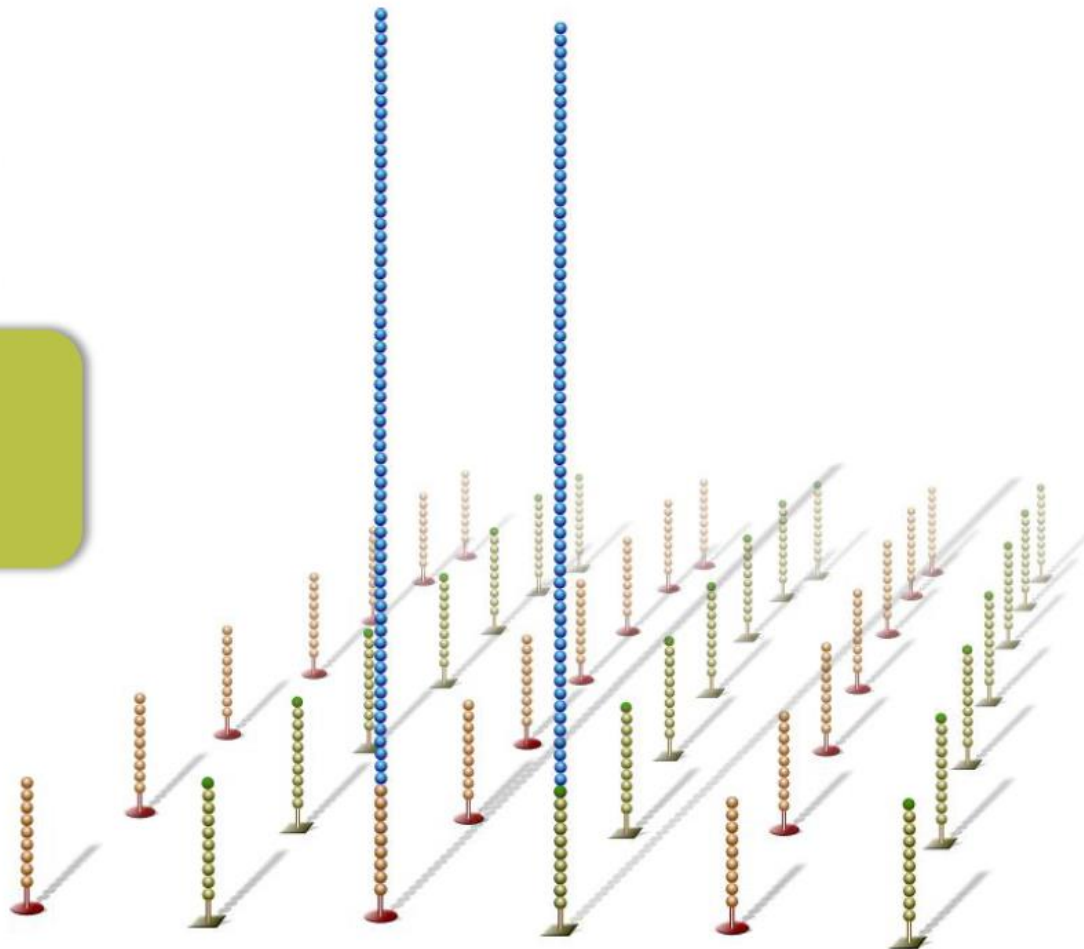


Sekvenátor II generace

2. Krok - Můstková PCR

Double-stranded bridge is denatured - 1st cycle denaturation

Result:
Two copies of covalently bound single-stranded templates



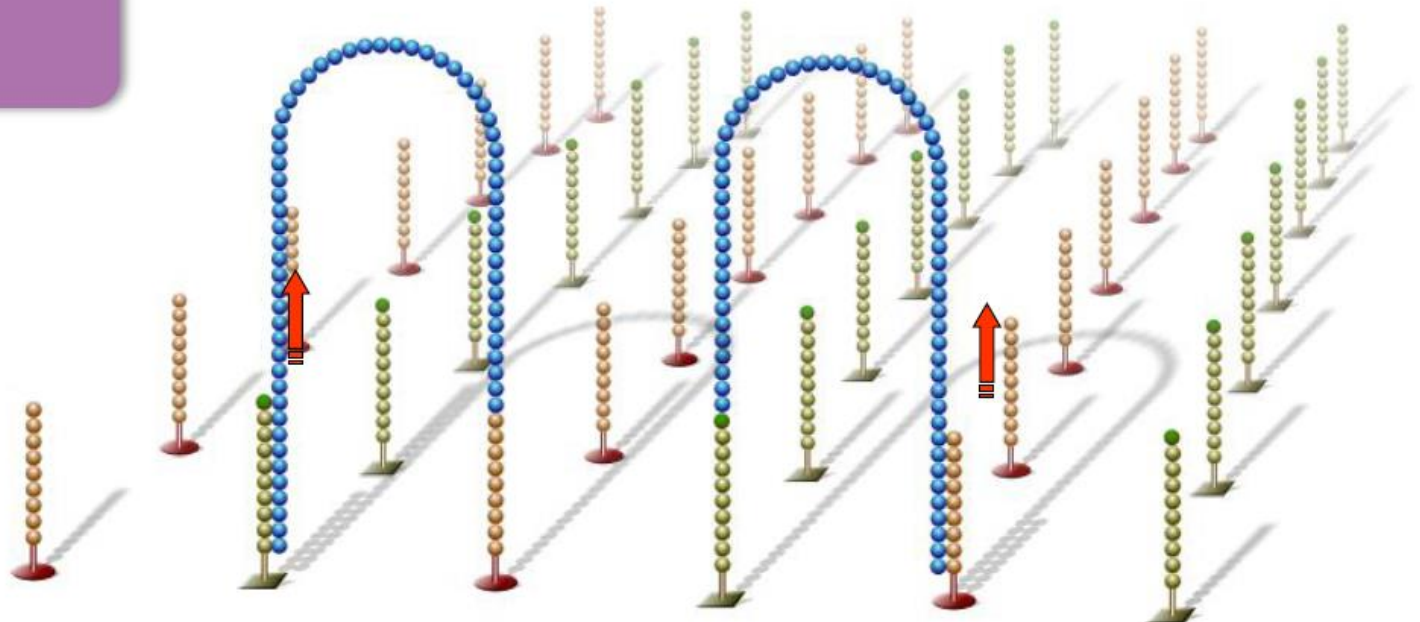


Sekvenátor II generace

2. Krok - Můstková PCR

Single-stranded molecules flip over to hybridize to adjacent primers

Hybridized primer is extended by polymerase

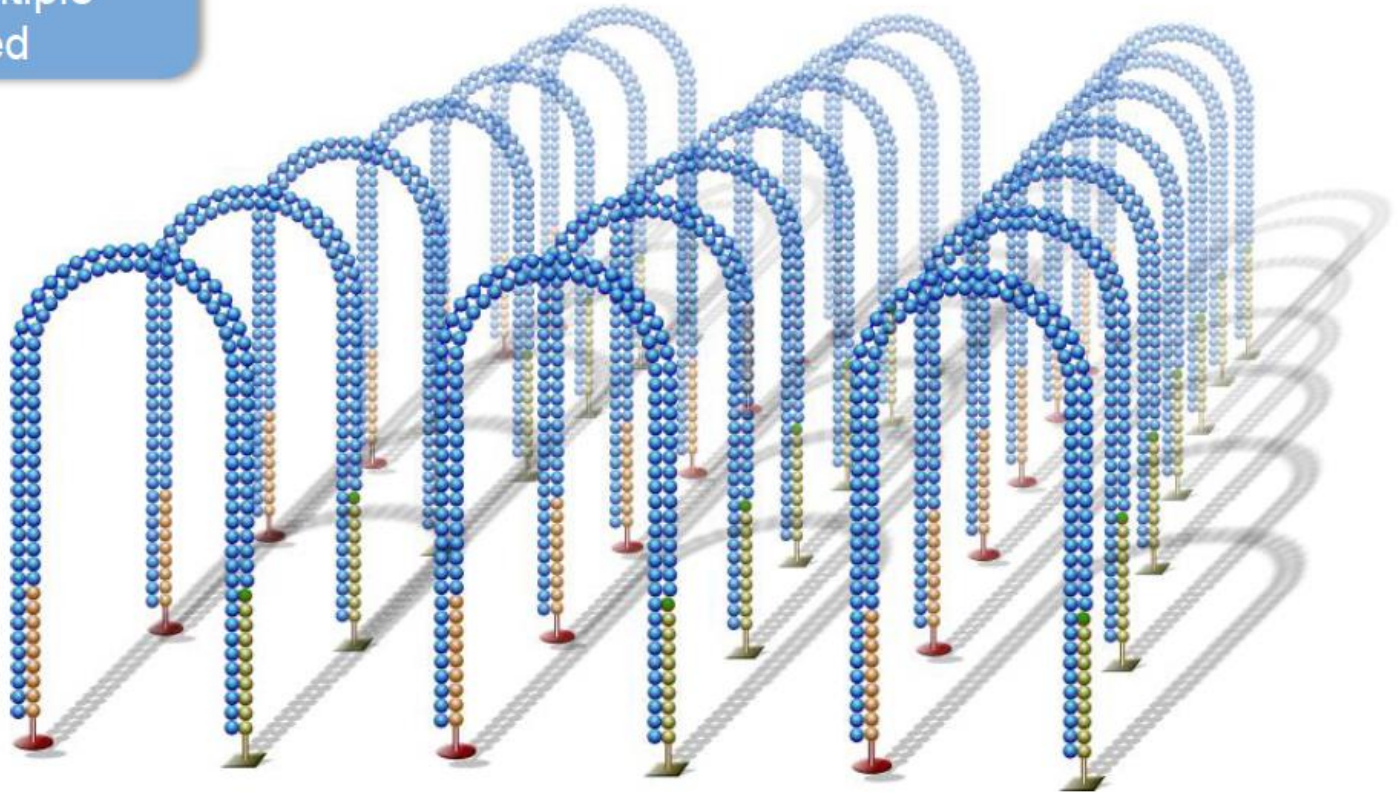




