

基于 F₁ 家系关联群体的野生大豆优异基因挖掘方法研究

常 玮,王 娟,于 洋,陈吉宝

(南阳师范学院 农业工程学院,河南 南阳 473061)

摘 要:为了研究 F₁ 家系关联群体在野生大豆优异基因挖掘中的效果,以大豆 HapMap 数据为基础,以大豆食心虫抗性为例,分别考虑不同遗传模型、基因数目、群体大小及遗传力大小,模拟自交系群体基因型及表型,并采用顶交的方式模拟生成 F₁ 家系关联群体,模拟其表型,在基因型与表型模拟的基础上,采用全基因组关联分析的方法比较二者的检测效力。经比较,F₁ 家系关联群体在 4 种遗传模型条件下的平均检出数目分别为 3.82,3.82,3.80 和 3.78,均显著高于自交系群体的平均检出数目(2.08,2.09,2.08 和 1.87)。分别在一定的群体大小、目标基因数目及遗传力大小下,比较自交系群体和 F₁ 家系关联群体的检测效力,结果表明,二者均随群体大小增大、基因数目的减少及遗传力的增大而提升,且在相同条件下,F₁ 家系关联群体的检测效力高于自交系群体。通过顶交的方式构建 F₁ 家系关联群体,可有效降低不良基因干扰,因此可大幅提高检测效力,从而使 F₁ 家系关联群体检测效果优于野生大豆资源群体。

关键词:野生大豆;优异基因;F₁ 家系关联群体;关联分析

中图分类号:S565.1 **文献标识码:**A **DOI:**10.11861/j.issn.1000-9841.2018.01.0050

Study of Mining Excellent Genes in *Glycine soja* by F₁ Family Association Population

CHANG Wei, WANG Juan, YU Yang, CHEN Ji-bao

(College of Agricultural Engineering, Nanyang Normal University, Nanyang 473061, China)

Abstract: In this study, we evaluated the detection effect of F₁ family association population based on soybean HapMap data, and made a case study of soybean pod borer resistance. The value of genotype and phenotype of both inbred line population and F₁ family association population were simulated under 4 different conditions, including genetic models (complete dominance and incomplete dominance), the number of genes, population size and heritability. Based on the simulation results, the detection effect of the two populations in GWAS was compared. As a result, the average numbers of detected genes by F₁ family association population in 4 different genetic models were 3.82, 3.82, 3.80 and 3.78, respectively, which were significantly higher than those of inbred line population (2.08, 2.09, 2.08 and 1.87). The result under different conditions also showed that the detection effect of both population were increased with the increasing of population size and heritability, reduced with the increasing of genes number, and the detection effect of F₁ population is higher than that of inbred population under the same conditions. Constructing the F₁ family association population by Top Cross could effectively reduce the interference of unfavorable allele, thereby make the F₁ population a better choice for genome-wide association study when mining the excellent alleles in *Glycine soja*.

Keywords: *Glycine soja*; Excellent genes; F₁ family association population; GWAS

一年生野生大豆 (*Glycine soja*) 是栽培大豆 (*Glycine max*) 的原始祖先种^[1]。由于在自然生境下几乎未受到人为选择的影响,因此很多丰富的功能基因并没有因为人为选择而丢失。野生大豆蕴藏着丰富的遗传多样性,近年来,研究人员对野生大豆所具有的诸如耐盐碱、干旱,多荚、多粒,高蛋白含量以及对大豆疫霉根腐病、大豆胞囊线虫、大豆灰斑病抗性等优异特性进行了深入研究^[2-6]。且由于野生大豆和栽培大豆之间不存在种间杂交障碍,无种间遗传隔离,因此,野生大豆可为栽培大豆培育提供丰富基因资源,是栽培大豆遗传改良不可多得的宝贵资源^[7]。尽管 GWAS 在野生大豆优异基因挖掘中取得了一定的成果,但由于野生大豆存在很多与优良性状连锁的不良性状,如果不能在关联群体表型鉴定时打破这种连锁,那么不仅会掩盖一些基因的效应,而且还会影响候选基因的应用价值。

针对上述问题,本研究提出了利用 F₁ 群体进行关联分析的设想,拟以大豆食心虫食率性状全基因组关联分析为例,通过假设与模拟,分别采用野生大豆资源群体和 F₁ 家系关联群体进行关联分析,最后通过比较二者的检测效力,验证 F₁ 关联分析群

体在挖掘野生大豆优异基因中的效果。 F_1 家系关联群体在关联分析中的成功应用,将为野生大豆优异等位基因挖掘提供新的策略。

1 材料与方法

1.1 数据下载及预处理

大豆 HapMap 数据(包含19 652个个体在52 041个位点上的分型结果),用于数据挖掘,来源于 Soy-Base (<http://soybase.org/snp>)^[8]。应用 Plink V 1.07 的质量控制功能来进行大豆 HapMap 数据的预处理^[9]。处理标准如下:剔除频率(minor allele frequency,MAF)小于1%稀有位点、缺失超过5%,且违背 Hardy-Weinberg 平衡($P < 1 \times 10^{-6}$)的 SNP 位点,以及缺失基因型、杂合基因型和未知基因型总数大于10%的个体。

