

玉米 SAUR 基因家族的鉴定与生物信息学分析

郭 栋^{1,2}, 杜 媚², 周宝元¹, 刘颖慧², 赵 明¹

(¹ 中国农业科学院作物科学研究所, 北京 100081; ² 河北北方学院, 张家口 075000)

摘要: 为研究玉米早期应答生长素基因 SAUR (Small auxin-up RNA) 家族, 本研究采用全基因组信息鉴定出 91 个玉米 SAUR 基因, 命名为 *ZmSAUR*。SAUR 家族成员基因结构、氨基酸特点、染色体定位及基因进化分析表明, SAUR 基因家族在染色体上呈现不均匀分布, 其中 2 号染色体上数量最多为 22 个, 基因的扩增模式为分散复制与片段复制。SAUR 基因家族具有相对保守的结构, 即包含 1 个保守的 Rna DNA 结构, SAUR 蛋白的 3D 结构含有 3 个 α 螺旋和 3 个 β 折叠。根据多物种 SAUR 蛋白进化树分析将其分为 9 个分支, 并分析发现玉米与物种相近的谷子聚在一起。这些信息为玉米 SAUR 基因家族功能分析奠定了一定的工作基础。

关键词: 玉米; 生长素; SAUR

Identification and Bioinformatics Analysis of Maize SAUR Gene Family

GUO Dong^{1,2}, DU Mei², ZHOU Bao-yuan¹, LIU Ying-hui², ZHAO Ming¹

(¹ Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing 100081;

² Hebei North University, Zhangjiakou 075000)

Abstract: In order to study the early auxin-up RNA family of maize, this study used genome-wide information to identify 91 maize SAUR genes, named *ZmSAUR*, based on the gene structure, amino acid characteristics, chromosomal location of SAUR family members. And genetic evolution analysis. The results showed that the SAUR gene family showed an uneven distribution on the chromosome, of which the number of chromosomes was up to 22, and the amplification pattern of the gene was dispersed replication and fragment replication. The SAUR gene family has a relatively conserved structure, ie contains a conserved Rna DNA structure, and the 3D structure of the SAUR protein contains three alpha helices and three beta sheets. According to the multi-species SAUR protein phylogenetic tree analysis, it was divided into 9 branches, and it was found that corn and clusters with similar species were clustered together. This information laid a foundation for the functional analysis of the maize SAUR gene family.

Key words: maize; auxin; SAUR

植物生长素广泛分布于各种植物中, 可以控制植物细胞分裂、扩张、分化等, 进而影响植物生长发育的各个方面^[1]。生长素早期应答基因由 GH3、AUX/IAA、SAUR3 基因家族构成, 其中 SAUR 与早期应答基因有密切的联系^[2]。GH3 对生长素信号途径与光信号途径都有重要作用, 同时对于防卫

反应也起到不同的作用^[3]; AUX/IAA 是转录抑制因子, 可以为生长素信号提供转导途径, 编码蛋白已被证实对生长素信号转导途径起着非常大的作用; SAUR 是早期应答基因, 是感应最迅速和最强烈的生长素基因, 许多植物编码中都存在有特异性的小分子蛋白质^[4-5]。SAUR 生长素在植物体内起

收稿日期: 2018-07-07 修回日期: 2018-07-23 网络出版日期: 2018-10-29

URL: <http://kns.cnki.net/kcms/detail/11.4996.S.20181025.1651.003.html>

第一作者研究方向为玉米栽培生理及生信研究, E-mail: 447114649@qq.com

通信作者: 赵明, 研究方向为玉米高效栽培利用研究, E-mail: zhaoming@caas.cn

基金项目: 国家重点研发计划 (2016YFD0300207, 2016YFD0300103)

Foundation project: National Key Research and Development Plan (2016YFD0300207, 2016YFD0300103)

到关键作用^[6],但只有小部分功能得到了鉴定。如玉米中生长素介导与 *ZmSAUR2* 有关^[7];拟南芥的 *SAUR63* 会让与生长素有所关联的基因来刺激不同器官的伸长,而 *SAUR41* 通过调节生长素使细胞进行转运^[8-9]; *SAUR39* 基因在水稻中起到生长素合成与运输的作用^[10]。

玉米 (*Zea mays* L.) 在我国具有悠久的栽培历史,是我国主要粮食产物之一。玉米是研究 C4 植物的模式植物。近年来,玉米基因组测序完成,为玉米基因组学的研究奠定了良好的基础,然而关于玉米 SAUR 基因家族的研究却相对少见。本研究利用生物信息学相关技术鉴定玉米的 SAUR 基因家族并对该家族基因结构进行分析,通过对其序列和蛋白的比较分析为玉米生长素相关基因的研究提供一定的参考。

1 材料与方法

1.1 玉米 SAUR 基因的鉴定及序列分析

首先利用 Pfam 数据库 (<http://pfam.xfam.org/>) 查找与下载 SAUR 蛋白结构域的隐马模型文件 (号码: PF00642)^[11],通过使用 Pfam 号码在 Gramene (<http://www.gramene.org/>) 在线网站进行相似性查找,找寻与玉米相关联的基因 ID 和有关联的蛋白序列,通过使用 SMART 与 CDD 在线网站对查找到的基因 ID 进行蛋白结构域的检查^[12-13],除去冗余。利用 Gramene 数据库得到所鉴定基因的相关信息,如蛋白 ID、染色体定位、蛋白序列以及外显子数量等信息。最后,用 ProtParam 在线网站获取玉米中 SAUR 蛋白的相关信息,包括分子量、氨基酸数量及等电点等。

1.2 染色体定位分析

通过 Ensemble Plants (<http://plants.ensembl.org/>) 在线网站查询玉米染色体的长度,并使用 MapChatr 软件进行绘制基因位置与染色体位置。

1.3 SAUR 蛋白系统分析与进化树的绘制

利用 ProSite (<http://prosite.expasy.org/>) 在线软件对 91 个 SAUR 蛋白序列进行分析,获得每个与 SAUR 蛋白相结合结构域位置^[14]。

通过 Clustal X 在线分析所得到的 SAUR 蛋白序列^[15],随后使用 MEGA6.0 软件功能中的邻接法绘制进化树,bootstrap 设置为 1000。并采用上述方法绘制不同物种间的 SAUR 蛋白进化树,包括玉米 (91 个)、水稻 (56 个)、二穗短柄草 (50 个)、小立碗藜 (17 个)、葡萄 (76 个)、拟南芥 (78 个) 和谷子 (73 个) 的共 441 个 SAUR 蛋白进行综合分析,采

