

## PRINCÍPIOS DE GENÉTICA DE POPULAÇÕES

### META

Mostrar ao aluno que a dinâmica dos genes nas populações pode ser compreendida por meio do estudo das suas frequências gênicas e genotípicas. Pretende-se fornecer os requisitos básicos para que o aluno possa entender os mecanismos relacionados a variação genética das populações biológicas ao longo do espaço e do tempo. Discutiremos os mecanismos básicos sobre os quais agem os fatores evolutivos.

### OBJETIVOS

Ao final desta aula, o aluno deverá:

saber identificar os principais mecanismos que afetam as frequências gênicas das populações no espaço e no tempo. Deverá saber calcular as frequências gênicas e genotípicas e entender a dinâmica dessas frequências e suas implicações numa população biológica.

### PRÉ-REQUISITOS

Noções de genética mendeliana, Binômio de Newton e noções de álgebra.

### INTRODUÇÃO

Prezado aluno, até agora os assuntos tratados em genética diziam respeito às características individuais e familiares e as análises utilizavam heredogramas e cruzamentos entre indivíduos. Nesse momento temos um novo desafio, como estudar uma população? Seria possível entender a dinâmica dos genes e dos genótipos em escala populacional? A resposta é positiva, mas muitos estudantes de biologia têm imensa dificuldade em entender os princípios básicos de genética de populações, principalmente em função dos cálculos envolvidos e por ser uma área da genética pouco trabalhada no ensino médio. No entanto, é possível contornar essa dificuldade inicial por meio de ilustrações e simulações disponíveis na internet. Nesse ponto o ensino a distância é facilitado pelo desenvolvimento da informática e da disponibilidade de softwares e sites especializados nessa área que serão comentados ao longo desse capítulo.

Uma população natural é formada por um conjunto de indivíduos de uma espécie contendo diversas famílias com muitos genótipos diferentes. Desse modo, a estrutura da população é definida pela frequência dos alelos que compõem os diferentes genótipos das diferentes famílias. Uma população natural é formada por todos os indivíduos que, ao se reproduzirem uns com os outros, compartilham de um pool gênico, que é formado pelo conjunto total de informações genéticas compartilhadas pelos membros reprodutivos da população.

Iniciaremos a aula descrevendo a variação genética populacional, o que determina essa variação genética e quais suas implicações. Depois iremos entender como seria a dinâmica dos genes e dos genótipos numa população ideal – População em equilíbrio de Hardy & Weinberg – Tal qual o zero na matemática, ou um ponto de referência na física, essa população ideal fornece um parâmetro básico para que possamos entender o perfil genético das populações e verificar como cada fator evolutivo pode interferir na variação populacional ao longo do tempo. A genética de populações estuda a origem da variação, a transmissão dos genes ao longo das gerações e a dinâmica da mudança na composição genética das populações ao longo do tempo por ação das forças evolutivas sistemáticas (determinísticas) e aleatórias (estocásticas).

### DEFINIÇÃO DE POPULAÇÃO

A população é formada por um grupo de indivíduos pertencentes a uma mesma espécie biológica. Em genética de populações, entretanto, é preciso especificar melhor este conceito. A palavra população não se refere a espécie como um todo, mas apenas a um grupo de organismos de uma mesma espécie que vive em uma determinada área geográfica suficientemente restrita para que qualquer membro possa se casar com qualquer outro do sexo oposto.

## VARIAÇÃO GENÉTICA POPULACIONAL

A variação é uma propriedade natural das populações biológicas. No entanto, essa concepção nem sempre foi assim. Antigamente pensava-se que cada espécie era formada por uma essência perfeita e que as diferenças entre os indivíduos de uma mesma espécie eram apenas erros, uma vez que a essência pertencia a um mundo perfeito, que seria o mundo das ideias para uns ou o mundo divino para outros. Esse pensamento que colocava a variação como erro persistiu por cerca de 2000 anos, desde Platão (na Grécia antiga) até meados do século XIX. Felizmente Charles Darwin revolucionou o mundo científico quando a defendeu que a variação é uma propriedade natural das espécies e não defeitos que afastam os meros mortais da essência perfeita. O reconhecimento da amplitude da variação fenotípica e de sua base hereditária levou Darwin à ideia da evolução por meio da Seleção Natural.

Para melhor compreender isso, vamos observar as pessoas ao nosso redor: parentes, vizinhos, conhecidos do bairro ou de outras regiões. Observe que todos variam entre si em relação à cor dos olhos, peso, cor da pele, altura, características faciais, tipos sanguíneos. É fascinante pensar que essa variação nos garante que nenhum indivíduo, excetuando os gêmeos univitelinos, é geneticamente igual ao outro e que não exista, dentre os mais de 6,5 bilhões de habitantes do nosso planeta, outro indivíduo igual a você e mais, que provavelmente nunca tenha existido outro e que nunca existirá outro igual a você no futuro. Portanto, a variação é uma realidade muito presente, não só na espécie humana, mas na maioria dos seres vivos. Por exemplo, as abelhas variam em tamanho, forma, padrões de coloração; camundongos variam em cor da pelagem; caramujos apresentam conchas com diferentes padrões de cores e listas e as plantas variam em relação à susceptibilidade diferencial a parasitas.

A nossa observação das espécies em nível fenotípico dá uma dimensão prática da existência de variação. No entanto, a variação genética é ainda maior que aquela que pode ser observada na comparação dos fenótipos. Ao nível molecular a variação existente é ainda maior, um fato comprovado com o avanço das técnicas moleculares desenvolvidas a partir da segunda metade do século XX e ainda em desenvolvimento. Uma boa parte da variação genética molecular não aparece no fenótipo em função de estar em regiões neutras do DNA, em outras áreas não codificantes ou em função da redundância do código genético que possibilita que códons diferentes decodifiquem um mesmo aminoácido. Desse modo, dois indivíduos podem produzir uma mesma proteína mesmo que suas sequências de DNA sejam diferentes.

