

基于随机交配设计的胚乳性状 QTL 定位方法

温永仙 吴为人*

(福建农林大学作物科学学院, 福州 350002. * 联系人, E-mail: wrwu2005@yahoo.com.cn)

摘要 提出了基于随机交配设计和三倍体遗传模型的胚乳性状 QTL 区间定位方法. 其基本思想是, 从一个已知标记基因型信息的作图群体出发, 将群体中的植株(或株系)进行随机交配, 产生由杂种家系组成的胚乳 QTL 定位群体; 对各杂种家系种子的胚乳性状进行混合测定, 得到家系平均值; 利用胚乳性状的家系平均值和亲本植株(或株系)的标记基因型信息对胚乳 QTL 进行定位和效应估计. 用计算机模拟对方法的可行性和有效性进行了检验. 结果表明, 该方法可以准确定位胚乳 QTL, 无偏、有效地估计出胚乳 QTL 的 3 种效应(加性效应、第一显性效应和第二显性效应).

关键词 胚乳性状 QTL 定位 随机交配 混合测定

禾谷类作物的种子主要由胚乳组成, 其中含有丰富的淀粉、蛋白质、脂肪等营养物质, 是人类的主食. 许多胚乳性状都与谷物的产量和品质有关, 因此, 研究胚乳性状的遗传基础对禾谷类作物的育种十分重要^[1~5]. 胚乳是三倍体, 就一对基因($Q-q$)而言, 它有 4 种可能的基因型, 即 QQQ , QQq , Qqq 和 qqq . 因此, 其基因效应模型与二倍体性状不同, 需定义 3 种效应, 即加性效应(a)、第一显性效应(d_1)和第二显性效应(d_2)^[6].

10 多年来, 借助分子标记, 人们对许多禾谷类作物的胚乳性状进行了 QTL 定位研究. 但这些研究并没有注意到胚乳性状与二倍体性状在遗传上的差异, 而是直接利用建立在二倍体遗传模型上的统计方法进行 QTL 定位分析^[7~12]. 近年来, 一些数量遗传学家已经注意到这个问题, 并试图发展出更为合理的胚乳性状 QTL 定位的模型和方法. 胚乳性状 QTL 定位模型与胚乳性状的测定方法有关, 目前提出的方法主要是针对单粒测定的^[13~17]. 但是, 单粒测定的方法存在两个问题: 一是技术上难度较大, 准确性较低, 许多性状难以进行单粒测定; 二是不能进行重复实验, 故而 QTL 与环境互作的效应无法估计. 所以, 在实际研究中, 常用的方法是将一棵植株上所结的多粒种子混合在一起进行测定(即混合测定), 这样得到的性状值是多粒种子的平均值, 亦即该植株的后代家系的平均值. 基于混合测定的胚乳性状 QTL 定位通常采用的是永久性群体(DH 或 RIL 群体). 在 DH 或 RIL 群体中, 每个胚乳 QTL 有两种可能的基因型(QQQ 和 qqq), 频率相等, 这与二倍体组织中的两种基因型(QQ 和 qq)相对应. 因此, 在 DH 或 RIL 群体中, 胚乳

性状的遗传行为与二倍体性状是相似的, 可以用基于二倍体模型的统计方法来分析胚乳 QTL. 但是, 用 DH 或 RIL 群体只能估计胚乳 QTL 的加性效应, 无法估计显性效应. 为了能同时估计胚乳 QTL 的加性效应和显性效应, Xu 等人^[15]提出了利用 F_2 植株标记信息和 $F_{2:3}$ 胚乳性状家系平均值定位胚乳 QTL 的方法. 然而, 虽然该方法是基于三倍体模型的, 但由于其实验设计的遗传特性, QTL 的两种显性效应(d_1 和 d_2)无法分开, 所以其结果与基于二倍体模型是等价的.

本研究对胚乳性状后代家系平均值的遗传模型进行探讨, 提出利用随机交配后代家系平均值进行胚乳 QTL 定位的方法, 并通过计算机模拟对新方法的可行性和有效性进行检验.

