

• 导读 •

开放科学（资源服务）标识码（OSID）：



## 限制性两阶段多位点全基因组关联分析法（RTM-GWAS） 的特点、常见提问与应用前景

盖钧锰，贺建波

（南京农业大学大豆研究所/国家大豆改良中心/农业部大豆生物学与遗传育种重点实验室/作物遗传与种质创新国家重点实验室/  
江苏省现代作物生产协同创新中心，南京 210095）

**摘要：**限制性两阶段多位点全基因组关联分析方法（RTM-GWAS）是新建立的一种可以全面检测自然群体和双（多）亲衍生群体中具有不同复等位变异 QTL 体系的关联分析方法。本文介绍了提出 RTM-GWAS 的出发点及其两大主要特点，包括建立适合自然群体和双（多）亲衍生群体特点的复等位变异标记和控制总体贡献率的多位点关联分析模型。一般读者和编者对 RTM-GWAS 的方法、原理，对复等位标记和多位点模型并无异议，提问与质疑主要分为两方面：一是 RTM-GWAS 检测到的 QTL 数量较多，大大多于单位点 MLM 模型所检出的 QTL 数目，怀疑增加的 QTL 是假阳性所致；另一是采用常规显著水准要求太低，不适于关联分析。本文对此做了严密释疑。最后介绍了关于 RTM-GWAS 的应用前景，包括遗传体系解析与重要基因克隆，双（多）亲杂交衍生群体遗传解析，群体遗传分化与进化和设计育种等方面。

**关键词：**限制性两阶段多位点全基因组关联分析；SNP 连锁不平衡区段；复等位变异；多位点模型；QTL-allele 矩阵；假阳性；模型显著性

## Major Characteristics, Often-Raised Queries and Potential Usefulness of the Restricted Two-Stage Multi-Locus Genome-Wide Association Analysis

GAI JunYi, HE JianBo

(Soybean Research Institute, Nanjing Agricultural University/National Center for Soybean Improvement/Key Laboratory of Biology and Genetic Improvement of Soybean (General), Ministry of Agriculture/State Key Laboratory for Crop Genetics and Germplasm Enhancement/Jiangsu Collaborative Innovation Center for Modern Crop Production, Nanjing 210095)

**Abstract:** Restricted two-stage multi-locus genome-wide association analysis (RTM-GWAS) is a novel GWAS procedure which provides a way to identify the QTL system with various multiple alleles in natural and bi- or multi-parental derived populations. The major purposes and its two major characteristics of the RTM-GWAS procedure were presented, including the establishment of the SNPLDB markers with multiple alleles fitting the property of the natural and bi- or multi-parental derived populations and the establishment of multi-locus model GWAS procedure with the total genetic contribution controlled within heritability value. Generally, the readers and editors do not doubt about the methods and principles, the multiple allele markers and the multi-locus model, but have questions and queries on the large amount of detected QTLs many more than those from single locus

收稿日期：2020-01-02；接受日期：2020-02-15

基金项目：国家自然科学基金（31701447）、国家作物育种重点研发计划（2017YFD0101500，2017YFD0102002）、长江学者和创新团队发展计划（PCSIRT\_17R55）、教育部 111 项目（B08025）、中央高校基本科研业务费项目（KYT201801）、农业部国家大豆产业技术体系 CARS-04、江苏省优势学科建设工程专项、江苏省 JCIC-MCP 项目

联系方式：盖钧锰，E-mail: sri@njau.edu.cn

MLM-GWAS procedure and on the general significance level without correction used in RTM-GWAS. These doubts were carefully and seriously explained and relieved. Furthermore, the potential usefulness of the RTM-GWAS procedure in genetic and evolutionary studies were summarized, including usefulness in relatively thorough identification of the QTL-allele system in populations and major gene finding and cloning, usefulness in relatively thorough identification of the QTL-allele system in bi-and multi-parental derived populations, usefulness in studies on population genetic differentiation and evolution and usefulness in breeding by genetic design.

