利用混合分离分析法和 Mapmaker/Exp 软件 定位互作基因的策略

吴为人 黄碧光

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

摘要 质量性状通常受单个基因控制, 但有时也受两个甚至更多基因的控制, 这种现象称为基因互作。用混合分离分析(bulked segregant analysis, BSA)快速寻找连锁分子标记, 然后用 Mapmaker/Exp 软件构建局域连锁图, 已成为定位单个主基因的常用方法。但由于基因互作的遗传模式与单基因差异很大, 因此针对单基因定位的方法和软件不能简单地应用于互作基因的定位。目前还没有提出快速筛选与互作基因连锁的分子标记的实验方法以及同时分析多个分子标记与互作基因间连锁关系的统计方法和相应的计算机软件。为解决这个问题, 本研究提出利用 BSA 和 Mapmaker/Exp 定位互作基因的策略。结果表明, 除个别情况需要用到 F_3 代之外, 绝大多数互作基因都可以直接在 F_2 代用“BSA + Mapmaker/Exp”的策略进行定位。由于 BSA 和 Mapmaker/Exp 已广泛应用于基因定位研究, 为许多学者所熟悉, 因此, 本研究提出的策略具有很好的实用性。

关键词 基因互作 基因定位 BSA Mapmaker/Exp

利用分子标记进行基因定位是现代遗传学的一项重要研究内容, 对单个主基因的定位已成为基因功能研究的常规工作。单基因定位通常包括两个步骤: 一是应用混合分离分析(bulked segregant analysis, BSA)方法^[1]快速寻找与目标基因连锁的分子标记; 二是在目标基因所在的染色体区段构建局域分子标记连锁图, 从而确定目标基因的位置。在第二步中, 通常须使用遗传作图软件, 目前最常用的作图软件是 Mapmaker/Exp^[2]。这种“BSA + Mapmaker/Exp”的方法已被证明是行之有效的, 为许多学者所熟悉和采用^[3-9]。

虽然实际研究中遇到的质量性状大多是受单个主基因控制的, 但有时也会遇到同时受两个主基因控制(亦即基因互作)的情况^[10-14]。由于基因互作时表型的分离比例取决于两对基因的分离情况, 因而对互作基因的定位要比单基因定位复杂得多。一种做法是, 沿用单基因定位的方法, 通过建立各对基因单独分离(另一对基因不发生分离)的群体, 对各对基因分别进行定位^[15]。然而, 这种做法不仅耗工费时, 而且有时也很难找到合适的亲本材料, 因此是不理想的。若能直接利用表现基因互作的 F_2 群体对互作基因进行定位, 则可大大减少所需的时间和人力物力。但目前还没有提出快速筛选与互作基因连锁的

分子标记的实验方法, 以及同时分析多个分子标记与互作基因间连锁关系的统计方法。虽然已经有利用 F_2 群体定位两个互补基因的实际研究报道^[16], 但该论文并没有说明是如何筛选与互作基因连锁的分子标记的, 而对互作基因的定位也只是采用单个标记与单个基因之间的连锁分析方法。

本文提出应用已为许多学者所熟悉的 BSA 方法和 Mapmaker/Exp 软件进行互作基因定位的策略。结果证明, 除个别情况需要用到 F_3 代外, 绝大多数互作基因都可以直接在 F_2 代用“BSA + Mapmaker/Exp”的方法进行定位。

1 互作基因的 BSA 方法

基因互作共有 6 种类型(表 1)^[17]。从表 1 可以看出, 在 F_2 代中, 受基因互作控制的性状分离成 2 或 3 种表现型(记为 Phen-1, Phen-2 和 Phen-3)。所谓 BSA, 就是将不同表现型分别建立 DNA 池, 然后比较不同池间分子标记的带型差异(即多态性), 由此判断被检标记是否与目标基因座连锁。为此, 在不同的 DNA 池间, 必须存在目标基因座的等位基因差异。根据这个原则, 可以建立各种互作类型的 BSA 方法。

