

烟草航天诱变突变体变异检测及分析

周世奇^{1,2}, 刘东阳³, 潘旭浩¹, 屈建康⁴, 程立锐¹, 任 民¹, 晁江涛¹, 张 玉¹, 罗成刚¹

(¹ 中国农业科学院烟草研究所 / 烟草遗传改良与生物技术重点开放实验室, 青岛 266101;

² 中国农业科学院研究生院, 北京 100081; ³ 四川省烟草公司凉山州公司, 西昌 615000; ⁴ 四川省烟草科学研究所, 成都 610041)

摘要: 为预测控制烟草重要性状的关键基因并研究航天诱变机理, 对经航天诱变选育的烟草突变体材料 NC89-M 与野生型 NC89 进行全基因组重测序, 测序深度 30×, 并对各种类型变异进行检测注释。NC89 中检测到单核苷酸多态性位点 (SNP, Single nucleotide polymorphism) 1848013 个, 小片段插入缺失 (Indel, insertion-deletion) 398922 个, 结构变异 (SV, Structure variation) 41969 个; NC89-M 中检测到 SNP 1876219 个, Indel 402011 个, SV 42699 个。NC89-M 和 NC89 对比共得到 271655 个 SNP, 分布在 8378 个基因上, 23450 个 Indel 造成 2156 个基因突变。分析结果表明, NC89-M 中 SNP 变异数目最多, 其转换类型和颠换类型的比值为 2.053, 说明航天诱变对烟草基因组的变异以单碱基突变为, 突变类型以转换为主; 在 Indel 中, 插入突变数目明显多于缺失突变, 证明航天诱变造成的 Indel 中以插入突变为; 在 SV 中, 航天诱变主要造成了插入、缺失、倒位、染色体内部迁移和染色体间的迁移 5 种结构变异类型; 变异基因 KEGG 注释表明, 与代谢通路和次生代谢产物合成两方面基因突变数目最多; 变异基因功能注释表明, 突变体中调控开花时间的 MADS-box 基因, 调控侧生器官发育与叶缘形状的 KNOX1 基因和萜类化合物合成相关基因等发生了变异。

关键词: 烟草; 航天诱变; 重测序; 变异分析

Sequence Modification Analysis of Tobacco Mutant Derived from Space Mutagenesis

ZHOU Shi-qi^{1,2}, LIU Dong-yang³, PAN Xu-hao¹, QU Jian-kang⁴, CHENG Li-rui¹, REN Min¹,
CHAO Jiang-tao¹, ZHANG Yu¹, LUO Cheng-gang¹

(¹ Tobacco Research Institute of CAAS/Key Laboratory of Tobacco Genetic Improvement and Biotechnology, Qingdao 266101;

² Graduate School, Chinese Academy of Agricultural Sciences, Beijing 100081; ³ Liangshan Tobacco Company of Sichuan Province, Xichang 615000; ⁴ Sichuan Province Tobacco Research Institute, Chengdu 610041)

Abstract: To explore candidate genes of tobacco important traits and provide insights of space mutagenesis, we conducted the whole genome re-sequencing (30x coverage) in tobacco mutant NC89-M that was generated by space mutagenesis, and wild-type NC89. In comparison to the reference genome, 1848013 single nucleotide polymorphisms (SNPs), 398922 Indels and 41969 structure variations (SVs) were observed in NC89, while 1876219 SNPs, 402011 Indels and 42699 SVs were detected in NC89-M. A total of 271655 SNPs and 23450 Indels differed between NC89-M and NC89, which resulted in sequence modifications on 8378 and 2156 genes, respectively. SNPs were often observed in mutant, suggesting that single nucleotide mutation is predominant using space mutagenesis. The ratio of conversion type to transversion type is 2.053. It is proved that the single-base mutations are the main mutation type in the tobacco genome induced by space and conversion type is the main mutation type among the single base mutation. The number of insertion was

收稿日期: 2018-06-05 修回日期: 2018-07-17 网络出版日期: 2018-11-13

URL: <http://kns.cnki.net/kcms/detail/11.4996.S.20181112.1031.005.html>

第一作者主要从事烟草遗传育种研究, E-mail: 13854291622@163.com

通信作者: 张玉, 主要从事烟草遗传育种工作, E-mail: zhangyu02@caas.cn

罗成刚, 主要从事烟草遗传育种工作, E-mail: luochenggang@caas.cn

基金项目: 四川省烟草专卖局科技重点项目 (SCYC201701, SCYC201501)

Foundation project: Major Science and Technology Project of Sichuan Province Tobacco Monopoly Bureau (SCYC201701, SCYC201501)

significantly more than that of the deletion. It is proved that the insertion is the main mutation type in Indel caused by space mutagenesis. Among the SV, space mutagenesis mainly results in five types of structural variation including insertion, deletion, inversion, intra-chromosomal translocation and inter-chromosomal translocation. KEGG annotation of mutant genes indicated that the number of mutant genes in the metabolic pathway and biosynthesis of secondary metabolites are the highest. Gene function annotation indicated that several functional genes were modified in mutant, including MADS-box genes that control flowering time, *KNOX1* gene that regulates development of lateral organ and leaf edge shape and also the genes that associate to the synthesis of terpenoids.

Key words: tobacco; space mutation; re-sequencing; variation analysis

我国烟草杂交育种始于 20 世纪 50 年代初^[1], 先后选育出了中烟 90、中烟 100、云烟 87 等优良烤烟品种。但是, 现有种质多年重复利用, 我国又是烟草非起源地, 原始种质受到限制, 致使我国烟草遗传资源匮乏, 育成品种的遗传基础狭窄。因此, 丰富烟草遗传多样性, 创新种质是我国烟草遗传育种的当务之急。

航天育种将农业与航天业相结合, 已被广泛应用于新品种的培育中。自 1987 年以来, 我国已多次利用返回式卫星和神舟飞船搭载植物种子等进行航天诱变育种研究^[2], 获得了一系列重要成果。1987 年, 黑龙江农业科学院利用航天育种培育出优良的辣椒品种太空甜椒 87-2 品种, 与其原种龙椒 2 号相比, 87-2 的 VC 含量和抗性均有明显提高^[3]。自 1992 年开始, 中国农业科学院烟草研究所首次从搭载的 5 个品种中选育出白花突变株, 并开展了相关的遗传研究, 积累了丰富的研究经验; 2006 年, 利用“实践八号”育种卫星装载中烟 100、NC89 等 10 个烤烟品种, 拉开了利用航天诱变开展烤烟种质创新及主栽品种定向改良的序幕。

