

• 药材与资源 •

远志 *bZIP* 基因家族的鉴定及表达分析

净易尧¹, 罗 瑶¹, 史嘉周¹, 陈晓颖¹, 党小琳², 张 岗¹, 高 静¹, 杨冰月¹, 胡本祥^{1,2*}, 彭 亮^{1*}

1. 陕西中医药大学 药学院/陕西省秦岭中草药应用开发工程技术研究中心/“秦药”研发重点实验室, 陕西 西安 712046

2. 陕西国际商贸学院, 陕西 西安 712046

摘要: 目的 鉴定远志 *Polygala tenuifolia bZIP* 基因, 为远志抗逆性改良及次生代谢调控研究提供参考。方法 基于远志三代全长转录组数据库, 采用生物信息学方法对远志 *bZIP* 基因家族成员进行鉴定和分析, 并利用实时荧光定量 PCR (qRT-PCR) 技术分析其在不同组织及处理中的表达特性。结果 共鉴定得到 27 个含有特征性保守结构域的 *PtbZIP* 基因, 编码蛋白质的氨基酸数量为 143 aa (*PtbZIP24*) ~ 846 aa (*PtbZIP9*), 相对分子质量范围为 16 201.52 ~ 92 932.30, 等电点 (pI) 介于 4.59 ~ 9.69, 其中 25 个家族成员为不稳定蛋白, 所有家族成员均为亲水性蛋白。蛋白二级结构主要由 α 螺旋和无规卷曲构成, 均无信号肽, 存在多种互作现象。进化树分析将 27 个 *PtbZIP* 蛋白分为 A、B、C、D、F、G、I、S 8 个亚组, 其中 G 亚组含有 *PtbZIP* 家族成员数量最多 (共有 8 个), 占总数的 29.63%, 没有 *PtbZIP* 基因被归类到 E 和 K 组。*PtbZIP* 家族基因的密码子偏好性较弱, 适应性较弱。表达模式分析显示, *PtbZIP4/15* 在叶中的表达量最高, 茎和根次之; *PtbZIP8/24* 在茎中的表达量最高, 叶和根次之; *PtbZIP1/17* 的表达量为根 > 叶 > 茎; 其余 21 个的表达模式均为根 > 茎 > 叶。qPCR 结果表明, *PtbZIP26* 基因受脱落酸、壳聚糖的诱导, 且能显著应答干旱和盐胁迫。结论 鉴定了远志 *bZIP* 家族基因及其分子特征, 为进一步研究 *PtbZIP* 基因在调控远志发育及次生代谢物合成方面的生物学功能奠定基础。

关键词: 远志; *bZIP* 基因家族; 生物信息学; 表达分析; 实时荧光定量 PCR

中图分类号: R286.12 文献标志码: A 文章编号: 0253-2670(2024)12-4140-10

DOI: 10.7501/j.issn.0253-2670.2024.12.021

Identification and expression analysis of *bZIP* gene family in *Polygala tenuifolia*

JING Yiyao¹, LUO Yao¹, SHI Jiazhou¹, CHEN Xiaoying¹, DANG Xiaolin¹, ZHANG Gang¹, GAO Jing¹, YANG Bingyue¹, HU Benxiang^{1,2}, PENG Liang¹

1. Key Laboratory of “Qin medicine” Research and Development, Shaanxi Qinling Engineering Research Center for Application Development of Chinese Herbal Medicine, College of Pharmacy, Shaanxi University of Chinese Medicine, Xi’an 712046, China

2. Shaanxi Institute of International Trade & Commerce, Xi’an 712046, China

Abstract: Objective To identify the *bZIP* gene in *Polygala tenuifolia* and provide a reference for the study of *P. tenuifolia* stress tolerance improvement and secondary metabolism regulation. **Methods** Based on the three-generation full-length transcriptome database of *P. tenuifolia*, a bioinformatics approach was used to identify and analyze the members of the *bZIP* gene family of *P. tenuifolia*, and their expression characteristics in different tissues and treatments were analyzed by using real-time fluorescence quantitative PCR (qRT-PCR) technology. **Results** A total of 27 *PtbZIP* genes containing characterized conserved structural domains were identified, the number of amino acid encoding proteins ranged from 143 aa (*PtbZIP24*) to 846 aa (*PtbZIP9*), relative molecular weight masses ranged from 16 201.52 to 92 932.30, and isoelectric points (pI) ranged from 4.59 to 9.69, a total of 25 family members were unstable proteins and all family members were hydrophilic proteins. The protein secondary structure mainly consisted of α -helices and random coiled-coils, both without signal peptides, and there were multiple

收稿日期: 2023-12-03

基金项目: 国家自然科学基金资助项目 (82003899); 陕西省科技厅项目 (2024SF-YBXM-457, S2024-JC-QN-1808, 2023-YBSF-036); 陕西省教育厅一般专项 (22JK0279); 榆林市科技局项目 (YF-2021-74)

作者简介: 净易尧 (2000—) 女, 在读硕士研究生, 从事中药资源质量评价与分子生药学研究。E-mail: 2502849506@qq.com

*通信作者: 胡本祥 (1960—) 男, 教授, 从事中药资源评价与开发利用。E-mail: hbx800823@126.com

彭 亮 (1985—) 男, 教授, 从事中药资源评价与开发利用、分子生药学研究。E-mail: ppengliang@126.com

interaction phenomena. Evolutionary tree analysis classified the 27 *PtbZIP* proteins into eight subgroups, A, B, C, D, F, G, I and S. Among them, subgroup G contained the highest number of *PtbZIP* family members totaling eight, accounting for 29.63% of the total, and no *PtbZIP* genes were categorized to groups E and K. The *PtbZIP* family genes had weak codon preferences and were less adaptive. Expression pattern analysis showed that *PtbZIP4/15* had the highest expression in leaves, followed by stems and roots; *PtbZIP8/24* had the highest expression in stems, followed by leaves and roots; *PtbZIP1/17* had the expression of roots > leaves > stems; and the rest of the 21 had the expression pattern of roots > stems > leaves. The qPCR results showed that the *PtbZIP26* gene was induced by abscisic acid and chitosan, and could significantly respond to drought and salt stress.

Conclusion In this study, we identified the *P. tenuifolia* *bZIP* family genes and their molecular characterization, laying the foundation for further studies on the biological functions of *PtbZIP* genes in the regulation of *P. tenuifolia* development and secondary metabolite synthesis.

