

Antonio Solé-Cava  
Edson Pereira da Silva  
Gisele Lôbo-Hajdu

Evolução







Fundação

**CECIERJ**

Consórcio **cederj**

Centro de Educação Superior a Distância do Estado do Rio de Janeiro

## Evolução

Volume 2 – Módulo 1

Antonio Solé-Cava

Edson Pereira da Silva

Gisele Lôbo-Hajdu



SECRETARIA DE  
CIÊNCIA E TECNOLOGIA

Ministério  
da Educação



Apoio:



Fundação Carlos Chagas Filho de Amparo  
à Pesquisa do Estado do Rio de Janeiro

# Fundação Cecierj / Consórcio Cederj

Rua Visconde de Niterói, 1364 – Mangueira – Rio de Janeiro, RJ – CEP 20943-001

Tel.: (21) 2334-1569 Fax: (21) 2568-0725

## Presidente

Masako Oya Masuda

## Vice-presidente

Mirian Crapez

## Coordenação do Curso de Biologia

UENF - Milton Kanashiro

UFRJ - Ricardo Iglesias Rios

UERJ - Celly Saba

## Material Didático

### ELABORAÇÃO DE CONTEÚDO

Antonio Solé-Cava

Edson Pereira da Silva

Gisele Lôbo-Hajdu

### COORDENAÇÃO DE DESENVOLVIMENTO INSTRUCIONAL

Cristine Costa Barreto

### DESENVOLVIMENTO INSTRUCIONAL E REVISÃO

José Meyohas

Maria Helena Hatschbach

### COORDENAÇÃO DE LINGUAGEM

Maria Angélica Alves

### REVISÃO TÉCNICA

Marta Abdala

## Departamento de Produção

### EDITORA

Tereza Queiroz

### COORDENAÇÃO EDITORIAL

Jane Castellani

### COPIDESQUE

Nilce Rangel Del Rio

Cristina Freixinho

### REVISÃO TIPOGRÁFICA

Kátia Ferreira dos Santos

Patrícia Paula

### COORDENAÇÃO DE PRODUÇÃO

Jorge Moura

### PROGRAMAÇÃO VISUAL

Fábio Guimarães

### ILUSTRAÇÃO

Fabiana Rocha

### CAPA

Fabiana Rocha

### PRODUÇÃO GRÁFICA

Oséias Ferraz

Patricia Seabra

Copyright © 2004, Fundação Cecierj / Consórcio Cederj

Nenhuma parte deste material poderá ser reproduzida, transmitida e gravada, por qualquer meio eletrônico, mecânico, por fotocópia e outros, sem a prévia autorização, por escrito, da Fundação.

S685e

Solé-Cava, Antonio.

Evolução v. 2 / Antonio Solé-Cava. – Rio de Janeiro:

Fundação CECIERJ, 2010.

194p.; 19 x 26,5 cm.

ISBN: 85-7648-099-9

1. Seleção natural. 2. Adaptacionismo. 3. Endocruzamentos. I. Silva, Edson Pereira da. II. Lôbo-Hajdu, Gisele. III. Título.

CDD: 576.8

2010/1

Referências Bibliográficas e catalogação na fonte, de acordo com as normas da ABNT.

# Governo do Estado do Rio de Janeiro

**Governador**  
Sérgio Cabral Filho

**Secretário de Estado de Ciência e Tecnologia**  
Alexandre Cardoso

## Universidades Consorciadas

**UENF - UNIVERSIDADE ESTADUAL DO  
NORTE FLUMINENSE DARCY RIBEIRO**  
Reitor: Almy Junior Cordeiro de Carvalho

**UERJ - UNIVERSIDADE DO ESTADO DO  
RIO DE JANEIRO**  
Reitor: Ricardo Vieiralves

**UFF - UNIVERSIDADE FEDERAL FLUMINENSE**  
Reitor: Roberto de Souza Salles

**UFRJ - UNIVERSIDADE FEDERAL DO  
RIO DE JANEIRO**  
Reitor: Aloísio Teixeira

**UFRRJ - UNIVERSIDADE FEDERAL RURAL  
DO RIO DE JANEIRO**  
Reitor: Ricardo Motta Miranda

**UNIRIO - UNIVERSIDADE FEDERAL DO ESTADO  
DO RIO DE JANEIRO**  
Reitora: Malvina Tania Tuttman



**SUMÁRIO**

<b>Aula 11</b> – Deriva gênica _____	<b>7</b>
<i>Antonio Solé-Cava</i>	
<b>Aula 12</b> – Exercício prático. Populus e deriva gênica _____	<b>27</b>
<i>Antonio Solé-Cava</i>	
<b>Aula 13</b> – Seleção Natural 1: conceito e modelos _____	<b>45</b>
<i>Edson Pereira da Silva</i>	
<b>Aula 14</b> – Seleção Natural 2: conceito e modelos _____	<b>65</b>
<i>Edson Pereira da Silva</i>	
<b>Aula 15</b> – Exercício prático: seleção natural _____	<b>81</b>
<i>Antonio Solé-Cava</i>	
<b>Aula 16</b> – Controvérsias evolutivas _____	<b>107</b>
<i>Edson Pereira da Silva</i>	
<b>Aula 17</b> – Exercício prático: interação seleção/deriva gênica _____	<b>123</b>
<i>Antonio Solé-Cava</i>	
<b>Aula 18</b> – Adaptação e adaptacionismo _____	<b>135</b>
<i>Edson Pereira da Silva</i>	
<b>Aula 19</b> – Endocruzamento: estimativas diretas e indiretas _____	<b>149</b>
<i>Gisele Lôbo-Hajdu</i>	
<b>Aula 20</b> – Simulação: migração e estruturação populacional _____	<b>173</b>
<i>Antonio Solé-Cava</i>	
<b>Referências</b> _____	<b>191</b>





## Deriva gênica

AULA

# 11

### Meta da aula

Estabelecer os principais fatores responsáveis pelas variações aleatórias das frequências gênicas nas populações naturais.

Ao final da aula, você deverá ser capaz de:

- Calcular o desvio padrão da frequência de um alelo.
- Discriminar distribuição normal e distribuição binomial.
- Relacionar o desvio padrão com a frequência gênica e com o tamanho populacional.

