

利用 P_1 、 P_2 和 DH 或 RIL 群体联合分离分析的拓展

章元明, 盖钧镒, 王永军

(南京农业大学大豆研究所; 农业部国家大豆改良中心, 南京 210095)

摘要: QTL 定位常报道主基因数多于 3 的情形。然而, 利用基因型杂合的分离群体拓展 3 对主基因+多基因混合遗传模型非常困难, 而 DH 或 RIL 群体是遗传分析的很好群体且拓展 3 对主基因+多基因模型相对容易些。在文献 1~4 的基础上, 拓展利用亲本和 DH 或 RIL 群体的 2 对连锁主基因、2 对连锁主基因+多基因、3 对主基因和 3 对主基因+多基因 4 类遗传模型。通过大豆科丰 1 号×1138 - 2 构成的 RIL 群体及其亲本研究了大豆开花期的遗传规律。

关键词: DH 群体; RIL 群体; 主基因+多基因混合遗传; 大豆; 开花期

中图分类号: Q348

文献标识码: A

文章编号: 0253-9772(2001)05-0467-04

An Expansion of Joint Segregation Analysis of Quantitative Trait for Using P_1 , P_2 and DH or RIL Populations

ZHANG Yuan-ming, GAI Jun-yi, WANG Yong-jun

(Soybean Research Institute, Nanjing Agricultural University, National Center of Soybean Improvement, Ministry of Agriculture, Nanjing 210095, China)

Abstract: More than three major genes controlling quantitative trait are often obtained in QTL mapping. However, it is very difficult to develop joint segregation analysis with three major genes but for DH or RIL population. Therefore, an expansion of the inheritance model, such as linkage between two major genes, three major genes, three major genes plus polygenes, was developed in this paper. Finally, an example of the inheritance of flowering date of soybean for using P_1 , P_2 and RIL populations derived from a cross between Keifeng 1 and Nannong 1138 - 2 was used to explain the procedures.

Key words: DH population; RIL population; major gene plus polygenes mixed inheritance; soybean; flowering date

在 QTL 定位前通过数量性状表型分布的分离分析了解其基因对数是十分有益的^[1]。QTL 作图显示控制数量性状的基因座数为 1~7 的占 80% 以上, 平均基因座数约为 3~4^[2]。因此, 很有必要拓展 3 对主基因及 3 对主基因+多基因模型。然而, 利用基因型杂合的分离群体拓展 3 对主基因以及 3 对主基因+多基因混合模型是十分困难的^[1]。近年来, 常见利用 DH 或 RIL 群体进行遗传分析的报道, 因此本文就利用 DH 或 RIL 群体对遗传模型进行拓展。

1998 年胡中立等首先报道了利用 DH 群体的 1~2 对主基因+多基因混合遗传模型^[3], 2 年后章元明等拓展了 2 对连锁主基因+多基因混合遗传模型^[4]。提高数量性状分离分析的精度, 主要有 3 种途径^[1]: 利用家系平均数; 采用适当的试验设计来控制试验误差; 采用亲本和 DH 或 RIL 群体的联合分析^[1]。然而, 利用亲本和 DH 或 RIL 群体的联合分析未涉及主基因间的连锁以及 3 对主基因。为此, 本文在此方面进行拓展。

收稿日期: 2000-09-29; 修回日期: 2001-02-26

基金项目: 国家 973 项目、重庆市科委应用基础研究项目

作者简介: 章元明(出生年 1965-), 男, 籍贯重庆永川, 硕士, 副教授, 从事数量遗传与生物统计教学与研究工作。

1 遗传模型及有关遗传参数的估计

根据文献[1]的假定, P_1 和 P_2 群体均表现为单一正态分布, DH 或 RIL 群体表现为 k 个主基因型所对应的 k 个正态分布的混合, 即

$$x_{1i} \sim N(\mu_{1i}, \sigma^2) \quad x_{2i} \sim N(\mu_{2i}, \sigma^2) \quad x_{3i} \sim \sum_{i=1}^k \pi_i N(\mu_{3i}, \sigma_{3i}^2)$$

