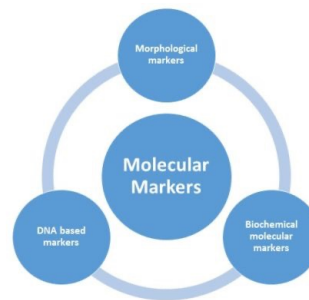
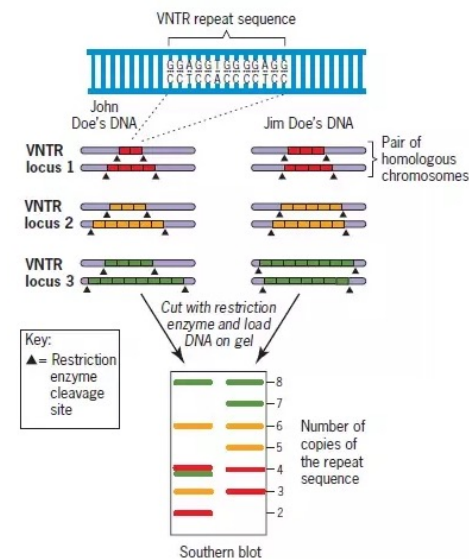


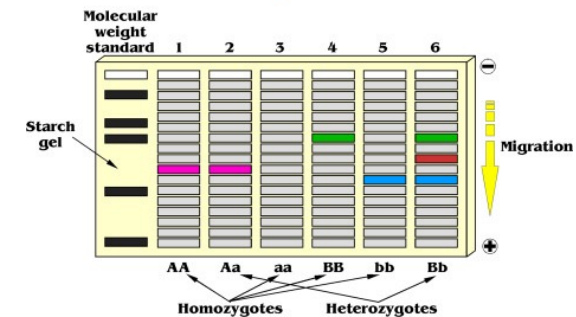
GENÉTICA E MELHORAMENTO DE PLANTAS

Licenciaturas em Eng^a Agronómica

Marcadores moleculares



Bibliography:
An Introduction to Genetic Analysis 11th Ed
Genetics: From Genes to Genomes 6th Ed



Manuela Gomes da Silva
manuelasilva@isa.ulisboa.pt

Marcadores Genéticos

Como medir a diversidade genética?

- Elevado polimorfismo
 - Detecção de grande número de loci
- Co-dominantes e de Hereditariedade simples
- De preferência ligados à característica de interesse (gene)
- Reprodutíveis
- Estabilidade em diferentes ambientes
- Robustos
- Baratos
- Fáceis de usar

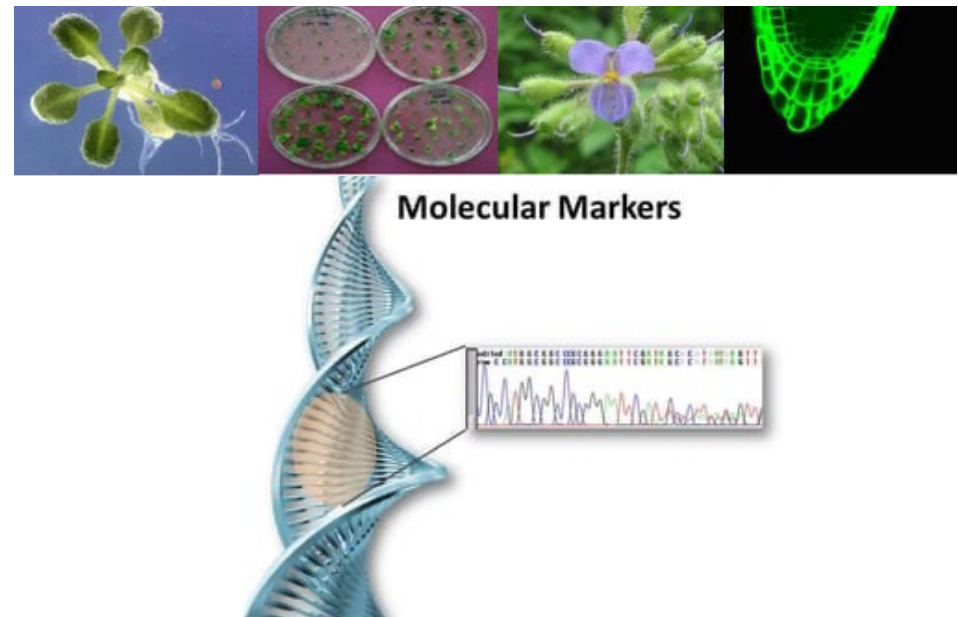
Tipos de Marcadores

Fenótipo

- Morfológicos
- Bioquímicos
(**proteínas**, enzimas, etc)

Genoma

- Cromossômicos
- Moleculares
(sequência de **DNA**)



Marcadores Morfológicos

As plantas podem ser selecionadas de acordo com características/**marcadores morfológicos**

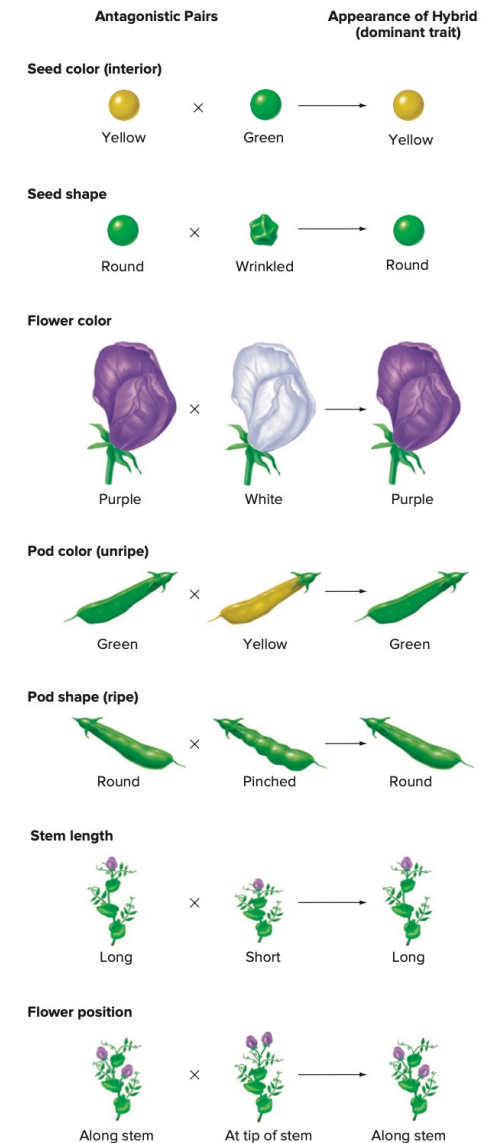
Ex: características da ervilha estudadas por Mendel

Crescimento vegetativo / reprodutivo

Tamanho, forma, cor de sementes

Tamanho, forma, cor de frutos

Hábitos de crescimento



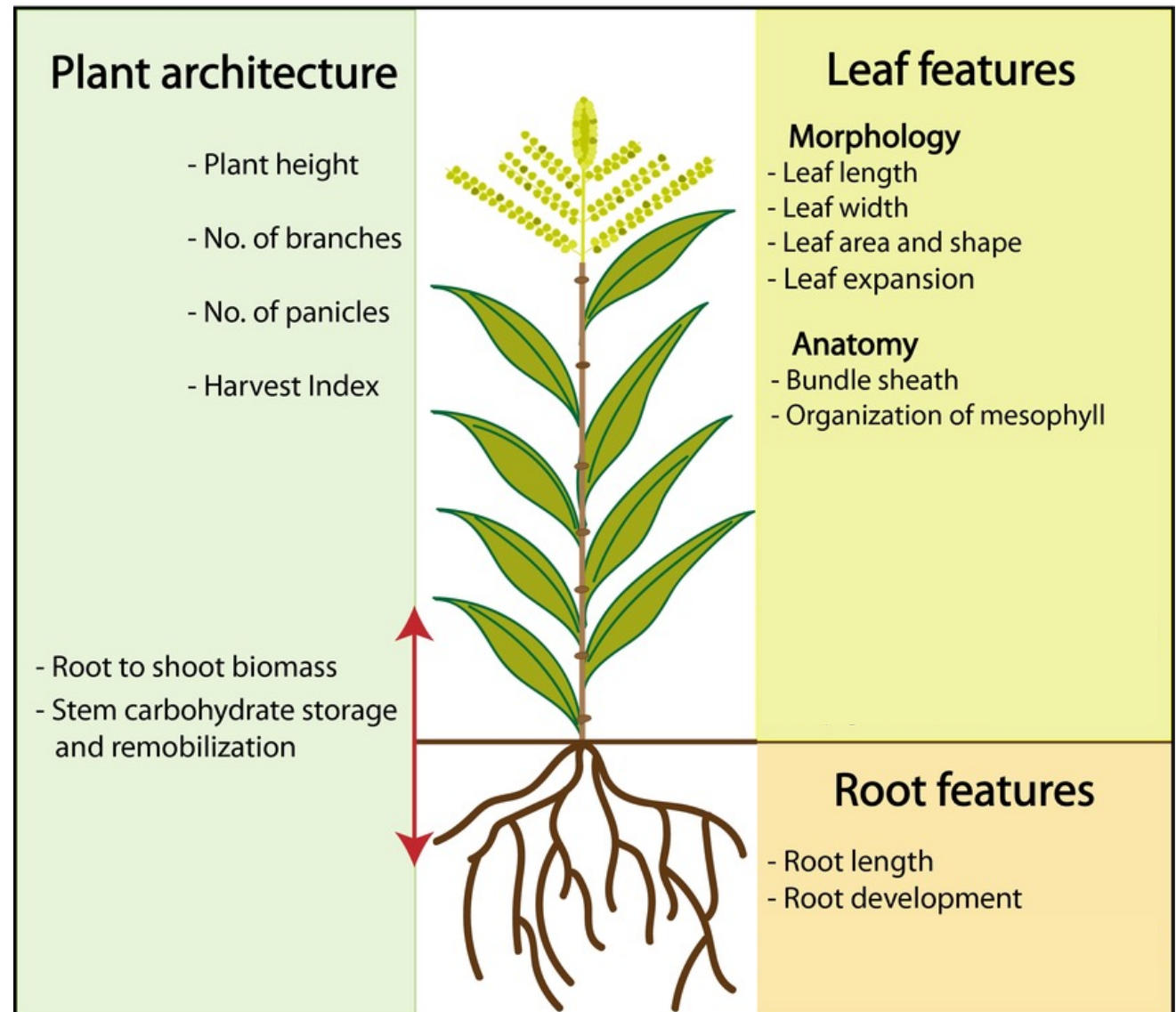
Morfológicos

Plant **morphological** traits relevant for high **yield**

Plant growth and development characteristics

Plant traits

- plant height
 - number of fruits/seeds
 - number of branches
 - root length
- crop performance and yield



Marcadores Morfológicos

Tamanho e peso corporal: o tamanho e o peso corporal são características importantes na produção de carne, leite e trabalho

Taxa de crescimento para a produção de carne

Conformação: forma e estrutura do animal de acordo com a função

Ex: Angus - conformação musculosa e compacta

Holstein - conformação leiteira

Capacidade de **produção de leite** de alta qualidade

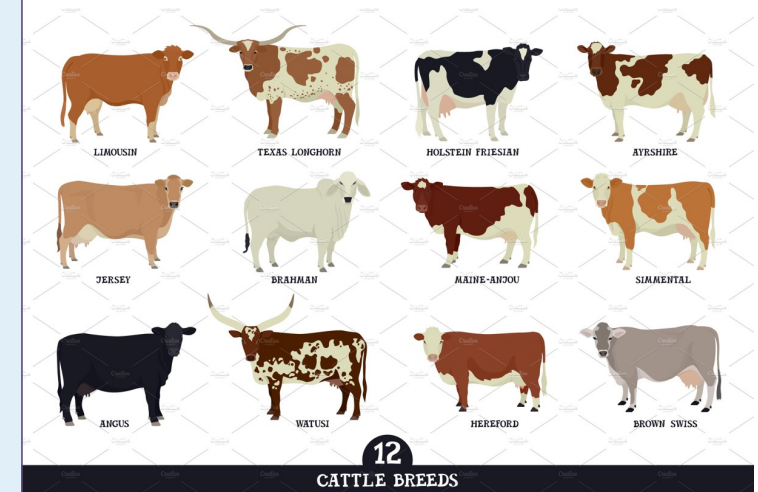
Adaptabilidade a diferentes condições ambientais e disponibilidade de alimentos

Resistência a doenças

Fertilidade importante em vacas reprodutoras

Longevidade para reduzir os custos de reposição do rebanho e garantir a produtividade ao longo do tempo

Ausência de cornos



Marcadores Morfológicos

Produção de carne:

Tamanho do corpo e taxa de crescimento

Produção de ovos:

Taxa de postura, tamanho do ovo, cor da casca do ovo

Plumagem:

Aparência, cor e padrão

Resistência a doenças

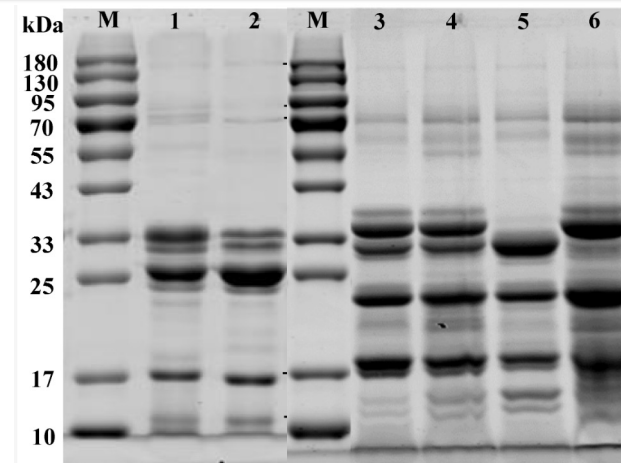
Capacidade de **adaptação** e sobrevivência em diferentes condições



Marcadores Bioquímicos

- **Isoenzimas** (co-dominantes)
- **Desvantagens:**
 - São instáveis
 - Dependem dos estados fisiológicos
 - São influenciados pelo ambiente
 - Têm uma baixa cobertura genómica pois a % de sequências génicas é baixa

Perfil electroforético de proteínas do leite



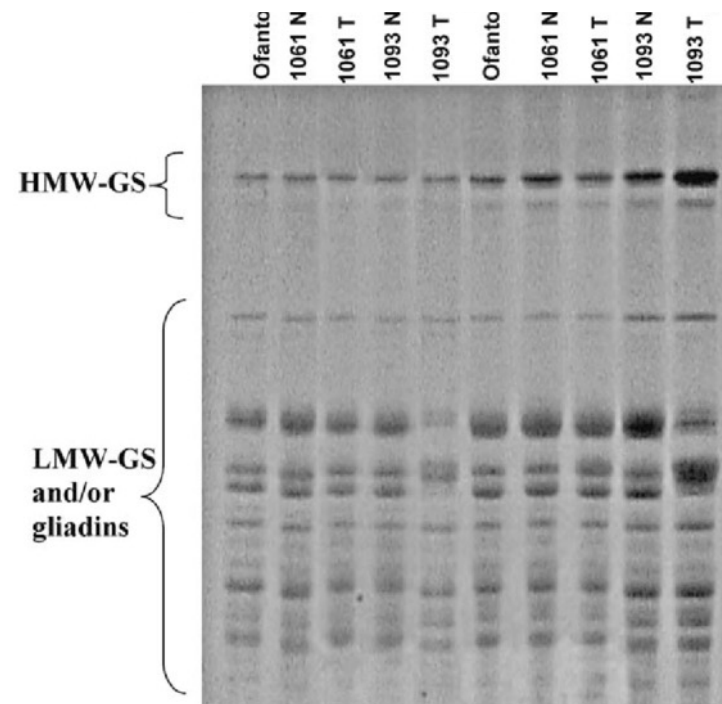
SDS-PAGE proteínas de leite canino, bovino e caprino
M: marcador de peso molecular (1) Vaca Holstein, (2) Cabra Saanen, (3) Golden Retriever, (4) Labrador, (5) Caucasian Sheepdog, (6) Rottweiler

Marcadores Bioquímicos

Proteínas de reserva de sementes de **trigo**

Glutenins consist of high-molecular weight (HMW) and low-molecular weight (LMW) subunits assembled into polymers stabilised by inter-chain disulphide and hydrogen bonds.

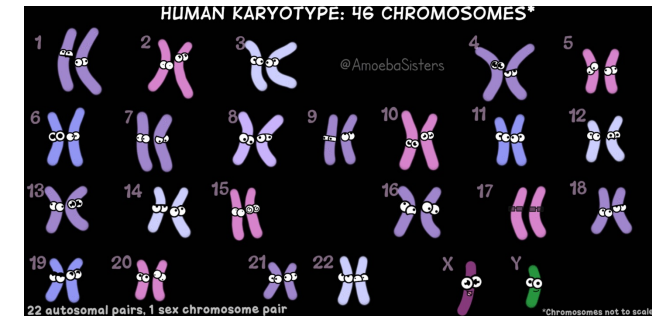
Perfil de gluteninas do grão de trigo



SDS-PAGE of gluten protein fractions from flours extracted from different durum wheat lines

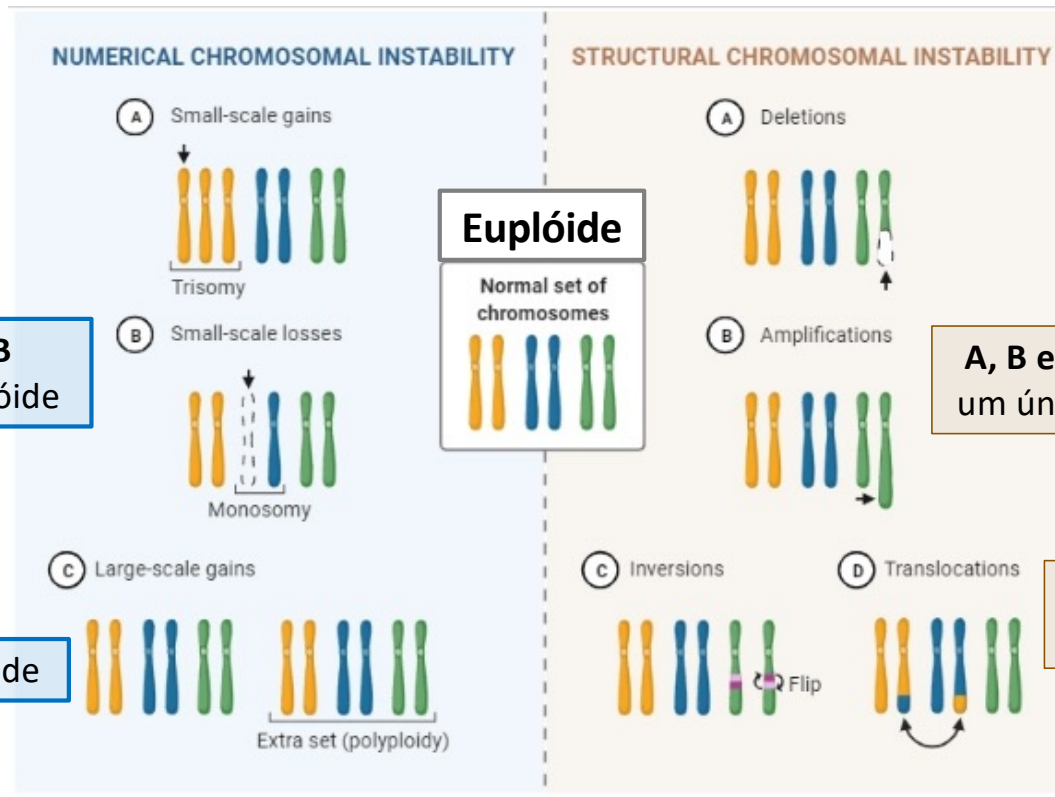
Marcadores Cromossômicos

Mutações Cromossômicas numéricas e estruturais



Desvantagens:

- Baixo polimorfismo
- Difíceis de analisar
(requerem grande especialização)



A e B
Aneuplóide

C poliplóide

A, B e C alterações de
um único cromossoma

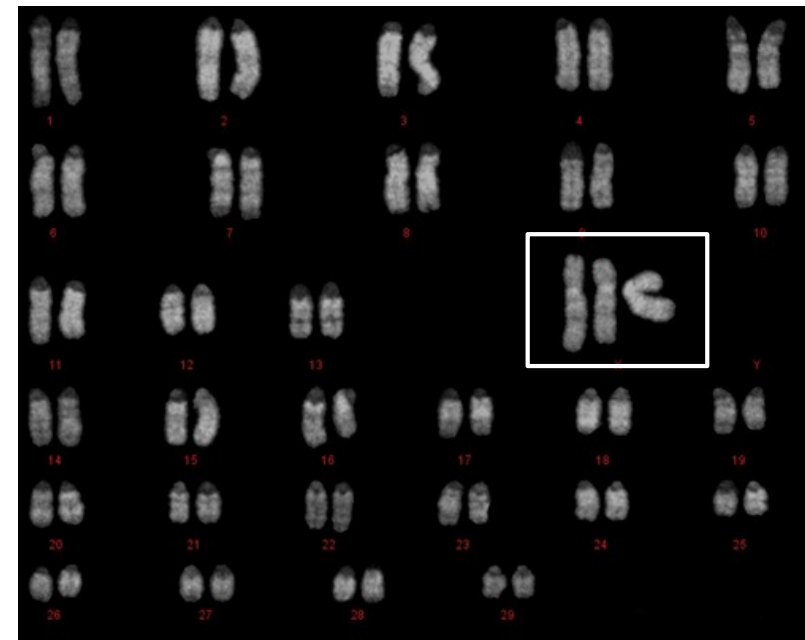
D alterações de
dois cromossomas

Marcadores Cromossômicos

- Variações numéricas
 - **Aneuploidias:**
trissomia, monossomia, nulissomia
- Variações estruturais
 - Delecções, inserções

Sex-chromosome aneuploidy in a large population of juvenile **cow**

<https://onlinelibrary.wiley.com/doi/10.1111/jbg.12866>



Karyotype of a **cow** carrying a trisomy on chromosome X

Marcadores Cromossômicos - Aneuploidias

Impacto na fertilidade animal

Mutação	Efeitos na fertilidade	Incidência em diferentes espécies
X monosomia	Esterilidade	Horse > cat and dog > cattle and pig
XXX trisomia	Infertilidade	Dog, horse
XXY trisomia	Esterilidade	Cat and cattle > dog and horse > pig
XYY trisomia	Infertilidade	Cattle

Szczerbal, I. and Switonski, M., 2016. Chromosome abnormalities in domestic animals as causes of disorders of sex development or impaired fertility. Insights from animal reproduction, pp.207-225.

