



# Metodologie molekulární fylogeneze a taxonomie hmyzu

## Bi7770

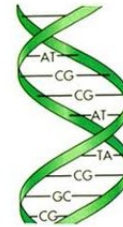
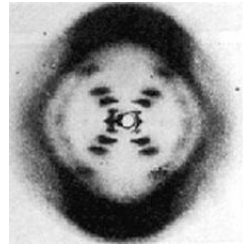
Andrea Tóthová



INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

# Metodické přístupy stanovení primární struktury DNA

- Historie:



- **1953**: struktura DNA (James Watson a Francis Crick; *Nature*, NC 1962)
- **1964**: NC 1968 rozluštění genetického kódu (Nirenberg M. a Matthaei H.)
- **1977**: sekvenování chemickou degradací – neenzymatické štěpení DNA (Allan **Maxam** a Walter **Gilbert**, *Proc.Natl.Acad.Sci.USA*)
- **1977**: sekvenování terminátorovou metodou (Frederick **Sanger**, *Proc.Natl.Acad.Sci.USA*)
- **1986**: objev PCR (**Mullis**; NC, dr.h.c. LF MU)
  - Mullis K., Faloona F., Scharf S., Saiki R., Horn G., and Erlich H. (1986). Specific enzymatic amplification of DNA *in vitro*: the polymerase chain reaction. *Cold Spring Harbor Symposium on Quantitative Biology*. 51 Pt 1:263–373

- Aktuální stav:
  - **1996: sekvenování druhé generace** – pyrosekvenování (Nyrén P., Ronaghi M. ; Stockholm), 454 sekvenování Roche/454, Illumina Genome Analyzer Iix, Life Technologies SOLiD 4 a další
  - **21. století: sekvenování třetí generace** – SMRT (single molecule real-time) – 2008; 2013 (Pacific Biosciences)??, nanotechnologie, hmotnostní spektrometrie, konfokální laserové spektrometrie a další
- Výběr technologie závisí od plánovaného využití analýz

# MAXAM - GILBERTOVA METODA

- A. M. Maxam a W. Gilbert-1977
- Sekvence DNA je vystavena určitým chemikáliím, které nasekají vlákno ve specifických místech (nukleotidech)



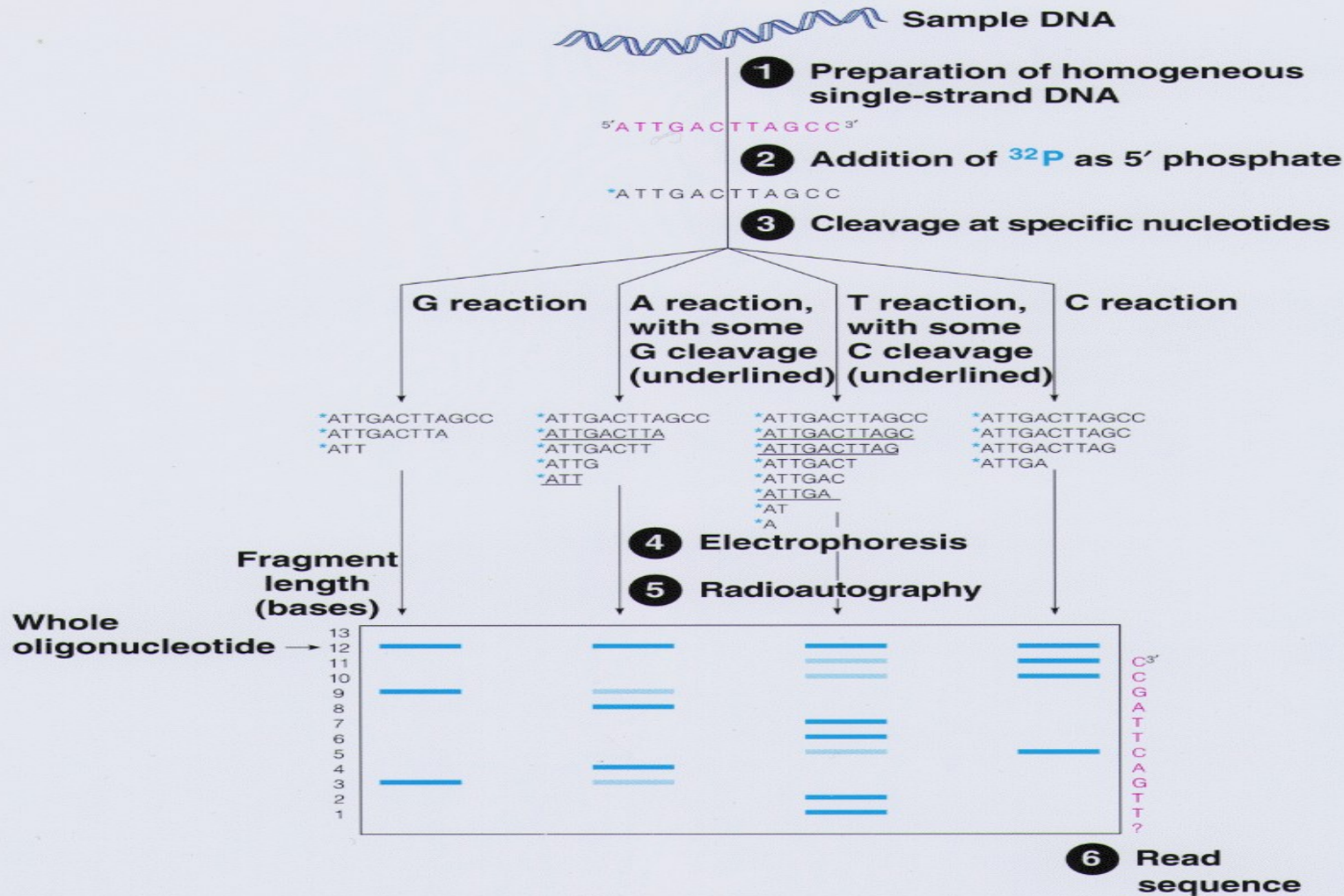
# Maxam-Gilbert

- ds DNA
- Radioaktivní značení 5 konce ds molekuly [ $\gamma$ - $^{32}\text{P}$ ] ATP (alkalická fosfatáza, polynukleotid kináza)
- Denaturace a separace vláken
- Chemická modifikace vlákna - metylace (G-dimetylsulfát, AT-kyselina mravenčí, CT-hydrazín, C-NaCl s hydrazínem)
- Selekční štěpení metylací piperidinem
- Elektroforéza
- Autoradiografie
- Vysoko radioaktivní látka, poločas rozpadu 14 dní, několik desítek bazí na gelu
  
- Dnes se již nepoužívá

**Table 10.1 Specific Base Reactions in Maxam-Gilbert Sequencing**

Chain breaks at:	Base Modifier	Reaction	Time (min at 25°C)
G	Dimethylsulphate	Methylates G	4
G + A	Formic acid	Protonates purines	5
T + C	Hydrazine	Splits pyrimidine rings	8
C	Hydrazine + salt	Splits only C rings	8

**Figure 4A.4 Sequencing an oligonucleotide by the Maxam-Gilbert method**



# Sanger – Coulson – (Nicklen)

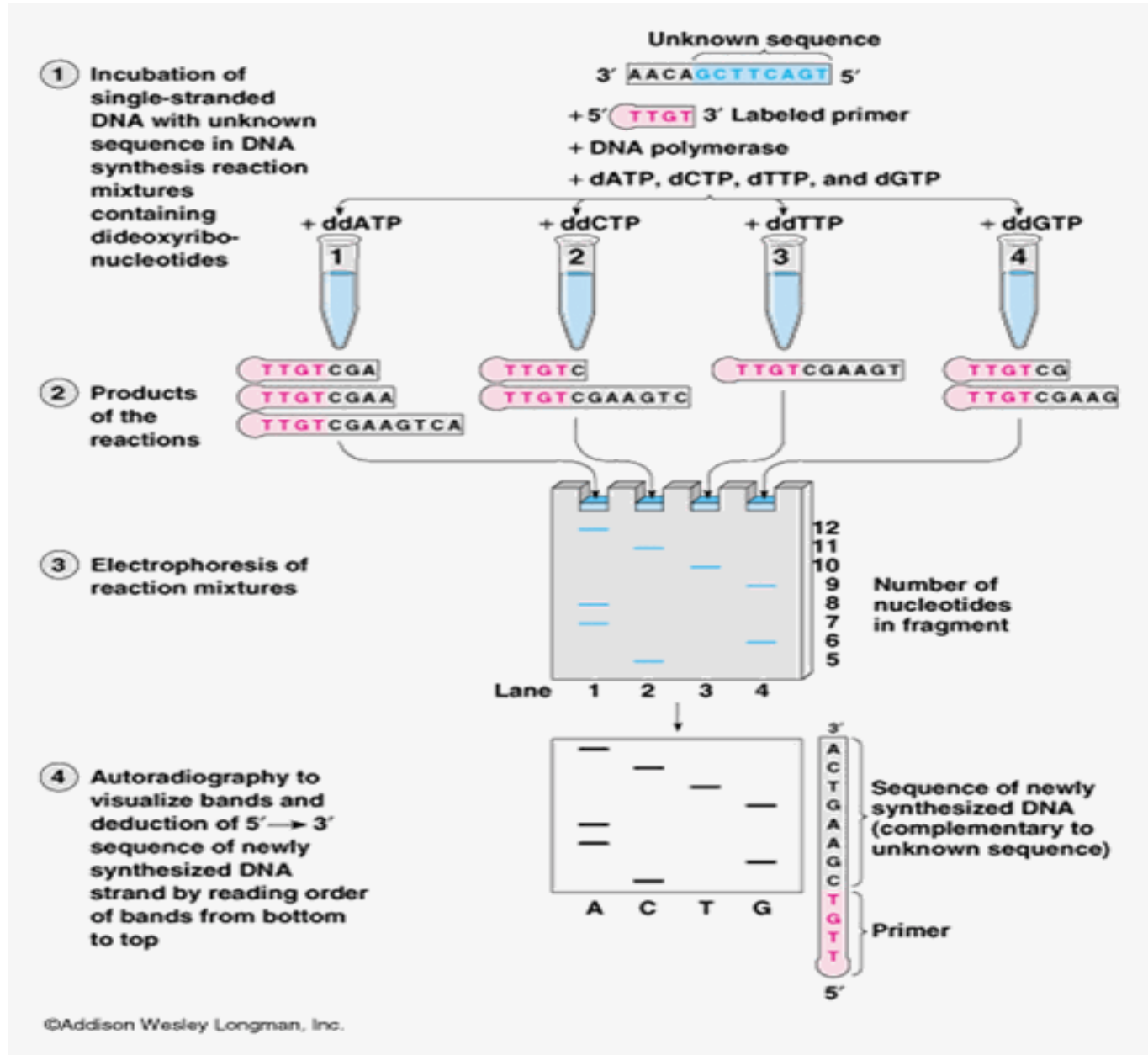
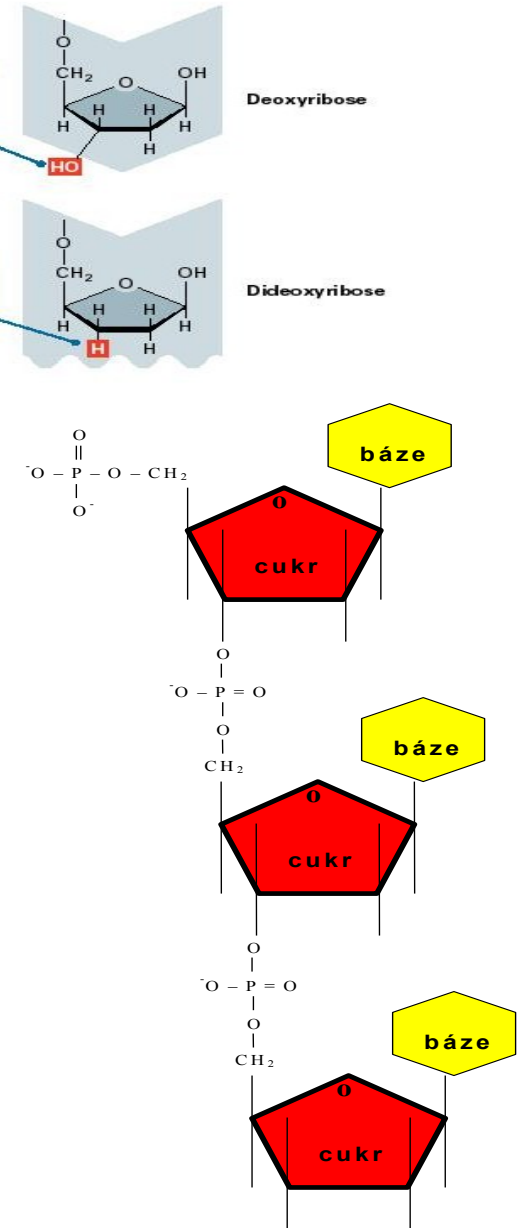
- Syntéza komplementárního vlákna z ss DNA templátu – restriční fragmenty (reakční pufr; dNTPs; DNA pol I – Klenow fragment, T7 DNA pol, Taq DNA pol)
- Radioaktivní značení vznikající molekuly [ $\alpha$ - $^{32}\text{P}$ ]dATP, [ $\alpha$ - $^{35}\text{S}$ ]dATP (alkalická fosfatáza, polynukleotid kináza)
- Složitá inkubace ve vícero krocích (annealing, amplifikace)
- Přidání terminátorů ddNTPs – postupná terminace (poměr dTTP:ddTTP)
- Nutnost terminace reakce (formaldehyd, chelatační činidla, barviva)
- Elektroforéza – 6-8% PAGE + TBE (TEMED, močovina, bis-akrylamid)
- Autoradiografie vysušeného gelu (24-48 hod) a následná detekce
- Radioaktivní látka, 30 nt od primeru – max 400 nt
  
- Prodloužení délky čteného vlákna
- Modifikace chemie (nahrazení radioaktivního značení)
- Modernizace metod analýzy syntetizovaných fragmentů, automatizace



# SANGEROVA metoda

- Nejpoužívanější způsob sekvenování
- Objeveno Frederickem Sangerem - 1977
- Nobelova cena - 1980





# Nahrazení radioaktivního značení fluorescenčními barvičkami

- Oligo – hybridizační sondy
- Radioaktivní nuklidy:  $^3\text{H}$ ,  $^{32}\text{P}$ ,  $^{35}\text{S}$ ,  $^{125}\text{I}$
- Detekce záření na scintilačním fotometru nebo autoradiografií na RTG filmu
- 3 i 5', celé fragmenty: DNA I pol Klenow fragment, T4 DNA pol, T4 polynukleotid kináza, alkalická fosfatáza (BAP, CIP)
  
- Fluorochromy – široké spektrum a využití v biologii
- 6-FAM, NED, PET, ROX, HEX, VIC, JOE, LIZ, BigDyes
- Detekce fluorescenčního záření: laser + indukční a emisní spektrum fluorochromu – excitace fluorochromu – detekce přes optický systém (CCD kamera; vzduchem chlazený CCD čip) – virtuální filtrování – vlnová délka se odečítá pouze v oblasti spektrálního maxima
- Ovlivnění mobility fragmentů

Radionuklid	Poločas rozpadu	Typ	Záření Max energie	Účinnost	
				voda	vzduch
$^{32}\text{P}$	14,3 dní	$\beta$	1,71 MeV	600 cm	7,8 mm
$^{35}\text{S}$	87,4 dní	$\beta$	0,167 MeV	26 cm	340 $\mu\text{m}$
$^{125}\text{I}$	60 dní	$\gamma$	0,35 MeV	1,5 cm	20 $\mu\text{m}$
$^3\text{H}$	12,43 roka	$\beta$	0,018 MeV	0,5 cm	5,5 $\mu\text{m}$

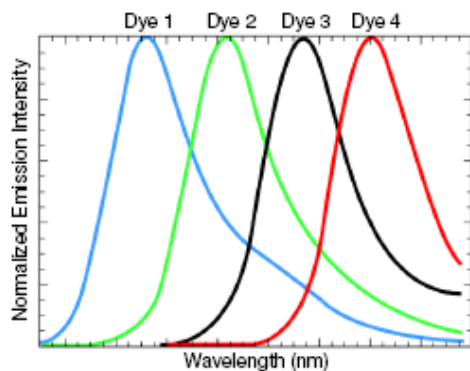


Figure 7 Emission spectra of the four BigDye dyes, where Dye 1 = Big-d110, Dye 2 = R6G, Dye 3 = Big-dTAMRA, and Dye 4 = Big-dROX

