

DOI:10.13275/j.cnki.lykxyj.2022.02.011

曼地亚红豆杉 NAC 基因家族 鉴定及表达分析

张恺恺¹, 杨立莹³, 丰美静^{1,2}, 张林凤¹, 陈段芬², 邱德有¹, 杨艳芳^{1*}

(1. 中国林业科学研究院林业研究所, 林木遗传育种国家重点实验室, 国家林业和草原局林木培育重点实验室, 北京 100091;

2. 河北农业大学园艺学院, 河北保定 071001; 3. 中国花卉协会, 北京 100020)

摘要: [目的] 筛选红豆杉 NAC 转录因子家族中影响根系形成的关键基因, 探索红豆杉根系生长发育分子机理。[方法] 利用曼地亚红豆杉全长转录组数据, 通过生物信息学方法鉴定 NAC 转录因子, 并对筛选出的 NAC 转录因子进行蛋白结构及基因组表达谱等分析。[结果] 共鉴定出 44 个 NAC 转录因子, 其在 N 端均具有典型的保守 NAC 结构域, 分别聚类到拟南芥的 7 个亚家族中, 蛋白三级结构较为相似, 且成员大多含有 5 个保守亚结构域。组织表达谱显示 *TmNAC15*、*16*、*18*、*21*、*22*、*29*、*39*、*40*、*41* 和 *44* 在根中表达水平高于茎和叶。[结论] 从曼地亚红豆杉中共鉴定出 44 个 NAC 转录因子, 其结构较为保守, 且聚类为 7 个亚家族, 其中成员 *TmNAC21*、*22*、*39*、*40* 和 *44* 极有可能参与红豆杉根系生长发育。

关键词: 红豆杉; NAC 基因; 基因鉴定; 表达分析

中图分类号: S718.46

文献标志码: A

文章编号: 1001-1498(2022)02-0097-07

红豆杉 (*Taxus* spp.) 是举世公认的珍贵林木, 在全世界一共有 11 种, 我国存在 4 种和 1 变种, 即东北红豆杉 (*T. cuspidata* Sieb. et Zucc.)、西藏红豆杉 (*T. wallichiana* Zucc.)、云南红豆杉 (*T. yunnanensis* Cheng et L. K. Fu)、中国红豆杉 (*T. chinensis* (Pilger) Rehd.) 以及南方红豆杉 (*T. chinensis* var. *mairei* (Lemee et Levl) Cheng et L. K. Fu) [1]。曼地亚红豆杉 (*Taxus × media* Rehder) 为一种天然杂交种, 于上世纪 90 年代由美国或加拿大引种于我国。

NAC (NAM/ATAF/CUC) 是植物中最常见的转录因子家族之一, 在植物生长发育过程中起到重要的调控作用, 如响应非生物胁迫、调控器官生长发育以及参与激素信号转导等 [2-4]。研究发现, 水稻 OsNAC2 蛋白能够通过影响植物生长素和细胞分

裂素 (cytokinin) 反应基因来调控根的发育 [5], 而拟南芥 ANAC092 可以与生长素应答因子 ARF8 和 PIN4 的启动子相结合, 通过控制生长素信号途径来影响拟南芥根系的发育 [6-8]。此外, 在拟南芥中异源表达 *BnNAC14*、*GmNAC109*、*CiNAC3* 和 *CiNAC4* 等其他物种的 NAC 基因, 也可以对侧根的形成起到促进作用 [8-11]。这些研究结果表明, NAC 转录因子在植物根系的形成过程中扮演重要角色。

近年来, 人们为了缓解红豆杉开发与利用的矛盾, 在红豆杉人工种植研究上取得了巨大进展, 但仍存在红豆杉扦插繁殖生根较慢 (大约需要 45~60 d) 的技术瓶颈。并且, 目前关于红豆杉生根的分子机制研究也较为少见。本研究从曼地亚红豆杉三代全长转录组数据中鉴定出了 44 个 NAC

收稿日期: 2021-04-06 修回日期: 2021-06-25

基金项目: “十三五”国家重点研发计划子课题“红豆杉高效培育技术研究”(2017YFD060070605); 国家林业和草原局科技发展中心项目 (KJZXS202037); 国家自然科学基金项目 (31570675) 项目资助

* 通讯作者: 杨艳芳, 博士, 副研究员. 主要研究方向: 林木次生代谢与调控. 电话: 010-62889638. E-mail: echoyyf@caf.ac.cn

