DNA markery ve šlechtění hospodářských plodin

Markery založené na polymorfismu DNA – DNA markery

- DNA marker představuje jakýkoliv dědičný rozdíl v DNA sekvenci mezi jedinci téhož druhu/příbuzných druhů a označuje (markeruje) specifickou oblast genomu, která se nachází v jeho bezprostředním okolí

- Lze využít k lokalizaci, detekci a izolaci neznámých genů a lokusů kvantitativních vlastností – QTL

- **DNA markery** jsou detekovány přímou analýzou DNA

Markery založené na polymorfismu DNA: základní vlastnosti

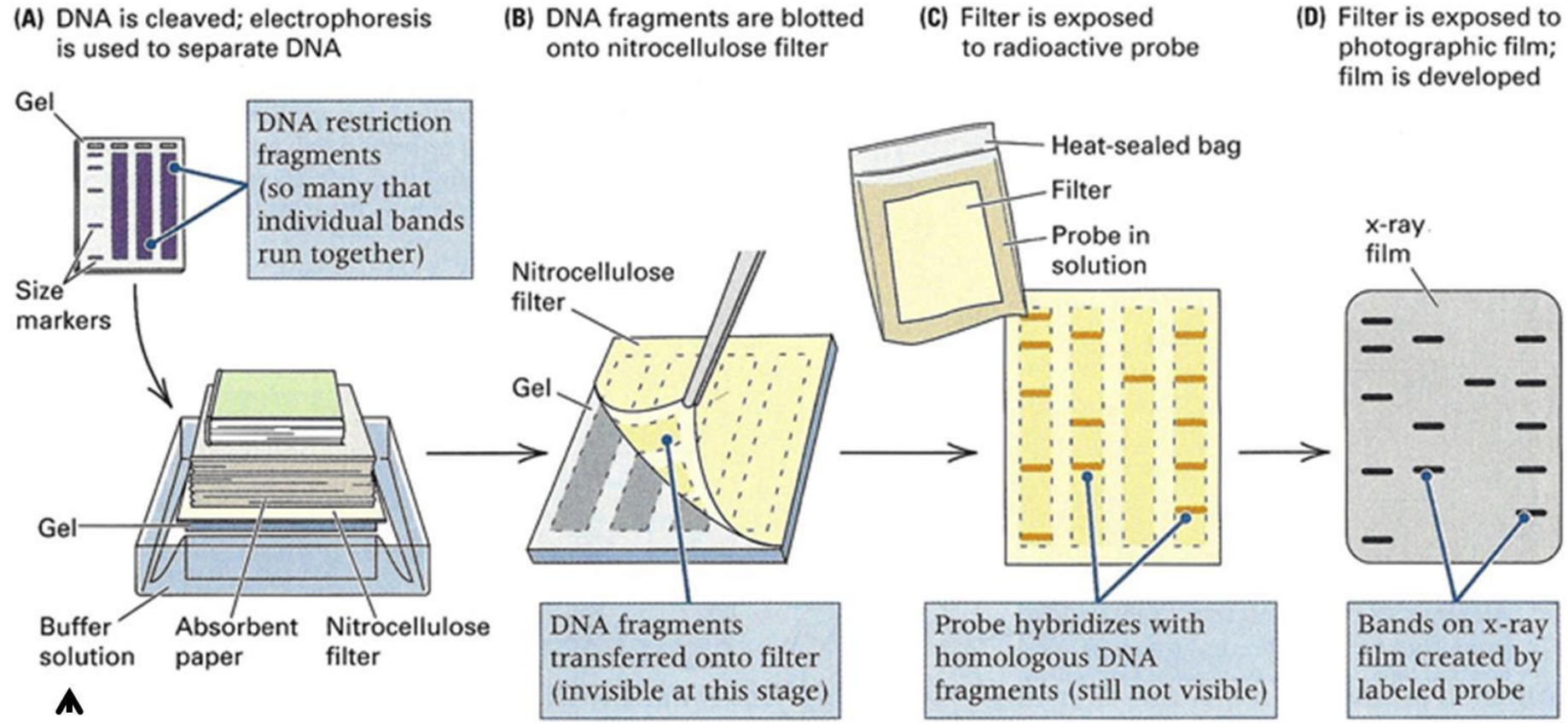
- jsou nedestruktivní - lze je aplikovat i při minimálním množství biologického materiálu
- mají velmi široké uplatnění: lze je aplikovat u všech organismů, kde je zvládnutá technika izolace DNA
- ze je využít pro charakterizaci velmi raných vývojových stádií organizmů
- nejsou závislé na podmínkách prostředí
- jejich počet je téměř neomezený

Základní typy DNA markerů

- **Markery založené na hybridizaci**
  - RFLP
- **Markery založené na PCR reakci**
  - RAPD
  - AFLP
  - SSR
- **Markery založené na sekvenování**
  - SNP (single nucleotide polymorphism), dělíme podle způsobu detekce:
    - SNP detekované pomocí biočipů (detekce založená na hybridizaci)
    - SNP detekované genotypováním pomocí sekvenování (GBS – genotype by sequencing)

RFLP – restriction fragment length polymorphism
Southernův přenos - schéma

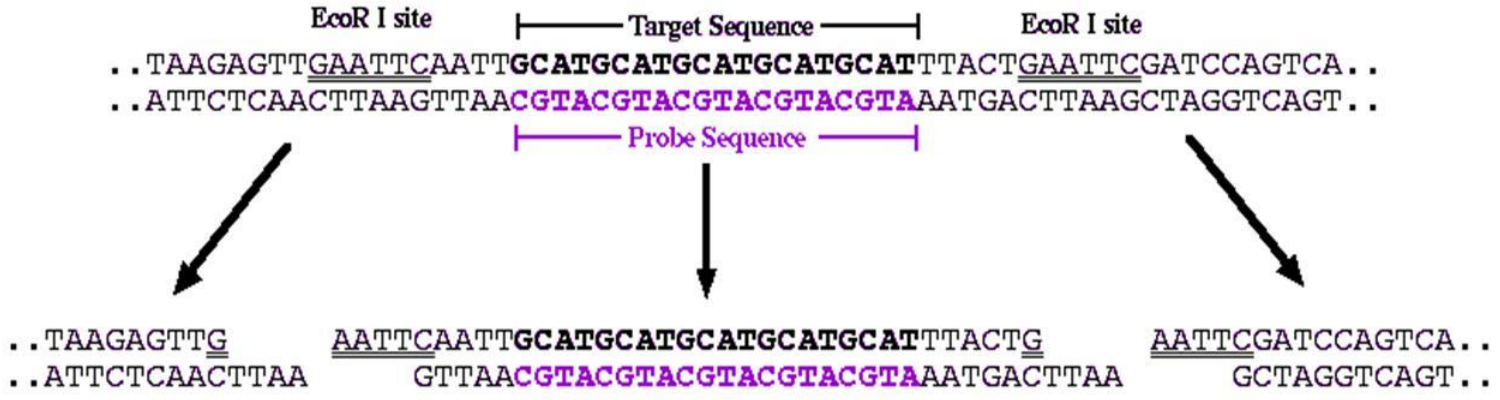
Polymorfizmus délky restrikčních fragmentů



Alkalický roztok denaturace DNA!

Sondy

- Sonda je tvořena úsekem jednořetězcové DNA molekuly, který má sekvenci komplementární k cílové DNA sekvenci
- Sondy mohou být značeny radioizotopy fosforu nebo fluorescenčně
- Sondy mohou mít různý původ:
  - Známé geny, - genomické, - odvozené od mRNA (EST, cDNA)



Sondy z hlediska funkce

- Cílené sondy: detekují konkrétní známý gen.
  - příklad: detekce počtu kopií transgenů v transgenních rostlinách

- genomické a odvozené od EST, cDNA
  - **Slouží pro detekci polymorfních RFLP markerů v genomech rostlin**

Jak zjistím, která sonda je polymorfní?

- sondy jsou testovány na analyzovaném souboru různých genotypů (dva a více): např. různé odrůdy, jedinci mapující populace (P, F1 a F2 jedinci)…
- Polymorfní sonda bude detekovat fragmenty o rozdílné velikosti a případně počtu v souboru testovaných genotypů
- Monomorfní sonda bude detekovat fragmenty o stejné velikosti a případně počtu u všech testovaných genotypů
- Zajímá nás **polymorfní** sonda = můžeme hledat asociaci mezi pozorovaným polymorfizmem (RFLP markerem) a vlastnostmi analyzovaných genotypů

Sondy z hlediska počtu detekovaných lokusů

- Jednolokusové sondy – detekují polymorfizmus v jednom lokusu, jejich polymorfizmus se snadno analyzuje, proto žádoucí
- Multilokusové sondy – detekují polymorfizmus na více lokusech, interpretace výsledků může být obtížná
- RFLP markery jsou tedy jednolokusové i vícelokusové

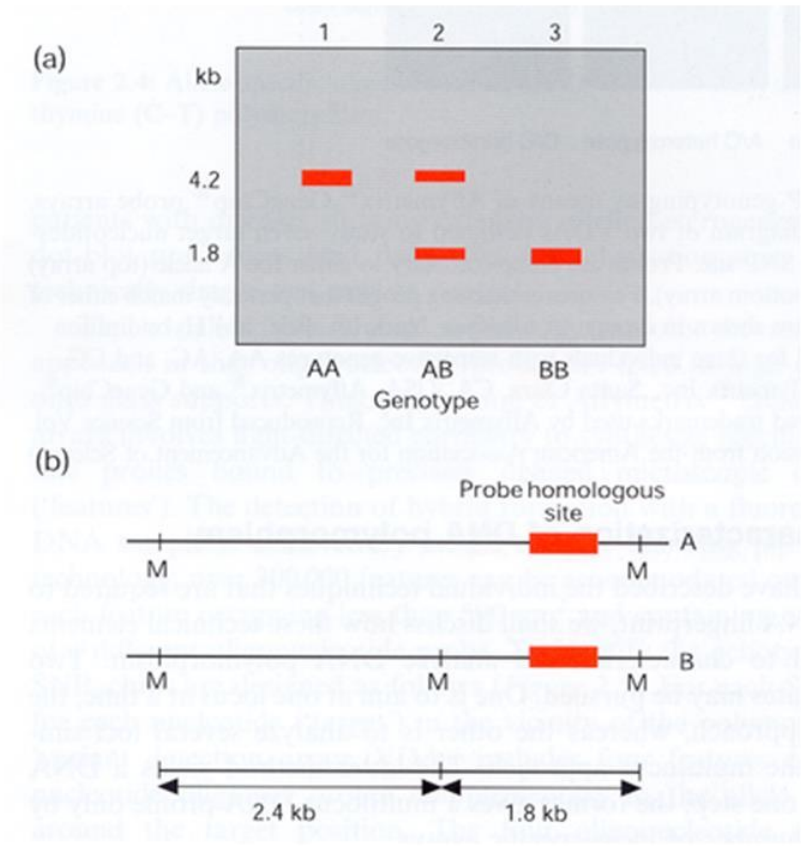
Zdroje RFLP polymorfizmů:

- nukleotidové substituce: zruší nebo vygenerují restrikční místo
- **přestavby** (inzerce/delece), mezi restrikčními místy
  - **Nejčastější typ RFLP polymorfizmů u rostlin** – empiricky zjištěno
  - Proto používáme restriktázy rozpoznávající 6 míst a né 4
- duplikace DNA úseků mezi restrikčními místy

4-cutter: 4^4 = každých 256bp, 6-cutter: 4^6 = 4,096bp,

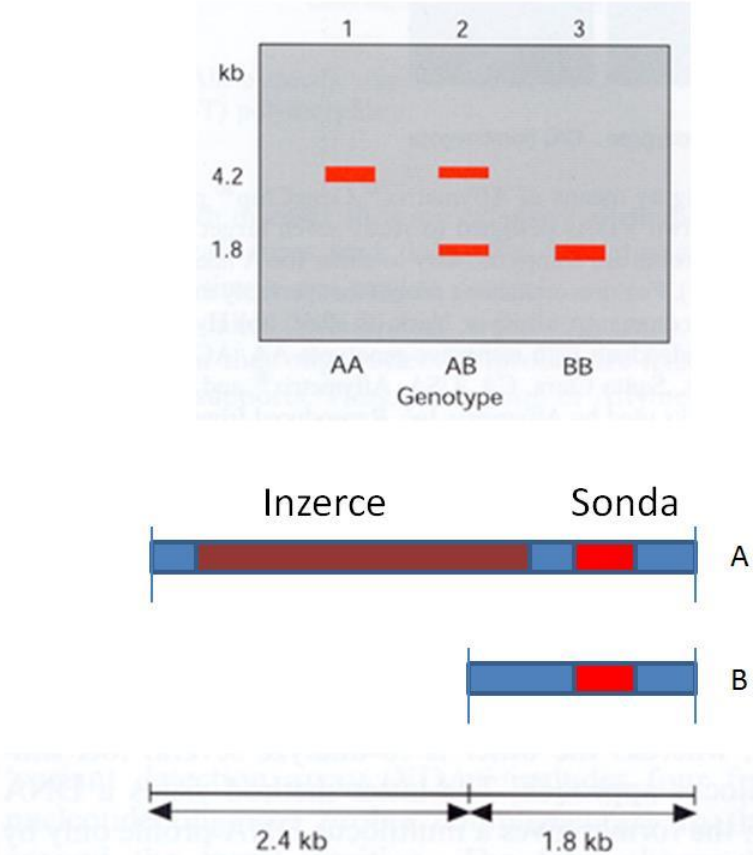
Typy RFLP polymorfizmů

Jednoduchá kodominance



Polymorfizmus způsobený nukleotidovou substitucí- v alele A zaniklo prostřední restrikční místo M

Jednoduchá kodominance



Polymorfizmus způsobený inzercí mezi restrikčními místy

Multilokusový RFLP polymorfizmus – jedna sonda více homologních genů

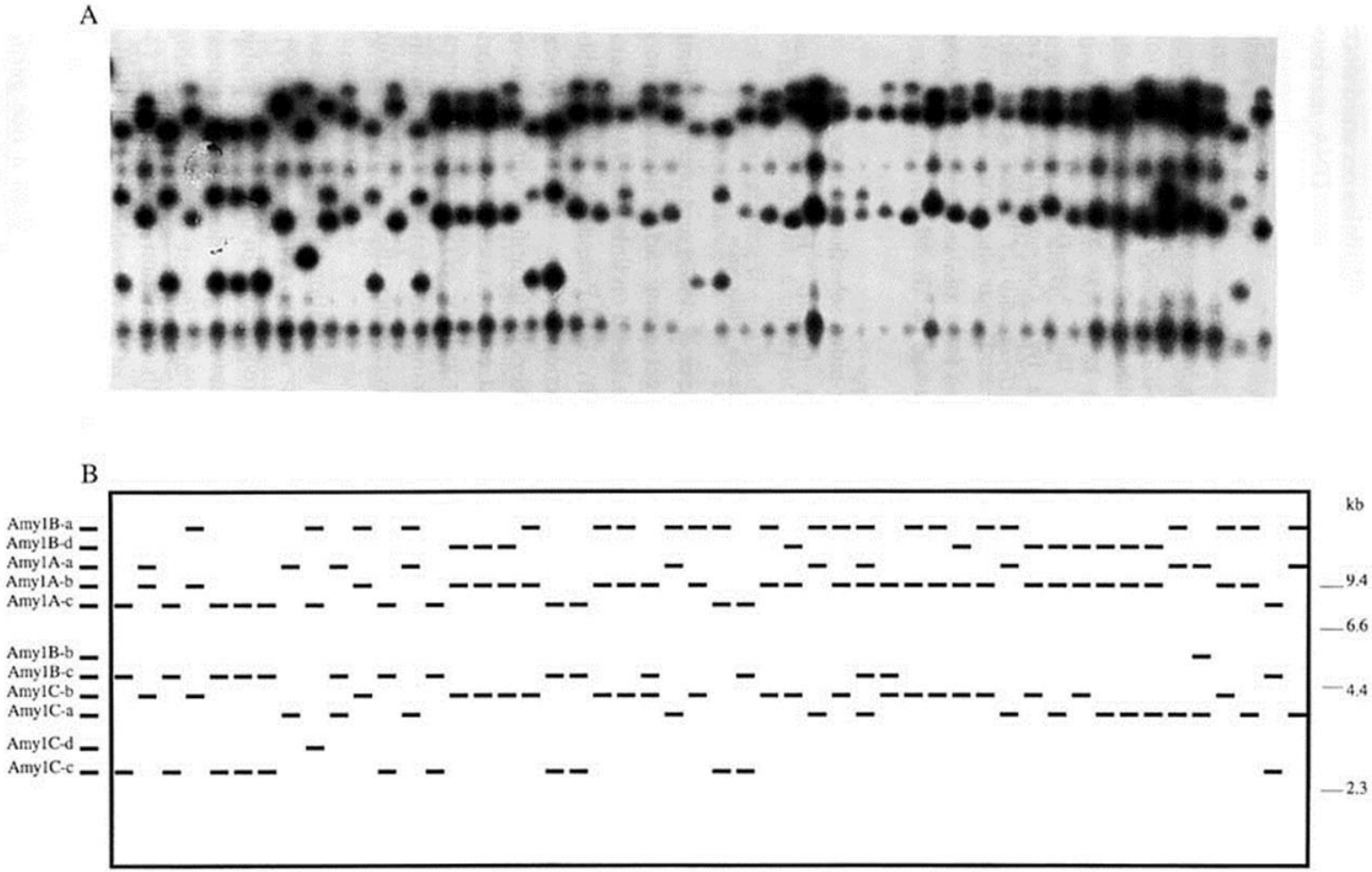
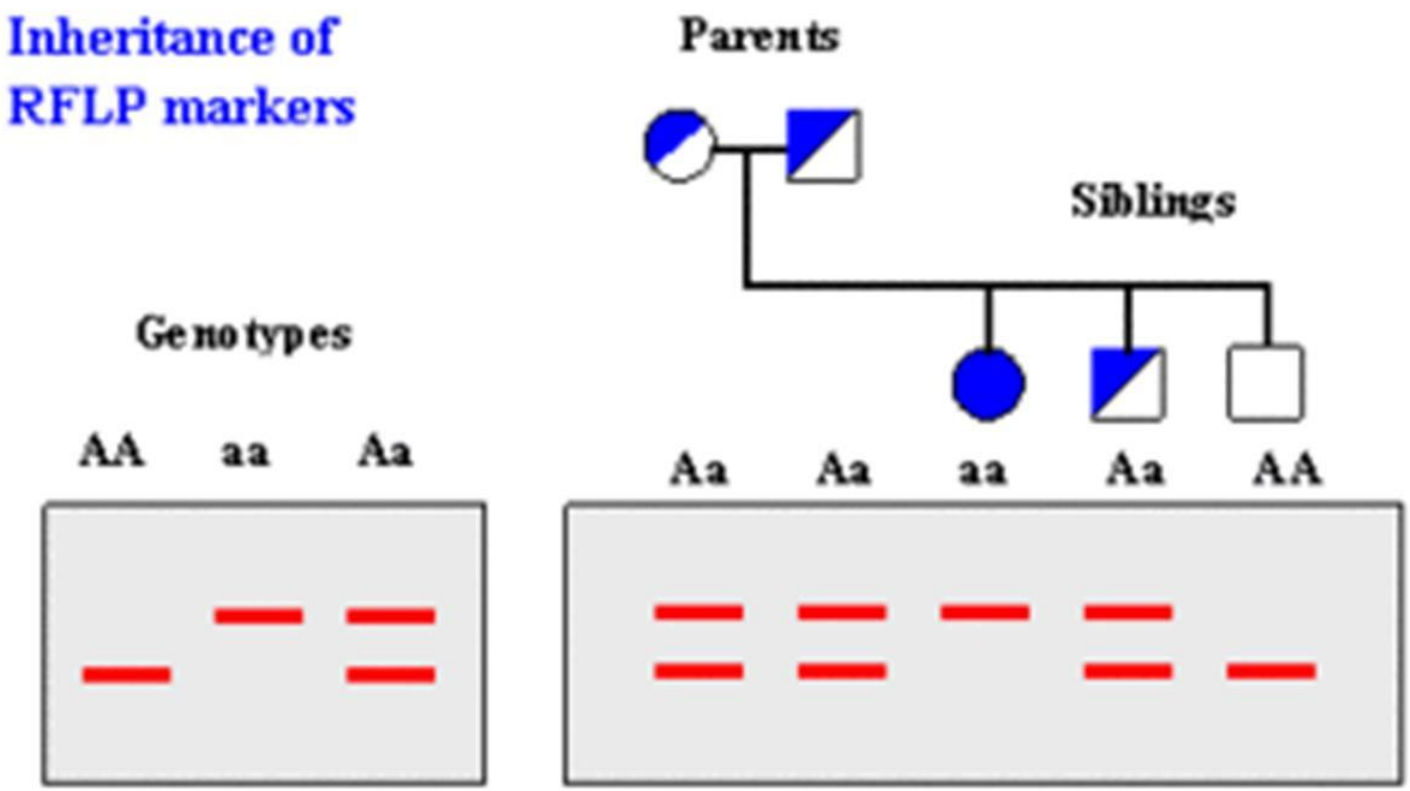


Fig. 2. Southern hybridization patterns of pOS103 on *Sca*I digest of total DNA from 50 rice cultivars (A) and their schematic representations (B). RFLP bands were associated with the genes *Amy1A*, *1B* and *1C*.

