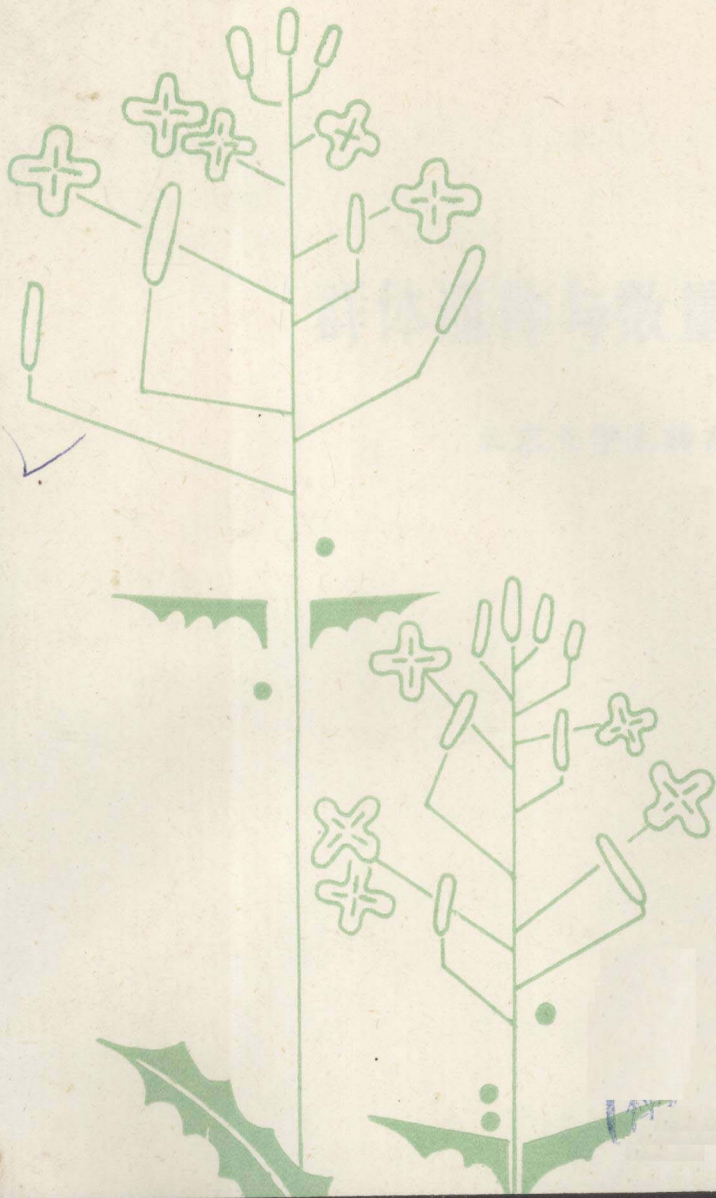




群体遗传与数量遗传学基础

山西省科学技术情报研究所
山西省玉米协作组

一九八〇年八月



中国植物志卷之四十四

被子植物门
双子叶植物纲
蔷薇目
蔷薇科
蔷薇亚科
蔷薇属

群体遗传与数量遗传学基础

北京大学生物系 戴灼华

山西省科学技术情报研究所
山西省玉米育种科研协作组

一九八〇年八月

目 录

第一章 群体遗传学基础	(1)
(一) 群体的基因型频率与基因频率	(1)
(二) Hardy—Weinberg (哈蒂—温伯)平衡法则.....	(3)
(三) 改变群体遗传组成的因素	(13)
(四) 自交群体的遗传组成	(17)
第二章 数量性状的遗传及其主要的遗传参数	(21)
(一) 数量性状改良的基本原理	(21)
(二) 数量性状的遗传力	(31)
(三) 数量性状的遗传参数之二——遗传相关	(51)
第三章 亲本配合力的测定	(57)
(一) 配合力的定义	(57)
(二) 配合力的分析测定方法	(57)
第四章 通径系数的理论及其在育种上的应用	(69)
(一) 通径系数的性质	(69)
(二) 通径系数的运算方法	(76)
(三) 通径分析在遗传育种上的应用	(85)

第一章 群体遗传学基础

(一) 群体的基因型频率与基因频率

所谓群体的遗传组成就是指孟德尔群体中的基因及基因型的种类和频率。

1. 孟德尔群体：享有一个共同的基因库的一群能够相互繁殖的个体叫做孟德尔群体 (Th. Dobzhansky 1955)。在有性繁殖的生物中，一个物种就是一个最大的孟德尔群体，而一个品种也可算是一个孟德尔群体，虽然它们在某些基因的频率上与其他类似的群有所不同。那么一个玉米的自交系可能占有一个小的“基因库”也被称为孟德尔群体。群体遗传学所研究的群体均系孟德尔群体。

