

葡萄 SAUR 基因家族鉴定与生物信息学分析

李 傲, 崔梦杰, 陈 珂, 许瀛之, 贾海峰, 房经贵

(南京农业大学园艺学院/江苏省果树品种改良与种苗繁育工程中心, 南京 210095)

摘要: 为了研究葡萄早期应答生长素基因 *SAUR* (Small auxin-up RNA) 家族, 本研究利用全基因组信息鉴定了葡萄 64 个 *SAUR* 家族成员, 并对 *SAUR* 家族成员的基因结构、氨基酸特性、染色体定位、基因进化、基因功能以及组织表达进行分析。结果表明, 葡萄全基因组上 64 个 *SAUR* 家族成员在 19 条染色体中的 8 条染色体上呈现簇状分布, 主要分布在 3、4 号染色体上, 其中 3 号染色体上数量最多为 37 个; 葡萄 *SAUR* 家族基因长度较短, 有 59 个基因是无内含子基因; 蛋白理化特征分析显示, 多数 *SAUR* 蛋白呈碱性, 结构稳定性较差, 蛋白脂溶指数高, 呈亲水性; 基因功能预测结果表明, 葡萄 *SAUR* 基因主要担当生长因子、结构蛋白、转录、转录调控以及响应胁迫应答和免疫应答 6 种功能, 其中更多参与生长调节功能; 根据系统进化分析将其分为 10 个分支, 另外不同组织表达谱的分析结果表明 *SAUR* 基因家族成员具有不同的组织表达模式, 对于非生物胁迫具有一定的调节作用。这些信息为葡萄 *SAUR* 基因家族功能分析奠定了一定的工作基础。

关键词: 葡萄; *SAUR* 基因家族; 系统进化; 功能预测; 基因表达; 定量 PCR

Identification and Bioinformatics Analysis of the *SAUR* Gene Family in Grape

LI Ao, CUI Meng-jie, CHEN Ke, XU Ying-zhi, JIA Hai-feng, FANG Jing-gui

(College of Horticulture, Nanjing Agricultural University/

Jiangsu Fruit Group Genetics Improvement and Seeding Propagation Engineering Center, Nanjing 210095)

Abstract: The phytohormone auxin exerts a pleiotropic effect on various aspects of plant growth and development, including cell elongation, cell division, differentiation, root initiation, apical dominance, and tropic responses. Auxin mediates these effects at the molecular level by altering the expression of numerous genes. The early auxin response genes are mostly classified into three families: *AUX/IAAs*, *GRETCHEN HAGEN3s (GH3s)*, and Small Auxin-Up RNA (*SAUR*). In order to recognize the *SAUR* gene from the whole genome of the grape, the *SAUR* gene was identified from the whole genome of the grape, and the gene structure, amino acid characterization, chromosome localization, gene evolution, functional analysis and tissue expression analysis of the *SAUR* gene family were carried out. A total of 64 members of the *SAUR* gene family were identified and showed cluster distribution on 8 chromosomes among 19 chromosomes. The genes were mainly distributed in Chr3 and Chr4. Chr3, exist the highest number of genes, 37 were distributed in it. The length of grape *SAUR* family genes is shorter, where 59 genes are intronless. Analysis of protein physical and chemical characteristics showed that most *SAUR* protein was alkaline, the structure stability was poor, the protein fat soluble index was high, and it was hydrophilic. The function prediction of genes showed that the *SAUR* gene mainly functioned as growth factors, structural proteins, transcription, transcriptional regulation and responded to stress and immunization response and immune response. Most of them were involved in growth regulation. According to the phylogenetic analysis, it was divided into 10 branches. Analysis of different tis-

收稿日期: 2017-08-14 修回日期: 2017-10-31 网络出版日期: 2018-02-09

URL: <http://kns.cnki.net/kcms/detail/11.4996.S.20180209.0901.018.html>

基金项目: 江苏省科技支撑计划 (BE2013431); 南京农业大学校青年基金 (KJ2013013)

第一作者研究方向为葡萄基因组学与分子生物学研究。E-mail: 2015104013@njau.edu.cn

通信作者: 房经贵, 研究方向为葡萄遗传育种与基因组学。E-mail: fanggg@njau.edu.cn

sue expression profiles showed that *SAUR* gene family members had different tissue expression patterns and had some regulation effect on abiotic stress. This information has laid a foundation for the functional analysis of *SAUR* gene family. This work will pave a new era of applying genetic information to a deeper understanding of grapes and be used to enhance the agricultural production. Making genetic information and genomics applied in many aspects of grape production, which means that the era of application of genetic information in crop production has arrived.

Key words: grape; *SAUR* genes family; system evolution; functional prediction; gene expression; quantitative PCR

植物生长素能促进植物生根,调节芽发育以及细胞分化与分裂,并且在植物的顶端优势、向光性、两重性、防止落花落果、果实的发育、胚的形成等方面也是必需的^[1]。生长素与生长素早期应答基因有着密切的关系,生长素早期应答基因可以分为 *AUX/IAA*、*GH3*、*SAUR* 三大基因家族^[2]。其中生长素早期应答基因家族(*SAUR*)的基因是受生长素感应最迅速和最强烈的一类基因,它们能在 2~5 min 内响应活跃生长素基因,并存在于许多植物中编码较特异的小分子蛋白质^[3-4]。在植物体内,*SAUR* 起着较为关键的作用^[5],如发现玉米中 *SAUR2* 与生长素介导的细胞伸长有关^[6];拟南芥中 *SAUR63* 促进了 40 个生长素相关基因来刺激器官伸长,*SAUR41* 调节细胞中生长素的转运,*SAUR76~78* 可能影响乙烯受体信号并促进植物生长^[7-8];水稻中 *SAUR39* 基因在生长素的合成和运输中起着负调节的作用^[9]。目前关于 *SAUR* 基因家族的报道还尚少,第 1 个 *SAUR* 基因是在大豆中发现的,随后在拟南芥中发现 72 个 *SAUR* 基因,水稻中有 58 个 *SAUR* 基因,番茄和马铃薯中分别鉴定出了 55 个和 134 个 *SAUR* 基因^[10-12]。近年来,葡萄基因组测序的完成成为葡萄基因组学的研究提供了更加全面的思路和条件^[13-15]。然而关于葡萄 *SAUR* 基因家族的研究却很少见。本文拟结合前人的研究,对葡萄 *SAUR* 基因家族进行鉴定并对其结构特征、功能特性、染色体定

位、蛋白质特性以及在不同组织中的表达情况进行研究,这些数据更加有助于我们理解葡萄 *SAUR* 基因家族,为今后葡萄 *SAUR* 基因家族的功能研究奠定基础。

1 材料与方法

1.1 材料

葡萄 (*Vitis vinifera* L.) 和拟南芥 (*Arabidopsis thaliana* (L.) Heynh.) 全基因组数据下载于 Phytozome 数据库 (<http://www.phytozome.net>)^[16]。葡萄芯片数据来源于 NCBI 的 GEO 数据库,芯片编号为 GSE36128。供试葡萄品种为 2 年生“魏可”盆栽苗。反转录试剂盒购自 Clontech 公司。荧光定量染料 SYBR Green I 购自于东洋纺生物科技有限公司。所用引物由上海生工生物工程股份有限公司合成(表 1)。

1.2 方法

1.2.1 葡萄 SAUR 基因家族数据集的建立和染色体定位分析 首先运用模式植物拟南芥的 *SAUR* 基因家族成员为靶序列^[17],对葡萄全基因组序列进行 BLAST 搜索,同时利用 Pfam 软件数据库工具建立葡萄全基因组蛋白结构域模型,使用 hmmer 程序筛选含有 *SAUR* 结构域(PF02519)的蛋白序列。使用 BioEdit 软件对葡萄 *SAUR* 蛋白进行多序列比对,并候选 $E \leq 10^{-5}$ 的序列作为候选蛋白序列,删除重复序列,最终获得了 64 个葡萄 *SAUR* 成员。

