

基于高密度 SSR 连锁群的烟草致香物质关联分析

任民, 张长静, 蒋彩虹, 程立锐, 贾兴华, 杨爱国

中国农业科学院烟草研究所, 烟草行业烟草基因资源利用重点实验室, 青岛 266101

摘要: 烟叶致香物质是烤烟品质及香型风格的物质基础, 是烟草品质改良的重点之一。采用气相色谱方法, 对 11 份不同香型风格的烟草种质进行了 52 种致香成分的检测, 同时采用 1 条高密度连锁群上的 62 个 SSR 位点对其进行了分子标记位点分型。通过标记 - 性状关联分析, 发掘到 17 种烟草致香物质的 24 个关联位点。大多数位点关联到了多种致香成分。对供试材料的连锁不平衡分析, 发现该连锁群存在 2 个单倍型, 单倍型 tsHap1_1 由位点 52429 和 53840 组成, 未发现其与致香物质相关联; 单倍型 tsHap1_2 由 3 个相邻的位点 61443、52305 和 55052 组成, 总长度 3.012cM, 研究发现该单倍型与 9 种致香物质关联。本研究发掘的关联位点和单倍型, 可用于烟草高香气育种分子标记辅助选择。

关键词: 烟草; 致香物质; 分子标记; 关联分析; 单倍型

doi: 10.3969/j.issn.1004-5708.2014.04.017

中图分类号: Q81 **文献标志码:** A **文章编号:** 1004-5708 (2014) 04-0088-06

Association analysis of tobacco aroma constituents based on high density SSR linkage group

REN Min, ZHANG Changjing, JIANG Caihong, CHENG Lirui, JIA Xinghua, YANG Aiguo

Key Laboratory for Tobacco Gene Resources, Tobacco Research Institute, Chinese Academy of Agricultural Sciences, Qingdao 266101, China

Abstract: Eleven tobacco varieties with different aroma types were selected as samples to examine 52 aroma constituents by Gas Chromatography. 62 simple-sequence repeat (SSR) markers which come from a high density tobacco SSR linkage group were scanned. Marker-trait association analysis showed that there were 24 loci significantly associated with 17 types of aroma constituents and each locus was often associated with several aroma constituents. LD analysis revealed two haplotypes at this linkage group, which were tsHap1_1 and tsHap1_2. The former consisted of locus 52429 and 53840 while the latter comprised locus 61443, 52305 and 55052, occupying 3.012cM. 9 aroma constituents were found to be associated with tsHap1_2. Therefore, the associated SSR loci and haplotype with aroma constituents are both beneficial for marker assistant selection in tobacco leaf quality improvement.

Keywords: tobacco; aroma constituents; molecular markers; association analysis; haplotype

烟草作为典型的嗜好性作物, 烟叶致香物质对其品质及香型风格有着非常重要的影响。烟叶在生长发育及其调制过程中会积累起大量的致香物质, 并在不同类型或品种间存在着显著的差异。致香物质的含量

及组成除受到生态条件、调制措施等外部环境因素影响外, 品种间的基因型差异是其遗传基础, 也是影响香吃味的第一因子^[1]。目前已经发现了一些影响烟叶致香物质的遗传性状^[2], 初步探明了部分致香物质的遗传规律^[3-5]。随着植物次生代谢网络的日渐清晰, 烟草多个致香物质代谢途径关键酶的基因也已被同源克隆, 并明确了表达模式^[6-9]。但是相对烟草致香物质的种类和数量, 目前已经发掘的相关基因仍远无法满足实际需要, 并极大的限制了分子育种技术在烤烟品质育种中的应用。此外致香物质成分复杂, 数量众多, 按照传统的方法进行基因发掘, 所需分离群体的群数将

基金项目: 农业部公益性行业科研专项 (201203091); 烟草重要性状基因发掘及功能标记开发 (110201301008)