随着新一代测序技术的发展, 高通量测序技术在作物基因研究上应用广泛。Cheng 等^[4]对水稻 9311 经 γ 射线诱变得到的 Red-1 进行全基因组重测序, 发现 Red-1 中 9.19% 的基因组发生了改变, 存在 381403 个 SNP、50116 个 Indel、1279 个拷贝数变异, 这些变异位于 14493 个基因中。Fu 等^[5]利用全基因组重测序对两个籼稻品种 RGD-7S 和 Taifeng B 进行分析, 检测到 2408845 个 SNP 和 349895 个 Indel, 验证了 Indel 中的杂种弱势基因 Hw3 和 Hw4, 并且获得了具有多态性的 Indel 标记用于克隆杂种弱势基因。Maldonado dos Santos 等^[6]对 28 个巴西大豆品种进行了基因组重测序, 在 20 个大豆染色体中发现了 5835182 个 SNP 和 1329844 个 Indel, 共检测到 327 个基因发生变异,

变异基因的功能与 DNA 转录、光合作用、细胞呼吸以及 ATP 合成 - 耦合电子转运等相关联。在烟草的基因组研究中, 已经完成了 3 个栽培种烟草 (TN90、K326 和 BasmaXanthi (BX)) 的全基因组测序工作, 为在基因水平上开展相关研究提供了支撑^[7]。

NC89-M 是烤烟品种 NC89 种子经“实践 8 号”育种卫星搭载诱变后, 经过 10 年选育, 获得的性状稳定纯合的突变品系。与 NC89 相比, NC89-M 具有株型舒展、上部叶开片好、叶面积增加、烤后烟叶颜色桔黄等优良性状。遗传分析表明, NC89-M 具有较好的配合力, 目前已作为核心亲本在烤烟遗传育种中利用, 并选育出 3 个优良品系, 参加了品系鉴定和品种比较试验。本研究运用基因组重测序技术, 对 NC89 和 NC89-M 全基因组进行分析, 并对检测出的变异位点进行验证, 为挖掘控制烟草关键性状基因及深入研究航天诱变机理奠定了基础。

1 材料与方法

1.1 材料

NC89 是来源于美国的烤烟品种, 由中国烟草种质资源平台提供。2006 年, NC89 经“实践八号”育种卫星搭载返回后, 对其突变性状进行鉴定、选择, 单株套袋自交, 经过 10 年连续选择, 育成突变性状纯合稳定的突变品系 NC89-M, NC89 和 NC89-M 均种植于中国农业科学院西南烟草试验基地, 于中心花开放期各取 10 株调查其农艺性状和植物学性状。

1.2 方法

1.2.1 DNA 的提取 NC89 和 NC89-M 烟苗生长至 5~6 叶期取样, 液氮速冻, 采用 CTAB 法提取 DNA。

1.2.2 NC89 和 NC89-M DNA 多态性检验 利

用已公布的 574 对均匀分布于烟草 24 条染色体的 SSR 标记^[8]对 NC89 和 NC89-M 进行 DNA 的多态性分析。PCR 反应体系:反应体系 10 μ L, 模板 DNA 1 μ L, 引物 2 μ L, 2 \times Taq DNA 聚合酶 5 μ L, ddH₂O 2 μ L, 冰上操作;PCR 反应程序:94 $^{\circ}$ C 预变性 5 min;94 $^{\circ}$ C 变性 15 s, 60 $^{\circ}$ C 退火 15 s, 72 $^{\circ}$ C 延伸 30 s, 35 个循环;72 $^{\circ}$ C 延伸 8 min, 10 $^{\circ}$ C 保存备用;8% 聚丙烯酰胺凝胶电泳检测 PCR 产物;电泳结束,将带有凝胶的玻璃板放入银染液(硝酸银 1.6 g、蒸馏水 1200 mL)中,缓慢摇动 8 min;将玻璃板放入蒸馏水中冲洗 1~2 次;将带有凝胶的玻璃板放入显影液(NaOH 15 g、甲醛 10 mL、蒸馏水 1200 mL),缓慢摇动直到出现明显条带。将带有凝胶的玻璃板放入胶片观察灯上,记录带型。

1.2.3 NC89 和 NC89-M 全基因组重测序 对质量检测合格的 DNA 样本通过使用 Covaris 破碎机随机打断成长度为 350 bp 的片段。采用 TruSeq Library Construction Kit 进行建库,对 DNA 片段进行末端修复、添加 ployA 尾、加测序接头、纯化、PCR 扩增等步骤完成整个文库的制备。文库构建完成后,先使用 Qubit2.0 进行初步定量,文库稀释至 1 ng/ μ L,随后使用 Agilent 2100 对文库的 insert size 进行检测,insert size 符合预期后,使用 Q-PCR 方法对文库的有效浓度进行准确定量(文库有效浓度 >2 nmol/L),保证文库质量。库检合格,进行 illumina HiSeq X Ten 测序。

1.2.4 烟草 NC89 和 NC89-M 基因组变异检测与注释 以烟草品种 K326 基因组作为参考基因组,将原始测序数据经数据过滤后产生的有效测序数据使用 BWA^[9]比对到参考基因组上,比对结果经

SAMTOOLS^[10]去除重复。使用 SAMTOOLS 和 ANNOVAR^[11]对个体 SNP 和 Indel 进行检测和注释,采用 BreakDancer^[12]和 ANNOVAR 对 SV 进行检测与注释。

