

## Endocruzamento: estimativas diretas e indiretas

### Meta da aula

Desenvolver o conceito de endocruzamento como força evolutiva em populações naturais.

Esperamos que, após o estudo do conteúdo desta aula, você seja capaz de:

- Definir endocruzamento, um caso especial de deriva.
- Analisar o conceito de identidade por descendência.
- Descrever e distinguir os coeficientes de endocruzamento.

### Pré-requisitos

Para acompanhar esta aula, é importante que você reveja os conceitos de heterozigidade e deriva gênica, Aulas 5 e 11.

## INTRODUÇÃO

### ENDOCRUZAMENTO

Em inglês, *inbreeding*, também denominado endogamia, são cruzamentos consanguíneos entre parentes que possuem um ancestral comum. Em populações humanas, estes tipos de casamentos são raros, restringidos por tradições culturais e religiosas.

Na Aula 11 desta disciplina, vimos o processo dispersivo que pode levar a uma alteração das frequências gênicas, conhecido como deriva genética, deriva gênica ou, simplesmente, deriva. Antes, abordamos o problema do ponto de vista da variação das frequências gênicas entre populações. Aqui, vamos abordar deriva sob um ponto de vista novo, isto é, como um processo de **ENDOCRUZAMENTO** decorrente do fato de as populações serem finitas em tamanho.

### VIOLAÇÃO DO PRESSUPOSTO DE CRUZAMENTO AO ACASO DO EHW: SISTEMAS DE ACASALAMENTO

O EHW aplica-se a um modelo de população idealizada (veja o verbete mais adiante nesta aula) que assume uma série de pressupostos como: tamanho populacional constante, a não-operação de qualquer outra força evolutiva e o acasalamento aleatório. De forma geral, nenhuma população real, seja natural ou de cativeiro/laboratório, satisfaz a todas as condições da população idealizada. Dessa forma, necessitamos trazer a teoria para o “mundo real”. Para isso, nesta aula, vamos ver o efeito de um sistema de acasalamento, o endocruzamento, em populações pequenas.

Demes (unidades básicas da alteração evolutiva, populações mendelianas; veja a Aula 5, disciplina Evolução) são caracterizadas, em parte, como populações cujos indivíduos compartilham um sistema comum de acasalamento. O modelo de Hardy-Weinberg assume um particular sistema desse tipo, denominado **acasalamento ao acaso**. Existem, porém, muitos **OUTROS SISTEMAS DE ACASALAMENTO**.



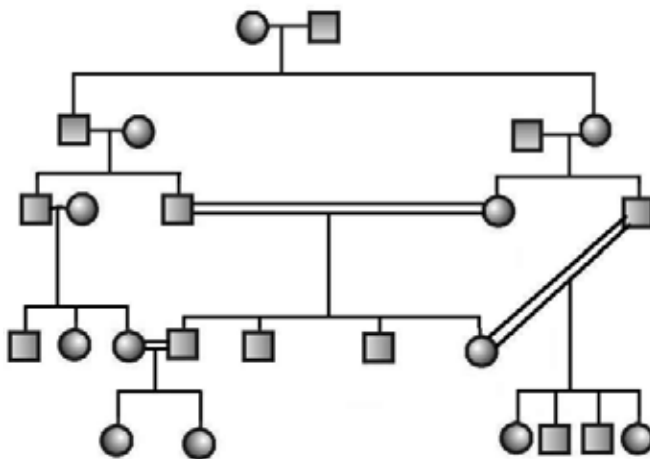
Você já reparou que utilizamos várias expressões para nos referirmos ao cruzamento ao acaso? Acasalamento aleatório, cruzamento randômico, contudo, todas as combinações entre essas palavras significam a mesma coisa: os casais não apresentam tendência na escolha do parceiro. Por exemplo, uma moça que só namora rapazes altos está sendo tendenciosa, escolhendo o parceiro pelo critério de altura. Um rapaz que namora moças gordinhas, morenas, loiras, de origem asiática etc., apesar de ser um desesperado por companhia, não apresenta tendência na escolha da parceira. Para ele, qualquer tipo físico serve para ser sua namorada.

### OUTROS SISTEMAS DE ACASALAMENTO

Acasalamento preferencial positivo: sob esse sistema de acasalamento, indivíduos com fenótipos similares são mais prováveis de se acasalarem entre si do que o esperado sob acasalamento aleatório (com respeito ao fenótipo) na população. Note a semelhança entre o acasalamento preferencial positivo e o endocruzamento, ambos os sistemas de acasalamento produzem um déficit de heterozigotos. A diferença é que endocruzamento se baseia na escolha de parceiros por intermédio de relacionamento genealógico, portanto, afetando o genoma como um todo! Por sua vez, acasalamento preferencial positivo (assim como o negativo) se baseia em uma escolha dos parceiros sexuais, a partir de algum fenótipo codificado por um único loco ou alguns poucos locos, mas não por todos os genes existentes em um genoma. Por exemplo, surdez nos humanos pode ser causada por homozigose para alelos em cerca de 35 locos distintos.

Acasalamento preferencial negativo: nesse sistema de acasalamento ocorre o oposto do preferencial positivo: há o aumento da frequência de heterozigotos e diminuição da frequência dos homozigotos. Sob acasalamento preferencial negativo, indivíduos com fenótipos distintos tendem a se acasalar mais frequentemente do que o esperado por acasalamento aleatório. Assim como no acasalamento preferencial positivo, as frequências genotípicas irão mudar a cada geração, mas, ao contrário do positivo, o acasalamento preferencial negativo também afeta as frequências gênicas. Nesse sentido, tal sistema de acasalamento poderia ser considerado uma força evolutiva (se definirmos forças evolutivas como processos que alteram as frequências gênicas).

Endocruzamento quer dizer, simplesmente, acasalamento ou reprodução entre indivíduos que possuem uma relação de parentesco ou ancestralidade; ou seja, implica cruzamento entre indivíduos que compartilham ancestrais comuns.



**Figura 19.1:** Heredograma apresentando cinco gerações e três casamentos consanguíneos: entre primos de primeiro grau e entre tio e sobrinha.

Você irá observar que o grau de relacionamento entre indivíduos de uma população depende do tamanho populacional ( $N$ ). Isso fica claro ao considerar o número de ancestrais possíveis. Em populações bissexuais, cada indivíduo tem dois pais, quatro avós, oito bisavós, 16 tataravós, e assim sucessivamente, ou, mais geralmente,  $2^t$  ancestrais na geração  $t$ .

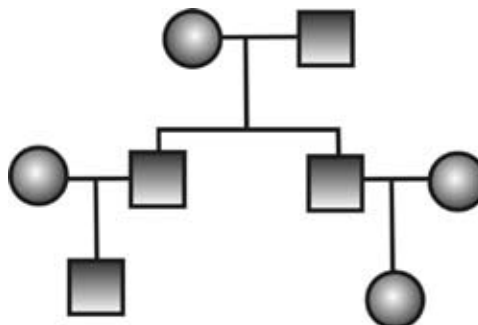
Em outras palavras, há não muitas gerações, o número de indivíduos requeridos para fornecer ancestrais separados para todos os indivíduos é maior do que o tamanho  $N$  que uma população real pode conter. Portanto, qualquer par de indivíduos deve estar relacionado, através de um ou mais ancestrais comuns no passado recente ou distante.

