

中国野生种葡萄 *mybA* 转录因子 SNP 特征分析

焦 健, 刘崇怀, 樊秀彩, 张 颖, 孙海生, 姜建福, 李 民

(中国农业科学院郑州果树研究所, 郑州 450009)

摘要: 以中国野生葡萄 14 个种为材料, 对控制花色苷合成的 *mybA* 转录因子进行克隆和序列分析, 获得 *VvmybA1* 和 *VlmybA2* 两个转录因子的全长基因序列, 共检测到 121 个 SNP, 表现出丰富的遗传多样性。3 种中性检测方法比较序列变异模式, 结果表明, 中国野生种葡萄 *VvmybA1* 和 *VlmybA2* 基因没有偏离中性模型, 反映出基因漂移和选择性中性突变之间的平衡。不同野生种材料的 *mybA* 基因结构存在很高的同源相似性。但是在启动子区、内含子区以及第 3 个外显子区存在不同程度碱基的缺失、插入和替换, 而且野生种葡萄 *mybA* 基因存在一些特有序列或突变, 这些突变可以作为分子标记区分不同的野生种材料。通过基因结构比对和系统进化树分析, 可将野生种葡萄细分为 5 个类群。初步推测桦叶葡萄和变叶葡萄进化地位较为原始。

关键词: 中国野生葡萄; *mybA*; 花色苷; SNP; 遗传多样性

Characterization of SNP Associated with *mybA* Transcription Factor in Chinese Wild Grapes

JIAO Jian, LIU Chong-huai, FAN Xiu-cai, ZHANG Ying, SUN Hai-sheng, JIANG Jian-fu, LI Min

(Zhengzhou Fruit Research Institute, Chinese Academy of Agricultural Sciences, Zhengzhou 450009)

Abstract: The transcription factor corresponding to the *mybA* that regulates anthocyanin biosynthesis was cloned and analyzed on a sample of 14 Chinese wild grapes species. Two complete *mybA*-related gene sequences, *VvmybA1* and *VlmybA2* were obtained, and a total of 121 SNPs were detected in the sequenced fragments which give a high level of polymorphism. Three neutrality tests had been used to compare the patterns of sequence variations, the result indicates a non departure from neutrality expectations for gene *VvmybA1* and *VlmybA2* of Chinese wild grape which reflected equilibrium between genetic drift and selectively neutral mutation. The *mybA* gene acquired form different wild accessions represented a highly homology. Yet extensive base deletion, insertion and substitution existed in promoter region, intron region and the third coding exon, which exhibited rich genetic diversity between species of Chinese wild grapes. Furthermore, there were several unique bases or mutations in the *mybA* gene of wild grapes that might be selected to be the molecular marker to distinguish themselves from others well. Comparing and analysing the gene structure and phylogenetic tree of the gene sequence, the studied Chinese wild grape species were divided into five groups. Moreover, *V. betulifolia* and *V. hancockii* were initially predicted to be primitive in the evolution.

Key words: Chinese wild grape; *mybA*; anthocyanin; SNP; genetic diversity

转录因子是能够与真核生物基因的顺式作用元件发生特异性结合, 并对转录有激活或抑制作用的 DNA 结合蛋白^[1]。根据 DNA 结合结构域重复片段 (R) 的数量, 可将 *Myb* 类转录因子分为单一 *Myb* 结构域蛋白 (R1/R2)、2R 蛋白 (R2R3) 和 3R 蛋白

(R1R2R3)^[2]。葡萄果皮存在白色、粉红、红色以及黑色等丰富的颜色表现型, 这种颜色差异是由果皮细胞内合成与积累的花色苷种类与数量引起的^[3-5], 而葡萄花色苷合成代谢途径中重要的调节基因 *mybA* 类转录因子, 如 *VvmybA1* 和 *VvmybA2*^[6] 都

收稿日期: 2012-12-31 修回日期: 2013-05-20 网络出版日期: 2013-08-09

URL: <http://www.cnki.net/kcms/detail/11.4996.S.20130809.1447.030.html>

基金项目: 现代农业产业技术体系建设专项资金 (CARS-30-yz-1); 农业部作物种质资源保护项目 (NB2012-2130135-34)

第一作者主要从事葡萄种质资源研究。E-mail: sdneyj2008@163.com

通信作者: 刘崇怀, 研究方向为葡萄种质资源。E-mail: liuchonghuai@caas.net.cn

是 R2R3 基因家族成员,它们通过调节 *UFGT* 的表达参与花色苷的代谢合成^[7]。

在基因组水平上,由单个碱基变异所引起的 DNA 序列多态性被称为单核苷酸多态性(SNP)^[8]。通过对种间不同样品的 DNA 进行测序比对,结果往往显示出丰富的 SNP 位点,并且在一些特殊位点存在不同程度的碱基插入和缺失。如果一种 SNP 普遍存在于一个种内(频率至少为 1%),它就可以作为一个分子标记,为研究种内亲缘关系提供依据^[9]。过去的十几年里,植物基因组 SNP 出现的频率和特性逐渐受到重视,在拟南芥^[10]、大豆^[11]、玉米^[12]和小麦^[13]等植物中,越来越多的研究通过 SNP 位点的差异分析基因组的多样性。葡萄作为世界上重要的经济作物之一,被广泛栽培并用于鲜食和酿酒,因此对其基因组的研究,可为认识和改良葡萄品种提供更多有用信息。就欧亚种葡萄而言,SNP 分析主要用于亲缘分析^[14]、基因连锁图绘制^[15]、关联分析^[16]以及探索基因的起源和进化^[17]。中国是葡萄属植物的主要起源地之一,也是世界葡萄属植物种类最多、遗传资源最为丰富的国家之一^[18]。原产我国的山葡萄、毛葡萄、桑叶葡萄、

