

研究报告 Original Papers

森林草莓全基因组MADS-box转录因子基因的鉴定及凤梨草莓果实 *FaMADS1* 基因克隆周厚成^{1,2,*}, 李刚¹, 赵霞¹, 康兆茹¹, 郭蔼光²¹中国农业科学院郑州果树研究所, 郑州450009; ²西北农林科技大学生命科学院, 陕西杨凌712100

摘要: 本文利用生物信息学方法对森林草莓(*Fragaria vesca*)基因组数据库中MADS-box基因的数量、结构类型、序列特征及染色体定位进行分析。结果表明, 获得70个含SRF-TF结构域的森林草莓*FvMADS*基因, DNA长度289~14 596 bp, 编码66~1 437个氨基酸残基, 有21个*FvMADS*没有内含子, 在7条染色体上呈不均匀分布; 68个*FvMADS*蛋白序列含有保守基序Motif 1, 最佳匹配序列为“RQVTF SKRRNGLLKKAYELSVLCDAEVALIIFSSRGKLYEF”。另外, 从栽培品种‘丰香’草莓果实中克隆了*FaMADS1*基因, 该基因属于MADS-box基因家族, cDNA全长1 167 bp, 编码区750 bp, 推导编码249个氨基酸, 具有MADS结构域和K-box结构域。

关键词: 草莓; MADS-box; 转录因子; 生物信息学; *FaMADS1*基因

Genome-Wide Sequence Identification of MADS-Box Transcription Factor Gene Family from *Fragaria vesca* and Cloning of *FaMADS1* gene from *F. ananassa* FruitZHOU Hou-Cheng^{1,2,*}, LI Gang¹, ZHAO Xia¹, KANG Zhao-Ru¹, GUO Ai-Guang²¹Zhengzhou Fruit Research Institute, Chinese Academy of Agricultural Sciences, Zhengzhou 450009, China; ²College of Life Sciences, Northwest A & F University, Yangling, Shaanxi 712100, China

Abstract: In this paper, MADS-box genes were obtained by *Fragaria vesca* genome database query and named *FvMADS*. MADS-box genes family, including the number of the *FvMADS* genes, structure type, sequence characteristic and chromosomal localization, was analyzed by using bioinformatics methods. The results showed that 70 *FvMADS* genes including SRF-TF structure domain were obtained. These *FvMADS* gene family DNA sequence were 289–14 596 bp in length and encoded 66–1 437 amino acid residues, and 21 *FvMADS* genes had no introns. These genes were not evenly distributed on 7 chromosomes. 68 *FvMADS* proteins contained a conserved base sequence Motif 1 of the best matching sequences for “RQVTF SKRRNGLLKKAYELSVLCDAEVALIIFSSRGKLYEF”. In addition, the *FaMADS1* gene from *F. ananassa* cv. ‘Toyonoka’ fruit was cloned. The gene which belonged to MADS-box gene family, the full-length cDNA of 1 167 bp encoding 249 amino acids, was derived by MADS and K-box structure domains.

Key words: *Fragaria*; MADS-box; transcription factor; bioinformatics; *FaMADS1* gene

植物MADS-box基因家族数量众多, 功能多样, 编码具有重要生物学功能的转录因子。植物特异MIKC型MADS-box转录因子具有MADS结构域(M区)以及I (intervening)区、K (keratin)区、C末端(C-terminal)等4个结构域。M区位于MADS-box蛋白的N末端, 是由大约60个氨基酸组成的高度保守结构, 具有与特异DNA结合和蛋白质二聚化等功能; I区是一个由大约30个氨基酸组成的序列保守性较低的区域, 作用是选择性帮助二聚体转录因子与DNA结合形成复合体; 中度保守的K区大约

有70个氨基酸, 具有3个 α 螺旋组成的coiled-coil结构, 参与蛋白质间的相互作用; C末端是在序列和长度上变化最大的区域, 在蛋白复合体的形成和转录激活中起重要作用(Smaczniak等2012)。MADS-box蛋白结构域在序列和结构上存在差异

收稿 2014-06-30 修定 2014-09-22

资助 河南省基础与前沿技术项目(132300410034)和公益性行业(农业)科研专项(201003064-7)。

* 通讯作者(E-mail: zhouhoucheng@caas.cn; Tel: 0371-65330972)。

与其功能作用密切相关。在植物中, MADS-box基因家族被认为是调节生殖生长和营养器官发育的关键基因家族, 广泛参与植物的根、叶、花和果实的发育以及果实的成熟进程(Schwarz-Sommer等1990; Becker和Theißen 2003; 刘菊华等2010; Smaczniak等2012)。近年来对MADS-box基因的研究主要集中于对花发育的调控, 对果实发育成熟的调节作用研究较少, 但MADS-box基因也参与果实的成熟调控(Vrevalov等2002; 刘菊华等2010)。目前已克隆出非乙烯调控途径调控果实成熟的MADS-box基因, 主要集中于番茄(Ito等2008; Vrevalov等2002)、葡萄和香蕉等跃变型果实(Poupin等2007; Liu等2009; 张远森等2013)。但MADS-box基因也参与调控草莓果实的成熟(Seymour等2011)。草莓是典型的非呼吸跃变型果实, 研究乙烯之外的调节因子对草莓果实的发育和成熟、延缓草莓果实成熟软化、增强果实耐贮运性能, 具有重要的理论和生产意义。