Quanto menor o tamanho da população das gerações anteriores, menos remoto é o ancestral comum ou maior é o seu número. Desta forma, pares acasalando aleatoriamente em uma população pequena são mais aparentados que pares acasalando ao acaso em populações grandes.

### Cálculo do coeficiente de parentesco ou consangüinidade

O parentesco genético entre dois seres humanos pode ser medido por intermédio da probabilidade de eles terem genes idênticos, herdados de um ancestral comum a ambos. Essa medida do parentesco, ou da correlação genética entre dois indivíduos, é denominada coeficiente de parentesco ou consangüinidade  $r$ , geralmente, é indicada pela letra  $r$ .

Para entendermos facilmente como se calcula o  $r$ , consideremos dois primos em primeiro grau, como os indivíduos III-1 e III-2 da **Figura 19.2**. Suponha que o indivíduo III-1 seja heterozigoto em relação a um par de alelos autossômicos  $A$  e  $a$ . Para calcular a probabilidade de sua prima III-2 também possuir o alelo  $a$ , por tê-lo herdado de um ancestral comum a ambos, temos de levar em conta que tal ancestral poderia ter sido o avô I-1 ou o avô I-2, já que os primos têm esses dois ancestrais em comum e um desses poderia ser heterozigoto  $Aa$ .



**Figura 19.2:** Heredograma de uma genealogia com primos em primeiro grau (III-1 e III-2).

Se o alelo autossômico 'a' presente em III-1 lhe tivesse sido transmitido por seu avô I-2, tal gene teria percorrido o caminho I-2→II-2→III-1. Se esse avô tivesse transmitido idêntico gene a III-2, está claro que o caminho percorrido seria I-2→II-3→III-2.

O alelo autossômico 'a' presente em III-1 poderia, entretanto, ter-lhe sido transmitido por sua avó I-1 e, nesse caso, esse gene teria percorrido o caminho I-1→II-2→III-1. E se essa avó tivesse transmitido o mesmo gene 'a' III-2, tal alelo teria percorrido o caminho I-1→II-3→III-2.

Considerando que cada passagem de um gene de uma geração a outra é chamado passo genético ou, simplesmente, passo, pode-se dizer que cada um dos quatro caminhos apontados para a transmissão do gene 'a' inclui dois passos. Tendo em mente que cada passo tem probabilidade de  $\frac{1}{2}$ , pois  $\frac{1}{2}$  é a probabilidade de o alelo 'a' passar de uma geração a outra, fica claro que a probabilidade de o gene 'a' ter sido transmitido pelo avô I-2 aos seus netos III-1 e III-2 é igual a  $(\frac{1}{2})^4 = 1/16$ , pois são quatro os passos existentes nos dois caminhos que esse gene teria de percorrer. Pelo mesmo motivo, tem-se que a probabilidade de o gene 'a' ter sido transmitido pela avó I-1 aos seus netos III-1 e III-2 é, também, igual a  $1/16$ . Pode-se concluir que a probabilidade de o gene autossômico 'a' ter sido transmitido por um ou outro ancestral comum aos dois primos em primeiro grau é igual a  $1/16 + 1/16 = 1/8$ . Visto que essa é a probabilidade de dois primos em primeiro grau herdarem o mesmo alelo de um ancestral comum a ambos, pode-se dizer, também, que o coeficiente de parentesco de primos de primeiro grau é  $1/8$ , ou seja,  $r = 1/8$ .

O cálculo do coeficiente de parentesco pode ser resumido pela fórmula  $r = \sum (\frac{1}{2})^N$ , na qual N é o número de passos que une dois consangüíneos a um ancestral comum. Evidentemente, quando dois consangüíneos têm em comum um único ancestral, não existe a possibilidade de somatório, de sorte que essa fórmula reduz-se a  $r = (\frac{1}{2})^N$ . Assim, por exemplo, no caso de dois meio-irmãos, a probabilidade de ambos serem portadores de um mesmo gene autossômico por origem comum é  $r = (\frac{1}{2})^2 = \frac{1}{4}$ , pois há um único ancestral comum a ambos (pai ou mãe) e os meio-irmãos estão ligados a ele por dois passos genéticos.

Usando o mesmo raciocínio para parentes consangüíneos colaterais com outros graus de consangüinidade, concluímos que o r de irmãos é  $\frac{1}{2}$ ; de tios e sobrinhos é igual ao de meio-irmãos ou de primos duplos

em primeiro grau, isto é,  $\frac{1}{4}$ ; de tios e meio-sobrinhos é  $\frac{1}{8}$ , como no caso de primos em primeiro grau; de primos em segundo grau é  $\frac{1}{16}$ ; de primos em terceiro grau é  $\frac{1}{32}$ ; e assim por diante.

### ATIVIDADE 1



Dentre os casais consangüíneos abaixo, qual oferece maior risco de gerar filhos com anomalias recessivas, supondo que todos esses casais são clinicamente normais e sem antecedentes familiares?

- a) primos em segundo grau;
- b) primos em primeiro grau;
- c) tio x meia-sobrinha;
- d) tio x sobrinha.

---

---

---

### RESPOSTA COMENTADA

Para responder essa questão, você precisa calcular o coeficiente de consangüinidade ( $r$ ) utilizando a fórmula  $r = (\frac{1}{2})^N$ . Lembre que  $N$  é o número de passos que unem dois consangüíneos a um ancestral comum e que  $\frac{1}{2}$  é a probabilidade de ocorrência de cada passo genético. Assim, temos  $r = \frac{1}{2}$  para irmãos ( $N = 1$ , eles possuem os mesmos pais);  $r = \frac{1}{4}$  para tios e sobrinhos, meio-irmãos ou primos duplos em primeiro grau ( $N = 2$ );  $r = \frac{1}{8}$  para tios e meio-sobrinhos e primos em primeiro grau ( $N = 3$ , o ancestral comum mais próximo é um avô ou avó);  $r = \frac{1}{16}$  para primos em segundo grau ( $N = 4$ ); e  $r = \frac{1}{32}$  para de primos em terceiro grau ( $N = 5$ ). Entre os casais apresentados no exercício, o que apresenta maior risco de gerar um filho com doença recessiva é o de tio com sobrinha, porque seu coeficiente de consangüinidade é o maior,  $r = \frac{1}{4}$ .

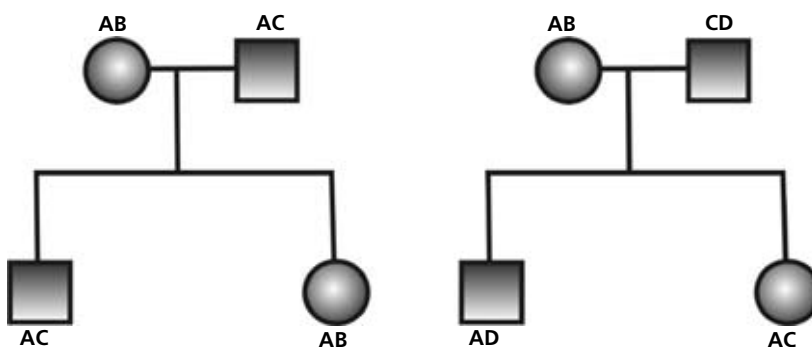
### Identidade por descendência e coeficiente de endocruzamento

A conseqüência essencial de dois indivíduos terem ancestrais comuns é que eles podem carregar réplicas de um dos genes presentes no seu ancestral; se esses indivíduos acasalam, eles podem passar essas réplicas para a prole. Ou seja, indivíduos endocruzados (prole gerada por endocruzamento) podem carregar dois genes em um loco, que são réplicas, e o mesmo gene de uma geração prévia.

