

与烟草细胞质雄性不育相关的线粒体 基因 *atp9* 的 mRNA 研究

周 玮^{1,2*}, 刘齐元^{2,3*}, 陈雪峰², 刘飞虎², 曹 槐², 刘次全^{2,4**}

(1. 湖南农业大学 生物安全科学技术学院, 长沙 410128; 2. 云南大学 生命科学学院, 昆明 650091;
3. 江西农业大学 农学院, 南昌 330045; 4. 中国科学院 昆明动物研究所, 昆明 650223)

摘 要: 已知导入未编辑 *atp9* mRNA 的烟草表现细胞质雄性不育(CMS), 因此认为线粒体基因 *atp9* 是引起高等植物 CMS 的主要基因。为了解 *atp9* 在 CMS 中的作用机制, 从 3 对烟草不育系及其同型保持系中提取 *atp9*, 利用实验与理论结合来分析其 mRNA 在编辑前后以及在不育系及其同型保持系中的一维、三维信息差别。结果表明, *atp9* mRNA 一维信息方面的差异, 更重要的是二级结构的差异和稳定性, 可能是影响 ATP 合成而导致 CMS 的根本原因。

关键词: *atp9*; 细胞质雄性不育; mRNA 二级结构; 烟草

中图分类号: Q943 **文献标识码:** A **文章编号:** 1000-3142(2007)01-0114-07

Studies on *atp9* mRNA related to cytoplasmic male sterility in tobacco

ZHOU Wei^{1,2*}, LIU Qi-Yuan^{2,3*}, CHEN Xue-Feng²,
LIU Fei-Hu², CAO Huai², LIU Ci-Quan^{2,4**}

(1. College of Bio-safety Science and Technology, Hunan Agricultural University, Changsha 410128, China; 2. College of Life Sciences, Yunnan University, Kunming 650091, China; 3. College of Agronomy college, Jiangxi Agricultural University, Nanchang 330045, China; 4. Kunming Institute of Zoology, The Chinese Academy of Sciences, Kunming 650223, China)

Abstract: It was reported that the transgenic tobacco plants transformed with an unedited copy of the mitochondrial *atp9* gene exhibited the trait of the cytoplasmic male sterility (CMS) phenotype. Therefore, it was believed that *atp9* gene functioned as a dominant gene on CMS in higher plants. To understand the CMS mechanism caused by unedited form of *atp9* deeper, *atp9* gene was cloned from three sterile tobacco lines and their isogenic fertile lines. The differences in one- or three-dimensional information between unedited states and edited ones, or between sterility and fertility were analyzed. The results suggested that the differences in one-dimensional sequences, especially in secondary structures and their stabilities might possibly be the underlying reason resulting in CMS because of affecting ATP synthesis.

Key words: *atp9*; cytoplasmic male sterility; mRNA secondary structure; *Nicotiana tabacum*

收稿日期: 2005-04-11 修回日期: 2005-12-23

基金项目: 国家自然科学基金(30160036); 云南省农业生物技术重点实验室开放项目(2003E01) [Supported by the National Natural Science Foundation of China (30160036); Open Foundation of Yunnan Key Laboratory of Agricultural Biological Technology (2003E01)]

作者简介: 周玮(1977-), 女, 湖南永顺县人, 博士, 主要从事生物信息学、植物生理生化、遗传育种, (E-mail) mengzhou@163.com.

* 共同作者(Coauthor)

** 通讯作者(Author for correspondence, E-mail; liucq@ynu.edu.cn)

大多数高等植物是两性的,雄性不育普遍被认为是发育过程中的意外事件。植物雄性不育最常见的是细胞质雄性不育(CMS),它在雌花两性花异株(或雌全异株)种类的进化中起到积极作用(Budar 等,2001)。CMS在高等植物界普遍存在,系细胞质基因与细胞核基因相互作用的结果,由于线粒体功能紊乱导致花粉发育遭到破坏(凌杏元等,2000;He 等,1996;季静等,1998)。CMS烟草表现出雄蕊花丝变短、花粉囊枯皱等特征(Bergman 等,2000)。

有研究认为,可能是由于某个基因或多个基因产物干扰了线粒体 F(0)F(1)-ATP 合成酶的功能或合成,从而削弱和影响高耗能的花粉发育过程(Dieterich 等,2003)。植物线粒体内膜上的 ATP 合成酶在能量产物的形成过程中起重要作用,由 F(0)和 F(1)两部分构成。在 F(0)F(1)-ATP 复合体中,已知的 9 个亚基已全部被鉴定。F(0)部分是酶蛋白的膜内区域,具有质子跨膜传输功能,共有 4 个不同的亚基:orf25、orfB、亚基 9 和亚基 6;F(1)部分是酶蛋白的膜外区域,具有催化 ATP 合成的功能,包括 5 个不同的亚基: α 、 β 、 γ 、 δ 、 ϵ 。大量研究表明 orf25、orfB、亚基 9 和亚基 6 都可能与植物 CMS 有关(李大东等,1990;Sabar 等,2003;Heazlewood 等,2003)。

RNA 的编辑及 CMS 是与高等植物线粒体有关的两个重要现象(Howad 等,1999)。Hernould 等人的一系列研究使这两种现象紧密联系起来。他们的研究显示,将未编辑的线粒体 *atp9* 基因的转录本导入烟草后,花粉囊畸形导致花粉败育,并且由此引起的 CMS 性状能够稳定遗传给后代。而编辑过的 *atp9* 基因表达后没有出现 CMS 现象,因此认为未编辑的 *atp9* 基因的表达与 CMS 的形成有关。此外,进一步的研究还清楚地表明,通过反义 RNA 来抑制导入烟草体内未编辑的 *atp9* 基因的表达可以使烟草的育性恢复。从而认为未编辑的 *atp9* 基因可能是通过线粒体能量转换过程中的 ATP 合成酶亚基 9 作用而成为 CMS 的关键因子(Hernould 等,1993,1998;Zabaleta 等,1996)。