基因库 (gene pool)：因为基因存在于孟德尔群体中，所以基因库即是指能够相互交换的那些共同的基因。因而孟德尔群体中基因的总量称为“基因库”。

2. 基因型频率及基因频率：

基因型频率与基因频率是群体遗传组成的基本标志，不同的群体的相同基因的频率往往不同，它们的变化决定了群体的遗传结构和行为的改变。

1) 基因型频率：群体中某特定基因型个体的数目占个体总数目的百分数即是某特定基因型的频率。

为简单起见，假定我们所讨论的是一个二倍体的植物或动物（体染色体）的位点上仅有一对等位基因存在于个体间。因此这个群体里的遗传组成用个体间三种基因型的数目或频率来表示，这些频率称为基因型频率。

例如：玉米籽粒的非糯性与糯性由一对等位基因 W 及 w 所控制。如果在某地区的玉米群体中没有糯性品种，那么非糯性的基因型频率为100%即1，而糯性的基因型频率为0，群体中关于糯与非糯基因型频率的总和则为1。

若已知某玉米的自然群体中糯性籽粒有250株；非糯性的有750株，而其中非糯的纯合体为250株，杂合非糯有500株。于是在下表中求出它三种基因型 WW , Ww , ww ，频率：

基 因 型	WW	Ww	ww	总 数
表 型	非糯	非糯	糯	
频 数 (f)	250	500	250	1000
频 率	0.25	0.50	0.25	1.00

群体的遗传组成不仅考虑基因型频率的大小，而且必须考虑在这一群体内的基因从

一代传递到下一代的问题。在传递过程中，亲代的基因型被拆散，分离成为配子中的基因，由配子所带的基因的结合又形成下一代新的基因型，于是，由群体所载的基因连续不断地代代相传着，可是包含它们的那些基因型却并不连续。因此在具有一对等位基因 W 与 w 的一个群体中究竟包含了多少个 (W) 基因和多少个 (w) 基因？这需要用基因频率来描述。

2) 基因频率：群体中某位点上所存在的某一基因的数目与该位点上等位基因总数目的比率或百分率。

例如： W 是在染色体某位点上的一个等位基因， W 基因的频率就是在这个位点上全部基因中 W 等位基因所占的比率。在任何一个位点上各等位基因的频率相加必须等于 1。

一般而论，每一个二倍体的个体的某位点上含有两个基因，这样就在该玉米群体中总共计数了 $2 \times 1000 = 2000$ 个基因。而每一个 (WW) 个体有 2 个 (W) 基因，每个 (Ww) 个体有一个 (W) 基因和一个 (w) 基因，所以在这个样品群体中共计有 $2 \times 250 + 500 = 1000$ 个 (W) 基因，每一个 (ww) 个体有 2 个 (w) 基因，于是 (w) 基因的数目是 $2 \times 250 + 500 = 1000$

$$\text{因而：} W \text{ 基因频率} = \frac{W \text{ 基因的数目}}{\text{该位点上等位基因总数}} = \frac{1000}{2000} = 0.5 \text{ 或 } 50\%$$

$$w \text{ 基因频率} = \frac{w \text{ 基因的数目}}{\text{该位点上等位基因总数}} = \frac{1000}{2000} = 0.5 \text{ 或 } 50\%$$

总结成下表可得：

基 因 型		WW	Ww	ww	总数	基因 频率
个 体 数		250	500	250	1000	
基 因	W	500	500	0	1000	0.5
	w	0	500	500	1000	0.5

为将上述计算用通式表示，以一对等位基因 A, a 为例：

设有一个 N 株的群体，于是某一位点共计有 $2N$ 个基因。如果群体有：

$n_1 AA + n_2 Aa + n_3 aa$ 个体，因而有

$n_1 + n_2 + n_3 = N$ ，于是此三种基因型频率为：

$$AA: P = \frac{n_1}{N}; \quad Aa: H = \frac{n_2}{N}; \quad aa: Q = \frac{n_3}{N}$$

而 $P + H + Q = 1$ 即： $\frac{n_1}{N} + \frac{n_2}{N} + \frac{n_3}{N} = \frac{N}{N} = 1$ 。

A 的基因频率为

$$p = \frac{2n_1 + n_2}{2N}$$

a的基因频率为:

$$q = \frac{n_2 + 2n_3}{2N}$$

而 $p + q = 1$ 亦即: $\frac{2n_1 + n_2}{2N} + \frac{n_2 + 2n_3}{2N} = \frac{2N}{2N} = 1$.