表 1 荧光定量 PCR 引物

Table 1 Primers used for real-time quantitative PCR

基因名称	正向引物序列	反向引物序列
Gene name	Forward primer sequence	Reverse primer sequence
<i>VtSAUR4</i>	GAAGAAAGCTGCTCAGCGTG	CGACTTGGGAAAACAGAGCC
<i>VtSAUR9</i>	TCCATCTGCAGAGTCCACAA	GTCAAACCCGAACCTTCCG
<i>VtSAUR17</i>	CGCTTGCCTGGGATTGTAA	CCCGAACTCTTCTCAGCCT
<i>VtSAUR25</i>	GGTTTTCGTTTGCTGGGAT	CCTGACTCAACAAGTCTGG
<i>VtSAUR35</i>	AGGAGGACTTTGAAGTGCCT	ATGGGAAGAGGACAGTGAGC
<i>VtSAUR39</i>	CATCAGTGGCAGATAAGGGC	TGATGGGGCCTTCACTTTGT
<i>VtSAUR47</i>	GCCCTCAATTCTTCATGCCA	GGAACCACAAACCGCTTCTT
<i>VtSAUR49</i>	CAGAAAGTCCCAAAAGGCCAC	CATTGCAGGGGATTGTGACA
<i>VtSAUR50</i>	TCAGCAACAACAGCAGAGGT	TTGCATGGGATTGTGACACC
<i>VtSAUR51</i>	CCTCAAAGTCAGTCTGTCC	AAACTCTTCTCTGCCTGCT
<i>ACTIN</i>	TACAATTCATCATGAAGTGTGATG	TTAGAAGCACTTCTGTGAACAATG

选取 *SAUR* 基因的染色体位置信息,使用 Map-Draw^[18] 绘制 *SAUR* 基因的染色体分布图。*SAUR* 基因家族成员包括如下信息:基因位点、染色体位置、基因组匹配区段、基因链信息。对获得的 64 个葡萄 *SAUR* 成员根据其染色体定位进行命名,命名方式参照拟南芥 *SAUR* 家族成员的方式,按照基因在染色体上的位置进行^[17]。

1.2.2 葡萄 *SAUR* 基因结构特征分析 使用 Gene Structure Display Server (<http://gsds.cbi.pku.edu.cn/index.php>) 工具进行基因结构绘图。使用在线软件 Multiple Em for Motif Elicitation (MEME) suite 4.11.1^[19] (<http://meme.nbcr.net/meme/>) 鉴定葡萄 *SAUR* 蛋白中保守的基序 (motif), 并通过 SMART 软件对预测的保守基序进行功能注释,最后用 VENN2.1 在线软件 (<http://bioinfo.gp.cnb.csic.es/tools/venny/index.html>) 绘制保守基序在不同基因的分布的韦恩图。

利用在线软件 ProtParam tool^[20] 分析 *SAUR* 基因家族成员氨基酸基本的理化性质:氨基酸个数、理论等电点、总平均疏水性 (GRAVY)、脂溶指数 (AI)、蛋白不稳定指数 (II)。利用软件 Prabi (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_hnn.html) 对 *SAUR* 基因家族的蛋白二级结构进行分析。

1.2.3 葡萄 *SAUR* 基因家族进化以及功能分析 使用 MEGA7 软件对葡萄 *SAUR* 蛋白进行多序列联配比对分析并构建系统发育树。通过在线预测蛋白质功能分类的工具 ProtFun^[21-22] (<http://www.cbs.dtu.dk/services/ProtFun/>) 进行葡萄 *SAUR* 基因的蛋白质 GO (gene ontology) 功能预测。选出最大概率值所指示的注释信息。

1.2.4 葡萄 *SAUR* 基因家族的表达分析 葡萄芯片数据来源于 NCBI 的 GEO 数据库芯片平台 GPL13936,芯片编号 GSE36128 葡萄芯片数据包括 54 个葡萄样品,覆盖了不同阶段的大部分葡萄器官^[23]。葡萄在铜、水以及氧胁迫下基因的表达情况由本实验室在上海汉宇生物技术公司测序完成。

将葡萄 2 年生盆栽苗分成 4 组,第 1 组用于对照,浇灌适量的水并用气泵进行正常通气;第 2 组用于水胁迫,用气泵进行正常通气并浇灌大量的水;第 3 组用于铜胁迫,在浇灌的适量的水中加入 1 mmol/L CuSO_4 并用气泵进行正常通气;第 4 组用于氧胁迫,浇灌适量的水并在正常通气的气泵中通入纯净的

N_2 ,在处理 1d 后采样,均选择从茎基部起第 5~7 片叶片。每个处理有 3 个生物学重复,每个重复 10 棵植株,所采取样品用液氮速冻处理,放 -80°C 冰箱备用。

1.2.5 RNA 的提取纯化与 cDNA 合成 使用传统 CTAB 方法^[24] 进行总 RNA 提取,提取后参照 TaKaRa 产品 DNase I (RNase free) 使用说明书,进行微量基因组 DNA 的去除。使用 One Drop OD-1000+ 超微量分光光度计检测样品 RNA 的纯度,然后取 3 μL 样品琼脂糖电泳检测 RNA 的完整性。以 RNA 样品为模板,参考 TaKaRa 产品 RNA 反转录试剂盒说明书,进行 cDNA 的合成,获得的 cDNA 产物直接用于 PCR 或 -20°C 贮藏备用。

1.2.6 PCR 引物设计及其荧光定量 PCR 方法步骤 以葡萄 *ACTIN* 为内参基因,按照 SYBR Premix ExTaqTM 试剂盒 (购自宝生物工程有限公司) 操作指导,应用 Bio-Rady-IQ2 采用实时荧光定量 PCR (RT-qPCR, real-time quantitative PCR) 方法,检测基因的相对表达量。反应体系按 SYBR Green I (TOYOBO) 说明书进行。扩增体系含 1 μL cDNA,上下游引物各 0.8 μL (引物信息详见表 1),10 μL 反应 MIX,7.4 μL ddH_2O ,总体积 20 μL 。反应程序为 95°C 变性 1 min, 95°C 变性 10 s, T_m 退火 20 s, 72°C 延伸 30 s,40 个循环,退火温度为 58°C ;反应结束后分析荧光值变化曲线以及融解曲线。试验设 3 次重复,试验数据用 Lin-RegPCR 和 Excel 软件分析,采用 CK 组的值为对照来计算相对值,用 $2^{-\Delta\Delta\text{CT}}$ 计算方法,相对表达量为处理组和对照组的相对值。以不加模板 cDNA 为阴性对照,标准品 cDNA 和待测样品均设置 3 次重复。

2 结果与分析

2.1 葡萄 *SAUR* 家族的全基因组鉴定

利用 BioEdit, hmmer 程序以及 Pfam 处理后获得了 64 个 *SAUR* 基因家族成员 (表 2)。进一步分析发现 *SAUR* 基因家族成员分布在 Chr 1、2、3、4、6、17、18、19 等 8 条染色体上 (图 1)。葡萄 *SAUR* 成员除了 *VvSAUR1*、4、42、59、60、64 等 6 个基因外,其他基因在染色体上都呈现簇状分布。如: Chr 3 分布的 37 个 *VvSAUR* 基因均在染色体的上端;其次是 Chr 4 上的 13 个 *VvSAUR* 基因中有 12 个都分布在染色体的中下端;而 Chr 6 上仅有的 2 个 *VvSAUR* 基因分布在染色体下端, Chr17 染色体上仅有 2 个 *VvSAUR* 基因分布在染色体的上端。

