

Medindo Seleção Natural, um guia teórico

Ana Paula A Assis, Gustavo Requena, Daniela M Rossoni, Bárbara MA
Costa, Lucas Medeiros, Vinícius Montagner

Departamento de Ecologia, Universidade de São Paulo, Brasil

Abstract

Por muito tempo a discussão sobre seleção natural era relacionada às mudanças nas frequências de alelos ao longo do tempo. Embora muito da nossa compreensão de como ocorrem mudanças evolutivas tenha sido alcançada com essa abordagem, dado que a maior parte das características são contínuas (peso, altura, tamanho corporal, etc) faltava uma teoria que ligasse o efeito de seleção natural a estes caracteres contínuos. Russell Lande em 1976 desenvolveu tal teoria em um artigo altamente influente. Neste artigo Lande se utilizou da abordagem de genética quantitativa que já era bem desenvolvida na época, devido a experimentos de melhoramento animal. A grande contribuição deste trabalho foi ampliar a visão de Wright sobre paisagens adaptativas de frequência de alelos para fenótipos. Ele partiu da equação de resposta à seleção usada em melhoramento animal e mostrou que ela é similar a equação que modela uma paisagem adaptativa proposta por Wright. Ele investigou casos onde seleção natural é tanto independente de frequência, isto é, as mudanças nas frequências de cada um dos alelos não afeta o valor relativo de fitness de cada classe fenotípica e modelou também como se dá a mudança de fitness na população com a mudança na média em um cenário onde seleção é dependente de frequência.

Keywords: Seleção, Seleção dependente de frequência, Deriva genética

1. Lande 1976:

1.1. Seleção de alelos e paisagens adaptativas

Sewall Wrigth desenvolveu a teoria sobre como alelos devem mudar ao longo do tempo se estiverem sobre seleção natural. Ele se utilizou de uma

analogia conhecida como paisagem adaptativa. Essa analogia mostra como o fitness médio da população muda em relação à mudança na frequência de um determinado genótipo na população. O fitness é descrito pela altura da paisagem adaptativa, sendo que a frequência alélica que possui maior valor adaptativo está associada a um pico adaptativo (maior altura no eixo de fitness), isto é um local na paisagem adaptativa que maximiza o fitness médio da população. Ele demonstrou que seleção natural sempre vai levar a população a atingir um pico adaptativo.

Uma população vai evoluir até atingir o pico adaptativo, levando-se em consideração as restrições de herança mendeliana. A beleza de usar a analogia de paisagens adaptativas é que ela consegue relacionar a geometria da paisagem (padrão de seleção) às frequências genotípicas da população. A mudança na frequência do alelo p de uma geração para a outra pode ser descrita, então, como:

$$\Delta p = \frac{1}{2} \frac{p(1-p)}{\bar{W}} \frac{\partial \bar{W}}{\partial p} \quad (1)$$

Na qual p é a frequência do alelo de interesse. \bar{W} é o fitness médio da população. A relação $p(1-p)$ deriva da variância da distribuição binomial, mostrando que a mudança na frequência do alelo é dependente de sua variância. A relação $\frac{\partial \bar{W}}{\partial p}$ é uma derivada parcial que estima como o fitness médio muda, quando a frequência de p muda, mantendo-se a frequência de todos os outros alelos constantes. Isto é, nos conta como o alelo p afeta o fitness médio da população. Agora vamos expandir esse raciocínio para mudanças fenotípicas.

1.2. Paisagem adaptativa para fenótipos

Assumindo-se que o fitness é constante e que as populações são infinitas (deriva genética não importa) podemos usar a equação de resposta à seleção de melhoramento animal, em gerações discretas (Falconer, 1960):

$$\begin{aligned} \Delta \bar{z} &= \bar{z}_{t+1} - \bar{z}_t \\ &= (z_{wt} - \bar{z}_t)h^2 \end{aligned} \quad (2)$$

Isso é, a mudança na média do caráter de uma geração a outra pode ser medida como a diferença entre a média na geração $(t+1)$ menos a da geração (t) , que equivale a diferença nas médias na geração t antes (\bar{z}) e depois do

episódio de seleção $(\bar{z}_w)^1$ vezes a herdabilidade do caráter².

