

Advances in Research on the Location and Cloning of Maize Nuclear Male Sterility Gene

Meng Ma

College of Plant Science of Huazhong Agricultural University Agronomy, Wuhan Hubei
Email: 1015371206@qq.com

Received: May 8th, 2020; accepted: May 22nd, 2020; published: May 29th, 2020

Abstract

Maize is a very important food crop and economic crop. In corn production, the main application is hybrids, and male sterility has gradually become the trend of hybrid seed production due to its own advantages. In recent decades, researchers have made unremitting efforts to study the male sterility mutants of maize nuclear nucleus, and have located, cloned and functionally studied male sterility genes. At present, sterility genes have been cloned. This article mainly reviews the location, cloning and function of maize nuclear male sterility genes, and briefly introduces the more commonly used gene mapping technology and the molecular markers involved in recent years.

Keywords

Maize, Nuclear Male Sterility, Gene Mapping, Gene Cloning, BSA, Molecular Markers

玉米细胞核雄性不育基因定位及克隆的研究进展

马 萌

华中农业大学, 植物科学技术学院, 湖北 武汉
Email: 1015371206@qq.com

收稿日期: 2020年5月8日; 录用日期: 2020年5月22日; 发布日期: 2020年5月29日

摘 要

玉米是一种非常重要的粮食作物和经济作物, 在玉米生产中, 主要应用的是杂交种, 而雄性不育因其自身的优势逐渐成为杂交制种应用的趋势。近几十年来, 科研人员在对玉米细胞核雄性不育突变体的研究

进行了持续不懈的努力,对雄性不育基因进行定位、克隆及功能研究,目前已经有不育基因被克隆出来。本文主要对玉米细胞核雄性不育基因定位克隆和功能等进行了综述,并简要的介绍近些年比较常用的基因定位技术以及涉及到的分子标记。

关键词

玉米, 细胞核雄性不育, 基因定位, 基因克隆, 集团分离分析法, 分子标记

Copyright © 2020 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

玉米(*Zea mays*)是一种相当重要的粮食作物,也是非常重要的经济作物。在玉米的种植中,应用的大多为杂交种。作物育种过程中普遍利用杂种优势,其可以明显的增加作物产量、质量和抗性等。在杂交制种过程中母本去雄是最关键的一个环节,目前常用的母本去雄技术有人工去雄、化学杀雄、机械去雄和雄性不育。其中人工去雄、化学杀雄和机械去雄成本比较高,而且去雄不及时不彻底,还会对母本造成一定程度的伤害。而雄性不育植株不需要去雄,成为杂交制种应用的趋势[1]。

植物雄性不育(male sterility, MS)是指高等植物雄性器官存在发育异常,雄配子(花粉)不能正常行使功能,但雌性器官能接受正常雄配子,从而受精结实,并将雄性不育性状遗传给后代的现象[2]。按照遗传或起源,可以将雄性不育分为细胞核雄性不育、细胞质雄性不育和核质互作雄性不育三类[1]。细胞核雄性不育(Genic Male Sterility, GMS)是指雄性不育是由核基因控制的,不受细胞质的母系遗传的影响,没有正反交的遗传效应。根据遗传特征可将细胞核雄性不育分为显性核雄性不育和隐性核雄性不育两类,目前报道的核不育材料大部分为隐性核不育类型。细胞质雄性不育(Cytoplasmic male sterility, CMS)是母系遗传,由线粒体或叶绿体基因控制。核质互作雄性不育,由细胞核基因和细胞质内的线粒体基因等共同控制[3]。

现如今,在玉米的生产中,核质互作雄性不育材料应用的更为广泛,被广泛使用,这种材料是利用三系法生产杂交种,但是此类材料存在一定的缺点,例如:细胞质单一,败育不够彻底,需要特定的恢复基因,存在比较大的病原小种专业化侵染风险。而细胞核雄性不育材料是非常好的雄性不育材料,细胞核雄性不育是由单基因控制的,不存在核质互作雄性不育材料的不足[4],如果能够得以应用将是杂交生产中的一大飞跃。

1921年 Eyster 首次发现玉米无花粉型雄性不育突变体 *ms1* 之后,陆续发现鉴定命名了 40 多个雄性不育突变体[5]。其中隐性核不育突变体占绝大部分,成为研究的热点,而最终能够得以应用的也主要为隐性核不育突变体,据不完全统计,至少有 15 个玉米隐性核不育基因完成了克隆和功能分析(表 1),分别为 *Ms1* [6]、*Ms7* [7]、*Ms8* [8]、*Ms9* [9]、*Ms22/Msca1* [10]、*Ms23* [11]、*Ms26* [12]、*Ms30* [13]、*Ms32* [14]、*Ms33* [15]、*Ms45* [16]、*Ocl4* [17]、*Mac1* [18]、*Ipe1* [19]、*Apv1* [20]。基因定位中常用 BSA 技术进行初步定位,再利用相应的分子标记技术进行精细定位,近几年常用的分子标记有 SSR、CAPS、SNP、InDel。

本文简要的对上述已完成克隆和功能分析的玉米隐性核不育基因进行综述,并简要的介绍近些年比较常用的基因定位技术以及涉及到的分子标记。为玉米细胞核雄性不育突变体的研究及相关基因的定位提供了一定的理论价值。

