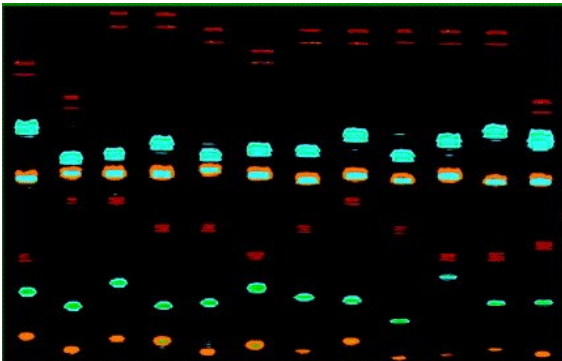


LGN0232 - Genética Molecular

Marcadores Moleculares: Do Melhoramento a Conservação

Aula 9



Antonio Figueira

CENA

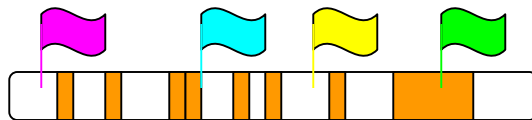
figueira@cena.usp.br

Genética e Melhoramento

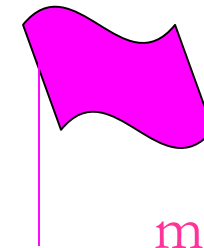
- “Arte” e seleção inconsciente
 - da invenção da Agricultura até século XIX
- 1900s - Descoberta dos princípios genéticos
- 1920-50 - Melhoramento genético científico
 - genética quantitativa e biometria
 - *(fenótipo é previsor ruim do valor genético!)*
- 1970-80s em diante - Utilização de marcadores genéticos moleculares

Genética e Melhoramento

- Sucesso no melhoramento depende da capacidade de distinguir fatores genéticos herdáveis dos ambientais
- Marcadores genéticos são unidades herdáveis simples
- Marcadores genéticos quando associados a características de interesse aumentam a eficiência de seleção

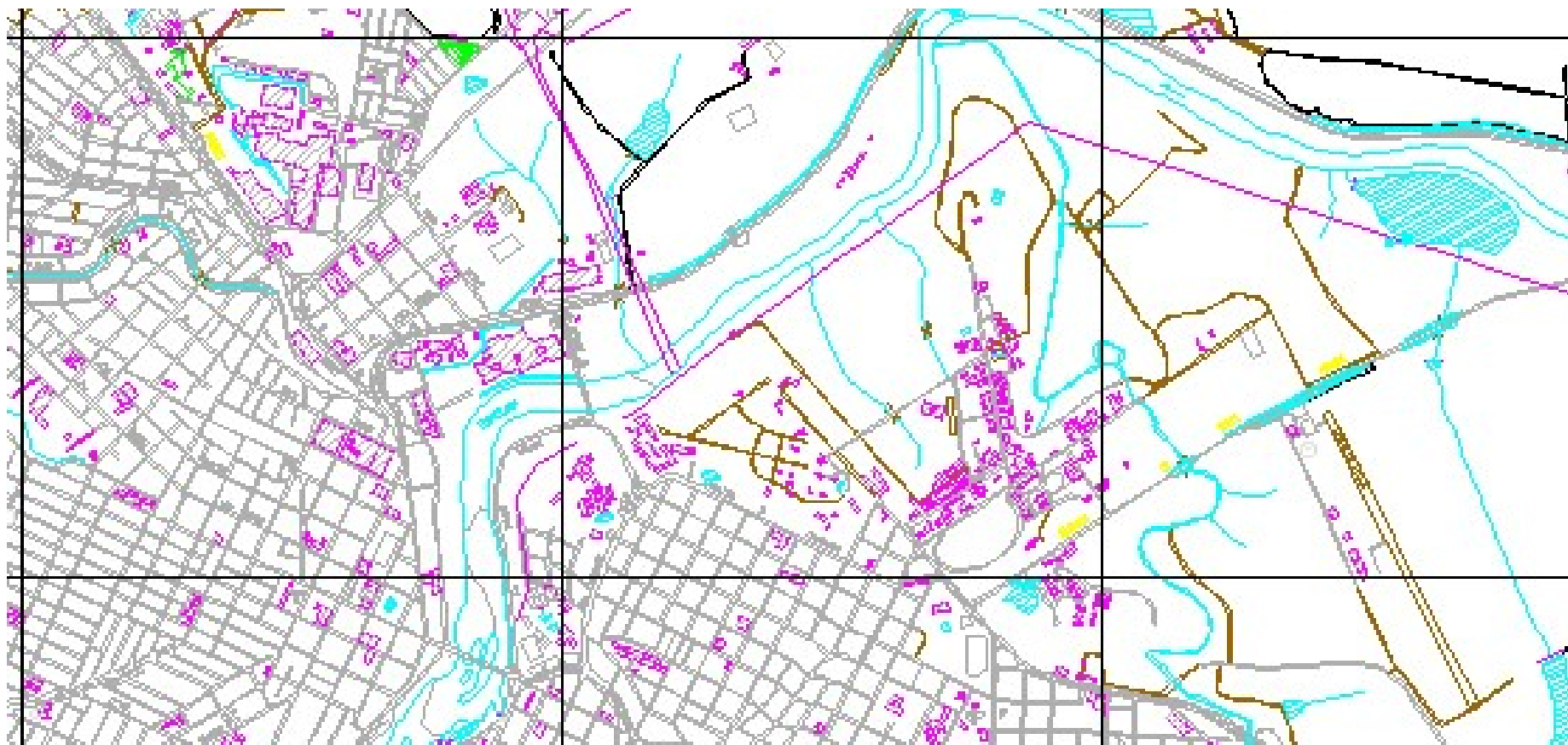


cromossomo



marcadores

Mapa de Piracicaba



“Marcadores” ???

Mapa de Piracicaba



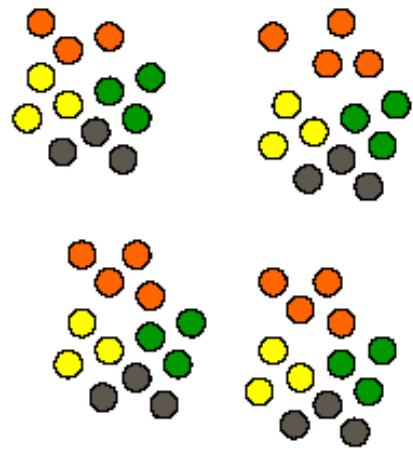
“Marcadores” ???

Aplicações no Melhoramento

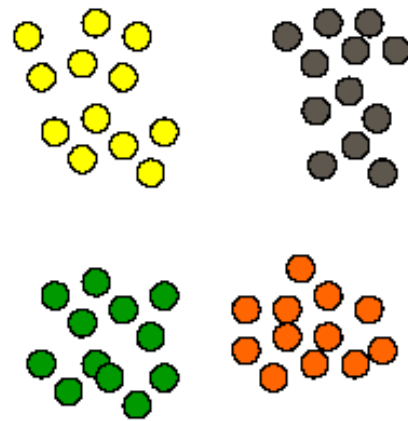
- **Diversidade genética e filogenia**
 - Conservação e caracterização de germoplasma
 - Identificação de acessos
 - Seleção de genitores para cruzamentos e populações
 - Filogenia e Evolução
- **Mapeamento genômico**
 - Seleção Assistida por Marcadores - MAS
 - Clonagem/Identificação de genes

Estrutura Genética de Populações:

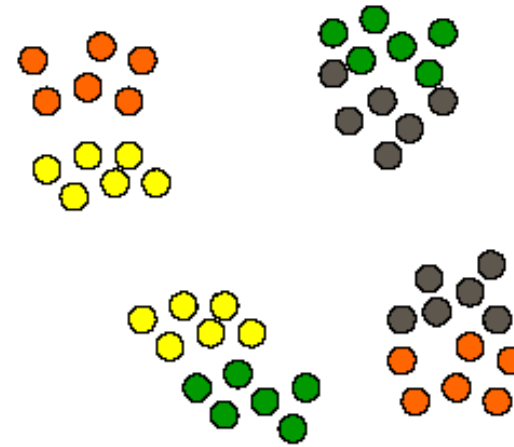
. É o modo em que variabilidade genética está organizada em populações



GRUPO 1
variabilidade
genética não
estruturada entre
populações
(uma só pop.)



GRUPO 2
Mesmo nível de
variabilidade
genética, mas esta
organizada entre
populações



GRUPO 3
Mesma magnitude de
variabilidade
genética, porém esta
organizada entre e
dentro de populações

Estrutura Genética de Populações

Importância para estudar a estrutura genética das populações

- **Conservação Genética:** organização da diversidade entre ou dentro das populações
- **Melhoramento Genético:** avaliar a diversidade genética da espécie disponível
- **Manutenção da Diversidade:** estratégias para manter a diversidade da espécie
- **Fragmentação de Populações:** recuperação de ambientes degradados

Estrutura Genética de Populações

Forças que afetam a estruturação genética das populações

- **Deriva genética:** aumenta a estruturação genética
- **Fluxo gênico:** reduz a estruturação genética
- **Mutação:** aumenta a estruturação genética
- **Seleção:** aumenta a estruturação genética

Histórico de Marcadores

1. **Karl Sax (1923):** propôs método para localização de QTLs

ligação entre genes de característica qualitativa (cor de semente) e quantitativa (peso de semente)

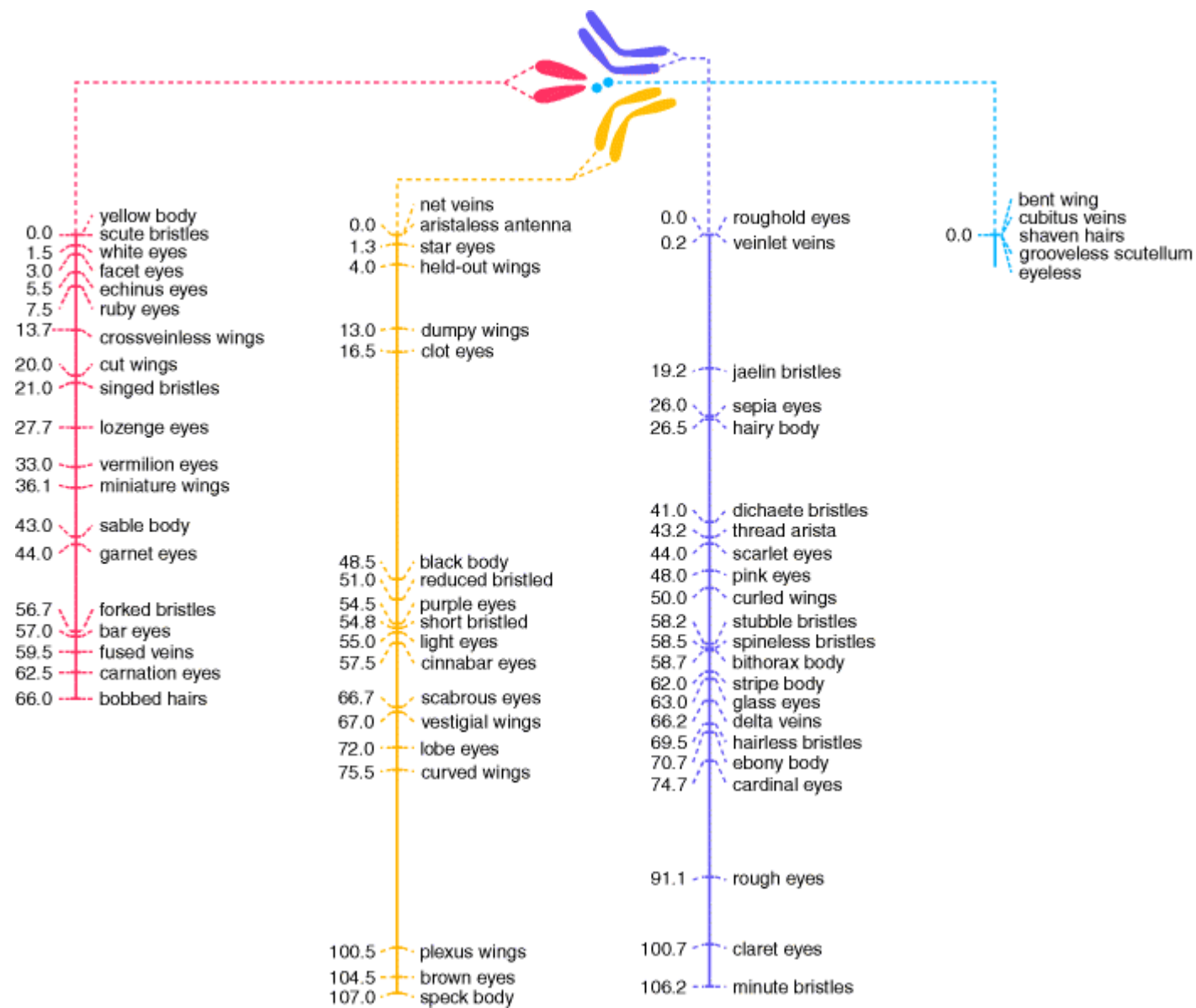
Problema: ausência de mutações múltiplas em estoque de elite, baixa viabilidade

2. **Hunter & Marker (1957)** - marcas bioquímicas

desenvolveram isoenzimas em gel de amido

3. **Hubby & Lewontin (1966)**

demonstraram que 30% de loci de isoenzimas exibiam polimorfismo em populações selvagens de *Drosophila*;



Drosophila – mapa genético com marcas morfológicas

Histórico de Marcadores

1970s - ferramentas moleculares

desenvolvimento de vetores de clonagem; enzimas de restrição; polimerases; ligases; Southern (1977)

4. RFLP proposto por Botstein *et al.* (1980)

descrito para humanos

5. **PCR** proposto por Mullis & Faloona (1987)

6. **VNTR** por Jeffrey (1987)

7. RAPD por Rafalski *et al.* (1990)

Histórico de Marcadores

8. **SSR** em plantas por Akkaya *et al.* (1992)
9. **AFLP** por Zabeau & Vos (1993)
10. **CAPS** por Konieczny & Ausubel (1993)
11. **SCAR** por Paran & Michelmore (1993)
12. **SNPs** em *Arabidopsis* por Cho *et al.* (1999)

