

Chương 11

Di truyền học quần thể

Di truyền học quần thể (population genetics) nghiên cứu sự phân bố của các gene trong quần thể cũng như sự duy trì và thay đổi tần số của các gene và kiểu gen qua các thế hệ. Di truyền học quần thể không những liên quan đến những yếu tố di truyền như đột biến, sinh sản mà còn liên quan đến những yếu tố xã hội và môi trường như sự chọn lọc, tình trạng di dân và đương nhiên liên quan đến việc nghiên cứu tần số và sự phân bố các bệnh di truyền trong gia đình và trong cộng đồng.

I. Tính đa dạng di truyền trong quần thể loài người

Trái đất với trên 6 tỷ con người được phân thành những chủng tộc khác nhau, những quần thể dân cư khác nhau. Mặc dù tất cả đều mang bộ nhiễm sắc thể giống nhau, sự phân bố của các gene trên NST cũng như nhau nhưng bản chất của các allele và tần số của các allele tại mỗi vị trí của gene trên NST là rất khác nhau giữa các dân tộc, các chủng tộc.

Sự chọn lọc của các đột biến có lợi dưới tác động của môi trường, cơ hội sống sót của các cá thể mang các đột biến có hại, các đột biến trung tính cùng với sự cách ly sinh sản giữa các nhóm tạo nên sự khác biệt di truyền rất lớn giữa các quần thể.

II. Cách xác định tần số của các allele trong quần thể

Nếu chúng ta có thể biết được kiểu gen thật sự của mỗi cá thể trong một gia đình đang cần tư vấn di truyền, chắc chắn chúng ta có thể thông báo cho họ về nguy cơ tái phát của bệnh đó với độ chính xác cao. Tuy nhiên trong thực tế nhiều trường hợp chỉ có kiểu hình của bệnh là cái duy nhất mà chúng ta có thể quan sát và đánh giá. Do đó tỷ lệ của một bệnh lí di truyền hoặc một đặc điểm di truyền thường được sử dụng để xác định tần số của một kiểu gen đặc hiệu rồi qua đó suy ra tần số của các allele đặc hiệu chịu trách nhiệm cho các kiểu gen khác nhau.

Dưới đây là ví dụ về một đặc điểm được chi phối bởi một cặp allele nằm trên NST thường minh họa cho cách thức xác định tần số của các allele trong quần thể.

Gen *CCR5* chịu trách nhiệm mã hóa cho một receptor cytokin trên bề mặt tế bào, receptor này đóng vai trò như một điểm vào của các chủng virus gây suy giảm miễn dịch ở người (HIV) gây hội chứng suy giảm miễn dịch mắc phải (acquired immunodeficiency syndrome, AIDS). Một đột biến mất 32 cặp nucleotide trong gen *CCR5* làm xuất hiện allele $\Delta CCR5$

mã hóa cho một protein không có chức năng do đã bị thay đổi trong cấu trúc và bị chấm dứt sớm việc tổng hợp.

Những người đồng hợp tử về allele $\Delta CCR5$ ($\Delta CCR5/\Delta CCR5$) sẽ không có loại receptor cytokin trên bề mặt tế bào và tạo nên khả năng đề kháng với HIV. Việc mất chức năng của gene $CCR5$ là một đặc điểm lành tính và được xác định dựa trên khả năng đề kháng với HIV. Allele bình thường và allele đột biến mất 32 cặp base $\Delta CCR5$ được phân biệt một cách dễ dàng qua phân tích gene bằng kỹ thuật PCR. Một mẫu nghiên cứu trên 788 người Âu châu đã cung cấp số lượng thông tin đầy đủ về số cá thể đồng hợp hoặc dị hợp của 2 allele nói trên

Bảng 1: Tần số gene và kiểu gene của allele bình thường $CCR5$ và allele đột biến $\Delta CCR5$

Kiểu gene	Số người	Tần số tương đối của các kiểu gene	Allele	Tần số của các allele
$CCR5/CCR5$	647	0,821		
$CCR5/\Delta CCR5$	134	0,1682	$CCR5$	0,906
$\Delta CCR5/\Delta CCR5$	7	0,0108	$\Delta CCR5$	0,094
Tổng	788	1,000		1,000

