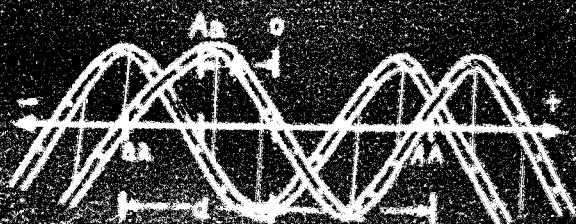


数量遗传学

高之仁 编著



四川大学出版社

内 容 提 要

本书系统介绍了数量遗传学的基本理论和分析方法。主要内容
包括：绪言，加性——显性效应，非等位基因的交互作用与连锁，
基因型与环境的交互作用，遗传交配设计，性状相关分析，遗传距
离与聚类分析，遗传力与选择效果，以及配合力分析等九章。可作
植物遗传育种专业研究生数量遗传学课程教材，亦可供农学类专业
高年级本科生选修本门学科和研究人员对有关数量性状进行遗传分
析时参考。

数 量 遗 传 学

高之仁编著

*

四川大学出版社（成都四川大学内）

四川省新华书店发行

雅安地区印刷厂印刷

*

开本785×1092毫米 1/32 印张16.5 插页2 字数33万字

1986年6月第一版 1986年12月第一次印刷

印数：1—5000册

统一书号：13404·7 定价：3.95元

前 言

本书系作者根据近年来数量遗传学的教学和科研实践以及本门学科的新发展，对原编教材经过数度修改和补充而成。本书在修改过程中荣廷昭副教授、明道绪讲师和龙漫远与潘光堂同志提出了很多宝贵意见。荣廷昭、龙漫远和潘光堂同志在分别讲授五、六与七、八章时均对原教材作了进一步的修改和补充。作者对上述四位同志表示衷心的感谢。

本书在写作过程中曾参阅中外大量文献和专著，限于篇幅，只列出主要参考书目。作者对这些文献和专著的作者致以诚挚的谢意。对于在本书出版过程中给予热情帮助的刘维庆等同志亦一并表示感谢。

由于水平有限，书中还存在不少缺点和错误。这些缺点和错误均由作者本人负责，并请读者批评指正。

高之仁

1985.12于 四川农业大学数量遗传研究室

目 录

绪 言	(1)
第一章 加性和显性效应	(7)
§ 1.1 表型均数的组成分量	(7)
§ 1.2 检验模型	(13)
一、单个尺度检验法	(14)
二、联合尺度检验法	(17)
§ 1.3 变异的组成成分	(28)
一、 F_2 和回交世代中的变异	(28)
二、由 F_2 衍生世代的变异分析	(34)
三、用二级统计量检验模型的适合度	(42)
第二章 非等位基因的互作和连锁	(48)
§ 2.1 非等位基因间互作	(48)
§ 2.2 用加性、显性和双基因互作参数所表示的 世代均数	(53)
§ 2.3 方差和协方差	(62)
§ 2.4 连锁	(68)
第三章 基因型与环境的交互作用	(79)
§ 3.1 基因型与环境交互作用的释义与探测	(79)
§ 3.2 多种环境中基因效应 $\times$ 环境互作的分析	(83)
§ 3.3 多种基因型在多种环境中的基因型 $\times$ 环境 互作分析	(91)

§ 3.4 稳定性分析·····	(102)
一、Eberhart—Russell法·····	(102)
二、戴乔治法·····	(115)
§ 3.5 杂种优势与基因型×环境互作·····	(119)
一、按加性——显性模型测定杂种优势·····	(120)
二、按双基因互作模型测定杂种优势·····	(121)
第四章 遗传交配设计·····	(126)
§ 4.1 双亲本杂交设计·····	(126)
§ 4.2 双列杂交设计·····	(131)
§ 4.3 NC 设计·····	(145)
一、NC I设计·····	(145)
二、NC II 设计·····	(150)
三、NC III 设计·····	(159)
§ 4.4 三重测交法·····	(165)
一、TTC试验设计(标准式)及其遗传学依据·····	(165)
二、统计分析方法·····	(174)
第五章 性状相关分析·····	(188)
§ 5.1 回归与相关·····	(189)
一、线性回归和单相关·····	(189)
二、多元回归与复相关·····	(192)
三、两个多元线性回归方程中回归系数差异显著性 检验·····	(194)
四、在多元线性回归中剔除作用不大的自变量·····	(196)
五、偏相关系数·····	(197)
§ 5.2 通径分析·····	(198)
一、通径系数的意义·····	(198)
二、通径系数的性质·····	(203)

