

胚乳性状的遗传模型及其分析方法<sup>\*,\*\*</sup>

朱军

许馥华

(浙江农业大学, 浙江杭州 310029)

## 提 要

提出分析谷类作物胚乳数量性状世代平均数的遗传模型。该模型把控制胚乳性状的总遗传效应  $G$  分解为胚乳直接效应  $G_e$ 、细胞质效应  $C$  和母体植株效应  $G_m$ 。而  $G_e$  又可分解为直接加性 ( $A$ ) 和直接显性 ( $D$ ) 遗传分量,  $G_m$  则可分解为母体加性 ( $A_m$ ) 和母体显性 ( $D_m$ ) 遗传分量。用 MINQUE (0/1) 法可以无偏地估算该模型中的直接遗传方差分量  $\sigma_a^2$  和  $\sigma_d^2$ 、细胞质遗传方差  $\sigma_c^2$  以及母体植株的遗传方差分量  $\sigma_{A_m}^2$  和  $\sigma_{D_m}^2$ , 也可估算胚乳基因与母体植株基因的遗传协方差分量  $\sigma_{A, A_m}$  和  $\sigma_{D, D_m}$ 。成对性状的遗传协方差分量同样可用 MINQUE (0/1) 法估算。该模型中的随机遗传效应则可用 MINQUE (0/1) 法作调整无偏预测。由于遗传变异是从若干组合的不同世代间估计, 因而有关胚乳数量性状的测定可以整个世代的籽粒为样本。

关键词 胚乳性状, 遗传模型, 遗传方差和协方差估算, 遗传效应值预测

谷类作物

胚乳是谷类作物种子的主要组成部分。谷类作物种子中的营养成分, 如蛋白质、必需氨基酸、油分等, 都主要储存在胚乳中。由于种子胚乳是三倍体, 而提供其营养物质的母体植株是二倍体。这一复杂的遗传体系给胚乳性状的遗传分析带来很大的困难。能否提出具有生物学意义的遗传模型和简易、实用的统计分析方法, 是决定胚乳数量性状遗传研究能否取得成效的关键所在。莫惠栋<sup>[1,2,9]</sup>提出了三倍体胚乳性状的遗传模型和详细的统计分析方法。Bogyo 等<sup>[5]</sup>也曾提出过类似的三倍体模型。这些遗传模型都没有包括植株母体效应对胚乳性状的作用。由于胚乳的营养物质由母体植株所提供, 有些胚乳性状的表现可能会同时受到胚乳的三倍体基因和母体植株的二倍体基因两套遗传体系的控制。另外细胞质基因也可能通过控制叶绿体(或线粒体)的合成而影响植株的光合(或呼吸)作用, 从而间接控制胚乳性状的表现。Foolad 和 Jones<sup>[6]</sup>曾于 1992 年提出了一个复杂的生物模型, 可以同时分析种子的胚、胚乳、母体和细胞质等项遗传效应。以上这些模型都采用固定模型(Fixed Model)分析世代平均数。朱军运用混合线性模(Mixed Linear Model)原理, 提出了无偏估算随机效应方差和协方差分量<sup>[3]</sup>以及预测随机效应值<sup>[3,4]</sup>的新方法, 可用于分析复杂的遗传模型。

本文的目的是: 提出一个胚乳数量性状的遗传模型, 可以用世代平均数同时分析胚乳、细胞质和母体植株三套遗传体系的基因效应; 提出估算种子性状遗传率和遗传相关的方法; 用蒙特卡罗模拟(Monte Carlo Simulation)的方法, 比较三种不同遗传交配设计的分析效益; 比较预测随机遗传效应值的二种方法, 即 BLUP (Best Linear Unbiased Prediction) 法<sup>[7]</sup>和 AUP (Adjusted Unbiased Prediction) 法<sup>[4]</sup>。

