

森林草莓 MYB 家族鉴定及生物信息学分析

袁华招, 于红梅, 夏瑾, 庞夫花, 陈晓东, 赵密珍

(江苏省农业科学院果树研究所 / 江苏省高效园艺作物遗传改良重点实验室, 南京 210014)

摘要: 为研究草莓 MYB 基因家族, 采用生物信息学方法检索森林草莓 (*Fragaria vesca* L.) 基因组中 MYB 基因, 并对其基因结构、蛋白保守结构域、染色体定位、基因进化及花药发育过程中表达水平进行分析。结果表明, 草莓基因组中可能含有 105 个 R2R3-MYB、4 个 MYB3R、1 个 MYB4R 基因, 它们不均匀分布在 7 条草莓染色体上, 其中 5 号染色体上数量最多为 26 个, 部分基因在染色体上形成基因簇。理化分析发现这些 MYB 蛋白长度和大小差异较大, 但均具有保守的 R2 和 R3 型结构域, R 重复单元均含有特征性的氨基酸, 包括最保守的有序间隔出现的色氨酸 (W) 残基。通过系统进化关系和基因结构信息, 草莓 MYB 基因家族被分成了 34 个亚组 (A1~A34), 同时利用拟南芥基因功能预测草莓中对应的直系同源基因的功能, 结果表明草莓 MYB 基因家族参与调控黄酮物质合成、细胞发育、生物/非生物胁迫。花药发育过程中表达水平分析发现 FvMYB2、FvMYB85、FvMYB74、FvMYB28 在花药发育的 5 个时期中表达量逐渐升高, 尤其是发育后期表达量较高, 表明这些基因可能调控花药发育过程。本研究结果为草莓 MYB 基因家族的功能研究提供参考。

关键词: 草莓; MYB; 花药发育

Genome-wide Identification and Bioinformatics Analysis of MYB Gene Families in *Fragaria vesca* L.

YUAN Hua-zhao, YU Hong-mei, XIA Jin, PANG Fu-hua, CHEN Xiao-dong, ZHAO Mi-zhen

(Institute of Pomology, Jiangsu Academy of Agricultural Sciences/Jiangsu Key Laboratory for Horticultural Crop Genetic Improvement, Nanjing 210014)

Abstract: To get insight of MYB genes in strawberry, we used bioinformatics methods to screen the MYB genes from the *Fragaria vesca* L. genome database, and further analyzed gene structure, conserved motif, chromosomal location, phylogenetic analysis and expression in the period of the development of anther. Our results showed that three types of MYB transcription factors, including 105 R2R3-MYB, 4 MYB3R, 1 MYB4R, were identified. These genes unevenly distributed on 7 chromosomes, e.g., that chromosome Fvb5 contained the largest number of 26 MYB genes. Some of MYBs formed distinct tandem duplicate gene clusters. Although the protein length of strawberry MYB genes differed, all of MYBs contained conserved R2R3 domain, with characteristic amino acids, especially orderly spaced conservative tryptophan residues (W). Phylogenetic and gene structure analysis showed that strawberry MYB genes were divided into 34 distinct groups (A1-A34). The function prediction of strawberry orthologous genes using *Arabidopsis thaliana* L. gene annotation showed strawberry MYB genes were most likely involved in flavone synthesis, cell fate, biotic and abiotic stresses. Temporal or spatial expression trends of MYB genes during strawberry anther development indicated that FvMYB2, FvMYB85, FvMYB74, and FvMYB28 had increased expression during 5 period of the development of strawberry anther, which were mainly expressed in the

收稿日期: 2018-09-27 修回日期: 2018-10-18 网络出版日期: 2018-11-29

URL: <http://kns.cnki.net/kcms/detail/11.4996.S.20181128.1028.002.html>

第一作者研究方向为果树学, E-mail: 544250348@qq.com

通信作者: 赵密珍, 研究方向为果树学, E-mail: njzhaomz@163.com

基金项目: 江苏省农业重大新品种创制项目 (PZCZ201721); 农业部作物种质资源保护项目 (111821301354051014)

Foundation project: Project for Agricultural Significant New Varieties Breeding of Jiangsu Province (PZCZ201721), National Crop Germplasm Resources Protection of Ministry of Agriculture of China (111821301354051014)

later period of anther development. These results suggested that those *MYB* genes may play important role during strawberry anther development. Thus, this paper provided information for functional identification of strawberry *MYB* genes.

Key words: strawberry; *MYB*; anther development

转录因子通常至少由一个 DNA 结合域, 转录激活域, 核定位信号以及蛋白质-蛋白质结合域组成^[1]。许多重要的生物进程是由转录因子在转录水平调控基因表达控制的^[2]。*MYB* 被认为是植物中最大的一个转录因子家族, 它参与调控许多生物进程, 如生物和非生物胁迫、发育、分化、代谢和防御等^[3]。*MYB* 蛋白均有非常保守的 DNA 结合结构域 (*MYB* 结构域)。*MYB* 结构域由 1~4 个 R 结构域组成。每个 R 结构域通常由 52 个氨基酸残基组成, 含有 3 个 α 螺旋结构, 第 2 和第 3 个 α 螺旋通过 3 个均匀分布的色氨酸可以形成一个 HTH 螺旋结构 (helix-turn-helix)^[4]。R 结构域根据序列同源性被分成 3 种类型: R1、R2 和 R3, 同时 *MYB* 基因家族可以通过 R 结构域的重复次数分成 4 个亚组: 1R-*MYB*、R2R3-*MYB*、3R-*MYB* 和 4R-*MYB*^[3,5,6]。4 种类型的 *MYB* 基因在植物中均有被发现, 其中 1R-*MYB* 和 R2R3-*MYB* 是植物中最主要的类型^[7-8]。

R2R3-*MYB* 拥有一个保守的 N 端 DNA 结合域 (*MYB* 结构域) 和一个 C 端转录激活或抑制结构域, 而其他区域变异较大^[9-10]。R2R3-*MYB* 具有功能多样性, 广泛参与初生和次生代谢、生长发育、细胞凋亡以及生物非生物胁迫^[4]。前人研究表明 *FaMYB10* 特异在草莓成熟果实中表达, 同时参与调控生长素合成路径基因表达; 而 *FaMYB1* 却作为生长抑制因子调控花色苷合成基因表达^[11-12]。玉米 G34 群体中 *ZmMYB021/RS2* 基因, 通过抑制 *KNOX* 基因在叶片中的表达参与调控器官正常生长及侧生组织器官发育途径^[13]。拟南芥 *AtMYB96* 参与脱落酸和水杨酸信号通路, 通过脱落酸信号通路促进水杨酸的合成从而调控气孔运动、耐旱性和抗病性^[14-16]。拟南芥中一些 *MYB* 基因调控花药发育, 包括 *AtMYB21*、*AtMYB24*、*AtMYB108/BOS1*、*AtMYB35/TDF1*、*AtMYB80* 和 *AtMYB125/DUO1*。*AtMYB24* 和 *AtMYB108* 具有功能冗余, 均为 *AtMYB21* 的下游基因^[17]。*AtMYB80* 作为 *AtMYB35* 下游调控基因共同调控花药外壁的形成^[18]。*AtMYB125/DUO1* 特异在花药中表达, 调控雄性生殖细胞的分裂和分化^[19]。

到目前为止, 对拟南芥^[4]、玉米^[20]、甜菜^[21]、

葡萄^[22]和桃^[23]等多个物种中的 *MYB* 家族基因进行了全基因组分析, 然而在基因组水平上研究草莓 *MYB* 基因及其在花药发育过程中的功能研究尚未见报道。草莓拥有注释较为完整的基因组使得利用生物信息学手段全基因组分析 *MYB* 基因成为可能^[24]。本研究利用生物信息学手段预测森林草莓 (*Fragaria vesca* L.) 基因组中 *MYB* 基因, 同时分析其基因结构、蛋白保守结构域、染色体定位、基因进化及花药发育过程中的表达模式, 为挑选及克隆草莓 *MYB* 基因提供信息, 并为进一步研究 *MYB* 转录因子在草莓花药发育过程中的功能提供参考。

1 材料与方法

1.1 材料

森林草莓全基因组数据文件来源于 GDR 数据库 (ftp://ftp.bioinfo.wsu.edu/species/Fragaria_vesca/Fvesca-genome.v2.0.a1/)。拟南芥 *MYB* 家族基因核酸序列和蛋白序列下载自 tair (<https://www.arabidopsis.org/>)。

1.2 全基因组预测森林草莓 *MYB* 基因

截取拟南芥 125 个 R2R3-*MYB* 蛋白的 R2R3 保守结构域 (包含约 104 个氨基酸), 通过 HMMER 3.0 (28 March 2010)^[25] 软件构建 R2R3 保守结构域的隐马尔可夫模型 (HMM), 然后检索草莓基因组蛋白 FASTA 文件, 任何可以匹配基因家族结构域的蛋白序列被初步认为是森林草莓 *MYB* 基因家族成员, Hmsearch 的 E 值设置 $<e-10$ 。手动剔除不包含完整 R2R3 保守结构域的序列, 同时通过 BLASTn 程序同源检索 NR 数据库确认其是否为 *MYB* 基因。通过 ExPASy 在线软件获得 *MYB* 蛋白的等电点、分子量。

