

## 小麦富含亮氨酸重复(LRR)蛋白基因 *TaLR11* 的克隆及其特征分析

孟凡荣<sup>1</sup>, 冉彩华<sup>1</sup>, 刘浩<sup>1</sup>, 陈雷<sup>2</sup>, 李金花<sup>2</sup>, 尹钧<sup>2</sup>, 李永春<sup>2,\*</sup>

河南农业大学<sup>1</sup> 生命科学学院, <sup>2</sup> 国家小麦工程技术研究中心, 郑州 450002

**摘要:** 在水分胁迫反应基因表达谱分析的基础上, 采用 RACR 技术, 从‘洛旱 2 号’小麦中克隆了一个编码富含亮氨酸蛋白的 cDNA 全长, 命名为 *TaLR11*。序列分析显示, *TaLR11* 的 cDNA 全长为 2 657 bp (GenBank 登录号为 GU593320), 其中 5'-非翻译区 47 bp, 3'-非翻译区 1 314 bp, 开放阅读框 1 296 bp, 可编码 431 个氨基酸。进一步分析显示, *TaLR11* 基因编码的蛋白具有典型的富含亮氨酸 N 末端保守域和富含亮氨酸的核酸酶抑制因子保守域, 该蛋白为弱酸性蛋白, 等电点为 5.69, 无明显的疏水/亲水区域。三维结构预测显示, *TaLR11* 蛋白以  $\alpha$ -螺旋、 $\beta$ -折叠和无规卷曲为骨架, 可形成弧状的螺线管结构以及一对侧臂凸起。RT-PCR 分析表明, 水分胁迫过程中, *TaLR11* 基因在根系中的表达量呈先升高后降低的趋势, 以胁迫处理 12 h 的表达量最高, 推测该基因在水分胁迫反应过程中发挥重要功能。

**关键词:** 小麦; *TaLR11*; 基因克隆; 序列特征

## Cloning and Characterization of a Leucine-Rich Repeat (LRR) Protein Gene *TaLR11* in Wheat (*Triticum aestivum* L.)

MENG Fan-Rong<sup>1</sup>, RAN Cai-Hua<sup>1</sup>, LIU Hao<sup>1</sup>, CHEN Lei<sup>2</sup>, LI Jin-Hua<sup>2</sup>, YIN Jun<sup>2</sup>, LI Yong-Chun<sup>2,\*</sup>

<sup>1</sup>College of Life Sciences, <sup>2</sup>National Engineering Research Centre for Wheat, Henan Agricultural University, Zhengzhou 450002, China

**Abstract:** Based on our previous study on the water stress responsive transcriptomes in wheat, a full-length cDNA, encoding a leucine-rich repeat (LRR) protein, was cloned via rapid amplification of cDNA ends (RACE) technology from the wheat (*Triticum aestivum*) cultivar ‘Luohan 2’, which was termed as *TaLR11* with the GenBank accession No. of GU593320. Sequence analysis showed that the full-length cDNA of *TaLR11* was 2 657 bp, which including 47 bp 5' UTR, 1 314 bp 3' UTR and 1 296 bp open reading frame (ORF) encoding 431 amino acid residues. The further analysis indicated that a typical LRR N-terminal domain and a LRR ribonuclease inhibitor (LRR\_RI) domain was included in the deduced amino acid sequence of *TaLR11*. The putative *TaLR11* protein was acidulous with a pI of 5.69 and obvious hydrophobic or hydrophilic domains were not found in it. The 3D structure prediction showed that the *TaLR11* was constructed by  $\alpha$ -helixes,  $\beta$ -strands and random loops, and could wrap around in a curved, right-handed solenoid with a pair of extended side arms. The expression pattern of *TaLR11* in roots under water stress was analyzed via RT-PCR and the results showed that the gene was gradually induced with the water-stress accumulation, while its expression level decreased after 12 h after water-stress treatment, which indicated that the *TaLR11* played important roles during the water-stress responding in wheat.