\* 国家教委资助项目。

\*\* 本文承蒙李道藩教授审阅, 提出宝贵意见, 特此致谢。

本文于 1992 年 10 月 12 日收到, 1993 年 3 月 24 日终审完毕。

## 1 遗传模型和统计分析方法

当胚乳性状同时受到胚乳、细胞质和母体植株的基因控制时,其总遗传效应( $G$ )可以分解为胚乳直接遗传效应( $G_e$ )、细胞质遗传效应( $C$ )和母体植株遗传效应( $G_m$ )<sup>[3]</sup>。胚乳直接遗传效应( $G_e$ )又可被进一步分解为直接加性( $A$ )和直接显性( $D$ )遗传分量。母体植株遗传效应( $G_m$ )则可进一步被分解为母体加性( $A_m$ )、母体显性( $D_m$ )遗传分量。假定(1)不存在父体效应(Paternal Effect), (2)细胞质基因只通过母体稳定遗传, (3)不存在上位性效应, (4)不存在高价显性效应(三个等位基因之间的互作)和(5)不存在基因型与环境的互作效应,胚乳数量性状表现型值可用以下线性模型表示。一组纯系材料作遗传交配,第  $i$  个母本与第  $j$  个父本的第  $k$  种交配类型世代平均数的第  $l$  次重复的表现型值 ( $y_{ij\mu l}$ ) 可以表示为

$$y_{ij\mu l} = \mu + G_{ij\mu} + e_{ij\mu l}$$

其中群体平均数  $\mu$  是固定效应; 剩余机误  $e_{ij\mu l}$  是随机效应,  $e_{ij\mu l} \sim (0, \sigma_e^2)$ ;  $G_{ij\mu}$  是世代平均数的遗传效应。 $G_{ij\mu}$  的遗传组成决定于不同亲本 ( $i, j$ ) 及交配类型 ( $k$ ), 交配类型总共可有七种 ( $k=0, 1, \dots, 6$ )。

亲本遗传效应分量 ( $k=0, i=j$ )

$$G_{ij0} = 3A_i + 3D_i + C_i + 2A_m + D_m$$

杂种一代  $F_{1ij}$  遗传效应分量 ( $k=1$ )

$$G_{ij1} = 2A_i + A_j + D_i + 2D_{ij} + C_i + 2A_m + D_m$$

杂种二代  $F_{2ij}$  遗传效应分量 ( $k=2$ )

$$G_{ij2} = 1.5A_i + 1.5A_j + D_i + D_{jj} + D_{ij} + C_i + A_m + A_{mj} + D_m$$

回交  $F_{1ij} \times P_i$  遗传效应分量 ( $k=3$ )

$$G_{ij3} = A_i + 2A_j + 0.5D_i + 1.5D_{jj} + D_{ij} + C_i + A_m + A_{mj} + D_m$$

回交  $F_{1ij} \times P_j$  遗传效应分量 ( $k=4$ )

$$G_{ij4} = 2A_i + A_j + 1.5D_i + 0.5D_{jj} + D_{ij} + C_i + A_m + A_{mj} + D_m$$

回交  $P_j \times F_{1ij}$  遗传效应分量 ( $k=5$ )

$$G_{ij5} = 0.5A_i + 2.5A_j + 2D_{jj} + D_{ij} + C_j + 2A_m + D_m$$

回交  $P_i \times F_{1ij}$  遗传效应分量 ( $k=6$ )

$$G_{ij6} = 2.5A_i + 0.5A_j + 2D_{ii} + D_{ij} + C_i + 2A_m + D_m$$

如果杂交亲本是从一个推断群体中抽取的随机样本, 以上各项遗传效应分量均为随机效应。随机遗传效应的定义分别为: 直接加性效应  $A_i$  或  $A_j \sim (0, \sigma_A^2)$ ; 直接显性效应  $D_i, D_{jj}$  或  $D_{ij} \sim (0, \sigma_D^2)$ ; 细胞质效应  $C_i \sim (0, \sigma_C^2)$ ; 母体加性效应  $A_m$  或  $A_{mj} \sim (0, \sigma_{A_m}^2)$ ; 母体显性效应  $D_{mi}, D_{mjj}$  或  $D_{mij} \sim (0, \sigma_{D_m}^2)$ 。由于胚乳基因的三分之二来自其母体植株, 因而直接遗传效应与母体遗传效应不是相互独立的, 存在加性遗传协方差  $\text{Cov}(A_i, A_m) = \sigma_{A, A_m}$  和显性遗传协方差  $\text{Cov}(D_{ij}, D_{mj}) = \sigma_{D, D_m}$ 。