用上述方法绘制 SAUR 蛋白进化树。

1.4 Motif 的获取和 SAUR 蛋白的 3D 结构预测

通过使用在线 MEME 工具分析获得玉米 SAUR 蛋白的 motif 基序^[16],随后使用 Weblogo 在线网获得 SAUR 的 3D 结合结构域的 motif。最后使用 SWISS—MODEoL 预测其三维结构^[17]。

2 结果与分析

2.1 玉米 SAUR 基因家族鉴定

通过 Pfam 数据库查找并下载 SAUR 基因家族中的隐马氏模型文件号码 (PF02519),搜索玉米基因组数据库,使用 SMART 和 CDD 网站进行蛋白序列结构域检测,除去冗余,最终得到 91 个玉米 SAUR 基因序列,命名为 *ZmSAUR1*~*ZmSAUR91* (表 1)。由表 1 可知,分析玉米 91 个 SAUR 蛋白序列发现不同蛋白差异很大:编码 SAUR 蛋白氨基酸长度范围从 79 aa (*ZmSAUR34*)~918 aa (*ZmSAUR88*),开放阅读框是从 240 bp (*ZmSAUR34*)~9300 bp (*ZmSAUR15*),分子量范围从 8.82 kD (*ZmSAUR77*)~101.18 kD (*ZmSAUR88*),等电点范围从 4.63 (*ZmSAUR81*)~11.8 (*ZmSAUR87*),蛋白含 1~18 个内含子。通过表 1 发现 SAUR 蛋白的特性变化与序列长度变化都很大。并且发现有 20 个蛋白等电点小于 7,说明部分 SAUR 蛋白编码弱酸性蛋白,会在酸性亚细胞的环境中发挥作用。

由表 1、图 1 可知,玉米的 91 个 SAUR 基因在 10 条染色体上呈现不均匀分布。其中 2 号染色体基因最多为 22 个,4 号、5 号染色体均为 12 个,1 号与 7 号染色体均为 11 个,10 号染色体 6 个,8 号与 9 号染色体各 5 个,3 号染色体 4 个,6 号染色体分布最少只有 3 个 SAUR 基因。在 Phytozome 数据库中利用 Gbrowse 进行查看成簇分布的 SAUR 基因家族成员与侧翼蛋白质编码基因的位置关系,同一成员如果位于同一个或相邻的基因间区域,则这些成员为串联重复关系,按照此标准,说明玉米 SAUR 基因家族扩增是无法通过串联重复来达到目的。

2.2 玉米 SAUR 蛋白的结构域分析

通过 ProSite 软件对每个玉米的 SAUR 蛋白结构域进行分析由图 2 得出,所有玉米 SAUR 都含有植物特异的 SAUR 结构域,同样大部分的序列都只含有 1 个 Rna 结构域,其中只有 *ZmSAUR1* 与 *ZmSAUR88* 含有 2 个结构域,分别是 Bhlh 结构域、Pkd 结构域。通过使用在线软件 MEME 对玉米的 SAUR 蛋白结构域分析,发现 SAUR 结构域的氨基酸组成较稳定。

表1 玉米 SAUR 基因家族的鉴定及特性

Table 1 The characteristics of identified SAUR gene family in maize

基因名 Gene name	转录名 Genomic accession No.	染色体 Chromosome distribution	位置 Location	外显子数 No. of extron	开放阅读框 长度 (bp) Open reading fragment length	推测蛋白质 Speculated protein		
						氨基酸 (aa) Amino acid	分子量 (kD) Molecular mass	等电点 Isoelectric point
<i>ZmSAUR1</i>	Zm00001 d027753_T001	1	12658925-12661298	6	2374	423	46.81	8.21
<i>ZmSAUR2</i>	Zm00001 d031146_T001	1	179600373-179600708	1	336	111	11.68	6.50
<i>ZmSAUR3</i>	Zm00001 d032088_T001	1	212207299-212207712	1	414	137	14.54	6.04
<i>ZmSAUR4</i>	Zm00001 d032091_T001	1	212248940-212249338	1	399	132	14.23	5.93
<i>ZmSAUR5</i>	Zm00001 d032094_T001	1	212314858-212315259	1	402	133	14.31	6.69
<i>ZmSAUR6</i>	Zm00001 d032475_T001	1	227990837-227991226	1	390	129	14.05	8.49
<i>ZmSAUR7</i>	Zm00001 d033460_T001	1	263312679-263313005	1	327	108	11.46	7.95
<i>ZmSAUR8</i>	Zm00001 d033462_T001	1	263334484-263334816	1	333	110	11.62	5.31
<i>ZmSAUR9</i>	Zm00001 d033463_T001	1	263437724-263438035	1	312	103	10.84	6.39
<i>ZmSAUR10</i>	Zm00001 d033464_T001	1	263441814-263442125	1	312	103	11.10	9.75
<i>ZmSAUR11</i>	Zm00001 d033729_T001	1	272362116-272364398	5	2283	366	39.21	5.01
<i>ZmSAUR12</i>	Zm00001 d001961_T001	2	3566924-3567634	1	711	236	25.51	10.73
<i>ZmSAUR13</i>	Zm00001 d001963_T001	2	3589962-3590393	1	432	143	15.58	11.07
<i>ZmSAUR14</i>	Zm00001 d001964_T001	2	3600255-3600683	1	429	142	15.37	9.36
<i>ZmSAUR15</i>	Zm00001 d002302_T001	2	9791380-9800679	8	9300	366	39.54	9.50
<i>ZmSAUR16</i>	Zm00001 d002304_T001	2	9802200-9802541	1	342	113	12.34	8.92
<i>ZmSAUR17</i>	Zm00001 d002374_T001	2	11368790-11369293	1	504	167	18.15	9.46
<i>ZmSAUR18</i>	Zm00001 d002468_T001	2	13393684-13393995	1	312	103	11.37	9.34
<i>ZmSAUR19</i>	Zm00001 d002826_T001	2	23922277-23922813	1	537	178	19.46	11.48
<i>ZmSAUR20</i>	Zm00001 d003018_T001	2	29562998-29563381	1	384	127	14.13	7.72
<i>ZmSAUR21</i>	Zm00001 d004384_T001	2	107710366-107710761	1	396	131	14.48	10.08
<i>ZmSAUR22</i>	Zm00001 d004578_T001	2	119261511-119268416	2	6906	124	13.62	8.46
<i>ZmSAUR23</i>	Zm00001 d005802_T001	2	189128516-189128986	1	471	156	16.39	7.74
<i>ZmSAUR24</i>	Zm00001 d005803_T001	2	189211330-189212136	2	807	181	20.15	5.17
<i>ZmSAUR25</i>	Zm00001 d006101_T001	2	198162667-198163122	1	456	151	16.22	9.59
<i>ZmSAUR26</i>	Zm00001 d006274_T001	2	203395420-203402757	2	7338	149	16.49	8.95
<i>ZmSAUR27</i>	Zm00001 d006275_T001	2	203397990-203399055	2	1066	137	15.23	8.18
<i>ZmSAUR28</i>	Zm00001 d006276_T001	2	203401394-203401852	1	459	152	17.10	8.89
<i>ZmSAUR29</i>	Zm00001 d006277_T001	2	203403246-203403638	1	393	130	13.87	6.27
<i>ZmSAUR30</i>	Zm00001 d006279_T001	2	203405094-203405564	1	471	146	15.71	8.66
<i>ZmSAUR31</i>	Zm00001 d006282_T001	2	203489297-203489769	1	473	152	16.43	8.09
<i>ZmSAUR32</i>	Zm00001 d006283_T001	2	203490871-203491346	1	476	144	15.30	6.93
<i>ZmSAUR33</i>	Zm00001 d006285_T001	2	203553961-203554453	1	493	149	16.06	8.67
<i>ZmSAUR34</i>	Zm00001 d039791_T001	3	14788097-14788336	1	240	79	9.06	8.89
<i>ZmSAUR35</i>	Zm00001 d041462_T001	3	121836564-121836986	1	423	140	15.21	6.18
<i>ZmSAUR36</i>	Zm00001 d042292_T001	3	159333398-159333859	1	462	153	16.23	9.24
<i>ZmSAUR37</i>	Zm00001 d044325_T001	3	225261649-225261975	1	327	108	11.93	6.89
<i>ZmSAUR38</i>	Zm00001 d049659_T001	4	38969055-38969360	1	306	101	10.81	8.62
<i>ZmSAUR39</i>	Zm00001 d050093_T001	4	65383709-65384116	1	408	135	14.51	8.93

