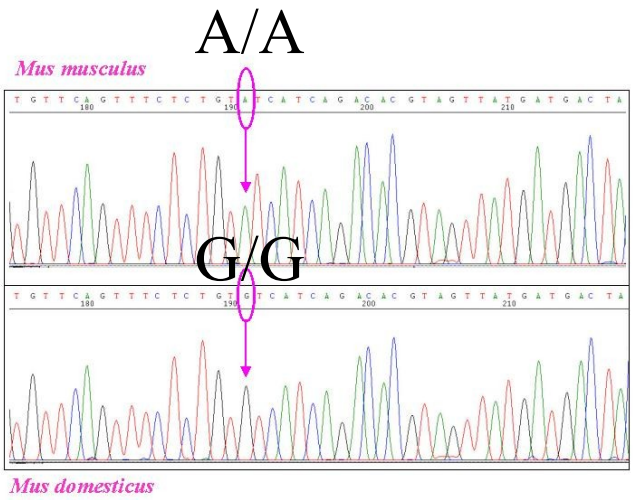
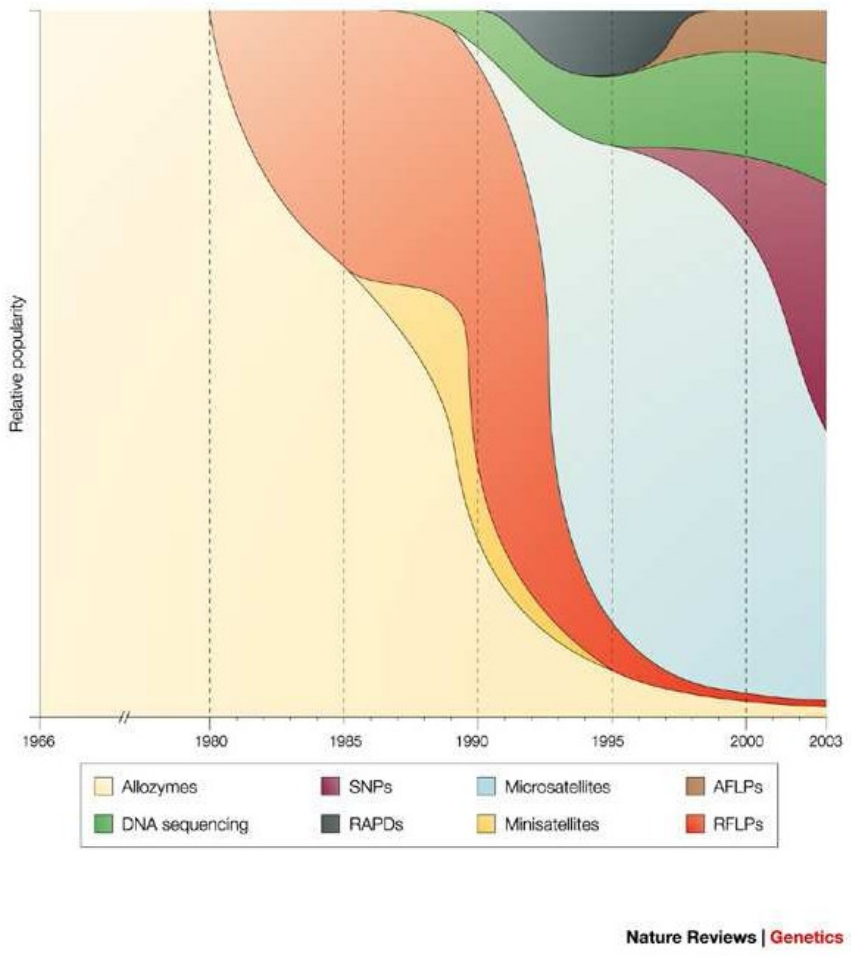
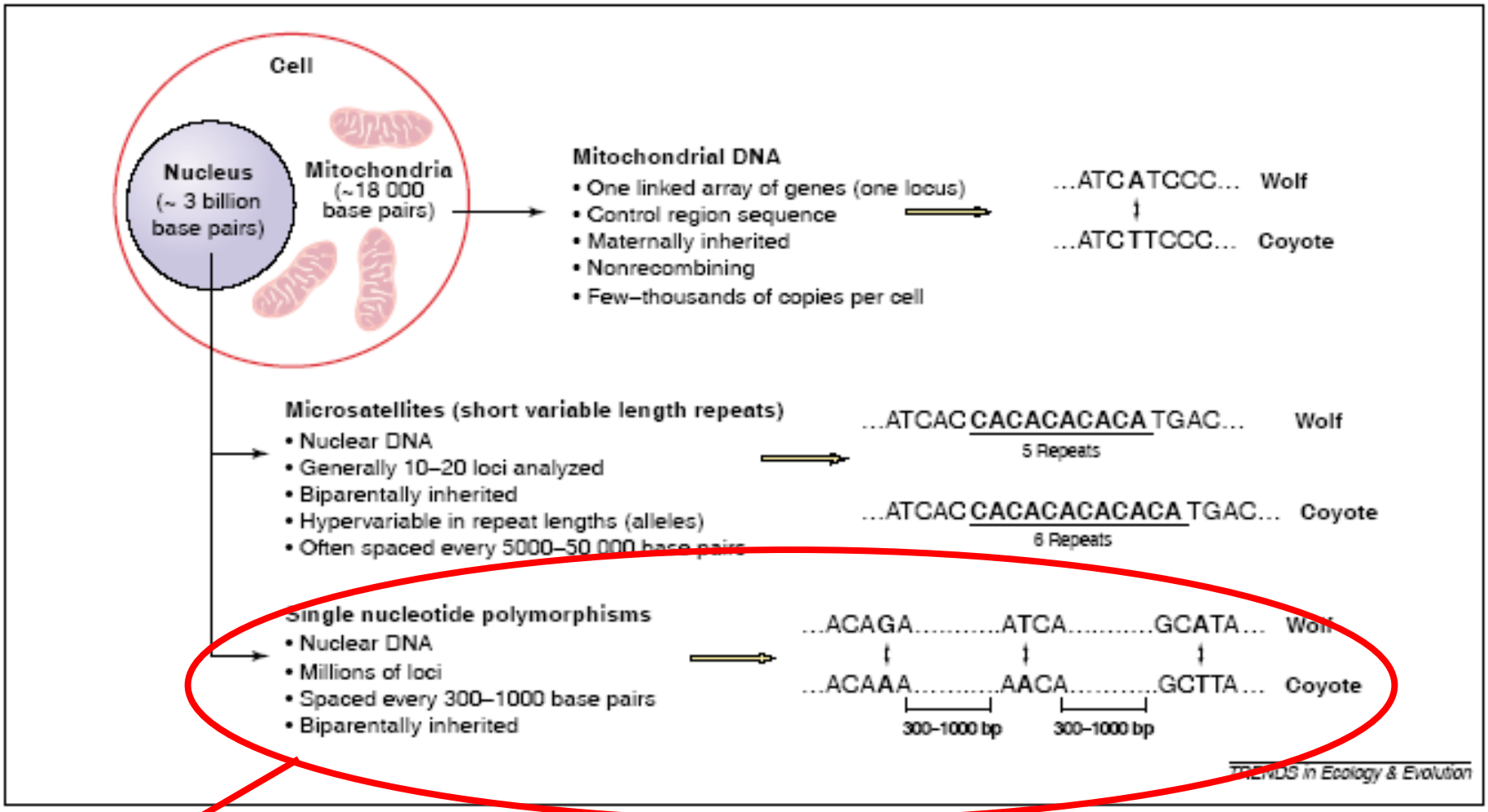
Single nucleotide polymorphisms (SNPs)





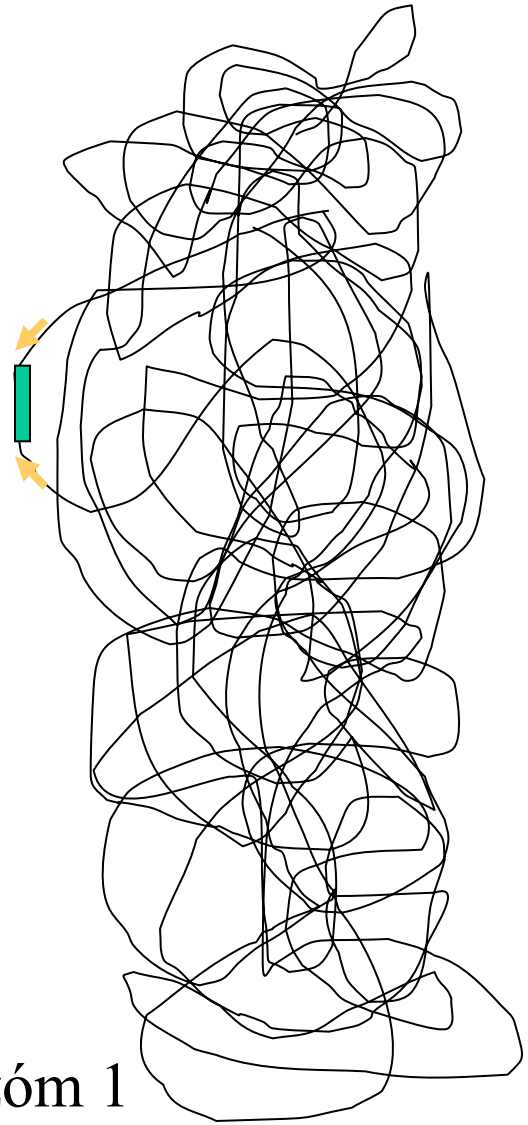
Single nucleotide polymorphisms (SNPs)



SNPs : nuclear genome (consensus)

Single-locus genetic markers

- SNPs (single nucleotide polymorphisms) – sekvenční polymorfismus
- kodominantní – je možné odlišit heterozygota (např. A/T) od homozygota (např. A/A)



CAAGTA
TGGACG

CATGTA
TGCACG

A/T

CAAGTA
TGGACG

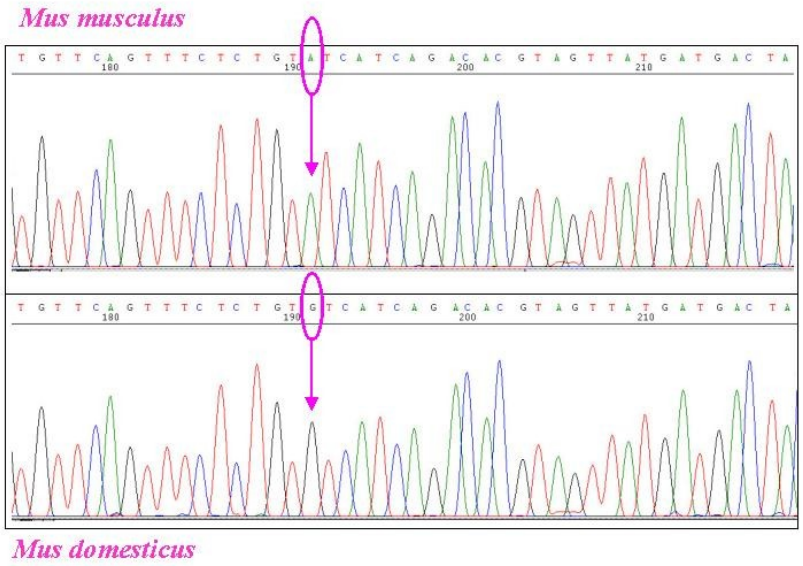
CAAGTA
TGGACG

A/A

Př.: chromozóm 1

Příklad informativního SNP znaku

- fixovaný polymorfismus (homozygoti) – využití např. při studiu hybridizací (hybridi = heterozygoti)



Značení heterozygotů

N = A, C, G, T
V = G, A, C
D = G, A, T
H = A, T, C
B = G, T, C
R = A, G
Y = C, T
M = A, C
K = G, T
S = G, C
W = A, T

transice

A ↔ G

transition: Pu→Pu or Py→Py transversion: Pu→Py or Py→Pu

Využití SNPs znaků

- obdobné jako u mikrosatelitů
- identifikace druhu (nebo genetické skupiny) - studium hybridizace
- fylogeografie
- populační genetika (genetická variabilita a struktura, tok genů, identifikace jedinců a vztahů mezi nimi, populační velikost a její změny atd.)
- mutace ve funkčních genech – i záměna jedné aminokyseliny může mít fatální dopad

Výhody

- početné a rozšířené v genomu (v kódujících i nekódujících oblastech) – milióny lokusů
- 1 SNP cca každých 300-1000 bp (v rámci druhu)
- Mendelovská dědičnost (vs. mtDNA)
- evoluce je dobře popsatelná jednoduchým mutačním modelem (vs. microsatellites)
- jsou analyzovány kratší fragmenty DNA – neinvazivní genetika