1.3 森林草莓 *MYB* 转录因子 R2R3 保守基序分析

截取最终确认的 105 个森林草莓 R2R3-*MYB* 转录因子的 R2R3 结构域序列 (大约 104 个氨基酸残基), 利用 ClustalW 软件^[26] 生成多序列比对文件, 在线软件 WebLogo3^[27] 生成 R2R3 结构域 logos 图片。

1.4 森林草莓 MYB 基因结构、染色体定位及系统进化分析

通过森林草莓基因组 GFF 文件获取森林草莓 MYB 结构信息,通过 GSDB 2.0^[28] 在线软件生成基因结构图;通过森林草莓基因组 GFF 文件获取染色体定位信息,手工标示每个基因在染色体上的位置,得到各个基因在染色体上的分布状况。利用 ClustalW 软件对基因家族进行多序列比对。MEGA 5 软件^[29] 构建邻接树 (NJ),设置 Bootstrap 为 1000 次重复。

1.5 森林草莓 MYB 基因花药发育过程中的表达谱分析

从 SGR (Strawberry genomic resources)^[30] 数据库获得森林草莓花药发育过程的转录组数据,包括 5 个时期,对应拟南芥花发育时期划分标准的 7-8、9、10、11、12。7-8: 雄蕊和雌蕊迅速发育伸长,花药体积迅速增大,4 个花粉囊结构形成;9: 花瓣开始发育,面积明显增加;10: 花瓣长度与弱雄蕊 (SS) 一致;11: 雌蕊顶部毛状细胞形成;12: 花瓣与强雄蕊 (LS) 长度一致。利用 HemI 1.0 软件^[31] 生成草莓 MYB 基因 RPMK 值对应的表达热图。

2 结果与分析

2.1 森林草莓 MYB 基因鉴定及保守结构域分析

通过 HMMER 3.0 程序,利用 125 个拟南芥 R2R3-MYB 蛋白的 DNA 结合域序列构建隐马尔科夫模型,利用这个模型从森林草莓基因组中检索到 127 个蛋白序列含有 2 个及以上 MYB 重复单元。这些蛋白序列来源于 110 个基因,其中 12 个基因发生可变剪切事件。通过 BLASTn 程序比对 NCBI 的 NR 数据库对基因进行注释,发现 1 个含有 4 个 MYB 重复单元的基因为 MYB4R,4 个含有 3 个 MYB 重复单元的基因为 MYB3R,剩下的 105 个含有 2 个 MYB 重复单元的基因为 R2R3-MYB。通过 ExPASy 工具,对检索到的 110 个基因进行了蛋白质长度、分子量及等电点等生化属性分析 (表 1)。蛋白质生化属性分析发现,这些基因蛋白长度和大小差异较大,R2R3-MYB 基因蛋白长度范围为 128~577 个氨基酸残基,等电点范围为 4.82~10.19,推测的蛋白分子量范围为 15~65 kD。MYB3R 和 MYB4R 基因含有多个 MYB 重复单元,对应的蛋白长度相对较长,范围为 375~1083 个氨基酸残基,等电点范围为 5.08~9.11,推测的蛋白分子量范围为 42~119 kD。

表 1 森林草莓 MYB 基因家族信息分析

Table 1 Information analysis of MYB gene family in *Fragaria vesca* L.

基因名称 Name	GDR 基因登录号 GDR gene accession No.	染色体定位 Chromosome location	GDR 蛋白登录号 GDR protein accession No.	蛋白大小 (aa) Protein length	等电点 PI	分子量 (kD) MW
<i>FvMYB1</i>	gene00185	Fvb2: 24788606-24795594	gene00185-v1.0-hybrid.t1	285	9.30	30308.02
<i>FvMYB1</i>	gene00185	Fvb2: 24788606-24795594	gene00185-v1.0-hybrid.t2	327	9.43	35107.70
<i>FvMYB1</i>	gene00185	Fvb2: 24788606-24795594	gene00185-v1.0-hybrid.t3	327	9.43	35107.70
<i>FvMYB2</i>	gene00431	Fvb7: 4474135-4479574	gene00431-v2.0.a2-hybrid.t1	471	5.39	51273.18
<i>FvMYB2</i>	gene00431	Fvb7: 4474135-4479574	gene00431-v2.0.a2-hybrid.t2	563	5.14	61175.23
<i>FvMYB2</i>	gene00431	Fvb7: 4474135-4479574	gene00431-v2.0.a2-hybrid.t3	563	5.14	61175.23
<i>FvMYB2</i>	gene00431	Fvb7: 4474135-4479574	gene00431-v2.0.a2-hybrid.t4	563	5.14	61175.23
<i>FvMYB3</i>	gene01013	Fvb2: 9731583-9733752	gene01013-v1.0-hybrid.t1	363	9.54	40641.94
<i>FvMYB4</i>	gene01457	Fvb3: 29369784-29371197	gene01457-v1.0-hybrid.t1	327	6.31	36325.78
<i>FvMYB5</i>	gene01992	Fvb6: 32975063-32976701	gene01992-v1.0-hybrid.t1	373	5.85	41324.22
<i>FvMYB6</i>	gene01997	Fvb6: 33006779-33008374	gene01997-v1.0-hybrid.t1	290	5.43	33031.15
<i>FvMYB7</i>	gene01998	Fvb6: 33011702-33012896	gene01998-v1.0-hybrid.t1	267	4.99	30401.82
<i>FvMYB8</i>	gene02214	Fvb5: 16065664-16067468	gene02214-v2.0.a2-hybrid.t1	357	7.65	39982.82
<i>FvMYB9</i>	gene02324	Fvb5: 28949396-28952220	gene02324-v2.0.a2-hybrid.t1	468	5.52	50864.90
<i>FvMYB9</i>	gene02324	Fvb5: 28949396-28952220	gene02324-v2.0.a2-hybrid.t2	468	5.52	50864.90

表 1(续)

基因名称 Name	GDR 基因登录号 GDR gene accession No.	染色体定位 Chromosome location	GDR 蛋白登录号 GDR protein accession No.	蛋白大小 (aa) Protein length	等电点 PI	分子量 (kD) MW
<i>FvMYB10</i>	gene02643	Fvb2: 24172549-24174247	gene02643-v2.0.a2-hybrid.t1	421	6.64	46573.26
<i>FvMYB11</i>	gene02652	Fvb2: 24118298-24119877	gene02652-v1.0-hybrid.t1	278	5.64	30674.36
<i>FvMYB12</i>	gene03321	Fvb3: 12784054-12786833	gene03321-v2.0.a2-hybrid.t1	554	6.88	61179.73
<i>FvMYB13</i>	gene03323	Fvb3: 12793051-12795669	gene03323-v2.0.a2-hybrid.t1	535	6.31	59945.23
<i>FvMYB14</i>	gene03712	Fvb4: 26918469-26920148	gene03712-v1.0-hybrid.t1	284	5.62	32085.98
<i>FvMYB15</i>	gene03817	Fvb6: 35605081-35607534	gene03817-v1.0-hybrid.t1	361	6.16	39442.95
<i>FvMYB16</i>	gene03993	Fvb4: 29347584-29353269	gene03993-v2.0.a2-hybrid.t1	480	6.45	54000.25
<i>FvMYB16</i>	gene03993	Fvb4: 29347584-29353269	gene03993-v2.0.a2-hybrid.t2	480	6.45	54000.25
<i>FvMYB17</i>	gene04057	Fvb4: 29637695-29639712	gene04057-v1.0-hybrid.t1	256	8.60	29224.91
<i>FvMYB18</i>	gene04232	Fvb6: 34452862-34454408	gene04232-v2.0.a2-hybrid.t1	334	6.36	37534.91
<i>FvMYB19</i>	gene04423	Fvb6: 35303926-35306311	gene04423-v2.0.a2-hybrid.t1	372	6.18	40672.03
<i>FvMYB20</i>	gene04710	Fvb4: 28980375-28982277	gene04710-v2.0.a2-hybrid.t1	417	6.81	45685.25
<i>FvMYB21</i>	gene04728	Fvb4: 29080544-29083398	gene04728-v1.0-hybrid.t1	411	6.61	45495.22
<i>FvMYB22</i>	gene04800	Fvb7: 9754784-9756871	gene04800-v1.0-hybrid.t1	326	4.97	36752.40
<i>FvMYB23</i>	gene05018	Fvb4: 21736625-21739563	gene05018-v1.0-hybrid.t1	341	5.99	38151.85
<i>FvMYB24</i>	gene05167	Fvb4: 22596858-22600215	gene05167-v1.0-hybrid.t1	384	5.37	42796.71
<i>FvMYB25</i>	gene05581	Fvb1: 11439868-11441833	gene05581-v1.0-hybrid.t1	364	7.63	40642.43
<i>FvMYB26</i>	gene06920	Fvb6: 26493064-26495265	gene06920-v1.0-hybrid.t1	384	6.12	42485.77
<i>FvMYB27</i>	gene07418	Fvb6: 27035601-27037088	gene07418-v1.0-hybrid.t1	229	8.72	26108.23
<i>FvMYB28</i>	gene07646	Fvb2: 5392968-5396196	gene07646-v1.0-hybrid.t1	267	9.08	29408.41
<i>FvMYB29</i>	gene08006	Fvb3: 25874348-25877063	gene08006-v2.0.a2-hybrid.t1	519	7.77	57813.79
<i>FvMYB30</i>	gene08084	Fvb2: 16990673-16993280	gene08084-v1.0-hybrid.t1	273	5.07	30450.84
<i>FvMYB31</i>	gene08535	Fvb2: 27363945-27367451	gene08535-v1.0-hybrid.t1	358	9.23	41253.09
<i>FvMYB31</i>	gene08535	Fvb2: 27363945-27367451	gene08535-v1.0-hybrid.t2	358	9.23	41253.09
<i>FvMYB31</i>	gene08535	Fvb2: 27363945-27367451	gene08535-v1.0-hybrid.t3	358	9.23	41253.09
<i>FvMYB32</i>	gene08694	Fvb2: 26593842-26595758	gene08694-v1.0-hybrid.t1	310	5.41	34060.82
<i>FvMYB33</i>	gene08880	Fvb5: 11336883-11339995	gene08880-v2.0.a2-hybrid.t1	221	7.72	24989.97
<i>FvMYB33</i>	gene08880	Fvb5: 11336883-11339995	gene08880-v2.0.a2-hybrid.t2	221	7.72	24989.97
<i>FvMYB33</i>	gene08880	Fvb5: 11336883-11339995	gene08880-v2.0.a2-hybrid.t3	232	8.81	26111.15
<i>FvMYB34</i>	gene09039	Fvb2: 25516528-25517939	gene09039-v1.0-hybrid.t1	247	8.86	27051.89
<i>FvMYB35</i>	gene09311	Fvb5: 10445706-10447831	gene09311-v1.0-hybrid.t1	278	8.86	32512.17
<i>FvMYB36</i>	gene09406	Fvb5: 9903871-9905702	gene09406-v1.0-hybrid.t1	192	8.46	21606.39
<i>FvMYB37</i>	gene09407	Fvb5: 9897392-9899486	gene09407-v1.0-hybrid.t1	189	8.99	21409.53
<i>FvMYB38</i>	gene09546	Fvb5: 9153829-9155881	gene09546-v1.0-hybrid.t1	367	5.07	40628.47
<i>FvMYB39</i>	gene09663	Fvb6: 11634509-11636563	gene09663-v1.0-hybrid.t1	318	7.56	35384.50
<i>FvMYB40</i>	gene09683	Fvb6: 11765352-11766742	gene09683-v2.0.a2-hybrid.t1	230	5.15	26615.51
<i>FvMYB41</i>	gene09845	Fvb5: 3502394-3504213	gene09845-v1.0-hybrid.t1	378	5.76	42382.98
<i>FvMYB42</i>	gene10181	Fvb1: 19658095-19659591	gene10181-v2.0.a2-hybrid.t2	341	7.18	37839.29

