

Typy genetických markerů

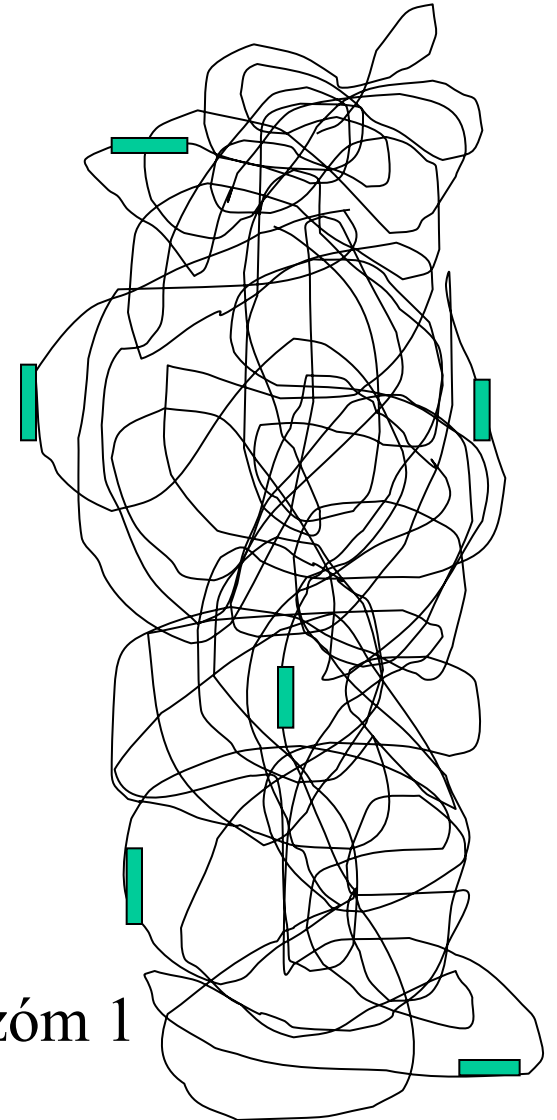
- **multilocus** markery – na více místech v genomu
- **single-locus** markery – jen na jednom místě genomu
- **dominantní** markery – present or absent
- **kodominantní** markery – identifikace homologních alel, tj. je možno rozlišit homozygotní a heterozygotní stav (umožňují stanovit frekvenci alel)

Typy genetických markerů

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

Multi-locus genetic markers

- Mnoho znaků náhodně rozmístěných v genomu – **celogenomový scan**
 - minisatellite DNA fingerprinting
 - RAPD (randomly amplified polymorphic DNA)
 - AFLP (amplified fragment length polymorphism)
- presence vs. absence = dominantní znaky (neodliší heterozygota)

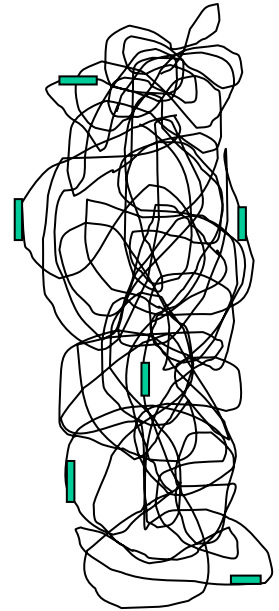


Př.: chromozóm 1

(Minisatellite) DNA fingerprinting

(Jeffreys et al. 1985)

- náhodně rozmístěné repetice (např. **Alu** sekvence, SINE, LINE)
- různě dlouhé minisatelity (různá délka repetice, 10-15 bp „core sequences“)
- restriční štěpení kompletní DNA – sekvenčně specifické **restriční endonukleázy**



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)

Restriction site

Palindrome

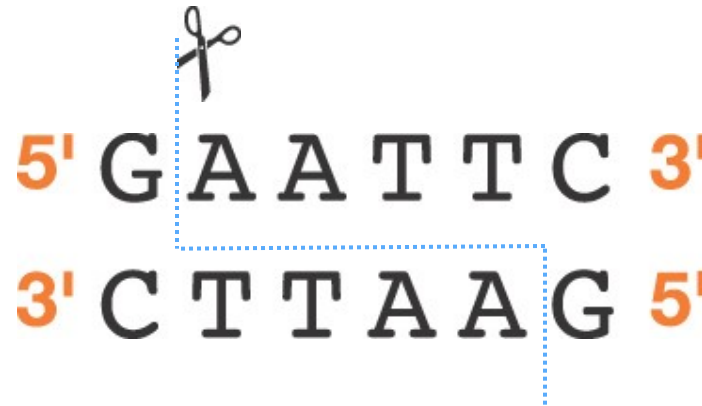
G T A G G A A T T C A T T T C A C G C A
C A T C T T A A G T A A G T G C G T