胚乳性状遗传模型可以进一步改写成矩阵形式的混合线性模型

$$y = 1\mu + U_A e_A + U_D e_D + U_C e_C + U_{A_m} e_{A_m} + U_{D_m} e_{D_m} + e_e = 1\mu + \sum_{u=1}^5 U_u e_u$$

其中常数向量  $1$  是  $1$  的向量,  $e_u$  是第  $u$  项随机效应向量,  $e_u \sim (0, \sigma_u^2 I)$ ;  $U_u$  是随机效应向量  $e_u$  的系数矩阵 ( $u=1, 2, \dots, 5$ ),  $U_u = I$  是单位矩阵。

胚乳数量性状表现型向量  $y$  的方差和协方差矩阵为

$$\begin{aligned} \text{Var}(y) = & \sigma_A^2 \mathbf{U}_A \mathbf{U}_A^T + \sigma_D^2 \mathbf{U}_D \mathbf{U}_D^T + \sigma_C^2 \mathbf{U}_C \mathbf{U}_C^T + \sigma_{A_m}^2 \mathbf{U}_{A_m} \mathbf{U}_{A_m}^T + \sigma_{D_m}^2 \mathbf{U}_{D_m} \mathbf{U}_{D_m}^T \\ & + \sigma_{A, A_m} (\mathbf{U}_A \mathbf{U}_{A_m}^T + \mathbf{U}_{A_m} \mathbf{U}_A^T) + \sigma_{D, D_m} (\mathbf{U}_D \mathbf{U}_{D_m}^T + \mathbf{U}_{D_m} \mathbf{U}_D^T) + \sigma_e^2 \mathbf{I} \end{aligned}$$

其中  $\sigma_A^2$  是直接加性方差,  $\sigma_D^2$  是直接显性方差,  $\sigma_C^2$  是细胞质方差,  $\sigma_{A_m}^2$  是母体加性方差,  $\sigma_{D_m}^2$  是母体显性方差,  $\sigma_{A, A_m}$  是直接加性效应与母体加性效应的协方差,  $\sigma_{D, D_m}$  是直接显性效应与母体显性效应的协方差,  $\sigma_e^2$  是剩余效应方差。

尽管模型中给出了七个遗传交配类型, 实际应用时并不一定要包括全部的交配类型。可以采用三种不同的交配设计: (I)  $k=0, 1, 2$ ; (II)  $k=0, 1, 2, 3$ , 或 (III)  $k=0, 1, \dots, 6$ 。

采用 MINQUE(0/1)法<sup>[3]</sup>, 可以无偏地估算混合线性模型中的各项方差分量。MINQUE(0/1)是一种最小范数二阶无偏估算法<sup>[12]</sup>(Minimum Norm Quadratic Unbiased Estimation, MINQUE法), 它设所有的方差先验值为常数1、协方差先验值为常数0。按以下公式可估算  $F_2$  种子的遗传方差和遗传协方差分量,  $V_A = 4.5\sigma_A^2$ ,  $V_D = 3\sigma_D^2$ ,  $V_C = \sigma_C^2$ ,  $V_{A_m} = 2\sigma_{A_m}^2$ ,  $V_{D_m} = \sigma_{D_m}^2$ ,  $C_{A, A_m} = 3\sigma_{A, A_m}$ ,  $C_{D, D_m} = \sigma_{D, D_m}$ ,  $V_e = \sigma_e^2$ 。胚乳数量性状表现型方差( $V_P$ )包括  $V_P = (V_A + V_D) + V_C + (V_{A_m} + V_{D_m}) + 2(C_{A, A_m} + C_{D, D_m}) + V_e = V_{G_0} + V_C + V_{G_m} + 2C_{G_0, G_m} + V_e$ 。

