

基于加性-显性效应的杂种表现分子标记预测模型

徐新福, 唐章林, 李加纳, 柴友荣, 王 瑞, 谌 利

(西南大学农学与生物科技学院/重庆市油菜工程技术研究中心, 重庆 400716)

摘要: 【目的】以甘蓝型油菜为研究材料, 探索杂种表现分子标记预测模式。【方法】以 6 份甘蓝型油菜隐性核不育两型系、11 份恢复系为亲本按 NC II 设计配制成 66 个杂交组合。利用 F_1 的 9 个性状的表型值对亲本材料的 SSR 和 AFLP 标记位点进行筛选, 建立标记效应和标记型值估算体系, 估算这些特异标记位点对性状表现的效应及杂种标记型值, 进而分析杂种标记型值与杂种表现的相关性, 应用逐步回归分析建立 9 个性状杂种表现的分子标记预测模型。【结果】114 个 SSR 和 205 个 AFLP 标记位点中, 在 0.01 显著水平下 9 个性状筛选到的特异性标记位点分别为 39~85 个; 不同标记位点对性状表现的效应大小、方向及作用方式存在广泛差异; 9 个性状的杂种 F_1 标记型值与性状表现间的相关性均达到极显著水平, 相关系数为 0.6824~0.8113。9 个预测模型中分别包括了 6~14 个标记位点, 可决系数 R^2 为 0.5191~0.6783, 预测模型稳定性强, 精确度较高。【结论】利用标记型值预测作物杂种表现是一种有效的途径, 可以利用较少的标记位点建立甘蓝型油菜杂种表现预测模型。

关键词: 甘蓝型油菜; 杂种表现; 分子标记; 预测模型

Prediction Model of Hybrid Performance Using Molecular Marker Based on Additive-Dominant Effects

XU Xin-fu, TANG Zhang-lin, LI Jia-na, CHAI You-rong, WANG Rui, CHEN Li

(College of Agronomy and Biotechnology, Southwest University, /Chongqing Engineering Research Center for Rapeseed, Chongqing 400716, China)

Abstract: 【Objective】 Mathematic model was explored to predict hybrid performance using molecular markers in *Brassica napus* L. 【Method】 66 hybrids were developed between six recessive nuclear male sterile lines and 11 restorer lines according to the design of North Carolina mating design II (NC II). Specific marker loci of SSR and AFLP with significant effects on nine agronomic traits of hybrids were chosen from DNA polymorphic loci among parents by ANOVA to estimate the effects of markers (additive or dominant) and the marker-type values of parents and their hybrids. The relation between the marker-type values and hybrid performance was analyzed. The mathematic models to predict hybrid performance were established by stepwise regression analysis. 【Result】 Marker loci with significant effects of additive or dominant ($P=0.01$), ranging from 39 to 85, were chosen from 114 SSR and 205 AFLP markers for nine agronomic traits, which exhibited different effects in the size and direction. High and significant ($P<0.01$) relations of marker-type values were found with hybrid phenotype values with r ranging from 0.6824 to 0.8113 for nine traits. The models established with 6-14 marker loci can be used to predict precisely the hybrid performance, with R^2 ranging from 0.5191 to 0.6783 for nine traits. 【Conclusion】 It is a new and effective way to predict hybrid performance using marker-type values. It is possible to establish model to predict hybrid performance using small amounts of molecular markers in *B. napus* L..

Key words: *Brassica napus* L.; Hybrid performance; Molecular marker; Prediction model

0 引言

【研究意义】油菜是世界范围内广泛种植的主要

油料作物之一。中国油菜种植面积和总产均约占世界的 1/3, 均居世界首位, 其中杂交种植面积占栽培总面积的 50%左右。中国的油菜杂种优势利用研究居国

收稿日期: 2007-07-20; 接受日期: 2008-05-06

基金项目: 国家“863”项目(20060110z1053), 重庆市自然科学基金(CSTC2007BB1361), 国家科技支撑计划(2006BAD01A04)

作者简介: 徐新福(1971-), 男, 河南遂平人, 副研究员, 研究方向为油菜遗传育种。E-mail: xinfuxu@126.com。通讯作者唐章林(1966-), 男, 四川罗江人, 研究员, 研究方向为油菜遗传育种。Tel: 023-68250701; E-mail: tangzhlin@126.com

际领先地位,但在杂优育种中,还没有一个简便而有效的预测指标来指导亲本选配。在多种作物上分子标记技术被用于杂种表现及杂种优势的预测研究,而在油菜上这方面的研究报道还比较少。【前人研究进展】利用分子标记预测杂种优势的研究,主要集中在利用 DNA 分子标记研究亲本的遗传差异与 F_1 杂种优势的关系(Smith 等^[1]、Godshalk 等^[2]、Dudley 等^[3]、Melchinger 等^[4,5]、Xiao 等^[6]、Dreisigacker 等^[7]和 Yadav 等^[8]),然而不同作物、不同研究者得出的结论迥异。多数研究者认为,分子标记检测到的亲本遗传差异与杂种优势没有相关或相关程度很低。沈金雄等^[9,10]和 Yu 等^[11]在对甘蓝型油菜的研究中也得到相似的结论。张启发等^[12,13]利用特殊杂合性分析提高了相关系数,但在随后的研究中发现,当研究材料不同时,这种高相关性也不复存在。何光华等^[14]把阳性位点分解成增效位点和减效位点,单独利用增效位点或减效位点计算的亲本间遗传距离与杂种表现的相关性得到很大提高。

【本研究切入点】Bernardo^[15]较早地将 QTL 效应应用到杂种表现的预测研究,Vuylsteke 等^[16]在对玉米杂种表现预测研究中应用 QTL 效应分析取得了较好的预测效果。而 Shen 等^[17]在甘蓝型油菜杂种优势研究中,通过对一组双列杂交 F_1 及其中一个组合的 $F_{2:3}$ 分