表 2 葡萄 SAUR 基因家族信息

Table 2 The information of SAUR gene family in grape

染色体 Chromo- some	基因位点 Gene locus	基因组匹配区段 Genome location	基因链 Strand	基因大小 (外显子 + 内含子)(bp)	蛋白的理化性质 Physicochemical properties					不同类型二级结构占 总二级结构百分比				功能 预测 Gene ontology category
				The length of genes						Percentage of different types of secondary structures to total secondary structure				
				(exon + intron)	A	B	C	D	E	1	2	3	4	
Chr1	VtSAUR1	6254753..6255088	1	335(335 +0)	111	5.72	0.195	99.91	31.71	0.4324	0	0.4234	0.1441	GF
	VtSAUR2	2196642..21967514	-1	1090(1090 +0)	153	9.32	-0.478	72.09	44.85	0.2418	0	0.5882	0.1699	GF
	VtSAUR3	22022590..22023096	1	506(506 +0)	154	9.62	-0.553	68.31	32.42	0.2532	0	0.5455	0.2013	TR
Chr2	VtSAUR4	4822437..4823283	-1	846(846 +0)	163	10.62	-0.585	71.04	38.08	0.3252	0	0.4724	0.2025	SP
Chr3	VtSAUR5	736435..737319	1	884(884 +0)	141	9.78	-0.284	83.62	68.92	0.3617	0	0.5248	0.1135	SP
	VtSAUR6	738127..738887	1	760(760 +0)	147	9.45	-0.162	88.78	71.83	0.4354	0	0.4286	0.1361	GF
	VtSAUR7	749716..750398	-1	682(682 +0)	104	7.74	-0.274	87.12	52.12	0.3654	0	0.5	0.1346	GF
	VtSAUR8	858886..859170	-1	284(284 +0)	94	6.56	-0.18	82.87	44.59	0.3191	0	0.5213	0.1596	SR
	VtSAUR9	862995..863398	1	403(403 +0)	95	8.56	-0.08	95.47	48.19	0.3789	0	0.4737	0.1474	IR
	VtSAUR10	864608..864876	1	268(268 +0)	88	5.71	-0.126	67.61	39.81	0.2614	0	0.5795	0.1591	TR
	VtSAUR11	866357..866897	1	540(540 +0)	95	8.55	-0.12	94.42	58.36	0.3579	0	0.5053	0.1368	SP
	VtSAUR12	867544..868366	1	822(822 +0)	88	5.71	-0.101	68.75	37.45	0.2841	0	0.5568	0.1591	TR
	VtSAUR13	869586..872380	1	2794(2794 +0)	95	7.76	-0.146	93.37	60.39	0.3579	0	0.4947	0.1474	SP
	VtSAUR14	880653..881315	-1	662(662 +0)	95	8.93	-0.069	93.37	62.88	0.3263	0	0.4947	0.1789	SP
	VtSAUR15	883689..883976	-1	287(287 +0)	95	6.82	-0.058	85.16	49.75	0.3789	0	0.4842	0.1368	TP
	VtSAUR16	885397..885684	-1	287(287 +0)	95	6.95	-0.138	89.26	55.07	0.3579	0	0.4947	0.1474	SP
	VtSAUR17	888635..888955	1	320(320 +0)	95	8.49	-0.182	81.05	35.47	0.3895	0	0.4842	0.1263	SR
	VtSAUR18	892009..892382	1	373(373 +0)	95	8.69	-0.129	97.47	51.11	0.3684	0	0.4842	0.1474	SP
	VtSAUR19	894783..895216	-1	433(433 +0)	95	8.93	-0.131	93.37	56.67	0.3263	0	0.5053	0.1684	SP
	VtSAUR20	897813..898180	1	367(367 +0)	95	6.55	-0.089	82	35.24	0.3263	0	0.5368	0.1368	IR
	VtSAUR21	900854..901605	1	751(751 +0)	95	8.55	-0.089	97.47	51.11	0.3684	0	0.4842	0.1474	SP
	VtSAUR22	903102..903604	-1	502(502 +0)	95	9.06	-0.192	93.37	54.95	0.3053	0	0.4842	0.2105	SP
	VtSAUR23	905892..906181	1	289(289 +0)	95	7.76	-0.209	78.95	42.79	0.3474	0	0.5053	0.1474	IR
	VtSAUR24	908817..909399	1	582(582 +0)	95	8.55	-0.146	93.37	57.57	0.3579	0	0.4947	0.1474	SP
	VtSAUR25	910867..911175	-1	308(308 +0)	88	5.72	-0.087	78.75	46.74	0.3636	0	0.4773	0.1591	SP
	VtSAUR26	913170..913875	1	705(705 +0)	95	8.55	-0.146	93.37	57.57	0.3579	0	0.4947	0.1474	SP
	VtSAUR27	915230..915499	-1	269(269 +0)	89	4.97	-0.104	77.87	38.84	0.2921	0	0.5056	0.2022	GF
	VtSAUR28	917424..917748	-1	324(324 +0)	95	8.8	-0.191	84.11	39.34	0.2421	0	0.5368	0.2211	IR
	VtSAUR29	921733..927965	1	6232(829 +5403)	144	9.69	-0.216	94.79	65.15	0.3194	0	0.4514	0.2292	TP
	VtSAUR30	2513584..2514003	1	419(419 +0)	139	6.71	-0.047	95.4	49.46	0.3741	0	0.4173	0.2086	GF
	VtSAUR31	2515009..2517162	1	2153(836 +1317)	276	8.59	-0.011	97.54	56.88	0.5362	0	0.3442	0.1196	GF
	VtSAUR32	2518613..2520751	1	2138(509 +1629)	167	9.3	-0.093	94.67	61.55	0.497	0	0.3533	0.1497	GF
	VtSAUR33	2524184..2524597	1	413(431 +0)	137	9.55	-0.129	89.05	48.36	0.4599	0	0.3723	0.1679	GF
	VtSAUR34	2525431..2525885	-1	454(454 +0)	139	7.73	-0.11	82.09	37.04	0.4173	0	0.4173	0.1655	GF

表 2(续)

染色体 Chromo- some	基因位点 Gene locus	基因组匹配区段 Genome location	基因链 Strand	基因大小 (外显子 + 内含子)(bp)	蛋白的理化性质					不同类型二级结构占 总二级结构百分比				功能 预测 Gene ontology category
				The length of genes (exon + intron)	Physicochemical properties					Percentage of different types of secondary structures to total secondary structure				
					A	B	C	D	E	1	2	3	4	
Chr4	<i>VaSAUR35</i>	2528365..2532361	1	3996(866 + 3130)	270	8.69	-0.037	92.15	54.44	0.4296	0	0.437	0.1333	GF
	<i>VaSAUR36</i>	2533266..2533720	-1	454(454 + 0)	148	9.07	-0.107	87.64	46	0.4865	0	0.3851	0.1284	GF
	<i>VaSAUR37</i>	2536233..2541680	1	5447(899 + 4548)	236	7.08	-0.058	87.54	50.59	0.4958	0	0.3517	0.1525	GF
	<i>VaSAUR38</i>	2746826..2748763	-1	1937(1937 + 0)	148	8.76	-0.093	88.99	57.18	0.5	0	0.3784	0.1216	GF
	<i>VaSAUR39</i>	2749512..2749985	-1	473(473 + 0)	139	7.62	-0.028	88.42	56.38	0.3957	0	0.4173	0.1871	GF
	<i>VaSAUR40</i>	2751683..2752647	-1	964(964 + 0)	148	8.84	-0.067	92.91	51.62	0.473	0	0.3514	0.1757	GF
	<i>VaSAUR41</i>	2755238..2755923	-1	685(685 + 0)	146	9.44	-0.23	82.81	61.37	0.4178	0	0.4041	0.1781	GF
	<i>VaSAUR42</i>	2328075..2329742	1	1667(1667 + 0)	162	9.22	-0.45	84.81	54.89	0.3272	0	0.537	0.1358	TP
	<i>VaSAUR43</i>	16733744..16734278	-1	534(534 + 0)	96	7.82	-0.228	81.25	47.2	0.375	0	0.4583	0.1667	SP
	<i>VaSAUR44</i>	16735498..16735794	-1	296(296 + 0)	89	8.64	-0.183	96.4	40.6	0.4157	0	0.4157	0.1685	SP
	<i>VaSAUR45</i>	16740348..16740890	-1	542(542 + 0)	96	9.21	-0.184	93.44	47.71	0.3438	0	0.4896	0.1667	GF
	<i>VaSAUR46</i>	16755865..16756167	-1	302(302 + 0)	96	7.84	-0.108	96.46	55.2	0.3854	0	0.4583	0.1562	SP
	<i>VaSAUR47</i>	16765725..16766094	-1	369(369 + 0)	103	9.25	-0.057	85.15	55.57	0.2524	0	0.5146	0.233	IR
	<i>VaSAUR48</i>	16767601..16767914	-1	313(313 + 0)	98	6.9	-0.205	77.55	48.64	0.2245	0	0.5714	0.2041	SP
	<i>VaSAUR49</i>	16769755..16770326	-1	571(571 + 0)	103	7.97	-0.055	90.87	49.28	0.3786	0	0.4272	0.1942	SP
	<i>VaSAUR50</i>	16775746..16776426	-1	680(680 + 0)	103	9.3	0.039	96.6	42.94	0.3301	0	0.4757	0.1942	GF
<i>VaSAUR51</i>	16779590..16780085	-1	495(495 + 0)	97	6.9	-0.115	94.43	66.64	0.4227	0	0.4227	0.1546	TP	
<i>VaSAUR52</i>	16782060..16782354	-1	294(294 + 0)	97	6.16	-0.186	88.35	51.61	0.3711	0	0.4536	0.1753	GF	
<i>VaSAUR53</i>	16785245..16785735	-1	490(490 + 0)	91	6.26	-0.168	91.98	46.83	0.3736	0	0.4615	0.1648	TP	
<i>VaSAUR54</i>	19785898..19787577	-1	1679(1679 + 0)	104	8.62	-0.269	83..4	45.13	0.3654	0	0.5096	0.125	SP	
Chr6	<i>VaSAUR55</i>	17809099..17809535	1	436(436 + 0)	143	6.43	-0.415	87.9	38.55	0.3636	0	0.5105	0.1259	GF
	<i>VaSAUR56</i>	17828336..17828767	1	431(431 + 0)	143	6.83	-0.427	83.78	31.52	0.2867	0	0.5804	0.1329	GF
Chr17	<i>VaSAUR57</i>	675175..676163	1	988(988 + 0)	102	9.93	-0.743	51.57	39.72	0.2941	0	0.549	0.1569	TR
	<i>VaSAUR58</i>	691438..691896	1	458(458 + 0)	152	9.38	-0.352	70.53	56.98	0.2368	0	0.5855	0.1776	SP
Chr18	<i>VaSAUR59</i>	1262287..1262685	1	398(398 + 0)	132	9.51	0.044	93.11	65.35	0.4697	0	0.3939	0.1364	GF
	<i>VaSAUR60</i>	4960794..4961129	1	335(335 + 0)	111	5.2	0.065	92.97	35.88	0.3604	0	0.4685	0.1712	GF
	<i>VaSAUR61</i>	11959623..11960345	1	722(722 + 0)	151	9.19	-0.278	77.42	66.8	0.4371	0	0.4636	0.0993	IR
	<i>VaSAUR62</i>	11987475..11987922	-1	447(447 + 0)	104	6.72	-0.257	87.12	54.18	0.375	0	0.5	0.125	GF
	<i>VaSAUR63</i>	12002927..12003389	1	462(462 + 0)	88	6.71	-0.131	90.91	36.89	0.4432	0	0.3977	0.1591	TP
Chr19	<i>VaSAUR64</i>	22279071..22279770	1	699(699 + 0)	131	8.93	-0.341	93.66	75.43	0.3511	0	0.4656	0.1832	TP

