

夏雯雯, 李 锦, 祝建波. 天山雪莲水孔蛋白基因(*AQP*)家族鉴定与生物信息学分析[J]. 江苏农业学报, 2016, 32(6): 1244-1250.
doi:10.3969/j.issn.1000-4440.2016.06.008

天山雪莲水孔蛋白基因(*AQP*)家族鉴定与生物信息学分析

夏雯雯, 李 锦, 祝建波

(石河子大学生命科学学院农业生物技术重点实验室, 新疆 石河子 832000)

摘要: 为研究新疆天山雪莲水孔蛋白基因(*AQP*)家族并对其基因家族成员进行鉴定及分类, 从雪莲转录组数据库中获得33条天山雪莲水孔蛋白基因序列, 利用生物信息学软件, 将其与拟南芥中38条水孔蛋白基因序列进行分析、分类和保守序列分析, 结果显示天山雪莲33条*AQP*基因家族序列与拟南芥38条*AQP*基因序列一样分成4个亚族, 即质膜水孔蛋白(Plasma membrane intrinsic protein, PIP)基因、液泡膜内在水孔蛋白(Tonoplast intrinsic protein, TIP)基因、类Nodulin(NOD26)膜内在水孔蛋白(NOD26-like membrane intrinsic protein, NIP)基因和膜内小基本蛋白(Small and basic membrane intrinsic protein, SIP)基因。拟南芥与天山雪莲*AQP*基因间具有较高的保守性。对天山雪莲*AQP*基因家族进行了系统进化树构建, 并分析了水孔蛋白氨基酸组成成分、理化性质, 预测了二级结构, 发现水孔蛋白不同亚族间氨基酸数目、氨基酸序列间的亲水性存在一定差异, 水孔蛋白的二级结构以 α 螺旋、 β 折叠、无规则卷曲为主。从这33条*AQP*基因序列中克隆了3个水孔蛋白基因, 测序结果证明天山雪莲转录组数据库基因序列准确率较高。

关键词: 天山雪莲; 拟南芥; 水孔蛋白基因家族; 生物信息学

中图分类号: Q811.4 **文献标识码:** A **文章编号:** 1000-4440(2016)06-1244-07

Identification and bioinformatics analysis of aquaporin gene (*AQP*) family in *Saussurea involucre* Kar.

XIA Wen-wen, LI Jin, ZHU Jian-bo

(Laboratory of Agricultural Biotechnology, College of Life Science, Shihezi University, Shihezi 832000, China)

Abstract: To identify the aquaporin (*AQP*) gene family of *Saussurea involucre* Kar., 33 sequences of the *AQP* were acquired from transcriptome database. Using a variety of bioinformatics softwares, the gene sequences of *S. involucre* Kar. *AQP* and *Arabidopsis thaliana* *AQP* were analyzed and classified. 33 *S. involucre* Kar. *AQP* genes as well as 38 *Arabidopsis* *AQP* genes were classified into four subfamilies, plasma membrane intrinsic protein gene (*PIP*), tonoplast intrinsic protein gene (*TIP*), NOD26-like membrane intrinsic protein gene (*NIP*) and small and basic membrane intrinsic protein gene (*SIP*). *AQP* genes were highly conservative between *Arabidopsis* and *S. involucre* Kar.. There were some differences in the number of amino acids and the hydrophilicity of amino acid sequence among different subfamilies of *AQP* gene. The secondary structure of *AQP* protein were dominated by

alpha helix, beta bridge, and random coil. The sequencing of three *AQP* gene cloned from 33 *AQP* gene sequences verified the accuracy of gene sequences in transcriptome database.

Key words: *Saussurea involucre* Kar.; *Arabidopsis thaliana*; *AQP* gene family; bioinformatics

收稿日期: 2015-10-22

基金项目: 国家自然科学基金项目(C020406)

作者简介: 夏雯雯(1988-), 女, 新疆乌鲁木齐人, 硕士研究生, 主要从事植物基因工程研究。(E-mail) 365689825@qq.com

通讯作者: 祝建波, (E-mail) 274831213@qq.com

水孔蛋白(Aquaporin, AQP)是能选择性地高效转运水分子的通道蛋白,分子质量一般为23 000~31 000,属于古老的膜内在蛋白超家族,具有典型的跨膜结构域^[1-2]。截止到目前,已经发现了大量的水孔蛋白基因,广泛存在于植物、动物、微生物等各类生物体中。已发现的水孔蛋白具有高度保守的结构特征与功能结构域。水孔蛋白功能的确定最早是在1992年,Preston等^[3]利用爪蟾卵母细胞异源表达系统成功证明了从人红细胞质膜分离的CHIP28的水转运活性,CHIP28作为第一个被证实的水通道蛋白被定名为AQP1。