三、确定通径链的原则·····	(215)
四、通径系数的一般计算程序·····	(217)
五、通径分析的统计检验·····	(220)
六、性状相关的通径分析实例·····	(223)
§5.3 典型相关分析 ·····	(226)
一、典型变量与典型相关系数·····	(227)
二、典型相关系数的显著性检验·····	(231)
三、典型相关分析的计算步骤·····	(232)
四、实例·····	(233)
第六章 遗传距离与聚类分析 ·····	(236)
§6.1 遗传距离计算的主成分法·····	(237)
§6.2 遗传距离计算的枢轴凝聚法 (Pivotal codensation method) ·····	(253)
一、综合效应的显著性检验·····	(255)
二、相关变量的转换·····	(262)
三、计算 D^2 值·····	(265)
四、检验 D^2 值的显著性·····	(267)
§6.3 聚类分析——系统聚类法·····	(267)
一、什么是系统聚类法·····	(267)
二、最短距离法·····	(269)
三、类平均法·····	(271)
四、实例分析·····	(275)
§6.4 聚类分析——Tocher聚类法·····	(281)
一、基本思想·····	(282)
二、实例·····	(282)
§6.5 聚类分析——典型根方法·····	(287)
§6.6 聚类分析——模糊聚类法·····	(295)

一、什么是模糊聚类分析.....	(295)
二、分析步骤.....	(299)
三、实例.....	(299)
§6.7 聚类分析——Scott—Knott 聚类法.....	(303)
一、Scott—Knott 法的基本原理和计算程序.....	(303)
二、实例.....	(306)
§6.8 遗传距离与聚类分析几个问题的讨论.....	(310)
一、关于遗传距离的测定方法.....	(310)
二、关于遗传距离应用于亲本选择.....	(311)
三、对系统聚类法的最短距离法使用价值的讨论.....	(313)
第七章 配合力分析	(314)
§7.1 完全双列杂交配合力分析.....	(316)
一、双列杂交类别.....	(316)
二、双列杂交设计配合力分析的统计模型及相应统计分析方法.....	(320)
三、双列杂交法 (Griffing法) 分析实例.....	(351)
§7.2 双列杂交配合力分析中的有关公式的推导.....	(368)
一、配合力平方和计算公式的推导.....	(369)
二、配合力之期望均方.....	(385)
§7.3 不完全双列杂交设计的配合力分析.....	(414)
§7.4 杨氏估算特殊配合力效应的简便方法及其理论基础.....	(434)
一、杨氏估算特殊配合力效应值的方法.....	(435)
二、杨氏简法的理论基础.....	(435)
三、杨氏简法与现行估算配合力方法之间的数学关系.....	(438)
四、三种估算特殊配合力方法的比较.....	(442)

第八章 遗传力与选择效果	(453)
§ 8.1 育种值	(453)
一、群体均数与基因平均效应	(454)
二、育种值 (基因加性值), 显性差值和上位性值	(457)
§ 8.2 遗传力	(466)
一、遗传力的一般概念	(466)
二、遗传力的估计原理	(467)
三、遗传力的估计方法	(469)
(一) 广义遗传力的估计方法	(469)
(二) 狭义遗传力的估计方法	(474)
四、遗传力在作物育种中的应用	(479)
§ 8.3 遗传相关	(480)
一、表型相关的分解	(480)
二、遗传相关的分析方法	(482)
§ 8.4 选择响应及其预测	(491)
一、选择响应的概念	(491)
二、预测选择响应的原理	(492)
三、选择强度	(496)
四、相关遗传进度	(498)
§ 8.5 选择指数	(503)

