Types of genetic markers

- **multilocus** markers (RAPD, AFLP, minisatelitte DNA fingerprinting)
- **single-locus** markers (allozymes, microsatelittes, SNPs)

- **dominant** markers – scored as present or absent (RAPD, AFLP, ...)
- **codominant** markers – identification of homologous alleles, i.e. scoring of homozygote and heterozygote states (allow estimation of allele frequencies – SNPs, microsatelittes, ...)

Main markers used in molecular ecology

	Single locus	Codominant	PCR assay	Overall variability
Mitochondrial DNA (or sex specific sequences like Y chromosome)				
Sequences	Yes	Haplotypes	Yes	Low-high
Nuclear multilocus				
Minisatellite fingerprints	No	No	No	High
RAPD	No	No	Yes	High
AFLP	No	No	Yes	High
Nuclear single locus				
Allozymes	Yes	Yes	No	Low-medium
Microsatellites	Yes	Yes	Yes	High
SNPs (sequences)	Yes	Yes	Yes	Low-high

Multi-locus genetic markers

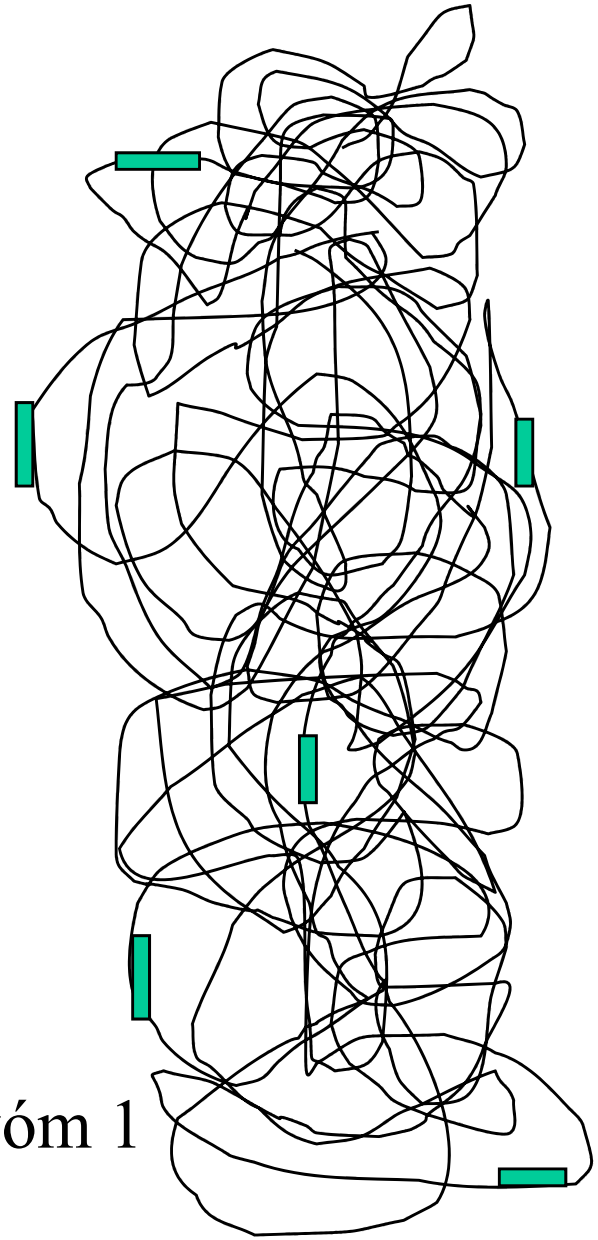
- screening of many loci distributed randomly throughout the genome

➢ minisatellite DNA fingerprinting

➢ RAPD (randomly amplified polymorphic DNA)

➢ AFLP (arbitrary or amplified fragment length polymorphism)

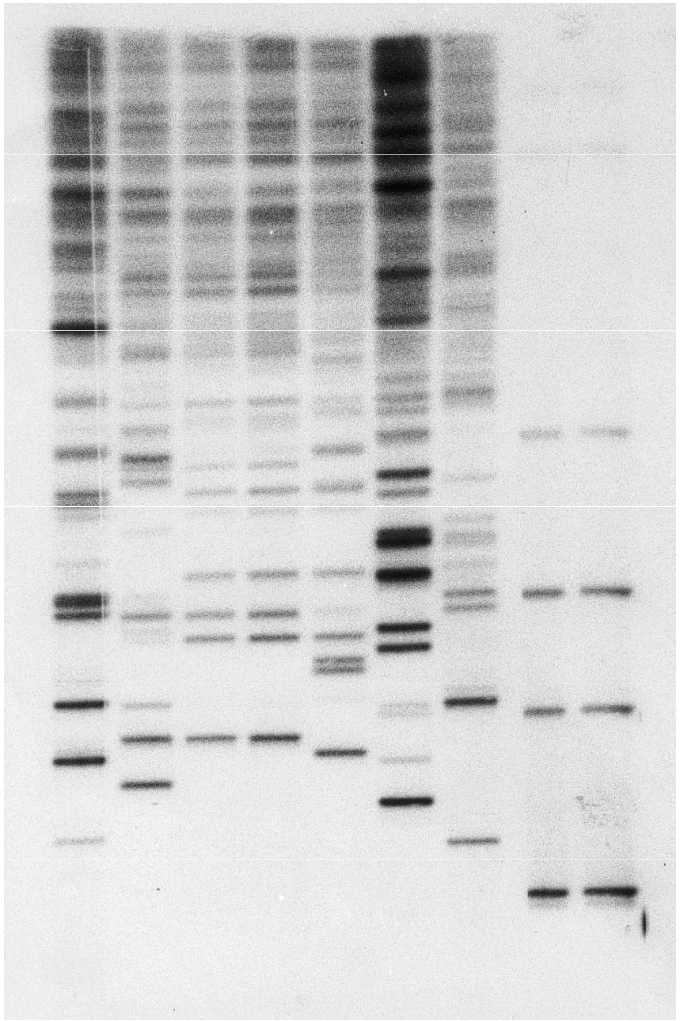
- presence vs. absence - codominant



Př.: chromozóm 1

Minisatellite DNA fingerprinting

- randomly distributed repetitions (Alu sekvence, SINE, LINE)
- restriction – sequence specific endonucleases
- electrophoresis
- blotting
- hybridization with probe
- over recent years – shift to PCR-based methods



RAPD (randomly amplified polymorphic DNA)

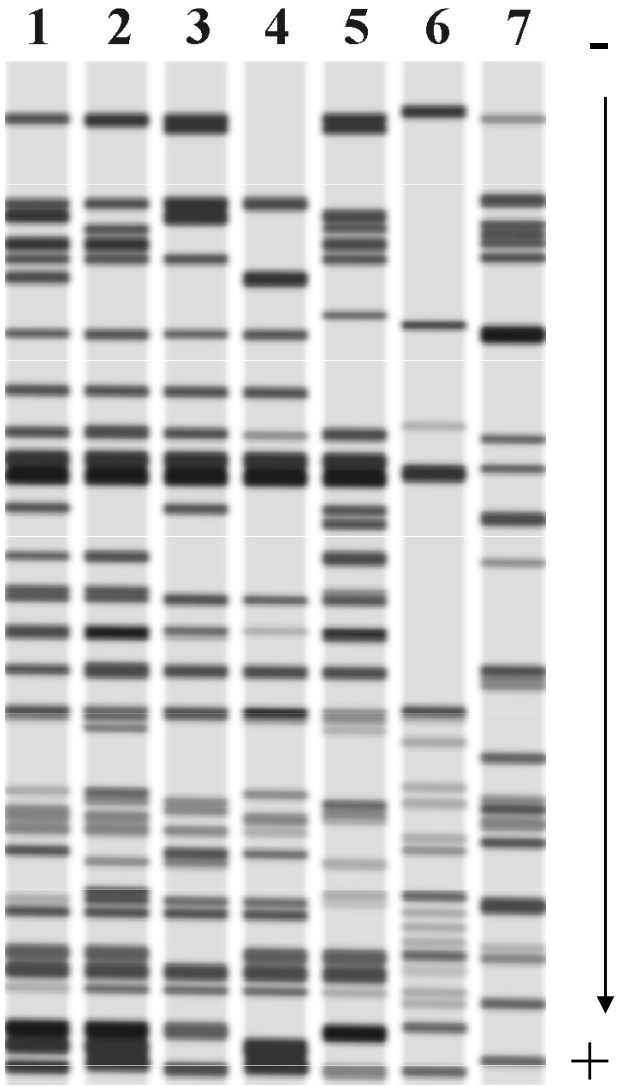
short random oligonucleotides as primers