线粒体中 *atp9* mRNA 编辑与否同烟草 CMS 的关系已被证实。那么作为 CMS 相关基因的 *atp9*,它的 mRNA 在烟草不育系及其同型保持系中的差异是否与 *atp9* mRNA 编辑前后的差异相似,并成为导致 CMS 的关键因子引起了我们的极大兴趣。鉴于以上原因,我们试图在实验提取和一维信息对比的基础上,结合理论方法预测二级结构来分析不育系及其同

型保持系中 *atp9* 编码区的转录本在 mRNA 序列、3-核苷酸模式、氨基酸和二级结构方面的差异。通过比较分析 *atp9* mRNA 在编辑前后、不育系及其同型保持系的序列、尤其是空间结构等方面的差异,希望了解 *atp9* 作为导致烟草 CMS 关键因子的一维乃至三维信息,并探讨其发挥重要功能的分子机制。CMS 是一个复杂的生物学现象,且具有重要的经济价值,关于 CMS 的探究无论对这种生物现象本身意义重大,还是对涉及核质互作、线粒体半自主遗传系统等研究都具有重要的作用。

1 材料和方法

1.1 材料

(1)栽培普通烟草(*Nicotiana tabacum*)线粒体中未编辑的 *atp9* mRNA(*uned-atp9* mRNA)和编辑的 *atp9* mRNA(*ed-atp9* mRNA)(Hernould 等,1992);(2)烟草 CMS 系云烟 85 *atp9* mRNA(*MS85-atp9* mRNA)及其同型保持系云烟 85 *atp9* mRNA(*85-atp9* mRNA);(3)烟草 CMS 系 03 *atp9* mRNA(*MS03-atp9* mRNA)及其同型保持系 03 *atp9* mRNA(*03-atp9* mRNA);(4)烟草 CMS 系 31 *atp9* mRNA(*MS31-atp9* mRNA)及其同型保持系 31 *atp9* mRNA(*31-atp9* mRNA)。

1.2 方法

分离克隆材料(2)~(4)中的 *atp9* 基因,并进行 DNA 序列测定;根据 Watson-Crick 碱基配对原则,把测定的 DNA 编码区序列转录成对应的未编辑 mRNA 序列;并根据 *ed-atp9* mRNA 提供的编辑位点,对不育系及其同型保持系中未编辑 mRNA 序列进行编辑位点替换,从而推测出材料(2)~(4)中的 *ed-atp9* mRNA 序列。先用 DNASIS 方法比较 *atp9* 编辑前后、不育系及其同型保持系中 *ed-atp9* mRNA 的序列差异。再对各 mRNA 序列作 3-核苷酸模式的提取分析,根据密码子—氨基酸的指配关系找出对应于各密码子的氨基酸类型,比较 *atp9* 编辑前后、不育系及其同型保持系中 *ed-atp9* mRNA 的 3-核苷酸模式和编码的氨基酸差异。最后利用 RNAstructure version 4.1 程序预测各 mRNA 编码区序列的二级结构,从所获得的 mRNA 二级结构中选出具有最小自由能的二级结构,并用我们自己的 RNA studio 程序(Zhang 等,2002)调整成 2D 图像,然后统计出二级结构中的结构单元数,比较 *atp9* 编辑前后、不育系及其同型保持系

中 *ed-atp9* mRNA 的二级结构差异。

2 结果

2.1 *atp9* mRNA 在编辑前后、不育系及其同型保持系中序列的差异

测序结果表明(图 1), 3 对材料(85-*atp9* 和 MS85-*atp9*、03-*atp9* 和 MS03-*atp9*、31-*atp9* 和 MS31-*atp9*)的 DNA 编码区的核苷酸长度均为 234nt。不育系及其同型保持系中的 *atp9* 要经过编辑后才能继续它们的信息流传递, 因此按照 Hernould 等人提供的编辑位点, 把我们测序所得的 DNA 序列替换为编辑后的 mRNA 序列。由于 3 个保持系的

ed-atp9 mRNA 序列与 Hernould 等人测定的 *ed-atp9* mRNA 序列完全一致, 因此统一使用 *ed-atp9* mRNA 作为保持系参照。各条 mRNA 序列对比后的结果(图 1, 表 1)显示, *ed-atp9* 有 10 个编辑位点且全部为碱基 C(MS)→U 的替换; 值得注意的是, 由于 223 位碱基 C 被 U 替换, 造成 CAA→UAA, 从而使终止子提前出现, 导致编辑后的序列长度缩短为 225nt; 不育系及其同型保持系 mRNA 的序列差异主要表现为, MS85-*atp9* 与 85-*atp9* 间仅有 1 个位点发生改变, 即 126 位碱基由 A→U(MS); MS03-*atp9* 与 03-*atp9* 间有 2 个位点发生改变, 即 126 位碱基由 A→U(MS)与 215 位碱基由 U→C(MS); MS31-*atp9* 与 31-*atp9* 间有 3 个位点发生改变, 即 53 位碱基由 C→U(MS)、