1: 正义链; -1: 反义链; A: 氨基酸个数; B: 理论等电点; C: 总平均疏水性 (GRAVY); D: 脂溶指数 (AI); E: 蛋白不稳定指数 (II); 1: α -螺旋; 2: β -转角; 3: 无规则卷曲; 4: 扩展链结构; GF: 生长因子; SP: 结构蛋白; TP: 转录; IR: 免疫应答; TR: 转录调控; SR: 胁迫应答

1: forward chains, -1: reverse chains, A: Number of amino acids, B: Theoretical pI, C: Grand average of hydropathicity, D: Aliphatic index, E: Instability index, 1: Alpha helix, 2: Beta brige, 3: Random coil, 4: Extended strand, GF: growth factor, SP: structural protein, TP: transcription, IR: immune response, TR: transcription regulation, SR: stress response

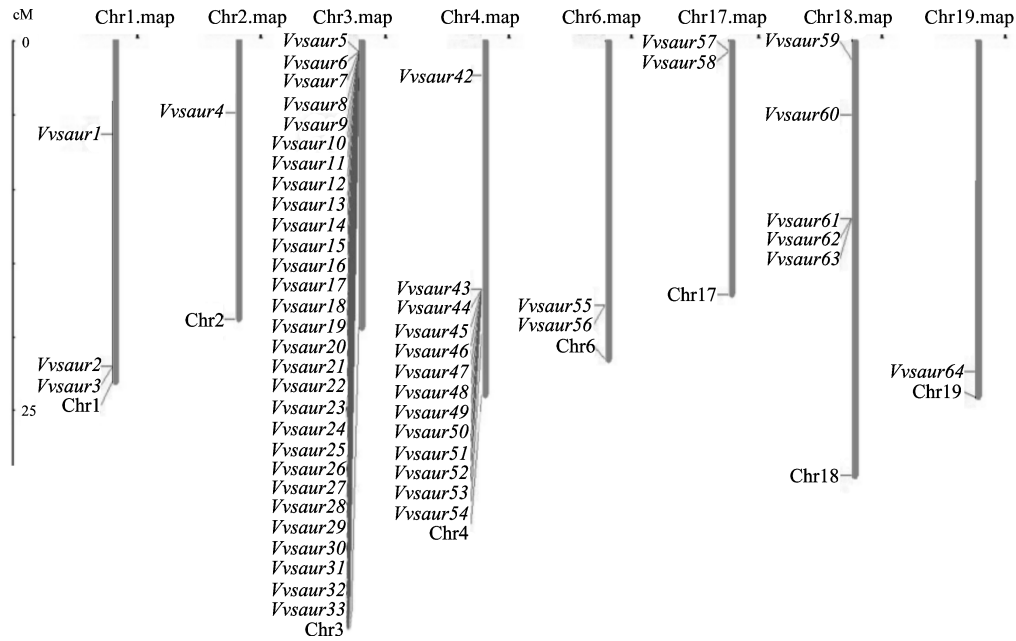


图1 葡萄 SAUR 基因家族在染色体上的位置

Fig. 1 The chromosomal distribution of SAUR gene family

2.2 葡萄 SAUR 基因家族结构特征分析

2.2.1 葡萄 SAUR 基因家族结构 葡萄 SAUR 基因家族均为无内含子基因或少内含子基因,其中 59 个为无内含子基因,仅有 5 个为少内含子基因,分别为 *VvSAUR29* 含有 3 个内含子, *VvSAUR35* 和 *VvSAUR37* 含有 2 个内含子, *VvSAUR31* 和 *VvSAUR32* 含有 1 个内含子(图 2)。另外这也决定了 SAUR 家族的基因长度明显小于葡萄基因组中其他有内含子基因。除了 *VvSAUR13*、*VvSAUR29*、*VvSAUR31*、*VvSAUR32*、*VvSAUR35* 以及 *VvSAUR37* 以外,其他基因长度均小于 2000bp,且主要集中在 0~1000bp 之间(表 2)。

2.2.2 葡萄 SAUR 家族蛋白二级结构以及蛋白质特性分析 对 SAUR 家族各基因的氨基酸序列所编码蛋白的理化性质进行分析,结果显示(表 2),葡萄 SAUR 家族的蛋白长度较短,大多数包含 90~150 个氨基酸残基。其中最长的葡萄 SAUR 蛋白(*VvSAUR31*)包含 276 个氨基酸残基,最短的蛋白(*VvSAUR10*、*VvSAUR12* 和 *VvSAUR63*)仅含有 88 个氨基酸残基。SAUR 基因家族成员对应的等电点范围从 4.97(*VvSAUR27*)到 10.62(*VvSAUR4*),大多数家族成员呈现碱性,其中有 36 个 SAUR 蛋白等电点大于 8.0,显碱性;仅有 9 个 SAUR 蛋白等电点小于 6.5,显酸性。通过分析不稳定指数发现,其中 15 个 SAUR 蛋白不稳定指数小于 40,为稳定蛋白;49 个 SAUR 蛋白不稳定指数大于 40,为不稳定蛋白。60 个 SAUR 蛋白为亲水蛋白;4 个 SAUR 蛋白为疏水蛋

白(*VvSAUR1*、*VvSAUR50*、*VvSAUR59* 和 *VvSAUR60*)。

分析发现,葡萄 SAUR 蛋白二级结构均由 α -螺旋、无规则卷曲和扩展链结构等 3 种形式组成,均未发现 β -转角结构存在。且全部氨基酸序列的二级结构都是扩展链结构所占比例最小。*VvSAUR34*、*VvSAUR44* 和 *VvSAUR51* 中各组成的百分比从大到小的顺序依次为无规则卷曲 = α -螺旋 > 扩展链结构;*VvSAUR1*、6、31、32、33、36、37、38、40、41、59 和 62 这 12 个 SAUR 蛋白序列中各组成的百分比从大到小的顺序依次为 α -螺旋 > 无规则卷曲 > 扩展链结构;其余 49 个 SAUR 蛋白序列中各组成的百分比从大到小的顺序依次为无规则卷曲 > α -螺旋 > 扩展链结构(表 2)。