⇩

PCR at low stringency

⇩

detection of PCR fragments by electrophoresis

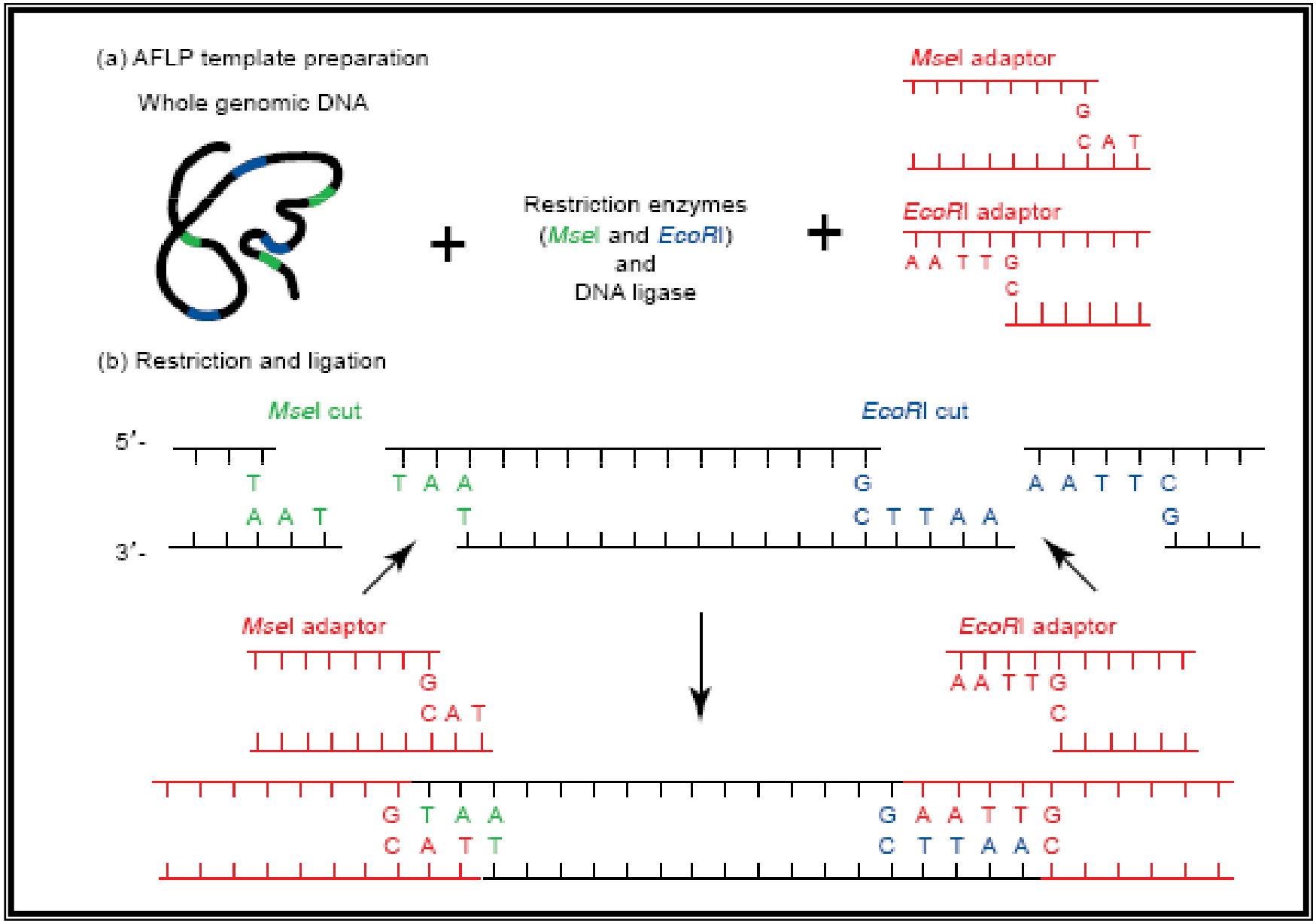


low repeatability due to many factors affecting PCR – is not more accepted as method for studies of population structure

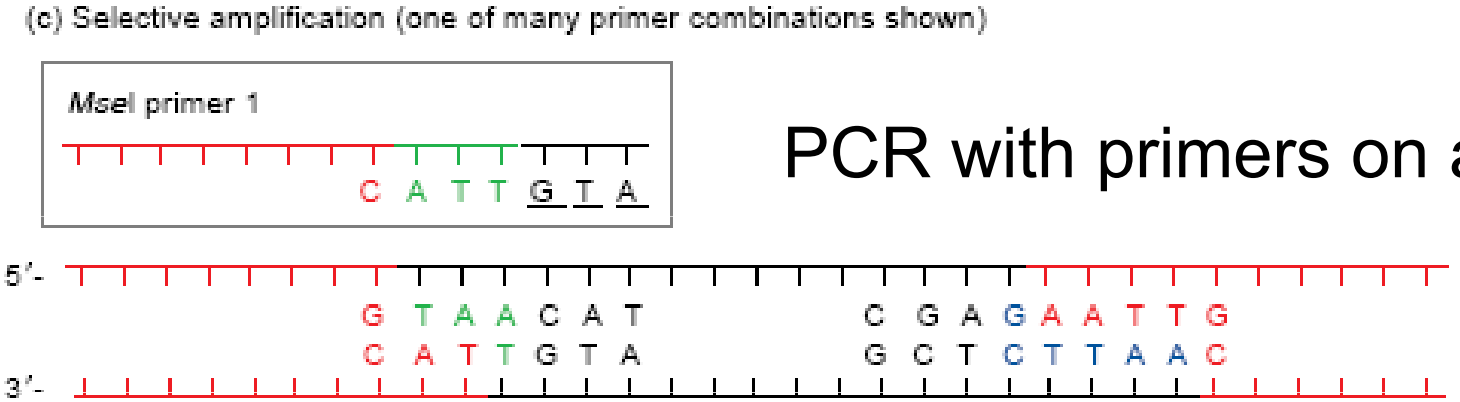
AFLP (amplified fragments length polymorphism)

- cheap, easy, fast and reliable method to generate hunderds of informative genetic markers
- simultaneous screening of many different DNA regions distributed randomly throughout the genome
- more reproducible banding pattern than RAPD