Dữ kiện của Martinson JJ, Chapman NH, Rees DC và cộng sự (1997): Global distribution of the CCR5 gene 32 basepair deletion. Nat Genet 16:100-103

Trên cơ sở tần số kiểu gen quan sát được chúng ta có thể trực tiếp xác định tần số của các allele một cách đơn giản thông qua việc đếm các allele. Khi chúng ta tính toán tần số trong quần thể của một allele, chúng ta sẽ xem vốn gen (gene pool) như là một tập hợp tất cả các allele ở một locus đặc hiệu trong toàn bộ quần thể. Đối với các locus trên NST thường, kích cỡ của vốn gene ở một locus sẽ gấp đôi số cá thể trong quần thể vì mỗi kiểu gen trên NST thường sẽ gồm có hai allele, nghĩa là người có kiểu gene $\Delta CCR5/\Delta CCR5$ sẽ có 2 allele $\Delta CCR5$, và người có kiểu gen $CCR5/\Delta CCR5$ sẽ có 1 allele $CCR5$ và 1 allele $\Delta CCR5$.

Trong ví dụ này tần số của các allele $CCR5$ sẽ là:

$$(2 \times 647 + 1 \times 134) / 788 \times 2 = 0,906$$

Tương tự, người ta có thể tính toán được tần số của allele $\Delta CCR5$ là 0,094, hoặc bằng cách tính trực tiếp từ số cá thể:

$(2 \times 7 + 1 \times 134) = 148$ trong tổng số 1576 allele, hoặc đơn giản là lấy 1 trừ cho tần số của allele $CCR5$: $1 - 0,906 = 0,094$.

III. Định luật Hardy - Weinberg

Một câu hỏi được đặt ra là làm thế nào để có thể tính được tần số của các kiểu gene khác nhau trong quần thể khi biết được tần số của các allele? Làm thế nào để có thể tính được tần số của các allele, tần số của các kiểu gene đồng hợp và dị hợp khi chỉ có thể biết được tỷ lệ của một kiểu hình hoặc một bệnh di truyền trong quần thể ?

Định luật Hardy-Weinberg, được nhà toán học Geoffrey Hardy (Anh) và bác sĩ Wilhelm Weinberg (Đức) đồng thời phát hiện năm 1908, cho phép đánh giá tần số của các kiểu gene từ tần số của các allele qua một công thức toán học đơn giản nếu quần thể thoả mãn một số điều kiện nhất định và đây là định luật đặt nền tảng cho di truyền học quần thể.

Định luật Hardy-Weinberg gồm 2 tính chất quan trọng:

(1) Giả sử gọi p là tần số của allele A, q là tần số của allele a và các allele kết hợp với nhau một cách ngẫu nhiên trong quần thể thì tần số của các kiểu gene AA, Aa, aa được thể hiện qua nhị thức:

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

Bảng 2: Tần số của các kiểu kết hôn và tần số các kiểu gene ở thế hệ sau theo định luật Hardy-Weinberg.

Kiểu kết hôn			Con		
Bố	Mẹ	Tần số	AA	Aa	aa
AA	AA	$p^2 \times p^2 = p^4$	$1(p^4)$		
AA	Aa	$p^2 \times 2pq = 2p^3q$	$1/2(2p^3q)$	$1/2(2p^3q)$	
Aa	AA	$2pq \times p^2 = 2p^3q$	$1/2(2p^3q)$	$1/2(2p^3q)$	
AA	aa	$p^2 \times q^2 = p^2q^2$		$1(p^2q^2)$	
aa	AA	$q^2 \times p^2 = p^2q^2$		$1(p^2q^2)$	
Aa	Aa	$2pq \times 2pq = 4p^2q^2$	$1/4(4p^2q^2)$	$1/2(4p^2q^2)$	$1/4(4p^2q^2)$
Aa	aa	$2pq \times q^2 = 2pq^3$		$1/2(2pq^3)$	$1/2(2pq^3)$
aa	Aa	$q^2 \times 2pq = 2pq^3$		$1/2(2pq^3)$	$1/2(2pq^3)$
aa	aa	$q^2 \times q^2 = q^4$			$1(q^4)$
Tổng			p^2	$2pq$	q^2

(2) Tần số của các kiểu gen sẽ không đổi qua các thế hệ miễn là tần số của các allele không đổi. Nghĩa là nếu tần số của các kiểu gene AA, Aa

và aa phân bố trong quần thể theo tỷ lệ $p^2 : 2pq : q^2$ thì tần số này ở các thế hệ sau cũng sẽ là $p^2 : 2pq : q^2$. Trạng thái này được gọi là trạng thái cân bằng của quần thể (Bảng 2).