Origem do Polimorfismo Molecular

1. Mutações pontuais - SNPs

A. ATCTCGTGATTCCTAGTCGTA
TAGAGCACTAAGGGATCAGCAT ex. Sítio *Eco*RI

B. ATCTCGTGATTATAGTCGTA
TAGAGCACTAAATATCAGCAT

2. Pequenas deleções e/ou inserções - InDels

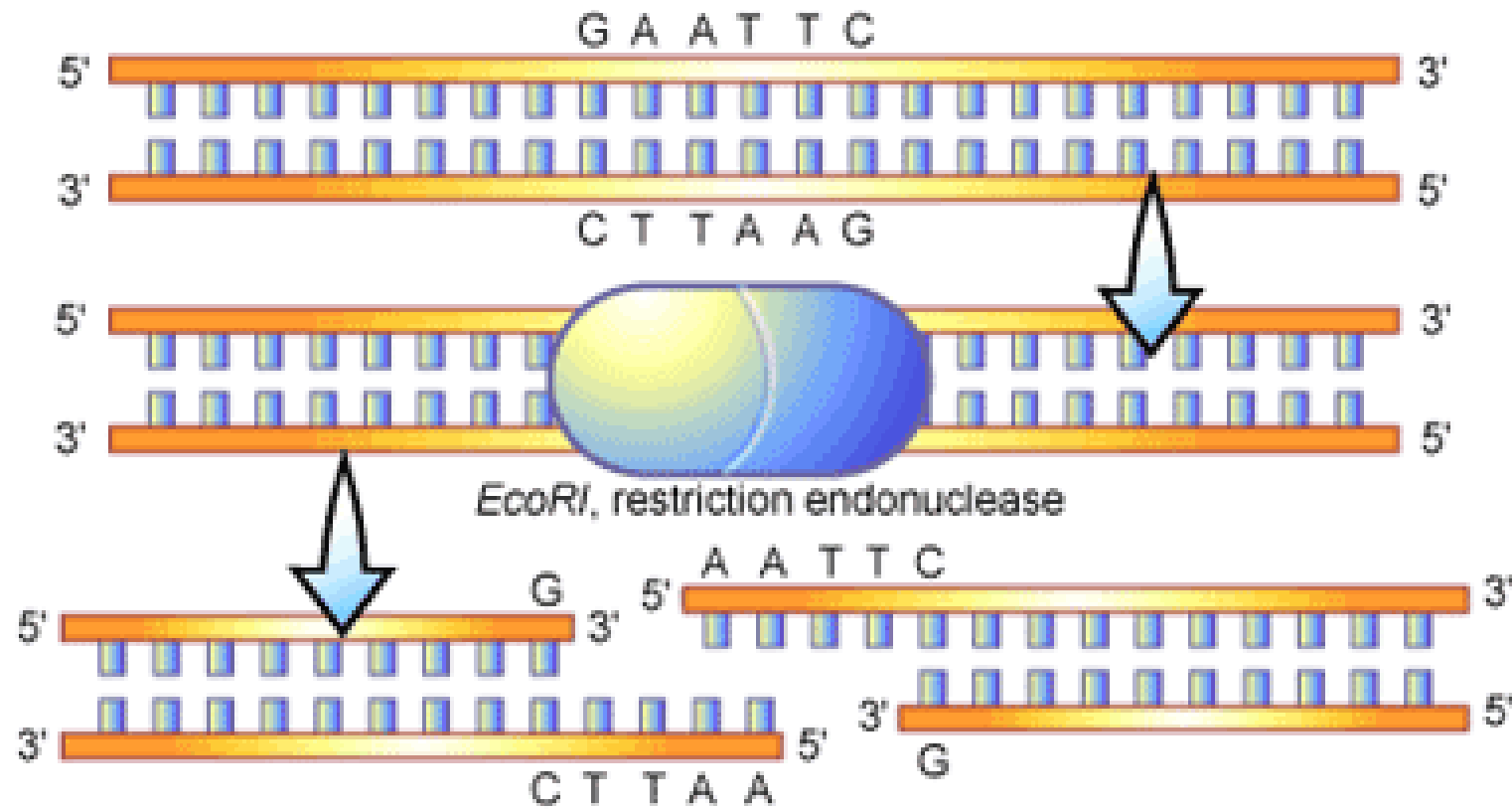
A. ATCTCGTCTAGTCGTA
TAGAGCAGATCAGCAT

B. ATCTCGT---GTCGTA
TAGAGCA---CAGCAT

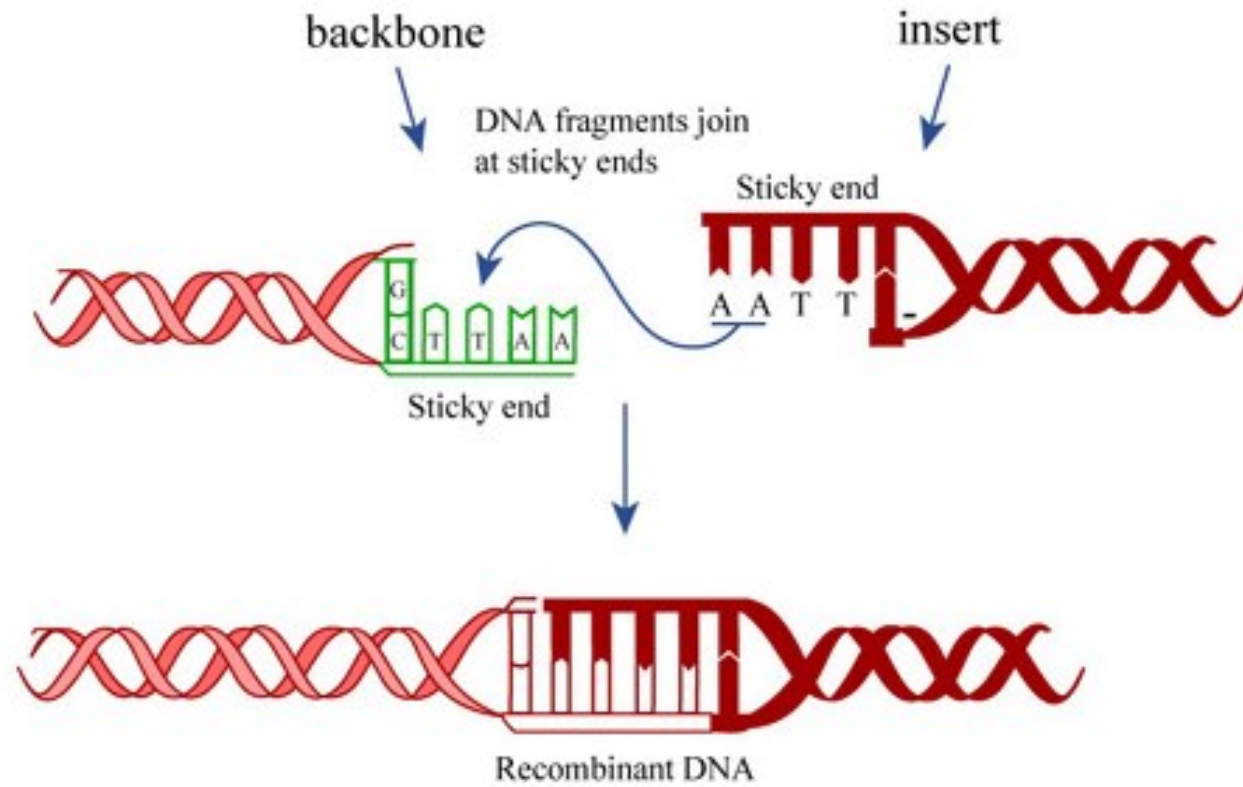
Relembrando...

Caixa de ferramentas da genética molecular...

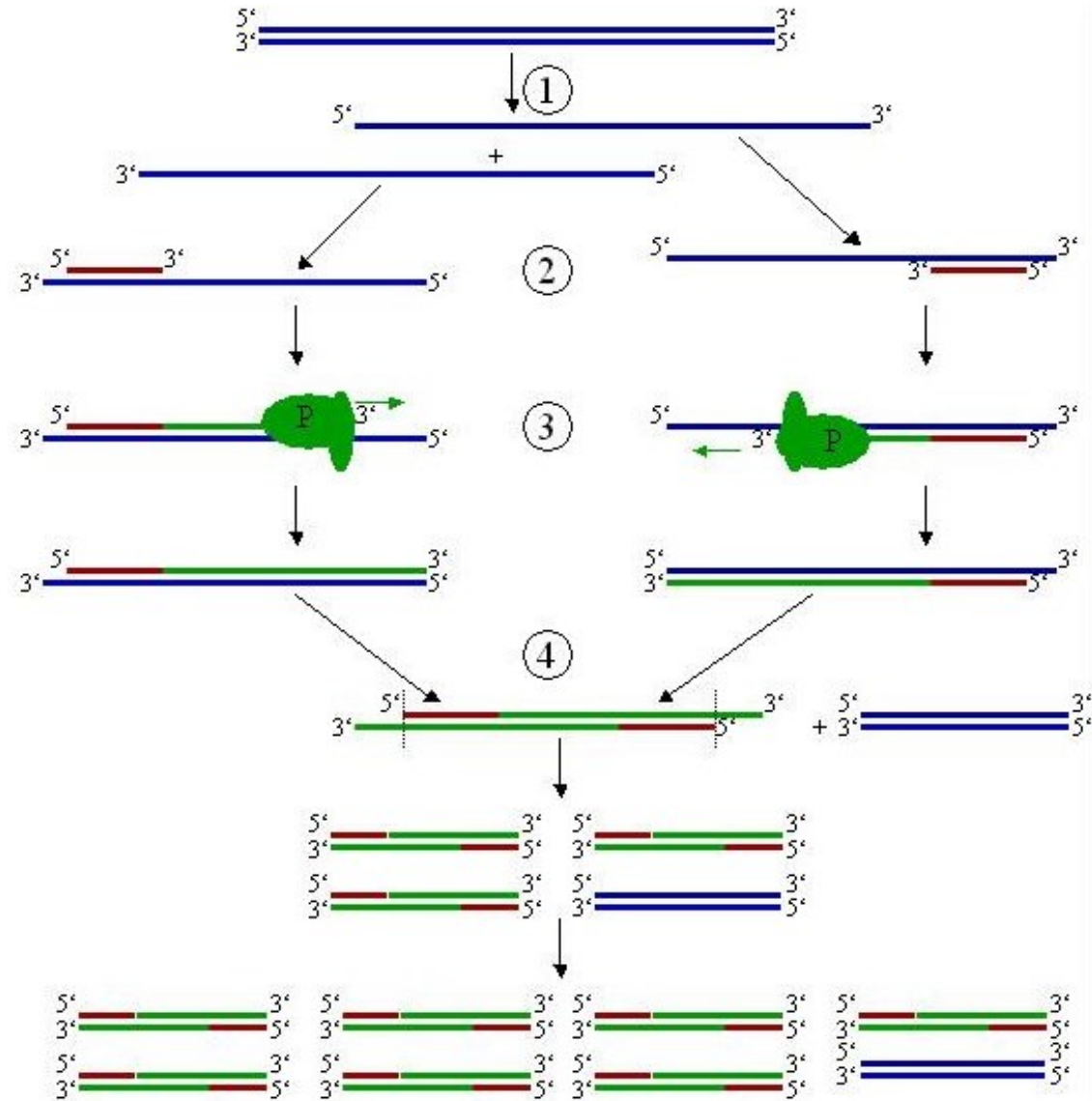
Enzimas de Restrição



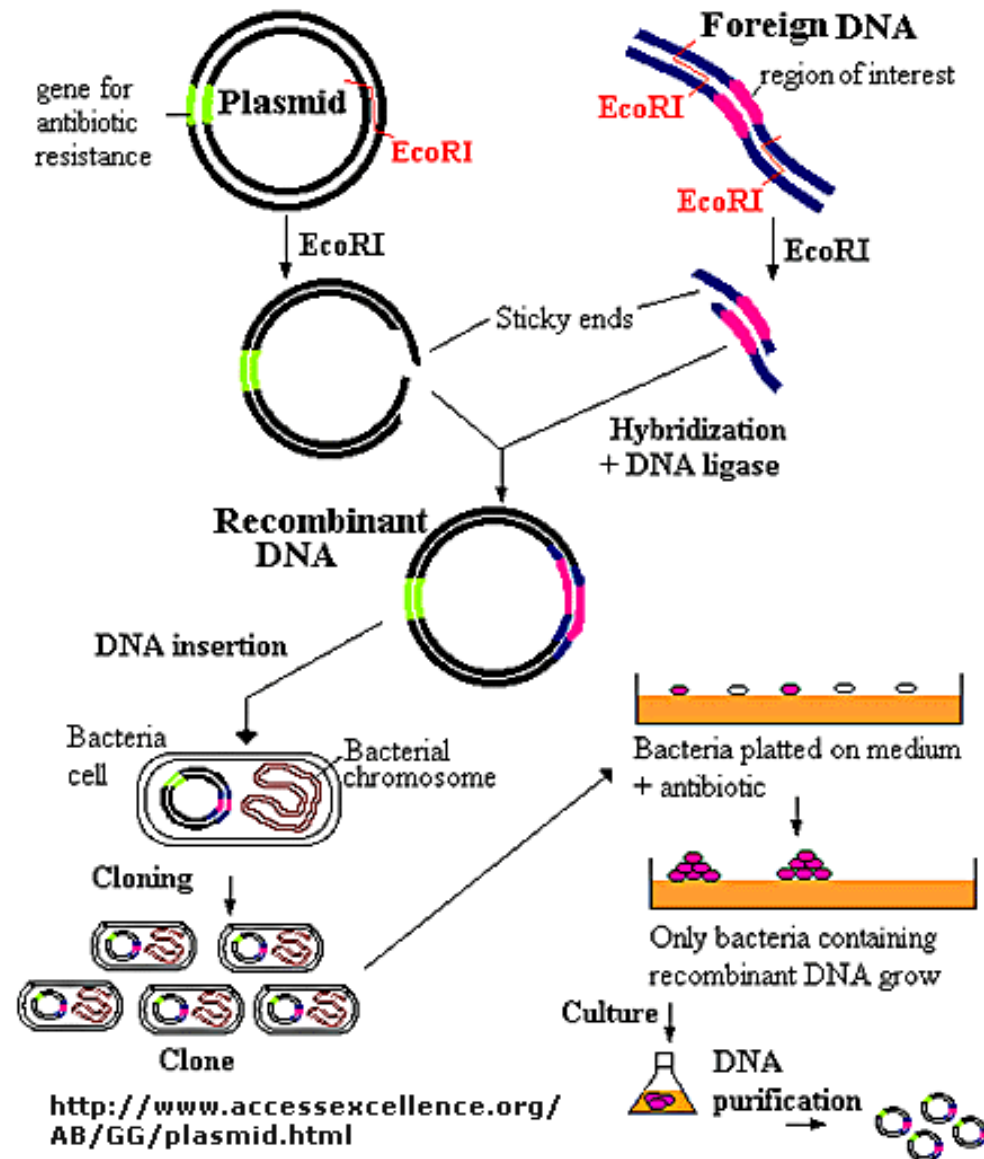
DNA Ligases



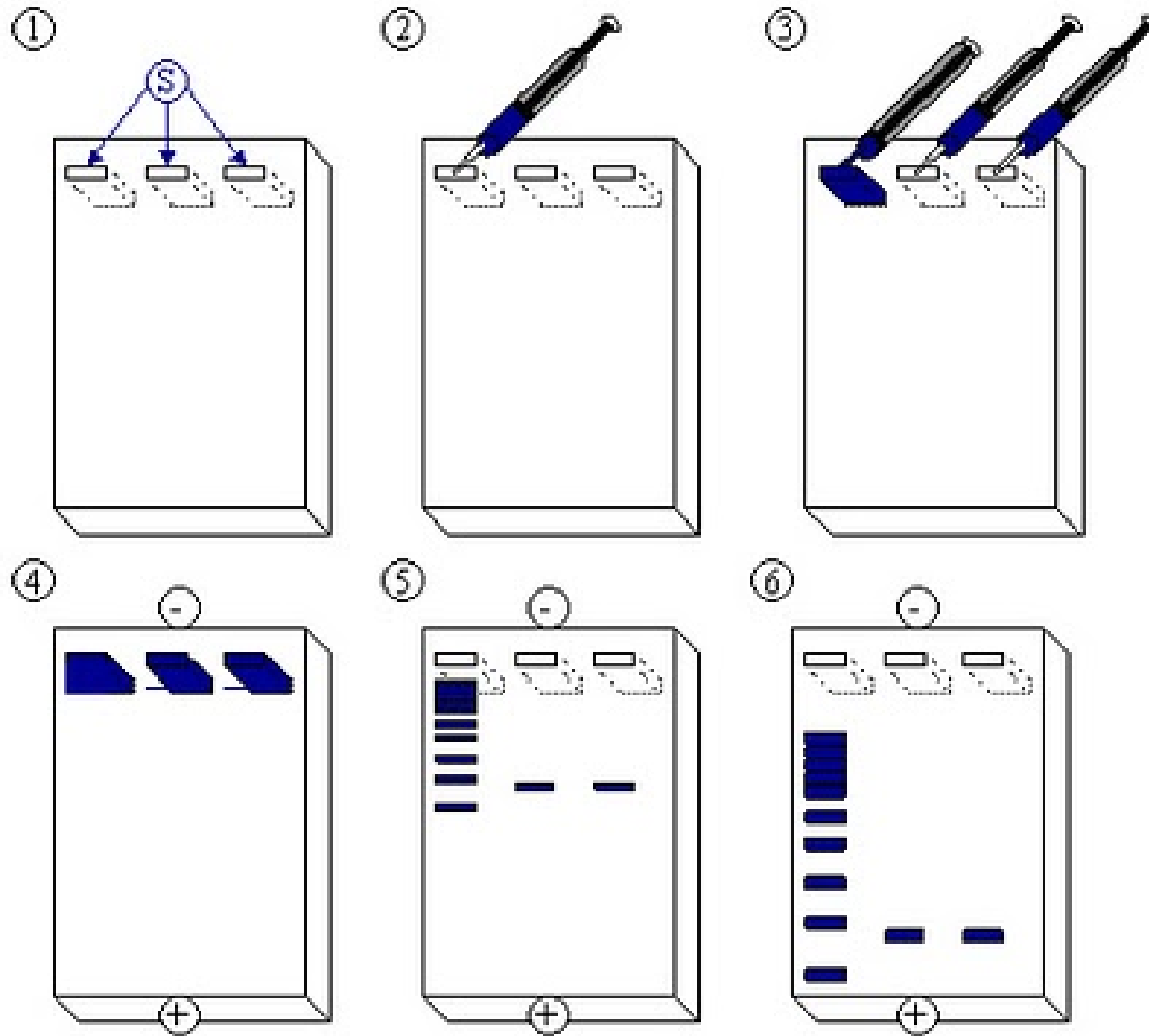
DNA Polymerases



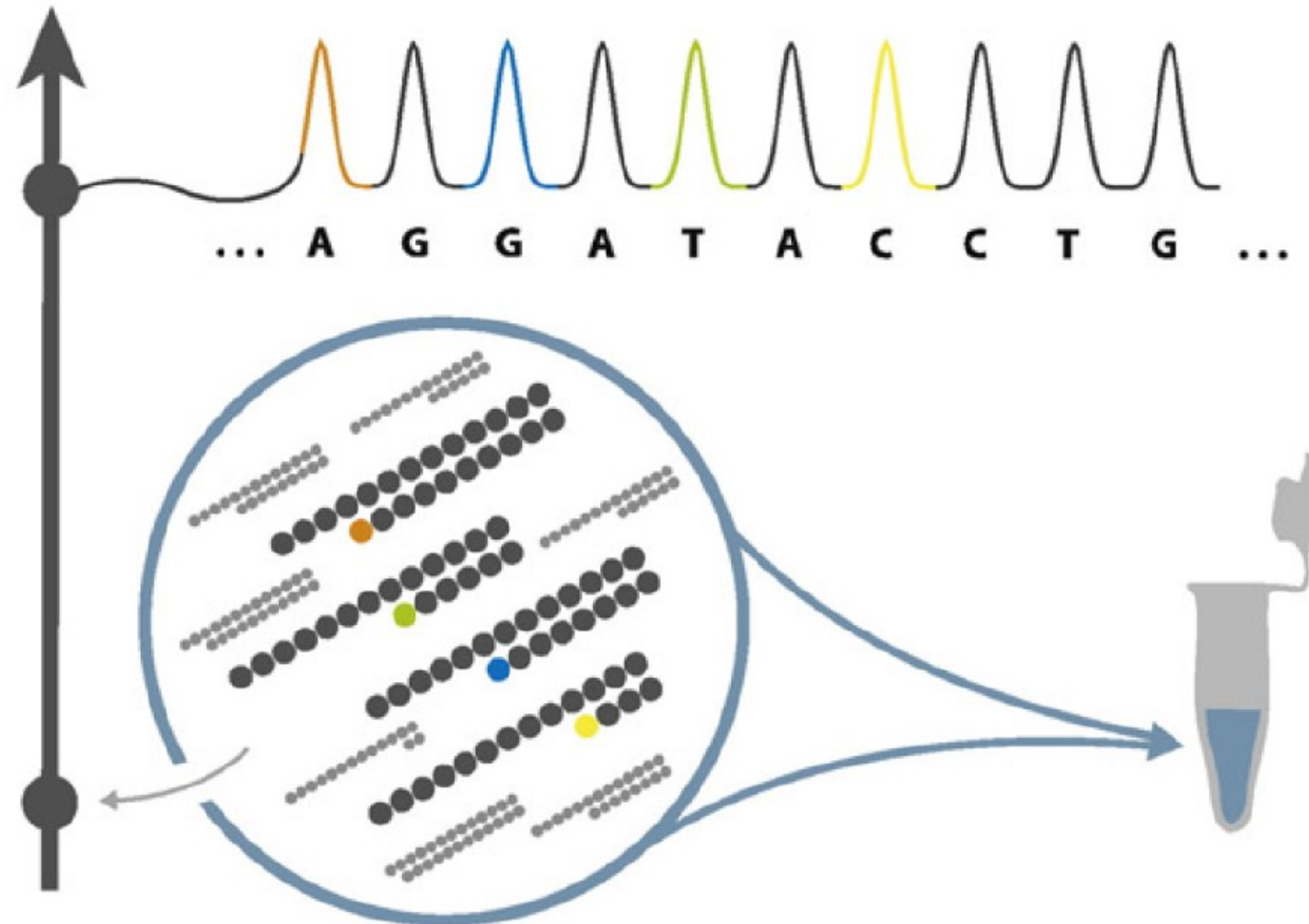
Clonagem Molecular



Eletroforese em Gel



Sequenciamento de DNA



Mas o que é um Marcador Genético?

