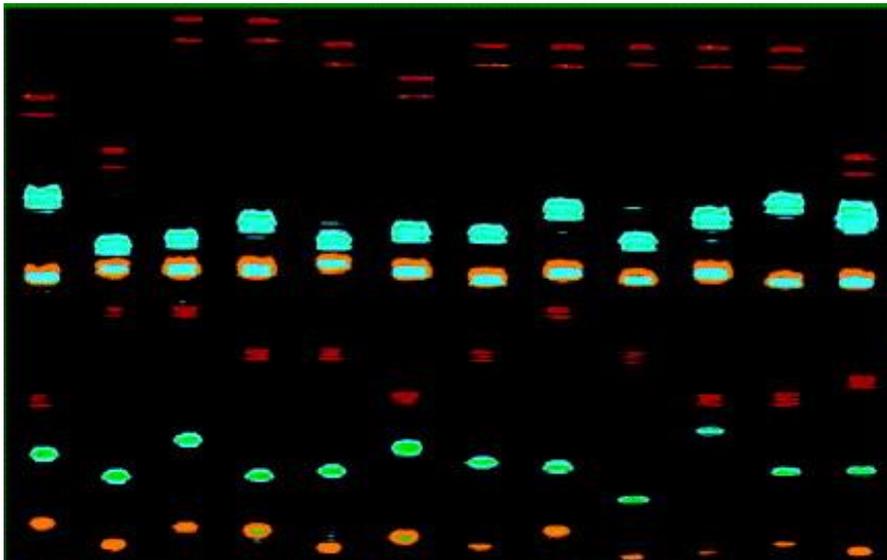


MARCADORES MOLECULARES E GENÉTICA DA CONSERVAÇÃO

Aula 10

LGN0218 – Genética Geral



Maria Carolina Quecine
mquecine@usp.br
Maria Lucia Carneiro Vieira
mlcvieir@usp.br
Departamento de Genética

Mas o que é um Marcado Molecular?

O ideal seria sequenciar todos os genótipos, mas.....

Se usa Marcador Genético ou Molecular como proxy!

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

MARCADORES MORFOLÓGICOS

- Monomorfismo



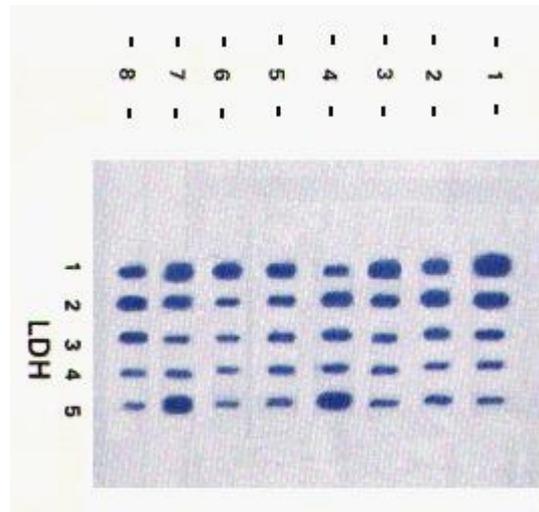
- Polimorfismo



MARCADORES BIOQUÍMICOS

- Variabilidade observada como resultado da tradução de RNA em proteínas

- Monomorfismo

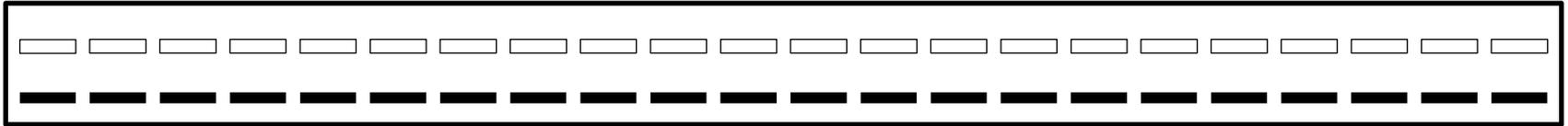


- Polimorfismo

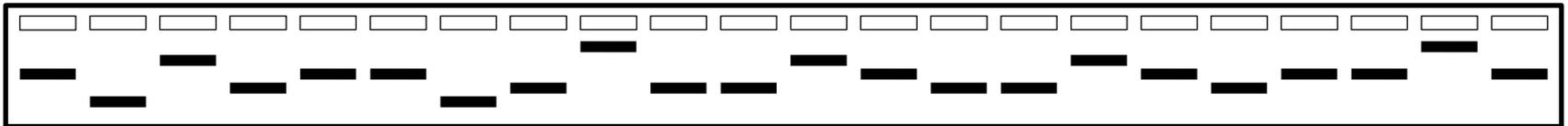


MARCADORES MOLECULARES

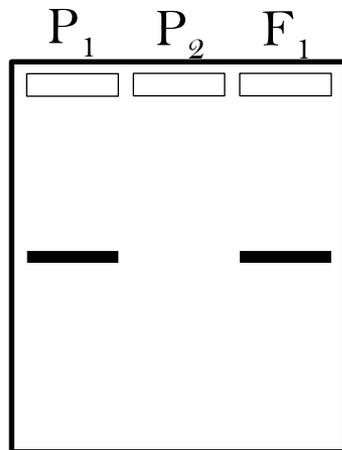
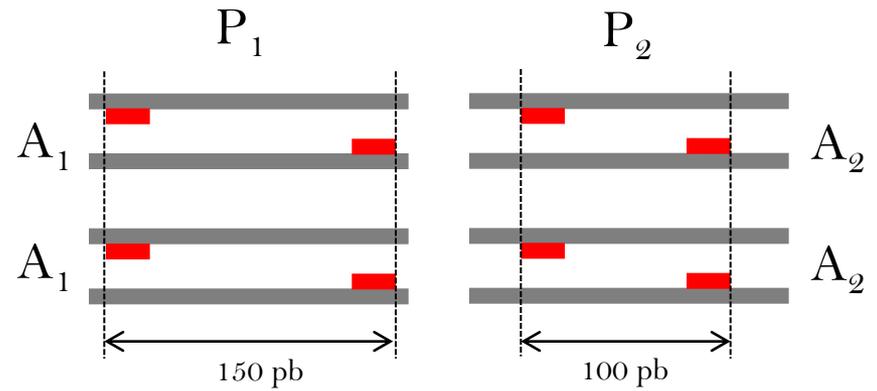
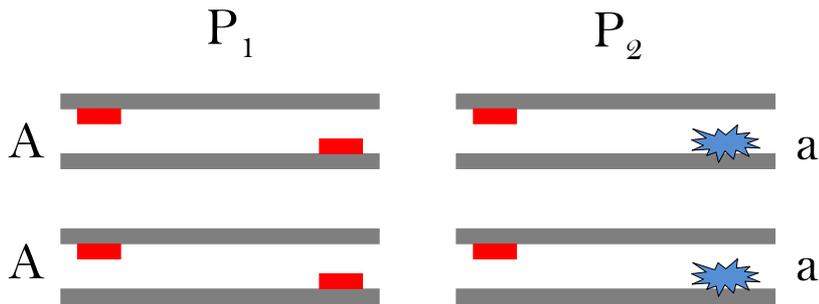
- Variabilidade surge por mutação, que é a base para identificação de marcadores
- Monomorfismo



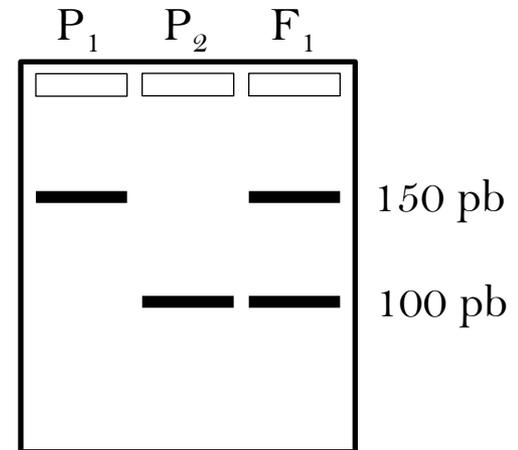
- Polimorfismo



MARCADORES DOMINANTES E CODOMINANTES?