综上所述列表说明如下:

基 因 型					基 因		
	AA	Aa	aa	总和	A	a	总和
频率	n_1	n_2	n_3	N	$2n_1 + n_2$	$2n_3 + n_2$	$2N$
频率	$\frac{n_1}{N}$	$\frac{n_2}{N}$	$\frac{n_3}{N}$	1	$\frac{2n_1 + n_2}{2N}$	$\frac{2n_3 + n_2}{2N}$	1
代号	P	H	Q		p	q	

3) 基因频率与基因型频率的关系: 基于上述计算可见一对等位基因在群体中遗传时便有:

$$p = \frac{2n_1 + n_2}{2N} = \frac{2n_1}{2N} + \frac{n_2}{2N} = \frac{n_1}{N} + \frac{1}{2} \frac{n_2}{N} = P + \frac{1}{2} H$$

$$q = \frac{2n_3 + n_2}{2N} = \frac{2n_3}{2N} + \frac{n_2}{2N} = \frac{n_3}{N} + \frac{1}{2} \frac{n_2}{N} = Q + \frac{1}{2} H.$$

所以基因频率与基因型频率的关系式:

$$\left. \begin{aligned} p &= P + \frac{1}{2} H \\ q &= Q + \frac{1}{2} H \end{aligned} \right\} \quad 1.1.$$

在动物中由于有伴性基因, 在计算性染色体上的基因频率时, 须将雌、雄个体分别看作两个群体, 在雄性异型 (或称异配性别) 的生物中, 例如很多昆虫, 某些鱼类、某些两栖类、全部哺乳动物等的性别决定都是 XY 型或 XO 型, 所以这类生物的雄性群体中只有两种基因型 $X^A Y$ 及 $X^a Y$ 。因而基因频率就等于基因型频率。而它们的雌性群体的基因型频率与基因频率的关系与上述常染色体基因一样可以直接利用上述关系式来求。雌性异型 (或称雌性异配性别) 的生物, 如鸟类、家禽等动物的雌性都是 XY 或 ZW 型, 而雄性却是 XX 或 ZZ 型, 这样计算伴性基因的频率时的情况正好相反。

(二) Hardy—Weinberg (哈蒂—温伯) 平衡法则

英国数学家Hardy和德国医生Weinberg经过各自独立的研究同时于1908年发现了

基因频率与基因型频率的重要规律，现在公称为Hardy—Weinberg平衡法则，它是群体遗传学与数量遗传学的一个重要理论基石。

这个法则说：在一个大的孟德尔群体中，个体进行随机交配，同时没有突变、迁移和选择发生时，下一代基因型的频率将和前一代一样，其基因频率亦是世代相传不发生变化。于是这个群体就被称为处于平衡状态中。

一、法则的证明：这个法则的证明包括三个步骤：

1. 从亲代的遗传组成到它们所产生的配子；
2. 从配子的随机结合（或个体的随机交配）到所产生的子代（合子）中的基因型频率；
3. 从合子的基因型到在子代中的基因频率。

以一对等位基因 A , a 为例，则群体中可能的基因型只有 AA , Aa 及 aa 。

1. 设亲代（0 世代）具有以下遗传组成，即基因频率和其可能的基因型频率如下：

A	a	AA	Aa	aa
p_0	q_0	P_0	H_0	Q_0
$p_0 = P_0 + \frac{1}{2}H_0$		$q_0 = Q_0 + \frac{1}{2}H_0$		

亲代群体所产生的配子有两种，一种带有基因 (A)，另一种带有基因 (a)，这些配子类型的频率是与产生它们的亲代中的基因 (A) 及 (a) 的频率相等的，即：

$$A \text{ 型配子频率} = A \text{ 基因频率} = p_0$$

$$a \text{ 型配子频率} = a \text{ 基因频率} = q_0$$

因为 AA 个体只产生 A 型配子， Aa 个体产生等量的 A 和 a 型配子， aa 个体只产生 a 型配子，所以群体中所产生的 A 型配子的频率可由 $P_0 + \frac{1}{2}H_0$ 式，及 a 型配子的频率可由 $Q_0 + \frac{1}{2}H_0$ 式求出，而这就是 A 基因及 a 基因的频率。

2. 个体间作随机交配（相当于它们所产生的配子间的随机结合）到它们所产生的子代（合子）中的基因型频率。

所谓随机交配就是说任何个体有同样的机会和任何其他（异性）个体交配。也就是说，各种个体交配的频率完全决定于其频率的大小，而不受任何其他因素的影响。同理，任何配子有同样的机会和任何其他（异性）配子结合的过程称为配子的随机结合。

我们可以设想，亲代群体中有大数量的配子，配子成对地随机结合而产生合子，在这些合子中的基因型频率是那些结合起来的配子频率的乘积，如下表所示：

亲代		雄体的配子及其频率	
		A (p_0)	a (q_0)
雌体的配子及其频率	$A(p_0)$	AA (p_0^2)	Aa (p_0q_0)
	$a(q_0)$	Aa (p_0q_0)	aa (q_0^2)