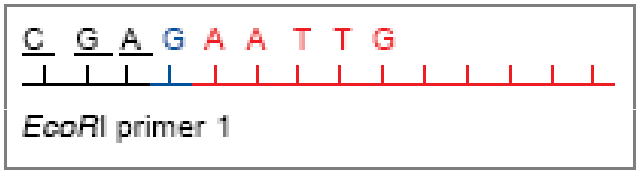
Generating AFLP markers

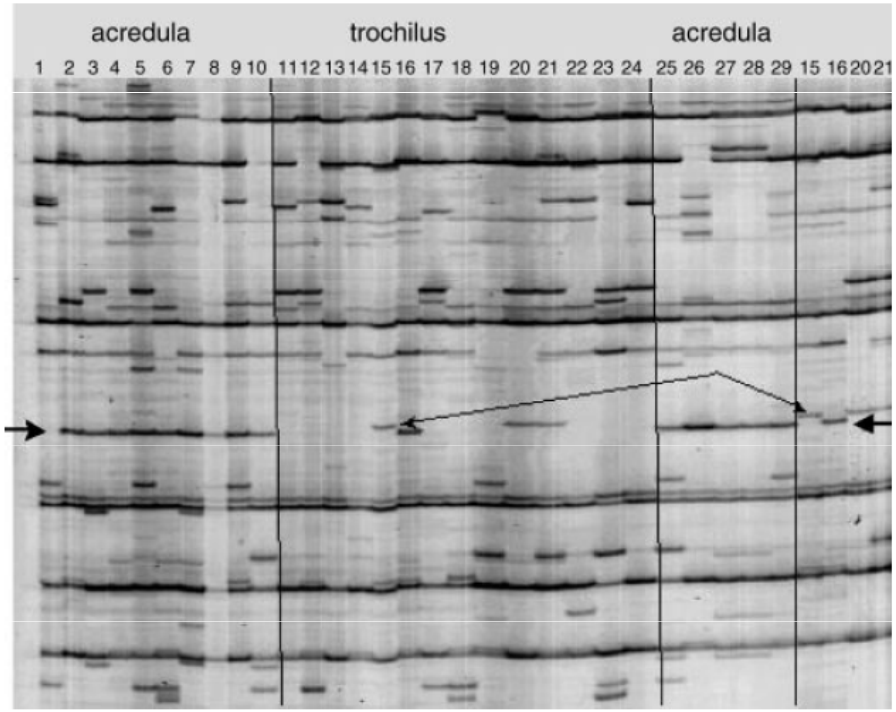


Generating AFLP markers

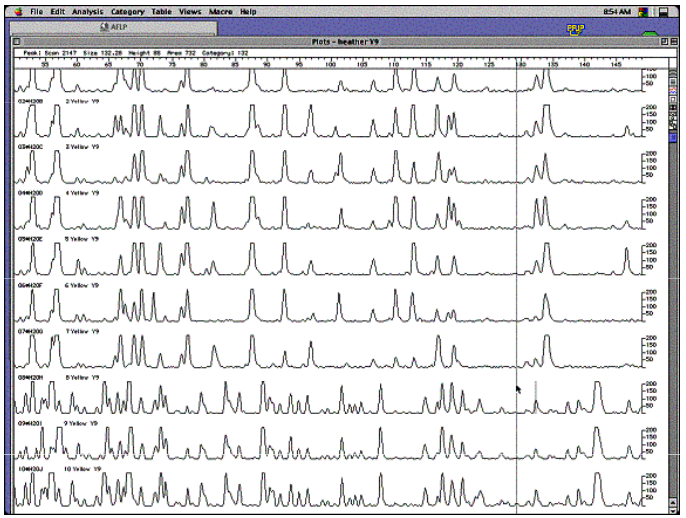


PCR with primers on adaptors





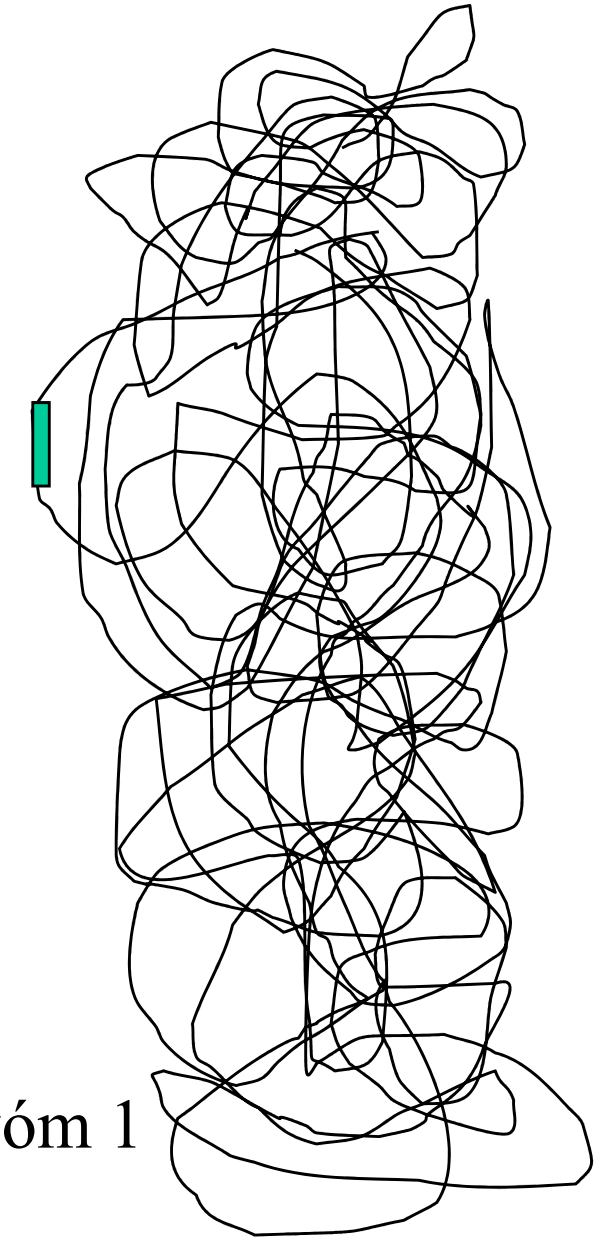
multi-locus genotype



„capillary version“

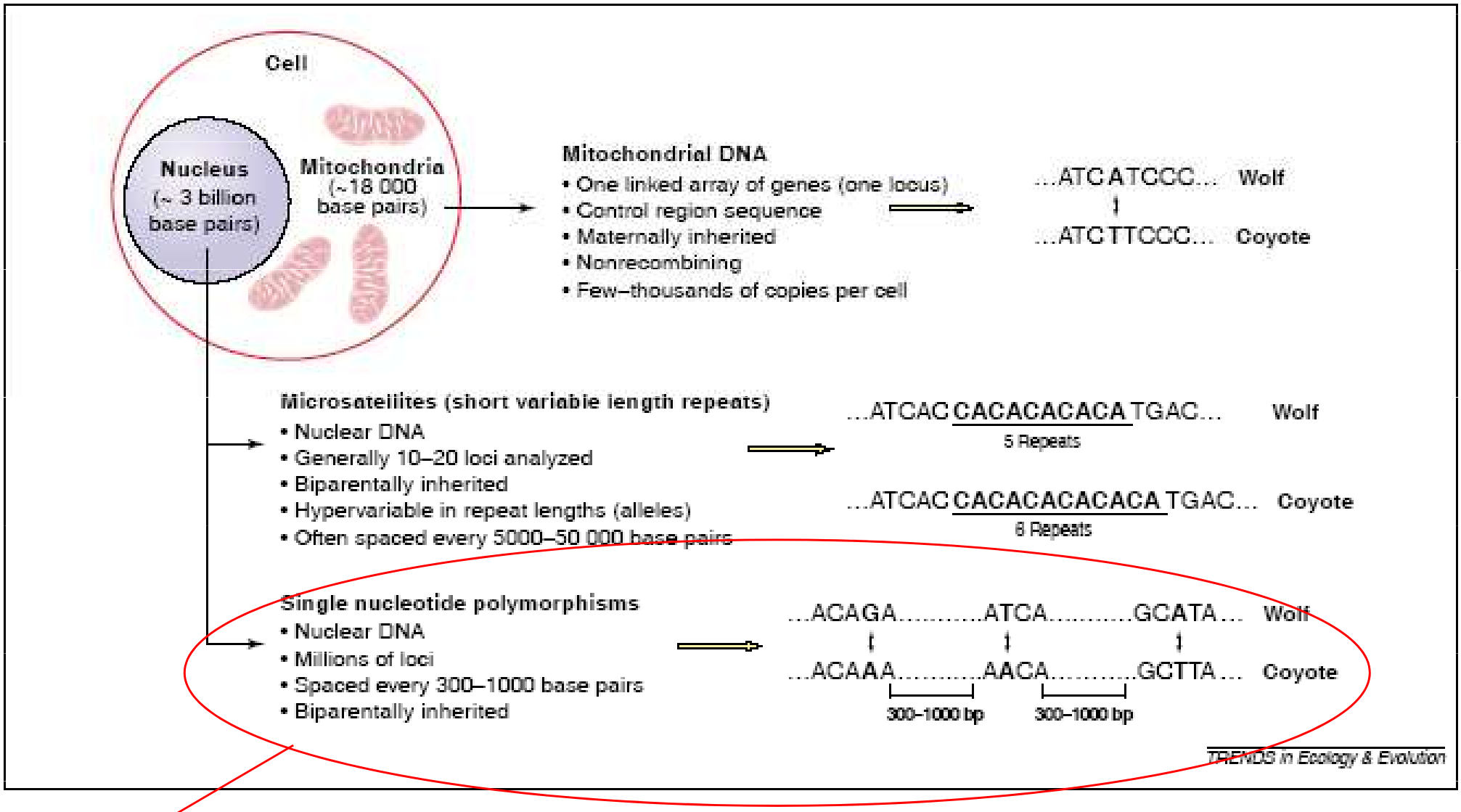
Single-locus genetic markers

- allozymes and other transcribed genes
- SNPs (single nucleotide polymorphisms)
- microsatelittes (length polymorphism)



