

doi:10.11937/bfyy.20174346

拟南芥 *AtOFP8* 的生物信息学分析及表达分析

唐尧¹, 张微^{1,2}, 尹艳莉¹, 冯鹏¹, 陈宇峰¹, 常缨¹

(1. 东北农业大学 生命科学院, 黑龙江 哈尔滨 150030; 2. 黑龙江省农业科学院 克山分院, 黑龙江 克山 160000)

摘 要: 以 Columbia-0 拟南芥为试材, 采用生物信息学方法对拟南芥转录因子 *AtOFP8* 进行生物信息学分析, 应用荧光定量 PCR 技术对 *AtOFP8* 基因进行组织表达分析和逆境表达分析, 为进一步研究其它 *AtOFP* 的功能奠定基础。结果表明: *AtOFP8* 基因全长 954 bp, 编码蛋白由 221 个氨基酸组成, 是一个分子量为 25.466 kDa 的不稳定蛋白, 该蛋白无跨膜区和信号肽, 定位在细胞核中, 亲水能力较强; *AtOFP8* 基因在根和花中表达较高, 在高温干旱等环境下转录水平明显升高。

关键词: 拟南芥; 转录因子; *AtOFP8*; 生物信息学

中图分类号: Q 947 **文献标识码:** A **文章编号:** 1001-0009(2018)18-0035-07

拟南芥 (*Arabidopsis thaliana*) 属十字花科双子叶草本植物, 因为其基因组小, 生长周期短, 而且很早以前就完成基因组测序, 所以一直以来备受研究人员关注, 成为研究其它双子叶植物基因功能的最佳工具, 但是, 对于作为模式植物的拟南芥自身而言, 还有太多的秘密等待研究人员去

探索^[1-2]。

转录因子 (transcriptional factor) 又称反式作用因子, 能够与下游靶基因启动子中的顺式作用元件结合, 从而调控靶基因的表达^[3]。典型的转录因子通常含有 DNA 结合区 (DNA-binding domain)、转录调控区 (activation domain)、寡聚化位点 (oligomerization site)、核定位信号 (nuclear localization signal) 等功能区域, 这些功能区域决定着转录因子的功能与特性, 使不同的转录因子各司其职, 调控植物体内基因的表达, 参与植物体生长发育、胁迫应答等各个方面, 确保植物体适应外界环境生存繁衍^[4-6]。

卵形家族蛋白 (ovate family protein, OFP) 是植物特有的转录因子家族, 均含有一个由约 70 个

第一作者简介: 唐尧 (1992-), 男, 硕士研究生, 研究方向为植物资源学与分子植物学。E-mail: 2427472516@qq.com.

责任作者: 常缨 (1970-), 女, 博士, 教授, 博士生导师, 研究方向为植物资源学与分子生物学。E-mail: changying@neau.edu.cn.

基金项目: 国家自然科学基金资助项目 (31370221)。

收稿日期: 2018-03-22

order to provide reference for the application of the strain in the biological control of grape disease. The results showed that 50 proteins were identified from the secreted proteins, and the relative molecular weight was all more than 10 kDa. By gene ontology of identified proteins were classified to the identification of proteins involved in carbohydrate metabolism, energy metabolism, and lipid metabolism, amino acid metabolism process, at the same time as the composition of cellular components involved in nucleic acids, membrane structure, cytoskeleton, and cell wall. Antibacterial proteins of chitinase and β -1,3-glucanase related to biological control were found, which provided a reference for the application of *Bacillus licheniformis* in the prevention and control of grape diseases.

Keywords: *Bacillus licheniformis*; biocontrol; gene ontology; secreted protein

氨基酸组成的 Ovate 保守结构域^[7]。这个转录因子家族存在于多种植物中,水稻(*Oryza sativa*)中共有 33 个,玉米(*Zea mays*)中共有 45 个,拟南芥中共有 19 个^[8]。对于拟南芥卵形家族蛋白(*Arabidopsis thaliana* ovate family protein, AtOFP)的研究很早就有报道,是一类新的转录因子,在植物生长发育的多个方面发挥作用^[9]。AtOFP1 能够抑制赤霉素生物合成过程中的关键酶基因 *AtGA20ox1* 的表达,从而抑制细胞的伸长^[10]。AtOFP4 能够和 KNAT7 相互作用,进而参与拟南芥次级细胞壁的形成^[11]。AtOFP5 能够在胚囊发育早期负调控 BLH1-KNAT3 蛋白复合物的活性^[12]。