由于我们不需要对 A 卵和 a 精子的结合与 A 精子和 a 卵的结合加以区别, 所以合子(子代)的基因型频率为:

$$\begin{array}{ccc} AA & Aa & aa \\ p_0^2 & 2p_0q_0 & q_0^2 \end{array}$$

$$\text{设: } P_1 = p_0^2; \quad H_1 = p_0q_0 + p_0q_0 = 2p_0q_0$$

$$Q_1 = q_0^2$$

由此可见, 子代的基因型频率只决定于亲代的基因频率, 而与亲代的基因型频率无关, 只要亲代是随机交配的。显然, 子代各种基因型频率为 $(p_0A + q_0a)^2 = 1$ 的展开。

3. 由子代合子的基因型频率计算其基因频率:

我们可以应用公式(1.1)求得:

$$\begin{aligned} p_1(A \text{ 的基因频率}) &= P_1 + \frac{1}{2} H_1 = p_0^2 + \frac{1}{2} \times 2p_0q_0 \\ &= p_0^2 + p_0q_0 = p_0(p_0 + q_0) \\ &= p_0 \end{aligned}$$

$$\begin{aligned} q_1(a \text{ 的基因频率}) &= Q_1 + \frac{1}{2} H_1 = q_0^2 + \frac{1}{2} (2p_0q_0) = q_0^2 + p_0q_0 \\ &= q_0(q_0 + p_0) = q_0 \end{aligned}$$

由此可见合子(子代)的基因频率与亲代的基因频率相等。即:

$$p_1 = p_0$$

$$q_1 = q_0$$

同样可以证明在随机交配的条件下, 只要没有其他因素的干扰由亲代(0世代)所产生的子二代, 以至 n 代群体中的基因频率亦有:

$$\begin{array}{cc} p_2 = p_0 & q_2 = q_0 \\ \vdots & \vdots \\ p_n = p_0 & q_n = q_0 \end{array}$$

基因频率代代不变。

如果使亲代个体行随机交配, 则各种交配的频率完全由其基因型频率所决定, 它们的频率为:

母 本 父 本	AA	Aa	aa
	P	H	Q
AA P	P ²	PH	PQ
Aa H	PH	H ²	HQ
aa Q	PQ	QH	Q ²

显然, 各种交配的频率为 $(P + H + Q)^2 = 1$ 的展开, 所以九种交配类型的频率 分别为下列六项

$$(P + H + Q)^2 = P^2 + H^2 + Q^2 + 2PH + 2PQ + 2HQ$$

这是因为我们不必要去区别亲体的性别。即正交与反交的频率相同, 可将同样基因型的交配合并在一起, 于是我们就有表1.1。

表1.1 交配类型的频率及其基因型的平衡

交 配		子 代 的 基 因 型 和 频 率		
类 型	频 率	AA	Aa	aa
AA × AA	P ²	P ²	—	—
AA × Aa	2PH	PH	PH	—
AA × aa	2PQ	—	2PQ	—
Aa × Aa	H ²	$\frac{1}{4}H^2$	$\frac{1}{2}H^2$	$\frac{1}{4}H^2$
Aa × aa	2HQ	—	HQ	HQ
aa × aa	Q ²	—	—	Q ²
总和 =		$(P + \frac{1}{2}H)^2$ p ²	$2(P + \frac{1}{2}H)(Q + \frac{1}{2}H)$ 2pq	$(Q + \frac{1}{2}H)^2$ q ² (依式 1.1)

可由上表看出: 子代的基因型已经达到平衡, 只要没有其他因素的影响, 以后一代一代的随机交配下去, 这种平衡状态永远保持不变。

现在我们举一个实际例子来加以说明。已知玉米籽粒的黄色和白色是由一对等位基因所控制, 其中黄色(Y)为显性, 白色(y)为隐性。若在某自然群体中有 $(YY)P_0 = 0.4$; $(Yy)H_0 = 0.4$; $(yy)Q_0 = 0.2$ 让其自由授粉, 后代的基因频率及基因型频率如何?

1) 亲代群体的基因频率以及由它们所产生的配子频率;

根据式(1.1)亲代基因Y及y的频率为:

$$p_0(Y) = P_0 + \frac{1}{2}H_0 = 0.4 + \frac{1}{2}(0.4) = 0.6$$

$$q_0(y) = Q_0 + \frac{1}{2}H_0 = 0.2 + \frac{1}{2}(0.4) = 0.4.$$

亲代群体产生的配子只有Y和y两种,配子的频率等于亲代基因频率,即Y类配子频率为0.6,y类配子的频率为0.4,于是我们有:

2)亲代配子的随机结合所产生的子₁代基因型频率及基因频率:

亲代♀ \ 亲代♂	子 ₁ 代	Y	y
		(0.6)	(0.4)
Y(0.6)		YY(0.36)	Yy(0.24)
y(0.4)		Yy(0.24)	yy(0.16)