Se considerarmos a possibilidade acima, podemos pensar em dois tipos de homozigotos e de identidade entre alelos. Até agora,

consideramos a identidade funcional, na qual, se dois genes não podem ser distinguidos por seus efeitos fenotípicos ou por qualquer outro critério funcional ou químico, eles são chamados alelos. Se um indivíduo qualquer tem essas duas cópias indistintas funcionalmente, ele é chamado homocigoto, no sentido mais comum do termo.

Mas vamos ver agora que existe outro tipo de identidade que se deve à origem (compartilhada ou não) entre alelos. Dois genes que se originaram da replicação de um único gene em uma geração prévia podem ser chamados **IDÊNTICOS POR DESCENDÊNCIA (IPD)** ou simplesmente idênticos. Inversamente, dois genes que não têm origem comum, ou seja, não descendem de replicação de um mesmo gene em um indivíduo ancestral, são caracterizados como idênticos por estado ou independentes na origem.



**Figura 19.3:** Heredogramas apresentando pares de irmãos não-IPD e IPD. Ambos os pares de irmãos compartilham o alelo A. O primeiro par de irmãos tem dois exemplares independentes de A (não-IPD). O segundo par de irmãos compartilha cópias do mesmo alelo A parental (IPD, alelo herdado da mãe). A diferença só é perceptível se os genótipos parentais forem conhecidos.

Homozigotos de genes idênticos por origem são denominados homocigotos idênticos ou **AUTOZIGOTOS**, e homocigotos de genes idênticos por estado são chamados homocigotos independentes ou **ALOZIGOTOS**. Em resumo, um indivíduo pode ter duas cópias de um gene que são funcionalmente indistintas (ou quimicamente indistintas), mas pode ainda ter duas possíveis relações de identidade por descendência ou origem, idênticos ou independentes. A produção de autozigotos gera o aumento de homocigotos como consequência do endocruzamento.

#### IDÊNTICOS POR DESCENDÊNCIA

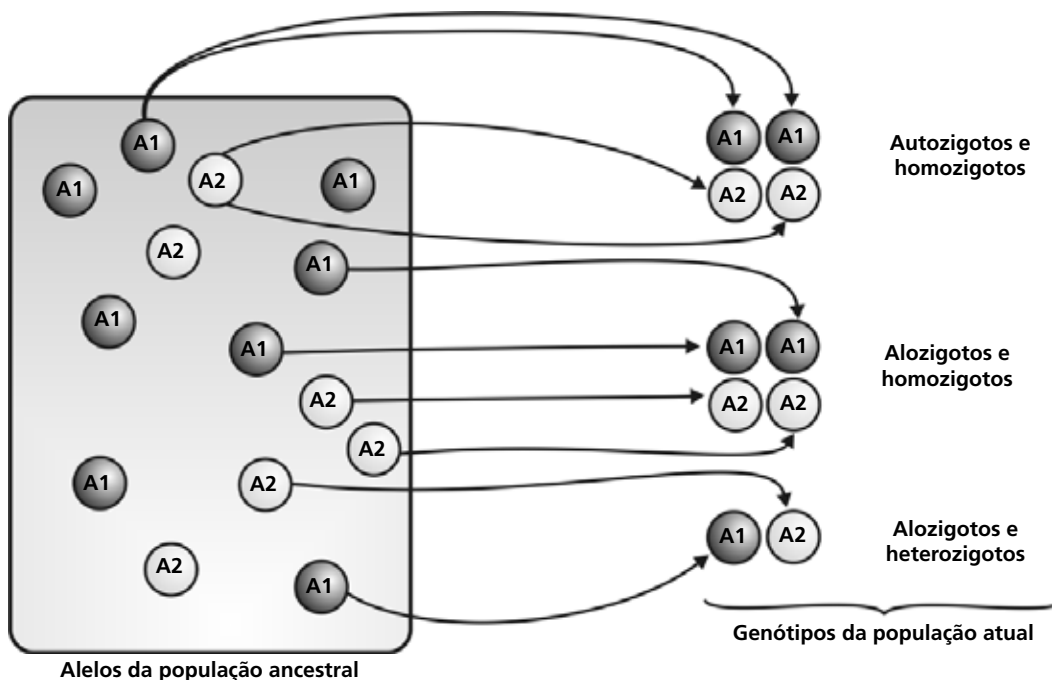
Em inglês, *identical by descent* ou *IBD*, ou por origem: alelos idênticos em duas pessoas, porque ambas são descendentes do mesmo ancestral.

#### AUTOZIGOTO

É um indivíduo homocigoto para alelos idênticos por descendência (IPD).

#### ALOZIGOTO

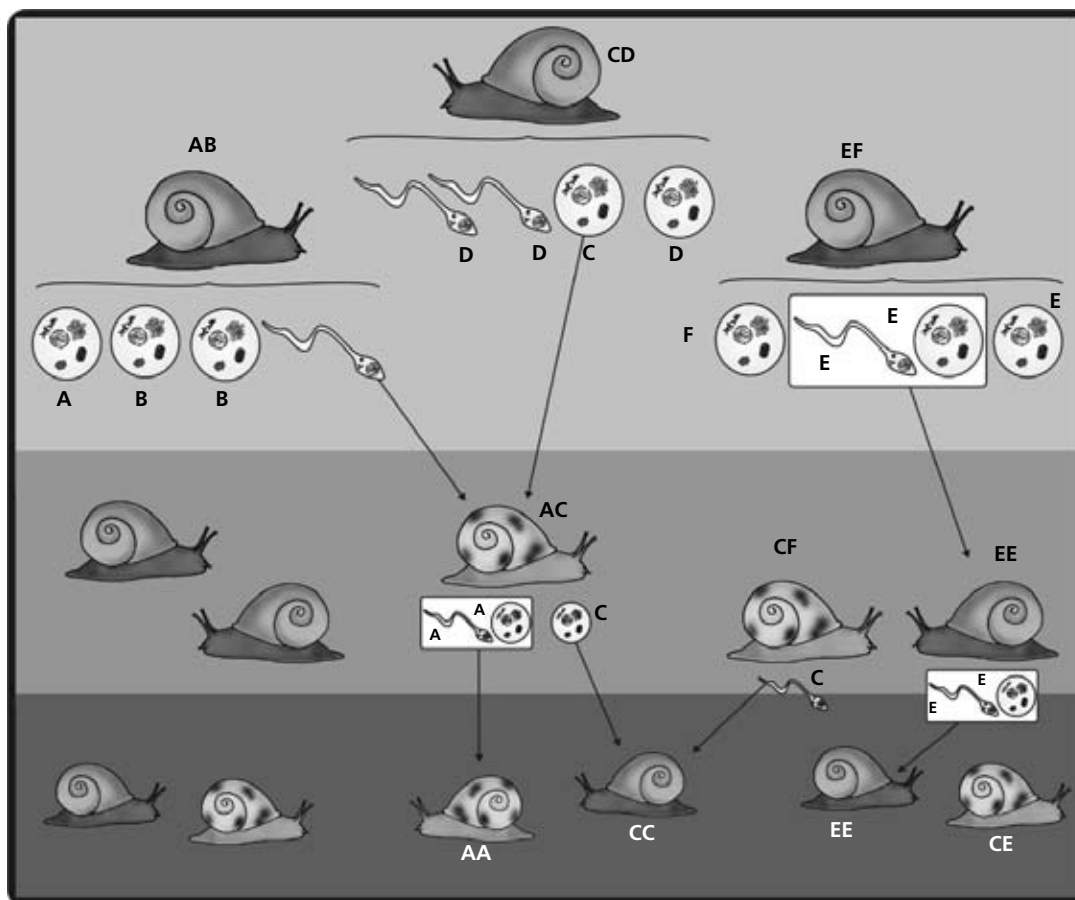
É um indivíduo homocigoto para alelos idênticos por estado ou condição; é a posse coincidente de alelos iguais, porém, independentes na origem (não-IPD).