	1						60
<i>ed-atp9</i>	AUGUUAGAAG	GUGCAAAAUU	AAUGGGUGCA	GGAGCUGCUA	CAAUUGCUIU	AGCGGGAGCU	
<i>uned-atp9</i>	AUGUUAGAAG	GUGCAAAAUU	AAUGGGUGCA	GGAGCUGCUA	CAAUUGCUIU	AGCGGGAGCU	
MS85- <i>atp9</i>	AUGUUAGAAG	GUGCAAAAUU	AAUGGGUGCA	GGAGCUGCUA	CAAUUGCUIU	AGCGGGAGCU	
MS03- <i>atp9</i>	AUGUUAGAAG	GUGCAAAAUU	AAUGGGUGCA	GGAGCUGCUA	CAAUUGCUIU	AGCGGGAGCU	
MS31- <i>atp9</i>	AUGUUAGAAG	GUGCAAAAUU	AAUGGGUGCA	GGAGCUGCUA	CAAUUGCUIU	AGUGGGAGCU	
	61						120
<i>ed-atp9</i>	GCUAUCGGUA	UUGGAAACGU	UUUUAGUUCU	UUGAUUCAUU	CCGUGGCGCG	AAAUCCAUCA	
<i>uned-atp9</i>	GCUAUCGGUA	UUGGAAACGU	UUUUAGUUCU	UUGAUUCAUU	CCGUGGCGCG	AAAUCCAUCA	
MS85- <i>atp9</i>	GCUAUCGGUA	UUGGAAACGU	UUUUAGUUCU	UUGAUUCAUU	CCGUGGCGCG	AAAUCCAUCA	
MS03- <i>atp9</i>	GCUAUCGGUA	UUGGAAACGU	UUUUAGUUCU	UUGAUUCAUU	CCGUGGCGCG	AAAUCCAUCA	
MS31- <i>atp9</i>	GCUAUCGGUA	UUGGAAACGU	UUUUAGUUCU	UUGAUUCAUU	CCGUGGCGCG	AAAUCCAUCA	
	121						180
<i>ed-atp9</i>	UUGGC AAAAC	AAUUAUUUGG	UUAUGCCA AU	UUGGGCUUUG	CUCUAACCGA	AGCUAUUGCA	
<i>uned-atp9</i>	UUGGC AAAAC	AAUUAUUUGG	UUAUGCCA AU	UUGGGCUUUG	CUCUAACCGA	AGCUAUUGCA	
MS85- <i>atp9</i>	UUGGC AAAAC	AAUUAUUUGG	UUAUGCCA AU	UUGGGCUUUG	CUCUAACCGA	AGCUAUUGCA	
MS03- <i>atp9</i>	UUGGC AAAAC	AAUUAUUUGG	UUAUGCCA AU	UUGGGCUUUG	CUCUAACCGA	AGCUAUUGCA	
MS31- <i>atp9</i>	UUGGC AAAAC	AAUUAUUUGG	UUAUGCCA AU	UUGGGCUUUG	CUCUAACCGA	AGCUAUUGCA	
	181						234
<i>ed-atp9</i>	UUGUUUGCCC	UAAUGAUGGC	CUUUUUGAUC	UUAUUCGU AU	UCUAAGUCCG	UUAG	
<i>uned-atp9</i>	UUGUUUGCCC	UAAUGAUGGC	CUUUUUGAUC	UUAUUCGU AU	UCUAAGUCCG	UUAG	
MS85- <i>atp9</i>	UUGUUUGCCC	UAAUGAUGGC	CUUUUUGAUC	UUAUUCGU AU	UCUAAGUCCG	UUAG	
MS03- <i>atp9</i>	UUGUUUGCCC	UAAUGAUGGC	CUUUUUGAUC	UUAUUCGU AU	UCUAAGUCCG	UUAG	
MS31- <i>atp9</i>	UUGUUUGCCC	UAAUGAUGGC	CUUUUUGAUC	UUAUUCGU AU	UCUAAGUCCG	UUAG	

图 1 烟草育性相关基因 *atp9* 编码区的 mRNA 序列在编辑前后及在烟草品系 MS85、85、MS03、03、MS31、31 中的比较

Fig. 1 Comparison of the coding regions of *atp9* mRNA between unedited states and edited ones, lines MS85 and 85, lines MS03 and 03, or lines MS31 and 31 in tobacco

123 位碱基由 G→U(MS)、158 位碱基由 U→A(MS)。
2.2 *atp9* mRNA 在编辑前后、不育系及其同型保持系中 3-核苷酸模式的差异

atp9 在编辑前后和不育系及其同型保持系中的 3-核苷酸模式使用频率上存在差异, 其差异统计结果如表 2 所示。由于存在 10 个编辑位点, *ed-atp9* 编辑

后的 3-核苷酸模式使用频率差异较显著, 分别体现在 CUU、UCG、CAA、CCA、GUC、GUU、UAA、UCA、UCC、UCU、CUA、UUA、UUG、UUU 共 14 种 3-核苷酸模式上; MS85-*atp9* 与 85-*atp9* 在 GCA 和 GCU 的 3-核苷酸模式使用频率存在差异; MS03-*atp9* 与 03-*atp9* 在 UCC、UUC、GCA 和 GCU 共 4 种 3-核苷酸模

式使用频率存在差异；MS31-*atp9* 与 31-*atp9* 在 GUG、UAU、GCG、UUG 和 UUU 共 5 种 3-核苷酸模式使用频率存在差异。其中 MS85-*atp9* 及其 85-*atp9*、MS03-*atp9* 及其 03-*atp9* 间有 2 种 3-核苷酸模式(GCA and GCU)的类型及数量差异完全相同。

表 1 烟草育性相关基因 *atp9* 的 mRNA 序列在编辑前后及在烟草品系 MS85、85、MS03、03、MS31、31 中的位点差异
Table 1 The differences of loci of *atp9* mRNA between unedited states and edited ones, lines MS85 and 85, lines MS03 and 03, or lines MS31 and 31 in tobacco

位点 Positions	<i>ed- atp9</i>	<i>uned- atp9</i>	MS85- <i>atp9</i>	MS03- <i>atp9</i>	MS31- <i>atp9</i>
20	U	<u>C</u>	U	U	U
50	U	<u>C</u>	U	U	U
53	C	<u>C</u>	C	C	<u>U</u>
81	U	<u>C</u>	U	U	<u>U</u>
82	U	<u>C</u>	U	U	U
90	U	<u>C</u>	U	U	U
92	U	<u>C</u>	U	U	U
123	G	<u>G</u>	G	G	<u>U</u>
126	A	<u>A</u>	<u>U</u>	<u>U</u>	A
158	U	U	<u>U</u>	<u>U</u>	<u>A</u>
182	U	<u>C</u>	U	U	<u>U</u>
191	U	<u>C</u>	U	U	U
212	U	<u>C</u>	U	U	U
215	U	<u>U</u>	U	<u>C</u>	U
223	U	<u>C</u>	U	<u>U</u>	U
编码区核苷酸长度 Nucleotide length of the coding region(nt)	225	234	225	225	225

斜体表示编辑位点, 正体表示差异位点。下划线+粗体表示与 *ed-atp9* 比较存在差异。下同。

The italic forms denoted editing sites; the normal forms denoted different sites; and the underlined and bold-faced forms denoted that there were differences in nucleotides comparing with *ed-atp9*. The same as follows.