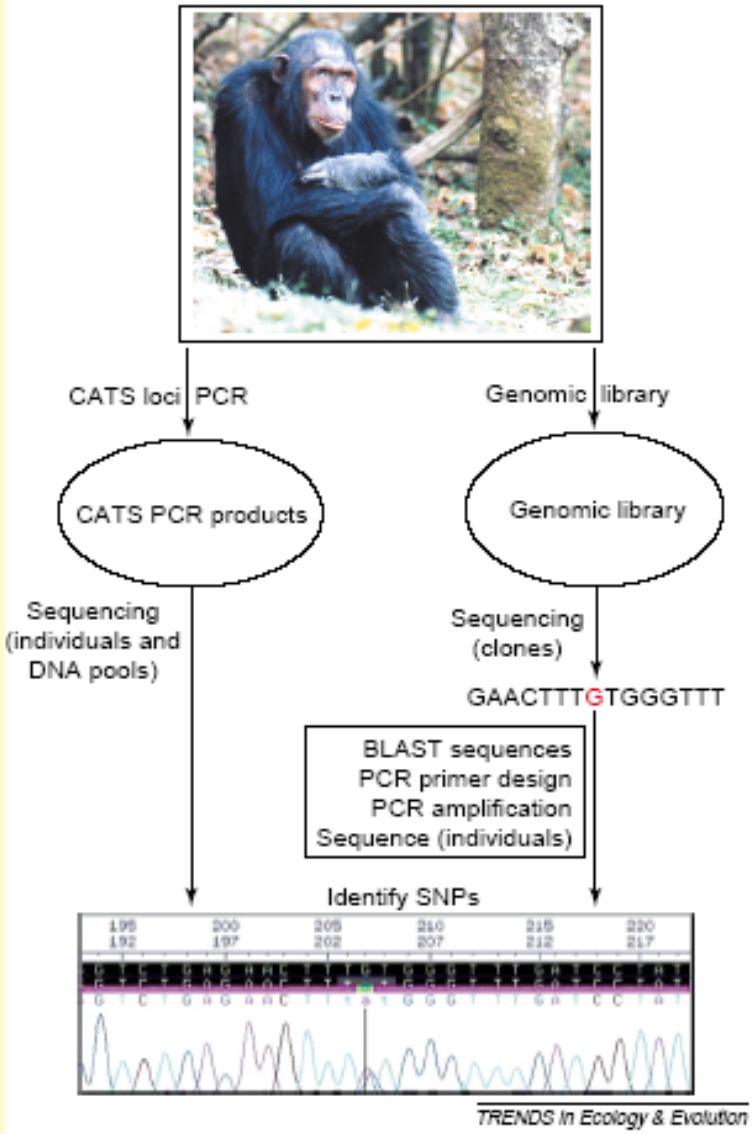
Nevýhody

- „ascertainment bias“ – výběr znaků se provádí na základě jen malého počtu jedinců a nemusí být reprezentativní
- nízká variabilita na lokus (většinou jen 2 alely)
- pro populační genetiku je vyžadován větší počet lokusů (4-10 krát více než u mikrosatelitů)

Metody analýzy

1. Nalezení lokusů („ascertainment“)
2. Genotypizace

1. Nalezení SNPs

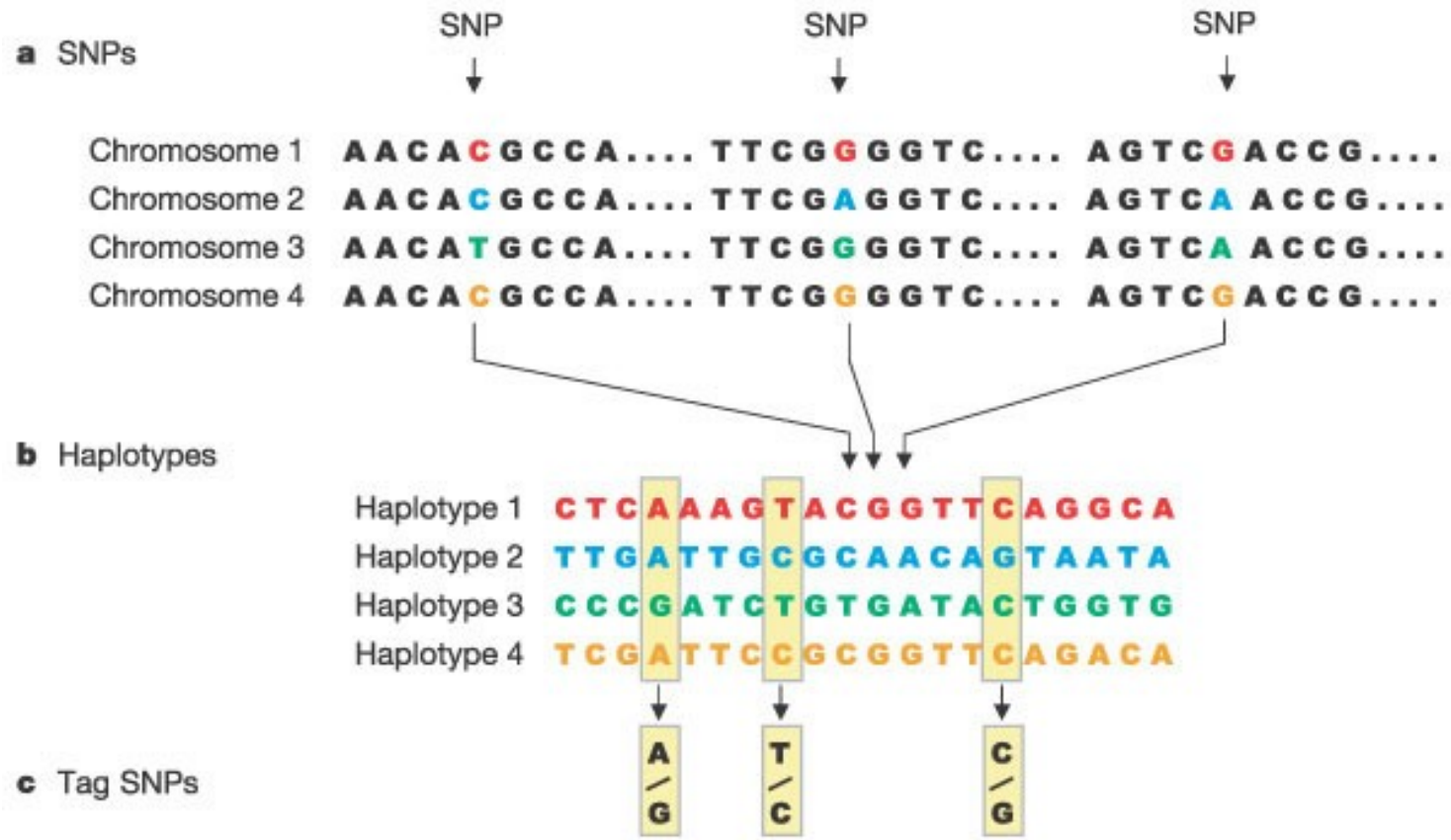


(1) CATS loci = comparative anchor tagged site loci (= cross amplification)

(2) Genomic library = genome restriction + cloning (náhodný výběr klonů – 1 SNP každých 300-1000 bp)

V současné době: Next-generation sequencing – sekvenování genomu více jedinců a hledání polymorfismů

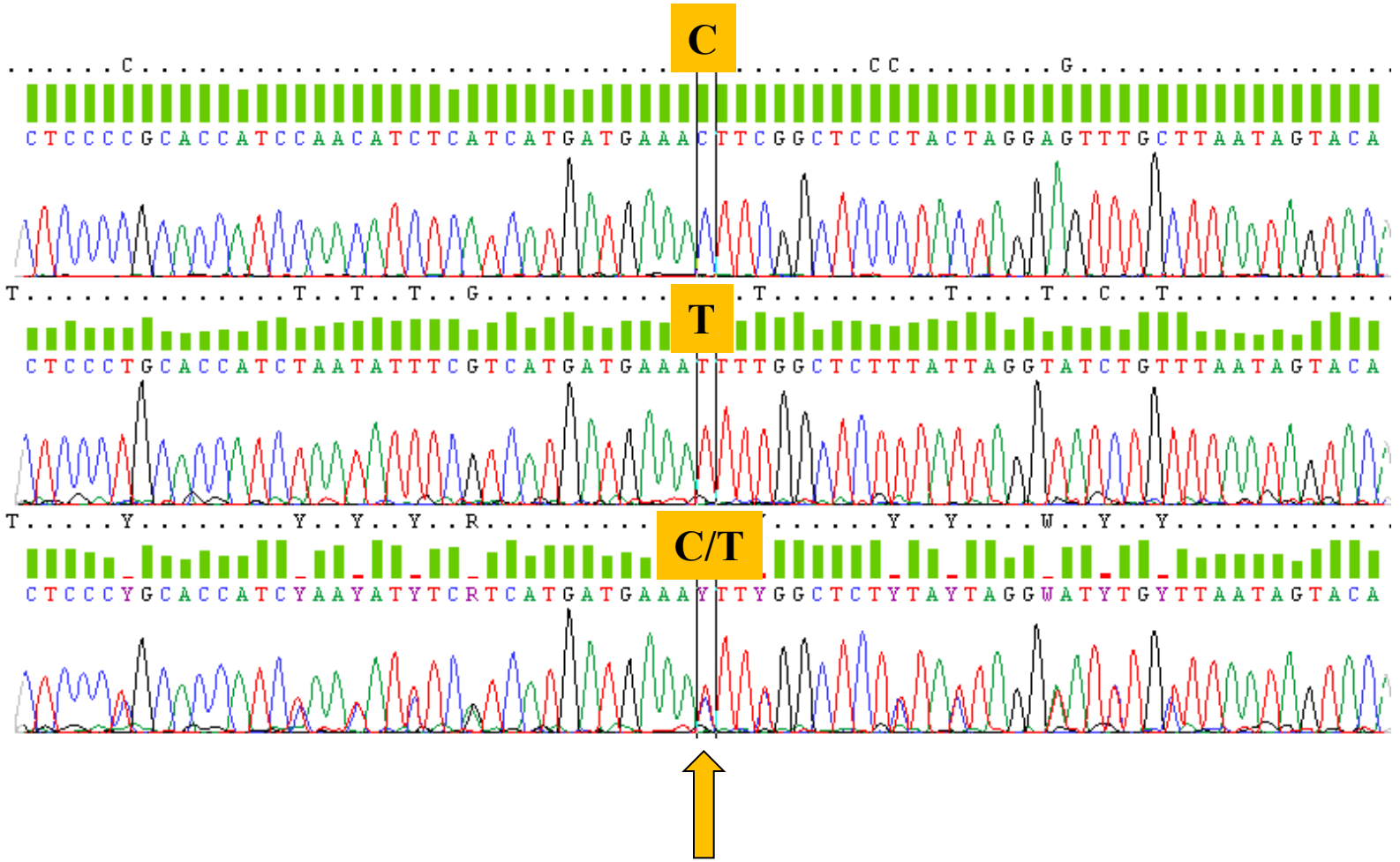
Analýza NGS dat:
Identifikace různých genotypů u různých jedinců
(= homologních chromozómů, tj. variabilita alel)