## DA CERVEJA AOS FEIJÕES, DOS FEIJÕES AOS GENES

Bom, vamos agora dar o último passo na cadeia de pensamentos que começou com o exemplo dos bêbados e das cervejas e passou para os feijões pretos e carioquinhas (os nossos alelos  $P$  e  $C$ ). Vimos na aula passada que, em Ciência, os modelos podem ser deterministas ou estocásticos. Com o exemplo dos bêbados, que nos modelos estocásticos o acaso é levado em conta a cada passo, enquanto que nos modelos deterministas o comportamento das variáveis é previsível a partir de regras gerais e constantes que, ao se repetirem os experimentos (Figura 10.7 da Aula 10), o modelo estocástico, na média, acaba se aproximando do modelo determinista. No exemplo dos feijões, percebemos que a variação das frequências de cada tipo dependia de quantos feijões eram amostrados de cada vez (Figuras 10.8 e 10.9 da Aula 10). Quando a variação era maior?



A variação era maior quando o número de feijões amostrados era menor. Isso é o esperado, mesmo pelo senso comum. Uma amostra pequena deve ser menos representativa do total dos feijões no saco do que uma amostra grande. Mas podemos ver isso de uma maneira mais formal. A oscilação dos valores de uma variável é chamada variância. A variância é calculada de forma diferente, de acordo com o tipo de distribuição da variável. Qual será o tipo de distribuição da frequência dos genes? No caso dos feijões, como só existem duas alternativas para nossa variável (o feijão OU é preto OU não é), temos uma **DISTRIBUIÇÃO BINOMIAL**.

### DISTRIBUIÇÃO BINOMIAL

Como você viu em Estatística, toda variável tem um tipo de distribuição. Por exemplo, a distribuição das alturas dos jovens do Estado do Rio de Janeiro, de 10 a 30 anos de idade, deve obedecer a uma distribuição normal, com uma média definida e uma distribuição de alturas em torno da média, formando uma curva em forma de sino. Na distribuição normal, as variáveis são contínuas, ou seja, podem assumir valores intermediários que dependem apenas da precisão do aparelho usado para medi-las. Assim, uma pessoa pode medir 1,6 metros; mas poderia ter, usando-se uma régua mais precisa, 1,62 metros, ou 1,621 metros, ou 1,6214 metros etc. Na distribuição binomial, só existem dois estados para cada variável, e as variáveis são discretas, ou seja, não têm valores intermediários. Os valores das variáveis na distribuição binomial também não são medidos. As variáveis são proporções. Assim, na distribuição normal, a variável é medida; na binomial, ela é contada. Na aula passada, vimos um exemplo disso: os feijões, que podiam ser carioquinhas ou não. As variáveis da distribuição binomial costumam ser descritas como tendo dois estados: 'p' e 'q'. Assim, um feijão carioquinha poderia ser chamado de 'p'; e não ser carioquinha poderia ser chamado de 'q'. Uma propriedade importante da distribuição binomial é que, como os valores de p e q são proporções, a soma dos dois deve sempre dar 1. Ou seja:  $p + q = 1$ . Os alelos dos genes têm distribuição binomial: o que importa são as proporções dos mesmos nas populações. Assim, quando falamos que a frequência do alelo  $Rh^+$  em uma população humana é 0,30, podemos imediatamente inferir que a proporção do alelo  $Rh^-$  será 0,70. São exemplos de distribuição normal as medidas populacionais de comprimento, altura, peso, porcentagem de atividade enzimática etc. São exemplos de distribuição binomial as frequências dos alelos, as proporções de torcedores dos times de futebol, as proporções das diferentes nacionalidades visitando o Rio de Janeiro no Carnaval etc.

Então, as frequências gênicas nas populações têm distribuição binomial. E qual é a variância da distribuição binomial? Ela é calculada para cada um dos dois estados da variável. Assim,

$$V(p) = pq / n$$

em que  $p$  e  $q$  são os dois estados da binomial, e  $n$  é o tamanho amostral. No caso de alelos de um organismo diplóide,  $n = 2N$ , em que  $N$  é o número de indivíduos amostrados. Por quê?



Porque, por ser diplóide, haverá dois alelos para cada loco. Assim, se amostramos  $N=10$  animais para estudar um grupo sanguíneo como o Rh, por exemplo, estaremos amostrando  $n=20$  alelos no total. Uma boa medida de quanto as frequências de  $p$  e  $q$  irão variar após uma amostragem é o desvio padrão (DP), que é calculado como

$$DP(p) = \sqrt{\frac{pq}{n}}$$

em que  $n$  é o número de alelos da população. Ou seja, o desvio padrão é a raiz quadrada da variância.

Então, vamos fazer um exercício. Em um estudo recente com genes responsáveis pela produção da enzima manose-fosfato isomerase, foram estimadas as frequências gênicas em uma amostra de 44 camarões do Rio de Janeiro (GUSMÃO, LAZOSKI E SOLÉ-CAVA, 2000). A partir dos dados apresentados na Tabela 11.1, calcule a variância e o desvio padrão de cada alelo.

**Tabela 11.1:** Frequências gênicas do loco da manose-fosfato isomerase, na espécie *Farfantepenaeus brasiliensis* do Rio de Janeiro.  $N = 44$

Alelo	Frequência	Variância	Desvio padrão
A	0,15		
B	0,76		
C	0,07		
D	0,02		



Para calcular as variâncias, você deve usar a fórmula  $V(p) = pq / n$ , lembrando que  $n = 2N$ ; ou seja, como foram amostrados 44 camarões,  $n = 88$ . Você faz esse cálculo para cada alelo, chamando-o de 'p', e chamando q o resto dos alelos, isto é,  $q = 1-p$ . Assim, por exemplo, a variância do alelo A será

$$V(A) = (0,15 \times 0,85) / 88 \text{ (já que } q = 1-0,15), \text{ ou}$$

$$V(A) = 0,1275 / 88 \text{ ou}$$

$$V(A) = 0,0014$$

Da mesma forma, o desvio padrão de A será

$$DP(A) = \text{Raiz quadrada de } V(A) \text{ ou}$$

$$DP(A) = \sqrt{0,0014} = 0,0381$$

Então, preenchendo a **Tabela 11.1** temos a **Tabela 11.2**. Verifique se você acertou as respostas!