Trong ví dụ về gene *CCR5* và *ΔCCR5* nói trên với tần số của allele *CCR5* là 0,906 và *ΔCCR5* là 0,094 thì theo định luật Hardy-Weinberg tần số tương đối của các tổ hợp allele sẽ là:

$$CCR5/CCR5: p^2 = 0,906 \times 0,906 = 0,821$$

$$CCR5/\Delta CCR5: 2pq = 2 \times 0,906 \times 0,094 = 0,170$$

$$\Delta CCR5/\Delta CCR5: q^2 = 0,094 \times 0,094 = 0,009$$

Sự phân bố này cũng tương tự với tần số thực tế (647 : 134 : 7). Như vậy chúng ta thấy định luật Hardy - Weinberg là một nhị thức với hai allele của một gene, trong đó $p + q = 1$ và $n = 2$ (n là số allele). Trong trường hợp gene nằm trên NST giới tính X thì tần số của các kiểu gene của người nữ (có 2 NST X) và người nam (có 1 NST X) sẽ được tính riêng (Bảng 3)

Nếu một gene có tới 3 allele ($n = 3$) : p, q và r thì sự phân bố tần số của các allele sẽ tuân theo công thức: $(p + q + r)^2$. Nói chung với n allele ($p_1, p_2, p_3 \dots p_n$) thì công thức sẽ là $(p_1 + p_2 + p_3 + \dots p_n)^2$.

1. Ứng dụng của định luật Hardy-Weinberg

Ứng dụng chính của định luật Hardy-Weinberg trong di truyền y học là tư vấn di truyền cho các bệnh di truyền gene lặn NST thường.

Ví dụ trong bệnh Phenylketonuria (PKU), tần số của người mắc bệnh ở trạng thái đồng hợp sẽ được xác định chính xác trong quần thể qua chương trình sàng lọc trên trẻ sơ sinh ở Ireland là 1/4.500, định luật Hardy-Weinberg cho phép xác định tần số của những người dị hợp tử có biểu hiện hoàn toàn bình thường:

$$\text{vì } q^2 = 1/4500 \text{ nên } q = 0,015$$

$$(p = 1 - 0,015 = 0,985 \text{ do đó } 2pq = 0,029 \neq 0,03.$$

Như vậy tần số của người dị hợp tử mang gene bệnh PKU ở quần thể người Ireland là khoảng 3%, nghĩa là nguy cơ để một người bố hoặc mẹ mang gene bệnh trong một hôn nhân sẽ là khoảng 3%.

2. Tần số gene và kiểu gene trong trường hợp gene liên kết với NST giới tính X

Đối với gene nằm trên NST giới tính X thì ở người nữ sẽ có 3 kiểu gene khác nhau trong khi đó ở người nam chỉ có 2 kiểu gene do NST Y không mang gene do đó việc áp dụng định luật Hardy - Weinberg có chút thay đổi. Lấy bệnh mù màu làm ví dụ với allele đột biến được kí hiệu là c_b

và allele bình thường được kí hiệu là (+). Tần số của các kiểu gene được trình bày trong bảng 3.

Qua bảng 3 có thể nhận thấy tần số mắc bệnh mù màu ở người nữ thấp hơn nhiều so với người nam, mặc dù tần số allele là như nhau ở cả hai giới. Chưa tới 1% người nữ mắc bệnh mù màu nhưng tần số này ở người nam là 8%. Tần số người nữ dị hợp mang gene bệnh là 15% những người này sẽ có nguy cơ sinh ra con trai mắc bệnh mù màu.

Bảng 3: Tần số gene và kiểu gen trong trường hợp bệnh mù màu, di truyền liên kết với NST giới tính X.