A quantidade de variação genética dentro de populações naturais e a dinâmica das forças limitantes e/ou moldadoras ao longo do tempo e do espaço formam a base de interesse do geneticista de populações. A vari-

abilidade genética, portanto, está diretamente relacionada à evolução das espécies e o grau de variabilidade dentro da espécie afeta diretamente o seu potencial de se adaptar a mudanças ambientais.

### O SURGIMENTO DA GENÉTICA DE POPULAÇÕES

Em 1900, com o redescobrimto dos trabalhos de Mendel, ficou cada vez mais claro que as características são determinadas por genes, que se segregam durante a formação de gametas que são transmitidos para a prole. A análise da transmissão genética em cruzamentos experimentais e heredogramas rapidamente deu origem a um novo tipo de análise que envolvia populações inteiras. Em 1930 a genética de populações já estava se estabelecendo como disciplina, contando com a grande contribuição de pesquisadores como Sewall Wright, R.A. Fisher e J.B.S. Haldane, passando a representar um alicerce fundamental para a teoria evolutiva e para estudos de variação genética populacional na área de ecologia. Não podemos negligenciar a influência do contínuo avanço tecnológico da biologia molecular, a partir da década de 1950 até os dias atuais, que permitiu a decifração de genomas inteiros, bem como com os avanços da bioinformática, no desenvolvimento da genética de populações, que se tornou uma disciplina com ampla sustentação teórica e uma enorme variedade de ferramentas práticas que permitem analisar as características genéticas das populações tanto em nível fenotípico, quanto ao nível molecular.

### FREQUÊNCIAS DE ALELOS E GENÓTIPOS

A princípio os termos frequências gênicas e frequências genotípicas parecem se tratar da mesma coisa, mas isso não é verdade, sendo de fundamental importância que o estudante saiba facilmente diferenciá-los.

Tomemos como base um locus qualquer, por exemplo: o locus A. Considerando que esse locus tenha apenas dois alelos diferentes (A e a), podemos concluir que os indivíduos diploides só poderão apresentar os genótipos AA, Aa e aa para esse locus.

Podemos concluir então que para se calcular a frequência genotípica, simplesmente somamos o número de indivíduos de um determinado genótipo na população e dividimos pelo número total de indivíduos (N). Então para o locus com três genótipos mencionado (AA, Aa e aa), a frequência (f) de cada genótipo é:

$$f(AA) = \frac{\text{número de indivíduos } AA}{N}$$

$$f(Aa) = \frac{\text{número de indivíduos } Aa}{N}$$

$$f(aa) = \frac{\text{número de indivíduos } aa}{N}$$

Dizer 0,5 ou  $\frac{1}{2}$  ou 50% é exatamente a mesma coisa, sendo apenas a escala diferente. 0,5 representa a metade do conjunto inteiro igual a 1, enquanto  $\frac{1}{2}$  é o mesmo que a metade de um conjunto inteiro de 2, ou seja, uma parte num total de duas partes. Do mesmo modo, utilizando uma escala de 100 partes (percentagem), 50% representa a metade de 100. É, portanto, apenas uma questão de escala.

Quando estamos falando de uma população de 100 indivíduos e verificamos que a metade deles é heterozigota para o locus A, ou seja, que 50 indivíduos sejam Aa, podemos dizer que a frequência genotípica dos heterozigotos é:

$$f(Aa) = \frac{50}{100}$$

$$Aa = 0,5 \text{ ou } \frac{1}{2} \text{ ou } 50\%$$

Considerando ainda que nessa população, 25 indivíduos sejam de genótipo aa e outros 25 de genótipo AA, podemos igualmente concluir que a frequência genotípica dos homozigotos recessivos é:

$$f(aa) = \frac{25}{100}$$

$$aa = 0,25$$

e, igualmente:

$$f(AA) = \frac{25}{100}$$

$$AA = 0,25$$

Observe que se somarmos as frequências de todos os genótipos teremos:

$$AA + Aa + aa = 1$$

ou seja, esses genótipos juntos englobam todos os indivíduos dessa população (100%). Podemos então concluir que as frequências genotípicas estão relacionadas à forma em que os alelos estão distribuídos aos pares nos indivíduos que formam a população.

Utilizando o mesmo exemplo anterior, então como seriam as frequências alélicas (também frequentemente chamadas de frequências gênicas)? Caso não tenhamos a necessidade de saber como os alelos estão distribuídos aos pares nos indivíduos, mas apenas querendo saber quais as suas frequências populacionais, podemos dizer que nessa população de 100 indivíduos, a soma de todos os alelos A e alelos a representam o conjunto de todos os alelos da população para esse locus. Então, podemos definir frequência alélica como a proporção de cópias gênicas de um determinado tipo alélico numa determinada população. A frequência alélica pode variar entre zero (0) e um (1), sendo zero quando o alelo está ausente e, em outro extremo, a frequência alélica é 1 quando o locus é monomórfico, ou seja, quando existe apenas um tipo alélico para o locus em questão na população.