**Tabela 11.2:** Freqüências gênicas do loco da manose-fosfato isomerase, na espécie *Farfantepenaeus brasiliensis* do Rio de Janeiro. N = 44

Alelo	Freqüência	Variância	Desvio padrão
A	0,15	0,0014	0,0381
B	0,76	0,0021	0,0455
C	0,07	0,0007	0,0272
D	0,02	0,0002	0,0149

E se esses valores de freqüência gênica tivessem sido obtidos com uma amostra muito menor (digamos, 10 camarões), como teriam sido os desvios padrões? Repita os cálculos da **Tabela 11.2** na **Tabela 11.3**, agora colocando apenas o desvio padrão, com N = 10 (não se esqueça,  $n = 2N$ ).



**Tabela 11.3:** Freqüências gênicas do loco da manose-fosfato isomerase, na espécie *Farfantepenaeus brasiliensis* do Rio de Janeiro. N = 10 (são fornecidos também os dados de desvio padrão com N = 44, extraídos da **Tabela 11.2**, para comparação)

Alelo	Freqüência	Desvio padrão (N = 10)	Desvio padrão (N = 44)
A	0,15		0,0381
B	0,76		0,0455
C	0,07		0,0272
D	0,02		0,0149

Você viu como os desvios padrões são maiores com  $N=10$ ? (Tabela 11.4). Por que isso acontece? Lembra da fórmula da variância?

$$V(p) = pq / n$$

Ou seja, como  $n$  está no denominador, quanto maior ele for, menor será a variância. Em outras palavras, **quanto maior for a população, menos suas frequências gênicas irão variar.**

**Tabela 11.4.** Frequências gênicas do loco da manose-fosfato isomerase, na espécie *Farfantepenaeus brasiliensis* do Rio de Janeiro.  $N=10$  (são fornecidos também os dados de desvio padrão com  $N=44$ , extraídos da Tabela 11.2, para comparação)

Alelo	Frequência	Desvio padrão ( $N=10$ )	Desvio padrão ( $N=44$ )
A	0,15	0,0798	0,0381
B	0,76	0,0955	0,0455
C	0,07	0,0571	0,0272
D	0,02	0,0313	0,0149

Dê mais uma olhada na Tabela 11.4 e responda: além do tamanho populacional, o que mais influencia o desvio padrão das frequências gênicas?

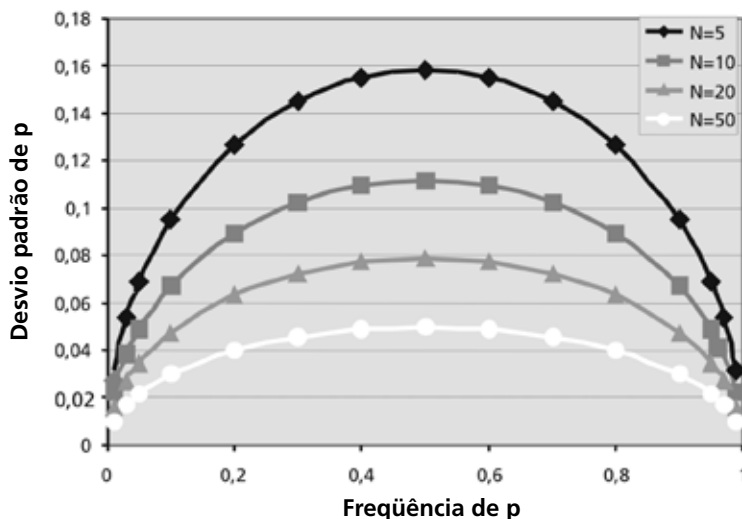


O que influencia é a própria frequência. À primeira vista, parece que, quanto maior a frequência, maior será o desvio padrão. No entanto, não é isso o que ocorre! Pois você viu que o desvio padrão quando  $A = 0,15$  e  $N = 44$  é 0,0381. Quanto será o desvio padrão quando  $A = 0,85$ , com o mesmo  $N$ ?



Será o mesmo! Afinal, o cálculo é feito como raiz quadrada de  $pq/N$ , em que  $q = 1-p$ , não é? Pois, então: quando  $A = 0,15$ ,  $pq = 0,15 \times (1-0,15) = 0,15 \times 0,85$ . Quando  $A = 0,85$ ,  $pq = 0,85 \times (1-0,85) = 0,85 \times 0,15$ . Como a ordem dos fatores não altera o produto, o  $pq$  será o mesmo quando  $A = 0,15$  e quando  $A = 0,85$ . Ou seja, o desvio padrão será o mesmo. Então, em que valor de  $p$  será maior o desvio padrão? Veja a Figura 11.1; nela registramos, em um gráfico, os valores do desvio padrão em função do valor de  $p$ . Também fizemos quatro linhas; uma para cada tamanho populacional. Nessa figura fica claro que o maior

desvio padrão ocorre quando  $p=0,50$ . Também fica claro, na **Figura 11.1**, que o desvio padrão vai ser maior com os menores tamanhos populacionais.



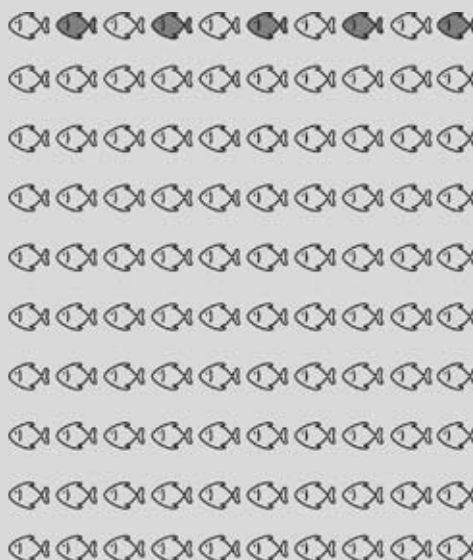
**Figura 11.1:** Desvio padrão da frequência gênica em função do tamanho populacional.

Pois bem, o que isso tudo tem a ver com a deriva gênica? Como vimos, os genes apresentam uma distribuição binomial, e o desvio padrão e a variância das frequências gênicas são dependentes do valor das frequências e do tamanho populacional. Como cada geração nova é produzida como uma amostragem da população anterior, a cada geração vai haver uma variação nas frequências gênicas, puramente ao acaso, ligada exclusivamente aos desvios padrões das frequências. Isso acontece porque, quando uma geração se reproduz, existe uma amostragem. Essa amostragem é como um sorteio, em que os alelos sorteados encontram-se nos gametas que se unem para formar os filhotes. Numa população infinita, todos os gametas vão se encontrar com outros, produzindo um número infinito de filhotes. Numa população infinita, o desvio padrão das frequências tende a zero (pois  $pq/n$  tende a zero quando  $n$  tende a infinito). Mas as populações não são infinitas, de modo que as gerações são produzidas como amostragens limitadas das gerações anteriores. Os genes, dessa forma, vão variar nas populações ao longo das gerações, por causa do desvio padrão das frequências. Mas você não precisa acreditar em mim! Basta fazer o exercício a seguir para ver como isso funciona...