**Figura 19.4:** Esquema mostrando como os alelos de uma população original foram agrupados formando os genótipos da população atual.

Exemplificando a geração de um organismo autozigoto, pense em um molusco hermafrodita, capaz de autofertilização, lançando ovos e espermatozoides na água (Figura 19.5). Existem  $N$  indivíduos, cada um lançando um número igual de gametas que se unem ao acaso. Todos os genes em um loco na população base são não idênticos por definição. Considerando apenas um loco, entre os gametas lançados pela população base existem  $2N$  tipos diferentes (definidos pela IPD, e não quimicamente), em números iguais, trazendo os genes A, B, C, D... Os gametas, quer sejam óvulos ou espermatozoides, podem carregar genes idênticos ou independentes. A população base tem  $F = 0$  e a única maneira de gerar um zigoto com dois alelos IPD, um autozigoto, na geração 1 é com ambos os gametas vindos de um mesmo ancestral (pai ou mãe).





**Figura 19.5:** Esquema apresentando um grupo de gastrópodos hermafroditas, os tipos de gametas produzidos e exemplos de geração de filhotes com genótipos autozígotos na F1 e na F2.

Identidade por descendência fornece a base para medir o processo dispersivo através do grau de relacionamento entre pares acasalando em uma população finita. Para isso, vamos definir o **COEFICIENTE DE ENDOCRUZAMENTO** ( $F$ , no sentido de *pedigree*), que é a probabilidade de dois genes em um loco, em um indivíduo, serem IPD. Esta definição deixa claro que  $F$  se refere ao indivíduo e expressa o grau de relacionamento entre seus pais. Se os pais foram acasalados ao acaso, o coeficiente de endocruzamento da prole tem a probabilidade de os dois gametas, tomados ao acaso da geração parental, carregarem genes idênticos em um loco.



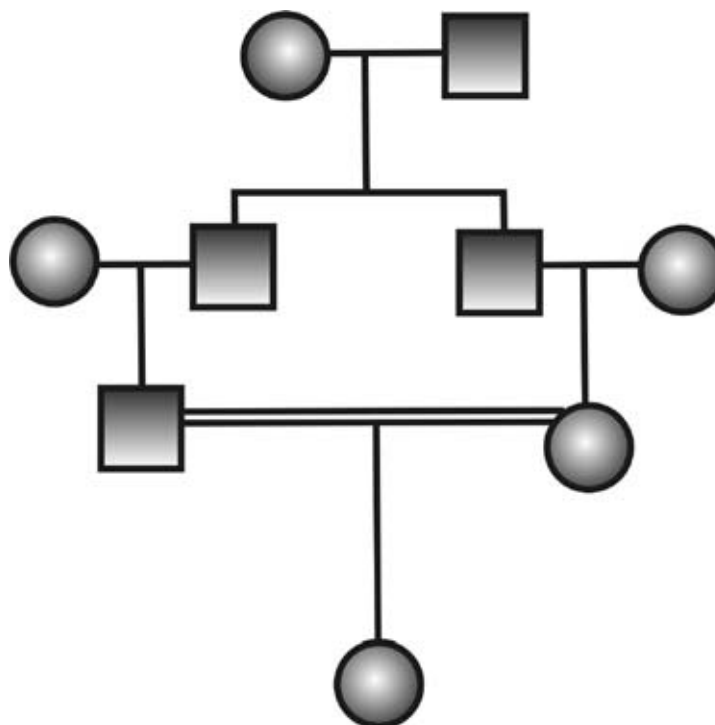
Neste texto, a palavra *pedigree* é utilizada quando nos referirmos à história genealógica de um indivíduo. Por exemplo: Felipe, filho de Rejane, filha de Joaquim, filho de Pedro. Assim, Felipe tem em seu *pedigree* um bisavô chamado Pedro.

**COEFICIENTE DE ENDOCRUZAMENTO (F)**

É a probabilidade de dois alelos em um indivíduo serem idênticos, um ao outro, por descenderem de um ancestral comum. F indica a probabilidade de um indivíduo ser autozigoto em decorrência do encontro de um par de alelos idênticos, derivados de um único gene herdado de um de seus ancestrais. Coeficiente de endocruzamento, também, é a correlação de genes dentro dos indivíduos (simbolizado por F ou FIT; coeficiente de endocruzamento total) ou é a correlação de genes dentro dos indivíduos dentro das populações (simbolizado por f ou FIS; coeficiente de endocruzamento dentro das populações). FIS também é conhecido como índice de fixação. Ambos FIS e FIT medem o desvio das proporções do EHW; valores positivos indicam deficiência de heterozigotos, enquanto valores negativos indicam excesso de heterozigotos. Já o coeficiente de coancestralidade é a correlação de genes de indivíduos diferentes de uma mesma população uma medida do parentesco de indivíduos dentro de uma população simbolizado por  $\theta$  ou FST.

Cada família pode diferir no seu valor de F, já que, com acasalamento aleatório, alguns casais serão mais próximos (aparentados) do que outros, diferindo no valor de F. Se, no entanto, tomamos o valor de F médio da população (a média de F entre todas as famílias de uma população) temos, agora, um coeficiente de endocruzamento F que se refere à população como um todo.

Para entendermos o cálculo do coeficiente de endocruzamento, tomemos o caso de um indivíduo que é filho de um casal de primos em primeiro grau, como a menina da **Figura 19.6**. Consideremos um par de alelos autossômicos 'B' e 'b' e suponhamos que a filha IV-1 é homocigota recessiva 'bb'.



**Figura 19.6:** Heredograma da genealogia de uma filha de um casal de primos em primeiro grau.

Se quisermos saber qual a probabilidade de IV-1 ser homocigota 'bb' por autozigose, teremos que calcular a probabilidade de um alelo 'b', presente em seu bisavô I-2 ou em sua bisavó I-1, que são os ancestrais comuns do casal de primos III-1 x III-2 genitores de IV-1, ter sido transmitido de modo tal que 'b' se encontre em autozigose em IV-1. Se designarmos o cromossomo autossomo qualquer que contém o loco do par de alelos 'B' e 'b' por C (de cromossomo), teremos que considerar a

possibilidade de o alelo 'b' estar em um dos dois cromossomos homólogos C da bisavó I-1, os quais poderão ser representados por  $C^1$  e  $C^2$  bem como a possibilidade de tal alelo estar em um dos dois autossomos homólogos C do bisavô I-2, os quais poderemos representar por  $C^3$  e  $C^4$ .