Giới	Kiểu gene	Kiểu hình	Tần số
Nam	X^+	bình thường	$p = 0,92$
	X^{cb}	Mù màu	$q = 0,08$
Nữ	X^+/X^+	Bình thường (đồng hợp)	$p^2 = (0,92)^2 = 0,8464$
	X^+/X^{cb}	Bình thường (dị hợp)	$2pq = 2.(0,92).(0,08) = 0,1472$
		Bình thường (tổng cộng)	$p^2 + 2pq = 0,9936$
	X^{cb}/X^{cb}	Mù màu	$q^2 = (0,08)^2 = 0,0064$

IV. Các yếu tố ảnh hưởng đến trạng thái cân bằng của định luật Hardy - Weiberg

1. Điều kiện nghiệm đúng cho định luật Hardy - Weinberg

- Quần thể phải lớn và sự kết hôn xảy ra một cách ngẫu nhiên
- Tần số của các allele được duy trì không đổi qua các thế hệ nghĩa là: (1) không xảy ra đột biến làm xuất hiện allele mới; (2) mọi cá thể đều có cơ hội như nhau trong kết hôn và sinh sản, nghĩa là tác động của chọn lọc xem như là đồng đều với mọi kiểu gen; (3) không xảy tình trạng nhập cư đáng kể từ những quần thể có vốn gene hoàn toàn khác.

2. Các yếu tố ảnh hưởng

Trong thực tế có rất nhiều yếu tố ảnh hưởng đến tần số của các allele trong quần thể.

- Sự khác biệt về chủng tộc trong một quần thể sẽ ảnh hưởng nhiều đến các quyết định tình trạng hôn nhân. Tình trạng kết hôn đồng huyết và sự lựa chọn bạn đời phù hợp cũng là những yếu tố làm cho sự kết hôn không thể diễn ra một cách ngẫu nhiên.

- Dưới tác động của chọn lọc, đột biến, sự di dân cũng sẽ làm ảnh hưởng ít nhiều đến định luật Hardy-Weinberg mặc dù xảy ra chậm.

- Sự phiêu bạt di truyền (genetic drift): Sự phiêu bạt di truyền dẫn đến sự thay đổi ngẫu nhiên tần số của các allele trong một quần thể nhỏ. Khi xảy ra một đột biến mới trong một quần thể nhỏ, ban đầu có thể chỉ ở một cá thể trong quần thể. Những yếu tố ngẫu nhiên không liên quan đến tình trạng mang gene đột biến có thể làm gia tăng khả năng sống, khả năng sinh sản của người mang đột biến và nhanh chóng làm gia tăng tần số allele đột biến này trong quần thể đó qua các thế hệ. Trong khi đó trong một quần thể lớn thì các yếu tố đó không ảnh hưởng nhiều đến tần số của các allele trong quần thể.

Hiện tượng tương tự cũng xảy ra khi một nhóm cá thể tách ra khỏi quần thể lớn hơn. Khi đó tần số của các allele trong quần thể mới này khác biệt ít nhiều so với tần số allele ở quần thể gốc. Sau khi trải qua vài thế hệ mặc dù kích thước của quần thể cũng chưa lớn nhưng sự khác biệt về tần số allele sẽ trở nên đáng kể. Nếu một trong số các cá thể trong nhóm cá thể tách ra đó mang một allele hiếm thì về sau tần số của allele này sẽ cao hơn trong quần thể mới so với tần số của allele đó trong quần thể cũ.

- Dòng chảy của gene (gene flow): Ngược với hiện tượng phiêu bạt di truyền, hiện tượng dòng chảy của gene là sự phát tán chậm chạp của gene qua biên giới (chữ biên giới được hiểu theo nghĩa rộng không chỉ là biên giới về địa lý mà có thể là sự cản trở trong sinh sản vì lý do chủng tộc, tôn giáo, văn hoá ...), một quá trình ảnh hưởng đến cả một quần thể lớn và làm thay đổi dần tần số của các allele. Các gene của quần thể nhập cư với những đặc trưng riêng của nó dần dần được đưa vào vốn gene của quần thể tiếp nhận và làm thay đổi tần số của các allele trong quần thể này.