

DOI: 10.3969/j.issn.1000-3142.2014.02.020

魏翔莺,张春英,付海辉,等. 云锦杜鹃 *NRT* 基因序列的生物信息学分析[J]. 广西植物,2014,34(2):242-247Wei XY,Zhang CY,Fu HH,et al. Bioinformatics analysis of *NRT* gene sequences in *Rhododendron fortunei*[J]. *Guihaia*,2014,34(2):242-247

# 云锦杜鹃 *NRT* 基因序列的生物信息学分析

魏翔莺<sup>1,2</sup>, 张春英<sup>2</sup>, 付海辉<sup>3</sup>, 潘东明<sup>1\*</sup>

(1. 福建农林大学 园艺学院, 福州 350002; 2. 上海市园林科学研究所, 上海 200232; 3. 厦门大学 环境与生态学院, 厦门 361101)

**摘要:** 通过转录组测序,获得在接种 ERM 真菌的云锦杜鹃苗根系中显著差异表达的基因,其中硝酸根转运蛋白(*NRT*)基因是硝态氮吸收转运的关键基因。利用生物信息学方法,分析云锦杜鹃根转录组的硝酸根转运蛋白(*NRT*)基因序列,对其推导的氨基酸的理化性质、亲水性/疏水性、跨膜结构、导肽、二级结构、高级结构进行预测,并对硝酸根转运蛋白的氨基酸做进化发育分析。为进一步了解 *NRT* 基因在云锦杜鹃接种苗根系氮素吸收的作用奠定了基础。

**关键词:** 云锦杜鹃; 硝酸根转运蛋白; 生物信息学; 进化树

**中图分类号:** Q81      **文献标识码:** A      **文章编号:** 1000-3142(2014)02-0242-06

## Bioinformatics analysis of *NRT* gene sequences in *Rhododendron fortunei*

WEI Xiang-Ying<sup>1,2</sup>, ZHANG Chun-Ying<sup>2\*</sup>, FU Hai-Hui<sup>3</sup>, PAN Dong-Ming<sup>1\*</sup>

(1. College of Horticulture, Fujian Agriculture and Forestry University, Fuzhou 350002, China; 2. Shanghai Landscape and Gardening Research Institute, Shanghai 200232, China; 3. College of Environment and Ecology, Xiamen University, Xiamen 361101, China)

**Abstract:** We used next generation sequencing technology to investigate the transcriptomes between inoculated roots and uninoculated roots of the *Rhododendron fortunei* and obtained many differentially expressed genes. In this paper, the nucleic acid sequences and amino acid sequences of nitrate transporters from *R. fortunei*, were analyzed by bioinformatics tools. Several parameters of these sequences, including sequences composition, physicochemical property, leader peptide, topological structure of transmembrane regions, hydrophobicity or hydrophilicity, secondary structures, functional domains and protein structures, were predicted. Phylogenetic tree was reconstructed for the nitrate transporters protein family. Provide the bioinformatics foundation to understand *NRT* gene's function in inoculated seedling roots.

**Key words:** *Rhododendron fortunei*; *NRT*; bioinformatics; phylogenetic tree

云锦杜鹃(*Rhododendron fortunei*)属杜鹃花科常绿灌木,生于海拔 400~1 900 m 山间林中,花大艳丽,极具观赏价值,主要产于浙江、江西、福建、广东、广西、湖南等省区(方云亿,1989)。目前对云锦杜鹃的研究主要在药用成分(李钧敏等,2004)、育种(朱春艳等,2006;边才苗等,2005)、遗传多样性(金则

新等,2006)、杜鹃花类菌根(王剑敏等,2011;张春英,2008;杨兵等,2010)等方面,而在分子生物学方面的研究较少。

硝酸盐转运体(Nitrate transporters, NRTs)的主要作用是完成根系中  $\text{NO}_3^-$  的吸收与转运(Zhen et al., 1991; Walker et al., 1995)。番茄(Lauter, et

收稿日期: 2013-10-22      修回日期: 2013-11-25

基金项目: 国家自然科学基金(30972409)

作者简介: 魏翔莺(1987-),女,福建建瓯人,在读博士,主要从事花卉育种和分子生物学等研究,(E-mail)weixiangying\_cool@126.com。

