

数量性状基因定位的混合线性模型分析方法

朱 军

(浙江农业大学生物数学研究中心, 杭州 310029)

作物产量、品质和抗逆等重要的农艺性状大多为数量性状基因所控制。近十年来, 随着分子标记检测技术的发展, 有关数量性状基因座位 (Quantitative Trait Loci, 简称 QTL) 的定位方法如区间作图法^[1]、复合区间作图法^[2]、标记回归法^[3]等已能将众多的数量基因定位在相应的连锁图上。一些学者利用这些定位方法已对水稻、小麦、玉米等农作物的许多重要性状进行了 QTL 定位研究。

国内外现有的 QTL 分析方法大都运用简单回归或多元回归的统计方法分析简单的数量遗传模型 (如加性-显性模型), 因此尚不能分析复杂的遗传现象及在不同时空下表达的基因效应, 比如基因型与环境的交互、不同发育阶段的基因表达。因此, QTL 分析方法的滞后, 使作物遗传育种工作者不能有效地对一些重要的农艺性状进行精细定位及遗传效应分析。本文提出了 QTL 基因定位的混合线性模型分析新方法。

1 遗传主效应及 GE 交互效应的 QTL 定位方法

如果在不同环境下实施定位 QTL 的遗传试验, 可以分析 QTL 的遗传主效应及基因型 \times 环境交互效应。第 j 种基因型在环境 h 中第 k 次重复中的表现型值可以用以下线性模型表示:

$$y_{hjk} = \mu + E_h + G_j + GE_{hj} + e_{hjk} \quad (1)$$

其中 y_{hjk} 是第 j 种基因型在第 h 个环境内第 k 次重复中的观察值; μ 是群体平均数, 固定效应; E_h 是环境效应, $E_h \sim (0, \sigma_E^2)$; G_j 是基因型效应, $G_j \sim (0, \sigma_G^2)$; GE_{hj} 是基因型 \times 环境交互效应, $GE_{hj} \sim (0, \sigma_{GE}^2)$; e_{hjk} 是剩余效应, $e_{hjk} \sim (0, \sigma_e^2)$ 。以上公式可以改写为矩阵形式的混合线性模型,

$$y = Xb + U_E e_E + U_G e_G + U_{GE} e_{GE} + e = Xb + \sum_{u=1}^4 U_u e_u = 1 U_u e_u \quad (2)$$

$$\sim N(Xb, V = \sum_{u=1}^4 U_u \sigma_u^2 U_u^T)$$

采用随机效应的预测方法 (如 BLUP 法、LUP 法、AUP 法)^[4] 可以获得基因型效应值 G_j 及基因型 \times 环境互作效应值 GE_{hj} 的无偏预测, 然后可进一步计算 $y_{j(G)} = \mu + G_j$ 和 $y_{hj(GE)} = \mu + GE_{hj}$ 的预测值。

如果采用区间作图法或复合区间作图法分析基因型效应值的数据 $\hat{y}_{j(G)}$,

区间作图: $\hat{y}_{j(G)} = \hat{0}_{(G)} + \hat{0}_{(G)} X_j^* + j_{(G)} \quad (3)$

复合区间作图: $\hat{y}_{j(G)} = \hat{0}_{(G)} + \hat{1}_{(G)} X_j^* + \hat{i}_{(G)} X_{ij} + j_{(G)} \quad (4)$

这时定位的 QTL, 其估算的效应为遗传主效应 (加性、显性效应)。如果采用区间作图法或复合区间作图法分析基因型与第 h 个环境的交互效应的数据 $\hat{y}_{hj(GE)}$,

区间作图: $\hat{y}_{hj(GE)} = \hat{0}_{(GEh)} + \hat{1}_{(GEh)} X_{hj}^* + \hat{h}_{j(GE)} \quad (5)$

本研究受国家自然科学基金的资助。

$$\text{复合区间作图: } \hat{y}_{j(GE)} = \mu_{(GEh)} + \sum_i^* X_{hj}^* X_{hij} + \mu_{hj(GE)} \quad (6)$$