其中  $V_{G_0}$  是胚乳直接遗传方差,  $V_C$  是胞质遗传方差,  $V_{G_m}$  是母体植株遗传方差,  $C_{G_0, G_m}$  是胚乳核基因效应与母体植株核基因效应的遗传协方差,  $V_e$  是剩余机误差。

由于胚乳性状同时受到三套遗传体系(胚乳核基因、细胞质基因和母体核基因)的控制, 狭义遗传率也应包括直接遗传率( $h_d^2$ )、胞质遗传率( $h_c^2$ )和母体遗传率( $h_m^2$ )三项。各项遗传率的计算公式为,

$$h_d^2 = (V_A + C_{A, A_m}) / V_P$$

$$h_c^2 = V_C / V_P$$

$$h_m^2 = (V_{A_m} + C_{A, A_m}) / V_P$$

而总的狭义遗传率则为

$$\begin{aligned} h^2 &= h_d^2 + h_c^2 + h_m^2 \\ &= (V_A + V_C + V_{A_m} + 2C_{A, A_m}) / V_P \end{aligned}$$

根据混合线性模型的原理, 胚乳成对性状( $y_1$  和  $y_2$ )之间的协方差矩阵为

$$\begin{aligned} \text{Cov}(y_1, y_2) = & \sigma_{A/A} \mathbf{U}_A \mathbf{U}_A^T + \sigma_{D/D} \mathbf{U}_D \mathbf{U}_D^T + \sigma_{C/C} \mathbf{U}_C \mathbf{U}_C^T + \sigma_{A_m/A_m} \mathbf{U}_{A_m} \mathbf{U}_{A_m}^T + \sigma_{D_m/D_m} \mathbf{U}_{D_m} \mathbf{U}_{D_m}^T + \sigma_{A, A_m} \\ & (\mathbf{U}_A \mathbf{U}_{A_m}^T + \mathbf{U}_{A_m} \mathbf{U}_A^T) + \sigma_{D, D_m} (\mathbf{U}_D \mathbf{U}_{D_m}^T + \mathbf{U}_{D_m} \mathbf{U}_D^T) + \sigma_{e/e} \mathbf{I} \end{aligned}$$

其中  $\sigma_{A/A}$  是直接加性协方差,  $\sigma_{D/D}$  是直接显性协方差,  $\sigma_{C/C}$  是细胞质协方差,  $\sigma_{A_m/A_m}$  是母体加性协方差,  $\sigma_{D_m/D_m}$  是母体显性协方差,  $\sigma_{A, A_m}$  是一个性状的直接加性效应与另一个性状的母体加性效应的协方差,  $\sigma_{D, D_m}$  是一个性状的直接显性效应与另一个性状的母体显性效应的协方差,  $\sigma_{e/e}$  是剩余效应协方差。

也可以用 MINQUE(0/1)法无偏估算  $F_2$  种子的成对性状之间的协方差分量<sup>[3]</sup>。成对性状之间的遗传协方差分量可按以下公式估算:  $C_A = 4.5\sigma_{A/A}$ ,  $C_D = 3\sigma_{D/D}$ ,  $C_C = \sigma_{C/C}$ ,  $C_{A_m} = 2\sigma_{A_m/A_m}$ ,  $C_{D_m} = \sigma_{D_m/D_m}$ ,  $C_{A, A_m} = 3\sigma_{A, A_m}$ ,  $C_{D, D_m} = \sigma_{D, D_m}$ ,  $C_e = \sigma_{e/e}$ 。成对性状的表现型协方差( $C_P$ )则包括

$$C_P = (C_A + C_D) + C_C + (C_{A_m} + C_{D_m}) + 2(C_{A, A_m} + C_{D, D_m}) + C_e = C_{G_0} + C_C + C_{G_m} + 2C_{G_0, G_m} + C_e$$

其中  $C_{G_0}$  是胚乳直接遗传协方差,  $C_C$  是胞质遗传协方差,  $C_{G_m}$  是母体植株遗传协方差,  $C_{G_0, G_m}$  是一个性状的胚乳核基因效应与另一个性状的母体植株核基因效应的遗传协方差,  $C_e$  是剩余机误协方差。