1.1 重叠作用

在重叠作用中, 两对基因(A-a 和 B-b)的作用是

表 1 基因互作的类型与遗传特征

互作类型	F ₂ 代表现型与基因型		
	Phen-1	Phen-2	Phen-3
重叠	9A_B_ + 3A_bb + 3aaB_	1aabb	
互补	9A_B_	3A_bb + 3aaB_ + 1aabb	
积加	9A_B_	3A_bb + 3aaB_	1aabb
显性上位	9A_B_ + 3A_bb	3aaB_	1aabb
隐性上位	9A_B_	3A_bb	3aaB_ + 1aabb
抑制	9A_B_ + 3aaB_ + 1aabb	3A_bb	

完全对称的, 因此我们只需讨论其中一种就可以了. 重叠作用在 F₂ 中只产生两种表现型. 就 A-a 基因座而言, Phen-1 池同时包含等位基因 A 和 a (我们称之为杂合池), 而 Phen-2 池只包含 a (相应地称为纯合池). 因此, 直接利用这两种表现型的 DNA 池即可进行 BSA 分析. 对于共显性的连锁标记, 杂合池将产生 2 条电泳条带, 而纯合池只产生 1 条电泳条带(图 1). 对于显性的连锁标记, 只要其显性等位基因(表现为有电泳条带)是与只存在于杂合池中的那个目标等位基因(这里即是 A)连锁的, 则也可检测出多态性, 这时杂合池将有 1 条电泳条带, 而纯合池没有条带(图 1). 这里有关连锁标记在两池中的表现的论述是普遍适用的. 在后述对其他互作类型的讨论中, 我们将不再重复.

图 1 重叠作用的情况下与目标基因 A-a 连锁的共显性标记(上)和显性标记(下)在两亲本(P₁和P₂)及两池(Phen-1和Phen-2)中的电泳带型示意图

1.2 互补作用

在互补作用中, 两对基因的作用也是完全对称的, 因此也只需讨论其中一种就可以了. 互补作用在 F₂ 中也只产生两种表现型, 但显然, 并不能直接用两种表现型进行 BSA, 因为两种表现型都同时包含了目标基因座的两种等位基因(A 和 a), 亦即两 DNA 池间不存在等位基因差异. 为解决这个问题, 必须将 Phen-1 个体自交, 由此鉴定出 AABB 基因型的个体(在 F₃ 代中不分离, 而其他基因型在 F₃ 代中皆有表现

型分离). 这样, 具 AABB 基因型的 Phen-1 与 Phen-2 就可分别构成纯合池和杂合池.

1.3 积加作用

在积加作用中, 两对基因的作用也是完全对称的, 因此也只需讨论其中一种. 积加作用在 F₂ 中产生 3 种表现型, 其中 Phen-1 和 Phen-2 皆同时含有两种等位基因(A 和 a), 而 Phen-3 只含有一种等位基因(a). 因此, 可以将 Phen-1 和 Phen-2 分别或混合构成杂合池, 将 Phen-3 构成纯合池.

1.4 显性上位作用

显性上位作用在 F₂ 中有 3 种表现型. 由于两对基因的作用是不对称的, 因而必须分别考虑. 就 A-a 而言, Phen-1 同时包含两种等位基因(A 和 a), 故可构成杂合池; 而 Phen-2 和 Phen-3 皆只包含一种等位基因(a), 故可分别或混合构成纯合池. 对于 B-b, Phen-1 和 Phen-2 皆同时包含两种等位基因(B 和 b), 故可分别或混合构成杂合池; 而 Phen-3 只包含一种等位基因(b), 故可构成纯合池.

1.5 隐性上位作用

对隐性上位作用的分析与显性上位作用相似. 就 A-a 而言, Phen-1 和 Phen-2 可分别或混合构成杂合池, 而 Phen-3 可构成纯合池. 对于 B-b, Phen-1 和 Phen-2 可分别或混合构成杂合池, 而 Phen-3 可构成纯合池.

