

# Metody sekvenování

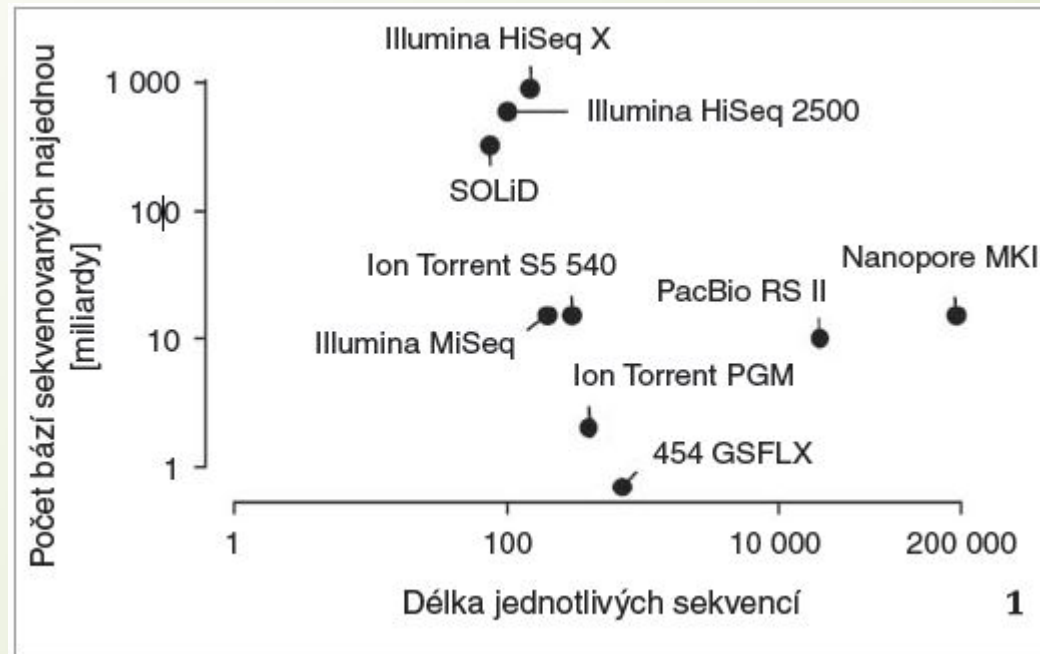
S. Meliorisová, M. Madrzyk, S. Sidorová, M. Šťastná, J. Polaček, A. Strunga



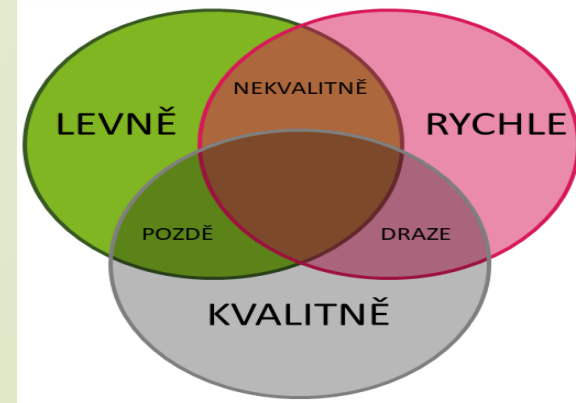
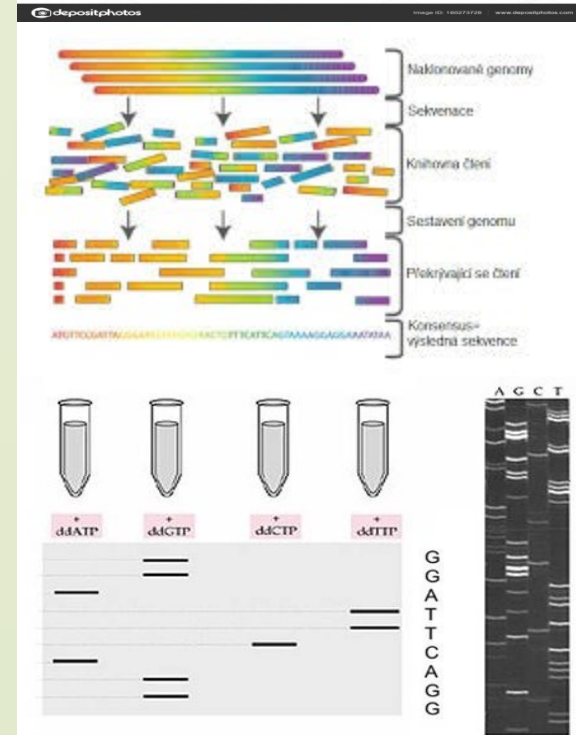
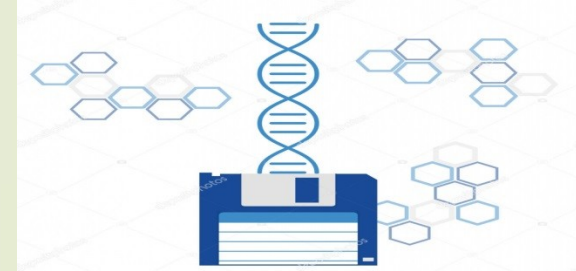
# Vysokokapacitní (high-throughput)analýzy

- Nástroje pro souběžné zpracování a testování velkého množství vzorků nebo velkého množství analytů.
- 1. Amplifikace pomocí PCR (kvantitativní PCR arraye)**
  - 2. Hybridizace (mikroarraye = mikročipy)**
  - 3. Sekvenační metody**

- = určení sekvence nukleotidů dané molekuly DNA
- V současné době nejrozšířenější způsob analyzování biologického materiálu
- Vyvinuto více technik
  - Tradiční – Sangerova metoda, Maxam – Gilbertova metoda
  - Sekvenování tzv. druhé generace
  - Sekvenování tzv. třetí generace



- Na začátku 70. let nepřímé sekvenování – RNA molekuly a proteiny
- 1975 – **Frederick Sanger**
- 1977 - **Walter Gilbert, Alan Maxam**
- 1995 – první zcela osekvenovaný genom bakterie *Haemophilus influenzae*
- Od konce 70. let 20.století je možné rychlé a účinné sekvenování





# Cesta k sekvenácii ľudského genómu

- **1865** Gregor Mendel – otec genetiky – hybridizácia rastlín
- **1869** Friedrich Miescher – objav nukleínu (**P**, H, O)
- **1952** Rosalind Franklinová – X-Ray difrakcia DNA – helikálna štruktúra
- **1953** J. Watson, F. Crick - DNA jako dvojšroubovica
- **1961** Marshall Nirenberg – prelomil gen. kód v zmysle proteosyntézy
- **1977** Frederick Sanger - rýchle DNA sekvenovanie
- **1983** Huntingtonova choroba – repetície CAG (izolácia o 10 rokov)
- **1989** **PCR** vyvinutá pre amplifikáciu DNA
- **1989** Identifikácia génovej mutácie pre cystickú fibrózu
- **1990** Prvý dôkaz o existencii BRCA1 génu
- **1990** „HUMAN GENOME PROJECT“