\*通讯作者: 潘东明,教授,主要从事园艺产品采后生理生化和分子生物学等研究,(E-mail)pdm666@126.com。

al., 1996; von Wiren *et al.*, 2000)、油菜 (Pearson *et al.*, 2002)、水稻 (Suenaga *et al.*, 2003)、白菜 (D'Apuzzo *et al.*, 2004) 等许多物种中克隆了该基因, 积累了许多序列数据。目前, 如何对这些序列数据进行科学的分析已经成为生物学工作者亟待解决的主要问题。随着生物信息学的出现给序列数据分析带来了极大的便利。本研究用生物信息学方法对云锦杜鹃根转录组的硝酸根转运蛋白基因的核苷酸序列以及所对应的氨基酸序列的理化性质、疏水性/亲水性、导肽、跨膜区域、二级结构及高级结构等进行预测和分析, 并对 *NRT* 蛋白构建系统进化树, 为研究硝酸转运蛋白的工作机制奠定分子基础。

## 1 材料与方 法

### 1.1 植物材料及 *NRT* 序列来源

植物材料来自浙江华顶山的云锦杜鹃种子。消毒后播种于 1/2 read 琼脂培养基中培养两个月, 将小苗转接至混合 MMN(氮源为  $\text{Ca}(\text{NO}_3)_2$ ) 营养液的基质中(草炭: 沙子=2:1), 其中一半实生苗转接时接种菌株 19(杜鹃花类菌根真菌), 另一半不进行接种处理。接种 6 周采集植物根系, 进行转录组测序。然后对接种与未接种的根系转录组做差异分析, 得到接种苗根系中显著差异表达的基因。

*NRT*(硝酸根转运蛋白)基因序列即来源于云锦杜鹃根系接种苗中显著差异表达的基因, unigeneID: CL5793.Contig1\_All。

### 1.2 方法

预测: *NRT* 基因编码蛋白的理化性质采用 ProtParam 程序; 疏水性/亲水性采用 ProtScale 程序; 导肽采用 SignalP 程序; 跨膜结构域采用 TM-Pred 程序; 蛋白质二级结构采用 GOR4 程序; 蛋白质亚细胞定位采用 WOLF PSORT 程序; 蛋白质功能域搜索用 BLAST 程序; 蛋白质三维结构分析用 Phyre2 程序和 PyMOL 软件; 核苷酸序列和氨基酸序列搜索用在线分析工具 blastn; 进化树的构建采用软件 clustalX 和 MEGA4.1。

## 2 结果与分析

### 2.1 氨基酸序列的理化性质分析

利用在线分析程序 ProtParam (Gasteiger *et al.*, 2005) 对云锦杜鹃的硝酸根转运蛋白基因所对

应的氨基酸序列的理化性质进行分析。结果表明, 硝酸根转运蛋白基因所对应的氨基酸数目为 200 个, 分子量为 22 073.0, 理论等电点为 8.15, 氨基酸组成如图 1 所示, 负电荷氨基酸残基总数 (Asp + Glu) 为 16, 正电荷氨基酸残基总数 (Arg + Lys) 为 18。不稳定系数为 23.53, 此等级认为是稳定蛋白。脂肪系数为 77.45, 总平均输水性为 -0.068。

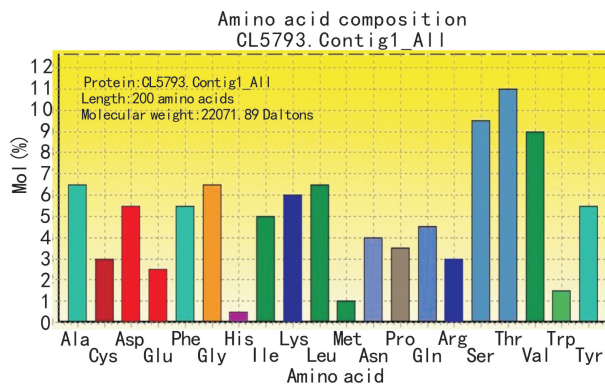


图 1 云锦杜鹃硝酸根转运蛋白的氨基酸组成  
Fig. 1 Amino acid composition of the deduced *NRT* peptide segment from *Rhododendron fortunei*

