

doi:10.11937/bfyy.20173398

苹果基因组中转录因子的鉴定与分析

许瑞瑞, 高明刚, 刘春香

(潍坊学院 生物与农业工程学院, 山东省高校生物化学与分子生物学重点实验室, 山东 潍坊 261061)

摘要: 转录因子是基因表达调控过程中的重要调节因子。为更好地了解蔷薇科植物苹果转录因子所编码的基因家族, 利用苹果基因组数据进行转录因子筛选鉴定和系统预测, 并与桃和草莓 2 个蔷薇科物种进行分析比较。结果表明: 在苹果、桃、草莓中分别鉴定到 3 039、1 527、1 506 个转录因子成员, 可分为 58 个转录因子家族; 基因染色体定位显示预测的转录因子以不同密度分布在所有染色体上; 所有的转录因子都具有类似的 GO 分析结果和亚细胞定位预测信息; 随机选择了与拟南芥 MYB 转录因子同源性较高的苹果基因进行了干旱胁迫处理下的表达量检测, 值得注意的是, 有 6 个基因经过 PEG 处理后在平邑甜茶和 T337 苹果品种中的表达量具有相似的变化趋势。该研究系统分析鉴定苹果全基因组中的所有转录因子基因家族, 为今后深入了解蔷薇科植物转录因子的分类和基因功能研究提供了一定的参考依据。

关键词: 苹果; 基因组; 转录因子; 鉴定分析; 干旱

中图分类号: S 661.103.6 **文献标识码:** A **文章编号:** 1001-0009(2018)09-0042-09

蔷薇科是双子叶植物中最大、最重要的家族之一, 分布十分广泛, 全世界有超过 3 300 种, 如人们所熟知的苹果、桃、梨、杏、草莓、樱桃、山楂和

李子等著名的水果品种, 其中扁桃仁和杏仁是著名的干果, 有许多优良品种在世界各地普遍栽培; 许多种类的水果含有维生素、糖和有机酸, 可用于食品或添加剂; 另一些则是美丽的花朵、水果和树叶, 是观赏植物的来源, 如玫瑰、中国玫瑰、梅花和樱花等^[1]。蔷薇科植物一直是果树研究领域的热点与重点, 其中苹果、桃和草莓的全基因组序列已测序完成, 给后续的转录因子基因家族的筛选鉴定提供了必要的数据基础^[1-5]。

植物在生长过程中会遭受多种逆境胁迫, 形

第一作者简介: 许瑞瑞(1982-), 女, 博士, 副教授, 研究方向为植物逆境分子生物学与基因表达调控。E-mail: xuruirui2006@163.com.

基金项目: 国家自然科学基金青年资助项目(31400225); 潍坊市科技发展计划资助项目(2016GX016); 山东省自然科学基金资助项目(2014ZRB01AAC)。

收稿日期: 2018-02-24

semithin section technology and their structures were observed by optical microscopy. The whole structures of ovaries, anthers and fruits were observed by stereomicroscope to explore the characteristics of the reproductive organs of seedless hawthorn. The results showed that the petals, anthers, ovaries and fruits of seedless hawthorn were smaller than ordinary hawthorn. The common hawthorn contained more active microspores, and the tapetum cells degrade normally in the process of development, while most germ cells of the seedless hawthorn were vacuolization and the tapetum cells degrade anomaly. Most of the ovules of seedless hawthorn get aborted in the process of development, only the bead was developed as seed coat.

Keywords: seedless *Crataegus pinnatifida* (hawthorn); abortion; ovary; anther; anatomical structure

成了十分复杂的逆境适应和抗逆机制,而基因转录水平上的调控是植物适应环境和抵御逆境胁迫的关键因素之一,转录因子在这一调控过程中扮演重要角色。转录因子(Transcription factors, TFs)可以结合到一个特定的基因序列上游,即基因 5'端,从而控制基因在一个特定时间和空间表达其遗传信息^[6]。转录因子一般包含 1 个或多个 DNA 结合域和 1 个转录调节区域,可以激活或抑制多个目标基因的转录效率^[7-8]。转录因子是激活或抑制编码或非编码基因表达的重要调控因子,在植物生长和逆境信号途径中发挥了多种功能^[9-11]。研究转录因子在蔷薇科植物生长发育及逆境响应过程中的作用具有重要的生物学意义。

研究表明,维管束植物基因组中有约 7% 的基因编码转录因子,目前,通过生物信息学的手段已有超过 50 种不同植物物种的转录因子家族被筛选鉴定分析^[8,12]。经过系统的生物信息学分析,分别在拟南芥、水稻和玉米全基因组序列中发现鉴定到 1 789、2 384 和 2 298 种转录因子家族^[13-15]。近年来,苹果基因组中的转录因子,如 MYB (myeloblastosis-related proteins)、NAC (NAM-ATAF-CUC2)、LBD (the lateral organ boundary domain)和 DREB(dehydration responsive element binding)等被分析鉴定,但是还有很多转录因子家族成员尚鲜见报道^[16-19]。鉴于转录因子在植物生长发育和生理生化代谢过程中的重要性,以及它们在改良植物抗逆性方面的潜在应用价值,该研究对苹果转录因子家族进行了全基因组分析,然后将其与桃和草莓中的转录因子家族进行了分析比较。