转录因子基因, 并对该基因家族进行了系统生物信息学及组织表达分析, 为探索 NAC 转录因子调控红豆杉根系的形成奠定理论基础, 有助于今后利用分子手段调控其根系生长、缩短扦插育苗时间。

1 材料与方法

1.1 红豆杉 NAC 转录因子家族的鉴定

通过隐马可夫模型 HMM(PF02365) (TBtools v1.064^[12]) 从课题组曼地亚红豆杉三代全长转录组数据(未发表)中获取 NAC 蛋白序列。利用在线工具 CD-HIT Suite 进行去冗余^[13], 利用 NCBI CD Search 工具进行结构域预测, 并筛选结构域完整的序列进行后续分析。

1.2 曼地亚红豆杉 NAC 家族理化性质及三级结构分析

利用 ExPASy 在线工具 ProtParam (<https://web.expasy.org/protparam/>) 分析 TmNAC 蛋白理化特性。根据进化树结果, 每个亚家族选取一个 TmNAC 蛋白作为代表, 利用 phyre2 在线网站 (<http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index>) 进行三级结构建模。

1.3 多序列比对, 系统进化树构建及保守基序分析

拟南芥 NAC 蛋白序列下载于 Plant TFDB (<http://planttfdb.cbi.pku.edu.cn/>) 数据库。从 138 个拟南芥 NAC 蛋白质序列中随机选取 6 个序列, 用 ClustalX2 软件将其与 44 个 TmNAC 序列进行多序列比对, 比对结果利用 EsPript 3.0 (<http://esprpt.ibcp.fr/ESPript/cgi-bin/ESPript.cgi>) 进行美化。利用 MEGA 7 软件中的 NJ (Neighbor-Joining) 法构建进化树, 重复 1 000 次, 其他参数为默认值。iTOL (<https://itol.embl.de/>) 在线网站美化进化树。

利用 MEME 在线网站对 TmNAC 蛋白进行保守基序 (Motif) 分析^[14], 最大 motif 检索数为 10, 其余为默认参数, 将结果通过 TBtools v1.064 软件进行可视化处理^[12]。

1.4 曼地亚红豆杉 NAC 基因在不同组织的表达模式分析

利用云南红豆杉植株根、茎和叶的 RNA-seq 数据 (SRP127697) 和 TBtools v1.064 软件绘制 TmNACs 基因在根、茎和叶组织中的表达热图, 对 NAC 成员基因表达量 FPKM 值使用 $\log_2(\text{FPKM} + 1)$ 对数转化处理^[12]。

2 结果与分析

2.1 TmNAC 转录因子家族成员理化特性分析

本研究共获得 44 条 NAC 蛋白序列, 命名为 TmNAC1~TmNAC44。对获得的 44 个 TmNAC 蛋白序列进行理化性质分析, 发现 TmNAC38 序列最长, 含有 1010 个氨基酸残基, 而 TmNAC9 序列最短, 含有 154 个氨基酸残基; 这些 TmNAC 蛋白等电点为 4.70~9.60, 相对分子质量为 18.13~113.57 kDa, GRAVY 值均为负值, 表明 TmNAC 蛋白均为亲水蛋白。

2.2 系统进化树分析

系统进化树结果显示, 44 个 TmNAC 蛋白被分为 NAM 等 7 个亚家族, 其他 NAC1、ONAC22 和 SEUN5 等拟南芥 9 个 NAC 亚家族中无 TmNACs 蛋白分布 (图 1)。

2.3 TmNAC 蛋白三级结构分析

从图 1 中可以看出, 各 TmNAC 成员的三级结构十分相似, 但也存在些许差异, 如 TmNAC1 蛋白比其他蛋白多了一个 β -折叠, TmNAC16 和 TmNAC21 蛋白比其他蛋白多一个 α 螺旋 (图 2)。

2.4 TmNAC 蛋白多序列比对

多序列比对结果发现, 所有的 NAC 蛋白在 N 端均具有保守的 NAC 结构域 (图 3), 其结果与拟南芥和水稻 NAC 转录因子家族分析结果相似^[15]。此外, TmNAC7 和 TmNAC42 蛋白具有不完整的亚结构域 D 与 E (图 3), 其余蛋白均具有 5 个保守的亚结构域 (A~E)。