刺葡萄等野生葡萄,在我国葡萄科技工作者收集筛选的基础上,选择出一些优良单株或品系直接应用于生产,形成了一定规模的鲜食和酿酒产业。以前针对中国野生葡萄亲缘关系及遗传多样性的研究主要采用形态学、孢粉学、细胞学、生物化学等方法^[19],近几年一系列分子标记技术,如 ISSR^[20]、SSR^[21-23]以及 RAPD^[24-25]等以快速、准确、灵敏度高等特点,逐渐取代传统的分类鉴定手段。如今,单核苷酸多态性(SNP)在欧亚种葡萄以及传统作物如大豆、玉米和小麦等植物中研究比较深入,但针对中国野生葡萄的相关研究较少。本研究主要针对葡萄花色苷合成途径的关键转录因子 *mybA* 基因进行克隆测序,利用 SNP 分析技术评价和描述中国野生种葡萄 14 个种的单核苷酸多态性以及种间的遗传进化关系。

1 材料与方法

1.1 试验材料

中国野生葡萄材料来源于中国农业科学院郑州果树研究所国家果树种质郑州葡萄圃,共 14 份种质(表 1)。采摘幼嫩叶片立即用液氮处理,保存于 -80 ℃ 条件下备用。

表 1 供试的样品以及相关特性

Table 1 Experimental samples and association attributes

编号 Code	种名 Species	品种(株系) Cultivar(Clones)	果皮颜色 Skin color	性别类型 Sex type	编号 Code	种名 Species	品种(株系) Cultivar(Clones)	果皮颜色 Skin color	性别类型 Sex type
1	菱叶葡萄 <i>V. hancockii</i>	菱叶 0945	黑	♀	8	毛葡萄 <i>V. heyneana</i>	都安毛葡萄	黑	♀
2	桑叶葡萄 <i>V. heyneana</i> subsp. <i>ficiifolia</i>	九里沟桑叶	黑	♀	9	网脉葡萄 <i>V. reticulata</i>	卢氏网脉	—	♂
3	刺葡萄 <i>V. davidii</i>	刺葡萄 940	黑	♀	10	秋葡萄 <i>V. romaneti</i>	青要山秋	—	♂
4	蓼蓂 <i>V. bryoniaefolia</i>	多裂叶蓼蓂	黑	♀	11	腺枝葡萄 <i>V. adenoclada</i>	双溪-1	黑	♀
5	华东葡萄 <i>V. pseudoreticulata</i>	万县葡萄	黑	♀	12	葛藟 <i>V. flexuosa</i>	葛藟 0943	—	♂
6	云南葡萄 <i>V. yunnanensis</i>	云南葡萄	—	♂	13	美丽葡萄 <i>V. bellula</i>	美丽 1104	—	♂
7	变叶葡萄 <i>V. piasezkii</i>	灵宝变叶	—	♂	14	桦叶葡萄 <i>V. betulifolia</i>	嵩县桦叶	—	♂

1.2 DNA 提取和质量检测

采用改良的 CTAB 法^[26]提取叶片基因组 DNA,用 1% 琼脂糖凝胶电泳检测 DNA 质量,采用核酸蛋白测定仪(Eppendorf)检测浓度后稀释至 40 ng/μL, -80 ℃ 冰箱保存备用。为了确保后续扩增的准确和真实性,首先采用 β-微管蛋白(β-tubulin, TUB)对 DNA 样品 PCR 条件进行优化。电泳结果显示扩增条带清晰,表明提取后稀释的 DNA 纯度和浓度都满

足 PCR 反应要求。

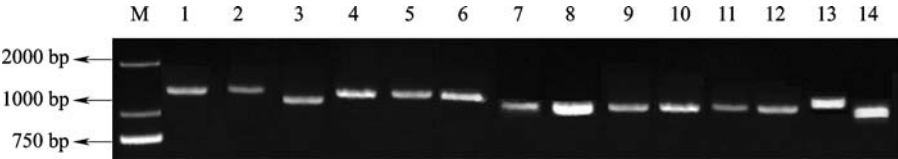
1.3 中国野生葡萄 *mybA* 基因克隆

根据已报道的 *VvmybA1* 序列(GeneBank 登录号 AB242302),分别从启动子区域和功能区末端序列设计上下游引物,扩增 *VvmybA1* 基因的上游引物为 PF(5'-GGACGTTAAAAAATGTTGCACGTG-3'),下游引物为 M(5'-CACACAAGTAAATCATTTGATCTGA-3'),引物均由上海生工合成。