离群体分析,认为无论是亲本杂合性还是产量性状的 QTL 都不能预测其杂种优势及杂种表现。【拟解决的关键问题】本研究在传统数量遗传学基础上,联合分析甘蓝型油菜生物学数据和分子标记数据,通过估算特异标记位点的标记效应计算亲本及杂种 F_1 的标记型值,分析杂种标记型值与表型值的关系,应用逐步回归分析建立基于加性—显性效应的甘蓝型油菜杂种表现的分子标记预测模型。

1 材料与方法

1.1 试验材料

研究中的 17 份甘蓝型油菜亲本材料是从西南大学重庆市油菜工程技术研究中心的骨干亲本中随机选取的,其中 6 份为隐性核不育两型系,11 份为恢复系(表 1)。所有研究材料均已自交或兄妹交选育 10 代以上。2003 年 3 月配制 6×11 的不完全双列杂交组合。

1.2 田间试验设计及性状考查

2003~2004 年把 66 份杂交组合、17 份杂交亲本以中油 821 为对照进行田间试验。试验设在西南大学试验农场,2003 年 9 月 23 日播种育苗,10 月 28 日移栽至大田。随机区组设计,2 次重复,3 行区,行距 0.4 m,行长 2.4 m,株距 0.24 m,每行 10 株。田间管

表 1 供试材料及来源

Table 1 Parent lines tested and their origins

亲本编号 Parent code	名称或系谱 Name or Genealogy	选育单位 Origin
P1	L18A×[194×(Expander×74-317)]AB	西南大学 Southwest University
P2	2018AB	上海市农科院 Shanghai Academy of Agricultural Sciences
P3	蜀杂 6 号选系 AB Shuza No.6 AB	四川大学 Sichuan University
P4	蜀杂 7 号选系 AB Shuza No.7 AB	四川大学 Sichuan University
P5	油研 9 号选系 AB Youyan No.9 AB	贵州思南油科所 Sinan Oil Research Institute, Guizhou
P6	绵油 12 号选系 AB Mianyou No.12 AB	四川绵阳市农科所 Mianyang Agricultural Research Institute
P7	SC94005×Andor	西南大学 Southwest University
P8	Tower×81008×81008×非野 Tower×81008×81008×Fei-yei	西南大学 Southwest University
P9	Tower×81008×81008×非野 Tower×81008×81008×Fei-yei	西南大学 Southwest University
P10	84-57N-9* ^{100K} L	西南大学 Southwest University
P11	84-57N-9* ^{100K} L×GH01	西南大学 Southwest University
P12	GH01×中双 1 号 GH01×Zhongshuang No.1	西南大学 Southwest University
P13	GH01×(L17C×油研 2 号) GH01×(L17C×Youyan No.2)	西南大学 Southwest University
P14	GH01×94005	西南大学 Southwest University
P15	6094×D2	西南大学 Southwest University
P16	88-14	西南大学 Southwest University
P17	955	华中农业大学 Huazhong Agricultural University

理按常规生产方式进行。成熟期在中间 1 行随机取样 6 株进行考种，考察性状包括：株高、一次有效枝起点、主序有效长、主序有效果数、一次有效分枝数、全株有效果数、每果粒数、千粒重、单株产量等 9 个。其中不育两型系在初花期对可育植株系绳标记，成熟期选可育株作性状考查。

1.3 SSR 和 AFLP 分析

对于每份亲本材料，于幼苗期从 10 株上摘取等量幼嫩叶片，用 CTAB 方法提取基因组 DNA。

用于 SSR 分析的 100 对甘蓝型油菜引物，来自 <http://ukcrop.net> 的 *Brassica* 数据库，引物编号为 NAS1~NAS100，由上海博亚公司合成。AFLP 分析采用鼎国生物技术公司 AFLP 试剂盒及其所提供的操作步骤。

扩增位点根据引物和扩增片段大小进行排序编码，SSR 扩增位点编码形式为：NAS x ，AFLP 扩增位点编码形式为：ExMy-z，其中 Ex、My 为 *EcoR* I 和 *Mse* I 酶切位点对应的引物编号，z 为按扩增片段大小所排序号。按亲本材料在 SSR 及 AFLP 某标记位点扩增带的有无记为“1”或“0”，称之为亲本在该位点的标记型。

1.4 标记筛选

本研究中的所有亲本材料都经过 10 代以上的自交或兄妹交，可以认为在各位点是纯合的，由杂交亲本的标记型可以推测出杂种 F_1 的标记型。根据 F_1 的标记型把 F_1 分为纯合 (0, 0) 组（双亲标记型都为 0），纯合 (1, 1) 组（双亲标记型都为 1）及杂合 (1, 0) 组（双亲标记型一个为 1，另一个为 0）。采用单向分组的方差分析法，对 3 组 F_1 性状表型均值进行 t 测验，筛选与该性状有关的标记位点（特异性标记）。

1.5 遗传距离 (GD) 计算

按 Nei 和 Li^[18]的方法计算遗传相似系数 (GS) 和遗传距离 (GD)：

$$GS = 2X_{12} / (X_1 + X_2)$$

$$GD = -\ln GS$$

其中， X_1 、 X_2 分别为成对比较的两个亲本的扩增带数， X_{12} 为这 2 个亲本共有带数。一般遗传距离 (GGD) 是指由所有扩增位点计算的遗传距离，特殊遗传距离 (SGD) 是指由筛选出的特异性标记位点计算的遗传距离。