1.2 提出假设

已知:大豆抗大豆食心虫的机制可分为不选择性(nonpreference,NP)、抗生性和耐害性(antibiosis and tolerance,A&T)^[10]。其中,抗生性和耐害性是指植物具有某些特性,对害虫的侵害表现出各种不利于害虫的反应,具有较高的利用潜力。不选择性主要包括了大豆一些形态上的性状,如茸毛有无、荚皮硬度、豆荚长度(pod length,PL)和百粒重(seed weight,SW)等性状。其中 PL 和 SW 与虫食率呈正相关,而这两个性状又是产量构成因子,与产量呈正相关。

根据上述研究基础,本研究提出如下假设:

(1)本研究以虫食率(herbivory rates,HR)来表示个体的抗虫能力。假设大豆的虫食率由 NP 和 A&T 引起,回归系数为 α_{NP} 和 $\alpha_{A\&T}$, $\alpha_{NP} + \alpha_{A\&T} = 1$;

(2)假设 NP 由 10 对显性等位基因决定,其效应相等,均为总效应的 1/10,A&T 由 n 对显性等位基因决定,其效应相等,均为总效应的 1/ n ,上述基因在基因组中均不存在连锁,且每个基因的两侧均各有一个遗传距离为 1 cM 的分子标记;

(3)假设 NP 相关基因在群体内的显性基因频率 $P = 0.2$,A&T 相关基因在群体内的显性基因频率 $P = 0.2$ 。裴颜龙等^[11]研究表明,野生大豆平均期望杂合度 $He = 0.133$,据此计算群体的基因型频率;

(4)本研究所涉及的群体均为遗传平衡群体,假设上述各基因、分子标记各有 2 个等位基因,其基因型频率符合 Hardy-Weinberg 平衡。

1.3 数据模拟

从19 652个个体中,随机抽选所要求数量的个

体组成目标群体。并在此基础上计算每个位点的基因频率和基因型频率,再从符合假设 iii,iv 的位点中分别抽选 10 个, n 个位点作为目标基因。最后随机选择在基因组中均匀分布的1 000个位点作为模拟群体的基因型。群体表型数据根据基因型数据模拟结果及假设 i,ii 进行计算。

用于比较的 F_1 家系群体是以上述目标群体为父本,以不具有抗性的栽培大豆为母本,采用顶交的方式获得。因此采用 F_1 家系关联群体进行关联分析时,基因型数据除目标基因外,其分子标记基因型与资源群体一致。再根据假设和模拟结果生成群体的基因型数据和表型数据,分别进行关联分析。

分别考虑下面 4 种遗传模型:第一,决定 NP 的 10 对基因和决定 A&T 的 n 对基因均为完全显性基因;第二,决定 NP 的 10 对基因为完全显性基因,决定 A&T 的 n 对基因均为共显性基因;第三,决定 NP 的 10 对基因为共显性基因,决定 A&T 的 n 对基因均为完全显性基因;第四,决定 NP 的 10 对基因和决定 A&T 的 n 对基因均为共显性基因。

在上述遗传模型的基础上,分别在群体数目为 100,200,300,500 的情况下, $\alpha_{A\&T}$ 分别为 0.2,0.4,0.6,0.8 的情况下以及 n 为 4,6,8,10 的情况下比较资源群体和 F_1 家系关联分析群体挖掘抗生性和耐害性相关基因的效果。本研究所涉及的 NP 相关基因在提高抗虫性的同时会降低大豆的产量且 A&T 相关基因具有较高的利用潜力,因此将以 A&T 相关基因定位准确性作为判断群体应用效果的判定标准。又由于本研究假设各个非等位基因具有相同的效应,故本研究以 A&T 相关基因被定位数目,即检出数目作为最终的标准。

上述各群体数目和 $\alpha_{A\&T}$ 的群体模拟分别在 4 种遗传模型下重复 50 次,并分别统计 50 次结果的平均值,即为平均检出数目,最后根据模拟结果和判定标准进行显著性检验。

1.4 全基因组关联分析

根据模拟结果,以每个性状的表型值进行性状与 SNP 标记的关联分析。SNP 标记/单倍型与性状的关联分析采用 TASSEL2.1 软件 MLM(mixed linear mode)程序,以 Structure 软件计算的 Q 值和 SPAGeDi1.2 软件计算的 Kinship 系数矩阵作为协变量使用 TASSEL 软件,根据所得 SNP 标记数据进行自交系性状与分子标记的关联分析,得到与品质及农艺性状相关联的分子标记;以 P_Maker 值来衡量关联标记的显著性。

2 结果与分析

2.1 大豆 HapMap 数据预处理

根据 Plink V 1.07 的质量控制功能来进行大豆 HapMap 数据的预处理。经预处理,最终可以进行模拟的数据包含 15 391 个个体在 42 449 个位点上的分型结果。将每个位点上比例 <50% 的位点作为 SNP,经统计,个体间 SNP 比率 R_{snp} 的范围为 0.089 ~ 0.531,平均值为 0.261。个体间 SNP 比率差异很大,变异系数为 27.3%,呈混合分布模式(图 1)。

2.2 群体基因型模拟

根据假设和群体大小进行大豆资源群体基因型数据模拟。当群体大小为 100,决定 A&T 的基因数 $n = 4$ 时,其中 NP 和 A&T 相关基因及其连锁 SNPs 在大豆基因组上分布且为 1 次重复的模拟结果如图 2 所示:上述 14 个基因分布于大豆全部 20 条染色体中的 11 条,且在基因组中不存在紧密连