Se o alelo 'b' estivesse localizado no cromossomo  $C^1$ , a probabilidade de a filha de primos em primeiro grau IV-1 da **Figura 19.6** ser autozigota 'bb' seria igual a  $1/64$ , porque seria necessário que o cromossomo  $C^1$  fosse transmitido percorrendo os seguintes dois caminhos: caminho 1 = I-1 → II-2 → III-1 → IV-1 e caminho 2 = I-1 → II-3 → III-2 → IV-1.

Visto que em cada um desses caminhos existem três passos, e que o cromossomo  $C^1$ , portando o alelo 'b', tem probabilidade  $1/2$  de percorrer um passo, então, a probabilidade de tal cromossomo efetuar os seis passos dos dois caminhos apontados é igual a  $(1/2)^6 = 1/64$ . Tendo em mente, porém, que o alelo 'b' poderia estar localizado no cromossomo  $C_2$  da bisavó I-1 ou no cromossomo  $C^3$  ou  $C^4$  do bisavô I-2, temos que aplicar o mesmo raciocínio a cada uma dessas possibilidades, resultando daí que a probabilidade de autozigose, em cada caso, é igual a  $1/64$ . Desse modo, concluímos que a probabilidade de uma filha de primos em primeiro grau ser autozigota de um gene autossômico presente em um de seus ancestrais é igual a  $1/16$ , pois  $1/64 + 1/64 + 1/64 + 1/64 = 1/16$ , sendo essa probabilidade o coeficiente de endocruzamento desse indivíduo ( $F = 1/16$ ).

Repetindo o mesmo raciocínio para filhos de casais com outros graus de consangüinidade, inclusive os resultantes de relações incestuosas, chegaremos à conclusão de que o coeficiente de endocruzamento de filhos de:

- a) pais com filhas, ou mães com filhos, ou de irmãos, é igual a  $1/4$ ;
- b) meio-irmãos, ou de primos duplos em primeiro grau, ou tios(as) e sobrinhos(as), é igual a  $1/8$ ;
- c) tios(as) e meia(o)-sobrinhos(os) é igual a  $1/16$ , como no caso de filhos de primos em primeiro grau;
- d) primos em segundo grau é igual a  $1/32$ ;
- e) primos em terceiro grau é igual a  $1/64$ .

Comparando os valores dos coeficientes de endocruzamento com os coeficientes de consangüinidade, concluímos que o coeficiente de endocruzamento em relação a genes autossômicos pode ser calculado rapidamente quando se conhece o coeficiente de consangüinidade dos genitores de um indivíduo, pois  $F = r/2$ .

Como medida do processo dispersivo, interessa-nos o coeficiente de endocruzamento médio, simbolizado simplesmente por F.



Fique atento para o contexto no qual F está sendo usado! Esse pode estar se referindo a uma família em particular (F de indivíduo) ou à população como um todo (F médio da população)!



### ATIVIDADE 2

João casou com sua prima Maria, que é filha do irmão gêmeo monozigótico de seu pai com a irmã gêmea monozigótica de sua mãe. A que equivale o parentesco genético entre João e Maria? Se João for heterozigoto de um gene autossômico raro que determina uma doença recessiva, qual o risco de João e Maria gerarem uma criança com tal anomalia?

---

---

---

### RESPOSTA COMENTADA

Perceba que João e Maria têm parentesco genético de irmãos, pois são filhos de dois casais que são gêmeos idênticos (univitelinos)! O risco de gerarem uma criança homozigota para um alelo recessivo é de  $1/8$ . Como calculamos esse risco? Esse valor é o produto de  $1/2$  (um passo genético)  $\times 1/4$  (F de filhos de irmãos). Recorde que há uma outra forma de calcular esse risco através da fórmula  $F = r/2$ , ou seja,  $1/4/2$  que resulta em  $1/8$ .

## Endocruzamento e heterozigosidade

Nossa discussão, de agora em diante, fará referência ao valor de F na população ou subpopulação como um todo, e não ao F de uma família em particular.

Note que F só tem sentido quando comparado com algum ponto de referência, e este ponto de referência é a população base ou principal, que originou as subpopulações. Sempre que se fala em um valor de F para uma linhagem, implica-se (explícita ou implicitamente) a comparação com a **POPULAÇÃO BASE** que, por ser infinita e com acasalamento aleatório por definição, tem um valor de  $F = 0$ .

**POPULAÇÃO BASE**

Também chamada inicial ou idealizada, é uma população de tamanho grande (infinito), na qual o acasalamento é aleatório, e que é subdividida em um número grande de subpopulações ou linhagens, cada uma de tamanho  $N$ , tamanho esse que permanece constante de geração para geração. A subdivisão pode surgir de causas geográficas ou ecológicas em populações naturais, ou de um programa de acasalamento controlado em cativeiro ou laboratório. Todas as subpopulações em conjunto constituem a população base, e cada linhagem é uma população pequena, na qual as frequências gênicas estão sujeitas à dispersão.

Se retrocedermos no tempo, geração por geração, todas as cópias gênicas descendem de um único gene ancestral. Precisamos, então, definir um tempo  $t$  (qualquer que seja ele; lembre-se de que isto é um modelo, portanto uma simplificação da realidade), no qual qualquer identidade por descendência não será mais obtida e, conseqüentemente, a população ancestral tem um valor de  $F = 0$ . Desta forma, o valor de  $F$  nas gerações subseqüentes expressa a quantidade do processo dispersivo que ocorreu desde a população base, e compara o grau de relacionamento entre indivíduos agora com aquele da população base.

O efeito do endocruzamento pode ser quantificado em termos da redução na heterozigidade. Isto significa que podemos determinar a quantidade de endocruzamento comparando a proporção real de genótipos heterozigotos na população com a proporção esperada de genótipos heterozigotos se ocorrer acasalamento aleatório.



Estudamos o conceito de heterozigidade com detalhes na Aula 5, Evolução! Heterozigidade ou heterozigose é a proporção de indivíduos heterozigotos de uma população. É usada como medida de variabilidade genética. Uma releitura seria interessante...

Concretamente, considere um loco com dois alelos  $A$  e  $a$ , com respectivas frequências  $p$  e  $q$  ( $p + q = 1$ ). Suponha que a frequência real de genótipos heterozigotos na população no presente momento é representada por  $H$ . Se a população estiver em EHW para este gene, a frequência de genótipos heterozigotos seria igual a  $2pq$ . Vamos representar este valor básico como  $H_0$ , assim  $H_0 = 2pq$ . O efeito do endocruzamento pode ser definido em termos da quantidade  $(H_0 - H)/H_0$ , a qual equivale ao coeficiente de endocruzamento ( $F$ ). Assim:

$$F = \frac{(H_0 - H)}{H_0}$$

Biologicamente,  $F$  mede a redução na heterozigozidade, calculada como a fração relativa à esperada em uma população com cruzamento ao acaso com as mesmas freqüências de alelos.

