



# **Sekvenování nukleových kyselin a analýza DNA sekvencí**

**doc. RNDr. Milan Bartoš, Ph.D.**

**bartosm@vfu.cz**

**Přírodovědecká fakulta MU, 2015**

# ***Obsah přednášky***

- 1) Pojem sekvenování**
- 2) Maxamovo-Gilbertovo sekvenování**
- 3) Sangerova metoda s využitím fluoroforů**
- 4) Pyrosekvenování**
- 5) Next generation sequencing**
- 6) Third generation sequencing**
- 7) Příklad reálné analýzy**



# *Doporučená literatura*

[www.farmakogenomika.cz](http://www.farmakogenomika.cz)

# ***Sekvenování***

## **Rozhodující metoda pro stanovení nukleotidových sekvencí**

- **Konečná fáze procesu individualizace jednotlivých izolátů**
- **Metoda je pro většinu mikrobiologických aplikací příliš přesná**

# ***Metody sekvenování nukleových kyselin***

**Chemická metoda sekvenování**  
(Maxamovo-Gilbertovo sekvenování)

**Enzymová metoda sekvenování**  
(Sangerovo sekvenování)

**Pyrosekvenování**

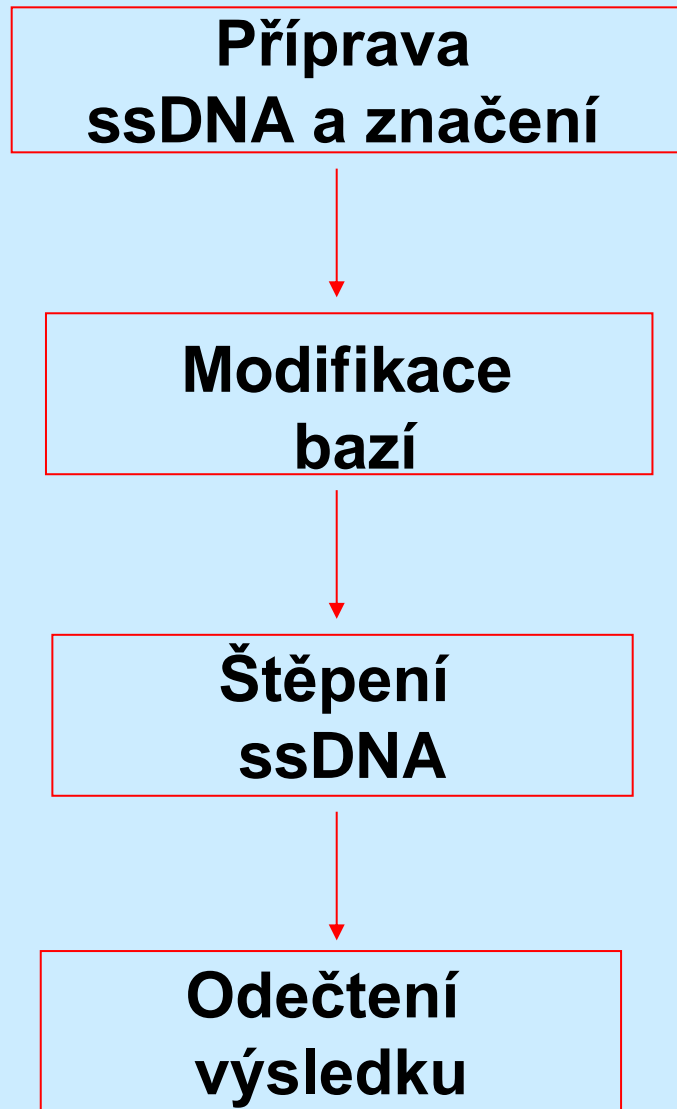
# ***Chemická metoda sekvenování*** ***(Maxamovo-Gilbertovo sekvenování)***

Podstatou je specifické štěpení molekuly **ssDNA** po modifikaci jednotlivých bází chemickými činidly s následnou elektroforetickou detekcí v denaturujícím polyakrylamidovém gelu.

Chemická činidla jsou **specifická** pro modifikaci určitých bází:

<b>G</b>	– DMS
<b>A+G</b>	– piperidin
<b>C+T</b>	– hydrazin
<b>C</b>	– hydrazin + NaCl

# *Chemická metoda sekvenování*





# Chemická metoda sekvenování

Příprava  
ssDNA a značení

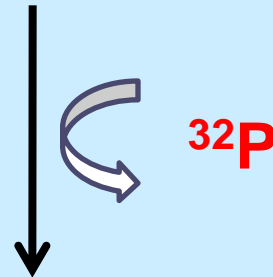


Modifikace  
bází

Asymetrická PCR

Využití vazby  
biotinylovaného primeru

Štěpení  
ssDNA



Odečtení  
výsledku



# Chemická metoda sekvenování

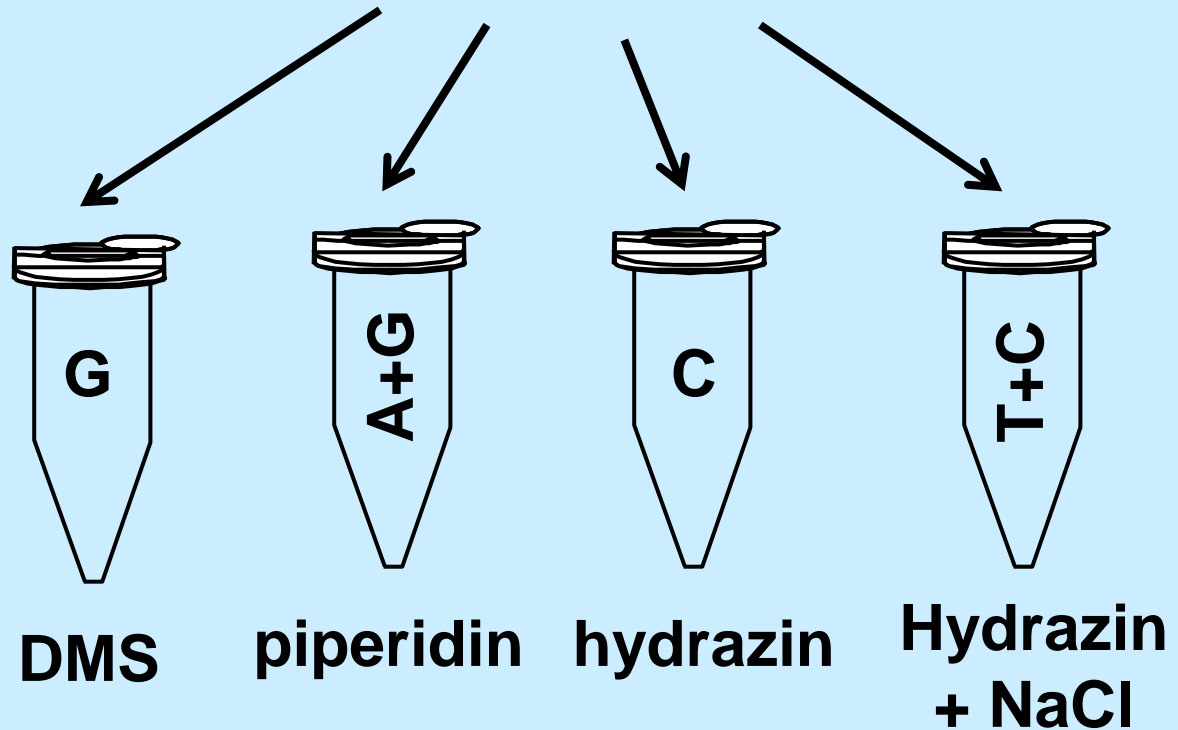
$^{32}\text{P}$  - GATCAGG - 3'

Příprava  
ssDNA a značení

Modifikace  
bází

Štěpení  
ssDNA

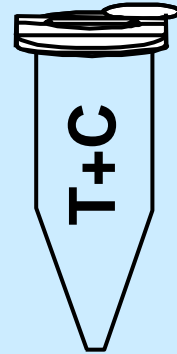
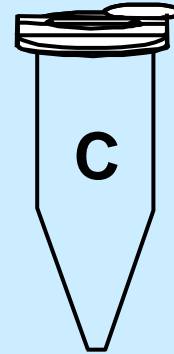
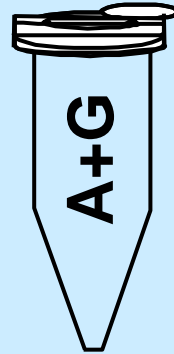
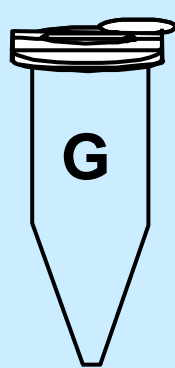
Odečtení  
výsledku



# Chemická metoda sekvenování

$^{32}\text{P}$  - GATCAGG - 3'

Příprava  
ssDNA a značení



Modifikace  
bází

DMS

piperidin

hydrazin

Hydrazin  
+ NaCl

Štěpení  
ssDNA

Štěpení piperidinem při vysoké teplotě

Odečtení  
výsledku

$^{32}\text{P}$  -GATCAG/G

$^{32}\text{P}$  -G/ATCA/G/G

$^{32}\text{P}$  -GATC/AGG

$^{32}\text{P}$  -GAT/C/AGG

**<sup>32</sup>P** - GATCAGG - 3'

**G**

DMS



**<sup>32</sup>P** - GATCAG  
**<sup>32</sup>P** - GATCAGG

**A+G**

piperidin



**<sup>32</sup>P** - GA  
**<sup>32</sup>P** - GATCA  
**<sup>32</sup>P** - GATCAG  
**<sup>32</sup>P** - GATCAGG

**T+C**

hydrazin



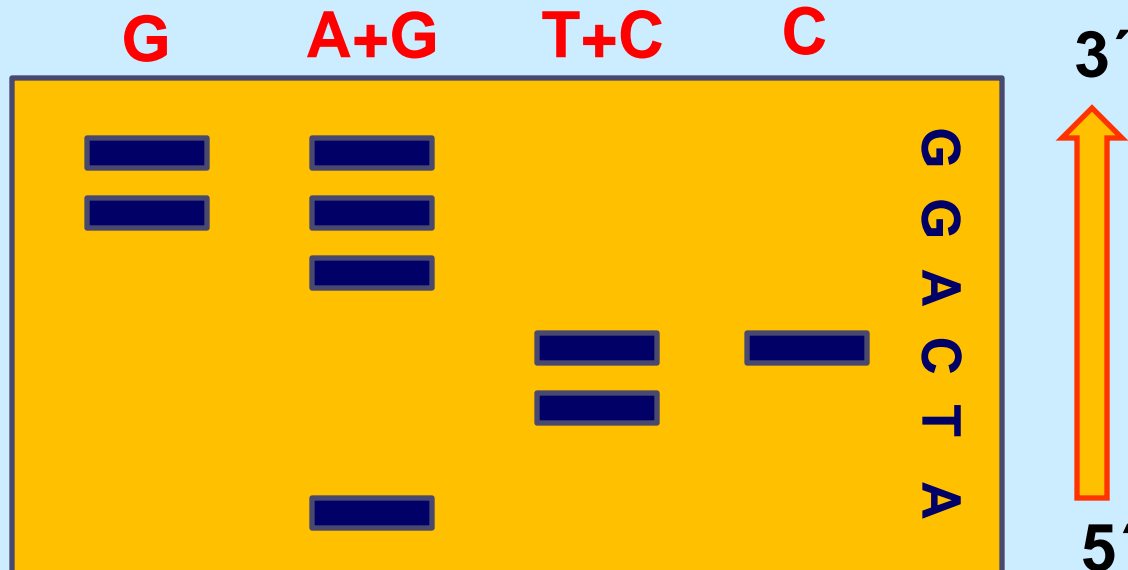
**<sup>32</sup>P** - GAT  
**<sup>32</sup>P** - GATC

**C**

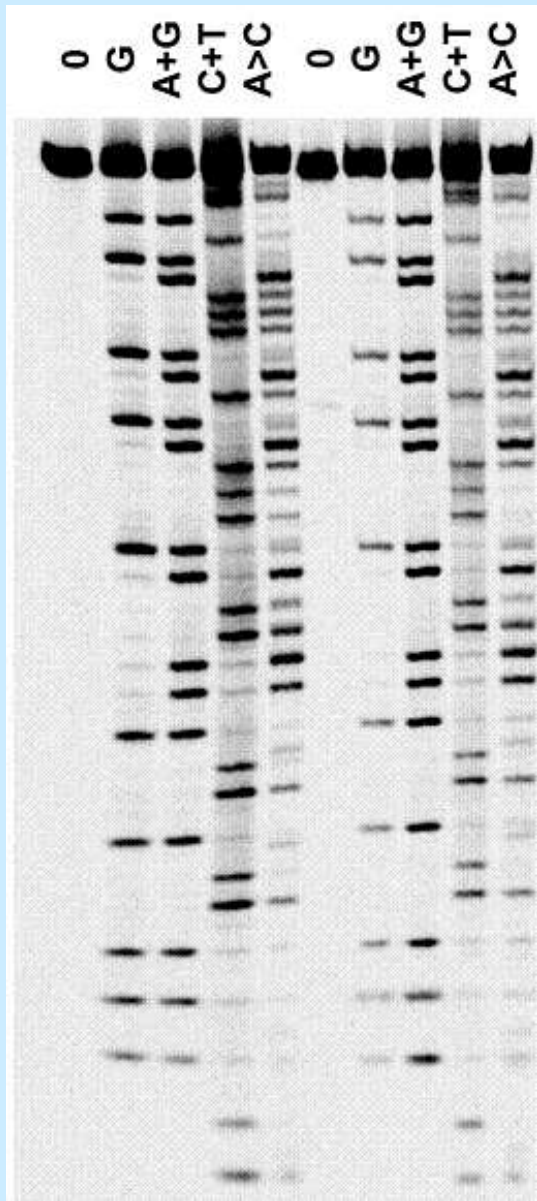
Hydrazin  
+ NaCl



**<sup>32</sup>P** - GATC



# Reálný výsledek chemické metody sekvenování



Převzato z:

Site-specific DNA transesterification catalyzed by a restriction enzyme

Giedrius Sasnauskas\*, Bernard A. Connolly†, Stephen E. Halford‡, and Virginijus Siksnys\*§

\*Institute of Biotechnology, Graiciuno 8, Vilnius, LT-02241, Lithuania; †Institute for Cell and Molecular Biosciences, University of Newcastle, Newcastle upon Tyne NE2 4HH, United Kingdom; and ‡Department of Biochemistry, School of Medical Sciences, University of Bristol, University Walk, Bristol BS8 1TD, United Kingdom