2.3 *atp9* mRNA 在编辑前后、不育系及其同型保持系中编码氨基酸的差异

mRNA 所包含的核苷酸序列通过三联体密码子决定了蛋白质的氨基酸类型和排列顺序。*atp9* mRNA 在编辑前后、不育系及其同型保持系中所编码的氨基酸也存在着如表 3 所示的差异。与 3-核苷酸模式使用频率相似, *ed-atp9* mRNA 在编辑后所编码的氨基酸频率差异也较显著, 分别体现在 Gln、Pro、Arg、Val、Ser、Phe 和 Leu 这 7 种氨基酸上; MS85-*atp9* 与 85-*atp9* 的 mRNA 序列间仅有一个碱基改变, 且作为第三位密码子使 3-核苷酸模式使用频率相异, 但所编码的氨基酸完全相同; MS03-*atp9* 与 03-*atp9* 间在 Ser 和 Phe 2 种氨基酸的编码频率上存在差异; MS31-*atp9* 与 31-*atp9* 间在 Tyr、Val、Leu 和

Ala 4 种氨基酸的编码频率上存在差异。

2.4 *atp9* mRNA 在编辑前后、不育系及其同型保持系中二级结构的差异

RNA 二级结构是 Watson-Crick 配对区及其未配对区的二维表现, 一般认为由 6 种二级结构元素组成: 双链螺旋区、单链区、发夹、膨胀环、内环和分支环 (Batey 等, 1999; 刘次全等, 2000)。分别将 4 对材料的 mRNA 编码区序列折叠成二级结构, 如图 2 所示, *ed-atp9* mRNA 的二级结构在编辑后变化非常显著, 而 MS85-*atp9*、MS03-*atp9* 的结构变化部位正好位于被替换的 126 位碱基附近, MS31-*atp9* 的二级结构变化也非常显著。

表 2 烟草育性相关基因 *atp9* 的 mRNA 序列在编辑前后及在烟草品系 MS85、85、MS03、03、MS31、31 中的 3-核苷酸模式频率差异
Table 2 The differences of 3-nucleotide sequence modes of *atp9* mRNA between unedited states and edited ones, lines MS85 and 85, lines MS03 and 03, or lines MS31 and 31 in tobacco

3-核苷酸模式 3-nucleotide sequence modes	<i>ed- atp9</i>	<i>uned- atp9</i>	MS85- <i>atp9</i>	MS03- <i>atp9</i>	MS31- <i>atp9</i>
CUU	0	<u>1</u>	0	0	0
UCG	0	<u>2</u>	0	0	0
CAA	1	<u>2</u>	1	1	1
CCA	1	<u>2</u>	1	1	1
GUC	1	<u>2</u>	1	1	1
GUG	1	1	1	1	<u>2</u>
GUU	1	<u>0</u>	1	1	1
UAA	1	<u>0</u>	1	1	1
UAU	1	1	1	1	<u>2</u>
UCA	1	<u>4</u>	1	1	<u>1</u>
UCC	1	<u>2</u>	1	<u>2</u>	1
UCU	1	<u>0</u>	1	1	1
CUA	2	<u>1</u>	2	2	2
GCG	2	<u>2</u>	2	2	<u>1</u>
UUC	2	2	2	<u>1</u>	2
GCA	4	4	<u>3</u>	<u>3</u>	4
UUA	5	<u>2</u>	5	5	5
UUG	5	<u>3</u>	5	5	<u>4</u>
UUU	5	<u>4</u>	5	5	<u>5</u>
GCU	7	7	<u>8</u>	<u>8</u>	7

MS31-*atp9* mRNA 序列第 123 位核苷酸由 G→U, 导致 3-核苷酸模式由 UUG→UUU; 而第 158 位核苷酸由 U→A, 导致 3-核苷酸模式由 UUU→UAU。所以最后综合的统计结果显示 UUU 的使用频率未变, 这正是因为中间出现一增一减互补的结果。

For MS31-*atp9* mRNA, the transition at position 123(G→U) resulted in the change of 3-nucleotide sequence mode from UUG to UUU; Meanwhile, the transition at position 158(U→A) resulted in the change of 3-nucleotide sequence mode from UUU to UAU. Therefore, the finally result showed that the number of UUU did not vary.