G T A G A A T T C A T T T C A C G C A
C A T C T T A A G T A A G T G C G T

Fragment 1

Fragment 2

Common Restriction Enzymes



EcoRI

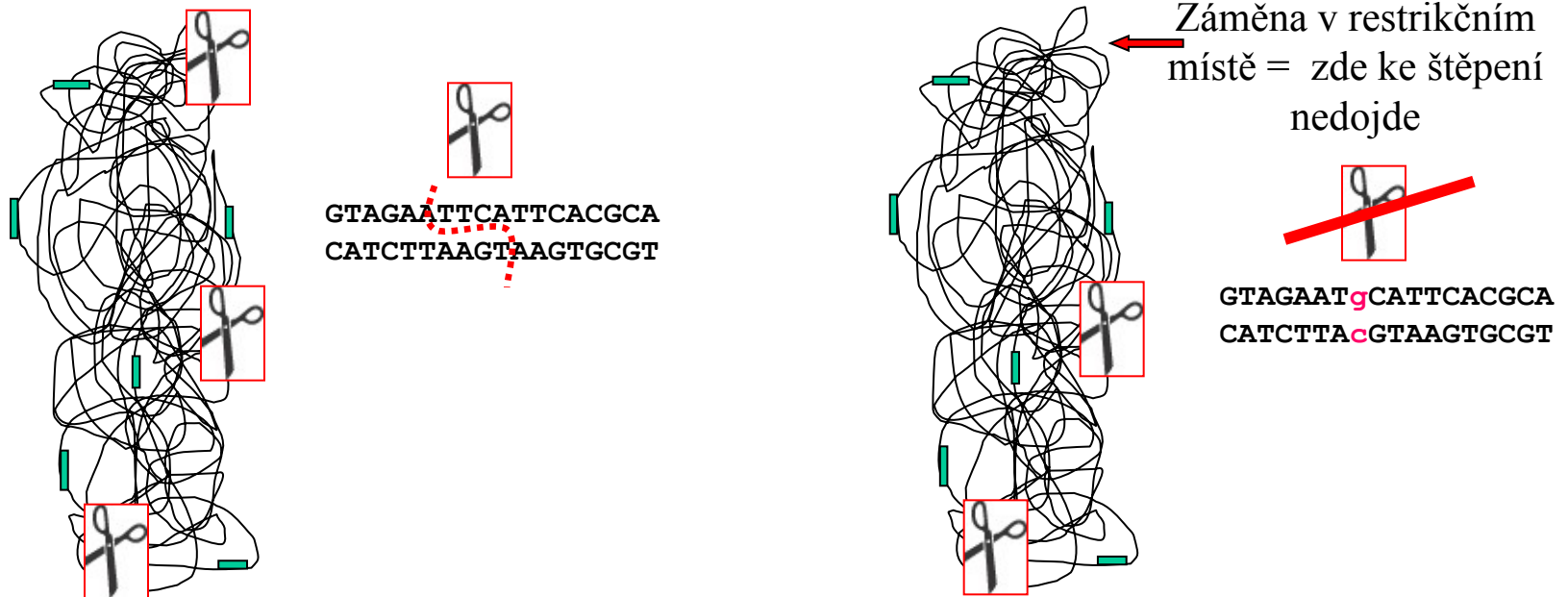
- Escherichia coli
- 5 prime overhang



PstI

- Providencia stuartii
- 3 prime overhang

Každý jedinec má jedinečný genom



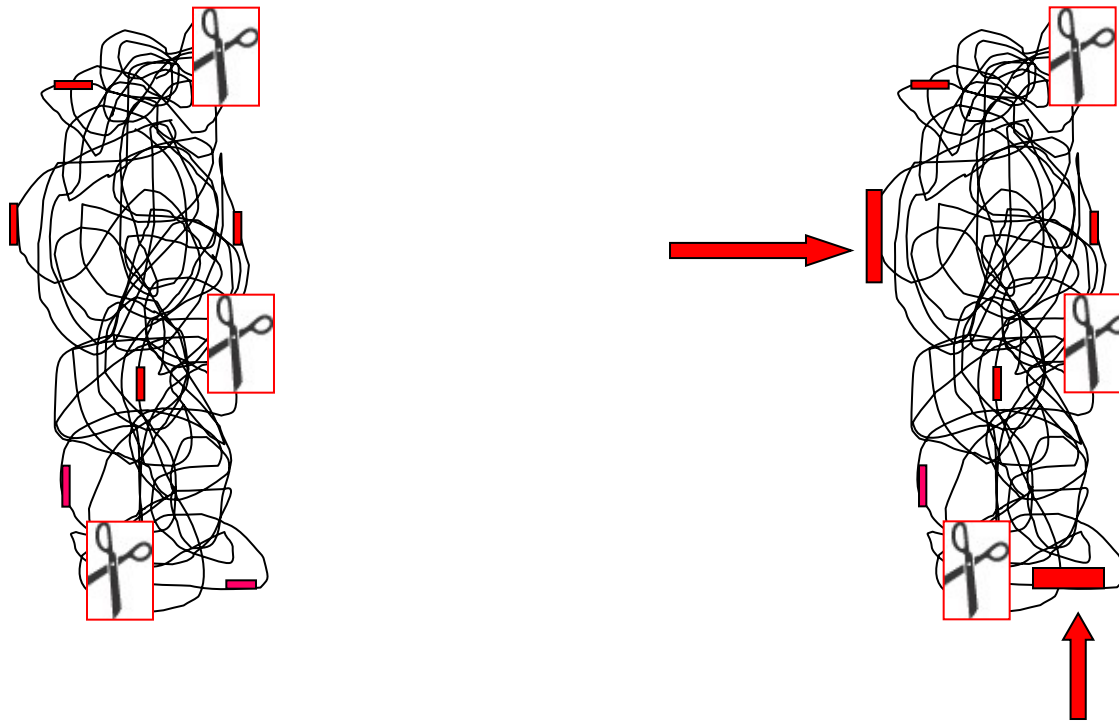
1. Ztráta nebo nabytí restričního místa

Každý jedinec má jedinečný genom



2. Ztráta nebo nabytí SINE (např. **Alu** sekvence) nebo LINE

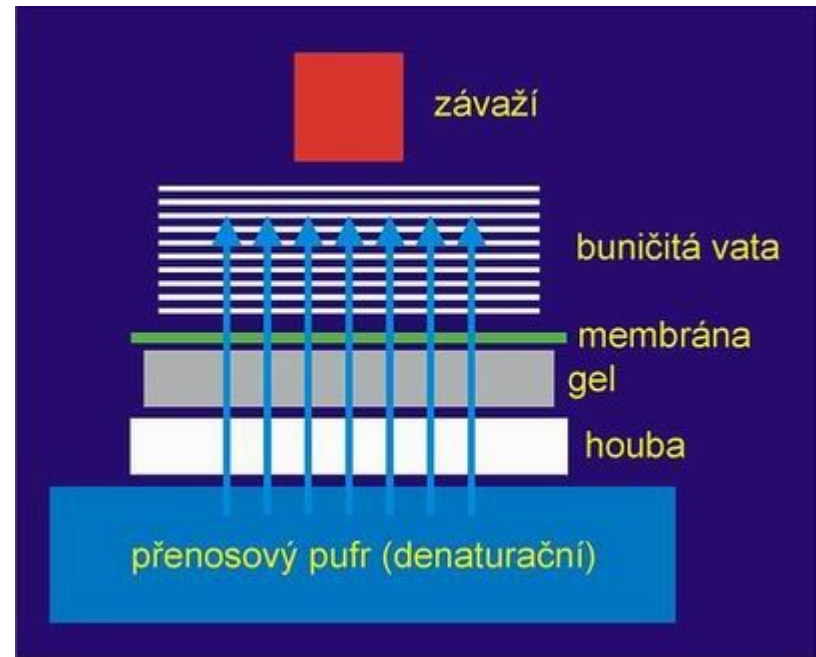
Každý jedinec má jedinečný genom



3. Vysoká mutační rychlost **minisatelitů** – rozdíly v počtu repeticí, tj. v délce daného úseku

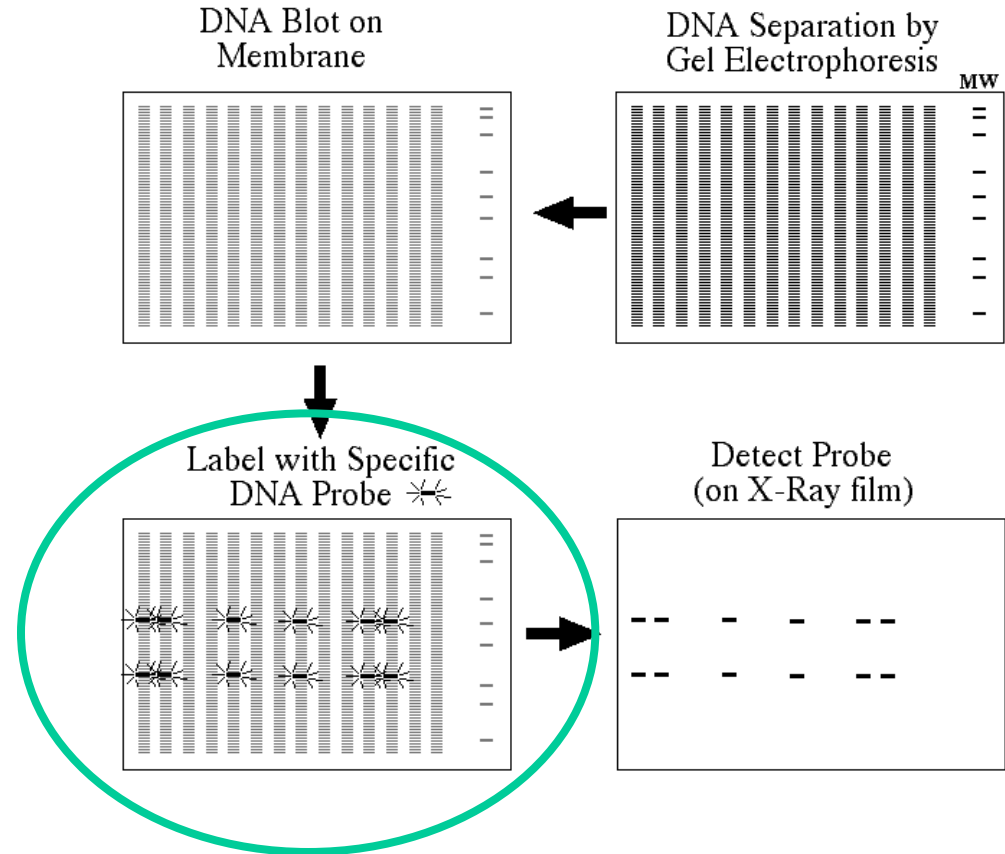
Minisatellite DNA fingerprinting

- elektroforéza rozštěpené DNA
- Southern blotting – přenesení DNA na membránu



Minisatellite DNA fingerprinting

- elektroforéza
- Southern blotting – přenesení DNA na membránu
- hybridizace se značenou sondou (nejčastěji radioaktivní značení), tj. specifickou sekvencí odpovídající danému minisatelitu



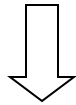
Minisatellite DNA fingerprinting

- elektroforéza
- Southern blotting – přenesení DNA na membránu
- hybridizace se značenou sondou, tj. specifickou sekvencí odpovídající danému minisatelitu
- zásadní objevy např. EPC u ptáků
- v posledních 10-15 letech – přesun k PCR-based metodám