Key words: *Polygala tenuifolia* Willd.; *bZIP* gene family; bioinformatics; expression analysis; qRT-PCR

远志系陕西道地药材,是“秦药”大宗道地药材品种之一^[1]。《中国药典》2020年版所收录的远志为远志科(Polygalaceae)植物远志 *Polygala tenuifolia* Willd.或卵叶远志 *P. sibirica* L.的干燥根^[2],具有镇静安神、祛痰开窍、解毒消肿等功效^[3]。现代研究表明,远志的主要活性成分有皂苷类、寡糖酯类、吡啶类等,具有抗记忆障碍、保护中枢神经系统、抗抑郁、抗心肌缺血和抗肿瘤等作用^[4]。目前,关于远志的研究多集中于含量研究^[5]、活性测定^[6]、遗传多样性分析等^[7]。随着分子生物学的发展,对药用植物相关活性成分生物合成途径相关调控基因、转录因子的挖掘已成为研究热点,基因组学、转录组学等技术,在远志上的成功应用,也为远志基因家族的筛选、鉴定与分析提供了技术支撑和数据基础^[8]。

碱性亮氨酸拉链(*bZIP*)基因家族作为真核生物中转录网络的重要开关,是植物中最大的转录因子家族之一。*bZIP*结构域由两个区域组成,即DNA结合基本区和亮氨酸拉链区^[9]。*bZIP*基因家族成员通过差异基因网络或生物过程,在调节植物发育、生长以及盐胁迫响应等方面发挥着重要作用^[10]。研究表明,拟南芥 *Arabidopsis thaliana* L.、番茄 *Solanum lycopersicum* L.、黄瓜 *Cucumis sativus* L.、李子 *Prunus salicina* L.和蓖麻 *Ricinus communis* L.等多种植物中的 *bZIP* 参与调控组织分化、细胞生长、糖代谢、生物和非生物胁迫等多个生物学过程^[11-12]。*bZIP* 基因家族成员还参与多种药用植物次生代谢产物合成调控,如丹参 *Salvia miltiorrhiza* Bunge.的 *SmbZIP1* 基因可抑制丹参酮的积累,大豆 *Glycine max* (Linn.) Merr.的 *GmbZIP123* 基因则参与大豆种子脂质积累的调控^[13]。同时, *bZIP* 表达受外源激素和胁迫诱导,壳聚糖处理葡萄 *Vitis vinifera* L.12 h

下,其 *VvLysM8* 和 *VvLysM9* 基因表达量显著提高^[14],糜子 *Panicum miliaceum* L.中的 *PmbZIP97* 不仅受到脱落酸(abscisic acid, ABA)、盐和干旱胁迫强烈诱导且参与调控萌发后的根系生长^[15]。

本实验利用远志三代转录组数据,以 *bZIP* 基因家族为研究对象,对其基因家族进行成员鉴定和生物信息学分析,并确定其在远志中的结构特点与进化特征,进一步通过实时荧光定量分析其在不同组织、不同处理条件下的表达模式,为后续深入研究 *bZIP* 的生物学功能奠定基础,同时为 *bZIP* 家族可能参与远志次生代谢成分生物合成途径研究提供思路。

1 材料及仪器

1.1 材料

2021年10月于陕西中医药大学药用植物园(陕西咸阳)采集3年生远志 *Polygala tenuifolia* Willd.及其成熟种子,经陕西中医药大学杨新杰副教授鉴定。选取5株三年生长势均匀的远志植株,将根、茎、叶等量混合后进行全长转录组测序分析。

1.2 试剂及仪器

ABA、壳聚糖(chitosan, CHT)均购自上海源叶生物科技有限公司,Trizol总RNA提取试剂盒、dd H₂O均购自生工生物工程(上海)股份有限公司, TB Green[®]Premix ExTaq[™] II (TliRNaseH Plus)、PrimeScript[™] II 1st strand cDNA Synthesis Kit 购自 TaKaRa 公司(日本),所用引物由武汉金开瑞生物公司合成。

StepOnePlus[™] Real-Time PCR (qPCR) 仪(美国 Applied Biosystems 公司), NanoDrop[™] 2000 分光光度计(美国 Thermo-fisher 公司), K5800 自动检测超微量分光光度计(凯奥公司), -80 °C 超低温冰箱(中科美菱公司)。

2 方法

2.1 样品的处理

选择大小均一, 颗粒饱满的远志种子, 用自来水冲洗 1 d, 10%双氧水消毒, 播种于装有泥炭土的花盆中, 在光周期 16/8 h, 光照强度 9 000 Lx 条件下培养^[16]。选取长势均一的 2 月幼苗, 喷 200 $\mu\text{mol/L}$ ABA、200 $\mu\text{mol/L}$ CTS, 干旱 (10% PEG 6000)、盐 (100 mmol/L NaCl) 20 mL, 以无菌水作为对照组; 以 0 h 为空白对照, 重复 3 次, 6、12、24 和 48 h 取样处理 (3 株), 于 $-80\text{ }^{\circ}\text{C}$ 冰箱储存, 采用 PacBio Seque III 进行上机测序, 获得远志全长转录组学文库^[17]。

2.2 远志 *bZIP* 家族基因鉴定及理化性质分析

基于远志转录组数据库, 筛选出注释结果为 *bZIP* 的序列, 将序列 gene id 对应的 fasta 结果输入 editseq 软件, 进一步获得具有完整开放阅读框 (open reading frame, ORF) 的基因, 通过 NCBI 中的 BlastX 进行比对与鉴定。ProtParam 分析目标蛋白的理化性质, ProtScale 预测不同氨基酸中的蛋白亲疏水性^[18]。

2.3 远志 *bZIP* 家族基因二级结构、信号肽、跨膜结构及亚细胞定位分析

用 ExPASy 分析基因编码蛋白质的结构域, CDD 验证; ProtParam 和 SOPMA 分析远志 *bZIP* 转录因子的二级结构; SignalP-5.0 和 TMHMM 预测信号肽和跨膜区域; WoLF PSORT 预测亚细胞定位^[19]。

2.4 远志 *bZIP* 家族基因进化树构建

从 Tair 网站下载拟南芥蛋白序列, 通过 MEGA 软件对远志、拟南芥 *bZIP* 氨基酸序列进行多序列比对, 利用 MEGA 的最大自然法构建系统发育树, 重复次数设置为 1 000 次^[20]。

2.5 远志 *bZIP* 家族基因密码子偏好性分析及蛋白互作预测分析

采用 CodonW、CUSP 和 Chips 分析密码子偏好性。蛋白互作预测分析利用 STRING 进行, 并以拟南芥筛选其同源基因后, 通过 Cytoscape 3.9.0 软件作图。

2.6 远志 *bZIP* 家族基因蛋白特征、保守基序分析及不同组织表达量热图

通过 chiplet 分析 *bZIP* 蛋白的结构域, MEME 获得 *bZIP* 蛋白的保守氨基酸基序, 并用 TBtools 进行可视化, Weblogo 分析蛋白序列位点。利用诺禾云平台将转录组数据库中 27 个 *PtbZIP* 基因在远志根、茎、叶 3 个部位的差异表达数据进行层级聚类