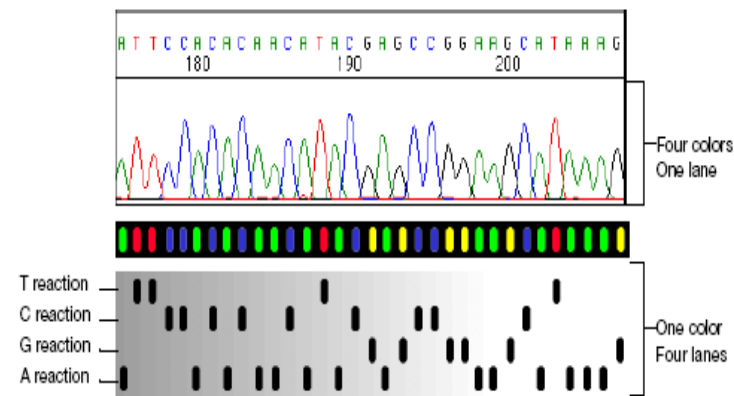
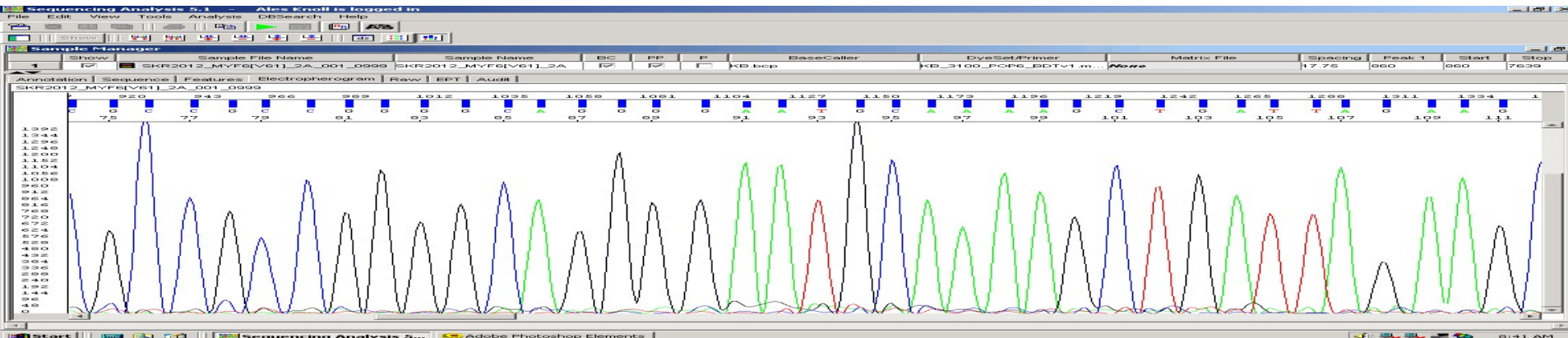


Figure 8 Fluorescent sequencing compared with radioactive sequencing



# 454- pyrosekvenování

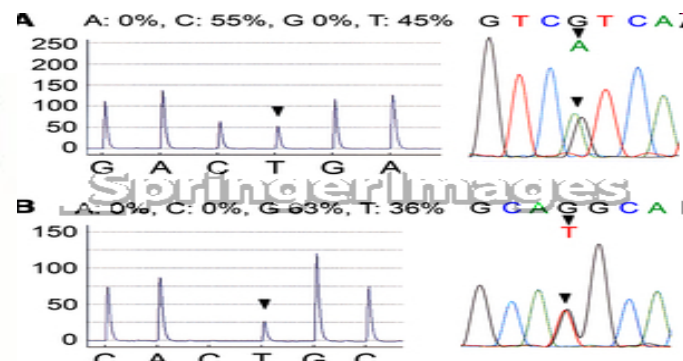
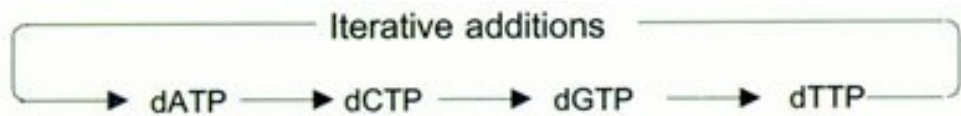
- 454 Life Sciences - komercializace
- Sekvenování druhé generace
- Celogenomové
- 4 enzymový systém
  - DNA polymeráza
  - ATP sulfuryláza
  - Luciferáza
  - Apyráza
- **Detekce nukleotidu začleněného do nově-syntetizovaného vlákna pomocí luminiscence**
- Roche – GS FLX Titanium, GSJunior

## SOLID chemie

- Solid – Life Technologies- Applied Biosystems
- **Detekce fluorescence**
- Kombinace dvou nukleotidů, vícero primerů
- Galaxy software

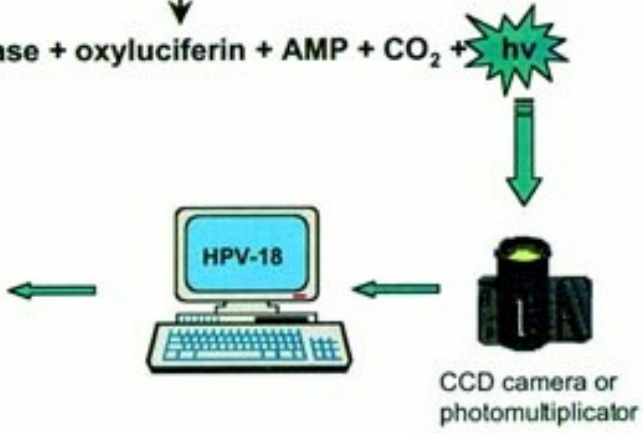
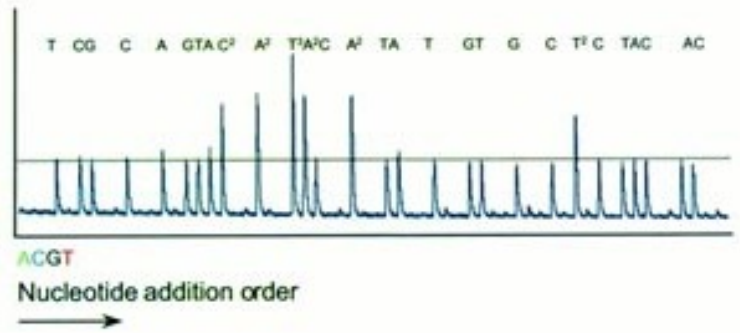
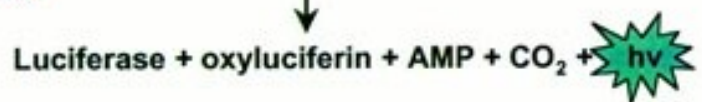
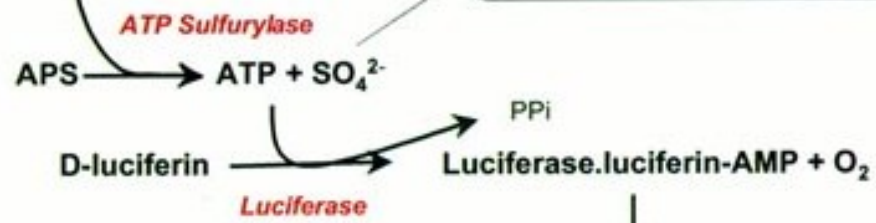
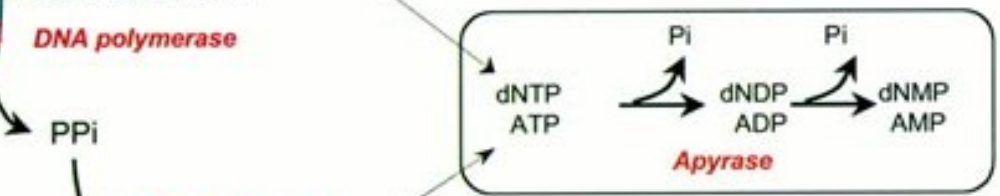
## Ion Torrent technologie

- Nejnovější technologie
- Celogenomový
- Personální využití – PGM systém
- **Detekce změny pH**

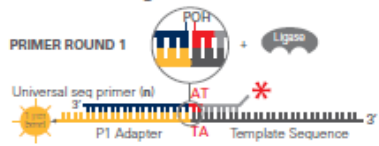


**...AGCGTCATGGTTAAATTG...**  
**TCGCAGTACC** DNA polymerase

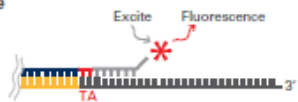
Primed Template



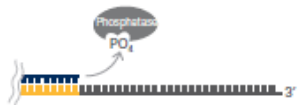
### 1. Prime and Ligate



### 2. Image



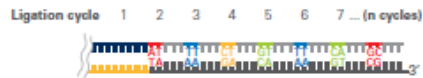
### 3. Cap Unextended Strands



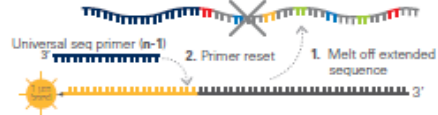
### 4. Cleave off Fluor



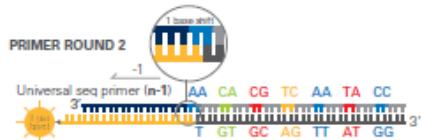
### 5. Repeat steps 1-4 to Extend Sequence



### 6. Primer Reset



### 7. Repeat steps 1-5 with new primer



### 8. Repeat Reset with , n-2, n-3, n-4 primers

	Read Position	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35		
1	Universal seq primer (n)	••••••••••	••																																				
2	Universal seq primer (n-1)		••••••••••	••																																			
3	Universal seq primer (n-2)			••••••••••	••																																		
4	Universal seq primer (n-3)				••••••••••	••																																	
5	Universal seq primer (n-4)					••••••••••	••																																

• Indicates positions of interrogation Ligation Cycle

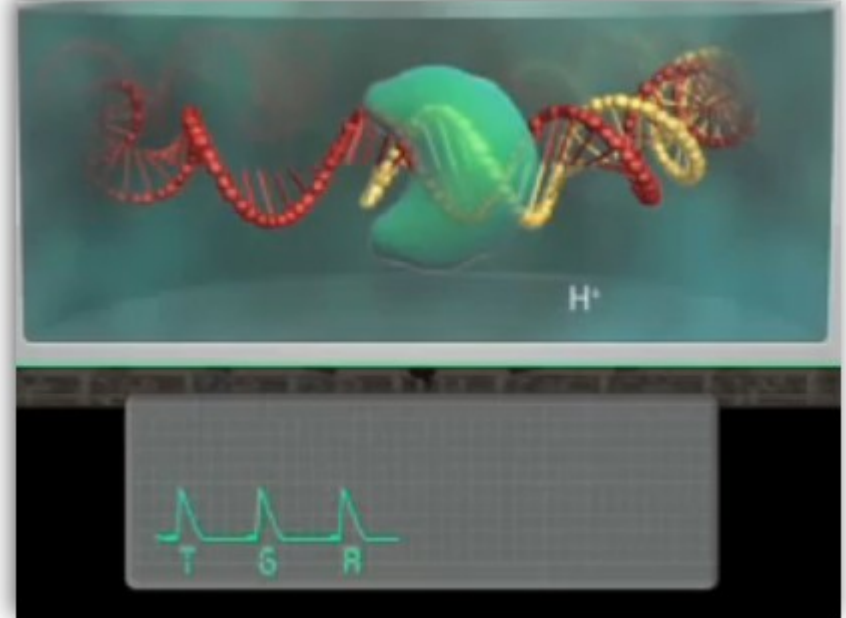
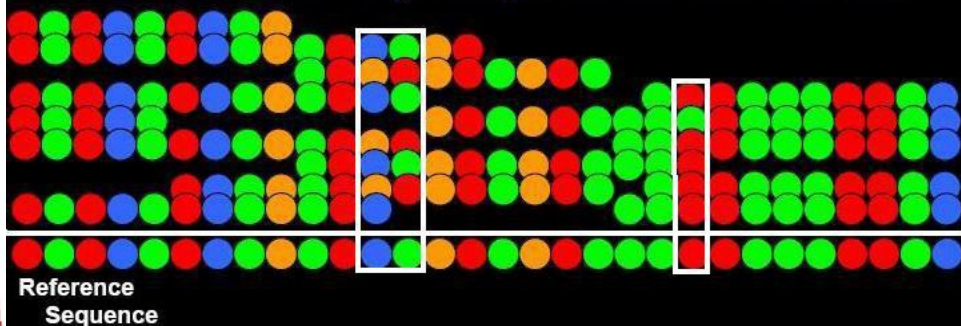
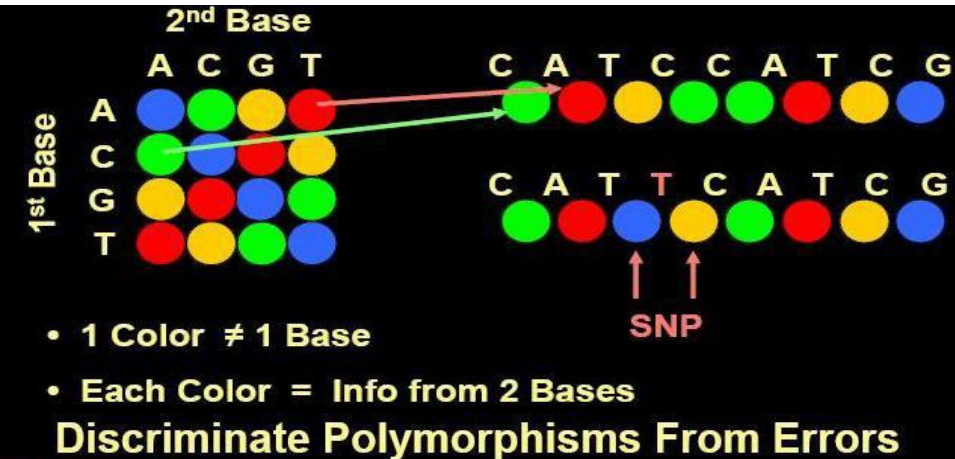
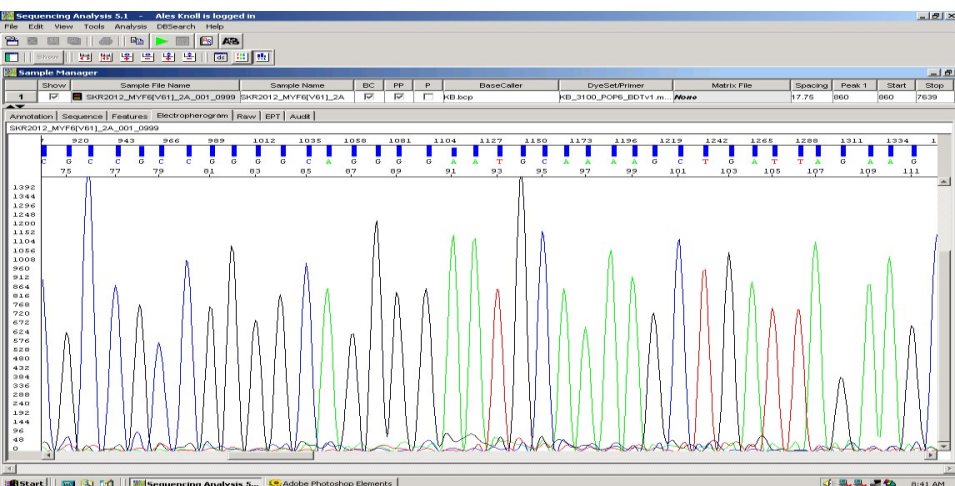
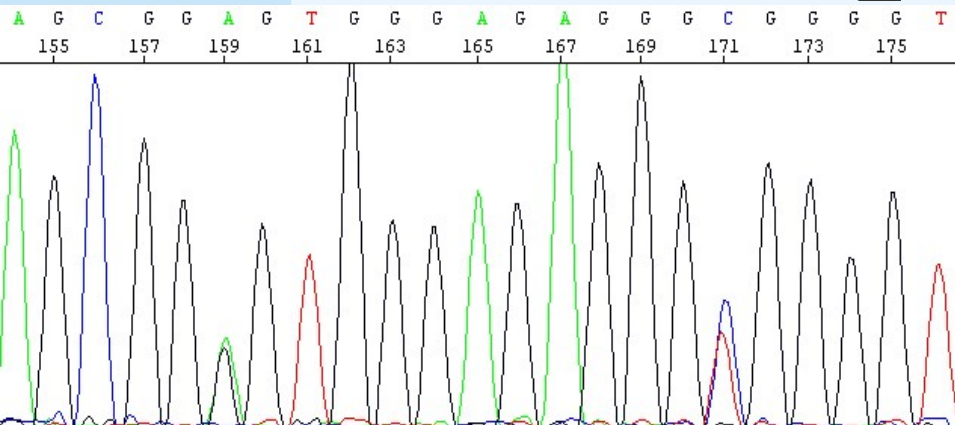
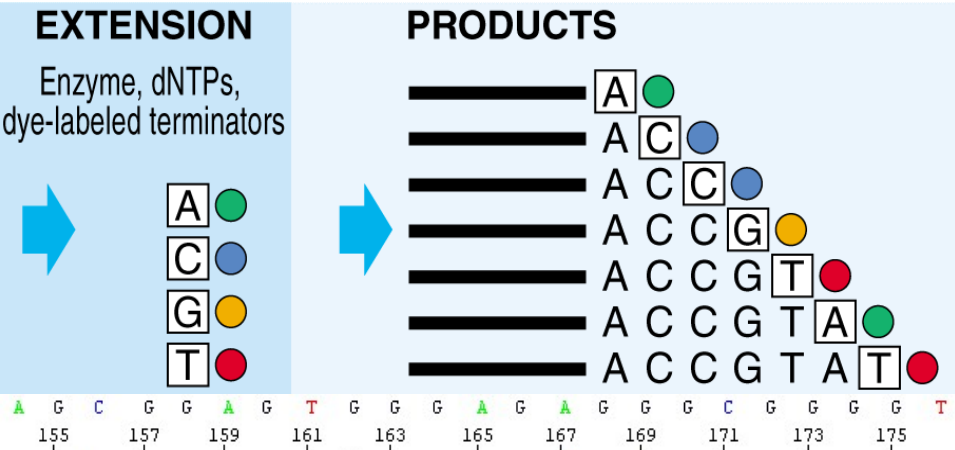