1 胚乳性状家系平均值的遗传模型

假设在某个群体中某个控制胚乳性状的 QTL 存在一对等位基因(Q 和 q), 3 种母体基因型 QQ , Qq 和 qq 的频率分别为 p_1 , p_2 和 p_3 . 后代家系可以通过自交或随机交配(植株(或株系) $i \times$ 植株(或株系) j , $i \neq j$)产生. 不论是自交还是随机交配, 后代家系的胚乳性状平均值皆可用下面的遗传模型表示:

$$y = \mu + ax + d_1z_1 + d_2z_2 + \varepsilon. \quad (1)$$

式中, y 为家系平均值, μ 为群体均值; ε 为随机误差, 服从正态分布 $N(0, \sigma_\varepsilon^2)$; x , z_1 和 z_2 为哑变量, 其取值决定于自交植株的基因型(表 1)或随机交配植株的基因型组合(表 2).

由表 1 可以看到, 在自交的情况下, 哑变量 z_1 和 z_2 之间存在线性关系: $z_1 = z_2$. 因此, 式(1)可以简化为:

$y = \mu + ax + dz + \varepsilon$. (2)

式中, $d = d_1 + d_2$, $z = z_1 = z_2$. 显然, 式(2)是一个二倍体遗传模型. 这就是说, 根据自交后代家系平均值, 只能估计出 1 个加性效应和 1 个显性效应, 其中显性效应是一个混合体, d_1 和 d_2 无法区分开. 这清楚地解释了 Xu 等人 [15] 所得结果的原因. 因此, 用自交后代来分析胚乳 QTL 的效应是不理想的. 从表 2 可以看出, 在随机交配的情况下, 3 个哑变量之间都不存在线性关系. 因此, 在随机交配的情况下, 模型(1)是成立的, 胚乳 QTL 的各种效应都可以估计出来. 下面我们讨论在随机交配的情况下, 胚乳 QTL 的分析方法.

表 1 自交情况下的哑变量取值

母体基因型	频率	x	z_1	z_2
QQ	p_1	3/2	0	0
Qq	p_2	0	1/4	1/4
qq	p_3	-3/2	0	0

表 2 随机交配情况下的哑变量取值

基因型组合	频率	x	z_1	z_2	编号(j)		
					F ₂	DH	BC ₁
QQ × QQ	p_1^2	3/2	0	0	1	1	1
QQ × Qq	p_1p_2	1	1/2	0	2		2
QQ × qq	p_1p_3	1/2	1	0	3	2	
Qq × QQ	p_1p_2	1/2	0	1/2	4		3
Qq × Qq	p_2^2	0	1/4	1/4	5		4
Qq × qq	p_2p_3	-1/2	1/2	0	6		
qq × QQ	p_1p_3	-1/2	0	1	7	3	
qq × Qq	p_2p_3	-1	0	1/2	8		
qq × qq	p_3^2	-3/2	0	0	9	4	

2 基于随机交配后代家系平均值的 QTL 区间定位

在任意实验群体中, 假设已知植株(或株系)的标

记基因型. 植株(或株系)间进行随机交配, 产生后代家系. 由式(1)可知, 胚乳 QTL 区间定位的遗传模型为:

$$y_i = \mu + ax_i + d_1z_{1i} + d_2z_{2i} + \varepsilon_i. \quad (3)$$

式中, i 表示第 i 个家系; 哑变量的值是未知的, 但可以根据 QTL 两侧的标记推得其取特定值的条件概率. 表 3 给出了 F₂ ($p_1 = p_3 = 1/4$, $p_2 = 1/2$)、BC₁ ($p_1 = p_2 = 1/2$, $p_3 = 0$)和 DH ($p_1 = p_3 = 1/2$, $p_2 = 0$)几种常见群体中二倍体植株(或株系)的 QTL 基因型的条件概率, 由此可以得到任一随机交配组合中各种可能的 QTL 基因型组合的条件概率, 它等于两杂交植株(或株系)的 QTL 基因型的条件概率的乘积. 例如, 在 DH 群体中, 任意 2 个株系杂交都存在 4 种可能的 QTL 基因型组合: QQ × QQ, QQ × qq, qq × QQ 和 qq × qq. 对于双侧标记基因型为 $M_1M_1M_2M_2$ 和 $M_1M_1m_2m_2$ 的两个株系杂交, QQ × QQ 的条件概率为 $s_1s_2/s \times s_1r_2/r = s_1^2r_2s_2/rs$, QQ × qq 的条件概率为 $s_1s_2/s \times r_1s_2/r = r_1s_1s_2^2/rs$, 其余类推. 根据表 2, 就可得到各哑变量取特定值的条件概率.