### ATIVIDADE

1. Vamos simular a evolução, por deriva gênica, de dois alelos (por exemplo, os alelos 'p' para peixes pintados e 'q' para peixes não pintados) em uma população pequena. Observe a **Figura 11.2**. Nela, vemos 10 fileiras com 10 peixes cada. Vamos estudar um gene mitocondrial, para facilitar as coisas, pois, como você viu no curso de Genética, os genes mitocondriais da maioria dos animais se comportam de maneira haplóide e são transmitidos apenas pelas mães. Assim, cada peixe terá apenas um alelo (em vez de dois, se fosse um gene nuclear, o que poderia complicar um pouco o desenho). Vamos considerar também apenas as fêmeas, já que os machos, como dissemos, em geral não transmitem suas mitocôndrias aos filhotes. Vamos dizer que, por alguma condição ambiental severa, apenas 10 peixes consigam sobreviver na nossa população, a cada geração. Cada uma de nossas fêmeas, na época da reprodução, pode deixar zero, um ou dois filhotes fêmeas. Assim, uma fêmea que deixe zero filhote não passará seu alelo para a próxima geração; uma fêmea que deixe um filhote passará seu alelo sem contribuir para aumentar sua frequência, e uma fêmea que deixe dois filhotes contribuirá para aumentar a frequência do alelo naquela geração. Vamos dizer que a chance de ter 0, 1 ou 2 filhotes seja a mesma. Como estamos fazendo um modelo estocástico, temos de definir a reprodução de cada fêmea a cada geração, assim como tínhamos definido a caminhada de nossos bêbados, na aula passada. Nesse caso, como temos três possibilidades igualmente prováveis, podemos deixar de lado o gerador de números aleatórios do computador, e vamos usar um outro também de números aleatórios, porém mais simples: um dado de seis lados. Vamos dizer que, se ao jogar um dado, der 1 ou 2, a fêmea não deixou filhotes; se der 3 ou 4, ela deixou um filhote, e se der 5 ou 6, ela deixou dois filhotes. Se por acaso as 10 vagas de filhotes da geração seguinte já tiverem sido preenchidas antes de completar a reprodução das 10 fêmeas da geração anterior, considere que as fêmeas que restavam não conseguiram reproduzir-se. Se as 10 vagas de filhotes não tiverem sido preenchidas, considere que, naquela geração, não foram produzidos peixes suficientes, e coloque um 'X' sobre os peixes que não foram gerados naquela geração.



**Figura 11.2:** Peixes à deriva! Um exercício em deriva gênica em um gene mitocondrial. Peixe pintado = alelo p; peixe sem pintar = alelo q.



Então vamos começar: você vai notar na **Figura 11.2** que as freqüências dos alelos  $p$  (pintado) e  $q$  (não pintado) começam iguais (ou seja, metade dos peixes começam pintados). Assim, comece jogando o dado e vendo quantos filhotes a primeira fêmea deixou. Se ela não deixou filhotes (ou seja, se você tirou 1 ou 2 no dado), coloque um "X" nela. Se ela deixou um filhote, coloque uma seta ligando-a a um dos filhotes vagos na geração seguinte e, se ela tiver o alelo ' $p$ ', pinte o filhote também. Se ela deixou dois filhotes, faça a mesma coisa, colocando duas setas que partam dela para os seus dois filhotes. Repita isso até que todas as fêmeas tenham se reproduzido (ou até que tenham acabado as 10 vagas de filhotes daquela geração).

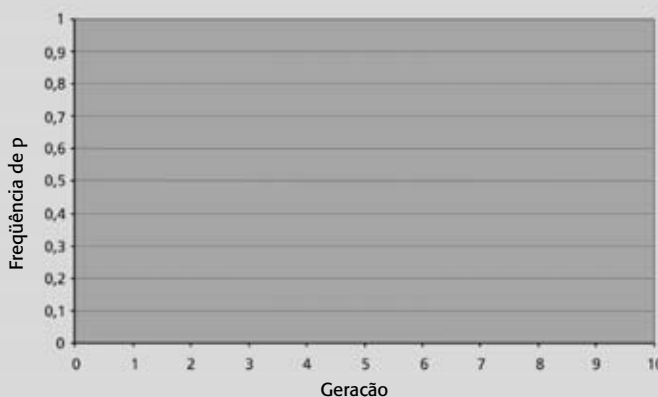
Quando você tiver acabado essa geração, recomece a nova geração, seguindo os mesmos passos anteriores. Repita isso até completar as 10 gerações de peixes.

Agora, preencha a **Tabela 11.5** com seus resultados, e use-os para desenhar uma linha na **Figura 11.3**, mostrando a evolução do alelo ' $p$ ' (peixes pintados) ao longo das 10 gerações. Não se esqueça de que a freqüência de  $p$  em cada geração é calculada como o número de peixes pintados dividido pelo total de peixes vivos naquela geração.

**Tabela 11.5:** Variação da freqüência do alelo ' $p$ ' nos peixes de nosso exercício

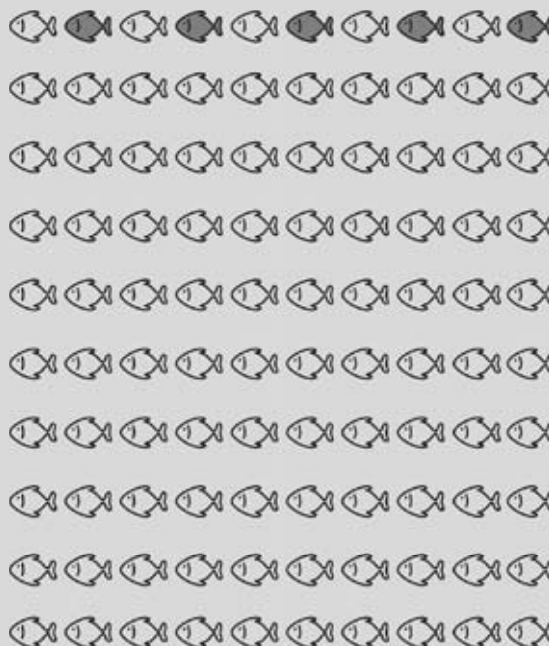
Geração	Peixes vivos	Peixes pintados	Freqüência de $p$
1	10	5	0,50
2			
3			
4			
5			
6			
7			
8			
9			
10			

**Figura 11.3:** Peixes à deriva! Registre, no gráfico, o valor de ' $p$ ' a cada geração, calculado na **Tabela 11.5**. Ligue todos os pontos com linhas retas.