MADS-box基因作为一个庞大的基因家族, 已发现拟南芥中有107个MADS-box基因(Paenicová等2003), 水稻中有75个MADS-box基因(Arora等2007), 黄瓜中有43个MADS-box基因(Hu和Liu 2012), 大豆中有106个MADS-box基因(Shu等2013)等。森林草莓(*Fragaria vesca*, $2n=2x=14$)基因组测序已经完成(Shulaev等2011), 为开展栽培种凤梨草莓(*F. ananassa*, $2n=8x=56$)的全基因组基因挖掘和分子育种提供了重要的参考信息。本文利用生物信息学方法搜索森林草莓基因组数据库中MADS-box基因家族, 分析MADS-box基因的数量、结构类型、序列特征及染色体定位, 为MADS-box基因的功能鉴定特别是在果实发育成熟过程中的作用分析奠定基础。并从草莓栽培品种果实中克隆了MADS-box基因*FaMADS1*, 旨在为进一步研究*FaMADS1*调控草莓果实发育和成熟的功能和机理奠定基础。

材料与方 法

1 基因数据与材料

森林草莓(*Fragaria vesca* L.)基因组数据来自植物基因组数据库网站<http://www.phytozome.net/>。*FaMADS9*基因序列从NCBI获得(登录号

AF484683)。凤梨草莓品种‘丰香’(*F. ananassa* Duch. cv. ‘Toyonoka’)果实采自中国农业科学院郑州果树研究所草莓试验基地, 液氮速冻后-80 °C保存用于RNA提取。

2 森林草莓MADS-box基因的筛选

首先利用Pfam数据库(<http://pfam.sanger.ac.uk/>)搜索关键词“MADS”, 获得登录号PF00319文件, 该文件含长度为51个氨基酸残基SRF-TF结构的MADS-box蛋白保守结构域序列10个。选择SEP1_ARATH/9-59 (KRIENKINRQVT-FAKRRNGLLKKAYELSVLCDAEVALIIFSNRGK-LYEFCS)与森林草莓全基因组数据库(<http://www.phytozome.net/>)进行Blastp (E-value=0.001)序列比对, 获得MADS-box基因序列。

3 MADS-box基因结构分析、染色体定位和模体预测

森林草莓MADS-box基因的相关信息, 包括登录号、所在染色体及基因组位置由基因组数据库获得; 登陆GSDS网站(Gene Structure Display Server, <http://gsds.cbi.pku.edu.cn/>)分析MADS-box基因结构。森林草莓MADS-box蛋白的模体(Motif)分析采用模体分析工具The MEME Suite (Motif-based sequence analysis tools, <http://meme.nbcr.net/meme/intro.html>) MEME模块(Multiple Em for Motif Elicitation, <http://meme.nbcr.net/meme/cgi-bin/meme.cgi>), 参数设置如下: 同一Motif在1条序列中出现的次数为0或1, Motif最大发现数目4个, 基序长度范围最大100个氨基酸残基, 其他参数为默认值。

4 *FaMADS1*基因克隆

果实总RNA的提取采用改良CTAB法。RT-PCR法克隆MADS-box基因片段, 以纯化的mRNA为模板, Oligo (dT15)为引物, 按照MLV反转录酶(TaKaRa公司)说明书合成cDNA, 然后以*FaMADS9* (NCBI序列号AF484683)基因序列设计2对引物P1:P2和P3:P4进行PCR扩增(表1), 反应程序为: 94 °C 5 min, 94 °C 30 s, 55 °C 40 s, 72 °C 90 s, 35个循环, 72 °C延伸10 min。1%琼脂糖凝胶电泳检测, 回收片段进行T/A克隆测序。

使用SMARTer™ RACE cDNA amplification kit (BD Bioscience Clontech Company, USA)提供的引物3' CDS, 反转录合成3' RACE Ready cDNA。

表1 PCR引物序列

Table 1 Sequence of PCR primers

引物名称	序列(5'→3')
P1	ATGGGGAGGGGAAGAGTGGAG
P2	GCAITTCATATAGAAGTGACGAA
P3	GGTAGGCAGAGGAAATTGGTGTAT
P4	GCTGCTGCATGCTATCTCA
3' CDS	AAGCAGTGGTATCAACGCAGAGTAC(T)30VN
GSP1	CAGCAGCAGCCTAGCAGATGCTCTTAT
UPM	long-CTAATACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGT short-CTAATACGAC TCACTATAGGGC

据以上所得MADS-box基因片段序列设计3' RACE特异引物GSP1, 结合通用引物UPM进行PCR。反应程序为94 °C 5 min, 94 °C 30 s, 68 °C 30 s, 72 °C 3 min、35个循环, 72 °C延伸10 min。利用DNAMAN 6.0软件进行序列拼接分析。

实验结果

1 森林草莓MADS-box基因的序列特征、基因结构与染色体定位

利用MADS-box蛋白保守基序SEP1_ARATH/9-59, 采用Blastp程序搜索森林草莓基因组数据库, 获得70个含SRF-TF结构域的MADS-box基因, 分别命名为*FvMADS1*~*FvMADS70*。*FvMADS*基因DNA长度289~14 596 bp, 其中*FvMADS19*的序列最长, 为14 596 bp, 而*FvMADS11*的序列最短, 仅为289 bp。70个*FvMADS*蛋白中, 最短的只含有66个氨基酸残基(*FvMADS11*和*FvMADS50*), 最长的由1 437个氨基酸残基组成(*Fv-*

MADS48) (表2)。

利用GSDS在线工具获得MADS-box基因结构, 部分基因的内含子和外显子分布如图1所示。森林草莓MADS-box基因家族成员的外显子数量为1~25, *FvMADS51*的外显子数量最多, 为25个; *FvMADS19*次之, 为23个; 有21个*FvMADS*仅有1个外显子。