于是, y_i 的分布可用一个混合正态分布来描述:

$$f(y_i) = \sum_{j=1}^m p_{ij} f_j(y_i). \quad (4)$$

式中, m 是可能的 QTL 基因型组合(子分布)数, 等于 4 (对于 DH 或 BC₁ 群体的情况)或 9 (对于 F₂ 群体的情况); p_{ij} 是第 i 家系第 j 种 QTL 基因型组合的条件概率; $f_j(y_i)$ 为第 j 个子分布的概率密度函数. 每一个子分布都是一个正态分布 $N(\mu_j, \sigma_\varepsilon^2)$, 其中 $\mu_j = ax + d_1z_1 + d_2z_2$ (表 2). 例如, 对于 F₂ 群体的情况, 第 2 个子分布(组合 QQ × Qq) $\mu_2 = a + d_1/2$.

从式(4)出发, 可以写出似然函数:

$$L(\Omega) = \prod_{i=1}^n \sum_{j=1}^m p_{ij} f_j(y_i). \quad (5)$$

表 3 3 种群体中二倍体植株 QTL 基因型依相邻标记基因型的条件概率^{a)}

相邻标记 基因型	F ₂			DH		BC ₁	
	QQ	Qq	qq	QQ	qq	QQ	Qq
$M_1M_1M_2M_2$	$s_1^2s_2^2/s^2$	$2r_1s_1r_2s_2/s^2$	$r_1^2r_2^2/s^2$	s_1s_2/s	r_1r_2/s	s_1s_2/s	r_1r_2/s
$M_1M_1M_2m_2$	$s_1^2r_2s_2/rs$	$r_1s_1(s_2^2+r_2^2)/rs$	$r_1^2r_2s_2/rs$			s_1r_2/r	r_1s_2/r
$M_1M_1m_2m_2$	$s_1^2r_2^2/r^2$	$2r_1s_1r_2s_2/r^2$	$r_1^2s_2^2/r^2$	s_1r_2/r	r_1s_2/r		
$M_1m_1M_2M_2$	$r_1s_1s_2^2/rs$	$r_2s_2(s_1^2+r_1^2)/rs$	$r_1s_1r_2^2/rs$			r_1s_2/r	s_1r_2/r
$M_1m_1M_2m_2$	$2r_1s_1r_2s_2/(s^2+r^2)$	$(s_1^2+r_1^2)(s_2^2+r_2^2)/(s^2+r^2)$	$2r_1s_1r_2s_2/(s^2+r^2)$			r_1r_2/s	s_1s_2/s
$M_1m_1m_2m_2$	$r_1s_1r_2^2/rs$	$r_2s_2(s_1^2+r_1^2)/rs$	$r_1s_1s_2^2/rs$				
$m_1m_1M_2M_2$	$r_1^2s_2^2/r^2$	$2r_1s_1r_2s_2/r^2$	$s_1^2r_2^2/r^2$	r_1s_2/r	s_1r_2/r		
$m_1m_1M_2m_2$	$r_1^2r_2s_2/rs$	$r_1s_1(s_2^2+r_2^2)/rs$	$s_1^2r_2s_2/rs$				
$m_1m_1m_2m_2$	$r_1^2r_2^2/s^2$	$2r_1s_1r_2s_2/s^2$	$s_1^2s_2^2/s^2$	r_1r_2/s	s_1s_2/s		

a) r_1, r_2 和 r 分别是左侧标记与 QTL 之间、QTL 与右侧标记之间以及两标记之间的重组率, $s = 1 - r, s_1 = 1 - r_1, s_2 = 1 - r_2$

www.scichina.com

1667

其中, n 为随机交配后代家系总数; Ω 为未知参数向量 $(\mu, a, d_1, d_2, r_1, \sigma_e^2)^T$, 这里 T 表示转置. 采用 EM 算法^[18], 可以求得各参数的极大似然估值.