PCR 反应体系为 20 μL : 10 \times Buffer 2.0 μL , 2.5 mmol/L dNTPs 1.6 μL , 25 mmol/L 的 Mg^{2+} 1.2 μL , DNA 模板 2.0 μL , *ExTaq* 酶(5 U/ μL)0.1 μL , 上下游引物(10 mmol/L)各 1.0 μL , 双蒸水 11.1 μL 。PCR 反应条件: 94 $^{\circ}\text{C}$ 预变性 3 min; 然后 94 $^{\circ}\text{C}$ 变性 30 s, 61 $^{\circ}\text{C}$ 退火 30 s, 72 $^{\circ}\text{C}$ 延伸 90 s, 共 34 个循环; 最后 72 $^{\circ}\text{C}$ 延伸 10 min, 4 $^{\circ}\text{C}$ 结束。扩增产物用 GoldViewTM 核酸染料染色后, 1.5% 琼脂糖凝胶中电泳 30 ~ 50 min, 最后用凝胶成像系统拍照。切胶回收纯化后载入 pMD18-T 载体, 16 $^{\circ}\text{C}$ 连接 3 h 后转化导入 DH5 α 感受态细胞中, 接种在含氨苄青霉素的 LB 培养基上培养 12 ~ 16 h。挑取白斑筛选活化, 摇菌 8 h 后进行菌液 PCR 检测, 将菌液送至上海生工, 使用 M13 通用引物双向测序, 每个菌液样品测 3 次确保序列正确。

1.4 *mybA* 基因序列分析

测序结果用 DNAMAN 4.0 生物软件进行拼接



M : DL2000; 数字对应表 1 中的材料编号 Figures correspond with the code of Table 1

图 1 不同材料 *mybA* 基因的 PCR 产物

Fig. 1 PCR product of *mybA* gene from different samples

2.2 不同材料 *mybA* 基因序列分析

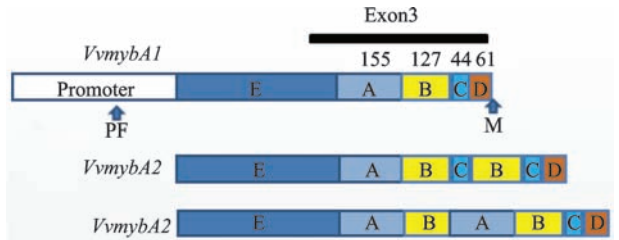
测序所得 14 个基因序列已提交 GeneBank, 登录号为 KC342651 ~ KC342664。NCBI 在线比对显示, 变叶葡萄、桦叶葡萄、菱叶葡萄、刺葡萄、网脉葡萄、腺枝葡萄、秋葡萄、美丽葡萄和葛藟扩增出的序列为 *VvmybA1*, 云南葡萄、桑叶葡萄、蓼莪、华东葡萄和毛葡萄扩增出的是 *VlmybA2*。 *VlmybA2* 等位基因最早在巨峰葡萄中得到, 瞬时表达结果显示此基因有调控花色苷合成代谢的功能^[29]。 *VlmybA2* (GeneBank 登录号 AB073013)、 *VvmybA1* (GeneBank 登录号 AB242302) 和 *VvmybA2* (GeneBank 登录号 DQ886419) 利用 MEGA 比对绘制基因结构图 (图 2), 序列差异主要在第 3 个外显子区。 *VlmybA2* 与 *VvmybA1* 相比, 在第 3 个外显子区多出 1 个 B 和 C 的重复区域, 大小为 171 bp。 *VvmybA2* 与 *VvmybA1* 相比, 在第 3 个外显子区多出 1 个 A 和 B 的重复区域, 大小为 282 bp。上述结果显示, 不同种质材料的扩增产物只有 1 种, 而且 *VvmybA1* 和 *VlmybA2* 是由相同的引物 PF/M 扩增得到, 但是由于 *VvmybA1* 和 *VlmybA2* 基因差异很大, 在此将扩增出 *VvmybA1* 基因的材料归于组 A, 扩增出 *VlmybA2* 基因材料归于组 B, 单独进行分析。

得到目的序列, 并与 GenBank 已提交的 *VvmybA1* 基因比对分析进行确定。采用 MEGA 5.05 软件对不同材料的 *mybA* 基因构建进化树和序列比对。利用 SNIPlay 在线工具 (<http://snipplay.cirad.fr/>)^[27] 对序列 SNP 出现的频率、突变位点以及多态性进行分析。单倍型多样性 (Hd)、核苷酸多样性 (Pi) 以及中性检测^[28] 采用 DnaSP 5.0 软件进行分析。

2 结果与分析

2.1 不同材料 *mybA* 基因扩增产物电泳分析

上游引物 PF 位于 *VvmybA1* 的基因启动子区域, 下游引物 M 位于 *VvmybA1* 的基因终止子末端。利用 PF/M 引物扩增不同种质材料 *mybA* 的基因全长。电泳结果显示, 不同种质材料扩增片段大小不同, 大小在 1200 ~ 1500 bp 之间, 这说明不同种质材料扩增出的 *mybA* 基因存在碱基序列的差异。



数字代表序列大小; 相同颜色区域表示相同的序列片段; 序列中的 SNP 未标示; 箭头位置为引物位点
Figures stand for the size of sequence. The same colored box indicates the same gene fragment. Single nucleotide polymorphisms within sequence are not shown.