作者简介: 任民, 博士, 副研究员, 从事烟草分子育种研究, Tel:0532-88702169, Email:renmin@caas.cn

通讯作者: 杨爱国, 博士, 副研究员, 从事烟草分子育种研究, Tel:0532-88702138, Email:yangaiquo@caas.cn

收稿日期: 2013-07-09

极为庞大,这也成为制约致香物质基因发掘研究的一个重要因素。

近年来,随着基因组学技术的发展,关联分析技术作为一种新的基因发掘理论与技术体系,在植物基因克隆领域得到了快速的发展。基于关联分析利用一套自然群体即可开展规模化的生物性状调控基因发掘,可有效的解决来自群体、亲本、性状等多方面的限制因素^[10]。同时,烟草分子标记的数量目前已较为丰富。2011年发布了烟草高密度 SSR 遗传连锁图谱^[11],随该图谱公开了 5000 多对 SSR 标记,其中 2700 多对分别位于 24 条连锁群,该图谱的发布为烟草基因发掘与克隆提供了丰富的分子标记。本研究利用该高密度遗传连锁图谱,探索基于关联分析发掘烟叶致香物质调控基因和遗传位点的技术途径与方法,

以期构建相关技术体系,为大规模发掘烟草致香物质和其他重要性状基因奠定理论与方法依据。

1 材料与方法

1.1 供试材料

本研究选用的供试品种见表 1,供试材料种植于中国农业科学院烟草研究所即墨试验基地,田间栽培及各项管理措施均一致,于烟叶成熟期采收单株叶片,在密集烤房内烘烤调制。

1.2 致香成分气相色谱(GC)检测

烤后烟叶致香物质检测采用气相色谱法^[12],共检测了 52 种致香物质(表 2),包括 18 种酸性成分,22 种中性成分和 12 种碱性成分。

表 1 本研究供试材料列表

Tab. 1 Experiment samples

代号	全国统一编号	品种名称	代号	全国统一编号	品种名称
PZ1	540	红花大金元	PZ7	2266	K326
PZ2	128	青梗	PZ8	2290	NC82
PZ3	493	大白筋 599	PZ9	2490	中烟 90
PZ4	1084	Speight G-28	PZ10	2503	CV87
PZ5	1307	翠碧一号	PZ11	3656	中烟 98
PZ6	1370	86-3002			

表 2 本研究所检测的 52 种致香物质及其属性

Tab. 2 52 aroma constituents detected in the experiment

致香物质	属性	致香物质	属性	致香物质	属性
2-甲基丙酸	酸性	2-甲基四氢呋喃-3-酮	中性	2-甲氧基吡嗪	碱性
丁酸	酸性	2-甲基-2-庚烯-6-酮	中性	2,5-二甲基吡嗪	碱性
异戊酸	酸性	2-糠醛	中性	2,3-二甲基吡嗪	碱性
戊酸	酸性	3-甲基-2-环戊烯-1-酮	中性	2-乙基-3-甲基吡嗪	碱性
2-甲基戊酸	酸性	苯甲醛	中性	2-乙酰吡啶	碱性
3-甲基戊酸	酸性	芳樟醇	中性	2-乙酰吡咯	碱性
4-甲基戊酸	酸性	异佛尔酮	中性	2-乙酰基-1-甲基吡咯	碱性
己酸	酸性	2-乙酰基-5-甲基呋喃	中性	2,3-二乙基吡嗪	碱性
庚酸	酸性	呋喃甲醇	中性	四甲基吡嗪	碱性
反-2-己烯酸	酸性	香芹酮	中性	3-乙酰吡啶	碱性
辛酸	酸性	香茅醇	中性	喹啉	碱性
壬酸	酸性	二氢大马酮	中性	吡啶	碱性
苯甲酸	酸性	2-乙酸苯乙酯	中性		
月桂酸(十二酸)	酸性	苯甲醇	中性		
肉豆蔻酸(十四酸)	酸性	β -紫罗兰酮	中性		
棕榈酸(十六酸)	酸性	金合欢醇	中性		
硬脂酸(十八酸)	酸性	新植二烯	中性		
油酸(十八烯酸)	酸性	巨豆三烯酮 1、2、3、4	中性		
		二氢猕猴桃内酯	中性		