1 材料与方法

1.1 苹果、桃和草莓中转录因子家族的鉴定

试验所用的基因组序列从蔷薇科基因组数据库(GDR)下载^[20];然后利用 PFAM 搜索工具进行 HMM 分析获得全部基因的蛋白结构域信息^[21];根据转录因子所包含的结构特征 DNA 结合域(DNA binding domain, DBD)鉴定预测所有的转录因子 TFs;最后根据 PlantTFDB 和 PlnTFDB 网站上的信息,使用基于 Perl 程序获得了所有 PFAM 搜索域中符合上述依据的序列^[8,12]。

1.2 转录因子的染色体定位和基因结构分析

从 GDR 数据库下载的苹果、桃和草莓基因组数据中检索染色体位置和基因信息^[20],利用 Circos 和 MapDraw 软件工具进行染色体定位作图^[22-23]。

1.3 转录因子的 GO 分析和亚细胞定位预测

首先,Blast2GO 工具被下载以创建本地数据库^[24]。GO 分析对试验结果有提示作用,比如通过差异基因的 GO 分析,可以找到富集差异基因的 GO 分类条目,寻找不同样品的差异基因可能相关的基因功能。因此,该研究所有转录因子序列都是使用 GO 分析搜索获得的,然后利用 WEGO 工具绘制^[25]。蛋白质粒度是一种新的理论,是提取蛋白质序列特征的有效方法^[26],利用该方法获得转录因子亚细胞定位信息,其预测结果可从网站 X-Loc (<http://202.194.131.176/>) 获得。

1.4 苹果材料处理和 qRT-PCR 分析

平邑甜茶(*M. hupehensis*)和 T337 幼苗在 26 °C 和 16 h 光照/8 h 黑暗周期的温室条件下生长 4 周后,选择均匀发育的幼苗用 10% 聚乙二醇(PEG 6000)进行干旱模拟处理 6 h,未处理的作为试验对照组,然后分别取材,液氮快速冷冻,并置于一 80 °C 超低温冰箱保存,用于 RNA 提取。苹果总 RNA 提取采用 RNA plant plus Reagent 试剂盒提取,根据 Takara 反转录试剂盒操作说明将 RNA 反转录得到苹果各种 cDNA,−20 °C 冰箱保存,用于 qRT-PCR 表达分析。

qRT-PCR 反应体系参照 SYBR® Green Real time PCR Master Mix 说明书;qRT-PCR 反应条件为:94 °C 1 min;94 °C 10 s,60 °C 10 s,72 °C 10 s,40 个循环,每次循环第 3 步进行荧光采集,每个样品重复 3 次,采用 $2^{-\Delta\Delta CT}$ 法对数据进行定量分析,用 Excel 软件作图。qRT-PCR 试验所需的引物通过 Beacon Designer 8 软件设计,*Md18S* 作为内参基因,试验所用的引物见表 1。

2 结果与分析

2.1 苹果、桃和草莓中转录因子家族的鉴定

目前,蔷薇科植物苹果、桃和草莓的测序工作

表 1 qRT-PCR 所用的引物序列
Table 1 Primers used for the qRT-PCR analyses

基因名称 Gene name	正向引物(5'-3') Forward primer (5'-3')	反向引物(5'-3') Reverse primer (5'-3')
MDP0000805075	AAGGAGGAGTTTGAGTGAAGGG	GGAGGAGGAGGATTGGTTGTTC
MDP0000265114	CACTCATAACCTCCTCCATTTC	TCACTTCACTTCCATTAGACTCAG
MDP0000230141	ACCAGGACGCACCGATAAC	TTGTTGATGTTGTTGATGATGTTG
MDP0000468201	GCAGCACAGAGGATGAGAAGG	TTATGAGACTTAGCCCACTTCCC
MDP0000149535	CAACAACAACAACAACGACAAC	AGTGCTCATTGAACGACCTC
MDP0000210970	CCAAGCACCAATCCTTCAACAAC	ACATATCGGATGGCAAGAGACTAG
Md18S	ACACGGGGAGGTAGTGACAA	CCTCCAATGGATCCTCGTTA

已全部完成,其基因组数据是公开免费的。为了筛选 3 个物种全基因组序列中编码的转录因子,该研究利用 DNA 结合结构域和基序对所有蛋白质序列进行了系统鉴定。最终,通过生物信息学手段在苹果基因组中鉴定出 3 039 个转录因子家族基因,约占苹果总编码基因总数的 4.87% (E 值<0.01),在桃和草莓全基因组中分别鉴定到 1 527 个(约占 5.31%)和 1 506 个(约占 4.31%)转录因子家族基因。根据 DNA 结合域

的同源性可以将所有的转录因子分为 58 个家族,如表 2 所示。其中,MYB 转录因子家族(包括 MYB 和 MYB 相关家族)成员最多,苹果、桃和草莓中各包含 375 个(12.33%)、170 个(11.13%)和 157 个(10.43%)基因成员;其次是 AP2-EREBP、NAC、bHLH、MADS 和 WRKY 转录因子家族;另外,在苹果基因组中未发现 NOZZLE 转录因子家族成员(表 2)。

表 2 苹果、桃和草莓转录因子的数目
Table 2 Number of transcription factors in apple,peach and strawberry

