

BERNARDO BEIGUELMAN

# Genética de Populações Humanas



Sociedade  
Brasileira de  
Genética

© 2008, dos autores

Direitos reservados desta edição  
Sociedade Brasileira de Genética

Editora SBG  
Sociedade Brasileira de Genética  
Ribeirão Preto, SP

Ilustração da capa:  
William Smellie, *A Sett of Anatomical Tables*, 1754.  
"Courtesy of Historical Collections & Services, Claude  
Moore Health Sciences Library, University of Virginia."

Capa  
cubo multimidia

Beiguelman , Bernardo

GENÉTICA DE POPULAÇÕES HUMANAS. /  
Bernardo Beiguelman - Ribeirão Preto: SBG, 2008.

235p.

I. Autor. II. Título.

# **GENÉTICA DE POPULAÇÕES HUMANAS**

**BERNARDO BEIGUELMAN**

**Professor Emérito da Universidade Estadual de Campinas (UNICAMP)**

**Membro Titular da Academia Latinoamericana e da Academia Brasileira de Ciências  
Membro Titular Fundador da Academia de Ciências do Estado de São Paulo**

**Para Sylvinha, nossos filhos Giselle, Evane, LÍlian e Luciano Heitor (*in memoriam*)  
e nossos netos Mayara, Noam Rafael, Amir, Luciana e Nicolle Beatriz.**

## GENÉTICA DE POPULAÇÕES HUMANAS

### Índice

<b>INTRODUÇÃO</b> .....	<b>8</b>
<b>CAPÍTULO 1. A LEI DE HARDY E WEINBERG</b> .....	<b>9</b>
Premissas para o estabelecimento da Lei de Hardy e Weinberg.....	9
As populações humanas e a lei de Hardy e Weinberg.....	14
Estimativa das frequências gênicas e de heterozigotos de genes autossômicos quando existe relação de dominância entre os fenótipos. ....	17
As relações entre as frequências alélicas e genotípicas em populações em equilíbrio de Hardy e Weinberg.....	21
Comparação de várias amostras.....	23
Questões e respostas.....	25
Referências. ....	28
<b>CAPÍTULO 2. EXTENSÃO DA LEI DE HARDY E WEINBERG AO POLIALELISMO AUTOSSÔMICO</b> .....	<b>29</b>
Cálculo da frequência gênica em casos de polialelismo autossômico quando existe codominância. ....	31
Cálculo da frequência gênica em casos de polialelismo autossômico quando existe dominância e recessividade.....	32
Anomalias recessivas e cálculo das frequências gênicas.....	37
Comparação de várias amostras.....	38
Questões e respostas.....	43
Referências. ....	45
<b>CAPÍTULO 3. OUTROS TIPOS DE EQUILÍBRIO GENÉTICO</b> .....	<b>47</b>
Genes do cromossomo X.....	47
Teste da hipótese de equilíbrio genético. ....	51
Estimativa das frequências de genes do cromossomo X quando existe relação de dominância entre os fenótipos frequentes na população.....	53
Estimativa das frequências de genes do cromossomo X quando existe relação de dominância em relação a um fenótipo raro. ....	55
Equilíbrio genético e herança poligênica.....	56
Equilíbrio genético e genes ligados.....	57
Questões e respostas.....	58
Referências. ....	60
<b>CAPÍTULO 4. A ANÁLISE FAMILIAL DE POLIMORFISMOS</b> .....	<b>61</b>
Distribuição familiar de polimorfismos autossômicos. ....	61
O método de Snyder.....	67
O método de Fisher.....	69
Distribuição familiar de polimorfismos ligados ao sexo. ....	73
A utilização de dados a respeito de mães e filhos. ....	78
A análise familiar de caracteres complexos.....	82
Questões e respostas.....	87
Referências. ....	91
<b>CAPÍTULO 5. O EFEITO DA CONSANGÜINIDADE</b> .....	<b>93</b>
As restrições aos casamentos entre parentes. ....	93
Coeficiente de consangüinidade ou coeficiente de parentesco. ....	97

Coeficiente de consangüinidade e risco de anomalias recessivas. ....	101
Coeficiente de endocruzamento. ....	105
Coeficiente médio de endocruzamento. ....	108
Coeficiente de endocruzamento e genes do cromossomo X. ....	110
Probabilidade de os filhos de casais consangüíneos serem homozigotos. ....	112
O equilíbrio de Wright. ....	116
Componente panmíctico e componente fixo da população. ....	119
O cálculo da proporção de filhos de casais consangüíneos entre homozigotos. ....	121
Cálculo da proporção de filhos de casais de primos em primeiro grau entre homozigotos quando não se conhece o coeficiente médio de endocruzamento da população. ....	126
Equivalente letal. ....	130
Os casamentos consangüíneos e o aconselhamento genético. ....	132
O efeito Wahlund. ....	134
Questões e respostas. ....	139
Referências. ....	146

## **CAPÍTULO 6. OS FATORES EVOLUTIVOS..... 149**

1. Mutaç�o e seleç�o. ....	149
Mutaç�es espont�neas e induzidas. ....	150
Mutaç�es som�ticas e gam�ticas. ....	150
Valor adaptativo e coeficiente seletivo. ....	151
Genes letais. ....	152
Persist�ncia m�dia. ....	153
O equil�brio entre mutaç�o e seleç�o. ....	154
O c�lculo da taxa de mutaç�o. ....	156
Seleç�o contra anomalias dominantes monog�nicas. ....	159
Seleç�o contra anomalias recessivas monog�nicas. ....	160
Seleç�o contra anomalias recessivas polig�nicas. ....	162
Seleç�o contra anomalias incompletamente recessivas. ....	164
Seleç�o contra heterozigotos. ....	165
Seleç�o a favor de heterozigotos. ....	167
A manutenç�o dos polimorfismos gen�ticos neutros. ....	168
A manutenç�o dos polimorfismos adaptativos. ....	171
2. A deriva gen�tica. ....	176
3. Fluxo g�nico de populaç�es migrantes. ....	180
Mistura racial e heredopatias recessivas. ....	184
4. O tamanho dos isolados. ....	185
Quest�es e respostas. ....	187
Refer�ncias. ....	194

## **CAPÍTULO 7. EFEITO DA PREVENÇ O DE DOENÇAS GEN TICAS SOBRE AS FREQU NCIAS G NICAS..... 197**

Doenç�as autoss�micas. ....	198
Doenç�as ligadas ao sexo. ....	202
Coment�rios. ....	208
Refer�ncia. ....	210

<b>CAPÍTULO 8. MOLÉSTIAS INFECCIOSAS E CONSTITUIÇÃO GENÉTICA DO HOSPEDEIRO.....</b>	<b>211</b>
Suscetibilidade e resistência genética a infecções. ....	212
Quando é permissível supor um componente genético importante do hospedeiro que condiciona resistência a uma infecção? .....	215
1. Ocorrência de indivíduos com resistência ao contágio. ....	216
2. Associação familiar. ....	219
3. Maior concordância da moléstia infecciosa em gêmeos monozigóticos do que em dizigóticos. ....	224
4. Risco empírico de contrair a moléstia infecciosa correlacionado ao coeficiente de consangüinidade entre o comunicante e o foco. ....	227
5. Variação racial da prevalência ou da incidência da moléstia infecciosa. ....	227
6. Possibilidade de sobrepor os resultados da experimentação em animais a observações na espécie humana.....	230
7. Associação entre polimorfismos e moléstias infecciosas. ....	231
O estudo de alguns modelos.....	233
Referências. ....	236



## INTRODUÇÃO

A Genética de Populações Humanas dedica-se ao estudo da distribuição de frequências gênicas e de caracteres hereditários normais e patológicos nas populações humanas, bem como aos fatores que mantêm ou, em oposição, alteram as frequências gênicas ou genóticas nessas populações. A busca pela compreensão da dinâmica populacional dos genes e dos genótipos normais e patológicos é a razão da importância da Genética de Populações para os estudiosos da Genética Humana e Médica, da Epidemiologia, da Antropologia Física e para aqueles que se dedicam à Biologia Humana e à Evolução, pois é por intermédio do estudo dos fatores evolutivos que entendemos como se faz a manutenção da carga hereditária deletéria através de gerações.

Os capítulos deste volume foram redigidos com a finalidade de atingir grande público. É por essa razão que, para sua boa compreensão, é necessário apenas conhecimentos rudimentares de Matemática e de Bioestatística básica. Muita atenção ao texto ajuda bastante. Boa sorte, pois.

## CAPÍTULO 1. A LEI DE HARDY E WEINBERG

Oito anos depois da redescoberta das leis de Mendel (1908), Wilhelm Weinberg e Godfrey Harold Hardy chegaram independentemente, e quase que simultaneamente, às mesmas conclusões a respeito daquilo que é considerado o fundamento da Genética de Populações, isto é, *o ramo da Genética que visa à investigação da dinâmica dos genes nas populações naturais, buscando a elucidação dos mecanismos que alteram a sua composição gênica* (efeito de *fatores evolutivos*, isto é, *mutações, seleção natural, deriva genética e fluxo gênico de populações migrantes*) *ou apenas a frequência genotípica pelo aumento da homozigose* (efeito dos *casamentos consanguíneos* ou da *subdivisão da população em grandes isolados*). As conclusões concordantes a que chegaram esses dois autores passaram a ser conhecidas como a *lei do equilíbrio de Hardy e Weinberg* ou, mais simplesmente, *lei de Hardy e Weinberg*.

Hardy foi um importante matemático inglês (1877-1947), mas sua contribuição à Genética restringiu-se ao assunto deste capítulo. Em oposição, Weinberg (1862-1937), além de ter sido um dos criadores da Genética de Populações, deu contribuições notáveis e pioneiras ao estudo de gêmeos, à correção de distorções causadas pelo tipo de averiguação, e à solução de numerosos problemas de estatística médica. O espantoso é que Weinberg conseguiu harmonizar seu duro trabalho de clínico geral e obstetra (mais de 3.500 partos), que exerceu durante 42 anos em Stuttgart, Alemanha, com suas atividades de criação científica original, reunidas em mais de 160 publicações.

### PREMISSAS PARA O ESTABELECIMENTO DA LEI DE HARDY E WEINBERG

Weinberg e Hardy perceberam que *se não existissem fatores evolutivos atuando sobre uma população, as frequências gênicas permaneceriam inalteradas e as proporções genotípicas atingiriam um equilíbrio estável, mostrando a mesma relação constante entre si ao longo do tempo*.

Para demonstrar esse princípio, imaginemos uma população teórica que não esteja sujeita a fatores evolutivos nem àqueles que alteram as frequências genotípicas aumentando a homozigose, isto é, uma população que obedeça às seguintes premissas:

1. A população é infinita.
2. Existe o mesmo número de homens e de mulheres na população.

3. A população está em *panmixia*, isto é, todos casam e os casamentos ocorrem aleatoriamente, não existindo, por conseguinte, casamentos preferenciais entre indivíduos por causa de seu genótipo, fenótipo, estratificação social ou consangüinidade. Aliás, por serem os casamentos realizados aleatoriamente, os casamentos consangüíneos podem existir, desde que ocorram aleatoriamente.
4. Todos os casais da população são igualmente férteis e geram o mesmo número de filhos.
5. Não há sobreposição de gerações na população, isto é, elas não se imbricam ao longo do tempo, porque todos os indivíduos devem ter a mesma idade ao casar.
6. Os genes da população não sofrem mutação.
7. A população não está sob pressão de seleção natural, porque todos os indivíduos são igualmente viáveis, não existindo fatores que aumentem ou diminuam a sobrevivência de indivíduos com determinado genótipo.
8. A população não recebe nem emite um fluxo gênico capaz de alterar a sua composição gênica original, porque ela não sofre miscigenação com uma população imigrante que apresenta freqüências gênicas diferentes da dela, nem há emigração diferencial, isto é, a saída de grupos de indivíduos com freqüência gênica distinta do resto da população.

Consideremos, agora, que, nessa população teórica, os genótipos  $AA$ ,  $Aa$  e  $aa$ , decorrentes de um par de alelos autossômicos  $A, a$ , se distribuem com a mesma freqüência nos indivíduos de ambos os sexos. As freqüências dos alelos  $A$  e  $a$  podem ser calculadas facilmente se tomarmos como ponto de partida os gametas que produziram os indivíduos da geração atual dessa população. Assim, o número de gametas com o alelo  $A$  deve ser igual ao dobro do número de indivíduos homocigotos  $AA$  dessa geração somado ao número de indivíduos heterocigotos  $Aa$ , pois cada indivíduo  $AA$  foi originado por dois gametas com o alelo  $A$  e cada indivíduo  $Aa$  foi formado por um gameta com o gene  $A$  e outro com o seu alelo  $a$ . Por raciocínio análogo conclui-se que o número de gametas com o alelo  $a$  que produziram os indivíduos da geração em estudo é igual ao dobro do número de indivíduos  $aa$  somado ao número de indivíduos heterocigotos  $Aa$ .

Em vista do exposto, se chamarmos as freqüências dos alelos  $A$  e  $a$  na população respectivamente de  $p$  e  $q = 1 - p$ , e simbolizarmos as freqüências dos indivíduos com

genótipos  $AA$ ,  $Aa$  e  $aa$  por  $AA$ ,  $Aa$  e  $aa$ , poderemos escrever que as frequências  $p$  e  $q$  dos alelos  $A$  e  $a$  na geração em estudo são:

$$p = \frac{2AA + Aa}{(2AA + Aa) + (2aa + Aa)}$$

$$q = \frac{2aa + Aa}{(2AA + Aa) + (2aa + Aa)}$$

Visto que o denominador dessas fórmulas, que representa a frequência total dos alelos, isto é,  $A + a$ , pode ser escrito como  $2(AA+Aa+aa)$ , e considerando que  $AA+Aa+aa = 1$  ou 100%, pode-se escrever, também, que:

$$p = \frac{2AA + Aa}{2} = AA + \frac{1}{2}Aa$$

$$q = \frac{2aa + Aa}{2} = aa + \frac{1}{2}Aa \text{ ou } q = 1 - p$$

Assim, por exemplo, se na geração inicial dessa população teórica os genótipos em discussão tivessem frequências  $AA = 0,30$ ,  $Aa = 0,50$  e  $aa = 0,20$ , as frequências  $p$  do gene  $A$  e  $q$  de seu alelo  $a$ , nessa geração, seriam iguais, respectivamente, a 55% e 45%, pois:

$$p = 0,30 + 0,25 = 0,55$$

$$q = 0,20 + 0,25 = 0,45 \text{ ou } q = 1 - 0,55 = 0,45$$

Evidentemente, a frequência do alelo  $A$  também pode ser obtida pela simples contagem dos indivíduos estudados. Assim, se somarmos o dobro do número de indivíduos que são homozigotos  $AA$  ao número de indivíduos que são heterozigotos  $Aa$  e dividirmos o resultado pelo dobro do total de indivíduos estudados ( $AA+Aa+aa$ ), pois estamos contando genes, também obteremos a frequência  $p$  do alelo  $A$ . A frequência  $q$  do alelo  $a$  poderá ser obtida por intermédio de  $q = 1 - p$  ou, é claro, pela soma do dobro do número de indivíduos homozigotos  $aa$  ao número de heterozigotos  $Aa$ , que deve ser dividida pelo dobro do total de indivíduos estudados. Tomemos um exemplo numérico, considerando uma amostra de 100 indivíduos dos quais 30 têm genótipo  $AA$ , 50 genótipo  $Aa$  e 20 genótipo  $aa$ . As frequências  $p$  e  $q$  dos alelos  $A$  e  $a$  também poderiam ser calculadas por intermédio de:

$$p = \frac{(2 \times 30) + 50}{200} = 0,55$$

$$q = \frac{(2 \times 20) + 50}{200} = 0,45$$

Consideremos, agora, uma população que, *na geração inicial*, tem os genótipos  $AA$ ,  $Aa$  e  $aa$ , respectivamente, com as frequências 30%, 60% e 10%, o que significa

que, nessa geração inicial,  $A = p = 0,60$  e  $a = q = 0,40$ . Considerando, ainda, que, por hipótese, a população teórica em apreço está em panmixia, as frequências dos diferentes tipos de casais segundo os genótipos  $AA$ ,  $Aa$  e  $aa$  serão aquelas calculadas na segunda coluna da Tabela 1.1-A. Por outro lado, levando em conta que todos os casais dessa população são, por hipótese, igualmente férteis, as proporções genótípicas esperadas entre os filhos dos diferentes tipos de casais são aquelas expressas nas três últimas colunas da Tabela 1.1-A, as quais mostram um total de 36% de indivíduos  $AA$ , 48% de indivíduos  $Aa$  e 16% de indivíduos  $aa$ . Esses totais permitem concluir pela validade de uma parte da lei de Hardy e Weinberg, isto é, daquela que diz que *as frequências gênicas se mantêm constantes ao longo das gerações* de uma população teórica como a que estamos considerando.

Tabela 1.1. Demonstração da equação do equilíbrio de Hardy e Weinberg partindo de uma população teórica panmíctica na qual, na geração inicial, os genótipos  $AA$ ,  $Aa$  e  $aa$  ocorrem com frequências iguais, respectivamente, a 30%, 60% e 10%.

#### 1.1.A Distribuição genotípica após uma geração de panmixia.

Casais (geração inicial)		Primeira geração filial		
Tipo	Frequência	$AA$	$Aa$	$aa$
$AA \times AA$	$0,30 \times 0,30 = 0,09$	0,09	-	-
$AA \times Aa$	$2 \times 0,30 \times 0,60 = 0,36$	0,18	0,18	-
$AA \times aa$	$2 \times 0,30 \times 0,10 = 0,06$	-	0,06	-
$Aa \times Aa$	$0,60 \times 0,60 = 0,36$	0,09	0,18	0,09
$Aa \times aa$	$2 \times 0,60 \times 0,10 = 0,12$	-	0,06	0,06
$aa \times aa$	$0,10 \times 0,10 = 0,01$	-	-	0,01
<b>Total</b>	<b>1,00</b>	<b>0,36</b>	<b>0,48</b>	<b>0,16</b>

#### 1.1.B. Distribuição genotípica na segunda geração filial em panmixia.

Casais (1ª. geração filial)		Segunda geração filial		
Tipo	Frequência	$AA$	$Aa$	$aa$
$AA \times AA$	$0,36 \times 0,36 = 0,1296$	0,1296	-	-
$AA \times Aa$	$2 \times 0,36 \times 0,48 = 0,3456$	0,1728	0,1728	-
$AA \times aa$	$2 \times 0,36 \times 0,16 = 0,1152$	-	0,1152	-
$Aa \times Aa$	$0,48 \times 0,48 = 0,2304$	0,0576	0,1152	0,0576
$Aa \times aa$	$2 \times 0,48 \times 0,16 = 0,1536$	-	0,0768	0,0768
$aa \times aa$	$0,16 \times 0,16 = 0,0256$	-	-	0,0256
<b>Total</b>	<b>1,0000</b>	<b>0,36</b>	<b>0,48</b>	<b>0,16</b>

**Lembrete:** Para o cálculo da frequência de casais com genótipos diferentes ( $AA \times Aa$ ,  $AA \times aa$  e  $Aa \times aa$ ), multiplicamos por dois o produto das frequências desses genótipos porque temos que levar em conta o sexo dos cônjuges. Assim, por exemplo, no caso dos casais  $AA \times Aa$  temos que levar em conta a probabilidade de o casal ser composto por marido  $AA$  e mulher  $Aa$ , bem como a probabilidade de o casal incluir marido  $Aa$  e mulher  $AA$ .

De fato, é fácil verificar que na geração filial as frequências  $p$  e  $q$  dos alelos  $A$  e  $a$  continuam iguais às da geração inicial, apesar de as frequências genótípicas terem sido

alteradas. Assim, as freqüências dos genótipos  $AA$ ,  $Aa$  e  $aa$  que eram, respectivamente, iguais a 30%, 60% e 10% passaram a ser iguais a 36%, 48% e 16%. Entretanto, as freqüências gênicas continuaram com os valores iniciais de 60% e 40%, pois:

$$p = AA + \frac{1}{2}Aa = 0,36 + 0,24 = 0,60$$

$$q = aa + \frac{1}{2}Aa = 0,16 + 0,24 = 0,40$$

Os resultados da Tabela 1.1.A não permitem concluir pela validade da outra parte da lei de Hardy e Weinberg, ou seja, daquela que afirma que *as proporções genotípicas atingirão equilíbrio estável, mostrando uma relação constante entre si através dos tempos*. Entretanto, em relação a um par de alelos autossômicos é simples demonstrar que tal equilíbrio é atingido *após uma única geração de panmixia*, mantendo-se constante, daí por diante, as freqüências genotípicas alcançadas na primeira geração filial, as quais se distribuem segundo:

$$(p + q)^2 = p^2 + 2pq + q^2$$

sendo  $p$  e  $q = 1 - p$  as freqüências dos alelos  $A$  e  $a$ . A distribuição  $p^2 + 2pq + q^2$  também é conhecida como a *equação do equilíbrio de Hardy e Weinberg*.

Numa população em equilíbrio de Hardy e Weinberg tudo se passa, portanto, como se os espermatozóides com o alelo  $A$ , cuja freqüência é  $p$ , e os espermatozóides com o alelo  $a$ , cuja freqüência é  $q$ , fertilizassem os ovócitos, respectivamente, com probabilidades  $p$  e  $q$ . Visto que os ovócitos com o alelo  $A$  também têm freqüência  $p$  e aqueles com o alelo  $a$  também têm freqüência  $q$  e considerando que estamos lidando com acontecimentos independentes, as probabilidades dos genótipos originados do encontro dos gametas devem ser o produto das probabilidades desses gametas. Tudo se passaria como no quadro abaixo onde as probabilidades estão assinaladas entre parênteses:

OVÓCITOS	ESPERMATOZÓIDES	
	$A$ ( $p$ )	$a$ ( $q$ )
$A$ ( $p$ )	$AA$ ( $p^2$ )	$Aa$ ( $pq$ )
$a$ ( $q$ )	$Aa$ ( $pq$ )	$aa$ ( $q^2$ )

Realmente, na Tabela 1.1.A constata-se, de imediato, que os genótipos  $AA$ ,  $Aa$  e  $aa$  se distribuem na geração filial como  $(p+q)^2 = 1$ , isto é,

$$AA + Aa + aa = p^2 + 2pq + q^2 = 1, \text{ pois:}$$

$$AA = p^2 = 0,60 \times 0,60 = 0,36$$

$$Aa = 2pq = 2 \times 0,60 \times 0,40 = 0,48$$

$$aa = q^2 = 0,40 \times 0,40 = 0,16$$

Do mesmo modo imediato é possível verificar na Tabela 1.1.B que a distribuição  $AA : Aa : aa :: p^2 : 2pq : q^2$  se mantém inalterada na segunda geração filial, pois nessa geração  $AA = 0,26$ ,  $Aa = 0,48$  e  $aa = 0,16$ .

Tabela 2.1. Distribuição das famílias de uma população teórica que está em equilíbrio de Hardy e Weinberg em relação aos genótipos determinados por um par de alelos autossômicos  $A,a$  com frequências iguais respectivamente a  $p$  e  $q = 1 - p$ .

Casais		Filhos		
Tipo	Frequência	$AA$	$Aa$	$aa$
$AA \times AA$	$p^2 \cdot p^2 = p^4$	$p^4$	-	-
$AA \times Aa$	$2(p^2 \cdot 2pq) = 4p^3q$	$2p^3q$	$2p^3q$	-
$AA \times aa$	$2(p^2 \cdot q^2) = 2p^2q^2$	-	$2p^2q^2$	-
$Aa \times Aa$	$2pq \cdot 2pq = 4p^2q^2$	$p^2q^2$	$2p^2q^2$	$p^2q^2$
$Aa \times aa$	$2(2pq \cdot q^2) = 4pq^3$	-	$2pq^3$	$2pq^3$
$aa \times aa$	$q^2 \cdot q^2 = q^4$	-	-	$q^4$
<b>Total</b>	<b><math>(p+q)^4 = 1</math></b>	<b><math>p^2</math></b>	<b><math>2pq</math></b>	<b><math>q^2</math></b>

A Tabela 2.1, por sua vez, generaliza a distribuição das famílias em uma população teórica que está em equilíbrio de Hardy e Weinberg em relação aos genótipos determinados por um par de alelos autossômicos  $A,a$  cujas frequências são iguais, respectivamente, a  $p$  e  $q = 1 - p$ . Nessa tabela é fácil constatar que a soma das frequências dos diferentes tipos de casais da população é igual a 1 ou 100%, pois sabemos que  $p+q = 1$  e que essa soma pode ser escrita como  $(p+q)^2(p+q)^2 = (p+q)^4 = 1$ . Também não é difícil verificar na Tabela 2.1 que as somas das frequências dos indivíduos com genótipos  $AA$ ,  $Aa$  e  $aa$  na geração filial resultam, respectivamente, em  $p^2$ ,  $2pq$  e  $q^2$ . De fato, lembrando que  $p^2+2pq+q^2 = 1$ , tem-se nas somas das três últimas colunas da Tabela 2.1:

$$p^4 + 2p^3q + p^2q^2 = p^2(p^2 + 2pq + q^2) = p^2$$

$$2p^3q + 4p^2q^2 + 2pq^3 = 2pq(p^2 + 2pq + q^2) = 2pq$$

$$p^2q^2 + 2pq^3 + q^4 = q^2(p^2 + 2pq + q^2) = q^2$$

### AS POPULAÇÕES HUMANAS E A LEI DE HARDY E WEINBERG

Se uma população humana, ou qualquer outra população de indivíduos com reprodução cruzada, obedecesse as oito premissas apresentadas no tópico anterior, ela apresentaria uma estabilidade genética que permaneceria inalterada através dos tempos.

Em outras palavras, se houvesse essa obediência, as populações humanas mostrariam uma **fixidez genética**, isto é, uma **inércia evolutiva**, pois não estariam sujeitas a uma série de fatores que serão estudados em detalhe em capítulos vindouros. Tais fatores são os **fatores evolutivos**, ou seja, aqueles capazes de alterar as freqüências gênicas (*mutações, seleção natural, fluxo gênico de populações migrantes e deriva genética*) e os **fatores que causam o aumento de homozigose** (*endocruzamento* ou, como nos referimos na espécie humana, *casamentos consanguíneos, e subdivisão da população em grandes isolados*).

Evidentemente, as oito condições estabelecidas para a obtenção do equilíbrio de Hardy e Weinberg não são satisfeitas completamente por nenhuma população real, seja ela humana ou não. Aliás, o que torna possível explicar o processo evolutivo dos seres vivos em termos mendelianos é, justamente, essa desobediência ao modelo teórico apresentado no tópico anterior. Apesar de nenhuma população humana obedecer às premissas enumeradas no tópico anterior, a prática tem demonstrado, num aparente paradoxo, em numerosas populações humanas e em relação a um grande número de caracteres monogênicos que não suscitam casamentos preferenciais, como é o caso dos grupos sanguíneos, que os genótipos se distribuem de acordo com a lei de Hardy e Weinberg. Para exemplificar, tomemos os dados obtidos em amostras de cinco populações humanas a respeito dos grupos sanguíneos M (genótipo *MM*), MN (genótipo *MN*) e N (genótipo *NN*) do sistema MNS (Tabela 3.1).

Tabela 3.1. Comparação entre as proporções genótípicas do sistema sanguíneo MNS observadas com o emprego de dois anti-soros (anti-M e anti-N) em várias amostras de populações e as esperadas em equilíbrio de Hardy e Weinberg. Os valores percentuais foram registrados entre parênteses.

AMOSTRA	GENÓTIPOS				FREQUÊNCIAS GÊNICAS			No. ESPERADO			$\chi^2_{(1)}$	Ref.*
	<i>MM</i>	<i>MN</i>	<i>NN</i>	Total	<i>p</i>	<i>q</i>	$\sigma$	$np^2$	$n2pq$	$nq^2$		
Norte-Americanos	125 (31,7)	193 (49,0)	76 (19,3)	394 (100)	0,562	0,438	0,018	124,5	193,8	75,7	0,006 0,90<P<0,95	1
Holandeses	52 (27,1)	88 (45,8)	52 (27,1)	192 (100)	0,500	0,500	0,026	48	96	48	1,354 0,20<P<0,30	2
Ingleses	363 (28,4)	634 (49,6)	282 (22,0)	1279 (100)	0,532	0,468	0,010	362	637	280	0,031 0,80<P<0,90	3
Xavantes	41 (51,9)	30 (38,0)	8 (10,1)	79 (100)	0,709	0,291	0,036	39,7	32,6	6,7	0,501 0,30<P<0,50	4
Brasileiros	30 (30)	50 (50)	20 (20)	100 (100)	0,550	0,450	0,035	30,25	49,5	30,25	0,011 0,90<P<0,95	5

\* 1- Wiener e Wexler, 1958; 2- Saldanha *et al.*, 1960; 3-Race e Sanger, 1962; 4- Neel *et al.*, 1964; 5- Beiguelman, dados não publicados.

As frequências dos alelos *M* e *N* das cinco amostras reunidas na Tabela 3.1 podem ser estimadas a partir dos percentuais assinalados nessa tabela ou por contagem dos indivíduos. Assim, por exemplo, no caso da população norte-americana mencionada nessa tabela, a frequência *p* e *q* dos alelos *M* e *N* pode ser obtida por intermédio de:

$$p = MM + \frac{1}{2} MN = 0,317 + 0,245 = 0,562$$

$$q = NN + \frac{1}{2} MN = 0,193 + 0,245 = 0,438 \text{ ou } q = 1 - 0,562 = 0,438$$

ou, por intermédio de:

$$p = \frac{(2 \times 125) + 193}{2 \times 394} = 0,562$$

$$q = \frac{(2 \times 76) + 193}{2 \times 394} = 0,438$$

O desvio padrão desses alelos é calculado por intermédio de  $\sigma = \sqrt{\frac{pq}{2n}}$  onde *n* é o número de indivíduos, sendo, portanto, *2n* o total de alelos.

Obtidas as frequências esperadas de indivíduos com grupos sanguíneos *M*, *MN* e *N* na hipótese de equilíbrio de Hardy e Weinberg por intermédio de  $p^2$ ,  $2pq$  e  $q^2$ , respectivamente, pode-se calcular o número esperado de pessoas com cada um desses grupos sanguíneos. Para tanto, basta multiplicar cada percentual pelo tamanho da amostra (*n*), ou seja, calcular  $np^2$ ,  $n2pq$  e  $nq^2$  e aplicar um teste de qui-quadrado ( $\chi^2$ ) para averiguar se a diferença entre os números observados e os esperados é significativa ou não. Assim, por exemplo, no caso da amostra de índios xavantes da Tabela 3.1 faríamos o cálculo abaixo:

Valores	M	MN	N	Total
Observados ( <i>o</i> )	41	30	8	79
Esperados ( <i>e</i> )	39,7	32,6	6,7	79
$\frac{(o - e)^2}{e}$	0,042	0,207	0,252	$\chi^2_{(1)} = 0,501$ $0,30 < P < 0,50$

Os valores de  $\chi^2$  calculados para cada uma das amostras da Tabela 3.1 tornam evidente que as diferenças entre números observados e esperados não são significativas, podendo todas as amostras serem consideradas como extraídas de populações em equilíbrio de Hardy e Weinberg quanto aos genótipos *MM*, *MN* e *NN*.

Aqui é importante chamar a atenção do leitor para o fato de os qui-quadrados da Tabela 3.1 terem apenas um grau de liberdade, apesar de existirem três classes esperadas. É que para o cálculo das classes esperadas são necessárias duas informações, isto é, o tamanho da amostra e a frequência de um dos alelos. Sabendo-se que ***o número***

*de graus de liberdade do qui-quadrado é igual ao número de classes esperadas menos o número necessário de informações da amostra para o cálculo das proporções nessas classes*, tem-se, no caso presente, que esse número é igual a um, pois  $3 - 2 = 1$ .

Os exemplos apresentados nas Tabela 3.1, aliados a uma infinidade de outros mencionados na literatura pertinente, servem para demonstrar de modo inequívoco que, apesar de a lei de Hardy e Weinberg ser baseada em um modelo teórico, ela descreve suficientemente bem o que ocorre com grande número de genótipos nas populações reais, pelo menos em um determinado intervalo de tempo. Evidentemente, quando não se consegue demonstrar que uma amostra representa uma população em equilíbrio de Hardy e Weinberg, somos obrigados a investigar a(s) causa(s) desse desvio entre os fatores evolutivos e entre aqueles que alteram as frequências genotípicas, como é o caso dos casamentos consanguíneos.

O aparente paradoxo da aceitação da hipótese de equilíbrio de Hardy e Weinberg nas populações humanas, apesar de elas estarem sabidamente expostas a fatores que deveriam afetar esse equilíbrio, encontra explicação na acomodação relativamente rápida dessas populações em um novo equilíbrio genético, toda a vez que elas são expostas à atuação de fatores evolutivos. Por outro lado, considerando que as taxas de mutação são bastante baixas, é de se prever que, durante um intervalo relativamente longo, a maioria das características monogênicas que não servem para estimular casamentos preferenciais deve mostrar equilíbrio de Hardy e Weinberg em grandes populações, que vivem em um ambiente relativamente estável, não sujeitas a migrações intensas, nem apresentando alta taxa de casamentos consanguíneos.

### **ESTIMATIVA DAS FREQUÊNCIAS GÊNICAS E DE HETEROZIGOTOS DE GENES AUTOSSÔMICOS QUANDO EXISTE RELAÇÃO DE DOMINÂNCIA ENTRE OS FENÓTIPOS**

A comprovação da validade da lei de Hardy e Weinberg para numerosos caracteres mendelianos sem relação de dominância permite admitir que aqueles com esse tipo de relação também devem obedecer essa lei em grandes populações que vivem em um ambiente relativamente estável, não sujeitas a migrações intensas e que não apresentam alta taxa de casamentos consanguíneos. Tal extrapolação tem um valor prático notável, pois, com base nela, se conhecermos a frequência populacional de indivíduos com fenótipo recessivo, poderemos estimar as frequências gênicas e, por

consequente, a frequência com que ocorrem nessa população os indivíduos com fenótipo dominante que são heterozigotos. Vejamos como isso pode ser feito.

Consideremos um par de alelos autossômicos  $A, a$  e que, em uma amostra aleatória de  $n$  indivíduos,  $x$  apresentam o fenótipo recessivo determinado pelo genótipo  $aa$ , enquanto  $y$  apresentam o fenótipo dominante determinado pelos genótipos  $AA$  ou  $Aa$ , o qual pode ser, por isso, representado por  $A_$ . Se aceitarmos que a população da qual foi extraída a amostra está em equilíbrio de Hardy e Weinberg em relação aos genótipos  $AA$ ,  $Aa$  e  $aa$ , ou seja, se admitirmos que  $AA = p^2$ ,  $Aa = 2pq$  e  $aa = q^2$  teremos que  $\frac{x}{n} = aa = q^2$ .

Nesse caso, a raiz quadrada dessa proporção estimará a frequência  $q$  do alelo  $a$ , podendo-se escrever, portanto, ser  $q = \sqrt{\frac{x}{n}}$  e calcular o desvio padrão de  $q$  por intermédio de  $\sigma = \sqrt{\frac{1-q^2}{4n}}$ , de acordo com Neel e Schull (1954) e Li (1972). Visto que  $p+q = 1$ , a frequência  $p$  do alelo  $A$  será estimada a partir de  $p = 1-q$  e as frequências dos heterozigotos  $Aa$  e dos homozigotos  $AA$  serão facilmente obtidas, já que  $Aa = 2pq$  e  $AA = p^2$ .

Para demonstrar a aplicação dessas fórmulas tomemos um exemplo numérico a respeito dos grupos sanguíneos D-positivo e D-negativo do sistema Rh, também conhecidos como Rh-positivo e Rh-negativo. A presença do antígeno D é condicionada por um gene  $D$  do loco  $RHD$  situado no braço superior do cromossomo número 1, mais precisamente em 1p36.2-p34.3 (Cherif-Zahar *et al.*, 1999). Em consequência de deficiências (*deletions*, em inglês) ou de outras alterações no gene  $D$  (Colin *et al.*, 1991), tem-se como resultado a ausência de atividade desse gene, a qual pode ser simbolizada pelo alelo  $d$ . Os indivíduos com genótipo  $DD$  ou  $Dd$  possuem fenótipo dominante D-positivo ou Rh-positivo, enquanto que aqueles com genótipo  $dd$  possuem o fenótipo recessivo D-negativo ou Rh-negativo.

Uma amostra da população do Estado de São Paulo, constituída de 2.039 indivíduos cujas hemácias foram testadas com um anti-soro anti-D (anti-Rh<sub>0</sub>) revelou que, dentre os indivíduos examinados, 1.848 eram D-positivo e 191 D-negativo (Beiguelman, 1963). Em vista disso, podemos dizer que, nessa amostra,  $\frac{191}{2.039} = 0,0937$

ou 9,37% eram homozigotos  $dd$  e que  $\frac{1.848}{2.039} = 0,9063$  ou 90,63% tinham genótipo

$D_+$ , isto é, os homozigotos  $DD$  e heterozigotos  $Dd$  constituíam 90,63% da amostra.

Com base nesses dados, para estimar a frequência de heterozigotos  $Dd$  na população da qual procede a amostra em questão, estimamos, inicialmente, a frequência  $q$  do alelo  $d$  como abaixo:

$$\begin{aligned} dd &= q^2 = 0,0937 \\ d &= q = \sqrt{0,0937} = 0,306 \\ \sigma &= \sqrt{\frac{1-0,0937}{4 \times 2.039}} = 0,011 \end{aligned}$$

e, a partir daí, resolvemos:

$$\begin{aligned} D &= p = 1 - 0,306 = 0,694 \\ Dd &= 2pq = 2 \times 0,694 \times 0,306 = 0,4274 \text{ ou } 42,47\%. \end{aligned}$$

Se, em vez de querer estimar qual a frequência de heterozigotos  $Dd$  quiséssemos saber qual a proporção de indivíduos D-positivo que são heterozigotos  $Dd$ , o cálculo teria que ser diverso, pois agora estamos diante de um problema de **probabilidade condicional**. Visto que queremos saber qual a fração de indivíduos heterozigotos ( $2pq$ ) dentre aqueles que são D-positivo ( $p^2 + 2pq$ ), temos que calcular a probabilidade de um indivíduo ser heterozigoto  $Dd$ , **dado que ele é  $D_+$** , ou seja, temos que calcular:

$$P(Dd | D_+) = \frac{2pq}{p^2 + 2pq} = \frac{2q}{p + 2q}$$

Lembrando que  $p$  pode ser substituído por  $1 - q$ , pode-se escrever, também, que:

$$P(Dd | D_+) = \frac{2q}{1 + q}$$

Usando os dados numéricos de nosso exemplo, tem-se que, na população estudada, a frequência de heterozigotos dentre aqueles que são D-positivo pode ser estimada em 46,86%, pois  $\frac{2 \times 0,306}{1 + 0,306} = 0,4686$ .

Em vista do exposto, se quisermos estimar qual a probabilidade de um casal constituído por cônjuges D-positivo gerar uma criança D-negativo, teremos que considerar a probabilidade de tais cônjuges serem heterozigotos  $Dd$ , dado que eles são D-positivo ( $D_+$ ). Considerando que os casais heterozigotos  $Dd \times Dd$  têm probabilidade igual a  $\frac{1}{4}$  de gerar uma criança D-negativo ( $dd$ ), tem-se que a probabilidade de um casal constituído por cônjuges D-positivo gerar uma criança D-negativo é calculada a partir da expressão:

$$P(Dd|D\_).P(Dd|D\_). \frac{1}{4} = \frac{2q}{1+q} \cdot \frac{2q}{1+q} \cdot \frac{1}{4} = \frac{q^2}{(1+q)^2} = \left( \frac{q}{1+q} \right)^2$$

de modo que, em relação à população em apreço, essa probabilidade é igual a 5,49%, pois  $\left( \frac{0,306}{1,306} \right)^2 = 0,0549$ .

Se quiséssemos saber qual a probabilidade de um casal constituído por um cônjuge D-positivo e outro D-negativo gerar uma criança D-negativo, teríamos que levar em conta a probabilidade de o cônjuge D-positivo ser heterozigoto  $Dd$  dado que ele é D-positivo ( $D\_$ ). De fato, sabemos que entre os casais D-positivo  $\times$  D-negativo somente podem gerar crianças D-negativo ( $dd$ ), com probabilidade  $\frac{1}{2}$ , aqueles cujo cônjuge D-positivo é heterozigoto  $Dd$ . Portanto, a probabilidade de um casal composto por um cônjuge D-positivo e outro D-negativo gerar uma criança D-negativo é estimada a partir da expressão:

$$P(Dd|D\_). \frac{1}{2} = \frac{2q}{1+q} \cdot \frac{1}{2} = \frac{q}{1+q}$$

de sorte que, em relação à população de nosso exemplo, essa probabilidade é igual a 23,4%, pois  $\frac{0,306}{1,306} = 0,234$ .

Quando lidamos com caracteres recessivos raros, como é o caso das hereditopatias recessivas com transmissão monogênica, e sabemos que a taxa de casamentos consangüíneos na população em estudo é tão pequena que pode ser desprezada, o cálculo da probabilidade de encontro de um indivíduo heterozigoto nessa população ou entre os indivíduos normais a ela pertencentes é bem mais simples. Tomemos um exemplo para demonstrar a veracidade dessa afirmação. Consideremos que, dentre 100.000 nascimentos ocorridos em um determinado período em uma certa população caucasóide, com taxa de casamentos consangüíneos desprezível, 5 foram de crianças com o erro inato do metabolismo decorrente da falta de hidroxilase de fenilalanina, ou seja, com fenilcetonúria clássica. Como se sabe, esse erro inato do metabolismo recessivo autossômico, quando não detectado e tratado precocemente com uma dieta pobre em fenilalanina provoca deficiência mental grave.

Pode-se, pois, estimar que a freqüência da fenilcetonúria entre os recém-nascidos da população estudada (incidência) é de 5:100.000 ou 1:20.000. Visto que a

fenilcetonúria é recessiva autossômica, podemos simbolizar o genótipo por  $aa$  e escrever:

$$aa = q^2 = \frac{1}{20.000}$$

$$q = \sqrt{\frac{1}{20.000}} \cong \frac{1}{141} \cong 0,0071$$

o que permite estimar a frequência  $p$  do alelo “normal” como sendo  $p = 1 - q = \frac{140}{141}$  e obter um desvio padrão muito pequeno ( $\sigma = 0,0016$ ).

Levando em conta que  $\frac{140}{141}$  é semelhante à unidade, a frequência de heterozigotos  $Aa$  poderá ser estimada pelo dobro da estimativa da frequência do alelo  $a$ , isto é,  $Aa = 2q \cong \frac{1}{70}$ . Em outras palavras, por causa da frequência muito alta do alelo  $A$ , não é necessário calcular  $2pq = 2 \cdot \frac{140}{141} \cdot \frac{1}{141}$  para estimar a frequência de heterozigotos, já que, calculando  $Aa = 2q = 2 \cdot \frac{1}{141}$  o resultado que se obtém é, praticamente, o mesmo, isto é, semelhante a  $\frac{1}{70}$ .

O valor obtido por intermédio de  $2q$  também serve para estimar a probabilidade de um indivíduo normal da população ser heterozigoto de um gene raro que, em homozigose, determina uma anomalia. Em outras palavras, para estimar essa probabilidade não é necessário calcular a probabilidade condicional  $P(Aa|A_) = \frac{2q}{1+q}$ , porque, por intermédio dessa fórmula, chegar-se-á a uma estimativa semelhante àquela que é dada por  $2q$ . De fato, empregando os dados de nosso exemplo sobre fenilcetonúria clássica, tem-se que, se a probabilidade de um indivíduo normal da população for calculada a partir de  $P(Aa|A_) = \frac{2q}{1+q}$ , obter-se-á um número muito semelhante a  $\frac{1}{70}$ :

$$\frac{\frac{2}{141}}{1 + \frac{1}{141}} = \frac{2}{142} = \frac{1}{71}$$

## AS RELAÇÕES ENTRE AS FREQUÊNCIAS ALÉLICAS E GENOTÍPICAS EM POPULAÇÕES EM EQUILÍBRIO DE HARDY E WEINBERG

A discussão feita no tópico anterior a respeito da estimativa da frequência de heterozigotos de genes autossômicos serve, também, para ressaltar que, à medida que

diminui a frequência do alelo determinante do fenótipo recessivo, maior se torna a razão entre a frequência de heterozigotos e homozigotos recessivos, isto é, a razão  $\frac{Aa}{aa}$ .

Realmente, no exemplo sobre a frequência de fenilcetonúricos podemos verificar que a proporção de heterozigotos do gene da fenilcetonúria clássica é, praticamente, 286 vezes maior do que a de fenilcetonúricos, pois, dentre cada 70 indivíduos da população em estudo, um é heterozigoto, enquanto que a proporção de fenilcetonúricos é de um em cada 20.000 indivíduos. Já em relação aos indivíduos D-positivo e D-negativo de nosso outro exemplo pode-se constatar que a frequência de heterozigotos  $Dd$  é apenas cerca de 4,5 vezes maior do que a de homozigotos  $dd$ , pois os primeiros ocorrem com frequência igual a 42,47% e os últimos com frequência igual a 9,37%.

A Fig.1.1 mostra, graficamente, as relações entre as frequências  $p$  e  $q$  de dois alelos autossômicos designados por  $A$  e  $a$ , e as frequências genóticas ( $AA$ ,  $Aa$  e  $aa$ ) em populações em equilíbrio de Hardy e Weinberg. Essa figura deixa claro que, se uma população estiver em equilíbrio de Hardy e Weinberg, a frequência máxima de heterozigotos ( $Aa$ ) é de 50% ou 0,5, e que isso ocorre quando os dois alelos têm a mesma frequência, isto é,  $p = q = 0,5$ . Nessa situação, os homozigotos têm a mesma frequência, que é igual a 25% ( $AA = aa = 0,25$ ). Por aí se vê que, a respeito da geração inicial do exemplo dado na Tabela 1.1, podemos dizer, de imediato, que ela não representa uma população em equilíbrio de Hardy e Weinberg, já que nela, os heterozigotos  $Aa$  aparecem com frequência igual a 60%. Nesse exemplo da Tabela 1.1, a distribuição dos genótipos  $AA : Aa : aa :: 30% : 60% : 10%$  foi proposital, para deixar claro que uma única geração de panmixia é suficiente para que a população passe a mostrar equilíbrio de Hardy e Weinberg.

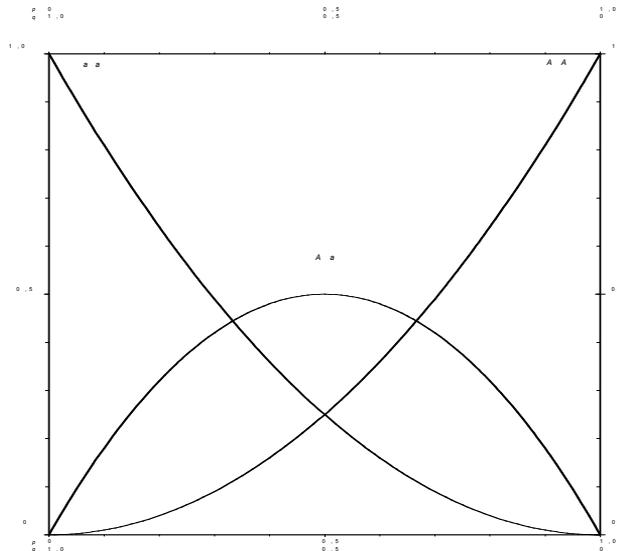


Fig.1.1 Gráfico das relações entre as frequências  $p$  e  $q$  de um par de alelos autossômicos  $A$ ,  $a$  e as frequências genotípicas  $AA$ ,  $Aa$  e  $aa$  em populações em equilíbrio de Hardy e Weinberg.

### COMPARAÇÃO DE VÁRIAS AMOSTRAS

Quando estudamos amostras de várias populações quanto a caracteres determinados por um par de alelos autossômicos, é muito comum que se queira comparar essas amostras, a fim de investigar se elas podem ser reunidas em uma única ou se isso não deve ser feito, em decorrência de uma diferença significativa entre elas. No caso de os alelos autossômicos não mostrarem relação de dominância e recessividade entre si, a comparação entre as diferentes amostras pode ser feita, facilmente, levando em conta a distribuição gênica. Se as amostras representarem populações em equilíbrio de Hardy e Weinberg poder-se-á usar a alternativa de comparar as distribuições genotípicas dessas amostras.

Para exemplificar, tomemos os dados da Tabela 3.1, que permitem fazer a contagem dos genes  $M = 2MM + MN$  e  $N = 2NN + MN$  como na Tabela 4.1 e aplicar um teste de qui-quadrado para esses dados. O valor do qui-quadrado com 4 graus de liberdade ( $\chi^2_{(4)} = 22,896$ ;  $P < 0,001$ ) assinalado abaixo dessa tabela indica que as cinco amostras não podem ser reunidas em uma única por diferirem significativamente quanto às proporções dos alelos  $M$  e  $N$ . Entretanto, se retirarmos da comparação feita na Tabela 4 a amostra de índios xavantes, verificaremos que as outras quatro amostras podem ser reunidas em uma única, já que, nelas, a proporção dos alelos  $M$  e  $N$  não diferem significativamente ( $\chi^2_{(3)} = 4,521$ ;  $0,20 < P < 0,30$ ).

Tabela 4.1. Comparação das cinco amostras da Tabela 3.1 quanto às proporções dos alelos

*M e N.*

AMOSTRA	<i>M</i>	<i>N</i>	TOTAL
N.- americanos	443	345	788
Holandeses	192	192	384
Inglêses	1.360	1.198	2.558
Xavantes	112	46	158
Brasileiros	110	90	200
Total	2.217	1.871	4.088

$$\chi^2_{(4)} = 22,896; P < 0,001$$

Visto que todas as amostras da Tabela 3.1 estão em equilíbrio de Hardy e Weinberg, chegaríamos a um resultado semelhante ao da análise feita na Tabela 4 se tivéssemos comparado a distribuição genotípica dessas amostras. Assim, também concluiríamos que as cinco amostras não podem ser reunidas em uma única ( $\chi^2_{(8)} = 25,897; P < 0,01$ ), mas que retirando os xavantes da comparação, as outras quatro amostras não mostrariam diferença significativa entre elas ( $\chi^2_{(6)} = 5,710; 0,30 < P < 0,50$ ).

Nos casos em que queremos comparar várias amostras quanto a fenótipos resultantes de um par de alelos que mostram relação de dominância e recessividade não temos alternativas, pois o único modo de compará-las é em relação à distribuição fenotípica. Assim, por exemplo, se quiséssemos comparar amostras de caucasóides, afro-descendentes e descendentes de japoneses quanto à frequência de indivíduos Rh-positivo (*D*\_) e Rh-negativo (*dd*) apresentadas na Tabela 5.1, teríamos que comparar as proporções ali apresentadas e concluir, em vista do alto valor do qui-quadrado com dois graus de liberdade, que elas diferem significativamente.

Tabela 5.1. Frequência de indivíduos D-positivo e D-negativo do sistema sanguíneo Rh em brasileiros caucasóides, afro-descendentes e descendentes de japoneses. Entre parênteses foram assinalados os valores em porcentagem.

AMOSTRA	D-positivo	D-negativo	TOTAL
Caucasóides	2247 (87,4)	324 (12,6)	2571
Afros	216 (91,5)	20 (8,5)	236
Japoneses	293 (97,7)	7 (2,3)	300

$$\chi^2_{(2)} = 30,343; P < 0,001$$

## QUESTÕES E RESPOSTAS

**Q 1.** O sangue de 200 indivíduos coletados aleatoriamente em uma população foi estudado eletroforeticamente no intuito de determinar os tipos de haptoglobina, tendo-se observado as seguintes freqüências genotípicas:

$$Hp^1Hp^1 = 32\%; \quad Hp^1Hp^2 = 46\%; \quad Hp^2Hp^2 = 22\%$$

Com base nesses dados estimar as freqüências dos alelos autossômicos  $Hp^1$  e  $Hp^2$  na população representada por essa amostra e o desvio padrão.

**R 1.**  $Hp^1 = p = 0,32 + 0,23 = 0,55$

$$Hp^2 = q = 1 - 0,55 = 0,45$$

$$\sigma = 0,025$$

**Q 2.** A amostra da questão anterior pode ser considerada como representativa de uma população em equilíbrio de Hardy e Weinberg?

**R 2.** Sim, porque  $\chi^2_{(1)} = 0,999; 0,30 < P < 0,50$ .

**Q 3.** Qual o percentual de casais heterozigotos  $Hp^1Hp^2 \times Hp^1Hp^2$  que devemos esperar na população representada pela amostra da questão número 1?

**R 3.** 24,5%, pois:

$$Hp^1Hp^2 = 2pq = 2 \times 0,55 \times 0,45 = 0,495$$

$$Hp^1Hp^2 \times Hp^1Hp^2 = 0,495 \times 0,495 = 0,245$$

**Q 4.** Em uma população 70% dos indivíduos apresentam o genótipo  $AA$  e 30% o genótipo  $aa$ , sendo os alelos  $A$  e  $a$  genes autossômicos. Com base nesses dados pergunta-se:

a) Qual a freqüência dos alelos  $A$  e  $a$  ?

b) A população está em equilíbrio de Hardy e Weinberg?

c) Qual a distribuição genotípica esperada em equilíbrio de Hardy e Weinberg?

**R 4.** a)  $A = p = 0,70; a = q = 0,30$

b) Não e nem é preciso fazer qualquer teste para demonstrar isso, porque essa população não apresenta heterozigotos.

c)  $AA = 0,49; Aa = 0,42; aa = 0,09$ .

**Q 5.** Com o emprego dos anti-soros anti-M e anti-N foram determinados os grupos sanguíneos M, MN e N de uma amostra aleatória de 100 indivíduos de uma população, encontrando-se a seguinte distribuição:

$$M = 20\%; \quad MN = 64\%; \quad N = 16\%$$

Visto que esses grupos sanguíneos são caracteres autossômicos e correspondem aos genótipos  $MM$ ,  $MN$  e  $NN$  quer-se saber quais as estimativas das frequências dos alelos  $M$  e  $N$  a partir dessa amostra e o seu desvio padrão.

**R 5.**  $M = p = 0,52$ ;  $N = q = 0,48$ ;  $\sigma = 0,035$

**Q 6.** Pode-se considerar que a amostra da questão anterior representa uma população em equilíbrio de Hardy e Weinberg quanto aos grupos sanguíneos  $M$ ,  $MN$  e  $N$ ?

**R 6.** Não, porque  $\chi^2_{(1)} = 7,955$ ;  $P < 0,01$ .

**Q 7.** Ao estudar a distribuição dos grupos sanguíneos  $M$ ,  $MN$  e  $N$  em uma amostra aleatória de indivíduos de uma população verificou-se a relação  $\frac{MN}{\sqrt{M \times N}} = 2$ . O que significa essa relação?

**R 7.** Essa relação significa que as frequências observadas correspondem exatamente às esperadas em equilíbrio de Hardy e Weinberg, pois  $\frac{2pq}{\sqrt{p^2 \times q^2}} = \frac{2pq}{pq} = 2$ .

**Q 8.** Em uma amostra de 200 pacientes com tuberculose pulmonar verificou-se que 128 eram acetiladores lentos de isoniazida, que, pelos métodos bioquímicos comuns, é um fenótipo recessivo em relação ao da acetilação rápida desse medicamento empregado no tratamento da tuberculose. Com base nesses dados estimar a frequência do alelo que determina a acetilação lenta da isoniazida e o desvio padrão.

**R 8.** Se  $q^2 = \frac{128}{200} = 0,64$ , tem-se  $q = \sqrt{0,64} = 0,80$  e  $\sigma = \sqrt{\frac{1-0,64}{800}} = 0,021$ .

**Q 9.** A frequência de indivíduos Rh-negativo ( $dd$ ) em uma população é estimada em 9%. Qual a estimativa da frequência de mulheres Rh-positivo heterozigotas ( $Dd$ ) nessa mesma população?

**R 9.** 42%, porque sendo  $d = q = \sqrt{0,09} = 0,30$  e  $D = p = 1 - 0,30 = 0,70$ , tem-se que  $Dd = 2pq = 2 \times 0,70 \times 0,30 = 0,42$ .

**Q 10.** Em uma população cuja frequência de indivíduos Rh-negativo ( $dd$ ) é estimada em 16%, quer-se saber qual a probabilidade de um casal Rh-positivo ter um filho Rh negativo?

**R 10.** 8,16%, porque  $q = \sqrt{0,16} = 0,40$  e  $P(Dd|D\_).P(Dd|D\_).$   $\frac{1}{4} = \left(\frac{0,40}{1,40}\right)^2 = 0,0816$ .

**Q 11.** Em uma cidade constituída basicamente por caucasóides de origem européia mediterrânea estima-se que 1% das crianças manifestam anemia de Cooley (talassemia  $\beta$  homozigótica). Sabendo-se que, nessa população, a taxa de casamentos

consangüíneos é desprezível, quer-se saber, sem fazer o rastreamento de heterozigotos, qual a estimativa da freqüência, nessa população :

- do gene da talassemia  $\beta$ ?
- de heterozigotos desse gene?
- de casais heterozigotos da talassemia  $\beta$ ?

**R 11.** a)  $q = \sqrt{0,01} = 0,10$  ou 10%

b)  $2pq = 0,18$  ou 18%.

c)  $4p^2q^2 = 0,324$  ou 32,4%.

**Q 12.** Ao estudar uma amostra de 2.520 brasileiros afro-descendentes adultos não se assinalou nenhum caso de anemia falciforme (homozigose do gene da hemoglobina S, isto é, SS), mas constatou-se que 202 apresentavam o traço siclêmico. Quer-se saber:

- Qual a estimativa da freqüência do gene da hemoglobina S na população representada por essa amostra?
- Qual a incidência de anemia falciforme por 100.000 crianças afro-descendentes, que se espera nessa população, se ela estiver em equilíbrio de Hardy e Weinberg?

**R 12.** a)  $S = q = \frac{202}{2 \times 2.520} = 0,04$ .

b) 160: 100.000, pois  $SS = q^2 = 0,0016$  ou  $P(AS \times AS) \frac{1}{4} = 0,08 \times 0,08 \times 0,25 = 0,0016$ .

**Q 13.** A freqüência de hiperplasia congênita das supra-renais por deficiência de 11- $\beta$ -hidroxilase foi estimada em 1: 10.000 entre as crianças de uma cidade. Se a taxa de casamentos consangüíneos na população dessa cidade for desprezível, qual será a estimativa da freqüência:

- do gene determinante da deficiência de 11- $\beta$ -hidroxilase?
- de heterozigotos desse gene?

**R 13.** a)  $q = 0,01$ .

b)  $2q = 0,02$ .

**Q 14.** Numa população altamente miscigenada, com predomínio do componente africano, a incidência de anemia falciforme na população é estimada em 16 por 10.000 nascimentos. Qual o risco que correm os(as) irmãos(ãs) clinicamente normais desses pacientes de gerar uma criança com anemia falciforme se vierem a casar com uma pessoa clinicamente normal da mesma população.

**R 14.** O risco é estimado em cerca de 1 : 78. De fato, sabemos que um(a) irmão(ã) clinicamente normal de um paciente com anemia falciforme tem, *a priori*, a

probabilidade igual a  $2/3$  de ser heterozigoto do gene da hemoglobina siclêmica. Por outro lado, tendo em mente que a frequência  $q$  do gene da hemoglobina siclêmica é a raiz quadrada de  $q^2 = 0,0016$ , isto é  $q = 0,04$ , conclui-se que a frequência de heterozigotos desse gene na população geral é igual a  $2pq = 2 \times 0,96 \times 0,04 = 0,0768$ . Finalmente, lembrando que um casal de heterozigotos corre o risco de  $1/4$  de gerar um filho homozigoto tem-se:

$$2/3 \times 0,0768 \times 1/4 = 0,0128 \text{ ou cerca de } 1: 78$$

porque levamos em conta a probabilidade de um(a) irmão(ã) de um paciente com anemia falciforme ser heterozigoto, a probabilidade de o cônjuge clinicamente normal oriundo da mesma população ser heterozigoto, e a probabilidade de, sendo ambos heterozigotos, gerarem uma criança homozigota.

## REFERÊNCIAS

- Beiguelman, B. Grupos sanguíneos e lepra. *Rev. Bras. Lepr.* 31: 34-44, 1963.
- Cherif-Zahar, B., Mattei, M.G., Le Van Kim, C., Bailly, P., Cartron, J.-P. & Colin, Y. Localization of the human Rh blood group gene structure to chromosome region 1p34.3-1p36.1 by *in situ* hybridization. *Hum Genet.* 86: 398-400, 1991.
- Colin, Y., Cherif-Zahar, B., Le Van Kim, C., Raynal, V., Van Huffel, V. & Cartron, J.P. Genetic basis of the RhD-positive and RhD-negative blood group polymorphism as determined by Southern analysis. *Blood* 78: 2747-2752, 1991.
- Hardy, G.H. Mendelian proportions in a mixed population. *Science* 78: 49-50, 1908.
- Li, C.C. *Population genetics*. Univ. Chicago Press, Chicago, 1972.
- Neel, J.V. & Schull, W.J. *Human heredity*. Univ. Chicago Press, Chicago, 1954.
- Neel, J.V., Salzano, F.M., Junqueira, P.C., Keiter, F. & Maybury-Lewis, D. Studies on the Xavante Indians of the Brazilian Mato Grosso. *Hum. Genet.* 16: 52-140, 1964.
- Race, R.R. & Sanger, R. *Blood groups in Man*. Blackwell Scient. Publ., Oxford, 1962.
- Saldanha, P.H., Frota-Pessoa, O., Eveleth, P., Ottensooser, F., Cunha, A.B., Cavalcanti, M.A.A. & outros. Estudo genético e antropológico de uma colônia de holandeses no Brasil. *Rev. Antrop. (Brasil)* 8: 1-42, 1960.
- Weinberg, W. Über den Nachweis der Vererbung beim Menschen. *Jahreshefte Verein f. vaterl. Naturk. in Wurtemberg* 64: 368-382, 1908.
- Wiener, A.S. & Wexler, I.B. *Heredity of the blood groups*. Grune and Stratton, N. York, 1958.

## CAPÍTULO 2. EXTENSÃO DA LEI DE HARDY E WEINBERG AO POLIALELISMO AUTOSSÔMICO

A lei de Hardy e Weinberg não está restrita apenas à ocorrência de um par de alelos em um loco autossômico, isto é, aos casos de *dialelismo*. Ela pode ser generalizada para os casos de *polialelismo autossômico*, ou seja, para aqueles em que consideramos vários alelos pertencentes a um loco autossômico. Em outras palavras, nos casos de polialelismo autossômico também se pode demonstrar que, se uma população obedecer as oito condições estabelecidas para a população teórica do capítulo anterior, *as freqüências gênicas dos diferentes alelos se manterão inalteradas e as proporções genotípicas atingirão equilíbrio estável após uma única geração de panmixia*. Assim, se um loco autossômico possuir  $n$  alelos ( $A, a, a_1, a_2, \dots, a_n$ ), que ocorrem em uma população teórica como a descrita no capítulo anterior com freqüências  $p, q, r, s, \dots, x$ , essas freqüências se manterão inalteradas e o equilíbrio estável dos genótipos será atingido após uma geração de panmixia, com os genótipos se distribuindo segundo:

$$(p + q + r + s + \dots + x)^2$$

Para demonstrar que esse equilíbrio se estabelece após uma única geração de panmixia consideremos um exemplo de trialelismo em uma população teórica que obedeça às oito premissas estabelecidas para a demonstração do equilíbrio de Hardy e Weinberg. Assim, se na geração inicial encontrarmos apenas indivíduos com os genótipos  $AA, aa$  e  $a_1a_1$ , determinados pelos alelos autossômicos  $A, a$  e  $a_1$  e se as freqüências desses indivíduos forem  $AA = 0,30, aa = 0,10$  e  $a_1a_1 = 0,60$ , tanto entre os homens, quanto entre as mulheres, é claro que as freqüências dos três alelos em questão serão  $A = p = 0,30, a = q = 0,10$  e  $a_1 = r = 0,60$ .

Evidentemente, a geração inicial não está em equilíbrio segundo a lei de Hardy e Weinberg, pois, para tanto, seria necessário que os genótipos decorrentes dos alelos  $A, a$  e  $a_1$  se distribuíssem segundo  $(p + q + r)^2 = 1$ . Em outras palavras, seria necessário observar as freqüências genotípicas abaixo, as quais, de acordo com a Tabela 1.2 são alcançadas após uma única geração em panmixia.

$$\begin{aligned} AA &= p^2 = 0,30 \times 0,30 = 0,09 \\ aa &= q^2 = 0,10 \times 0,10 = 0,01 \\ a_1a_1 &= r^2 = 0,60 \times 0,60 = 0,36 \\ Aa &= 2pq = 2 \times 0,30 \times 0,10 = 0,06 \\ Aa_1 &= 2pr = 2 \times 0,30 \times 0,60 = 0,36 \end{aligned}$$

$$aa_1 = 2qr = 2 \times 0,10 \times 0,60 = 0,12$$

Tabela 1.2. Distribuição dos genótipos decorrentes de trialelismo após uma única geração de uma população teórica em panmixia. A geração inicial é constituída por indivíduos homocigotos *AA* (30%), *aa* (10%) e *a<sub>1</sub>a<sub>1</sub>* (60%).

Casais (geração inicial)		Primeira Geração Filial					
Tipo	Frequência	<i>AA</i>	<i>aa</i>	<i>a<sub>1</sub>a<sub>1</sub></i>	<i>Aa</i>	<i>Aa<sub>1</sub></i>	<i>aa<sub>1</sub></i>
<i>AA</i> × <i>AA</i>	0,30 × 0,30 = 0,09	0,09	-	-	-	-	-
<i>AA</i> × <i>aa</i>	2 × 0,30 × 0,10 = 0,06	-	-	-	0,06	-	-
<i>AA</i> × <i>a<sub>1</sub>a<sub>1</sub></i>	2 × 0,30 × 0,60 = 0,36	-	-	-	-	0,36	-
<i>aa</i> × <i>aa</i>	0,10 × 0,10 = 0,01	-	0,01	-	-	-	-
<i>aa</i> × <i>a<sub>1</sub>a<sub>1</sub></i>	2 × 0,10 × 0,60 = 0,12	-	-	-	-	-	0,12
<i>a<sub>1</sub>a<sub>1</sub></i> × <i>a<sub>1</sub>a<sub>1</sub></i>	0,60 × 0,60 = 0,36	-	-	0,36	-	-	-
Total	1,00	0,09 ( <i>p</i> <sup>2</sup> )	0,01 ( <i>q</i> <sup>2</sup> )	0,36 ( <i>r</i> <sup>2</sup> )	0,06 (2 <i>pq</i> )	0,36 (2 <i>pr</i> )	0,12 (2 <i>qr</i> )

A Tabela 2.2 mostra a distribuição de famílias em uma população teórica que está em equilíbrio de Hardy e Weinberg em relação aos genótipos decorrentes de três alelos autossômicos *A*, *a* e *a<sub>1</sub>* com frequências *p*, *q* e *r*.

Tabela 2.2. Distribuição das famílias em uma população teórica que está em equilíbrio genético em relação aos genótipos determinados por três alelos autossômicos *A*, *a* e *a<sub>1</sub>* cujas frequências são *p*, *q* e *r*.

Casais		Filhos					
Tipo	Frequência	<i>AA</i>	<i>aa</i>	<i>a<sub>1</sub>a<sub>1</sub></i>	<i>Aa</i>	<i>Aa<sub>1</sub></i>	<i>aa<sub>1</sub></i>
<i>AA</i> × <i>AA</i>	<i>p</i> <sup>2</sup> · <i>p</i> <sup>2</sup> = <i>p</i> <sup>4</sup>	<i>p</i> <sup>4</sup>	-	-	-	-	-
<i>AA</i> × <i>aa</i>	2 <i>p</i> <sup>2</sup> · <i>q</i> <sup>2</sup> = 2 <i>p</i> <sup>2</sup> <i>q</i> <sup>2</sup>	-	-	-	2 <i>p</i> <sup>2</sup> <i>q</i> <sup>2</sup>	-	-
<i>AA</i> × <i>a<sub>1</sub>a<sub>1</sub></i>	2 <i>p</i> <sup>2</sup> · <i>r</i> <sup>2</sup> = 2 <i>p</i> <sup>2</sup> <i>r</i> <sup>2</sup>	-	-	-	-	2 <i>p</i> <sup>2</sup> <i>r</i> <sup>2</sup>	-
<i>AA</i> × <i>Aa</i>	2 <i>p</i> <sup>2</sup> · 2 <i>pq</i> = 4 <i>p</i> <sup>3</sup> <i>q</i>	2 <i>p</i> <sup>3</sup> <i>q</i>	-	-	2 <i>p</i> <sup>3</sup> <i>q</i>	-	-
<i>AA</i> × <i>Aa<sub>1</sub></i>	2 <i>p</i> <sup>2</sup> · 2 <i>pr</i> = 4 <i>p</i> <sup>3</sup> <i>r</i>	2 <i>p</i> <sup>3</sup> <i>r</i>	-	-	-	2 <i>p</i> <sup>3</sup> <i>r</i>	-
<i>AA</i> × <i>aa<sub>1</sub></i>	2 <i>p</i> <sup>2</sup> · 2 <i>qr</i> = 4 <i>p</i> <sup>2</sup> <i>qr</i>	-	-	-	2 <i>p</i> <sup>2</sup> <i>qr</i>	2 <i>p</i> <sup>2</sup> <i>qr</i>	-
<i>aa</i> × <i>aa</i>	<i>q</i> <sup>2</sup> · <i>q</i> <sup>2</sup> = <i>q</i> <sup>4</sup>	-	<i>q</i> <sup>4</sup>	-	-	-	-
<i>aa</i> × <i>a<sub>1</sub>a<sub>1</sub></i>	2 <i>q</i> <sup>2</sup> · <i>r</i> <sup>2</sup> = 2 <i>q</i> <sup>2</sup> <i>r</i> <sup>2</sup>	-	-	-	-	-	2 <i>q</i> <sup>2</sup> <i>r</i> <sup>2</sup>
<i>aa</i> × <i>Aa</i>	2 <i>q</i> <sup>2</sup> · 2 <i>pq</i> = 4 <i>pq</i> <sup>3</sup>	-	2 <i>pq</i> <sup>3</sup>	-	2 <i>pq</i> <sup>3</sup>	-	-
<i>aa</i> × <i>Aa<sub>1</sub></i>	2 <i>q</i> <sup>2</sup> · 2 <i>pr</i> = 4 <i>pq</i> <sup>2</sup> <i>r</i>	-	-	-	2 <i>pq</i> <sup>2</sup> <i>r</i>	-	2 <i>pq</i> <sup>2</sup> <i>r</i>
<i>aa</i> × <i>aa<sub>1</sub></i>	2 <i>q</i> <sup>2</sup> · 2 <i>qr</i> = 4 <i>q</i> <sup>3</sup> <i>r</i>	-	2 <i>q</i> <sup>3</sup> <i>r</i>	-	-	-	2 <i>q</i> <sup>3</sup> <i>r</i>
<i>a<sub>1</sub>a<sub>1</sub></i> × <i>a<sub>1</sub>a<sub>1</sub></i>	<i>r</i> <sup>2</sup> · <i>r</i> <sup>2</sup> = <i>r</i> <sup>4</sup>	-	-	<i>r</i> <sup>4</sup>	-	-	-
<i>a<sub>1</sub>a<sub>1</sub></i> × <i>Aa</i>	2 <i>r</i> <sup>2</sup> · 2 <i>pq</i> = 4 <i>pqr</i> <sup>2</sup>	-	-	-	-	2 <i>pqr</i> <sup>2</sup>	2 <i>pqr</i> <sup>2</sup>
<i>a<sub>1</sub>a<sub>1</sub></i> × <i>Aa<sub>1</sub></i>	2 <i>r</i> <sup>2</sup> · 2 <i>pr</i> = 4 <i>pr</i> <sup>3</sup>	-	-	2 <i>pr</i> <sup>3</sup>	-	2 <i>pr</i> <sup>3</sup>	-
<i>a<sub>1</sub>a<sub>1</sub></i> × <i>aa<sub>1</sub></i>	2 <i>r</i> <sup>2</sup> · 2 <i>qr</i> = 4 <i>qr</i> <sup>3</sup>	-	-	2 <i>qr</i> <sup>3</sup>	-	-	2 <i>qr</i> <sup>3</sup>
<i>Aa</i> × <i>Aa</i>	2 <i>pq</i> · 2 <i>pq</i> = 4 <i>p</i> <sup>2</sup> <i>q</i> <sup>2</sup>	<i>p</i> <sup>2</sup> <i>q</i> <sup>2</sup>	<i>p</i> <sup>2</sup> <i>q</i> <sup>2</sup>	-	2 <i>p</i> <sup>2</sup> <i>q</i> <sup>2</sup>	-	-
<i>Aa</i> × <i>Aa<sub>1</sub></i>	2 · 2 <i>pq</i> · 2 <i>pr</i> = 8 <i>p</i> <sup>2</sup> <i>qr</i>	2 <i>p</i> <sup>2</sup> <i>qr</i>	-	-	2 <i>p</i> <sup>2</sup> <i>qr</i>	2 <i>p</i> <sup>2</sup> <i>qr</i>	2 <i>p</i> <sup>2</sup> <i>qr</i>
<i>Aa</i> × <i>aa<sub>1</sub></i>	2 · 2 <i>pq</i> · 2 <i>qr</i> = 8 <i>pq</i> <sup>2</sup> <i>r</i>	-	2 <i>pq</i> <sup>2</sup> <i>r</i>	-	2 <i>pq</i> <sup>2</sup> <i>r</i>	2 <i>pq</i> <sup>2</sup> <i>r</i>	2 <i>pq</i> <sup>2</sup> <i>r</i>
<i>Aa<sub>1</sub></i> × <i>Aa<sub>1</sub></i>	2 <i>pr</i> · 2 <i>pr</i> = 4 <i>p</i> <sup>2</sup> <i>r</i> <sup>2</sup>	<i>p</i> <sup>2</sup> <i>r</i> <sup>2</sup>	-	<i>p</i> <sup>2</sup> <i>r</i> <sup>2</sup>	-	2 <i>p</i> <sup>2</sup> <i>r</i> <sup>2</sup>	-
<i>Aa<sub>1</sub></i> × <i>aa<sub>1</sub></i>	2 · 2 <i>pr</i> · 2 <i>qr</i> = 8 <i>pqr</i> <sup>2</sup>	-	-	2 <i>pqr</i> <sup>2</sup>	2 <i>pqr</i> <sup>2</sup>	2 <i>pqr</i> <sup>2</sup>	2 <i>pqr</i> <sup>2</sup>
<i>aa<sub>1</sub></i> × <i>aa<sub>1</sub></i>	2 <i>qr</i> · 2 <i>qr</i> = 4 <i>q</i> <sup>2</sup> <i>r</i> <sup>2</sup>	-	<i>q</i> <sup>2</sup> <i>r</i> <sup>2</sup>	<i>q</i> <sup>2</sup> <i>r</i> <sup>2</sup>	-	-	2 <i>q</i> <sup>2</sup> <i>r</i> <sup>2</sup>
Total	( <i>p</i> + <i>q</i> + <i>r</i> ) <sup>4</sup> = 1	<i>p</i> <sup>2</sup>	<i>q</i> <sup>2</sup>	<i>r</i> <sup>2</sup>	2 <i>pq</i>	2 <i>pr</i>	2 <i>qr</i>

### CÁLCULO DA FREQUÊNCIA GÊNICA EM CASOS DE POLIALLELISMO AUTOSSÔMICO QUANDO EXISTE CODOMINÂNCIA

Nos casos de polialelismo em que não existe relação de dominância e recessividade entre os fenótipos, as frequências gênicas podem ser facilmente calculadas por contagem. Assim, por exemplo, num caso de trialelismo, levando em conta os alelos  $A$ ,  $a$  e  $a_1$ , teríamos que considerar a existência de seis genótipos  $AA$ ,  $aa$ ,  $a_1a_1$ ,  $Aa$ ,  $Aa_1$ ,  $aa_1$ . Em uma amostra de  $n$  indivíduos podemos calcular as frequências de um gene após somar o número de vezes em que ele aparece nos diferentes genótipos e dividir a soma pelo número total de genes, isto é, por  $2n$ . Tem-se, desse modo, que as frequências  $p$ ,  $q$  e  $r$  dos alelos  $A$ ,  $a$  e  $a_1$  podem ser obtidas por intermédio de:

$$p = \frac{2AA + Aa + Aa_1}{2n}$$

$$q = \frac{2aa + Aa + aa_1}{2n}$$

$$r = \frac{2a_1a_1 + Aa_1 + aa_1}{2n}$$

Se em vez de trabalhar com o número de genótipos quisermos trabalhar com a sua frequência relativa, assinaladas em **negrito**, o denominador dessas fórmulas se reduzirá ao número 2, porque, nesse caso, o total da amostra passa a ser a unidade. Nesse caso, as frequências  $p$ ,  $q$  e  $r$  dos alelos  $A$ ,  $a$  e  $a_1$  são obtidas por intermédio de:

$$p = \frac{2AA + Aa + Aa_1}{2}$$

$$q = \frac{2aa + Aa + aa_1}{2}$$

$$r = \frac{2a_1a_1 + Aa_1 + aa_1}{2}$$

Tomemos um exemplo numérico a respeito da fosfatase ácida eritrocitária, cujos alelos mais frequentes são os genes  $P^a$ ,  $P^b$  e  $P^c$  responsáveis pelos fenótipos A, B, C, AB, AC e BC os quais correspondem, respectivamente, aos genótipos  $P^aP^a$ ,  $P^bP^b$ ,  $P^cP^c$ ,  $P^aP^b$ ,  $P^aP^c$ ,  $P^bP^c$ . Assim, consideremos que numa amostra de 480 indivíduos caucásóides observou-se a seguinte distribuição fenotípica, com as frequências em porcentagem assinaladas entre parênteses:

$$A = 35 (7,3); B = 256 (53,3); C = 0 (0); AB = 161 (33,5); AC = 8 (1,7); BC = 20 (4,2)$$

Se levássemos em conta o número de indivíduos, as frequências  $p$ ,  $q$  e  $r$  dos alelos  $P^a$ ,  $P^b$  e  $P^c$  seriam calculadas como abaixo.

$$p = P^a = \frac{(2 \times 35) + 161 + 8}{2 \times 480} = 0,249 \text{ ou } 24,9\%$$

$$q = P^b = \frac{(2 \times 256) + 161 + 20}{2 \times 480} = 0,722 \text{ ou } 72,2\%$$

$$r = P^c = \frac{8 + 20}{2 \times 480} = 0,029 \text{ ou } 2,9\%$$

Levando em conta as frequências relativas dos fenótipos, que correspondem aos genótipos, chegaríamos aos mesmos resultados a partir das operações abaixo:

$$p = P^a = \frac{(2 \times 0,073) + 0,335 + 0,017}{2} = 0,249 \text{ ou } 24,9\%$$

$$q = P^b = \frac{(2 \times 0,533) + 0,335 + 0,042}{2} = 0,722 \text{ ou } 72,2\%$$

$$r = P^c = \frac{0,017 + 0,042}{2} = 0,029 \text{ ou } 2,9\%$$

Para investigar se a amostra estudada representa uma população em equilíbrio de Hardy e Weinberg basta calcular o número de indivíduos nas seis classes fenotípicas esperadas segundo  $A = np^2$ ;  $B = nq^2$ ;  $C = nr^2$ ;  $AB = n2pq$ ;  $AC = n2pr$ ;  $BC = n2qr$ . O qui-quadrado resultante tem três graus de liberdade, porque para calcular as seis classes esperadas são necessárias três informações, ou seja, o total da amostra e a frequência de dois dos três alelos, visto que, conhecendo as frequências de dois deles, a do terceiro alelo será automaticamente conhecida, pois  $p+q+r = 1$ . Em decorrência do qui-quadrado obtido (2,396) com três graus de liberdade, pode-se concluir que a amostra analisada representa uma população em equilíbrio de Hardy e Weinberg quanto aos genótipos da fosfatase ácida eritrocitária.