Table 1. The genic male sterility genes that has been cloned in maize
表 1. 已克隆的玉米细胞核雄性不育基因

基因 Gene	基因模型 Gene model	染色体 Chromosome	突变体 Mutant	自交系 Inbred line	标记类型/遗传距离 Marker type/Genetic distance	
<i>Ms1</i>	GRMZM2G180319	chr6(L)	<i>ms1-6050</i>	昌 7-2(C7-2)	SSR/233kb	
<i>Ms7</i>	GRMZM5G890224	chr7(L)	<i>ms7gl1 ms7-6007</i>	昌 7-2(C7-2)	SSR/4.8cM	CAPs/180kb
<i>Ms8</i>	GRMZM2G119265	chr8(L)	/	/	/	/
<i>Ms9</i>	GRMZM2G308204	chr1(S)	<i>ms9-ref ms9-AD62A</i>	/	/	/
<i>Ms22/Msca1</i>	GRMZM2G442791	chr7(S)	<i>msca1-ref msca1-mg12 msca1-6036</i>	/	/	/
<i>Ms23</i>	GRMZM2G021276	chr8(S)	<i>ms23-ref ms23-6027 ms23-6059</i>	B73 A619	/	SSR/0.75mb
<i>Ms26</i>	ZEMMB73_004940	chr1(S)	/	/	/	/
<i>Ms30</i>	GRMZM2G174782	chr4(L)	<i>ms30-6028</i>	昌 7-2(C7-2)	SSR/2.9cM/41.2kb	
<i>Ms32</i>	GRMZM2G163233	chr2(L)	<i>ms32-ref ms32-6066</i>	/	/	/
<i>Ms33</i>	GRMZM2G070304	chr2(L)	<i>ms33-6029</i>	昌 7-2(C7-2)	SSR/11.5cM/349kb	
<i>Ms45</i>	GRMZM2G139372	chr9(L)	/	/	/	/
<i>Ocl4</i>	GRMZM2G123140	chr1(S)	<i>ocl4-1 ocl4-2</i>	A188	/	/
<i>Mac1</i>	GRMZM2G027522	chr10(S)	<i>mac1-1 mac1-Y211</i>	W23 B73	/	/
<i>Ipe1</i>	GRMZM2G434500	chr1(S)	<i>ipe1 ipe1-1 ipe1-2 ipe1-3</i>	郑 58	/	InDel/290kb
<i>Apv1</i>	GRMZM5G830329	chr10(L)	<i>apv1</i>	B73	/	/

2. 已克隆玉米隐性核不育基因

2.1. *Ms1*

Ms1 被定位在 6 号染色体长臂。Zhou 等[6]将突变体 *ms1-6050* 与自交系昌 7-2(C7-2)杂交, 构建了 F₂ 代分离群体。利用 F₂ 代群体的 218 个雄性不育个体, 对 *ms1* 位点进行遗传分析。已知 *ms1* 位点与位于 6 号染色体的 Y1 基因紧密连锁[22], 筛选 Y1 基因附近位于玉米 6 号染色体 6.02bin 的 4 个多态性 SSR 分子标记, 并进行 PCR 扩增以及凝胶电泳分析, 表明 4 个多态性标记与 *Ms1* 基因紧密连锁。利用 F₂ 代群体的 654 个雄性不育个体进行精细定位, 基于 B73 基因组序列开发了 8 对与上述 4 对标记相邻的多态性标记, 构建了 *ms1* 位点所在区间的连锁图谱。根据表型的数据, 进行连锁交换分析, 最终将 *ms1* 位点定位在 sj10-sj30 之间, 遗传距离为 233 kb。该区域有 11 个候选基因, 其中 GRMZM2G180319 基因具有与花器官发育相关的功能, 为 *Ms1* 基因。*Ms1* 基因的长度为 1286 bp, 包括有 2 个外显子和 1 个内含子, 编码 1 个 218 个氨基酸的蛋白。该蛋白中存在一个 DUF260 保守结构域, 属于次生器官发育相关基因(Lateral Organ Boundary, LOB)家族。突变基因 *ms-6050* 在+468~+474 处存在一个碱基的缺失, 导致编码区移码突变, 造成基因的翻译在+494 处提前终止。基因在 *ms1-6050* 突变体中编码的蛋白少了 74 个氨基酸, 从而造成其功能的丧失。

2.2. *Ms7*

Ms7 被定位在 7 号染色体长臂。Zhang 等[7]将突变体 *ms7gl1*、*ms7-6007* 分别与自交系昌 7-2(C7-2)杂交, 并借助分子标记辅助鉴定, F₁ 后代均为雄性可育, F₂ 代群体雄性可育植株与不育植株符合 3:1 分离比, 表明 *ms7gl1* 和 *ms7-6007* 均为单基因隐性遗传。*ms7-6007/ms7-6007* 纯合植株用 *ms7gl1*/+杂合植株