1.3 SSR 引物

本研究 SSR 位点及其引物序列来自烟草高密度 SSR 遗传连锁图谱的 1 号连锁群^[11], 共计 62 个遗传位点, 总长度 130.043 cM, 平均距离 2.097 cM, 最大间距 8.626 cM, 最小间距 0.269 cM。

1.4 DNA 提取

采用 TIANGENE™ Plant Genomic DNA Kit 提取供试材料全基因组 DNA。

1.5 PCR 反应体系

PCR 总反应体系为 10 μ L, 其中 30~50 ng/ μ L DNA 模板 1 μ L, 2 \times DreamTaq™ Green PCR Master Mix 5 μ L, 2 μ mol/L 正反向引物混合工作液 1 μ L, 加 ddH₂O 至 10 μ L。所用试剂购自 MBI。PCR 反应在 96 孔 ABI Veriti 梯度基因扩增仪上按以下程序运行: 首先 94 $^{\circ}$ C 预变性 5 min; 94 $^{\circ}$ C 15s, 55 $^{\circ}$ C 15 s, 72 $^{\circ}$ C 30 s 共 32 个循环; 最后 72 $^{\circ}$ C 延伸 7 min, 4 $^{\circ}$ C 保存。

1.6 电泳检测

完成 PCR 扩增 48 h 内, 取 2 μ L PCR 产物在 8% 非变性聚丙烯酰胺凝胶上, 恒电压 850 V 电泳 90 min。用稍加改进的 NaOH 银染方法^[13]对电泳后的

聚丙烯酰胺凝胶进行染色显影。

1.7 统计分析

供试群体遗传多样性指数的计算采用 POPGENE 软件^[14]。材料群体遗传结构分析采用 STRUCTURE 2.3.2 软件^[15], 其 Length of burn-in period 和 Number of MCMC Peps after Burn-in 均设为 100,000 次。分组方式即 K 值从 1 到 10, 迭代 10 次。通过计算 ΔK 值^[16], 确定最终的分组方式。标记 - 性状关联分析采用 TASSEL 软件 (v2.1)^[17] 基于“MLM”模型进行运算, 并代入群体结构, 遗传位点间连锁不平衡性分析采用“Link. Diseq.”模块。

2 结果与分析

2.1 遗传多样性

本研究在 63 个 SSR 位点上, 共得到 37 个多态性位点, 多态性比例 58.7%。总共扩增出 82 个多样性条带, 平均每个位点 2.216 个等位。编号为 61342 的位点多态性最高, 达 0.6400, 编号为 52969 的位点多样性最低, 仅为 0.095, 37 个多样性位点的平均多样性指数为 0.3314。

表 3 多态性位点的遗传多样性指数
Tab. 3 Genetic diversity index of polymorphic loci

位点	位置	多样性指数	位点	位置	多样性指数	位点	位置	多样性指数
50172	2.756	0.1975	60553	71.306	0.1653	61049	102.897	0.5000
61521	4.415	0.1800	55109	79.932	0.5744	60761	103.985	0.5124
54339	6.654	0.4628	50778	81.286	0.5124	51438	107.015	0.1653
52429	12.451	0.3200	50460	85.221	0.4800	51665	108.103	0.3750
53840	13.003	0.3200	61025	87.176	0.1653	53110	114.328	0.2550
61343	22.254	0.1653	54092	91.567	0.4959	50909	119.699	0.3967
60830	27.795	0.1653	54794	92.404	0.3200	52969	119.968	0.0950
54277	28.627	0.1800	61342	95.978	0.6400	53304	120.783	0.1653
52565	36.705	0.1653	50793	96.536	0.5312	61443	126.207	0.2975
53990	48.165	0.4959	53327	97.084	0.3200	52305	127.843	0.2975
50988	49.839	0.1653	50467	99.059	0.1653	55052	129.219	0.2975
61382	55.837	0.3967	52573	100.165	0.4938			
52548	56.389	0.3704	50457	100.716	0.1800	平均		0.3314