# Úkol



Z výše uvedeného záznamu odečtěte výslednou nukleotidovou sekvenci

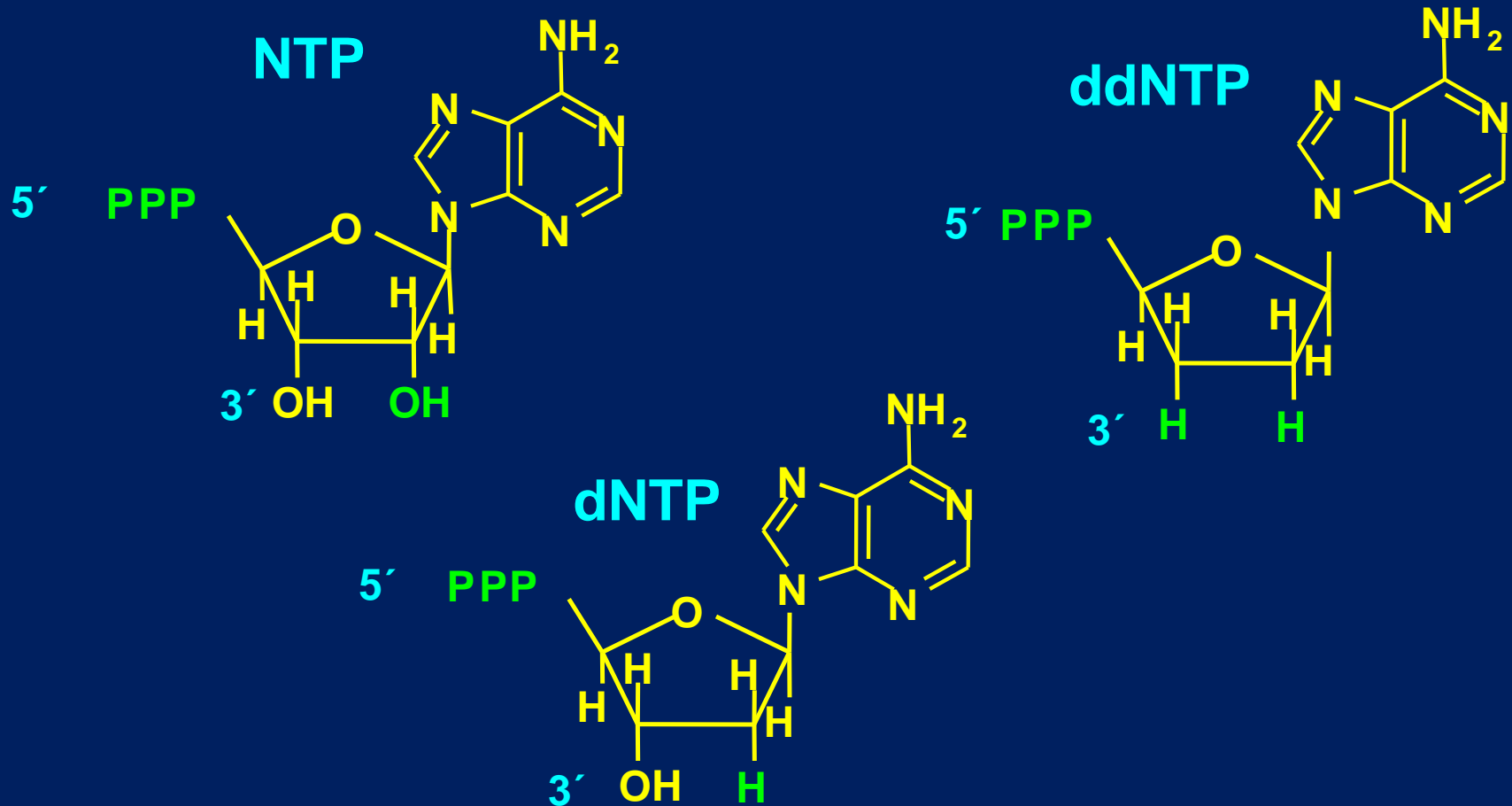
Výsledek bude tedy asi tento

CTg ggC TgC TgA ACC AgC CCA gCA  
gCC Cag Tg

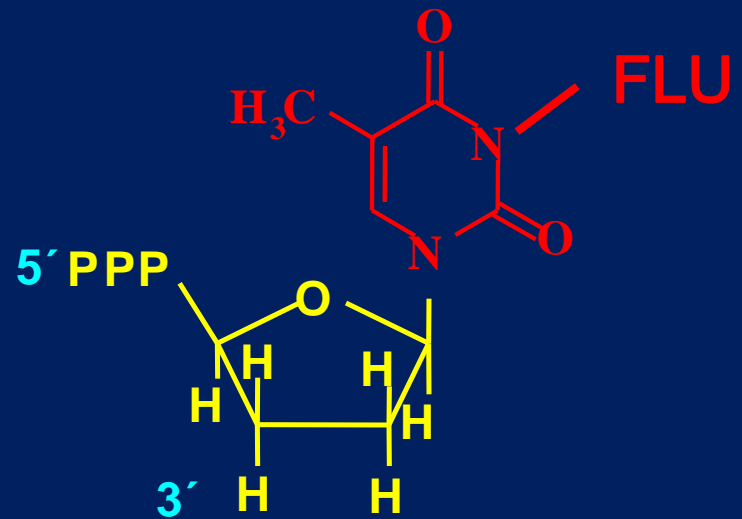
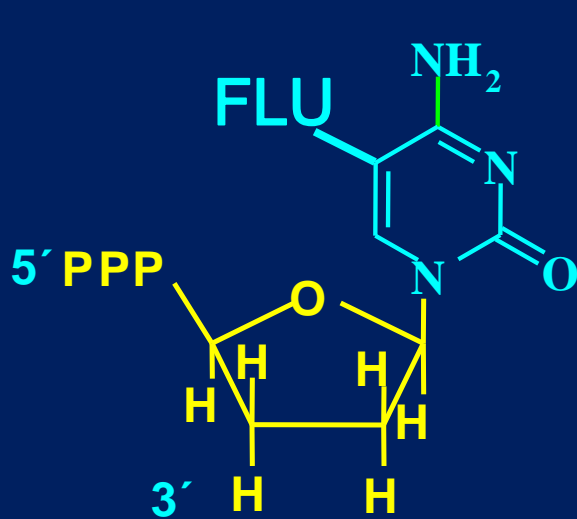
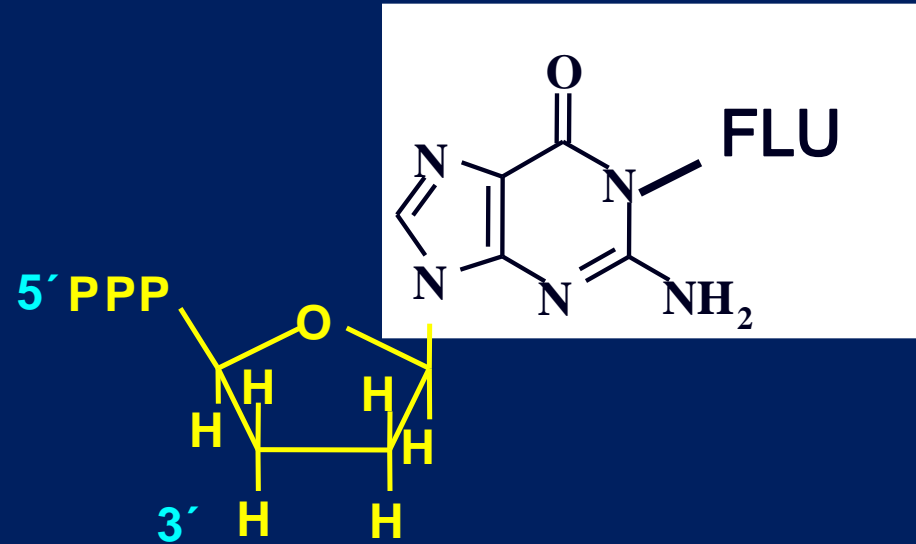
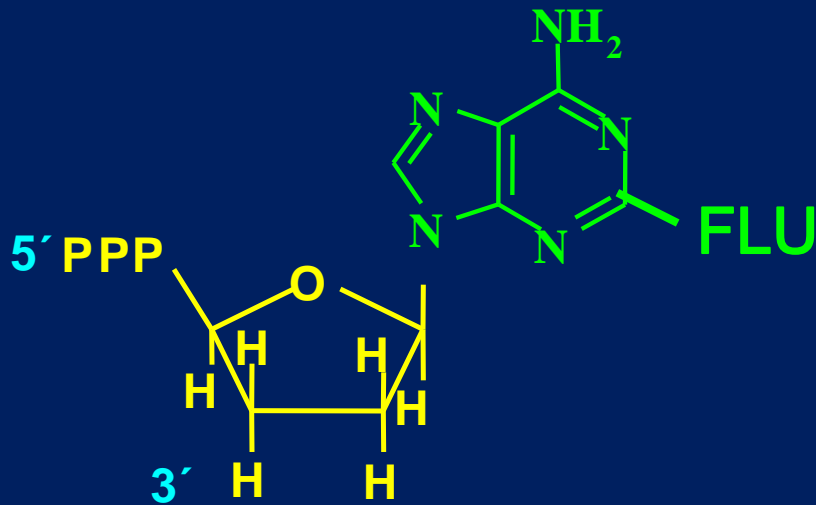


# Sangerova metoda

## dideoxyterminátory



# Dideoxyterminátory





# *Průběh sekvenování*

1. denaturace (92-96°C)



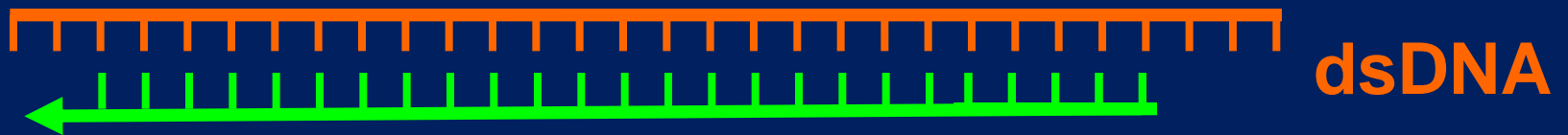
2. annealing (45-72°C)

3. extenze (72°C)



# *Průběh sekvenování*

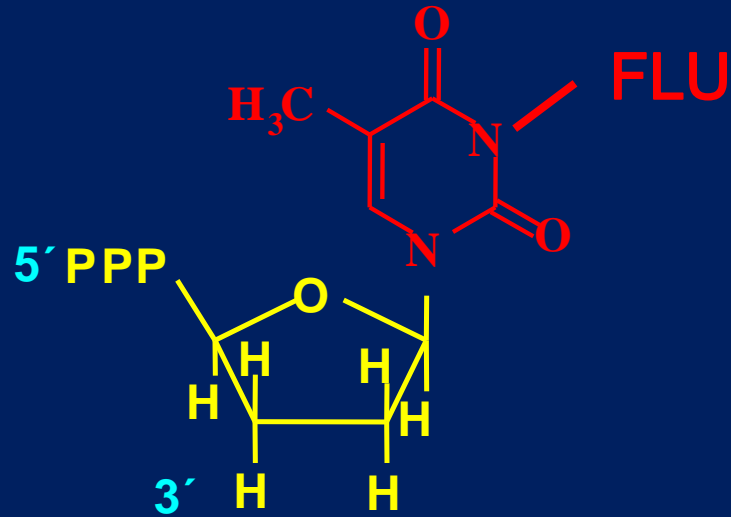
1. denaturace (92-96°C)



2. annealing (45-72°C)

3. extenze (72°C)

# Zařazování ddTTP



TTTTTGTCAAATCGGTGTAAACGGT

TTTTTGTCAAATCGGTGT

TTTTTGTCAAATCGGT

TTTTTGTCAAA

ATGCTACGTTTAGCCACATTTGCCATATGAACCT



# Zařazování dalších ddNTP

TTTTTGTCAAATCGGTGTA

TTTTTGTCAA

TTTTTGTCAA

TTTTTGTCA

ATGCTACGTTTAGCCACATTTGCCATATGAACCT



TTTTTGTCAAATCGGTGTAAC

TTTTTGTCAAATCC

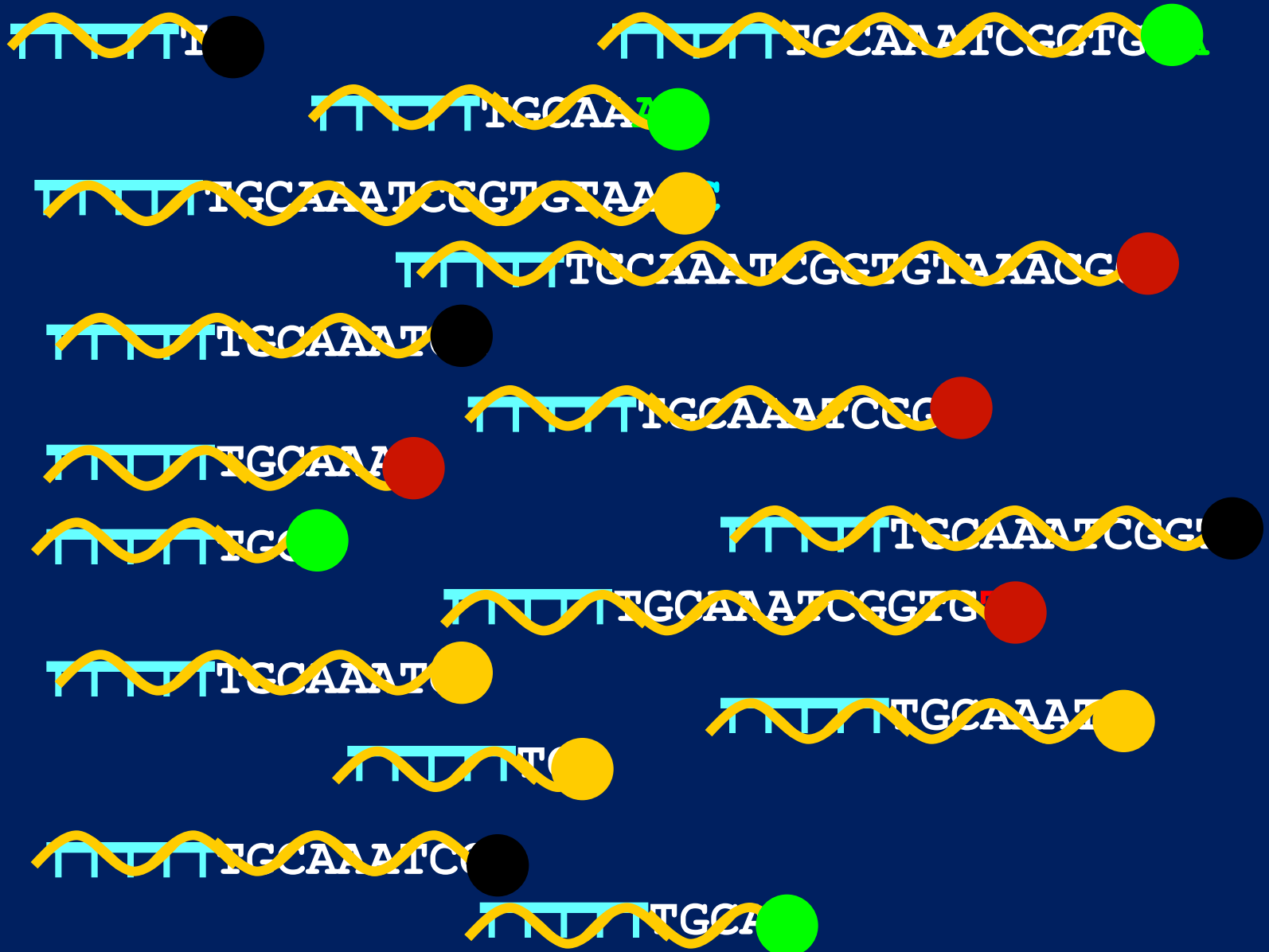
TTTTTGTCAAATC

TTTTTGTCA

ATGCTACGTTTAGCCACATTTGCCATATGAACCT



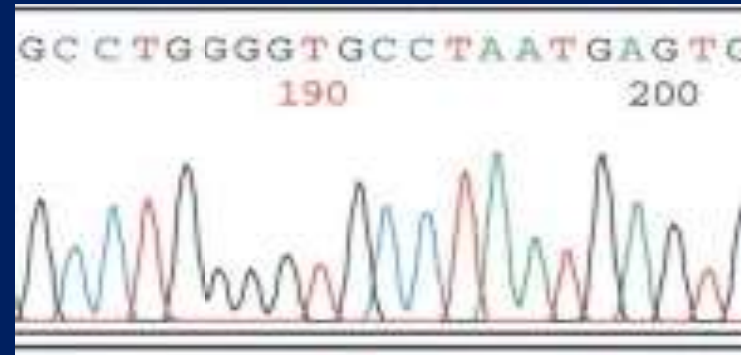
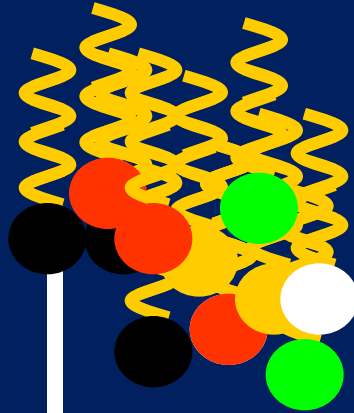
# Výsledek sekvenování



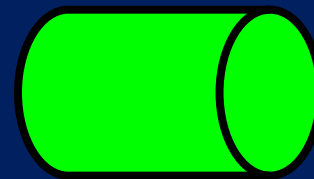
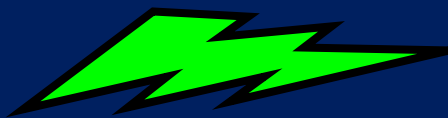
# *Následuje rozdělení fragmentů*



# Následuje rozdělení fragmentů

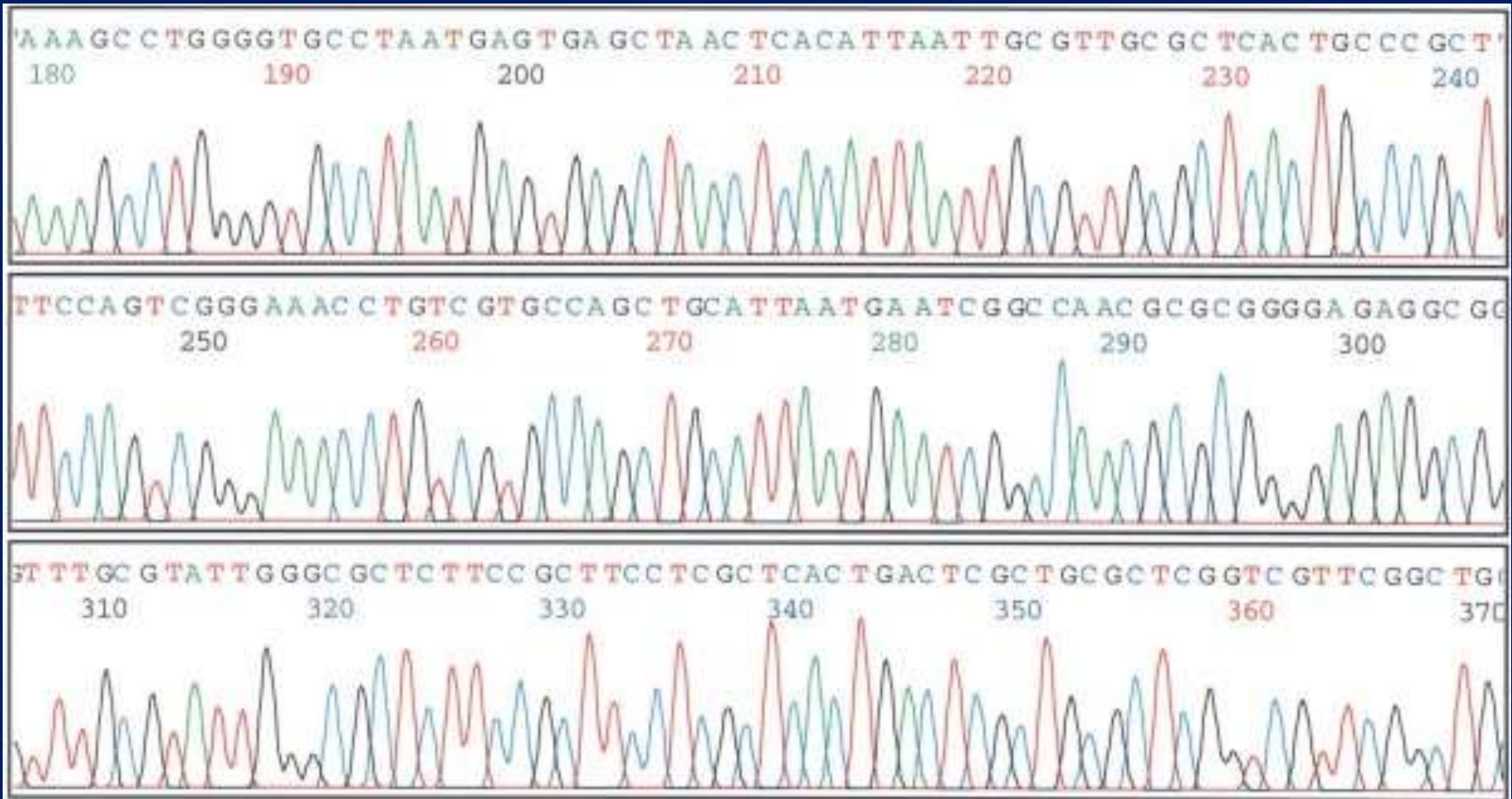


LASER



DETEKTOR

# Sekvenování - záznam





# *Kapilární gelová elektroforéza*



- **rozdělí produkty sekvenování podle velikosti**
- **detekuje fluorofory laserem**