### 2.2 疏水性/亲水性的预测和分析

氨基酸的亲疏水性在一定程度上反映蛋白质的折叠情况, 此外, 蛋白质疏水区域可以作为评判潜在跨膜区的参考依据。因而对氨基酸的疏水性/亲水性的预测和分析具有重要的意义。利用在线分析程序 ProtScale (Gasteiger *et al.*, 2005) 的 Kyte & Doolittle 算法对云锦杜鹃硝酸根转运蛋白的进行亲水/疏水性分析(正值表示疏水性, 负值表示亲水性, 介于 +0.5 ~ -0.5 之间主要为两性氨基酸)。图 2 预测结果表明, 云锦杜鹃硝酸根转运蛋白有两个明显的疏水区, 分别为 5~21, 174~191; 4 个明显的亲水区, 分别为 58~71, 85~101, 112~114, 151~158。

### 2.3 导肽的预测和分析

利用在线分析工具 SignalP (Henrik *et al.*, 1997) 的神经网络算法对云锦杜鹃的硝酸根转运蛋白进行预测。表 1 结果表明, 云锦杜鹃的硝酸根转运蛋白序列包含信号肽, 长度为 22 个氨基酸, 为分泌途径蛋白。

### 2.4 跨膜结构的预测和分析

利用在线工具 TMPred (Hofmann *et al.*, 1993) 对云锦杜鹃硝酸根转运蛋白序列进行跨膜结构分析,

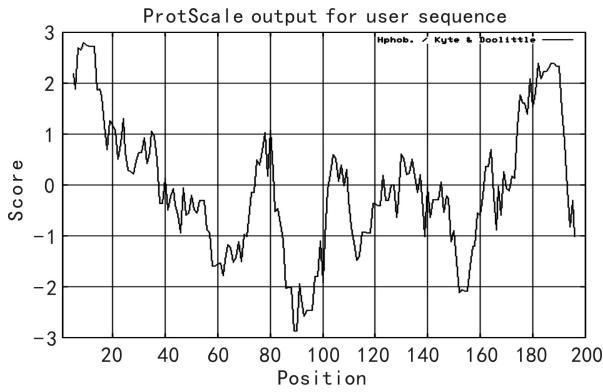


图 2 云锦杜鹃硝酸根转运蛋白的亲疏水性分析

Fig. 2 Predicted hydrophobicity/hydrophilicity for NRT in *Rhododendron fortunei*

表 1 云锦杜鹃硝酸根转运蛋白序列导肽的预测

Table 1 Predicted signal peptide of NRT protein in *Rhododendron fortunei*

Name	Len	cTP	mTP	SP	other	Loc	RC	TPlen
Sequence	200	0.007	0.065	0.968	0.063	S	1	22
Cutoff		0.000	0.000	0.000	0.000			

结果表明(图 3),该蛋白序列存在一种可靠的跨膜结构模型。该模型认为,N 端处在膜里面,第 1 个跨膜螺旋从第 3 位氨基酸开始,到第 21 位氨基酸结束,相对膜的取向是由里到外;第 2 个跨膜螺旋从第 175 位氨基酸开始,到第 194 位氨基酸结束,相对膜的取向是由外到里。

### 2.5 亚细胞定位预测

利用程序 wolf psort 分别对云锦杜鹃的 NRT 蛋白的亚细胞定位进行分析,结果表明如表 2 所示,NRT 蛋白在细胞质中的分值最高,推测该蛋白定位细胞质膜,行使氮转运功能。

表 2 云锦杜鹃的 NRT 蛋白的亚细胞定位进行分析

Table 2 Protein subcellular localization analysis of NRT from *Rhododendron fortunei*

Loc	plas	nucl	chlo	vacu	extr
Score	5.0	3.0	2.0	2.0	1.0

### 2.6 蛋白质二级结构的预测和分析

采用 GOR4(Garnier *et al.*, 1978)分析葡萄的类黄酮-O-葡萄糖基转移酶蛋白的二级结构,结果表明(图 4),该蛋白由 11%  $\alpha$ -螺旋(alpha helix)、29.50%延伸链(extended strand)和 59.50%无规则卷曲(random coil)所组成。