2.2 烤后烟叶致香成分检测结果

本研究利用气相色谱技术共对 52 种致香成分进

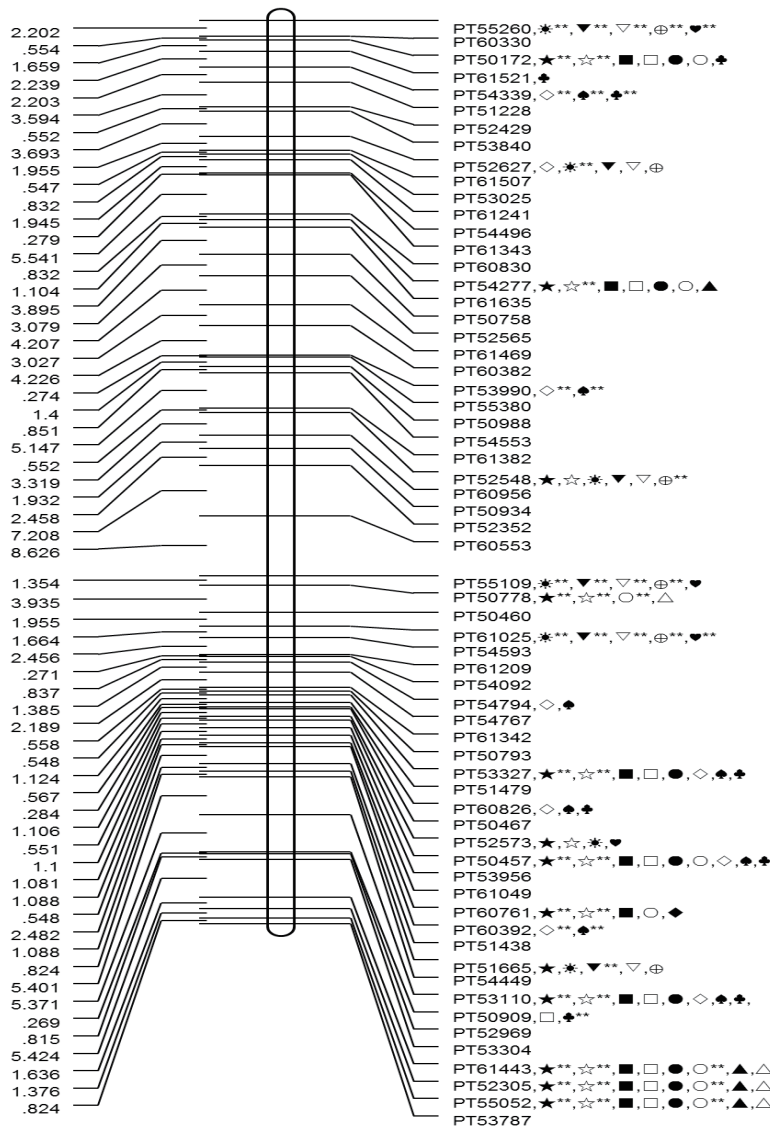
行了检测。在供试材料中 86-3002 (PZ6) 的致香物质总含量最高, 达到 672.99 μ g/g; 由于青梗 (PZ2)