4 对材料 *atp9* 的 mRNA 编码区二级结构单元, 分类统计结果如表 4 所示。未编辑时翻译区核苷酸数目相同, 但编辑后由于终止子提前产生, 编码区数目减少了 9 个核苷酸。导致不育的 *uned-atp9* 螺旋区有 138 个碱基、单链区有 96 个碱基, 而 MS85 和 MS03 中双链区碱基数目减少为 126 个, 单链区碱基数目增加到 99 个, MS31 螺旋区及单链区碱基数则与保持系一致。 *uned-atp9*、MS85-*atp9*、MS03-*atp9* 的发夹数与 *ed-atp9* 一致, MS31-*atp9* 则增加 1 个发夹。热力学研究表明含 4 或 5 个核苷酸的发夹环是最稳定的(刘次全等, 2000), *uned-atp9* 中减少 1 个 4 核苷酸发夹环(核苷酸组成为 5'-UUCU-3'), 而 MS31-*atp9* 中减少 2 个(核苷酸组成为 5'-UUCU-3' 及 5'-UUUAU-3'), 但增加 1 个 5 核苷酸发夹环(核苷酸组成为 5'-UUCAU-3')。相对于 *ed-atp9*, *uned-atp9* 增加 3 个膨胀环及 1 个对称内环; MS31-*atp9* 的

膨胀环数量不变, 增加 1 个对称内环的同时减少 1 个非对称内环; 其它 2 个不育系的膨胀环和内环数量均

表 3 烟草育性相关基因 *atp9* 的 mRNA 在编辑前后及在烟草品系 MS85、85、MS03、03、MS31、31 中编码的氨基酸差异

Table 3 The differences of amino acid of *atp9* mRNA between unedited states and edited ones, lines MS85 and 85, lines MS03 and 03, or lines MS31 and 31 in tobacco

Amino acid	<i>ed-atp9</i>	<i>uned-atp9</i>	MS85- <i>atp9</i>	MS03- <i>atp9</i>	MS31- <i>atp9</i>
Gln	0	<u>2</u>	0	0	0
Tyr	1	1	1	1	<u>2</u>
Pro	1	<u>2</u>	1	1	1
Arg	1	<u>2</u>	1	1	1
Val	3	<u>4</u>	3	3	<u>4</u>
Ser	4	<u>9</u>	4	<u>5</u>	4
Phe	7	<u>6</u>	7	<u>6</u>	7
Leu	12	<u>7</u>	12	12	<u>11</u>
Ala	16	16	16	16	<u>15</u>

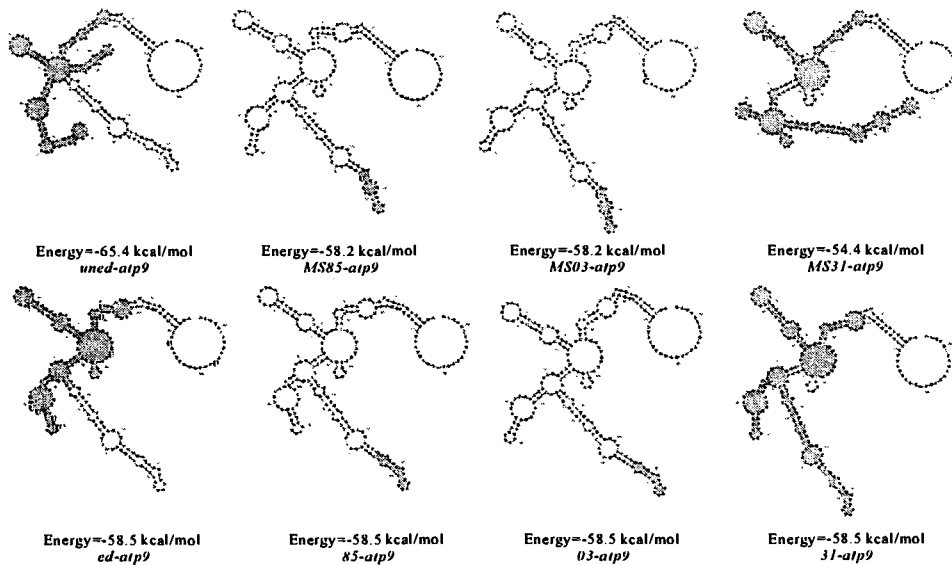


图 2 *uned-atp9* 和 *ed-atp9*、MS85-*atp9* 和 85-*atp9*、MS03-*atp9* 和 03-*atp9*、MS31-*atp9* 和 31-*atp9* 的 mRNA 二级结构。上排 4 个为与烟草 CMS 相关的 mRNA 二级结构, 下排对应为保持系的 mRNA 二级结构。相对于 *ed-atp9*, 与 CMS 相关的 mRNA 二级结构中改变的碱基被放大, 深色部分表示碱基改变导致的二级结构变化的区域。

Fig. 2 The secondary structures of mRNA of *uned-atp9*, *ed-atp9*, MS85-*atp9*, 85-*atp9*, MS03-*atp9*, 03-*atp9*, MS31-*atp9* and 31-*atp9*. Four upper secondary structures of mRNA denoted the secondary structures of mRNA related to sterility in tobacco, and four lower secondary structures of mRNA denoted the secondary structures of mRNA related to fertility in tobacco. Comparing with *ed-atp9*, the replaced bases were magnified, and the changed regions of secondary structures filled deep color.

一致。 *atp9* 具有 1 个三分支环和 1 个四分支环, MS85 与 MS03 与之一致, MS31 则具有 2 个四分支环, 而 *uned-atp9* 仅具 1 个五分支环。

如图 3 所示, 二级结构单元数量相对于各自序列核苷酸数目所占的百分比非常固定, 并且数量上的变化无规律, 但从二级结构上可以看出, 不育系与其保

持系的改变主要涉及三个复合发夹(保持系中的复合发夹 1, 端环碱基为 5'-CAAUUGCU-3'; 复合发夹 2, 端环碱基为 5'-UUCU-3'; 复合发夹 3, 端环碱基为 5'-UUUAU-3')的改变。

uned-atp9 mRNA 折叠结构的最小自由能相对于 *ed-atp9* 略低, 而另 3 个不育系都略高。

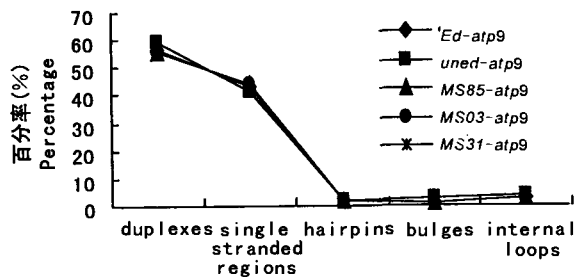


图 3 mRNA 二级结构单元的出现频率

Fig. 3 The percentage of the secondary structure motifs

3 讨论

已知有超过 150 个物种的 CMS 是由于线粒体基因异常编码从而阻止花粉发育 (Bentolila 等, 2002)。mRNA 在介导细胞信息从编码 DNA 的基因到功能蛋白的传递过程中起关键作用, 而 *atp9* mRNA 的编辑与否与烟草 CMS 密切相关 (Hernould 等, 1993)。