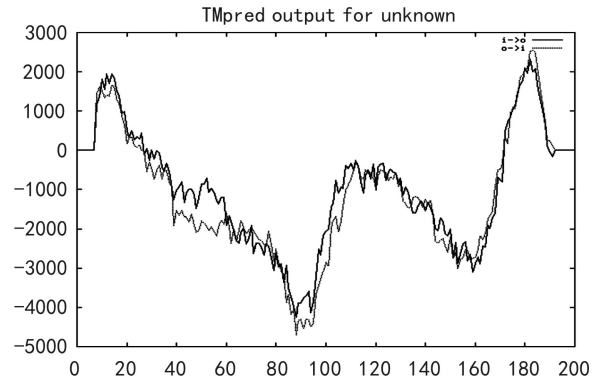


图 3 云锦杜鹃 NRT 蛋白的跨膜区域预测

Fig. 3 Predicted transmembrane segments for NRT in *Rhododendron fortunei*

### 2.7 蛋白质三维结构分析

利用在线分析工具 Phyre 2 (Kelley *et al.*, 2009),对云锦杜鹃 NRT 蛋白的氨基酸序列进行线串法预测(该方法是将待测序列“穿”入已知蛋白质的基本骨架内,通过计算各种折叠的概率,来预测蛋白质核心结构)蛋白质折叠,用 PyMOL 对建模结果进行处理,结果如图 5:A 所示。

### 2.8 植物中 NRT 核苷酸序列相似性比对与蛋白系统进化分析

通过 blastn 对 NRT 核苷酸序列在 NCBI 上进行同源性搜索,结果表明(图 6)云锦杜鹃的 NRT 核苷酸序列与杨毛果(*Populus trichocarpa*)、蓖麻(*Ricinus communis*)、野草莓(*Fragaria vesca* subsp. *Vesca*)、葡萄(*Vitis vinifera*)、黄瓜(*Cucumis sativus*)、大豆(*Glycine max*)、鹰嘴豆(*Cicer arietinum*)、百脉根(*Lotus japonicas*)和二柄蕙叶草(*Brachypodium distachyon*)等植物的 NRT 核苷酸序列有较高的同源性,分别为 69%、68%、67%、70%、67%、68%、67%、67%、76%。用 clustalX (Thompson *et al.*, 1997)和 MEGA4.0 (Tamura *et al.*, 2007)对云锦杜鹃、杨毛果 (XM\_002336362)、蓖麻(XM\_002511259)、野草莓(XM\_004300968)、葡萄(XM\_002279825)、黄瓜(XM\_004157728)、大豆(XM\_003549774)、鹰嘴豆(XM\_004515804)、百脉根(AB353299)和二柄蕙叶草(XM\_003575234)的 NRT 氨基酸进行多序列比对,构建系统进化树,采用 default 参数,自检举 1 000 次,结果如图 7 所示。10 个物种的 NRT 基因分为 3 个大类。其中杨毛果、蓖麻、野草莓、葡萄和云锦杜鹃聚为一大类,在