E 步: 计算矩阵 $\hat{P}$, 它是一个包含元素 $\{P_{ij}\}$ 的 $n \times m$ 矩阵, P_{ij} 为第 i 个家系第 j 种 QTL 基因型组合的后验概率, 定义为

$$P_{ij} = \frac{p_{ij} f_j(y_i)}{\sum_{j=1}^m p_{ij} f_j(y_i)}. \quad (6)$$

M 步: 在矩阵 $\hat{P}$ 的条件下, 解下列方程:

$$Y^T \hat{P} E_k = I^T \hat{P} Q_k X \hat{B}, \quad (k = 1, 2, 3, 4); \quad (7)$$

$$\hat{\sigma}_e^2 = \frac{1}{n} [Y^T Y - 2 Y^T \hat{P} X \hat{B} + I^T \hat{P} A X \hat{B}]. \quad (8)$$

其中 Y 是由 y_i 组成的 $n \times 1$ 向量; $\hat{B}$ 是由参数 μ, a, d_1 和 d_2 的估计值所组成的 4×1 向量; I 是所有元素为 1 的 $n \times 1$ 向量; $E_1 \sim E_4$ 是 $m \times 1$ 向量, 其中 E_1 所有元素皆为 1, E_2, E_3 和 E_4 分别对应于表 2 中的 x, z_1 和 z_2 列; Q_k 是一个 4×4 对角矩阵, 其对角线上的元素是向量 E_k 上的元素; A 是一个 $m \times m$ 对角阵, 其对角线上的元素是 $\mu_1 \sim \mu_m$ 的估计值; $X = (E_1, E_2, E_3, E_4)$. M 步结束后, 返回到 E 步. 迭代循环, 直至收敛.

QTL 的显著性可以用似然比 (LOD 值) 来检验:

$$\text{LOD} = \lg \left[\frac{L(\hat{\Omega})}{L(\tilde{\Omega})} \right]. \quad (9)$$

这里, $L(\tilde{\Omega})$ 和 $L(\hat{\Omega})$ 分别为在零假设 ($H_0: a = d_1 = d_2 = 0$) 和备择假设 ($H_1: a, d_1$ 和 d_2 不全为零) 情况下的极大似然函数值. 可以通过排列检验^[19] 或模拟抽样来估算合适的显著阈值.

3 模拟研究

为了检验上面提出的方法, 我们进行了计算机模拟. 假设在一条长为 100 cM 的染色体上有 11 个标记等距排列, 1 个 QTL 位于 55 cM 处, 其加性效应为 1, 第一显性效应为 0.8, 第二显性效应为 0.5, 群体均值为 10. 亲代群体以 F_2 和 DH 群体为例, 假设群体大小分别为 200 和 100. 随机交配的方法是: 将亲代群体随机平分成两组 (即母本组和父本组), 随机打乱各组中个体的排列顺序, 然后将两组相应序号的个体进行交配. 该过程可多次重复, 直至获得足够多的杂种家系. 考虑了不同情况, 每种情况重复模拟 100 次. 显著阈值通过 5000 次零假设 (即不存在 QTL 的情况) 模拟实验进行估计. 表 4 给出了部分模拟结果.

从表 4 可以看到, 不论用于随机交配的是 F_2 群体还是 DH 群体, 都能有效地检测和准确地定位 QTL. 对 QTL 效应的估计也是无偏的, 估值的准确性随着 QTL 遗传率和样本容量 (杂种家系数以及每个杂种家系测定的种子数) 的增大而提高. 对加性效应的估计要好于显性效应, 这可能与所假设的显性效应值相对较小有关. 相对而言, DH 群体随机交配对 QTL 效应的估计效果要优于 F_2 群体, 这可能是因为 F_2 随机交配产生的杂种家系遗传上不是同质的, 家系内存在分离, 误差较大. 因此, 为了减小误差, F_2 随机交配产生的杂种家系必须测定较多的种子.