# Sangerova metoda

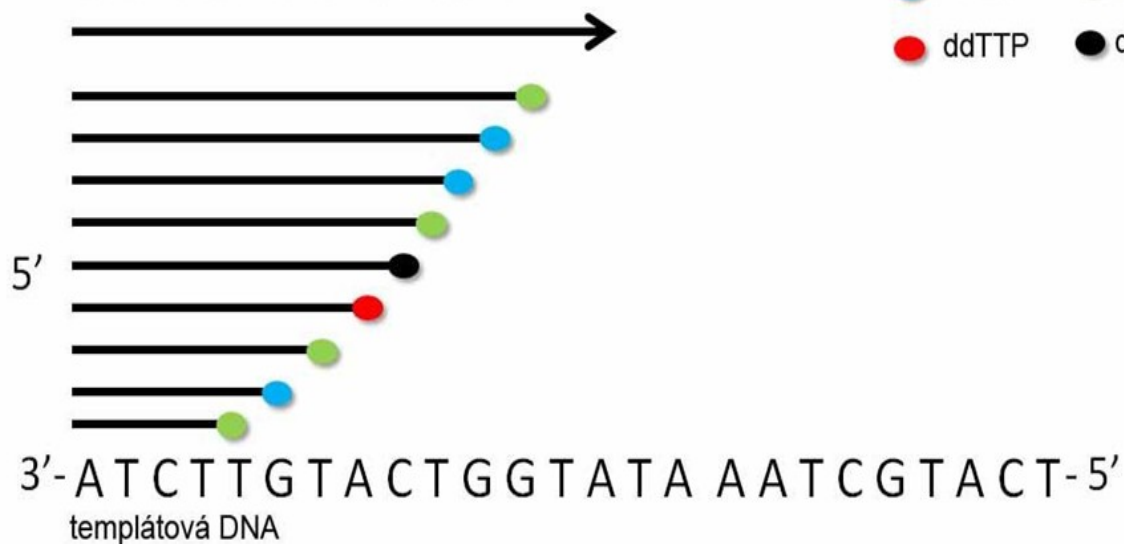
- „Dideoxy metoda“,
- syntéza řetězce (až tisíc bp)
- Použita v projektu sekvenace kompletního lidského genomu v roce 2001
- Vzorek DNA je rozdělen do čtyř paralelních reakcí
  - Směs obsahuje: DNA polymerázu, primer, deoxynukleotidy a jednotlivý dideoxynukleotid
  - dideoxynukleotidy se náhodně začlení do syntetizovaného řetězce místo příslušného dNTP a v tomto místě syntéza končí
- Začlenění ddNTP se děje v náhodném místě ---> vznikají různě dlouhé fragmenty DNA, které se analyzují pomocí elektroforézy

# SANGEROVA METODA SEKVENOVÁNÍ (KAPILÁRNÍ PROVEDENÍ)

Fluorescenčně značené ddNTP

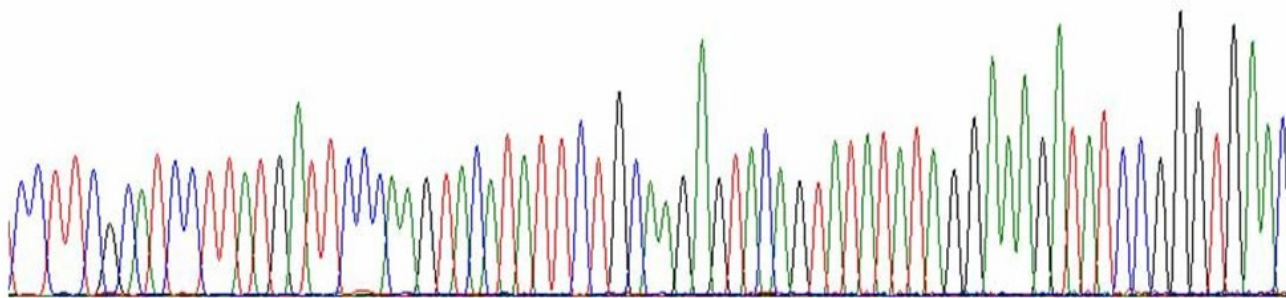
- ddCTP
- ddATP
- ddTTP
- ddGTP

Syntéza DNA za přítomnosti ddNTP

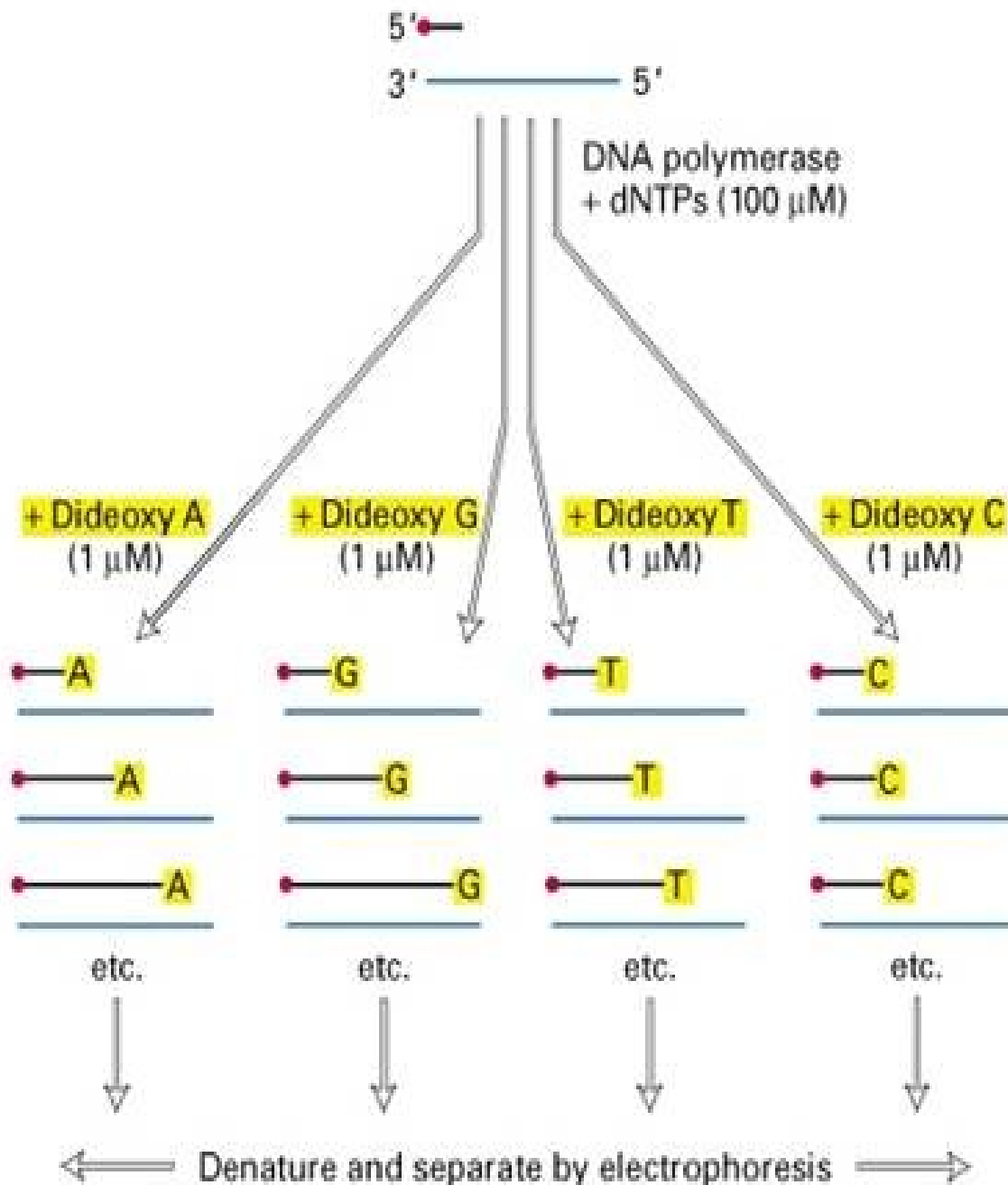


Výsledný sekvenogram

30 40 50 60 70 80 90 100  
 CCTTCG CAT CCT TATG ATTC CCAAG TACATAT TCTG CAAAGAGTACAG TATATATA G GAAA GATAT CC GGGT GAAC



(a)





# Maxam–Gilbertovo sekvenování

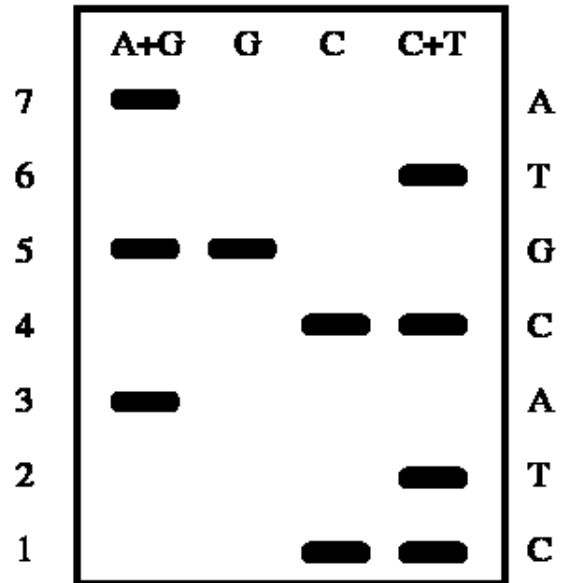
- využívá chemického štěpení jednotlivých typů bází
- pracuje s jednovláknovou DNA, která je na svém jednom konci (3' nebo 5') radioaktivně značená
- Probíhá ve čtyřech paralelních reakcích, kdy se štěpí pouze určitý typ bází
- Vznikají různě dlouhé fragmenty DNA končící v místě dané báze
- Sekvenci analyzujeme za pomoci elektroforézy, kdy odečtením pozice jednotlivých bází ve všech čtyřech reakcích stanovíme sekvenci daného úseku.