分别获得胚乳成对性状的方差分量和协方差分量的无偏估计值以后,可以进一步估算各项相关系数:直接加性相关  $r_A = C_A / \sqrt{(V_{A(1)}V_{A(2)})}$ , 直接显性相关  $r_D = C_D / \sqrt{(V_{D(1)}V_{D(2)})}$ , 细胞质相关  $r_c = C_C / \sqrt{(V_{C(1)}V_{C(2)})}$ , 母体加性相关  $r_{Am} = C_{Am} / \sqrt{(V_{Am(1)}V_{Am(2)})}$ , 母体显性相关  $r_{Dm} = C_{Dm} / \sqrt{(V_{Dm(1)}V_{Dm(2)})}$ , 剩余效应相关  $r_e = C_e / \sqrt{(V_{e(1)}V_{e(2)})}$ 。

国际上通常采用 BLUP 法<sup>[7]</sup>预测随机遗传效应值。当用方差分量的估计值代替 BLUP 预测公式中的方差参数,预测遗传效应值便失去了线性,也不再保证能获得无偏预测。我们建议采用调整无偏预测(简称 AUP)方法<sup>[4]</sup>预测随机遗传效应值。

可采用 Jackknife 的方法<sup>[8,3]</sup>计算各项遗传参数的估计值(或预测值)及其标准误。每个世代平均数的  $r$  次观察值可作为 Jackknife 的重复抽样单位。Jackknife 估计值(或预测值)具有  $t$ -分布,因而可用  $t$  测验对遗传参数作统计检验。

## 2 蒙特卡罗模拟结果

分别以 200 次蒙特卡罗模拟的结果计算遗传参数( $\hat{\Phi}$ )平均估计值( $\hat{\Phi}$ )、遗传参数的估计偏差( $\text{Bias} = \hat{\Phi} - \Phi$ )、均方误差( $\text{MSE} = \text{Bias}^2 + \text{Var}(\hat{\Phi})$ )和功效值(在 5% 水平下测出参数显著性的概率值)。并用这些统计量检验不同方法估算遗传方差、遗传协方差或预测遗传效应值的无偏性和有效性。模型中随机遗传效应的预测值向量  $\hat{e}$  与实际值向量  $\bar{e}$  之间的距离( $\|\hat{e} - \bar{e}\|$ )是比较预测效果的主要指标。用该指标比较了 BLUP 和 AUP 两种预测方法的效益。

为了用蒙特卡罗模拟比较三种交配设计的分析效果,可以人为地设置三种交配设计,并使它们的试验群体相接近(约有 35 个世代平均数)。交配设计 I 包括了 6 个亲本, 30 个正反交  $F_{1ij}$  和 30 个  $F_{2ij}$ ; 交配设计 II 包括了 5 个亲本, 20 个正反交  $F_{1ij}$ , 20 个  $F_{2ij}$  和 20 个  $F_{1ij} \times P_j$ ; 交配设计 III 包括了 5 个亲本, 10 个正交  $F_{1ij} (i < j)$ , 10 个  $F_{2ij} (i < j)$ , 10 个  $F_{1ij} \times P_j$ , 10 个  $F_{1ij} \times P_i$ , 10 个  $P_j \times F_{1ij}$  和 10 个  $P_i \times F_{1ij}$ 。

蒙特卡罗模拟结果表明,当世代平均数有一次或二次重复时,均能获得各项遗传参数的无偏估计。与只重复一次的相比,重复二次的均方误差很接近,但功效值略高。因此,增加重复观察值能提高遗传参数的估计效益。采用三种交配设计所估算的遗传参数在偏差、均方误差和功效值等方面均不存在明显的差异。这表明只要试验的遗传群体大小相似,采用不同的交配设计均可获得统计效益相等的参数估计值。但是不同交配设计当年需要做的杂交世代数却相差很多,设计 I 为 30 个,设计 II 为 40 个,设计 III 为 50 个。交配设计 I 所需要的杂交制种工作量比其它两种交配设计要少。因此采用只包括亲本、 $F_1$  和  $F_2$  的交配设计 I, 便能有效地估计胚乳数量性状的遗传参数。