的花粉授粉, F1 代群体雄性可育植株与不育植株符合 1:1 分离比, 表明 *ms7-6007* 的突变基因与 *ms7gl1* 突变基因是等位基因。利用突变体 *ms7gl1* 与自交系昌 7-2(C7-2)杂交得到的 F2 代群体的 154 个雄性不育个体进行初步定位, 利用 BSA 技术, 筛选 SSR 标记, 并对突变位点进行遗传分析, 有 9 对 SSR 标记与 *Ms7* 基因连锁, *Ms7* 基因被初定位在 SSR 标记 umc2617 和 bnlg1808 之间, 遗传区段为 4.8 cM。利用 F2 代群体的 611 个雄性不育个体进行精细定位, 在 umc2617 和 bnlg1808 之间设计了 6 对 CAPS 引物, 将 *Ms7* 定位在 EP299 和 EP302 之间, 遗传距离为 180 kb。在这个区间中有 6 个候选基因, 其中标记 EP239 与不育性状表现为共分离, 该序列来源基因为 GRMZM5G890224, 该基因编码一个 PHD-finger 结构域的蛋白。*Ms7* 基因全长 2435 bp, 由三个外显子和两个内含子组成, 编码 670 个氨基酸蛋白质, 第三个外显子编码的氨基酸能够形成保守的 PHD-finger 结构域。通过对突变体 *ms7-6007*、*ms7gl1* 中目的基因 DNA 进行测序, 突变基因 *ms7-6007* 全长 2453 bp, 在第一个外显子的+22 位点插入了 3 bp, 在第二个外显子的+814 位点插入了 7 bp, 蛋白翻译发生移码在+1426 处终止导致缺乏亮氨酸拉链区域和 PHD 结构域; 突变基因 *ms7gl1* 全长 3543 bp, 在第三个外显子的+1179 位点插入了 1136 bp 转座子(DTA_ZM00023), 翻译在+1225 处提前终止, 并导致缺乏亮氨酸拉链区域和 PHD 结构域。这些结果表明 GRMZM5G890224 为引起 *ms7* 突变体不育基因。

2.3. *Ms9*

Ms9 被定位在 1 号染色体短臂。Albertsen 等[9]利用由大约 450 个个体组成的小群体鉴定基因连锁的两侧标记, 利用从小群体中鉴定的两侧标记和新设计的标记鉴定重组子, 确定 *ms9* 基因在 1 号染色体的物理间隔。*Ms9* 的一个候选基因被发现是一个 R2/R3 植物特异性 myb 转录因子。突变基因 *ms9-ref* 在第一个外显子中有 4 个碱基对插入, 导致翻译移码突变。这种突变发生在 R2 结合域; 突变基因 *ms9-AD62A* 在第三个外显子中有一个 16-bp 缺失, 扰乱了 R3 结合结构域。

2.4. *Ms22/Msca1*

Ms22/Msca1 被定位在 7 号染色体短臂。Chaubal 等[21]利用 *mscal-ref* 等位基因, 用标记物对该群体进行定位, 并用跨越 2 个 BAC(细菌人工染色体)克隆的标记 pco144723 和 b49p15.f 确定遗传距离为 155 kb。BACs 的测序和其他标记的发展将遗传距离缩小到 9 kb。这个区域的序列显示只有一个开放的阅读框, 编码一个植物特异性谷胱甘肽还蛋白基因。突变基因 *mscal-ref* 谷胱甘肽还蛋白基因缺失, 有 7823 bp 缺失, 谷胱甘肽下游 4 kb 处有 1268 bp 插入。对 *mscal-mg12* 等位基因的克隆和序列分析显示, 谷胱甘肽基因 3'区有 490 bp 的缺失。*Mscal1* 基因中存在一个氧化还原区的序列 CCMC, 不育植株缺少 GSH 结合位点, 而在可育植株中存在的 GSH 结合区 LPVVVGGRLLG 在不育突变体中不存在。在 *mscal-6036* 等位基因的克隆和测序中, 检测到了有 850 bp 的插入, 这一插入创造了一个 8 bp 的宿主位点重复(GTCGAGA), 并且它似乎还包含小的完美的 TIRs [10]。

2.5. *Ms23*

Ms23 被定位在 8 号染色体短臂。Nan 等[11]利用 SNP-genotyping 技术对 *Ms23* 进行高通量遗传定位, *Ms23* 最初被定位在 6 号和 8 号染色体上。检测两个相关区域两侧的 PCR 标记, 并确认该雄性不育表型与 8 号染色体上 8.00~8.01 bins 附近的多态性标记共同分离。利用 *ms23-ref*、*ms23-6027* 或 *ms23-6059* 突变体各 250 个以上雄性不育个体进行定位。*Ms23* 被定位在 8S 染色体末端的 0.75 mb 范围内。该区域有 15 个候选基因, 利用 qTeller 检索 RNA 序列数据, 发现只有 GRMZM2G021276, 跨越位置 95823 到 98367, 编码预测的 bHLH 蛋白, 显示 tassel 特异表达。*Ms23* 基因全长 2361 bp, 有四个外显子, 编码 365 个氨