Sekvenátor II generace

2. Krok - Můstková PCR

Bridge amplification cycle repeated until multiple bridges are formed

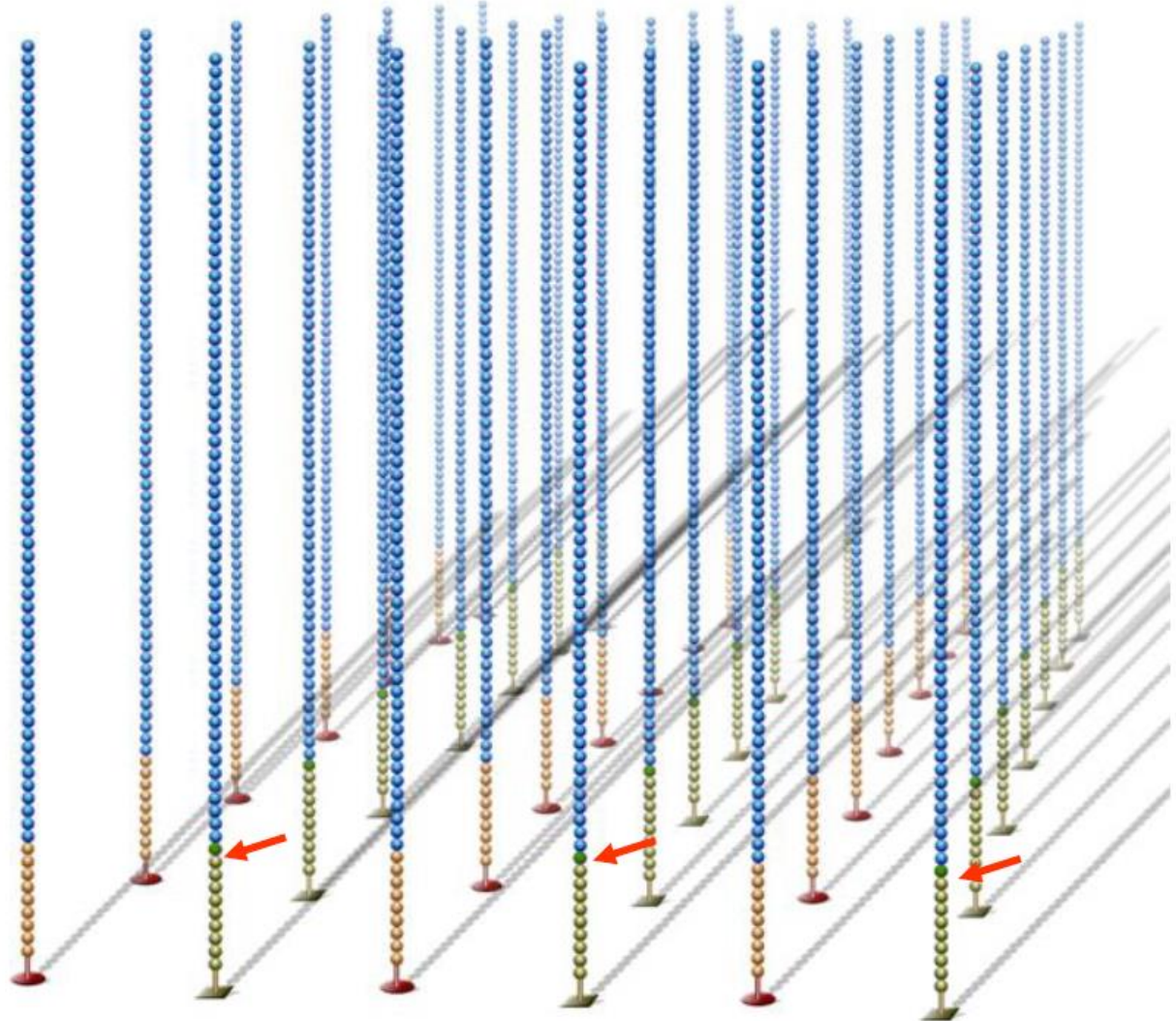




Sekvenátor II generace

3. Krok - Linearizace

dsDNA bridges are denatured

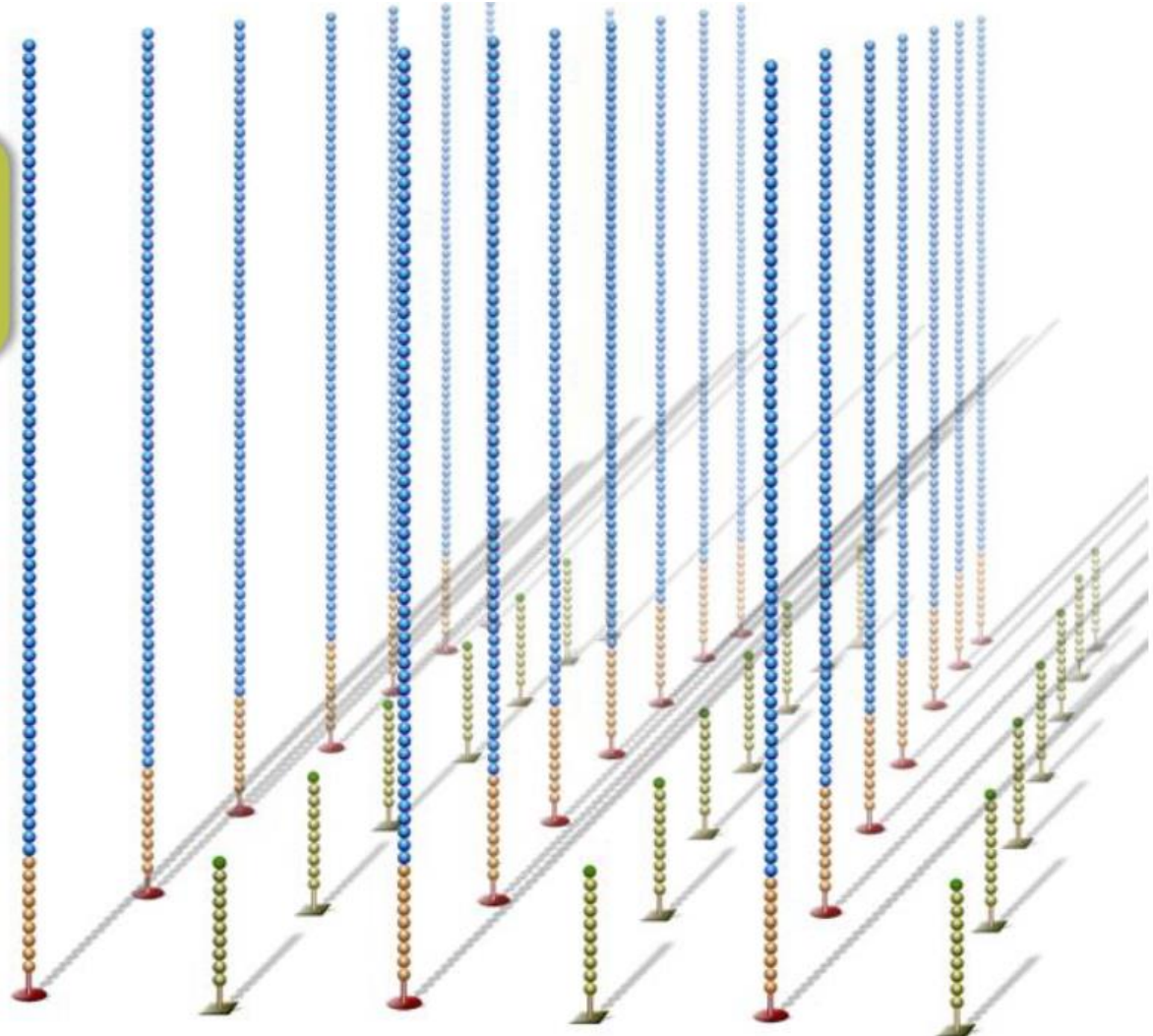




Sekvenátor II generace

4. Krok - Odštěpení reverzního vlákna

Reverse strands cleaved and washed away, leaving a cluster with forward strands only