Valor	A	B	C	AB	AC	BC	Total
Observado (o)	35	256	0	161	8	20	480
Esperado (e)	29,8	250,2	0,4	172,6	6,9	20,1	480
$\frac{(o - e)^2}{e}$	0,907	0,134	0,4	0,780	0,175	0,0005	$\chi^2_{(3)} = 2,396$ $0,30 < P < 0,50$

### CÁLCULO DA FREQUÊNCIA GÊNICA EM CASOS DE POLIALELISMO AUTOSSÔMICO QUANDO EXISTE DOMINÂNCIA E RECESSIVIDADE

Dentre os casos de polialelismo autossômico, a situação que parece mais interessante para ser aqui discutida é a de trialelismo, com um dos alelos se manifestando apenas em homozigose, como ocorre no sistema sanguíneo ABO clássico. Nesse sistema levamos em conta os alelos  $A$ ,  $B$  e  $O$ , sendo que o gene  $A$  determina a

produção do antígeno A, o gene *B* a produção do antígeno B, enquanto o gene *O*, quando em homozigose, é responsável pela falta dos antígenos A e B. Esses alelos, sabe-se atualmente, pertencem a um loco situado no braço inferior do cromossomo 9, mais precisamente em *9q31.3-qter* (Cook *et al.*, 1978; Narahara *et al.*, 1986). Os alelos *A* e *B* mantêm uma relação de codominância entre si, de modo que os heterozigotos *AB* são responsáveis pelo grupo sanguíneo AB. Esses dois alelos apresentam relação de dominância sobre o alelo *O*. Por isso, os indivíduos homozigotos *AA* ou os heterozigotos *AO* apresentam grupo sanguíneo A, aqueles com genótipos homozigoto *BB* e heterozigoto *BO* apresentam grupo sanguíneo B, enquanto os indivíduos do grupo sanguíneo O são os homozigotos *OO*.

Designando as freqüências dos alelos *A*, *B* e *O* por, respectivamente, *p*, *q* e *r*, as proporções genóticas e fenóticas do sistema ABO esperadas nas populações em equilíbrio genético estável são aquelas calculadas por intermédio da expansão do trinômio  $(p + q + r)^2$ , o que permite apresentar um quadro como o da Tabela 3.2.

Tabela 3.2 Freqüências genóticas e fenóticas do sistema ABO clássico que são esperadas em populações em equilíbrio genético estável, sendo *p*, *q* e *r* as freqüências dos genes *A*, *B* e *O*.

Grupo Sanguíneo (Fenótipo)	Genótipos	Freqüência Genotípica	Freqüência Fenotípica
A	<i>AA</i>	$p^2$	$p^2 + 2pr$
	<i>AO</i>	$2pr$	
B	<i>BB</i>	$q^2$	$q^2 + 2qr$
	<i>BO</i>	$2qr$	
AB	<i>AB</i>	$2pq$	$2pq$
O	<i>OO</i>	$r^2$	$r^2$

Com base nos elementos fornecidos na Tabela 3.2 é possível estimar as freqüências populacionais dos alelos do sistema ABO, a partir do conhecimento da distribuição fenotípica, isto é, de seus grupos sanguíneos. As melhores estimativas dessas freqüências são obtidas pelo método dos escores de verossimilhança máxima, com o auxílio de programas computacionais (Cabello e Krieger, 1997). Essas estimativas, entretanto, também podem ser obtidas de modo fácil com o emprego de métodos algébricos, dentre os quais, o que fornece valores muito semelhantes aos calculados pelo método dos escores de verossimilhança máxima é o proposto por Bernstein (1924,1925,1930). O próprio Stevens (1950) que, pela primeira vez, aplicou o método da verossimilhança máxima para estimar as freqüências dos genes do sistema ABO (Stevens, 1938), reconheceu que o método de Bernstein é muito eficiente, apesar

de sua simplicidade.

No método de Bernstein parte-se do princípio de que, se a amostra representar uma população em equilíbrio de Hardy e Weinberg, ter-se-á  $(p+q+r)^2 = 1$  e, para obter a frequência do alelo  $O$ , basta extrair a raiz quadrada da frequência observada de indivíduos do grupo  $O$ , pois,  $r = \sqrt{r^2} = \sqrt{O}$ . Depois disso, deve-se calcular:

$$p = 1 - \sqrt{B + O} = 1 - \sqrt{q^2 + 2qr + r^2} = 1 - (q + r)$$

$$q = 1 - \sqrt{A + O} = 1 - \sqrt{p^2 + 2pr + r^2} = 1 - (p + r)$$

Visto que a soma dessas estimativas raramente fornece um valor exatamente igual a 1, elas são chamadas de *estimativas preliminares* e representadas por  $p'$ ,  $q'$  e  $r'$ . Para corrigi-las deve-se calcular o desvio ( $D$ ) entre a unidade e a soma das estimativas, isto é, é necessário obter:

$$D = 1 - (p' + q' + r')$$

e calcular as *estimativas corrigidas*  $p$ ,  $q$  e  $r$  por intermédio de:

$$p = p' \left(1 + \frac{D}{2}\right)$$

$$q = q' \left(1 + \frac{D}{2}\right)$$

$$r = \left(r' + \frac{D}{2}\right) \left(1 + \frac{D}{2}\right) \text{ ou } r = 1 - (p + q)$$

Aplicamos essas fórmulas aos dados a respeito de uma amostra de 2.571 brasileiros caucasoídes que apresentaram a seguinte distribuição de grupos sanguíneos do sistema ABO clássico, com os valores em porcentagem entre parênteses:

$$A = 1001 (38,93); B = 250 (9,72); AB = 92 (3,58); O = 1228 (47,76)$$

Com base nesses dados, calcularíamos:

$$p' = 1 - \sqrt{B + O} = 1 - \sqrt{0,0972 + 0,4776} = 0,2418$$

$$q' = 1 - \sqrt{A + O} = 1 - \sqrt{0,3893 + 0,4776} = 0,0689$$

$$r' = \sqrt{O} = \sqrt{0,4776} = 0,6911$$

$$\text{Total} = 1,0018$$

$$D = 1 - 1,0018 = -0,0018$$

$$\frac{D}{2} = -0,0009$$

$$1 + \frac{D}{2} = 0,9991$$

$$p = 0,2418 \times 0,9991 = 0,242$$

$$q = 0,0689 \times 0,9991 = 0,069$$

$$r = 1 - (0,242 + 0,069) = 0,689$$

Para verificar se a distribuição dos fenótipos na amostra estudada está de acordo com a hipótese genética, isto é, se a amostra representa, realmente, uma população em equilíbrio de Hardy e Weinberg, existe a alternativa de calcular um qui-quadrado a partir da razão entre o quadrado do desvio  $D$  e a sua variância. Visto que a variância do desvio  $D$  é  $\sigma^2_{(D)} = \frac{1}{2n\left(1 + \frac{r}{pq}\right)}$  onde  $n$  é o tamanho amostral, o qui-quadrado pode ser

obtido como abaixo:

$$\chi^2 = 2n\left(1 + \frac{r}{pq}\right)D^2$$

Esse qui-quadrado terá apenas um grau de liberdade porque existem quatro classes fenotípicas (grupos A, B, AB e O) e são necessárias três informações da amostra para calcular os valores esperados nessas classes, isto é, o tamanho da amostra e as frequências de dois genes. É importante salientar que nessa fórmula de qui-quadrado não é necessário empregar as estimativas corrigidas das frequências gênicas, podendo-se partir das estimativas preliminares. Usando as estimativas preliminares das frequências dos alelos  $A$ ,  $B$  e  $O$  da amostra caucasóide de nosso exemplo obtemos a indicação de que ela representa uma população em equilíbrio de Hardy e Weinberg, pois:

$$\chi^2 = 2 \times 2.571 \left(1 + \frac{0,6911}{0,2418 \times 0,0689}\right) 0,00000324 = 0,707; 1 \text{ G.L.}; 0,30 < P < 0,50.$$

Uma outra alternativa é a de calcular um qui-quadrado depois de comparar os números observados com os esperados em cada classe, isto é, a partir de:

$$\chi^2 = \frac{[A - (p^2 + 2pr)n]^2}{(p^2 + 2pr)n} + \frac{[B - (q^2 + 2qr)n]^2}{(q^2 + 2qr)n} + \frac{(AB - 2pqn)^2}{2pqn} + \frac{(O - r^2n)^2}{r^2n}$$

onde A, B, AB e O são os números observados de indivíduos em cada um desses grupos sanguíneos. Assim, teríamos:

$$\chi^2 = \frac{(1001 - 1007,93)^2}{1007,93} + \frac{(250 - 256,70)^2}{256,70} + \frac{(92 - 85,86)^2}{85,86} + \frac{(1228 - 1220,51)^2}{1220,51} = 0,707$$

Como se pode constatar, os dois valores de qui-quadrado obtidos são idênticos, o que indica que qualquer das duas alternativas para calcular o qui-quadrado pode ser aceita. A primeira alternativa é menos trabalhosa, mas, nos casos em que a amostra não representa uma população em equilíbrio genético, a segunda alternativa é melhor do que a primeira porque permite detectar qual(ais) o(s) grupo(s) sanguíneo(s) que mostram

uma diferença significativa entre o observado e o esperado. Assim, por exemplo, se a amostra estudada não estivesse em equilíbrio por causa de um excesso de indivíduos do grupo sanguíneo AB, a aplicação da fórmula da primeira alternativa indicaria apenas que a amostra não representa uma população em equilíbrio genético, mas nada informaria a respeito do número excessivo observado de indivíduos do grupo sanguíneo AB.

Um bom expediente para facilitar o cálculo do qui-quadrado usando a segunda alternativa é o de trabalhar apenas com as freqüências relativas, deixando de multiplicá-las pelo tamanho da amostra (*n*) **mas não se deve esquecer, ao final das operações, de multiplicar o valor do qui-quadrado obtido pelo tamanho amostral**, porque o valor obtido é igual ao do qui-quadrado dividido por *n*. Assim, usando os dados de nosso exemplo, teríamos:

$$\frac{(0,3893 - 0,3920)^2}{0,3920} + \frac{(0,0972 - 0,0998)^2}{0,0998} + \frac{(0,0358 - 0,0334)^2}{0,0334} + \frac{(0,4776 - 0,4747)^2}{0,4747} = 0,000276$$

e, multiplicando esse resultado por 2.571, obtém-se um valor de qui-quadrado muito próximo aos calculados anteriormente, pois  $0,000276 \times 2.571 = 0,7096$ ;  $0,30 < P < 0,50$ .

Quando os subgrupos sanguíneos A<sub>1</sub>, A<sub>2</sub>, A<sub>1</sub>B e A<sub>2</sub>B são levados em conta, o sistema ABO passa a incluir seis fenótipos explicados à custa de dez genótipos que, numa população em equilíbrio genético se distribuem como na Tabela 4.2.

Tabela 4.2. Freqüências genotípicas e fenotípicas do sistema ABO em uma população em equilíbrio genético, quando se consideram os subgrupos A<sub>1</sub>, A<sub>2</sub>, A<sub>1</sub>B e A<sub>2</sub>B.

Fenótipo	Genótipo	Freqüência Genotípica	Freqüência Fenotípica
A <sub>1</sub>	A <sub>1</sub> A <sub>1</sub>	p <sub>1</sub> <sup>2</sup>	p <sub>1</sub> <sup>2</sup> + 2p <sub>1</sub> (p <sub>2</sub> +r)
	A <sub>1</sub> A <sub>2</sub>	2p <sub>1</sub> p <sub>2</sub>	
	A <sub>1</sub> O	2p <sub>1</sub> r	
A <sub>2</sub>	A <sub>2</sub> A <sub>2</sub>	p <sub>2</sub> <sup>2</sup>	p <sub>2</sub> (p <sub>2</sub> +2r)
	A <sub>2</sub> O	2p <sub>2</sub> r	
B	BB	q <sup>2</sup>	q(q+2r)
	BO	2qr	
A <sub>1</sub> B	A <sub>1</sub> B	2p <sub>1</sub> q	2p <sub>1</sub> q
A <sub>2</sub> B	A <sub>2</sub> B	2p <sub>2</sub> q	2p <sub>2</sub> q
O	OO	r <sup>2</sup>	r <sup>2</sup>

As freqüências preliminares p<sub>1</sub>', p<sub>2</sub>', q' e r' dos alelos A<sub>1</sub>, A<sub>2</sub>, B e O são estimadas a partir das freqüências fenotípicas A<sub>1</sub>, A<sub>2</sub>, A<sub>2</sub>B e O como abaixo:

$$p_1' = 1 - \sqrt{A_2 + B + A_2B + O}$$

$$p_2' = \sqrt{A_2 + B + A_2B + O} - \sqrt{B + O}$$

$$q' = 1 - \sqrt{A_1 + A_2 + O}$$

$$r' = \sqrt{O}$$

Depois de determinar o desvio (D) entre a unidade e a soma das estimativas preliminares, isto é, depois de calcular:

$$D = 1 - (p_1' + p_2' + q' + r')$$

as estimativas corrigidas  $p_1$ ,  $p_2$  e  $q$  das frequências dos genes  $A_1$ ,  $A_2$  e  $B$  serão obtidas pela multiplicação das estimativas preliminares  $p_1'$ ,  $p_2'$  e  $q'$  por  $\left(1 + \frac{D}{2}\right)$ . A estimativa corrigida  $r$  do alelo  $O$  pode ser obtida simplesmente pela diferença  $1 - (p_1 + p_2 + q)$  ou calculando  $\left(r' + \frac{D}{2}\right)\left(1 + \frac{D}{2}\right)$ .

Para estimar as frequências dos genes do sistema ABO em populações com o alelo  $A_{int}$  pode-se fazer o cálculo das estimativas preliminares como abaixo:

$$\begin{aligned} p_1' &= 1 - \sqrt{A_{int} + A_2 + B + A_{int}B + A_2B + O} \\ p_{int}' &= \sqrt{A_{int} + A_2 + B + A_{int}B + A_2B + O} - \sqrt{A_2 + B + A_2B + O} \\ p_2' &= \sqrt{A_2 + B + A_2B + O} - \sqrt{B + O} \\ q' &= 1 - \sqrt{A_1 + A_2 + A_{int} + O} \\ r' &= \sqrt{O} \end{aligned}$$

Depois de calcular o desvio  $D = 1 - (p_1' + p_{int}' + p_2' + q' + r')$ , todas as estimativas corrigidas, com exceção de  $r$ , serão obtidas pela multiplicação das estimativas preliminares por  $\left(1 + \frac{D}{2}\right)$ . A estimativa corrigida  $r$  será obtida por intermédio de  $1 - (p_1 + p_{int} + p_2 + q)$  ou por intermédio de  $\left(r' + \frac{D}{2}\right)\left(1 + \frac{D}{2}\right)$ .

## ANOMALIAS RECESSIVAS E CÁLCULO DAS FREQUÊNCIAS GÊNICAS

Não é raro termos que lidar com anomalias recessivas resultantes da homozigose ou heterozigose de diferentes mutações. Em outras palavras, não é raro nos depararmos com várias mutações, que poderemos designar por  $a$ ,  $a_1$ ,  $a_2$ ,  $a_3$  etc. de um gene “normal”  $A$  e que tanto os homozigotos  $aa$ ,  $a_1a_1$ ,  $a_2a_2$  ou  $a_3a_3$  quanto os heterozigotos  $aa_1$ ,  $aa_2$ ,  $aa_3$ ,  $a_1a_2$ ,  $a_1a_3$  ou  $a_2a_3$  determinem um quadro patológico semelhante. Em situações como essa, em que os diferentes genótipos podem ser detectados por técnicas de genética molecular, mas que, do ponto de vista clínico, resultam em um quadro semelhante, podemos, para fins de epidemiológicos, reunir os homozigotos  $aa$ ,  $a_1a_1$ ,

$a_2a_2$ ,  $a_3a_3$  e os heterozigotos  $aa_1$ ,  $aa_2$ ,  $aa_3$ ,  $a_1a_2$ ,  $a_1a_3$  e  $a_2a_3$  em uma única categoria, designada por  $aa$ .

Para exemplificar, consideremos o caso da *fibrose cística*, que é o defeito genético mais freqüente em populações caucasóides, decorrente de muitas mutações caracterizáveis por técnicas de estudo de DNA. Considerando, porém, que essas mutações em homozigose ou em heterozigose dos diferentes mutantes resulta num quadro patológico de fibrose cística, podemos tratá-las em conjunto para fins epidemiológicos. Assim, numa população caucasóide na qual a fibrose cística ocorre em 4 de cada 10.000 nascimentos, poderíamos estimar que, dentre cerca de 25 indivíduos dessa população, um seria portador (heterozigoto) de um alelo mutante responsável pela fibrose cística. Realmente, a partir do conhecimento da incidência de fibrose cística, calcularíamos  $q = \sqrt{0,0004} = 0,02$  e  $p = 1 - 0,02 = 0,98$ . Isso nos permitiria estimar a freqüência de portadores como sendo  $2pq = 0,0392$  ou  $1: 25,5 \cong 1: 25$ .

## COMPARAÇÃO DE VÁRIAS AMOSTRAS

Quando lidamos com polialelismo com codominância e queremos comparar várias amostras, a fim de investigar se elas podem ser reunidas em uma única ou se isso não deve ser feito em decorrência de uma diferença significativa entre elas, podemos levar em conta a distribuição gênica, como nos casos de dialelismo. Contudo, se as amostras representarem populações em equilíbrio de Hardy e Weinberg poder-se-á usar a alternativa de comparar as distribuições genótípicas dessas amostras. Para exemplificar, consideremos que em amostras de três populações estudadas quanto à fosfatase ácida eritrocitária tivéssemos encontrado os resultados apresentados na Tabela 5.2. Com esses dados pode-se elaborar a Tabela 6.2 e, pelo valor de qui-quadrado encontrado, concluir a favor da hipótese de que as três amostras podem ser reunidas como se fossem procedentes de uma mesma população, já que as diferenças entre as proporções dos três alelos estudados não diferem significativamente.

Tabela 5.2 Distribuição de três amostras populacionais quanto aos genótipos de fosfatase ácida eritrocitária.

Amostra	A	B	C	AB	AC	BC	Total
População 1	19	54	0	53	1	3	130
População 2	10	22	0	16	0	2	50
População 3	11	34	0	41	2	6	94

Tabela 6.2. Comparação das três amostras da Tabela 5.2 quanto às proporções dos alelos  $P^a$ ,  $P^b$  e  $P^c$ .

Amostra	$P^a$	$P^b$	$P^c$	Total
População 1	92	164	4	260
População 2	36	62	2	100
População 3	65	115	8	188

$$\chi^2_{(4)} = 3,402; 0,30 < P < 0,50$$

Nos casos de polialelismo em que existem relações de dominância e recessividade, a comparação de várias amostras pode ser feita usando o recurso de comparar as suas distribuições fenotípicas. No caso de trialelismo, como o dos grupos sanguíneos do sistema ABO clássico, Stevens (1950) idealizou um método mais preciso, que denominou de *método dos círculos equivalentes*, o qual permite a comparação da frequência dos alelos. Esse método foi publicado quando o eminente professor inglês, precocemente falecido, ministrava aulas de Estatística na Universidade de São Paulo. Ele leva em conta o sistema ABO, mas pode ser estendido para outros sistemas análogos

O método em questão é gráfico e permite a comparação rápida de um grande número de amostras, cada uma das quais é representada por um círculo, cujo raio é função da frequência  $r$  do gene  $O$ , ou da soma das frequências  $p$  e  $q$  dos alelos  $A$  e  $B$ . As coordenadas retangulares que servem para determinar a posição de cada círculo são dadas pelas raízes quadradas das frequências dos alelos  $A$  e  $B$  ( $\sqrt{p}$  e  $\sqrt{q}$ ).

No método de Stevens (1950) a hipótese de que duas ou mais amostras não diferem quanto às frequências gênicas do sistema ABO é aceita quando os círculos que as representam se interpenetram. Quando isso não ocorre, rejeita-se tal hipótese e aceita-se a alternativa de que as amostras têm frequências gênicas que diferem significativamente.

Para a obtenção dos raios dos círculos em função de  $r$ , Stevens (1950) preparou uma tabela que fornece os elementos para cálculo de raios de círculos que incluem diferentes probabilidades, parte da qual está apresentada na Tabela 7.2. Com base nos dados dessa tabela pode-se achar o raio do círculo equivalente se dividirmos pela raiz quadrada do tamanho amostral ( $\sqrt{n}$ ) o valor tabulado em função da frequência  $r$ . Assim, por exemplo, se o valor estimado de  $r$  em uma amostra de 100 indivíduos for 0,7, o raio do círculo de 80% de probabilidade será  $\frac{0,607}{10} = 0,0607$  e o de 95% de

probabilidade será  $\frac{0,829}{10} = 0,0829$ . Tais valores serão representados graficamente na

mesma escala utilizada para indicar os valores de  $\sqrt{p}$  e de  $\sqrt{q}$ .

Tabela 7.2. Elementos para o cálculo de raios de círculos equivalentes que incluem 80% e 95% de probabilidade em função da frequência  $r$  do gene  $O$  do sistema ABO.

$r$	80%	95%
0,5	0,581	0,793
0,6	0,596	0,813
0,7	0,607	0,829
0,8	0,617	0,842
0,9	0,626	0,854
1,0	0,634	0,865

Se o valor de  $r$  não estiver representado na Tabela 7.2, poder-se-á fazer uma interpolação para obter o raio do círculo equivalente. Assim, por exemplo, se em uma amostra de 100 indivíduos o valor de  $r$  for estimado em 0,65 considerar-se-á que, quando  $r$  é igual a 0,6 o valor tabulado para obter o raio do círculo de 80% de probabilidade é 0,596, sendo igual a 0,607 quando  $r$  é igual a 0,7. Assim, a um aumento de 0,1 no valor de  $r$  corresponde um aumento de 0,011 no valor tabulado. Portanto, o valor 0,0055 deve corresponder a um aumento de 0,05, isto é, à diferença entre 0,65 e 0,6. Desse modo, o raio do círculo equivalente de 80% de probabilidade será  $\frac{0,596 + 0,0055}{10} = 0,0602$ . No caso de se querer obter o raio do círculo equivalente de 95% de probabilidade sendo o valor de  $r$  igual a 0,65 ter-se-ia, pelos mesmos motivos  $\frac{0,813 + 0,008}{10} = 0,0821$ . Do ponto de vista prático, contudo, as interpolações para a obtenção dos valores dos raios dos círculos equivalentes constituem apenas refinamentos. Portanto, do ponto de vista prático, quando  $r$  não estiver representado na Tabela 7.2 bastará tomar o seu valor mais próximo.

Para entender melhor a aplicação do método de Stevens (1950), consideremos os dados apresentados na Tabela 8.2 a respeito de cinco amostras, distribuídas segundo os grupos sanguíneos do sistema ABO clássico (Beiguelman e Marchi, 1962; Beiguelman, 1963). Os dados dessa tabela permitem obter os elementos necessários à construção dos círculos equivalentes, de modo que, a partir deles pode-se preparar uma outra tabela como a Tabela 9.2 e desenhar os gráficos das Figuras 1.2 e 2.2.

Tabela 8.2. Distribuição de cinco amostras segundo os grupos sanguíneos do sistema ABO clássico e estimativas corrigidas das frequências dos alelos *A*, *B* e *O*. Entre parênteses estão assinalados os valores em porcentagem. 1. Brasileiros caucasóides com ancestrais brasileiros até bisavós; 2. brasileiros negróides; 3. italianos e brasileiros descendentes não-miscigenados de italianos; 4. espanhóis e brasileiros descendentes não miscigenados de espanhóis; 5. japoneses e descendentes não miscigenados de japoneses.

Amostra	A	B	AB	O	Total	<i>p</i>	<i>q</i>	<i>r</i>
1	228 (37,44)	52 (8,54)	15 (2,46)	314 (51,56)	609 (100)	0,225	0,057	0,718
2	56 (31,11)	23 (12,78)	8 (4,44)	93 (51,67)	180 (100)	0,196	0,090	0,714
3	276 (38,76)	69 (9,69)	33 (4,64)	334 (46,91)	712 (100)	0,247	0,074	0,679
4	47 (35,60)	13 (9,85)	5 (3,79)	67 (50,76)	132 (100)	0,221	0,071	0,708
5	111 (37,00)	68 (22,67)	40 (13,33)	81 (27,00)	300 (100)	0,293	0,198	0,509

Tabela 9.2. Elementos para a construção dos círculos equivalentes de 80% e 95% de probabilidade das amostras descritas na Tabela 7.

Amostra	$\sqrt{p}$	$\sqrt{q}$	$\sqrt{n}$	Raio (80%)	Raio (95%)
1	0,474	0,239	24,678	0,025	0,034
2	0,443	0,300	13,416	0,046	0,062
3	0,497	0,274	26,683	0,023	0,031
4	0,470	0,266	11,489	0,054	0,072
5	0,713	0,455	17,321	0,034	0,046

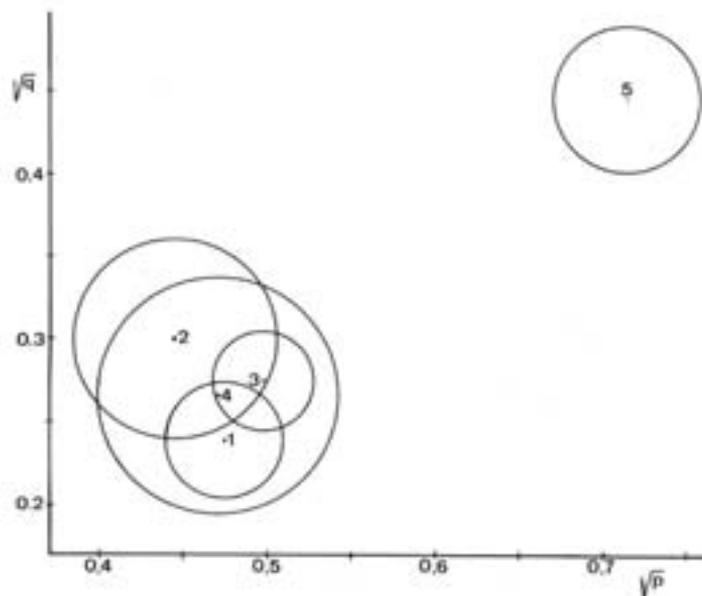


Fig.1.2. Círculos equivalentes de 95% de probabilidade das cinco amostras relacionadas na Tabela 8.2.

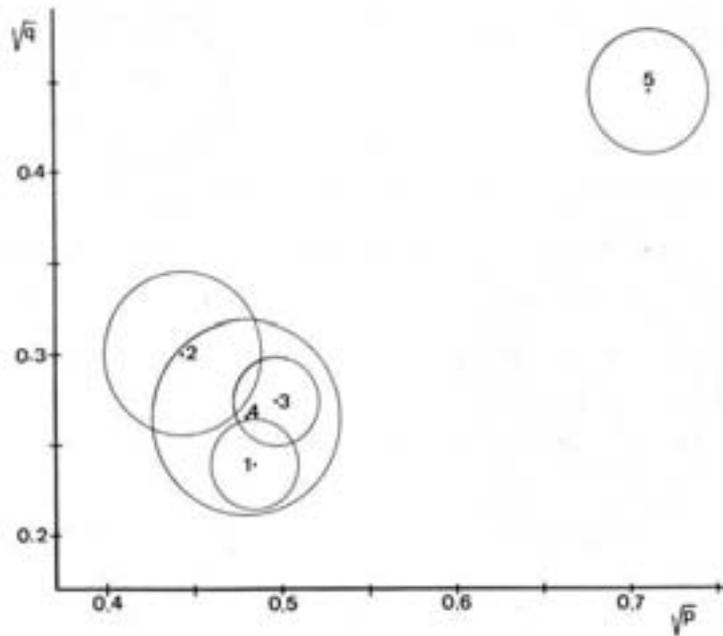


Fig.2.2. Círculos equivalentes de 80% de probabilidade das cinco amostras relacionadas na Tabela 8.2.

A análise da Fig.1.2 permite concluir que, no concernente às frequências gênicas do sistema ABO, as amostras 1, 2, 3 e 4 não diferem significativamente entre si, mas todas elas são bem diferentes da amostra 5, constituída por japoneses e seus descendentes não miscigenados. Um analista mais rigoroso construiria o gráfico da Fig.2.2 e, com base nos círculos equivalentes de 80% de probabilidade aceitaria que a amostra 2 difere da amostra 1 e da amostra 5, e que as amostras 1, 3 e 4 não diferem entre si, mas diferem da amostra 5.

**QUESTÕES E RESPOSTAS**

**Q 1.** Um sistema genético é constituído por quatro alelos autossômicos designados pelas letras *A*, *B*, *C* e *D*. Se em uma população as frequências desses alelos forem  $A = 10\%$ ,  $B = 20\%$ ,  $C = 30\%$  e  $D = 40\%$  quais as frequências genotípicas esperadas em equilíbrio genético?

**R 1.**

$$\begin{array}{ll} AA = 0,01 & AC = 0,06 \\ BB = 0,04 & AD = 0,08 \\ CC = 0,09 & BC = 0,12 \\ DD = 0,16 & BD = 0,16 \\ AB = 0,04 & CD = 0,24 \end{array}$$

**Q 2.** Os grupos sanguíneos do sistema ABO foram estudados em duas amostras de indivíduos caucasóides do Estado de São Paulo. Os resultados percentuais foram os apresentados abaixo. Com base neles:

- calcular as estimativas preliminares das frequências dos alelos *A*, *B* e *O*;
- verificar se as amostras representam populações em equilíbrio genético;
- calcular as estimativas corrigidas das frequências dos alelos *A*, *B* e *O*.

Amostra	No.	A	B	AB	O
1	712	38,76	9,69	4,64	46,91
2	609	37,44	8,54	2,46	51,56

**R 2. a)** Amostra 1

$$p' = 1 - \sqrt{0,0969 + 0,4691} = 0,2477$$

$$q' = 1 - \sqrt{0,3876 + 0,4691} = 0,0744$$

$$r' = \sqrt{0,4691} = 0,6849$$

$$D = -0,0070$$

Amostra 2

$$p' = 1 - \sqrt{0,0854 + 0,5156} = 0,2248$$

$$q' = 1 - \sqrt{0,3744 + 0,5156} = 0,0566$$

$$r' = \sqrt{0,5156} = 0,7181$$

$$D = 0,0005$$

b) As duas amostras representam populações em equilíbrio genético porque:

$$\text{Na amostra 1: } \chi^2_{(1)} = 2 \times 712 \left( 1 + \frac{0,5849}{0,2477 \times 0,0744} \right) 0,000049 = 2,667 ; 0,10 < P < 0,20.$$

$$\text{Na amostra 2: } \chi^2_{(1)} = 2 \times 609 \left( 1 + \frac{0,7181}{0,2248 \times 0,0566} \right) 0,00000025 = 0,017 ; 0,80 < P < 0,90$$

c) Amostra 1:  $p = 0,247$ ;  $q = 0,074$ ;  $r = 0,079$

Amostra 2:  $p = 0,225$ ;  $q = 0,057$ ;  $r = 0,718$

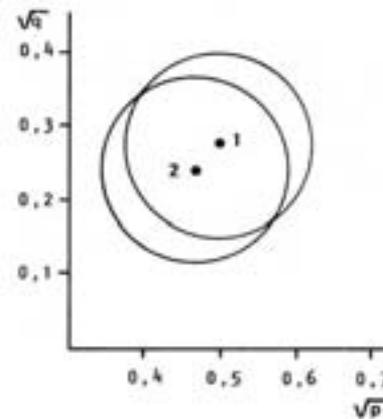
**Q 3.** Verificar graficamente se as duas amostras da questão anterior não diferem significativamente quanto às frequências dos genes do sistema ABO. Se elas não diferirem significativamente, calcular as estimativas combinadas das frequências dos alelos *A*, *B* e *O* nas amostras reunidas em uma única.

**R 3.** Tendo em vista a inexistência de diferença significativa entre as amostras estudadas pode-se calcular:

$$p = \frac{(712 \times 0,247) + (609 \times 0,225)}{1.321} = 0,237$$

$$q = \frac{(712 \times 0,074) + (609 \times 0,057)}{1.321} = 0,066$$

$$r = \frac{(712 \times 0,679) + (609 \times 0,718)}{1.321} = 0,697$$



**Q 4.** As amostras de sangue de praticamente todos os indivíduos (180) de uma comunidade ribeirinha de Rondônia (Portuchuelo, RO) foram investigadas eletroforeticamente para estudar a distribuição genotípica da fosfatase ácida eritrocitária, tendo-se observado o seguinte resultado:

$$AA = 2; BB = 118; CC = 0; AB = 50; AC = 1; BC = 9$$

Quer-se saber: as estimativas das freqüências dos alelos A, B e C, e se a distribuição genotípica observada pode ser considerada como representativa de uma população em equilíbrio de Hardy e Weinberg.

**R 4.** As freqüências  $p$ ,  $q$  e  $r$  dos alelos A, B e C podem ser obtidas como abaixo:

$$p = \frac{4 + 50 + 1}{360} = 0,153 \quad q = \frac{236 + 50 + 9}{360} = 0,819 \quad r = \frac{0 + 1 + 9}{360} = 0,028$$

Essas freqüências permitem calcular os valores esperados em cada classe genotípica e, com base no valor do qui-quadrado obtido (2,070) com três graus de liberdade, concluir que as proporções observadas não diferem significativamente daquelas esperadas em equilíbrio de Hardy e Weinberg

Valores	AA	BB	CC	AB	AC	BC
Obs. ( $o$ )	2	118	0	50	1	9
Esp. ( $e$ )	4,2	120,7	0,1	45,1	1,5	8,3
$\frac{(o - e)^2}{e}$	1,152	0,060	0,100	0,532	0,167	0,059

$$\chi^2_{(3)} = 2,070; 0,50 < P < 0,70.$$

**Q 5.** As hemácias de 862 indivíduos da população de Monte Negro, RO, foram tratadas com os anti-soros anti-Fy<sup>a</sup> e anti-Fy<sup>b</sup> do sistema sanguíneo Duffy, tendo sido observado o seguinte resultado:

Fenótipo	Fy(a+b-)	Fy(a-b+)	Fy(a+b+)	Fy(a-b-)
Genótipo	$Fy^a Fy^a$ ou $Fy^a Fy$	$Fy^b Fy^b$ ou $Fy^b Fy$	$Fy^a Fy^b$	$Fy Fy$
No.	236	328	243	55
%	27,38	38,05	28,19	6,38

Quer-se saber as estimativas das freqüências dos alelos  $Fy^a$ ,  $Fy^b$  e  $Fy$  e se a amostra estudada representa uma população em equilíbrio genético.

**R 5.** Aplicando o método de Bernstein para a análise dos dados do sistema Duffy, obtemos, inicialmente, as estimativas preliminares  $p'$ ,  $q'$  e  $r'$  dos alelos  $Fy^a$ ,  $Fy^b$  e  $Fy$ , bem como o desvio D:

$$p' = 0,3334$$

$$q' = 0,4190$$

$$r' = 0,2526$$

$$D = 1 - 1,005 = - 0,005$$

Com esses elementos obtemos as estimativas corrigidas  $p$ ,  $q$  e  $r$  :

$$p = 0,3326$$

$$q = 0,4180$$

$$r = 0,2494$$

e, pelo baixo valor do qui-quadrado ( $\chi^2_{(1)} = 0,120$ ;  $0,70 < P < 0,80$ ) concluímos que a amostra de Monte Negro representa uma população em equilíbrio genético em relação aos grupos sanguíneos do sistema Duffy.

## REFERÊNCIAS

Beiguelman, B. Grupos sanguíneos e lepra. *Rev. Bras. Lepr.* 31: 34-44, 1963.

Beiguelman, B. & Marchi, A. Blood groups among Japanese in Brazil. *Am. J. Phys. Anthropol.* 20: 29-31, 1962

Bernstein, F. Ergebnisse einer biostatistischen zusammenfassenden Betrachtung über die erblichen Blutstrukturen des Menschen. *Klin. Wschr.* 3: 1495-1497, 1924.

Bernstein, F. Zusammenfassende Betrachtungen über die erbliche Blutstrukturen des Menschen. *Z. indukt. Abstamm. u. VererbLehre.* 37: 237-270, 1925.

Bernstein, F. Fortgesetzte Untersuchungen aus der Theorie der Blutgruppen. *Z. indukt. Abstamm. u. VererbLehre* 56: 233-237, 1930.

Cabello, P.H. & Krieger, H. *Genioc. Sistema para análises de dados de Genética*. Publicação Técnica, Departamento de Genética, Instituto Oswaldo Cruz-FIOCRUZ, Rio de Janeiro, 1997.

Cook, P.J., Robson, E.B., Buckton, K.E., Slaughter, C.A., Gray, J.E., Blank, C.E., James, F.E., Ridler, M.A., Insley, J. & Hulten, M. Segregation of ABO, AK(1) and ACONs in families with abnormalities of chromosome 9. *Ann. Hum. Genet.* 41: 365-378, 1978.

Narahara, K., Takahashi, Y., Kikkawa, K., Wakita, Y., Kimura, S. & Kimoto, H. Assignment of ABO locus to 9q31.3qter by study of a family in which an intrachromosomal shift involving chromosome 9 is segregating. *Japan. J. Hum. Genet.* 31: 289-296, 1986.

Stevens, W.L. Estimation of blood-group gene frequencies. *Ann. Eugenics* 8: 362-375, 1938.

Stevens, W.L. Statistical analysis of the ABO blood groups. *Human Biol.* 22:191-217, 1950.

## CAPÍTULO 3. OUTROS TIPOS DE EQUILÍBRIO GENÉTICO

As considerações feitas nos capítulos anteriores a respeito de dialelismo e polialelismo autossômico não podem ser estendidas aos caracteres ligados ao sexo, nem aos caracteres poligênicos, nem a genes em ligação, como teremos oportunidade de verificar no presente capítulo.

### GENES DO CROMOSSOMO X

Na espécie humana, a esmagadora maioria das mulheres tem dois cromossomos X em seu cariótipo, de modo que, quando a população atinge equilíbrio genético, apenas elas podem apresentar os genótipos se distribuindo segundo  $(p + q)^2 = 1$ , nos casos de dialelismo, ou segundo  $(p + q + r + \dots + x)^2 = 1$  nos casos de polialelismo. Isso não pode acontecer nos homens, visto que eles, na esmagadora maioria, possuem cariótipo com um único cromossomo X, de modo que os genes desse cromossomo ficam em hemizigose. Como teremos oportunidade de constatar neste capítulo, se houver equilíbrio genético em relação aos caracteres monogênicos ligados ao sexo, as freqüências dos genótipos na população masculina deverão ser iguais às dos alelos que os determinam.

Um outro ponto interessante que deve ser levado em conta é que a estabilidade da distribuição genotípica em relação a genes do cromossomo X somente é atingida após uma única geração em panmixia quando, na geração inicial, as freqüências genotípicas dos homens correspondem às freqüências gênicas da população. *Se isso não ocorrer serão necessárias várias gerações em panmixia para que a estabilidade da distribuição genotípica seja alcançada.*

Para ilustrar essas afirmações consideremos um par de alelos  $A, a$  do cromossomo X e uma amostra de 110 mulheres e 100 homens de uma população teórica que apresente na geração inicial a seguinte distribuição genotípica:

Valor	Mulheres				Homens		
	$X^A X^A$	$X^A X^a$	$X^a X^a$	TOTAL	$X^A Y$	$X^a Y$	TOTAL
No.	56	20	34	110	60	40	100
%	50,9	18,2	30,9	100	60	40	100

As freqüências  $p$  e  $q$  dos alelos  $A$  e  $a$  do cromossomo X nessa população podem ser calculadas a partir da contagem do número de cromossomos X com esses alelos nos gametas que serviram para constituir essa geração inicial. Assim, para estimar a freqüência  $p$  do gene  $A$  poderemos, inicialmente, fazer a soma do número de homens que têm o genótipo  $X^A Y$  com o número de mulheres que possuem o genótipo heterozigoto  $X^A X^a$  e com o dobro do número de mulheres com o genótipo homozigoto  $X^A X^A$ , pois os dois cromossomos do cariótipo dessas últimas apresentam esse alelo. Em seguida, dividimos o resultado dessa contagem pelo número total de

cromossomos X nos gametas que deram origem à amostra. Esse número, evidentemente, é dado pela soma do total de indivíduos do sexo masculino (M) ao dobro do total de indivíduos do sexo feminino (F). Empregando os dados de nosso exemplo teríamos, pois:

$$p = \frac{X^A Y + X^A X^a + 2X^A X^A}{M + 2F} = \frac{60 + 20 + (2 \times 56)}{100 + (2 \times 110)} = 0,60$$

Para estimar a frequência  $q$  do alelo  $a$  ligado ao cromossomo X poderíamos calcular  $q = 1 - p = 1 - 0,60 = 0,40$  ou seguir o mesmo tipo de contagem anterior, isto é, calcular:

$$q = \frac{X^a Y + X^A X^a + 2X^a X^a}{M + 2F} = \frac{40 + 20 + (2 \times 34)}{100 + (2 \times 110)} = 0,40$$

Por ser o desvio padrão das frequências desses alelos estimada por intermédio de

$$\sigma = \sqrt{\frac{pq}{2F + M}}, \text{ teríamos em nosso exemplo que } \sigma = \sqrt{\frac{0,60 \times 0,40}{220 + 100}} = 0,027$$

Um outro modo de estimar as frequências dos alelos  $A$  e  $a$  ligados ao cromossomo X é aquela baseada nas frequências dos diferentes genótipos. Para isso, levamos em conta que a frequência de cada genótipo masculino multiplicada pelo total de homens é o número de homens com cada um dos genótipos. Por outro lado, levamos em conta, também, que a frequência de mulheres heterozigotas multiplicada pelo total de mulheres é o número de mulheres heterozigotas e que o dobro do total de mulheres multiplicado pela frequência de mulheres homozigotas é o dobro do número de mulheres com esse genótipo. Em vista disso, fica claro que as fórmulas de cálculo da frequência  $p$  mencionada acima poderia ter sido escritas como abaixo:

$$p = \frac{M \cdot X^A Y + F \cdot X^A X^a + 2F \cdot X^A X^A}{M + 2F}$$

Lembrando, porém que ao trabalharmos com porcentagens, é claro que tornamos o tamanho da amostra masculina artificialmente igual ao da feminina, pois, em ambos os casos a soma de todas as frequências parciais é 1 ou 100%. Pode-se, portanto, escrever que, ao lidarmos com frequências tem-se  $M = F$  e, portanto, a última fórmula pode ser escrita como:

$$p = \frac{F \cdot X^A Y + F \cdot X^A X^a + 2F \cdot X^A X^A}{3F}$$

a qual, depois de simplificada, passa a ser escrita como:

$$p = \frac{X^A Y + X^A X^a + 2X^A X^A}{3}$$

Aplicando raciocínio análogo, a frequência  $q$  do alelo  $a$  pode ser calculada pela fórmula abaixo:

$$q = \frac{X^a Y + X^A X^a + 2X^a X^a}{3}$$

ou, mais facilmente, por intermédio de  $q = 1 - p$ .

Usando os dados de nosso exemplo teríamos :

$$p = \frac{0,60 + 0,182 + (2 \times 0,509)}{3} = 0,60 ; q = 1 - p = 1 - 0,60 = 0,40$$

Como se vê, a freqüência genotípica dos homens corresponde exatamente à freqüência gênica, mas a distribuição genotípica entre as mulheres difere significativamente daquela esperada em equilíbrio genético estável, ou seja,  $X^A X^A = Fp^2$ ;  $X^A X^a = F2pq$ ;  $X^a X^a = Fq^2$ . De fato, comparando as proporções genotípicas esperadas com as observadas, apenas entre as mulheres, tem-se:

Valores	$X^A X^A$	$X^A X^a$	$X^a X^a$	Total
Observado (o)	56	20	34	110
Esperado (e)	39,6	52,8	17,6	110
$\frac{(o - e)^2}{e}$	6,792	20,376	15,282	$\chi^2_{(1)} = 42,45$ $P << 0,001$

Apesar de a distribuição genotípica observada nas mulheres diferir significativamente daquela esperada segundo  $(p+q)^2$ , tem-se que tal equilíbrio genético estável deverá ser atingido após uma única geração de panmixia porque, em nosso exemplo, **as freqüências genotípicas dos homens são iguais às freqüências gênicas**. De fato, é isso o que se pode observar na Tabela 1.3, pois, na primeira geração filial os homens continuam mostrando freqüências genotípicas idênticas às freqüências gênicas, enquanto os genótipos das mulheres passam a se distribuir segundo  $(p+q)^2 = 1$ .

Tabela 1.3. Demonstração de que a distribuição genotípica estável é atingida após panmixia de uma geração inicial composta por homens  $X^A Y$  (60%) e  $X^a Y$  (40%), e mulheres  $X^A X^A$  (50,9%  $\cong$  51%),  $X^A X^a$  (18,2%  $\cong$  18%) e  $X^a X^a$  (30,9%  $\cong$  31%).

Casais (geração inicial)		Primeira Geração Filial				
Tipo	Freqüência	$X^A X^A$	$X^A X^a$	$X^a X^a$	$X^A Y$	$X^a Y$
$X^A X^A \times X^A Y$	$0,51 \times 0,60 = 0,306$	0,153	-	-	0,153	-
$X^A X^A \times X^a Y$	$0,51 \times 0,40 = 0,204$	-	0,102	-	0,102	-
$X^A X^a \times X^A Y$	$0,18 \times 0,60 = 0,108$	0,027	0,027	-	0,027	0,027
$X^A X^a \times X^a Y$	$0,18 \times 0,40 = 0,072$	-	0,018	0,018	0,018	0,018
$X^a X^a \times X^A Y$	$0,31 \times 0,60 = 0,186$	-	0,093	-	-	0,093
$X^a X^a \times X^a Y$	$0,31 \times 0,40 = 0,124$	-	-	0,062	-	0,062
Total	=1,000	0,180	0,240	0,080	0,300	0,200
Freqüência em relação a cada sexo		0,360 ( $p^2$ )	0,480 ( $2pq$ )	0,160 ( $q^2$ )	0,600 ( $p$ )	0,400 ( $q$ )

Se a geração inicial da população teórica tomada para exemplo tivesse mostrado as freqüências genotípicas abaixo:

Mulheres:  $X^A X^A = 0,43$ ;  $X^A X^a = 0,24$ ;  $X^a X^a = 0,33$

Homens:  $X^A Y = 0,70$ ;  $X^a Y = 0,30$

o equilíbrio genético não seria atingido após uma única geração de panmixia, como se pode verificar na Tabela 2.3. Tal equilíbrio demoraria a ser atingido porque, apesar de as frequências dos alelos  $A$  e  $a$  serem as mesmas da população anterior, isto é,  $p = \frac{0,70 + 0,24 + 0,86}{3} = 0,60$  e  $q = 1 - 0,60 = 0,40$ , elas diferem das frequências genotípicas observadas nos homens da geração inicial.

Tabela 2.3. Demonstração de que a distribuição genotípica estável não é alcançada após panmixia de uma geração inicial composta por homens  $X^A Y$  (70%) e  $X^a Y$  (30%) e mulheres  $X^A X^A$  (43%),  $X^A X^a$  (24%) e  $X^a X^a$  (33%).

Casais (geração inicial)		Primeira Geração Filial				
Tipo	Frequência	$X^A X^A$	$X^A X^a$	$X^a X^a$	$X^A Y$	$X^a Y$
$X^A X^A \times X^A Y$	0,301	0,1505	-	-	0,1505	-
$X^A X^A \times X^a Y$	0,129	-	0,0645	-	0,0645	-
$X^A X^a \times X^A Y$	0,168	0,0420	0,0420	-	0,0420	0,0420
$X^A X^a \times X^a Y$	0,072	-	0,0180	0,0180	0,0180	0,0180
$X^a X^a \times X^A Y$	0,231	-	0,1155	-	-	0,1155
$X^a X^a \times X^a Y$	0,099	-	-	0,0495	-	0,0495
Total	1,000	0,1925	0,2400	0,0675	0,2750	0,2250
<b>Frequência em relação a cada sexo</b>		<b>0,3850</b>	<b>0,4800</b>	<b>0,1350</b>	<b>0,5500</b>	<b>0,4500</b>
<b>Frequência esperada em equilíbrio genético</b>		<b>0,3600</b> ( $p^2$ )	<b>0,4800</b> ( $2pq$ )	<b>0,1600</b> ( $q^2$ )	<b>0,6000</b> ( $p$ )	<b>0,4000</b> ( $q$ )

A Tabela 3.3, por sua vez, generaliza o que foi discutido acima ao mostrar a distribuição familiar em uma população teórica que está em equilíbrio genético em relação aos genótipos decorrentes de um par de alelos  $A, a$  do cromossomo X, com frequências  $p$  e  $q = 1 - p$ .

Tabela 3.3. Distribuição das famílias de uma população teórica que está em equilíbrio genético estável em relação aos genótipos determinados por um par de alelos  $A, a$  do cromossomo X, que ocorrem com frequências  $p$  e  $q = 1 - p$ , respectivamente.

Casais		Filhas			Filhos	
Tipo	Frequência	$X^A X^A$	$X^A X^a$	$X^a X^a$	$X^A Y$	$X^a Y$
$X^A X^A \times X^A Y$	$p^2 \cdot p = p^3$	$p^3$	-	-	$p^3$	-
$X^A X^A \times X^a Y$	$p^2 \cdot q = p^2 q$	-	$p^2 q$	-	$p^2 q$	-
$X^A X^a \times X^A Y$	$2pq \cdot p = 2p^2 q$	$p^2 q$	$p^2 q$	-	$p^2 q$	$p^2 q$
$X^A X^a \times X^a Y$	$2pq \cdot q = 2pq^2$	-	$pq^2$	$pq^2$	$Pq^2$	$pq^2$
$X^a X^a \times X^A Y$	$q^2 \cdot p = pq^2$	-	$pq^2$	-	-	$pq^2$
$X^a X^a \times X^a Y$	$q^2 \cdot q = q^3$	-	-	$q^3$	-	$q^3$
<b>Total</b>	<b><math>p+q = 1</math></b>	<b><math>p^2</math></b>	<b><math>2pq</math></b>	<b><math>q^2</math></b>	<b><math>p</math></b>	<b><math>q</math></b>

Nos casos de polialelismo ligados ao sexo que não mostram relações de dominância a estimativa das frequências gênicas também pode ser feita levando em conta a contagem do número

de cromossomos X com esses alelos nos gametas que serviram para constituir essa geração inicial ou a frequência dos genótipos. Assim, num caso de trialelismo em que são analisados os alelos  $A$ ,  $a$  e  $a_1$  de um loco do cromossomo X, poderemos estimar a frequência  $p$ ,  $q$  e  $r$  desses alelos a partir de:

$$p = \frac{X^A Y + X^A X^a + X^A X^{a_1} + 2X^A X^A}{M + 2F}$$

$$q = \frac{X^a Y + X^A X^a + X^a X^{a_1} + 2X^a X^a}{M + 2F}$$

$$r = \frac{X^{a_1} Y + X^A X^{a_1} + X^a X^{a_1} + 2X^{a_1} X^{a_1}}{M + 2F} \quad \text{ou } r = 1 - (p+q)$$

Se levarmos em conta as frequências relativas dos diferentes genótipos, essas frequências gênicas poderão ser estimadas a partir das fórmulas abaixo, onde os símbolos em negrito indicam as frequências relativas:

$$p = \frac{\mathbf{X^A Y} + \mathbf{X^A X^a} + \mathbf{X^A X^{a_1}} + 2\mathbf{X^A X^A}}{3}$$

$$q = \frac{\mathbf{X^a Y} + \mathbf{X^A X^a} + \mathbf{X^a X^{a_1}} + 2\mathbf{X^a X^a}}{3}$$

$$r = \frac{\mathbf{X^{a_1} Y} + \mathbf{X^A X^{a_1}} + \mathbf{X^a X^{a_1}} + 2\mathbf{X^{a_1} X^{a_1}}}{3} \quad \text{ou } r = 1 - (p+q)$$

### TESTE DA HIPÓTESE DE EQUILÍBRIO GENÉTICO

Quando, em um estudo de uma amostra, levamos em conta apenas um par de alelos codominantes do cromossomo X e queremos saber se a distribuição dos genótipos determinados por esses alelos é estável, isto é, representa uma população em equilíbrio genético poderemos comparar as proporções observadas com as esperadas nos homens ( $Mp$  e  $Mq$ ) e nas mulheres ( $Fp^2$ ,  $F2pq$  e  $Fq^2$ ) por intermédio de um teste de qui-quadrado. Esse qui-quadrado terá dois graus de liberdade, porque para calcular as cinco classes esperadas (duas masculinas e três femininas) valemo-nos de três informações, a saber, o número de homens, o número de mulheres e a frequência de um dos alelos.

Para exemplificar, consideremos que a desidrogenase de 6-fosfato de glicose (G-6PD) foi estudada eletroforéticamente numa amostra de 210 homens e 200 mulheres de origem europeia mediterrânea e que, dentre os homens, 198 (94,3%) apresentaram a variante normal ( $B^+$ ) e 12 (5,7%) mostraram a variante deficiente ( $B^-$ ) dessa enzima. Nas mulheres foi possível distinguir 182 (91%) com variante normal ( $B^+$ ), uma (0,5%) deficiente ( $B^-$ ) e 17 (8,5%) com atividade intermediária, por serem heterozigotas. Para facilitar a notação, indiquemos os genótipos masculinos por  $B^+$  e  $B^-$  e os femininos por  $B^+B^+$ ,  $B^+B^-$  e  $B^-B^-$ .

Com base nesses dados podemos calcular as frequências gênicas e as frequências genóticas esperadas para compará-las com as observadas, obtendo:

$$p = B^+ = \frac{0,943 + 0,085 + 1,820}{3} = 0,949$$

$$q = B^- = 1 - p = 1 - 0,949 = 0,051$$

$$\sigma = 0,009$$

Valores	Mulheres				Homens		
	$B^+B^+$	$B^+B^-$	$B^-B^-$	Total	$B^+$	$B^-$	Total
Observados (o)	182	17	1	200	198	12	210
Esperados (e)	180,1 ( $Fp^2$ )	19,4 ( $F2pq$ )	0,5 ( $Fq^2$ )	200	199,3 ( $Mp$ )	10,7 ( $Mq$ )	210
$\frac{(o - e)^2}{e}$	0,020	0,297	0,5		0,008	0,158	$\chi^2_{(2)} = 0,983$ $0,50 < P < 0,70$

O valor do qui-quadrado com dois graus de liberdade (0,983) permite-nos, pois, concluir que a amostra representa uma população em equilíbrio genético quanto aos genótipos estudados.

Consideremos, agora, um exemplo de trialelismo ligado ao sexo, analisando os resultados de estudos eletroforéticos de G-6PD em africanos (141 homens e 100 mulheres), nos quais se observou a seguinte distribuição:

Homens:  $A^+ = 31$  (22%),  $B^+ = 79$  (56%),  $A^- = 31$  (22%)

Mulheres:  $A^+ = 4$ ,  $B^+ = 31$ ,  $A^- = 4$ ,  $A^+B^+ = 28$ ,  $A^+A^- = 10$ ,  $A^-B^+ = 23$ .

Com essas dados podemos calcular as frequências  $p$ ,  $q$  e  $r$  dos alelos  $A^+$ ,  $B^+$  e  $A^-$  e as frequências genóticas esperadas para compará-las com as observadas, como abaixo:

$$p = A^+ = \frac{0,22 + 0,28 + 0,10 + 0,08}{3} = 0,227$$

$$q = B^+ = \frac{0,56 + 0,28 + 0,23 + 0,62}{3} = 0,563$$

$$r = A^- = 1 - (0,227 + 0,563) = 0,210$$

Valores	Mulheres							Homens			
	$A^+A^+$	$B^+B^+$	$A^-A^-$	$A^+B^+$	$A^+A^-$	$A^-B^+$	Total	$A^+$	$B^+$	$A^-$	Total
Obs. (o)	4	31	4	28	10	23	100	31	79	31	141
Esp. (e)	5,2	31,7	4,4	25,6	9,5	23,6	100	32	79,4	29,6	141
$\frac{(o - e)^2}{e}$	0,277	0,015	0,036	0,225	0,026	0,015		0,031	0,002	0,066	$\chi^2_{(5)} = 0,693$ $0,98 < p < 0,99$

O valor do qui-quadrado obtido (0,693) com cinco graus de liberdade, pois são necessárias quatro informações (a frequência de dois alelos e os totais de homens e de mulheres) para calcular as nove classes esperadas, permite concluir que as amostras de homens e mulheres estudadas representam uma população em equilíbrio genético quanto aos genótipos estudados.

## ESTIMATIVA DAS FREQUÊNCIAS DE GENES DO CROMOSSOMO X QUANDO EXISTE RELAÇÃO DE DOMINÂNCIA ENTRE OS FENÓTIPOS FREQUENTES NA POPULAÇÃO

Consideremos os alelos  $A, a$  de um loco do cromossomo X e que os fenótipos ligados ao sexo por eles determinados apresentam relação de dominância, isto é o fenótipo  $X^A X^A$  não pode ser distinguido do  $X^A X^a$ . Se o fenótipo recessivo nas mulheres ( $X^a X^a$ ) não for raro na população, a frequência do gene responsável por esse fenótipo pode ser estimada de modo relativamente simples a partir de uma amostra aleatória da população. Nessa amostra consideramos que o número de homens (M) representa a contribuição de M cromossomos X da geração anterior e, se o número de mulheres for F, ter-se-á que 2F representa a contribuição de 2F cromossomos X dessa geração.

Se  $q$  for a frequência do cromossomo  $X^a$  e  $p = 1 - q$  for a frequência do cromossomo  $X^A$  espera-se, numa população em equilíbrio genético estável, que:

1. O número de mulheres com o fenótipo recessivo ( $X^a X^a$ ) seja igual ao produto do número de mulheres pela frequência do genótipo  $X^a X^a$ , isto é  $Fq^2$ . Assim, representando o número de mulheres com o genótipo  $X^a X^a$  por  $F_{aa}$ , pode-se escrever  $F_{aa} = Fq^2$ , de onde se tira  $F = \frac{F_{aa}}{q^2}$ .

2. O número de homens com o genótipo  $X^a Y$  seja igual ao produto do número de homens pela frequência  $q$ , isto é  $Mq$ . Representando o número de homens com genótipo  $X^a Y$  por  $M_a$  e o daqueles com genótipo  $X^A Y$  por  $M_A$  pode-se escrever  $M_a = Mq$  de onde se tira  $M = \frac{M_a}{q}$ , bem como  $M_A = M - M_a$ .

Numa amostra aleatória de M homens e F mulheres o número total de cromossomos  $X^a$  será

$$(2F+M)q = 2Fq + Mq$$

e, levando em conta o exposto acima, essa expressão pode ser apresentada como uma equação do segundo grau (Pinto Jr., 1966), como se verifica abaixo:

$$\begin{aligned} (2F+M)q &= 2 \frac{F_{aa}}{q^2} \cdot q + M - M_A \\ (2F+M)q &= \frac{F_{aa}}{q} + \frac{M_a}{q} - M_A \\ (2F+M)q + M_A &= \frac{2F_{aa} + M_a}{q} \\ (2F+M)q^2 + M_A q - (2F_{aa} + M_a) &= 0 \end{aligned}$$

Essa última equação permite o cálculo da estimativa da frequência de  $X^a$ , porque uma das raízes de  $q$  em uma equação do segundo grau é  $q = \frac{-b + \sqrt{b^2 - 4ac}}{2a}$ . Pode-se, pois, escrever que a frequência  $q$  de  $X^a$  pode ser estimada por intermédio de:

$$q = \frac{-M_A + \sqrt{M_A^2 - 4(2F + M)(2F_{aa} + M_a)}}{2(2F + M)}$$

enquanto que a frequência  $p$  de  $X^A$  pode ser calculada por intermédio de  $p = 1 - q$ . O desvio padrão será calculado, de acordo com Neel e Schull (1954) pela fórmula abaixo:

$$\sigma = \sqrt{\frac{1}{\frac{4F}{1-q^2} + \frac{M}{pq}}}$$

Para ilustrar a aplicação dessas fórmulas consideremos os dados de Noades *et al.* (1966) a respeito de 2.082 indivíduos (1.013 homens e 1.069 mulheres) caucasóides da Grã-Bretanha cujas hemácias foram testadas com um anti-soro contendo anticorpos anti- $Xg^a$ , o qual é capaz de classificar os seres humanos em  $Xg(a+)$  e  $Xg(a-)$ , conforme as suas hemácias aglutinem ou não após a ação desse anti-soro e da prova indireta da antiglobulina humana (prova indireta de Coombs). Dentre os 1.013 homens examinados por Noades e colaboradores (1966), 667 foram classificados como  $Xg(a+)$  e 346 como  $Xg(a-)$ , enquanto que das 1.069 mulheres 967 foram classificadas como  $Xg(a+)$  e 102 como  $Xg(a-)$ .

Sabendo-se que a reação  $Xg(a-)$  é determinada pelo alelo  $Xg$  do cromossomo X em hemizigose nos homens ( $Xg$ ) e em homozigose nas mulheres ( $XgXg$ ), enquanto que a reação  $Xg(a+)$  é determinado pelo alelo  $Xg^a$  em hemizigose nos homens ( $Xg^a$ ) e em homozigose ou heterozigose nas mulheres ( $Xg^aXg^a$  ou  $Xg^aXg$ ), podemos estimar as frequências dos alelo  $Xg$  e  $Xg^a$  na população da Grã-Bretanha a partir desses dados. Assim, tem-se:

$M_A = 667$	$F_{aa} = 102$
$M_a = 346$	$F = 1.069$
$M = 1.013$	$2F_{aa} = 204$
	$2F = 2.138$

$$q = \frac{-667 + \sqrt{(667)^2 + 4(2.138 + 1.013)(204 + 346)}}{2(2.138 + 1.013)} = 0,3251$$

$$p = 1 - 0,3251 = 0,6749$$

$$\sigma = \sqrt{\frac{1}{\frac{4 \times 1.069}{1 - 0,106} + \frac{1.013}{0,675 \times 0,325}}} = 0,0103$$

Com base nessas estimativas podemos dizer que o número esperado de mulheres heterozigotas na amostra estudada é igual à frequência  $2pq$  multiplicada pelo total de mulheres, isto é  $F2pq = 1.069 \times 0,4388 = 469,1$ . Além disso, podemos dizer, também que o número esperado de homens e mulheres com os grupos sanguíneos  $Xg(a+)$  e  $Sg(a-)$  deve ser calculado como abaixo:

$$\begin{aligned} \text{Homens } Xg(a+) &= Mp &= 1.013 \times 0,6749 = 683,7 \\ \text{Homens } Xg(a-) &= Mq &= 1.013 \times 0,3251 = 329,3 \\ \text{Mulheres } Xg(a+) &= F(p^2 + 2pq) &= 1.069 \times 0,8943 = 956 \\ \text{Mulheres } Xg(a-) &= Fq^2 &= 1.069 \times 0,1057 = 113 \end{aligned}$$

Esses números esperados permitem pôr à prova a hipótese nula de que a amostra estudada pode ser considerada como extraída de uma população em equilíbrio genético por intermédio de um qui-quadrado com um grau de liberdade. Isso porque existem quatro classes esperadas e são três as informações da amostra necessárias para calculá-las (o total de homens, o total de mulheres e a frequência de um dos alelos). A análise dos dados de nosso exemplo feita na Tabela 4.3 permite concluir pela aceitação dessa hipótese, já que o qui-quadrado obtido (2,453) indica que as diferenças entre as proporções observadas e esperadas não são significativas.

Tabela 4.3. Verificação de que a amostra estudada por Noades *et al.* (1966) pode ser considerada como extraída de uma população em equilíbrio genético.

Sexo	Grupo Sanguíneo	Número Observado	Número Esperado	$\frac{(o-e)^2}{e}$
Masculino	Xg(a+)	667	683,7	0,408
	Xg(a-)	346	329,3	0,847
Feminino	Xg(a+)	967	956	0,127
	Xg(a-)	102	113	1,071
Total		2.082	2.082	$\chi^2_{(1)} = 2,453$ $0,10 < P < 0,20$

### ESTIMATIVA DAS FREQUÊNCIAS DE GENES DO CROMOSSOMO X QUANDO EXISTE RELAÇÃO DE DOMINÂNCIA EM RELAÇÃO A UM FENÓTIPO RARO

Do mesmo modo que em relação aos genes autossômicos é bem mais simples estimar a frequência de heterozigotas de genes do cromossomo X quando o fenótipo recessivo é raro. Quando isso acontece, aceitamos que a população está em equilíbrio genético, de modo que consideramos a frequência dos homens que exibem o fenótipo recessivo raro em estudo como sendo a estimativa da frequência  $q$  do alelo que o determina. Não há a necessidade de levar em conta o número de mulheres homozigotas desse gene porque a sua ocorrência eventual, como é sabido, pode ser desprezada.

Em vista do exposto, tem-se que, em uma população como a de Porto Alegre, RS, na qual a frequência de homens com hemofilia A, que apresentam deficiência do fator VIII de coagulação

(globulina anti-hemofílica), foi estimada em 1: 10.000 (Roisenberg, 1968), pode-se aceitar ser esse valor a estimativa da frequência  $q$  do gene da hemofilia A, ou melhor, das diferentes mutações tomadas em conjunto, que resultam na hemofilia A ( $q = 0,0001$ ). A frequência das mulheres heterozigotas do gene alterado em Porto Alegre pode, pois, ser estimada em 1: 5.000 seja a partir de  $Aa = 2pq$  ou de  $Aa = 2q$ , pois a frequência  $p = 1 - q$  do alelo que condiciona níveis normais do fator VIII é, praticamente, a unidade ( $p = 0,9999$ ). Em outras palavras, em relação a um alelo raro do cromossomo X, a proporção de mulheres portadoras desse alelo (heterozigotas) será, praticamente, o dobro da frequência dos homens que possuem tal alelo em populações em equilíbrio genético.

## EQUILÍBRIO GENÉTICO E HERANÇA POLIGÊNICA

Quando lidamos com caracteres que dependem de mais de um par de alelos (herança poligênica) pode-se demonstrar que, se a população obedecer as oito condições estabelecidas para as populações teóricas em equilíbrio de Hardy e Weinberg, os genótipos decorrentes de cada par de alelos do conjunto poligênico atingem, isoladamente, distribuição estável após uma única geração de panmixia, isto é, os genótipos decorrentes de cada par de alelos obedecerão a lei de Hardy e Weinberg. Entretanto, *os genótipos que incluem todo o conjunto poligênico necessitarão de várias gerações de casamentos aleatórios para alcançar uma distribuição genotípica estável, a menos que a geração inicial seja composta exclusivamente por heterozigotos de todos os alelos do conjunto poligênico.*

Entretanto, o número de gerações necessárias para que a população atinja o equilíbrio genético em relação aos caracteres poligênicos não depende apenas da composição genética inicial, mas também do número de pares de alelos implicados. De fato, se o caráter depender de dois pares de alelos o equilíbrio genético será atingido quando  $(p_1+q_1)^2(p_2+q_2)^2 = 1$ , mas se ele depender de três pares de alelos o equilíbrio genético somente será alcançado quando tivermos  $(p_1+q_1)^2(p_2+q_2)^2(p_3+q_3)^2 = 1$ . Para generalizar, pode-se escrever que, para os caracteres poligênicos, o equilíbrio genético é atingido quando

$$(p_1+q_1)^2(p_2+q_2)^2(p_3+q_3)^2 \dots (p_x+q_x)^2 = 1$$

Para exemplificar, consideremos apenas dois pares de alelos autossômicos  $A,a$  e  $B,b$  com frequências  $A = p_1 = 0,5$ ;  $a = q_1 = 0,5$ ;  $B = p_2 = 0,5$ ;  $b = q_2 = 0,5$ . Se a geração inicial de uma população for composta apenas por indivíduos com genótipos  $AABB$ ,  $AaBb$  e  $aabb$ , cujas frequências são iguais, respectivamente, a 25%, 50% e 25%, ela estará em equilíbrio em relação aos genótipos decorrentes dos alelos  $A, a$  e dos alelos  $B,b$  considerados separadamente, mas não em relação aos dois pares de alelos considerados simultaneamente.

De fato, quando se consideram esses dois pares de alelos ao mesmo tempo é necessário que os genótipos se distribuam segundo  $(p_1+q_1)^2(p_2+q_2)^2 = 1$  para que o equilíbrio genético seja

alcançado. Em outras palavras, é necessário que os nove genótipos possíveis ocorram com as frequências abaixo, as quais foram calculadas levando em conta que, sendo  $p_1 = p_2 = 0,50$  e  $q_1 = q_2 = 0,50$  é possível escrever  $p_1 = p_2 = p$  e  $q_1 = q_2 = q$ .

$$\begin{aligned}
 AABB &= p^2 \cdot p^2 = p^4 = 0,0625 \\
 AABb &= p^2 \cdot 2pq = 2p^3q = 0,1250 \\
 AAbb &= p^2 \cdot q^2 = p^2q^2 = 0,0625 \\
 AaBB &= 2pq \cdot p^2 = 2p^3q = 0,1250 \\
 AaBb &= 2pq \cdot 2pq = 4p^2q^2 = 0,2500 \\
 Aabb &= 2pq \cdot q^2 = 2pq^3 = 0,1250 \\
 aaBB &= q^2 \cdot p^2 = p^2q^2 = 0,0625 \\
 aaBb &= q^2 \cdot 2pq = 2pq^3 = 0,1250 \\
 aabb &= q^2 \cdot q^2 = q^4 = 0,0625
 \end{aligned}$$

Na Tabela 5.3 pode-se constatar que a primeira geração filial, resultante da panmixia da geração inicial em discussão, continua em equilíbrio apenas em relação a cada par de alelos considerados isoladamente, mas não em relação aos dois pares analisados em conjunto. Assim, os genótipos *AA*, *Aa* e *aa* continuam se distribuindo segundo 25%, 50% e 25%, o mesmo ocorrendo com os genótipos *BB*, *Bb* e *bb*. Na distribuição genotípica levando em conta os dois pares de alelos verifica-se na Tabela 5.3 que certos genótipos aparecem com frequência menor que as esperadas em equilíbrio genético, ocorrendo o inverso com outros. Mais algumas gerações em panmixia serão necessárias, portanto, para que a distribuição genotípica estável seja alcançada.

Tabela 5.3. Demonstração de que a distribuição genotípica estável em relação a caracteres poligênicos não é alcançada após uma única geração em panmixia se ela não for iniciada exclusivamente por heterozigotos. Para essa demonstração foram considerados apenas dois pares de alelos autossômicos (*A,a* e *B,b*) e a existência de três genótipos na geração inicial, *AABB* (25%), *AaBb* (50%) e *aabb* (25%).

Geração Inicial		Primeira Geração Filial								
Casais	Freq.	<i>AABB</i>	<i>AABb</i>	<i>AAbb</i>	<i>AaBB</i>	<i>AaBb</i>	<i>Aabb</i>	<i>aaBB</i>	<i>aaBb</i>	<i>aabb</i>
<i>AABB</i> × <i>AABB</i>	0,0625	0,062500	-	-	-	-	-	-	-	-
<i>AABB</i> × <i>AaBb</i>	0,2500	0,062500	0,062500	-	0,062500	0,062500	-	-	-	-
<i>AABB</i> × <i>aabb</i>	0,1250	-	-	-	-	0,125000	-	-	-	-
<i>AaBb</i> × <i>AaBb</i>	0,2500	0,015625	0,031250	0,015625	0,031250	0,062500	0,031250	0,015625	0,031250	0,015625
<i>AaBb</i> × <i>aabb</i>	0,2500	-	-	-	-	0,062500	0,062500	-	0,062500	0,062500
<i>aabb</i> × <i>aabb</i>	0,0625	-	-	-	-	-	-	-	-	0,062500
Total	1,0000	0,140625	0,093750	0,015625	0,093750	0,312500	0,093750	0,015625	0,093750	0,140625
Frequência esperada em equilíbrio genético		0,0625	0,1250	0,0625	0,1250	0,2500	0,1250	0,0625	0,1250	0,0625

### EQUILÍBRIO GENÉTICO E GENES LIGADOS

Se dois pares de alelos *A,a* e *B,b* pertencerem a um mesmo grupo de ligação e ocorrerem com frequências  $p_1$ ,  $q_1$  e  $p_2$ ,  $q_2$ , poderemos supor que na geração inicial de uma população teórica somente existissem indivíduos com os genótipos *AB/AB*, *AB/ab* e *ab/ab*. Se esses dois pares de

alelos se recombinarem em consequência de permuta cromossômica na gametogênese dos indivíduos duplamente heterozigotos  $AB/ab$ , tais indivíduos introduzirão novas combinações gênicas na população, ou seja, as combinações  $Ab$  e  $aB$ , inexistentes anteriormente. Em consequência disso, além dos indivíduos com os genótipos  $AB/AB$ ,  $AB/ab$  e  $ab/ab$ , a população passará a conter aqueles com genótipos  $AB/Ab$ ,  $AB/aB$ ,  $ab/Ab$ ,  $ab/aB$ ,  $Ab/Ab$ ,  $aB/aB$  e  $Ab/aB$ .

É evidente que os indivíduos com genótipo  $Ab/aB$  poderão reintroduzir as combinações gênicas originais na população, isto é, as combinações  $AB$  e  $ab$ , por intermédio de permutas durante a gametogênese. Mas é claro, também, que esse processo deverá ser menos intenso, pelo menos no início, já que a frequência de indivíduos com genótipo  $Ab/aB$  deverá ser pequena. De qualquer modo, a estabilidade da distribuição genotípica, isto é, o equilíbrio genético será atingido quando as frequências das combinações  $AB$ ,  $Ab$ ,  $aB$  e  $ab$  ocorrerem aleatoriamente na população, isto é, quando

$$AB = p_1p_2, Ab = p_1q_2, aB = q_1p_2 \text{ e } ab = q_1q_2$$

Nesse caso, a frequência dos indivíduos duplamente heterozigotos com genótipo  $AB/ab$  será igual à daqueles com genótipo  $Ab/aB$ , pois:

$$2(p_1p_2)(q_1q_2) = 2(p_1q_2)(p_2q_1)$$

e a distribuição estável dos genótipos se fará como se houvesse segregação independente, isto é segundo:

$$(p_1 + q_1)^2(p_2 + q_2)^2$$

### QUESTÕES E RESPOSTAS

**Q 1.** As hemácias de uma amostra de 350 indivíduos de uma população foram testadas com um anti-soro anti- $Xg^a$ , observando-se a seguinte distribuição:

Sexo	No.	$Xg(a+)$	$Xg(a-)$
Masculino	150	90 (60%)	60 (40%)
Feminino	200	178 (89%)	22 (11%)

Estimar a frequência dos alelos  $Xg^a$  e  $Xg$  na população representada por essa amostra e calcular o desvio padrão das frequências gênicas.

$$\mathbf{R 1. } Xg = q = \frac{-90 + \sqrt{8.100 + 4(400 + 150)(44 + 60)}}{2(400 + 150)} = 0,361$$

$$Xg^a = p = 1 - 0,361 = 0,639$$

$$\sigma = \sqrt{\frac{1}{\frac{800}{1-0,1303} + \frac{150}{0,2307}}} = 0,025$$

**Q 2.** No concenrente ao sistema sangüíneo Xg a amostra da questão anterior pode ser considerada como em equilíbrio genético?

**R 2.** Sim, porque o qui-quadrado obtido com um grau de liberdade (1,749) indica que não há diferenças significativas entre as proporções observadas e as esperadas nas quatro classes.

Sexo	Fenótipo	Obs.	Esp	$\frac{(o-e)^2}{e}$
M	Xg(a+)	90	96	0,375
	Xg(a-)	60	54	0,667
F	Xg(a+)	178	174	0,092
	Xg(a-)	22	26	0,615

$$\chi^2 = 1,749; 0,10 < P < 0,20$$

**Q 3.** O gene responsável pela produção de distrofina, uma proteína que se localiza na membrana das fibras musculares estriadas e cardíacas das pessoas normais, está localizado no braço superior do cromossomo X, na região Xp21.2. Várias mutações nesse gene podem impedir a produção dessa proteína e, como consequência, determinar a distrofia muscular do tipo Duchenne, que afeta os indivíduos do sexo masculino e os impede de atingir a idade reprodutiva. Numa população na qual a incidência dessa heredopatia ligada ao sexo tem incidência igual a 4 por 100.000 nascimentos, quer-se saber qual a estimativa da freqüência de mulheres heterozigotas do gene da distrofina alterado.

**R 3.** Tendo em mente que no caso das heredopatias recessivas ligadas ao sexo aceitamos que elas estão em equilíbrio genético, a freqüência dos homens com uma doença ligada ao cromossomo X é aceita como sendo a freqüência do gene alterado. Em nosso caso, portanto,  $q = 0,00004$ . Por outro lado, considerando que, em equilíbrio genético, a freqüência de mulheres heterozigotas é  $2pq$  teríamos em nosso caso que essa freqüência pode ser estimada em, aproximadamente, 8 por 100.000 nascimentos ou 1 por 12.500 nascimentos.

**Q 4.** Um caráter poligênico é determinado por dois pares de alelos  $A,a$  e  $B,b$ . Em uma certa população as freqüências dos alelos  $A$  e  $a$  são respectivamente  $p_1 = 0,60$  e  $q_1 = 0,40$ , enquanto as dos alelos  $B$  e  $b$  são respectivamente  $p_2 = 0,30$  e  $q_2 = 0,70$ . Se essa população estiver em equilíbrio genético em relação ao caráter em questão, qual será a distribuição esperada das freqüências genótípicas?

**R 4.**  $AABB = 0,0324$      $AaBB = 0,0432$      $aaBB = 0,0144$   
 $AABb = 0,1512$      $AaBb = 0,2016$      $aaBb = 0,0672$   
 $AAbb = 0,1764$      $Aabb = 0,2352$      $aabb = 0,0784$

**Q 5.** Quando dois caracteres estão em ligação, a associação entre eles é sempre observada ao nível familiar, populacional ou em ambos os níveis?

**R 5.** Somente a associação familiar é sempre observada.

**Q 6.** Sabemos que os grupos sangüíneos dos sistemas Duffy e Rh são determinados por genes do cromossomo número 1. Empregando os anti-soros anti-Fy<sup>a</sup>, anti-Fy<sup>b</sup> e anti-D, as seguintes

freqüências gênicas foram estimadas em uma certa população  $Fy^a = 0,43$  e  $Fy^b = 0,57$  no sistema Duffy, e  $D = 0,60$  e  $d = 0,40$  no sistema Rh. Se a população em apreço estiver em equilíbrio genético qual será a freqüência esperada de indivíduos duplamente heterozigotos  $Fy^aD/Fy^bd$  e  $Fy^ad/Fy^bD$ ?

**R 6.** A freqüência esperada de cada tipo de heterozigotos deve ser 11,76% pois deve-se ter  $2(p_1p_2)(q_1q_2) = 2(p_1q_2)(p_2q_1)$ , isto é,  $2(0,43 \times 0,60)(0,57 \times 0,40) = 2(0,43 \times 0,40)(0,57 \times 0,60) = 0,1176$ .

**Q 7.** As freqüências dos genes  $M$  e  $N$  determinadores dos grupos sanguíneos M, MN e N foram estimadas em 60% e 40% em uma população. Nessa mesma população as freqüências dos genes  $Lu^a$  e  $Lu^b$ , responsáveis pela produção dos antígenos  $Lu^a$  e  $Lu^b$  do sistema sanguíneo Lutheran foram estimadas, respectivamente, em 6% e 94%. Na hipótese de equilíbrio genético, qual a porcentagem de indivíduos com o fenótipo MN  $Lu(a+b+)$ , isto é, com genótipo  $MN Lu^aLu^b$  que se espera nessa população?

**R 7.** 5,41% pois  $(2 \times 0,60 \times 0,40)(2 \times 0,06 \times 0,94) = 0,0541$ .

## REFERÊNCIAS

Neel, J.V. & Schull, W.J. *Human Heredity*. Univ. Chicago Press, Chicago, 1a. ed., 1954.