Mendelistická dědičnost RFLP markeru

RFLP markery se vykazují jednoduchou mendelistickou dědičnost (obrázek)



Vlastnosti RFLP markerů

Výhody:

- **Kodominance**, vhodné pro genetické mapování
- Většinou velké rozdíly v polymorfizmech-snadno detekovatelný
- Možnost opětovného použití stejné membrány pro novou sondu
- Vysoce reproducibilní

Nevýhody:

- Nižší pokryvnost celkové genomické variability
- Dostupnost sond, pro nové druhy připravit nové
- Velké množství gDNA potřeba (5 – 10 mikrogramů) na jedno štěpení
- Radioaktivní detekce, pracné, nákladné, nemožnost automatizace

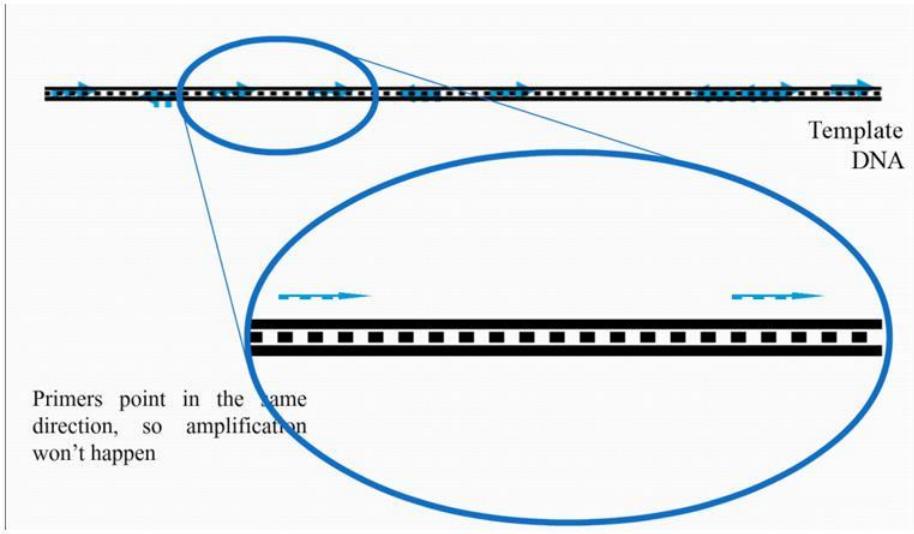
Přehled markerů založených na PCR

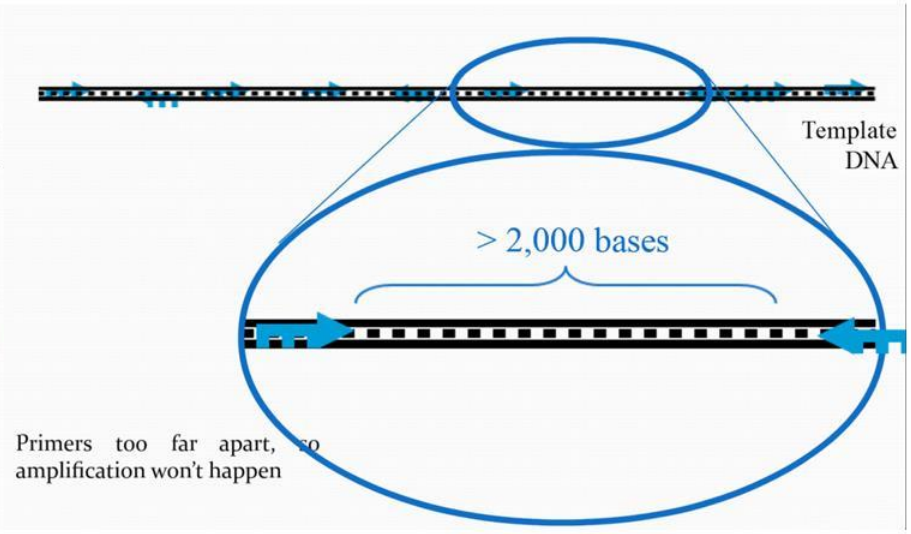
1. RAPD (Random Amplified Polymorphism DNA)
2. AFLP (Amplified Fragment Length Polymorphism)
3. SSR (Simple Sequence Repeat)

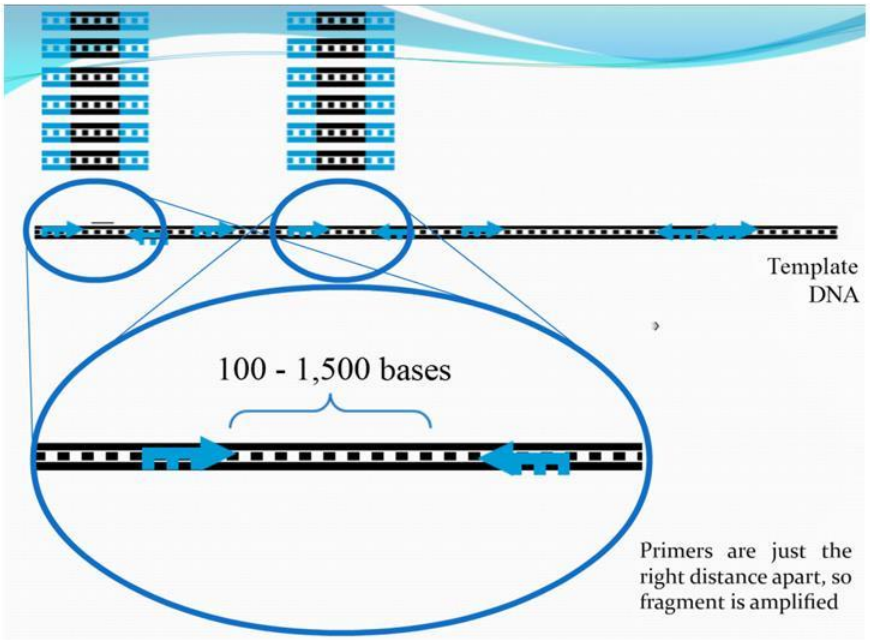
1. RAPD markery

- RAPD = Random Amplified Polymorphic DNA, Náhodně amplifikovaná polymorfní DNA
- RAPD je modifikací PCR která využívá pouze **jednoho náhodného primeru (délka 8-10bp !!)**. Také se označuje jako asymetrická PCR
- K amplifikaci dochází v případech kdy náhodný primer v průběhu „annealingu“ nasedne ve směru 5‘ – 3‘ „proti sobě“ v dostatečně blízkosti (< 2Kb) na obě komplementární vlákna gDNA
- RAPD polymorfizmus je výsledkem změny sekvence nukleotidů v místě navázání primeru nebo inzerce/delece
- Jedná se o **multilokusový anonymní** a **dominantní typ markeru**

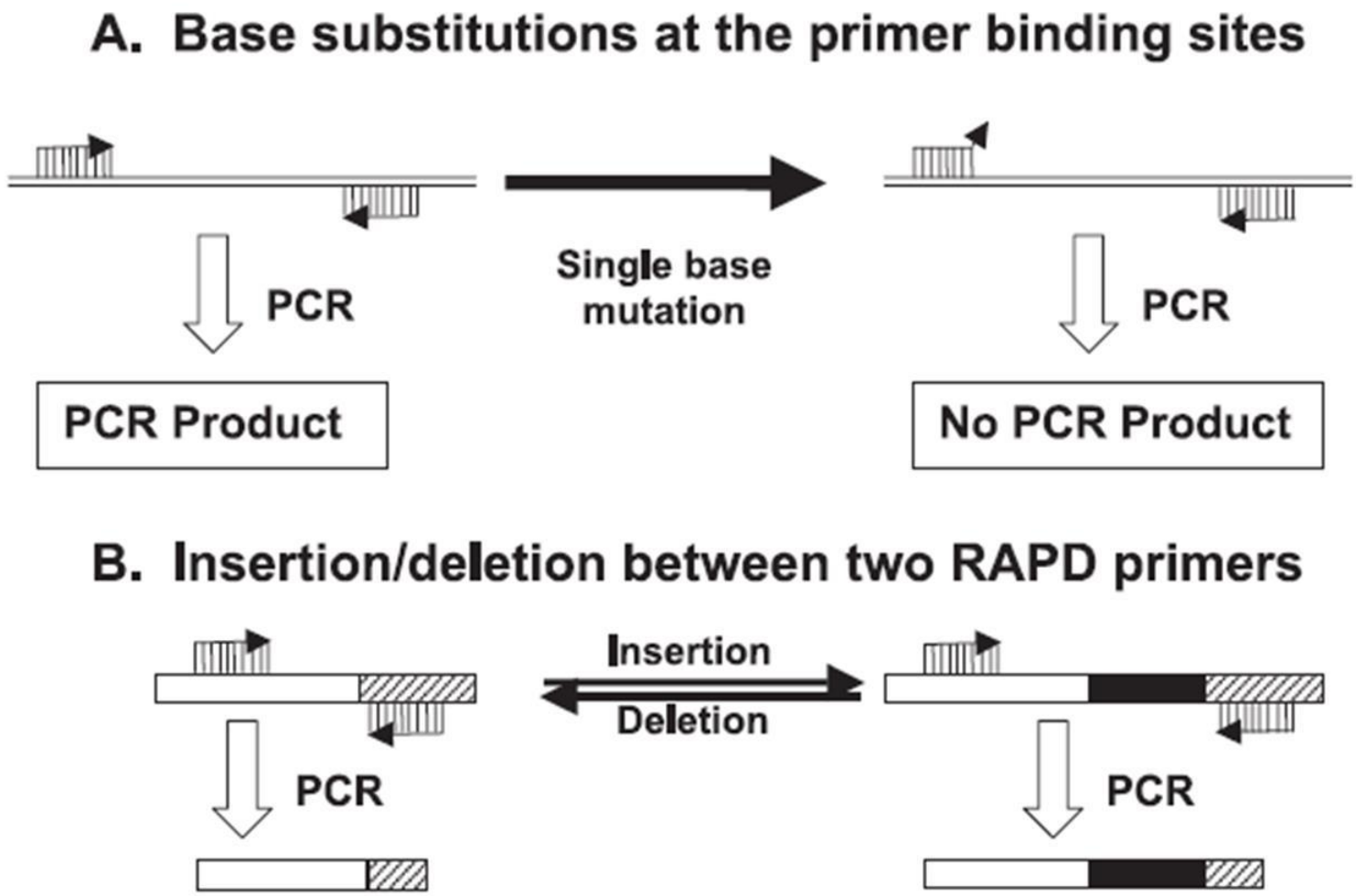
RAPD možné produkty amplifikace







Zdroje RAPD polymorfizmu



RAPD premix, reakce a elektroforeogram

Protocol

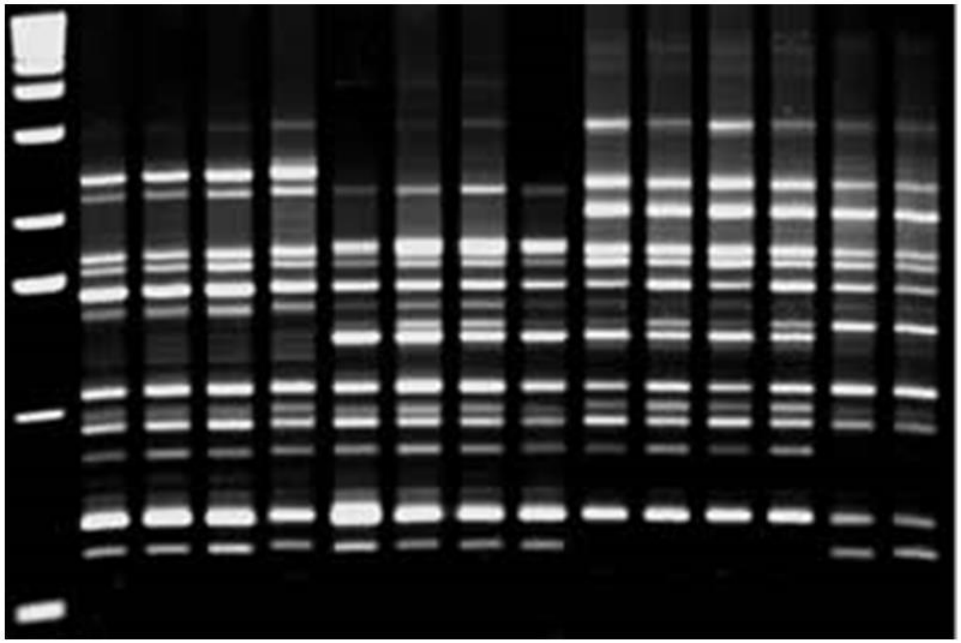
1) Master Stock Mixture

	Conc. stock solution	Vol	Final concentration
PCR-buffer + $MgCl_2$	10x (15 mM)	3µl	1x (1.5 mM)
Primer	15 µM	0.4 µl	0.2 µM
dNTP-mix	10 mM	0.51 µl	0.17 mM
Taq-polymerase	5 U/µl	0.25 µl	1.25 U
ddH_20		20.84 µl	
Final volume		25 µl	

2) Add 25µl of master mix to 5µl of your DNA in a sterile tube

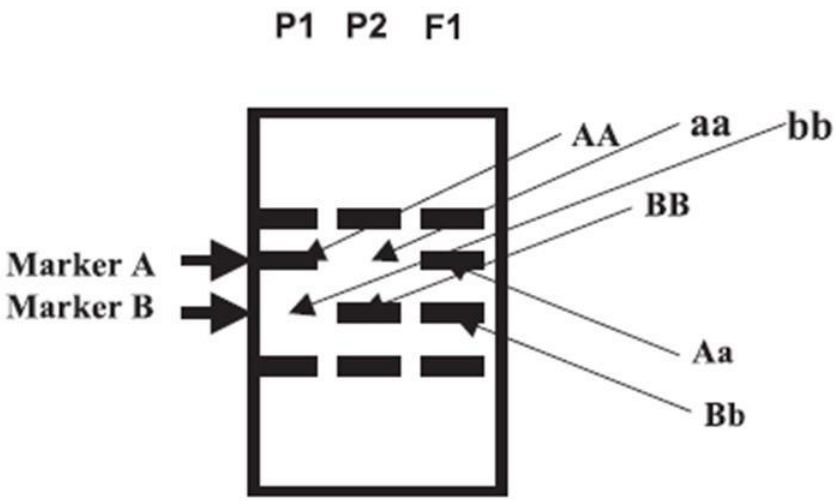
Note: In each PCR run you conduct, include 2 sample, one of control DNA without primer (3µl DNA), and one sample without DNA (5µl ddH_2O)

94° C for 1 minute
36° C for 1 minute
72° C for 2 minutes
45 cyklů



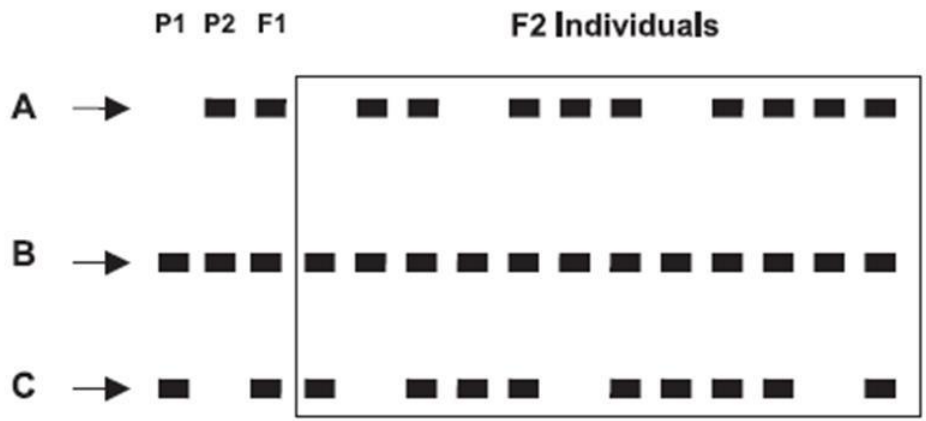
Dědičnost RAPD markerů

A. Identification of alleles of dominant markers



The alternative allele of a RAPD marker band is "no band"or absence of the band.