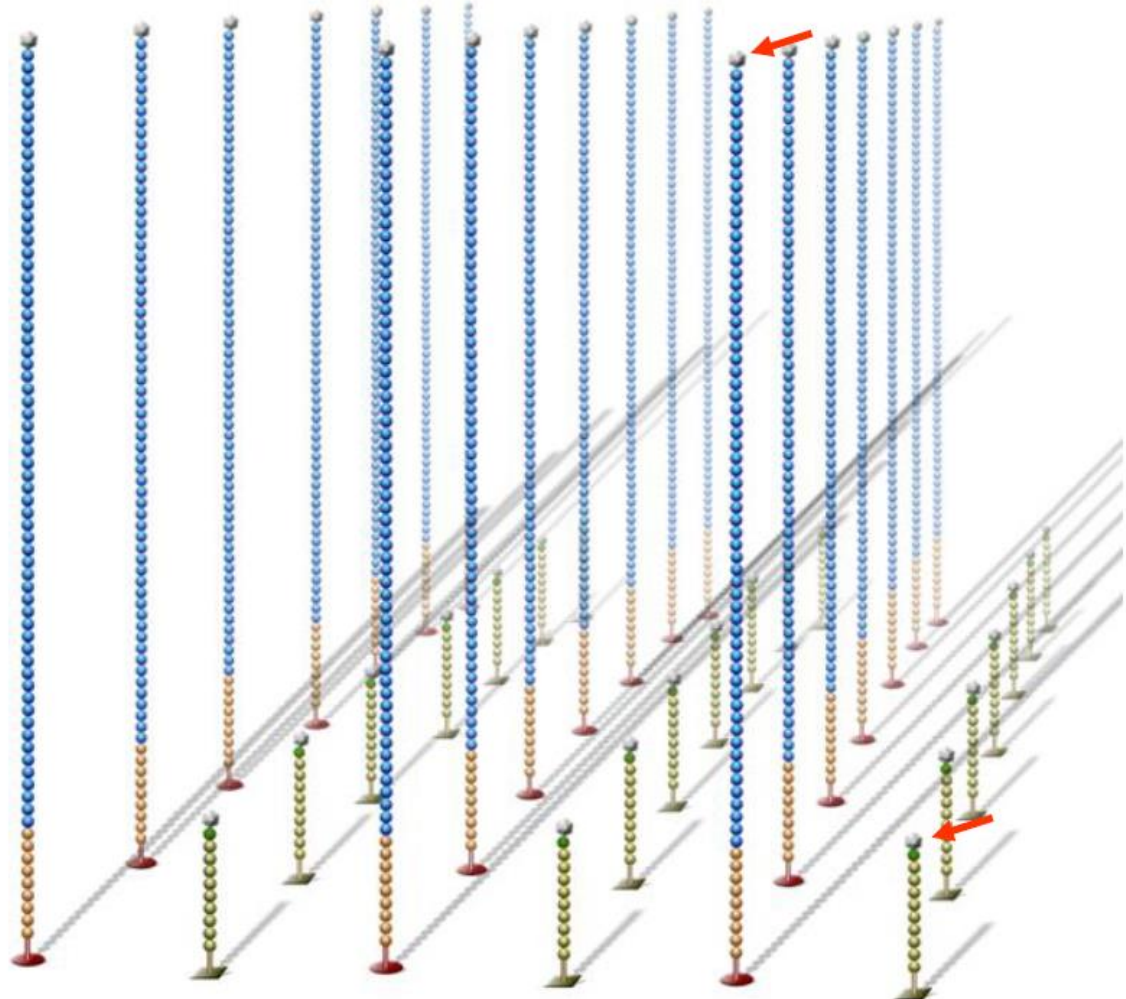




Sekvenátor II generace

5. Krok - Blokace volného 5' konce

Free 3' ends are blocked to prevent unwanted DNA priming

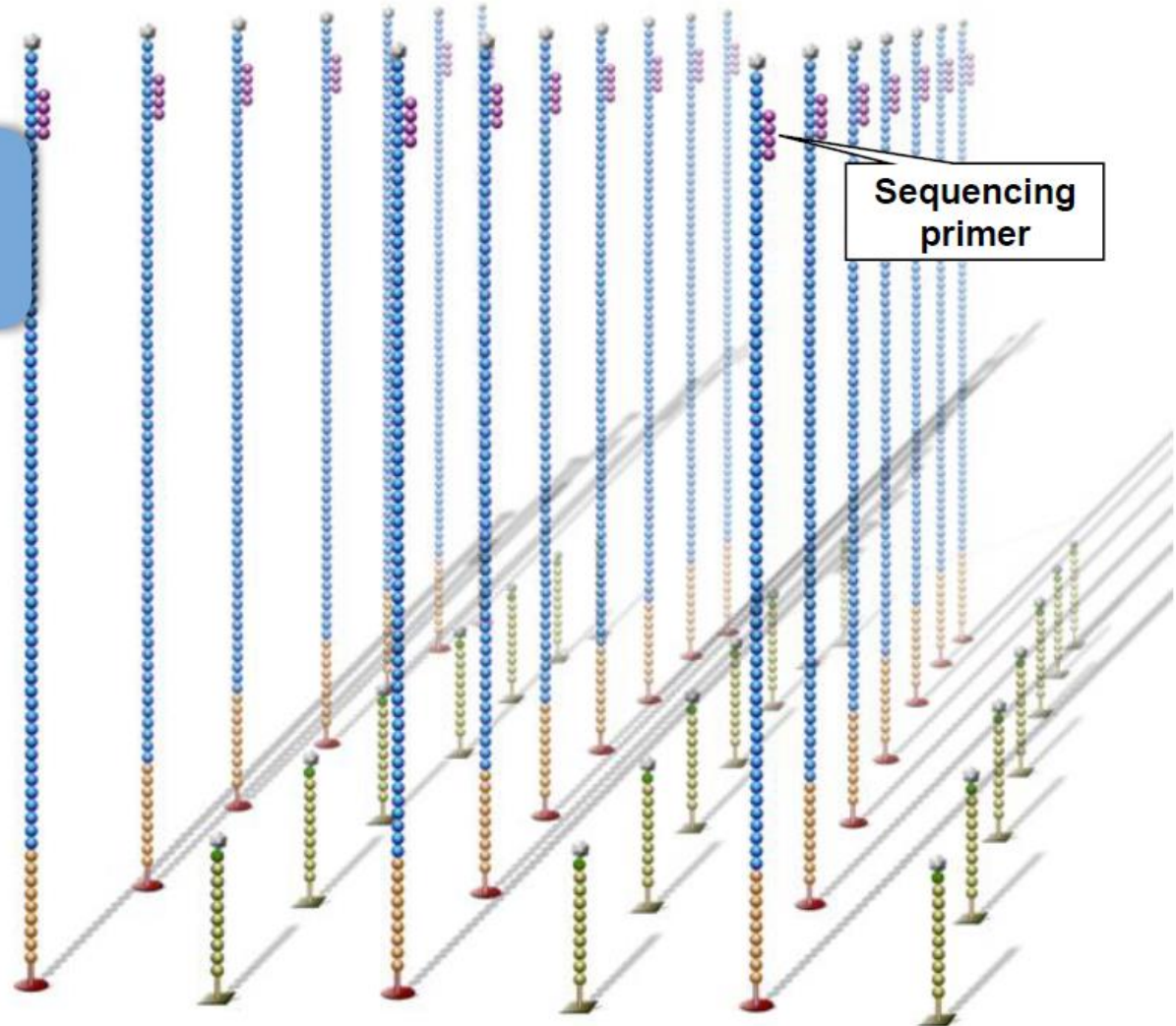




Sekvenátor II generace

6. Krok: Čtení - hybridizace sek. primeru

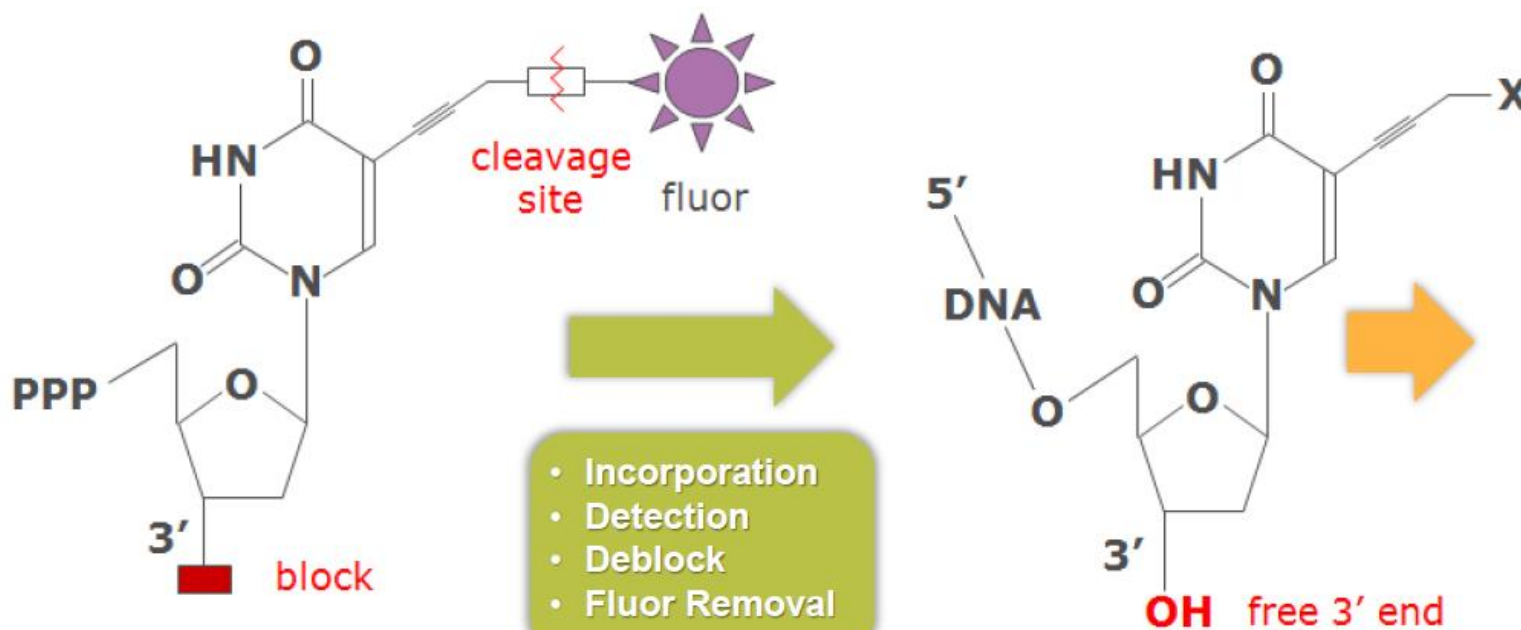
Sequencing primer is hybridized to adapter sequence



Sekvenátor II generace

Chemie reverzibilních terminátorů

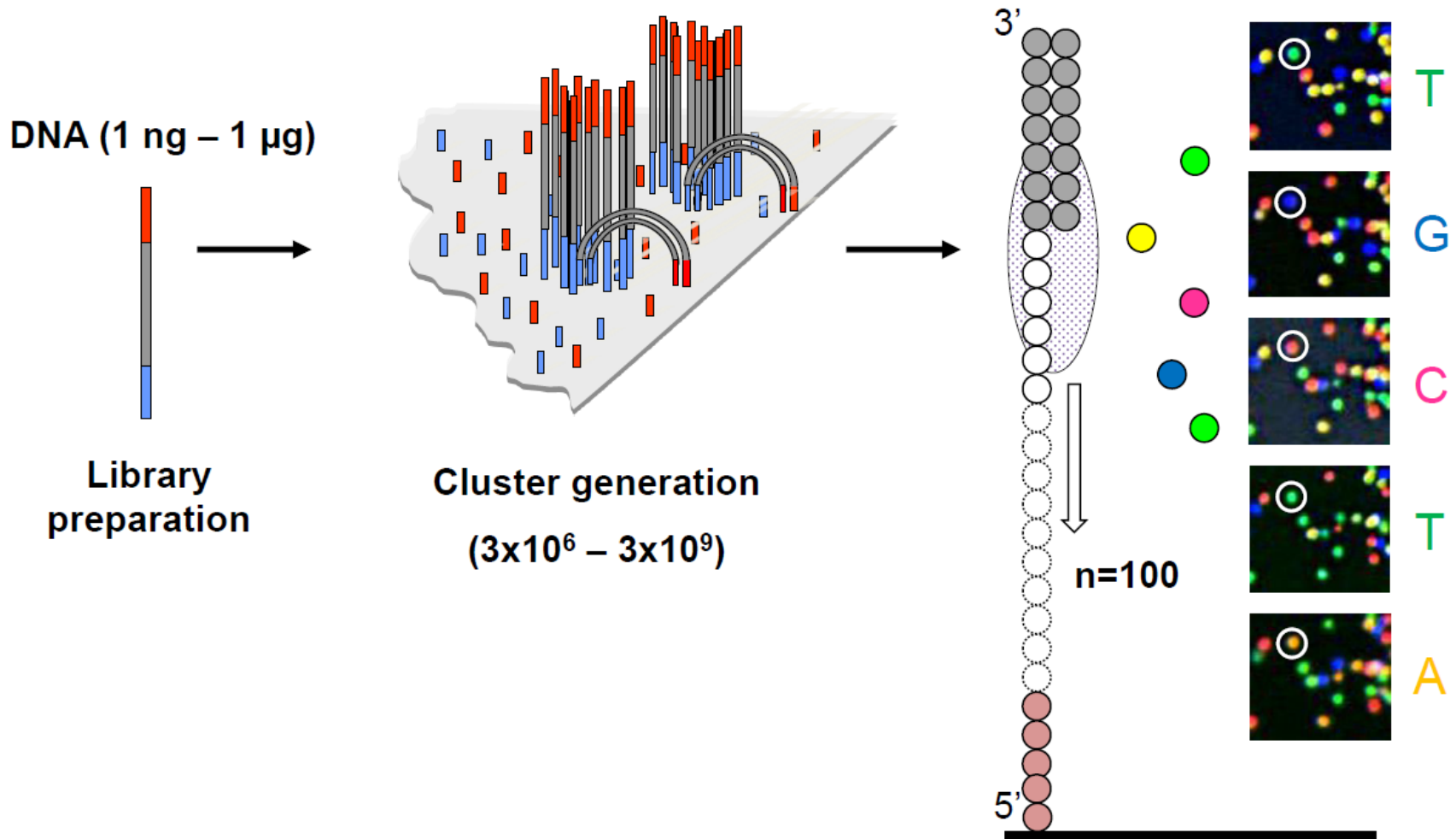
- All 4 nucleotides in 1 reaction
- Higher accuracy
- No problems with homopolymer repeats



- Incorporation
- Detection
- Deblock
- Fluor Removal

Sekvenátor II generace

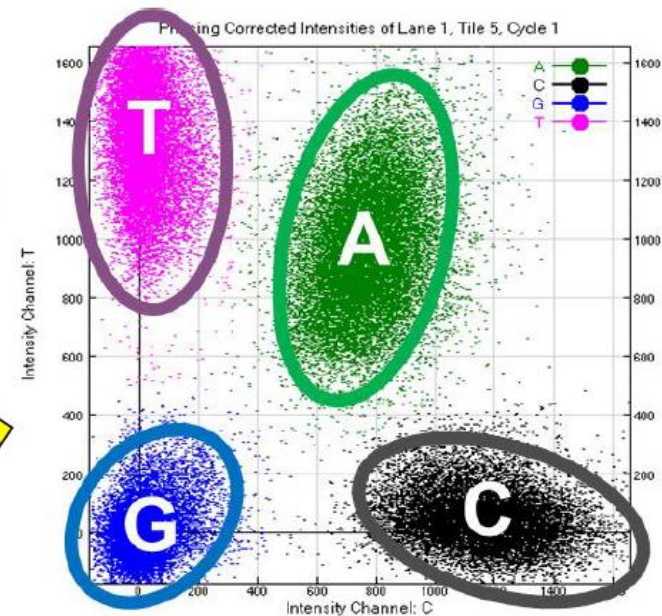
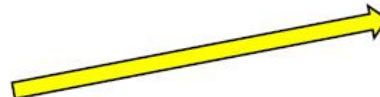
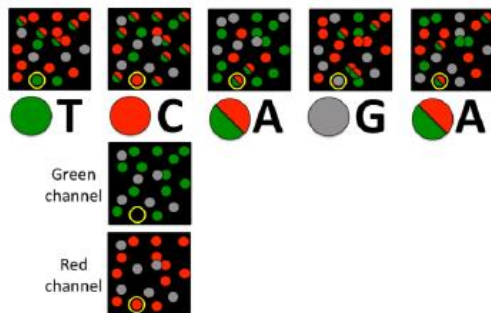
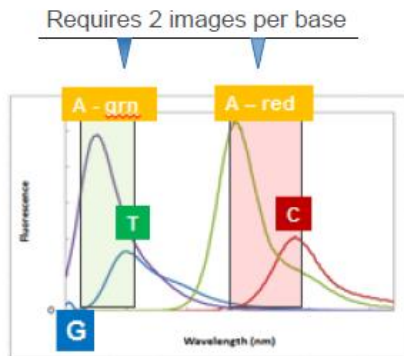
Sekvenace pomocí syntézy (SBS)