**Key words:** restricted two-stage multi-locus genome-wide association analysis (RTM-GWAS); SNP linkage disequilibrium block (SNPLDB); multiple alleles; multi-locus model; QTL-allele matrix; false positive; model significance

为克服全基因组关联分析 (genome-wide association study, GWAS) 在遗传育种研究中的局限性, HE 等<sup>[1]</sup>提出了限制性两阶段多位点全基因关联分析 (restricted two-stage multi-locus genome-wide association analysis, RTM-GWAS) 方法以全面解析数量性状 QTL-等位变异构成。目前, RTM-GWAS 方法已应用于多个群体遗传研究<sup>[2-9]</sup>。受编辑部邀约, 本课题组将以 RTM-GWAS 在大豆遗传与种质资源中的研究为例, 用以说明该方法的应用前景。该专题包括限制性两阶段多位点全基因组关联分析法在遗传育种中的应用<sup>[10]</sup>、东北大豆种质群体百粒重 QTL-等位变异的全基因组解析<sup>[11]</sup>、RTM-GWAS 方法应用于大豆 RIL 群体百粒重 QTL 检测的功效<sup>[12]</sup>、大豆巢式关联作图群体蛋白质含量的遗传解析<sup>[13]</sup>和大豆重组自交系群体异黄酮含量 QTL 连锁定位与关联定位的比较研究<sup>[14]</sup>等 5 篇应用文章。除第一篇简单介绍 RTM-GWAS 的原理外, 其余 4 篇介绍了在资源群体、双 (多) 亲衍生群体 (RIL, NAM) QTL 及其等位变异的定位与检测以及所获 QTL-Allele 数据在基因发掘和群体遗传研究中的应用。为便于读者了解 RTM-GWAS 方法及其应用, 本文着重介绍提出 RTM-GWAS 的出发点及其两大主要特点、编审过程中对 RTM-GWAS 的质疑及辩解以及 RTM-GWAS 方法的应用前景。

## 1 提出 RTM-GWAS 的出发点及其两大主要特点

全基因组关联分析利用自然群体广泛存在的遗传变异, 通过测验分子标记与表型间相关性来检测数量性状基因座 (quantitative trait loci, QTL), 为全面解析数量性状遗传体系提供了有效手段。GWAS 通常基于单核苷酸多态性 (single nucleotide polymorphism, SNP) 分子标记, 单个 SNP 标记一般仅有 2 个等位变异, 因此也只能完全拟合仅有 2 个等位变异的 QTL。然而, 自然群体 (包括种质资源群体) 长期经历多种环境的影响, 同一基因座上会产生新的等位变异, 形

成复等位变异。多亲衍生群体因亲本间的异质性, 也会有复等位变异出现。这与双亲分离群体 (例如重组自交系群体) 中每个 QTL 仅有 2 个等位变异不同。因此, 利用 SNP 标记不能检测自然群体和多亲衍生群体中广泛存在的复等位变异, 这可能会降低 GWAS 的检测功效。另外, 植物常规育种是聚合优异等位变异的遗传操作过程<sup>[15]</sup>, 获得每个 QTL 的复等位变异及其效应估计是分子标记辅助选择的必要前提, 因此, 利用仅有 2 个等位变异的 SNP 标记一定程度限制了 GWAS 在育种中的应用。

常用的 GWAS 方法, 例如混合线性模型 (mixed linear model, MLM) 方法<sup>[16]</sup>, 一般基于单位点模型检测 QTL, 即每个标记位点与表型的相关性测验彼此独立进行。然而实际上数量性状受大量的 QTL 控制, 单位点模型中位点效应和贡献率的估计必然受到相邻 QTL 的影响, 进而使得 GWAS 受到干扰。单位点效应 (贡献) 过高估计可能导致 QTL 总体表型变异解释率超过性状遗传率或甚至超过 100%<sup>[7]</sup>。为了控制这种膨胀, 统计学家提出在个别位点测验的基础上进行全试验的总体测验, 例如 Bonferroni 方法将测验阈值设为显著水平除以标记数目。鉴于 GWAS 涉及全基因组高密度 SNP 分子标记, 要使全试验显著阈值保持常规水准 ( $P=0.05$ ), Bonferroni 方法就必然对每一个标记的阈值设置很高 ( $-\lg P$  值很大)。尽管严格的测验阈值有效降低了全试验错误率, 但同时也导致了较高的假阴性, 以至于 GWAS 往往仅能检测到少数 QTL, 只能解释遗传变异的很小部分, 不能充分检出性状的全基因组 QTL。

RTM-GWAS 是新建立的一种可以全面检测自然群体和双 (多) 亲衍生群体中具有不同复等位变异 QTL 体系的关联分析方法。该方法采用了复等位变异标记和多位点模型作为解决上述两大问题的关键, 形成 RTM-GWAS 的两大主要特点, 目的在于全面解析群体数量性状 QTL-等位变异的遗传构成<sup>[1,17]</sup>。首先, RTM-GWAS 通过构建具有复等位变异的 SNP 连锁不平衡区段 (SNP linkage disequilibrium block, SNPLDB) 标记以检测自然群体中的复等位变异。SNPLDB 标记