采用生物信息学方法对 AtOFP8 蛋白的理化性质、蛋白结构、组织定位等情况进行全面分析,并且检测正常生长环境下各组织的转录水平及逆境胁迫下的转录水平,初步确定 AtOFP8 在拟南芥中所起的作用,为之后研究其具体功能提供帮助,也为研究同家族其它 AtOFP 提供线索和方向,对研究植物 OFP 具有重大意义。

1 材料与方法

1.1 试验材料

试验选用 Columbia-0 作为野生型拟南芥进行试验。拟南芥种子用 75% 酒精浸泡 3 min 消毒,换用 95% 酒精润洗后晾干,种在装有 1/2MS 培养基(含 3% 蔗糖、0.8% 琼脂)的平板上,用封口膜密封后放在 4℃ 冰箱中低温处理 60 h,然后拿到光下培养。待幼苗长出 4 片真叶,移栽到含蛭石的土壤($v:v=1:1$)中。组织表达选用生长

到 35 d 的拟南芥进行试验,逆境表达选用生长到 21 d 的拟南芥进行试验。

RNA 提取试剂盒、反转录试剂盒、荧光定量试剂盒均购自北京全式金生物技术有限公司。

1.2 试验方法

1.2.1 生物信息学分析

应用 NCBI 和 TAIR 数据库查询基因的基本信息;预测蛋白质理论等电点及分子质量应用 ExPASy 数据库中的 ProtParam;蛋白质二级结构预测应用 SOPMA;预测蛋白质三级结构应用 Phyre2 软件;对蛋白质的亲水性分析应用 ProtScale 软件进行预测;亚细胞定位预测采用 SUBA4 软件;跨膜区域使用 TMHMM Server 2.0 软件进行预测;信号肽使用 SignalP 4.1 Server 软件分析;保守结构域使用 InterPro 软件进行分析;网站链接见表 1。

1.2.2 *AtOFP8* 基因的荧光定量试验

应用 RNA 提取试剂盒及反转录试剂盒进行 RNA 的提取,并反转录成 cDNA。荧光定量引物应用 Primer Premier 5.0 软件进行设计,上游引物 AtOFP8-F: 5'-CCCACCAAACCCTTTC-TACGA-3';下游引物 AtOFP8-R: 5'-GGGT-CAGAACCAAACCTGCGAT-3';以基因 *AtActin2* 作为内参基因,上游引物 AtActin2-F: 5'-TGGGTTTTTACTTACGTCTGCG-3';下游引物 AtActin2-R: 5'-GGGAACAAAAGGAATA-AAGAGGC-3'。根据荧光定量试剂盒说明书进行加样,PCR 程序为:95℃、30 s;95℃、5 s,57℃、15 s,72℃、10 s;共 40 个循环。每个样本进行 3 次独立重复试验,应用 $2^{-\Delta\Delta Ct}$ 方法计算基

表 1 生物信息学研究的软件网址

Table 1 Website of software using for bioinformatics

软件 Software	网址 Website
NCBI	https://www.ncbi.nlm.nih.gov/
TAIR	http://www.arabidopsis.org/
ExPASy	https://www.expasy.org/
ProtParam	https://web.expasy.org/protparam/
SOPMA	https://npsa-prabi.ibcp.fr/cgi-bin/
Phyre2	http://www.sbg.bio.ic.ac.uk/phyre2/
ProtScale	https://web.expasy.org/protscale/
SUBA4	http://suba.live/
TMHMM Server 2.0	http://www.cbs.dtu.dk/services/TMHMM/
SignalP 4.1 Server	http://www.cbs.dtu.dk/services/SignalP/
InterPro	http://www.ebi.ac.uk/interpro/

因的相对表达量,用 Excel 软件进行绘图。

2 结果与分析

2.1 *AtOFP8* 基因的分析

从 NCBI 及 TAIR 数据库中了解到, *AtOFP8* 基因 (At5g19650) 位于拟南芥第 5 号染色体上, 全长 954 bp, 无内含子, 其中 CDS 区 666 bp, 5'UTR 长 111 bp, 3'UTR 长 177 bp。