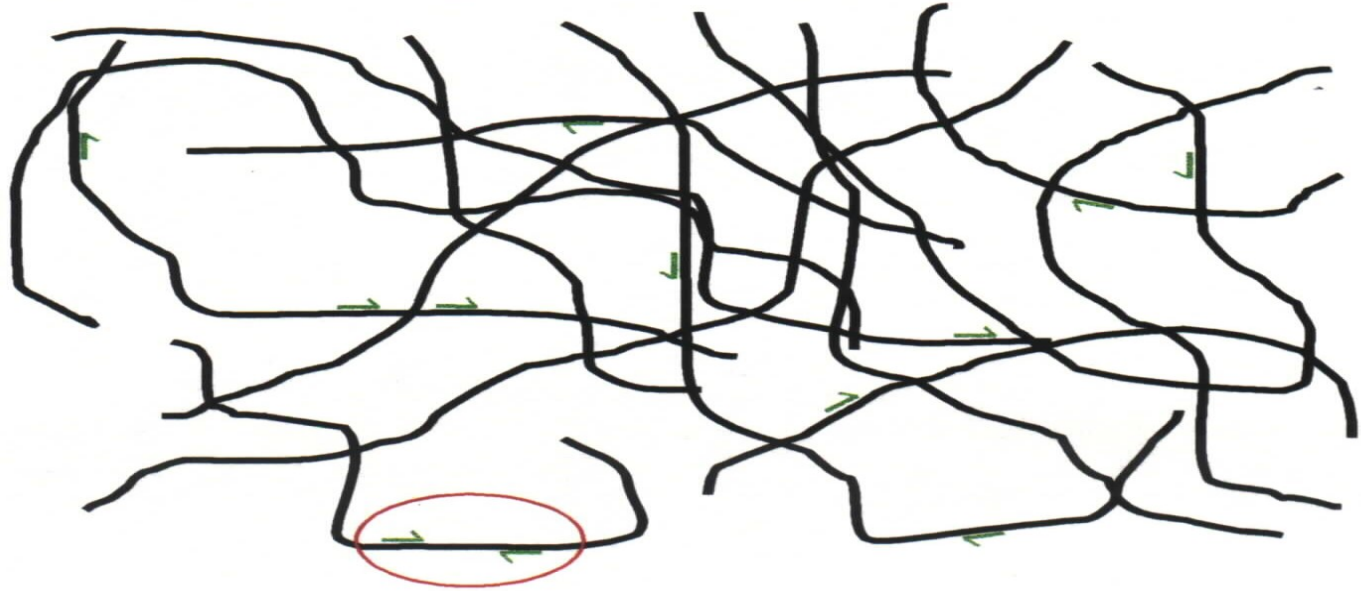
RAPD (randomly amplified polymorphic DNA)

Krátké náhodné oligonukleotidy
(~ 10 bp) jako primery

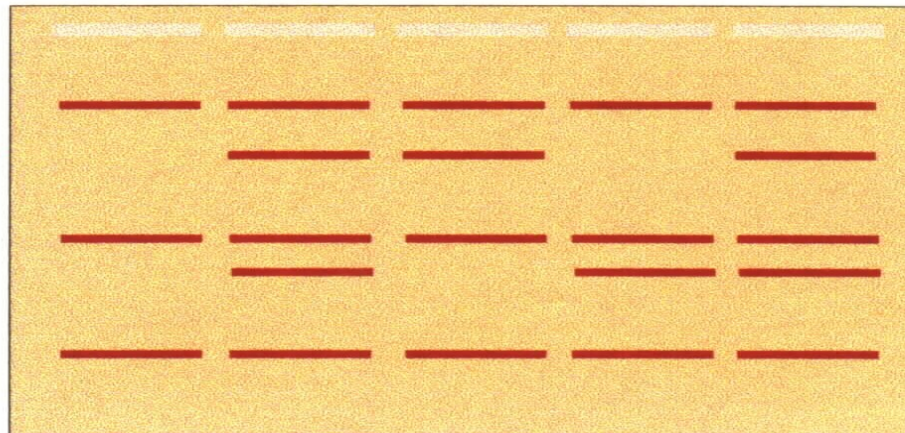


PCR za málo specifických podmínek

genomic DNA



- 1) PCR
- 2) Separation by size on agarose gel

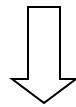


Variabilní DNA detekovaná metodou RAPD je důsledkem ztráty RAPD lokusů v důsledku:

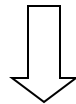
- a) Změna sekvence v místě nasedání primeru
- b) Delece místa nasedání primeru
- c) Velká inzerce mezi dvěma místy nasedání primeru

RAPD - review

Krátké náhodné oligonukleotidy
(~ 10 bp) jako primery

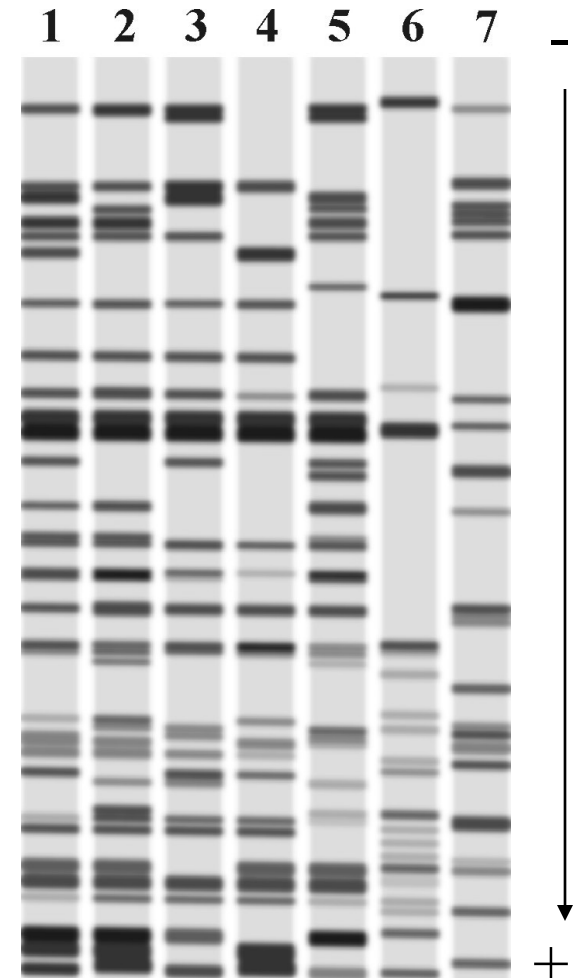


PCR za málo specifických podmínek



Detekce PCR produktů elektroforézou

Nízká opakovatelnost v důsledku
mnoha faktorů ovlivňujících PCR –
dnes již není akceptována jako
metoda např. pro studium populační
struktury

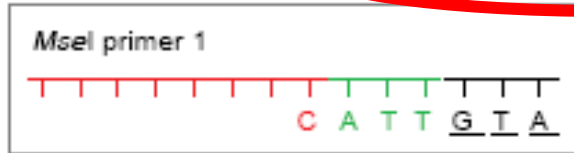


AFLP (amplified fragments length polymorphism)

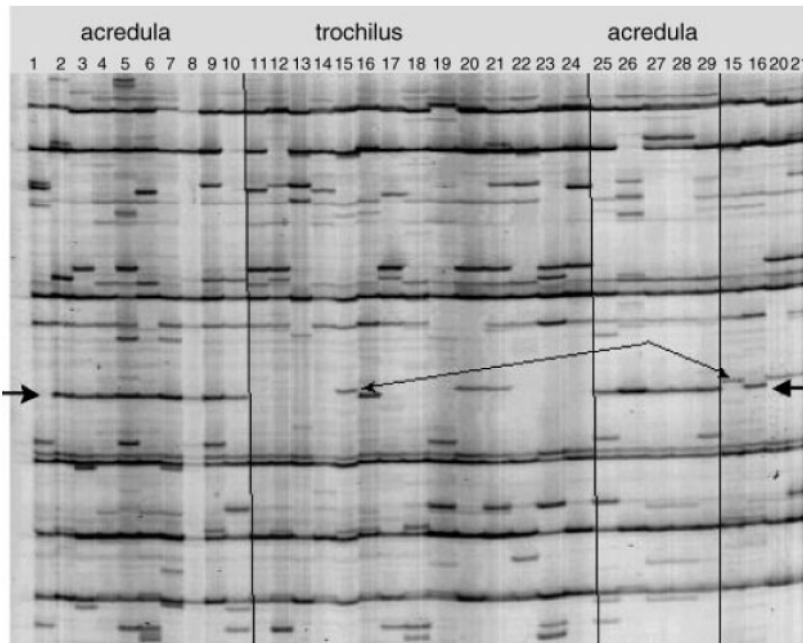
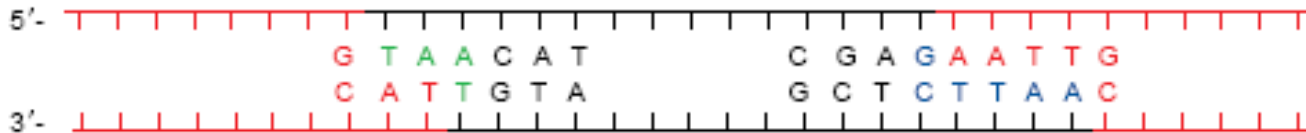
- levná, jednoduchá, rychlá a spolehlivá metoda na generování stovek informativních genetických markerů
- současný screening mnoha různých DNA oblastí distribuovaných náhodně v genomu
- lépe reprodukovatelná než RAPD
- založená na polymorfismu v restričním štěpení
- „genome scan“ – hledání asociací s fenotypovými znaky