# Úkol

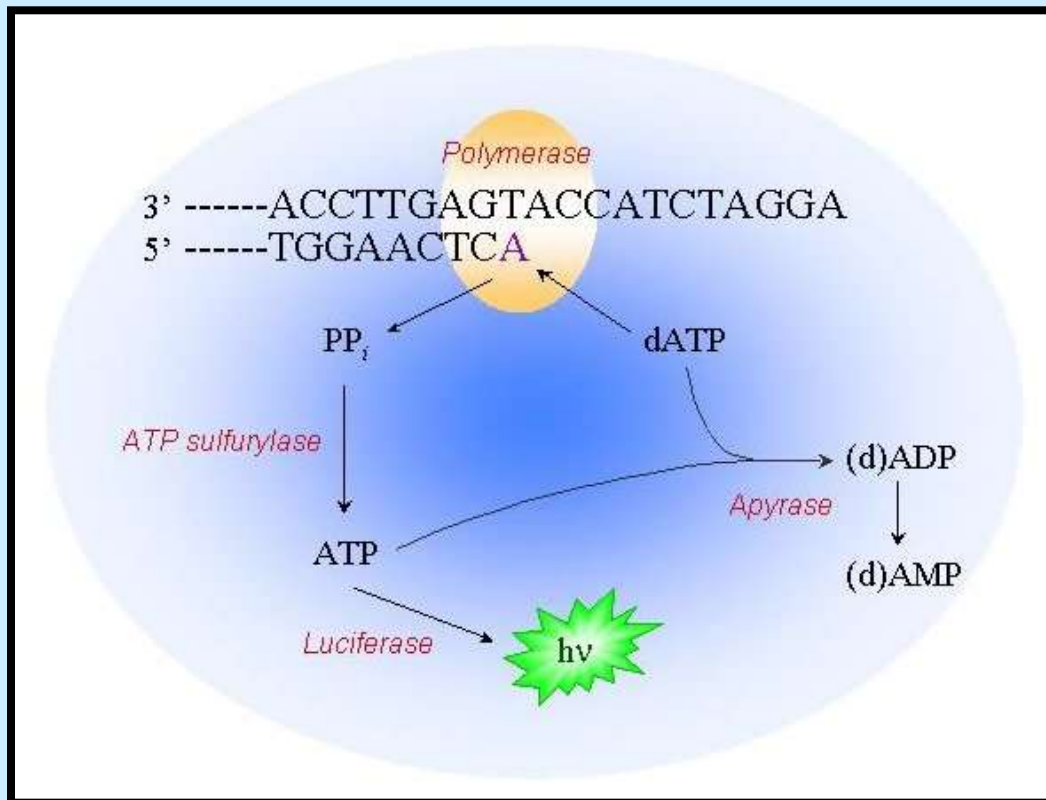


Na základě výsledků sekvenování  
zařad'te izoláty bakterií čeledi  
*Pasteurellaceae*

Použijte výukový materiál

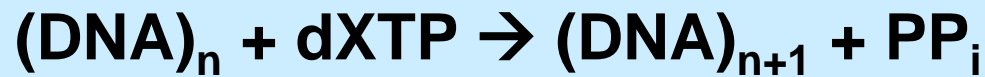
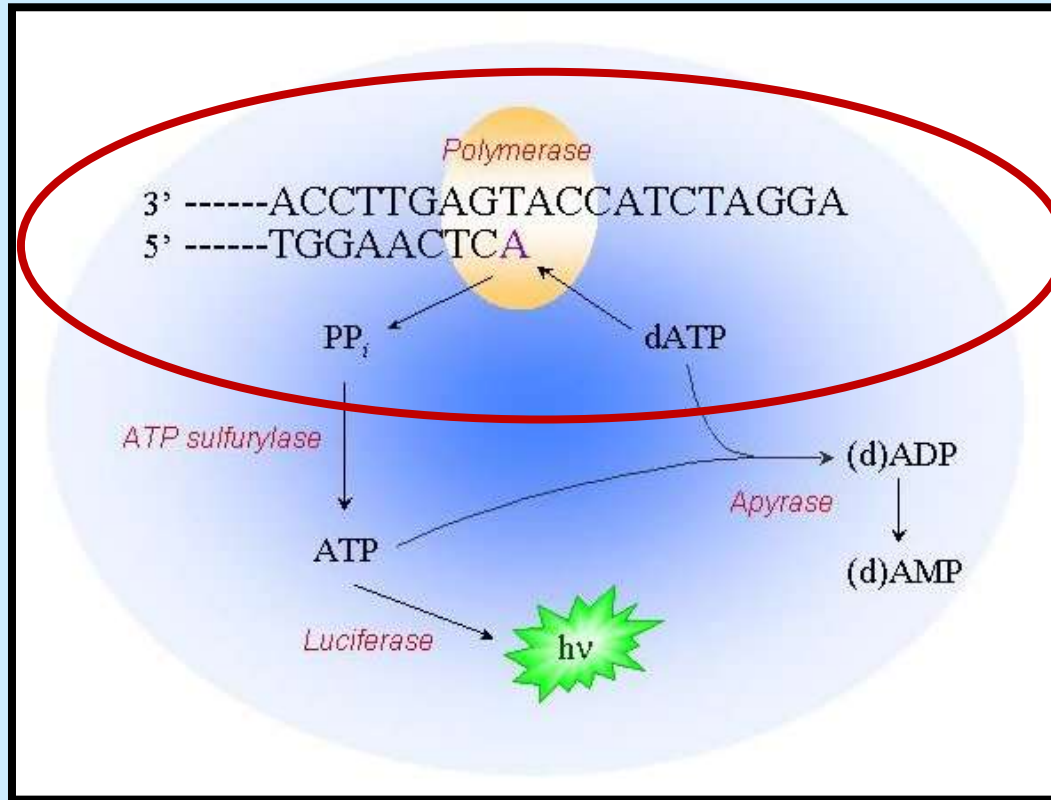
# Pyrosekvenování

- Metoda popsaná Mostafa Ronaghi et al. v roce 1990
- Určena pro krátké úseky DNA, SNP a úseky metylované



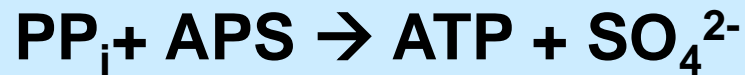
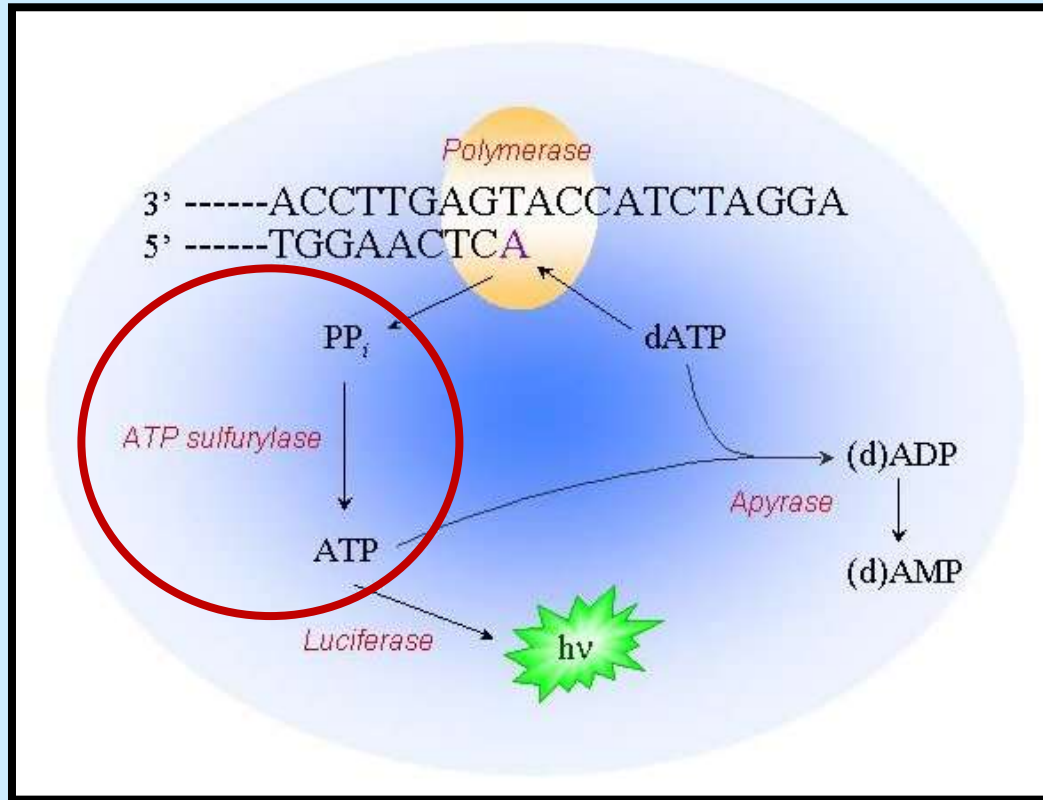
- 4 enzymy
- velmi přesná
- reakce se záhy „zahltí“

# Průběh pyrosekvenování



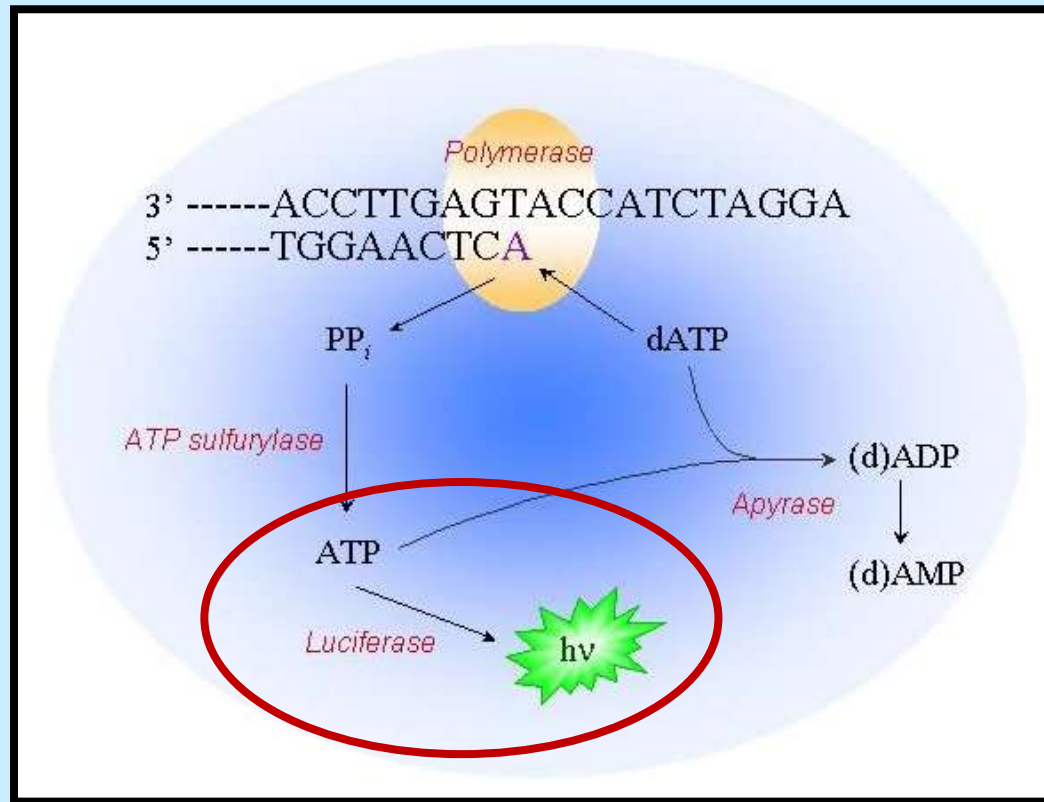
DNA polymeráza

# Průběh pyrosekvenování



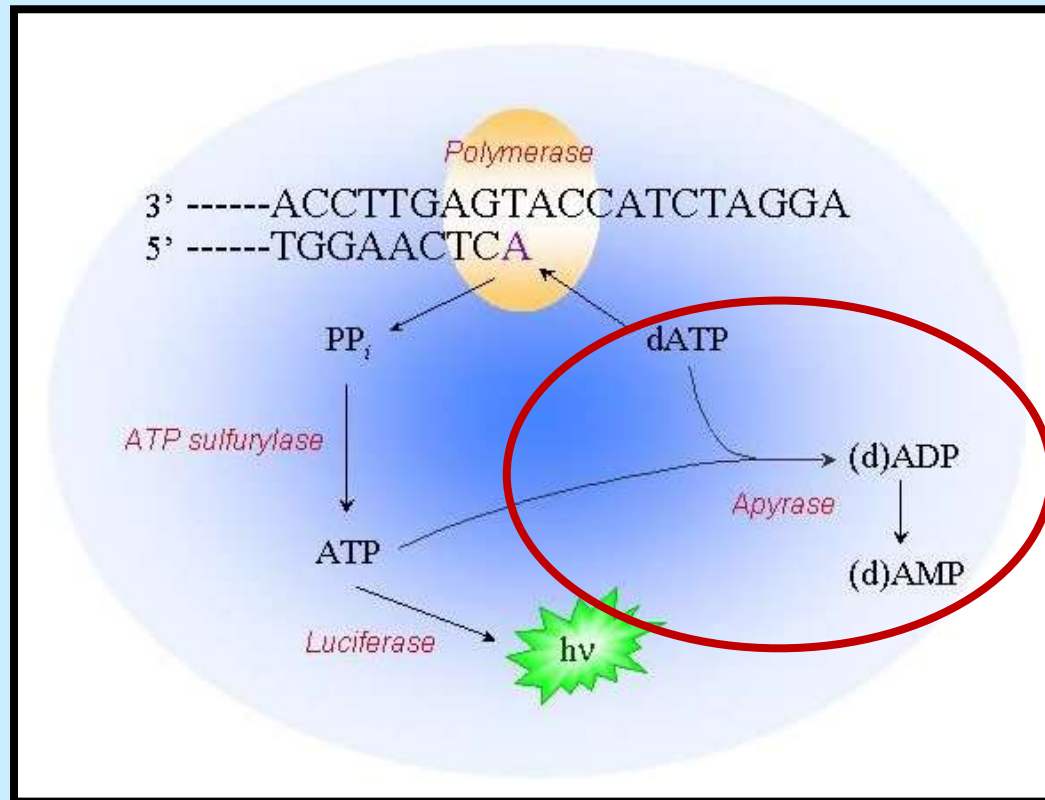
ATP sulfuryláza

# Průběh pyrosekvenování



Luciferáza

# Průběh pyrosekvenování



**dXTP → dXMP**

**apyráza**

# Parametry pyrosekvenování

➤ žádné značené primery ani značené nukleotidy

➤ žádná elektroforéza

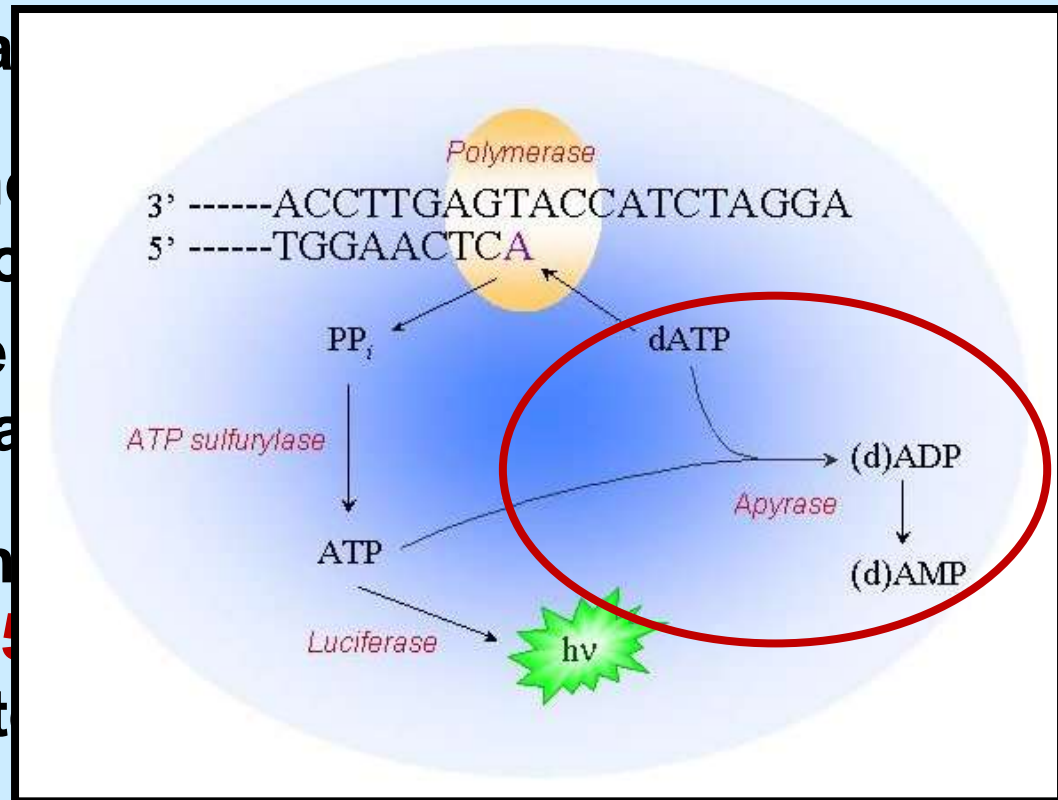
➤ Principem je uvolnění  
a emise viditelného světla