根据水孔蛋白不同的亚细胞定位区域,可以将水孔蛋白分为4类:质膜水孔蛋白(Plasma membrane intrinsic protein, PIP)、液泡膜内水孔蛋白(Tonoplast intrinsic protein, TIP)、类Nodulin(NOD26)膜内水孔蛋白(NOD26-like membrane intrinsic protein, NIP)和膜内小基本蛋白(Small and basic membrane intrinsic protein, SIP)^[4-5]。根据水孔蛋白的氨基酸序列特征又可将4类水孔蛋白次分,如质膜水孔蛋白分为PIP1、PIP2和PIP3共3种类型。不同类型的水孔蛋白,在植物生长发育过程和响应胁迫刺激时功能也有所不同^[6]。

植物水孔蛋白在实现水分的快速跨膜运输,调节细胞内外的渗透平衡中起着重要作用,具有生物调节的功能。Sahara^[7]以及Tanghe^[8]采用全基因组基因表达分析对不同的酵母耐冻菌株和冷冻敏感型菌株抗冻性进行分析,发现了水通道蛋白基因(AQP1和AQP2)与抗冻性有很大相关性。进一步在酵母细胞中删除这些基因,菌株对冰冻更敏感,而互补这些基因表达或过表达一个外源的AQP基因则提高了酵母的抗冻性能,表明水孔蛋白在酵母的抗冻性方面起着重要作用。

新疆天山雪莲(*Saussurea involucreata* Kar.)是新疆特有的珍奇名贵植物,又名新疆雪莲、大苞雪莲等,俗名雪兔子,属菊科(Compositae)凤毛菊属(*Saussurea*)雪莲亚属植物,为多年生一次性开花结实的高山冰缘草本植物^[9],天山雪莲常年生长在海拔雪线附近的悬崖峭壁,为适应这种严寒、缺氧、高辐射等极端的自然环境,新疆雪莲形成了一套与环境相适应的特殊生理和生化机制,主要表现为在低温条件下能够正常生长发育。由于水孔蛋白在植物的生长、发育及抗逆中发挥着不可忽视的作用,因

此,本研究拟利用已分离的拟南芥AQP基因编码的氨基酸序列,根据雪莲转录组数据库信息,查找比对雪莲AQP基因序列,对雪莲水孔蛋白基因家族进行分类、序列同源性分析、进化关系研究,为新疆天山雪莲AQP基因功能分析提供信息参考,为阐明AQP基因家族在雪莲生长发育中的调控作用奠定理论基础。

1 材料与方法

1.1 天山雪莲AQP基因家族成员的确定

从雪莲转录组数据库中分离所有水孔蛋白基因序列,利用在线工具Pfam(<http://pfam.janelia.org/>)和SMART(<http://smart.embl-heidelberg.de/>)^[9]对获得的候选雪莲AQP家族成员的氨基酸序列进行保守MIP蛋白结构域预测,存在该结构域即属于AQP基因家族。

1.2 天山雪莲AQP家族蛋白质结构分析

天山雪莲AQP家族蛋白质二级结构分析是利用在线工具SOPMA(http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_sopma.html)^[10]进行,预测其 α -螺旋、 β -折叠和无规则卷曲等在整体结构中所占的比例。

1.3 天山雪莲和拟南芥中AQP蛋白的保守氨基酸序列分析

利用MegAlign生物信息分析软件压缩包分别对拟南芥和天山雪莲NIP、PIP、TIP以及SIP进行多序列比对,分析比对结果。

1.4 天山雪莲与拟南芥AQP蛋白质系统发生分析

由于拟南芥和天山雪莲AQP蛋白的氨基酸序列具有高度保守的功能结构域,因此可以进行多序列的比对和构建系统进化树。拟南芥AQP基因及其编码的氨基酸序列来源于TAIR(<http://www.arabidopsis.org>)数据库。AQP氨基酸序列多重比对由ClustalX软件完成,参数为默认值^[11]。采用邻接法(Neighbor-joining method)构建系统发生树^[12],其输出用MEGA软件^[13]完成。