The arrows indicate locus of primers

图 2 *mybA* 基因对比结构图

Fig. 2 Comparison of *mybA* gene structure

2.2.1 *VvmybA1* 和 *VlmybA2* 序列 SNP 特征分析

变叶葡萄、桦叶葡萄等 9 个野生种质材料克隆得到了 *VvmybA1*, 长度在 1266 ~ 1422 bp 之间。测序得到包括启动子、2 个内含子以及 3 个外显子在内的基因全长序列。总共检测出 84 个 SNP 位点, 其中启动子区域 36 个、内含子区域 12 个、外显子区域 36 个。在外显子 SNP 当中, 有 29 处碱基突变使得编码的氨基酸发生了改变 (表 2)。

表 2 *VvmybA1* 基因序列 SNP 特征

Table 2 Characteristics of detected SNP in *VvmybA1*

编号	突变	SNP 频率 (%)	多数等位基因	少数等位基因	染色体位点	密码子位点	密码子	氨基酸	同义突变
Code	Variation	SNP frequency	Majority allele	Minority allele	Position on chromosome	Position in codon	Codon	Amino acid	Synonymous
1	[A/G]	22.2	A	G	14240778	2	aAg/aGg	K/R	NO
2	[T/C]	22.2	T	C	14240766	2	aTc/aCc	I/T	NO
3	[C/G]	22.2	G	C	14240729	3	gaG/gaC	E/D	NO
4	[A/C]	11.1	A	C	14240728	1	Aaa/Caa	K/Q	NO
5	[A/T]	11.1	A	T	14240593	1	Aat/Tat	N/Y	NO
6	[T/C]	11.1	C	T	14240558	3	ctC/ctT	L/L	YES
7	[A/G]	22.2	G	A	14240529	2	gGa/gAa	G/E	NO
8	[T/C]	11.1	T	C	14240338	2	aTt/aCt	I/T	NO
9	[A/C]	11.1	A	C	14240318	1	Agg/Cgg	R/R	YES
10	[T/C]	11.1	T	C	14240291	1	Tgg/Cgg	W/R	NO
11	[A/G]	22.2	A	G	14240285	1	Agt/Ggt	S/G	NO
12	[C/G]	44.4	C	G	14240274	3	ttC/ttG	F/L	NO
13	[A/G]	22.2	A	G	14240267	1	Gag/Aag	E/K	NO
14	[C/G]	11.1	G	C	14240259	3	caG/caC	Q/H	NO
15	[T/C]	11.1	T	C	14240220	3	tcT/tcC	S/S	YES
16	[T/G]	22.2	G	T	14240166	3	agG/agT	R/S	NO
17	[T/C]	11.1	T	C	14240165	1	Ttt/Ctt	F/L	NO
18	[T/C]	22.2	T	C	14240143	2	gTg/gCg	V/A	NO
19	[A/C]	22.2	A	C	14240121	3	gtC/gtA	V/V	YES
20	[A/G]	11.1	A	G	14240108	1	Aag/Gag	K/E	NO
21	[A/G]	11.1	A	G	14240100	3	tcA/tcG	S/S	YES
22	[T/C/G]	11.1	C	G	14240079	3	aaT/aaC	N/N	NO
23	[T/C]	11.1	C	T	14240055	3	agC/agT	S/S	YES
24	[T/C]	11.1	C	T	14240047	2	gCt/gTt	A/V	NO
25	[A/G]	11.1	G	A	14240009	1	Gct/Act	A/T	NO
26	[A/G]	22.2	G	A	14239997	1	Atg/Gtg	M/V	NO
27	[A/T]	22.2	T	A	14239979	1	Agg/Tgg	R/W	NO
28	[A/G]	11.1	A	G	14239961	1	Aca/Gca	T/A	NO
29	[A/C]	22.2	C	A	14239954	2	aAa/aCa	K/T	NO
30	[A/G]	22.2	G	A	14239931	1	Att/Gtt	I/V	NO
31	[A/T]	22.2	A	T	14239929	3	atT/atA	I/I	YES
32	[A/G]	11.1	A	G	14239914	3	ggA/ggG	G/G	YES
33	[T/C]	11.1	T	C	14239880	1	Tgg/Cgg	W/R	NO
34	[A/T]	44.4	T	A	14239874	1	Aca/Tca	T/S	NO
35	[T/C]	44.4	T	C	14239873	2	aCa/aTa	T/I	NO
36	[A/T]	44.4	T	A	14239872	3	acA/acT	T/T	YES

在云南葡萄等 5 个野生种中得到的是 *VlmybA2* 基因,长度在 1471 ~ 1484 bp 之间。测序结果同样得到的是包括启动子、2 个内含子和 3 个外显子在内的基因全长序列。总共检测出 37 个