转录因子家族 TF Family	苹果 Apple		桃 Peach		草莓 Strawberry	
	数目 Number	百分比 Percentage/%	数目 Number	百分比 Percentage/%	数目 Number	百分比 Percentage/%
NOZZLE	0	0.00	1	0.07	1	0.07
LFY	2	0.07	1	0.07	4	0.27
HRT	3	0.10	1	0.07	2	0.13
PBF-2-like	3	0.10	2	0.13	2	0.13
S1Fa-like	4	0.13	1	0.07	1	0.07
ULT	4	0.13	2	0.13	1	0.07
VOZ	4	0.13	3	0.20	2	0.13
SAP	5	0.16	2	0.13	2	0.13
BBR/BPC	7	0.23	4	0.26	3	0.20
CSD	8	0.26	1	0.07	4	0.27
SRS	9	0.30	6	0.39	5	0.33
TUB	10	0.33	8	0.52	9	0.60
Alfin-like	11	0.36	7	0.46	6	0.40
ARR-B	11	0.36	9	0.59	6	0.40
CAMTA	11	0.36	4	0.26	4	0.27
TAZ	11	0.36	5	0.33	6	0.40
GRF	12	0.39	9	0.59	10	0.66
PLATZ	12	0.39	9	0.59	11	0.73
C2C2-YABBY	13	0.43	7	0.46	6	0.40
Sigma70-like	13	0.43	7	0.46	6	0.40
GeBP	14	0.46	8	0.52	6	0.40
CCAAT	15	0.49	6	0.39	6	0.40
E2F-DP	15	0.49	6	0.39	7	0.46
C2C2-CO-like	16	0.53	9	0.59	6	0.40
EIL	16	0.53	4	0.26	6	0.40
RWP-RK	17	0.56	8	0.52	10	0.66
DBP	19	0.63	6	0.39	7	0.46
BSD	21	0.69	10	0.65	9	0.60

表 2(续)
Table 2(Continued)

转录因子家族 TF Family	苹果 Apple		桃 Peach		草莓 Strawberry	
	数目	百分比	数目	百分比	数目	百分比
	Number	Percentage/%	Number	Percentage/%	Number	Percentage/%
LIM	23	0.76	15	0.98	10	0.66
BES1	29	0.95	9	0.59	6	0.40
Tify	29	0.95	10	0.65	12	0.80
zf-HD	29	0.95	10	0.65	10	0.66
ARF	31	1.02	18	1.18	18	1.20
FHA	33	1.09	15	0.98	15	1.00
OFP	34	1.12	15	0.98	13	0.86
C2C2-GATA	36	1.18	22	1.44	19	1.26
SBP	42	1.38	17	1.11	16	1.06
C2H2	43	1.41	27	1.77	28	1.86
HSF	48	1.58	21	1.38	16	1.06
Trihelix	57	1.88	27	1.77	31	2.06
TCP	58	1.91	19	1.24	19	1.26
C2C2-Dof	60	1.97	26	1.70	24	1.59
mTERF	62	2.04	41	2.69	48	3.19
G2-like	66	2.17	37	2.42	36	2.39
FAR1	67	2.20	79	5.17	84	5.58
ABI3VP1	80	2.63	69	4.52	77	5.11
LOB	80	2.63	42	2.75	36	2.39
bZIP	112	3.69	49	3.21	50	3.32
C3H	114	3.75	45	2.95	34	2.26
HB	121	3.98	49	3.21	56	3.72
GRAS	127	4.18	49	3.21	53	3.52
MYB-related	129	4.24	46	3.01	45	2.99
WRKY	139	4.57	61	3.99	62	4.12
MADS	142	4.67	80	5.24	86	5.71
bHLH	192	6.32	113	7.40	92	6.11
MYB	246	8.09	124	8.12	112	7.44
NAC	253	8.33	115	7.53	133	8.83
AP2-EREBP	271	8.92	131	8.58	117	7.77
总计	3 039	100.00	1 527	100.00	1 506	100.00

2.2 转录因子家族基因的染色体定位

通过 MapDraw 工具将苹果 3 039 个转录因子中的 2 887 个(95%)基因定位到染色体上,苹果的每条染色体上都有转录因子基因分布,其中第 15 条染色体含有最多的转录因子基因,有 267 个;第 4 条染色体仅含有 113 个转录因子基因,分布最少(图 1、2A)。桃和草莓分别有 1 506 个(98.6%)和 1 383 个(91.8%)转录因子基因定位到染色体上;且每条染色体上基因分布密度不同,其中桃的第 1 条(315 个)和第 8 条(129 个)染色体上分别含有最多和最少的转录因子基因分布;草莓的第 6 条(295 个)和第 4 条(129 个)染色体上分别含有最多和最少的转录因子基因分布(图 2B、2C)。

2.3 转录因子家族基因的 GO 分析

GO 分析用于分类预测转录因子的功能,结

果显示所有转录因子可以分为 40 个功能组,如 x 轴所示,包括 3 个主要分类类别:细胞成分(cellular component)、分子功能(molecular function)和生物过程(biological progress);y 轴左边和右边分别代表基因的百分比和数量;用不同的颜色分别表示苹果、桃和草莓物种(图 3)。