1.5 天山雪莲PIP基因的克隆与序列分析

1.5.1 *sikPIP1-5*、*sikPIP2-1*、*sikPIP2-7*基因引物设计 从天山雪莲水孔蛋白基因家族中选取了3条基因(*sikPIP1-5*、*sikPIP2-1*、*sikPIP2-7*)序列,利用Primer 5.0设计PCR引物(表1)进行克隆,并对克隆结果进行测序及序列分析。

表 1 PCR 分析用引物

Table 1 Gene primers for PCR

基因	引物	引物序列 (5'→3')
<i>sikPIP1-5</i>	P1	TGCGGATCCGCATGAGTTTTAGAGAGAGAAATGTCG
	P2	TGCGTCGACATTAATTTGTAGCATTGCTTCTGA
<i>sikPIP2-1</i>	P1	TGCGGATCCTAATGGAGGGTAAGGAAGA
	P2	TGCGTCGACATAAGTTTAAGATCTGCTC
<i>sikPIP2-7</i>	P1	TGCGGATCCATGTCGAAAGAGGTGAGTGAG
	P2	TGCGTCGACATTTGTAGCATTGCTTCTGAA

1.5.2 *sikPIP1-5*、*sikPIP2-1*、*sikPIP2-7* 基因克隆

采用 Trizol 法提取天山雪莲总 RNA,用反转录试剂盒将 RNA 反转录成 cDNA。用上下游引物进行 PCR 扩增。同时以去离子水为模板作为阴性对照。PCR 反应程序为 95 ℃ 预变性 5 min;95 ℃ 变性 1 min,59 ℃ 复性 1 min,72 ℃ 延伸 1 min,30 个循环;72 ℃ 延伸 7 min,4 ℃ 保温。PCR 产物经琼脂糖凝胶电泳后,用胶回收试剂盒回收目的片段,在 16 ℃ 用 pGM-Teasy 载体连接 16~24 h,转化 TOP10 感受态细胞。经抗性筛选及 PCR 鉴定后,保存阳性克隆,并送华大基因进行测序。对测序结果进行生物信息学分析。

2 结果与分析

2.1 天山雪莲 AQP 基因家族成员的确定

利用 Pfam 和 SMART 对天山雪莲转录组数据库中的七十几条水孔蛋白基因序列进行功能域分析,通过 MIP 最终确定了 33 条天山雪莲 AQP 基因家族成员。从以往的研究结果来看,水稻有 37 条水孔蛋白基因^[14],其中 *PIP* 基因序列 14 条,*TIP* 基因序列 10 条,*NIP* 基因序列 11 条和 2 条 *SIP* 基因序列。拟南芥水孔蛋白基因家族包含 38 条基因序列^[15],其中 *PIP* 14 条,*TIP* 11 条,*NIP* 10 条,*SIP* 3 条。玉米含有 13 条水孔蛋白基因家族序列^[16],苔藓含 23 个 AQP,番茄含 37 个 AQP,白杨含 55 个 AQP 等^[17]。将这 33 条水孔蛋白氨基酸序列与拟南芥水孔蛋白氨基酸比对,根据其亚细胞定位将雪莲 AQP 基因家族成员分为 4 类:*sikPIP*、*sikNIP*、*sikTIP*、*sikSIP*,其中 *sikPIP* 中又可分为 *sikPIP1* 和 *sikPIP2* 亚类,共 17 条,*sikPIP1* 有 6 条,*sikPIP2* 有 11 条;*sikNIP* 共 4 条,为 *sikPIP1*、*sikPIP4*、*sikPIP5*、*sikPIP6*; *sikTIP* 共 10

条,其中 *sikTIP1* 有 4 条,*sikTIP2* 有 4 条,*sikTIP3* 和 *sikTIP5* 各 1 条。*sikSIP* 包括 *sikSIP1* 和 *sikSIP2* 2 条序列,他们编码的蛋白质长度不等,最短的含 225 个氨基酸,最长的含 888 个氨基酸。4 类水孔蛋白的等电点(*pI*)值最高的达到 5.42,最低的为 5.07。

2.2 天山雪莲 AQP 家族蛋白质结构分析

对 33 条天山雪莲 AQP 家族氨基酸序列分析表明,AQP 蛋白二级结构绝大多数由 α 螺旋、 β 折叠、 β 转角和无规则卷曲 4 种形式组成。其中 *sikPIP* 编码的蛋白质二级结构中 β 折叠、 β 转角和无规则卷曲略多于 α 螺旋。*sikNIP*、*sikSIP* 编码的蛋白质则以 α 螺旋、 β 折叠形式为主。*sikTIP* 编码的蛋白质二级结构中 β 折叠、 β 转角和无规则卷曲远多于 α 螺旋形式。33 条雪莲 AQP 基因编码的氨基酸序列分析显示其亲水性也存在细微差别(表 2)。