SNP 位点、其中启动子区域 6 个、内含子区域 12 个、外显子区域 19 个。在外显子 SNP 当中,有 11 处碱基突变导致氨基酸序列改变(表 3)。

表 3 *VlmybA2* 基因序列 SNP 特征

Table 3 Characteristics of detected SNP in *VlmybA2*

编号	突变	SNP 频率 (%)	多数等位基因	少数等位基因	染色体位点	密码子位点	密码子	氨基酸	同义突变
Code	Variation	SNP frequency	Majority allele	Minority allele	Position on chromosome	Position in codon	Codon	Amino acid	Synonymous
1	[A/G]	20.0	G	A	14180707	2	gCt/gAt	G/D	NO
2	[A/G]	20.0	A	G	14180659	2	aAa/aGa	K/R	NO
3	[T/C]	20.0	C	T	14180628	3	ccC/ccT	P/P	NO
4	[T/C]	20.0	T	C	14180262	3	ggT/ggC	G/G	YES
5	[T/C]	20.0	T	C	14180223	3	taT/taC	Y/Y	YES
6	[T/C]	20.0	T	C	14180214	3	ggT/ggC	G/G	YES
7	[A/G]	40.0	A	G	14180171	1	Aaa/Gaa	K/E	NO

表 3(续)

编号 Code	突变 Variation	SNP 频率 (%) SNP frequency	多数等位基因 Majority allele	少数等位基因 Minority allele	染色体位点 Position on chromosome	密码子位点 Position in codon	密码子 Codon	氨基酸 Amino acid	同义突变 Synonymous
8	[A/G]	20.0	A	G	14180161	2	cAa/cGa	Q/R	NO
9	[C/G]	20.0	C	G	14180112	3	tcC/tcG	S/S	YES
10	[A/G]	40.0	G	A	14180098	2	aGg/aAg	R/K	NO
11	[A/C]	20.0	A	C	14180086	2	aAa/aCa	K/T	NO
12	[A/G]	20.0	A	G	14180016	3	caA/caG	Q/Q	YES
13	[T/G]	20.0	G	T	14180013	3	ccG/ccT	P/P	YES
14	[T/C]	20.0	T	C	14180010	3	aaT/aaC	N/N	YES
15	[A/G]	20.0	A	G	14180000	1	Ata/Gta	I/V	NO
16	[A/G]	20.0	G	A	14179941	3	tcG/tcA	S/S	YES
17	[C/G]	40.0	G	C	14179929	3	gaG/gaC	E/D	NO
18	[A/G]	20.0	A	G	14179928	1	Atg/Gtg	M/V	NO
19	[A/T]	20.0	A	T	14179910	1	Tgg/Agg	W/R	NO

2.2.2 基因的遗传多样性分析和中性检测

衡量种群遗传多样性(变异程度)的重要参数是核苷酸多样性度(Pi)和单倍型多样性度(Hd)。Pi 值

和 Hd 值越大,种群多样性程度越高,遗传多样性就越丰富。将组 A 和组 B 的基因序列采用 DnaSP 5.0 软件计算 Pi 值和 Hd 值,结果如表 4 所示。

表 4 供试样品的基因多态性

Table 4 Genetic diversity for the studied samples

<i>VvmybA1</i>					<i>VlmybA2</i>				
	NPS	NH	Pi	Hd		NPS	NH	Pi	Hd
组 A	68	9	0.01743(±0.02018)	0.00274(±0.052)	组 B	35	5	0.01006±(0.01142)	0.016±(0.126)

NPS:多态位点数量;NH:单倍型数量;Pi:核苷酸多样性度;Hd:单倍型多样性度

NPS: Number of polymorphic sites, NH: Number of haplotypes, Pi: Nucleotide genetic diversity, Hd: Haplotype genetic diversity

利用 DnaSP 5.0 软件,通过 3 种中性检测方法(Tajima’s D, Fu and Li’s D*, Fu and Li’s F*),评价自然选择和中性选择对 *VvmybA1* 和 *VlmybA2* 基因位点进化的影响。结果显示,检测值为负且不显著,表明供试的中国野生种葡萄 *VvmybA1* 和

VlmybA2 基因没有偏离中性模型,反映出基因漂移和选择性中性突变之间的平衡,从而推测供试的野生葡萄种质资源群体在过去较近的历史时期内没有经历种群扩张事件,群体大小保持相对稳定(表 5)。

表 5 供试样品 *VvmybA1* 和 *VlmybA2* 基因中性检测

Table 5 Neutrality tests for the genes *VvmybA1* and *VlmybA2* based on the studied samples

<i>VvmybA1</i>				<i>VlmybA2</i>			
	Tajima’s D	Fu and Li’s D*	Fu and Li’s F*		Tajima’s D	Fu and Li’s D*	Fu and Li’s F*
组 A	-0.69919 (NS, $P>0.10$)	-0.57570 (NS, $P>0.10$)	-0.68043 (NS, $P>0.10$)	组 B	-0.89303 (NS, $P>0.10$)	-0.89303 (NS, $P>0.10$)	-0.96408 (NS, $P>0.10$)

2.2.3 基因结构和系统进化树分析

将所得到的 *mybA* 序列与黑比诺 *VvmybA1* (GeneBank 登录号 AB242302) 基因进行进化树分析和基因结构比对(图 3),14 个中国野生种可细分为 A、B、C、D 和 E 5 组,基因结构分为 6 种。A 组的基因结构主要是在第 3 个外显子区域存在 171 bp 的插入,得到的是 *VlmybA2* 基因,其中云南葡萄在第 1 个内含子区域存在 10 bp 的缺失。B、C、D 和 E 组种质材料得到的序列是 *VvmybA1* 基因,基因结构的主