的新植二烯含量在所有供试品种中最低，导致其致香物质总含量仅为 282.46 $\mu\text{g/g}$ ，为所有供试材料的最低值。在各种致香成分中新植二烯的平均含量最高，达 284.60 $\mu\text{g/g}$ ，2-甲基丙酸的平均含量最低，平均仅为：0.0189 $\mu\text{g/g}$ 。平均含量最高的前十种成分依次为：新植二烯、棕榈酸（十六酸）、吲哚、2-乙酰吡咯、2-乙酸苯乙酯、巨豆三烯酮 3、金合欢醇、苯甲醇、2-乙酰吡啶、巨豆三烯酮 2。

2.3 标记 - 性状关联分析

本研究利用 62 个 SSR 标记对 52 种致香物质进行了致香物质标记 - 性状关联分析。首先进行了群体遗传结构分析，当 $K=4$ 时， ΔK 值最高。然后代入上述遗传结构，进行了标记 - 性状关联分析。在供试

连锁群上，共发掘到 17 种致香成分的 24 个关联位点（图 1），相关致香成分分别为：丁酸（★）、3-甲基戊酸（☆）、2,3-二甲基吡嗪（■）、月桂酸（□）、新植二烯（●）、2-甲基丙酸（○）、2-甲基-2-庚烯-6-酮（▲）、苯甲醛（△）、2-甲基四氢呋喃-3-酮（◆）、异戊酸（◇）、巨豆三烯酮 2（♠）、巨豆三烯酮 4（♣）、庚酸（☼）、2-乙酰基-1-甲基吡咯（▼）、喹啉（▽）、2,5-二甲基吡嗪（⊕）、2-乙酰吡咯（♥）。各关联 SSR 位点的关联成分和数量各不相同，但多数与多种致香成分相关，只有 6 个位点关联到的致香成分在 3 种以内。其中 3 个相邻的位点 61443、52305 和 55052 关联到的致香成分最多，达到 9 种，且成分相同。



注：符号后标有“**”表示不同物质之间的相关性有高度统计学意义 ($P < 0.01$)。

图 1 致香成分遗传连锁图

Fig. 1 Genetic linkage map of aroma constituents

2.4 SSR 位点单倍型 (Haplotype) 分析

通过对供试群体 SSR 分型结果进行连锁不平衡性 (LD) 分析, 在供试连锁群上发掘到 2 个 SSR 单倍型, 涉及 5 个 SSR 位点。单倍型 tsHap1_1 由位点 52429 和 53840 组成, 在本研究中未发现关联性状。单倍型 tsHap1_2 由 3 个相邻的位点 61443、52305 和 55052 组成, 总长度 3.012 cM, 在本研究中发现与 9 种致香物质关联 (图 1)。该单倍型在每个组成位点上均扩增出 2 种等位变异, 按带型大小分别记为 Allel-1 和 Allel-2。单倍型 tsHap1_2 在供试群体中共检测到 2 种组合方式, 分别是 tsHap1_2^A (61443^{Allel-2}

+ 52305^{Allel-1} + 55052^{Allel-2}) 和 tsHap1_2^B (61443^{Allel-1} + 52305^{Allel-2} + 55052^{Allel-1})。在品种 PZ1 和 PZ2 中为 tsHap1_2^A, 在其他供试品种中为 tsHap1_2^B。由图 2 可知, tsHap1_2^A 对致香物质 2-甲基-2-庚烯-6-酮、2-甲基丙酸、丁酸、月桂酸和 2,3-二甲基吡嗪含量的遗传贡献为正向, 即利用 tsHap1_2 单倍型对群体分类后, 含有 tsHap1_2^A 单倍型的群体上述成分的含量高于 tsHap1_2^B 群体; 对 2-甲基四氢呋喃-3-酮、苯甲醛、3-甲基戊酸和新植二烯含量的遗传贡献为负向。该单倍型组成位点 52305 在群体扩增中带型较为稳定、清晰, 可以作为 tag SSR 标记使用。