Como  $H_0 = 2pq$ , a freqüência real de genótipos heterozigotos ( $Aa$ ) na população em endocruzamento pode ser escrita em termos de  $F$  como:

$$H = H_0 - H_0F = 2pq - 2pqF$$

A freqüência de genótipos homozigotos ( $AA$ ) na população em endocruzamento também pode ser expressa em termos de  $F$ . Suponha que a proporção de genótipos  $AA$  seja igual a  $D$ . Porque a freqüência do alelo  $A$  é  $p$ , devemos ter que  $D + H/2 = p$ . Assim:

$$D = p - H/2$$

$$D = p - \frac{(2pq - 2pqF)}{2} = p^2 + pqF$$

Da mesma forma a freqüência de genótipos  $aa$  ( $R$ ) será:

$$R = q - H/2$$

$$R = q - H/2$$

$$R = q - \frac{(2pq - 2pqF)}{2} = q^2 + pqF$$

Após uma pequena manipulação algébrica, as freqüências genotípicas com endocruzamento serão:

$$AA: \quad p^2 (1 - F) + pF$$

$$Aa: \quad 2pq (1 - F)$$

$$aa: \quad q^2 (1 - F) + qF$$

Perceba que essas fórmulas nos mostram como os homozigotos aumentam à custa dos heterozigotos. Considere que  $(1-F)$  refere-se a genes que são independentes em origem e são, portanto, equivalentes a pares de gametas tomados ao acaso da população como um todo; enquanto que  $pF$  (ou  $qF$ ) refere-se a genes que são IPD.

Podemos notar que a freqüência de homozigotos idênticos entre ambos os genótipos em conjunto é, por definição, o coeficiente de

endocruzamento  $F$  [ $pF + qF$ , ou  $F(p + q)$ , que com  $p + q = 1$ , leva a  $F$ ]. Além disso, a divisão de homozigotos idênticos entre os dois genótipos é em proporção às frequências gênicas iniciais, então  $pF$  é a frequência de homozigotos  $AA$  idênticos e  $qF$  é a frequência de homozigotos  $aa$  idênticos.

Como o valor  $(1-F)$  é diminuído das frequências genotípicas esperadas pelo EHW, a diferença fica por conta do que é acrescentado aos homozigotos ( $pF$  ou  $qF$ ). Dessa forma, o aumento dos homozigotos às custas dos heterozigotos em um processo de dispersão das frequências alélicas é gerado pelo aumento de autozigotos.

### **Endocruzamento: um desvio do esperado pelo EHW (cruzamento ao acaso)**

Podemos perceber que o endocruzamento é um processo resultante do acasalamento entre indivíduos aparentados. Esse tipo de endocruzamento resulta do *pedigree* entre indivíduos componentes de uma população. Dessa forma, o coeficiente de endocruzamento é a probabilidade de que a prole seja homozigota devido à identidade por descendência (IPD) em um loco autossômico qualquer, escolhido ao acaso. Já que  $F$  é uma probabilidade, este varia de 0 (nenhum loco IPD) a 1 (todos os locos IPD). Note que  $F$  se refere a um indivíduo qualquer em uma genealogia vindo de uma determinada união. Estendemos esse conceito para a população como um todo, simplesmente considerando o valor de  $F$  médio na população.

Entretanto, sistema de acasalamento é um conceito populacional aplicado em nível de organização biológica da deme ou população. Em certo sentido, a definição de endocruzamento, via *pedigree*, é completamente irrelevante para entender o problema dos sistemas de acasalamento. Para entender o impacto do endocruzamento nos sistemas de acasalamento e no nível de organização das populações, nós devemos examinar os desvios das frequências genotípicas do EHW devido ao acasalamento não aleatório. Primeiro, vamos reconsiderar o caso do acasalamento ao acaso para um loco com dois alelos:

**Quadro 19.1:** Freqüências alélicas e genotípicas resultantes do cruzamento entre dois heterozigotos  $A_1A_2$

		Gametas femininos		Freqüências alélicas na deme
		$A_1$	$A_2$	
Freqüência alélica		$p$	$q$	
Gametas masculinos	$A_1$	$A_1A_1$ $p \times p$ $p^2$	$A_1A_2$ $p \times q$ $pq$	$p^2+pq=p$
	$A_2$	$A_2A_1$ $q \times p$ $qp$	$A_2A_2$ $q \times q$ $q^2$	
				$qp+q^2=q$
	Freqüências alélicas na subpopulação	$p^2+qp=p$	$pq+q^2=q$	

Agora, vamos supor que os gametas sejam unidos de tal forma que exista um desvio do produto esperado de freqüências genotípicas pela regra do EHW, mas que as freqüências alélicas na população permaneçam as mesmas. Definindo-se  $\lambda$  (do grego: *lambda*) como o parâmetro que mede este desvio, temos as freqüências como representadas no **Quadro 19.2**.

**Quadro 19.2:** Freqüências alélicas e genotípicas, resultantes do cruzamento entre dois heterozigotos  $A_1A_2$ , corrigidas para o efeito do endocruzamento

		Gametas femininos		Freqüências alélicas na deme
		$A_1$	$A_2$	
Freqüência alélica		$p$	$q$	
Gametas masculinos	$A_1$	$A_1A_1$ $p \times p$ $p^2 + \lambda$	$A_1A_2$ $p \times q$ $pq - \lambda$	$p^2+pq=p$
	$A_2$	$A_2A_1$ $q \times p$ $qp - \lambda$	$A_2A_2$ $q \times q$ $q^2 + \lambda$	
				$qp+q^2=q$
	Freqüências alélicas na subpopulação	$p^2+qp=p$	$pq+q^2=q$	



Estatisticamente, o termo  $\lambda$  corresponde à **CO-VARIÂNCIA** entre gametas em união. Portanto, se  $\lambda > 0$ , existe uma **CORRELAÇÃO** positiva entre gametas em união, o que intuitivamente quer dizer que existe um excesso (de homocigotos) da expectativa por cruzamento aleatório e, contrariamente, se  $\lambda < 0$ , existe uma correlação negativa entre gametas em união. Acasalamento ao acaso corresponde à situação em que os gametas em união não estão correlacionados ( $\lambda = 0$ ).

A correlação verdadeira entre gametas em união, na verdade, é dada por  $\lambda/pq$ , tanto que podemos definir o **COEFICIENTE DE ENDOCRUZAMENTO VOLTADO PARA O SISTEMA DE ACASALAMENTO** como sendo  $f = \lambda/pq$ . Então, utilizando esta definição de  $f$ , as frequências genotípicas podem ser expressas como:

Genótipo:	$A_1A_1$	$A_1A_2$	$A_2A_2$
Frequência:	$p^2+pqf$	$2pq-2fpq=2pq(1-f)$	$q^2+pqf$

#### COEFICIENTE DE ENDOCRUZAMENTO NO SENTIDO DO SISTEMA DE ACASALAMENTO (f)

É uma medida que varia de  $-1$  a  $1$ , sendo negativo quando o sistema de cruzamento evita o endocruzamento ( $f < 0$ ), igual a zero quando ocorre cruzamento ao acaso ( $f = 0$ ) e positivo quando o sistema de acasalamento envolve cruzamento entre indivíduos relacionados ( $f > 0$ ).