1.2.5 NC89 和 NC89-M 变异基因分析 对 NC89 与 NC89-M 之间的 SNP、Indel 进行分析,选取变异位点最多的 2000 个基因与 KEGG^[13]数据库 BLAST^[14],并对功能注释的基因进行分类。

1.2.6 SNP 突变位点的验证 根据重测序结果,随机选取 16 个纯合突变的 SNP 进行验证,变异位点上下游 DNA 序列于 <https://www.solgenomics.net> 下载,根据突变 SNP 上下游 DNA 序列使用 Primer5 软件设计引物。PCR 反应体系:模板 DNA 2 μ L,上、下游引物各 2 μ L,高保真酶 25 μ L, ddH₂O 19 μ L。PCR 反应程序:72 $^{\circ}$ C 预变性 5 min;98 $^{\circ}$ C 变性 10 s, Tm 退火 20 s, 72 $^{\circ}$ C 延伸 1 min, 35 个循环;72 $^{\circ}$ C 延伸 5 min。琼脂糖凝胶检测 PCR 产物,挑选条带单一且符合扩增片段长度的 PCR 产物测序。

2 结果与分析

2.1 NC89 和 NC89-M 表型鉴定

为进一步鉴定 NC89 和 NC89-M 的表型差异(图 1),分别对其农艺性状和植物学性状进行了测量和调查。植物学性状调查表明(表 1),与 NC89 相比,NC89-M 叶形宽椭圆,叶色深绿,叶缘为锯齿状,茎叶角度小,主脉略粗,侧枝少,花序集中。农艺性状结果分析表明(表 2),与 NC89 相比,NC89-M 的株高、节距及大田生育期存在极显著差异,叶数、茎围存在显著差异;腰叶长、宽略有增加,开花时间平均延迟 5 d。



A: 花序比对; B: 叶部比对; a: NC89; b: NC89-M

A: Comparison of inflorescences, B: Comparison of leaves, a: NC89, b: NC89-M

图 1 NC89 和 NC89-M 比对图

Fig.1 Phenotypic variation between NC89 and NC89-M

表 1 NC89 和 NC89-M 植物学性状
Table 1 Botanical Traits of NC89 and NC89-M

品种	株型	叶形	叶色	叶缘	茎叶角度	主脉粗细	侧枝	花序	花色
Varieties	Plant type	Leaf shape	Leaf color	Leaf Margin	Angle between stem and leaf	Main vein thickness	Lateral branch	inflorescence	Flower color
NC89	塔型	长椭圆	绿	浅波	大	细	多	分散	粉红
NC89-M	塔型	宽椭圆	深绿	锯齿	小	略粗	少	集中	红

表 2 NC89 和 NC89-M 主要农艺性状
Table 2 Major agronomic traits of NC89 and NC89-M

品种 Varieties	株高 (cm) Plant height	叶数 Leaf number	茎围 (cm) Stem circumfe rence	节距 (cm) Plant pitch	腰叶		上二棚叶		下二棚叶		大田生育期 (d) Field growth period
					Waist leaf		Upper second leaf		Lower second leaf		
					长 (cm) Length	宽 (cm) Width	长 (cm) Length	宽 (cm) Width	长 (cm) Length	宽 (cm) Width	
NC89	148.8	25	7.7	6.1	61.8	25.5	56.1	20.9	62.4	27.8	118
NC89-M	160.3**	22*	9.7*	7.3**	62.4	26.1	55.6	22.3	51.7	31.2	123**

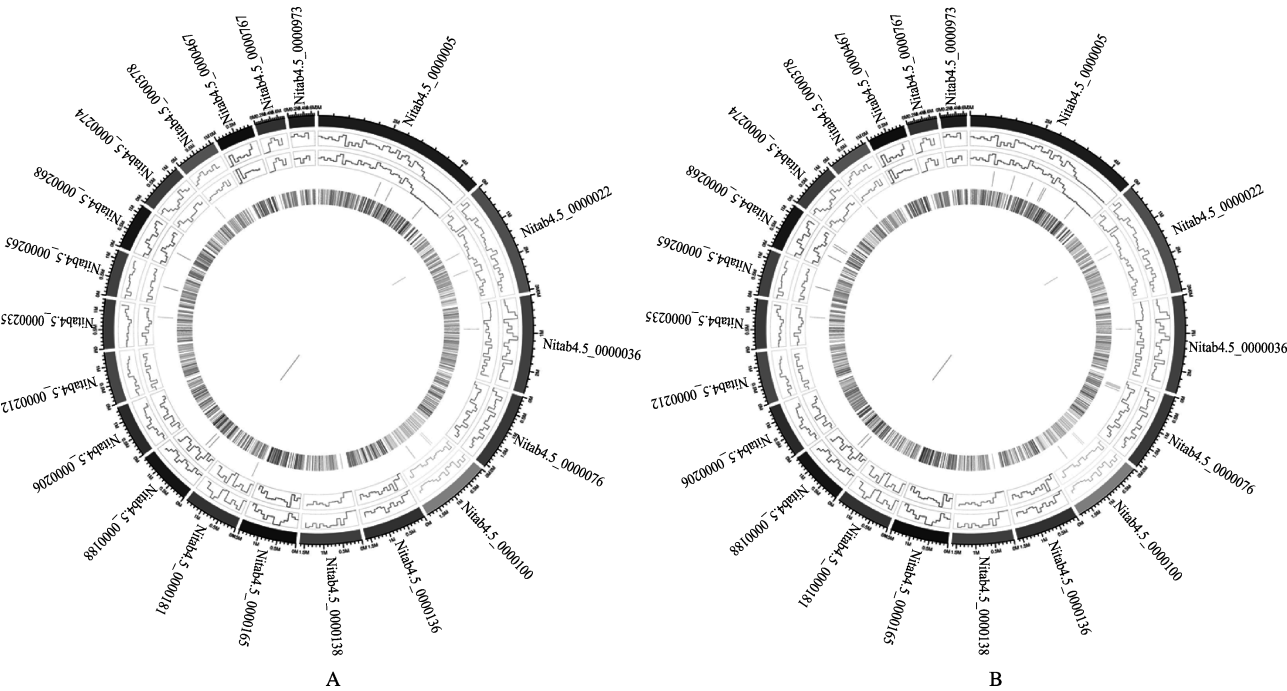
** 差异极显著 (P<0.01); * 差异显著 (P<0.05)
** extremely significant difference at 0.01 probability level, * significant difference at 0.05 level