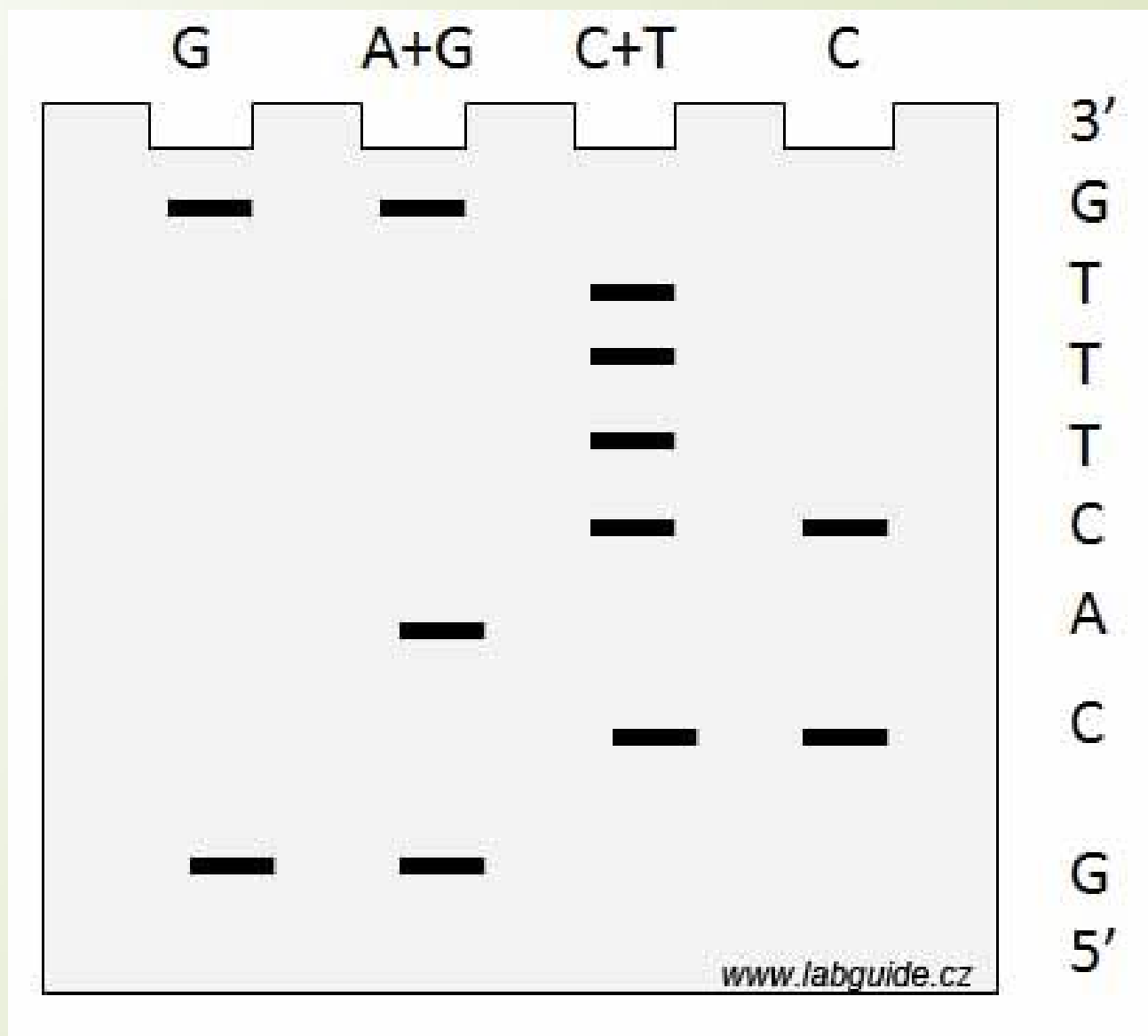
5' <sup>32</sup>P GCTACGTA 3'

Cleavage at: A+G      G      C      C+T

<sup>32</sup>P GCT  
<sup>32</sup>P GCTAC  
<sup>32</sup>P GCTACGT  
<sup>32</sup>P GCTAC    <sup>32</sup>P G  
<sup>32</sup>P GCTA  
<sup>32</sup>P G  
<sup>32</sup>P GC  
<sup>32</sup>P GCTA  
<sup>32</sup>P GCTACG



Sequencing Gel





# Sekvenování nové generace

- PCR se provádí paralelně pro všechny fragmenty DNA
  - Paralelní sekvenování několika milionů sekvencí
  - Délka získaných sekvencí cca 50 – 600 bp
  - Sekvenační výtěžek jednoho běhu až několik tisíc Gb (až o 6 řádů vyšší než u kapilárního sekvenování)
  - Cena sekvenace za bázi o řád až dva nižší než u kapilárního sekvenování
- 
- Druhá generace: Illumina, 454 sekvenování, Solid sekvenování,
  - Třetí generace: SMRT, Nanopor, Microarray

# PROJEKT ĽUDSKÉHO GENÓMU METÓDY SEKVENÁCIE



# HGP (1990)

- Medzinárodný výskumný program (13 r.)
- Sponzori : National Institute of Health, U.S. Department of Energy
- skepticizmus

## Ciele:

- Medzinárodná spolupráca (IHGSC)
- Vyvinúť technológie pre analýzu DNA, zmapovať a sekvenovať ľudský genóm
- Všetky informácie verejne dostupné do 24 hod
- 200 laboratórií v U.S., 18 krajín z celého sveta

# CELERA Genomics – Craig Venter

- Craig Venter - zakladateľ Inštitútu genomického výskumu (Maryland)
- Celera – 1998
- sekvenovať ľudský genóm za 3 roky
- „preteky“
- Informácie dostupné iba pre platiacich zákazníkov
- Predbežné patenty na 6000 génov
- Plné patenty na približne 100 génov
- (kontrola nad tým, kto získa prístup k informáciám)
- HGP súčasne zvyšuje financovanie Sangerovho inštitútu (GB)

# Sekvenované genómy

- *Heamophilus influenzae* (1995) - shotgun
- *Homo sapiens* – chromozóm 22 (1999)
- *Drosophila melanogaster* (2000)
- *Homo sapiens* (2001) – súčasne Celera a HGP
- *Mus musculus* (2002)

HGP – hierarchiálne shotgun sekvenovanie (klon po klone) – DNA fragmenty

Celera – shotgun celého genómu  
- rýchlejšie sekv. vďaka technike



# Konflikt – verejný a súkromný výzkum

- 2000 – tlak Bieleho Domu na urovnanie rozdielov
- 26 jún 2006 – vyhlásenie dvoch rôznych konceptov sekvenácie ľudského genómu
- Preteky sú na konci, obe strany vyhrali
- Craig Venter zostúpil z prezidentského kresla spoločnosti, keďže Celera začala smerovať viac k FARMÁCIÍ

