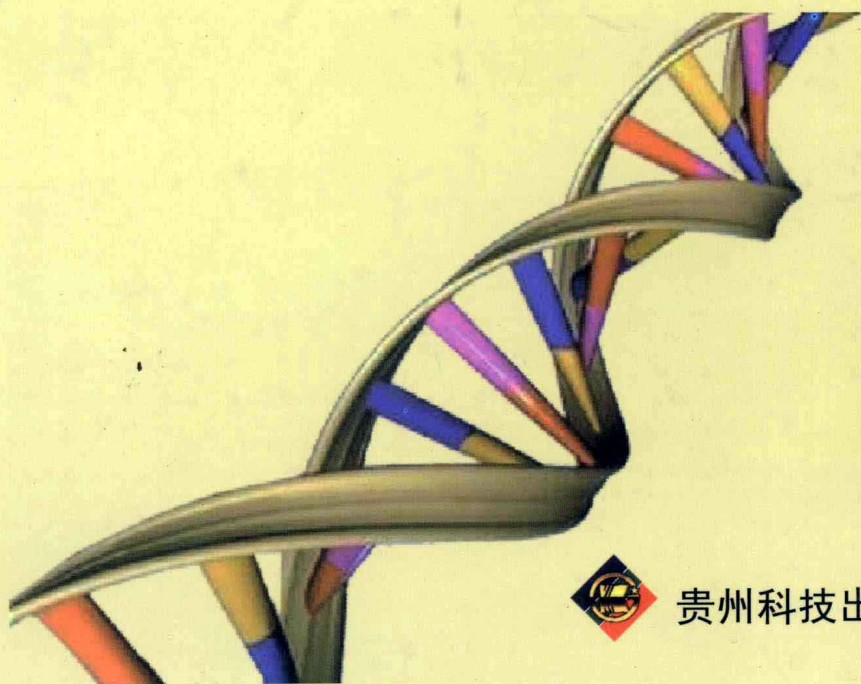


群体与数量遗传学

陈泽辉 编著

$$y = ax^2 + bx + c$$



贵州科技出版社

二次函数

二次函数 二次函数

$$y = ax^2 + bx + c$$
$$y = ax^2 + bx + c$$



二次函数

群体与数量遗传学

陈泽辉 编著

贵州科技出版社

· 贵阳 ·

图书在版编目(CIP)数据

群体与数量遗传学/陈泽辉编著. —贵阳:贵州科技出版社, 2009. 8

ISBN 978 - 7 - 80662 - 793 - 8

I. 群… II. 陈… III. ①群体遗传学②数量遗传学
IV. Q347 Q348

中国版本图书馆 CIP 数据核字(2009)第 133795 号

群体与数量遗传学

QUNTI YU SHULIANG YICHUANXUE

出 版
发 行 贵州科技出版社

出版人 周维莉

地 址 贵阳市中华北路 289 号 邮政编码 550004

经 销 贵州省新华书店

印 刷 贵州新华印刷厂

开 本 787mm × 1092mm 1/16

印 张 22.75 印张

字 数 300 千字

版 次 2009 年 8 月第 1 版

印 次 2009 年 8 月第 1 次

印 数 1 ~ 1000 册

书 号 ISBN 978 - 7 - 80662 - 793 - 8/Q · 030

定 价 39.00 元

前 言

遗传学主要是研究基因的传递和表达,而群体遗传学是以个体集中的某一群体为对象,研究群体中遗传组成及其变化。群体遗传学的研究历程可以追溯到 1859 年达尔文的自然选择学说、孟德尔遗传规律的发现和上世纪初的纯系学说及突变学说,随后通过对果蝇的研究使群体遗传学得到较快的发展。虽然群体遗传学的研究内容丰富,但其研究核心是群体中基因频率的变化。

随着分子遗传学的发展,对群体中的遗传变异研究,逐渐进行到蛋白质和 DNA 水平,并得出了群体中有惊人程度的遗传变异现象,发现了遗传多态现象及其保持机制和分子进化机制,提出了基于遗传漂变理论的中性基因的分子进化学说,从而推动了群体遗传学的发展。

传统的数量遗传学是建立在群体中基因的效应和基因频率基础之上的基因型方差的研究;是建立在一定假设前提下的遗传模型、遗传方差、遗传协方差、遗传力、主基因多基因分析、配合力选择响应和杂种优势等的研究,为育种应用提供了理论基础和方法手段,并成功应用于动植物的育种实践。