虽然遗传模型中包括了胚乳、细胞质和母体植株三套遗传体系的基因效应,这并不意味着所有胚乳数量性状都受这三套遗传体系控制。有些性状可能只明显受其中的一套或二套遗传体系控制。如果性状只受细胞质和母体植株的基因控制时,用全模型分析总是可以正确地鉴别出不存在直接遗传方差( $\sigma_A^2 = \sigma_D^2 = 0$ )及协方差( $\sigma_{A,Am} = \sigma_{D,Dm} = 0$ ),同时细胞质和母体植株的遗传方差却仍可无偏地估算。不存在细胞质效应时,所估算的细胞质方差不显著的概率约为 99% (1.00-功效值)。其它遗传方差和遗传协方差分量仍然可以无偏估算。当母体效应不存在时,胚乳的直接显性方差、以及细胞质遗传方差不但可以无偏估算,而且还更容易测出其显著性(有

较高的功效值)。母体效应方差分量( $\sigma_{Am}^2$ 、 $\sigma_{Bm}^2$ )和直接效应与母体效应的协方差分量( $\sigma_{A_m}^2$ 、 $\sigma_{B_m}^2$ )不显著的概率也可高达94%以上。蒙特卡罗模拟结果证明,全模型可以无偏地估算胚乳数量性状的各项遗传方差分量和协方差分量。

胚乳性状遗传模型不但适用于单一性状的遗传分析,还可用于成对性状间的遗传协方差分量估算。采用三种遗传交配设计,对所有的遗传协方差分量均可以无偏地估算,并可获得相似的估算效果。因而这个遗传模型及其分析方法,为胚乳不同性状之间遗传相关性分析提供了有效的手段。

本文所提出的遗传模型要求杂交亲本是一个推断群体的随机样本。通过对样本材料的统计分析,可以推断这个群体中胚乳数量性状的遗传规律。但是在实际应用时,育种工作者往往还希望对杂交亲本的育种价值也能有所了解。因此有必要预测各项遗传效应值。两种预测方法(BLUP和AUP)都能获得遗传效应的无偏预测值,AUP的均值( $\approx 10^{-6}$ )比BLUP的均值( $\approx 10^{-2}$ )更接近遗传效应的均值参数(零)。用预测值计算的方差是否接近方差参数,是判断预测有效性的一项指标。对于所有的遗传效应,用BLUP值计算的方差都略低于其方差参数(估计值/参数 $\approx 0.8$ )。这表明,遗传效应的BLUP绝对值比实际遗传效应的绝对值小。而采用AUP法所获得的遗传效应预测值,其方差是模型方差参数的无偏估计(估计值/参数 $\approx 1.0$ ),因而遗传效应的AUP绝对值与实际遗传效应的绝对值非常接近。

遗传效应的预测值与实际值之间的距离是度量预测有效性的最重要指标,距离越小,表明预测值越接近实际值。在所有的线性无偏预测中,BLUP值的距离应该是最小的。本试验中各项遗传效应AUP值的距离都非常接近(略大于)BLUP值的距离。在实际应用时,由于方差参数是未知的,因而BLUP值是无法算得的。而在估算遗传方差分量的同时,不需太多的计算便可获得AUP值。AUP值虽然不是遗传效应的线性预测,却是无偏和有效的预测。

### 3 讨 论

如果胚乳性状主要受胚乳遗传效应控制,采用莫惠栋提出的三倍体胚乳性状遗传模型及其统计分析方法<sup>[1,2,9]</sup>,能有效地分析胚乳核基因的加性和显性效应。已有研究表明<sup>[11,13]</sup>,谷类作物胚乳的一些数量性状可能还受到细胞质和母体植株的遗传影响。当胚乳性状还受到细胞质和母体植株的基因控制时,用三倍体模型分析可能会导致有偏的估算。蒙特卡罗模拟结果表明,本文所提出的模型可以无偏地分析三套遗传体系的基因效应。