表 4 本研究发掘到的 2 个烟草 SSR 单倍型

Tab. 4 tsHap1_1 and tsHap1_2

单倍型	位点 1	位置 1	位点 2	位置 2	距离 /cM	R ²	D 值	p 值
tsHap1_1	53840	13.003	52429	12.451	0.552	1	1	0.022
tsHap1_2	52305	127.843	61443	126.207	1.636	1	1	0.018
	55052	129.219	61443	126.207	3.012	1	1	0.018
	55052	129.219	52305	127.843	1.376	1	1	0.018

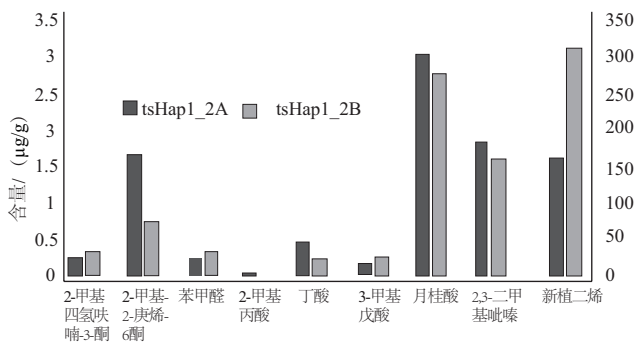


图 2 tsHap1_2 单倍型对关联致香成分的遗传效应

Fig. 2 Genetic effects of tsHap1_2 on associated aroma constituents

3 结论与讨论

本研究通过一组自然群体, 利用标记-性状关联分析的技术途径开展了烟草致香物质关联位点发掘研究, 共找到 17 种致香物质的 24 处关联位点。该研究初步探索了一套复杂性状调控基因或遗传位点发掘的新途径。即: 1) 根据研究目标构建自然群体, 该群体应该尽可能多的覆盖目标性状变异范围, 同时兼顾到其他性状的研究; 2) 开展目标性状的精准鉴定; 3) 对供试群体进行标记位点基因分型, 标记密度越高, 分析效果越好; 4) 在进行关联分析之前, 应开展群

体遗传结构分析, 以消除遗传结构对表型变异的影响^[18]; 5) 结合表型鉴定开展标记-性状关联分析, 关联位点的选择除了可以利用显著性水平判定外, 还可利用关联位点的重复性以及其它数据的规律性协助选择。

致香物质是一类复杂的表型变异, 涉及成分众多, 且以多基因调控为主。本研究发现利用关联分析技术进行类似复杂性状的基因发掘, 不但技术上可行, 而且具有一定的优势。首先, 一套研究群体即可满足不同性状的基因发掘研究, 能够显著降低研究成本。如果利用人工群体, 每个群体仅能针对有限的性状和成分, 面对致香物质这类复杂性状时, 必须构建多套人工群体, 因此研究成本的增加将会非常明显; 其次, 自然群体是天然的稳定群体, 省去通过杂交、自交构建群体的步骤, 因此可以有效缩短研究周期^[19]; 最后, 自然群体相当于一个丰富的自然变异基因池, 每个基因座基本可以囊括全部等位变异, 从而可以突破连锁分析每个基因座只有 2 种等位变异的限制, 进而可以更加精细的研究每种等位变异的遗传贡献, 结合育种目标找到最优等位变异。

当前, 致香物质的检测虽然实现了较高的自动化, 但成本依然偏高, 成为了限制这类研究大规模开展的

一个重要因素。因此,在今后的研究中应当进一步优化检测技术,以大幅降低检测成本。还应不断探索新型检测方法,如利用电子鼻等新型仪器设备,开展田间的快速检测。在扩大群体的基础上,还有必要不断提高标记密度和基因组覆盖度,随着烟草基因组测序工作的基本完成^[20],更高密度的分子标记,如 SNP 将日趋成熟^[21]。可以预计随着致香物质检测技术和高通量分子标记的迅速发展,烟草致香物质基因发掘将会迎来新的突破。

