BIO-103 Biologia evolutiva

Excerto da apostila de Genética de populações para as disciplinas BIO-305 e BIO-212 (1986-2006), vários autores.

EQUILÍBRIO DE HARDY-WEINBERG

Um dos aspectos importantes do estudo da Evolução é a análise da variabilidade genética das populações e do seu comportamento ao longo das gerações. Esses aspectos constituem a preocupação fundamental da Genética de populações, que procura descrever a composição genética das populações bem como sua resposta frente à atuação de fatores tais como o tipo de cruzamento, o tamanho da população, a mutação, a migração e os vários tipos de seleção. A Genética de populações, por quantificar os fenômenos evolutivos, fornece parâmetros para a análise da variabilidade genética das populações, sua origem e manutenção.

Vamos, portanto, mostrar como essa variabilidade é caracterizada para fins do estudo da Genética de populações.

Frequências gênicas

A fim de conceituar frequência gênica, vamos considerar inicialmente o que ocorre com um par de genes autossômicos em organismos diplóides. Supondo que não haja dominância, poderemos distinguir os três genótipos possíveis, representados por AA, Aa e aa. Esses três genótipos corresponderão a três classes fenotípicas diferentes. Assim, em uma população constituída de N indivíduos poderemos contar D indivíduos AA, H indivíduos Aa e R indivíduos aa.

Os valores D, H e R são chamados de frequências absolutas (note as letras maiúsculas) enquanto que esses valores, divididos pelo total de indivíduos da população (N) nos dão as frequências relativas (representadas pelas mesmas letras, só que minúsculas).

AA	Aa	aa
$d = \frac{D}{N}$	$h = \frac{H}{N}$	$r = \frac{R}{N}$

A soma das frequências relativas é sempre 1.

$$\frac{D}{N} + \frac{H}{N} + \frac{R}{N} = \frac{D + H + R}{N} = \frac{N}{N} = 1$$

As frequências relativas podem ser interpretadas, no caso de amostragens muito grandes, como probabilidades, ou seja, d é a probabilidade de se tomar "ao acaso" um indivíduo AA desta população.

Podemos agora procurar saber quais as frequências gênicas nessa população. Neste caso, o método direto para estimar as frequências gênicas é o da contagem simples.

Dada a população:

$$\begin{bmatrix} AA & Aa & aa \\ & & \\ D & H & R \end{bmatrix} N$$

contamos o número de alelos A e a e estimamos as frequências gênicas. Por se tratar de uma população diplóide vamos atribuir a cada indivíduo dois genes. A população toda terá, pois, 2N genes. Os indivíduos AA terão 2D genes A e os indivíduos Aa terão H genes A, perfazendo um total de 2D + H genes A em uma população com um total de 2N genes; logo, a frequência do alelo A será:

$$f(A) = \frac{2D + H}{2N} = \frac{2D}{2N} + \frac{H}{2N} = d + \frac{h}{2} = p$$

Com o mesmo raciocínio veremos que a frequência do alelo a será:

$$f(a) = \frac{2R + H}{2N} = \frac{2R}{2N} + \frac{H}{2N} = r + \frac{h}{2} = q$$

pode-se verificar que:

$$p + q = d + \frac{h}{2} + \frac{h}{2} + r = 1$$

Uma vez calculada a frequência de um alelo, a frequência do outro pode ser obtida pela diferença em relação à unidade, uma vez que

$$p + q = 1$$

$$p = 1 - q \text{ e } q = 1 - p$$

Exercícios:

1. Determine nos casos que se seguem as frequências relativas das classes fenotípicas e as frequências gênicas:

a) AA 25 · Aa 60 · aa 15 N = d= h= r= f(A) = p = f(a) = q =	b) AA 320 · Aa 0 · aa 80 N = d= h= r= f(A) = p = f(a) = q =
c) AA 0 · Aa 120 · aa 80 N = d= h= r= f(A) = p = f(a) = q =	d) AA 10 · Aa 180 · aa 810 N = d= h= r= f(A) = p = f(a) = q =

2. Determine, nos casos que se seguem, as frequências absolutas das classes genotípicas e as frequências gênicas (para uma população de 1000 indivíduos):

a) AA D= 0,30 · Aa H= 0,60 · aa R= 0,10 p= q=	b) AA D= 0,36 · Aa H= 0,48 · aa R= 0,16 p= q=
c) AA D= 0,20 · Aa H= 0,80 · aa R= 0,00 p= q=	d) AA D= 0,58 · Aa H= 0,04 · aa R= 0,38 p= q=

O que você pode concluir a partir dos resultados do exercício 2?

EQUILÍBRIO (OU LEI) DE HARDY-WEINBERG

O que fizemos até o momento foi representar um par de genes autossômicos, sem dominância, em uma população diplóide e estimar as frequências dos alelos. Agora verificaremos o que acontecerá com uma população desse tipo na geração seguinte.

Por isso consideramos uma população com reprodução sexuada, que se reproduza por fecundação cruzada.