Generating AFLP markers

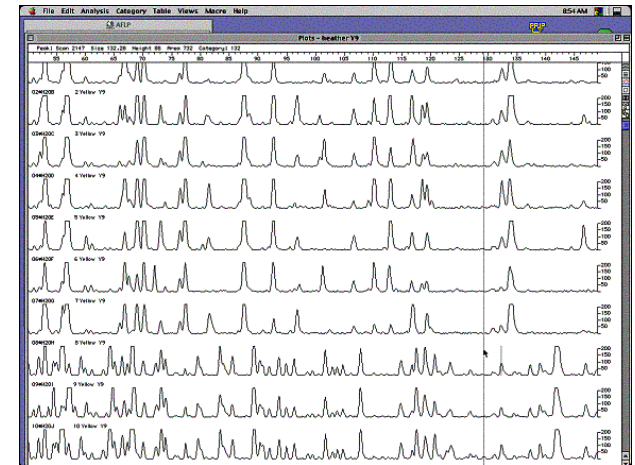
(c) Selective amplification (one of many primer combinations shown)



PCR with primers on adaptors



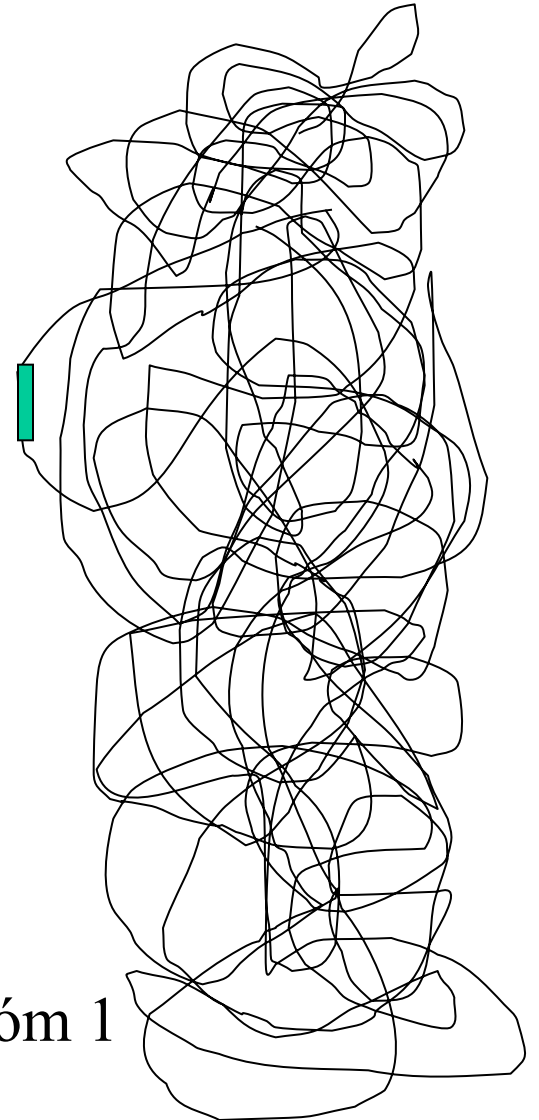
multi-locus genotype



„capillary version“

Single-locus genetic markers

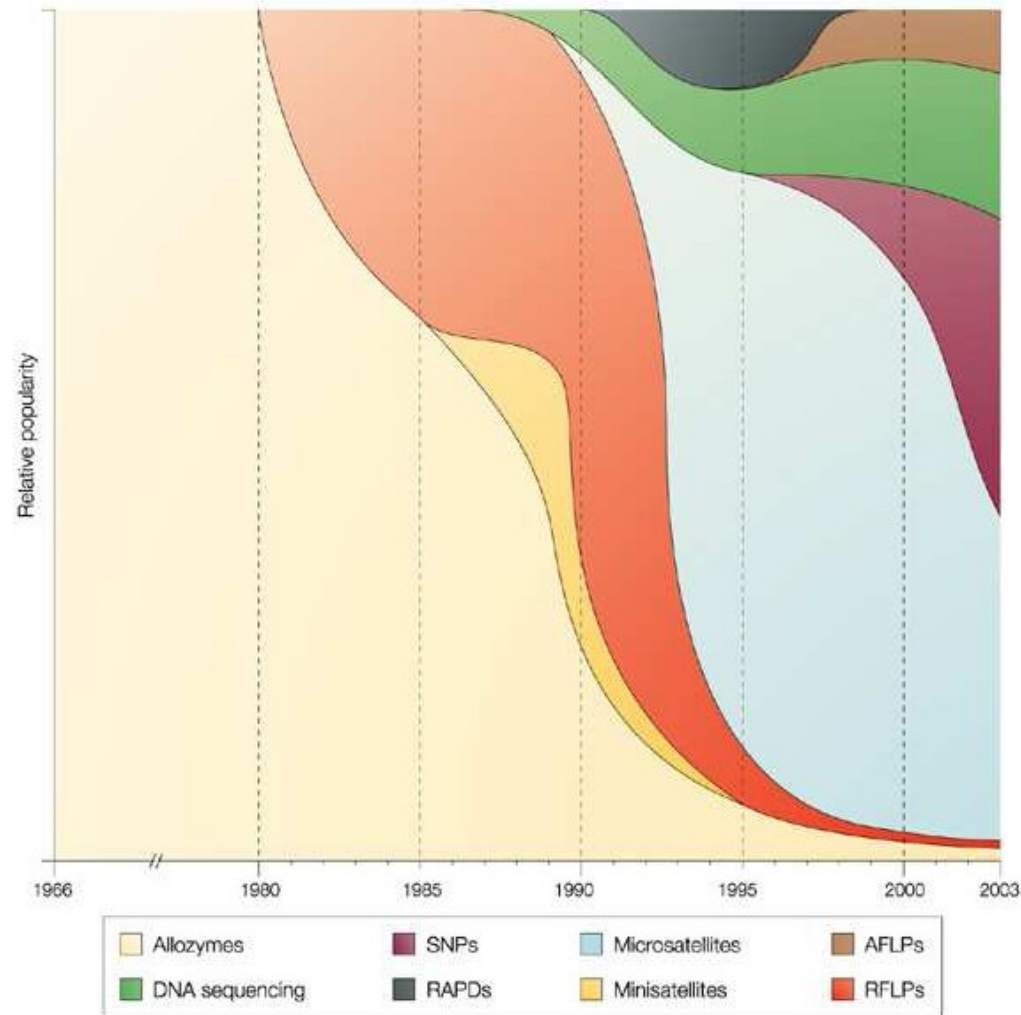
- kodominantní – možno stanovovat frekvence alel (= lze odlišit homo- a heterozygota)
- allozomy a jiné funkční geny - MM
- mikrosatelity – délkový polymorfismus
- SNPs (single nucleotide polymorphisms) – sekvenční polymorfismus
- SINE, LINE



Př.: chromozóm 1

Mikrosatelite

Mikrosatelity jsou stále nejpoužívanější markery v molekulární ekologii



Mikrosatelity

- VNTR („variable number of tandem repetitions“), SSR („simple sequence repeats“)
- jednotlivé alely se liší délkou