緒 言

数量遗传学，是以生物群体为对象，研究个体间属于程度或数量上而不是属于种类或质量上差异的遗传现象的科学。这是一门具有鲜明的思辩色彩，并在植物育种实践中具有重大意义的科学，同时也是引起强烈论争但正为越来越多的生物及农业科学工作者所接受和发展的科学。

数量遗传学的诞生应追溯到十九世纪下半叶，那正是整个生物科学蓬勃兴起的年代。1859年，Darwin 出版了不朽的名著《物种起源》，揭示了自然界生物进化的规律。但在进化链条的上下代遗传关系上，他却留下了空白。Mendel 在这方面作出了划时代的贡献，发现了遗传因子的传递规律。可惜1865年他在布隆（Brünn）自然历史学会宣读的论文被冷落了近半个世纪。与此同时，F. Galton 也在努力建立他的遗传学，并于1889年出版《自然遗传》（Natural Inheritance）一书。Galton和 Mendel最大的不同，在于前者观察的是数量性状，后者则以质量性状为研究对象。因此他建立遗传学的目的没有达到。但他发现的亲子遗传（Ancestral Heredity）定律以及创造的回归与相关的工具却是认识和分析数量性状的宝贵财富。1900年，Galton的理论由Weldon及其挚友、著名的应用数学家 Karl Pearson 等人发展成一个非常活跃的学派，即生物统计学派（Biometricians）。这一学派很推崇运用统计学方法研究遗传与进化的作法。

1900年, de Vries、Tschermak和Correns各自又独立地重新发现 Mendel 的理论, 使遗传学的发展跨入新的飞跃时代。但是孟德尔的方法对于数量性状显然是无能为力的, 数量性状在变幅为两个极端之间表现出连续的变异, 其各种表现的频率分布常常近于正态分布曲线, 而不能象 Mendel 观察的质量性状那样可以进行明确而又简单的分组。这就构成了数量性状的一个显著的特征。而 Galton 与 Pearson 等人关于人类体高的工作虽然没能发现遗传传递的方式, 但却发现数量性状至少部分是遗传的。这时的情况, 无论 Galton 还是 Mendel 的方法都不能为数量性状的遗传提供满意的答案。对连续变异的理解还有待于两个学派方法的融汇和互补。Mendel 学派提供了数量性状分析必需依据的原理, 而生物统计学派则表明了分析连续变异的途径和工具。但是这种融汇却在世纪之交两个学派的激烈的论争中被搁置起来了, 任何想要平息论争结合两种思想的作法都是不为任何一方所欢迎的。一方认为连续变异在进化中是没有意义的, 数量性状是不遗传的; 另一方则认为 Mendel 的不连续变异与连续变异是不相容的。

就在两个学派的思想融汇之前, 发生了两件具有重大意义的的事件。

一是 Johanssen (1909) 提出了著名的纯系理论(Pure line theory)。他以菜豆 (*Phaseolus vulgaris*) 为材料, 研究呈数量变异的种子重量性状的遗传。在这个研究中特别重要的是, 他发现数量性状同时受遗传与非遗传因素的控制, 并且二者的效应处于同一数量级, 除了采取繁育措施以外, 没有办法把二者对数量变异的贡献区别开来, 由此可以

清楚基因型与表现型的关系。基因型的不连续的效应可以为环境效应所修饰 (Smooth out) 而在表型上表现为连续变异。运用今天的统计语言, Johanssen 的理论说明一个数量性状的表型值 (P) 应该是支配该性状的基因型值 (G) 与环境效应 (e) 共同作用的结果表现。若 G 和 e 彼此独立地作用于 P, 且 $e \sim N(0, \sigma^2)$ 则有