B. Dominant marker inheritance



M \ F	A	a
A	AA	Aa
a	Aa	aa

Vlastnosti RAPD markerů

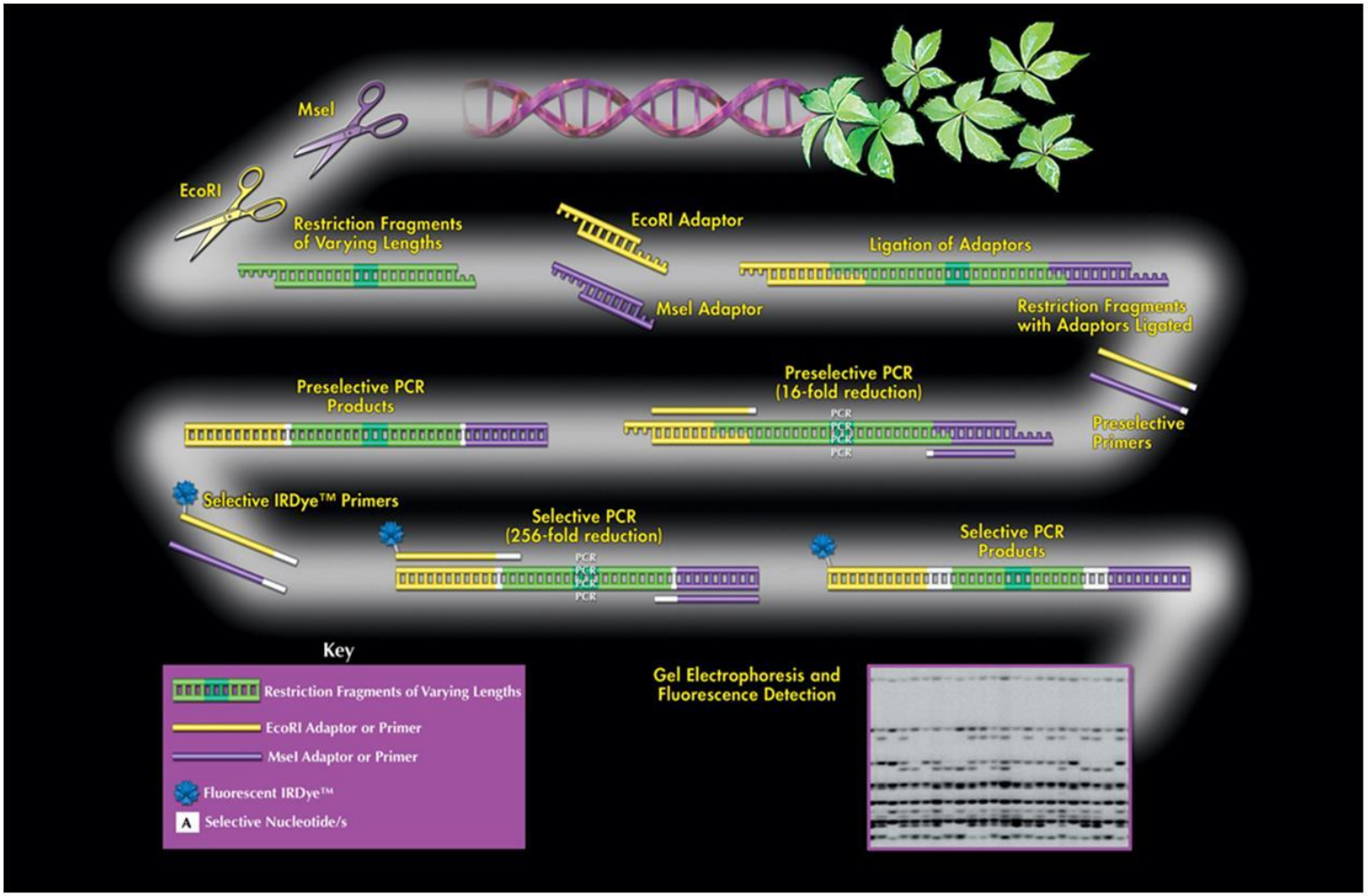
Výhody:

- Primery komerčně dostupné, nevyžadují znalost DNA sekvence
- rychlá detekce na agarózovém nebo akrylamidovém gelu
- Možno analyzovat velké množství jedinců a markerů

Nevýhody:

- Obtížně reprodukovatelné mezi laboratořemi i v laboratoři
- Nerozliší mezi homozygoty a heterozygoty

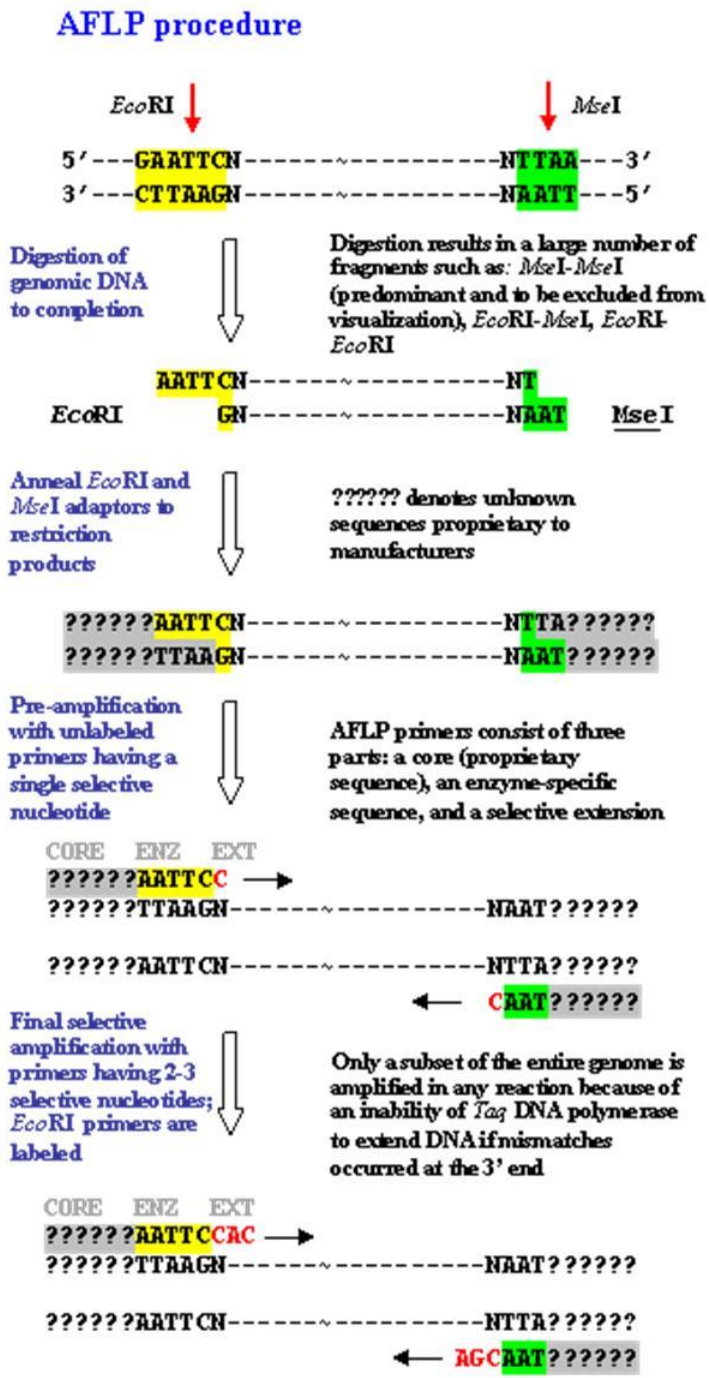
AFLP (Amplified Fragment Lenght Polymorfism), AFLP – polymorfizmus délky amplifikovaných fragmentů



Metodické kroky AFLP

- Metoda zahrnuje kombinaci RFLP a PCR
- Vlastní kroky AFLP:

1. Restrikční štěpení genomické DNA
2. Ligace adaptorů
3. Preselektivní amplifikace
4. Selektivní amplifikace
5. Separace AFLP produktů a jejich detekce
6. Vyhodnocení polymorfizmů

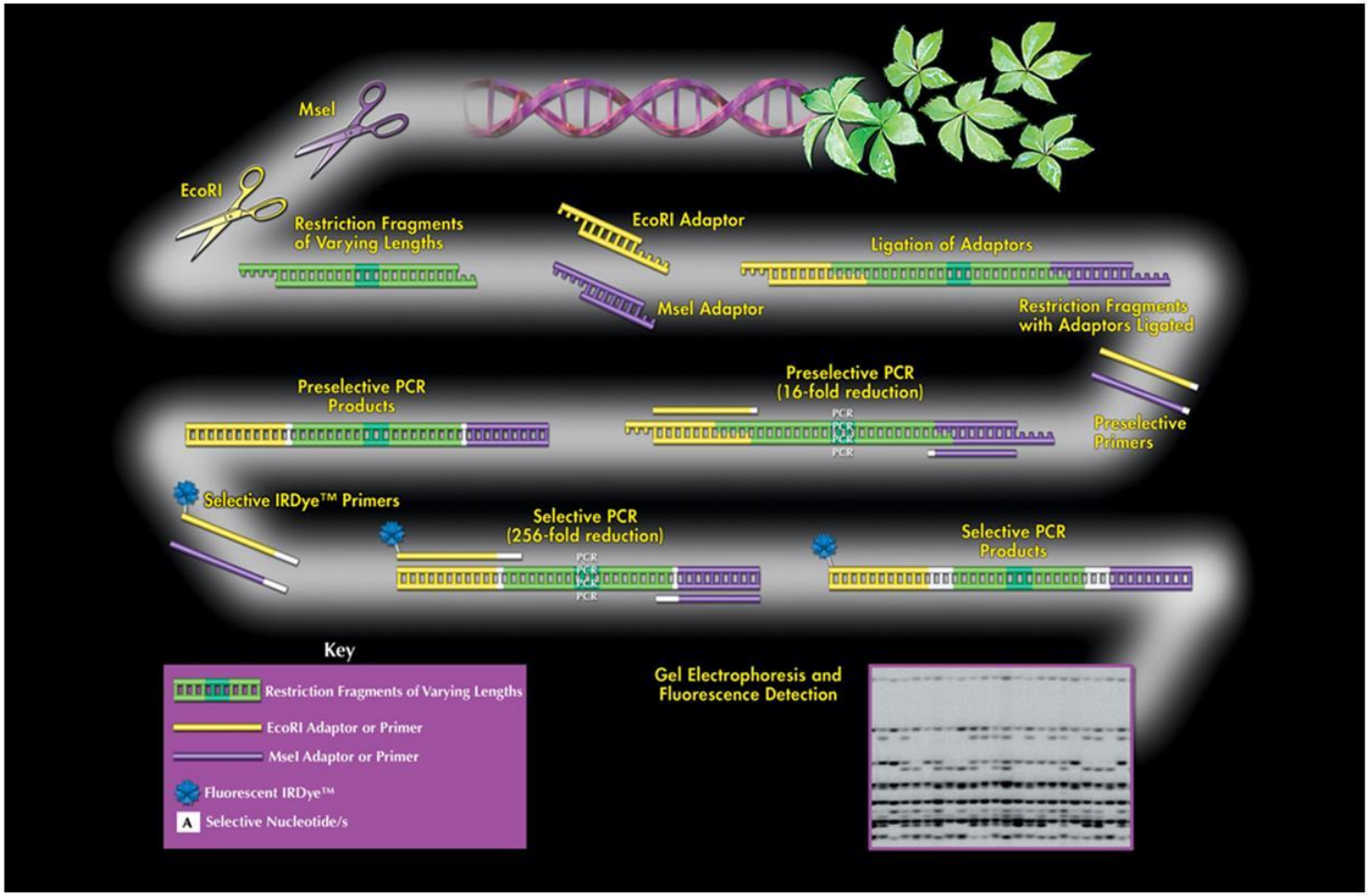


Shrnutí metodických kroků AFLP

Důležité je si uvědomit obrovský potenciál AFLP pro generování markerů

V kitu máme 64 EcorI a 64 MseI adaptorových primerů s tří nukleotidovými přesahy = 64 x 64 = **4096** různých AFLP profilů * cca 100 markerů na profil = **409 600 teoretických markerů** pro danou populaci/jedince

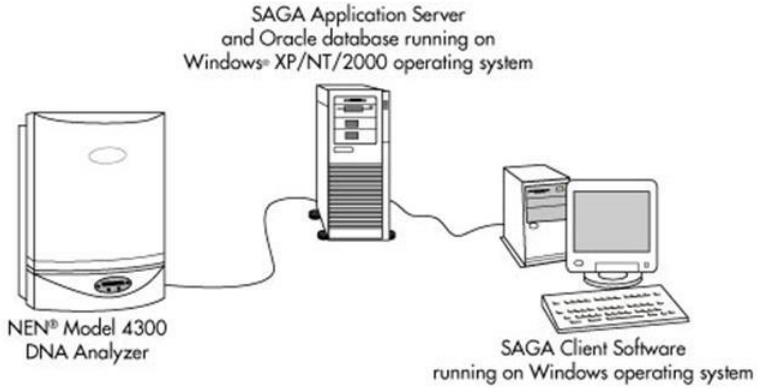
Detekce a analýza AFLP produktů pomocí systému LI-COR1

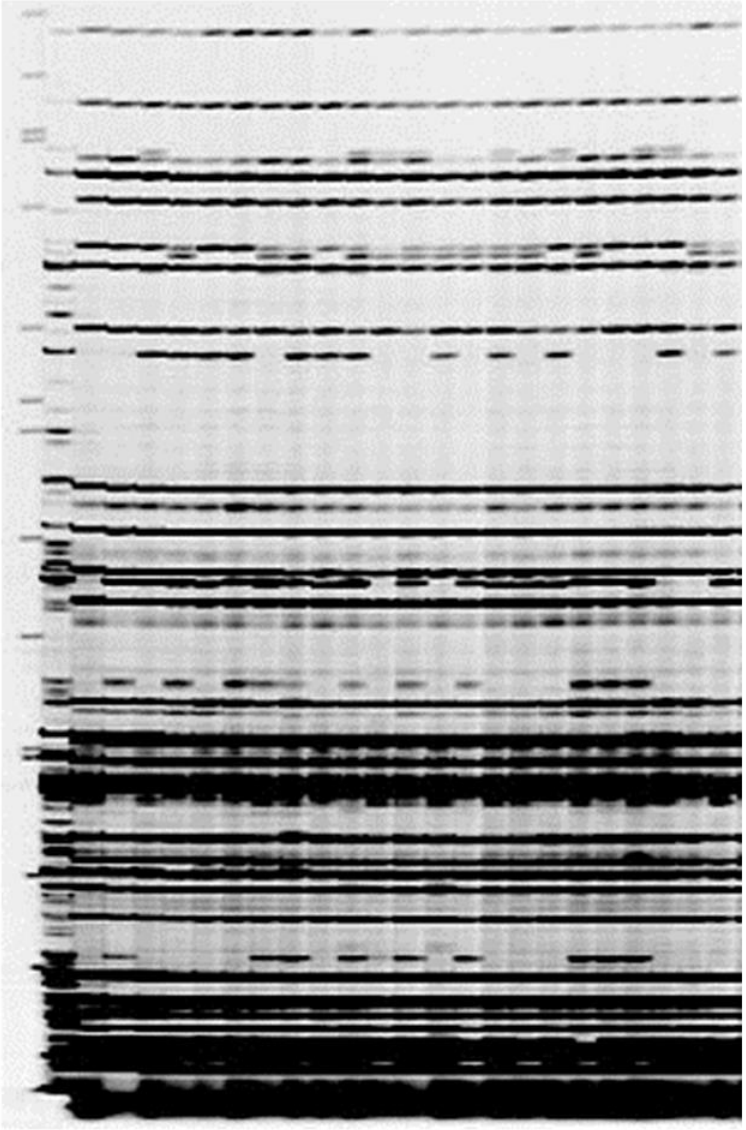


Výhody detekce AFLP produktů pomocí emise infračerveného spektra

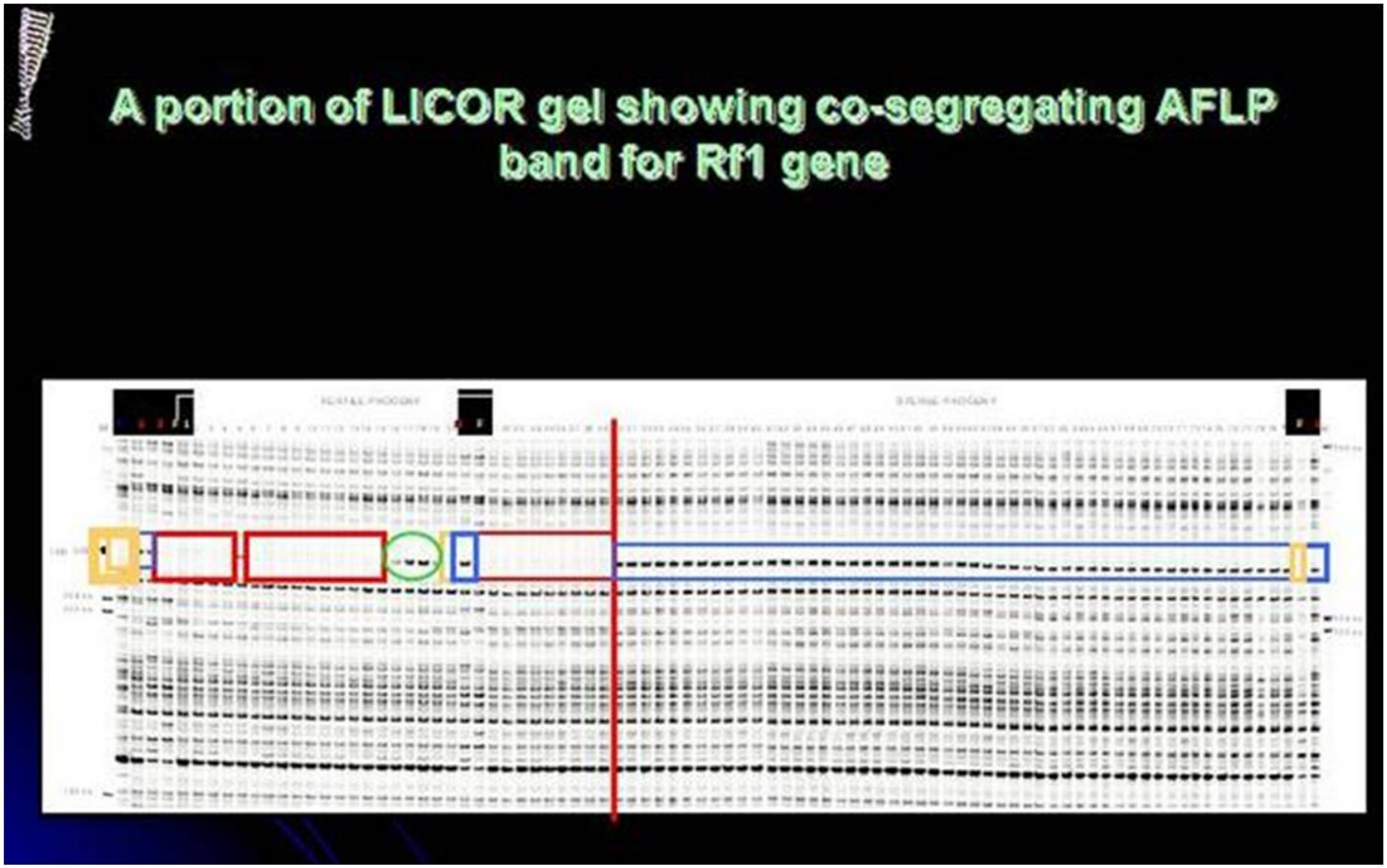
- Vysoké rozlišení, naprostá většina biomolekul neemituje fluorescenci v infračerveném spektru
- Proto je každý gel zobrazován jako reálný obraz
- IRdye ® infrared dye-labeled (either IRDye ® 700 or IRDye ® 800) EcoR I primers
- Dva excitační lasery a dva detektory
- možno analyzovat produkty dvou různých AFLP reakcí během jednoho „runu“ na 96 jamkové elektroforetické cele

Detekce a analýza AFLP produktů pomocí systému LI-COR 2





Kosegregace AFLP markeru s genem rezistence



Vlastnosti AFLP markerů

Výhody:

- Vysoká informativnost = vysoká detekční schopnost polymorfizmů genomu, až 100 lokusů jeden run
- Nevyžadují informaci o sekvenci
- Primery komerčně dostupné, nevyžadují znalost sekvence
- Poloautomatizovaný systém separace a detekce – LI-COR

Nevýhody:

- Finanční náročnost metody
- Dominantní markery

3. SSR (Simple Sequence Repeat), Opakování jednoduchých sekvencí:

- Označovány také jako mikrosatelitní markery, Short Tandem Repeats (STRs) nebo Sequence-Tagged Microsatellite (STMS)
- Mikrosatelity jsou krátké několika nukleotidové tandemově se opakující motivy:
  - 1-6 bp motiv = mikrosatelit
  - > 6 bp motiv = minisaltelit
  - Rostliny jsou bohaté na AT motiv

Příklady různých typů mikrosatelitních lokusů:

Microsatellites – Types

Based on repeat pattern

- Perfect – CACACACACACACACACACACACA
- Imperfect – CACACACACA CACACACA CACACACA

6. Compound – CACACACACACACA CATACATACATA CATACATACATA

- Complex – CACACACACACACACACA AATAATAATAATAATAATAAT

Další charakteristiky mikrosatelitů:

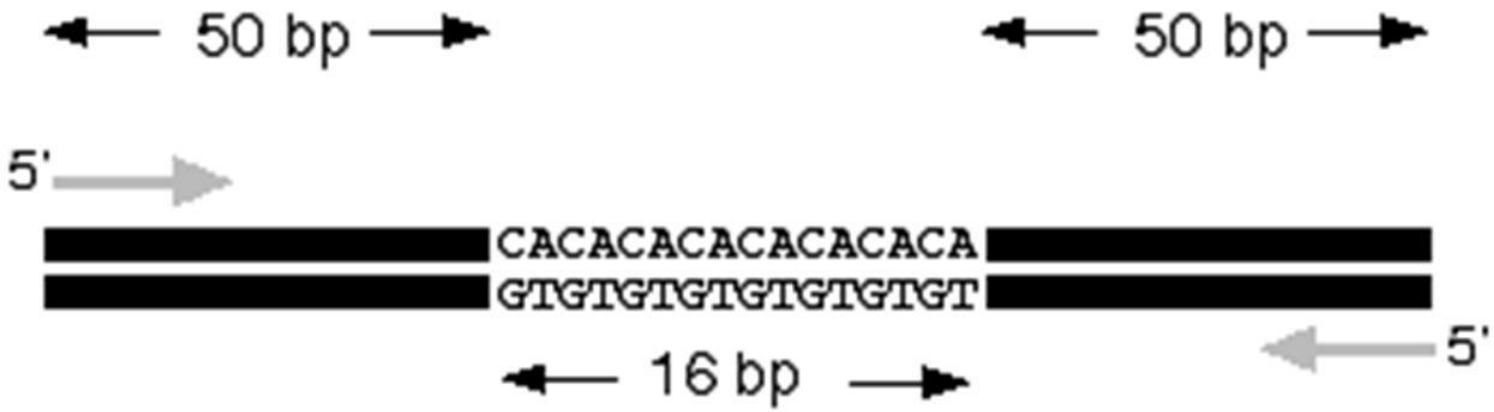
- 5-100 repeticí na lokus
- Náhodně distribuovány v genomu, jsou v nekódujících i kódujících oblastech
- 10000-100000 na genom
- Vysoký stupeň polymorfizmu, který je dán různým počtem opakování jednotlivých repetic
- Polymorfizmus pravděpodobně vzniká prokluzováním DNA polymerázy v průběhu replikace

Další charakteristiky mikrosatelitů:

- Kodominance
- Jednoduchá mendelistická dědičnost
- Lemující oblasti, ty na které se navrhují primery jsou vysoce konzervované mezi příbuznými druhy i rody

Detekce SSR polymorfizmu

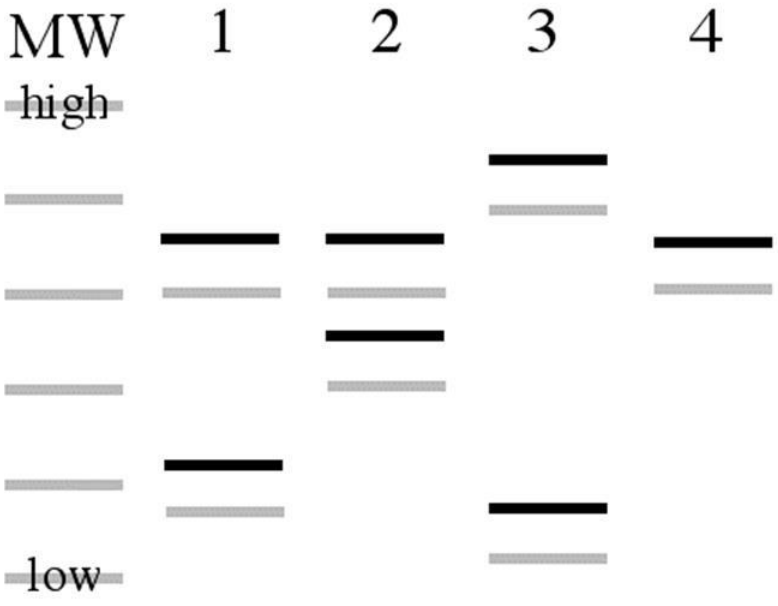
- Unikátní primery navrženy na lemující oblasti mikrosatelitů - jeden primer pár jeden lokus



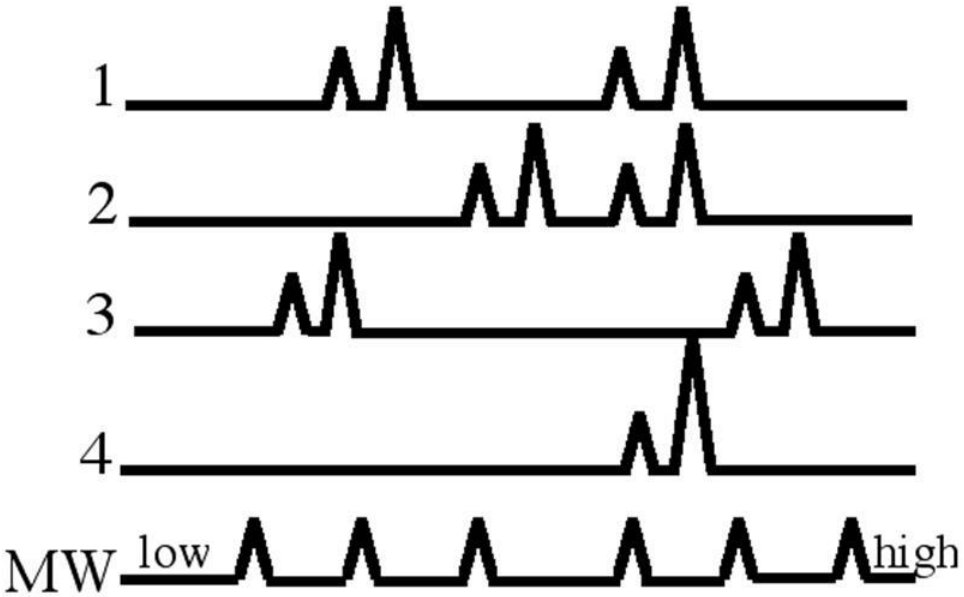
Detekce SSR polymorfizmů

„Slutter“ pík, malý pík (nebo šedý proužek na gelu) - následuje bezprostředně po hlavním píku, který odpovídá alele mikrosatelitu. Vzniká prokluzováním DNA polymerázy – PCR produkt obvykle kratší o 1-4 repetice. Pokud je „slutter“ pík příliš vysoký může značně ztížit detekci mikrosatelitních alel

Gel:

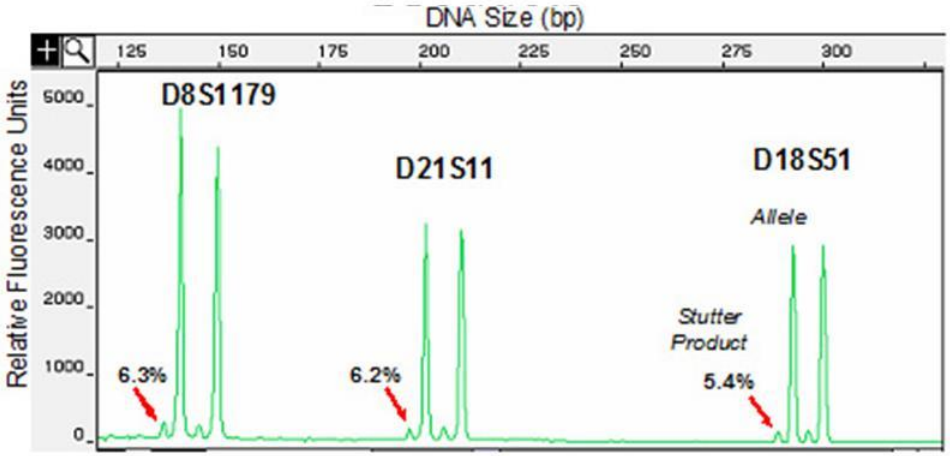


Kapilární elfo:



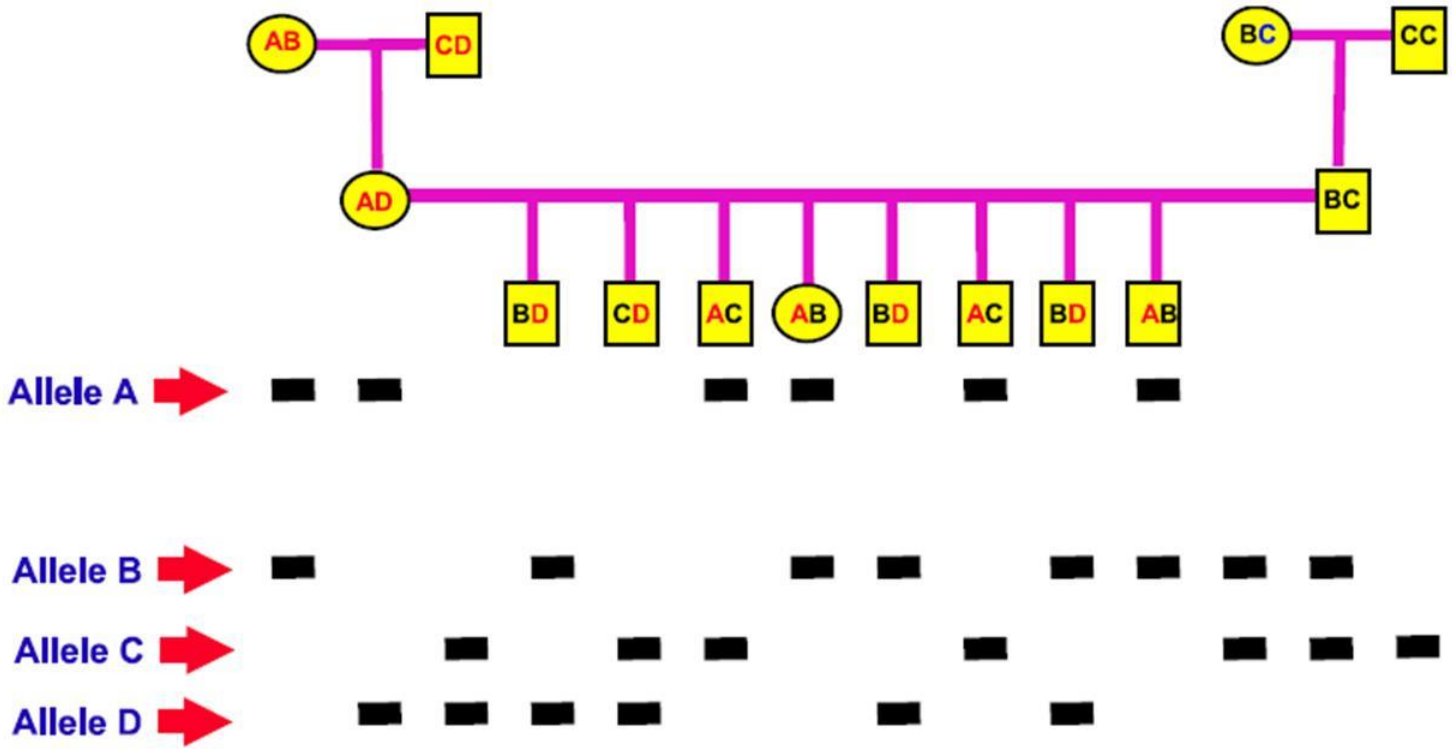
Automatizovaná detekce SSR polymorfizmů

- Vysokokapacitní genotypování SSR markerů, multilokusová analýza – využíváme fluorescenčně značené primery,
- Přístroje: LICOR, sekvenátory



Simultánní analýza tří různých SSR lokusů: kapilární elektroforéza

Mendelistická dědičnost SSR markerů

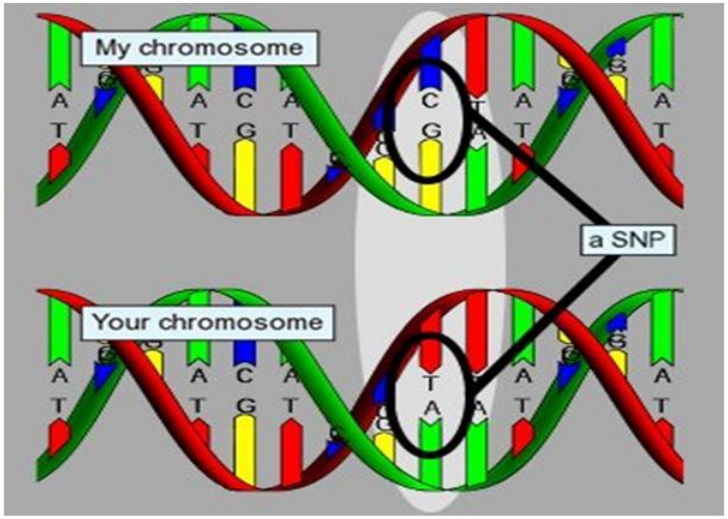


Shrnutí vlastností SSR markerů

- Výhody:
  - Kodominance
  - Vysoká abundance v genomu
  - Rovnoměrná distribuce v genomu
  - Malá velikost lokusu (můžu snadno amplifikovat pomocí PCR a rozlišit na gelu/kapilární elektroforéze)
  - Vysoký stupeň polymorfizmu
  - Automatizovatelná detekce
- Nevýhody:
  - Obtížná *de novo* identifikace mikrosatelitů, vývoj SSR markerů je pracovně a časově náročný

Markery založené na sekvenování

- Zahrnují především analýzu **SNP** (Single Nucleotide Polymorphism) - jednonukleotidový polymorfizmus
- SNP = polymorfizmus způsobený záměnou jedné báze v dané nukleotidové pozici - lokusu (obrázek SNP)



Vlastnosti SNP (jednonukleotidových polymorfizmů)

- Nejvíc abundantní polymorfizmus v genomu = např. genom *Arabidopsis* ekotyp L*er* se liší od genomu ekotypu Col celkem 349171 SNP!!, 1SNP na každých 300-1000bp

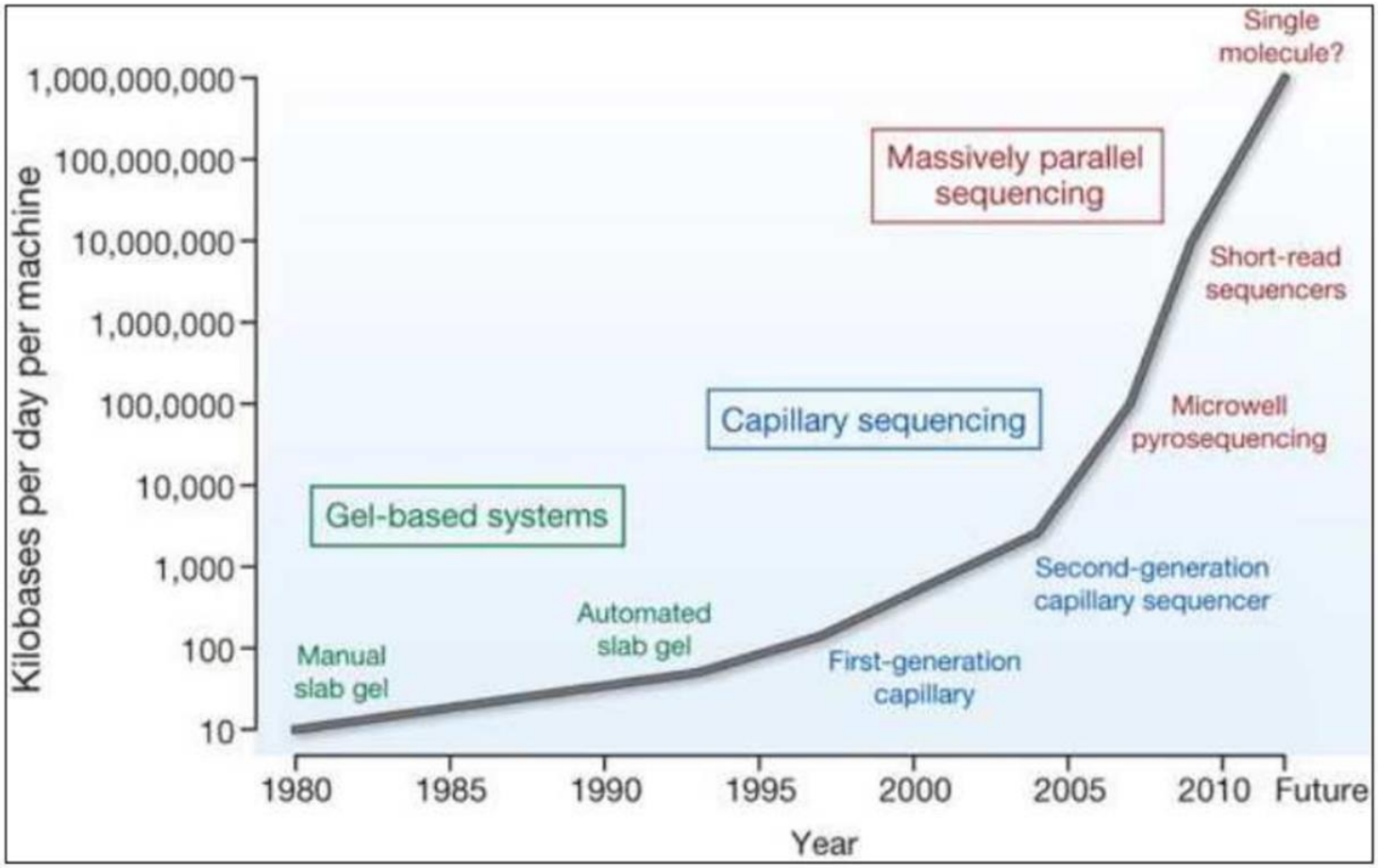
- Z toho plyne: nejvíce „hi-tech“ markery

- SNPs jsou jednolokusové, kodominantní markery

- Teoreticky mohou existovat v daném lokusu max. 4 SNP

- Obvykle však detekujeme v analyzované populaci 2 SNP v jednom lokusu

Sekvenování - základní metoda pro získávání a detekci SNP



Typy a výkonost sekvenačních přístrojů Illumina

MiSeq

Run Mode	N/A
Flow Cells per Run	1
Output Range	0.3-15 Gb
Run Time	5-55 hours
Reads per Flow Cell†	25 million‡
Maximum Read Length	2 x 300 bp
System Overview	Speed and simplicity for targeted and small genome sequencing.

Obrovská výkonost od 0,3 do 1800 Gb!! na jeden run!!

HiSeq X System

Run Mode	N/A	N/A
Flow Cells per Run	1 or 2	1 or 2
Output Range	900-1800 Gb	900-1800 Gb
Run Time	<3 days	<3 days
Reads per Flow Cell†	3 billion	3 billion
Maximum Read Length	2 x 150 bp	2 x 150 bp
System Overview	Maximum throughput for production-scale human whole-genome sequencing.	Maximum throughput and lowest cost population-scale human whole-genome sequencing.
	Learn More	Learn More

Lidský genom pod 1000 dolarů!! = HiSeq X Ten →

Table 1: HiSeq X System Sequencing Capacity.

	HiSeq X Ten System	HiSeq X Five System
Minimum Number of Instruments	10	5
Annual Genome Capacity	> 18,000	> 9000
Price per 30x genome	< $1000	< $1500

When operating at scale, HiSeq X Ten and HiSeq X Five Systems generate a staggering level of throughput, sequencing thousands of genomes per year.