mRNA 序列的比较显示, *uned-atp9* mRNA 序列中的 10 个 C 替换为 U 后就成为 *ed-atp9* mRNA, 并且由于 223 位碱基替换使终止子提前出现, 这使得编辑前后的差异不仅体现在编辑位点上也体现在序列长度上。而 3 对不育系及其同型保持系材料的 *ed-atp9* mRNA 间的碱基差异就体现在与编辑位点不重复的 1-3 个位点发生碱基改变, 这些碱基改变也都涉及了 U 的替换或被替换, 其中不育系序列出现更多的 U 代替其同型保持系相同位点的其它碱基。U 的替换频率非常高, 这可能是与烟草 CMS 相关的最重要的可能一维信息。

由于编辑的缘故, *uned-atp9* mRNA 与 *ed-atp9* mRNA 间的 3-核苷酸模式使用频率差异较显著, 共计有 14 种; 另 3 对材料分别在 2 种、4 种、5 种 3-核苷酸模式使用频率存在差异。其中 *MS85-atp9* 及其 *85-atp9* 的 3-核苷酸模式使用差异在 *MS03-atp9* 及其 *03-atp9* 间都有出现, *uned-atp9* mRNA 与 *ed-atp9* mRNA、*MS31-atp9* 及其 *31-atp9* 间也有 2 种 3-核苷酸模式的差异完全相同, 但在 4 对材料中没有发现它们之间共同的差别。

差异显著的 3-核苷酸模式使用频率导致 *atp9* mRNA 编辑前后氨基酸的编码频率差异也较显著, 共体现在 7 种氨基酸上; *MS85-atp9* 与 *85-atp9* 的 mRNA 序列及 3-核苷酸模式使用频率相异, 但所编

码氨基酸完全相同; 另 2 对材料分别在 2 种、4 种氨基酸编码频率存在差异。虽然 *MS85-atp9* 及其 *85-atp9* 在 mRNA 序列及 3-核苷酸模式使用频率上都存在差异, 且它们所有的差异在 *MS03-atp9* 及其 *03-atp9* 中都存在, 但这种差异并没体现在氨基酸上。

表 4 烟草育性相关基因 *atp9* 的 mRNA 在编辑前后及在烟草品系 MS85、85、MS03、03、MS31、31 中二级结构单元差异

Table 4 The differences of secondary structure motifs of *atp9* mRNA between unedited states and edited ones, lines MS85 and 85, lines MS03 and 03, or lines MS31 and 31 in tobacco

二级结构单元 Motifs	<i>ed- atp9</i>	<i>uned- atp9</i>	<i>MS85- atp9</i>	<i>MS03- atp9</i>	<i>MS31- atp9</i>
双链螺旋区 Duplexes/nt	128	<u>138</u>	<u>126</u>	<u>126</u>	128
单链区 Single stranded regions/nt	97	<u>96</u>	<u>99</u>	<u>99</u>	97
发夹 Hairpins/Entries	4	4	4	4	<u>5</u>
4 碱基环发夹 Tetra-loops hairpins/Entries	3	<u>2</u>	3	3	<u>1</u>
5 碱基环发夹 Penta-loops hairpins/Entries	0	0	0	0	<u>1</u>
膨胀环 Bulges/Entries	3	<u>6</u>	3	3	3
内环 Internal loops/Entries	7	<u>8</u>	7	7	7
对称内环 Symmetric internal loops/Entries	4	<u>5</u>	4	4	<u>5</u>
非对称内环 Asymmetric internal loops/Entries	3	3	3	3	<u>2</u>
三分支环 Three-way junctions/Entries	1	<u>0</u>	1	1	<u>0</u>
四分支环 Four-way junctions/Entries	1	<u>0</u>	1	1	<u>2</u>
五分支环 Five-way junctions/Entries	0	<u>1</u>	0	0	0
最小自由能 Minimum free energy/kcal · mol ⁻¹	-58.5	<u>-65.4</u>	<u>-58.2</u>	<u>-58.2</u>	<u>-54.4</u>

根据以上 3 项的一维比较结果, 我们很难从中寻找和判断导致烟草 CMS 的原因所在。因此要了解 RNA 间的相似性或者核心问题, 简单的序列比较是不全面的, 因为 RNA 链还能够形成螺旋区、发夹、膨胀环等结构 (Kin 等, 2002)。如前所述, *atp9* mRNA 编辑前后的一维差异几乎没有出现在不育系及其同型保持系之间; 更有甚者, *MS85-atp9* 与 *85-atp9* 的 mRNA 序列及 3-核苷酸模式使用频率均不同, 但由于密码子的简并特性致使所编码的氨基酸完全相同, 核苷酸的变化并未引起氨基酸的改变。序列的改变不一定导致二级结构的改变, 但二级结构的改变肯定依赖于—维序列的改变。这些现象提醒我们, 仅仅考虑与序列相联系的线性信息是

不够的,mRNA 三维结构特征应该成为一个值得考虑的重要因素。毕竟 RNA 的生物功能多样性源于 RNA 是遗传信息的携带者,并且能够折叠成复杂的三维结构从而提供化学催化位点(Batey 等,1999)。