$$P = G + e$$

在无限群体中, $E(e) = 0$, 于是有

$$P = G$$

在随机样本中为

$$P = \hat{G} + e$$

后来的遗传学工作者发展了这一思想, 提出另一种假定: G 和 e 除各自对 P 发挥作用外, 还存在联合效应, 即互作, 则有

$$P = G + e + G \times e$$

Johanssen 的理论清楚地说明了表现型、基因型和环境三者的关系。

另一件事则是由 Yule (1902、1906) 最先提出并为 Nilsson—Ehle (1908)、Emerson 和 East (1910) 分别在小麦和玉米的试验中得到证实的多基因假说的产生。该假说认为个别作用较小的遗传因子可以解释连续的数量变异, 每一个因子都按 Mendel 的方式遗传, 其变异是不连续的, 因子的作用是累加的。如果控制同一性状的因子数目较多, 就可能产生许多不同的剂量, 从而产生许多中间类型。如果表型的表现与因子的剂量成比例, 变异就会是数量的, 几乎连续的, 并遵从 Galton 的频率曲线。而且非遗传因素的影响会使连续性表现得更为充分。以后 K. Mather 提出“微效多

基因”的概念来表示这类基因以同Mendel 基因相区别。应该注意的是由多基因系统所支配的数量性状，由于基因效应微小，并以累积方式影响相同的表型，且受环境影响大，无法个别地区分单个基因的作用，因而只能按基因作用的类别进行分析研究，探索性状的遗传规律。数量性状受这样的多基因系统支配，为数量性状的第二个基本特征。十多年以后，杰出的数量生物学家Fisher (1918) 发表了一篇划时代的文献“根据孟德尔遗传假设的亲属间相关的研究”成功地运用多基因假说分析资料，首次将数量变异划分为各个分量，开创了数量性状遗传研究的思想方法。

此后，S. Wright、J.B.S. Haldane 以及 K.Mather 等人为数量遗传学体系的形成作出了重大的贡献。而以 Iowa 和 North Carolina 为首的应用数量遗传学派则在对植物育种实际问题的研究中取得丰硕的成果。从60年代末期起，一批批科学工作者进一步探索结合分子生物学、生物化学与数量遗传学进行研究的可能性，为数量遗传学打开了一个新的研究领域。

考查数量遗传学的产生及其发展过程，可以看到迄今为止的数量遗传学，历史上每一次重大的进展，都是由于生物学本身的内容发生了飞跃造成的。围绕这些飞跃，数量遗传学家们也在不断完善和创造新的认识工具与新的形势相适应。因而生物学基本观念的变革促进了数量遗传学发展的根本变革。另一方面，对更大范围的科学领域的考查，又使我们看到，是定性走向定量构成当代科学发展的必然趋势。这里涉及质和量关系的认识。按照辩证唯物主义的观点：一定的质规定着一定的量，一定的量反映一定的质。因此，在生物

学领域中亦必须坚持定量研究的方向。在这方面数量遗传学主要是遵循下列一些步骤分析问题的：

1. 根据事前的假定，对观察数据（遗传现象）用数学公式表示，即建立模型。这里所说事前的假定有两个最基本的，一是e和G对P的作用是否彼此独立，二是各类基因效应是否独立无关。

2. 检验模型，即依据模型所估计的理论值是否与观察值相配合，据此确定拒绝或接受事前的假定。

3. 估算参数。

4. 检验显著性（检验参数估计值是否存在）。

5. 结合实际问题（如育种问题），提出参数解释。

由前面的叙述可见，数量遗传学是和统计思想相联系的。数量遗传学又称为生统遗传学(Biometrical Genetics)或统计遗传学 (Statistical Genetics)。这是由数量性状特征所决定的。因此，在对结果进行解释时，我们始终要考虑到所计算和推测得的各类遗传或非遗传统计量，始终是以一定的概率出现的，而不能以决定论的方式对结果进行解释。运用统计方法和思想分析事物常有两类情况。一是我们对研究对象的细节不感兴趣，而只探讨这些细节构成的宏观表现（或集体表现），因而运用统计思想对这些宏观表现进行描述。二是我们对研究对象的细节无法感兴趣，例如无法把多基因系统中一个个成员的作用区分开来，只能运用统计思想对各类基因效应进行总的研究。虽然存在这两类情况，但就数量性状问题的研究而言，以多基因假说作为考虑问题的一个出发点则是有用的。