*FvMADS*在染色体的物理定位分析表明, 70个*FvMADS*在森林草莓7条染色体上呈不均匀分布, 除5个*FvMADS*基因未能定位外, 其余65个*FvMADS*基因分布在7条染色体上, 其中6个*FvMADS*基因位于1号染色体, 8个分布在2号染色体, 3号、4号染色体分别分布7个, 5、6号染色体分别分布14个, 7号染色体分布9个(图2)。

2 森林草莓MADS-box蛋白的基序分布及序列特征分析

利用在线MEME模体(Motif)分析工具对森林草莓MADS-box蛋白的保守性进行分析。结果表明, 70个森林草莓MADS-box蛋白中有4类Motif,

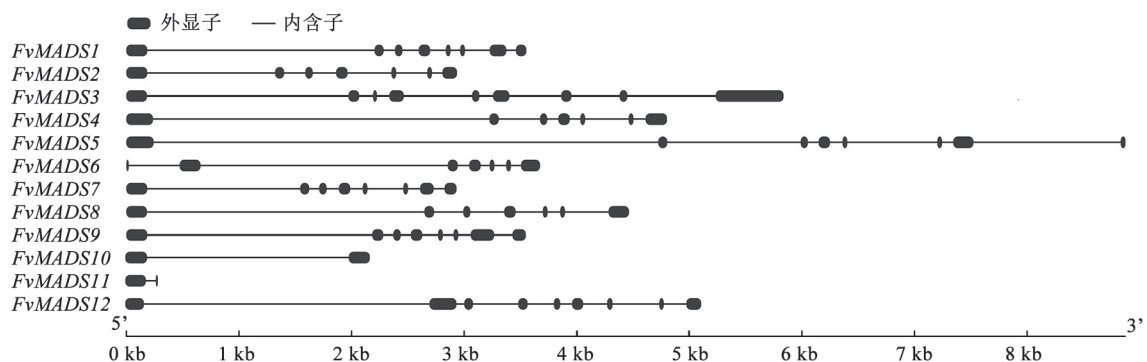


图1 森林草莓MADS-box部分基因家族成员的基因结构

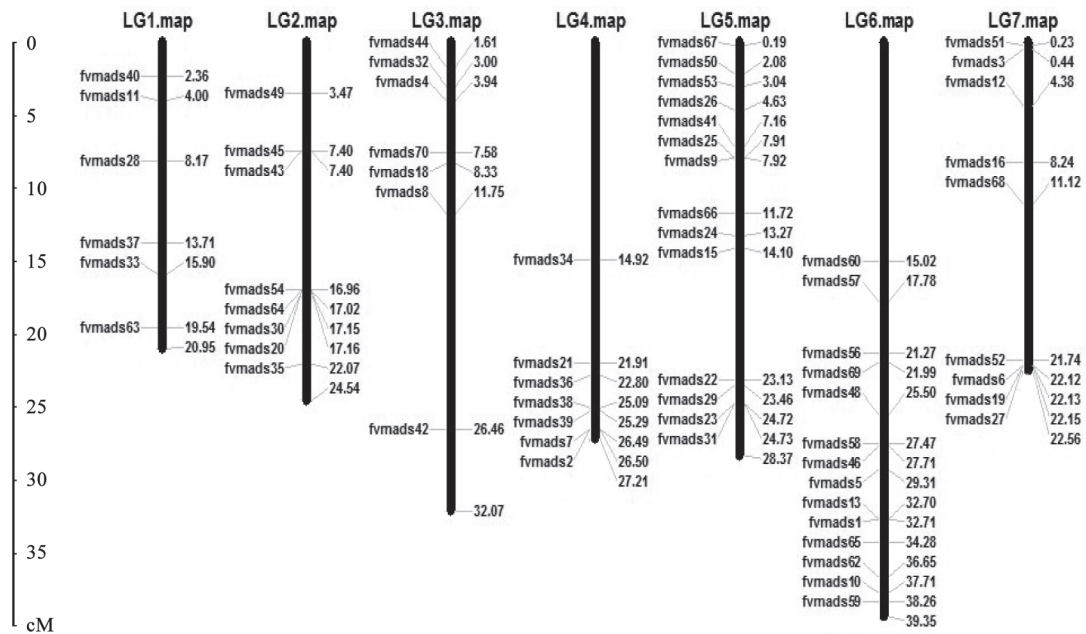
Fig.1 Gene structures of partial MADS-box gene family members of *F. vesca*

表2 森林草莓*FvMADS*基因的基本信息Table 2 Basic information of the *FvMADS* genes from *F. vesca*