从 CMS 相关基因 *atp9* 的 mRNA 二级结构的比较结果看,都有不同程度的改变,其中 *uned-atp9* 和 *MS31-atp9* 变化非常大,这种结构上的显著差异也使得二级结构单元差异显著。相比之下的 *MS85-atp9* 与 *MS03-atp9* 变化较简单,在序列方面它们都存在 126 位 A→U 的碱基替换,这种变化由于密码子的简并性并没有引起它们一维信息的改变,但却部分地改变了 *atp9* mRNA 的二级结构。而 *MS03-atp9* mRNA 215 位 U→C 的碱基替换引起所编码氨基酸的改变但并不改变 mRNA 的二级结构。*uned-atp9* 和 *MS31-atp9* 序列中存在较多的碱基替换,所以引起的结构变化也非常明显,但导致终止子提前出现的碱基替换并没有参与 *atp9* mRNA 二级结构的形成。从二级结构的总体上看,保持系中的复合发夹 1、复合发夹 2 和复合发夹 3 的改变可能是导致保持系丧失育性最重要的原因。值得注意的是,*atp9* mRNA 二级结构的变化无论大小,其二级结构单元始终具有非常固定的比率,这可能是由于自然界的 RNA 需要保持相同的结构参数,这些参数又与总体平衡和结构稳定性有关(Zorn 等,2004)。

在能力学的描述上,不育系 *atp9* 二级结构的最小自由能都略高或略低于保持系,这就意味着不育系的 *atp9* mRNA 折叠结构稳定性发生了改变。

Zabaleta 等(1996)认为转基因烟草中的未编辑 *atp9* 的表达会引起线粒体功能紊乱,进而影响花粉囊,尤其是绒粘层细胞的正常发育,减少花粉的形成。根据 *atp9* 三维结构的理论分析结果,我们推测不育系线粒体功能紊乱的原因可能与 *atp9* mRNA 折叠结构(尤其是复合发夹 1-3)及其稳定性密切相关。

即使在芸苔、向日葵、油菜、玉米等植物中已经找到一些与 CMS 相关的基因或开放阅读框,但引起花粉败育的机制仍不清楚(Dieterich 等,2003; Sabar 等,2003; Heazlewood 等,2003; Landgren 等,1996; Bellaoui 等,1997; Gallagher 等,2002; 陈庆富等,2005)。高等植物线粒体 DNA 的突变、编辑或重排似乎都能导致 CMS(Gallagher 等,2002; Araya 等,1998; Handa 等,1992; Akagi 等,1994; Kempken

等,1998),但在在一维信息改变基础上的三维信息却一直被忽视。本文的研究结果清晰地表明,通过理论预测并分析 *atp9* mRNA 二级结构特征来寻找 CMS 的原因是一条值得探索的途径。

4 结论

未编辑的 *atp9* mRNA 能导致烟草不育已被证实,我们所分析的烟草不育系及其同型保持系中的 *atp9* 确实存在差异。我们分析后认为 *atp9* mRNA 在序列、3-核苷酸模式、氨基酸等方面的一维信息的差异,更重要的是在保持物种稳定基础上的二级结构差异及稳定,尤其是三个重要复合发夹的改变可能是影响线粒体 F(0)F(1)-ATP 合成酶正常功能的内在原因,从而导致 ATP 合成受影响,其结果是引起花粉败育,产生 CMS 性状。

参考文献:

- 刘次全,曹恩华,白春礼,等. 2000. 核酸结构多态性[M]. 北京: 高等教育出版社:131
- Akagi H, Sakamoto M, Shinjo C, et al. 1994. A unique sequence located downstream from the rice mitochondrial *atp6* may cause male sterility[J]. *Curr Genet*, 25(1):52-58
- Araya A, Zabaleta E, Blanc V, et al. 1998. RNA editing in plant mitochondria, cytoplasmic male sterility and plant breeding[J]. *Journal of Biotechnology*, 1:31-39
- Batey RT, Rambo RP, Doudna JA. 1999. Tertiary motifs in RNA structure and folding[J]. *Angew Chem Int Ed*, 38:2 326-2 343
- Bellaoui M, Pelletier G, Budar F. 1997. The steady-state level of mRNA from the Ogura cytoplasmic male sterility locus in *Brassica* hybrids is determined post-transcriptionally by its 3' region[J]. *EMBO J*, 16(16):5 057-5 068
- Bentolila S, Alfonso AA, Hanson MR. 2002. A pentatricopeptide repeat-containing gene restores fertility to cytoplasmic male-sterile plants[J]. *Proc Natl Acad Sci*, 9(16):10 887-10 892
- Bergman P, Edqvist J, Farbos I, et al. 2000. Male-sterile tobacco displays abnormal mitochondrial *atp1* transcript accumulation and reduced floral ATP/ADP ratio[J]. *Plant Mol Biol*, 42(3):531-544
- Budar F, Pelletier G. 2001. Male sterility in plants: occurrence, determinism, significance and use[J]. *C R Acad Sci III*, 324(6):543-550
- Chen QF(陈庆富), Qin Y(覃亚), Tan WF(谭武芳), et al. 2005. Comparative study of fertility restoring properties on various male-sterile lines of common wheat(几种小麦雄性不育系育性恢复性的比较研究)[J]. *Guihaia*(广西植物), 25(1):62-65
- Dieterich JH, Braun HP, Schmitz UK. 2003. Alloplasmic male sterility in *Brassica napus*(CMS'Tournefortii-Stiewe') is asso-
- (下转第 39 页 Continue on page 39)