2.2 NC89 和 NC89-M 的 DNA 多态性分析结果

对 NC89 和 NC89-M 进行 SSR 标记筛选,共选出 12 对多态性标记,根据 Bindler 等^[8]的研究,这 12 对多态性标记依次位于烟草的 1、2、4、4、14、15、15、17、19、20、20 和 22 号染色体上,说明 NC89 和 NC89-M 基因组存在差异。

2.3 NC89 和 NC89-M 全基因组重测序结果

对烟草 NC89 和 NC89-M 使用 illuminaHiSeq X Ten 测序,NC89 共得到 clean reads154792428300 个,其中正确识别率大于 Q30 的碱基占比为 92.43%;NC89-M 共获得 clean reads 148510617000 个,正确识别率大于 Q30 的碱基占比为 90.19%。选择长度最长的前 20 个 Scaffolds 上的各种变异类型分布如图 2 所示。



A: NC89 各类型变异在 scaffolds 的分布; B: NC89-M 各类型变异在 scaffolds 的分布
A: distribution of variations in scaffolds on NC89 chromosomes, B: distribution of variations in scaffolds on NC89-M chromosomes

图 2 NC89 和 NC89-M 变异分析
Fig.2 Variation analysis of NC89 and NC89-M

2.4 NC89 和 NC89-M 基因组 SNP 检测与注释

采用 SAMTOOLS 进行基因组 SNP 的检测,与参考基因组 K326 相比,NC89 共发现 1848013 个 SNP,NC89-M 发现 1876219 个 SNP,对 SNP 检测结果进行注释(表 3)。分析结果得到 6 种类型的 SNP 发生(图 3),NC89 与 NC89-M 对比,发现转换类型 T:A>C:G 数目最多,颠换类型 C:G>G:C 数目最少。为进一步比较 NC89 与 NC89-M 的差异,根据测序结果对 NC89 与 NC89-M 之间的 SNP 再次进行统计,结果发现位于基因上游 1 kb 区的 SNP 共 5072 个,基因下游 1 kb 区 4286 个,基因间区突变 SNP 位点 238963 个,外显子区域突变 SNP 位点 5528 个,内含子区域突变 SNP 位点 17339 个。在所有突变 SNP 中,转换类型(transitions)182676 个,颠换类型(transversions)88979 个,转换与颠换的比率为 2.053。所有突变 SNP 位点共位于 8378 个基因上。

表 3 SNP 统计注释结果
Table 3 SNP annotations

变异类型 Type	NC89	NC89-M	区域 Region
基因上游 1 kb 区 Upstream	15684	16659	
同义变异 Synonymous	6858	8008	CDS
非同义变异 Non-synonymous	8870	9961	CDS
终止密码子获得 Stop gain	196	209	CDS
终止密码子丢失 Stop loss	64	61	CDS
内含子区域突变 Intronic	55799	59633	CDS
剪切位点突变 Splicing	123	138	
基因下游 1 kb 区 Downstream	14234	15095	
基因间区突变 Intergenic	1744602	1764708	
全基因组杂合比率 Het rate	0.295	0.292	
SNP 位点总数 SNP total number	1848013	1876219	

2.5 NC89 和 NC89-M 基因组 Indel 检测与注释

利用 SAMTOOLS 软件检测 Indel,NC89 与参考基因组 K326 对比共检测到 398922 个 Indel;NC89-M 检测到 Indel 数目为 402011 个。使用 ANNOVAR 软件对检测出的 Indel 进行注释(表 4)。统计编码区与全基因组 Indel 的分布(图 4),结果表明,在 Indel 中,NC89 和 NC89-M 均以单碱基插入缺失为主要类型。对 NC89 和 NC89-M 之间发生的 Indel 进行统计,发现 NC89-M 共产生 23450 个 Indel,插入突变 18606 个,缺失突变 4844 个,插入类型明显多于缺失类型。