基因名称	序列号	DNA长度/bp	CDS长度/bp	氨基酸数目	染色体定位	正(+)/负(-)链
<i>FvMADS1</i>	gene04229.1	3 550	750	249	LG6: 32713572~32717121	-
<i>FvMADS2</i>	gene04563.1	2 935	642	213	LG4: 26502083~26505017	-
<i>FvMADS3</i>	gene19428.1	5 838	1 398	465	LG7: 4394291~4400128	-
<i>FvMADS4</i>	gene24852.1	4 799	753	250	LG3: 3939971~3944769	+
<i>FvMADS5</i>	gene24494.1	8 874	789	262	LG6: 29313839~29322712	-
<i>FvMADS6</i>	gene13354.1	3 672	645	214	LG7: 22117084~22120755	+
<i>FvMADS7</i>	gene04562.1	2 931	738	245	LG4: 26497226~26500156	-
<i>FvMADS8</i>	gene25070.1	4 461	693	230	LG3: 11745926~11750386	-
<i>FvMADS9</i>	gene26119.1	3 546	855	284	LG5: 7921634~7925179	+
<i>FvMADS10</i>	gene01336.1	2 170	378	125	LG6: 37712160~37714329	-
<i>FvMADS11</i>	gene26429.1	289	201	66	LG1: 4034162~4034450	+
<i>FvMADS12</i>	gene19425.1	5 116	930	309	LG7: 4380104~4385219	+
<i>FvMADS13</i>	gene04228.1	4 698	774	257	LG6: 32701767~32706464	-
<i>FvMADS14</i>	gene25618.1	2 303	642	213	unanchored: 6827902~6830204	-
<i>FvMADS15</i>	gene31290.1	3 366	1 053	350	LG5: 14099242~14102607	-
<i>FvMADS16</i>	gene29503.1	3 918	930	309	LG7: 8235519~8239436	+
<i>FvMADS17</i>	gene20808.1	3 960	675	224	unanchored: 2475365~2479324	+
<i>FvMADS18</i>	gene20134.1	6 161	831	276	LG3: 8329328~8335488	+
<i>FvMADS19</i>	gene13356.1	14 596	2 790	929	LG7: 22131204~22145799	-
<i>FvMADS20</i>	gene11268.1	2 090	429	142	LG2: 17155740~17157829	+
<i>FvMADS21</i>	gene05027.1	3 832	825	274	LG4: 21916620~21920451	+
<i>FvMADS22</i>	gene27105.1	2 262	591	196	LG5: 23134576~23136837	-
<i>FvMADS23</i>	gene12119.1	10 797	708	235	LG5: 24719403~24730199	+
<i>FvMADS24</i>	gene22696.1	4 545	795	264	LG5: 13273715~13278259	+
<i>FvMADS25</i>	gene26118.1	6 371	630	209	LG5: 7910922~7917292	+
<i>FvMADS26</i>	gene06301	1 185	474	157	LG5: 4625966~4627150	-
<i>FvMADS27</i>	gene13358.1	1 836	468	155	LG7: 22152939~22154774	-
<i>FvMADS28</i>	gene14896.1	1 761	699	232	LG1: 8179391~8181151	-
<i>FvMADS29</i>	gene25669.1	678	687	225	LG5: 23457653~23458330	-
<i>FvMADS30</i>	gene11267.1	2 417	606	201	LG2: 17151913~17154329	+
<i>FvMADS31</i>	gene12120.1	7 012	708	235	LG5: 24734690~24741701	+
<i>FvMADS32</i>	gene30741.1	669	669	222	LG3: 2995375~2996043	-
<i>FvMADS33</i>	gene05564.1	612	612	203	LG1: 15877291~15877902	+
<i>FvMADS34</i>	gene00615.1	684	684	227	LG4: 14922036~14922719	+
<i>FvMADS35</i>	gene15090.1	675	675	224	LG2: 22071433~22072107	+
<i>FvMADS36</i>	gene05180.1	1 571	297	98	LG4: 22804817~22806387	+
<i>FvMADS37</i>	gene31594.1	847	504	197	LG1: 13712298~13713144	+
<i>FvMADS38</i>	gene03714.1	5 318	675	224	LG4: 25089184~25094501	-
<i>FvMADS39</i>	gene03748.1	5 360	675	224	LG4: 25288329~25293588	+
<i>FvMADS40</i>	gene31068.1	636	636	211	LG1: 2362822~2363457	-
<i>FvMADS41</i>	gene26028.1	7 302	2 103	700	LG5: 7161605~7168906	+
<i>FvMADS42</i>	gene14058.1	555	555	184	LG3: 26460131~26460685	-
<i>FvMADS43</i>	gene01788.1	696	696	231	LG2: 7400967~7401662	-
<i>FvMADS44</i>	gene30567.1	549	549	182	LG3: 1609369~1609917	+
<i>FvMADS45</i>	gene01789.1	663	663	220	LG2: 7405163~7405825	-
<i>FvMADS46</i>	gene05451.1	702	702	233	LG6: 27709858~27710559	+
<i>FvMADS47</i>	gene07364.1	738	738	245	unanchored: 7465583~7466320	-
<i>FvMADS48</i>	gene25753.1	10 062	4 434	1 437	LG6: 25500708~25510769	+
<i>FvMADS49</i>	gene24978.1	3 524	1 407	468	LG2: 3477474~3480997	-
<i>FvMADS50</i>	gene11073.1	319	201	66	LG5: 20793873~20794191	+

表2 (续)