表 1(续)

基因名 Gene name	转录名 Genomic accession No.	染色体 Chromosome distribution	位置 Location	外显子数 No. of extron	开放阅读框 长度 (bp) Open reading fragment length	推测蛋白质 Speculated protein		
						氨基酸 (aa) Amino acid	分子量 (kD) Molecular mass	等电点 Isoelectric point
<i>ZmSAUR40</i>	Zm00001 d050507_T001	4	94918988-94919425	1	438	145	16.02	9.13
<i>ZmSAUR41</i>	Zm00001 d051127_T001	4	144744938-144745333	1	396	131	14.64	9.73
<i>ZmSAUR42</i>	Zm00001 d051302_T001	4	152451610-152452209	1	600	199	21.92	10.68
<i>ZmSAUR43</i>	Zm00001 d051802_T001	4	170029713-170030060	1	348	115	12.55	9.02
<i>ZmSAUR44</i>	Zm00001 d051920_T001	4	174649398-174649715	1	318	105	11.85	8.52
<i>ZmSAUR45</i>	Zm00001 d051962_T001	4	175386555-175387938	3	1384	160	17.73	5.23
<i>ZmSAUR46</i>	Zm00001 d052148_T001	4	181202168-181202603	1	436	117	12.86	9.51
<i>ZmSAUR47</i>	Zm00001 d053311_T001	4	225550062-225551773	4	1712	342	37.27	7.20
<i>ZmSAUR48</i>	Zm00001 d053815_T001	4	241609820-241610200	1	381	126	13.58	11.02
<i>ZmSAUR49</i>	Zm00001 d053884_T001	4	242957976-242958374	1	399	132	14.71	9.18
<i>ZmSAUR50</i>	Zm00001 d013142_T001	5	5377049-5377345	1	297	98	11.19	10.89
<i>ZmSAUR51</i>	Zm00001 d013616_T001	5	15567224-15567550	1	327	108	11.43	8.71
<i>ZmSAUR52</i>	Zm00001 d013617_T001	5	15578282-15578608	1	327	108	11.43	8.71
<i>ZmSAUR53</i>	Zm00001 d013619_T001	5	15654693-15655013	1	321	106	11.07	6.82
<i>ZmSAUR54</i>	Zm00001 d013869_T001	5	23456902-23457294	1	393	130	14.47	7.90
<i>ZmSAUR55</i>	Zm00001 d014682_T001	5	58760829-58761221	1	393	130	14.38	9.39
<i>ZmSAUR56</i>	Zm00001 d014774_T001	5	62765407-62765844	1	438	145	15.25	7.90
<i>ZmSAUR57</i>	Zm00001 d015354_T001	5	85946140-85946524	1	385	126	13.57	10.79
<i>ZmSAUR58</i>	Zm00001 d015661_T001	5	104161775-104162116	1	342	113	12.10	9.69
<i>ZmSAUR59</i>	Zm00001 d016582_T001	5	168259095-168259583	1	489	162	17.36	9.12
<i>ZmSAUR60</i>	Zm00001 d017397_T001	5	194705728-194706324	1	597	198	21.83	11.49
<i>ZmSAUR61</i>	Zm00001 d018200_T001	5	216447005-216447600	1	596	122	13.40	9.64
<i>ZmSAUR62</i>	Zm00001 d036415_T001	6	87428736-87429128	1	393	130	13.62	7.88
<i>ZmSAUR63</i>	Zm00001 d036463_T001	6	89312912-89313355	1	444	147	16.58	6.82
<i>ZmSAUR64</i>	Zm00001 d036623_T001	6	95057231-95057635	1	405	134	13.99	9.15
<i>ZmSAUR65</i>	Zm00001 d020605_T001	7	124482034-124482507	1	474	157	16.39	8.41
<i>ZmSAUR66</i>	Zm00001 d020606_T001	7	124627553-124628462	2	910	193	21.18	5.91
<i>ZmSAUR67</i>	Zm00001 d020782_T001	7	132495245-132495577	1	333	110	12.00	9.18
<i>ZmSAUR68</i>	Zm00001 d021062_T001	7	141857876-141858286	1	411	136	14.59	9.54
<i>ZmSAUR69</i>	Zm00001 d021237_T001	7	146666554-146668407	5	1854	284	31.02	5.44
<i>ZmSAUR70</i>	Zm00001 d021337_T001	7	149226369-149226713	1	345	114	12.29	9.69
<i>ZmSAUR71</i>	Zm00001 d021454_T001	7	152609402-152609851	1	450	149	15.98	8.88
<i>ZmSAUR72</i>	Zm00001 d021455_T001	7	152642179-152642637	1	459	152	16.00	8.88
<i>ZmSAUR73</i>	Zm00001 d021456_T001	7	152644791-152645267	1	477	158	16.78	9.58
<i>ZmSAUR74</i>	Zm00001 d021457_T001	7	152649238-152649669	1	432	143	15.43	8.93
<i>ZmSAUR75</i>	Zm00001 d021459_T001	7	152719949-152720386	1	438	145	16.40	8.23
<i>ZmSAUR76</i>	Zm00001 d008420_T001	8	8273413-8275055	4	1643	235	26.01	5.84
<i>ZmSAUR77</i>	Zm00001 d010775_T001	8	127477547-127477795	1	249	82	8.82	7.96
<i>ZmSAUR78</i>	Zm00001 d010776_T001	8	127492890-127493392	2	503	124	13.72	11.03
<i>ZmSAUR79</i>	Zm00001 d011978_T001	8	165552780-165558195	2	5416	183	19.22	5.46