2.2 AtOFP8 蛋白质的理化性质分析

AtOFP8 蛋白由 221 个氨基酸组成,应用 ProtParam 软件对蛋白质理化性质进行分析,结

果表明, AtOFP8 蛋白的分子式为 $C_{1118}H_{1755}N_{315}O_{348}S_9$, 分子量为 25.466 kDa, 理论等电点为 9.05, 脂肪指数为 56.88, 稳定系数为 67.78, 根据稳定系数数值小于 40 属于稳定蛋白的标准, 推测 AtOFP8 蛋白质属于不稳定蛋白^[13]。通过对 *AtOFP8* 基因编码的氨基酸进一步分析发现, AtOFP8 蛋白质由 20 种氨基酸组成, 其中 Ser 含量最多, 占氨基酸总数的 16.70%, Arg、Glu、Lys 次之, Trp 含量最少, 仅占氨基酸总数的 0.90%; 带正电荷的氨基酸残基(Arg+Lys)总数为 37 个, 带负电荷的氨基酸残基(Asp+Glu)总数为 31 个(图 1)。

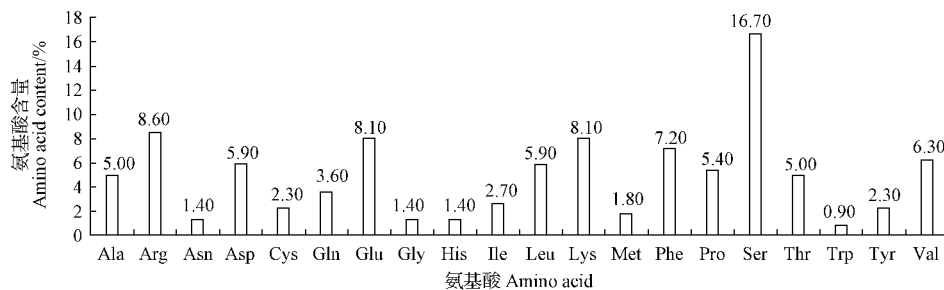


图 1 AtOFP8 蛋白质的氨基酸组成分析

Fig. 1 Amino acid composition analysis of the AtOFP8 protein.

2.3 AtOFP8 蛋白质的二级结构及三级结构分析

运用 SOPMA 对 *AtOPF8* 编码的蛋白质进行二级结构分析,结果表明,在 *AtOPF8* 蛋白中, α 螺旋(Alpha helix)和无规卷曲(Random coil)含量最多,分别占 40.27%和 43.44%,是 *AtOPF8* 蛋白质二级结构的主要构成元件,延伸链

(Extended strand) 占 9.95%, β 转角 (Beta turn) 占 6.33% (图 2); 将 AtOFP8 蛋白的氨基酸序列发送至 Phyre2 进行蛋白质三级结构预测, 并用分子图像观察软件 rasmol 对 AtOFP8 的三级结构进行分析, 结果显示 AtOFP8 蛋白质三级结构中二级结构构成元件与二级结构预测结果基本一致 (图 3)。