- Variação ou polimorfismo que, numa população segregante, se comportam de acordo com as leis Mendelianas!
- Baseiam-se na existência de variabilidade e na possibilidade de sua detecção
- **Tipos:**
 - Morfológicos
 - Bioquímicos (isoenzimas)
 - Moleculares (AFLP, microsatélites, SNPs, etc.)

Marcadores Morfológicos

- Monomorfismo



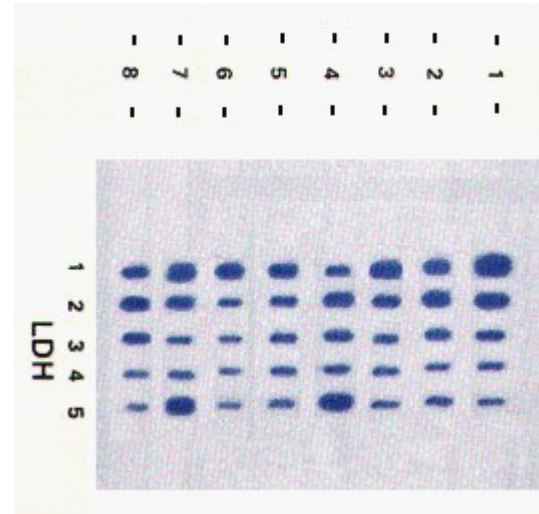
- Polimorfismo



Marcadores Bioquímicos

- Variabilidade observada como resultado da tradução de RNA em proteínas – ex. isoenzimas

- Monomorfismo



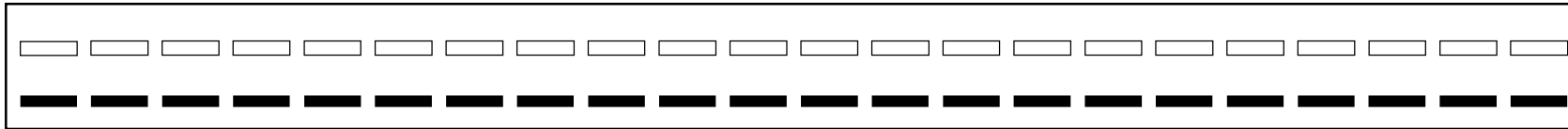
- Polimorfismo



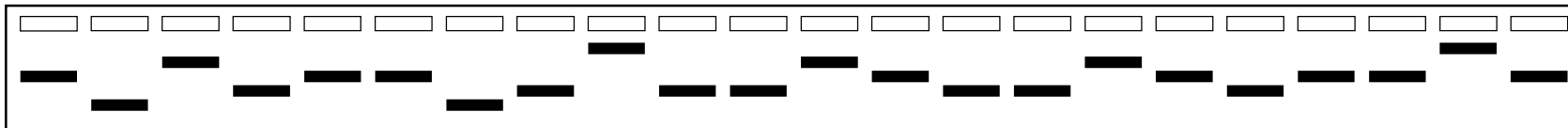
Marcadores Moleculares

- Variabilidade surge por mutação, que é a fundamento para identificação de marcadores moleculares

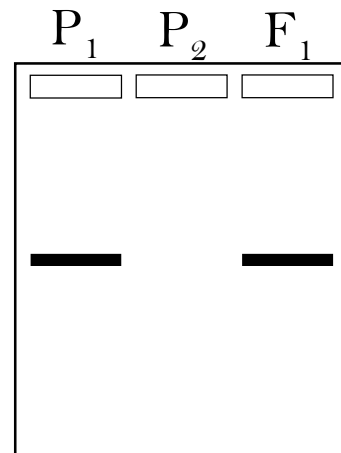
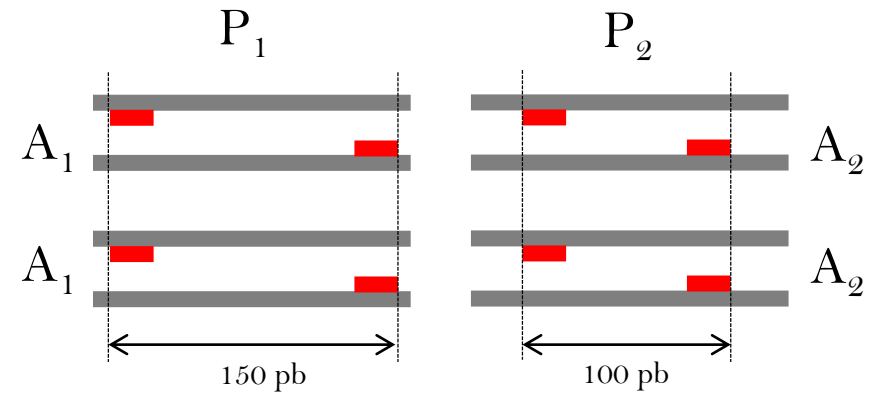
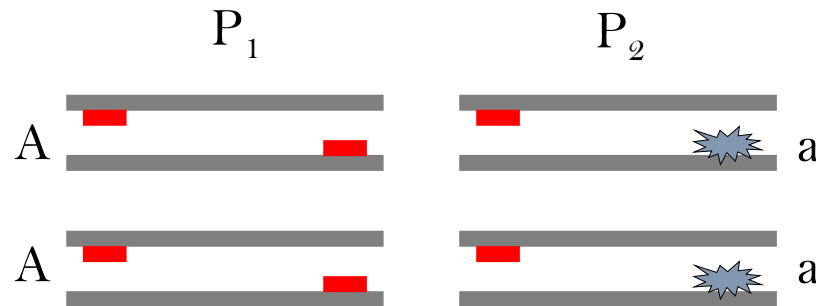
- **Monomorfismo**



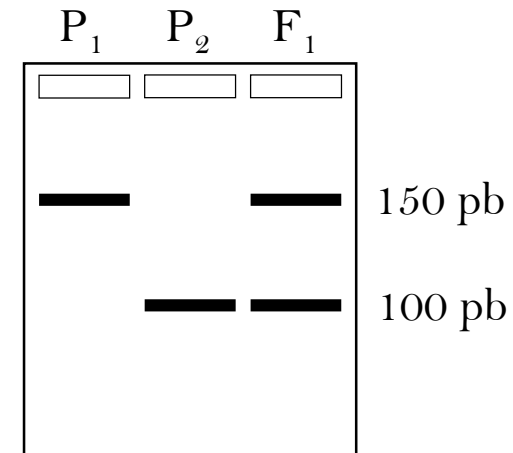
- **Polimorfismo**



Marcadores Dominantes ou Codominantes



Dominante



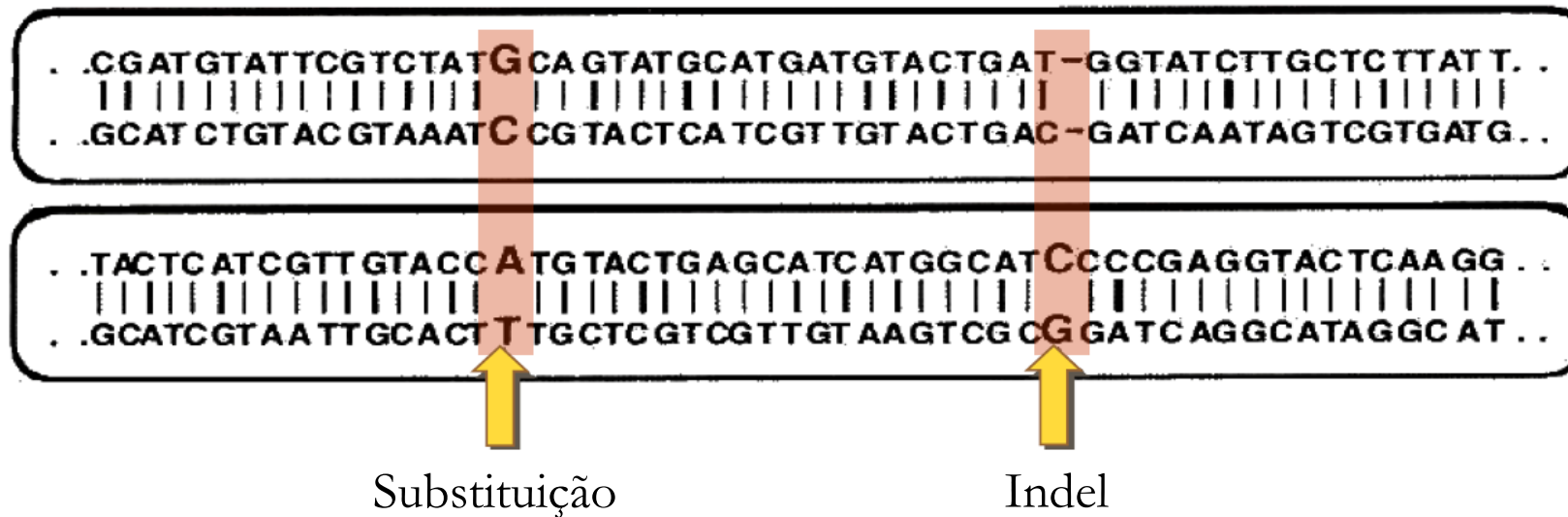
Codominante

Marcadores Moleculares

Origem do polimorfismo

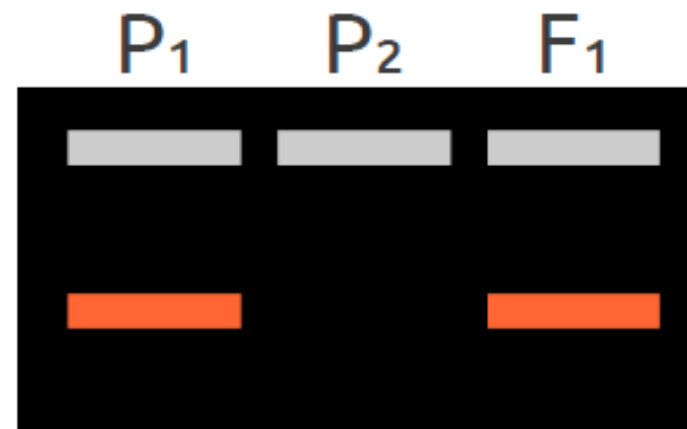
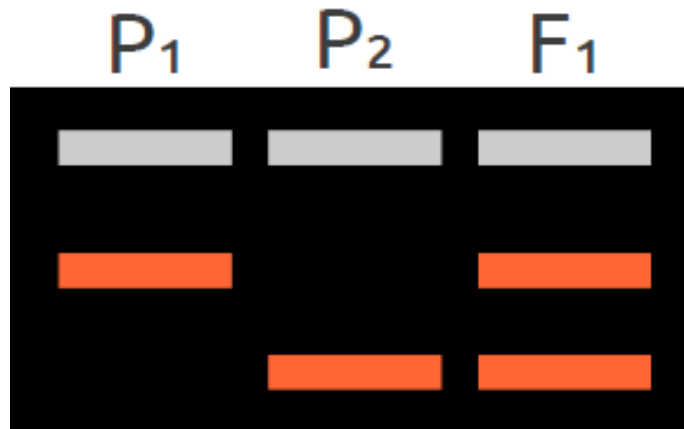
Polimorfismos de tamanho: originados por indels (inserções ou deleções)

Polimorfismos de base: causados por substituição de bases nitrogenadas



Marcadores Moleculares

- Herança
 - Dominantes: não se identificam heterozigotos
 - Codominantes: identificam-se heterozigotos

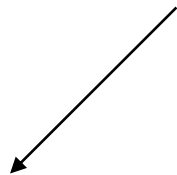
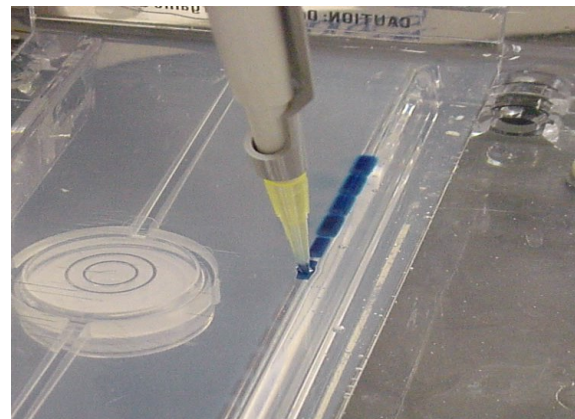


Alguns Exemplos de Marcadores Moleculares

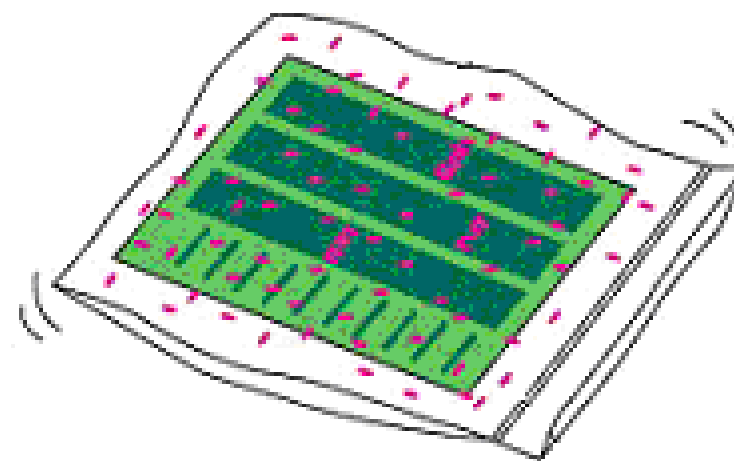
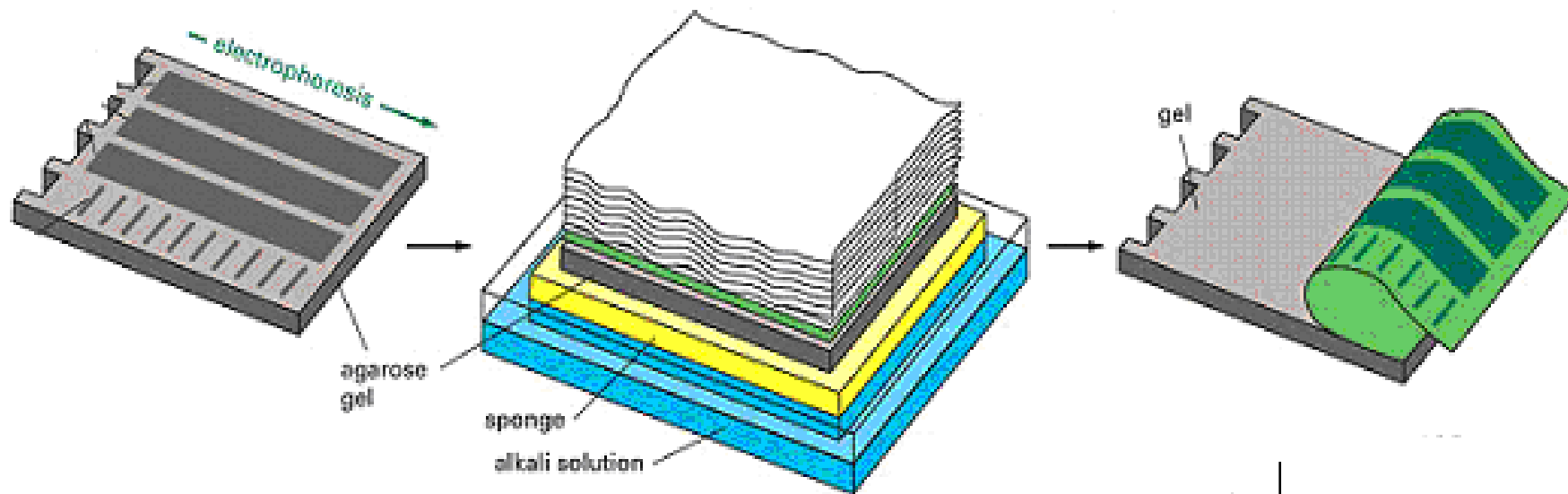
Metodologia de RFLP

1. Digerir DNA com enzimas de restrição em fragmentos pequenos
2. Separação dos fragmentos por gel eletroforese
3. Transferência de fragmentos de DNA para filtro
4. Visualização dos fragmentos de DNA
 - sondas marcadas (^{32}P) ou a frio – clones em *E. coli*
5. Análise dos resultados
 - bandas analisadas para alelos e/ou presença/ausência
 - diferenças em padrão de bandas reflete diferenças genéticas