1.6 杂种表现的预测模型

1.6.1 标记效应估计 对于筛选到的特异性标记，根据三组标记型的杂种 F_1 性状表型均值差异显著性，分

析可能存在的遗传效应，按如下式估算该标记位点的加性效应 a 和显性效应 d ：

$$a = \frac{1}{2}(\bar{x}_{11} - \bar{x}_{00}) / se$$

$$d = [\bar{x}_{10} - \frac{1}{2}(\bar{x}_{11} + \bar{x}_{00})] / se$$

其中， \bar{x}_{11} 是 (1, 1) 组 F_1 的性状表型均值， \bar{x}_{00} 是 (0, 0) 组 F_1 的性状表型均值， \bar{x}_{10} 是 (1, 0) 组 F_1 的性状表型均值， se 为标准误。当某标记位点加性效应存在时，该位点 (1, 1) 标记型对性状的效应为 a ，(0, 0) 标记型对性状的效应为 $-a$ ，当显性效应存在时，该位点 (1, 0) 标记型对性状的效应为 d 。

根据第 i 和第 j 个杂交亲本在第 k 个标记位点的标记型和该位点的加性、显性效应，估算该标记位点对该杂种 F_1 的实际效应 g_{Mk} ：

$$g_{Mk} = a(M_{ik} + M_{jk} - 1) + d|M_{ik} - M_{jk}|$$

其中， M_{ik} 和 M_{jk} 分别是第 i 和第 j 个杂交亲本在第 k 个标记位点的标记型，取值为“0”或“1”。

在不考虑互作的情况下，各位点间的效应是可以累加的，即符合线性可加性原则，从而可以得到杂交亲本及杂种的标记型值估计公式：

$$p_i = \sum_{k=1}^n a(2M_{ik} - 1)$$

$$F_{ij} = \sum_{k=1}^n (a(M_{ik} + M_{jk} - 1) + d|M_{ik} - M_{jk}|)$$

式中， p_i 是亲本 i 的标记型值， F_{ij} 是亲本 i 和亲本 j 的杂种 F_1 的标记型值， n 为与性状相关的特异标记位点总数。

1.6.2 基于加性—显性效应的杂种表现预测模型 在加性—显性遗传模型下，杂种基因型值与表型值可以分别表示为：

$$G = U + \sum_{l=1}^h g_{Ql}$$

$$P = U + \sum_{l=1}^h g_{Ql} + E$$

式中， G 为性状基因型值， P 为性状表型值， U 为常数， g_{Ql} 为第 l 个 QTL 位点的效应， h 为与该性状有关的 QTL 数目， E 为随机误差。

当每个 QTL 选择一个与其最紧密连锁的标记位点时，杂种基因型值与表型值可以用分子标记效应分别表示为：

$$G = U' + \sum_{k=1}^h g_{Mk}$$

$$P = U' + \sum_{k=1}^h g_{M_k} + E$$

由于生物生长发育的复杂性,与某个性状有关的 QTL 可能会很多,因此利用方差分析所筛选到的特异标记位点也会较多;而且这些 QTL 往往存在复杂的网状关系,有些 QTL 可能只是一些间接作用;另外与同一个 QTL 连锁的标记往往不只 1 个,这些都不利于在实际研究中应用,需要对特异标记位点作出进一步选择。

在本研究中,利用所有相关位点的实际标记效应对性状表型值进行逐步回归分析建立多元回归预测模型:

$$y = X\beta + \varepsilon$$

1.7 预测模型 Jackknife 检验

Jackknife 抽样技术是 Quenouille 于 1956 年提出的,近年来广泛用于模型的评价和参数估计^[19-21],其中可决系数 R^2 和预测残差(或剩余平方和)是模型评价的重要参数。在本研究中,每次从 66 个组合中剔除一个,计算可决系数 R^2 和预测残差,评价模型的稳定性和精确度。

2 结果与分析

100 对 SSR 引物对 17 个杂交亲本 DNA 模板扩增结果,有 45 对引物具有多态性,共得到 114 条多态性位点。从 17 个杂交亲本中随机选取两份材料用于对

64 对 AFLP 引物组合进行筛选,获得扩增效果好的引物组合 34 对,用于 17 个杂交亲本 AFLP 标记研究。对 100~2 000 bp 的扩增位点进行统计,得到 205 个具有多态性的 AFLP 标记位点。

2.1 特异性标记筛选

表 2 给出了与性状相关的特异性标记筛选结果。在 0.01 的显著水平下,319 个标记位点中筛选到 268 个与 9 个性状相关的位点,单株产量筛选到的相关位点最少,仅 39 个,而千粒重的相关位点则多达 85 个,其它性状的相关位点在 42~58。部分标记位点同时在多个性状中被筛选出来,其中 NAS10 标记位点与 7 个性状有关,NAS41、NAS82、E5M4-4、E7M5-7 分别与不同的 6 个性状相关,平均每个位点与 2.32 个性状相关。这些位点有的只表现加性作用(加性位点),有的只表现显性作用(显性位点),而有的位点则既具有加性作用又具有显性作用(加显位点)。对每个性状来说,显性位点所占比例较高,加性位点较低,加显位点所占比例最低。

在 0.001 的显著水平下,319 个标记位点中筛选到 121 个与 9 个性状相关的位点。与在 0.01 的显著水平下筛选结果相比,各性状的相关位点数目大大减少,单株产量相关位点仅有 9 个,与千粒重相关的位点数目下降到 46 个。