表 1(续)

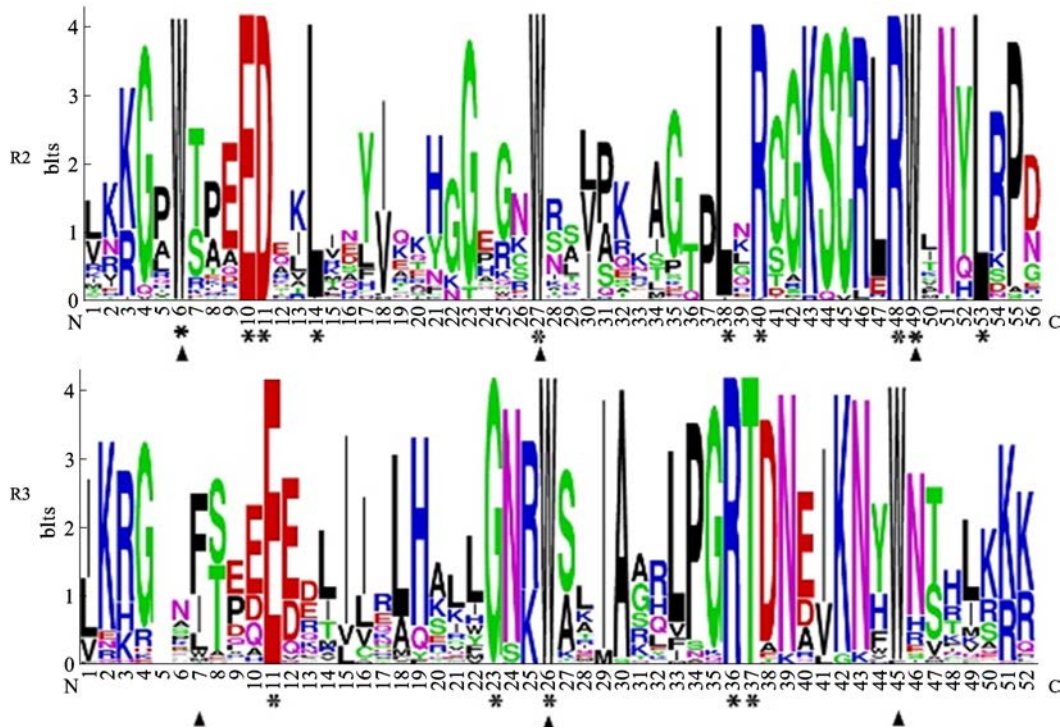
基因名称 Name	GDR 基因登录号 GDR gene accession No.	染色体定位 Chromosome location	GDR 蛋白登录号 GDR protein accession No.	蛋白大小 (aa) Protein length	等电点 PI	分子量 (kD) MW
<i>FvMYB43</i>	gene10329	Fvb6: 33872081-33873496	gene10329-v1.0-hybrid.t1	340	9.18	38450.65
<i>FvMYB44</i>	gene11356	Fvb1: 3498817-3500318	gene11356-v2.0.a2-hybrid.t1	271	8.42	30090.65
<i>FvMYB45</i>	gene11462	Fvb4: 3265822-3267872	gene11462-v1.0-hybrid.t2	257	5.06	28718.01
<i>FvMYB46</i>	gene12261	Fvb1: 12720453-12722439	gene12261-v1.0-hybrid.t1	316	5.22	35793.56
<i>FvMYB47</i>	gene12564	Fvb7: 23507179-23508456	gene12564-v1.0-hybrid.t1	401	6.48	44822.81
<i>FvMYB48</i>	gene12736	Fvb1: 4358634-4360698	gene12736-v2.0.a2-hybrid.t1	271	5.14	30921.39
<i>FvMYB49</i>	gene12738	Fvb1: 4370385-4371870	gene12738-v1.0-hybrid.t2	297	6.15	33075.97
<i>FvMYB50</i>	gene13970	Fvb3: 24674958-24676749	gene13970-v1.0-hybrid.t1	272	8.13	30739.45
<i>FvMYB51</i>	gene14666	Fvb3: 22074437-22077500	gene14666-v1.0-hybrid.t1	336	5.05	37348.33
<i>FvMYB52</i>	gene15383	Fvb2: 23459466-23460748	gene15383-v1.0-hybrid.t1	224	8.83	25813.05
<i>FvMYB53</i>	gene15386	Fvb2: 23470665-23471822	gene15386-v1.0-hybrid.t1	229	7.64	26036.05
<i>FvMYB54</i>	gene15389	Fvb2: 23478176-23479998	gene15389-v2.0.a2-hybrid.t1	301	5.52	33498.23
<i>FvMYB55</i>	gene15390	Fvb2: 23487420-23489146	gene15390-v1.0-hybrid.t1	295	6.34	32987.97
<i>FvMYB56</i>	gene15391	Fvb2: 23491398-23492870	gene15391-v1.0-hybrid.t1	283	5.14	32045.75
<i>FvMYB57</i>	gene15392	Fvb2: 23498361-23499483	gene15392-v1.0-hybrid.t1	281	6.34	31021.82
<i>FvMYB58</i>	gene16287	Fvb1: 17042306-17047139	gene16287-v1.0-hybrid.t1	350	5.82	38797.57
<i>FvMYB58</i>	gene16287	Fvb1: 17042306-17047139	gene16287-v1.0-hybrid.t2	350	5.82	38797.57
<i>FvMYB59</i>	gene18421	Fvb6: 8264548-8267586	gene18421-v1.0-hybrid.t1	358	9.38	41297.17
<i>FvMYB60</i>	gene18690	Fvb7: 14090537-14091793	gene18690-v2.0.a2-hybrid.t1	285	9.18	32319.44
<i>FvMYB61</i>	gene18756	Fvb6: 7526842-7530795	gene18756-v1.0-hybrid.t1	387	7.59	42508.72
<i>FvMYB62</i>	gene19715	Fvb3: 8376976-8378760	gene19715-v2.0.a2-hybrid.t1	340	6.43	37958.55
<i>FvMYB63</i>	gene19934	Fvb3: 7881667-7882694	gene19934-v1.0-hybrid.t1	311	5.90	35224.47
<i>FvMYB64</i>	gene20888	Fvb7: 18016611-18019735	gene20888-v1.0-hybrid.t1	399	6.77	43822.40
<i>FvMYB65</i>	gene21001	Fvb7: 18602643-18604015	gene21001-v1.0-hybrid.t1	256	5.16	27923.23
<i>FvMYB66</i>	gene21264	Fvb7: 20291604-20292820	gene21264-v1.0-hybrid.t1	251	5.21	27934.30
<i>FvMYB67</i>	gene22961	Fvb6: 4932256-4934402	gene22961-v1.0-hybrid.t1	442	7.28	49138.29
<i>FvMYB68</i>	gene23113	Fvb6: 25707552-25710686	gene23113-v2.0.a2-hybrid.t2	250	5.04	28632.58
<i>FvMYB69</i>	gene23946	Fvb6: 32479026-32482513	gene23946-v1.0-hybrid.t1	540	5.72	58612.42
<i>FvMYB70</i>	gene24027	Fvb1: 10333571-10335193	gene24027-v1.0-hybrid.t1	310	8.67	33816.88
<i>FvMYB71</i>	gene25060	Fvb5: 22753746-22755138	gene25060-v1.0-hybrid.t1	199	9.81	22876.00
<i>FvMYB72</i>	gene25098	Fvb5: 6682846-6683906	gene25098-v1.0-hybrid.t1	224	8.23	25348.54
<i>FvMYB73</i>	gene25149	Fvb5: 6949074-6950665	gene25149-v1.0-hybrid.t1	285	5.82	32987.63
<i>FvMYB74</i>	gene25685	Fvb5: 22474335-22476050	gene25685-v1.0-hybrid.t1	314	9.10	35605.97
<i>FvMYB75</i>	gene25982	Fvb5: 8728990-8730021	gene25982-v2.0.a2-hybrid.t1	286	8.81	32311.35
<i>FvMYB76</i>	gene26007	Fvb5: 8625381-8627894	gene26007-v1.0-hybrid.t1	359	6.72	39848.05
<i>FvMYB77</i>	gene26044	Fvb5: 8470810-8472995	gene26044-v2.0.a2-hybrid.t1	318	8.22	35334.80
<i>FvMYB78</i>	gene26045	Fvb5: 8462522-8464247	gene26045-v2.0.a2-hybrid.t1	290	9.02	32130.21
<i>FvMYB79</i>	gene26289	Fvb2: 387196-388723	gene26289-v1.0-hybrid.t1	271	8.31	30074.76