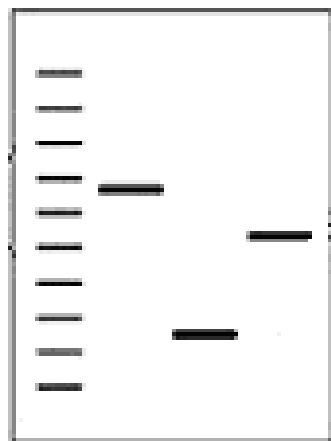
A escolha de sonda/enzima de restrição é crucial



Eletroforese

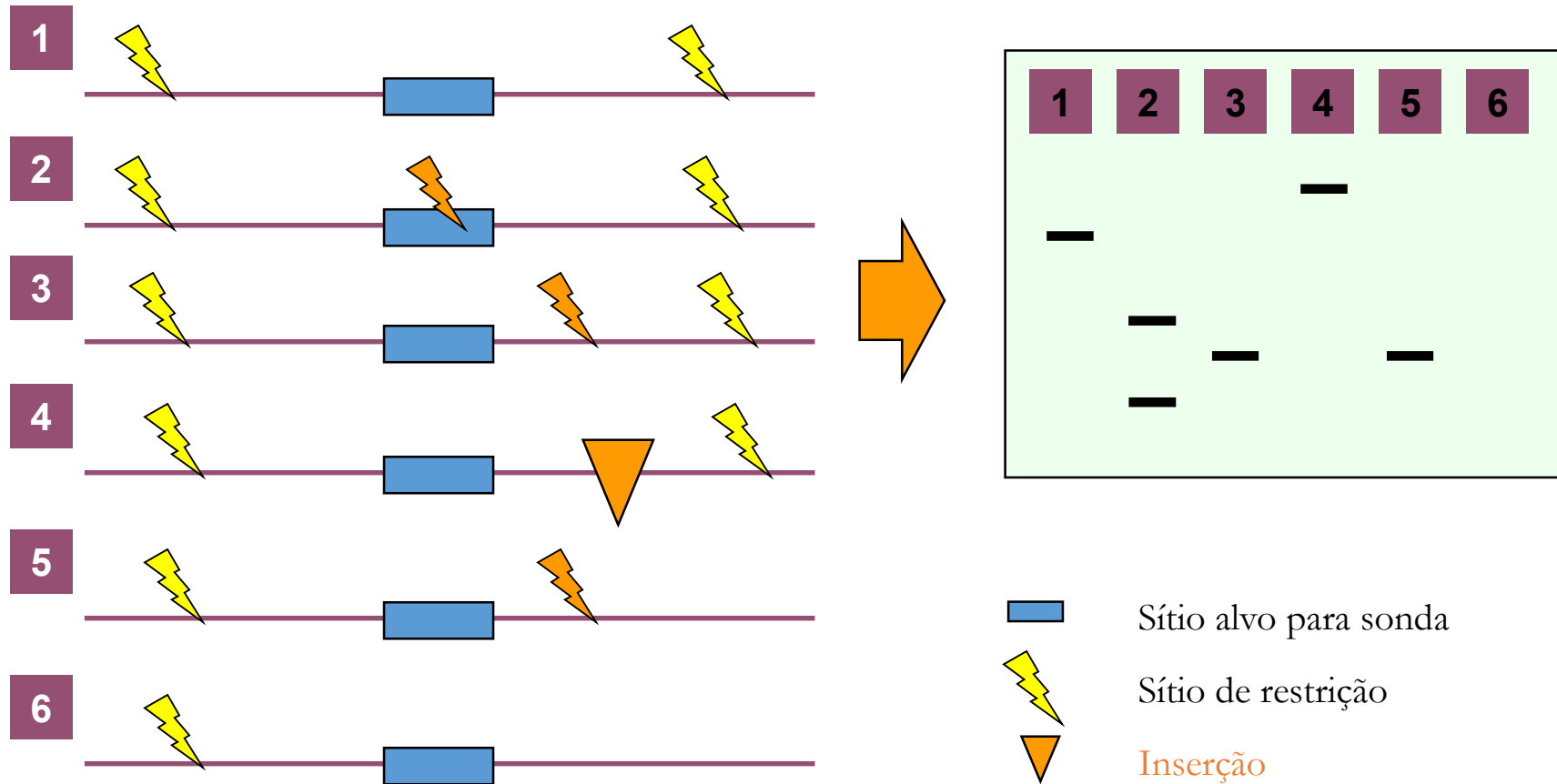


Hibridização com sonda marcada

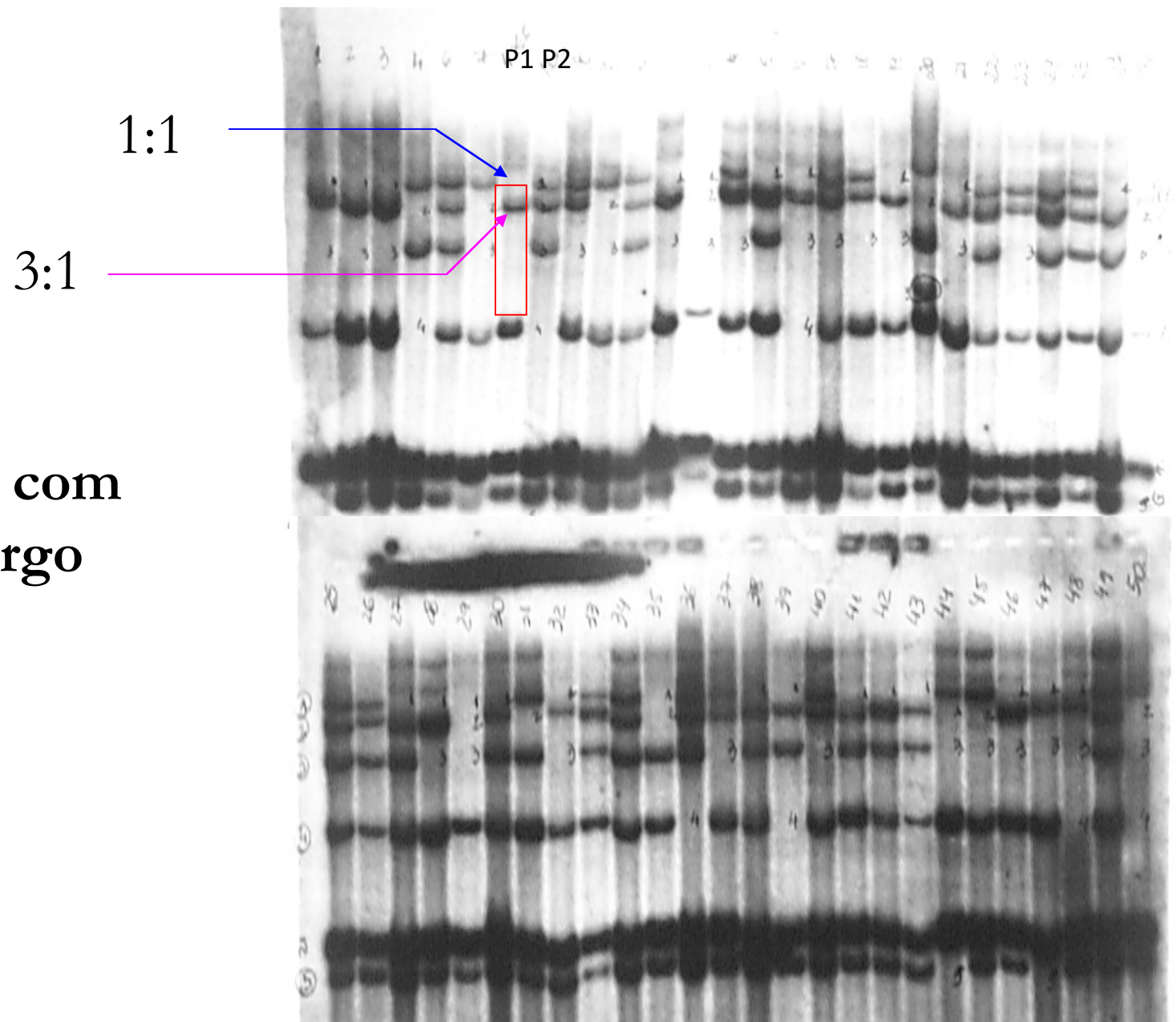


Exposição em filme de Raios-X

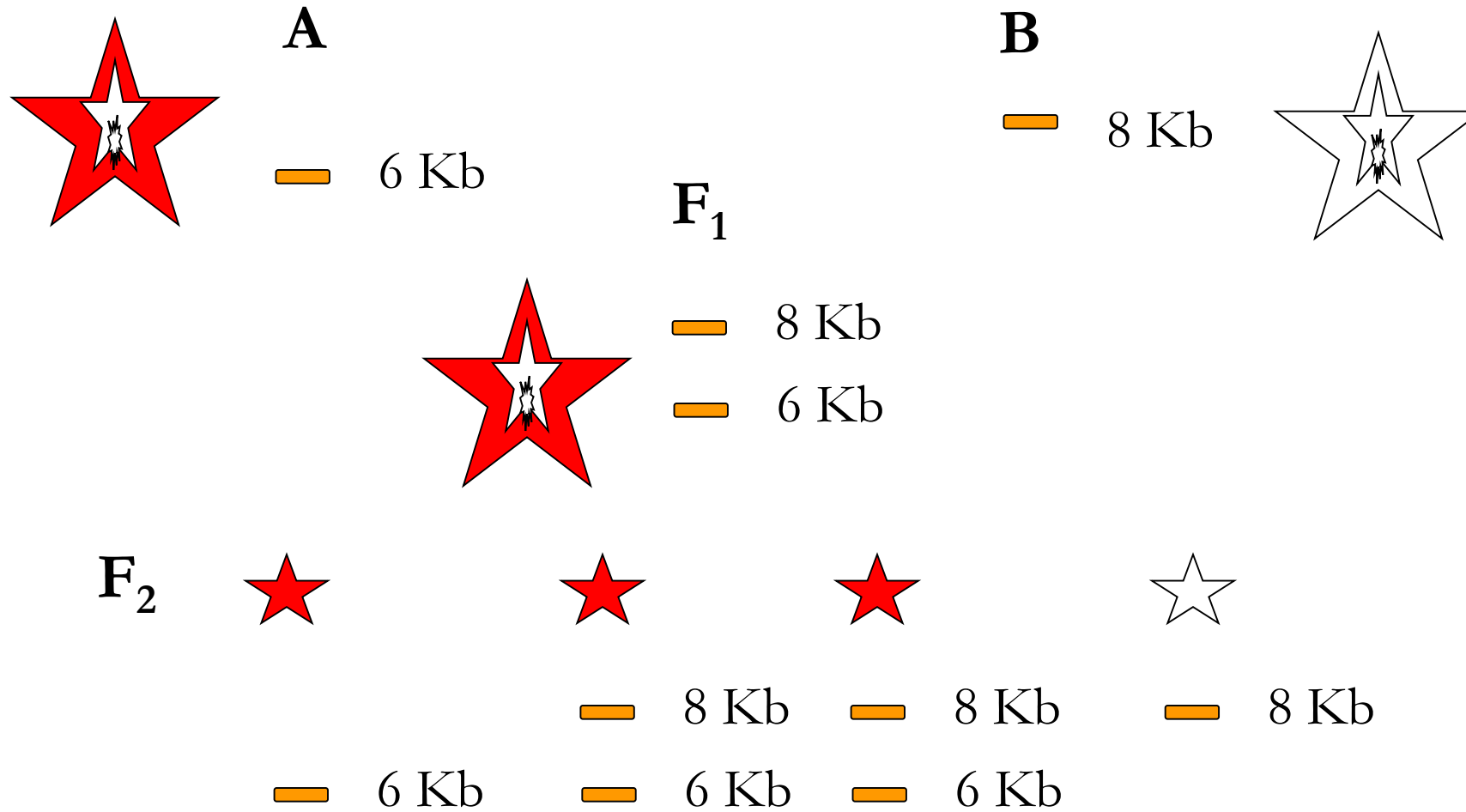
Interpretação de resultados RFLP



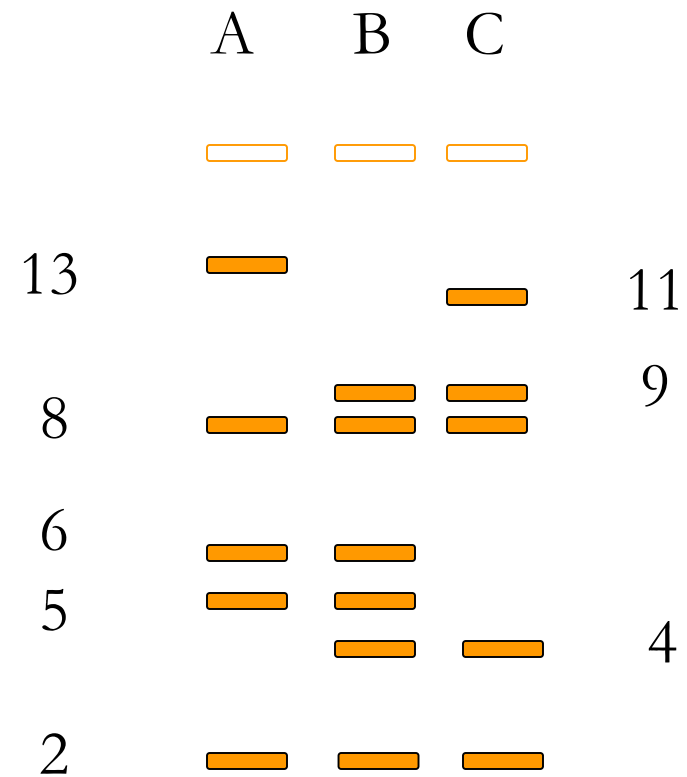
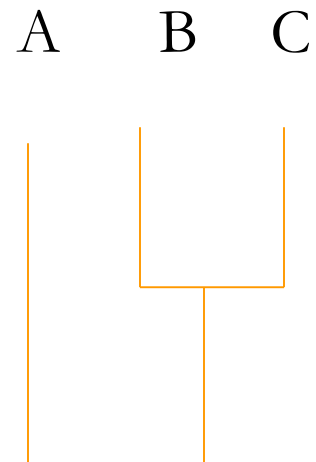
Hibridização com sonda de sorgo



Herança de RFLPs



Análise de Diversidade e Filogenia por RFLP



8 bandas total

Quantas compartilhadas entre AB, BC e AC?

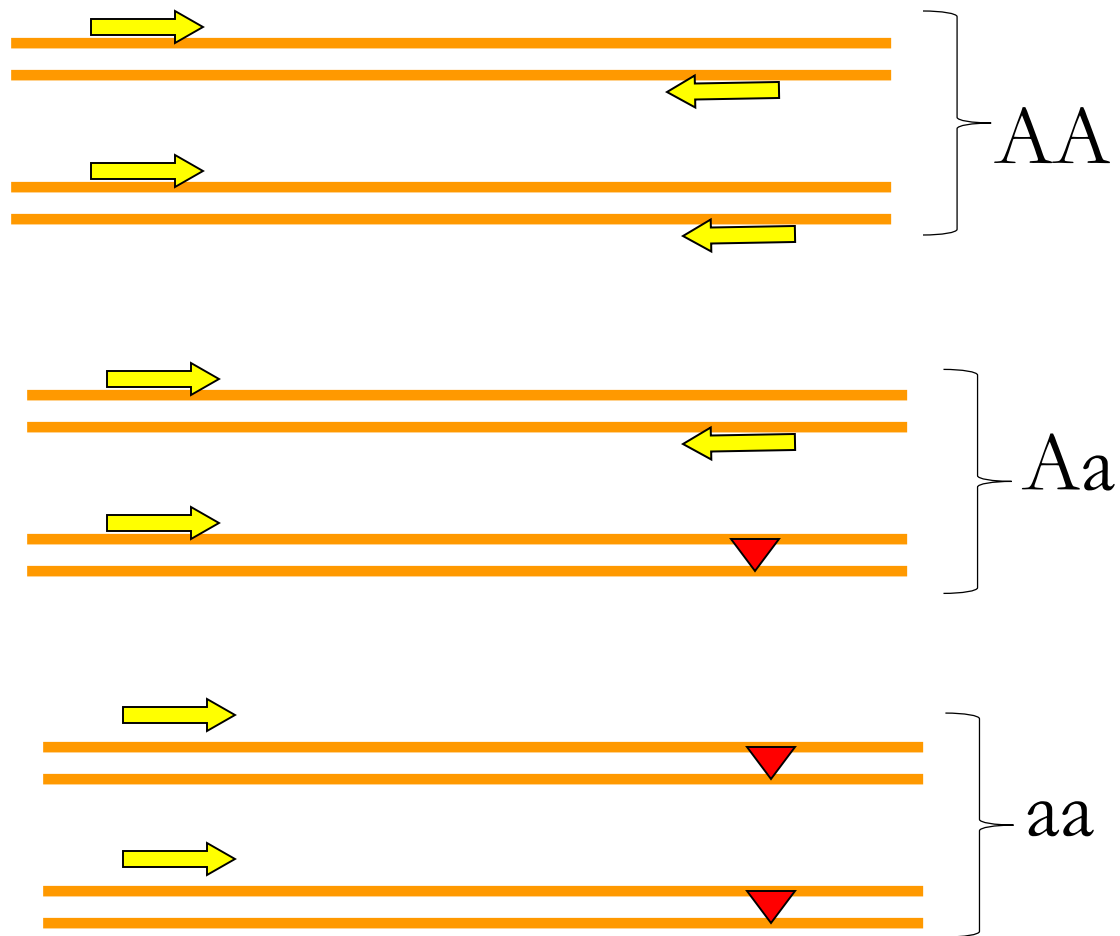
RAPD

Random Amplified Polymorphic DNA

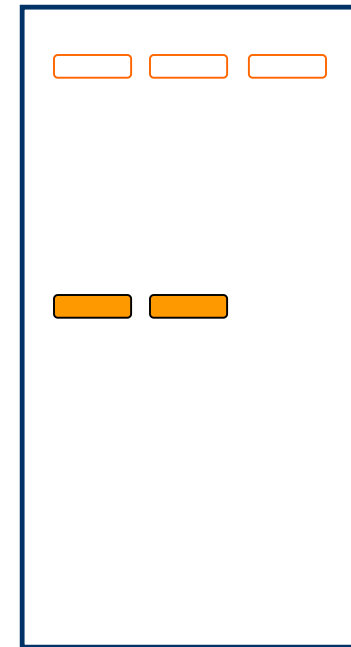
Polimorfismo de DNA amplificado ao acaso