基因名称	序列号	DNA长度/bp	CDS长度/bp	氨基酸数目	染色体定位	正(+)/负(-)链
<i>FvMADS1</i>	gene18483.1	9 245	2 592	863	LG7: 226718~235962	+
<i>FvMADS2</i>	gene13277.1	4 680	1 941	646	LG7: 21736485~21741164	-
<i>FvMADS3</i>	gene32155.1	675	675	224	LG5: 3040980~3041654	+
<i>FvMADS4</i>	gene11233.1	945	945	314	LG2: 16968321~16969265	-
<i>FvMADS5</i>	gene07365.1	6 393	1 425	474	unanchored: 7468076~7474468	-
<i>FvMADS6</i>	gene17829.1	2 056	978	325	LG6: 21269834~21271889	+
<i>FvMADS7</i>	gene18094.1	498	498	165	LG6: 17778486~17778983	-
<i>FvMADS8</i>	gene05411.1	4 168	1 587	528	LG6: 27471686~27475853	-
<i>FvMADS9</i>	gene25871.1	4 775	1 296	431	LG6: 38261461~38266235	-
<i>FvMADS60</i>	gene15563.1	597	597	198	LG6: 15023108~15023704	+
<i>FvMADS61</i>	gene07361.1	531	531	176	unanchored: 7431451~7431981	-
<i>FvMADS62</i>	gene28409.1	651	651	216	LG6: 36645777~36646427	+
<i>FvMADS63</i>	gene16491	498	498	165	LG1: 19537208~19537705	+
<i>FvMADS64</i>	gene11244.1	3 263	1 524	507	LG2: 17020653~17023915	+
<i>FvMADS65</i>	gene23924.1	2 692	1 356	451	LG6: 34281039~34283730	-
<i>FvMADS66</i>	gene10739.1	735	735	244	LG5: 11718000~11718734	-
<i>FvMADS67</i>	gene24650.1	453	453	150	LG5: 19474~19926	+
<i>FvMADS68</i>	gene04949.1	510	510	169	LG7: 11118553~11119062	+
<i>FvMADS69</i>	gene15899.1	6 411	891	296	LG6: 21989895~21996305	-
<i>FvMADS70</i>	gene20015.1	1 816	765	254	LG3: 7575627~7577442	+

图2 森林草莓*FvMADS*基因的物理定位Fig.2 Physical locations of *FvMADS* genes of *F. vesca*

其中有68个蛋白序列含有Motif 1, 仅*FvMADS36*和*FvMADS57*不含有Motif 1; 54个蛋白序列含有Motif 2, 33个蛋白序列含有Motif 3, 53个蛋白序列含有Motif 4 (图3)。进一步分析发现, Motif 1由41个

氨基酸残基组成, 最佳匹配序列为*FvMADS3* (RQVTFSKRRNGLLKKAYELSVLCDAE-VALIIFSSRGKLYEF); Motif 2由15个氨基酸残基组成, 最佳匹配序列为*FvMADS8* (MGRGKIE-

IKRIENTT), Motif 2与Motif 1中间仅隔1个氨基酸残基,二者串联组成57个氨基酸残基的MADS-box蛋白(MIKC类型) M区。在M区后紧接1个保守性较低的I区(Motif 4),最佳匹配序列为FvMADS31 (GHPCVESVIERFLTRNPPH), I区后是次级保守区K区,包含有Motif 3,最佳匹配序列为Fv-

MADS45 (GEDLEGLNIDELQKLEKQIEGGLSRVLQTK) (图3和4,表3)。从图4可以看出, FvMADS48是由2个MADS-box基因转录串联而成, FvMADS10、FvMADS11、FvMADS27、FvMADS50和FvMADS67等仅有M区而缺少K、I区等结构域。

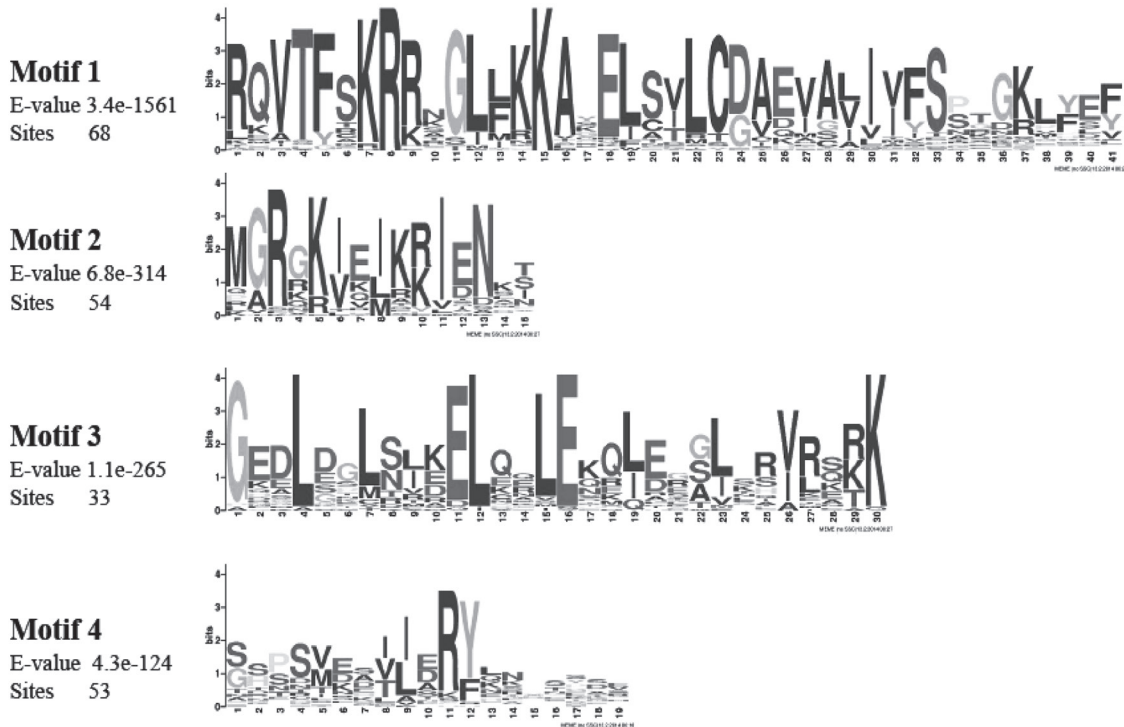


图3 森林草莓MADS-box转录因子的模体

Fig.3 Motifs of MADS-box transcription factors from *F. vesca*

表3 森林草莓MADS-box蛋白的模体组成

Table 3 Motif composition of MADS-box proteins from *F. vesca*

模体	基序氨基酸数目	最佳匹配
Motif 1	41	RQVTFsKRRNGLLKKAYELSVLCDAEVALIIFSSRGKLYEF
Motif 2	15	MGRGKIEIKRIENTT
Motif 3	30	GEDLEGLNIDELQKLEKQIEGGLSRVLQTK
Motif 4	19	GHPCVESVIERFLTRNPPH