Note que, embora este tipo de endocruzamento altere as frequências genotípicas em relação ao esperado por EHW, ele não altera as frequências alélicas. A frequência de  $A_1$ , por exemplo, é  $(p^2+pqf) + \frac{1}{2}[2pq(1-f)] = p$ . Portanto,  $f$  não é uma força evolutiva por si só, isto é, endocruzamento (no sentido de sistemas de acasalamento) sozinho não altera o *pool* gênico. Note, também, que as alterações nas frequências sugerem outra interpretação para  $f$ , além de ser uma correlação;  $f$  também é uma medida direta de desvio das frequências genotípicas dos heterocigotos do esperado por EHW, isto é:

$$1-f = H_o/H_e$$

onde:  $H_o$  = frequência observada de heterocigotos e  $H_e$  = frequência esperada de heterocigotos por EHW.

#### CO-VARIÂNCIA E CORRELAÇÃO

Uma noção estatística que é usada para estudar genética quantitativa é a associação, ou correlação, entre variáveis.

Como resultado de vias complexas de causa, muitas variáveis na Natureza variam juntas, mas de um modo imperfeito ou aproximado.

A medida usual de precisão de uma relação entre duas variáveis  $x$  e  $y$  é o coeficiente de correlação ( $r_{xy}$ ). Este coeficiente é calculado, em parte, pelo produto do desvio de cada observação de  $x$  da média dos valores de  $x$  e pelo desvio de cada observação de  $y$  da média dos valores de  $y$ , uma quantidade chamada de co-variância de  $x$  e  $y$  ( $cov_{xy}$ ). Procure mais exemplos destas variáveis no seu livro de Elementos de Matemática e Estatística, Aulas 19 e 21.

Portanto, uma correlação positiva entre gametas em união leva a uma deficiência de heterozigotos na população, a não correlação gera frequências de EHW, e a correlação negativa leva a um excesso de heterozigotos na deme.

Note que  $f$  (coeficiente de endocruzamento do sistema de acasalamento) e  $F$  (coeficiente de endocruzamento de *pedigree*) são muito diferentes:  $f$  refere-se à deme, não a indivíduos;  $f$  mede desvios das frequências genóticas ao esperado por HW, e não a probabilidade de IPD;  $f$  varia de  $-1$  a  $+1$ , e, portanto, engloba como casos especiais o evitamento do endocruzamento ( $f < 0$ ), cruzamento ao acaso ( $f = 0$ ) e endocruzamento ( $f > 0$ ). É matematicamente impossível medir o evitamento do endocruzamento com  $F$ .

Para ilustrar essas diferenças, vamos considerar o exemplo das ‘gazelas de Speke’, as quais tiveram um rebanho estabelecido no Zoológico de Saint Louis, Estados Unidos da América, no começo dos anos 1970, com apenas um macho e três fêmeas (a população inteira conhecida em cativeiro naquela época). Como só existia um macho, todos os animais

nascidos neste rebanho eram relacionados (aparentados), sendo que meio-irmãos era o relacionamento mais distante possível na prole derivada dos quatro animais originais.

Em 1982, todos os animais tinham um valor de  $F > 0$ , e o valor médio de  $F$  era de  $0,149$ , indicando um valor bastante alto de ‘endocruzamento’. Um estudo de alozimas (alozimas) também foi feito nesses animais, em 1982, e em um determinado loco, G6PD por exemplo, a heterozigosidade observada foi de  $0,5$ , mas a heterozigosidade esperada sob acasalamento casualizado seria de  $0,375$ . Portanto, o valor de  $f$  para este loco é de  $f = -0,333$ . Vários

outros locos polimórficos de alozimas foram estudados, todos gerando um valor de  $f < 0$ , com o valor médio de  $f$  sobre todos os locos sendo de  $-0,291$ . Esse valor de  $f$ , altamente negativo, fornece uma indicação forte de que o sistema de acasalamento era um no qual se evitava o endocruzamento. Assim sendo, este rebanho de gazelas era ao mesmo tempo um dos mais altamente endocruzados (no sentido genealógico) em termos de populações de grandes mamíferos conhecidas e, ao mesmo tempo, uma das mais altamente não-endocruzadas (no sentido de sistema de acasalamento). Não existe paradoxo aqui!



Figura 19.7: Gazela de Speke.



Lembre-se de que os dois tipos de coeficientes de endocruzamento estão medindo atributos biológicos completamente distintos.

O valor negativo de  $f$  indica que os criadores geriram o rebanho de forma a evitar, ao máximo, o endocruzamento dentro das restrições impostas pelo número limitado de indivíduos desta manada. Tenha em mente que, no acasalamento ao acaso, machos e fêmeas são acasalados par a par, a despeito do seu grau de parentesco (genealógico). Em qualquer população finita existe uma probabilidade finita de que dois indivíduos relacionados sejam pareados como reprodutores sob cruzamento aleatório.

Quanto menor a população, maior será esta probabilidade. Portanto, cruzamento aleatório ( $f = 0$ ) implica que alguns acasalamentos vão gerar um  $F > 0$  em qualquer população finita. A população das gazelas de Speke era muito pequena; então, este efeito era grande. De fato, mesmo evitando o endocruzamento ativamente ( $f < 0$ ), ao parrear somente os indivíduos menos relacionados, ainda assim a maioria dos cruzamentos era entre meio-irmãos ( $F = 0,125$ ).

Deste modo, mesmo um forte evitamento do endocruzamento, em termos do sistema de acasalamento, pode resultar em muitos indivíduos endocruzados no sentido de *pedigree*. De muitas formas, o valor médio de  $F$  de uma população é devido mais ao seu tamanho finito do que ao seu sistema de acasalamento. De fato, o valor médio de  $F$  para uma deme é usado como uma medida de uma força evolutiva que já estudamos, a deriva gênica, sendo um reflexo do valor de  $N$ .

### Impacto evolutivo de $f$

Mesmo pequenos desvios do EHW podem ter impactos maiores nas populações. Vamos considerar endocruzamento (no sentido de sistema de acasalamento) e a incidência de uma doença rara, recessiva (como é o caso de muitas doenças genéticas humanas). Nos exemplos dos Quadros 19.1 e 19.2, considere  $A_2$  como um alelo recessivo. A frequência de indivíduos com o fenótipo recessivo será, portanto, de  $q^2 + pqf$ . Suponha que  $q = 0,001$  (os alelos de doenças genéticas recessivas são raros) e  $f = 0$  (acasalamento ao acaso). Então, a frequência de indivíduos afetados é  $(0,001)^2 = 0,000001$  (1 em um milhão). Agora, considere um valor de  $f = 0,01$ , um desvio aparentemente

mínimo de EHW. Então,  $q^2 + pqf = 0,000001 + (0,999)(0,001)(0,01) = 0,000011$ . Assim, um endocruzamento de 1% causa um aumento de 1100% na incidência de uma doença (aumento de 11x). Portanto, mesmo um desvio menor do acasalamento ao acaso pode causar uma mudança profunda nas frequências genotípicas, quando alelos raros estão envolvidos.