TTCAGGCACACACA**ATCTCTAGCTTCGA**

27 bp

TTCAGGCACACA**ATCTCTAGCTTTGA**

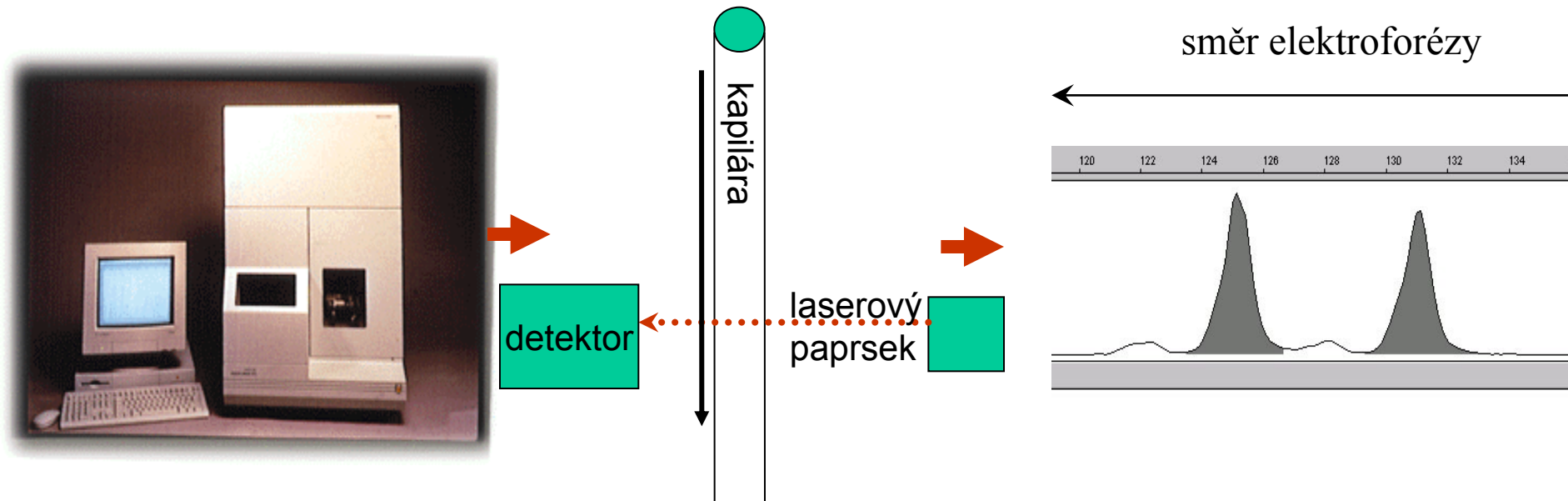
25 bp

genotyp diploidního jedince: **25/27**

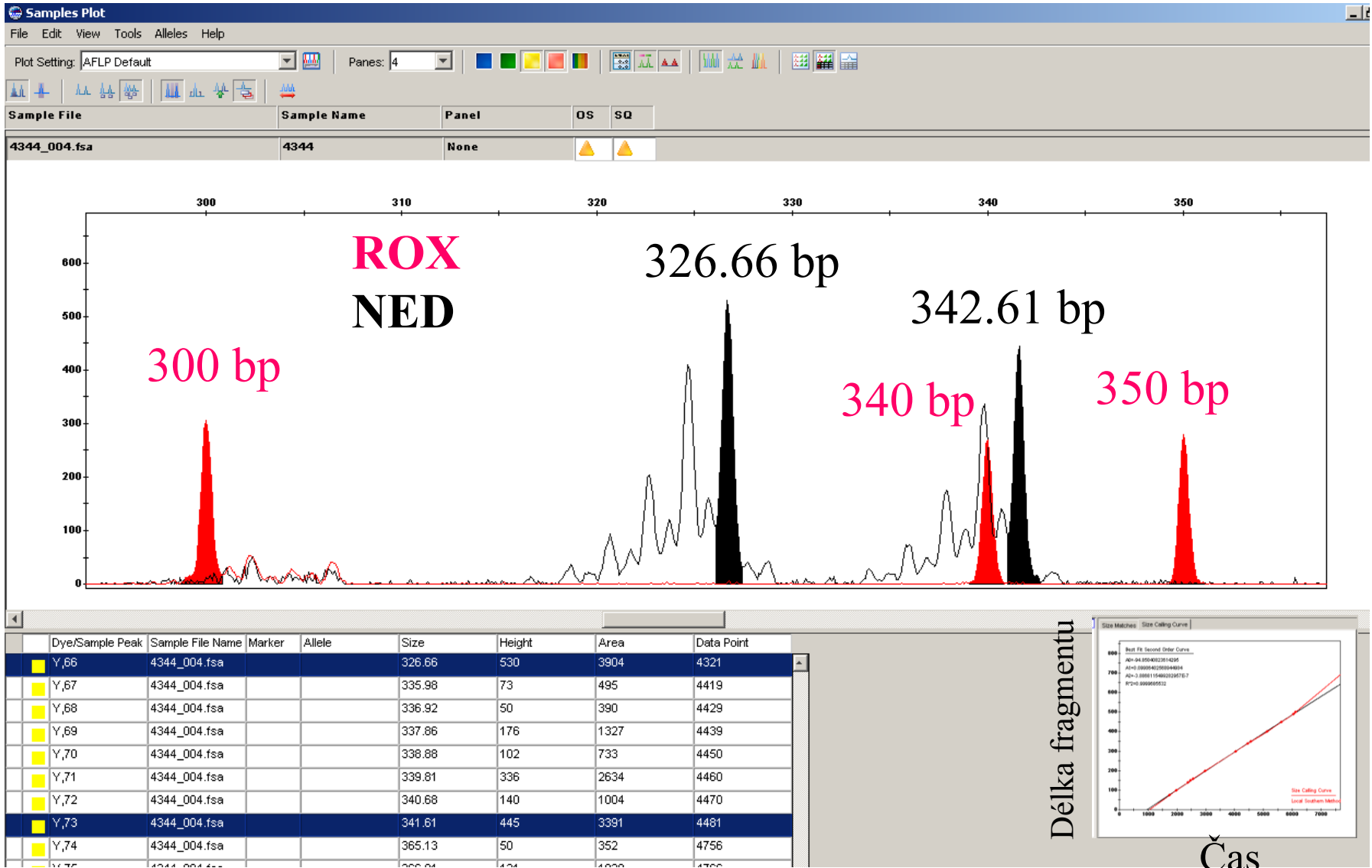
Mikrosatelity

- 1-6 (nejč. 2-4) bp motiv
- početné po celém genomu
- vysoká úroveň polymorfismu (běžně 15 alel v populaci)
- Mendelovská dědičnost (autosomy) - kodominance
- ideální pro studium populační struktury a příbuzenských vztahů

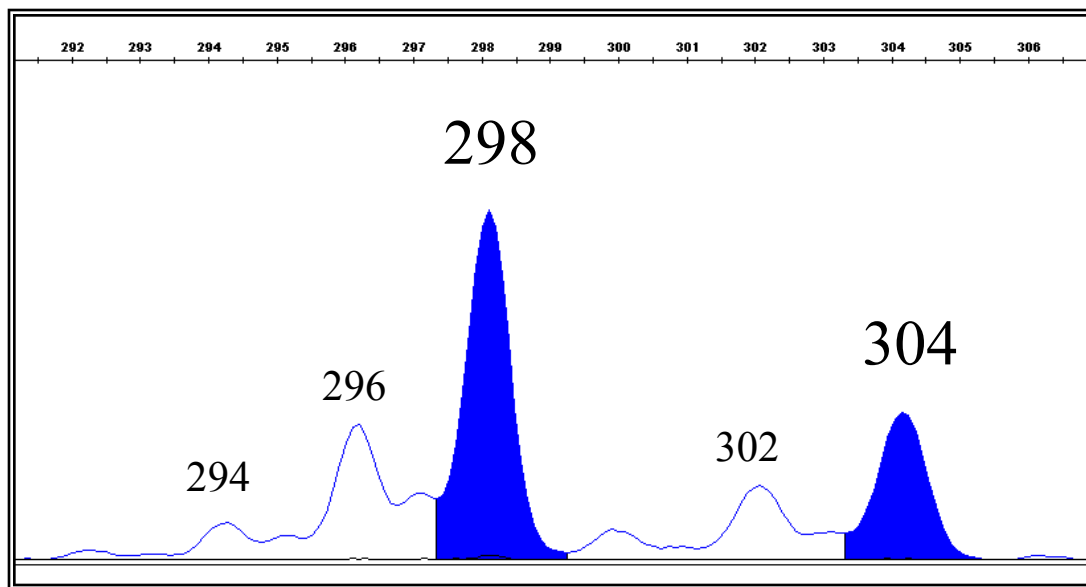
Kapilární elektroforéza ~ Fragmentační analýza