As distribuições do caráter (z) na população no tempo t (antes da seleção) será descrita como $p(z,t)$. O fenótipo médio, antes do episódio de seleção é portanto:

$$\bar{z}_t = \int z p(z, t) dz \quad (3)$$

Isto é, o fenótipo médio é igual a integral de z em relação a distribuição de todos os valores de z na população. Embora pareça complicada, essa definição vem da própria definição de média como a integral de uma função de probabilidade de densidade (PDF, *probability density function*) em relação a uma variável. Se considerarmos o fitness do indivíduo com o caráter z na população igual $W(z)$, o fitness médio desta população será igual a:

$$\bar{W}_t = \int p(z, t) W(z) dz \quad (4)$$

que é a integral de todos os valores de $W(z)$ dada a distribuição dos valores do caráter $p(z,t)$, de novo uma integral de uma variável (neste caso $W(z)$) em relação a sua PDF (probability density function). Desta forma, o valor do fenótipo depois do episódio de seleção será igual a:

$$\bar{z}_{wt} = \frac{1}{\bar{W}_t} \int z p(z, t) W(z) dz \quad (5)$$

Isso é, uma função do fitness e da distribuição de z na população. Até essa equação nós não tínhamos nenhuma grande premissa para considerar

¹Essa diferença é o que nós chamamos de diferencial de seleção (S), ou o que Lande chamou de força da seleção.

²Herdabilidade é uma medida que descreve quanto da variação existente em um caráter, de uma dada população, é devida à variação genética. Outras fontes de variação são efeitos ambientais ou erros de mensuração. Na literatura é comum ver a distinção entre *narrow-sense heritability* e *broad-sense heritability*. A diferença é que a primeira considera só os efeitos aditivos (isto é mecanismos de herança nos quais os efeitos combinados de alelos são iguais as somas de seus efeitos individuais) e a segunda considera efeitos de dominância e epistáticos. Na maioria dos casos, biólogos evolutivos estão interessados nos efeitos aditivos, pois eles que determinam o quanto os filhos são parecidos com os pais (uma vez que tendo efeitos de dominância e epistáticos, os filhos podem ser muito diferentes dos pais).

a mudança evolutiva. Nós apenas consideramos que cada valor fenotípico (z) tinha associado a ele um valor de fitness (W) e dada a distribuição dos valores fenotípicos na população nós conseguiríamos calcular o fitness médio e o diferencial de seleção, nesta população, de uma geração a outra. A partir de agora, nós teremos uma premissa bastante importante, na qual a maior parte dos trabalhos que medem/estimam/trabalham com seleção de caracteres contínuos fazem, isto é, vamos considerar que os caracteres tem distribuição normal³ na população. Embora essa premissa pareça sensata, é importante sempre termos em mente que ao aplicarmos modelos de seleção baseados em genética quantitativa, essa importante premissa, está por trás de todos nossos resultados. Além disso, o Lande chama atenção que mesmo caracteres que não possuam distribuição normal, é possível transformá-los para que assumam essa distribuição. Partindo-se então dessa premissa, nós podemos descrever a distribuição de z na população como:

$$p(z, t) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(z-\bar{z})^2}{2\sigma^2}} \quad (6)$$

Esta função nada mais nos conta que para estimarmos a distribuição de z na população nós só precisamos saber duas coisas sobre essa população, sua média (\bar{z}) e sua variância (σ^2). Por isso, a suposição de normalidade facilita muito o estudo de seleção. Outra suposição importante é de que a variância é independente da média. Nem sempre isso ocorre em populações naturais, especialmente se formos estudar animais que possuam escalas muito díspares, como por exemplo um elefante e um ratinho. Neste caso, a média está associada a variância, uma vez que ratinhos terão médias muito menores assim como variâncias, por outro lado os elefantes terão médias em todos os caracteres muito maiores, e consequentemente variâncias. No entanto, na maior parte dos estudos evolutivos, se estivermos interessados em mudanças microevolutivas (isto é dentro da mesma espécie) as variâncias de fato serão independentes da média. Se aplicarmos as equações 3 a 6 nós temos que:

³A distribuição normal também pode ser chamada de gaussiana. A função representada pela equação 6 nada mais é do que a função que descreve uma distribuição normal (a função de densidade de probabilidade- PDF).

$$\begin{aligned}
\frac{\partial \bar{W}}{\partial \bar{z}_t} &= \int \frac{\partial p(z, t)}{\partial \bar{z}_t} W_t dz \\
&= \int \frac{z - \bar{z}}{\sigma^2} p(z, t) W_t dz \\
&= \frac{\bar{W}}{\sigma^2} (\bar{z}_{Wt} - \bar{z}_t)
\end{aligned} \tag{7}$$