2.2 亲本遗传距离与杂种表现的相关性

表 2 特异标记位点筛选结果

Table 2 Screening results of specific molecular marker loci

	显著水平 $\alpha=0.01$ 0.01 level of significance				显著水平 $\alpha=0.001$ 0.001 level of significance			
	加性位点 Additive loci	显性位点 Dominant loci	加显位点 Additive and dominant loci	总位点 All loci	加性位点 Additive loci	显性位点 Dominant loci	加显位点 Additive and dominant loci	总位点 All loci
株高 Plant height	22	28	7	43	7	12	4	15
一次有效枝起点 Branch height	19	34	2	51	2	11	0	13
主序有效长 Length of main inflorescence	20	34	12	42	8	13	3	18
一次有效枝数 No. of primary branches	25	44	11	58	6	12	1	17
全株有效果数 Siliques per plant	18	38	12	44	9	16	5	20
主序有效果数 Siliques of main inflorescence	20	45	10	55	8	14	0	22
每果粒数 Seeds per silique	21	34	11	44	9	11	5	15
千粒重 1000-seed weight	40	71	26	85	18	36	8	46
单株产量 Yield per plant	15	33	9	39	2	9	2	9

从表 3 可以看出，杂交亲本分子标记的一般遗传距离 (GGD) 与杂种 F_1 性状表型值的相关性都不高，其中与主序有效数的表型值呈极显著负相关 (-0.4294)，与一次有效枝起点及每果粒数表型值的相关性达 0.05 的显著水平，而与其它性状表型值的相关均不显著。

亲本特异标记遗传距离 (SGD) 与杂种性状表型值的相关性除一次有效枝起点和主序有效数外其它性状上的相关性都有不同程度的提高。杂种 F_1 株高、一次有效枝数和单株产量表型值与其亲本的两种遗传距离的相关性都不显著。

表 3 亲本遗传距离与杂种表现相关性分析

Table 3 Correlations between hybrid performance and genetic distance of their parents

	株高 Plant height	一次有效 枝起点 Branch height	主序有效长 Length of main inflorescence	一次有效枝数 No. of primary branches	全株有效果数 Siliques per plant	主序有效果数 Siliques of main inflorescence	每果粒数 Seeds per silique	千粒重 1000-seed weight	单株产量 Yield per plant
GGD	-0.085	-0.28*	-0.1396	0.0274	-0.1567	-0.4294**	0.2495*	0.209	0.0548
SGD	-0.0987	-0.1591	-0.2433*	0.0763	-0.5165**	-0.3821**	0.4785**	0.5039**	-0.1527

* $r_{0.05}=0.232$, ** $r_{0.01}=0.301$

2.3 标记型值与杂种表现的相关性分析

2.3.1 标记效应估计 在 9 个性状中都存在一些位点只具有加性效应，另一些位点只具有显性效应，只有少数位点同时具有加性、显性效应。同时具有加性、显性效应的位点中，又以部分显性效应为主，在 9 个

性状中，只发现 2 个同时具有加性、显性效应的位点是超显性。无论是加性效应，还是显性效应，不同位点间的效应大小存在差异，作用方向也不一致。

与单株产量相关的 39 个标记位点中，15 个位点具有加性效应 (表 4)，效应值在 $-1.0732 \sim 0.8049$ ，

表 4 单株产量相关的标记位点及效应

Table 4 Marker loci and their effects related to yield per plant

标记位点 Loci	显性效应 Dominant effect	加性效应 Additive effect	标记位点 Loci	显性效应 Dominant effect	加性效应 Additive effect
NAS4	0	-0.892	E4M2-8	-1.2944	0
NAS6	-0.6946	0	E4M2-14	-0.7445	-1.0723
NAS8	-0.8351	0	E4M3-3	1.7253	0
NAS10	0	0.5174	E4M6-4	-0.7086	0
NAS23-1	0.2323	0	E5M3-3	-0.2565	0.7868
NAS23-2	0.1527	0	E5M5-3	0.1757	0.6755
NAS26	-0.4647	0	E5M7-3	0	0.6105
NAS41-1	0.2429	-0.8015	E5M7-4	-1.0261	0
NAS41-2	0.7564	0	E5M8-5	0.9392	0
NAS45	1.7253	0	E5M8-7	-0.9376	0
NAS61	-0.9309	0	E5M8-10	0.7573	0
NAS64	-0.3333	0	E6M1-2	0	0.8049
NAS82	0.312	-0.6586	E6M3-7	-1.3496	0
NAS99-1	0	-0.4677	E6M3-9	-0.2834	0
NAS99-2	-0.154	0.7843	E7M3-4	0.8273	0
E3M2-10	0	0.8043	E7M5-9	0.2407	-0.6759
E3M2-12	0	0.7174	E7M6-2	0.2668	-1.0019
E3M2-16	-1.0966	0	E8M1-1	0.8972	0
E3M7-4	0.7148	0	E8M8-7	-0.2507	0
E4M2-6	-1.2367	0			

7 个位点效应值小于 0，8 个位点效应值大于 0，也就是说 15 个有加性效应的位点中，8 个标记型为 1 的位点具有增效作用，7 个标记型为 0 的位点具有增效作用，其中 E4M2-14 对单株产量的加性作用最强。33 个具有显性效应的位点，效应值在 -1.3496~1.7253。16 个显性效应大于 0，17 个显性效应小于 0，可见在杂种表现中并非杂合都是有益的。

2.3.2 杂种标记型值与杂种表现的相关性 表 5 给

出了各亲本及 F₁ 利用 0.01 显著水平下筛选到的特异性标记位点估算的单株产量标记型值。17 个杂交亲本的单株产量标记型值在 -11.271~11.271，平均为 -0.8534；66 个杂种 F₁ 的单株产量标记型值在 -20.672~18.661，平均为 -2.806。杂种 F₁ 间单株产量标记型值的变异幅度大于杂交亲本间的变异幅度。