Marcadores cromossômicos

Wheat aneuploid lines

Examples of two types of aneuploidy

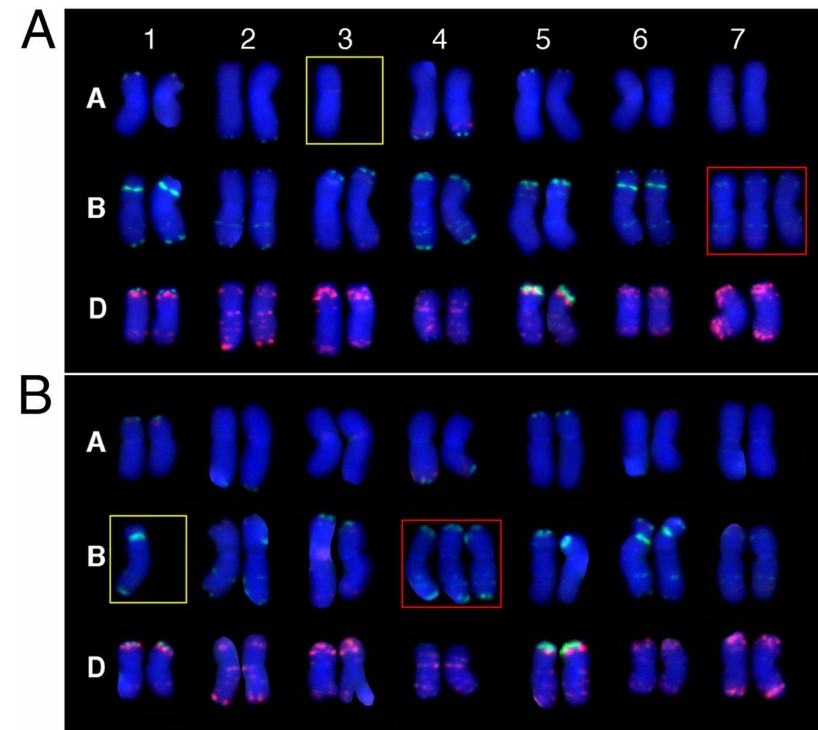
(A) Karyotype of $2n = 42$

genomic constitution of $38 + 1(3A) + 3(7B)$.

(B) Karyotype of $2n = 42$

genomic constitution of $38 + 1(1B) + 3(4B)$

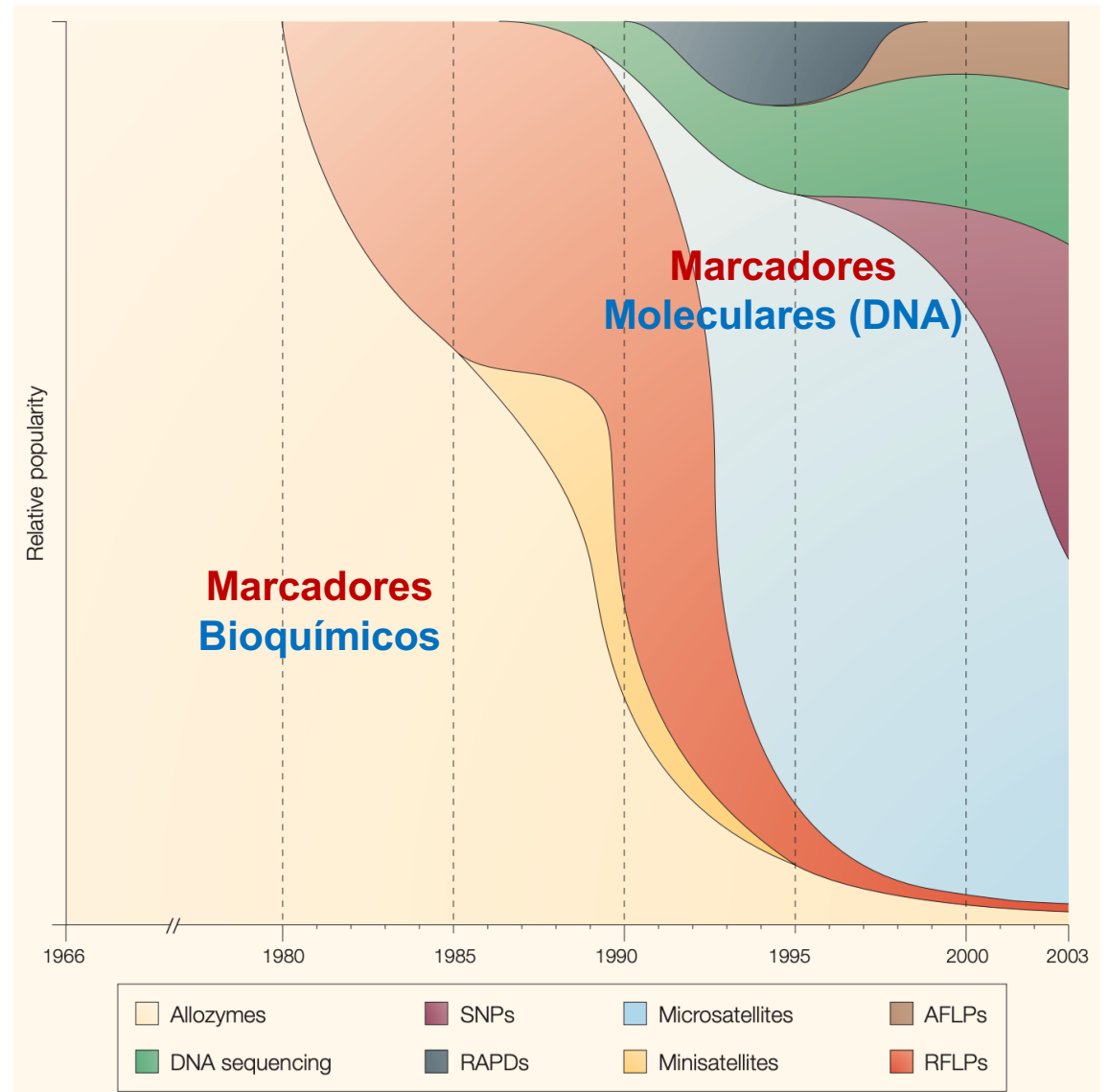
(Yellow and red boxes respectively mark the lost and gained chromosomes)



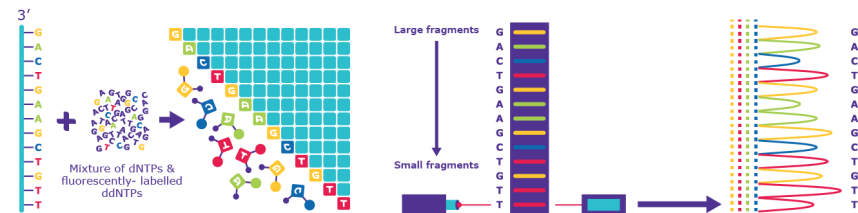
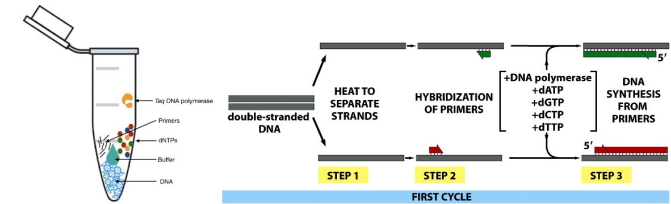
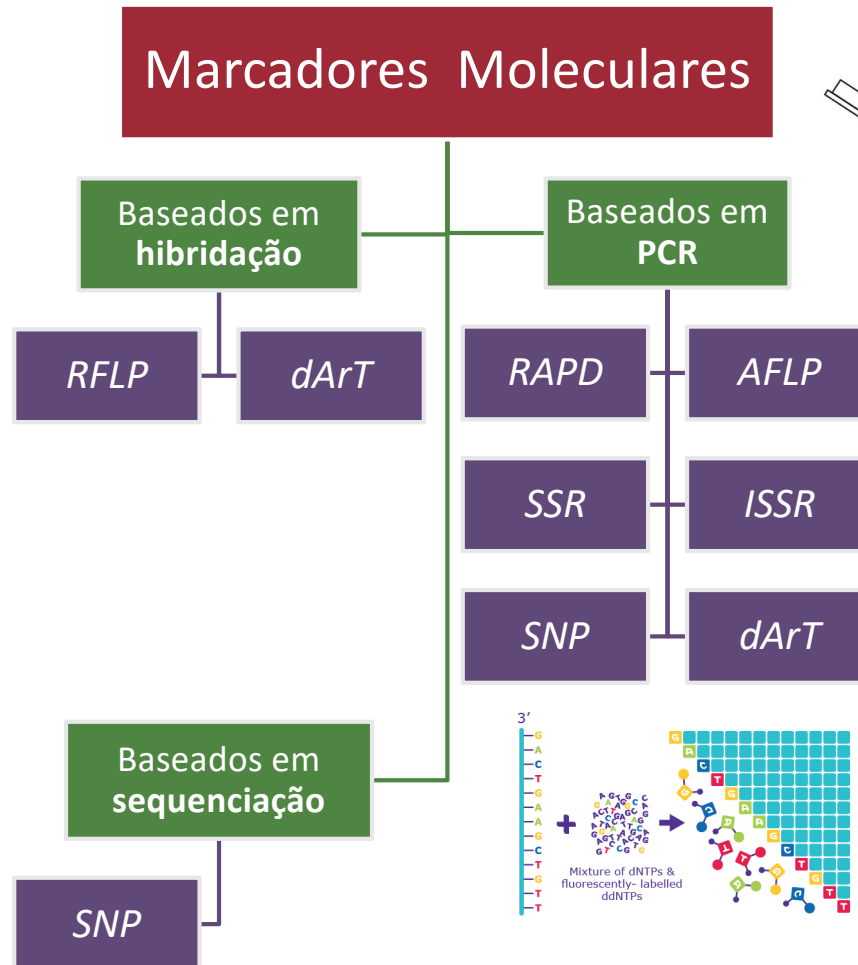
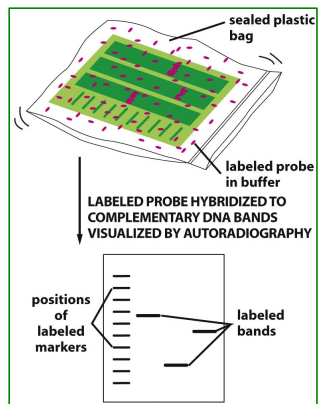
Marcadores moleculares

DNA manipulation techniques
→ **DNA-based**
molecular markers

- Morfológicos
- **Bioquímicos**
proteínas, enzimas, etc
- Cromossómicos
- **Moleculares**
sequência de DNA



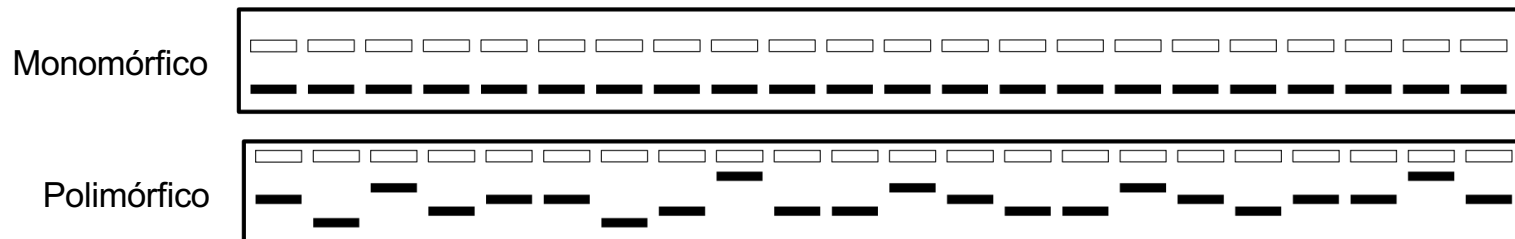
Marcadores Moleculares



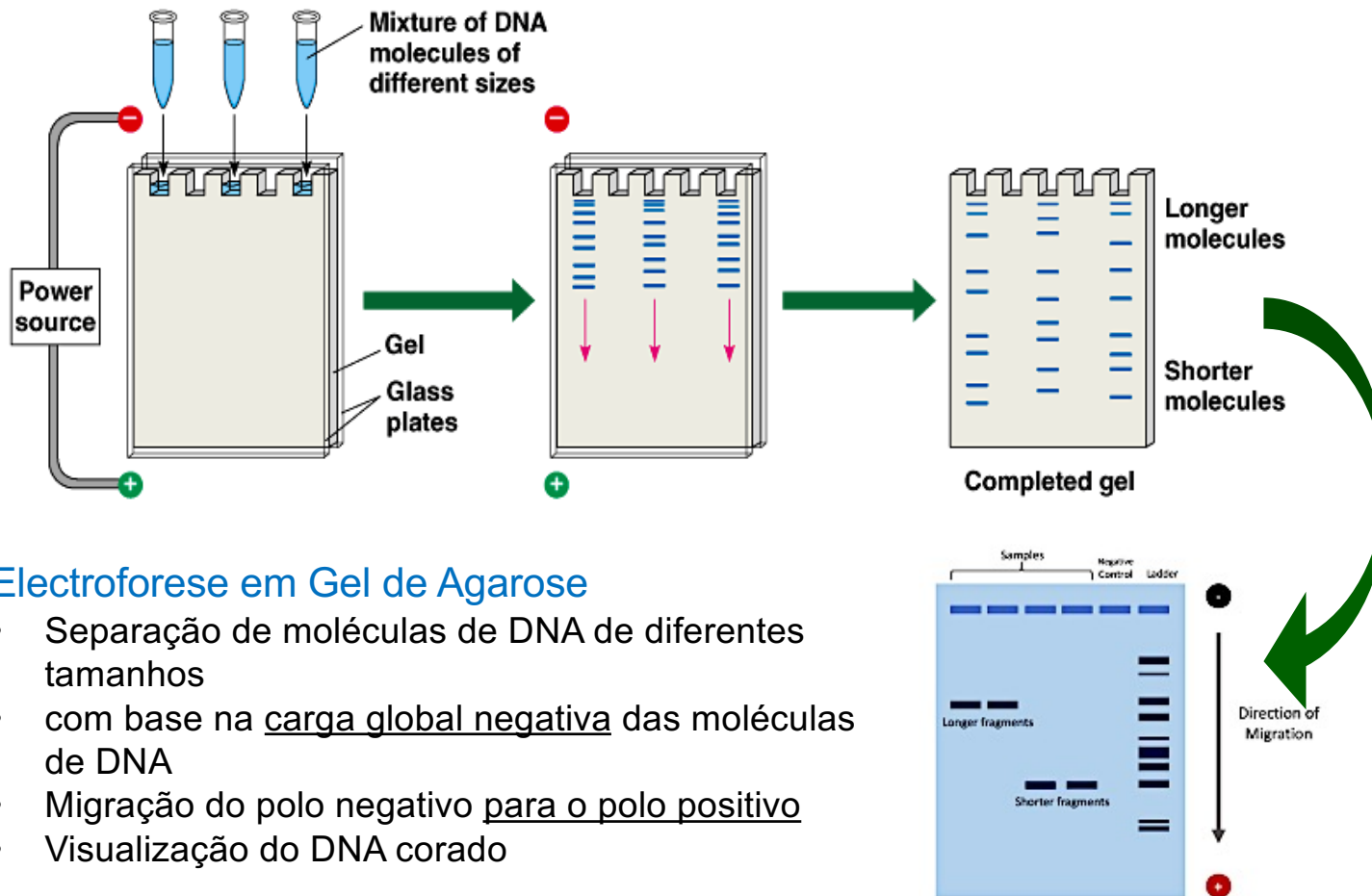
Marcadores moleculares

- Revelam **variação a nível do DNA**
 - Muitos são co-dominantes
 - Podem ser obtidos muito precocemente pois não necessitam que o fenótipo de expresse
 - Muito polimórficos
 - Distribuição aleatória em todo o genoma
 - Metodologias automatizadas

Análise de DNA por electroforese



Electroforese de Ácidos Nucleicos



Electroforese em Gel de Agarose

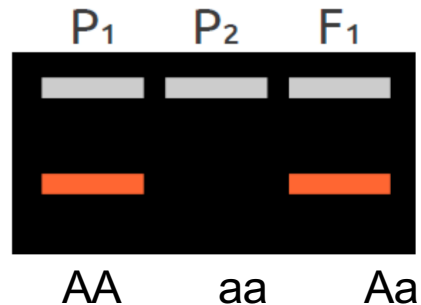
- Separação de moléculas de DNA de diferentes tamanhos
- com base na carga global negativa das moléculas de DNA
- Migração do polo negativo para o polo positivo
- Visualização do DNA corado

Marcadores moleculares

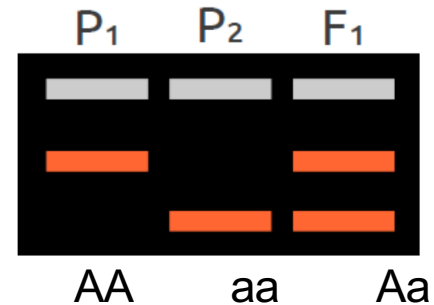
- **Vantagens**

- O **DNA** pode ser isolado de qualquer tecido, em qualquer fase do ciclo de vida
- O **DNA** pode ser armazenado durante longos períodos
- Não é afectado pelo ambiente nem por outras condições

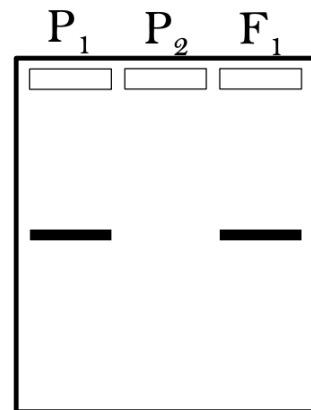
Dominantes
não se identificam
heterozigotos



