

Origem da variação

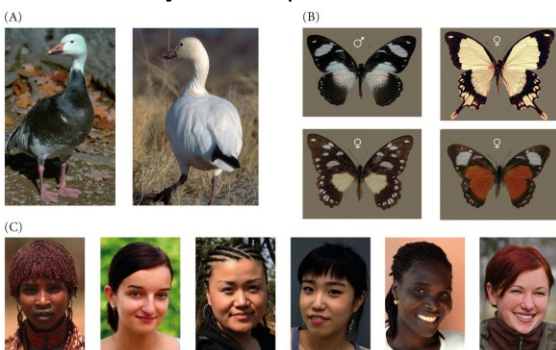
Professor Fabrício R Santos
fsantos@icb.ufmg.br
Departamento de Biologia Geral, UFMG
2012

Variação genética



Chapter 1 Opener: Evolutionary Analysis, 4/e
© 2007 Pearson Prentice Hall, Inc.

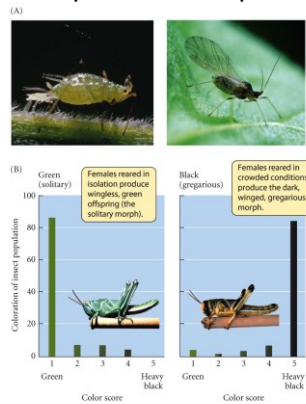
Variação fenotípica hereditária



EVOLUTION 2e, Figure 9.1

© 2008 Sinauer Associates, Inc.

Variação fenotípica causada pelo ambiente



EVOLUTION 2e, Figure 9.2

© 2008 Sinauer Associates, Inc.

Conceitos importantes

Variação genética: variantes alélicas originados por mutação e/ou recombinação

Diversidade ou variabilidade genética: medida da variabilidade existente dentro ou entre populações e/ou espécies.

Mutação: o evento na meiose (de organismos sexuais) que gera variantes alélicas.

Diversidade Genética

- Necessária para a continuidade evolutiva;
- Fundamento de toda Biodiversidade;
- A variação genética é originada por mutação e intensificada por mecanismos recombinatórios e sexuais;
- Manifestada no nível genotípico e fenotípico (minoria);
- Pode ser neutra (*dinâmica influenciada apenas pela Deriva*) ou adaptativa (*dinâmica influenciada pela Deriva e Seleção Natural*).

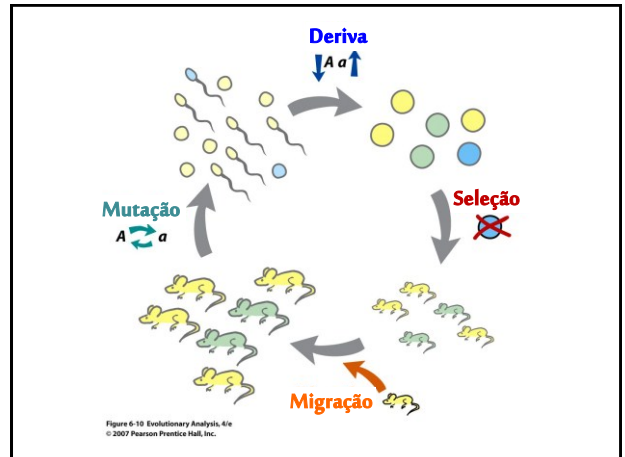
Variação genética e Evolução

Não há evolução (seleção, deriva, etc) sem variação genética pré-existente.

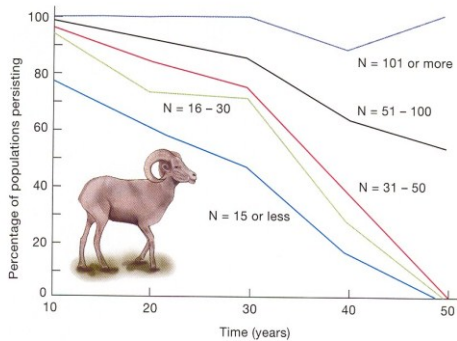
É a matéria-prima para as mudanças evolutivas e adaptativas, originada por mutação e recombinação.

É necessária para as populações serem capazes de lidar com mudanças ambientais. Assim, a perda da diversidade genética está, muitas vezes, associada com a redução do valor adaptativo populacional.

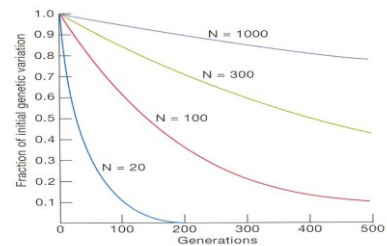
A diversidade pode ser medida a partir de variantes fenotípicas (morfologia, proteínas etc) e genotípicas (DNA).



A diversidade genética está relacionada com o tamanho populacional e o potencial evolutivo

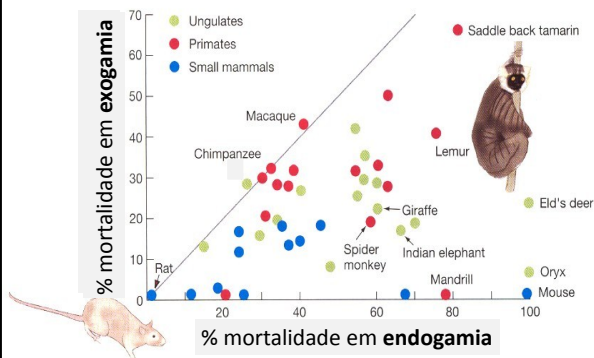


Efeitos da deriva genética na diversidade intrapopulacional

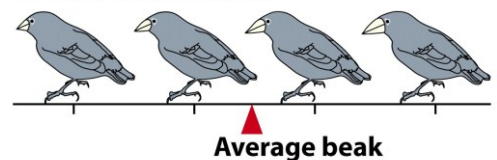


Quanto menor é o tamanho populacional (N), maior é a perda de diversidade genética por processos estocásticos (deriva e endogamia).