分子标记等生物技术的发展,对传统的数量遗传研究形成很大的冲击,同时也使数量性状的遗传研究进入了现代数量遗传学研究时代。现代数量遗传学继承了传统数量遗传学的核心内容,并使其得到新的发展,使数量性状的表现观测与分子标记技术相结合,并能对数量性状进行基因定位、克隆和转育。这些新的发展,将会大幅度地提高动植物育种效益。

笔者 1991 年在国际小麦玉米改良中心(CIMMYT)学习时,以

Hallauer 的“Quantative genetics in maize breeding”和 Falconer 的“Introduction to Quantitative Genetics”为教材学习群体遗传学和数量遗传学,这对我后来的工作和学习产生了重要影响。1994 年 CIM-MYT 为我供奖学金在菲律宾大学和 IRRI 攻读植物育种专业硕士和博士学位,期间主要在 Carpena 教授的严格要求和精心指导下学习群体遗传学和数量遗传学。通过学习,认为这两门学科是紧密联系的,并且群体遗传学是学好数量遗传学的基础。1998 年回国后主要从事玉米育种,同时在贵州大学为研究生讲授“数量遗传学”。在课程内容安排上,先介绍部分群体遗传学知识,然后讲授数量遗传学和 QTL 定位,使学生更容易理解所学内容。

根据多年育种和教学工作经历及学生的要求,部分遗传育种工作者的建议,我编写出版了这本《群体与数量遗传学》。它既可作为研究生教材,也可为遗传育种工作者参考。本书在编排上分为 16 章,其中第 1~4 章介绍群体遗传学,第 5 章简要介绍有关的数学和统计学知识,第 6~13 章重点介绍数量遗传学的基础知识和应用,第 14~16 章简要介绍数量遗传学的新发展,包括主基因-多基因分析、QTL 定位和遗传多样性分析。

本书在撰写过程中大量引用了前人著作中的精彩内容,也引用了不少经典例题,在此对这些著作的作者表示衷心的感谢。由于作者水平所限,不当之处在所难免,诚望读者批评指正,以便今后修正补充。

作 者

2009 年 7 月 18 日

目 录

第一章	群体的遗传组成与平衡群体	(1)
第一节	基因型频率和基因频率	(2)
第二节	Hardy - Weinberg 定律	(7)
第三节	复等位基因情况下群体的遗传平衡	(13)
第四节	性连锁下群体的遗传平衡	(14)
第五节	多个位点下群体的遗传平衡	(16)
第六节	多倍体的遗传平衡	(22)
第七节	同型交配	(25)
第二章	群体中基因频率的变化	(28)
第一节	迁移	(28)
第二节	突变	(31)
第三节	选择	(36)
第四节	在突变与选择之间的平衡	(48)
第五节	中性学说	(50)
第三章	理想小群体	(53)
第一节	理想化群体	(53)
第二节	近亲交配	(55)
第三节	近交系数	(58)
第四节	理想小群体内的遗传效应	(68)
第四章	非理想小群体	(80)
第一节	有效群体含量及其应用	(80)
第二节	迁移、突变、选择	(93)
第三节	在自然群体里的随机漂移	(97)

第五章 统计学基础	(100)
第一节 定义与概念	(100)
第二节 平均数和变异数	(101)
第三节 直线回归和相关	(102)
第四节 线性函数	(104)
第五节 数学期望	(105)
第六节 多元线性回归	(106)
第七节 极大似然估计	(111)
第八节 矩阵及应用	(113)
第六章 数量性状的群体平均与遗传效应	(118)
第一节 数量性状及其遗传特点	(118)
第二节 表现型值与基因型值	(121)
第三节 随机交配平衡群体的均值和遗传效应	(122)
第四节 双亲后代群体的世代均值与遗传效应	(138)
第五节 环境偏差和基因型与环境的互作	(143)
第七章 群体方差和协方差	(146)
第一节 方差和协方差的概念及计算	(146)
第二节 随机交配平衡群体的方差	(148)
第三节 亲属间相似与遗传协方差	(158)
第四节 双亲后代群体的方差和协方差	(169)
第八章 遗传交配设计与遗传方差成分估计	(175)
第一节 单因素遗传方差设计	(175)
第二节 多世代方差分析	(178)
第三节 NCI 设计	(179)
第四节 NCII 设计	(182)
第五节 NCIII 设计	(188)
第六节 TTC 设计	(193)
第七节 双列杂交设计	(197)