分析。

2.7 远志 *bZIP* 家族基因表达模式验证与分析

Trizol 法提取各样品总 RNA, 凝胶电泳检测后测定总 RNA 浓度。使用 Prime ScriptTM II 1st strand cDNA Synthesis Kit 合成 cDNA, 检测浓度后于 $-20\text{ }^{\circ}\text{C}$ 保存备用。设计荧光定量引物, 并送生工生物工程 (上海) 股份有限公司合成。以甘油醛-3-磷酸脱氢酶 (F: 5'-ACAGCAACGTGCTTCTCACC-3', R: 5'-CCCTTCATCCACCACCGACTA-3') 为内参基因, 验证 *PtbZIP26* (F: 5'-GCACTGATGGGAAGGCTGAA-3', R: 5'-GATTGCCCAACACTTGAGGG-3')、*PtbZIP27* (F: 5'-GTCGGATGGTAGTGAACGGG-3', R: 5'-CACCATTTCCTGAACCTGA-3') 在不同部位样本中的表达量。选择表达量较高的 *PtbZIP26* 进行不同激素、胁迫处理下的表达量分析。qRT-PCR 反应体系为 TB Green Premix Ex Taq II (2 \times) 5.0 μL ; 上下游引物各 0.4 μL ; 50 \times ROX Reference Dye 0.2 μL , cDNA 1.0 μL ; ddH₂O 3.0 μL 。PCR 反应程序参照 TB Green Premix Ex Taq II 试剂说明书进行, 每个反应重复 3 次。基因相对表达量采用 $2^{-\Delta\Delta C_t}$ 法计算, SPSS 27.0 统计分析。

3 结果与分析

3.1 远志 *bZIP* 基因家族成员的鉴定和蛋白理化性质分析

基于远志全长转录组数据库, 共筛选得到 63 个注释为 *bZIP* 基因的序列 ID, 进一步分析后获得 39 个包含完整 ORF 的序列。整理 ORF 差异位点并合并重复, 最终得到 27 个全长 *bZIP* 转录因子, 编号 *PtbZIP1*~*PtbZIP27* (表 1)。该转录因子的氨基酸个数 143~846, 相对分子质量介于 16 201.52~92 932.3, 等电点 4.59~9.69。除 *PtbZIP1* 和 *PtbZIP22* 的不稳定指数小于 40, 系稳定蛋白质外, 其余 *PtbZIP* 均为不稳定蛋白。*bZIP* 基因家族脂肪系数介于 48.31~92.66, 所有 *bZIP* 蛋白的平均亲水性数值是负值, 为亲水性蛋白。

3.2 远志 *bZIP* 基因家族成员的二级结构、信号肽、跨膜结构及亚细胞定位分析

二级结构分析结果 (表 2) 表明, 远志 *bZIP* 家族蛋白均具有 α 螺旋、延伸链、 β 转角和无规卷曲, 主要由 α 螺旋和无规卷曲构成, 延伸链和 β -折叠所占比例较小, 散布于整个蛋白中。SignalP-5.0 和 TMHMM 在线分析结果一致, 所有远志 *bZIP* 蛋白信号肽分值都低于 0.5, 说明其均无信号肽, 不属于分泌

表 1 远志 *bZIP* 基因家族基本信息

Table 1 Basic information of *P. tenuifolia bZIP* gene family

基因名	序列	氨基酸数量/aa	相对分子质量	等电点	不稳定指数	脂肪指数/%	亲水性平均值
<i>PtbZIP1</i>	transcript_HQ_PT_transcript7151/f3p0/2561	245	27 057.36	6.00	31.74	61.35	-0.688
<i>PtbZIP2</i>	transcript_HQ_PT_transcript5115/f2p0/2842	338	36 355.88	5.87	50.66	48.31	-0.920
<i>PtbZIP3</i>	transcript_HQ_PT_transcript12430/f3p0/1953	453	49 588.25	5.96	52.30	75.06	-0.531
<i>PtbZIP4</i>	transcript_HQ_PT_transcript18930/f4p0/1176	196	22 378.55	4.59	84.67	57.70	-1.071
<i>PtbZIP5</i>	transcript_HQ_PT_transcript18476/f2p0/1299	378	40 960.14	7.98	52.56	55.87	-0.945
<i>PtbZIP6</i>	transcript_HQ_PT_transcript4199/f2p0/3046	445	48 343.07	9.51	49.18	69.46	-0.685
<i>PtbZIP7</i>	transcript_HQ_PT_transcript16391/f2p0/1581	552	60 334.70	6.72	67.22	59.42	-0.896
<i>PtbZIP8</i>	transcript_HQ_PT_transcript12135/f4p0/1933	782	85 663.78	6.30	51.24	71.39	-0.574
<i>PtbZIP9</i>	transcript_HQ_PT_transcript4190/f3p0/3034	846	92 932.30	6.16	50.77	86.84	-0.097
<i>PtbZIP10</i>	transcript_HQ_PT_transcript13706/f2p0/1780	363	41 454.22	8.27	49.83	78.82	-0.518
<i>PtbZIP11</i>	transcript_HQ_PT_transcript4292/f2p0/3009	395	42 407.86	6.43	60.93	51.95	-0.913
<i>PtbZIP12</i>	transcript_HQ_PT_transcript5871/f3p0/2719	280	30 641.57	9.00	62.12	69.21	-0.629
<i>PtbZIP13</i>	transcript_HQ_PT_transcript10950/f2p0/2122	448	49 186.99	6.36	55.48	77.88	-0.507
<i>PtbZIP14</i>	transcript_HQ_PT_transcript13326/f2p0/1851	828	91 347.66	5.83	46.49	85.08	-0.136
<i>PtbZIP15</i>	transcript_HQ_PT_transcript13007/f2p0/1912	163	18 261.57	9.00	67.28	73.68	-0.514
<i>PtbZIP16</i>	transcript_HQ_PT_transcript3608/f2p0/3166	291	32 206.00	7.89	42.07	72.68	-0.675
<i>PtbZIP17</i>	transcript_HQ_PT_transcript7075/f6p0/2529	162	18 421.90	9.69	59.14	73.46	-0.628
<i>PtbZIP18</i>	transcript_HQ_PT_transcript14890/f2p0/1729	397	42 584.82	5.47	54.12	51.21	-0.899
<i>PtbZIP19</i>	transcript_HQ_PT_transcript8067/f2p0/2441	252	27 473.66	6.52	72.32	73.13	-0.726
<i>PtbZIP20</i>	transcript_HQ_PT_transcript17492/f2p0/1433	558	61 025.67	6.56	65.30	62.10	-0.820
<i>PtbZIP21</i>	transcript_HQ_PT_transcript13341/f2p0/1905	420	45 406.46	9.16	53.76	57.24	-0.839
<i>PtbZIP22</i>	transcript_HQ_PT_transcript13925/f4p0/1798	221	25 296.66	7.79	37.78	69.37	-0.645
<i>PtbZIP23</i>	transcript_HQ_PT_transcript12834/f3p0/1910	344	37 436.69	9.03	48.74	75.49	-0.640
<i>PtbZIP24</i>	transcript_HQ_PT_transcript13910/f3p0/1812	143	16 201.52	6.15	58.69	92.66	-0.550
<i>PtbZIP25</i>	transcript_HQ_PT_transcript15099/f2p0/1704	404	43 326.58	5.85	49.17	52.70	-0.899
<i>PtbZIP26</i>	transcript_HQ_PT_transcript17693/f12p0/1396	412	44 539.21	8.56	49.29	54.08	-0.902
<i>PtbZIP27</i>	transcript_HQ_PT_transcript19569/f2p0/1019	771	83 644.45	5.85	50.98	72.53	-0.498