Noades, J., Gavin, J., Tippett, P., Sanger, R. & Race, R.R. The X-linked blood group system Xg. Tests on British, Northern American and Northern European unrelated people and families. *J. Med. Genet.* 3: 162-168, 1966.

Roisenberg, I. *Hemofilia e estados hemofilóides no Rio Grande do Sul*. Tese de Doutorado, Univ. Fed. R.G.do Sul, 1968.

## CAPÍTULO 4. A ANÁLISE FAMILIAL DE POLIMORFISMOS

Atualmente aceita-se que um gene com frequência entre 1% e 99% deve ser classificado como *gene polimorfo*, aquele com frequência inferior a 1% deve ser denominado *gene idiomorfo*, enquanto um gene com frequência superior a 99% deve ser classificado como *gene monomorfo* (Morton, 1976/1977). Entretanto, é conveniente deixar bem claro que um loco polimórfico pode incluir entre os alelos a ele pertencentes um ou vários genes idiomorfos. Assim, por exemplo, se os alelos  $A$ ,  $a$  e  $a_1$  de determinado loco tiverem frequências gênicas iguais, respectivamente, a 0,600, 0,395 e 0,005 diremos que os alelos  $A$  e  $a$  são polimorfos, enquanto o alelo  $a_1$  será dito idiomorfo. Por sua vez, os caracteres que resultam de locos que incluem pelo menos dois alelos polimórficos são denominados *polimorfismos genéticos* ou *sistemas genéticos polimórficos*.

No presente capítulo discutiremos os métodos clássicos de análise familiar de caracteres polimórficos para testar uma hipótese monogênica que explique a associação familiar de um par de fenótipos alternativos, um dos quais mostra dominância sobre o outro. O final deste capítulo será dedicado a uma abordagem da metodologia utilizada para o estudo familiar de caracteres complexos.

### DISTRIBUIÇÃO FAMILIAL DE POLIMORFISMOS AUTOSSÔMICOS

Para acompanhar melhor os métodos clássicos de análise da distribuição familiar de um caráter polimórfico nada melhor do que partir de um exemplo concreto. Assim, consideremos os dados apresentados na Figura 1.4 a respeito de 50 famílias, cujos componentes foram classificados segundo a presença ou ausência na saliva de substâncias grupo-específicas hidro-solúveis do sistema ABO.

Suponhamos que não se soubesse que a capacidade de secretar ou não essa substância sob a forma hidro-solúvel na saliva e em outros líquidos do corpo é um caráter autossômico monogênico. Nesse caso, a observação da Figura 1.4 nos permitiria dizer, apenas, que os fenótipos *secretor* e *não-secretor* mostram uma associação familiar, pois a proporção de indivíduos com o fenótipo *não-secretor* entre os filhos de casais *secretor* × *secretora* (14 em 83 ou 16,9%) foi menor do que entre os filhos de casais

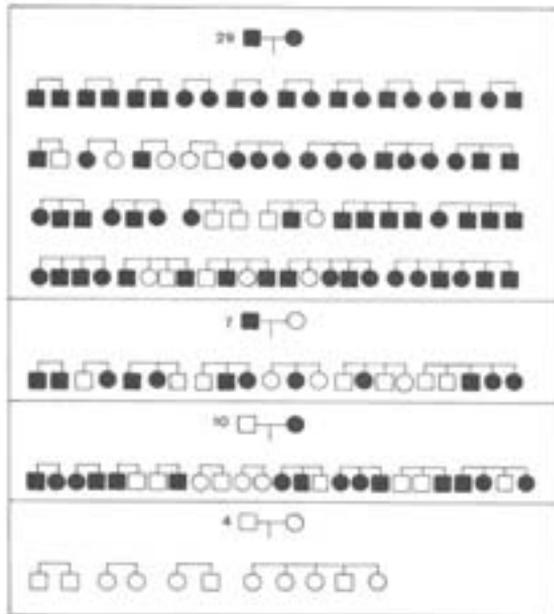


Fig. 1.4. Dados a respeito de 50 famílias cujos elementos foram classificados segundo a capacidade de secretar substâncias grupo-específicas do sistema ABO na saliva. Os indivíduos *secretores* estão representados por símbolos escuros e os *não-secretores* por símbolos claros.

*secretor(a) × não-secretor(a)* (20 em 47 ou 42,6%), enquanto que todos os 11 filhos de casais *não-secretor × não-secretora* (100%) manifestaram o fenótipo *não-secretor*. Em outras palavras, a proporção de indivíduos *secretores* e *não-secretores* na geração filial depende do fenótipo dos genitores como se depreende do alto valor de qui-quadrado com dois graus de liberdade, calculado como no quadro abaixo, onde os valores esperados na hipótese de ausência de associação familiar estão entre parênteses:

Casais	Filhos		
	Secretor	Não-Secretor	Total
Secretor × Secretora	69 (56,5)	14 (26,5)	83
Secret. × Não-Secret.	27 (32,0)	20 (15,0)	47
Não-Secret.× Não-Secret.	- ( 7,5)	11 ( 3,5)	11
Total	96	45	141

$$(\chi^2_{(2)} = 34,680; P < 0,001)$$

Visto que a capacidade de secretar substâncias grupo-específicas do sistema ABO sob a forma hidro-solúvel é uma característica endógena, é permissível supor que a associação familiar observada nos dados da Fig. 1.4 tem determinação genética. Nesse caso, deve-se admitir que o fenótipo *não-secretor* é recessivo, já que dos casais *não-secretor × não-secretora* nasceram apenas filhos *não-secretores*. Por outro lado, pode-se propor uma hipótese monogênica autossômica para explicar a transmissão hereditária desses caracteres. Assim, é permissível supor que o fenótipo *não-secretor* resulta de um gene autossômico *se*

em homozigose (*sese*), enquanto que o fenótipo secretor resulta de seu alelo *Se* em homozigose ou heterozigose (*Se\_*), isto é, de genótipos *SeSe* ou *Sese*.

A hipótese de que os genes devem ser autossômicos resulta da constatação de que as proporções de *secretores* e de *não-secretores* observadas nos indivíduos do sexo masculino não diferiram significativamente daquelas verificadas nos indivíduos do sexo feminino. De fato, nos 50 casais que constituíram a amostra, 36 homens eram *secretores* e 14 *não-secretores*, e 39 mulheres eram *secretoras* e 11 *não-secretoras* ( $\chi^2_{(1)} = 0,480; 0,30 < P < 0,50$ ). Por outro lado, de acordo com a Tabela 1.4, pode-se concluir que, também entre os filhos dos diferentes tipos de casais da amostra, não houve diferença sexual significativa quanto às proporções de *secretores* e de *não-secretores*. Isso permite, pois, analisar os dados a respeito da geração filial sem distinção de sexos (M + F).

Além disso, podemos reunir em um só grupo os dados a respeito das famílias constituídas pelos casais com fenótipos diferentes porque o número de casais *secretor* × *não-secretora* (7) e o de casais *não-secretor* × *secretora* (10) não diferem significativamente ( $\chi^2_{(1)} = 0,529; 0,30 < P < 0,50$ ), do mesmo modo que não diferem significativamente as proporções de *secretores* e de *não-secretores* dos filhos desses dois tipos de casais ( $\chi^2_{(1)} = 0,142; 0,70 < P < 0,80$ ).

Tabela 1.4. Verificação de que entre os filhos dos diferentes tipos de casais da Fig. 1.4 não há diferença sexual significativa quanto às proporções de secretores e de não-secretores.

Casais	Filhos				$\chi^2_{(1)}$
	Sexo	Secret.	Ñ-Secret.	Total	
Secretor × Secretora (29)	M	39	7	46	0,200; 0,50 < P < 0,70
	F	30	7	37	
	M+F	69	14	83	
Secretor × Não-Secret. (7)	M	9	7	16	0,069; 0,70 < P < 0,80
	F	3	3	6	
	M+F	12	10	22	
Não-secret. × Secretora (10)	M	10	7	17	0,031; 0,80 < P < 0,90
	F	5	3	8	
	M+F	15	10	25	
Não-Secret × Não-Secret. (4)	M	-	4	4	
	F	-	7	7	
	M+F	-	11	11	
Total (50)	M	58	25	83	0,299; 0,50 < P < 0,70
	F	38	20	58	
	M+F	96	45	141	

Tendo em vista essas considerações, podemos testar a hipótese monogênica autossômica, isto é, a hipótese de que os caracteres em estudo são determinados um par de alelos autossômicos  $Se, se$ . Para facilitar a notação, denominemos de  $A$  ao alelo  $Se$  e de  $a$  ao alelo  $se$  e consideremos que eles ocorrem na população com frequências iguais, respectivamente, a  $p$  e  $q$ , sendo  $p + q = 1$ . Consideremos, ainda, que o gene  $a$  somente se manifesta em homozigose ( $aa$ ) e que os genótipos  $AA$  e  $Aa$  se expressam de modo idêntico (fenótipo dominante  $A_{-}$ ). Se a população estiver em equilíbrio de Hardy e Weinberg em relação aos genótipos  $AA$ ,  $Aa$  e  $aa$ , é claro que os indivíduos  $A_{-}$  ocorrerão com frequência  $p^2 + 2pq$ , enquanto aqueles com o fenótipo recessivo ( $aa$ ) serão encontrados com frequência  $q^2$ .

Na população será possível a distinção de três tipos de casais quanto aos fenótipos  $A_{-}$  e  $aa$ , ou seja, casais  $A_{-} \times A_{-}$ ,  $A_{-} \times aa$  e  $aa \times aa$ , com frequências que dependerão das frequências dos dois fenótipos alternativos em discussão. Assim, a probabilidade de sortear um casal da população que seja do tipo  $A_{-} \times A_{-}$  é:

$$P(A_{-} \times A_{-}) = (1 - q^2)^2$$

porque  $P(A_{-}) = p^2 + 2pq = 1 - q^2$ .

A probabilidade de sortear um casal da população que seja do tipo  $A_{-} \times aa$ , o qual tanto pode ser constituído por homem  $A_{-}$  e mulher  $aa$  ou homem  $aa$  e mulher  $A_{-}$  é:

$$P(A_{-} \times aa) = P(A_{-} \times aa) + P(aa \times A_{-}) = 2q^2 (1 - q^2)$$

pois sabemos que  $P(aa) = q^2$  e  $P(A_{-}) = 1 - q^2$ .

Obviamente, a probabilidade de sortear um casal da população que seja do tipo  $aa \times aa$  é:

$$P(aa \times aa) = q^2 \cdot q^2 = q^4$$

Em vista do exposto pode-se dizer que, ao lidarmos com um caráter polimórfico determinado por um par de alelos autossômicos com relação de dominância, se tomarmos uma amostra aleatória de  $N$  casais da população, os números observados dos três tipos de casais não deverão diferir significativamente dos valores calculados a partir das fórmulas:

$$\begin{aligned} A_{-} \times A_{-} &= N(1 - q^2)^2 \\ A_{-} \times aa &= 2Nq^2(1 - q^2) \\ aa \times aa &= Nq^4 \end{aligned}$$

Portanto, quando se trabalha com uma amostra de famílias, o primeiro passo para pôr à prova uma hipótese monogênica autossômica, a fim de explicar a associação familiar de um par de fenótipos alternativos, um dos quais mostrando dominância sobre o outro, consiste em comparar a distribuição observada de casais com aquela esperada segundo os cálculos apresentados acima. Evidentemente, esse passo somente deve ser dado depois de demonstrar, como foi feito acima, que as proporções de cônjuges do sexo masculino que manifestam os fenótipos dominante e recessivo não diferem significativamente daquelas observadas entre os cônjuges femininos, bem como, que o número de casais constituídos por homem  $A_-$  e mulher  $aa$  não difere significativamente do número de casais compostos por homem  $aa$  e mulher  $A_-$ .

Aqui é importante tecer algumas considerações sobre a maneira de estimar a frequência  $q$  do alelo  $a$  e, por diferença, a frequência  $p$  do alelo  $A$ , visto que  $p = 1 - q$ . Muitos autores têm estimado a frequência  $q$  a partir da própria amostra de famílias por eles coletadas. Assim, com base em uma amostra de  $N$  famílias, contendo  $n_1$  casais  $A_- \times A_-$ ,  $n_2$  casais  $A_- \times aa$ ,  $n_3$  casais  $aa \times aa$  e um total de  $f$  filhos, parte dos quais tem o fenótipo recessivo ( $f_{aa}$ ) e os restantes o fenótipo dominante ( $f_{A_-}$ ), esses autores têm estimado a frequência  $q$  do alelo  $a$  a partir de  $q = \sqrt{\frac{2n_3 + n_2 + f_{aa}}{2N + f}}$ . Outros pesquisadores têm levado em conta apenas as  $2N$  pessoas que constituem os  $N$  casais, de sorte que estimam a frequência  $q$  do alelo  $a$  a partir de  $q = \sqrt{\frac{2n_3 + n_2}{2N}}$ . Obviamente, em ambos os casos a estimativa da frequência  $p$  do alelo  $A$  é obtida por intermédio de  $p = 1 - q$ .

Entretanto, essas não são as maneiras mais corretas de se estimar as frequências  $p$  e  $q$  para se fazer a análise familiar. Isso porque os indivíduos da geração filial não são independentes da geração paterna, nem são, freqüentemente, independentes entre si, visto que, muitos deles, são irmãos. Portanto, a maneira mais correta para se estimar as frequências dos alelos  $A$  e  $a$  é a de verificar as frequências dos fenótipos alternativos em uma amostra aleatória da população da qual procedem as famílias. Desse modo, a frequência dos indivíduos com o fenótipo recessivo será considerada como  $q^2$ , a frequência do alelo  $a$  será a raiz quadrada desse valor, e a frequência do alelo  $A$  será  $p = 1 - q$ .

Apesar de ser mais correto estimar as frequências dos alelos a partir de uma amostra aleatória da população da qual procedem as famílias, é importante ressaltar que, se a amostra das famílias for grande e a frequência dos genes for alta não se correrá o risco de estimar erradamente as frequências gênicas, a partir dos próprios casais coletados.

Retomando, agora, os dados de nosso exemplo, consideremos que em uma amostra aleatória de 100 indivíduos procedentes da população na qual foram coletadas as 50 famílias da Fig. 1.4, constatou-se que 73 eram *secretores* e 27 *não-secretores*. Isso permite escrever que  $q^2 = 0,27$  e, a partir daí estimar  $q = \sqrt{0,27} = 0,52$  e  $p = 1 - q = 0,48$ . Temos, pois, os elementos para calcular a distribuição esperada de casais segundo os fenótipos *secretor* e *não-secretor*, de acordo com a hipótese monogênica, como abaixo:

$$\begin{aligned} Se_ \times Se_ &= N(1-q^2)^2 = 26,64 \\ Se_ \times sese &= 2Nq^2(1-q^2) = 19,71 \\ sese \times sese &= Nq^4 = 3,65 \end{aligned}$$

Essa distribuição é comparada, em seguida, com a observada por intermédio de um teste de qui-quadrado. Visto que para calcular as três classes esperadas necessitamos apenas de uma informação da amostra, isto é, do total de casais, o qui-quadrado, nesse caso, terá dois graus de liberdade ( $3 - 1 = 2$ ). Entretanto, se tivéssemos estimado a frequência  $q$  do alelo *se* a partir dos casais ou dos casais e seus filhos, o qui-quadrado teria, evidentemente, apenas um grau de liberdade, porque, além do tamanho amostral estaríamos nos valendo de uma segunda informação da amostra.

Comparando as proporções observadas com as esperadas como no quadro abaixo, concluímos, pelo valor do qui-quadrado com dois graus de liberdade (0,616) as diferenças entre elas não são significativas, o que indica que a distribuição dos casais está de acordo com a esperada segundo a hipótese monogênica.

Casais	Obs.	Esp.	$\frac{(o - e)^2}{e}$
<i>Se</i> × <i>Se</i>	29	26,64	0,209
<i>Se</i> × <i>sese</i>	17	19,71	0,373
<i>sese</i> × <i>sese</i>	4	3,65	0,034

$$\chi^2_{(2)} = 0,616; 0,70 < P < 0,80.$$

O segundo passo a ser dado para pôr à prova uma hipótese monogênica autossômica, a fim de explicar a associação familiar de um par de fenótipos alternativos,

um dos quais mostrando dominância sobre o outro, consiste em comparar a distribuição fenotípica observada entre os elementos que constituem as irmandades geradas pelos três tipos de casais ( $A_-\times A_-$ ,  $A_-\times aa$  e  $aa\times aa$ ), com aquela esperada de acordo com a hipótese monogênica. Para tal investigação o geneticista pode valer-se de métodos simples, como os sugeridos por Snyder (1932) e por Fisher (1939).

### O MÉTODO DE SNYDER

No método de Snyder (1932) considera-se que se  $f_1$  for o número total de filhos dos casais  $A_-\times A_-$  da amostra, e se  $q$  for a frequência do alelo  $a$ , determinante do fenótipo recessivo  $aa$ , o número esperado de indivíduos com o fenótipo recessivo entre os filhos de casais  $A_-\times A_-$  deve ser,  $\left(\frac{q}{1+q}\right)^2$  de acordo com a hipótese monogênica. De fato, sabemos que a probabilidade de um casal  $A_-\times A_-$  gerar um filho  $aa$  é igual a  $\frac{1}{4}$  da probabilidade de os dois cônjuges serem  $Aa$ , dado que são  $A_-$ , isto é,

$$P(Aa|A_-).P(Aa|A_-). \frac{1}{4} = \left(\frac{2q}{1+q}\right) \left(\frac{2q}{1+q}\right) \frac{1}{4} = \left(\frac{q}{1+q}\right)^2$$

Multiplicando essa probabilidade por  $f_1$ , isto é,  $\left(\frac{q}{1+q}\right)^2 f_1$ , obtém-se o número esperado de indivíduos  $aa$  entre os filhos de casais  $A_-\times A_-$ .

Se  $f_2$  for o número total de filhos de casais  $A_-\times aa$  da amostra, espera-se, de acordo com a hipótese monogênica, que  $\frac{q}{1+q}f_2$  manifestem o fenótipo recessivo, pois sabemos que a probabilidade de um casal  $A_-\times aa$  gerar um filho  $aa$  é igual à metade da probabilidade de o cônjuge  $A_-$  ser  $Aa$ , isto é,

$$P(Aa|A_-). \frac{1}{2} = \frac{q}{1+q}$$

Assim, multiplicando essa probabilidade por  $f_2$ , isto é,  $\frac{q}{1+q}f_2$ , obtém-se o número esperado de indivíduos  $aa$  entre os filhos de casais  $A_-\times aa$ .

Evidentemente, se  $f_3$  for o número total de filhos de casais  $aa\times aa$ , espera-se que o fenótipo recessivo ocorra em todos os  $f_3$  indivíduos, já que os cônjuges devem ser sempre

homozigotos *aa*. A Tabela 2.4 apresenta de modo sucinto as frequências esperadas dos diferentes tipos de casais e de seus filhos quando se analisa a distribuição familiar de dois fenótipos alternativos de um carácter polimórfico determinado por um par de alelos com relação de dominância e recessividade.

Tabela 2.4. Distribuição familiar esperada numa amostra aleatória de *N* famílias, segundo os fenótipos *A\_* e *aa* determinados por um par de alelos autossômicos *A,a* com frequências *p* e *q*, sendo *p + q = 1*. Presume-se que a população da qual procede a amostra está em equilíbrio de Hardy e Weinberg.

Casais		Filhos		
Tipo	No.	<i>aa</i>	<i>A_</i>	Total
<i>A_ × A_</i>	$N(1 - q^2)^2$	$\left(\frac{q}{1+q}\right)^2 f_1$	$f_1 - \left(\frac{q}{1+q}\right)^2 f_1$	$f_1$
<i>A_ × aa</i>	$2Nq^2(1 - q^2)$	$\frac{q}{1+q} f_2$	$f_2 - \frac{q}{1+q} f_2$	$f_2$
<i>aa × aa</i>	$Nq^4$	$f_3$	-	$f_3$

Voltemos, agora, ao nosso exemplo numérico da Fig. 1.4, para comparar, pelo método de Snyder (1932) as proporções observadas com as esperadas de filhos *secretores* e *não-secretores* nos três tipos de famílias. Essa comparação pode ser feita dispondo os dados como na Tabela 3.4, na qual os valores observados foram obtidos na Tabela 1.4 e os esperados foram calculados a partir das fórmulas apresentadas na Tabela 2.4.

Tabela 3.4. Aplicação do método de Snyder aos dados familiares da Fig. 1.4.

Casais	Filhos Não-Secretores		Filhos Secretores		Total	$\chi^2_{(1)}$
	OBS.	ESP.	OBS.	ESP.		
Secretor × Secretora	14	9,7	69	73,3	83	2,158; 0,10 < P < 0,20
Secret × Não-Secret.	20	16,1	27	30,9	47	1,437; 0,20 < P < 0,30
Não-Secret × Não-Secret.	11	11	-	-	11	

Os valores expressos na Tabela 3.4 tornam evidente que as proporções fenotípicas esperadas e observadas entre os filhos dos três tipos de casais não diferiram significativamente. Tais resultados permitem, pois, aceitar a hipótese monogênica apresentada para explicar a distribuição dos fenótipos *secretor* e *não-secretor* nas famílias e

na população. Nessa análise não se aplicou um teste estatístico em relação aos filhos dos casais *não-secretor* × *não-secretora* porque se tem, no caso, uma única classe esperada (filhos não-secretores), o que impede o cálculo do qui-quadrado. Se, por acaso, entre os filhos dos casais *não-secretor* × *não-secretora* tivesse sido detectada um pequeno número de *secretores* sua ocorrência poderia ser explicada pela admissão de hipóteses tais como ilegitimidade, falta de penetrância ou erro na determinação do fenótipo. Entretanto, se a proporção de indivíduos *secretores* entre os filhos de casais *não-secretores* tivesse sido alta, seria necessário buscar outras explicações para poder aceitar a hipótese monogênica.

Vejamos, agora, no próximo tópico, como aplicar o método de Fisher (1939) aos dados a respeito das irmandades.

### O MÉTODO DE FISHER

O método de Fisher (1939) para pôr à prova a hipótese monogênica pela investigação das irmandades geradas pelos três tipos de casais considera que, se a hipótese monogênica puder ser aceita para explicar a associação familiar dos fenótipos  $A_{-}$ , dominante, e  $aa$ , recessivo, os casais  $A_{-} \times A_{-}$  incluirão três grupos de casais quanto ao genótipo ( $AA \times AA$ ,  $AA \times Aa$  e  $Aa \times Aa$ ) e os casais  $A_{-} \times aa$  incluirão dois grupos ( $AA \times aa$  e  $Aa \times aa$ ).

Em vista disso, a probabilidade de um filho de um casal  $A_{-} \times A_{-}$  manifestar o fenótipo  $A_{-}$  dependerá da probabilidade de tal casal ser  $Aa \times Aa$  dado que ele é  $Aa \times Aa$  bem como de isso não ser verdadeiro, isto é, da probabilidade de esse casal ser  $AA \times AA$  ou  $AA \times Aa$ , dado que ele é  $A_{-} \times A_{-}$ . Ora, sabendo que a probabilidade de um casal ser  $Aa \times Aa$  dado que ele é  $A_{-} \times A_{-}$  é igual a  $\left(\frac{2q}{1+q}\right)^2$  quando  $q$  é a frequência do alelo  $a$ , pois

$P(Aa|A_{-}) = \frac{2q}{1+q}$ , tem-se que a probabilidade de isso não ocorrer, ou seja, de o casal não ser

$Aa \times Aa$  porque é  $AA \times AA$  ou  $AA \times Aa$  dado que é  $A_{-} \times A_{-}$  será igual a  $1 - \left(\frac{2q}{1+q}\right)^2$ .

Visto que os casais  $AA \times AA$  e os casais  $AA \times Aa$  geram, com certeza, somente filhos  $A_{-}$  e que os casais  $Aa \times Aa$  têm probabilidade igual a  $\frac{3}{4}$  de gerar um filho  $A_{-}$ ,

conclui-se que a probabilidade de um filho de um casal  $A_+ \times A_+$  manifestar o fenótipo dominante  $A_+$  será, de acordo com a hipótese monogênica, igual a

$$1 - \left(\frac{2q}{1+q}\right)^2 + \frac{3}{4} \left(\frac{2q}{1+q}\right)^2$$

Em uma irmandade de  $i$  irmãos gerados por um casal  $A_+ \times A_+$ , a probabilidade de todos os  $i$  irmãos terem o fenótipo dominante será, portanto, igual a

$$1 - \left(\frac{2q}{1+q}\right)^2 + \left(\frac{3}{4}\right)^i \left(\frac{2q}{1+q}\right)^2$$

Se lidarmos com  $n$  irmandades de  $i$  irmãos, é claro que o número esperado de irmandades constituídas apenas por indivíduos que manifestam o fenótipo  $A_+$  será igual a

$$n \left[ 1 - \left(\frac{2q}{1+q}\right)^2 + \left(\frac{3}{4}\right)^i \left(\frac{2q}{1+q}\right)^2 \right]$$

Em uma amostra incluindo irmandades de vários tamanhos poder-se-á, portanto, calcular o número esperado de irmandades constituídas somente por indivíduos com o fenótipo dominante  $A_+$  por intermédio de

$$\sum n \left[ 1 - \left(\frac{2q}{1+q}\right)^2 + \left(\frac{3}{4}\right)^i \left(\frac{2q}{1+q}\right)^2 \right]$$

Obviamente, o número esperado de casais  $A_+ \times A_+$  com pelo menos um filho  $aa$  será obtido pela diferença entre o número total de irmandades geradas pelos casais  $A_+ \times A_+$  e o número esperado daquelas que têm apenas indivíduos com o fenótipo dominante  $A_+$ .

No concernente aos filhos de casais  $A_+ \times aa$  têm-se, de acordo com a hipótese monogênica, que a probabilidade de um filho de tal tipo de casal manifestar o fenótipo  $A_+$  depende da probabilidade de o cônjuge com fenótipo dominante ser homocigoto, isto é,

$$P(AA|A_+) = \frac{p}{1+q}, \text{ bem como da probabilidade de ele ser heterocigoto, isto é,}$$

$$P(Aa|A_+) = \frac{2q}{1+q}. \text{ Se o cônjuge com fenótipo dominante for homocigoto, é evidente que o}$$

filho do casal  $A_+ \times aa$  terá, certamente, fenótipo dominante. Se, entretanto, o cônjuge com o fenótipo dominante for heterocigoto é claro, também, que a probabilidade de um filho do casal  $A_+ \times aa$  ter fenótipo dominante será igual a  $\frac{1}{2}$ . Portanto, pode-se dizer, que a

probabilidade de um filho de um casal  $A_-\times aa$  manifestar fenótipo dominante é, de acordo com a hipótese monogênica igual a  $\frac{p}{1+q} + \frac{1}{2} \cdot \frac{2q}{1+q}$ .

Em uma irmandade de  $i$  irmãos gerados por um casal  $A_-\times aa$  a probabilidade de todos os  $i$  irmãos terem o fenótipo dominante será, pois, igual a  $\frac{p}{1+q} + \left(\frac{1}{2}\right)^i \frac{2q}{1+q}$ . Lidando com  $n$  irmandades de  $i$  irmãos gerados por casais  $A_-\times aa$  é óbvio que o número esperado daquelas constituídas somente por indivíduos com o fenótipo dominante será calculado a partir de:

$$n \left[ \frac{p}{1+q} + \left(\frac{1}{2}\right)^i \frac{2q}{1+q} \right]$$

Em uma amostra que inclua irmandades de diferentes tamanhos, geradas por casais  $A_-\times aa$  calcular-se-á o número esperado daquelas compostas somente por indivíduos com o fenótipo dominante por intermédio de

$$\Sigma n \left[ \frac{p}{1+q} + \left(\frac{1}{2}\right)^i \frac{2q}{1+q} \right]$$

Tal fórmula também pode ser escrita como

$$\Sigma n \left[ \frac{p}{1+q} + \left(\frac{1}{2}\right)^i \left(1 - \frac{p}{1+q}\right) \right]$$

pois  $\frac{2q}{1+q}$  é igual a  $1 - \frac{p}{1+q}$ .

O número esperado de casais  $A_-\times aa$ , com pelo menos um filho apresentando o fenótipo recessivo  $aa$ , será obtido pela diferença entre o número total de irmandades geradas pelos casais  $A_-\times aa$  e o número esperado daquelas que têm apenas indivíduos com o fenótipo dominante  $A_-$ .

Podemos, agora, analisar a geração filial pelo método de Fisher (1939), dispondo os dados da Figura 1.4 como na Tabela 4.4. Nessa tabela os números esperados de irmandades geradas por casais  $Se_-\times Se_-$  e por casais  $Se_-\times sese$  que deveriam ser constituídas por indivíduos  $Se_-$  foram calculados com o emprego das fórmulas mencionadas neste tópico. Assim, por exemplo, para obter o número esperado de irmandades constituídas por dois

Tabela 4.4. Análise dos dados familiares da Fig.1 pelo método de Fisher (1939). *i* - tamanho da irmandade; *n* = número de irmandades de tamanho *i*; P - probabilidade de uma irmandade de tamanho *i* ser constituída somente por *secretores* (*Se*<sub>-</sub>); ESP. - número esperado de irmandades constituídas apenas por indivíduos *secretores*; OBS. – número observado de irmandades compostas apenas por indivíduos *secretores*.

Casais	<i>i</i>	<i>n</i>	P	Esp.	Obs.
<i>Se</i> <sub>-</sub> × <i>Se</i> <sub>-</sub>	2	14	0,795	11,130	10
	3	8	0,729	5,832	6
	4	5	0,680	3,400	3
	5	1	0,643	0,643	-
	6	1	0,615	0,615	1
	<b>Total</b>	<b>29</b>			<b>21,620</b>
<i>Se</i> <sub>-</sub> × <i>sese</i>	2	8	0,487	3,896	3
	3	6	0,402	2,412	4
	4	2	0,359	0,718	1
	5	1	0,337	0,337	1
	<b>Total</b>	<b>17</b>			<b>7,363</b>
<i>sese</i> × <i>sese</i>	2	3			
	5	1			
	<b>Total</b>	<b>4</b>			

Casais	Irmandades	Obs.	Esp.	$\chi^2_{(1)}$
<i>Se</i> <sub>-</sub> × <i>Se</i> <sub>-</sub>	Apenas <i>Se</i> <sub>-</sub>	20	21,620	0,477; 0,30 < P < 0,50
	Com <i>sese</i>	9	7,380	
<i>Se</i> <sub>-</sub> × <i>sese</i>	Apenas <i>Se</i> <sub>-</sub>	9	7,363	0,642; 0,30 < P < 0,50
	Com <i>sese</i>	8	9,637	

indivíduos *Se*<sub>-</sub> filhos de casais *Se*<sub>-</sub> × *Se*<sub>-</sub>, calculou-se, inicialmente, a probabilidade de encontro de tais irmandades por intermédio de  $0,532 + (0,5625 \times 0,468) = 0,795$ , pois:

$$\left(\frac{2q}{1+q}\right)^2 = \left(\frac{2 \times 0,520}{1,520}\right)^2 = 0,468 \qquad 1 - \left(\frac{2q}{1+q}\right)^2 = 0,532 \qquad \left(\frac{3}{4}\right)^2 = 0,5625$$

Multiplicando a probabilidade 0,795 pelo número total de irmandades com dois filhos, que, em nosso caso é 14, obtém-se o número esperado daquelas constituídas apenas por indivíduos *Se*<sub>-</sub> (11,130). Operações análogas foram feitas para obter o número esperado de irmandades com 3, 4, 5 e 6 indivíduos, todos apresentando o fenótipo *secretor*.

Para obter, por exemplo, o número esperado de irmandades constituídas por dois indivíduos *Se*<sub>-</sub>, filhos de casais *Se*<sub>-</sub> × *sese*, calculou-se, inicialmente, a probabilidade de encontro de tais irmandades por intermédio de  $0,316 + (0,25 \times 0,684) = 0,487$ , pois:

$$\frac{p}{1+q} = \frac{0,480}{1,520} = 0,316 \qquad 1 - \frac{p}{1+q} = 0,684 \qquad \left(\frac{1}{2}\right)^2 = 0,25$$

Multiplicando a probabilidade 0,487 por 8 que, em nosso caso, é o número total de irmandades com dois filhos, obtém-se o número esperado daquelas constituídas apenas por indivíduos *secretores* (3,896). Operações análogas foram realizadas para obter o número esperado de irmandades com 3, 4 e 5 indivíduos.

De acordo com os valores de qui-quadrado expressos na Tabela 4.4, tem-se, portanto, que, também pelo método de Fisher (1939), pode-se aceitar a hipótese monogênica para explicar a distribuição dos fenótipos *secretor* e *não-secretor* nas populações e nas famílias.

Para finalizar o presente tópico é importante salientar que o método de Fisher (1939) deve ser usado, de preferência ao de Snyder (1932), quando as amostras de famílias não são grandes, nem constituídas por irmandades de poucos indivíduos. Isso porque as pessoas com fenótipo dominante ou recessivo que compõem a geração familiar não são amostras independentes de uma população homogênea, mas dependem da composição genética da geração paterna. Em outras palavras, as proporções de indivíduos com os fenótipos alternativos entre os filhos de casais do tipo  $A_ \times A_$  dependem das proporções de casais  $AA \times AA$ ,  $AA \times Aa$  e  $Aa \times Aa$  existentes entre esses últimos. Do mesmo modo, as proporções fenotípicas observadas entre os filhos de casais do tipo  $A_ \times aa$  dependem da proporção de casais  $AA \times aa$  e de casais  $Aa \times aa$  entre eles.

### DISTRIBUIÇÃO FAMILIAL DE POLIMORFISMOS LIGADOS AO SEXO

Consideremos um par de alelos  $A, a$  do cromossomo X e que o gene  $a$  somente se manifeste nas mulheres quando em homozigose ( $X^aX^a$ ), isto é, que o gene  $A$  seja responsável pelo fenótipo dominante  $X^AX^$  entre elas (genótipos  $X^AX^A$  e  $X^AX^a$ ). Se a frequência do gene  $A$  for  $p$  e a de seu alelo  $a$  for  $q$ , tendo  $p$  e  $q$  valores altos e sendo  $p + q = 1$ , é claro que em uma população em equilíbrio genético as mulheres com o fenótipo recessivo ocorrerão com frequência  $q^2$  e aquelas com o fenótipo dominante serão encontradas com frequência  $p^2 + 2pq = 1 - q^2$ . Entre os homens, os indivíduos  $X^AY$  e  $X^aY$  terão frequências  $p$  e  $q$ , respectivamente.

Os casais de tal população poderão ser classificados em quatro tipos, isto é,  $X^AY \times X^AX^$ ,  $X^AY \times X^aX^a$ ,  $X^aY \times X^AX^$  e  $X^aY \times X^aX^a$ , e a frequência de cada um desses tipos dependerá, evidentemente, das frequências dos fenótipos alternativos entre os homens

e as mulheres, Isso porque as probabilidades de tais casais serem sorteados da população serão as seguintes:

$$\begin{aligned} P(X^A Y \times X^A X^-) &= p(1 - q^2) \\ P(X^A Y \times X^a X^a) &= pq^2 \\ P(X^a Y \times X^A X^-) &= q(1 - q^2) \\ P(X^a Y \times X^a X^a) &= q^3 \end{aligned}$$

Em consequência disso, em uma amostra aleatória de  $N$  casais da população, os números esperados dos quatro tipos de casais em discussão não deve diferir significativamente daqueles calculados como abaixo:

$$\begin{aligned} X^A Y \times X^A X^- &= Np(1 - q^2) \\ X^A Y \times X^a X^a &= Npq^2 \\ X^a Y \times X^A X^- &= Nq(1 - q^2) \\ X^a Y \times X^a X^a &= Nq^3 \end{aligned}$$

Portanto, se a partir de uma amostra de famílias, quisermos pôr à prova uma hipótese monogênica para explicar a transmissão de fenótipos alternativos ligados ao sexo, um dos quais mostra dominância sobre o outro, deve-se, de início, verificar se a distribuição dos diferentes tipos de casais da amostra não se desvia significativamente daquela esperada e calculada como acima.

O segundo passo a ser seguido consiste do exame da distribuição dos indivíduos com os fenótipos dominante e recessivo nas irmandades geradas pelos diferentes tipos de casais. Isso porque nos casos de herança monogênica ligada ao sexo, com relação de dominância e recessividade entre os fenótipos alternativos, verifica-se que:

1. Todas as filhas dos casais  $X^A Y \times X^A X^-$  e  $X^A Y \times X^a X^a$  apresentam fenótipo dominante pois, em condições normais, todas as filhas de homens  $X^A Y$  têm, obrigatoriamente, fenótipo dominante, já que um de seus cromossomos X é de origem paterna.

2. Todos os filhos e todas as filhas de casais  $X^a Y \times X^a X^a$  apresentam fenótipo recessivo.

3. Todos os filhos de casais  $X^A Y \times X^a X^a$  manifestam fenótipo recessivo, pois, normalmente, eles somente podem herdar o seu cromossomo X de suas mães.

4. Se o número total de filhos do sexo masculino gerados pelos casais  $X^A Y \times X^A X^-$  for  $m$  e a frequência do gene  $a$  for  $q$ , o número esperado de filhos  $X^a Y$  desses casais não deve desviar-se significativamente de  $\frac{qm}{1+q}$ . Isso porque consideramos que a probabilidade de um filho de um casal  $X^A Y \times X^A X^-$  ser  $X^a Y$  é igual à metade da probabilidade de sua mãe ser heterozigota, dado que ela tem o fenótipo dominante. De fato, a probabilidade de uma mulher heterozigota  $X^A X^a$  transmitir o cromossomo  $X^a$  a um filho é  $\frac{1}{2}$ , enquanto que a probabilidade de uma mulher  $X^A X^-$  ser heterozigota é  $P(X^A X^a | X^A X^-) = \frac{2pq}{p^2 + 2pq} = \frac{2q}{1+q}$ .

Portanto,  $\frac{1}{2} \cdot \frac{2q}{1+q} = \frac{q}{1+q}$ .

5. Se  $f$  for o número de filhas e  $m$  for o número de filhos de casais  $X^A Y \times X^A X^-$ , o número esperado de filhas com o fenótipo recessivo será  $\frac{qf}{1+q}$  e o número esperado de filhos com esse fenótipo será  $\frac{qm}{1+q}$ , pelas mesmas razões expostas no item anterior. Os números esperados de filhos e de filhas com o fenótipo dominante serão obtidos pelas diferenças entre  $f$  e  $m$  e os números esperados de indivíduos com o fenótipo recessivo de cada sexo. A Tabela 5.4 resume o que foi discutido até agora no presente tópico.

Tabela 5.4. Distribuição esperada em uma amostra aleatória de  $N$  famílias segundo os fenótipos  $A_-$  e  $aa$  determinados por um par de alelos  $A, a$  do cromossomo X, com frequências iguais, respectivamente, a  $p$  e  $q = 1 - p$ , quando a população da qual procede a amostra está em equilíbrio genético.

Casais		Filhas			Filhos		
Tipo	No.	$X^a X^a$	$X^A X^-$	Total	$X^a Y$	$X^A Y$	Total
$X^A Y \times X^A X^-$	$Np(1 - q^2)$	-	$f_1$	$f_1$	$\frac{qm_1}{1+q}$	$m_1 - \frac{qm_1}{1+q}$	$m_1$
$X^A Y \times X^a X^a$	$Npq^2$	-	$f_2$	$f_2$	$m_2$	-	$m_2$
$X^a Y \times X^A X^-$	$Nq(1 - q^2)$	$\frac{qf_3}{1+q}$	$f_3 - \frac{qf_3}{1+q}$	$f_3$	$\frac{qm_3}{1+q}$	$m_3 - \frac{qm_3}{1+q}$	$m_3$
$X^a Y \times X^a X^a$	$Nq^3$	$f_4$	-	$f_4$	$m_4$		$m_4$

Uma outra alternativa para verificar se a distribuição dos filhos de casais  $X^A Y \times X^A X^-$  e dos filhos e filhas dos casais  $X^a Y \times X^A X^-$  está de acordo com a hipótese

monogênica, consiste em comparar os números observados e esperados de irmandades constituídas apenas por indivíduos com o fenótipo dominante, bem como aquelas que incluem pelo menos um indivíduo com o fenótipo recessivo, por intermédio, respectivamente, de

$$\sum n \left[ \frac{p}{1+q} + \left(\frac{1}{2}\right)^i \left(1 - \frac{p}{1+q}\right) \right] \quad \text{e} \quad \sum n - \sum n \left[ \frac{p}{1+q} + \left(\frac{1}{2}\right)^i \left(1 - \frac{p}{1+q}\right) \right]$$

No caso de irmandades geradas por casais  $X^A Y \times X^A X^-$ ,  $n$  indica o número de irmandades com  $i$  indivíduos do sexo masculino, enquanto que, em relação às irmandades geradas por casais  $X^a Y \times X^A X^-$ ,  $n$  se refere ao total de indivíduos nas irmandades.

Um exemplo poderá esclarecer melhor o que foi discutido até agora no presente tópico. Assim, suponhamos que não soubéssemos que o grupo sanguíneo  $Xg(a-)$  é determinado por um gene do cromossomo X, que se manifesta nas mulheres homocigotas e nos homens hemizigotos, enquanto que o grupo sanguíneo  $Xg(a+)$  é determinado por um alelo desse gene. Apesar disso, poderíamos com base nos dados da Fig. 2.4, a respeito de 50 famílias estudadas por Mann *et al.* (1962), sugerir a hipótese da existência de um par de alelos  $Xg^a, Xg$  para explicar esses grupos sanguíneos, sendo o alelo  $Xg^a$  determinante do grupo  $Xg(a+)$  e o alelo  $Xg$  em homocigose nas mulheres e em hemizigose nos homens o responsável pelo grupo  $Xg(a-)$ .

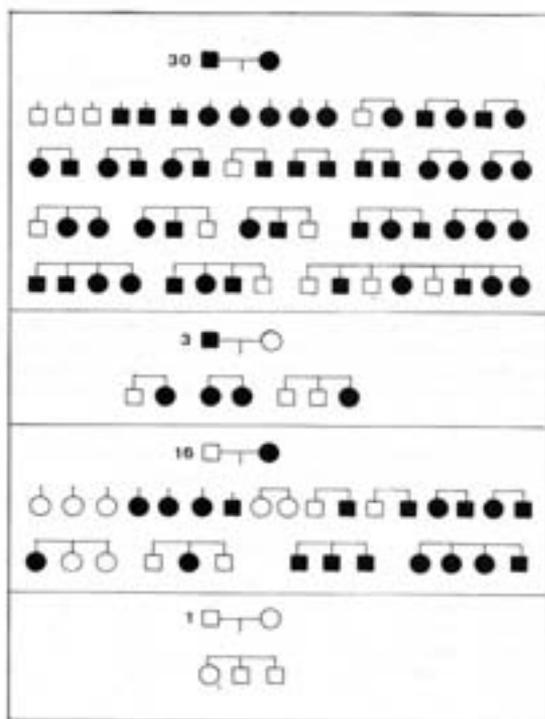


Fig. 2.4 Dados a respeito de 50 famílias cujos elementos tiveram suas hemácias classificadas com um anti-soro anti- $Xg^a$  por Mann *et al.* (1962). Os indivíduos  $Xg(a+)$  estão representados por símbolos escuros e os  $Xg(a-)$  por símbolos claros.

De fato, a sugestão de que esses fenótipos dependem do sexo dos indivíduos é dada pelo fato de que o grupo sanguíneo Xg(a+) é mais freqüente nas mulheres. Assim, levando em conta apenas os casais da Figura 2.4 tem-se que, 33 dentre os 50 homens (66%) eram Xg(a+), ao passo que, dentre 50 mulheres, 46 manifestavam esse grupo sanguíneo (92%). Por outro lado, fala a favor de que o grupo sanguíneo Xg(a-) é um caráter recessivo ligado ao sexo. a verificação, na Figura 2.4, de que:

1. O único casal Xg(a-) × Xg(a-) gerou uma filha e dois filhos Xg(a-).
2. Todas as filhas de pai Xg(a+) eram Xg(a+).
3. Todos os filhos de mãe Xg(a-) eram Xg(a-);

Para pôr à prova a hipótese em discussão precisamos, evidentemente, da estimativa das freqüências dos supostos alelos Xg<sup>a</sup> e Xg. Assim, partindo de dados de Mann *et al.* (1962) que, em 342 caucasóides (154 homens e 188 mulheres), encontraram 95 homens Xg(a+), 59 homens Xg(a-), 167 mulheres Xg(a+) e 21 mulheres Xg(a-), podemos estimar a freqüência *q* do alelo Xg em 0,356, pois:

$$q = \frac{-95 + \sqrt{(95)^2 + 4[(2 \times 188) + 154][(2 \times 21) + 59]}}{2[(2 \times 188) + 154]} = 0,356$$

o que equivale a dizer que a estimativa da freqüência do gene Xg<sup>a</sup> é *p* = 1 - 0,356 = 0,644

Com base nas estimativas *p* e *q* calculamos a distribuição esperada dos diferentes tipos de casais e seus filhos, de acordo com a hipótese monogênica, dispondo os dados como na Tabela 6.4. Nessa tabela os números observados foram obtidos a partir da Figura 2.4 e os esperados foram calculados segundo as fórmulas expressas na Tabela 5.4.

Tabela 6.4. Números observados e esperados de casais e de filhos Xg(a+) e Xg(a-) na amostra de 50 famílias estudadas por Mann *et al.* (1962). Entre parênteses estão assinalados os valores esperados.

Casais		Filhas			Filhos		
Marido × Mulher	No.	Xg(a-)	Xg(a+)	Total	Xg(a-)	Xg(a+)	Total
Xg(a+) × Xg(a+)	30 (28,1)	- (-)	29 (29,0)	29	12 (9,2)	23 (25,8)	35
Xg(a+) × Xg(a-)	3 ( 4,1)	- (-)	4 ( 4,0)	4	3 (3,0)	-	3
Xg(a-) × Xg(a+)	16(15,5)	7 (4,5)	10 ( 2,5)	17	4 (3,4)	9 (9,6)	13
Xg(a-) × Xg(a-)	1 ( 2,3)	1 (1,0)	- ( - )	1	2 (2,0)	-	2

Podemos, agora, comparar as proporções observadas e esperadas dos diferentes tipos de casais por intermédio de um qui-quadrado com três graus de liberdade, já que as quatro classes esperadas são obtidas a partir de uma única informação da amostra, isto é, o número  $N$  de casais. Assim, obtemos  $\chi^2_{(3)} = 1,174$ ;  $0,70 < P < 0,80$ , que satisfaz a hipótese monogênica.

No concernente à geração filial temos que comparar apenas as proporções observadas e esperadas de filhos dos casais  $Xg(a+) \times Xg(a+)$  e de filhas e filhos dos casais constituídos por marido  $Xg(a-)$  e mulher  $Xg(a+)$ . Tais comparações nos fornecem, respectivamente:

$$\begin{aligned}\chi^2_{(1)} &= 1,156; 0,20 < P < 0,30 \\ \chi^2_{(1)} &= 1,889; 0,10 < P < 0,20 \\ \chi^2_{(1)} &= 0,144; 0,70 < P < 0,80\end{aligned}$$

o que nos leva, evidentemente, a aceitar a hipótese monogênica proposta.

Um outro tipo de análise da geração filial pode ser feito a partir da distribuição dos indivíduos do sexo masculino nas irmandades geradas pelos casais  $Xg(a+) \times Xg(a+)$ , e de todos os componentes das irmandades geradas por homens  $Xg(a-)$  e mulheres  $Xg(a+)$ . Assim, levando em conta apenas os indivíduos do sexo masculino constatamos que as irmandades geradas por casais  $Xg(a+) \times Xg(a+)$  são, agora, 22, distribuídas como na Tabela 7.4, dentre as quais verificamos que 12 são constituídas somente por indivíduos  $Xg(a+)$ . Considerando os indivíduos de ambos os sexos das irmandades geradas pelos casais marido  $Xg(a-) \times$  mulher  $Xg(a+)$ , constatamos que 8 dentre as 16 são compostas somente por indivíduos  $Xg(a+)$ .

Os valores de qui-quadrado apresentados na Tabela 7.4 indicam, novamente, que a hipótese monogênica deve ser aceita para explicar a distribuição dos grupos sanguíneos  $Xg(a+)$  e  $Xg(a-)$  nas populações e nas famílias.

### **A UTILIZAÇÃO DE DADOS A RESPEITO DE MÃES E FILHOS**

Um método mais prático do que o estudo de famílias completas para se fazer a análise genética de caracteres polimórficos é o estudo de pares de indivíduos, constituídos por mãe e filho(a), se o caráter é supostamente autossômico, ou por mãe e filho, se ele é supostamente ligado ao sexo. Tal tipo de investigação é particularmente interessante

quando os caracteres sob investigação são congênitos, isto é, já se expressam ao nascimento, pois o trabalho pode ser facilmente executado dentro de maternidades.

Tabela 7.4. Análise das irmandades geradas pelos casais  $Xg(a+) \times Xg(a+)$  levando em conta apenas os indivíduos do sexo masculino, e das irmandades geradas pelos casais constituídos por homem  $Xg(a-)$  e mulher  $Xg(a+)$ .  $i$  - tamanho da irmandade;  $n$  - número de irmandades de tamanho  $i$ ;  $P$ - probabilidade de uma irmandade de tamanho  $i$  ser constituída somente por indivíduos  $Xg(a+)$ .

Casais	$i$	$n$	$P$	Esp.	Obs.
Marido $Xg(a+)$ $\times$ Mulher $Xg(a+)$	1	13	0,738	9,594	8
	2	7	0,606	4,242	4
	3	1	0,541	0,541	-
	5	1	0,491	0,491	-
	Total	22		14,868	12
Marido $Xg(a-)$ $\times$ Mulher $Xg(a+)$	1	7	0,738	5,166	4
	2	5	0,606	3,030	2
	3	3	0,541	1,623	1
	4	1	0,508	0,508	1
	Total	16		10,327	8

Casais	Irmandades	Obs.	Esp.	$\chi^2_{(1)}$
Marido $Xg(a+) \times$ Mulher $Xg(a+)$	Apenas $Xg(a+)$	12	14,868	1,706; $0,10 < P < 0,20$
	Com $Xg(a-)$	10	7,132	
Marido $Xg(a-) \times$ Mulher $Xg(a+)$	Apenas $Xg(a+)$	8	10,327	1,479; $0,20 < P < 0,30$
	Com $Xg(a-)$	8	5,673	

Na Tabela 8.4 foram reunidos quatro quadros que mostram as frequências esperadas dos pares mãe-filho(a) em uma população em equilíbrio genético. Para exemplificar o modo pelo qual foram construídos os quadros da Tabela 8.4 calcularemos as frequências dos pares de mãe-filho que devem ser encontrados quando a mãe é homozigota de um gene autossômico  $A$  (mãe  $AA$ ). Evidentemente, as mães homozigotas  $AA$  não podem gerar filhos  $aa$ , mas apenas filhos  $AA$  e  $Aa$  com frequências que dependerão das frequências com que ocorrem os genótipos  $AA$ ,  $Aa$  e  $aa$  na população.

Assim, se  $p$  for a frequência do gene  $A$ , e  $q$  a do seu alelo  $a$ , ter-se-á em uma população em equilíbrio de Hardy e Weinberg que os casais compostos por mulher  $AA$  e marido  $AA$ ,  $Aa$  ou  $aa$  deverão ocorrer nas seguintes proporções:

$$\begin{aligned} \text{mulher } AA \times \text{marido } AA &= p^4 \\ \text{mulher } AA \times \text{marido } Aa &= 2p^3q \\ \text{mulher } AA \times \text{marido } aa &= p^2q^2 \end{aligned}$$

Tabela 8.4. Freqüências dos tipos de pares mãe-filho(a) em uma população em equilíbrio genético, levando-se em conta caracteres autossômicos e ligados ao sexo, com e sem relação de dominância.

**A. Caracteres autossômicos sem dominância.**

Mãe	Filho (a)			Total
	AA	Aa	aa	
AA	$p^3$	$p^2q$	-	$p^2$
Aa	$p^2q$	$pq$	$pq^2$	$2pq$
aa	-	$pq^2$	$q^3$	$q^2$
Total	$p^2$	$2pq$	$q^2$	1,0

**B. Caracteres autossômicos com dominância.**

Mãe	Filho (a)		Total
	A <sub>-</sub>	aa	
A <sub>-</sub>	$p(1+pq)$	$pq^2$	$P^2+2pq$
Aa	$pq^2$	$q^3$	$q^2$
Total	$p^2+2pq$	$q^2$	1,0

**C. Caracteres ligados ao sexo sem dominância.**

Mãe	Filho		Total
	X <sup>A</sup> Y	X <sup>a</sup> Y	
X <sup>A</sup> X <sup>A</sup>	$p^2$	-	$p^2$
X <sup>A</sup> X <sup>a</sup>	$pq$	$pq$	$2pq$
X <sup>a</sup> X <sup>a</sup>	-	$q^2$	$q^2$
Total	$p$	$q$	1,0

**D. Caracteres ligados ao sexo com dominância.**

Mãe	Filho		Total
	X <sup>A</sup> Y	X <sup>a</sup> Y	
X <sup>A</sup> X <sup>-</sup>	$p^2 + pq$	$pq$	$p^2 + 2pq$
X <sup>a</sup> X <sup>a</sup>	-	$q^2$	$q^2$
Total	$p$	$q$	1,0

Os filhos desses casais, por sua vez, se distribuirão como abaixo:

$$\begin{aligned} \text{mulher } AA \times \text{marido } AA &= p^4 \text{ filhos } AA \\ \text{mulher } AA \times \text{marido } Aa &= p^3q \text{ filhos } AA \text{ e } p^3q \text{ filhos } Aa \\ \text{mulher } AA \times \text{marido } aa &= p^2q^2 \text{ filhos } Aa \end{aligned}$$

Portanto, a frequência total de filhos *AA* de mães *AA* será  $p^4 + p^3q = p^3(p + q) = p^3$  e o de filhos *Aa* dessas mães será  $p^2q^2 + p^3q = p^2q(p + q) = p^2q$ , pois  $p + q = 1$ .

Para demonstrar como a análise dos dados de pares mãe-filho(a) poderia ser utilizada no estudo de um caráter supostamente autossômico monogênico, consideremos na Figura 1.4 apenas as mães e o primeiro indivíduo de cada irmandade. Com tal conduta, obtemos os dados necessários à construção da Tabela 9.4 na qual os valores esperados são obtidos a partir das fórmulas expressas na Tabela 8.4-B. As estimativas de  $p$  e  $q$  são aquelas obtidas por intermédio de dados populacionais ( $q = 0,520$  e  $p = 0,480$ ).

Tabela 9.4. Proporções observadas e esperadas de pares mãe-filho(a) com os fenótipos *secretor* e *não-secretor*. Entre parênteses foram assinalados os valores esperados.

Mãe	F i l h o ( a )		
	Secretor(a)	Não-Secretor(a)	Total
Secretora	32 (30,0)	7 (6,5)	39 (36,5)
Não-Secretora	4 (6,5)	7 (7,0)	11 (13,5)
Total	36 (36,5)	14(13,5)	50 (50,0)

Os dados da Tabela 9.4 permitem constatar que as proporções observadas de mães *secretoras* (39) e *não-secretoras* (11) não se desviam significativamente das esperadas ( $\chi^2_{(1)} = 0,634$ ;  $0,30 < P < 0,50$ ). Do mesmo modo, não é possível constatar desvio significativo entre as proporções observadas e esperadas de *secretores* e *não-secretores* nos filhos ( $\chi^2_{(1)} = 0,026$ ;  $0,80 < P < 0,90$ ). Além disso, as proporções desses fenótipos alternativos nas mães não difere daquela verificada nos filhos ( $\chi^2_{(1)} = 0,480$ ;  $0,30 < P < 0,50$ ).

No concernente às quatro classes de mães e filhos, a comparação dos valores observados com os esperados permite concluir, também, pela falta de significação dos desvios, pois o qui-quadrado com três graus de liberdade é igual a 1,133 ( $0,98 < P < 0,99$ ). Tais comparações, portanto, permitiriam aceitar a hipótese de herança monogênica sem a necessidade de estudar famílias completas.

A título de exercício o leitor pode aplicar este tipo de investigação aos dados da Figura 2.4, considerando os dados formados pelo primeiro filho do sexo masculino das irmandades e suas respectivas mães.

## A ANÁLISE FAMILIAL DE CARACTERES COMPLEXOS

Nos tópicos anteriores tivemos a oportunidade de discutir a análise familiar de caracteres que tinham uma transmissão estritamente mendeliana, ou seja, para explicar a manifestação e transmissão hereditária desses caracteres podíamos desprezar o efeito do ambiente e da constelação gênica individual, isto é, não levávamos em conta a participação multifatorial. Em relação a numerosos outros caracteres humanos, denominados *caracteres complexos*, essa simplicidade, entretanto, não pode ser aceita *a priori*. Tais caracteres requerem, para sua análise familiar, métodos mais poderosos, que exigem programas de computação, sendo o mais utilizado atualmente aquele que é denominado *método unificado de análise de segregação* (Lalouel *et al.*, 1983).

Esse método de análise de segregação parte de um modelo, denominado *modelo mendeliano misto*, o qual considera que a manifestação do caráter em estudo depende da contribuição independente dos alelos de um loco autossômico principal, de um componente poligênico e, ainda, de fatores aleatórios do ambiente. No loco autossômico principal supõe-se a existência de um par de alelos  $A, a$  com frequências  $p$  e  $q = 1 - p$ , sendo a distribuição populacional dos genótipos resultantes ( $AA$ ,  $Aa$  e  $aa$ ) feita segundo a lei de Hardy e Weinberg, isto é, segundo  $(p+q)^2$ , isto é,  $AA: Aa: aa :: p^2 + 2pq + q^2$ .

A posição da média do genótipo heterozigoto ( $Aa$ ) em relação às médias dos homozigotos é designada por *grau de dominância* e representada por  $d$ . No caso de um dos fenótipos em estudo ser completamente dominante ( $AA = Aa$ ),  $d$  será igual a 1 e, em havendo codominância,  $d$  será igual a 0,5. Se a dominância for parcial ter-se-á  $0,5 < d < 1$ . A distância entre as médias dos dois genótipos homozigotos ( $AA$  e  $aa$ ) é denominada *deslocamento* e representada por  $t$ .

Aceita-se que a variação em torno da média de cada um desses genótipos tem distribuição normal, com uma variância comum  $C + E$  a todos, sendo  $C$  a variância devida a efeitos multifatoriais transmissíveis e  $E$  o componente da variância residual do ambiente que não é transmitido dentro das famílias. A variância fenotípica total é designada por  $V$  e a razão  $\frac{C}{V} = H$  é a herdabilidade, que reflete a transmissão poligênica.

No modelo misto são quatro os parâmetros estimados, isto é, o grau de dominância ( $d$ ), a distância entre as médias dos homozigotos ( $t$ ), a frequência de um dos alelos ( $q$ ) e a

herdabilidade (**H**). Nesse modelo, as probabilidades de os genótipos  $AA$ ,  $Aa$  e  $aa$  transmitirem o alelo  $A$ , simbolizados por  $\tau_{AA}$ ,  $\tau_{Aa}$  e  $\tau_{aa}$ , não são estimadas, e sim fixadas em, respectivamente, 1,  $\frac{1}{2}$  e 0, mas num outro modelo, aqui chamado de **misto completo**, as probabilidades de os genótipos  $AA$ ,  $Aa$  e  $aa$  transmitirem o alelo  $A$  são parâmetros estimados. No modelo misto completo, portanto, estimam-se sete parâmetros. Num outro modelo, oposto ao modelo misto completo, considera-se que as probabilidades de os genótipos  $AA$ ,  $Aa$  e  $aa$  transmitirem o alelo  $A$  não diferem entre si, razão pela qual estima-se a mesma probabilidade para os três genótipos, isto é,  $\tau_{AA} = \tau_{Aa} = \tau_{aa}$ . Neste modelo são, pois, cinco os parâmetros estimados. Dependendo do carácter analisado pode-se empregar variações do modelo misto completo, ao deixar de estimar o grau de dominância ( $d$ ) e(ou) a herdabilidade (**H**).

Em outros modelos, usados para comparação com o modelo misto, o número de parâmetros estimados é menor. Assim, no modelo que prega a inexistência de um gene principal, denominado **modelo sem gene principal**, o único parâmetro estimado é a herdabilidade. No modelo **sem componente multifatorial** são três os parâmetros estimados ( $d$ ,  $t$  e  $q$ ), enquanto que no **modelo esporádico**, nenhum parâmetro é estimado. Há ainda outros modelos usados para comparação com o modelo misto, a saber, o **modelo recessivo**, o **modelo aditivo** e o **modelo dominante**. Nesses modelos tanto os parâmetros  $\tau_{AA}$ ,  $\tau_{Aa}$  e  $\tau_{aa}$  quanto o grau de dominância e a herdabilidade são fixados em consonância com o que eles pregam, de sorte que os parâmetros estimados são apenas dois ( $t$  e  $q$ ).

Os testes das hipóteses realizados são avaliados por intermédio do **quociente de verossimilhança (q.v.)** o qual é menos duas vezes o logaritmo natural da verossimilhança máxima, isto é  $q.v. = -2\ln L$ . O quociente de verossimilhança também pode ser representado por  $q.v. = -2\ln L + c$ , sendo  $c$  uma constante, que consiste no menor dos  $q.v.$  calculados. Em qualquer caso o  $q.v.$  pode ser lido como um qui-quadrado. Também se deve levar em conta a probabilidade de averiguação ( $\pi$ ) das famílias estudadas.

Para melhor entender o método unificado de análise de segregação, tomemos um exemplo tirado do trabalho de Feitosa *et al.* (1996) sobre **reação de Mitsuda**, que é o nome dado à resposta tardia, lida 28 a 30 dias após a inoculação intradérmica de 0,1 ml de uma suspensão autoclavada de *Mycobacterium leprae*, extraído de hansenomas de pacientes virgens de tratamento ou de tecidos de tatus infectados por *M. leprae*. Essa suspensão é

denominada *lepromina* ou *mitsudina* e a reação por ela provocada depende dos eventos que sucedem a fagocitose dos bacilos pelos macrófagos da pele (histiócitos). A ausência de resposta clínica após 30 dias da inoculação da lepromina é denominada reação negativa, enquanto que uma infiltração franca, pápula ou nódulo com mais de 3 milímetros de diâmetro é denominada reação positiva, a qual, de acordo com a intensidade da resposta pode ser classificada como + (3 a 5 mm), ++ (mais de 5 mm) e +++ (infiltração nodular ulcerada). Um infiltração discreta com menos de 3 milímetros de diâmetro é considerada duvidosa ( $\pm$ ). No estudo de Feitosa *et al.* (1996), do qual foi extraída a Tabela 1.4 foram consideradas apenas duas classes de reação de Mitsuda, a reação negativa (- e  $\pm$ ) e a positiva (+, ++ e +++).

A reação de Mitsuda tem grande importância em hansenologia porque a resposta positiva indica imunidade à hanseníase do tipo virchowiano (lepra lepromatosa). Aliás, a resposta negativa ao teste de Mitsuda é característica dos pacientes virchowianos. O leitor interessado em maiores informações sobre essa reação e o seu valor prognóstico e diagnóstico pode consultar Beiguelman (1999, 2001).

Estudos feitos na década de 60 já haviam mostrado que a reação de Mitsuda apresentava associação familiar tanto em famílias sem hanseníase (Beiguelman, 1962; Beiguelman e Quagliato, 1965) quanto em famílias de hansenianos (Beiguelman, 1965). Essas análises familiares indicaram que a reação de Mitsuda positiva ( $\geq +$ ) era o fenótipo dominante sobre a resposta negativa (- e  $\pm$ ), mas, com os recursos estatísticos existentes à época em que foram realizados tais trabalhos, a interpretação monogênica ficava sujeita a críticas. Vejamos, pois, o que resultou da retomada do estudo dessa reação por Feitosa *et al.* (1996) com o emprego do método unificado de análise de segregação (Lalouel *et al.*, 1983).

Os dados da análise de segregação reunidos na Tabela 10.4 permitem várias **comparações levando em conta as diferenças entre os quocientes de verossimilhança (q.v.)**, os quais, como foi mencionado, são lidos como qui-quadrado, sendo obtidos os graus de liberdade de cada qui-quadrado pela diferença entre os parâmetros estimados dos modelos comparados. Visto que o modelo mendeliano misto pressupõe que, para a manifestação do caráter em estudo, devem ser levadas em conta as contribuições independentes de um loco principal, de um componente multifatorial e do efeito do ambiente, está claro que esse modelo também prega a existência de agregação familiar do

caráter em estudo. Isso porque, embora a associação familiar não implique, obrigatoriamente, na hereditariedade de um caráter, o oposto é verdadeiro, isto é, todo o caráter hereditário é familiar. O **q.v.** do modelo mendeliano misto deve, por isso, ser comparado com o **q.v.** do modelo esporádico, que propõe a inexistência de agregação familiar e, em seguida, com o q.v. do modelo que prega a inexistência de um gene principal e com o q.v. do modelo que fala a favor da inexistência de um componente multifatorial.

Tabela 10.4. Análise de segregação da reação de Mitsuda, levando em conta 544 famílias nucleares com 2.925 indivíduos.

*d* = grau de dominância; *t* = deslocamento; *q* = frequência do alelo *a*; H = herdabilidade;  $\tau_{AA}$ ,  $\tau_{Aa}$  e  $\tau_{aa}$  = probabilidades de os genótipos *AA*, *Aa* e *aa* transmitirem o alelo *A*; q.v. = quociente de verossimilhança (-2lnL+c), P.E. = parâmetros estimados; G.L. = graus de liberdade. Em itálico negrito foram assinalados os parâmetros estimados. (Modificado de Feitosa *et al.*, 1996).

MODELO	<i>d</i>	<i>t</i>	<i>q</i>	H	$\tau_{AA}$	$\tau_{Aa}$	$\tau_{aa}$	Q.v.	P.E.	Teste	$\chi^2$	G.L.	P
1. Misto	<b>0,81</b>	<b>1,98</b>	<b>0,47</b>	<b>0,0*</b>	1	½	0	0,07	4				
2. Esporádico	0	0	0	0	-	-	-	238,57	0	1×2	238,50	4	<0,001
3. Sem gene principal	0	0	0	<b>0,66</b>	-	-	-	27,87	1	1×3	27,80	3	<0,001
4. Sem comp.mulfatorial	<b>0,81</b>	<b>1,98</b>	<b>0,47</b>	0	1	½	0	0,07	3	1×4	0	1	≅ 1
5. Misto completo	<b>0,80</b>	<b>1,99</b>	<b>0,47</b>	<b>0,0*</b>	<b>1,0*</b>	<b>0,49</b>	<b>0,0*</b>	0,00	7	1×5	0,07	3	>0,99
6. $\tau_{AA} = \tau_{Aa} = \tau_{aa}$	<b>0,35</b>	<b>1,60</b>	<b>0,17</b>	<b>0,0*</b>	<b>1,0*</b>			198,28	5	5×6	198,28	2	<0,001
7. Mendeliano recessivo	0	<b>1,86</b>	<b>0,88</b>	0	1	½	0	35,52	2	1×7	35,45	2	<0,001
8. Mendeliano aditivo	½	<b>2,44</b>	<b>0,47</b>	0	1	½	0	13,56	2	1×8	13,49	2	<0,01
9. Mendeliano dominante	1	<b>1,65</b>	<b>0,45</b>	0	1	½	0	9,15	2	1×9	9,08	2	<0,02

\*Atingiu o limite.

Na primeira dessas comparações, isto é, entre os modelos mendeliano misto (modelo No. 1) e esporádico (modelo No. 2), testamos a hipótese nula de inexistência de associação familiar contra a hipótese alternativa de que, para a manifestação do caráter em estudo, devemos aceitar que ela existe. Visto que o q.v. do modelo esporádico foi 238,57 e o do modelo mendeliano misto foi 0,07 tem-se  $\chi^2_{(4)} = 238,57 - 0,07 = 238,50$ ;  $P \ll 0,001$ . Esse qui-quadrado tem 4 graus de liberdade porque a diferença entre os parâmetros estimados é 4 (4 no modelo mendeliano misto e 0 no modelo esporádico). O valor do qui-quadrado encontrado indica, pois, que devemos rejeitar a hipótese nula que fala a favor do modelo esporádico, isto é, a inexistência de associação familiar, e devemos optar pela hipótese alternativa, que pede a aceitação do modelo mendeliano misto e, com isso, a aceitação da existência de agregação familiar.

Na segunda comparação, feita entre o q.v. do modelo mendeliano misto (modelo No. 1) e o q.v. do modelo que propõe a inexistência de um gene principal (modelo No. 3), a hipótese nula é a de que não há gene principal contra a alternativa de que ele existe. Por ser o q.v. do modelo sem gene principal igual a 27,87 e o do modelo mendeliano misto igual a 0,07, tem-se  $\chi^2_{(3)} = 27,87 - 0,07 = 27,80$ ;  $P < 0,001$ , possuindo esse qui-quadrado 3 graus de liberdade, porque o modelo mendeliano misto tem 4 parâmetros estimados enquanto o modelo que prega a inexistência de um gene principal tem um único ( $4 - 1 = 3$ ). Esse valor de qui-quadrado também fala a favor da rejeição da hipótese nula e pela aceitação da hipótese alternativa, que prega a existência de um gene principal.

Na terceira comparação, feita entre o q.v. do modelo mendeliano misto (modelo No.1) e o q.v. do modelo que prega a inexistência de um componente multifatorial (modelo No.4), a hipótese nula é a de que não existe componente multifatorial e a alternativa é a de que ele existe. Por serem os q.v.s de ambos modelos iguais a 0,07, e por ter o modelo que propõe a inexistência de um componente multifatorial 3 parâmetros estimados, a hipótese nula pode ser aceita, isto é, pode-se rejeitar a importância da participação de um componente multifatorial na manifestação do caráter em estudo, pois  $\chi^2_{(1)} = 0,07 - 0,07 = 0$ ;  $P \cong 1$ .

A coerência dos resultados apresentados na Tabela 1.40 é ressaltada pelo fato de que a hipótese nula de transmissão mendeliana do modelo que chamamos de misto completo (modelo No. 5) é facilmente aceita quando o seu q.v. é comparado ao q.v. do modelo mendeliano misto (modelo No. 1), fornecendo um pequeno qui-quadrado com três graus de liberdade ( $\chi^2_{(3)} = 0,07 - 0,0 = 0,07$ ;  $P > 0,99$ ), pois, levando em conta os parâmetros estimados, tem-se  $7 - 4 = 3$ . Na Tabela 1.40 é fácil constatar que, no modelo misto completo, os valores estimados de  $\tau_{AA}$ ,  $\tau_{Aa}$  e  $\tau_{aa}$  ficaram muito próximos daqueles fixados pela teoria nos modelos No.s 1, 4, 7, 8 e 9, isto é, muito semelhantes a 1,  $\frac{1}{2}$  e 0. Em oposição, a hipótese nula que afirma não haver transmissão mendeliana de um gene principal (modelo No. 6) contra a hipótese de que ela existe (modelo No. 5) é facilmente rejeitada ( $\chi^2_{(2)} = 198,28 - 0,00 = 198,28$ ;  $P \ll 0,001$ ).

Finalmente, comparando os valores dos q.v.s dos modelos estritamente recessivo (modelo No.7), aditivo (modelo No.8) e dominante (modelo No. 9) com o q.v. do modelo mendeliano misto (modelo No.1) conclui-se que as hipóteses propostas por esses três

modelos devem ser rejeitadas. De fato, levando em conta os valores de qui-quadrado, todos com 2 graus de liberdade, porque no modelo No. 1 foram estimados 4 parâmetros enquanto que nos outros três modelos só foram estimados 2 parâmetros, verifica-se que o modelo estritamente recessivo não pode ser aceito porque  $\chi^2_{(2)} = 35,52 - 0,07 = 35,45$ ;  $P < 0,001$  e o modelo mendeliano aditivo deve ser rejeitado porque  $\chi^2_{(2)} = 13,56 - 0,07 = 13,49$ ;  $P < 0,01$ . O modelo estritamente dominante também não pode ser aceito porque  $\chi^2_{(2)} = 9,15 - 0,07 = 9,08$ ;  $P < 0,02$  o que permite aceitar o modelo mendeliano misto, que fala a favor de um efeito parcialmente dominante da reação Mitsuda-positivo sobre a reação Mitsuda-negativo, com  $d = 0,81$  e desvio padrão estimado em 0,041.

### QUESTÕES E RESPOSTAS

**Q 1.** Ao estudar 304 famílias dinamarquesas com o auxílio do anti-soro anti- $P_1$ , Henningsen (1950) encontrou a distribuição apresentada na tabela abaixo. Nessa tabela os indivíduos do grupo sanguíneo  $P_1$  são aqueles cujas hemácias aglutinam quando suspensas no anti-soro anti- $P_1$ , sendo os do grupo  $P_2$  aqueles em que isso não ocorre.

Casais		Filhos		
Tipo	No.	$P_1$	$P_2$	Total
$P_1 \times P_1$	194	471	53	524
$P_1 \times P_2$	93	169	71	240
$P_2 \times P_2$	17	1	38	39

Considerando que na população da qual procedem as famílias estudadas a frequência do grupo sanguíneo  $P_2$  pode ser estimada em 20,25%, teste a hipótese de que os grupos sanguíneos  $P_1$  e  $P_2$  decorrem de um par de alelos autossômicos e que do fenótipo  $P_1$  é dominante, empregando o método de Snyder (1932). Para tanto, parta do pressuposto de que já foi demonstrado que:

- As proporções de indivíduos  $P_1$  e  $P_2$  entre os cônjuges do sexo masculino não diferem significativamente das observadas entre os do sexo feminino.
- O número de casais compostos por marido  $P_1$  e mulher  $P_2$  e o de casais constituídos por marido  $P_2$  e mulher  $P_1$  não diferem significativamente.
- A proporção de indivíduos  $P_1$  e  $P_2$  entre os filhos dos diferentes tipos de casais é independente do sexo.

**R 1.** Pode-se aceitar a hipótese de que os grupos sanguíneos  $P_1$  e  $P_2$  decorrem de um par de alelos autossômicos sendo o fenótipo  $P_1$  dominante. De fato, se  $q^2 = 0,2025$ , tem-se  $q = \sqrt{0,2025} = 0,45$  e, comparando, inicialmente, as proporções observadas dos três tipos de casais com as esperadas numa população em equilíbrio de Hardy e Weinberg obtém-se um qui-quadrado com dois graus de liberdade ( $\chi^2_{(2)} = 1,897$ ;  $0,30 < P < 0,50$ ) que indica que as diferenças entre essas proporções não são significativas.

Casais	Observado	Esperado	
$P_1 \times P_1$	194	$304 (1 - 0,2025)^2$	= 193,3
$P_1 \times P_2$	93	$608 \times 0,2025 (1 - 0,2025)$	= 98,2
$P_2 \times P_2$	17	$304 \times 0,0410$	= 12,5
Total	304		304,0

Em seguida, comparando os números observados de filhos  $P_1$  e  $P_2$  resultantes dos casais  $P_1 \times P_1$  e  $P_1 \times P_2$  com os esperados segundo o método de Snyder (1932) obtém-se valores de qui-quadrado que indicam não serem significativas as diferenças encontradas.

Casais	Filhos $P_1$		Filhos $P_2$		$\chi^2_{(1)}$
	Obs.	Esp.	Obs.	Esp.	
$P_1 \times P_1$	471	473,7	53	50,3	0,160; $0,50 < P < 0,70$
$P_1 \times P_2$	169	165,5	71	74,4	0,225; $0,50 < P < 0,70$
$P_2 \times P_2$	1	0	38	39,0	

**Q 2.** Ao estudar 50 famílias quanto ao fenótipo secretor de substâncias grupo-específicas do sistema ABO, Schiff e Sasaki (1932) encontraram a distribuição apresentada na tabela abaixo.

Casais	Irmandades		Número de Irmandades	
	<i>i</i>	<i>n</i>	Apenas Secretores	Com Não-Secretores
$Se_ \times Se_$	1	4	3	1
	2	9	7	2
	3	3	3	-
	4	3	3	-
	5	3	2	1
	6	2	1	1
	7	1	1	-
$Se_ \times sese$	1	3	3	-
	2	8	2	6
	3	5	2	3
	4	2	1	1
	5	1	-	1
	6	1	-	1
	7	1	1	-
$sese \times sese$	2	3	-	3
	6	1	-	1

Considerando que, na população da qual procedem essas famílias a frequência de indivíduos não-secretorees pode ser estimada em 29%, teste a hipótese de que os fenótipos secretor e não-secretor decorrem de um par de alelos autossômicos e que o fenótipo secretor é dominante, empregando o método de Fisher (1939).

Para tanto, parta do pressuposto de que já foi demonstrado que:

a) As proporções de indivíduos secretorees e não-secretorees entre os cônjuges do sexo masculino não diferem significativamente das observadas entre os do sexo feminino.

b) O número de casais constituídos por marido secretor e mulher não-secretora e o de casais compostos por marido não-secretor e mulher secretora não diferem significativamente.

c) A proporção de indivíduos secretorees e não-secretorees entre os filhos dos diferentes tipos de casais é independente do sexo.

**R 2.** Pode-se aceitar a hipótese de que os fenótipos *secretor* e *não-secretor* são determinados por um par de alelos autossômicos sendo o fenótipo secretor dominante por causa dos resultados abaixo ( $q^2 = 0,29$ ;  $q = \sqrt{0,29} = 0,54$ ):

Casais	<i>i</i>	<i>n</i>	<b>P</b>	<b>Esp.</b>	<b>Obs.</b>
<i>Se</i> × <i>Se</i>	1	4	0,88	3,52	3
	2	9	0,79	7,11	7
	3	3	0,72	2,16	3
	4	3	0,67	2,01	3
	5	3	0,63	1,89	2
	6	2	0,60	1,20	1
	7	1	0,58	0,58	1
	Total	25		18,47	20
<i>Se</i> × <i>sese</i>	1	3	0,65	1,95	3
	2	8	0,48	3,84	2
	3	5	0,39	1,95	2
	4	2	0,34	0,68	1
	5	1	0,32	0,32	-
	6	1	0,31	0,31	-
	7	1	0,31	0,31	1
	Total	21		9,36	9
<i>sese</i> × <i>sese</i>	2	3			
	6	1			
	Total	4			

Casais	Irmandades	Obs.	Esp.	$\chi^2_{(1)}$
<i>Se</i> × <i>Se</i>	Apenas <i>Se</i>	20	18,47	0,485; 0,30 < P < 0,50
	Com <i>sese</i>	5	6,53	
<i>Se</i> × <i>sese</i>	Apenas <i>Se</i>	9	9,36	0,025; 0,80 < P < 0,90
	Com <i>sese</i>	12	11,64	

**Q 3.** Um pesquisador verificou que os seres humanos podiam ser classificados em um grupo positivo ou negativo conforme as suas hemácias aglutinassem ou não quando suspensas em um anti-soro por ele descoberto. As proporções de homens e de mulheres com respostas positiva (+) e negativa (-) não diferiram significativamente, permitindo estimar que 36% dos indivíduos dão resposta positiva e 64% resposta negativa. Os casais positivo  $\times$  positivo geraram filhos positivos e negativos, mas os casais negativo  $\times$  negativo geraram apenas filhos com resposta negativa.

O referido pesquisador estudou a frequência do caráter em questão, analisando 100 pares de mães e filhos(as), entre os quais anotou 25 pares ++, 9 pares +-, 10 pares -+ e 56 pares --. Verificar se é possível aceitar a hipótese de que os grupos positivo e negativo têm determinação monogênica autossômica, sendo o grupo negativo recessivo em relação ao grupo positivo.