2.5 TmNAC 蛋白的保守基序分析

基序分析结果发现 motif 2、4、5 几乎存在于所有的 NAC 蛋白成员中。另外, 有的亚家族具有或缺失某些蛋白基序, 如 motif 8、10 仅存在于 ANAC011 亚家族中, motif 7 仅存在于 NAM 与 ATAF 亚家族中; 而 ANAC011 亚家族中的 TmNAC36 和 TmNAC15 则分别缺少 motif 2 和 motif 5 (图 4)。

2.6 红豆杉 TmNAC 基因组织表达分析

组织表达分析结果显示, TmNAC15、16、18、21、22、29、39、40、41 和 44 在根中均有较高的表达量, 而在叶和茎中的表达量都相对较低。另外, TmNAC2、3、8、17、25 和 42 在根与叶中表达量相近, 而 TmNAC27 和 TmNAC43 在茎中表达量最高, 在叶和根中的表达量均较低 (图 5)。

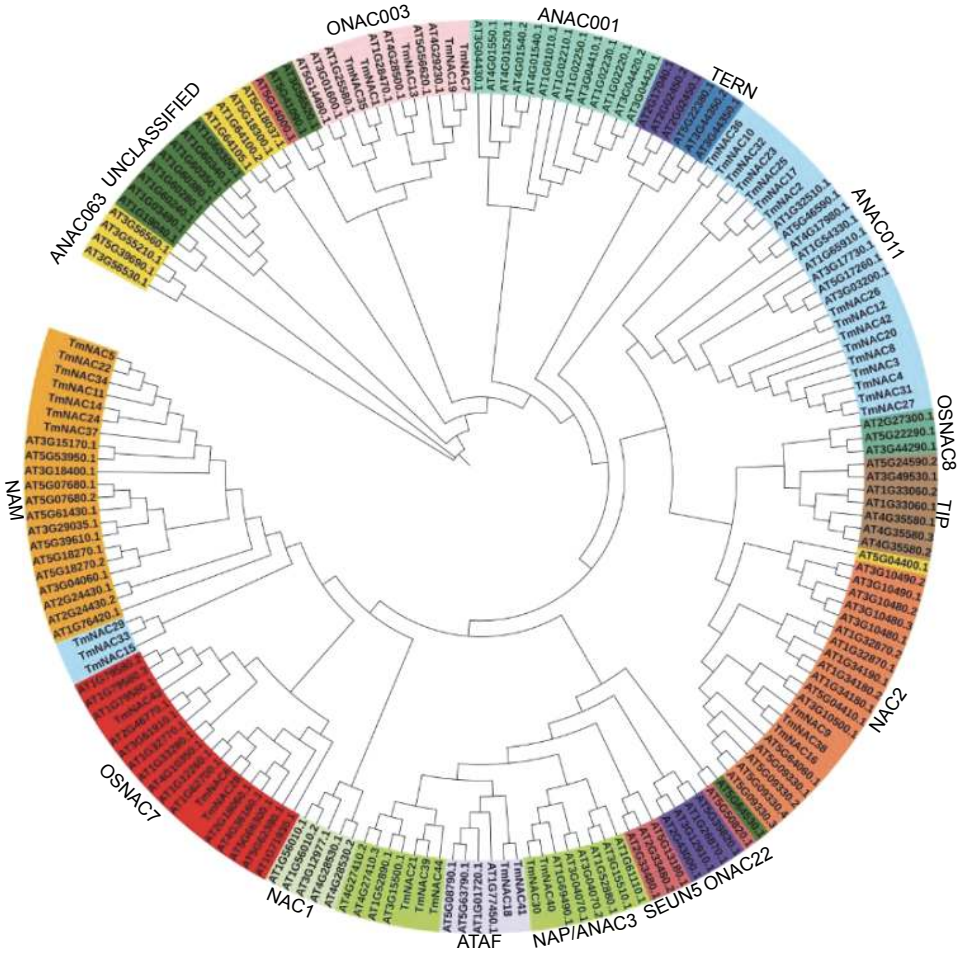
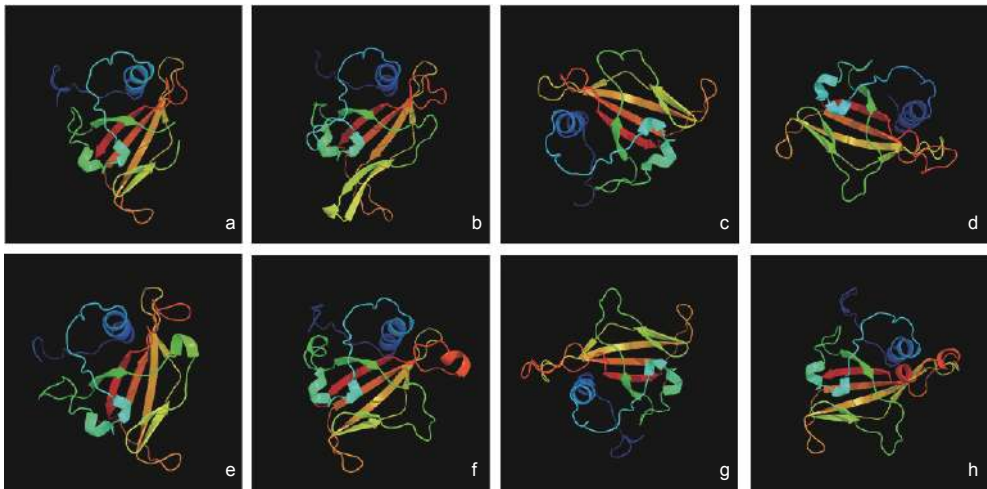