表 2 远志 *bZIP* 基因家族二级结构、信号肽、跨膜结构及亚细胞定位分析

Table 2 Secondary structure, signal peptide, transmembrane structure and subcellular localization analysis of *P. tenuifolia*

bZIP gene family

基因名	α -螺旋	β -转角	无规则卷曲	信号肽/%	跨膜结构域	亚细胞定位
<i>PtbZIP1</i>	0.306 1	0.093 9	0.567 3	0.06	无	细胞核
<i>PtbZIP2</i>	0.263 3	0.047 3	0.665 7	0.09	无	细胞核
<i>PtbZIP3</i>	0.496 7	0.075 1	0.401 8	0.54	无	细胞核
<i>PtbZIP4</i>	0.382 7	0.112 2	0.454 1	0.20	无	细胞核、叶绿体
<i>PtbZIP5</i>	0.301 6	0.052 9	0.611 1	0.39	无	细胞核、叶绿体、质膜
<i>PtbZIP6</i>	0.303 4	0.092 1	0.577 5	0.79	无	细胞核、叶绿体
<i>PtbZIP7</i>	0.320 7	0.034 4	0.610 5	0.09	无	细胞核
<i>PtbZIP8</i>	0.180 3	0.131 7	0.656 0	0.26	无	细胞核
<i>PtbZIP9</i>	0.398 3	0.153 7	0.390 1	0.11	1	细胞核、质膜、内质网
<i>PtbZIP10</i>	0.617 1	0.060 6	0.297 5	0.13	无	细胞核
<i>PtbZIP11</i>	0.202 5	0.053 2	0.721 5	0.07	无	细胞核、细胞质、质膜
<i>PtbZIP12</i>	0.460 7	0.042 9	0.482 1	0.15	1	细胞核
<i>PtbZIP13</i>	0.508 9	0.073 7	0.395 1	0.08	无	细胞核
<i>PtbZIP14</i>	0.419 1	0.152 2	0.365 9	0.68	无	细胞核
<i>PtbZIP15</i>	0.601 2	0.049 1	0.300 6	0.49	无	细胞核
<i>PtbZIP16</i>	0.323 0	0.068 7	0.587 6	0.04	无	细胞核、叶绿体
<i>PtbZIP17</i>	0.561 7	0.049 4	0.333 3	0.13	无	细胞核、细胞骨架
<i>PtbZIP18</i>	0.201 5	0.055 4	0.730 5	0.03	无	细胞核、细胞质
<i>PtbZIP19</i>	0.440 5	0.027 8	0.500 0	0.10	无	细胞核
<i>PtbZIP20</i>	0.308 2	0.037 6	0.634 4	0.12	无	细胞核
<i>PtbZIP21</i>	0.283 3	0.050 0	0.635 7	0.17	无	细胞核
<i>PtbZIP22</i>	0.497 7	0.081 4	0.393 7	0.47	无	细胞核
<i>PtbZIP23</i>	0.369 2	0.034 9	0.572 7	0.22	无	细胞核
<i>PtbZIP24</i>	0.622 4	0.014 0	0.342 7	0.10	无	细胞核
<i>PtbZIP25</i>	0.190 6	0.064 4	0.730 2	0.22	无	细胞核
<i>PtbZIP26</i>	0.281 6	0.055 8	0.628 6	0.02	无	细胞核、细胞质
<i>PtbZIP27</i>	0.182 9	0.124 5	0.662 8	0.000 67	无	细胞核、细胞质、线粒体、细胞壁

蛋白。跨膜结构域分析则显示,仅 *PtbZIP9* 和 *PtbZIP13* 有跨膜结构域。亚细胞定位结果表明,远志 *bZIP* 家族成员主要定位在细胞核。

3.3 远志 *bZIP* 基因家族成员系统进化分析

利用 MEGA7.0 构建远志与拟南芥 *bZIP* 转录因子家族系统进化树。结果表明,27 个 *PtbZIP* 蛋白

分为 A、B、C、D、F、G、I、S 8 个组,没有 *bZIP* 蛋白分到 E 和 K 组中。其中 G 是最大的 1 个亚组,含有 *PtbZIP* 家族成员共 8 个,占总数的 29.63%; A、F、I 和 S 组均含 3 个 *PtbZIP* 家族成员, B 组含 2 个 *PtbZIP* 家族成员, C 组含 1 个 *PtbZIP* 家族成员, D 组含 4 个 *PtbZIP* 家族成员 (图 1)。

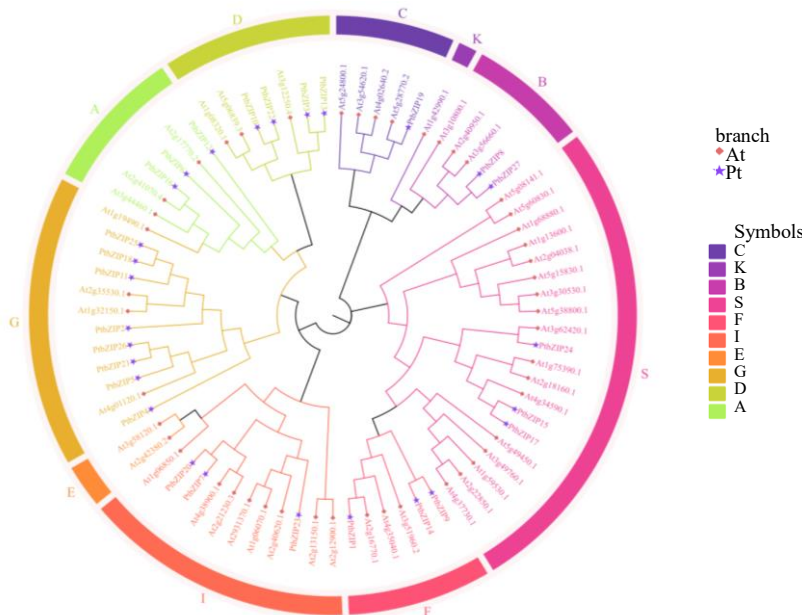


图 1 远志和拟南芥 *bZIP* 蛋白系统进化树

Fig. 1 Phylogenetic tree of *P. tenuifolia* and *A. thaliana* *bZIP* proteins

3.4 远志 *bZIP* 基因家族成员蛋白结构域分析

BRLZ、MFMR 和 DOG1 为 *bZIP* 蛋白中的常见结构域, BRLZ 参与调控果生炭疽菌的营养生长, MFMR 涉及蛋白与蛋白之间的相互作用, DOG1 则与种子休眠相关^[21-22]。远志 *bZIP* 的结构域分析结果表明: 10 个蛋白存在 BRLZ 结构域, 9 个蛋白存在 MFMR 结构, 6 个蛋白存在 DOG1 结构域 (图 2)。 *PtbZIP3* 和 *PtbZIP13* 含有大小相近的 CCDC 158 superfamily, *PtbZIP26*、*PtbZIP21* 和 *PtbZIP5* 则均含有 BRLZ、MFMR 及 homeobox 结构, 结合进化树结果可知 *PtbZIP3* 和 *PtbZIP13* 聚在一起, *PtbZIP26*、*PtbZIP21* 和 *PtbZIP5* 三者亲缘关系较近。