Depressão endogâmica



Before selection



After selection

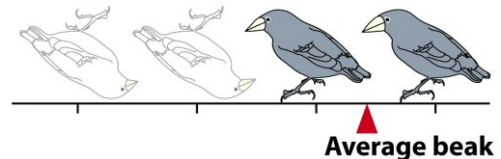
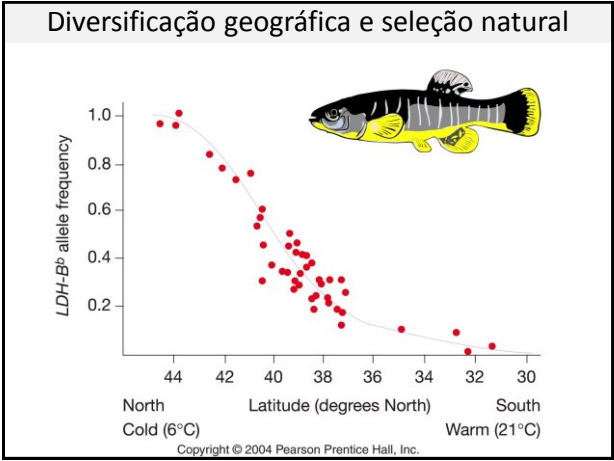
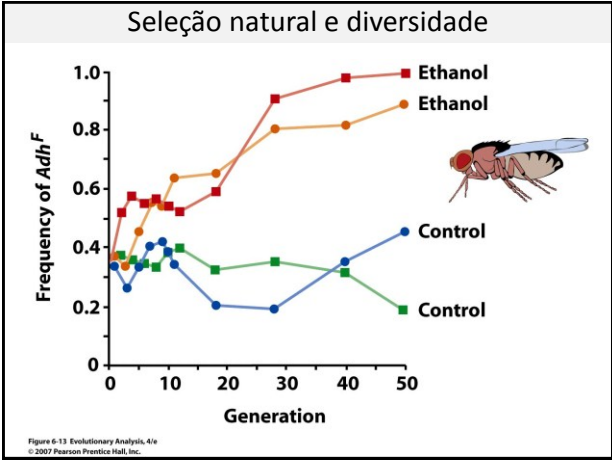
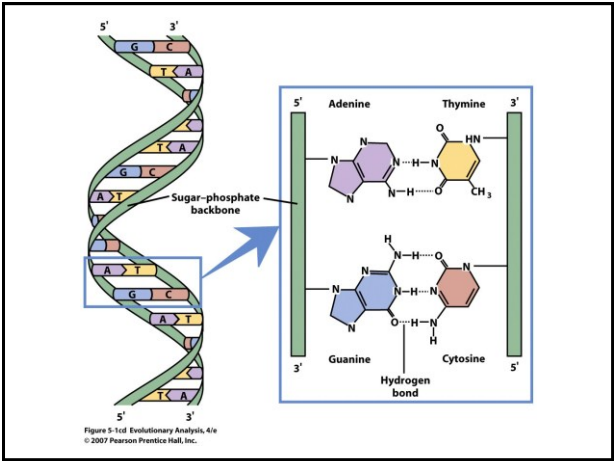
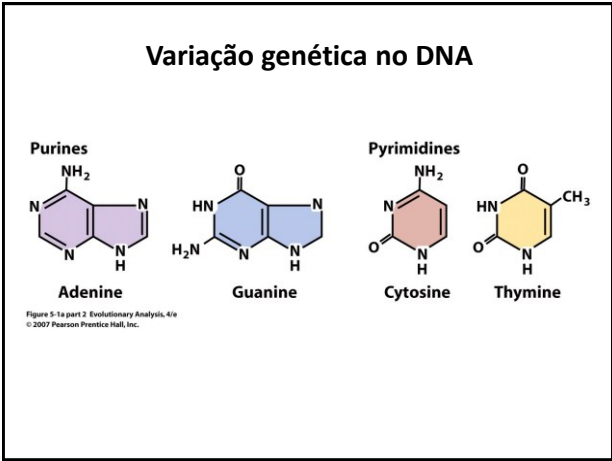
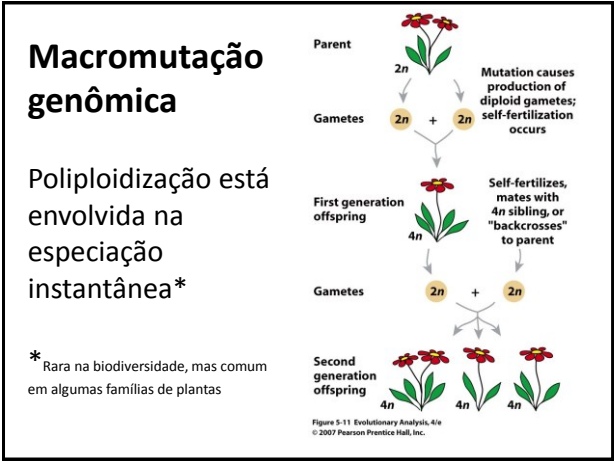


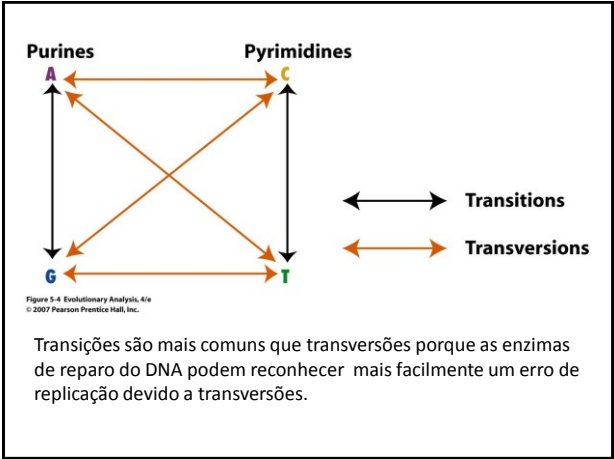
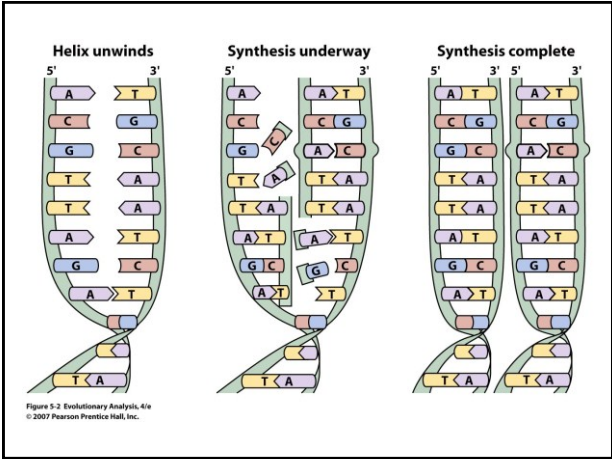
Figure 3-16: Evolutionary Analysis, 4/e
© 2007 Pearson Prentice Hall, Inc.