Dominante



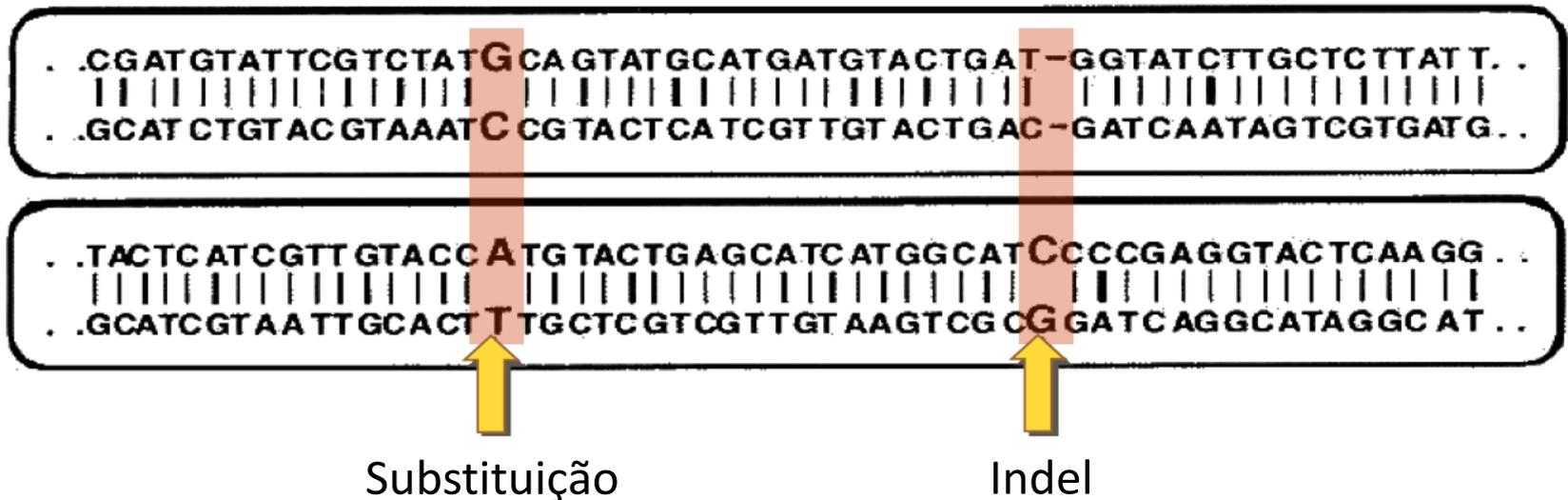
Codominante

MARCADORES MOLECULARES

Origem do polimorfismo

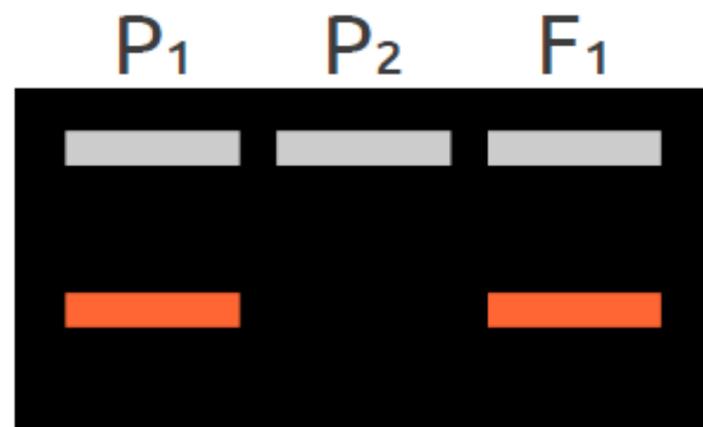
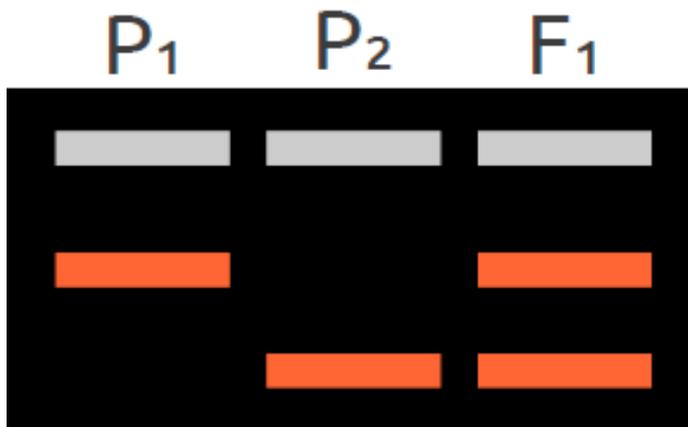
Polimorfismos de tamanho: originados por indels (inserções ou deleções)

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

MICROSSATÉLITES ou SSR

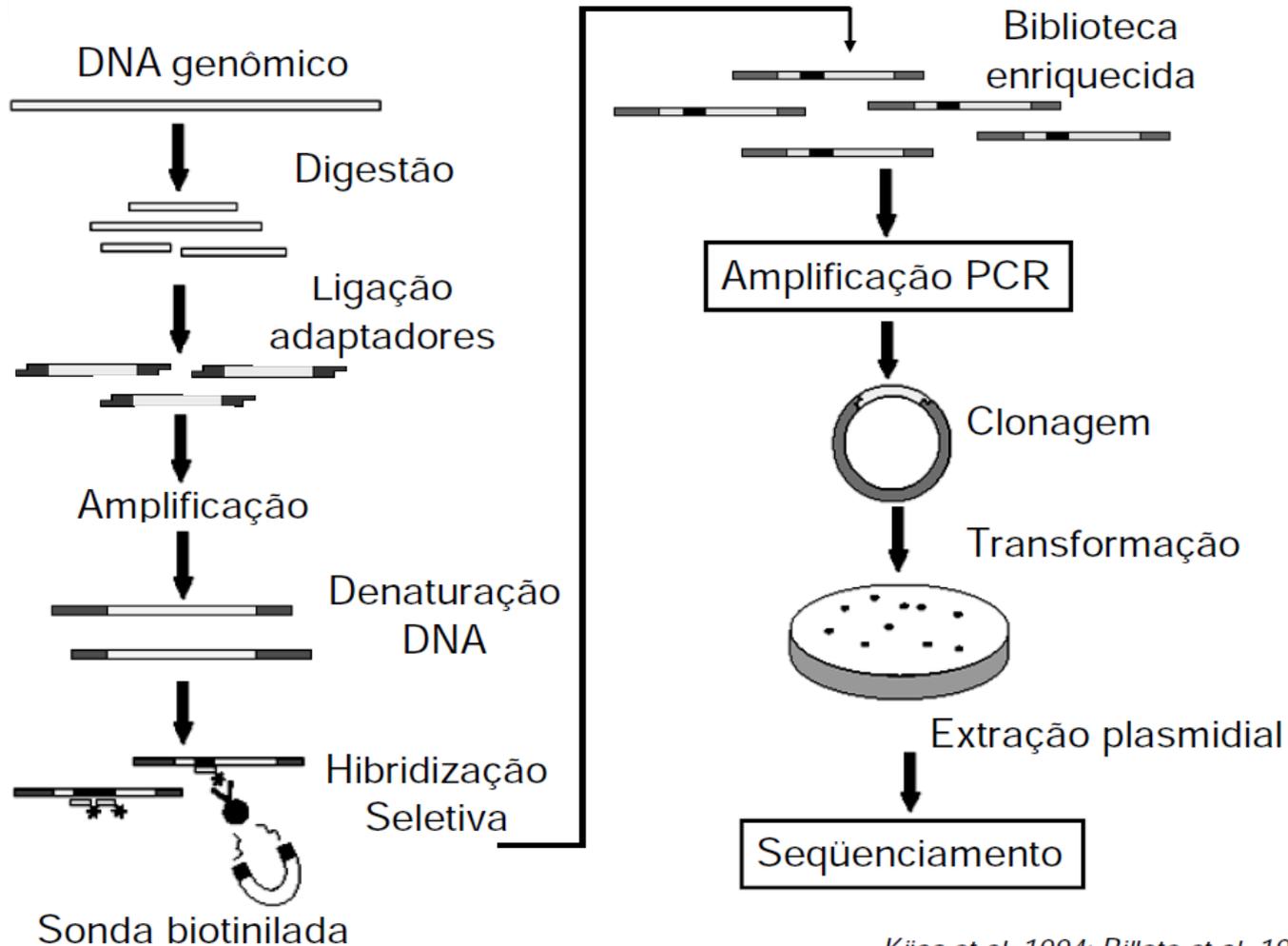
Simple Sequence Repeats

Sequências simples repetidas

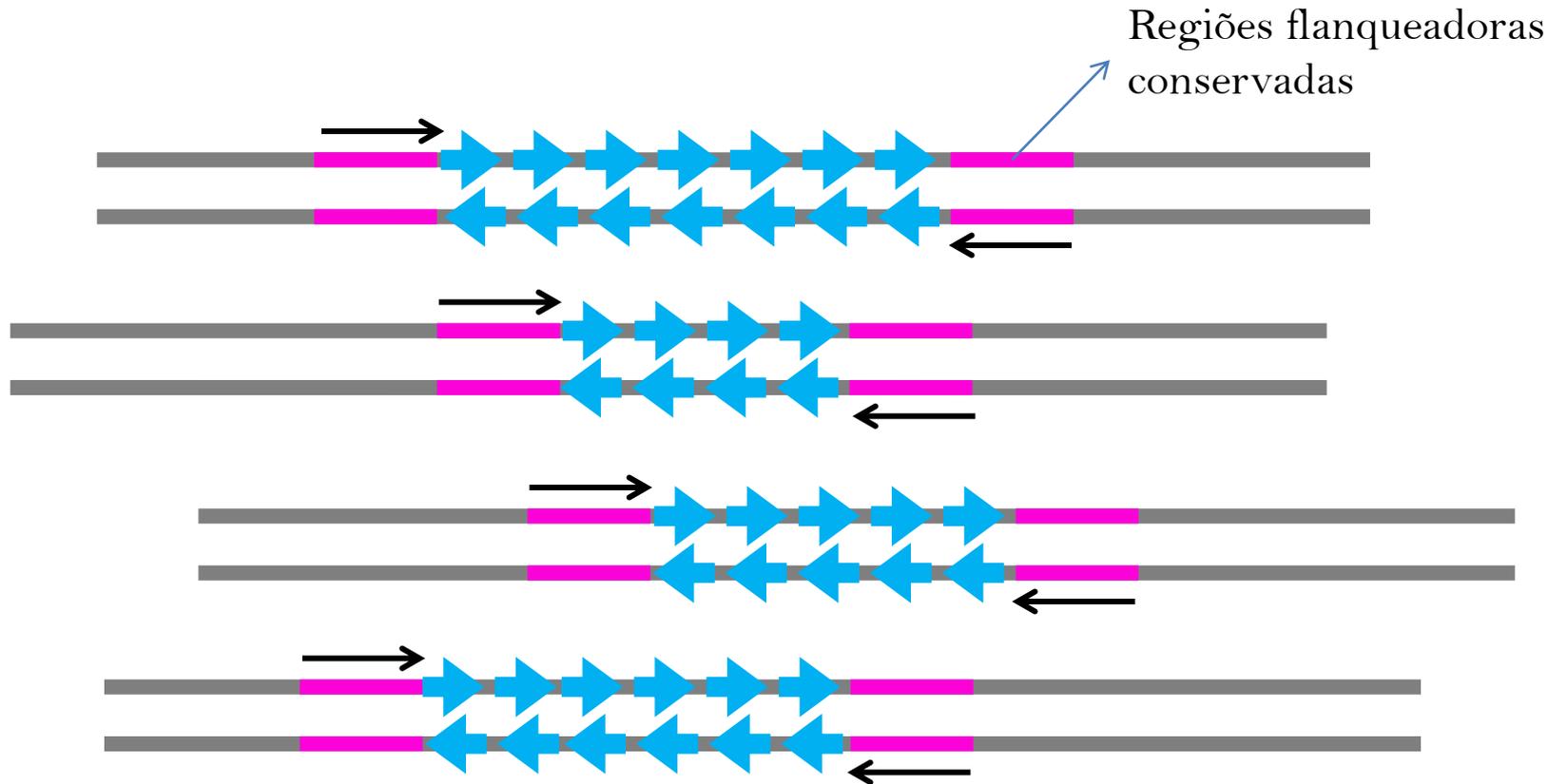
- Pequenas sequências de DNA (1 a 6 nucleotídeos) repetidos em tandem
- **Codominante**
 - Indivíduos heterozigotos são detectados
- Multialélicos e amplamente distribuído no genoma
- **Base genética:**
 - Amplificação de locos específicos de sequências repetitivas
 - Baseados em PCR com primers específicos