第九章 遗传力分析	(215)
第一节 遗传力概念和定义	(215)
第二节 遗传力估计的原理	(217)
第三节 遗传力估算方法及实例	(219)
第四节 遗传力在育种中的应用	(226)
第五节 重复力	(230)
第十章 遗传模型检验与基因效应分析	(231)
第一节 尺度效应	(231)
第二节 单一尺度检验	(233)
第三节 联合尺度检验	(238)
第四节 世代平均值分析	(245)
第五节 阈性状	(251)
第十一章 选择及其响应	(254)
第一节 选择及其效应	(254)
第二节 选择差和选择强度	(256)
第三节 选择响应	(261)
第四节 选择响应的度量	(264)
第五节 选择方法比较	(265)
第六节 群体的轮回选择及选择响应	(268)
第七节 选择对群体方差的影响	(271)
第十二章 遗传相关与选择	(273)
第一节 遗传相关分析	(273)
第二节 相关响应与间接选择	(283)
第三节 选择指数	(286)
第十三章 交配效应与配合力分析	(290)
第一节 近交衰退	(290)
第二节 杂种优势	(292)
第三节 配合力分析	(299)

第十四章	数量性状的主基因-多基因分析	(312)
第一节	主基因及其鉴别	(312)
第二节	主基因和多基因混合遗传模型	(314)
第三节	遗传方差的分解	(315)
第四节	混合遗传模型的分离分析方法	(317)
第十五章	数量性状基因定位及效应分析	(323)
第一节	概述	(323)
第二节	QTL 定位的基本方法	(330)
第三节	QTL 的精细定位	(338)
第十六章	数量性状的遗传多样性分析	(341)
第一节	遗传距离计算的主成分分析	(341)
第二节	利用分子标记数据测定遗传距离	(346)
第三节	系统聚类	(347)
第四节	数量性状遗传多样性的评价指标	(349)
第五节	遗传多样性分析实例	(350)
参考文献		(355)

第一章 群体的遗传组成与平衡群体

在遗传学上把一群可以相互交配的个体所组成的集团称为群体。群体的遗传结构、基因频率及基因型频率的变化,就是群体遗传学的研究范围。

为了了解各种因素对群体遗传结构的影响,首先要知道群体在没有外部因素影响时的遗传特性。当然这样的群体在自然界很难找到,所以我们必须定义一个理论上的理想群体。理想群体就是一定时间和空间内存在在一起的相互交配而繁殖的一群个体。

下面我们进一步来限定这种理想群体。假定群体中的两性个体,在性成熟个体构成群体的阶段,且雄性和雌性个体所有可能的交配发生概率相同(即不受交配个体间距离、基因型、年龄等的影响),那么这个群体称为随机交配群体。如果是从基因型(或表现型)来看,随机交配群体的特点是:一种基因型个体与另一种基因型个体间交配率等于两种基因型在群体中概率的乘积。

群体遗传学的很多理论都是建立在大的随机交配群体的基础之上的。实际上,对密度较大的昆虫或异花授粉植物,这种理想群体可能是接近真实的群体。而当非随机交配存在时,如部分自花授粉植物的群体,理想群体就不符合实际了,必须作出相应的修正。当然,为了引出群体遗传的概念,我们将从大的随机交配群体开始讨论,以后再逐渐引入其他因素,如非随机交配或小群体等。这里我们先从随机交配平衡群体开始进行讨论。

第一节 基因型频率和基因频率

如果描述一个群体的遗传组成,当然首先就应该详细说明群体的基因型,并说出每种基因型的多少,这可能是一个完整的描述。假定我们讨论的是某一常染色体上的基因位点 A ,并且在这个位点上存在两个不同的等位基因 A_1 和 A_2 。群体的遗传结构,就是用各种基因型的个体数在整个群体中的比率来加以描述,这种比率或频率称为基因型频率。例如我们发现在一个群体中有 $1/4$ 是属于 A_1A_1 型的,则这一种基因型的频率就是 0.25 。当然,把全部基因型频率加在一起应该是 1.00 。

例 1.1 人的 $M-N$ 血型是由在一个位点上的两个等位基因所决定的,三种基因型相当于三种血型: M 型、 MN 型和 N 型。根据 Mourant (1954) 提供的东格陵兰爱斯基摩人和冰岛人中的血型频率比较:

	血型频率			人数
	M 型	MN 型	N 型	
格陵兰	0.835	0.156	0.009	569
冰 岛	0.312	0.515	0.173	747

显然这两个群体在这些基因型频率上是不同的,在格陵兰 N 血型很少见,但在冰岛却相当普遍。在这个位点上,不但两个群体各自内部基因型频率不一致,而且同一基因型在群体间的频率也不一致。

在遗传学意义上,一个群体不但是一群个体,而且是繁育着的一群个体。各个基因通过减数分裂和受精作用,并不发生融合而维持其各自的个体特性。因此群体的遗传学不仅是要考虑到个体

的遗传结构,而且还要考虑到基因从一代传递到下一代的问题。在传递的过程中亲代的基因型被拆散,在后代中又重新由那些配子传递来的基因组成基因型。这样由群体所载有的基因连绵不断地从一代传到下一代,但包含它们的基因型却并不连续。

基因除了以极低的频率发生突变或重组外,其频率是不会发生变化的。因此,只要没有选择、迁移等因素的作用,这个基因在群体中的相对频率(基因频率)也不会发生变化。与此相反,作为二倍体生物的相对频率的基因型频率,在即使基因频率不变,如在交配方式上发生微小的变化,也能使其受到很大的影响。极端的情况是,进行随机交配的植物,一旦进行了自交,在后代中,纯合体比例将不断增加,而杂合体将会减少。如考虑更多基因位点时,与等位基因数相比,等位基因间形成的基因型数目是相当大的。因此,除了考虑群体的基因型频率外,以基因频率为基准表示群体的遗传特性更为重要。群体遗传学主要是研究群体的基因频率变化。

一个群体的遗传组成,依据它所载有的各种基因,可以用对应的基因频率来描述,那就是对在每个位点上所有存在的等位基因及其数目或比率加以详细说明。例如 A_1 是在 A 位点上的一个等位基因,则基因 A_1 的频率或 A_1 的基因频率,乃是在该位点上全部等位基因中 A_1 等位基因所占的比率。在任何一个位点上全部等位基因的频率相加必须等于 1。

在一群个体中,某一特定定位点上的基因频率可以从已知的基因型频率推算。设有两个等位基因 A_1 和 A_2 ,对 100 个个体按各种基因型分类,得:

表 1.1 群体的基因型频率与基因频率

基因型	A_1A_1	A_1A_2	A_2A_2	总 数
个体数	30	60	10	100

续表

基因型	A_1A_1	A_1A_2	A_2A_2	总 数
基因数 A_1	60	60	0	120 80) 总数 200
A_2	0	60	20	

表 1.1 中每个个体含有两个基因, 这样在该位点上共计数了 200 个基因。每个 A_1A_1 个体包含两个 A_1 基因, 每个 A_1A_2 包含一个 A_1 基因。所以在这个样品中共有 120 个 A_1 基因和 80 个 A_2 基因。 A_1 基因的频率为 0.6, A_2 基因的频率是 0.4。为了把这种关系用更通用的方式表示, 可把基因的频率和基因型的频率用符号表示如下:

基因		基因型		
A_1	A_2	A_1A_1	A_1A_2	A_2A_2
频率 p	q	P	H	Q

因此 $p + q = 1$, $P + H + Q = 1$ 。由于每个个体包含两个基因, A_1 基因的频率为 $\frac{1}{2}(2P + H)$ 。群体中基因频率和基因型频率间的关系如下:

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

例 1.2 为了说明如何从基因型频率计算基因频率, 可用上例中所给的 MN 血型的频率来进行。M 血型 and N 血型代表两个纯合的基因型, MN 型则代表杂合的基因型。在格陵兰 M 基因的频率按照式(1.1)计算为 $0.835 + \frac{1}{2}(0.156) = 0.913$, N 基因的频率则为

$0.009 + \frac{1}{2}(0.156) = 0.087$, 这两个频率之和为 1.000。同时对冰岛样品的基因频率计算后获得以下在两群体的基因频率:

	基因频率	
	M	N
格陵兰	0.913	0.087
冰 岛	0.570	0.430

计算结果显示,这两个群体不但在基因型频率上不同,而且在基因频率上也不相同。

从群体的数量也可直接计算基因频率。在 N 个个体组成的群体中, A_1 基因总数为 $2D + H$, A_2 基因总数 $H + 2R$ 。基因频率:

$$p = (2D + H)/(2N) = (D + \frac{H}{2})/N$$

$$q = (H + 2R)/(2N) = (\frac{H}{2} + R)/N$$

例 1.3 如果在 100 个个体的群体中, A_1A_1 、 A_1A_2 、 A_2A_2 基因型分别是 50、32 和 18 个,则:

$$p = (50 + 16)/100 = 0.66$$

$$q = (16 + 18)/100 = 0.34$$

同时,也可先算出基因型频率(0.50、0.32、0.18),后求算基因频率。

对复等位基因位点,基因频率的计算,也可用相同的方式进行。

例 1.4 William(1973)对美国佛罗里达鳗鱼群体己醇脱氢酶位点进行电泳分析,得到以下结果:

基因型	aa	ab	bb	ac	bc	cc	总数
个 数	2	10	44	9	32	12	109

该位点有 3 个等位基因和 6 种基因型,基因型频率如下:

$$p(aa) = 2/109 = 0.0183$$

$$P(ab) = 10/109 = 0.0917$$

$$p(bb) = 44/109 = 0.4037$$

$$p(ac) = 9/109 = 0.0826$$

$$P(bc) = 32/109 = 0.2936$$

$$p(cc) = 12/109 = 0.1001$$

总数为 1.0000。

三个等位基因的频率为:

$$p(a) = p(aa) + \frac{1}{2}p(ab) + \frac{1}{2}p(ac) = 0.1054$$

$$p(b) = p(bb) + \frac{1}{2}p(bc) + \frac{1}{2}p(ab) = 0.5964$$

$$p(c) = p(cc) + \frac{1}{2}p(ac) + \frac{1}{2}p(bc) = 0.2982$$

总数为 1.0000。

一个群体的遗传特性受多种因素的作用,使它在从一代到下一代的传递过程中受到影响,这些起作用的因素形成了研究群体遗传的重要内容。能够使群体的遗传特性发生改变的是以下一些因素。

1. 群体的含量 从一代传到下一代的基因是亲代基因的一个样品。因此在相继世代基因频率间存在着抽样的误差,而且亲代的数目如果愈少,抽样的误差也就愈大。抽样误差的效应将在后面章节讨论,在此之前我们一直先假定是处在一个大群体中,在这里暂对这种由于群体的含量所引起的效应不予讨论,这简单的意味着抽样误差小到可以不计的一种情况。在实际工作中,一个大群体相当于含有以百计的成体而不是以十计成体的群体。

2. 生殖力和生活力的差异 目前我们虽然不去考虑那些基因

的表现型效应,但我们不能忽视它们的生殖力和生活力上的各种效应。因为这些效应影响到下一代的遗传组成,在亲代中不相同的基因型可能具有不相同的生殖力。如果是这样,它们对形成下一代用的配子的供给量也就不会相同。按照这种方式,就有可能在传递过程中使基因频率发生变化。不仅如此,在新形成的合子的那些基因型间,也可能具有不相同的成活率,当个体发育为成体,并且本身成为亲体时,在新一代中的基因频率也可能发生变化,这些过程被称为选择。

3. 迁移和突变 群体中的基因频率还可能由于个体从另一群体的迁入或者通过基因突变而发生变化。

4. 交配系统 在子代中的基因型取决于那些由配子成对结合所形成的合子,而配子的结合又受到亲代交配的影响。因此在子代中的基因型频率受到那些在亲代中交配配偶的基因型的影响。在此处假定,就所讨论的基因型而论,交配是随机的。随机交配,或者随机交配群体,意味着任何一个个体跟群体中任何另一个个体都有均等的交配机会。

第二节 Hardy - Weinberg 定律

群体世代交替中,上一代个体之间不产生选择的交配,是指任一雌配子与任一雄配子结合的概率相同,这时各种基因型间交配频率就是基因型频率的乘积,如 $A_1A_1 \times A_1A_2$ 的频率就等于 PH , $A_1A_2 \times A_2A_2$ 的频率就等于 HQ ,余下类同。

有关交配组合、交配频率以及基因频率与基因型频率在随机交配前提下的平衡,列于表 1.3、1.4。从表中可以看到,在子代群体中基因型比例与上一代的比例相同,这意味着在一个群体中从上一代到下一代基因型频率保持不变,决定基因型频率的基因频率也保持不变。以此类推,再下一代的结果亦相同,这样的群体就