

# 水分胁迫下小麦 *TaWSI18* 基因的电子克隆及生物信息学分析

樊磊,陈娟,张艳娥,张林生,朱维宁,张大鹏

(西北农林科技大学生命科学院 旱区作物逆境生物学国家重点实验室,陕西 杨凌 712100)

**摘要:** 利用电子克隆和 RT-PCR 技术,从干旱胁迫的小麦叶片中分离出 4 条新的具有高度同源性的 *WSI18* 基因,即 *TaWSI18-1*、*TaWSI18-2*、*TaWSI18-3* 和 *TaWSI18-4* (GeneBank 登陆号分别为: KP226849、KP226850、KP226851 和 KP226852);其开放阅读框均为 678 bp,编码 225 个氨基酸。生物信息学分析表明,该类蛋白具有亲水性,存在多个磷酸化位点,无信号肽结构域及剪切位点,定位于细胞质中,无跨膜结构域;二级结构分析表明,该蛋白  $\alpha$  螺旋和无规则卷曲的比例达到 90% 以上;同源性比较及聚类分析表明,小麦 *TaWSI18* 蛋白同无芒雀麦 (*Bromus inermis*, BAN15016) *WSI18* 蛋白和二穗短柄草 (*Brachypodium distachyon*, XP\_003569641) *LEA3* 蛋白具有高度的同源性,相似性分别为 94% 和 70%。小麦 *TaWSI18* 基因编码蛋白的氨基酸序列的 N 端存在与 *LEA3* 蛋白的 N 端区域高度保守的序列。表明小麦 *TaWSI18* 蛋白的性质与 *LEA3* 蛋白的相似,说明了小麦的 *TaWSI18* 蛋白可能与 *LEA3* 蛋白存在相同或相似的功能,为进一步研究其在小麦水分胁迫下的功能奠定基础。

**关键词:** 小麦;水分胁迫;*TaWSI18* 基因;电子克隆;生物信息学

**中图分类号:** S512.1 **文献标志码:** A

## Identification and characterization of water stress-induced *TaWSI18* gene in silico cloning from *Triticum aestivum* and its bioinformatic analysis

FAN Lei, CHEN Juan, ZHANG Yan-e, ZHANG Lin-sheng, ZHU Wei-ning, ZHANG Da-peng

(State Key Laboratory of Crop Stress Biology in Arid Area, College of Life Sciences,

Northwest A & F University, Yangling, Shaanxi 712100, China)

**Abstract:** *WSI18*, a water stress-induced gene, plays an important role in plants under water deficit. Silico cloning and RT-PCR were taken to isolate four new highly homologous *WSI18* genes: *TaWSI18-1*, *TaWSI18-2*, *TaWSI18-3* and *TaWSI18-4* (GenBank accession number: KP226849, KP226850, KP226851 and KP226852) in wheat, ORFs of which are all 678 bp encoding 225 amino acids. Bioinformatic analysis showed that they are highly hydrophilic and have multiple phosphorylation sites. They are located to cytoplasm and have no transmembrane domain signal peptides, as well as no splicing sites. The proportion of  $\alpha$  helix and random coil are over 90% by a secondary structure prediction. Homology comparison and phylogenetic analysis showed that *TaWSI18* proteins are highly homologous to *WSI18* of *Bromus inermis* and *LEA3* *Brachypodium distachyon*, sharing 94% and 70% similarities respectively. In addition, *TaWSI18* proteins have conserved amino acid sequences common to *LEA3* protein in the N-terminal region. They also have a similar character to *LEA3* protein by bioinformatic analysis that *TaWSI18* proteins may have the same or similar function to *LEA3* protein, which lies a foundation for the future study on its function under water-stress.