E assim essa relação nos permite expressar a mudança observada na equação 2 como:

$$\Delta \bar{z} = \frac{h^2 \sigma^2}{\bar{W}} \frac{\partial \bar{W}}{\partial \bar{z}_t} = h^2 \sigma^2 \frac{\partial \ln(\bar{W})}{\partial \bar{z}_t} \tag{8}$$

Desta forma a mudança no caráter z é proporcional a derivada parcial de como o fitness médio muda em relação às mudanças no caráter z , isto é, a mudança é tangente à paisagem adaptativa (\bar{W}) apontando para a direção de mudança de z que maximize a mudança no fitness da população. Esta fórmula é extremamente similar a forma derivada por Wrigth (equação 1), onde a mudança evolutiva é na direção que máxima o fitness da população e é dependente da variação existente na população. A partir disso podemos afirmar que fenótipos também podem ser descritos pela mesma analogia de paisagem adaptativa como mudanças nas frequências de alelos.

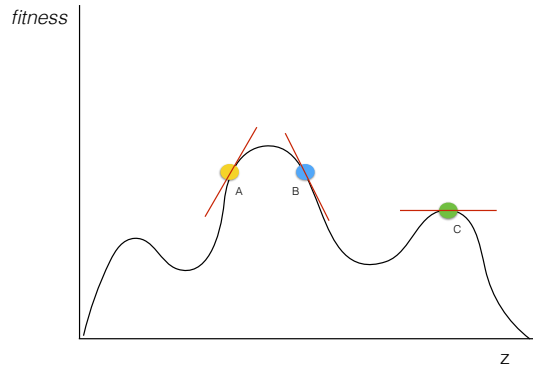


Figure 1: relação entre o fitness

1.3. Seleção dependente de frequência

Até agora nós consideramos que a frequência de um determinado fenótipo na população não afeta o fitness das diferentes classes fenotípicas, isto é, o fitness é independente de frequência. Agora vamos ver como esperamos que ocorra a mudança fenotípica caso seleção seja dependente de frequência.

2. Lande, 1979

Até agora nós vimos como a evolução de uma única característica evolui em resposta a seleção natural. No entanto, os organismos são formados por grupos de caracteres que interagem e são correlacionados, afetando como cada um destes caracteres responde a seleção natural. Em 1979, Lande, então desenvolveu a teoria de como esperamos que estes conjuntos de caracteres evoluam e respondam a seleção natural, dada a correlação observada entre diferentes caracteres.

Vamos começar definindo um vetor \bar{z} formado por um conjunto de caracteres $(\bar{z}_1, \bar{z}_2, \bar{z}_3 \dots \bar{z}_n)$, sendo que cada elemento deste vetor representa a média de um determinado caráter em uma dada população. Assumindo-se que estes caracteres são determinados por fatores genéticos **aditivos** e ambientais, **independentes** um do outro. Desta forma, podemos representar o valor do caráter z_i na população como:

$$z_i = x_i + e_i \quad (9)$$

No qual x_i representa o valor de acasalamento ⁴ do caráter z_i e e_i representa a variação ambiental. Mais uma vez vamos assumir que estes diferentes caracteres possuem distribuição multivariada normal na população. Desta forma, a distribuição pode ser representada para um único caráter como a equação 6, mas como estamos lidando com múltiplos caracteres podemos descrever essa relação como:

$$g(x) = \sqrt{2\pi^{-m}|G^{-1}|} e^{-\frac{1}{2}(x-\bar{x})^T G^{-1}(x-\bar{x})} \quad (10)$$

Na qual $|G^{-1}|$ representa o determinante da inversa da matriz de efeitos genéticos aditivos. Da mesma forma, podemos representar

⁴Valor de acasalamento refere-se a soma dos valores aditivos afetando o caráter

Appendix A. Como sair da equação 2 e chegar na equação 8

Vamos passo por passo verificar como ele chegou no resultado expresso na equação 7. Primeiramente, nós queremos encontrar como o fitness médio muda em relação ao valor médio da característica na população, lembrando que o fitness médio é $\int p(z, t)W(z) dz$ isso quer dizer que a derivada parcial do fitness médio sobre \bar{z} pode ser reescrita como:

$$\frac{\partial \bar{W}}{\partial \bar{z}_t} = \frac{\partial \int p(z, t)W(z) dz}{\partial \bar{z}_t} \quad (\text{A.1})$$