从图中可以看出，杂种标记型值与性状表现的相关性相当高，9 个性状的相关系数在 0.6824~0.8113

表 5 杂交亲本及 F₁ 单株产量标记型值估计

Table 5 Marker-type values of yield per plant of hybrids and their parents

亲本编号 Parent code	P1	P2	P3	P4	P5	P6
	1.41	-6.617	8.901	7.664	5.682	-2.618
P7	5.944	2.069	-3.304	4.776	4.498	4.828
P8	-3.335	-2.989	-7.884	1.816	3.797	2.064
P9	-1.331	-2.287	-6.583	3.085	5.067	2.832
P10	-11.271	-11.983	-20.672	-10.566	-8.936	-10.318
P11	-5.208	-6.981	-11.046	-4.489	-3.324	-3.936
P12	3.198	-0.863	-3.383	4.509	6.49	4.256
P13	-5.484	-4.26	-10.381	1.464	3.093	0.281
P14	11.271	12.587	9.652	18.661	15.795	15.195
P15	-5.721	-6.856	-11.751	-1.699	-0.07	-1.452
P16	-11.271	-11.983	-20.672	-10.566	-8.936	-10.318
P17	-5.721	-6.856	-11.751	-1.699	-0.07	-1.452

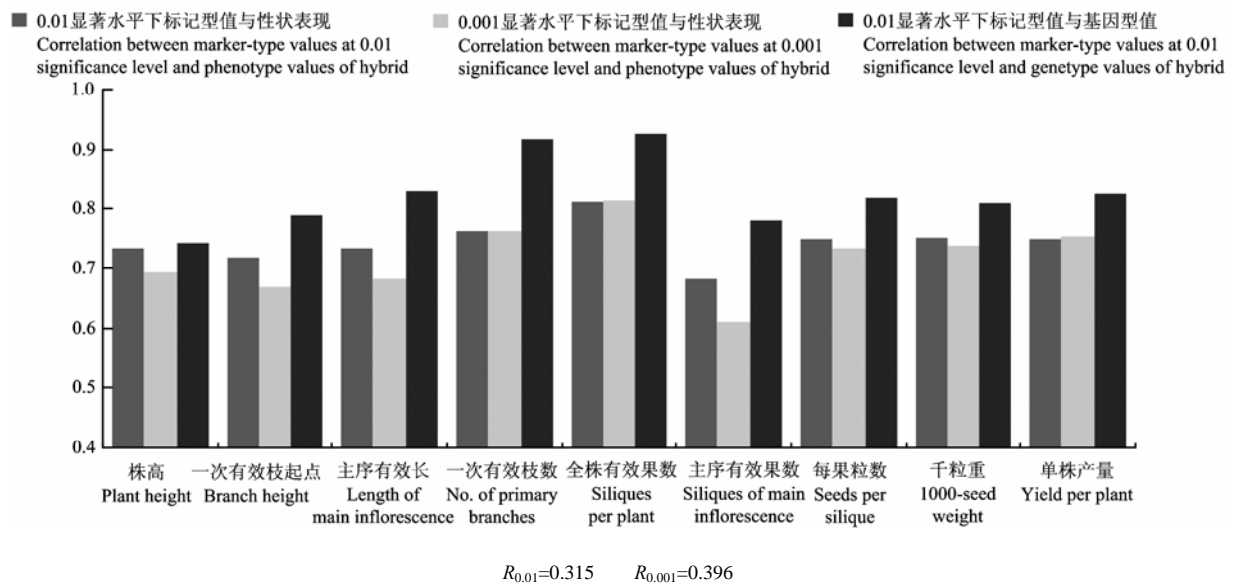


图 杂种标记型值与性状表型值及基因型值的相关性

Fig. Correlation between marker-type values of hybrid and their phenotype and genotype values

之间, 决定系数最高达到 67.44%, 可以根据杂种标记型值来预测杂种性状表现。而利用杂种标记型值预测性状基因型值(朱军^[22,23])的效果更好, 全株有效果数标记型值可以解释该性状基因型值 85.66%的变异, 单株产量标记型值解释其基因型值 68.11%的变异。

不同的筛选方式, 得到的标记型值与表型值或基因型值间的相关性存在一定差异, 筛选的显著水平对相关系数也有一定影响, 与 0.01 水平下相比, 在 0.001 水平下筛选的标记, 多数性状的相关系数略有下降。

以基因型值进行筛选时, 标记型值与基因型值间的相关性在 9 个性状上都有较大提高, 相关系数在 0.7433~0.9255, 这可能是因为与性状表型值相比基因型值排除了部分随机误差的结果。

2.4 杂种表现预测模型

以杂种 F₁ 各标记位点的标记效应为自变量, 通过逐步回归分析建立了 9 个研究性状的预测方程(表 6)。由于在 0.01 显著水平下千粒重筛选到的标记位点数多达 85 个, 超过了杂交组合数目, 因此在建立千粒重的