3.5 远志 *bZIP* 基因家族成员保守基序分析

利用 MEME 对远志 27 个 *bZIP* 蛋白序列进行保守基序分析的结果显示, 不同 *bZIP* 转录因子基因包含的保守元件数量及种类存在差异, 其中 *bZIP14* 基因包含的保守元件数量最少 (2 个), *bZIP18/25* 基因包含的保守元件数量最多 (11 个),

说明 *bZIP* 成员具有功能冗余现象, 也具有功能差异性 (图 3)。

bZIP 蛋白结合位点序列分析结果表明, *bZIP* 转录因子的每个重复结构域约为 65 aa, 均含有 1 个保守的 *bZIP* 结构域, 其中 N 端一般具有高度保守的 N-X7-R 蛋白基序和碱性亮氨酸区域 (图 4)。

3.6 远志 *bZIP* 基因家族成员密码子偏好性分析

密码子可用于推断基因组内部或基因组之间的进化关系, 而不同种类或同一种类的基因对密码子使用有不同的偏好模式^[23]。由 *bZIP* 基因家族中的 27 条核苷酸序列中密码子 GC 的总含量 (GC) 以及同义密码子第 1 位 (GC1s)、第 2 位 (GC2s)、第 3 位的 (GC3s) 的 GC 含量分析结果可知: 27 条 *PtbZIP* 基因序列的 GC1s、GC2s 和 GC3s 的均值分别为 52.24%、44.90% 和 40.93%, 不同位置的 GC 含量存在差异; 它们的 GC 平均值为 46.11%, 小于 50%, 表明其更偏向于 A 或 U 结尾的密码子^[24] (表 3)。

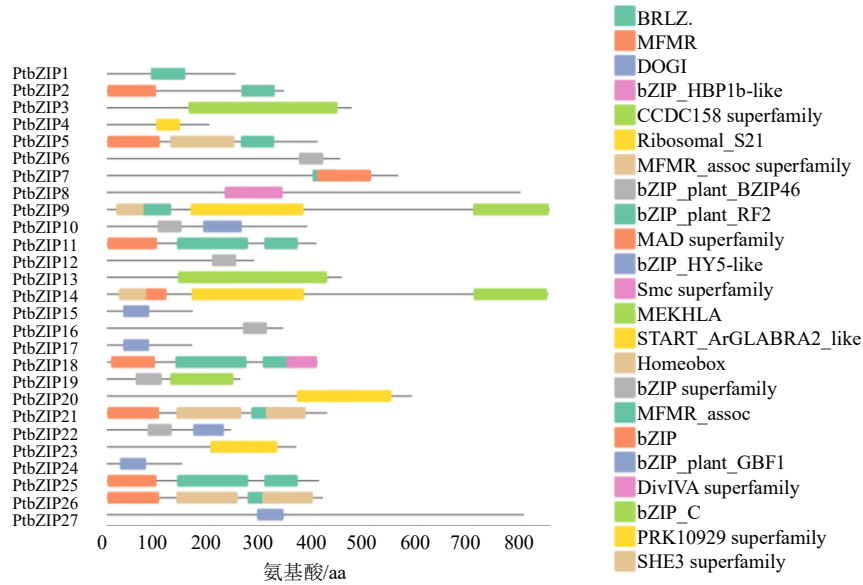


图2 远志 bZIP 家族蛋白结构域分析

Fig. 2 Structural domain analysis of *P. tenuifolia* bZIP family proteins

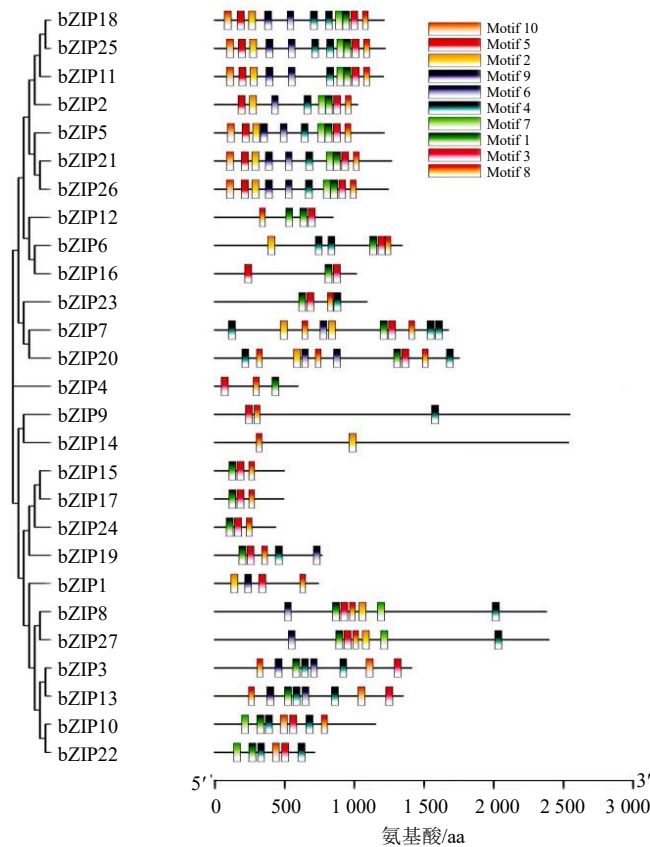


图3 bZIP 蛋白保守基序

Fig. 3 Conserved motifs of bZIP proteins

有效密码子 (effective number of codon, ENC) 反映了密码子偏离随机选择的结果, 它是对同义密码子非均衡使用偏好程度的一个重要指标^[25], ENC 数值一般在 20~61 范围内, 当 ENC>35 则表示密码子

偏好性较弱。密码子适应指数 (codon adaption index, CAI) 是指编码该蛋白的所有密码子相对于这条基因都使用最优密码子情况下的适应系数^[24]。由表 3 可知, 远志 bZIP 家族成员的 ENC 数值为 43.088~