**Key words:** wheat; *TaLR11*; gene cloning; sequential characterization

干旱是影响作物持续高产稳产的主要逆境因子之一, 关于作物水分胁迫反应分子调控机理的研究是现代作物分子生物学领域的重要课题(Shinozaki 和 Yamaguchi-Shinozaki 2007; 王转等 2004)。研究表明, 植物对水分胁迫的感知和调节机制非常复杂, 涉及到许多胁迫反应相关基因的表达与调控(杨洪强等 2001; Tardif 等 2007; Xiang 等 2007)。富

含亮氨酸重复(leucine-rich repeat, LRR)的蛋白是一类包含由 20~29 个氨基酸组成的重复序列的蛋白,

收稿 2010-03-01 修定 2010-05-31

资助 国家转基因生物新品种培育科技重大专项(2009ZX08002-011B)和河南省教育厅自然科学研究计划(2009A210012)。

\* 通讯作者(E-mail: yongchunli71@yahoo.com.cn; Tel: 0371-63558215)。

可形成由 $\alpha$ -螺旋和 $\beta$ -折叠组成的规则螺线管空间结构。自1985年首次报道人类LRR蛋白(Takahashi等1985)以来,已经在动物、真菌以及植物中相继发现了一批具有LRR的蛋白。研究发现,LRR的保守域为包含11个氨基酸残基的LxxLxLxx(N/C)xL结构(其中,“x”可以为任意氨基酸)(Kobe和Deisenhofer 1994),该结构在蛋白互作及植物逆境胁迫信号转导过程中发挥着重要的调节功能(Baker等1997; Vadassery和Oelmuller 2009)。我们在前期分析小麦水分胁迫反应基因表达谱时,发现了一个受水分胁迫诱导的基因(TaAffx.99551.1)被注释为编码LRR蛋白的基因(李永春等2008)。为了进一步分析该基因在小麦中的存在情况,本文以探针组的代表序列为基础,通过RACE技术从小麦中克隆了该基因的cDNA全长(命名为*TaLR11*),并系统分析了其编码蛋白的结构特性,为进一步探讨小麦*TaLR11*基因在逆境胁迫反应过程中的生物学功能奠定了基础。

## 材料与amp;方法

将‘洛旱2号’小麦(*Triticum aestivum* L.)种子在培养皿中水培至两叶一心期,选取生长发育一致的幼苗,用20%的PEG6000进行水分胁迫处理。剪取胁迫处理0、0.5、1.0、2.0、6.0、12.0、24.0和48.0 h的根系,迅速用液氮冷冻后保存于-80℃备用。

总RNA提取采用Trizol RNA提取试剂盒, RNA纯化参照Meng等(2007)的方法进行。依据Affymetrix小麦基因芯片探针组TaAffx.99551.1的代表序列(GB: CA653329)设计特异性引物GP1 (5' ATGATAACAGCACGGACAT 3')和GP2 (5' TGTGGAGACGGTTGAGAGG 3'),通过RT-PCR技术从‘洛旱2号’小麦中克隆到该基因的部分cDNA片段,并进行测序分析。根据克隆片段的序列信息再设计嵌套引物GP3 (5' GCCATTTGCGGAC-AAGTTCA 3')、GP4 (5' GCTTGCCTTGGGTCG-TTGATGA 3')、GP5 (5' CCATCATCAACGAC-CCAAG 3')和GP6(5' ATCAACGACCCAAGGCAAG

3')用于该基因5'端和3'端cDNA序列的克隆。5'和3' RACE试剂盒采用TaKaRa公司的产品,具体操作步骤根据试剂盒说明进行。引物组合情况为:GP3和GP5分别与5'和3' RACE试剂盒提供的Outer Primer组合进行扩增,而GP4和GP6则与相应的Inner Primer组合。扩增DNA片段采用北京天为时代DNA凝胶回收试剂盒回收后连接于pGEM-T Easy载体(Promega公司)中,并转化大肠杆菌DH5 $\alpha$ ,然后送北京博尚生物技术有限公司进行测序分析;cDNA序列的初步分析、氨基酸翻译使用DNAMAN软件,蛋白质一级结构以及三级结构分析通过ExPaSy (<http://www.expasy.org>)分析(Arnold等2006),二级结构的预测采用PredictProtein (<http://www.predictprotein.org/>)分析(Rost等2004)。

提取水分胁迫处理不同时间后的小麦根系RNA,反转录获得的cDNA作为半定量RT-PCR分析的模板。用于半定量分析的特异引物为GP1和GP2。为防止扩增达到平台期,分别进行28、30和35个循环的扩增,最后选用30个循环进行表达特性分析。内标基因*Actin*的扩增引物序列为:Actin-1, 5' CAGCAACTGGGATGATATGG 3'; Actin-2, 5' ATTTTCGCTTTCAGCAGTGGT 3'。