复等位变异丰富, 可以拟合 QTL 上多个等位变异, 比仅有 2 个等位变异的 SNP 标记更符合自然群体特征, 从而提高了检测功效。基于 SNPLDB 标记的连锁不平衡衰减距离比 SNP 标记更短, 因此, 还可能提高检测精度。其次, RTM-GWAS 通过多位点复等位变异模型以检测全基因组 QTL 及其复等位变异, 以遗传率值作为检出 QTL 总贡献率的上限, 使假阳性事件得到合理的控制。为提高 QTL 检测效率, RTM-GWAS 还通过两阶段分析策略以降低多位点模型运算量, 并利用基于 SNPLDB 标记的遗传相似系数矩阵控制群体结构偏差导致的假阳性。鉴于多位点模型的显著性测验代表了全模型测验, 不需要再做全试验多重测验, 无需对显著水平进行矫正, 因而同样采用常规显著水平(例如 0.01 或 0.05), 实际的 QTL 检出数大大提高, 一定程度上避免了由于过严格显著水平导致的假阴性问题。此外, RTM-GWAS 还与表型鉴定试验设计紧密结合, 可直接用试验设计原始数据(包括环境和区组)进行分析, 因而 RTM-GWAS 将试验设计严格的误差控制和关联分析紧密结合, 降低了试验误差, 提高了 QTL 检出能力。详细说明请参考文献<sup>[1-9,17]</sup>。

## 2 RTM-GWAS 方法的常见质疑与辩解

RTM-GWAS 方法及其应用文章发表过程中, 审稿人对上述方法、原理以及对复等位标记和多位点模型的 2 个主要特点一般并无异议, 但有怀疑与质疑。质疑最多的问题, 主要有两方面: 一是 RTM-GWAS 检测到的 QTL 数量较多, 大大多于传统的单位点 MLM 模型所检出的 QTL 数目, 怀疑多出来的 QTL 是假阳性所致; 二是采用常规显著水准, 阈值太低, 不适于关联分析。

### 2.1 关于检测的 QTL 数量多, 怀疑假阳性高的辩解

与其他常用 GWAS 方法相比, RTM-GWAS 方法通常能检测更多的 QTL, 这符合设计 RTM-GWAS 方法的初衷。该方法的目的是要将遗传率所反映的 QTL 遗传变异尽可能多的挖掘出来, 但总贡献率不应超过全试验的遗传率值。这是通过多位点模型分析中检测模型的显著性实现的。只要模型的贡献率不超过遗传率值, 所有检测到的 QTL 都应是合理的, 非假阳性的。另一方面, 如上所述, RTM-GWAS 与精细试验设计紧密结合, 降低了误差, 提高了检出功效, 这是其他 GWAS 方法没有关注到的。此外, 大豆基因组有 4.5 万—5.0 万个基因, 数量性状属于复杂性状, 其遗传构成是一套基因网络体系, 涉及大量效应不等相互影响的遗传位点。因此, 检测的 QTL 数量多并无错误, 这正好从侧

面反应了数量性状遗传调控网络体系的复杂性。实际试验中, QTL 总贡献率通常还达不到性状遗传率值, 说明还有一部分 QTL 未被检测出来, 改进试验精确度, 增加标记密度, 还有可能多挖掘出一批小贡献率的 QTL。

### 2.2 关于所用显著水平不够严格的辩解

为了平衡假设测验中的假阳性和假阴性, 统计学上普遍采用 0.01 或 0.05 作为显著水平, 更严格的显著性水平可用于满足特定目的, 这取决于重要性和成本。因此, RTM-GWAS 也使用常规的显著水平 0.01 或 0.05 测验模型的显著性, 控制全试验错误率。如前所述, 由于 RTM-GWAS 基于多位点模型, 模型的显著性本身就说明了模型的合理性和模型中所包括 QTL 的合理性, 并不需要进行额外的多重测验校正。这与单位点模型不同, 单位点模型的显著性仅说明该位点是显著的合理的, 并不说明全部检测到的位点整体的显著性和合理性, 因而要加做全试验所检测到位点整体的显著性测验, 例如 Bonferroni 校正。简言之, 多位点模型下, 模型的显著性已包含了全部入选的 QTL, 不需另作矫正, 反之, 单位点模型下, 入选单个位点的全体须经整体的显著性测验, 要另做多重测验或 Bonferroni 校正。另外, RTM-GWAS 同时还给出每个检出 QTL 的统计测验  $P$  值, 研究者还可以选择特定阈值进一步筛选 QTL, 而无需重新计算。这是因为 RTM-GWAS 检测的 QTL 以  $P$  值从小到大依次排列,  $P$  值较小的往往贡献率越大。如果实际研究有更严格的要求, 例如克隆候选基因费时费事, 可以使用严格的阈值来筛选相对更重要的位点。