Considerando que cada locus possui uma representação específica de seus alelos, convencionou-se representar as frequências de alelos pelos símbolos p e q (substituindo A e a, no nosso exemplo). As frequências dos alelos podem ser calculadas pelas seguintes fórmulas gerais:

$$p = f(A) = \frac{2n_{AA} + n_{Aa}}{2N}$$

$$q = f(a) = \frac{2n_{aa} + n_{Aa}}{2N}$$

Em que  $n_{AA}$ ,  $n_{Aa}$  e  $n_{aa}$  representam o número de indivíduos para cada um dos genótipos (AA, Aa e aa) e N representa o número total de indivíduos de todos os genótipos. Divide-se por 2N porque cada indivíduo diploide possui dois alelos em um locus (exceto os hemizigotos). Como apresentado anteriormente, a soma de todos os alelos é sempre igual a 1, ou seja,  $p+q=1$ . Considerando um locus com dois alelos, uma vez que tenhamos obtido p, podemos obter q por subtração:  $q=1-p$

Voltando ao exemplo inicial em que as frequências alélicas de A e a foram calculadas a partir das frequências genotípicas, utilizando a fórmula geral temos:

$$p = f(A) = f(AA) + \frac{1}{2}f(Aa)$$

$$q = f(a) = f(aa) + \frac{1}{2}f(Aa)$$

Considerando as frequências genóticas exemplificadas acima ( $AA=0,25$ ;  $Aa=0,5$ ;  $aa=0,25$ ), podemos dizer que a frequência alélica de  $A$  nessa população é igual ao total dos homozigotos somado à metade dos heterozigotos, logo:

Utilizando o número de indivíduos, temos:

$$p = f(A) = \frac{2n_{AA} + n_{Aa}}{2N} \rightarrow p = f(A) = \frac{2(25) + 50}{2(100)} = 0,5$$

Ou, alternativamente, podemos utilizaras frequências genóticas para chegarmos ao mesmo resultado:

$$p = f(A) = f(AA) + \frac{1}{2}f(Aa) \rightarrow p = f(A) = (0,25) + \frac{1}{2}(0,5) = 0,5$$

Podemos dizer então que nessa população a frequência alélica de  $e$ , por correspondência a frequência alélica do alelo  $a$  pode ser obtida por diferença, considerando que:

Se

$$p+q=1 \rightarrow A+a=1$$

Então sendo

$$p = f(A) = 0,5$$

$$f(a) = 1 - p \rightarrow f(a) = 1 - 0,5 = 0,5$$

Dessa forma também podemos descrever o pool gênico de uma população por meio de suas frequências alélicas. Lembrando que em uma população biológica de reprodução sexuada, os genótipos são apenas reuniões temporárias de alelos nos indivíduos, que se desfazem a cada geração quando os alelos individuais são passados aos descendentes através dos gametas. Portanto, são os tipos de alelos, e não os genótipos, que têm uma real continuidade de uma geração para a outra, e que constituem o conjunto gênico de uma população.

## O PRINCÍPIO DE HARDY-WEINBERG

Há mais de 100 anos atrás, em 1908, G. H. Hardy, um matemático inglês e W. Weinberg, um médico alemão, desenvolveram independentemente um conceito matemático relativamente simples, hoje denominado princípio de Hardy-Weinberg, considerado o fundamento básico da genética de populações. Esse princípio diz que numa população panmítica (de reprodução ao acaso), infinita e na ausência de fatores evolutivos, como o fluxo gênico, mutação, seleção e deriva genética, as frequências alélicas e genotípicas permanecem constantes ao longo do tempo. O princípio de Hardy-Weinberg pode ser utilizado para determinar a frequência de cada alelo de um par ou de uma série, bem como frequências genotípicas de homocigotos e heterocigotos na população.

Vamos considerar novamente o nosso conhecido locus A, com os genótipos AA, Aa e aa, que serão aqui tratados como, respectivamente D, H e R, uma alusão para genótipo dominante, heterocigoto e recessivo, mesmo considerando que nem todos os loci apresentam essa forma de dominância. Do mesmo modo, as frequências alélicas de A e a serão tratadas como p e q. Essas letras (D, H e R para genótipos e p e q para alelos) são utilizadas pelos livros convencionalmente para que seja possível desenvolver fórmulas gerais que possam ser utilizadas para qualquer locus com dois alelos.

Hardy & Weinberg perceberam que a distribuição dos alelos numa população, considerando um locus uma população mendeliana, pode ser ajustada perfeitamente ao binômio de Newton, estudado por todos nós ainda no ensino fundamental, quem se lembra de  $(p+q)^2$ ? Então, pelos produtos notáveis sabemos que o cálculo fica simples, apenas lembrando “o quadrado do primeiro mais duas vezes o primeiro pelo segundo mais o quadrado do segundo” =  $(p+q)^2 = p^2 + 2pq + q^2$ , mas como reunir isso que eu aprendi em matemática à genética de populações?

	<b>A</b> p	<b>a</b> q	
<b>A</b> p	AA $p \times p = p^2$	Aa $q \times p = pq$	$f(A) = p$ $f(a) = q$
<b>a</b> q	Aa $p \times q = pq$	aa $q \times q = q^2$	
			$f(AA) = p^2$ $f(Aa) = 2pq$ $f(aa) = q^2$

Figura 1: Binômio de Newton ajustado ao quadrado de Punnett.

É surpreendente o quanto essa distribuição binomial se ajusta perfeitamente à distribuição dos alelos e dos genótipos numa população, vejamos: um locus autossômico em um indivíduo diploide apresenta dois alelos, que podem ser iguais (quando o indivíduo é homocigoto) ou diferentes (nos heterocigotos). De qualquer modo, eles estão aos pares e durante a formação dos gametas eles se separam e são distribuídos de forma independente para os descendentes, então voltando ao nosso velho exemplo, o locus A, numa população onde esse locus apresenta dois alelos (A+a). A distribuição desses alelos se dá aos pares, uma vez que os indivíduos são diploides, logo  $(A+a)^2$ ; Os indivíduos, por sua vez, podem ser homocigotos dominantes (AA), heterocigotos (Aa) e homocigotos recessivos (aa), os quais podemos chamar, respectivamente de D, H e R.