2.2.3 葡萄 SAUR 基因家族的保守基序分析 采用 MEME 在线软件分析 64 个葡萄 SAUR 蛋白的序列特征,共得到 3 个保守基序,分别是 motif1 (KRFVVPISYLKHPFQNLISQAEEEFQFDHPMG)、motif2 (PSAESTBVPKGHFVYVGETQ)与 motif3 (LTIPCREEAIDLTS),基序长度分别为 33、21 和 15。其中 motif2 位置分布靠前,motif1、motif3 位置分布靠后(图 3A)。从葡萄 SAUR 3 个基序所在基因的韦恩图来看,34 个基因同时含有 motif1、motif2、motif3。10 个基因同时含有 motif2 与 motif3;4 个同时含有 motif1 与 motif2;3 个仅含有 motif1 与 motif3,为 *SAUR1*、*SAUR6* 和 *SAUR56*;8 个仅含有 motif1,分别为 *SAUR30*、*SAUR31*、*SAUR32*、*SAUR36*、*SAUR37*、*SAUR38*、*SAUR39* 与 *SAUR40*;1 个仅含有 motif2;2 个仅含有 motif3,为 *SAUR5* 与 *SAUR63*(图 3B)。

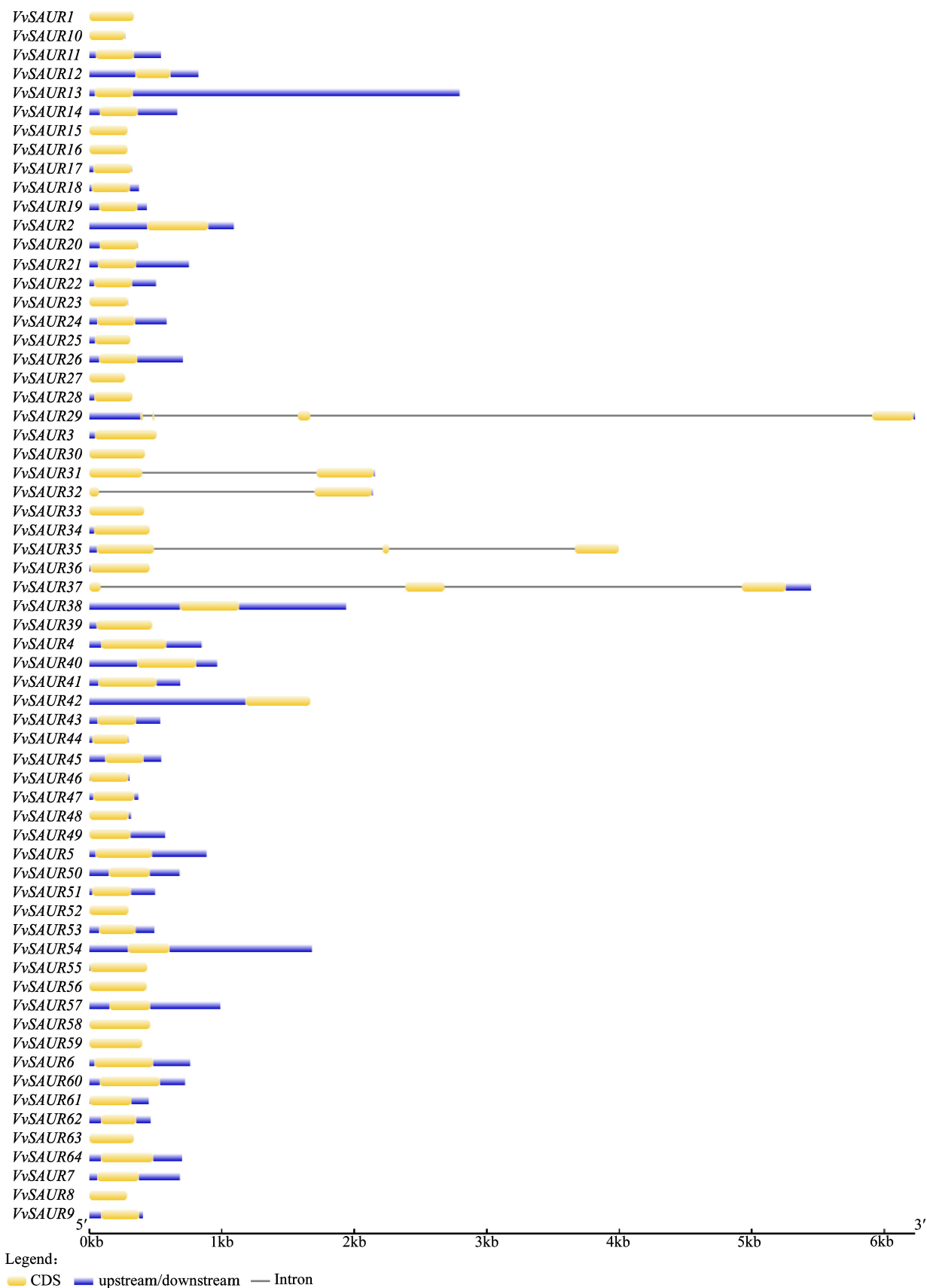
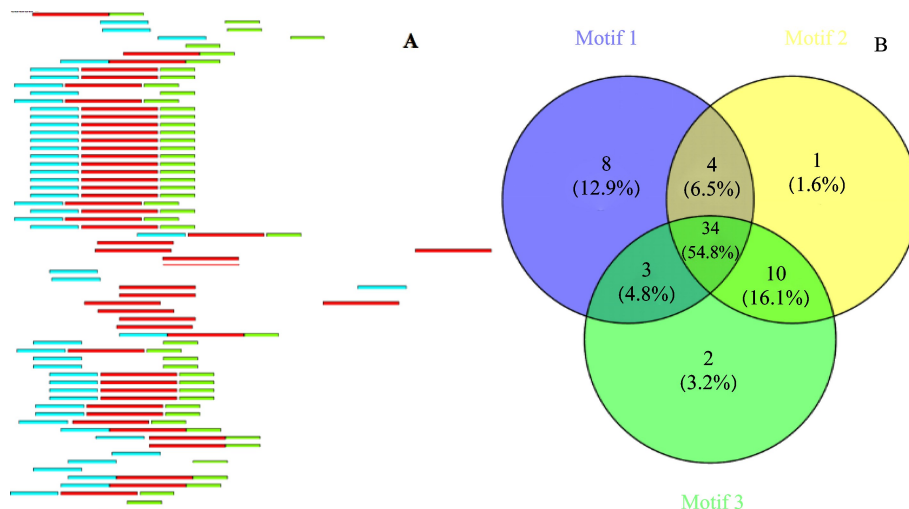


图2 葡萄 SAUR 基因家族基因结构

Fig. 2 The gene structure of SAUR gene family in grape



A: 从上到下依次为 *VvSAUR1* ~ 64, 其中红色方框: KRFVVPISYLKHPFQNLLSQAEFEFGFDHPMG;

蓝色方框: PSAESTBVPKGHFAVYVGETQ; 绿色方框: LTIPCREEAIDLTS;

B: 不同基序在基因中分布的韦恩图, 括号内数据指含有相应 Motif 的无内含子基因的数量

A: From top to bottom on behalf of *VvSAUR1* ~ 64, Red frame: KRFVVPISYLKHPFQNLLSQAEFEFGFDHPMG,

Blue frame: PSAESTBVPKGHFAVYVGETQ, Green frame: LTIPCREEAIDLTS, B: The Venn Figure shows the

distribution of different motifs in the gene, the data in brackets refers to the number of intronless genes containing the corresponding Motif

图 3 葡萄 SAUR 基因家族的保守基序分布

Fig. 3 Motifs identified from SAUR gene family in grape

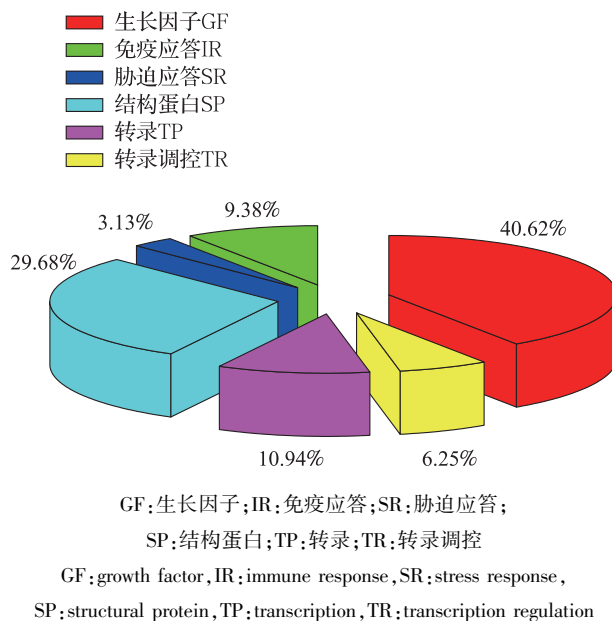
用 SMART 软件对预测的保守基序进行分析, 结果显示 motif1 属于 Auxin-inducible 结构域, 葡萄 SAUR 基因家族中有 49 个基因含有 motif1, 为 SAUR 非常重要的保守基序。由于葡萄 SAUR 基因大多具有 Auxin-inducible 结构域, 所以这类基因为应答生长素基因, 参与植物体内生长素调节活动。另外其余 motif 功能尚未被确定。