子₁代基因型频率:

$$P_1 = YY \text{ 基因型频率} = (0.6)^2 = 0.36$$

$$H_1 = Yy \text{ 基因型频率} = (0.4)(0.6) + (0.4)(0.6) = 0.48$$

$$Q_1 = yy \text{ 基因型频率} = (0.4)^2 = 0.16$$

子₁代的基因频率根据式1.1可求:

$$p_1 = Y \text{ 基因频率} = P_1 + \frac{1}{2}H_1 = 0.36 + \frac{1}{2}(0.48) = 0.6$$

$$q_1 = y \text{ 基因频率} = Q_1 + \frac{1}{2}H_1 = 0.16 + \frac{1}{2}(0.48) = 0.4$$

可见:子₁代的基因频率与亲代相同。

$$p_1 = p_0 = 0.6 \quad q_1 = q_0 = 0.4.$$

3)子一代个体随机交配,相当于它们所产生的配子随机结合,而产生的孙代(即子₂代)群体:由于子₁代的配子频率仍是Y=0.6、y=0.4。所以又有:

子 ₁ 代♀ \ 子 ₁ 代♂	孙代(子 ₂ 代)	Y _{P₁} (0.6)	y _{q₁} (0.4)
Y _{P₁} (0.6)		YY _{P₁} ² (0.36)	Yy _{P₁q₁} (0.24)
y _{q₁} (0.4)		Yy _{P₁q₁} (0.24)	yy _{q₁} ² (0.16)

孙代(子₂代)基因型频率:

$$P_2 YY \text{基因型频率} = p_1^2 = (0.6)^2 = 0.36$$

$$H_2 Yy \text{基因型频率} = 2p_1q_1 = 2(0.24) = 0.48$$

$$Q_2 yy \text{基因型频率} = q_1^2 = (0.4)^2 = 0.16.$$

在此, 亲代的基因型频率 $P_0 = 0.4$; $H_0 = 0.4$; $Q_0 = 0.2$. 经过一代的随机交配, 子₁代的基因型频率则被确定为:

$$P_1 = 0.36. \quad H_1 = 0.48. \quad Q_1 = 0.16.$$

而随机交配后的孙代 (子₂代) 群体的基因型频率则与上一代相等, 并且以后各代保持不变:

$$P_2 = 0.36 = P_1; \quad H_2 = 0.48 = H_1 \quad Q_2 = 0.16 = Q_1$$

于是我们可以推知随机交配至 n 代: $P_1 = P_2 = P_3 \cdots = P_n = p^2 = 0.36$

$$H_1 = H_2 = H_3 \cdots = H_n = 2p_1q_1 = 0.48$$

$$Q_1 = Q_2 = Q_3 \cdots = Q_n = q^2 = 0.16$$

而孙代 (子₂代) 的基因频率:

$$p_2 = Y \text{ 基因频率} = P_2 + \frac{1}{2}H_2 = 0.36 + \frac{1}{2}(0.48) = 0.6$$

$$q_2 = y \text{ 基因频率} = Q_2 + \frac{1}{2}H_2 = 0.16 + \frac{1}{2}(0.48) = 0.4.$$

可知基因频率保持不变:

$$p_0 = p_1 = p_2 = 0.6; \quad q_0 = q_1 = q_2 = 0.4.$$

如果让其自由授粉以至 n 代, 则仍然保持:

$$p_0 = p_1 = p_2 = \cdots = p_n = 0.6$$

$$q_0 = q_1 = q_2 = \cdots = q_n = 0.4.$$

二、Hardy—Weinberg平衡法则的要点:

综上所述我们将平衡法则的要点概括如下:

1. 在随机交配的大群体中, 若没有其他因素的影响 (如基因的突变、迁移和选择), 基因频率代代保持不变。

2. 在任何一个大群体中, 如果只有一对等位基因在遗传时, 无论亲代的基因型频率如何, 只要经过一代的随机交配, 其常染色体基因型频率的平衡就被确立下来。若没有其他因素的影响, 继续随机交配, 其基因型频率则从此代代保持不变。

3. 由亲本个体间的随机交配所产生的子代基因型频率可根据亲代基因频率按下列二项式展开式来计算:

$$(pA + qa)^2 = p^2(AA) + 2pq(Aa) + q^2(aa) \quad 1.2.$$

这就是处于Hardy—Weinberg平衡状态下的群体, 在任何一代中的基因频率和基因型频率之间的关系。

在动物群体中, 当常染色体某位点有三个等位基因 (称为复等位基因), 其频率各为 p, q, r 时, 子代的基因型频率将为 $(p + q + r)^2 = 1$ 的展开。与一对等位基因的情况相同, 平衡只需要一代的随机交配就可以达到。