- Representa regiões codificadoras ou não do genoma; de distribuição aleatória
- **Dominante:**
 - Apenas presença ou ausência da banda (não identifica heterozigotos)
- **Base genética:**
 - Amplificação via PCR
 - Utiliza apenas um primer curto e aleatório (arbitrário - >50% G+C)
 - *Primer* se anela em diversas partes do genoma
 - Polimorfismo se origina da mutação no sítio de anelamento do *primer*

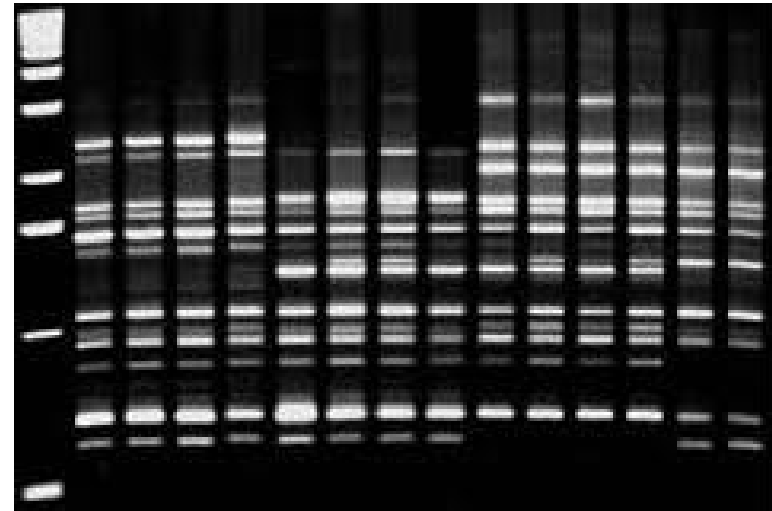
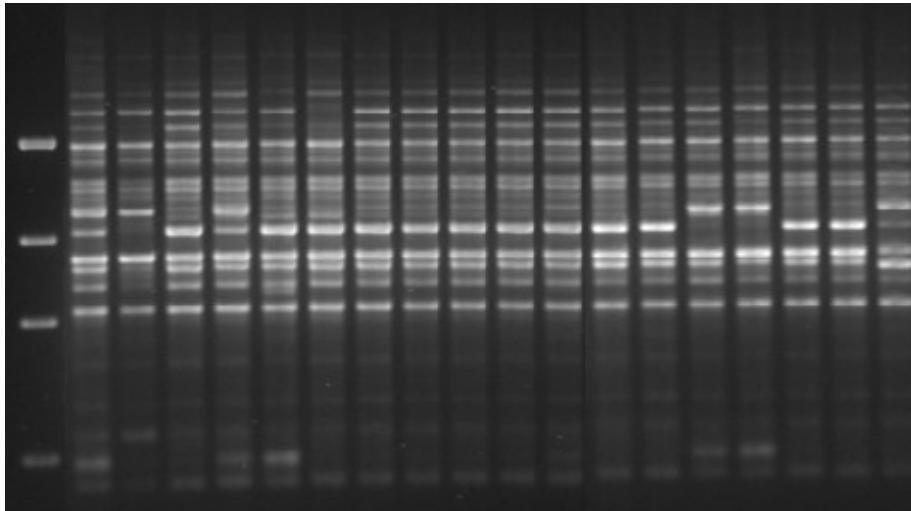
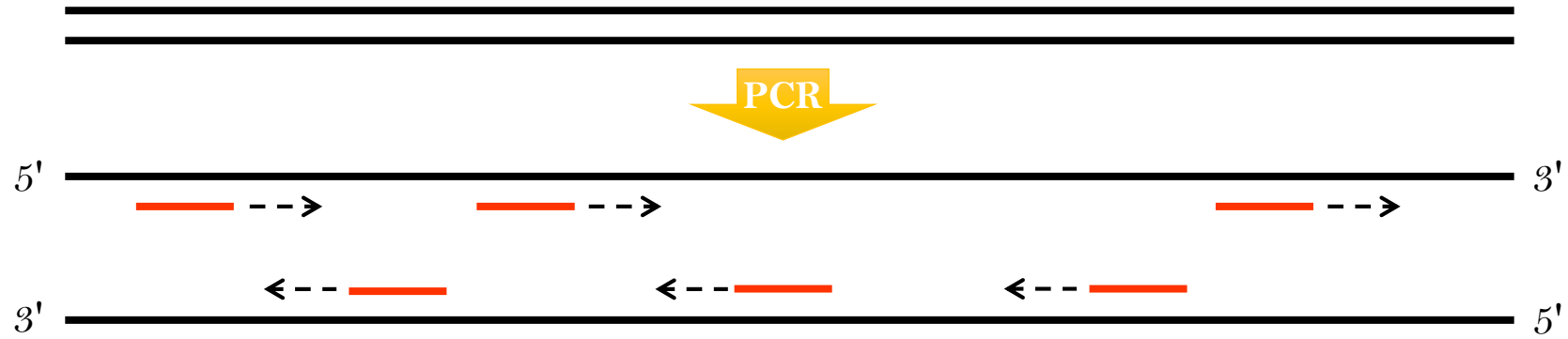
RAPD



AA Aa aa



RAPD



RAPD

Vantagens:

- ❑ Prático e simples
- ❑ Barato, quando já se tem os *primers*
- ❑ Gera muita informação em pouco tempo

Desvantagens:

- ❑ Marcadores Anônimos
- ❑ Marcador dominante (informação limitada)
- ❑ **Baixa reprodutibilidade**
- ❑ **Problemas na interpretação** – mesma banda -> mesmo fragmento?
- ❑ **Bandas podem ser consideradas homólogas???**

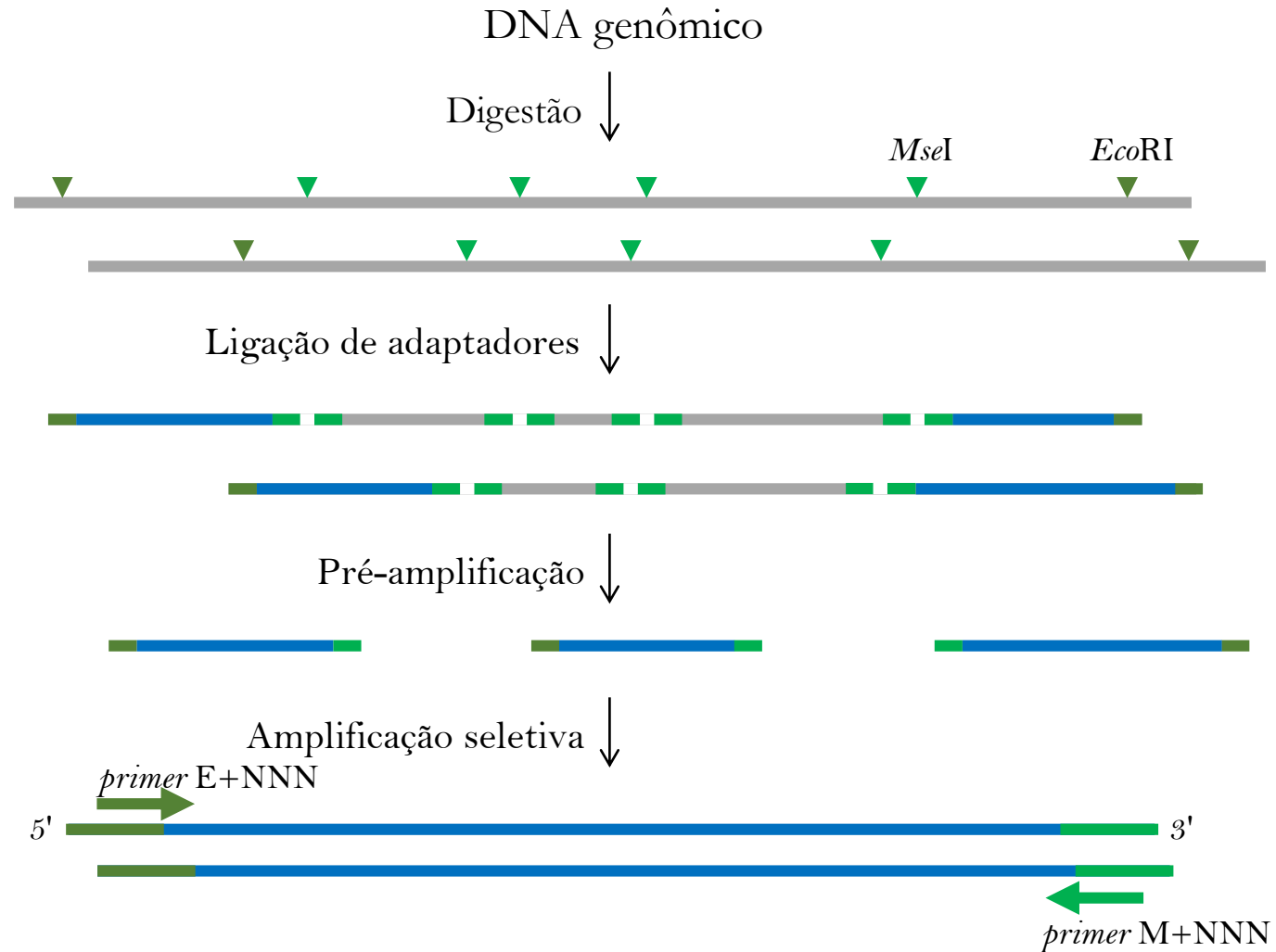
AFLP

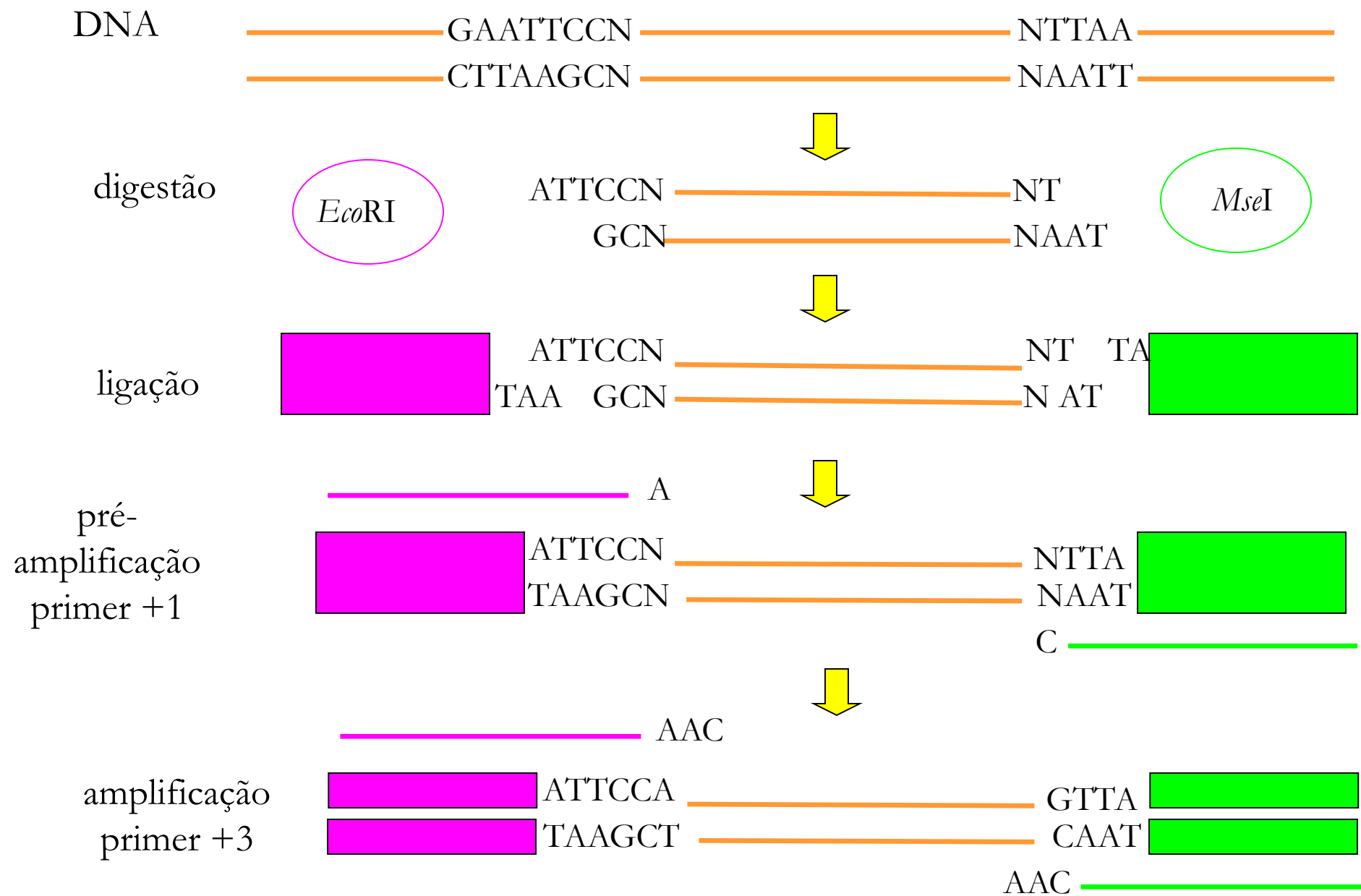
Amplified Fragment Length Polymorphism

Polimorfismo de tamanho de fragmento amplificado

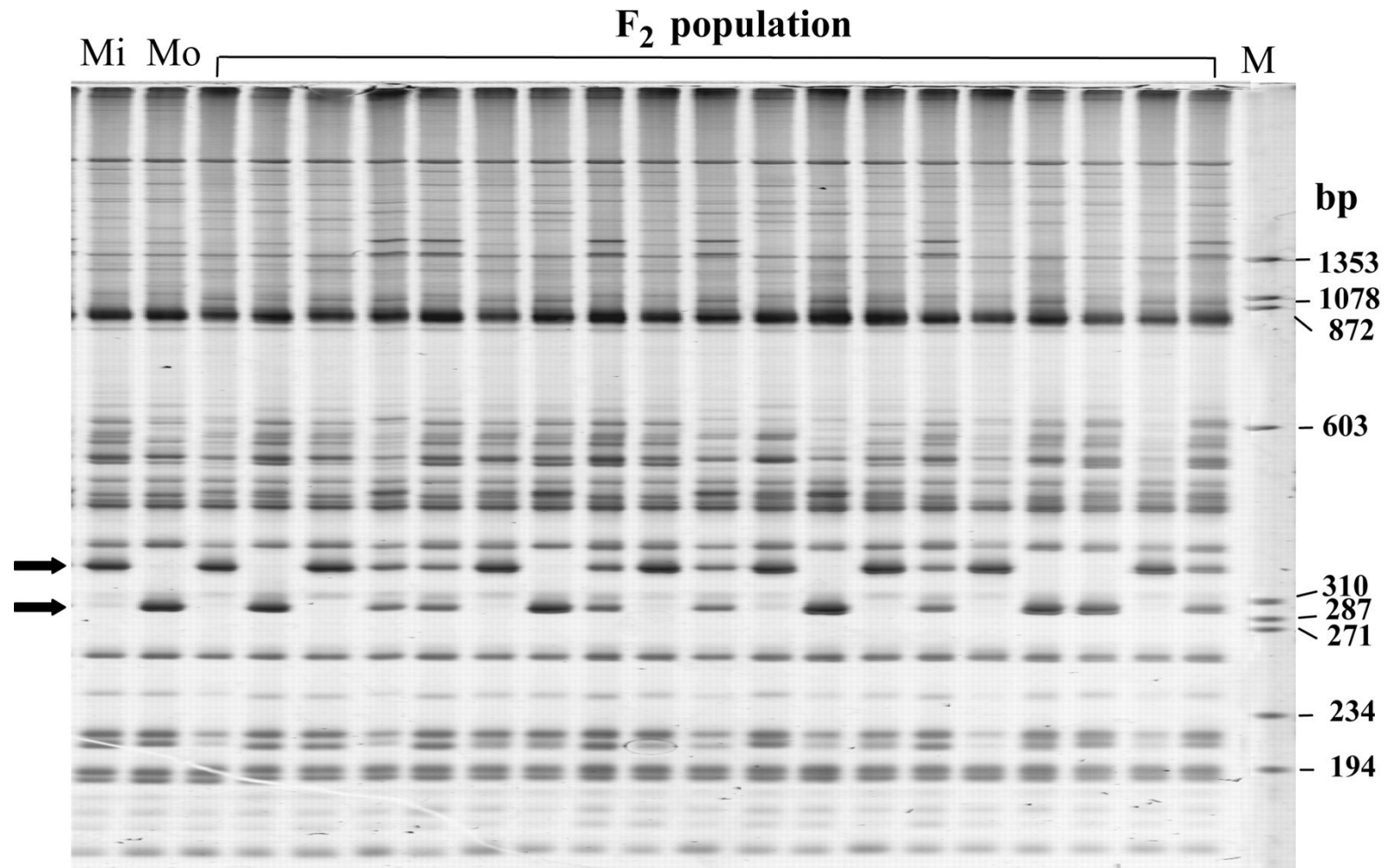
- Na maioria, representam regiões não expressas do genoma
- **Dominante:**
 - Apenas ausência e presença da banda (não identifica heterozigotos)
- **Base genética**
 - Clivagem do DNA por enzimas de restrição
 - Ligação de adaptadores
 - Amplificação seletiva dos fragmentos
 - Análise dos fragmentos amplificados em gel de poliacrilamida