值得注意的是,生物过程中的“代谢过程”(metabolic process)、“运动”(locomotion)、“定位”(localization)、“死亡”(death)和“细胞组分生物发生”(cellular component biogenesis)等在 3 个物种中的基因比例是基本一致的;GO 分析的分子功能中也有差异,例如,“抗氧化剂(antioxidant)”“辅助转运蛋白质(auxiliary transport protein)”“电子载体(electron carrier)”和“酶调节子(enzyme regulator)”在桃和草莓中没有预测到有基因行使以上功能;苹果中没有发现有与“节律过程”(rhythmic process)相关的基因;此外,“包

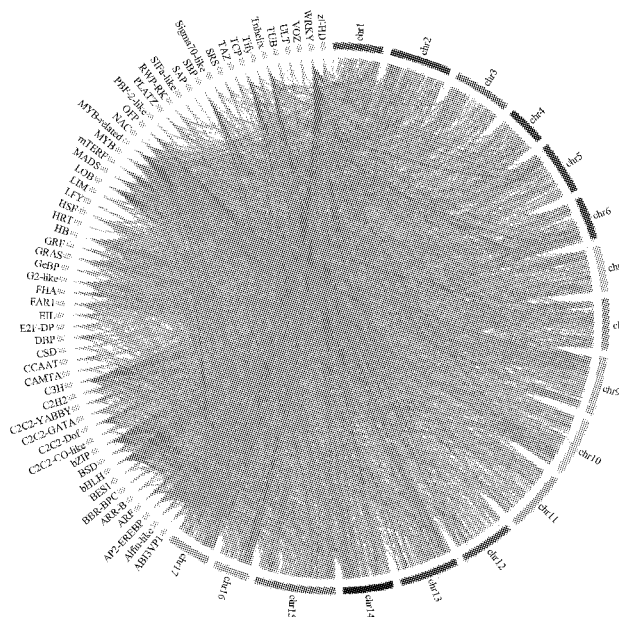


图1 苹果转录因子的染色体定位图

Fig.1 Chromosomal mapping of the transcription factors in apple

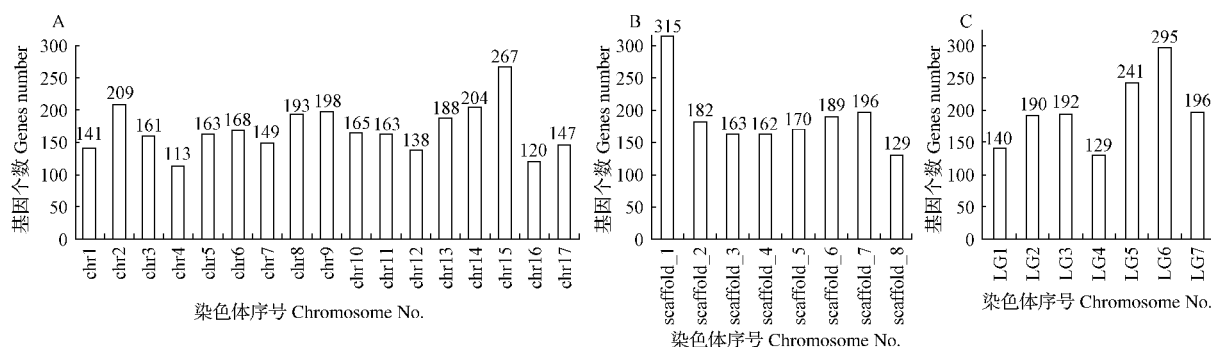


图2 每条染色体上转录因子 TFs 的基因数量分布(A 苹果、B 桃、C 草莓)

Fig.2 Genes number and distribution of the transcription factors in each chromosome(A apple, B peach, C strawberry)

膜(envelope)”“生长过程(growth)”和“多/生物过程(multi/organism process)”的基因比例与数目差异非常明显(图3)。

2.4 转录因子的亚细胞定位预测

蛋白质亚细胞定位预测是生物信息学预测蛋白质功能的重要组成部分。该研究通过蛋白质粒度法预测了苹果、桃和草莓中转录因子在12种(内质网、细胞外基质、过氧化物酶体、质体、高尔基体、细胞壁、液泡、细胞膜、线粒体、细胞质、叶绿体和细胞核)主要细胞器中的定位情况。如图4所示,蔷薇科植物苹果、桃和草莓中发现了类似的

亚细胞定位情况,内质网、细胞外基质、过氧化物酶体、质体、高尔基体、细胞壁和液泡中均未预测到转录因子的定位;苹果、桃和草莓中的转录因子定位在细胞核中分别占到总数的66%、70%和72%;其次是定位在叶绿体中的3个物种转录因子较多,占20%左右;定位在细胞质、线粒体和细胞膜中的转录因子较少,低于10%。