Figure 1: SOLiD™ System – Sequencing by ligation using di-base labeled probes



## SOLiD read and quality format

```
>1279_22_371_F3
T0133231.3220220312223121011022022023113131.1121021
>1279_35_1729_F3
T10110133100320311131212213221102021310210310331113
```

```
>1279_22_371_F3
23 20 5 13 11 22 28 -1 16 8 14 21 8 16 5 18 23 2 27 18 5 8 3
>1279_35_1729_F3
24 18 13 23 7 7 20 7 13 7 5 8 3 11 6 26 28 20 5 25 10 8 5 3
```



# Přístrojové vybavení

- Klasická vertikální elektroforéza
- Gelové sekvenátory – plošné PAGE gely, ruční příprava
  - Applied Biosystems, Bio-Rad, Beckman
- Kapilární sekvenátory – komerčně dostupné polymery – denaturační / nendenaturační, automatické „nanášení“ vzorků, elektroforéza v kapiláře
  - (Applied Biosystems, Amersham Pharmacie BioTech, Beckman)
    - ABI PRISM 310, 3100, 3100-*Avant*, 3730, ABI 3500 Genetic Analyzer; MegaBACE 500, 750, 1000, 15000, 4000; CEQ 8000, CEQ 8800

- 1986 – 1. sekvenátor (Perkin-Elmer)
- Gelové poloautomatické sekvenátory
- Vývoj úrovně automatizace – princip kapilární elektroforézy
- 1 – 4 – 8 – 16 – 24 – 96 kapilár
- Vývoj ovládacího softwaru
- Vývoj kvality, kvantity a rychlosti zpracování vzorků
- Sekvenování **první generace** – **Sanger** – do 1000 bp
- Sekvenování **druhé generace** – **celogenomové** – paralelní sekvenování amplifikovaného DNA templátu – 3 Gbp/run → 20 Gbp/run → 100 Gbp/run
- Sekvenování **třetí generace** – rychlé a dlouhé čtení – **sekvenování jedné molekuly**



ABI PRISM®  
310 Genetic



Applied Biosystems  
3130 Genetic Analyzer



Applied Biosystems  
3130xl Genetic Analyzer



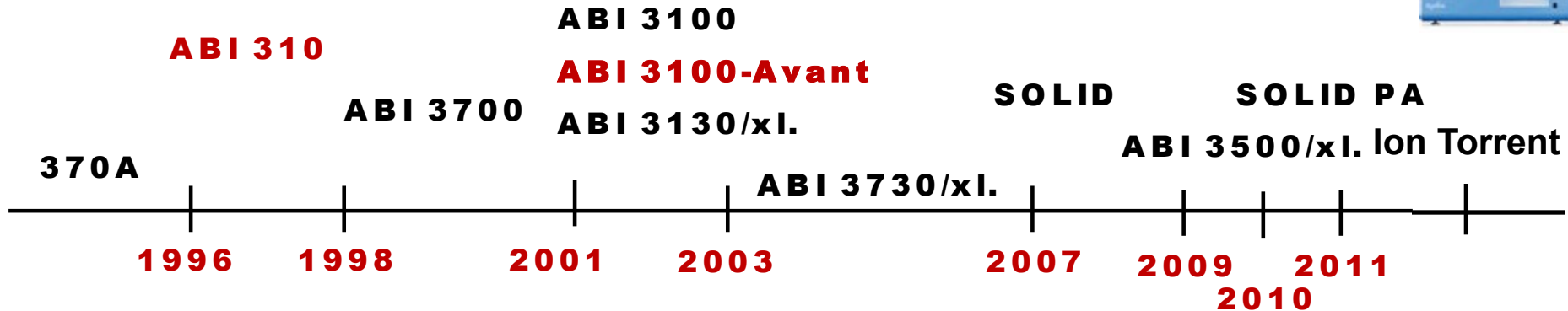
Applied Biosystems  
3730 DNA Analyzer



Applied Biosystems  
3730xl DNA Analyzer



**Key applications:** *De novo* sequencing • Resequencing • Mutation/heterozygote detection  
SNP genotyping • Relative fluorescent quantitation • Microsatellite analysis • AFLP® analysis  
Methylation analysis • T-RFLP analysis • MLST • BAC fingerprinting • SAGE™





**SOLID 3 Plus System  
Applied Biosystems**

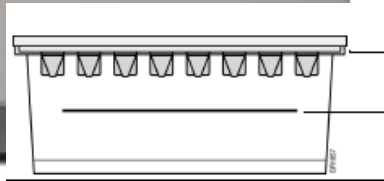
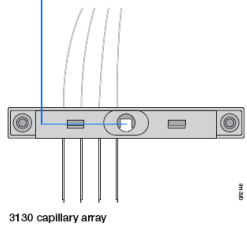


**Illumina Genome Analyzer IIx**

**CEQ 8000 Beckman Coulter**



**Genome Sequencer FLX System  
Roche Diagnostics**



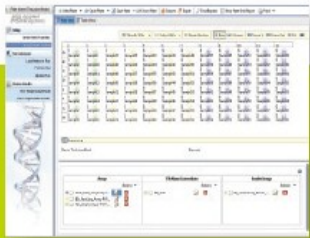
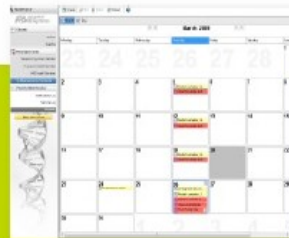


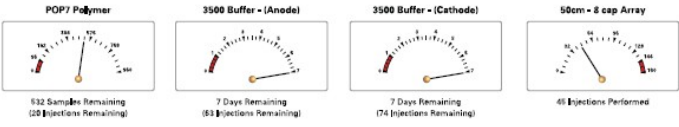
Plate Setup



Monitor Run



Maintenance Scheduling Calendar



Capillary Array



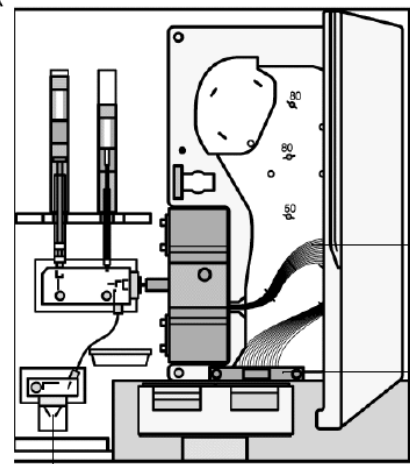
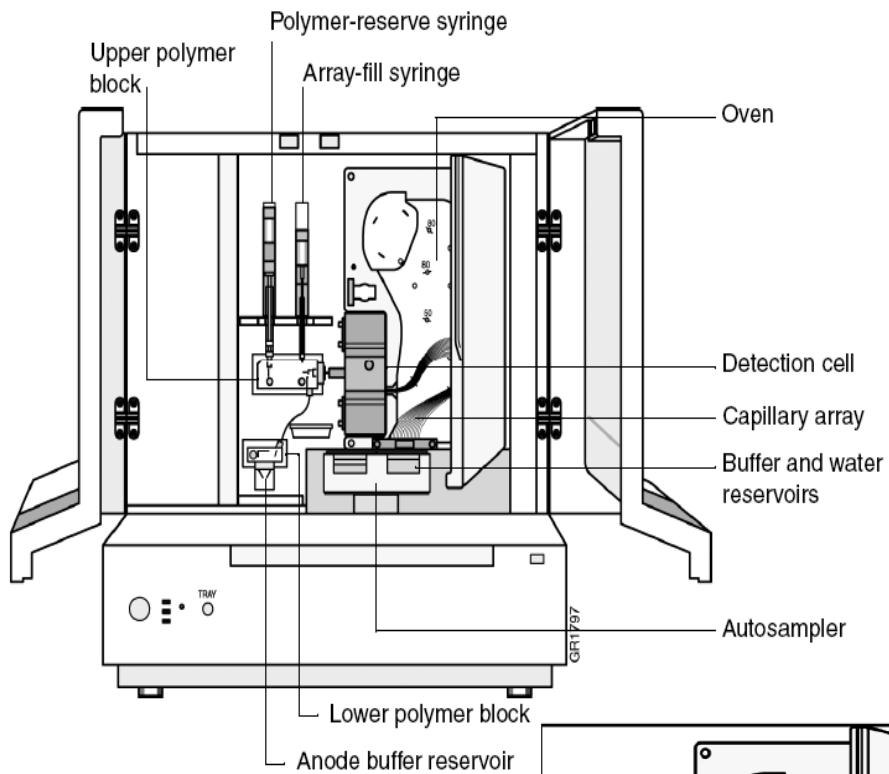
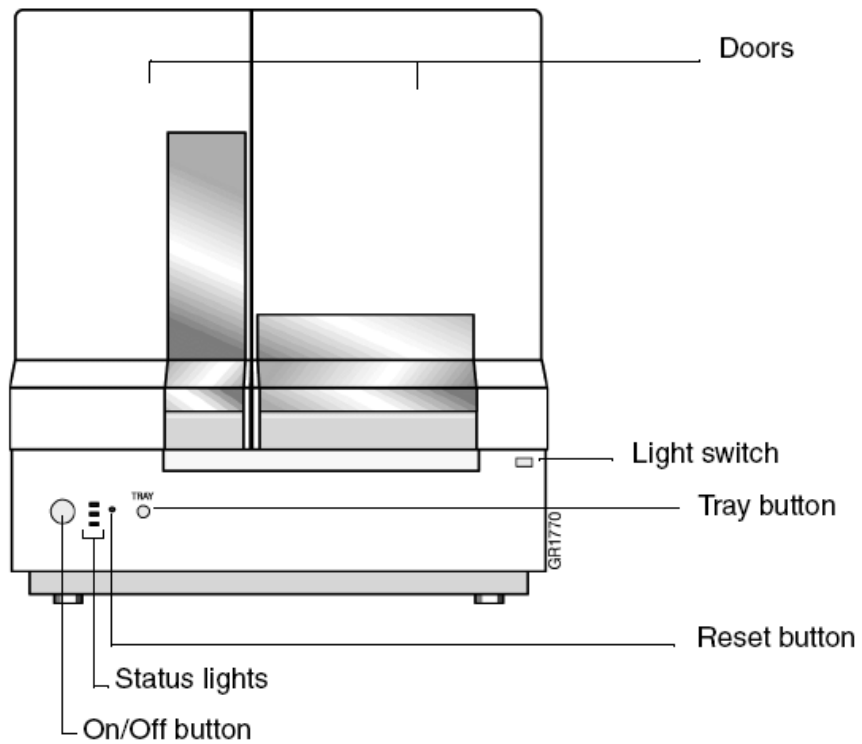
Cathode Buffer Container



Anode Buffer Container

# ABI Prism 3100-Avant Genetic Analyzer a metoda sekvenování

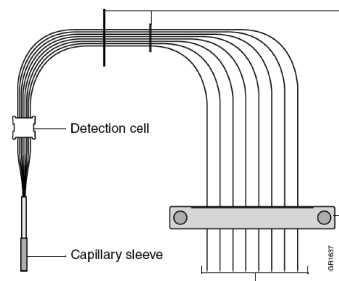
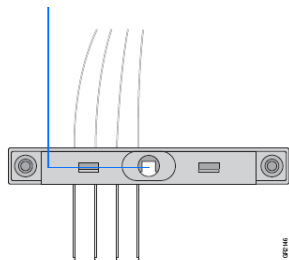
- Automatický autosampler
- Plata pro 96 vzorků
- 4 kapiláry – paralelní runy
- Píčka
- Detekční prostor (laser, optika, CCD kamera, okénko kapiláry)
- Katody a anoda
- Elektroforetický pufr
- Dávkování polymeru – systém stříkaček
- Ovládací software





# Možnosti ABI Prism 3100- Avant

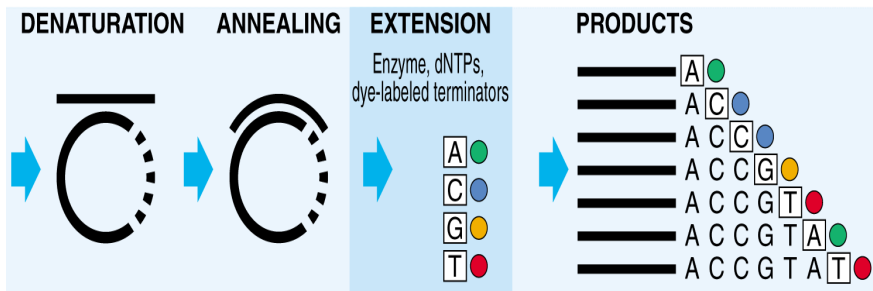
Capillary Length	Polymer	Application	Run Module	Run Time	Resolution	Performance	KB Q20 LOR**
22 cm	POP-4™	High throughput SNP analysis	SNP22_POP4	15 min	250 bp	0.50 SD†	--
		High throughput, small size fragment analysis	FragmentAnalysis22_POP4	20 min	400 bp		
36 cm		Standard SNP analysis	SNP36_POP4	30 min	250 bp	0.15 SD†	--
		Standard fragment analysis	FragmentAnalysis36_POP4	45 min	400 bp		
		Ultra rapid sequencing	UltraSeq36_POP4	40 min	500 bp	98.5% *	400 bp
POP-6™	Rapid sequencing	RapidSeq36_POP6	1 hr	500 bp	500 bp		
50 cm	POP-4™	Standard sequencing	StdSeq50_POP4	1 hr 40 min	--	-	600 bp
		Long fragment analysis	FragmentAnalysis50_POP4	1 hr 5 min	500 bp		0.15 SD†
	POP-6™	Long fragment analysis	FragmentAnalysis50_POP6	1 hr 30 min		650 bp	
		Standard sequencing	StdSeq50_POP6	2 hr 30 min	600 bp		
80 cm	POP-4™	Long read sequencing	LongSeq80_POP4	3 hr 40 min	950 bp		700 bp



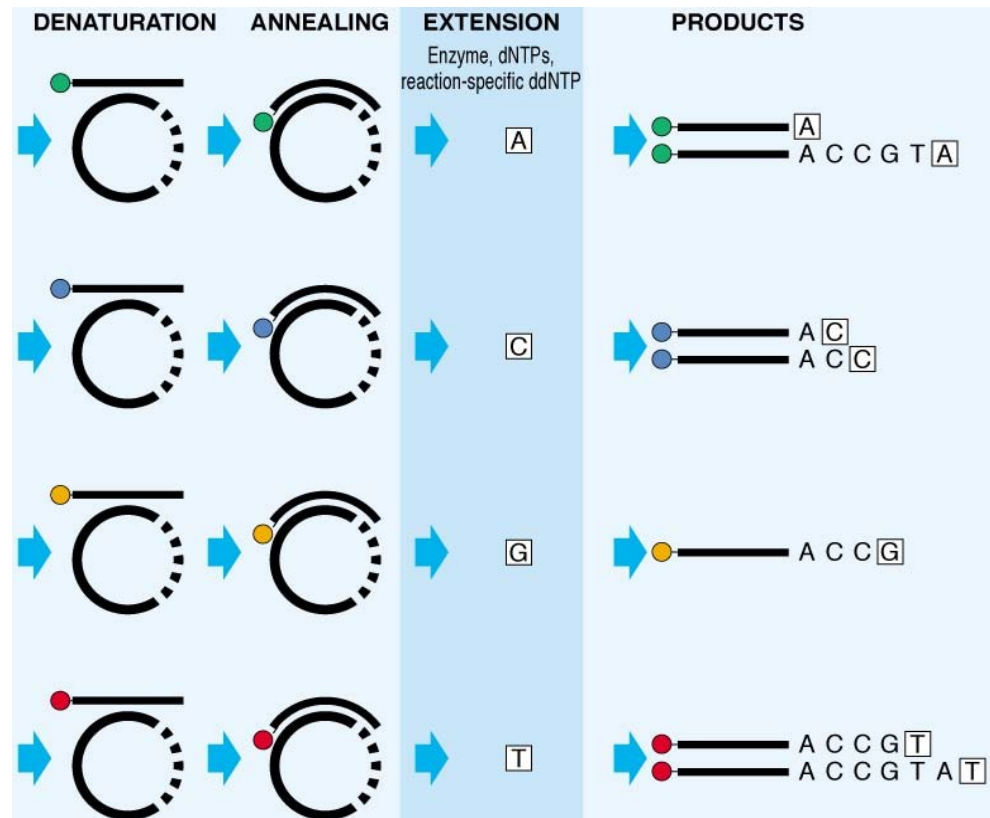
# Princip metody sekvenování na ABI

Postup: purifikace DNA templátu – PCR amplifikace a přečištění produktu – cyklické sekvenování – přečištění sekvenační reakce – kapilární elektroforéza – analýza dat