表 1(续)

基因名 Gene name	转录名 Genomic accession No.	染色体 Chromosome distribution	位置 Location	外显子数 No. of extron	开放阅读框 长度 (bp) Open reading fragment length	推测蛋白质 Speculated protein		
						氨基酸 (aa) Amino acid	分子量 (kD) Molecular mass	等电点 Isoelectric point
ZmSAUR80	Zm00001 d012222_T001	8	170385558-170386010	1	453	150	16.83	9.09
ZmSAUR81	Zm00001 d045039_T001	9	10832838-10834164	2	1327	236	25.72	4.63
ZmSAUR82	Zm00001 d045423_T001	9	21633058-21633435	1	378	125	13.60	9.29
ZmSAUR83	Zm00001 d046986_T001	9	114297761-114298180	1	420	139	14.39	8.33
ZmSAUR84	Zm00001 d046993_T001	9	114347644-114347928	1	285	94	10.33	9.49
ZmSAUR85	Zm00001 d048316_T001	9	154456440-154456811	1	372	123	13.42	5.62
ZmSAUR86	Zm00001 d023518_T001	10	8597735-8598124	1	390	129	13.54	8.88
ZmSAUR87	Zm00001 d025947_T001	10	134516141-134516668	1	528	175	19.26	11.80
ZmSAUR88	Zm00001 d026246_T001	10	141781998-141789379	18	7382	918	101.18	7.26
ZmSAUR89	Zm00001 d026262_T001	10	142339502-142340011	1	510	169	18.50	9.36
ZmSAUR90	Zm00001 d026308_T001	10	143309469-143309843	1	375	124	12.50	8.61
ZmSAUR91	Zm00001 d026530_T001	10	147743187-147743636	1	450	149	16.08	9.54

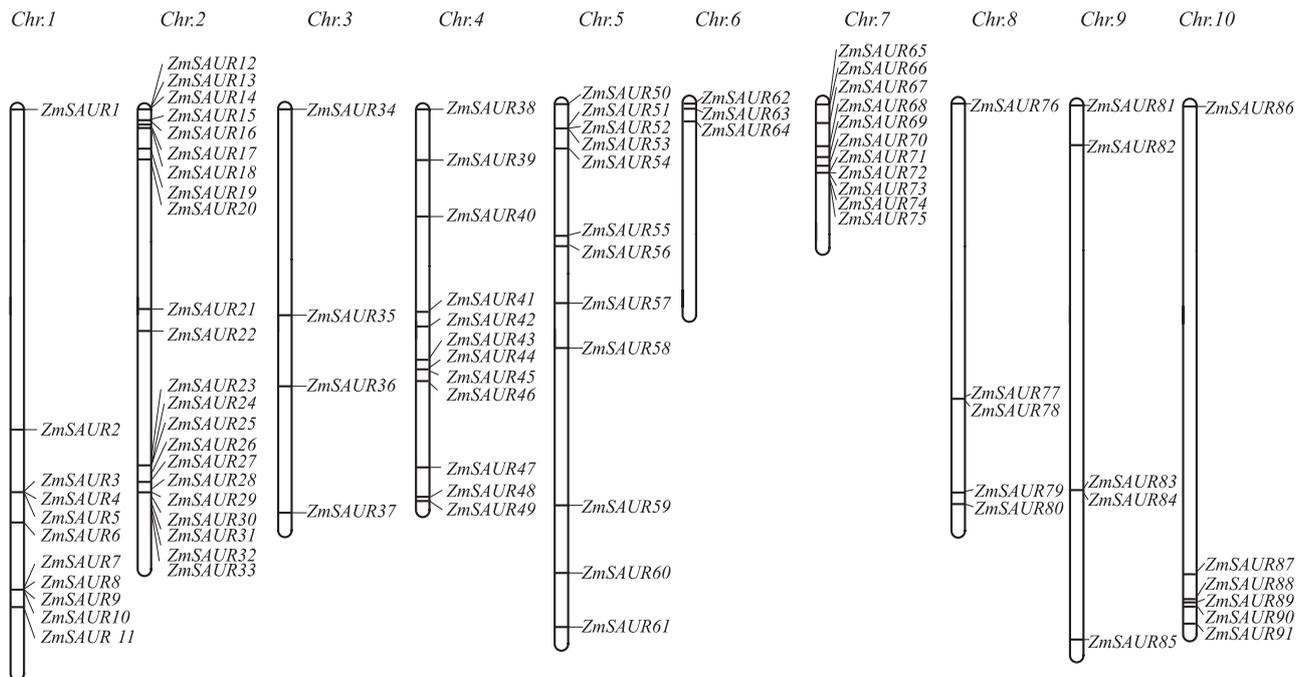


图 1 玉米 SAUR 基因的染色体分布

Fig. 1 Chromosomal distribution of maize SAUR gene

2.3 玉米 SAUR 基因家族分子进化树的构建

对 SAUR 蛋白序列进行比对, 绘制出玉米 SAUR 进化树(图 3)。据进化树分支, 可以将这些蛋白分为 3 组, 其中 I 组的 SAUR 蛋白有 6 个保守结构域分别: ZmSAUR58、ZmSAUR70、ZmSAUR41、ZmSAUR44、ZmSAUR46、ZmSAUR61、ZmSAUR54、ZmSAUR6、ZmSAUR56、ZmSAUR62、ZmSAUR24、ZmSAUR66; II 组