AFLP



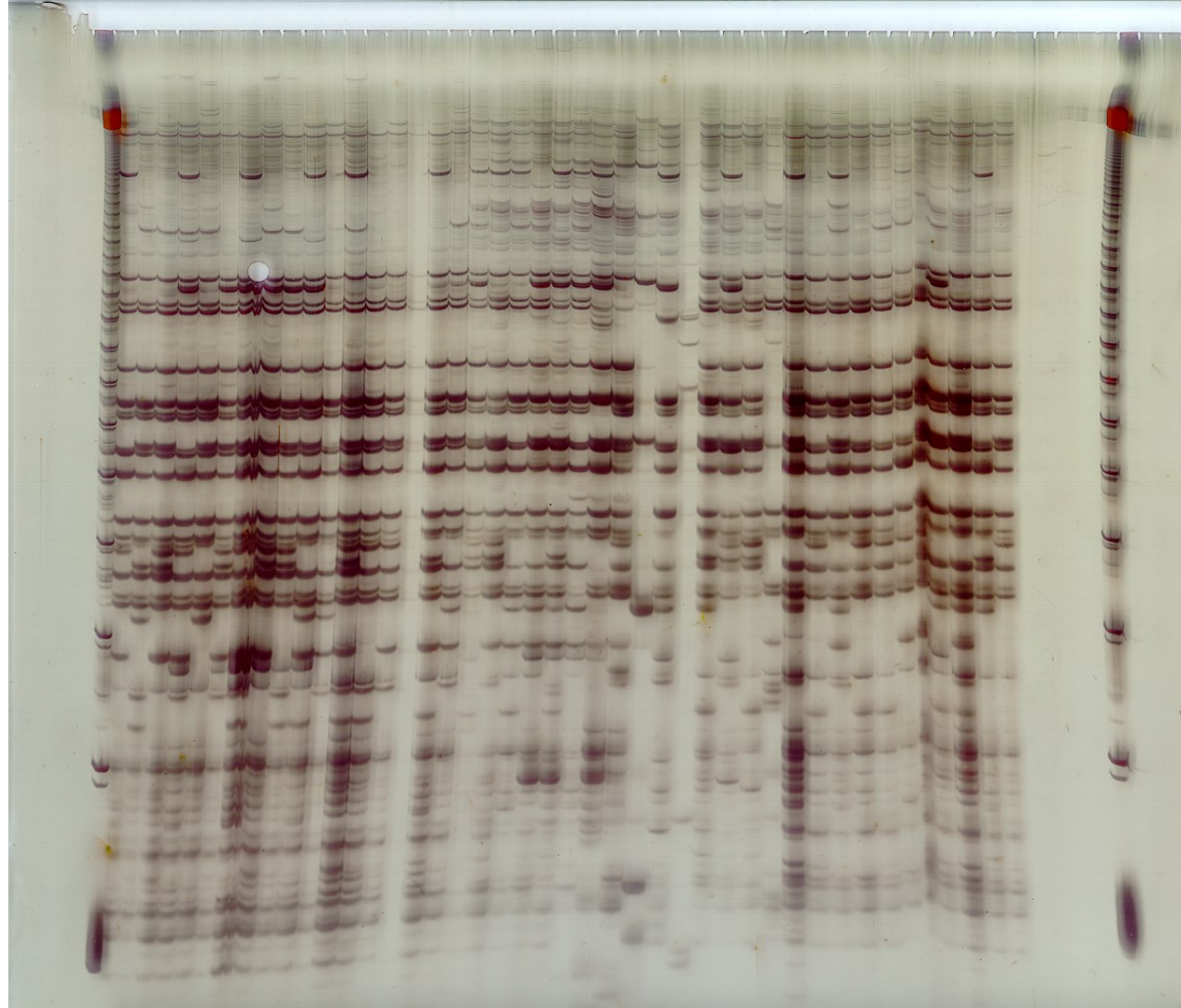


AFLP



AFLP de feijão

gel desnaturante corado com prata



AFLP

Vantagens

- ❑ Baseado em PCR
- ❑ Alta reprodutibilidade
- ❑ Altamente variável

Desvantagens

- ❑ Dominante
- ❑ Custo elevado
- ❑ Trabalhoso, por envolver muitas etapas

MICROSSATÉLITES ou SSR

Simple Sequence Repeats = SSR

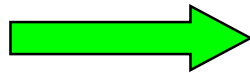
Sequências simples repetidas

- Pequenas sequências de DNA (1 a 6 nucleotídeos) repetidos em linha (*tandem*)
- Presentes em Eucariotos e Procariotos
- Presentes em região codificadora ou não codificadoras
- **Codominante**
 - Indivíduos heterozigotos são detectados
- Multialélicos e amplamente distribuído no genoma
- **Base genética:**
 - Amplificação de locos específicos de sequências repetitivas
 - Baseados em PCR com *primers* específicos

MICROSSATÉLITES ou SSR

ACTTCATTAA TGTCTTAGGT GGCAGAATAC TTTGATGCAA CAATTTCAAG GGGTGCTGAC ATTAAGTTGG CTGCCAATTG
GATAATGGGT GATATTGCTG CCTATATGAA AAATGAAAAG CTTTCTATTA ATGAGATCAA ACTTATGCCA GAAGAGCTAG
TAGAGCTGAT AGCTTCCATT AAAGGTGGGA CTATCAGTGG AAAGATTGGA AAGGAGGTAA GCATTGCTT CTTNACTGA
TGCCACTTTC ATGTTCAAAC ATTTGTTAGT AATCCTGTCT ATTTATTTTC ATGGAAGAAT TTTACTAGCT ATTATTCTCC
ACTGTGTAGA TGTGATTTTA TAGTTTGTTT GGATATATAA TTATTGTTTC TGTTTTTTT TTTAATCCAA ACTTTATAAT
CTTTCCAAGT GCTTTTCTCC TCCCTGTCTT TTCCTCTAC CACACACACA CACACACACA CACA CTCATA CAGAAAAAGG
AAAAAGGAAA GAAAGAGAAC AGGAGATATA AAAACCTTTT TTCTTTATCA ATTAGATAAT TAGTTATAAA AGTTTTTCTC
CTGCTTCTTA TCCCTCCGCT AAATGCCTGA TTAACTTTCT GCTGGTAAAG ATTTAAAATA ACTTGTTAAT TTTGGACATG

Primer FOR



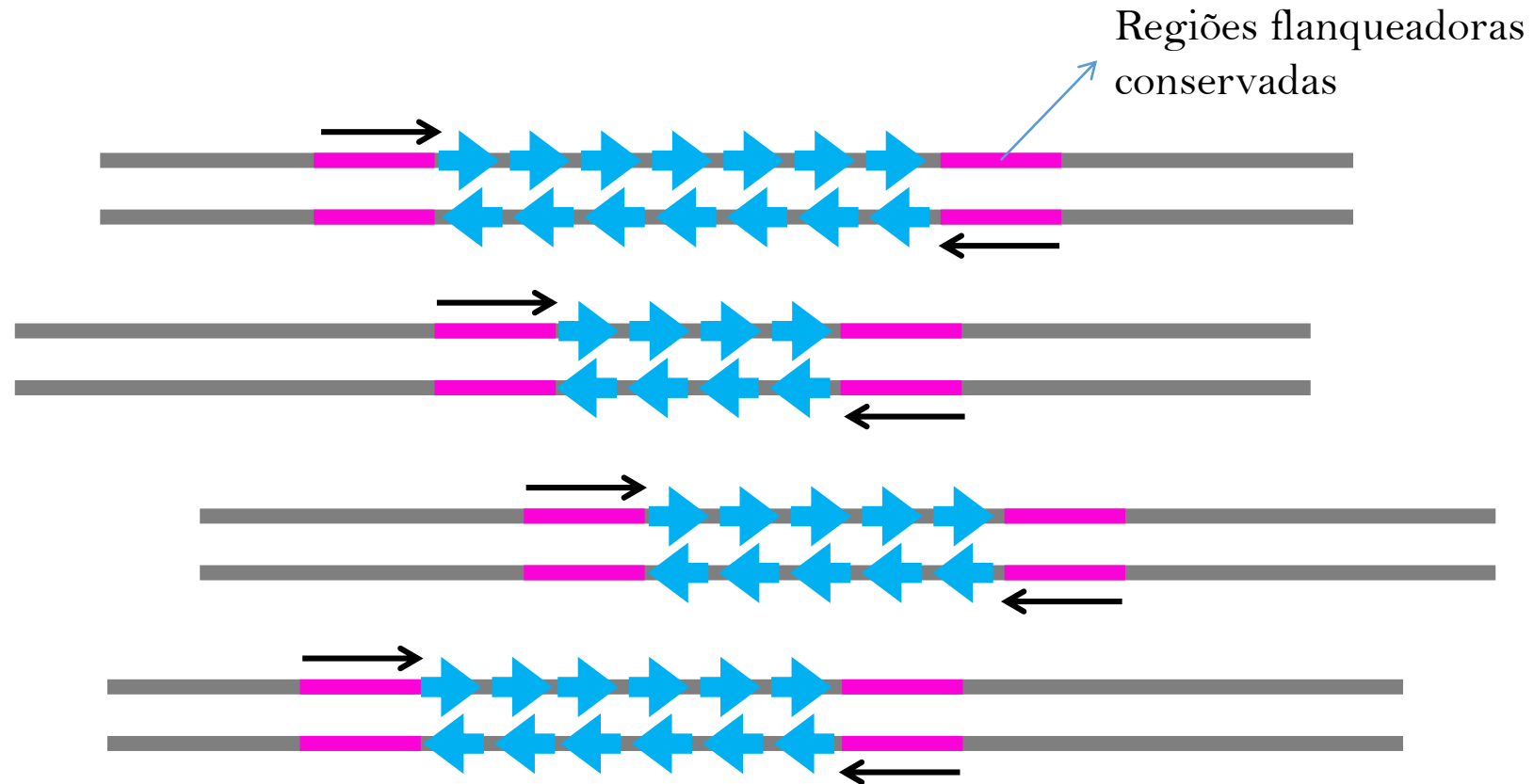
NNNNNNNNNNNNNNNNNNNNNNNNNN CACACACACACACACACAC NNNNNNNNNN NNNNN
NNNNNNNNNNNNNNNNNNNNNNNNNN GTGTGTGTGTGTGTGTG NNNNNNNNNNNNNNNNNNN

Repetição CA/GT



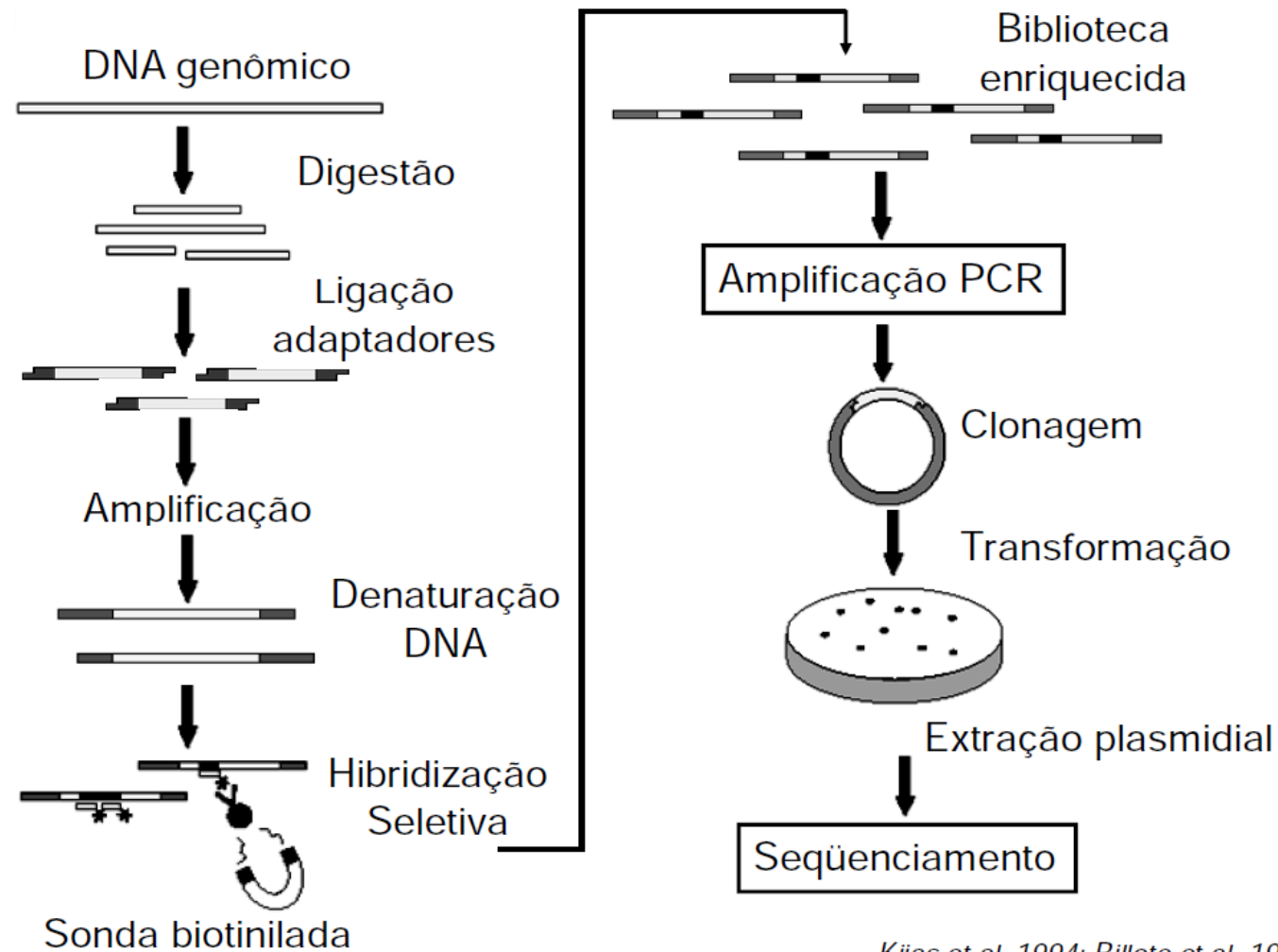
Primer REV

MICROSSATÉLITES (SSR)

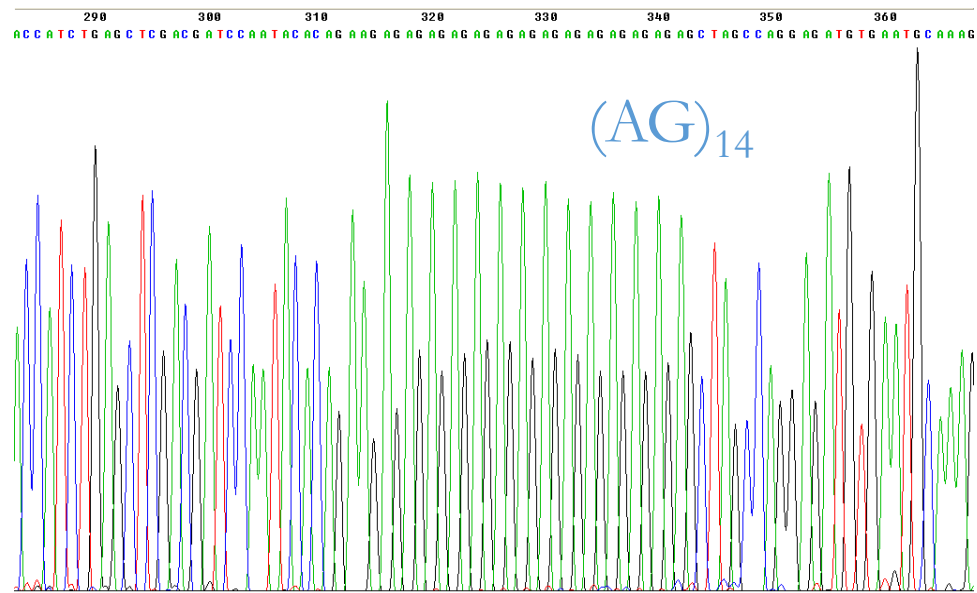
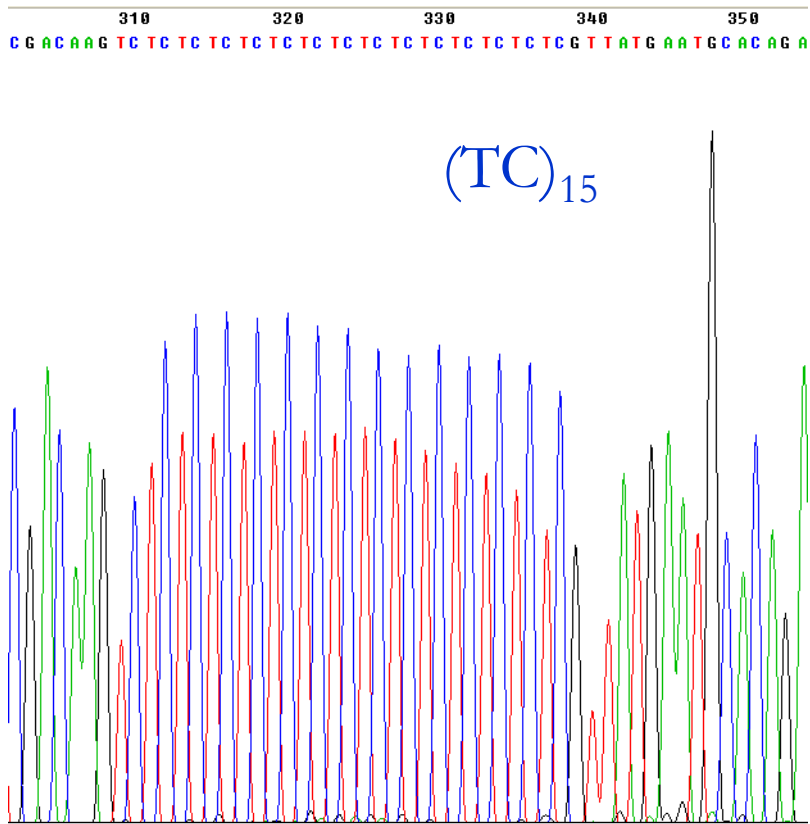