Vamos considerar, para simplificar o problema, um modelo com "gerações discretas", ou seja, uma população na qual não haja cruzamentos entre indivíduos pertencentes a duas ou mais gerações diferentes. Assim, temos uma população de N indivíduos adultos:

$$\begin{bmatrix} AA & Aa & aa \\ & & \\ D & H & R \end{bmatrix} N$$

Vamos supor que os cruzamentos nesta população ocorram ao acaso. Tal fenômeno é conhecido como pan-mixia e diz-se que a população que se reproduz assim é pan-mítica.

Pan-mixia significa que a probabilidade de um indivíduo de qualquer genótipo cruzar com outro pertencente a qualquer genótipo depende apenas das frequências genotípicas. Isso é o mesmo que dizer que não há preferências, seja ela por genótipos iguais ou diferentes, na escolha de parceiros.

Considerando as frequências relativas dos genótipos como probabilidades, podemos dizer que a probabilidade dos indivíduos AA, com frequência relativa d, cruzar com outro indivíduo de mesmo genótipo é simplesmente dXd = d^2.

Assim, podemos construir um quadro com as probabilidades, ou frequências, dos cruzamentos "ao acaso".

QUADRO 1 - Frequências de cruzamentos "ao acaso".

	machos	AA	Aa	aa
fêmeas	frequências genotípicas	d	h	r
AA	d	d^2	dh	dr
Aa	h	hd	h^2	hr
aa	r	rd	rh	r^2

Lembre-se que a soma dessas probabilidades ou frequências de cruzamentos será sempre igual a 1, pois

$$(d + h + r) \times (d + h + r) = 1 \times 1 = 1$$

Agora vamos verificar qual a descendência deixada por cada um desses cruzamentos. Novamente teremos de fazer algumas considerações a respeito de como calcular o número de descendentes. O número de descendentes por casal é variável, mas podemos admitir que este número não dependa dos genótipos dos indivíduos que formam o casal, sendo, em média, o mesmo.

Assim, podemos apresentar a frequência de descendentes de cada classe de casal pela própria frequência dos cruzamentos. Podemos assim determinar os genótipos dos descendentes de cada casal e suas respectivas frequências.

Exercício: Complete o quadro abaixo, determinando os genótipos dos descendentes de cada tipo de cruzamento e suas respectivas frequências.

	machos	AA	Aa	aa
fêmeas		d	h	r
AA	d	d^2 AA	dh/2 AA dh/2 Aa	
Aa	h			
aa	r			

Para determinar as frequências dos diferentes genótipos na nova geração basta somar as frequências das 3 classes genotípicas dos descendentes.

tipo de cruzamento	frequência de cruzamento	Descendentes		
		AA	Aa	aa
AA X AA	d^2	d^2		
AA X Aa	2dh	dh	dh	
AA X aa	2dr		2dr	
Aa X Aa	h^2	$h^2/4$	$h^2/2$	$h^2/4$
Aa X aa	2hr		hr	hr
aa X aa	r^2			r^2
Total	1	$(d+h/2)^2$	$2(d+h/2)(r+h/2)$	$(r+h/2)^2$

Lembrando que $(d + h/2) = p$ e $(h/2 + r) = q$ verificamos imediatamente que a distribuição das frequências genotípicas poderá ser expressa assim:

AA	Aa	aa	
p^2	$2pq$	q^2	ou
p^2	$2p(1-p)$	$(1-p)^2$	ou

$(1-q)^2$ $2q(1-q)$ q^2

As duas últimas representações têm a vantagem de ter apenas uma variável de frequência gênica.

As frequências gênicas não mudam, pois:

$$p_1 = d_1 + \frac{h_1}{2} = p^2 + pq = p^2 + p - p^2 = p$$

Note que uma série de condições foi imposta na elaboração do modelo, o que levou a população ao equilíbrio. Estas condições são:

- população de tamanho infinito;
- reprodução sexuada, por fecundação cruzada;
- pan-mixia;
- ausência de mutação;
- ausência de migração diferencial;
- ausência de seleção.

Nestas condições, uma população não sofre alterações em suas frequências gênicas, ao longo das gerações, nas proporções:

p^2 $2pq$ q^2

Estas proporções serão atingidas em uma única geração.

Alelos Múltiplos

O princípio visto acima pode ser estendido para qualquer número de alelos.

Sejam os alelos:

A_1 A_2 A_3 ... A_N, com as frequências gênicas:
p_1 p_2 p_3 ... p_N

No equilíbrio, as frequências dos genótipos homozigotos serão:

$$f(A_1A_1) = p_1^2;\quad f(A_2A_2) = p_2^2;\quad f(A_3A_3) = p_3^2 \quad \ldots \quad f(A_NA_N) = p_N^2$$

e as frequências dos genótipos heterozigotos serão:

$$f(A_1A_2) = 2p_1p_2;\ f(A_1A_3) = 2p_1p_3;\ f(A_2A_3) = 2p_2p_3 \ldots\ f(A_{N-1}A_N) = 2p_{N-1}p_N$$

Teste para verificar se a população está ou não em equilíbrio de H.W.