3 *FaMADS1*基因序列分析

利用RT-PCR和RACE技术从‘丰香’草莓果实总RNA中克隆得到910 bp的中间片段和853 bp的5'端片段以及301 bp的3'端片段(图5),经拼接获得*FaMADS1*基因cDNA序列,NCBI登录号为HQ602763。*FaMADS1*基因cDNA全长1 167 bp,包

含1个750 bp的完整ORF、1个100 bp的5' UTR和1个287 bp的3' UTR及30个poly(A),ORF推导编码249个氨基酸。核苷酸序列相似性分析表明,*FaMADS1* (1~1 136位)与*FaMADS9* (279~1 410位)相似性为99%,与*FaMADS9*相比,*FaMADS1*在3' UTR的poly(A)之前缺少了82 bp的片段。

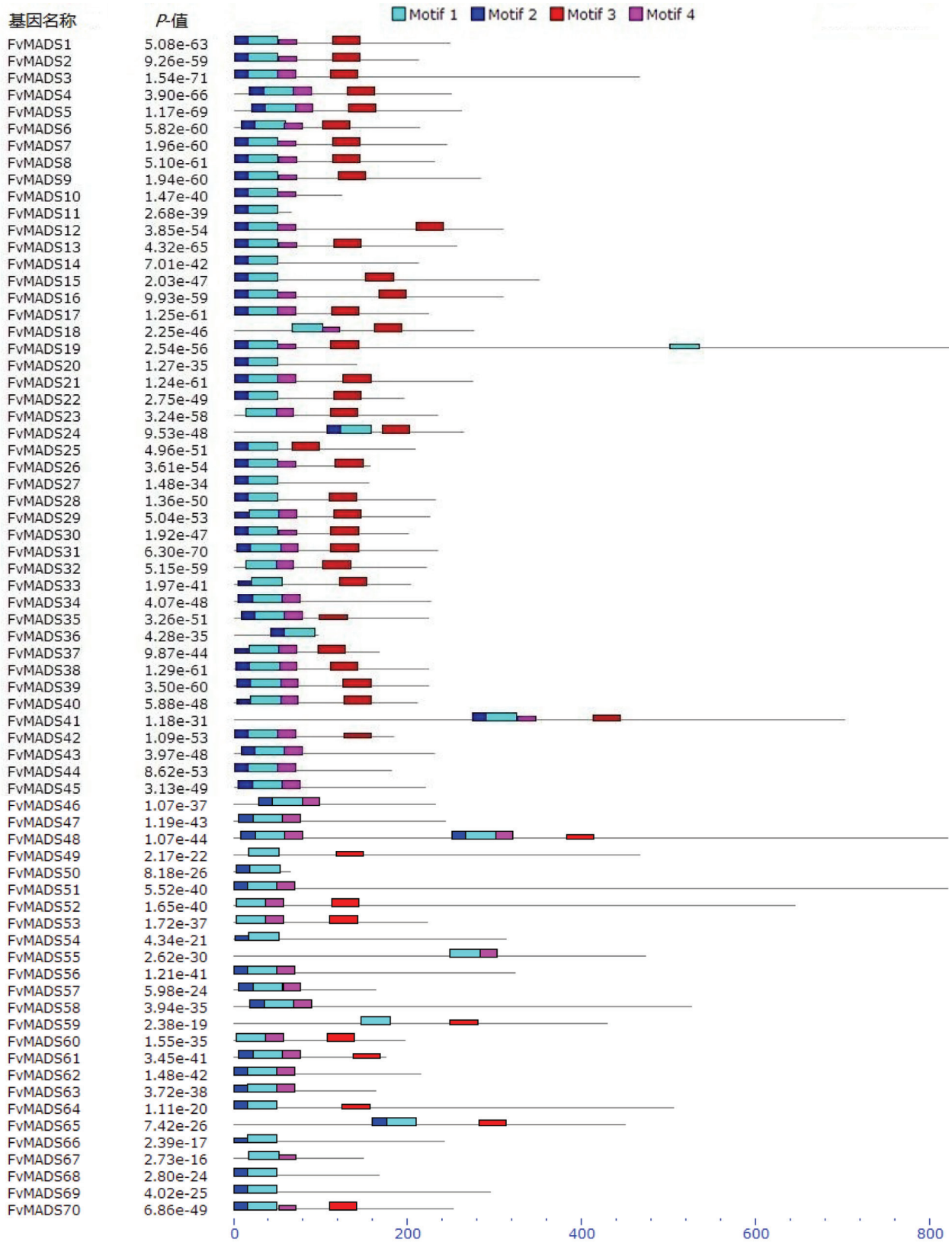
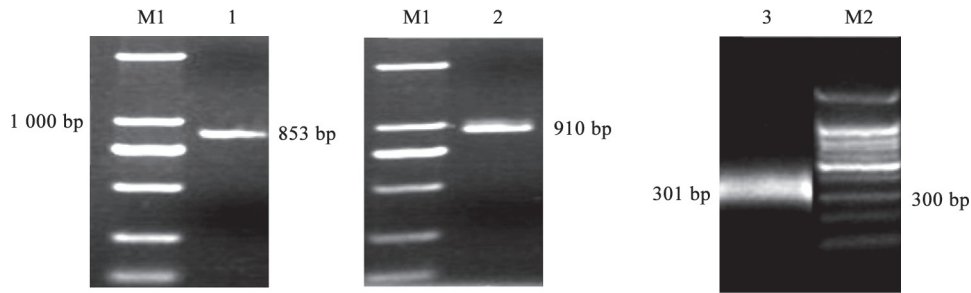