表 1(续)

基因名称 Name	GDR 基因登录号 GDR gene accession No.	染色体定位 Chromosome location	GDR 蛋白登录号 GDR protein accession No.	蛋白大小 (aa) Protein length	等电点 PI	分子量 (kD) MW
<i>FvMYB80</i>	gene26513	Fvb5: 25411700-25413075	gene26513-v2.0.a2-hybrid.t1	128	10.19	15016.27
<i>FvMYB80</i>	gene26513	Fvb5: 25411700-25413075	gene26513-v2.0.a2-hybrid.t2	128	10.19	15016.27
<i>FvMYB81</i>	gene27001	Fvb7: 914954-917555	gene27001-v1.0-hybrid.t1	335	6.18	37648.00
<i>FvMYB82</i>	gene27379	Fvb4: 12387971-12389812	gene27379-v1.0-hybrid.t1	364	6.01	41073.85
<i>FvMYB83</i>	gene28021	Fvb6: 37722831-37724012	gene28021-v1.0-hybrid.t1	282	4.91	32009.81
<i>FvMYB84</i>	gene28293	Fvb4: 10760934-10762325	gene28293-v1.0-hybrid.t1	260	8.97	29916.45
<i>FvMYB85</i>	gene28435	Fvb6: 36919007-36925216	gene28435-v1.0-hybrid.t1	208	7.73	23207.98
<i>FvMYB86</i>	gene28443	Fvb6: 36971717-36975168	gene28443-v2.0.a2-hybrid.t1	550	8.33	62778.73
<i>FvMYB87</i>	gene28902	Fvb3: 15585030-15587155	gene28902-v1.0-hybrid.t1	377	8.03	42547.53
<i>FvMYB88</i>	gene29254	Fvb5: 18345666-18347971	gene29254-v1.0-hybrid.t1	343	8.32	37966.87
<i>FvMYB88</i>	gene29254	Fvb5: 18345666-18347971	gene29254-v1.0-hybrid.t2	341	8.32	37796.66
<i>FvMYB89</i>	gene29961	Fvb3: 9566920-9569656	gene29961-v1.0-hybrid.t1	378	5.08	42127.79
<i>FvMYB90</i>	gene30024	Fvb5: 27258682-27261096	gene30024-v1.0-hybrid.t1	323	5.95	36208.52
<i>FvMYB91</i>	gene30520	Fvb1: 21129958-21131928	gene30520-v1.0-hybrid.t1	285	4.82	32565.16
<i>FvMYB92</i>	gene30725	Fvb3: 2092924-2094535	gene30725-v2.0.a2-hybrid.t1	367	5.67	42024.57
<i>FvMYB93</i>	gene30736	Fvb3: 2028312-2029906	gene30736-v1.0-hybrid.t1	260	5.35	29444.70
<i>FvMYB94</i>	gene31098	Fvb1: 1518634-1520058	gene31098-v1.0-hybrid.t2	320	6.64	36116.90
<i>FvMYB95</i>	gene31413	Fvb1: 15405667-15407900	gene31413-v1.0-hybrid.t1	269	8.54	31384.64
<i>FvMYB96</i>	gene31835	Fvb5: 1823181-1825643	gene31835-v1.0-hybrid.t1	354	5.69	39997.76
<i>FvMYB97</i>	gene31836	Fvb5: 1829971-1831738	gene31836-v1.0-hybrid.t1	294	6.66	33361.48
<i>FvMYB98</i>	gene32198	Fvb5: 3125877-3127877	gene32198-v1.0-hybrid.t1	500	5.58	56421.46
<i>FvMYB99</i>	gene32268	Fvb5: 297286-299254	gene32268-v1.0-hybrid.t1	280	4.98	31253.96
<i>FvMYB100</i>	gene32611	Fvb4: 2217300-2222864	gene32611-v1.0-hybrid.t1	577	5.96	65023.98
<i>FvMYB101</i>	gene36520	Fvb3: 32656522-32658181	gene36520-v2.0.a2-hybrid.t1	347	7.68	37023.11
<i>FvMYB102</i>	gene37047	Fvb5: 3026296-3028276	gene37047-v2.0.a2-hybrid.t1	257	6.00	29297.30
<i>FvMYB102</i>	gene37047	Fvb5: 3026296-3028276	gene37047-v2.0.a2-hybrid.t2	257	6.00	29297.30
<i>FvMYB103</i>	gene37838	Fvb6: 30638633-30640095	gene37838-v2.0.a2-hybrid.t1	213	6.14	24010.03
<i>FvMYB104</i>	gene37997	Fvb6: 36082254-36084374	gene37997-v2.0.a2-hybrid.t1	415	6.46	45397.19
<i>FvMYB105</i>	gene38309	Fvb7: 11944500-11947975	gene38309-v2.0.a2-hybrid.t3	364	6.28	40851.03
<i>FvMYB3R1</i>	gene08793	Fvb2: 26089472-26095781	gene08793-v2.0.a2-hybrid.t1	1083	5.08	119487.22
<i>FvMYB3R2</i>	gene19114	Fvb7: 11681804-11686348	gene19114-v2.0.a2-hybrid.t1	552	9.11	62369.05
<i>FvMYB3R3</i>	gene37595	Fvb6: 8304421-8311334	gene37595-v2.0.a2-hybrid.t1	670	7.22	73808.26
<i>FvMYB3R4</i>	gene32418	Fvb5: 1099328-1103035	gene32418-v2.0.a2-hybrid.t1	375	8.56	42007.77
<i>FvMYB3R4</i>	gene32418	Fvb5: 1099328-1103035	gene32418-v2.0.a2-hybrid.t2	435	8.38	48780.23
<i>FvMYB4R</i>	gene09412	Fvb5: 9867185-9873902	gene09412-v2.0.a2-hybrid.t1	920	8.03	105052.61
<i>FvMYB4R</i>	gene09412	Fvb5: 9867185-9873902	gene09412-v2.0.a2-hybrid.t2	927	8.03	105740.32

为研究森林草莓 R2R3-MYB 保守结构域,共截取 105 个森林草莓 R2R3-MYB 蛋白序列中 MYB 重复单元序列 (R2R3 结构域, 大约 104 个氨基酸残基), 利用 ClustalW 软件生成多序列比对文件, 在线软件 WebLogo 3 生成 R2R3 结构域 logos 图片 (图 1)。

森林草莓 MYB 具有和拟南芥、玉米、葡萄、大豆相似的 R2R3 结构域序列, 均含有特征性的氨基酸, 包括最保守的有序间隔出现的色氨酸 (W) 残基, 这些色氨酸残基在构成 MYB 蛋白螺旋结构以及特异的 DNA 位点方面起重要作用。



*: 代表 105 个 R2R3 蛋白序列在该位置上氨基酸 100% 保守; ▲: 代表一连串的保守色氨酸 (W) 残基

Asterisk: represents the amino acid in 105 R2R3 protein sequences that are 100% conserved at this location, Triangle: represents a series of conservative tryptophan (W) residues