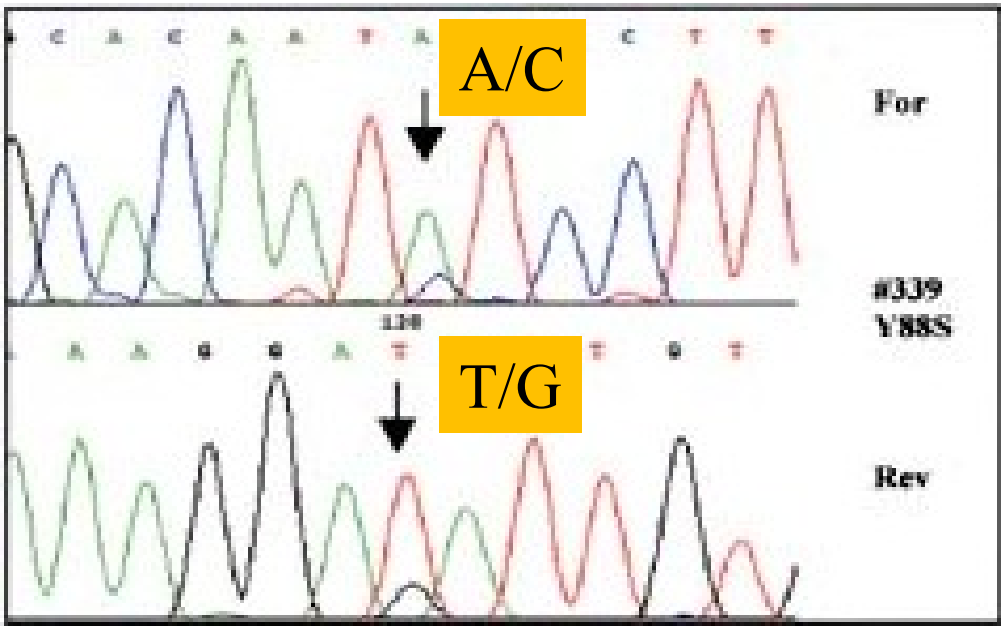
2. SNPs genotyping

= zjištění genotypu daného jedince

SNPs genotyping – sekvenování?
Je drahé a nejasné u heterozygotů

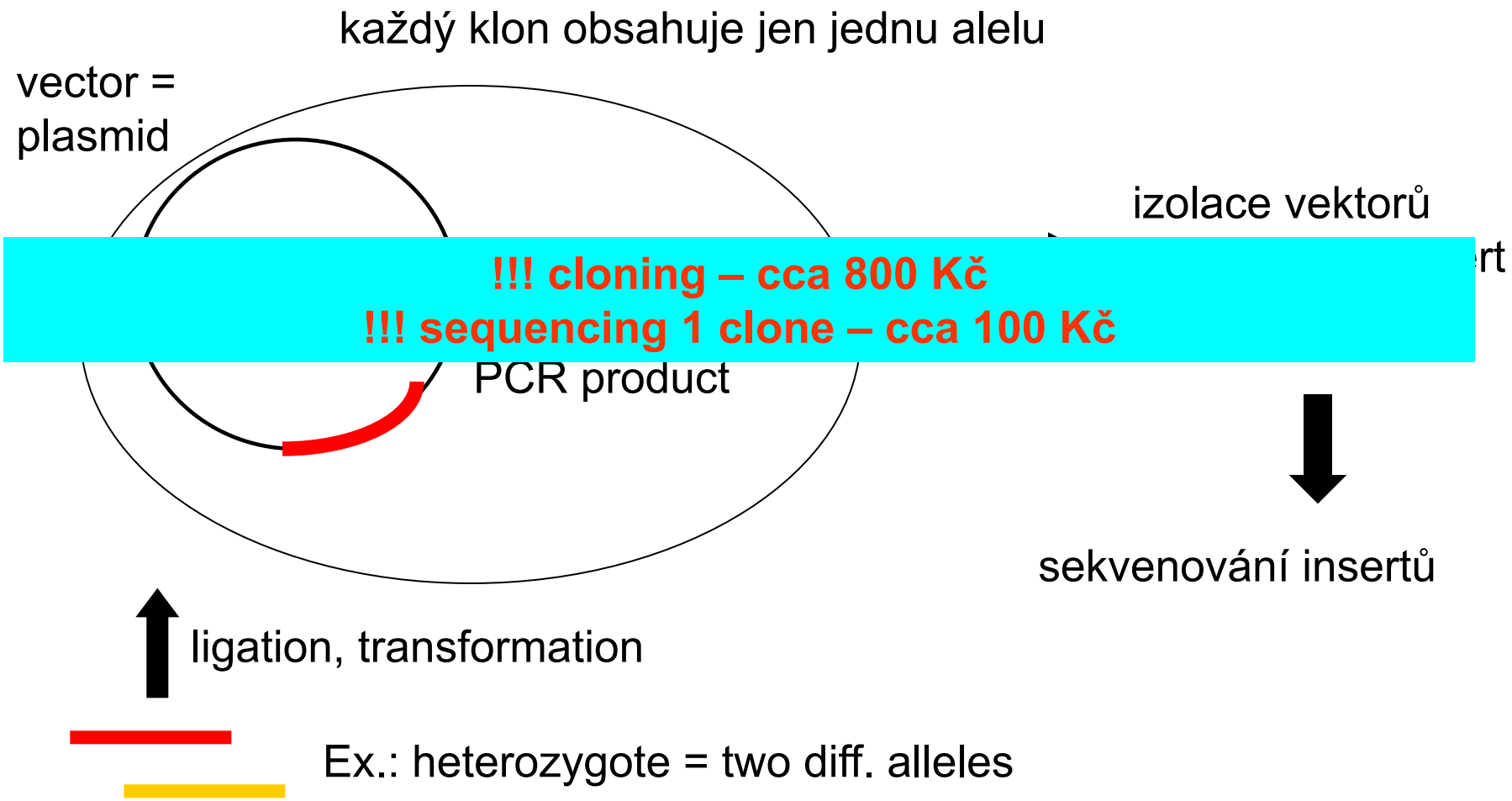


Heterozygotes?



Sekvenování z obou stran – are you really sure?

SNPs genotyping – klonování a následné sekvenování? - rozdělení dvou alel (či více u duplikovaných genů)

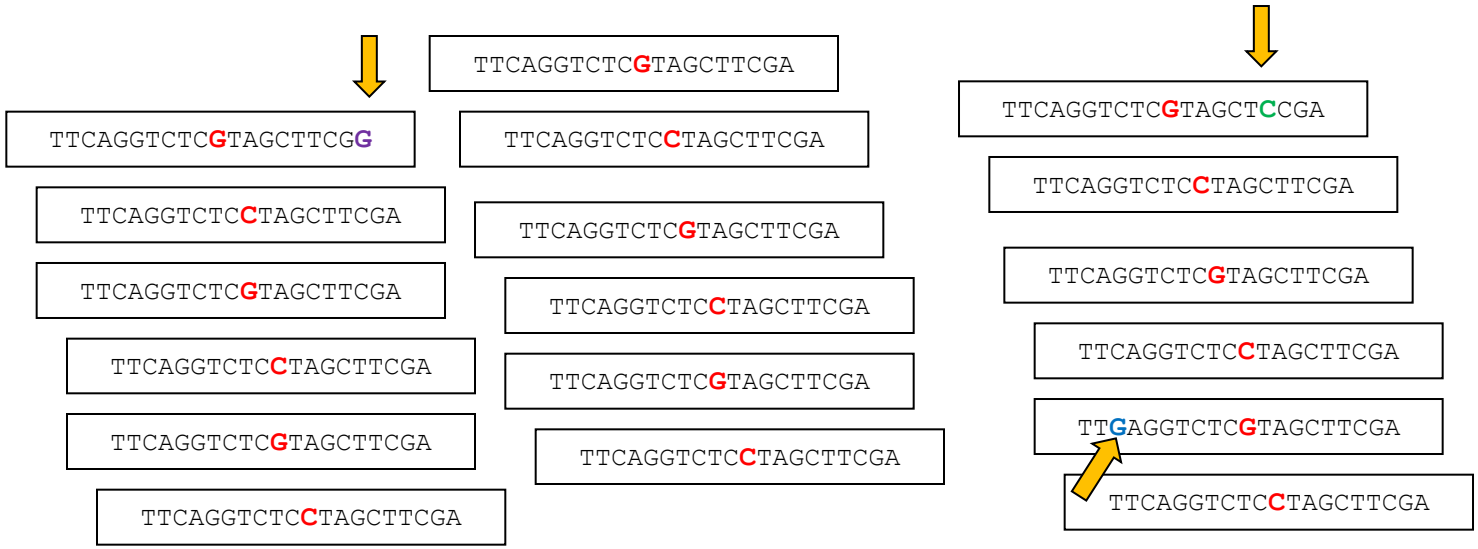


PCR is making substitution errors that **are visualised by cloning (!)**

TTCAGGTCTCGTAGCTTCGA

TTCAGGTCTCCTAGCTTCGA

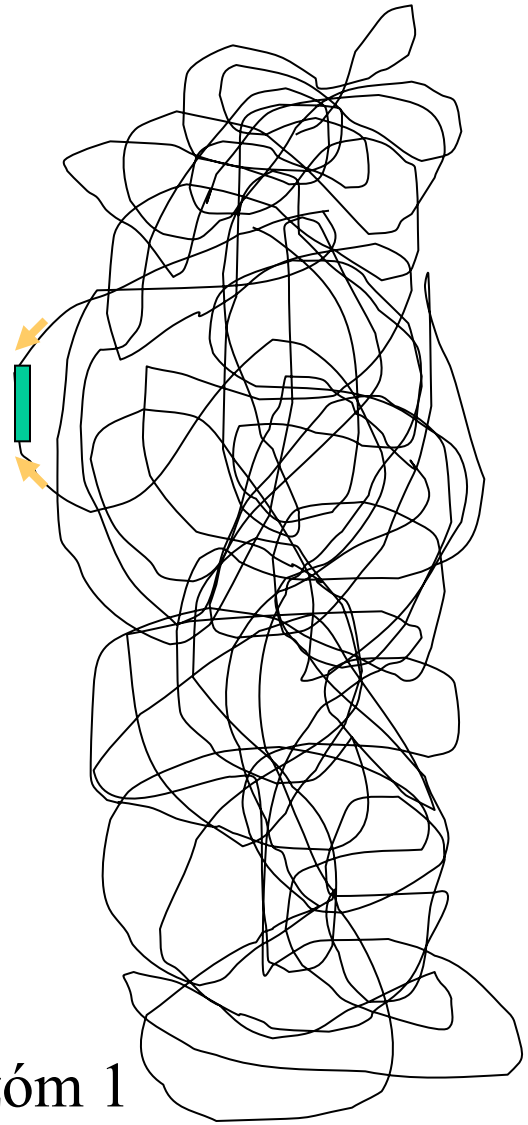
... před PCR = heterozygot G/C



PCR artefacts

Single-locus genetic markers

- SNPs (single nucleotide polymorphisms) – sekvenční polymorfismus
- kodominantní – je možné odlišit heterozygota (např. A/T) od homozygota (např. A/A)



CAAGTA
TGGACG

CATGTA
TGCACG

A/T

CAAGTA
TGGACG

CAAGTA
TGGACG

A/A

Př.: chromozóm 1

SNPs genotyping

1. **Old standards (PCR-based)**

- RFLP: PCR + štěpení + standardní elfo
- DGGE, TGGE, SSCP: PCR + nestandardní elfo
- původně detekce geneticky podmíněných chorob, např. cystická fibróza

2. **New methods (not based on standard PCR)**

- HRM: high-resolution melting (real-time PCR)
- real-time PCR se specifickými sondami (TaqMan, molecular beacon)
- ASPE: allele-specific primer extension
- SBE: single base extension
- SNP microarrays (GeneChip method)

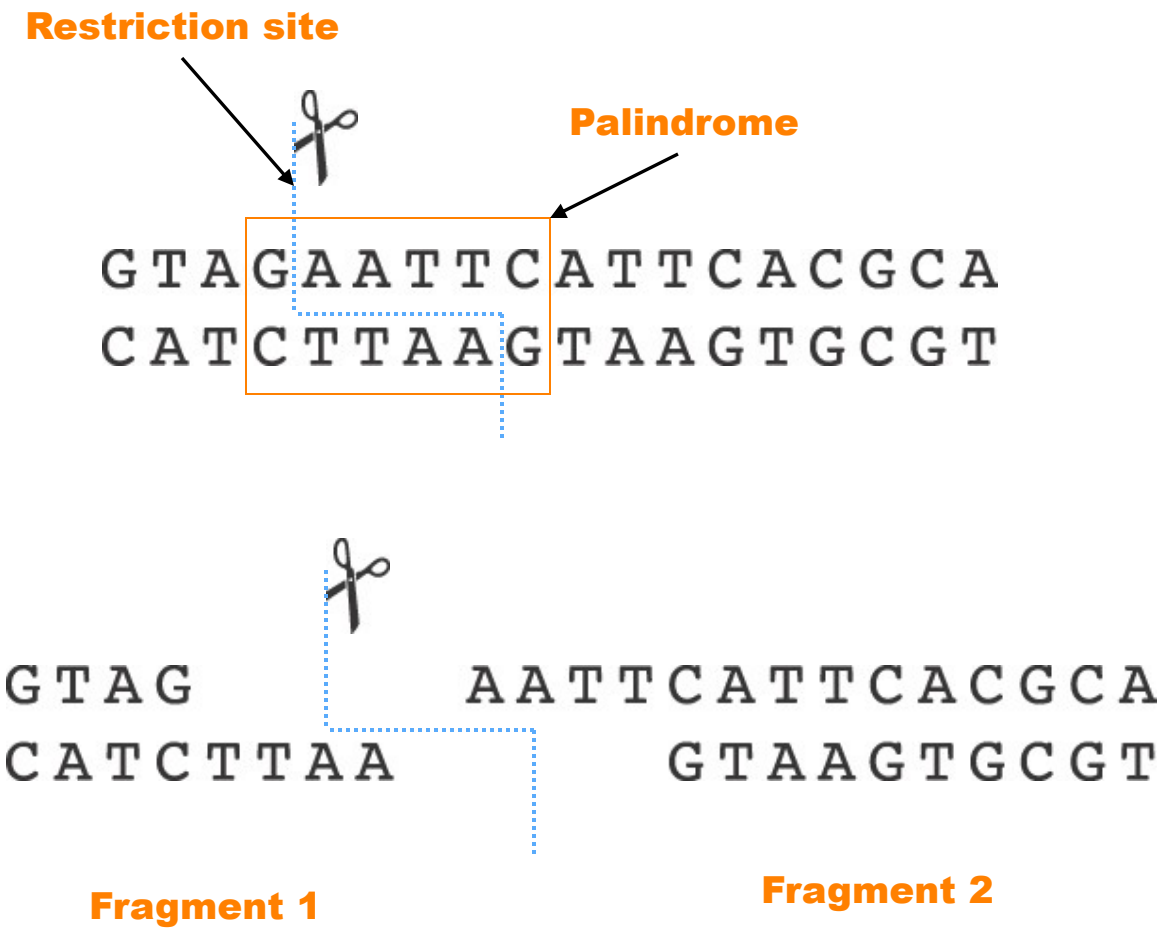
SNP genotyping - old standards

PCR-RFLP

(restriction fragments length polymorphism)

Enzyme Site Recognition

- Each enzyme digests (cuts) DNA at a specific sequence = **restriction site**
- Enzymes recognize 4- or 6- base pair, **palindromic sequences (eg GAATTC)**



Běžné restrikční enzymy

EcoRI
– *Eschericha coli*
– 5 prime overhang

PstI
– *Providencia stuartii*
– 3 prime overhang

SNP genotyping - old standards

PCR-RFLP

Allele A

CCGATCA**A**TGCGGCAA
GGCTAGT**T**ACGCCGTT

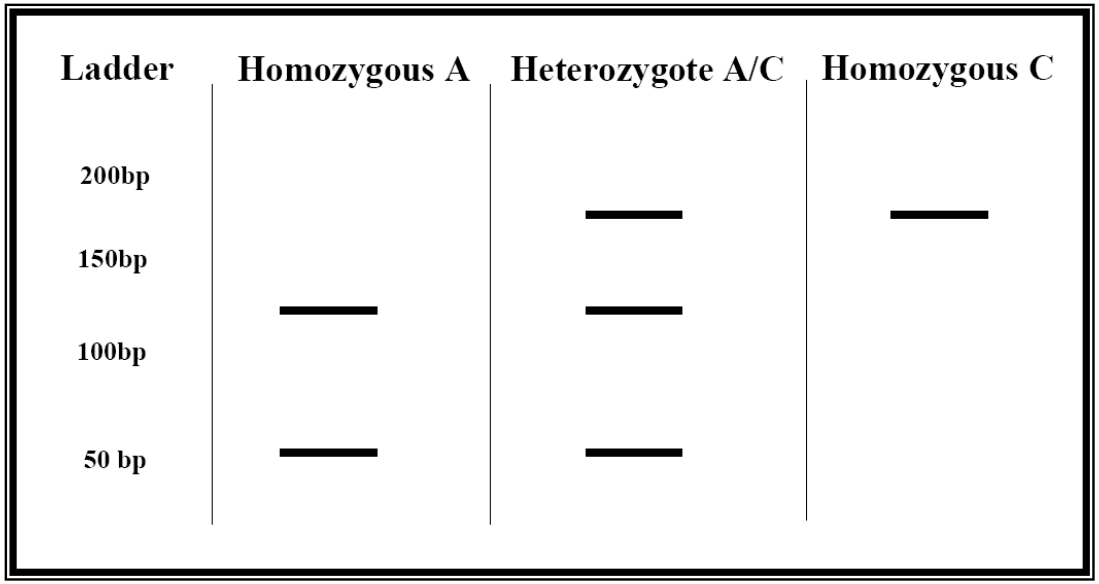
cutting by restriction endonuclease

Allele C

CCGATCA**C**TGCGGCAA
GGCTAGT**G**ACGCCGTT

no cut

- neumožní nalézt novou variantu daného SNP (odliší pouze 2 formy daného znaku: +/-)