图 1 曼地亚红豆杉与拟南芥 NAC 家族进化树分析

Fig. 1 Phylogenetic relationship among NAC superfamily genes of of *Taxus x media* and *Arabidopsis thaliana*

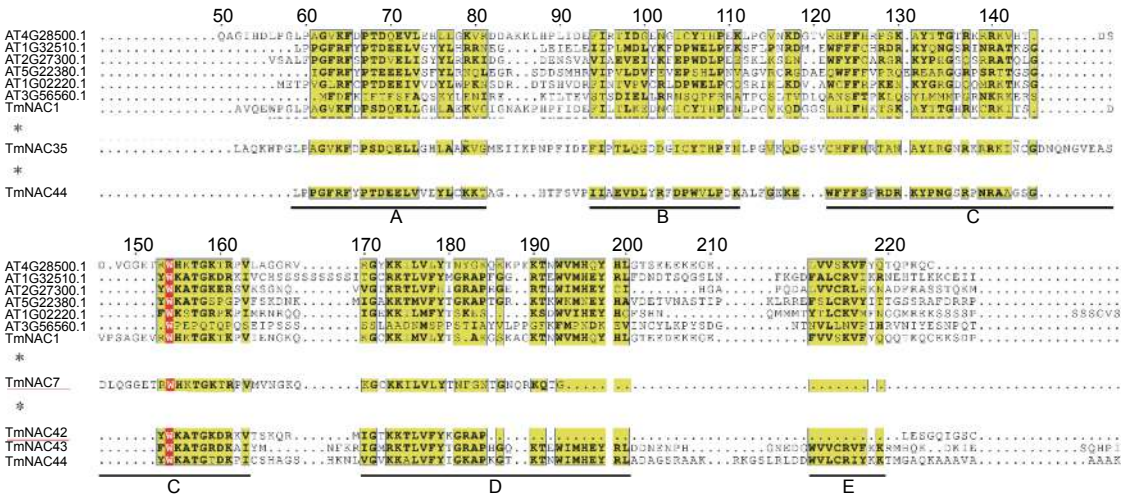


a: c3ulxA (模型); b: TmNAC1; c: TmNAC2; d: TmNAC5; e: TmNAC6; f: TmNAC16; g: TmNAC18; h: TmNAC21.

a: c3ulxA (model); b: TmNAC1; c: TmNAC2; d: TmNAC5; e: TmNAC6; f: TmNAC16; g: TmNAC18; h: TmNAC21.