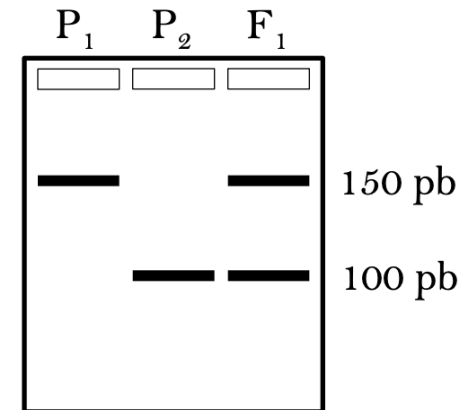
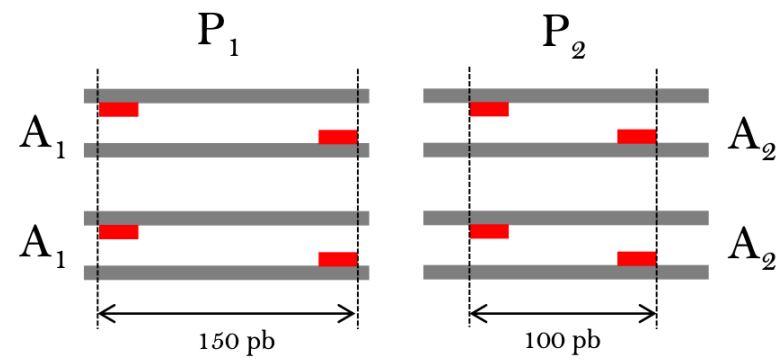
Co-dominantes
identificam-se
heterozigotos



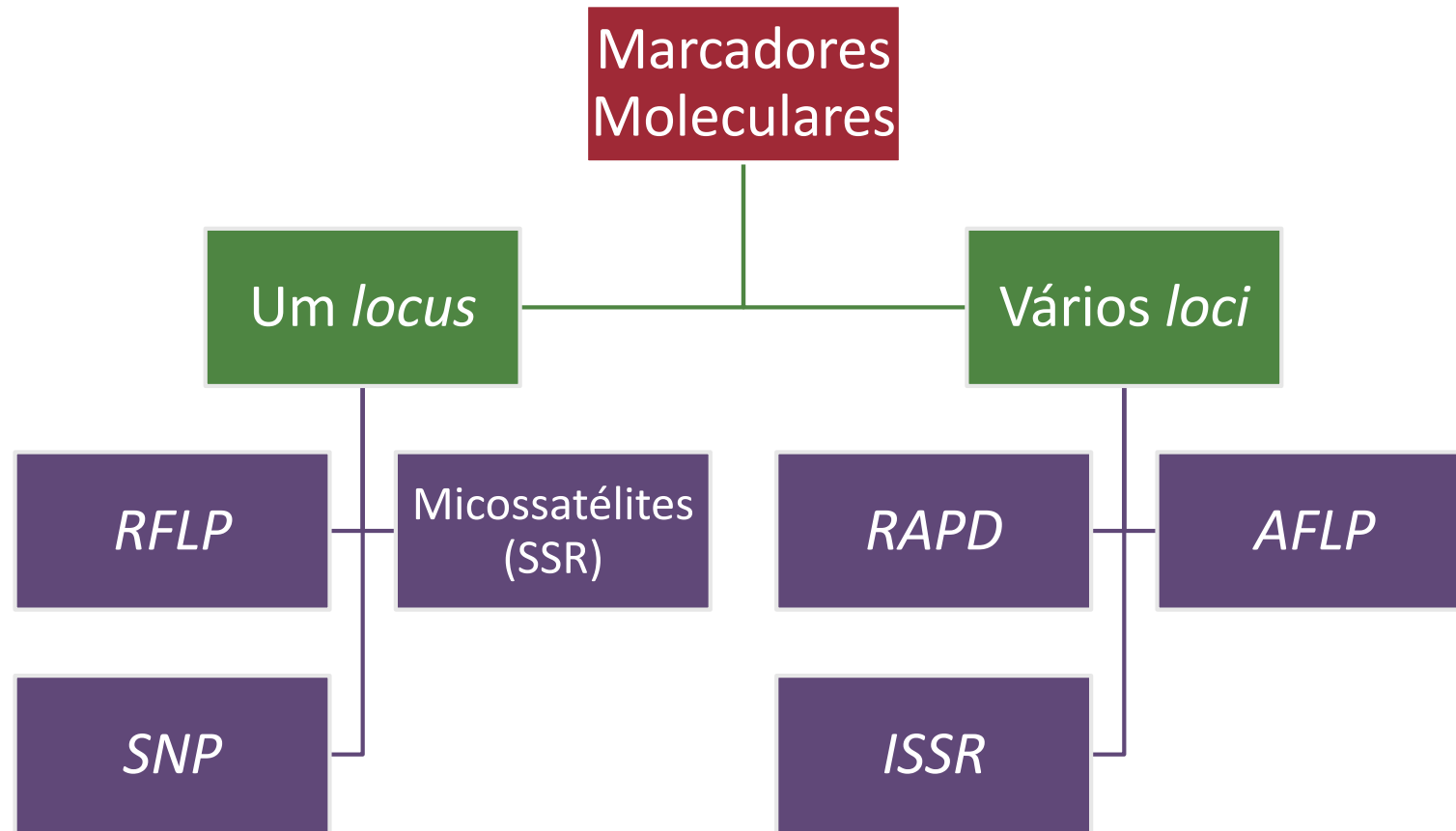
Dominantes vs Codominantes



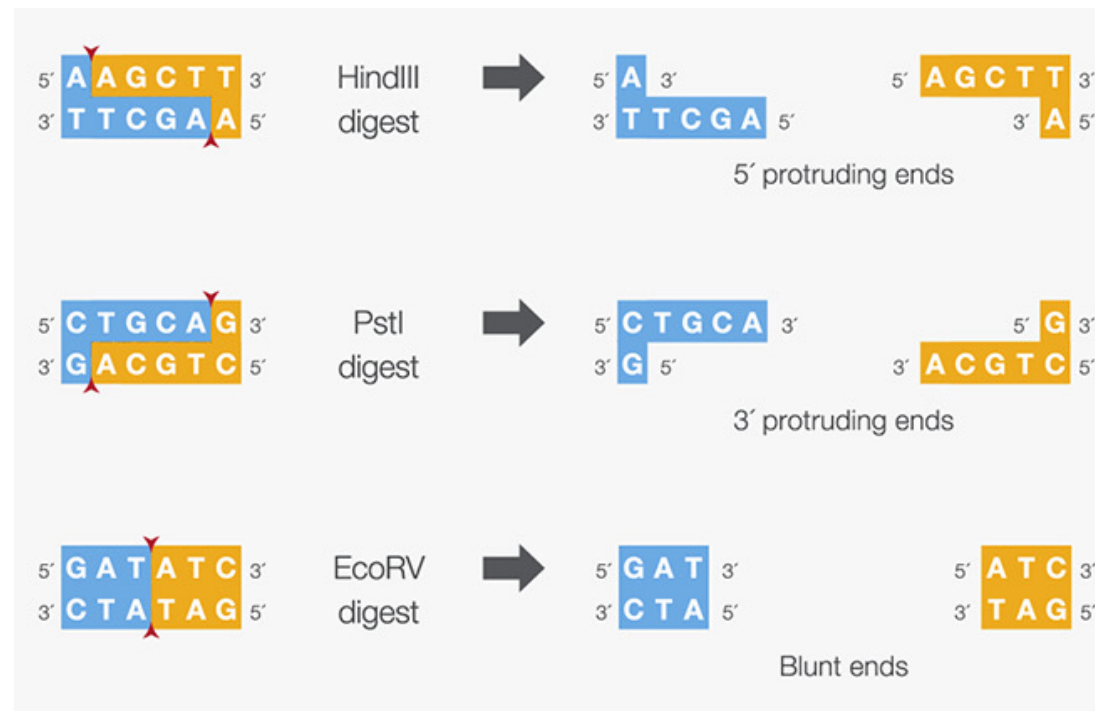
Dominante



Co-dominante



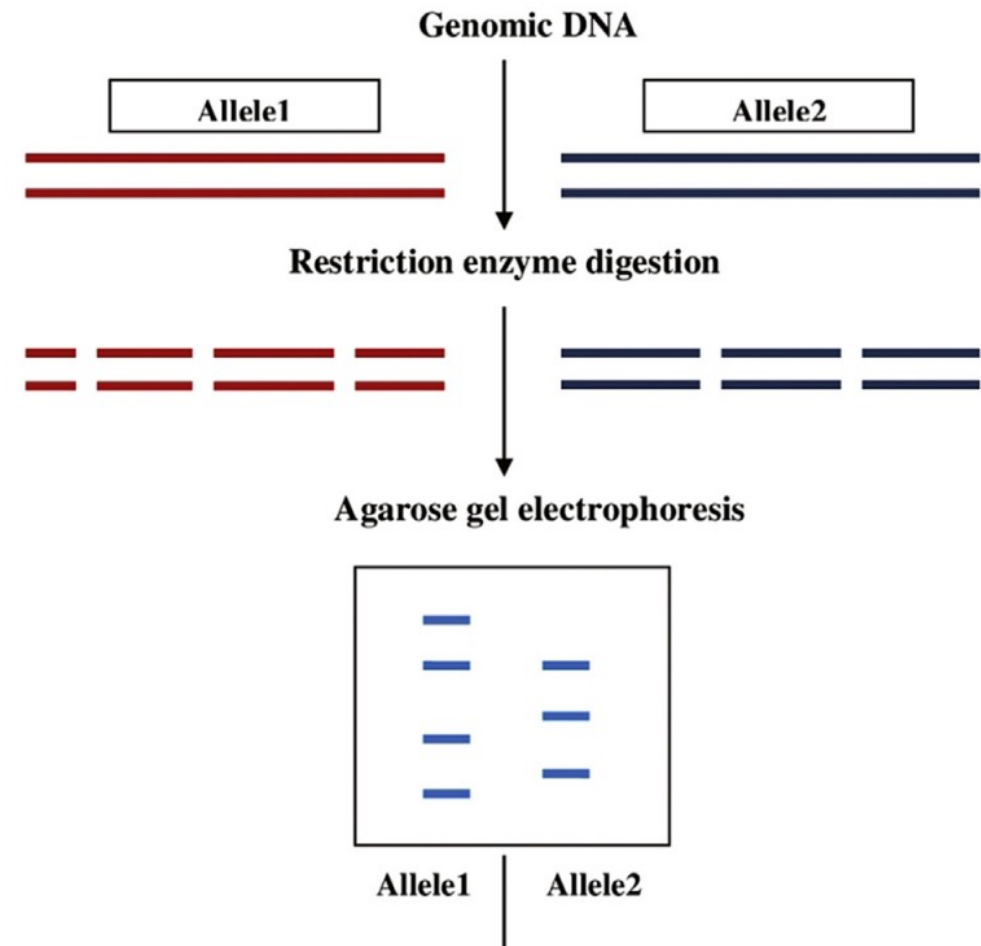
Enzimas de restrição - endonucleases



- **Endonucleases de Restrição** clivam o DNA em sequências específicas, produzindo **fragmentos de DNA em cadeia dupla** de **tamanhos estritamente definidos** pela sua especificidade

RFLP - Restriction Fragment Length Polymorphism

1. Enzimas de **restrição**
2. **Electroforese**
3. **Hibridação** com sonda
4. Detecção



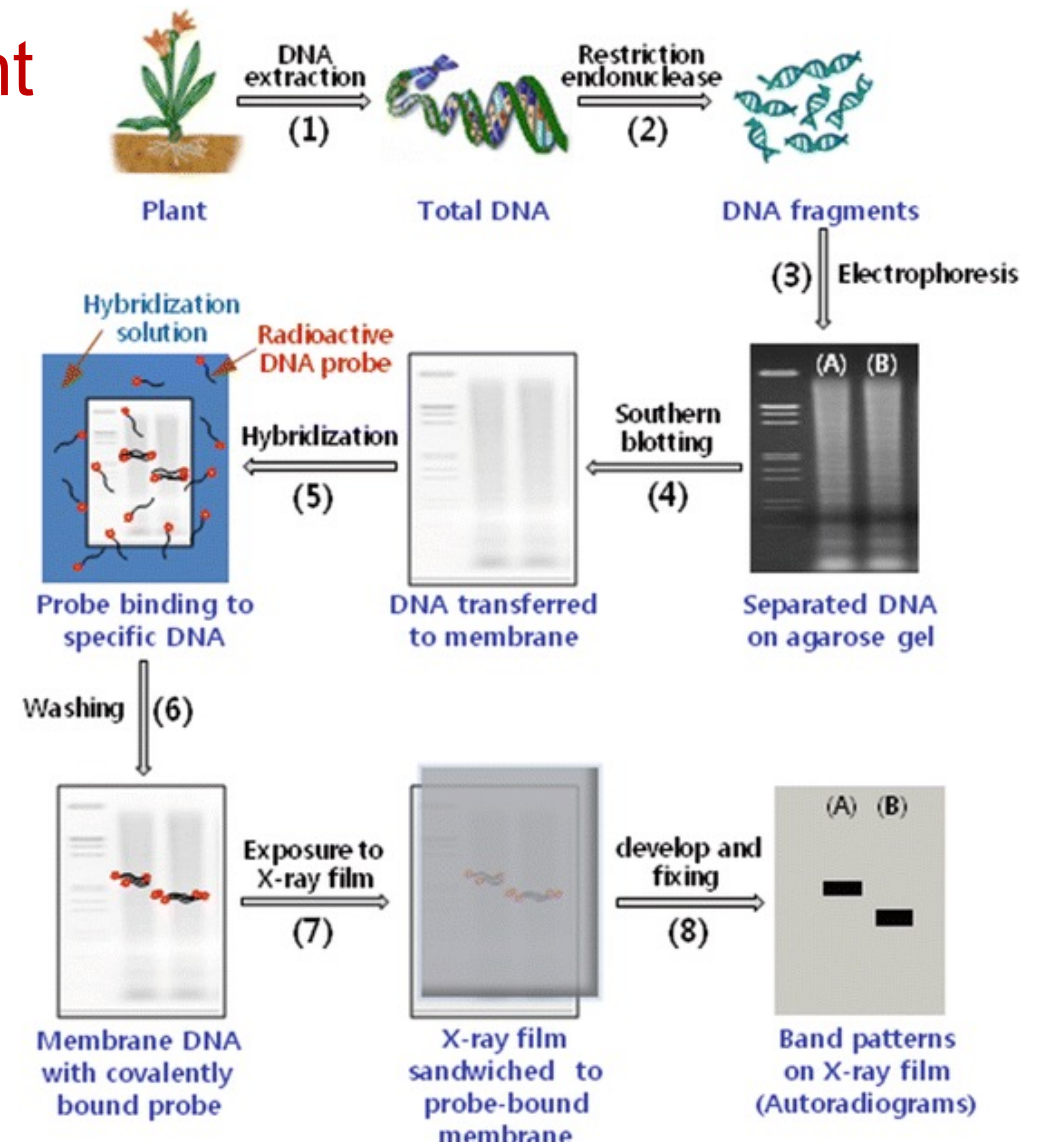
Um *locus*

RFLP - Restriction Fragment Length Polymorphism

RFLP procedure

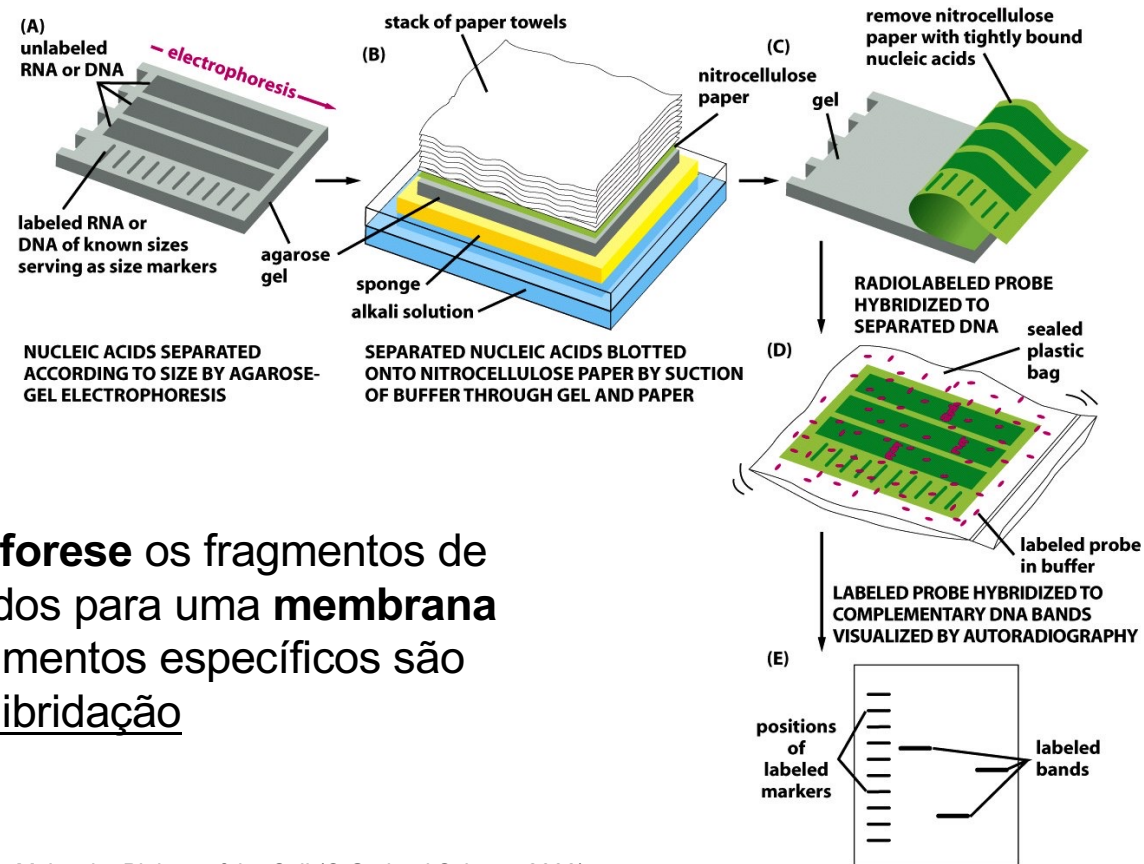
- (1) Extraction of **DNA** from individuals A and B
- (2) DNA cut by **restriction enzymes**
- (3) DNA **electrophoreses**
- (4) DNA Transfer to a membrane by **Southern blot**
- (5) **Hybridization** of labeled DNA fragments (probes)
- (6) Washing to remove nonspecifically bound probes
- (7,8) Detection to observe **DNA polymorphisms**

Um locus



RFLP - Restriction Fragment Length Polymorphism

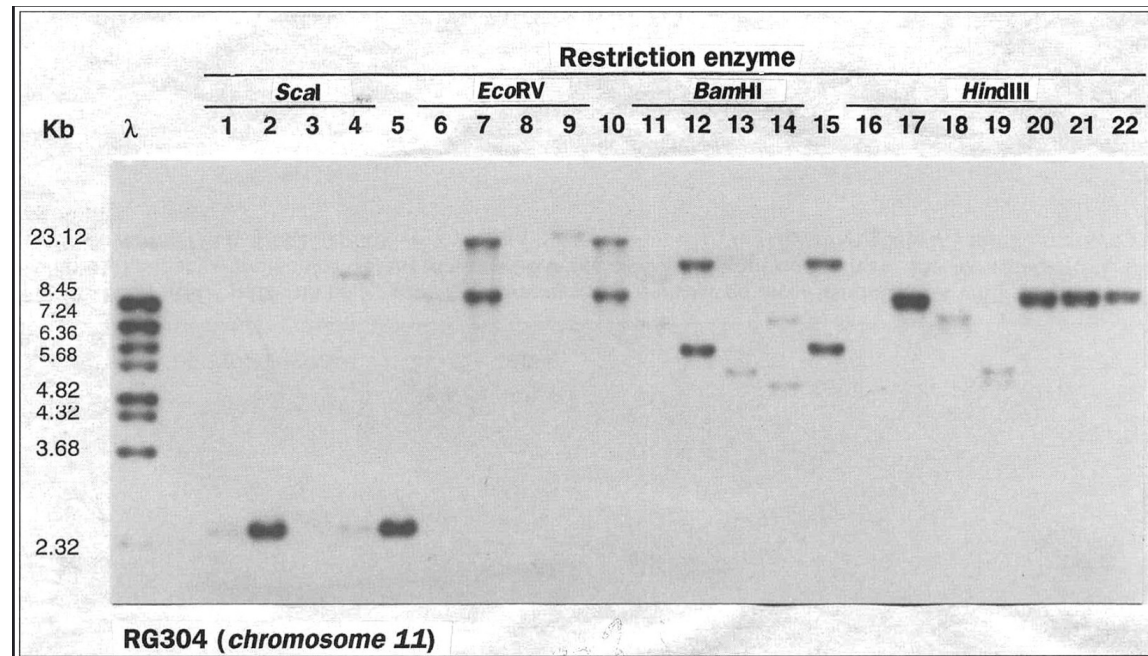
Blotting: Detecção de moléculas específicas de DNA