基酸的 bHLH 蛋白。在突变基因 *ms23-ref* 中, 整个基因被删除, 这个删除包含高达 96.5 kb。突变基因 *ms23-6027* 在 1076 位插入 2-bp(AT), 使在预测的 bHLH 区域的正上游位置 1091 处有一个早期终止密码子, 引起移码突变。*Ms23* 编码一个花药特异性的、预测为绒毡层分化所需的碱性螺旋环螺旋(bHLH)转录因子。RNA-seq 和蛋白质组学数据以及酵母双杂交分析表明, *Ms23* 与 *Ms32*、bHLH122 和 bHLH51 一起依次作为同二聚体或异二聚体来指导绒毡层发育。其中, *Ms23* 是最早的作用因子, 位于 bHLH51 和 bHLH122 的上游, 控制着绒毡层的规格和成熟度。相比之下, *Ms32* 是组成性的, 独立调节的, 在绒毡层分化中比 *Ms23* 晚。

2.6. *Ms30*

Ms30 被定位在 4 号染色体长臂。Zhou 等[13]将突变体 *ms30-6028* 与自交系昌 7-2(C7-2)杂交, 构建了 F₂ 代分离群体。利用 F₂ 代群体的 118 个雄性不育个体, 利用 BSA 技术, 按照 maizeGDB 中标记数据库信息, 筛选 SSR 分子标记, 对突变位点进行遗传分析, 有 5 对多态性分子标记与 *ms30* 位点连锁, *Ms30* 基因初步定位在 *umc1808* 和 *umc2046* 之间。另取 F₂ 代群体 284 株, 对其进行育性调查, 雄性可育株与不育株比符合 3:1 的单基因分离比。从 maizeGDB 中检索到 *umc1808* 至 *umc2046* 之间的 9 对多态标记(包括 *umc1808* 和 *umc2046*), 对该 F₂ 群体进行了基因型鉴定, 构建了 *ms30* 位点遗传连锁图谱, 并根据单株的育性数据, *Ms30* 被初定位在 *idp1991* 至 *tidp9194* 之间, 遗传距离为 2.9 cM。对 *Ms30* 进行精细定位, 调取 B73 基因组序列并设计了多态性标记 *ep462*, 将 *Ms30* 定位在 *ep462* 和 *tidp3747* 之间, 遗传距离为 41.2 kb。该区域有 6 个候选基因, 其中一个预测为 GDSL 脂肪酶家族基因 GRMZM2G174782 具有与花器官发育相关的功能, 为 *Ms30* 基因。该基因有 2370 bp 长, 含有 3 个外显子和 2 个内含子, 有 1277bp 的蛋白质编码区序列。DNA 测序结果表明, 突变基因 *ms30-6028* 的蛋白质编码序列有 4 处核苷酸存在插入或缺失, 突变基因在 1、2、4 处分别插入 1、3、9 个碱基, 在第 3 处突变缺失了 4 个碱基。

2.7. *Ms32*

Ms32 被定位在 2 号染色体长臂。Chaubal [21]等将突变基因 *ms32-ref* 初部定位在 2 号染色体 UNC139 标记的远端。Moon 等[14]利用来自 1:1 分离群体的 8 个雄性可育个体和 9 个雄性不育个体的 gDNA 池, 利用 22 对 InDel 引物进行初定位, 遗传距离为 10 cM, 其中 2 号染色体上的引物对 IDP792 表现出多态性。利用 249 个雄性不育个体组成的更大的定位群体, 利用 BSA 法, 利用 24 对 IDP 引物进行定位, 将 *Ms32* 定位在标记 *umc1049* 和 *JM16* 之间, 遗传距离为 0.5 mb。在该区域有 15 个候选基因。在 GRMZM2G1633233 基因模型中发现 *ms32-ref* 突变基因携带大于 1.6 kb 的缺失, GRMZM2G1633233 基因的长度约为 10 kb, 包含四种带注释的转录类型, 其中一种转录类型与 bHLH 蛋白的 N-末端结构域和液泡 ATP 合酶亚基 H 的 C-末端相似, 而水稻、苞片和高粱基因组中 bHLH 基因和液泡 ATP 合酶亚单位 H 基因被注释为单独的基因。GRMZM2G1633233_T02 和 GRMZM2G1633233_T03 对应于 bHLH 结构域, 而 GRMZM2G1633233_T04 对应于液泡 ATP 合酶亚基, 另外对 EST 序列和 PlantGDB 组装的唯一转录本研究, 结果表明 GRMZM2G1633233 是两个基因的嵌合体, 其中一个是 *Ms32*。*ms32-ref* 中的缺失被证实包括 bHLH 基因的 C 末端。GRMZM2G1633233 是两个基因: GRMZM2G1633233_T02 是编码 bHLH 蛋白的 *Ms32* 基因, GRMZM2G1633233_T04 是液泡 ATP 合酶亚单位 H 基因。*Ms32* 基因编码一种基本的螺旋-环-螺旋(bHLH)转录因子, 在花药发育过程中起着重要的分裂和分化调控作用。*Ms32* 基因由四个外显子和三个内含子构成。候选基因(GRMZM2G1633233_T02)由四个外显子组成, 它们被预测编码 219 个氨基酸蛋白。*ms32-ref* 中的缺失在第 2 外显子之后立即开始, 并延伸到第 4 外显子之后, 从而使 N-末端 bHLH 域保持完整, 只有该基因剩余的 N-末端区域(外显子 1 和 2)仍在 *ms32-ref* 中表达。