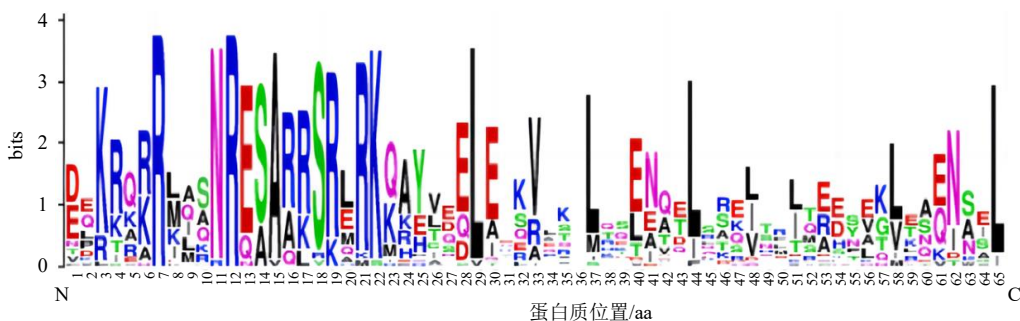


图 4 远志 bZIP 蛋白的 DNA 结合域

Fig. 4 DNA-binding domain of *P. tenuifolia* bZIP proteins

表 3 远志 bZIP 基因家族密码子偏好分析

Table 3 Codon preference analysis of *P. tenuifolia* bZIP gene family

基因名	ENC	CAI	CBI	GC/%	GC1/%	GC2/%	GC3/%
<i>PtbZIP1</i>	50.434	0.234	-0.124	44.04	52.44	42.68	36.99
<i>PtbZIP2</i>	49.859	0.209	-0.089	45.29	51.18	49.71	35.00
<i>PtbZIP3</i>	55.472	0.184	-0.054	46.06	55.44	44.14	38.59
<i>PtbZIP4</i>	56.275	0.225	0.040	50.08	55.33	42.13	52.79
<i>PtbZIP5</i>	45.940	0.190	-0.088	45.24	52.11	50.12	33.50
<i>PtbZIP6</i>	52.291	0.164	-0.052	46.86	51.35	43.27	45.96
<i>PtbZIP7</i>	52.715	0.205	0.003	45.30	51.71	47.40	36.80
<i>PtbZIP8</i>	48.556	0.206	-0.045	45.34	51.33	42.86	41.85
<i>PtbZIP9</i>	54.351	0.205	-0.060	46.36	54.66	45.22	39.20
<i>PtbZIP10</i>	52.463	0.185	-0.068	43.78	53.52	39.16	38.64
<i>PtbZIP11</i>	47.517	0.204	-0.035	44.72	44.72	51.12	31.17
<i>PtbZIP12</i>	47.366	0.214	0.038	50.53	55.16	45.91	50.53
<i>PtbZIP13</i>	55.914	0.183	-0.058	46.10	56.12	44.32	37.86
<i>PtbZIP14</i>	56.152	0.203	-0.026	47.95	53.55	45.50	44.79
<i>PtbZIP15</i>	57.195	0.228	0.103	46.14	45.12	40.24	53.05
<i>PtbZIP16</i>	54.542	0.185	-0.062	43.82	49.85	41.54	40.06
<i>PtbZIP17</i>	53.310	0.196	0.009	46.01	45.40	37.42	55.21
<i>PtbZIP18</i>	43.088	0.196	-0.082	46.48	53.10	49.88	36.48
<i>PtbZIP19</i>	47.308	0.185	-0.065	46.54	56.08	44.71	38.82
<i>PtbZIP20</i>	49.311	0.193	-0.017	45.28	50.77	47.68	37.39
<i>PtbZIP21</i>	48.975	0.180	-0.082	45.68	54.39	51.78	30.88
<i>PtbZIP22</i>	54.679	0.177	-0.089	42.05	50.21	39.24	36.71
<i>PtbZIP23</i>	51.577	0.194	0.067	50.37	53.04	46.96	51.10
<i>PtbZIP24</i>	51.951	0.184	0.032	50.00	55.56	36.81	57.64
<i>PtbZIP25</i>	44.107	0.195	-0.097	45.73	52.71	49.75	34.73
<i>PtbZIP26</i>	49.929	0.197	-0.080	44.71	54.48	50.61	29.06
<i>PtbZIP27</i>	49.131	0.211	-0.036	44.46	51.07	42.03	40.28

57.195 个，平均值为 51.13 个，密码子偏好性较弱。CBI 值较低说明其外源基因在目的宿主中表达较弱。CAI 值较低，则说明其适应性较弱。

3.7 远志 bZIP 基因家族成员蛋白互作网络分析

为深入了解远志 bZIP 蛋白的潜在功能和家族成员之间的相互作用，利用 STRING 软件，基于拟南芥数据库，对远志的 27 个 bZIPs 蛋白进行了互作网络分析。由图 5 可知，调控网络中共有 27 个

节点（代表 bZIPs 蛋白），104 条边（代表蛋白质之间的相互作用），表明远志的 bZIPs 蛋白存在多种互作现象，且 26 个 bZIPs 成员之间存在潜在的互作关系，为进一步验证远志 bZIP 的功能提供了重要依据。

3.8 远志 bZIP 基因家族成员不同组织表达量热图和验证

根据远志转录组数据，对 27 个 *PtbZIP* 基因在

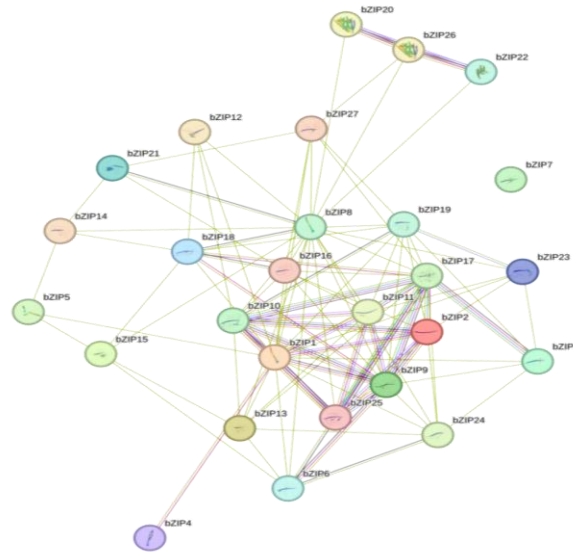


图5 远志 bZIP 蛋白互作预测分析

Fig. 5 Predictive analysis of *P. tenuifolia* bZIP protein interactions

远志根、茎、叶中的 FPKM 差异表达数据进行了双向聚类分析。通过表达量热图分析可知，绝大部分基因的表达不恒定，在不同组织具有相对较高的表达量，根、茎和叶中表达量较高的基因数分别为 23、2 和 2。PtbZIP4/15 在叶中的表达量最高，茎和根次之；PtbZIP8/24 在茎中的表达量最高，叶和根次之；剩下 23 个除 PtbZIP1/17 的表达量为根>叶>茎，其余表达模式为根>茎>叶（图 6-A）。基于 RT-qPCR 验证转录组数据结果显示，PtbZIP26、PtbZIP27 在根中的表达量最高，茎、叶次之，与转录组结果一致（图 6-B）。

3.9 PtbZIP26 不同处理下的表达模式

为了探究 bZIP 家族基因在远志不同处理条件下的表达模式，以 PtbZIP26 为代表，对其进行了激素和干旱、盐胁迫处理条件下的表达模式分析。结果发现，以 0 h 为空白对照（CK），PtbZIP26 的表达量在 ABA 处理 6 h 内迅速上升，在 24 h 达到峰值；CTS 处理分别持续上调至峰值为 CK 的 5.3 倍（24 h）后逐渐下调（图 7-A）。PEG 处理 6 h 迅速下降后又随着处理时间增加缓慢恢复上调，NaCl 处理 6 h 后上调明显（图 7-B）。