Depois da **electroforese** os fragmentos de DNA são transferidos para uma **membrana** (*blotting*) onde fragmentos específicos são identificados por hibridação

Figure 8-38 Molecular Biology of the Cell (© Garland Science 2008)

RFLP - Genetic diversity of Rice and wild relatives



RFLP DNA restriction with four restriction enzymes (Sca I, Eco RV, Bam HI, and Hin d III)

- hybridization with RG304 (located on chromosome 11) used in studies of **spikelet length**

→ patterns of *O. sativa* and four wild species (*O. rufipogon*, *O. brachyantha*, *O. granulata* and *O. australiensis*)

RFLP - Genetic diversity of Chicken

RFLP multilocus analysis

- Identification of **intrapopulation differences**
- calculate **genetic distances**
 - **phylogenetic tree**
 - origin and degree of relationship
 - in **18** divergently selected **breeds**
- four main branches and several subclusters

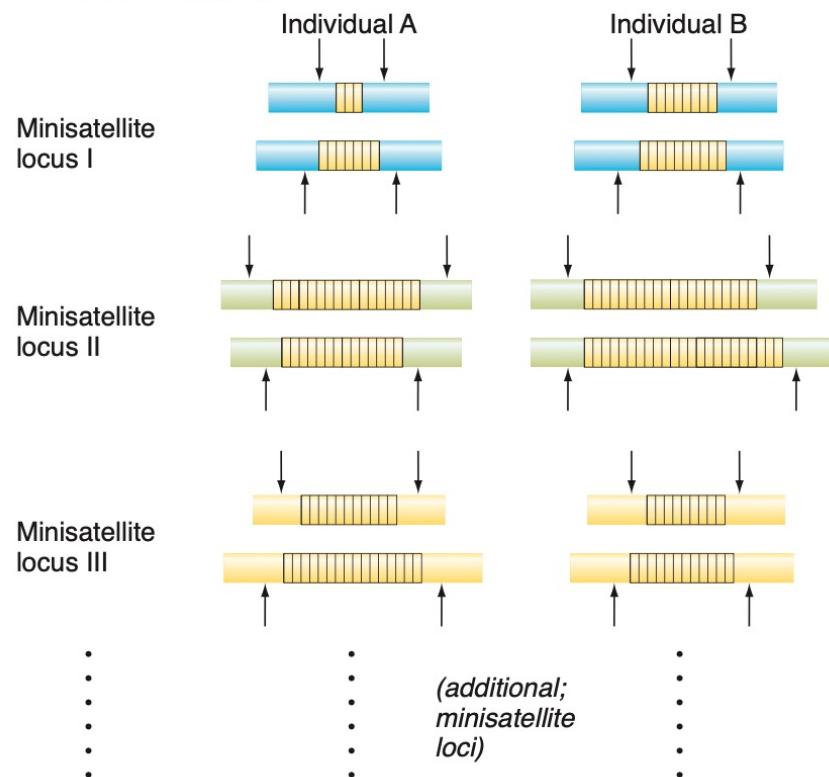
Dementieva et al 2022. Comparative Analysis of Molecular RFLP and SNP Markers in Assessing and Understanding the Genetic Diversity of Various Chicken Breeds



RFLP Minisatellites (500 bp -20 Kb)

Restriction sites

(a) Digest DNA with restriction enzyme that does not cut inside minisatellite.



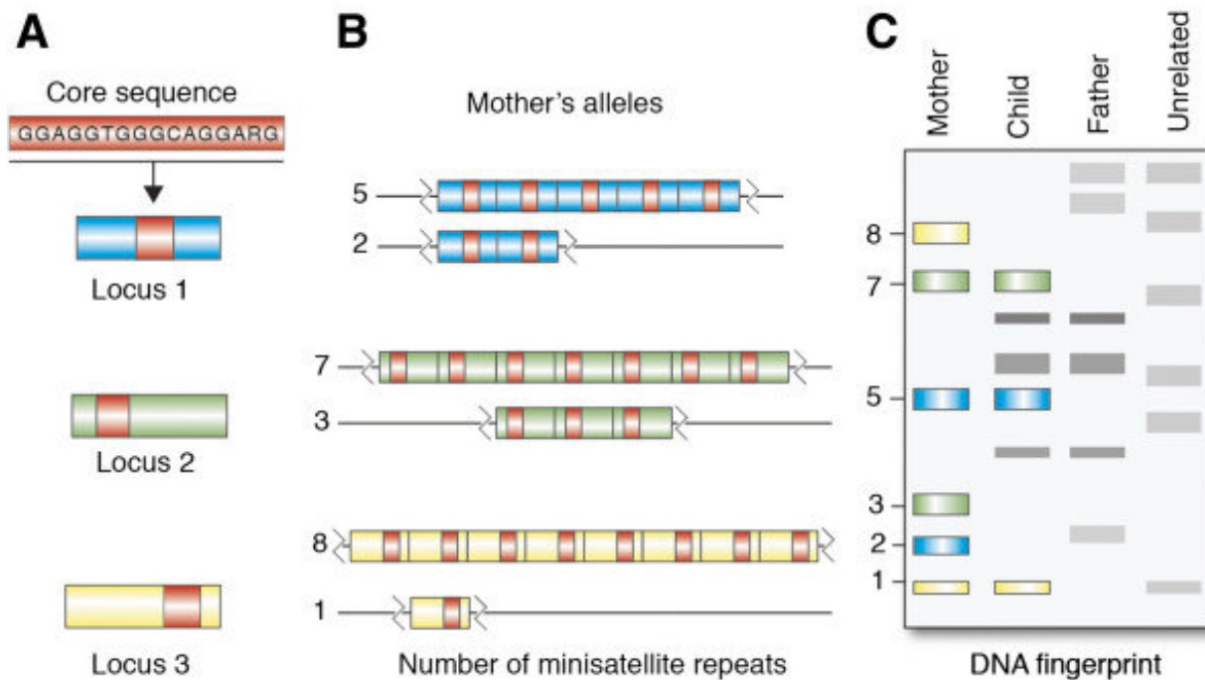
- Digestão com enzimas de restrição
- Eletroforese
- Southern blotting (RFLP)
- Hibridização com a sonda para o minissatélite

DNA fingerprinting

Individuals
A B C D



RFLP Minisatellites fingerprinting



Minisatellite repeat (16 bp core sequence) present in humans and other animals

(A) Minisatellite repeat present in 3 loci

(B) Mother is heterozygous at each of the three loci

Locus 1 genotype: 5, 2

Locus 2 genotype: 7, 3

Locus 3 genotype: 8, 1

(C) Restriction fragment profiles

of four individuals at these three loci

Child - at each locus, one allele shared with the mother and the other with the father

Unrelated individual

shares only a small number of bands

PCR - Amplificação de DNA

1985 - Kary Mullis e colaboradores

Reacção de PCR

(Polymerase Chain Reaction)

Permite **amplificar exponencialmente**
fragmentos específicos de DNA



Kary Mullis

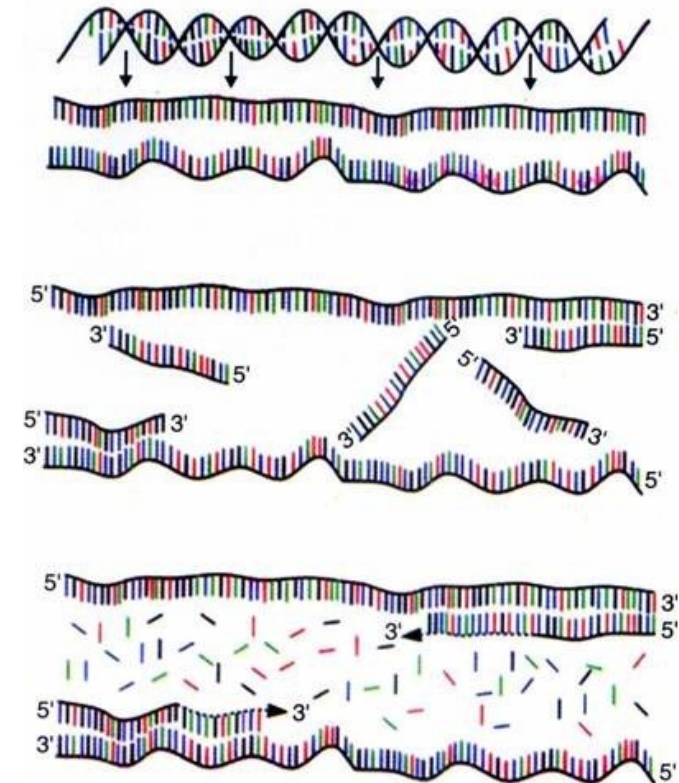
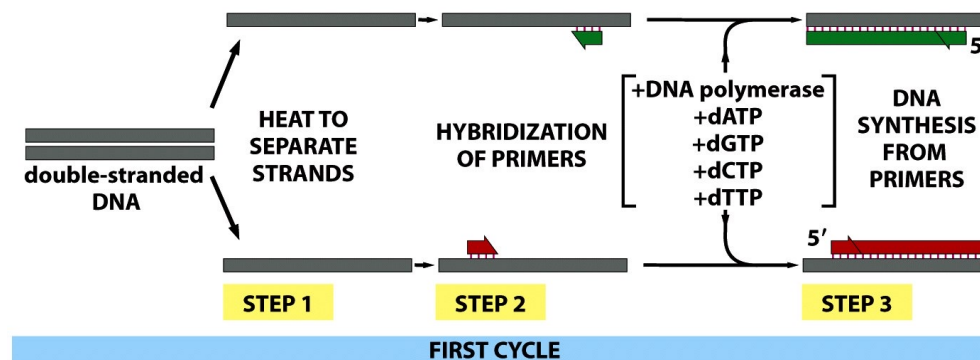
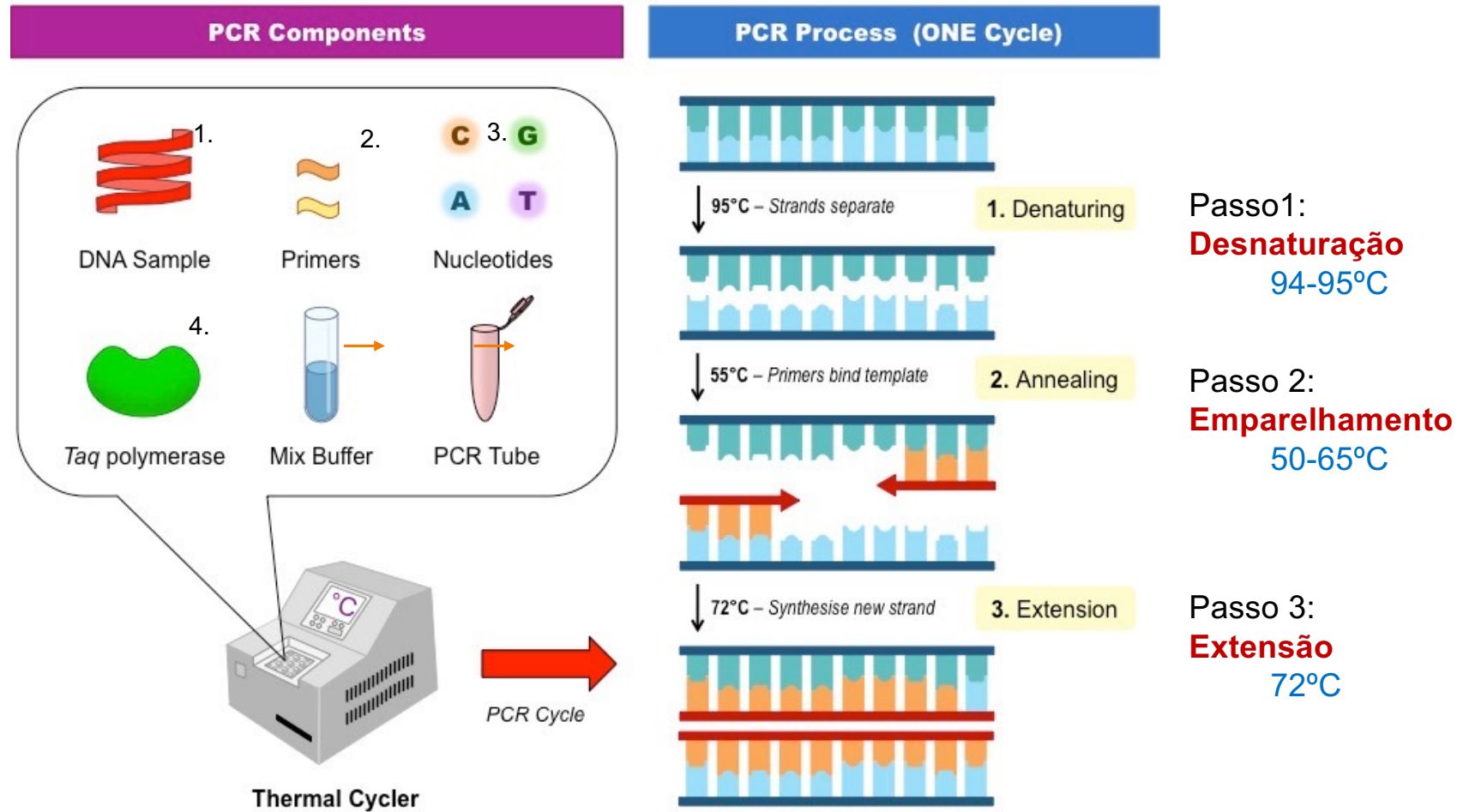
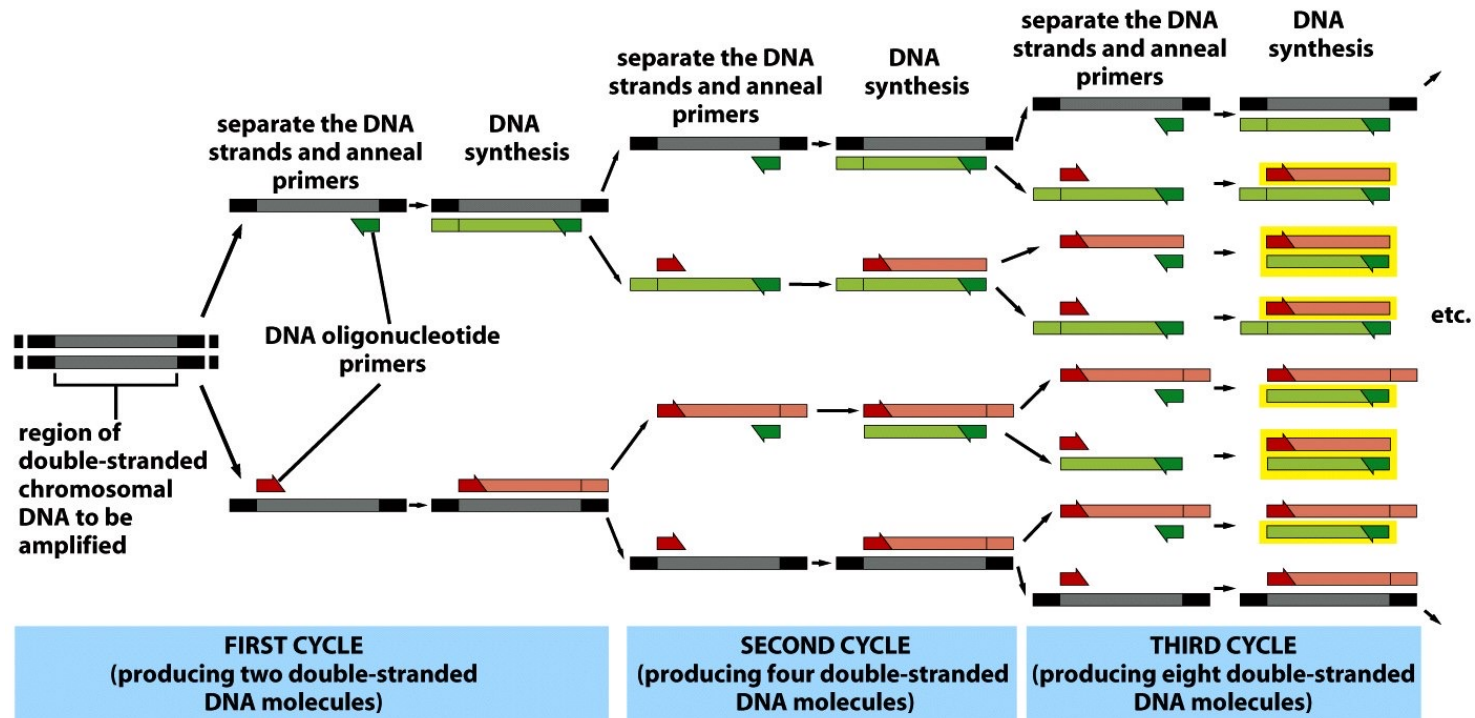


Figure 8-45a Molecular Biology of the Cell (© Garland Science 2008)



Taq polymerase: isolated from *Thermus aquaticus* (thermophilic bacteria) - optimal temperature: 72°C

PCR - Amplificação de DNA



Cada ciclo duplica a quantidade de DNA sintetizada no ciclo anterior

No ciclo seguinte as cadeias previamente sintetizadas servem de molde

→ ao fim de poucos ciclos a molécula predominante é um **fragmento de DNA** com uma **dimensão** igual à distância entre os dois primers originais

PCR - Polymerase Chain Reaction

PCR Demo

The Polymerase Chain Reaction (PCR)

© W. H. Freeman & Company

RAPD - Randomly Amplified Polymorphic DNA

Amplificação por **PCR**

Sem necessidade de conhecer o genoma

Marcador **dominante**

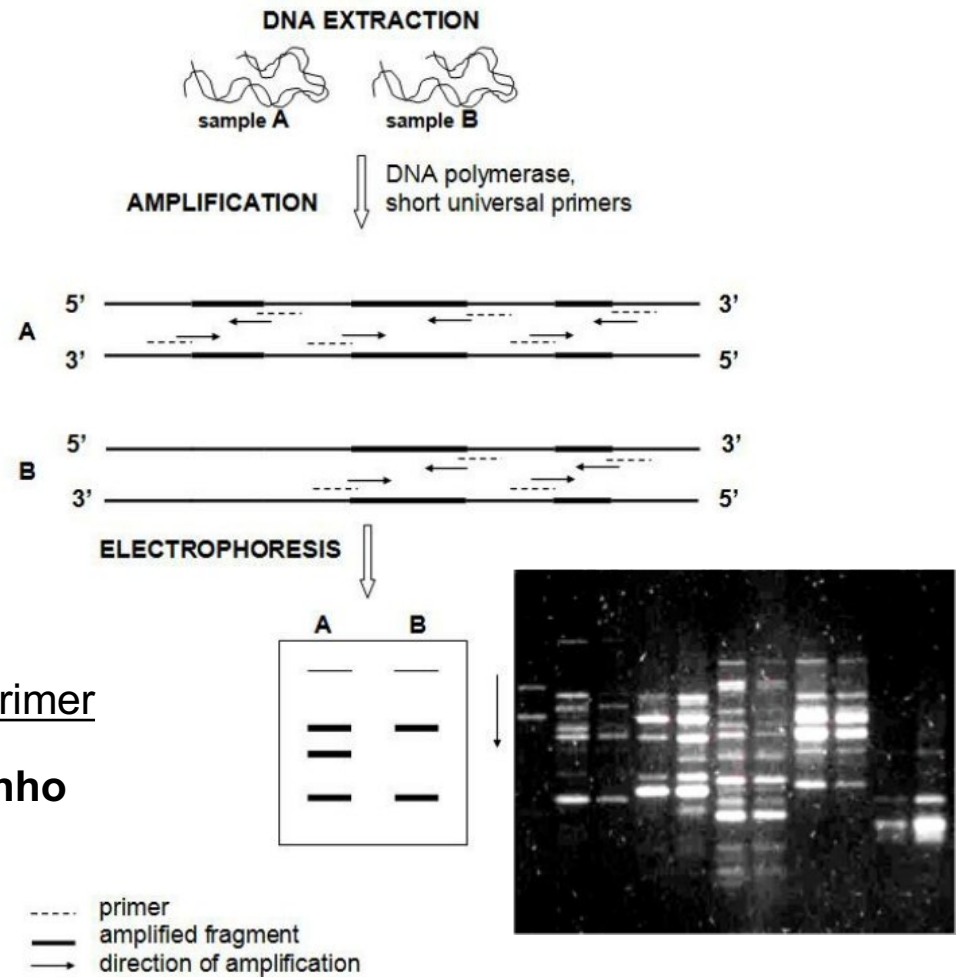
Apenas presença ou ausência da banda
(não identifica heterozigotos)