Obecné schéma získávání SNP

1a) Sekvenace kompletních genomů *de novo* – dideoxy metoda (starší genomy) a NGS (Next Generation Sequencing) v kombinaci s dideoxy metodou (genomy od roku 2005) = <u>získání referenčního genomu</u>

1b) Komparativní re-sekvenace kompletních nebo redukovaných genomů jedinců populace daného druhu u kterého je dostupný referenční genom = NGS metody

2) Sekvenace redukovaných genomů u jedinců populace daného druhu u kterých není dostupný referenční genom = NGS metody

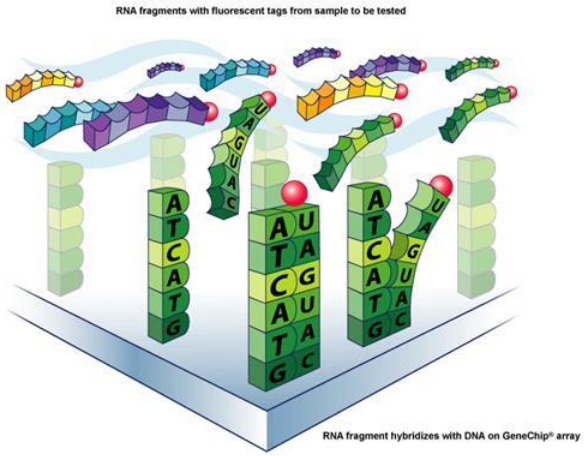
3) Analýzou transkriptomu: NGS sekvenace kompletních transkriptomů u jedinců v populaci - RNA-seq, analýzy transkriptomických databází (EST, cDNA)

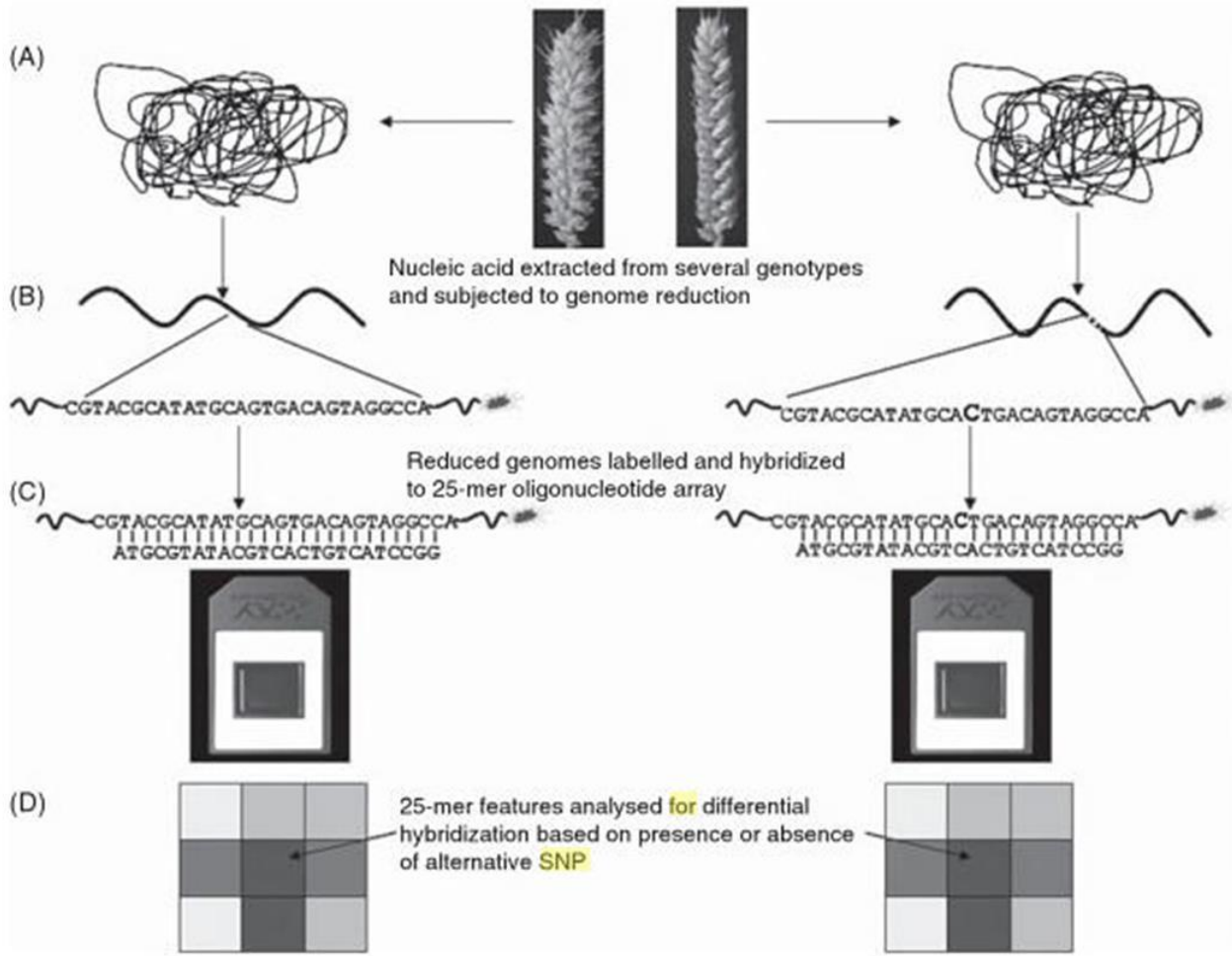
4) Starší zdroj: sekvenace STS markerů

Genotypování **známých** SNP

- Existují tři možné způsoby genotypování známých SNP v populaci jedinců:
  - Genotypování pomocí biočipů (detekce založená na hybridizaci)
  - Genotypování pomocí sekvenování (GBS – genotype by sequencing)
  - Genotypování pomocí STS markerů (starší způsob)

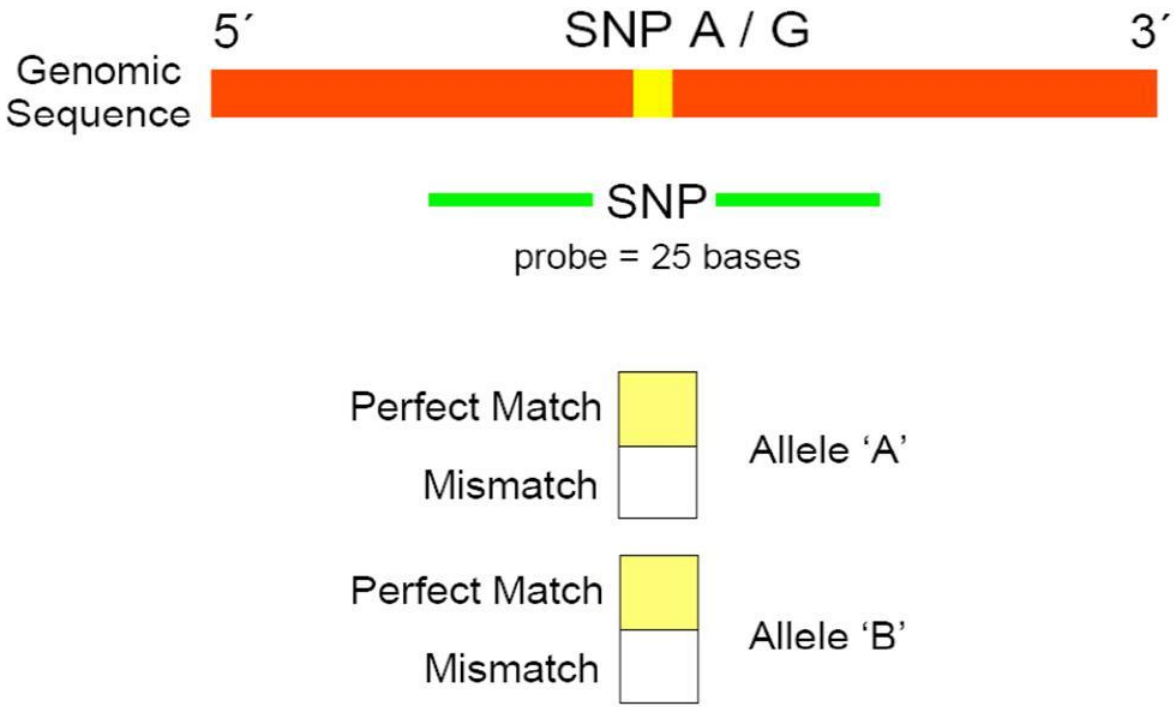
Schéma detekce SNPs u hospodářských plodin pomocí biočipů





Nekomplementární/
komplementární alela

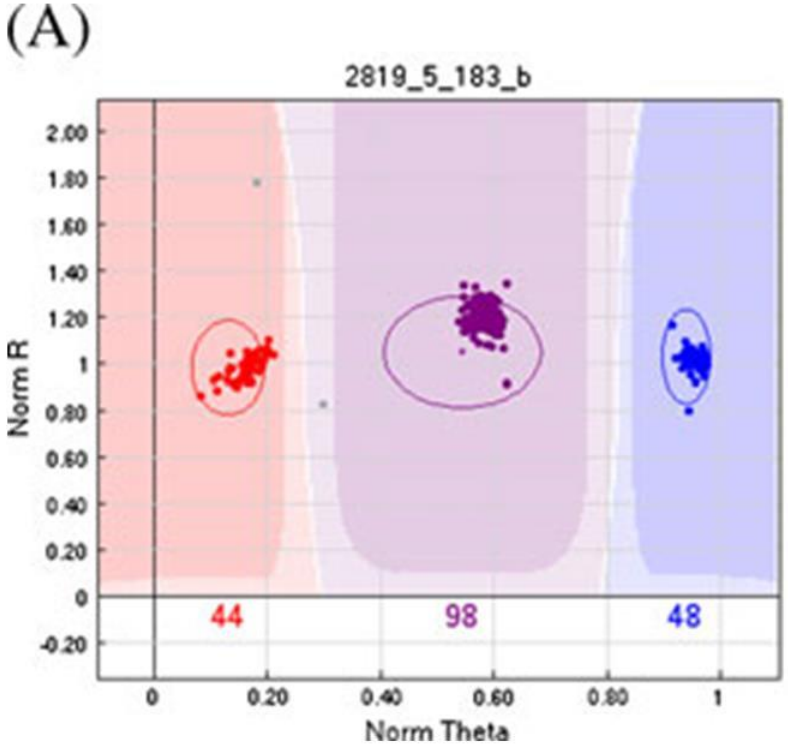
Detekce SNPs pomocí biočipů - detail



- V čipu se nachází sondy pro oba polymorfizmy SNP A i SNP G
- „Mismatch“ jsou kontrolní „nekomplentární“ sondy, které slouží jako negativní kontroly – pozadí
- Můžeme detekovat i heterozygoty! SNPs jsou kodominantní!

Detekce heterozygotů v SNPs u diploidních rostlin

- Detekujeme např. A/G SNP polymorfizmus
- Červená oblast (cluster) na obrázku představuje jedince, kteří jsou homozygoti v alele A, tj. AA (jedna tečka = jeden jedinec)
- Fialová oblast zahrnuje jedince, kteří jsou heterozygoti AG
- Modrá oblast pak jedince, kteří jsou homozygoti v G, tj. GG



- Detekce je obtížnější u polyploidních druhů, větší počet shluků, př. Autotetraploid – brambor = 5 shluků pro dvě alely (AAAA, AAAB, AABB, ABBB and BBBB)

Přehled biočipů pro SNP genotypování v plodinách

Table 1. Overview on large genotyping array development in major crop plants

Crop species	Array size*	Reference
Apple (*Malus domestica*)	8 K	Chagné *et al.* (2012)
Grape (*Vitis vinifera*)	9 K	Myles *et al.* (2011)
Maize (*Zea mays*)	60 K	Ganal *et al.* (2011)
Rice (*Oryza sativa*)	44 K	Zhao *et al.* (2011)
Rye (*Secale cereale*)	5 K	Hasenmeyer *et al.* (2011)
Sunflower (*Helianthus annuus*)	10 K	Bachlava *et al.* (2012)
Barley (*Hordeum vulgare*)	9 K	unpublished
Cherry (*Prunus spec.*)	6 K	unpublished *http://www.illumina.com/applications/agriculture.ilmn#ag_consortia*
Grape (*Vitis vinifera*)	20 K	in preparation *http://www.illumina.com/applications/agriculture.ilmn#ag_consortia*
Peach (*Prunus persica*)	8 K	unpublished *http://www.illumina.com/applications/agriculture.ilmn#ag_consortia*
Potato (*Solanum tuberosum*)	10 K	unpublished *http://solcap.msu.edu/*
Tomato (*Solanum lycopersicum*)	10 K	unpublished *http://solcap.msu.edu/*
Wheat (*Triticum aestivum*)	9 K	unpublished *http://www.triticeaecap.org/*
Oilseed rape (*Brassica napus*)	60 K	in preparation *http://www.illumina.com/applications/agriculture.ilmn#ag_consortia*
Wheat (*Triticum aestivum*)	90 K	in preparation *http://www.illumina.com/applications/agriculture.ilmn#ag_consortia*

* Number of features on the array. Note that this does not necessarily correspond to the number of SNPs that can be analysed (e.g. with the 60 K maize array, approximately 50000 functional SNPs can be analysed).

Genotypování SNPs pomocí sekvenování (GBS)

- Využívá NGS technologie pro **(re) sekvenaci** genomů jedinců analyzované populace (obvykle mapovací) čistě za účelem genotypování (přímá detekce SNPs u těchto jedinců)
- Není sekvenován celý genom (stále drahé pro většinu plodin) = vyžaduje pokryvnost cca 10 x genom
- Místo toho redukce pokryvnosti genomu sekvenování na 0,10 X pro mapovací populace (např. rýže)
- V případě GBS se na jeden „run“ sekvenátoru sekvenuje mnoho genomů najednou – každý označen jinými adapory.

Genotypování SNP pomocí sekvenování (GBS)

- Největším problémem zůstávají vysoké nároky na výpočetní techniku, analýzu dat a také uchování obrovského množství dat získaných NG sekvenací rozsáhlých mapovacích populací
- Zatím není jasné která technologie genotypování převládne, zda GBS nebo detekce SNP pomocí micorarray

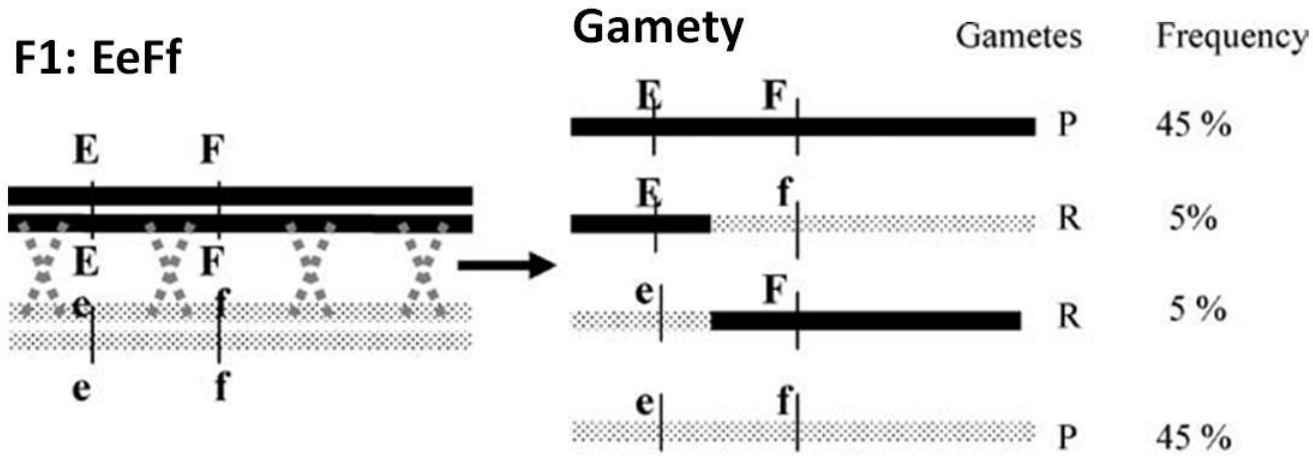
Tvorba genetických (rekombinačních) map = hledání asociace mezi polymorními DNA markery a vlastnostmi plodin

Tvorba genetických (rekombinačních) map

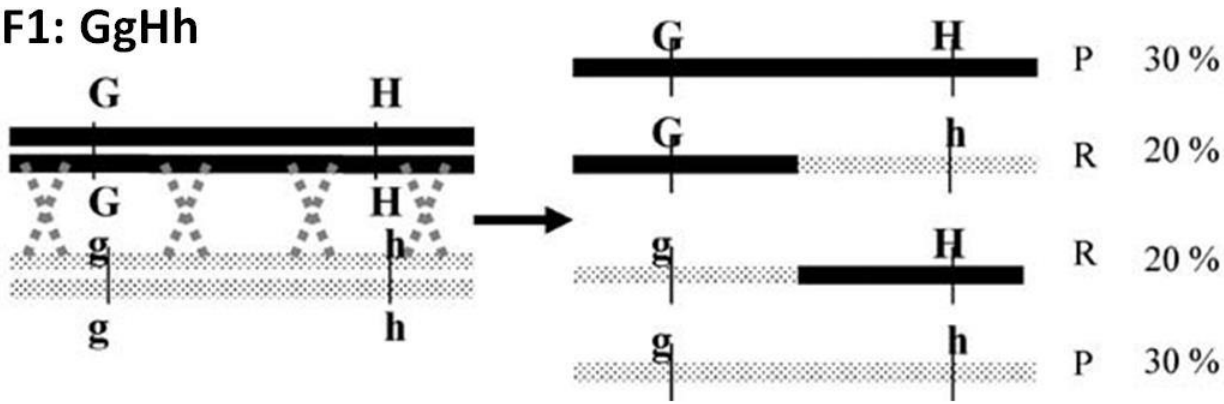
- Tvorba genetických map principiálně vychází ze segregace (oddělování) markerů a major genů/QTL lokusů v důsledku crossing-overu mezi chromozómy
- Ke crosssing-overu (rekombinacím) chromozómů dohází během redukčního dělení – meióze
- Segregaci markerů a genů/QTL lokusů je proto možno analyzovat v generaci potomků získaných pohlavní rozmnožováním

Crossing-over mezi dvěma geny/markery

A) **P: EEFF x eeff** **F1: EeFf**



B) **P: GGHH x gghh** **F1: GgHh**



Geny a markery, které se nachází v těsné blízkosti jsou přenášeny do další generace častěji než geny a markery, mezi kterými je větší vzdálenost

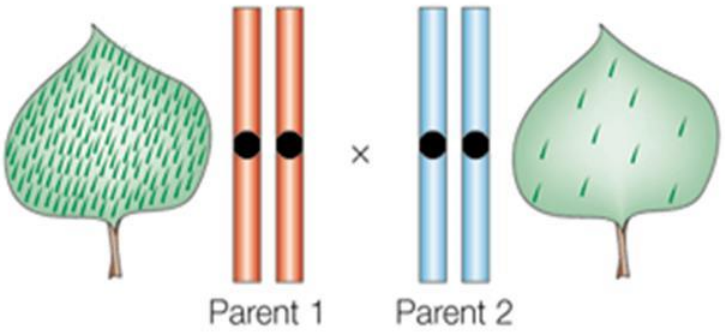
Platí že…

- Informaci o frekvenci (procentu) výskytu rekombinantních genotypů v populaci potomků lze využít pro odvození genetické vzdálenosti mezi markery.
- Čím větší je frekvence rekombinací mezi dvěma markery, tím větší vzdálenost je mezi nimi na chromozómu a obráceně
- Dva markery, které se na chromozómu nacházejí ve vzdálenosti větší, nebo rovno 50% (cM) nejsou ve vazbě = volná kombinovatelnost vloh
- Analýzou segregace markerů v populaci určíme pořadí markerů na chromozómu a genetické vzdálenosti mezi jednotlivými markery – sestavíme mapu