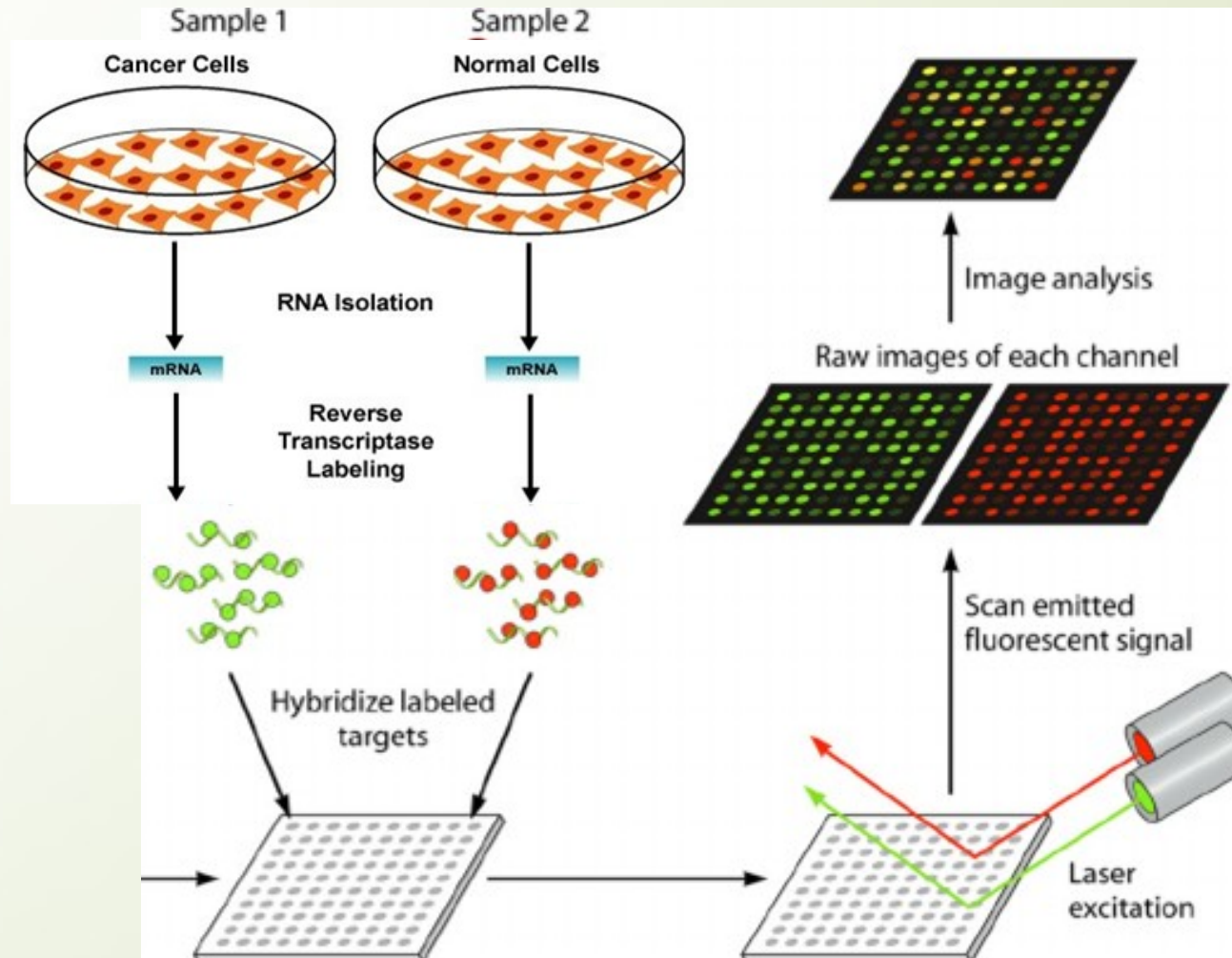


# Real-time PCR arraye

- Přesný nástroj pro paralelní analýzu genové exprese a je využitelná pro ostatní aplikace kvantitativní PCR
- Používají se reagenty jako v běžné PCR, detekce probíhá pomocí hybridizačních (TaqMan) nebo LNA sond, případně SYBR green I
- Výhody:
  - Vysokokapacitní zpracování analýz
  - Přesnost, specificita kvantitativní PCR
  - Design arrayí je flexibilní, velké množství panelů



# DNA microarrays (čipy)- postup

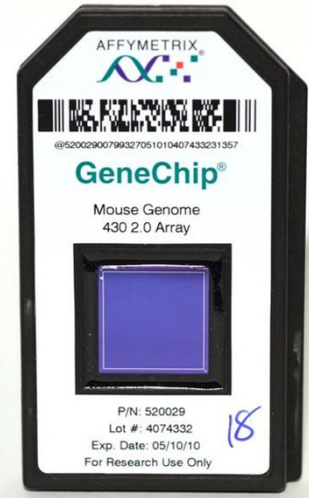
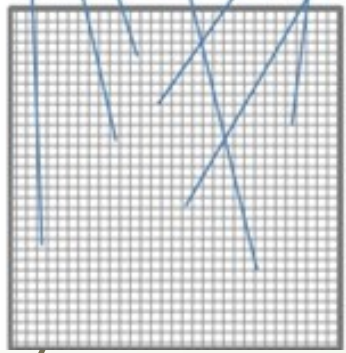
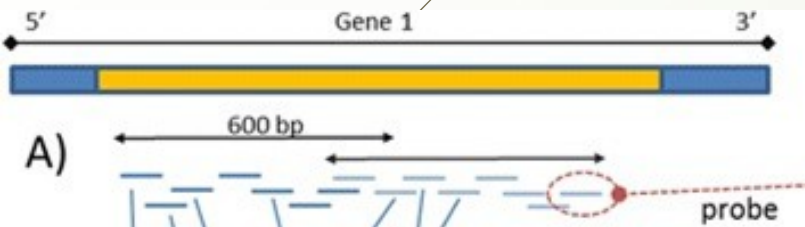




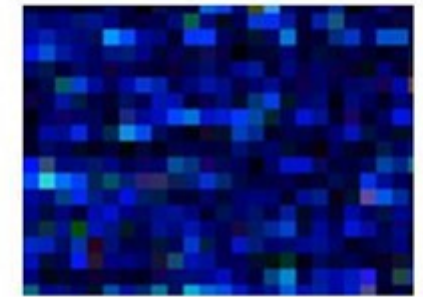
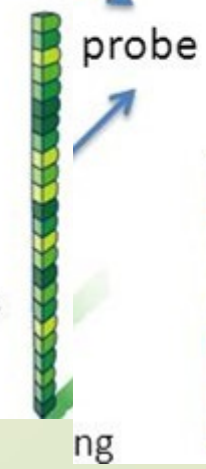
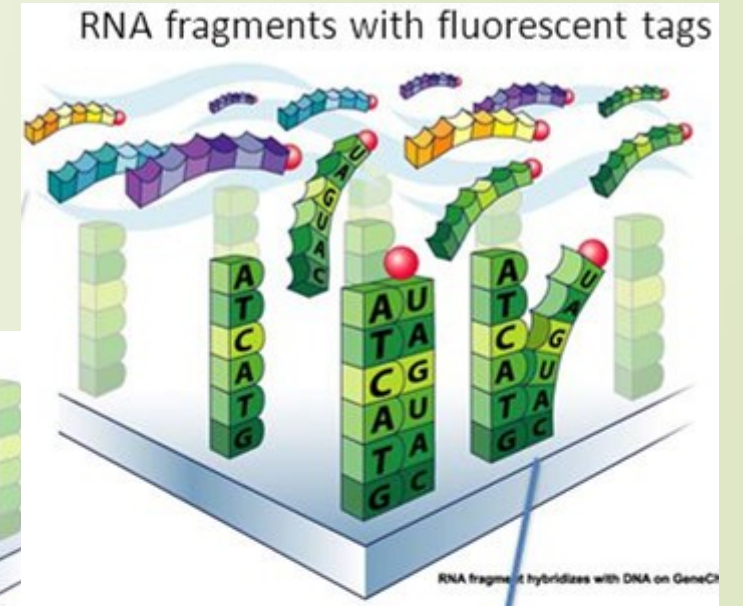
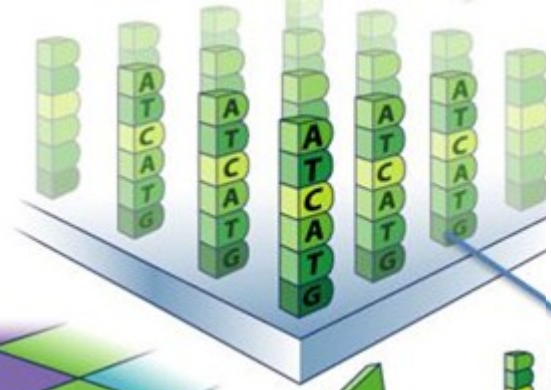
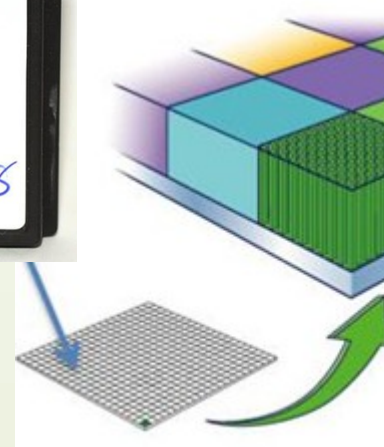
# DNA microarrays (čipy)