SNPs genotyping – old standards electrophoresis methods of mutation detection

- Thermal gradient gel electrophoresis (**TGGE**)
- Denaturing gradient gel electrophoresis (**DGGE**)
- Single-strand conformation polymorphism (**SSCP**)

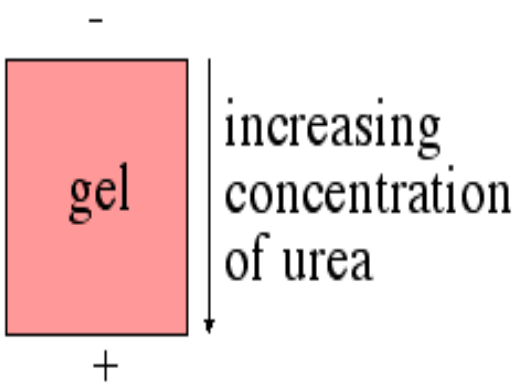
= special electrophoresis methods based on differences in mobility of different DNA sequences

Denaturing gradient gel electrophoresis (DGGE) (TGGE – podobné, ale gradient teploty)

Krátké PCR fragmenty (200-700 bp) jsou separovány v denaturačním gradientu (PAGE = polyakrylamidový gel)

Fragmenty se stejnou délkou denaturují v odlišnou dobu v závislosti na koncentraci močoviny

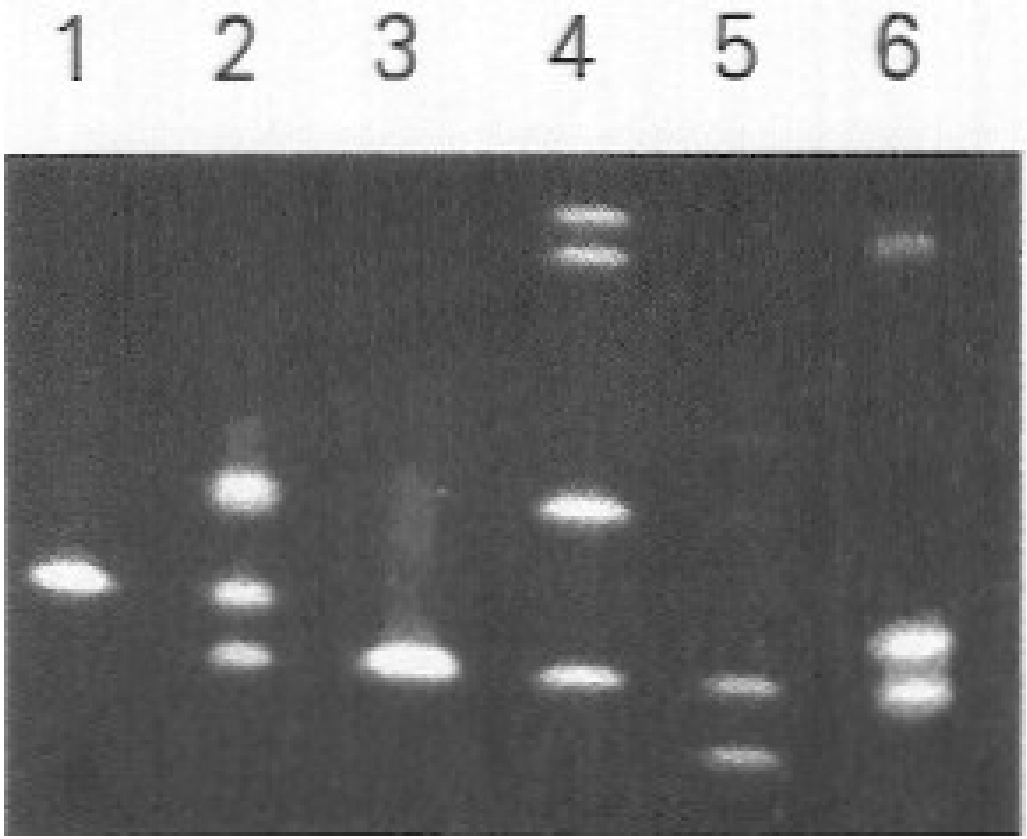
=A POINT where the DNA BEGINS TO MELT



Denaturované fragmenty putují v gelu pomaleji

Po obarvení lze vidět rozdílné pozice PCR produktů v závislosti na jejich sekvenci

Detekce nových mutací – např. v diagnostice genetických chorob



1- normal homozygote
3- homozygous mutations will yield one band on a different position
2, 4, 5, 6 – heterozygous mutations will yield 4 bands (2 homozygous and 2 heterozygous)

NOT ALL BANDS ARE SEEN !!!!!

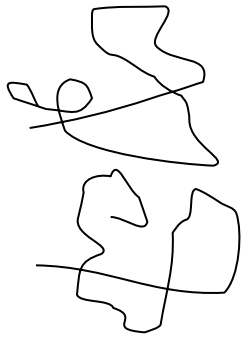
www.leveninc.com/cftr ex.gif

Single strand conformation polymorphism (SSCP)

Allele 1 - C

...CGCTTCAGG ...

...GCGAAGTCC...



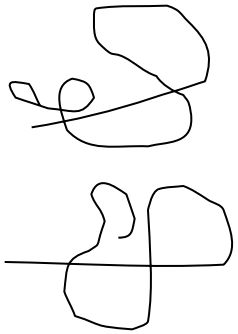
heating - denaturation

snap-cooling → partial renaturation

Allele 2 - A

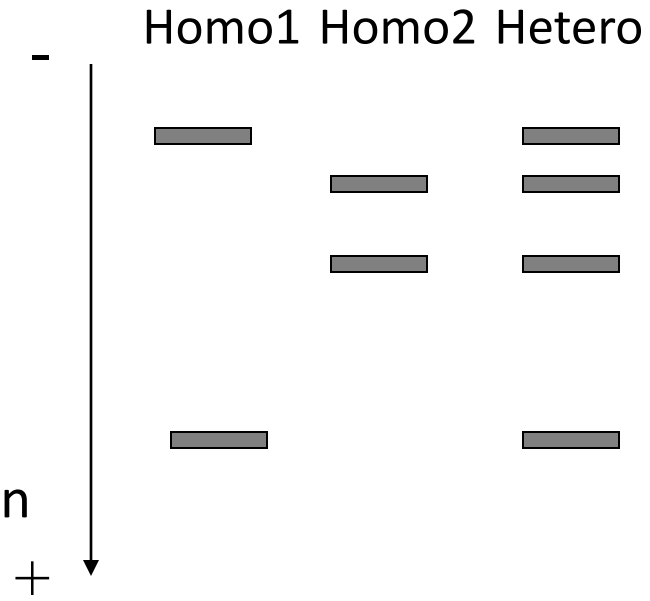
...CGCTTAAGG ...

...GCGAATTCC...



sequence-specific

ssDNA conformations



!!! non-denaturing PAGE

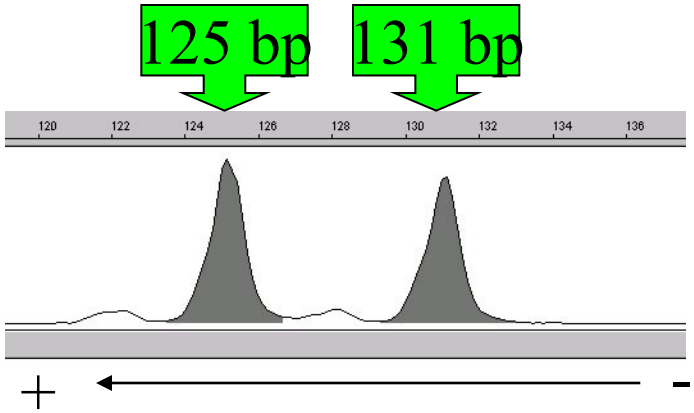
radioisotopes

silver-staining

fluorescent dyes (SYBR gold)

Použití automatických sekvenátorů

(**denaturing** polymer POP7 – ssDNA, e.g. microsatellites – one labelled primer)



Well controlled electrophoresis parameters, high sensitivity

Použití automatických sekvenátorů

Why not non-denaturing electrophoresis?
e.g. CAP (conformation analysis polymer)