2.3 天山雪莲和拟南芥中 AQP 蛋白的保守序列分析

通过比对发现,拟南芥和天山雪莲的 *NIP* 氨基酸序列在 260 至 290 氨基酸之间存在 2 段共有的高度保守序列,该序列位于 MIP 保守功能域。拟南芥和天山雪莲的 *PIP1* 和 *PIP2* 在 80 至 90 氨基酸之间存在差异,天山雪莲 *sikPIP1* 和 *sikPIP2* 同样存在差异。拟南芥 *PIP* 和天山雪莲 *PIP* 的氨基酸序列存在 4 段高度保守区域。这种序列上的相似性说明天山雪莲水孔蛋白可能具有与拟南芥水孔蛋白相似的功能。拟南芥 3 条 *SIP* 和天山雪莲的 2 条 *SIP* 氨基酸序列比对结果显示,其高度保守区域集中在 130~240 氨基酸。将拟南芥和天山雪莲的 *TIP* 氨基酸序列比对可以发现,在 200~280 氨基酸两者表现出一致的差异性,其高度保守的序列也集中在这部分(表 3)。

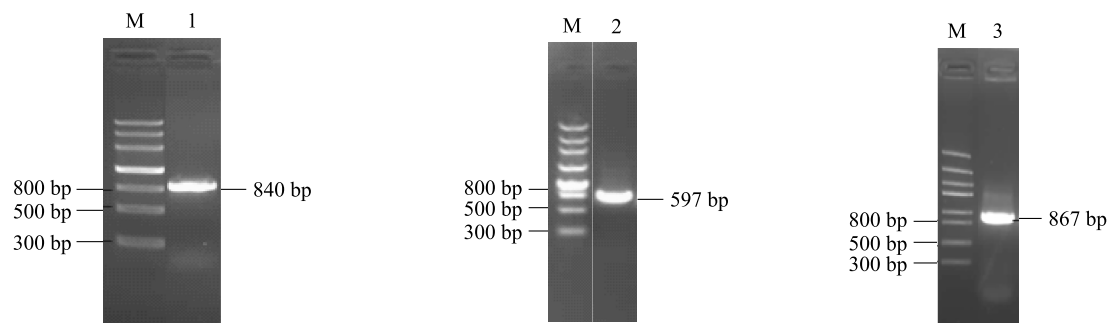
表2 天山雪莲 AQP 家族蛋白质结构分析

Table 2 The structure of proteins encoded by AQP family in *Saussurea involucre*

水孔蛋白	α -螺旋	β -折叠	β -转角	无规则卷曲	亲水性区域
sikNIP1	15(25.8%)	23(39.6%)	10(17.2%)	10(17.2%)	头、尾
sikNIP4	11(22.9%)	19(39.5%)	7(14.5%)	11(22.9%)	头、中后尾
sikNIP5	15(20.2%)	28(37.8%)	17(22.9%)	14(18.9%)	头尾
sikNIP6	0(0)	6(33.3%)	7(38.8%)	5(27.7%)	尾部
sikPIP1-1	15(23.0%)	21(32.3%)	15(23.0%)	14(21.5%)	头、中后
sikPIP1-2	17(25.7%)	21(31.8%)	15(22.7%)	13(19.7%)	头、中后尾
sikPIP1-3	2(14.2%)	7(50.0%)	4(28.5%)	1(7.1%)	中间偏后
sikPIP1-4	0(0)	8(40.0%)	7(35.0%)	5(25.0%)	中后部
sikPIP1-5	3(13.6%)	7(31.8%)	6(27.2%)	6(27.2%)	中前头部
sikPIP1-6	1(2.2%)	16(35.5%)	12(26.6%)	16(35.5%)	头、中后
sikPIP2-1	5(29.4%)	5(29.4%)	6(35.2%)	1(5.8%)	头、中间
sikPIP2-2	11(22.9%)	18(37.5%)	11(22.9%)	8(16.6%)	头、中后尾
sikPIP2-3	6(19.3%)	11(35.4%)	6(19.3%)	8(25.8%)	中前
sikPIP2-4	10(17.8%)	21(37.5%)	11(19.6%)	12(21.4%)	头、中后
sikPIP2-5	8(25.8%)	15(48.3%)	8(25.8%)	10(32.2%)	中间
sikPIP2-6	12(20.0%)	23(38.3%)	15(25.0%)	10(16.6%)	头、中后尾
sikPIP2-7	11(19.2%)	20(35.1%)	15(26.3%)	11(19.2%)	头、中后尾
sikPIP2-8	0(0)	3(27.2%)	5(45.5%)	3(27.2%)	中后尾
sikPIP2-9	0(0)	6(35.3%)	5(29.4%)	6(35.3%)	尾部
sikPIP2-10	3(16.7%)	8(44.4%)	5(27.8%)	2(11.1%)	头、中后