1.6 抑制作用

抑制作用在 F₂ 中只产生两种表现型. 就 A-a 而言, Phen-1 和 Phen-2 皆同时包含 A 和 a, 故无法直接用于 BSA 分析. 为解决这个问题, 可将 F₂ 中 Phen-2 个体自交, 其中基因型 AAbb 的自交后代不分离, 而基因型 Aabb 的自交后代将分离成 1 Phen-1(aabb) 3 Phen-2 (A_bb). 因此, 利用 Aabb 的自交后代中 Phen-1 和 Phen-2 即可分别建成纯合池和杂合池. 对于 B-b, 情况则很简单, 只需将 F₂ 中 Phen-1 和 Phen-2 分别建

成杂合池和纯合池即可。
为便于查询, 将上述互作基因的 BSA 方法总结于表 2.

2 用 Mapmaker/Exp 定位互作基因

Mapmaker/Exp 软件的主要功能是用于构建分子标记连锁图. 由于质量性状通常可以作为遗传标记, 因此 Mapmaker/Exp 也可以用来定位控制质量性状的单个主基因. 对于受单个基因(A-a, A 为显性)控制的质量性状, F₂ 中表现型分离比例为 3A_ 1aa. 因此, 对于任意一个基因, 只要其表现型在 F₂ 中呈“3 显性 1 隐性”的分离比例, 就可以像标记座位一样用 Mapmaker/Exp 进行定位. 从表 1 可以看出, 在显性上位和隐性上位中可以找出两个基因表现“3 显性 1 隐性”的分离. 在显性上位中, Phen-1 (Phen-2 + Phen-3) = 3A_ 1aa; Phen-2 Phen-3 = 3B_ 1bb. 在隐性上位中, (Phen-1 + Phen-2) Phen-3 = 3A_ 1aa; Phen-1 Phen-2 = 3B_ 1bb. 因此, 对于显性上位和隐性上位的基因, 可以基于整个 F₂ 群体(对于 A-a)或部分 F₂ 群体(对于 B-b), 用 Mapmaker/Exp 进行定位. 但对于所有其他互作基因, 都无法找到“3 显性 1 隐性”的分离. 那么, 对这些不表现“3 显性 1 隐性”分离的基因是否也能够用 Mapmaker/Exp 进行定位呢?

为了回答这个问题, 我们来看看单基因定位的极大似然法. 假设有一个共显性标记(M-m)与一个目标基因(A-a, A 为显性)连锁, 二者间重组率为 *r*. 对于

任意一个群体, 最多可分离出 6 种标记基因型和性状表现型的组合, 即 MMA_, MmA_, mmA_, MMaa, Mmaa 和 mmaa; 记各组合的频率分别为 *p*₁₁, *p*₁₂, *p*₁₃, *p*₂₁, *p*₂₂ 和 *p*₂₃, 实际观察数分别为 *n*₁₁, *n*₁₂, *n*₁₃, *n*₂₁, *n*₂₂ 和 *n*₂₃. 于是, 可以写出似然函数:

$$L(r) = L_1(r) \cdot L_2(r) = \prod_{i=1}^3 p_{1i}^{n_{1i}} \cdot \prod_{i=1}^3 p_{2i}^{n_{2i}} \tag{1}$$

式(1)是针对完整群体的. 若只利用显性表现型进行分析(即令 *n*₂₁ = *n*₂₂ = *n*₂₃ = 0), 则式(1)变成

$$L(r) = L_1(r) = \prod_{i=1}^3 p_{1i}^{n_{1i}} \tag{2}$$

若只利用隐性表现型进行分析(即令 *n*₁₁ = *n*₁₂ = *n*₁₃ = 0), 则式(1)变成

$$L(r) = L_2(r) = \prod_{i=1}^3 p_{2i}^{n_{2i}} \tag{3}$$

显然, 不论是式(1), 式(2)还是式(3), 都存在一个 *r* 值使 *L*(*r*) 达到极大值, 亦即都能够得到 *r* 的极大似然估计值. 例如, 在一个 MMAA × mmaa 的 F₂ 代中, 用隐性表现型个体分析得到的 *r* 的极大似然估计值为 $\hat{r} = (n_{21} + n_{22}/2)/(n_{21} + n_{22} + n_{23})$ ^[18]. 这说明, 在应用极大似然法进行单基因定位分析中, 并不一定需要完整的群体, 利用单种表现型也是可以的.