Espero que você não esteja cansado demais de ficar jogando dados, pois agora quero que repita o experimento, usando um novo grupo de peixes. É importante você fazer essa repetição, pois poderá observar de que modo a deriva pode ser forte e como ela é imprevisível em populações pequenas. Então, tome um cafezinho e mãos à obra! Repita a evolução dos peixes pintados e não pintados na **Figura 11.4**, e preencha a **Tabela 11.6** com os dados.



**Figura 11.4:** Peixes à deriva III! Um exercício em deriva gênica em um gene mitocondrial. Peixe pintado = alelo  $p$ ; peixe sem pintar = alelo  $q$ .

**Tabela 11.6:** Deriva nos peixes II

Geração	Peixes vivos	Peixes pintados	Frequência de $p$
1	10	5	0,50
2			
3			
4			
5			
6			
7			
8			
9			
10			



Agora, registre os dados da **Tabela 11.6** na **Figura 11.3**, que você já havia feito, sobre a evolução da frequência de  $p$  no tempo. Faça uma nova linha ligando os pontos desse segundo experimento (se quiser, use uma cor diferente, para contrastar melhor as duas linhas).

Então, o que você observou? As duas linhas são parecidas? Pode até ser que sejam. Afinal, o acaso faz de tudo! Mas o mais provável é que elas sejam diferentes em vários pontos. Talvez o alelo  $p$  tenha atingido a frequência de 1,0 (nesse caso, dizemos que ele ‘se fixou’) ou tenha atingido a frequência de zero (nesse caso, ele ‘se extinguiu’). Observe que, quando um alelo se fixa ou se extingue, sua variância vira zero (já que a variância =  $pq/n$ ; se  $p=1$ , então  $q=(1-p) = 0$ ; e  $pq = (1 \times 0) = 0$ ), ou seja, a frequência do alelo não varia mais.

Quanto tempo esperamos que leve, em média, para um alelo se fixar na população? Apesar de o cálculo desse número ser um tanto complicado e fora do objetivo desta Aula, o resultado final é muito simples:  $2N$  gerações. Ou seja, assim que um alelo aparece numa população por mutação, a deriva gênica pode fazer com que ele se extinga ou se fixe. Se ele se fixar, o tempo que vai levar para isso, em número de gerações, é cerca de duas vezes o tamanho populacional. Assim, em populações de espécies ameaçadas de extinção, os alelos vão se fixar muito mais rapidamente, pois se uma espécie está ameaçada, é porque já deve ter um tamanho populacional bem pequeno. Por exemplo, os tamanhos populacionais dos guepardos têm se mantido baixos há muito tempo. Assim, a maior parte dos seus locos gênicos só possui um alelo, visto que os outros se extinguiram quando aquele alelo se fixou. Isso quer dizer que as espécies ameaçadas de extinção possuem uma variabilidade gênica (heterozigosidade) muito menor do que a das outras, o que é um problema evolutivo importante para elas, como veremos na Aula 28 de nosso curso, sobre Genética da conservação das espécies.

Como dissemos, e como você mesmo pôde ver, o comportamento da frequência gênica no exercício que fizemos obedece a um modelo estocástico (daí o comportamento ‘bêbado’ da frequência de  $p$ ). Mas qual seria a evolução esperada da frequência de  $p$  nas mesmas condições iniciais, embora com uma população de tamanho infinito de peixes (ou seja, com um modelo determinista)? Para calcular isso, devemos levar em conta as Esperanças das frequências gênicas. Para calcular as Esperanças, você precisa multiplicar a probabilidade e o resultado de cada possibilidade. Em nosso caso, temos três possibilidades (ter 0, 1 ou 2 filhos), cada uma com a mesma probabilidade ( $1/3$ ). Então, as Esperanças seriam:

$$E(\text{filhos}) = E(0; 1/3) + E(1; 1/3) + E(2; 1/3)$$

Que é igual a

$$(0 \times 1/3) + (1 * 1/3) + (2 * 1/3) =$$

$$0 + 1/3 + 2/3 =$$

$$3/3 = 1$$

Ou seja, em um modelo determinista, cada mãe teria, a cada geração, exatamente um filhote. O que aconteceria com a frequência do alelo  $p$  após 10 gerações, sob uma evolução determinista?



A frequência de  $p$  (e de  $q$ ) não mudaria, já que todas as fêmeas passariam para a geração seguinte apenas uma cópia de seu alelo. Ou seja, na ausência de deriva gênica (que é o componente estocástico da evolução), e na ausência das outras forças evolutivas (veja que o alelo  $p$  não tem nenhuma vantagem ou desvantagem seletiva em relação a  $q$ , e nós tampouco consideramos a possibilidade de mutação na população), as frequências tenderiam a permanecer constantes.

Logo, deriva gênica é isso: são as variações que as frequências dos genes sofrem porque, como as populações não são infinitas, cada ser vivo deixa um número variável de filhotes a cada geração. Matematicamente, a deriva gênica é o resultado direto do efeito do desvio padrão das frequências na amostragem dos alelos que acontece a cada ciclo reprodutivo. Como o desvio padrão das frequências depende do tamanho da população, a deriva gênica vai ser maior nas populações menores. Quando as populações são muito grandes, a deriva gênica fica muito pequena. Em populações de milhões de indivíduos (como nos seres humanos e em várias espécies de insetos e peixes), a deriva gênica pode tornar-se quase zero, a não ser que a população esteja dividida em vários grupos menores. Nesse caso, a deriva vai atuar independentemente em cada uma, dependendo das taxas de migração entre elas. Mas isso já é história para outra aula...