锁,且每个基因的两侧均各有 1 个遗传距离为 1 cM 的分子标记。

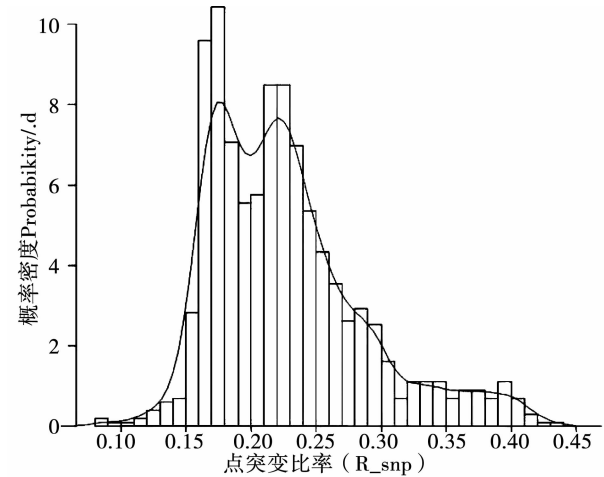


图 1 15 391 份大豆种质 SNP 比率分布图

Fig. 1 The distribution of SNP rate among 15 391 soybean accessions

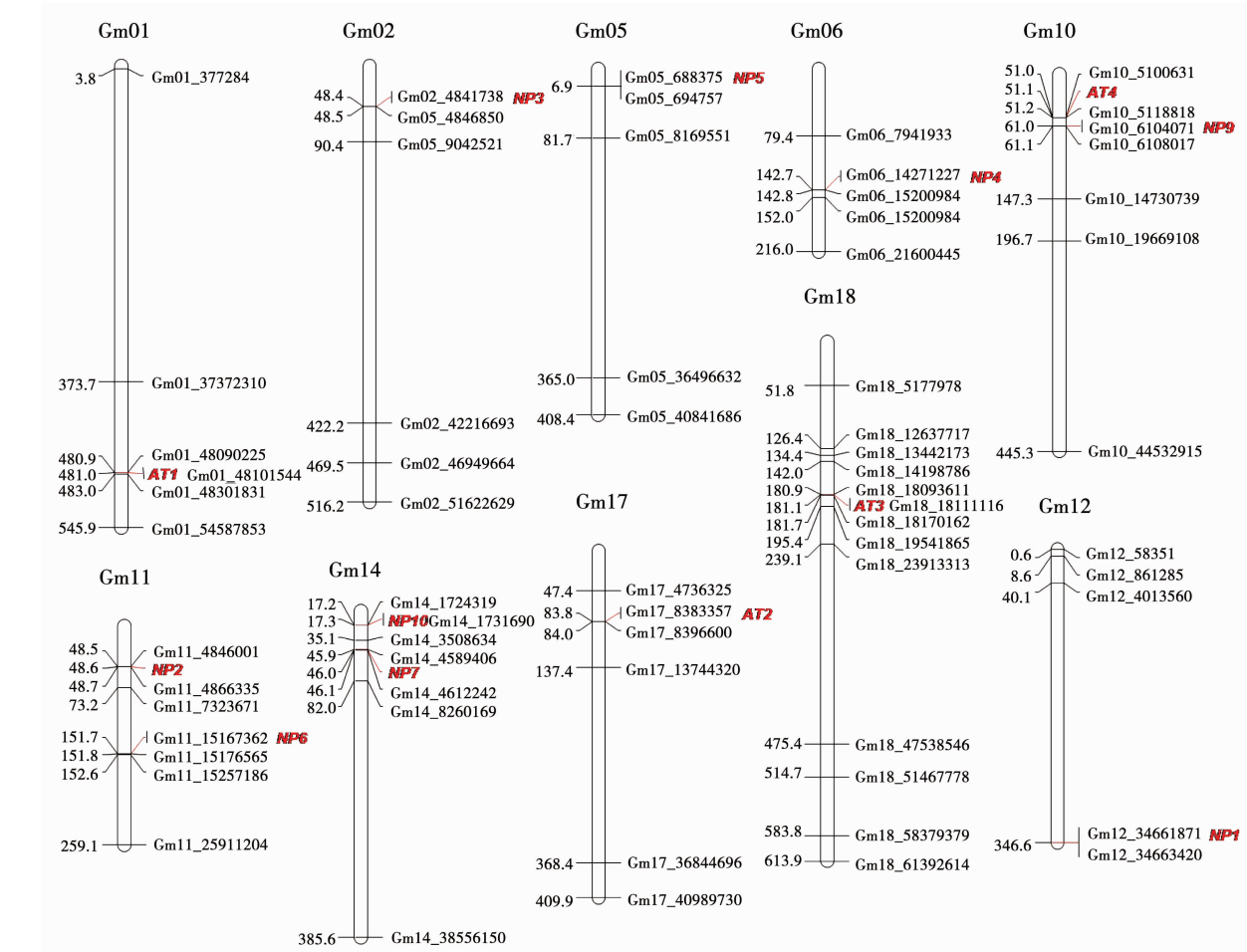


图 2 NP and A&T 在大豆基因组上的模拟结果

Fig. 2 The simulation of NP and A&T genes distribution on the soybean genome

根据上述基因在基因组上的分布模拟结果,100 份野生大豆资源基因型相关信息如表 1 所示,对于 4 个 A&T 相关基因,每个个体平均包含 0.76 个显性纯合基因、0.07 个杂合基因以及 3.17 个隐性纯