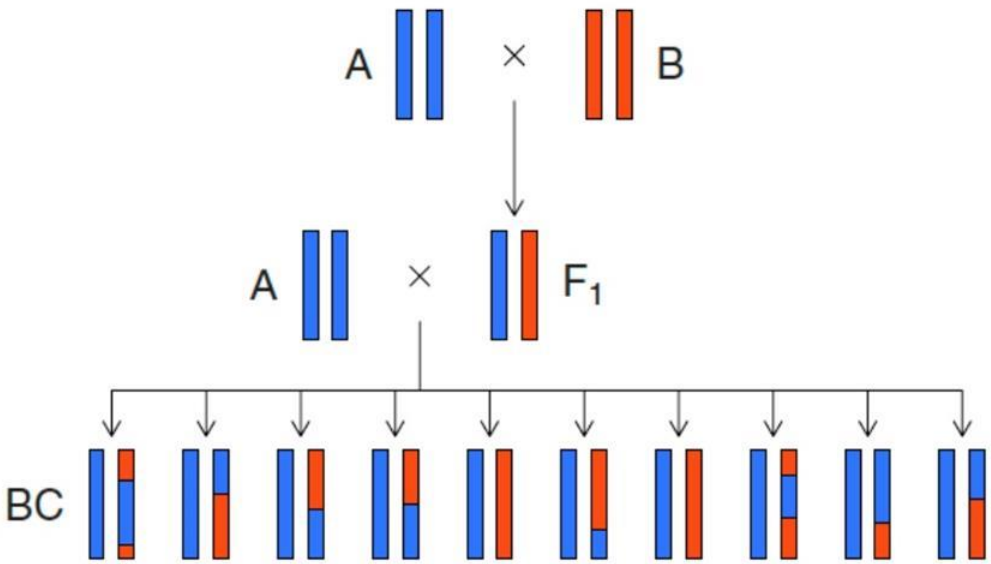
Čtyři základní kroky pro vytvoření genetické mapy

1) Příprava mapovací populace
2) Identifikace polymorfních DNA markerů a genotypování mapovací populace
3) Analýza vazby mezi markery (přítomnost, absence vazby, vzdálenost mezi markery, relativní pořadí markerů)
4) Mapování genů/QTL lokusů

1. Příprava mapovací populace. Na příkladu back-cross populace u samosprašných rostlin



Příklad jedinců kontrastních ve znaku



Křížení rodičů (homozygotní linie kontrastní ve sledovaném znaku)

Zpětné křížení F_1 hybridů s s rodičem A

BC hybridi vykazují rekombinace pouze v rámci jedné sady chromozómů

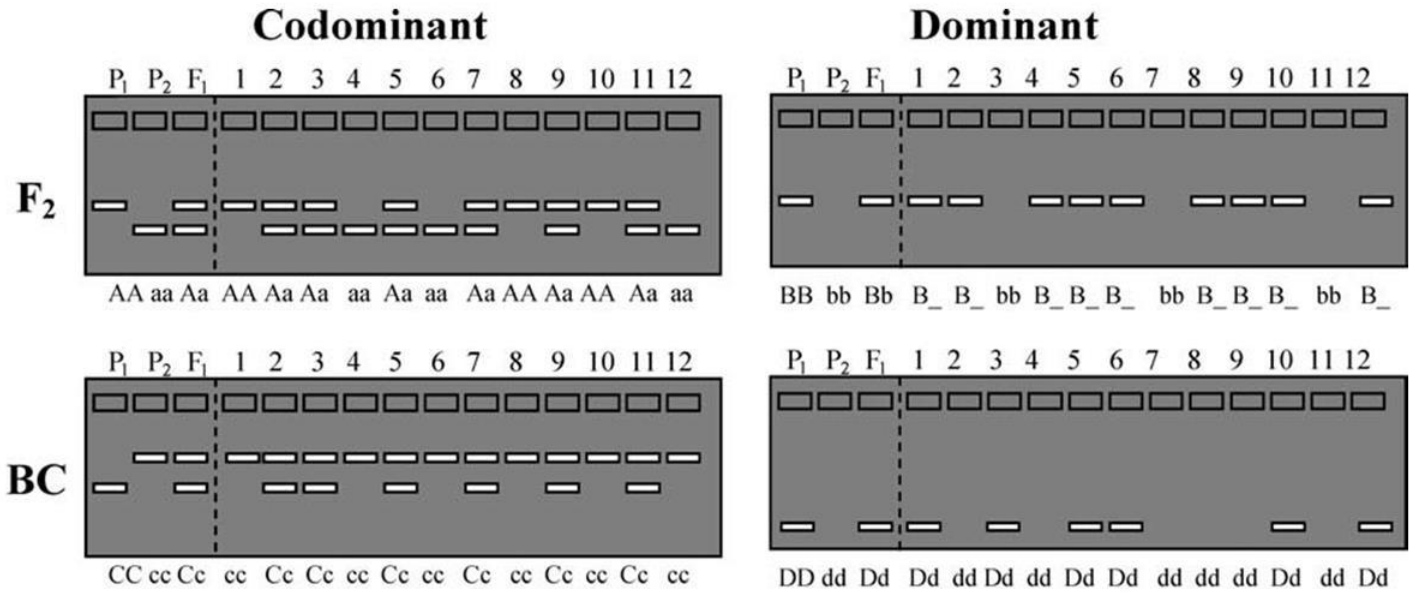
2. Identifikace polymorfních DNA markerů a genotypování mapovací populace

- Základem je existence dostatečných množství DNA polymorfizmů mezi rodiči populace
- cizosprašné rostliny = vysoký stupeň polymorfizmu
- U samosprašných rostlin je obvykle nutný výběr rodičů populace na základě vyhodnocení míry polymorfizmu pomocí DNA markerů
- **Polymorfní (informativní) DNA markery jsou identifikovány genotypováním rodičů pomocí souboru DNA markerů**

Možné výstupy genotypování mapovacích populací podle typu populace a polymorfního DNA markeru:

F2 populace

BC populace



Očekávané segregační (štěpné) poměry DNA markerů v populacích

Population type	Codominant markers	Dominant markers
F_2	1: 2:1 (AA:Aa:aa)	3:1 (B_:bb)
Backcross	1:1 (Cc:cc)	1:1 (Dd:dd)
Recombinant inbred or doubled haploid	1:1 (EE: ee)	1:1 (FF:ff)

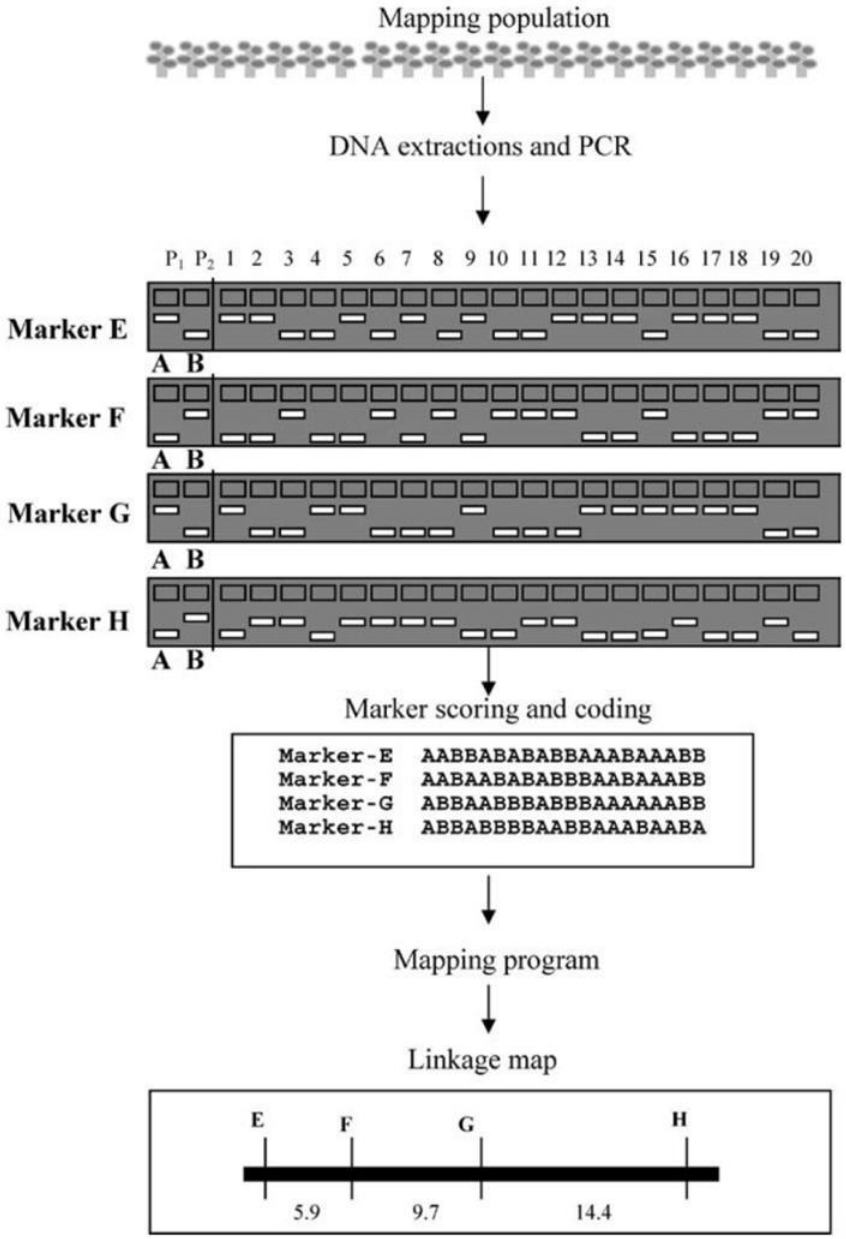
Verifikují se chí kvadrát testem

3. Analýza vazby mezi markery

Zahrnuje:

- kódování dat pro každý DNA marker a pro každého jedince v populaci
- Sestavení vazbových skupin markerů pomocí metody LOD score
- Určení genetické vzdálenosti mezi markery = konverze rekombinačních frekvencí na genetickou vzdálenost v cM pomocí mapovacích funkcí

Výstup mapovacího programu: Dílčí vazbová skupina



První rodič je kódován jako „A“

Druhý rodič je kódován jako „B“

Kódování umožní vazbovou analýzu pomocí programu

Kolik jedinců s rekombinací mezi markery EF se nachází v mapovací populaci?

Výstup mapovacího programu: genetická mapa

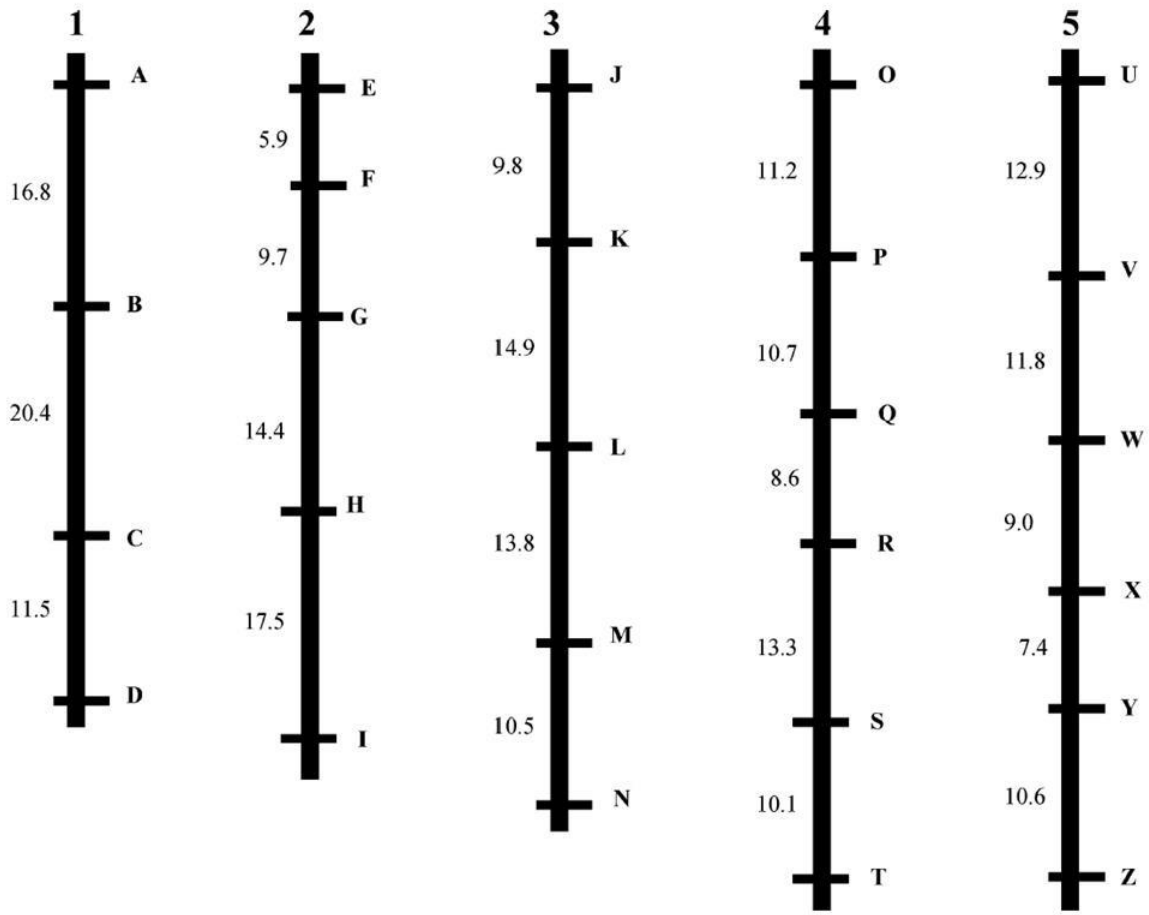


Figure 7. Hypothetical 'framework' linkage map of five chromosomes (represented by linkage groups) and 26 markers. Ideally, a framework map should consist of evenly spaced markers for subsequent QTL analysis. If possible, the framework map should also consist of anchor markers that are present in several maps, so that they can be used to compare regions between maps.

4. mapování major genů/QTL lokusů

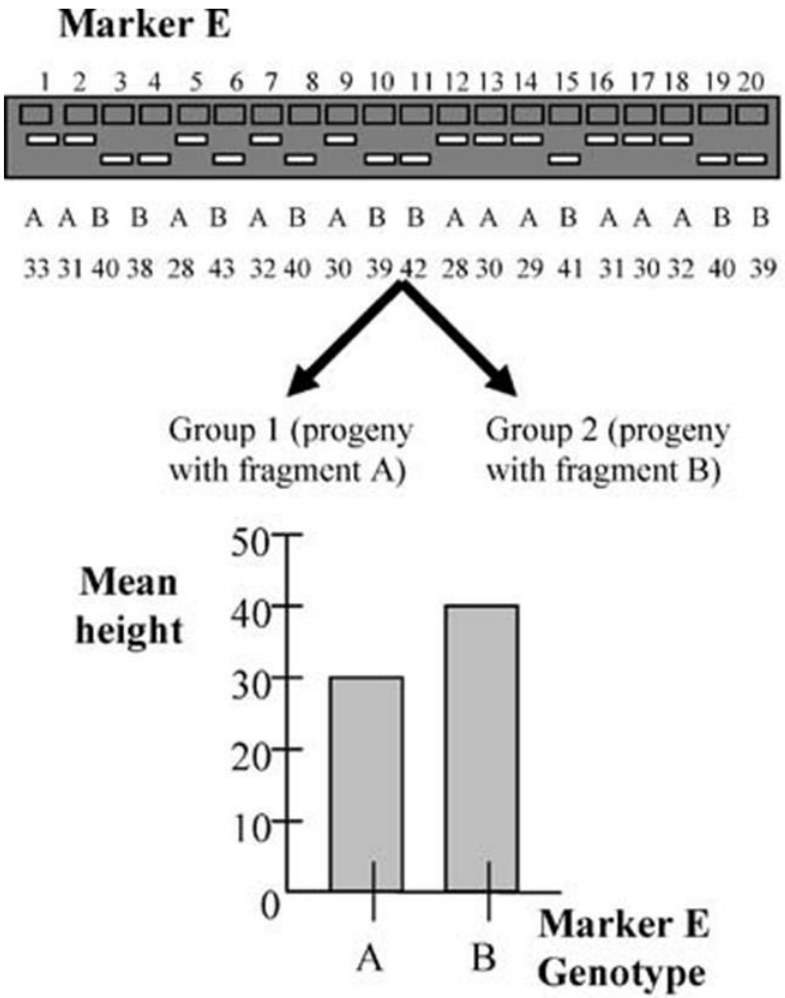
- V tomto kroku je vazbová mapa využita k určení chromozomální lokalizace (příslušnosti k vazbové skupině) genů řídících sledované znaky:
  - Major geny - kvalitativní znaky
  - QTL – lokusy kvantitativních vlastností

- Zároveň jsou **identifikovány DNA markery, které se nachází ve vazbě s QTL, major genem**

Využítí vazbových map pro identifikaci QTL: QTL analýza

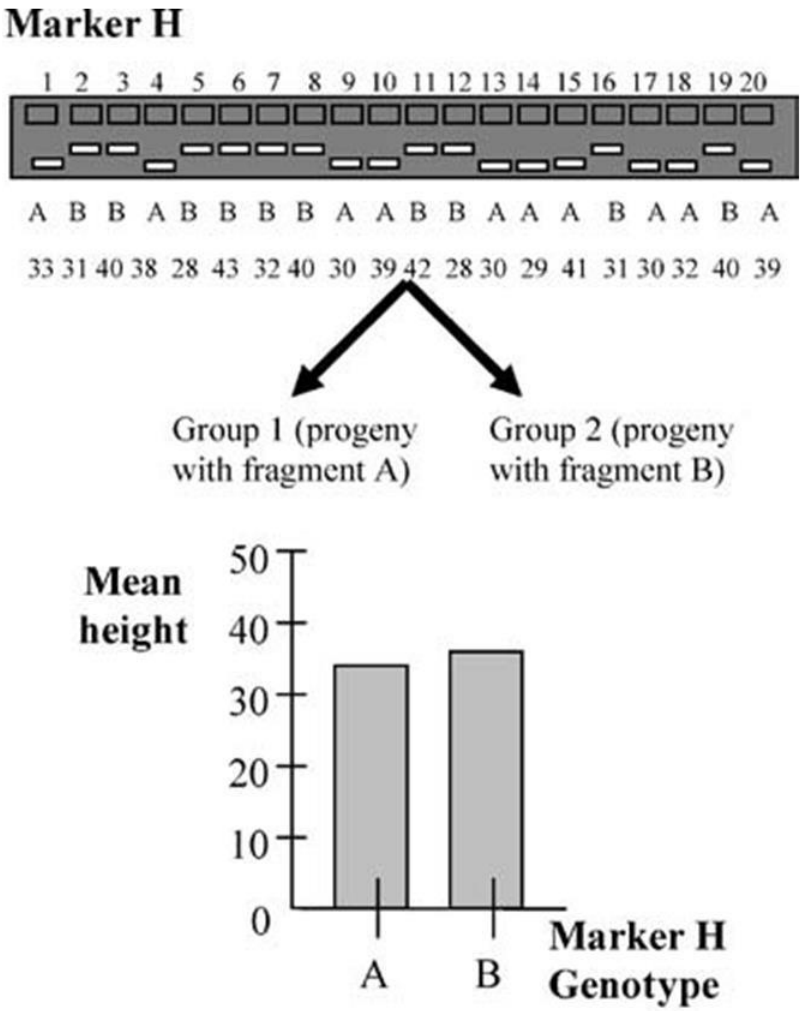
- Je založena na principu detekce asociace mezi fenotypem sledovaného znaku a sestavou „genotypem“ DNA markeru
- DNA markery jsou využity pro rozdělení mapovací populace do genotypových skupin (podle typu sestavy markeru) a posléze je statisticky vyhodnocen rozdíl v průměrné hodnotě sledované vlastnosti mezi skupinami
- Pokud existuje statisticky významný rozdíl v průměrné hodnotě vlastnosti mezi skupinami (2-3 skupiny – Proč?) pak je daný marker ve vazbě s QTL pro sledovaný znak – viz další slide