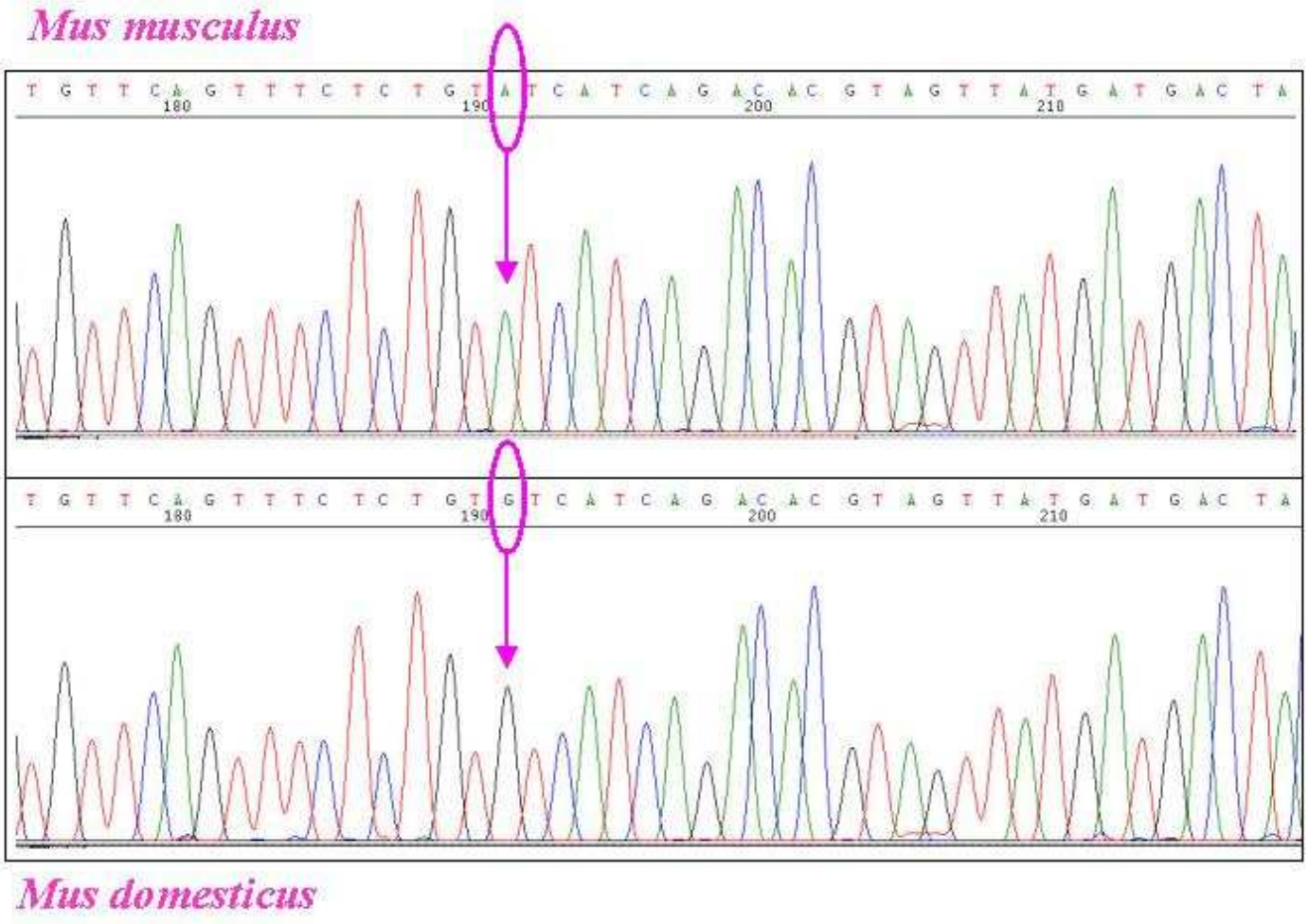
Př.: chromozóm 1

Single nucleotide polymorphisms (SNPs)



SNPs : nuclear genome (consensus)

Example of SNP marker



transice
A ↔ G

transition: Pu→Pu or Py→Py

transversion: Pu→Py or Py→Pu

Use of SNPs markers

- species (or genetical group) identification and analysis of hybridization
- phylogeography
- population genetics (genetic variation, individual identification – parentage, relatedness, population structure, population size, changes in population size)

Advantages

- abundant and widespread in many genomes (in both coding and non-coding regions) – milions of loci
- spaced every 300-1000 bp
- biparentaly inherited (vs. mtDNA)
- evolution is well described by simple mutation models (vs. microsatellites)
- shorter fragments are needed – using in non-invasive methods

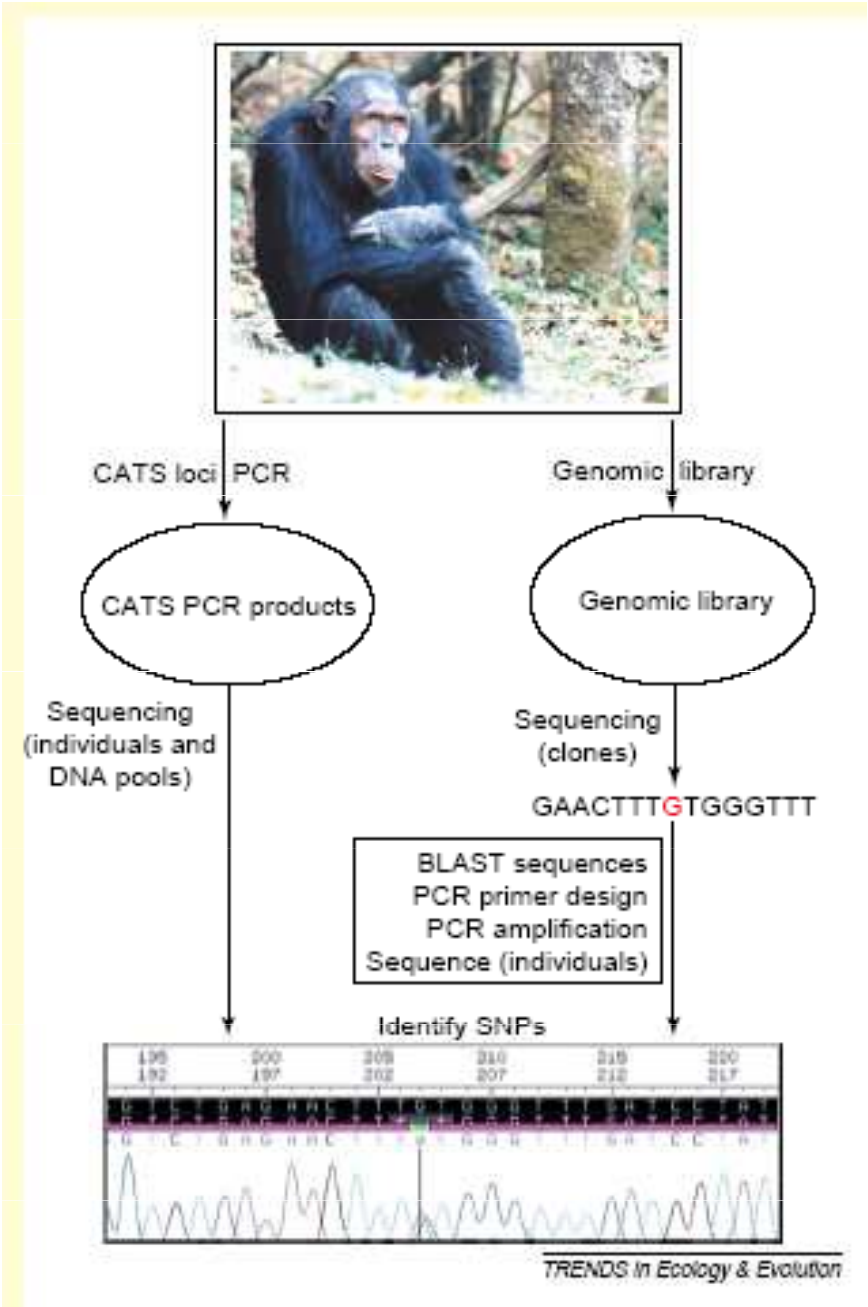
Disadvantages

- ascertainment bias – selection of loci from an unrepresentative sample of individuals
- low variability per locus (usually bi-allelic)
- higher number of loci is needed in population genetic applications (4-10 times more loci)