2.3 葡萄 SAUR 家族功能分析

为了更进一步了解葡萄 SAUR 家族基因的功能, 我们利用 ProtFun 分析软件预测了这些基因的编码产物的 GO 功能分类。结果显示在 13 个不同的 GO 功能中, 葡萄 SAUR 家族参与了其中 6 个功能, 分别为生长因子、免疫应答、胁迫应答、结构蛋白、转录以及转录调控。被归类为生长因子的蛋白所占的比例最多为 40.62%, 而被归类为胁迫应答的蛋白比例最少为 3.13% (图 4)。

2.4 葡萄 SAUR 家族进化关系

我们利用 64 个葡萄 SAUR 家族的全长蛋白序列构建进化树, 将 64 个家族成员分成了 10 个分支, 分别用 A、B、C、D、E、F、G、H、I、J 表示不同的分支。其中 A、C 和 J 3 个分支中家族成员较多, 分别含有 13 个、13 个和 12 个家族成员; H 和 I 两个分支中分别含有 8 和 9 个家族成员; D 分支含有 3 个成员, 为 *VvSAUR7*、*VvSAUR54* 和 *VvSAUR61*; B 分支中有 2 个



GF: 生长因子; IR: 免疫应答; SR: 胁迫应答;

SP: 结构蛋白; TP: 转录; TR: 转录调控

GF: growth factor, IR: immune response, SR: stress response,

SP: structural protein, TP: transcription, TR: transcription regulation

图 4 葡萄 SAUR 基因家族在 13 个不同的 GO 分类中的数量分布情况

Fig. 4 Distribution of grape SAUR genes among 13 GO categories

成员, 为 *VvSAUR55* 和 *VvSAUR56*; E 分支中有 2 个成员, 为 *VvSAUR51* 和 *VvSAUR52*; F、G 两个分支中都仅含有 1 个家族成员, 分别为 *VvSAUR29* 和 *VvSAUR62* (图 5)。

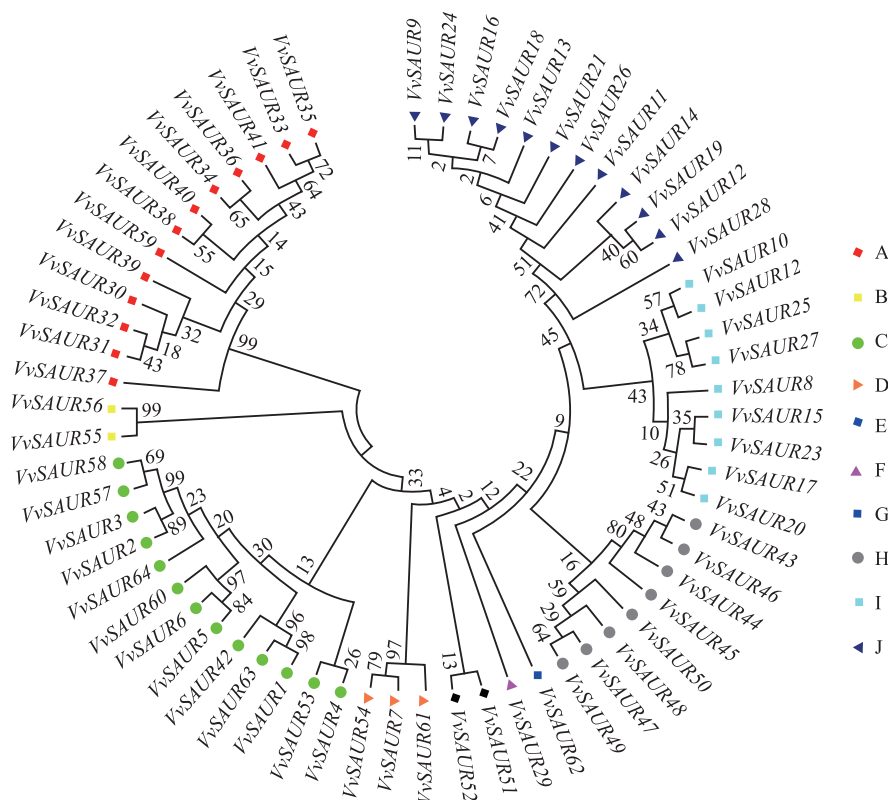


图5 葡萄 SAUR 的系统发育关系

Fig. 5 Phylogenetic relationship of SAUR gene family members in grape

从基因结构来看,所有的含有内含子的基因都分布在 A 与 F 两个分支里面,其他分支的基因都属于无内含子基因。从功能来看,A 与 B 分支内基因均被归类于生长因子;C 分支内主要参与转录以及转录调控;J 与 H 分支大多属于结构蛋白;I 分支内基因主要参与免疫与胁迫应答(表 2)。

2.5 葡萄 SAUR 家族的表达分析

2.5.1 葡萄 SAUR 基因家族在不同组织中的表达情况

通过葡萄基因芯片平台,将葡萄 SAUR 基因家族基因表达的芯片数据进行表达分析,在聚类图中用红色的深度表示基因表达的深度,绿色表示信号强度较弱,染色越深代表信号越强。结果显示(图 6),大量葡萄 SAUR 家族的基因表达强度较弱,其中 *VvSAUR1* ~ 7, *VvSAUR38*、*VvSAUR53*、*VvSAUR54*、*VvSAUR57* 及 *VvSAUR58* 在所有组织中表达量均较高。葡萄 SAUR 基因家族在不同的组织中的表达量差异不大,不过相对来说 SAUR 基因家族在花和卷须中的表达量较高。

2.5.2 葡萄 SAUR 基因家族在不同逆境处理下的表达模式

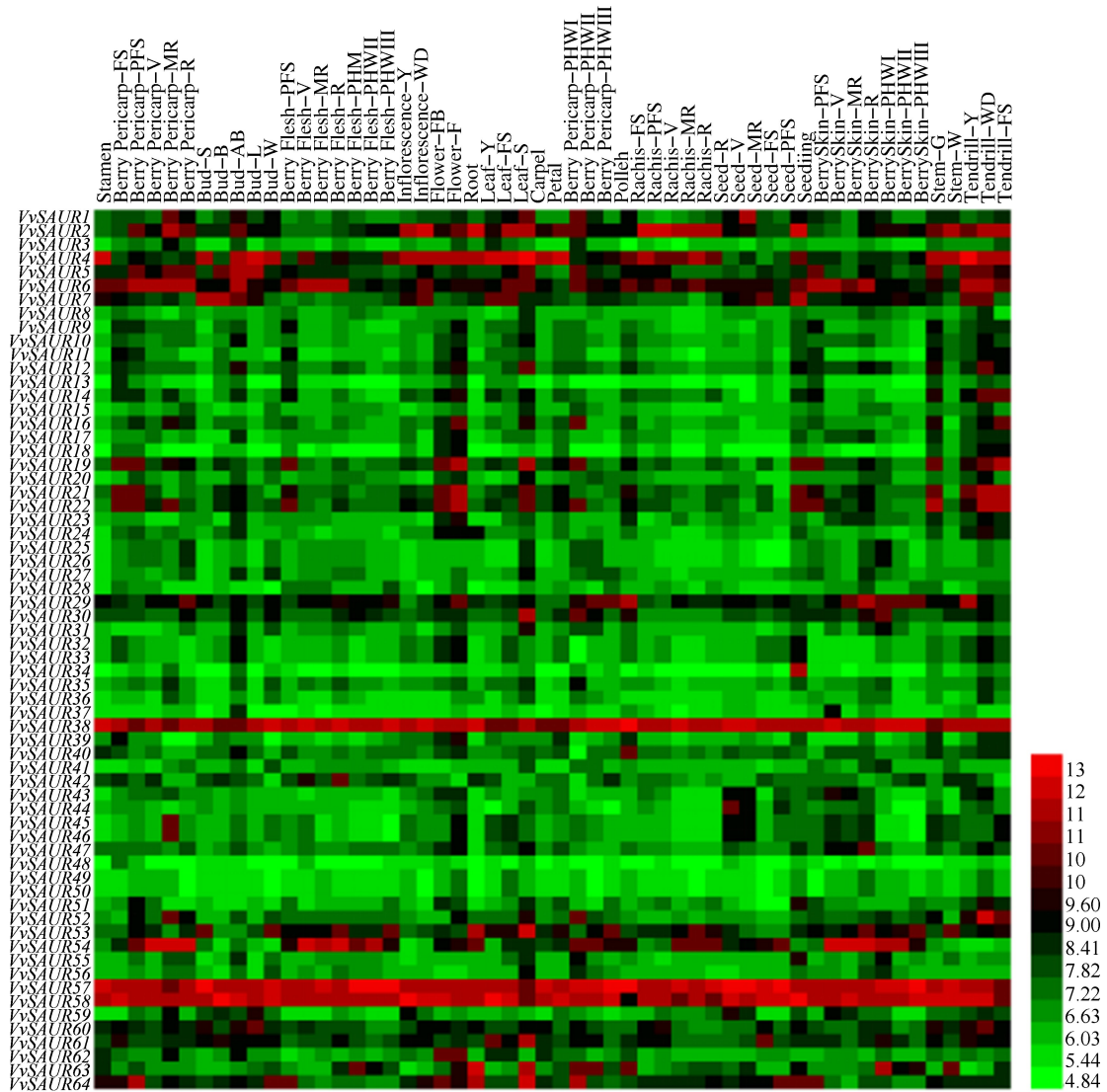
葡萄在栽培过程中会遇到涝害,土壤重金属污染以及低氧这 3 种常见胁迫。葡萄 SAUR 基因家族在水、铜、氧 3 种胁迫条件下基因的表达情