➤ Množství uvolněné  
množství zabudované

➤ Namísto standardní  
**2'-deoxyadenosin-5'**  
který není substrát

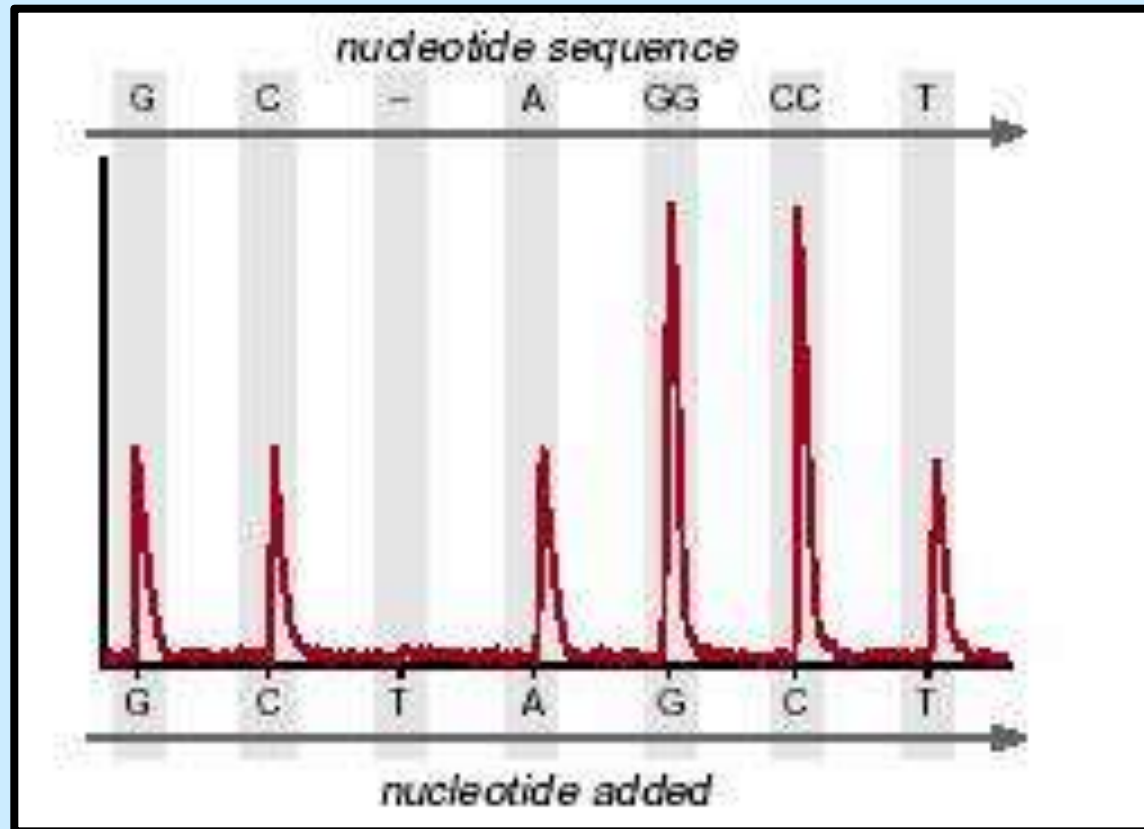
➤ Rychlost sekvenování = 1 báze za minutu

➤ Délka stanovené sekvence = 100 nukleotidů





# Pyrosekvenování - záznam



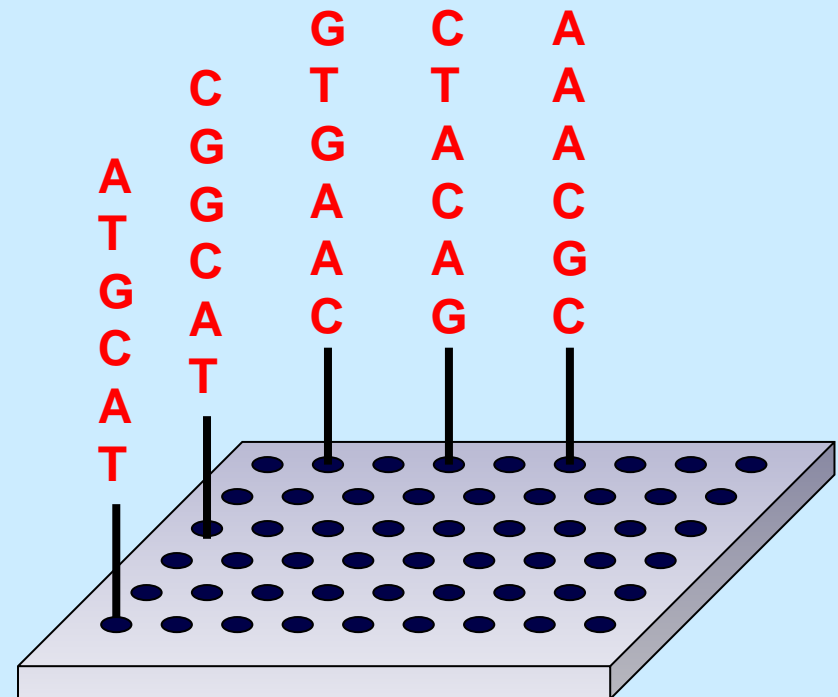
Intenzita signálu odpovídá počtu začleněných nukleotidů

# *Sekvenování pomocí hybridizace*

- **Nepřímá metoda využívající DNA čipů**
- **Výsledek odečten konfokálním mikroskopem**
- **Sekvence odečtena dedukcí**



T  
A  
C  
G  
T  
A



# Otázka



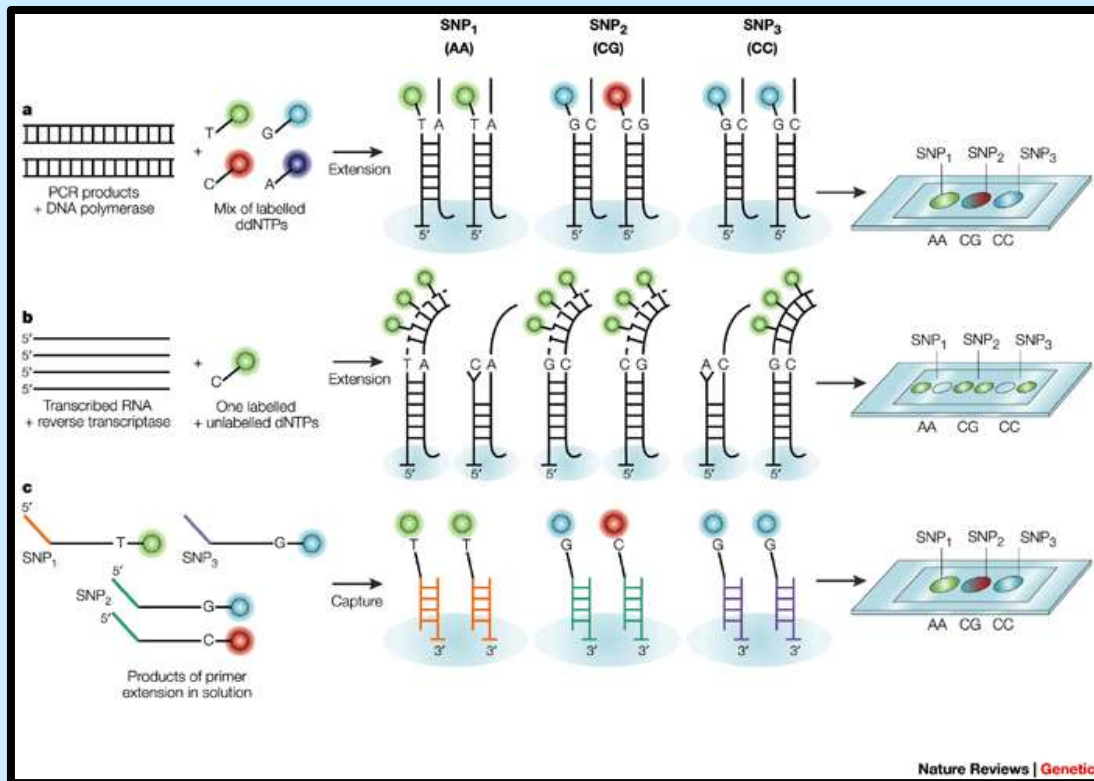
Kdyby byly na čipu naneseny 20 mery, kolik políček by musel obsahovat čip pro sekvenování 1 Mb?

Prý celkem  $1,112 \times 10^6$



# Minisekvenování

- Metoda využívající technologie DNA čipu
- Využívá se pro stanovení jednoduchých nukleotidových polymorfismů
- Více variant, metoda je automatizovaná, multiplexní



# ***Next generation sequencing (NGS) methods***

***(Vysokokapacitní  
sekvenování)***

# Základní pojmy

- **Read** (čtení) - souvislá DNA sekvence produkovaná sekvenačním procesem (sekvenátorem)
- **Coverage** (pokrytí) - počet vzájemně se překrývajících čtení v dané oblasti genomu (kolikrát je přečten příslušný úsek nebo baze)
- **Konvenční sekvence** (consensus sequence) – idealizovaná sekvence, ve které je každá pozice reprezentována nejčastěji nacházenou bází při porovnávání více různých sekvencí dané oblastí
- **Contig** (kontig) – sada překrývajících se segmentů DNA (čtení) tvořících souvislou (nepřetržitou) konvenční sekvenci
- **Assembly** (skládání) – zarovnávání a překrývání DNA sekvencí (čtení, kontigů) za účelem rekonstrukce původní sekvence DNA molekuly v plné délce
- **Scaffold** – sada propojených nesouvislých (přetržitých) sérií genomových sekvencí skládajících se z kontigů oddělených mezerami (gaps) známých délek

# ***Charakteristika NGS***

- **Moderní, vysokokapacitní (high-throughput) technologie sekvenování DNA**
- **Masivní, paralelní, rychlé ...**
- **Snižující se cena, čas, složitost protokolů, četnost chyb ...**
- **Zvyšující se množství a kvalita dat, počet a délka čtení (kapacita uložení dat), repertoár potřebných bioinformatických metod**
- **Široká škála aplikací**



**Čím se liší „Next generation sequencing“ od „Third generation sequencing“ (TGC)?**

**Sekvenování z jedné molekuly, v reálném čase,  
*in situ*, méně (bio)chemie**



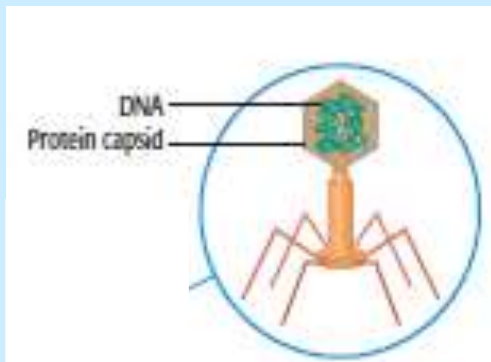


# Vstupní materiál pro genomiku

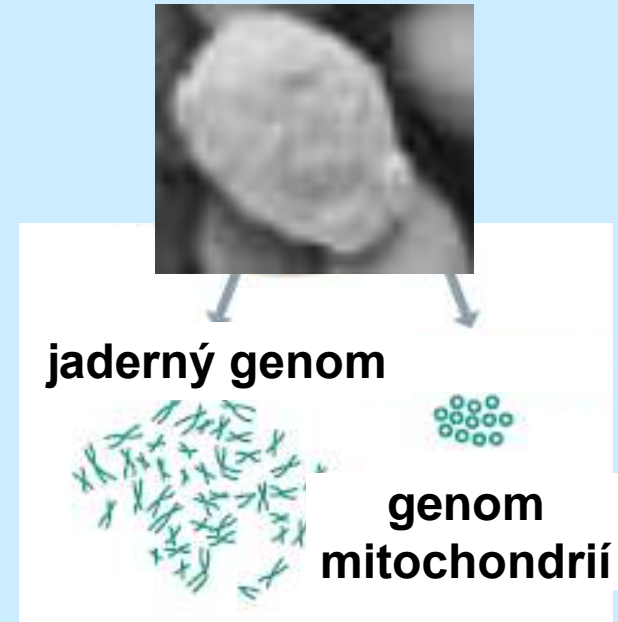
## DNA

- *De novo* sekvenování genomu
- Resekvenování (ChIP-Seq)
- Sekvenování ampliconů (16S)
- Sequence capture
- Detekce modifikace bází
- Variace genomové informace

## virová



## eukaryotická



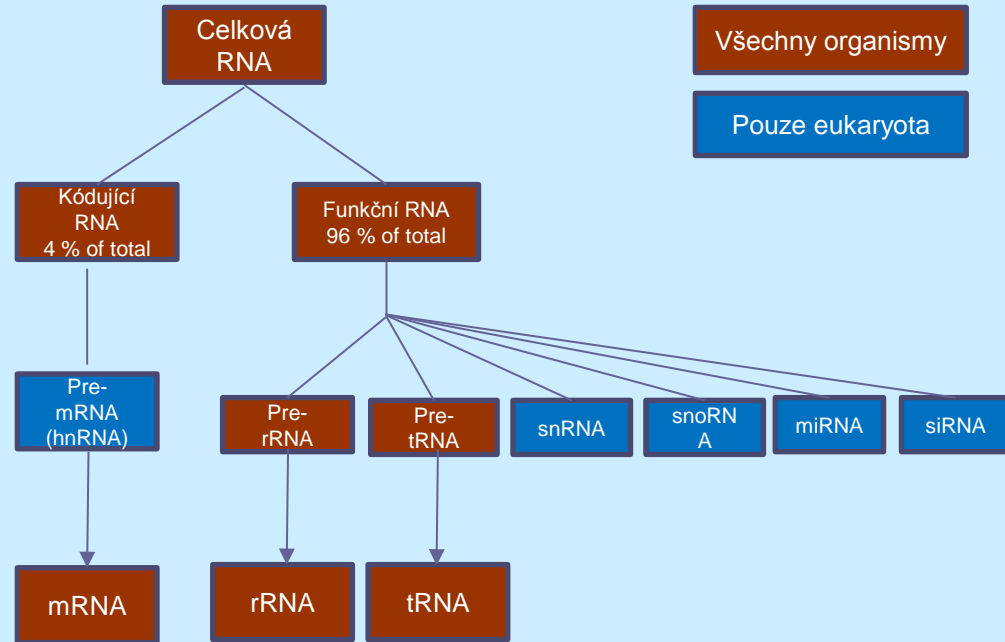
## prokaryotická

chromozóm plasmidy



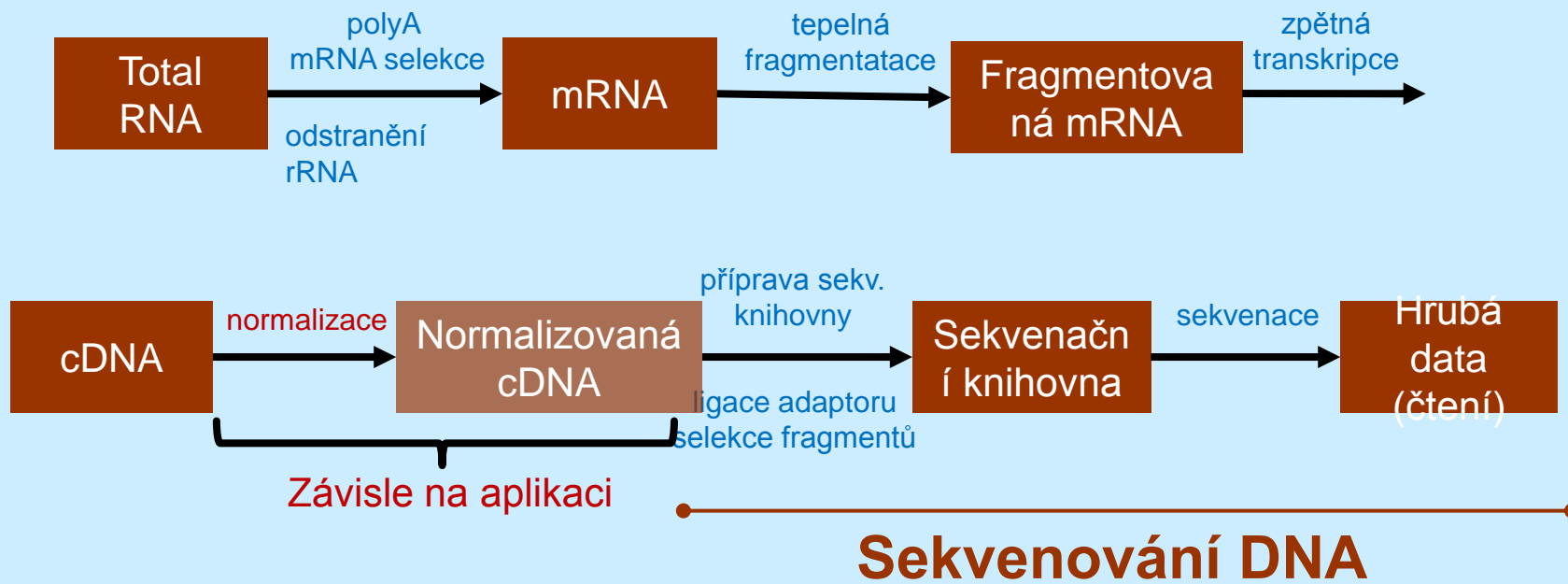
# Vstupní materiál pro transkriptomiku

## RNA



- Sekvenování mRNA
- (Whole Transcriptome Shotgun Seq – WTSS)
- Detekce polymorfismů (SNP)
- Sekvenování dalších typů RNA (miRNA)
- Kvantitativní expresní analýza

# Proces sekvenování mRNA

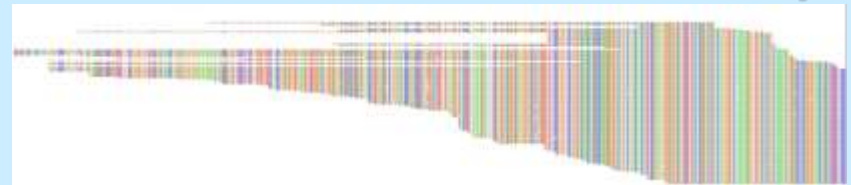


# Kvalita RNA

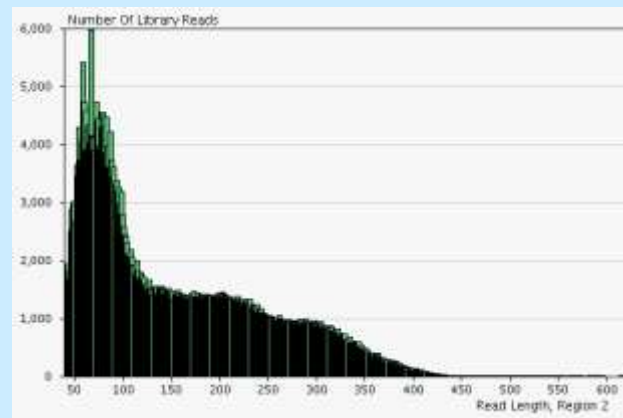
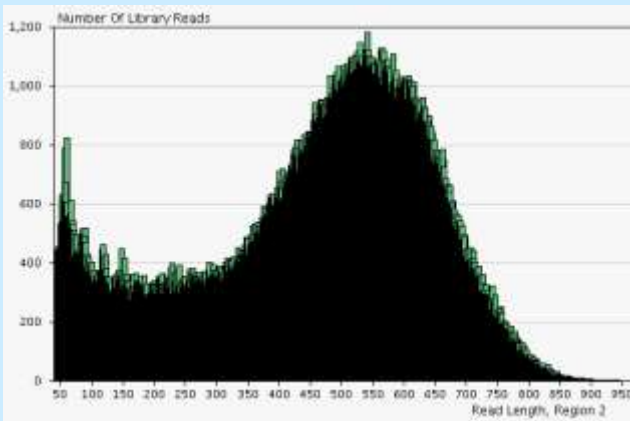
- Kvalita vstupní celkové (total) RNA - RNA integrity number (RIN)
- $RIN < 7 \Rightarrow$  nerovnoměrná distribuce čtení podél transkriptu (mezi 5' a 3'- konci)