Origem da Variação Genética:

mutação e recombinação





Como medir a taxa de mutação no DNA?

Sequenciamento genômico de *C. elegans*

Acúmulo de mutações nas linhagens

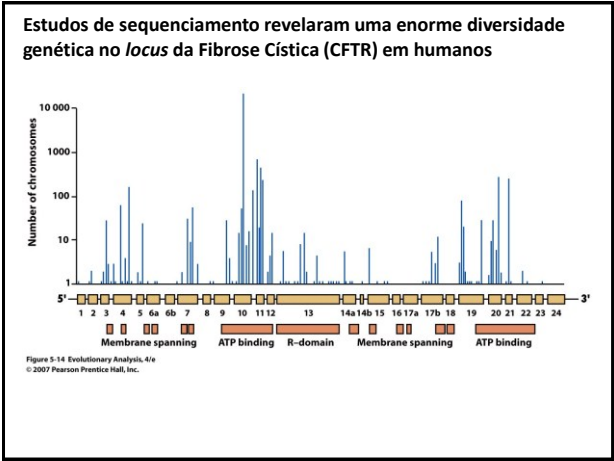
Denver et al. 2004

Geração 280: 29.561 pb em 72 linhagens
 Geração 353: 14.550 pb em 68 linhagens
 Geração 396: 18.718 pb em 58 linhagens

Detectadas: 30 mutações (17/30 indels)
 Transições >> Transversões 1,6 : 1

Taxa de mutação: $2,1 \times 10^{-8}$ /sítio/geração

400 gerações
 ≈ 1 mutação/gameta ou 2/zigoto

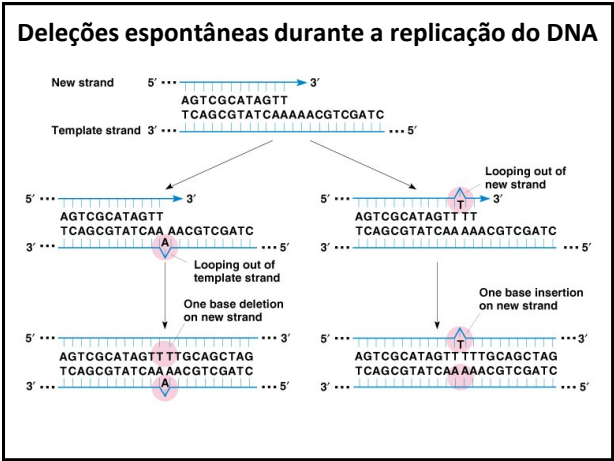


Diversidade no gene CCR5 em várias populações humanas

Population	Number of people tested	Number with each genotype			Allele frequency (%)	
		+/+	+/ $\Delta 32$	$\Delta 32/\Delta 32$	CCR5+	CCR5- $\Delta 32$
Europe						
Ashkenazi	43	26	16	1	79.1	20.9
Iceland	102	75	24	3	85.3	14.7
Britain	283	223	57	3		
Italy	91	81	10	0		
Middle East and Asia						
Saudi Arabia	241	231	10	0		
Yemen	34	34	0	0		
Russia (Udmurtia)	46	38	7	1		
Pakistan	34	32	2	0		
Hong Kong	50	50	0	0		
Mongolia	59	59	0	0		
Philippines	26	26	0	0		
Africa						
Nigeria	111	110	1	0		
Central African Republic	52	52	0	0		
Kenya	80	80	0	0		

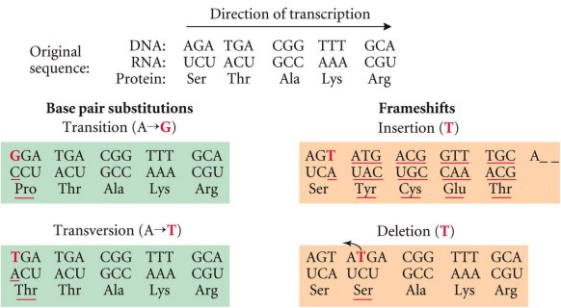
Excerpted from Martinson et al. 1997. Copyright © 1997, Nature Genetics. Reprinted by permission of the Nature Publishing Group, New York, NY

Table 5-3 Evolutionary Analysis, 4/e
 © 2007 Pearson Prentice Hall, Inc.



$\Delta 32$ – alelo que confere resistência à AIDS

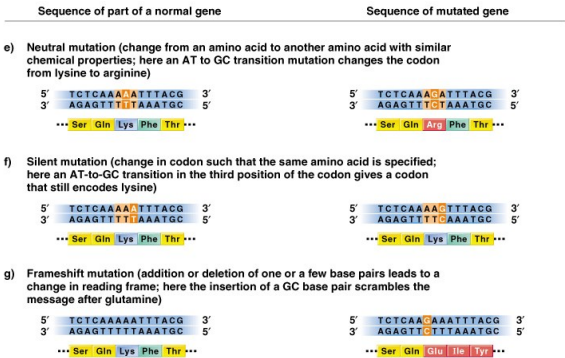
Exemplos de mutações de ponto



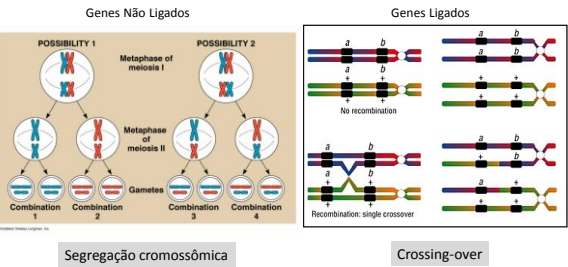
EVOLUTION 2e, Figure 8.4

© 2008 Sinauer Associates, Inc.