有 3 个保守结构域分别: ZmSAUR77、ZmSAUR78、ZmSAUR16、ZmSAUR18、ZmSAUR51、ZmSAUR52。从 3 组蛋白结构域可知 I 组的蛋白数量最多, 约占总数的 57%(图 3)。通过进化树的聚类分析可以得出, 相同蛋白结构域与相同基因扩增或者复制的蛋白相聚在一起, 如: ZmSAUR58 和 ZmSAUR70 分散复制聚在一起, 具有相同结构域的聚为一组。

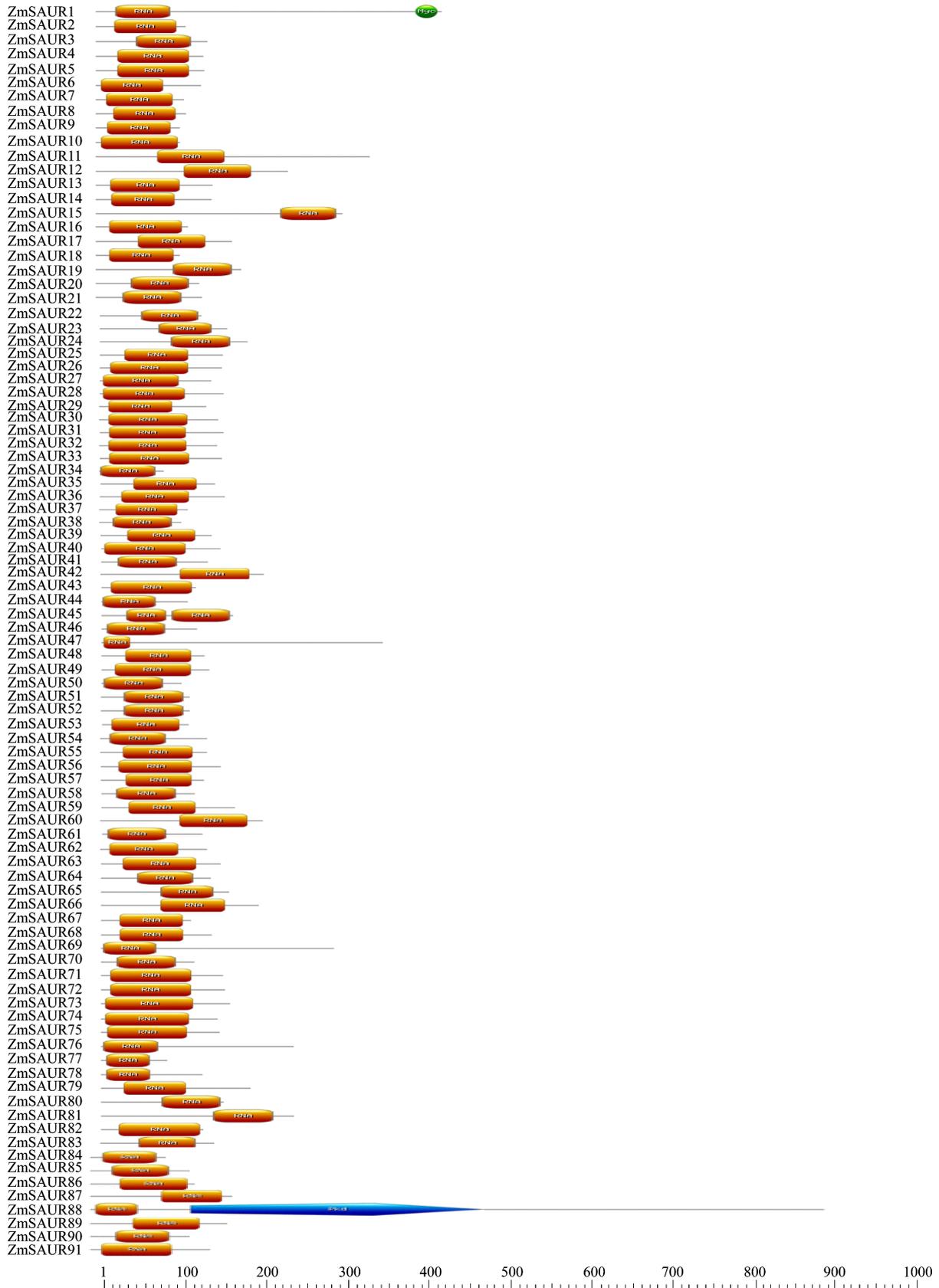


图 2 玉米 SAUR 蛋白的结构域示意图

Fig. 2 Schematic diagram of the domain of maize SAUR protein