图 1 森林草莓 MYB 蛋白的 R2R3 保守结构域序列标签

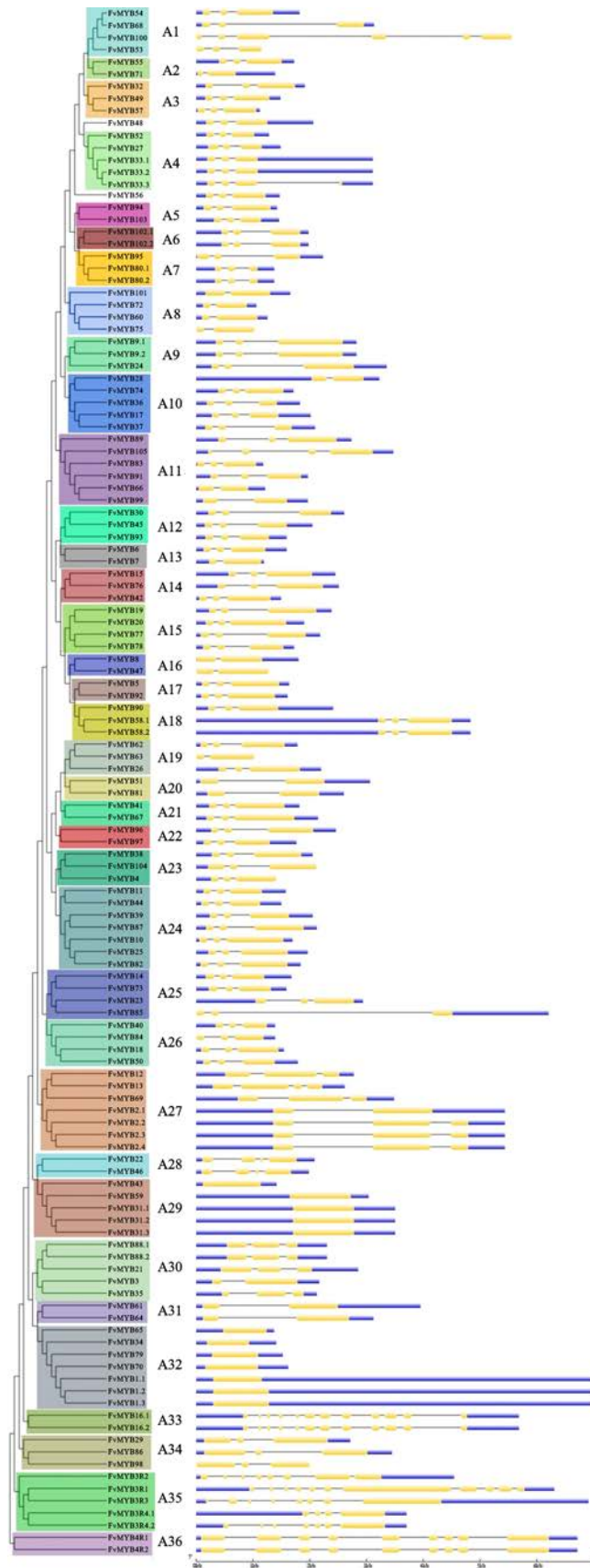
Fig.1 The conserved domain logos of the R2R3 domain in 105 *Fragaria vesca* L. R2R3-MYB proteins

2.2 森林草莓 MYB 基因结构及系统进化分析

为研究森林草莓 MYB 进化关系, 我们利用 110 个森林草莓 MYB 基因 (包括可变剪切) 家族的全长蛋白序列构建进化树; 同时通过整理森林草莓基因组基因结构 GFF 文件获得森林草莓 MYB 内含子和外显子结构信息, 利用在线软件 GSDS 2.0 生成基因结构图; 结合系统进化树和基因结构信息, 将森林草莓 MYB 基因分成了 36 个亚组 (Group) (图 2)。从图中可以发现 4 个 MYB3R (*FvMYB3R1*, *FvMYB3R2*, *FvMYB3R3*, *FvMYB3R4*) 和 1 个 MYB4R (*FvMYB4R*) 基因含有较多的外显子数目, 分别被单独分为 A35 和 A36 组。草莓 R2R3-MYB 基因外显子数量差异较大, 被分成了 34 个亚组 (A1~A34), 同一个组中的成员大多具有相似的基因结构, 如 A29 和 A32 中所有的基因均有唯一的外显子, 而 A33 中的基因具有 12 个外显子。

2.3 森林草莓 MYB 基因功能推测

拟南芥中含有 126 个 R2R3-MYB 基因, 其中大部分基因具有较为完善的基因功能注释, 根据系统进化关系和基因功能注释, 拟南芥 R2R3-MYB 被分为 25 个亚组 (S1~S25)^[3]。为进一步研究森林草莓中 MYB 的基因功能, 我们将森林草莓和拟南芥的 MYB 共同构建系统进化树。从图 3 可以发现 A1、A2、A3、A4、A13、A23 亚组基因为森林草莓所特有; 而草莓中 4 个 MYB3R 和拟南芥的 5 个 MYB3R 被分在同一分支; 同样森林草莓和拟南芥的 MYB4R 也具有非常近的同源关系。直系同源基因通常具有相同的功能, 我们可以通过拟南芥基因功能推测森林草莓中对应的直系同源基因的功能。如黄酮合成方面: 拟南芥 *AtMYB11/PFG1*、*AtMYB12/PFG1* 和 *AtMYB111/PFG3* (S7) 调控各个组织中黄酮物质的合成^[32], 而森林草莓 A9 亚组和拟南芥 S7 亚

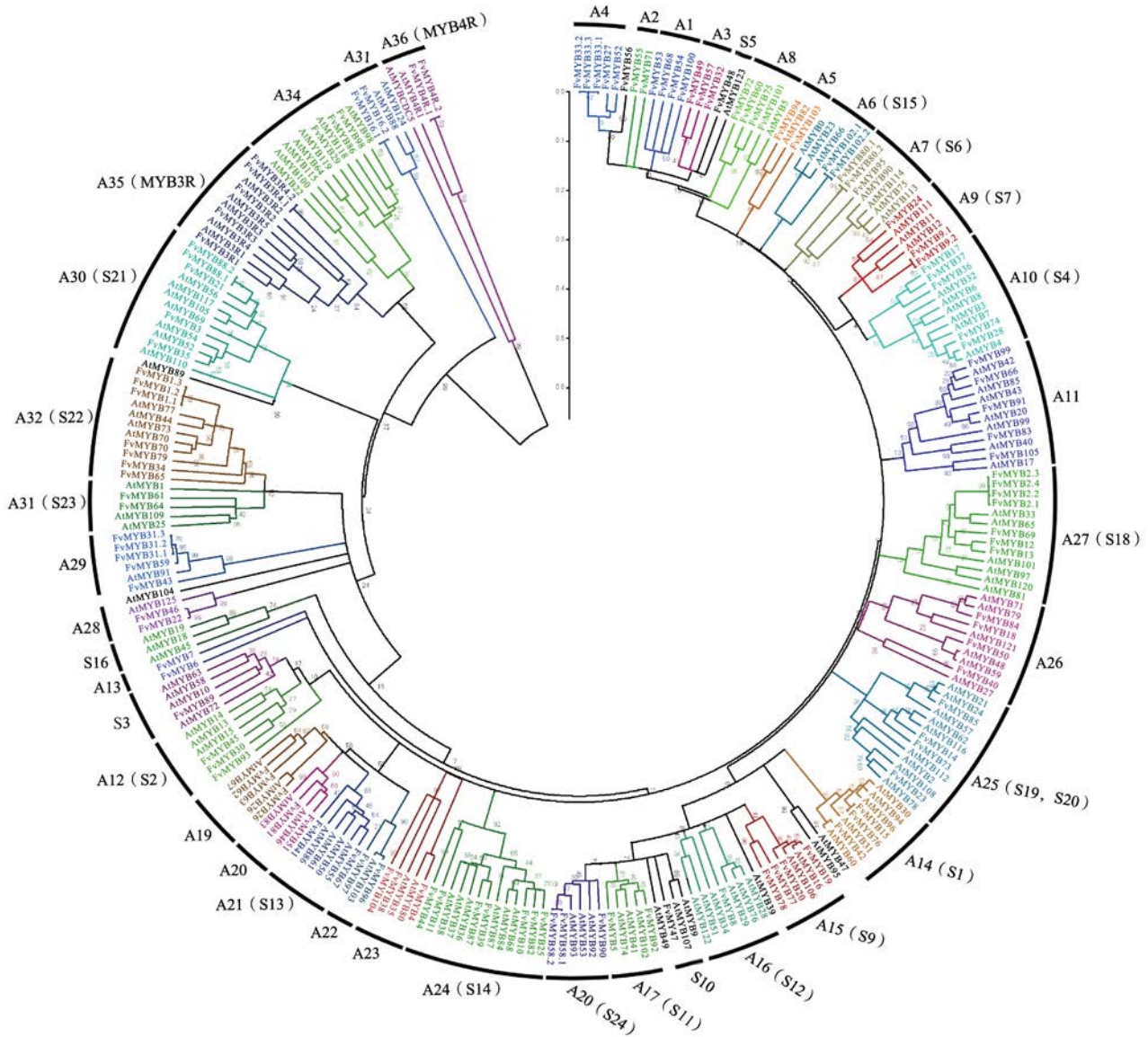


蓝线代表基因 UTR 区,黄线代表外显子,黑线代表内含子

The blue line represents the UTR region of the gene, the yellow line represents exons, and the black line represents introns