4 讨论

由于技术上的限制, 许多胚乳性状难以进行单粒测定, 必须多粒混合测定. 因此, 发展基于多粒混合测定值 (亦即家系平均值) 的胚乳 QTL 定位方法是十分必要的. 为了估计胚乳 QTL 效应, 必须发展基于三倍体遗传模型的 QTL 定位方法. 然而, 从前面的已经看到, 胚乳 QTL 效应的估计与实验设计有关. 采用自交的实验设计 (即基于自交后代家系平均值) 无法估计出胚乳 QTL 的所有效应 (加性效应、第一显性效应、第二显性效应), 因为这时三倍体模型实际上是和二倍体模型等价的. 本研究提出的随机交配的实验设计 (即基于杂种家系平均值) 能够提供更多的遗传信息, 从而能够估计出胚乳 QTL 的所有效应. 此外, 由于群体中植株 (或株系) 间所有可能的杂交组合数非常大 (如一个含 100 个个体的群体, 所有可能的杂交组合数为 4950), 远远超出实际研究的需要, 而且每个杂种家系的标记基因型都可以从其亲本植株 (或株系) 推知, 因此, 随机交配实验设计还有一个突出优点, 就是可以产生很大的杂种家系群体而无需增加分子标记检测的费用, 因而可以有效提高 QTL 分析的统计功效. 从模拟结果看, 利用永久性群体进行随机交配要比利用 F_2 群体好, 对 QTL 效应的估计更准确. 从实际应用角度看, 永久性群体也更为好用, 因为每个株系都可以提供大量的植株用于杂交, 而且株系间杂交可以反复进行, 从而可以进行重复实验. 另外, 目前在主要禾谷类作物中都已构建了许多永久性群体和相应的分子标记图谱. 因此, 利用永久性群体进行随机交配的实验设计和相应的胚乳 QTL 分析方法具有很大的应用价值. 当然, 随机交配实验设计的杂交工作量大, 这对于自花授粉作物来说有较大难度, 有赖于杂交技术的发展, 但对于异花

表 4 利用 F₂ 和 DH 群体随机交配对胚乳 QTL 的定位和效应估计

遗传率 (%) ^{a)}	杂种家 系数	测定种 子数	群体	QTL 位置 ^{b)} /cM	QTL 效应 ^{b)}			统计功效(%)	
					<i>a</i>	<i>d</i> ₁	<i>d</i> ₂		
5	250	5	F ₂	55.29(5.20)	1.02(0.15)	0.76(0.62)	0.55(0.69)	97	
			DH	55.46(3.63)	1.00(0.17)	0.84(0.39)	0.52(0.36)	100	
		10	F ₂	55.00(2.60)	1.02(0.12)	0.76(0.44)	0.48(0.43)	100	
			DH	55.12(2.68)	0.99(0.10)	0.78(0.25)	0.48(0.30)	100	
	500	5	F ₂	55.40(2.47)	1.00(0.12)	0.80(0.47)	0.54(0.45)	100	
			DH	54.79(2.26)	1.02(0.09)	0.75(0.29)	0.50(0.27)	100	
		10	F ₂	55.06(1.69)	0.99(0.10)	0.82(0.32)	0.53(0.32)	100	
			DH	54.65(1.73)	1.00(0.09)	0.80(0.20)	0.51(0.19)	100	
	10	250	5	F ₂	54.56(3.13)	0.99(0.14)	0.83(0.44)	0.48(0.46)	100
				DH	55.02(2.52)	0.99(0.09)	0.81(0.30)	0.49(0.32)	100
			10	F ₂	54.89(1.91)	0.99(0.09)	0.82(0.25)	0.47(0.30)	100
				DH	54.65(2.58)	0.99(0.07)	0.79(0.18)	0.49(0.21)	100
500		5	F ₂	54.89(2.08)	1.00(0.07)	0.82(0.30)	0.48(0.27)	100	
			DH	54.76(1.86)	1.00(0.07)	0.80(0.21)	0.53(0.19)	100	
		10	F ₂	55.08(1.08)	1.01(0.07)	0.80(0.20)	0.50(0.19)	100	
			DH	54.61(1.58)	0.99(0.05)	0.78(0.13)	0.48(0.13)	100	
20		250	5	F ₂	54.86(1.76)	1.00(0.09)	0.77(0.26)	0.45(0.29)	100
				DH	55.14(2.28)	1.00(0.06)	0.79(0.20)	0.50(0.18)	100
			10	F ₂	54.88(1.63)	0.99(0.06)	0.80(0.20)	0.52(0.21)	100
				DH	54.69(1.37)	1.00(0.03)	0.80(0.10)	0.50(0.10)	100
	500	5	F ₂	55.12(1.00)	1.00(0.04)	0.78(0.14)	0.52(0.15)	100	
			DH	54.79(1.34)	1.00(0.05)	0.81(0.13)	0.51(0.11)	100	
		10	F ₂	55.08(1.08)	0.99(0.03)	0.79(0.10)	0.49(0.11)	100	
			DH	55.07(1.21)	0.99(0.04)	0.80(0.10)	0.49(0.08)	100	
	真值				55	1	0.8	0.5	