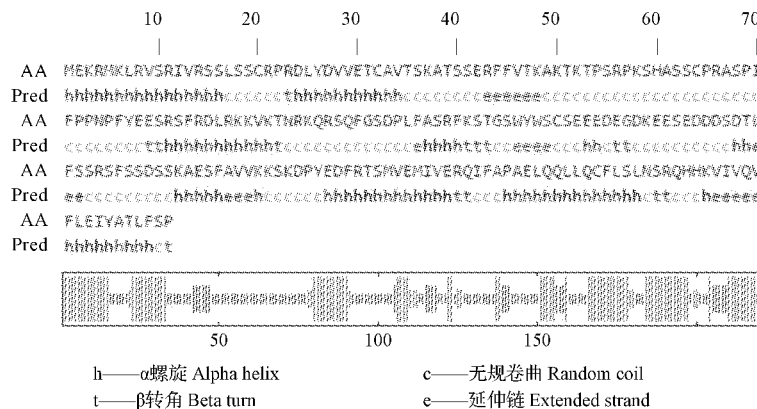


图 2 AtOFP8 蛋白质的二级结构预测

Fig. 2 Secondary structure prediction of the AtOFP8 protein

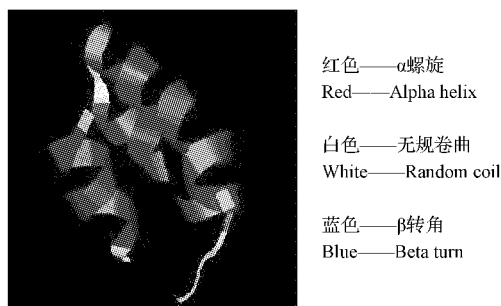


图3 AtOFP8蛋白质的三级结构预测

Fig. 3 Tertiary structure prediction of the AtOFP8 protein

2.4 AtOFP8蛋白质的亲水性分析

分析 AtOFP8 蛋白质亲水性的结果表明,位于该蛋白第 210 位的缬氨酸(Val)分值最高,为 2.356;位于该蛋白第 132 位的丝氨酸(Ser)分值最低,为-3.244。一般亲水能力强的氨基酸分值低,疏水能力强的氨基酸分值高,所以第 210 位的缬氨酸疏水能力最强,第 132 位的丝氨酸亲水能力最强。从整体看,AtOFP8 蛋白质的亲水性氨基酸数目多于疏水性氨基酸,总的亲水性平均系数(grand average of hydropathicity, GRAVY)为-0.717,亲水能力较强(图 4)。

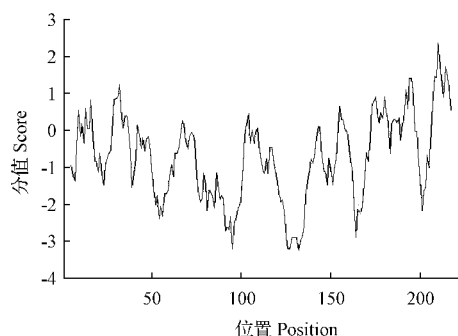


图4 AtOFP8蛋白质的亲水性分析

Fig. 4 Hydrophilic analysis of the AtOFP8 protein

2.5 AtOFP8蛋白质的保守结构域分析

应用 InterPro 对 AtOFP8 蛋白质的氨基酸序列进行分析发现,作为 AtOFP 家族中的一员,AtOFP8 蛋白与家族中其它成员一样,在 C 末端有一个近 70 AA 长的 Ovate 保守结构域,这是这个家族的特征结构域,但是除了这个结构域外,AtOFP8 蛋白中未预测到其它已知的结构域(图 5)。

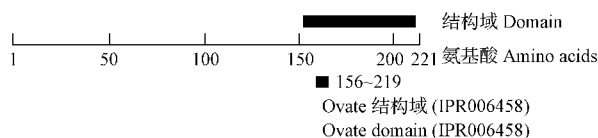


图5 AtOFP8蛋白质保守结构域的分析

Fig. 5 Analysis of conserved domains about the AtOFP8 protein

2.6 AtOFP8蛋白质的亚细胞定位预测分析

AtOFP8 是转录因子 AtOFP 家族中的一员,应用最新的拟南芥亚细胞定位预测网站 SUBA4 对其进行亚细胞定位预测分析,结果显示,AtOFP8 蛋白定位在细胞核中的 SUBAcon 数值达到 0.975,主要定位在细胞核中(图 6)。

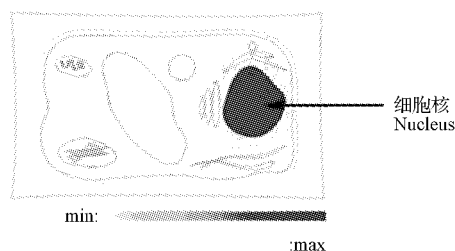


图6 AtOFP8蛋白质的亚细胞定位预测

Fig. 6 Subcellular localization prediction of the AtOFP8 protein

2.7 AtOFP8蛋白质的跨膜区域及信号肽分析

利用 TMHMM Server 2.0 在线分析软件对 AtOFP8 蛋白进行在线分析,结果显示,AtOFP8 蛋白预测的跨膜螺旋(transmembrane helices, TMHs)数量为 0,说明 AtOFP8 蛋白不是跨膜蛋