Methods

1. Locus discovery (ascertainment)
2. Genotyping

SNPs discovery

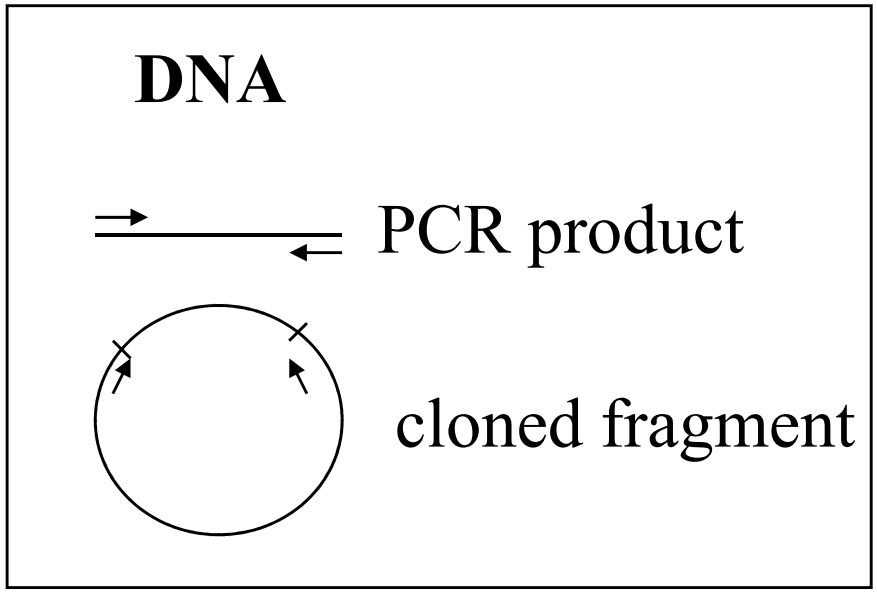


CATS loci = comparative anchor tagged site loci (= cross amplification)

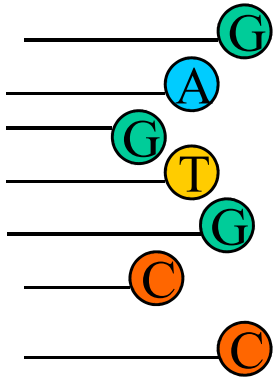
Genomic library = genome restriction + cloning

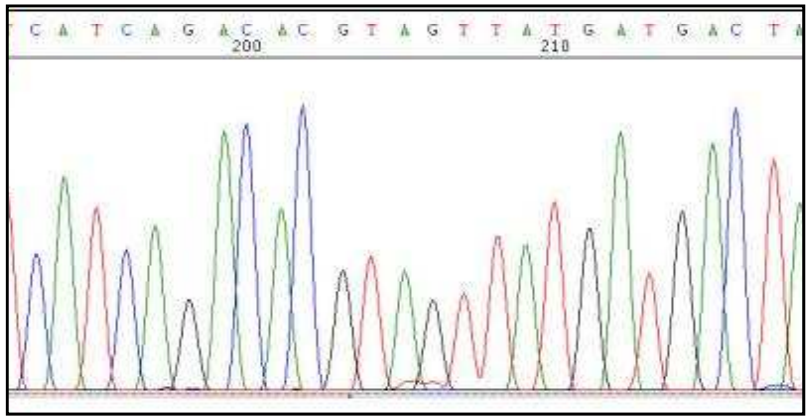
AFLP = alternative to the genomic library construction (provide PCR fragments, can be transformed to informative SNP)

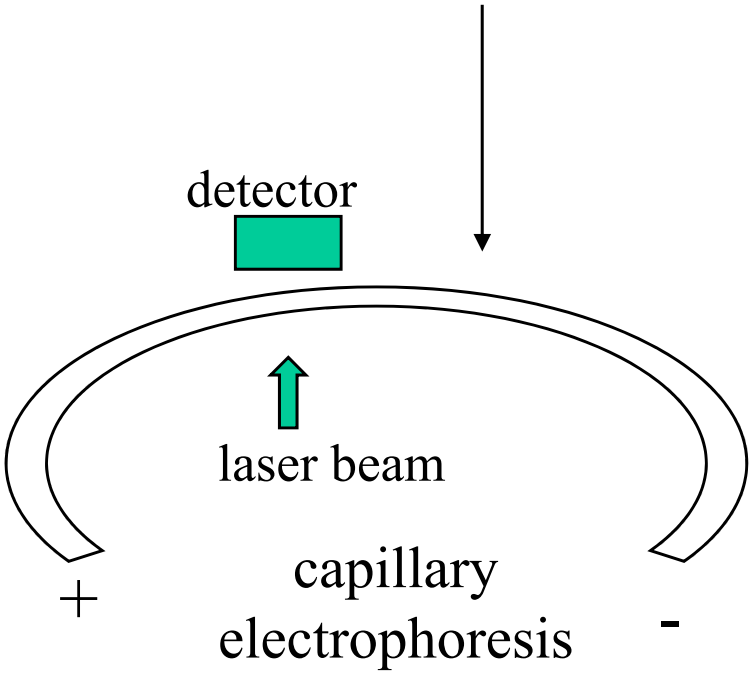
Sequencing

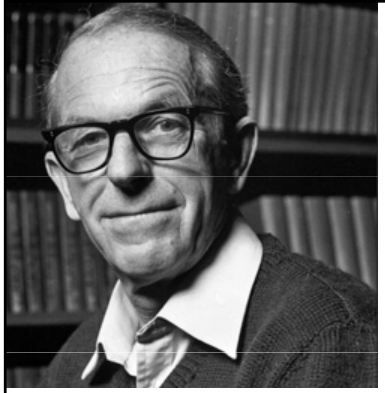


sequencing reaction with marked dideoxynucleotides and specific or universal primers