Sekvenátor II generace

2-Channel SBS Chemistry: MiniSeq, NextSeq



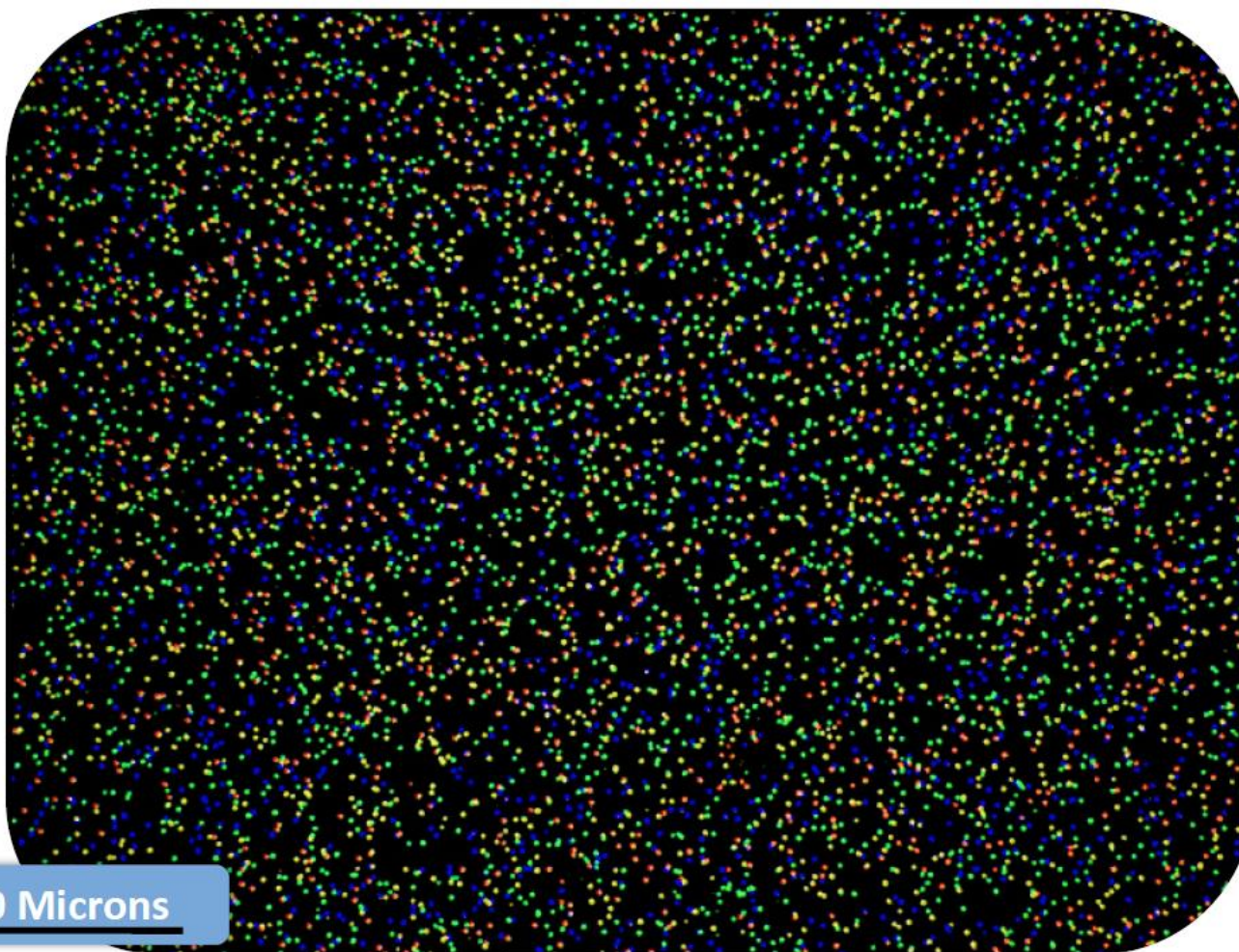
High performance with
half the pictures

50%
fewer images



Sekvenátor II generace

Clusters



100 Microns

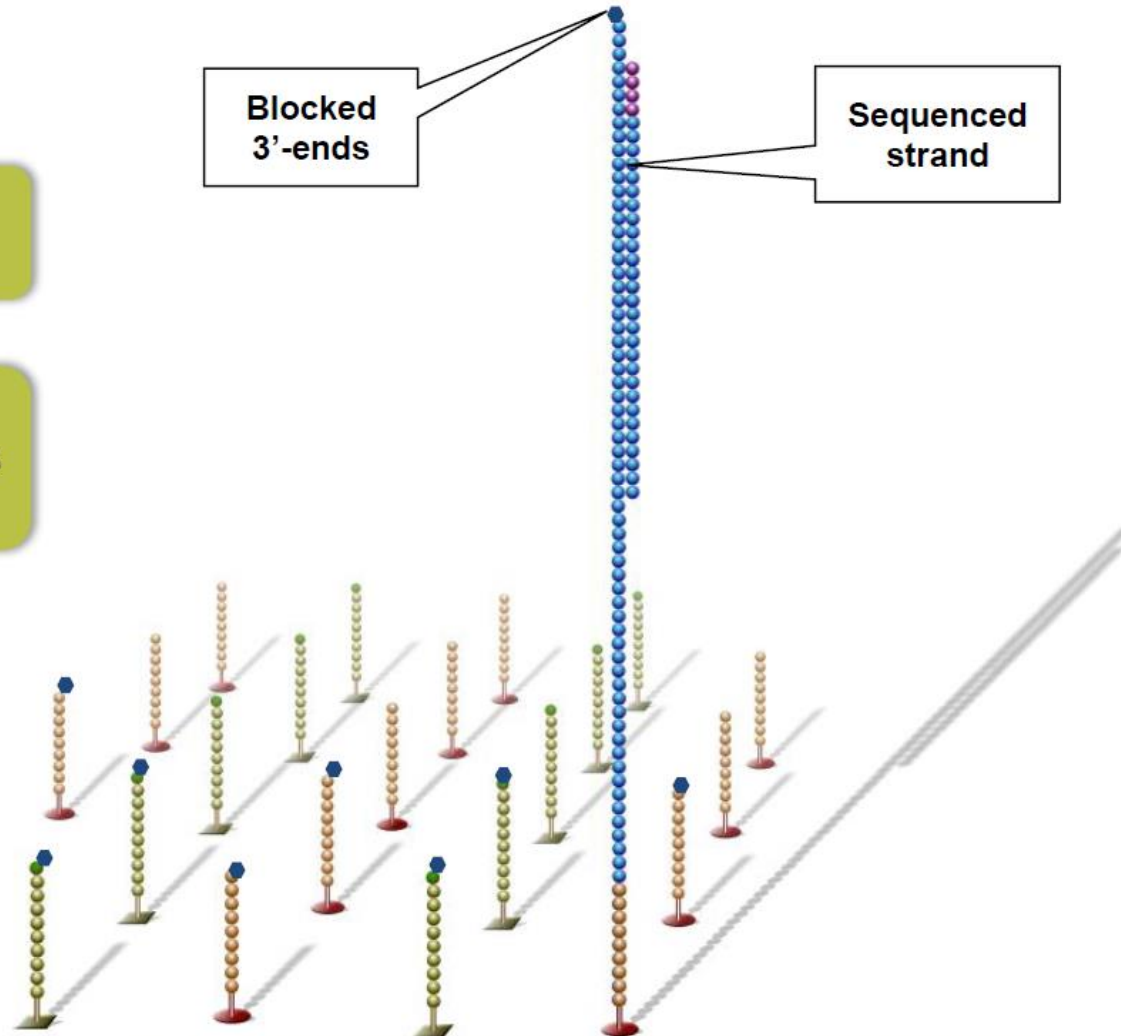


Sekvenátor II generace

7. Krok - Pair-End sekvenace

Sequenced strand is stripped off

3'-ends of template strands and lawn primers are unblocked



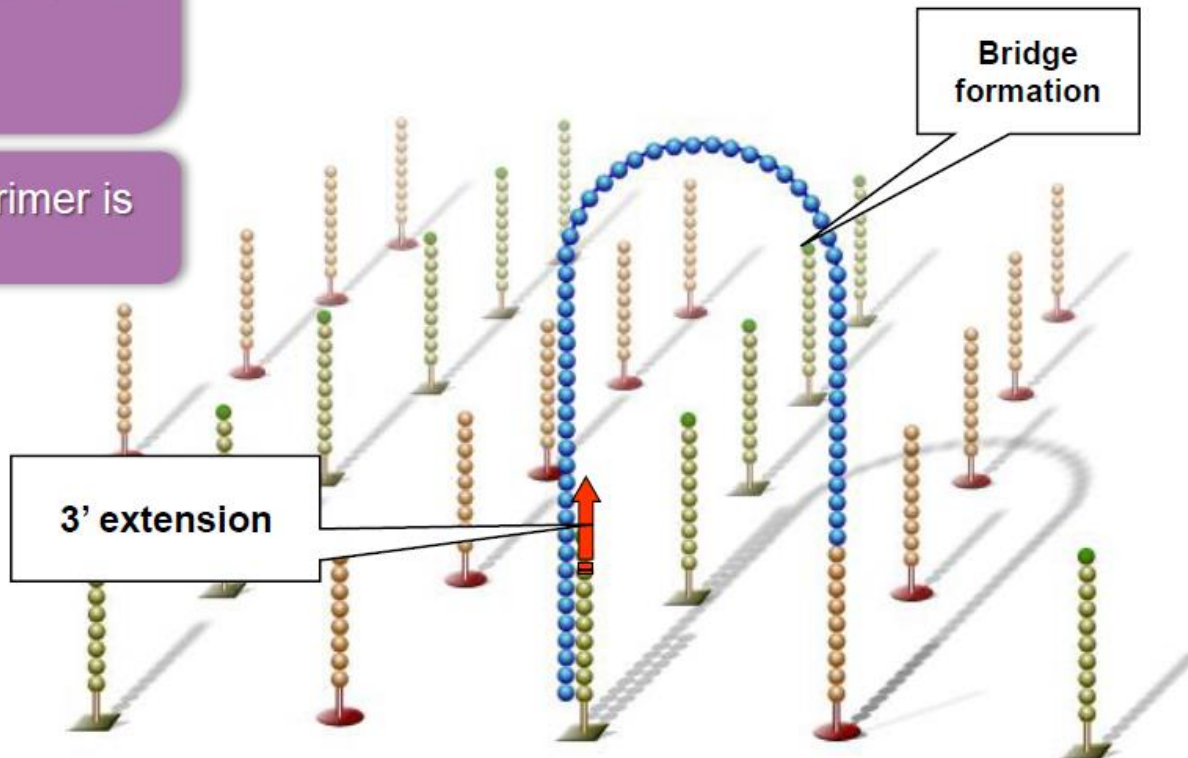


Sekvenátor II generace

7. Krok - Pair-End sekvenace

Single-stranded template loops over to form a bridge by hybridizing with a lawn primer