2.5 苹果干旱相关转录因子的 qRT-PCR 分析

转录因子在植物响应生物和非生物胁迫过程中起着十分重要的作用。拟南芥中发现66个转录因子家族,包括DREB、NAC、MYB、WRKY和

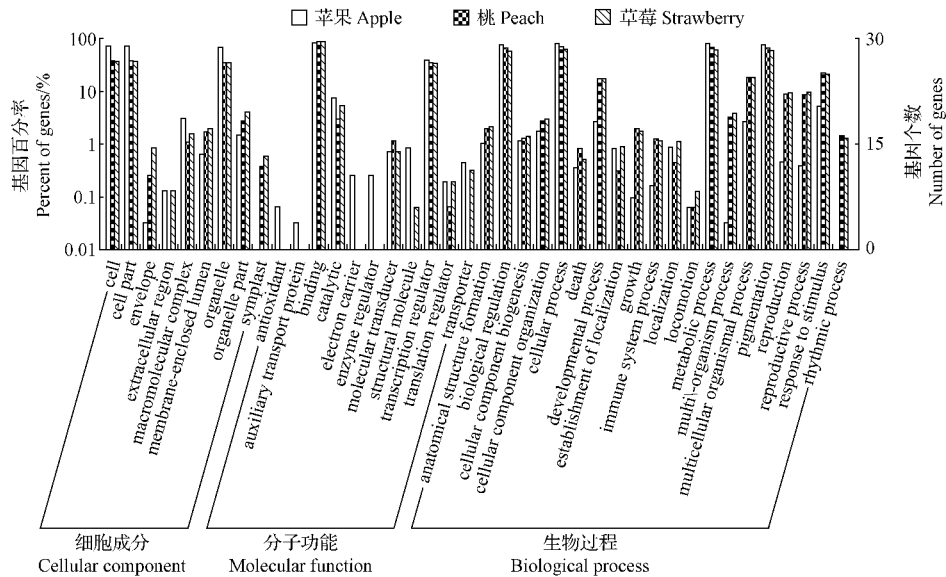


图3 苹果、桃和草莓转录因子的GO分析

Fig. 3 The GO terms of the transcription factors in apple, peach and strawberry

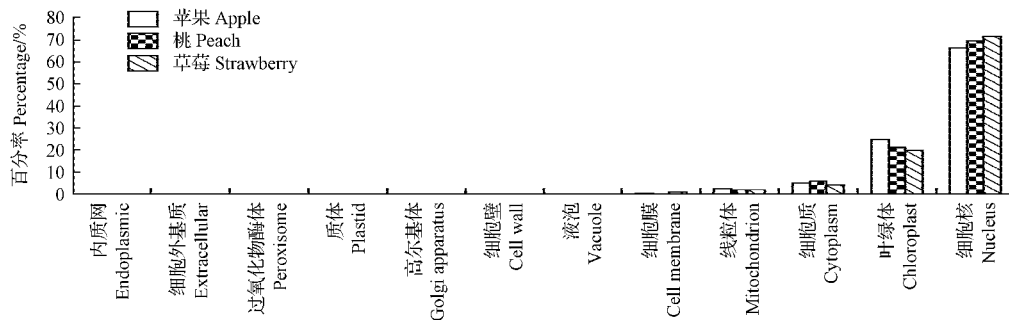


图4 苹果、桃和草莓转录因子的亚细胞定位统计

Fig. 4 Subcellular localization statistics of the transcription factors in apple, peach and strawberry

AP2等,这些家族基因的遗传转化能调节植物的抗旱性。例如,超表达 *DREB1A* 基因可以提高烟草的干旱胁迫耐受性^[27]; MYB 转录因子家族成员 *AtMYB60*、*AtMYB52*、*MYB15*、*AtMYB41*、*AtMYB96* 和 *AtMYB88* 均参与调控拟南芥抗旱响应途径^[28-33]。根据前人研究结果,推测与拟南芥 MYB 转录因子同源的苹果转录因子家族基因可能在干旱逆境响应过程中具有相似的功能,该研究寻找了与上述 MYB 转录因子同源的苹果基因,检测了平邑甜茶和 T337 苹果品种幼苗经 PEG 模拟干旱处理后 6 个基因的表达量变化(表 3、图 5)。

qRT-PCR 结果表明,平邑甜茶幼苗中除 MDP0000230141 和 MDP0000468201 经过干旱处理后基因表达量变化不很明显外,其余 4 个基

因的表达量变化均高于 2.0 倍; T337 幼苗中该 6 个基因表达量的变化趋势相似(图 5)。由此推测,苹果 MYB 转录因子家族成员在调控干旱胁迫信号通路中可能起着较重要的调控作用。