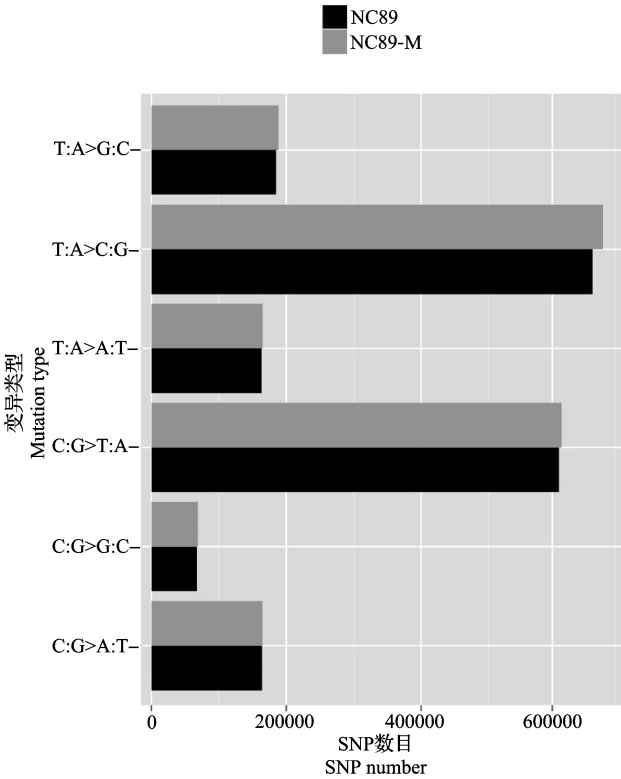
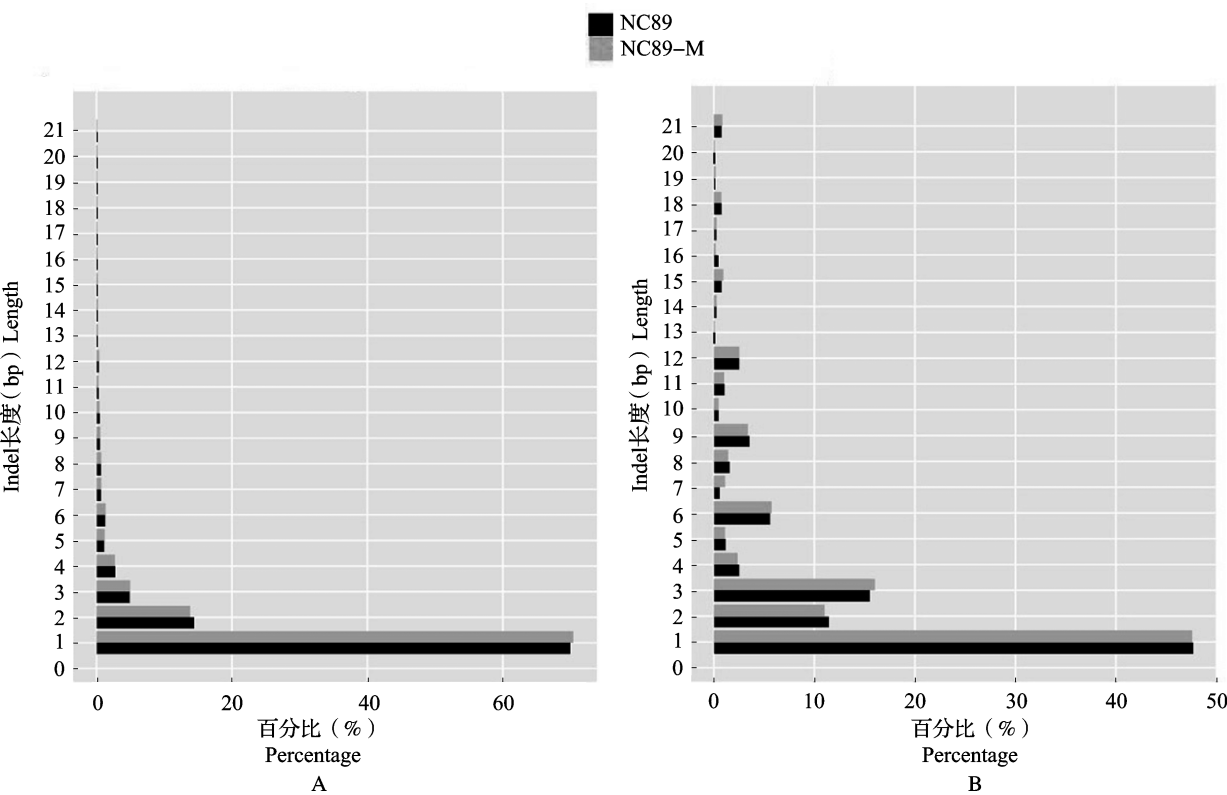


图 3 不同类型 SNP 分布图
Fig.3 Distribution of different types of SNP

表 4 Indel 统计注释结果
Table 4 Indel annotations

变异类型 Type	NC89	NC89-M	区域 region
基因上游 1kb 区 Upstream	9043	9410	
基因下游 1kb 区 Downstream	6903	7212	
终止密码子获得 Stop gain	17	22	CDS
终止密码子丢失 Stop loss	6	5	CDS
蛋白编码框改变的移码突变 (非 3 的整数倍插入) Frameshift insertion	362	382	CDS
蛋白编码框改变的移码突变 (非 3 的整数倍缺失) Frameshift deletion	324	327	CDS
非蛋白编码框改变的突变 (3 的整数倍插入) Non-Frameshift inserion	151	164	CDS
非蛋白编码框改变的突变 (3 的整数倍缺失) Non-Frameshift deletion	153	161	CDS
内含子区域突变 Intronic	22337	23295	CDS
剪接位点突变 Splicing	59	57	
基因间去突变 Intergenic	354820	356223	
Indel 杂合率(%)Hybrid Indel rate	0.020	0.019	
Indel 总数 Indel total number	398922	402011	



A: 全基因组 Indel 长度分布图; B: 编码区 Indel 长度分布图
A: Length distribution of Indels in whole genome, B: Length distribution of Indels in CDS
图 4 全基因组和编码区 Indel 长度分布图
Fig.4 Length distribution of Indels in whole genome and CDS

2.6 NC89 和 NC89-M 基因组 SV 检测与注释

利用 BreakDancer 对 SV 进行检测(表 5), 与参考基因组 K326 对比, NC89 共检测到 41969 个

表 5 SV 统计注释结果
Table 5 SV annotations

变异类型 Type	NC89	NC89-M
基因上游 1 kb 区 Upstream	393	404
基因下游 1 kb 区 Downstream	291	305
外显子区域突变 Exonic	593	613
内含子区域突变 Intronic	1131	1151
剪接位点突变 Splicing	3	3
插入 Insertion	741	778
缺失 Deletion	15668	16685
倒位 Inversion	1085	1161
染色体内部迁移 Intra-chromosomal translocation	23228	23004
染色体间的迁移 Inter-chromosomal translocation	1187	1071
SV 总数 SV total number	41969	42699

SV, NC89-M 共检测到 42699 个 SV。结果表明, 航天诱变共造成了插入、缺失、倒位、染色体内部迁移和染色体间的迁移共 5 种结构变异, 其中缺失与染色体内部迁移的数量高于其他类型变异。

2.7 NC89 和 NC89-M 变异基因分析

变异基因的 KEGG 分类图表明(图 5), NC89-M 与代谢通路相关基因发生突变数目最多, 次生代谢产物生物合成相关基因次之, 该两种基因变异数目明显多于其他类型变异基因。分析 NC89-M 变异基因, 染色体 Nitab4.5_0001021 发生一段 32331 bp 的缺失突变, 致使基因 Nitab4.5_0001021g0050.1 被敲除, 该基因属于 KNOX1^[15] 家族基因, 对植物侧生器官形态建成与叶缘形状具有调控作用。基因 Nitab4.5_0001113g0010.1 编码区的第 28 个核苷酸发生单碱基突变(图 6), 突变后恰好产生终止密码子, 翻译过程提前终止, 该基因为 MADS-box 家族基因, 编码一类具有 4 个保守序列的蛋白转录因子, 对植物的生长发育和调控开花时间有重要作用^[16]。基因 Nitab4.5_0007906g0030.1 发生突变, 该基因编码的环焦磷酸合酶参与植物萜类化合物合成的赤藓糖 (MEP)