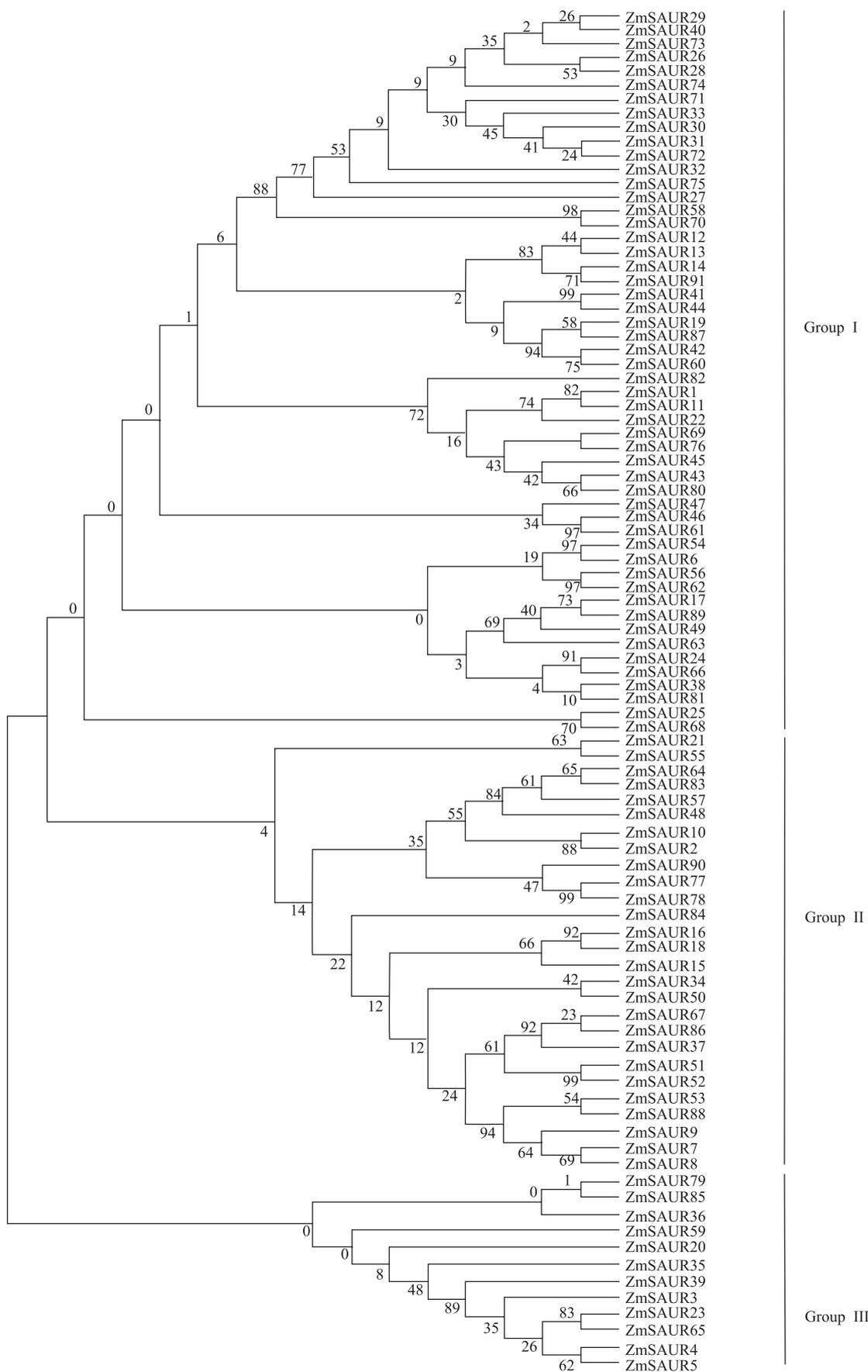
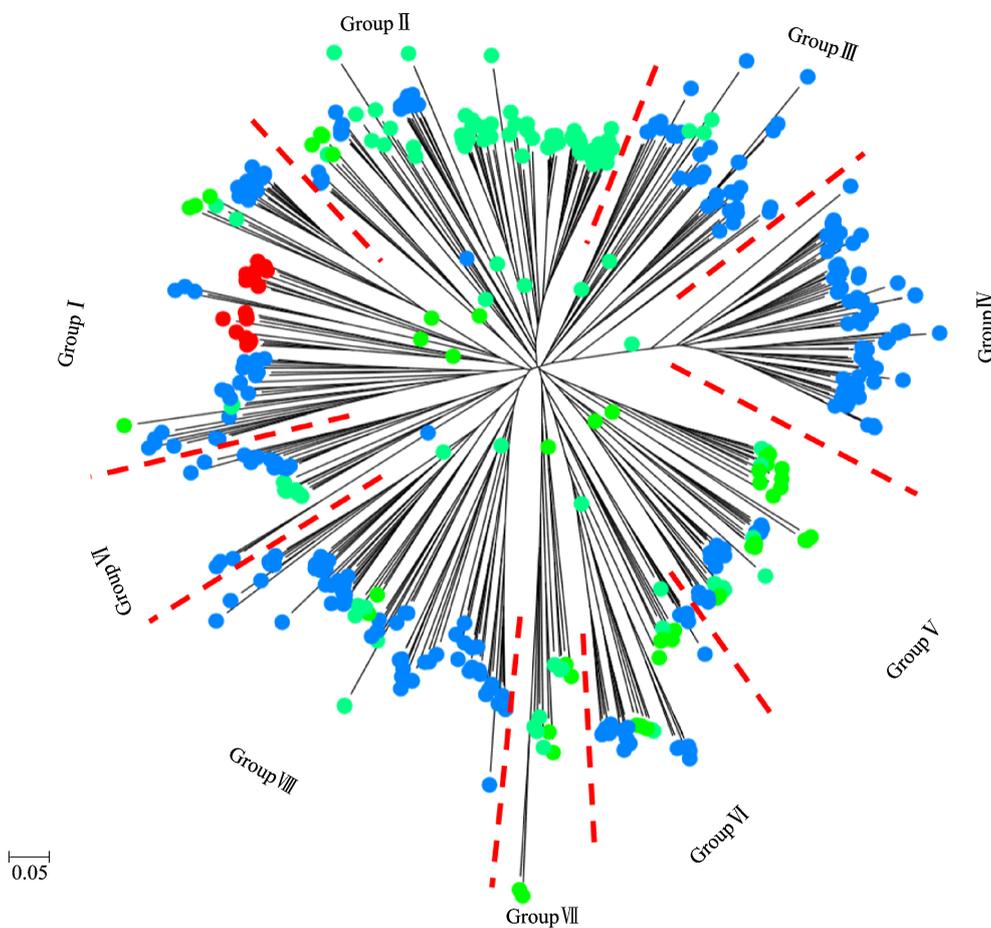


图3 玉米 SAUR 蛋白的系统发育进化树
 Fig. 3 Phylogenetic tree of maize SAUR protein

2.5 不同植物 SAUR 蛋白进化关系

为充分揭示玉米 SAUR 基因家族进化的关系,利用谷子(73 个)、葡萄(76 个)、拟南芥(78 个)、二穗短柄草(50 个)、小立碗藓(17 个)及水稻(56 个)进行绘制 SAUR 蛋白构建系统发育树(图 4)