3'-ends of lawn primer is extended

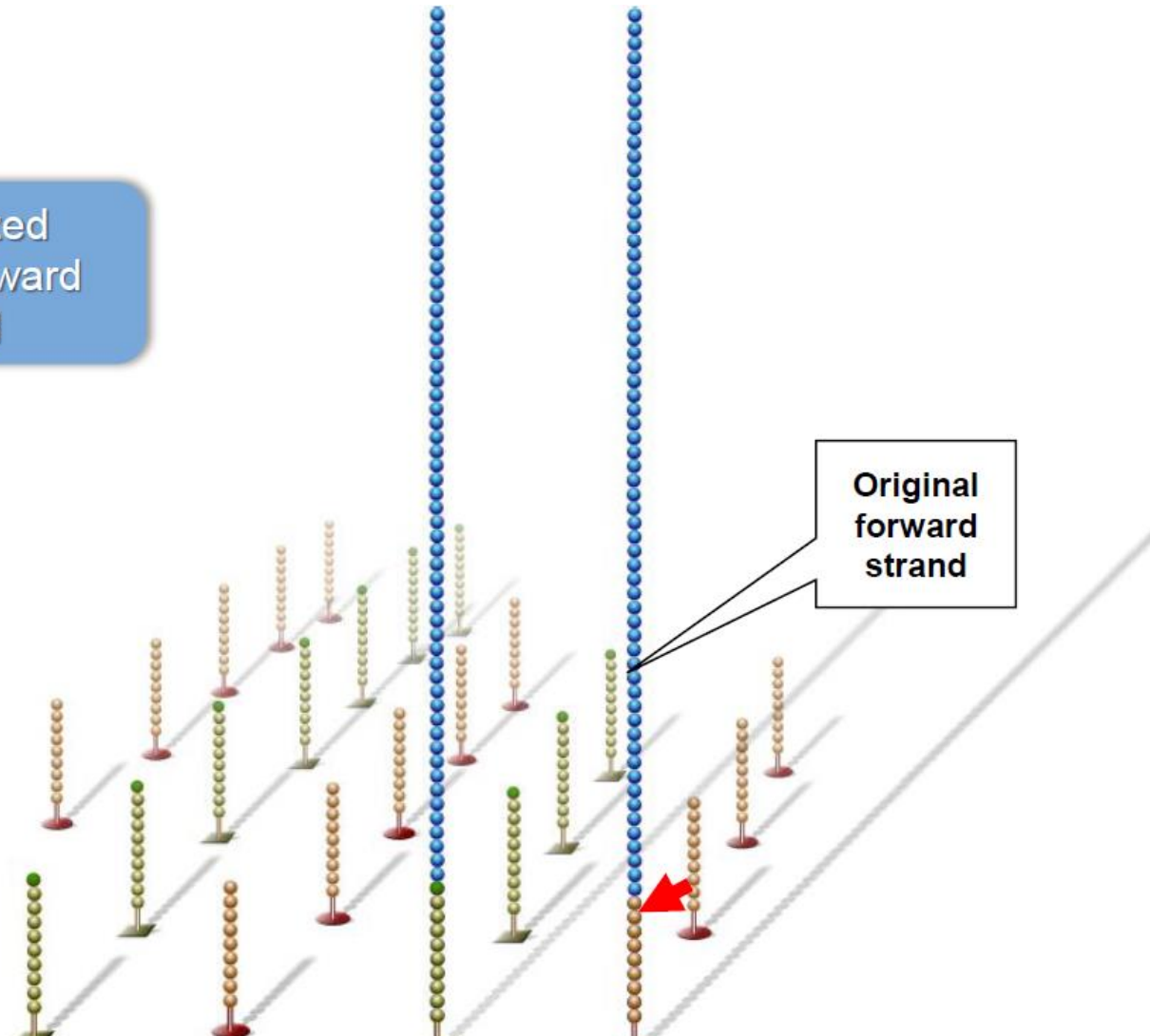




Sekvenátor II generace

7. Krok - Pair-End sekvenace

Bridges are linearized and the original forward template is cleaved



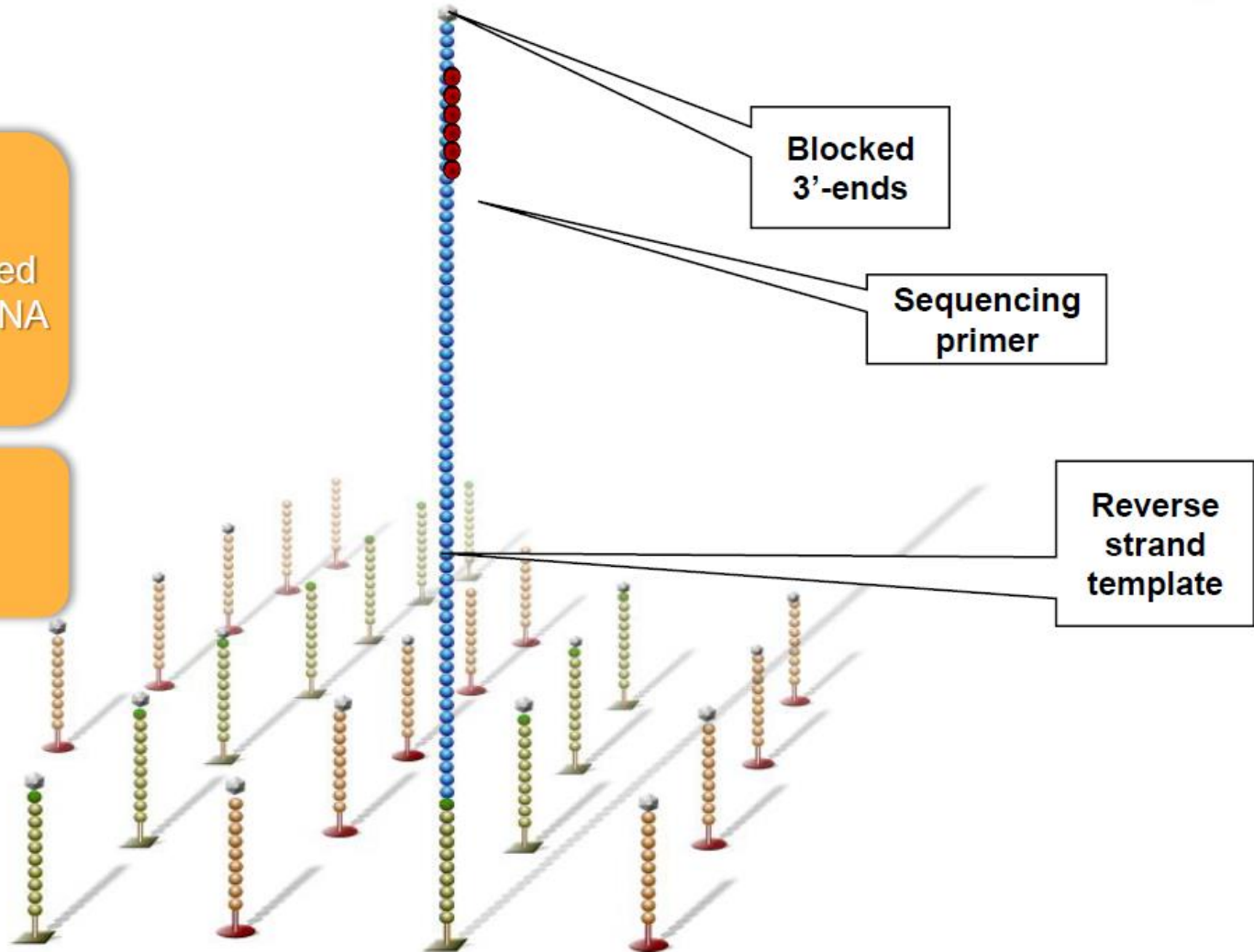


Sekvenátor II generace

7. Krok - Pair-End sekvenace

Free 3' ends of the reverse template and lawn primers are blocked to prevent unwanted DNA priming

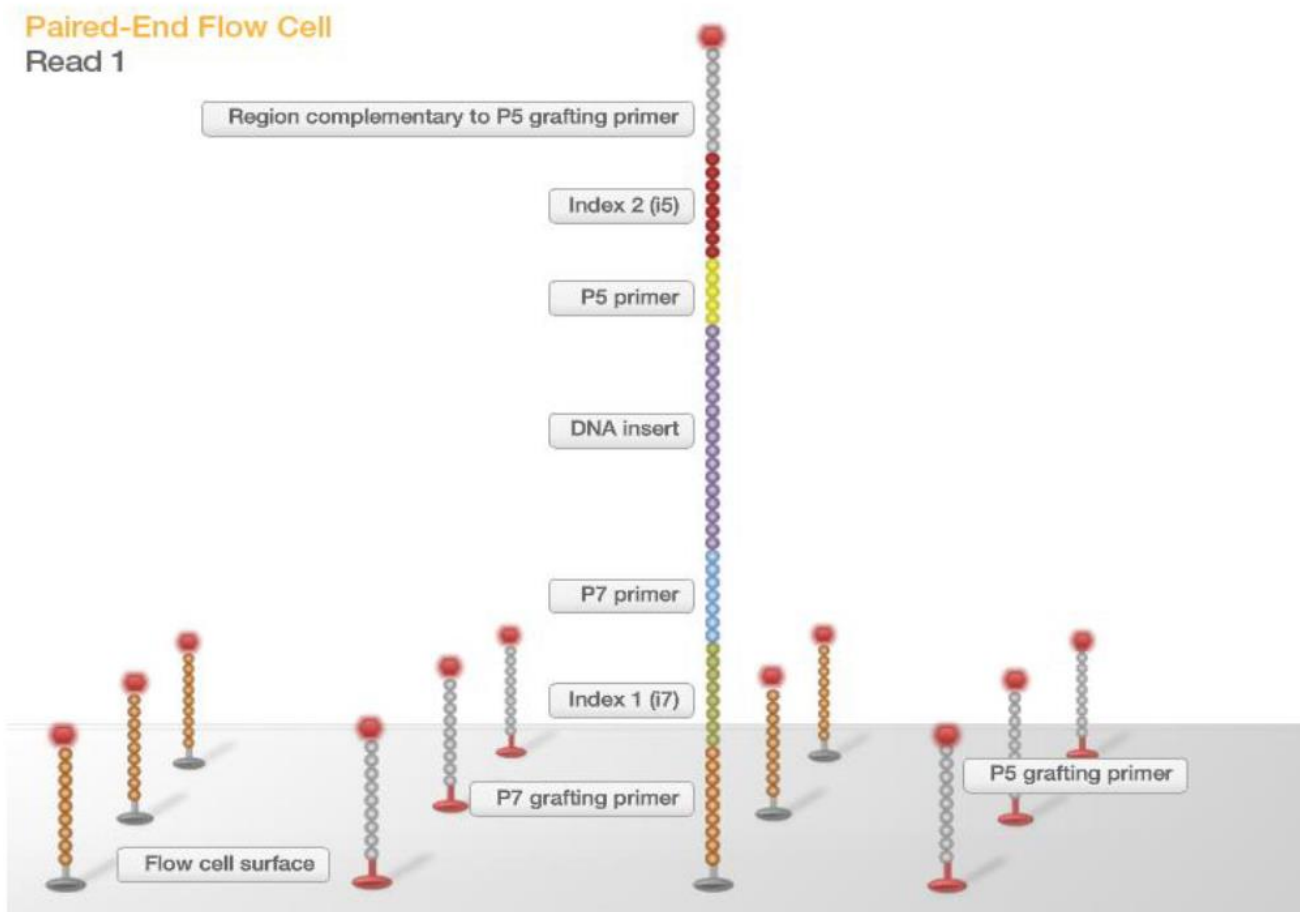
Sequencing primer is hybridized to adapter sequence