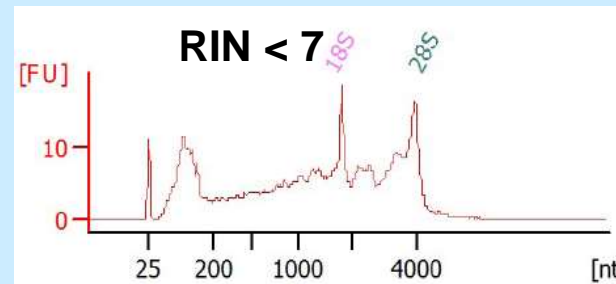
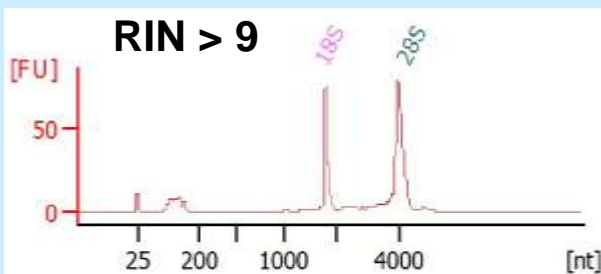
$\Rightarrow$  snížená kvalita sekvenačních dat



Number of reads



Distribuce délek čtení u 454 technologie



Záznam kvality RNA Agilent Bioanalyzer

# Syntéza cDNA

Celková RNA (ug)

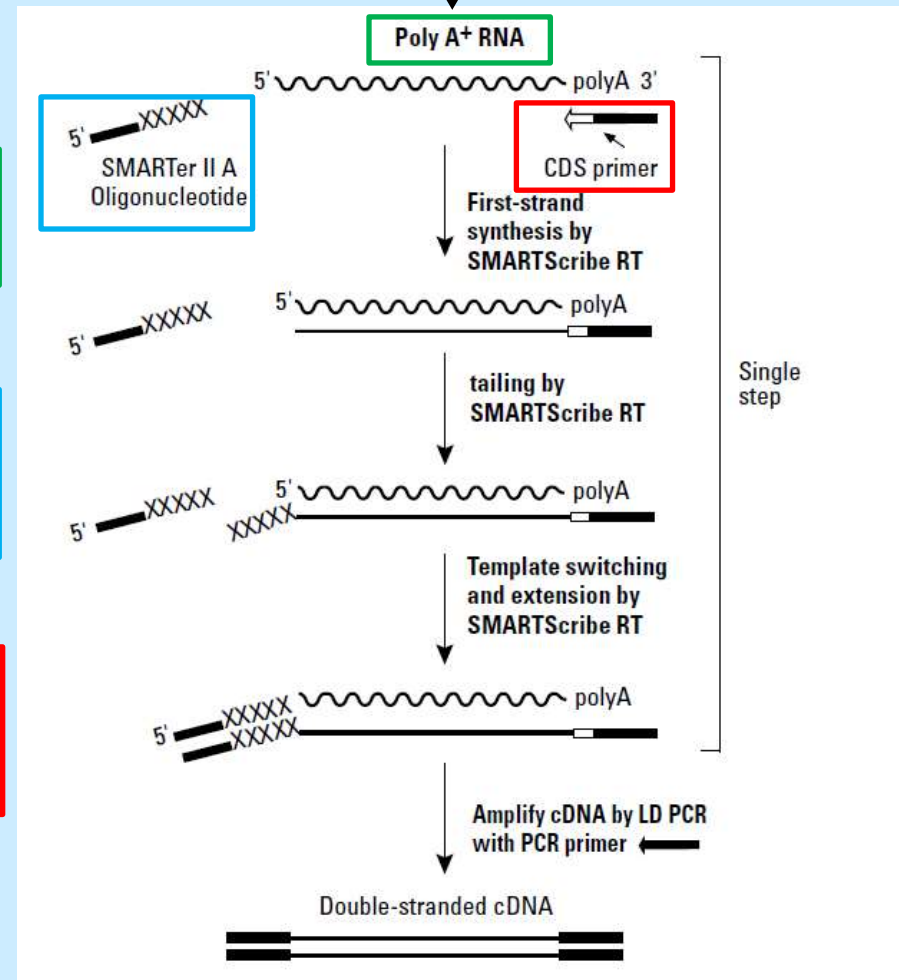
mRNA with polyA 3'end

**SMARTer II A Oligo:**

5'-AAGCAGTGGTATCAACGCAGAGTACGCGGG-3'

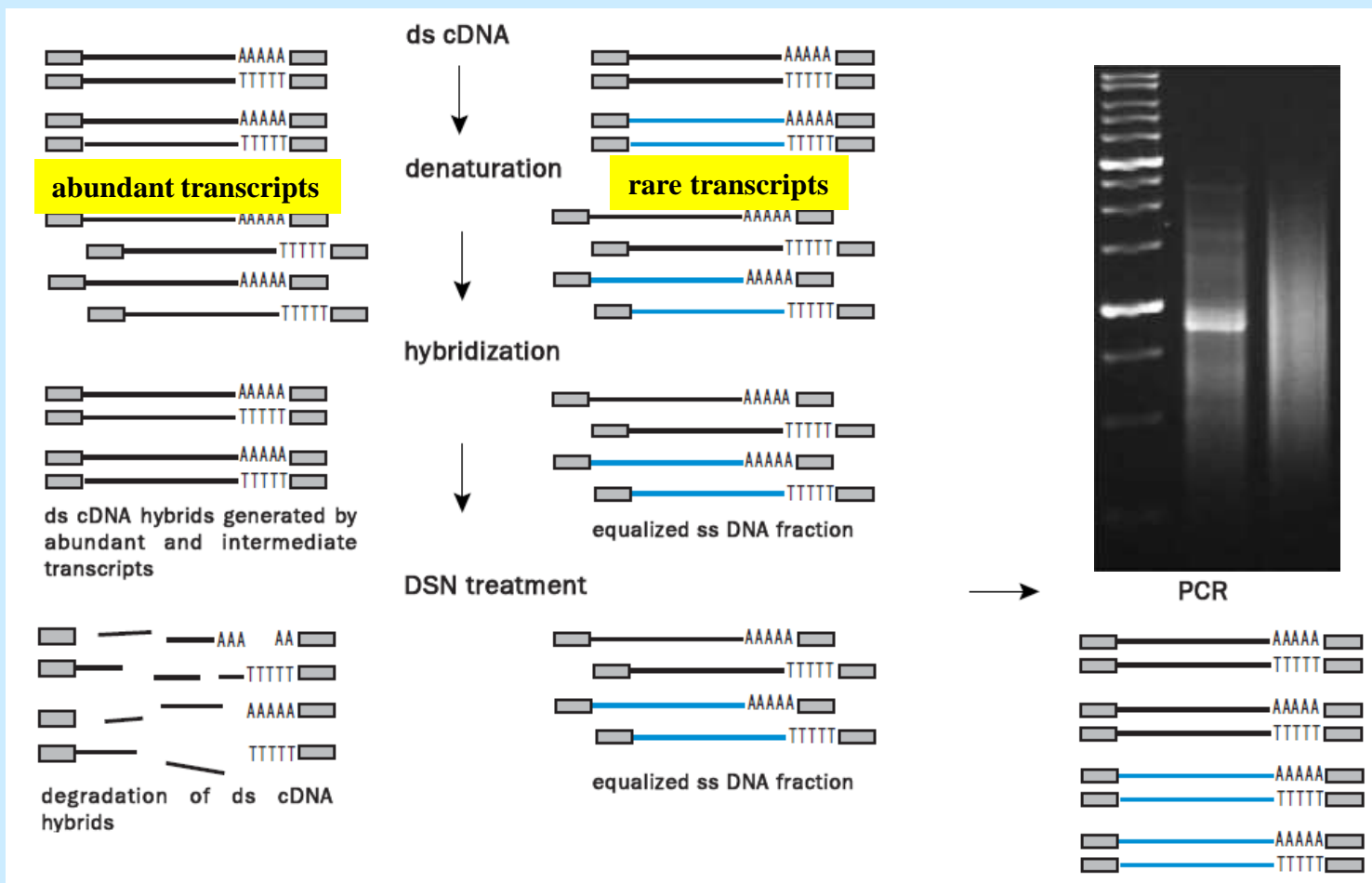
**Modified CDS Primer:**

5'-AAGCAGTGGTATCAACGCAGAGTTTTTGTTCCTTTTTTTTTTNN-3'



# Normalizace cDNA

- Za účelem získu vyrovnaného počtu čtení všech transkriptů nezávisle na míře jejich exprese
- Analýza méně četných transkriptů (rare transcripts)



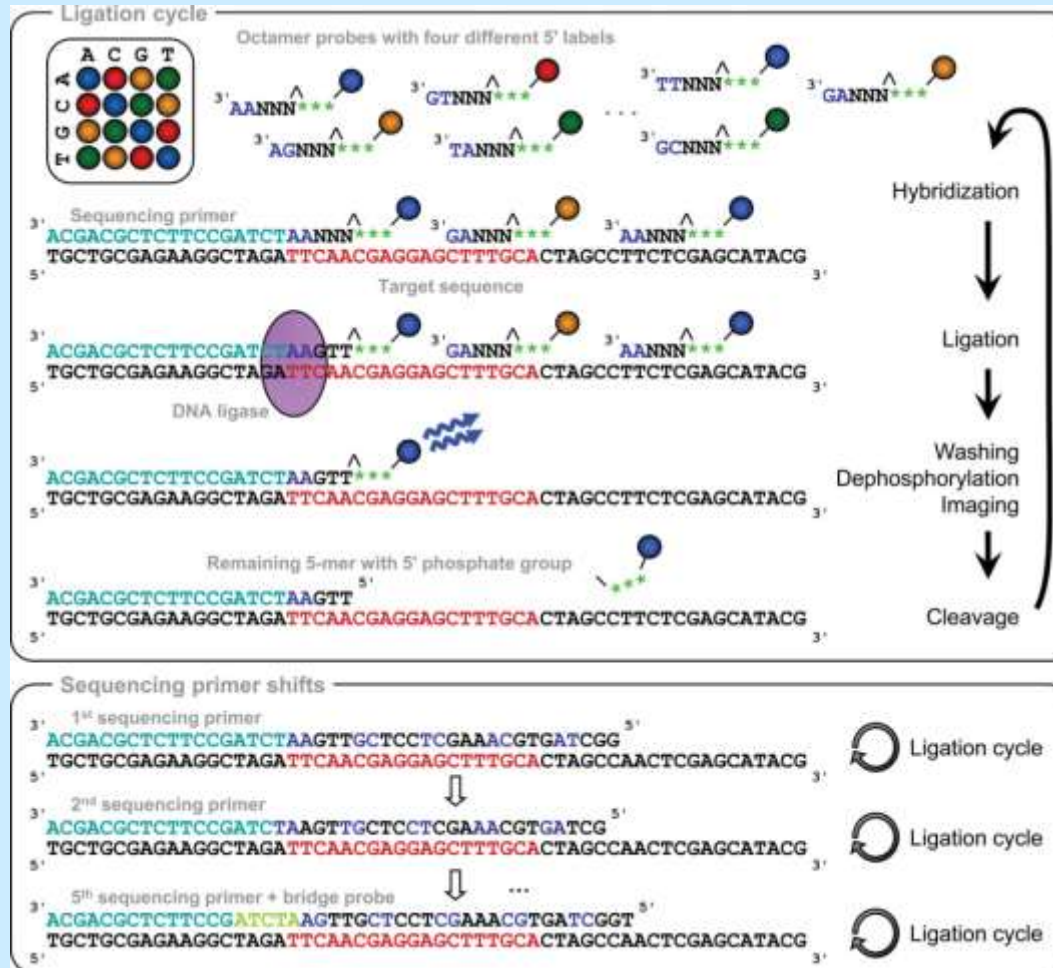
# *Současné sekvenační platformy*

- Roche/454 (GS FLX+/GS Junior)
- Illumina Genome Analyzer (HiSeq/MiSeq/NextSeq)
- Life Technologies (3500 Genetic Analyzer, Ion Torrent Proton/PGM)
- Pacific Biosciences (PACBIO RSII)
- Applied Biosystems (SOLiD, 3730x/ DNA Analyzer )



# Sekvenování ligací

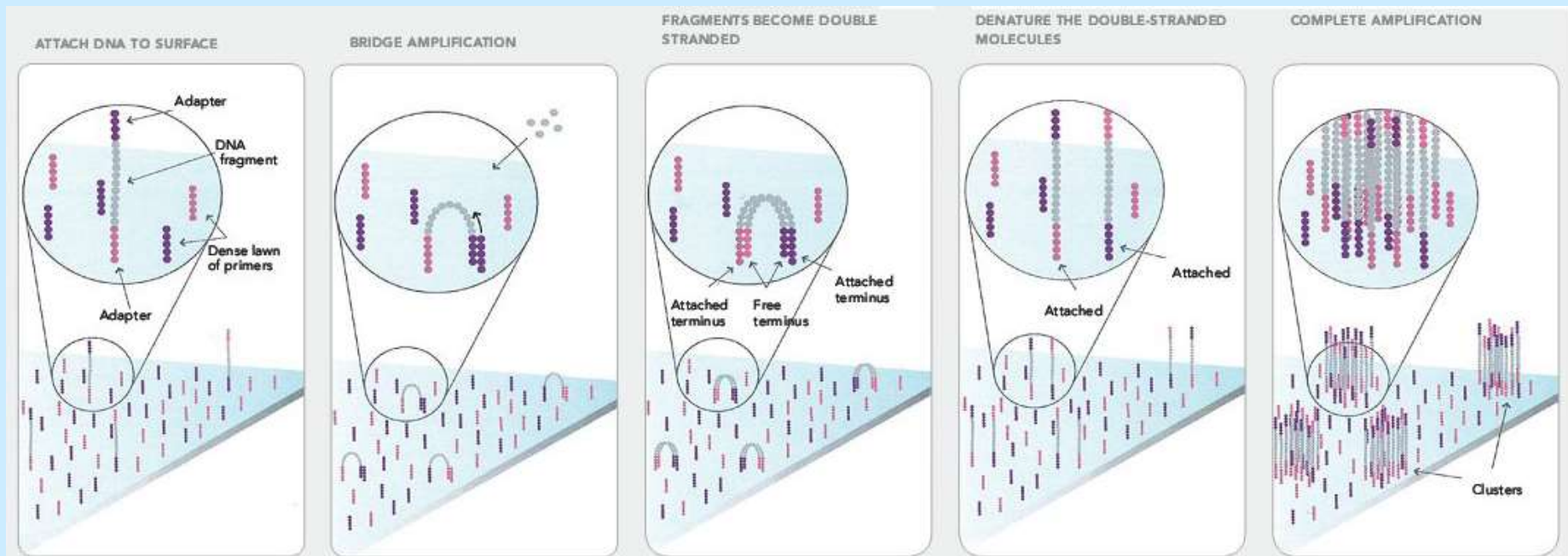
## Oligo Ligation Detection (SOLiD)





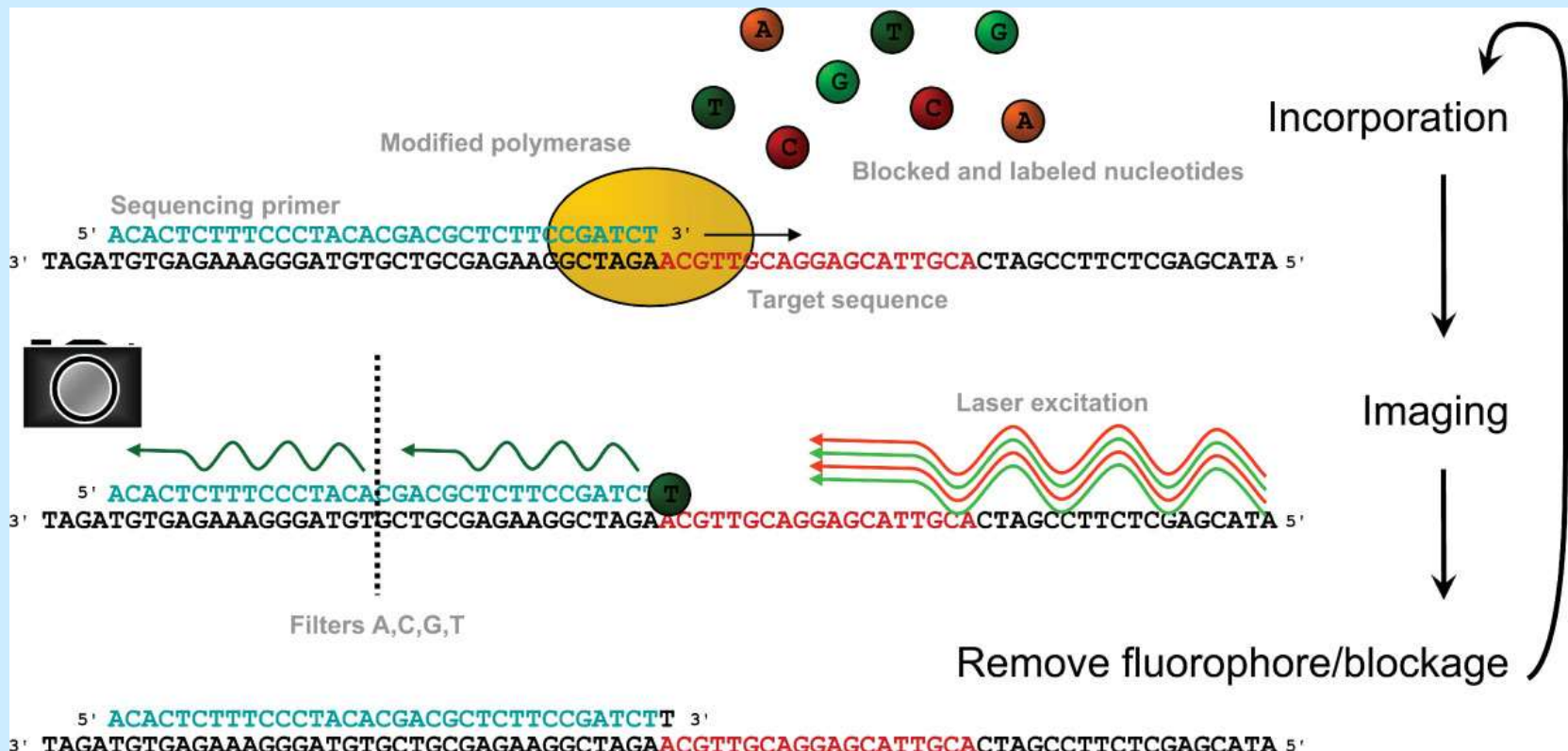
# Reversible Terminator (HiSeq, MiSeq, NextSeq)