sikPIP2-11	3(33.3%)	3(33.3%)	2(22.2%)	1(11.1%)	头、中后
sikSIP1	13(24.1%)	21(38.8%)	9(16.7%)	11(20.3%)	中后尾
sikSIP2	13(24.1%)	21(38.8%)	9(16.7%)	11(20.3%)	中后尾
sikTIP1-1	8(17.4%)	17(37.0%)	9(19.6%)	12(26.1%)	后尾
sikTIP1-2	1(2.3%)	16(36.4%)	11(25.0%)	16(36.4%)	头尾
sikTIP1-3	11(21.1%)	18(34.6%)	10(19.2%)	13(25.0%)	头尾
sikTIP1-4	0(0)	17(40.4%)	13(30.9%)	12(28.5%)	头尾
sikTIP2-1	0(0)	13(39.4%)	12(36.4%)	8(24.2%)	中后尾
sikTIP2-2	8(16.3%)	19(28.8%)	11(22.4%)	11(22.4%)	中后尾
sikTIP2-3	11(20.8%)	21(39.6%)	12(22.6%)	9(16.9%)	中后尾
sikTIP2-4	10(25.0%)	18(45.0%)	11(27.5%)	11(27.5%)	中后尾
sikTIP3	9(24.3%)	15(40.5%)	8(21.6%)	5(13.5%)	头部
sikTIP5	1(2.2%)	16(35.5%)	12(26.7%)	16(35.5%)	头、中后

左右的 DNA 片段(图 2)。测序结果经生物信息学分析,3 条质膜内在水孔蛋白基因序列均有 MIP 保守功能域。与天山雪莲转录组数据库中的序列比

对,*sikPIP1-5*、*sikPIP2-1*、*sikPIP2-7* 序列一致性分别为 99.52%、100.00%、97.64%。



M; DNA marker; 1~3 分别代表基因 *sikPIP1-5*、*sikPIP2-1*、*sikPIP2-7*。

图 2 水孔蛋白基因 *sikPIP1-5*、*sikPIP2-1*、*sikPIP2-7* 的 PCR 扩增

Fig. 2 PCR amplification of aquaporin genes *sikPIP1-5*, *sikPIP2-1* and *sikPIP2-7*

3 讨论

植物体内的水孔蛋白(AQP)家族的功能主要是关于水、甘油和可溶性小分子的转运,其转运机制目前并不是十分清楚。目前已知不同类型的水孔蛋白在植株生长各个阶段不同部位有序调控表达,从而完成植物本身一个完整生命过程。

前期我们从已构建的天山雪莲 cDNA 文库中筛选出 3 个水孔蛋白基因 *sikPIP1-5*、*sikPIP2-1* 和 *sikPIP2-7*,其中 *sikPIP2-7* 在转基因烟草中表现出了较强的抗冻和抗旱效果^[18]。该基因与拟南芥的 *AtPIP2-7* 同源性最高。*AtPIP2-7* 基因在拟南芥中不同器官和不同逆境胁迫中的响应是不同的^[16],该基因在拟南芥叶片中受干旱胁迫时表达下降,而在根中受低温诱导表达下调,却能被盐胁迫诱导表达上调。这种表达模式也可以用来推测 *sikPIP2-7* 基因抗冻和抗旱特性。

本研究首次通过生物信息学手段获得 33 个雪莲 AQP 基因。利用模式植物中功能已知的拟南芥 AQP 蛋白家族为参考序列,对天山雪莲测序数据结果进行 AQP 基因家族的鉴定,AQP 蛋白家族的组分和结构分析,AQP 基因家族进化和保守基序分析等系统研究,并克隆了 3 个天山雪莲质膜内在水孔蛋白基因,证明天山雪莲转录组数据库提供的基因序列与测序结果一致性达 95.00% 以上。本研究为将来揭示天山雪莲 AQP 基因家族的功能奠定了重要