图 2 TmNAC 蛋白的三级结构分析

Fig. 2 3D structures of TmNAC proteins



A~E 代表保守亚结构域; *代表序列较为相似, 进行省略

A~E represent conservative subdomains; * means the other TmNAC protein sequences are omitted as they are very similar

图 3 曼地亚红豆杉和拟南芥 NAC 蛋白保守结构域分析

Fig. 3 Analysis of conserved domains of the NAC proteins of *Taxus x media* and *Arabidopsis thaliana*

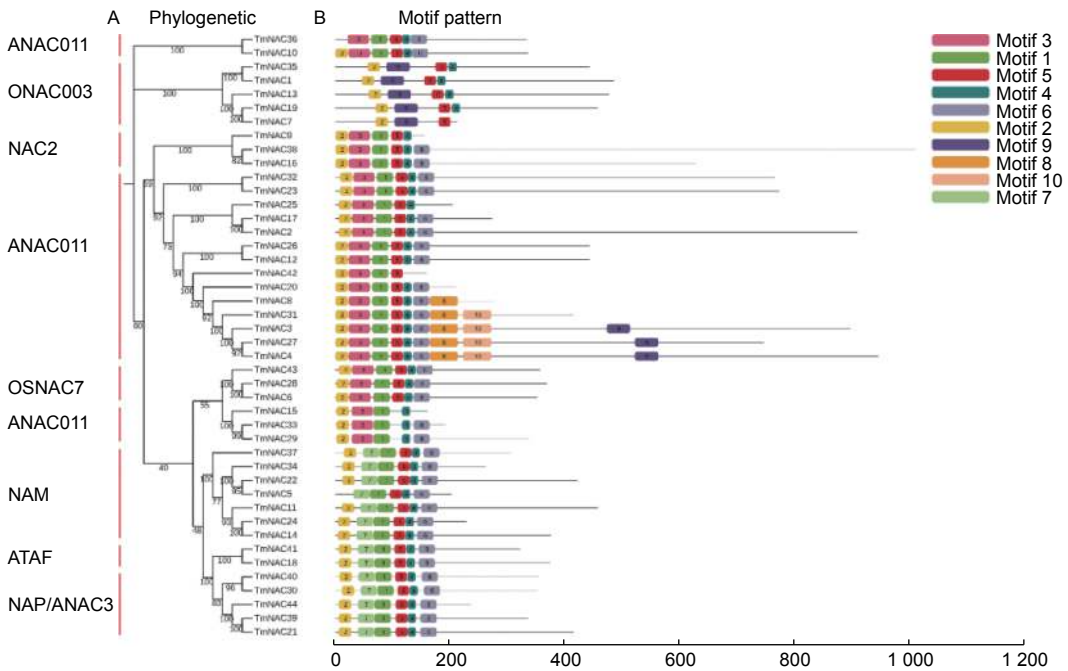


图 4 曼地亚红豆杉 TmNAC 蛋白 motif 分析

Fig. 4 Motif analysis of TmNAC proteins of *Taxus x media*

3 讨论

NAC 基因家族具有保守的结构域和功能域, 对植物的生长发育具有重要的作用^[15-17]。本研究从曼地亚红豆杉转录组数据中鉴定出 44 个 NAC 转录因子, 数目明显少于拟南芥和水稻基因组中的 138 个和 140 个^[18-19]。究其原因, 很有可能是由于本研究的 NAC 来源于三代全长转录组测序数据,

也有可能是因为物种不同造成。目前, 已经有西藏红豆杉和南方红豆杉基因组发表^[20-21], 未来可以利用这些基因组信息对 NAC 基因家族进一步进行分析, 红豆杉 NAC 家族的完整性可以得到进一步完善和确定。