共 441 个, SAUR 蛋白聚为 9 个组,主要是以亲缘关系与进化时间为聚类,其中单子叶植物、双子叶植物与苔藓类植物中的 SAUR 蛋白为家族性聚集,植物间不同纲目的 SAUR 蛋白同样具有较高同源性。



蓝色为单子叶植物;绿色为双子叶;红色为苔藓
Blue is monocotyledon, green is dicotyled, red is moss

图 4 多物种 SAUR 蛋白进化树

Fig. 4 Multi-species SAUR protein phylogenetic tree

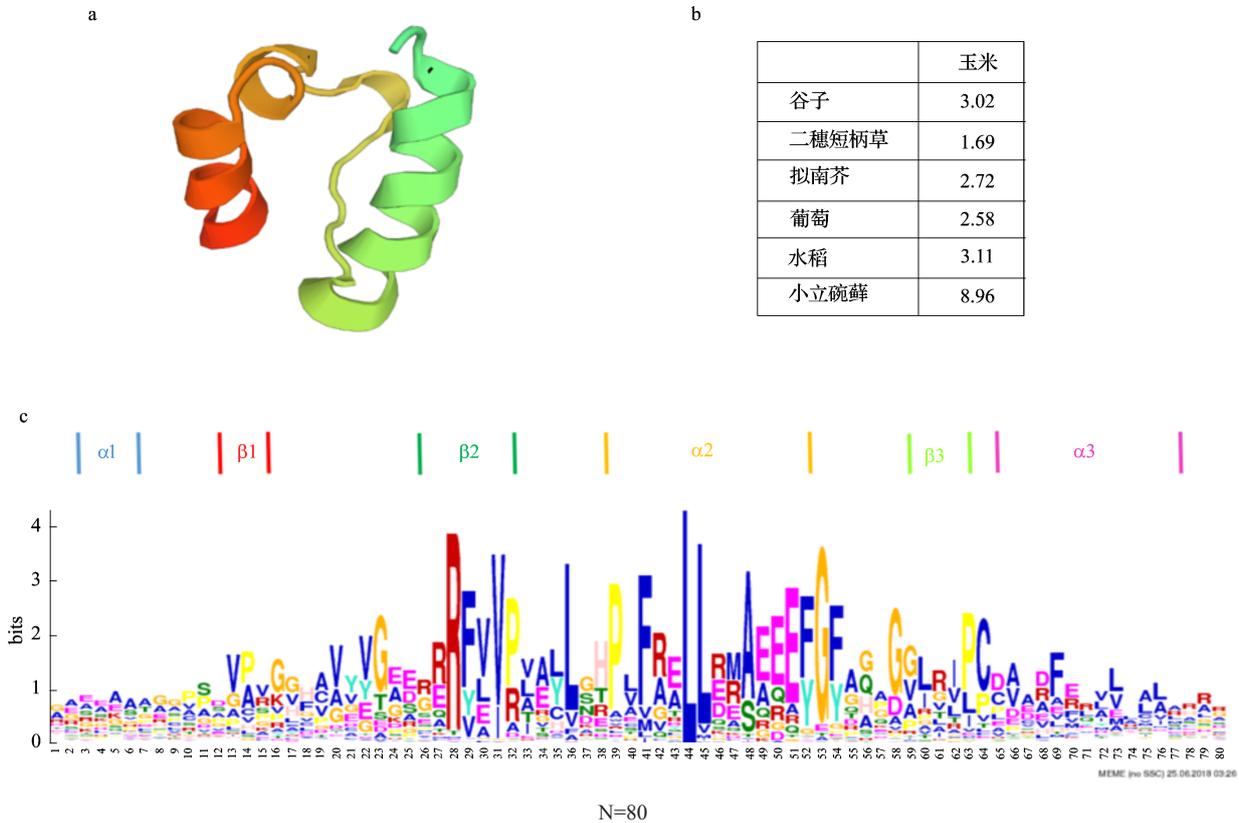
2.6 玉米 SAUR 蛋白 3D 结构与序列分析

通过利用 SWISS-MODEL 软件对 SAUR 蛋白进行构建 Rna 3D 结构的 DNA 结构域。结果发现, SAUR 蛋白的 Rna DNA 结构域内有 3 个 α 螺旋和 3 个 β 折叠。3 个 β 折叠分别位于 α_1 、 α_2 与 α_3 螺旋之间(图 5)。通过 β 折叠与连接形状的结构,可以看出 DNA 的大沟结合留下大的空隙。同时玉米 SAUR 蛋白的 3D 结构域与其他物种相应的 3D 结构域进行比对,如图 5 显示,比值越小相似度越相近,如单子叶植物二穗短柄草(1.69),其次是双子叶植物葡萄(2.58),其相差最远的是葫芦藓科中的小立碗藓(8.96),SAUR 蛋白中的 3D 结构与物种的

亲缘关系是具有相关性,即表明相似性越高亲缘关系越近。

3 讨论

SAUR 基因家族是生长素信号转导的重要转录因子,对其研究也备受关注,前人已经对许多植物进行了 SAUR 的全基因组分析,如玉米、高粱、水稻、番茄与拟南芥,其中水稻含有 58 个 SAUR 基因,拟南芥与番茄分别含有 72 个与 55 个 SAUR 家族成员,如今在玉米中鉴定出 91 个 SAUR 家族成员^[18-21],但目前对于玉米 SAUR 基因家族的系统进化与基因结构没有全面解开,因此本研究对玉米 SAUR



a: SAUR 3D 结构蛋白; b: SAUR 蛋白与其他物种的 3D 结构比对;
c: SAUR 蛋白的 3D 结构域 (N 代表氨基酸数量)
a: SAUR 3D structural protein; b: alignment of SAUR protein with other species 3D structures;
c: 3D domain of SAUR protein (N represents the number of amino acids)

图 5 玉米 SAUR 蛋白的 Rna 结构域分析
Fig. 5 Rna domain analysis of maize SAUR protein

基因家族分析,为进一步了解玉米基因信息作出了贡献。

通过染色体定位分析发现,玉米 SAUR 基因呈簇状分布,这可能与家族的重复基因有关,同样基因的扩张动力主要来自于基因重复^[22-23]。与本研究玉米的基因扩张相一致,基因家族数量增加主要是以基因重复为主要增加方式,基因家族中普遍有相似基因,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,说明基因重复是玉米 SAUR 基因扩张的主要动力和特点,基因重复在植物演化过程中扮演着重要角色,同样在生物学中普遍存在。通过 MEME 查询以及分析表明,发现玉米 SAUR 基因家族是具有很强保守型的蛋白结构域。通过保守基序与进化树的比对,可以得出玉米 SAUR 蛋白在进化形成新的一族时,玉米的 SAUR 蛋白保守基序也随变化而变化。