**R 3.** Pode-se aceitar a hipótese de que o caráter em questão tem determinação monogênica autossômica e que o grupo negativo é recessivo, porque:

$$q^2 = 0,64; q = 0,8; p = 1 - 0,8 = 0,2.$$

Mães	Filhos +	Filhos -	Total
+	25 (23,2)	9 (12,8)	34
-	10 (12,8)	56 (51,2)	66
Total	35	65	100

$$\text{Mães: } \chi^2_{(1)} = 0,174; 0,50 < P < 0,70$$

$$\text{Filhos: } \chi^2_{(1)} = 0,043; 0,80 < P < 0,90$$

$$\text{Mães } \times \text{ Filhos: } \chi^2_{(1)} = 0,022; 0,80 < P < 0,90$$

$$\text{Pares: } \chi^2_{(3)} = 2,330; 0,50 < P < 0,70$$

**Q 4.** Um pesquisador verificou que os seres humanos podem ser classificados em *normais* e *deficientes* conforme apresentem ou não uma enzima eritrocitária ativa em certas condições experimentais. Verificou, também, que dentre 100 homens examinados, 12 foram classificados como deficientes, enquanto que, dentre 100 mulheres estudadas, apenas uma não apresentou atividade da enzima eritrocitária em apreço.

O estudo de 100 famílias quanto aos fenótipos normal e deficiente mostrou a distribuição apresentada na tabela abaixo. Com base nesses dados é possível aceitar a

hipótese de que esses fenótipos são determinados por um par de alelos do cromossomo X, e que o fenótipo normal é dominante nas mulheres?

Casais		Filhos			Filhas		
Marido×Mulher	No.	Normais	Deficientes	Total	Normais	Deficientes	Total
Normal×Normal	87	117	15	132	128	-	128
Normal×Deficiente	1	-	1	1	3	-	3
Deficiente×Normal	12	24	2	26	19	3	22

**R 4.** Pode-se aceitar a hipótese de herança monogênica ligada ao sexo e que o fenótipo normal é dominante nas mulheres porque:

$$M_a = 12 \quad F_{aa} = 1$$

$$M_A = 88 \quad F_{A_} = 99$$

$$q = \frac{-88 + \sqrt{7.744 + (1.200 \times 14)}}{600} = 0,11$$

$$p = 1 - 0,11 = 0,89$$

Marido×Mulher	Obs.	Esp.
Normal × Normal	87	87,9
Normal × Deficiente	1	1,1
Deficiente × Normal	12	10,9
Normal × Deficiente	-	0,1

$$\chi^2_{(3)} = 0,229; 0,95 < P < 0,98$$

Casais	Filhos			Filhas		
Marido × Mulher	Normais	Deficientes	Total	Normais	Deficientes	Total
Normal × Normal	117 (118,9)	15 (13,1)	132	128	-	128
	$\chi^2_{(1)} = 0,306; 0,50 < P < 0,70$					
Deficiente × Normal	24 (23,4)	2 (2,6)	26	19 (19,8)	3 (2,8)	22
	$\chi^2_{(1)} = 0,154; 0,50 < P < 0,70$			$\chi^2_{(1)} = 0,323; 0,50 < P < 0,70$		

## REFERÊNCIAS

Beiguelman, B. Hereditariedade da reação de Mitsuda. *Rev. Brasil. Leprol.* 30: 153-172, 1962.

Beiguelman, B. The genetics of the resistance to leprosy. *Internat. J. Lepr.* 33: 808-812, 1965.

Beiguelman, B. & Quagliato, R. Nature and familial character of the lepromin reactions. *Internat.J. Lepr.* 33: 800-807, 1965.

- Beiguelman B. A reação de Mitsuda oitenta anos depois. *Hansen. Internat.* 24: 144-16, 1999.
- Beiguelman, B. Genética e hanseníase. *Ciência & Saúde Coletiva* 7: 11-128, 2002.
- Feitosa, M., Krieger, H., Borecki, I., Beiguelman, B. & Rao, D.C. Genetic epidemiology of the Mitsuda reaction in leprosy. *Hum. Hered.* 46: 32-35, 1996.
- Fisher, R.A. em Taylor, G.L. & Prior, A.M. Blood groups in England. III-Discussion of the family material. *Ann. Eugen.* 9: 18-44, 1939.
- Henningsen, K. Étude d'ensemble du facteur sanguin P. *Rev. Hémat.* 5: 276-284, 1950.
- Lalouel, J.M., Rao, D.C., Morton, N.E. & Elston, R.C. A unified model for complex segregation analysis. *Am. J. Hum. Genet.* 35: 816-826, 1983.
- Li, C.C. - *Population genetics*. Univ. Chicago Press, Chicago, 7a. reimpressão, 1972.
- Mann, J.D., Cahan, A., Gelb, A.G., Fisher, N., Hamper, J., Tippett, P., Sanger, R. & Race, R.R. A sex-linked blood group. *Lancet* 1: 8-10, 1962.
- Morton, N.E. Forces maintaining polymorphisms. *Acta Anthropogenetica* 1: 3-14, 1976/1977.
- Schiff, F. & Sasaki, H. Der Ausscheidungstypus, ein auf serologischen Wege nachweisbares mendelndes Merkmal. *Klin. Wschr* 11: 1426-1429, 1932.
- Snyder, L.H. Studies in human inheritance. IX. The inheritance of taste deficiency in man. *Ohio J. Sci.* 32: 436-440, 1932.

## CAPÍTULO 5. O EFEITO DA CONSANGÜINIDADE

Quando se estuda o padrão de herança das anomalias recessivas enfatiza-se que, entre os indivíduos afetados por elas, a porcentagem dos que são filhos de casais consangüíneos pode atingir valores muito altos, porque tais casais têm maior probabilidade de gerar filhos homocigotos do que os não-consangüíneos. Mas, se os casamentos consangüíneos aumentam a probabilidade de homocigose em seus filhos, está claro, também, que a elevação da taxa desses casamentos poderá perturbar o equilíbrio de Hardy e Weinberg, ao alterar as frequências genóticas sem modificar as frequências alélicas desse sistema. No presente capítulo teremos a oportunidade de analisar o modo pelo qual essa desestabilização do equilíbrio de Hardy e Weinberg ocorre e de abordar outras questões relacionadas ao efeito da consangüinidade. No final ainda abordaremos o fenômeno de aumento da homocigose sem necessidade de consangüinidade, conhecido como *efeito Wahlund*.

Os fundamentos para o entendimento do efeito dos casamentos consangüíneos sobre a composição genotípica das populações foram lançados por um número muito restrito de geneticistas na primeira metade do século vinte. Dentre os trabalhos por eles publicados é obrigatório destacar os de Wright (1921,1922,1951), Dahlberg (1929,1938,1948), Haldane e Moshinsky (1939) e Malécot (1948), sem os quais o presente capítulo não poderia ter sido escrito. No Brasil, os estudos sobre consangüinidade foram realizados, predominantemente, por Newton Freire-Maia, que conduziu, sozinho ou com a colaboração de geneticistas oriundos de sua escola, numerosas pesquisas pioneiras nessa área, as quais foram revistas, em boa parte, em uma de suas obras (1974).

### AS RESTRIÇÕES AOS CASAMENTOS ENTRE PARENTES

As restrições aos casamentos entre parentes existem não apenas nas sociedades civilizadas, mas também nas primitivas. Nem sempre, porém, os critérios para o impedimento dessas uniões têm ou tiveram por objetivo evitar o maior risco de manifestação de anomalias recessivas. De fato, o Código Civil brasileiro impede o casamento de ascendentes com descendentes e de consangüíneos colaterais até o terceiro grau, o que estende a restrição a irmãos, meios-irmãos e a tios(as) e sobrinhas(os). Contudo,

ele impede, também, o casamento de parentes afins em linha reta, o casamento do(a) adotante com o cônjuge do(a) adotado(a) e do(a) adotado(a) com o cônjuge do(a) adotante, bem como o casamento do(a) adotado(a) com o filho(a) superveniente ao pai ou à mãe adotiva. Evidentemente, ao impedir esses casamentos entre parentes afins o Legislador tinha em mente apenas a organização e a proteção da família.

Nem sempre, porém, os critérios adotados para o impedimento de casamentos entre parentes têm uma base lógica. Assim, por exemplo, durante cerca de um milênio a Igreja Católica proibiu o casamento entre parentes até o sétimo grau, com base no preceito bíblico de que o mundo foi criado em sete dias. Mais tarde, ela passou a proibir os casamentos até o quarto grau tomando como base os quatro elementos, terra água, ar e fogo (Almström, 1958).

Nos países atingidos pelo movimento luterano, as leis canônicas foram abolidas e a proibição dos casamentos entre parentes consangüíneos passou a ser feita de acordo com o Levítico (18:6-18). Dessa maneira, tais proibições atingiram alguns consangüíneos (pai × filha, mãe × filho, irmão × irmã, meio-irmão × meia-irmã, avô × neta, tio × sobrinha, tia × sobrinho) e alguns não-consangüíneos com afinidade próxima (filho × madrasta, sobrinho × esposa do tio, sogro × nora, cunhado × cunhada). Visto que a Bíblia, surpreendentemente, não faz restrição aos casamentos entre primos, verificou-se, nos países que aceitaram a reforma da Igreja, que a freqüência de casamentos entre primos em primeiro grau aumentou rapidamente a partir do século dezesseis.

Apesar de existirem restrições aos casamentos consangüíneos, mesmo em sociedades primitivas, existe um fenômeno que nos leva a ter que admitir a sua existência, em alta proporção, em épocas remotas da história da humanidade. É o chamado fenômeno da *perda dos ancestrais*. De fato, se não existissem casamentos consangüíneos todo indivíduo deveria ter dois genitores, quatro avós, oito bisavós, 16 trisavós, 32 tetravós, e, de uma maneira geral  $2^n$  ancestrais, sendo  $n$  o número de gerações que antecedem o indivíduo. Se aceitarmos que em cada século houve, em média, quatro gerações, tem-se que admitir a existência de 40 gerações em um milênio, e, também, que, há mil anos, o número de ancestrais de um indivíduo da época atual deveria ser igual a  $2^{40}$ , isto é, 1.099.511.627.776.

Essa estimativa, entretanto, é inverossímil, em vista das informações que temos sobre a pequena população humana terrestre há um milênio. Portanto, a hipótese da

existência de casamentos consanguíneos em alta proporção em eras progressas é aquela que melhor serve para explicar o fenômeno da perda de ancestrais. Assim, por exemplo, enquanto os filhos de não-consanguíneos têm oito bisavós, os filhos de primos em primeiro grau têm seis e os filhos de primos duplos em primeiro grau têm quatro bisavós. O fenômeno da perda dos ancestrais nos leva, pois, a concluir com Stern (1960) que, do ponto de vista evolutivo, ***a fraternidade da espécie humana não é um mero conceito espiritual, mas uma realidade genética.***

No Levítico existem indicações de que entre os povos anteriores e contemporâneos dos antigos hebreus os casamentos incestuosos poderiam não ter sido circunscritos apenas às famílias reais, como as de algumas dinastias do Egito (Figura 1.5). É o que se depreende das restrições que a Lei Mosaica faz no Levítico, 18:26-27, aos casamentos consanguíneos e a outras uniões, com a advertência: "Guardai as minhas leis e os meus mandamentos, e não cometais nenhuma dessas abominações, tanto os naturais quanto os estrangeiros entre vós. Porque todas essas execrações cometeram os que foram, antes de vós, habitantes desta terra, e a contaminaram". Curiosamente, o casamento entre irmãos também foi a regra entre os reis incas (Rothhammer e Cruz-Coke, 1977). Assim, nas onze gerações que antecederam à do rei Huáscar, que Pizarro mandou assassinar em 1532, nove foram de casamentos entre irmãos, uma de casamento de primos em primeiro grau e uma de tia com sobrinho (Figura 2.5).

Quando as populações são muito pequenas, os casamentos consanguíneos podem ocorrer não porque sejam preferenciais, mas porque o tamanho reduzido da população faz com que aumente a probabilidade de parentesco consanguíneo próximo entre os cônjuges. Isso tem sido observado em aldeias localizadas em ilhas, em tribos beduínas e em pequenas comunidades religiosas. Existem comunidades pequenas, entretanto, como as encontradas nos Alpes suíços, em que a frequência de casamentos consanguíneos é bem menor do que a esperada ao acaso, o que indica que, nesses agrupamentos humanos, tais uniões são evitadas.

No cômputo geral, porém, predominam as evidências de que os casamentos consanguíneos não ocorrem aleatoriamente na maioria das populações humanas (Haldane e Moshinsky, 1939; Hajnal, 1963). No Brasil, por exemplo, constatou-se que, entre os casais de primos em primeiro grau, aqueles em que o pai do marido é irmão do pai da mulher

ocorrem mais freqüentemente do que aqueles em que a mãe do marido é irmã da mãe da mulher, o que fala a favor da influência direta ou indireta dos genitores masculinos no casamento de seus filhos (Freire-Maia, 1958; Freire-Maia e Freire-Maia, 1961).

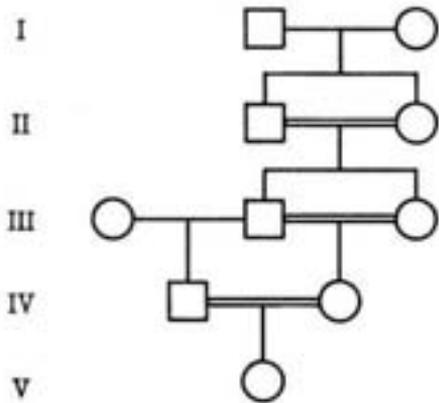


Fig.1.5. Casamentos entre irmãos e meio-irmãos na 18ª. dinastia do Egito (1580-1350 A.C.).

- I-1×I-2 : Sequenerra III × Aahotep
- II-1×II-2 : Aahmes × Ahmes Nefértari
- III-1×III-2: Senseneb × Amenhotep
- III-2×III-3: Amenhotep × Aahotep II
- IV-1×IV-2: Thotmes × Aahmes
- V-1 : Hatsheput

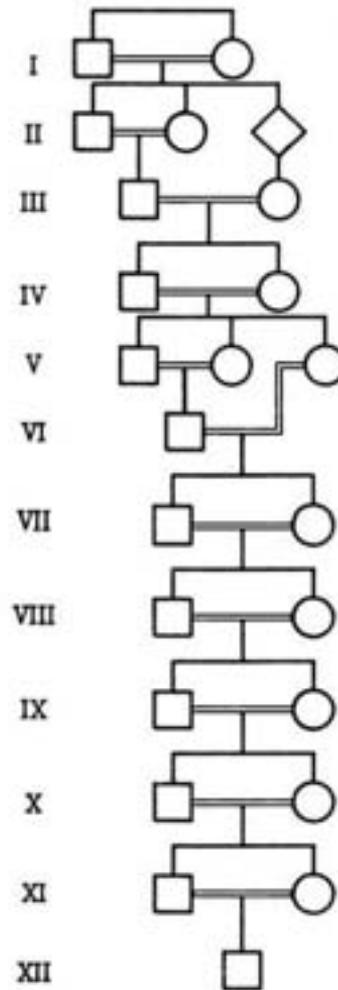


Fig.2.5. Casamentos consangüíneos entre os ancestrais do rei inca Huáscar.

- I-1 × I-2 : Manco Capac × Coya Oello
- II-1 × II-2 : Sinchi Roca × Chimpo
- III-1 × III-2 : Lloque Yupanqui × Anac Varqui
- IV-1 × IV-2 : Mayta Capac × Chimpo Urma
- V-1 × V-2 : Capac Yupanqui × Chimpo Oollo
- VI-1 × V-3 : Inca Roca × Cusi Chimpo
- VII-1 × VII-2 : Yahuar Huacac × Ipavaco
- VIII-1× VIII-2: Viracocha Inca×Yunto Cayan
- IX-1 × IX-2 : Pachacutec Inca Yupanqui × Anahuarque
- X-1 × X-2 : Tupac Inca Yupanqui × Oello Coya
- XI-1 × XI-2 : Huayna Capac × Ravaallo
- XII-1 : Huáscar

É difícil analisar os fatores que determinam a prevalência de casamentos consanguíneos nas populações humanas, mas os de ordem econômica devem influir de modo prioritário, mormente em comunidades rurais, com a finalidade de manter a integridade da propriedade na família sem herdeiros do sexo masculino. Essa solução, aliás, não deixa de ser bíblica, como se pode ler em Números, 36: 10-12: "E as filhas de Salfaad fizeram como lhes tinha sido mandado e Maala, e Tersa, e Helga, e Melca, e Noa, casaram com os filhos de seu tio paterno, da família de Manassés, que foi filho de José; e a possessão que lhes tinha sido adjudicada permaneceu na tribo e família de seu pai".

Na maior parte dos países, entretanto, a tendência atual é a diminuição acelerada das taxas de casamentos consanguíneos, em decorrência do processo de urbanização, com a formação de grandes aglomerados humanos, das conquistas femininas, que tornam a mulher cada vez mais independente, e das facilidades de locomoção e comunicação, que propiciam maior contato entre pessoas de origem diversa. Assim, por exemplo, a Arquidiocese de Florianópolis, que registrava 4,5% de casamentos consanguíneos em 1920, passou a assinalar 1 % desses matrimônios em 1980 (Agostini e Meireles-Nasser, 1986).

### COEFICIENTE DE CONSANGÜINIDADE OU COEFICIENTE DE PARENTESCO

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, independentemente de tais genes condicionarem fenótipos dominantes ou recessivos, normais ou anômalos. Essa medida do parentesco, ou da correlação genética entre dois indivíduos, é denominada *coeficiente de consangüinidade* ou *coeficiente de parentesco* e, geralmente, indicada pela letra  $r$ , inicial da palavra inglesa *relationship* = parentesco.

Para entendermos facilmente como se chega ao valor de  $r$ , consideremos dois primos em primeiro grau, como os indivíduos III-1 e III-2 do heredograma da Fig. 3.5, e suponhamos que, em relação a um par de alelos autossômicos  $A,a$ , o indivíduo III-1 é heterozigoto ( $Aa$ ). Para calcular a probabilidade de seu primo (ou prima) III-2 também possuir o alelo  $a$ , por tê-lo herdado de um ancestral comum a ambos, temos que levar em conta que tal ancestral poderia ter sido o avô I-1 ou a avó I-2, já que os primos III-1 e III-2 têm esses dois ancestrais em comum e um desses ancestrais poderia ser heterozigoto  $Aa$ .

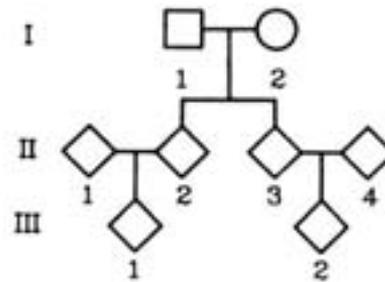


Fig. 3.5. Heredograma de uma genealogia com primos em primeiro grau (III-1 e III-2).

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

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

Considerando que cada passagem de um gene de uma geração a outra é chamado de *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  $\frac{1}{2}$ , 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-1 aos seus netos III-1 e III-2 é igual a  $\left(\frac{1}{2}\right)^4 = \frac{1}{16}$ , pois são quatro os passos existentes nos dois caminhos que esse gene teria que percorrer. Pelo mesmo motivo, tem-se que a probabilidade de o gene  $a$  ter sido transmitido pela avó I-2 aos seus netos III-1 e III-2 também é igual a  $\frac{1}{16}$ . Pode-se, pois, 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  $\frac{1}{16} + \frac{1}{16} = \frac{1}{8}$ . 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 consangüinidade de primos em primeiro grau é  $\frac{1}{8}$ , ou seja,  $r = \frac{1}{8}$ .

Quando existem mais de dois ancestrais em comum, todos eles devem ser levados em conta no cálculo do coeficiente de consangüinidade. Assim, por exemplo, no caso de primos duplos em primeiro grau devemos considerar que tais primos têm quatro ancestrais em comum (I-1, I-2, I-3 e I-4 da Fig. 4.5) e que existem quatro passos que devem ser levados em conta em relação a cada um deles. Desse modo,  $r = \frac{1}{4}$  para primos duplos em primeiro grau, porque  $r = \left(\frac{1}{2}\right)^4 + \left(\frac{1}{2}\right)^4 + \left(\frac{1}{2}\right)^4 + \left(\frac{1}{2}\right)^4 = \frac{4}{16} = \frac{1}{4}$ .

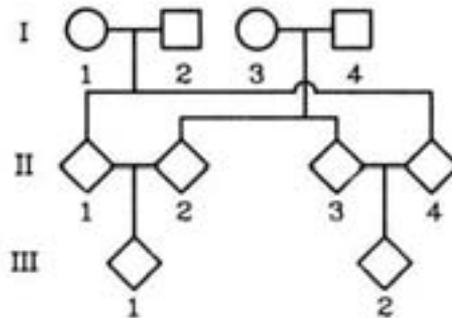


Fig. 4.5. Heredograma de uma genealogia com primos duplos em primeiro grau (III-1 e III-2)

A indicação do cálculo do coeficiente de consangüinidade pode ser resumida pela fórmula  $r = \Sigma \left(\frac{1}{2}\right)^N$ , na qual  $N$  é o número de passos que unem 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  $r = \left(\frac{1}{2}\right)^N$ . Assim, por exemplo, no caso de dois meio-irmãos, a probabilidade de ambos serem portadores de um mesmo gene autossômico por origem comum é  $\left(\frac{1}{2}\right)^2 = \frac{1}{4}$ , 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 coeficiente de consangüinidade 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.

O conhecimento do coeficiente de consangüinidade de parentes colaterais permite constatar que o parentesco genético entre esses consangüíneos está freqüentemente em desacordo com o grau de parentesco estabelecido no Direito Civil brasileiro e no Direito Canônico, porque, em Genética, a contagem dos graus de consangüinidade entre parentes colaterais é diferente daquela utilizada em Direito. Aliás, até 27 de novembro de 1983 as contagens dos graus de consangüinidade colateral em Direito Civil e em Direito Canônico também eram discordantes entre si. A identidade dos dois Códigos a esse respeito somente foi alcançada com a entrada em vigor de um novo texto do Código Canônico (Freire-Maia, 1986). A partir daí, tanto em Direito Civil quanto em Direito Canônico, o grau de consangüinidade entre dois colaterais é a soma do número de gerações contadas nas linhas colaterais ascendente e descendente, tomando como referência o *tronco* da genealogia, isto é, os ancestrais comuns mais próximos dos parentes consangüíneos.

Tem-se, assim, que, em Direito, tanto os irmãos quanto os meio-irmãos são consangüíneos em segundo grau. Tios e sobrinhos são parentes em terceiro grau. Aqueles que, em Genética, são primos em primeiro grau, em Direito são denominados parentes em quarto grau, enquanto que os primos que, em Genética, são ditos de segundo grau, são, em Direito, considerados parentes em quinto grau. Os primos em terceiro grau são considerados parentes em sexto grau à luz do Direito, e assim por diante. Já em relação aos consangüíneos em linha reta sempre houve concordância entre os Códigos Civil e Canônico em considerar os graus de consangüinidade entre dois indivíduos como o número de gerações que os separam. Desse modo, a consangüinidade entre pais e filhos é dita em primeiro grau, entre avós e netos é dita em segundo grau, entre bisavós e bisnetos é dita em terceiro grau e assim por diante.

Essa discordância entre a Genética e o Direito é que faz com que, por exemplo, os meio-irmãos, os primos duplos em primeiro grau e os tios e sobrinhos, com o mesmo parentesco genético, pois possuem, em média, a mesma quantidade de genes idênticos por origem comum ( $r = \frac{1}{4}$ ) sejam classificados diferentemente em Direito. Assim, no Direito Civil e no Direito Canônico, os meio-irmãos são parentes colaterais em segundo grau, os primos duplos em primeiro grau são parentes em quarto grau e os tios e sobrinhos são parentes colaterais em terceiro grau. Por outro lado, de acordo com o Direito Civil brasileiro e com o Direito Canônico não existe diferença quanto ao grau de parentesco entre irmãos ( $r = \frac{1}{2}$ ) e os meio-irmãos ( $r = \frac{1}{4}$ ), que são tratados como parentes em segundo grau. Os primos em primeiro grau ( $r = \frac{1}{8}$ ) e os primos duplos em primeiro grau ( $r = \frac{1}{4}$ ), são, também, tratados do mesmo modo, já que, em ambos os casos, são considerados como parentes em quarto grau.

### **COEFICIENTE DE CONSANGÜINIDADE E RISCO DE ANOMALIAS RECESSIVAS**

Uma das maneiras de demonstrar numericamente que os casais consangüíneos estão, de fato, sujeitos a um risco maior do que os outros de gerar uma criança com uma anomalia recessiva baseia-se no conhecimento do coeficiente de consangüinidade. Assim, por exemplo, consideremos um indivíduo heterozigoto de um gene autossômico  $a$ , isto é,  $Aa$ , que, em homozigose ( $aa$ ), determina uma anomalia rara, com incidência estimada em 1 por 40.000. Suponhamos, ainda, que esse heterozigoto deseja casar com uma mulher que não é sua parente consangüínea e quer saber qual a probabilidade de gerar com ela uma criança com a anomalia recessiva determinada pelo genótipo  $aa$ .

Se a taxa de casamentos consangüíneos na população da qual procede o heterozigoto for desprezível, diremos que tal probabilidade é estimada em  $1/400$ , pois a frequência estimada do gene  $a$  é  $q = \sqrt{\frac{1}{40.000}} = \frac{1}{200}$ , de sorte que a frequência estimada de heterozigotos na população é  $2pq = \frac{1}{100}$ . Sabendo-se que um casal  $Aa \times Aa$  tem probabilidade igual a  $\frac{1}{4}$  de gerar uma criança homozigota  $aa$ , levamos, portanto, em conta

a probabilidade de a mulher ser heterozigota como o seu futuro marido que, com certeza (probabilidade 1) é  $Aa$ , bem como a probabilidade de um casal  $Aa \times Aa$  gerar uma criança  $aa$  ( $\frac{1}{4}$ ). Calculamos, por isso,  $1 \cdot \frac{1}{4} \cdot \frac{1}{100} = \frac{1}{400}$  ou 0,0025.

Se, entretanto, o heterozigoto  $Aa$  desejasse casar com uma de suas primas em primeiro grau, poderíamos admitir que o gene  $a$ , por ser raro, não deve ter entrado na genealogia desse heterozigoto mais de uma vez. Isso equivale a dizer que poderíamos admitir que a ocorrência do gene  $a$  em consangüíneos desse heterozigoto  $Aa$  teria sempre origem em um ancestral comum a esses parentes. Nesse caso, a informação que daríamos ao heterozigoto consulente é a de que a probabilidade de ele gerar uma criança com o genótipo  $aa$  é estimada em  $\frac{1}{32}$  ou 0,03125, isto é, 12,5 vezes mais alta do que a que ele estaria sujeito se casasse com uma mulher não-consangüínea dele, pois  $\frac{0,03125}{0,0025} = 12,5$ .

Realmente, visto que a prima em primeiro grau do heterozigoto  $Aa$  tem probabilidade igual a  $\frac{1}{8}$  de ser, também, heterozigota  $Aa$  por ter recebido o gene  $a$  de um ancestral comum a ambos ( $r = \frac{1}{8}$ ), pode-se calcular a probabilidade de um casal de primos em primeiro grau gerar uma criança homozigota  $aa$ , dado que um dos cônjuges é, com certeza  $Aa$ , resolvendo  $1 \cdot \frac{1}{4} \cdot \frac{1}{8} = \frac{1}{32}$ . Ao calcularmos, desse modo, a probabilidade de um filho do casal consangüíneo ser homozigoto, dado que um dos cônjuges é, sabidamente, heterozigoto, estamos, pois, na realidade, admitindo a possibilidade de ele ser o que se pode chamar de **autozigoto** ou **isozigoto**, isto é, um homozigoto que possui dois alelos idênticos, derivados de um único gene herdado de um de seus ancestrais.

Se a anomalia recessiva  $aa$  tivesse freqüência bem menor na população, digamos a metade (1 : 80.000), a probabilidade de o heterozigoto  $Aa$  gerar com sua prima em primeiro grau um filho com o genótipo  $aa$  seria, proporcionalmente, ainda maior (17,6 vezes), quando comparada àquela que ele teria em se casando com uma mulher que não lhe é consangüínea. Isso porque essa última teria probabilidade praticamente igual a  $\frac{1}{141}$  de ser

heterozigota, de modo que a probabilidade de ela gerar uma criança *aa* com o heterozigoto *Aa* seria igual a  $\frac{1}{4} \cdot \frac{1}{141} = \frac{1}{564} = 0,00177$ , o que fornece  $\frac{0,03125}{0,00177} = 17,6$ .

Esse último exemplo serve para ressaltar que a razão entre a probabilidade de encontrar um gene em homozigose em filhos de consangüíneos e a de encontrá-lo desse modo em filhos de não-consangüíneos aumenta com a raridade do gene. Evidentemente, essa razão também aumenta com o grau de consangüinidade dos casais, visto que aqueles com grau de consangüinidade mais próximo têm coeficiente de consangüinidade mais alto. Mais adiante, no tópico sobre "O equilíbrio de Wright", neste mesmo capítulo, teremos a oportunidade de chegar a essas mesmas conclusões a partir de um modelo mais geral.

A essa altura, o leitor atento deve ter percebido que um heterozigoto de um determinado gene autossômico, casado com uma de suas primas em primeiro grau, tem probabilidade muito maior de **não gerar** um filho homozigoto por autozigose de tal gene do que de gerá-lo. Realmente, já vimos que tal indivíduo terá probabilidade de 3% de gerar uma criança autozigota desse gene, pois  $\frac{1}{32} \cong 0,03$ , o que equivale a dizer que a probabilidade de isso não acontecer será de 97%, porque  $1 - 0,03 = 0,97$ . Entretanto, quando lembramos que as pessoas são heterozigotas em relação a um grande conjunto de genes, tem-se que a probabilidade de um indivíduo, que é casado com uma de suas primas em primeiro grau, não gerar um autozigoto de pelo menos um dos genes do conjunto diminui à medida que o número dos genes desse conjunto aumenta. Assim, se um indivíduo é heterozigoto de  $n$  genes autossômicos e é casado com uma prima em primeiro grau, a probabilidade de nenhum de tais genes estar em autozigose em um dos filhos desse casal é  $\left(\frac{31}{32}\right)^n$  e a de *pelo menos um desses genes* ser encontrado nesse estado é  $1 - \left(\frac{31}{32}\right)^n$ , pois a distribuição dos genótipos dos filhos desse casal em relação a esses genes será dada pelo binômio  $\left(\frac{31}{32} + \frac{1}{32}\right)^n$ .

Para exemplificar, consideremos um indivíduo que é heterozigoto de cinco genes autossômicos ( $n = 5$ ), cada qual com efeito deletério quando em homozigose. Se esse indivíduo casar com uma de suas primas em primeiro grau, ele passará a ter alta probabilidade de gerar um filho afetado por *pelo menos uma* das cinco anomalias recessivas

determinadas por esses genes, ainda que tais genes sejam muito raros na população, pois a probabilidade de que pelo menos um deles ocorra em autozigose será 14,7%, pois

$$1 - \left(\frac{31}{32}\right)^5 = 0,147.$$

Antes de encerrar este tópico é importante lembrar que, ao analisarmos casais consanguíneos em relação a um determinado gene, devemos estar atentos para a possibilidade de existir uma indicação de ele ter entrado mais de uma vez na genealogia. Se isso acontecer, a estimativa da probabilidade de ocorrência desse gene no casal não poderá ser o coeficiente de consangüinidade dos cônjuges, como se poderá constatar no exemplo dado na Figura 5.5.

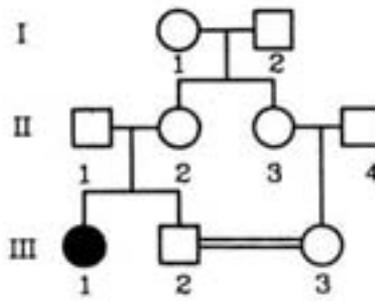


Fig. 5.5. Heredograma de uma genealogia na qual houve ocorrência de uma anomalia recessiva autossômica (III-1).

No heredograma da Figura 5.5, que descreve uma genealogia na qual ocorreu uma anomalia recessiva autossômica monogênica manifestada pela mulher III-1 ( $aa$ ), fica evidente que o gene  $a$  entrou duas vezes em tal genealogia, pois a homozigose de III-1 não teria sido possível se o gene  $a$  não estivesse presente em heterozigose tanto em II-1 quanto em II-2, que não têm parentesco consanguíneo entre si. A mulher III-1 é, pois, uma homozigota, mas não uma autozigota. Em vista do exposto, para calcularmos a probabilidade de III-3 ter o gene  $a$  em heterozigose, não poderemos tomar o coeficiente de consangüinidade do casal III-2  $\times$  III-3, mesmo porque não sabemos se III-2 é ou não heterozigoto  $Aa$ . O que devemos fazer, então, é usar o coeficiente de consangüinidade de II-2 e III-3, isto é da tia, reconhecidamente heterozigota  $Aa$ , e de sua sobrinha, o qual é, como sabemos,  $\frac{1}{4}$ . Por outro lado, visto que não sabemos se o cônjuge III-2 é heterozigoto, temos que calcular a probabilidade de ele ser  $Aa$ , dado que ele é filho normal de um casal

reconhecidamente heterozigoto, pois gerou uma filha  $aa$ . Tal probabilidade é, como sabemos,  $\frac{2}{3}$  (probabilidade condicional). Portanto, se o casal III-2  $\times$  III-3 quisesse saber qual a probabilidade de gerar uma criança homozigota  $aa$ , a qual pode ser autozigota ou não, responderíamos  $\frac{1}{24}$ , pois calcularíamos  $\frac{1}{4} \cdot \frac{2}{3} \cdot \frac{1}{4} = \frac{1}{24} \cong 0,04$  ou 4%.

### COEFICIENTE DE ENDOCRUZAMENTO

O *coeficiente de endocruzamento*, geralmente simbolizado pela letra  $F$ , indica a probabilidade de um indivíduo ser homozigoto, em decorrência do encontro, nele, de um par de alelos idênticos, derivados de um único gene herdado de um de seus ancestrais. Em outras palavras, o coeficiente de endocruzamento indica a probabilidade de um indivíduo ser autozigoto de um gene qualquer presente em um de seus ancestrais. 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 o indivíduo IV.1 da Figura 6.5. Consideremos um par de alelos autossômicos  $B,b$  e suponhamos que o indivíduo IV.1 é homozigoto recessivo  $bb$ .

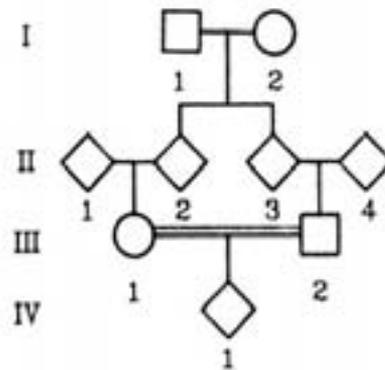


Fig.6.5. Heredograma de uma genealogia de um(a) filho(a) de primos em primeiro grau.

Se quisermos saber qual a probabilidade de o filho de primos em primeiro grau (IV-1) ser homozigoto  $bb$  por autozigose, teremos que calcular a probabilidade de um alelo  $b$ , presente em seu bisavô (I-1) ou em sua bisavó (I-2), que são os ancestrais comuns do casal de primos (III-1  $\times$  III-2), genitores de IV-1, ter sido transmitido de modo tal a se encontrar em autozigose em IV-1. Se designarmos o autossomo qualquer que contém o loco do par de alelos  $B,b$  por  $A$ , teremos que considerar a possibilidade de o alelo  $b$  estar em um

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

Se o alelo  $b$  estivesse localizado no cromossomo  $A^1$ , a probabilidade de o filho de primos em primeiro grau IV-1 da Figura 6.5 ser autozigoto  $bb$  seria igual a  $\frac{1}{64}$ , porque seria necessário que o cromossomo  $A^1$  fosse transmitido percorrendo os seguintes dois caminhos:

Caminho 1 : I-1  $\rightarrow$  II-2  $\rightarrow$  III-1  $\rightarrow$  IV-1

Caminho 2: I-1  $\rightarrow$  II-3  $\rightarrow$  III-2  $\rightarrow$  IV-1

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

Repetindo o mesmo raciocínio para os filhos de casais com outros graus de consangüinidade, inclusive os resultantes de ligações incestuosas, chegar-se-á à conclusão de que o coeficiente de endocruzamento de filhos de:

a) pais com filhas, ou de mães com filhos, ou de irmãos é igual a  $\frac{1}{4}$  ;

b) meio-irmãos, ou de primos duplos em primeiro grau, ou de tios(as) e sobrinhos(as) é igual a  $\frac{1}{8}$  ;

- c) tios(as) e meia(o)-sobrinhas(os) é igual a  $\frac{1}{16}$ , como no caso de filhos de primos em primeiro grau;
- d) primos em segundo grau é igual a  $\frac{1}{32}$ ;
- e) primos em terceiro grau é igual a  $\frac{1}{64}$ .

Comparando os valores dos coeficientes de endocruzamento com os coeficientes de consangüinidade é fácil concluir 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  $F = \frac{r}{2}$ . O coeficiente de endocruzamento também pode ser calculado com base na contagem do número de passos nas linhas de consangüinidade que unem um indivíduo a cada um dos ancestrais comuns a seus genitores. Assim, designando o número de tais passos por  $N$ , resolve-se a fórmula  $F = \Sigma \left(\frac{1}{2}\right)^{N-1}$ . No caso de filhos de primos em primeiro grau, por exemplo, tem-se  $F = \left(\frac{1}{2}\right)^5 + \left(\frac{1}{2}\right)^5 = \frac{1}{16}$ , pois há dois ancestrais comuns a seus genitores e seis passos nas linhas que os unem a esses ancestrais. No caso de filhos de primos duplos em primeiro grau,  $F = \left(\frac{1}{2}\right)^5 + \left(\frac{1}{2}\right)^5 + \left(\frac{1}{2}\right)^5 + \left(\frac{1}{2}\right)^5 = \frac{1}{8}$  porque são quatro os ancestrais comuns a seus genitores e existem seis passos em cada uma das quatro linhas que os unem a eles.

Chamando de  $a$  o número de ascendentes que existem nas linhas que unem um filho de consangüíneos a cada um dos ancestrais comuns a seus pais, o valor de  $F$  também poderá ser calculado por intermédio de  $F = \Sigma \left(\frac{1}{2}\right)^a$  pois  $a = N - 1$ , quando  $N$  é o número de passos nas linhas de consangüinidade que unem um indivíduo a um ancestral comum a seus genitores.

Quando um ancestral de um filho de consangüíneos também é filho de consangüíneos, o coeficiente de endocruzamento desse último ( $F_a$ ) também deve ser levado em conta na fórmula geral. Assim, em relação a tal ancestral deve-se calcular  $(1 + F_a) \left(\frac{1}{2}\right)^{N-1}$  ou  $(1 + F_a) \left(\frac{1}{2}\right)^a$ .

Na Figura 7.5 foram reunidos heredogramas de genealogias com alguns tipos de casamentos consangüíneos e assinalados os coeficientes de consangüinidade desses casais, bem como os coeficientes de endocruzamento de seus filhos em relação a genes autossômicos.

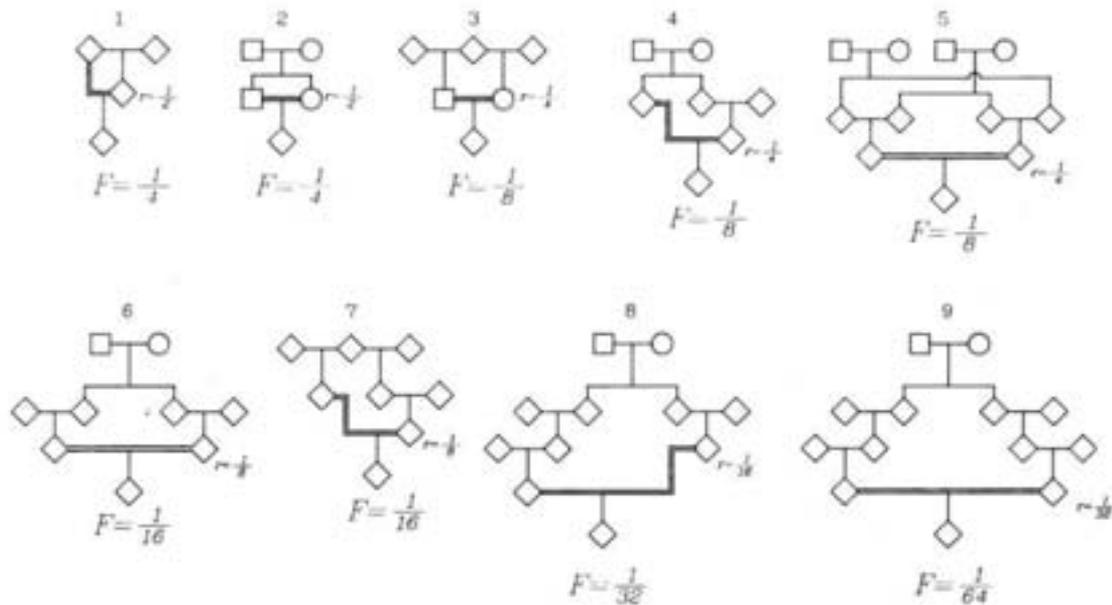


Fig. 7.5. Heredogramas de genealogias em que há casamentos consangüíneos, inclusive incestuosos ( $r$  = coeficiente de consangüinidade;  $F$  = coeficiente de endocruzamento).  
 1- pai(mãe) × filha(o); 2- irmão × irmã; 3- meio-irmão × meia-irmã; 4- tio(a) × sobrinha(o); 5- primos duplos em primeiro grau; 6- primos em primeiro grau; 7- tio(a) × meia(o) sobrinha(o); 8- primos em segundo grau; 9- primo × prima em terceiro grau

### COEFICIENTE MÉDIO DE ENDOCRUZAMENTO

Se conhecermos as freqüências com que os casais com diferentes graus de consangüinidade ocorrem em uma população, poderemos obter uma estimativa do seu *coeficiente médio de endocruzamento*, o qual pode ser simbolizado por  $\bar{F}$ . Numa população predominante católica, como a brasileira, essas freqüências são obtidas de modo relativamente fácil nas paróquias, porque os casamentos consangüíneos requerem autorização do bispo.

Se aceitarmos que, em média, todos os tipos de casais têm a mesma fertilidade, poderemos calcular o coeficiente médio de endocruzamento de uma população a partir da média ponderada obtida com a soma dos produtos dos números dos casais com diferentes

graus de consangüinidade pelos respectivos coeficientes de endocruzamento de seus futuros filhos ou, mais simplesmente, a partir da soma dos produtos das freqüências dos casais com diferentes graus de consangüinidade pelos respectivos coeficientes de endocruzamento de seus futuros filhos.

Para exemplificar, consideremos uma cidade brasileira qualquer (A) na qual, num determinado período, tivessem ocorrido 1.000 casamentos, dos quais 853 (85,3%) entre não-consangüíneos ( $F = 0$ ), 50 (5%) entre primos em terceiro grau ( $F = \frac{1}{64}$ ), 35 (3,5%) entre primos em segundo grau ( $F = \frac{1}{32}$ ), 60 (6%) entre primos em primeiro grau ( $F = \frac{1}{16}$ ) e 2 (0,2%) entre tios(as) e sobrinhas(os) ( $F = \frac{1}{8}$ ). Com base nesses dados estimariamos o coeficiente médio de endocruzamento da população dessa cidade em  $\bar{F} = 0,0059$ , porque calcularíamos  $\bar{F}$  a partir de qualquer das expressões abaixo:

$$\bar{F} = \frac{\frac{50}{64} + \frac{35}{32} + \frac{60}{16} + \frac{2}{8}}{1.000} = 0,0059 \quad \text{ou} \quad \frac{0,05}{64} + \frac{0,035}{32} + \frac{0,06}{16} + \frac{0,002}{8} = 0,0059.$$

Se na população de outra cidade brasileira, que designaremos por B, tivessem ocorrido, no mesmo período, 800 casamentos, dos quais 624 (78%) entre não-consangüíneos ( $F = 0$ ), 72 (9%) entre primos em terceiro grau ( $F = \frac{1}{64}$ ), 32 (4%) entre primos em segundo grau ( $F = \frac{1}{32}$ ), 64 (8%) entre primos em primeiro grau ( $F = \frac{1}{16}$ ), 6 (0,75%) entre primos duplos em primeiro grau ( $F = \frac{1}{8}$ ) e 2 (0,25%) entre tio(as) e sobrinhas(os) ( $F = \frac{1}{8}$ ), estimariamos o seu coeficiente médio de endocruzamento em  $F = 0,0089$  porque:

$$\bar{F} = \frac{0,09}{64} + \frac{0,04}{32} + \frac{0,08}{16} + \frac{0,0075}{8} + \frac{0,0025}{8} = 0,0089$$

Sabendo-se que o coeficiente de endocruzamento de filhos de casais de primos em quarto grau é  $F = \frac{1}{128} = 0,0078$ , podemos, pois, dizer que na cidade B o grau médio de consangüinidade de seus casais é um pouco superior ao dos casais de uma população hipotética na qual todos os casamentos se dessem entre primos em quarto grau. Em relação à cidade A dir-se-ia o oposto, isto é, que o grau médio de consangüinidade de seus casais é inferior ao dos casais dessa cidade hipotética.

**COEFICIENTE DE ENDOCRUZAMENTO E GENES DO CROMOSSOMO X**

Quando o coeficiente de endocruzamento diz respeito a genes do cromossomo X é óbvio que se deve levar em conta apenas as filhas de consangüíneos, pois somente elas possuem, normalmente, dois cromossomos X em seu cariótipo, o que lhes dá a oportunidade de apresentar esses genes em homozigose. Para entender o cálculo do coeficiente de endocruzamento em relação aos genes do cromossomo X, que é, geralmente, simbolizado por  $F'$ , consideremos os quatro tipos de casamento entre primos em primeiro grau (Figura 8.5).

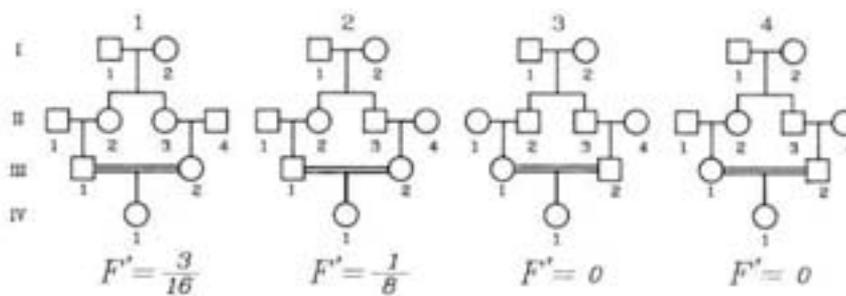


Fig. 8.5. Coeficientes de endocruzamento das filhas de primos em primeiro grau em relação a genes do cromossomo X.

Analisemos, inicialmente, a probabilidade de os dois cromossomos X de uma filha de cada um desses tipos de primos serem do seu bisavô I-1, que é ancestral comum a seus pais. Evidentemente, para que isso aconteça é necessário que o cromossomo X do indivíduo I-1 seja transmitido seguindo dois caminhos:

Caminho 1: I-1 → II-2 → III-1 → IV-1

Caminho 2: I-1 → II-3 → III-2 → IV-1

A observação, ainda que rápida, da Figura 8.5 permite constatar logo que nos heredogramas 2, 3 e 4 não existe a possibilidade de os dois cromossomos X da mulher IV-1 terem sido herdados do bisavô I-1, porque, em tais heredogramas, ao menos um dos caminhos apresenta pelo menos dois homens em sucessão, o que inviabiliza a transmissão do cromossomo X. Desse modo, em relação às genealogias representadas pelos heredogramas 2, 3 e 4 pode-se dizer que é nula a probabilidade de uma filha de primos em primeiro grau (IV-1) ter os genes do cromossomo X de seu bisavô (I-1) em autozigose.

Já na genealogia representada pelo heredograma 1 da Figura 8.5 essa probabilidade é igual a  $\frac{1}{8}$ , pois existe a certeza de que o cromossomo X do bisavô I-1 foi transmitido às suas filhas II-2 e II-3, podendo-se, ainda, dizer que, se tal cromossomo tiver sido herdado pelo indivíduo III-1, ele será, com certeza, transmitido à mulher IV-1. Portanto, aos passos I-1 → II-2, I-1 → II-3 e III-1 → IV-1 corresponde uma probabilidade 1. A cada um dos três passos restantes dos caminhos 1 e 2 corresponde, evidentemente, uma probabilidade  $\frac{1}{2}$  de o cromossomo X do bisavô I-1 ter sido transmitido. Com isso, tem-se que a probabilidade de a mulher IV-1 do heredograma 1 da Figura 8.5 ter os genes do cromossomo X de seu bisavô I-1 em autozigose é igual a  $\left(\frac{1}{2}\right)^3 = \frac{1}{8}$ .

Vejam, agora, qual a probabilidade de os dois cromossomos X da mulher IV-1 de cada um dos heredogramas da Figura 8.5 serem idênticos a um dos dois cromossomos sexuais presentes na bisavó I-2, os quais podem ser designados por  $X^1$  e  $X^2$ . Para que isso aconteça é necessário que o cromossomo  $X^1$  ou o cromossomo  $X^2$  sejam transmitidos percorrendo os dois caminhos seguintes:

Caminho 3: I-2 → II-2 → III-1 → IV-1

Caminho 4: I-2 → II-3 → III-2 → IV-1

Novamente, a simples observação da Figura 8.5 mostra logo que a probabilidade de tal acontecimento é nula nas genealogias representadas pelos heredogramas 3 e 4, porque no caminho 4 existe a sucessão de dois homens, ou seja, ao passo II-3 → III-2 corresponde uma probabilidade zero. Isso permite dizer que as filhas de casais de primos em primeiro grau dos tipos 3 e 4 da Fig.8.5 têm  $F' = 0$ , visto que elas não podem possuir em duplicata o cromossomo X de seu bisavô I-1, do mesmo modo que não podem possuir em duplicata um dos dois cromossomos X de sua bisavó I-2.

No concernente à genealogia representada pelo heredograma 1 da Figura 8.5, deve-se considerar que, com exceção do passo III-1 → IV-1 do caminho 3, que tem probabilidade 1, cada um de todos os restantes correspondem a uma probabilidade  $\frac{1}{2}$ . Portanto, a probabilidade de os dois cromossomos X da mulher IV-1 desse heredograma serem  $X^1$  é

igual a  $\left(\frac{1}{2}\right)^5 = \frac{1}{32}$ , pois existem cinco passos com probabilidade  $\frac{1}{2}$  nos caminhos 3 e 4 tomados conjuntamente. Obviamente, a probabilidade de os dois cromossomos da mesma mulher IV-1 serem  $X^2$  será, também, igual a  $\frac{1}{32}$ .

Por isso, recordando que  $\frac{1}{8}$  é a probabilidade de os dois cromossomos  $X$  dessa mulher serem do bisavô I-1, tem-se que as filhas de casais de primos em primeiro grau do tipo 1 da Figura 8.5 têm coeficiente de endocruzamento em relação a genes do cromossomo  $X$  igual a  $\frac{3}{16}$ , pois  $F' = \frac{1}{8} + \frac{1}{32} + \frac{1}{32} = \frac{3}{16}$ . Aplicando o mesmo raciocínio na análise da genealogia representada pelo heredograma 2 da Figura 8.5, concluiremos que o valor de  $F'$  da mulher IV-1 nele representada é igual a  $\frac{1}{8}$ , pois  $F' = 0 + \frac{1}{16} + \frac{1}{16} = \frac{1}{8}$ .

O valor de  $F'$  pode ser obtido por intermédio da fórmula  $F' = \Sigma(1+F_a)\left(\frac{1}{2}\right)^n$  na qual  $n$  é o número de mulheres em cada linha que une os consangüíneos e  $F_a$  é O coeficiente de endocruzamento dos ancestrais do sexo feminino.

Na Figura 9.5 foram representados heredogramas de genealogias com alguns tipos de casais com diferentes graus de consangüinidade e assinalados os coeficientes de endocruzamento das filhas desses casais em relação a genes do cromossomo  $X$ .

### PROBABILIDADE DE OS FILHOS DE CASAIS CONSAGÜÍNEOS SEREM HOMOZIGOTOS

Os filhos de casais consangüíneos podem ser homozigotos por serem autozigotos mas, evidentemente, eles também podem ser homozigotos sem que ocorra autozigose, isto é, os genes em homozigose nos filhos de consangüíneos podem ser oriundos de dois ancestrais e não de um único. Tendo isso em mente, vejamos como calcular a probabilidade de homozigose de filhos de casais consangüíneos considerando um par de alelos autossômicos  $A, a$ , que ocorrem na população com freqüências iguais, respectivamente, a  $p$  e  $q$ , sendo  $p + q = 1$ .

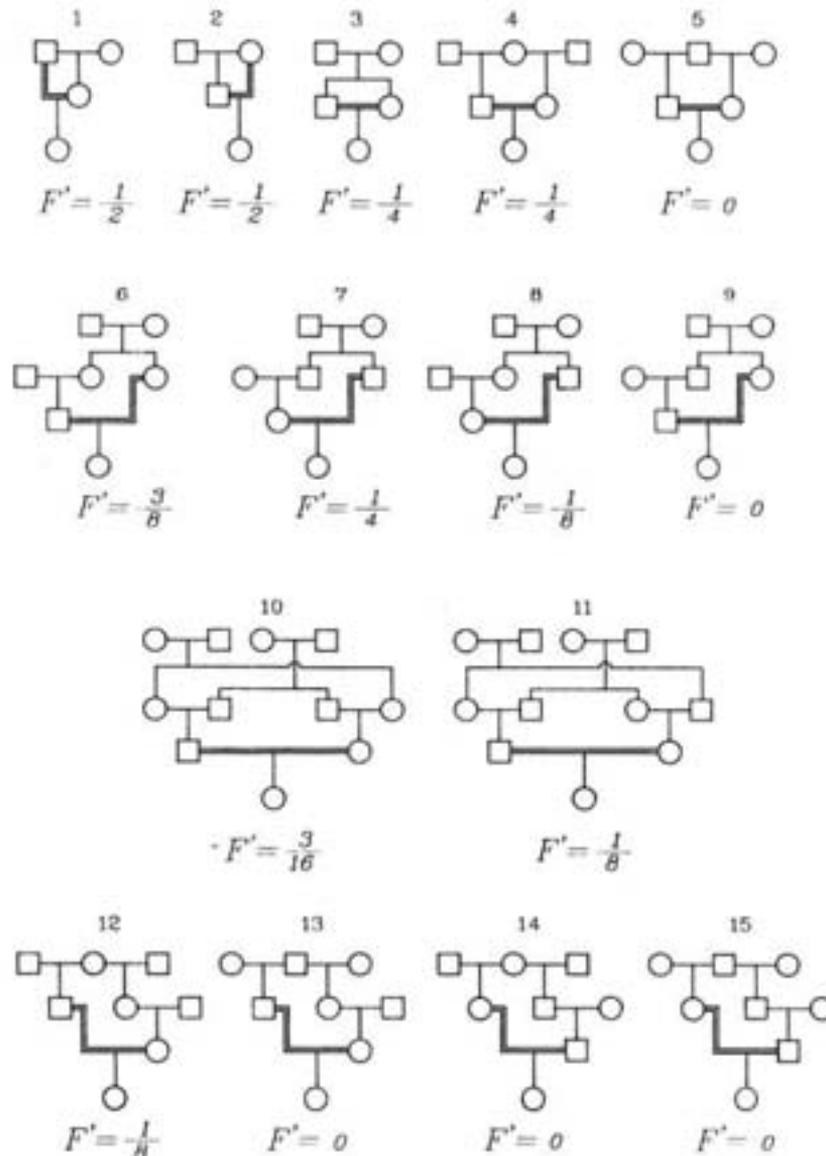


Fig. 9.5. Heredogramas de genealogias que incluem casais consangüíneos, inclusive incestuosos, e coeficientes de endocruzamento em relação a genes do cromossomo X das filhas desses casais. 1- pai × filha; 2- mãe × filho; 3- irmãos; 4, 5- meio-irmãos; 6, 9 - tia × sobrinho; 7, 8 - tio × sobrinha; 10, 11- primos duplos em primeiro grau; 12, 13 - tio × meia-sobrinha; 14, 15- tia × meio sobrinho.

Para calcular a probabilidade de um filho de um casal com um determinado grau de consangüinidade ser homozigoto  $AA$  por autozigose temos que conhecer a probabilidade de o alelo  $A$  estar presente em um dos ancestrais comuns ao casal consangüíneo. Essa probabilidade é  $p$ , pois  $p$  é a freqüência do gene  $A$  na população. Multiplicando  $p$  pelo coeficiente de endocruzamento ( $F$ ) do filho do casal consangüíneo em estudo, obtemos a

probabilidade de esse indivíduo ser autozigoto  $AA$ . Assim, por exemplo, se o casal consanguíneo fosse composto por primos em primeiro grau e a frequência do alelo  $A$  fosse igual a 0,80, essa probabilidade seria igual a 0,05, porque  $Fp = \frac{1}{16} \cdot 0,80 = 0,05$ . Se o casal consanguíneo fosse composto por primos duplos em primeiro grau e a frequência do alelo  $A$  continuasse a mesma, teríamos  $Fp = \frac{1}{8} \cdot 0,80 = 0,10$ .

Se  $F$  é a probabilidade de um filho de um casal consanguíneo ser autozigoto de um gene presente em um de seus ancestrais, é claro que  $1 - F$  será a probabilidade de isso não ocorrer, isto é,  $1 - F$  será a probabilidade de tal filho não ser autozigoto de um gene presente em um dos seus ancestrais. Portanto, para calcular a probabilidade de um filho de um casal consanguíneo ser homozigoto  $AA$  sem que haja autozigose, temos que multiplicar a probabilidade  $p$  de o gene  $A$  encontrar outro alelo  $A$  de origem diversa, ou seja,  $(1 - F)p \cdot p = (1 - F)p^2$ . Em uma população na qual o alelo  $A$  tem frequência  $p = 0,80$ , um filho de um casal de primos em primeiro grau teria, pois, 60% de probabilidade de ser homozigoto  $AA$  sem autozigose porque  $(1 - F)p^2 = \frac{15}{16} \cdot 0,64 = 0,60$ .

Desse modo, pode-se afirmar que a probabilidade de um filho de um casal com um determinado grau de consangüinidade ser homozigoto de um alelo  $A$ , cuja frequência na população é  $p$ , é igual à probabilidade de ele ser homozigoto por autozigose, isto é,  $Fp$ , somada à probabilidade de ele ser homozigoto sem autozigose, isto é,  $(1 - F)p^2$  ou, resumidamente,

$$P(AA) = Fp + (1 - F)p^2$$

Tendo em mente que o mesmo raciocínio se aplica ao alelo  $a$  pode-se, afirmar, também, que a probabilidade de um casal com determinado grau de consangüinidade ser homozigoto de um alelo  $a$ , cuja frequência na população é  $q$ , é igual à probabilidade  $Fq$  de ele ser homozigoto por autozigose somada à probabilidade  $(1 - F)q^2$  de ele ser homozigoto sem autozigose, o que permite escrever

$$P(aa) = Fq + (1 - F)q^2$$

Em consequência do exposto, somos levados a concluir que a probabilidade de um filho de um casal com determinado grau de consangüinidade ser heterozigoto em relação a alelos  $A, a$  com frequências  $p$  e  $q$ , respectivamente, na população, é

$$P(Aa) = (1 - F)2pq$$

De fato, o gene  $A$  pode estar presente em um dos ancestrais comuns ao casal consanguíneo (probabilidade  $p$ ), não se encontrar por autozigose com um alelo  $A$  idêntico no filho desse casal (probabilidade  $1 - F$ ), mas encontrar-se com um alelo  $a$  (probabilidade  $q$ ). A probabilidade de isso acontecer, isto é, de haver heterozigose, é, pois,  $(1 - F)pq$ . Por outro lado, o gene  $a$  pode estar presente em um dos ancestrais comuns ao casal consanguíneo (probabilidade  $q$ ), não se encontrar por autozigose com um alelo  $a$  idêntico no filho desse casal (probabilidade  $1 - F$ ), mas encontrar-se com um alelo  $A$  (probabilidade  $p$ ) havendo, também neste caso, probabilidade  $(1 - F)pq$  de heterozigose. Portanto, a probabilidade total de heterozigose  $Aa$  de um filho de um casal com um determinado tipo de consangüinidade passa a ser  $P(Aa) = (1 - F)pq + (1 - F)pq = (1 - F)2pq$ , como havíamos assinalado acima.

As expressões que indicam as probabilidades de homozigose  $AA$  e  $aa$ , e de heterozigose  $Aa$  de um filho de um casal com determinado tipo de consangüinidade também podem ser apresentadas como abaixo:

$$P(AA) = p^2 + Fpq$$

$$P(aa) = q^2 + Fpq$$

$$P(Aa) = 2pq - 2Fpq$$

porque:

$$P(AA) = Fp + (1 - F)p^2 = Fp + p^2 - Fp^2 = p^2 + Fp(1 - p) = p^2 + Fpq$$

$$P(aa) = Fq + (1 - F)q^2 = Fq + q^2 - Fq^2 = q^2 + Fq(1 - q) = q^2 + Fpq$$

$$P(Aa) = (1 - F)2pq = 2pq - 2Fpq$$

A exposição feita no presente tópico fornece elementos para demonstrar, de outro modo, a razão pela qual os casais consanguíneos estão sujeitos a maior risco do que os não-consanguíneos de gerar uma criança com uma anomalia recessiva. Para exemplificar, consideremos uma anomalia recessiva autossômica monogênica que ocorre em uma população em equilíbrio de Hardy e Weinberg com freqüência igual a 1 por 10.000, ou seja, 0,0001. A freqüência do gene  $a$  determinador dessa anomalia será, portanto, estimada como  $q = 0,01$ , o que equivale a dizer que a freqüência de seu alelo  $A$ , determinador de normalidade, é estimada como  $p = 0,99$ . Desse modo, pode-se afirmar que **um filho de**

*primos em primeiro grau corre o risco sete vezes mais alto do que os filhos de casais não-consangüíneos de manifestar a anomalia recessiva em questão.* Realmente, enquanto os filhos desses últimos correm o risco de 1:10.000 de manifestar a anomalia recessiva em apreço ( $q^2 = 0,0001$ ), um filho de primos em primeiro grau está sujeito a um risco de 0,0007 de manifestar essa mesma anomalia, pois

$$P(aa) = q^2 + Fpq = 0,0001 + \frac{1}{16} \times 0,99 \times 0,01 = 0,0007$$

Um filho de primos duplos em primeiro grau correria risco mais alto de manifestar a anomalia em questão porque, nesse caso

$$P(aa) = 0,0001 + \frac{1}{8} \times 0,99 \times 0,01 = 0,0013$$

## O EQUILÍBRIO DE WRIGHT

As considerações feitas no tópico anterior podem ser generalizadas para populações humanas que tiverem coeficiente médio de endocruzamento ( $\bar{F}$ ) maior do que zero. Assim, em relação a uma população com  $\bar{F}$  diferente de zero não se afirmará que os genótipos  $AA$ ,  $Aa$  e  $aa$  decorrentes de um par de alelos autossômicos  $A,a$  com frequências  $p$  e  $q$ , respectivamente, se distribuem segundo a lei de Hardy e Weinberg, isto é, com  $AA = p^2$ ,  $Aa = 2pq$  e  $aa = q^2$ . Levando em conta o valor de  $\bar{F}$ , tais genótipos se distribuirão de acordo com a *lei do equilíbrio de Wright*, isto é, do seguinte modo:

$$AA = p^2 + \bar{F}pq, \quad Aa = 2pq - 2\bar{F}pq \quad \text{e} \quad aa = q^2 + \bar{F}pq.$$

Esse tipo de equilíbrio genético, que tem seu nome associado ao do insigne geneticista norte-americano Sewall Wright, permite estabelecer duas conclusões importantes para a Genética de Populações. A primeira conclusão é a de que os casamentos consangüíneos são capazes de alterar as frequências genotípicas, sem afetar, por si só, as frequências gênicas, pois, se não existir seleção preferencial contra ou a favor de algum genótipo, as frequências gênicas manter-se-ão constantes, independentemente do coeficiente médio de endocruzamento da população.

Para demonstrar essa conclusão consideremos um par de alelos autossômicos  $A,a$  com frequências, respectivamente,  $p = 0,99$  e  $q = 0,01$  em uma população que está em equilíbrio de Hardy e Weinberg, isto é, com  $\bar{F} = 0$ . Em tal população, portanto, os genótipos decorrentes desses alelos serão encontrados com frequências  $AA = p^2 = 0,9801$ ,  $Aa = 2pq = 0,0198$  e  $aa = 0,0001$ . Se em uma geração dessa população os casamentos

consangüíneos começarem a ocorrer de modo a provocar um coeficiente médio de endocruzamento  $\bar{F} = 0,01$ , a distribuição genotípica em relação aos alelos  $A, a$  passará a ser  $AA = p^2 + \bar{F}pq = 0,9802$ ,  $Aa = 2pq - 2\bar{F}pq = 0,0196$  e  $aa = q^2 + \bar{F}pq = 0,0002$ . Se o coeficiente médio de endocruzamento continuar aumentando, de modo a atingir  $\bar{F} = 0,03$ , a distribuição genotípica passará a ser  $AA = 0,9804$ ,  $Aa = 0,0192$  e  $aa = 0,0004$ .

Como se vê, apesar de, na população utilizada para exemplo, as frequências dos alelos  $A$  e  $a$  serem sempre as mesmas, isto é  $p = 0,99$  e  $q = 0,01$ , as frequências genotípicas ficaram alteradas com a elevação do coeficiente médio de endocruzamento. Essa alteração consiste no aumento dos homozigotos  $AA$ , cuja frequência passou de 0,9801 para 0,9802 e, em seguida, para 0,9803, e dos homozigotos  $aa$ , cuja frequência passou de 0,0001 para 0,0002 e, depois, para 0,0004, bem como na diminuição dos heterozigotos  $Aa$ , cuja frequência passou de 0,0198 para 0,0196 e para 0,0192.

A segunda conclusão que se pode tirar do equilíbrio de Wright é a de que a frequência de homozigotos de um determinado gene é diretamente proporcional ao coeficiente médio de endocruzamento e inversamente proporcional à frequência desse gene. Realmente, no exemplo utilizado, se nos reportarmos ao alelo raro ( $a$ ) e compararmos a frequência de homozigotos  $aa$  quando  $\bar{F} = 0,01$  com aquela que a população mostrou quando  $\bar{F} = 0$ , notaremos que a primeira frequência é duas vezes mais alta do que a

segunda, pois  $\frac{q^2 + \bar{F}pq}{q^2} = \frac{0,0002}{0,0001} = 2$ . Em outras palavras, considerando que  $\frac{\bar{F}pq}{q^2}$  mede o aumento relativo de homozigotos que uma população apresenta quando  $\bar{F} = 0$  passa a  $\bar{F} > 0$ , pode-se dizer que na população de nosso exemplo os homozigotos  $aa$  aumentaram 100% quando  $\bar{F}$  passou de zero a 0,01, pois  $\frac{\bar{F}pq}{q^2} = 1$  ou 100%. Quando  $\bar{F}$  cresceu para 0,03, a

frequência de homozigotos  $aa$  passou a ficar quatro vezes mais alta do que a apresentada pela população com  $\bar{F} = 0$ , isto é, houve um aumento de 300% dos homozigotos, porque  $\frac{q^2 + \bar{F}pq}{q^2} = \frac{0,0004}{0,0001} = 4$  e  $\frac{\bar{F}pq}{q^2} = \frac{0,0003}{0,0001} = 3$  ou 300%.

Já em relação ao alelo  $A$ , o crescimento do coeficiente médio de endocruzamento afetou muito pouco a frequência de homozigotos, porque esse gene tem frequência alta na população. Assim, quando  $\bar{F}$  passou de zero a 0,01, a frequência de homozigotos teve um

aumento relativo de apenas 0,01% porque  $\frac{\bar{F}pq}{q^2} = \frac{0,0001}{0,9801} = 0,0001$  ou 0,01%. Em razão da frequência alta do gene  $A$ , o aumento relativo foi desprezível mesmo quando  $\bar{F} = 0,03$ , pois foi igual a 0,03%, visto que  $\frac{\bar{F}pq}{q^2} = \frac{0,0003}{0,9801} = 0,0003$  ou 0,03%.

Como se vê, tanto no caso de  $\bar{F} = 0,01$  quanto no caso de  $\bar{F} = 0,03$  o aumento relativo de homozigotos  $aa$  foi 10.000 vezes maior do que o de homozigotos  $AA$ , pois  $\frac{1}{0,0001} = \frac{3}{0,0003} = 10.000$ . É por causa desse efeito diferente do coeficiente médio de endocruzamento sobre os genes freqüentes e os raros que, em relação aos polimorfismos genéticos, as taxas de casamentos consangüíneos usualmente encontradas nas populações humanas afetam pouco a distribuição das freqüências genotípicas esperadas segundo a lei de Hardy e Weinberg.

Para demonstrar mais claramente essa última afirmação, consideremos que um pesquisador estivesse estudando a distribuição dos grupos consangüíneos  $M$ ,  $MN$  e  $N$  em uma população caucasóide do Sul do Brasil, na qual as estimativas das freqüências dos alelos  $M$  e  $N$  são, respectivamente,  $p = 0,53$  e  $q = 0,47$ . Consideremos, ainda, que essa população apresente um coeficiente médio de endocruzamento  $\bar{F} = 0,02$ , que é muito alto para populações dessa região brasileira, as quais, na época atual, apresentam, geralmente, valores de  $\bar{F}$  cerca de 50 a 100 vezes menores do que o de nosso exemplo (Bassi e Freire-Maia, 1985; Agostini e Meireles-Nasser, 1986).

Mesmo desconhecendo tão alto valor do coeficiente médio de endocruzamento, as conclusões do pesquisador a respeito da distribuição dos grupos sangüíneos  $M$ ,  $MN$  e  $N$  não seriam afetadas, porque a distribuição esperada na população em equilíbrio de Wright ( $M = 0,2859$ ;  $MN = 0,4882$ ;  $N = 0,2259$ ) difere muito pouco daquela esperada em equilíbrio de Hardy e Weinberg ( $M = 0,2809$ ;  $MN = 0,4982$ ;  $N = 0,2209$ ). Ainda que a população estudada mostrasse exatamente os valores esperados em equilíbrio de Wright, seria necessário que a amostra coletada pelo pesquisador fosse muito grande, isto é, de cerca de 10.000 indivíduos, para que se pudesse notar que a distribuição dos grupos sangüíneos  $M$ ,  $MN$  e  $N$  se desvia significativamente daquela esperada em equilíbrio de Hardy e Weinberg.

De fato, mesmo que a amostra fosse constituída por 9.000 indivíduos, dos quais 2.573 fossem do grupo M, 4.394 do grupo MN e 2.033 do grupo N, não seria possível rejeitar a hipótese de que a população da qual foi extraída a amostra está em equilíbrio de Hardy e Weinberg porque, com esses dados, estimariamos  $p = \frac{2.573 + 2.197}{9.000} = 0,53$  e  $q = 1 - p = 0,47$  e faríamos o cálculo abaixo, obtendo um qui-quadrado inferior ao crítico.

Valores	M	MN	N	Total
Observados ( <i>o</i> )	2.573	4.394	2.033	9.000
Esperados ( <i>e</i> )	2.528	4.484	1.988	9.000
$\frac{(o-e)^2}{e}$	0,801	1.806	1,019	$\chi^2_{(1)} = 3,626$ 0,05 < P < 0,10

Consideremos, agora, que o pesquisador, em vez de um polimorfismo genético, como o dos grupos sanguíneos M, MN e N, estivesse estudando uma anomalia recessiva autossômica com incidência igual a 1:10.000 nessa mesma população. Nesse caso, o desconhecimento do alto valor de seu coeficiente médio de endocruzamento provocaria grande distorção nos resultados, pois, sem levar em conta que  $\bar{F} = 0,02$ , isto é, aceitando  $\bar{F}$  como igual ou próximo de zero, o pesquisador aceitaria 0,0001 como estimativa de  $q^2$ , de modo que estimaria a freqüência do gene responsável pelo caráter recessivo como sendo  $q = 0,01$ . Contudo, levando em conta o coeficiente médio de endocruzamento concluiria que a freqüência desse gene é bem menor. De fato, tendo em mente que  $0,0001 = q^2 + \bar{F}pq = \bar{F}q + (1 - \bar{F})q^2$  substitui-se  $\bar{F}$  por seu valor (0,02), obtendo-se a equação do segundo grau  $0,98q^2 + 0,02q - 0,0001 = 0$ , cuja raiz positiva fornece  $q = 0,004$ .

### COMPONENTE PANMÍCTICO E COMPONENTE FIXO DA POPULAÇÃO

A expressão da lei do equilíbrio de Wright permite concluir que em certos seres, como muitos vegetais, em que a autofecundação é possível, os heterozigotos desaparecem da população quando esse tipo de reprodução é a regra. Nesse caso extremo de endogamia, a população é rapidamente conduzida à **fixação dos genótipos homozigotos** ou, simplesmente, à **fixação**. Em relação a um par de alelos *A, a* tal população que atinge a fixação mostra apenas os genótipos *AA* e *aa*, cujas freqüências serão iguais a *A* e *a*. Nesse caso,  $\bar{F} = 1$ .

$$AA = p^2 + \bar{F}pq = p^2 + p(1-p) = p$$

$$Aa = 2pq - 2\bar{F}pq = 2pq - 2pq = 0$$

$$aa = q^2 + \bar{F}pq = q^2 + (1-q)q = q$$

Evidentemente, a fixação dos homozigotos ocorrerá não apenas em relação a um par de alelos, mas afetará todos os locos gênicos. Assim, por exemplo, se considerarmos um outro par de alelos ( $B, b$ ), encontraremos, quando houver fixação completa, apenas os genótipos  $BB$  e  $bb$  e, quando eles forem analisados simultaneamente com os alelos  $A, a$  poderemos encontrar os genótipos  $AABB$ ,  $AAbb$ ,  $aaBB$  e  $aabb$ , mas não os genótipos  $AABb$ ,  $AaBB$ ,  $AaBb$ ,  $Aabb$  nem  $aaBb$ .

A diminuição dos heterozigotos em consequência da endogamia é o princípio empregado na produção de linhagens *isogênicas* de animais e vegetais, porque, a menos que haja mutações, tem-se, numa população estritamente endogâmica, que cada indivíduo acabará por gerar uma prole cujos genes são idênticos aos dele. É o caso das raças de animais de laboratório, obtidas a partir do acasalamento de irmãos durante muitas gerações. Desse modo, fica aumentada não apenas a probabilidade de identidade genética dos indivíduos, isto é, sua isogenicidade, mas, também a sua homozigosidade. Está claro, porém, que o sistema de acasalamento entre irmãos requer maior número de gerações do que a autofecundação para se conseguir o desaparecimento dos heterozigotos.

Aqui é importante lembrar que a manutenção da isogenicidade dos animais de laboratório tem importância crucial para a interpretação dos resultados de pesquisas que visam à avaliação dos efeitos do ambiente, pois, trabalhando com tais animais, os diferentes resultados obtidos serão atribuídos às variações do ambiente, às técnicas adotadas etc., e não às variações genóticas. Na espécie humana, os gêmeos monozigóticos, também se prestam a esse tipo de estudo por causa de sua isogenicidade, mas a isogenicidade de tais gêmeos não é comparável à das linhagens isogênicas animais e vegetais, porque nos gêmeos monozigóticos ela não implica em homozigosidade.

Conhecendo-se o fenômeno da fixação, é possível apresentar o equilíbrio de Wright, não como um desvio da panmixia, como se fez até agora, mas como um desvio da fixação completa, quando  $AA = p$  e  $aa = q$ . De fato, consideremos uma população teórica, com  $\bar{F} = 1$ , a qual passasse a apresentar um coeficiente médio de endocruzamento entre zero e a unidade ( $0 < \bar{F} < 1$ ). Nesse caso, a frequência de homozigotos  $AA$  deixaria de ser  $p$ , pois

diminuiria para  $p - (1 - \bar{F})pq$ , enquanto a frequência de homozigotos  $aa$  passaria de  $q$  para  $q - (1 - \bar{F})pq$ . Disso resulta que a frequência de heterozigotos  $Aa$  crescerá de zero para  $2(1 - \bar{F})pq$ .

É simples demonstrar que as expressões das proporções genótípicas sob a forma de desvio da fixação e de desvio da panmixia se equivalem, como se pode ver abaixo:

$$AA = p - (1 - \bar{F})pq = p - pq + \bar{F}pq = p - (1 - p)p + \bar{F}pq = p^2 + \bar{F}pq$$

$$Aa = 2(1 - \bar{F})pq = 2pq - 2\bar{F}pq$$

$$aa = q - (1 - \bar{F})pq = q - pq + \bar{F}pq = q - (1 - q)q + \bar{F}pq = q^2 + \bar{F}pq$$

Um outro modo de expressar o equilíbrio de Wright pode ser conseguido quando consideramos, simultaneamente, o grau de endogamia da população, isto é, o seu coeficiente médio de endocruzamento ( $\bar{F}$ ) e o seu grau de panmixia, indicado por  $1 - \bar{F}$ . Dessa maneira, pode-se avaliar um **componente fixo**, decorrente de  $\bar{F}$ , e um **componente panmíctico** ( $1 - \bar{F}$ ) da população em estudo. Isso pode ser obtido por intermédio das transformações abaixo:

$$AA = p^2 + \bar{F}pq = p^2 + \bar{F}p(1 - p) = p^2 + \bar{F}p - \bar{F}p^2 = \bar{F}p + (1 - \bar{F})p^2$$

$$Aa = 2pq - 2\bar{F}pq = 2(1 - \bar{F})pq$$

$$aa = q^2 + \bar{F}pq = q^2 + \bar{F}q(1 - q) = q^2 + \bar{F}q - \bar{F}q^2 = \bar{F}q + (1 - \bar{F})q^2$$

A Tabela 1.5 resume a maior parte das considerações feitas neste tópico.

Tabela 1.5 Frequências genótípicas esperadas em uma população panmíctica ( $\bar{F} = 0$ ), com fixação completa ( $\bar{F} = 1$ ) e com  $0 < \bar{F} < 1$ . Nesse último caso são indicados os três modos de representar o equilíbrio de Wright.