Dada uma amostra populacional com a identificação completa dos genótipos feita, surge o problema: Esta pode ser uma amostra representativa de uma população que está em equilíbrio de Hardy-Weinberg? Vamos estudar primeiramente o caso da ausência de dominância, em um loco com dois alelos:

$$\begin{bmatrix} AA & Aa & aa \\ & & \\ D & H & R \end{bmatrix} N$$

onde a frequência do alelo A é $p = \dfrac{2D + H}{2N}$ e a do a é $q = \dfrac{2R + H}{2N}$

Note que esta estimativa não requer que a população esteja em equilíbrio. Conhecidos os valores de p e q, podemos então estimar as frequências das três classes genotípicas, como se a população estivesse em equilíbrio. Essas frequências absolutas, como já vimos, são:

AA	Aa	aa
$p^2.N$	$2pq.N$	$q^2.N$, onde

$N = D + H + R$ e $p = 1 - q$.

Exercício:

Calcule as frequências gênicas e as frequências genotípicas esperadas (relativas e absolutas) para as seguintes amostras:

1.	AA	Aa	aa		**3.**	A	Aa	aa	
	204	494	302			40	220	240	
N =					N =				
p =			q =		p =			q =	
p^2=		2pq=		q^2=	p^2=		2pq=		q^2=
p^2.N=		2pq.N=		q^2.N=	p^2.N=		2pq.N=		q^2.N=
2.	AA	Aa	aa		**4.**	AA	Aa	aa	
	20	50	30			80	100	120	
N =					N =				
p =			q =		p =			q =	
p^2=		2pq=		q^2=	p^2=		2pq=		q^2=
p^2.N=		2pq.N=		q^2.N=	p^2.N=		2pq.N=		q^2.N=

Como temos os valores obtidos e os esperados, podemos ter uma idéia se a população da qual a amostra foi retirada estava em equilíbrio ou não. Como foi visto no exercício acima, aparece uma dificuldade: se os valores forem exatamente iguais (o que, na prática, é pouco provável) concluimos que a população encontra-se em equilíbrio, mas se isto não acontece, temos que lançar mão de um outro método para tomar a decisão de forma objetiva. A estatística nos fornece a ferramenta apropriada para a comparação de proporções: o teste do X^2 (que se lê qui-quadrado). **Este teste só pode ser aplicado com os valores absolutos.**

	AA	Aa	aa	Total
Valores observados (O)	D	H	R	N
Valores esperados (E)	$p^2.N$	$2pq.N$	$q^2.N$	N

O valor do X^2 é dado por:

$$X^2 = \sum_{i=1}^{k} \frac{(observado_i - esperado_i)^2}{esperado_i}$$

ou seja: $$X^2 = \frac{\left(D - p^2.N\right)^2}{p^2.N} + \frac{\left(H - 2pq.N\right)^2}{2pq.N} + \frac{\left(R - q^2.N\right)^2}{q^2.N}$$

O valor obtido através dessa soma corresponde ao X^2. Para saber se a população está ou não em equilíbrio, ou seja, se a amostra foi ou não retirada de uma população em equilíbrio, basta comparar este valor com os valores dados na tabela do X^2. O estabelecimento de um determinado valor de X^2 dependerá, inicialmente, da escolha do "nível de significância", que consiste no risco que se corre, em termos de probabilidade, de rejeitar uma hipótese verdadeira. Esse nível deve ser fixado previamente e, em geral, para os experimentos biológicos usa-se o nível de significância (alfa) de 0,05 (ou 5%).

Temos, entretanto, vários valores de X^2 para 5%, cada um correspondendo a um determinado número de graus de liberdade (primeira coluna da tabela). É através desse valor que vamos procurar o X^2 na tabela.

O número de graus de liberdade é dado pela quantidade de classes independentes envolvidas na análise. No caso de três genótipos, temos apenas 1 classe independente para o mesmo N e o mesmo p. Por exemplo, se existem 34 indivíduos AA em um total de 100 e a frequência do alelo A é 0,6, então as outras duas classes estarão já determinadas: 52 Aa e 14 aa. Quaisquer outros valores de H e R alterarão as frequências gênicas ou o total, o que não é permitido. Então existe uma única classe independente e, portanto, um único grau de liberdade.

Na prática, o número de graus de liberdade pode ser calculado assim:

g.l.= Número de classes-(número de parâmetros da amostra utilizados no teste para calcular os esperados das classes)

No nosso caso, temos 3 classes e usamos dois parâmetros: a frequência de um alelo e o total. g.l. = 3 - 1 - 1 = 1.

Assim, vamos procurar na tabela simplificada, anexa, na linha correspondente ao número de graus de liberdade <u>1</u>, o valor do X^2 para 5%, que é 3,841. Se o valor obtido for maior que 3,841, rejeitamos a hipótese de equilíbrio; se for menor aceitamos a hipótese.

<u>Exercício</u>:

Aplique o teste do X^2 para as populações do último exercício, e verifique quais estão em equilíbrio.