况,在聚类图中用红色的深度表示基因表达的深度,绿色表示信号强度较弱,染色越深代表信号越强。结果显示(图 7.1),葡萄 SAUR 基因家族响应胁迫较为强烈。其中 *VvSAUR56* ~ 64 在 3 种胁迫下都表现出明显的下降趋势。在水胁迫条件下大多数 SAUR 基因表达下调;在铜和氧胁迫下大多数 SAUR 基因表达上调,且在氧胁迫下葡萄 SAUR 基因表达高于铜胁迫。

葡萄 64 个 SAUR 基因中挑选 10 个基因在水、铜、氧 3 种胁迫条件下做定量分析,结果表明(图 7.2),葡萄 SAUR 基因家族响应胁迫较为强烈。其中氧与铜胁迫条件下基因表达量上调且氧胁迫条件下上调幅度较大,在水胁迫条件下大部分 SAUR 基因表达量下调。这与图 7.1 结果相一致。

3 讨论

生长素早期应答基因家族(SAUR)基因是受生长素感应最迅速和最强烈的一类基因。前人的报告已经在许多模式植物中对 SAUR 基因进行了全基因组分析,如拟南芥、水稻、高粱、茄科和玉米^[10-12,25],其中拟南芥含有 72 个 SAUR 基因成员,水稻和番茄中分别有 58 个和 55 个 SAUR 家族成员。如今我们在葡萄



Bub-X, Root, Stem-X, Leaf-X, Tendri-X, Carpel, Petal, Pollen, Rachis, Inflorescence-X, Stamen, Flower-X, BerryPericarp-X, Berryflesh-X, BerrySkin-X, Seed-X, Seeding 依次代表不同时期的芽,根,茎,叶,卷须,心皮,花瓣,花粉,花丝,花序,花药,花,葡萄果皮,葡萄果肉,葡萄表皮,种子,幼苗
Bub-X, Root, Stem-X, Leaf-X, Tendri-X, Carpel, Petal, Pollen, Rachis, Inflorescence-X, Stamen, Flower-X, BerryPericarp-X, Berryflesh-X, BerrySkin-X, Seed-X, Seeding represent different periods of Bud, Root, Stem, Leaf, Tendri, Carpel, Petals, Pollen, Rachis, Inflorescence, Stamen, Flower, BerryPericarp, Berryflesh, BerrySkin, Seed, Seedling, respectively

图 6 葡萄 SAUR 基因在不同组织中的表达模式

Fig. 6 The expression pattern of SAUR gene in different tissues

中鉴定出了 64 个 SAUR 家族成员。但是目前关于葡萄 SAUR 基因家族的系统进化分析和其基因结构特征分析未见报道,这与现在基因信息学的快速发展不符。因此在这项研究中我们对葡萄 SAUR 基因家族进行了分析,这项工作为进一步打开利用基因信息更加深入了解葡萄的新时代作出了很重要的贡献。

本研究根据葡萄综合系统发育树确定了 64 个 SAUR 基因,分为 10 组,并对 SAUR 家族成员进行基因结构、氨基酸特性分析、染色体定位、基因进化、功能分析以及组织表达分析。结果表明,外显子/内含

子结构和基序具有较高的保守性,且 64 个 SAUR 家族中有 59 个都属于无内含子基因,这一特征也表现在其他物种的 SAUR 基因家族中。例如,在拟南芥的 72 个 SAUR 基因中,除了 *AtSAUR11* 其他成员均为无内含子基因;从大豆胚轴区分离获得的 3 种 SAUR 基因都为无内含子基因;水稻中的 58 个 *Os-SAUR* 基因家族成员也都不含有内含子阵列^[17,26-27]。MEME 软件分析结果表明,葡萄大量 SAUR 家族成员都具有 motif1、motif2、motif3,且排列顺序按照 2-1-3 进行排列,这表明葡萄中 SAUR 基因家族的结构域具有较强的保守性,这与赵敬会等^[28]

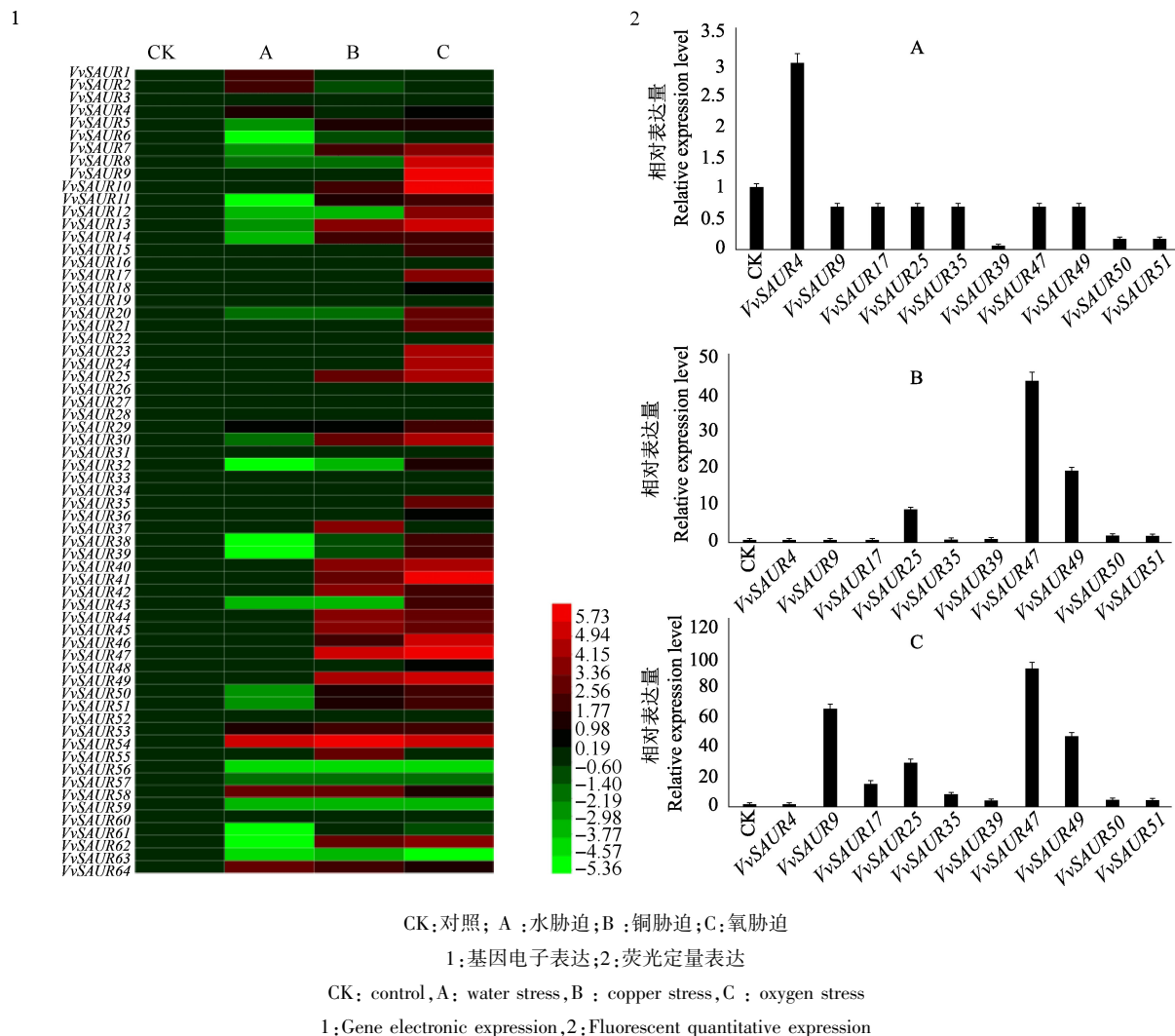


图7 葡萄 SAUR 基因在 3 种不同胁迫条件下的基因表达情况

Fig. 7 Gene expression of SAUR genes of grape under three different stress conditions