Genótipo	$(\bar{F} = 0)$	$(\bar{F} = 1)$	$0 < \bar{F} < 1$		
	Panmixia	Fixação	Desvio da Panmixia	Desvio da Fixação	Componente Fixo e Panmíctico
$AA$	$p^2$	$p$	$p^2 + \bar{F}pq$	$p - (1 - \bar{F})pq$	$\bar{F}p + (1 - \bar{F})p^2$
$Aa$	$2pq$	$0$	$2pq - 2\bar{F}pq$	$2(1 - \bar{F})pq$	$2(1 - \bar{F})pq$
$aa$	$q^2$	$q$	$q^2 + \bar{F}pq$	$q - (1 - \bar{F})pq$	$\bar{F}q + (1 - \bar{F})q^2$
Total	$(p+q)^2 = 1$	$p+q = 1$	$(p+q)^2 = 1$	$p+q = 1$	$(p+q)^2 = 1$

### O CÁLCULO DA PROPORÇÃO DE FILHOS DE CASAIS CONSANGÜÍNEOS ENTRE HOMOZIGOTOS

Consideremos que, em uma certa população, a frequência de casais com um determinado grau de consangüinidade é  $c$ , e que nela ocorra um par de alelos autossômicos  $A, a$  com frequências  $p$  e  $q$ , respectivamente. Levando em conta apenas o alelo  $a$  tem-se,

entre os filhos de consangüíneos, cujo coeficiente de endocruzamento é  $F$ , que a probabilidade de homozigose  $aa$  pode ser calculada por intermédio de  $Fq + (1 - F)q^2$ . Portanto, entre os filhos de casais com determinado grau de consangüinidade que ocorrem com frequência  $c$  e cujos filhos têm coeficiente de endocruzamento  $F$ , a frequência esperada de homozigotos  $aa$  deve ser  $c[Fq + (1 - F)q^2]$ . Por outro lado, se a população em estudo apresentar coeficiente médio de endocruzamento  $\bar{F}$ , ela deverá conter uma frequência de homozigotos  $aa$  que não deve desviar-se significativamente de  $\bar{F}q + (1 - \bar{F})q^2$ . Daí se pode concluir que, para estimar a proporção esperada de filhos de casais com um determinado grau de consangüinidade entre os homozigotos  $aa$  (proporção  $k$ ) bastará resolver a fórmula:

$$k = \frac{c[Fq + (1 - F)q^2]}{\bar{F}q + (1 - \bar{F})q^2} = \frac{c[F + (1 - F)q]}{\bar{F} + (1 - \bar{F})q}$$

Assim, por exemplo, se quisermos saber qual a proporção esperada de filhos de primos em primeiro grau entre os homozigotos  $aa$ , a fórmula acima passará a ser escrita como abaixo, pois o valor de  $F$  será substituído por  $\frac{1}{16}$ , que é o coeficiente de endocruzamento de filhos de primos em primeiro grau:

$$k = \frac{c(1 + 15q)}{16[\bar{F} + (1 - \bar{F})q]}$$

Quando o valor de  $\bar{F}$  é muito baixo, a fórmula geral para calcular a proporção  $k$  pode ser mais simplificada ainda, visto que a diferença  $1 - \bar{F}$  passará a ser próxima da unidade. Assim, poder-se-á calcular  $k$  a partir de:

$$k = \frac{c[F + (1 - F)q]}{\bar{F} + q}$$

Na Tabela 2.5 estão apresentadas as fórmulas de cálculo de  $k$  em relação a filhos de casais com diferentes graus de consangüinidade. Tais fórmulas, juntamente com os dados da Tabela 3.5 indicam, claramente, que a proporção  $k$  de filhos de consangüíneos entre homozigotos é:

1. diretamente proporcional à frequência  $c$  dos casamentos consangüíneos sob investigação;
2. inversamente proporcional à frequência do gene que é objeto de estudo;

3. inversamente proporcional ao coeficiente médio de endocruzamento da população.

Tabela 2.5. Fórmulas para calcular a proporção  $k$  esperada de filhos de casais com diferentes graus de consangüinidade entre os homozigotos de um gene autossômico  $a$  com freqüência  $q$ .

Casal	$\bar{F}$ Alto	$\bar{F}$ Baixo
Pai × Filha Mãe × Filho Irmãos	$k = \frac{c(1+3q)}{4[\bar{F} + (1-\bar{F})q]}$	$k = \frac{c(1+3q)}{4(\bar{F} + q)}$
Meio irmãos Tio(a) × Sobrinha(o) Primos duplos em 1º. grau	$k = \frac{c(1+7q)}{8[\bar{F} + (1-\bar{F})q]}$	$k = \frac{c(1+7q)}{8(\bar{F} + q)}$
Primos em 1º. grau Tio(a) × Meia(o) sobrinha(o)	$k = \frac{c(1+15q)}{16[\bar{F} + (1-\bar{F})q]}$	$k = \frac{c(1+15q)}{16(\bar{F} + q)}$
Primos em 2º. grau	$k = \frac{c(1+31q)}{32[\bar{F} + (1-\bar{F})q]}$	$k = \frac{c(1+31q)}{32(\bar{F} + q)}$
Primos em 3º. grau	$k = \frac{c(1+63q)}{64[\bar{F} + (1-\bar{F})q]}$	$k = \frac{c(1+63q)}{64(\bar{F} + q)}$

Tabela 3.5. Proporções esperadas ( $k$ ) de filhos de primos em primeiro grau entre os homozigotos de um gene autossômico com diferentes freqüências  $q$ , segundo várias taxas de casamento entre primos em primeiro grau ( $c$ ) e diferentes valores de coeficiente médio de endocruzamento ( $\bar{F}$ ).

$q$	$c = 0,005$			$c = 0,01$			$c = 0,02$		
	$\bar{F} = 0,005$	$\bar{F} = 0,01$	$\bar{F} = 0,02$	$\bar{F} = 0,005$	$\bar{F} = 0,01$	$\bar{F} = 0,02$	$\bar{F} = 0,005$	$\bar{F} = 0,01$	$\bar{F} = 0,02$
0,001	0,0529	0,0290	0,0152	0,1058	0,0580	0,0304	0,2116	0,1155	0,0605
0,005	0,0338	0,0226	0,0136	0,0677	0,0452	0,0271	0,1347	0,0899	0,0540
0,01	0,0242	0,0182	0,0122	0,0481	0,0361	0,0241	0,0962	0,0722	0,0482
0,02	0,0163	0,0136	0,0103	0,0326	0,0273	0,0205	0,0653	0,0545	0,0410
0,05	0,0100	0,0092	0,0080	0,0200	0,0184	0,0159	0,0400	0,0368	0,0317
0,10	0,0075	0,0072	0,0066	0,0150	0,0143	0,0132	0,0299	0,0287	0,0265

Se a proporção  $k$  de filhos de casais com um determinado grau de consangüinidade entre os homozigotos não for significativamente mais alta do que a freqüência  $c$  desses casais na população é evidente que a razão  $\frac{k}{c}$  também não diferirá significativamente da unidade. A investigação dessa razão com base nos dados da Tabela 3.5 permite notar que, para um determinado coeficiente médio de endocruzamento, a razão  $\frac{k}{c}$  é muito mais alta quando o gene em estudo é raro do que quando ele é freqüente, independentemente de a

frequência de casamentos consangüíneos ser alta ou baixa. Além disso, à medida que aumenta a frequência do gene sob análise, a razão  $\frac{k}{c}$  tende para a unidade.

Realmente, usando os dados da Tabela 3.5, verifica-se que, numa população com coeficiente médio de endocruzamento  $\bar{F} = 0,005$ , ter-se-á, independentemente de a frequência de casais de primos em primeiro grau na população ser igual a 0,5 %, 1 % ou 2%, isto é,  $c = 0,005$ ,  $c = 0,01$  ou  $c = 0,02$ , que a proporção  $k$  de filhos de primos em primeiro grau entre os homozigotos de um gene com frequência  $q = 0,001$  será 10,6 vezes maior do que a frequência de casais consangüíneos nessa mesma população, porque:

$$\frac{k}{c} = \frac{0,0529}{0,005} = \frac{0,1058}{0,01} = \frac{0,2116}{0,02} = 10,6$$

Quando, para idêntica comparação, se considera a frequência do gene como  $q = 0,10$ , a razão  $\frac{k}{c}$  baixa para 1,5, pois:

$$\frac{k}{c} = \frac{0,0075}{0,005} = \frac{0,0150}{0,01} = \frac{0,0299}{0,02} = 1,5$$

e tenderá para a unidade, à medida que a frequência do gene aumentar.

Uma aplicação importante derivada do conhecimento da proporção de filhos de casais consangüíneos entre homozigotos é a de se poder estimar, com base nessa proporção, a frequência de genes responsáveis pela manifestação de anomalias recessivas autossômicas monogênicas. Para exemplificar suponhamos que desconhecêssemos que a frequência de um certo gene autossômico que, em homozigose, determina uma anomalia rara, é  $q = 0,005$ . Se, entretanto, com base em dados anamnésticos, tivéssemos a informação de que, na população em estudo, 13,47% dos indivíduos com essa anomalia são filhos de primos em primeiro grau ( $k = 0,1347$ ), e que, nessa mesma população, a taxa de casamentos entre primos em primeiro grau é  $c = 0,02$ , sendo o coeficiente médio de endocruzamento  $\bar{F} = 0,005$ , resolveríamos a fórmula de  $k$  para primos em primeiro grau, apresentada no início deste tópico, em função de  $q$ , escrevendo:

$$q = \frac{c - 16k\bar{F}}{16k(1 - \bar{F}) - 15c}$$

e obteríamos  $q = 0,005$ . Visto que o valor de  $\bar{F}$  é muito baixo, poderíamos considerar  $1 - \bar{F} = 1$  e chegaríamos, praticamente, ao mesmo resultado a partir da fórmula:

$$q = \frac{c - 16k\bar{F}}{16k - 15c}$$

Outra grande aplicação do conhecimento do valor de  $k$  em Genética Médica é a de, por intermédio dele, muitas vezes, podermos obter a indicação de que uma heredopatia recessiva autossômica monogênica, aparentemente homogênea, é, na realidade constituída por mais de uma entidade genético-clínica, isto é, por duas ou mais genocópias. Para demonstrar isso, consideremos uma população com coeficiente médio de endocruzamento  $\bar{F} = 0,002$  e taxa de casamentos entre primos em primeiro grau  $c = 0,01$ , na qual a frequência de recém-nascidos com uma anomalia recessiva autossômica é igual a 1:10.000. Consideremos, também, que 17% dos indivíduos afetados pela anomalia em discussão são filhos de primos em primeiro grau.

Se a análise das irmandades com essa anomalia recessiva apoiar a hipótese monogênica, o nosso primeiro impulso será o de considerar a existência de um único gene ( $a$ ) responsável por ela e o de estimar que a sua frequência é  $q = 0,009$ . De fato, com base nos dados à nossa disposição, e a partir da equação  $aa = \bar{F}q + (1 - \bar{F})q^2$ , podemos calcular:

$$\begin{aligned} 0,0001 &= 0,002q + 0,998q^2 \\ 0,998q^2 + 0,002q - 0,0001 &= 0 \\ q &= \frac{-0,002 + \sqrt{0,000004 + 0,0003992}}{1,996} = 0,009 \end{aligned}$$

Entretanto, se isso fosse verdadeiro, deveríamos esperar que a frequência de filhos de casais de primos em primeiro grau entre os indivíduos afetados pela anomalia fosse  $k = 6,49\%$  e não 17%, porque:

$$k = \frac{c(1 + 15q)}{16[\bar{F} + (1 - \bar{F})q]} = \frac{0,01(1 + 0,135)}{16(0,002 + 0,009)} = 0,0625$$

Como explicar, então, o encontro de tão alta taxa de filhos de primos em primeiro grau entre os recém-nascidos que apresentam a anomalia em discussão? Uma hipótese que poderia ser aventada é a de que a taxa estimada de casais de primos em primeiro grau na população ( $c = 0,01$ ) está errada, isto é, que ela constitui uma sub-estimativa do valor real. Assim, se o valor de  $c$  estivesse entre 0,026 e 0,027, o valor de  $k$ , obtido pela fórmula acima seria 17%. Tal hipótese, contudo, deve ser rejeitada porque ela conduz a um raciocínio incoerente. Realmente, se o valor de  $c$  fosse mais alto, isso significaria que na população existem grupos endogâmicos, responsáveis pela ocorrência da anomalia na população.

Entretanto, se isso fosse verdadeiro, dever-se-ia aceitar que a frequência do gene em discussão é mais alta do que a estimada, pois para obter essa estimativa somente deveríamos levar em conta esses grupos endogâmicos nos quais o eventual gene  $a$  ocorre e não a população geral. Mas, se a frequência do gene  $a$  fosse alta, o valor de  $k$  esperado deveria ser baixo e não tão alto como o encontrado, o que invalida o raciocínio anterior.

O melhor, portanto, é aceitar uma outra hipótese, segundo a qual se supõe a existência de genocópias responsáveis pela manifestação da anomalia em discussão. Assim, por exemplo, se considerássemos que a anomalia recessiva autossômica em apreço é determinada por dois genes não alelos ( $a$  e  $b$ ) em homozigose, que ocorrem com igual frequência na população ( $q_1 = q_2$ ), teríamos que a incidência de recém-nascidos  $aa$  seria igual a 1: 20.000 ou 0,00005, o mesmo ocorrendo com a incidência de recém-nascidos  $bb$ . Em conseqüência disso, diríamos que  $q_1 = q_2 = 0,006$ , pois:

$$\begin{aligned}
 aa = bb &= \bar{F}q + (1-\bar{F})q^2 \\
 0,00005 &= 0,002q + 0,998q^2 \\
 0,998q^2 + 0,002q - 0,00005 &= 0 \\
 q &= \frac{-0,002 + \sqrt{0,000004 + 0,0001996}}{1,996} = 0,006
 \end{aligned}$$

Com base nos valores de  $q_1$  e  $q_2$  podemos calcular a proporção  $k$  esperada de filhos de primos em primeiro grau dentre os homozigotos  $aa$  e  $bb$ , e verificar que, em cada caso, ela é igual a 8,45%. Disso resulta que 16,9% é a proporção esperada de homozigotos  $aa$  e  $bb$  que são filhos de tais casais. Visto que esse valor é praticamente igual ao valor observado (17%), toma-se plausível aceitar a hipótese da existência de genocópias para explicar o excesso de filhos de consangüíneos entre os homozigotos.

### **CÁLCULO DA PROPORÇÃO DE FILHOS DE CASAIS DE PRIMOS EM PRIMEIRO GRAU ENTRE HOMOZIGOTOS QUANDO NÃO SE CONHECE O COEFICIENTE MÉDIO DE ENDOCRUZAMENTO DA POPULAÇÃO**

Os casamentos entre primos em primeiro grau costumam ser acontecimentos bem mais freqüentes do que as ligações incestuosas, os casamentos entre tios(as) e sobrinhas(os) e os casamentos entre primos duplos em primeiro grau. Em vista disso, e por ser difícil a averiguação dos casamentos de indivíduos com consangüinidade mais distante do que a de

primos em primeiro grau, costuma-se aceitar que, nas populações com coeficiente médio de endocruzamento baixo, a maioria dos casais consangüíneos é de primos em primeiro grau.

Tendo isso em mente, quando não é possível conhecer o valor do coeficiente médio de endocruzamento de uma população, apesar de desconfiarmos que ele é baixo, poderemos obter uma estimativa da proporção  $k$  de filhos de casais de primos em primeiro grau sem levar em conta o valor de  $\bar{F}$ . Tal estimativa será tanto mais próxima daquela obtida com o concurso de  $\bar{F}$  quanto menor for o coeficiente médio de endocruzamento da população. Para demonstrar como isso pode ser feito, suponhamos que conhecemos a freqüência  $c$  com que os casais de primos em primeiro grau ocorrem na população, bem como a freqüência  $q$  de um gene autossômico  $a$  que, em homozigose, determina uma anomalia. Visto que a probabilidade de um filho de primos em primeiro grau ser homozigoto de um gene com freqüência  $q$  é  $\frac{q}{16}(1+15q)$  porque:

$$Fq + (1-F)q^2 = \frac{q}{16} + \frac{15}{16}q^2 = \frac{q}{16}(1+15q)$$

pode-se concluir que, se a freqüência de casais de primos é  $c$ , a freqüência esperada de homozigotos  $aa$  que são filhos de casais de primos será  $\frac{cq}{16}(1+15q)$ .

Por considerarmos que, praticamente, todos os casais consangüíneos são casais de primos em primeiro grau,  $1-c$  será a freqüência dos casais restantes, entre os filhos dos quais a probabilidade de encontro de homozigotos é  $q^2$ . Disso resulta que  $(1-c)q^2$  será a freqüência esperada dos homozigotos  $aa$  que não são filhos de casais de primos em primeiro grau. Desse modo, a freqüência total de homozigotos  $aa$  esperada na população será:

$$\frac{cq}{16}(1+15q) + (1-c)q^2$$

expressão essa que pode ser reduzida a  $\frac{cq(1-q) + 16q^2}{16}$ .

Em conseqüência disso, a proporção  $k$  esperada de filhos de primos em primeiro grau entre os homozigotos poderá ser calculada a partir de  $k = \frac{c(1+15q)}{c(1-q) + 16q}$ , porque:

$$k = \frac{\frac{cq}{16}(1+15q)}{\frac{cq(1-q)+16q}{16}} = \frac{c(1+15q)}{c(1-q)+16q}$$

Evidentemente, essa fórmula também pode servir para estimar a frequência de genes responsáveis por anomalias recessivas autossômicas, bastando, para isso, resolvê-la em função de  $q$ . Assim, por exemplo, consideremos que tivéssemos observado que 7,5% dos pacientes com uma determinada anomalia recessiva são filhos de casais de primos em primeiro grau ( $k = 0,075$ ) e que, na população à qual pertencem esses pacientes, esses casais de primos ocorrem com frequência igual a 0,6% ( $c = 0,006$ ). Se não soubéssemos que o coeficiente médio de endocruzamento dessa população é  $\bar{F} = 0,0004$ , poderíamos resolver

$$q = \frac{c(1-k)}{16k - (15+k)c}$$

e assim, estimar em  $q = 0,005$  a frequência do gene responsável pela anomalia recessiva em estudo. Se tivéssemos levado em conta o valor de  $\bar{F}$  e estimado  $q$  a partir da fórmula

$$q = \frac{c - 16k\bar{F}}{16k(1-\bar{F}) - 15c},$$

apresentada no tópico anterior, teríamos chegado, praticamente, à mesma estimativa.

Consideremos, agora, uma situação em que o valor de  $\bar{F}$  não é tão baixo, digamos  $\bar{F} = 0,005$ , e que 13,5% dos pacientes com uma determinada anomalia recessiva são filhos de casais de primos em primeiro grau ( $k = 0,135$ ), enquanto na população geral a frequência desses casais é 2% ( $c = 0,02$ ). Nesse caso, o desconhecimento do coeficiente médio de endocruzamento conduziria a uma estimativa muito alta da frequência do gene, pois, por intermédio da fórmula que não leva em conta  $\bar{F}$  obteríamos  $q = 0,009$ , enquanto que pela fórmula que leva em conta o coeficiente médio de endocruzamento obteríamos  $q = 0,005$ .

O problema de estimar a proporção de filhos de casais de primos em primeiro grau entre homozigotos, quando não se conhece o coeficiente médio de endocruzamento, pode ser tratado de outra forma. Para demonstrá-la, consideremos um gene autossômico  $a$  que, em homozigose, determina uma anomalia rara. Se esse gene tiver frequência  $q$  e seu alelo  $A$  tiver a frequência  $p = 1 - q$ , poderemos escrever que a probabilidade de ocorrência de um heterozigoto  $Aa$  na população será  $2pq$  e que a probabilidade de um primo desse

heterozigoto ser, também, heterozigoto do gene raro  $a$  é  $\frac{1}{8}$ , que é o coeficiente de consangüinidade de primos em primeiro grau.

Desse modo, a probabilidade de um casal de primos em primeiro grau ser heterozigoto ( $Aa \times Aa$ ) é  $\frac{2pq}{8} = \frac{pq}{4}$  e, nesse caso, tal casal terá probabilidade igual a  $\frac{1}{4}$  de gerar um filho homozigoto  $aa$ . Portanto, a probabilidade de um casal de primos em primeiro grau dessa população ser heterozigoto do gene raro  $a$  e gerar uma criança homozigota  $aa$  é  $\frac{pq}{4} \cdot \frac{1}{4} = \frac{pq}{16}$ . Em conseqüência disso, pode-se, pois, dizer que, se a freqüência de casais de primos em primeiro grau na população for  $c$ , a freqüência esperada de homozigotos  $aa$  gerados por esses casais será  $\frac{cpq}{16}$ .

Entre os casais não consangüíneos, que aceitamos como tendo freqüência  $1 - c$ , por considerarmos que, praticamente, todos os casais consangüíneos são de primos em primeiro grau, a probabilidade de um casal ser heterozigoto ( $Aa \times Aa$ ) e ter um filho homozigoto  $aa$  é, como, se sabe,  $2pq \cdot 2pq \cdot \frac{1}{4} = p^2q^2$ . Portanto a freqüência esperada de homozigotos  $aa$  gerados por casais não consangüíneos passa, nesse caso, a ser estimada por intermédio de  $(1 - c)p^2q^2$ , enquanto que a freqüência total de homozigotos  $aa$  esperada na população passa a ser estimada por  $\frac{cpq}{16} + (1 - c)p^2q^2$ .

Com esses elementos, a fórmula da proporção  $k$  de filhos de casais de primos em primeiro grau esperada entre os homozigotos  $aa$  pode ser escrita como:

$$k = \frac{c}{c + 16pq(1 - c)}$$

porque  $k = \frac{\frac{cpq}{16}}{\frac{cpq}{16} + p^2q^2(1 - c)} = \frac{c}{c + 16pq(1 - c)}$ .

Quando conhecemos  $k$  e  $c$ , essa fórmula também pode servir para estimar a freqüência  $q$  de um gene que em homozigose determina uma anomalia autossômica, pois ela pode ser transformada na equação do segundo grau, apresentada abaixo, depois que a freqüência  $p$  é substituída por  $1 - q$ :

$$16q^2(kc - k) + 16q(k - kc) - c(1 - k) = 0$$

Usando os mesmos dados do exemplo em que  $k = 0,075$  e  $c = 0,006$  numa população em que  $\bar{F} = 0,0004$ , a raiz positiva dessa equação nos forneceria, aproximadamente,  $q = 0,005$ , isto é, a mesma estimativa que obteríamos com a fórmula que leva em conta o valor de  $\bar{F}$ . Contudo, se  $\bar{F}$  não for tão baixo, mas, digamos,  $\bar{F} = 0,005$ , como no exemplo dado anteriormente, o desconhecimento desse coeficiente conduzirá a uma super-estimativa da frequência do gene em estudo. Assim, por exemplo, no caso de  $k = 0,135$  e  $c = 0,02$ , a frequência do gene  $a$  seria estimada como  $q = 0,008$ , ao passo que, levando em conta  $\bar{F} = 0,005$ , obtemos  $q = 0,005$ .

### EQUIVALENTE LETAL

Os casamentos consangüíneos não são incluídos entre os fatores evolutivos (*mutação, seleção, deriva genética, migração*) porque eles não podem por si só contribuir para o aumento ou a perda de genes de uma população. De fato, visto que os casamentos consangüíneos apenas alteram as frequências genotípicas, mas não as gênicas, eles não podem contribuir para a *evolução* de uma espécie, pois a evolução implica, obrigatoriamente, na *substituição de um conjunto gênico de uma espécie por outro conjunto*.

É evidente, porém, que, ao aumentar a probabilidade de homozigose, os casamentos consangüíneos aumentam a probabilidade de eliminação de genes que, ao determinar doenças recessivas, ficam sujeitos à seleção mais intensa do que seus alelos, promovendo, assim, *indiretamente*, alterações das frequências gênicas. Foi o conhecimento dessa realidade que levou o Prêmio Nobel de Fisiologia e Medicina de 1946, o geneticista norte-americano Hermann Joseph Muller, a utilizar os casamentos consangüíneos para estimar, pela primeira vez, o número de genes que os seres humanos possuem em heterozigose e que, quando em homozigose, determinam a morte de um indivíduo (Muller, 1948, 1956).

A estimativa de Muller baseou-se em dados familiares coletados por Arner (1908), nos quais constatou que 83,2% dentre 672 filhos de casais de primos em primeiro grau alcançaram os 20 anos de idade, atingindo, portanto, idade reprodutiva, ao passo que uma porcentagem maior (88,4%) de 3.184 filhos de casais não-consangüíneos chegou até essa idade. Esses filhos de casais não-consangüíneos tomados para comparação eram sobrinhos

dos cônjuges consangüíneos, para que os efeitos do ambiente pudessem ser considerados como equivalentes nos dois grupos estudados.

Comparando a taxa de sobrevivência dos filhos de primos em primeiro grau com a dos filhos de não-consangüíneos, estimou Muller (1948,1956) que a viabilidade relativa dos primeiros era 94%, pois  $\frac{83,2}{88,4} = 0,94$ . Isso equivale a dizer, portanto, que 6% dos filhos de casais de primos em primeiro grau haviam falecido antes da idade reprodutiva, em decorrência do efeito de genes que se encontravam em homozigose por causa desse tipo de consangüinidade.

Por ser o coeficiente de endocruzamento ( $F$ ) dos filhos de casais de primos em primeiro grau igual a  $\frac{1}{16}$ , considerou Muller que os 6% de óbitos até os 20 anos de idade que observou entre os filhos desses casais seriam causados pela autozigose de  $\frac{1}{16}$  dos genes deletérios que ocorrem em heterozigose na população composta de indivíduos viáveis. Em outras palavras, o efeito desses genes deveria ser 16 vezes maior do que aquele que foi capaz de ser avaliado, de sorte que, se pudesse haver um coeficiente de endocruzamento  $F = 1$ , o efeito desses genes deveria corresponder a, aproximadamente, um letal por gameta, pois  $16 \times 0,06 = 0,96 \cong 1$ .

Com base nesse raciocínio, Muller concluiu que os gametas humanos conteriam aquilo que, em seguida, veio a ser chamado de *equivalente letal*, isto é, um gene com efeito letal quando em homozigose ou um conjunto de genes que podem, cada um deles, determinar a morte de uma certa proporção de homozigotos e que, se distribuídos em homozigose em diferentes indivíduos teriam o efeito global de um letal. Por outro lado, se existisse um equivalente letal por gameta, os zigotos que dão origem aos seres humanos deveriam, segundo Muller (1956), conter dois equivalentes. Como se vê, um equivalente letal tanto pode corresponder a um único gene com efeito letal em homozigose quanto, por exemplo, a 5 genes, cada qual com 20% de probabilidade de causar morte quando em homozigose, ou a 20 genes, cada qual com apenas 5% de probabilidade de causar a morte dos homozigotos, e assim por diante.

O trabalho pioneiro de Muller estimulou a realização de outro por Morton, Crow e Muller (1956), no qual foram levados em conta dados coletados na França por Sutter e

Tabah (1952,1953) sobre natimortalidade e mortalidade neonatal, infantil e juvenil dos filhos de casais de primos em primeiro grau e de casais não-consangüíneos, além dos dados de Arner (1908) e de informações ainda mais antigas (Bemiss, 1858), obtidas nos Estados Unidos. Esse trabalho de Morton, Crow e Muller (1956), no qual foi criada a designação equivalente letal, permitiu-lhes estimar que os seres humanos possuiriam, em média, de 3 a 5 equivalentes letais em heterozigose. Posteriormente, Morton (1960) criou o conceito de *equivalente detrimental*, para designar o efeito de genes deletérios que produzem anomalias, mas que não são avaliados pela mortalidade. Nesse trabalho, após analisar dados obtidos por Schull (1958) no Japão, Sutter (1958) na França, Slatk, Reis e Hoene (1958) nos Estados Unidos, Bõök (1957) na Suécia, além de dados mais antigos obtidos por Bemiss (1858) nos Estados Unidos, Morton chegou à conclusão de que nos seres humanos existem, em média, 3 equivalentes deletérios por zigoto.

### OS CASAMENTOS CONSANGÜÍNEOS E O ACONSELHAMENTO GENÉTICO

Regra geral, os noivos que são parentes consangüíneos e buscam aconselhamento genético não estão interessados em saber o risco que correm de ter um filho que manifeste uma heredopatia específica, pois, na maioria das vezes, desconhecem a ocorrência de doenças genéticas em seus ancestrais ou colaterais. O que eles desejam saber é se, de fato, correm maior risco de gerar crianças com anomalias genéticas ou com menor desenvolvimento físico e mental do que os filhos de casais não-consangüíneos. Como dar a esses casais uma resposta razoável, se não conhecemos as frequências gênicas da maioria das doenças recessivas nem em nossa nem em outras populações humanas? Por outro lado, como levar em conta a grande variabilidade da gravidade das manifestações clínicas dessas heredopatias?

Uma solução para esse problema pode ser buscada nos trabalhos de Morton, Crow e Muller (1956) e Morton (1960), segundo os quais existe uma relação exponencial entre o efeito do endocruzamento ( $M$ ) e o coeficiente de endocruzamento ( $F$ ), a qual pode ser expressa por:

$$M = 1 - e^{-(A+BF)}$$

Nessa equação  $M$  indica a taxa de mortalidade precoce e morbidade,  $e$  é a base dos logaritmos neperianos,  $A$  é a estimativa do dano global, isto é, do dano genético e não-genético, sofrido pelos filhos de não-consangüíneos, isto é, com  $F = 0$ ,  $B$  é a estimativa do

número médio de equivalentes letais e detrimenais por gameta que se expressam nos filhos de consangüíneos e  $F$  é o coeficiente de endocruzamento. Nas populações com alto nível de vida, nas quais a taxa de mortalidade infantil e juvenil é muito baixa, pode-se aceitar que  $A$  seja a estimativa do dano genético sofrido pelos filhos de não-consangüíneos.

Tendo em mente que  $A + BF$  é uma regressão, na qual  $A$  é o intercepto e  $B$  é o coeficiente de regressão, pode-se estimar, para diferentes valores de  $F$ , o aumento do risco de mortalidade precoce e morbidade ao qual estão sujeitos os filhos de consangüíneos, em consequência de homozigose de genes deletérios. Para isso basta obter a diferença entre  $1 - e^{-(A+BF)}$  e  $1 - e^{-A}$ .

Tomemos um exemplo numérico para ilustrar essa afirmação, considerando os dados de Freire-Maia (1984) sobre a população brasileira de classe média e alta, para a qual estimou  $B = 2$ , pois aceitou a existência de 1,5 equivalente letal e 0,5 equivalente detrimenais por gameta, e  $A = 0,22$ . Com base nesses dados podemos informar aos casais de primos em primeiro grau ( $F = \frac{1}{16}$ ) que seus filhos correm um risco de mortalidade precoce e morbidade 9% mais alto do que o dos filhos de não-consangüíneos, porque calculamos:

$$1 - e^{-(0,22 + 2 \cdot \frac{1}{16})} - (1 - e^{-0,22}) = 0,2918 - 0,1975 = 0,0943 \cong 9\%.$$

Aos casais formados por primos duplos em primeiro grau ou por tio(a)-sobrinha(o), isto é, com  $F = \frac{1}{8}$ , podemos informar, com base nesses mesmos dados, que seus filhos correm um risco de mortalidade precoce e morbidade 18% mais alto do que o dos filhos de não-consangüíneos, porque, nesse caso, calculamos

$$(1 - e^{-(0,22 + 2 \cdot \frac{1}{8})}) - (1 - e^{-0,22}) = 0,3750 - 0,1975 = 0,1775 \cong 18\%.$$

À medida que o coeficiente de endocruzamento diminui, cai bastante a diferença entre os riscos de mortalidade e morbidade em filhos de consangüíneos e de não-consangüíneos, decorrentes da expressão de genes recessivos. Assim, com base nos dados de Freire-Maia (1984), esse risco seria menos do que 3% mais alto para os filhos de primos em terceiro grau do que para filhos de não-consangüíneos.

De acordo com Morton (1978), no aconselhamento genético dado a casais consangüíneos, deve-se levar em conta, ainda, que os filhos desses casais estão sujeitos a risco mais alto de manifestar deficiência mental do que os filhos de não-consangüíneos.

Segundo ele, esse risco para os filhos de casais normais não-consangüíneos é de 1,2%, enquanto que para os filhos de primos em primeiro grau, também normais, é de 6,2 %.

### O EFEITO WAHLUND

O termo isolado é comumente empregado em Genética para definir um conjunto humano que está separado dos outros por alguma barreira, seja ela geográfica, política, sócio-econômica, religiosa ou cultural, que impede ou dificulta a troca de genes com outro conjunto, por intermédio do casamento de seus elementos. Nessa definição está implícito, pois, que, na espécie humana, também existem mecanismos de isolamento que impedem o intercâmbio de genes entre populações que vivem na mesma área geográfica.

Os efeitos genéticos resultantes do isolamento dependem do tamanho dos isolados. Se uma população for dividida em isolados grandes, o resultado desse isolamento (*efeito Wahlund*) será semelhante ao dos casamentos consangüíneos, isto é, haverá aumento da frequência de homozigotos na população (Wahlund, 1928). Em outras palavras, quando a população é dividida em grandes isolados, o isolamento não terá efeito evolutivo, pois somente haverá alteração das frequências genotípicas, mas não das frequências gênicas.

Para demonstrar o efeito Wahlund consideremos uma população na qual os alelos autossômicos  $A, a$  tenham frequências  $p_o$  e  $q_o$ , respectivamente, e que essa população tenha sido dividida em  $n$  isolados grandes e panmícticos, de sorte que, em cada um deles, os genótipos  $AA$ ,  $Aa$  e  $aa$  se distribuam de acordo com a lei de Hardy e Weinberg. Suponhamos, ainda, que, nesses  $n$  isolados os alelos  $A$  e  $a$  tenham frequências diferentes, designadas por  $p_1, p_2, p_3, \dots, p_n$  e  $q_1, q_2, q_3, \dots, q_n$ .

Evidentemente, as frequências médias  $\bar{p}$  e  $\bar{q}$  dos alelos  $A$  e  $a$  nos  $n$  isolados serão iguais às frequências gênicas da população antes de sua divisão em subpopulações. Assim, se os  $n$  isolados tiverem o mesmo tamanho, poderemos escrever que as frequências médias dos alelos  $A$  e  $a$  serão

$$\bar{p} = p_o = \frac{\sum p}{n} \quad \text{e} \quad \bar{q} = q_o = \frac{\sum q}{n} \quad \text{ou} \quad \bar{q} = q_o = 1 - \bar{p}$$

Por outro lado, a variância de  $\bar{q}$ , que é igual a variância de  $\bar{p}$ , será obtida por intermédio de  $\frac{\sum q^2}{n} - \bar{q}^2$  porque:

$$\frac{\Sigma(q - \bar{q})^2}{n} = \frac{\Sigma q^2 - \frac{(\Sigma q)^2}{n}}{n} = \frac{\Sigma q^2}{n} - \left(\frac{\Sigma q}{n}\right)^2 = \frac{\Sigma q^2}{n} - \bar{q}^2$$

No caso de os isolados terem tamanhos diferentes e  $x_1, x_2, x_3, \dots, x_n$  representarem as suas proporções, as frequências médias dos alelos  $A$  e  $a$  serão médias ponderadas, isto é, serão obtidas a partir de

$$\bar{p} = p_o = \Sigma xp \quad \text{e} \quad \bar{q} = q_o = \Sigma xq \quad \text{ou} \quad \bar{q} = q_o = 1 - \bar{p}$$

A variância de  $\bar{q}$  ou de  $\bar{p}$  será obtida por meio de  $\sigma^2 = \Sigma xq^2 - \bar{q}^2$ .

Na população dividida em  $n$  isolados grandes, a distribuição dos genótipos  $AA$ ,  $Aa$  e  $aa$  não se fará, entretanto, de acordo com  $\bar{p}^2$ ,  $2\bar{p}\bar{q}$  e  $\bar{q}^2$ , isto é, segundo  $p_o^2$ ,  $2p_oq_o$ ,  $q_o^2$ . Tal distribuição somente voltará a ser atingida quando houver miscigenação completa entre os isolados, ou seja, quando houver quebra de todos os isolados com panmixia. Enquanto a população estiver dividida em grandes isolados, sejam eles de mesmo tamanho ou não, as frequências dos genótipos  $AA$ ,  $Aa$  e  $aa$  serão as médias das frequências desses genótipos nos  $n$  isolados. Tais médias corresponderão à seguinte distribuição, no caso de os isolados terem tamanho igual:

$$\begin{aligned} AA &= \frac{\Sigma p^2}{n} = \bar{p}^2 + \sigma^2 \\ Aa &= \frac{\Sigma 2pq}{n} = 2\bar{p}\bar{q} - 2\sigma^2 \\ aa &= \frac{\Sigma q^2}{n} = \bar{q}^2 + \sigma^2 \end{aligned}$$

Ao mesmo tipo de distribuição se chega quando os isolados têm tamanhos diferentes, como se vê abaixo:

$$\begin{aligned} AA &= \Sigma xp^2 = \bar{p}^2 + \sigma^2 \\ Aa &= \Sigma x2pq = 2\bar{p}\bar{q} - 2\sigma^2 \\ aa &= \Sigma xq^2 = \bar{q}^2 + \sigma^2 \end{aligned}$$

Como se vê, nas populações subdivididas em grandes isolados com frequências gênicas diferentes, os genótipos homozigotos passam a ter frequência mais alta do que a encontrada na população sem subdivisões. Obviamente, quando ocorre o oposto, isto é, quando os isolados se unem até constituir uma única população panmíctica, a frequência dos homozigotos decresce.

Os exemplos numéricos da Tabela 4.5 servem bem para ilustrar o que foi mencionado até agora. Na Tabela 4.5- A temos uma população subdividida em quatro grandes isolados de mesmo tamanho, nos quais as freqüências dos alelos autossômicos  $A, a$  mostram grande variação. Nos dados dessa tabela, as freqüências médias desses alelos ( $\bar{p}$  e  $\bar{q}$ ), que são, também, é claro, as freqüências desses mesmos alelos antes de a população ser dividida em quatro grandes isolados ( $p_o$  e  $q_o$ ), são  $\bar{p} = p_o = 0,6$  e  $\bar{q} = q_o = 0,4$ , mas as freqüências médias dos genótipos não se distribuem segundo

$$AA = \bar{p}^2 = 0,36 \quad Aa = 2\bar{p}\bar{q} = 0,48 \quad aa = \bar{q}^2 = 0,16$$

mas, segundo:

$$AA = \frac{\Sigma p^2}{n} = 0,41 \quad Aa = \frac{\Sigma 2pq}{n} = 0,38 \quad aa = \frac{\Sigma q^2}{n} = 0,21$$

A diferença entre as freqüências genótípicas nessas duas distribuições é, exatamente, a variância, no caso de cada genótipo homozigoto (0,05), e o dobro da variância no caso do genótipo heterozigoto (0,10). Pode-se, pois, escrever que a população de nosso exemplo, quando dividida em quatro grandes isolados de mesmo tamanho, passará a ter a distribuição genotípica  $AA = 0,36 + 0,05 = 0,41$ ;  $Aa = 0,48 - 0,10 = 0,38$ ;  $aa = 0,16 + 0,05 = 0,21$ , mas as freqüências gênicas iniciais não sofrerão qualquer alteração, porque  $A = 0,41 + 0,19 = 0,60$  e  $a = 0,21 + 0,19 = 0,40$ .

Na Tabela 4.5.B a população na qual os alelos  $A, a$  mostravam, inicialmente, as freqüências  $p_o = 0,6$  e  $q_o = 0,4$ , também foi subdividida em quatro isolados grandes, com a mesma variação das freqüências gênicas observada na Tabela 4.5.A. Na Tabela 4.5.B, entretanto, os isolados têm tamanhos diferentes, correspondendo, os de números 1 e 4, cada qual, a 30% da população total ( $x = 0,3$ ), e os de números 2 e 3, cada qual, a 20% da população total ( $x = 0,2$ ). O efeito final é, contudo, semelhante ao observado na Tabela 4.5.A, porque as diferenças entre os tamanhos dos isolados não foram muito acentuadas.

Tendo em mente que, nas populações subdivididas em grandes isolados, o aumento da freqüência de homozigotos depende do valor da variância das freqüências gênicas, está claro que, nessas populações, tal aumento será tanto mais pronunciado quanto maior for a variação da freqüência gênica dos isolados, pois é essa variabilidade que determina o valor da variância. A Tabela 4.5.C ilustra isso de modo claro. De fato, apesar de essa tabela referir-se a uma população com as freqüências gênicas iniciais idênticas às da Tabela 4.5.A e de ela também ter sido subdividida em quatro isolados de mesmo tamanho, as freqüências

de homozigotos na população da Tabela 4.5.C são menores do que as da Tabela 4.5.A, porque as frequências gênicas variaram menos nos isolados da população da Tabela 4.5.C.

Tabela 4.5. Exemplos numéricos demonstrativos de que a subdivisão de uma população aumenta a frequência de homozigotos.

**A. Os isolados têm o mesmo tamanho.**

Isolado	$p$	$q$	$p^2$	$2pq$	$q^2$
1	0,9	0,1	0,81	0,18	0,01
2	0,7	0,3	0,49	0,42	0,09
3	0,5	0,5	0,25	0,50	0,25
4	0,3	0,7	0,09	0,42	0,49
População dividida	0,6	0,4	0,41	0,38	0,21
População inicial	0,6	0,4	0,36	0,48	0,16
Diferença			0,05 ( $\sigma^2$ )	-0,10 ( $-2 \sigma^2$ )	0,05 ( $\sigma^2$ )

**B. Os isolados têm tamanho diferente, mas as frequências gênicas variam como em A.**

Isolado	Proporção (x)	$P$	$q$	$P^2$	$2pq$	$q^2$
1	0,3	0,9	0,1	0,81	0,18	0,01
2	0,2	0,7	0,3	0,49	0,42	0,09
3	0,2	0,5	0,5	0,25	0,50	0,25
4	0,3	0,3	0,7	0,09	0,42	0,49
População dividida		0,6	0,4	0,418	0,364	0,218
População inicial		0,6	0,4	0,360	0,480	0,160
Diferença				0,058 ( $\sigma^2$ )	-0,116 ( $-2 \sigma^2$ )	0,058 ( $\sigma^2$ )

**C. Os isolados têm o mesmo tamanho, mas as frequências gênicas variam menos do que em A e B.**

Isolado	$p$	$Q$	$p^2$	$2pq$	$q^2$
1	0,8	0,2	0,64	0,32	0,04
2	0,7	0,3	0,49	0,42	0,09
3	0,5	0,5	0,25	0,50	0,25
4	0,4	0,6	0,16	0,48	0,36
População dividida	0,6	0,4	0,385	0,430	0,185
População inicial	0,6	0,4	0,360	0,480	0,160
Diferença			0,025 ( $\sigma^2$ )	-0,050 ( $-2 \sigma^2$ )	0,025 ( $\sigma^2$ )

Se compararmos as proporções genóticas resultantes do efeito Wahlund com aquelas que são observadas em populações que estão em equilíbrio de Wright,

concluiremos logo que é possível comparar o acréscimo de homozigotos que resultam desses dois processos em relação às frequências esperadas em equilíbrio de Hardy e Weinberg. Em outras palavras, pode-se comparar a variância resultante do efeito Wahlund com  $Fpq$ , que é o acréscimo de  $p^2$  e de  $q^2$ , o que permitirá avaliar a qual valor de  $F$  corresponde o aumento da frequência de homozigotos por efeito do isolamento.

Tomemos para exemplo os dados da Tabela 4.5.A. Se o acréscimo na frequência de cada genótipo homozigoto ( $\sigma^2$ ) fosse igual ao acréscimo  $Fpq$ , observado nas populações em equilíbrio de Wright, teríamos  $\sigma^2 = Fpq$ . Nos dados da Tabela 4.5.A tem-se  $\sigma^2 = 0,05$ , enquanto que  $p$  e  $q$  correspondem, respectivamente, a  $p_o = \bar{p} = 0,6$  e  $q_o = \bar{q} = 0,4$  pois são as frequências dos alelos autossômicos  $A,a$  na população sem subdivisão em grandes isolados. Em vista disso, extrairíamos o valor de  $F$  a partir da igualdade  $\sigma^2 = Fpq$ , obtendo:

$$F = \frac{\sigma^2}{pq} = \frac{0,05}{0,6 \times 0,4} = 0,21$$

isto é, um valor próximo de 0,25 ou  $\frac{1}{4}$ .

Lembrando que  $\frac{1}{4}$  é o coeficiente de endocruzamento de filhos de irmãos, ou de pais e filhas ou de mães e filhos, podemos, pois, dizer que o aumento de homozigotos por efeito do isolamento na população da Tabela 4.5.A corresponde àquele que seria conseguido se, praticamente, todos os casamentos da população fossem incestuosos. Como se vê, a subdivisão da população em grandes isolados com frequências gênicas diferentes é mais eficiente do que os casamentos consangüíneos para aumentar as frequências dos genótipos homozigotos. Essa eficiência será tanto maior quanto maior for a variância da frequência dos genes.

**QUESTÕES E RESPOSTAS**

**Q 1.** O pai de João é irmão da mãe de Maria e a mãe de João é irmã do pai de Maria. João e Maria querem casar. Considerando que eles são normais e sem antecedentes dignos de nota, e supondo que João é heterozigoto de cinco genes autossômicos muito raros com efeito deletério em homozigose, qual o risco que João e Maria correm de gerar uma criança com pelo menos uma das cinco anomalias determinadas pelos genes autossômicos deletérios?

**R 1.** O risco é de 27,58% porque  $1 - \left(\frac{15}{16}\right)^5 = 0,2758$ .

**Q 2.** João e Maria são casados. A mãe de João e a mãe de Maria são irmãs. A avó materna de João e Maria é filha de primos em primeiro grau e o avô materno de João e Maria é filho de primos duplos em primeiro grau. Qual a probabilidade de João e Maria gerarem uma criança autozigota de um gene autossômico?

**R 2.** A probabilidade é estimada em 6,8% porque  $\left(\frac{1}{5}\right)^5 \left(1 + \frac{1}{8}\right) + \left(\frac{1}{2}\right)^5 \left(1 + \frac{1}{16}\right) = 0,068$ .

**Q 3.** O pai de João tem um irmão gêmeo monozigótico, que é casado com a irmã da mãe de João, a qual, por sua vez, é gêmea monozigótica de sua mãe. João casou com sua prima Maria, que é filha do irmão gêmeo de seu pai com a irmã gêmea de sua mãe. A que equivale o parentesco genético entre João e Maria?

**R 3.** Ao parentesco entre irmãos.

**Q 4.** Se na questão anterior 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?

**R 4.** O risco é de  $\frac{1}{8}$ , pois  $\frac{1}{2} \cdot \frac{1}{4} = \frac{1}{8}$ .

**Q 5.** Se na questão 3 o pai e o tio de João fossem gêmeos dizigóticos, mas a mãe e a tia de João fossem monozigóticas, a que equivaleria o parentesco genético entre João e Maria?

**R 5.** A um parentesco maior do que o existente entre meio-irmãos comuns, pois  $r = \frac{3}{8}$ .

**Q 6.** Se na questão 3 tanto o pai e o tio de João, quanto a sua mãe e sua tia fossem gêmeos dizigóticos, a que equivaleria o parentesco genético entre João e Maria?

**R 6.** Ao parentesco entre primos duplos em primeiro grau.

**Q 7.** Em amostras de 1.000 casamentos ocorridos durante o mesmo período em duas populações (A e B) foram constatadas as distribuições abaixo. Quais os coeficientes médios de endocruzamento dessas duas populações?

Casais	População A	População B
Não-consangüíneos	970	967
Primos em 3 <sup>o</sup> . grau	-	2
Primos em 2 <sup>o</sup> . grau	4	2
Primos em 1 <sup>o</sup> . grau	15	17
Primos duplos em 1 <sup>o</sup> . grau	6	8
Tio e meia-sobrinha	2	-
Tia e meio-sobrinho	1	2
Tio e sobrinha	1	2
Tia e sobrinho	1	-

**R 7.**  $\bar{F} = 0,0022$  na população A e  $\bar{F} = 0,0025$  na população B.

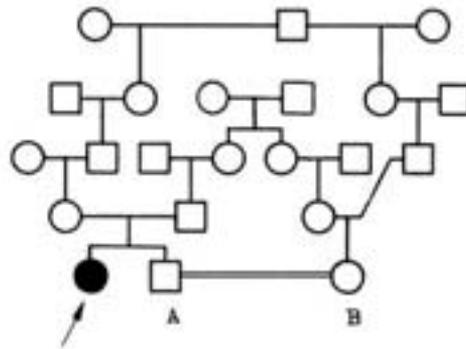
**Q 8.** Se a proporção de casais de primos em primeiro grau entre os genitores de pacientes com fenilcetonúria for 5% na população A e 15% na população B, em qual população devemos supor que o gene da fenilcetonúria é mais freqüente?

**R 8.** Na população A, porque as populações A e B têm coeficiente médio de endocruzamento semelhante e freqüência de casais de primos em primeiro grau semelhante (1,5% na A e 1,7% na B). No entanto, a relação  $\frac{k}{c}$  é muito mais alta na população B

( $\frac{5}{1,5} = 3,3$  na população A e  $\frac{15}{1,7} = 8,8$  na população B).

**Q 9.** No heredograma abaixo a propósita manifesta síndrome adrenogenital por deficiência de 21-hidroxilase. Qual a probabilidade de:

- o indivíduo A da genealogia representada por esse heredograma ser heterozigoto do gene dessa deficiência enzimática?
- o indivíduo B da genealogia representada por esse heredograma ser heterozigoto do gene dessa deficiência enzimática, por tê-lo herdado de um ancestral comum aos indivíduos A e B?
- o casal A x B gerar uma filha com síndrome adrenogenital ou um filho com macrogenitossomia em decorrência da homozigose do gene da deficiência de 21-hidroxilase?



**R 9.** a)  $\frac{2}{3}$ ; b)  $\frac{5}{64}$ ; c)  $\frac{10}{78}$  ou 1,3%.

**Q 10.** Se o casal A x B da questão anterior já tivesse gerado uma menina com síndrome adrenogenital, qual seria o risco de ele gerar outra criança com deficiência de 21-hidroxilase?

**R 10.** O risco seria de 25%.

**Q 11.** Um casal que tem ancestrais normais gerou quatro filhos, três dos quais normais e um com a doença de Tay-Sachs. A irmã do marido quer casar com o irmão da mulher e deseja saber qual o risco que correm de gerar uma criança com essa doença. Antes de fazer a pesquisa para verificar se a irmã do marido e o irmão da mulher são heterozigotos do gene que, em homozigose, causa a doença de Tay-Sachs, qual é esse risco? Depois de fazer essa pesquisa e constatar que os dois eventuais cônjuges são heterozigotos do gene em questão, qual é esse risco?

**R 11.** Antes da pesquisa é 6,25%, pois  $\frac{1}{2} \cdot \frac{1}{2} \cdot \frac{1}{4} = \frac{1}{16}$  ou 0,0625. Depois da pesquisa é  $\frac{1}{4}$ , pois teremos a certeza de que ambos são heterozigotos.

**Q 12.** Se na questão anterior a pesquisa para o rastreamento de heterozigotos da doença de Tay-Sachs, por intermédio da determinação do nível de hexosaminidase A nos leucócitos, tivesse demonstrado que apenas a irmã do marido é heterozigota, qual seria o risco de o futuro casal gerar uma criança com a doença em discussão?

**R 12.** Risco nulo.

**Q 13.** Dentre os casais consanguíneos abaixo, qual aquele que oferece maior risco de gerar filhos com anomalias recessivas, supondo que todos esses casais são clinicamente normais e sem antecedentes familiares dignos de nota?

Primos em segundo grau.

Primos em primeiro grau.

Tio × meia-sobrinha.

Tio × sobrinha.

**R 13.** Tio × sobrinha porque seu coeficiente de consangüinidade é  $r = \frac{1}{4}$ .

**Q 14.** A incidência de uma anomalia autossômica recessiva monogênica em uma população com coeficiente médio de endocruzamento tão pequeno que pode ser desprezado é estimada em 1:19.600. Qual o risco de um casal normal dessa população, sem antecedentes familiares dessa anomalia, gerar uma criança que a manifeste, se ele for constituído por primos em primeiro grau?

**R 14.** O risco é de 0,05% porque  $\frac{1}{19.600} + \frac{1}{16} \cdot \frac{139}{140} \cdot \frac{1}{140} = 0,0005$ .

**Q 15.** Qual o aumento relativo do risco de ocorrência dessa anomalia entre os filhos de casais de primos em primeiro grau da população da questão anterior?

**R 15.** 869% porque  $q = \frac{1}{19.600} = 0,000051$  e  $Fpq/q^2 = \frac{0,000443}{0,000051} = 8,69$ .

**Q 16.** Em duas populações, A e B, a incidência de fenilcetonúria é estimada em 1: 12.100. Na população A o coeficiente médio de endocruzamento é estimado em 0,002, enquanto que na população B ele é muito alto, sendo estimado em 0,01. Qual a freqüência do gene da fenilcetonúria na população A e na população B?

**R 16.** Na população A a freqüência do gene é 0,008, enquanto que na população B essa freqüência é 0,005. Tais freqüências são obtidas por intermédio do cálculo da raiz positiva de  $q$  nas seguintes equações:

$$0,002q + 0,998q^2 = \frac{1}{12.100}$$

$$0,01q + 0,99q^2 = \frac{1}{12.100}$$

**Q 17.** Em uma população os alelos autossômicos  $A, a$  têm freqüências iguais a  $p = 0,9$  e  $q = 0,1$ , respectivamente. Quais as freqüências esperadas dos genótipos  $AA$ ,  $Aa$  e  $aa$  nessa população no caso de: a) panmixia?

b) fixação completa?

c)  $\bar{F} = 0,02$ ?

R 17.

Genótipos	a)	b)	c)
AA	0,81	0,9	0,8118
Aa	0,18	-	0,1764
aa	0,01	0,1	0,0118

**Q 18.** A incidência da doença de Niemann-Pick em duas cidades (A e B) foi estimada em 1:40.000. O coeficiente médio de endocruzamento das populações dessas cidades é, praticamente, igual a  $\bar{F} = 0,004$ , mas a porcentagem de casais de primos em primeiro grau é 1,5% na cidade A e 3% na cidade B. Qual a proporção esperada de filhos de primos em primeiro grau entre os pacientes com a doença recessiva autossômica em discussão nessas duas cidades?

**R 18.** Na cidade A 14,02% e na cidade B 28,04%, porque a frequência  $q$  do gene responsável pela doença é estimada em 0,003, por intermédio da resolução de  $0,004q + 0,996q^2 = \frac{1}{40.000}$ . Com isso, obtém-se:

$$k = \frac{0,015(1+15 \times 0,003)}{16(0,004 + 0,996 \times 0,003)} = 0,1402 \quad \text{e} \quad k = \frac{0,03(1+15 \times 0,003)}{16(0,004 + 0,996 \times 0,003)} = 0,2804$$

**Q 19.** Se na questão anterior o coeficiente médio de endocruzamento não tivesse sido levado em conta, quais seriam as estimativas das proporções esperadas de filhos de primos em primeiro grau entre os pacientes com a doença de Niemann-Pick?

**R 19.** Na cidade A 16,99% e na cidade B 33,97%, porque calcularíamos:

$$q = \sqrt{\frac{1}{40.000}} = \frac{1}{200} \text{ ou } 0,005$$

$$k = \frac{0,015(1+15 \times 0,005)}{(0,015 \times 0,995) + (16 \times 0,005)} = 0,1699 \quad \text{e} \quad k = \frac{0,03(1+15 \times 0,005)}{(0,015 \times 0,995) + (16 \times 0,005)} = 0,3397$$

**Q 20.** A incidência de anemia falciforme numa população afro-descendente foi estimada em

1,6: 1.000. Visto que nessa população a frequência de casais de primos em primeiro grau é estimada em 1%, quer-se saber qual a proporção esperada de filhos de primos em primeiro grau entre as crianças com anemia falciforme:

a) levando em conta o coeficiente médio de endocruzamento dessa população, que é  $\bar{F} = 0,0035$ .

b) sem levar em conta o coeficiente médio de endocruzamento.

**R 20.** a)  $k = 0,0237$  ou 2,37%.

b)  $k = 0,0246$  ou 2,46%.

**Q 21.** Em uma cidade, na qual a freqüência de casamentos entre primos em primeiro grau é igual, praticamente, a 2%, e na qual o coeficiente médio de endocruzamento é estimado em  $\bar{F} = 0,005$ , verificou-se que 10% dos pacientes fenilcetonúricos eram filhos de primos em primeiro grau. Qual a freqüência do gene da fenilcetonúria que podemos estimar para essa cidade?

**R 21.** 0,9% porque  $\frac{0,02-16 \times 0,10 \times 0,005}{(16 \times 0,10) - (15 \times 0,02)} = 0,009$ .

**Q 22.** Um pediatra estudou uma anomalia que se manifesta na infância e que ocorre na população com uma freqüência de, praticamente, 1: 19.600. Essa anomalia tem transmissão hereditária recessiva autossômica monogênica e é considerada como uma entidade genético-clínica única. Tendo em vista que, na casuística estudada pelo pediatra, 25 % dos pacientes eram filhos de primos em primeiro grau, enquanto que na população da qual eles procedem, e que tem coeficiente médio de endocruzamento  $\bar{F} = 0,0015$ , a freqüência desses casais é 1,5%, pergunta-se:

a) Qual a freqüência esperada do gene determinador da anomalia em questão, no caso de ela constituir, de fato, uma entidade genético-clínica?

b) Os dados do pediatra permitem aceitar a hipótese de que a anomalia por ele estudada é uma entidade genético-clínica única?

**R 22.** a) 0,006, valor obtido a partir de  $0,0015q + 0,9985q^2 = \frac{1}{19.600}$ ;

b) Não, porque se a anomalia fosse decorrente de um único gene em homozigose a proporção esperada de filhos de primos em primeiro grau entre os pacientes seria, praticamente, a metade da porcentagem observada, pois teríamos  $k = 0,1364$  ou 13,64%.

**Q 23.** Numa grande cidade não se conhece a incidência da fenilcetonúria. Sabe-se, contudo, que 10% dos pacientes são filhos de casais de primos em primeiro grau, enquanto que na população a taxa desses casamentos é de 1,5%. Admitindo-se que nas grandes cidades o

coeficiente médio de endocruzamento é baixo, qual a estimativa da frequência do gene da fenilcetonúria nessa cidade?

**R 23.** 1 %, porque  $q = \frac{0,015(1-0,10)}{16 \times 0,10 - (15 + 0,10)0,015} = 0,0098 \cong 0,01$ .

**Q 24.** Se tivéssemos constatado que, dentre 1.250 filhos de primos duplos em primeiro grau e de casais formados por tio(a) e sobrinha(o), 338 (27%) faleceram antes de atingir a idade reprodutiva, enquanto que, dentre 2.500 filhos de não-consangüíneos, 250 (10%) foram a óbito antes de atingir essa idade, qual o número médio de equivalentes letais por indivíduo que poderíamos estimar com base nesses dados?

**R 24.** Três equivalentes letais por indivíduo, porque a viabilidade relativa dos filhos desses casais consangüíneos seria estimada em 81%, pois,  $\frac{73}{90} = 0,81$ . Visto que tais indivíduos possuem  $F = \frac{1}{8}$  e sendo  $1 - 0,81 = 0,19$ , tem-se que o número de equivalentes letais por gameta é estimado em  $8 \times 0,19 = 1,52 \cong 1,5$ , de sorte que o número médio desses equivalentes por zigoto é estimado em 3.

**Q 25.** Uma população de uma certa área geográfica está subdividida em três isolados genéticos (A, B e C), constituídos, respectivamente, por 1.200, 1.800 e 3.000 indivíduos em idade reprodutiva. Investigando a distribuição dos grupos sangüíneos M, MN e N nesses três isolados foi possível estimar as frequências dos genes *M* e *N* em  $M = 0,70$  e  $N = 0,30$  no isolado A;  $M = 0,60$  e  $N = 0,40$  no isolado B, e  $M = 0,50$  e  $N = 0,50$  no isolado C. Calcular as estimativas das frequências dos:

- genes *M* e *N* na população total da área geográfica estudada;
- genótipos *MM*, *MN* e *NN* na população total da área geográfica estudada;
- genótipos *MM*, *MN* e *NN* na população da área geográfica estudada, na hipótese de ocorrência de uma quebra completa dos isolados, com panmixia.

**R 25.** Visto que o isolado A constitui 20% da população da área estudada e que os isolados B e C compõem, respectivamente, 30% e 50% dessa população, conclui-se:

a)  $M = p = (0,20 \times 0,70) + (0,30 \times 0,60) + (0,50 \times 0,50) = 0,57$   
 $N = q = 1 - 0,57 = 0,43$

b)  $MM = (0,20 \times 0,49) + (0,30 \times 0,36) + (0,50 \times 0,25) = 0,331$   
 $MN = (0,20 \times 0,42) + (0,30 \times 0,48) + (0,50 \times 0,50) = 0,478$

$$NN = (0,20 \times 0,09) + (0,30 \times 0,16) + (0,50 \times 0,25) = 0,191$$

c)  $MM = (0,57)^2 = 0,325$

$$MN = 2 \times 0,57 \times 0,43 = 0,490$$

$$NN = (0,43)^2 = 0,185$$

**Q 26.** Verificar se, nos isolados da questão anterior, o aumento de homozigotos por efeito do isolamento pode ser considerado igual, superior ou inferior ao que seria conseguido se todos os casamentos tivessem ocorrido entre primos em terceiro grau.

**R 26.** Tendo em mente que as diferenças abaixo indicam que a variância da frequência do

gene  $M$ , ou do gene  $N$  é 0,006 conclui-se que  $F = \frac{\sigma^2}{pq} = \frac{0,06}{0,57 \times 0,43} = 0,0245$ .

MM	MN	NN
0,331	0,478	0,191
<u>0,325</u>	<u>0,490</u>	<u>0,185</u>
0,006	-0,012	0,006

Lembrando que  $F = \frac{1}{64} = 0,0156$  é o coeficiente de endocruzamento de filhos de primos em terceiro grau, conclui-se, também, que o aumento de homozigotos por isolamento é maior do que aquele que seria obtido se todos os casais fossem primos em terceiro grau.

## REFERÊNCIAS

- Agostini, J.M. & Meirelles-Nasser, C. Consanguineous marriages in the archdiocese of Florianópolis, South Brazil. *Rev. Brasil. Genet.* 9: 479-486, 1986.
- Almström, C.H. First-cousin marriages in Sweden 1750-1844 and a study of the population movement in some Swedish subpopulations from the genetic-statistical viewpoint. *Acta Genet.* 8: 295-369, 1958.
- Arner, G.B.L. Consanguineous marriages in the American population. *Studies in History, Economics and Public Law, Columbia University*, 31, n°. 3: 1-100, 1908, cf. Muller, H.J., 1956, *op cit.*
- Bassi, R.A. & Freire-Maia, N. Marriage age and inbreeding in Curitiba, Southern Brazil. *Rev. Brasil. Genet.* 8: 199-203, 1985.
- Bemiss, S.M. Report on influence of marriages of consanguinity upon offspring. *Tr. Am. Med. Ass.* 11: 319-425, 1858, cf. Morton, N.E., 1960, *op. cit.*
- Böök, J.A. Genetical investigations in a North Swedish population. The offspring of first cousin marriages. *Ann. Hum. Genet.* 21: 191-221, 1957.
- Dahlberg, G. Inbreeding in man *Genetics* 14: 421-454, 1929.

- Dahlberg, G. On rare defects in human populations with particular regard to inbreeding and isolate effects. *Proc. Roy. Soc. Edinburgh*, 58, Part II: 213-232, 1938.
- Dahlberg, G. *Mathematical methods for Population Genetics*. Interscience Publ., London, N. York, 1948.
- Freire-Maia, N. Consanguineous marriages in Brazil. I -Structure of such marriages. II. Factors of geographic distribution. *Eugen. Quart.* 5: 105-114, 1958.
- Freire-Maia, N. *Genética de populações humanas*. HUCITEC, Editora da USP, S. Paulo, 1974.
- Freire-Maia, N. Effects of consanguineous marriages on morbidity and precocious mortality: genetic counseling. *Am. J. Med. Genet.* 18: 401-406, 1984.
- Freire-Maia, N. Os casamentos consanguíneos e o novo Código de Direito Canônico. *Rev. Brasil. Genet.* 9: 565-567, 1986.
- Freire-Maia, N. & Freire-Maia, A. The structure of consanguineous marriages and its genetic implications. *Ann. Hum. Genet.* 25: 29-39, 1961.
- Freire-Maia, N. & Elisbão, T. Inbreeding effect on morbidity: III. A review of the world literature. *Am. J. Med. Genet.* 18: 391-400, 1984.
- Hajnal, J. Concepts of random mating and the frequency of consanguineous marriages. *Proc. Roy. Soc. B.* 159: 125-177, 1963.
- Haldane, J.B.S. & Moshinsky, P. Inbreeding in mendelian populations with special reference to human cousin marriage. *Ann. Eugen.* 9: 321-340, 1939.
- Malécot, G. *Les mathématiques de l'hérédité*, Masson & Cie., Paris, 1948.
- Morton, N.E. Non-randomness in consanguineous marriage. *Ann. Hum. Genet.* 20: 116-124, 1955.
- Morton, N.E. The mutational load due to detrimental genes in man *Am. J. Hum. Genet.* 12: 348-364, 1960.
- Morton, N.E. Effect of inbreeding on IQ and mental retardation. *Proc. Natl. Acad. Sci. USA*, 75: 3906-3908, 1978.
- Morton, N.E., Crow, J.F. & Muller, H.J. An estimate of mutational damage in man from data on consanguineous marriages. *Proc. Natl. Acad. Sci. USA* 42: 855-863, 1956.
- Muller, H.J. *Mutational prophylaxis*. *Bull. N.Y. Acad. Med.* 24: 447-469, 1948.
- Muller, H.J. Further studies bearing on the load of mutations in man. *Acta Genet.* 6: 157-168, 1956.
- Rothhammer, F. & Cruz-Coke, R. *Curso Básico de Genética Humana*. Editorial Universitária, Santiago Chile, 1977.
- Schull, W.J. Empirical risks in consanguineous marriages: sex ratio, malformations, and variability. *Am. J. Hum. Genet.* 10: 294-349, 1958.
- Statis, H.M., Reis, R.H. & Hoene, R.E. Consanguineous marriages in the Chicago region. *Am. J. Hum. Genet.* 10: 446-464, 1958.
- Stern, C. *Principles of Human Genetics*. W.H. Freeman and Co., San Francisco, 1960.
- Sutter, J. Recherches sur les effets de la consanguinité chez l'homme. Lons-le-Saunier, Declume Press, 1958, cf. Morton, N.E., 1960, *op. cit.*

- Sutter, J. & Tabah, L. Effets de la consanguinité et de la endogamie. *Population* 7: 249-266, 1952.
- Sutter, J. & Tabah, L. Structure de la mortalité dans les familles consanguines. *Population* 8: 511-526, 1953.
- Wahlund, S. Zusammensetzung von Populationen und Korrelationserscheinungen von Standpunkt der Vererbungslehre aus betrachtet. *Hereditas* 11: 65-106, 1928, cf. Li, C.C. *Population genetics*. Univ. Chicago Press, Chicago, 7ª reimpressão, 1972.
- Wright, S. Systems of mating. I-V. *Genetics* 6: 111-178, 1921.
- Wright, S. Coefficients of inbreeding and relationship. *Amer. Nat.* 56: 330-338, 1922.
- Wright, S. -The genetic structure of populations. *Ann. Eugen.* 15: 323-354, 1951.

## CAPÍTULO 6. OS FATORES EVOLUTIVOS

Ao discutirmos o equilíbrio de Hardy e Weinberg, ficou bastante claro que a estabilidade das frequências gênicas numa população somente pode ser mantida se nela não estiverem atuando os fatores evolutivos, ou seja, se a população não estiver sujeita a **mutações**, **seleção natural**, **deriva genética** e a **fluxos gênicos de populações migrantes**. Os efeitos de tais fatores serão analisados no presente capítulo. Antes, porém, é necessário salientar que a **evolução** é definida, atualmente, como a **alteração das frequências de alelos pertencentes ao conjunto gênico da população estudada**. Em vista disso, pode-se concluir que, apesar de a evolução afetar indivíduos, não são eles que evoluem e sim a população como um todo. Assim, por exemplo, se um par de alelos  $A, a$ , que ocorre em uma determinada geração de uma população com frequências  $p = 0,95$  e  $q = 0,05$ , passar a mostrar, em uma geração descendente, frequências  $p = 0,92$  e  $q = 0,08$ , não diremos que houve evolução dos indivíduos portadores do alelo  $a$  e sim que o conjunto gênico da população evoluiu num determinado sentido, que, no caso, foi o aumento da frequência do alelo  $a$ .

### 1. MUTAÇÃO E SELEÇÃO

É sabido que toda a alteração do genótipo, que surge repentina e aleatoriamente é denominada **mutação**. Em consequência disso, sob esse nome estão incluídas tanto as **alterações submicroscópicas** do material genético, isto é, as alterações nas seqüências de bases do DNA, quanto as alterações do material genético visíveis ao microscópio, isto é, as **aberrações cromossômicas**. As alterações submicroscópicas do material genético, que são comumente chamadas de **mutações gênicas**, ocorrem durante a replicação do DNA, na fase S do ciclo celular, sem que sejam reparadas, podendo tornar-se, por isso, permanentes.

As mutações gênicas, mesmo quando pontuais, isto é, resultantes da substituição de uma única base (púrica ou pirimídica), podem provocar alterações fenotípicas importantes. Não é difícil, pois, imaginar que as aberrações cromossômicas, por serem modificações mais grosseiras do material genético, notadas ao microscópio, possam ter efeitos dramáticos, ao provocar profunda desorganização da informação genética, que pode chegar a impedir a viabilidade do organismo no qual se manifestam. Essas anomalias do cariótipo são classificadas como **numéricas**, quando existe a adição ou a perda de cromossomos, e como **estruturais** (**translocação**, **inversão**, **deficiência**, **duplicação**) quando ocorrem rearranjos do material genético de um cromossomo ou entre cromossomos.

Seguindo a tendência atual, no presente capítulo não discutiremos as aberrações cromossômicas, e nos ateremos apenas às mutações gênicas, como provedoras da variabilidade genética que é submetida à **seleção natural**, definida por Darwin (1859) como sendo  **$a$**

*preservação das variações favoráveis e a supressão das prejudiciais*, definição essa que continua a ser aceita. Em outras palavras, a permanência na população dos alelos surgidos por mutação depende da ação seletiva exercida pelo ambiente contra os portadores desses alelos. Parece claro, pois, que as duas questões – *mutação* e *seleção* - estão ligadas indissolavelmente, sendo essa a razão de ambas serem discutidas em conjunto. Antes, porém, vale a pena breves comentários sobre mutações espontâneas e induzidas, e sobre mutações somáticas e gaméticas.

### MUTAÇÕES ESPONTÂNEAS E INDUZIDAS

As experiências com microrganismos revelaram que, mesmo em condições normais controladas de cultura, podem ocorrer mutações, as quais, por não terem causa aparente, são denominadas *mutações espontâneas*, para diferenciá-las daquelas que são provocadas por *agentes mutagênicos* e, por isso, denominadas *mutações induzidas*. A verdade é que, a não ser em casos em que é possível controle laboratorial, dificilmente encontramos um ambiente desprovido de agentes mutagênicos, os quais incluem, principalmente, as radiações ionizantes e numerosas substâncias químicas como os pesticidas e fungicidas empregados na agricultura, os peróxidos, os conservantes de alimentos, o ciclamato usado como adoçante, o benzopireno presente no ar poluído das grandes cidades, as alterações de pH etc. Acredita-se que a dose máxima permitida para as pessoas que trabalham sob risco de radiações (0,3 r por semana) prevê apenas riscos individuais, mas não os riscos de indução de mutações, pois essa dose estaria acima daquela capaz de aumentar as taxas de mutação.

Outros agentes mutagênicos são as radiações alfa e beta, os nêutrons, a luz ultravioleta e a idade dos genitores. A elevação da idade materna está bastante associada ao aumento do risco de aberrações cromossômicas, mas o aumento da idade dos genitores também está associada a ocorrência de mutações submicroscópicas do material genético. Assim, por exemplo, com a elevação da idade paterna cresce a probabilidade de nascimento de crianças com nanismo acondroplásico e de aparecimento de mutações em genes do cromossomo X.

### MUTAÇÕES SOMÁTICAS E GAMÉTICAS

Evidentemente, as mutações podem ocorrer tanto nas células somáticas (*mutações somáticas*) quanto nas células germinativas (*mutações gaméticas*). As primeiras podem ter sérios reflexos a nível individual, já que a célula somática onde ocorre a mutação pode sofrer alterações significativas, como tornar-se cancerosa ou morrer. Entretanto, do ponto de vista genético, as mutações somáticas não são importantes, porque elas não são transmissíveis hereditariamente e desaparecem com a morte das células nas quais elas ocorreram ou com a morte do indivíduo ao qual essas células pertencem.

Em oposição, as mutações gênicas gaméticas têm grande importância genética, porque elas podem introduzir novas formas alélicas na população, pois podem ser transmitidas hereditariamente. Em outras palavras, se o zigoto que contém a mutação for viável e der origem a um indivíduo, todas as células de tal indivíduo apresentarão a mutação em apreço, a qual poderá ser transmitida a seus descendentes por intermédio dos gametas desse indivíduo. Já a maioria das aberrações cromossômicas que ocorrem na linhagem germinativa não tem importância do ponto de vista genético, porque, geralmente, não são transmissíveis hereditariamente, a exemplo das mutações somáticas.

### VALOR ADAPTATIVO E COEFICIENTE SELETIVO

Visto que as mutações são alterações acidentais de um material genético que se mostrou, de certo modo, satisfatório, após milênios de seleção natural, é compreensível que numerosas mutações sejam reconhecidas pelo efeito deletério que provocam, apesar de poderem existir aquelas que, no mesmo ambiente, têm efeito superior ou idêntico ao do alelo mais antigo.

Se um alelo condicionar um efeito fenotípico deletério, seus portadores encontrarão dificuldade para se manter na população, o que prejudicará a transmissão desse gene às gerações seguintes, impedindo, assim, o aumento de sua frequência. Em outras palavras, se um mutante tiver efeito fenotípico deletério, seus portadores mostrarão menor adequação biológica ao ambiente, isto é, menor *adaptação*, quando comparados aos portadores do alelo mais antigo, o qual pode ser chamado de *alelo normal*.

Essa adequação biológica, que recebe o nome de *valor adaptativo*, geralmente simbolizado por  $f$ , inicial da palavra inglesa *fitness*, pode ser definida como *a razão entre a fecundidade média dos portadores de uma mutação e a fecundidade média dos indivíduos com o fenótipo tomado como referência e considerado normal*. Como se vê, o valor adaptativo compara a capacidade que os portadores de uma mutação têm de transmiti-la à geração seguinte, com a capacidade que os não-portadores dessa mutação têm em transmitir o alelo mais antigo. Desse modo, o genótipo que propiciar maior número de filhos será considerado melhor adaptado ao ambiente. O valor adaptativo mede, portanto, o sucesso reprodutivo, que depende da *sobrevivência até a idade reprodutiva*, da *seleção sexual*, isto é, do sucesso no acasalamento, e do *número de descendentes*.

Apesar de a seleção natural ser menos rigorosa na espécie humana, em consequência, principalmente, da melhoria das condições sanitárias e dos avanços da medicina, devemos nos lembrar que apenas cerca de 30% dos zigotos alcançarão o estágio de indivíduos aptos a se reproduzir, porque se estima que cerca de 30% das gestações são abortadas espontaneamente, 5%

delas são natimortos ou em óbito neonatal, 3% falecem na infância, 20% sobrevivem até a idade adulta, mas não casam nem deixam descendentes e, dentre os que casam, 10% não têm filhos.

Para o cálculo do valor adaptativo, que é um número que varia de zero a um, o mais correto é levar em conta apenas os filhos que alcançaram a idade reprodutiva. Assim, por exemplo, se os indivíduos com uma determinada anomalia genética tiverem, em média, 1,5 filhos que alcançaram a idade reprodutiva, enquanto os indivíduos de uma amostra controle tiverem, em média, 4 filhos que atingiram essa idade, diremos que o valor adaptativo dessa anomalia é 0,375, porque  $f = \frac{1,5}{4} = 0,375$ .

O valor adaptativo de uma anomalia também pode ser calculado tomando-se como referência a estimativa do número médio de filhos de casais da população geral que alcançam a idade de 20 anos. Assim, por exemplo, admitamos que se estima em 2,5 o número médio de filhos dos casais da população geral e que 80% dos indivíduos dessa população atingem 20 anos de idade. Aceitemos, também, que os indivíduos com uma determinada heredopatia deixam, em média, 1,5 filhos, e que a taxa de sobrevivência até os 20 anos dos indivíduos com essa doença é 20%. Nesse caso, calcularíamos o valor adaptativo das anomalia em estudo por intermédio de  $f = \frac{1,5 \times 0,20}{2,5 \times 0,80} = 0,15$ .

A recíproca do valor adaptativo é o **coeficiente seletivo**, geralmente simbolizado pela letra  $s$ , que serve para medir a intensidade da seleção natural sobre um determinado fenótipo. Assim, por exemplo, se uma anomalia genética tiver valor adaptativo  $f = 0,20$ , poderemos dizer que seu coeficiente seletivo é  $s = 0,80$ , porque  $s = 1 - f = 1 - 0,20 = 0,80$ . Se o valor adaptativo for  $f = 0,15$ , o coeficiente seletivo valerá, obviamente,  $s = 0,85$ .

É bastante sabido que a seleção natural opera sobre os fenótipos e não sobre os genes. Apesar disso, no caso de mutações com efeito dominante e penetrância completa, pode-se dizer que os conceitos de valor adaptativo e de coeficiente seletivo podem ser estendidos aos genes mutantes, porque, nesse caso, os fenótipos deles resultantes e, conseqüentemente, os genes mutantes, ficam expostos à seleção natural desde o momento do surgimento dessas mutações. Não é esse, porém, o caso das mutações dominantes com penetrância incompleta, nem daquelas que se manifestam apenas em homozigose (fenótipo recessivo). Nesse último caso, os mutantes somente se expõem à seleção natural quando atingem freqüências que permitem o aparecimento de homozigotos.

## GENES LETAIS

Se uma mutação produzir um alelo que confere a seus portadores um valor adaptativo nulo ( $f = 0$ ), deixando-os, portanto, sujeitos à seleção total ( $s = 1$ ), tal alelo será denominado

*gene letal*. Evidentemente, sua frequência na população será baixa e mantida apenas pela taxa de mutação. Os genes letais não têm, obrigatoriamente, que provocar a morte precoce de seus portadores, porque a seleção natural pode operar em vários níveis, como já mencionamos acima, de sorte que *o efeito deletério de uma mutação pode manifestar-se nos gametas, nos zigotos, no embrião, nos recém-nascidos ou em qualquer outro momento do desenvolvimento do indivíduo*. Por isso, do ponto de vista genético, tanto os genes que causam esterilidade quanto os que impedem seus portadores de alcançar a idade reprodutiva são considerados letais, pois, nessas situações  $s = 1$ .

Obviamente, tanto os genes letais quanto as mutações que conferem baixo valor adaptativo a seus portadores serão encontrados em proporção muito pequena na população. Essa a razão pela qual a esmagadora maioria dos genes que determinam efeitos deletérios são *genes idiomorfos*, isto é, têm frequências inferiores a 1%, e não chegam a ser *genes polimorfos*, que atingem frequências entre 1% e 99%.

### PERSISTÊNCIA MÉDIA

Se um mutante puder ser considerado um *gene neutro*, isto é, se, praticamente, os seus portadores não estiverem sujeitos à seleção ( $s = 0$ ), ele permanecerá indefinidamente ou quase indefinidamente na população. Por isso, se um mutante conferir a seus portadores maior adaptabilidade ao ambiente do que seus alelos mais antigos, a sua frequência passará a aumentar na população, enquanto a frequência dos alelos mais antigos diminuirá.

Para tornar mais compreensíveis essas afirmações, consideremos um alelo surgido por mutação, que tenha efeito dominante e penetrância completa. A *persistência média* desse gene, simbolizada por  $i$ , indica o número médio de gerações que esse mutante permanecerá na população, e é inversamente proporcional ao coeficiente seletivo de seus portadores, podendo-se, pois, escrever  $i = \frac{1}{s}$ . Se esse alelo surgido por mutação for letal, isto é, se o coeficiente seletivo de seus portadores for  $s = 1$ , a persistência média desse gene na população também será a unidade ( $i = \frac{1}{1} = 1$ ), ou seja, ele somente conseguirá permanecer na população durante uma única geração, sendo, por isso, a sua frequência igual à taxa dessa mutação. Se, ao contrário, o alelo surgido por mutação for neutro, de modo a conferir valor adaptativo  $f = 1$ , isto é,  $s = 0$  a seus portadores, esse gene permanecerá indefinidamente na população porque  $i = \frac{1}{0} = \infty$ .

A introdução de modificações no ambiente pode fazer com que um alelo, que, até então, conferia pouca adaptabilidade a seus portadores, passe a sofrer pouca seleção. Desse modo, aumentará a persistência média de tal alelo e, em consequência, a sua frequência na população. A

Medicina moderna está se tomando um poderoso agente anti-seletivo, já que, por intermédio dos recursos que ela vem utilizando, numerosos defeitos genéticos que provocavam morte precoce ou que, de algum outro modo, diminuía as possibilidades matrimoniais dos portadores desses defeitos estão sendo corrigidos atualmente.