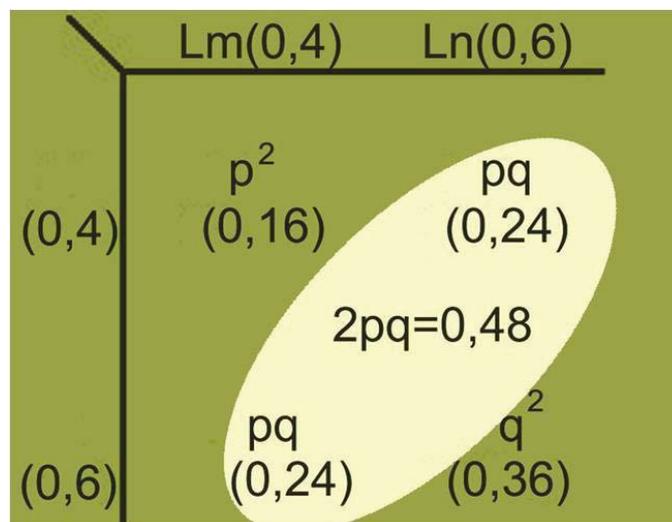


Figura 2: Cálculo das frequências genotípicas para o locus A.

Suponha que cada genótipo seja igualmente representado por machos e fêmeas, e que os cruzamentos ocorram inteiramente ao acaso na população. Nesse caso não estamos mais preocupados em saber quais indivíduos estão sendo cruzados, como fazíamos nos cruzamentos encontrados nos exercícios de genética mendeliana elementar, uma vez que o foco é populacional, queremos apenas estimar quais as frequências genotípicas populacionais baseados nas frequências gênicas, ou vice-versa, quais são as frequências gênicas da população quando já sabemos a priori suas frequências genotípicas.

Para compreender melhor esse ajuste entre a distribuição binomial e a genética de populações, vamos utilizar uma população humana hipotética de 1000 indivíduos, igualmente distribuídos entre homens e mulheres. Utilizaremos o grupo sanguíneo MN, que é controlado por dois alelos, Lm e Ln. Cada alelo codifica uma molécula de polissacarídeo na superfície das hemácias e pode ser distinguido por meio de reagentes químicos apropriados. Os tipos de moléculas são designados M e N respectivamente e os

alelos  $L^mL^n$  são codominantes, o que quer dizer que se expressam nos heterozigotos. Um exame para o grupo sanguíneo MN constatou que nessa população a frequência do alelo  $L^m$  é 0,4. Utilizando o que já aprendemos anteriormente para calcular frequências alélicas, podemos deduzir facilmente que a frequência do alelo  $L^n$  pode ser obtida por diferença, já que  $L^m + L^n = 1$  ( $p + q = 1$ ), então qual seria o número esperado de indivíduos do tipo M (genótipo  $L^mL^m$ ), MN (genótipo  $L^mL^n$ ) e do tipo N (genótipo  $L^nL^n$ ) nessa população? Considerando que essa população está em equilíbrio de Hardy-Weinberg, podemos calcular facilmente esses valores utilizando o binômio de Newton  $(p+q)^2$ , vejamos:

Se

$$p = f(L^m) = 0,4$$

e

$$q = f(L^n) = 1 - p = 1 - 0,4 = 0,6$$

Logo:

$$D = f(L^mL^m) = p^2$$

$$D = 0,4^2 = 0,16$$

Considerando que são 1000 indivíduos e que 0,16 são do tipo M (homozigotos, D), então basta multiplicar o número total de indivíduos pela sua frequência:

$$D = f(L^mL^m) = 0,16 \times 1000 = 160 \text{ Indivíduos}$$

Para calcular o número estimado de heterozigotos procedemos assim:

Se

$$H = f(L^mL^n) = 2pq$$

Então:

$$H = f(L^mL^n) = 2 \times 0,4 \times 0,6 = 0,48$$

Considerando que são 1000 indivíduos e que 0,48 são do tipo MN (heterozigotos, H), então basta multiplicar o número total de indivíduos pela frequência de heterozigotos:

$$D = f(L^m L^n) = 0,48 \times 1000 = 480 \text{ indivíduos}$$

Finalmente, para calcular o número esperado de indivíduos homozigotos do tipo N, basta calcular:

$$R = f(L^n L^n) = q^2$$

$$R = f(L^n L^n) = 0,6^2 = 0,36$$

Considerando que são 1000 indivíduos e que 0,48 são do tipo MN (heterozigotos, H), então basta multiplicar o número total de indivíduos pela respectiva frequência:

$$R = f(L^n L^n) = 0,36 \times 1000 = 360 \text{ indivíduos}$$

Portanto:

$$p^2 + 2pq + q^2 = 1, \text{ assim como } D + H + R = 1$$

Nesse exemplo:

$$160M + 480MN + 360N = 1000 \text{ indivíduos}$$

Também podemos calcular as frequências gênicas quando temos apenas as frequências genotípicas, utilizando o inverso do raciocínio anterior. Utilizando os resultados da questão anterior, vejamos como conseguimos chegar às frequências alélicas:

Se considerarmos a frequência do alelo  $L^m$  como  $p$  e a frequência de  $L^n$  como  $q$ , então podemos estimar as frequências alélicas da seguinte forma:

Fenótipo	N	$f$ Genotípica
M	160	0,16
MN	480	0,48
N	360	0,36
Total	1000	1

$$f(p) = f(L^m) = D + \frac{1}{2}(H) \rightarrow 0,16 + \frac{1}{2}(0,48) = 0,4$$

$$f(q) = f(L^n) = R + \frac{1}{2}(H) \rightarrow 0,36 + \frac{1}{2}(0,48) = 0,6$$

É fácil perceber, portanto, que existe uma relação direta entre as frequências alélicas e genotípicas. De tal modo que sabendo as frequências genotípicas é possível estimar as frequências alélicas e vice-versa, desde que a população considerada esteja em equilíbrio de Hardy-Weinberg.

No entanto, algumas vezes não temos como saber a priori as frequências de todos os genótipos, como é o caso onde existe relação de dominância entre os alelos em que não é possível diferenciar homozigotos dominantes de heterozigotos. Nesse caso, temos como saber estimar esses valores partindo da frequência do homozigoto recessivo. Para exemplificar essa situação, vamos imaginar uma população em equilíbrio em que a frequência de indivíduos albinos seja de 0,1 ( $aa=0,1$ ). Nesse caso quais seriam as frequências alélicas de  $a$  e  $A$ ? e quais seriam as frequências genotípicas de  $AA$  e  $Aa$ ?