## Generování klastrů na povrchu sekvenační flow-celly



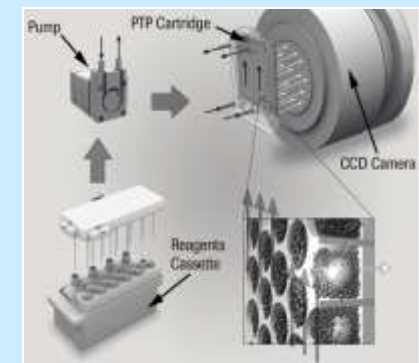
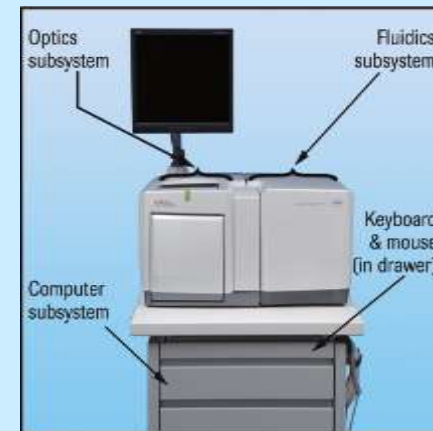
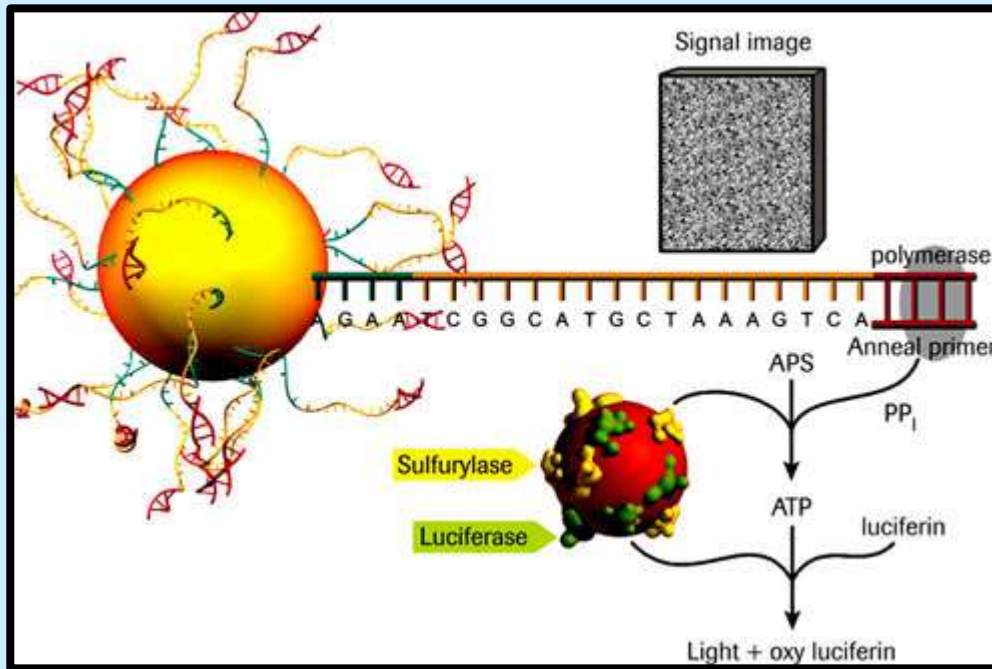
# Reversible Terminator (HiSeq, MiSeq, NextSeq)

## Sekvenování syntézou



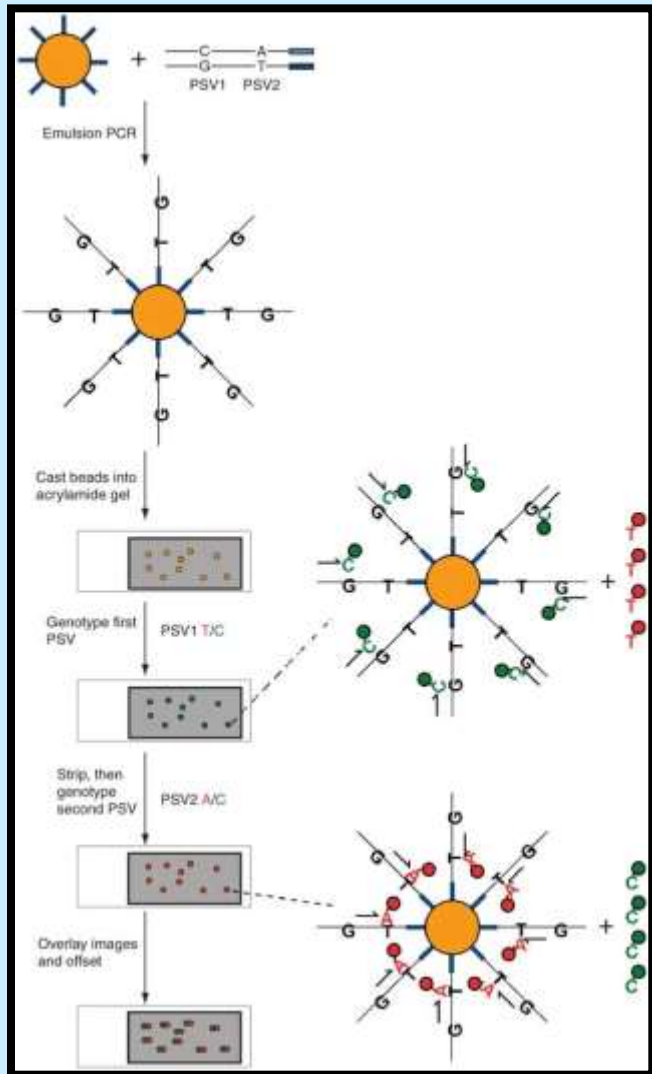
# Pyrosekvenování (GS FLX, GS Junior)

## 454 sekvenční systém firmy Roche



- Sekvenují se produkty emulsní PCR

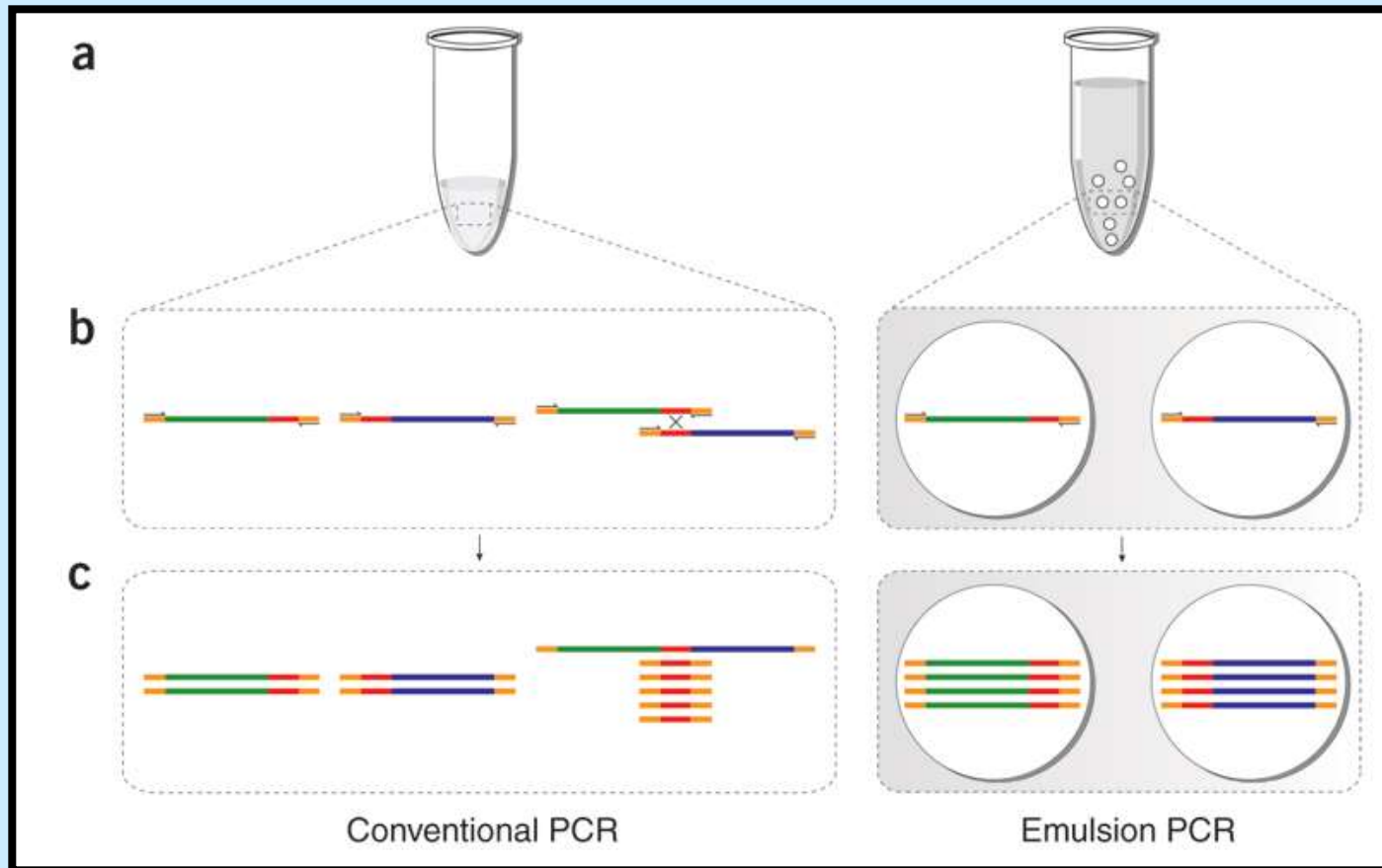
# Emulsní PCR



- Amplifikace probíhá v emulsi (voda-olej)
- DNA je fragmentována na úseky dlouhé 300-800 bp
- Jednotlivé matrice jsou navázány samostatně na povrch jednotlivých kapek obalených primery

# Emulsní PCR

- Amplikon je výsledkem jedinečné PCR
- Méně nespecifických produktů



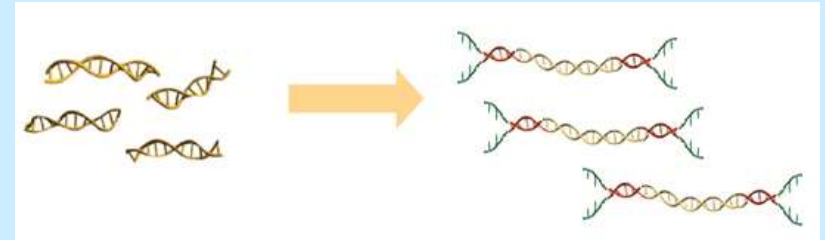
# Proces přípravy vzorku pro 454/Roche

## Příprava sekvenační knihovny

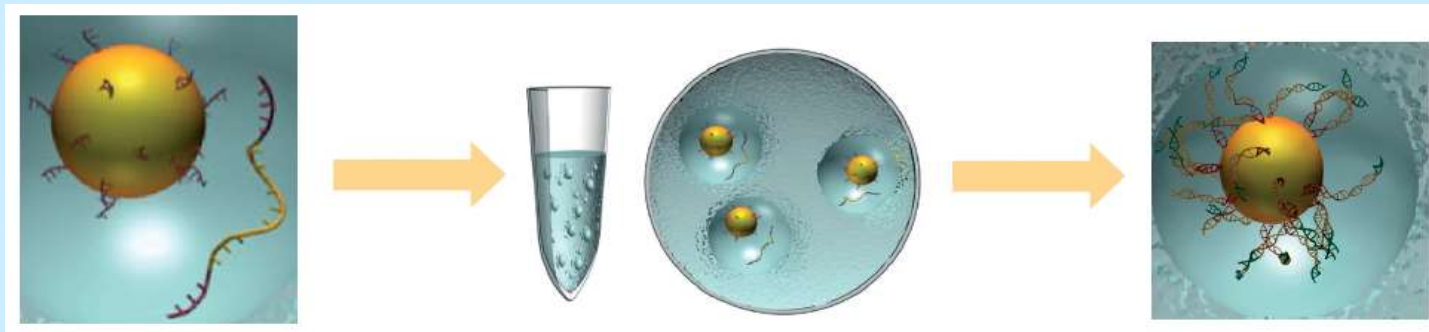
### Fragmentace



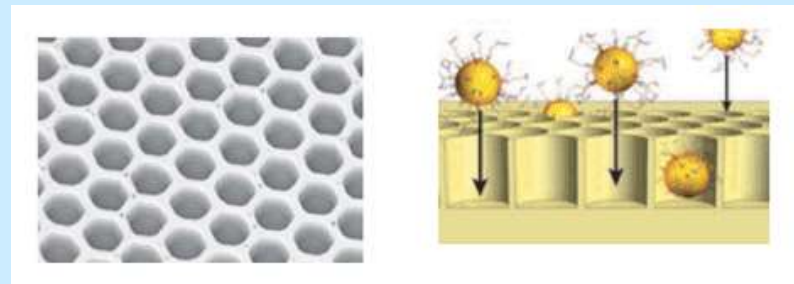
### Ligace adaptoru



### Emulzní PCR - amplifikace



### Nanášení kuliček na sekvenační matici (PTP)



# ***454 sekvenční systém***

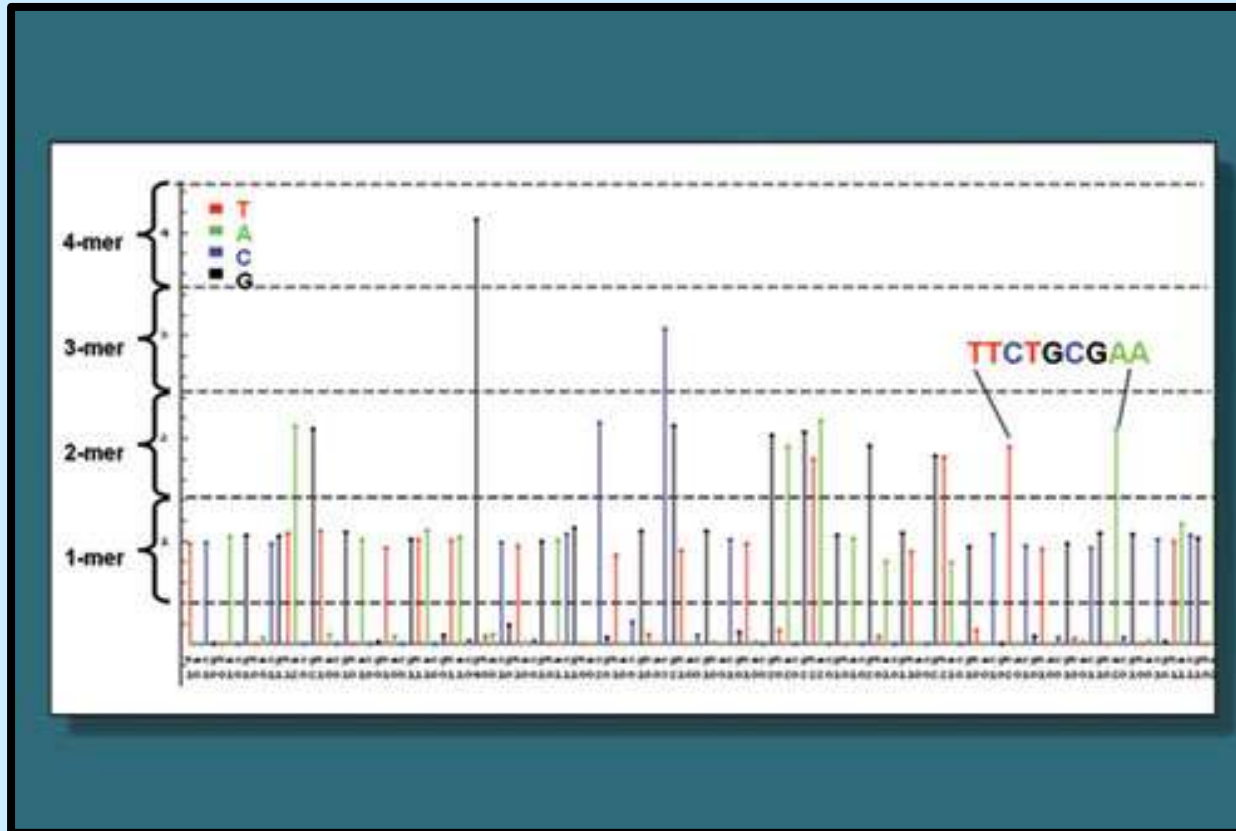
## **Základní kroky**

- **Tvorba jednořetězcových DNA matic**
- **Připojení adaptérů a vazba na pevné částice**
- **Amplifikace DNA matic v emulzi**
- **Vytvoření sekvenčních dat**
- **Analýza sekvencí různými nástroji bioinformatiky**