Mapmaker/Exp 软件采用的是极大似然法, 因此有理由推测, Mapmaker/Exp 可以利用单种表现型进行单基因的定位. 事实上, 在实际研究中, 一些学者

表 2 互作基因的 BSA 方法		
互作类型	目标基因	合适的 DNA 池对
重叠	A-a/B-b	杂合池: Phen-1
		纯合池: Phen-2
互补	A-a/B-b	杂合池: Phen-2
		纯合池: Phen-1 中的 AAbb(通过 F ₃ 鉴定)
积加	A-a/B-b	杂合池: Phen-1/Phen-2/Phen-1+Phen-2
		纯合池: Phen-3
显性上位	A-a	杂合池: Phen-1
		纯合池: Phen-2/Phen-3/Phen-2+Phen-3
	B-b	杂合池: Phen-1/Phen-2/Phen-1+Phen-2
		纯合池: Phen-3
隐性上位	A-a	杂合池: Phen-1/Phen-2/Phen-1+Phen-2
		纯合池: Phen-3
	B-b	杂合池: Phen-1/Phen-3/Phen-1+Phen-3
		纯合池: Phen-2
抑制	A-a	在来自 Phen-2(Aabb)的 F ₃ 中:
		杂合池: Phen-2(A ₁ bb)
	B-b	纯合池: Phen-1(aabb)
		杂合池: Phen-1
		纯合池: Phen-2

已经利用Mapmaker/Exp分析单种表现型来定位单基因 [3~5], 尽管这种做法的可行性尚未经过严格验证. 为了验证利用Mapmaker/Exp分析单种表现型来定位单基因的可行性, 我们进行了计算机模拟实验.

假设目标基因座(A-a)与 3 个标记座位(M₁-m₁, M₂-m₂, M₃-m₃)连锁, 排列顺序为 M₁-M₂-A-M₃, 相邻座位间距相等(为 5 或 10 cM), 重组率与图距间的关系服从 Kosambi 作图函数. 用计算机随机抽样产生 F₂ 的完整群体、显性表现型群体(A₋)或隐性表现型群体(aa), 大小皆为 200, 每种情况皆产生 50 个样本. 然后用 Mapmaker/Exp (3.0 版)进行连锁分析, 并对各相邻座位间图距的估计值与真值之间的偏差进行 *t* 检验. 结果表明(表 3), 不论是用显性表现型群体还是用隐性表现型群体, 估得的相邻座位间的遗传图距与其真值之间皆不存在显著差异, 而且在相同的样本容量下, 其估计效果不亚于使用完整的 F₂ 群体. 因此看来, 用 Mapmaker/Exp 分析单种表现型进行单基

因定位是完全可行的.

现在, 回到利用 Mapmaker/Exp 定位互作基因的问题上来. 从表 1 可以看出, 对于每一个互作基因, 至少存在一种对应其“显性”(A₋)或“隐性”(aa)的表现型. 例如, 在重叠作用中, Phen-2 对于基因 A-a 和 B-b 来说都是“隐性表现型”; 在互补作用中, Phen-1 对于基因 A-a 和 B-b 来说则都是“显性表现型”; 在抑制作用中, Phen-2 对于基因 A-a 来说是“显性表现型”, 而对于基因 B-b 来说则是“隐性表现型”. 因此, 根据 Mapmaker/Exp 可以分析单种表现型进行基因定位的性质, 我们就可以利用它来定位互作基因了.

综上所述, 我们可以建立利用 Mapmaker/Exp 定位互作基因的策略: 所有互作基因皆可利用单种表现型进行定位; 但对于显性上位和隐性上位互作基因, 由于其表现型分离比例符合“3 显性:1 隐性”, 也可以使用完整的或部分的 F₂ 群体进行定位. 各种互作类型的基因定位策略总结于表 4.