2.8. *Ms33*

Ms33 被定位在 2 号染色体长臂。已知 *ms33* 位点位于 2 号染色体长臂 [22]。Zhang 等 [15] 按照 maizeGDB 中标记数据库信息, 筛选出 5 个在 *ms33* 位点呈多态性的 SSR 分子标记, 利用 228 株雄性不育植株进行 PCR 扩增和凝胶电泳分析, 检测不育基因与标记物之间的连锁反应。结果表明, 这 5 个标记与 *ms33* 位点密切相关。利用从 maizeGDB 检索到的 9 对来自 mmc0381 到 umc2214 之间的多态性标记, 对 F₂ 群体的单株进行基因分型, 构建了 *ms33* 位点所在区域的遗传连锁图谱。根据单株的育性资料, *Ms33* 基因被初步定位在 EP97 和 bnlgl1893 之间, 遗传距离为 11.5 cM。检索 B73 基因组序列, 设计了 4 对多态性标记 EP198、EP603、EP605 和 EP480。最后, *Ms33* 被定位于 EP603 和 EP605 之间, 遗传距离为 349kb。该区间有 15 个候选基因, 其中 GRMZM2G070304 基因为 GPAT 家族基因, 具有与花器官发育相关的功能, 为 *Ms33* 基因。该基因全长 2198 bp, 含有 2 个外显子和 1 个内含子。突变体 *ms33-6029* 中突变基因 *ms33-6029* 全长 3566 bp, 在第一个外显子+223 处插入了一个长为 1368 bp 的转座子(DTA_ZM00216), 基因翻译在 +288 处提前终止。在野生型 B73 中, *Ms33* 基因的蛋白质编码序列为 1578 bp 长。通过该基因的 cDNA 序列的模拟翻译, 发现该基因编码 525 个氨基酸的蛋白质。蛋白质序列和 GenBank SwissProt 蛋白质数据库的 blastx 分析显示, 该蛋白质含有 Lysophospholipidacyltransferases (LPLAT) 保守结构域。在突变体 *ms33-6029* 中, 由于转座子插入了 *Ms33* 基因序列的核苷酸 223 导致氨基酸 75 后的改变, 编码的蛋白质仅包含 95 个氨基酸, 并且突变基因编码的蛋白质不再具有任何功能域。该实验结果显示突变体中 *Ms33* 基因的序列的改变直接导致翻译氨基酸的序列变异并提前终止翻译, 尤其会破坏基因的 LPLAT 结构域。

2.9. *Ocl4*

Ocl4 被定位在 1 号染色体短臂。*Ocl4* 基因中包括 10 个外显子, 编码 881 个氨基酸的假定蛋白, 包含 HD-ZIP-IV 蛋白的四个特征结构域, 即 home、leu-cine-zipper、START 和 HD-SAD 结构域。突变基因 *ocl4-1* 在外显子 6 中插入了转座子 Mu8, 在玉米(TUSC)的先锋性状实用系统中发现, 而突变基因 *ocl4-2* 在外显子 5 和内含子 5 的接合处插入了 Mu8, 源于生物膜突变型集合 [17]。

2.10. *Mac1*

Mac1 被定位在 10 号染色体短臂。Sheridan [23] 等将 *Mac1* 定位到 10S 染色体, 在遗传标记 oy1 和 T9-10b 之间。该基因被定位在标记 TIDP8865 和 IDP8390 之间, 遗传距离为 900 kb。该区域有 14 个候选基因, 其中基因模型 GRMZM2G027522 与水稻 *TDL1A* 和拟南芥 *TPD1* 基因相似, 缺失等位基因导致与 *mac1* 突变体有相似的表型, GRMZM2G027522 模型为 *Mac1* 基因。利用引物 P222、P223 和 P220、P223 对突变基因 *mac1-1* 进行基因型鉴定, 利用引物 P222、P223、P258 和 P292 对突变基因 *mac1-Y211* 进行基因型鉴定。*Mac1* 基因中包括 4 个外显子和 3 个内含子; 通过 DNA 序列分析, 发现 *mac1-1* 等位基因中有一个 1263 bp 的缺失, 从内含子 2 开始, 延伸到外显子 4; *mac1-Y211* 等位基因内含子 2 中插入了 4.9kb 的自主 MuDR 元件 [18]。

2.11. *Ipe1*

Ipe1 被定位在 1 号染色体短臂。Chen 等 [19] 前期将 *Ipe1* 基因初步定位在 SSR 分子标记 S1 和 S11 之间。后期对 *Ipe1* 基因进行精细定位, 开发了 9 对多态性 InDel 标记, 利用 F₂ 群体的 4021 个突变植株, 最终将目的基因组区域缩小到 InDel 标记 S4 和 S8 之间, 遗传距离约 290kb。该区域有 6 个候选基因, 通过对 *ipe1* 基因组 DNA 进行测序以及利用基因特异性引物 P123F 和 Mu1 转座子特异引物 I10R 的组合, 表明突变体 GRMZM2G434500 基因中 Mu1 转座子插入在了第三个外显子开始的第 64 个碱基对处。通过