要差异在于启动子区域是否存在 43 bp 和 111 bp 的插入。另外,网脉葡萄和刺葡萄在第 2 个内含子区域相同位置存在 22 bp 的缺失。通过对 *VvmybA1* 基因构建进化树分析,结果显示黑比诺 *VvmybA1* 基因与中国野生种桦叶葡萄、变叶葡萄的 *VvmybA1* 基因关系较近,并且在启动子区域都没有 23 bp、19 bp 和 111 bp 的插入,表现出比较原始的基因类型。黑比诺是欧亚种葡萄最古老的品种之一^[30],这或许说明桦叶葡萄、变叶葡萄是相对古老的中国野生种葡萄。

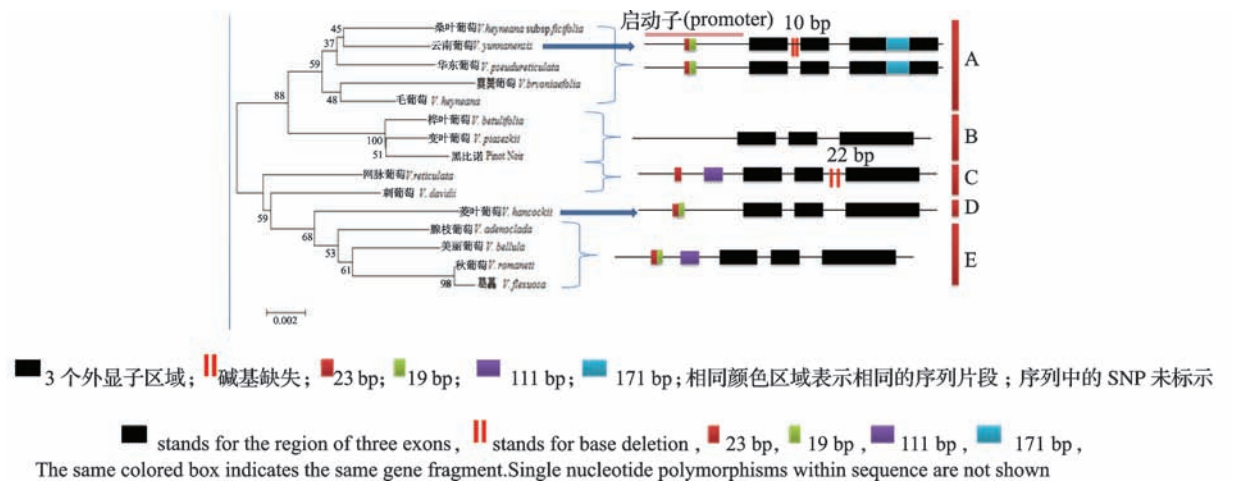


图 3 不同种质材料 *mybA* 基因的结构和系统进化树分析

Fig. 3 Phylogenetic tree and gene struction of *mybA* gene form difernent samples

2.3 不同材料特征序列和 SNP 标记筛选

通过序列比对和 SNP 分析,结果显示,不同野生种质材料在 *mybA* 基因结构上存在很高的相似性。与欧亚种葡萄相比,中国野生种葡萄之间 *VvmybA1* 和 *VlmybA2* 基因在启动子区、内含子区以及第 3 个外显子区存在不同程度碱基的缺失、插入和替换,表现出丰富的遗传多样性。不同野生种质的 *VvmybA1* 和 *VlmybA2* 基因存在一些特有序列或突变,这些碱基序列差异可以很好地将自身与其他种区分开:桑叶葡萄在第 3 个外显子区有 GAC 碱基插入;云南葡萄在第 1 个内含子区有 ACTTTGTCTG 的缺失;刺葡萄与网脉葡萄在第 2 个内含子存在 22 bp 的缺失,所不同的是刺葡萄在启动子区又存在 21 bp 的缺失;菱叶葡萄在第 1 个内含子区有 CATG 的缺失。其他种质材料 *mybA* 基因序列虽然没有出现缺失和插入,但存在特异性 SNP 位点,如表 2 和表 3 中变异频率在 11.1% 和 20.0% 的变异位点都可以为 SNP 标记的筛选作参考。

3 讨论

本试验在 *VvmybA1* 基因中共检测到 84 个 SNP 位点,平均每 17 bp 存在 1 个 SNP, *VlmybA2* 基因共检测出 37 个 SNP 位点,平均每 43 bp 存在 1 个 SNP。D. Lijavetzky 等^[31]通过对欧亚种酿酒葡萄黑比诺测序显示,每 64 bp 存在 1 个 SNP 位点,而 R. Velasco 等^[32]研究表明黑比诺基因组每 100 bp 存在 1 个 SNP 位点。这或许是由于中国野生葡萄绝大多数为雌雄异株,长期的种内杂交繁殖使得其基因组产生更多的 SNP 位点,因为在以往的研究中显示,异型杂交的玉米比大豆、大麦和小麦等自交的植物显