图 2 森林草莓 MYB 基因家族基因结构和系统进化分析

Fig.2 Gene structure and evolutionary tree of *Fragaria vesca* L. MYB gene family



At 代表拟南芥基因, Fv 代表森林草莓基因

The blue line represents the UTR region of the gene, the yellow line represents exons, and the black line represents introns

At represents Arabidopsis gene, and Fv represents Fragaria vesca L. gene

图 3 森林草莓和拟南芥 MYB 基因家族系统进化分析和亚组分类

Fig.3 Phylogenetic relationships and subgroup division of MYB gene family from *Fragaria vesca* L. and *arabidopsis*

组处于同一分支,森林草莓 A9 亚组可能参与黄酮物质合成;拟南芥 *AtMYB75/PAP1*、*AtMYB90/PAP2*、*AtMYB113* 和 *AtMYB114* (S6) 调控花色苷合成^[33], 同样森林草莓 A7 亚组和拟南芥 S6 亚组处于同一分支,森林草莓 A7 亚组可能参与花色苷物质合成;细胞发育方面: *AtMYB0/GL1*、*AtMYB23*、*AtMYB66/WER* (S15) 控制拟南芥嫩芽的毛状体初始化^[34], 森林草莓 A6 亚组和拟南芥 S15 亚组处于同一分支可能参与毛状体发育;生物和非生物胁迫方面: 拟南芥 S1 被认为参与生物和非生物胁迫, *AtMYB30* 可以激活病原菌引起的超敏细胞程序死亡^[34], 而

AtMYB60 和 *AtMYB96* 参与 ABA 介导的气孔运动、干旱胁迫^[15,35], 森林草莓 A14 亚组和拟南芥 S1 亚组处于同一分支可能参与生物和非生物胁迫。

2.4 森林草莓 MYB 基因染色体定位分析

根据基因位置信息得到 MYB 基因在森林草莓染色体上的定位图(图 4), *R2R3-MYB* 基因不均匀分布在 7 条森林草莓染色体上,其中 5 号染色体上数量最多为 25 个,7 号染色体上数量最少为 9 个。部分 *R2R3-MYB* 基因在染色体上同一区域集中分布,尤其是 7 个 *R2R3-MYB* 基因(*FvMYB52*、*FvMYB53*、*FvMYB54*、*FvMYB55*、*FvMYB56*、

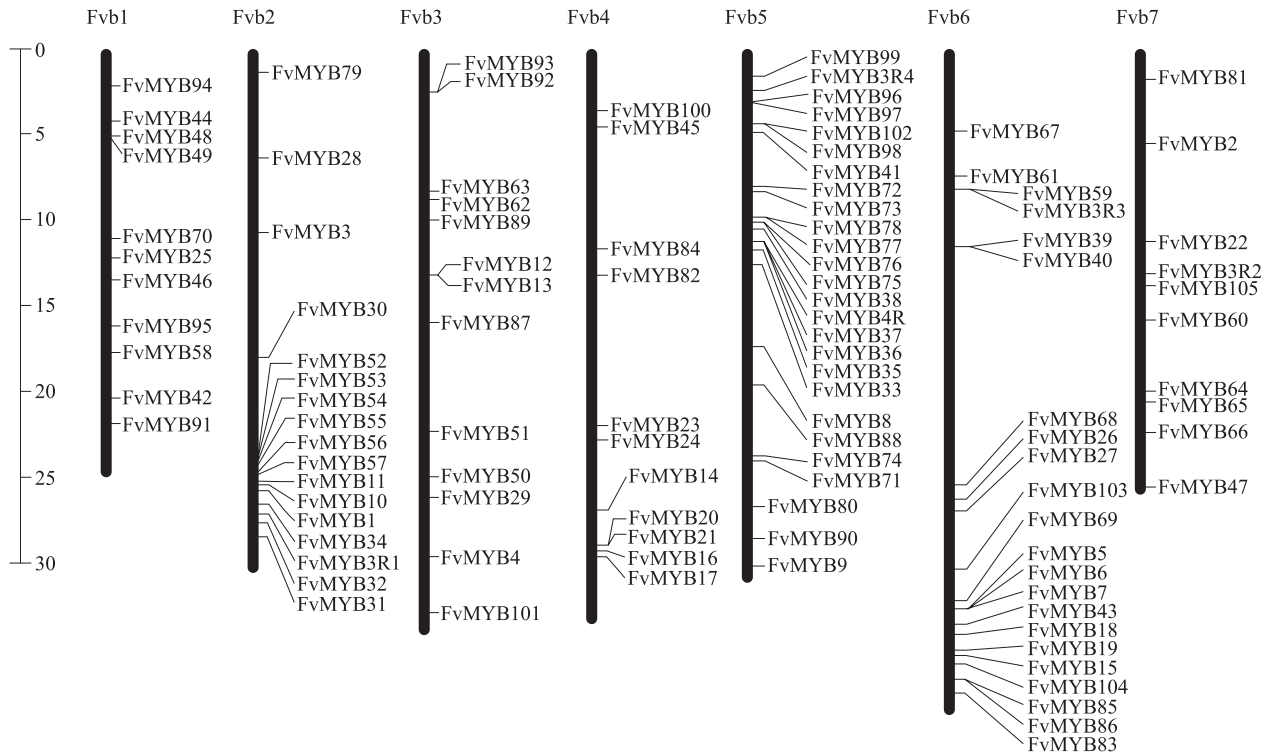


图4 MYB基因的染色体定位图

Fig.4 Chromosomal location information of MYB family

FvMYB57) 在 2 号染色体下端紧密连锁形成基因簇。有趣的是这 7 个基因均位于森林草莓所特有亚组 (A1、A2、A3、A4), 拟南芥中没有。4 个 *MYB3R* 基因均匀分布在 2、5、6 和 7 号染色体上。*MYB4R* 基因定位在 5 号染色体上。

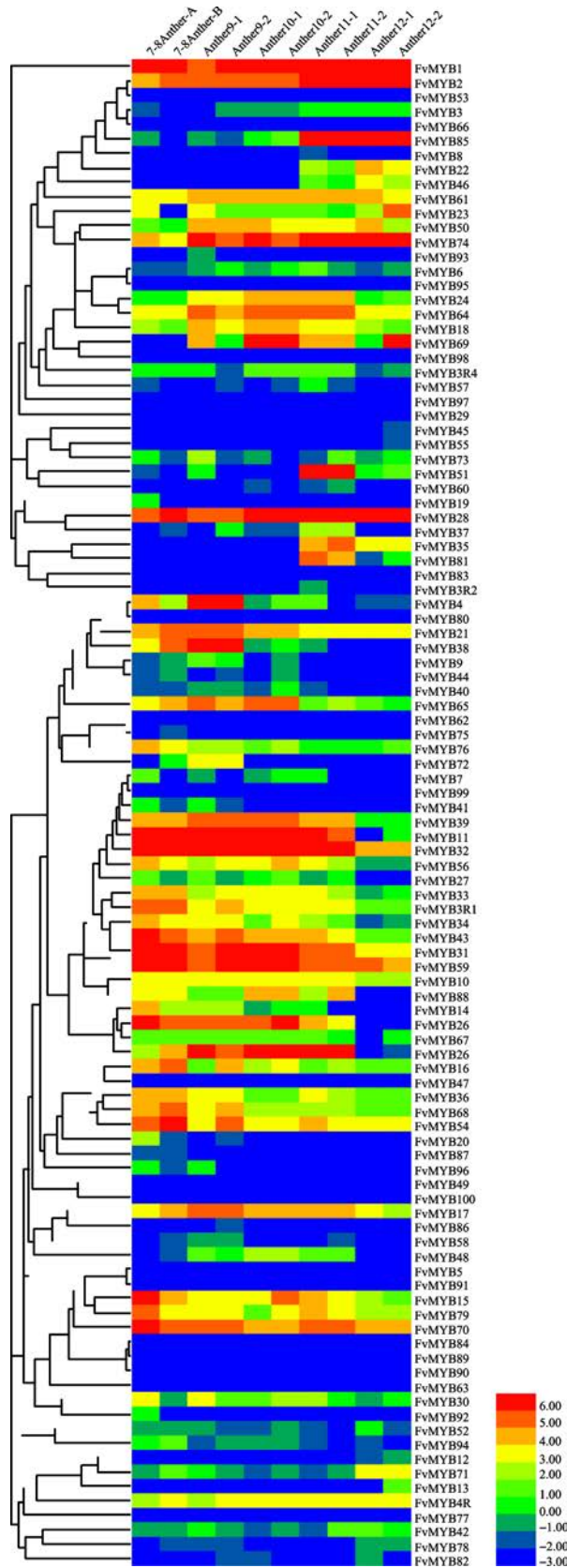
2.5 森林草莓花药发育过程中 MYB 基因的表达模式分析

为了解草莓花药发育过程中 MYB 基因的表达变化, 我们从 SGR (strawberry genomic resources) 数据库获得草莓花药发育过程中的转录组数据, 包括 5 个时期 (对应拟南芥花发育时期划分标准的 7-8、9、10、11、12), 每个时期 2 个重复。获取 110 个森林草莓 MYB 基因在每个时期对应 RPMK 值并利用 HemI 1.0 软件生成对应的基因表达热图, 其中 5 个基因 (*FvMYB101*、*FvMYB102*、*FvMYB103*、*FvMYB104*、*FvMYB105*) 未找到对应的转录组数据。从图 5 可以看出, 大量森林草莓 MYB 基因在花药中表达量较低, 一些森林草莓 MYB 基因在花药中有较高的表达量, 尤其是 *FvMYB1* 在花药发育 5 个时期均有很高的表达量。基因 *FvMYB2*、*FvMYB85*、*FvMYB74*、*FvMYB28* 在花药发育的 5 个时期中表达量逐渐升高, 尤其是在 11 和 12 两个时期具有非常高的表达量, 而且 *FvMYB74* 和 *FvMYB28* 具有非