Foolad和Jones的遗传模型<sup>[6]</sup>也可以分别估算细胞质效应和母体加、显性效应。但是他们的模型要求分析18个世代(其中包括9个需要杂交的世代),并且需测定每个世代的单粒种子。与Foolad和Jones的遗传模型<sup>[6]</sup>相比,本文所提出的遗传模型及分析方法有以下几条优越性:(1)可以无偏估算成对性状的遗传协方差分量;(2)在估算遗传方差分量和遗传协方差分量的同时,可以无偏预测全部遗传体系中的基因效应值;(3)不需要分析单粒或单株的胚乳品质性状,每个世代类型(包括分离世代)植株上的种子可以混合后抽样分析;(4)不需要众多的杂交类型,采用交配设计I只需若干组合的亲本、 $F_1$ 和 $F_2$ ,便可有效地分析胚乳数量性状;(5)对于一些组合或交配类型缺失的非平衡数据,仍然可以获得各项遗传参数的无偏估计(或预测)。

我们所提出的遗传模型是分析作为随机样本的一组亲本之间的遗传变异。而现有的胚乳遗传模型大多采用固定模型,分析特定组合内不同遗传世代,推断双亲间的遗传变异,其分析结果适用于所分析的特定组合。如果需要推断某个遗传群体胚乳性状的一般遗传规律,可在该

群体中挑选一些亲本相互杂交,分析这些组合的综合表现,从而推断该群体内这些亲本之间的遗传变异。当这些亲本是群体中的一个随机样本时,采用随机模型和固定模型的分析结果应该是基本一致的。国内已有不少学者采用世代平均数的分析方法,分析了胚乳性状的遗传规律。如果其试验分析的一组亲本是某个遗传群体中有代表性的遗传材料,那么本文提出的遗传模型仍然适用于分析该试验的资料,并能提供更全面的遗传信息。

分析数量性状的线性模型,通常将等位基因之间的互作定义为显性效应。如果遗传材料包括了亲本或分离世代,模型中将有纯合显性( $D_{ii}, D_{mii}$ )和杂合显性( $D_{ij}, D_{mij}$ )二种显性效应。由于显性效应的期望值为零,因而可以用纯合显性效应的总和( $\sum_i D_{ii}, \sum_i D_{mii}$ )分别度量胚乳和母体遗传的自交衰退(或杂种优势)。 $\sum_i \hat{D}_{ii} < 0$ (即  $\sum_i \Sigma_{j>i} \hat{D}_{ij} > 0$ )表明胚乳核遗传存在杂种优势。 $\sum_i \hat{D}_{mii} < 0$ (即  $\sum_i \Sigma_{j>i} \hat{D}_{mij} > 0$ )则表明母体核遗传存在杂种优势。

采用本文提出的遗传模型分析胚乳性状,要求杂交亲本数目不宜太少。亲本数目多、试验群体大,能提高分析的效益(有较高的功效值)。但是如果亲本数目太多,也会增加很多杂交制种和品质分析的工作量,因而会增加试验成本。本文蒙特卡罗模拟结果表明,采用不少于五至六个亲本的双列杂交,三种交配设计均可望获得各项遗传参数的无偏估计。当需要分析很微弱的遗传效应时,有必要采用更多的杂交亲本。

交配设计 I 包括了一些杂交组合的亲本、 $F_1$  和  $F_2$ , 其中  $F_1$  种子需要人工杂交产生。杂交制种不仅会增加试验的工作量,还可能会损坏花器、影响  $F_1$  种子的正常发育。如果遗传试验种植杂交组合的亲本、 $F_1$  和  $F_2$  植株,然后自交产生亲本、 $F_2$  和  $F_3$  种子,便可全部采用自交种子进行遗传分析。 $F_{3ij}$  胚乳性状的平均基因型值 =  $1.5A_i + 1.5A_j + 1.25D_{ii} + 1.25D_{jj} + 0.5D_{ij} + C_i + A_{mi} + A_{mj} + 0.25D_{mii} + 0.25D_{mjj} + 0.5D_{mij}$ 。由于  $F_3$  种子长在基因型分离的  $F_2$  植株上,因此需要注意采集众多单株上的种子,分析较大的种子样本,才能避免抽样偏差。另一种更简单的试验设计只需种植双列杂交的亲本和  $F_2$  植株,并自交产生亲本和  $F_2$  种子。采用这种简易交配设计时,需用求广义逆的方法才能解 MINQUE(0/1) 方程组,可分析不包括细胞质效应的其它各项遗传效应。