研究表明 NAC 蛋白具有典型结构特征, 均含有 NAC 功能域, 且 NAC 区域包含 A~E 共 5 个 NAC 亚结构域。本研究蛋白多序列比对结果显

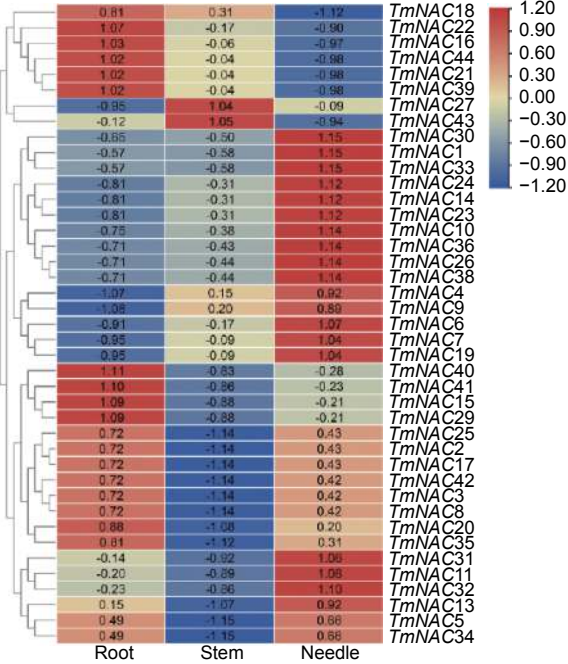


图5 TmNAC 基因在根、茎和叶中表达模式

Fig. 5 The expression patterns of *TmNAC* genes in roots, leaves and stems

示, 44 个 *TmNAC* 蛋白在 N 端均含有 NAC 结构域, 但 *TmNAC7* 和 *TmNAC42* 蛋白具有不完整的 D 和 E 亚结构域。这 5 个亚结构域的保守性顺序为 A>C>D>B>E, 亚结构域 A 与 C 对蛋白质的稳定发挥着重要作用。此外, 亚结构域 E 是 NAC 蛋白中一个重要的 DNA 结合区域, 可能参与了发育时期调控, 并且前人也推测存在由亚结构域 D 和 E 组成的 DNA 结合区^[22]。蛋白的功能与其结构密切相关, 由此可以推测, *TmNAC* 家族有很大可能与其他植物的 NAC 蛋白功能相似, 而 *TmNAC7* 和 *TmNAC42* 蛋白功能则极其可能与其他 *TmNAC* 蛋白功能有所不同。

进化树分析显示红豆杉 *TmNAC* 成员与大多数拟南芥 NAC 家族各亚家族成员具有较高同源性, 表明 NAC 家族在物种间进化保守。因此, 可以通过其他物种 NAC 蛋白功能来预测红豆杉 NAC 同源蛋白的功能。此外, 拟南芥中 NAC1、ONAC22 和 SEUN5 等 9 个亚家族并不包含红豆杉 NAC 同源蛋白, 这些情况可能是由于曼地亚红豆杉的 NAC 家族序列是从转录组数据中获得, 并未全部获取到基因组中其他亚家族 NAC 蛋白序列所造成。

多序列比对、motif 及蛋白质三级结构分析均发现, 44 个曼地亚红豆杉 NAC 蛋白具有较高的相

似性和保守性, 该结果与桃树、蓖麻、茶树、番茄中的报道相类似, 说明其在进化上非常保守^[23-26]。值得注意的是, *TmNAC1*、*TmNAC16* 和 *TmNAC21* 三者所在的亚家族与其他亚家族成员结构上存在显著差异, 且不同亚家族或同一亚家族不同成员之间 motif 类型和数目也存在差异, 这些结构差异是否导致它们功能发生变异尚不清楚, 值得研究人员进一步深入研究。

研究表明, 不同的物种间同源基因在进化过程中会保留相同或相似的功能, 而基因的组织表达模式往往与其功能密切相关^[27]。拟南芥和水稻的一些 NAC 转录因子的功能已经得到验证, 这为预测红豆杉 *TmNAC* 转录因子的功能提供了线索。*TmNAC* 家族 *TmNAC15*、*16*、*18*、*21*、*22*、*29*、*39*、*40*、*41* 和 *44* 基因在红豆杉根中表达量较高, 同时 *TmNAC21*、*30*、*39*、*40* 和 *44* 蛋白与拟南芥 *AtNAC2/AT3G15510* 蛋白在系统发育树中聚为一支, 而 *AtNAC2* 恰恰参与侧根的形成与发育^[28]; *TmNAC5*、*11*、*14*、*22*、*24*、*34* 和 *37* 蛋白与拟南芥 *AtNAC1/AT3G15170* 蛋白在系统发育树中聚为一支, *AtNAC1* 介导生长素信号以促进侧根的形成^[29]。综合上述, 推测红豆杉 *TmNAC21*、*39*、*40*、*44* (*NAP/ANAC3* 亚家族) 和 *TmNAC22* (*NAM* 亚家族) 可能在红豆杉根系生长发育中起到重要调控作用。

4 结论

本研究从红豆杉转录组中鉴定出 44 个 NAC 成员, 发现红豆杉 NAC 基因编码的 44 个蛋白在 N 端均具有保守的 NAC 结构域, *TmNAC* 家族有很大可能与其他植物的 NAC 蛋白功能相似。蛋白功能结构域分析结果预示着红豆杉 NAC 家族成员可能在亚家族间以及亚家族内部具有功能的特异性。表达量比较分析显示, *TmNAC* 基因在红豆杉不同组织中的表达量各不相同, *TmNAC21*、*22*、*39*、*40* 和 *44* 可能与红豆杉根的生长发育有关, 为进一步探索红豆杉 NAC 转录因子调控红豆杉根系生长提供理论基础。