表 6 性状表现预测模型

Table 6 Prediction model of hybrid performance

性状 Traits	回归方程 Regression equation	复相关系数 Multiple correlation coefficient	剩余误差 Remainder error	标记位点 Loci	
株高 Plant height	$y = 222.50 + 10.517X_1 + 10.031X_2 + 4.188X_3 - 6.073X_4 + 3.665X_5 + 4.248X_6 + 4.737X_7 + 7.103X_8 + 5.475X_9 - 8.989X_{10}$	0.835	5.9346	NAS22 NAS94-1 E3M3-4 E5M5-3 E7M5-7	NAS82 NAS94-2 E4M2-6 E6M3-9 E7M6-2
一次有效枝起点 Branch height	$y = 91.48 + 5.949X_1 + 3.455X_2 + 4.008X_3 + 6.038X_4 - 20.337X_5 + 5.615X_6 + 5.566X_7$	0.792	6.0572	NAS21 NAS41 E4M2-7 E7M1-2	NAS28 NAS82 E6M3-7
主序有效长 Length of main inflorescence	$y = 63.23 + 3.116X_1 + 3.525X_2 + 1.795X_3 + 1.609X_4 + 3.213X_5 + 4.064X_6 - 2.803X_7 + 1.929X_8 + 5.014X_9 + 1.635X_{10} - 3.090X_{11}$	0.839	2.8077	NAS8 NAS41 NAS96-2 E4M3-8 E7M2-10 E8M8-8	NAS21 NAS96-1 E4M2-14 E7M1-6 E7M6-1
一次有效枝数 No. of primary branches	$y = 9.77 + 0.253X_1 - 0.560X_2 + 0.938X_3 + 0.496X_4 - 1.742X_5 + 0.979X_6 + 0.360X_7 + 0.397X_8 + 0.668X_9 + 0.733X_{10} + 0.359X_{11} - 0.704X_{12} - 1.419X_{13} - 1.026X_{14}$	0.873	0.5016	NAS17 NAS41 NAS64 NAS99-1 E4M2-7 E4M4-12 E7M5-9	NAS33 NAS61 NAS98 E3M3-2 E4M2-14 E4M6-3 E7M6-2
全株有效果数 Siliques per plant	$y = 474.59 + 29.12X_1 - 58.990X_2 + 63.453X_3 + 54.359X_4 + 30.704X_5 + 43.150X_6$	0.859	43.5665	NAS99-1 E3M7-6 E5M5-6	NAS99-2 E5M4-7 E5M8-5
主序有效果数 Siliques of main inflorescence	$y = 84.67 + 3.960X_1 + 5.037X_2 + 6.807X_3 + 2.894X_4 + 8.485X_5 + 6.412X_6$	0.782	6.4280	NAS94 E3M3-11 E5M2-3	NAS98 E5M1-2 E5M8-5
每果粒数 Seeds per silique	$y = 20.33 - 2.649X_1 + 1.587X_2 + 1.156X_3 + 1.041X_4 + 1.5395X_5 + 1.209X_6 + 1.487X_7 + 1.024X_8$	0.819	1.4532	NAS41 E3M2-9 E4M3-2 E5M2-4	NAS82 E4M2-4 E4M6-4 E6M5-3
千粒重 1000-seed weight	$y = 3.71 + 0.144X_1 - 0.241X_2 + 0.181X_3 - 0.147X_4 + 0.288X_5 + 0.0898X_6 + 0.2983X_7 + 0.3028X_8$	0.856	0.2326	NAS31 NAS48 NAS99-3 E5M8-5	NAS34 NAS94 E8M2-9 E5M5-6
单株产量 Yield per plant	$y = 18.86 + 3.682X_1 + 1.251X_2 + 3.405X_3 - 3.239X_4 + 2.087X_5 + 2.788X_6$	0.796	3.0120	NAS45 E5M8-5 E6M3-7	E4M2-8 E5M8-7 E7M6-2

预测模型时,采用 0.001 显著水平下筛选的标记位点,其它性状的预测模型均采用 0.01 显著水平下筛选的标记位点。

9 个性状的预测方程中,标记效应与性状表现的复相关系数在 0.782~0.873。9 个性状预测方程涉及的标记位点在 6~14 个,单株产量、主序有效果数和全株有效果数的预测方程只含有 6 个标记位点,一次有效枝数的预测方程涉及了 14 个位点。千粒重、全株有效果数与单株产量的预测方程中有 1 个共同标记位点 E5M8-5,全株有效果数与千粒重的预测方程中有 2 个共同位点 E5M8-5 和 E5M5-6。

2.5 预测模型的评价

表 7 预测模型的评价

Table 7 Evaluation of prediction models

项目 Item	R^2	预测残差 Prediction residual	CV%	随机误差 Random error
株高 Plant height	0.5765	6.5104	3.02	6.8899
一次有效枝起点 Branch height	0.5322	6.4411	7.27	8.8271
主序有效长 Length of main inflorescence	0.5618	3.1518	5.02	4.7751
一次有效枝数 No. of primary branches	0.6238	0.5673	7.13	0.8424
全株有效果数 Siliques per plant	0.6783	46.0705	10.05	95.208
主序有效果数 Siliques of main inflorescence	0.5191	6.8401	7.78	8.8669
每果粒数 Seeds per silique	0.5426	1.6176	8.22	2.3083
千粒重 1000-seed weight	0.6508	0.2501	6.68	0.3059
单株产量 Yield per plant	0.5510	3.1865	14.00	5.975

R^2 为性状预测值对表型值的可决系数; CV% 为预测残差的变异系数; 随机误差指杂种 F_1 性状表现的试验误差

R^2 was determination coefficient of predictive value to phenotypic value. CV% was variation coefficient of prediction residual error. Random error was experimental error of hybrid performance

3 讨论

分子标记是生物遗传物质承载信息的外在表现,对某一个标记位点,其外在特征具有一定的稳定性和可考察性。本研究正是基于这一认识展开的。分子标记自身不是表达单元,但当某一标记位点与数量性状位点(QTL)相连锁时,则可以通过分子标记检测到相关 QTL 的效应。这里所说的分子标记的效应并不等同于 QTL 的效应,它应是 QTL 效应在分子标记与 QTL 连锁紧密程度影响下的一个结果。从某种意义上说,本研究是针对表达的性状相关 QTL 展开的。

研究中发现,无论杂交亲本分子标记一般遗传距离(GGD)还是特殊遗传距离(SGD)只在部分性状上与杂种表现的相关性达到显著或极显著水平,SGD 比 GGD 与杂种表现的相关性更强,但决定系数均较