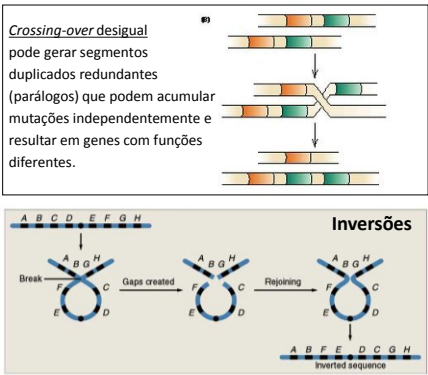
Mutações de ponto em éxons



Recombinação



Recombinação e rearranjos genômicos

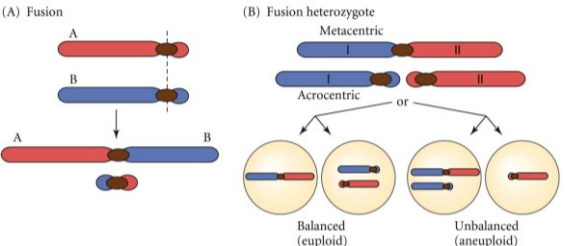


Tipos de mutações e impacto na evolução

Name	Description	Mechanism	Significance
Point mutation	Base pair substitutions in DNA sequences	Chance errors during DNA synthesis or during repair of damaged DNA	Creates new alleles
Chromosome inversion	Flipping of a chromosome segment, so order of genes along the chromosome changes	Breaks in DNA caused by radiation or other insults	Alleles inside the inversion are likely to be transmitted together, as a unit
Gene duplication	Duplication of a short stretch of DNA, creating an extra copy of the sequence	Unequal crossing-over during meiosis or retrotransposition	Redundant new genes may acquire new functions, by mutation
Genome duplication	Addition of a complete set of chromosomes	Errors in meiosis or (in plants) mitosis	May create new species; massive gene duplication

Além de outras: Translocação, transposição de elementos...

Fusão de cromossomos



EVOLUTION 2e, Figure 8.25

© 2008 Sinauer Associates, Inc.

Muntiacus reevesii (2N = 46)

$$\begin{array}{c} \text{XX} \quad \text{YY} \quad \text{nn} \quad \text{XX} \\ \text{X} \quad \text{Y} \quad \text{X} \quad \text{Y} \end{array} \quad \left. \vphantom{\begin{array}{c} \text{XX} \quad \text{YY} \quad \text{nn} \quad \text{XX} \\ \text{X} \quad \text{Y} \quad \text{X} \quad \text{Y} \end{array}} \right\}$$

EVOLUTION 2e, Figure 8.26

Figure S-6. Evolutionary Analysis, 4/e
© 2002 Pearson Benjamin Cummings

Size of gene family	Number of families		
	Yeast	Fruit flies	Roundworms
2	415	404	665
3	56	113	188
4	23	46	93
5	9	21	71
6-10	19	52	104
>10	8	38	98

Table 5-1 Evolutionary Analysis, 4/e
© 2007 Pearson Prentice Hall, Inc.

- DNA Ribossomal
- Globinas
- Genes Homeobox

Goat: ϵ^I , ϵ^{II} , $\psi\beta$, β^C , ϵ^{III} , ϵ^{IV} , $\psi\beta$, β^A , ϵ^V , ϵ^VI , $\psi\beta$, β^B

Mouse: $\psi\beta_{h0}$, β_{h1} , ψ_{h2} , ψ_{h3} , β^{maj} , β^{min}

Rabbit: ϵ , γ , $\psi\beta$, β

Chicken: ρ , β^H , β , ϵ

Embryonic genes: ϵ, γ
Adult genes: β

Diferentes genomas apresentam estruturas duplicadas de genes das globinas herdados dos ancestrais comuns