4 讨论

bZIP 基因家族在植物中广泛分布，参与植物的多个生长过程，如生长发育、应激反应以及次生代谢物的生物合成^[26]。现阶段，bZIP 基因家族已在多个物种有过相关的鉴定和研究，使得对 bZIP 的生物功能了解更透彻。本实验基于远志三代全长转录

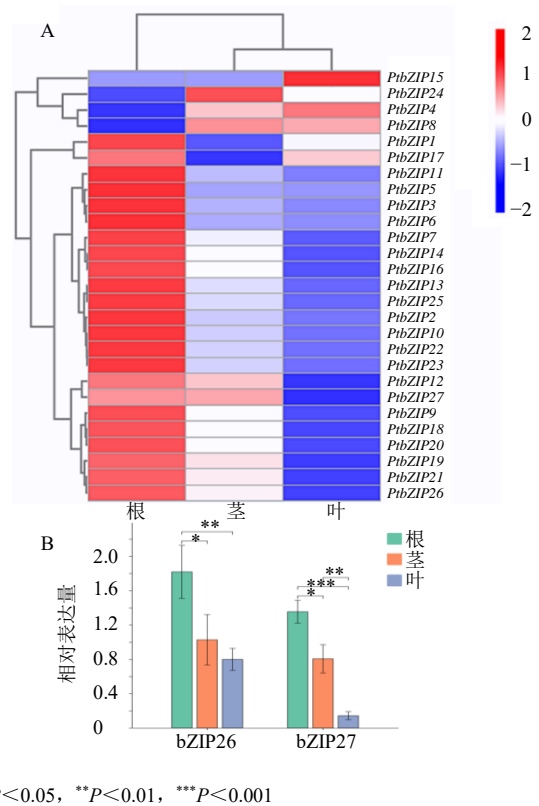
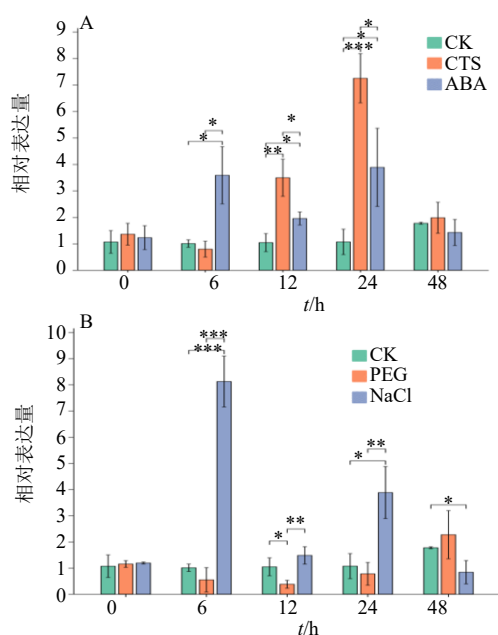


图6 PtbZIP 基因在远志不同组织中的表达模式分析 (A) 和 RT-qPCR 验证 PtbZIP26 和 PtbZIP27 在远志不同组织中的表达 (B)

Fig. 6 *PtbZIP* gene expression pattern analysis in different tissues of *P. tenuifolia* (A) and RT-qPCR validation of *PtbZIP26* and *PtbZIP27* expression in different tissues of *P. tenuifolia* (B)



* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.

图7 *PtbZIP26* 外源激素处理不同时间点的表达分析 (A) 和 *PtbZIP26* 非生物胁迫处理不同时间点的表达分析 (B)

Fig. 7 Expression analysis of *PtbZIP26* exogenous hormone treatments at different time points (A) and expression analysis of *PtbZIP26* abiotic stress treatments at different time points (B)

组数据库, 找到 39 个 *bZIP* isoforms, 通过完整开放阅读框与 BlastX 分析找出具有完整 ORF 的基因, 去除重复的 isoforms, 筛选并鉴定得到 27 个 *PtbZIP* 基因家族成员。理化性质分析显示, 27 个成员均为亲水性蛋白, 且除 *PtbZIP1* 和 *PtbZIP22* 外均为不稳定蛋白; 理论等电点小于 7 的蛋白有 16 个, 属酸性蛋白, 其余均为碱性蛋白。*PtbZIP* 蛋白信号肽分值都低于 0.5, 说明其均无信号肽, 信号肽是分泌蛋白的决定因子, 推测 *PtbZIP* 蛋白不属于分泌蛋白。亚细胞定位结果显示, 远志 *bZIP* 蛋白主要定位于细胞核, 这与转录因子主要在细胞核中发挥作用一致。*PtbZIP* 家族成员的蛋白二级结构也有明显的特点, 主要有 α -螺旋、无规卷曲。

系统进化分析显示, 27 个 *PtbZIP* 蛋白分为 A、B、C、D、F、G、I、S 8 个组, 其中含有 8 个 *PtbZIP* 家族成员的 G 亚组系最大亚组。*PtbZIP11/18/25* 与拟南芥 At1g32150.1、At2g35530.1 高度同源, 且包含的保守元件数量最多, 推测 *PtbZIP11/18/25* 可能在远志干旱应答的分子机制中起重要作用^[27]。研究表明, A 类别的大多数功能信息提示在 ABA 或应

激信号中的作用, *PtbZIP6/12/16* 被分在 A 组, 推测该基因可能参与到远志 ABA 信号转导途径^[28]。S 类别是拟南芥最大的 *bZIP* 类别之一, 在胁迫处理后也被转录激活或在花的特定部分特异表达。研究证实, 拟南芥 *bZIP* 家族中的 S 类别的基因在响应干旱有重要作用, 本研究中共有 3 个 *PtbZIP* 基因被分到 S 类别下, 其中 *PtbZIP15* 在叶中表达量高, *PtbZIP24* 在茎中表达量高, 可能参与调控远志对干旱的响应。同时, 27 个 *PtbZIP* 基因家族成员的蛋白二级结构预测结果十分相似, 但序列间同源性相对较低, 表明 *PtbZIP* 基因可能在远志生长发育方面发挥广泛的生物功能。表达模式分析发现, 大部分 *PtbZIP* 在根中表达最高, qPCR 结果验证与转录组数据一致, 推测它们主要在远志地下部分发挥作用。

植物中转录因子的表达与激素密切相关, 研究发现葡萄 *VvLysM8* 和 *VvLysM9* 在壳聚糖处理 12 h、脱落酸处理 3 h 时相对表达量最高^[14]。马铃薯 *StHXX* 家族基因在 ABA 诱导下表达均显著上调, 且在 10%PEG 胁迫处理下也呈不同程度的上调表达^[29]。陆地棉 *GhKIN* 基因家族的鉴定和分析发现, 干旱和盐胁迫处理后 *GhKIN14* 和 *GhKIN27* 表达出现下调, 而 *GhKIN18* 等在一定时间点表现为表达上调^[30]。本研究选择一个在根中高表达的 *PtbZIP26* 基因, 通过不同激素、胁迫处理探讨了其是否受到相关激素和胁迫调控, 结果表明激素处理 (ABA 和 CTS) 远志幼苗后, *PtbZIP26* 表达水平显著提高; 同时, 盐胁迫和干旱胁迫处理也可诱导 *PtbZIP26* 基因的表达发生改变且随胁迫时间的变化呈现出差异性, 说明 *PtbZIP26* 可能通过不同信号通路参与远志应对逆境胁迫的表达, 具体作用机制有待深入研究。