Stanovení délky PCR fragmentů srovnáním se známým standardem

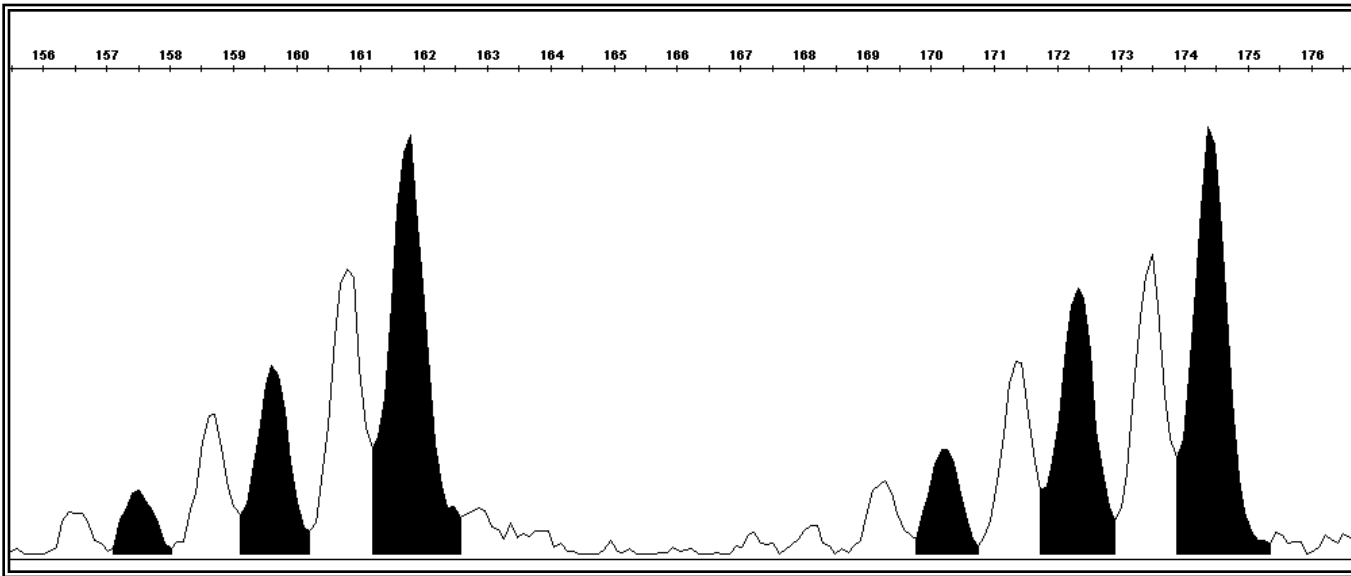


Genotyp mikrosatelitu na lokusu NED = 326/342 nebo 327/343



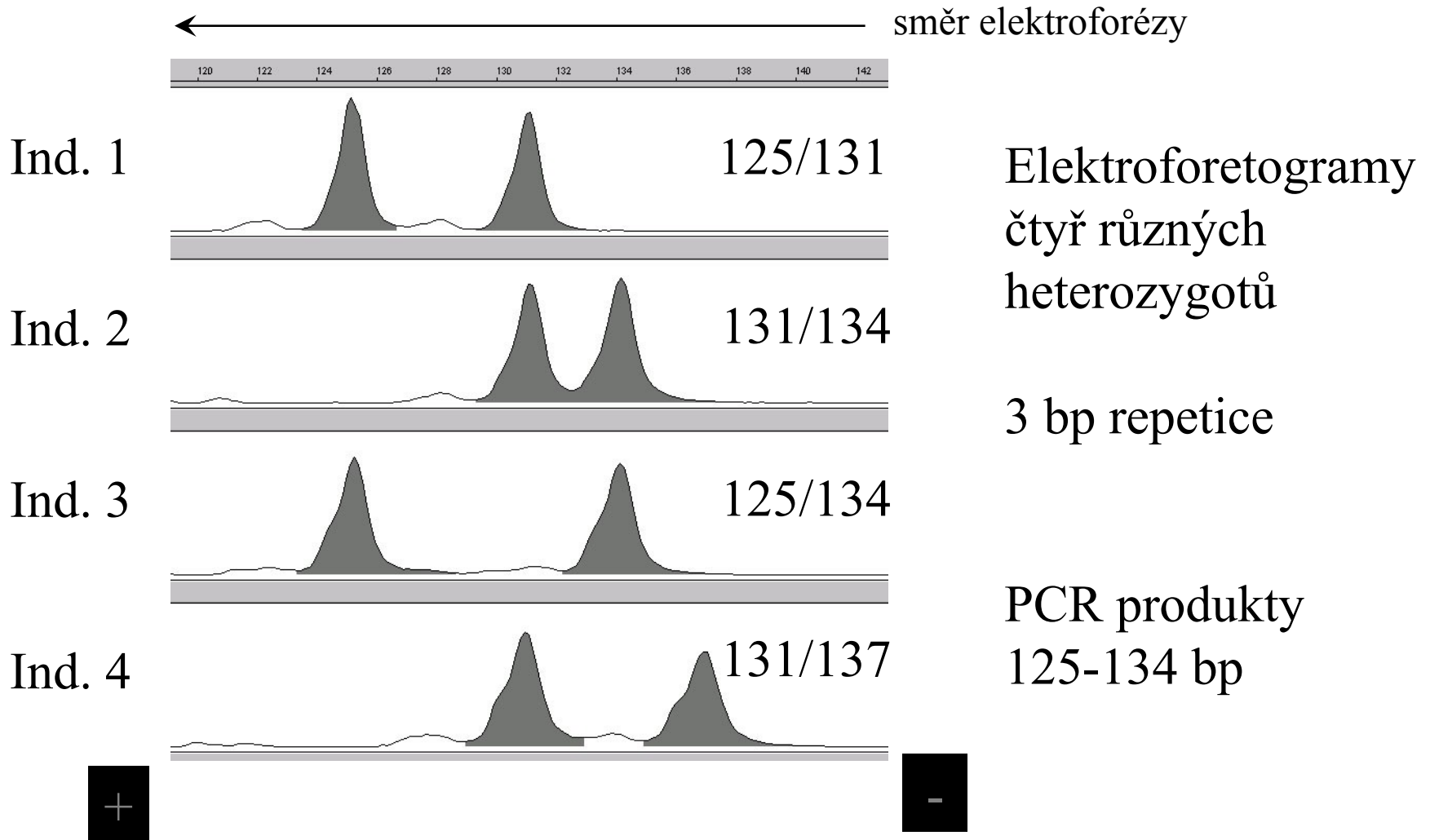
„stutters“ – chyby v důsledku „sklouznutí“ polymerázy při PCR

- často odlišují mikrosatelity od nespecifických PCR produktů
- rozdíl mezi alelou a „stutter“ je délka repetice (zde 2 bp)



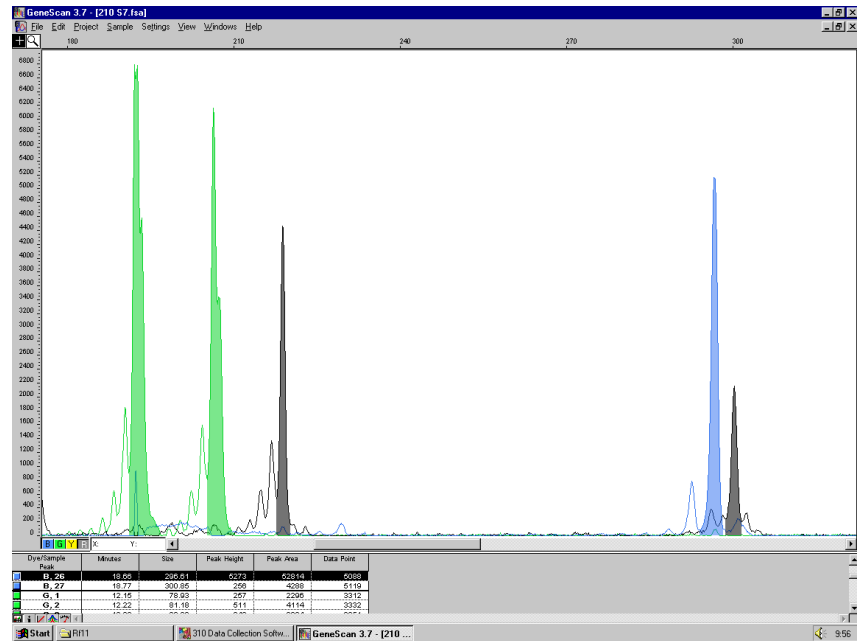
- alely a jejich stuttery jsou černě (rozdíl mezi nimi je 2 bp)
- bílé píky jsou tzv. „minus A-alely“ a jejich stuttery = výsledek jiné chyby polymerázy, a to nepřidání koncového adeninu
- rozdíl mezi černým a sousedním bílým píkem je 1 bp (tj. chybějící adenin)
- pattern daného lokusu je vždy specifický a často záleží na PCR podmínkách

Srovnání různých jedinců - analýzy příbuznosti



Různé značení různých znaků

- Snížení časových a finančních nákladů
- = „multiplex set“
- Až 4 různé barvy – analýza až 4 lokusů o stejné velikosti alel



Mikrosatelity - omezení

- nalezení lokusů (navržení primerů) je pracné a nákladné u volně žijících druhů (genomová knihovna, klonování, screening, sekvencování)