MICROSSATÉLITES

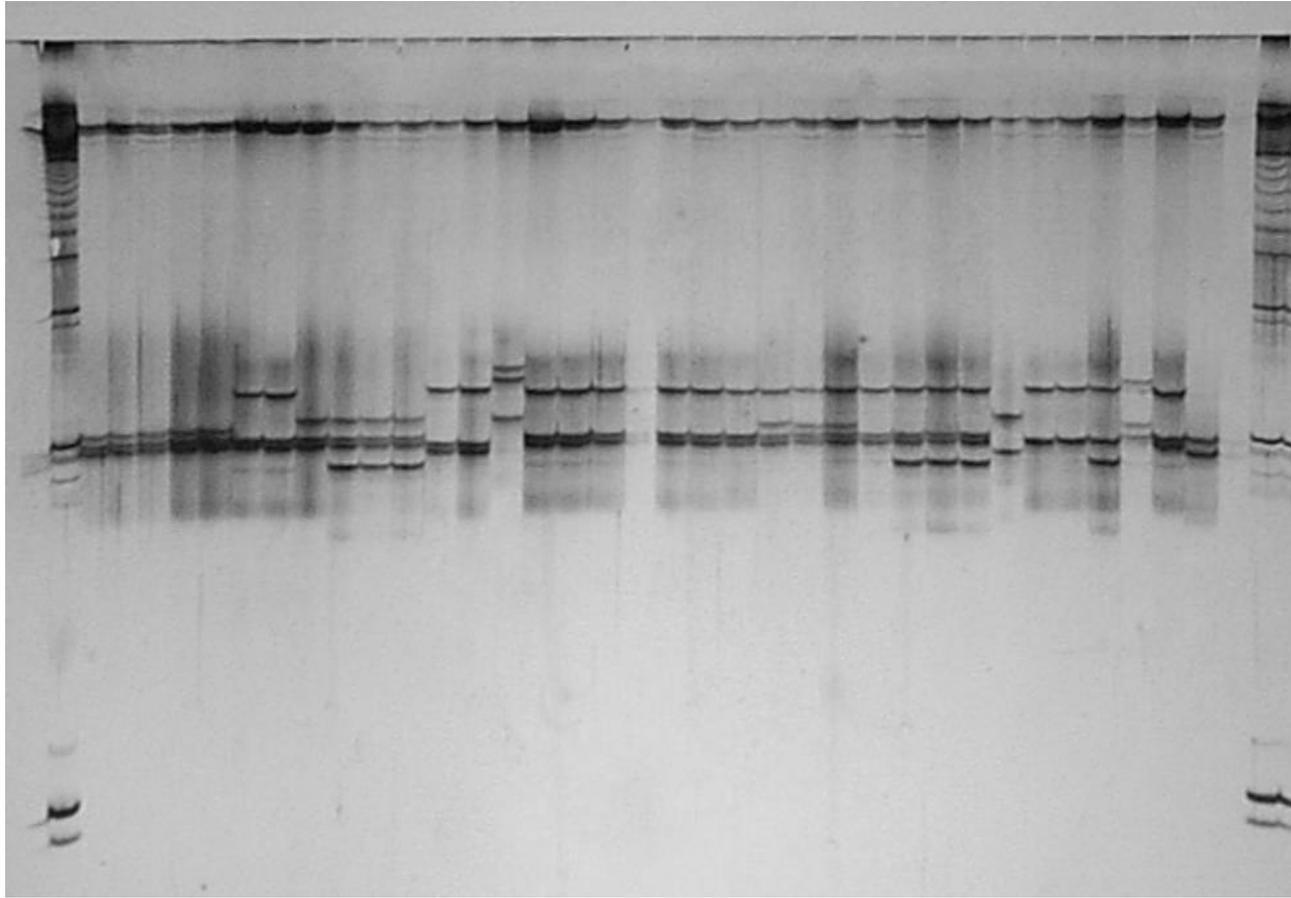
Biblioteca genômica enriquecida com microssatélites



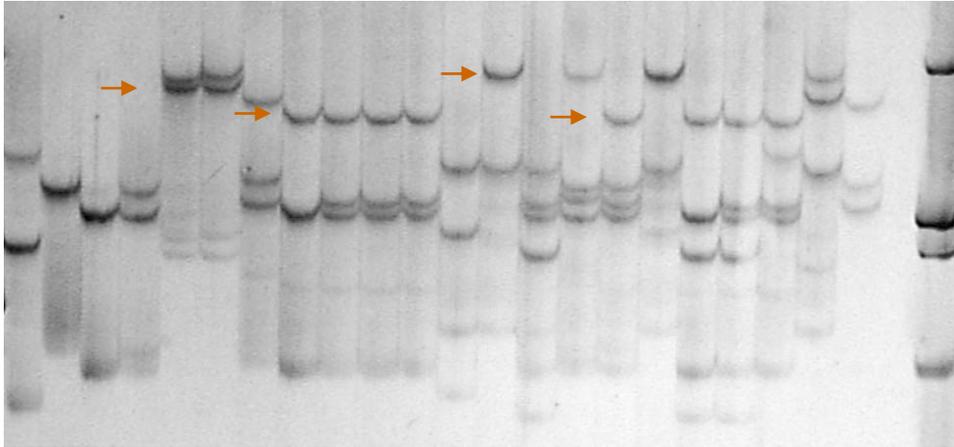
MICROSSATÉLITES



Microsatélites (SSR)

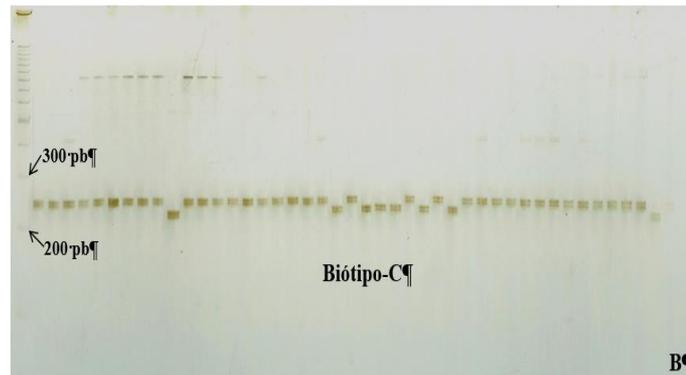
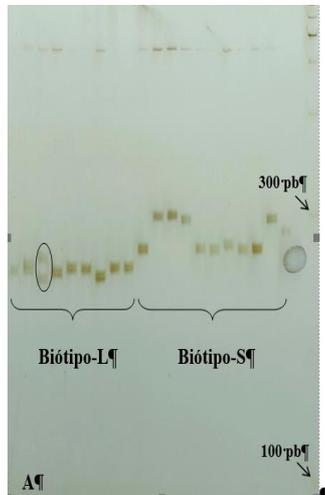


Microssatélites (SSR)



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

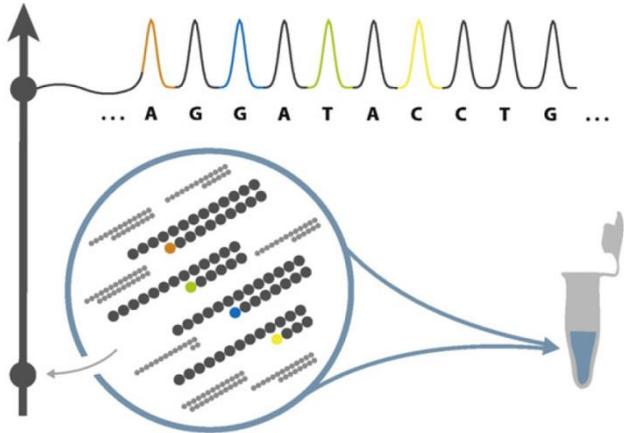
SNP

Single Nucleotide Polymorphism Polimorfismo de base única

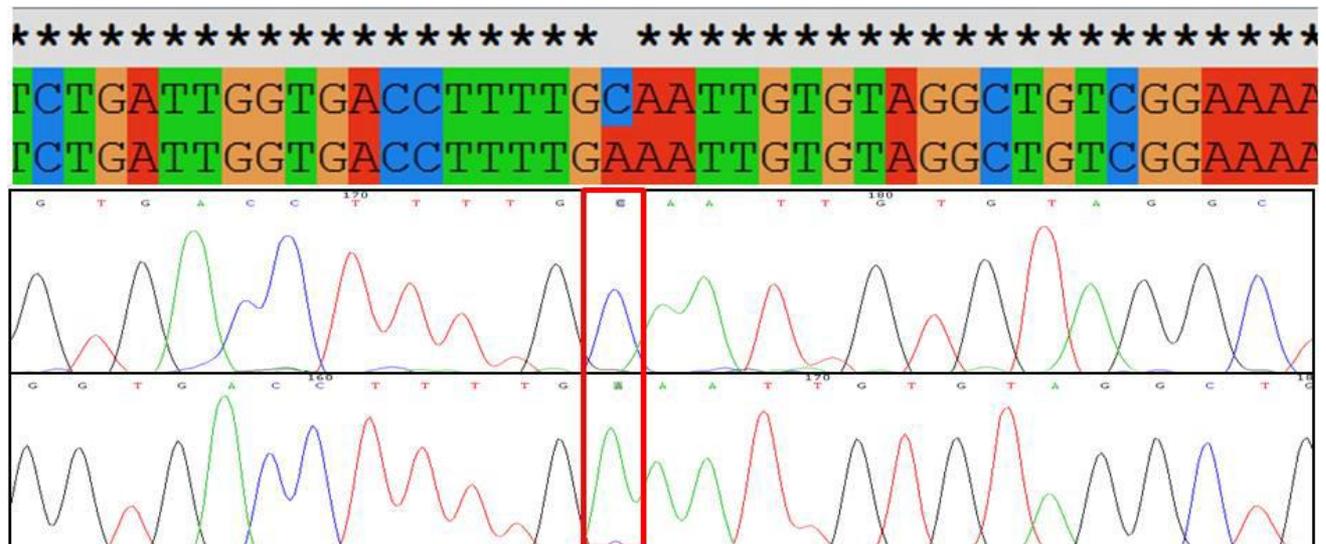
- Polimorfismo resultante da alteração de uma única base
- Abundantes no genoma
- **Codominante e bialélico**
- Há diversas maneiras de se avaliar a variação:
 - Alinhamento e comparação de sequências
 - Métodos baseados em géis
 - Métodos baseados em chips
 - Sequenciamento de 2ª geração, etc.