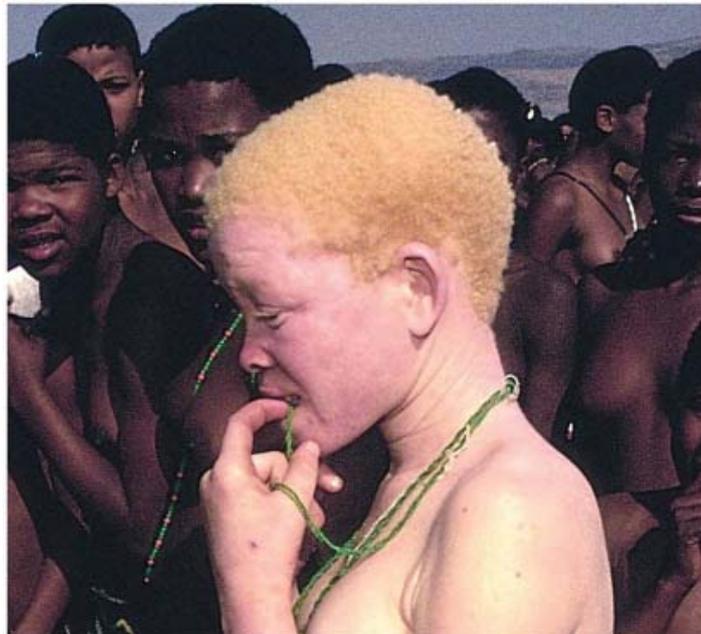


Figura 3: Indivíduo albino em uma população.  
(Fonte: [http://www.mun.ca/biology/scarr/Albinism\\_in\\_Humans.html](http://www.mun.ca/biology/scarr/Albinism_in_Humans.html)).

Se

$$aa = 0,1 \rightarrow q^2 = 0,1 \rightarrow R = 0,1$$

Então, se sabemos a frequência genotípica do homozigoto recessivo e sabendo que essa frequência é igual a  $q^2$ , basta raciocinar: “como obter  $q$  a partir de  $q^2$ ?”. Caso você tenha pensado na raiz quadrada, é isso mesmo, então vejamos:

$$q^2 = 0,1$$

$$\sqrt{q^2} = \sqrt{0,1}$$

$$q = 0,01$$

Agora basta seguir o mesmo princípio utilizado anteriormente:

Se

$$q=001$$

Então

$$p = 1 - q \rightarrow p = 1 - 0,01 = 0,99$$

Sabendo as frequências alélicas podemos calcular facilmente as frequências genotípicas esperadas pelo equilíbrio de Hardy-Weinberg:

$$D = f(AA) = p^2 = (0,99)^2 = 0,9801$$

$$H = f(Aa) = 2pq = (2 \times 0,99 \times 0,01) = 0,0198$$

$$R = f(aa) = q^2 = (0,01)^2 = 0,0001$$

Esses resultados indicam que se essa população tivesse 10.000 indivíduos, poderíamos estimar, multiplicando as frequências genotípicas calculadas pelo total de indivíduos, que de um indivíduo seria albino (aa), 198 indivíduos seriam heterozigotos para albinismo (Aa) e 9801 indivíduos seriam homozigotos normais (AA).

A figura a seguir ilustra a relação entre frequências gênicas e genotípicas para um locus autossômico com dois alelos em populações diploides;

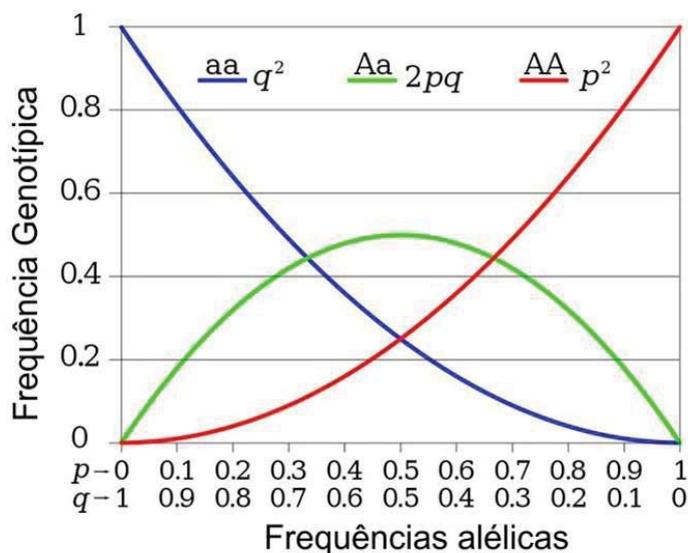


Figura 4: Equilíbrio de Hardy-Weinberg

Note que quando a frequência alélica no eixo x atinge 0,5 a frequência de genótipos heterozigotos atinge o ponto máximo. Portanto, populações em equilíbrio nunca terão taxas de heterozigotos significativamente acima de 50%. Para as frequências de homozigotos (AA e aa) veja que a distribuição é inversamente proporcional. Quando a frequência de alelos A atinge seu nível máximo (1,0) a frequência de alelos recessivos é zero. Nesse caso houve a fixação de alelos A e todos os indivíduos da população teriam genótipo AA. No outro extremo, temos a situação inversa com a fixação do alelo recessivo em que todos os indivíduos da população seriam de genótipo aa. Entre esses dois extremos existem todos os valores possíveis de frequências alélicas, sempre obedecendo a equação  $p+q=1$ .