如果我们考虑两对等位基因, 而基因频率为 $pA:qa$ 和 $rB:tb$ 时, 子代的基因型频率

可以由 $[(p+q)^2 \cdot (r+t)^2]$ 的展开式求得。基因型频率的平衡不是一代随机交配就能达到，但速度也相当快。

若所涉及的基因的对数越多，平衡的速度越慢，如果两对基因有连锁，则比非连锁的基因达到平衡的速度更慢，连锁强度越大，接近平衡的速度越小。

三、Hardy—Weinberg平衡法则的生物学例证：

最明显的例子是人类的 MN 血型。据研究，人的 MN 血型是由一对常染色体基因 L^M 和 L^N 决定的。 L^M 基因决定 M 抗原的存在，它们的基因型为 $L^M L^M$ ，表现为 M 血型，N 型个体的红血细胞有 N 抗原，是由 L^N 基因决定的，它们的基因型为 $L^N L^N$ ， L^M 与 L^N 基因决定 M，N 抗原的存在，表现为 MN 血型，其基因型则为 $L^M L^N$ ，这一对等位基因的两个成员在杂合体中都显示出来，具有并显性(codominance)现象，所以基因型与表现型一致，易于从表现型分辨基因型。这是选用这个性状作为例证的原因之一。其次是因为这个性状一般在婚配时是不加选择的，而且与适应性基本无关，也就是不受自然选择的影响，因此就这个性状来说，一般都是随机婚配的。第三个原因是由于人类的群体（民族）有的是很大的，因此，有可能在相当大的群体中进行调查。

1977年在我国某大城市调查了1788人的 MN 血型，结果如下表

表1.2 某城市1788人的 MN 血型的调查

血 型	基 因 型	观 察 人 数	L^M 基因数	L^N 基因数	合 计
M	$L^M L^M$	397	794	—	
MN	$L^M L^N$	861	861	861	
N	$L^N L^N$	530	—	1060	
基 因 数			1655	1921	3576
基因频率			$p = 0.4628$	$q = 0.5372$	1.0000

$$L^M \text{ 基因频率: } p = \frac{397 \times 2 + 861}{3576} = 0.4628$$

$$L^N \text{ 基因频率: } q = \frac{861 + 530 \times 2}{3576} = 0.5372$$

由于就这个性状来说，基本上是随机交配的，而且群体较大，所以根据平衡法则应该处于平衡状态， $P = p^2$ ， $H = 2pq$ ， $Q = q^2$ 。

$$M \text{ 型频率} = P = p^2 = (0.4628)^2 = 0.2142$$

$$MN \text{ 型频率} = H = 2pq = 2 \times 0.4628 \times 0.5372 = 0.4972$$

$$N \text{ 型频率} = Q = q^2 (0.5372)^2 = 0.2886$$

这便是一个平衡群体中，预计的三种基因型频率的值，我们将它们分别乘以总人数（N=1788）则可得到其予期人数，再与实得人数比较，进行测验，计算结果如下：（表1.3）

表1.3 MN血型资料的理论频数的 χ^2 测验

表 现 型	M	MN	N	合 计
基 因 型	$L^M L^M$	$L^M L^N$	$L^N L^N$	
予期基因型频率	$(0.4628)^2$	$2 \times 0.4628 \times 0.5327$	$(0.5327)^2$	1
予期频数(C)	382.96	889.05	515.99	1788
实得频数(O)	397.00	861.00	530.00	1788
$\frac{(O-C)^2}{C}$	0.51	0.88	0.38	$\chi^2 = 1.77$

$$\chi^2 = \sum \frac{(O-C)^2}{C} = 1.77.$$

当自由度 = 3 - 2 = 1时（由于在计算予期频数时又占去了一个自由度）， $\chi^2_{0.05} = 3.841$ ，所以 $P > 0.05$ ，表明这三种血型的基因型频率符合 Hardy-Weinberg 平衡法则，此群体的基因频率及基因型频率确认处于平衡状态之中。

这样的实例，在动植物界（包括人类）群体中还有很多。因此平衡法则具有普遍意义。

四、基因频率的计算：

1. 无显性或显性不完全时：这是最简单的情况，因为这时由表现型直接可以识别基因型，上述人类的 MN 血型亦是如此。我们只需要统计表现型的频率就可以得到基因型频率，也就可以按平衡法则的公式（公式1.1）算出基因频率。

例如安达鲁西鸡有三种毛色：黑色、兰色和白花。据研究它们是由一对基因 B 及 b 控制的，B 对 b 为部分显性。黑色的基因型为 BB，兰色的基因型为 Bb，白花的基因型为 bb。若调查大群安达鲁西鸡的结果，黑色鸡占49%，兰色鸡占42%，白花鸡占 9 %。我们立刻就知道：BB 的基因型频率为0.49，Bb 的基因型频率为0.42，bb 的基因型频率为0.09，于是可以算出：

$$B \text{ 的基因频率} = 0.49 + \frac{1}{2} 0.42 = 0.7 \dots (p)$$

$$b \text{ 的基因频率} = 0.09 + \frac{1}{2} 0.42 = 0.3 \dots (q)$$

或 直接由 $1 - p = q$ $q = 1 - 0.7 = 0.3$ 。算得 b 的基因频率。

2. 完全显性时：由于一对等位基因之间显性完全，其可能的基因型有三种，而表型

却只有两种，显性纯合子和杂合子的表型相同，不能直接识别。我们只能得到隐性纯合子的比率，如果是一个随机交配的大群体，根据平衡法则，它应该处于基因频率平衡状态，它的隐性纯合子的基因型频率应该等于 q^2 ，即 $Q = q^2$ ，所以

$$q = \sqrt{Q}, \text{ 而 } p = 1 - q$$

例如：某地区某玉米品种的果穗上约有1/10的子粒是紫红糊粉层，由 R 显性基因决定的。紫红糊粉层的表型必有两种基因型，我们无法直接计算出 R 基因频率。但是无色糊粉层的子粒是隐性纯合体 (rr) 占9/10，即 $q^2 = 0.9$ 。