## 3 RTM-GWAS 的应用前景

以往研究者常将 GWAS 方法用于寻找个别基因, 因而并不注重性状遗传体系中全部 QTL 的检出。如上所述, 建立 RTM-GWAS 方法着眼于解析自然群体或遗传群体的遗传体系或 QTL-等位变异 (QTL-allele) 矩阵。该方法自建立以来已做了多方面的应用尝试, 现归纳于后。更多的应用还有待今后在使用中发展。

### 3.1 遗传体系解析与重要基因克隆

对多个大豆数量性状的遗传解析显示, 与以往 GWAS 方法相比, RTM-GWAS 能检测到较多的 QTL 和相应的等位变异, QTL 总贡献率也更接近性状的遗传率值, 为研究数量性状完整的遗传体系提供了途径。RTM-GWAS 除应用在单环下解析群体 QTL-等位变异体系外, 还可对多环境试验表型数据做基因与环境互作效应的解析。检测到的 QTL 可以有主效应有互作效

应,有主效应无互作效应,无主效应有互作效应等不同类型;分析的结果可以归纳成主效 QTL-等位变异和互作效应 QTL-等位变异 2 个矩阵,为等位变异环境效应的研究提供了通路。另外,RTM-GWAS 分析结果中 QTL 按概率值(显著程度)由小到或贡献率由大到小依次检出,因此,除可以考察整个遗传体系外,还可以按各个 QTL 的重要程度分别进行研究,包括检出重要位点做基因克隆研究等。

### 3.2 双(多)亲杂交衍生群体遗传解析

RTM-GWAS 特别适合于双(多)亲杂交后代群体,例如重组自交系群体和巢式关联作图群体。与自然群体不同,由于双亲群体遗传构成规则,群体偏差干扰小,检测功效更高。对大豆重组自交系群体的分析显示,RTM-GWAS 除了能检测到传统复合区间作图法检测的 QTL,还能检测更多的 QTL,解释更多的表型变异。另外,对 4 个重组自交系群体组成的大豆巢式关联作图群体的分析显示,RTM-GWAS 可以检测到复等位变异数目不等的 QTL,每个 QTL 包含 2—5 个等位变异。以往巢式关联作图群体分析方法尽管将多个重组自交系群体联合分析,但仍将重组自交系群体相互独立处理<sup>[18]</sup>,因此,对多个亲本的复等位变异估计不确切。而 RTM-GWAS 视多个重组自交系群体为一个整群体,利用 SNPLDB 标记可以估计 QTL 上多个亲本间不同的等位变异,更符合群体遗传特征。

### 3.3 群体遗传分化与进化

RTM-GWAS 可以估计出所有 QTL 上每个等位变异的效应,据此可建立数量性状的 QTL-等位变异(QTL-allele)矩阵,即群体内每个材料在每个 QTL 上的效应矩阵。QTL-allele 矩阵包括了性状在群体中的全部遗传组成,不仅能用于候选基因发掘,还特别适合于群体遗传分化和进化分析。例如,在对中国东北地区不同熟期组大豆主茎节数的遗传研究中(Crop Science 即将发表),大豆主茎节数从晚熟组(MGI, MGII)的 17.89 个减少到早熟组(MG0, MG00, MG000)的 13.11 个。在东北大豆种质群体中,RTM-GWAS 共检测到 76 个主茎节数 QTL,包括 183 个等位变异,共解释了 65.63%的表型变异。在晚熟组到早熟组的进化过程中,有 28.42%的等位变异产生变化,其中新生等位变异占 6.56%,淘汰等位变异占 21.86%,而 71.58%的等位变异直接从晚熟组传递到早熟组。说明东北大豆主茎节数进化过程中,遗传是首要动力,其次是淘汰或选择(淘汰正效等位变异),第三是新生或突变(新生负效等位变异),最后通过所剩 QTL 等位变异