合基因。对于 10 个 NP 相关基因,每个个体平均包含 1.86 个显性纯合基因、0.11 个杂合基因以及 6.03 个隐性纯合基因。

表 1 100 份野生大豆资源基因型频率统计表

Table 1 Summary of genotype frequency for 100 *Glycine soja* individuals

基因 Genes		基因型频率 Genotype frequency			平均显性基因数 Mean number of dominant genes		
		AA	Aa	aa	AA	Aa	aa
A&T	A&T 1	0. 18	0. 03	0. 79	0. 76	0. 07	3. 17
	A&T 2	0. 18	0. 02	0. 80			
	A&T 3	0. 21	0. 01	0. 78			
	A&T 4	0. 19	0. 01	0. 80			
NP	NP1	0. 18	0. 01	0. 81	1. 86	0. 11	6. 03
	NP2	0. 18	0. 01	0. 81			
	NP3	0. 19	0. 01	0. 80			
	NP4	0. 20	0. 01	0. 79			
	NP5	0. 22	0. 01	0. 77			
	NP6	0. 18	0. 01	0. 81			
	NP7	0. 17	0. 01	0. 82			
	NP8	0. 19	0. 01	0. 80			
	NP9	0. 18	0. 02	0. 80			
	NP10	0. 17	0. 01	0. 82			

上述模拟结果为父本的 F₁ 关联分析群体基因型相关信息(表 1)。根据表 2 可知,对于 4 个 A&T 相关基因,杂交后,纯合显性基因仍为纯合显性,杂合基因发生分离,其中一半变为显性纯合,而纯合

隐性基因变为杂合基因。此时,上述 3 种基因平均数目分别 0. 80,3. 20 和 0. 00。NP 相关基因有 3 种基因型,每个个体所包含的显性纯合、杂合及隐性纯合平均基因数目与资源群体相同。

表 2 F₁ 关联群体基因型频率统计表

Table 2 Summary of genotype frequency for F₁ association population

基因 Genes		基因型频率 Genotype frequency			平均显性基因数 Mean number of dominant genes		
		AA	Aa	aa	AA	Aa	aa
NP	A&T1	0. 20	0. 80	0. 00	0. 80	3. 20	0. 00
	A&T2	0. 18	0. 82	0. 00			
	A&T3	0. 18	0. 82	0. 00			
	A&T4	0. 17	0. 83	0. 00			

2.3 群体表型模拟

根据基因型模拟结果,当群体大小为 100,决定 A&T 的基因数 $n = 4$ 、 $\alpha_{A\&T} = 0.5$ 时分别考虑 4 种不同的遗传模型,分别进行资源群体和 F₁ 关联群体的表型值模拟。当基因为不完全显性基因时,杂合型基因的效应为纯合显性基因的 1/2;当基因为完全显性基因时,杂合型基因的效应与纯合显性基因的效应相等。此时,在 4 种模型下群体表型值模拟结果如图 3 所示,由于野生大豆资源群体的纯合显性基因和杂合基因频率较小,因此 4 种遗传模型下的表型值在群体中的分布无明显变化,平均值为

5. 00,标准差为 1. 10。F₁ 群体中 NP 性状杂合基因频率显著增加,当基因为不完全显性时,NP 性状对表型的贡献部分降低,此时群体表型平均值为 3. 01,标准差为 0. 99;当基因为完全显性时,NP 性状对表型无贡献,此时群体表型平均值为 1. 02,标准差为 1. 00。对 2 种群体在 4 种遗传模型下的变异系数进行比较,方差分析结果表明,F₁ 群体的在 4 种遗传模型下的平均变异系数 65. 4% 显著高于野生大豆资源群体平均变异系数 22. 1% ($P = 0. 032\ 4$)。上述结果表明,F₁ 群体具有减少 NP 性状干扰和增大群体变异的作用。