白(图 7)。利用 SignalP 4.1 Server 在线分析软件对 AtOFP8 蛋白 N 端前 70 个氨基酸进行信号肽预测分析,结果显示,在 N 端前 70 个氨基酸中,最大的 S 值为 0.206,最大的平均 S 值为

0.133,根据平均 S 值大于 0.5 的蛋白质是分泌蛋白的标准,推测 AtOFP8 蛋白不具有信号肽,是非分泌蛋白(图 8)。

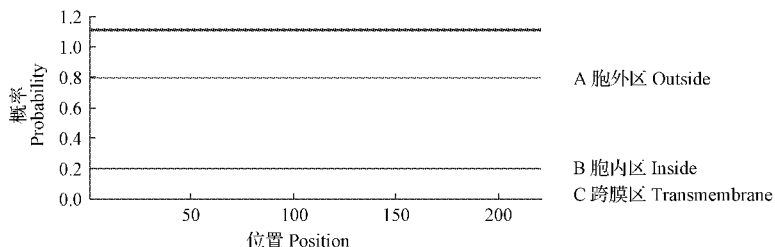


图 7 AtOFP8 蛋白质的跨膜区域预测

Fig. 7 Transmembrane region prediction of the AtOFP8 protein

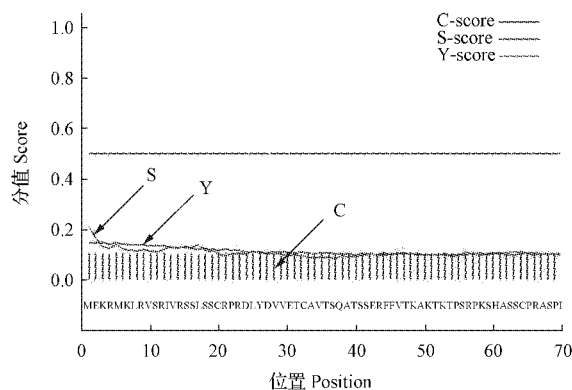


图 8 AtOFP8 蛋白质的信号肽预测

Fig. 8 Signal peptide prediction of the AtOFP8 protein

2.8 *AtOFP8* 基因的组织表达

对生长到 35 d 的拟南芥植株进行各组织的表达分析,结果显示,*AtOFP8* 基因在拟南芥根中的转录水平最高,是在莲座叶中转录水平的 2.8 倍,在花中的转录水平次之,约是莲座叶中的 2.0 倍,在茎和茎生叶中的转录水平比莲座叶略低(图 9)。

2.9 *AtOFP8* 基因在逆境条件下的表达

干旱条件处理下,拟南芥莲座叶中 *AtOFP8* 基因的转录水平明显上升,干旱处理 9 d 时达到最大,转录水平上升近 3 倍(图 10)。35 ℃ 高温处理下,*AtOFP8* 基因的转录水平先上升后下降,然后再次上升,处理 12 h 时上升最多,是正常环境生长的莲座叶转录水平的 2 倍(图 11)。4 ℃ 环境