间的遗传重组使群体产生遗传分化和进化。

### 3.4 QTL-allele 矩阵应用于设计育种

亲本组配和后代选择是常规育种的 2 个主要步骤,对于复杂性状的遗传改良,背景选择和前景选择同等重要,仅通过少数几个主效位点的重组选择很可能无法创造出突破性新品种。由 RTM-GWAS 建立的 QTL-allele 矩阵为亲本组配和后代选择提供了遗传依据。基于 QTL-allele 矩阵可以对所有亲本组合的后代纯合群体进行预测,从而筛选最优亲本组合。因此,基于 QTL-allele 矩阵的选择是对目标性状位点进行的直接选择,更符合实际育种需求。同时,根据 QTL-allele 矩阵还可以设计最佳基因型(各位点最佳等位变异的组合),根据相应标记对后代做标记辅助选择。这与 MEUWISSEN 的全基因组选择方法(genome-wide selection)<sup>[19]</sup>有本质不同。全基因组选择首先基于参考子群体建立分子标记与多个目标性状表型的综合线性关系,这种关系对每个性状来说是黑箱关系。然后利用个体的全基因组分子标记信息预测候选个体的综合育种值,凭综合育种值对个体做选择。鉴于作物育种通常涉及许多组合,一个组合涉及上千后代个体,采用全基因组选择时大量全基因组分子标记数据花费高昂,而利用性状的 QTL-allele 矩阵只涉及目标性状的标记,即便有多个性状,涉及的标记总数也只是全基因组标记的极小部分,因而 QTL-allele 矩阵除可用以进行优化组合设计外,对组合后代个体也可能是一种有效的标记辅助设计和选择的途径,不过还有待于实践的检验。

### References

- [1] HE J, MENG S, ZHAO T, XING G, YANG S, LI Y, GUAN R, LU J, WANG Y, XIA Q, YANG B, GAI J. An innovative procedure of genome-wide association analysis fits studies on germplasm population and plant breeding. *Theoretical and Applied Genetics*, 2017, 130(11): 2327-2343.
- [2] ZHANG Y, HE J, WANG Y, XING G, ZHAO J, LI Y, YANG S, PALMER R G, ZHAO T, GAI J. Establishment of a 100-seed weight quantitative trait locus-allele matrix of the germplasm population for optimal recombination design in soybean breeding programmes. *Journal of Experimental Botany*, 2015, 66(20): 6311-6325.
- [3] MENG S, HE J, ZHAO T, XING G, LI Y, YANG S, LU J, WANG Y, GAI J. Detecting the QTL-allele system of seed isoflavone content in Chinese soybean landrace population for optimal cross design and gene system exploration. *Theoretical and Applied Genetics*, 2016, 129(8): 1557-1576.