- Hlavním principem: hybridizace komplementárních úseků DNA, z nichž jeden pochází z biologického vzorku, zatímco druhý je uměle připravený a funguje jako **sonda (próba)**, která je fixována na podložním sklíčku čipu, v přesně definované pozici
- Nejvýznamnější výrobci Affymetrix a Agilent
- DNA molekuly vzorku musí být před nanesením na čip označeny (fluorescenční značka), aby mohla být identifikována jejich pozice a kvantita.
- **Hybridizační jednotka (spot)** = plocha určená pro detekci jedné cílové sekvence
- Každý spot obsahuje miliony kopií sondy

# Uspořádání čipu

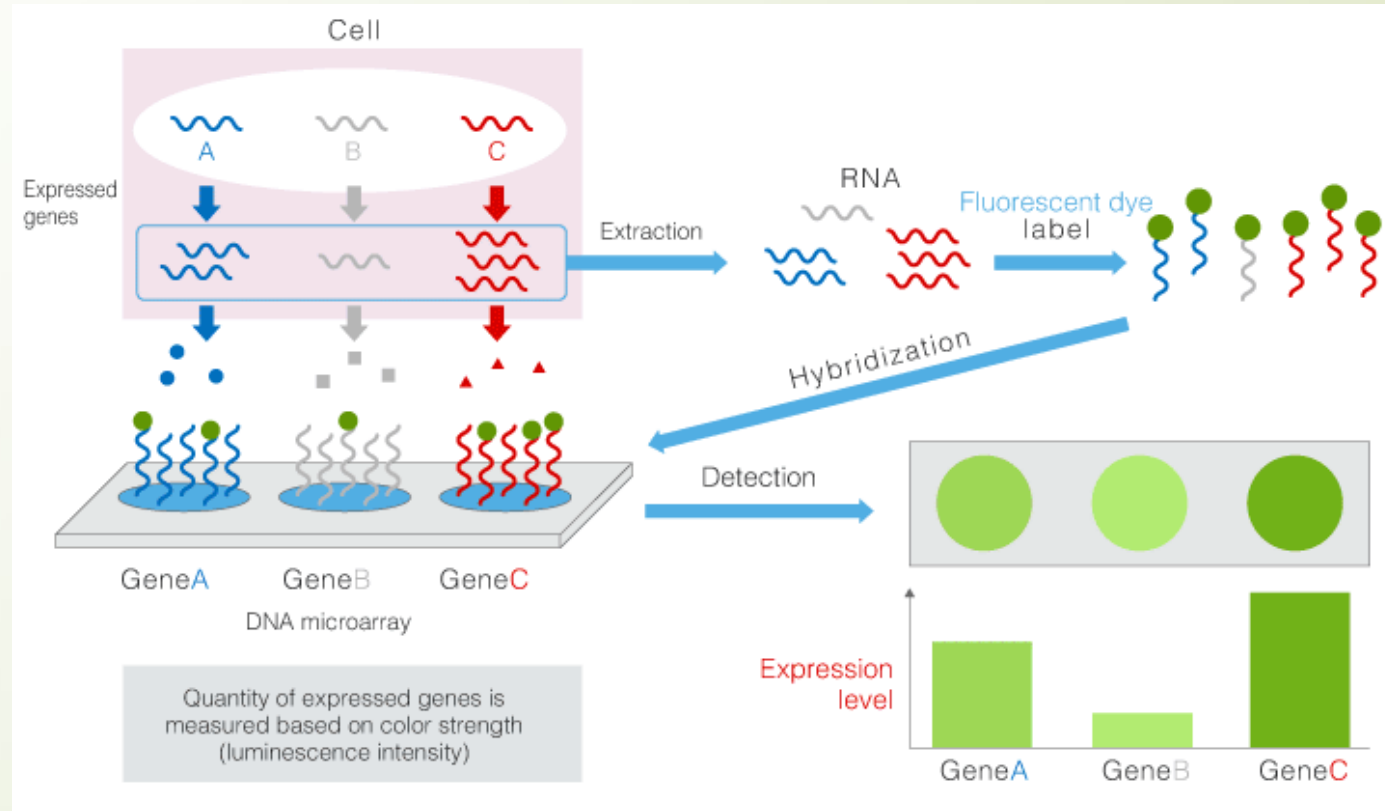


1.28 cm  
1.28 cm  
actual size of GeneChip™



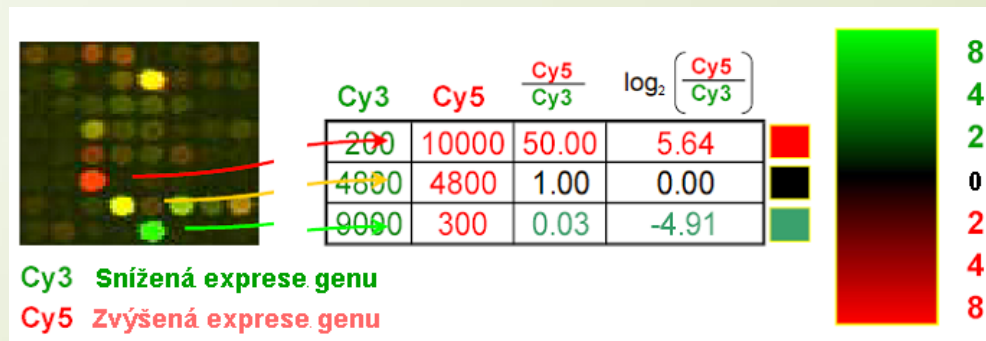
ng

# Jednobarevný čipový experiment



# Dvoubarevný čipový experiment (Agilent)

- U cDNA čipů hybridizujeme na jednom čipu značenou ss-cDNA získanou z testovaného a referenčního vzorku. Proto je nutno každou ss-cDNA označit rozdílným fluorochromem.
- Standardně: referenční ss-cDNA je Cy-3 – zeleně a testovaná ss-cDNA Cy5 – červeně
- DNA z obou vzorků soutěží o sondy na jednom spotu a výsledný poměr nahybridizovaných molekul (=poměr intenzit signálů) bude odpovídat rozdílu exprese daného genu v obou vzorcích
- Po hybridizaci následuje vymytí přebytečného materiálu a čtení (skenování) čipu pomocí fluorescenčního skeneru





# Aplikace mikročipových technologií

- **Exonové čipy**
  - identifikace stříhových variant transkriptů, kvantifikace jednotlivých stříhových variant
- **SNP čipy**
  - Genotypizace, forenzní analýza, predispozice k onemocněním
- Chromatinová imunoprecipitace na čipu (**ChIP on Chip**)
  - Analýza interakcí proteinů s DNA *in vivo* na celogenomové úrovni
  - Využívá se kombinace imunoprecipitace chromatinu s vysokokapacitními mikročipy
  - Identifikuje vazebná místa pro studovaný protein v celém genomu současně
- vs **ChIP-seq**
  - přesnější, větší citlivost, specificita, přímá identifikace získaných fragmentů, není limitován studovaným organismem
  - Nevýhodou je nemožnost analýzy pouze vybraných oblastí genomu, analýza dat je náročnější




# Výhody a nevýhody mikročipů

- Výhody: využití v klinické diagnostice
  - ampliChip CYP450 - variantní alely genů pro cytochromy P450
- Nevýhody:
  - Čipové technologie jsou překonány sekvenováním nové generace



# Second-generation sequencing





# Metódy sekvenovania druhej generácie

- Od roku 2005 vyvinuté nové metódy sekvenovania (rýchlejšie, lacnejšie)
- Mnohonásobné paralelné sekvenovanie genómu v priebehu niekoľkých týždňov
- Využívajú syntézu DNA podľa templátu
- ILLUMINA, 454 sekvenovanie, Ion torrent a SOLiD sekvenovanie