Na Aula 12, vamos usar um **SIMULADOR POPULACIONAL** em computador; ou seja, o programa Populus, para fixar melhor o que você viu nesta aula. A Aula 12 deve ser feita no Pólo Regional, pois você vai precisar de ajuda dos monitores para aprender a usar o programa. Aprendendo a usá-lo, você poderá simular, dentro do computador, a evolução das populações naturais sob o efeito da deriva gênica, da seleção natural e da migração.

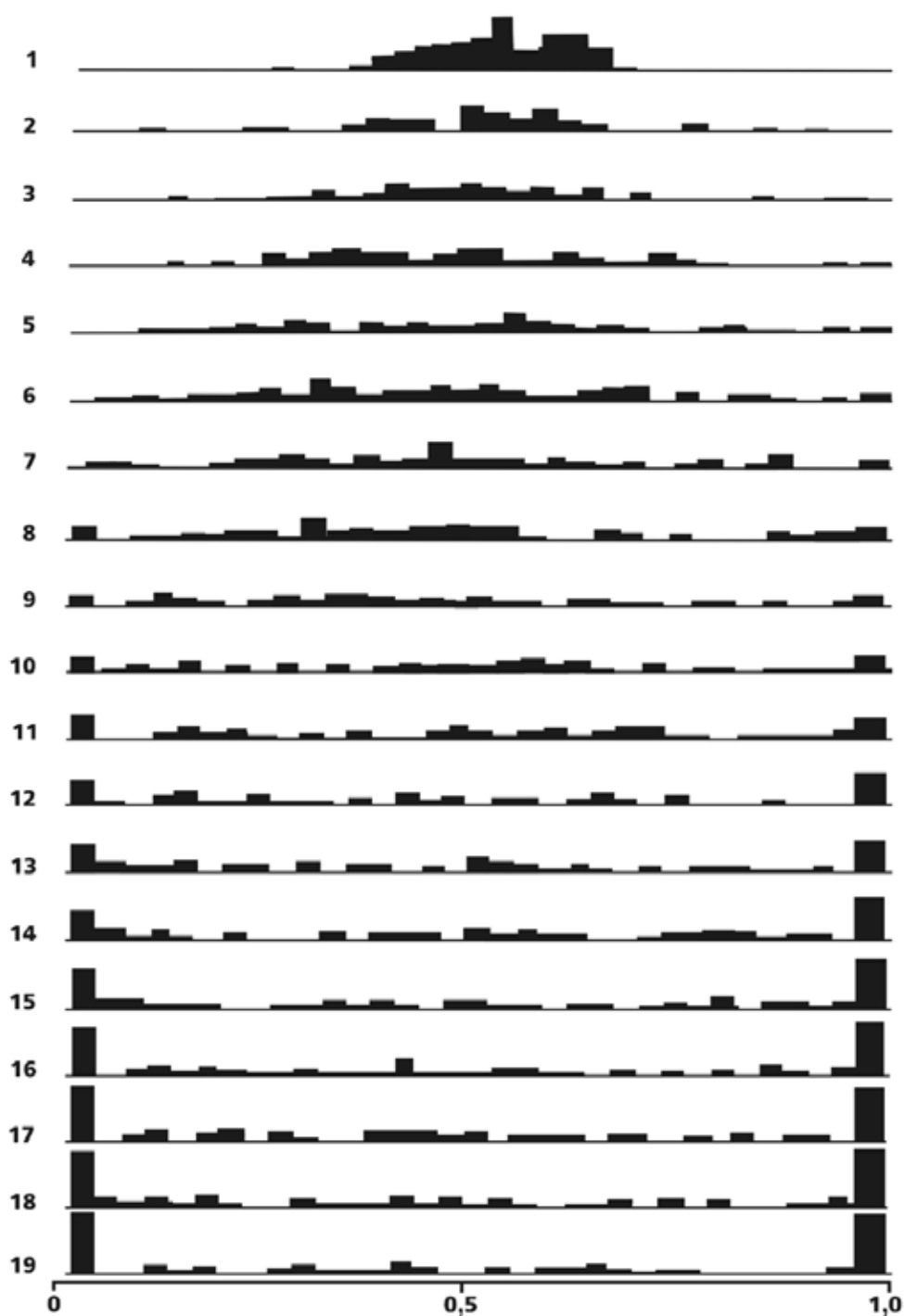
A evolução acontece de maneira muito lenta, do ponto de vista dos seres humanos, pois depende das mudanças das frequências gênicas ao longo das gerações de cada espécie. Assim, quando queremos observar a evolução de uma ou mais populações sob o efeito de alguma força evolutiva (como a deriva ou a seleção natural), temos que estudar essas populações, às vezes, durante anos. Esse é um dos motivos que fizeram com que os pesquisadores de Evolução gostassem tanto das drosófilas: elas têm um ciclo de vida muito curto e ocupam pouco espaço, de modo que em um ano podem ser observadas em laboratório várias gerações de um número grande de indivíduos. Imagine quantos anos levaria para completar o trabalho um pesquisador que quisesse estudar a evolução dos genes dos elefantes! Uma maneira alternativa de estudar a evolução dos genes em populações é criar populações virtuais, que existam apenas dentro da memória dos computadores. Esses programas são chamados **SIMULADORES POPULACIONAIS**. Neles, os pesquisadores podem simular a evolução dessas populações virtuais, controlando os parâmetros populacionais (por exemplo, o número de indivíduos que a população terá ou a taxa de seleção natural que ocorrerá sobre cada genótipo ou ainda a taxa de migração dos indivíduos imaginários entre essas subpopulações virtuais). Simuladores populacionais são úteis, tanto na pesquisa como no ensino de Evolução. Neles, você pode ver, ao toque de um botão, a evolução de centenas ou milhares de gerações de populações imaginárias que podem variar de alguns indivíduos a um número infinito.