### O EQUILÍBRIO ENTRE MUTAÇÃO E SELEÇÃO

A aceitação da existência de um equilíbrio entre mutação e seleção não é difícil, se partimos do princípio de que, se tal equilíbrio não existisse, as hereditopatias monogênicas que ocorrem atualmente deveriam ter freqüências muito altas entre os nossos ancestrais mais distantes.

Assim, por exemplo, uma anomalia dominante autossômica sujeita a um coeficiente seletivo  $s = 0,80$ , isto é, com valor adaptativo  $f = 0,20$ , deveria, na ausência de mutação e de alterações sérias do ambiente, ter a sua freqüência diminuída de 80% em cada geração. Desse modo, aceitando a existência de quatro gerações humanas em cada século, ter-se-ia, em apenas dois séculos, que a freqüência de tal anomalia ficaria reduzida a  $f^8 = (0,2)^8 = 0,00000256$  de seu valor original. Em outras palavras, teríamos que contrariar as evidências históricas e admitir o absurdo de que, há alguns séculos, numerosas anomalias dominantes autossômicas teriam freqüências centenas ou milhares de vezes mais altas do que as que apresentam atualmente. Como corolário, teríamos, ainda, que admitir estarem essas anomalias em processo de extinção. Visto que essa hipótese deve ser rejeitada, a melhor alternativa é a admissão da existência de um equilíbrio entre o processo de eliminação dos genes com efeito deletério (*seleção*) e a sua *taxa de mutação*, isto é, a freqüência com a qual eles são introduzidos na população em cada geração.

Analisemos, agora, de modo muito simplificado, a maneira pela qual esse equilíbrio dinâmico é atingido. Para tanto admitamos a existência de uma população teórica vivendo em um ambiente que não sofre alterações. Consideremos, ainda, que essa população mantém um número constante de um milhão de indivíduos em cada geração, e aceitemos que na geração inicial todos os indivíduos têm genótipo  $aa$  e que a taxa de mutação ( $\mu$ ) do gene  $a$  para o seu alelo  $A$ , cujo efeito em estudo é dominante, é igual a 1:50.000 ou 0,00002. Finalmente, admitamos que os indivíduos  $Aa$  estão sujeitos a um coeficiente seletivo  $s = 0,80$ .

Visto que a população em apreço tem um milhão de indivíduos por geração, pode-se dizer que na geração inicial existiriam dois milhões de genes  $a$  e que tal geração transmitiria à seguinte  $\frac{1}{50.000} \times 2.000.000 = 40$  genes mutantes  $A$ . Tendo em vista o alto coeficiente seletivo ao qual estão sujeitos os indivíduos  $Aa$ , é claro que essa primeira geração filial não transmitiria os 40 mutantes para a segunda, mas apenas 8 deles, pois o valor adaptativo dos indivíduos  $Aa$  é  $f = 1-s = 0,20$ , de sorte que se tem  $0,2 \times 40 = 8$ . Em consequência disso, e não levando em conta,

para facilidade de cálculo e exposição, que a primeira geração filial tem  $2.000.000 - 40 = 1.999.960$  alelos  $a$ , mas considerando que ela e todas as seguintes continuam a ter dois milhões desses genes, concluiríamos que a segunda geração filial deveria conter 48 alelos  $A$ , ou seja, 8 da geração anterior e 40 novos mutantes.

Lembrando que, por hipótese, o coeficiente seletivo contra os portadores do gene  $A$  se mantém constante, tem-se, então, que a terceira geração filial conteria 49,6 genes  $A$ , pois ela receberia 40 novos mutantes da segunda geração filial, 8 dos mutantes originários da primeira geração filial ( $0,2 \times 40$ ) e 1,6 mutantes oriundos da geração inicial ( $0,2 \times 8$ ). Operando de modo análogo nas gerações seguintes chega-se à conclusão de que, na população teórica em apreço, o número de alelos  $A$  tenderia para o valor fixo de 50 alelos, ou seja, para a frequência

$$p = \frac{50}{2.000.000} = 0,000025, \text{ o que equivale a } p = \frac{\mu}{s} = \frac{0,00002}{0,80} = 0,000025.$$

Se no mesmo exemplo admitíssemos que a taxa de mutação fosse a metade da anterior, isto é, 1: 100.000, concluiríamos que o aumento da frequência do alelo  $A$  na população seria menor, enquanto que o equilíbrio seria atingido quando o número de genes  $A$  atingisse o limite de 25, ou seja, a frequência  $p = \frac{25}{2.000.000} = 0,0000125$ .

Ainda em relação ao mesmo exemplo, é fácil constatar que se o coeficiente seletivo que opera contra os portadores do mutante  $A$  fosse menor, a frequência atingida por esse alelo, no momento em que fosse alcançado o equilíbrio entre a taxa de mutação e a seleção, seria mais alta do que as mencionadas anteriormente. Assim, se o coeficiente seletivo fosse igual ao valor adaptativo, isto é, se  $s = f = 0,5$ , a frequência do alelo  $A$  tenderia para o valor  $p = 0,00004$  quando a taxa de mutação fosse 1:50.000, e para o valor  $p = 0,00002$  quando ela fosse 1:100.000. Por outro lado, se a seleção fosse total ( $s = 1$ ) é claro que a frequência do alelo  $A$  na população seria igual à taxa de mutação. Em outras palavras, se a taxa de mutação for constante e a seleção natural aumentar, a frequência do gene mutante diminuirá, ocorrendo o oposto quando houver um relaxamento do processo seletivo.

Do exposto, portanto, pode-se concluir que, em relação a uma heredopatia dominante autossômica monogênica, que é mantida somente por mutação, a frequência  $p$  do gene que a condiciona atingirá, no máximo, o valor dado pela razão entre a taxa de mutação ( $\mu$ ) e o seu coeficiente seletivo ( $s$ ), ou seja,  $p = \frac{\mu}{s}$ .

Por analogia se conclui que, no caso de uma anomalia recessiva autossômica, isto é, quando o mutante tem efeito detrimental apenas em homozigose, a frequência  $q$  de tal gene atingirá, no máximo, o valor dado por  $q = \sqrt{\frac{\mu}{s}}$ . De fato, considerando que, no caso das anomalias

recessivas autossômicas mantidas por mutação, são eliminados os homozigotos, que devem ocorrer com frequência  $q^2$ , tem-se  $q^2 = \frac{\mu}{s}$ , de onde se tira  $q = \sqrt{\frac{\mu}{s}}$ .

No caso de heredopatias recessivas ligadas ao sexo, pode-se aceitar, também, que a frequência  $q$  do gene ligado ao cromossomo X responsável por elas atingirá, no máximo,  $q = \frac{3\mu}{s}$ .

O mais correto, na realidade, seria escrever  $q \cong \frac{3\mu}{s}$  porque a taxa de mutação de um gene ligado ao cromossomo X pode não ser idêntica nas mulheres ( $\mu_F$ ) e nos homens ( $\mu_M$ ), o que conduz a  $q = \frac{2\mu_F + \mu_M}{s}$ , porque as mulheres têm dois cromossomos X e os homens apenas um. Como se vê, somente quando  $\mu_F = \mu_M$  tem-se  $q = \frac{3\mu}{s}$ .

### O CÁLCULO DA TAXA DE MUTAÇÃO

A aceitação do equilíbrio entre mutação e seleção permite entrever a possibilidade de se estimar a taxa com que são produzidos genes mutantes responsáveis por heredopatias com transmissão monogênica, quando se sabe apenas a sua *incidência*, isto é, a *freqüência com que surgem casos novos na população em um determinado intervalo de tempo*.

Em relação às anomalias autossômicas dominantes, a incidência ( $x$ ) de anômalos diz respeito, na prática, apenas aos heterozigotos do gene que as determina ( $x = Aa$ ), pois sabemos que, nesses casos, é nula ou quase nula a ocorrência de indivíduos homozigotos ( $AA$ ). Em vista disso, se a frequência do alelo  $A$  for designada por  $p$  é claro que seu valor será igual à metade da incidência dos indivíduos  $Aa$ , ou seja,  $p = \frac{Aa}{2} = \frac{x}{2}$ . Considerando, por outro lado, que  $p = \frac{\mu}{s}$  podemos escrever que  $\mu = sp$ , ou, ainda, que no caso das anomalias autossômicas dominantes  $\mu$  é calculado segundo:

$$\mu = \frac{sx}{2}$$

A essa mesma conclusão se pode chegar por intermédio de outro tipo de raciocínio. Assim, consideremos que  $N$  é o número de indivíduos de uma população composta predominantemente por pessoas com o genótipo  $aa$ , isto é, por homozigotos de um gene autossômico  $a$ ,  $\mu$  é a taxa de mutação do gene  $a$  para  $A$ , que condiciona uma anomalia dominante ( $Aa$ ) sujeita a um coeficiente seletivo  $s$ ,  $x$  é a incidência de tal anomalia e  $xN$  é o número de recém-nascidos com a anomalia dominante.

Nesse caso,  $N-xN$  será o número total de indivíduos homozigotos  $aa$  entre os recém-nascidos;  $2(N-xN)$  será o número de genes dos indivíduos homozigotos  $aa$  e  $xN$  será tanto o número de genes  $a$ , quanto o de genes  $A$  dos indivíduos  $Aa$ . Portanto, em uma geração, o

número de genes  $a$  capazes de produzir alelos  $A$  por mutação será  $N(2-x)$  pois  $2(N-xN) + xN = 2N - 2xN + xN = 2N - xN = N(2-x)$ , de sorte que  $\mu N(2-x)$  será o número de novos mutantes por geração. Lembrando, porém, que  $x$  tem valor muito baixo pode-se escrever que tal número será obtido por intermédio de  $2N\mu$ .

Visto que, dos  $xN$  anômalos de cada geração,  $sxN$  são eliminados, e considerando a existência de equilíbrio entre o número de mutações introduzidas na população e aquelas eliminadas, chega-se à igualdade  $2N\mu = sxN$ , da qual se extrai  $\mu = \frac{sx}{2}$ .

No caso das anomalias recessivas autossômicas já vimos no tópico anterior que  $q^2 = \frac{\mu}{s}$  e, tendo em mente que a frequência  $x$  de anômalos recessivos corresponde a  $q^2$  conclui-se, pois que  $\mu = sx$ .

No concernente às anomalias recessivas ligadas ao sexo sabemos que  $q = \frac{3\mu}{s}$ , sendo  $q$ , praticamente, igual à frequência  $x$  dos homens anômalos. Em vista disso, a estimativa da taxa de mutação passa a ser obtida por intermédio de  $\mu = \frac{sx}{3}$ .

Apesar da validade teórica dessas fórmulas, elas estão sujeitas a muitos riscos de erro, pois, para a sua aplicação é necessário que entre os casos anômalos não sejam, incluídas fenocópias, genocópias, nem filhos ilegítimos. Além disso, os genes ou genótipos em estudo devem ter penetrância completa. De fato, em relação às doenças dominantes, tanto a sua penetrância incompleta nos genitores dos anômalos quanto as mutações em locos diferentes que produzem fenótipo semelhante (genocópias) provocariam super-estimativas das taxas de mutação. Os riscos de erro são, evidentemente, maiores no caso das doenças recessivas autossômicas, pois sabemos que muitos dos genes que as causam exercem um efeito seletivo também nos heterozigotos. Por outro lado, pelo fato de as populações humanas estarem, atualmente, quebrando os isolados de modo muito intenso, em consequência das maiores facilidades de comunicação e transporte, é fácil vislumbrar que elas não estão em equilíbrio em relação a muitos genes. Por isso, a possibilidade de decréscimo da frequência de homozigotos recessivos entre elas, impedindo a sua averiguação e eliminação, teria que ser levada em conta. Em resumo, pode-se dizer, portanto, que das fórmulas citadas acima para estimar a taxa de mutação, a que fornece maior risco de erro é aquela que pretende calcular a taxa de mutação de genes autossômicos responsáveis por anomalias recessivas.

Para calcular a taxa de mutação, no caso das doenças dominantes autossômicas, não é necessário aplicar o método indireto aqui descrito, podendo-se, para esse fim, empregar o método direto, que consiste em calcular a metade da incidência dos casos esporádicos que ocorrem em

um determinado intervalo de tempo. Evidentemente, o emprego do método direto não exclui os riscos de erro do método indireto, ou seja, a inclusão de fenocópias, de genocópias e de filhos ilegítimos, bem como a inclusão, entre os mutantes, de filhos de pessoas nas quais o gene não se expressou clinicamente, isto é, não teve penetrância.

Para demonstrar a aplicação dos métodos direto e indireto de cálculo da taxa de mutação de genes com efeito dominante, suponhamos que, em 100.000 nascimentos ocorridos numa população, em um determinado período, tenha sido diagnosticada uma anomalia autossômica dominante em 6 crianças, duas das quais representado casos familiares, isto é, com um dos genitores manifestando a mesma anomalia. Visto que os quatro recém-nascidos restantes são casos esporádicos, isto é, gerados por genitores normais, a incidência desses casos pode ser estimada em  $4: 100.000$  ou  $0,00004$ .

Aceitando-se a hipótese de que, nos casos esporádicos, a anomalia dominante foi determinada por um único gene autossômico raro com penetrância completa e resultante de mutação, consideramos cada caso esporádico como oriundo de um zigoto no qual o mutante estava acompanhado de um alelo condicionador de normalidade. Com isso, a taxa da mutação que determina a anomalia em estudo passa a ser estimada em  $0,00002$ , pois devemos levar em conta a metade da frequência de anômalos, isto é,  $\mu = \frac{0,00004}{2} = 0,00002$  ou  $\mu = 2 \times 10^{-5}$ . Isso equivale a dizer que, em cada geração, 2 em cada 100.000 gametas apresentariam a mutação em apreço.

Para calcular a taxa de mutação que origina um gene responsável por uma anomalia dominante empregando o método indireto, precisamos, inicialmente, estimar o valor adaptativo da anomalia, o que é feito pela comparação da fecundidade média dos anômalos com a fecundidade média de uma amostra controle. Suponhamos que tivéssemos constatado que 100 indivíduos com a anomalia em estudo tiveram um total de 63 filhos que alcançaram a idade reprodutiva, enquanto que 400 pessoas normais de uma amostra controle geraram 840 filhos que ultrapassaram a maioridade. Nesse caso, a fecundidade média dos anômalos seria estimada em  $0,63$ , pois  $\frac{63}{100} = 0,63$ , e a dos normais seria estimada em  $2,10$ , pois  $\frac{840}{400} = 2,10$ , o que, por sua vez, permite estimar o valor adaptativo da anomalia em  $0,30$ , pois  $f = \frac{0,63}{2,10} = 0,30$ . Isso equivale a dizer que o coeficiente seletivo dessa anomalia é alto, pois  $s = 1 - f = 0,70$ .

Se a anomalia dominante autossômica em estudo for a mesma do exemplo anterior, poderemos estimar sua incidência na população em  $x = 0,00006$ , pois  $\frac{6}{100.000} = 0,00006$ . Como

conseqüência, a taxa de mutação do gene da anomalia passa a ser estimada em, praticamente,  $2 \times 10^{-5}$  porque  $\mu = \frac{sx}{2} = \frac{0,7 \times 0,00006}{2} = 0,000021$ .

Saldanha (1962<sub>a</sub>) sugeriu um método alternativo para calcular o valor adaptativo de anomalias dominantes autossômicas, o qual consiste em verificar a freqüência de casos familiares em uma amostra de anômalos, pois tais casos indicariam a freqüência com que os genes mutantes passariam de uma geração a outra. Desse modo, ter-se-ia uma avaliação direta da razão entre as mutações transmitidas e as mutações produzidas. Assim, por exemplo, se dentre 100 propósitos com uma determinada anomalia autossômica dominante 10 forem casos familiares, o valor adaptativo dessa anomalia seria estimada em 0,10, pois  $f = \frac{10}{100} = 0,10$ , disso resultando que a razão entre as mutações eliminadas e as produzidas seria estimada em  $s = 0,90$ . Nesse caso, o intervalo de confiança de 95% de probabilidade que contém o valor do coeficiente seletivo está entre 0,84 e 0,96, pois o desvio padrão de  $s$  é estimado em 0,03, visto que

$$\sigma = \sqrt{\frac{0,90 \times 0,10}{100}} = 0,03.$$

### SELEÇÃO CONTRA ANOMALIAS DOMINANTES MONOGÊNICAS

A velocidade com que uma anomalia dominante autossômica monogênica é eliminada da população depende do seu coeficiente seletivo ( $s$ ), bem como da penetrância ( $P$ ) do gene que a determina, porque a persistência média ( $i$ ) desse gene será estimada por  $i = \frac{1}{Ps}$ .

Se a seleção for total ( $s = 1$ ) e a penetrância do gene autossômico for completa ( $P = 1$ ), a freqüência da anomalia dominante por ele determinada será, evidentemente, igual apenas à taxa de mutação que re-introduz esse gene na população, pois os mutantes não terão oportunidade de deixar prole. Cada gene resultante de mutação persistirá por uma única geração, pois  $i = 1$ .

Se a seleção for total ( $s = 1$ ), mas a penetrância do gene autossômico que determina a anomalia dominante não for completa ( $P < 1$ ), a velocidade com que essa anomalia será eliminada da população dependerá apenas do valor da penetrância do gene. Assim, por exemplo, consideremos dois genes mutantes determinadores de anomalias autossômicas dominantes e que um deles ( $A$ ) tem penetrância igual a 70% e o outro ( $B$ ) penetrância igual a 90%. Se ambos estiverem sujeitos a seleção total ( $s = 1$ ), a anomalia determinada pelo gene  $B$  será eliminada mais rapidamente da população do que aquela determinada pelo gene  $A$ , porque a penetrância do gene  $B$ , sendo maior, propiciará persistência média menor dos mutantes. De fato, enquanto a

persistência média do gene  $A$  é estimada em 1,43 gerações, pois  $i = \frac{1}{0,7}$ , a do gene  $B$  é estimada em 1,11 gerações, visto que  $i = \frac{1}{0,9}$ .

Se a seleção não for total ( $s < 1$ ), mas a penetrância do gene autossômico que determina a anomalia dominante for completa ( $P=1$ ), é claro que a velocidade de eliminação dessa anomalia dependerá apenas do valor de seu coeficiente seletivo. Assim, por exemplo, consideremos dois genes mutantes ( $A$  e  $B$ ) com penetrância completa ( $P = 1$ ), que determinam anomalias autossômicas dominantes. Se o gene  $A$  estiver sujeito a coeficiente seletivo  $s = 0,8$  e o gene  $B$  a  $s = 0,4$ , a anomalia determinada pelo gene  $A$  será eliminada mais rapidamente da população porque a persistência média desse gene será  $i = \frac{1}{0,8} = 1,25$ , enquanto que a do gene  $B$  será  $i = \frac{1}{0,4} = 2,5$ .

Quando não há seleção total ( $s < 1$ ) contra um gene sem penetrância completa ( $P < 1$ ) que determina uma anomalia autossômica dominante, a velocidade de eliminação dessa anomalia será inversamente proporcional ao produto dos valores dessas duas variáveis. Para exemplificar, consideremos dois genes autossômicos resultantes de mutação ( $A$  e  $B$ ) que determinam anomalias dominantes, e aceitemos que o mutante  $A$ , com penetrância de 60%, confere  $s = 0,6$  a seus portadores, ao passo que o mutante  $B$ , com penetrância de 90%, confere a seus portadores  $s = 0,4$ . Nesse caso, as duas anomalias serão eliminadas com a mesma velocidade, pois os dois mutantes terão a mesma persistência média na população (2,78 gerações). De fato, em relação ao gene  $A$ ,  $i = \frac{1}{0,6 \times 0,6} = 2,78$ , e, em relação ao gene  $B$ ,  $i = \frac{1}{0,4 \times 0,9} = 2,78$ .

Quando a penetrância é incompleta e a seleção não é total, o efeito da seleção contra anomalias dominantes ligadas ao sexo pode diferir daquele que resulta da ação seletiva contra anomalias dominantes autossômicas porque nas ligadas ao sexo, tanto o coeficiente seletivo quanto a penetrância dos genes que as determinam podem não ser iguais em ambos os sexos.

## SELEÇÃO CONTRA ANOMALIAS RECESSIVAS MONOGÊNICAS

A seleção natural contra as anomalias recessivas autossômicas monogênicas tem efeito muito pobre, quando comparado ao que ela exerce sobre as anomalias dominantes. Para demonstrar isso, consideremos uma geração qualquer de uma população, que chamaremos de inicial, na qual os genótipos determinados por um par de alelos autossômicos  $A, a$ , com frequências  $p$  e  $q = 1 - p$  se distribuem segundo a lei de Hardy e Weinberg. Aceitemos, ainda, que o gene  $a$  somente se expressa em homozigose, provocando uma anomalia sujeita a seleção total.

Partindo de uma geração inicial na qual as freqüências genotípicas se distribuem segundo  $AA = p^2$ ,  $Aa = 2pq$ ,  $aa = q^2$  ter-se-á que, na primeira geração filial, a freqüência do alelo  $a$  dependerá da proporção de indivíduos com genótipo  $Aa$  nas gerações seguintes. A freqüência  $q$  do alelo  $a$  na primeira geração filial será, então,  $q_1 = \frac{q}{1+q}$ , pois

$$q_1 = \frac{\frac{1}{2}Aa}{AA + Aa} = \frac{pq}{p^2 + 2pq} = \frac{q}{p + 2q} = \frac{q}{(1-q) + 2q} = \frac{q}{1+q}$$

Se o mesmo processo seletivo continuar na primeira geração filial é óbvio que, não levando em conta a ocorrência de mutações, a freqüência do alelo  $a$  na segunda geração filial

$$\text{será calculada por intermédio de } q_2 = \frac{q}{1+2q}, \text{ pois } q_2 = \frac{p_1 q_1}{p_1^2 + 2p_1 q_1} = \frac{q_1}{1+q_1} = \frac{\frac{q}{1+q}}{1 + \frac{q}{1+q}} = \frac{q}{1+2q}.$$

Admitindo que o mesmo processo persista durante  $n$  gerações, chega-se à conclusão de que na  $n$ ésima geração a freqüência  $q_n$  do alelo  $a$  será obtida por intermédio de

$$q_n = \frac{q}{1+nq}$$

Essa fórmula permite estimar, também, o número de gerações necessárias para que a freqüência  $q$  do alelo  $a$  na geração inicial baixe para uma determinada freqüência  $q_n$ , bastando, para isso, resolvê-la em função de  $n$ , obtendo

$$n = \frac{q - q_n}{qq_n}$$

Além disso, pode-se concluir que, não levando em conta a taxa de mutação, a freqüência de um gene que condiciona uma anomalia recessiva autossômica sujeita a seleção total somente ficará reduzida à metade quando o número de gerações ( $n$ ) for igual a  $\frac{1}{q}$ , pois, quando isso acontece,  $q_n = \frac{q}{2}$ .

Assim, por exemplo, lembrando que, em média, existem quatro gerações humanas por século, e não levando em conta a taxa de mutação, que é sempre muito baixa, tem-se que, para reduzir a incidência de uma anomalia autossômica recessiva monogênica de 1: 10.000 para 2,5: 100.000 seriam necessárias 100 gerações, isto é, 2.500 anos de seleção total. Somente assim a freqüência do gene responsável pela anomalia baixaria de 0,01 para 0,005.

A pouca eficiência da seleção natural sobre as anomalias recessivas autossômicas se acentua sensivelmente quando ela não é total ou quando a penetrância é incompleta. Realmente, se a seleção dos homozigotos  $aa$  não for total, ter-se-á, em uma geração inicial, que as

freqüências genotípicas que devem ser levadas em conta para estimar a freqüência do gene  $a$  na geração seguinte serão  $Aa = 2pq$  e  $aa = q^2 - sq^2$ . Desse modo, a freqüência  $q_1$  do alelo  $a$  na primeira geração filial será obtida por intermédio de  $q_1 = \frac{q - sq^2}{1 - sq^2}$ , pois:

$$q_1 = \frac{pq + q^2 - sq^2}{1 - sq^2} = \frac{(1 - q)q + q^2 - sq^2}{1 - sq^2} = \frac{q - sq^2}{1 - sq^2}$$

No caso de o valor do coeficiente seletivo ser pequeno, o denominador da última fórmula pode ser considerado como igual à unidade, sendo possível derivar uma outra, escrita como abaixo (Li, 1972), que permite calcular o número de gerações necessárias para que a freqüência de um gene que apenas se expressa em homozigose baixe de  $q$  até  $q_n$ :

$$n = \frac{\frac{q - q_n}{qq_n} + 2,303 \log_{10} \frac{qP_n}{q_nP}}{s}$$

De acordo com essa fórmula tem-se, por exemplo, que um gene condicionador de uma anomalia recessiva, sujeito a um coeficiente seletivo  $s = 0,10$  necessitaria de 1.007 gerações para reduzir a sua freqüência de 0,01 para 0,005 pois:

$$n = \frac{\frac{0,01 - 0,005}{0,01 \times 0,005} + 2,303 \log_{10} \frac{0,01 \times 0,995}{0,05 \times 0,99}}{0,10} = 1.007$$

A seleção natural opera de modo mais eficiente quando as heredopatias recessivas são determinadas por genes do cromossomo X. Realmente, se uma anomalia recessiva ligada ao sexo ficar sujeita à seleção total, tem-se, na hipótese de ausência de mutação, que o gene com efeito letal eliminaria, em cada geração, um terço dos genes da geração anterior.

### SELEÇÃO CONTRA ANOMALIAS RECESSIVAS POLIGÊNICAS

Quando uma anomalia recessiva tem determinação poligênica, o efeito da seleção contra elas é ainda menos eficiente do que a que opera contra as anomalias recessivas com transmissão monogênica. Para facilitar a demonstração dessa afirmação consideremos apenas dois pares de alelos autossômicos,  $A, a$  e  $B, b$  que segregam independentemente, e aceitemos que o genótipo  $aabb$  determina uma anomalia sujeita à seleção total. Consideremos, ainda, para maior facilidade de exposição, a existência de uma população na qual os genes  $A, a, B, b$  têm, em uma determinada geração, considerada inicial, freqüências  $A = p_1$ ,  $a = q_1 = 1 - p_1$ ,  $B = p_2$ ,  $b = q_2 = 1 - p_2$ , as quais são todas iguais a 0,5, de sorte que seja possível escrever  $A = a = B = b = 0,5 = q$ .

Se nessa geração inicial os genótipos decorrentes desses dois pares de alelos estiverem distribuídos segundo a lei de Hardy e Weinberg ter-se-ia, pois, que as frequências genotípicas esperadas seriam as seguintes:

$$\begin{array}{lll}
 AABB = p_1^2 p_2^2 = q^4 & AaBB = 2p_1 q_1 p_2^2 = 2q^4 & aaBB = q_1^2 p_2^2 = q^4 \\
 AABb = 2 p_1^2 p_2 q_2 = 2q^4 & AaBb = 4p_1 q_1 p_2 q_2 = 4q^4 & aaBb = 2 q_1^2 p_2 q_2 = 2q^4 \\
 Aabb = p_1^2 q_2^2 = q^4 & Aabb = 2p_1 q_1 q_2^2 = 2q^4 & aabb = q_1^2 q_2^2 = q^4
 \end{array}$$

Visto que os indivíduos com genótipo *aabb*, que ocorrem com frequência  $q^4 = 6,25\%$  na geração inicial, estão sujeitos a seleção total é evidente que eles não poderão contribuir com seus genes para a geração seguinte. Desse modo, a população geneticamente ativa passará a contar com  $1 - q^4$  indivíduos, dentre os quais, apenas os seis tipos de casais apresentados na Tabela 1.6 poderão dar origem a indivíduos *aabb* na primeira geração filial.

Tabela 1.6. Frequências dos casais que dão origem a indivíduos *aabb* e dos filhos com tal genótipo quando existe seleção completa contra eles. Nesse exemplo  $A = a = B = b = q$ .

Casais		Filhos <i>aabb</i>	
Tipo	Frequência	Proporção familiar	Frequência
<i>AaBb</i> × <i>AaBb</i>	$\frac{16q^8}{(1-q^4)^2}$	$\frac{1}{16}$	$\frac{q^8}{(1-q^4)^2}$
<i>AaBb</i> × <i>Aabb</i>	$\frac{16q^8}{(1-q^4)^2}$	$\frac{1}{8}$	$\frac{2q^8}{(1-q^4)^2}$
<i>AaBb</i> × <i>aaBb</i>	$\frac{16q^8}{(1-q^4)^2}$	$\frac{1}{8}$	$\frac{2q^8}{(1-q^4)^2}$
<i>Aabb</i> × <i>Aabb</i>	$\frac{4q^8}{(1-q^4)^2}$	$\frac{1}{4}$	$\frac{q^8}{(1-q^4)^2}$
<i>Aabb</i> × <i>aaBb</i>	$\frac{8q^8}{(1-q^4)^2}$	$\frac{1}{4}$	$\frac{2q^8}{(1-q^4)^2}$
<i>aaBb</i> × <i>aaBb</i>	$\frac{4q^8}{(1-q^4)^2}$	$\frac{1}{4}$	$\frac{q^8}{(1-q^4)^2}$

Com base nos dados da Tabela 1.6, na qual a soma da última coluna fornece  $\frac{9q^8}{(1-q^4)^2}$  conclui-se, pois, que, na população utilizada para exemplo, a frequência de indivíduos com o genótipo *aabb* baixará, após uma geração na qual há seleção total contra eles, de 6,25% para 4%, visto que  $\frac{9q^8}{(1-q^4)^2} = \frac{9 \times 0,003906}{(1-0,0625)^2} = 0,04$ . Entretanto, a frequência dos genes *a* e *b* não sofrerá decréscimo tão grande, pois cada um deles passará a ser encontrado na primeira geração filial com frequência igual a 46,67% visto que  $\frac{q - q^4}{1 - q^4} = \frac{0,4375}{0,9375} = 0,4667$ .

A partir dessa geração o resultado do processo seletivo será cada vez menos marcante, mesmo que prevaleçam as condições de seleção total contra os indivíduos com genótipo *aabb*. Assim, por exemplo, na segunda geração filial, a frequência dos alelos *a* e *b* passará a ser 44,01%, na terceira 41,83%, na quarta 39,99% e assim por diante. Em outras palavras, a diminuição relativa da frequência dos alelos *a* e *b*, que na primeira geração filial é de 6,66%, pois  $\frac{0,5 - 0,4667}{0,5} = 0,0666$ , passará a ser de 5,70% na segunda geração filial, 4,95% na terceira geração, 4,40% na quarta, e assim por diante, sendo, por isso, muito difícil a eliminação de tais genes.

Essa demonstração serve, portanto, de argumento poderoso para falar contra o valor que as medidas drásticas de esterilização de indivíduos com anomalias recessivas com determinação poligênica teriam para as populações humanas. De fato, mesmo supondo que essas medidas pudessem ser aplicadas a todos os anômalos, o que é pouco provável, seria remota a possibilidade de extinguir novos casos nas gerações seguintes, porque sempre haveria uma alta proporção de portadores dos genes cuja eliminação se pretende. Em nosso exemplo, pode-se constatar que a grande maioria dos indivíduos da população inicial possui os genes *a* ou *b* em heterozigose, pois  $1 - (AABB + aabb) = 1 - (0,0625 + 0,0625) = 0,875$  ou 87,5%.

### SELEÇÃO CONTRA ANOMALIAS INCOMPLETAMENTE RECESSIVAS

Consideremos dois pares de alelos autossômicos *A,a* e *B,b* que, numa geração inicial de uma população em equilíbrio de Hardy e Weinberg apresentam as frequências  $A = B = p$  e  $a = b = q = 1 - p$ . Suponhamos, ainda, que os genes *a* e *b* provocam, cada qual, uma anomalia, quando em homozigose (*aa* e *bb*) e aceitemos que apenas o gene *a* possa se expressar em parte dos heterozigotos *Aa*. Em outras palavras, consideremos que uma fração dos heterozigotos *Aa* possa manifestar alguns ou todos os sinais da anomalia presente nos homozigotos *aa*, mas que fenômeno análogo não ocorra nos heterozigotos *Bb*.

Tendo em vista que a anomalia causada pelo gene *b* é, por hipótese, completamente recessiva, e que o gene *a* exerce efeitos desvantajosos nos heterozigotos *Aa*, é claro que, se os homozigotos *aa* e *bb* forem selecionados com a mesma intensidade, o efeito seletivo contra o gene *a* será mais eficiente do que contra o gene *b*.

Designando por *h* a frequência de heterozigotos *Aa* ou *Bb* sujeitos à seleção em decorrência da capacidade de expressão dos alelos *a* ou *b*, ter-se-á  $h > 0$  em relação aos heterozigotos *Aa* e  $h = 0$  em relação aos heterozigotos *Bb*. Em vista disso considerar-se-á que a seleção natural operará selecionando tanto os homozigotos *aa* quanto os *bb* com frequência  $sq^2$  e os heterozigotos *Aa* com frequência  $sh2pq$ .

Com base no exposto não é difícil demonstrar que a eliminação do gene  $a$  pela seleção que opera contra os heterozigotos será tanto mais eficiente, quanto maior for a frequência de heterozigotos em que esse gene se expressa ( $h$ ), e quanto menor for a frequência desse alelo na população. De fato, a relação entre os efeitos da seleção por intermédio dos heterozigotos  $Aa$  e homozigotos  $aa$  permite escrever  $\frac{sh2pq}{sq^2} = \frac{2hp}{q}$ .

Por outro lado, sabendo que essa razão exprime a comparação entre a seleção contra os zigotos  $Aa$  (numerador) e os zigotos  $aa$  (denominador), tem-se que a eliminação gamética, isto é, de genes, será igual à metade desse valor. Isso permite escrever que a eliminação do gene  $a$  pelos heterozigotos  $Aa$  em comparação aos homozigotos  $aa$  se faz segundo  $\frac{hp}{q}$ . Essa relação pode ser mais simplificada, ainda, se levarmos em conta que no caso de anomalias recessivas, a frequência  $p$  pode ser considerada semelhante à unidade, pois  $q = 1 - p$  tem valor muito baixo.

Para tornar ainda mais claro o que foi exposto tomemos um exemplo numérico. Assim, consideremos um par de alelos autossômicos  $A, a$ , com frequência  $p = 0,99$  e  $q = 0,01$ , e que o alelo  $a$  determine uma anomalia letal em todos os homozigotos  $aa$  e em metade dos heterozigotos  $Aa$ . Nesse caso, a eliminação do gene  $a$  por intermédio dos heterozigotos  $Aa$  seria cerca de 50 vezes mais alta do que por meio dos homozigotos  $aa$  pois  $\frac{hp}{q} = \frac{0,5 \times 0,99}{0,01} = 49,5$  ou

$\frac{h}{q} = \frac{0,5}{0,01} = 50$ . Se, nesse exemplo, o valor de  $h$  fosse menor é evidente que a eliminação por

intermédio dos heterozigotos seria menos intensa. Assim, se  $h$  fosse igual a 25% ter-se-ia

$\frac{hp}{q} = \frac{0,25 \times 0,99}{0,01} = 24,75$  ou  $\frac{h}{q} = \frac{0,25}{0,01} = 25$ . Consideremos, agora, que a frequência do gene  $a$  é

$q = 0,005$ . Nesse caso, para valores de  $h = 0,5$  e  $h = 0,25$  teríamos que a eliminação do gene  $a$  por intermédio dos heterozigotos seria mais eficiente que no exemplo anterior, pois  $\frac{h}{q}$  passa a ser

igual a 100 quando  $h = 0,5$  e igual a 50 quando  $h = 0,25$ .

## SELEÇÃO CONTRA HETEROZIGOTOS

Para analisar as conseqüências genéticas da seleção contra heterozigotos, consideremos um par de alelos autossômicos  $A, a$  e aceitemos que os homozigotos  $AA$  e  $aa$  estão sujeitos à mesma pressão seletiva, enquanto os heterozigotos  $Aa$  sofrem maior seleção do que os homozigotos.

Se a frequência desses dois alelos for igual, isto é, se  $p = q = 0,5$ , mesmo que os heterozigotos sofram seleção total, as frequências gênicas permanecerão inalteradas, porque, com

a eliminação de cada heterozigoto haverá a eliminação simultânea dos dois alelos, isto é, não haverá seleção preferencial de um deles. Contudo, se  $p > q$  ou  $p < q$ , a seleção contra os heterozigotos tenderá a eliminar o alelo com menor frequência na população, pois **a taxa de seleção será a mesma** em relação a alelos que ocorrem com **freqüências diferentes**.

Para exemplificar, consideremos que um par de alelos autossômicos  $A,a$  ocorra com frequência  $p = 0,70$  e  $q = 0,30$  em uma população em equilíbrio genético, de sorte que  $AA = 0,49$ ,  $Aa = 0,42$ ,  $aa = 0,09$ . Se os heterozigotos  $Aa$  passarem a sofrer seleção total, tem-se que, em uma geração, a frequência do alelo  $a$  passará de  $q = 0,30$  para  $q_1 = 0,1552$  porque, com a eliminação dos heterozigotos, a frequência de  $aa$  será  $\frac{0,09}{0,58} = 0,1552$  enquanto que a frequência de  $AA$  será relativamente maior, isto é,  $\frac{0,49}{0,58} = 0,8448$ . Se o processo de seleção total dos heterozigotos  $Aa$  continuar, a frequência do alelo  $a$  baixará rapidamente para  $q_2 = 0,0327$  e  $q_3 = 0,0012$  até que se dê sua eliminação da população.

A seleção contra os heterozigotos  $Dd$  do sistema Rh que, antes do advento da globulina anti-Rh<sub>0</sub>, operou nas populações humanas, se enquadra no tipo de seleção que acabamos de expor, pois durante milênios as crianças Rh-positivo  $Dd$  geradas por mães múltiparas Rh-negativo ( $dd$ ) estiveram sujeitas à eritroblastose fetal. É, pois, surpreendente que o alelo  $d$  desse sistema ainda seja mantido com frequências altas nas populações humanas, e que elas variem em torno de 35% a 40% em numerosas populações caucasóides. Várias hipóteses já foram propostas para explicar esse achado, as quais, entretanto, não são mutuamente exclusivas.

Uma delas é a de que em eras passadas certas populações humanas teriam sido compostas predominantemente por indivíduos Rh positivo e outras por indivíduos Rh negativo. Assim, as populações atuais teriam surgido por miscigenação entre aqueles grupamentos humanos, de sorte que a seleção ainda estava se processando em muitas delas, com a finalidade de eliminar o alelo  $d$ . A favor dessa hipótese tem-se as populações da China e do Japão, nas quais a frequência do alelo  $d$  do sistema Rh é muito baixa ou mesmo nula, e os bascos, nos quais a frequência desse alelo é maior que a do gene  $D$ . Aliás, os bascos são considerados como remanescentes de uma população européia primitiva.

Uma outra hipótese considera que os indivíduos Rh positivo heterozigotos ( $Dd$ ) teriam maior valor adaptativo que os homozigotos  $DD$  ou  $dd$ , de modo que a perda dos heterozigotos por eritroblastose fetal seria compensada pelo maior valor adaptativo dos que passassem pelo crivo da seleção natural.

Para explicar a manutenção de altas frequências do alelo  $d$  do sistema Rh também se supôs que as mães Rh negativo casadas com homens Rh positivo teriam a necessidade

psicológica de gerar muitos filhos, para compensar a perda daqueles com eritroblastose fetal. Assim, as mulheres Rh negativo casadas com heterozigotos  $Dd$  teriam, a possibilidade de repor os alelos  $d$ , perdidos por intermédio dos heterozigotos  $Dd$ , ao gerar mais filhos Rh negativo ( $dd$ ).

Finalmente, apesar de não existir qualquer evidência a favor, existe a hipótese de que a taxa de mutação do gene  $D$  para seu alelo  $d$  seria alta.

## SELEÇÃO A FAVOR DE HETEROZIGOTOS

Para compreendermos o resultado da seleção a favor de heterozigotos, consideremos uma população em equilíbrio de Hardy e Weinberg em relação a um par de alelos  $A, a$ , cujas freqüências são iguais, respectivamente, a  $p$  e  $q = 1 - p$ . Consideremos, ainda, que o genótipo homozigoto  $AA$  passe a sofrer seleção de intensidade  $s_1$  e o genótipo homozigoto  $aa$  fique sujeito a seleção de intensidade  $s_2$ . A seleção contra o genótipo heterozigoto  $Aa$  será considerada nula, porque a viabilidade desse genótipo será tomada como referência para a dos genótipos homozigotos  $AA$  e  $aa$ .

Nesse caso, a freqüência do alelo  $a$  na primeira geração após seleção será:

$$q_1 = \frac{q^2 - s_2q^2 + pq}{1 - (s_1p^2 + s_2q^2)} = \frac{q(p+q) - s_2q^2}{1 - (s_1p^2 + s_2q^2)} = \frac{q - s_2q^2}{1 - (s_1p^2 + s_2q^2)}$$

de sorte que a diminuição da freqüência desse alelo após uma geração de seleção poderá ser escrita como  $q - q_1 = q \cdot \frac{q - s_2q^2}{1 - (s_1p^2 + s_2q^2)}$  ou, mais simplificada,  $q - q_1 = \frac{pq(s_2q - s_1p)}{1 - (s_1p^2 + s_2q^2)}$ .

Nessa equação é fácil constatar que, se persistirem as mesmas condições de seleção, as freqüências gênicas cessarão sua variação, atingindo equilíbrio quando  $s_1p$  for igual a  $s_2q$ , o que promoverá a nulidade da variação da freqüência do alelo  $a$  através das gerações. Quando isso acontecer pode-se escrever  $s_1p = s_2q$  ou  $s_1(1 - q) = s_2q$ , de onde se tira  $q = \frac{s_1}{s_1 + s_2}$ . Também se

pode escrever, é claro,  $s_1p = s_2(1 - p)$  de onde se extrai  $p = \frac{s_2}{s_1 + s_2}$ .

Essas últimas fórmulas permitem verificar facilmente que, se os homozigotos  $AA$  e  $aa$  estiverem sujeitos à mesma pressão seletiva, isto é, se  $s_1 = s_2$ , as freqüências dos alelos  $A$  e  $a$  se igualarão ( $p = q = 0,5$ ), independentemente do valor dos coeficientes seletivos, ou seja, de a seleção ser total ou parcial.

## A MANUTENÇÃO DOS POLIMORFISMOS GENÉTICOS NEUTROS

As anomalias com transmissão monogênica, dominante ou recessiva, não oferecem dificuldades para explicar sua ocorrência nas populações, porque é fácil admitir que os genes raros que as determinam são consequência de mutação dos alelos normais, sendo os genes raros mantidos com frequência baixa pela ação da seleção natural. Entretanto, como explicar a ocorrência de sistemas de alelos frequentes, em que nenhum dos genes polimórficos produz qualquer efeito anormal evidente, estando, assim, seus portadores sujeitos a coeficiente seletivo nulo (*polimorfismos neutros*)?

Do ponto de vista teórico é fácil demonstrar que um polimorfismo neutro pode ser mantido graças, apenas, à pressão de mutação (*polimorfismo neutro mutacional*). Para tanto, basta considerar uma população na qual os alelos autossômicos  $A, a$ , com frequências, respectivamente,  $p$  e  $q = 1 - p$ , conferem o mesmo valor adaptativo a seus portadores. Se designarmos por  $\mu$  a taxa de mutação de  $a$  para  $A$ , e por  $\nu$  a taxa de mutação inversa, isto é, de  $A$  para  $a$ , o equilíbrio genético estável será atingido quando a frequência de alelos  $a$  surgidos por mutação ( $\mu p$ ) for igual à de alelos  $A$  surgidos por mutação inversa ( $\nu q$ ). Quando isso acontecer ter-se-á  $\mu p = \nu q$  ou  $\mu(1 - q) = \nu q$ , fórmula essa que permite escrever  $\mu - \mu q = \nu q$ , de onde se tira  $q(\mu + \nu) = \mu$  e, finalmente,  $q = \frac{\mu}{\mu + \nu}$  o que permite dizer que o equilíbrio genético estável da

população será atingido quando a frequência do alelo  $a$  for  $q = \frac{\mu}{\mu + \nu}$ . Assim, por exemplo, se a taxa de mutação de  $A$  para  $a$  fosse  $\mu = 1,4 \times 10^{-6}$  e a taxa de mutação inversa fosse  $\nu = 1,1 \times 10^{-6}$ , o equilíbrio genético estável seria atingido quando a frequência do alelo  $a$  fosse igual a 56%, pois  $q = \frac{1,4}{1,4 + 1,1} = 0,56$ .

Na prática, porém, certos dados fazem restrições à aceitação da existência de polimorfismos neutros mutacionais. Assim, se os pequenos grupamentos humanos do passado não tivessem sofrido tão intensamente os efeitos da deriva genética, que serão estudados em outro tópico deste capítulo, as populações atuais não deveriam diferir significativamente entre si quanto às frequências de genes aparentemente neutros. No entanto, os vários sistemas de grupos sanguíneos, de proteínas séricas e de enzimas eritrocitárias, que, como se sabe, são polimórficos, são prova eloqüente de que as populações humanas diferem muito entre si quanto à distribuição das frequências dos alelos que compõem esses sistemas.

Por outro lado, admitindo a ação da deriva genética no passado, deveríamos observar nas populações humanas atuais variação muito maior das freqüências gênicas dos sistemas polimórficos estudados. A clássica análise feita por Alice Brues (1954), a respeito da limitação da distribuição das freqüências dos genes *A* e *B* do sistema sanguíneo eritrocitário ABO em 251 populações, serve bem para ilustrar essa afirmação. A representação gráfica da distribuição dessas freqüências na Figura 1.6 evidencia logo que somente cerca de 20% da área do triângulo correspondente às freqüências possíveis dos alelos *A* e *B*, foram ocupados pelas freqüências observadas, o que não deixa de ser surpreendente. Uma situação semelhante foi assinalada por Saldanha (1964) em relação ao polimorfismo da reação gustativa à feniltiouréia, pois a freqüência do gene responsável pela insensibilidade a essa substância nunca ultrapassa 60%.

Em conseqüência desses argumentos, aventou-se a hipótese de que a existência e a manutenção dos genes polimórficos que não produzem efeito anormal evidente também deveriam ser conseqüência de um processo seletivo. Em outras palavras, os polimorfismos neutros não existiriam, porque os portadores de genes polimórficos que, aparentemente, não têm efeito anormal, também estariam sujeitos à seleção, apesar de o processo seletivo contra eles não ser tão drástico quanto o que opera contra os portadores de genes responsáveis por anomalias.

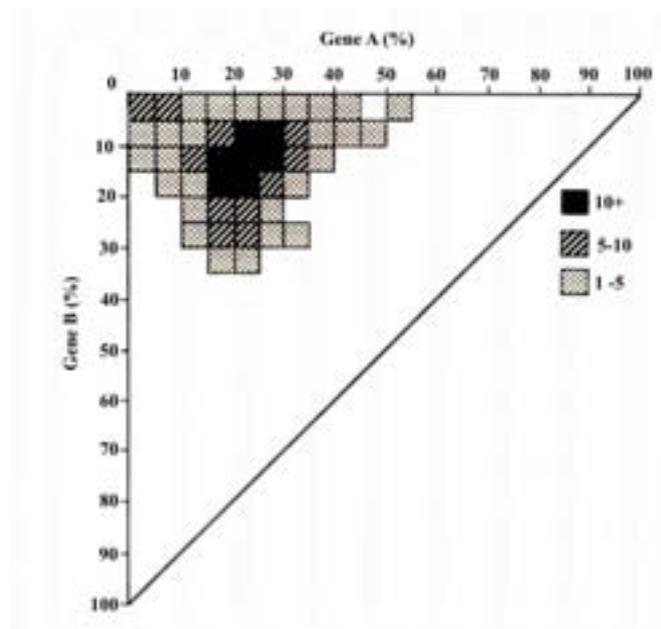


Fig. 1.6. Variação das freqüências dos genes *A* e *B* do sistema sanguíneo eritrocitário ABO observadas em 215 populações, dentro da área correspondente à variação possível (Brues, 1954)..

Essa hipótese começou a ganhar força quando se constatou que os grupos sanguíneos do sistema ABO mostravam associação com doenças, isso indicando que, entre os efeitos pleiotrópicos desses genes, estaria incluída a predisposição a diversas moléstias. A associação do grupo sanguíneo A com carcinoma gástrico e outros tipos de câncer, e com a anemia perniciosa, ou a associação do grupo O com a úlcera gástrica e a úlcera duodenal (Aird *et al.*, 1953; Roberts,

1957) não resultariam em efeito seletivo, porque essas doenças incidem, geralmente, em pessoas que já ultrapassaram a idade reprodutiva. Entretanto, a associação do grupo sanguíneo A e a hanseníase virchowiana (lepromatosa) e do grupo O com a hanseníase tuberculóide (Beiguelman, 1963, 1964<sub>a</sub>; Yankah, 1965; Vogel e Chakravarti, 1966) resultariam em efeito seletivo, porque a hanseníase pode causar esterilidade dos indivíduos do sexo masculino, em consequência de orqui-epididimite.

Situações semelhantes foram observadas no polimorfismo da reação gustativa à feniltiouréia. Assim, em pacientes com bócio nodular, a proporção de insensíveis à feniltiouréia foi significativamente mais alta do que a esperada (Harris *et al.*, 1949; Kitchin *et al.*, 1959), ao passo que os sensíveis a essa tiouréia foram mais freqüentes do que se esperava em indivíduos com bócio difuso, em tuberculosos, em hansenianos e em hanseníase com tuberculose (Saldanha, 1956, 1964; Kitchin *et al.*, 1959; Beiguelman, 1962<sub>b</sub>, 1964<sub>b</sub>, 1964<sub>c</sub>).

Outras observações falaram a favor de que forças seletivas de pouca intensidade operariam sobre os alelos aparentemente neutros de vários sistemas. Assim, por exemplo, em relação ao sistema sanguíneo ABO, verificou-se que dos casais compostos por mulher sem anticorpos séricos correspondentes aos antígenos ABH das hemácias dos maridos, isto é, dos **casais compatíveis** no sistema ABO resultaram mais filhos do que dos **casais incompatíveis** nesse sistema (Tabela 2.6).

Tabela 2.6. Casais compatíveis e incompatíveis no sistema sanguíneo ABO.

Compatíveis Marido × Mulher		Incompatíveis Marido × Mulher	
O	O	A	O
O	A	A	B
O	B	B	O
O	AB	B	A
A	A	AB	O
A	AB	AB	A
B	B	AB	B
B	AB		
AB	AB		

Visto que os casais incompatíveis podem gerar filhos com antígenos eritrocitários para os quais as mães têm anticorpos séricos correspondentes, a maior taxa de abortos desses casais poderia ser explicada por incompatibilidade sanguínea materno-fetal no sistema ABO. Assim, no Japão, verificou-se que a taxa de mortes intra-uterinas antes do sexto mês de gestação foi de 11,8% nos casais constituídos por marido O e mulher A, enquanto que nos casais compostos por marido A e mulher O essa taxa foi de 17,1% (Matsunaga, 1955; Matsunaga e Itoh, 1958).

Apesar dos argumentos aqui expostos, existe, atualmente, a tendência a aceitar a ocorrência de polimorfismos neutros mutacionais, porque um gene pode ser considerado neutro se o produto do coeficiente seletivo de seus portadores pelo tamanho efetivo da população for muito inferior à unidade (Kimura, 1968). Desse modo, um mesmo gene pode ser neutro ou praticamente neutro em uma população com pequeno tamanho efetivo e não ser neutro em outra, com grande tamanho efetivo.

Um outro tipo de polimorfismo neutro, que sempre teve aceitação fácil, é o **polimorfismo neutro migratório**. Realmente, não existem obstáculos à admissão de que uma população composta, apenas ou quase totalmente, por homozigotos de um gene neutro qualquer  $A$  se tome polimórfica porque recebe alta proporção de um alelo  $a$ , cujos portadores têm o mesmo valor adaptativo do gene  $A$ , em decorrência de miscigenação com outra, que possui o alelo  $a$  em proporção elevada. Assim, por exemplo, nas populações européias ou de origem européia, com exceção dos judeus sefaraditas da Turquia, a frequência de indivíduos com o antígeno eritrocitário  $Di^a$  do sistema Diego pode ser considerada nula. Em índios, japoneses e chineses, entretanto, esse antígeno é polimórfico, pois pode atingir frequências iguais a 36% em índios, 8% a 12% em japoneses e 2,5% em chineses (Layrisse *et al.*, 1955; Layrisse e Arends, 1956; Lewis *et al.*, 1956). A presença do gene  $Di^a$  em frequência alta nas populações caucasóides de alguns países das Américas é, sem dúvida, o resultado do fluxo gênico de índios nessas populações.

## A MANUTENÇÃO DOS POLIMORFISMOS ADAPTATIVOS

Na espécie humana reconhecem-se alguns polimorfismos cuja manutenção somente pode ser explicada se admitirmos a existência de pressão seletiva. Dentre tais polimorfismos, denominados **polimorfismos adaptativos**, o mais estudado na espécie humana foi o **polimorfismo equilibrado**, cuja manutenção somente pode ser explicada levando em conta a desvantagem ou a vantagem seletiva dos heterozigotos em relação aos homozigotos. Esse tipo de polimorfismo também tem sido chamado por alguns de **polimorfismo balanceado**, numa tradução imperfeita do inglês (*balanced polymorphism*).

Os aspectos teóricos da seleção contra e a favor dos heterozigotos já foram discutidos em tópicos anteriores. Por isso, basta enfatizar, aqui, que a seleção contra heterozigotos somente conduz a uma estabilidade nas frequências genotípicas, isto é, a um polimorfismo equilibrado, no caso particular de os dois alelos de um par terem frequências iguais. Já a seleção a favor dos heterozigotos conduz sempre a um polimorfismo equilibrado.

O exemplo mais divulgado de polimorfismo equilibrado por seleção a favor de heterozigotos é o da hemoglobina S nas populações africanas e em algumas áreas do sul da Europa (Sicília, Calábria, Grécia) e do sul da Ásia. Visto que, nessas populações, os indivíduos

com anemia falciforme (homozigotos  $SS$ ) nunca atingiam a idade reprodutiva, é claro que a frequência do gene responsável por essa hemoglobinopatia deveria corresponder apenas à taxa de mutação. No entanto, esse gene atinge proporções muito altas em algumas populações africanas (Allison, 1954<sub>b</sub>), entre as quais é comum o encontro de frequências iguais ou superiores a 10%. Isso equivale a dizer, pois, que, para explicar a manutenção do gene da hemoglobina S nessa frequência tão alta, deveríamos supor a taxa de mutação de 1%, o que, evidentemente, é um absurdo.

Em vista disso, recorreu-se à hipótese de que os indivíduos com o traço ciclêmico (heterozigotos  $AS$ ) teriam uma vantagem seletiva em relação aos homozigotos do gene da cadeia beta de hemoglobina normal ( $AA$ ) e, é claro, em relação aos indivíduos afetados pela anemia falciforme ( $SS$ ). Essa vantagem ocorreria em áreas tropicais e subtropicais onde a malária causada pelo *Plasmodium falciparum* é ou era frequente. Como se sabe, nas áreas holo-endêmicas, a malária causada por *P. falciparum* é o tipo mais letal em crianças entre os seis meses e os três anos de idade. Antes dos seis meses de vida as crianças dessas regiões são imunes à malária, porque apresentam imunidade adquirida passivamente das mães. Por outro lado, após o terceiro ano de idade, o quadro clínico tende a ser menos drástico, mesmo em indivíduos com alta concentração de parasitas no sangue. Em outras palavras, nessas regiões, as crianças com anemia falciforme (homozigotos  $SS$ ) estariam condenadas à morte precoce por causa da própria hereditária, enquanto que as crianças com hemoglobina totalmente normal (homozigotos  $AA$ ) teriam maior probabilidade do que as heterozigotas  $AS$  de ir a óbito por malária causada por *P. falciparum*.

Essa hipótese ganhou apoio porque:

1. no período entre os seis meses e os três anos de idade as crianças de regiões endêmicas de *P. falciparum* revelaram que os heterozigotos  $AS$  tinham menor parasitemia que os homozigotos  $AA$  (Allison, 1954<sub>b</sub>);

2. em crianças com mais de mil parasitas por milímetro cúbico de sangue, a frequência de heterozigotos  $AS$  (13,7% em 386 crianças) foi, praticamente, a metade da esperada (24,4% em 1.180 controles) na hipótese de inexistência de seleção (Vandepitte e Delaisse, 1957);

3. reunindo os dados de vários autores a respeito de óbitos infantis em diversas populações africanas, causados por malária predominantemente do tipo cerebral, verificou-se que, dentre 100 desses óbitos, havia apenas um caso com traço ciclêmico, o que contrasta com o número teórico esperado de heterozigotos  $AS$  (22,6%) na hipótese de inexistência de seleção preferencial (Motulsky, 1964).

Além desses, existem, ainda, os dados de Allison (1954<sub>a</sub>) sobre os resultados das inoculações de *P. falciparum* em 30 negros africanos de uma área hiper-endêmica de malária.

(Teriam eles se submetido voluntária e conscientemente a essas inoculações?). Dentre os indivíduos inoculados 15 eram heterozigotos *AS* e 15 eram homozigotos *AA* (controles). Dos 15 portadores do traço siclêmico apenas dois contraíram malária e, mesmo assim, a contagem de parasitas no sangue foi pequena, ao passo que todos os controles manifestaram malária e alguns casos precisaram de tratamento. Beutler *et al.* (1955), entretanto, não conseguiram obter resultado semelhante ao repetir o mesmo experimento em prisioneiros negros norte-americanos, já que não encontraram diferença significativa entre o grupo controle e o grupo com traço siclêmico.

Uma explicação plausível para a proteção conferida pela hemoglobina S contra a malária causada pelo *P. falciparum* pode ser a de que as hemácias com os trofozoítos ficariam retidas nos capilares durante um tempo suficiente para causar hipóxia, devida tanto à perda de oxigênio para os tecidos quanto para os parasitas. Esse acontecimento, aliado ao fato de que as hemácias com os trofozoítos de *P. falciparum* tendem a aderir às paredes dos capilares dos tecidos periféricos, facilitaria a falciformação das hemácias parasitadas, as quais, por isso, seriam destruídas (Miller *et al.*, 1956).

A explicação dada para a manutenção do polimorfismo da hemoglobina S foi estendida à hemoglobina C, que ocorre em alta freqüência na África Ocidental, principalmente em Gana, e à hemoglobina E, que tem freqüência alta no Sudoeste da Ásia, apesar de os homozigotos CC e EE não manifestarem complicações tão graves quanto os homozigotos SS. Do mesmo modo, a alta freqüência da talassemia beta nas populações vizinhas ao Mar Mediterrâneo e nas asiáticas tem sido explicada como uma conseqüência da melhor viabilidade dos heterozigotos em regiões onde a malária causada pelo *P. falciparum* foi endêmica.

Aliás, é importante lembrar que foi a coincidência da distribuição geográfica da talassemia beta e, no passado, do *P. falciparum* (Figura 2.6) que levou o grande geneticista, bioquímico e humanista inglês John Burdon Sanderson Haldane a sugerir a hipótese de que os indivíduos com talassemia heterozigótica teriam vantagem seletiva sobre os homozigotos (Haldane, 1949). Como se sabe, os indivíduos com talassemia beta homozigótica apresentam uma anemia grave (*anemia de Cooley* ou *talassemia major*) com alterações eritrocitárias (alvócitos, microcitose, anisocitose, hipocromia, aumento da resistência osmótica), esplenomegalia, hepatomegalia, além de outras alterações que podem prejudicar a sobrevivência até a idade reprodutiva. Os homozigotos do gene que confere ritmo de síntese normal da cadeia beta de hemoglobina seriam, por sua vez, mais selecionados pela malária do que os heterozigotos assintomáticos (*talassemia mínima*) ou que aqueles com uma forma pouco grave de talassemia (*talassemia minor*). Tal proteção poderia decorrer das necessidades nutricionais do *P.*

*falciparum*, o qual, eventualmente, poderia encontrar obstáculos ao seu desenvolvimento nas hemácias dos heterozigotos do gene da talassemia beta.

A ocorrência de populações com alta frequência de indivíduos com deficiência de G-6PD também foi atribuída à malária causada pelo *P. falciparum* (Motulsky, 1960<sub>a</sub>, 1960<sub>b</sub>), por causa da grande correlação entre a distribuição geográfica das antigas áreas malarígenas e das populações com alta frequência de homens com deficiência de G-6PD. Essa correlação chega a ser tão grande que, na Sardenha, a frequência de homens com deficiência de G-6PD é relativamente baixa nos locais montanhosos (3% a 4%) e muito alta no litoral dessa grande ilha (Siniscalco *et al.*, 1961), onde a malária foi endêmica, podendo atingir 35%. Além disso, nas populações africanas, a frequência da variante A<sup>-</sup> de G-6PD mostra forte correlação positiva com o gene da hemoglobina S e, de modo análogo, na Sardenha, a frequência da variante mediterrânea de G-6PD está correlacionada positivamente com a do gene da talassemia beta (Siniscalco *et al.*, 1961; Motulsky, 1964). Essas correlações ganham grande significação pelo fato de que a deficiência de G-6PD é ligada ao sexo, ao passo que as hemoglobinopatias em discussão são autossômicas.

A maior preservação dos portadores da deficiência de G-6PD pelo *P. falciparum* talvez seja uma consequência de esse parasito desenvolver-se nas hemácias mais velhas, as quais, nos indivíduos com a variante africana deficiente de G-6PD, são as que mostram enzimopenia. Assim, sabendo-se que as hemácias deficientes de G-6PD podem ter níveis muito baixos de glutatião reduzido (Beutler, 1959), que o *P. falciparum* se utiliza da via oxidativa direta da glicólise (Geiman, 1951), que tal parasito requer glutatião reduzido para crescer *in vitro* (McGhee e Trager, 1950) e que 50% desse glutatião das hemácias contribuem para a produção da cisteína necessária a ele (Fulton e Grant, 1956), pode-se supor que o *P. falciparum*, por causa de suas necessidades nutricionais, também tenha preservado não só os homens deficientes de G-6PD, mas, também, as mulheres heterozigotas, do gene da deficiência, já que elas possuem duas populações de hemácias (normais e deficientes de G-6PD). Num ambiente com *P. falciparum*, as mulheres heterozigotas poderiam ter uma vantagem seletiva em relação aos homens com deficiência de G-6PD porque estes estariam sujeitos a crises hemolíticas e elas não.

Os polimorfismos equilibrados, decorrentes da seleção a favor dos heterozigotos não implicam, obrigatoriamente, em maior valor adaptativo dos heterozigotos em todos os ambientes, ou, o que dá no mesmo, em um ambiente constante (**polimorfismo equilibrado univalente**). De fato, se os homozigotos de um dos alelos de um par autossômico *A,a* forem favorecido em um grupo de nichos ecológicos, enquanto os homozigotos do outro alelo forem favorecidos em outro grupo de nichos, pode acontecer que, apesar de os heterozigotos não terem o maior valor adaptativo em nenhum dos nichos ecológicos, a média ponderada da viabilidade dos

heterozigotos será, no conjunto, superior à dos homozigotos. Nesse caso, o polimorfismo será dito *polimorfismo equilibrado ambivalente* e será classificado como *espacial* se as diferenças seletivas se distribuírem no espaço e *estacional* ou *sazonal* se elas se distribuírem no tempo (Saldanha, 1964).

Contrastando com o que ocorre nos polimorfismos adaptativos do tipo equilibrado, que são mantidos por um mecanismo homeostático, existem os polimorfismos adaptativos que são fadados à extinção (*polimorfismos transitórios*). Um polimorfismo transitório poderá existir quando os heterozigotos estão em desvantagem seletiva em relação aos homozigotos e os alelos têm freqüências diferentes. Nesse caso, como sabemos, o alelo menos freqüente desaparecerá da população. Um polimorfismo transitório pode, também, ser conseqüência de mutações recorrentes, tendo o mutante vantagem seletiva sobre seu alelo. Durante a fase de substituição do alelo mais antigo pelo mutante existirá o polimorfismo. É importante assinalar, porém, que nem sempre um polimorfismo transitório deve ser adaptativo. Realmente, pode-se supor que um polimorfismo transitório seja conseqüência de deriva genética ou de migração diferencial, sem que essas causas estejam associadas, obrigatoriamente, a um processo seletivo.

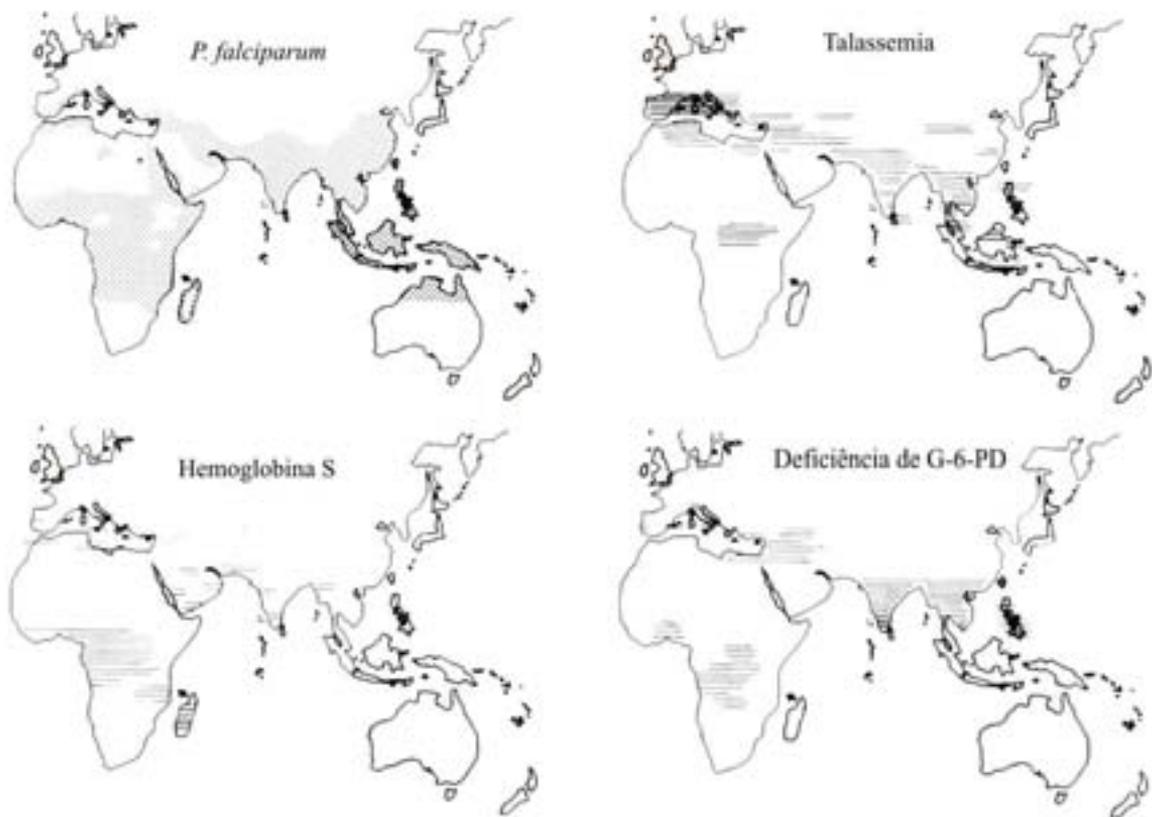


Fig.2.6. Antiga distribuição geográfica do *Plasmodium falciparum*, da talassemia, da hemoglobina S e da deficiência de G-6PD.

## 2. A DERIVA GENÉTICA

No tópico final do capítulo sobre “Efeito da consangüinidade” foi frisado que os efeitos genéticos resultantes do isolamento dependem do tamanho dos isolados. Nessa ocasião, tivemos oportunidade de verificar que, se uma população for dividida em isolados grandes, o efeito do isolamento é semelhante ao dos casamentos consangüíneos, isto é, há aumento da freqüência de homozigotos na população (*efeito Wahlund*). Nesse caso, portanto, o isolamento não tem efeito evolutivo, pois somente há alteração das freqüências genotípicas, mas não das gênicas.

Nos isolados pequenos, contudo, o isolamento tem efeito evolutivo porque tais grupamentos não conseguem manter um equilíbrio estável da distribuição genotípica, segundo a lei de Hardy e Weinberg, mesmo supondo panmixia na ausência de migração, mutação e seleção. Nessas pequenas subpopulações, a variação aleatória das freqüências gênicas ao longo das gerações pode provocar a eliminação ou, ao contrário, a fixação casual de um gene, independentemente de seu coeficiente seletivo. Em outras palavras, nos pequenos isolados, um gene com alto valor adaptativo pode desaparecer, enquanto outro, com baixo valor adaptativo pode ser fixado. Tal fenômeno, que, em inglês, é denominado *genetic drift*, foi chamado, em português, durante algum tempo, de *oscilação genética*, e depois, renomeado, mais apropriadamente, por Freire-Maia (1974), de *deriva genética*.

Não é difícil demonstrar, algebricamente, que ao longo das gerações de uma pequena população possa ocorrer a eliminação ou a fixação completa de um gene como decorrência, simplesmente, da variação casual de sua freqüência. De fato, cada geração pode ser considerada como uma amostra aleatória de  $2n$  gametas retirados da geração precedente de  $n$  indivíduos geneticamente ativos. Desse modo, por exemplo, se os alelos  $A$  e  $a$  estiverem presentes em uma geração de uma população, com freqüências  $p$  e  $q=1-p$ , respectivamente, eles terão 95% de probabilidade de se apresentar na geração seguinte com freqüências contidas nos intervalos praticamente iguais a dois desvios padrão acima e abaixo das freqüências  $p$  e  $q$ , ou seja, nos intervalos  $p \pm 2\sigma$  e  $q \pm 2\sigma$ . (O desvio padrão de  $p$  ou de  $q$  é calculado por intermédio de  $\sigma = \sqrt{\frac{pq}{2n}}$ , no caso de os alelos serem autossômicos, ou por intermédio de  $\sigma = \sqrt{\frac{pq}{3n}}$  se eles pertencerem ao cromossomo X).

Visto que o desvio padrão de  $p$  ou de  $q$  depende do valor de  $n$ , está claro que, se  $n$  for grande será pouco provável que a variação das freqüências gênicas provoque desvios significativos das freqüências originais nas gerações que se sucedem, desde que se admita ausência de migração, seleção e mutação. Entretanto, se  $n$  for pequeno será bastante provável

que, entre uma geração e outra, ocorram desvios significativos das freqüências  $p$  e  $q$  em relação às da geração anterior.

Com a finalidade de ilustrar essa discussão consideremos uma geração inicial de duas populações teóricas (A e B), nas quais os alelos autossômicos  $A$  e  $a$  têm freqüências  $p = q = 0,5$ . Consideremos, ainda, que as populações A e B são constituídas por um número constante de casais igualmente férteis, que geram o mesmo número de filhos ao longo das gerações, mas que o número de casais da população A é 500.000, enquanto na população B existem apenas 25 casais.

Visto que na população A  $n = 1.000.000$  e na população B  $n = 50$ , sendo  $n$  a população geneticamente ativa, tem-se, portanto, que, na primeira, o desvio padrão de  $q$  será

$$\sigma = \sqrt{\frac{0,5 \times 0,5}{2 \times 1.000.000}} = 0,00035, \text{ enquanto na segunda ele será estimado como } \sigma = \sqrt{\frac{0,5 \times 0,5}{2 \times 50}} = 0,05.$$

Em conseqüência disso, na geração seguinte à inicial da população A, a freqüência do alelo  $a$  terá 95% de probabilidade de estar contida no intervalo muito estreito de 0,4993 e 0,5007. Em oposição, na primeira geração filial da população B, a freqüência do alelo  $a$  tem 95 % de probabilidade de estar contida em intervalo mais amplo, isto é, de 0,4 a 0,6.

Em outras palavras, enquanto a freqüência do alelo  $a$  tem pouca probabilidade de ficar significativamente alterada ao longo das gerações da população A, tal probabilidade é grande na população B. Assim, se na primeira geração filial da população B a freqüência do alelo  $a$  fosse, por acaso,  $q = 0,45$ , ter-se-ia que, na segunda geração filial dessa população, o gene  $a$  teria 95% de probabilidade de ocorrer com uma freqüência contida no intervalo entre 0,35 e 0,55, pois

$$\sigma = \sqrt{\frac{0,45 \times 0,55}{2 \times 50}} = 0,0497.$$

Diante do exposto, não é difícil vislumbrar que, se o mesmo processo continuar, pode-se chegar à extinção do gene  $a$ , com a conseqüente fixação de seu alelo  $A$ . Por outro lado, considerando a possibilidade de ocorrência de um aumento casual do gene  $a$ , pode-se imaginar, também, um fenômeno oposto, ou seja, a fixação desse gene e a eliminação de seu alelo  $A$ . Quanto menor o isolado, maior será a probabilidade de as variações casuais das freqüências gênicas provocarem a extinção ou a fixação de um dos alelos, podendo-se dizer que ***nos isolados pequenos as flutuações de amostragem deixam as freqüências gênicas sujeitas a variações aleatórias comparáveis às alterações de rumo de um barco abandonado à deriva***. Além disso, pode-se dizer que ***nos pequenos isolados ocorre uma queda da variabilidade genética, pois parte dos alelos é fixada e parte é eliminada, sendo a probabilidade de eliminação igual à de fixação***.

É evidente que nas populações reais o fenômeno da deriva genética pode ser um processo bem mais drástico do que aquele apresentado em nosso exemplo, baseado em populações

teóricas. Isso porque nas populações reais, além de ter que levar em conta a influência do **número de casais férteis**, devemos considerar a ocorrência de outros fatores, um dos quais, por exemplo, é o **tamanho das famílias**. Realmente, se dois pequenos isolados, A e B, tiverem o mesmo número de casais e o mesmo número médio de filhos por casal, mas se no isolado A metade dos casais tiver 2 filhos e a outra metade 6, enquanto no isolado B todos os casais tiverem quatro filhos, este último terá maior probabilidade de apresentar maior heterogeneidade genética do que o primeiro.

Um outro fator importante que deve ser levado em conta é o sistema de casamentos da população, isto é, se ele é monogâmico ou poligâmico, ou se a monogamia está restrita ou não a uma parcela da população ou, ainda, se a monogamia é constante ou se ela se alterna com períodos de promiscuidade sexual como acontecia com os esquimós do Alaska. Além disso, deve-se levar em conta a existência ou não de fatores que favorecem o aumento da consangüinidade entre os cônjuges. Obviamente, tanto a poligamia quanto o aumento da consangüinidade entre os cônjuges favorecem a queda da variabilidade genética.

Tendo em vista a influência desses fatores foi que Wright (1931) criou o conceito de **tamanho efetivo da população**. Assim, o tamanho efetivo da população, obtido por intermédio de fórmulas que levam em conta tais fatores, informará que ela se comporta como se fosse uma população na qual existem  $n$  indivíduos reprodutores, com o mesmo número de homens e de mulheres ( $\frac{n}{2}$ ) que se acasalam de modo aleatório, onde a variância do desvio das freqüências de genes autossômicos é  $(\frac{pq}{2n})$  e a queda da variabilidade genética é  $\frac{1}{2n}$ . Tais fórmulas permitiram verificar que o tamanho efetivo das populações humanas varia em torno de  $\frac{1}{4}$  do seu tamanho total, havendo estimativas desde  $\frac{1}{6}$  até  $\frac{1}{3}$  desse tamanho.

A fim de tomar mais claro o entendimento do conceito de tamanho efetivo de uma população, consideremos que num isolado com 200 habitantes se tivesse estimado seu tamanho efetivo em 46 indivíduos. Esse número nos indicaria que a queda da variabilidade genética desse isolado, isto é, o decréscimo da heterozigose por geração seria idêntico ao de um isolado onde houvesse apenas 23 casais geneticamente ativos (46 indivíduos). Além disso, poderíamos estimar a queda da variabilidade genética nesse isolado à razão de  $\frac{1}{92}$  ou 1,09% por geração, o que equivale a dizer que, após 92 gerações, toda a população seria constituída por homozigotos.

A distribuição das freqüências de certos genes em algumas populações humanas tem grande probabilidade de ter sido consequência da deriva genética, a qual deve ter tido papel evolutivo muito mais relevante do que a seleção natural nas populações primitivas, constituídas

por pequenos isolados genéticos, cujo tamanho efetivo não deveria exceder a 100. Nesses pequenos isolados existia a possibilidade de um único indivíduo, altamente fecundo, ou com descendentes muito fecundos, provocar a predominância de um ou mais genes na população originária de tais isolados, ainda que um desses genes conferisse baixo valor adaptativo a seus portadores. Quando se reconhece a existência desse processo, costuma-se falar em *efeito do fundador*.

Nas populações primitivas também era freqüente a diminuição abrupta de seu tamanho, pela redução do número de pessoas em consequência de guerras, epidemias, fome e outras catástrofes, dizimando famílias, independentemente do valor adaptativo de seus genes. As freqüências gênicas dos sobreviventes nem sempre correspondiam à da população original, de sorte que a população oriunda deles podia mostrar uma composição genética diversa da que existia anteriormente. Esse efeito, decorrente do “estreitamento da passagem” de genes de uma população original para outra por intermédio de uma geração reduzida por uma catástrofe, que afetou uma parte do isolado e que, evidentemente, nada tem a ver com o processo de seleção natural, costuma ser denominado *efeito do gargalo*.

Dentre os exemplos de possíveis consequências de deriva genética pode-se mencionar o caso dos indígenas sul-americanos não miscigenados, os quais apresentam, praticamente, apenas o grupo sangüíneo O. Assim, por exemplo, todos os 1.622 índios do Amazonas e do Mato Grosso estudados por vários pesquisadores (Biocca e Ottensooser, 1944; Ottensooser e Pasqualin, 1949; Lima, 1950; Junqueira e Wishart, 1956) mostraram ser do grupo sangüíneo O. A ausência dos genes *A* e *B* do sistema ABO entre eles poderia ser explicada como consequente de deriva genética ocorrida nos pequenos grupos que compunham as gerações ancestrais que lhes deram origem.

Existem, entretanto, outros exemplos em que o efeito da deriva genética está bem mais evidente, porque os genes que se manifestam com alta freqüência conferem a seus portadores valor adaptativo baixo. É o caso da porfiria hepática do tipo sul-africano (dominante autossômica), que é rara na maioria das populações humanas, mas ocorre em alta proporção nos africânderes (cerca de dois milhões de descendentes de pequeno número de holandeses e de franceses, que se estabeleceram na África do Sul em fins do século XVII). As investigações genealógicas realizadas levam a crer que a maioria dos casos de porfiria hepática entre os africânderes seja constituída por descendentes de um holandês que chegou à Cidade do Cabo em 1686.

Outro caso interessante é o da população dos atóis de Pinguelape e Moki e de seus migrantes para a ilha vizinha de Ponape. Nessas ilhas, que fazem parte das Ilhas Carolinas, a freqüência da acromatopsia associada à miopia é altíssima, pois está em torno de 5%. A

explicação para tão alta frequência dessa anomalia autossômica recessiva baseia-se no conhecimento de que a população atual dessas ilhas descende dos poucos sobreviventes (cerca de 30) que restaram na ilha de Pinguelape depois que, em 1775, o tufão Lengkieki e o período de fome que a ele se seguiu devastaram a ilha, matando a maioria dos seus habitantes. Um desses sobreviventes (Muahuele), que deve ter sido heterozigoto do gene raro que produz acromatopsia, deixou prole exageradamente grande, fazendo com que tal alelo se acumulasse na população e aparecesse na forma homozigótica após cinco gerações (Morton, 1973; Maumenee, 1976).

A alta frequência de albinos entre os habitantes da Ilha de Lençóis, no Maranhão (Freire-Maia e Cavalli, 1972,1973) tem, também, todas as características de ser consequência de deriva genética, pois num local tão ensolarado como essa ilha o albinismo tem, evidentemente, valor adaptativo muito baixo.

### **3. FLUXO GÊNICO DE POPULAÇÕES MIGRANTES**

Do mesmo modo que a deriva genética, as migrações têm um efeito evolutivo porque são capazes de promover alterações das frequências gênicas tanto nas populações da qual se originam os imigrantes, quanto naquelas que os recebem. De fato, se de uma população X emigrar um grupo que não é uma amostra representativa dessa população no concernente a certas frequências gênicas, esse grupo emigrante poderá, dependendo do seu tamanho e da intensidade da corrente migratória, passar a acusar as perdas gênicas provocadas pela emigração. Por outro lado, se o grupo que emigra da população X se estabelecer no território ocupado por uma população Y e diferir desta última quanto à frequência de certos genes, tal situação também provocará uma alteração da frequência desses genes na população que recebe os imigrantes.