# Výhody a nevýhody

- Umožňujú sledovať tisícky ľudských genómov, vzniknuté mutácie, možné riziká ochorení
- Detekujú pridávanie báz jednu po druhej
- Sekvenujú tisíce až milióny DNA naraz
- Nevýhodou je krátka maximálna dĺžka výsledných sekvencií (ca 100-500 báz)
- Oproti Sangerovmu sekvenovaniu menšia presnosť a častejšie chyby pri čítaní DNA

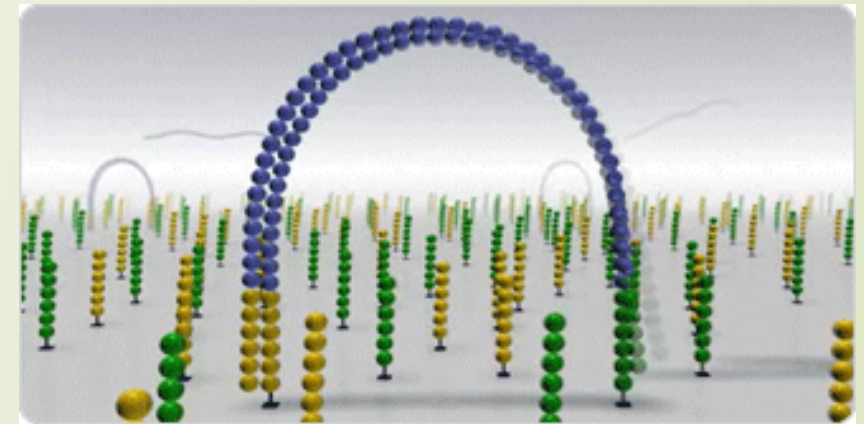


# Postup pri sekvenovaní druhej generácie

1. DNA štiepená na kratšie reťazce
2. Naviazaný adaptér na konce týchto reťazcov (krátka DNA o presnej sekvencii)
3. Uchytenie DNA na pevný povrch (adaptér komplementárny ku reťazcom na sekvenačnom povrchu)
4. Amplifikácia uchytených reťazcov (vznik klastrov)
5. Zachytenie signálu kamerou

# ILLUMINA sekvenovanie

- Nebulizace – vplyvom stlačeného vzduchu DNA rozdelená na menšie fragmenty (1-1000 bp), enzymatický upravené
- Na konce sa naviažu dva adaptéry
- Na pevnej doske DNA prichytená adaptérom, namnoží sa
- Vzniká mozaika klastrov (zhlukov) identických DNA
- Do rastúceho reťazca pridávané bázy s naviazanou fluorescenčnou farbou na 3' OH koniec (báza = špecifická farba)
- Záznam vysoko citlivou kamerou, následne enzymatické odstránenie fluorescenčného značenia a blokujúcej časti molekuly
- Podľa rozdielnej fluorescencie kamera rozozná typ pridanej bázy



Obr. Bridge amplification



# ILLUMINA

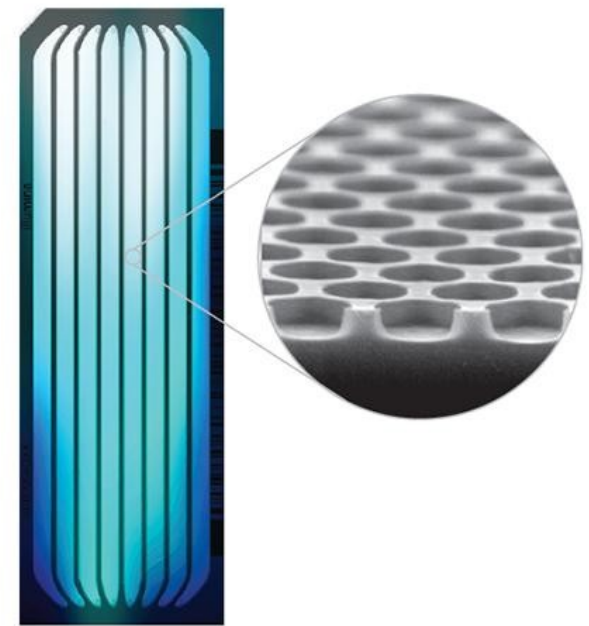
Analýza fluorescenčných signálov



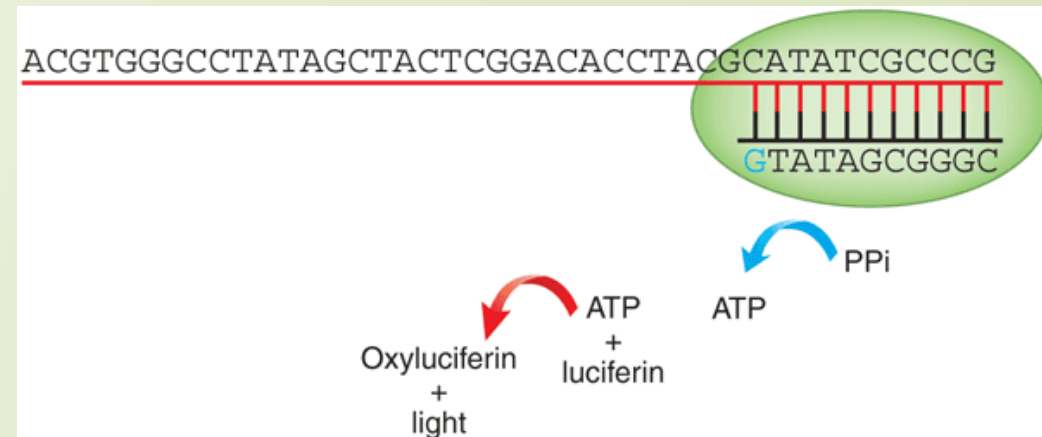
# 454 sekvenovanie a Ion torrent

- Dokáže analyzovať viac než 1 mil DNA naraz a dĺžka sekvencií je ca 700-1000 báz (1000 je max)
- DNA prichytená na malú „guličku“ na ktorej sa amplifikuje, klony DNA pokrývajú celý povrch
- Guličky umiestnené do komôrky sekvenačnej doštičky, princíp pyrosekvenovania
- Pri zachytení novej báze sa uvoľní pyrofosfát, enzymaticky spracovaný luciferázou => svetelný záblesk spracovaný citlivým detektorom (CCD kamera = charge-coupled device)
- Do reakčnej zmesi je stále pridaný iba jeden typ báze

- Pri pridaní rovnakých bází za sebou (pr. sekvencia DNA templátu je AAA) po pridaní troch T je vyžiarené 3x viac svetla
- Kvalitatívna analýza – kde na doske sa uvoľnilo svetlo
- Kvantitatívna analýza – koľko svetla sa uvoľnilo
- Do zmesi pridávané všetky báze, medzi jednotlivými krokmi sú prebytočné báze premývané
- Na podobnom princípe funguje **Ion Torrent sekvenovanie**, nemeria sa množstvo pyrofosfátu ale zmena pH