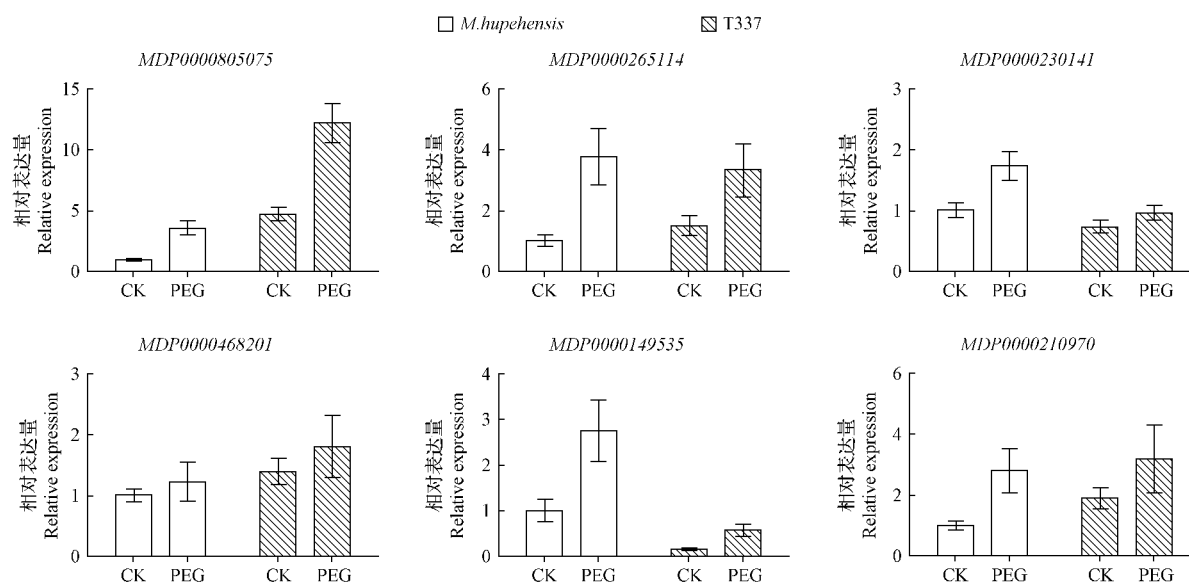
3 讨论

转录因子可以通过与 DNA 或其它蛋白相互作用来激活或抑制调控基因的表达,它们在植物的许多生理生化代谢过程中起着重要作用。研究证明, *AtMYB28*、*AtMYB29* 和 *AtMYB76* 参与调节葡萄糖苷的生物合成, *AtMYB16* 可诱导表皮细胞形态发生^[34-36]; *MdMYB121* 参与调控苹果响应非生物胁迫过程, *MdMYB3* 参与了花青素生

表3 苹果和拟南芥干旱相关的 MYB 转录因子信息

Table 3 Information on MYB transcription factors associated with drought stress in apple and *Arabidopsis*

基因 ID Gene ID	功能 Function	苹果同源基因 Best orthology in apple
AT1G08810	AtMYB60 is specifically expressed in guard cells, and its expression is negatively modulated during drought ^[28]	MDP0000805075
AT1G17950	Overexpression of AtMYB52 confers ABA hypersensitivity and drought tolerance ^[29]	MDP0000265114
AT3G23250	Over-expression of MYB15 improves drought and salt tolerance ^[30]	MDP0000230141
AT4G28110	Over-expression of the <i>Arabidopsis</i> AtMYB41 gene alters cell expansion and leaf surface permeability ^[31]	MDP0000468201
AT5G62470	Over-expression of <i>Arabidopsis</i> MYB96 confers drought resistance in <i>Camelina sativa</i> via cuticular wax accumulation ^[32]	MDP0000149535
AT1G14350	A new function of FLP/MYB88 in sensing and/or transducing abiotic stress ^[33]	MDP0000210970



注:CK 为对照,PEG 为聚乙二醇。

Note:CK is control,PEG is polyethylene glycol.

图5 qRT-PCR 检测干旱条件下平邑甜茶和 T337 中 *MdMYB* 基因的表达分析Fig. 5 Expression analysis of the six identified *MdMYB* genes under PEG conditions in *M. hupehensis* and T337

物合成和花发育过程的调控^[16,37]; *AtWRKY6* 和 *AtWRKY70* 在调节衰老反应时发挥作用^[38-39]。

该研究分别在苹果、桃和草莓全基因组中分别鉴定到 3 039、1 527 和 1 506 个编码转录因子的基因成员,说明转录因子在蔷薇科植物中分布也是十分广泛。该研究结果显示,MYB、AP2、NAC、bHLH、MADS 和 WRKY 转录因子家族是蔷薇科中成员最多的 6 个家族,以上转录因子在拟南芥、水稻和玉米等物种基因组中也是分布较多^[13-15]。

研究表明,MYB 转录因子家族成员在干旱胁迫响应过程中起着十分关键的调控作用。*Md-*

MYB10 可以提高转基因拟南芥对渗透胁迫的耐受性^[40]; *OsMYB3R-2* 可以增加拟南芥耐受冻害、干旱和盐害胁迫^[41];超表达 *AtMYB52* 和 *MYB15* 都能够提高拟南芥对于 ABA 的敏感性和对干旱胁迫的耐受性^[29-30];另外,研究也发现 *AtMYB60*、*AtMYB41*、*AtMYB96* 和 *AtMYB88* 也都参与调节干旱响应过程^[28,31-33]。通过检测干旱处理条件下与上述 MYB 转录因子同源的苹果 *MdMYB* 基因的表达量变化,得知 6 个 *MdMYB* 基因均不同程度地响应 PEG 处理,在平邑甜茶和 T337 苹果品种中表达量上调趋势基本一致,只是具体表达量变化倍数存

在差异(图5)。推测这些 *MdMYB* 基因在对干旱的响应过程中可能起调控作用,今后在苹果抗干旱育种工作中可能具有较大潜力作为新基因资源进行分子育种研究。该研究系统分析了苹果基因组中的转录因子家族,并与蔷薇科其它物种桃和草莓基因组中转录因子家族进行了分析比较,分析了苹果 *MdMYB* 干旱相关转录因子在干旱胁迫条件下的表达情况,研究结果为今后研究蔷薇科植物尤其是苹果中转录因子的功能和作用机理提供了参考依据。