Primers aleatórios curtos (**10-12 bp**) ligam-se a
regiões codificadoras ou não do genoma de
distribuição aleatória

→ DNA de tamanhos diferentes

Polimorfismo - mutação no local de *annealing* do primer

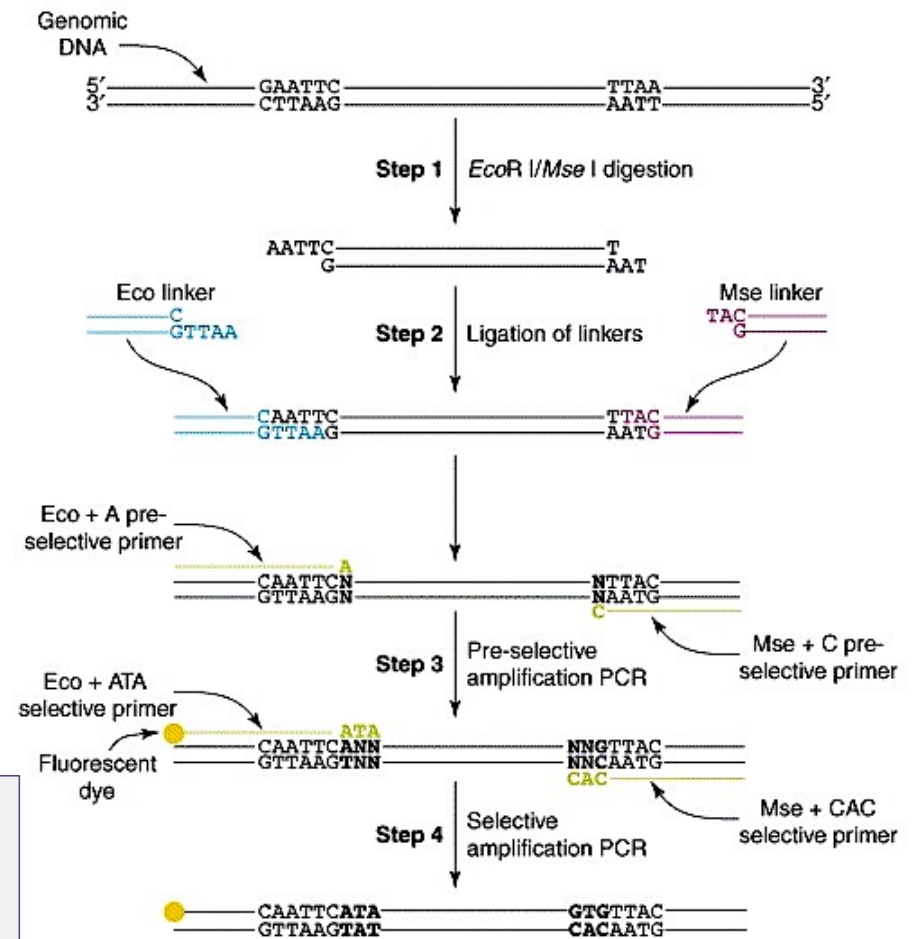
Difícil reprodutibilidade devido ao **reduzido tamanho**
dos primers que se ligam de modo instável



AFLP - Amplified Fragment Length Polymorphism

1. [Enzimas de restrição](#)
2. Ligação com [adaptadores](#)
3. [Amplificação](#) por **PCR** selectiva dos fragmentos
4. Análise dos fragmentos amplificados em gel de poliacrilamida ([Electroforese](#))
5. Visualização

AFLP são **estáveis** e têm **boa reprodutibilidade**
Dominante → Apenas ausência e presença da banda
(não identifica heterozigotos)



AFLP - Phylogeny of bovine species

AFLP

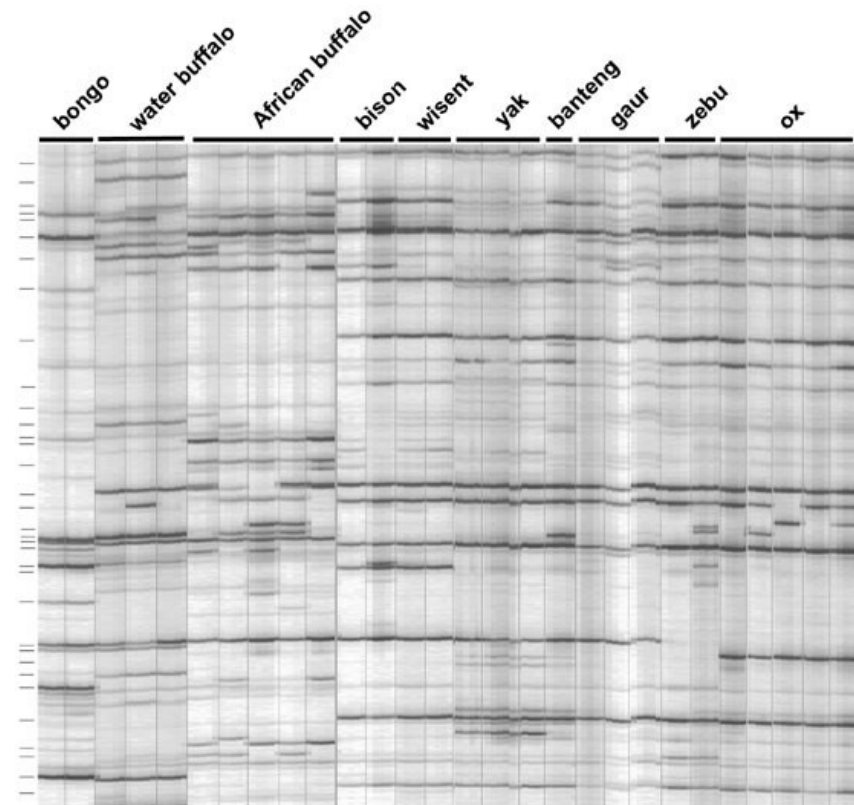
Individuals from different **Bovini** species

PCR with EcoRI and TaqI primers

Computer-assisted scoring of electrophoretic fingerprinting patterns → **361** markers

→ genomic distances between species

→ genetic exchange between ancestral populations



AFLP – rice germplasm diversity

AFLP assessment of rice germplasm

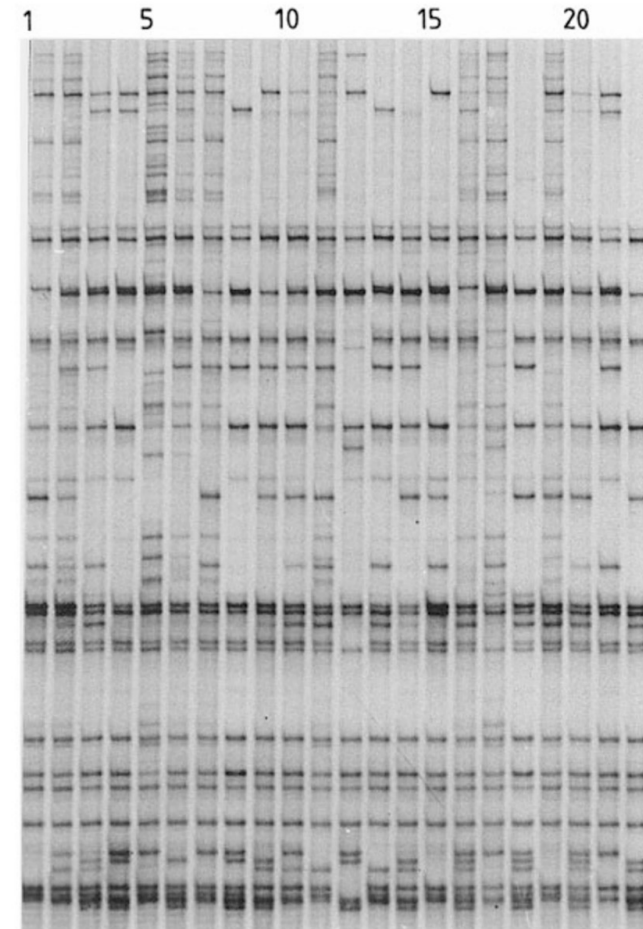
151 markers

1/3 association with indica–japonica differentiation

PCR with 4 primer combinations

→ Identification of 5 rice genotypes

<https://www.nature.com/articles/6884410.pdf>



SSR - Simple Sequence Repeat or Microsatellite

Amplificação por **PCR** com primers específicos de 18-25 bp

Polimorfismos baseados no nº de unidades repetidas

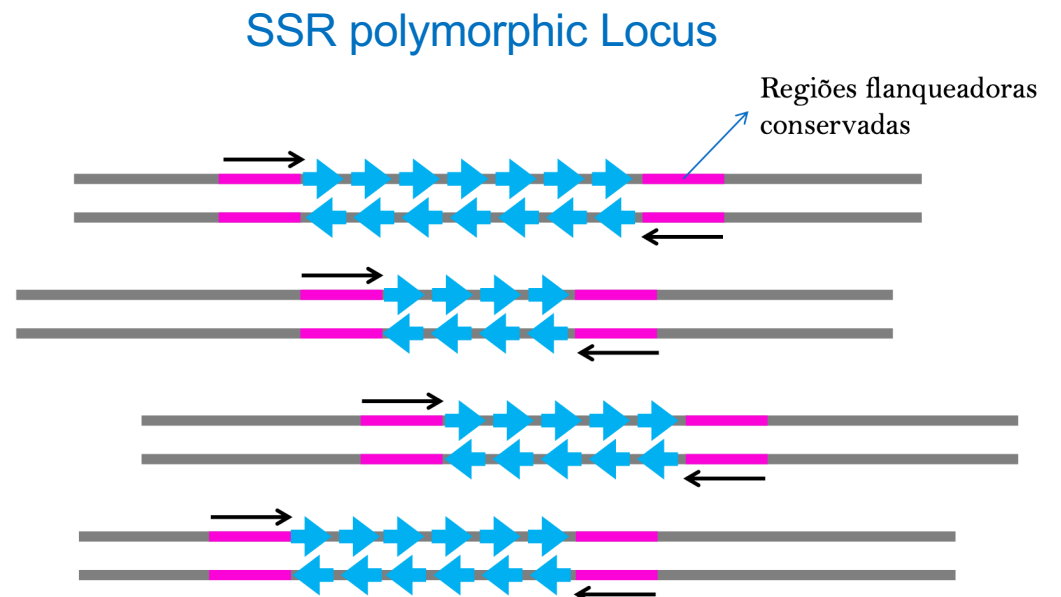
Co-dominante

Hipervariáveis

Têm amplificação estável e boa repetibilidade

Fáceis de executar e automatizar

Ex: genoma de mamíferos - SSR (CA)_n ocorre em média uma vez a cada 30 000 bp

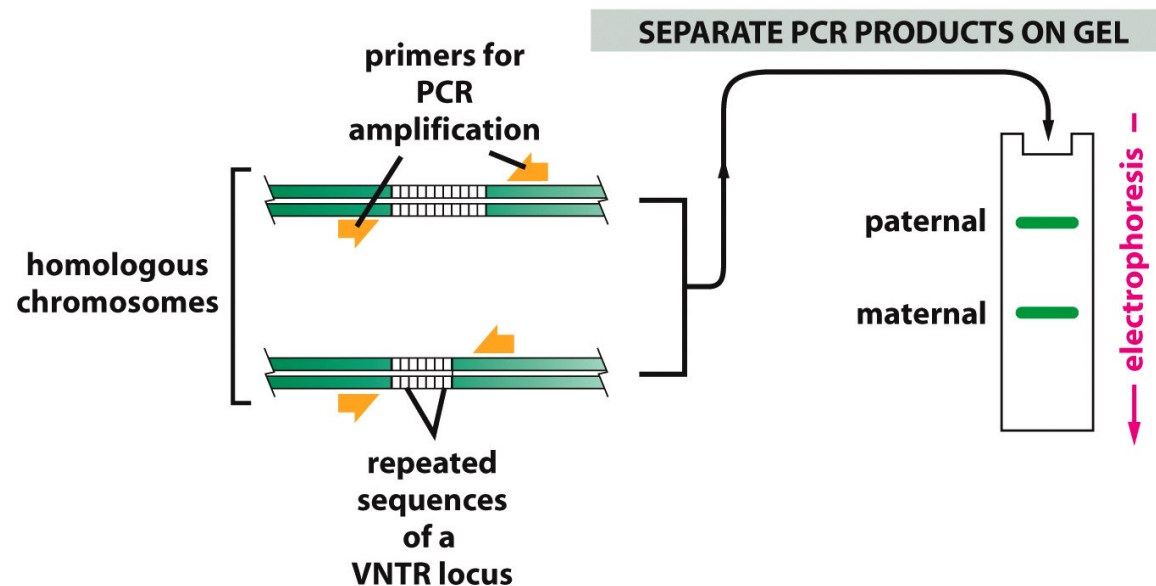


SSR - Simple Sequence Repeat or Microsatellite

Microsatélites – sequências repetidas de nucleótidos

VNTR - *variable number of tandem repeat*

...TTGTATGTATGTATGTATGTATGTGATG...



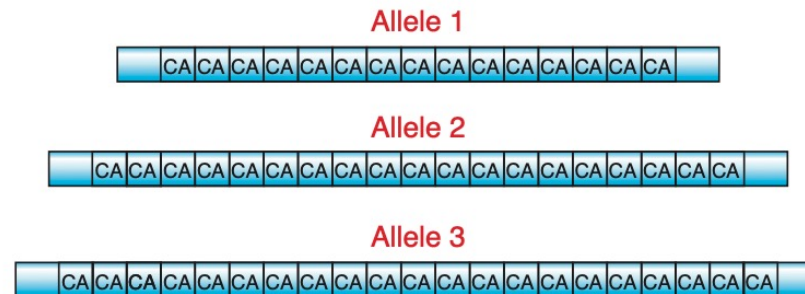
- **Polimorfismos** - pequenas variações/mutações no nº de nucleótidos
- **Amplificação do DNA por PCR** com *primers* para as regiões adjacentes
- Análise do nº de nucleótidos por **electroforese/sequenciação**

SSR – variabilidade numa população

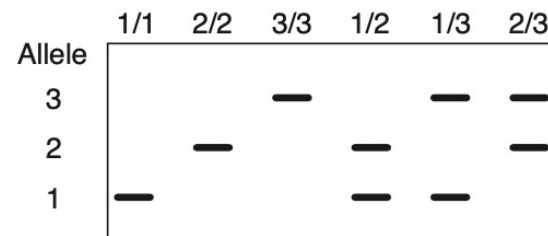
Variabilidade natural de um SSR numa população com três alelos diferentes num locus com repetição de um dinucleótido

→ **Co-dominante**

Example of population with three alleles



Six diploid genotypes are present in this population.

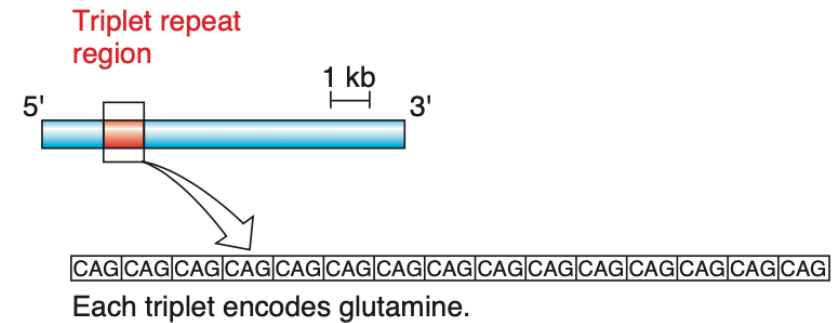


SSR associados a doenças

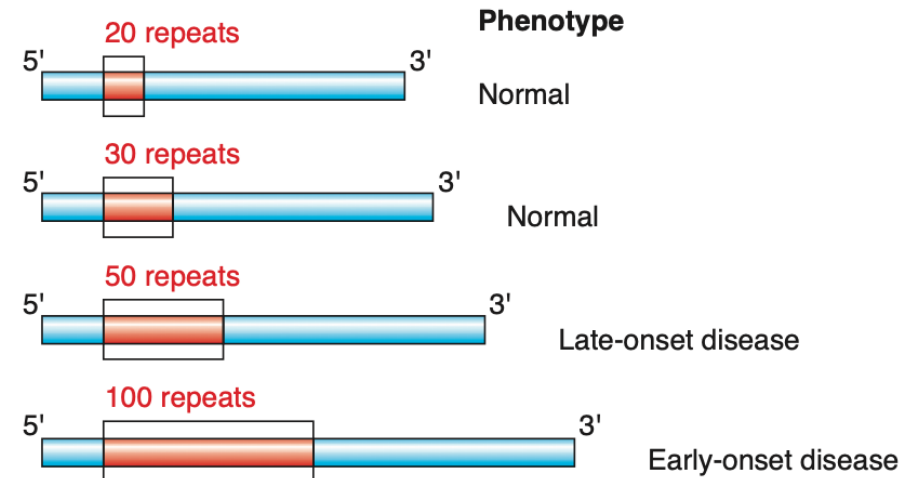
As mutações causadoras da
doença de Huntington

Causada por expansão de um tripleto
repetido numa região codificadora

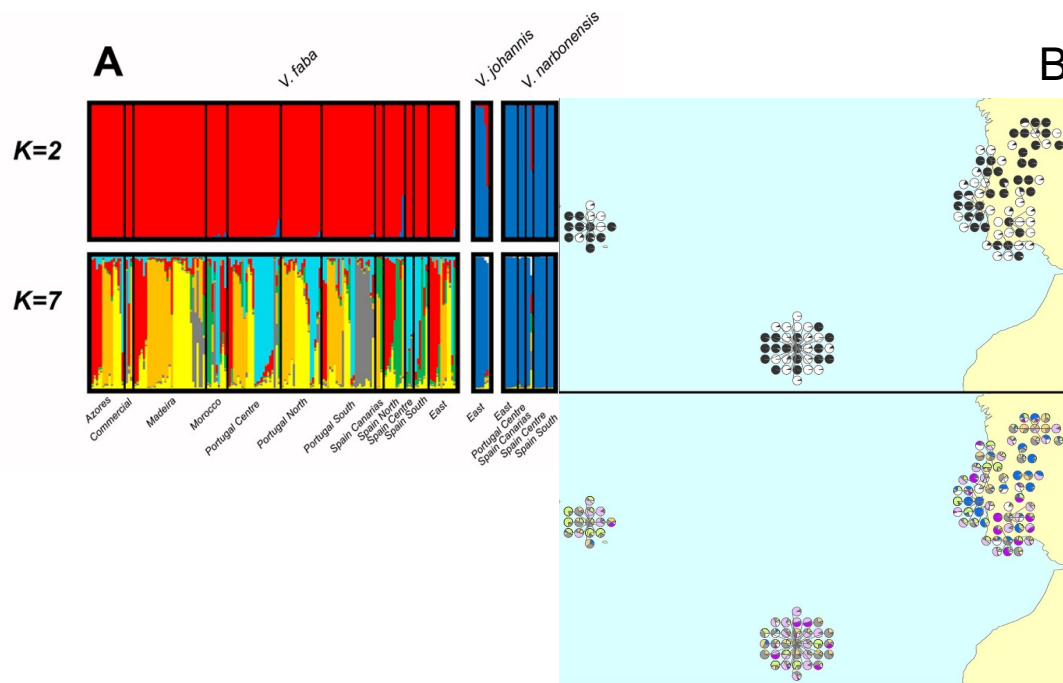
(a) Basic structure of the *HD* gene's coding region