基础,也为天山雪莲其他基因家族的研究和其他拥有大量测序数据的植物基因家族研究提供了参考信息。

参考文献:

- [1] AGRE P. Aquaporin water channels (Nobel lecture) [J]. Angewandte Chemie International Edition, 2004, 43(33): 4278-4290.
- [2] BENG A. Birth of water channel proteins—the aquaporins [J]. Cell Biology International, 2003, 27(9): 701-709.
- [3] PRESTON G M, CARROLL T P, GUGGINO W B, et al. Appearance of water channels in *Xenopus* oocytes expressing red cell CHIP28 protein [J]. Science, 1992, 256(5055): 385-387.
- [4] JOHANSSON U, KARLSSON M, JOHANSSON I, et al. The complete set of genes encoding major intrinsic proteins in *Arabidopsis* provides a framework for a new nomenclature for major intrinsic proteins in plants [J]. Plant Physiology, 2001, 126(4): 1358-1369.
- [5] JOHANSSON U, GUSTAVSSON S. A new subfamily of major intrinsic proteins in plants [J]. Mol Biol Evol, 2002, 19(4): 456-461.
- [6] AROCA R, PORCEL R, RUIZ-LOZANO J M. Regulation of root water uptake under abiotic stress conditions [J]. J Exp Bot, 2012, 63(1): 43-57.
- [7] SAHARA T, GODA T, OHGIYA S. Comprehensive expression analysis of time-dependent genetic responses in yeast cells to low temperature [J]. J Biol Chem, 2002, 277: 50015-50021.
- [8] TANGHE A. Aquaporin expression correlates with freeze tolerance in baker's yeast, and overexpression improves freeze tolerance in industrial strains [J]. Appl Environ Microbiol, 2002, 68: 5981-5989.

- [9] TOURNAIRE-ROUX C, SUTKA M, JAVOT H, et al. Cytosolic pH regulates root water transport during anoxic stress through gating of aquaporins[J]. *Nature*, 2003, 425(6956): 393-397.
- [10] UEHLEIN N, LOVISOLO C, SIEFRITZ F, et al. The tobacco aquaporin NtAQPI is a membrane CO₂ pore with physiological functions[J]. *Nature*, 2003, 425(6959): 734-737.
- [11] GE Y X, ANGENENTG C, DAHLHAUS E, et al. Partial silencing of the *NECI* gene results in early opening of anthers in *Petunia hybrida* [J]. *Mol Genet Genomics*, 2001, 265(3): 414-423.
- [12] SAITOU N, NEI M. The neighbor-joining method: a new method for reconstructing phylogenetic trees[J]. *Mol Biol Evo*, 1987, 4(4):406-425.
- [13] TAMURA K, DUDLEY J, NEI M, et al. MEGA4: molecular evolutionary genetics analysis (MEGA) software version 4.0[J]. *Mol Biol and Evo*, 2007, 24(8): 1596-1599.
- [14] WUDICK M M, LUU D T, MAUREL C. A look inside: localization patterns and functions of intracellular plant aquaporins [J]. *New Phytologist*, 2009, 184(2): 289-302.
- [15] PENG Y H, LI W L, CAI W M, et al. Overexpression of a *Panax ginseng* tonoplast aquaporin alters salt tolerance, drought tolerance and cold acclimation ability in transgenic *Arabidopsis* plants [J]. *Planta*, 2007, 226: 729-740.
- [16] LI X J, WANG X H, YANG Y, et al. Single-molecule analysis of PIP2-1 dynamics and partitioning reveals multiple modes of arabidopsis plasma membrane aquaporin regulation [J]. *The Plant Cell*, 2011, 23: 3780-3797.
- [17] BOURSIAIC Y B J, POSTAIRE O, LUU D T, et al. Stimulus-induced downregulation of root water transport involves reactive oxygen species-activated cell signalling and plasma membrane intrinsic protein internalization[J]. *The Plant Journal*, 2008, 56(2):207-218.
- [18] 焦天奇,孙 辉,刘瑞娜,等. 转天山雪莲 *sikPIP3* 基因烟草的获得及抗逆性鉴定[J]. *西北植物学报*, 2012, 32(3):431-438.

(责任编辑:孙 宁)