常近的同源关系, 共同属于 A10 亚组, 表明这些基因可能在花药发育过程中起重要作用; *FvMYB11*、*FvMYB32*、*FvMYB31*、*FvMYB59* 在花药发育的 5 个时期中表达量逐渐降低, 在花药发育的 7-8、9、10 时期表达水平较高, 而在 12 时期表达量很低, 表明这些 MYB 基因可能是花药发育的抑制因子。

3 讨论

MYB 基因家族被认为是植物中最大的一个转录因子家族, 通过 R 结构域的数目可以简单的分为 4 个亚组: 1R-MYB 具有 1 个 R 结构域, 可能是一类重要的端粒结合蛋白^[36]; 2R3-MYB 具有 2 个 R 结构域, 该亚组成员数目众多, 参与次生代谢调节、细胞分化、生物/非生物胁迫应答等^[3]; 3R-MYB 具有 3 个 R 结构域, 主要参与调控细胞周期和细胞分化^[37]; 4R-MYB 在植物中多以单基因存在, 目前基因功能研究不清晰。多年来, MYB 基因的功能研究一直是分子生物研究的一个热点。到目前为止, 拟南芥、葡萄、梨、毛果杨、水稻等多种植物中 MYB 基因家族被鉴定和分析 (表 2)。本文从森林草莓基因组中预测到 105 个 2R3-MYB、4 个 3R-MYB 和 1 个 4R-MYB, 同其他物种中数目类似。



图中色块颜色代表 MYB 基因的表达量,红色越深代表基因表达量越高,蓝色越深代表基因表达量越低
 The color blocks in the figure represent the expression level of MYB genes. The red color represents the higher expression level, and the blue color represents the lower expression level

图 5 森林草莓花药发育过程中 MYB 基因的表达谱
 Fig.5 The expression profile of MYB gene family during *Fragaria vesca* L. anther development

表2 森林草莓、拟南芥、葡萄、梨、毛果杨和水稻基因组中 MYB 基因家族成员数目

Table 2 The number of MYB gene family members in *Fragaria vesca* L., *Arabidopsis*, grape, pear, poplar and rice genome

MYB 基因 MYB gene	森林草莓 <i>Fragaria vesca</i>	拟南芥 <i>Arabidopsis</i>	葡萄 Grape	梨 Pear	毛果杨 Populus	水稻 Rice
R2R3-MYB	105	126	108	105	192	109
3R-MYB	4	5	5	2	5	5
4R-MYB	1	1	1	n.d	n.d	1

n.d 为未检测到

n.d means not detectable

蛋白质生化属性分析发现,森林草莓 MYB 转录因子具有较大的差异性,R2R3-MYB 蛋白长度范围为 128~577 个氨基酸残基,对应的蛋白分子量范围为 15~65 kD;而 MYB3R 和 MYB4R 蛋白长度范围为 375~1083 个氨基酸残基,对应的蛋白分子量范围为 42~119 kD。基因结构分析也同样表明森林草莓 MYB 具有较大的变异,部分森林草莓 R2R3-MYB 基因具有唯一的外显子,而另外一部分 R2R3-MYB 基因具有 12 个外显子。R2R3-MYB 通常含有非常保守的 R2R3 结构域,森林草莓 R2R3-MYB 也同样都包含 R2R3 结构域所具有的特征性的氨基酸,尤其是最保守的有序间隔出现的色氨酸(W)残基。从 R2R3 保守结构域序列标签可以发现,拟南芥的 R2R3 结构域包含 104 个左右的氨基酸残基,而森林草莓 R2R3 却包含有 107 个左右的氨基酸残基;拟南芥中 MYB 基因根据系统关系和基因功能 R2R3-MYB 被分成了 25 个亚组,而森林草莓中 MYB 基因根据系统进化和基因结构分析分成了 36 个亚组(Group),包括 3R-MYB 和 4R-MYB 亚组(图 2)。同一个亚组中的成员大多具有相似的基因结构,这个结果与袁华招等^[38]对葡萄 CHS 和 STS 基因结构的研究结论类似,即在进化树中进化距离越近,基因结构越类似。

直系同源基因通常具有相同的功能,通过共同构建拟南芥和森林草莓 MYB 基因系统进化树,我们发现大部分拟南芥 MYB 亚组可以和森林草莓亚组处于同一分支中。通过同源基因功能预测,森林草莓 MYB 基因可能参与调控类黄酮物质合成、细胞发育、生物/非生物胁迫。为进一步确认森林草莓 MYB 基因功能推测的准确性,分析了森林草莓花药发育过程中 MYB 基因的表达变化。拟南芥 *AtMYB108* 和 *AtMYB24* 为功能冗余基因,共同作为 *AtMYB21* 下游基因,通过茉莉酸(JA)途径调控雄蕊和花粉的成熟^[39]。图 3 可以发现 *AtMYB24*、*AtMYB21* 与 *FvMYB85* 为直系同源基因;同

时 *FvMYB85* 在花药发育的 5 个时期中表达量逐渐升高,尤其是在 11 和 12 两个时期具有非常高的表达量。以上结果表明 *FvMYB85* 可能也具有调控花粉发育的功能。

参考文献

- [1] Latchman D S. Transcription factors: an overview. *The International Journal of Biochemistry & Cell Biology*, 1997, 29 (12): 1305-1312
- [2] Riechmann J L, Heard J, Martin G, Reuber L, Jiang C, Keddie J, Adam L, Pineda O, Ratcliffe O J, Samaha R R, Creelman R, Pilgrim M, Broun P, Zhang J Z, Ghandehari D, Sherman B K, Yu G. *Arabidopsis* transcription factors: Genome-wide comparative analysis among eukaryotes. *Science*, 2000, 290 (5499): 2105-2110
- [3] Dubos C, Stracke R, Grotewold E, Weisshaar B, Martin C, Lepiniec L. MYB transcription factors in *Arabidopsis*. *Trends in Plant Science*, 2010, 15 (10): 573-581
- [4] Ogata K, Kaneiishii C, Sasaki M, Hatanaka H, Nagadoi A, Enari M, Nakamura H, Nishimura Y, Ishii S, Sarai A. The cavity in the hydrophobic core of Myb DNA binding domain is reserved for DNA recognition and trans-activation. *Nature Structural & Molecular Biology*, 1996, 3 (2): 178-187
- [5] 魏麦玲,裴丽丽,刘佳明,闵东红,陈明,李连城,马有志,徐兆师,张小红. 大豆 MYB 转录因子基因 *GmMYB174* 的克隆及分子特性分析. *植物遗传资源学报*, 2015, 16 (1): 94-99
Wei M L, Pei L L, Liu J M, Min D H, Chen M, Li L C, Ma Y Z, Xu Z S, Zhang X H. Isolation and molecular characteristics analysis of soybean transcription factor gene *GmMYB174*. *Journal of Plant Genetic Resources*, 2015, 16 (1): 94-99
- [6] 魏潇,曹珂,王力荣,朱更瑞,方伟超,陈昌文. 红根甘肃桃根系 MYB 基因片段的克隆及序列分析. *植物遗传资源学报*, 2011, 12 (3): 455-459, 463
Wei X, Cao K, Wang L R, Zhu G R, Fang W C, Chen C W. Cloning and characterization of MYB genes from root system of Honggengansutao peach (*Prunus kansuensis* L.). *Journal of Plant Genetic Resources*, 2011, 12 (3): 455-459, 463
- [7] Katiyar A, Smita S, Lenka S K, Rajwanshi R, Chinnusamy V, Bansal K C. Genome-wide classification and expression analysis of MYB transcription factor families in rice and *Arabidopsis*. *BMC Genomics*, 2012, 13: 544