(b) Some alleles at the *HD* locus



SSR - Genetic Diversity of *Vicia faba* and wild related



Genetic diversity evaluation and population structure

- Geographical distribution (Western Mediterranean basin)
- Wild related species

53 *Vicia faba*

2 *V. johannis*

7 *V. narbonensis*

Accessions from Portugal, Spain and Morocco

28 faba bean Single Sequence Repeats (**SSRs**)

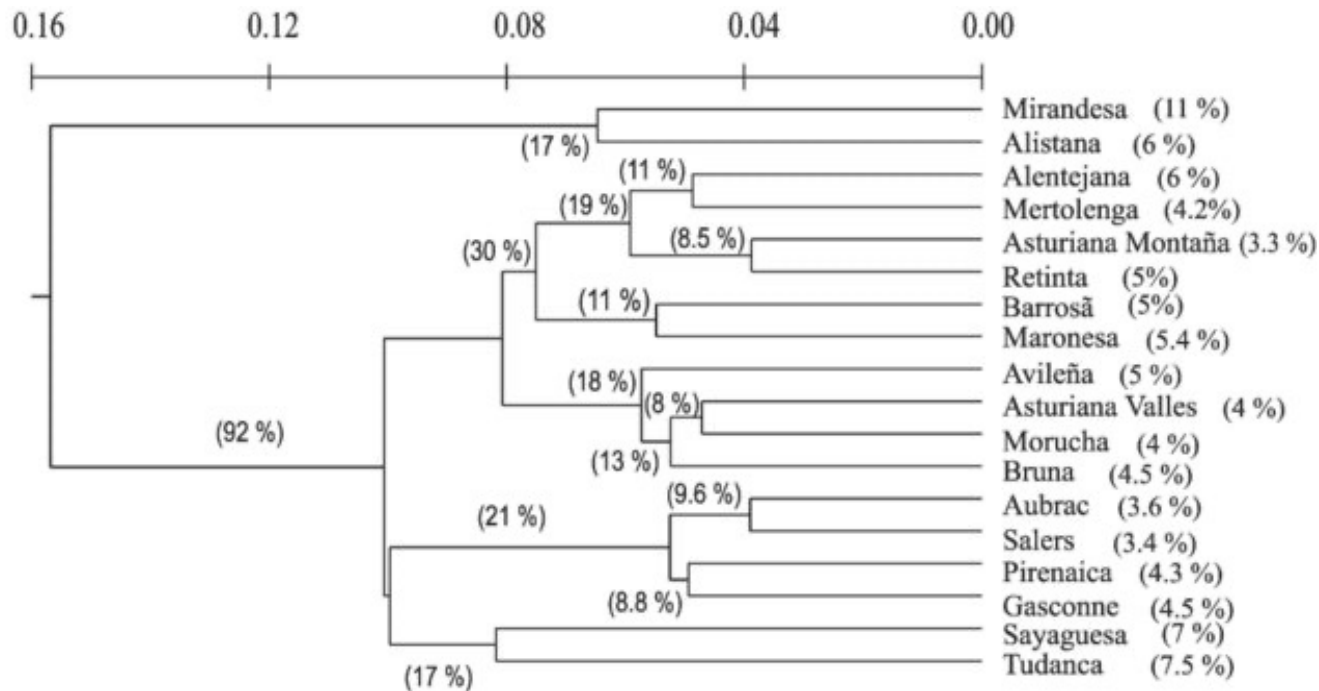
Oliveira et al 2016

<https://journals.plos.org/plosone/article?id=10.1371/journal.pone.015480>

SSR - Genetic diversity of European beef cattle

Tree of 18 local cattle breeds

Values in brackets represent the loss of diversity caused by the extinction of a breed or a set of breeds



Estudo de relações evolutivas e estrutura e diversidade genética de 18 raças locais de bovinos Espanha, Portugal e França usando 16 microssatélites

- Distinção de 4 grupos de raças

- Classificação de raças importantes para a diversidade

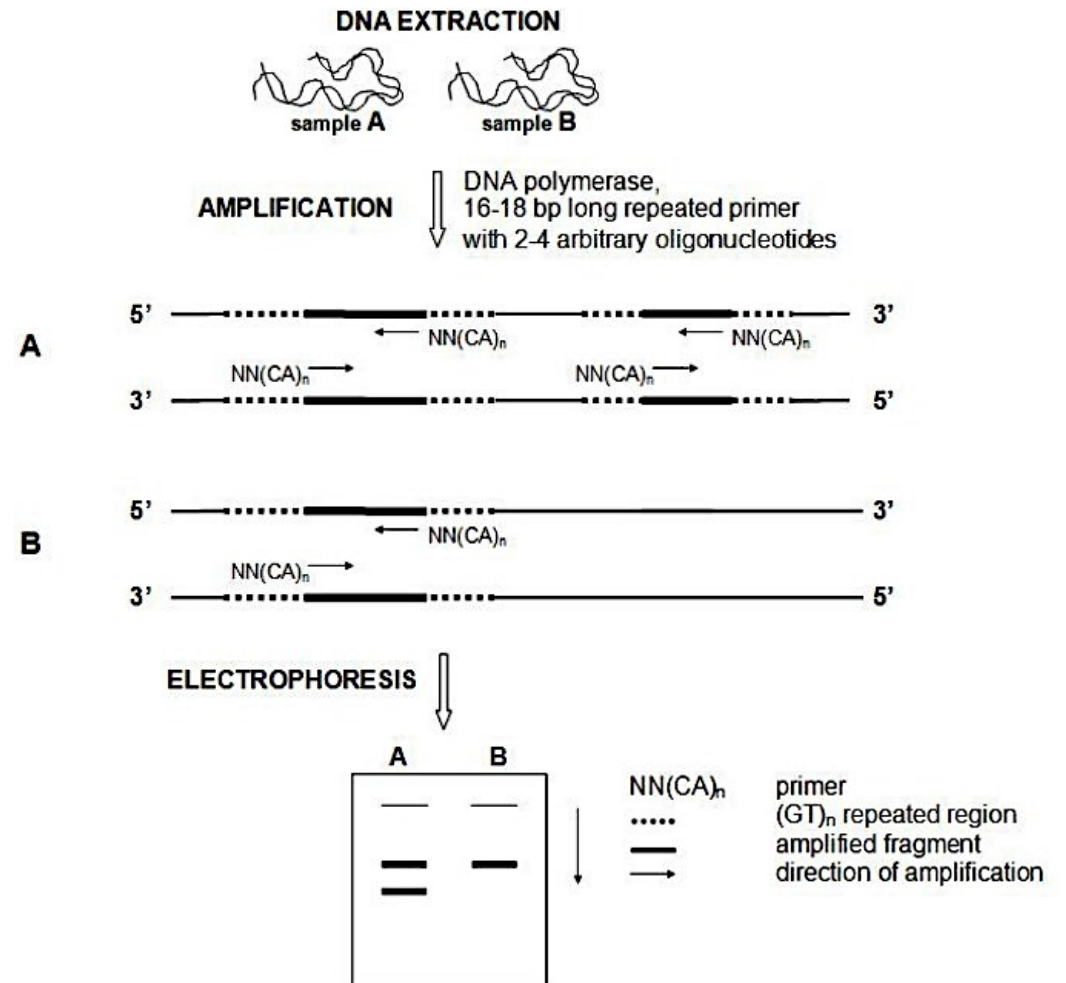
Cañón et al 2011

<https://pubmed.ncbi.nlm.nih.gov/11403750/>

ISSR – Inter Simple Sequence Repeat

Amplificação de regiões do genoma
flanqueadores de microssatélites

1. **PCR**, usando um 1 primer
2. Uso de **eletroforese** em gel de agarose ou poliacrilamida
3. **Avaliação das bandas** ISSR
4. Análise de dados

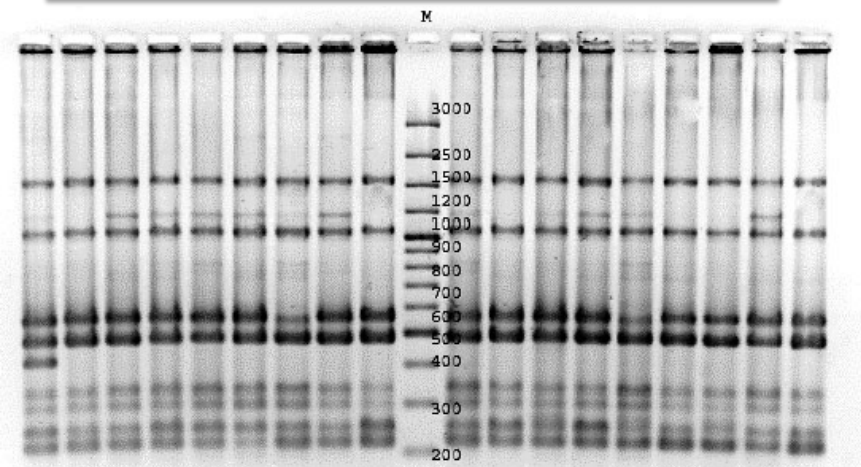


ISSR – Diversidade Genética na raça Mazandaranian

Este estudo realizado para investigar a diversidade genética na raça nativa do Irão (**Mazandarania**) em comparação com a raça **Holstein** usando o marcador Inter Simple Sequence Repeats (**ISSR**)



ISSR-PCR pattern - GA-ISSR marker



175 animais, 71 Mazandaranian e 104 Holstein
PCR com primer (GA)9C

Maior diversidade genética em Mazandaranian
(menor seleção)

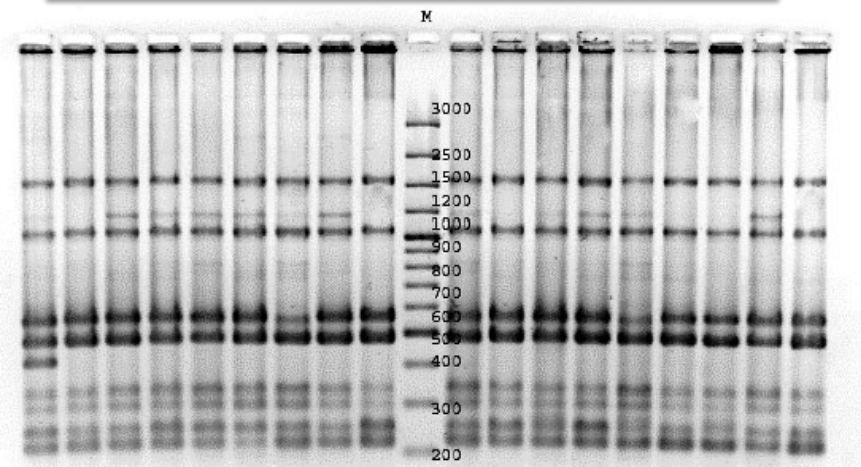
Menor diversidade genética em Holstein
(programas de seleção intensivos)

ISSR – Diversidade Genética na raça Mazandaranian

Este estudo realizado para investigar a diversidade genética na raça nativa do Irão (Mazandarania) em comparação com a raça Holstein usando o marcador Inter Simple Sequence Repeats (ISSR)



ISSR-PCR pattern - GA-ISSR marker



Cálculo de parâmetros básicos

$$\% \text{ of polymorphic bands} = \frac{\text{No. of polymorphic bands}}{\text{Total no. of bands}} \times 100\%$$

Nº de bandas **total** = ??

Nº de bandas **polimórficas** = ??

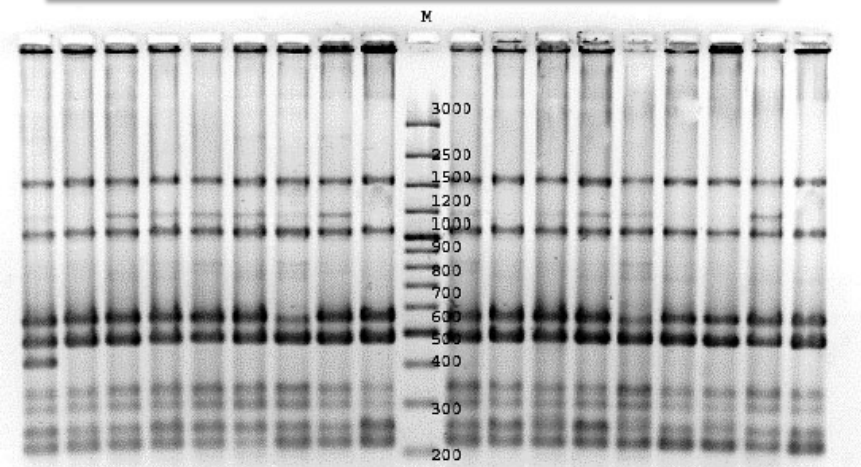
% de bandas **polimórficas** = ??

ISSR – Diversidade Genética na raça Mazandaranian

Este estudo realizado para investigar a diversidade genética na raça nativa do Irão (Mazandarania) em comparação com a raça Holstein usando o marcador Inter Simple Sequence Repeats (**ISSR**)



ISSR-PCR pattern - GA-ISSR marker



Cálculo de parâmetros básicos

$$\% \text{ of polymorphic bands} = \frac{\text{No. of polymorphic bands}}{\text{Total no. of bands}} \times 100\%$$

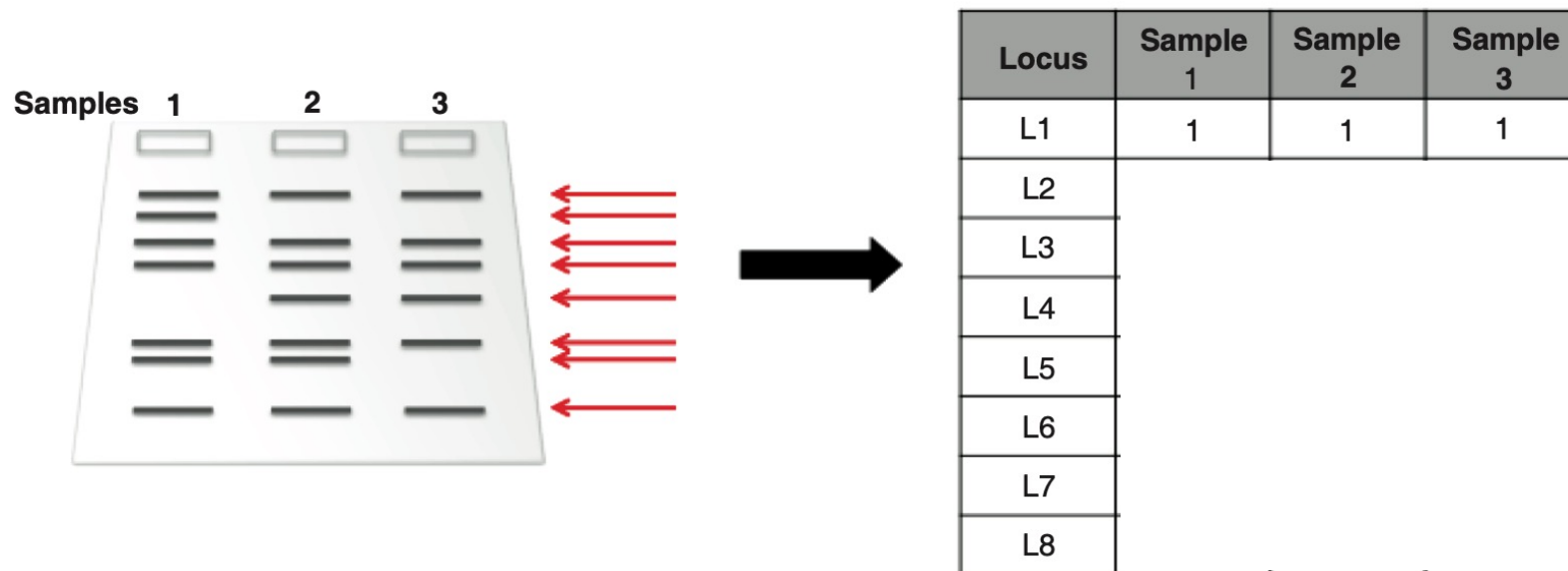
Nº de bandas **total** = 10

Nº de bandas **polimórficas** = 2

% de bandas **polimórficas** = $2/10 = 20\%$

Avaliação da diversidade genómica

análise de perfis de bandas

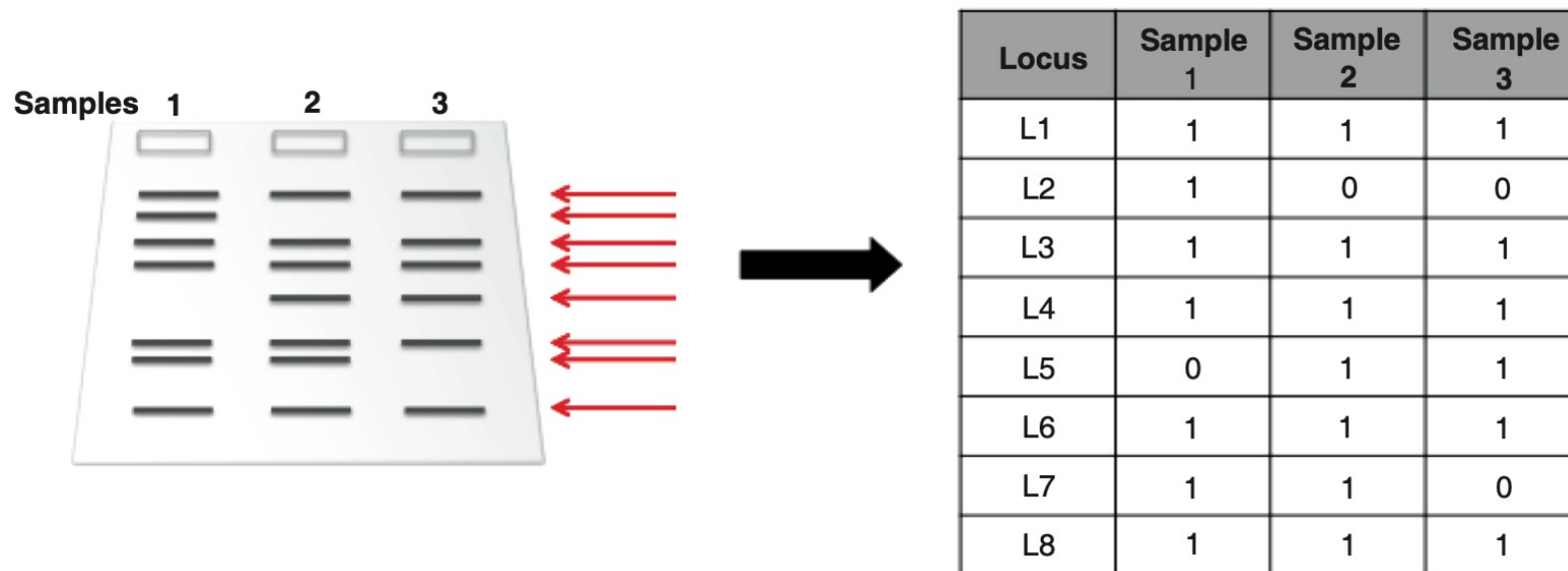


Análise de perfis de bandas - método para comparar a **semelhança** com base na presença e ausência de bandas num gel

- construir **árvores de similaridade** (dendrogramas)
- calcular índices de **diversidade**