SNP

Sequenciamento



Alinhamento de
sequências



SNP

Exemplos de plataforma de genotipagem por SNPs

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

	SSR	SNP
Base genética	PCR com primers específicos	Sequencia-mento
Tipo de herança	Codominante	Codominante
Número de locos	Único	Único
Número de alelos	Vários	Dois

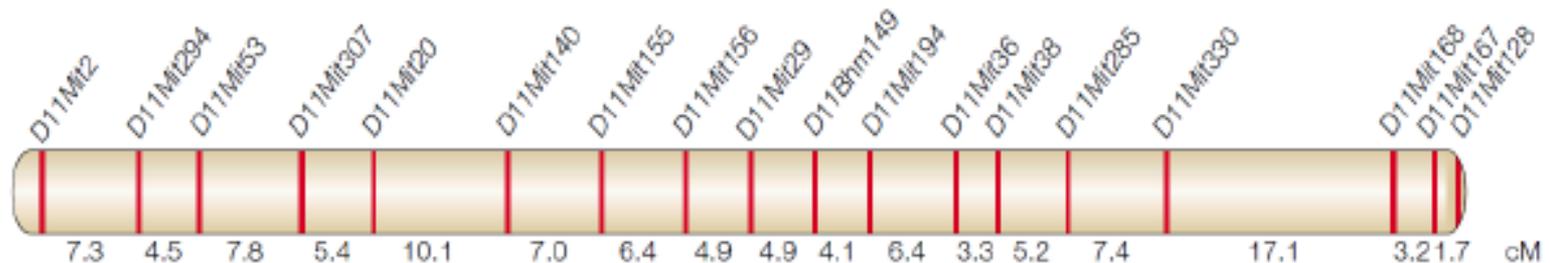
USO DOS MARCADORES EM ESTUDOS GENÉTICOS

CONSTRUÇÃO DE MAPAS GENÉTICOS

Mapa genético: representação de ordem e distância entre marcadores genéticos

Ligação genética

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

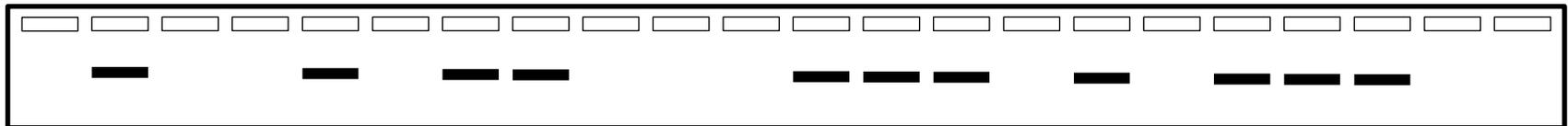
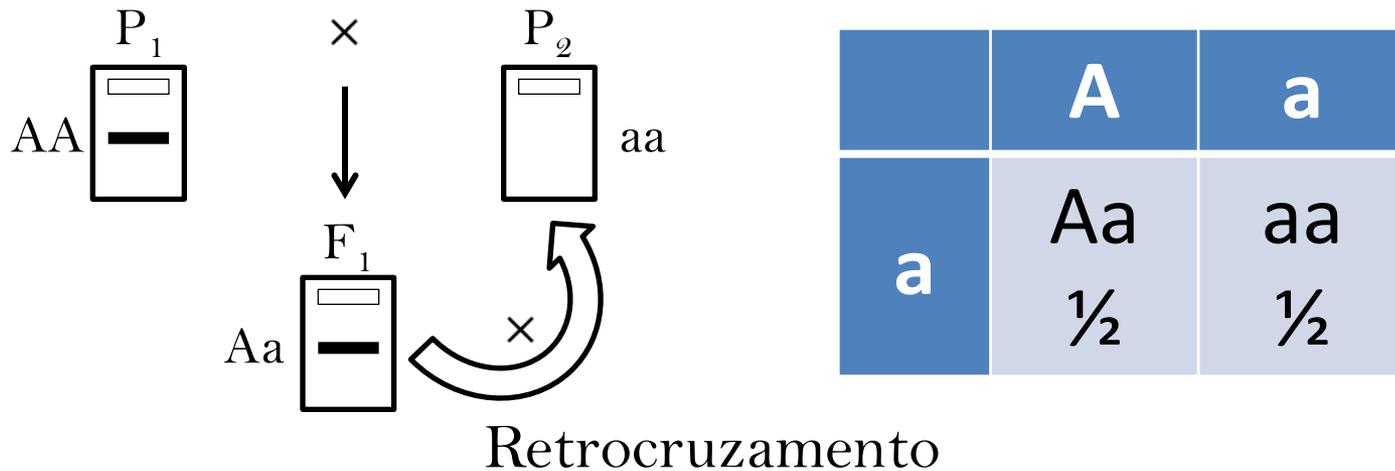


Mus musculus L.

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

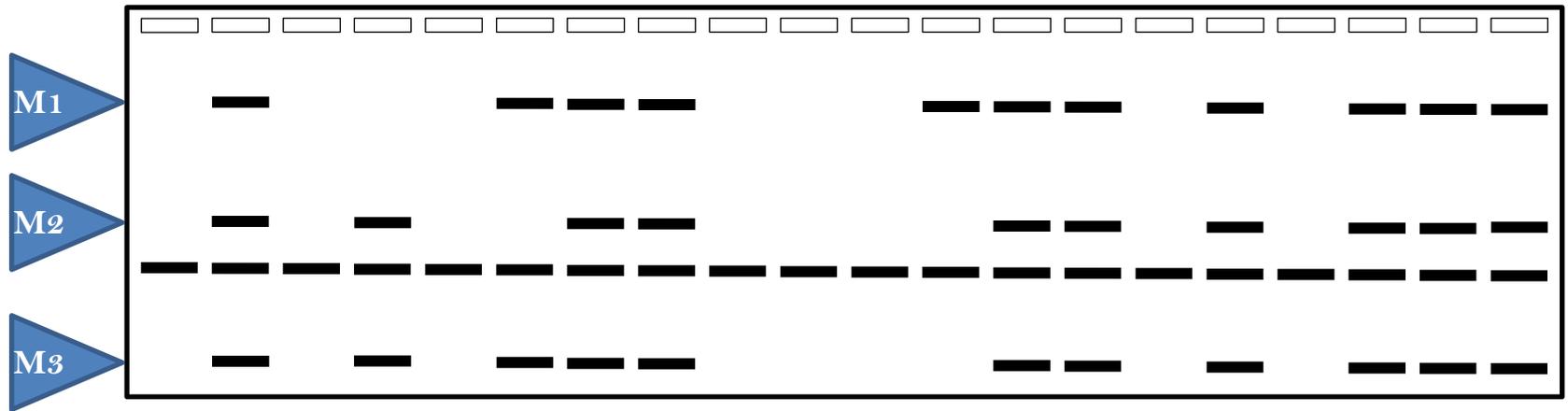
CONSTRUÇÃO DE MAPAS GENÉTICOS

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



CONSTRUÇÃO DE MAPAS GENÉTICOS

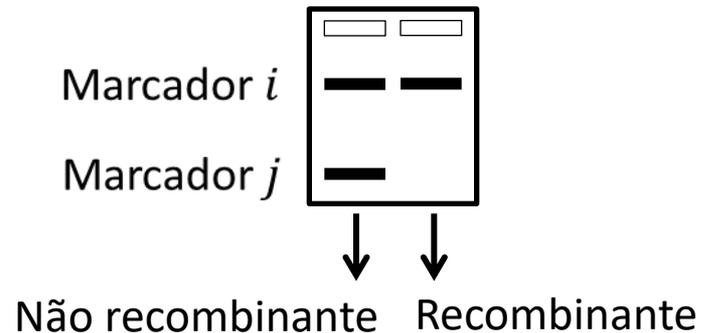
População de retrocruzamento genotipada com três locos AFLP



- Cálculo da fração de recombinação:

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

Comparando dois locos polimórficos



CONSTRUÇÃO DE MAPAS GENÉTICOS

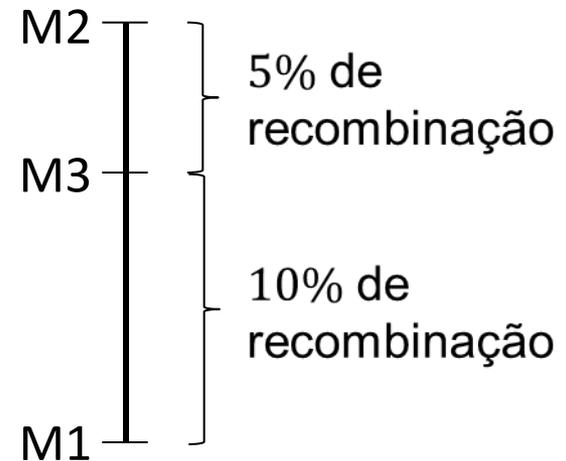
Leitura do gel

	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

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

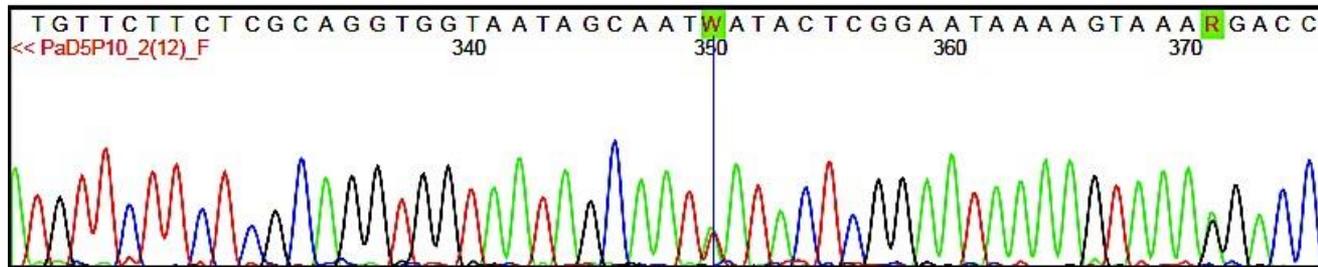
$$r_{1,3} = \frac{2}{20} = 0,10$$

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

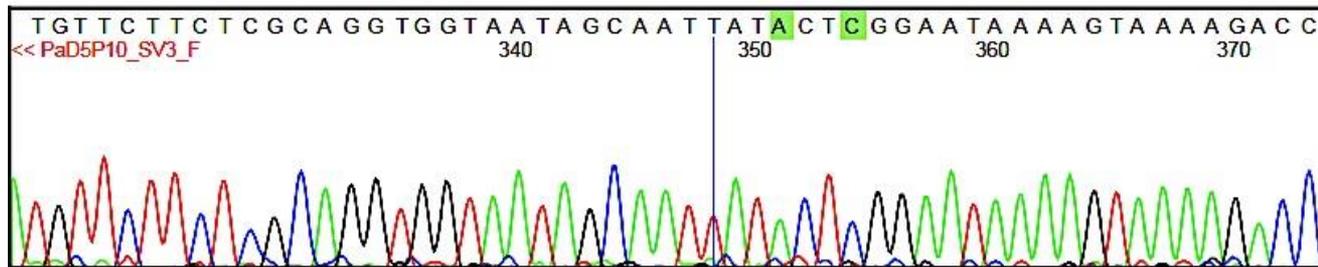


Mapa genético

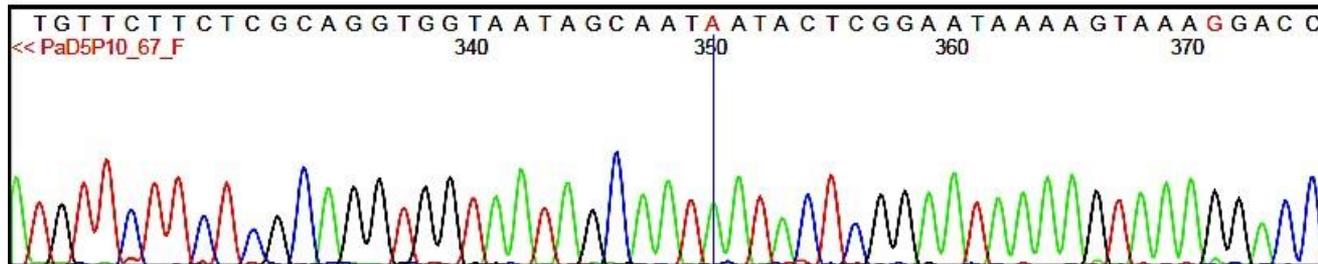
ESTUDOS DE GENÉTICA DE POPULAÇÕES



W (A/T)
= 200



T = 400



A = 100

Como se define a **estrutura genética** de uma população:

- Frequências genotípicas (por loco)

$$P = f(T) = 400 / 700 = 57,14\%$$

$$PQ = f(AT) = 200 / 700 = 28,58\%$$

$$Q = f(A) = 100 / 700 = 14,28\%$$

$$P + PQ + Q = 100\%$$

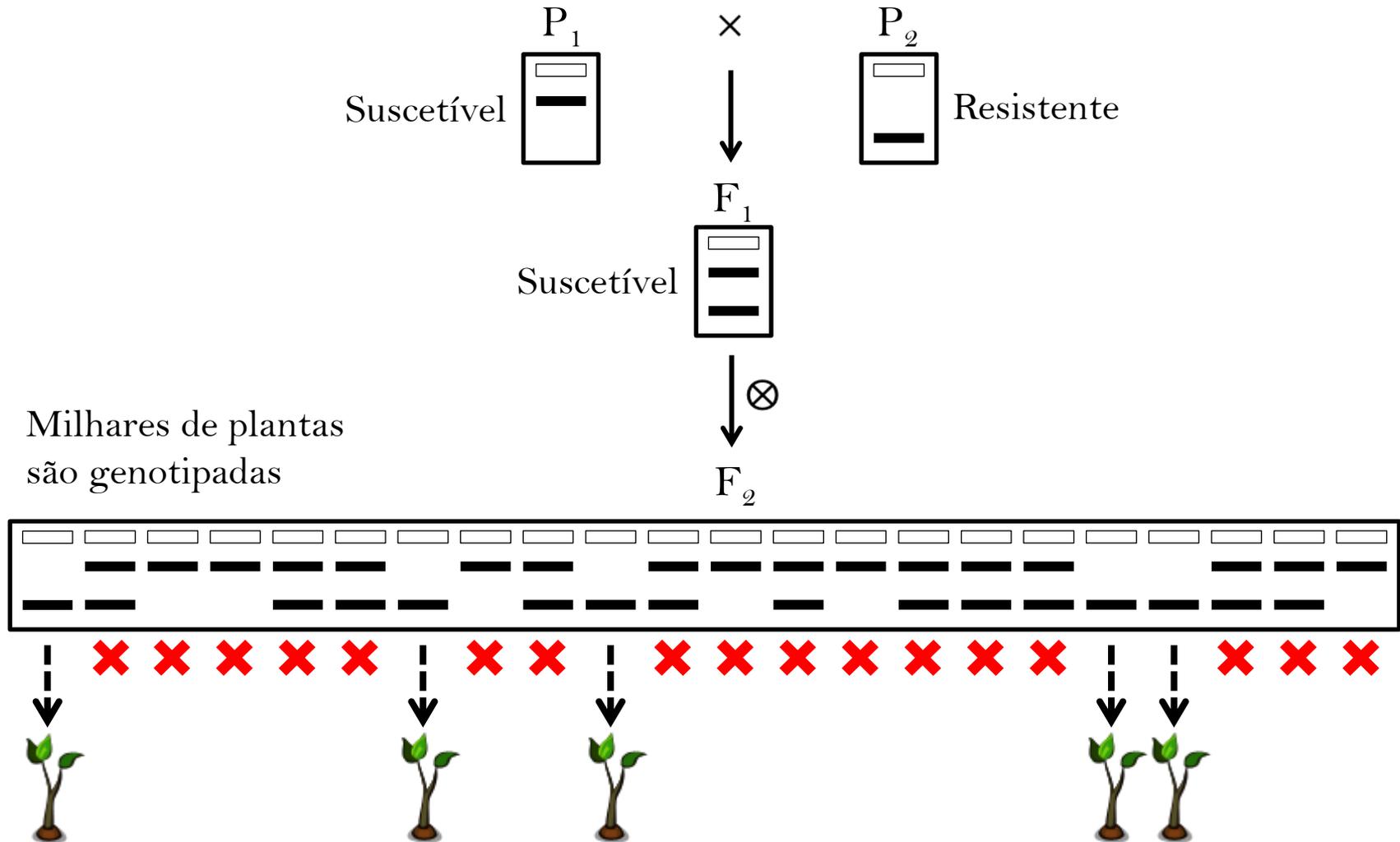
- Frequências alélicas ou gaméticas (por loco)

$$p = f(T) = P + \frac{1}{2} PQ / N = 400 + 100 / 700 = 0,7143$$

$$q = f(A) = Q + \frac{1}{2} PQ / N = 100 + 100 / 700 = 0,2857$$

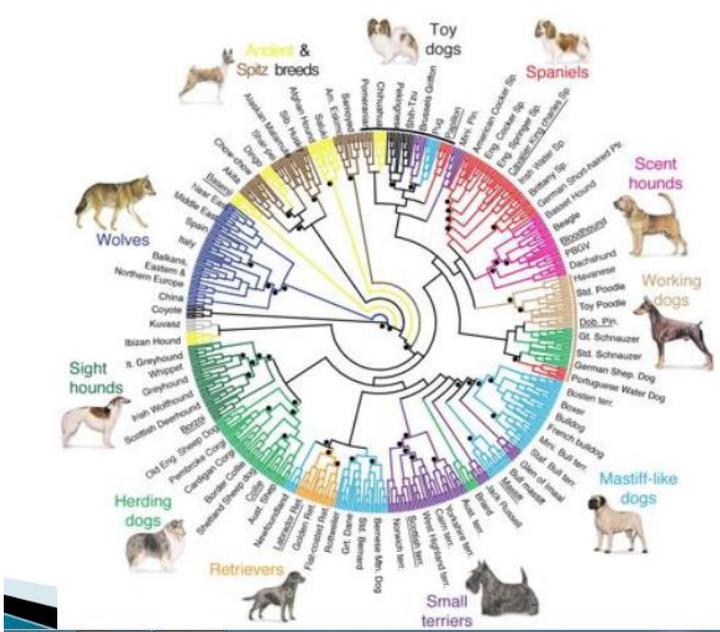
$$p + q = 1$$

SELEÇÃO ASSISTIDA POR MARCADORES

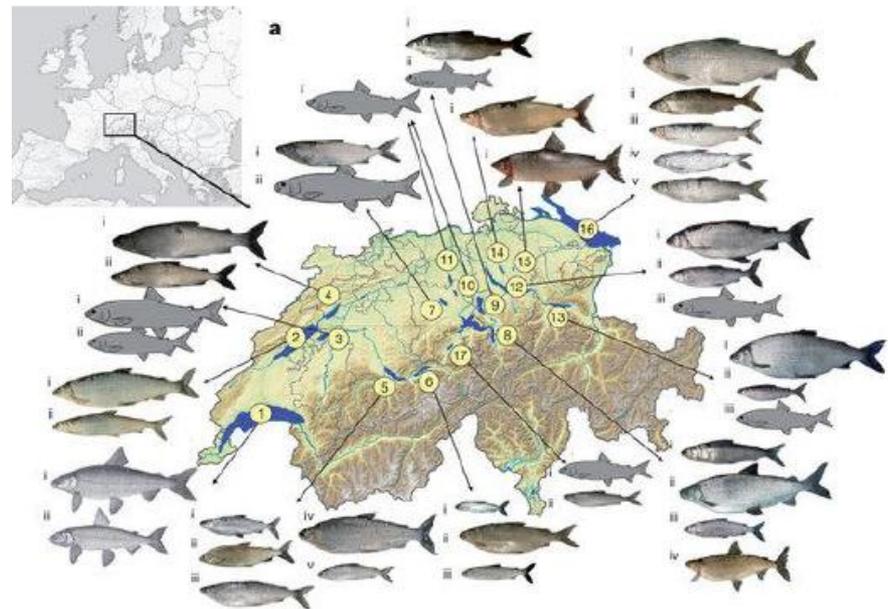


A seleção fenotípica é baseada em marcadores de DNA

USO DOS MARCADORES EM ESTUDOS DE DIVERSIDADE E CONSERVAÇÃO



Filogênia



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

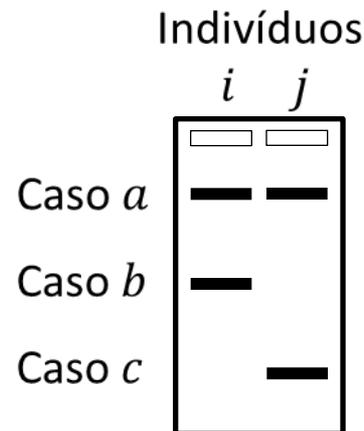
MARCADORES DOMINANTES

- Cálculo de similaridade genética: coeficiente de Jaccard

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

Em que:

- $a = \text{n}^{\circ}$ de locos com alelos presentes nos indivíduos i e j
- $b = \text{n}^{\circ}$ de locos com alelos presentes em i e ausentes em j
- $c = \text{n}^{\circ}$ de locos com alelos ausentes em i e presentes em j



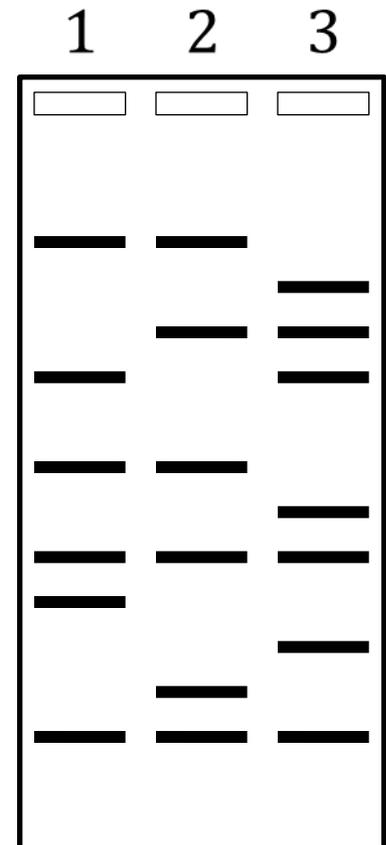
MARCADORES DOMINANTES

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

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

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

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



Genética da Conservação

A genética da conservação é um subcampo interdisciplinar da genética que visa compreender a dinâmica dos genes em populações principalmente para evitar a extinção. Por isso, aplica métodos genéticos para a conservação e restauração da biodiversidade.

A **Biologia da Conservação** estuda os indivíduos e as populações que tem sido afetadas por perda do seu habitat, exploração antrópica e mudanças ambientais. O conhecimento advindo dessas populações orienta as **decisões** que possam garantir a sua sobrevivência. A **Genética** estuda a transmissão dos caracteres de geração a geração e os genes responsáveis pela herança. As duas áreas da ciência formam a **Genética da Conservação**.

<https://learn.genetics.utah.edu/content/science/conservation/>

**A Genética da Conservação
combina abordagens da
ecologia, genética de
populações, modelagem
matemática, e da taxonomia
evolutiva.**



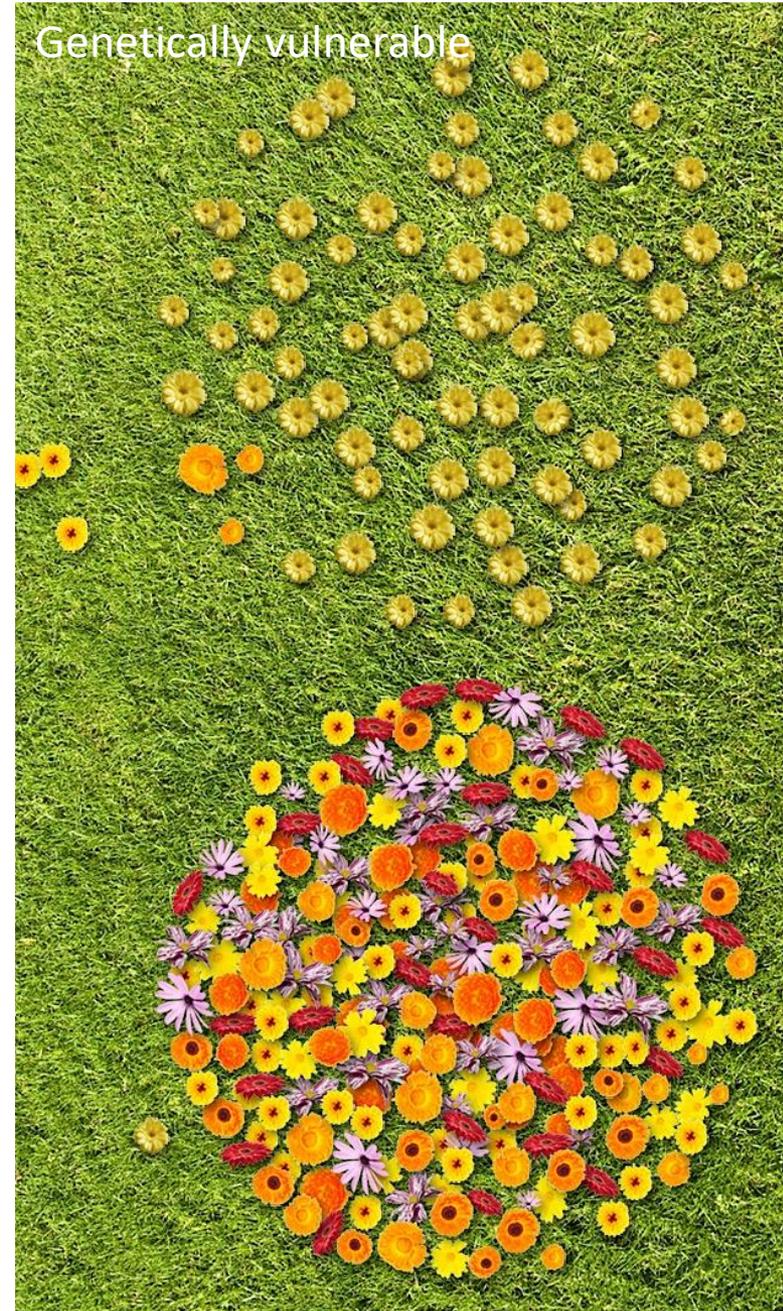
É uma ciência **fundamental**
mas também **aplicada**.
Primeiramente, os cientistas
devem entender as relações
genéticas e taxonômicas entre
os organismos sob estudo.

Os organismos estudados pelos conservacionistas pertencem a **populações ameaçadas.**

O que levou essas populações ao risco de **extinção**, e quais medidas devem ser tomadas para **reverter** esta tendência?

O conhecimento sobre a diversidade genética ajuda aos cientistas a reverter essa situação. Sem a **análise genética**, recursos valiosos são perdidos.

À direita: há alelos distintos que codificam diferentes cores da flor.



Principais fatores que interferem nas estratégias de conservação da diversidade natural:

- 1. Mudanças no tamanho das populações
- 2. Isolamento geográfico
- 3. Adequada identificação e inventário dos organismos
- 4. Análise da diversidade com base em informação científica
- 5. Interpretação e manejo
- 5. Mudanças climáticas drásticas
- 5. Decisão política

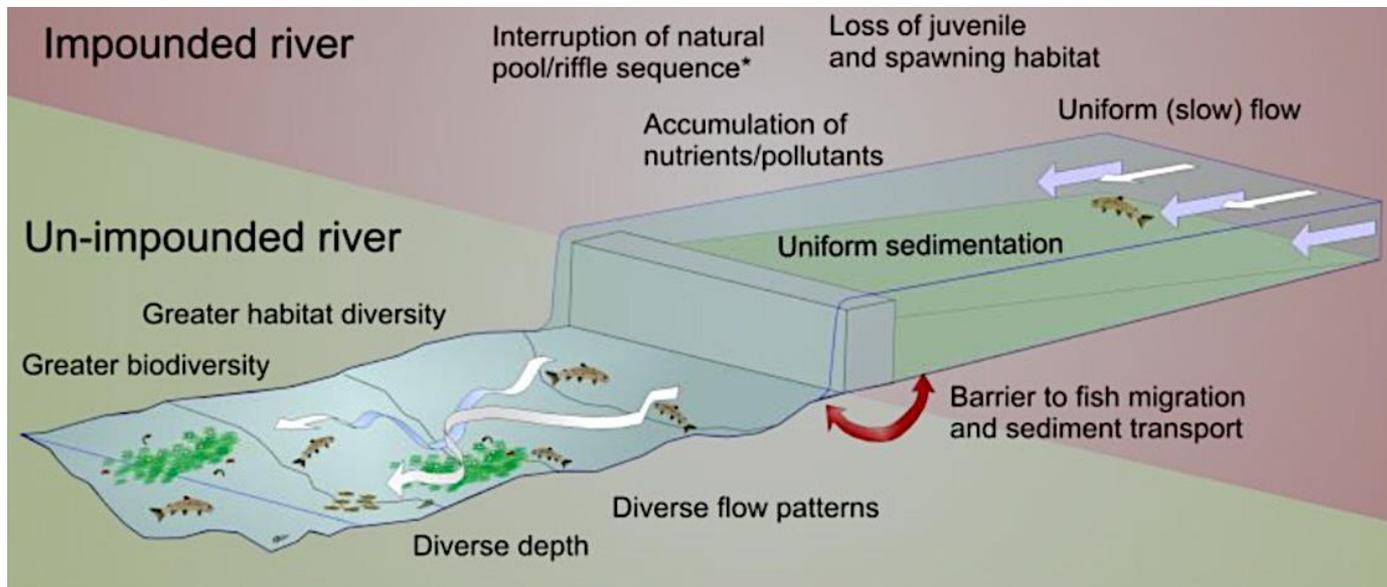
1. Tamanho da população :

A vigilância de **pequenas populações** é fundamental, porque elas são particularmente sensíveis à mudança. Catástrofes naturais, mudanças ambientais ou mutações genéticas podem causar uma diminuição no tamanho da população. Quando a população de uma espécie é pequena para começar, uma maior redução de seus números remanescentes pode **reduzir a diversidade genética**.