人对白菜 SAUR 的研究一致。通过葡萄中无内含子基因 GO 分类进行比较,我们发现葡萄有一些部分被分类到免疫应答,这与前人在玉米中的研究相一致^[29]。葡萄无内含子基因中参与免疫应答的蛋白比例揭示了这些基因帮助葡萄增强它在生物胁迫和非生物胁迫下的反应。另外葡萄 SAUR 基因家族在不同组织中的表达量较低,但在 3 种胁迫条件下表达量较为显著,这可能是由于在胁迫作用下快速应答的基因更趋向于没有内含子的基因^[30]。

通过对葡萄 SAUR 家族成员进行染色体定位分析,发现大量的 SAUR 基因呈现簇状分布,我们推断这可能是由于家族的基因重复造成的。基因重复是基因家族数量增加的主要方式,在基因家族中通常具有相似的基因,如不平等的交换、逆转录或全基因重复,以产生与原始基因相似的基因或碱基序列^[31]。在白菜的 SAUR 基因家族中共发现了 51 对

基因重复现象,说明了基因重复是白菜 SAUR 基因扩张的主要动力^[28]。基因重复在植物演化过程中扮演着重要角色,在生物学中普遍存在^[32]。

研究葡萄 SAUR 基因家族可以帮助了解相关基因和基因组的进化模式。通过基因组学和生物信息学验证得出葡萄 SAUR 基因家族的基因结构、序列特征及不同物种间的进化分析,为今后更加深入的了解葡萄基因组打下坚实的基础。从生物信息学角度预测了其潜在的功能,预期本研究报道的结果对于研究这些基因可能涉及的性状的植物生物学家以及葡萄或葡萄科其他成员的植物将是特别有价值的。

参考文献

- [1] 张娟. 生长素信号转导途径及参与的生物学功能研究进展[J]. 生命科学研究, 2009, 13(3): 272-277
- [2] 李亚男, 冯霞, 陈大清. *ARF*、*Aux*/*IAA* 和生长素受体对基因表达的调控[J]. 安徽农学通报, 2008, 14(7): 36-39

- [3] Spartz A K, Lee S H, Wenger J P, et al. The SAUR19 subfamily of SMALL AUXIN UP RNA genes promote cell expansion[J]. Plant J, 2012, 70(6): 978-990
- [4] Hagen G, Guilfoyle T. Auxin-responsive gene expression: genes, promoters and regulatory factors[J]. Plant Mol Biol, 2002, 49(3-4): 373-385
- [5] Ren H, Gray W M. SAUR proteins as effectors of hormonal and environmental signals in plant growth[J]. Mol Plant, 2015, 8(8): 1153-1164
- [6] Knauss S, Rohrmeier T, Lehle L. The auxin-induced maize gene ZmSAUR2 encodes a short-lived nuclear protein expressed in elongating tissues[J]. J Biol Chem, 2003, 278: 23936-23943
- [7] Chae K, Isaacs C G, Reeves P H, et al. Arabidopsis SMALL AUXIN UP RNA63 promotes hypocotyl and stamen filament elongation[J]. Plant J, 2012, 71(4): 684-697
- [8] Kong Y, Zhu Y, Gao C, et al. Tissue-specific expression of SMALL AUXIN UP RNA41 differentially regulates cell expansion and root meristem patterning in Arabidopsis[J]. Plant Cell Physiol, 2013, 54: 609-621
- [9] Kant S, Bi Y M, Zhu T, et al. SAUR39, a small Auxin-up RNA gene, acts as a negative regulator of auxin synthesis and transport in Rice[J]. Plant Physiol, 2009, 151(2): 691-701
- [10] Li Y, Gretchen H, Tom J, Guilfoyle. An Auxin-responsive promoter is differentially induced by auxin gradients during tropisms[J]. Plant Cell, 1991, 3(11): 1167-1175
- [11] Mukesh J, Akhilesh K T, Jitendra P K. Genome-wide analysis, evolutionary expansion, and expression of early auxin-responsive SAUR gene family in rice (*Oryza sativa*) [J]. Genomics, 2006, 88(3): 360-371
- [12] Wu J, Liu S, He Y, et al. Genome-wide analysis of SAUR gene family in Solanaceae species[J]. Gene, 2012, 509(1): 38-50
- [13] Jaillon O, Aury J M, Noel B, et al. The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm-phyla[J]. Nature, 2007, 449(7161): 463-467
- [14] Zenoni S, Delledonne M. Characterization of transcriptional complexity during berry development in *Vitis vinifera* using RNA-Seq[J]. Plant Physiol, 2010, 152(4): 1787-1795
- [15] Venturini L, Ferrarini A, Zenoni S, et al. De novo transcriptome characterization of *Vitis vinifera* cv. Corvina unveils varietal diversity[J]. BMC Genomics, 2013, 14(1): 41
- [16] Goodstein D M, Shu S, Howson R, et al. Phytozome: a comparative platform for green plant genomics[J]. Nucleic Acids Res, 2012, 40: 1178-1186
- [17] Markakis M N, Boron A K, Loock B V, et al. Characterization of a Small Auxin-Up RNA (SAUR)-Like gene involved in Arabidopsis thaliana development[J]. Plos One, 2013, 8(11): e82596
- [18] Liu R H, Meng J L. Map Draw: a microsoft excel macro for drawing genetic linkage maps based on given genetic linkage data[J]. Hereditas, 2003, 25(3): 317-321
- [19] Tek Chand B. Motif Design for Nitrilases[J]. Journal of Data Mining in Genomics & Proteomics, 2012, 3: 3
- [20] Kennedy J F, Knill C J. The Protein Protocols Handbook[J]. Bio-separation, 1998, 7(1): 61-62
- [21] Punta M, Ofra Y. The rough guide to in silico function prediction, or how to use sequence and structure information to predict protein function[J]. Plos Comput Biol, 2008, 4(10): e000160
- [22] Jensen L J, Gupta R, Blom N, et al. Prediction of human protein function from post-translational modifications and localization features[J]. J Mol Biol, 2002, 319(5): 1257-1265
- [23] Fasoli M, Santo D, Zenoni S, et al. The grapevine expression atlas reveals a deep transcriptome shift driving the entire plant into a maturation program[J]. Plant Cell, 2012, 24: 3489-3505
- [24] 刘克禄, 陈卫国. 植物耐热相关基因研究进展[J]. 植物遗传资源学报, 2015, 16(1): 127-132
- [25] Chen Y, Hao X, Cao J. Small auxin upregulated RNA (SAUR) gene family in maize: Identification, evolution, and its phylogenetic comparison with Arabidopsis, rice, and sorghum[J]. J integr plant biol, 2014, 56: 133-150
- [26] Rensing S A, Lang D, Zimmer A D, et al. The physcomitrella genome reveals evolutionary insights into the conquest of land by plants[J]. Science, 2008, 319(5859): 64-69
- [27] Wang S, Bai Y, Shen C, et al. Auxin-related gene families in abiotic stress response in Sorghum bicolor[J]. Funct Integr Genomic, 2010, 10(4): 533-546
- [28] 赵敬会, 张涛. 白菜 SAUR 基因家族的生物信息学分析[J]. 中国农学通报, 2012, 28(22): 130-137
- [29] 严涵薇. 玉米及其近缘物种无内含子基因的数据库构建与进化研究[D]. 合肥: 安徽农业大学, 2014
- [30] Jeffares D C, Penkett C J, Bahler J. Rapidly regulated genes are intron poor[J]. Trends Genet, 2008, 24(8): 375-378
- [31] Zhang J. Evolution by gene duplication: an update[J]. Trends Ecol Evol, 2003, 18: 292-298
- [32] Ohno S. Evolution by gene duplication[J]. Am J Hum Genet, 1971, 23(5): 541.