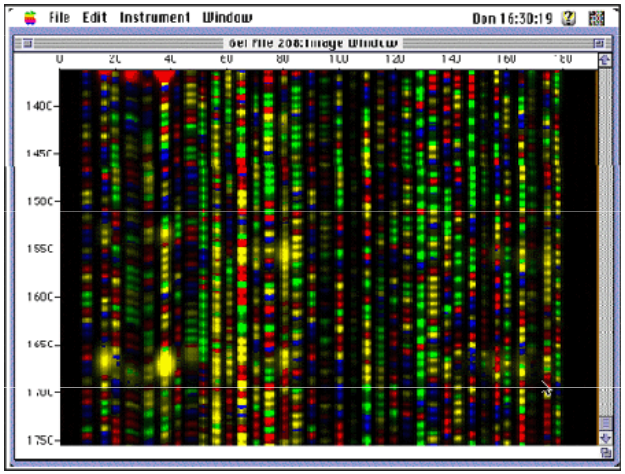


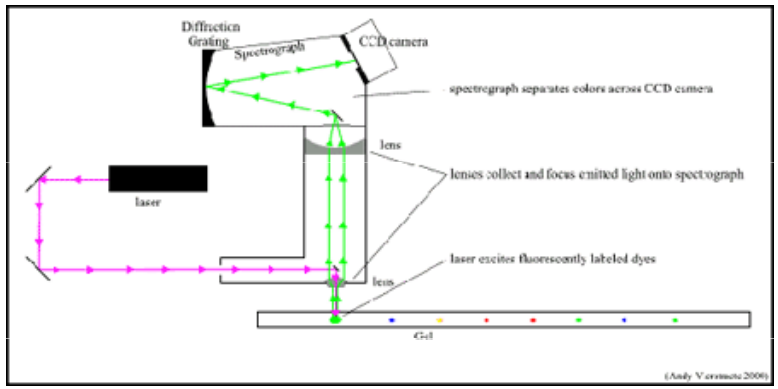


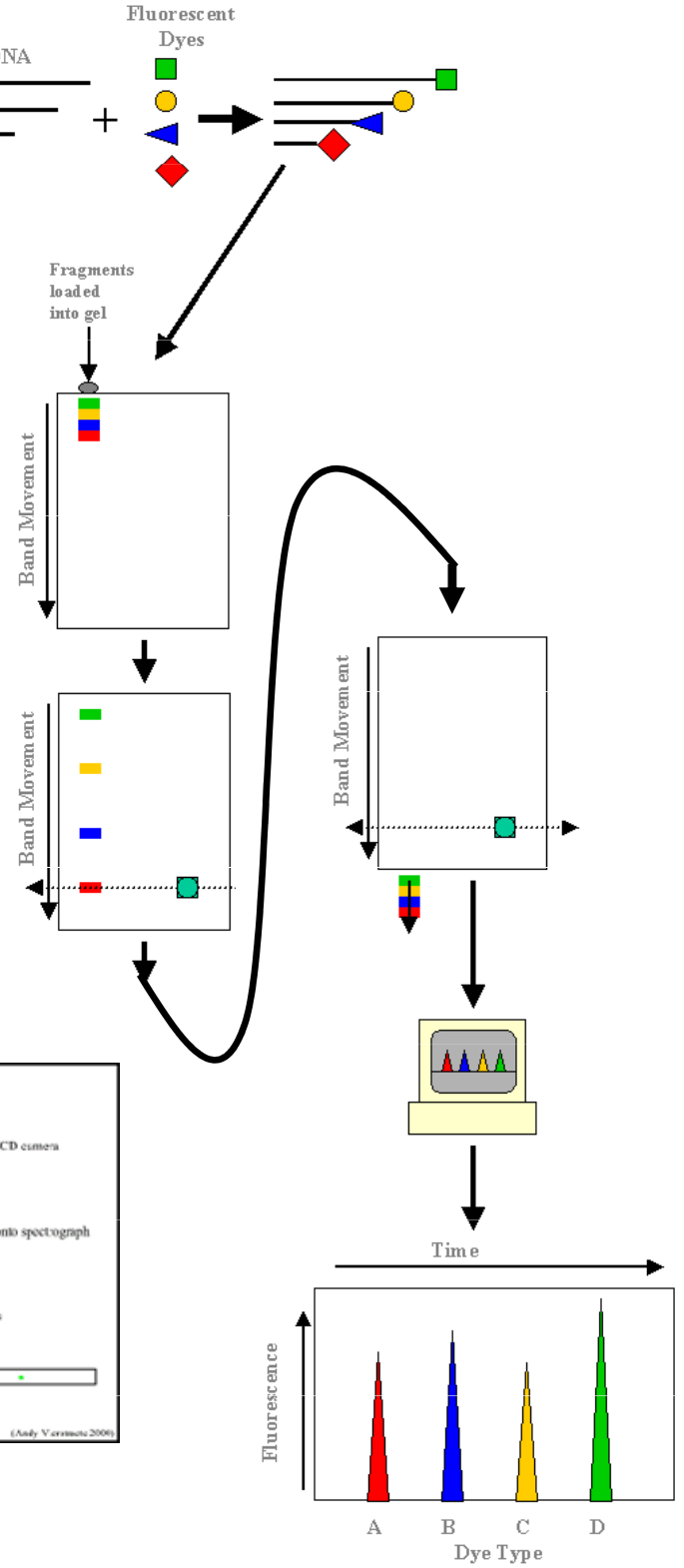


Sangrova dideoxy metoda

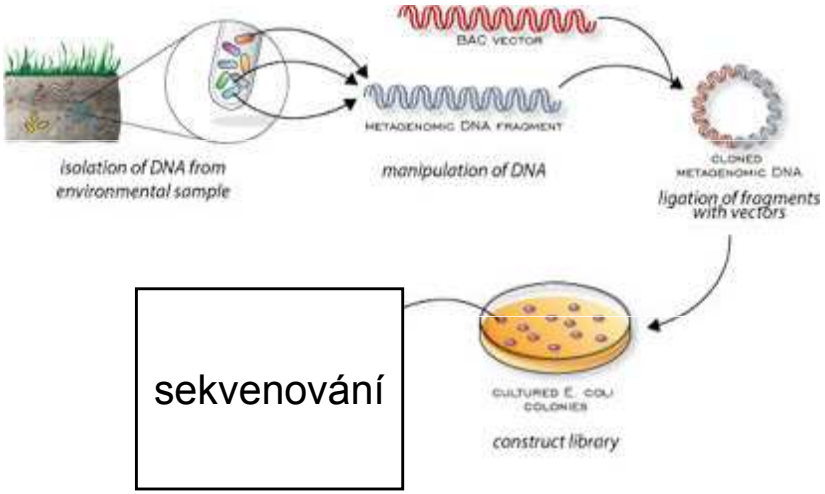
- Sekvence délky 500 – 1000 bp
- 4 kapiláry - destička s 96 vzorky za noc
- Jsou i sekvenátory s 96 kapilárami



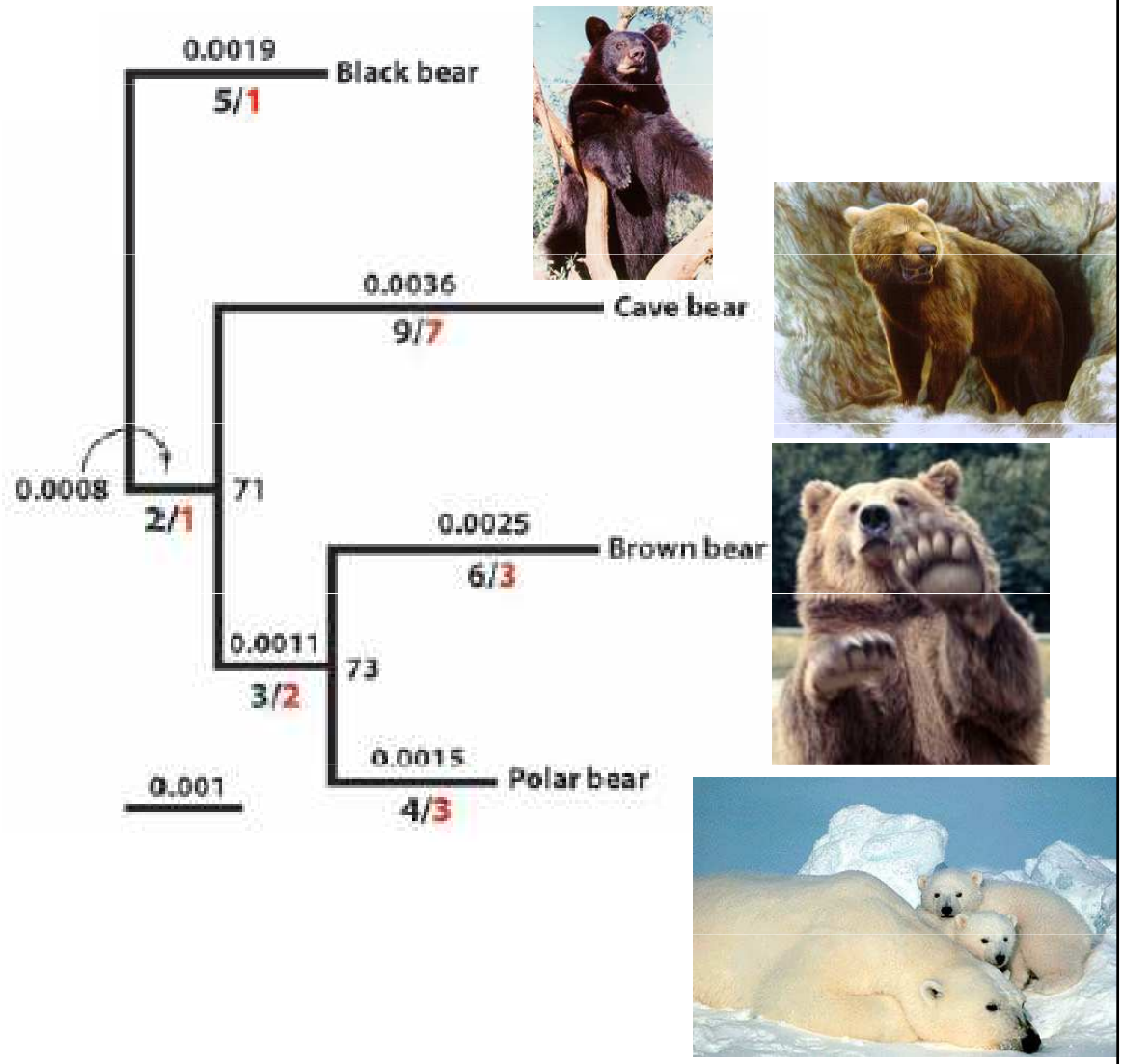




Použití nových přístupů



sekvenování



Metagenomické knihovny

Ursus spelaeus

> 28 000 bp

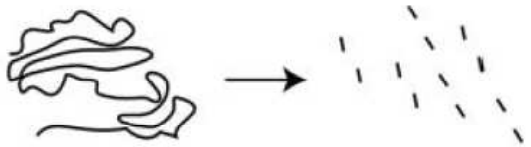
(jaderná i mitochondriální DNA)

(*Noonan et al. 2005*)

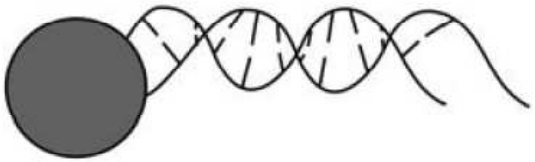
„Next generation“ sequencing

(Hudson 2008)

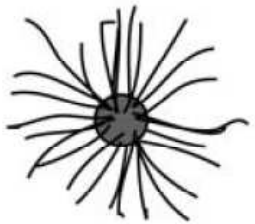
1) Randomly fragment many molecules of target DNA