参考文献:

- [1] 王亚飞, 王 强, 阮 晓, 等. 红豆杉属植物资源的研究现状与开发利用对策[J]. 林业科学, 2012, 48 (5): 116-125.
- [2] 赵 灿, 刘任重, 柳展基, 等. 亚洲棉全基因组中NAC类转录因子

- 基因的鉴定与分析[J]. 石河子大学学报:自科版, 2015, 33(4): 397-404.
- [3] Yang Y F, Zhu K, Wu J, *et al.* Identification and characterization of a novel NAC-like gene in chrysanthemum (*Dendranthema lavandulifolium*)[J]. Plant Cell Reports, 2016, 35(8): 1783-98.
- [4] 王立国, 傅明川, 李 浩, 等. 陆地棉 NAC 转录因子基因 *GhSNAC1* 的克隆及其抗旱耐盐分析[J]. 农业生物技术学报, 2019, 27(4): 571-580.
- [5] Mao C, He J, Liu L, *et al.* OsNAC2 integrates auxin and cytokinin pathways to modulate rice root development[J]. Plant Biotechnology Journal, 2020, 18(2): 429-442.
- [6] Xi D, Chen X, Wang Y, *et al.* Arabidopsis ANAC092 regulates auxin-mediated root development by binding to the *ARF8* and *PIN4* promoters[J]. Journal of Integrative Plant Biology, 2019, 61(9): 1015-1031.
- [7] Chen D, Chai S, McIntyre C L, *et al.* Overexpression of a predominantly root-expressed NAC transcription factor in wheat roots enhances root length, biomass and drought tolerance[J]. Plant Cell Reports, 2018, 37(2): 225-237.
- [8] Hao Y J, Wei W, Song Q X, *et al.* Soybean NAC transcription factors promote abiotic stress tolerance and lateral root formation in transgenic plants[J]. Plant Journal, 2011, 68(2): 302-13.
- [9] Han X, Feng Z, Xing D, *et al.* Two NAC transcription factors from *Caragana intermedia* altered salt tolerance of the transgenic Arabidopsis[J]. BMC Plant Biology, 2015, 15: 208.
- [10] Yang X, Kim M Y, Ha J, *et al.* Overexpression of the soybean NAC gene *GmNAC109* increases lateral root formation and abiotic stress tolerance in transgenic Arabidopsis plants[J]. Frontiers in Plant Science, 2019, 10: 1036.
- [11] Hegedus D, Yu M, Baldwin D, *et al.* Molecular characterization of *Brassica napus* NAC domain transcriptional activators induced in response to biotic and abiotic stress[J]. Plant Molecular Biology, 2003, 53(3): 383-97.
- [12] Chen C, Chen H, Zhang Y, *et al.* TBtools: An integrative toolkit developed for interactive analyses of big biological data[J]. Molecular Plant, 2020, 13(8): 1194-1202.
- [13] Huang Y, Niu B, Gao Y, *et al.* CD-HIT Suite: a web server for clustering and comparing biological sequences[J]. Bioinformatics, 2010, 26(5): 680-2.
- [14] Bailey T L, Boden M, Buske F A, *et al.* MEME SUITE: tools for motif discovery and searching[J]. Nucleic Acids Research, 2009, 37 (Web Server issue): W202-8.
- [15] Sun H, Hu M, Li J, *et al.* Comprehensive analysis of NAC transcription factors uncovers their roles during fiber development and stress response in cotton[J]. BMC Plant Biology, 2018, 18(1): 150.
- [16] Zhuo X, Zheng T, Zhang Z, *et al.* Genome-wide analysis of the NAC transcription factor gene family reveals differential expression patterns and cold-stress responses in the woody plant *Prunus mume*[J]. Genes (Basel), 2018, 9(10): 494.
- [17] Gong X, Zhao L, Song X, *et al.* Genome-wide analyses and expression patterns under abiotic stress of NAC transcription factors in white pear (*Pyrus bretschneideri*)[J]. BMC Plant Biology, 2019, 19(1): 161.
- [18] Ooka H, Satoh K, Doi K, *et al.* Comprehensive analysis of NAC family genes in *Oryza sativa* and *Arabidopsis thaliana*[J]. DNA Research, 2003, 10(6): 239-47.
- [19] Fang Y, You J, Xie K, *et al.* Systematic sequence analysis and identification of tissue-specific or stress-responsive genes of NAC transcription factor family in rice[J]. Molecular Genetics and Genomics, 2008, 280(6): 547-63.
- [20] Cheng J, Wang X, Liu X, *et al.* Chromosome-level genome of Himalayan yew provides insights into the origin and evolution of the paclitaxel biosynthetic pathway[J]. Molecular Plant, 2021, S1674-2052(21): 00160-X.
- [21] Xiong X, Gou J, Liao Q, *et al.* The Taxus genome provides insights into paclitaxel biosynthesis[J]. BioRxiv, 2021: 2021.04.29.441981.
- [22] Ernst H A, Olsen A N, Larsen S, *et al.* Structure of the conserved domain of ANAC, a member of the NAC family of transcription factors[J]. EMBO Reports, 2004, 5(3): 297-303.
- [23] 张春华, 上官凌飞, 俞明亮, 等. 桃 NAC 基因家族生物信息学分析[J]. 江苏农业学报, 2012, 28(2): 406-414.
- [24] 代梦媛, 高 梅, 李文昌. 蓖麻 NAC 转录因子家族的鉴定及生物信息学分析[J]. 分子植物育种, 2020, 18(6): 1808-1817.
- [25] 韩雅彭, 程 琳, 杨凌霄. 茶树 NAC 转录因子家族的鉴定及生物信息学分析[J]. 河南大学学报:自然版, 2017(47): 301-309.
- [26] 姜秀明, 牛义岭, 许向阳. 番茄 NAC 基因家族的系统进化及表达分析[J]. 分子植物育种, 2016, 14(8): 1948-1964.
- [27] 王 震, 米要磊, 孟祥霄, 等. 中药火麻仁基原植物大麻 *LBD* 基因家族成员的鉴定与表达分析[J]. 中国中药杂志, 2020, 45(22): 175-184.
- [28] He X J, Mu R L, Cao W H, *et al.* AtNAC2, a transcription factor downstream of ethylene and auxin signaling pathways, is involved in salt stress response and lateral root development[J]. Plant Journal, 2005, 44(6): 903-16.
- [29] Xie Q, Frugis G, Colgan D, *et al.* Arabidopsis NAC1 transduces auxin signal downstream of TIR1 to promote lateral root development[J]. Genes & Development, 2000, 14(23): 3024-3036.