- [8] Chen Y H, Yang X Y, He K, Liu M H, Li J G, Gao Z F, Lin Z Q, Zhang Y F, Wang X X, Qiu X M, Shen Y P, Zhang L, Deng X H, Luo J C, Deng X W, Chen Z L, Gu H Y. The MYB transcription factor super family of *Arabidopsis*: expression analysis and phylogenetic comparison with the rice MYB family. *Nature Structural & Molecular Biology*, 2006, 60: 107-124
- [9] Dias A P, Braun E L, McMullen M D, Grotewold E. Recently duplicated maize R2R3 Myb genes provide evidence for distinct mechanisms of evolutionary divergence after duplication. *Plant Physiology*, 2003, 131 (2): 610-620
- [10] Zimmermann I M, Heim M A, Weisshaar B, Uhrig J F. Comprehensive identification of *Arabidopsis thaliana* MYB transcription factors interacting with R/B-like BHLH proteins. *The Plant Journal*, 2004, 40: 22-34
- [11] Medina-Puche L, Cumplido-Laso G, Amil-Ruiz F, Hoffmann T, Ring L, Rodríguez-Franco A, Caballero J L, Schwab W, Muñoz-Blanco J, Blanco-Portales R. *MYB10* plays a major role in the regulation of flavonoid/phenylpropanoid metabolism during ripening of *Fragaria x ananassa* fruits. *Journal of Experimental Botany*, 2015, 65: 401-417
- [12] Kadomura-Ishikawa Y, Miyawaki K, Takahashi A, Noji S. RNAi-mediated silencing and overexpression of the *FaMYB1* gene and its effect on anthocyanin accumulation in strawberry fruit. *Biologia Plantarum*, 2015, 59 (4): 677-685
- [13] Tropf S, Lanz T. Evidence that stilbene synthases have developed from chalcone synthases several times in the course of evolution. *Journal of Molecular Evolution*, 1994, 38: 610-618
- [14] Cominelli E, Galbiati M, Vavasseur A, Conti L, Sala T, Vuylsteke M, Leonhardt N, Dellaporta S L, Tonelli C. A guard-cell-specific MYB transcription factor regulates stomatal movements and plant drought tolerance. *Current Biology*, 2005, 15: 1196-1200
- [15] Seo P J, Xiang F, Qiao M, Park J Y, Lee Y N, Kim S G, Lee Y H, Park W J, Park C M. The MYB96 transcription factor mediates abscisic acid signaling during drought stress response in *Arabidopsis*. *Plant Physiology*, 2009, 151: 275-289
- [16] Seo P J and Park C M. *MYB96*-mediated abscisic acid signals induce pathogen resistance response by promoting salicylic acid biosynthesis in *Arabidopsis*. *The New Phytologist*, 2010, 186: 471-483
- [17] Mandaokar A and Browse J. *MYB108* acts together with *MYB24* to regulate jasmonate-mediated stamen maturation in *Arabidopsis*. *Plant Physiology*, 2009, 149: 851-862
- [18] Zhang Z B, Zhu J, Gao J F, Wang C, Li H, Li H, Zhang H Q, Zhang S, Wang D M, Wang Q X, Huang H, Xia H J, Yang Z N. Transcription factor *AtMYB103* is required for anther development by regulating tapetum development, callose dissolution and exine formation in *Arabidopsis*. *The Plant Journal*, 2007, 52: 528-538
- [19] Brownfield L, Hafidh S, Borg M, Sidorova A, Mori T, Twell D. A plant germline-specific integrator of sperm specification and cell cycle progression. *PLoS Genetics*, 2009, 5 (3): e1000430
- [20] Du H, Feng B R, Yang S S, Huang Y B, Tang Y X. The R2R3-MYB transcription factor gene family in maize. *PLoS One*, 2012, 7 (6): e37463
- [21] Stracke R, Holtgräwe D, Schneider J, Pucker B, Sørensen T R, Weisshaar B. Genome-wide identification and characterisation of R2R3-MYB genes in sugar beet (*Beta vulgaris*). *BMC Plant Biology*, 2014, 14 (1): 249-266
- [22] Matus J T, Aquea F, Arce-Johnson P. Analysis of the grape *MYB R2R3* subfamily reveals expanded wine quality-related clades and conserved gene structure organization across *Vitis* and *Arabidopsis* genomes. *BMC Plant Biology*, 2008, 8 (1): 83-98
- [23] Zhou H, Peng Q, Zhao J, Owiti A, Ren F, Liao L, Wang L, Deng X, Jiang Q, Han Y. Multiple R2R3-MYB transcription factors involved in the regulation of anthocyanin accumulation in peach flower. *Frontiers in Plant Science*, 2016, 7 (28): 1557-1568
- [24] Shulaev V, Sargent D J, Crowhurst R N, Mockler T C, Folkerts O, Delcher A L, Jaiswal P, Mockaitis K, Liston A, Mane S P, Burns P, Davis T M, Slovin J P, Bassil N, Hellens R P, Evans C, Harkins T, Kodira C, Desany B, Crasta O R, Jensen R V, Allan A C, Michael T P, Setubal J C, Celton J M, Rees D J G, Williams K P, Holt S H, Rojas J J R, Chatterjee M, Liu B, Silva H, Meisel L, Adato A, Filichkin S A, Troggio M, Viola R, Ashman T L, Wang H, Dharmawardhana P, Elser J, Raja R, Priest H D, Bryant D W, Fox S E, Givan S A, Wilhelm L J, Naithani S, Christoffels A, Salama D Y, Carter J, Girona E L, Zdepski A, Wang W Q, Kerstetter R A, Schwab W, Korban S S, Davik J, Monfort A, Denoyes-Rothan B, Arus P, Mittler R, Flinn B, Aharoni A, Bennetzen J L, Salzberg S L, Dickerman A W, Velasco R, Borodovsky M, Veilleux R E, Folta K M. The genome of woodland strawberry (*Fragaria vesca*). *Nature Genetics*, 2010, 43: 109-116
- [25] Eddy S R. Profile hidden Markov models. *Bioinformatics*, 1998, 14 (9): 755-763
- [26] Thompson J D, Higgins D G, Gibson T J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, 1994, 22 (22): 4673-4680
- [27] Crooks G E, Hon G, Chandonia J M, Brenner S E. WebLogo: A sequence logo generator. *Genome Research*, 2004, 14: 1188-1190
- [28] Hu B, Jin J P, Guo A Y, Zhang H, Luo J, Gao G. GSDS 2.0: an upgraded gene feature visualization server. *Bioinformatics*, 2015, 31 (8): 1296-1297
- [29] Tamura K, Peterson D, Peterson N, Stecher G, Nei M, Kumar S. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Molecular Biology and Evolution*, 2011, 28 (10): 2731-2739
- [30] Darwish O, Slovin J P, Kang C, Hollender C A, Geretz A, Houston S, Liu Z, Alkharouf N W. SGR: an online genomic resource for the woodland strawberry. *BMC Plant Biology*, 2013, 13 (1): 223
- [31] Deng W K, Wang Y B, Liu Z X, Cheng H, Xue Y. HemI: A toolkit for illustrating heatmaps. *PLoS One*, 2014, 9 (11): e111988
- [32] Stracke R, Ishihara H, Huep G, Barsch A, Mehrtens F, Niehaus K, Weisshaar B. Differential regulation of closely related R2R3-

- MYB transcription factors controls flavonol accumulation in different parts of the *Arabidopsis thaliana* seedling. *Plant Journal*, 2007, 50: 660-677
- [33] Gonzalez A, Zhao M, Leavitt J M, Lloyd A M. Regulation of the anthocyanin biosynthetic pathway by the TTG1/bHLH/Myb transcriptional complex in *Arabidopsis* seedlings. *The Plant Journal*, 2008, 53: 814-827
- [34] Kang Y H, Kirik V, Hulskamp M, Nam K H, Hagely K, Lee M M, Schiefelbein J. The *MYB23* gene provides a positive feedback loop for cell fate specification in the *Arabidopsis* root epidermis. *Plant Cell*, 2009, 21: 1080-1094
- [35] Cominelli E, Galbiati M, Vavasseur A, Conti L, Sala T, Vuylsteke M, Leonhardt N, Dellaporta S L, Tonelli C. A guard-cell-specific MYB transcription factor regulates stomatal movements and plant drought tolerance. *Current Biology*, 2005, 15: 1196-1200
- [36] Rosinski J A, Atchley W R. Molecular evolution of the Myb family of transcription factors: evidence for polyphyletic origin. *Journal of Molecular Evolution*, 1998, 46: 74-83
- [37] Haga N, Kato K, Murase M, Araki S, Kubo M, Demura T, Suzuki K, Müller I, Voss U, Jürgens G, Ito M. R1R2R3-Myb proteins positively regulate cytokinesis through activation of KNOLLE transcription in *Arabidopsis thaliana*. *Development*, 2007, 134: 1101-1110
- [38] 袁华招, 赵密珍, 吴伟民, 于红梅, 王庆莲. 葡萄 CHS 和 STS 基因家族生物信息学鉴定和表达分析. *植物遗传资源学报*, 2016, 17(4): 756-765
- Yuan H Z, Zhao M Z, Wu W M, Yu H M, Wang Q L. Genome-wide identification and expression analysis of CHS and STS gene families in grape. *Journal of Plant Genetic Resources*, 2016, 17(4): 756-765
- [39] Mandaokar A, Browse J. *MYB108* acts together with *MYB24* to regulate jasmonate-mediated stamen maturation in *Arabidopsis*. *Plant Physiology*, 2009, 149: 851-862