参考文献

- [1] 李天福,陈萍. 烤烟不同耐肥品种的肥料利用率与烟叶品质[J]. 烟草科技, 1999(4): 33-34.
- [2] 史宏志,官春云. 烟草腺毛分泌物的化学成分及遗传[J]. 作物研究, 1995, 9(3): 46-49.
- [3] Johnson J C, Nielsen M T, Collins G B. Inheritance of glandular trichomes in tobacco[J]. *Crop Science*, 1988, 28: 241-244.
- [4] Gwynn G R, Severson R F, Jackson D M, et al. Inheritance of sucrose esters containing beta-methylvaleric acid in tobacco[J]. *Tobacco Science*, 1985, 29: 79-81.
- [5] Nielsen M T. Altering flavor and aroma constituents of burley tobacco[J]. *Tobacco Science*, 1991, 35: 69-73.
- [6] Fukasawa-Akada T, Kung S D, Watson J C. Phenylalanine ammonia-lyase gene structure, expression, and evolution in *Nicotiana*[J]. *Plant Molecular Biology*, 1996, 30: 711-722.
- [7] Nagai N, et al. Cloning and sequencing of a full-length cDNA coding for phenylalanine ammonia-lyase from tobacco cell culture[J]. *Plant Physiology*, 1993, 33(6): 1264-1268.
- [8] Facchini P, Chappell J. Gene family for an elicitor-induced sesquiterpene cyclase in tobacco[J]. *Proc Natl Acad Sci*, 1992, 89: 3088-3092.
- [9] 徐宜民,王树声,赖禄祥,等. 烟草生物碱的研究现状[J]. *中国烟草科学*, 2003, 24(2): 12-16.
- [10] 杨小红,严建兵,郑艳萍,等. 植物数量性状关联分析研究进展[J]. *作物学报*, 2007, 33(4): 523-530.
- [11] Bindler G, Plieske J, Bakaer N, et al. A high density genetic map of tobacco (*Nicotiana tabacum* L.) obtained from large scale microsatellite marker development[J]. *Theor Appl Genet*, 2011, 123: 219-230.
- [12] 任民,王日新,贾兴华,等. 普通烟草种内主要栽培类型间烟叶香味成分的比较与分析[J]. *中国烟草科学*, 2008, 29(06): 36-42.
- [13] 任民,贾兴华,蒋彩虹,等. Bassam 和 Sanguinetti 银染方法在 SRAP 和 TRAP 标记中的比较研究[J]. *生物技术通报*, 2008, 1: 113-116
- [14] Yeh F C, Boyle T J B. Population genetic analysis of co-dominant and dominant markers and quantitative traits[J]. *Belgian Journal of Botany*, 1997, 129: 157.
- [15] Pritchard J K, Stephens M, Donnelly P. Inference of population structure using multilocus genotype data[J]. *Genetics*, 2000, 155: 945-959.
- [16] Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals using the software structure: a simulation study[J]. *Molecular Ecology*, 2005, 14(8): 2611-2620.
- [17] Bradbury P J, Zhang Z, Kroon D E, et al. TASSEL: Software for association mapping of complex traits in diverse samples[J]. *Bioinformatics*, 2007, 23: 2633-2635.
- [18] Larsson S J, Lipka A E, Buckler E S. Lessons from Dwarf8 on the strengths and weaknesses of structured association mapping[J]. *PLoS Genetics*, 2013, 9(2): e1003246.
- [19] Yu J M, Buckler E S. Genetic association mapping and genome organization of maize[J]. *Current Opinion in Biotechnology*. 2006, 17:155-160.
- [20] 王元英. 烟草基因组知识篇: 1. 基因组与烟草基因组计划[J]. *中国烟草科学*, 2010, 31(1): 81-82.
- [21] Elshire R J, Glaubitz J C, Sun Q, A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species[J]. *PLoS One*, 2011, 6(5): e19379.