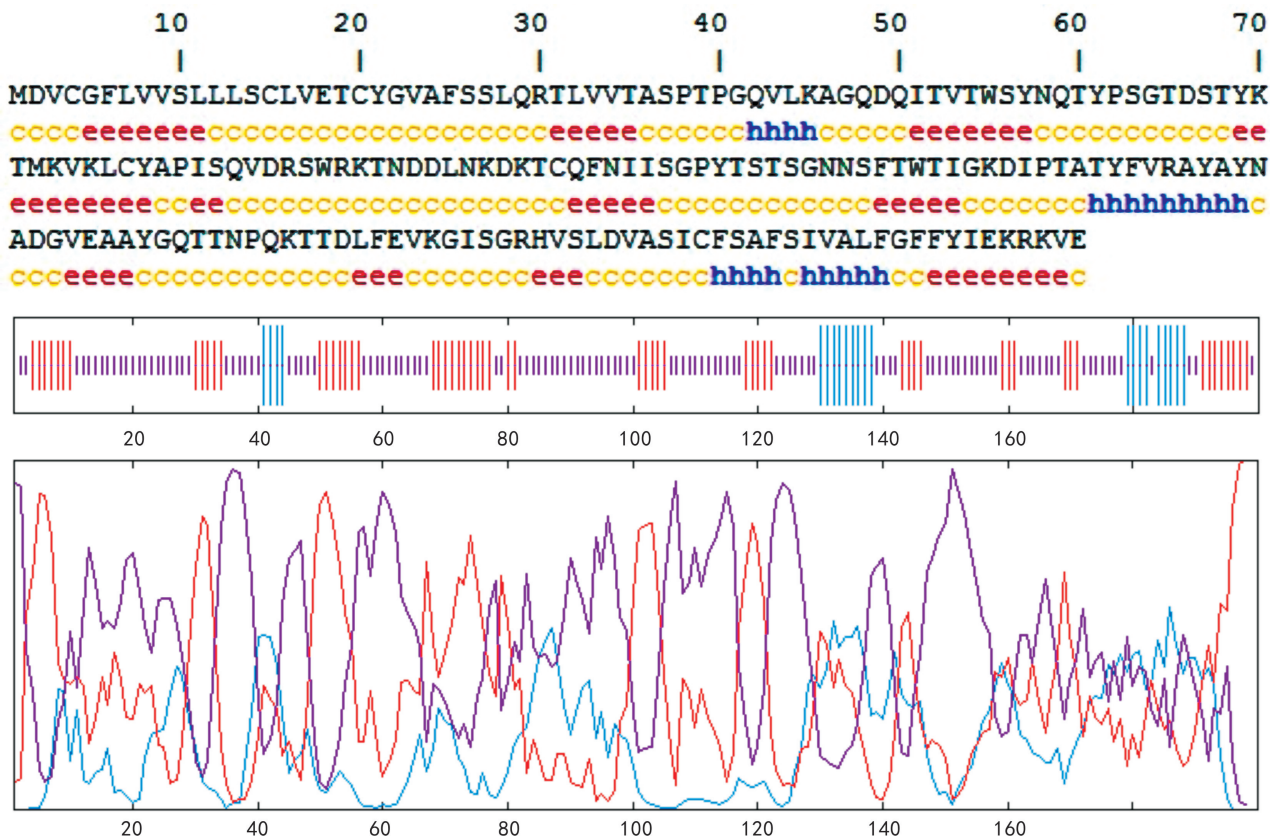


图 4 云锦杜鹃 *NRT* 蛋白的二级结构预测 h.  $\alpha$ -螺旋; e.  $\beta$ -折叠延伸链; L.  $\beta$ -转角; c. 无规则卷曲。

Fig. 4 Secondary structure prediction of *NRT* in *Rhododendron fortunei*

h. Alpha helix; e. Beta sheet Extended strand; L. Beta turn; c. Random coil.

这个大类中又分为两个亚类,云锦杜鹃和葡萄为一亚个类,杨毛果和蓖麻为一个亚类。黄瓜、大豆、鹰嘴豆以及百脉根聚为另一大类,其中大豆、鹰嘴豆以及百脉根为大类里面的一个亚类。而禾本科的二柄蕙叶草单独为一类。

### 3 结论与讨论

硝酸盐转运体蛋白在植物体内行使硝酸根吸收与转运的功能,因而被广泛的研究和关注。本研究用生物信息学方法对云锦杜鹃根转录组的硝酸根转运蛋白基因的核苷酸序列以及所对应的氨基酸序列进行分析。结果表明云锦杜鹃 *NRT* 蛋白为稳定蛋白,推测主要因为该蛋白在植物根系充当硝酸盐转运功能。跨膜结构分析表明,该蛋白存在两个跨膜结构,一个位于 N 端第 3 至第 21 个氨基酸内,另一个位于 175~194 个氨基酸内。跨膜结构分析结果与 Phyre 2 分析结果一致,亲疏水分析表明,两个跨膜区域的位置存在两个明显的疏水区;此外,预测的