Uma grande parte das populações biológicas, as frequências genotípicas na maioria dos loci se adequam muito bem na distribuição de Hardy-Weinberg. No entanto, o princípio de Hardy-Weinberg apresenta diversos pressupostos para que seja válido:

### **PRESSUPOSTOS PARA O PRINCÍPIO DE HARDY-WEINBERG**

#### 1. População Panmítica

São populações de acasalamento ao acaso, ou seja, nesse caso a probabilidade de um determinado alelo participar do processo reprodutivo é uma questão inerente apenas à sua frequência populacional. Um dos fatores que alteram esse pressuposto é a endogamia, que está associada ao acasalamento preferencial entre indivíduos aparentados.

#### 2. População infinita

É impossível que uma população biológica seja efetivamente infinita, mas populações muito grandes se comportam praticamente como se fossem infinitas. No entanto, quanto menor for o tamanho efetivo da população, maior serão os desvios ao acaso em suas frequências gênicas e genotípicas ao longo do tempo. Esse efeito é consequência da deriva genética, que causa variações ao acaso (estocásticas) em populações finitas, sendo tanto mais forte sua influência quanto menor for a população;

#### 3. Que não exista fluxo gênico entre populações

O equilíbrio de Hardy-Weinberg pressupõe que as populações sejam fechadas, ou seja, que não haja intercâmbio de genes em função da migração. A saída e/ou entrada de indivíduos, portanto, pode afetar as frequências genéticas populacionais.

#### 4. Genes não mudam de um estado alélico para outro

Essa alteração é chamada de mutação e está associada a entrada de variação genética na população, que por consequência alteraria as suas frequências genéticas;

5. Todos os indivíduos possuem probabilidades iguais de sobrevivência e/ou reprodução

Esse pressuposto diz respeito à completa ausência de qualquer tipo de seleção. Caso haja diferença de probabilidade de sobrevivência e/ou reprodução as frequências dos genótipos e dos alelos podem se alterar ao longo das gerações.

Caso todos os pressupostos mencionados anteriormente sejam atendidos, as frequências gênicas e genotípicas populacionais permanecerão constantes ao longo do tempo, como afirma o princípio de Hardy-Weinberg.

Imagino que você deva estar pensando a essa altura que esse princípio de Hardy-Weinberg é extremamente teórico e não possui qualquer aplicação prática. Entretanto, esse princípio é verdadeiramente o principal alicerce para o estudo da genética de populações. Isso é possível porque o princípio de Hardy-Weinberg permite gerar parâmetros genéticos básicos de uma população. Numa população ideal que atenda a todos os pressupostos mencionados não existe qualquer influência da evolução e, por isso, as condições genéticas da população continuariam constantes ao longo do tempo. Por se tratar de uma super simplificação da realidade, o Princípio de Hardy-Weinberg pode ser considerado como um modelo de referência para uma população ideal e quando verificamos se uma população biológica encontra-se ou não em equilíbrio, isso nos dá uma idéia clara sobre o grau de influência do processo evolutivo sobre essa população. Portanto, a distância de uma população ideal em relação a uma população biológica formada pelos ingredientes da evolução. É importante ressaltar que as suposições do equilíbrio de Hardy-Weinberg se aplicam a apenas um locus. Nenhuma população natural se reproduz aleatoriamente para todas as características; nem a população está livre dos fatores evolutivos para todas as características. O princípio de Hardy-Weinberg, portanto não exige que a população seja panmítica, ausência de seleção, migração, deriva genética e mutações para todos os loci. Ela exige essa condição apenas para o locus que está sendo considerado na análise. Uma população pode estar em equilíbrio para um locus e não para outros.

## **AVALIAÇÃO DO EQUILÍBRIO POPULACIONAL**

Utilizando as informações anteriores é possível agora avançar para as populações naturais e verificar se tais populações estariam ou não em equilíbrio de Hardy-Weinberg. Com as frequências alélicas de uma população real construímos uma população teórica e comparamos as duas distribuições de frequências genotípicas. Esse passo é fundamental porque a distribuição teórica tem propriedades matemáticas que nos possibilitam tirar conclusões de grande importância. Se ao analisarmos a população real se apresentar estatisticamente semelhante à população teórica, poderemos estender as

conclusões para a população real e compreender seu comportamento genético. Para compararmos duas populações iniciamos verificando suas diferenças quanto as frequências genotípicas.

Utilizando novamente como exemplo os grupos sanguíneos do sistema MN nos seres humanos, vamos considerar os dados de Race e Sanger (1975) que realizaram amostras de sangue em 1000 ingleses e encontraram: 298 M, 489 MN e 213 N. Essa população estaria em equilíbrio de Hardy-Weinberg? Para responder a essa pergunta precisamos comparar os dados dessa população real com os valores que seriam esperados em uma população ideal, então vejamos:

Fenótipos	Genótipos	Número observado de indivíduos na população real	Frequências genotípicas esperadas
M	$L^mL^m$	298	$D=p^2$
MN	$L^mL^n$	489	$H=2pq$
N	$L^nL^n$	213	$R=q^2$

A tabela acima resume as informações necessárias para a comparação entre a população real e a população ideal. A estatística apropriada para essa comparação é o teste de qui-quadrado, que permite avaliar os desvios entre frequências observadas ( $f_o$ ) e frequências esperadas ( $f_e$ ). A frequência esperada é a distribuição de frequências genotípicas da população teórica e frequência observada é a distribuição de frequências genotípicas da população real. A fórmula do qui-quadrado é a seguinte:

$$\chi^2 = \sum \frac{(f_o - f_e)^2}{f_e}$$

Para calcular fe basta voltar ao conteúdo anterior que nos ensina a calcular as frequências alélicas:

$$p = f(L^m) = \frac{2n_M + n_{MN}}{2N}$$

$$p = f(L^m) = \frac{2(298) + 489}{2(1000)} = 0,543$$

$$q = f(L^n) = 1 - p = 1 - 0,543 = 0,458$$

Agora sabendo as frequências alélicas, podemos estimar as frequências esperadas de indivíduos do tipo M:

$$f_e(M) = p^2 + \frac{1}{2}(2pq) \cdot (N_t)$$

Onde  $f_e(M)$  é a frequência esperada de indivíduos de fenótipo M e  $N_t$  é o número total de indivíduos, então:

$$f_e(M) = (0,543)^2 + \frac{1}{2}(2 \times 0,543 \times 0,458) \times (1000) = 294,3$$