Sekvenátor II generace

Čtení indexů

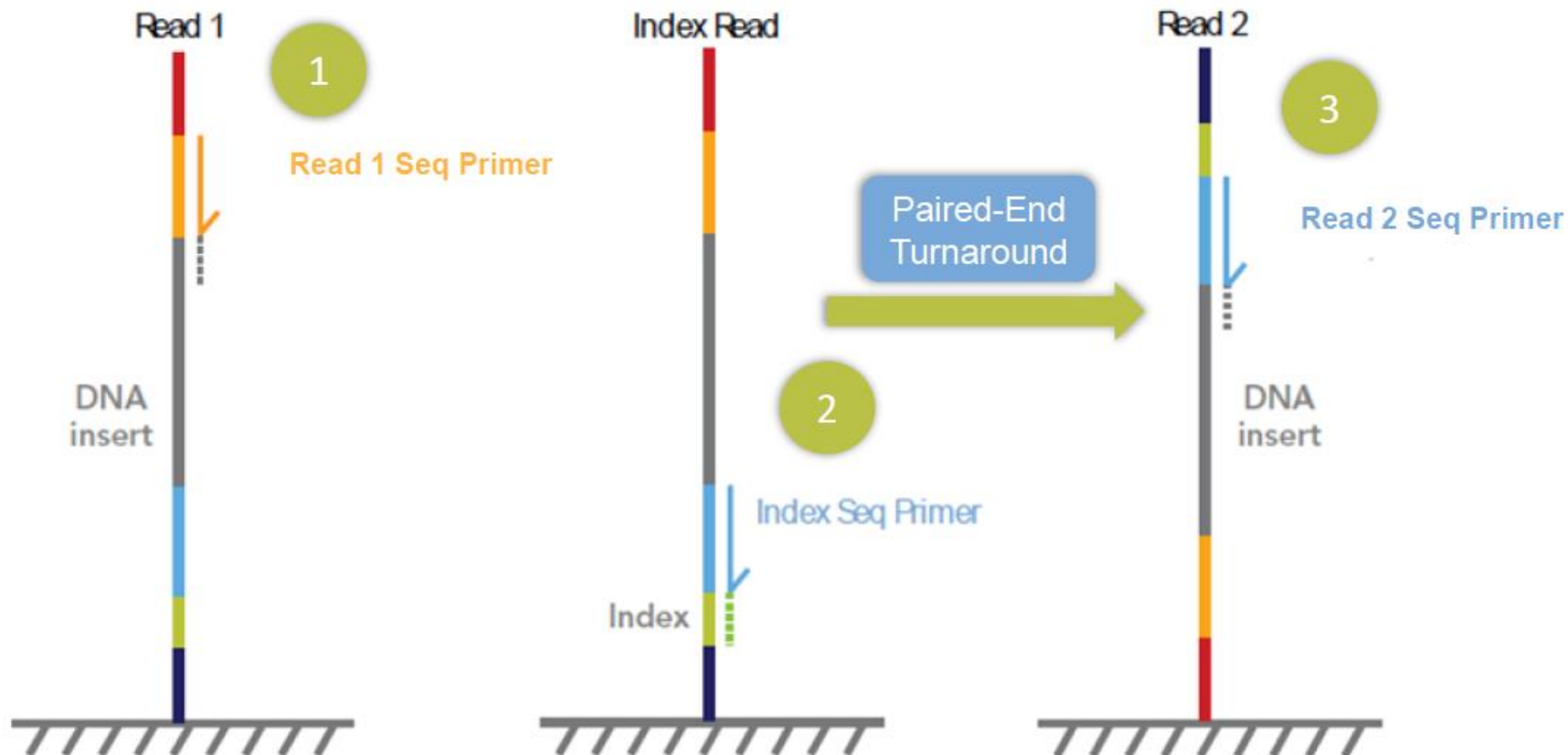




Sekvenátor II generace

Single index čtení

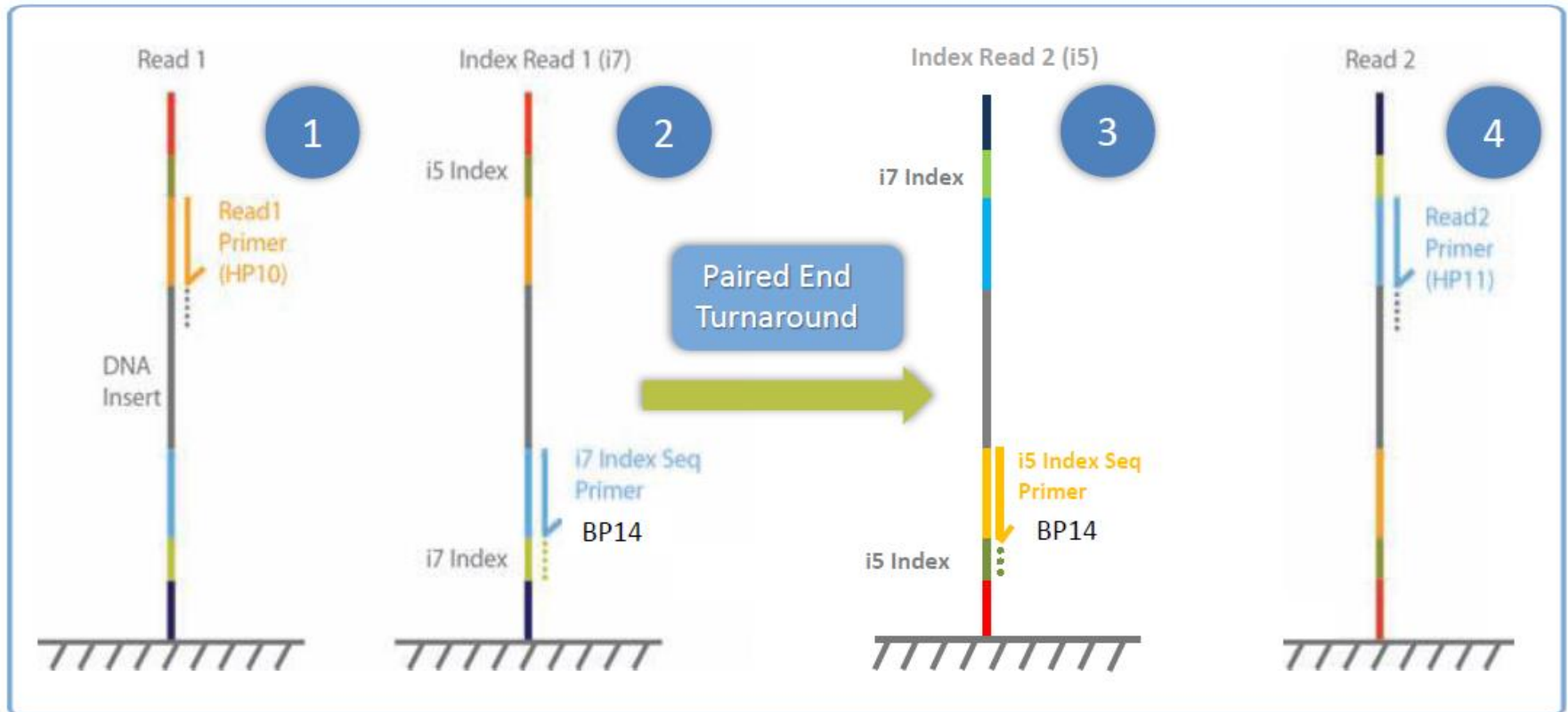
Single indexed sequencing utilizes three sequencing reads





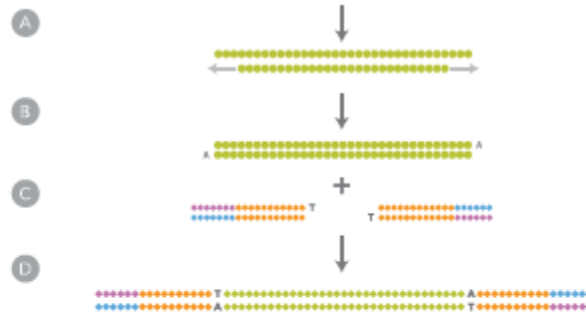
Sekvenátor II generace

Dual index čtení (iSeq, MiniSeq, NextSeq, HiSeq)



1 LIBRARY PREPARATION

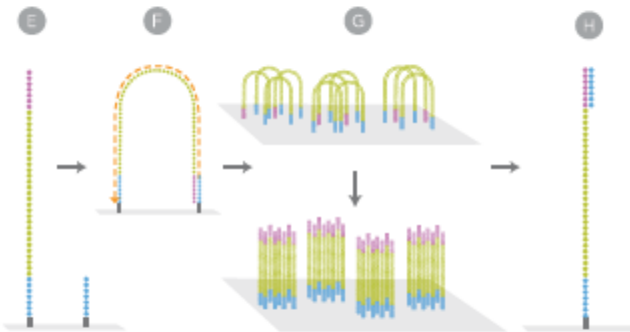
6 hours
3 hours hands-on time



- A Fragment DNA
- B Repair ends
Add A overhang
- C Ligate adapters
- D Select ligated DNA

2 CLUSTER GENERATION

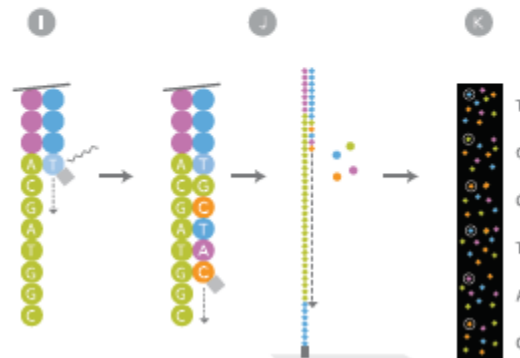
4 hours
< 10 minutes hands-on time
1-96 samples



- E Attach DNA to
flow cell
- F Perform bridge
amplification
- G Generate clusters
- H Anneal sequencing
primer

3 SEQUENCING

1-3 days single-read run
3-9 days paired-end run
30 minutes hands-on time
8 lanes, up to 96 samples
per flow cell (run)



- I Extend first base,
read, and deblock
- J Repeat step above
to extend strand
- K Generate base calls



Sekvenátor II generace - SOLiD

See the difference
The SOLiD™ 4 System

[Learn More →](#)

AB applied biosystems
by life technologies™

SOLiD™ 4
SYSTEM SEQUENCING

Library Type	Read Length	Number of Slides/Run	Typical Mappable* Output/Slide	Typical Mappable* Output/Run (2 slides)	Total Tags/Run	Number of Days/Run
Fragment Library	35 bases	1-2	1.5-2 GB	3-4 GB	86M - 114M	6 days/run
Mate-Paired Library	2 x 25 bases	1-2	2.5-3 GB	5-6 GB	200M - 240M	10 days/run

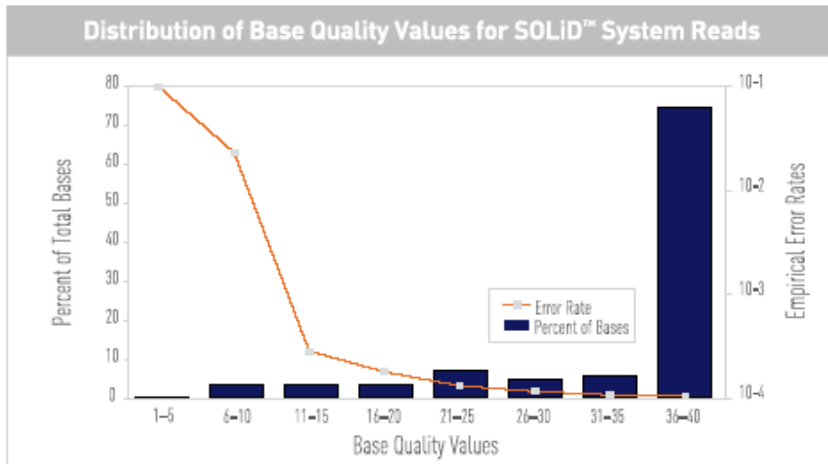
Sekvenátor II generace - SOLiD



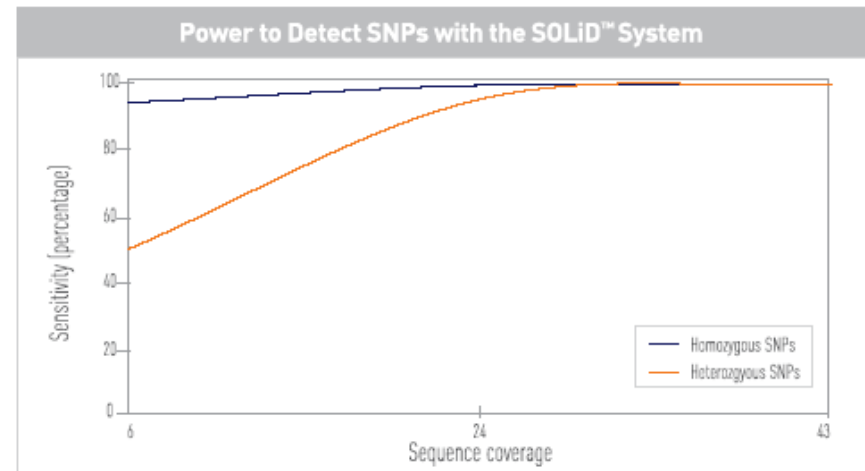
New SOLiD™ Systems			
Specifications	SOLiD™ 4 System	SOLiD™ 4hq System*	SOLiD™ PI System*
Throughput/Run**	Up to 100 Gb	Up to 300 Gb	Up to 50 Gb
System Accuracy	99.94%	99.99%	99.99%
Cost/Genome***	As low as \$6,000	As low as \$3,000	As low as \$8,000
Read Length	<ul style="list-style-type: none"> Fragment: 50 bp Mate-pair: 2 x 50 bp Paired-end: 50 x 25 bp 	<ul style="list-style-type: none"> Fragment: 75 bp Mate-pair: 2 x 75 bp Paired-end: 75 x 35 bp 	<ul style="list-style-type: none"> Fragment: 75 bp Mate-pair: 2 x 75 bp Paired-end: 75 x 35 bp
Multiplexing	<ul style="list-style-type: none"> 48 RNA barcodes 96 DNA barcodes 	<ul style="list-style-type: none"> 96 RNA barcodes 96 DNA barcodes 	<ul style="list-style-type: none"> 96 RNA barcodes 96 DNA barcodes
Run Time	<ul style="list-style-type: none"> 3 days for 35 bp 11 days for 50 x 25 bp 12 days for 2 x 50 bp 	<ul style="list-style-type: none"> 3 days for 35 bp 12 days for 75 x 35 bp 14 days for 2 x 75 bp 	<ul style="list-style-type: none"> 1 day for 35 bp