信号肽也落在了 N 端的跨膜结构区域,四个分析软件的结果相互支持。Brian *et al.*(2000)曾对大麦的 *NRT* 蛋白做过预测分析,认为大麦的 *NRT* 蛋白存在 12 个跨膜结构。最近,我国著名结构生物学家施一公研究团队,解析细菌硝酸盐转运蛋白晶体结构,认为细菌硝酸盐转运蛋白存在 10 个跨膜结构共同组成硝酸盐运输路径(Hanchi *et al.*, 2013)。本次数据分析与大麦以及细菌的 *NRT* 蛋白的跨膜结构存在如此差异,推测可能由于 *NRT* 蛋白是一个家族蛋白,由于可变剪切的结果。二级结构分析表明,云锦杜鹃 *NRT* 蛋白质二级结构元件主要由  $\alpha$ -螺旋、延伸链和无规则卷曲所组成,这一分析结果跟用 Phyre 2 在二级结构分析结果比较一致。通过 blastn 对 *NRT* 核苷酸序列在 NCBI 上进行同源性搜索,结果表明云锦杜鹃的 *NRT* 核苷酸序列与杨毛果、蓖麻、野草莓、葡萄、黄瓜、大豆、鹰嘴豆、百脉根和二柄蕙叶草等植物的 *NRT* 核苷酸序列有很高的同源性。系统进化分析表明,云锦杜鹃的 *NRT* 蛋白在进化上与葡萄关系最近,与禾本科的单子叶