图4 森林草莓MADS-box蛋白的模体分布

Fig.4 Distribution of conserved motifs of MADS-box proteins of *F. vesca*

图5 *FaMADS1* cDNA扩增片段Fig.5 cDNA fragments amplification of *FaMADS1* gene

1: 5'端RT-PCR片段; 2: RT-PCR中间片段; 3: 3' RACE片段; M1: DL2000 DNA marker; M2: 100 bp ladder marker.

经森林草莓全基因组搜索, *FaMADS1*与森林草莓MADS-box转录因子基因*FvMADS1*匹配, 二者为同源基因, 这表明*FaMADS1*可能为*FvMADS1*经多次复制形成的。*FvMADS1*位于第6条染色体上(32713572~32717121), DNA全长3 550 bp, 由8个外显子和7个内含子组成, 编码区750 bp, 编码249个氨基酸残基。

*FaMADS1*推导的氨基酸结构域分析表明, 该序列具有MADS结构域(MADS_MEF2_like)和K-box结构域。在MADS结构域中有DNA结合位点、二聚体相互作用位点和翻译后修饰的磷酸化位点, K-box结构域为卷曲螺旋结构(coiled-coil), 参与形成多聚体。这些结构特征表明*FaMADS1*基因属于MADS-box转录因子基因家族。

讨 论

1 森林草莓全基因组*FvMADS*基因分析

利用森林草莓基因组序列, 搜索MADS-box蛋白的保守结构域(SRF-TF结构), 获得70条MADS-box基因, 表明*FvMADS*基因为一庞大基因家族。*FvMADS*基因推导编码的*FvMADS*蛋白由66~1 437个氨基酸残基组成(表2), 基因结构及蛋白结构变化很大, 表明MADS-box基因可能具有的功能多样性。

根据基因结构、蛋白结构和系统进化关系的不同, 植物MADS-box基因分为I型(type I)和II型(type II)两大类型。植物I型MADS-box基因通常含有1~2个外显子, 编码蛋白含有1个高度保守的SRF-like MADS结构域, 但缺少K结构域。经分析, 70个*FvMADS*中有34个*FvMADS*缺少K域, 其中有16个有1个外显子, 5个有2个外显子, 其它在3~25

个之间。植物II型MADS-box基因, 也叫MIKC型基因, 一般包含6个内含子和7个外显子。但*FvMADS*基因结构较为复杂, 有7个外显子的*FvMADS*仅有8个, 其它在1~23个之间。MIKC型基因具有MADS-box基因共有的MADS-box结构域(M区)以及I区、K区、C末端等4个结构域。K区是MIKC型MADS-box基因的特征序列, *FvMADS*中有36个具有K区, 应归属于MIKC型。MIKC型基因的这4个结构域在序列和结构上存在差异与其功能作用密切相关, 由此编码多种具有重要生物学功能的转录调节因子。

2 草莓果实*FaMADS1*基因分析

MADS-box转录因子的作用几乎涉及到了植物生长发育的各个方面, 其中对果实发育成熟的调控功能是近年来MADS-box基因研究的一个热点领域。研究发现, 番茄中RIN是MADS-box家族的重要成员, *rin*突变体中有2个串联的MADS-box基因*LeMADS-RIN*与*LeMADS-MC*, 由于*LeMADS-RIN*最后一个外显子缺失, 与*LeMADS-MC*发生融合, 导致RIN蛋白功能丧失, 果实不能成熟(Vrebalov等2002)。Moore等(2002)进一步确定RIN转录因子对果实成熟具有重要的调控作用, RIN蛋白只在成熟的果实中积累(Ito等2008)。RIN与MADS-box基因一样, 其功能具有多效性。RIN作为转录因子, 其直接调控的下游基因主要包括其它转录因子、乙烯合成及信号转导组分、细胞壁代谢、类胡萝卜素代谢等多条途径中的关键酶或相关组分的基因(Fujisawa等2011; Li等2011; Martel等2011)以及直接调控果实成熟过程中芳香物质的合成(Qin等2012)。

Vrebalov等(2002)从草莓果实中分离出了同

源基因*FaMADS9*, 该基因呈现果实专一的表达模式。Seymour等(2011)进一步采用基因沉默法使草莓*SEP1/2-like (FaMADS9)*基因沉默, 结果表明, 草莓花瓣、种子和果实等发育成熟受阻, 证明了SEP为果实成熟的调控因子, 同时调控跃变型果实和非跃变型果实的成熟。我们利用RT-PCR和RACE技术从软质草莓品种‘丰香’果实中克隆了*FaMADS1*基因, 获得了其cDNA全长序列。*FaMADS1*具有MADS结构域和K-box结构域, 属于MADS-box转录因子基因家族。与*FaMADS9*相比, *FaMADS1*在3' UTR的poly(A)之前缺少了82 bp的片段。真核生物mRNA的3' UTR在基因表达的转录后调控中起着重要作用, *FaMADS1*基因这种缺失导致的功能变化有待进一步分析。*FaMADS1*基因的获得为进一步分析*FaMADS1*调控草莓果实发育和成熟的功能机理奠定基础。