**Keywords:** *Triticum aestivum*; water stress; *TaWSI18*; silico cloning; bioinformatics

小麦是我国主要的粮食作物,非生物胁迫会导致细胞缺水,影响植物的生长发育,诱导多种基因表达<sup>[1]</sup>,产生一系列重要的功能蛋白和调控蛋白,减轻

水分胁迫对作物的损害。调控蛋白如转录因子、蛋白激酶等参与信号转导和基因表达调控,调节细胞的新陈代谢、转录并对各种刺激产生应答反应<sup>[2-3]</sup>。

收稿日期:2015-05-22

基金项目:国家自然科学基金资助项目(31071349);高等学校博士学科点专项基金(20120204110033)

作者简介:樊磊(1989—),女,山西运城人,硕士研究生,研究方向为小麦抗旱相关基因的研究。E-mail: echo18735439872@163.com。

通讯作者:张林生(1958—),教授,博士生导师,主要从事小麦抗旱方面的研究。E-mail: linszhang@nwsuaf.edu.cn。

胚胎发育晚期丰富蛋白(Late embryogenesis abundant protein, LEA)自 1981 年首次在棉花中发现以来<sup>[4-5]</sup>,人们在许多植物中都发现了此类蛋白的存在,尤其是在干旱等逆境胁迫下,发现 LEA 蛋白表达量升高<sup>[6-8]</sup>。在水分胁迫条件下,LEA 蛋白对酶(如乳酸脱氢酶,苹果酸脱氢酶等)<sup>[9-10]</sup>和膜结构具有保护作用,增强植物的抗性。张瑞越<sup>[11]</sup>等从小麦中克隆的水分胁迫诱导基因 *W89* 受干旱、低温和 ABA 的诱导表达。从水稻中分离到水分胁迫诱导基因 *wsi18*, *wsi76*, *wsi724*, 其中 *wsi18* 在水分胁迫时表达量增加<sup>[12-13]</sup>。植物的水分胁迫应答涉及复杂的调控网络,胁迫机制尚有待进一步研究。本研究运用电子克隆的方法对小麦 EST 序列数据库进行比对分析,拼接克隆得到 *WSII8* 的 cDNA 序列,利用生物信息学分析其结构与功能,并进行同源性比对,为深入研究小麦水分胁迫诱导基因的功能具有重要意义。

## 1 材料与方法

### 1.1 材料

供试小麦品种为陕合 6 号,由西北农林科技大学生命科学学院中心实验室提供。

试剂:大肠杆菌 JM109,由西北农林科技大学生命科学学院中心实验室保存;PrimeSTAR<sup>®</sup> HS DNA, pMD18-T 载体, DNA Marker 均购自 TaKaRa 公司;凝胶回收试剂盒及小量质粒提取试剂盒,均购自天根公司;RNA 提取试剂 Trizol 购自 Invitrogen 公司。序列测定由华大基因完成。

### 1.2 引物的设计合成

以电子克隆序列为基础,采用 OLIGO7 软件设计 RT-PCR 扩增的 1 对引物,上游引物 F: 5' 'ACTTTTCGAGCCCTATCCAGT 3', 下游引物 R: 5' 'AAGATACATCAGGGTCGCCTT 3'。由华大基因合成。

### 1.3 电子克隆的技术路线和方法

目前电子克隆最常用的方法是利用 EST 信息进行同源性检索、聚类、序列拼装等获得部分乃至全长 cDNA 序列。本研究参考张德礼等的方法<sup>[14]</sup>,略有改动。在 NCBI 网站的 GenBank 数据库中检索小麦或与小麦亲缘关系较近的 cDNA 序列,以该序列作为种子序列,对小麦 EST 数据库进行 Blast 检索,提取高同源性 *WSII8* 基因的 EST 序列,利用 CEx-press 软件进行拼接。再以拼接后的 EST 作为新的查询探针,重复以上步骤,继续搜索小麦 EST 数据

库,直到没有新的 *WSII8* 基因的 EST 序列可供拼接为止。对拼接后的 EST 进行 ORF 预测、Blast 及完整性分析。

### 1.4 RNA 提取及 RT-PCR 扩增

使用 Trizol 试剂盒提取干旱处理 24 h 的小麦幼苗的总 RNA,置于 -80℃ 冰箱中备用。按照 TaKaRa 公司 PrimeScript<sup>®</sup> RT reagent Kit Perfect Real Time 的操作,将 2  $\mu$ l 总 RNA 反转录成 cDNA,作为 PCR 反应的模板。利用 PrimeSTAR<sup>®</sup> HS DNA 高保真酶进行 PCR 反应,体系为 20  $\mu$ l,程序为 98℃ 3 min, 98℃ 10 s, 56℃ 20 s, 72℃ 40 s, 35 个循环;72℃ 延伸 10 min。最后 1.2% 琼脂糖凝胶电泳检测 RT-PCR 扩增结果。