a) 指随机交配后代群体中 QTL 的遗传率; b) 100 次模拟的平均值, 括号中为标准差

授粉作物(如玉米)来说是容易做到的. 因此, 本方法特别适用于异花授粉作物的研究.

从模拟研究结果可以看到, 增加随机交配后代家系数和每家系胚乳数都能提高 QTL 定位的功效和效应估计的准确性. 但从统计上看, 家系数和胚乳数的作用是不同的. 增加胚乳数主要只是减小环境误差, 而增加家系数除了能够减小环境误差外, 还能够减小基因型抽样的误差, 提高对连锁座位间的分辨率. 所以, 增加家系数具有更多的益处. 由于随机交配实验设计可以产生很大的杂种家系群体而无需增加分子标记检测费用, 因此当杂交不太困难时(如异花授粉作物), 宜尽量建立大的杂种家系群体.

F_{2:3} 家系群体是比较常用的 QTL 定位实验群体. F_{2:3} 家系内存在基因型分离, 因此家系内误差服从混合分布. Zhang 和 Xu^[20]对 F_{2:3} 的家系内误差分布进行了分析, 将其整合到基于 F_{2:3} 群体的 QTL 定位模型中. 模拟研究显示, 考虑家系内误差分布的信息可以提高 QTL 检测的功效和效应估计的准确性. 本研究提出的利用 F₂ 群体进行随机交配的实验设计, 后代家

系内也存在基因型分离. 因此, 利用家系内误差分布的信息, 也可能提高胚乳 QTL 分析的效果, 是值得进一步研究的.

种子相对于母体植株而言是下一代, 但它是生长在母体植株上的, 因此胚乳性状除了受自身基因型的控制之外, 还可能受母体植株基因型的影响^[17]. 这意味着一个控制胚乳性状的 QTL 除了在胚乳中具有效应(称为胚乳效应)之外, 还可能在母体植株中产生效应(称为母体效应). 本研究讨论的方法只涉及对 QTL 胚乳效应的估计. 如果存在母体效应, 就应该在遗传模型中将母体效应考虑进去, 否则估计出来的胚乳效应就会与母体效应相混杂. 但是, 从理论上可以证明, 基于家系平均值(混合测定)的胚乳 QTL 分析方法是无法将胚乳效应和母体效应完全区分开的. 事实上, 若假设存在母体加性效应(*a_m*)和显性效应(*d_m*), 则式(1)可以拓展为:

$$y = \mu + ax + d_1z_1 + d_2z_2 + a_mx_m + d_mz_m + \varepsilon. \quad (10)$$

自交的实验设计连胚乳效应都无法全部估计出来, 故不必考虑, 只需考虑随机交配的情形. 可以推得,

母体效应的哑变量 x_m 和 z_m 与表2中9种交配组合对应的取值分别为(1, 1, 1, 0, 0, 0, -1, -1, -1)和(0, 0, 0, 1, 1, 1, 0, 0, 0)。将它们与表2中其他哑变量的取值进行比较,可以发现存在一个线性关系: $3x_m = 2x + 2z_1 - 2z_2$ 。这就是说,式(10)中有一个自变量是多余的,因此不可能估计出所有的效应。所以,在基于家系平均值(混合测定)进行胚乳QTL分析时,必须假设不存在母体效应。要估计母体效应,则必须采用对胚乳性状单粒测定的方法^[17]。因此,对于可以进行单粒测定的胚乳性状,可以考虑采用基于单粒测定的QTL分析方法,以便检验是否存在母体效应。

最后,有必要指出,虽然本研究只给出基于随机交配实验设计的胚乳QTL区间定位方法,但它很容易拓展成复合区间定位法^[21]和多区间定位法^[22]。限于篇幅,不作讨论。