本研究通过生物信息学和基因组学得出玉米 SAUR 基因的功能和结构,可以为今后研究玉米与单子叶植物基因家族的提供重要的理论依据。

参考文献

- [1] Ren H, Gray W M. SAUR proteins as effectors of hormonal and environmental signals in plant growth. *Molecular Plant*, 2015, 8 (8): 1153-1164
- [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. *安徽农学通报*, 2008 (7): 36-39
Li Y N, Feng X, Chen D Q. Regulation of gene expression by ARF, Aux/IAA and auxin receptors. *Anhui Agricultural Science Bulletin*, 2008 (7): 36-39
- [3] Gretchen H, Gary M, Yi L, Tom J, Guilfoyle. Auxin-induced expression of the soybean GH3 promoter in transgenic tobacco plants. *Plant Molecular Biology*, 1991, 17 (3): 567-579
- [4] Spartz A K, Lee S H, Wenger J P. The Auxin-responsive gene expression: genes, promoters and regulatory factors of *SMALL AUXIN UP RNA* genes promote cell expansion. *The Plant Journal*, 2012, 70 (3): 978-990
- [5] Hagen G, Guilfoyle T. Auxin-responsive gene expression: genes, promoters and regulatory factors. *Plant Molecular Biology*, 2002, 49 (34): 373-385
- [6] Ren H, Gray W M. SAUR proteins as effectors of hormonal and environmental signals in plant growth. *Molecular Plant*, 2015, 8 (8): 1153-1164
- [7] Knauss S, Rohrmeier T, Lehle L. The auxin-induced maize gene *ZmSAUR2* encodes a short-lived nuclear protein expressed in elongating tissues. *Journal of Biological Chemistry*, 2003, 278

- (5): 23936-23943
- [8] Chae K, Isaacs C G, Reeves P H, Maloney G S, Muday G K, Reed J W. *Arabidopsis SMALL AUXIN UP RNA63* promotes hypocotyl and stamen filament elongation. *The Plant Journal*, 2012, 71 (4): 648-697
- [9] Kong Y, Zhu Y, Gao C. Tissue specific expression of *SMALL AUXIN UP RNA41* differentially regulates cell expansion and root meristem patterning in *Arabidopsis*. *Plant and Cell Physiology*, 2013, 54 (8): 609-621
- [10] Kant S, Bi Y M, Zhu T, Rothstein S J. *SAUR39*, a small Auxin-up RNA gene, acts as a negative regulator of auxin synthesis and transport in Rice. *Plant Physiology*, 2009, 151 (2): 691-701
- [11] Finn R D, Mistry J, Tate J, Coggill P C, Sammut S J, Hotz H R, Ceric G, Forslund K, Eddy S R, Sonnhammer E L L, Bateman A. The Pfam protein families database. *Nucleic Acids Research*, 2008, 36 (1): 281-288
- [12] Schultz J, Milpetz F, Bork P, Ponting C P. SMART, a simple modular architecture research tool: Identification of signaling domains. *Proceedings of the National Academy of Sciences of the United States of America*, 1998, 95 (11): 5857-5864
- [13] Marchler-Bauer A, Anderson J B, Chitsaz F, Derbyshire M K, Fong J H, Geer L Y, Geer R C, Gonzales N R, Gwadz M, He S, Hurwitz D I, Jackson J D, Ke Z X, Gabriele H, Mullokandov M, Song J S, Tasneem A, Thanki N, Yamashita R A, Zhang D, Zhang N, Bryant H S. CDD: specific functional annotation with the Conserved Domain Database. *Nucleic Acids Research*, 2009, 37 (1): 205-210
- [14] Sigrist C J, Cerutti L, de Castro E, Bulliard V, Bairoch A, Hulo N. PROSITE, a protein domain database for functional characterization and annotation. *Nucleic Acids Research*, 2010, 38 (1): 161-166
- [15] Larkin M A, Blackshields G, Brown N P, Chenna R, McGettigan P A, McWilliam H, Valentin F, Wallace I M, Wiln A, Lopez R, Thompson J D, Gibson T J, Higgins D G. Clustal W and Clustal X version 2.0. *Bioinformatics*, 2007, 23 (21): 2947-2948
- [16] Bailey T L, Boden M, Buske F A, Frith M, Grant C E, Clementi L, Ren J, Li W, Noble W S. MEME-SULTE: tools for motif discovery and searching. *Nucleic Acids Research*, 2009, 37 (1): 202-208
- [17] Kiefer F, Arnold K, Bordoli L, Schwede T. The SWISS-MODEL repository and associated resources. *Nucleic Acids Res*, 2009, 37 (1): 387-392
- [18] Li Y, Gretchen H, Tom J, Guilfoyle A. An Auxin responsive promoter is differentially induced by auxin gradients during tropisms. *The Plant Cell*, 1991, 3 (11): 1167-1175
- [19] 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*). *Genomics*, 2006, 88 (3): 360-371
- [20] Wu J, Liu S, He Y, Guan X Y, Zhu X F, Cheng L, Wang J, Lu G. Genome-wide analysis of SAUR gene family in Solanaceae species. *Gene*, 2012, 509 (1): 38-50
- [21] 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. *Journal of Integrative Plant Biology*, 2014, 56 (2): 133-150
- [22] Zhang J. Evolution by gene duplication: an update. *Trends in Ecology and Evolution*, 2003, 18 (6): 292-298
- [23] 李傲, 崔梦杰, 陈珂, 许瀛之, 贾海峰, 房经贵. 葡萄 SAUR 基因家族鉴定与生物信息学分析. *植物遗传资源学报*, 2018, 19 (2): 326-337
- Li A, Cui M J, Chen W, Xu W Z, Jia H F, Fang J G. Identification and Bioinformatics Analysis of Grape SAUR Gene Family. *Journal of Plant Genetic Resources*, 2018, 19 (2): 326-337