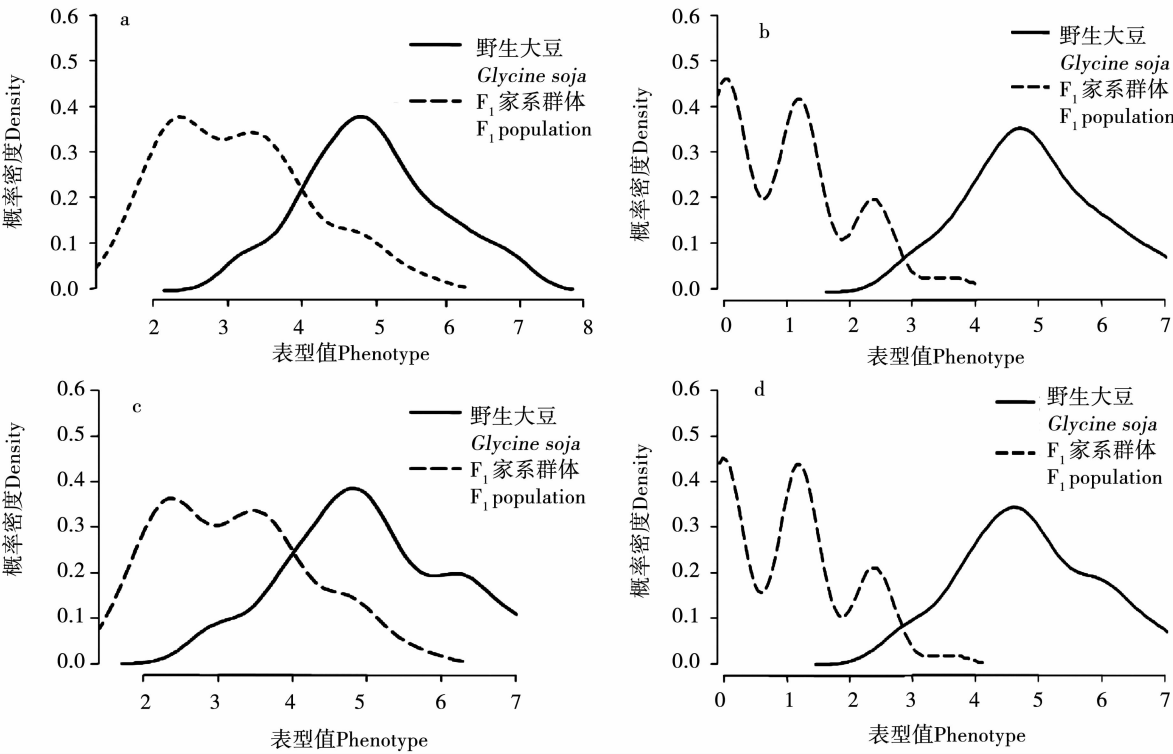


图3 野生大豆资源及 F₁ 群体 4 种遗传模型表型值模拟结果

Fig. 3 The phenotype simulation of *Glycine soja* and F₁ population under 4 genetic models

2.4 全基因组关联结果比较分析

采用 TASSEL V 2.1 软件 MLM 模型程序,分别采用 100 份野生大豆资源和 F₁ 群体进行食心虫抗性的 GWAS,2 个群体在 4 种遗传模型下重复 50 次,结果见表 3。采用野生大豆资源群体,在 4 种遗传模型条件下,分别检测到 7,5,8 和 5 个与抗虫性显

著关联($P < 0.0001$)的 SNP,其中与 A&T 性状相关的位点数分别为 2,2,2 和 2 个(表 3)。采用 F₁ 群体,在 4 种遗传模型条件下,分别检测到 5,5,5 和 5 个与抗虫性显著关联($P < 0.0001$)的 SNP,其中与 A&T 性状相关的位点数分别为 4,4,3 和 4 个(表 3)。

表 3 资源群体及 F₁ 群体关联分析结果

模型 Model	资源群体 Inbred line population		F ₁ 家系关联群体 F ₁ population	
	标记 Marker	基因 Gene	标记 Marker	基因 Gene
完全 + 完全 Complete + complete	Gm07_12925872		Gm01_48087483	A&T1
	Gm08_40256738		Gm10_5113345	A&T4
	Gm15_11909660		Gm16_36305973	
	Gm16_36305973		Gm17_8384370	A&T2
	Gm17_8384370	A&T2	Gm18_18104071	A&T3
	Gm18_18104071	A&T3		
完全 + 不完全 Complete + incomplete	Gm18_2684576			
	Gm15_11909660		Gm01_48087483	A&T1
	Gm16_36305973		Gm10_5113345	A&T4
	Gm17_8384370	A&T2	Gm16_36305973	
	Gm18_18104071	A&T3	Gm17_8384370	A&T2
	Gm18_2684576		Gm18_18104071	A&T3
不完全 + 完全 Incomplete + complete	Gm07_12925872		Gm01_48087483	A&T1
	Gm07_14230590		Gm07_12925872	

续表 3

模型 Model	资源群体 Inbred line population		F_1 家系关联群体 F_1 population	
	标记 Marker	基因 Gene	标记 Marker	基因 Gene
不完全 + 不完全 Incomplete + incomplete	Gm08_40256738		Gm16_36305973	
	Gm15_11909660		Gm17_8384370	A&T2
	Gm16_36305973		Gm18_18104071	A&T3
	Gm17_8384370	A&T2		
	Gm18_18104071	A&T3		
	Gm18_2684576			
	Gm15_11909660		Gm01_48087483	A&T1
	Gm16_36305973		Gm10_5113345	A&T4
	Gm17_8384370	A&T2	Gm16_36305973	
	Gm18_18104071	A&T3	Gm17_8384370	A&T2
	Gm18_2684576		Gm18_18104071	A&T3

斜体加粗字体表示相关基因及相邻分子标记。
Bold font means related genes and associated molecular markers.