**Podívejte se na komerční prezentaci na stránkách  
<http://454.com/products/technology.asp>**

# 454 sekvenční systém

## Výsledný záznam





**Další moderní přístupy  
najdete na**

**<http://grf.lshtm.ac.uk/sequencing.htm>**

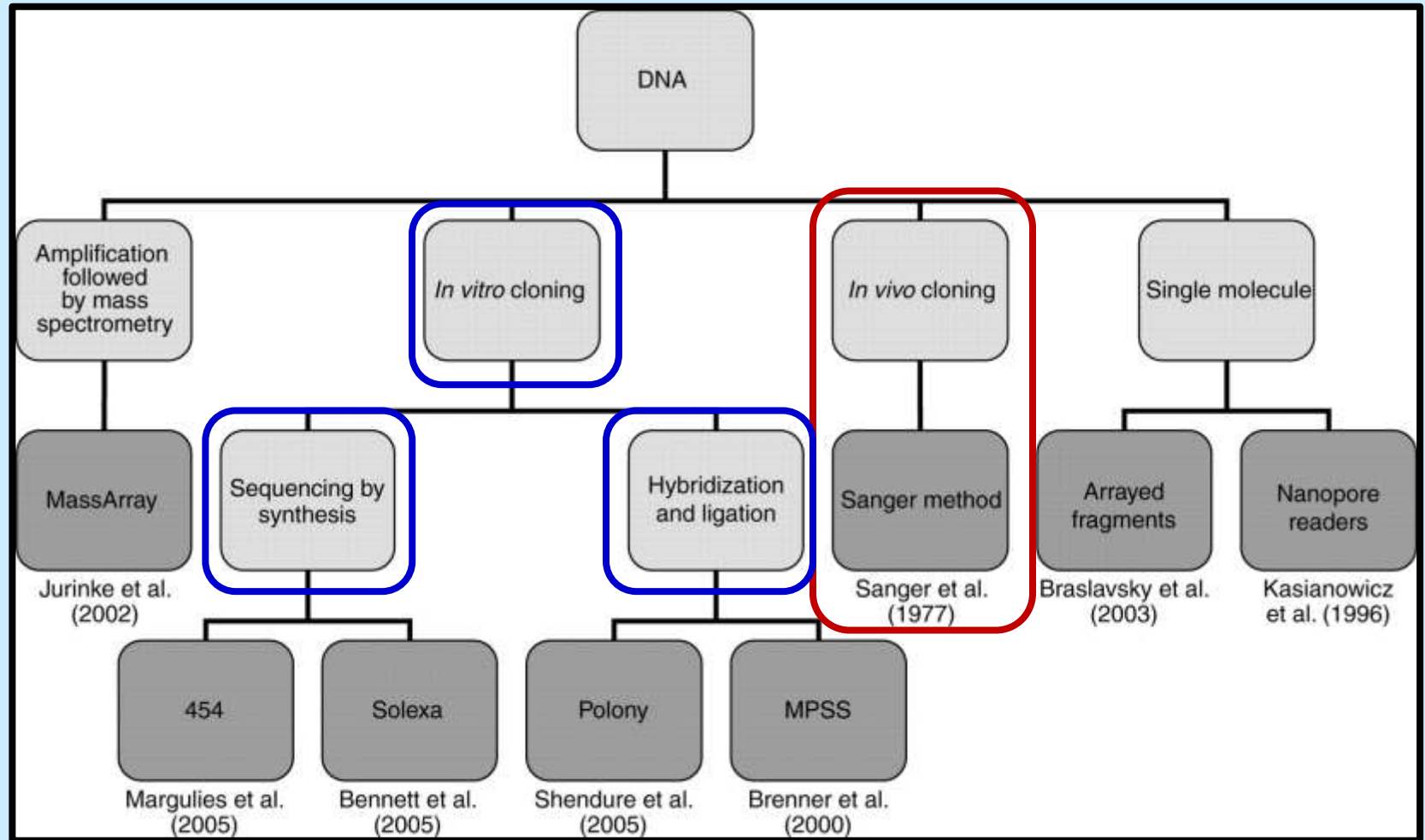
**SOLiD  
(Applied Biosystems)**

**454/Pyrosequencing  
(Roche)**

**SOLEXA  
(Illumina)**

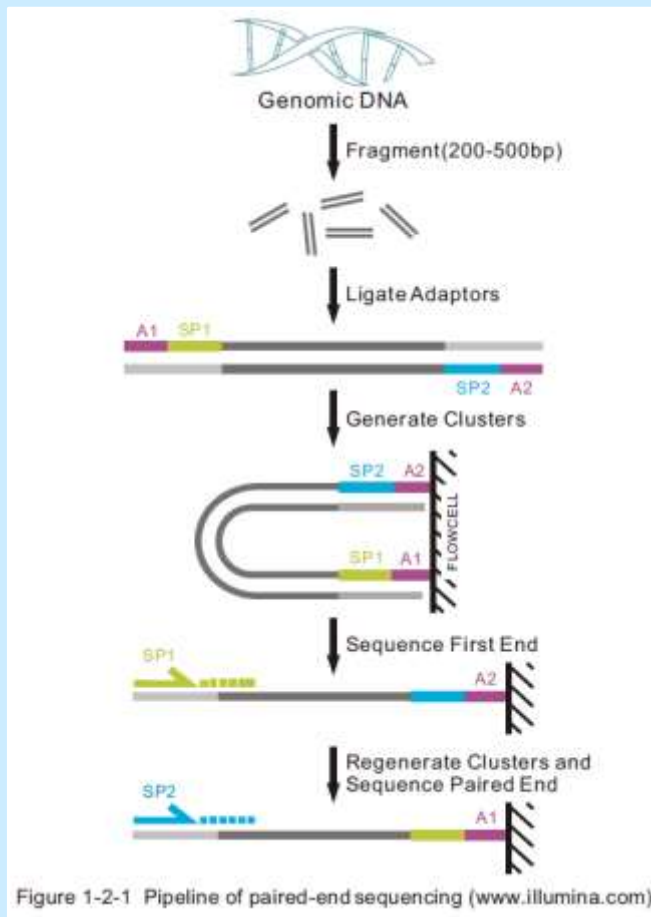


# Přehled stávajících a nových technik celogenomového sekvenování

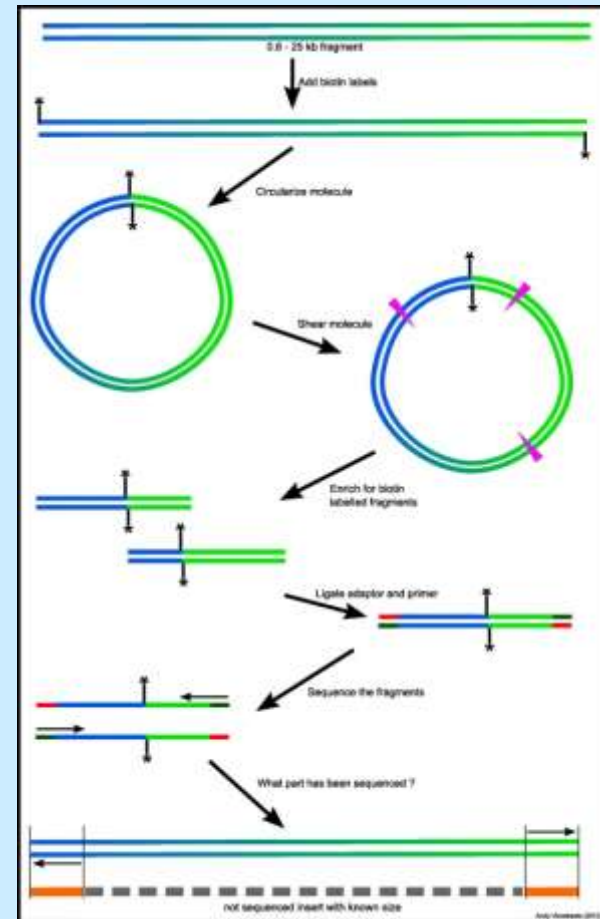


# Paired-end x Mate-pair

- Paired-end – sekvenování z obou konců fragmentu DNA (< 1 kb)
- Mate-pair – molekuly (3-20 kb) circularizované prostřednictvím vnitřního adaptoru

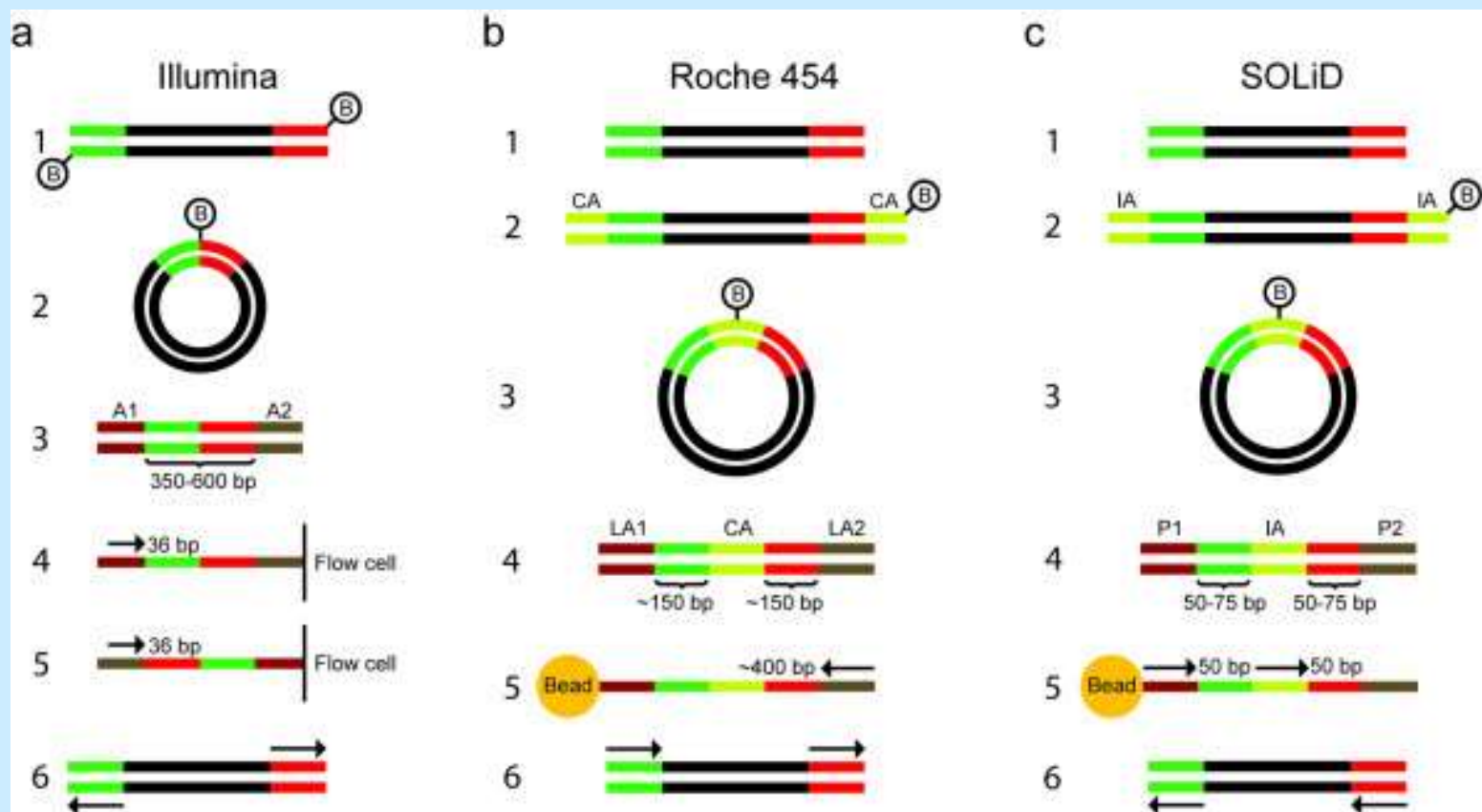


X



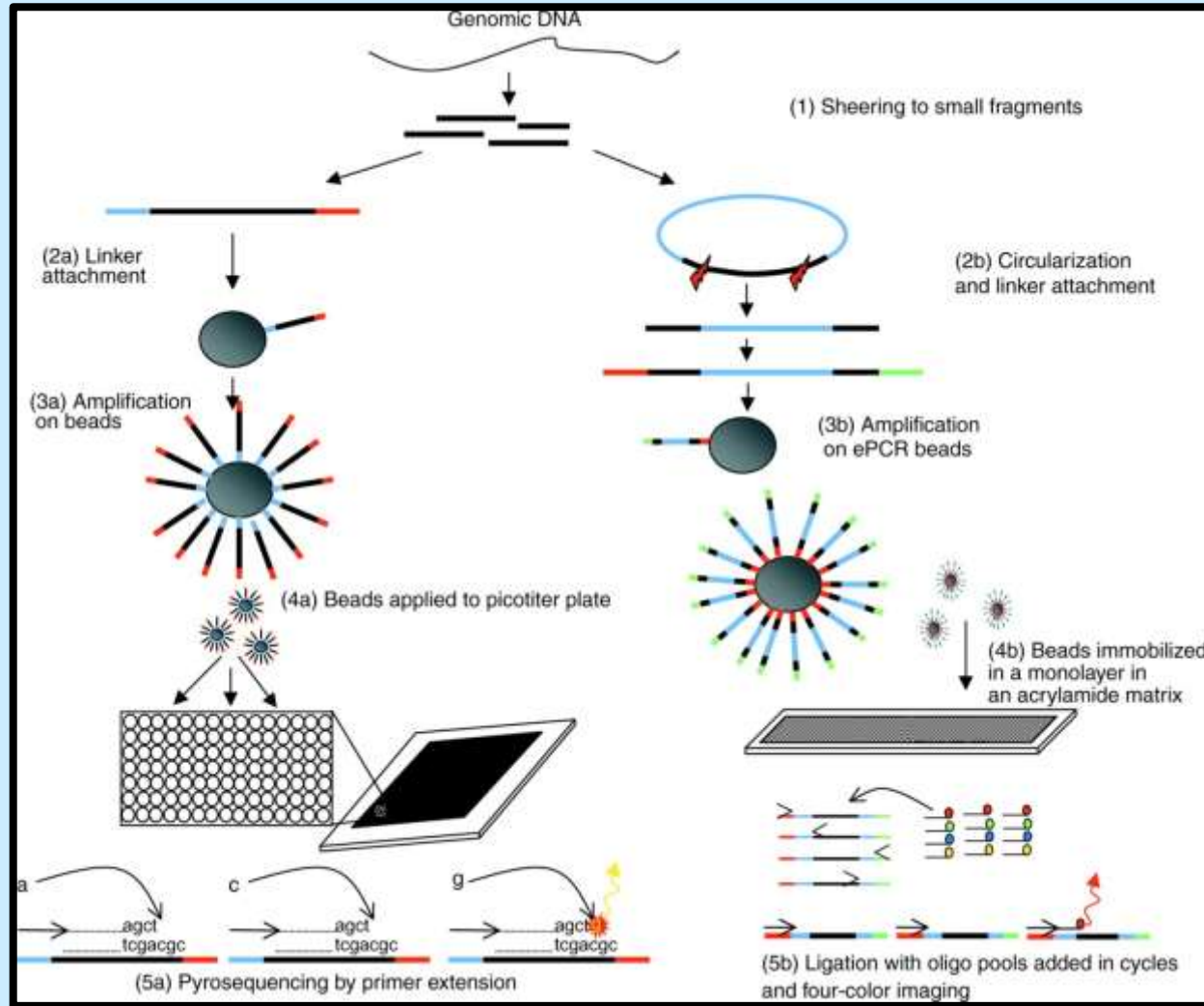
# Paired-end x Mate-pair

- Mate-pair – molekuly (3-20 kb) circularizované prostřednictvím vnitřního adaptoru
- Usnadnění skládání (u De Novo genomových projektů)



# Porovnání 454 a polony (SOLiD)

přímé  
připojení  
adapterů



cirkularizace  
a linearizace

destička

pyrosekve  
nování

monovrstva  
na PAGE

ligace  
značených  
oligo

# Porovnání parametrů – princip, výkon, přesnost, doba

Sequencer	454 GS FLX	HiSeq 2000	SOLiDv4	Sanger 3730xl	PacBio RSII
Sequencing mechanism	Pyrosequencing	Sequencing by synthesis	Ligation and two-base coding	Dideoxy chain termination	Sequencing by synthesis
Read length	700 bp	50SE, 50PE, 101PE	50 + 35 bp or 50 + 50 bp	400~900 bp	> 4000 bp
Accuracy	99.9%*	98%, (100PE)	99.94% *raw data	99.999%	99,999%
Reads	1 M	3 G	1200~1400 M	—	0.06 M
Output data/run	0.7 Gb	600 Gb	120 Gb	1.9~84 Kb	1.6 GB
Time/run	24 Hours	3~10 Days	7 Days for SE 14 Days for PE	20 Mins~3 Hours	30 Min – 3 Hours
Advantage	Read length, fast	High throughput	Accuracy	High quality, long read length	Read length, fast, no amplification, real time record
Disadvantage	Error rate with polybase more than 6, high cost, low throughput	Short read assembly	Short read assembly	High cost low throughput	Low throughput, low accuracy

Liu et al. 2012. Comparison of Next-Generation Sequencing Systems. Journal of Biomedicine and Biotechnology. 251364.