### 1.5 RT-PCR 产物胶回收、基因克隆及测序分析

按照胶回收和基因克隆试剂盒说明,进行 RT-PCR 产物胶回收和基因克隆,获得的目的片段,通过加“A”反应与 pMD18-T 载体连接,并转化大肠杆菌 JM109,在含有氨苄青霉素(Amp)、异丙基- $\beta$ -D 硫代半乳糖苷(IPTG)和 5-溴-4-氯-3-吡啶- $\beta$ -D-半乳糖苷(X-Gal)的 LB 平板上筛选阳性克隆,对菌落进行 PCR 鉴定。将鉴定的阳性菌在 37℃ 220 r/min 培养过夜,收菌提取质粒,PCR 鉴定,然后进行基因测序,将测序结果与电子克隆序列进行比对分析。

### 1.6 小麦 *TaWSII8* 基因的生物信息学分析

利用 ORF Finder 软件对基因 ORF 进行预测;使用 ExPASy 在线程序 ProtParam 对蛋白序列的分子量、等电点、氨基酸组成等基本性质进行预测;ProtScale 进行亲/疏水性预测;NetPhos 2.0 在线软件对磷酸化位点进行分析;应用 SignalP 3.0 Server 工具进行信号肽预测;TargetP 和 PSORT 在线分析亚细胞定位;TMHMM 工具预测跨膜结构域;并利用神经网络 HNN 进行二级结构分析。使用 ClustalX 软件进行氨基酸序列比对,并结合 MEGA 5.0 软件,基于 ML 方法及 500 次 bootstrap 检测构建系统进化树。

## 2 结果与分析

### 2.1 小麦 *TaWSII8* 基因的电子克隆

以无芒雀麦 *WSII8* 基因(登录号 AB813835)为探针,从 NCBI 的小麦 EST 库中,筛选到两条序列,其登录号分别为 BQ166490 和 GH727079。利用 CEx-press 软件拼接,再以该 EST 为探针,在小麦 EST 数据库中 Blast 检索,重复以上过程,直至没有更多的重叠 EST 检出。通过 ORF Finder 软件预测,被筛选

而拼接的序列具有完整的 ORF(90 – 767 bp), 编码 225 个氨基酸。以起始密码子的碱基为第 1 位(标记为 +1), 在 +4 位核苷酸为 G, –3 和 –6 位核苷酸为 G, 符合 kozak 原则(ATG 附近的核苷酸序列以

ANNATGG 和 GNNATGG 利用率较高)。在 3'UTR 区有典型的 poly A, 表明拼接获得了 1 条完整的 cDNA 序列, 命名为 *TaWSII8 – EST*(图 1)。

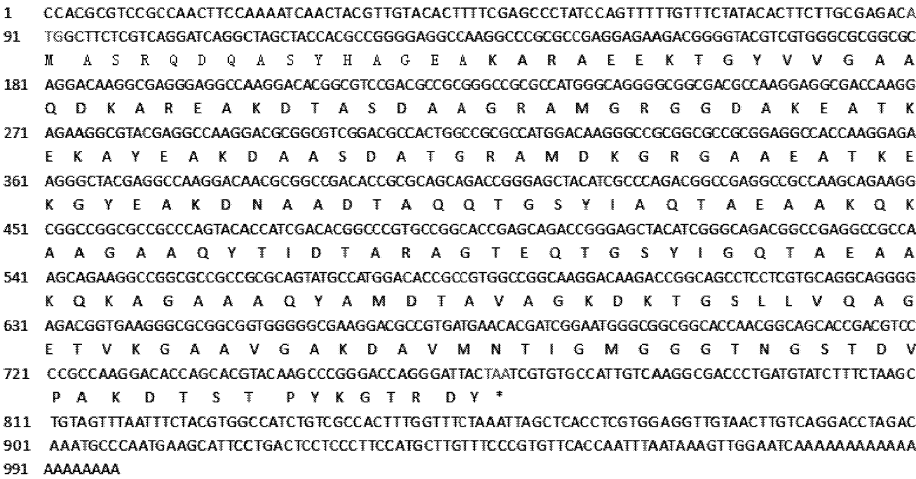


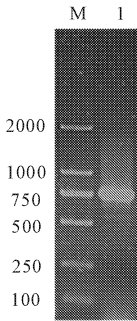
图 1 电子克隆序列 *TaWSII8 – EST* 及预测的氨基酸序列

Fig.1 Electronic cloning sequence of *TaWSII8 – EST* and the prediction of its amino acid sequence