由此, 建立样本似然函数,

$$L(Y|\theta) = \prod_{i=1}^{n_1} f(x_{1i}; \mu_1, \sigma^2) \prod_{i=1}^{n_2} f(x_{2i}; \mu_2, \sigma^2) \prod_{i=1}^{n_3} \sum_{i=1}^k \pi_i f(x_{3i}; \mu_{3i}, \sigma_{3i}^2)$$

其中, 其分布参数与遗传参数间的关系因遗传模型的不同而异。因 2 对主基因连锁遗传模型和 3 对主基因遗传模型目前尚未建立, 所以下面将建立这 2

大类遗传模型。

1.1 2 对主基因的连锁遗传模型

当两主基因的重组率 $r=0.5$ 时表现为独立遗传。对于连锁遗传模型, 文献[4]已经论及, 但未得到 RIL 群体重组率的估计公式; 对于 P_1 、 P_2 、DH 或 RIL 群体的联合分析, 文献[5]已经论及, 但未涉及连锁遗传模型。对于 2 对连锁主基因和 2 对连锁主基因+多基因模型, 与文献[5]相比较, 主要在于前者多出一个重组率参数, 其余均与文献[5]一致。因此, 这里只给出重组率参数的估计方法, 详见表 1。2 对独立和连锁主基因模型记号分别用 B-1-X 和 B-2-X 表示; 2 对独立和连锁主基因+多基因混合遗传模型记号分别用 E-1-X 和 E-2-X 表示。

表 1 利用 P_1 、 P_2 、DH 或 RIL 群体鉴定数量性状 2 对连锁主基因遗传模型的重组率
Table 1 The estimate of recombination value for identifying two linkage major genes plus polygenes inheritance model using P_1 , P_2 , DH or RIL populations

模型 Model	DH		RIL	
	成分分布比例 Proportion of Component Distributions	重组率 Recombination value	成分分布比例 Proportion of Component Distributions	重组率 Recombination value
加性-上位性 Additive-epistatic	$\frac{1-r}{2}, \frac{r}{2}, \frac{r}{2}, \frac{1-r}{2}$	$\Sigma(w_{21} + w_{31})/n$	$\frac{1-R}{2}, \frac{R}{2}, \frac{R}{2}, \frac{1-R}{2}$	$\Sigma(w_{21} + w_{31})/\Sigma 2(w_{11} + w_{13})$
等加性 Additive	$\frac{1-r}{2}, \frac{r}{2}, \frac{r}{2}, \frac{1-r}{2}$	$\Sigma(w_{21} + w_{31})/n$	$\frac{1-R}{2}, \frac{R}{2}, \frac{R}{2}, \frac{1-R}{2}$	$\Sigma(w_{21} + w_{31})/\Sigma 2(w_{11} + w_{13})$
等加性 Equal additive	$\frac{1-r}{2}, r, \frac{1-r}{2}$	$\Sigma w_{21}/n$	$\frac{1-R}{2}, R, \frac{1-R}{2}$	$\Sigma w_{21}/\Sigma 2(w_{11} + w_{13})$
显性上位 Epistatic dominance	$\frac{1}{2}, \frac{r}{2}, \frac{1-r}{2}$	$\Sigma w_{21}/\Sigma(w_{21} + w_{31})$	$\frac{1}{2}, \frac{R}{2}, \frac{1-R}{2}$	$\Sigma w_{21}/\Sigma 2w_{13}$
隐性上位 Epistatic recessive	$\frac{1-r}{2}, \frac{r}{2}, \frac{1}{2}$	$\Sigma w_{21}/\Sigma(w_{11} + w_{21})$	$\frac{1-R}{2}, \frac{R}{2}, \frac{1}{2}$	$\Sigma w_{21}/\Sigma 2w_{11}$
累加作用 Additive effect	$\frac{1-r}{2}, r, \frac{1-r}{2}$	$\Sigma w_{21}/n$	$\frac{1-R}{2}, R, \frac{1-R}{2}$	$\Sigma w_{21}/\Sigma 2(w_{11} + w_{13})$
互补作用 Complementary	$\frac{1-r}{2}, \frac{1+r}{2}$	$\Sigma(w_{21} - w_{11})/n$	$\frac{1-R}{2}, \frac{1+R}{2}$	$\Sigma(w_{21} - w_{11})/\Sigma 4w_{11}$
重叠作用 Duplicate effect	$\frac{1+r}{2}, \frac{1-r}{2}$	$\Sigma(w_{11} - w_{21})/n$	$\frac{1+R}{2}, \frac{1-R}{2}$	$\Sigma(w_{11} - w_{21})/\Sigma 4w_{12}$
抑制作用 Inhibiting effect	$\frac{2-r}{2}, \frac{r}{2}$	$2\Sigma w_{21}/n$	$\frac{2-R}{2}, \frac{R}{2}$	$\Sigma w_{21}/\Sigma(w_{11} - w_{12})$