# Porovnání parametrů – náklady, výkon, aplikace

Sequencers	454 GS FLX	HiSeq 2000	SOLiDv4	3730xl
Instrument price	Instrument \$500,000, \$7000 per run	Instrument \$690,000, \$6000/(30x) human genome	Instrument \$495,000, \$15,000/100 Gb	Instrument \$95,000, about \$4 per 800 bp reaction
CPU	2* Intel Xeon X5675	2* Intel Xeon X5560	8* processor 2.0 GHz	Pentium IV 3.0 GHz
Memory	48 GB	48 GB	16 GB	1 GB
Hard disk	1.1 TB	3 TB	10 TB	280 GB
Automation in library preparation	Yes	Yes	Yes	No
Other required device	REM e system	cBot system	EZ beads system	No
Cost/million bases	\$10	\$0.07	\$0.13	\$2400

Sequencers	454 GS FLX	HiSeq 2000	SOLiDv4	3730xl
Resequencing		Yes	Yes	
<i>De novo</i>	Yes	Yes		Yes
Cancer	Yes	Yes	Yes	
Array	Yes	Yes	Yes	Yes
High GC sample	Yes	Yes	Yes	
Bacterial	Yes	Yes	Yes	
Large genome	Yes	Yes		
Mutation detection	Yes	Yes	Yes	Yes

**Liu et al. 2012. Comparison of Next-Generation Sequencing Systems. Journal of Biomedicine and Biotechnology. 251364.**

# Porovnání parametrů – stolní varianty přístrojů

	PGM	MiSeq	Junior
Output	10 MB–100 MB	120 MB–1.5 GB	70 Mb
Read length	~200 bp	Up to 2 × 150 bp	700 bp
Sequencing time	2 hours for 1 × 200 bp	3 hours for 1 × 36 single read 27 hours for 2 × 150 bp pair end read	18 hours
Sample preparation time	8 samples in parallel, less than 6 hrs	As fast as 2 hrs, with 15 minutes hand on time	2 days
Sequencing method	semiconductor technology with a simple sequencing chemistry	Sequencing by synthesis (SBS)	Pyrosequencing
Potential for development	Various parameters (read length, cycle time, accuracy, etc.)	Limited factors, major concentrate in flowcell surface size, insert sizes, and how to pack cluster in tighter	Minimize hand on time, increase emPCR reproducibility
Input amount	μg	Ng (Nextera)	μg
Data analysis	Off instrument	On instrument	On/Off instrument



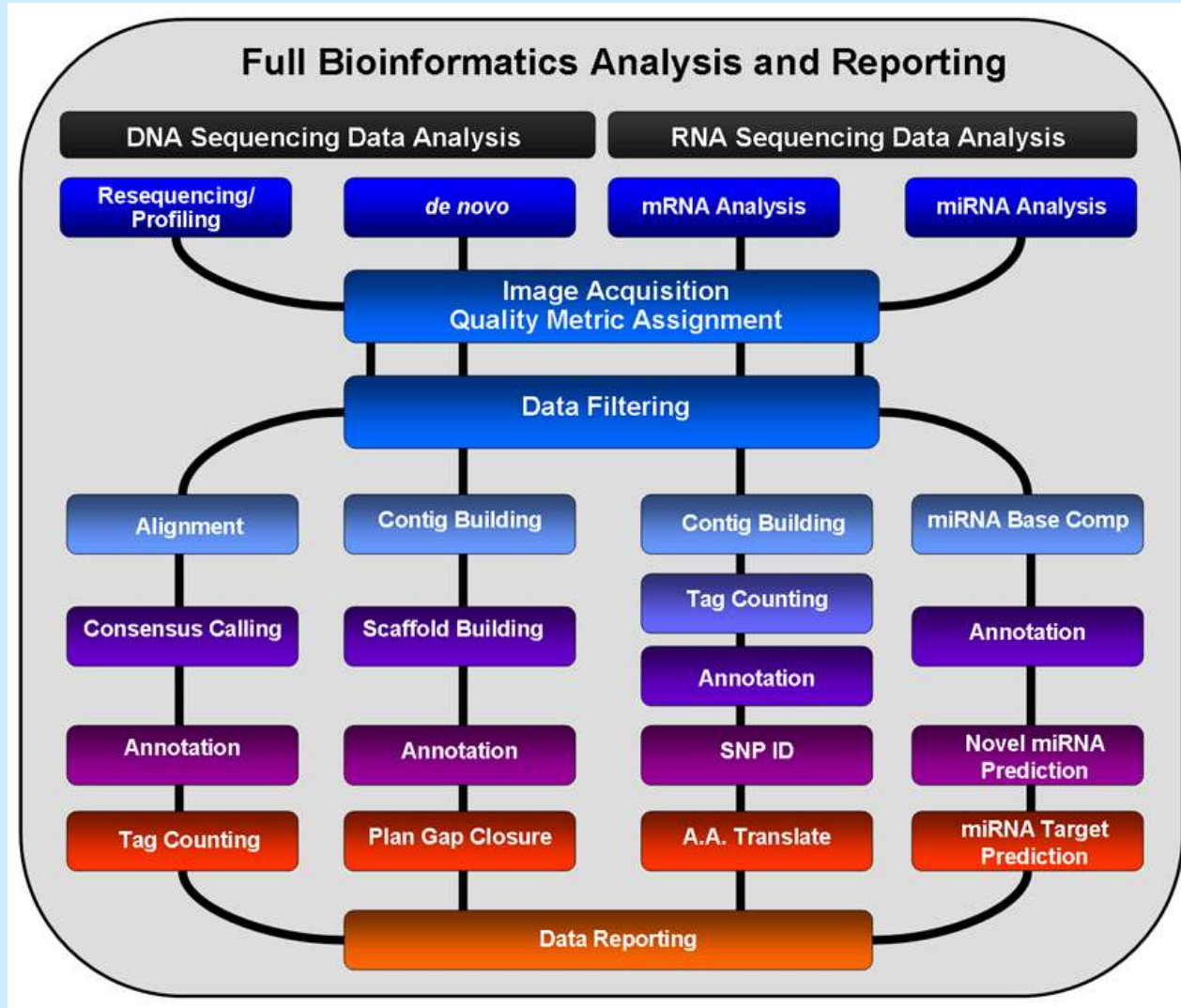
**Liu et al. 2012. Comparison of Next-Generation Sequencing Systems. Journal of Biomedicine and Biotechnology. 251364.**



# ***Aplikace a využitelnost***

- **de novo DNA/RNA sekvenace** – Illumina, Roche/454 (PE), PacBio
- **Resekvenování** – SOLiD, Illumina
- **Detekce SNPs** – Roche/454, PacBio (x InDels variation – Illumina, SOLiD)
- **Sequence capture** - Illumina
- **Sanger** – sekvenování konkrétních regionů s nízkým pokrytím (diagnostické genotypování), sekvenování malých genomů (plasmidy, viry, fágy ...)
- **Ion Torrent** – kratší amplikony
- **SOLiD** – kvantitativní aplikace, malé RNA, epigenomika
- **HeliScope** – kvantitativní aplikace, malé vstupní množství (single molecule)
  
- **Kombinace metod?**

# Analýza dat, skládání, anotace



# ***Analýza dat, skládání, anotace***

- **Software od výrobce dané technologie (uživ. jednoduché, neefektivní)**
- **Obecné, volně dostupné programy (najít ten správný, porovnat více)**
- **Vyvinutý uživatelem pro danou aplikaci (kvalifikovaní bioinformatičtí)**
  
- **Kombinace dat z různých platforem x problém pro assembly**
- **Specifické chyby technologií, specifické nastavení parametrů analýz**
- **Vícestupňové filtrování dat**

# ***Sekvenování mikroorganismů***

- První virus – fíX 174 v roce 1977
- První sekvenovaný mikroorganismus = *Haemophilus influenzae*, 1995, TIGR
- Následovaly *Mycoplasma genitalium* 1995, *Methanococcus jannaschii* 1996, *Helicobacter pylori* 1997, *Escherichia coli* 1997, *Mycobacterium tuberculosis* 1998
- První eukaryotický organismus *Saccharomyces cerevisiae* 1996
- V současnosti jsou v databázích stovky sekvencí
- Podívejte se na <http://www.ncbi.nlm.nih.gov/genome>

# Úkol



Podívejte se do databáze a zjistěte, u kolika halobakterií (třída *Halobacteria*) je v současné době známa nukleotidová sekvence genomu, a to jak kompletní, tak i částečná

Zjistěte totéž u čeledi *Pasteurellaceae*

A ještě u cyanobakterií (kmen *Cyanobacteria*)

# Další příklady I

## Genome Gallery

A selection of notable genomes that have been sequenced.

*Φ X 174*

(1977) 5386 bp

First genome sequenced, a bacteriophage

*Haemophilus influenzae*

(1995) 1 830 000 bp

First genome of a free-living organism

*Mycoplasma genitalium*

(1995) 580 000 bp

Smallest genome of any free-living organism

*Saccharomyces cerevisiae*

(1996) 12 100 000 bp

First genome of a 'eukaryotic' (nucleus-containing) organism, the yeast used by brewers and bakers

*Methanococcus jannaschii*

(1996) 1 660 000 bp

The first genome from the third kingdom, *Archae*, which comprises microbes that live in harsh environments, for example thermal springs

*Escherichia coli*

(1997) 4 670 000 bp

Workhorse bacterium for biologists

*Helicobacter pylori*

(1997) 1 660 000 bp

Bacterium associated with gastric disease



Dr Linda Stannard, UCT / Science Photo Library

# Další příklady II

## Genome Gallery



Wellcome Photo Library

*Mycobacterium tuberculosis*  
(1998) 4 400 000 bp  
Cause of the disease  
tuberculosis

*Deinococcus radiodurans*  
(1999) 2 600 000 bp Highly  
radiation-resistant bacterium

First human chromosomes  
(1999 and 2000)

Chromosomes 22,  
48 000 000 bp and 21,  
45 000 000 bp respectively



National Human Genome  
Research Institute, NIH

*Caenorhabditis elegans*  
(1998) 97 000 000 bp  
The first genome sequence  
of an animal, the roundworm

AVP, LSHTM/Wellcome Photo Library



*Drosophila melanogaster*  
(2000) 180 000 000 bp  
Fruit fly, an important labor-  
atory organism in genetics

# Další příklady III

## Genome Gallery



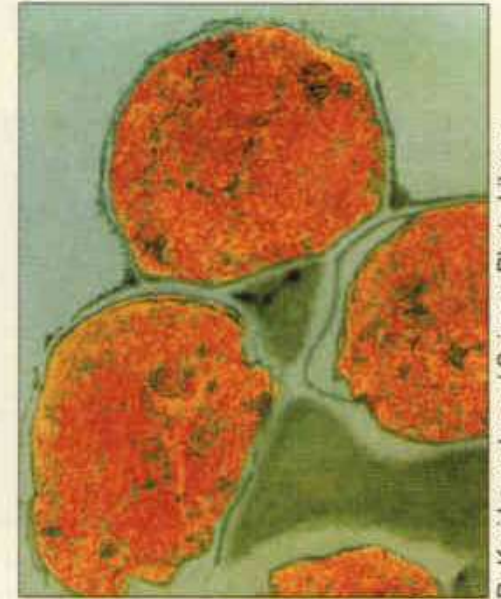
Eye of Science / Science Photo Library

*Vibrio cholerae*  
(2000) 4 030 000 bp  
Cause of the disease  
cholera



Images courtesy of J. Berger, T. Laux & E. Meyerowitz

*Arabidopsis thaliana*  
(2000) 120 000 000 bp  
The first genome of a plant,  
the mustard weed



Dr Kari Lounatmaa / Science Photo Library

*Mycobacterium leprae*  
(2001) 3 270 000 bp  
Cause of the disease  
leprosy



***Metody sekvenování  
tzv. 3. generace***

# Metody, které nevyžadují amplifikaci

**Helicos (bankrot)**

**Pacific Biosciences**

**Nanopore**



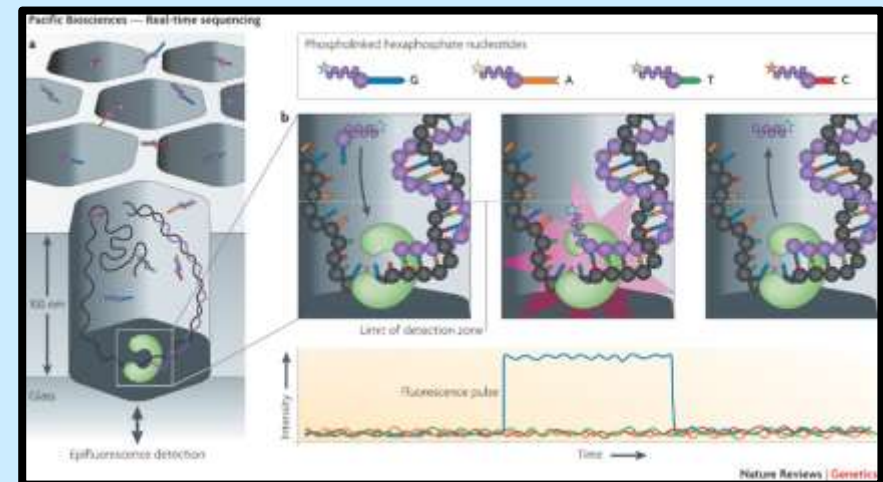
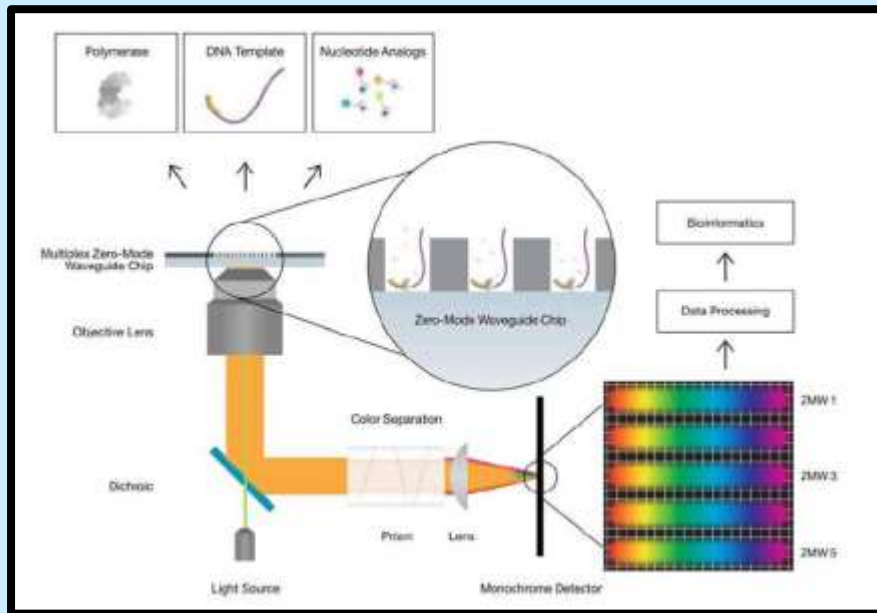
# ***Technologie třetí generace***

- Na úrovni jedné molekuly, v reálném čase, případně *in situ*
- Využití mikro a nanofluidních technologií
- Bez potřeby celogenomové amplifikace pomocí MDA -  $\Phi$ 29 DNA polymerázové reakce nebo PCR amplifikace knihoven
- Redukce enzymových reakcí (využití charakteristických vlastností jednotlivých bází – ovlivnění toku elektrického proudu)
- Sekvenování lidského genomu (pokrytí 30x) za cenu nižší než 1000 \$ ohlášena společností Illumina (platforma HiSeq X Ten)

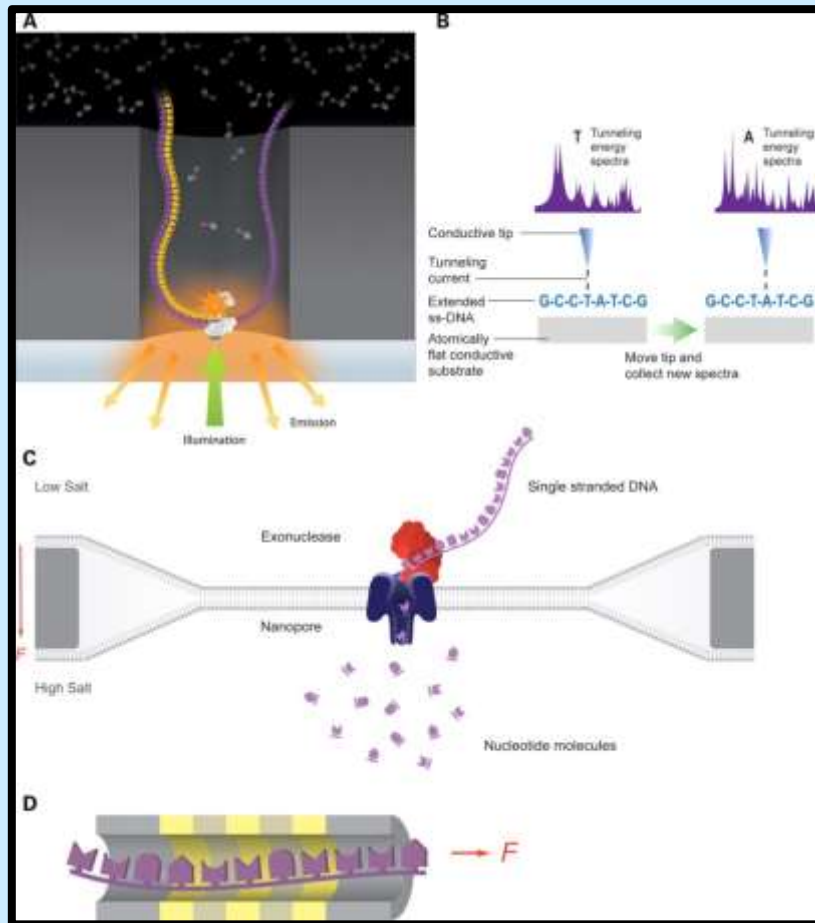
***Jak sekvencuje SMRT?***

# Technologie třetí generace

## Single-Molecule Real-Time seq – SMRT Pac Bio (without amplification necessary for signal detection)



# Technologie třetí generace



## Pacific Biosciences technology

direct observation of DNA synthesis on single DNA molecules in real time. A DNA polymerase is confined in a zero-mode waveguide and base additions measured with fluorescence detection of gamma-labeled phosphonucleotides.

## Reveo technology

direct inspection of stretching ssDNA using electric current change detection (electron microscopy)

## Oxford Nanopores technologies

measuring translocation of nucleotides cleaved from a DNA molecule across a pore, driven by the force of differential ion concentrations across the membrane.

## IBM – silicon based nanopores

DNA transistor technology reads individual bases of ssDNA molecules as they pass through a narrow aperture based on the unique electronic signature of each individual nucleotide. Gold bands represent metal and gray bands dielectric layers of the transistor.

# *Technologie třetí generace*

**Minion**

**versus**

**MinION**



**A loyal servant of another, usually more powerful, being.**



**A single-molecule sensing instrument based on nanopores technology.**

# ***Plánování sekvenačního experimentu***

- **Vstupní materiál (dostatečné množství, dobrá kvalita)**
- **Výběr aplikace => metody/platformy**
- **Vlastní zpracování x servis**
- **Výpočetní kapacita (uložiště, operační prostor, zálohy)**
- **Bioinformatické zázemí**



# ***A teď se podíváme na něco praktického***



**Použijte připravený výukový materiál ze sekvenování zástupců čeledi *Pasteurellaceae***

# *Shrnutí*

- 1) Pojem sekvenování**
- 2) Maxamovo-Gilbertovo sekvenování**
- 3) Sangerova metoda s využitím fluoroforů**
- 4) Pyrosekvenování**
- 5) Next generation sequencing**
- 6) Third generation sequencing**
- 7) Příklad reálné analýzy**