Avaliação da diversidade genómica

análise de perfis de bandas



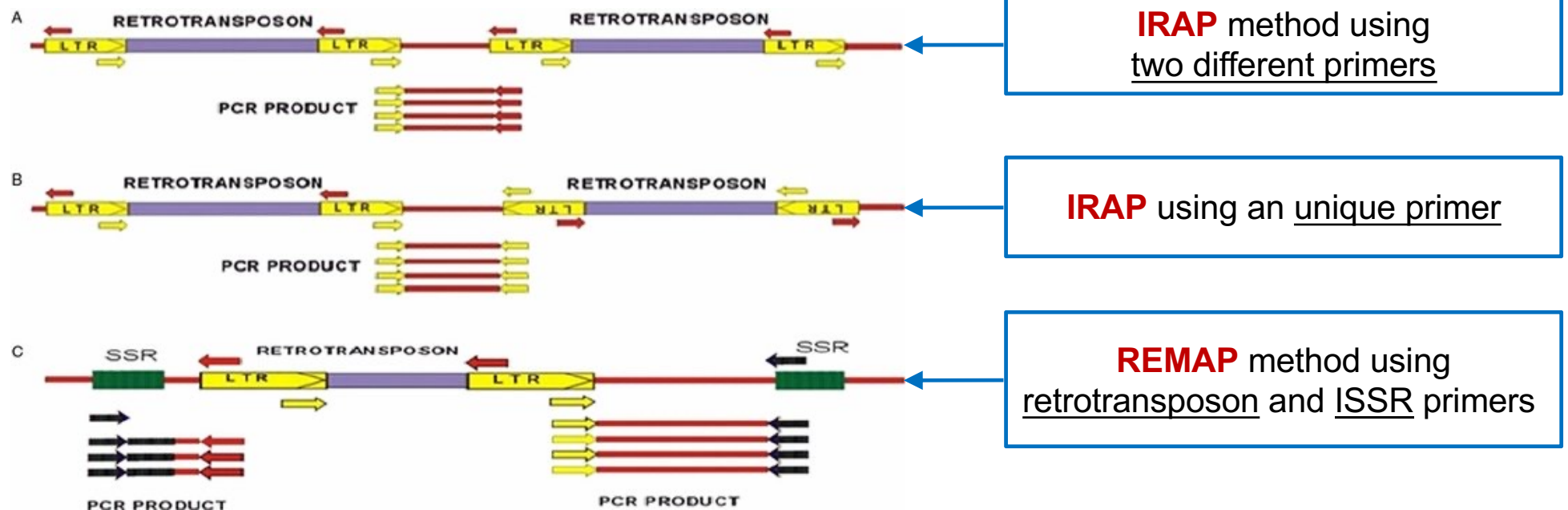
Análise de perfis de bandas - método para comparar a **semelhança** com base na presença e ausência de bandas num gel

- construir **árvores de similaridade** (dendrogramas)
- calcular índices de **diversidade**

Marcadores baseados em Retroelementos

IRAP - inter-retrotransposon amplified polymorphism

REMAP - retrotransposon-microsatellite amplified polymorphism

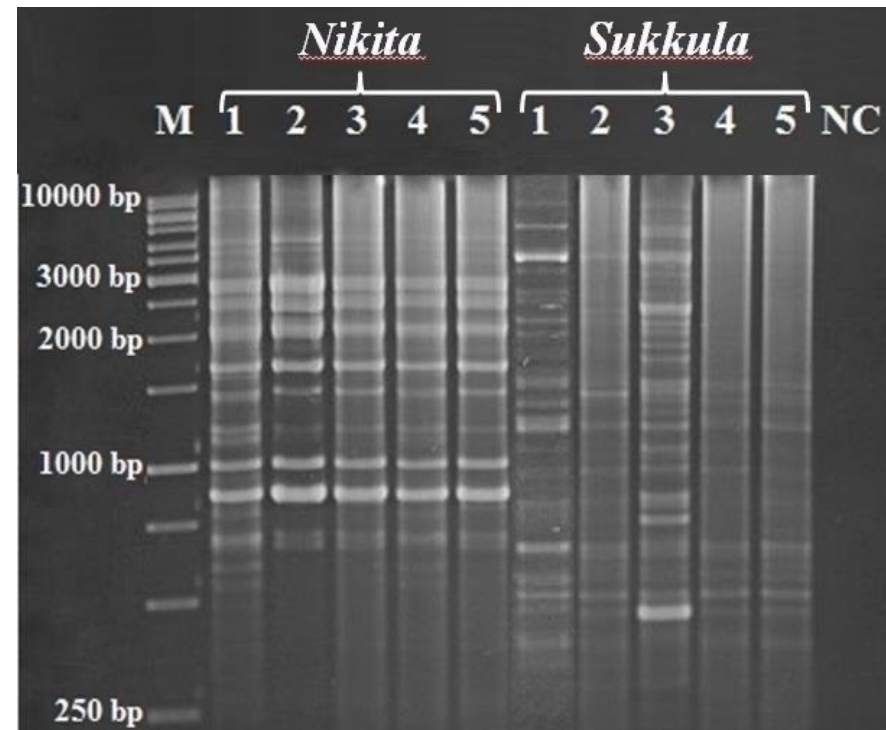


Amplification of genome regions between **Retrotransposons** and/or **SSR**

Sukkula IRAP - *Cucumis sativus* genome

IRAP (Inter-Retrotransposon
Amplified Polymorphism)
amplify genome regions between
barley-specific retrotransposons
Nikita and Sukkula

→ Discrimination of *Cucumis* sp.

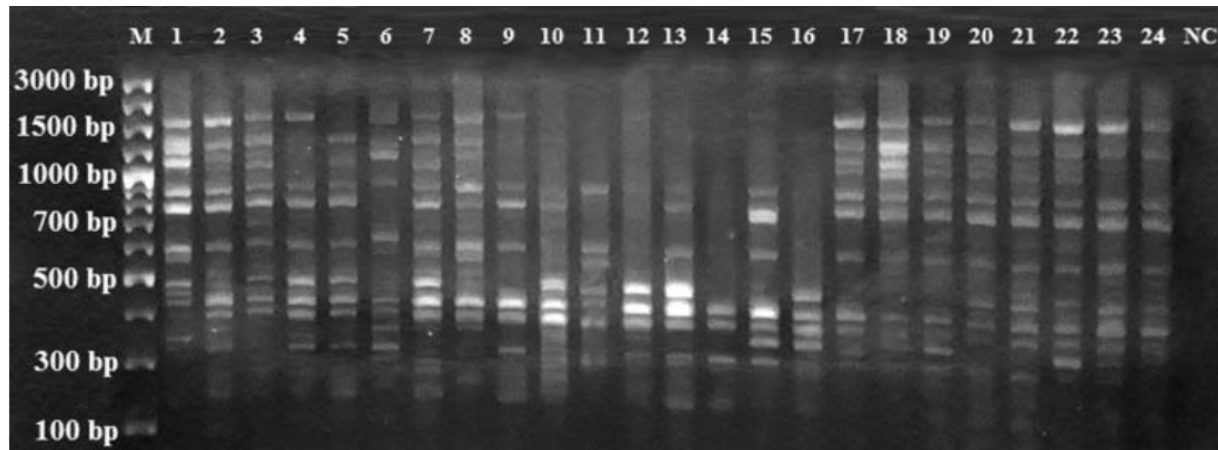


Nikita and *Sukkula* IRAP-PCR results of *Cucumis sativus* M: marker; NC: negative control. N°s indicate five biological replicates

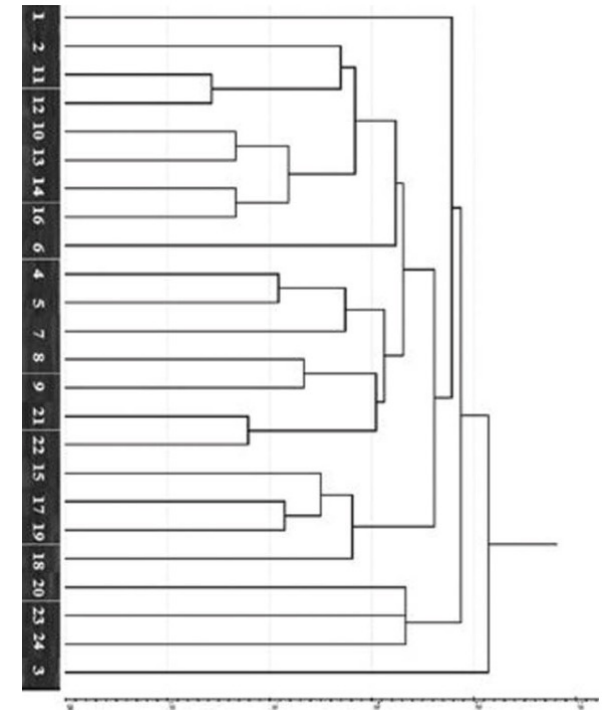
https://www.researchgate.net/publication/341320775_Retrotransposon_Analyses_in_Cucurbitaceae_family

Sukkula IRAP - human genome

IRAP amplification using Sukkula-specific primers
M - marker, NC, negative control (no template DNA)

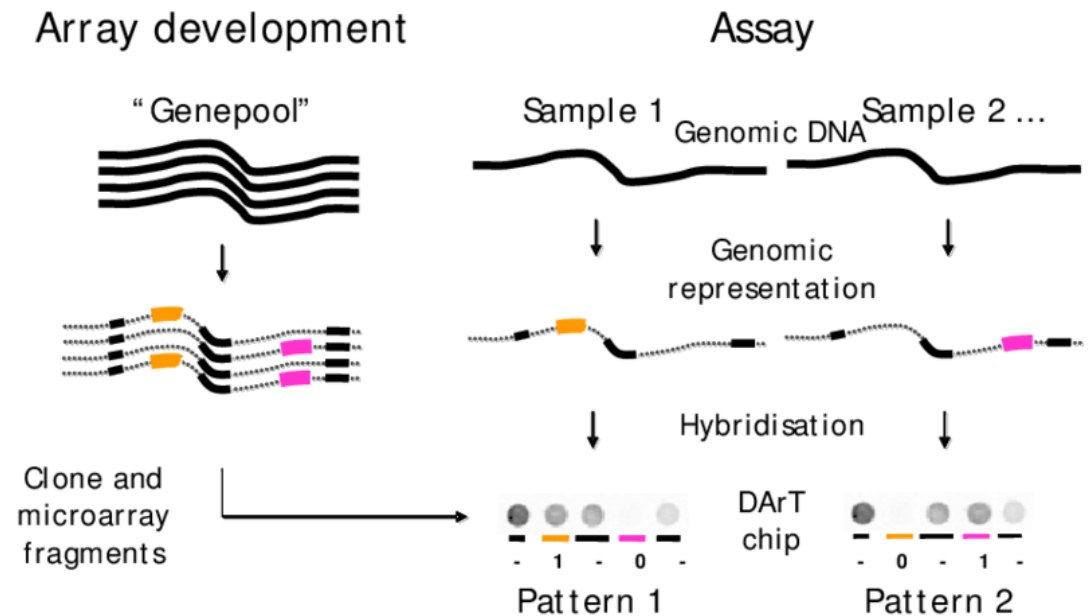


Plant-specific retrotransposon (Sukkula) identification in the human genome (24 DNA samples) by inter-retrotransposon amplified polymorphism (IRAP)
Polymorphism ratios: 8%–100 in females and 13%–100% in males
→ First reports on plant-specific retrotransposons in the human genome



DArT - Diversity Array Technology

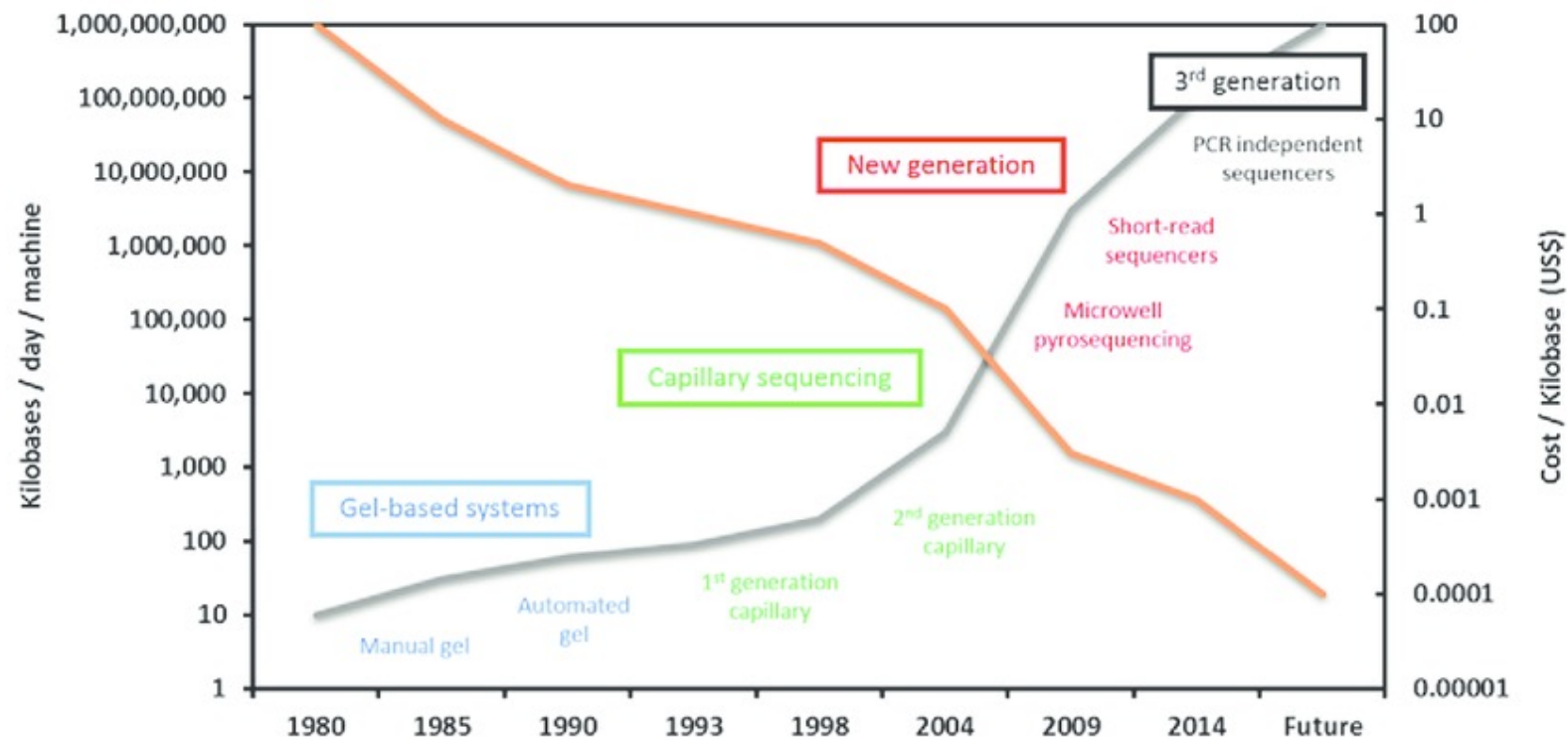
- Não usa electroforese
- Usa **hibridação** de fragmentos de DNA numa superfície sólida
→ **array** ou **microChip**
- Sistema de alto débito, ou seja, analisa uma enorme quantidade de fragmentos simultaneamente
- Detecta a presença ou ausência de fragmentos individuais



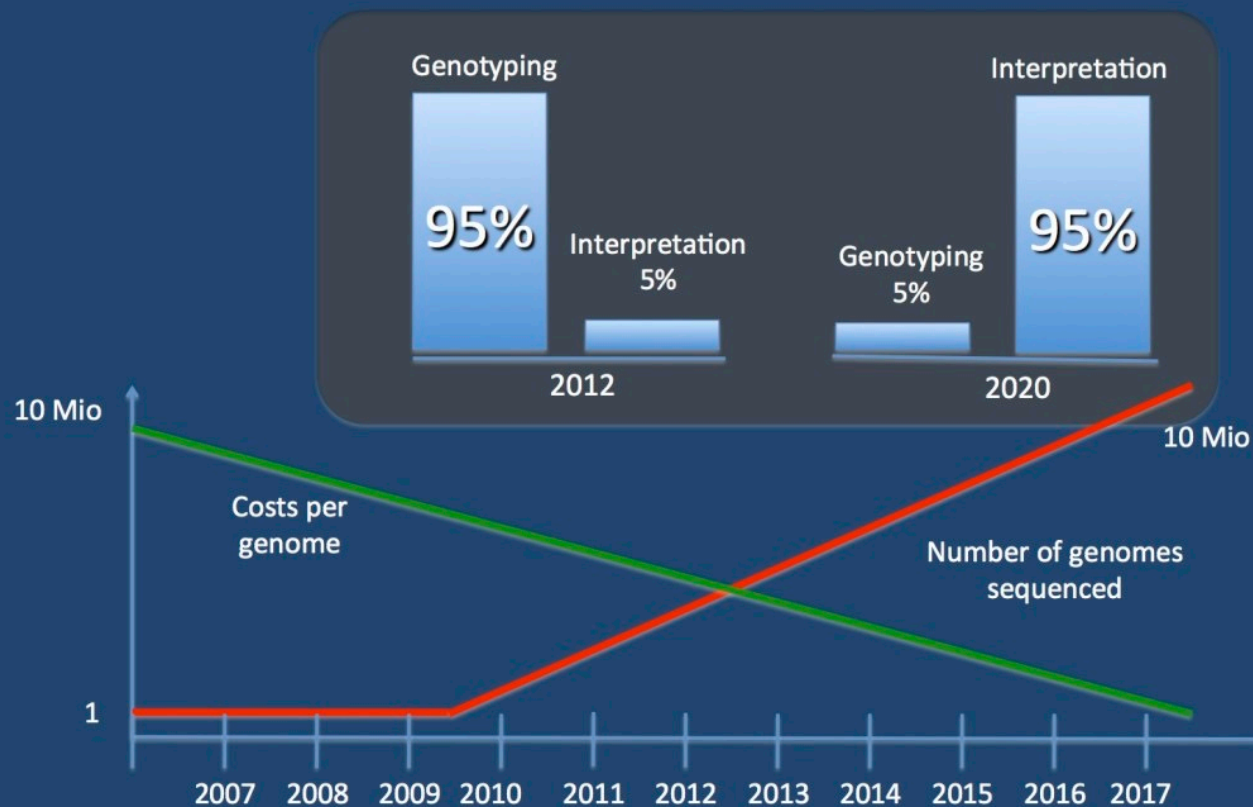
Resumo das características dos marcadores

Features	RFLP	AFLP	RAPD	SSR	ISSR	SNP	DArT	Retrotransposons
Detection method	Hybridization	PCR	PCR	PCR	PCR	PCR / Seq	Seq/Hibridação	PCR
Type of probe/primer used	g DNA/ cDNA sequence of structural genes	Sequence specific primers	Random design primer	Sequence specific primers	SSR unique primer	Sequence specific primers	gDNA	Sequence specific primers
Dependent of genomic knowledge	Limited	Limited	Limited	Extensive	Limited	Extensive	Extensive	Limited
Degree of polymorphisms	Low	High	Medium to High	High	High	Very High	High	High
Phenotype expression	Co dominant	Dominant	Dominant	Multiallelic Co- Dominante	Multiallelic Dominante	Bi-allelic Co- dominant	Dominante	Multialélico Dominante
Possibility of automation	No	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Sequencing yield and cost



Genome vs. genome analysis



Genome vs Genome analysis:

- Costs for genome **sequencing** are **decreasing** significantly
- In parallel, costs for genome **interpretation** are **increasing**

DNA barcoding

Técnica molecular utilizada para **identificação** de espécies/variedades

Análise de uma **sequência de DNA específica** padronizada
(“bioidentificador”)

Gene ou **região específica** de DNA para cada grande grupo taxonômico
- identificador universal

Após amplificação por **PCR** e **sequenciação**

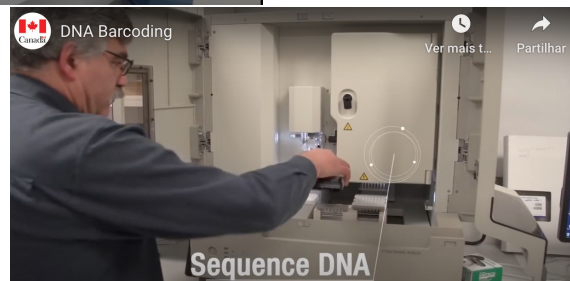
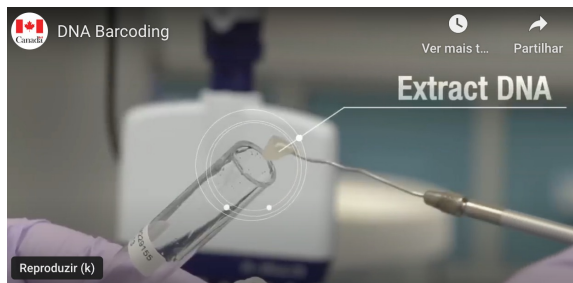
Comparação das sequências (barcodes) com bases de dados de referência
→ → **identificação** (ID)

Marcadores DNA barcode

Grupo	DNA Barcode	Localização genômica
Animais	<i>COI</i>	Mitocôndria
Plantas	<i>MatK + rbcL</i>	Cloroplasto
Fungos	ITS-rDNA	Núcleo
Bactérias	16S-rDNA	Mitocôndria