* $R = 2r/(1+2r)$

1.2 3 对主基因和 3 对主基因+多基因混合遗传模型

为初步了解数量性状是否由 3 对主基因控制, 本文初步构造了 3 对主基因的加性-上位性 ($d_a, d_b, d_c, i_{aa}, i_{bb}, i_{cc}$)、加性 (d_a, d_b, d_c)、等加性 1 ($d = d_a = d_b = d_c$) 和等加性 2 ($d_1 = d_a = d_b, d_2 = d_c$) 模

型, 详见表 2。

1.3 遗传模型的选择与遗传参数的估计

遗传模型的选择按文献[1]的 AIC 准则和适合性检验进行。一阶遗传参数是利用一阶分布参数与一阶遗传参数间的关系通过最小二乘法进行估计; 二阶遗传参数按文献[4]的方法进行估计。笔者利

用 Turbo C++ 语言编写了样本似然函数中分布参数极大似然估计 IECM 算法的计算程序,程序名为

MGDH. EXE。欢迎读者来函联系。

表 2 3 对主基因遗传模型

Table 2 Three major genes models

主基因 major genes	多基因不存在 without polygenes			多基因存在 with polygenes		
	代号 Code	一阶参数 First order Parameter	二阶参数 2nd order Parameter	代号 Code	一阶参数 First order Parameter	二阶参数 2nd order Parameter
全模型 All	—	—	—	G - 0	$\mu_1, \mu_2, \mu_{31} \sim \mu_{438}$	σ^2, σ_{31}^2
加性-上位性 Additive epistatic	F - 1	$m, d_a, d_b, d_c, i_{ab}, i_{ac}, i_{bc}, i_{abc}$	σ^2	G - 1	$m, d_a, d_b, d_c, i_{ab}, i_{ac}, i_{bc}, i_{abc}, [d]$	σ^2, σ_{31}^2
加性 Additive	F - 2	m, d_a, d_b, d_c	σ^2	G - 2	$m, d_a, d_b, d_c, [d]$	σ^2, σ_{31}^2
全等加性 Equal additive	F - 3	m, d	σ^2	G - 3	$m, d, [d]$	σ^2, σ_{31}^2
部分等加性 Partly equal additive	F - 4	$m, d_1 = d_a = d_b, d_2 = d_c$	σ^2	G - 4	$m, d_1 = d_a = d_b, d_2 = d_c, [d]$	σ^2, σ_{31}^2

2 实例分析

用大豆科丰 1 号 (P_1) × 南农 1138 - 2 (P_2) 形成的 F_2 群体经一粒传所构成的 RIL 群体及其亲本的花期为例进行分析。RIL 群体共 206 个家系,为使群体不出现偏分离而通过 SPSC 法(另文