Princip mapování QTL



P value <0.0001 = significant

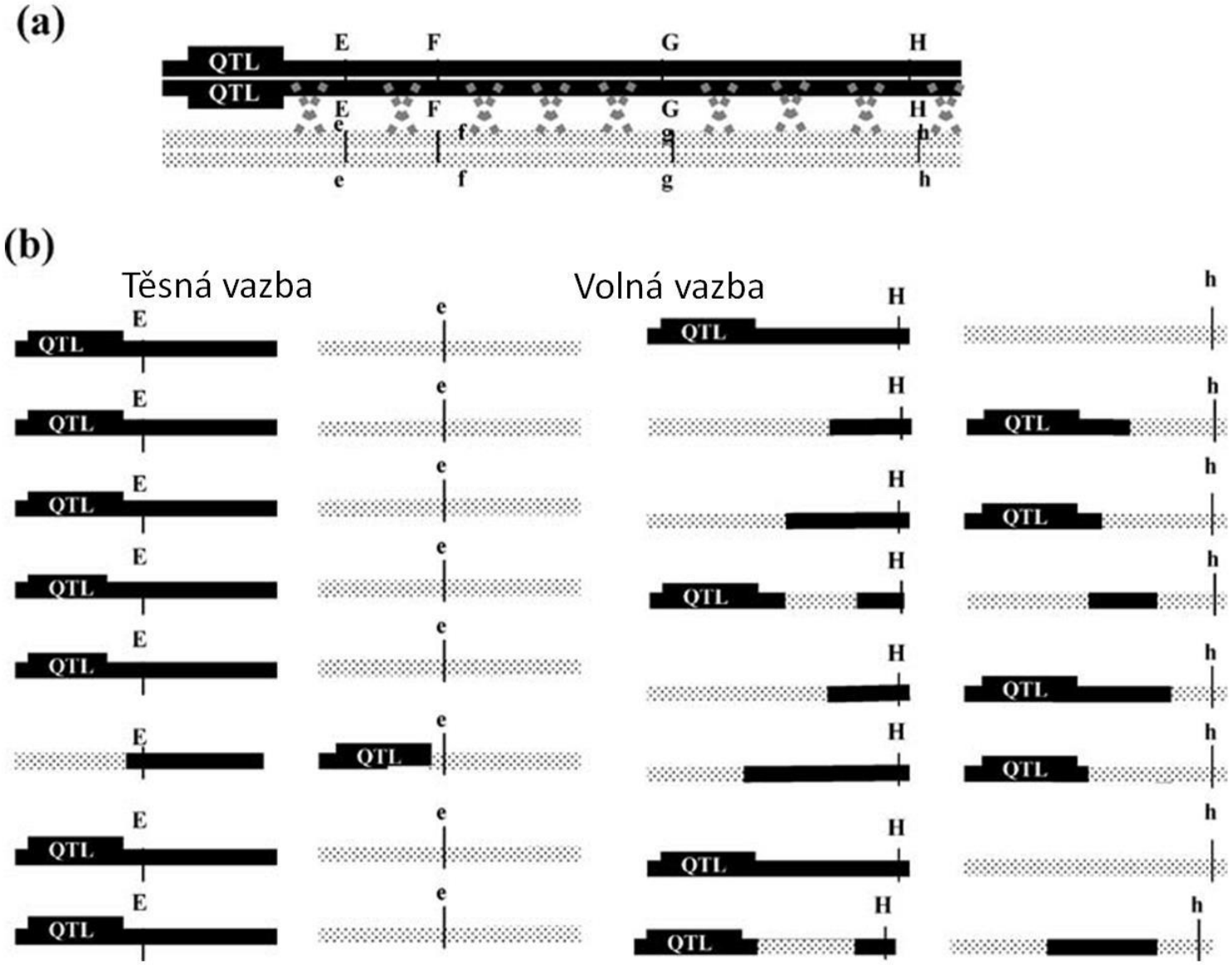
Conclusion: marker is linked to a QTL



P value 0.5701 = not significant

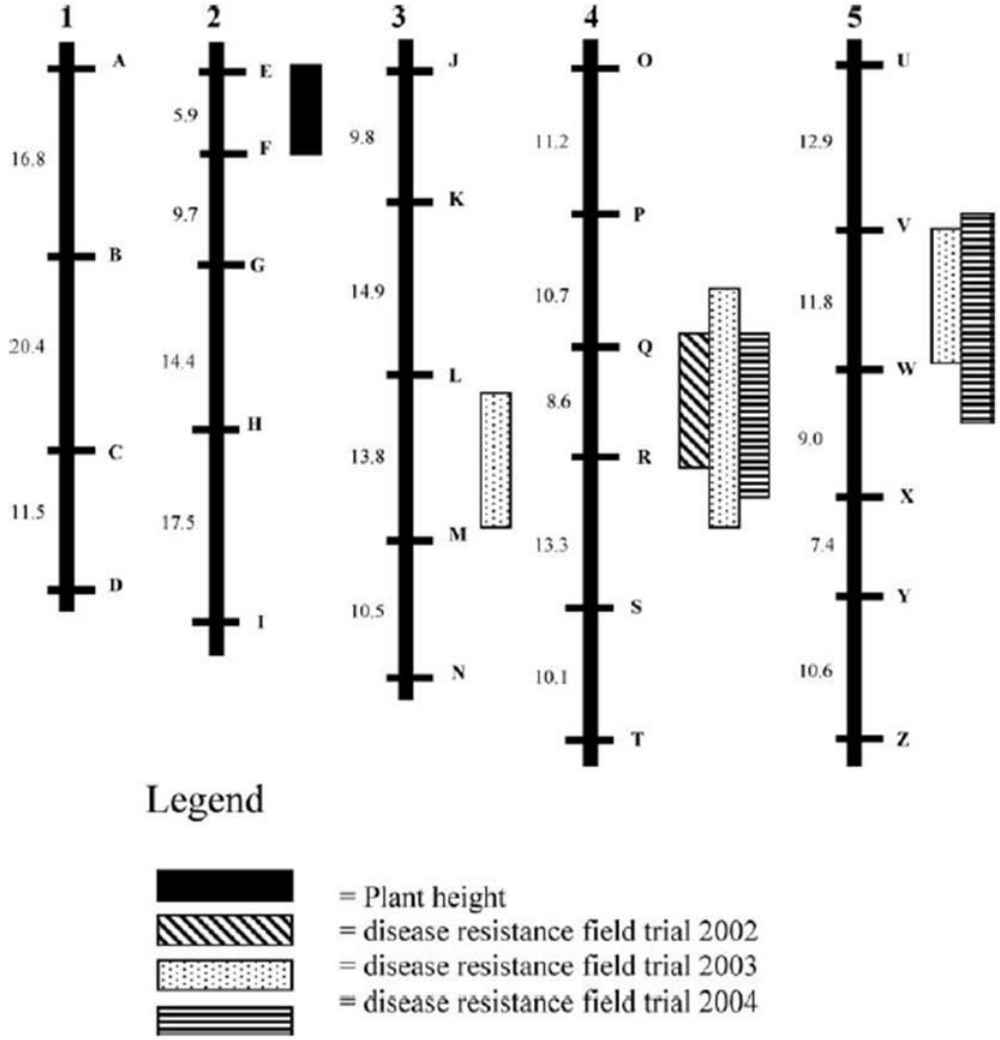
Conclusion: marker is unlinked to a QTL

Těsná a volná vazba mezi markerem a QTL



V potomstvu dominantní alela téměř výhradně v kombinaci s QTL a recesivní téměř výhradně v kombinaci s absencí QTL

Příklad výstupu mapování QTL



- Výška rostliny = jedno majoritní QTL,
- Odolnost k chorobě = dvě majoritní QTL na chromozómu 4 a 5 a jedno minoritní na chromozomu 3

Výstupy mapování major genů/QTL

- Získání DNA markerů pro proces **markery asistovaná selekce (MAS)**
- Marker vhodný pro MAS je:
  - co nejblíže selektovanému major genu/QTL (< 5cM)
  - případně se nachází přímo v markerovaném genu/QTL (100% kosegregace s DNA markerem) – pokud je gen/QTL získán pozičním klonováním
  - Polymorfní a informativní v různých souborech genotypů, v různých populacích
- Klonovaný major/minor gen – lze využít pro přípravu transgenních rostlin

Markery asistovaná selekce (MAS)

Využití DNA markerů ve šlechtění hospodářských plodin s cílem zefektivnění šlechtitelského procesu

Výhody MAS oproti klasické selekci

- Uplatní se zejména u znaků, které se obtížně fenotypují: šetří čas, peníze a úsilí. Př. rezistence k cystotvorným háďátkům u obilnin, kvalita zrna - nákladná na vyhodnocení

- Selekce ve fázi klíčních rostlin:
  - Uplatní se například pro znaky jako kvalita zrna, = nemusím čekat a pěstovat řadu rostlin až do zralosti, vyselektuji pomocí MAS a pěstuji pouze ty, které budou mít kvalitní zrna, kvalita zrna je navíc drahá na vyhodnocení fenotypově

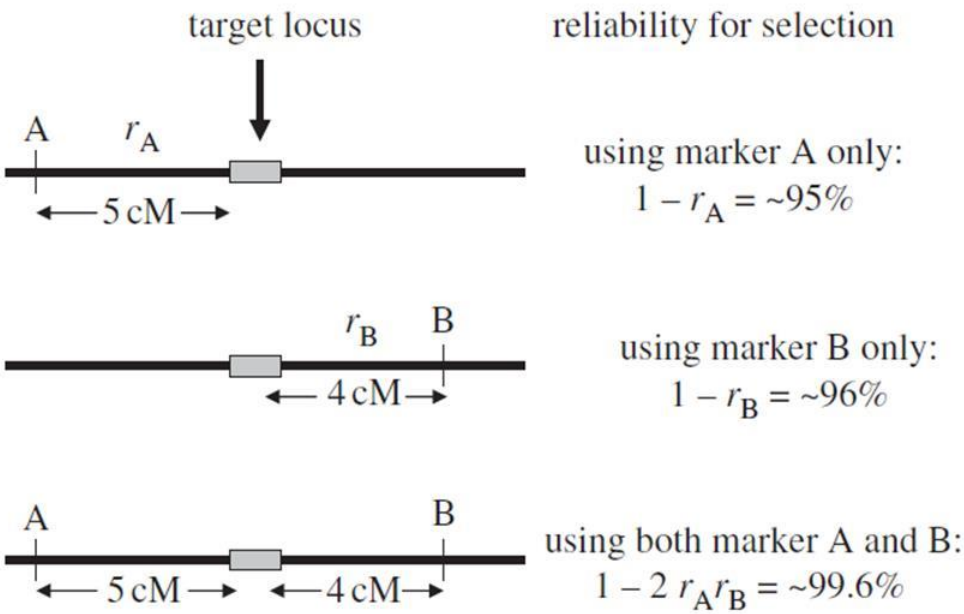
- Zvýšená spolehlivost selekce:
  - Eliminuje vliv vnějšího prostředí (především u kvantitativních znaků, které se obtížně fenotypují: například polní rezistence vůči patogenu)
  - Rozlišíme homozygoty a heterozygoty

- Zpřesněná a zefektivněná selekce specifických genotypů-**jednotlivých rostlin**:
  - Zrychlení šlechtitelského procesu, respektive tvorby odrůd.

Kroky k MAS

1. Mapování major genů/QTL a získání spolehlivých markerů pro MAS:

2. Spolehlivý marker pro MAS má být co nejblíže cílovému lokusu (major genu/QTL), ideálně přímo v genu/ QTL



3. Použití dvou lemujících markerů zvyšuje spolehlivost selekce – Obrázek, předpokládáme 0 interferenci crossing-overů. Dvojitý crosing-over=velmi nízká pravděpodobnost

Použití dvou markerů však zároveň zvyšuje náklady a čas

Uplatnění MAS

- MAS v převodném křížení
- MAS v pyramidování genů/QTL
- MAS v charakterizaci šlechtitelského materiálu

MAS v převodném křížení: Příklad klasického převodného křížení

P_1 x P_2

- High yielding
- Susceptible for 1 trait
- Called recurrent parent (RP)

Elite cultivar → Donor

Desirable trait

e.g. disease resistance

P_1 x F_1

P_1 x BC1

Discard ~50% BC1

Visually select BC1 progeny that resemble RP

P_1 x BC2

Repeat process until BC6

P_1 x BC3

P_1 x BC4

P_1 x BC5

P_1 x BC6

⊗

BC6F2

Pozor! Podíl genomu rekurentího rodiče na genomu hybridů v jednotlivých generacích BC představuje průměrnou hodnotu. V každé BC generaci se nacházejí jedinci, kteří mají větší nebo menší podíl rekurentího genomu než je průměr.

Generation	Recurrent parent genome (%)
BC1	75.0
BC2	87.5
BC3	93.8
BC4	96.9
BC5	98.4
BC6	99.2

Recurrent parent genome recovered

Additional backcrosses may be required due to linkage drag

Vazebné táhnutí

MAS v převodném křížení:

1. Využití markerů pro selekci cílového major genu/QTL = foreground selection
2. Využití markerů pro selekci rekombinantů mezi cílovým genem/QTL a lemujícími markery= recombinant selection
3. Využití markerů pro selekci vůči genomu donora = background selection
4. Ve šlechtitelském programu lze samozřejmě využívat kombinace všech tří typů MAS

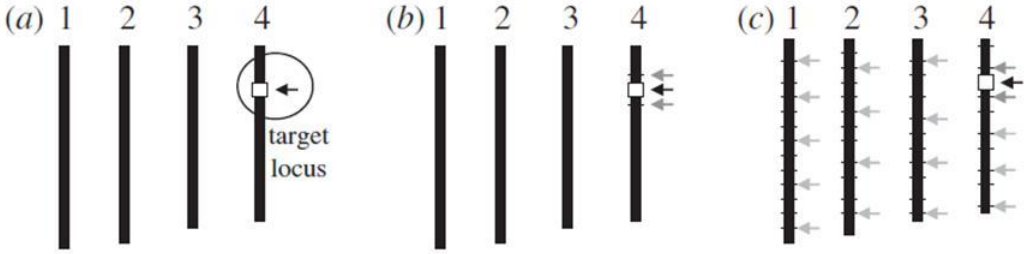
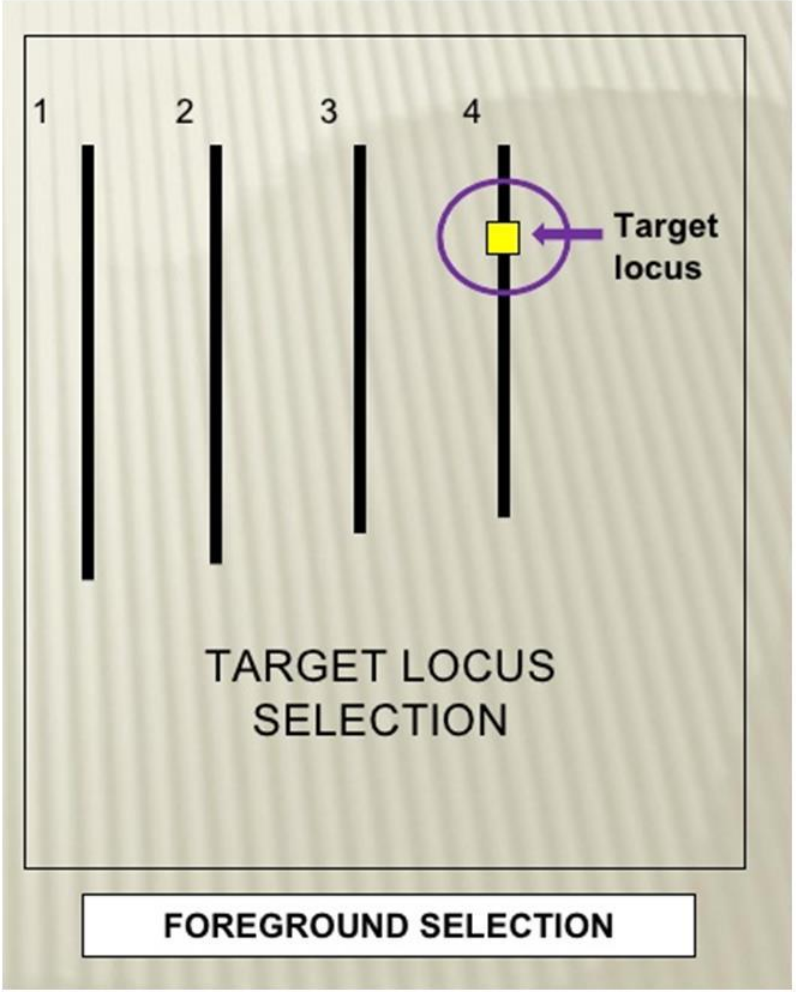


Figure 3. Levels of selection during marker-assisted backcrossing. A hypothetical target locus is indicated on chromosome 4. (*a*) Foreground selection, (*b*) recombinant selection and (*c*) background selection.

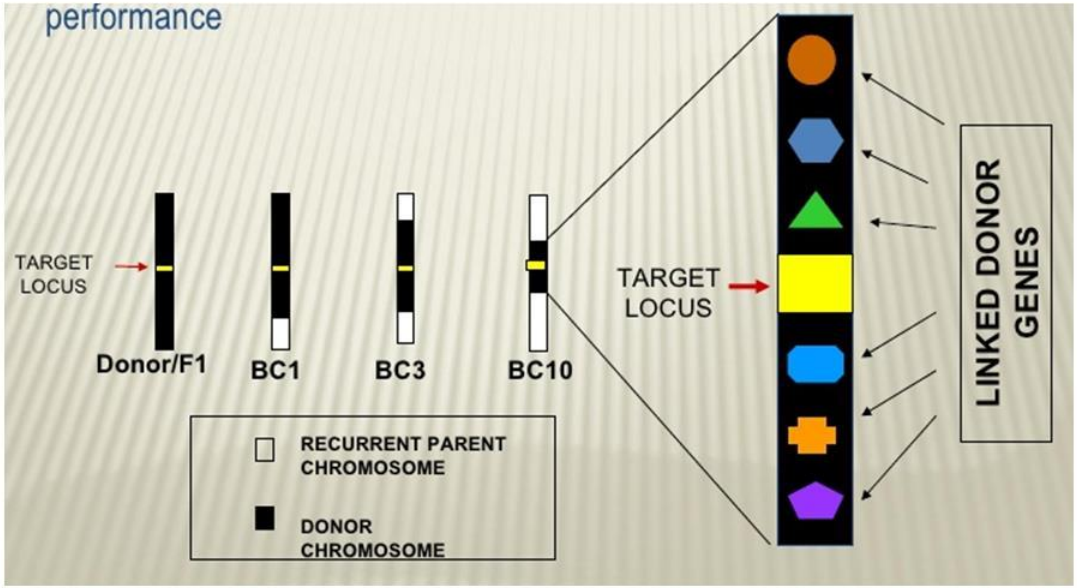
1. „Foreground“ selekce



- Využití markerů pro selekci cílového major genu nebo QTL.
- Výhodné zejména v případě vlastností, které se špatně fenotypují.
- Při selekci recesivně determinovaného znaku. Ten se projeví pouze u recesivního homozygota. Selekce genotypů s recesivními alelami je obtížná pomocí konvenčních metod.