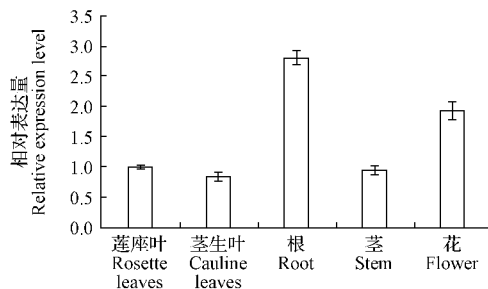
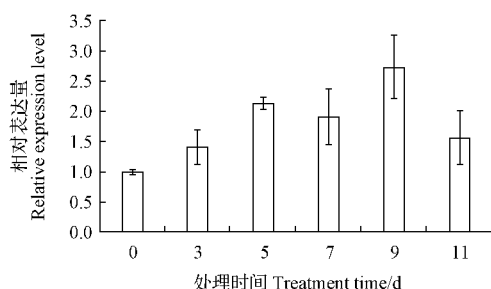
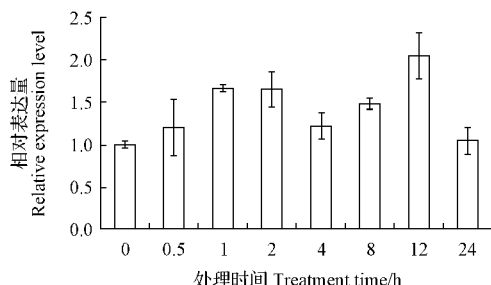
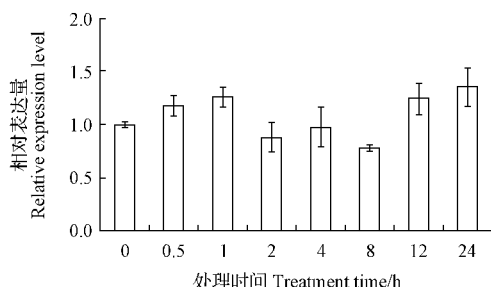
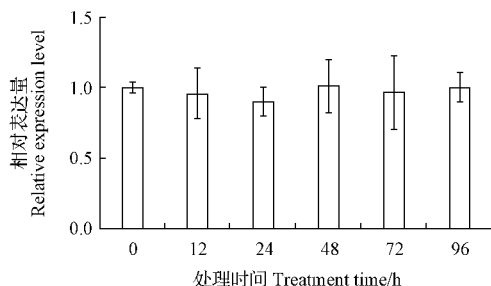


图 9 *AtOFP8* 基因的组织表达

Fig. 9 Tissue expression of the *AtOFP8* gene

下,*AtOFP8* 基因的转录水平有些波动,变化不大(图 12)。黑暗条件下,*AtOFP8* 基因的转录水平基本不发生变化(图 13)。

图 10 *AtOFP8* 基因在干旱条件下的表达Fig. 10 Expression of the *AtOFP8* gene under drought conditions图 11 *AtOFP8* 基因在 35 °C 环境下的表达Fig. 11 Expression of the *AtOFP8* gene at 35 °C图 12 *AtOFP8* 基因在 4 °C 环境下的表达Fig. 12 Expression of the *AtOFP8* gene at 4 °C图 13 *AtOFP8* 基因在黑暗条件下的表达Fig. 13 Expression of the *AtOFP8* gene in dark conditions

3 结论与讨论

随着计算机技术的发展以及大数据时代的到来,各种生物信息数据库以及预测软件得到更全面的更新,应用生物信息学技术分析出来的结果也更令人信服^[14]。应用生物信息学对 *AtOFP8* 分析发现,*AtOFP8* 蛋白是一个分子量为 25.466 kDa 的亲水性蛋白,既不是跨膜蛋白又不是分泌蛋白的不稳定蛋白。*AtOFP8* 作为 *AtOFP* 转录因子家族中的一员,具有该家族的典型特征,在 C 末端有一个大概 70 AA 大小的 Ovate 保守结构域,但是遗憾的是结构域预测未找到其它已知功能的结构域,未能为 *AtOFP8* 功能研究提供新的方向。转录因子是在细胞核中结合下游靶基因启动子从而调控其表达的蛋白质^[15],亚细胞定位预测显示 *AtOFP8* 定位在细胞核中,与 *AtOFP8* 作为转录因子在细胞核中行使功能的观点相一致,另外,跨膜区域预测结果证明其并非跨膜蛋白,信号肽预测结果显示其不是分泌蛋白,这 2 个结果也与其作为转录因子发挥功能的结果相符。

在生长发育过程中,植物不断遭遇各种逆境胁迫的干扰,最终进化产生一系列能够适应环境的转录因子,受到逆境侵害时就会作出应答调控下游靶基因的表达,抵抗逆境维持生命^[16]。为探究转录因子 *AtOFP8* 是否能够帮助拟南芥抵抗逆境,于是进行了拟南芥在逆境下的表达分析,结果表明 *AtOFP8* 基因在干旱条件和 35 °C 高温条件下转录水平明显升高,在 4 °C 低温环境处理下转录水平略有波动,而在黑暗条件下转录水平基本不变,推测转录因子 *AtOFP8* 可能参与抵抗干旱高温等恶劣环境的生物过程,但是 *AtOFP8* 在其中如何发挥作用还需进一步研究。

参考文献

- [1] 周太炎. 中国植物志(33 卷)[M]. 北京:科学出版社,1987.
- [2] 杨淑萍,马伟琴,刘阳,等. 模式植物拟南芥研究进展的文献计量分析[J]. 北方园艺,2015(19):195-201.
- [3] 李杰. 植物转录因子与基因调控[J]. 生物学通报,2004,39(3):9-10.
- [4] 刘强,张贵友,陈受宜,等. 植物转录因子的结构与调控作用[J]. 科学通报,2000,45(14):1465-1474.
- [5] 曹志方,李洁,刘强. 植物转录因子的结构和功能[J]. 植物科学进展,2000(3):95-108.