Identification and Expression Analysis of NAC Gene Family in *Taxus × media* Rehder

ZHANG Kai-kai¹, YANG Li-ying³, FENG Mei-jing^{1,2}, ZHANG Lin-feng¹,
CHEN Duan-fer², QIU De-you¹, YANG Yan-fang¹

(1. Research Institute of Forestry, Chinese Academy of Forestry, State Key Laboratory of Forest Genetics and Breeding, Key Laboratory of Forest Cultivation, National Forestry and Grassland Administration, Beijing 100091, China; 2. College of Horticulture, Hebei Agricultural University, Baoding 071001, Hebei, China; 3. China Flower Association, Beijing 100020, China)

Abstract: [Objective] To obtain the key NAC genes involving in regulating the root growth and development of yew trees. [Method] Basing on the full-length transcriptome data of *Taxus × media*, bioinformatics methods were used to identify NAC transcription factors. The protein structure and gene tissue expression profile of the selected NAC genes were also analyzed. [Result] A total of 44 NAC transcription factors were identified, which have typical conserved NAC domains at the N-terminal, and most of the members contain 5 conserved structural functional subdomains. Phylogenetic tree analysis showed that 44 *T. × media* NAC (TmNAC) proteins were clustered into 7 subfamilies of *Arabidopsis*. Sequence structure analysis indicated that the tertiary structures of the TmNAC proteins were similar, but the structures of the representative member of different subfamily also showed obvious differences. The expression results showed that *TmNAC15, 16, 18, 21, 22, 29, 39, 40, 41, and 44* had higher expression level in roots than stems and needles. [Conclusion] In this study, 44 TmNAC transcription factors were identified from *T. × media* and they were clustered into 7 subfamilies. Among them, *TmNAC21, 22, 39, 40, and 44* may be involved in the root growth and development progresses of yew trees.

Keywords: *Taxus* spp.; NAC gene; gene identification; expression analysis

(责任编辑: 彭南轩)