示出更丰富的 SNP 位点^[33]。

检测到的 SNP 位点中,36 个(大约 42%)存在于 *VvmybA1* 编码区,19 个(大约 51%)存在于 *VlmybA2* 编码区。以往的研究表明,非编码区 SNP 位点明显要多于编码区^[34],然而有些研究结果与之相反^[31]。本研究显示,SNP 位点在编码区和非编码区之间出现的几率差异并不显著,这或许是因为中国野生种葡萄在进化和传播过程中,自然选择和有性繁殖占主导地位,欧亚种葡萄的人工选择和无性繁殖使得不同品种间的基因功能区序列相对稳定。在编码区突变位点中,71% 的突变引起了氨基酸序列的变化,甚至夔夔的 *VlmybA2* 基因在编码区存在 3 碱基 GAC 的插入,导致一个额外的氨基酸的插入。过去的研究显示,*mybA* 基因的一些位点突变会影响其功能表达,如欧亚种葡萄 *VvmybA1* 启动子区域逆转录转座子 *Gert1* 的插入^[35],以及 *VvmybA2* 编码区碱基的突变^[36],都抑制了其调控功能。中国野生葡萄果实为黑色,果皮细胞合成和积累大量的花色苷,尽管上述试验显示, *VvmybA1* 和 *VlmybA2* 编码区存在许多非同义突变,但是这些突变并没有影响到野生葡萄花色苷的合成与积累。或许这些氨基酸的变化并没有影响到转录因子的功能,或许花色苷合成的调控有更多未知的潜在机制。

D. Lijavetzky 等^[31]研究显示,欧亚种葡萄基因组核苷酸多样性 Pi 值为 0.0051, L. Riahi 等^[33]研究突尼斯地区的野生种和欧亚种葡萄 *mybA* 基因中 Pi 值在 0.001 ~ 0.009 之间。本试验对中国野生葡萄的研究显示, *VvmybA1* 基因 Pi 值为 0.01743 (± 0.02018), *VvmybA2* 基因 Pi 值为 0.01006 ± (0.01142),数值明显大于以往的研究报道。这说明雌雄异株的中国野生

葡萄在自然杂交过程中会产生更多的变异位点,表现在更高的遗传多样性。

本研究获得的序列是由 PF/M 引物扩增出,特异性引物按照 *VvmybA1* 的序列设计,但是在云南葡萄、桑叶葡萄、蓼萼、华东葡萄和毛葡萄 5 个野生种基因组中扩增出条带单一的 *VlmybA2* 基因。*VlmybA2* 最早在欧美杂交种巨峰中发现^[29],本研究中 *VlmybA2* 比 *VvmybA1* 多出 B 和 C 的重叠序列,共 171 bp。在进化过程中,从序列结构上, *VlmybA2* 可能是由 *VvmybA1* 的 B、C 序列复制后插入形成的。但为何同一对引物在其他 9 个野生种葡萄中扩增得到的是 *VvmybA1* 基因,而在云南葡萄等 5 个野生种葡萄得到的是 *VlmybA2* 基因,或许这 5 个野生种葡萄基因组内只存在 *VlmybA2* 基因,也可能 *VvmybA1* 的基因位点发生了改变,这些猜测需要后续试验进行验证。

参考文献

[1] Jin H, Martin C. Multifunctionality and diversity within the plant MYB-gene family [J]. Plant Molecular Biology, 1999, 41: 577-585

[2] 陈清,汤浩茹,董晓莉,等. 植物 Myb 转录因子的研究进展[J]. 基因组学与应用生物学,2009,28(2):365-372

[3] This P, Lacombe T, Cadle-Davidson M, et al. Wine grape (*Vitis vinifera* L.) color associates with allelic variation in the domestication gene *VvmybA1* [J]. Theor Appl Genet, 2007, 114: 723-730

[4] Azuma A, Kobayashi S, Mitani N, et al. Genomic and genetic analysis of *Myb*-related genes that regulate anthocyanin biosynthesis in grape berry skin [J]. Theor Appl Genet, 2008, 117: 1009-1019

[5] Ramazzotti S. Expression of genes associated with anthocyanin synthesis in red-purplish, pink, pinkish-green and green grape berries from mutated 'sangiovese' biotypes; A case study [J]. Vitis, 2008, 47(3): 147-151

[6] Pelsy F. Molecular and cellular mechanisms of diversity within grapevine varieties [J]. Heredity, 2009, 104: 331-340

[7] Kobayashi S, Ishimaru M, Ding C K, et al. Comparison of UDP-glucose: flavonoid 3-O-glucosyltransferase (UFGT) gene sequences between white grapes (*Vitis vinifera*) and their sports with red skin [J]. Plant Sci, 2001, 160: 543-550

[8] Taillon-Miller P, Gu Z, Li Q, et al. Overlapping genomic sequences: a treasure trove of single-nucleotide polymorphisms [J]. Genome Res, 1998, 8: 748-754

[9] Brookes A J. The essence of SNPs [J]. Gene, 1999, 234: 177-186

[10] Aguadé M. Nucleotide sequence variation at two genes of the phenylpropanoid pathway, the FAH1 and F3H genes, in *Arabidopsis thaliana* [J]. Mole Biol Evo, 2001, 18: 1-9

[11] Zhu Y L, Song Q J, Hyten D L, et al. Single-nucleotide polymorphisms in soybean [J]. Genetics, 2003, 163: 1123-1134