SOLiD - základní parametry čtení

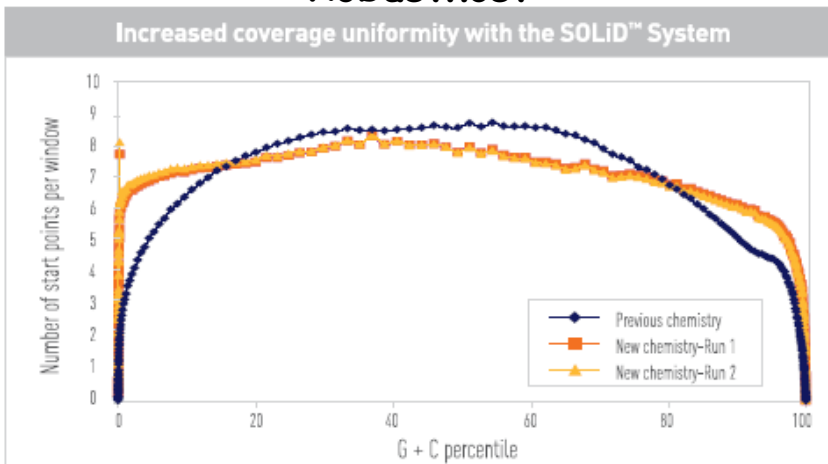
Kvalita



Přesnost



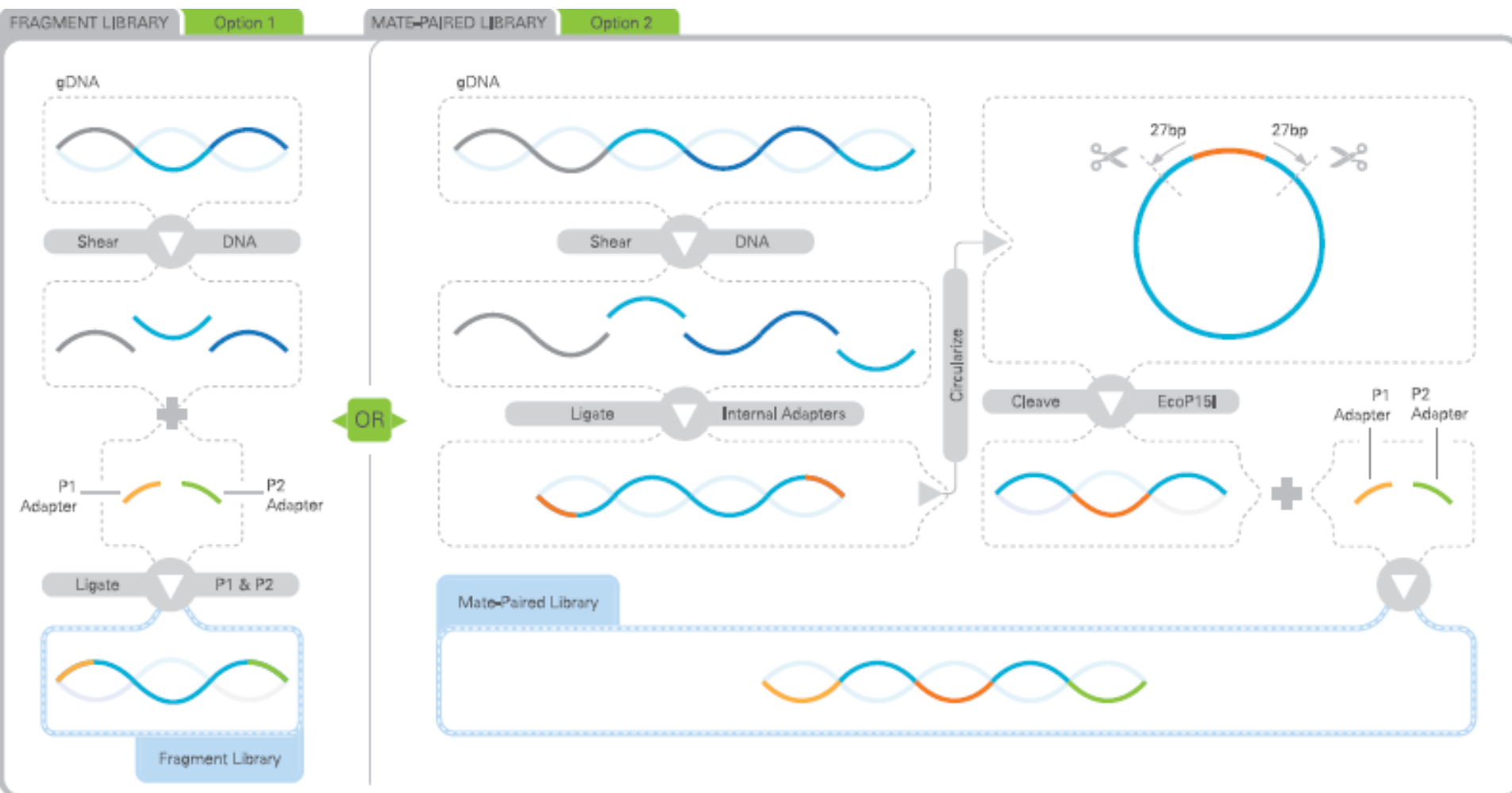
Robustnost



- Pro detekci polymorfismů nutno opakovat minimálně 20x
- Pro detekci somatických mutací 30x