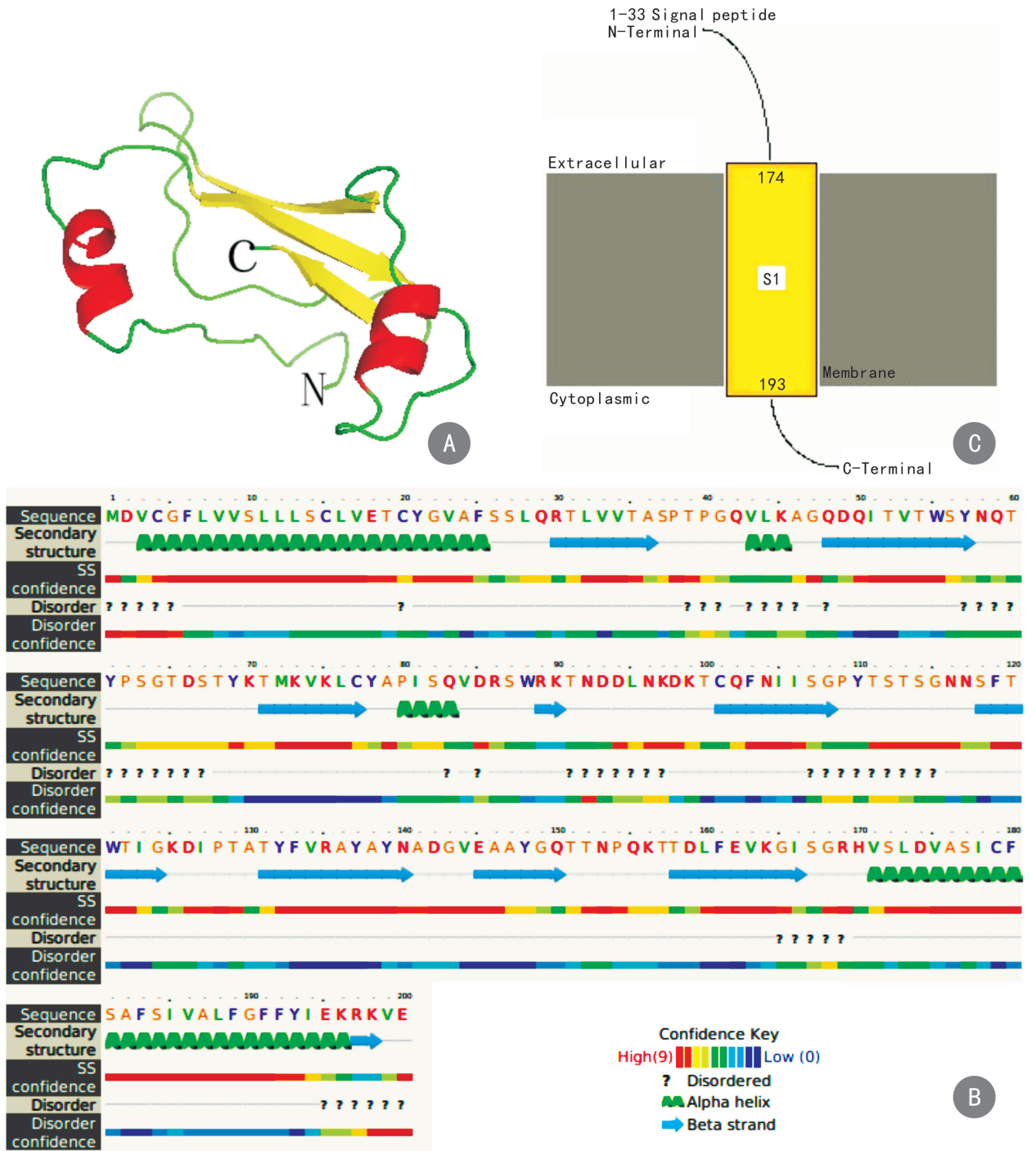


图 5 利用 Phyre 2 程序对推导的云锦杜鹃 *NRT* 蛋白结构预测 A. 3D 模型; B. 二级结构; C. 跨膜结构

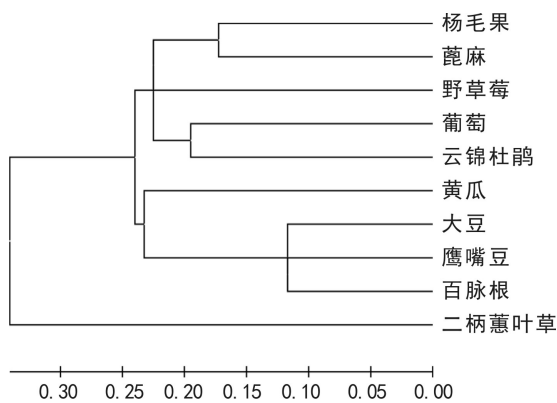
Fig. 5 Structure prediction of the deduced *NRT* proteins from *Rhododendron fortunei* using Phyre 2 A. 3D MODEL; B. Secondary structure; C. Transmembrane structure

植物关系最远,这一点与植物学传统分类相似。

本文通过对 *NRT* 基因的分析,得到了该基因蛋白结构特征信息,初步了解了该基因在云锦杜鹃根系中行驶的功能,这预示了该基因在根系中主要从事对硝酸盐的转运,对云锦杜鹃根系吸收硝酸

盐的重大意义。而硝酸盐是早作土壤中的主要无机氮源,可以推测该基因可能是植物对生命元素——氮素吸收的关键基因,对植物的生长发挥举足轻重的作用。为后续基因功能验证提供了有力的生物信息学基础。

Alignments							
Download GenBank Graphics Distance tree of results							
	Description	Max score	Total score	Query cover	E value	Ident	Accession
<input type="checkbox"/>	PREDICTED: <i>Vitis vinifera</i> high-affinity nitrate transporter 3.2-like (LOC100258771). mf	248	248	53%	9e-62	70%	<a href="#">XM_002279825.2</a>
<input type="checkbox"/>	<a href="#">Vitis vinifera clone SS0AEB5YC11</a>	248	248	53%	9e-62	70%	<a href="#">FQ397962.1</a>
<input type="checkbox"/>	<a href="#">Vitis vinifera clone SS0AEB25YF05</a>	248	248	53%	9e-62	70%	<a href="#">FQ387794.1</a>
<input type="checkbox"/>	<a href="#">Vitis vinifera clone SS0AEB5YH05</a>	248	248	53%	9e-62	70%	<a href="#">FQ397866.1</a>
<input type="checkbox"/>	<a href="#">Vitis vinifera clone SS0AEB5YO19</a>	248	248	53%	9e-62	70%	<a href="#">FQ397717.1</a>
<input type="checkbox"/>	<a href="#">Vitis vinifera clone SS0AEB7YB04</a>	248	248	53%	9e-62	70%	<a href="#">FQ397359.1</a>
<input type="checkbox"/>	<a href="#">Vitis vinifera clone SS0AEB30YL03</a>	248	248	53%	9e-62	70%	<a href="#">FQ385898.1</a>
<input type="checkbox"/>	<a href="#">Vitis vinifera clone SS0AEB12YH05</a>	248	248	53%	9e-62	70%	<a href="#">FQ378795.1</a>
<input type="checkbox"/>	<a href="#">Vitis vinifera clone SS0AEB31YM20 &gt;embIFQ387645.1 Vitis vinifera clone SS0AEB25</a>	248	248	53%	9e-62	70%	<a href="#">FQ385571.1</a>