途径^[17],对烟草萜类物质的合成具有调控作用。

2.8 重测序 SNP 变异位点验证分析

在变异检测结果中随机挑选 16 个 SNP 纯合变

异位点进行验证,结果表明 16 个位点中有 14 个位点与检测结果一致,2 个位点未检测到变异发生,假阳性率为 12.5%(表 6)。

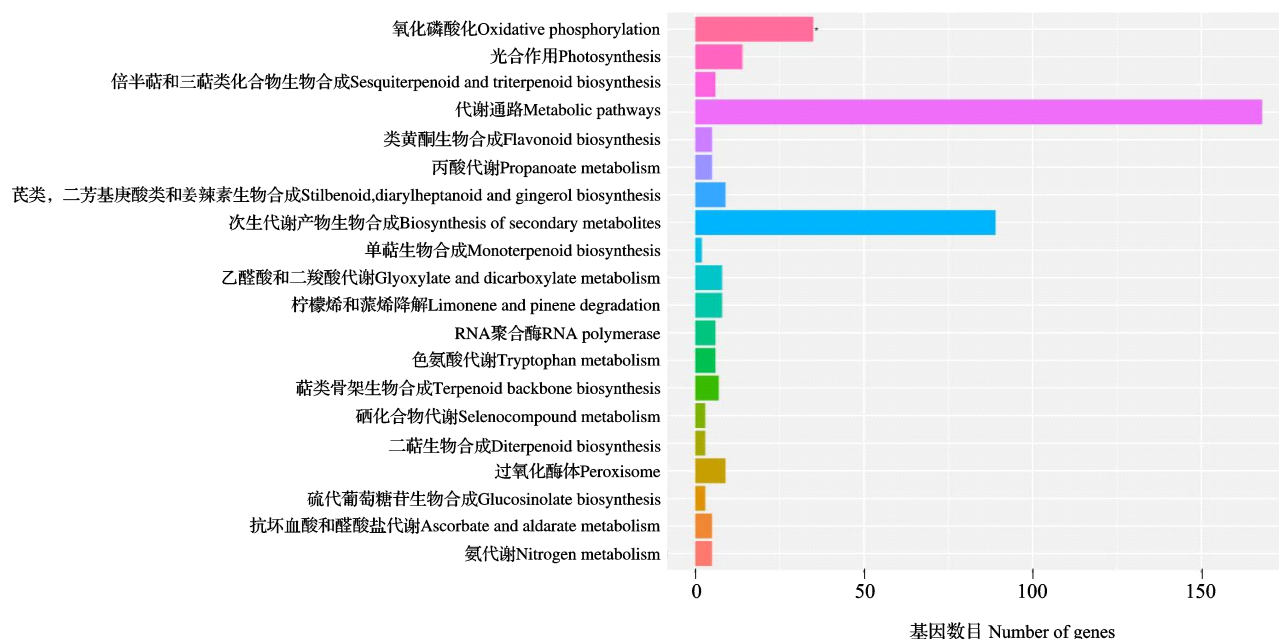
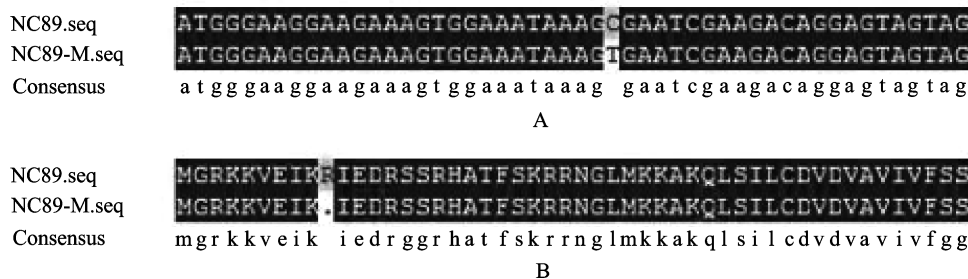


图 5 变异基因的 KEGG 注释分类图

Fig.5 Classification of mutant genes blasted by KEGG database



A: 基因序列比对; B: 蛋白序列比对

A: DNA sequence alignment, B: peotein sequence alignment

图 6 基因 Nitab4.5_0001113g0010.1 序列比对

Fig.6 Sequence alignment of gene Nitab4.5_0001113g0010.1

3 讨论

烟草为异源四倍体,基因组大,遗传组成复杂,加大了变异检测分析的难度。为得到高质量的检测数据及结果,本研究采用的全基因组重测序采取了 30× 的测序深度,测序深度高,克服了之前由于测序深度低无法深入挖掘突变材料遗传信息的不足。分析表明,野生型和突变体共检测到 271655 个 SNP, 23450 个 Indel,结果表明 SNP 数量多于其他类型变异数量,证明航天诱变对烟草基因组的改变

主要以单核苷酸变异为主,与张志勇等^[18]关于水稻航天诱变的研究结论相符。与 EMS 诱变等传统诱变方式变异结果不同,航天诱变突变体检测到大量结构变异,其中以染色体内部迁移和大片段缺失两种类型居多,这可能与航天诱变的机理有关。在变异 SNP 中,转换类型数量明显多于颠换类型,证明航天诱变因子对烟草基因组造成的单核苷酸突变以转换为主。Indel 突变中插入突变类型数目明显多于缺失突变,证明航天诱变产生的插入缺失突变中插入突变是主要因素。

表 6 纯合变异位点验证
Table 6 Verification of homozygous variant sites

项目 Item	染色体(或基因)ID Chromosome(gene)ID	位置 Location	变异发生区域 Mutation area	变异类型 Mutation type	编码氨基酸改变 Amino acid changes
正确验证 Correct verification	Nitab4.5_0000651	353767	基因间区	A>G	
	Nitab4.5_0002829	108928	基因间区	C>A	
	Nitab4.5_0000058	621772	基因间区	T>C	
	Nitab4.5_0000058	622652	基因间区	C>T	
	Nitab4.5_0000002	5157934	基因间区	G>A	
	Nitab4.5_0000164	1306958	基因间区	C>A	
	Nitab4.5_0001653	464337	基因间区	G>A	
	Nitab4.5_0000263	1044444	基因间区	A>G	
	Nitab4.5_0001952g0170.1	339159	内含子	A>G	
	Nitab4.5_0001952g0090.1	155895	外显子	A>G	无
	Nitab4.5_0001008g0010.1	13711	外显子	A>G	无
	Nitab4.5_0004457g0020.1	133009	外显子	T>C	无
	Nitab4.5_0006573g0010.1	79295	外显子	T>C	无
	Nitab4.5_0001562g0030.1	183310	外显子	A>C	Asn413His
	Nitab4.5_0003660	57160	基因间区	A>C	
未正确验证 Incorrect verification	Nitab4.5_0000314g0090.1	1053900	外显子	G>T	