对 *ipe1-1*、*ipe1-2*、*ipe1-3* 基因组测序组测序进一步证明 GRMZM2G434500 为引起 *ipe1* 突变体不育基因。利用 cDNA 测序, 表明 *Ipe1* 基因由三个外显子和两个内含子组成, 编码一个含 582 个氨基酸蛋白。利用 SMART 进行蛋白质结构域分析以及与 NCBI 数据库中全长蛋白进行比对分析, 表明在 *Ipe1* 基因 N 端氨基酸和 C 端氨基酸包含两个 GMC 氧化还原酶功能域, 即 *Ipe1* 编码 GMC 氧化还原酶。

2.12. *Apv1*

Apv1 被定位在 10 号染色体长臂。Zhang 等[20]将突变体 *apv1* 与自交系 B73 杂交, F_1 后代均为雄性可育, BC_1F_1 (以突变体 *apv1* 为杂交亲本)中 249 个个体雄性可育植株与不育植株符合 1:1 分离比, F_2 代群体中 311 个个体雄性可育植株与不育植株符合 3:1 分离比, 表明 *apv1* 为单基因隐性遗传。利用 F_2 代群体的 480 个雄性不育个体进行初定位, 设计了一系列多态性标记对, 利用分子标记 YS-38 和 YS-39 将基因定位区域缩小到 1.32 Mbps。利用 2012 年得到的 BC_1F_1 的不育个体组成的更大的定位群体进行精细定位, 利用分子标记 YS-48 和 YS-75 将 *Apv1* 基因定位区域缩小到 660 kb。有 21 个基因在这个区间中, 而在减数分裂雄穗中表达的有 5 个基因, 其中 GRMZM5G830329 在减数分裂雄穗中高表达。对目的基因进行 DNA 测序, *apv1* 突变基因有 848 bp 的缺失, 该缺失跨越以下两个相邻基因: GRMZM5G830329 和 GRMZM2G439268。该缺失始于 GRMZM5G830329 的外显子 2, 延伸至 GRMZM2G439268 的外显子 1。利用 CRISPR/Cas9 对 GRMZM5G830329 基因和 GRMZM2G439268 基因敲除, 并对 GRMZM5G830329 敲除和对 GRMZM2G439268 敲除个体花药使用 I2-KI 染色观察, 前者不能观察到黑色染色的花粉粒, 后者观察到有染色良好的花粉粒。GRMZM5G830329 基因敲除系(ms/ms)与 *apv1*(ms/+)的杂交以 1:1 的比例产生雄性不育和可育植株($v_2 = 0.05$, $0.5 < P < 0.9$)。这一结果证实了 GRMZM5G830329 的缺失导致 *apv1* 雄性不育, 为 *Apv1*。*Apv1* 编码 P450 亚区 CYP703A2-Zm 的一个成员, CYP703A2-Zm 含有 530 个氨基酸。

上述很多基因定位都用到了 BSA 这项技术, 以及相关的分子标记技术, 接下来简单介绍所用到的基因定位方法和分子标记技术。

3. BSA 技术

BSA 技术在基因定位的初步定位中广泛应用, 上述基因的定位中大部分都应用了 BSA 技术, 下面简要介绍最近大热的 BSA 技术

集团分离分析法(Bulk-segregant analysis, BSA)原理是: 在目的基因表型存在差异的亲本后代中选择性状分离的的植株, 构成两个亚组或群体; 将每个群体的 DNA 进行混合, 形成了两个性状相关的“基因池”; 利用分子标记对这两个基因池进行分析, 找到与目的基因性状连锁的分子标记; 利用得到的分子标记, 分析作图群体, 分子标记和目的基因的连锁得到了进一步的检测, 完成了标记在基因上的定位。特点是: 利用特定的分离群体; 只选择目的基因性状; 不存在环境和人为因素的影响; 省时省力[24]。

步骤为:

- 1) 分别提取两个群体的 DNA, 得到了两个分离群体的 DNA 样品, 按相同比例把各群体中各样本 DNA 混合;
- 2) 降低高群体基因组的复杂性, 以获得复杂性降低的 DNA 样本;
- 3) 以同等比例混合降低复杂性的分离群体 DNA, 采用高通量测序方法进行测序;
- 4) 对两组群体 DNA 样本的测序结果进行比较, 找到了序列标记, 然后根据序列标记的多样性差异, 检测出与性状相关的标记[25]。

4. 分子标记

4.1. SSR 标记

简单序列重复(Simple Sequence Repeats, SSR)即微卫星标记,又被称作简单序列重复标记。原理是:对重复序列两端特定短序列设计引物;进行 PCR 扩增;进行琼脂糖凝胶电泳;进行放射自显影;最终在其重复单位数不同的 DNA 区域中检测出多态性。特点是:共显性遗传;有简单的标记带型,记录的条带也比较一致且比较明确和客观;没有生物学方面的数量方面的限制;不需要很多 DNA 用量,且不需要很高的质量要求;每个位点都有很多种等位形式;多态性比较高;实验过程简单易操作。局限性是:建立、筛选基因组文库才能获得 SSR 标记,还需要测序、设计引物等,所需要实验过程多且复杂,费时费力[26]。