Para a frequências esperada de indivíduos MN temos:

$$f_e(MN) = (2pq) \cdot (N_t)$$

Logo:

$$f_e(MN) = (2 \times 0,543 \times 0,458) \times (1000) = 496,4$$

Finalmente, para a frequência esperada de N, temos:

$$f_e(N) = q^2 + \frac{1}{2}(2pq) \cdot (N_t)$$

Então:

$$f_e(N) = (0,458)^2 + \frac{1}{2}(2 \times 0,543 \times 0,458) \times (1000) = 209,3$$

Então, resumindo a tabela anterior:

Aplicando a fórmula de qui-quadrado teremos:

Fenótipo	$f_o$	$f_e$	$f_e$ Indiv.
M	298	$p^2$	294,3
MN	489	$2pq$	496,4
N	213	$q^2$	209,3
Total	1000	1	1000

$$\chi^2 = \sum \frac{(f_o - f_e)^2}{f_e}$$

$$\chi^2 = \frac{(298 - 294,3)^2}{294,3} + \frac{(489 - 496,4)^2}{496,4} + \frac{(213 - 209,3)^2}{209,3} = 0,221$$

Os resultados desta fórmula são comparados com os resultados de uma tabela de qui-quadrado de referência, levando-se em conta a margem de erro considerada no cálculo estatístico (geralmente  $\alpha = 0,05$  e o número de graus de liberdade, que para populações é igual ao número de classes fenotípicas menos 2 ( $GL=NC-2$ ). Se o valor do qui-quadrado calculado for menor que o valor correspondente de qui-quadrado da tabela podemos dizer que as diferenças entre as duas distribuições não são significativas e podemos considerar que a população real comporta-se como uma população teórica em equilíbrio. O valor de qui-quadrado de referência, na tabela de qui-quadrado nesse caso é 3,841. Podemos concluir, portanto, que a população estudada por Race e Sanger está em equilíbrio de Hardy-Weinberg para o locus MN.

Você deve estar se perguntando como é possível que um locus esteja em equilíbrio numa população humana, já que sabemos que as pessoas definitivamente não escolhem seus parceiros ao acaso. O que ocorre é que as pessoas de fato escolhem seus parceiros por diversos caracteres fenotípicos, com altura, peso, cor da pele, nível social, etc. No entanto, ninguém escolhe o parceiro pelo tipo sanguíneo. Com isso, pelo menos para o locus MN os acasalamentos são de fato aleatórios nessa população.

### CRUZAMENTOS NÃO ALEATÓRIOS

Existem diversos motivos pelos quais o princípio de HW pode ser afetado em uma população em particular. A não atenção a qualquer um dos pressupostos mencionados anteriormente podem influenciar na variação das frequências gênicas e genotípicas ao longo do tempo, como é o caso dos cruzamentos não aleatórios.

A relação entre as frequências alélicas e genotípicas se desfaz se o cruzamento entre os indivíduos da população não ocorre ao acaso. Portanto, um dos principais fatores que pode afetar o equilíbrio de HW ocorre quando a população não é panmítica. A endogamia e a reprodução preferencial são os principais fatores associados aos cruzamentos não aleatórios. A endogamia ocorre quando temos acasalamento entre indivíduos aparentados, o que aumenta a probabilidade de que os descendentes tenham dois alelos idênticos por descendência. A endogamia resulta no aumento da frequência de genótipos homocigotos e redução no número de heterocigotos em relação ao que seria esperado por acasalamentos aleatórios. O efeito da endogamia é tanto maior quando mais próximos geneticamente forem os indivíduos acasalados. Se em uma determinada população o acasalamento entre indivíduos aparentados for comum, isso resultará na probabilidade de encontrar um indivíduo que tenha dois alelos idênticos por descendência quando fazemos uma amostragem aleatória da população. Esta probabilidade é chamada de  $F$ , ou coeficiente de endogamia. Se voltarmos o suficiente no tempo, muitos alelos provavelmente serão idênticos por descendência,

mas, para calcular os efeitos da endogamia consideramos a identidade por descendência levando em conta apenas algumas gerações. O ponto mais importante que deve ser observado é que a endogamia altera as frequências genotípicas, mas as frequências gênicas permanecem inalteradas. O coeficiente de endogamia ( $F$ ) pode variar de 0 a 1. Um valor de 0 indica que a reprodução em uma população é totalmente aleatória, enquanto um valor de 1 indica que todos os alelos são idênticos por descendência. Uma das formas mais fortes de endogamia ocorrem na auto-fecundação, fenômeno comum em plantas, que reduz em 50% o número de heterozigotos por geração, até que todos os genótipos na população sejam homozigotos. Para a maioria das espécies panmíticas, a endogamia é prejudicial porque aumenta a probabilidade de alelos deletérios e letais recessivos se manifestarem na população em função do aumento do número de homozigotos.

No caso dos cruzamentos preferenciais, eles ocorrem quando os membros da população preferem se reproduzir com parceiros fenotipicamente parecidos (cruzamento preferencial positivo), como exemplo, indivíduos de estatura alta procuram outros de mesma estatura para se relacionar e o mesmo pode ocorrer com a cor da pele e outros caracteres fenotípicos. Apesar de serem diferentes, a endogamia e os cruzamentos preferenciais positivos resultam no mesmo efeito qualitativo. Elas reduzem a frequência de heterozigotos e aumentam a frequência de homozigotos em comparação ao que seria esperado pelo equilíbrio de Hardy-Weinberg. No entanto, também existem cruzamentos preferenciais negativos onde existe uma tendência de reprodução entre indivíduos diferentes, o que pode aumentar a frequência de heterozigotos e reduzir o número de homozigotos.