对每个遗传模型下的资源群体和 F_1 群体的 A&T 性状相关位点检测效率进行比较, F_1 群体在 4 种遗传模型条件下的平均检出数目分别为 3.82, 3.82, 3.80 和 3.78, 均显著高于资源群体的平均检

出数目, 分别为 2.08, 2.09, 2.08 和 1.87(图 4)。上述结果表明, 在满足上述假设的前提下, F_1 关联群体对于 A&T 性状相关位点检测效果优于野生大豆资源群体。

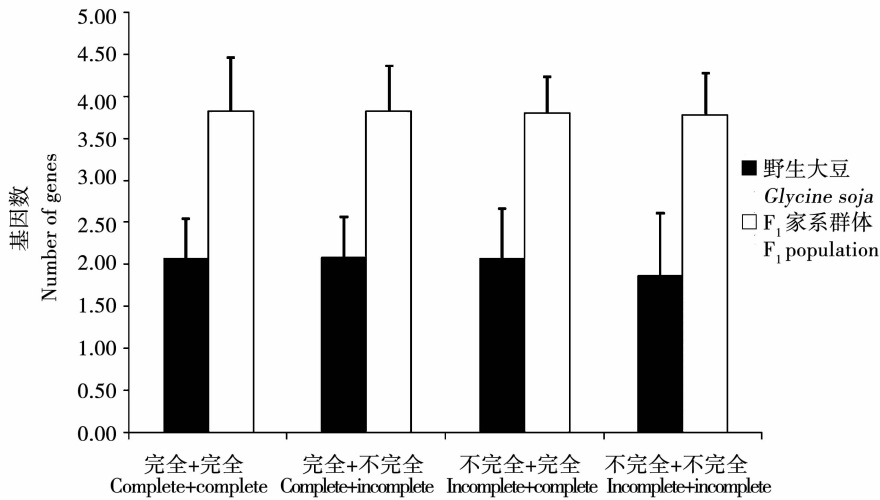


图 4 野生大豆资源及 F_1 群体 4 种遗传模型位点检测效率比较

Fig. 4 The comparison of detection efficiency for *Glycine soja* and F_1 population under 4 genetic models

2.5 影响 F_1 家系关联群体检测效果的因素

为了阐明影响 F_1 关联群体检测效果的因素, 本研究分别从群体大小、目标基因数目及贡献率大小 3 个方面来对两种群体检测效果进行比较, 结果如图 5 所示。

当遗传力为 0.5, 目标基因数目为 4 个时, 随着群体大小的增大, 自交系群体的检测效力分别为 0.52, 0.61, 0.70 和 0.79, F_1 群体的检测效力分别 0.96, 1, 1 和 1(图 5a), 表明两个群体的检测效力都

随群体大小的增大而提升, 且 F_1 群体的检测效力均高于自交系群体。

当群体大小为 100, 遗传力为 0.5 时, 随着目标基因数目的增多, 自交系群体的检测效力分别为 0.52, 0.43, 0.34 和 0.26, F_1 群体的检测效力分别 0.96, 0.91, 0.87 和 0.72(图 5b), 表明两个群体的检测效力都随基因数目的增加而下降, 且 F_1 群体的检测效力均高于自交系群体。

当群体大小为 100, 目标基因数目为 4 个时, 随

着遗传力的增大,自交系群体的检测效力分别为 0.26,0.53,0.65 和 1.00, F_1 群体的检测效力分别为 0.70,0.96,1.00 和 1.00 (图 5c),表明两个群体的检测效力都随遗传力的增加而增大,且 F_1 群体的检测效力均高于自交系群体。

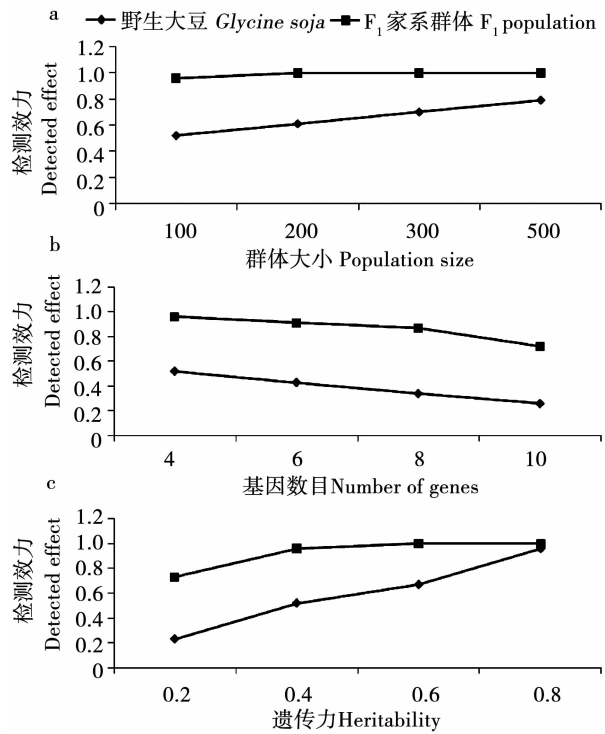


图 5 野生大豆资源及 F_1 群体分别在不同群体大小、基因数目及遗传力下检测效力比较
Fig. 5 The comparison of detection efficiency for *Glycine soja* and F_1 population under different population size, gene numbers and heritability