参考文献

- [1] JUNG S, MAIN D. Genomics and bioinformatics resources for translational science in *Rosaceae*[J]. Plant Biotechnology Reports, 2014(8): 49-64.
- [2] SHULAEV V, KORBAN S S, SOSINSKI B, et al. Multiple models for *Rosaceae* genomics[J]. Plant Physiology, 2008, 147: 985-1003.
- [3] SHULAEV V, SARGENT D J, CROWHURST R N, et al. The genome of woodland strawberry(*Fragaria vesca*)[J]. Nature Genetics, 2011, 43: 109-116.
- [4] VELASCO R, ZHARKIKH A, AFFOURTIT J, et al. The genome of the domesticated apple (*Malus × domestica* Borkh.) [J]. Nature Genetics, 2010, 42: 833-839.
- [5] VERDE I, ABBOTT A G, SCALABRIN S, et al. The high-quality draft genome of peach (*Prunus persica*) identifies unique patterns of genetic diversity, domestication and genome evolution [J]. Nature Genetics, 2013, 45: 487-494.
- [6] ALVES M S, DADALTO S P, GONCALVES A B, et al. Plant bZIP transcription factors responsive to pathogens: A review[J]. International Journal of Molecular Sciences, 2013(14): 7815-7828.
- [7] RIECHMANN J L, HEARD J, MARTIN G, et al. *Arabidopsis* transcription factors: Genome-wide comparative analysis among eukaryotes[J]. Science, 2000, 290: 2105-2110.
- [8] ZHANG L, ZHAO G, XIA C, et al. A wheat R2R3-MYB gene, *TaMYB30-B*, improves drought stress tolerance in transgenic *Arabidopsis*[J]. Journal of Experimental Botany, 2012, 63: 5873-5885.
- [9] BAKSHI M, OELMULLER R. WRKY transcription factors: Jack of many trades in plants[J]. Plant Signaling Behavior, 2014(9): e27700.
- [10] DUBOS C, STRACKE R, GROTEWOLD E, et al. MYB transcription factors in *Arabidopsis*[J]. Trends in Plant Science, 2010(15): 573-581.
- [11] RUSHTON D L, TRIPATHI P, RABARA R C, et al. WRKY transcription factors: Key components in abscisic acid signalling[J]. Plant Biotechnology Journal, 2012(10): 2-11.
- [12] PEREZ-RODRIGUEZ P, RIANO-PACHON D M, CORREA L G, et al. PlnTFDB: Updated content and new features of the plant transcription factor database[J]. Nucleic Acids Research, 2010, 38: 822-827.
- [13] GUO A, HE K, LIU D, et al. DATF: A database of *Arabidopsis* transcription factors[J]. Bioinformatics, 2005, 21: 2568-2569.
- [14] GAO G, ZHONG Y, GUO A, et al. DRTF: A database of rice transcription factors[J]. Bioinformatics, 2006 (22): 1286-1287.
- [15] JIANG Y, ZENG B, ZHAO H, et al. Genome-wide transcription factor gene prediction and their expressional tissue-specificities in maize[J]. Journal of Integrative Plant Biology, 2012, 54: 616-630.
- [16] CAO Z H, ZHANG S Z, WANG R K, et al. Genome wide analysis of the apple MYB transcription factor family allows the identification of *MdMYB121* gene conferring abiotic stress tolerance in plants[J]. PLoS One, 2013(8): e69955.
- [17] SU H Y, ZHANG S Z, YUAN X W, et al. Genome-wide analysis and identification of stress-responsive genes of the NAM-ATAF1, 2-CUC2 transcription factor family in apple[J]. Plant Physiology Biochemistry, 2013(1): 11-21.
- [18] WANG X, ZHANG S, SU L, et al. A genome-wide analysis of the LBD (LATERAL ORGAN BOUNDARIES domain) gene family in *Malus domestica* with a functional characterization of *MdLBD11*[J/OL]. PLoS One, 2013(8): e57044.
- [19] ZHAO T, LIANG D, WANG P, et al. Genome-wide analysis and expression profiling of the DREB transcription factor gene family in *Malus* under abiotic stress[J]. Molecular Genetics and Genomics, 2012, 287: 423-436.
- [20] JUNG S, FICKLIN S P, LEE T, et al. The genome database for rosaceae (GDR): Year 10 update[J]. Nucleic Acids Research, 2014, 42: 1237-1244.
- [21] FINN R D, BATEMAN A, CLEMENTS J, et al. Pfam: The protein families database[J]. Nucleic Acids Research, 2014, 42: D222-230.
- [22] LIU R H, MENG J L. MapDraw: A microsoft excel macro for drawing genetic linkage maps based on given genetic linkage data[J]. Hereditas, 2003, 25: 317-321.
- [23] KRZYWINSKI M, SCHEIN J, BIROL I, et al. Circos: An information aesthetic for comparative genomics[J]. Genome Research, 2009(19): 1639-1645.
- [24] CONESA A, GÖTZ S. Blast2GO: A comprehensive suite for functional analysis in plant genomics[J]. International Journal of Plant Genomics, 2008: 1-12.
- [25] YE J, FANG L, ZHENG H, et al. WEGO: A web tool for plotting GO annotations[J]. Nucleic Acids Research, 2006, 34: 293-297.
- [26] LIU Z X, LIU S L, YANG H Q, et al. Using protein granularity to extract the protein sequence features[J]. Journal of Theoretical Biology, 2013, 331: 48-53.