表 3 单种表现型群体连锁分析的模拟实验结果

F ₂ 群体	相邻座位	图距/cM		<i>t</i>	<i>P</i> (大于)
		真值	估计值		
完整	M ₁ -M ₂	5	5.06 ± 1.17	0.36	0.50
	M ₂ -A	5	4.86 ± 1.35	0.73	0.40
	A-M ₃	5	5.22 ± 1.63	0.95	0.20
显性	M ₁ -M ₂	5	5.14 ± 1.09	0.91	0.20
	M ₂ -A	5	4.99 ± 1.94	0.04	0.50
	A-M ₃	5	4.96 ± 1.67	0.17	0.50
隐性	M ₁ -M ₂	5	4.73 ± 0.98	1.94	0.05
	M ₂ -A	5	5.02 ± 1.06	0.13	0.50
	A-M ₃	5	4.92 ± 1.22	0.46	0.50
完整	M ₁ -M ₂	10	9.80 ± 1.89	0.75	0.40
	M ₂ -A	10	10.32 ± 2.03	1.11	0.20
	A-M ₃	10	10.14 ± 1.92	0.51	0.50
显性	M ₁ -M ₂	10	9.84 ± 1.83	0.62	0.50
	M ₂ -A	10	9.67 ± 2.64	0.89	0.20
	A-M ₃	10	10.14 ± 2.28	0.43	0.50
隐性	M ₁ -M ₂	10	10.11 ± 1.36	0.57	0.50
	M ₂ -A	10	10.07 ± 1.30	0.38	0.50
	A-M ₃	10	10.20 ± 1.53	0.92	0.20

表 4 用 Mapmaker/Exp 定位互作基因的策略

互作类型	适合的作图群体 ^{a)}	
	A-a	B-b
重叠	Phen-2(R)	Phen-2(R)
互补	Phen-1(D)	Phen-1(D)
积加	Phen-1(D)/Phen-3(R)	Phen-1(D)/Phen-3(R)
显性上位	完整 F ₂ /Phen-1(D)/Phen-2+Phen-3(R)	Phen-2+Phen-3/Phen-2(D)/Phen-3(R)
隐性上位	完整 F ₂ /Phen-1+Phen-2(D)/Phen-3(R)	Phen-1+Phen-2/Phen-1(D)/Phen-2(R)
抑制	Phen-2(D)	Phen-2(R)

a) 括号中的大写英文字母表示目标基因的“表现型”. D: 显性(A₋或 B₋); R: 隐性(aa 或 bb)

3 讨论

BSA 是快速鉴定与目标基因连锁的分子标记的有效方法。BSA 要求两个 DNA 池之间在目标基因座上存在等位基因的差异。根据这个原理,我们建立了针对互作基因的 BSA 策略。从表 2 可以看出,每对 DNA 池都包含一个杂合池和一个纯合池。然而,需指出的是,BSA 并不一定需要建立两个 DNA 池。事实上,将纯合池(AA 或 aa)与亲本 DNA 比较,也能够检测出连锁标记。一个共显性标记若不与目标基因连锁,则在纯合池中会出现两亲本的条带;若与目标基因连锁,则只会出现一个亲本的条带。对于一个连锁显性标记,只要它与目标基因的连锁相是合适的,即在隐性纯合池(aa)中为相引连锁,在显性纯合池(AA)中为相斥连锁,则也能够检测到。这样,与目标基因连锁的显性标记在纯合池中将没有多态性条带,而不与目标基因连锁的显性标记在纯合池中将出现多态性条带。因此,构建纯合池是 BSA 的关键。

我们已经证明极大似然法和 Mapmaker/Exp 软件可以利用单种表现型定位单基因。这种性质使得我们可以利用 Mapmaker/Exp 来定位互作基因。这种性质不应是 Mapmaker/Exp 所独有的,其他基于极大似然法的遗传作图软件应该也具有这种性质,因此它们也应该能够用于互作基因的定位。这种利用现有遗传作图软件来定位互作基因的策略的优点是无需发展新的统计方法和软件。由于 Mapmaker/Exp 软件可以免费获得,并为许多学者所熟悉,已得到广泛应用,因此在实际研究中用它来定位互作基因是非常方便的。

本研究只考虑了双基因互作的情形。然而,利用 BSA 和 Mapmaker/Exp 来定位互作基因的原理同样适用于多个基因互作的情形。例如,考虑 3 个重叠基因(A-a, B-b, C-c)的情形。F₂ 代将分离出两种表现型,即 63 Phen-1 (除 aabbcc 之外的所有其他基因型) 1 Phen-2 (aabbcc)。显然,对于其中任意一个基因,都可以用 Phen-1 和 Phen-2 构建 BSA 所需的杂合池和纯合池,并用 Phen-2 进行 Mapmaker/Exp 的定位分析。