3 讨论

本研究采用顶交的方式来构建 F_1 家系群体,与自交系群体相比,改变了群体的基因频率和基因型频率。如表 1 和 2 所示,杂交使 F_1 群体 NP 基因的显性基因频率大幅增加,其最终结果表现为减少 NP 性状干扰从而增加了群体的变异。Sesay 等^[12]对玉米 10 个性状的研究结果也表明,通过顶交可以增大多数性状的变异度,从而为选育新品种提供了更大的遗传改良尺度。资源群体及 F_1 群体关联分析的结果中, F_1 群体表型值的分布曲线可明显看出有 4 个峰,表明由 4 个基因控制该性状,其在 4 种遗传模型下的平均变异系数 65.4%,显著高于野生大豆资源群体平均变异系数 22.1% ($P = 0.0324$),这就是为什么野生大豆资源及 F_1 群体 4 种遗传模型位点检测效率比较中 F_1 关联群体对于 A&T 性状相关位点检测效力优于野生大豆资源群体的主要原因;第二,对比野生大豆资源及 F_1 群体在不同群体大小、基因数目及遗传力条件下的定位结果,我们可以看

出,与自交系群体相比, F_1 家系关联群体较小群体大小、较多目的基因以及较小遗传力的条件下也能获得较高的检测效力;第三,一旦通过 F_1 家系关联群体定位了目的基因,马上即可以选择具有目的基因的家系进行自交,并应用由其衍生的 F_1 大群体进行定位研究^[13]。

从野生大豆资源及 F_1 群体分别在不同群体大小下检测效力比较的结果中我们可以发现,无论是采用自交系群体,还是 F_1 家系群体,群体的大小、目标基因数目以及目标性状遗传力大小,都会影响关联分析的结果,这与前人的研究结果是一致的。Wang 等^[14]系统研究了群体大小对大麦 GWAS 结果的影响。其研究表明,随着群体数目从 766 逐渐减小为 672,576,480,384 及 96,检测到的位点数目逐渐减少。当群体大小为 96 时,只有位点 QHd2H.064 可以被检测到,并得出群体数大于 384 时才能检测到全部的位点。在本研究中随着群体大小的增大,自交系群体及 F_1 群体的检测效力都随群体大小的增大而提升,且 F_1 群体可在群体较小时获得较高的检测效力。

野生大豆资源及 F_1 群体分别在不同基因数目的检测效力比较结果表明无论是自交系群体还是 F_1 家系群体的检测效力,都随性状遗传力的增加而提高。何小红等(2001)对数量性状基因作图精度的主要影响因子进行了研究,表明随着遗传力提高,发现 QTL 的能力也会随之提高^[15]。赵洪波等^[16]关于性状遗传力对 F_2 设计下 QTL 定位效果的影响研究,以及王孝义等^[17]采用 GBLUP (gene-assisted best linear unbiased prediction, GBLUP) 对不同遗传力性状选择效果研究结果,均表明数量性状基因检测效力会随性状遗传力的增加而提升。本研究的结果也证实了性状遗传力对关联分析结果产生影响,同样表现为基因检测效力会随性状遗传力的增加而提升。

从野生大豆资源及 F_1 群体分别在不同遗传力的检测效力比较结果表明自交系群体和 F_1 家系群体的检测效力,都随着目标性状基因数目的增加而降低。其主要原因是多基因性状易受到环境的影响,且不同基因之间的互作效应也会增强,因此目标性状基因越多,其检测效力会越低。王军等^[18]对大麦 DH 群体的数量性状进行遗传分析的结果也证实了这一结论。

本研究的结果是在一系列的假设的基础上进行模拟的。首先,假定大豆的虫食率仅由 NP 和 A&T 性状决定。而实际中还会有由稀有变异、结构变异、上位性互作及基因与环境互作而造成的丢失遗传率 (missing heritability)。其次,每个性状的相

关基因也不可能都具有相同的效应。本研究为计算方便而假设其具有相同的遗传效应。此外本研究的遗传模型也只假设了完全显性和不完全显性两种情况。尽管如此,由于本研究是在相同的条件下对自交系群体及其衍生的 F₁ 家系群体进行的比较,因此在实际应用中具有一定参考价值。

本研究所提及的方法可具有以下两个方面的应用价值:第一,由于该方法可有效降低不良基因的干扰,从而大幅提高检测效力,这对野生大豆等具有许多优良基因同时又有许多与之连锁的不良基因的材料来说是非常好的选择;第二,由于 F₁ 家系群体可在群体大小较小时具有较高的检测效力,因此该方法也适用于群体数量较小或基因型及表型性状调查工作量大的研究。

最后,在应用该方法时,选择合适的母本材料进行顶交试验设计也是该方法能够成功实施的前提。

4 结 论

通过顶交的方式构建 F₁ 家系关联群体,可使大豆食心虫 NP 基因的显性基因频率大幅增加,降低了 NP 性状的表型方差,从而有效降低不良基因的干扰,因此可大幅提高检测效力。关联分析的结果表明 F₁ 关联群体对于 A&T 性状相关位点检测效果优于野生大豆资源群体,且 F₁ 家系关联群体的检测效力随群体大小及性状遗传力的的增加而提高,而随着目标性状基因数目的增加而降低。该方法可有效应用于野生大豆优异基因的挖掘研究。

参考文献

[1] 徐豹. 中国野生大豆(*G. soja*)研究十年[J]. 吉林农业科学, 1989, 1: 5-13. (Xu B. A ten years studying of *Glycine soja* (*G. soja*) in China[J]. Journal of Jilin Agricultural Sciences, 1989, 1: 5-13.)