## 结果与amp;讨论

### 1 小麦*TaLR11*全长cDNA的克隆和序列分析

采用Affymetrix小麦芯片中探针组TaAffx.99551.1对应代表序列的特异引物GP1和GP2,通过RT-PCR技术从‘洛旱2号’中克隆到了大小为285 bp的DNA片段,测序分析显示克隆序列和代表性序列(GB: CA653329)高度同源。用已克隆片段设计的特异引物GP3和GP5分别与5'和3' RACE试剂盒提供的Outer Primer组合进行5'和3' RACE的Outer PCR扩增,进一步以5'和3' RACE的Outer PCR产物为模板,分别进行Inner PCR扩增,结果发现5' RACE获得了约300 bp的特异片段,3' RACE获得了约2 300、800和700 bp的3个片段(图1)。将上述片段回收、克隆并测序后发现,5' RACE获得的300 bp片段和3' RACE获得的2 300 bp片段

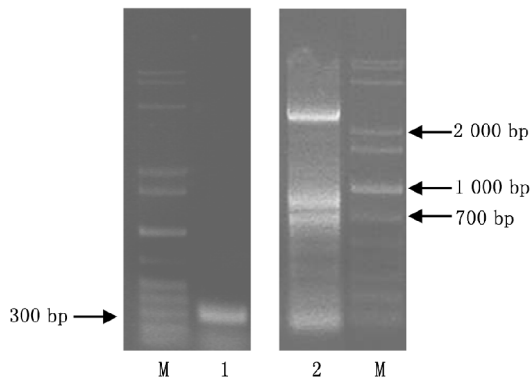


图1 *TaLRII* 基因 5' RACE 和 3' RACE 的巢式 PCR 扩增  
Fig.1 Inner PCR amplification for 5' RACE  
and 3' RACE of *TaLRII* gene

M: DNA 分子量标准; 1: 5' RACE inner PCR 产物; 2: 3' RACE inner PCR 产物。

为 LRR 基因的 cDNA 片段, 而 3' RACE 获得的 800 和 700 bp 两个片段均为非特异扩增产物, 与该基因无关。通过 DNAMAN 软件对获得的序列进行拼接组装后获得了小麦 LRR 基因的 cDNA 全长。序列分析显示, 该基因的 cDNA 全长为 2 657 bp, 其中 5'

UTR (非翻译区) 47 bp、3' UTR 1 314 bp、ORF (开放阅读框) 1 296 bp。通过 NCBI 进行的 Blastx 分析表明, 该基因编码蛋白与 LRR 蛋白的同源性较高, 其中与小麦 (GenBank 登录号为 ACY30448)、水稻 (*Oryza sativa*) (GenBank 登录号为 ABA91654 和 ABA91652) 的氨基酸序列相似性分别为 64%、63% 和 62%, 表明克隆基因为编码富含亮氨酸蛋白的基因, 命名为 *TaLRII* (GenBank 登录号为 GU593320)。

## 2 小麦 *TaLRII* 编码蛋白质的一级结构特性分析

分析表明, *TaLRII* 基因的蛋白包含 431 个氨基酸, 其中含量较高的氨基酸有 Leu (19.3%)、Ser (11.1%) 和 Gly (10.2%); *TaLRII* 蛋白的分子式为  $C_{2012}H_{3276}N_{560}O_{623}S_{16}$ , 分子量为 45.79 kDa, 理论等电点为 5.69, 负电荷氨基酸 (Asp+Glu) 总数为 31, 正电荷氨基酸 (Arg+Lys) 总数为 24; *TaLRII* 蛋白的半衰期理论值为 30 h, 不稳定参数为 33.85, 属于稳定蛋白。进一步分析发现, *TaLRII* 蛋白还包含多个翻译后修饰位点, 包括 10 个 N 糖基化位点 (第 47、114、