9 个性状的预测值与表型值间的相关性都比较高,均高于杂种标记型值与表型值的相关性。利用 Jackknife 抽样技术对预测模型的稳定性进行检验。从 66 个杂种 F_1 中依次剔除一个组合,用其余 65 个组合重建回归方程,并预测剔除杂种 F_1 的性状表现,计算预测模型的可决系数 R^2 和预测残差(表 7)。9 个性状预测方程的可决系数 R^2 在 0.5191~0.6783,预测残差在 0.2501~46.0705,预测残差变异系数在 3.02%~14.00%。各性状预测方程的剩余误差(表 6)和预测残差都小于调查数据的随机误差。也说明了预测模型的稳定性强,预测误差较小,精确度高,具有实用价值。

小(最大 0.2668),不能用于杂种优势预测,这与前人的研究结果相一致^[9-11,17]。

利用亲本遗传距离(或遗传差异)预测杂种优势(距离模式)的基础是人们长期育种实践中的一种认知:在一定范围内亲本遗传差异越大,杂种优势越强。从以前的研究结论和本文的结果来看,这种模式并不理想。Bernardo^[15]在用计算机模拟的基础上认为,这种基于分子标记的杂交亲本间遗传差异预测杂种优势的准确性在很大程度上取决于 QTL 覆盖率,及与 QTL 连锁的标记比例。何光华等^[14]对遗传距离与杂种表现的相关性较低的原因作了分析,认为出现这一现象的原因大致有以下几点:①标记位点仅仅反应了基因组的很少一部分,导致只有有限比例的遗传因子的抽样是有效的;②大多数标记座位的等位基因可能与性状相关不密切;③环境效应的影响,QTL 存在环境效应。

在本研究中发现,在产量组成性状中 SGD 与每果粒数和千粒重极显著正相关,但与有效角果数极显著负相关,可能是这种相互抵消的结果导致最终与产量无相关。进一步研究中通过分析标记位点的效应,发现基因对性状表现的作用存在多种作用模式,有些位点只表现加性效应,而有些位点只表现显性效应,另一些位点则同时具有加性和显性效应。在同一模式中,不同位点的效应大小、作用方向也不一致,杂合对性状表现并非都是有利的。在遗传距离(差异)分析模式中,无论利用一般杂合性还是特殊杂合性,都是将所分析的位点等同看待,没有考虑各位点的效应大小和作用方向,这应该是预测效果不佳的主要原因。

与亲本分子标记遗传距离相比,杂种标记型值与杂种表型值的相关性有了很大提高,9个性状的相关系数在 0.6824~0.8113;可见通过估计特异标记位点的标记效应分析标记型值与表型值的关系,充分考虑了各作用位点的效应大小和作用方向,预测效果较好。

利用逐步回归分析建立的预测模型中,9个性状的预测方程涉及的标记位点在 6~14 个,大大减少了用于性状表现预测所需的标记数目,从而使利用分子标记预测性状表现变得更具有可操作性。预测值与杂种表现的相关系数在 0.782~0.873,普遍高于由筛选的全部相关位点得到的杂种标记型值与性状表现的相关性。说明利用方差分析所筛选的标记位点中确实存在一些干扰因子,影响了标记型值与杂种表现的相关性,利用逐步回归分析进一步对标记位点进行筛选是有效的。

这种通过 QTL 效应分析的方法由 Vuylsteke 等^[16]在对玉米杂种优势的研究中进行了探索,但在他的研究中把特异标记的效应等同于 QTL 效应,而且没有分析某种效应(加性或显性)存在的可能性,直接计算其效应值,这与传统的数量遗传学理论不一致。本研究提出了标记效应和标记型值的概念,在建立的杂种优势预测模型中首先应用了逐步回归分析,有效的排除了多种干扰因子,提高了预测效果,而且减少了预测模型中所涉及的标记位点,使之在育种实践中更具有实用性。

研究中 9 个性状的预测模型都没有考虑基因间的互作。在进一步的研究中如能在预测模型中包含互作位点,则不但会提高对杂种表现的预测效果,而且可以对杂交亲本的性状表现进行预测,从而把预测模型引入到常规育种中去,预测杂交组合后代纯合个体可能的表现,指导亲本选配及分子标记辅助选择。另外

研究中所建立的预测模型只是一个特定生长环境下的结果,没有进行多环境条件下的分析,模型中没有包含环境与基因型之间的互作效应,这也是今后要研究的重点。

4 结论

在估计特异标记位点效应的基础上,利用逐步回归分析建立了株高、一次有效枝起点、主序有效长、一次有效枝数、全株有效果数、主序有效果数、每果粒数、千粒重、单株产量等 9 个甘蓝型油菜产量相关性状杂种表现的预测模型。Jackknife 检验分析表明这些预测模型稳定性强,预测效果好。预测模型中涉及的标记位点少,有利于今后在育种实践中的应用。