## CONCLUSÃO

Conforme afirmado anteriormente, a evolução pode ser definida de forma simplificada como a alteração nas frequências genéticas em uma população ao longo do tempo. Para compreender um cenário onde as frequências gênicas mudam, é imprescindível que o estudante de biologia entenda um cenário um pouco mais simples, no qual as frequências genéticas simplesmente não mudam, sendo essa compreensão a principal contribuição dessa aula. A genética de populações forma o alicerce básico para o estudo da evolução e também vem sendo muitíssimo utilizada na área de ecologia e no estudo de populações humanas e estudos epidemiológicos. Diante do que foi visto nessa aula, você pode concluir que uma população pode ser caracterizada geneticamente por suas frequências gênicas e genotípicas e que tais frequências se ajustam bem ao modelo matemático do binômio de Newton por meio do princípio de Hardy-Weinberg, que por sua vez nos fornece um parâmetro básico para reconhecer o grau de influência dos fatores evolutivos numa população biológica.



### RESUMO

A genética de populações estuda a origem da variação, a transmissão dos genes ao longo das gerações e a dinâmica da mudança na composição genética das populações ao longo do tempo por ação das forças evolutivas. A variação é uma propriedade natural das populações biológicas. O reconhecimento da amplitude da variação fenotípica e de sua base hereditária levou Darwin à idéia da evolução por meio da Seleção Natural. A nossa observação das espécies em nível fenotípico dá uma dimensão prática da existência de variação. No entanto, a variação genética é ainda maior que aquela que pode ser observada na comparação dos fenótipos. Ao nível molecular a variação existente é ainda maior. A quantidade de variação genética dentro de populações naturais e a dinâmica das forças limitantes e/ou moldadoras ao longo do tempo e do espaço formam a base de interesse do geneticista de populações. A variabilidade genética, portanto, está diretamente relacionada à evolução das espécies e o grau de variabilidade dentro da espécie afeta diretamente o seu potencial de se adaptar a mudanças ambientais. As frequências genotípicas populacionais estão relacionadas à forma em que os alelos estão distribuídos aos pares nos indivíduos que formam a população, enquanto a frequência alélica está relacionada à proporção de cópias gênicas de um determinado tipo alélico numa determinada população. em 1908, Hardy e Weinberg desenvolveram, independentemente, um conceito matemático relativamente simples, hoje denominado princípio de Hardy-Weinberg, considerado o fundamento básico da genética de populações. Esse princípio diz que numa população de reprodução ao acaso, infinita e na ausência de fatores evolutivos, as frequências alélicas e genotípicas permanecem constantes ao longo do tempo. O princípio de Hardy-Weinberg pode ser utilizado para determinar a frequência de cada alelo de um par ou de uma série, bem como frequências genotípicas de homocigotos e heterocigotos na população.

### ATIVIDADES



1. Considere essas 5 populações. Calcule as frequências alélicas de cada uma delas e verifique quais estão em equilíbrio de Hardy-Weinberg utilizando um teste de qui-quadrado.

	AA	Aa	aa
a. População 1	320	480	200
b. População 2	380	460	160
c. População 3	45	300	665
d. População 4	184	432	384
e. População 5	351	418	230

### COMENTÁRIO SOBRE AS ATIVIDADE

1. Para facilitar o seu trabalho, utilize o simulador offline gratuito que você poderá encontrar no site do Evotutor (nesse link: <http://www.evotutor.org>). Na página do Evotutor, que é um simulador para genética de populações e processos evolutivos, você pode escolher entrar em “Offlineapplication” para instalar o simulador no seu computador. Na página você escolherá o sistema operacional da sua máquina (provavelmente Windows) e fará o download do programa numa versão compactada (é bem pequeno, 1.3Mb). Basta descompactá-lo numa pasta do seu computador e executar o arquivo evotutor.bat. Quando abrir o programa entre em “calculators” na parte de cima da página e escolha o item “Chi-Square”, que é uma calculadora para o teste de qui-quadrado. Da mesma forma na guia “calculators” você pode entrar em HW-equilibrium que calcula as frequências esperadas para o equilíbrio de Hardy-Weinberg. Com essas ferramentas você fará todos os cálculos rapidamente, mas eu sugiro que você faça ao menos dois exercícios utilizando as fórmulas estudadas na aula para que você possa obter maior domínio sobre o assunto.

### AUTOAVALIAÇÃO

Após estudar esta aula, consigo:

1. Definir o que são frequências alélicas?
2. Definir o que são frequências genotípicas?
3. Definir quais são os pressupostos para que uma população esteja em equilíbrio de Hardy-Weinberg?
4. Explicar e Justificar ma população em equilíbrio pode ter uma frequência de heterozigotos maior que 50% no locus analisado? Justifique.



5. Explicar se uma população humana de uma determinada localidade possui 5000 habitantes e desses 10 são albinos (aa), quantos indivíduos heterozigotos para o albinismo (Aa) seriam esperados nessa população se considerarmos que esse locus esteja em equilíbrio de Hardy-Weinberg?
6. Definir o que é endogamia e como ela pode afetar as frequências genotípicas?
7. Definir o que são cruzamentos preferenciais?
8. Explicar quando consideramos que houve fixação de um alelo na população e qual a consequência direta disso?

### REFERÊNCIAS

- GRIFFITS, A. J. F. e Col – **Introdução à genética** – Editora Guanabara Koogan, 9a edição, Rio de Janeiro, RJ, 2008.
- SNUSTAD, D. P. – **Fundamentos à Genética** - Editora Guanabara Koogan, 6a edição, Rio de Janeiro, RJ, 2008.
- FUTUYMA, D. **Biologia evolutiva** – Editora Funpec, 3ª edição, Ribeirão Preto, SP, 2009