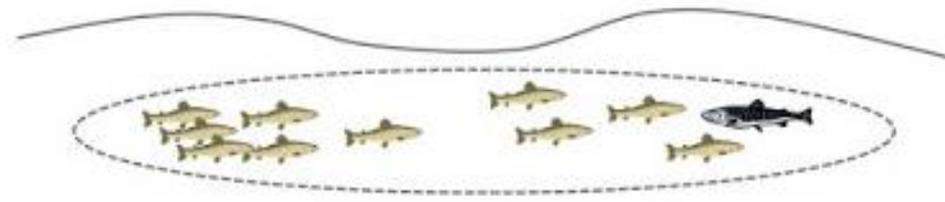
2. Isolamento Geográfico

Mesmo uma grande população pode perder diversidade genética. O isolamento geográfico pode acontecer se uma nova barreira for imposta através de um habitat. Por exemplo, se um reservatório de rio é construído por uma barragem, uma população de peixes pode ser dividida em dois grupos. Por acaso, o conjunto de variantes genéticas nas duas populações separadas pode diferir uma da outra.

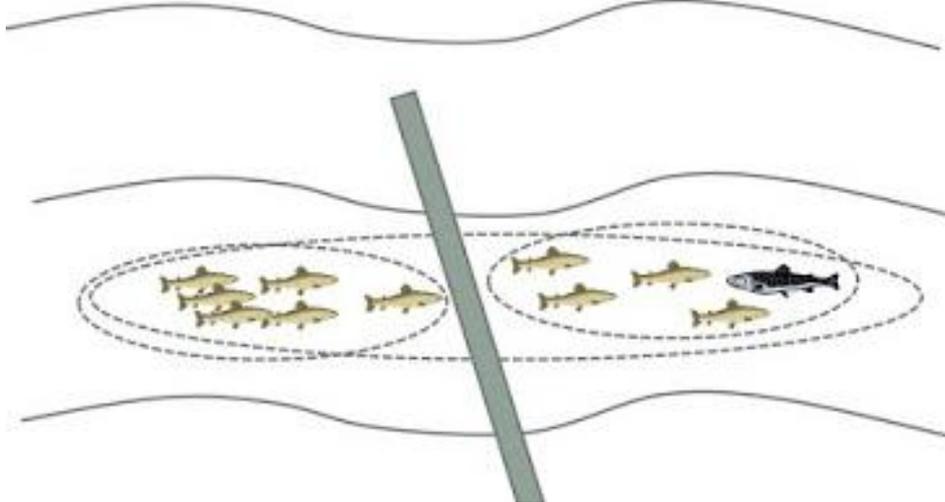




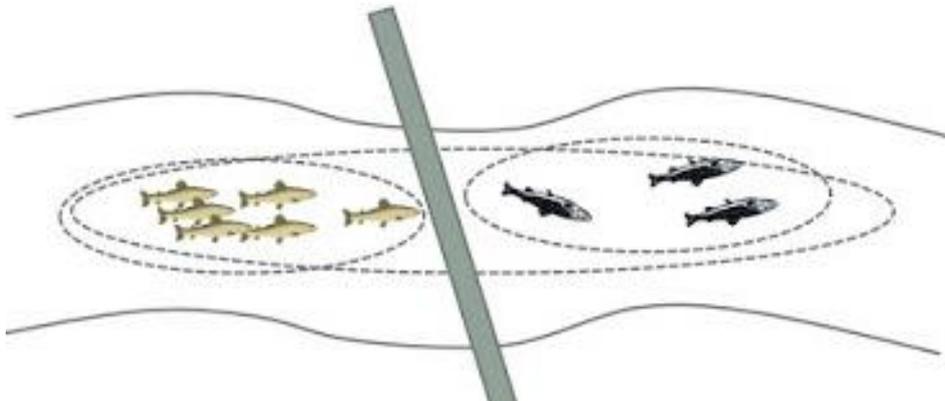
O isolamento geográfico pode causar distintos efeitos em diferentes grupos de espécies: plantas, insetos, aves, mamíferos



- Ocorrência de uma espécie abundante e uma rara



- Construção de uma barragem



- Isolamento da espécie rara e possível extinção

3. Identificação, inventário e análises

a. Definir populações e áreas de interesse. Como há muitos organismos, as espécies ameaçadas de extinção geralmente têm prioridade.

b. Observe a população. Quais são as formas conhecidas da espécie? Quais são os parentes conhecidos da espécie? Quais são as características físicas utilizadas para classificar as diferentes formas e espécies?



c. Formar hipóteses sobre relações entre populações ou espécies e teste essas hipóteses examinando características genéticas (dados de DNA ou proteínas).

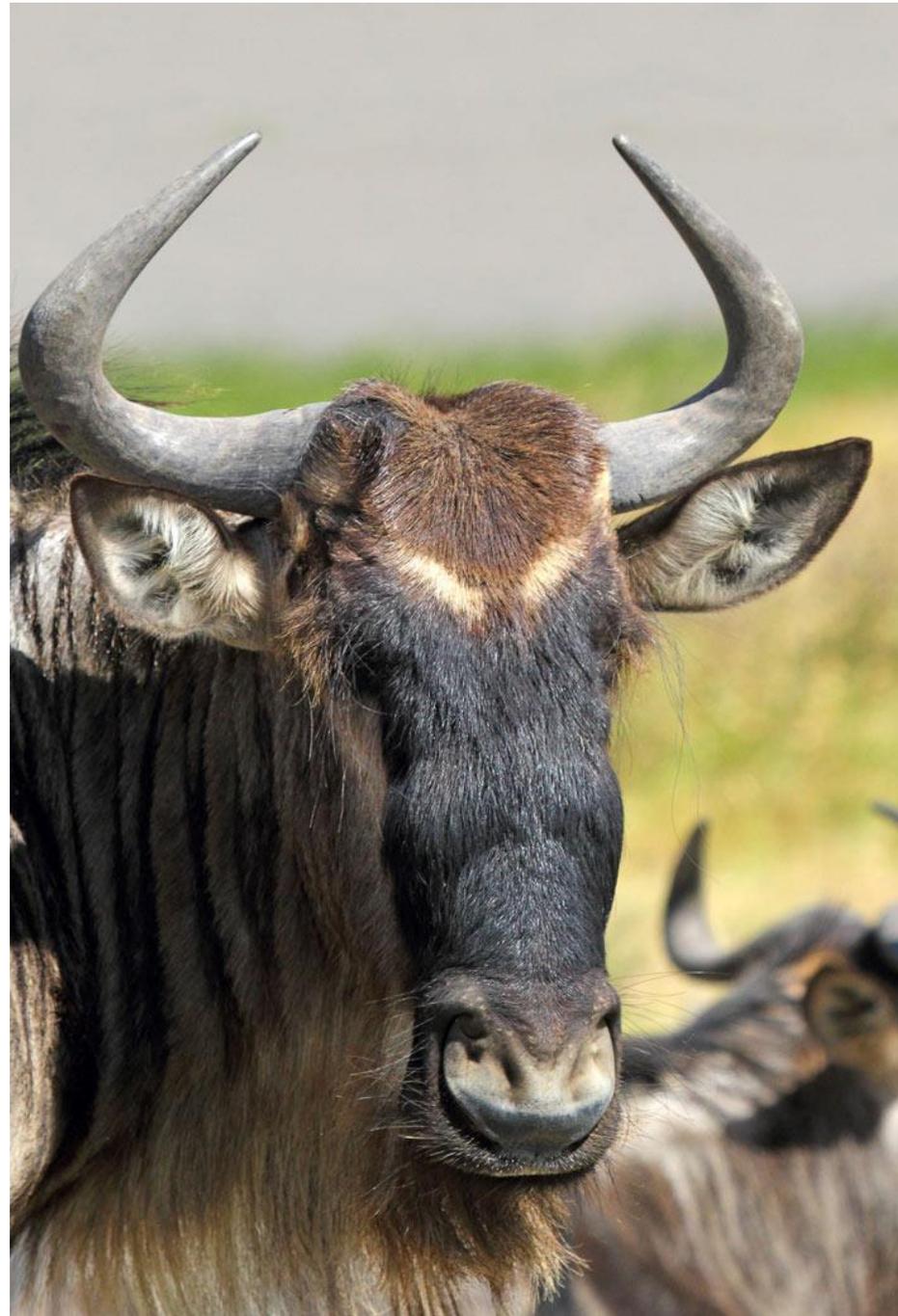
d. Utilizar modelos matemáticos para analisar os dados. Determinar quanta diversidade existe em populações separadas da espécie, bem como a taxa em que os genes são trocados entre as populações (fluxo gênico).



4. Interpretação e manejo

Cientistas e gestores trabalham juntos para identificar organismos ameaçados. Para desenvolver uma estratégia de manejo, eles investigam o habitat do organismo:

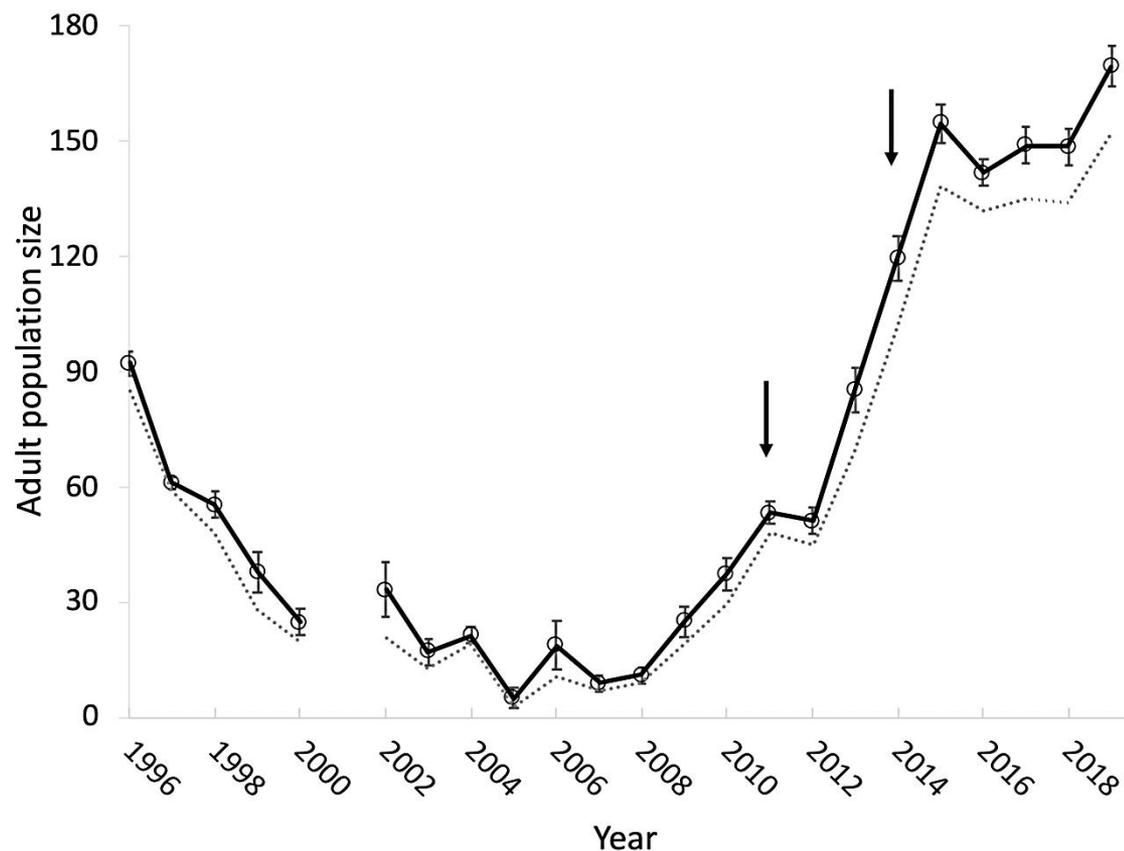
a. Determinar o grau em que o organismo é adaptável a várias temperaturas, solos e condições de água.



b. Examinar fatores que influenciam a diversidade genética, como polinizadores de plantas. A saúde das espécies polinizadoras pode ser crítica para a sobrevivência de uma espécie vegetal ameaçada de extinção.

c. Estudar ameaças à integridade do habitat da espécie, incluindo fatores humanos e climáticos. Uma vez que todos os aspectos da população e seu ambiente são compreendidos, os cientistas podem desenvolver um plano de preservação inteligente.

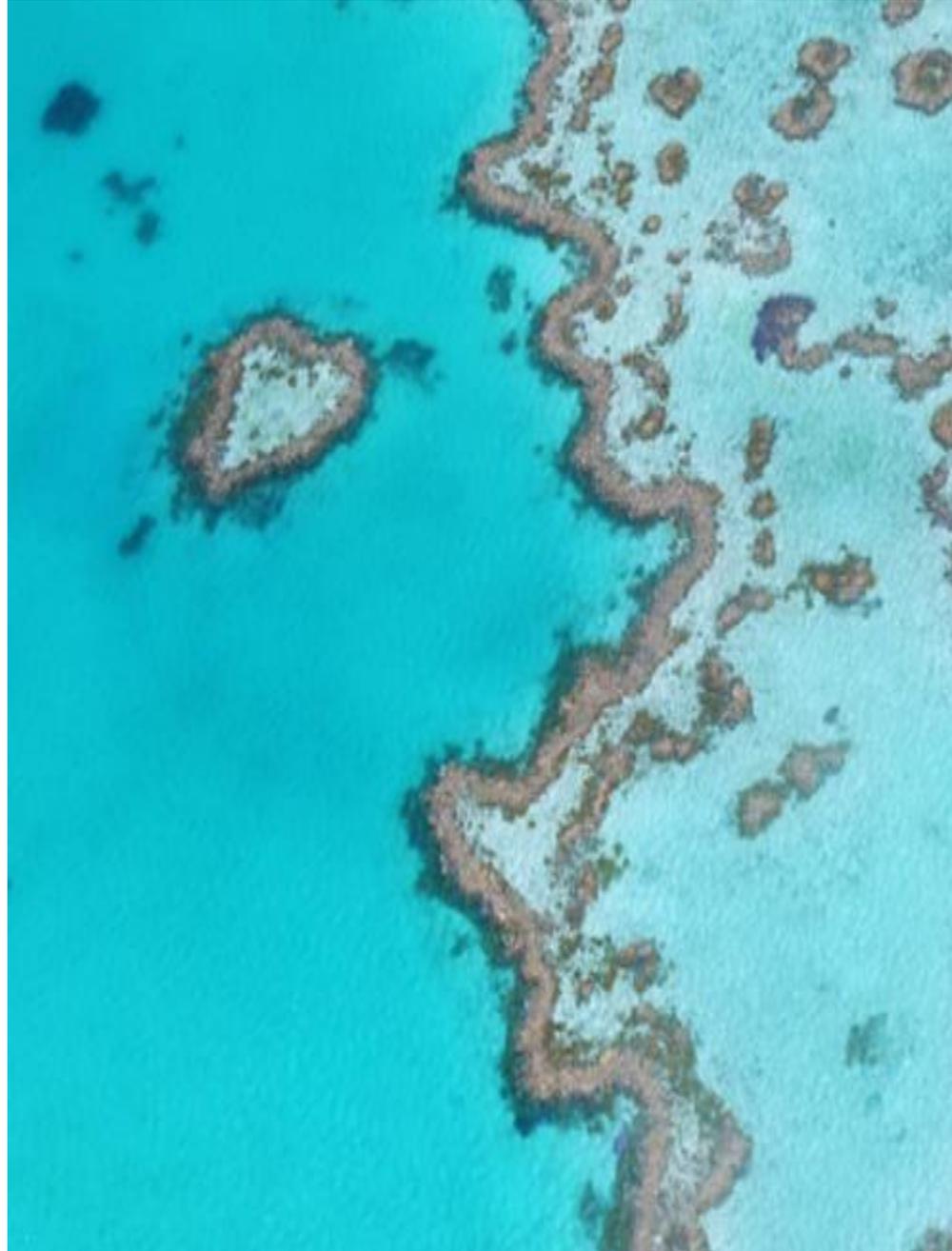




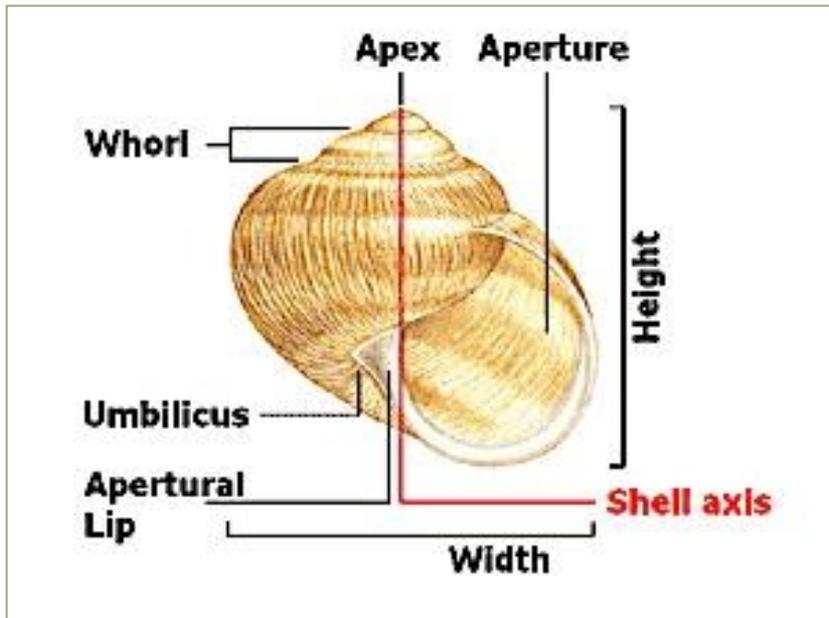
Tamanho da população adulta de um marsupial em Mount Buller, Austrália. As médias ao longo dos anos (barras são desvios) são conectadas pela linha sólida. As setas indicam os anos em que seis machos foram introduzidos das populações do Monte Higginbotham para a população do Monte Buller.

5. Mudanças climáticas

As mudanças climáticas drásticas e que ocorrem em curto período de tempo afetam a diversidade genética, resultando em efeito gargalo e podendo extinguir a espécies. Como exemplo podemos citar os recifes de corais.



Estimativa de parâmetros estatísticos usados para mensurar a variação em um **caráter quantitativo**: largura da concha de um caracol (filo *Mollusca*) ameaçado de extinção no Taiti (Polinésia francesa, no Pacífico Sul)



Largura média (em cm) da concha de 40 casais e filhos

	Parental means (P)	Offspring means (O)		Parental means (P)	Offspring means (O)		Parental means (P)	Offspring means (O)
1	6,8	7,3	15	7,5	7,3	28	7,8	7,5
2	6,9	7,4	16	7,6	7,7	29	7,9	7,6
3	6,9	7,6	17	7,6	7,7	30	7,9	7,7
4	7,1	7,5	18	7,6	7,9	31	7,9	7,7
5	7,3	7,3	19	7,6	7,4	32	7,9	7,7
6	7,3	7,2	20	7,6	7,5	33	7,9	7,8
7	7,3	7,4	21	7,6	7,4	34	8,0	7,7
8	7,4	7,7	22	7,7	7,6	35	8,0	7,9
9	7,5	7,6	23	7,7	7,9	36	8,0	7,8
10	7,5	7,5	24	7,8	7,5	37	8,1	7,8
11	7,5	7,7	25	7,8	7,8	38	8,1	7,8
12	7,5	7,4	26	7,8	7,9	39	8,1	7,9
13	7,5	7,8	27	7,8	7,6	40	8,5	8,1
14	7,5	7,6						

1. Calcular a média dos pais (*Parents*)
2. Calcular a média dos filhos (*Offsprings*)
3. Calcular a variância fenotípica dos pais (*Parents*)
4. Calcular a variância fenotípica dos filhos (*Offsprings*)
5. Calcular o desvio padrão dos pais (*Parents*)
6. Calcular o desvio padrão dos filhos (*Offsprings*)
7. Calcular o valor da covariância entre filhos e pais = $Cov(PO)$
8. Calcular o valor da correlação entre filhos e pais = $r(PO)$
9. Calcular o valor da regressão entre filhos e pais =
 $b = \underline{\text{estimativa da herdabilidade}}$