利用诱变方法创制突变体材料,是种质资源创新的一种重要手段。张德建等^[19]使用 Co- γ 射线处理水稻品种“9311”种子,对 M₂ 筛选得到株型、生育期等性状改变明显的突变体材料。本研究利用航天诱变获得突变体材料 NC89-M,与 NC89 相比,NC89-M 的叶型、叶色及花序变异明显。检测到的突变基因 MADS-box 基因对植物花器官调控具有重要作用。Bowman 等^[20]发现在 *apl* (MADS-boxA 类基因)单突变体中,叶片结构取代了花萼,而几乎没有花瓣生成。Mao 等^[21]研究发现 *JOINTLESS* (MADS-box 基因)过表达引起番茄花序无限生长。Yoo 等^[22]研究发现 I 类 MADS-box 基因 *AGL28* 过量表达会造成提前开花。突变体中基因 Nitab4.5_0001113g0010.1 编码区第 28 个碱基由胞嘧啶突变为胸腺嘧啶,产生终止密码子,造成翻译过程提前终止。将该基因经 NCBI 数据库 Blast,结果表明该基因为 *AGL27*,推测该基因的变异可能是导致突变体开花时间延长的因素;突变基因 Nitab4.5_0001021g0050.1 属于 homeodomain

家族 *KNOX1* 亚家族基因,*KNOX1* 基因广泛分布于植物中,*KNOX1* 基因家族在拟南芥中有 4 个成员: *STM*、*BP*、*KNAT2* 和 *KNAT6*,在番茄中有 kn2 和 Tkn1,有研究表明 *KNOX1* 基因能够促进细胞分裂素的合成进而促进植物的横向生长^[23-25]。刘青等^[26]利用麻竹中分离出的 *DIKNOX1* 基因转化拟南芥,转基因植株出现叶片缺刻、扭曲,植株分枝增多等表型;目前在烟草中对 *KNOX1* 基因的研究尚无报道,推测该基因的缺失可能与突变体侧生器官发育受到抑制及叶缘形状的改变有关。在 NC89-M 的突变基因中,发现与烟草萜类化合物合成相关基因 Nitab4.5_0007906g0030.1 发生突变,而萜类化合物是烟草香气品质的重要影响因素^[27]。因此推测 NC89-M 可能在香气品质方面较 NC89 存在差异,可补充试验进行验证。

烟草育种以抗病、优质、稳产作为主要育种目标。产量指标包括单株叶数、单叶重和株型等要素^[28]。本研究的突变体材料 NC89-M 株型叶形变异较大。对 NC89-M 田间性状调查结果表明,其植

株略高,叶数适中,节距均匀,腰叶面积略有增大,株型合理舒展,烟株受光良好;烤后烟叶性状调查结果表明,NC89-M 烤后烟叶颜色桔黄、亮度适中、油分好,较 NC89 烤后烟叶外观品质有所改善。对 NC89-M 烤烟评吸鉴定后发现,NC89-M 香气量足,香气质较 NC89 烤烟有明显改善,质量档次显著提高。因此,NC89-M 可作为性状优良的突变体材料,在改良烟草株型、叶形品质等方面具有重要的利用价值。

本研究通过航天诱变创制了优良种质资源,并在表型鉴定的基础上,利用基因组重测序进行了变异检测分析,在基因组水平上分析了变异位点,挖掘可能存在的变异基因,为基因功能验证及优良变异资源在烟草育种中的利用奠定了基础。