## Dye terminator chemistry



## Dye primer chemistry



Templát – koncentrace, čistota

Primer - specifita

Pufr

dNTP

ddNTP

Taq DNA polymeráza



# Princip práce stroje

- Ovládání softwarem – DataCollection
- Elektrokinetická injektáž – objem vzorku zůstává nezměněn
- Pohyb fragmentů DNA gelem (polymerem) v elektrickém poli
- Definování správných podmínek runu (modul) – bere v úvahu směs použitých fluorochromů, délku kapiláry (50-80 cm), polymer (POP6)...
- Detekce emise fluoroforu při přechodu detekčním okénkem
- Převod dat: spektra – raw data – elektroforetogram
- Vyhodnocení a analýza

**SequencingAnalysis Plate Editor**

Plate Name: SeqA\_2  
 Plate Sealing: Septs

Well	Sample Name	Comment	Priority	Res
D07	LRS		100	Seq
C07	LRS		100	Seq
D07	LRS		100	Seq
E07	LRS		100	Seq
F07	LRS		100	Seq
G07	LRS		80	Seq
H07	LRS		100	Seq
A08	LRS		100	Seq
B08	LRS		100	Seq
C08	LRS		100	Seq
D08	LRS		100	Seq

Description:

---

**Foundation Data Collection Version 3.0** - No User is logged in

GA Instruments > ga3130d > DawdPT4 > Run Scheduler > Run View

Run	RunID	Run Type	Module	Status
1	Run_Dawd PT4_20...	Regular	SMSeq50_POP6_1	Validated
2	Run_Dawd PT4_20...	Regular	SMSeq50_POP6_1	Validated
3	Run_Dawd PT4_20...	Regular	SMSeq50_POP6_1	Validated
4	Run_Dawd PT4_20...	Regular	SMSeq50_POP6_1	Validated
5	Run_Dawd PT4_20...	Regular	SMSeq50_POP6_1	Validated
6	Run_Dawd PT4_20...	Regular	SMSeq50_POP6_1	Validated

GA\_2 has been linked to Day 0

Application: SequencingAnalysis

Sample Sheet

Done

---

**Run Module Editor**

Run Module Description

Name:

Type: REGULAR

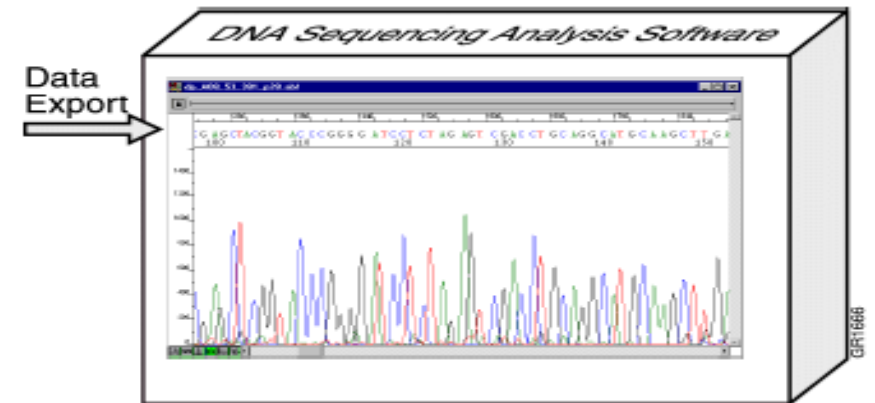
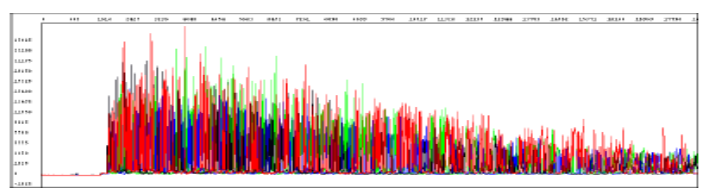
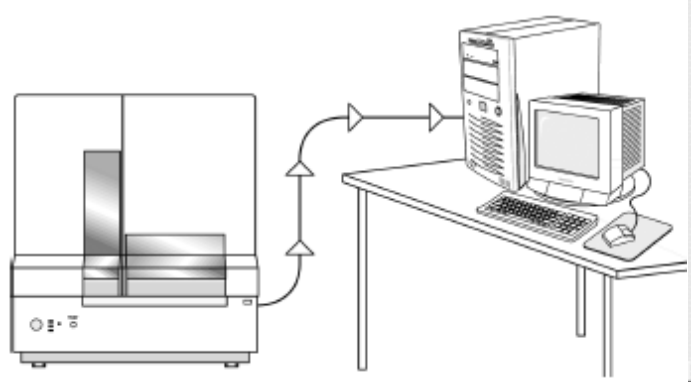
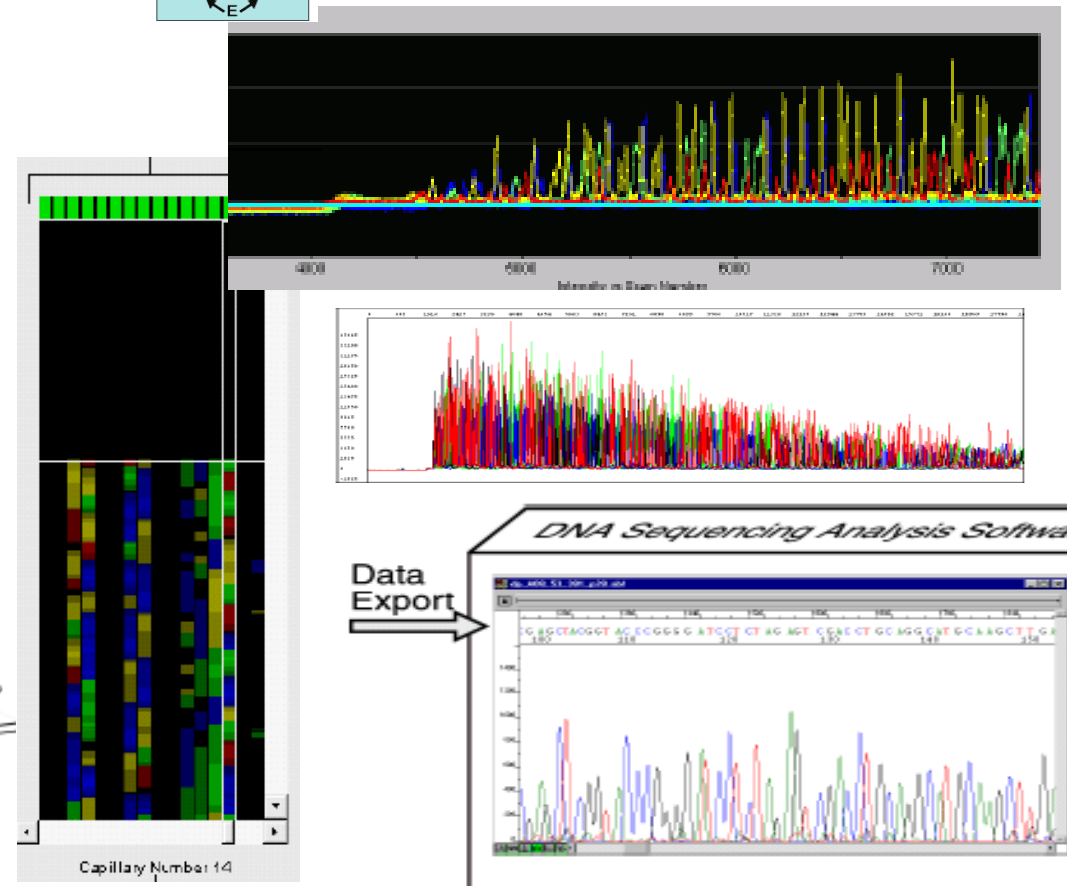
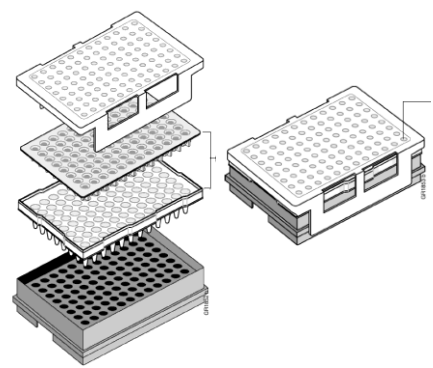
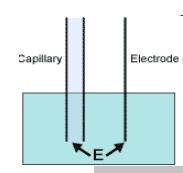
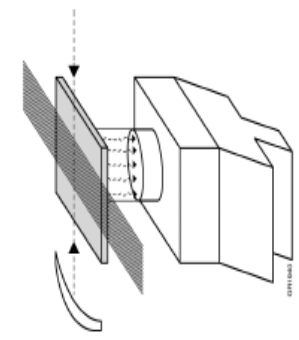
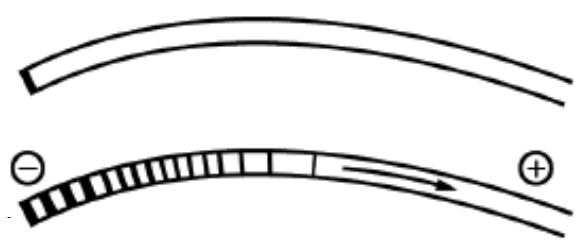
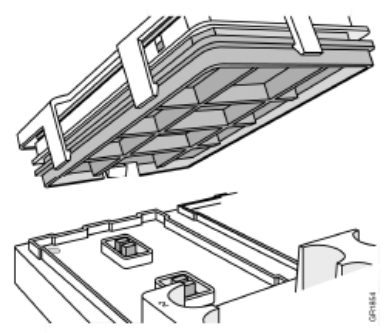
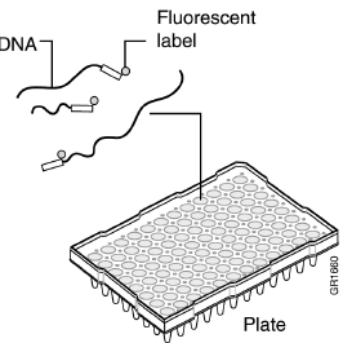
Template: SMSeq50\_POP6

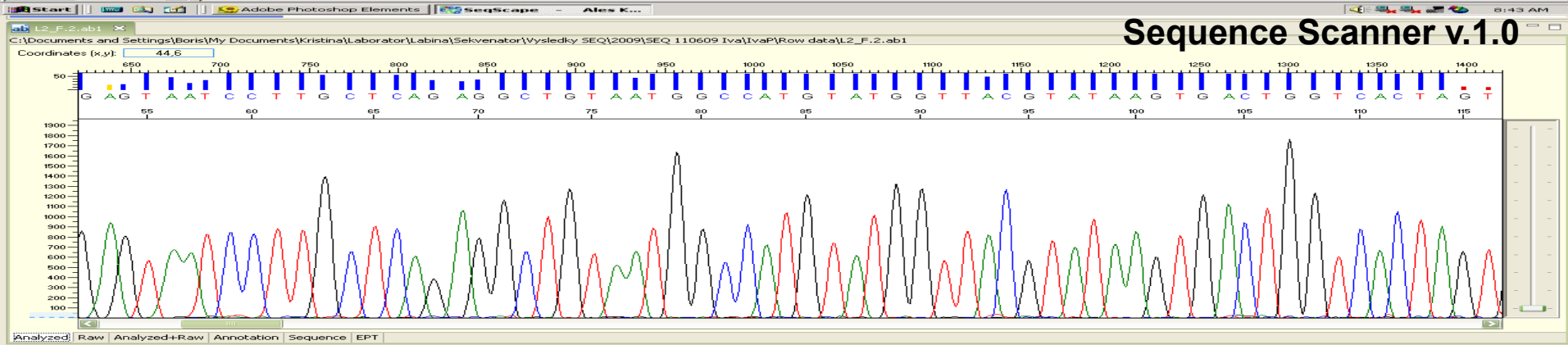
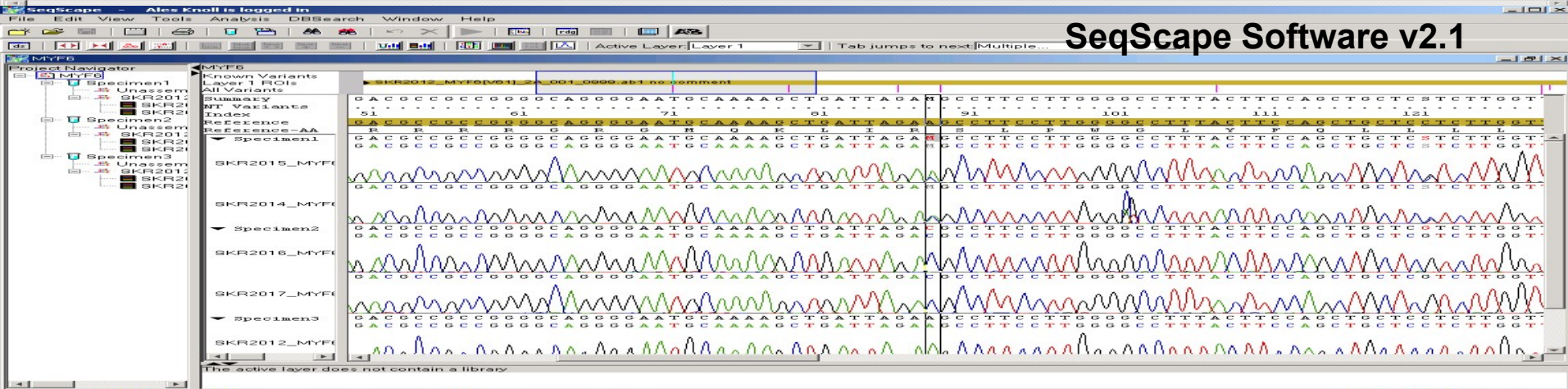
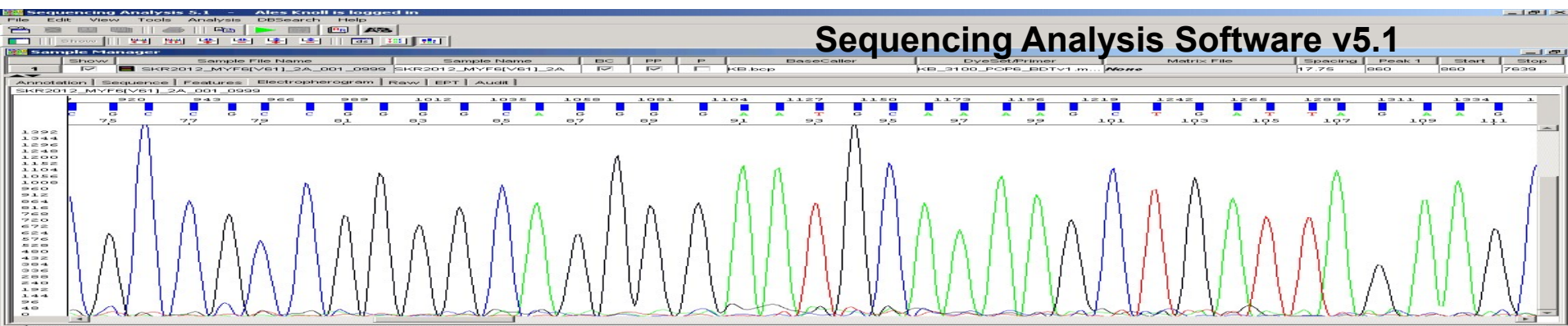
Description:

Run Module Settings

Name	Value	Range
Overn_Temperature	50	18..65 Deg. C
PreRun_Voltage	12.2	0..15 kVolts
Pre_Run_Time	180	1..1000 sec.
Injection_Voltage	1.0	1..15 kVolts
Injection_Time	22	1..600 sec.
Voltage_Number_Of_Steps	10	1..100 nk
Voltage_Step_Interval	60	1..60 sec.
Data_Delay_Time	1200	1..3600 sec.
Run_Voltage	12.2	0..15 kVolts
Run_Time	6500	300..14000 sec.