2.2 扩增产物序列与电子克隆序列比对分析

以干旱处理的小麦叶片的 cDNA 为模板, 采用基因特异性引物 F 和 R 扩增目的片段, 此目的片段长度为 764 bp(图 2), 测序获得 4 条高度同源的 *TaWSII8* 基因, 分别命名为 *TaWSII8 – 1*、*TaWSII8 – 2*、*TaWSII8 – 3* 和 *TaWSII8 – 4*, Blast 分析表明, 克隆的 4 条序列与 *TaWSI – EST* 序列相似性分别为 99%、99%、99% 和 98%。ClutalX 比对分析与 *TaWSII8 – EST* 序列的差异, 以克隆序列的第一个碱基为第一位, *TaWSII8 – 1* 基因在 539、599、647 和 699 位点与其存在碱基差异, 752 和 753 位存在两个碱基缺失差异。 *TaWSII8 – 2* 基因在 289、290、291、293、539、599、647 和 699 位点存在碱基差异, 752 和 753 位点两个碱基缺失差异。 *TaWSII8 – 3* 基因碱基差异位点与 *TaWSI – like2* 基因的差异位点相同, 但 752 和 753 位点与 *TaWSII8 – EST* 序列相同。 *TaWSII8 – 4* 基因与 *TaWSII8 – EST* 序列的差异较其他 3 个基因的大, 在 25 ~ 39、209、288 ~ 296、346、362、365、370、499、536 和 566 位点存在差异, 40 ~ 57 位点缺失差异(表 1)。利用 ORF Finder 软件分析, 结果表明 *TaWSII8 – 1*、*TaWSII8 – 2*、*TaWSII8 – 3* 和 *TaWSII8 – 4* 基因 ORF 框均为 678 bp, 编码 225 个氨基酸。Blast 分析表明, 克隆的基因与 *TaWSII8 – EST* 蛋白序列相似性分别为 99%、99%、99% 及 97%。ClustalX 比对分析, *TaWSII8 – 2* 和 *TaWSII8 – 3* 基因

ORF 框一致, 编码相同的蛋白。小麦中的 *WSII8* 蛋白与 *TaWSII8 – EST* 氨基酸序列的差异性如表 2 所示。



注 Note: M—Marker; I—PT – PCR 产物 PT – PCR product

图 2 干旱条件下小麦叶片中 *TsWSII8* 基因的 RT – PCR 电泳结果

Fig.2 The RT – PCR of *TsWSII8* from wheat leaves under drought conditions

在 NCBI 网站通过 Blast 分析, 本文克隆的基因与已公布的无芒雀麦的 *BiWSII8* 基因(AB813835)、二穗短柄草的 *lea3* 基因(XM-003569593.1)、野生水稻(*Oryza brachyacha*)的 *lea3* 基因(XM-006644519)和日本水稻(*Oryza sativa Japonica*)的 *WSII8* 基因(D26536)相似性分别约为 94%、90%、90%、90% 和 90%。并与其他植物的水分胁迫相关的基因具有一定的相似性, 证实所得到的基因为水分胁迫诱导基因。

表 1 小麦的 *TaWSH18-1*、*TaWSH18-2*、*TaWSH18-3*、*TaWSH18-4* 基因与电子克隆序列 *TaWSH18-EST* 的差异碱基

Table 1 The sequence differences between *TaWSH18-1*, *TaWSH18-2*, *TaWSH18-3*, *TaWSH18-4* and *TaWSH18-EST* in wheat