References

- [1] Smith O S, Smith J S C, Bowen S L, Tenborg R A, Wall S J. Similarities among a group of elite maize inbreds as measured by pedigree: F₁ grain yield, grain yield, heterosis, and RFLPs. *Theoretical and Applied Genetics*, 1990, 80: 833-840.
- [2] Godshalk E B, Lee M, Lamkey K R. Relationship of restriction fragment length polymorphisms to single-cross hybrid performance of maize. *Theoretical and Applied Genetics*, 1990, 80: 273-280.
- [3] Dudley J W, Saghai Maroof M A, Rufenner G K. Molecular marker and grouping of parents in maize breeding programs. *Crop Science*, 1991, 31: 718-723.
- [4] Melchinger A E, Lee M, Lamkey K R, Woodman W L. Genetic diversity for restriction fragment length polymorphisms: relation to estimated genetic effects in maize inbreds. *Crop Scienc.*, 1990, 30: 1033-1040.
- [5] Melchinger A E, Lee M, Lamkey K R, Hallauer A R, Woodman W L. Genetic diversity for restriction fragment length polymorphisms and heterosis for two diallel sets of maize inbreds. *Theoretical and Applied Genetics*, 1990, 80: 488-496.
- [6] Xiao J, Li J, Yuan L, McCouch S R, Tanksley S D. Genetic diversity and its relationship to hybrid performance and heterosis in rice as revealed by PCR-based markers. *Theoretical and Applied Genetics*, 1996, 92: 637-643.
- [7] Dreisigacker S, Melchinger A E, Zhang P, Ammar K, Flachenecker C, Hoisington D, Warburton M L. Hybrid performance and heterosis in spring bread wheat, and their relations to SSR-based genetic distances and coefficients of parentage. *Euphytica*, 2005, 144: 51-59.
- [8] Yadav H K, Shukla S, Singh S P. Genetic divergence in parental genotypes and its relation with heterosis, F₁ performance and general

- combining ability (GCA) in opium poppy (*Papaver somniferum* L.). *Euphytica*, 2007, 157: 123-130.
- [9] 沈金雄, 陆光远, 傅廷栋, 杨光圣. 甘蓝型油菜遗传多样性及其与杂种表现的关系. 作物学报, 2002, 28(5): 622-627.
Shen J X, Lu G Y, Fu T D, Yang G S. Relationships between genetic diversity and hybrid performance in oilseed rape (*Brassica napus* L.). *Acta Agronomica Sinica*, 2002, 28(5): 622-627. (in Chinese)
- [10] 沈金雄, 傅廷栋, 杨光圣. 甘蓝型油菜 SSR、ISSR 标记的遗传多样性及其与杂种表现的关系. 中国农业科学, 2004, 37(4): 477-483.
Shen J X, Fu T D, Yang G S. Relationship between hybrid performance and genetic diversity based on SSR and ISSR in *Brassica napus* L.. *Scientia Agricultura Sinica*, 2004, 37(4): 477-483. (in Chinese)
- [11] Yu C Y, Hu S W, Zhao H X, Guo A G, Sun G L. Genetic distances revealed by morphological characters, isozymes, proteins and RAPD markers and their relationships with hybrid performance in oilseed rape (*Brassica napus* L.). *Theoretical and Applied Genetics*, 2005, 110: 511-518.
- [12] Zhang Q F, Gao Y J, Yang S H, Ragab R A, Saghai Maroof M A, Li Z B. A diallel analysis of heterosis in elite hybrid rice based on RFLPs and microsatellites. *Theoretical and Applied Genetics*, 1994, 89: 185-192.
- [13] Zhang Q F, Zhou Z Q, Yang G P, Xu C G, Liu K D, Saghai Maroof M A. Molecular marker heterozygosity and hybrid performance in indica and japonica rice. *Theoretical and Applied Genetics*, 1996, 93: 1218-1224.
- [14] 何光华, 侯磊, 李德谋, 罗小英, 牛国清, 唐梅, 裴炎. 利用分子标记预测杂交水稻产量及其构成因素. 遗传学报, 2002, 29(5): 438-444.
He G H, Hou L, Li D M, Luo X Y, Niu G Q, Tang M, Pei Y. Prediction of yield and yield components in hybrid rice by using molecular markers. *Acta Genetica Sinica*, 2002, 29(5): 438-444. (in Chinese)
- [15] Bernardo R. Relationship between single-cross performance and molecular marker heterozygosity. *Theoretical and Applied Genetics*, 1992, 83: 628-634.
- [16] Vuylsteke M, Kuiper M, Stam P. Chromosomal regions involved in hybrid performance and heterosis: their AFLP(R)-based identification and practical use in prediction models. *Heredity*, 2000, 85: 208-218.
- [17] Shen J X, Fu T D, Yang G S, Tu J X, Ma C Z. Prediction of heterosis using QTLs for yield traits in rapeseed (*Brassica napus* L.). *Euphytica*, 2006, 151: 165-171.
- [18] Nei M, Li W. Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceeding of the National Academy of Sciences of the United States of America*, 1979, 76: 5269-5273.
- [19] You J H, Chen G. Delete-group Jackknife estimate in partially linear regression models with heteroscedasticity. *Acta Mathematicae Applicatae Sinica* (English Series), 2003, 19(4): 599-610.
- [20] 许国. Jackknife 估计及其应用浅探. 山东师范大学学报(自然科学版), 2000, 15(4): 454-455.
Xu G. Exploration on the Jackknife estimating and its application. *Journal of Shandong Normal University* (Natural Science), 2000, 15(4): 454-455. (in Chinese)
- [21] 张南松, 陈军杰, 祝增荣. 利用多重抽样估计稀有物种的种类数. 生物数学学报, 2002, 17(4): 455-460.
Zhang N S, Chen J J, Zhu Z R. Estimating the number of rare species by multisampling. *Journal of Biomathematics*, 2002, 17(4): 455-460. (in Chinese)
- [22] 朱军. 作物杂种后代基因型值和杂种优势的预测方法. 生物数学学报, 1993, 8(1): 32-44.
Zhu J. Methods of predicting genotype value and heterosis for offspring of hybrids. *Journal of Biomathematics*, 1993, 8(1): 32-44. (in Chinese)
- [23] 朱军. 数量性状遗传分析的新方法及其在育种中的应用. 浙江大学学报: 农业与生命科学版, 2000, 26(1): 1-6.
Zhu J. New approaches of genetic analysis for quantitative traits and their application in breeding. *Journal of Zhejiang University* (Agricultural & Life Sciences), 2000, 26(1): 1-6. (in Chinese)

(责任编辑 于 竞)