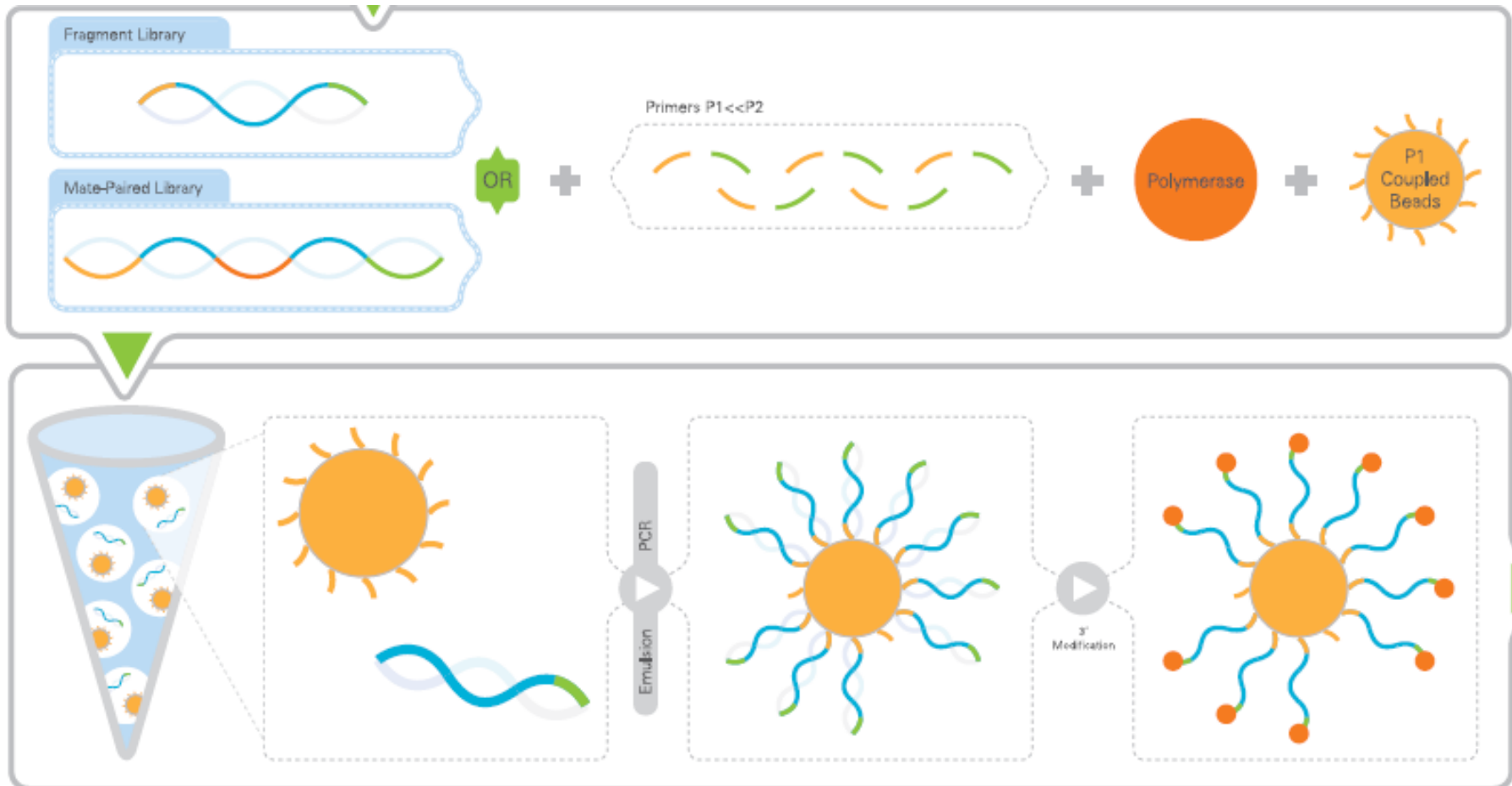
Sekvenátor II generace - SOLiD

Vytvoření tzv. Sekvenační knihovny



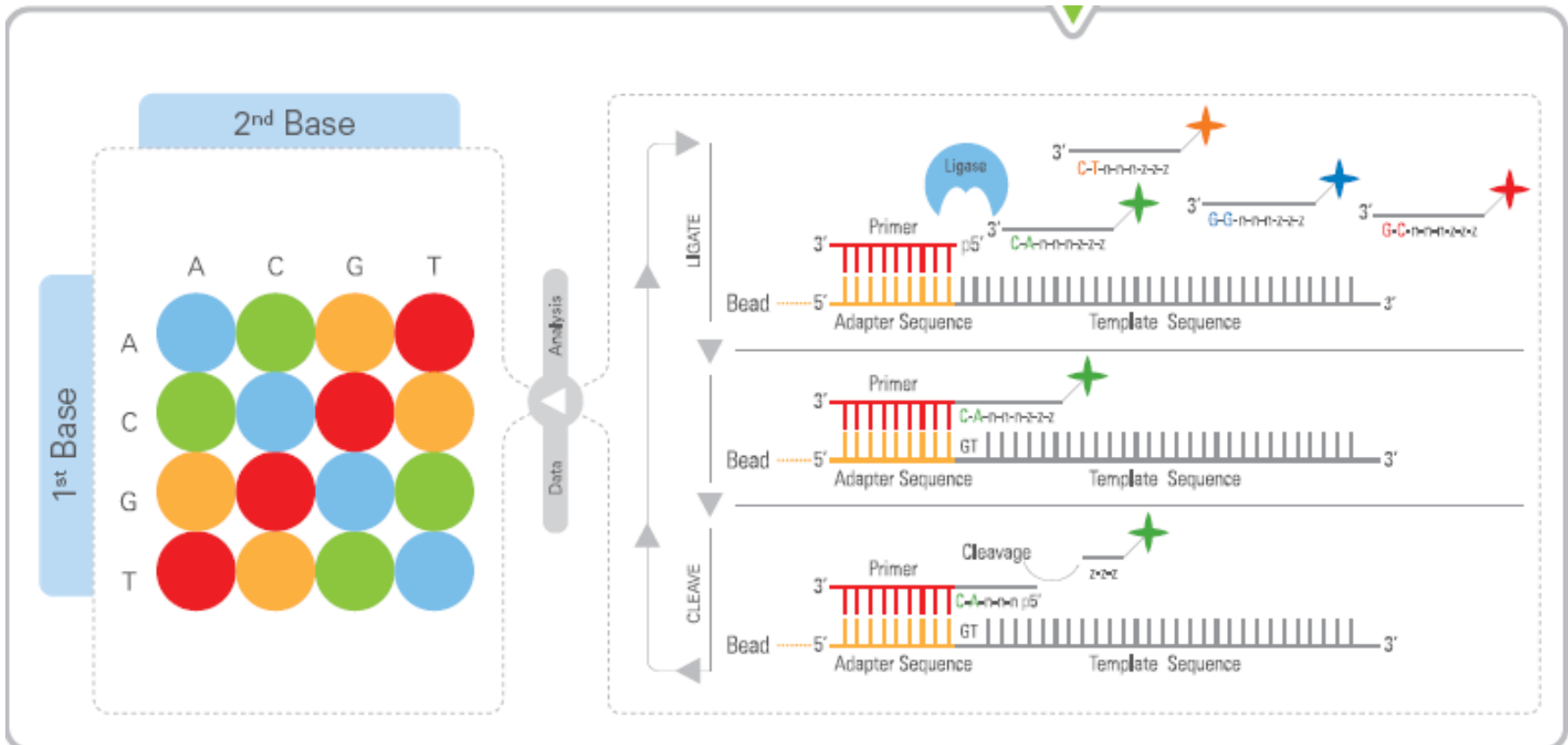
Sekvenátor II generace - SOLiD

Namnožení Sekvenační knihovny (em-PCR)



Sekvenátor II generace - SOLiD

Sekvenace pomocí ligační reakce





Sekvenátor II generace - Ion Torrent

Ion 314™ Chip v2:

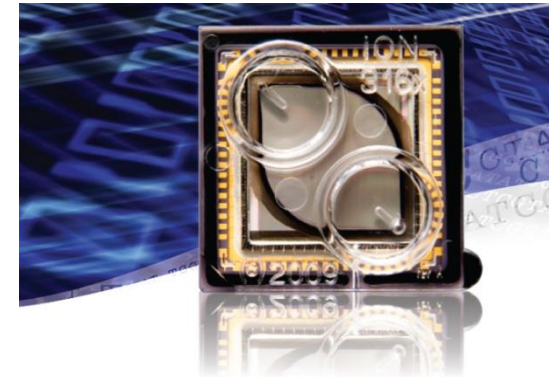
30-100 Mb sekvenčních dat s dobou čtení 2-4 hodiny

Ion 316™ Chip v2:

300 Mb-1.0 Gb sekvenčních dat s dobou čtení 3-5 hodin

Ion 318™ Chip v2:

600 Mb-2.0 Gb sekvenčních dat s dobou čtení 4-7 hodin



Sekvenátor II generace - Ion Torrent

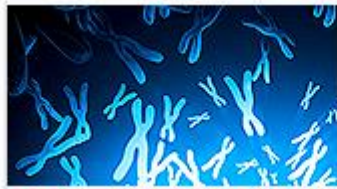
Postup práce



Aplikace (Sekvenace)



Cílená



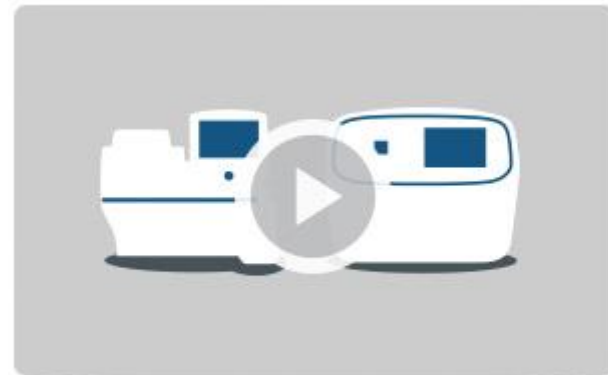
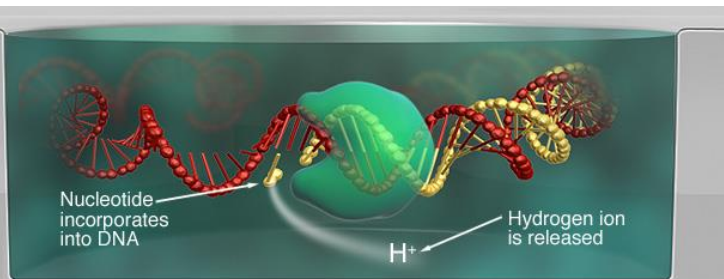
Exomu



Transkriptomu



Genomu



Kvantifikace NGS knihovny

Elektroforetické metody

- Fragment Analyzer (Adv. Anal.)
- TapeStation (Agilent)
- BioAnalyzer (Agilent)



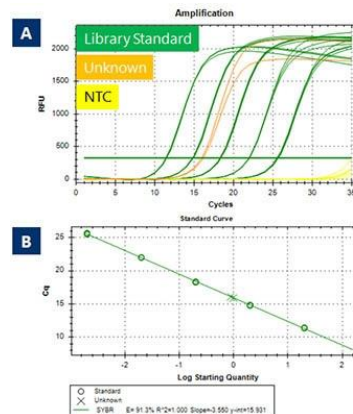
Fluorimetrické metody

- Qubit (Thermo Scientific)



Real-Time PCR

- KapaBiosystem
- NEB





Sekvenátor III generace - PacBio



Sequel System



PacBio RS II

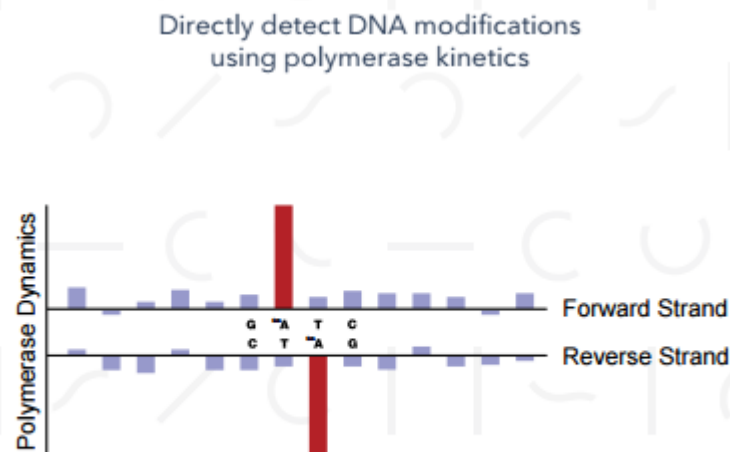
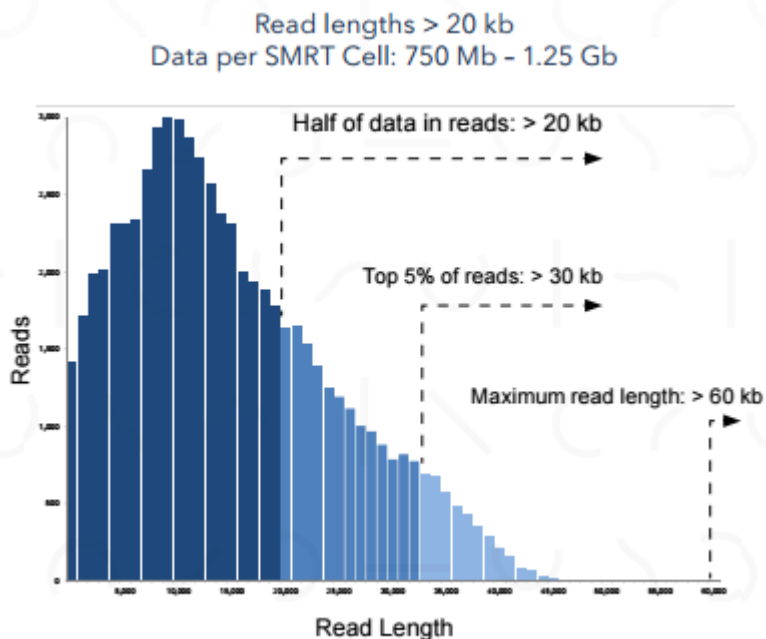




Sekvenátor III generace - PacBio

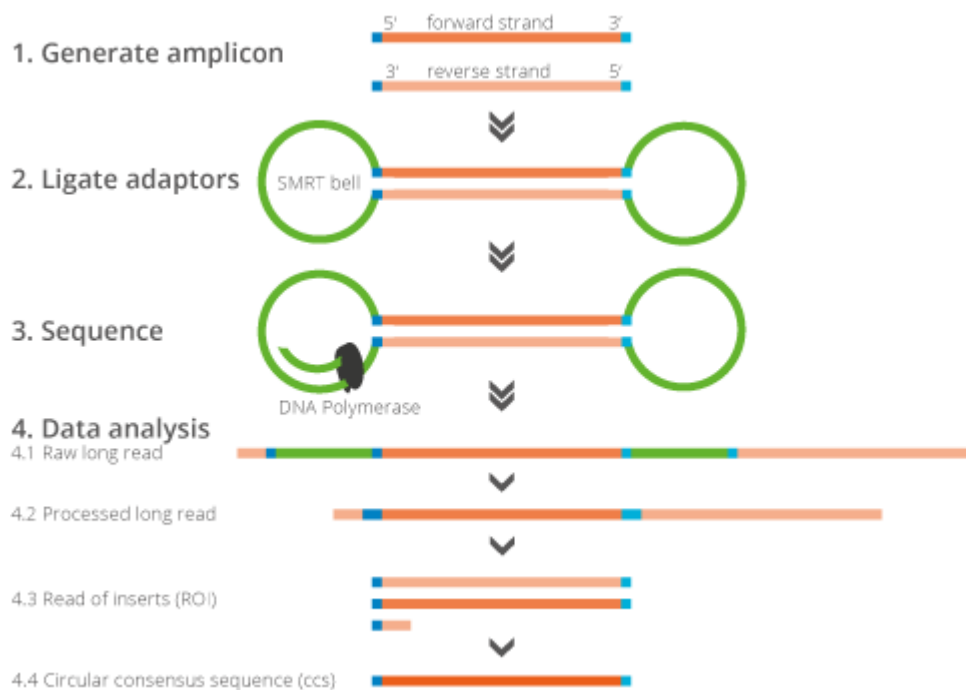


- Sekvence založena na Single Molecule, Real-Time (SMRT®) technologii
- Vyžívá tzv. Zero-Mode Waveguides (ZMWs) umožňující osvětlení pouze spodní části jamky, ve které je dole imobilizována DNA polymeráza
- Hlavní výhoda je možnost dlouhého čtení (až 20 kb)
- Další výhoda je možnost přímé detekce methylovaných bazí (epigenom)





Příprava knihovny



<https://www.youtube.com/watch?v=v8p4ph2MAvI>



Sekvenátor III generace Oxford Nanopores



Minion

- Základem technologie jsou nanopóry (nanodíry)
- Na začátku sekvenace je NK navázána na nanopór tvořený proteinem
- Poté je rozpletena a prochází přes nanopór, což generuje změnu proudu
- Na základě pozorování změny jsou odečítány v reálném čase jednotlivé báze
- Umožňuje sekvenaci velmi dlouhých úseků (desítky až stovky kilobází)
- Nevýhodou je vyšší chybovost, správnost >90%