α e β globins

Cap site AUG Poly A site

Human globins

ε 54 93 122 222 850 126 118

γ 53 93 122 222 866 126 87

δ 49 93 128 222 886 126 128

β 50 93 130 222 850 126 132

α 37 93 113 204 141 129 112

Mouse globins

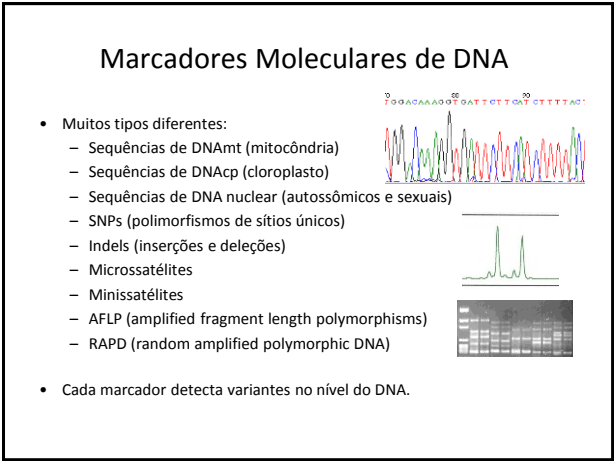
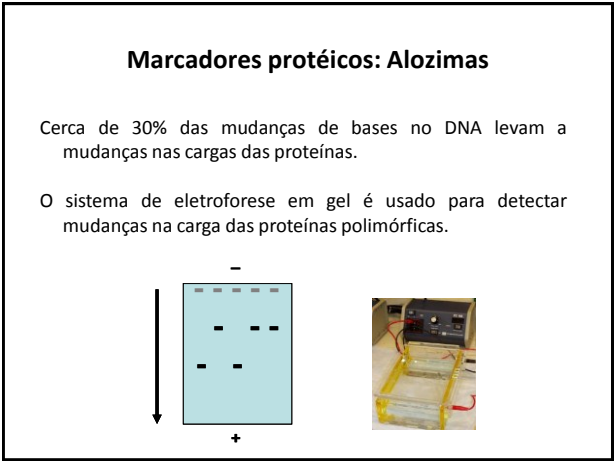
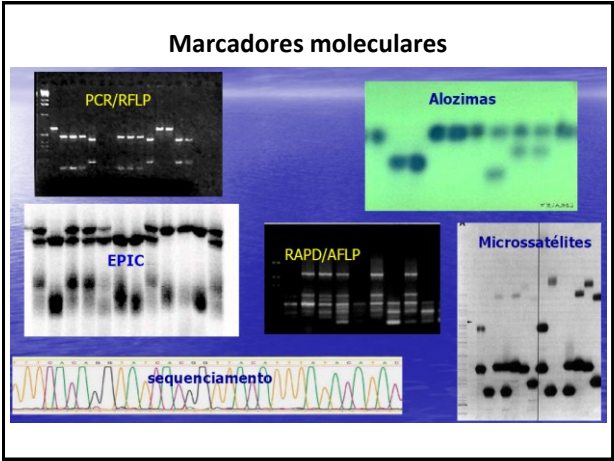
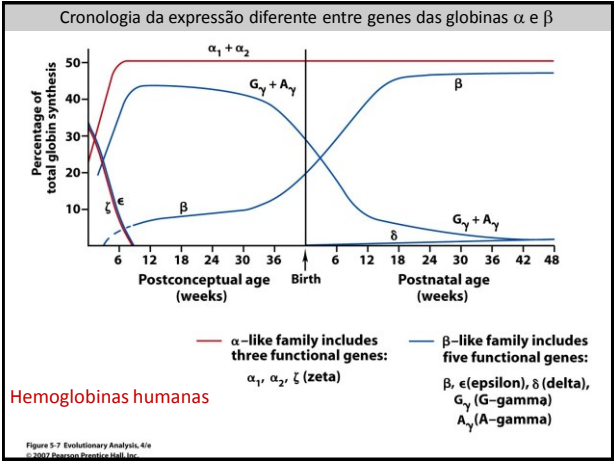
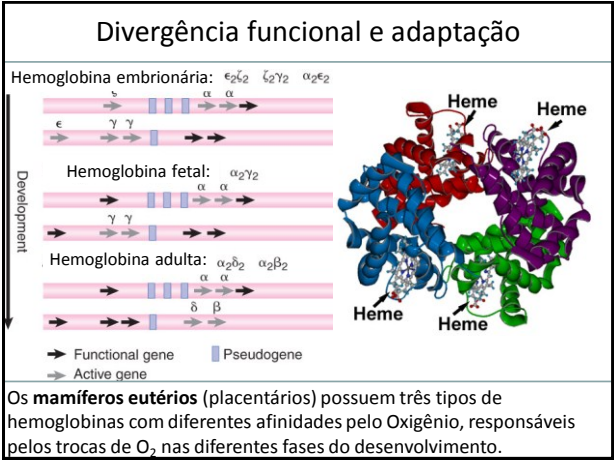
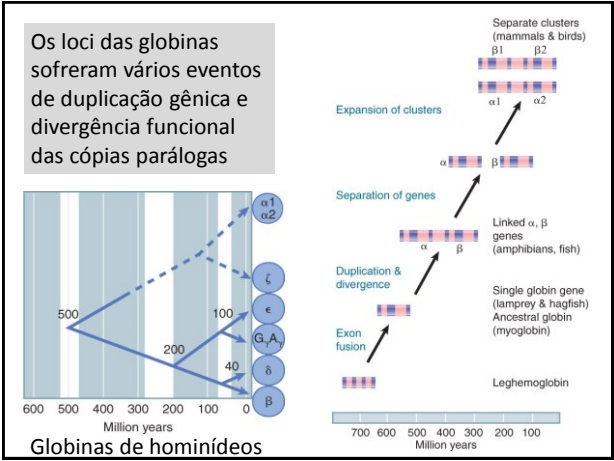
β 51 93 116 222 646 126 135

α 51 96 122 204 135 126 105

Untranslated coding sequence coding sequence Intron

Figure 5-B Evolutionary Analysis, 4b

análogos- *loci* duplicados que podem ser divergentes
 ólogos - *loci* que são homólogos e podem ser divergentes

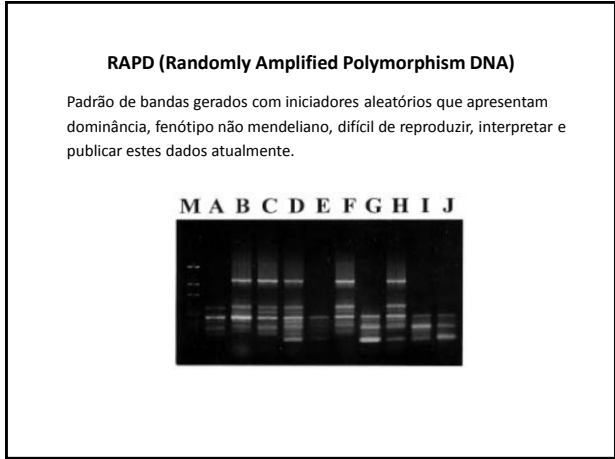
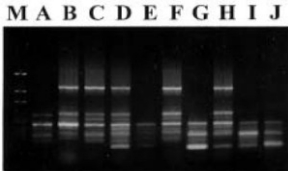


RAPD (Randomly Amplified Polymorphism DNA)

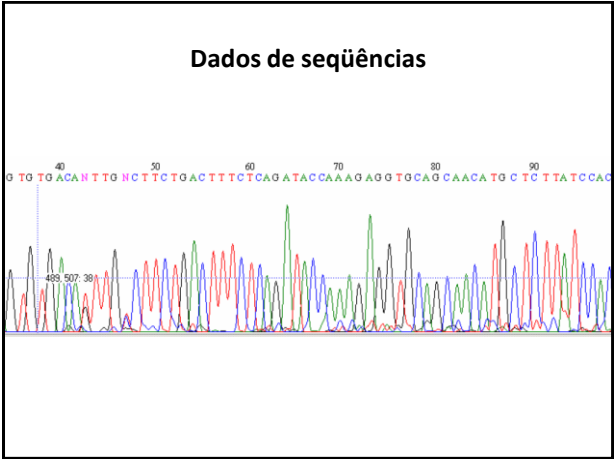
Padrão de bandas gerados com iniciadores aleatórios que apresentam dominância, fenótipo não mendeliano, difícil de reproduzir, interpretar e publicar estes dados atualmente.