2) Immoblize individual DNA molecules on solid support

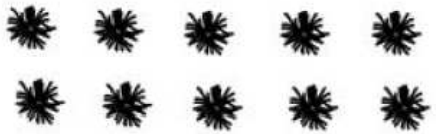


3) Amplify DNA in clonal ‘polymerase colony’

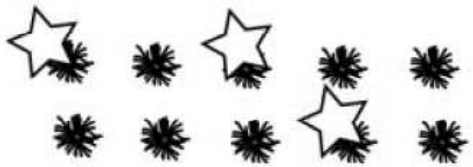


„polonies“
(polymerase colonies)

4) Sequence DNA by adding liquid reagents to immoblized DNA colonies



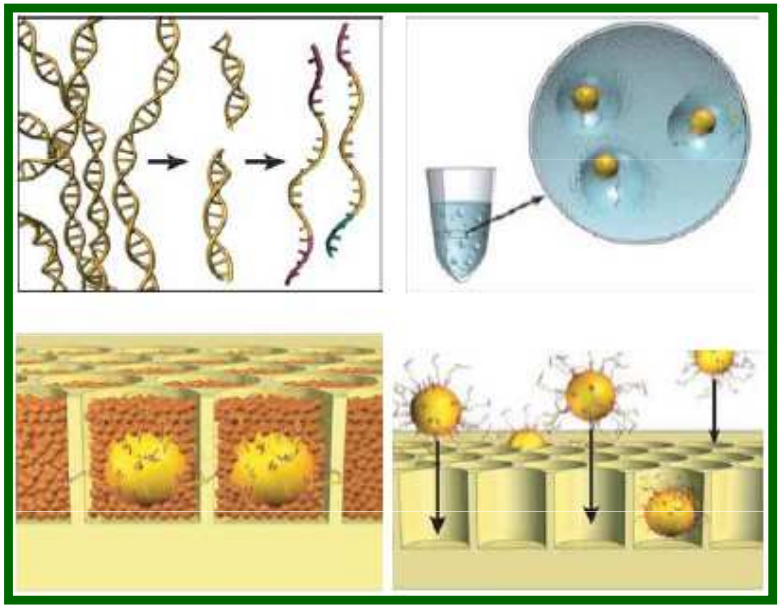
5) Interrogate sequence incorporation *in situ* after each cycle using fluorescence scanning or chemiluminescence

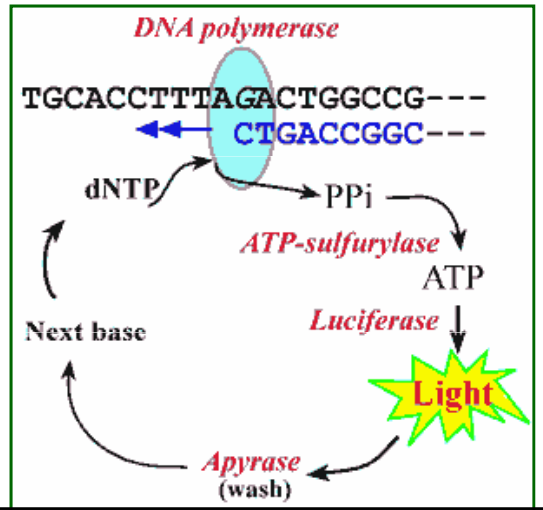


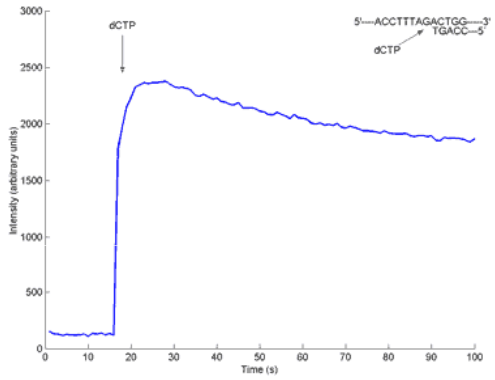
454 pyrosequencing

- emulzní techniky amplifikace
  pikolitrové objemy
- simultánní sekvenování
  na destičce z optických vláken
  detekce pyrofosfátů uvolňovaných při inkorporaci bazí
- První generace GS20
  → 200 000 reakcí najednou
  (zhruba 20 milionů bp)
  dnes FLX → 400 000 reakcí najednou
- Problémy s homopolymery
- Délka jednotlivých sekvencí 100 – 400

1 600 000 well plate

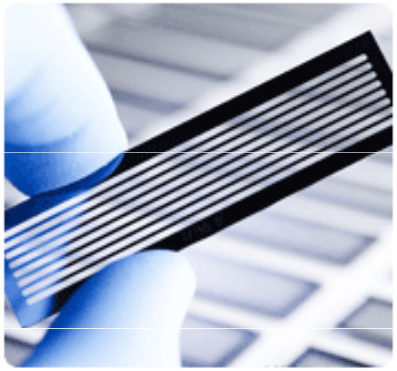






Solexa/Illumina 1G SBS technology

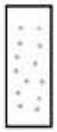
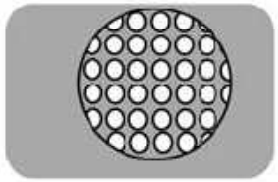
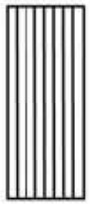
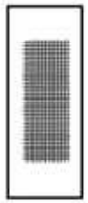
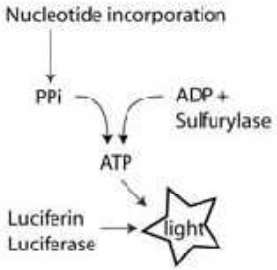
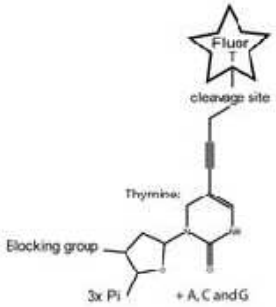
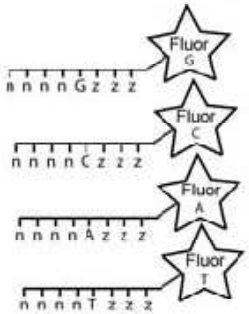
(SBS = sequencing by synthesis)



- 1 Gb (šestinásobek genomu *Drosophily*)
- Výrazně levnější
- Sekvence délky 35 bp
- Flourescence, reversibilní terminátory
- Spíš pro resequencing

SOLiD

(sequencing by Oligonucleotide Ligation and Detection)

	454 pyrosequencing	Solexa SBS sequencing	Agencourt / ABI SOLiD polony sequencing
	All methods ligate single, randomly sheared DNA molecules to support		
DNA support	 25–36 μm bead	surface of flow cell 	 ~1 μm bead
Amplification	 emulsion-phase PCR	 *in situ* PCR on solid surface	 emulsion-phase PCR
Sequencing surface	 1 600 000 well plate one bead per well	 8-channel flow cell clusters of DNA randomly located	 Single slide imaged in panels beads random
Sequencing chemistry	 pyrosequencing	 reversible-terminator sequencing by synthesis	 Ligation of sequence-specific labeled oligos
Sequence detection	Chemiluminescence (one channel)	Fluorescence (four channel)	Fluorescence (four channel)
Read length and number	100–400 bp $> 2 \times 10^5$ reads	35 bp $\sim 4 \times 10^7$ reads	25 bp (paired) $> 10^7$ reads

SNP genotyping - old standards

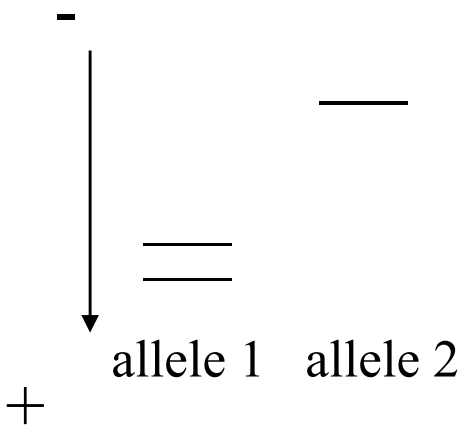
PCR-RFLP
(restriction fragments length polymorphism)