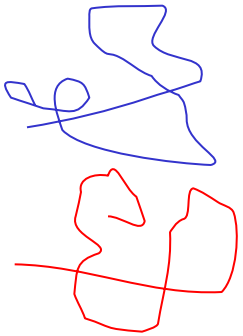
- well controlled electrophoresis

- two fluorescent labels

- high sensitivity

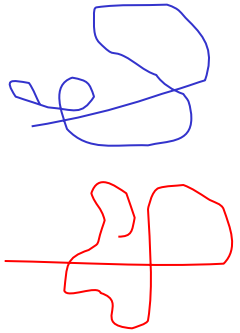
Allele 1

FAM... **C**GCTT**C**AGG ...
... GCGAA**G**TC**C** ***...HEX***

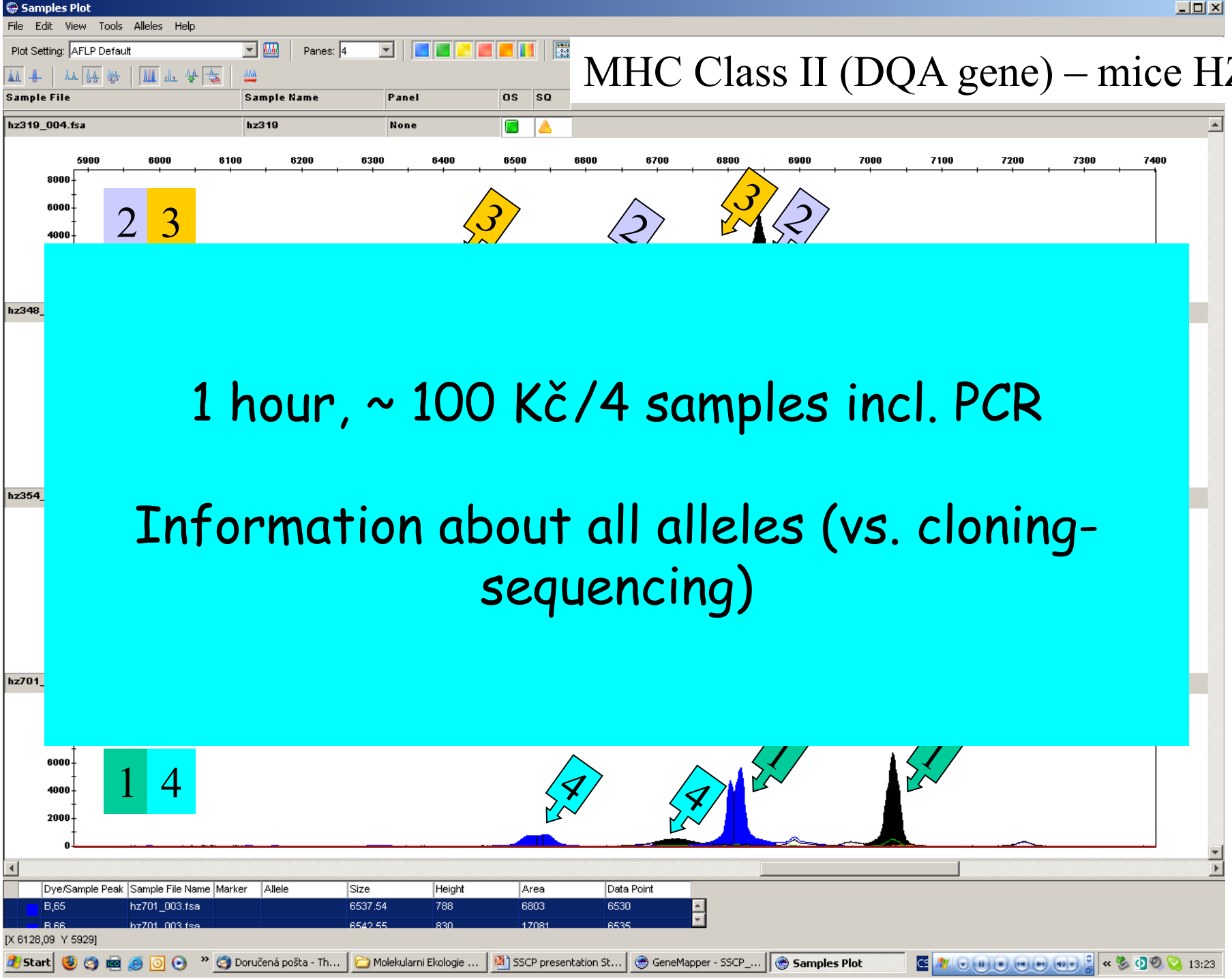


Allele 2

FAM... **C**GCTT**A**AGG ...
... GCGAA**T**TC**C** ***...HEX***



MHC Class II (DQA gene) – mice HZ



1 hour, ~ 100 Kč/4 samples incl. PCR

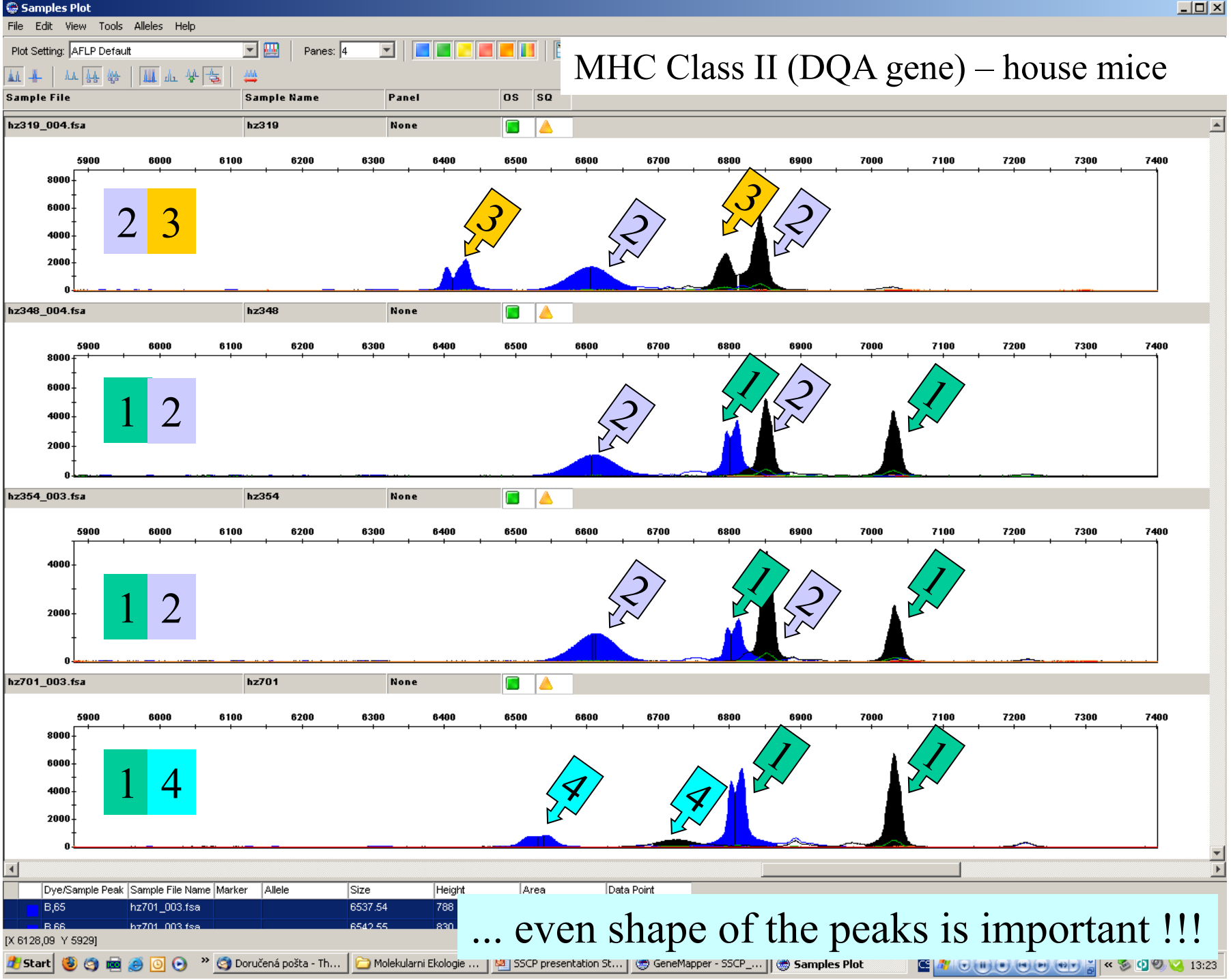
Information about all alleles (vs. cloning-sequencing)

Analýza elektroforetogramů

- např. GeneMapper (Applied Biosystems)
- specifický „Size+Conformation Standard“ pro každou teplotu
- srovnání více vzorků
- umožňuje detekci krátkých odlišných sekvencí s více SNPs (užitečné např. pro genotypizaci MHC, tj. vysoce variabilních genů)

Použití

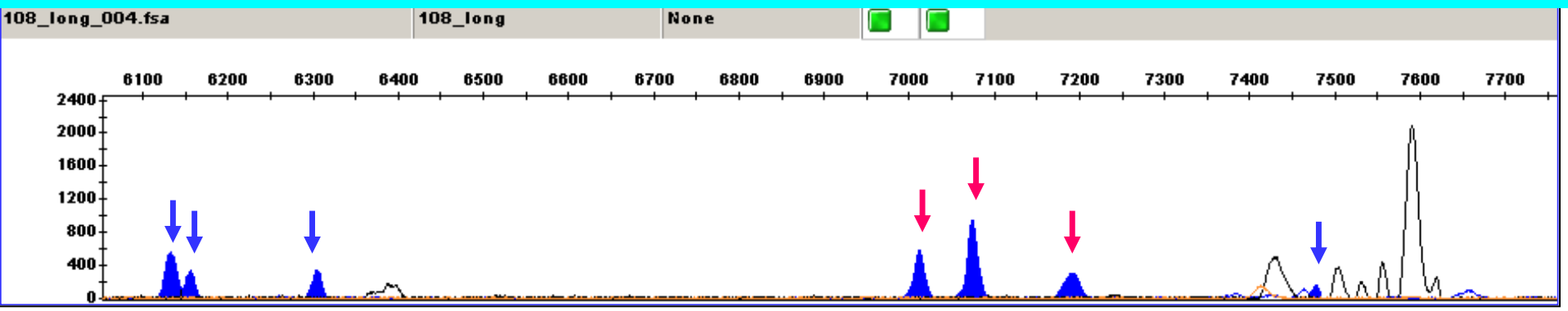
1) Genotyping of codominant markers (e.g. single copy MHC genes)


Samples Plot
File Edit View Tools Alleles Help
Plot Setting: AFLP Default
Panes: 4
MHC Class II (DQA gene) – house mice
Sample File
Sample Name
Panel
OS
SQ
hz319_004.fsa
hz319
None
2 3
hz348_004.fsa
hz348
None
1 2
hz354_003.fsa
hz354
None
1 2
hz701_003.fsa
hz701
None
1 4
Dye/Sample Peak
Sample File Name
Marker
Allele
Size
Height
Area
Data Point
B,65
hz701_003.fsa
6537.54
788
... even shape of the peaks is important !!!
Start
Doručená pošta - Th...
Molekularni Ekologie ...
SSCP presentation St...
GeneMapper - SSCP_...
Samples Plot
13:23

Použití

1) Genotyping of codominant markers (e.g. single copy MHC genes)
2) Identification of number of genes (e.g. duplicated MHC genes)

Sedm píků stejné barvy=
= alespoň čtyři kopie daného genu !!!



SSCP of three individuals: - different alleles - same alleles

Carpodacus erythrinus – MHC Class I (Promerová et al. 2009)

Použití

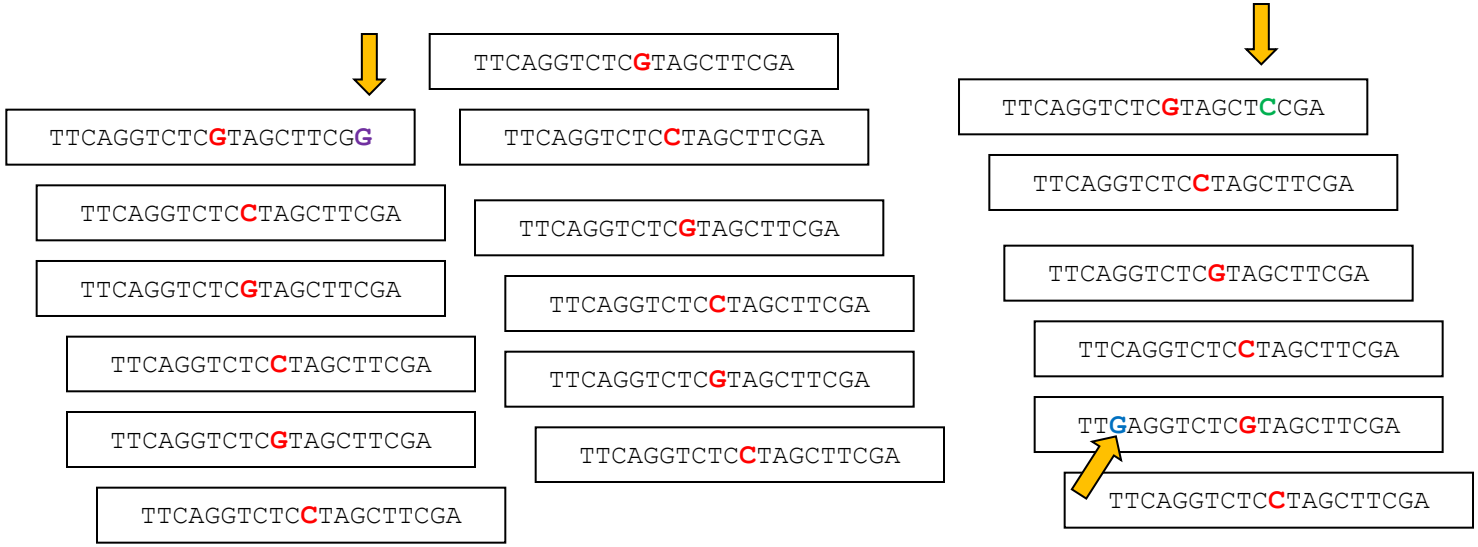
1) Genotyping of codominant markers (e.g. single copy MHC genes)
2) Identification of number of genes (e.g. duplicated MHC genes)
3) Detection of PCR artefacts during cloning

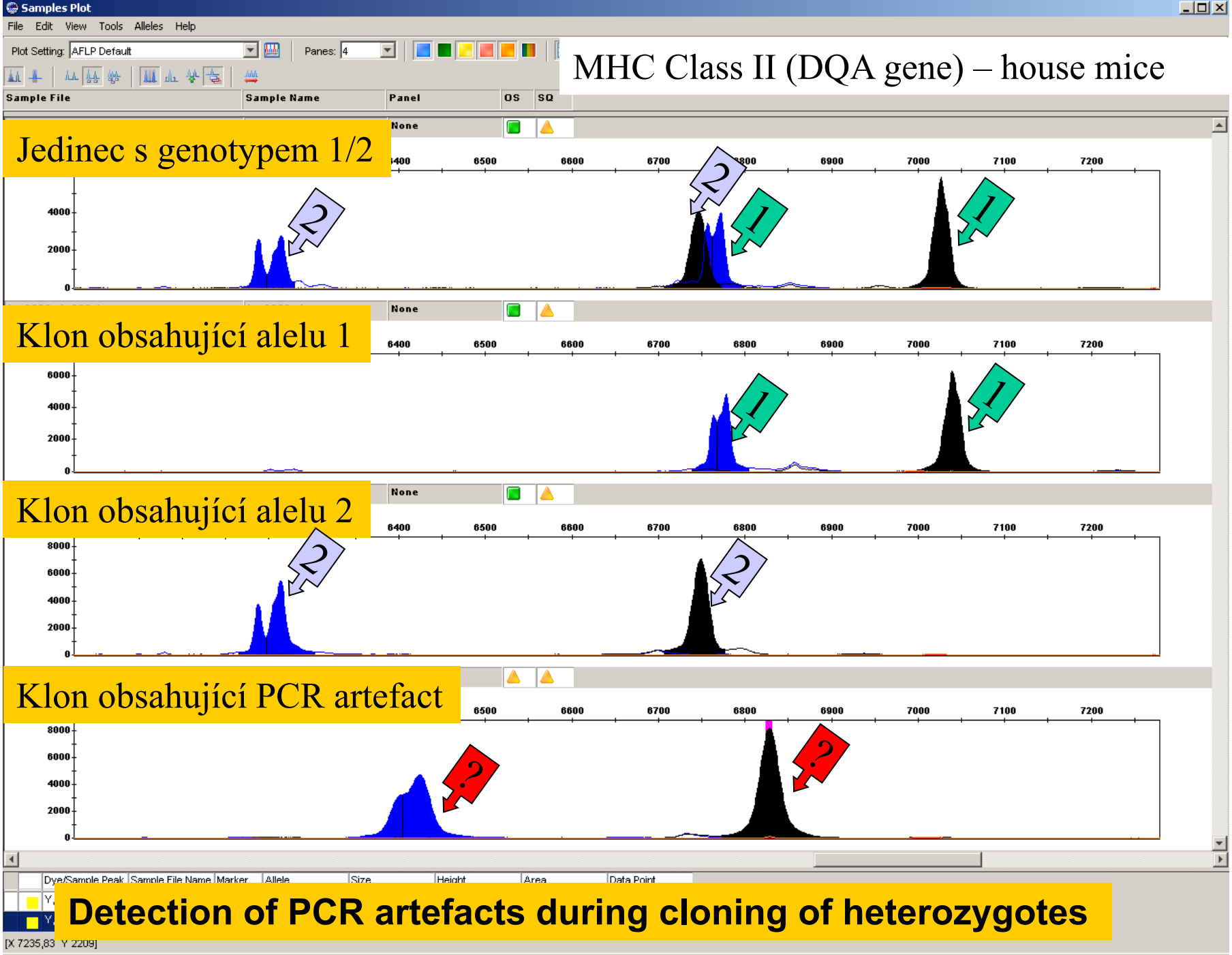
Detection of PCR artefacts during cloning

TTCAGGTCTCGTAGCTTCGA

TTCAGGTCTCCTAGCTTCGA

... před PCR = heterozygot G/C



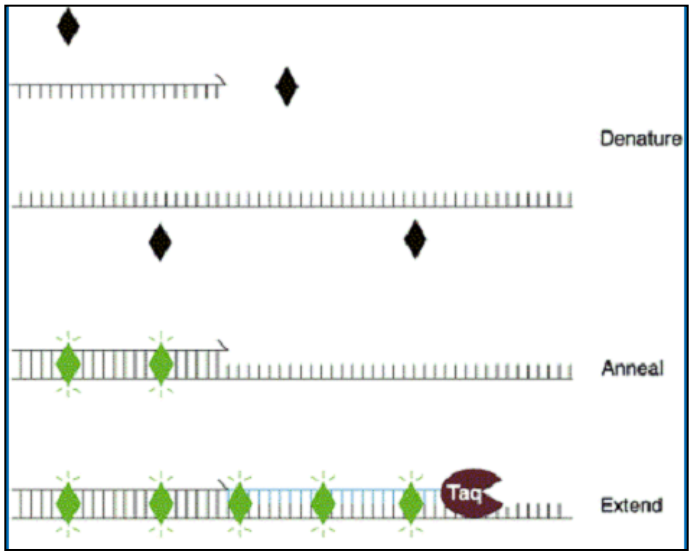

Samples Plot
File Edit View Tools Alleles Help
Plot Setting: AFLP Default
Panes: 4
Sample File
Sample Name
Panel
OS
SQ
MHC Class II (DQA gene) – house mice
Jedinec s genotypem 1/2
Klon obsahující alelu 1
Klon obsahující alelu 2
Klon obsahující PCR artefact
Detection of PCR artefacts during cloning of heterozygotes
Start
Samples Plot
17:08