最后，应当指出的是，由于历史的成因和科学传统的差

异，当今世界上存在多个数量遗传学派。各个学派都有自己的一套研究方法及略有不同的研究方向和目的，学术观点自然也不尽相同。我国的数量遗传学正处于一个引进，消化，在消化的基础上提高的阶段。为了使读者全面了解国际数量遗传学的理论体系，我们本着“兼容并包”的原则，在本书中尽可能地综合国际上各个学派数量遗传学研究的主要成果，作较全面的介绍，以飨读者。

第一章 加性和显性效应

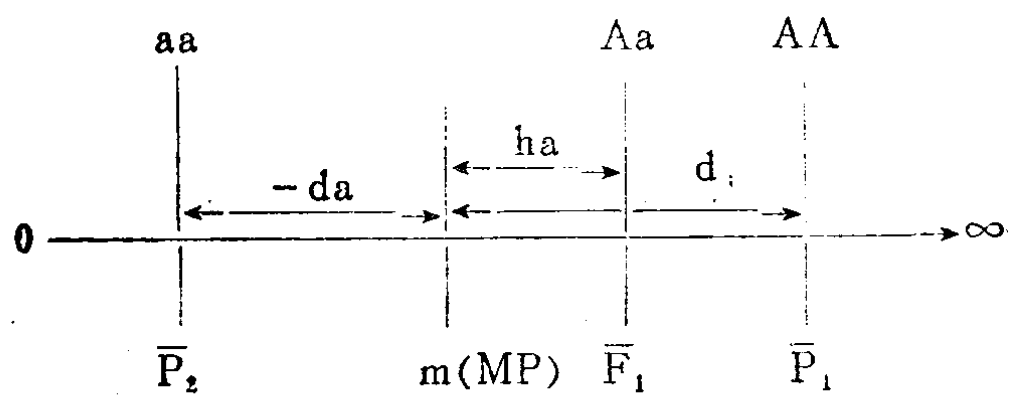
控制数量性状的多基因系统，因单个基因效应甚微，且易受环境影响，难于按表型明确划分不同基因型间的界限，因而不能按传统的孟德尔方法进行遗传分析。但是，任一数量性状均有其特定的多基因系统与环境共同作用所形成的表型频率分布。这样，就可通过生物统计方法，如用均数、方差、协方差等统计量进行描述，并按一定的遗传假设，对多基因系统的基因效应进行分析。在分析中，由于单个基因的作用不能确定，因而所得结果只能反映各类基因效应的总的平均的性质。

§ 1.1 表型均数的组成分量

绪论中已介绍数量性状的表型值为控制该性状的基因型值 (G) 与环境效应 (e) 共同作用的结果表现。如前所述， G 和 e 作用于 P 有两种假设，一是 G 和 e 的作用彼此无关，另是除 G 和 e 各自的作用外，还存在两者的联合效应，即 $G \times e$ 。目前暂时将 e 撇开，只就 G 的剖分进行讨论。对 G 的剖分主要有两个假定，一是各个位点的基因效应对 G 的贡献，彼此独立无关，称为独立性假定 (assumption of independence)，另一个是各位点除本身作用外，还有在位点间的联合效应，称为非独立性假定 (assumption of non-independence)。

本章只就第一种假定讨论G的剖分。这种假定的含义是指各位点上的基因效应是可加的，纯结合位点上的基因效应称为加性效应，杂结合位点的基因效应称为位点内的互作效应或显性效应。所有位点的基因效应对G的贡献为累积的，同时又由于位点的基因效应有正向和负向之分，因而这种累积效应为代数和效应。按照这一假定，作用于G的位点效应只有加性效应和显性效应，因此合称为加性和显性效应，据此建立的遗传模型称为加性——显性模型。