致谢 本工作为福建省科技重大专项(批准号:2004NZ01-2)资助项目。

参 考 文 献

- Sadimantara G R, Abe T, Sasahar T. Genetic analysis of high molecular weight protein in rice (*Oriza sativa* L) endosperm. *Crop Sci*, 1997, 37: 1177—1180
- Mazur B, Krebbe E, Tingey S. Gene discovery and product development for grain quality traits. *Science*, 1999, 285: 372—375[DOI]
- Tan Y F, Li J X, Yu S B, et al. The three important traits for cooking and eating quality of rice grains are controlled by a single locus in an elite rice hybrid, Shanyou 63. *Theor Appl Genet*, 1999, 99: 642—648[DOI]
- Wang X L, Larkins B A. Genetic analysis of amino acid accumulation in opaque-2 maize endosperm. *Plant Physiol*, 2001, 125: 1766—1777[DOI]
- Lou X Y, Zhu J. Analysis of genetic effects of major genes and polygenes on quantitative traits. . Genetic models for seed traits of crops. *Theor Appl Genet*, 2002, 105: 964—971[DOI]
- Mo H D. Genetic expression for endosperm traits. *Proceedings of the 2nd International Conference on Quantitative Genetics*. Sinauer Associates, Sunderland, MA, 1987. 478—487
- Sourdille P, Perretant M R, Charme G, et al. Linkage between RFLP markers and genes affecting kernel hardness in wheat. *Theor Appl Genet*, 1996, 93: 580—586
- Parker G D, Chalmers K J, Rathjen A J, et al. Mapping loci associated with flour colour in wheat (*Triticum aestivum* L). *Theor Appl Genet*, 1998, 97: 238—245 [DOI]
- Araki E, Miura H, Sawada S. Identification of genetic loci affecting amylose content and agronomic traits on chromosome 4A of wheat. *Theor Appl Genet*, 1999, 98: 977—984[DOI]
- Sene M, Thevenot C, Hoffmann D, et al. QTL for grain dry milling properties, composition and vitreousness in maize recombinant inbred lines. *Theor Appl Genet*, 2001, 102: 591—599[DOI]
- Wang X L, Woo Y M, Kim C S, et al. Quantitative trait locus mapping of loci influencing elongation factor 1 alpha content in maize endosperm. *Plant Physiol*, 2001, 125: 1271—1282[DOI]
- Tan Y F, Sun M, Xing Y Z, et al. Mapping quantitative trait loci for milling quality, protein content and color characteristics of rice using a recombinant inbred line population derived from an elite rice hybrid. *Theor Appl Genet*, 2001, 103: 1037—1045[DOI]
- Wu R L, Lou X Y, Ma C X, et al. An improved genetic model generates high-resolution mapping of QTL for protein quality in maize endosperm. *Proc Natl Acad Sci USA*, 2002, 99: 11281—11286[DOI]
- Wu R L, Ma C X, Gallo-Meagher M, et al. Statistical methods for dissecting triploid endosperm traits using molecular markers: an autogamous model. *Genetics*, 2002, 162: 875—892
- Xu C, He X, Xu S. Mapping quantitative trait loci underlying triploid endosperm traits. *Heredity*, 2003, 90: 228—235[DOI]
- Kao C H. Multiple-interval mapping for quantitative trait loci controlling endosperm traits. *Genetics*, 2004, 167(4): 1987—2002 [DOI]
- 胡治球, 徐辰武. 谷物胚乳性状数量基因定位新方法. *科学通报*, 2005, 50(14): 1480—1485
- Dempster A P, Laird N M, Rubin D B. Maximum likelihood from incomplete data via EM algorithm. *J R Stat Soc Ser B*, 1977, 39: 1—38
- Churchill G A, Doerge R W. Empirical threshold values for quantitative trait mapping. *Genetics*, 1994, 138: 963—971
- Zhang Y M, Xu S Z. Mapping quantitative trait loci in F_2 incorporating phenotypes of F_3 progeny. *Genetics*, 2004, 166: 1981—1993[DOI]
- Zeng Z B. Precision mapping of quantitative trait loci. *Genetics*, 1994, 136: 1457—1468
- Kao C H, Zeng Z B, Teasdale R D. Multiple interval mapping for quantitative trait loci. *Genetics*, 1999, 152: 1203—1216

(2006-04-02 收稿, 2006-06-29 接受)