- [6] 杨致荣,王兴春,李西明,等. 高等植物转录因子的研究进展[J]. 遗传,2004,26(3):403-408.
- [7] HACKBUSCH J, RICHTER K, MULLER J, et al. A central role of *Arabidopsis thaliana* ovate family proteins in networking and subcellular localization of 3-aa loop extension homeodomain proteins[J]. Proc Natl Acad Sci USA, 2005, 102: 4908-4912.
- [8] LIU D, SUN W, YUAN Y W, et al. Phylogenetic analyses provide the first insights into the evolution of OVATE family proteins in land plants[J]. Ann Bot, 2014, 113: 1219-1233.
- [9] WANG S C, CHANG Y, GUO J J, et al. *Arabidopsis* Ovate Family Proteins, a novel transcriptional repressor family, control multiple aspects of plant growth and development[J]. PLoS One, 2011(6): e23896.
- [10] WANG S C, CHANG Y, GUO J J, et al. *Arabidopsis* Ovate Family Protein 1 is a transcriptional repressor that suppresses cell elongation[J]. Plant J, 2007(50): 858-872.
- [11] LI E Y, WANG S C, LIU Y Y, et al. OVATE FAMILY PROTEIN4 (OFP4) interaction with KNAT7 regulates secondary cell wall formation in *Arabidopsis thaliana* [J]. Plant J, 2011, 67: 328-341.
- [12] PAGNUSSAT G C, YU H J, SUNDARESAN V. Cell-fate switch of synergid to egg cell in *Arabidopsis* eostre mutant embryo sacs arises from misexpression of the BEL1-like homeodomain gene BLH1[J]. Plant Cell, 2007(19): 3578-3592.
- [13] 范晶,胥成浩,张西玉,等. 番茄 LeNHX3 基因的生物信息学分析[J]. 湖北农业科学, 2009, 48(12): 2917-2921.
- [14] 高媛. 后基因组时代的生物信息学发展[J]. 中国科技信息, 2009(10): 225-226.
- [15] 曹丽霞,马有志,杨俊英. 应答非生物胁迫的植物转录因子[J]. 内蒙古农业科技, 2005(5): 10-14.
- [16] 李红霞,汪好,张战凤. 植物转录因子与作物抗逆胁迫关系的研究进展[J]. 麦类作物学报, 2013, 33(3): 613-618.

Bioinformatics Analysis and Expression Analysis of *AtOFP8* in *Arabidopsis thaliana*

TANG Yao¹, ZHANG Wei^{1,2}, YIN Yanli¹, FENG Peng¹, CHEN Yufeng¹, CHANG Ying¹

(1. College of Life Sciences, Northeast Agricultural University, Harbin, Heilongjiang 150030; 2. Keshan Branch, Heilongjiang Academy of Agricultural Sciences, Keshan, Heilongjiang 160000)

Abstract: The Columbia-0 *Arabidopsis thaliana* was used as material. Bioinformatics analysis of *Arabidopsis* transcription factor AtOFP8 was carried out by using bioinformatics methods and AtOFP8 gene expression analysis was used by fluorescence quantitative PCR technology. Both of them laid the foundation for further research on the function of other AtOFPs. The results showed that the full-length of AtOFP8 gene was 954 bp, AtOFP8 protein was composed of 221 amino acids, its molecular weight was 25.466 kDa, it had no transmembrane domain and signal peptide, it was located in the nucleus and it had strong hydrophilicity. In addition, AtOFP8 gene was highly expressed in roots and flowers, and its transcriptional level was obviously elevated in the environment of high temperature and drought.

Keywords: *Arabidopsis thaliana*; transcription factor; AtOFP8; bioinformatics