本文所介绍的胚乳遗传模型的分析,是建立在混合线性模型统计分析的基础上的。这些统计分析方法虽然优胜于传统的方差分析方法,但也存在一些局限性。其中最大的局限性是:这些统计分析方法需要矩阵的复杂运算,因而只能在内存大、运算快的电脑上才能完成分析。目前我们已经研制成适用于分析各种交配设计的平衡或非平衡数据的计算机软件,可供分析单位使用。

## 参 考 文 献

- [1] 莫惠栋, 1989a, 遗传学报, 16, 111—117.
- [2] 莫惠栋, 1989b, 遗传学报, 16, 333—341.
- [3] 朱军, 1992, 生物数学学报, 7, 1—11.
- [4] 朱军, 1993, 生物数学学报, 8, 32—44.
- [5] Bogyo, T. P., et al., 1988, Heredity, 60, 61—67.
- [6] Foolad, M. R. & R. A. Jones, 1992, Theor. Appl. Genet., 83, 360—366.
- [7] Henderson, C. R., 1963, In: Hanson, W. D. & H. F. Robinson, (eds.), Statistical Genetics and Plant Breeding. Natl. Acad. Sci. — N R C 982: 141—163.

- [ 8 ] Miller, R. G. , 1974, *Biometrika*, 61, 1—15.
- [ 9 ] Mo, H. D. , 1988, In: Weir B. S. et al. , (eds), *Proc. of the 2nd International Conference on Quantitative Genetics*. Sinauer, Sunderland/MA, pp. 478—487.
- [10] Poneleit, C. G. & D. B. Egli, 1983, *Crop Sci.* , 23, 871—875.
- [11] Rao, A. P. & A. A. Fleming, 1978, *Crop Sci.* , 18, 935—937.
- [12] Rao, C. R. , 1971, *J. Multivar. Anal.* , 1, 257—275.
- [13] Weber, E. J. , 1983, *Biochemical Genet.* , 21, 1—13.

## A Genetic Model and Analysis Methods for Endosperm Traits

Zhu Jun      Xu Fu-hua

(*Agronomy Department, Zhejiang Agricultural University, Hangzhou 310029*)

### Abstract

A Genetic model is proposed for analyzing generation means of endosperm quantitative traits in cereal crops. The total genetic effect ( $G$ ) of endosperm quantitative traits is partitioned into endosperm direct genetic effect ( $G_o$ ), cytoplasm genetic effect ( $C$ ), and maternal plant genetic effect ( $G_m$ ). Endosperm direct genetic effect can be further partitioned into direct additive ( $A$ ) and direct dominance ( $D$ ) genetic components. Maternal plant genetic effect ( $G_m$ ) can also be partitioned into maternal additive ( $A_m$ ) and maternal dominance ( $D_m$ ) genetic components. MINQUE (0/1) method can be used for estimating endosperm direct genetic variance components ( $\sigma_A^2$  and  $\sigma_D^2$ ), cytoplasm genetic variance ( $\sigma_C^2$ ) and maternal plant genetic variance components ( $\sigma_{A_m}^2$  and  $\sigma_{D_m}^2$ ). Genetic covariance components between endosperm genes and maternal plant genes ( $\sigma_{A, A_m}$  and  $\sigma_{D, D_m}$ ) are estimable. Unbiased estimation for covariance components between two traits can also be obtained by the MINQUE(0/1) method. The random genetic effects in the model are able to be predicted by the Adjusted Unbiased Prediction (AUP) approach with MINQUE (0/1) method. Since genetic variation is analyzed for different generations in several crosses, only sampled seeds (neither individual seed nor seeds of individual plant) from the whole generation are measured for endosperm traits.

**Key words**      Endosperm traits, Genetic model, Estimation of genetic variances and covariances, Prediction of genetic effects