报道)剔除个别家系,用作分析家系数 184 个。随机区组设计,重复 8 次。方差分析结果见表 3。从表 3 可知,亲本间差异不显著,RIL 群体家系间差异极显著,可作遗传分析。RIL 群体开花期的分布见图 1,可见开花期呈多峰分布,有主基因的存在。

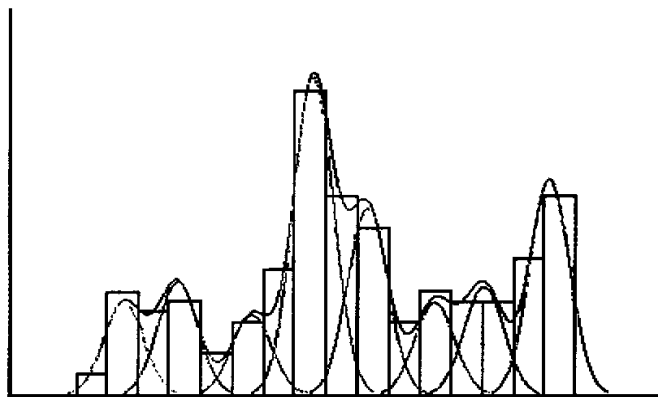


图 1 大豆科丰 1 号 × 南农 1138 - 2 形成的 RIL 群体开花期的分离分析结果

Fig. 1 The results of segregating analysis of date to flower of a soybean RIL population derived from a cross between Keifeng 1 and Nannong 1138 - 2

通过 IECM 算法,得到 52 种遗传模型的 AIC 值(见表 4)。从表 4 可知,G - 1 模型的 AIC 值最低,其适合性检验(图 1)也较好。因此,G - 1 模型是较优的遗传模型。根据该模型,估计出开花期的

遗传参数(表 5)。从表 5 可知,两个主基因使花期提早 4 天和 1 天,不过这两主基因间的互作却使花期延后 0.46 天;另一主基因的效应较小,该主基因与最大效应主基因的互作使花期延后 1.8 天;三个

主基因间的互作使花期延后 3.4 天;多基因的加性效应却很大,使花期提早 4.5 天。

表 3 方差分析表

Table 3 Analysis of variances

变异来源		df	P ₁		P ₂		RIL		
Sources of variation	SS		F	SS	F	df	SS	F	
区组	Block	7	7.0	1.75	1.4375	0.21	7	13.1742	1.10
家系	Family	1	1.0	1.00	0.5625	0.57	205	60820.0273	174.03
误差	Error	7	7.0		6.9375		1435	2446.3259	
总	Total	15	15.0		8.9375		1647	63279.5273	

表 4 不同遗传模型的 AIC 值

Table 4 Akaike's information criterion (AIC) values under various genetic models

模型	AIC	模型	AIC	模型	AIC	模型	AIC
Model		Model		Model		Model	
A-1	1207.77	B-2-4	1216.11	E-1-3	1172.31	E-2-6	1151.48
B-1-1	1172.30	B-2-5	1162.32	E-1-4	1149.47	E-2-7	1202.45
B-1-2	1172.65	B-2-6	1172.29	E-1-5	1149.47	E-2-8	1202.45
B-1-3	1170.74	B-2-7	1208.00	E-1-6	1195.22	E-2-9	1198.71
B-1-4	1216.41	B-2-8	1208.00	E-1-7	1204.62	F-1	1151.99
B-1-5	1209.50	B-2-9	1216.68	E-1-8	1204.62	F-2	1205.02
B-1-6	1170.29	C	1200.44	E-1-9	1200.48	F-3	1174.46
B-1-7	1218.88	C-1	1214.79	E-2-0	1202.09	F-4	1171.90
B-1-8	1209.58	D	1196.10	E-2-1	1153.25	G	1151.10
B-1-9	1218.94	D-1	1209.38	E-2-2	1213.39	G-1	1148.42
B-2-1	1174.30	E-1-0	1200.09	E-2-3	1174.31	G-2	1218.55
B-2-2	1174.68	E-1-1	1151.71	E-2-4	1151.26	G-3	1216.57
B-2-3	1172.73	E-1-2	1211.26	E-2-5	1198.11	G-4	1174.32