Ok Cancel





# Aplikace a využití metody sekvenování

- Sekvenování *de novo*
- Resekvenování
- detekce mutací – SNP, INDELs (nutné rozklonování PCR produktů k detekcím jednotlivých alel)

### X5 s alelami

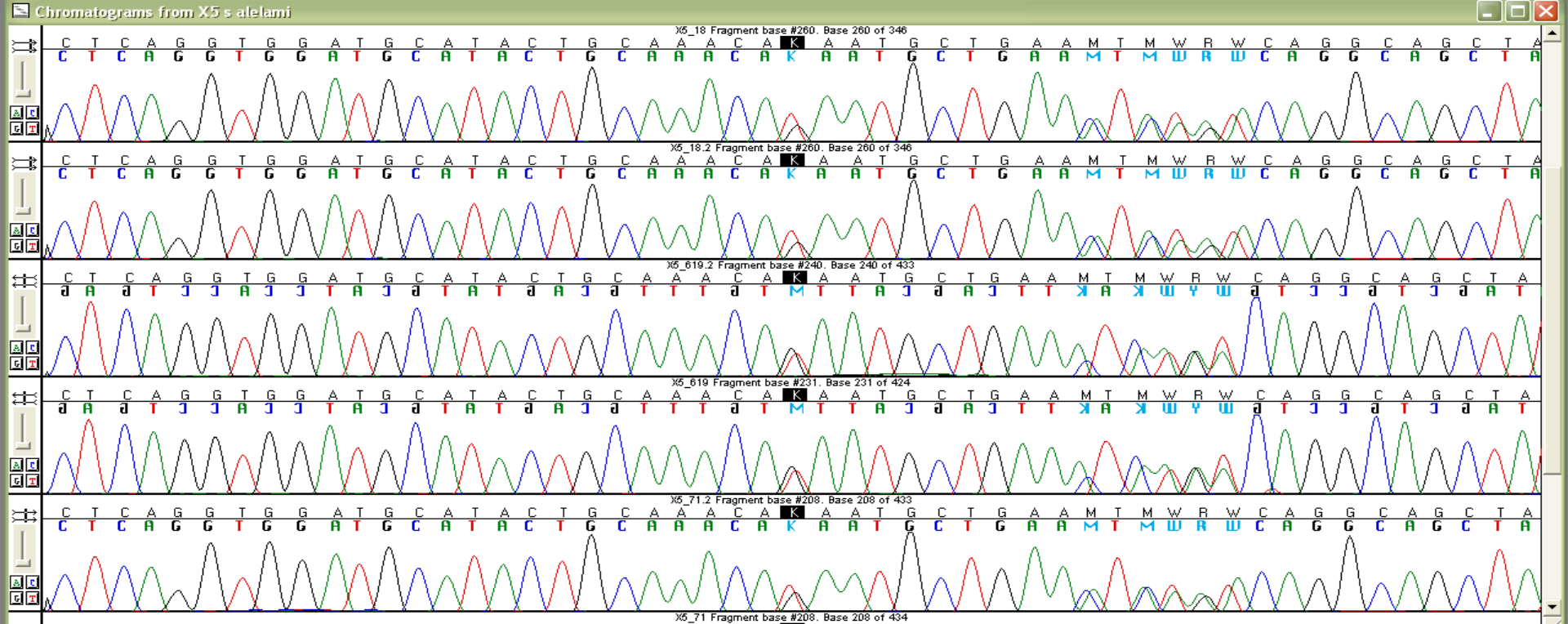
Overview Summary Cut Map Find Show Chromatograms ReAligner

DAB1\*20  
X5\_374.2  
X5\_374  
X5\_18  
X5\_18.2  
X5\_619.2  
X5\_619  
X5\_71.2  
X5\_71

```
GGGGTACACTGAAC TTGGAGTATATAAATGCACGAAGATTAACGACGATCCCAAC TTCTGCAGTCAGGAGGCTCAGGTGGATGCATACTGCCAAACA GAATGCTGAAATCAGACAGGCAGCTATCGG
GGGGTACACTGAAC TTGGAGTATATAAATGCACGAAGATTAACGACGATCCCAAC TTCTGCAGTCAGGAGGCTCAGGTGGATGCATACTGCCAAACA AATGCTGAATM CAGGCAGCTATCGG
GGGGTACACTGAAC TTGGAGTATATAAATGCACGAAGATTAACGACGATCCCAAC TTCTGCAGTCAGGAGGCTCAGGTGGATGCATACTGCCAAACA AATGCTGAA T CAGGCAGCTATCGG
GGGGTACACTGAAC TTGGAGTATATAAATGCACGAAGATTAACGACGATCCCAAC TTCTGCAGTCAGGAGGCTCAGGTGGATGCATACTGCCAAACA AATGCTGAA T CAGGCAGCTATCGG
GGGGTACACTGAAC TTGGAGTATATAAATGCACGAAGATTAACGACGATCCCAAC TTCTGCAGTCAGGAGGCTCAGGTGGATGCATACTGCCAAACA AATGCTGAA T CAGGCAGCTATCGG
GGGGTACACTGAAC TTGGAGTATATAAATGCACGAAGATTAACGACGATCCCAAC TTCTGCAGTCAGGAGGCTCAGGTGGATGCATACTGCCAAACA AATGCTGAA T CAGGCAGCTATCGG
GGGGTACACTGAAC TTGGAGTATATAAATGCACGAAGATTAACGACGATCCCAAC TTCTGCAGTCAGGAGGCTCAGGTGGATGCATACTGCCAAACA AATGCTGAA T CAGGCAGCTATCGG
```

10 frag bases selected at consensus position 293

```
GGGGTACACTGAACWTGGAGTATATAAATGCACGAAGATVTAACGACGATCCCAACWTTCTGCAGTCAGMGAGAGGCTCAGGTGGATGCATACTGCCAAACA KAATGCTGAATMWRWCAGGCAGCTATCGG
```





### Contig[0001]

Overview Summary Cut Map Find Show Chromatograms ReAligner

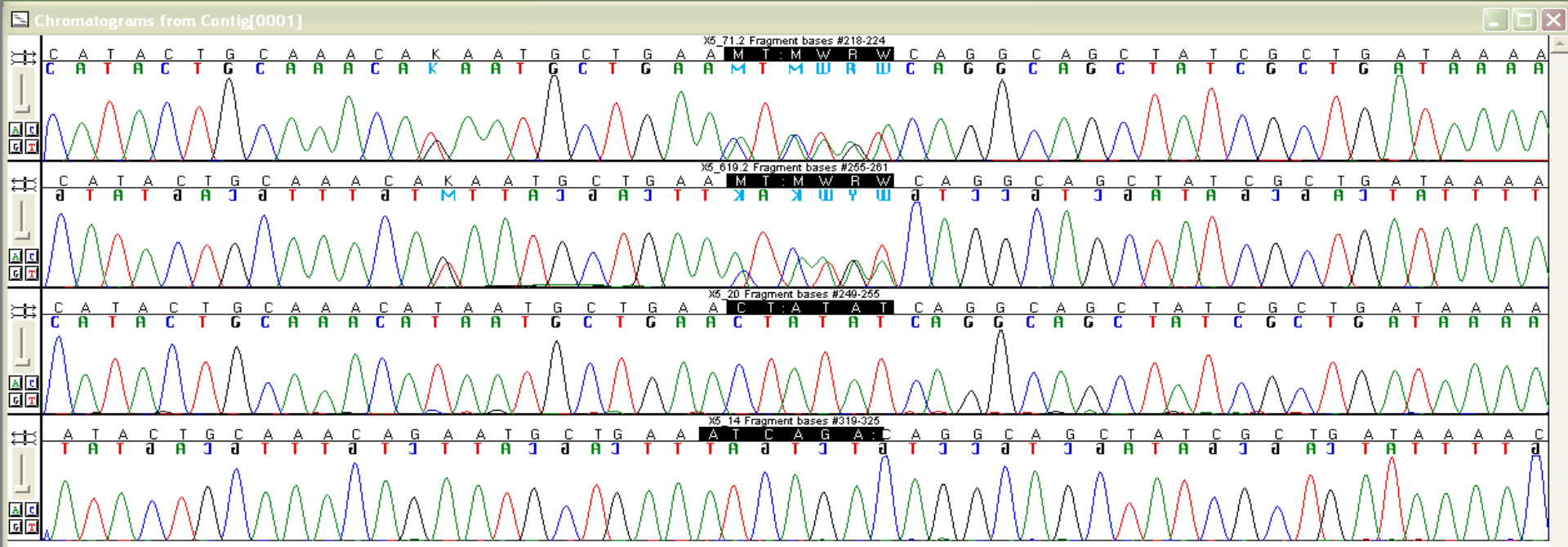
x5\_71.2  
x5\_619.2  
x5\_20  
x5\_14

```
ACGACGATCCCAAC TTCTGCAGTCAG GAGAGCTCAGGTGG ATGCATACTGCAAAACA AATGCTGAA T: CAGGCAGCTATCGCTGATAAAACAGGTACAGCACAGGCTTCTGATCCTCCTG
ACGACGATCCCAAC TTCTGCAGTCAG GAGAGCTCAGGTGGATGCATACTGCAAAACA AATGCTGAA T: CAGGCAGCTATCGCTGATAAAACAGGTACAGCACAGGCTTCTGATCCTCCTG
ACGACGATCCCAAC TTTCTGCAGTCAGAGAGAGCTCAG GTGGATGCATACTGCAAAACA AATGCTGAA CT: ATAT CAGGCAGCTATCGCTGATAAAACAGGTACAGCACAGGCTTCTGATCCTCCTG
ACGACGATCCCAAC ATTTCTGCAGTCAGGCAGAGCTCAG GTGGATGCATACTGCAAAACAGAAATGCTGAA ATCAGA: CAGGCAGCTATCGCTGATAAAACAGGTACAGCACAGGCTTCTGATCCTCCTG
```

28 frag bases  
8.7 consensus  
bases selected  
at consensus  
position 319

250 260 270 280 290 300 310 320 330 340 350 360 370

```
ACGACGATCCCAACWTTCTGCAGTCAGMGAGAGCTCAGGTGGATGCATACTGCAAAACA KAATGCTGAA MT: ATAT CAGGCAGCTATCGCTGATAAAACAGGTACAGCACAGGCTTCTGATCCTCCTG
```



### Sequencher

File Edit Select Contig Sequence View Window Help

#### Contig[0014]

Overview Summary Cut Map Find Show Chromatograms ReAligner Maximize

```

SKR5392_MSTN_1_1B_002_3432 TT C TCT TG T A TT G T C C AA A ATA CT GAG GTGA: A: TTTGTGTTTTACAAAAGTATCCTCA
SKR5391_MSTN_1_1A_001_3433 TTACCCCTCTAACTGTGGACTTTGGAAGCTTTTGGATGGGATTGGATTATTGCACCCAAAAGATATAAGGCCAATTACTGCTCTGGAGAGTGA
SKR5406_MSTN_9_1B_004_3436 T C CT TG T A TT G G T C C AA A ATA CT GAG GTGA: A: TTTGTGTTTTACAAAAGTATCCTCA
SKR5405_MSTN_9_1A_003_3436 TTACCCCTCTAACTGTGGACTTTGGAAGCTTTTGGATGGGATTGGATTATTGCACCCAAAAGATATAAGGCCAATTACTGCTCTGGAGAGTGC T TGT TTT
  
```

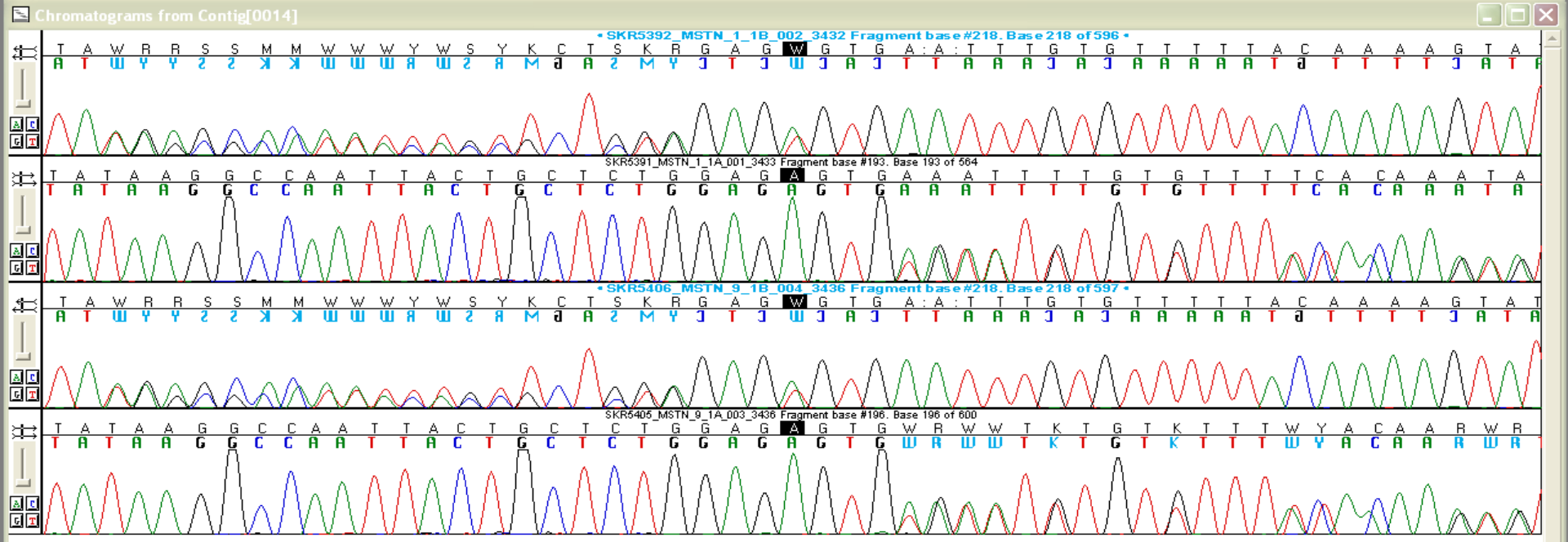
4 frag bases selected at consensus position 218

30 140 150 160 170 180 190 200 210 220 230 240 250

```

TTACCCCTCTAACTGTGGACTTTGGAAGCTTTTGGATGGGATTGGATTATTGCACCCAAAAGATATAAGGCCAATTACTGCTCTGGAGAGTGA
  
```

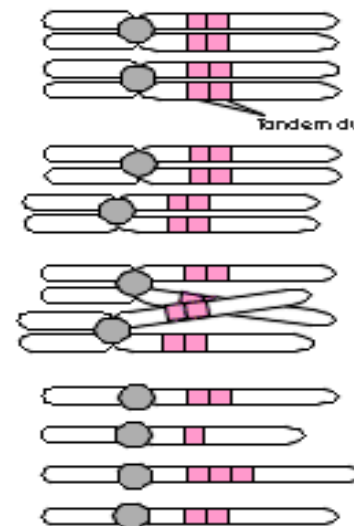
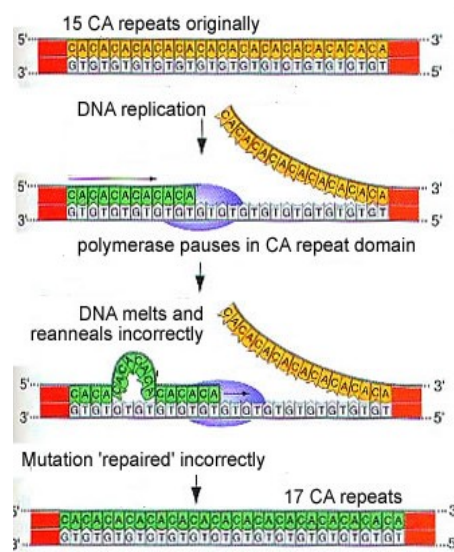
++ . +++++