MICROSSATÉLITES (SSR)

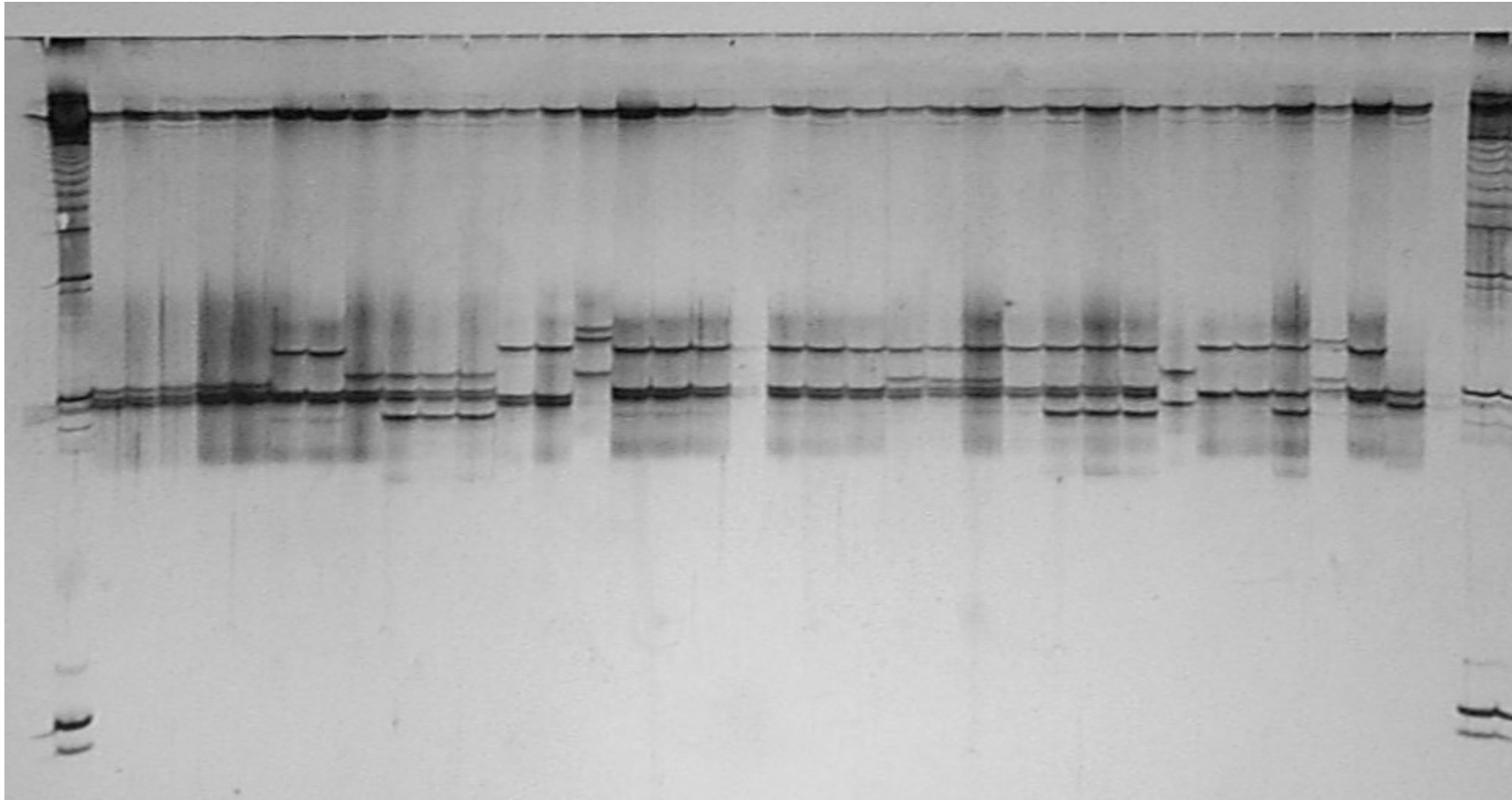
Biblioteca genômica enriquecida com microssatélites



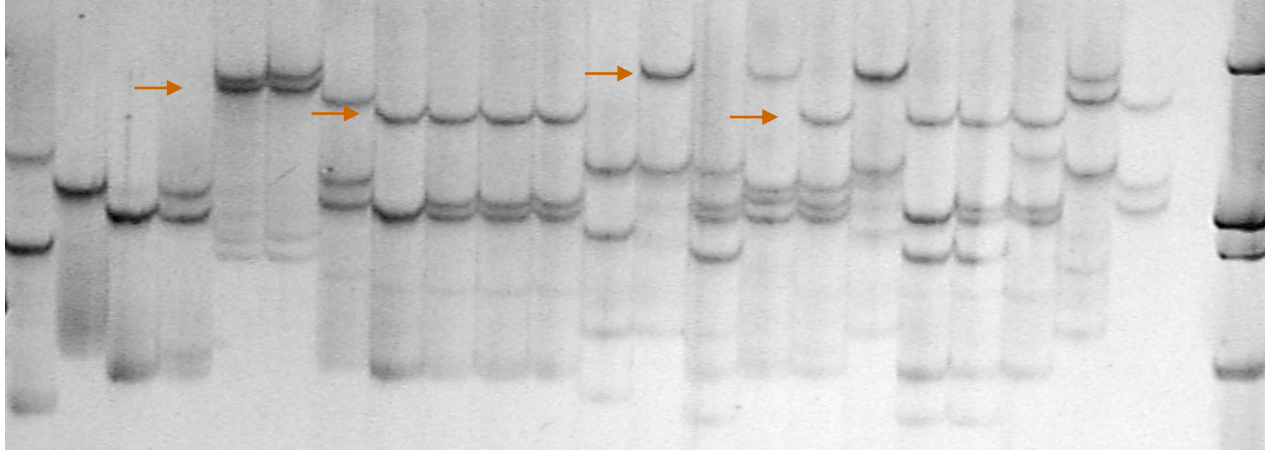
Microsatélites (SSR)



MICROSSATÉLITES

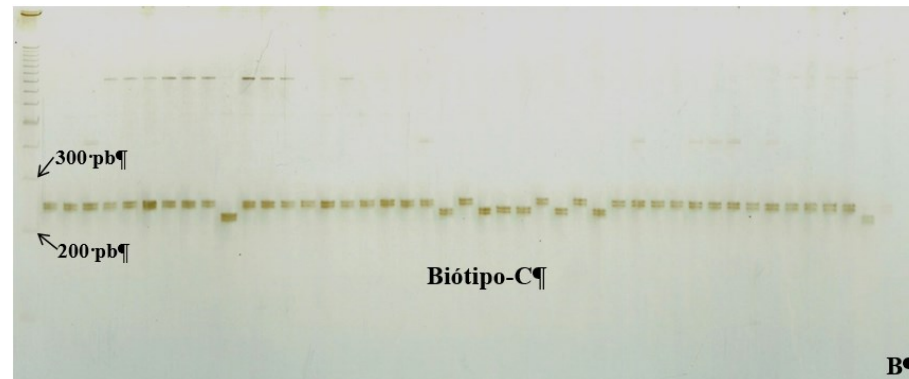
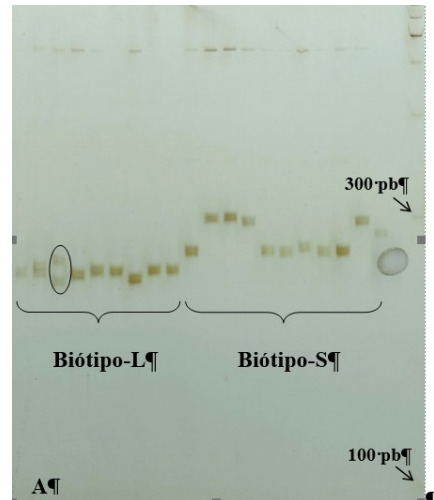


MICROSSATÉLITES



Seta: alelos genoma B

Musa balbisiana



MICROSSATÉLITES

Vantagens

- ❑ Baseiam-se em PCR
- ❑ Altamente reprodutíveis
- ❑ Codominantes e multialélicos

Desvantagens

- ❑ Há necessidade de informações de sequências oriundas de bibliotecas genômicas ou de cDNA
- ❑ Elevado custo inicial

SNP

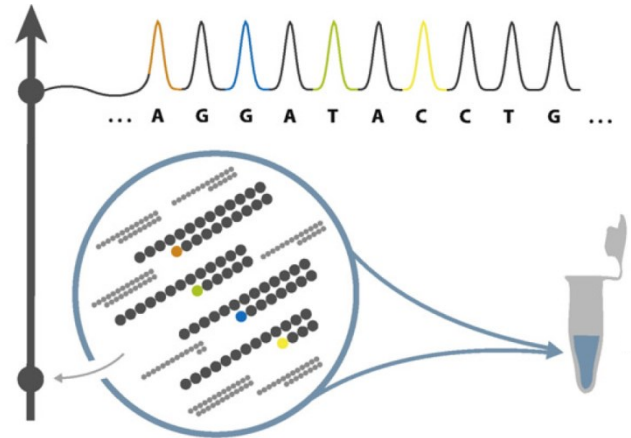
Single Nucleotide Polymorphism

Polimorfismo de base única

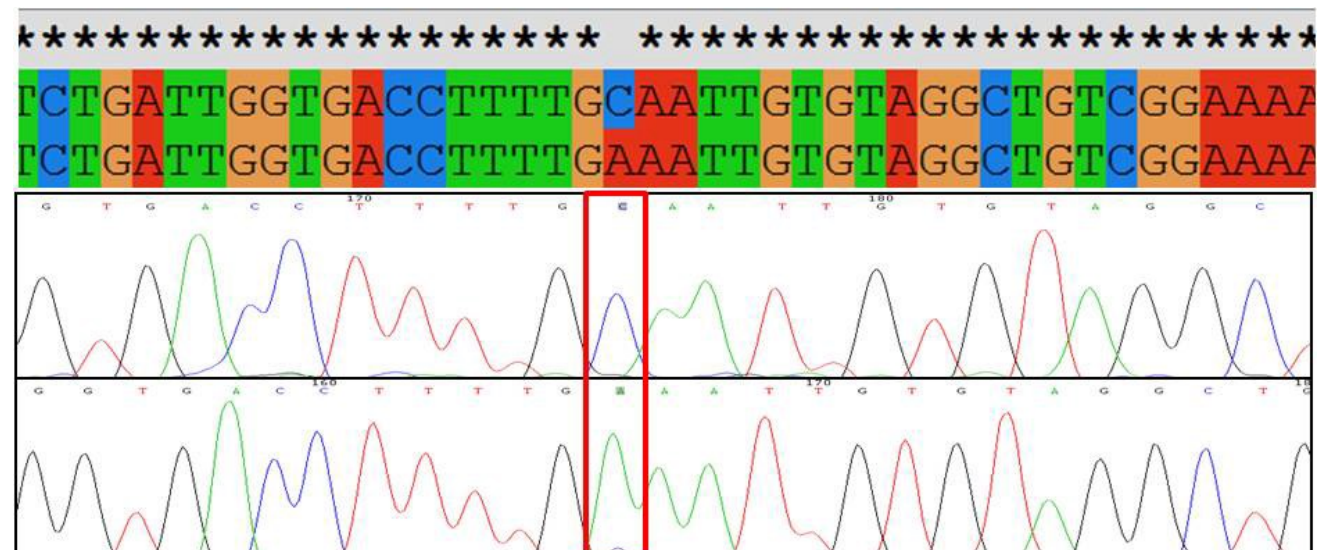
- Polimorfismo resultante da alteração de uma única base
- Abundantes no genoma
- Codominante e bialélico
- Alta processividade – baixo custo por unidade de dado
- Há diversas maneiras de se avaliar a variação:
 - Alinhamento e comparação de sequências
 - Métodos baseados em chips (*arrays*) – *Infinium*, *Axiom*, *GoldenGate*,..
 - Plataformas de genotipagem (*TaqMan*, *KASP*, *Fluidigm Dynamic Arrays*, *LGC KASP*, *ArrayTape*,... .
 - NGS – sequenciamento do genoma, *Genotype-by-Sequencing*, *RAD-Seq*, *DArTseq*, *RESTseq*, *RESCAN*,...

SNP

Sequenciamento



Alinhamento de
sequências



SNP

Target



Subgroup	id5000197	id5000200	id5000204	id5000205	id5000025	id5000209	id5000217	id5000223
IND	A	T	G	T	T	A	A	C
IND	A	T	G	T	T	A	A	G
IND	A	T	G	N	T	A	N	C
IND	A	T	G	T	T	A	A	C
IND	A	T	G	T	T	A	A	C
IND	A	C	G	T	T	A	A	G
IND	A	T	G	T	T	A	A	G
IND	A	T	G	T	T	A	N	C
IND	A	T	G	T	T	A	A	G
IND	A	T	G	T	T	A	A	G
IND	A	T	G	T	T	A	A	C
AUS	A	C	G	T	T	A	A	G
AUS	A	C	A	C	T	A	C	C
AUS	C	C	G	T	C	G	C	C
AUS	A	C	G	T	T	A	A	G
AUS	C	C	G	T	C	G	C	C
AROMATIC	A	C	G	T	T	A	A	G
AROMATIC	A	C	G	T	T	A	A	G
AROMATIC	A	C	G	T	T	A	A	G
AROMATIC	A	C	G	T	T	A	A	G
AROMATIC	A	C	G	T	T	A	A	G
AROMATIC	A	C	G	T	T	A	A	G
TEJ	A	C	A	C	T	A	C	C
TEJ	A	C	A	C	T	A	C	C
TEJ	A	C	A	C	T	A	C	C
TEJ	A	C	A	C	T	A	C	C
TEJ	A	C	A	C	T	A	C	C
TEJ	A	C	A	C	T	A	C	C

SNP

Table 1. Examples of high-throughput SNP genotyping technologies.

Genotyping Platform	Technology	SNP x sample combinations	Capital investment	Cost per sample	Advantages
Illumina Infinium iSelect HD	Fixed array	3,072 – 700K SNPs x 24 samples	High (iScan)	Moderate to high	Highly multiplexed
Affymetrix Axiom	Fixed array	50K SNPs x 384 samples; 650K SNPs x 96 samples	High (GeneTitan)	Moderate to high	Highly multiplexed
Douglas Array Tape	Flexible, PCR-based	1 SNP/sample x 76,800 reactions/reel	Very High (Nexar, Soellex, Araya)	Very low	Ultra high-throughput
Fluidigm Dynamic Arrays	Flexible, PCR-based	96 SNPs x 96 samples; 24 SNPs x 192 samples	Moderate (IFC Controller, FC1, EP1)	Low	High-throughput
RE-based GBS	Genotyping by sequencing	~10K-100K SNPs x 96 or 384 samples	Low to moderate (NGS outsourced or in-house)	Low to moderate	Lots of data relative to the cost
Amplicon sequencing	Genotyping by sequencing	Variable (e.g. 20-500 SNPs x 48-384 samples)	Low to moderate (NGS outsourced or in-house)	Low to moderate	Multiple targeted loci at once

SNP

Vantagens:

- ❑ Abundante no genoma
- ❑ Cobertura em alta densidade do genoma
- ❑ Gera muita informação em pouco tempo
- ❑ Marcador codominante (bialélico)

Desvantagens:

- ❑ Elevado custo total (mas por dado, é barato)

COMPARAÇÃO DOS MARCADORES

	RAPD	AFLP	SSR	SNP
Base genética	PCR com <i>primers</i> arbitrários	Digestão e PCR com <i>primers</i> seletivos	PCR com <i>primers</i> específicos	Sequenciamento
Tipo de herança	Dominante	Dominante	Codominante	Codominante
Número de locos	Vários	Vários	Único	Único
Número de alelos	Dois	Dois	Vários	Dois

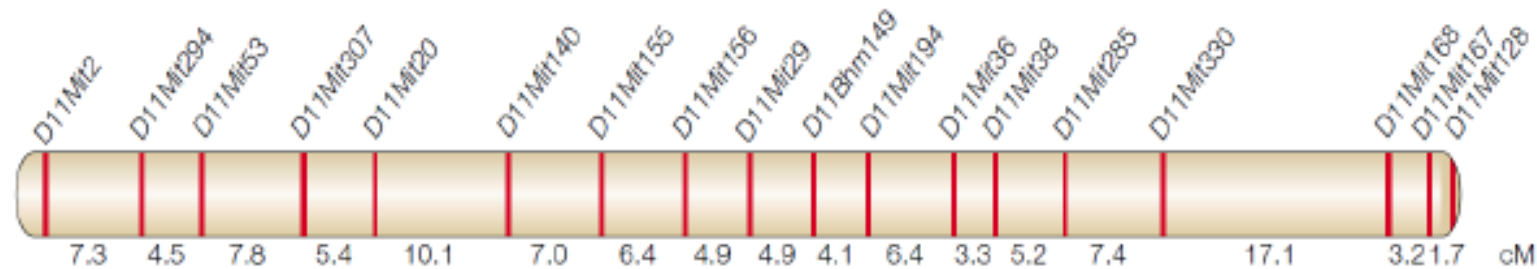
Uso dos Marcadores no Melhoramento de Plantas

Construção de Mapas Genéticos

Mapa genético: representação de ordem e distância entre marcadores genéticos equivalente ao cromossomo