为简便计，先讨论一对等位基因。设在二倍体的体细胞染色体上，某一位点有A和a两个等位基因，其可能的基因型为AA，Aa和aa三种。这三种基因型的表型均数可用一条直线上三点表示如下：



图中 $\overline{P}_2$ 表示aa的平均表型值， $\overline{P}_1$ 表示AA的平均表型值， $\overline{F}_1$ 表示Aa的平均表型值。通常用 $\overline{P}_1$ 表示较大或增效亲本表型平均值， $\overline{P}_2$ 表示较小或减效亲本表型平均值。 m 为双亲的平均值，称为中亲值，记为 m 或MP， $m = \frac{1}{2}(\overline{P}_1 + \overline{P}_2)$ 。 $\pm da$ 为 $\overline{P}_1$ 和 $\overline{P}_2$ 离 m 的偏差， ha 为 $\overline{F}_1$ 离 m 之差，其值有正有负。亲本离中亲值之差称为加性效应， $\overline{F}_1$ 离中亲值之差称为显性效应。这里所采用大写字母A并不含有显性之意，只表

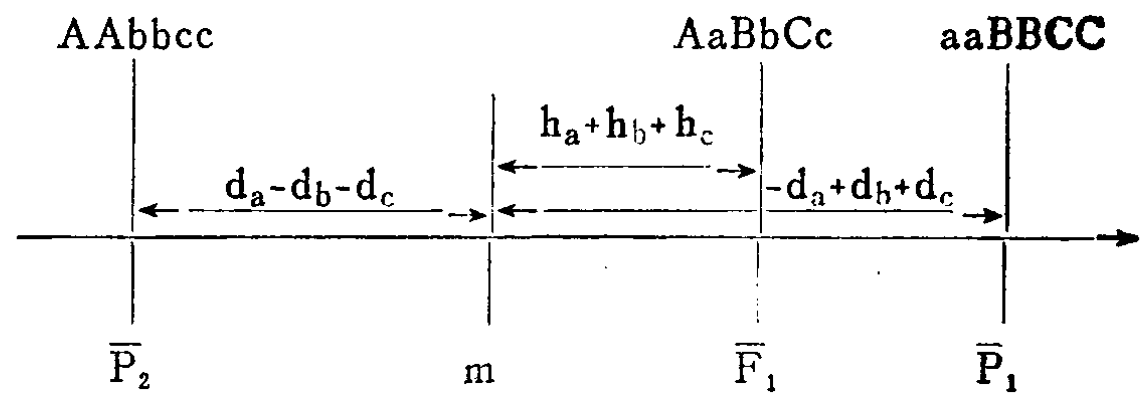
示增大或增效。例如 $Aa = m$ 时, $h_a = 0$ 表示无显性; 如 $-d_a < h_a < 0$, 则 a 对 A 为部分显性; 如 $0 < h_a < d_a$ 则 A 对 a 为部分显性; 如 $h_a = \pm d_a$, 则为完全显性; 如 $h_a > |d_a|$, 则 $\bar{F}_1$ 超出两纯合子所规定的界限。就一个位点而言, 这种现象称为超显性。从上述图解还可看到, 一个位点上的三种表型值均数能用下列代数式表示之。

$$\bar{P}_1 = m + d_a$$

$$\bar{P}_2 = m - d_a$$

$$\bar{F}_1 = m + h_a$$

如涉及三对独立基因, 则有



从上述图解可以看出

$$\bar{P} = m + (-d_a + d_b + d_c) = m + (d_b + d_c - d_a)$$

$$\bar{P}_2 = m + (d_a - d_b - d_c) = m - (d_b + d_c - d_a)$$

$$\bar{F}_1 = m + (h_a + h_b + h_c)$$

显然括弧内为各位点效应的代数, 如涉及 K 对独立基因, 则用 $[d]$ 和 $[h]$ 表示两种效应的代数。这里 $[d] = \sum d_+ - \sum d_-$, 式中 $\sum d_+$ 表示增效位点效应和, $\sum d_-$ 表示减效应位点效应和。