- Light and electron microscopic observation on the developing pattern of cotyledons of *Nelumbo nucifera* (莲子叶发育过程的光学和电子显微镜观察)[J]. *Acta Bot Sin* (植物学报), 31(10):750-756
- Thorne RF. 2000. The classification and geography of the flowering plants; dicotyledons of the class angiospermae[J]. *Bot Review*, 66(4):442-624
- Tillich HJ. 1990. Die keimpflanzen der Nymphaeaceae-monocotyl oder dicotyl[J]. *Flora*, 184:169-176
- Titova GE, Batygina TB. 1996. Is the embryo of Nymphaealean plants (Nymphaeales s. l.) a dicotyledonous[J]. *Phytomorphology*, 46(2):171-190
- Wei PH (韦平和), Chen WP (陈维培), Chen RY (陈瑞阳). 1994. Study on the karyotype analysis of Nymphaeaceae and its taxonomic position (睡莲科的核型分析及其分类学位置的探讨)[J]. *Acta Phytotax Sin* (植物分类学报), 32(4):293-300
- Williamson PS, Moseley MF. 1989. Morphological studies of the Nymphaeaceae sensu lato. XVII. Floral anatomy of *Ondinea purpurea* subspecies *purpurea* (Nymphaeaceae)[J]. *Am J Bot*, 76(12):1779-1794
- Wu ZY (吴征镒), Lu AM (路安民), Tang YC (汤彦承), et al. 2003. The families and genera of angiosperms in China-A comprehensive analysis (中国被子植物科属综论)[M]. Beijing: Science Press:6-12, 36, 47-51
- Yan SZ (严素珍). 1986. The development of embryo and endosperm of *Nelumbo nucifera* Gaertn (莲胚与胚乳的发育)[J]. *Acta Bot Sin* (植物学报), 28(4):355-360
- Zanis MJ, Soltis DE, Soltis PS, et al. 2002. The root of the angiosperms revisited[J]. *Proceedings of the National Academy of Sciences of the United States of America*, 99(10):6848-6853

(上接第 120 页 Continue from page 120)

- ciated with a special gene arrangement around a novel *atp9* gene[J]. *Mol Genet Genomics*, 269(6):723-731
- Gallagher LJ, Betz SK, Chase CD. 2002. Mitochondrial RNA editing truncates a chimeric open reading frame associated with S male-sterility in maize[J]. *Curr Genet*, 42(3):179-184
- Handa H, Nakajima K. 1992. Different organization and altered transcription of the mitochondrial *atp6* gene in the male-sterile cytoplasm of rapeseed (*Brassica napus* L.) [J]. *Curr Genet*, 21(2):153-159
- He S, Abad AR, Gelvin SB, et al. 1996. A cytoplasmic male sterility-associated mitochondrial protein causes pollen disruption in transgenic tobacco[J]. *Proc Natl Acad Sci*, 93(21):11763-11768
- Heazlewood JL, Whelan J, Millar AH. 2003. The products of the mitochondrial *orf25* and *orfB* genes are FO components in the plant F1FO ATP synthase[J]. *FEBS Lett*, 540(1-3):201-205
- Hernould M, Mouras A, Litvak S, et al. 1992. RNA editing of the mitochondrial *atp9* transcript from tobacco[J]. *Nucleic Acids Res*, 20(7):1809
- Hernould M, Suharsono S, Litvak S, et al. 1993. Male-sterility induction in transgenic tobacco plants with an unedited *atp9* mitochondrial gene from wheat[J]. *Proc Natl Acad Sci*, 90(6):2370-2374
- Hernould M, Suharsono, Zabaleta E, et al. 1998. Impairment of tapetum and mitochondria in engineered male-sterile tobacco plants[J]. *Plant Mol Biol*, 36(4):499-508
- Howard W, Tang HV, Pring DR, et al. 1999. Nuclear genes from Tx CMS maintainer lines are unable to maintain *atp6* RNA editing in any anther cell-type in the sorghum bicolor A3 cytoplasm[J]. *Curr Genet*, 36(1-2):62-68
- Ji J (季静), Wang G (王罡), Series H, et al. 1998. Characterization of *atpA* loci of CMS mitochondrial genome in sunflower (向日葵细胞质雄性不育系线粒体基因组 *atpA* 位点的研究)[J]. *Acta Genet Sin* (遗传学报), 25(4):367-374
- Kempken F, Howard W, Pring DR. 1998. Mutations at specific *atp6* codons which cause human mitochondrial diseases also lead to male sterility in a plant[J]. *FEBS Lett*, 441(2):159-160
- Kin T, Tsuda K, Asai K. 2002. Marginalized kernels for RNA sequence data analysis[J]. *Genome Informatics*, 13:112-122
- Landgren M, Zetterstrand M, Sundberg E, et al. 1996. Alloplasmic male-sterile *Brassica* lines containing *B. tournefortii* mitochondria express an ORF 3' of the *atp6* gene and a 32 kDa protein[J]. *Plant Mol Biol*, 32(5):879-890
- Li DD (李大东), Wang B (王斌). 1990. Gene cloning and relationship with cytoplasmic male sterility (水稻线粒体 *atpA* 基因的克隆及其与细胞质雄性不育的关系)[J]. *Hereditas* (遗传), 12(4):1-4
- Ling XY (凌杏元), Zhou PJ (周培疆), Zhu YG (朱英国). 2000. Progress of the studies on molecular mechanism of cytoplasmic male sterility (植物细胞质雄性不育分子机理研究进展)[J]. *Chin Bull Bot* (植物学通报), 17(4):319-332
- Sabar M, Gagliardi D, Balk J, et al. 2003. ORFB is a subunit of F1F(O)-ATP synthase; insight into the basis of cytoplasmic male sterility in sunflower[J]. *EMBO Rep*, 4(4):381-386
- Zabaleta E, Mouras A, Hernould M, et al. 1996. Transgenic male-sterile plant induced by an unedited *atp9* gene is restored to fertility by inhibiting its expression with antisense RNA[J]. *Proc Natl Acad Sci*, 93:11259-11263
- Zhang H, Liu CQ. 2002. RNA Studio, a full-featured object-oriented program for visualizing RNA secondary structures [J]. *Journal of Molecular Graphics and Modelling*, 5308:1-2
- Zorn J, Gan HH, Shiffeldrim N, et al. 2004. Structural motifs in ribosomal RNAs; Implications for RNA design and genomics[J]. *Biopolymers*, 73:340-347