参考文献

- [1] 尚志强. 中国烟草育种进展及发展对策. 内蒙古农业科技, 2010(5): 1-3
Shang Z Q. Advances and development strategies of burley tobacco breeding in China. Inner Mongolia Agricultural Science and Technology, 2010(5): 1-3
- [2] 钟波, 朱列书, 贺鹏, 谢慧玲, 林跃平. 浅谈航天诱变育种. 作物研究, 2007(51): 511-516
Zhong B, Zhu L S, He P, Xie H L, Lin Y P. Discussion on space mutation breeding. Crop Research, 2007(51): 511-516
- [3] 刘敏, 李金国, 王亚林, 张赞, 薛淮, 张纯花, 李社荣, 郝连元, 李广元. 卫星搭载的甜椒 87-2 过氧化物同工酶检测和 RAPD 分子检测初报. 核农学报, 1999, 13(5): 291-294
Liu M, Li J G, Wang Y L, Zhang Z, Xue H, Zhang C H, Li S R, Hao L Y, Li G Y. Preliminary study on peroxidase isoenzyme detection and RAPD molecular verification for sweet pepper 87-2 Carried by a recoverable satellite. Journal of Nuclear Agricultural Sciences, 1999, 13(5): 291-294
- [4] Cheng Z, Lin J, Lin T, Xu M, Huang Z, Yang Z, Huang X, Zheng J. Genome-wide analysis of radiation-induced mutations in rice (*Oryza sativa* L. ssp. indica). Molecular Biosystems, 2014, 10(4): 795
- [5] Fu C Y, Liu W G, Liu D L, Li J H, Zhu M S, Liao Y L, Liu Z R, Zeng X Q, Wang F. Genome-wide DNA polymorphism in the indica rice varieties RGD-7S and Taifeng B as revealed by whole genome re-sequencing. Genome, 2016, 59(3): 197-207
- [6] Maldonado dos Santos J V, Valliyodan B, Joshi T, Khan S M, Liu Y, Wang J, Vuong T D, de Oliveira M F, Marcelino-Guimarães F C, Xu D, Nguyen H T, Abdelnoor R V. Evaluation of genetic variation among Brazilian soybean cultivars through genome resequencing. BMC Genomics, 2016, 17(1): 110
- [7] Sierro N, Battey J N, Ouadi S, Bakaher N, Bovet L, Willig A, Goepfert S, Peitsch M C, Ivanov N V. The tobacco genome sequence and its comparison with those of tomato and potato. Nature Communications, 2014, 5(5): 3833
- [8] Bindler G, Plieske J, Bakaher N, Gunduz I, Ivanov N, Van der Hoeven R, Ganai M, Donini P. A high density genetic map of tobacco (*Nicotiana tabacum* L.) obtained from large scale microsatellite marker development. Theoretical & Applied Genetics, 2011, 123(2): 219
- [9] Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics, 2009, 25(14): 1754-1760
- [10] Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R; 1000 Genome Project Data Processing Subgroup. The Sequence Alignment/Map format and SAMtools. Bioinformatics, 2009, 25(16): 2078-2079
- [11] Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. Nucleic Acids Research, 2010, 38(16): e164
- [12] Chen K, Wallis J W, McLellan M D, Larson D E, Kalicki J M, Pohl C S, McGrath S D, Wendl M C, Zhang Q, Locke D P, Shi X, Fulton R S, Ley T J, Wilson R K, Ding L, Mardis E R. BreakDancer: an algorithm for high-resolution mapping of genomic structural variation. Nature Methods, 2009, 6(9): 677-681
- [13] Kanehisa M, Goto S, Kawashima S, Okuno Y, Hattori M. The KEGG resource for deciphering the genome. Nucleic Acids Research, 2004, 32(Nucleic Acids Research): D277-D280
- [14] Altschul S F, Madden T L, Schäffer A A, Zhang J, Zhang Z, Miller W, Lipman D J. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Research, 1997, 25(17): 3389-3402
- [15] Hake S, Smith H M, Holtan H, Magnani E, Mele G, Ramirez J. The role of knox genes in plant development. Annual Review of Cell and Developmental Biology, 2004, 20(1): 125-151
- [16] Messenguy F, Dubois E. Role of MADS box proteins and their cofactors in combinatorial control of gene expression and cell development. Gene, 2003, 316(1): 1-21
- [17] Rodrı́a-Guez-Concepci3 N M, Boronat A. Elucidation of the methylerythritol phosphate pathway for isoprenoid biosynthesis in bacteria and plastids. A metabolic milestone achieved through genomics. Plant Physiology, 2002, 130(3): 1079-1089
- [18] 张志勇, 蒲志刚, 王平, 向跃武, 蔡平钟, 张志雄. 水稻航天诱变突变体全基因组测序研究. 西南农业学报, 2014, 27(2): 469-475
Zhang Z Y, Pu Z G, Wang P, Xiang Y W, Cai P Z, Zhang Z X. Sequencing research on whole genome of rice mutant induced by space. Southwest China Journal of Agricultural Sciences, 2014, 27(2): 469-475
- [19] 张德建, 梁发茂, 胡刚, 邢永忠, 李志新. γ 射线辐射“9311”水稻突变体的筛选. 植物遗传资源学报, 2015, 16(4): 876-882
Zhang D J, Liang F M, Hu G, Xing Y Z, Li Z X. Screening of mutant for rice cultivar “9311” by γ -ray irradiation. Journal of Plant Genetic Resources, 2015, 16(4): 876-882
- [20] Bowman J L, Alvarez J, Weigel D, Meyerowitz E M, Smyth D R. Control of flower development in *Arabidopsis thaliana* by APETALA1 and interacting genes. Development, 1993, 119

- (3): 721-743
- [21] Mao L, Begum D, Chuang H W, Budiman M A, Szymkowiak E J, Irish E E, Wing R A. JOINTLESS is a MADS-box gene controlling tomato flower abscission zone development. *Nature*, 2000, 406 (6798): 910-913
- [22] Yoo S K, Lee J S, Ahn J H. Overexpression of AGAMOUS-LIKE 28 (AGL28) promotes flowering by upregulating expression of floral promoters within the autonomous pathway. *Biochem Biophys Res Commun*, 2006, 348 (3): 929-936
- [23] Ori N, Juarez M T, Jackson D, Yamaguchi J, Banowitz G M, Hake S. Leaf senescence is delayed in tobacco plants expressing the maize homeobox gene knotted1 under the control of a senescence-activated promoter. *Plant Cell*, 1999, 11 (6): 1073-1080
- [24] Hewelt A, Prinsen E, Thomas M, Van Onckelen H, Meins F Jr. Ectopic expression of maize knotted1 results in the cytokinin-autotrophic growth of cultured tobacco tissues. *Planta*, 2000, 210 (6): 884-889
- [25] Frugis G, Mariotti D. Overexpression of KNAT1 in lettuce shifts leaf determinate growth to a shoot-like indeterminate growth associated with an accumulation of isopentenyl-type cytokinins. *Plant Physiology*, 2001, 126 (4): 1370-1380
- [26] 刘青, 汪玉凤, 赵韩生, 陈颖, 高志民. 麻竹同源异型盒基因 DIKNOX1 的克隆及功能初步分析. *林业科学*, 2014, 50 (2): 56-62
- Liu Q, Wang Y F, Zhao H S, Chen Y, Gao Z M. Molecular characteristics and primary functional analysis of DIKNOX1 gene from dendrocalamus latiflorus. *Scientia Silvae Sinicae*, 2014, 50 (2): 56-62
- [27] 申培林, 杨铁钊, 张小全, 李亚培, 贾燕超, 宋文峰, 吕军朋. 烟草类萜生物合成途径中关键酶基因的克隆与表达调控. *江苏农业科学*, 2012, 40 (1): 37-40
- Shen P L, Yang T Z, Zhang X Q, Li Y P, Jia Y C, Song W F, Lv J P. Cloning and expression regulation of key enzyme genes in tobacco terpenoids biosynthesis pathway. *Jiangsu Agricultural Sciences*, 2012, 40 (1): 37-40
- [28] 张崇范. 烤烟育种目标刍议. *中国烟草科学*, 1986 (3): 21-23
- Zhang C F. Discussion on breeding objectives of flue-cured tobacco. *Chinese Tobacco Science*, 1986 (3): 21-23