RAPD (Randomly Amplified Polymorphism DNA)

Padrão de bandas gerados com iniciadores aleatórios que apresentam dominância, fenótipo não mendeliano, difícil de reproduzir, interpretar e publicar estes dados atualmente.



Dados de seqüências



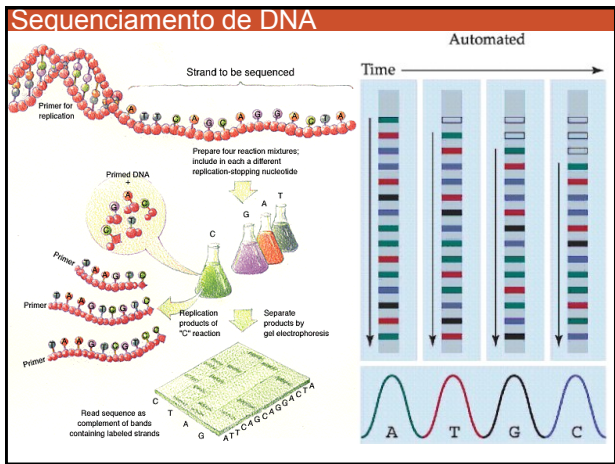
The diagram illustrates the Sanger sequencing process, comparing manual and automated methods.

Manual Sanger Sequencing:

- Strand to be sequenced:** A double-stranded DNA molecule is shown. The top strand is the template, and the bottom strand is the strand to be sequenced.
- Primer for replication:** A short DNA sequence (3'-TAC-5') is used to initiate synthesis.
- Prepare four reaction mixtures:** Four separate reaction mixtures are prepared, each containing a different replication-stopping nucleotide (A, T, G, or C).
- Primer extension:** The primer is extended using DNA polymerase and the four reaction mixtures.
- Separate products by gel electrophoresis:** The DNA fragments are separated by size using gel electrophoresis.
- Read sequence as complement of bands:** The sequence is determined by reading the bands from the gel. The sequence shown is 5'-ATT-CAG-CGG-CA-3'.

Automated Sanger Sequencing:

- Time:** The process is automated, allowing for faster sequencing.
- Fluorescently labeled nucleotides:** The reaction mixtures contain fluorescently labeled nucleotides (A, T, G, C).
- Fluorescence detection:** The DNA fragments are detected by a fluorescence detector, which identifies the terminal nucleotide of each fragment.
- Sequence determination:** The sequence is determined by reading the fluorescence signal.



DNA Mitochondrial

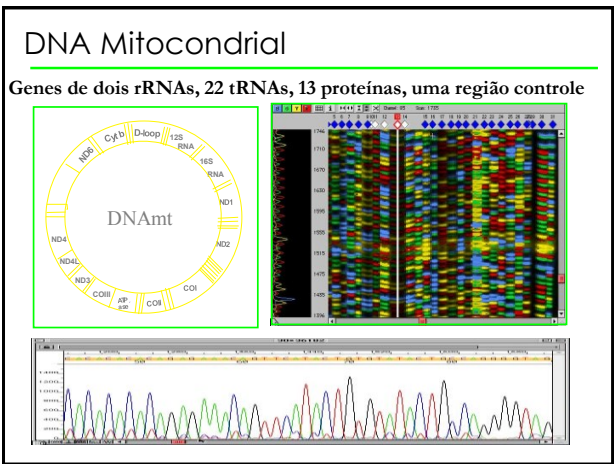
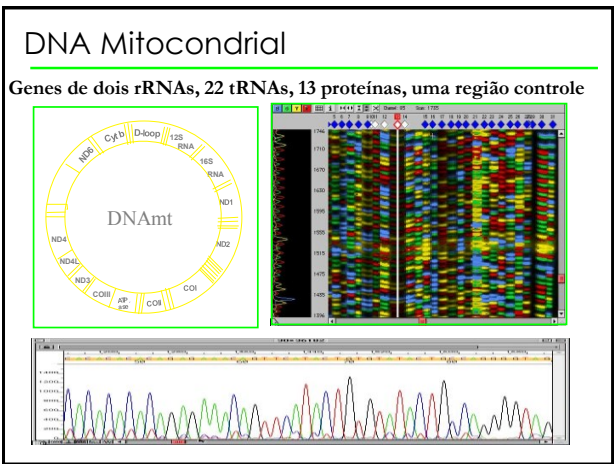
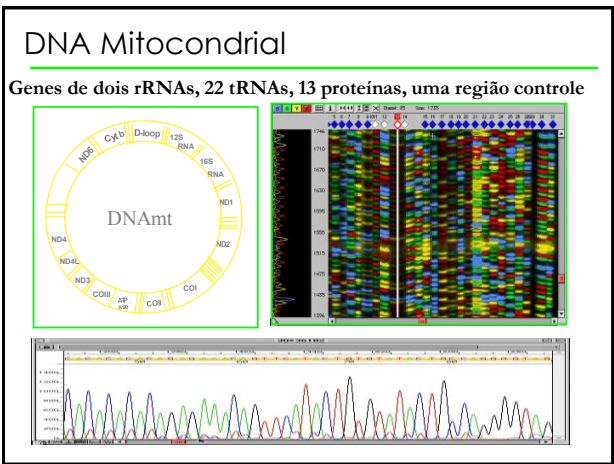
Genes de dois rRNAs, 22 tRNAs, 13 proteínas, uma região controle

The image displays the structure and analysis of mitochondrial DNA (mtDNA). On the left, a circular diagram of the mtDNA molecule is shown, labeled 'DNAmt'. The circle is divided into segments representing various genes: 12S rRNA, 16S rRNA, ND1, ND2, COI, COII, COIII, ND4, ND4L, ND5, ND6, Cyt b, and D-loop. The D-loop is highlighted as a control region. To the right, a sequence alignment plot is shown, displaying multiple sequence reads (indicated by colored bars) aligned to a reference sequence (indicated by a black line). The alignment is labeled 'Seq Name: 01' and 'Size: 17000'. Below the alignment, a chromatogram plot shows the sequence data, with peaks corresponding to the bases A, C, G, and T. The chromatogram is labeled 'Seq: 01' and 'Size: 17000'.