参考文献

- 刘菊华, 徐碧玉, 张静, 金志强(2010). MADS-box转录因子的相互作用及对果实发育和成熟的调控. 遗传, 32 (9): 1~10
- 张远森, 秦晓萌, 严金平, 张昆林, 杨恩(2013). 香蕉MADS-Box基因家族的生物信息学分析. 植物生理学报, 49 (10): 1030~1040
- Arora R, Agarwal P, Ray S, Singh AK, Singh VP, Tyagi AK, Kapoor S (2007). MADS-box gene family in rice: genome-wide identification, organization and expression profiling during reproductive development and stress. BMC Genomics, 8: 242
- Becker A, Theißen G (2003). The major clades of MADS-box genes and their role in the development and evolution of flowering plants. Mol Phylogenet Evol, 29: 464~489
- Fujisawa M, Nakano T, Ito Y (2011). Identification of potential target genes for the tomato fruit-ripening regulator RIN by chromatin immunoprecipitation. BMC Plant Biol, 11: 26
- Hu L, Liu S (2012). Genome-wide analysis of the MADS-box gene family in cucumber. Genome, 55 (3): 245~256
- Ito Y, Kitagawa M, Ihashi N, Yabe K, Kimbara J, Yasuda J, Ito H, Inakuma T, Hiroi S, Kasumi T (2008). DNA-binding specificity, transcriptional activation potential, and the *rin* mutation effect for the tomato fruit-ripening regulator RIN. Plant J, 55 (2): 212~223
- Li L, Zhu B, Yang P, Fu D, Zhu Y, Luo Y (2011). The regulation mode of RIN transcription factor involved in ethylene biosynthesis in tomato fruit. J Sci Food Agric, 91: 1822~1828
- Liu JH, Xu BY, Hu L, Su W, Wu J, Li M, Yang J, Jin Z (2009). Involvement of a banana MADS-box transcription factor gene in ethylene-induced fruit ripening. Plant Cell Rep, 28 (1): 103~111
- Martel C, Vrebalov J, Tafelmeyer P, Giovannoni JJ (2011). The tomato MADS-box transcription factor RIPENING INHIBITOR interacts with promoters involved in numerous ripening processes in a COLORLESS NONRIPENING-dependent manner. Plant Physiol, 157: 1568~1579
- Moore S, Vrebalov J, Payton P, Giovannoni J (2002). Use of genomics tools to isolate key ripening genes and analyse fruit maturation in tomato. J Exp Bot, 53 (377): 2023~2030
- Paenicová L, de Folter S, Kieffer M, Horner DS, Favalli C, Busscher J, Cook HE, Ingram RM, Kater MM, Davies B et al (2003). Molecular and phylogenetic analyses of the complete MADS-box transcription factor family in *Arabidopsis*: new openings to the MADS world. Plant Cell, 15 (7): 1538~1551
- Pimentel P, Salvatierra A, Moya-León MA, Herrera R (2010). Isolation of genes differentially expressed during development and ripening of *Fragaria chiloensis* fruit by suppression subtractive hybridization. J Plant Physiol, 167: 1179~1187
- Poupin MJ, Federici F, Medina C, Matus JT, Timmermann T, Arce-Johnson P (2007). Isolation of the three grape sub-lineages of B-class MADS-box *TM6*, *PISTILLATA* and *APETALA3* genes which are differentially expressed during flower and fruit development. Gene, 404: 10~24
- Qin GZ, Wang YY, Cao BH, Wang WH, Tian SP (2012). Unraveling the regulatory network of the MADS box transcription factor RIN in fruit ripening. Plant J, 70: 243~255
- Schwarz-Sommer Z, Huijser P, Nacken W, Saedler H, Sommer H (1990). Genetic control of flower development by homotic genes in *Antirrhinum majus*. Science, 250: 931~936
- Seymour GB, Ryder CD, Cevik V, Hammond JP, Popovich A, King GJ, Vrebalov J, Giovannoni JJ, Manning K (2011). A *SEPAL-LATA* gene is involved in the development and ripening of strawberry (*Fragaria × ananassa* Duch.) fruit, a non-climacteric tissue. J Exp Bot, 62 (3): 1179~1188
- Shu Y, Yu D, Wang D, Guo D, Guo C (2013). Genome-wide survey and expression analysis of the MADS-box gene family in soybean. Mol Biol Rep, 40 (6): 3901~3911
- Shulaev V, Sargent DJ, Crowhurst RN, Mockler TC, Folkerts O, Delcher AL, Jaiswal P, Mockaitis K, Liston A, Mane SP et al (2011). The genome of woodland strawberry (*Fragaria vesca*). Nat Genet, 43: 109~116
- Smaczniak C, Immink RG, Angenent GC, Kaufmann K (2012). Developmental and evolutionary diversity of plant MADS-domain factors: insights from recent studies. Development, 139 (17): 3081~3098
- Vrebalov J, Ruezinsky D, Padmanabhan V, White R, Medrano D, Drake R, Schuch W, Giovannoni JJ (2002). A MADS-box gene necessary for fruit ripening at the tomato *ripening-inhibitor (rin)* locus. Science, 296: 343~346