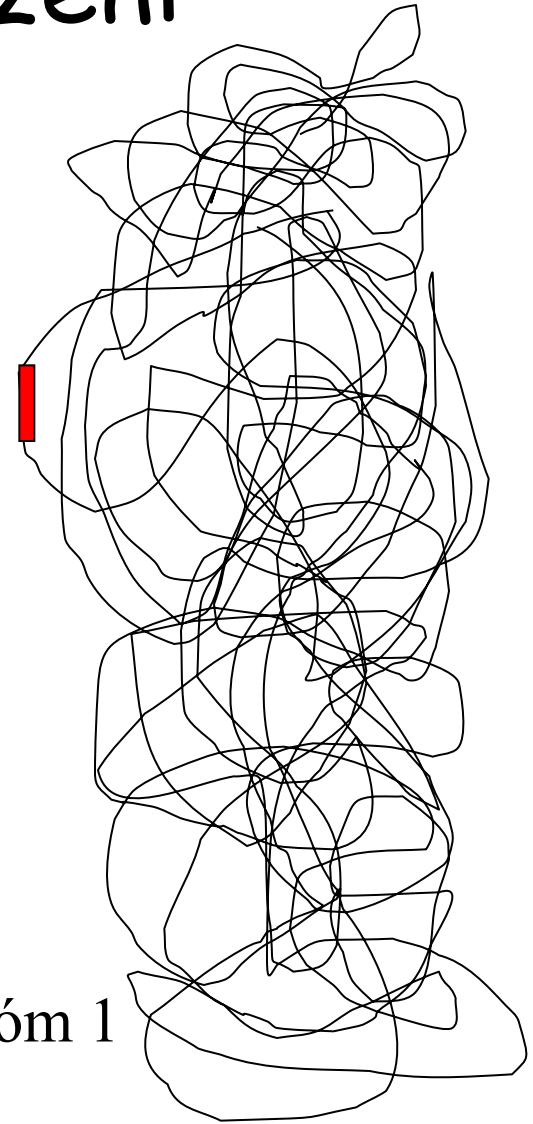


TTCAGG**CACACACA**TCTCTAGCTTCGA

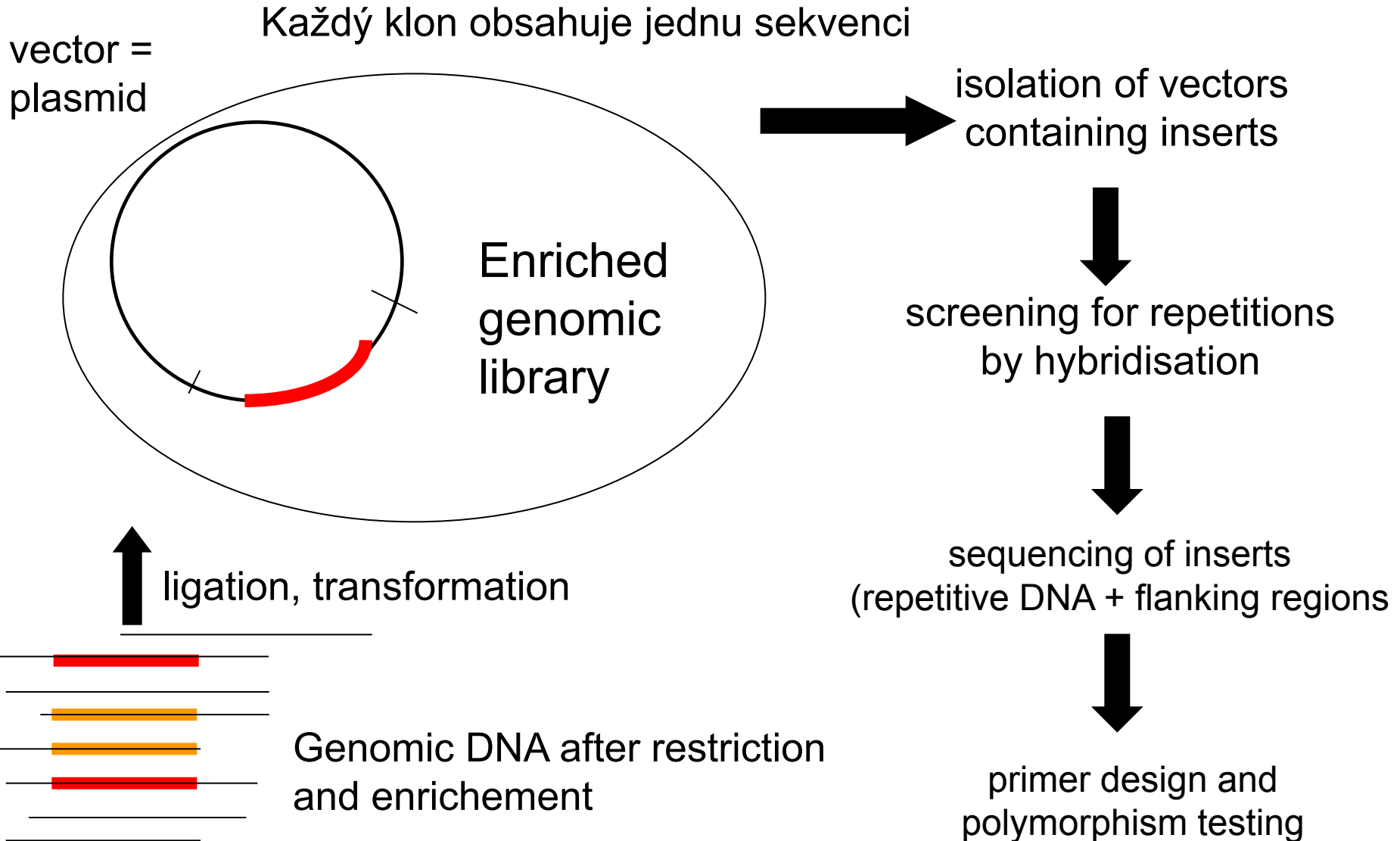


„flanking regions“ – ohraničují repetici a zde musí být navrženy primery pro PCR

Př.: chromozóm 1

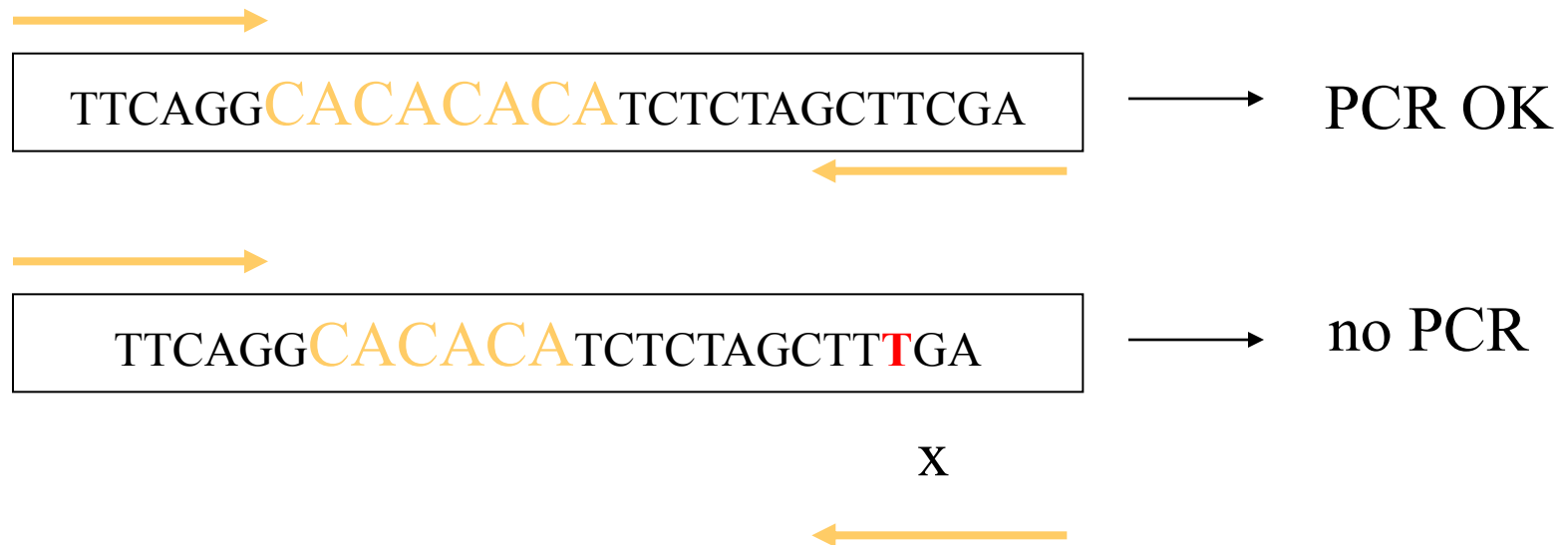


Restriction, enrichment, cloning, and sequencing



Mikrosatelity - omezení

- „cross-amplification“ – úspěšnost klesá s fylogenetickou vzdáleností
- nulové alely (mutace v primerových sekvencích) → vyšší proporce „homozygotů“



Mikrosatelity - budoucnost (?)

- „next-generation sequencing“ – velice rychlá sekvenace stovek tisíců fragmentů z jakéhokoliv genomu
- vyhledání repetitivních sekvencí vhodným softwarem a navržení primerů
- identifikace nových mikrosatelitů rychle, elegantně a relativně levně (1000 EUR)

Velikost alel a evoluce mikrosatelitů

- identita alel (stejně x různé) vs. rozdíl délek alel (tj. počet opakování)?
- Indikují podobně dlouhé alely na jednom lokusu bližšího společného předka než alely rozdílných délek?