4.2. CAPS 标记

酶切扩增多态性(Cleaved Amplified Polymorphism Sequences, CAPS)技术又被称作 PCR-RFLP 法,综合了 PCR 技术和 RFLP 技术。原理是:设计特异性引物,利用 PCR 扩增某个基因或片段,利用特定的限制性内切酶切割扩增产物,利用凝胶电泳分离片段,然后进行 RFLP 分析,从而显示出扩增产物 DNA 片段的多态性。特异性 PCR 引物是根据已知位点的核苷酸序列而设计的,引物长度大概为 19~27 bp。特点是:共显性;DNA 用量少;所需引物较长;操作简单。局限性是:必须利用限制性内切酶筛选,费时费力;突变位点少[27]。

4.3. SNP 标记

单核苷酸多态性(Single Nucleotide Polymorphism, SNP)是指基因组中发生突变,该突变由单个核苷酸碱基的改变(包括单个碱基的缺失、插入、置换或颠换)引起的,从而引起基因组 DNA 序列的多态性。SNP 标记是第三代 DNA 分子标记。特点是:二等位基因性;位点比较多,分布比较广;遗传稳定性高;多态性比较高;检测快速,能够进行自动化分析。局限性是:成本高,标记少[28]。

4.4. InDel 标记

InDel 是指在亲缘关系较近的物种或同一物种的不同个体之间,在相同的基因位点的核酸序列中插入或缺失大小不同的 DNA 片段。InDel 是一种同源基因进行序列比对而产生空位的现象。原理是:设计特异性引物,进行 PCR 扩增,进行凝胶电泳。特异性引物的设计需要依据插入/缺失位点两侧的序列而设计,这样才能得到多态性标记。根据多态性标记的长度不同,利用凝胶电泳进行分型。特点是:InDel 标记准确性高和稳定性高;分布比较广,数量比较多;不会产生因为特异性和复杂性造成的分析歧义。局限性是:二等位基因性,没有很多的遗传信息量;对于基因组序列还没有完全测序的作物,设计开发引物比较难[29]。

近些年来,众多学者对玉米细胞核雄性不育基因进行定位及克隆,他们大多采用 BSA 技术进行初步定位,将不育基因定位在具体的染色体的某一区段,并筛选分子标记,再利用筛选出的分子标记对单株进行鉴定,对不育基因进行精细定位,而近些年来使用的分子标记大多是 SSR、CAPS、SNP、InDel。

5. 存在问题与展望

近些年来逐渐有玉米细胞核雄性不育基因被克隆出来,但是仍有大量的基因未被克隆出来,克隆出来的基因也很少得到应用,玉米细胞核雄性不育材料是一种非常好的材料,如果大量的得以应用,必然会对玉米的杂交制种有重要的意义。

本文总结了近些年已经克隆的细胞核雄性不育基因和相应的基因定位的方法等,以期望对之后的研

究有一定的借鉴意义, 希望更多的研究学者从事这方面的研究, 更多的基因被克隆, 而细胞核雄性不育也能更广泛的应用到实际生产当中, 这需要更多的学者前仆后继的努力。