[2] Pazdernik D L, Hartman G L, Huang Y H, et al. A greenhouse technique for assessing *Phytophthora* root rot resistance in *Glycine max* and *G. soja*[J]. Plant Disease, 1997, 81(10): 1112-1114.

[3] Wang D, Diers B W, Arelli P R, et al. Loci underlying resistance to race 3 of soybean cyst nematode in *Glycine soja* plant introduction 468916[J]. Theoretical and Applied Genetics, 2001, 103(4): 561-566.

[4] Shi H. Studies on the drought resistance of wild soybean germplasm[J]. Soybean Science, 2003, 22(4): 5.

[5] Lee J D, Shannon J G, Vuong T D, et al. Inheritance of salt tolerance in wild soybean (*Glycine soja* Sieb. and Zucc.) accession PI483463[J]. Journal of Heredity, 2009: esp027.

[6] 马淑梅, 韩新华. 野生大豆资源对灰斑病抗性鉴定与评价[J]. 中国农学通报, 2015, 31(30): 86-91. (Ma S M, Han X H. Evaluation and identification of resistance of wild soybean germplasm to *Cercospora sojina*[J]. Chinese Agricultural Science Bulletin, 2015, 31(30): 86-91.)

[7] Concibido V, La Vallee B, McIaird P, et al. Introgression of a quantitative trait locus for yield from *Glycine soja* into commercial soybean cultivars[J]. Theoretical and Applied Genetics, 2003, 106(4): 575-582.

[8] Song Q, Hyten D L, Jia G, et al. Development and evaluation of Soy SNP50K, a high-density genotyping array for soybean[J]. PLoS One, 2013, 8(1): e54985.

[9] Purcell S, Neale B, Todd-Brown K, et al. PLINK: A tool set for whole-genome association and population-based linkage analyses[J]. American Journal of Human Genetics, 2007, 81(3): 559-575.

[10] 赵桂云, 李红丽, 王宇, 等. 大豆抗大豆食心虫机制研究进展[J]. 吉林农业, 2013(3): 73, 72. (Zhao G Y, Li H L, Wang Y, et al. Advances in study on the mechanism of pod borer resistance in soybean[J]. Jilin Agriculture, 2013(3): 73, 72.)

[11] 裴颜龙, 王岚, 葛颂, 等. 野生大豆遗传多样性研究 I: 4 个天然居群等位酶水平的分析[J]. 大豆科学, 1996, 25(4): 302-309. (Pei Y L, Wang L, Ge S, et al. Study on the genetic diversity of wild soybean I: Allozyme analysis of four natural populations[J]. Soybean Science, 1996, 25(4): 302-309.)

[12] Sesay S, Ojo D K, Ariyo O J, et al. Genetic variability, heritability and genetic advance studies in top-cross and three-way cross maize (*Zea mays* L.) hybrids[J]. Maydica, 2016, 61(2): M12, 1-7.

[13] 周坤华, 雷刚, 方荣, 等. 利用辣椒种间 F₂ 和 F_{2:3} 两个群体进行其主要农艺性状 QTL 分析[J]. 园艺学报, 2015, 42(5): 879-889. (Zhou K H, Lei G, Fang R, et al. Detection of QTLs for main agronomic traits using F₂ and F_{2:3} inter specific populations in pepper[J]. Acta Horticulturae Sinical, 2015, 42(5): 879-889.)

[14] Wang H, Smith K P, Combs E, et al. Effect of population size and unbalanced data sets on QTL detection using genome-wide association mapping in barley breeding germplasm[J]. Theoretical and Applied Genetics, 2012, 124(1): 111-124.

[15] 何小红, 徐辰武, 蒯建敏, 等. 数量性状基因作图精度的主要影响因子[J]. 作物学报, 2001, 28(4): 469-475. (He X H, Xu C W, Kuai J M, et al. Principal factors affecting the power of detection and accuracy of QTL mapping[J]. Acta Agronomica Sinica, 2001, 28(4): 469-475.)

[16] 赵洪波, 李明丽, 鲁绍雄, 等. 群体规模和性状遗传力对 F₂ 设计下 QTL 定位效果的影响[J]. 云南农业大学学报, 2007, 22(2): 159-163. (Zhao H B, Li M L, Lu S X, et al. Study on the effects of population size and trait heritability on the accuracy of QTL mapping under F₂ design[J]. Journal of Yunnan Agricultural University, 2007, 22(2): 159-163.)

[17] 王孝义, 李明丽, 刘刚, 等. 基因型选择和基因辅助 BLUP 对不同遗传力性状的选择效果[J]. 云南农业大学学报(自然科学), 2013, 28(6): 796-803. (Wang X Y, Li M L, Liu G, Lu S X. Effects of genotype selection and gene-assisted BLUP for selection on the traits with different heritabilities[J]. Journal of Yunnan Agricultural University (Natural Sciences), 2013, 28(6): 796-803.)

[18] 王军, 周美学, 黄祖六, 等. 大麦 DH 群体若干数量性状的遗传分析[J]. 扬州大学学报(农业与生命科学版), 2006, 27(3): 65-69. (Wang J, Zhou M X, Huang Z L, et al. Genetic analysis of quantitative traits of a doubled haploid population in barley[J]. Journal of Yangzhou University (Agricultural and Life Science Edition), 2006, 27(3): 65-69.)