2. Rekombinační selekce a koncept „linkage drag“

- Cílem rekombinační selekce je redukce velikosti chromosomálního segmentu donora, který obsahuje cílový lokus, respektive velikosti introgrese
- Proč? Rychlost zmenšování velikosti tohoto donorového fragmentu je pomalejší než u chromozomálních regionů donora, které nejsou ve vazbě s cílovým lokusem a navíc často nese řadu nežádoucích genů donora, které jsou ve vazbě s cílovým lokusem (nezapomenout, že donor je např. planý druh, který má řadu negativních vlastností – např. nízkou produktivitu, kvalitu, pozdnost…) = **„linkage drag“**
- Při použití klasických selekčních metod může donorový segment zůstat značně velký dokonce i za použítí velkého počtu BC generací (více než 10) = obrázek



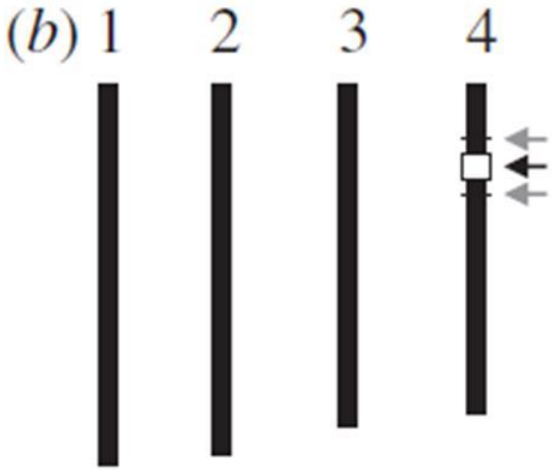
Rekombinační selekce = minimalizace „linkage drag“

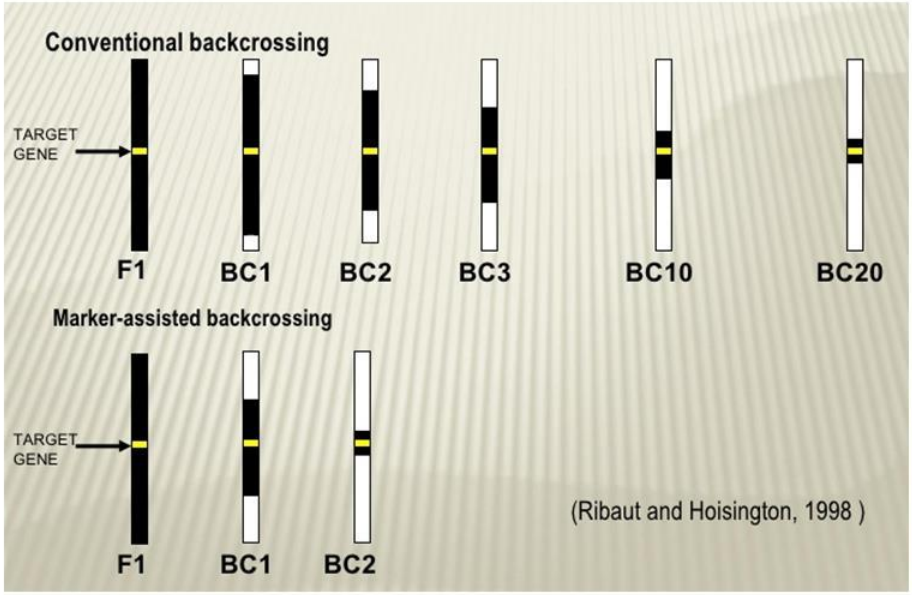
- Při použití markerů, které lemují cílový gen (např. vzdálenost < 5cM na každé straně) lze problém „linkage drag“ minimalizovat (obrázek (b))

- Selektovány jsou genotypy, které vykazují rekombinaci mezi cílovým lokusem a lemujícím markerem

- Obvykle je zapotřebí minimálně dvou BC generací k vyselektování genotypů, které vykazují rekombinaci cílového lokusu s oběma lemujícími markery a dostatečně velkých populací

- Rychlost minimalizace „linkage drag“ je však nesrovnatelně větší než v případě konvenčního převodného křížení (obrázek dole)



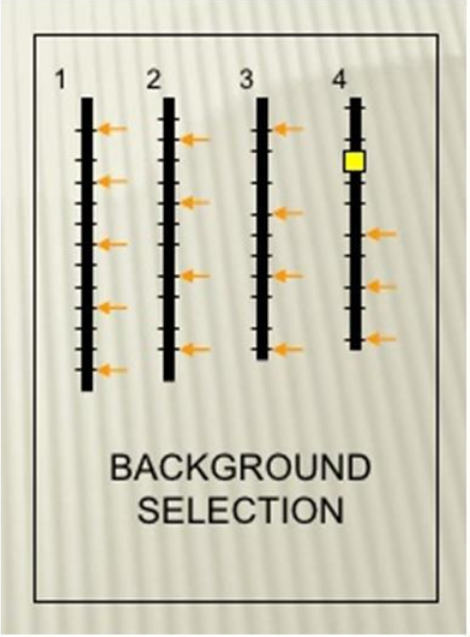


3. „Bacground“ selekce

- Znovu připomínáme, že podíl genomu rekurentího rodiče na genomu hybridů v jednotlivých generacích BC představuje průměrnou hodnotu. V každé BC generaci se nacházejí jedinci s cílovým lokusem, kteří mají větší nebo menší podíl rekurentího genomu než je průměr.

Generation	Recurrent parent genome (%)
BC1	75.0
BC2	87.5
BC3	93.8
BC4	96.9
BC5	98.4
BC6	99.2

- Pomocí markerů, které nejsou ve vazbě s cílovým lokusem a jsou distribuovány na všech chromozómech rekurentího rodiče (žluté šipky – spodní obrázek) lze v BC generacích vybrat jedince s největším podílem rekurentního genomu a ty dále křížit s rekurentním rodičem



- Touto cestou lze „ušetřit“ 2 až dokonce 4 generace zpětného křížení, tj. oproti obvyklé BC6 lze již v BC2 získat jedince s genotypem rekurentního rodiče a cílovým lokusem

- „Bacground“ selekce je obvykle kombinována s rekombinační selekcí

Starší tabulka příkladů využití MAS v převodném křížení u obilnin

species	trait(s)	gene/QTLs	foreground selection	background selection	reference
barley	barley yellow dwarf virus	*Yd2*	STS	not performed	Jefferies *et al.* (2003)
barley	leaf rust	*Rphq6*	AFLP	AFLP	van Berloo *et al.* (2001)
barley	stripe rust	QTLs on 4H and 5H	RFLP	not performed	Toojinda *et al.* (1998)
barley	yield	QTLs on 2HL and 3HL	RFLP	RFLP	Schmierer *et al.* (2004)
maize	corn borer resistance	QTLs on chromosomes 7, 9 and 10	RFLP	RFLP	Willcox *et al.* (2002)
maize	earliness and yield	QTLs on chromosomes 5, 8 and 10	RFLP	RFLP	Bouchez *et al.* (2002)
rice	bacterial blight	*Xa21*	STS[a]	RFLP	Chen *et al.* (2000)
rice	bacterial blight	*Xa21*	STS[a]	AFLP	Chen *et al.* (2001)
rice	bacterial blight	*xa5, xa13* and *Xa21*	STS, CAPS	not performed	Sanchez *et al.* (2000)
rice	bacterial blight	*xa5, xa13 and Xa21*	STS	not performed	Singh *et al.* (2001)
rice	bacterial blight + quality	*xa13, Xa21*	STS and SSR	AFLP	Joseph *et al.* (2004)
rice	blast	*Pi1*	SSR	ISSR[b]	Liu *et al.* (2003)
rice	deep roots	QTLs on chromosomes 1, 2, 7 and 9	RFLP and SSR	SSR	Shen *et al.* (2001)
rice	quality	waxy	RFLP[a]	AFLP	Zhou *et al.* (2003*a*)
rice	root traits and aroma	QTLs on chromosomes 2, 7, 8, 9 and 11	RFLP and SSR	RFLP and SSR	Steele *et al.* (2006)
rice	submergence tolerance	*Sub1* QTL	phenotyping and SSR[a]	SSR	Mackill *et al.* (2006)
rice	submergence tolerance, disease resistance, quality	*Subchr9* QTL, Xa21, Bph and blast QTLs and quality loci	SSR and STS	not performed	Toojinda *et al.* (2005)
wheat	powdery mildew	22 *Pm* genes	phenotyping	AFLP	Zhou *et al.* (2005)

[a] Indicates recombinant selection performed to minimize linkage drag around target locus.
[b] ISSR and inter SSRs.

MAS v pyramidování genů/QTL

- Pyramidování představuje kombinaci více genů/QTL v jednom genotypu.
- Zdrojem genů/QTL pro pyramidování je 2 a více různých rodičů
- MAS má velké uplatnění v pyramidování genů, protože současné fenotypování každé rostliny z potomstva křížení na několik znaků je obtížné a prakticky nemožné např. u znaků vyhodnocených pomocí metod destruktivních pro rostlinu – př. odolnost k chorobám.
- Klasickým příkladem je pyramidování různých major genů rezistence vůči jednomu nebo více patogenům. Pyramidování major genů rezistence může poskytnout stabilnější rezistenci vůči patogenům

Starší příklady pyramidování genů u plodin

Table 1. Selected examples of MAS based gene pyramiding for important traits in major crops.

Crop	Trait	Pyramided genes	Reference
Rice	Blight resistance	*Xa4,xa5,xa13,Xa21*	Huang et al., 1997, Singh et al., 2001, Narayanan et al., 2002
	Blast resistance	*Pi(2)t,Piz5,Pi(t)a*	Hittalmani et al., 2000
	Gallmidge resistance	*Gm1,Gm4*	Kumaravadivel et al., 2006
Wheat	Leaf rust resistance	*Lr41, Lr42, Lr43*	Cox et al., 1994
	Powdery mildew resistance	*Pm-1, Pm-2*	Liu et al., 2000
Cotton	Insect pest resistance	*Cry 1Ac, Cry 2Ac*	Jackson et al., 2003, Gahan et al., 2005
Pea	Nodulation ability	*Sym9, Sym10*	Schneider et al., 2002
Barley	Yellow mosaic virus resistance	*rym4, rym5, rym9, rym11*	Werner et al., 2005
Soybean	Soybean mosaic virus resistance	*Rsv*1, *Rsv*3, *Rsv*4	Zhu et al., 2006

Starší příklady pyramidování genů/QTL u obilnin

Table 2. Examples of gene or QTL pyramiding in cereals.

species	trait(s)	genes from parent 1	genes from parent 2	selection stage	DNA marker(s) used	reference
barley	barley yellow mosaic virus	*rym1*	*rym5*	F_2	RFLP, CAPS	Okada *et al.* (2004)
barley	barley yellow mosaic virus	*rym4, rym9, rym11*	*rym4, rym9, rym11*	F_1-derived doubled haploids	RAPD, SSR	Werner *et al.* (2005)
barley	stripe rust	*Rspx* *Rspx*	QTLs 4, 7 QTL 5	F_1-derived doubled haploids	SSR	Castro *et al.* (2003)
rice	bacterial blight	*xa5, xa13*	*Xa4, Xa21*	F_2	RFLP, STS	Huang *et al.* (1997)
rice	bacterial blight, yellow stem borer, sheath blight	*Xa21, Bt*	*RC7* chitinase gene, *Bt*	F_2	STS	Datta *et al.* (2002)
rice	blast disease	*Pi1, Piz-5*	*Pi1, Pita*	F_2	RFLP, STS	Hittalmani *et al.* (2000)
rice	brown plant hopper	*Bph1*	*Bph2*	F_4	STS	Sharma *et al.* (2004)
rice	insect resistance and bacterial blight	*Xa21*	*Bt*	F_2	STS	Jiang *et al.* (2004)
wheat	powdery mildew	*Pm2*	*Pm4a*	F_2	RFLP	Liu *et al.* (2000)

MAS v charakterizaci šlechtelského materiálu

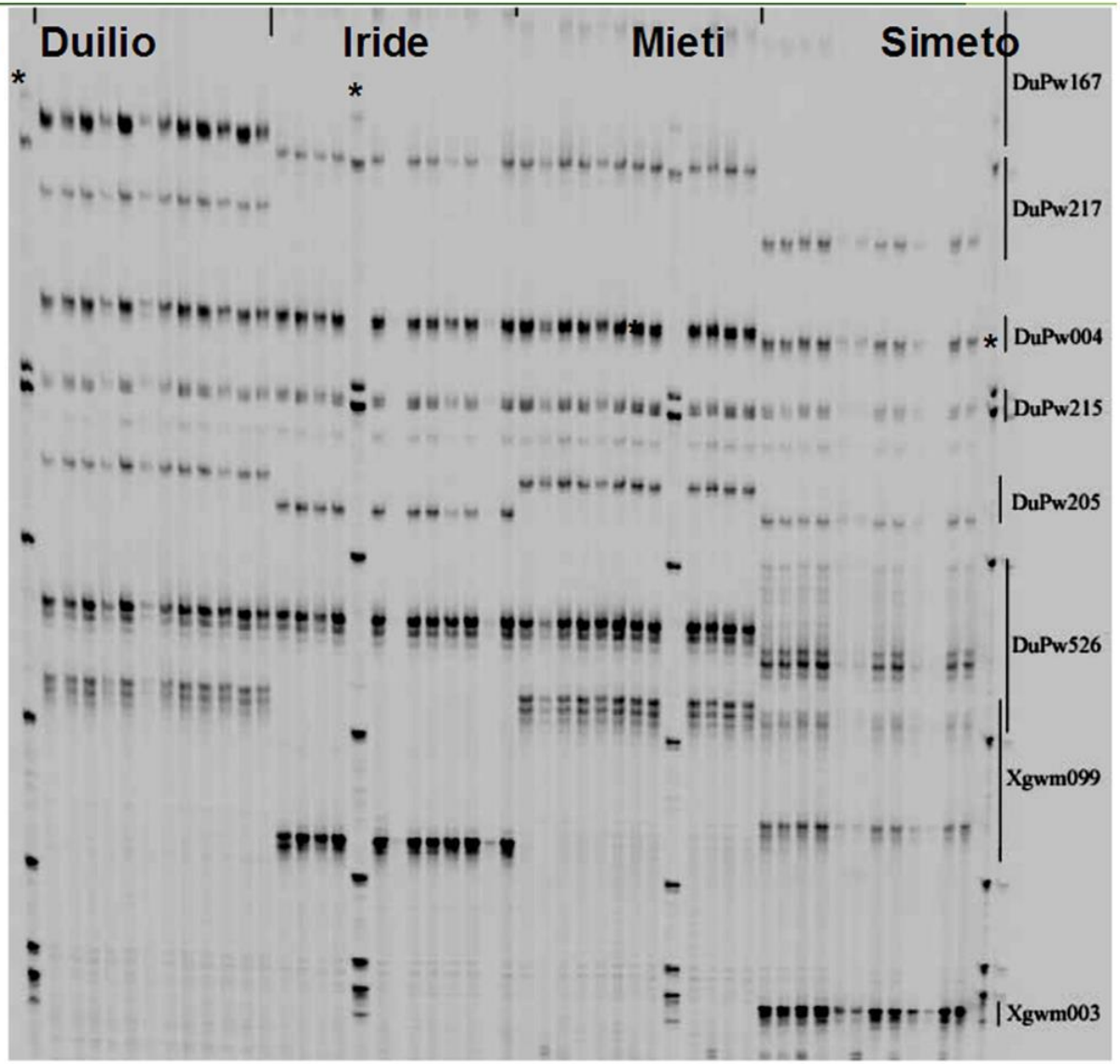
1. Ověření identity kultivaru a jeho čistoty
2. Ověření genetické diverzity a selekce rodičů
3. Studie heteróze

Ověření identity kultivaru a jeho čistoty

Sample of SSR results for wheat:

- **8 markers shown**
- **Multiplexed PCR protocol**
- **Visualization system: LiCor**

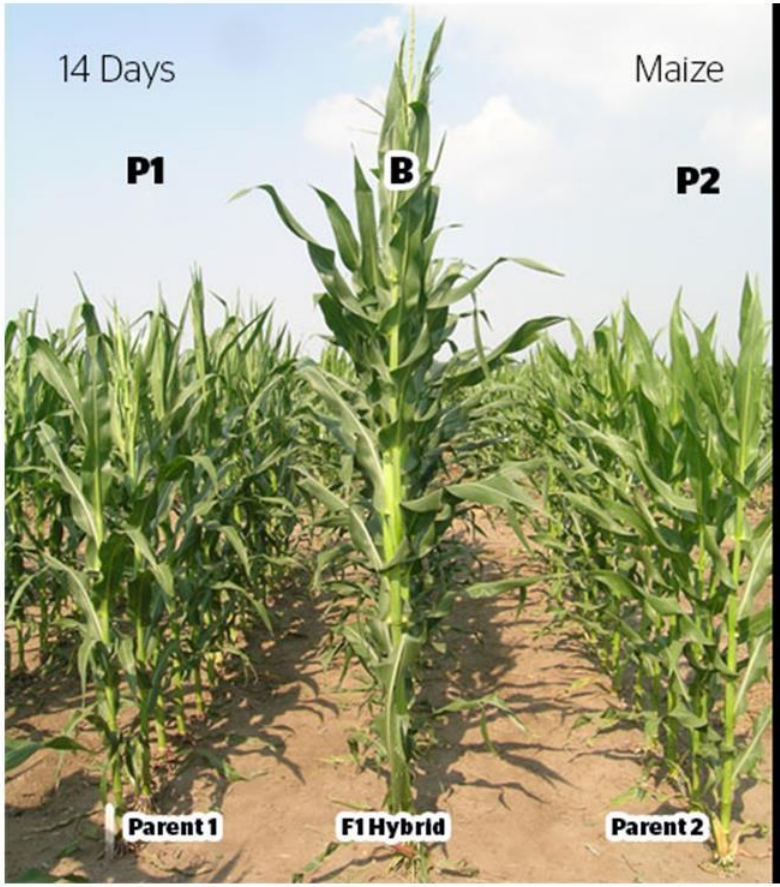
**indicates lane for size marker*



Ověření genetické diverzity a selekce rodičů

- Úspěšnost šlechtitelských programů je závislá vysokém stupni genetické diverzity výchozího šlechtitelského materiálu, jen tak lze dosáhnout selekčního efektu
- Rozšíření genetické diverzity výchozího šlechtitelského materiálu vyžaduje identifikaci rozdílných kmenových rostlin (genových zdrojů) – rodičů, kteří budou kříženi s elitními kultivary
- MAS úspěšně asistuje při výběru těchto rodičů-více detailní informace pro jejich výběr

Studie heteróze

- U plodin jejichž produktivita je založena na získávání hybridů: kukuřice, čirok
- Využívá se heterozního efektu
- Markery se v tomto případě používájí pro charakterizaci diverzity rodičovských linií
- Heterózní efekt nelze predikovat, zatím ani pomocí MAS

Heterózní křížení kukuřice