参考文献

- [1] 陈晓阳. 玉米雄性不育基因 IPE1 克隆与功能分析[D]: [博士学位论文]. 北京: 中国农业大学, 2017.
- [2] 王超, 安学丽, 张增为, 等. 植物隐性核雄性不育基因育种技术体系的研究进展与展望[J]. 中国生物工程杂志, 2013, 33(10): 124-130.
- [3] 张丹凤. 玉米雄性不育基因 MS7 的图位克隆与应用[D]: [博士学位论文]. 长沙: 湖南农业大学, 2017.
- [4] 吕洪坤, 郑军. 图位克隆技术在玉米基因分离中的应用[J]. 分子植物育种, 2013(3): 460-468.
- [5] 柳双双, 吴锁伟, 饶力群, 等. 玉米核雄性不育的分子机制研究与应用分析[J]. 中国生物工程杂志, 2018, 38(1): 100-107.
- [6] 万向元, 吴锁伟, 周岩, 等. 植物花粉发育调控基因 *Ms1* 及其编码蛋白[P]. 中国专利, 201410381072.5.2. 2017-06-16.
- [7] Zhang, D., Wu, S., An, X., *et al.* (2017) Construction of a Multi-Control Sterility System for a Maize Male-Sterile Line and Hybrid Seed Production Based on the *ZmMs7*, Gene Encoding a PHD-Finger Transcription Factor. *Plant Biotechnology Journal*, **16**, 1-13. <https://doi.org/10.1111/pbi.12786>
- [8] Wang, D., Skibbe, D.S. and Walbot, V. (2013) Maize Male Sterile 8(Ms8), a Putative β -1,3-galactosyltransferase, Modulates Cell Division, Expansion, and Differentiation during Early Maize Anther Development. *Plant Reproduction*, **26**, 329-338. <https://doi.org/10.1007/s00497-013-0230-y>
- [9] Albertsen, M., Fox, T., Leonard, A., *et al.* (2016) Cloning and Use of the *ms9* Gene from Maize. US Patent 20150191743A1.
- [10] Albertsen, M.C., Fox, T., Trimnell, M., *et al.* (2009) *Msc1* Nucleotide Sequences Impacting Plant Male Fertility and Method of Using Same. US Patent 20090038027A1.
- [11] Nan, G.L., Zhai, J., Arikiti, S., *et al.* (2017) MS23, a Master Basic Helix-Loop-Helix Factor, Regulates the Specification and Development of the Tapetum in Maize. *Development*, **144**, 163-172. <https://doi.org/10.1242/dev.140673>
- [12] Djukanovic, V., Smith, J., Lowe, K., *et al.* (2013) Male-Sterile Maize Plants Produced by Targeted Mutagenesis of the Cytochrome P450-Like Gene (MS26) Using a Re-Designed I-CreI Homing Endonuclease. *Plant Journal*, **76**, 888-899. <https://doi.org/10.1111/tpj.12335>
- [13] An, X., Dong, Z., Tian, Y., *et al.* (2019) *ZmMs30* Encoding a Novel GDSL Lipase Is Essential for Male Fertility and Valuable for Hybrid Breeding in Maize. *Molecular Plant*, **12**, 343-359.
- [14] Moon, J., Skibbe, D., Timofejeva, L., *et al.* (2013) Regulation of Cell Divisions and Differentiation by MALE STERILITY32 Is Required for Anther Development in Maize. *Plant Journal*, **76**, 592-602. <https://doi.org/10.1111/tpj.12318>
- [15] 万向元, 吴锁伟, 张丹凤, 等. 玉米花粉发育调控基因 *Ms33* 的 DNA 序列及其编码蛋白[P]. 中国专利, 201610880590.0. 2017-02-15.
- [16] Cigan, A.M., Unger, E., Xu, R.J., *et al.* (2001) Phenotypic Complementation of *ms45* Maize Requires Tapetal Expression of MS45. *Sexual Plant Reproduction*, **14**, 135-142. <https://doi.org/10.1007/s004970100099>
- [17] Vernoud, V., Laigle, G., Rozier, F., *et al.* (2009) The HD-ZIP IV Transcription Factor OCL4 Is Necessary for Trichome Patterning and Anther Development in Maize. *The Plant Journal*, **59**, 883-894. <https://doi.org/10.1111/j.1365-3113X.2009.03916.x>
- [18] Wang, C.J.R., Nan, G.L., Kelliher, T., *et al.* (2012) Maize Multiple Archesporial Cells 1 (*mac1*), an Ortholog of Rice TDL1A, Modulates Cell Proliferation and Identity in Early Anther Development. *Development*, **139**, 2594-2603. <https://doi.org/10.1242/dev.077891>
- [19] Chen, X., Zhang, H., Sun, H., *et al.* (2017) Irregular Pollen Exine1 Is a Novel Factor in Anther Cuticle and Pollen Exine Formation. *Plant Physiology*, **173**, 307-325. <https://doi.org/10.1104/pp.16.00629>
- [20] Somaratne, Y., Tian, Y., Zhang, H., *et al.* (2017) Abnormal Pollen Vacuolation1 (APV1) Is Required for Male Fertility by Contributing to Anther Cuticle and Pollen Exine Formation in Maize. *Plant Journal*, **90**, 96-110. <https://doi.org/10.1111/tpj.13476>
- [21] Chaubal, R., Zanella, C., Trimnell, M.R., *et al.* (2000) Two Male-Sterile Mutants of *Zea mays* (Poaceae) with an Extra Cell Division in the Anther Wall. *American Journal of Botany*, **87**, 1193-1201. <https://doi.org/10.2307/2656657>
- [22] 吴锁伟, 方才臣, 邓联武, 等. 玉米隐性核雄性不育基因研究进展及其育种应用途径分析[J]. 分子植物育种网络

版, 2012, 10(1): 1001-1011.

- [23] Sheridan, W.F., Avalkina, N.A., Shamrov, I.I., *et al.* (1996) The Mac1 Gene: Controlling the Commitment to the Meiotic Pathway in Maize. *Genetics*, **142**, 1009-1020.
- [24] 白凤虎, 李德芳, 陈安国, 等. 基于 BSA 分析法的分子标记基因定位技术在农作物中的应用[J]. 中国麻业科学, 2006, 28(6): 282-288.
- [25] 王晓武, 郑洪坤. 基于测序和 BSA 技术的性状相关的分子标记筛选方法[P]. 中国专利, 21010225405.7.2. 2013-09-18.
- [26] 刘志雄. SSR 分子标记在玉米遗传育种中的应用[D]: [硕士学位论文]. 呼和浩特: 内蒙古农业大学, 2009.
- [27] 邢延豪, 周延清, 楚素霞, 等. CAPS 标记技术及其应用进展[J]. 江苏农业科学, 2011, 39(5): 74-76.
- [28] 库丽霞. 玉米株型相关性状分子遗传机理研究[D]: [博士学位论文]. 郑州: 河南农业大学, 2010.
- [29] 杨洁, 赫佳, 王丹碧, 等. InDel 标记的研究和应用进展[J]. 生物多样性, 2016, 24(2): 117-123.