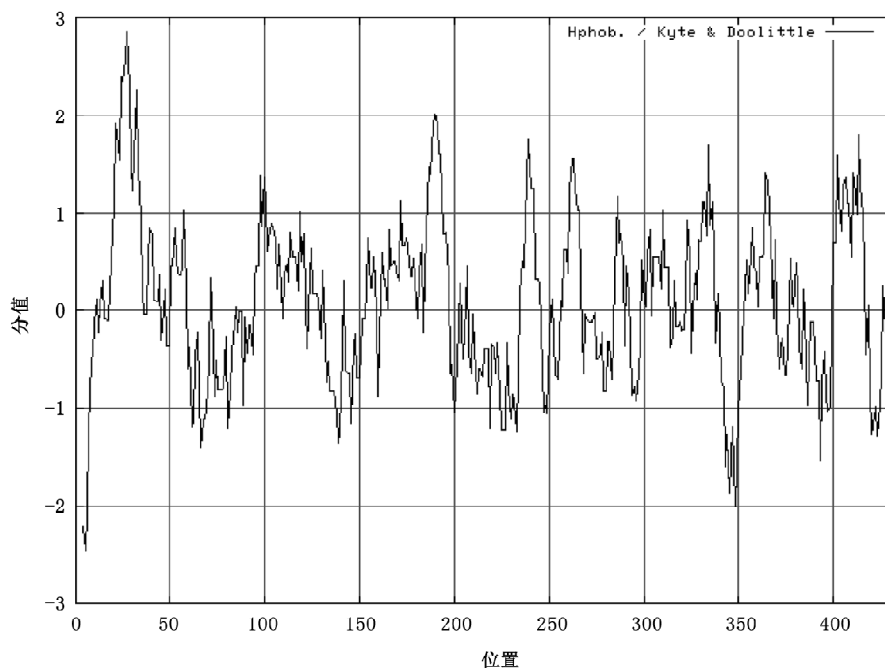


图2 *TaLRII* 的疏水性分析

Fig.2 Hydrophobicity profile of *TaLRII*

横坐标为编码蛋白的氨基酸顺序; 纵坐标为亲、疏水值, 0 值以上的部分表示疏水区域, 0 值以下的部分表示亲水区域。

122、161、198、209、256、305、359、408位氨基酸)、4个蛋白激酶C磷酸化位点(第18、200、292、344位氨基酸)、4个酪蛋白激酶II磷酸化位点(第53、227、344、413位氨基酸)和1个酪氨酸激酶磷酸化位点(第247位氨基酸)。

采用 ProtScale 进行的疏水性 / 亲水性分析显示: TaLRII 蛋白的亲水性最高分值为-2.34, 疏水性最高分值为 2.86, 且疏水性 / 亲水性氨基酸分布比较均匀, 没有明显的疏水或亲水区域(图 2)。

### 3 小麦 TaLRII 编码蛋白质的高级结构特性分析

通过 PredictProtein 对小麦 TaLRII 基因编码蛋白的二级结构预测显示, 该蛋白可形成由  $\alpha$ -螺旋

(占整个氨基酸序列的 19.03%)、 $\beta$ -折叠(占 13.46%)和无规卷曲(占 67.52%)组成的特定结构(图 3)。

NCBI 的保守域分析显示(图 4), TaLRII 蛋白包含有典型的富含亮氨酸 N 末端保守域(从第 47~87 位氨基酸)和 LRR 核酸酶抑制因子(LRR ribonuclease inhibitor, LRR\_RI)保守域(从第 140~349 位氨基酸)。LRR\_RI 是一类能对细胞内核酸酶起抑制作用的功能蛋白, 可以通过延缓细胞内 RNA 的降解速度直接参与真核基因的表达调控。另外, LRR\_RI 保守域也是 GTPase 激活蛋白(GTPase-activating proteins)的重要特征(Marchler-Bauer 等 2009), 而 GTPase 激活蛋白是调控 G 蛋白活性的重