Para exemplificar essas afirmações consideremos o caso de uma corrente migratória de uma população X para o território ocupado por uma população Y, que é incentivada pelas autoridades governamentais dessa última, as quais, entretanto, a troco do incentivo, exigem a seleção médica dos imigrantes. Suponhamos, que na população X existe uma frequência apreciável dos genes da talassemia beta e da deficiência de G-6PD, e que na população Y tais genes estejam ausentes ou, praticamente ausentes.

Se as autoridades da população Y somente autorizarem a imigração de pessoas que estão livres dos genes da talassemia beta e da deficiência de G-6PD, é evidente que as frequências desses genes crescerão ainda mais na população X, já que ela sofrerá emigração diferencial. O grupo migrante, por sua vez, poderá passar a constituir um grupo racial diverso do original, mesmo que não entre em miscigenação com outras populações. De fato, ele será um conjunto humano que difere da população da qual procede em relação às frequências de certos genes. Se,

entretanto, a seleção mencionada não existisse, estabelecer-se-ia um fluxo de genes da talassemia beta e da deficiência de G-6PD da população X para a população Y.

Os resultados do fluxo gênico de um grupamento humano em outro, pode ser detectado facilmente em populações de áreas nas quais as barreiras geográficas, sócio-culturais, políticas e religiosas são pouco acentuadas. Esse ingresso de genes também pode ser apreciado em populações humanas separadas por grandes distâncias geográficas, desde que as diferenças gênicas entre elas sejam acentuadas e que essas distâncias sejam ocupadas por populações pelas quais os genes possam fluir. Assim, por exemplo, a frequência do gene que determina a produção de antígeno B nas hemácias é mais alta nas populações mongólicas do que nas populações caucasóides européias, ocorrendo o inverso em relação ao gene da insensibilidade gustativa à feniltiouréia (PTC). Entretanto, quando se examina a distribuição das frequências do gene B nas populações européias e asiáticas, pode-se notar um gradiente que diminui no sentido da Ásia para a Europa, e um gradiente no sentido oposto no concernente ao gene da insensibilidade gustativa à PTC. As invasões mongólicas a partir do ano 500 poderiam explicar essas observações.

Nas populações em que há miscigenação com grupos imigrantes, o efeito do fluxo gênico oriundo desses últimos depende, evidentemente, da diferença entre as frequências gênicas das populações que se miscigenam, bem como do tamanho relativo delas. Para exemplificar, consideremos uma cidade predominantemente caucasóide do Estado de São Paulo, na qual se estabeleceu um isolado japonês, que passou a constituir 10% de sua população. Tomemos como marcadores genéticos dessas populações o gene *d*, responsável, quando em homozigose, pelo grupo sanguíneo Rh-negativo, e o gene *t* que, em homozigose, determina a insensibilidade gustativa à feniltiouréia. Sabendo-se que as frequências dos genes *d* e *t* nos caucasóides de São Paulo são estimados, respectivamente, em 0,304 e 0,497, enquanto que nos japoneses desse Estado são estimados em 0,055 e 0,356, respectivamente (Beiguelman e Marchi, 1962; Beiguelman, 1962, 1963, 1964), vejamos o que aconteceria na população da cidade de nosso exemplo se nela houvesse miscigenação completa.

Visto que o componente caucasóide corresponde a 90% da população total e o japonês a 10% dela, é claro que as frequências dos genes *d* e *t* passariam a ser iguais a 27,9 % e 48,3 %, pois o valor médio da frequência do gene *d* é  $\bar{q} = \sum xq = (0,9 \times 0,304) + (0,1 \times 0,058) = 0,279$ , enquanto o valor médio da frequência do gene *t* é  $\bar{q} = \sum xq = (0,9 \times 0,497) + (0,1 \times 0,356) = 0,483$ . No cálculo da frequência do gene *d* não levamos em conta a possibilidade de efeito seletivo resultante da incompatibilidade sanguínea materno-fetal, porque no Estado de São Paulo as mulheres Rh-negativo que geram uma criança Rh-positivo têm o direito de receber a globulina anti-D para impedir sua sensibilização pelo antígeno D.

Quando uma população é o resultado da miscigenação de duas outras, que diferem entre si quanto à frequência de um determinado gene, é evidente que a frequência desse gene na população miscigenada ( $\bar{q}$ ) será obtida por intermédio de  $\bar{q} = xq_1 + (1-x)q_2$ , onde  $q_1$  e  $q_2$  representam as frequências do gene nas populações originais, enquanto  $x$  e  $1-x$  representam as proporções com que cada urna delas participou para a formação da população miscigenada. Por isso, conhecendo-se as frequências gênicas  $\bar{q}$ ,  $q_1$  e  $q_2$  pode-se estimar as proporções  $x$  e  $1-x$ , pois desenvolvendo  $\bar{q} = xq_1 + (1-x)q_2$  obtém-se, sucessivamente:

$$\begin{aligned}\bar{q} &= xq_1 + q_2 - xq_2 \\ \bar{q} - q_2 &= x(q_1 - q_2) \\ x &= \frac{\bar{q} - q_2}{q_1 - q_2}\end{aligned}$$

Tomemos um exemplo numérico para demonstrar a aplicabilidade do que foi exposto, considerando a frequência do gene da insensibilidade gustativa à PTC em afro-brasileiros (0,359 segundo Saldanha,1965), em africanos negros (0,188, em média, segundo vários dados da literatura) e em caucasóides brasileiros (0,497, segundo Beiguelman,1964). Sabendo-se que a população afro-brasileira sofreu miscigenação com brancos, pode-se, com base nesse marcador genético, tentar estimar qual foi a intensidade da contribuição caucasóide para a composição do grupo afro-brasileiro. Para tanto escrevemos  $\bar{q} = 0,359$ ,  $q_1 = 0,497$  e  $q_2 = 0,188$  e calculamos

$$x = \frac{0,359 - 0,188}{0,497 - 0,188} = 0,553$$

o que nos levar a estimar que, para a constituição da população afro-brasileira, os caucasóides contribuíram com, praticamente, 55%. Visto que o componente caucasóide da população afro-norte-americana varia entre 10% e 30%, pode-se, pois, concluir que os afro-brasileiros são mais caucasóides do que os afro-norte-americanos.

Com base na informação de que os índios brasileiros, do mesmo modo que outros índios sul-americanos, eram todos do grupo sanguíneo O, pode-se empregar um método muito simples (Beiguelman, 1980) para estimar em que proporção os índios, os africanos negros e os caucasóides participaram da composição das populações do nordeste brasileiro, sabidamente tri-híbridas.

Se as frequências dos genes  $A$ ,  $B$  e  $O$  do sistema ABO dos caucasóides e dos negros africanos que contribuíram para a formação das populações nordestinas forem designadas, respectivamente, por  $p_1$ ,  $q_1$  e  $r_1$ , e  $p_2$ ,  $q_2$ , e  $r_2$ , entre os índios deveremos levar em conta apenas a frequência  $r_3 = 1$ , pois os índios não miscigenados não possuem os genes  $A$  e  $B$ . Na população nordestina atual as frequências dos genes  $A$ ,  $B$  e  $O$  devem, pois, ser representadas por  $\bar{p}$ ,  $\bar{q}$  e  $\bar{r}$ .

Se as proporções de caucasóides, africanos e índios que entraram na composição dessa população forem indicadas, respectivamente, por  $x_1$ ,  $x_2$  e  $x_3$ , poderemos escrever:

$$\begin{aligned}\bar{p} &= x_1 p_1 + x_2 p_2 \\ \bar{q} &= x_1 q_1 + x_2 q_2 \\ \bar{r} &= x_1 r_1 + x_2 r_2 + x_3 r_3\end{aligned}$$

Na primeira equação  $x_2$  pode ser substituído por seu significado obtido a partir da segunda equação, isto é, por  $x_2 = \frac{\bar{q} - x_1 q_1}{q_2}$ , o que permitirá estimar a contribuição da população caucasóide para a população nordestina por intermédio de:

$$x_1 = \frac{\bar{p} q_2 - p_2 \bar{q}}{p_1 q_2 - p_2 q_1}$$

Depois de calcular o valor de  $x_1$ , pode-se estimar facilmente os valores de  $x_2$  e de  $x_3$  por meio de:

$$x_2 = \frac{\bar{q} - x_1 q_1}{q_2}$$

$$x_3 = 1 - (x_1 + x_2) \quad \text{ou} \quad x_3 = \frac{\bar{r} - (x_1 r_1 + x_2 r_2)}{r_3}$$

Para ilustrar a aplicação desse método, consideremos as frequências gênicas do sistema ABO estimadas por Saldanha (1962) em migrantes procedentes de vários Estados do nordeste brasileiro, que, numa determinada época, passaram pela Hospedaria de Imigrantes de São Paulo ( $\bar{p} = 0,210$ ;  $\bar{q} = 0,083$ ;  $\bar{r} = 0,707$ ). Como representativas dessas frequências gênicas em caucasóides, tomemos as fornecidas por Cunha e Morais (1959) para a população de Portugal ( $p_1 = 0,304$ ;  $q_1 = 0,066$ ;  $r_1 = 0,630$ ), já que foram os imigrantes portugueses os caucasóides que predominaram no nordeste do Brasil. Para os negros africanos consideremos como representativas as fornecidas por Mourant *et al.* (1958) ( $p_2 = 0,157$ ;  $q_2 = 0,150$ ;  $r_2 = 0,693$ ). Visto que os índios brasileiros possuíam apenas o grupo sanguíneo O no sistema ABO, escreveremos  $r_3 = 1$ .

Com base nesses dados podemos estimar que os caucasóides, os negros africanos e os índios contribuíram, respectivamente, com 53%, 32% e 15% de seus genes para a população nordestina representada pelos migrantes estudados por Saldanha (1962), pois:

$$\begin{aligned}x_1 &= \frac{(0,210 \times 0,150) - (0,157 \times 0,083)}{(0,304 \times 0,150) - (0,157 \times 0,066)} = 0,53 \\ x_2 &= \frac{0,083 - (0,528 \times 0,066)}{0,150} = 0,32 \\ x_3 &= 1 - (0,53 + 0,32) = 0,15\end{aligned}$$

Esses percentuais estão muito próximos daqueles obtidos por Saldanha (1962), empregando outro método (48% de caucasóides, 34% de negros africanos e 18% de índios). Esse

método consistiu em avaliar a contribuição desses três grandes grupos raciais a partir de um marcador genético ausente em caucasóides e negros africanos, mas presente em índios (gene  $Di^a$  do sistema sanguíneo Diego), e de outro praticamente ausente em caucasóides e em índios, mas muito freqüente em africanos (gene  $R$  ou combinação  $cDe$  do sistema Rh). Desse modo, Saldanha (1962) pôde tratar a população como di-híbrida, usando a fórmula  $x = \frac{\bar{q} - q_2}{q_1 - q_2}$  e considerando, no caso do sistema Diego, os grupos raciais índios e "não-índios", e, no caso do sistema Rh, os grupos raciais afro e "não-afro".

Para finalizar este tópico é importante assinalar que, quando pretendemos avaliar a contribuição de duas ou mais populações para outra, resultante de miscigenação, é evidente que a alternativa mais correta é o estudo de vários sistemas genéticos simultaneamente, porque, desse modo, aumentamos a precisão da estimativa da contribuição genômica das populações originais. Foi o que fizeram Krieger *et al.* (1965) quando analisaram 17 sistemas alélicos em outro conjunto de migrantes nordestinos que passaram pela Hospedaria de Imigrantes de São Paulo. Isso lhes permitiu estimar as contribuições dos caucasóides, afros e índios para a população nordestina brasileira em, respectivamente, 59%, 30% e 11%, percentuais esses que, entretanto, não divergem muito daqueles estimados por Saldanha (1962) e pelo autor deste volume.

### MISTURA RACIAL E HEREDOPATIAS RECESSIVAS

Em relação às heredopatias recessivas que mostram grande variação de freqüência em diferentes populações, a mistura racial tem um efeito benéfico, pelo menos durante um longo período, pois ela provoca uma diminuição da freqüência de homozigotos na população miscigenada. Para exemplificar, consideremos que em uma área geográfica convivem duas populações (A e B) que constituem isolados de mesmo tamanho. Suponhamos, ainda, que duas anomalias autossômicas recessivas monogênicas ( $a$  e  $b$ ) ocorrem com freqüências, respectivamente, 1: 10.000 e 1: 1.000.000 no isolado A e com freqüências inversas, isto é, 1: 1.000.000 e 1: 10.000 no isolado B.

Se houver quebra dos isolados A e B, com miscigenação completa e panmixia dos indivíduos que os compõem é evidente que as anomalias  $a$  e  $b$  passarão a ocorrer na população miscigenada, cada qual com freqüência igual a 3: 100.000. Em outras palavras, a miscigenação beneficiará o isolado A ao baixar a freqüência da anomalia  $a$  e o isolado B ao diminuir a freqüência da anomalia  $b$ .

O fluxo de genes responsáveis por heredopatias, em consequência de mistura racial, também pode fazer com que uma anomalia recessiva, que é considerada relativamente freqüente, mostre, paradoxalmente, alta proporção de casais consangüíneos entre os genitores de pessoas afetadas por ela.

Para exemplificar, tomemos o caso da fibrose cística do pâncreas, a qual é, sabidamente, decorrente de numerosos mutantes alelos, pertencentes a um mesmo loco do cromossomo 7. Visto que tais mutantes determinam essa anomalia recessiva tanto em homozigose quanto em heterozigose entre si, tudo se passa como se houvesse um único alelo responsável por esse defeito genético. Desse modo, em decorrência da alta frequência dessa heredopatia entre os caucasóides (1:1.000 a 1:3.000), não se deve esperar alta proporção de filhos de casais consangüíneos entre os pacientes caucasóides com fibrose cística do pâncreas, pois, nesse caso, a heterogeneidade genética, como vimos, não diz respeito a diferentes locos gênicos, mas a um único.

Em países onde há grande mistura racial, como no Brasil, pode acontecer, entretanto, que se assinale alta proporção de casais consangüíneos entre os genitores de pacientes com fibrose cística do pâncreas. De fato, sabendo-se que essa heredopatia é praticamente ausente tanto em negros africanos quanto nas populações de origem mongólica, está claro que nas populações resultantes da miscigenação de caucasóides com negros ou com mongolóides a frequência do “gene” responsável pela fibrose cística do pâncreas estará diminuída. Dependendo do componente caucasóide na população miscigenada, a frequência desse “gene” poderá ser muito baixa e, assim, propiciar o aparecimento de uma alta proporção de filhos de consangüíneos entre os pacientes com essa heredopatia.

#### 4. O TAMANHO DOS ISOLADOS

Os isolados podem crescer por reprodução de seus indivíduos (*crescimento dentro de suas fronteiras*) ou por miscigenação com imigrantes (*quebra dos isolados*), mas eles podem, ao invés, diminuir ou, ao menos, manter seu tamanho, em consequência da pouca reprodutibilidade e(ou) de emigração de seus elementos. Por isso, a medida do tamanho dos isolados deve sempre referir-se a um período definido. Além disso, considerando que os geneticistas estão interessados na população geneticamente ativa, da qual depende a variação do tamanho dos isolados, a medida desse tamanho deve ser feita com o objetivo de estimar não o número de indivíduos que os constituem, mas sim o número de pessoas do isolado que, em um determinado período, estão se casando (Dahlberg, 1929) ou que estão se unindo e deixando prole (Frota-Pessoa, 1957,1963).

Para estimar esse número Dahlberg (1929) partiu do princípio de que, em panmixia, os casamentos entre consangüíneos ocorrem aleatoriamente e, nesse caso, a frequência desses casamentos dependerá apenas do tamanho médio das famílias e do tamanho do isolado. Assim, se  $i$  for o tamanho médio das irmandades,  $i - 1$  será o número médio de irmãos de cada indivíduo e  $2(i - 1)$  será o número médio de tios de cada indivíduo, isto é,  $i - 1$  tios paternos e  $i - 1$  tios maternos. Portanto, cada indivíduo da população terá, em média,

$2i(i-1)$  primos em primeiro grau e  $i(i - 1)$  primos em primeiro grau do sexo oposto, dentre os quais poderá escolher seu cônjuge.

Desse modo, a probabilidade  $c$  de ocorrência casual de um casamento entre primos em primeiro grau será  $c = \frac{i(i-1)}{\frac{n}{2}} = \frac{2i(i-1)}{n}$  se  $n$  for o tamanho do isolado e  $\frac{n}{2}$  for o número de

indivíduos do sexo oposto dentre os quais um indivíduo escolhe o seu cônjuge. Por isso, resolvendo a última fórmula em relação a  $n$ , o tamanho do isolado é calculado, segundo Dahlberg (1929), a partir de  $n = \frac{2i(i-1)}{c}$ .

Apliquemos essa fórmula a uma comunidade que constitui um isolado, a qual, em determinado período apresentava um número médio de filhos por casal  $i = 2,2$  e uma taxa de casamentos entre primos em primeiro grau  $c = 5\%$ . Nesse caso, diríamos que, no período em estudo, o tamanho do isolado era de 106 nubentes, pois  $n = \frac{2 \times 2,2 \times 1,2}{0,05} = 105,6$ . Se, posteriormente, o mesmo isolado continuasse apresentando um número médio de filhos por casal  $i = 2,2$ , mas a taxa de casamentos entre primos em primeiro grau fosse menor,  $c = 1\%$ , por exemplo, diríamos que o isolado cresceu porque  $n$  passaria a ser constituído por 528 nubentes, já que  $n = \frac{2 \times 2,2 \times 1,2}{0,01} = 528$ .

Frota-Pessoa (1963) melhorou a fórmula de Dahlberg porque levou em conta as frequências  $x_1, x_2, x_3, \dots, x_k$  das irmandades com 1, 2, 3 ...  $k$  indivíduos, de modo que o tamanho do isolado passa a ser medido por intermédio de  $n = \frac{2 \sum xi(i-1)}{c}$ .

Para ilustrar a aplicação dessa última fórmula, consideremos que, no exemplo anterior, soubéssemos que o número de indivíduos nas irmandades geradas pelos casais desse isolado varia de 1 a 5, e que tais irmandades ocorrem na população com frequências iguais, respectivamente, a 30%, 35%, 25%, 5% e 5%. Diante dessas informações operaríamos como no quadro abaixo e concluiríamos que o isolado em questão teve, nos dois períodos, tamanho maior do que o calculado segundo a fórmula de Dahiberg, pois no primeiro período  $n = \frac{2 \times 3,8}{0,05} = 152$  e no segundo  $n = \frac{2 \times 3,8}{0,01} = 760$ .

$x$	$i$	$xi$	$xi(i-1)$
0,30	1	0,30	-
0,35	2	0,70	0,70
0,25	3	0,75	1,50
0,05	4	0,20	0,60
0,05	5	0,25	1,00
1,00		2,20	3,80

Quando essas técnicas de mensuração dos isolados são aplicadas a pequenos grupos humanos bem definidos (populações de pequenas ilhas, pequenas comunidades rurais isoladas, ou pequenos grupos religiosos) é claro que o tamanho do isolado dá uma idéia de uma entidade real. Contudo, quando essas mesmas técnicas são aplicadas a isolados de limites indefinidos, como é o caso das subpopulações estratificadas que constituem os grupamentos urbanos, também é óbvio que as medidas do tamanho dos isolados devem ser aceitas como mensurações de entidades abstratas. Essa é a razão pela qual, nesses casos, se fala em **tamanho médio dos isolados**.

Outras fórmulas, que levam em conta outros tipos de casamentos consangüíneos, também foram apresentadas na literatura especializada com a finalidade de estimar o tamanho dos isolados. Entretanto, a sua aplicabilidade é muito pequena, seja porque é mais difícil obter dados censitários a respeito de casais com consangüinidade distante (primos em terceiro grau e em graus superiores), seja porque se referem a casamentos entre colaterais de gerações diferentes (casamentos entre tios(as) e sobrinhas(os) ou entre primos em segundo grau, por exemplo).

A crítica mais importante que se faz ao emprego das fórmulas aqui apresentadas para estimar o tamanho dos isolados é a de que elas partem do princípio de que os casamentos consangüíneos podem ser considerados como ocorrendo aleatoriamente, quando, na realidade, eles não ocorrem ao acaso (Morton, 1955). Assim, em populações nas quais existam fatores favorecendo os casamentos entre primos em primeiro grau, a aplicação dessas fórmulas conduzirá a sub-estimativas do tamanho dos isolados, pois  $c$  terá valores altos. Em oposição, nos isolados constituídos por imigrantes, onde as possibilidades matrimoniais entre consangüíneos podem ficar diminuídas pela redução do número de parentes, as fórmulas em discussão darão super-estimativas do seu tamanho.

## QUESTÕES E RESPOSTAS

**Q 1.** A incidência de acondroplásicos entre os recém-nascidos de uma população foi estimada em 1: 18.000. Qual a taxa de mutação da acondroplasia (anomalia autossômica dominante monogênica) que pode ser estimada nessa população, se o coeficiente seletivo dessa anomalia for  $s = 0,90$ .

**R 1.** 2,5: 100.000 ou  $2,5 \times 10^{-5}$ , porque  $p = \frac{1}{2} \cdot \frac{1}{18.000} = \frac{1}{36.000}$ , de sorte que  $\mu = sp = 0,90 \cdot \frac{1}{36.000} = 0,000025$ .

**Q 2.** A incidência de uma doença recessiva autossômica monogênica em uma população é estimada em 1:25.000. Visto que o coeficiente seletivo dessa anomalia foi estimado em 0,98,

qual a taxa de mutação que pode ser calculada para o gene determinante da anomalia em questão?

**R 2.**  $3,9 \times 10^{-5}$  porque  $q^2 = \frac{1}{25.000}$ ;  $\mu = sq^2 = 0,98 \cdot \frac{1}{25.000} = 0,000039$ .

**Q 3.** A taxa de mutação do gene da distrofia muscular do tipo Duchenne (anomalia recessiva ligada ao cromossomo X) foi estimada em  $9 \times 10^{-5}$  em uma determinada população. Visto que o coeficiente seletivo dessa anomalia é igual à unidade, qual a frequência de indivíduos com essa miopatia que pode ser estimada nessa população?

**R 3.** 27:100.000 porque  $q = 3 \times 0,00009 = 0,00027$ , pois  $q = \frac{3\mu}{s}$ .

**Q 4.** Em uma amostra de 100 indivíduos com uma anomalia autossômica dominante monogênica verificou-se que eles tiveram 30 filhos que chegaram até a idade reprodutiva. Em uma amostra controle de 500 indivíduos normais constatou-se um total de 1.000 filhos que alcançaram essa idade. Com base nesses dados estimar o coeficiente seletivo da anomalia em discussão.

**R 4.** Visto que a fecundidade média dos indivíduos com a anomalia é  $\frac{30}{100} = 0,3$  e a dos indivíduos normais é  $\frac{1000}{500} = 2$ , pode-se calcular  $s = 0,85$ , pois  $f = \frac{0,3}{2} = 0,15$ .

**Q 5.** Em uma amostra de 40 indivíduos com uma determinada anomalia autossômica dominante monogênica verificou-se que 8 eram casos familiares, sendo os restantes casos esporádicos. Qual o coeficiente seletivo que se pode estimar para essa anomalia?

**R 5.**  $s = 0,80$ ; pois  $f = \frac{8}{40} = 0,20$ .

**Q 6.** Uma anomalia autossômica dominante monogênica está sujeita a um coeficiente seletivo  $s = 0,5$ . Visto que o gene que a determina tem 70% de penetrância, quer-se saber qual a persistência de um mutante causador dessa anomalia.

**R 6.** 2,86 gerações, porque  $i = \frac{1}{0,70 \times 0,5} = 2,86$ .

**Q 7.** Certas publicações preconizam a esterilização de indivíduos que apresentam anomalia genéticas. Se tal prática fosse adotada, quantos anos seriam necessários para que uma anomalia recessiva autossômica monogênica tivesse a sua frequência reduzida de 1:22.500 para 1:90.000?

**R 7.** 3.750 anos, porque  $q^2 = \frac{1}{22.500}$  e  $q = \frac{1}{150}$  na geração inicial. Na  $n$ -ésima geração  $q_n^2 = \frac{1}{90.000}$  e  $q_n = \frac{1}{300}$ . Portanto, para reduzir  $q$  de  $\frac{1}{150}$  para  $\frac{1}{300}$  isto é, à metade, é necessário que o número de gerações ( $n$ ) seja igual a  $\frac{1}{q} = 150$  gerações ou  $\frac{150}{4} = 37,5$  séculos.

**Q 8.** A esterilização compulsória de todos os indivíduos com uma anomalia dominante autossômica monogênica pode deixar as gerações futuras livres dessa anomalia?

**R 8.** Não, porque novos casos surgirão, por mutação.

**Q 9.** Numa população os alelos autossômicos  $A, a$  têm frequências iguais, respectivamente, a  $p = 0,80$  e  $q = 0,20$ . Se em uma determinada geração dessa população, e na seguinte, ocorrer seleção completa dos homozigotos  $aa$  antes da idade reprodutiva, qual será a distribuição dos genótipos  $AA, Aa$  e  $aa$  entre os recém-nascidos da segunda geração filial?

**R 9.**  $AA = 73,96\%$ ,  $Aa = 24,08\%$  e  $aa = 1,96\%$  porque  $q_2 = \frac{1}{1+(2 \times 0,2)} = 0,14$  e, portanto  $p_2 = 0,86$ .

**Q 10.** Se na população da questão anterior, em vez de seleção completa dos homozigotos  $aa$ , ocorresse seleção completa contra os dois tipos de homozigotos ( $AA$  e  $aa$ ), qual seria a distribuição dos genótipos  $AA, Aa$  e  $aa$  entre os recém-nascidos da segunda geração filial?

**R 10.**  $AA = 25\%$ ,  $Aa = 50\%$  e  $aa = 25\%$ , pois apenas os indivíduos  $Aa$  se reproduziriam.

**Q 11.** Numa população de negros africanos, a frequência de indivíduos com anemia falciforme que alcançam a idade reprodutiva é nula. Nessa mesma população, a incidência do traço siclêmico é 10%. Se 8 % dos indivíduos com o traço siclêmico não atingirem a idade reprodutiva por causa de complicações decorrentes da hemoglobina S, a eliminação do gene da hemoglobina S por intermédio dos heterozigotos será mais eficiente do que aquela feita por intermédio dos homozigotos?

**R 11.** Sim, porque ela será 1,52 vezes maior. De fato,  $q = \frac{0,10}{2} = 0,05$  e  $p = 0,95$ , de sorte que  $\frac{hp}{q} = \frac{0,08 \times 0,95}{0,05} = 1,52$ .

**Q 12.** Suponha que um caráter é condicionado por dois pares de alelos autossômicos ( $A, a$  e  $B, b$ ), cujas frequências em uma população em equilíbrio de Hardy e Weinberg são  $A = p_1 = 0,90$ ;  $a = q_1 = 0,10$ ;  $B = p_2 = 0,50$ ;  $b = q_2 = 0,50$ . Se os indivíduos duplamente homozigotos  $aabb$  dessa população passassem a sofrer seleção completa antes da idade reprodutiva, qual seria a frequência dos genes  $a$  e  $b$  após a primeira geração de seleção natural?

**R 12.**  $a = 0,0977$ ;  $b = 0,4987$ , porque  $a = \frac{0,10 - (0,01 \times 0,25)}{1 - (0,01 \times 0,25)} = 0,0977$  e  $b = \frac{0,50 - (0,01 \times 0,25)}{1 - (0,01 \times 0,25)} = 0,4987$ .

**Q 13.** Considere um par de alelos autossômicos  $A, a$  em uma população, e que os homozigotos estejam sujeitos a um mesmo processo seletivo, enquanto os heterozigotos estão sujeitos ao dobro da seleção que opera contra os homozigotos  $AA$  e  $aa$ . Se a frequência do gene  $A$  for  $p$  e a do gene  $a$  for  $q$ , o que deve acontecer com essas frequências gênicas se:

- a)  $p > q$
- b)  $p < q$
- c)  $p = q$

- R 13.** a) A frequência do gene  $a$  deve diminuir.  
b) A frequência do gene  $A$  deve diminuir.  
c) Não deve haver alteração das frequências dos alelos  $A$  e  $a$ .

**Q 14.** Suponha uma região, na qual a malária por *Plasmodium falciparum* é endêmica, habitada por uma população na qual o gene da hemoglobina S está presente. Se nessa população se tiver observado que chegaram à idade reprodutiva:

- a) 540 dentre 1.000 indivíduos com hemoglobina normal ( $AA$ );  
b) 600 dentre 1.000 indivíduos com o traço siclêmico ( $AS$ );  
c) Nenhum indivíduo com anemia falciforme ( $SS$ );

Qual deverá ser a estimativa da frequência do gene S na população em equilíbrio genético sob seleção natural?

**R 14.** Pode-se estimar a frequência do gene  $S$  em 9%. De fato, o valor adaptativo dos homozigotos  $AA$  pode ser estimado em 0,54, o dos heterozigotos  $AS$  em 0,60 e a dos homozigotos  $SS$  pode ser considerado nulo. Desse modo, se o valor adaptativo dos heterozigotos for considerado como a unidade de comparação, o valor adaptativo dos homozigotos  $AA$  passará a ser  $\frac{0,54}{0,60} = 0,9$  de sorte que seu coeficiente seletivo será estimado em 0,1. Visto que o

coeficiente seletivo dos homozigotos  $SS$  é 1, pode-se calcular  $q = \frac{0,1}{0,1+1} = 0,09$  ou 9%.

**Q 15.** Um gene determina uma anomalia neurológica dominante autossômica, que se manifesta por volta dos 30 anos de idade e causa a morte do seu portador por volta dos 40 anos de idade. Tal gene pode ser considerado letal?

**R 15.** Não, porque ele permite a reprodução de seu portador, podendo, por isso, ser transmitido às gerações futuras.

**Q 16.** Existem genes com coeficiente seletivo praticamente igual a 1, que são considerados como determinadores de anomalias dominantes autossômicas. De que decorre a aceitação desse padrão de herança?

**R 16.** A aceitação desse padrão de herança decorre do estudo dos descendentes de alguns poucos casos esporádicos que alcançam a idade reprodutiva.

**Q 17.** Os poucos casos de epilóia que se reproduzem indicam que essa anomalia é determinada por um gene autossômico e que ela é transmitida de modo dominante. Isso nos permite concluir que a esmagadora maioria dos casos de epilóia é constituída por mutantes.?

**R 17.** Sim.

**Q 18.** A distrofia muscular do tipo Duchenne, determinada por um gene do cromossomo X, impede, como se sabe, que os indivíduos do sexo masculino hemizigotos desse gene cheguem à idade reprodutiva. Isso nos permite concluir que a esmagadora maioria nos casos de distrofia muscular do tipo Duchenne é composta por mutantes?

**R 18.** Não, porque esses casos podem ser gerados por mulheres heterozigotas.

**Q 19.** Em uma pequena comunidade tribal africana, completamente isolada de outras por barreiras geográficas, viviam, em uma determinada época, 10 casais. Dentre eles, 8 eram constituídos por cônjuges com hemoglobina normal do adulto ( $AA \times AA$ ), um era composto por um cônjuge com o traço ciclêmico e outro com hemoglobina normal do adulto ( $AS \times AA$ ) e um era composto por cônjuges com o traço ciclêmico ( $AS \times AS$ ). Os 8 casais  $AA \times AA$  geraram um total de 40 filhos que chegaram à idade reprodutiva, o casal  $AS \times AA$  gerou três filhos que chegaram à idade reprodutiva, todos com hemoglobina normal, enquanto o casal  $AS \times AS$  revelou-se estéril, em consequência de uma orqui-epididimite infecciosa que afetara o cônjuge masculino. Pergunta-se:

- Qual a frequência do gene que condiciona a produção da cadeia  $\beta^S$  de hemoglobina na geração paterna?
- Qual a frequência do gene que condiciona a produção da cadeia  $\beta^S$  de hemoglobina na geração filial?
- O que ocorreu com os genes  $A$  e  $S$  na comunidade tribal em apreço?
- Qual a causa dessa ocorrência?

**R 19.** a) 7,5 %, pois  $AS = \frac{3}{20} = 0,15$ , de modo que  $S = \frac{0,15}{2} = 0,075$ .

b) Nula.

c) O gene  $A$  foi fixado e o gene  $S$  foi eliminado da população que constitui o isolado.

d) A deriva genética.

**Q 20.** Se, na questão anterior, não soubéssemos o que ocorreu na geração filial, nem tivéssemos informações sobre a orqui-epididimite que afetou um dos indivíduos  $AS$  da geração paterna, o que se poderia prever a respeito da frequência do gene responsável pela produção das cadeias  $\beta^S$  de hemoglobina?

**R 20.** A eliminação do gene  $S$ , a manutenção de sua frequência original ou o aumento de sua frequência, pois, na geração paterna, o desvio padrão da frequência do gene  $S$  é estimado em 4%

porque  $\sigma = \sqrt{\frac{0,925 \times 0,075}{40}} = 0,04$ .

**Q 21.** Um isolado A é constituído por 260 habitantes e tem tamanho efetivo de 42 indivíduos. Outro isolado B, de estrutura genética, social e econômica similar a A é constituído por 150 habitantes e tem tamanho efetivo de 60 indivíduos. Se as mesmas condições de isolamento persistirem nos dois isolados, qual deles mostrará maior velocidade de diminuição da variabilidade genética? Por que?

**R 21.** O isolado A, porque, nele, a queda da variabilidade genética por geração é feita na taxa de 1,19 %, pois  $\frac{1}{84} = 0,0119$ , enquanto que no isolado B essa taxa é de 0,83%, pois  $\frac{1}{120} = 0,0083$ .

**Q 22.** O papel da deriva genética nas populações humanas deve ter sido mais importante no passado ou no presente?

**R 22.** No passado, porque as populações humanas eram menores e o isolamento genético mais acentuado.

**Q 23.** Dois isolados, A e B, têm o mesmo tamanho efetivo, mas no isolado A o sistema de casamentos é poligâmico, enquanto no isolado B ele é monogâmico. Em qual dos dois o aumento da homozigose se fará mais rapidamente?

**R 23.** No isolado A, porque a poligamia favorece a identidade genética dos descendentes desses casamentos.

**Q 24.** Em uma pequena povoação de uma ilha isolada do continente verificou-se que 8% dos casais são constituídos por primos em primeiro grau. Nessa povoação as irmandades variam de um a seis indivíduos e são encontradas segundo o percentuais abaixo:

Irmandades	%
1	25
2	40
3	15
4	10
5	5
6	5

Estimar o tamanho do isolado dessa ilha segundo as fórmulas de:

- a) Dahlberg.
- b) Frota-Pessoa.

**R 24.** a) 88,8 nubentes.  
b) 135 nubentes.

Realmente, operando como abaixo, tem-se:

<i>i</i>	<i>x</i>	<i>xi</i>	<i>xi(i-1)</i>
1	0,25	0,25	-
2	0,40	0,80	0,80
3	0,15	0,45	0,90
4	0,10	0,40	1,20
5	0,05	0,25	1,00
6	0,05	0,30	1,50
Total		2,45	5,40

$$n = \frac{2 \times 2,45 \times 1,45}{0,08} = 88,8 \qquad n = \frac{2 \times 5,40}{0,08} = 135$$

**Q 25.** Em uma cidade foram registrados 100.000 nascimentos em um determinado período. Dentre eles, 10.000 eram crianças pertencentes a um isolado genético, duas das quais manifestaram uma anomalia recessiva autossômica monogênica. Qual a frequência de heterozigotos do gene dessa anomalia que se estimaria:

- a) Se não soubéssemos que a anomalia ocorreu apenas nas crianças do isolado?
- b) Conhecendo essa informação?

**R 25.** a) 0,89%, pois  $q^2 = 0,00002$  permite estimar  $q = 0,00447$ , de sorte que  $Aa = 2pq = 2 \times 0,99553 \times 0,00447 = 0,0089$

b) 2,78% no isolado, porque  $q^2 = 0,0002$ ;  $q = 0,0141$ ;  $Aa = 2pq = 2 \times 0,9859 \times 0,0141 = 0,0278$ . Fora do isolado a frequência de heterozigotos é nula.

**Q 26.** Sabemos que uma população C é resultado da miscigenação de dois grupos raciais distintos, A e B, e que na população C um determinado fenótipo recessivo autossômico monogênico ocorre com frequência igual a 16%. Sabendo-se que a frequência desse fenótipo é estimada em 64% no grupo racial A e em 9 % no grupo racial B, quer-se estimar a proporção de indivíduos do grupo racial A que interveio na formação da população C.

**R 26.** 20%, porque:

$$\begin{aligned} \bar{q}^2 &= 0,16; \bar{q} = 0,40 \\ q^2_A &= 0,64; q_A = 0,80 \\ q^2_B &= 0,09; q_B = 0,30 \\ x &= \frac{0,40 - 0,30}{0,80 - 0,30} = 0,20 \end{aligned}$$

**Q 27.** Uma pequena cidade recebeu um contingente de imigrantes que equivalia a cerca de 20% da população receptora. O exame dos grupos sanguíneos de uma amostra aleatória de 100 indivíduos da população imigrante e de 100 da população receptora revelou a seguinte distribuição quanto aos grupos sanguíneos M, MN e N:

População	M	MN	N
Receptora	35	50	15
Migrante	40	40	20

Se houver miscigenação entre as populações receptora e imigrante, ela afetará as frequências dos genes *M* e *N* na população resultante da miscigenação?

**R 27.** Não, porque os genes *M* e *N* ocorrem com frequências idênticas nas duas populações ( $M = 0,60$ ;  $N = 0,40$ )

**Q 28.** Uma anomalia recessiva autossômica monogênica ocorre com frequência igual a 1:10.000 na população A e com frequência 0,01:10.000 na população B. Se a população B se miscigenar

com a A em uma proporção tal que A constitua 20% da população total, qual será a frequência da anomalia em discussão na população híbrida?

**R 28.** 0,08: 10.000 porque:

$$q^2_A = 0,0001; q_A = 0,01$$

$$q^2_B = 0,000001; q_B = 0,001$$

$$\bar{q} = (0,20 \times 0,01) + (0,80 \times 0,001) = 0,0028$$

$$\bar{q}^2 = 0,000008$$

## REFERÊNCIAS

- Aird, J., Bentall, H.H. & Roberts, J.A.F. A relationship between cancer of stomach and ABO blood groups. *Brit. Med. J. I*: 794-801, 1953.
- Allison, A.C. Protection afforded by sickle-cell trait against subtertian malarial infection. *Brit. Med. J. I*: 290-294, 1954a.
- Allison, A.C. The distribution of the sickle-cell trait in East Africa and elsewhere and its apparent relationship to the incidence of subtertian malaria. *Tr. Roy. Soc. Trop. Med. Hyg.* 48: 312-318, 1954b.
- Beiguelman, B. Taste sensitivity to PTC among Japanese immigrants in Brazil. *Rev. Bras. Biol.* 22: 93-97, 1962a.
- Beiguelman, B. Reação à fenil-tio-carbamida (PTC) e lepra. *Rev. Bras. Lepr.* 30: 111-124, 1962 b.
- Beiguelman, B. Grupos sanguíneos e lepra. *Rev. Bras. Lepr.* 31: 34-44, 1963.
- Beiguelman, B. Sistema ABO e epidemiologia de lepra. *Rev. Paul. Med.* 65: 80-86, 1964a.
- Beiguelman, B. Taste sensitivity to phenylthiourea among patients affected with both tuberculosis and leprosy. *Acta Genet. Med. Gemellol.* 13: 190-192, 1964b.
- Beiguelman, B. Taste sensitivity to phenylthiourea and leprosy. *Acta Genet. Med. Gemellol.* 13: 193-196, 1964c.
- Beiguelman, B. Um método simples para o estudo da mistura racial no Brasil. *Ciência e Cultura* 32(Supl.): 745, 1980.
- Beiguelman, B. & Marchi, A. The frequency of blood groups among Japanese immigrants in Brazil. *Am. J. Phys. Anthropol.* 20:29-31, 1962.
- Beutler, E. The hemolytic effect of primaquine and related compounds. A review. *Blood* 14: 103-139, 1959.
- Beutler, E., Dern, R.J. & Flanagan, C.L. Effect of sickle-cell trait on resistance to malaria. *Brit. Med. J. I*: 1189-1191, 1955.
- Biocca, E. & Ottensooser, F. Estudos etnobiológicos sobre os índios da região do Alto Rio Negro: Amazonas. I. Grupos sanguíneos comuns e fatores MN. *Arq. Biol. (São Paulo)* 28: 111-118, 1944.
- Brues, A.M. Selection and polymorphism in the A-B-0 blood groups. *Am. J. Phys. Anthropol.* 12: 559-557, 1954.
- Cunha, A.X. & Morais, M.X Os grupos sanguíneos dos portugueses. Contribuição ao Estudo da Antropologia Portuguesa, 7:17-36, 1959, cf. Saldanha, P.H. *op. cit.*
- Dahlberg, G. Inbreeding in man. *Genetics* 14: 421-454, 1929.
- Darwin, C. *Origem das espécies*. Ed. Itatiaia Ltda./EDUSP, S. Paulo, 1985. Tradução de AMADO, E., da 1ª edição de *On the origin of species by means of natural selection*, John Murray, Abennale Street, London, 1859.

- Frota-Pessoa, O. The estimation of the size of isolates based on census data. *Am. J. Hum. Genet.*, 9: 9-16, 1957.
- Frota-Pessoa, O. *O tamanho médio dos isolados das populações brasileiras*. Tese de Livre Docência, Universidade de São Paulo, 1963.
- Freire-Maia, N. *Genética de populações humanas*. HUCITEC, Editora da USP, São Paulo, 1974.
- Freire-Maia, N. & Cavalli, I.J. Estudos médicos e genéticos na Ilha dos Lençóis, Maranhão. III -Aspectos genéticos. *Ciência e Cultura*, 24(Supl.): 179, 1972.
- Freire-Maia, N. & Cavalli, I.J. Breeding structure and genetic drift in the population of Lençóis Island. *Ciência e Cultura*, 25(Supl.): 229-230, 1973.
- Fulton, J.D. & Grant, J.P. The sulphur requirements of the erythrocytic form of *Plasmodium knowlesi*. *Biochem. J.* 63: 274-282, 1956.
- Geiman, P.M. The cultivation of malarial parasites. Em Most, H. (Ed.) *Parasite infections in man*. Columbia University Press, N. York, 1951.
- Glass, B. The action of selection on the principal Rh alleles. *Am. J. Hum. Genet* 2: 269-278, 1950.
- Haldane, J.B.S. Disease and evolution. *Supp. La Ricerca Sci.* 19: 68-76, 1949.
- Harris, H., Kalmus, H. & Trotter, W.R. Taste sensitivity to phenylthiourea in goitre and diabetes. *Lancet* 2: 1038, 1949.
- Junqueira, P.C. & Wishart, R.J. Blood groups of Brazilian Indians (Carajas). *Nature*, 177: 40, 1956.
- Kimura, M. Genetic variability maintained in a finite population due to mutational production of neutral and nearly neutral isoalleles. *Genet. Res.* 11: 247-269, 1968.
- Kitchin, F.D.; Howel-Evans, W.; Clarke, C.A.; McConnell, R.B. & Sheppard, P.M. PTC taste response and thyroid disease. *Brit. Med. J.* 1: 1069-1084, 1959.
- Krieger, H., Morton, N.E., Mi, M.P., Azevedo, E., Freire-Maia, A. & Yasuda, N. Racial amixture in north-eastern Brazil. *Ann. Hum. Genet.*, 29: 113-125, 1965.
- Layrisse, M. & Arends, T. The Diego blood factor in Chinese and Japanese. *Nature* 177: 1083-1084, 1956.
- Layrisse, M.; Arends, T. & Dominguez Sisco, R. Nuevo grupo sanguíneo encontrado en descendientes de índios. *Acta Med. Venezolana* 3: 132-138, 1955.
- Lewis, M.; Ayukawa, H.; Chown, B. & Levine, P. The blood group antigen Diego in North American Indians and in Japanese. *Nature* 177: 1084, 1956.
- Li, C.C. *Population genetics*. Univ. Chicago Press, Chicago, 7a. reimpressão, 1972.
- Lima, P.E. Grupos sanguíneos dos índios do Xingu. *Bol. Museu Nac. (Rio de Janeiro)*, 11: 1-4, 1950.
- Matsunaga, E. Intra-uterine selection by the ABO incompatibility of mother and foetus. *Am. J. Hum. Genet.* 7: 66-71, 1955.
- Matsunaga, E. & Itoh, S. Blood groups and fertility in a Japanese population, with special reference to intra-uterine selection due to maternal-foetal incompatibility. *Ann. Hum. Genet.* 22: 111-131, 1958.
- Maumenee, I.H. Um padrão genético. *A Saúde do Mundo (fevereiro e março)*: 30-31, 1976.
- McGhee, R. & Trager, W. The cultivation of *Plasmodium lophurae* *in vitro* in chicken erythrocyte suspensions and the effects of some constituents of the culture medium upon its growth and multiplication. *J. Parasitol.* 36: 123-127, 1950.

- Miller, M.J., Neel, J.V. & Livingstone, F.B. Distribution of parasites in the red cells of sickle-cell trait carriers infected with *Plasmodium falciparum*. *Tr. Roy. Soc. Trop. Med., Hyg.* 50: 294-196, 1956.
- Morton, N.E. Non-randomness in consanguineous marriage. *Ann. Hum. Genet.*, 20: 116-124, 1955.
- Morton, N.E. Population structure and historical genetics of isolates. *Israel J. Med. Sci.*, 9: 1299-1307, 1973.
- Motulsky, A.G. Population genetics of glucose-6-phosphate dehydrogenase deficiency of the red cell. *Proc. Conf Genet. Polym. Geog. Variations Disease, U.S. Dept. Health Educ. Welf.* 258-292, 1960a.
- Motulsky, A.B. Metabolic polymorphisms and the role of infectious diseases in human evolution. *Human Biol.* 32: 28-62, 1960b.
- Motulsky, A.G. Hereditary red cell traits and malaria. *Am. J. Trop. Med. Hyg.* 13: 147-158, 1964.
- Mourant, A.E., Kopeck, A.C. & Domaniewska-Sobczak, K. *The ABO blood groups*. Blackwell Scient. Publ., Oxford, 1958.
- Ottensouper, F. & Pasqualin, R. Blood types of Brazilian Indians (Mato Grosso). *Am. J. Hum. Genet.*, 1: 141-155, 1949.
- Raper, A.B. Further observations on sickling and malaria. *Tr. Roy. Soc. Trop. Med., Hyg.* 53: 110, 1959.
- Roberts, J.A.F. Blood groups and susceptibility to disease. *Brit. J. Prev. Soc. Med.* 11: 107-125, 1957.
- Saldanha, P.H. Apparent pleiotropic effect of genes determining taste thresholds for phenylthiourea. *Lancet* 271: 74, 1956.
- Saldanha, P.H.. The estimation of relative fitness of gene determining achondroplasia based on the equilibrium between mutation and selective elimination rates. *J. Génét. Hum.* 11: 314-324, 1962a.
- Saldanha, P.H. Race mixture among northeastern Brazilian populations. *Am. Anthrop.*, 64: 751-759, 1962b.
- Saldanha, P.H. Significação genética do polimorfismo da sensibilidade à fenil-tio-uréia. *Ciência e Cultura* 16: 3 - 27, 1964.
- Saldanha, P.H. *Contribuição ao estudo da mistura racial no Brasil*. Tese, Universidade de São Paulo, 1965.
- Siniscalco, M., Bernini, L., Latte, B. & Motulsky, A.G. Favism and thalassaemia in Sardinia and their relationship to malaria. *Nature* 190: 1170-1180, 1961.
- Vandepitte, J. & Delaisse, J. Sicklemie et paludisme. *Ann. Soc. Belge Med. Trop.* 37: 703-735, 1957.
- Vogel, F. & Chakravarti, M.R. ABO blood groups and the type of leprosy in an Indian population. *Humangenet.* 3: 186-188, 1966.
- Wright, S. Evolution in mendelian populations. *Genetics* 16: 97-159, 1931.
- Yankah, J.A.K. Observation on the frequency of ABO blood groups in leprosy and non-leprosy people in Ghana. *Lep. Rev.* 36: 73-74, 1965.

## CAPÍTULO 7. EFEITO DA PREVENÇÃO DE DOENÇAS GENÉTICAS SOBRE AS FREQUÊNCIAS GÊNICAS

O aconselhamento genético e os programas de prevenção de doenças genéticas têm objetivos distintos. Uma das metas prioritárias do aconselhamento genético é ajudar indivíduos ou famílias que estão, ou supõem estar, sob risco de ocorrência ou recorrência de defeitos genéticos a tomar decisões racionais a respeito da procriação. Além disso, de acordo com o princípio bioético da autonomia, o aconselhamento genético não deve ser diretivo e deve proporcionar aos consulentes ampla compreensão de todas as implicações dessas doenças ou riscos genéticos e as opções oferecidas pela medicina atual para sua terapia ou profilaxia. A finalidade do aconselhamento genético em relação a um casal que está sob risco de gerar uma criança com determinada doença genética será, portanto, alcançada, tanto se o casal decidir evitar o nascimento dessa criança quanto, ao contrário, se esse casal permitir o nascimento dela.

Em outras palavras, o aconselhamento genético não tem por objetivo a prevenção dessas enfermidades, nem leva em conta se a opção reprodutiva dos consulentes resultará em efeitos eugênicos ou disgênicos, desde que eles tomem sua decisão livremente, depois de suficientemente informados. Ao contrário dos princípios eugênicos, que visam, primordialmente, a defesa dos interesses da sociedade, o aconselhamento genético procura a defesa do bem-estar de famílias e indivíduos, ao tentar ajudá-los a resolver problemas de natureza genética e esclarecer suas dúvidas, diminuindo ou evitando sofrimentos e preocupações.

Para exemplificar, consideremos que foi dado aconselhamento genético a dois casais igualmente jovens, que incluem um cônjuge com cegueira hereditária dominante autossômica monogênica, de manifestação precoce, em relação à qual não existe, no momento, possibilidades de correção. Conscientes do risco de 50% de transmitir essa hereditopatía, um dos casais toma a decisão de não ter filhos biológicos, porque julga que essa conduta evitará sofrimento e preocupações a seus descendentes e ao próprio casal. O outro par, contudo, decide ter filhos biológicos, porque o cônjuge com deficiência visual completa considera-se uma pessoa feliz e útil, sendo sua opinião compartilhada pelo cônjuge com visão normal. Esse casal desdenha o alto risco de gerar uma criança que se tornará cega, pois considera que ela poderá ser tão feliz e útil quanto o genitor com deficiência visual. Apesar de o primeiro casal

---

Este capítulo teve a colaboração do Prof. Dr. Henrique Krieger, do Instituto de Ciências Biomédicas da Universidade de São Paulo.

ter optado por uma medida eugênica e o segundo por uma solução disgênica, tem-se que, em ambos os casos o aconselhamento genético atingiu seu objetivo, visto que os dois casais do exemplo, além de terem demonstrado ampla compreensão de todas as implicações do defeito genético em discussão, tomaram decisões racionais em relação à procriação.

Diferentemente do que ocorre no aconselhamento genético, os programas de prevenção de hereditopatias têm, infelizmente, em graus variados, cunho diretivo que pode chegar à coerção, como é o caso da Lei de Proteção à Saúde Materna e Infantil, promulgada no dia 1º de junho de 1995 na República Popular da China. Tais programas de prevenção, que pretendem um resultado eugênico, podem resultar, ao contrário de sua intenção, em efeito disgênico, como tentaremos demonstrar no presente capítulo.

### DOENÇAS AUTOSSÔMICAS

Para iniciar, consideremos um par de alelos autossômicos  $A,a$  e que o genótipo homozigoto  $aa$  determina uma hereditopia recessiva, cuja gravidade do quadro clínico impede àqueles(as) que as manifestam de chegar à idade reprodutiva. Visto que essas pessoas  $aa$  não atingem a idade reprodutiva, está claro que somente existirão três tipos de casais na população no que diz respeito aos alelos  $A,a$  ( $AA \times AA$ ,  $AA \times Aa$  e  $Aa \times Aa$ ) e que os homozigotos  $aa$  serão gerados apenas pelos casais heterozigotos ( $Aa \times Aa$ ).

Admitamos, ainda, que o diagnóstico dos homozigotos  $aa$  é possível *in útero* e que os casais  $Aa \times Aa$  podem ser detectados em exame pré-nupcial. Se tais casais, ***que planejaram ter um determinado número de filhos***, interromperem a gestação toda a vez que o feto gerado for homozigoto  $aa$  e, em oposição, permitirem a geração de um feto não- $aa$ , isto é, com genótipo  $AA$  ou  $Aa$ , haverá ***compensação reprodutiva***, porque as perdas dos fetos  $aa$  serão repostas pela gestação de crianças com genótipo  $AA$  ou  $Aa$ . Essa prevenção da doença causada pelo genótipo  $aa$  poderá provocar um efeito disgênico, pois é fácil demonstrar que a ***freqüência do alelo  $a$  na população diminuirá menos rapidamente do que quando não se tomam medidas preventivas***.

Realmente, quando os casais  $Aa \times Aa$  não interrompem a gestação de fetos homozigotos  $aa$ , eles podem gerar crianças com genótipo  $AA$ ,  $Aa$  ou  $aa$  com probabilidades respectivamente iguais a  $\frac{1}{4}$ ,  $\frac{1}{2}$  e  $\frac{1}{4}$ , sendo os homozigotos  $aa$  geneticamente inativos, pois não alcançam a idade reprodutiva. Entretanto, quando os casais  $Aa \times Aa$  optam por

interromper a gravidez toda a vez que é gerado um feto com genótipo *aa*, mas permitem a gestação de fetos com genótipo *AA* ou *Aa*, a probabilidade de esses casais gerarem um feto com genótipo *AA* passa a ser  $\frac{1}{3}$  e a de gerar um feto *Aa* passa a ser  $\frac{2}{3}$ . De acordo com as Tabelas 1.7 e 2.7, os casais *Aa* × *Aa* que optam por interromper a gestação após diagnóstico pré-natal de um homocigoto *aa* e a fazer compensação reprodutiva, ao permitir a gestação de indivíduos *AA* e *Aa*, transmitem maior quantidade de genes *a* para a geração seguinte do que aqueles que geram o mesmo número de filhos, mas não adotam essa medida preventiva. Assim, por exemplo, casais *Aa* × *Aa* com três crianças, que adotaram compensação reprodutiva depois de interromper a gestação de fetos *aa* detectados por diagnóstico pré-natal, transmitirão dois alelos *a* à geração seguinte (Tabela 2.7) enquanto os casais com o mesmo número de filhos, que não se valem de medidas preventivas, transmitem 1,5 alelos *a* (Tabela 1.7). As Tabelas 1.7 e 2.7 permitem constatar, ainda, que a quantidade de alelos *a* transmitida à geração seguinte aumenta com o número de filhos.

Tabela 1.7. Número de alelos *a* transmitidos por casais *Aa* × *Aa* a seus filhos quando as gestantes não se submetem ao diagnóstico pré-natal.

No. de Filhos	Genótipos	Probabilidade	Alelos <i>a</i>	Alelos <i>a</i> transmissíveis
1	<i>AA</i>	1/4	-	-
	<i>Aa</i>	1/2	1	1/2
	<i>aa</i>	1/4	-	-
	<b>Total</b>			$\frac{1}{2} = 0,5$
2	<i>AA-AA</i>	1/16	-	-
	<i>AA-Aa</i>	4/16	1	4/16
	<i>Aa-Aa</i>	4/16	2	8/16
	<i>AA-aa</i>	2/16	-	-
	<i>Aa-aa</i>	4/16	1	4/16
	<i>aa-aa</i>	1/16	-	-
	<b>Total</b>			$\frac{16}{16} = 1$
3	<i>AA-AA-AA</i>	1/64	-	-
	<i>AA-AA-Aa</i>	6/64	1	6/64
	<i>AA-AA-aa</i>	3/64	-	-
	<i>AA-Aa-Aa</i>	12/64	2	24/64
	<i>AA-Aa-aa</i>	12/64	1	12/64
	<i>Aa-Aa-Aa</i>	8/64	3	24/64
	<i>Aa-Aa-aa</i>	12/64	2	24/64
	<i>AA-aa-aa</i>	3/64	-	-
	<i>Aa-aa-aa</i>	6/64	1	6/64
	<i>aa-aa-aa</i>	1/64	-	-
	<b>Total</b>			$\frac{96}{64} = 1,5$

Tabela 2.7. Número de alelos  $a$  transmitidos por casais  $Aa \times Aa$  a seus filhos quando esses casais adotam a compensação reprodutiva depois do diagnóstico pré-natal e interrupção da gestação de fetos  $aa$ .

No. de Filhos	Genótipos	Probabilidade	Alelos $a$	Alelos $a$ transmissíveis
1	$AA$	1/3	-	-
	$Aa$	2/3	1	2/3
	<b>Total</b>			<b>2/3 = 0,67</b>
2	$AA-AA$	1/9	-	-
	$AA-Aa$	4/9	1	4/9
	$Aa-Aa$	4/9	2	8/9
	<b>Total</b>			<b>12/9 = 1,33</b>
3	$AA-AA-AA$	1/27	-	-
	$AA-AA-Aa$	6/27	1	6/27
	$AA-Aa-Aa$	12/27	2	24/27
	$Aa-Aa-Aa$	8/27	3	24/27
	<b>Total</b>			<b>54/27 = 2</b>

Em vista do exposto, quando os três tipos de casais geneticamente ativos da população ( $AA \times AA$ ,  $AA \times Aa$  e  $Aa \times Aa$ ) são levados em conta, pode-se prever que, se esses casais tiverem, em média, o mesmo número de filhos, a frequência do alelo com efeito deletério ( $a$ ) diminuirá mais lentamente na população em que se tomam aquelas medidas preventivas do que naquelas em que isso não é feito. Realmente, consideremos que, numa geração inicial, a frequência dos indivíduos com genótipo homozigoto dominante ( $AA$ ) seja  $D$  e a dos heterozigotos ( $Aa$ ) seja  $H$ . Nesse caso, está claro que a frequência dos casais  $AA \times AA$  será  $D^2$ , a dos casais  $AA \times Aa$  será  $2DH$  e a dos casais  $Aa \times Aa$  será  $H^2$ .

Na Tabela 3.7 fica evidente que, se não houver medidas preventivas, a frequência do gene  $a$  na geração resultante desses casais será estimada a partir de  $H\left(\frac{D}{2} + \frac{H}{4}\right)$ . Se, entretanto, os casais  $Aa \times Aa$  optarem por interromper a gravidez toda a vez que for gerado um feto com genótipo  $aa$ , com reposição por um feto com genótipo  $AA$  ou  $Aa$ , a frequência do gene  $a$  passará a ser estimada por  $H\left(\frac{D}{2} + \frac{H}{3}\right)$ . A Fig. 1.7 dá uma noção gráfica da diferença de intensidade de eliminação de um gene autossômico com efeito deletério em homozigose quando não se tomam medidas preventivas e quando elas consistem de interrupção da gestação dos homozigotos recessivos com compensação reprodutiva.

Tabela 3.7. Distribuição de famílias numa população teórica na qual se leva em conta um par de alelos autossômicos ( $A, a$ ) e um dos genótipos ( $aa$ ) está sob seleção total. Estima-se a frequência do alelo  $a$  quando os casais não tomam medidas preventivas e quando eles fazem compensação reprodutiva depois de interromper a gestação de fetos  $aa$ .

Casais	Filhos					
	Sem medidas Preventivas			Com medidas preventivas		
	$AA$	$Aa$	$aa$	$AA$	$Aa$	$aa$
$AA \times AA (D^2)$	$D^2$	-	-	$D^2$	-	-
$AA \times Aa (2DH)$	DH	DH	-	DH	DH	-
$Aa \times Aa (H^2)$	$\frac{H^2}{4}$	$\frac{H^2}{2}$	$\frac{H^2}{4}$	$\frac{H^2}{3}$	$\frac{2H^2}{3}$	-
<b>Frequência do alelo <math>a</math> na geração filial</b>	$\frac{(DH)}{2} + \frac{H^2}{4} = H\left(\frac{D}{2} + \frac{H}{4}\right)$			$\left(\frac{DH}{2}\right) + \frac{1}{2} \cdot \frac{2H^2}{3} = H\left(\frac{D}{2} + \frac{H}{3}\right)$		

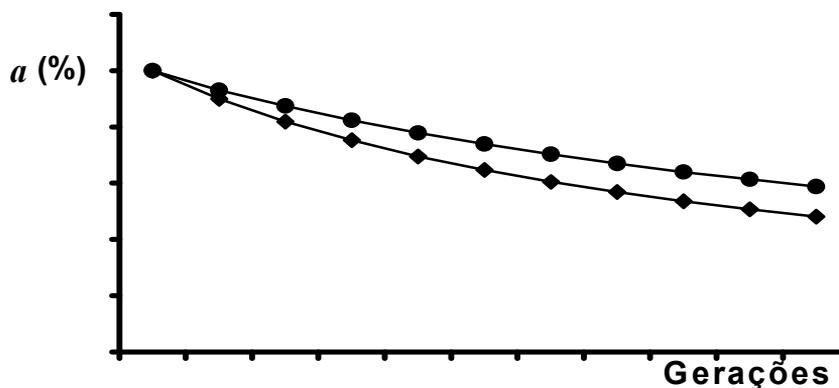


Fig. 1.7. Intensidade da eliminação de um gene autossômico com efeito deletério em homozigose ( $aa$ ) quando não se tomam medidas preventivas ( $\blacklozenge$ ) e quando se opta por compensação reprodutiva após interromper( $\bullet$ ) a gestação de fetos  $aa$ .

Visto que à medida que aumenta o número de filhos de casais  $Aa \times Aa$  também aumenta a quantidade do alelo  $a$  transmitido à geração seguinte (Tabelas 1.7 e 2.7), pode-se antever que se houver **super-compensação reprodutiva**, isto é, se as crianças não- $aa$  repostas forem mais numerosas que as perdas fetais  $aa$ , de modo que os casais  $Aa \times Aa$  tenham, em média, mais filhos que os outros casais, a frequência desse alelo nas gerações seguintes será ainda mais alta do que nos casos de compensação reprodutiva.

Evidentemente, o efeito disgênico da prevenção de hereditopatias recessivas

autossômicas será menor se a identificação dos casais  $Aa \times Aa$  somente for feita depois de haver nascido uma criança com genótipo  $aa$ . A Tabela 4.7 exemplifica essa situação, ao levar em conta o que ocorreria no caso de casais  $Aa \times Aa$  que planejaram ter dois ou três filhos (Tabelas 4.7-1 e 4.7-2), optando por interromper a gestação de fetos  $aa$  após o nascimento do primeiro filho com esse genótipo e de permiti-la no caso de o feto não ser  $aa$  ( $AA$  ou  $Aa$ ). Nelas fica claro que o aumento da transmissão do alelo  $a$  é menor do que nas irmandades correspondentes da Tabela 2.7, oriundas de casais  $Aa \times Aa$  identificados pré-matrimonialmente. Parece evidente, também, que, para a elaboração da Tabela 4.7 aceitou-se que a manifestação do genótipo  $aa$  se dá ao nascimento. Se isso não ocorrer, é claro que o efeito disgênico desse tipo de prevenção de hereditopatias recessivas diminuirá à medida que aumentar o tempo transcorrido entre ao nascimento e a manifestação da hereditopatia.

### DOENÇAS LIGADAS AO SEXO

Consideremos um par de alelos  $A, a$  do cromossomo X e que o alelo  $a$  determina uma doença grave nos homens hemizigotos ( $X^aY$ ) que os impede de alcançar a idade reprodutiva. Isso equivale a dizer que não há mulheres homozigotas  $X^aX^a$  na população. Casais que incluem uma mulher portadora do gene  $a$ , isto é, uma heterozigota  $X^AX^a$ , podem gerar filhas  $X^AX^A$  ou  $X^AX^a$  e filhos  $X^AY$  ou  $X^aY$ , tendo cada genótipo probabilidade de ocorrência igual a  $\frac{1}{4}$ . A interrupção da gestação de fetos masculinos, detectados por diagnóstico pré-natal, também poderá ter efeito disgênico e a intensidade desse efeito dependerá, por sua vez, da possibilidade de se fazer o diagnóstico pré-natal dos fetos masculinos hemizigotos  $X^aY$  e da ocorrência de compensação reprodutiva.

Consideremos, agora, casais na qual a mulher é heterozigota  $X^AX^a$  e que pretendem ter dois filhos. Nesse caso, as filhas geradas pelas mulheres  $X^AX^a$  transmitirão para a geração seguinte, em média, 0,5 alelo  $a$  ligado ao sexo (Tabela 5.7). O outro 0,5 gene  $a$  será perdido por intermédio dos filhos  $X^aY$ , que não alcançam a idade reprodutiva. Se, entretanto, todas as heterozigotas  $X^AX^a$  puderem ser identificadas pré-nupcialmente, optarem por diagnóstico pré-

Tabela 4.7 Número médio de alelos *a* transmitidos a irmandades com dois ou três indivíduos gerados por casais *Aa* × *Aa* que, depois de gerarem uma criança homocigota *aa*, decidem interromper a gestação de um próximo feto se ele for homocigoto *aa* e a adotar compensação reprodutiva.

4.7-1. Irmandades com dois indivíduos.

Irmandade	Probabilidade	Alelos <i>a</i> de heterocigotos	Alelos <i>a</i> transmissíveis
<i>AA - AA</i>	$(1/4)(1/4) = 1/16$	-	-
<i>AA - Aa</i>	$2(1/4)(2/4) = 2/16$	1	2/16
<i>Aa - AA</i>	$(2/4)(1/4) = 2/16$	1	2/16
<i>AA - aa</i>	$(1/4)(1/4) = 1/16$	-	-
<i>aa - AA</i>	$(1/4)(1/3) = 1/12$	-	-
<i>Aa - Aa</i>	$(2/4)(2/4) = 4/16$	2	8/16
<i>Aa - aa</i>	$(2/4)(1/4) = 2/16$	1	2/16
<i>aa - Aa</i>	$(1/4)(2/3) = 2/12$	1	2/12
<b>Total</b>			<b>1,04</b>

4.7-2. Irmandades com três indivíduos.

Irmandade	Probabilidade	Alelos <i>a</i> de heterocigotos	Alelos <i>a</i> transmissíveis
<i>AA-AA-AA</i>	$(1/4)(1/4)(1/4) = 1/64$	-	-
<i>AA-AA-Aa</i>	$3(1/4)(1/4)(2/4) = 6/64$	1	6/64
<i>AA-Aa-Aa</i>	$3(1/4)(2/4)(2/4) = 12/64$	2	24/64
<i>Aa-Aa-Aa</i>	$(2/4)(2/4)(2/4) = 8/64$	3	24/64
<i>AA-Aa-aa</i>	$(1/4)(2/4)(1/4) = 2/64$	1	2/64
<i>AA-aa-Aa</i>	$(1/4)(1/4)(2/3) = 2/48$	1	2/48
<i>aa-AA-Aa</i>	$(1/4)(1/3)(2/3) = 2/36$	1	2/36
<i>Aa-AA-aa</i>	$(2/4)(1/4)(1/4) = 2/64$	1	2/64
<i>Aa-aa-AA</i>	$(2/4)(1/4)(1/3) = 2/48$	1	2/48
<i>aa-Aa-AA</i>	$(1/4)(2/3)(1/3) = 2/36$	1	2/36
<i>AA-AA-aa</i>	$(1/4)(1/4)(1/4) = 1/64$	-	-
<i>AA-aa-AA</i>	$(1/4)(1/4)(1/3) = 1/48$	-	-
<i>aa-AA-AA</i>	$(1/4)(1/3)(1/3) = 1/36$	-	-
<i>Aa-Aa-aa</i>	$(2/4)(2/4)(1/4) = 4/64$	2	8/64
<i>Aa-aa-Aa</i>	$(2/4)(1/4)(2/3) = 4/48$	2	8/48
<i>aa-Aa-Aa</i>	$(1/4)(2/3)(2/3) = 4/36$	2	8/36
<b>Total</b>			<b>1,61</b>

natal para a interrupção da gestação toda a vez que for identificado um feto masculino com genótipo  $X^aY$  e fizerem compensação reprodutiva de modo a gerar uma filha ( $X^AX^A$  ou  $X^AX^a$ ) ou um menino com genótipo  $X^AY$ , ocorrerá um efeito disgênico, pois as filhas heterocigotas

$X^A X^a$  dessas mulheres transmitirão, em média, 0,667 gene  $a$  para a geração seguinte, como se pode constatar na Tabela 6.7.

Tabela 5.7. Pares de irmãos que podem ser gerados por heterozigotas de um alelo  $a$  ligado ao sexo ( $X^A X^a$ ) que impede os hemizigotos  $X^a Y$  de alcançar a idade reprodutiva. Número médio de alelos  $a$  transmitidos por intermédio das filhas heterozigotas na ausência de medidas preventivas.

Pares de irmãos	Probabilidade	Alelos $a$ de filhas heterozigotas	Alelos $a$ transmissíveis
$X^A X^A - X^A X^A$	$(1/4)(1/4) = 1/16$	-	-
$X^A X^A - X^A X^a$	$2(1/4)(1/4) = 2/16$	1	2/16
$X^A X^a - X^A X^a$	$(1/4)(1/4) = 1/16$	2	2/16
$X^A X^A - X^A Y$	$2(1/4)(1/4) = 2/16$	-	-
$X^A X^a - X^A Y$	$2(1/4)(1/4) = 2/16$	-	-
$X^A X^a - X^a Y$	$2(1/4)(1/4) = 2/16$	1	2/16
$X^A X^a - X^a Y$	$2(1/4)(1/4) = 2/16$	1	2/16
$X^A Y - X^A Y$	$(1/4)(1/4) = 1/16$	-	-
$X^A Y - X^a Y$	$2(1/4)(1/4) = 2/16$	-	-
$X^a Y - X^a Y$	$(1/4)(1/4) = 1/16$	-	-
<b>Total</b>			<b>8/16 = 0,5</b>

Tabela 6.7. Pares de irmãos que podem ser gerados por heterozigotas de um alelo  $a$  ligado ao sexo ( $X^A X^a$ ) que impede os hemizigotos  $X^a Y$  de alcançar a idade reprodutiva. Número médio de alelos  $a$  transmitidos por intermédio das filhas heterozigotas quando os casais optam por compensação reprodutiva após a interrupção da gestação de fetos  $X^a Y$ .

Pares de irmãos	Probabilidade	Alelos $a$ de filhas heterozigotas	Alelos $a$ transmissíveis
$X^A X^A - X^A X^A$	$(1/3)(1/3) = 1/9$	-	-
$X^A X^A - X^A X^a$	$2(1/3)(1/3) = 2/9$	1	2/9
$X^A X^a - X^A X^a$	$(1/3)(1/3) = 1/9$	2	2/9
$X^A X^A - X^A Y$	$2(1/3)(1/3) = 2/9$	-	-
$X^A X^a - X^A Y$	$2(1/3)(1/3) = 2/9$	1	2/9
$X^A Y - X^A Y$	$(1/3)(1/3) = 1/9$	-	-
<b>Total</b>			<b>6/9 = 0,667</b>

Consideremos, agora, o caso em que os fetos hemizigotos  $X^a Y$  não podem ser detectados por diagnóstico pré-natal. Se os casais em que a mulher é heterozigota  $X^A X^a$  e que pretendem ter dois filhos, não querendo ter um filho com a doença ligada ao sexo em discussão, optarem por interromper a gravidez toda a vez que gerarem um feto masculino, o número médio do gene  $a$  transmitido para a geração seguinte dobrará, pois tais casais passam a gerar apenas filhas  $X^A X^A$  ou  $X^A X^a$ , tendo cada um desses genótipos probabilidade de ocorrência igual a  $\frac{1}{2}$  (Tabela 7.7).

Tabela 7.7. Pares de irmãs que podem ser geradas por heterozigotas de um gene *a* ligado ao sexo ( $X^A X^a$ ) que impede os hemizigotos  $X^a Y$  de atingir a idade reprodutiva. Número médio de alelos *a* transmitidos por intermédio das filhas heterozigotas quando os casais optam por compensação reprodutiva após interrupção da gestação de fetos masculinos.

Pares de irmãos	Probabilidade	Alelos <i>a</i> de filhas heterozigotas	Alelos <i>a</i> transmissíveis
$X^A X^A - X^A X^A$	$(1/2)(1/2) = 1/4$	-	-
$X^A X^A - X^A X^a$	$2(1/2)(1/2) = 2/4$	1	2/4
$X^A X^a - X^A X^a$	$(1/2)(1/2) = 1/4$	2	2/4
<b>Total</b>			<b>4/4 = 1</b>

Em conseqüência do exposto, quando são levados em conta os dois tipos de casais da população em relação a esses alelos ligados ao cromossomo X ( $X^A X^A \times X^A Y$  e  $X^A X^a \times X^A Y$ ) pode-se prever que, se esses casais tiverem, em média, o mesmo número de filhos, o alelo com efeito deletério (*a*) diminuirá menos na população em que se tomam aquelas medidas preventivas do que naquelas em que isso não é feito. Realmente, consideremos que, numa geração inicial, a freqüência das mulheres com genótipo homozigoto dominante ( $X^A X^A$ ) seja **D** e a das heterozigotas ( $X^A X^a$ ) seja **H**. Nesse caso, está claro que a freqüência dos casais  $X^A X^A \times X^A Y$  será **D** e a dos casais  $X^A X^a \times X^A Y$  será **H**, porque todos os homens geneticamente ativos são  $X^A Y$ .

De acordo com a Tabela 8.7 fica evidente que, se não houver medidas preventivas, a freqüência do gene *a* ligado ao cromossomo X na geração seguinte será estimada a partir de  $\frac{H}{6}$ . Se, entretanto, os casais  $X^A X^a \times X^A Y$  optarem por interromper a gravidez toda a vez que for gerado um feto masculino com genótipo  $X^a Y$ , com compensação reprodutiva para gerar uma filha ou um filho com genótipo  $X^A Y$ , a freqüência do gene *a* ligado ao cromossomo X na geração seguinte será estimada a partir de  $\frac{2H}{3(3+H)}$ , como se pode constatar na Tabela 9.7. Finalmente, se os casais  $X^A X^a \times X^A Y$  interromperem a gestação de fetos masculinos, com compensação reprodutiva de fetos femininos, a freqüência do alelo *a* será estimada por intermédio de  $\frac{H}{3(1+H)}$ , de acordo com a Tabela 10.7.

Tabela 8.7. Distribuição de famílias e frequência de um alelo  $a$  ligado ao sexo em uma população teórica na qual não se tomam medidas preventivas e os hemizigotos  $X^aY$  não atingem a idade reprodutiva, sofrendo, portanto, seleção total.

Casais	Filhas		Filhos	
	$X^AX^A$	$X^AX^a$	$X^AY$	$X^aY$
$X^AX^A \times X^AY$ (D)	$\frac{D}{2}$	-	$\frac{D}{2}$	-
$X^AX^a \times X^AY$ (H)	$\frac{H}{4}$	$\frac{H}{4}$	$\frac{H}{4}$	$\frac{H}{4}$
Total ao nascer	$\frac{D}{2} + \frac{H}{4}$	$\frac{H}{4}$	$\frac{D}{2} + \frac{H}{4}$	$\frac{H}{4}$
Total ao casar, segundo o sexo	$D + \frac{H}{2}$	$\frac{H}{2}$	1	-
Frequência do alelo $a$ na geração filial	$\frac{1}{3} \cdot \frac{H}{2} = \frac{H}{6}$			

Tabela 9.7. Distribuição de famílias e frequência de um alelo  $a$  ligado ao sexo em uma população teórica na qual os hemizigotos  $X^aY$  sofrem seleção total. Os casais optam por compensação reprodutiva após interromper a gestação de fetos  $X^aY$ .

Casais	Filhas		Filhos
	$X^AX^A$	$X^AX^a$	$X^AY$
$X^AX^A \times X^AY$ (D)	$\frac{D}{2}$	-	$\frac{D}{2}$
$X^AX^a \times X^AY$ (H)	$\frac{H}{3}$	$\frac{H}{3}$	$\frac{H}{3}$
Total ao nascer	$\frac{D}{2} + \frac{H}{3}$	$\frac{H}{3}$	$\frac{D}{2} + \frac{H}{3}$
Total ao casar, segundo o sexo	$\frac{3D+2H}{3D+4H}$	$\frac{2H}{3D+4H}$	1
Frequência do alelo $a$ na geração filial	$\frac{1}{3} \cdot \frac{2H}{3D+4H} = \frac{2H}{3(3+H)}$		

Tabela 10.7. Distribuição de famílias e frequência de um gene *a* ligado ao sexo numa população teórica na qual os hemizigotos  $X^aY$  sofrem seleção completa. Os casais optam por compensação reprodutiva após interromper a gestação de fetos masculinos.

Casais	Filhas		Filhos
	$X^AX^A$	$X^AX^a$	$X^AY$
$X^AX^A \times X^AY$ (D)	$\frac{D}{2}$	-	$\frac{D}{2}$
$X^AX^a \times X^AY$ (H)	$\frac{H}{2}$	$\frac{H}{2}$	-
Total ao nascer	$\frac{D+H}{2}$	$\frac{H}{2}$	$\frac{D}{2}$
Total ao casar, segundo o sexo	$\frac{1}{1+H}$	$\frac{H}{1+H}$	1
Frequência do alelo <i>a</i> na geração filial	$\frac{H}{3(1+H)}$		

A Fig. 2.7 dá uma noção gráfica da diferença de intensidade de eliminação de um gene ligado ao cromossomo X com efeito deletério em hemizigose quando não se tomam medidas preventivas e quando elas consistem de compensação reprodutiva após interrupção da gestação dos hemizigotos do gene deletério ou dos fetos masculinos.

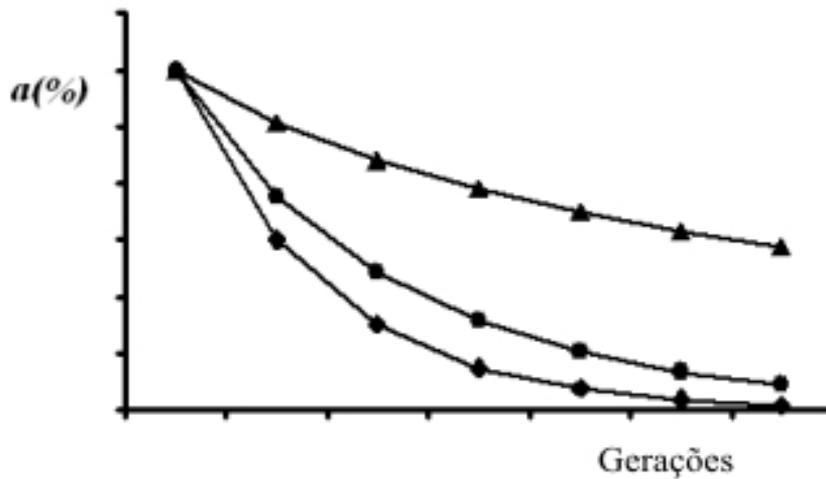


Fig. 2.7 Intensidade de eliminação de um gene ligado ao sexo que tem efeito deletério em hemizigose ( $X^aY$ ) quando os casais não tomam medidas preventivas (◆) e quando os casais optam por compensação reprodutiva depois de interromper a gestação de fetos  $X^aY$  (●) ou de fetos do sexo masculino (▲).