Ligação genética

- A herança conjunta de diferentes locos dá-se por conexão física

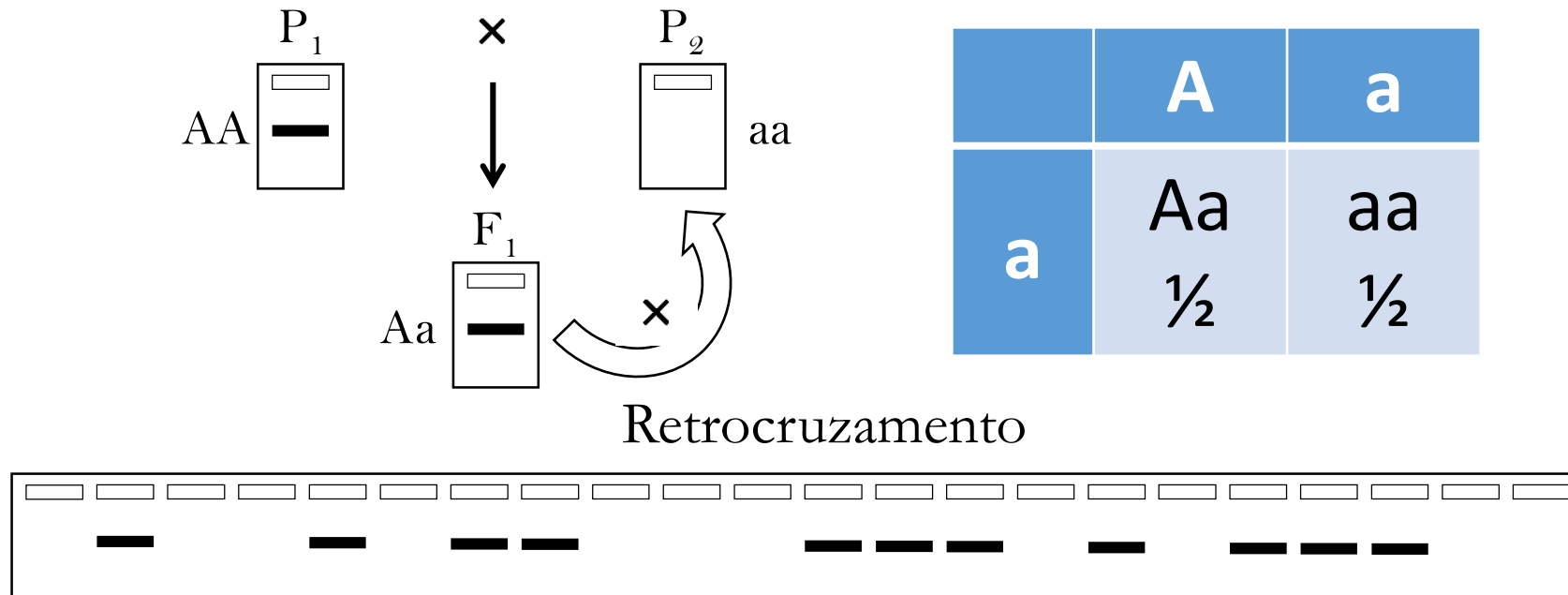


Mus musculus L.

Doerge (2002) *Nature Reviews* 3:43-52 adaptado de:
Butterfield et al. (1999) *The Journal of Immunology* 162:3096-3102

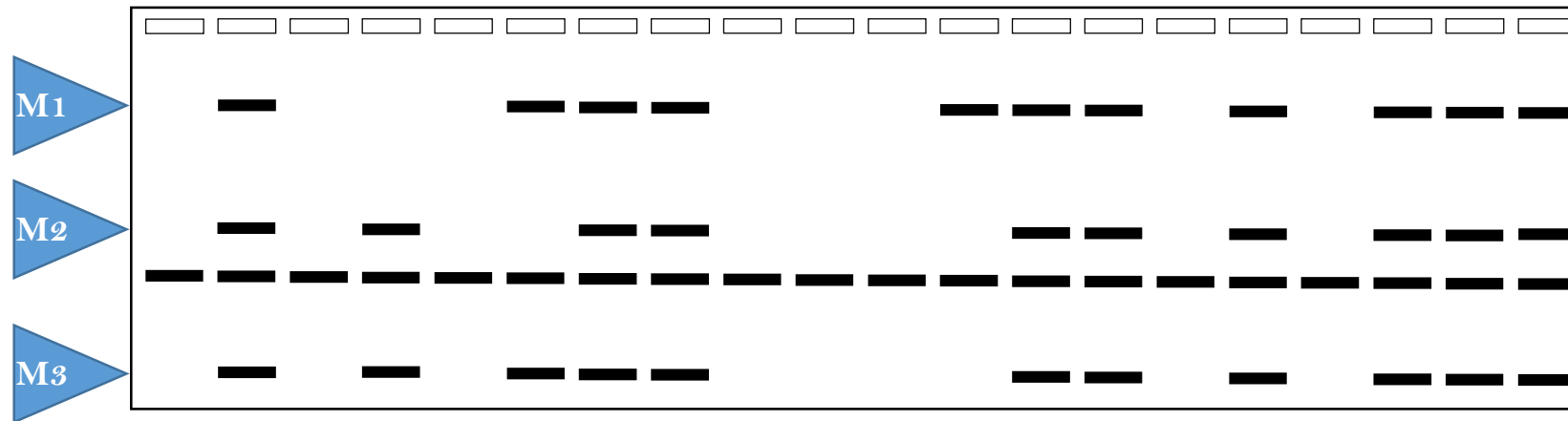
Construção de Mapas Genéticos

Consiste na avaliação de uma população segregante (retrocruzamentos, F_2 , por exemplo) por meio de vários locos marcadores



Construção de Mapas Genéticos

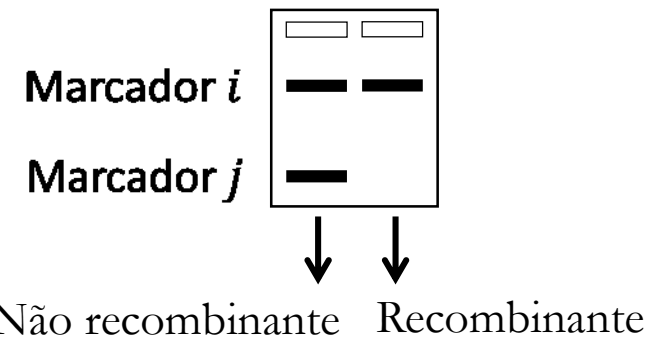
População de retrocruzamento genotipada com três locos AFLP



- Cálculo da fração de recombinação:

$$r_{ij} = \frac{n_{\text{recombinantes}}}{n}$$

Comparando dois locos polimórficos



Construção de Mapas Genéticos

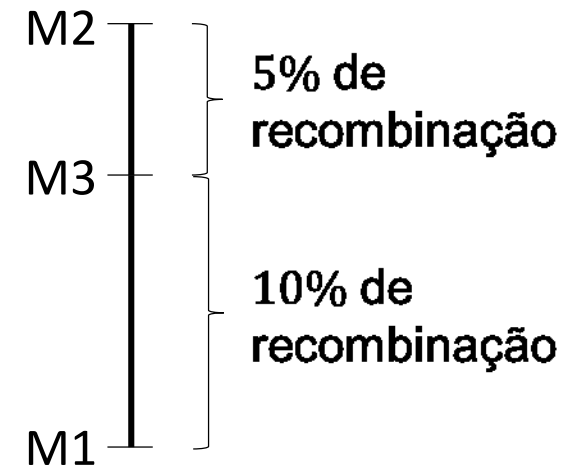
Leitura do gel

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
M1	0	1	0	0	0	1	1	1	0	0	0	1	1	1	0	1	0	1	1	1
M2	0	1	0	1	0	0	1	1	0	0	0	0	1	1	0	1	0	1	1	1
M3	0	1	0	1	0	1	1	1	0	0	0	0	1	1	0	1	0	1	1	1

$$r_{1,2} = \frac{3}{20} = 0,15$$

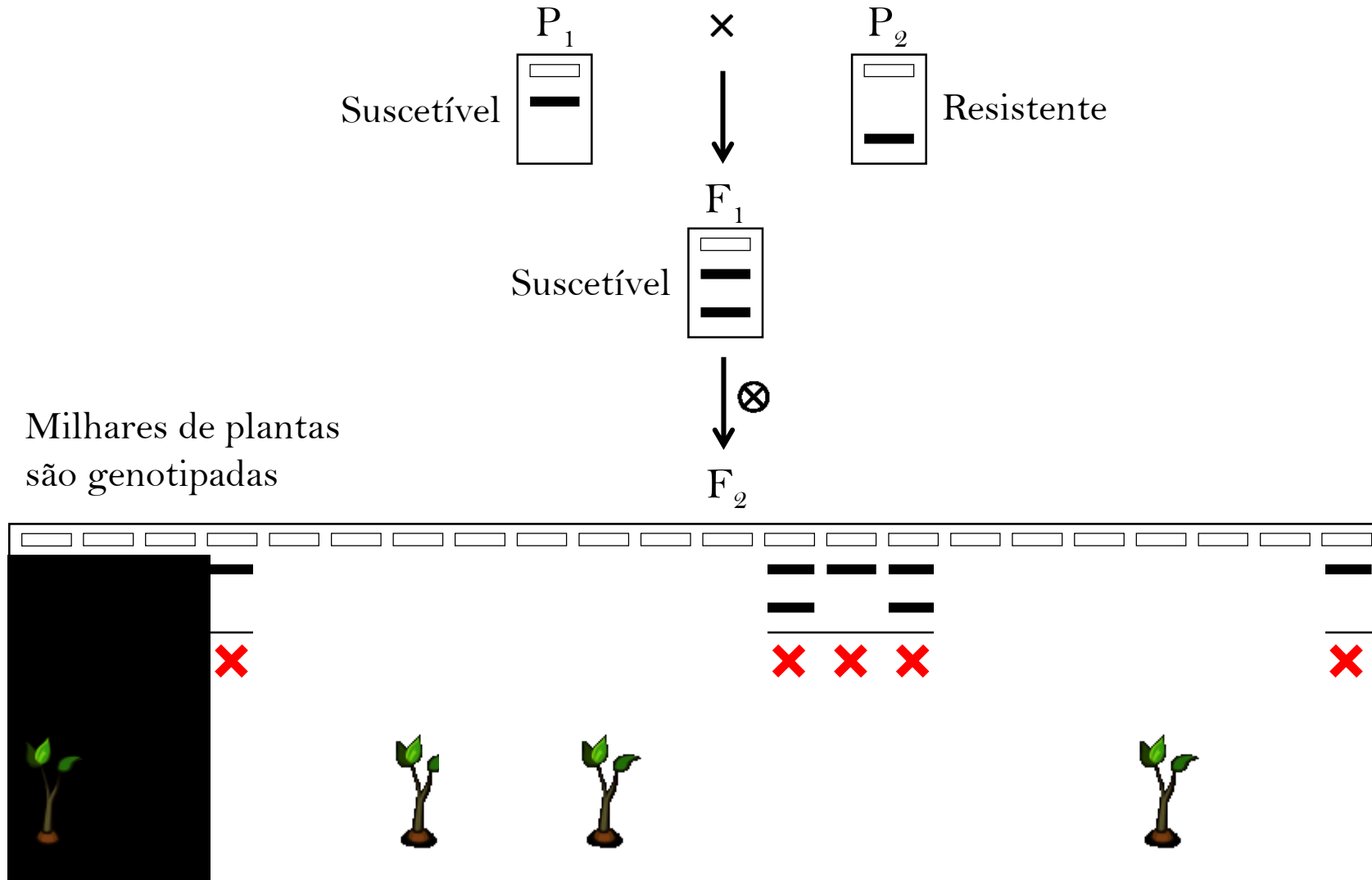
$$r_{1,3} = \frac{2}{20} = 0,10$$

$$r_{2,3} = \frac{1}{20} = 0,05$$



Mapa genético

Seleção Assistida por Marcadores



A seleção fenotípica é baseada em marcadores de DNA

Marcadores Dominantes

- Cálculo de similaridade genética: coeficiente de Jaccard

$$S_{ij} = \frac{a}{a + b + c}$$

Em que:

- a = nº de locos com alelos presentes nos indivíduos i e j
- b = nº de locos com alelos presentes em i e ausentes em j
- c = nº de locos com alelos ausentes em i e presentes em j

	Indivíduos	
	i	j
Caso a		
Caso b		
Caso c		

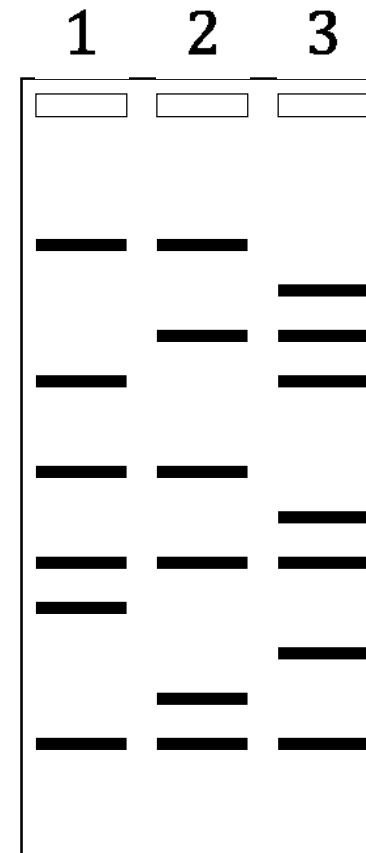
Marcadores Dominantes

$$S_{12} = \frac{a}{a + b + c} \Rightarrow$$

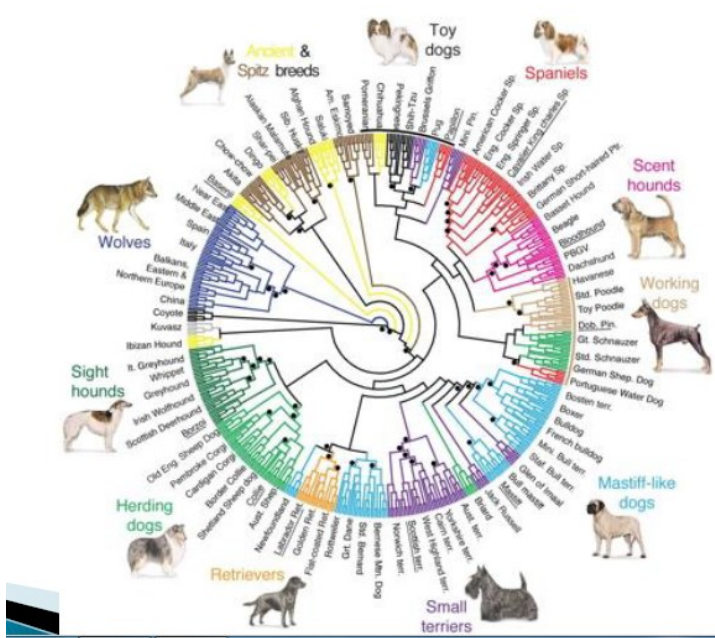
$$S_{13} = \frac{a}{a + b + c} \Rightarrow$$

$$S_{23} = \frac{a}{a + b + c} \Rightarrow$$

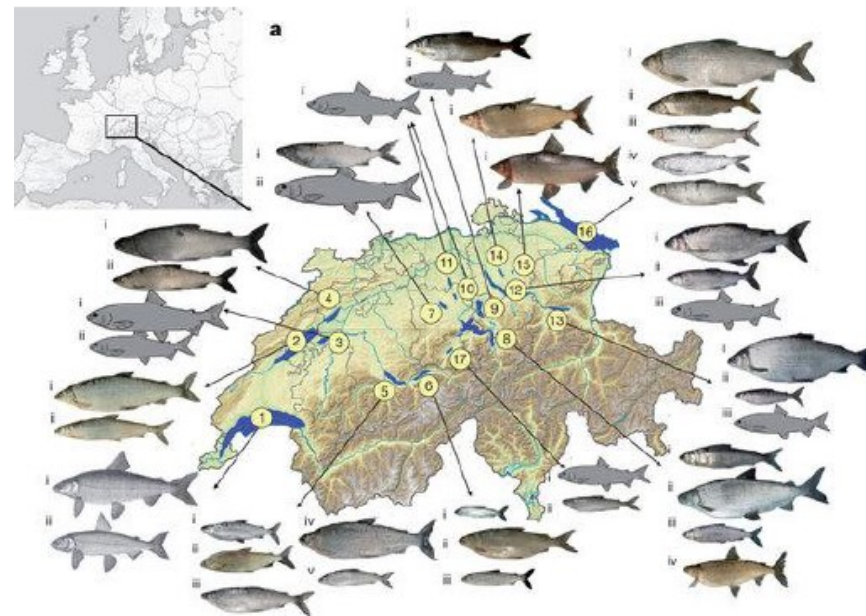
- Conclusão: são mais similares os indivíduos e mais diferentes os indivíduos



Uso de Marcadores em Estudos de Diversidade e Conservação



Filogenia



Distribuição geográfica

Por que usar marcadores?

Fornece informações a respeito de:

- ❑ Diversidade genética
- ❑ Endogamia e sistemas de cruzamento
- ❑ Fluxo gênico
- ❑ Paternidade
- ❑ Sexagem

Auxilia na definição de estratégias de manejo e conservação

Ajuda a organizar coleções de germoplasma

Estudo Dirigido

1. O que são marcadores moleculares?
2. Importância de enzimas (de restrição, ligases e polimerases), clonagem molecular e eletroforese na obtenção de marcadores.
3. Definir marcadores moleculares.
4. Que são marcadores dominantes e codominantes?
5. Quais os tipos de marcadores moleculares?
6. Qual a utilidade de marcadores moleculares no mapeamento genético?
7. Qual a utilidade de marcadores moleculares nos estudos de diversidade genética e conservação das espécies?