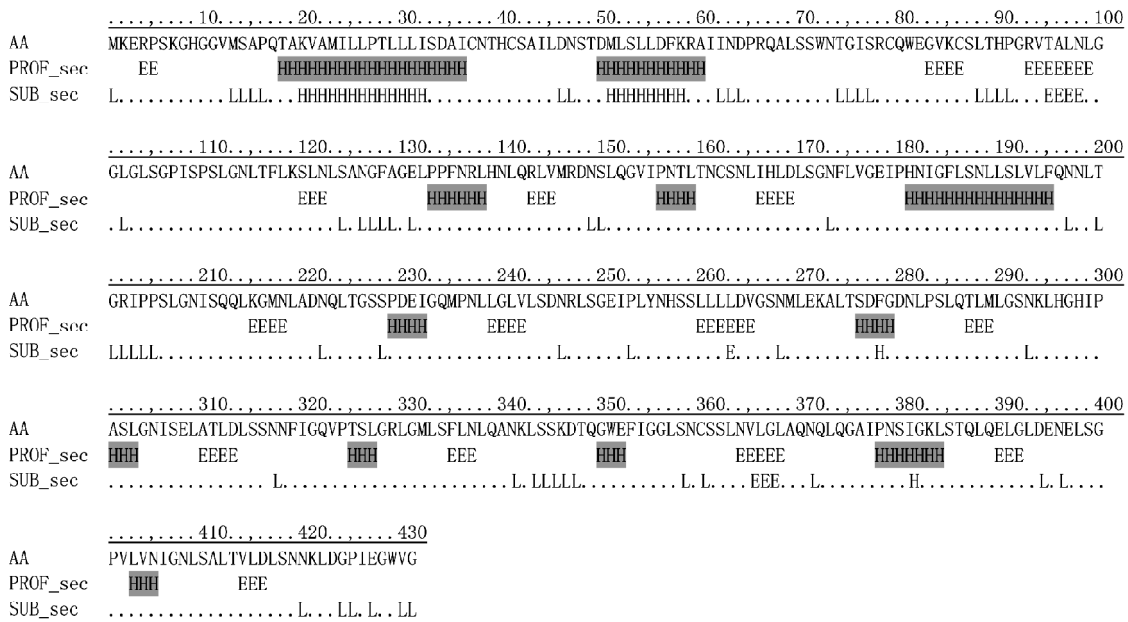


图 3 TaLRII 蛋白的二级结构预测

Fig.3 The predicted secondary structure of TaLRII

H:  $\alpha$ -螺旋; E:  $\beta$ -折叠; L: 无规卷曲。

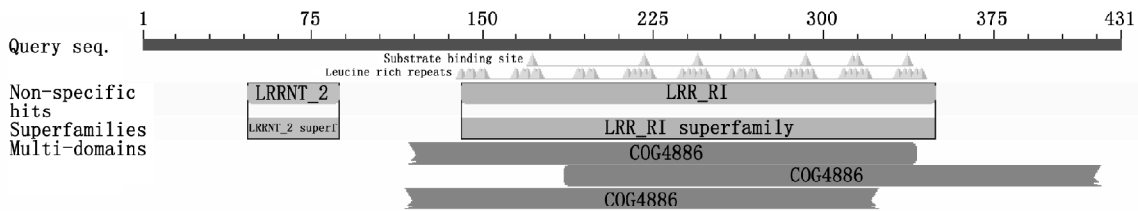


图 4 TaLRII 的保守域分析

Fig.4 Conserved domains of TaLRII

要因子, 通过激活G蛋白自身的GTPase活性, 使得G蛋白从激活态(GTP-结合形式)转变为失活态(GDP-结合形式), 在逆境胁迫信号转导过程中发挥着重要的调控功能。

蛋白三级结构预测结果显示, TaLR11可以形成

LRR蛋白的典型螺线管空间结构(图5), 而且由 $\alpha$ -螺旋和 $\beta$ -折叠组成的螺线管呈弧状弯曲, 在螺线管弧状弯曲的侧面及背面分别由 $\beta$ -折叠和 $\alpha$ -螺旋构成了2个柄状长臂, 双臂夹角约为 $120^\circ$ (图5)。该结构是否会赋予小麦TaLR11新的生物学功能还有

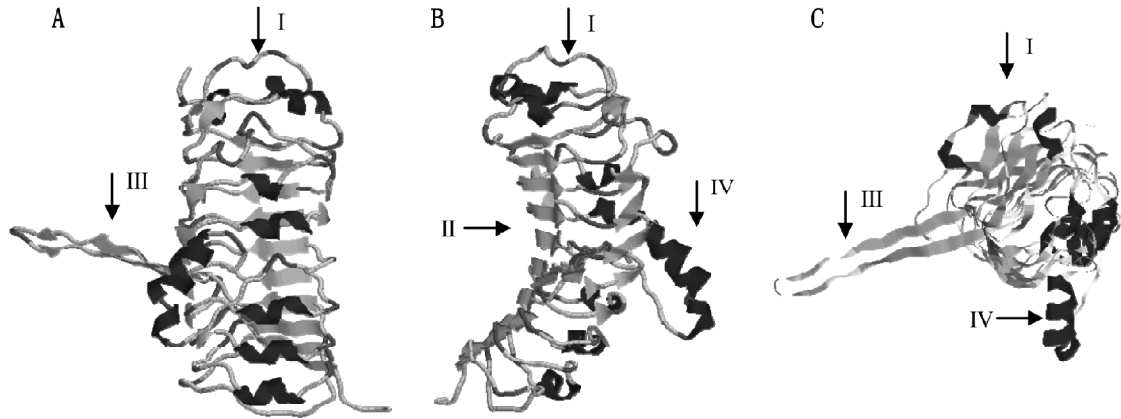


图5 小麦 TaLR11 的三维结构预测图

Fig.5 The predicted 3D structure of TaLR11

A: 正视图; B: 侧视图; C: 俯视图。I: 典型螺线管空间结构; II: 螺线管呈弧状弯曲; III:  $\beta$ -折叠; IV:  $\alpha$ -螺旋。