- [27] KASUGA M, MIURA S, SHINOZAKI K, et al. A combination of the *Arabidopsis* DREB1A gene and stress-inducible rd29A promoter improved drought- and low-temperature stress tolerance in tobacco by gene transfer[J]. Plant Cell Physiology, 2004, 45: 346-350.
- [28] COMINELLI E, GALBIATI M, VAVASSEUR A, et al. A guard-cell-specific MYB transcription factor regulates stomatal movements and plant drought tolerance[J]. Curr Biology, 2005, 15(13): 1196-1200.
- [29] PARK M Y, KANG J Y, KIM S Y. Overexpression of AtMYB52 confers ABA hypersensitivity and drought tolerance[J]. Molecular Cells, 2011, 31(5): 447-454.
- [30] DING Z, LI S, AN X, et al. Transgenic expression of MYB15 confers enhanced sensitivity to abscisic acid and improved drought tolerance in *Arabidopsis thaliana*[J]. Journal of Genetics Genomics, 2009, 36(1): 17-29.
- [31] COMINELLI E, SALA T, CALVI D, et al. Over-expression of the *Arabidopsis* AtMYB41 gene alters cell expansion and leaf surface permeability[J]. Plant Journal, 2008, 53(1): 53-64.
- [32] LEE SB, KIM H, KIM R J, et al. Overexpression of *Arabidopsis* MYB96 confers drought resistance in *Camelina sativa* via cuticular wax accumulation[J]. Plant Cell Reports, 2014, 33(9): 1535-1546.
- [33] XIE Z, LI D, WANG L, et al. Role of the stomatal development regulators FLP/MYB88 in abiotic stress responses[J]. Plant Journal, 2010, 64(5): 731-739.
- [34] BAUMANN K, PEREZ-RODRIGUEZ M, BRADLEY D, et al. Control of cell and petal morphogenesis by R2R3 MYB transcription factors[J]. Development, 2007, 134: 1691-1701.
- [35] GIGOLASHVILI T, ENGQVIST M, YATUSEVICH R, et al. HAG2/MYB76 and HAG3/MYB29 exert a specific and coordinated control on the regulation of aliphatic glucosinolate biosynthesis in *Arabidopsis thaliana* [J]. New Phytologist, 2008, 177: 627-642.
- [36] GIGOLASHVILI T, YATUSEVICH R, BERGER B, et al. The R2R3-MYB transcription factor HAG1/MYB28 is a regulator of methionine-derived glucosinolate biosynthesis in *Arabidopsis thaliana* [J]. Plant Journal, 2007, 51: 247-261.
- [37] VIMOLMANGKANG S, HAN Y, WEI G, et al. An apple MYB transcription factor, MdMYB3, is involved in regulation of anthocyanin biosynthesis and flower development[J]. BMC Plant Biology, 2013(13): 176.
- [38] ULKER B, SHAHID MUKHTAR M, SOMSSICH I E. The WRKY70 transcription factor of *Arabidopsis* influences both the plant senescence and defense signaling pathways[J]. Planta, 2007, 226: 125-137.
- [39] EULGEM T, RUSHTON P J, ROBATZEK S, et al. The WRKY superfamily of plant transcription factors[J]. Trends in Plant Science, 2000(5): 199-206.
- [40] GAO J J, ZHANG Z, PENG R H, et al. Forced expression of *Mdmyb10*, a myb transcription factor gene from apple, enhances tolerance to osmotic stress in transgenic *Arabidopsis* [J]. Molecular Biology Reports, 2011, 38: 205-211.
- [41] DAI X, XU Y, MA Q, et al. Overexpression of an R1R2R3 MYB gene, *OsMYB3R-2*, increases tolerance to freezing, drought, and salt stress in transgenic *Arabidopsis* [J]. Plant Physiology, 2007, 143: 1739-1751.

Identification and Analysis of Transcription Factors in Apple Genome

XU Ruirui, GAO Minggang, LIU Chunxiang

(College of Biological and Agricultural Engineering, Weifang University/Key Laboratory of Biochemistry and Molecular Biology in Universities of Shandong, Weifang, Shandong 261061)

Abstract: Transcription factors (TFs) are essential regulators of gene expression. To better understand TFs encoding genes in *Rosaceae*, especially in apple, a silico genomic analysis of TFs prediction was performed using apple genomes, as well as compared with peach and strawberry. The results showed that a total of 3 039, 1 527 and 1 506 TFs genes were identified in apple, peach and strawberry, respectively, which could be classified into 58 families. The genomic location showed that the predicted TFs were distributed across all chromosomes with different densities. Meanwhile, all the TFs had the similar GO terms and subcellular localization information. Then, the drought-related transcription factors in apple homologous with *AtMYBs* were predicted and detected in *Arabidopsis*. Encouragingly, six of the selected genes were found to respond to PEG treatments in *M. hupehensis* and T337 with the similar trends. To the best of our knowledge, this was the first report of a genome-wide analysis of the TFs gene family in three rosaceous plants. This study provided valuable information for understanding the classification and putative functions of the TFs in rosaceous plant.

Keywords: apple; genome; transcription factor; identification and analysis; drought