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

- 1 Michelmore R W, Paran I, Kesseli R V. Identification of markers linked to disease-resistance genes by bulked segregant analysis: A rapid method to detect markers in specific genomic regions by using segregating populations. *Proc Natl Acad Sci USA*, 1991, 88: 9828—9832[DOI]
- 2 Lincoln S E, Daly M J, Lander E S. Constructing genetic linkage maps with MAPMARKER/EXP Version 3.0: A tutorial and reference manual, 3rd ed. Cambridge MA: Whitehead Institute for Biometrical Research, 1993
- 3 Chen D H, dela Viña M, Inukai T, et al. Molecular mapping of the blast resistance gene, *Pi44(t)*, in a line derived from a durably resistant rice cultivar. *Theor Appl Genet*, 1999, 98: 1046—1053[DOI]
- 4 Ferreira C F, Borém A, Carvalho G A, et al. Inheritance of angular leaf spot resistance in common bean and identification of a RAPD marker linked to a resistance gene. *Crop Sci*, 2000, 40: 1130—1133
- 5 Wu J Y, Wu H K, Chung M C. Co-dominant RAPD markers closely linked with two morphological genes in rice (*Oryza sativa* L). *Bot Bull Acad Sin*, 2002, 43: 171—180
- 6 何光华, 王文明, 刘国庆, 等. 利用 SSR 标记定位明恢 63 的 2 对恢复基因. *遗传学报*, 2002, 29: 798—802
- 7 段远霖, 吴为人, 张丹凤, 等. 水稻幼穗分化受阻突变体 *lhd* 的遗传分析与基因定位. *科学通报*, 2003, 48: 1662—1665[摘要]
- 8 段远霖, 李维明, 吴为人, 等. 水稻小穗分化调控基因 *fzp(t)* 的遗传分析和分子标记定位. *中国科学C辑: 生命科学*, 2003, 33: 27—32[摘要]
- 9 Mammadov J A, Zwonitzer J C, Biyashev R M, et al. Molecular mapping of leaf rust resistance gene *Rph5* in barley. *Crop Sci*, 2003, 43: 388—393
- 10 Bingham E T, Groose R W, Woodfield D R, et al. Complementary gene interactions in alfalfa are greater in autotetraploids than diploids. *Crop Sci*, 1994, 34: 823—829
- 11 Li H, Liu A, Tang L, et al. Complementary genes controlling temperature-sensitive sterility in hybrids between indica and japonica types. *Rice Genet Newsl*, 1995, 12: 184
- 12 Jonsson R, Sall T, Kraft T, et al. Inheritance of resistance to *Pyrenophora teres f teres* in spring barley. *Plant Breed*, 1999, 118: 313—317
- 13 黄碧光, 卢泳全. 杂交水稻 V 优 77 色素与育性的遗传关系研究. *福建农业学报*, 2002, 17(4): 215—216
- 14 Ragiba M, Prabhu V K, Singh R B. Monosomic analysis of *Helminthosporium* leaf blight resistance genes in wheat. *Plant Breed*, 2004, 123: 405—409[DOI]
- 15 Lee J M, Bush A L, Specht J E, et al. Mapping of duplicate genes in soybean. *Genome*, 1999, 42: 829—836[DOI]
- 16 吴建宇, 丁俊强, 杜彦修, 等. 两个玉米矮花叶病显性互补抗病基因的发现和定位. *遗传学报*, 2002, 29: 1095—1099
- 17 Phillips P C. The language of gene interaction. *Genetics*, 1998, 149: 1167—1171
- 18 Zhang Q, Shen B Z, Dai X K, et al. Using bulked extremes and recessive class to map genes for photoperiod-sensitive genic male sterility in rice. *Proc Natl Acad Sci USA*, 2004, 91: 8675—8679[DOI]

(2006-05-24 收稿, 2006-08-15 接受)

<https://engine.scichina.com/doi/10.1360/972006-790>