所以 $q = \sqrt{q^2} = \sqrt{0.9} = 0.9487$ 这就是 r 的基因频率， R 基因频率为 $1 - q = 1 - 0.9487 = 0.0513$ 。由基因频率还可以推算出各种基因型的频率：

$$\left. \begin{aligned} RR \text{ 基因型频率} &= 0.0513 \times 0.0513 = 0.0026 \\ Rr \text{ 基因型频率} &= 0.0513 \times 0.9487 = 0.0487 \\ rR \text{ 基因型频率} &= 0.9487 \times 0.0513 = 0.0487 \\ rr \text{ 基因型频率} &= 0.9487 \times 0.9487 = 0.900 \end{aligned} \right\} = 0.1 \text{ (紫红糊粉层)}$$

(无色糊粉层)

又如：人的遗传性疾病，如白化病、先天性聋哑、镰形细胞贫血是隐性的、常染色体遗传病。每当我们知道了这些病在人群中的发病率后，便可推算出它们的杂合体携带者的频率。

据统计，在英国人群中，白化病（先天性皮肤毛发色素缺乏的疾病，全身白化症的患者全身黑色素细胞均缺乏酪氨酸酶，不能使酪氨酸变成黑色素。不含有黑色素的皮肤呈白色或淡红色，毛发银白或淡黄色、瞳孔呈浅红色，视力差）的发病率约为1/20,000。据人类遗传学研究，此病由隐性 c 基因所决定，表现为白化者必为 cc 基因型，即 $cc = q^2 = 1/20,000$

$$\text{所以 } c \text{ 基因频率} = \sqrt{q^2} = \sqrt{1/20,000} = 1/141$$

$$\text{因而 } C \text{ 基因频率} = 1 - q = 1 - 1/141 = 140/141$$

危害较大的是隐性白化基因 (c) 的携带者，它们的表型为正常但基因型却是杂合体 (Cc)，那么此人群中杂合子携带者的频率究竟有多大？我们根据平衡法则便可推算得知：

$$Cc \text{ 的基因型频率即为 } 2pq = 2 \times 1/140 \times 140/141 = 1/70$$

这个推算告诉我们：尽管白化病的发病率很低，只有1/20,000，但是群体中的白化隐性基因的携带者却高达1/70，即为发病率的280倍，这不能不引起遗传学工作者和医学工作者的足够注意。

3. 复等位基因：人的 A、B、O 血型决定于三个等位基因系列（称为复等位基因），其中 I^A 对 i 呈显性， I^B 对 i 也是显性，而 I^A 与 I^B 间无显性。基因型 $I^A I^A$ 和 $I^A i$ 表现为 A 型血， $I^B I^B$ 和 $I^B i$ 表现为 B 型血， $I^A I^B$ 表现为 AB 血型， ii 则为 O 型。

设 I^A 的基因频率为 p ， I^B 的基因频率为 q ， i 的基因频率为 r ，A 型人的比率为 A ，O 型人的比率为 O 。

$$\text{由于 } O = r^2 \quad \therefore r = \sqrt{O}$$

$$\text{而 } A + O = p^2 + 2pr + r^2 = (p + r)^2$$

$$\therefore p + q + r = 1 \quad \therefore p + r = 1 - q$$

$$(p+r)^2 = (1-q)^2$$

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

于是: $1-q = \sqrt{A+O}$

$$\therefore q = 1 - \sqrt{A+O}$$

$$\therefore p = 1 - q - r$$

例如: 在昆明调查6000个人的血型: O型1846人, A型1920人, B型1627人, AB型为607人, 求各基因 (I^A , I^B , i) 的频率。