SNP genotyping – new methods

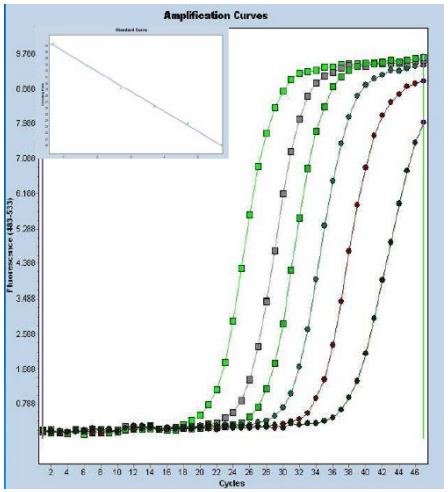
= not based on standard PCR

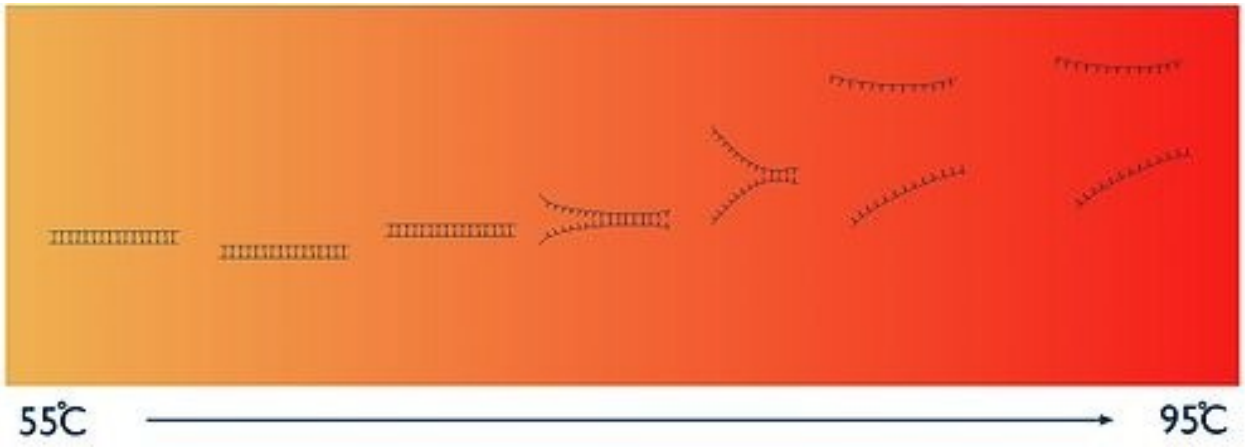
1. high-resolution melting temperature (HRMT)
2. real-time PCR se specifickými sondami (TaqMan, molecular beacon)
3. ASPE: allele-specific primer extension
4. SBE: single base extension
5. SNP microarrays (GeneChip method)

1. High-resolution melting temperature (HRMT)



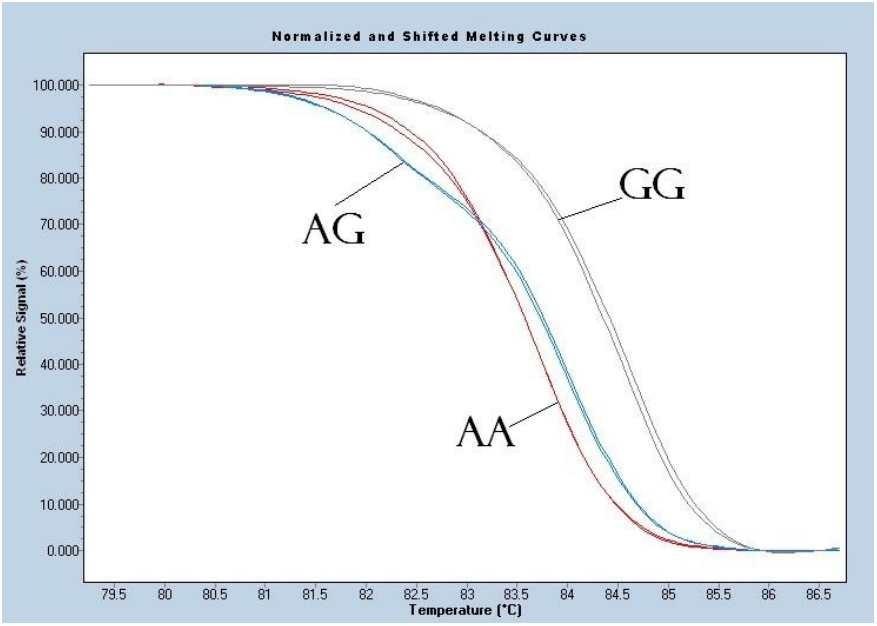
Step 1: real-time PCR = increase of fluorescence

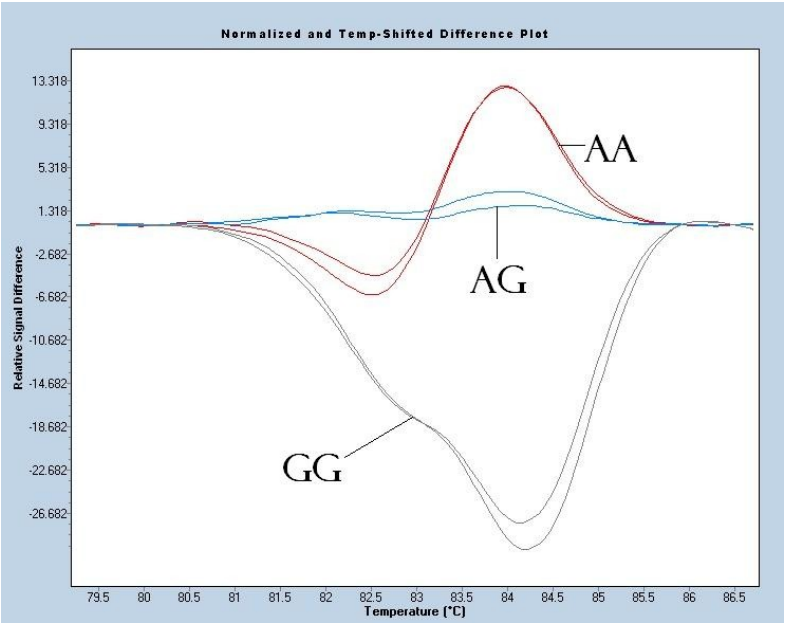




Step 2: measuring melting after PCR = decrease of fluorescence

HRMT genotyping

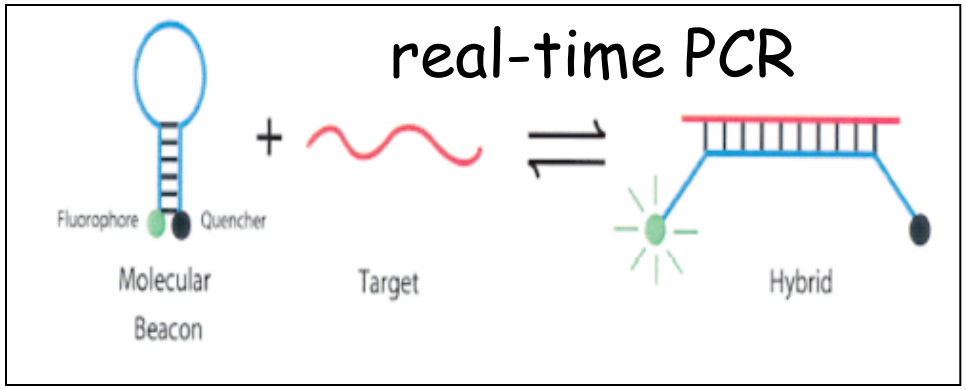




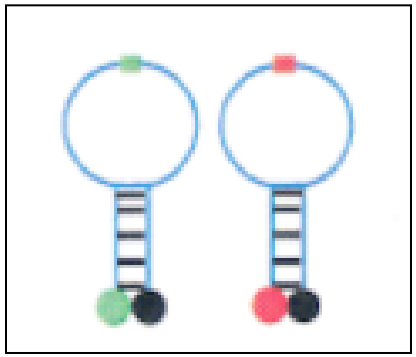
Detekce heterozygotů

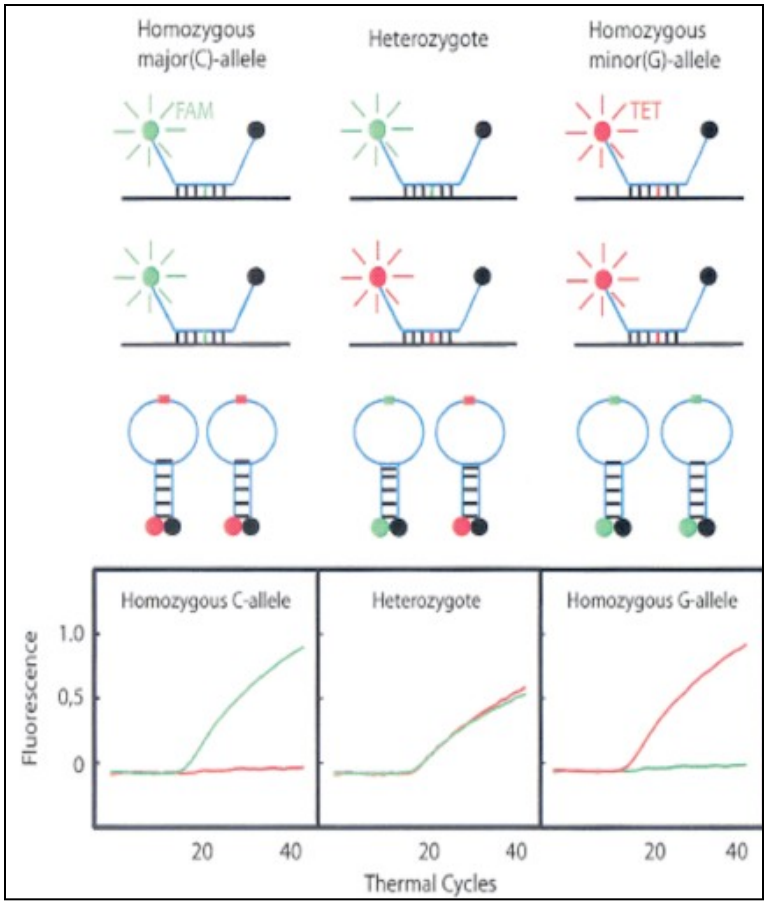
- velmi levná a jednoduchá metoda – v podstatě jen qPCR
- vhodná na genotypizace jednoduchých SNP u velkého množství vzorků

2. Real-time PCR se specifickou sondou



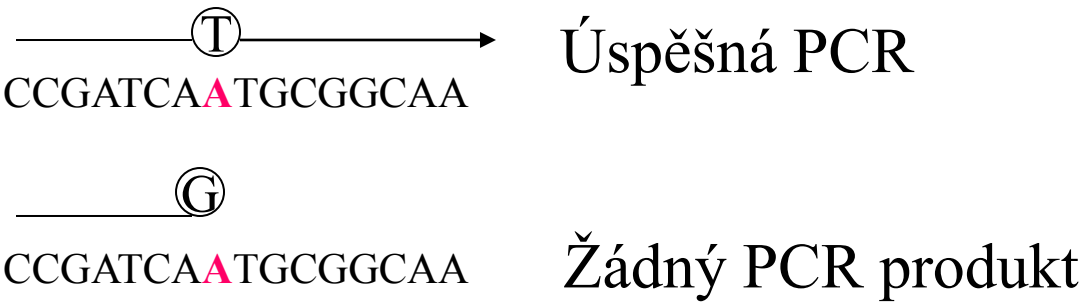
sondy specifické pro jednotlivé alely





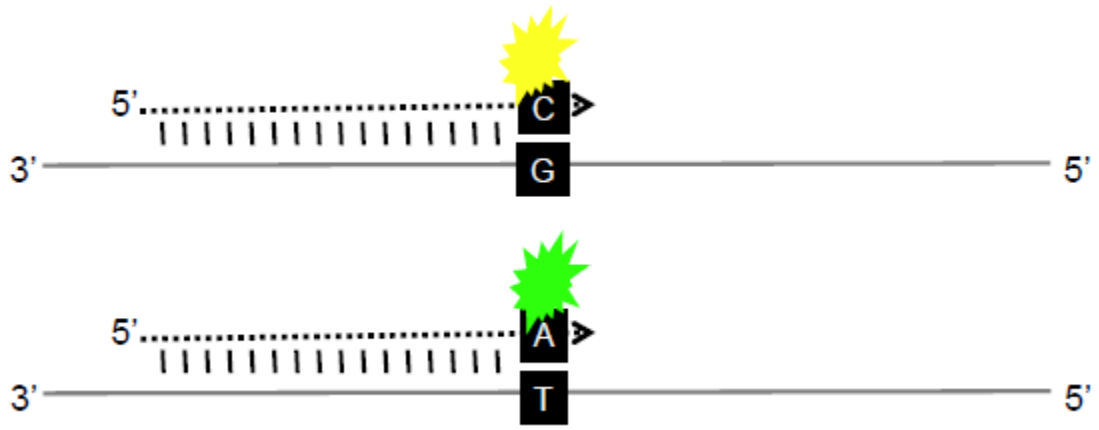
1) TaqMan sondy
2) Molecular Beacons („maják“)