### Impacto evolutivo de F

Da mesma forma que descrito para  $f$ , endocruzamento no sentido de *pedigree* também apresenta uma interação forte com alelos raros. Por exemplo, considere a incidência de doenças genéticas na população dos Estados Unidos. Casamento entre primos representa apenas 0,05% dos casais entre os norte-americanos com cor branca de pele, mas casamentos entre primos respondem de 18 a 24% dos albinos e de 27 a 53% dos casos de Tay-Sachs, na população americana. Este mesmo padrão é visto nas outras doenças genéticas recessivas. Portanto, mesmo uma quantidade pequena de endocruzamento (no sentido de *pedigree*) em uma população com acasalamento ao acaso pode aumentar a incidência de alguns tipos de doenças genéticas em ordens de magnitude na subpopulação endocruzada.

Resumidamente, endocruzamento produz um déficit de heterozigotos. O principal efeito do endocruzamento é aumentar a frequência de genótipos homozigotos relativamente ao esperado em uma população em acasalamento aleatório. Diferentemente do cruzamento ao acaso, que afeta alguns genes, mas não outros, o endocruzamento afeta todos os genes do genoma. O efeito do endocruzamento pode ser quantificado em termos da redução na heterozigosidade. Isto significa que podemos determinar a quantidade de endocruzamento comparando a proporção real de genótipos heterozigotos na população com a proporção esperada de genótipos heterozigotos, se ocorrer cruzamento ao acaso.

**RESUMO**

Endocruzamento significa acasalamento de indivíduos que são relacionados entre si por laços de ancestralidade. O principal efeito do endocruzamento é aumentar a frequência de genótipos homozigotos relativamente ao esperado em uma população em EHW. Diferentemente do cruzamento ao acaso, que afeta alguns genes, mas não outros, o endocruzamento afeta todos os genes do genoma.

Para qualquer par de alelos presente em um único indivíduo gerado por endocruzamento, os alelos são chamados idênticos por descendência (IPD), se ambos derivarem por replicação do DNA de um único alelo presente em alguma população ancestral.

Se dois alelos em um indivíduo gerado por endocruzamento forem IPD, o genótipo neste loco é dito autozigoto. Caso os dois alelos não sejam IPD, o genótipo é dito alozigoto. O ponto essencial é que dois alelos podem ser quimicamente idênticos (em termos de possuírem a mesma seqüência de nucleotídeos do DNA) sem serem idênticos por descendência; o conceito de identidade por descendência refere-se à origem ancestral de um alelo e não à sua identidade bioquímica.

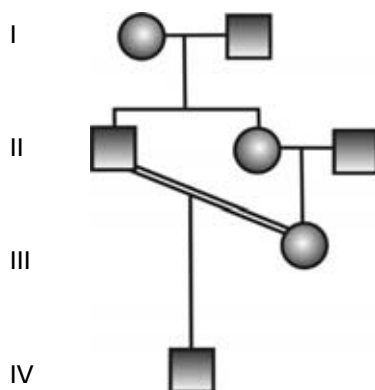
O coeficiente de endocruzamento ( $F$ ) varia de 0 a 1 e pode ser quantificado por  $(H_0 - H)/H_0$ . As frequências genotípicas com endocruzamento são: AA:  $p^2(1 - F) + pF$ ; Aa:  $2pq(1 - F)$  e aa:  $q^2(1 - F) + qF$ .

O coeficiente de endocruzamento do sistema de acasalamento ( $f$ ) difere do coeficiente de endocruzamento de *pedigree* ( $F$ ). O  $f$  refere-se sempre à deme, mede desvios das frequências genotípicas ao esperado por EHW e varia de -1 a +1.

### ATIVIDADES FINAIS

1. Desenhe um heredrograma com três gerações em que esteja representado um casamento consangüíneo. Calcule os coeficientes de consangüinidade e de endocruzamento do casal consangüíneo e de um filho gerado por eles. Lembre-se de que você pode escolher os parceiros, assim, não há só uma resposta...

**RESPOSTA COMENTADA**



No heredrograma está representado o casamento de um tio (indivíduo II-1) com sua sobrinha (III-1). Este tipo de casamento apresenta coeficiente de parentesco ou consangüinidade ( $r$ ) de  $1/4$ , visto que dois passos genéticos são necessários para achar um ancestral em comum (indivíduo I-1, avó ou I-2, avô). O coeficiente de endocruzamento do casal consangüíneo pode ser calculado pela fórmula  $F = r/2$ ; como o  $r$  é igual a  $1/4$ ,  $F = 1/4/2 = 1/8$ .

2. O que são alelos idênticos por descendência (IPD)? Dica: veja a figura dos gastrópodos (19.5), que poderá ajudá-lo a responder. Qual a diferença entre um organismo homozigoto alozigoto e um homozigoto autozigoto?

---



---



---

**RESPOSTA COMENTADA**

Alelos idênticos por descendência (IPD) são cópias, geradas por replicação, do DNA de um único alelo presente em um ancestral. Um organismo homozigoto alozigoto tem dois alelos iguais, por exemplo,  $D_1D_1$ , que têm origem independente, um veio do pai e outro da mãe (que não são parentes e têm famílias bem diferentes, um paulista casado com uma chinesa). Um organismo homozigoto autozigoto possui alelos iguais ( $D_1D_1$ ), originados por cópia de um mesmo alelo, por exemplo, se este indivíduo herdou os alelos dos pais que são primos e, por sua vez, herdaram o mesmo alelo  $D_1$  de um avô em comum.

3. Diferencie o coeficiente de endocruzamento no sentido de *pedigree* ( $F$ ) do coeficiente de endocruzamento no sentido de sistema de acasalamento ( $f$ ). Relembre o exemplo das gazelas de Speke!

---

---

---

#### RESPOSTA COMENTADA

*O coeficiente de endocruzamento no sentido de pedigree ( $F$ ) varia de 0 a 1 e é a probabilidade de que dois alelos em um indivíduo sejam idênticos, um ao outro, por descenderem de um ancestral comum (serem IPD). O coeficiente de endocruzamento do sistema de acasalamento ( $f$ ) refere-se sempre à população, mede desvios das frequências genotípicas ao esperado por EHW e varia de  $-1$  a  $+1$ .*

### AUTO-AVALIAÇÃO

Você percebeu que endocruzamento pode ser estudado de formas diferentes? É um caso especial de deriva e, também, um tipo de sistema de cruzamento. Contudo, o efeito final é o mesmo. Qual? Você está com o aproveitamento máximo se pensou que endocruzamento leva à perda de variabilidade genética (parabéns!). Isso mesmo: cruzamento entre parentes leva à homogeneização dos alelos nos locos; fica todo mundo igual! Pense assim: se você é um fazendeiro e tem uma vaca Mimosa que produz trinta litros de leite por dia, certamente vai querer que todas as suas vacas sejam iguais a ela. Como alcançar um plantel de vacas com produção média diária de 30 litros de leite? Cruze a Mimosa e suas filhas com um filho dela. A tendência será que todas as vacas produzam mais leite, pois todas serão parentas. Se você ainda tem dificuldades, veja de novo o exemplo das gazelas no Zoológico. Incrível, não? Acredito que o efeito do endocruzamento, aumentando as doenças autossômicas recessivas em humanos, seja um bom motivo para você pensar antes de decidir namorar seu (sua) primo(a)...

### INFORMAÇÕES SOBRE A PRÓXIMA AULA

Na próxima aula, analisaremos os princípios e as conseqüências do processo sistemático de migração ou fluxo gênico. Essa troca de organismos, migrantes, vai causar alterações nas frequências gênicas em populações de imigrantes e de emigrantes.