$$\text{O型频率} = \frac{1846}{6000} = 0.30766$$

$$\text{A型频率} = \frac{1920}{6000} = 0.32000$$

$$\text{B型频率} = \frac{1627}{6000} = 0.27116$$

$$\text{AB型频率} = \frac{607}{6000} = 0.10116$$

$$\therefore r = \sqrt{O} = \sqrt{0.30766} = 0.5547$$

$$q = 1 - \sqrt{0.32000 + 0.30766} = 1 - \sqrt{0.62766} = 0.2077$$

$$p = 1 - 0.2077 - 0.5547 = 0.2376$$

4. 伴性基因: 一对伴性基因在雄性异型(XY)生物的雄性群体中基因型只有两种: $X^A Y$ 和 $X^a Y$, 表现型也是两种, 因此由表现型就能识别基因型, 基因型频率也就是基因频率。例如人的红绿色盲为伴性隐性遗传病。男人中色盲者所占的比率就是色盲基因在整个群体中的频率。雌性异型的生物的情况正相反(雌性为XW), 例如鸡的芦花斑纹是显性伴性性状, 黑色为其相对性状。在一个只有芦花与黑色两种毛色的随机交配大鸡群中, 雌鸡中芦花鸡所占的比率就是芦花基因在这个鸡群中的频率, 黑鸡的比率就是黑色基因的频率。

五、Hardy—Weinberg平衡法则的意义:

这个法则所揭示的基因频率及基因型频率的规律在研究动物、植物及人类群体的遗传结构和行为时都具有极其重要的理论和实践意义。因而平衡法则已成为群体遗传学和数量遗传学的重要理论基石之一。

平衡法则揭示了基因频率和基因型频率在一定条件下的相对稳定性, 因而群体的遗传特性才能保持相对的稳定。基因和基因型的差异导致生物体的遗传变异, 而基因频率和基因型频率的差异则必然造成群体的遗传变异。稳定性是相对的。如果群体失去了平衡的条件, 亦即在有基因的突变、迁移和选择等因素的影响下改变了群体的基因频率, 从而改变了群体的遗传结构, 这就是目前动植物育种工作的最主要的手段之一。若上述因素不再继续起作用, 换句话说, 只要我们稍稍放松我们的育种工作, 让其自由授粉(随机交配), 群体中的基因频率和其可能的基因型的频率又将自动地保持新的平衡。

平衡法则揭示了在一个随机交配的大群体中, 基因频率和基因型频率的一般关系,

特别是隐性纯合体的频率与隐性基因的关系。这就为认识群体的性质，分析研究基因的动态和行为，进行各种隐性遗传疾病的遗传咨询和防治提供重要的理论依据。

(三) 改变群体遗传组成的因素

平衡条件能够得到保证的群体被认为是一种理想群体。由于下列因素的影响，使基因频率和基因型频率发生改变，从而改变群体的遗传组成。

一、突变：(mutation)

基因突变对于群体遗传组成的改变有两个重要的作用：1. 突变提供选择的材料，如果突变的方向与选择的方向一致，基因频率改变的速度就更快。2. 突变本身就改变了基因频率。各种基因都有一定的自然突变率，虽然自然突变率一般很低，大约在每代 10^{-4} — 10^{-8} 之间，而回复突变率只有正向突变率的1/10，虽然如此，但仍是一个经常起作用的因素。

设 a 基因的频率为 q ，A 基因的频率为 $p = 1 - q$ ，A 突变为 a 的突变率为 u ，a 回复突变为 A 的突变率为 v ，则每一代中共有 $(1 - q)u$ 的 A 突变为 a，有 qv 的 a 突变为 A。若 $(1 - q)u > qv$ ，则 a 的频率增加，若长时期由 A → a，没有其他因素的阻碍，最后群体中的 A 将完全为 a 替代，这就是突变所产生的突变压 (mutation pressure)。若 $(1 - q)u < qv$ 则 A 的频率逐渐增加。若 $(1 - q)u = qv$ ，则基因频率 q 保持不变，即处于平衡状态。所以在平衡时：

$$(1 - q)u = qv$$

$$u - qu = qv$$

$$\therefore q = \frac{u}{u + v}$$

同理可有：

$$p = \frac{v}{u + v}$$

1.3.

因而，基因频率完全由突变率 u 和 v 决定。

基因频率未达到平衡时，则群体内 A 的频率的 Δ 改变 (p)，将是基因 a 的突变率 qv 减去基因 A 的突变率 pu ，即： $\Delta p = qv - pu$ 。

例如：如由 A → a 的突变率是每代一百万个配子中有一个突变 ($u = 0.000001$)，而由 a → A 的突变率是 $v = 0.0000005$ 。则按式 (1.3) 可求出

$$p = \frac{0.0000005}{0.000001 + 0.0000005} = 0.33$$

这说明群体中 33% 的染色体具有基因 A，而 67% 具有基因 a。当 $\Delta p = 0$ 时，即表示 $qv = pu$ 亦即群体达到了平衡。所以就一对等位基因而言，其正反突变率相等 (即 $u = v$) 时， p 和 q 的平衡值是 0.5。

如果由 A → a 的突变率比 a → A 的突变率大一倍，即 $u = 2v$ 时，则 a 的基因频率逐代增加，直到 $q = \frac{u}{u + v} = \frac{2v}{v + 2v} = \frac{2v}{3v} = \frac{2}{3} = 0.67$ 时，这个群体的 A 和 a 的基因频率