3. ASPE: allele-specific primer extension



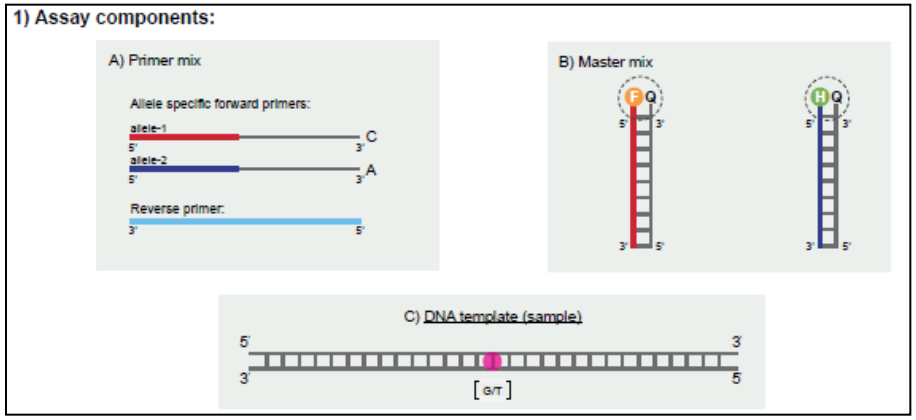
- dvě PCR se specifickými primery
- 3’ terminální nukleotid na primerech je komplementární k SNP nukleotidu
- alelově-specifická amplifikace je umožněna vysoce specifickou polymerázou

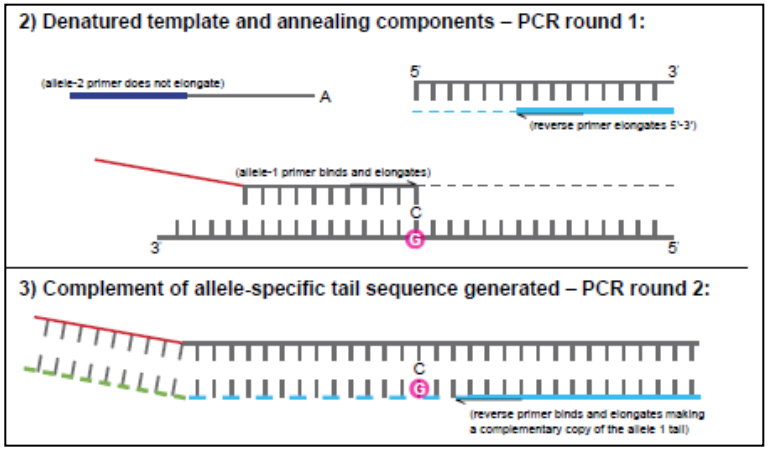
ASPE: allele-specific primer extension (automatizovaná verze)

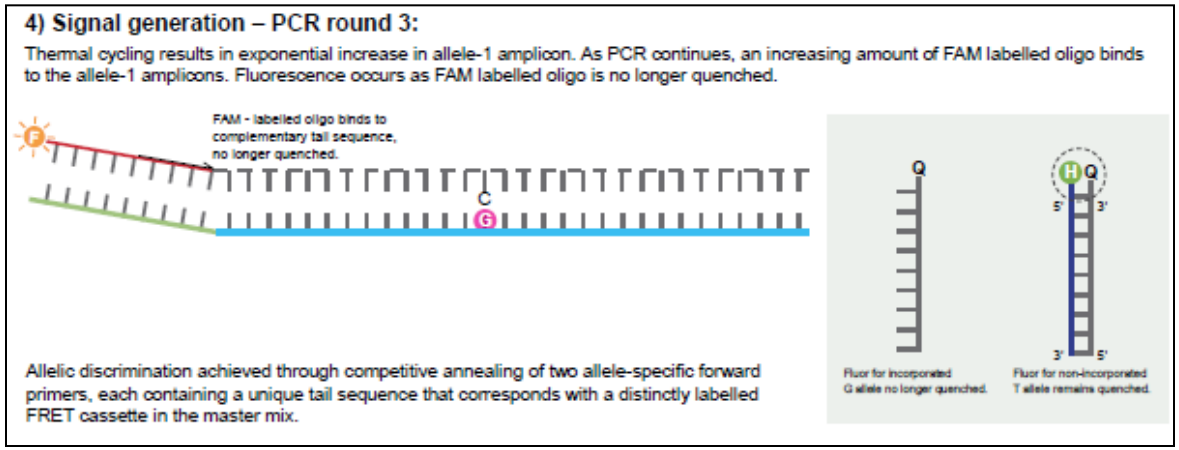


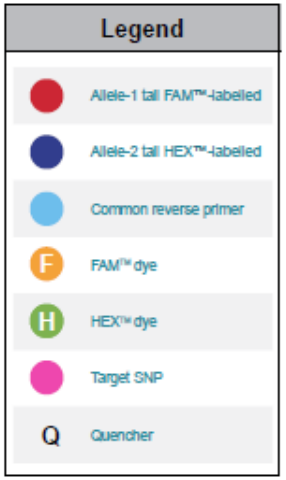
- existují zoptimalizované multiplexy pro modelové druhy (např. člověk 1536 SNPs)
- fluorescenční detekce (Illumina)

Kompetitive Alelle Specific PCR









The KASP™ genotyping assay from **LGC Genomics**

LGC Genomics incorporating biosearch

Cena analýzy („outsourcing")

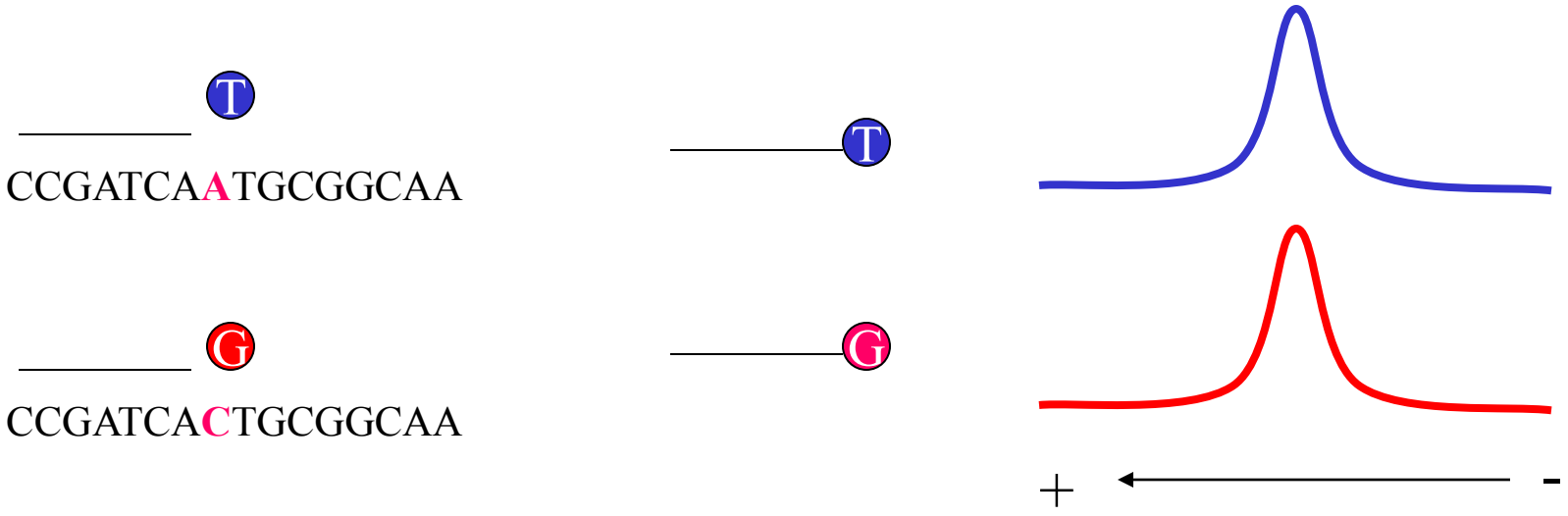
Small scale study of 15 SNPs genotyped over 96 samples where no Assay on Demand (an alternative type of assay from ABI) SNP exists

	LGC Genomics cost	ABI Taqman® Assay by Design
SNP assay design costs (validation)	**£1,620.00**	£6,750.00
Genotyping cost	**£701.50**	£388.80
Total	**£2,321.50**	£7,138.80

The KASP™ genotyping assay from **LGC Genomics**

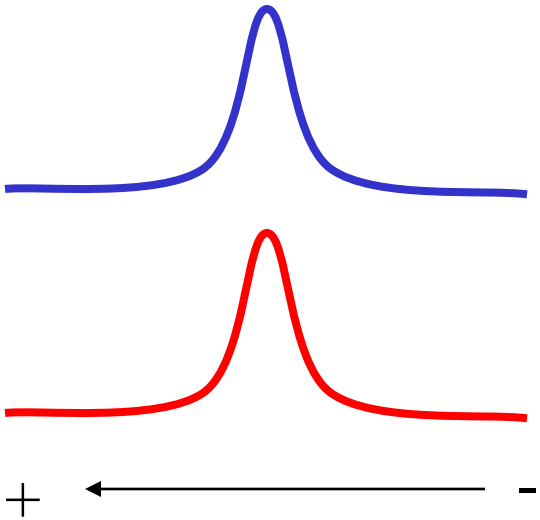
LGC Genomics incorporating KBioscience

4. SBE: single base extension

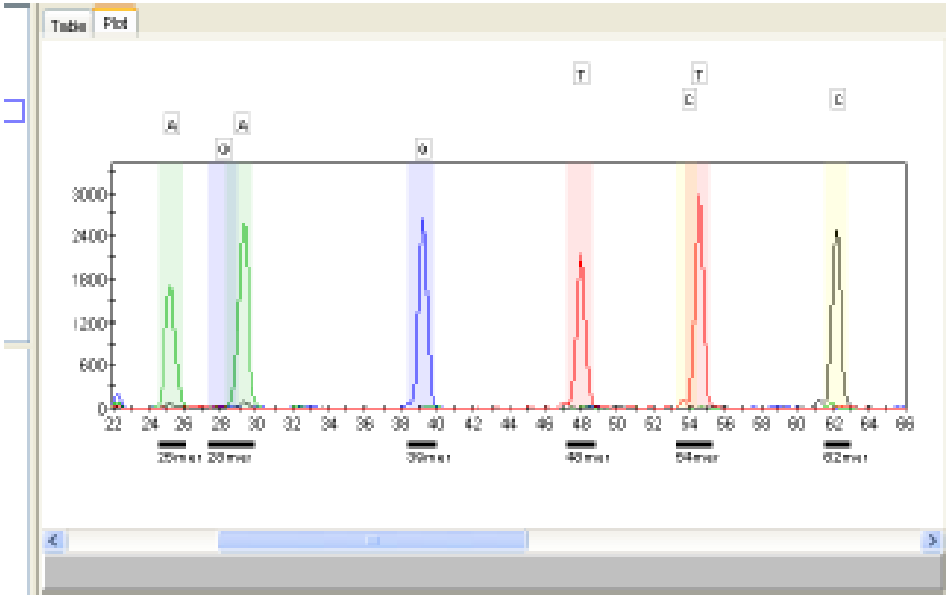


- pouze jeden dideoxynukleotid je přidán k primeru
- detekce různými metodami

Detekce SBE produktů kapilární elektroforézou



kapilární elektroforéza SNaPShot Multiplex Kit (Life Technologies)

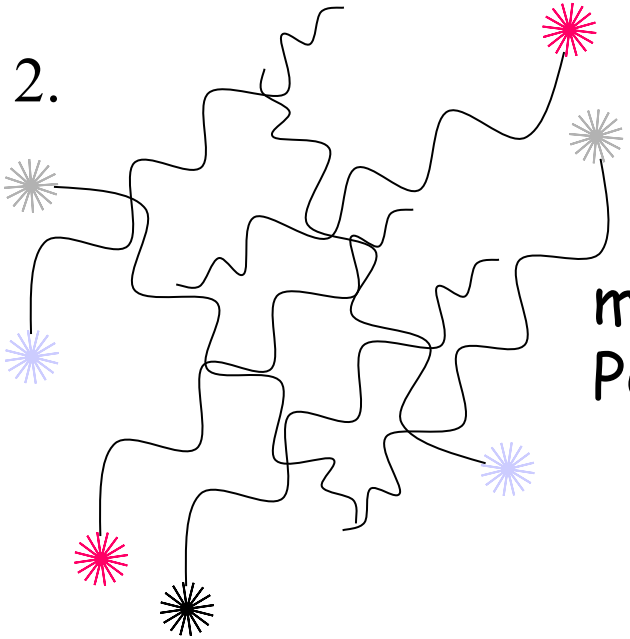


„multiplex version" – různě dlouhé primery, aby bylo možné odlišit různé lokusy

Detekce SBE produktů přes „microarray" (tj. hybridizace)

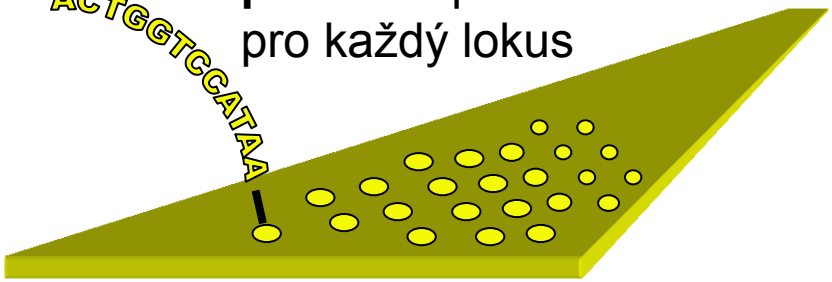
1. **tag** – specifický pro každý lokus

2.