图 6 利用 blastn 对云锦杜鹃 *NRT* 核苷酸序列比对结果Fig. 6 Results of alignment of *NRT* from *Rhododendron fortunei* using blastn图 7 植物中 *NRT* 蛋白系统进化分析Fig. 7 Phylogenetic analysis of *NRT* proteins in plants

## 参考文献:

- Bian CM(边才苗), Jin ZX(金则新). 2005. Studies on floral dynamic and breeding system of *Rhododendron fortunei* (云锦杜鹃的开花动态与繁育系统研究)[J]. *Guihaia* (广西植物), **25**(2): 169-173
- Brian G, Forde. 2000. Nitrate transporters in plants: structure, function and regulation[J]. *Biochem Biophys Acta*, **1** 465: 219-235
- D'Apuzzo E, Rogato A, Simon-Rosin U, et al. 2004. Characterization of three functional high-affinity ammonium transporters in *Lotus japonicus* with differential transcriptional regulation and spatial expression[J]. *Plant Physiol*, **134**: 1 763-1 774
- Fang YY(方云亿). 1989. The Flora of Zhejiang Vol.5(浙江植物志第 5 卷)[M]. Hangzhou(杭州): Zhejiang Science and Technology Press(浙江科学技术出版社): 9 210
- Garnier J, Osguthorpe DJ, Robson B. 1978. Analysis of the accuracy and implications of simple methods for predicting the secondary structure of globular proteins[J]. *J Mol Biol Mar*, **120**(1): 97-120
- Gasteiger E, Hoogland C, Gattiker A, et al. 2005. Protein Identifi-

- cation and Analysis Tools on the ExpASy Server//John M. Walker (ed). The Proteomics Protocols Handbook[M]. Humana Press: 571-607
- Henrik N, Jacob E, Søren, et al. 1997. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites[J]. *Prot Eng*, **10**: 1-6
- Hofmann K, Stoffel W. 1993. TMbase-A database of membrane spanning proteins segments[J]. *Biol Chem*, **374**: 166
- Jin ZX(金则新), Li JM(李钧敏), Gu QP(顾奇萍). 2006. Genetic diversity in the natural populations of rhododendron fortunei revealed by ISSR molecular markers(云锦杜鹃自然居群遗传多样性的 ISSR 分析)[J]. *Acta Hort Sin*(园艺学报), **33**(6): 1 263-1 267
- Kelley LA, Sternberg MJE. 2009. Protein structure prediction on the web: a case study using the Phyre server[J]. *Nature Protocols*, **4**: 363-371
- Lauter FR, Ninnemann O, Bucher M, et al. 1996. Preferential expression of an ammonium transporter and of two putative nitrate transporters in root hairs of tomato[J]. *Proc Nat Acad Sci (USA)*, **93**: 8 139-8 144
- Li JM(李钧敏), Jin ZX(金则新), Yang BF(杨蓓芬). 2004. The content and analysis of the components of flavonoids in *Rhododendron fortunei* (云锦杜鹃总黄酮含量及成分分析)[J]. *J Northwest For Univ*(西北林学院学报), **19**(1): 110-112
- Pearson JN, Finnemann J, Schjoerring JK. 2002. Regulation of the high-affinity ammonium transporter (BnAMT1,2) in the leaves of Brassica napus by nitrogen status[J]. *Plant Mol Biol*, **49**: 483-490
- Suenaga A, Moriya K, Sonoda Y, et al. 2003. Constitutive expression of a novel type ammonium transporter OsAMT2 in rice plants[J]. *Plant Cell Physiol*, **44**: 206-211
- Tamura K, Dudley J, Nei M, et al. 2007. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4. 0[J]. *Mol Biol Evol*, **24**: 1 596-1 599
- Yan HC, Huang WY, Yan CY, et al. 2013. Structure and mechanism of a nitrate transporter[J]. *Cell Reports* **3**, **28**: 716-723
- (下转第 147 页 Continue on page 147)