DNA barcoding



<https://www.youtube.com/watch?v=9YGM6L8bzxg&t=58s>
3'35"

Spinach barcoding

Analyzing local spinach plant diversity
with rpoC1 and matK DNA barcode technique

Standard DNA barcode markers

Maturase K (**matK**)

RNA polymerase subunit beta (**rpoC1**)

Set of Primers used for PCR of the extracted DNA

Rpoc1 F: GGCAAAGAGGGAAGATTTCG

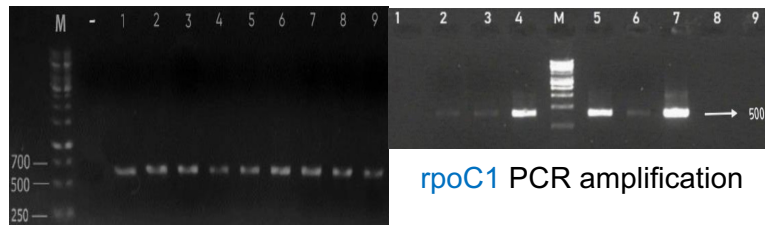
Rpoc1 R: GAGAAAACATAAGTAAACGGGC

MatK F: CATCCATCTGGAAATCTTGGTTC

MatK R: TATGATAATGAGAAAGATTCTGC



PCR amplification from spinach plants

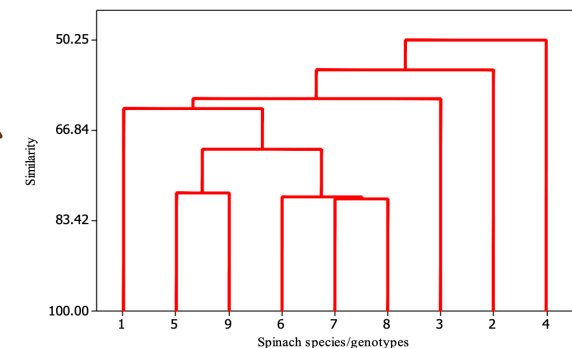


matK PCR amplification

rpoC1 PCR amplification

Sequencing

Spinach diversity



<https://www.cabidigitallibrary.org/doi/pdf/10.5555/20230398404>

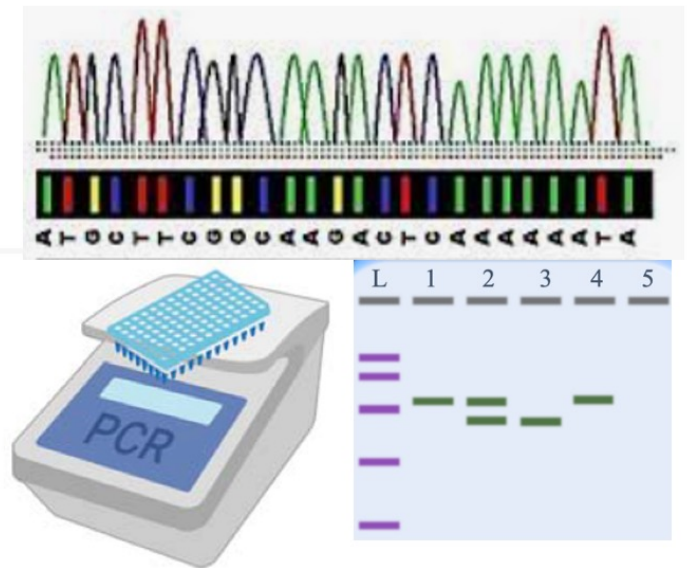
SNP - Single Nucleotide Polymorphism

Single nucleotide polymorphism (SNP)

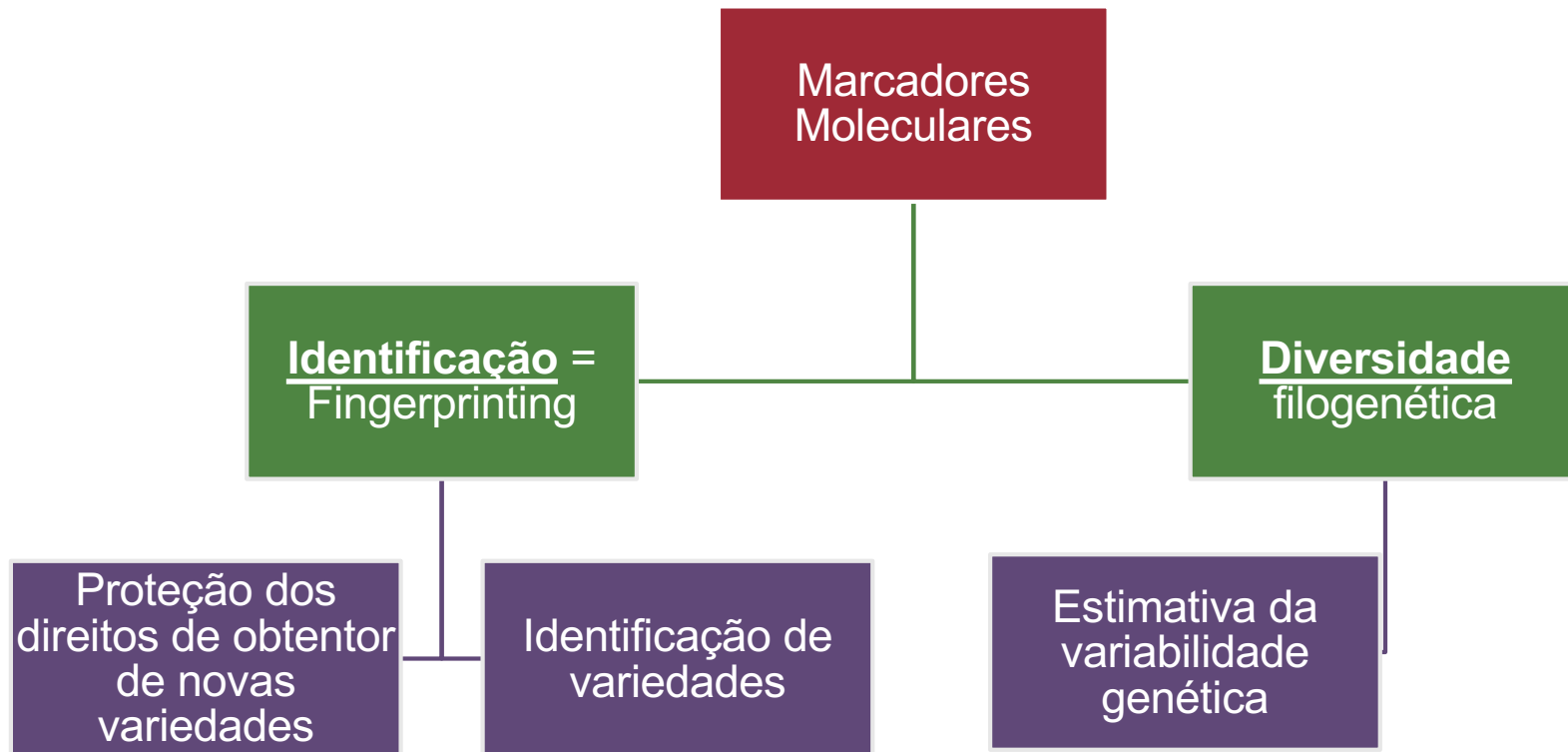
- variation at a single position in a DNA sequence – 1 bp
- SNP **within a gene**
 - gene with more than one allele
 - potential variations in the amino acid sequence
- SNPs in **noncoding** regions of DNA

Detected by:

- Sequencing
- PCR



Aplicações dos Marcadores Moleculares

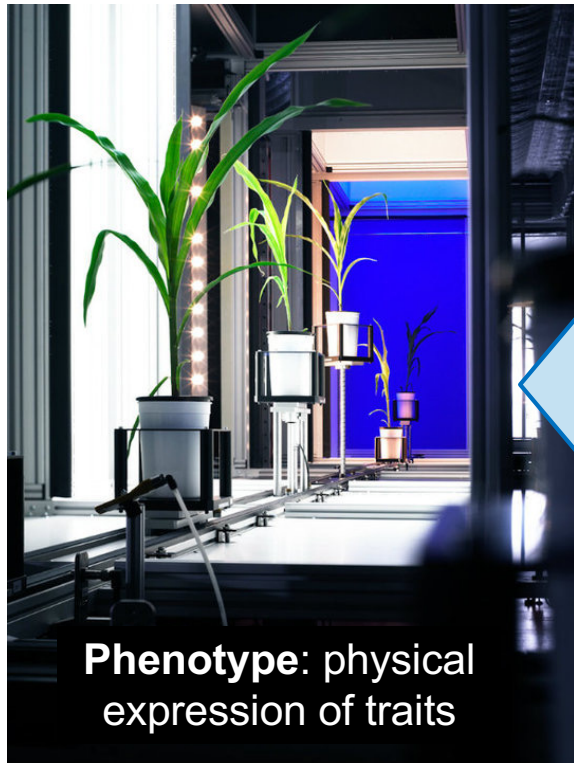


Using nuclear science in marker-assisted plant breeding



https://www.youtube.com/watch?v=H3DtOOTkjRc&t=1s&ab_channel=FoodandAgricultureOrganizationoftheUnitedNations
3'07"

Marker assisted selection (**MAS**)



Selecting for **DNA markers** is faster than selecting for **phenotype**

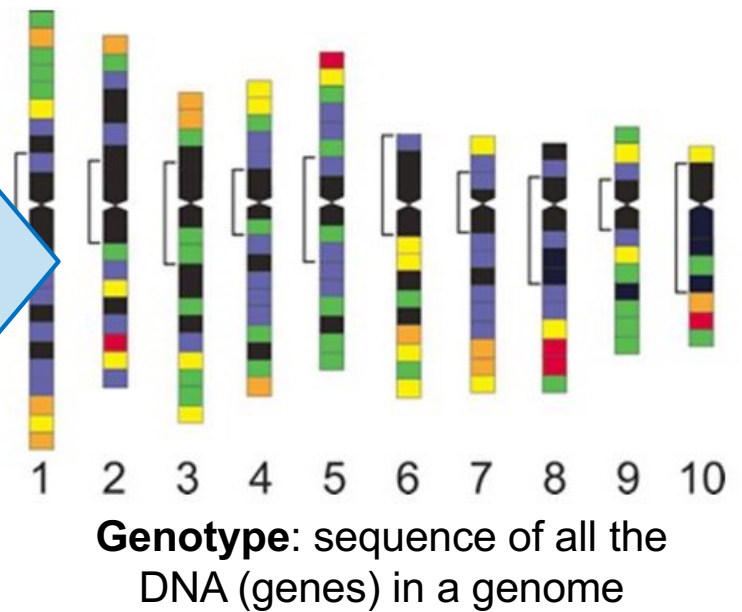


Photo credit [LemnaTec](#); Anderson, L.K., Lai, A., Stack, S.M., Rizzon, C. and Gaut, B.S. (2006). Uneven distribution of expressed sequence tag loci on maize pachytene chromosomes. *Genome Research*. 16: 115-122.

© 2013 American Society of Plant Biologists

Example: Introgression of a disease resistance gene



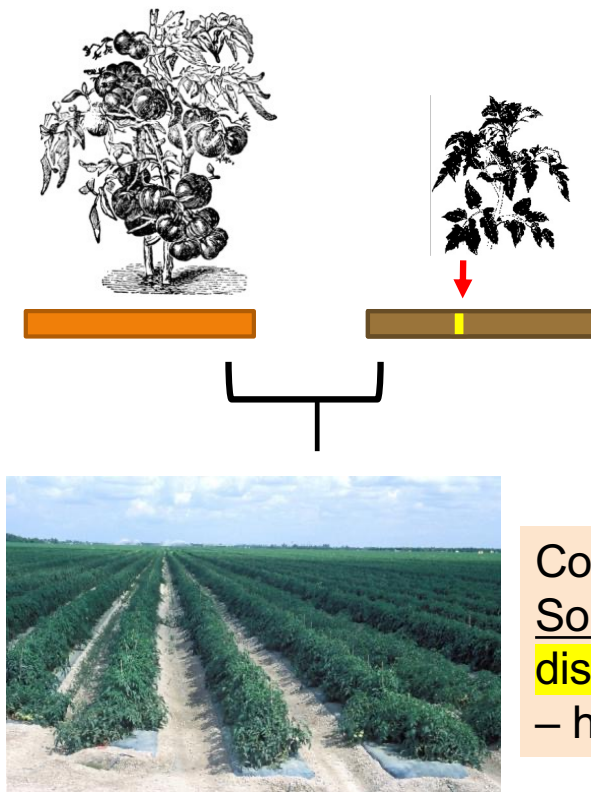
Elite tomato

How to add a
disease resistance
trait to an “elite”
tomato plant?



Poor tomato but
disease resistant
(resistance gene indicated)

Example: Introgression of a disease resistance gene

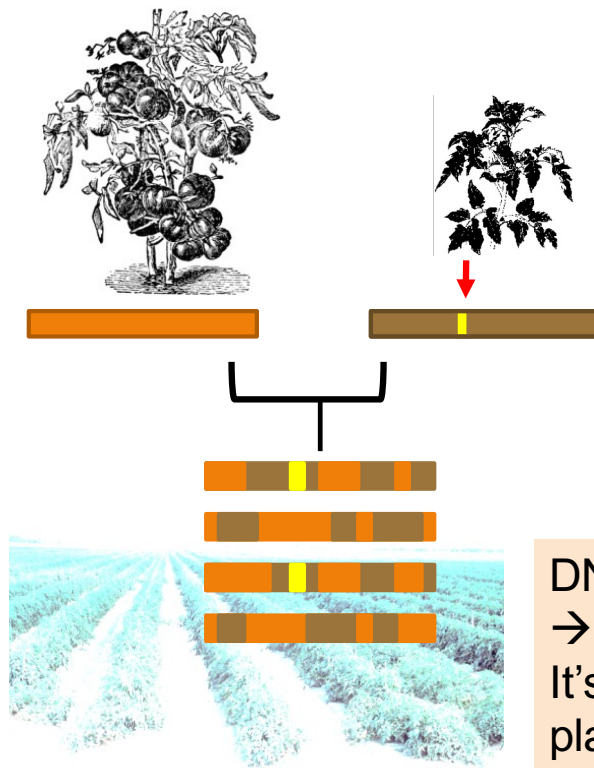


Cross the two plants
Some of their progeny inherit the
disease resistance trait, some don't
– how to identify the difference?

Photo by [Stephen Ausmus](#) USDA

© 2013 American Society of Plant Biologists

Example: Introgression of a disease resistance gene

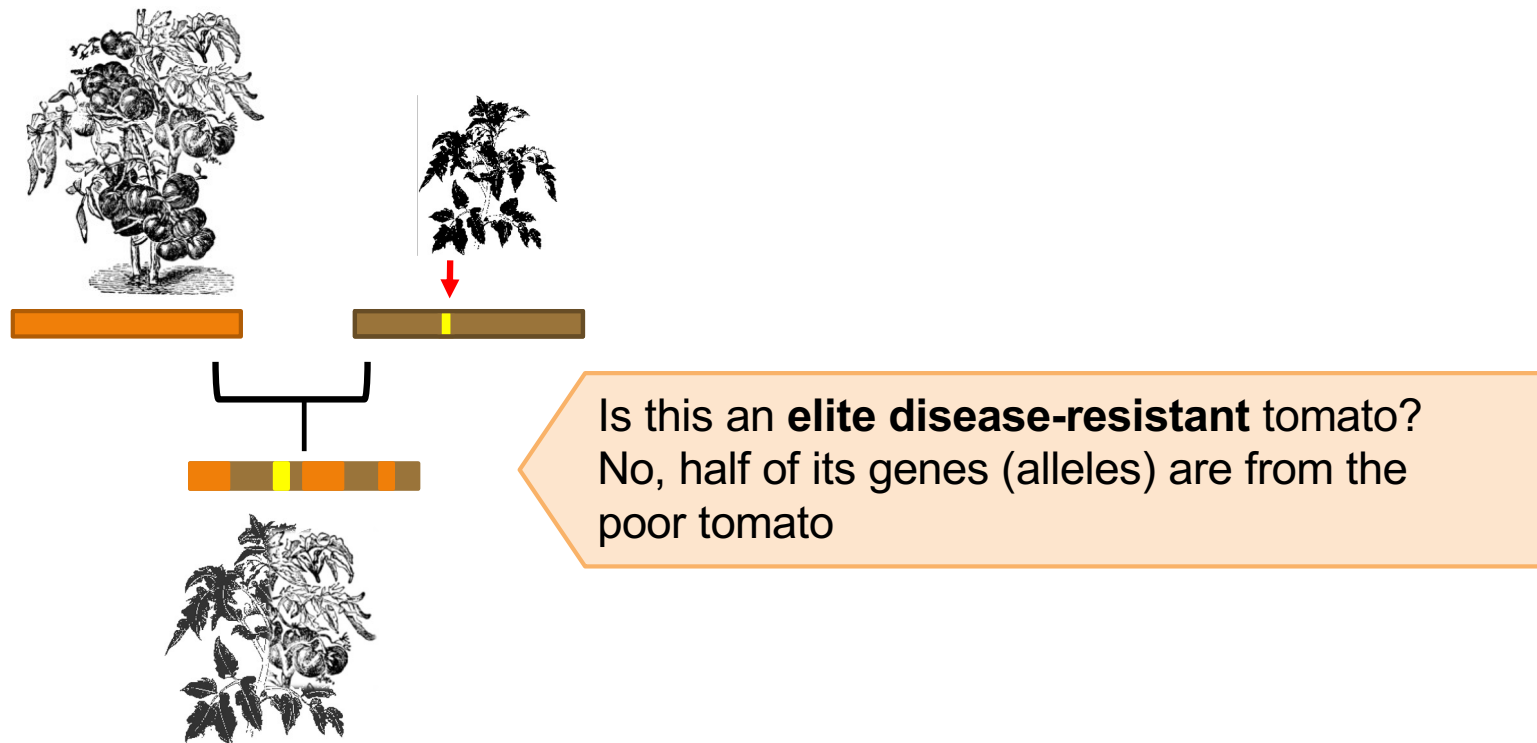


DNA molecular markers

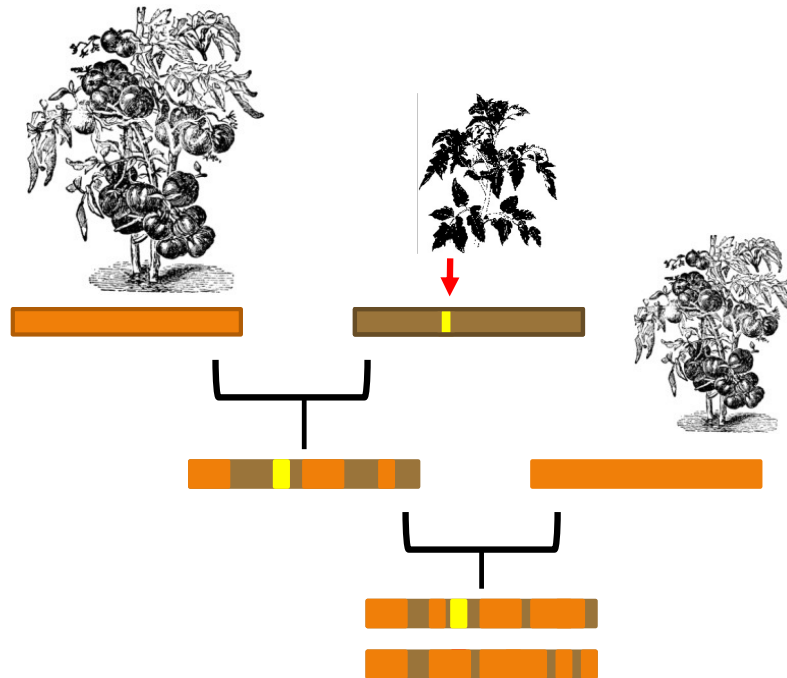
→ identify plants with the **resistance gene**

It's faster and easier than infecting all plants to see the phenotypes

Example: Introgression of a disease resistance gene



Example: Introgression of a disease resistance gene



Repeatedly cross back to the elite tomato, using **markers** to identify plants with the disease **resistance gene**

Example: Introgression of a disease resistance gene