Rodiče nepříbuzní



Rodiče příbuzní



- → k zodpovězení je nutno znát něco o evoluci mikrosatelitů (tedy jak vznikají nové alely)

Teoretické mutační modely

Dva extrémy

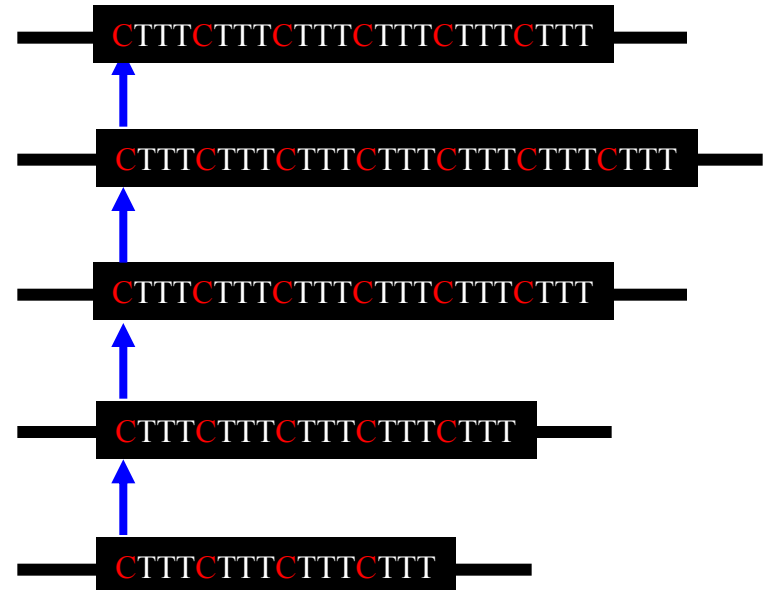
- **IAM – infinitive allele model**

(Při mutaci ztráta nebo získání libovolného počtu opakování. Vzniká nová alela, která doposud v populaci nebyla)



- **SMM – stepwise mutation model**

(Mutace způsobeny pouze ztrátou nebo získáním jediného opakování motivu. Mutací může vzniknout alela, která je již v populaci přítomna)



Další modely

- **TPM – two phase model**

(ztráta nebo zisk X opakování, při $X=1$ získáme SMM)

- **K-allele model**

(V populaci je přesně K myslitelných alel. Alela s konstantní pravděpodobností mutuje ke stavu jedné z $K-1$ dalších alel.)

- A další...

Indels

- inserce nebo delece 1bp či delších úseků – použití pouze pro modely vyžadující „identity“

TTCAGG CACACACA TCTCTAGCTTCGA

27 bp

SMM model – možno kvantifikovat podobnost alel

TTCAGG CACACACA CA TCTCTAGCTTCGA

29 bp

TTCAGG CACACACA TCTC G TAGCTTCGA

28 bp

TTCAGG CACAC GACA TCTCTAGCTTCGA

28 bp

„Indels“ – pouze pro analýzy, kde je vyžadována „identity“ a nikoliv podobnost

TTCAGG CACACCA TCTCTAGCTTCGA

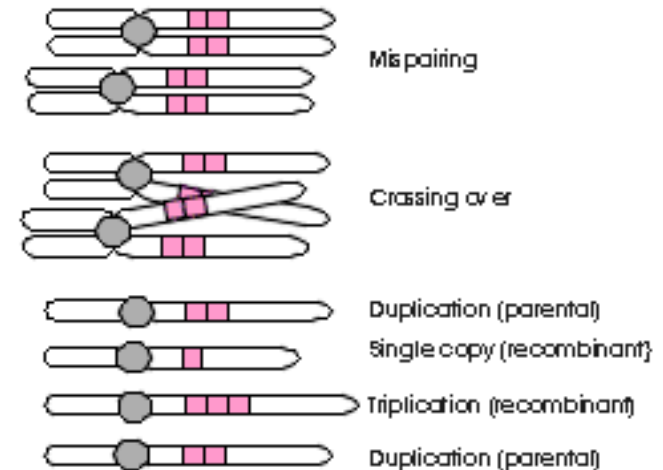
26 bp

TTCAGG CACACACA TCTCTAGTTCGA

26 bp

Proč je tolik alel? (microsatellite instability)

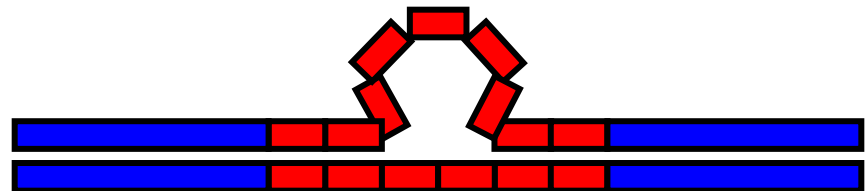
- **Nerovnoměrný (Unequal) crossing-over**
(díky špatnému alignmentu)



- **Sklouznutí polymerázy při replikaci**

Slip-strand mispairing

(při replikaci nejprve polymeráza sklouzne a vyrobí odlišný počet opakujícího se motivu mikrosatelitu, při alignmentu je pak část opakování vykloněna mimo dvoušroubovici, flanking regions tedy párují)



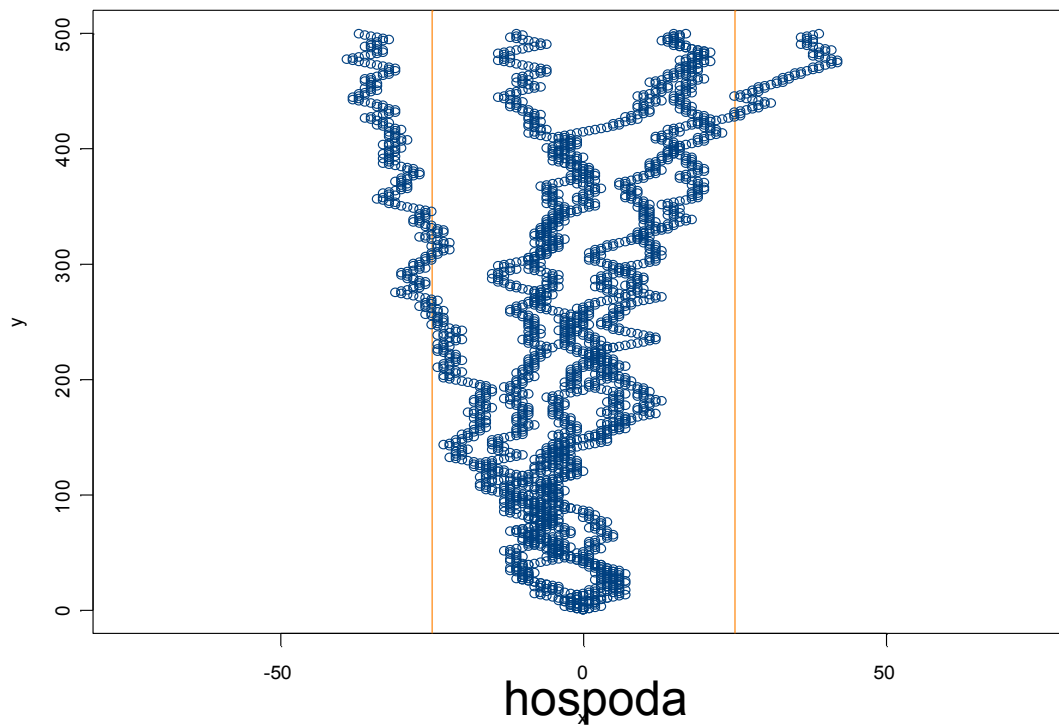
Brownův pohyb

- Simulace náhodné evoluce mikrosatelitů SMM mechanismem
- Pohyb opilého námořníka Browna po molu
- 1 krok vpřed
- 1 krok náhodně vlevo nebo vpravo
- Výchyvky od rovného směru se postupně zvětšují
- Dlouhé molo → pravděpodobný pád do moře



Výsledek 5ti pokusů opilého Browna dojít od hospody k majáku.

maják



Aplikace na evoluci mikrosatelitů

- Krok **vlevo** – **ztráta** jedné repeat unit
Krok **vpravo** – **nabytí** jedné repeat unit
- Náhoda → Random walking (Brownův pohyb)
Náhodné přidávání a odebrání jednotek opakování
- → vede k neúměrnému prodloužení či zkrácení
- Proti tomu svědčí častý výskyt stejných mikrosatelitových lokusů u příbuzných druhů
- → **proces prodlužování a zkracování asi není zcela náhodný**

Bias (skutečná data)

- **Kratší mikrosatelity** (s malým počtem opakování motivu) **mají zřejmě tendenci se spíše prodlužovat** (slabě převládají adice nad delecemi)
- **Delší mikrosatelity se spíše zkracují** (náchylnější k velkým delecím)
- **Delší mikrosatelity rychleji mutují** (díky více opakováním je vyšší pravděpodobnost pro sklouznutí polymerázy (SSM) – mají více alel)

Závěr

- Mechanismy evoluce mikrosatelitů stále nepříliš objasněny
- Stepwise mutation model SMM platí jen omezeně
- = nevýhoda v populační genetice (jsou rychle nahrazovány jinými markery, např. SNPs)
- = tolik nevádí při identifikaci jedinců a analýzy příbuznosti (paternity)