**RESUMO**

As freqüências gênicas possuem uma distribuição binomial, pois a freqüência de cada alelo é calculada pela divisão do número de vezes que esse alelo aparece na população pelo número total de alelos. Assim, para cada freqüência  $p$  de um alelo haverá uma freqüência  $q$  de todos os outros, de modo que  $p = 1 - q$ . A variância da distribuição binomial depende do valor da freqüência do alelo e do tamanho da população, pela fórmula  $\text{variância} = pq/n$ . Como  $n$  é o número total de alelos, nos organismos diplóides  $n$  deve ser calculado com o dobro do número amostrado, pois cada organismo vai carregar dois alelos para cada *locus*. A variância (ou o desvio padrão, que é a raiz quadrada da variância) vai ser maior nas freqüências gênicas intermediárias (quando  $p$  é próximo de 0,50) e nos tamanhos populacionais pequenos. Assim, populações pequenas vão variar suas freqüências gênicas muito mais rapidamente do que populações grandes. A expressão mais simples da variância é a reprodução diferencial a cada geração, em que alguns organismos deixam poucos filhotes, enquanto outros deixam muitos, apenas ao acaso. A variação aleatória das freqüências gênicas no tempo é chamada deriva gênica.

**ATIVIDADES FINAIS**

1. Em 1956, foi publicado um artigo fundamental na verificação experimental da deriva gênica. O biólogo Buri estudou a evolução, por deriva, de dois alelos que não tinham vantagem adaptativa em relação um ao outro: os alelos  $bw$  e  $bw^{75}$ . Esses alelos são codominantes, e seus genótipos são facilmente distinguidos em laboratório: as moscas  $bw\ bw$  têm olhos brancos, as  $bw\ bw^{75}$  têm olhos vermelho-escuros, e as moscas  $bw^{75}\ bw^{75}$  têm olhos vermelho-claros. Buri fundou 212 populações, cada uma com oito machos e oito fêmeas. Na primeira geração, ele manteve as freqüências de  $bw$  e  $bw^{75}$  aproximadamente iguais, em 0,50 cada. A partir daí, ele deixou as moscas em cada uma das populações se reproduzirem livremente; a cada geração, porém, ele registrava as freqüências dos alelos  $bw$  e  $bw^{75}$  e retirava todos os filhotes, exceto oito machos e oito fêmeas, que escolhia ao acaso e deixava reproduzir para a próxima geração. Buri fez isso por 19 gerações (você pode imaginar a quantidade enorme de moscas que ele analisou! Foram 212 populações, cada uma com 16 moscas, durante 19 gerações). O resultado que ele obteve, nas 19 gerações, aparece na **Figura 11.5**.



**Figura 11.5:** O resultado de uma cadeia de Markov na freqüência do alelo *bw<sup>75</sup>* de *Drosophila*. Os números representam as várias gerações. O eixo das ordenadas indica a freqüência gênica do alelo, e o eixo das abscissas indica a quantidade de subpopulações que tinham cada freqüência do alelo (desenhado a partir de BURI, 1956).

a) Sabendo que os dois alelos são seletivamente neutros, ou seja, que a seleção natural não favorece nem um nem outro alelo, o que deveria acontecer com as frequências dos alelos  $bw$  e  $bw^{75}$  das populações, se elas tivessem tamanhos infinitos, isto é, se elas seguissem um modelo determinista?

---

---

---

**RESPOSTA**

As frequências deveriam permanecer as mesmas. A evolução das populações, ou seja, as mudanças em suas frequências gênicas, só acontecem na presença de alguma força evolutiva. Como dissemos que os dois alelos são seletivamente neutros (ou seja, nenhum tem vantagem sobre o outro), e como a mutação não é uma força importante em um número tão pequeno de gerações, a única força evolutiva que poderia afetar as frequências gênicas seria a deriva gênica. No entanto, como a deriva gênica depende da variância das frequências, e como a variância é zero em uma população de tamanho infinito, a deriva também seria zero, e as frequências gênicas, então, não mudariam de uma geração para outra.

b) Por que, com o tempo, os picos nos pontos de frequência de  $bw^{75}$  igual a zero e de frequência igual a um foram aumentando?

---

---

---

**RESPOSTA**

Porque nesses pontos a variância é igual a zero, pois só resta um alelo. Lembre-se de que a variância é calculada como  $pq/n$ . Então, nas populações em que a frequência de  $p$  ficou igual a zero,  $pq$  também será zero. Da mesma forma, nas populações em que a frequência de  $p$  ficou igual a um, a frequência de  $q$  será zero (pois  $q = 1 - p$ ), e  $pq$ , mais uma vez, será igual a zero. Nos dois casos a variância será  $pq/n = 0/n = 0$ . Na ausência de variância (e de outras forças evolutivas, como vimos na resposta anterior), as frequências alélicas não irão mudar.

c) Compare as freqüências das populações na primeira geração (em que a maioria delas tinha freqüências de  $bw^{75}$  próximo de 0,50) com as da geração 19 (em que a maioria delas tem freqüências em 0 ou em 1). A média das freqüências do alelo  $bw^{75}$  (considerando as 212 populações ao mesmo tempo) na primeira geração é 0,50. Mas a média da freqüência de  $bw^{75}$  (considerando as 212 populações ao mesmo tempo), na geração 19, também é 0,50. Por quê?

---



---



---

### RESPOSTA

Porque a média total das populações funciona como se fosse uma única população muito grande (no caso, de  $212 \times 16$  moscas, ou seja, 3392 moscas). Populações grandes têm pouca deriva gênica, pois a variância é muito pequena (como  $n$  é muito grande,  $pq/n$  fica muito pequeno). No nosso caso, como começamos com  $p=0,5$ , a variância inicial era, nas populações de 16 moscas =  $(0,5 \times 0,5) / 32$ , isto é, variância =  $0,25/32 = 0,0078$ , o que dá um desvio padrão de  $\sqrt{0,0078} = 0,0884$ ; ou melhor, uma variação de 8,8% na freqüência de  $p$  a cada geração. No caso do conjunto de populações, com  $N=3392$  (e  $n=2 \times 3392 = 6784$ ), o desvio padrão fica sendo  $\sqrt{\frac{0,5 \times 0,5}{6784}} = \sqrt{0,000037} = 0,0061$ , quer dizer, uma variação de menos de 1% no valor de  $p$  a cada geração.

2. No estudo de camarões brasileiros, de Gusmão, Lazoski e Solé-Cava (2000), foram observadas as seguintes freqüências gênicas para o loco responsável pela síntese da enzima peptidase da população da espécie *Farfantepenaeus paulensis*, do Rio Grande do Sul (foram analisados 26 camarões dessa população):

Alelo A = 0,29; Alelo B = 0,65; Alelo C = 0,06

a) Qual a variância do alelo A?