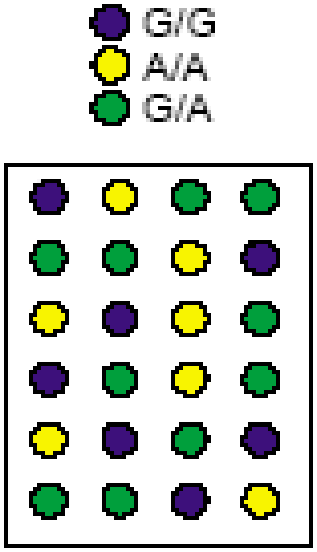


multiplex PCR

3. **tag-complementary probe –** specifická pro každý lokus



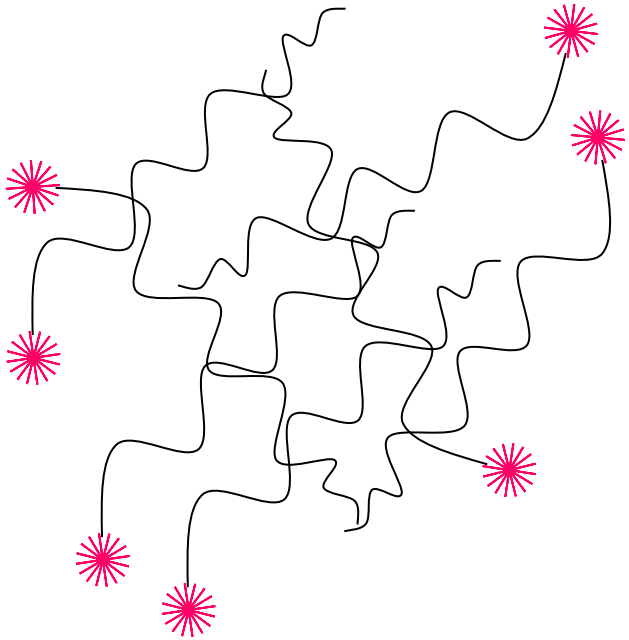
4.

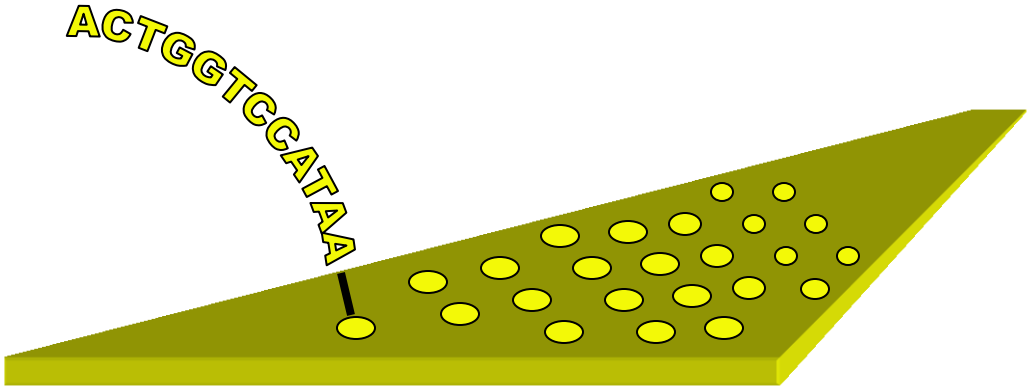


Small-scale "in house" SNP genotyping

multicolor detection (using of 5' oligonucleotide tags on SBE primers)

5. Microarray analýza SNPs
(celogenomový přístup – „chip technology")

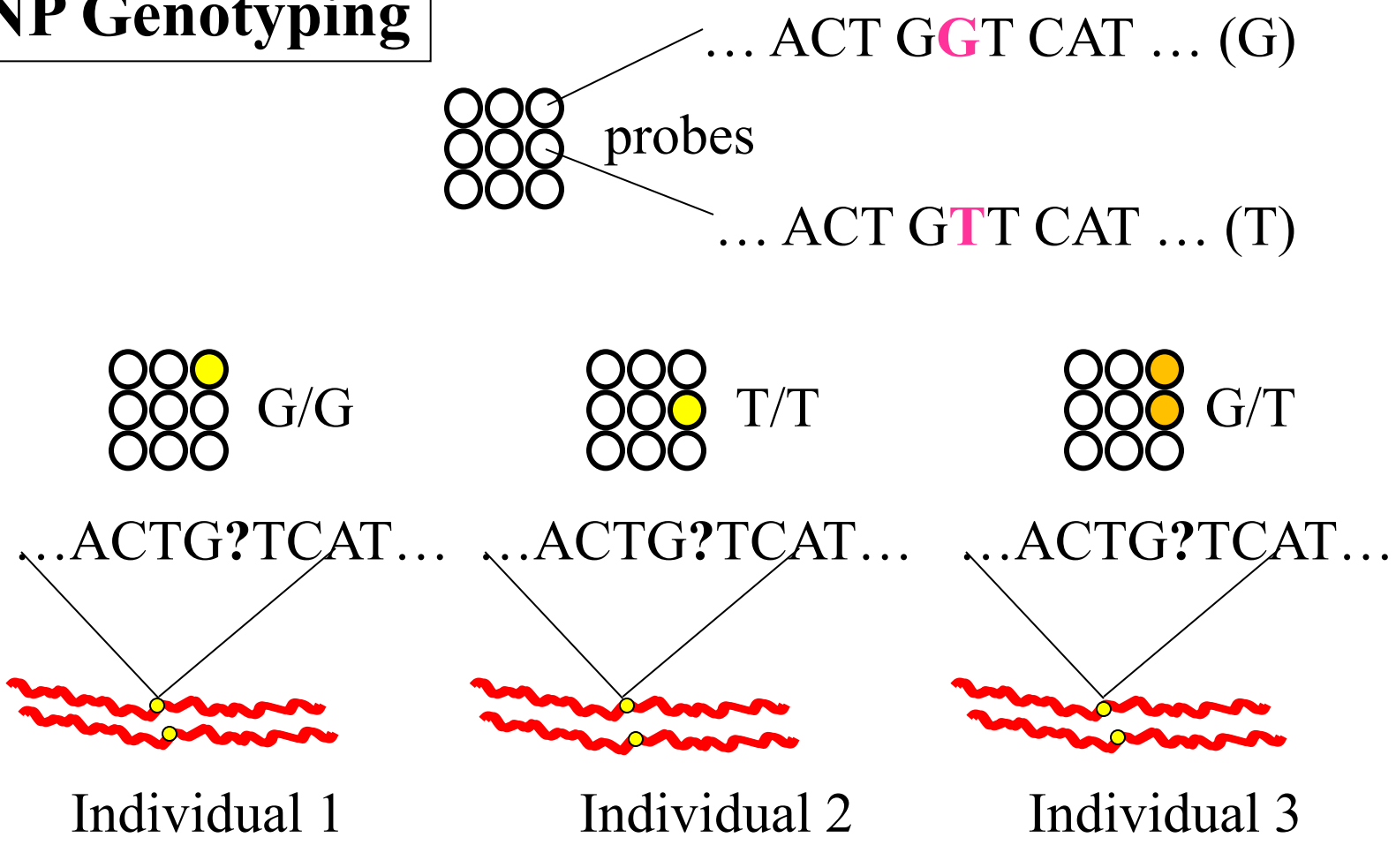




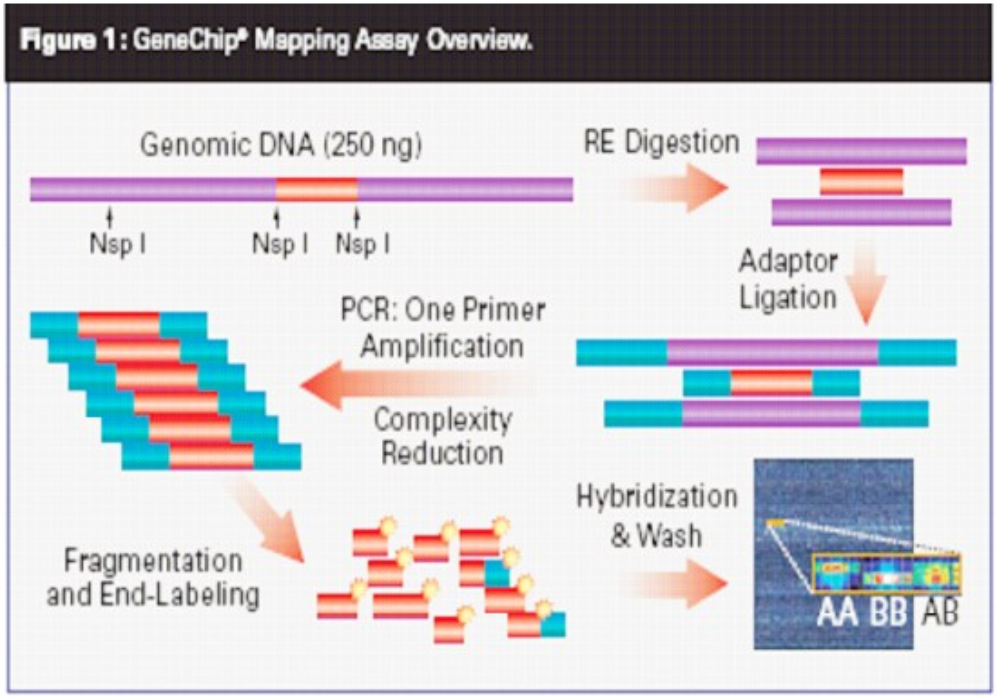
Target (genomická DNA rozštěpená restrikčními enzymy)

Probe
(specifická sonda pro každou alelu)

Microarray SNP Genotyping



Detekce: např. Affymetrix



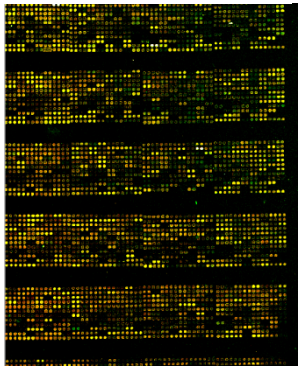
Affymetrix® Mouse Diversity Genotyping Array

- 10 tisíc – 500 tisíc SNP znaků najednou – „chip technology“

- např. Mouse Diversity Genotyping Array – 623 tisíc SNPs (je známa pozice každého z nich na genomu)

- je možné si navrhnout vlastní Array

Affymetrix®
Revolutionize life



Fees - Whole Genome Genotyping									
Platform	SNP multiplex	# samples per array	# genotypes	array $	reagent $	core fee $	Project price per sample	Project price per genotype	volume discount bins
Affymetrix 10K	10,000	1	10,000	185	50	255	**$490.00**	$0.0490	
Affymetrix 50K	50,000	1	50,000	210	50	255	**$515.00**	$0.0103	
Affymetrix 100K (50K x2)	100,000	1	100,000	420	100	510	**$920.00**	$0.0092	
Affymetrix 250K	250,000	1	250,000	470	55	255	**$780.00**	$0.0031	
Affymetrix 500K (250K x2)	500,000	1	500,000	940	110	510	**$1,560.00**	$0.0031	
Affymetrix 500K (250K x2)	500,000	1	500,000	800	110	510	**$1,420.00**	$0.0028	1000-2000 samples
Affymetrix 500K (250K x2)	500,000	1	500,000	700	110	510	**$1,320.00**	$0.0026	2001-5000 samples
Illumina Human-1	109,000	1	109,000	800	na	110	**$910.00**	$0.0083	1-256 samples
Illumina Human-1	109,000	1	109,000	720	na	110	**$830.00**	$0.0076	257-496 samples
Illumina Human-1	109,000	1	109,000	640	na	110	**$750.00**	$0.0069	497-736 samples
Illumina Human-1	109,000	1	109,000	560	na	110	**$670.00**	$0.0061	737-976 samples
Illumina Human-1	109,000	1	109,000	480	na	110	**$590.00**	$0.0054	977+ samles
Illumina HumanHap300	317,000	1	317,000	1100	na	110	**$1,210.00**	$0.0038	1-256 samples
Illumina HumanHap300	317,000	1	317,000	990	na	110	**$1,100.00**	$0.0035	257-496 samples
Illumi									ples
Illumi									ples
Illumi									mles
Illumi									ples
Illumi									ples
Illumi									ples
Illumina HumanHapS	240,000	1	240,000	700	na	110	**$810.00**	$0.0034	737-976 samples
Illumina HumanHapS	240,000	1	240,000	600	na	110	**$710.00**	$0.0030	977+ samles
Illumina HumanHap550	550,000	1	550,000	1600	na	110	**$1,710.00**	$0.0031	1-256 samples
Illumina HumanHap550	550,000	1	550,000	1440	na	110	**$1,550.00**	$0.0028	257-496 samples
Illumina HumanHap550	550,000	1	550,000	1280	na	110	**$1,390.00**	$0.0025	497-736 samples
Illumina HumanHap550	550,000	1	550,000	1120	na	110	**$1,230.00**	$0.0022	737-976 samples
Illumina HumanHap550	550,000	1	550,000	960	na	110	**$1,070.00**	$0.0019	977+ samles
HumanHap300 + HumanHapS	550,000	1	550,000	1750	na	220	**$1,970.00**	$0.0036	1-256 samples
HumanHap300 + HumanHapS	550,000	1	550,000	1575	na	220	**$1,795.00**	$0.0033	257-496 samples
HumanHap300 + HumanHapS	550,000	1	550,000	1400	na	220	**$1,620.00**	$0.0029	497-736 samples
HumanHap300 + HumanHapS	550,000	1	550,000	1225	na	220	**$1,445.00**	$0.0026	737-976 samples
HumanHap300 + HumanHapS	550,000	1	550,000	1050	na	220	**$1,270.00**	$0.0023	977+ samles

Použití u příbuzných druhů je možné, ale je tam velmi silný „ascertainment bias“