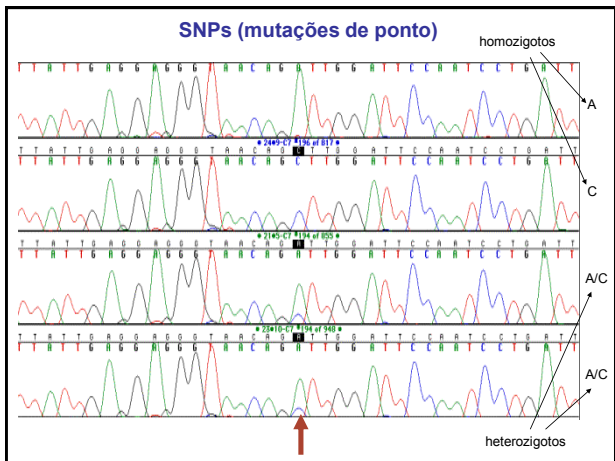
DNA Mitochondrial

Genes de dois rRNAs, 22 tRNAs, 13 proteínas, uma região controle

The image displays the structure and sequencing of mitochondrial DNA (mtDNA). On the left, a circular diagram of the mtDNA molecule is shown, labeled 'DNAmt'. The circle is divided into segments representing various genes: 12S rRNA, 16S rRNA, ND1, ND2, COI, COII, COIII, ND4, ND4L, ND5, ND6, Cyt b, and D-loop. The D-loop is highlighted as a control region. On the right, a sequencing chromatogram is shown, displaying multiple lanes of DNA sequence data. The chromatogram shows peaks corresponding to the four nucleotides (A, C, G, T) across different positions. The top of the chromatogram is labeled with '12S rRNA' and '16S rRNA'.



The diagram illustrates the detection of heterozygotes using SNP genotyping. It shows four stacked chromatograms (A, C, A/C, A/C) representing different genotypes. The top two chromatograms (A and C) show homozygotes, while the bottom two (A/C and A/C) show heterozygotes. A red arrow points to the heterozygous peak in the bottom chromatogram.



SNPs

- **SNPs** (Single Nucleotide Polymorphisms)
Polimorfismos de sítios únicos.

GTGGACGTGCTT[G/C]TCGATTACCTAG

- O tipo mais simples e comum de polimorfismo.
- Abundantes, 1 a cada 200/1000 bases no genoma humano.

- # SNPs
- **SNPs** (Single Nucleotide Polymorphisms)
Polimorfismos de sítios únicos.
- GTGGACGTGCTT[G/C]TCGATTACCTAG
- O tipo mais simples e comum de polimorfismo.
 - Abundantes, 1 a cada 200/1000 bases no genoma humano.

SNPs

- **SNPs** (Single Nucleotide Polymorphisms)
Polimorfismos de sítios únicos.

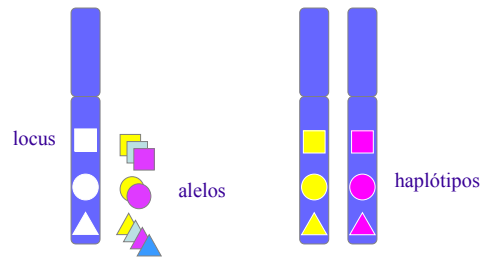
GTGGACGTGCTT[G/C]TCGATTACCTAG

- O tipo mais simples e comum de polimorfismo.
- Abundantes, 1 a cada 200/1000 bases no genoma humano.

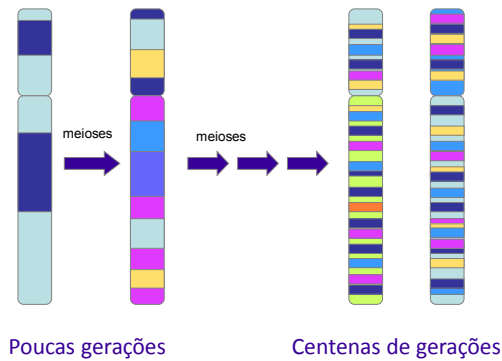
- # SNPs
- **SNPs** (Single Nucleotide Polymorphisms)
Polimorfismos de sítios únicos.
- GTGGACGTGCTT[G/C]TCGATTACCTAG
- O tipo mais simples e comum de polimorfismo.
 - Abundantes, 1 a cada 200/1000 bases no genoma humano.

Haplótipo

- Múltiplos *loci* no mesmo cromossomo cujos alelos são herdados juntos, pela ausência de recombinação;
- Geralmente é um conjunto de SNPs ligados.

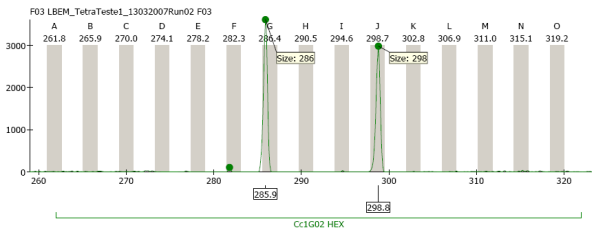


Formação de blocos haplotípicos

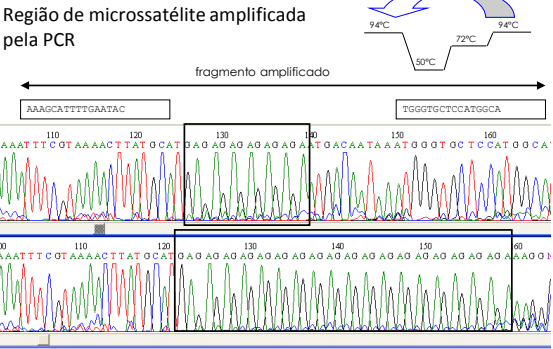


Microssatélites

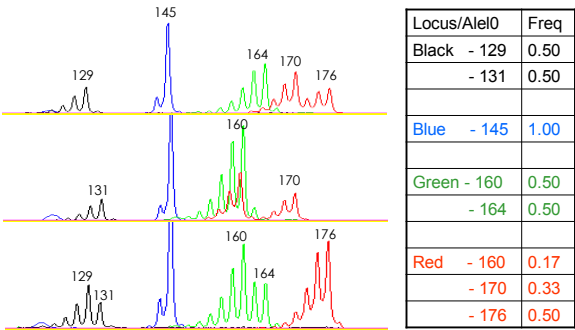
- Características:**
- Codominantes
 - Alta taxa de mutação (indicado para estudos populacionais)