这时定位的 QTL, 其估算的效应为基因型 \times 环境互作效应 (加性 \times 环境互作效应、显性 \times 环境互作效应)。

2 发育数量遗传的 QTL 定位方法

数量性状的最终表现是生物体不同发育阶段基因表达的综合结果。研究基因在特定时刻的表达及其对数量性状的影响, 是发育数量遗传学的一项重要研究内容。

环境 h 的第 k 次重复中第 j 种基因型在 t 时刻 ($t = 1, 2, \dots$) 的表现型值可以用以下线性模型表示,

$$y_{hjk}(t) = \mu_{(t)} + E_{h(t)} + G_{j(t)} + GE_{hj(t)} + e_{hjk(t)} \quad (7)$$

以上公式可以改写为矩阵形式的混合线性模型,

$$y_{(t)} = Xb_{(t)} + U_E e_{E(t)} + U_G e_{G(t)} + U_{GE} e_{GE(t)} + e_{(t)} = Xb_{(t)} + \sum_u^4 U_u e_{u(t)} \quad (8)$$

$$\sim N(Xb_{(t)}, V_{(t)} = \sum_u^4 U_u U_u^T)$$

采用随机效应的预测方法^[4]可以获得 $G_{j(t)}$ 及 $GE_{hj(t)}$ 的无偏预测, 然后进一步计算 $y_{j(G)(t)} = \mu_{(t)} + G_{j(t)}$ 和 $y_{hj(GE)(t)} = \mu_{(t)} + GE_{hj(t)}$ 的预测值。用区间作图法 (3) 或复合区间作图法 (4) 分析 t 时刻基因型效应值的数据 $Y_{T_{j(G)(t)}}$, 定位的 QTL 具有初始时刻至 t 时刻 ($0 \sim t$) 的遗传主效应。采用区间作图法 (5) 或复合区间作图法 (6) 分析基因型与第 h 个环境在 t 时刻的互作效应数据 $Y_{T_{hj(GE)(t)}}$, 可以分析在发育阶段 ($0 \sim t$) 具有基因型 \times 环境互作效应的 QTL。

给定 $t-1$ 时刻的表现型值, t 时刻的条件表现型值是条件随机变量,

$$y_{hjk(t|t-1)} = \mu_{(t|t-1)} + E_{h(t|t-1)} + G_{j(t|t-1)} + GE_{hj(t|t-1)} + e_{hjk(t|t-1)}$$

具有条件变量分布,

$$\begin{aligned} y_{(t|t-1)} &= Xb_{(t|t-1)} + U_E e_{E(t|t-1)} + U_G e_{G(t|t-1)} + U_{GE} e_{GE(t|t-1)} + e_{(t|t-1)} \\ &= Xb_{(t|t-1)} + \sum_{u=1}^4 U_u e_{u(t|t-1)} \end{aligned} \quad (9)$$

$$\sim N(Xb_{(t|t-1)}, V_{(t|t-1)} = \sum_u^4 U_u U_u^T)$$

采用随机效应的预测方法^[5]可以获得条件遗传效应 $G_{j(t|t-1)}$ 及条件互作效应 $GE_{hj(t|t-1)}$ 的无偏预测。用区间作图法 (3) 或复合区间作图法 (4) 分析条件变量 $y_{j(G)(t|t-1)} = \mu_{(t|t-1)} + G_{j(t|t-1)}$ 的预测值, 所定位 QTL 的可以分析特定发育阶段 ($t-1 \sim t$) 的净遗传主效应。采用区间作图法 (5) 或复合区间作图法 (6) 分析条件变量 $y_{hj(GE)(t|t-1)} = \mu_{(t|t-1)} + GE_{hj(t|t-1)}$, 则可以分析 ($t-1 \sim t$) 阶段的 QTL 及净基因型 \times 环境互作效应。

参 考 文 献

- 1 Lander E S, Botstein D. Mapping Mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics*, 1989, 121: 185 ~ 189
- 2 Zeng Z B. Precision mapping of quantitative trait loci. *Genetics*, 1994, 136: 1457 ~ 1468
- 3 Moreno - Gonzalez J. Estimates of marker-associated QTL effects in Monte Carlo backcross generations using multiple regression. *Theor. Appl. Genet.*, 1992, 85: 423 ~ 434
- 4 Zhu J, Weir B S. Diallel analysis for sex-linked and maternal effects. *Theor. Appl. Genet.*, 1996, 92 (1): 1 ~ 9
- 5 Zhu J. Analysis of conditional genetic effects and variance components in developmental genetics. *Genetics*, 1995, 141: 1633 ~ 1639

1997 - 08 - 18 收稿, 1997 - 12 - 26 修回.