表 5 大豆开花期遗传参数估计值

Table 5 The estimates of genetic parameters of soybean

一阶参数	估计值	一阶参数	估计值	2阶参数	估计值
1st order parameter	Estimate	1st order parameter	Estimate	1st order parameter	Estimate
m	51.4092	i_{ir}	-0.0613	σ_p^2	38.4211
d_a	-4.3403	i_{ab}	3.3971	σ_{ab}^2	35.0252
d_b	0.1840	[d]	-4.5198	σ_{da}^2	3.2465
d_c	-1.1272			σ_d^2	0.1493
i_{ab}	1.7688			$h_{ag}^2(\%)$	91.16
i_{ir}	0.4575			$h_{br}^2(\%)$	8.45

3 讨论

若要比表型遗传分析结果与 QTL 定位结果,拓展现有的遗传模型是十分必要的。这是因为 QTL 定位认为控制数量性状的基因座数多数介于 1~7 个之间,平均大约为 3~4 个^[2]。然而,若利用 F₂ 和 F_{2,3} 等世代拓展 3 对主基因+多基因混合遗传模型,则成分分布数过多,参数估计值误差较大,致使遗传参数估计结果可信度降低;有幸的是,广泛

用作遗传分析的基因型纯合的 DH 或 RIL 群体在 3 对主基因情况下其成分分布数较少。因此,本文就拓展 DH 或 RIL 群体的 3 对主基因模型。

利用该 RIL 群体进行 QTL 作图发现,其结果^[1]与本文的结果基本是一致的。因此,表型数据遗传分析结果与 QTL 定位结果有一定的一致性。这也说明本文发展的 3 对主基因遗传模型是有益的。

参考文献 (References):

- [1] 董钧熹,章元明,王建康著. 植物数量性状遗传体系[M]. 北京: 科学出版社,2001:99~103.
- [2] Kearsey M J, Farquhar A G L. QTL analysis in plants: where are we now? [J] Heredity, 1998, 80, 137~142
- [3] 胡中立,章志宏. 质量-数量性状的遗传参数估计 II. 利用 DH 或 RIL 群体[J]. 武汉大学学报(自然科学版), 1998, 44 (6):784~788.
- [4] 章元明,董钧熹. 利用 DH 或 RIL 群体检测 QTL 体系并估计其遗传效应[J]. 遗传学报, 2000, 27(7):634~640.
- [5] 胡中立,章志宏,章元明. 质量-数量性状的遗传参数估计的 P₁, P₂, DH 联合分析方法[J]. 作物学报, 2000, 26(5):631~634.
- [6] 章元明,董钧熹. 数量性状分离分析中成分分布参数估计的 IECM 算法[J]. 作物学报, 2000, 26(6):699~706.

· 书 讯 ·

《作物改良理论与方法》一书出版

刘秉华研究员的专著《作物改良理论与方法》已由中国农业科技出版社出版。全书以作物非整倍体、细胞核雄性不育及其在遗传育种研究中应用的总体论述为前导,先后引入减数分裂、非整倍体、雄性不育、轮回选择、基因定位、等基因系、性状标记、分子标记、种质创新、育种方法、杂优利用和良种良法十二专题,既包括深化传统育种方法的介绍,又含有品种改良新策略的研究,将作物遗传育种知识的广度与深度、实用性与新颖性有机地结合于一体,构成一个比较完整而系统的知识体系。全书 32 万字,图 50 幅,文字精练,问题论述深入浅出,图文并茂,是作物育种工作者、农业院校有关专业师生必备的参考书。每本定价 32 元,欲购者,请直接与作者联系。

地址:北京市中关村南大街 12 号;中国农业科学院作物所 刘秉华收 邮编:100081

电话:010-68918628