Obr.: Amplifikačná doska

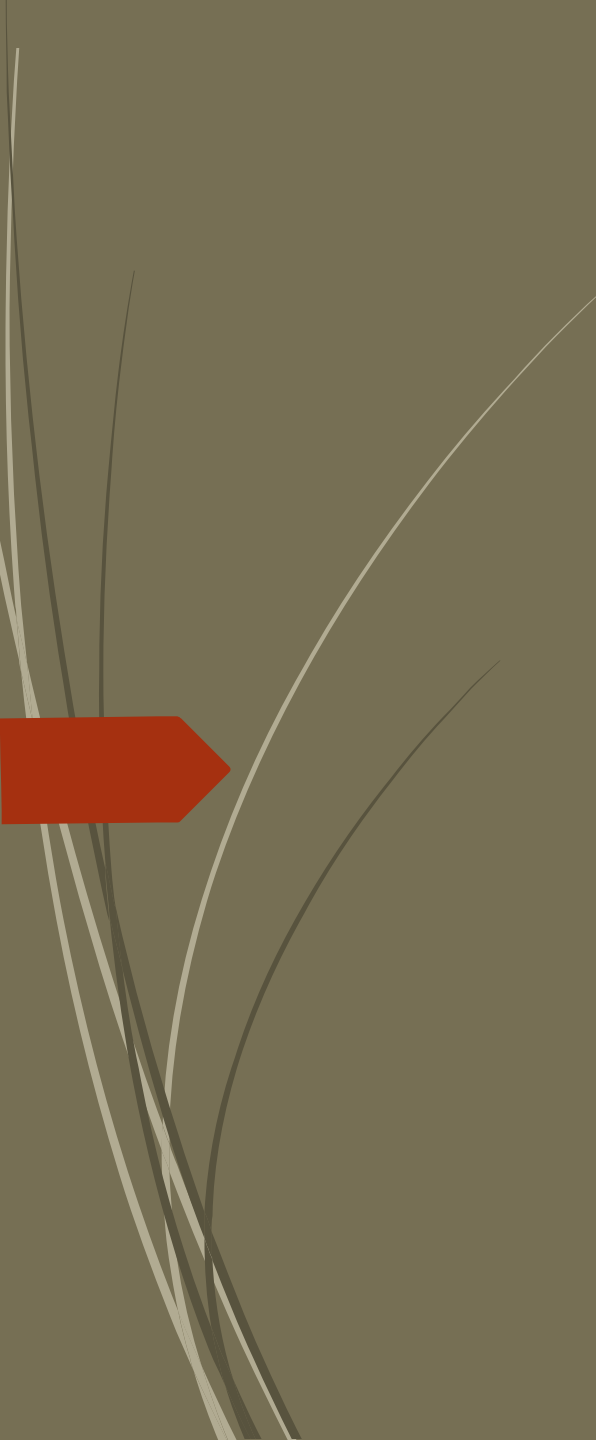




# SOLiD sekvenovanie (Sequencing by Oligonucleotide Ligation and Detection)

- Amplifikácia prebieha ako u pyrosekvenovania na mikroguličkách
- Dochádza tu k sekvenácii ligáciou
- K templátu sú pridávané sondy (krátke DNA tvorené dvojkombináciami nukleotidov = 16 rôznych sond)
- Každá sonda nesie fluorescenčné značenie (4 rôzne)
- Začiatok čítania sa vždy posúva o 1 nukleotid
- Zo znalosti sekvencie adaptéru a výsledného signálu možno odvodiť výslednú DNA sekvenciu
- Výstup podobný ako ILLUMINA
- Problémom sú palindromické úseky






# Aktívne translatovaný transkriptom (RIBO-SEQ)



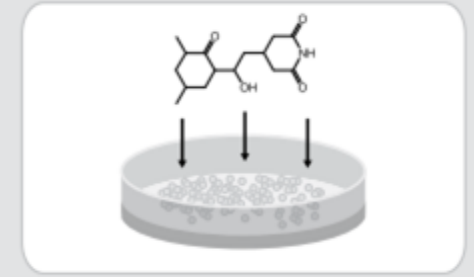
# Ribozómové profilovanie

- Sekvenovanie mRNA translatovanej vo vnútri ribozómov
  - Ribozóm chráni tzv. ribozómový odtlačok
- 

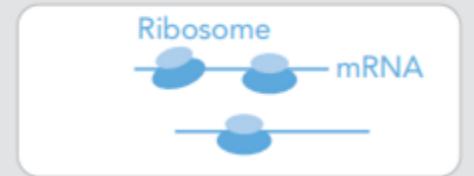
# Priebeh profilovanie

- I. Imobilizácia RNA-ribozómových komplexov
- II. Izolácia komplexov
- III. Naštiepenie nukleázou
- IV. Extrakcia mRNA fragmentov na gélovej elektroforéze
- V. Reverzná traskripcia na cDNA
- VI. Identifikácia sekvencií

## Immobilize Ribosome Complexes on mRNA



## Isolate RNA/Ribosome Complexes



## Nuclease Digestion




## Prepare Sequencing Library: Denaturation, RNA Size Selection, Polyadenylation, Reverse Transcription



## Illumina Sequencing




# Využitie

- Skúmanie počiatkov translácie
  - Zistiť rýchlosť syntézy proteínov
  - Vplyv vonkajších faktorov
  - Predikcia translačných produktov
- 

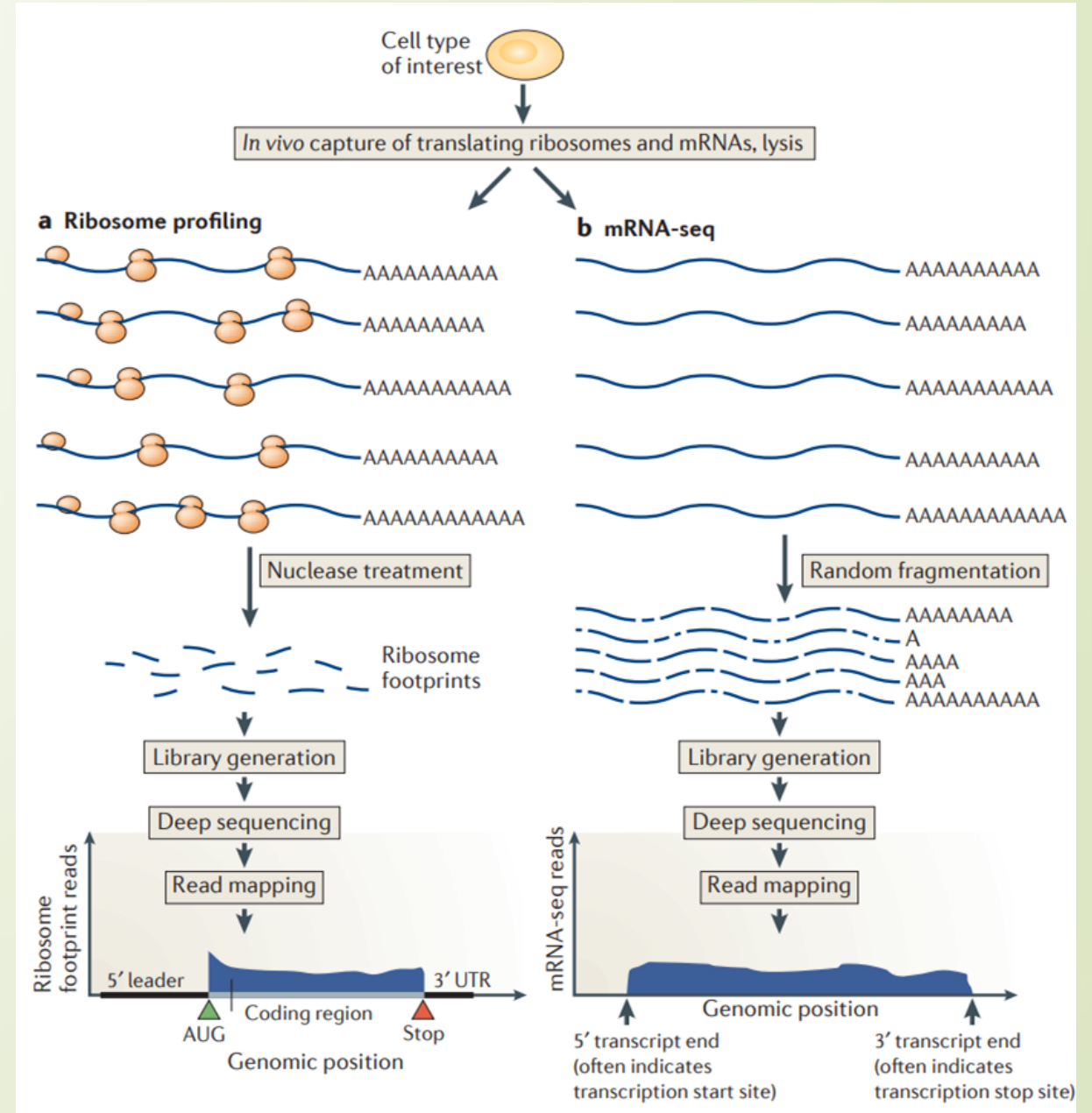


# Výhody

- Senzitivita a presnosť kvantifikácie
  - Presný popis translačných miest
  - Popis aktuálnej situácie v bunke
- 