Vamos focar, no momento, apenas para o numerador desta equação. Nós podemos passar a derivada para dentro da integral:

$$\int \partial p(z, t)W(z) dz \quad (\text{A.2})$$

Uma vez que temos a derivada de um produto, isto é $p(z, t)$ e $W(z)$, nós podemos usar a regra do produto de cálculo e assim obtermos:

$$\int \partial p(z, t)W(z) dz = \int \partial[p(z, t)]W(z) dz + \int p(z, t)\partial[W(z)] dz \quad (\text{A.3})$$

Uma vez que o fitness de cada valor de z é independente do valor médio de z , a derivada parcial de $\partial[W(z)]$ é igual a zero. Desta forma, a parte à direita da equação A.3 assume valor zero. Como temos como premissa que a distribuição do caráter na população é normal a expressão $\partial[p(z, t)]$ é igual a $\partial \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(z-\bar{z})^2}{2\sigma^2}}$.

$$\frac{\partial \bar{W}}{\partial \bar{z}_t} = \int \partial \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(z-\bar{z})^2}{2\sigma^2}} W(z) dz \quad (\text{A.4})$$

Uma vez que $\frac{1}{\sqrt{2\pi\sigma^2}}$ é uma constante, nós podemos retirá-la da derivada ficando: $\frac{1}{\sqrt{2\pi\sigma^2}} \partial(e^{-\frac{(z-\bar{z})^2}{2\sigma^2}})$. Usando a regra da cadeia nós vemos que a derivada desta expressão é igual a: $\partial(e^{-\frac{(z-\bar{z})^2}{2\sigma^2}}) \partial - \frac{(z-\bar{z})^2}{2\sigma^2}$. A derivada da função exponencial é sempre ela mesma, portanto nossa função fica:

$$\frac{\partial \bar{W}}{\partial \bar{z}_t} = \int \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(z-\bar{z})^2}{2\sigma^2}} \partial - \frac{(z-\bar{z})^2}{2\sigma^2} W(z) dz \quad (\text{A.5})$$

A distribuição do caráter na população $[p(z,t)]$ é igual a $\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(z-\bar{z})^2}{2\sigma^2}}$. Portanto, podemos substituir essa expressão na equação A.5. Calculando a derivada da expressão $\partial - \frac{(z-\bar{z})^2}{2\sigma^2}$ temos:

$$\partial\left(-\frac{z^2 - 2z\bar{z} + \bar{z}^2}{2\sigma^2}\right) = \frac{1}{2\sigma^2}\partial(-z^2 + 2z\bar{z} - \bar{z}^2) = \frac{2z - 2\bar{z}}{2\sigma^2} = \frac{z - \bar{z}}{\sigma^2} \quad (\text{A.6})$$

O que faz com que a equação 7 possa ser expressa como :

$$\frac{\partial \bar{W}}{\partial \bar{z}_t} = \int p(z,t) \frac{z - \bar{z}}{\sigma^2} W(z) dz \quad (\text{A.7})$$

Como a variância é uma constante nós podemos retirá-la da integral, distribuindo o produto de $(z - \bar{z})$ e re-expressar a equação como:

$$\frac{\partial \bar{W}}{\partial \bar{z}_t} = \frac{1}{\sigma^2} \int p(z,t) z W(z) dz - \int p(z,t) \bar{z} W(z) dz \quad (\text{A.8})$$

O fitness médio \bar{z} é uma constante que podemos retirar da integral, o que faz com que $\int p(z,t) \bar{z} W(z) dz$ seja igual a $\bar{z} \bar{W}$. Além disso, olhando a equação vemos que a expressão $\int p(z,t) z W(z) dz$ é igual a $\bar{z}_{W(t)} \bar{W}$. Desta forma, a equação de como o fitness médio muda em relação a mudanças na média do caráter é igual a:

$$\frac{\partial \bar{W}}{\partial \bar{z}_t} = \frac{\bar{W}}{\sigma^2} (\bar{z}_{Wt} - \bar{z}_t) \quad (\text{A.9})$$

A qual representa a forma final da equação 7. A partir desta equação podemos ver que $(\bar{z}_{Wt} - \bar{z}_t) = \frac{\partial \bar{W}}{\partial \bar{z}_t} \frac{\bar{W}}{\sigma^2}$. E assim essa relação nos permite expressar a mudança observada na equação 2 como:

$$\Delta \bar{z} = \frac{h^2 \sigma^2}{\bar{W}} \frac{\partial \bar{W}}{\partial \bar{z}_t} = h^2 \sigma^2 \frac{\partial \ln(\bar{W})}{\partial \bar{z}_t} \quad (\text{A.10})$$