## COMENTÁRIOS

Se a prevenção baseada no diagnóstico pré-natal de uma determinada hereditária recessiva monogênica, autossômica ou ligada ao sexo, incluísse a interrupção da gestação dos heterozigotos do gene causador dessa hereditária, o problema da queda da eliminação desse gene, como resultado do programa de prevenção, certamente deixaria de existir. Entretanto, essa prática é absurda no aspecto lógico e indefensável do ponto de vista ético, porque a expectativa é a de que os heterozigotos serão saudáveis. Do ponto de vista científico ela é irracional, porque todos os seres humanos são, com probabilidade confinante à certeza, heterozigotos de vários genes deletérios. Desse modo, a eliminação de um feto heterozigoto de um determinado gene  $a$ , pode resultar, após compensação reprodutiva, na gestação de um outro indivíduo que, embora não sendo heterozigoto do gene  $a$ , pode ser heterozigoto de um ou mais genes não-alelos de  $a$  ( $b$ ,  $c$ ,  $d$ ,  $e$ , etc.), com efeitos muito mais graves do que o desse gene, quando se expressam em homozigose.

A exposição feita até aqui pode dar a impressão de que estamos preocupados com o efeito disgênico advindo de programas de prevenção de doenças genéticas. Nada mais falso. Temos plena consciência de que a medicina atual age em oposição à seleção natural e que o efeito disgênico resultante dos programas de prevenção de hereditárias monogênicas recessivas é semelhante àquele alcançado quando se consegue tratamento clínico ou cirúrgico para elas. De fato, consideremos, novamente, para exemplificar, um par de alelos autossômicos  $A, a$  com frequências iguais, respectivamente, a  $p$  e  $q = 1-p$ . Suponhamos, ainda, que os homozigotos  $aa$  não atingem a idade reprodutiva ou, mesmo que a atinjam, não se reproduzem. Em situações como essa, não levando em conta a ocorrência de mutações e supondo que os heterozigotos  $Aa$  não têm vantagem reprodutiva, é fácil demonstrar que, após  $n$  gerações, a frequência  $q$  do alelo  $a$  baixaria até  $q_n = \frac{q}{1+nq}$ , onde  $q_n$  é a frequência desse gene após  $n$  gerações.

Admitamos, agora, que, como resultado de um tratamento ou de uma dieta específica, cessasse o processo seletivo contra os homozigotos  $aa$ , de sorte que eles passassem a ter as mesmas possibilidades reprodutivas dos indivíduos com genótipo  $AA$  ou  $Aa$ . Tal tratamento resultaria num efeito disgênico não só porque o alelo  $a$  deixaria de ser eliminado, mas porque mais genes  $a$  seriam introduzidos na população, em decorrência da taxa de mutação do gene  $A$  para seu alelo  $a$ . Desse modo, mesmo não supondo uma vantagem seletiva para os

heterozigotos  $Aa$ , a frequência do alelo  $a$  deveria alcançar um patamar mais elevado, cujo valor ficaria na dependência da frequência de alelos  $a$  surgidos por mutação do gene  $A$  para seu alelo  $a$  na taxa  $\mu$ , isto é,  $\mu p$ , e na de alelos  $A$  surgidos por mutação inversa, na taxa  $\nu$ , ou seja  $\nu q$ .

No caso das hereditopatias monogênicas dominantes, o efeito disgênico decorrente da possibilidade de tratamento clínico ou cirúrgico dos indivíduos afetados será tanto mais acentuado quanto maior for a diferença entre o coeficiente seletivo  $s$  que esses indivíduos estavam sujeitos antes e depois da criação do tratamento. Para exemplificar, consideremos uma hereditopatia dominante que determina óbito precoce ou que impossibilita a reprodução dos indivíduos afetados por ela, o que, do mesmo modo, resulta em inatividade genética. Nesse caso, está claro que todos os indivíduos com essa hereditopatia que surgem na população são portadores do gene que sofreu mutação, o qual fica sujeito a seleção completa ( $s = 1$ , sendo  $s$  o coeficiente seletivo). A persistência média desse gene, simbolizada por  $i$ , isto é, o número médio de gerações que ele permanece na população, e que é inversamente proporcional ao coeficiente seletivo ( $i = \frac{1}{s}$ ), será, pois, igual a uma única geração, pois, com  $s = 1$  tem-se, também,  $i = 1$ . Em outras palavras, a frequência desse gene será igual à taxa de mutação.

Se, em decorrência da criação de um tratamento, o valor do coeficiente seletivo diminuir, a persistência do gene aumentará, podendo, até, permanecer indefinidamente na população. Essa situação extrema acontecerá, evidentemente, quando os portadores desse gene alcançarem valor adaptativo máximo, ou seja, quando o coeficiente seletivo for nulo. De fato, quando  $s = 0$  tem-se  $i = \frac{1}{0} = \infty$ .

Conscientes que estamos da ação anti-seletiva da medicina atual, claro está que o presente capítulo não pretende posicionar-se contra o efeito disgênico resultante da prevenção de hereditopatias recessivas, já que, nele, não se questionam as decisões que redundam em efeito disgênico, quando elas são tomadas por casais, após aconselhamento genético não-diretivo. Além disso, esse efeito não deve ser preocupante, porque o desenvolvimento espetacular das aplicações da genética molecular coloca-nos muito próximos do momento em que a substituição de genes com efeito deletério será um recurso comum do arsenal terapêutico à disposição dos clínicos.

Ao escrever o presente capítulo sobre um assunto que já foi abordado de modo diverso por Fraser (1972), a intenção foi a de enfatizar que a suposta finalidade eugênica, usada como argumento para a realização dos programas de prevenção de doenças genéticas é totalmente equivocada, mormente quando são programas coercivos. Nada justifica o ferimento do princípio da autonomia, pois ninguém deve violar o direito das pessoas de tomar decisões conscientes sobre seu estado de saúde e sobre sua reprodução ou impedir a recusa a exames genéticos e a detecção pré-sintomática de doenças, mesmo as que expõem os indivíduos a risco de morte precoce. Nenhum resultado de teste genético pode ser utilizado para influenciar a interrupção da gravidez e ninguém é obrigado a saber o resultado de um exame genético. Nada justifica a violação do princípio da privacidade, pois a informação genética e os resultados dos exames das pessoas devem ser confidenciais e continuar como tais mesmo após a morte. Nada justifica a violação do princípio da justiça porque as crianças, os deficientes mentais, os deficientes culturais, aqueles com problemas psiquiátricos, enfim, todas as pessoas juridicamente incapazes devem ter seus direitos preservados. Esperamos estar próximos do dia em que o equilíbrio justo entre o direito dos indivíduos que compõem a sociedade humana e os deveres individuais perante a saúde coletiva venham a reger o comportamento ético de todos aqueles dedicados à genética médica e constituir a meta primordial desses profissionais em todo o mundo.

## REFERÊNCIA

Fraser, G.R.- The implications of prevention and treatment of inherited disease for the genetic future of mankind. *J. Génét. Hum.* 20: 185-205, 1972.

## CAPÍTULO 8. MOLÉSTIAS INFECCIOSAS E CONSTITUIÇÃO GENÉTICA DO HOSPEDEIRO

As moléstias infecciosas devem ter começado a desempenhar um papel importante como agentes da seleção natural quando os seres humanos deixaram de constituir grupos nômades pequenos e passaram a se associar em aglomerados rurais e urbanos, mais ou menos sedentários, o que deve ter acontecido há 5.000 ou 7.000 anos. Desse modo, ao mesmo tempo que a disseminação das infecções era facilitada pelo aumento da contigüidade entre as pessoas, a seleção natural tinha oportunidade de operar, eliminando os indivíduos geneticamente suscetíveis e poupando aqueles que eram geneticamente resistente aos agentes infecciosos.

Nos países desenvolvidos as moléstias infecciosas perderam quase toda a sua importância como agentes da seleção natural. Entretanto, até há pouco tempo, numerosas populações da África, da Ásia e das Américas exibiam taxas de mortalidade antes da idade reprodutiva semelhantes às registradas na Europa do século 18. Nesse tempo, eram comuns as taxas de 200 óbitos por mil até a idade de um ano e cerca de 500 por mil até os 20 anos de idade, provocados, na maior parte, por moléstias infecciosas e carenciais.

Se tomarmos por referência a mortalidade até os 12 meses de idade, fica fácil demonstrar que nessas populações subdesenvolvidas a maioria dos óbitos na infância é conseqüência de doenças infecciosas intestinais, doenças imunopreveníveis, infecções respiratórias agudas, deficiências nutricionais, anemias carenciais e peso muito abaixo do normal ao nascer. De fato, sabendo-se, de há muito, que a incidência de anomalias genéticas graves varia de 1% a 5,5% (Neel, 1958; Stevenson, 1959; Saldanha *et al.*, 1963) pode-se estimar que, pelo menos, aproximadamente 10 por mil, isto é, 1% dos recém nascidos terão alterações patológicas constitucionais que poderão provocar óbito antes de um ano de idade.

Em outras palavras, a taxa esperada de óbitos antes de um ano de idade por causas exógenas que, no caso, são as moléstias infecciosas e carenciais, será tudo aquilo que exceder 10 por mil, pois nessa faixa etária os acidentes são raros. Por sinal, nos países desenvolvidos, bem como naqueles que, apesar de estarem em desenvolvimento, possuem

assistência médico-sanitária exemplar a taxa de óbitos até os 12 meses é menor, ainda, do que 10 por mil nascimentos.

Os dados oficiais brasileiros mostram grandes variações regionais da mortalidade de crianças até um ano de idade, as quais são um reflexo das diferenças sócio-econômicas entre as regiões do País. Assim, no ano 2000, a taxa de óbitos até os 12 meses de idade na região sul do Brasil era pouco menor do que a média geral dessa taxa no Brasil (cerca de 30 por mil nascimentos), enquanto que no nordeste do País essa taxa de óbitos era, praticamente, o dobro da observada no sul do Brasil. Apesar de a taxa de mortalidade infantil no Brasil mostrar uma forte tendência de queda, é constrangedor verificar que, em relação a esse importante indicador, ainda estamos distantes de países com economia muito inferior à do Brasil, como é o caso de Cuba e de Costa Rica, onde a taxa de mortalidade entre as crianças até um ano de idade varia entre 6 e 8 por mil. As diferenças regionais observadas no Brasil não são mantidas, entretanto, quando se analisam apenas as famílias com rendimento mensal acima de 10 salários mínimos, porque, nessas famílias, a proporção de partos institucionais é comparável à dos países os mais desenvolvidos, girando em torno de 99% em todas as regiões do Brasil (Faúndes *et al.*, 1987).

A ação seletiva das moléstias infecciosas, contudo, não tem se manifestado nos seres humanos apenas por intermédio da mortalidade antes da idade reprodutiva. No caso de infecções crônicas elas conseguem atuar como agentes seletivos por intermédio da diminuição da capacidade reprodutiva dos indivíduos por ela afetados, seja por causas de ordem biológica, que determinam esterilidade ou diminuição da fecundidade, seja por injunções sociais que restringem as possibilidades matrimoniais dos afetados. A hanseníase, por exemplo, além de poder provocar esterilidade masculina em consequência de orqui-epididimite (Beiguelman *et al.*, 1966), durante muitos anos manteve contra si grande preconceito, o que limitava as possibilidades matrimoniais dos hansenianos com pessoas sem hanseníase.

### **SUSCETIBILIDADE E RESISTÊNCIA GENÉTICA A INFECCÕES**

As investigações a respeito dos mecanismos genéticos que interferem na determinação da suscetibilidade e resistência a infecções esbarram em sérias dificuldades porque, em relação a elas, não se pode considerar o ambiente como constante, como se faz no caso de estudos genéticos sobre moléstias constitucionais e degenerativas. Por outro

lado, além da variação do ambiente e da variabilidade genética do organismo infectado, deve-se levar em conta a variação genética dos microrganismos infectantes. A simples consideração do binômio organismo infectado e microrganismo patogênico já torna o problema bastante complexo, pois sabemos que, dentre os microrganismos de uma determinada espécie, aqueles que provocam menos complicações patológicas serão selecionados favoravelmente, em virtude de sua benignidade favorecer a sobrevivência do organismo hospedeiro. O esquema tridimensional apresentado na Fig. 1.8 ilustra bem a complexidade do problema, quando se considera a taxa de sobrevivência do organismo hospedeiro como uma função do seu grau de resistência a uma bactéria patogênica que, por sua vez, apresenta diferentes graus de patogenicidade.

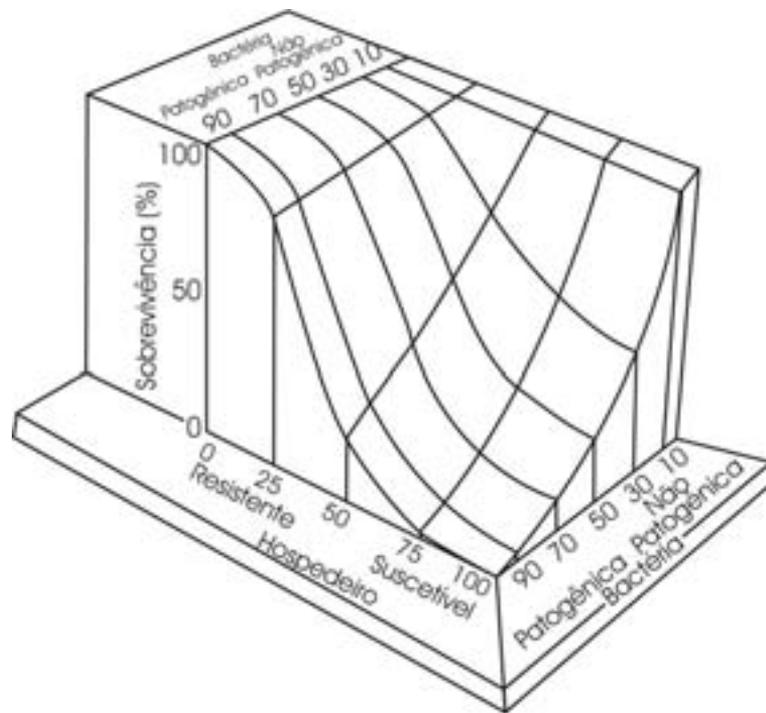


Fig. 1.8. Modelo ilustrativo do efeito da interação da constituição genética de uma bactéria infectante com a do organismo hospedeiro sobre a sua taxa de mortalidade em consequência da infecção (Adaptado de Gowen, 1952).

A validade do esquema da Fig.1.8 também se estende aos vírus, como se poderá concluir do relato sobre a mixomatose provocada nos coelhos australianos a partir de 1950, e que pode servir de modelo para a compreensão de situações semelhantes na espécie humana. A mixomatose é uma doença contagiosa e mortal no coelho europeu (*Oryctolagus cuniculus*). Ela é causada pelo ultravírus *Molitor mixomae*, que provoca edemas na cara do

animal infectado e inflamações purulentas das pálpebras, que o cegam. No coelho sul-americano (*Sylvillagus brasiliense*), que é o hospedeiro natural do vírus, a doença é benigna e caracterizada por nódulos cutâneos. Essa ação seletiva da mixomatose em coelhos europeus havia sido notada em 1898 pelo microbiologista italiano Giuseppe Sanarelli (1864-1940), quando ele trabalhou no Uruguai.

O coelho europeu, que foi introduzido na Austrália no século 19, proliferou tanto nessa região, que passou a constituir uma verdadeira praga para os agricultores e a preocupar seriamente a população. Em vista disso, o médico e pesquisador brasileiro Henrique Beaurepaire de Aragão (1879-1956), do Instituto Oswaldo Cruz do Rio de Janeiro, sugeriu que as autoridades australianas empregassem o vírus da mixomatose para controlar a multiplicação dos coelhos (Aragão, 1927). Tal conselho, dado antes de 1927, entretanto, somente foi seguido em 1950. Nessa ocasião, o vírus se espalhou rapidamente na Austrália, tendo o mosquito por vetor natural, e atingiu proporções panzoóticas, destruindo 95% dos coelhos infectados.

Após a instalação da mixomatose na Austrália verificou-se que, na região do lago Urana, na Nova Gales do Sul, passou a ocorrer, por ocasião de cada verão (setembro a novembro), uma epizootia espontânea, que incidia sobre os descendentes dos sobreviventes das epizootias anteriores. Analisando, anualmente, o efeito da mixomatose nos coelhos que, na Austrália, têm tempo médio de vida igual a um ano, pôde-se constatar um decréscimo da taxa de mortalidade por essa afecção, ano após ano. Tais estudos demonstraram que a diminuição da taxa de mortalidade decorria de duas causas, isto é, do efeito da seleção sobre o vírus, produzindo uma linhagem menos virulenta, e do efeito da seleção sobre os coelhos, produzindo uma linhagem mais resistente à infecção.

O efeito seletivo sobre os vírus pôde ser constatado pela comparação da virulência daqueles colhidos após cada surto, com o vírus-padrão, a qual mostrou, claramente, uma atenuação do efeito. Por outro lado, a seleção de uma linhagem resistente de coelhos ficou demonstrada pelos seguintes fatos:

a) houve uma relação inversa entre a taxa de mortalidade dos coelhos e o número de epizootias ocorridas entre os ancestrais desses animais;

b) os coelhos sobreviventes de cada epizootia, quando inoculados com o vírus-padrão da mixomatose, mostraram um aumento da resistência, proporcional ao número de surtos de mixomatose entre os seus ancestrais;

c) os cruzamentos entre os descendentes de ancestrais sujeitos a maior número de epizootias produziram prole mais resistente à mixomatose do que os cruzamentos entre descendentes de ancestrais sujeitos a poucas epizootias.

Uma outra fonte de dificuldades para os estudos dos mecanismos genéticos que interferem na determinação da suscetibilidade e resistência às infecções é consequência da falta de diferenciação, em numerosas situações, entre o que é **resistência genética** e o que é **resistência** ou **imunidade adquirida**. Tal problema adquire maior complexidade quando se lembra que é possível reconhecer a atuação de fatores genéticos que interferem no desenvolvimento da imunidade adquirida.

Em outras palavras, é difícil diferenciar a proteção total ou parcial contra determinada moléstia infecciosa, independentemente de contato prévio com seu agente causal (**resistência genética**) daquela adquirida permanente ou temporariamente, ativa ou passivamente (**imunidade adquirida**). A imunização ativa é aquela que depende de contato com o agente causador da infecção, enquanto que a imunização passiva é aquela adquirida pela simples introdução de anticorpos no organismo, como acontece, por exemplo, no caso das gestantes, que transmitem anticorpos a seus filhos.

#### **QUANDO É PERMISSÍVEL SUPOR UM COMPONENTE GENÉTICO IMPORTANTE DO HOSPEDEIRO QUE CONDICIONA RESISTÊNCIA A UMA INFECÇÃO?**

Já vimos no tópico anterior que a investigação de um mecanismo genético que interfere na determinação da suscetibilidade e resistência a uma infecção esbarra em sérias dificuldades. Menos difícil, porém, é a realização de pesquisas para saber se uma moléstia infecciosa satisfaz ou não a um conjunto de condições que permitam aceitar a hipótese de que os seres humanos possuem um componente genético importante que confere resistência ou suscetibilidade ao agente etiológico dessa doença. Essas condições, que analisaremos neste tópico, podem ser relacionadas como abaixo:

1. Ocorrência de indivíduos com resistência ao contágio.
2. Associação familiar da moléstia infecciosa.

3. Maior concordância da manifestação da moléstia infecciosa em gêmeos monozigóticos do que em dizigóticos.

4. Risco empírico de contrair a moléstia infecciosa correlacionado ao coeficiente de consangüinidade entre os comunicantes e o foco.

5. Variação racial da prevalência ou da incidência da moléstia infecciosa.

6. Possibilidade de sobrepor os resultados da experimentação em animais a observações na espécie humana

7. Associação entre polimorfismos e moléstias infecciosas.

É recomendável que, ao investigar a participação de um componente genético do hospedeiro na manifestação de uma moléstia infecciosa, essas condições sejam satisfeitas, pelo menos em parte, antes de submeter os dados familiares a respeito da doença em estudo a uma análise de segregação como a mencionada no tópico sobre “Análise familiar de caracteres complexos” (Lalouel *et al.*, 1983) do capítulo sobre “A análise familiar de polimorfismos”. De fato, se uma moléstia infecciosa não obedecer à maioria das condições aqui enumeradas, será pouco provável que a análise de segregação traga resultados indicadores da participação de um componente genético importante do hospedeiro na manifestação dessa moléstia.

### **1. Ocorrência de indivíduos com resistência ao contágio**

Evidentemente, a primeira condição que deve ser satisfeita para poder supor a existência de indivíduos geneticamente suscetíveis e de outros geneticamente resistentes a uma infecção, seja ela crônica ou aguda, é o encontro de pessoas que, aparentemente, sem a interferência de fatores exógenos, não se contagiam, mesmo após exposição prolongada ao agente infeccioso.

A hanseníase é um bom exemplo para demonstrar a obediência a essa condição, mas, antes, é necessário que se faça um comentário, ainda que muito breve, sobre as diferentes formas dessa doença, pois ela não é monomórfica, apesar de depender de um único agente etiológico, o *Mycobacterium leprae*, descoberto por Hansen em 1874. Assim, dependendo do cortejo fisiopatológico que suceder a infecção por esse bacilo, o hanseniano poderá ser classificado em um dos dois tipos polares da moléstia, isto é, no **tipo virchowiano** (lepra lepromatosa) ou no **tipo tuberculóide**, ou, ainda, em um dos dois grupos, isto é, no **grupo indeterminado** ou no **grupo dimorfo**.

As lesões dos pacientes virchowianos mostram diferenças antitéticas em relação às dos pacientes tuberculóides, pois enquanto as primeiras são infiltrados nos quais predominam as células de Virchow, ricas em bacilos de Hansen, as lesões tuberculóides são infiltrados do tipo sarcóide ou folicular, freqüentemente bacteriologicamente negativos. Em outras palavras, os macrófagos dos indivíduos com hanseníase virchowiana transformam-se em células de Virchow porque são incapazes de destruir o *M. leprae* que fagocitam e permitem sua proliferação e acúmulo de gordura no interior da célula. Tal incapacidade está restrita especificamente ao *M. leprae* pois, regra geral, os macrófagos de pacientes virchowianos são capazes de destruir outras micobactérias após sua fagocitose (Beiguelman, 1968<sub>b</sub>).

Em oposição, os macrófagos dos pacientes com hanseníase do tipo tuberculóide têm a capacidade de destruir os bacilos de Hansen que fagocitam, razão pela qual as lesões tuberculóides, quando examinadas ao microscópio, mostram-se abacilares ou paucibacilares. A diferença fundamental quanto à resistência tecidual à infecção hanseniana observada nesses dois tipos de hanseníase explica a contagiosidade dos pacientes virchowianos que, na ausência de tratamento, se tornam bacilíferos, bem como a falta de contagiosidade dos pacientes do tipo tuberculóide, que, regra geral, são abacilíferos. Foi por causa dessas diferenças e pelo fato de os pacientes virchowianos não poderem transformar-se em tuberculóides e vice-versa, isto é, pelo fato de eles se manterem estáveis, que essas duas formas de hanseníase foram denominadas **tipos polares**.

O grupo dimorfo inclui os pacientes que mostram similaridade com o tipo virchowiano ou com o tipo tuberculóide da hanseníase, mas não podem histologicamente ser classificados em um desses tipos polares, porque suas lesões mostram características tanto da hanseníase virchowiana quanto da hanseníase tuberculóide em reação. Os pacientes desse grupo interpolar oferecem risco de contágio porque são bacilíferos.

Alguns pacientes dimorfos, que manifestam episódios reacionais sucessivos, podem, ao final desses episódios, apresentar-se transformados clínica e histologicamente em pacientes virchowianos. Tais pacientes são os responsáveis pelo conceito de instabilidade atribuído à hanseníase dimorfa, o que faz considerá-la como um grupo e não como um tipo. Opromolla (1981), entretanto, é de opinião que os dimorfos também deveriam constituir um tipo porque, segundo suas observações, quando um paciente

dimorfo se transforma em virchowiano, essa transformação não é definitiva, pois, se depois de curados recidivarem eles voltarão a exibir lesões dimorfas e não lesões virchowianas.

O grupo indeterminado é composto por pacientes que apresentam as manifestações iniciais da hanseníase. Suas lesões são infiltrados inflamatórios simples, compostos basicamente por linfócitos, e localizados, geralmente, em torno dos pequenos filetes nervosos cutâneos e músculos *arrectores pilorum*. Nesses infiltrados os bacilos de Hansen são raros ou não demonstráveis. Quando não tratados, os pacientes indeterminados, regra geral, podem evoluir para um dos tipos polares de hanseníase ou para o grupo dimorfo, ao passo que aqueles sob tratamento podem sarar ou permanecer no grupo indeterminado durante anos. Como se vê, a designação **indiferenciado** para esse grupo, dada por Rabello (1976) parece mais adequada.

Em uma amostra de 430 cônjuges de hansenianos do tipo virchowiano, que conviveram com esses pacientes durante mais de cinco anos após o início da doença, o autor pôde constatar que apenas 4% manifestaram esse tipo de hanseníase. Dos cônjuges restantes, 86,7% não exibiram qualquer sinal dessa moléstia, enquanto que 9,3% manifestaram formas abacilíferas, pois 5,8% eram do tipo tuberculóide e 3,5% pertenciam ao grupo indeterminado (Beiguelman, 1971, 1972). Ora, tendo em mente que um indivíduo que se mantém sadio após cinco anos de contato íntimo com um hanseniano bacilífero tem pouca probabilidade de vir a se contagiar (Quagliato, 1957), os dados a respeito da manifestação da hanseníase em cônjuges de pacientes virchowianos são uma forte indicação de que a grande maioria das pessoas exibe uma resistência tecidual à proliferação do *M. leprae*. Realmente, além da alta porcentagem de cônjuges de hansenianos que não se contagiaram (86,7%), observou-se que 5,8% manifestaram o tipo tuberculóide de hanseníase. Dito de outro modo, ainda que todos os cônjuges com hanseníase indeterminada (3,5%) evoluíssem para o tipo virchowiano da doença, o que não é verossímil, poderíamos afirmar que 92,5 % das pessoas mostram, seguramente, resistência tecidual à proliferação do *M. leprae*.

Os insucessos das numerosas tentativas de infecção experimental de hanseníase em *anima nobile* realizadas no século 19 (Miguez-Alonso, 1966; Peterson e Skinsnes, 1973) falam, igualmente, a favor da hipótese de que a maioria dos seres humanos apresenta resistência à infecção pelo *M. leprae*. Nessas tentativas foram utilizados fragmentos de

hansenomas, sangue, exsudato pleural e pús de úlceras hansenóticas, que eram injetadas ou mantidos em contato com áreas escarificadas da pele de indivíduos submetidos ao experimento. De todas essas tentativas, a única inoculação que resultou em contágio foi feita em um indivíduo que era parente consanguíneo de hansenianos. Tratava-se de um assassino havaiano condenado à morte, de nome Keanu, que, aos 48 anos de idade, em 1864, optou por ser submetido a uma experiência de inoculação experimental de hanseníase, feita por Edward Arning (1854-1936), como alternativa para não ser executado. Em 1887 Keanu mostrou sinais de hanseníase virchowiana, vindo a falecer em 1892.

Muito ilustrativo da resistência tecidual que a maioria dos seres humanos oferece à proliferação do bacilo de Hansen são, também, os casos de inoculação acidental desse bacilo, que resultaram em hanseníase do tipo tuberculóide. Assim, por exemplo, é bem conhecido o relato a respeito de dois soldados norte-americanos, que haviam sido tatuados no antebraço esquerdo, num mesmo dia de junho de 1943, por um mesmo tatuador de Melbourne, Austrália (Porrit e Olsen, 1947). No período de janeiro a junho de 1946 apareceu no local das tatuagens dos dois soldados, uma lesão anestésica e eritematosa, com estrutura tuberculóide típica, sendo interessante ressaltar que um dos soldados já tinha outras tatuagens, mas a lesão tuberculóide somente apareceu naquela feita em Melbourne. Terencio de las Aguas (1967), por sua vez, descreveu a manifestação de hanseníase do tipo tuberculóide em um par de gêmeas com três anos de idade, que haviam recebido três transfusões de sangue de um doador virchowiano quando tinham a idade de 20 meses.

## **2. Associação familiar**

A constatação de que uma moléstia infecciosa apresenta associação familiar não constitui argumento forte para apoiar a hipótese de existência de um importante componente genético dos seres humanos, responsável pela resistência ou suscetibilidade ao microrganismo infeccioso, visto que essa associação pode depender mais da exposição diferencial ao agente etiológico da doença do que de uma predisposição hereditária à infecção. Contudo, apesar de a associação familiar de uma doença infecciosa não ser condição suficiente, ela é necessária para poder supor a participação de um componente hereditário humano na sua manifestação.

Os dados da Tabela 1.8, extraídos de Stern (1960), com modificações, a respeito da frequência de tuberculose pulmonar em filhos de casais sem e com essa doença em um ou

nos dois cônjuges, mostra, claramente, que a tuberculose é uma doença familiar. De fato, a proporção de doentes entre os filhos de pai tuberculoso e mãe sadia (14%) não difere significativamente daquela observada entre os filhos de pai sadio e mãe tuberculosa (13,5%) porque  $\chi^2_{(1)} = 0,106$ ;  $0,70 < P < 0,80$ . Contudo, a proporção de tuberculosos entre os filhos de casais com um dos cônjuges tuberculoso (13,5%) é significativamente maior  $\chi^2_{(1)} = 16,229$ ;  $P < 0,001$  do que a observada entre os filhos de casais sadios (8,3 %) e significativamente menor  $\chi^2_{(1)} = 36,660$ ;  $P < 0,001$  do que a observada entre os filhos de pai e mãe tuberculosos (35,7 %).

Tabela 1.8. Distribuição familiar da tuberculose pulmonar.

Genitor tuberculoso	Filhos	
	No.	Tuberculosos
Nenhum (327)	1.501	125 ( 8,3%)
Pai (88)	372	52 (14,0%)
Mãe (122)	492	65 (13,5%)
Pai e Mãe (27)	115	41 (35,7%)

O estudo do contágio intrafamiliar de hanseníase, quando um ou ambos genitores manifestavam o tipo virchowiano dessa doença, permitiu observar a distribuição apresentada na Tabela 2.8 (Beiguelman, 1971,1972). Nessas famílias somente foram incluídas pessoas com pelo menos cinco anos de coabitação com o foco virchowiano.

Tabela 2.8. Contágio em filhos e em cônjuges de focos virchowianos após pelo menos cinco anos de coabitação com o foco depois do início da doença. (V - virchowiano; T - tuberculóide; I - indeterminado; D - dimorfo). Dados de Beiguelman (1971, 1972).

Foco Virchowiano	Comunicantes	% de contagiados				
		V	T	I	D	Total
Pai (167)	Filhos (346)	11,0	1,2	4,9	0,3	17,4
	Filhas (334)	7,5	1,8	3,0	-	12,3
	Total (680)	9,3	1,4	4,0	0,2	14,9
Mãe (92)	Filhos (180)	11,1	1,7	3,9	-	16,7
	Filhas (176)	7,4	2,2	5,1	-	14,7
	Total (356)	9,3	1,9	4,5	-	15,7
Pai e Mãe (30)	Filhos (74)	25,7	5,4	8,1	-	39,2
	Filhas (55)	20,0	1,8	1,8	-	23,6
	Total (129)	23,2	3,8	5,4	-	32,4
Marido (271) Mulher (159)	Mulher (271)	2,9	6,3	3,7	-	12,9
	Marido (159)	5,0	5,0	3,1	-	13,8
	Total (430)	4,0	5,8	3,5	-	13,3

Na Tabela 2.8 fica logo evidente que, ao lidar com o risco intrafamiliar de contágio da hanseníase, devemos distinguir os tipos e grupos dessa doença. Realmente, se tal distinção não tivesse sido feita seríamos induzidos a concluir que a taxa de contágio da hanseníase entre cônjuges (13,3%), com contato íntimo e prolongado, não difere significativamente daquela observada em filhos de pai (14,9 %) ou de mãe (15,7 %) com o tipo virchowiano da doença. Entretanto, quando se levam em conta os tipos e grupos de hanseníase torna-se claro que os parentes consangüíneos de pacientes virchowianos têm maior probabilidade de mostrar o mesmo tipo polar de hanseníase (9,3 %) do que os cônjuges (não-consangüíneos) desses pacientes (4%), apesar do maior contato entre esses últimos. Essa conclusão foi confirmada por Smith *et al.* (1978), que observaram ser a hanseníase virchowiana, em famílias filipinas, cerca de três vezes mais prevalente em indivíduos com pai ou mãe virchowianos do que em filhos de pais sem hanseníase. Essa diferença não foi constatada quando os filhos de pacientes com outras formas de hanseníase foram comparados aos filhos de casais não-hansenianos.

Outra conclusão que pode ser extraída dos dados da Tabela 2.8 é que a frequência de hanseníase virchowiana nos filhos de casais com um dos cônjuges afetado por esse tipo de hanseníase não depende do sexo do genitor hanseniano. Por coincidência, a proporção de indivíduos virchowianos nas famílias em que o pai manifestava hanseníase virchowiana foi exatamente a mesma observada nos casos em que o genitor virchowiano era a mãe (9,3%).

A frequência significativamente mais alta de pacientes virchowianos entre os filhos de pai e mãe virchowianos (23,2%) está de acordo com a hipótese de que deve existir um componente genético do hospedeiro responsável pela proliferação do *M. leprae*. Entretanto, está claro que esses dados também podem ser usados para enfatizar a importância que a maior exposição ao *M. leprae* teria para a manifestação da hanseníase virchowiana.

A associação familiar de uma moléstia infecciosa, em uma população que habita uma região onde ela tem prevalência alta, também pode ser demonstrada pelo estudo de irmandades que incluem pelo menos um indivíduo com a doença (Beiguelman, 1968<sub>a</sub>). Esse tipo de análise leva em conta que, nas irmandades, se a probabilidade de encontro de um indivíduo com a moléstia infecciosa em estudo for semelhante à prevalência dessa doença na população, então o número observado de irmandades com um e com mais de um

caso afetado por tal doença não deve diferir daquele esperado numa distribuição binomial truncada. Desse modo, em  $N$  irmandades com um determinado tamanho  $n$ , o número esperado daquelas com somente um doente será  $S = \frac{N(np^{n-1}q)}{1-p^n}$ , enquanto que o número

esperado de irmandades com mais de um doente será  $M = \frac{N[1-(p^n + np^{n-1}q)]}{1-p^n}$ . Nessas

fórmulas  $q$  é a probabilidade de ocorrência de um doente e  $p = 1 - q$  é a probabilidade de isso não acontecer. Evidentemente, para conhecer o total de irmandades com somente um doente e com mais de um doente bastará obter os somatórios de  $S$  e de  $M$ , respectivamente.

Esse método foi aplicado em 1968, num estudo de 961 irmandades, que constituíam, praticamente, todas aquelas com 2 até 15 indivíduos residentes em Campinas, SP, que incluíam pelo menos um hanseniano (Beiguelman, Dall'Áglio e Silva, 1968). Depois de atribuir diferentes valores a  $q$ , constatou-se que a hipótese de recorrência familiar aleatória da hanseníase somente ficava satisfeita se a prevalência dessa moléstia tivesse um valor entre 80 a 100 por 1.000 habitantes (Tabela 3.8). Visto que isso é um absurdo, pois a prevalência da hanseníase em Campinas nunca havia atingido atingiu 3 por 1.000 habitantes, a hipótese de que a recorrência familiar dessa doença é aleatória teve que ser rejeitada, aceitando-se a hipótese alternativa de que ela apresenta associação familiar.

Tabela 3.8. Distribuição de 961 irmandades com pelo menos um hanseniano (  $n$  = tamanho da irmandade;  $N$  = número de irmandades;  $q$  = prevalência da moléstia). Dados de Beiguelman, Dall'Áglio e Silva (1968).

Irmandades		Nº. de hansenianos							Nº. esperado de irmandades com um único hanseniano		
$N$	$N$	1	2	3	4	5	6	7	$q = 0,003$	$q = 0,080$	$q = 0,100$
2	65	63	2	-	-	-	-	-	64,8*	62,3	61,6
3	109	92	10	7	-	-	-	-	108,3*	100,0*	97,7
4	132	109	19	4	0	-	-	-	130,8*	116,0	111,9
5	132	106	17	6	3	0	-	-	130,4*	111,0	105,7
6	140	107	26	5	2	0	0	-	137,9*	112,5	105,8
7	110	84	18	5	3	0	0	0	108,0*	84,4	78,4
8	108	74	18	9	4	2	0	1	105,7*	79,2	72,6
9	67	49	12	3	1	1	1	0	65,4	46,9	42,4*
10	61	44	6	4	2	2	2	1	59,4*	40,7	36,3*
11	18	15	2	1	0	0	0	0	17,5*	11,5	10,1
12	13	10	2	0	1	0	0	0	12,6*	7,9	6,8
13	4	2	2	0	0	0	0	0	3,8	2,3	2,0
14	1	0	0	1	0	0	0	0	1,0 *	0,5	0,5
15	1	1	0	0	0	0	0	0	1,0	0,5	0,4
Total	961	756	134	45	16	5	3	2	946,6*	775,8	732,2

\* Desvios significativos.

Essa conclusão não pôde ser confirmada por Morton *et al.* (1972), que estudou irmandades do Atol de Pinguelape, na Micronésia, onde a prevalência da hanseníase era extremamente alta, pois, de acordo com Sloan *et al.* (1972) ela era de 66: 1.000 e, de acordo com Morton *et al.* (1972), ela atingia 81,6: 1.000. De fato, na Tabela 4.8 pode-se constatar que a hipótese nula de recorrência aleatória da hanseníase nas irmandades estudadas por Morton *et al.* (1972) tem que ser aceita, uma vez que as proporções observadas não se desviam significativamente, seja quando se emprega  $q = 0,066$ , seja quando se usa  $q = 0,0816$ , isto é, quando se tomam as prevalências citadas, respectivamente, por Sloan *et al.* (1972) e por Morton *et al.* (1972).

Tabela 4.8. Distribuição de 79 irmandades da Micronésia com pelo menos um hanseniano ( $n$  = Tamanho da irmandade;  $N$  = número de irmandades;  $q$  = prevalência da moléstia).  
Dados de Morton *et al.* (1972).

Irmandades		Nº. de hansenianos		Nº. esperado de irmandades com um único hanseniano	
$n$	$N$	1	2	$q = 0,066$	$q = 0,0816$
2	15	15	0	14,49	14,36
3	11	8	3	10,26	10,08
4	9	8	1	8,10	7,88
5	13	12	1	11,29	10,89
6	10	8	2	8,38	7,99
7	7	5	2	5,65	5,34
8	8	6	2	6,22	5,83
9	2	0	2	1,50	1,39
10	1	1	0	0,72	0,66
11	2	1	1	1,39	1,26
12	1	0	1	0,67	0,60
Total	79	64	15	68,67	66,28

Essa discrepância decorre, evidentemente, do surpreendente número reduzido de irmandades com mais de um caso de hanseníase, no estudo de Morton *et al.* (1972). Aliás, numa população com tão alta prevalência de hanseníase, como a estudada no Atol de Pinguelape, chama a atenção não apenas esse fato, mas a constatação de que nas irmandades estudadas houve, no máximo, dois indivíduos afetados por essa moléstia.

Sabendo-se que:

1. à época das investigações dessas famílias, a frequência de pacientes virchowianos entre os hansenianos do Atol de Pinguelape era, praticamente, a metade (22%, segundo Sloan *et al.*, 1972) da observada em hansenianos brasileiros (45%, segundo Mattos, 1964);

2. nas famílias de pacientes virchowianos a probabilidade de encontro daquelas com mais de um hanseniano é o dobro da estimada para hansenianos com outra forma da moléstia (Kapoor, 1963);

parece plausível supor que a discrepância entre os resultados observados em Campinas e no Atol de Pinguelape possam ser conseqüência da menor proporção de pacientes virchowianos na população dessa ilha da Micronésia.

De qualquer modo, atualmente, todos os trabalhos que se seguiram a esses aqui mencionados vieram a confirmar a existência de associação familiar na hanseníase.

### **3. Maior concordância da moléstia infecciosa em gêmeos monozigóticos do que em dizigóticos**

Quando se constata que a proporção de pares de gêmeos concordantes quanto a uma moléstia constitucional é significativamente mais alta entre os monozigóticos do que entre os dizigóticos, aceita-se logo que tal doença depende de um componente genético importante. Essa conclusão a respeito da participação de um componente genético importante dos seres humanos também pode ser estendida a muitas moléstias infecciosas que apresentam maior concordância em pares monozigóticos do que em dizigóticos. Assim, por exemplo, em dados a respeito de 308 tuberculosos que tinham um irmãos gêmeo verificou-se que 78 eram monozigóticos e 230 dizigóticos (Kallmann e Reisner, 1943). Ao constatar que 68 dos 78 pares monozigóticos (87,2%) eram concordantes quanto à tuberculose, enquanto apenas 59 dos 230 pares dizigóticos (25,6%) mostravam concordância quanto a essa doença, ficava fácil concluir que a suscetibilidade e a resistência ao *Mycobacteriutn tuberculosis* dependem de um componente genético importante dos seres humanos, porque  $\chi^2 = 90,997$ ; 1 g.l.;  $P \ll 0,001$ .

Os dados a respeito da poliomielite paralisante em gêmeos (Herndon e Jennings, 1951) também permitem concluir a favor da importância de um componente genético humano para a manifestação dessa doença, pois a proporção de concordância de poliomielite paralisante em 14 pares monozigóticos foi de 35,7%, enquanto que em 33 pares dizigóticos apenas 6,1 % foram concordantes.

O método dos gêmeos não pode, entretanto, ser estendido a todas moléstias infecciosas, sem que sejam tomadas algumas cautelas. De fato, em relação a doenças como a hanseníase, que se manifestam sob diferentes expressões clínico-patológicas, os estudos de gêmeos não podem ficar restritos a uma simples comparação da proporção de

concordância da moléstia em pares de gêmeos monozigóticos e dizigóticos coletados aleatoriamente, sendo necessário que essa coleta obedeça algumas condições (Beiguelman, 1972, 1974, 1978)

A primeira dessas condições é a de que tanto os pares monozigóticos quanto os dizigóticos tenham as mesmas oportunidades de exposição ao agente infeccioso. A melhor maneira de satisfazer essa condição é, pois, averiguar os gêmeos a partir de pacientes que oferecem risco de contágio. Se a doença estudada fosse a hanseníase, os gêmeos deveriam ser averiguados a partir de pacientes bacilíferos. Os pacientes virchowianos são a melhor opção para servirem de caso-índice, porque os dimorfos são instáveis. Os pacientes com hanseníase tuberculóide em reação, apesar de poderem mostrar alto índice baciloscópico, também não podem ser considerados como caso-índice, porque muitos deles podem ficar bacteriologicamente negativos sem tratamento (Bechelli e Guinto, 1970).

Se a moléstia infecciosa ocorrer mais freqüentemente em indivíduos de um dos sexos, como é o caso da hanseníase, que, pelo menos em pessoas com mais 14 anos, incide mais freqüentemente naquelas do sexo masculino (Beiguelman, Silva e Dall’Aglío, 1968), deve-se comparar pares masculinos monozigóticos e dizigóticos separadamente dos femininos. Os pares de gêmeos dizigóticos de sexos diferentes não devem ser incluídos para comparação.

A terceira condição de importância fundamental é a de incluir somente pares informativos na amostra de pares monozigóticos e dizigóticos a serem investigados quanto à concordância da moléstia infecciosa. No caso da hanseníase, os pares de gêmeos que incluem pelo menos um paciente com hanseníase indeterminada ou dimorfa não podem ser considerados como informativos, porque a concordância ou discordância observada pode ser espúria, visto que tanto os pacientes indeterminados quanto os dimorfos são instáveis. Além disso, os pacientes com hanseníase indeterminada não podem ser incluídos entre os casos-índice para estudo de gêmeos, porque apresentam índice baciloscópico baixo, não sendo, por isso, contagiosos.

Os pares compostos por gêmeos com hanseníase tuberculóide também não podem ser utilizados para comparação da proporção de concordância da hanseníase, seja por causa de seu baixo índice baciloscópico, seja pelo viés que podem provocar. Realmente, pelo fato de que, às vezes, as lesões tuberculóides não são percebidas, os casos de hanseníase

tuberculóide que são esporádicos são menos freqüentemente detectados do que aqueles que ocorrem em famílias com mais de um hanseniano. Por isso, certas situações de amostragem poderiam propiciar um excesso de pares concordantes em um ou outro grupo de gêmeos, o que distorceria as conclusões em um ou outro sentido, sendo essa distorção impossível de ser avaliada.

A essa altura parece claro que, no caso da hanseníase, os gêmeos devem ser averiguados a partir de pacientes virchowianos, sendo incluídos para estudo apenas os pares que têm um gêmeo do mesmo sexo afetado por hanseníase virchowiana ou tuberculóide. Quando os dois elementos de um par são virchowianos eles são classificados como concordantes, sendo discordantes os pares compostos por um gêmeo virchowiano e outro tuberculóide. Também podem ser classificados como discordantes os pares de gêmeos que incluem um paciente virchowiano e um indivíduo sadio, desde que este último tenha convivido com o doente durante mais de cinco anos após o início da hanseníase. É claro que a inclusão desses pares para análise depende, também, da gravidade da hanseníase virchowiana e da regularidade do tratamento do paciente, visto que a contagiosidade depende desses fatores.

Lamentavelmente, os estudos de gêmeos para investigar a importância da variabilidade genética humana na determinação da suscetibilidade à infecção pelo *M. leprae*, além de poucos (Spickett, 1962; Mohamed-Ali, 1965; Mohamed-Ali e Ramanujam, 1966; Chakravarti e Vogel, 1973, 54), não levaram em consideração os requisitos aqui apresentados.

Antes de encerrar este tópico sobre estudo de gêmeos é importante assinalar que, se a proporção de concordância de pares monozigóticos for alta em relação a uma moléstia infecciosa, mas não diferir significativamente daquela observada nos pares dizigóticos, isso não servirá para concluir que não existe um componente genético importante na determinação dessa moléstia. Realmente, se a grande maioria dos indivíduos de uma população tiver um componente genético que confere suscetibilidade a um certo agente infeccioso, o fato de os gêmeos monozigóticos terem o mesmo genótipo terá pouca influência sobre os resultados, pois os dizigóticos também terão alta probabilidade de possuir o genótipo que determina suscetibilidade à infecção.

#### **4. Risco empírico de contrair a moléstia infecciosa correlacionado ao coeficiente de consangüinidade entre o comunicante e o foco.**

Se o *risco empírico* de um indivíduo contrair uma moléstia infecciosa, que é uma probabilidade calculada com base nas freqüências observadas, for proporcional ao grau de parentesco consangüíneo entre ele e o foco, esse resultado poderá ser utilizado como um argumento a favor da existência de um componente genético importante que intervém no condicionamento da suscetibilidade à infecção em estudo. Contudo, não podemos ignorar que a consanguinidade próxima está quase sempre associada com a contigüidade estreita e com período mais longo de coabitação.

#### **5. Variação racial da prevalência ou da incidência da moléstia infecciosa**

Visto que as diferenças raciais decorrem da diversidade das freqüências gênicas, a variação racial da prevalência ou da incidência de uma moléstia infecciosa poderá ser um indício da existência de um componente genético importante na determinação da suscetibilidade a seu agente etiológico, mormente se os diferentes grupos raciais analisados viverem em um mesmo habitat e não houver diferenças sócio-culturais e econômicas entre eles. Por isso, os dados a respeito da incidência média anual de poliomielite no Havaí, durante o período entre 1939 e 1947, juntamente com os estudos de gêmeos, falam a favor da existência de um componente genético determinante da suscetibilidade ao vírus causador dessa doença (Sabin, 1951). Assim, entre crianças que viviam na mesma vizinhança e frequentavam escolas sem segregação racial foram constatados os seguintes valores de incidência por 100.000: caucasóides = 10,2; parcialmente havaianos = 9,0; japoneses = 3,9; chineses = 2,7; filipinos = 1,6; havaianos = 1,3.

Em relação à tuberculose é bastante conhecido o fato de que os judeus asquenazitas (oriundos da Europa central) apresentavam menor taxa de mortalidade por tuberculose do que os não-judeus que com eles conviviam, apesar de os dois grupos mostrarem a mesma proporção de positividade à prova tuberculínica (Perla e Marmorston, 1941). O mesmo, entretanto, não ocorre com os judeus de origem iemenita, cuja alta suscetibilidade à tuberculose é atribuída ao fato de não terem estado sujeitos à pressão seletiva dessa doença, que grassou durante séculos nos guetos europeus (Dubos e Dubos, 1952).

Igualmente conhecida é a baixíssima resistência que as populações indígenas das Américas tinham à varíola, atribuindo-se a essa doença a responsabilidade pela dizimação de cerca da metade dos Aztecas. Acredita-se, inclusive, que a varíola tenha sido transmitida

a esse povo em 1520 por um dos soldados de Cortez e que um dos primeiros exemplos da guerra biológica tenha sido dado por colonizadores norte-americanos no século 19. Conhecendo a alta suscetibilidade dos índios à varíola, alguns desses colonizadores teriam distribuído, propositalmente, mantas infectadas por indivíduos variolosos aos indígenas, causando, com isso, o extermínio de grande número de índios Mandan, Assinboin e Crow (Motulsky, 1960).

A observação de que uma moléstia infecciosa ocorre com frequência elevada ou que ela causa maior taxa de mortalidade em populações que constituem pequenos isolados genéticos, não só pode falar a favor de que existe a intervenção de um componente genético na determinação da suscetibilidade à moléstia, mas que o mesmo pode ser recessivo. Isso porque nos isolados de pequeno tamanho o coeficiente médio de endocruzamento é elevado, proporcionando, assim, maior probabilidade de encontro de homozigotos. Os dados a respeito das altas taxas de contágio, de óbito e de paralisia causadas pela poliomielite em isolados canadenses da Groenlândia e em algumas ilhas vêm ao encontro das sugestões de recessividade da suscetibilidade ao vírus dessa enfermidade que foram apresentadas na literatura pertinente a partir de estudos familiares (Addair e Snyder, 1942; Aycock, 1942; Sabin, 1951).

Foi a observação de que entre os negros africanos era altíssima a proporção daqueles que mostram resistência completa à malária causada pelo *Plasmodium vivax* que conduziu à pesquisa da causa genética dessa resistência. Sabendo-se, também, que em caucasóides é um acontecimento raro o encontro de indivíduos sem os antígenos  $Fy^a$  e  $Fy^b$  do sistema sangüíneo Duffy, isto é, com o grupo sangüíneo  $Fy(a-b-)$ , decorrente do genótipo  $FyFy$ , e que em negros africanos a frequência desse grupo sangüíneo é muito alta, podendo atingir valores, muitas vezes até de 100%, como se pode constatar nos dados coletados por Mourant *et al.* (1976), passou a ser de crucial importância investigar a participação desse sistema sangüíneo na determinação da resistência ao *P. vivax*.

Graças ao trabalho de Miller *et al.* (1976) ficou-se sabendo que a presença dos antígenos  $Fy^a$  ou  $Fy^b$  ou de ambos na superfície das hemácias é condição necessária para que os merozoítos penetrem nessas células. Desse modo, os indivíduos com genótipo,  $FyFy$ , isto é, com grupo sangüíneo  $Fy(a-b-)$  ficam preservados da malária por *P. vivax*, pois suas hemácias não oferecem receptores aos merozoítos. Essa proteção somente é

encontrada em indivíduos homozigotos  $F_yF_y$ , mas não em heterozigotos do gene  $F_y$ , conforme ficou demonstrado numa investigação dos grupos sanguíneos do sistema Duffy ao nível molecular em uma área de malária endêmica (Cavasini *et al.*, 2001).

Apesar de a infecção de indivíduos saudáveis pelo *P. vivax* raramente resultar em morte, pode-se supor que ela tivesse provocado óbito ao afetar crianças subnutridas e sujeitas a outras doenças infecciosas endêmicas, como é o caso, até hoje, de numerosas crianças africanas. Desse modo poder-se-ia explicar a eliminação dos genes  $F_y^a$  e  $F_y^b$  das populações negras da África e o aumento do alelo  $F_y$ , até o *P. vivax* desaparecer dessa região. É claro, porém, que esse fenômeno notado na África pode ter outra versão. Assim, pode-se supor que as populações negras africanas eram quase todas  $F_y(a-b-)$  e que, por isso, não foi possível à malária causada pelo *P. vivax* tomar-se endêmica entre elas (Livingstone, 1984).

Aqui parece importante assinalar que o reconhecimento da explicação monogênica para a resistência dos seres humanos à infecção pelo *P. vivax* ajudou a abalar a tendência de generalizar a interpretação poligênica para todos os mecanismos de resistência e suscetibilidade às moléstias infecciosas, que existia entre os geneticistas. Por sinal, essa tendência era descabida, visto que os estudos de experimentação em animais, ao demonstrar que algumas infecções em mamíferos dependiam de sistemas monogênicos dos animais experimentais, permitiam supor que o mesmo mecanismo pudesse atuar nos seres humanos.

Realmente, de há muito já se sabia que a inoculação intracerebral do vírus da febre amarela em camundongos de quatro semanas provoca 100% de mortalidade nos da linhagem *Swiss* e nenhuma letalidade nos da linhagem *Pri* (Sabin, 1952). Tal inoculação também não provoca mortalidade nos camundongos obtidos por cruzamento *Swiss* × *Pri* (geração  $F_1$ ), nem naqueles obtidos dos cruzamentos com os da linhagem *Pri*. Contudo, a inoculação da mesma dose provoca uma proporção de mortalidade que não se desvia significativamente de 25% nos filhos de híbridos de *Swiss* com *Pri* cruzados entre si (geração  $F_2$ ), e de 50% nos híbridos de  $F_1$  cruzados com *Swiss*. Como se vê, tais resultados, resumidos na Tabela 5.8, permitem aceitar que a suscetibilidade ao vírus da febre amarela nos camundongos é determinada monogenicamente e supor que uma situação semelhante exista na espécie humana.

Tabela 5.8. Resultados da inoculação intracerebral de 10.000 vírus de febre amarela (linhagem UD) em camundongos de quatro semanas (Sabin, 1952).

Camundongos		Mortalidade (%)		Genótipo	
Linhagem	No.	Obs.	Esp.	Pais	Prole
Swiss	300	100	100	$aa \times aa$	$aa$
Pri	100	-	-	$AA \times AA$	$AA$
F <sub>1</sub>	51	-	-	$aa \times AA$	$Aa$
F <sub>2</sub>	213	28,2	25	$Aa \times Aa$	$AA: 2Aa: aa$
F <sub>1</sub> × Pri	79	-	-	$Aa \times AA$	$AA: Aa$
F <sub>2</sub> × Swiss	90	50	50	$Aa \times aa$	$Aa: aa$

Outro dado interessante obtido da experimentação animal e que pode ser estendido à espécie humana diz respeito ao fato de que certos fatores hereditários conferem resistência a vários microrganismos filogeneticamente relacionados. Assim, por exemplo, o gene que condiciona o fator depressor da multiplicação de vírus no camundongo protege-o contra todo um grupo de vírus afins (febre amarela, febre oriental do Nilo, encefalite St. Louis e encefalite tumoral primaveril), mas não lhe dá proteção a infecções por outros vírus que atacam o sistema nervoso central (encefalomielite oriental e ocidental, poliomielite, raiva, meningite linfocítica, herpes vírus e febre do Rift Valley (Sabin, 1954).

#### 6. Possibilidade de sobrepor os resultados da experimentação em animais a observações na espécie humana

A possibilidade de sobrepor os resultados da experimentação em animais a observações feitas na espécie humana pode ser usada para indicar a existência de um componente genético importante relacionado à resistência e suscetibilidade a moléstias infecciosas. Aos exemplos mencionados acima podemos acrescentar as observações feitas na Índia, a respeito da taxa de mortalidade de ratos por peste bubônica (Sokhey e Chitre, 1937), as quais têm sido usadas como argumento falando a favor da existência de um componente genético que interviria na determinação da suscetibilidade à *Pasteurella pestis*.

Esses autores inocularam uma quantidade padrão de *P. pestis* em ratos capturados em diferentes cidades da Índia e verificaram que a taxa de mortalidade por peste entre esses animais era inversamente proporcional ao período de tempo em que a cidade esteve exposta à doença. De fato, nas cidades em que não havia sido assinalada a ocorrência de peste nos 30 anos anteriores à captura dos animais, a taxa de mortalidade dos ratos atingiu 91%,

enquanto que nas cidades onde a ocorrência de peste havia sido recente a taxa de mortalidade era de 10%.

Se essa conclusão for extrapolada para as populações humanas, as observações de que na África do Sul os descendentes de europeus são mais resistentes à peste pulmonar do que os negros, e estes mais resistentes do que os de origem asiática, poderiam ser interpretadas como uma decorrência de causas genéticas. Assim, os descendentes de europeus seriam aqueles cujos ancestrais passaram pelo crivo da seleção natural causada pela peste que, no século 14, exterminou, no mínimo, um quarto da população européia. Essa interpretação, entretanto, pode ser criticada, pois as condições sócio-econômicas das populações caucasóides e não-caucasóides da África do Sul são bem diferentes.

O trabalho de Lurie e de sua escola (Lurie *et al.*, 1952), a respeito da seleção de linhagens de coelhos resistentes e suscetíveis à tuberculose, permitem tomar os resultados dessa experimentação como modelo para interpretar a suscetibilidade à tuberculose no homem. De acordo com esses trabalhos, as linhagens de coelhos altamente resistentes ao *Mycobacterium tuberculosis* cruzadas com as linhagens altamente suscetíveis ao bacilo, produziram uma geração  $F_1$  com resistência intermediária. O retrocruzamento dos indivíduos da geração  $F_1$  com os da linhagem altamente resistente deu origem a indivíduos com resistência semelhante a essa última, enquanto que o retrocruzamento dos indivíduos da geração  $F_1$  com os da linhagem suscetível deu origem a indivíduos mais resistentes do que os da linhagem ancestral suscetível. Lurie e seus colaboradores concluíram que na resistência natural à tuberculose intervêm fatores múltiplos, complexos e cumulativos, mas que os fatores que condicionam resistência à tuberculose podem ser interpretados como dominantes em relação àqueles que condicionam suscetibilidade.

### **7. Associação entre polimorfismos e moléstias infecciosas**

A investigação de possíveis associações entre polimorfismos genéticos e moléstias infecciosas é uma linha de trabalho que já foi mais utilizada por geneticistas do que ela o é hoje. Nesse tipo de pesquisa, o objetivo principal pode ser o de investigar se uma determinada moléstia infecciosa estaria incluída entre as forças seletivas que, eventualmente, contribuem para a manutenção de um ou de mais polimorfismos, ou pode ser o oposto, isto é, o de averiguar se é possível explicar a suscetibilidade e a resistência a

uma certa moléstia infecciosa por intermédio da associação entre ela e um determinado polimorfismo genético.

Aparentemente, deveria dar na mesma ser um ou outro o objetivo principal desse tipo de trabalho, porque a metodologia de investigação é idêntica. Isso, porém, não é verdade, porque o nível de exigência em relação aos resultados difere muito segundo a finalidade da pesquisa. De fato, no primeiro caso, a detecção de uma associação, ainda que pequena, entre o polimorfismo e a moléstia, desde que significativa, já satisfaz o pesquisador, porque tal resultado pode contribuir para elucidar a manutenção do polimorfismo estudado. Entretanto, quando o objetivo principal é o encontro de uma explicação para a suscetibilidade ou resistência a uma moléstia infecciosa, a associação entre o polimorfismo que está sob análise e a doença deve, evidentemente, ser muito alta, para que tal associação possa ter valor diagnóstico e(ou) prognóstico.

As pesquisas que perseguiram esses objetivos mostraram um número apreciável de resultados negativos ou contraditórios (Beiguelman, 1967, 1982, 1983) porque, na maioria das vezes, os polimorfismos genéticos investigados foram escolhidos aleatoriamente, isto é, sem que houvesse qualquer indicação lógica de que a suscetibilidade à doença estudada pudesse depender dos genes polimórficos sob análise. Isso não significa, é claro, que mesmo quando se tem uma indicação lógica para investigar a associação entre um polimorfismo e uma doença infecciosa deve-se esperar, obrigatoriamente, o encontro de um resultado positivo.

Realmente, sabe-se, por exemplo, que muitos microrganismos têm especificidades serológicas semelhantes aos aglutinógenos A, B e H do sistema sangüíneo ABO. Seria lógico, pois, que as pessoas do grupo sangüíneo A ou AB fossem mais suscetíveis a infecções por microrganismos com antígenos A-símiles, porque elas são incapazes de produzir aglutinina anti-A e, dessa maneira, se defender dessas infecções. Por razões análogas, as pessoas do grupo sangüíneo B ou AB deveriam ser mais suscetíveis a microrganismos com antígenos B-símiles, e as do grupo sangüíneo O às bactérias com antígenos H-símiles. Essas hipóteses, no entanto, não puderam ser comprovadas, talvez porque os microrganismos têm múltiplas especificidades antigênicas dentre as quais as A, B e H-símiles não são, necessariamente, as mais importantes (Springer e Wiener, 1962).

Um outro tipo de abordagem, no campo de estudos aqui em discussão, é o de procurar averiguar se um dentre os vários caracteres que definem a norma de reação dos seres humanos a um agente infeccioso constitui um polimorfismo genético (Beiguelman, 1967). Esse enfoque, que requer o conhecimento da fisiopatologia da moléstia, deu certo na hanseníase, pelo estudo da distribuição familiar da reação de Mitsuda (Beiguelman, 1962, 1965, 1971, 1982, 1983; Beiguelman e Quagliato, 1965; Feitosa *et al.*, 1996), como tivemos oportunidade de verificar no capítulo sobre “A análise familiar de polimorfismos” ao estudarmos essa reação cuja positividade indica resistência à hanseníase virchowiana (Dharmendra e Chatterjee, 1955; Quagliato, 1962).

### O ESTUDO DE ALGUNS MODELOS

As dificuldades existentes para investigar os mecanismos genéticos que interferem na determinação da resistência e suscetibilidade às infecções não devem impedir a análise de modelos a respeito da seleção natural, que levem em conta as moléstias infecciosas. Ao contrário, tal tipo de discussão deve ser estimulado, pois os genes que determinam anomalias constitucionais têm dinâmica bastante diferente daquela que apresentam os genes responsáveis pela suscetibilidade a infecções.

De fato, quando se estuda o efeito seletivo contra anomalias constitucionais, considera-se que elas sempre estiveram sujeitas a alto coeficiente seletivo, o que equivale a aceitar que os genes por elas responsáveis sempre tiveram baixa frequência na população. Contudo, em relação a genes que conferem suscetibilidade a infecções, pode-se aceitar que, em um determinado ambiente e durante um certo tempo, os caracteres determinados por tais genes não estiveram, praticamente, sujeitos a seleção, o que pode permitir que eles sejam encontrados com frequência elevada na população, pelo menos até o início do processo seletivo.

Iniciemos, pois, as discussões deste tópico supondo um par de alelos autossômicos  $A, a$  encontrados com frequências  $p$  e  $q$  iguais ( $p = q = 0,5$ ), em uma população que está em equilíbrio de Hardy e Weinberg em relação aos genótipos  $AA, Aa$  e  $aa$ . Suponhamos, ainda, que o alelo  $A$ , tanto em homozigose quanto em heterozigose confere suscetibilidade a uma infecção por um microrganismo  $X$ , que provoca a morte dos indivíduos a ele suscetíveis, e que está ausente do ambiente onde vive a população de nosso exemplo.

Se o microrganismo X for introduzido no ambiente em que vive a população teórica em apreço, várias situações poderão ser vislumbradas. A mais extrema é aquela na qual todos os indivíduos  $AA$  e  $Aa$  são eliminados da população em uma única geração, após uma epidemia que exterminaria praticamente 75% dos indivíduos que a compõem, pois admitimos que os alelos  $A$  e  $a$  têm frequências idênticas ( $p = q = 0,5$ ). Outras situações que podemos entrever levam em conta a possibilidade de o coeficiente seletivo contra os portadores do gene  $A$  não aumentar de  $s = 0$  para  $s = 1$ , mas que  $s$  tome valores intermediários ( $0 < s < 1$ ) em consequência de:

a) terapêutica que impeça a mortalidade de parte dos portadores do gene  $A$ , mas não impeça as seqüelas da doença provocada pelo organismo X em parte ou em todos os indivíduos  $AA$  e  $Aa$ ;

b) ocorrência de mutantes do microrganismo X, que provocam uma moléstia sem as características de malignidade do microrganismo original;

c) distribuição geográfica ou estratificação da população, capaz de proporcionar diferentes oportunidades de contágio aos portadores do alelo  $A$ .

Nessas condições, é evidente que a frequência do gene  $A$  diminuiria mais lentamente, sendo o processo seletivo sustado somente após a erradicação do microrganismo X ou pela vacinação de todos os indivíduos suscetíveis.

Consideremos, agora, que, na população de nosso exemplo, o gene  $a$  fosse o responsável pela suscetibilidade à infecção fatal provocada pelo microrganismo X. Nesse caso, admitindo seleção total de todos os indivíduos homozigotos  $aa$ , ter-se-ia, no momento em que se iniciasse o processo seletivo, que a eliminação do gene  $a$  seria feita com grande intensidade durante as primeiras gerações de seleção. Entretanto, à medida que esse gene diminuísse a sua frequência, o efeito da seleção seria menos intenso, como se pode verificar na Figura 2.8. Assim, duas gerações após a inicial a frequência do alelo  $a$  se reduziria à metade, pois passaria de  $q = 0,5$  para  $q_2 = \frac{q}{1+2q} = 0,25$ . Já, por exemplo, a diferença da frequência do alelo  $a$  entre a 10<sup>a</sup> e a 11<sup>a</sup> geração sob seleção total dos indivíduos  $aa$  seria mínima porque  $q_{10} = 0,0833$  e  $q_{11} = 0,0769$ , o que equivale a dizer que a frequência do alelo  $a$  na 11<sup>a</sup> geração corresponderia a 92% da observada na geração anterior.

Se admitirmos a atuação de fatores que tornem a seleção dos indivíduos  $aa$  incompleta, concluiremos que tais fatores serão mais eficazes para a preservação do gene  $a$  do que para a preservação do alelo  $A$  quando os portadores de  $A$  estão sujeitos a processo seletivo. Em condições semelhantes, a seleção contra portadores de genes com efeito dominante é mais eficiente do que aquele que opera contra indivíduos com um fenótipo recessivo.

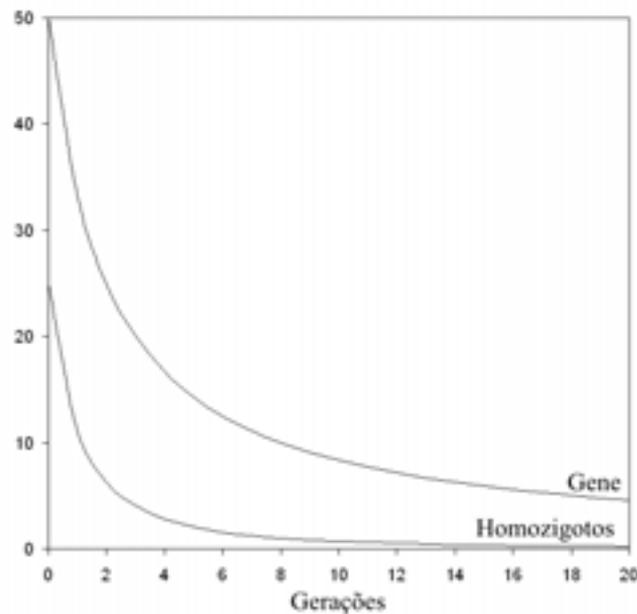


Figura 2.8. Efeito da seleção total, durante 20 gerações, contra um gene que somente se expressa em homozigose e que, na geração inicial, ocorre com alta frequência na população.

Consideremos, agora, o caso de a suscetibilidade à infecção pelo microrganismo X ter determinação poligênica. A título de exemplificação levemos em conta apenas dois pares de alelos ( $A,a$  e  $Bb$ ), em uma população em equilíbrio genético em relação aos genótipos por eles determinados, e analisemos os resultados de duas situações alternativas, isto é, que somente os indivíduos com o genótipo  $aabb$  são suscetíveis ou que somente eles são resistentes à infecção pelo microrganismo X.

No caso de apenas os indivíduos  $aabb$  serem suscetíveis, o efeito seletivo seria comparável àquele em que o genótipo  $aabb$  determina uma anomalia constitucional sujeita à seleção. Em outras palavras, o processo seletivo contra o sistema poligênico será menos eficiente do que aquele que opera contra sistemas monogênicos. Por outro lado, em

decorrência da lentidão da eliminação dos genes de sistemas poligênicos, haverá mais tempo para o desenvolvimento de mutantes do microrganismo X que provocam infecções com características benignas.

Ao considerar a segunda alternativa, isto é, a de que somente os indivíduos *aabb* são resistentes à infecção fatal pelo microrganismo X, tem-se que, em uma população na qual os genes *A,a* e *B, b* ocorrem com a mesma frequência ( $p_1 = q_1 = p_2 = q_2 = 0,5$ ) haveria uma verdadeira catástrofe quando esse microrganismo fosse introduzido na população. De fato, em uma única geração poderia haver a eliminação de 93,75% dos indivíduos, porque  $1 - q^4 = 1 - 0,0625 = 0,9375$ .

É óbvio que as conclusões a respeito de dois pares de alelos podem ser extrapoladas para situações em que se deve aceitar a participação de mais de dois pares. Desse modo, considerando a primeira alternativa, isto é, a de que somente os indivíduos homocigotos em relação a todos os pares de alelos estariam sujeitos a seleção, ter-se-ia que, em relação a um grande número de pares de alelos, a seleção seria tanto menos eficiente quanto maior fosse o número desses pares. Levando em conta a segunda alternativa, isto é, a de que somente os homocigotos em relação a todos os pares de alelos seriam preservados, ter-se-ia que os resultados seriam tanto mais catastróficos quanto maior fosse o número de pares de genes em jogo na determinação do fenótipo. Evidentemente, é possível vislumbrar outras situações que levam em conta a seleção contra vários estados heterocigóticos.

## REFERÊNCIAS

- Addair, J. & Snyder, L.H. Evidence for an autosomal recessive gene for susceptibility to paralytic poliomyelitis. *J. Hered.* 33: 307-309, 1942.
- Aragão, H.B. Myxoma dos coelhos. *Mem. Inst. Oswaldo Cruz* 20: 225-236, 1927. (Myxoma of rabbits. *Mem. Inst. Oswaldo Cruz* 20:237-247, 1927).
- Aycock, W.L. Familial aggregation in poliomyelitis. *Am. J. Med. Sci.* 203: 452-465, 1942.
- Bechelli, L.M. & Guinto, R.S. Some recent laboratory findings of *Mycobacterium leprae*. *Bull. W. H. O.* 43: 559-569, 1970.
- Beiguelman, B. Hereditariedade da reação de Mitsuda. *Rev Bras. Leprol.* 30: 153-172, 1962.
- Beiguelman, B. The genetics of resistance to leprosy. *Int. J. Lepr.* 33: 808-812, 1965.
- Beiguelman, B. Leprosy and Genetics. A review of past research with remarks concerning future investigations. *Bull. W.H.O.* 37: 461-476, 1967.

- Beiguelman, B. *Dinâmica dos genes nas populações e nas famílias*. EDART S. Paulo Livraria Editora Ltda., 1968a.
- Beiguelman, B. Some remarks on the genetics of leprosy resistance. *Acta Genet. Med. Gemellol.* 17:584-594, 1968b.
- Beiguelman, B. Lepromin reaction. Genetic studies including twin pair analysis. *Acta Leprol.* 44: 5-65, 1971.
- Beiguelman, B. An appraisal of genetic studies on leprosy. *Acta Genet. Med. Gemellol.* 21: 21-52, 1972.
- Beiguelman, B. Um programa multinacional de investigação leproológica utilizando o estudo, de gêmeos. *Ciência e Cultura* 26: 459-468, 1974.
- Beiguelman, B. *Genetics in leprosy*. Em Chatterjee, B.R. (Ed.) *A window on leprosy*, Gandhi Memorial Leprosy Foundation, The Statesman Commercial Press, Calcutta, India: 71-81, 1978.
- Beiguelman, B. Lepra e Genética. *Ciência e Cultura* 34: 1121-1146, 1982.
- Beiguelman, B. Leprosy and Genetics: a review. *Rev. Bras. Genet.* 6: 109-172, 1983.
- Beiguelman, B. & Quagliato, R. Nature and familial character of the lepromin reactions. *Int. J. Lepr.* 33: 800-807, 1965.
- Beiguelman, B., Dall'Aglio, F.F. & Silva, E. da Análise da recorrência familiar da lepra. *Rev. Paul. Med.* 72: 105-110, 1968.
- Beiguelman, B., Silva, E. da & Dall'Aglio, F.F. Lepra e sexo. *Rev. Paul. Med.* 72:120-129, 1968.
- Beiguelman, B., Marchi, A., Hama, T., Amin, C.C., Godoi, M.N.C. & Baptista, T.A. Fecundidade e lepra. *Rev. Paul. Med.* 66: 207-213, 1966.
- Cavasini, C.E., Pereira, F.J.T., Ribeiro, W.L., Wunderlich, G. & Ferreira, M.U. Duffy blood group genotypes among malaria patients in Rondônia, Western Brazilian Amazon. *Rev. Soc. Bras. Med. Trop.* 34: 591-595, 2001.
- Chakravarti, M.R. & Vogel, F. *A twin study on leprosy*. Geord Thieme Publisliers, Stuttgart, 1973.
- Dharmendra & Chatterjee, B.R. Prognostic value of the lepromin test in contacts of leprosy cases. *Leprosy in India* 27: 149-152, 1955.
- Dubos, R. & Dubos, J. *The white plague: tuberculosis, man and society*. Little, Brown & Co., Boston, 1952.
- Faúndes, A., Cecatti, J.G., Silva, J.L.C.P. & Pinotti, J.A. O problema da mortalidade materna. *Femina* 15: 25-31, 1987.
- Feitosa, M., Krieger, H., Borecki, I., Beiguelman, B. & Rao, D.C. Genetic epidemiology of the Mitsuda reaction in leprosy. *Hum. Hered.* 46:32-35, 1996.
- Gowen, J. Humoral and cellular elements in natural and acquired resistance to typhoid. *Am. J. Hum. Genet.* 4: 285-302, 1952.
- Herndon, C.N. & Jennings, R.G. A twin-family study of susceptibility to poliomyelitis. *Am. J. Hum. Genet.* 3: 17-46, 1951.
- Kallmann, F.J. & Reisner, D. Twin studies on significance of factors in tuberculosis. *Ann. Rev. Tuberc.* 47: 549-574, 1943.

- Kapoor, P. Epidemiological survey of leprosy in Maharashtra State (India). *Leprosy in India* 35: 83-89, 1963.
- Lalouel, J.M., Rao, D.C., Morton, N.E. & Elston, R.C. A unified model for complex segregation analysis. *Am. J. Hum. Genet.* 35: 816-826, 1983.
- Livingstone, F.B. The Duffy blood groups, vivax malaria, and malaria selection in human populations: a review. *Hum. Biol.* 56: 413-425, 1984.
- Lurie, M.B., Zappasodi, P., Dannenberg Jr., A.M. & Weiss, G.H. On the mechanism of genetic resistance to tuberculosis and its mode of inheritance. *Am. J. Hum. Genet.* 4: 302-314, 1952.
- Mattos, O. Campanha nacional contra a lepra. Relatório da Assessoria Técnica. *Bol. Serv. Nac. Lepra* 24: 239-275, 1964.
- Miguez-Alonso, A. *Lepra dimorfa. Fundamentos de sua conceituação*. Editora Livro S.A., Rio de Janeiro, 1966.
- Miller, L.H., Mason, S.J., Clyde, D.T. & McGinniss, M.H. The resistance factor to *Plasmodium vivax* in Blacks. The Duffy-blood-group genotype *FyFy*. *N. Engl. J. Med.* 295: 302-304, 1976.
- Mohamed-Ali, P. Genetic influence in leprosy. *Leprosy in India* 37:252-267,1965.
- Mohamed-Ali, P. & Ramanujam, K. Leprosy in twins. *Int. J. Lepr.* 34: 405-407,1966.
- Morton, N.E., Lew, R., Hussels, I.E. & Little, G.F. Pingelap and Mokil Atolls: historical genetics. *Am. J. Hum. Genet.* 24: 277-289,1972.
- Motulsky, A.G. Metabolic polymorphisms and the role of infectious diseases in human evolution. *Human Biol.* 32: 28-62, 1960.
- Mourant, A.E., Kopec, A.C. & Domaniewska-Sobczak, K. - *The distribution of the human blood groups and other polymorphisms*, 2a. ed., Oxford University Press, London, 1976.
- Neel, J.V. The study of major congenital defects in Japanese infants. *Am. J. Hum. Genet.* 10: 398-445, 1958.
- Opromolla, D.V.A. *Noções de Hansenologia*. Centro de Estudos Dr. Reynaldo Quagliato, Hospital Lauro Souza Lima, Bauru, SP, Brasil, 1981.
- Perla, D. & Marmorston, J. *Natural resistance and Clinical Medicine*. Little, Brown & Co., Boston, 1941.
- Peterson, L.W. & Skinsnes, O.K. Photographic gallery of contributors of the century. *Int. J. Lepr.* 41: 156-176,1973.
- Porrit, R.J. & Olsen, R.E. Two simultaneous cases of leprosy developing in tatoos. *Am. J. Pathol.* 23: 805-817,1947.
- Quagliato, R. Lepra conjugal. *Rev. Bras. Leprol.* 25: 59-68, 1957.
- Quagliato, R. Interpretação das reações limítrofes ou duvidosas do teste lepromínico. *Bol. Serv. Nac. Lepra* 21: 13-34, 1962.
- Rabello, F.A. The indeterminate group of hanseniasis, and its basic connotation: the polar concept. An evaluation and a refutation of the so-called "spectral" approach. *Hansen. Int.* 1: 111- 119, 1976.

- Sabin, A.B. Paralytic consequences of poliomyelitis infection in different parts of the world and in different population groups. *Am. J. Publ. Health.* 41: 1215-1230, 1951.
- Sabin, A.B. Genetic, hormonal and age factors in natural resistance to certain viruses. *Ann. N. Y Acad. Sci.* 54: 936-944, 1952.
- Sabin, A.B. Genetic factors affecting susceptibility and resistance to virus diseases of the nervous system. Em Hooker, D. & Hare, C.C. (Eds.) *Genetics and the inheritance of integrated neurological and psychiatric patterns*. Williams & Wilkens Co., Baltimore, 1954.
- Saldanha, P.H., Cavalcanti, M.A. & Lemos, M.L. Incidência de defeitos congênitos na população de São Paulo. *Rev. Paul. Med.* 63: 211-229, 1963.
- Sloan, N.R., Worth, R.M., Jano, B., Fasal, P. & Shepard, C.C. Acedapsona in leprosy treatment: trial in 68 active cases in Micronesia. *Int. J. Lepr.* 40: 48-52, 1972.
- Smith, D.G., Blumberg, B.S., Guinto, R.S. & Wittenstein, FS. Genetics in Leprosy. Em Chatterjee, B.R. (Ed.) *A window in leprosy*, Gandhi Memorial Leprosy Foundation, The Statesman Commercial Press, Calcutta, India: 82-102, 1978.
- Sokhey, S.S. & Chitre, R.B.G.D. L'immunité des rats sauvages de l'Inde vis-à-vis de la peste. *Bull. Off. Int. Hyg., Publ.* 29: 2093-2096, 1937.
- Spickett, S.G. Genetics and the epidemiology of leprosy. II - The form of leprosy. *Lepr. Rev.* 33: 173-181, 1962.
- Springer, G.F. & Wiener, A.S. Alleged causes of the present-day world distribution of the human ABO blood groups. *Nature* 194: 444-451, 1962.
- Stern, C. *Principles of Human Genetics* W.H. Freeman and Co., San Francisco, USA, 1960.
- Stevenson, A.C. The load of hereditary defects in human population. *Radiat. Res. (Suppl. 1)*: 306-325, 1959.
- Terencio De Las Aguas, J. Inoculación accidental de la lepra por transfusión sanguínea en gemelos univitelinos. *Fontilles* 6: 603-611, 1967.