- [4] LI S, CAO Y, HE J, ZHAO T, GAI J. Detecting the QTL-allele system conferring flowering date in a nested association mapping population of soybean using a novel procedure. *Theoretical and Applied Genetics*, 2017, 130(11): 2297-2314.
- [5] KHAN M A, TONG F, WANG W, HE J, ZHAO T, GAI J. Analysis of QTL-allele system conferring drought tolerance at seedling stage in a nested association mapping population of soybean [*Glycine max* (L.) Merr.] using a novel GWAS procedure. *Planta*, 2018, 248(4): 947-962.
- [6] PAN L, HE J, ZHAO T, XING G, WANG Y, YU D, CHEN S, GAI J. Efficient QTL detection of flowering date in a soybean RIL population using the novel restricted two-stage multi-locus GWAS procedure. *Theoretical and Applied Genetics*, 2018, 131(12): 2581-2599.
- [7] ZHANG Y, HE J, WANG H, MENG S, XING G, LI Y, YANG S, ZHAO J, ZHAO T, GAI J. Detecting the QTL-allele system of seed oil traits using multi-locus genome-wide association analysis for population characterization and optimal cross prediction in soybean. *Frontiers in Plant Science*, 2018, 9: 1793.
- [8] ZHANG Y H, HE J B, MENG S, LIU M F, XING G N, LI Y, YANG S P, YANG J Y, ZHAO T J, GAI J Y. Identifying QTL-allele system of seed protein content in Chinese soybean landraces for population differentiation studies and optimal cross predictions. *Euphytica*, 2018, 214(9): 157.
- [9] KHAN M A, TONG F, WANG W, HE J, ZHAO T, GAI J, WILLENBORG C. Using the RTM-GWAS procedure to detect the drought tolerance QTL-allele system at the seedling stage under sand culture in a half-sib population of soybean [*Glycine max* (L.) Merr.]. *Canadian Journal of Plant Science*, 2019, 99(6): 801-814.
- [10] 贺建波, 刘方东, 王吴彬, 邢光南, 管荣展, 盖钧镒. 限制性两阶段多位点全基因组关联分析方法的特点与计算程序. 中国农业科学, 2020, 53(9): 1699-1703.
- HE J B, LIU F D, WANG W B, XING G N, GUAN R Z, GAI Y. Restricted two-stage multi-locus genome-wide association analysis and its applications to genetic and breeding studies. *Scientia Agricultura Sinica*, 2020, 53(9): 1699-1703. (in Chinese)
- [11] 郝晓帅, 傅蒙蒙, 刘再东, 贺建波, 王燕平, 任海祥, 王德亮, 杨兴勇, 程延喜, 杜维广, 盖钧镒. 东北大豆种质群体百粒重 QTL-等位变异的全基因组解析. 中国农业科学, 2020, 53(9): 1704-1729.
- HAO X S, FU M M, LIU Z D, HE J B, WANG Y P, REN H X, WANG D L, YANG X Y, CHENG Y X, DU W G, GAI J Y. Genome-wide QTL-allele dissection of 100-seed weight in the Northeast China soybean germplasm population *Scientia Agricultura Sinica*, 2020, 53(9): 1704-1729. (in Chinese)
- [12] 潘丽媛, 贺建波, 赵晋铭, 王吴彬, 邢光南, 喻德跃, 张小燕, 李春燕, 陈受宜, 盖钧镒. RTM-GWAS 方法应用于大豆 RIL 群体百粒重 QTL 检测的功效. 中国农业科学, 2020, 53(9): 1730-1742.
- PAN L Y, HE J B, ZHAO J M, WANG W B, XING G N, YU D Y, ZHANG X Y, LI C Y, CHEN S Y, GAI J Y. Detection power of RTM-GWAS applied to 100-seed weight QTL identification in a recombinant inbred lines population of soybean *Scientia Agricultura Sinica*, 2020, 53(9): 1730-1742. (in Chinese)
- [13] 李曙光, 曹永策, 贺建波, 王吴彬, 邢光南, 杨加银, 赵团结, 盖钧镒. 大豆巢式关联作图群体蛋白质含量的遗传解析. 中国农业科学, 2020, 53(9): 1743-1755.
- LI S G, CAO Y C, HE J B, WANG W B, XING G N, YANG J Y, ZHAO T J, GAI J Y. Genetic dissection of protein content in a nested association mapping population of soybean. *Scientia Agricultura Sinica*, 2020, 53(9): 1743-1755. (in Chinese)
- [14] 刘再东, 孟珊, 贺建波, 邢光南, 王吴彬, 赵团结, 盖钧镒. 大豆重组自交系群体异黄酮含量 QTL 连锁定位与关联定位的比较研究. 中国农业科学, 2020, 53(9): 1756-1772.
- LIU Z D, MENG S, HE J B, XING G N, WANG W B, ZHAO T J, GAI J Y. A comparative study on linkage and association QTL mapping for seed isoflavone contents in a recombinant inbred line population of soybean. *Scientia Agricultura Sinica*, 2020, 53(9): 1756-1772. (in Chinese)
- [15] GAI J Y, CHEN L, ZHANG Y, ZHAO T J, XING G N, XING H. Genome-wide genetic dissection of germplasm resources and implications for breeding by design in soybean. *Breeding Science*, 2012, 61(5): 495-510.
- [16] SUL J H, MARTIN L S, ESKIN E. Population structure in genetic studies: Confounding factors and mixed models. *PLoS Genetics*, 2018, 14(12): e1007309.
- [17] 贺建波, 刘方东, 邢光南, 王吴彬, 赵团结, 管荣展, 盖钧镒. 限制性两阶段多位点全基因组关联分析方法的特点与计算程序. 作物学报, 2018, 44(9): 1274-1289.
- HE J B, LIU F D, XING G N, WANG W B, ZHAO T J, GUAN R Z, GAI J Y. Characterization and analytical programs of the restricted two-stage multi-locus genome-wide association analysis. *Acta Agronomica Sinica*, 2018, 44(9): 1274-1289. (in Chinese)
- [18] BUCKLER E S, HOLLAND J B, BRADBURY P J, *et al.* The genetic architecture of maize flowering time. *Science*, 2009, 325(5941): 714-718.
- [19] MEUWISSEN T H, HAYES B J, GODDARD M E. Prediction of total genetic value using genome-wide dense marker maps. *Genetics*, 2001, 157(4): 1819-1829.

(责任编辑 李莉)