# Nevýhody

- Riziko pomalej inhibície translácie
- Kontaminácia rRNA
- Vyžaduje väčšie množstvo materiálu ako mRNA-seq
- Zložité vyhodnocovanie





# Sekvenovanie tretej generácie




# Nedostatky sekvenovania druhej generácie

- Čítanie dlhých sekvencií
- Sekvenovanie s dostatočnou presnosťou
- Sekvenovanie z 1 molekuly DNA
  - 1 nukleotid – slabý signál
  - **Amplifikácia** – chybovosť polymerázy

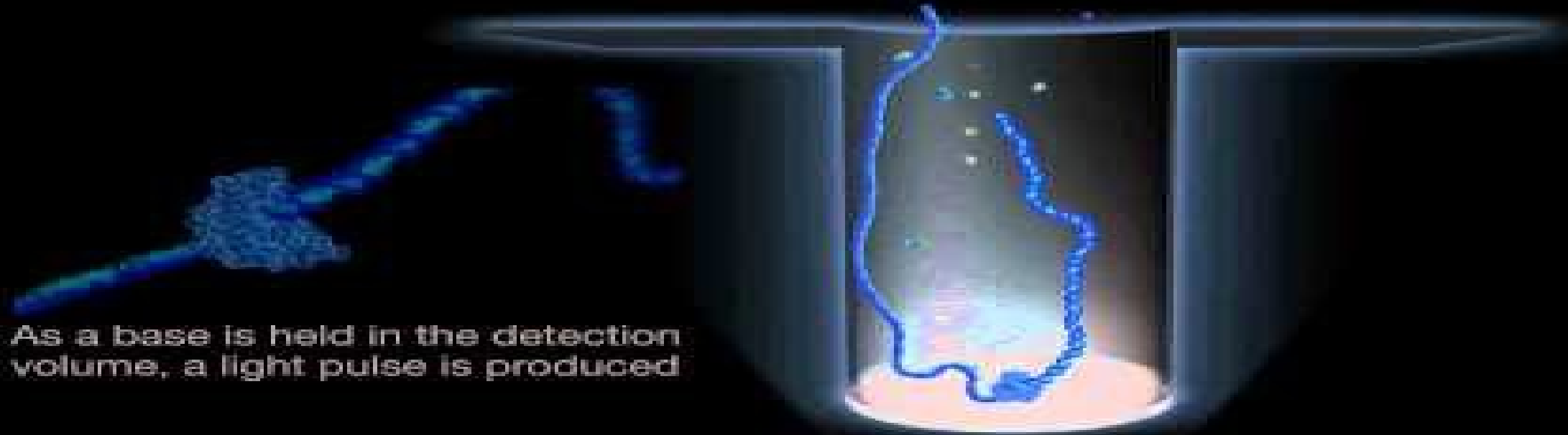




# Pacific BioSciences - SMRT

- SMRT = single molecule real time sequencing
  - Syntéza komplementárneho vlákna DNA polymerázou
  - Hairpin adaptér – cirkulárna DNA – coverage
  - 4 nukleotidy – 4 fluorofory
  - ZMW – Zero Mode Waveguids – svetlovody usmerňujúce svetelnú energiu do malého objemu → zeptoliter ( $10^{-21}$ )
- 

# Pacific BioSciences - SMRT

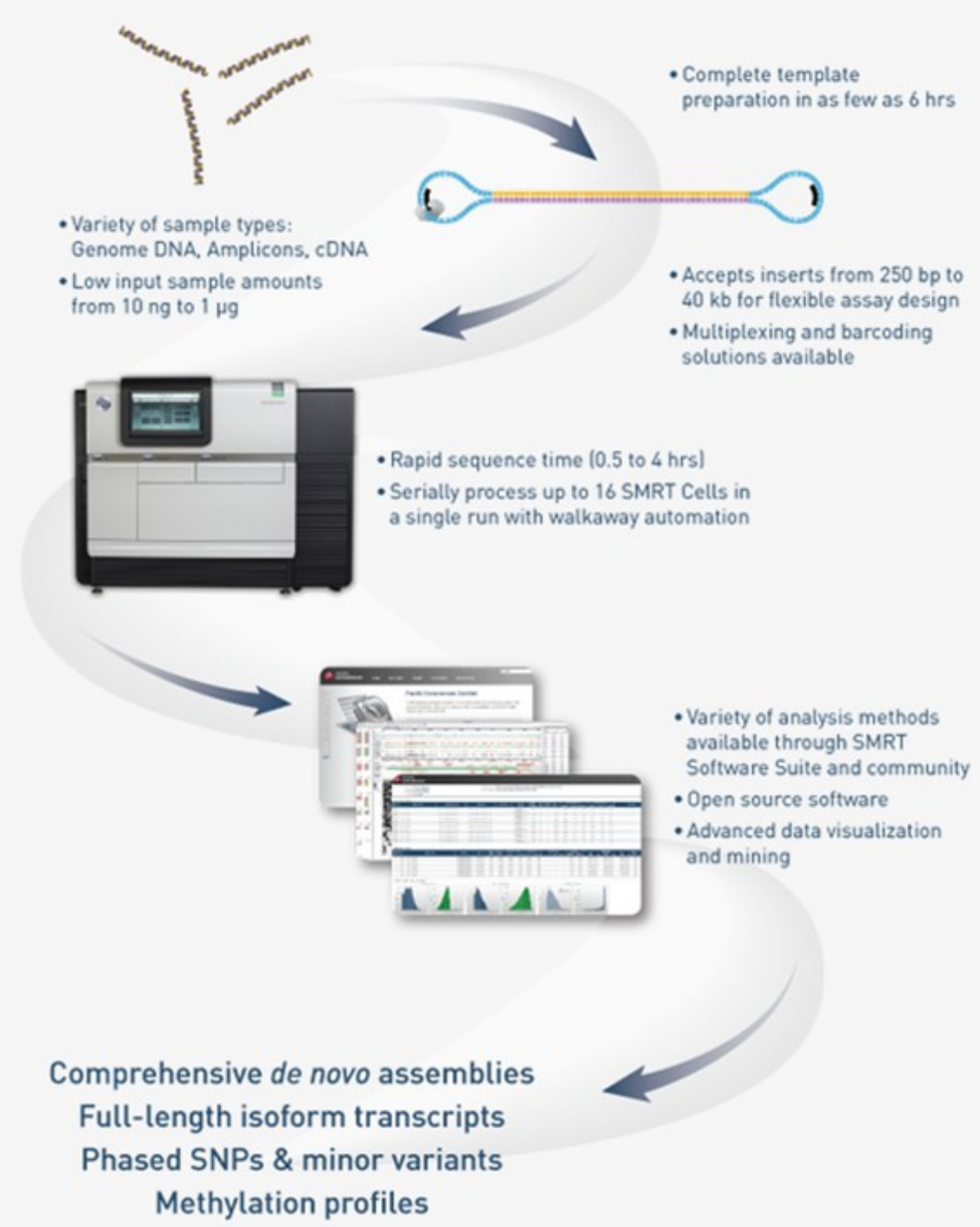


As a base is held in the detection volume, a light pulse is produced

# Pacific BioSciences

	RS II (P6-C4)	Sequel
Run time	up to 240 min	up to 240 min
Total output	~500 Mb - 1 Gb	5 Gb - 10 Gb
Output/day	~2 Gb	20 Gb
Mean read length	10 -15 kb	10 -15 kb
Single pass accuracy	~86%	~86%
Consensus (30X) accuracy	>99.999%	>99.999%
# of reads	~50k	~500k
Instrument price	~\$700k	\$350k
Run price	~\$400	~\$850






# Oxford Nanopore Technology



- Proteínová nanopórová membrána s vysokým elektrickým odporom
- DNA
  - neznačená
  - Prechod vďaka vloženému napätiu
- Meranie zmien elektrického prúdu
- MinION, GridION, PromethION
- DNA, RNA, proteíny





# Sekvenovanie tretej generácie

- Nevyužíva amplifikáciu za účelom zvýšenia signálu (vyššia presnosť – accuracy)
  - Produkuje dlhé čítania
  - Dobrá presekvenovanosť GC bohatých oblastí
  - Epigenetika
- 



Ďakujeme za pozornosť