Variação em Microssatélites



Cromatograma: dados de Microssatélites



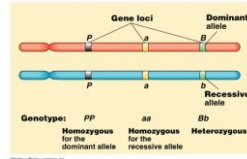
Medidas de Diversidade Genética

Populações naturais podem ter sua diversidade genética medida de diferentes formas a partir dos marcadores moleculares:

- Número de alelos ou haplótipos,
- Número de sítios ou *loci* polimórficos ou segregantes,
- Nível de polimorfismo,
- Heterozigidade,
- Fst ou equivalentes (Gst, Rst, fst),
- Variância,
- Diversidade haplotípica,
- Diversidade nucleotídica,
- Valores q etc.

Alelos (ou variantes)

- Formas alternativas de um gene ou sequência qualquer do DNA em uma posição cromossômica específica (*locus*)
- Em cada *locus* diplóide um indivíduo possui dois alelos (de cada parental), e isto constitui um genótipo.
- À combinação de alelos ligados em uma região sem recombinação, chamamos haplótipo.



Locus	1	2	3
Alelo M	A	G	G
Alelo P	T	C	G

Os níveis hierárquicos da variabilidade

- Indivíduo:** quantos alelos possui e quantos *loci* são polimórficos?
- Família:** quem contribuiu quais alelos para os genótipos do filho na geração seguinte? (estudos de parentesco)
- População:** quais as frequências alélicas dos *loci* polimórficos e o número médio e efetivo de alelos por *locus*?
- Espécie:** como a variabilidade genética total da espécie está repartida entre suas populações?

Número de alelos

Média do número de alelos (A) ou número efetivo (A_e) por *locus* é usada para caracterizar a extensão da diversidade genética em uma população.

Número médio de alelos (A)

$$A = \frac{\sum \text{alelos}}{NL} \quad \leftarrow \text{Número de loci}$$

Número efetivo de alelos (A_e):

$$A_e = \frac{1}{\sum x_i^2} \quad \leftarrow \text{Frequência de cada alelo}$$

Locus de microsatélite em Chimpanzés

Locus com 5 alelos com frequências:

$$A = 0,026$$

$$B = 0,316$$

$$C = 0,026$$

$$D = 0,079$$

$$E = 0,553$$



$$A_e = \frac{1}{(0,026^2 + 0,316^2 + 0,026^2 + 0,079^2 + 0,553^2)} = 2,42$$

Diversidade gênica ou alélica

É a variedade de alelos e genótipos presentes em populações, espécies ou grupos de espécies

- Pode ser medida de várias maneiras:
- Quanto heterozigotos a população tem (H_o)/deveria ter (H_e), em função das frequências gênicas? (heterozigosidade observada e esperada / diversidade haplotípica)
- Se levarmos em conta também a diferença molecular entre os haplótipos chamamos de diversidade nucleotídica (p).

Medidas de heterozigosidade ou diversidade gênica

A composição genética de uma população é representada em termos de frequências alélicas, número de alelos e heterozigosidades.

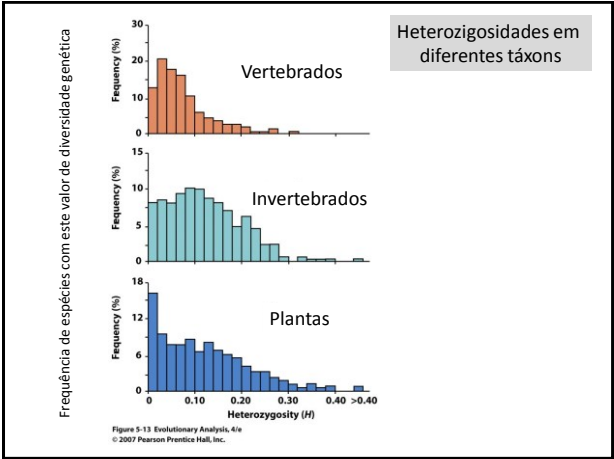
Heterozigosidade observada x esperada:

A diversidade genética em um *locus* é caracterizada pela heterozigosidade observada H_o , esperada H_e e a diversidade alélica

$$H_o = \text{freq. heterozigotos}$$

$$H_e = 2pq \text{ ou } 1 - \sum p_i^2$$

- A medida mais comum para a variabilidade de frequência é a **heterozigosidade** (H), ou diversidade haplotípica (h) em casos de regiões ligadas ou haplóides.
- A variabilidade de seqüências pode ser medida pelo número de sítios polimórficos (NP) e pela distância molecular entre seqüência (p).
- A diversidade nucleotídica (p) entre seqüências incorpora além da frequência de cada seqüência também a diferença molecular entre elas.



Diversidade genética em lobos da Etiópia

	A	H_o	H_e	A_e	N amostral
Lobo da Etiópia					
Sanetti Plateu	2,0	0,150	0,138	0,121	16
Web Valley	2,8	0,313	0,271	1,37	23
Cão	6,4	0,516	0,679	3,11	35
Lobo cinza	4,5		0,620	2,63	18
Coioote	5,9		0,675	3,08	17

Ho – heterozigosidade observada
He – heterozigosidade esperada
Ae – Alelos efetivos

Resultados:

- A , H_o , H_e e A_e são menores no lobo da Etiópia do que nas espécies não-ameaçadas, portanto tem menor potencial evolutivo.

- A H_e da população de Sanetti foi menor que a de Web => hibridização com cães em Web (presença de alelos de cães).