Allele 1 CCGATCAATGCGGCAA
GGCTAGTTACGCCGTT

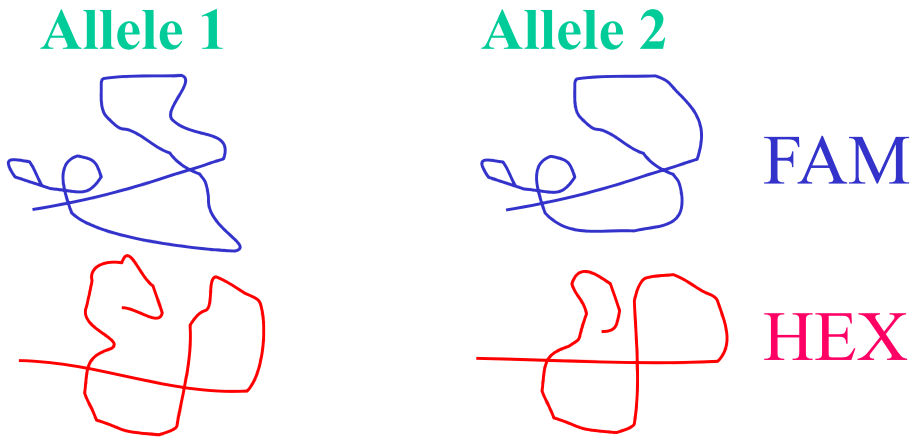
cutting by restriction endonuclease

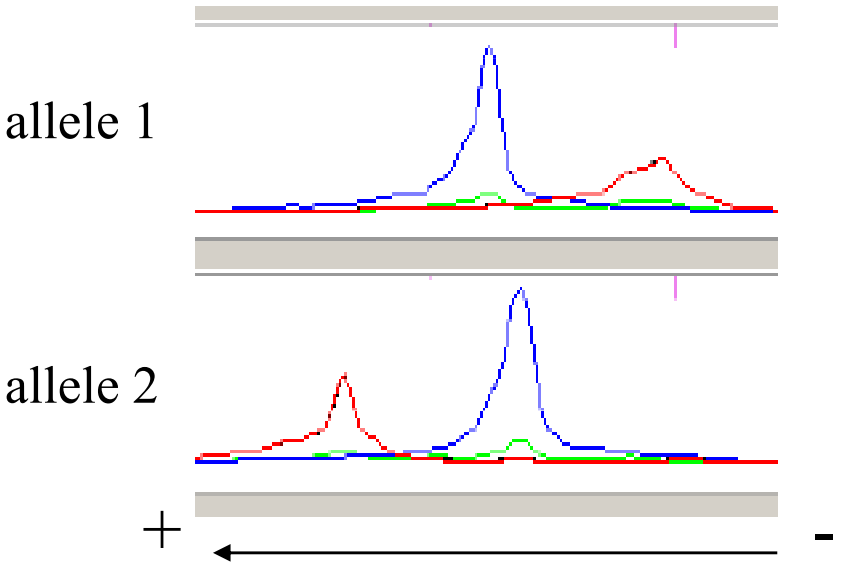
Allele 2 CCGATCACTGCGGCAA
GGCTAGTGACGCCGTT

no cut



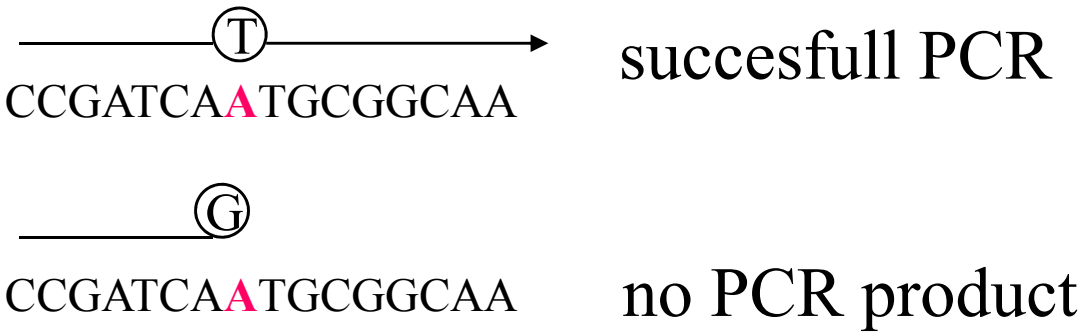
PCR-SSCP (gel or capillary)
(single strand conformation polymorphism)





SNP genotyping – new methods

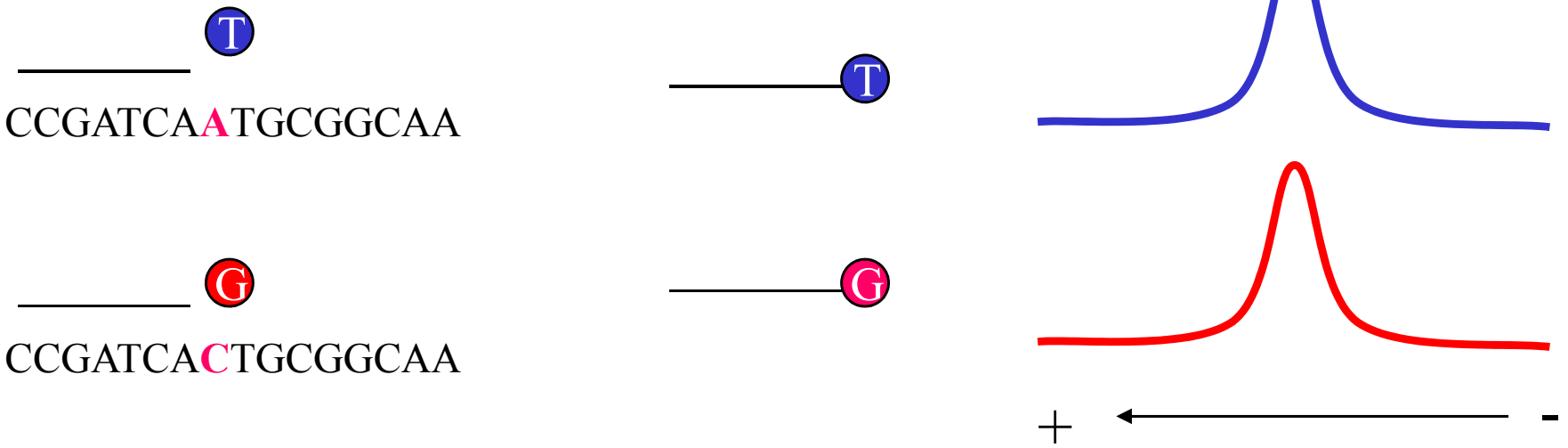
1) ASPE: allele-specific primer extension



- primer extension with deoxy nucleotides and highly specific polymerase enables allele-specific amplification
- 3' terminal nucleotide of the two primers contains the SNP nucleotide
- two PCRs with specific primers are necessary

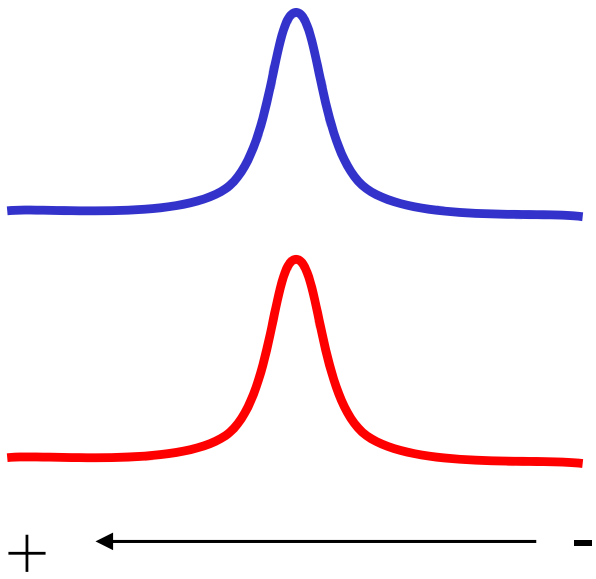
SNP genotyping – new methods

2) SBE: single base extension

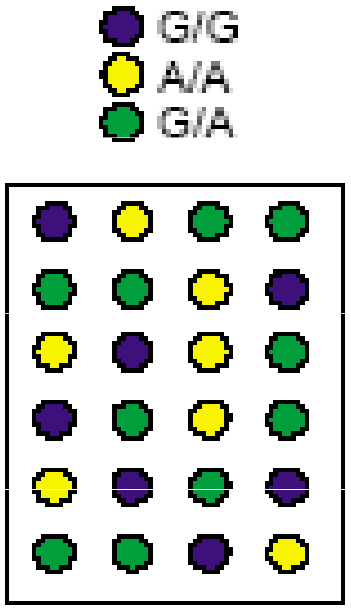


only one dideoxy nucleotide is added to the primer

Detection or SBE products



electrophoresis in a capillary



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

other methods: flow cytometry, fluorescence polarization