---



---



---

**RESPOSTA**

A variância de A é  $\frac{0,29 \times (1 - 0,29)}{2 \times 26}$ , ou seja  $(0,29 \times 0,71) / 52 = 0,00396$ .

b) Qual o seu desvio padrão?

---

---

---

**RESPOSTA**

O desvio padrão de A é  $\sqrt{0,00396} = 0,0629$ , ou 6,3%.

c) O desvio padrão dos alelos determinava, nas populações de drosófila do experimento de Buri (2000), qual seria a variação dos mesmos de uma geração à outra (ou seja, sua deriva gênica). No caso dos camarões, o desvio padrão que você calculou não determina a Deriva gênica desses alelos. Por quê?

---

---

---

**RESPOSTA**

A variância que você estimou em 2b é a da amostragem. No caso do exercício 1c, você viu a variância da população. A variância da amostragem me diz que, ao estimar a frequência do alelo A na população, eu posso ter feito um erro de 6,3%. O efeito da deriva gênica nesse alelo, na população do camarão, vai depender do número de camarões totais na população. No caso do exemplo das drosófilas, o tamanho amostral (8 machos e 8 fêmeas) era o mesmo que o tamanho da população. Ou seja, eu analisava os alelos de todos os indivíduos que participavam da reprodução da próxima geração. Assim, podemos dizer que, no caso dos camarões, existem duas medidas de n: uma é o tamanho da população total (que determina a variância da frequência dos genes na produção da próxima geração); a outra é o tamanho da amostra que usamos para estimar as frequências gênicas da população (que determina a variância da minha medida das frequências de cada alelo). Essa distinção é muito importante. A não ser quando se medem todos os indivíduos da população, o que é raro, (exceto em experimentos de laboratório, como o das drosófilas que Buri usou), sempre existirão duas variâncias: a populacional e a amostral. A primeira é a que tem importância evolutiva; a segunda é apenas uma medida do meu erro amostral.



3. O teste estatístico de Student (Teste  $t$ ) só deve ser usado para comparar coisas que têm uma distribuição normal. O teste do qui-quadrado só deve ser usado para comparar coisas com distribuição binomial. Se eu estivesse querendo saber se as mulheres cariocas são mais altas que as mulheres paulistas, e tivesse medido 200 cariocas e 200 paulistas, que teste eu deveria usar para comparar as medidas?

---

---

---

**RESPOSTA**

*Como a distribuição de alturas é do tipo normal, deve-se usar o teste  $t$ . Lembre-se de que as variáveis de distribuição contínua (como a altura das cariocas e das paulistas) obedecem, em geral, a uma distribuição normal. Teria sido usado um Teste de Student também se fossem medidos os pesos ou o tamanho dos pés dessas mulheres. Por outro lado, deveríamos usar um teste do qui-quadrado se tivéssemos comparado as mulheres paulistas e as cariocas em termos da proporção de loiras ou morenas, ou as proporções de seguidoras das várias religiões. Nesses casos, por se tratarem de proporções, as distribuições seriam binomiais.*

4. Compare a **Figura 11.3** com a **Figura 10.7** (da aula passada). Qual a semelhança entre as duas? Qual o motivo dessa semelhança?

---

---

---

**RESPOSTA**

*As duas figuras apresentam linhas que não são retas. Elas são quebradas em vários segmentos de retas, que aumentam ou diminuem a cada ponto de maneira aleatória. Os dois gráficos têm figuras com esse comportamento errático, porque eles retratam o comportamento de variáveis sob modelos estocásticos.*

5. Quando um alelo tem maior variância, quando ele tem freqüência alta, baixa ou média?

---

---

---

**RESPOSTA**

Com as freqüências médias. A variância com dois alelos é máxima quando  $p = q = 0,5$ . É só você observar a **Figura 11.1**. Nela você vê que os desvios padrões (as curvas das variâncias seriam parecidas) aumentam quando as freqüências são próximas de 0,5. Assim, por exemplo, em uma amostra com  $n = 10$ , se  $p = 0,2$  e  $q = 0,8$ , a variância será  $(0,2 \times 0,8) / 10 = 0,016$ . Essa variância será a mesma se  $p = 0,8$  e  $q = 0,2$ , certo? (já que a ordem dos fatores não altera o produto). Com o mesmo  $n$ , se  $p = 0,5$  e  $q = 0,5$ , a variância será de 0,025, ou seja, maior que quando  $p = 0,2$  ou  $p = 0,8$ .

### AUTO-AVALIAÇÃO

Esta aula é uma continuação clara da aula passada sobre modelos estocásticos. A deriva gênica é, afinal, o resultado direto da variação estocástica. O mais importante que você deve ter aprendido nesta aula é que, quanto menor a população, maior será a deriva gênica. É importante que você não tenha apenas guardado isso. Se você fez os exercícios com os alelos dos peixes (**Figuras 11.2 e 11.4**), deve ter aprendido, de verdade, como as freqüências dos genes variam ao acaso, mesmo quando a regra é constante. Você viu que, se o modelo fosse determinista, as freqüências dos alelos nos peixes não mudariam. Uma outra coisa importante é a formalização matemática da deriva gênica, ou seja, ela é o resultado do desvio padrão das freqüências alélicas a cada geração, por causa da amostragem dos alelos que ocorre durante a reprodução. A primeira atividade final desta aula, baseada no trabalho com drosófilas (BURI, 1956) é importante por ser um marco clássico nos estudos de populações, e por preparar você para a aula sobre endocruzamento e migração.

Uma outra coisa importante, que aparece na Atividade 2c, é a distinção entre variância populacional e variância amostral. Essa distinção não foi citada no texto, para que você tivesse a oportunidade de descobri-la ao tentar resolver a atividade. Repare que, na maioria das vezes, quando cientistas estudam as populações naturais, eles têm de lidar com a variância amostral. As freqüências gênicas estimadas com a amostragem são boas aproximações das freqüências gênicas das

populações estudadas (e o que indica a precisão dessas aproximações é justamente o desvio padrão amostral). Mas o que importa, na evolução das populações por deriva gênica, é a variância populacional. Para os cientistas, a única maneira de ter uma idéia da variância populacional é usar as freqüências gênicas estimadas pela amostragem para o cálculo da variância populacional, usando estimativas do tamanho populacional a partir de outras fontes (como contagens diretas dos organismos da população, técnicas ecológicas de captura e recaptura, ou mesmo estimativas indiretas, a partir dos dados de variabilidade gênica).