[12] Batley J. Mining for single nucleotide polymorphisms and insertions/deletions in maize expressed sequence tag data [J]. Plant Physiol, 2003, 132: 84-91

[13] Feltus F A, Wan J, Schulze S R, et al. An SNP resource for rice genetics and breeding based on subspecies indica and japonica genome alignments [J]. Genome Res, 2004, 14: 1812-1819

[14] Hunt H V, Lawes M C, Bower M A, et al. A banned variety was

the mother of several major wine grapes [J]. Biol Lett, 2010, 6: 367-369

[15] Vezzulli S, Troggio M, Coppola G, et al. A reference integrated map for cultivated grapevine (*Vitis vinifera* L.) from three crosses, based on 283 SSR and 501 SNP-based markers [J]. Theor Appl Genet, 2008, 117: 499-511

[16] Fournier-Level A, Le Cunff L, Gomez C, et al. Quantitative genetic bases of anthocyanin variation in grape (*Vitis vinifera* L. ssp. *sativa*) berry: a quantitative trait locus to quantitative trait nucleotide integrated study [J]. Genetics, 2009, 183: 1127-1139

[17] Fournier-Level A, Lacombe T, Le Cunff L, et al. Evolution of the *VvMybA* gene family, the major determinant of berry colour in cultivated grapevine (*Vitis vinifera* L.) [J]. Heredity, 2010, 104: 351-362

[18] 刘崇怀,冯建灿,姜建福. 中国葡萄属植物形态学聚类分组研究 [J]. 植物遗传资源学报, 2011, 12(6): 847-852

[19] 王姣. 河南境内桑叶葡萄 (*Vitis ficifolia* Bge.) 种内形态和遗传多样性分析 [D]. 北京: 中国农业科学院, 2008

[20] 张永辉,刘崇怀,樊秀彩,等. ISSR 标记在中国野生葡萄分类中的应用 [J]. 果树学报, 2011, 28(3): 406-412

[21] 方连玉,王军,许雷. 15 份葡萄种质遗传多样性的 SSR 分析 [J]. 分子植物育种, 2010, 8(3): 511-515

[22] 温景辉. 基于 SSR 分子标记的山葡萄种质遗传多样性研究与核心种质构建 [D]. 长春: 吉林农业大学, 2011

[23] 温景辉,申海林,邹利人,等. 20 份葡萄种质亲缘关系的 SSR 分析 [J]. 果树学报, 2011, 28(5): 782-786

[24] 王玉娟,张彦,房经贵,等. 利用基于 RAPD 标记的 MCID 法快速鉴定 72 个葡萄品种 [J]. 中国农业科学, 2012, 45(14): 2913-2922

[25] 张旭彤. 中国野生葡萄种质资源的亲缘关系研究 [D]. 杨凌: 西北农林科技大学, 2012

[26] 慕茜,吴伟民,房经贵,等. 不同葡萄品种的 *VvmybA1* 基因型及其特征性 DNA 片段的序列分析 [J]. 园艺学报, 2011, 38(11): 2075-2084

[27] Dereeper A, Nicolas S, Le Cunff L, et al. SNIPlay: a web-based tool for detection, management and analysis of SNPs. Application to grapevine diversity projects [J]. BMC Bioinformatics, 2011, 12: 1-14

[28] 周琦,王文. DNA 水平自然选择作用的检测 [J]. 动物学研究, 2004, 25(1): 73-80

[29] Kobayashi S, Ishimaru M, Hiraoka K, et al. *Myb*-related genes of the Kyoho grape (*Vitis labruscana*) regulate anthocyanin biosynthesis [J]. Planta, 2002, 215: 924-933

[30] Vouillamoz J F, Grando M S. Genealogy of wine grape cultivars: 'Pinot' is related to 'Syrah' [J]. Heredity, 2006, 97: 102-110

[31] Lijavetzky D, Cabezas J, Ibanez A, et al. High throughput SNP discovery and genotyping in grapevine (*Vitis vinifera* L.) by combining a re-sequencing approach and SNPlex technology [J]. BMC Genomics, 2007, 8: 424

[32] Velasco R, Zharkikh A, Troggio M, et al. A high quality draft consensus sequence of the genome of a heterozygous grapevine variety [J]. PloS one, 2007, 2(12): e1326

[33] Riahi L, Zoghalmi N, Fournier-Level A, et al. Characterization of single nucleotide polymorphism in Tunisian grapevine genome and their potential for population genetics and evolutionary studies [J]. Genet Resour Crop Evol, 2013, 60: 1139-1151

[34] An C, Saha S, Jenkins J, et al. Cotton (*Gossypium* spp.) R2R3-MYB transcription factors SNP identification, phylogenomic characterization, chromosome localization, and linkage mapping [J]. Theor Appl Genet, 2008, 116: 1015-1026

[35] Kobayashi S, Goto-Yamamoto N, Hirochika H. Retrotransposon-Induced mutations in grape skin color [J]. Science, 2004, 304: 982

[36] Walker A R, Lee E, Bogs J, et al. White grapes arose through the mutation of two similar and adjacent regulatory genes [J]. Plant J, 2007, 49: 772-785