# Mikrosatelity – definice a využití

- STR, SSR – jednoduché motivy, VNTR – složité motivy
- Různé typy motivů v tandemovém opakování
  - Mononukleotidový motiv – jednoduchá repetice (polyA) **(A)<sub>n</sub>**
  - Dinukleotidový motiv – nejčastější v panelech v určování původu hospodářských zvířat **(GT)<sub>n</sub>**
  - Trinukleotidový motiv – vhodnější na odečet (panely u psů) – kombinace s di-nt motivy v panelech **(GTC)<sub>n</sub>**
  - Tetranukleotidový motiv – panely pro určování původu u zvířat i lidí, vhodné na odečet, ne tak časté **(ATCT)<sub>n</sub>**
  - Složené motivy – složité sekvence – zejména u nižších živočichů
    - *Bythinella*  
GA(CA)<sub>3</sub>(GACA)<sub>4</sub>(GA)<sub>2</sub>(CA)<sub>2</sub>(GA)<sub>22</sub>CA(GA)<sub>7</sub>CA(GA)<sub>4</sub>CA(GA)<sub>4</sub>CA(GA)<sub>7</sub>
    - *Gammarus*  
[(CAT)<sub>2</sub>CACC(CAT)<sub>2</sub>C]<sub>2</sub>(CAT)<sub>2</sub>CACC(CAT)<sub>5</sub>G(CAT)<sub>2</sub>CACC(CAT)<sub>2</sub>

- Polymorfismus na základě variability opakování – vysoce polymorfní
- Multialelické – zdroj genetické variability
- Klasické Mendelovské křížení a segregace
- Vznik nových alel – DNA pol. slippage, chyby v crossing-overu během meiózy
- Zejména v nekódujících oblastech – intergenové oblasti (genetický balast) a intragenové oblasti (UTR, introny)



# Funkce a využití MS

- Funkce neověřená - ochotně rekombinují, tvoří sekundární struktury (vliv na replikaci DNA a buněčný cyklus), možná regulace genové aktivity (transkripce a translace)
- Využití zejména:
  - v populačních studiích (struktura populace, fylogenetické analýzy, geografická vazba)
  - forensní genetika (15 MS+amylogenin)
  - identifikace jedince (paternita, parentita, původ – i u zvířat – nutnost stanovení genetického profilu)
  - diagnostika a určování onemocnění (zvířata i lidi, vazbové markery)
  - konstrukce vazbových map (potřeba rodin a populací – existují již komerční kity např. ABI Prism Human Linkage Mapping Site)

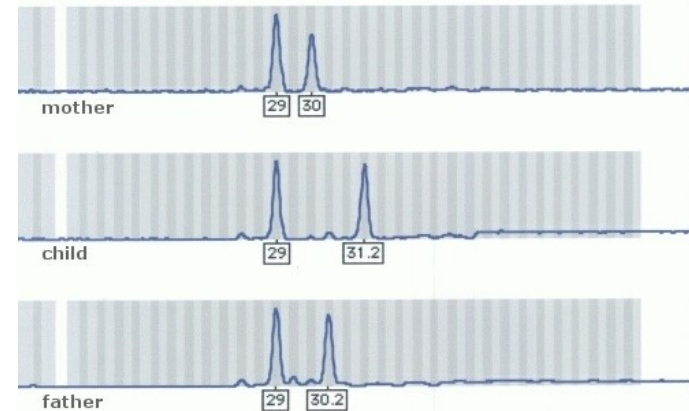
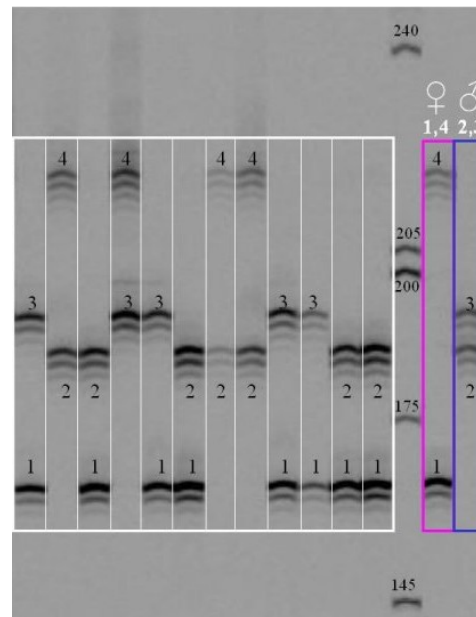
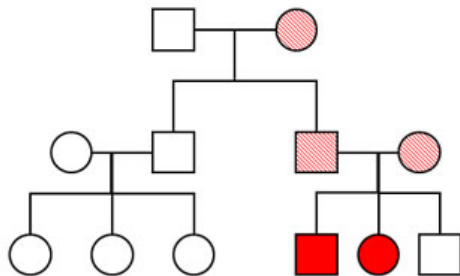
# Velikost alel mikrosatelitních markerů a jejich variabilita

- Variabilita v opakování motivu – variabilní délka amplifikovaného fragmentu
- Možná modifikace (mutace) v místě nasedání primeru – falešná homozygotita, nulové alely
- Možné chyby při PCR amplifikaci – zklouznutí DNA polymerázy, amplifikace nespecifického místa
- Přidávání adeninu na konec fragmentu
- Nestabilita při amplifikaci – polymeráza dělá čím dál tím kratší fragmenty

**CGTAGCCTTGCATCCTTCTCTCTCTCTCTCTCTCTCTATCGGTACTACGTGG**  
**CGTAGCCTTGCATCCTTCTCTCTCTCTATCGGTACTACGTGG**  
**CGTAGCCTTGCATCCTTCTCTATCGGTACTACGTGG**

# Historie hodnocení variability mikrosatelitních markerů

- Genealogie – tvorba rodokmenů
- PCR amplifikace variabilních míst v genomu – horizontální gelová elektroforéza (EtBr)
- Fragmentační analýza kapilární gelovou elektroforézou (fluorofory)

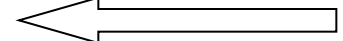


# Princip fragmentační analýzy kapilární elektroforézou FA-CE

- Nezajímá nás sekvence fragmentů
- Důležité – počet bazí (délka fragmentů), množství DNA (určuje výška píku)
- Relativní srovnávací metoda – potřeba interního standardu (alignment by time scale/size scale)
- Víceero píků+artefakty – nutno odlišit konkrétní alelu
- 1 vs. víceero markerů – počet rozhoduje
- Amplifikace polymorfního místa PCR (možnost multiplex) – důležitá kvalita a kvantita templátu (empirické stanovení)
- Značení fragmentů pomocí značeného primeru (5 modifikace fluoroforem)
- Separace amplifikovaných fragmentů v elektrickém poli kapilární gelovou elektroforézou



**CGTAGCCTTGCATCCTTCTCTCTCTCTCTCTCTCTCTCTATCGGTACTACGTGG**





- Příprava vzorku pro FA – denaturační činidlo (formamid)+interní standard velikosti fragmentů
- Sekvenátor
- Kapilára – 36 cm, polymer POP4, matrice a spektrální kalibrace pro daný modul FA
- Ovládání softwarem – DataCollection
- Elektrokinetická injektáž – objem vzorku zůstává nezměněn
- Pohyb fragmentů DNA gelem
- Definování správných podmínek runu (modul) – bere v úvahu směs použitých fluorochromů, délku kapiláry
- Detekce emise fluoroforu při přechodu detekčním okénkem
- Převod dat – spektra – raw data – elektroforetogram
- Vyhodnocení a analýza – GeneScan + Genotyper, GeneMapper+PeakScanner

**3100 Data Collection Software**

File View Instrument Tools Service Help

Plate View Run View Status View Array View Capillary View

**Instrument Condition**

- Laser: On
- EP: Off
- Over: On
- Front Doors: Closed
- Oven Door: Closed
- Autosampler: Return

Time Remaining this run: 00:42:08  
Run Time: 00:44:24

EP Voltage: 20.0 kV  
EP Current: 800.0 µA

Laser Power: 15.2 mW  
Laser Current: 4.5 A

Oven Temp: 39.5 °C

Capillary Array Serial Number: demo  
Capillary Array Usage: 49

**Events**

```
[Wed Apr 12 16:58:24 PDT 2000]
Set run status to STARTING
[Wed Apr 12 16:58:24 PDT 2000]
Started a Gene Scan Run: Run_demo_3100_2000-04-12_46
[Wed Apr 12 16:58:23 PDT 2000]
Sending run module to the instrument.
[Wed Apr 12 16:37:15 PDT 2000]
%% dis-connect data communication port
[Wed Apr 12 17:02:00 PDT 2000]
%% RUN PARAMETERS command reply
[Wed Apr 12 17:02:03 PDT 2000]
%% BEGIN command reply
[Wed Apr 12 17:02:36 PDT 2000]
OVEN-TEMPERATURE-TOLERANCE 3.0 %% Oven temperature b
```

**Errors**

```
[Wed Apr 12 16:58:24 PDT 2000]
Instrument offline
```

Waiting for Oven to Stabilize at Run Temperature

Set Application

Application: SequencingAnalysis

Sample Sheet

Done

5a

5b

**Matrices for dye set D**

Intensity vs. Bin Number

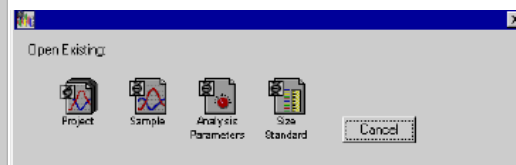
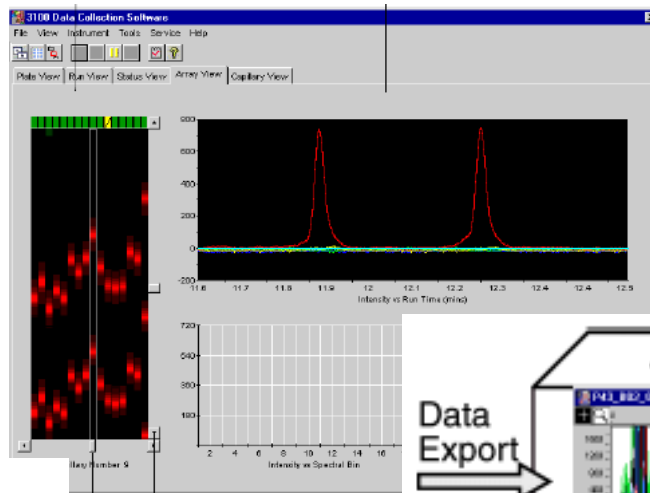
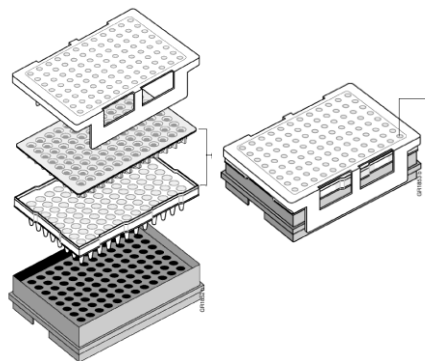
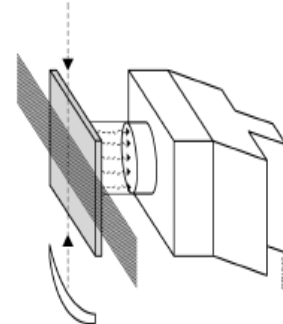
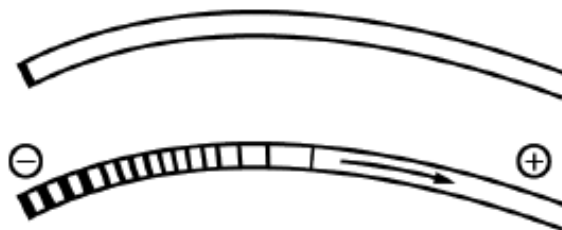
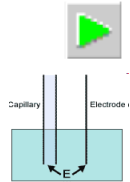
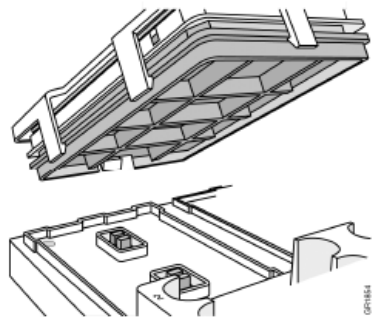
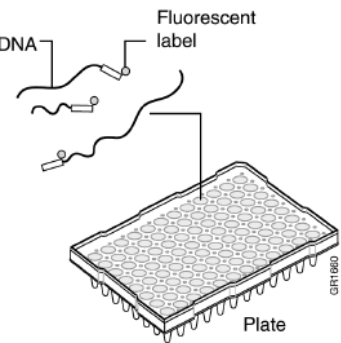
Intensity vs. Scan Number

Capillary Number 1  
Condition Number: 5.18  
Q Value: 0.994  
Data Source:

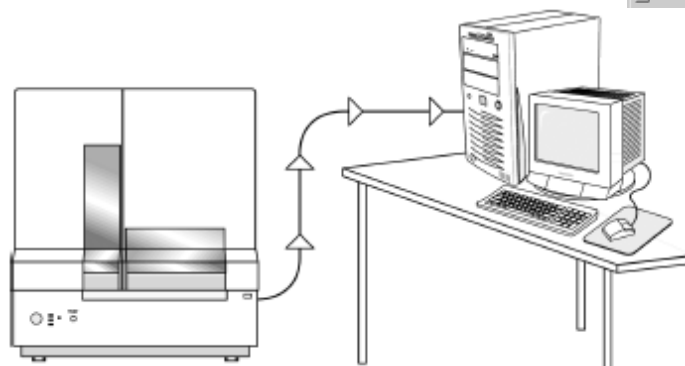
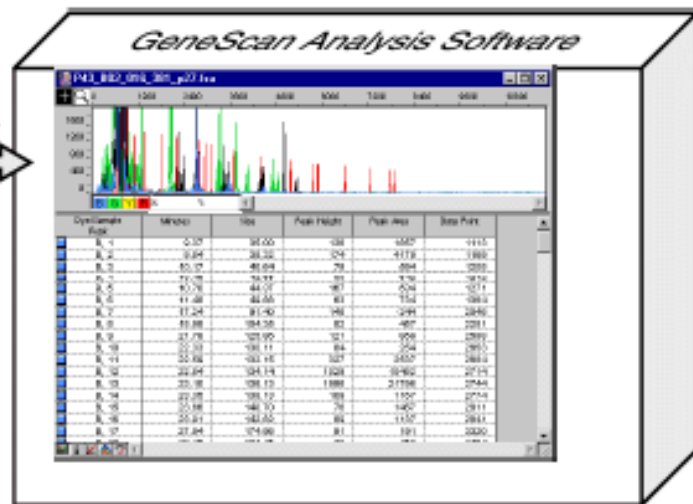
OK Cancel

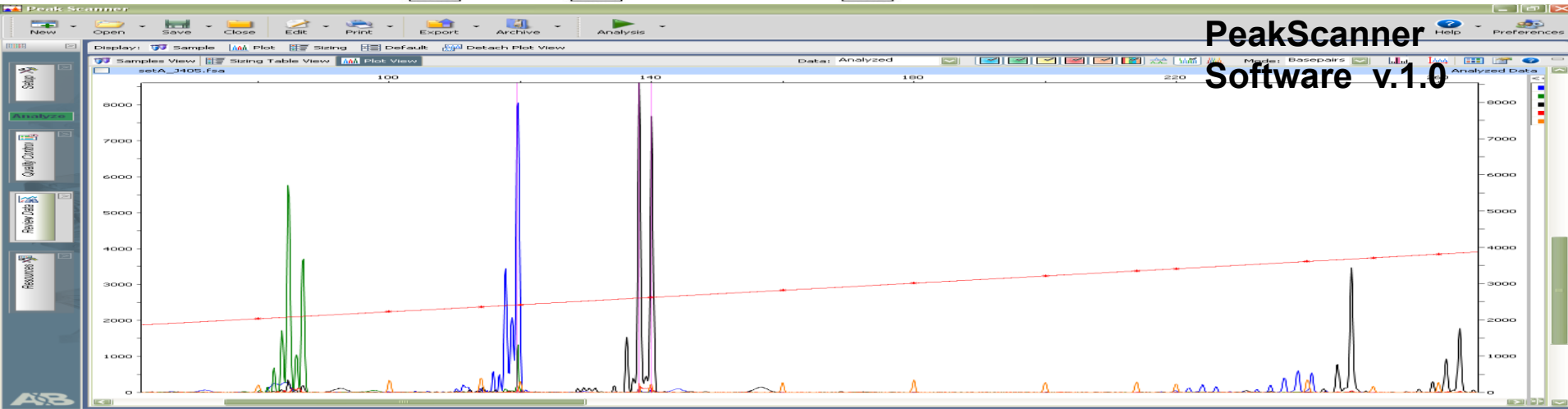
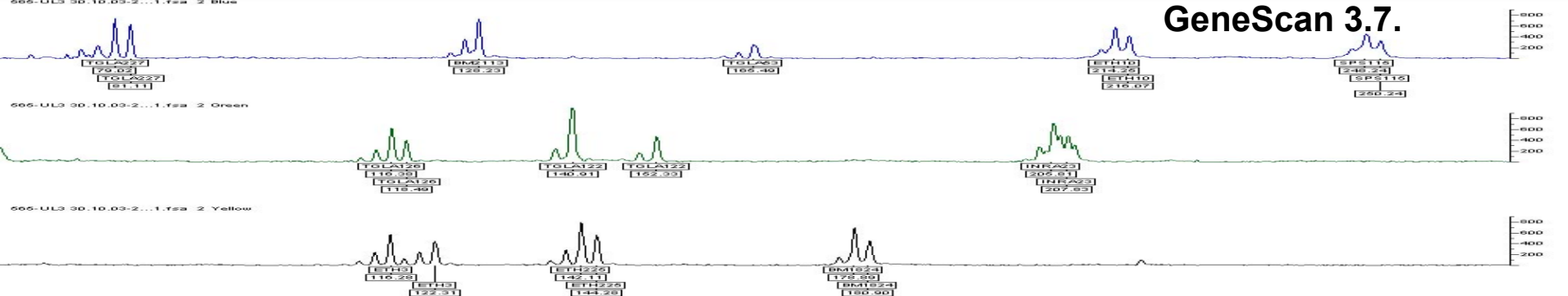
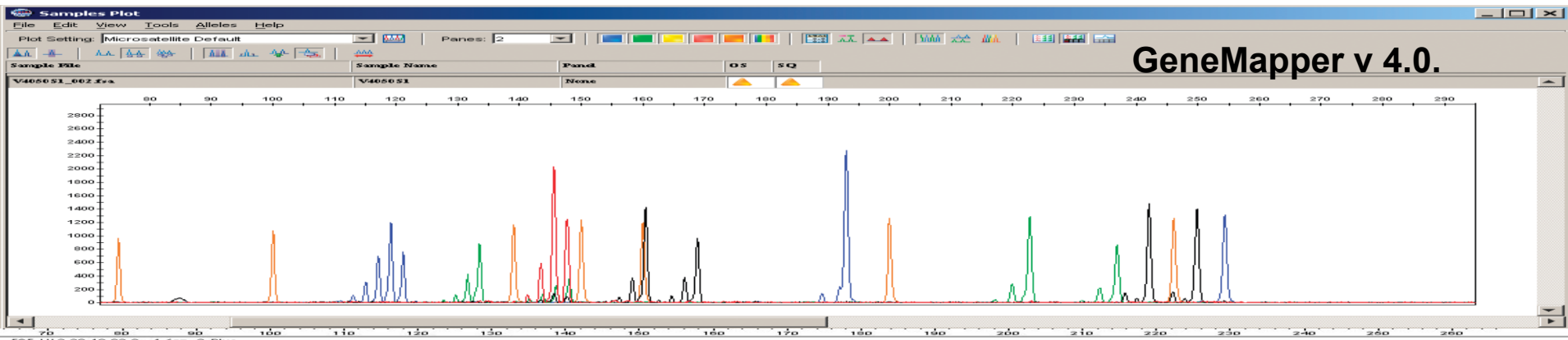
Dye Sets, Dye Chemistry, and Applications:

Dye Set	Dye Chemistry	Application Kit
DS-30 (D)	6-FAM™ (blue), HEX™ (green), NED™ (yellow), ROX™ (red)	<ul style="list-style-type: none"> <li>Linkage Mapping Sets-LD20, -MD10, and -HD5</li> <li>Custom Oligos</li> </ul>
DS-31 (D)	New 4-Dye Chemistry: 6-FAM (blue), VIC™ (green), NED (yellow), ROX (red)	<ul style="list-style-type: none"> <li>Linkage Mapping Custom Oligos</li> <li>Mouse Mapping Markers</li> <li>Custom Oligos</li> </ul>
DS-32 (F)	5-FAM™, JOE™, NED, ROX	<ul style="list-style-type: none"> <li>AmpFISTR® products</li> <li>Stockmarks™</li> </ul>
DS-33 (G5) 5-dye chemistry for high throughput genotyping	6-FAM, VIC, NED, PET™, LIZ™	<ul style="list-style-type: none"> <li>AmpFISTR® Identifier™</li> <li>Custom Oligos</li> </ul>
DS-02 (E5)	dR110, dR6G, dTAMRA™, dROX™, LIZ	SNAPSHOT™ Multiplex Kit

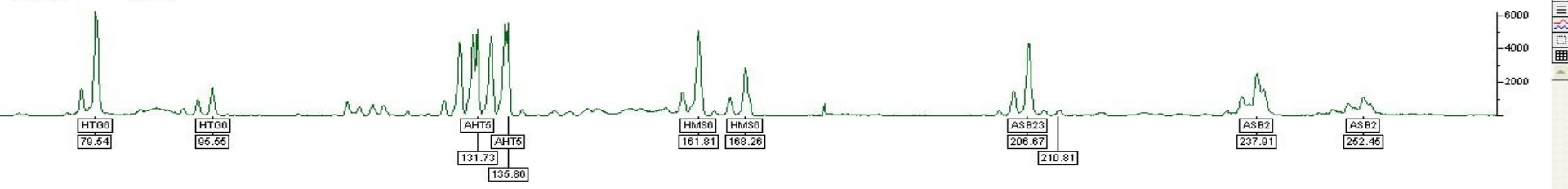


Data Export

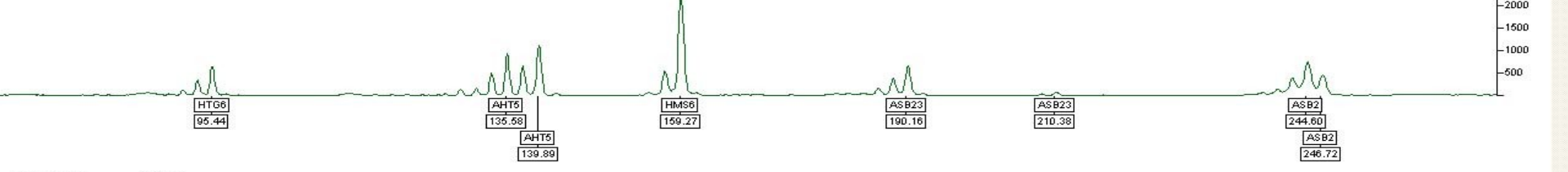




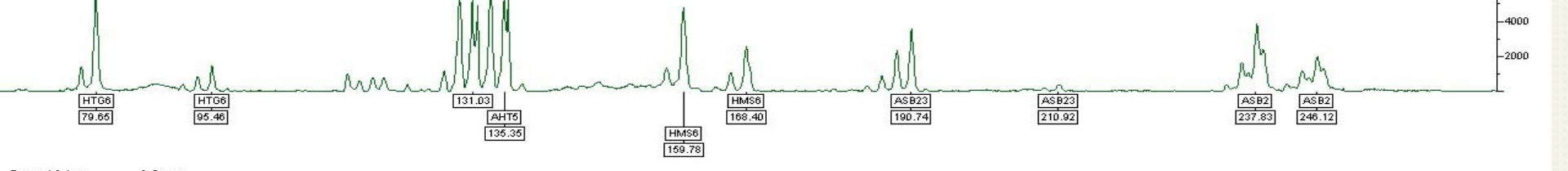
Matka1.fsa 20 Green



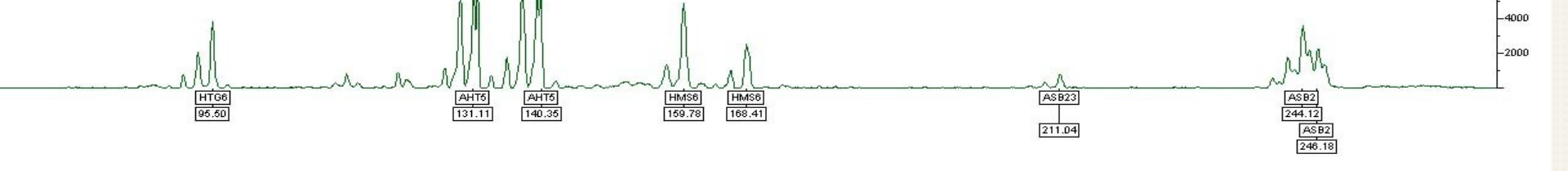
Otec.fsa 6 Green



Potomek1.fsa 19 Green



Potomek2.fsa 9 Green



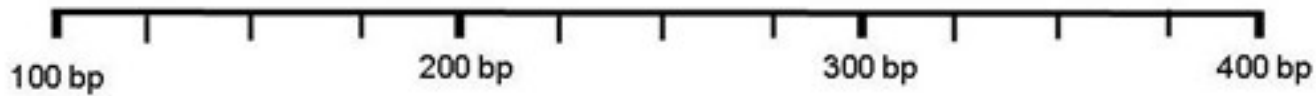
# Kritéria pro tvorbu panelů MS markerů

- Počet je rozhodující
- Pokrytí celého genomu (téměř všechny chromosomy)
- Co největší variabilita, množství alel
- Vysoký PIC (polymorfní informační obsah)
- Minimum nulových alel
- Vhodnost pro multiplex-PCR
- Velikostní rozpětí fragmentů alel, vliv na použití fluorochromu na značení
- 6-FAM, NED, PET, ROX, HEX, VIC, JOE, LIZ – možnosti 4 a 5 dye systémů

# Typy MS panelů

- Člověk
- Skot, Prase, Kůň
- Psi (10), dravci – sokol (tmavý, raroh, lovecký, stěhovavý), poštolka, orel (5), jelen, kočka
- Ryby (různé rody i druhy)
- Bezobratlí – šneci, blešivci, mravenci, mouchy etc...malé panely

# AmpFSTR® Identifiler™



D8S1179    D21S11    D7S820    CSF1PO

D3S1358    TH01    D13S317    D16S539    D2S1338

D19S433    vWA    TPOX    D18S51

A    D5S818    FGA



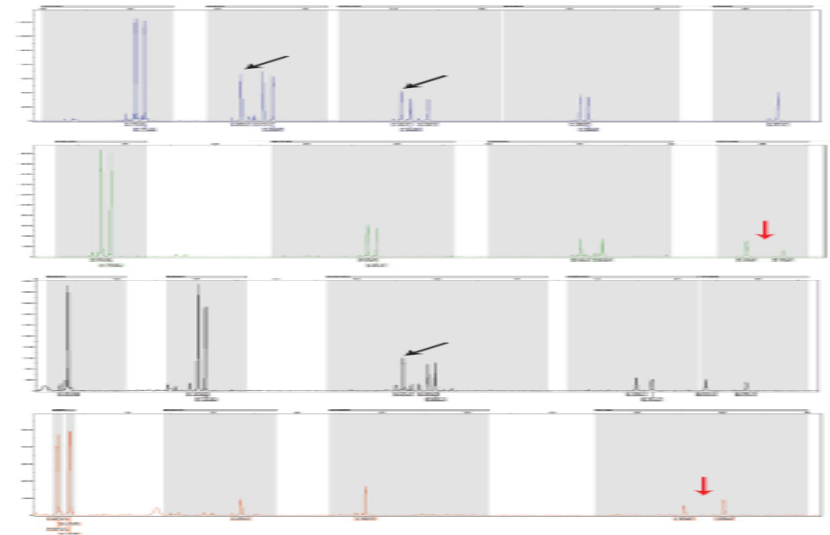
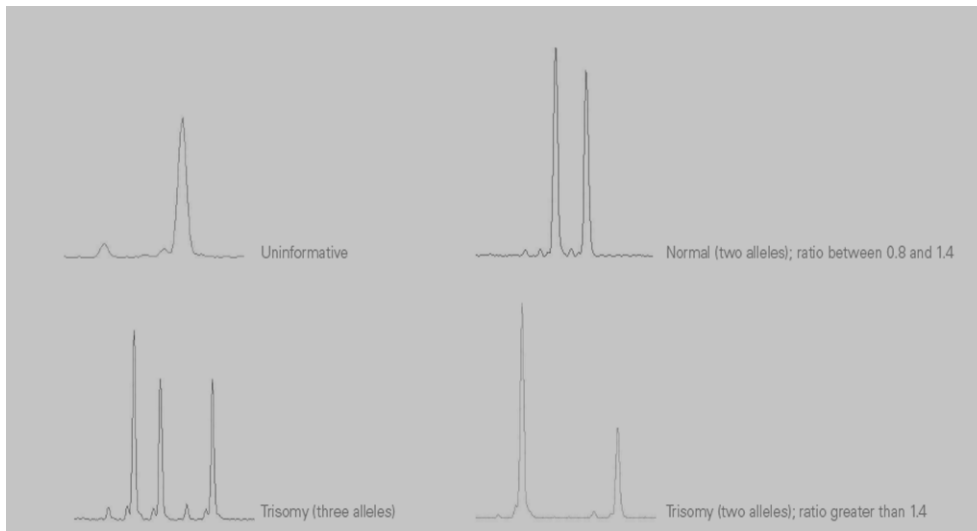
GS500-internal lane standard



lokus	chromozom	fluorescenční značka	barva	rozsah (v bp)
<b>VHL20</b>	30	FAM	modrá	83-102
<b>HTG4</b>	9	FAM	modrá	116-137
<b>AHT4</b>	24	FAM	modrá	140-166
<b>HMS7</b>	1	FAM	modrá	167-187
<b>HTG6</b>	15	VIC	zelená	74-103
<b>AHT5</b>	6	VIC	zelená	126-147
<b>HMS6</b>	4	VIC	zelená	154-170
<b>ASB23</b>	3	VIC	zelená	176-212
<b>ASB2</b>	15	VIC	zelená	237-268
<b>HTG10</b>	21	NED	žlutá	83-110
<b>HTG7</b>	4	NED	žlutá	114-128
<b>HMS3</b>	9	NED	žlutá	146-170
<b>HMS2</b>	10	NED	žlutá	215-236
<b>ASB17</b>	2	PET	červená	104-116
<b>LEX3</b>	X	PET	červená	137-160
<b>HMS1</b>	15	PET	červená	166-178
<b>CA425</b>	28	PET	červená	224-247

# Možnost stanovit množství DNA pomocí FA

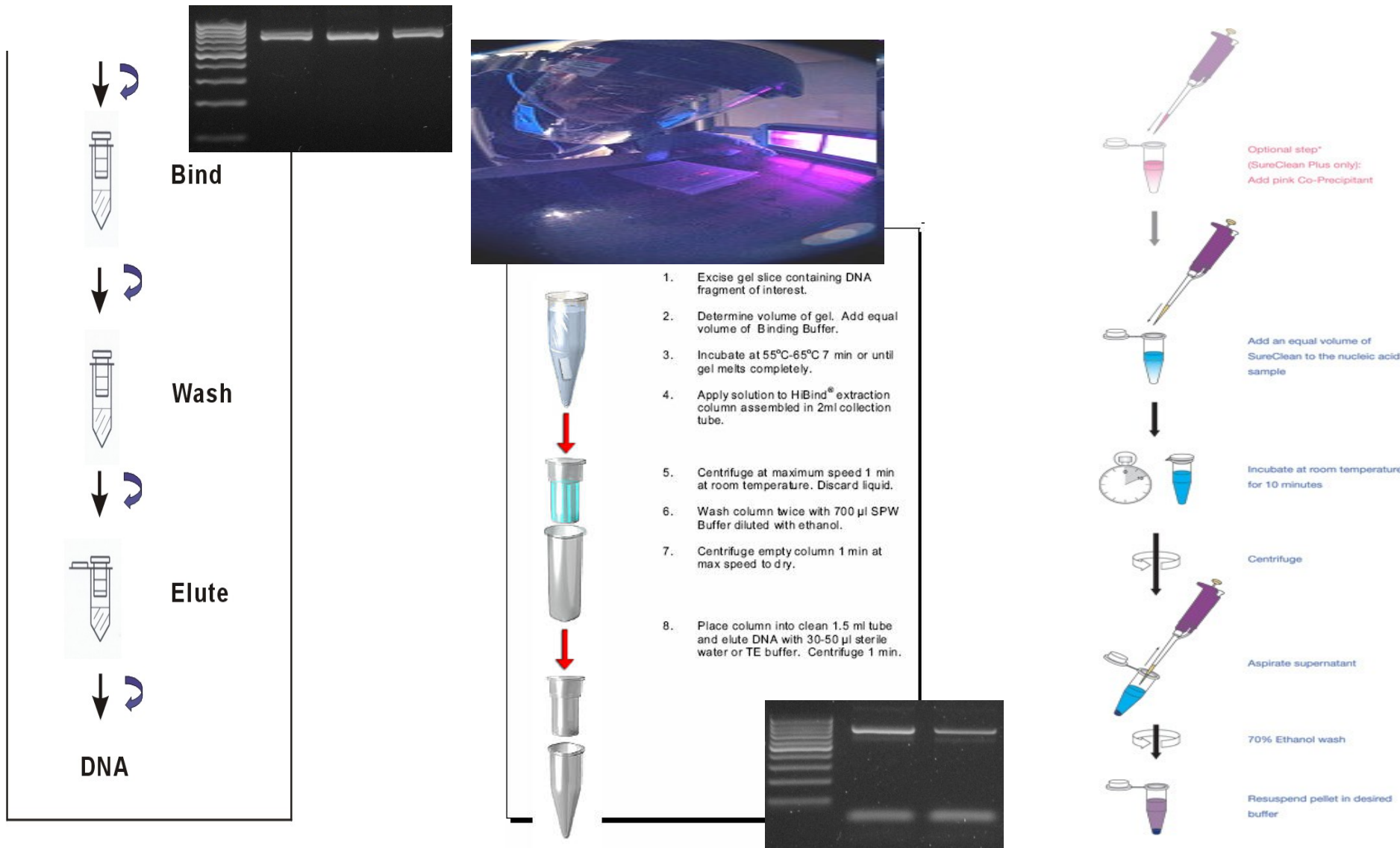
- Nezaměňovat s kvantifikací pomocí qPCR
- Určení množství ampliconu pomocí výšky a plochy píku
- Nepřesné a relativní – hrozí „plato efekt“ (vyčerpání systému)
- Využívané při detekci onemocnění způsobenými polyploidii
- Trisomie chromosomu 21 u lidí – Downův syndrom