本实验基于远志三代转录组数据, 以远志 *bZIP* 基因家族为研究对象, 对其家族成员进行鉴定和生物信息学预测分析, 明确了相关结构特点与进化特征, 进一步通过 qPCR 分析其在不同组织、不同处理下的表达模式, 为探究 *PtbZIP* 参与生长发育、代谢过程及非生物胁迫的调控机制提供参考依据, 为后期的基因功能研究奠定了基础。

利益冲突 所有作者均声明不存在利益冲突

参考文献

- [1] 胡本祥, 彭亮, 杨冰月, 等. “秦药”的现代研究概况 [J]. 中草药, 2018, 49(21): 4949-4959.
- [2] 中国药典 [S]. 一部. 2020: 163-164.
- [3] 姬海月, 胡本祥, 杨冰月, 等. 芸苔素对铅胁迫下远志

- 种子萌发和幼苗生长生理特征的影响 [J]. 西北植物学报, 2022, 42(12): 2093-2102.
- [4] 姚辛敏, 周晓洁, 周妍妍, 等. 远志化学成分及药理作用研究进展 [J]. 中医药学报, 2022, 50(2): 103-107.
- [5] 王世晖, 苏琪辉, 邓凡莹, 等. 一测多评结合指纹图谱分析远志挥发油物质的变化规律 [J]. 药物分析杂志, 2023, 43(6): 1044-1052.
- [6] 姬海月, 罗瑶, 净易尧, 等. 不同土壤环境对远志标志性成分积累与体外抗氧化抗炎活性及乙酰胆碱酯酶抑制能力的影响 [J]. 时珍国医国药, 2023, 34(1): 166-169.
- [7] 罗瑶, 胡本祥, 张晗, 等. 卵叶远志叶绿体基因组序列特征与系统发育分析 [J]. 中草药, 2023, 54(18): 6065-6073.
- [8] Meng F B, Chu T Z, Feng P M, *et al.* Genome assembly of *Polygala tenuifolia* provides insights into its karyotype evolution and triterpenoid saponin biosynthesis [J]. *Hortic Res*, 2023, 10(9): 139.
- [9] Zhang B Y, Feng C, Chen L, *et al.* Identification and functional analysis of *bZIP* genes in cotton response to drought stress [J]. *Int J Mol Sci*, 2022, 23(23): 14894.
- [10] Zhao K, Chen S, Yao W J, *et al.* Genome-wide analysis and expression profile of the *bZIP* gene family in poplar [J]. *BMC Plant Biol*, 2021, 21(1): 122.
- [11] Rahman S, Rehman A, Waqas M, *et al.* Genome-wide exploration of *bZIP* transcription factors and their contribution to alkali stress response in *Helianthus annuus* [J]. *Plant Stress*, 2023, 10: 100204.
- [12] Fuhrmann-Aoyagi M B, de Fátima Ruas C, Barbosa E G G, *et al.* Constitutive expression of *Arabidopsis* *bZIP* transcription factor AREB1 activates cross-signaling responses in soybean under drought and flooding stresses [J]. *J Plant Physiol*, 2021, 257: 153338.
- [13] 怀浩, 宁康, 侯聪, 等. 火麻仁基原植物 *bZIP* 基因家族鉴定及其调控油脂代谢的功能初探 [J]. 药学学报, 2022, 57(8): 2528-2542.
- [14] 褚明宇, 李婉莹, 左存武, 等. 葡萄 LysM 类受体激酶基因家族的鉴定及表达分析 [J]. 农业生物技术学报, 2023, 30(2): 282-297.
- [15] 安沛沛. 糜子耐盐相关 *bZIP* 转录因子的筛选与功能研究 [D]. 杨凌: 西北农林科技大学, 2022.
- [16] 赵霞, 李元敏, 李依民, 等. 基于全长转录组测序的掌叶大黄 R1-MYB 基因家族鉴定分析 [J]. 药学学报, 2023, 58(5): 1354-1363.
- [17] 陈泓宇, 董树廷, 郭妙弦, 等. 茯苓基因组中 MYB 转录因子基因家族鉴定及表达分析 [J]. 中草药, 2023, 54(1): 245-253.
- [18] 王思嘉, 孙嘉莹, 刘美琦, 等. 人参 *bZIP* 基因家族生物信息学分析 [J]. 中草药, 2022, 53(9): 2786-2794.
- [19] 蔡伟, 肖敬忠, 余小丽, 等. 鱼腥草 *bZIP* 基因的鉴定与分析 [J]. 中草药, 2023, 54(13): 4295-4305.
- [20] 李司政, 姚权, 李河. 果生炭疽菌转录因子 CfHac1 的 BRLZ 结构域生物学功能研究 [J]. 北京林业大学学报, 2021, 43(9): 70-76.
- [21] 朱芸晔, 薛冰, 王安全, 等. 番茄 *bZIP* 转录因子家族的生物信息学分析 [J]. 应用与环境生物学报, 2014, 20(5): 767-774.
- [22] 傅武祥, 郭雨桐, 张雪梅, 等. 九翅豆蔻叶绿体基因组特征及其密码子偏好性分析 [J]. 基因组学与应用生物学, 2024, 43(1): 17-30.
- [23] 黄长兵, 沈萍, 侯哲. 15 种蔷薇属 *matK* 基因的密码子偏好性分析 [J]. 福建农林大学学报: 自然科学版, 2023, 52(4): 466-472.
- [24] 郭佳星, 黄祥, 杨梅花, 等. 桦木科叶绿体基因组密码子偏好性及系统发育分析 [J]. 中国农业科技导报, 2023, 25(10): 74-83.
- [25] 甘颖, 袁滨, 韩星, 等. 三种侧耳属真菌基因组密码子偏好性比较分析 [J]. 菌物学报, 2023, 42(6): 1285-1297.
- [26] 蔡伟, 肖敬忠, 余小丽, 等. 鱼腥草 *bZIP* 基因的鉴定与分析 [J]. 中草药, 2023, 54(13): 4295-4305.
- [27] 马斯琦. 水稻转录因子 OsbZIP46 调控干旱应答的分子机制 [D]. 武汉: 华中农业大学, 2018.
- [28] 邓璇, 陈春兵, 刘练练, 等. 白桑 *bZIP* 基因家族的全基因组鉴定及表达谱分析 [J]. 蚕业科学, 2022, 48(6): 477-488.
- [29] 王言, 何雷, 刘涛涛, 等. 马铃薯 *HXX* 基因家族鉴定及表达特征分析 [J]. 西北植物学报, 2023, 43(12): 2018-2029.
- [30] Guo L J, Wu F, Ci P Y, *et al.* Genome-wide identification and characterization of the kinesin gene family in upland cotton [J]. *Indust Crops Prod*, 2023, 204: 23.

[责任编辑 时圣明]