待于进一步研究。

#### 4 水分胁迫过程中 TaLR11 基因的表达分析

半定量 RT-PCR 分析结果表明, 在‘洛旱2号’幼苗的根系中, TaLR11基因在水分胁迫过程中随着胁迫时间的增加其表达量逐渐增强, 胁迫处理12 h的表达量达到最强, 但胁迫24 h的表达量又开始下降(图6)。可见, TaLR11 与小麦水分胁迫反应密切相关。关于该基因在水分胁迫信号转导及基因表达调控过程中的具体生物学功能仍有待于进一步深入研究。

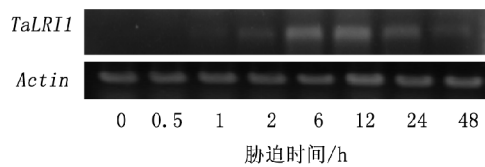


图6 水分胁迫下小麦 TaLR11 的表达

Fig.6 Expression of TaLR11 under water-stress in wheat

#### 参考文献

- 李永春, 孟凡荣, 王潇, 陈雷, 任江萍, 牛洪斌, 李磊, 尹钧(2008). 水分胁迫条件下‘洛旱2号’小麦根系的基因表达谱. 作物学报, 34 (12): 2126~2133
- 王转, 臧庆伟, 郭志爱, 景蕊莲(2004). 小麦幼苗期水分胁迫所诱导基因表达谱的初步分析. 遗传学报, 31 (8): 842~849
- 杨洪强, 贾文锁, 张大鹏(2001). 植物水分胁迫信号识别与转导. 植物生理学通讯, 37 (2): 149~154
- Arnold K, Bordoli L, Kopp J, Schwede T (2006). The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics, 22 (2): 195~201
- Baker B, Zambryski P, Staskawicz B, Dinesh-Kumar SP (1997). Signaling in plant-microbe interactions. Science, 276: 726~733
- Kobe B, Deisenhofer J (1994). The leucine-rich repeat: a versatile binding motif. Trends Biochem Sci, 19 (10): 415~421
- Marchler-Bauer A, Anderson JB, Chitsaz F, Derbyshire MK, DeWeese-Scott C, Fong JH, Geer LY, Geer RC, Gonzales NR, Gwadz M et al (2009). CDD: specific functional annotation with the Conserved Domain Database. Nucleic Acids Res, 37 (Database issue): D205~D210

- Meng F, Ni Z, Wu L, Sun Q (2007). Identification of 17 differentially expressed cDNAs between wheat reciprocal cross-fertilized kernels and their parents. *South Afr J Bot*, 73 (4): 522~529
- Rost B, Yachdav G, Liu J (2004). The PredictProtein server. *Nucleic Acids Res*, 32 (Web Server issue): W321~W326
- Shinozaki K, Yamaguchi-Shinozaki K (2007). Gene networks involved in drought stress response and tolerance. *J Exp Bot*, 58 (2): 221~227
- Takahashi N, Takahashi Y, Putnam FW (1985). Periodicity of leucine and tandem repetition of a 24-amino acid segment in the primary structure of leucine-rich  $\alpha_2$ -glycoprotein of human serum. *Proc Natl Acad Sci USA*, 82 (7): 1906~1910
- Tardif G, Kane NA, Adam H, Labrie L, Major G, Gulick P, Sarhan F, Laliberte JF (2007). Interaction network of proteins associated with abiotic stress response and development in wheat. *Plant Mol Biol*, 63 (5): 703~718
- Vadassery J, Oelmuller R (2009). Calcium signaling in pathogenic and beneficial plant microbe interactions: what can we learn from the interaction between *Piriformospora indica* and *Arabidopsis thaliana*. *Plant Signal Behav*, 4 (11): 1024~1027
- Xiang Y, Huang Y, Xiong L (2007). Characterization of stress-responsive *CIPK* genes in rice for stress tolerance improvement. *Plant Physiol*, 144 (3): 1416~1428