# Sekvenační reakce - příprava

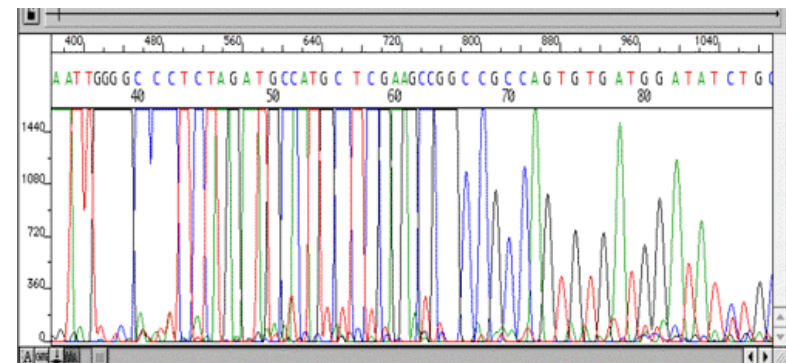
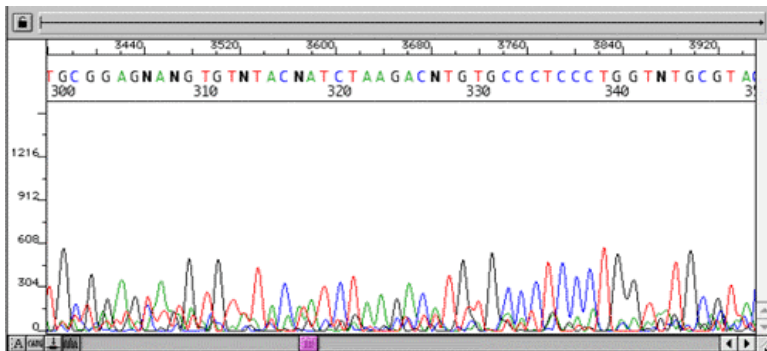
- Templát – PCR, BAC, plazmidový vektor,
- Možnosti přečištění PCR
  - Kolony
  - Purifikace fragmentu z gelu – nespecifity al. dimery primerů
  - SureClean
- Možnosti stanovení koncentrace PCR produktu po purifikaci
  - Spektrofotometricky
  - Gelová elektroforéza – porovnání velikosti a koncentrace s markerem
  - Na základě fluorescence – Qubit
  - Nanodrop – kvalita i kvantita (koncentrace DNA při 260 nm, do up to 3700 ng/ul bez ředění)
- Reagencie sekvenační PCR
  - Typy sekvenačních kitů
  - Tabulka využití
- Teplotní profil reakce – klasický vs. fast

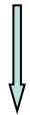
# Přečištění PCR produktu



# Stanovení koncentrace

Template	Cycle Sequencing Chemistry				
	BigDye Terminator v1.1 and v3.1	BigDye XTerminator Purification Kit	dGTP BigDye Terminator v3.0	dRhodamine Terminator	BigDye Primer
PCR product:					
100 to 200 bp	1 to 3 ng	0.5 to 3 ng	1 to 3 ng	1 to 3 ng	2 to 5 ng
200 to 500 bp	3 to 10 ng	1 to 10 ng	3 to 10 ng	3 to 10 ng	5 to 10 ng
500 to 1000 bp	5 to 20 ng	2 to 20 ng	5 to 20 ng	5 to 20 ng	10 to 20 ng
1000 to 2000 bp	10 to 40 ng	5 to 40 ng	10 to 40 ng	10 to 40 ng	20 to 50 ng
>2000 bp	40 to 100 ng	10 to 50 ng	40 to 100 ng	40 to 100 ng	50 to 150 ng
Bisulfite converted genomic DNA-PCR product	3 to 10 ng	3 to 10 ng	not recommended	not recommended	not recommended
Single-stranded DNA	50 to 100 ng	10 to 50 ng	50 to 100 ng	50 to 100 ng	150 to 400 ng
Double-stranded DNA	200 to 500 ng	50 to 300 ng	200 to 500 ng	200 to 500 ng	200 to 800 ng
Cosmid, BAC	0.5 to 1.0 µg	200 to 1,000 ng	0.5 to 1.0 µg	not recommended	0.5 to 1.0 µg
Bacterial genomic DNA	2 to 3 µg	1,000 to 3,000 ng	2 to 3 µg	not recommended	not recommended





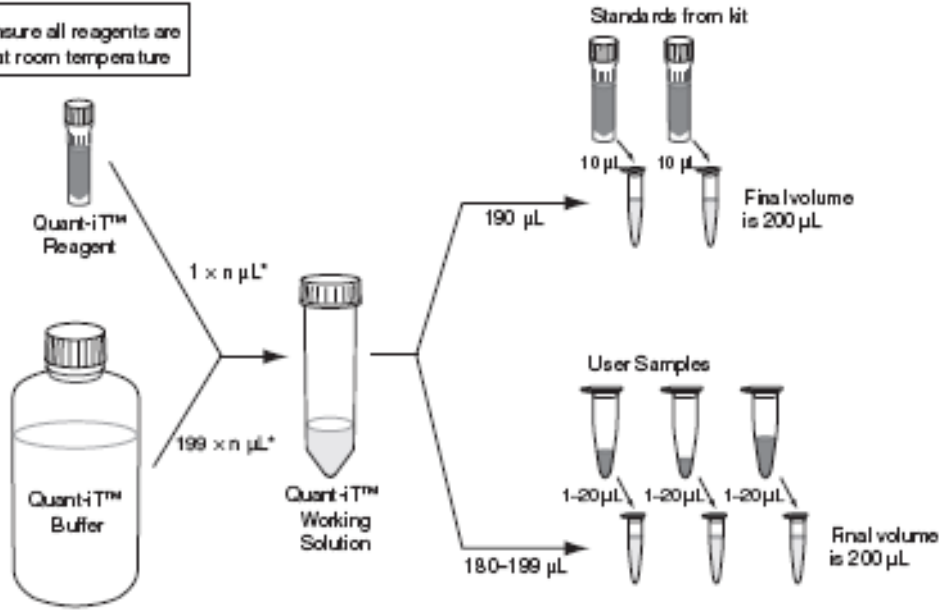
marker (v  $\mu\text{l}$ )

PCR 20 10 5 2,5 1,25 0,625



fragmenty	množství DNA v ng na gelu při nanášení ředěného markeru:						
	20 $\mu\text{l}$	10 $\mu\text{l}$	5 $\mu\text{l}$	2,5 $\mu\text{l}$	1,25 $\mu\text{l}$	0,625 $\mu\text{l}$	0,312 $\mu\text{l}$
1031	206	103	51,5	25,75	12,875	6,4375	3,2136
900	180	90	45	22,5	11,25	5,625	2,808
800	160	80	40	20	10	5	2,496
700	140	70	35	17,5	8,75	4,375	2,184
600	120	60	30	15	7,5	3,75	1,872
500	200	100	50	25	12,5	6,25	3,12
400	80	40	20	10	5	2,5	1,248
300	60	30	15	7,5	3,75	1,875	0,936
250	50	25	12,5	6,25	3,125	1,5625	0,78
200	80	40	20	10	5	2,5	1,248
150	30	15	7,5	3,75	1,875	0,9375	0,468
100	60	30	15	7,5	3,75	1,875	0,936
50	40	20	10	5	2,5	1,25	0,624

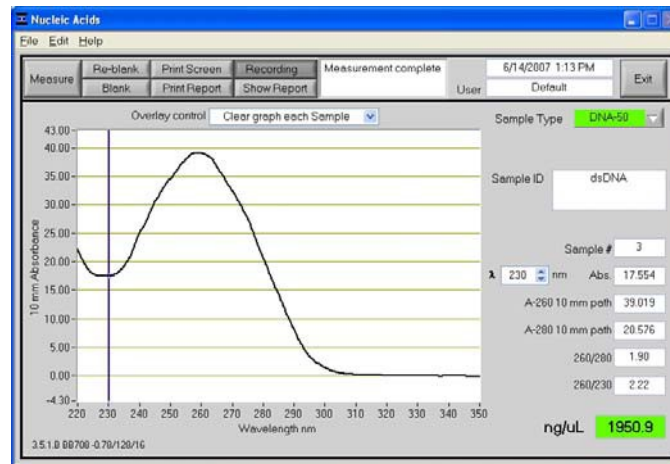
Ensure all reagents are at room temperature



\* where n = number of Standards plus number of Samples

invitrogen

Read tubes in Qubit™ fluorometer



# Reagence sekvenační PCR a možnosti modifikace teplotního profilu

## Chemistry Options

Applications	BigDye® Terminator v3.1 Kit	BigDye® Terminator v1.1 Kit
<i>de novo</i> sequencing	+	✓
Resequencing	+	✓
Sequencing difficult templates	+	+
Long-read sequencing	+	✓
Sequencing across all template types (plasmids, PCR products, BACs, and fosmids)	+	✓
Mixed-base detection	+	✓
Sequencing short PCR products using rapid electrophoresis run modules	✓	+

+ Recommended ✓ Satisfactory

## Nové sekvenační kity

## BigDye Direct Cycle Sequencing Kit

## M13-tailed PCR primery

## Výhody vs. nevýhody

Application	BigDye Terminator v3.1	BigDye Terminator v1.1	dGTP BigDye Terminator	dRhodamine Terminator	BigDye Primer
<b>DNA Sequencing Application</b>					
Bisulfite sequencing	+	++	-	-	-
Comparative sequencing (germline mutations 50:50 heterozygotes)	++	++	-	+	++
Comparative sequencing (somatic mutations 10:90 heterozygotes)	-	-	-	-	+
Comparative sequencing (somatic mutations 30:70 heterozygotes)	+	+	-	-	++
<i>De novo</i> sequencing	++	++	+	+	+
Gap closure (custom primers)	++	++	++ <sup>‡</sup>	+	-
Gene walking (custom primers)	++	++	+	+	-
Shotgun sequencing (universal primers, M13)	++	++	+	+	++
<b>DNA Sequence Context</b>					
AT-rich >65%	++	++	-	++	++
GC-rich >65%	++	++	+	+	+
GT-rich regions	+	+	++	++	++
Homopolymer A or T >25 bp <sup>§</sup>	-	-	-	++	-
<b>Template Type</b>					
BAC, cosmid, fosmid, lambda, large PCR product	++	++	-	+	+
Bacterial genomic DNA	++	++	-	-	-
Bisulfite-treated genomic DNA	+	++	-	-	-
PCR amplicon	++	++	+	++	++
PCR amplicon (heterozygous 10:90)	-	-	-	-	+





**Traditional Method** 690 min (11 hr 30 min)



AmpliTaQ Gold Fast  
PCR Master Mix

BigDye®  
XTerminator™ Kit

**Fast Method** 245 min (4 hr 5 min)

Activation of Enzyme	PCR			PCR Final (Step)	
	Cycle (35 Cycles)			Hold	Hold
Hold	Denature	Annealing	Extension	Hold	Hold
95°C	96°C	Primer <sup>TM</sup> *	68°C	72°C	4°C
10 Min	3 sec	3 sec	See Table Below	10 sec	∞

Denaturation	Cycle Sequencing			Hold
	Cycle (25 Cycles)			
Hold	Denature	Annealing	Extension	Hold
96°C	96°C	50°C	60°C	4°C
1 Min	10 sec	3 sec	75 sec	∞

Recommended Extension Times:

Length (Kb)	Extension Time
0.5	5 sec
1.0	15 sec
1.5	30 sec

# Možnosti přečištění sekvenační reakce a příprava vzorku k analýze





- Ethanol (inhibitor)
- Kombinace kolony a ethanol – odstraní víc neinkorporovaných terminátorů
- BigDye X-Terminator Kit
  - Není potřeba pracovat s formamidem (teratogen)
  - když zkrátíme dobu třepání, neubere tolik krátkých fragmentů
  - Platíčkové verze – minimum pipetování

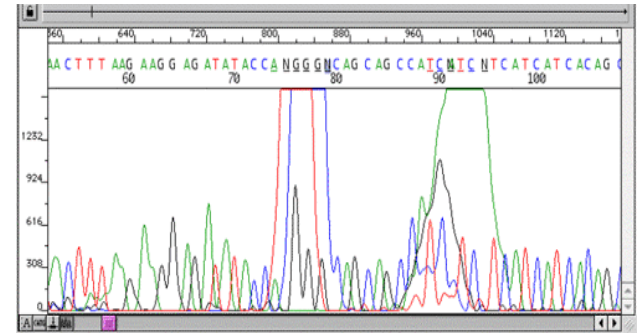
Start



Separation

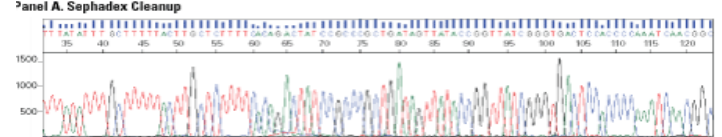
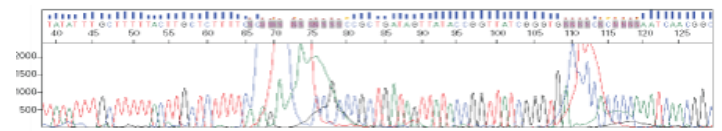
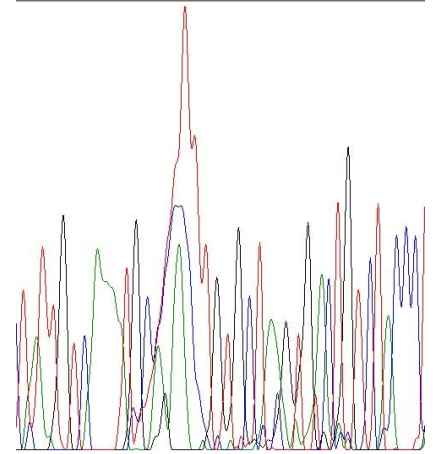
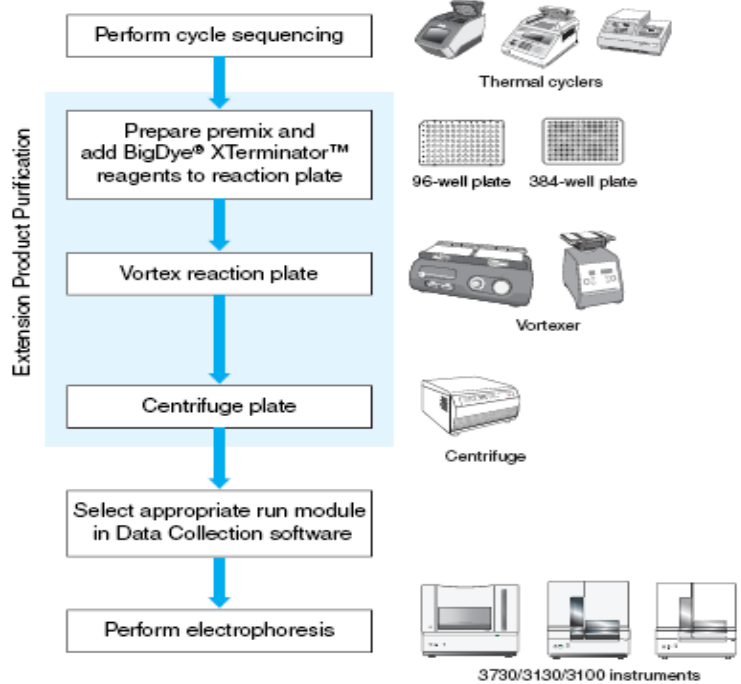


-  Gel-filtration material
-  Dye terminator
-  DNA
-  Dye terminator inside gel-filtration material



:TAT TGT CAA TGCCTNT TTTG TCT ANG TG ACTGT CTACCC:

**BigDye®  
XTerminator™  
Purification Kit  
Workflow**



3730/3130/3100 instruments

# Příprava stroje před runem – běžný uživatel