| 基因<br>Gene           | 位点 Positive |    |    |    |    |    |    |    |    |                    |  |  |  |  |     |     |     |     |
|----------------------|-------------|----|----|----|----|----|----|----|----|--------------------|--|--|--|--|-----|-----|-----|-----|
|                      | 25          | 26 | 27 | 31 | 32 | 33 | 34 | 36 | 39 | 40 ~ 57            |  |  |  |  | 209 | 288 | 289 | 290 |
| <i>TaWSH18 - 2</i>   | T           | G  | T  | T  | A  | T  | A  | A  | T  | ————               |  |  |  |  | C   | A   | A   | C   |
| <i>TaWSH18 - 3</i>   | T           | G  | T  | T  | A  | T  | A  | A  | T  | ————               |  |  |  |  | C   | A   | A   | C   |
| <i>TaWSH18 - 1</i>   | T           | G  | T  | T  | A  | T  | A  | A  | T  | ————               |  |  |  |  | C   | A   | C   | T   |
| <i>TaWSH18 - EST</i> | T           | G  | T  | T  | A  | T  | A  | A  | T  | ————               |  |  |  |  | C   | A   | C   | T   |
| <i>TaWSH18 - 4</i>   | G           | C  | G  | G  | T  | G  | C  | G  | C  | GATCCAATCCTACCTTAG |  |  |  |  | T   | C   | C   | C   |

| 基因<br>Gene           | 位点 Positive |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |     |
|----------------------|-------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
|                      | 291         | 292 | 293 | 294 | 295 | 296 | 346 | 362 | 365 | 370 | 499 | 536 | 539 | 566 | 599 | 647 | 699 | 752 | 753 |
| <i>TaWSH18 - 2</i>   | T           | G   | G   | C   | G   | C   | A   | C   | G   | A   | G   | C   | A   | G   | A   | G   | A   | —   | —   |
| <i>TaWSH18 - 3</i>   | T           | G   | G   | C   | G   | C   | A   | C   | G   | A   | G   | C   | A   | G   | A   | G   | A   | T   | G   |
| <i>TaWSH18 - 1</i>   | G           | G   | C   | C   | G   | C   | A   | C   | G   | A   | G   | C   | A   | G   | A   | G   | A   | —   | —   |
| <i>TaWSH18 - EST</i> | G           | G   | C   | C   | G   | C   | A   | C   | G   | A   | G   | C   | C   | G   | G   | C   | C   | T   | G   |
| <i>TaWSH18 - 4</i>   | C           | C   | G   | G   | C   | G   | G   | G   | C   | G   | A   | G   | C   | C   | G   | C   | C   | T   | G   |

表 2 小麦的 *TaWSH18-1*、*TaWSH18-2*/*TaWSH18-3*、*TaWSH18-4* 蛋白与电子克隆序列 *TaWSH18-EST* 的编码蛋白的氨基酸差异

Table 2 The amino acid sequence differences between *TaWSH18-1*, *TaWSH18-2*, *TaWSH18-3*, *TaWSH18-4* and *TaWSH18-EST* in *Triticum aestivum*

| 蛋白<br>Protein                       | 位点 Positive |  |    |  |    |  |    |  |
|-------------------------------------|-------------|--|----|--|----|--|----|--|
|                                     | 74          |  | 75 |  | 76 |  | 93 |  |
| <i>TaWSH18-3</i> / <i>TaWSH18-2</i> | N           |  | W  |  | R  |  | Y  |  |
| <i>TaWSH18-EST</i>                  | T           |  | G  |  | R  |  | Y  |  |
| <i>TaWSH18-1</i>                    | T           |  | G  |  | R  |  | Y  |  |
| <i>TaWSH18-4</i>                    | P           |  | P  |  | A  |  | C  |  |

2.3 小麦 *TaWSH18* 蛋白特性分析

利用 ProtParam 在线分析,结果表明 *TaWSH18-1*、*TaWSH18-2*/*TaWSH18-3* 和 *TaWSH18-4* 蛋白的相对分子质量分别为 22.75 kD、22.90 kD 和 22.67 kD,等电点均为 6.64。正(Arg + Lys)、负(Asp + Glu)电荷氨基酸残基数目均为 31 个;不稳定指数分别为 4.45、4.37 和 13.3,表明蛋白均具有稳定性;用 Prot Scale 预测亲疏水性,结果显示 4 个蛋白多肽链的第 35 位表现为最低值 - 2.522,亲水性最强,第 179 位表现为最高值 1.078,疏水性最强,多肽链的多数位点值均小于 0,表现为亲水性(图 3);GRAVY 值分别为 - 0.771、- 0.786 和 - 0.741,表明蛋白为亲水性。这些性质与水分胁迫诱导性蛋白的稳定性、亲水性的特点相一致,有利于植物在受到胁迫时对细胞的保护作用。

2.4 小麦 *TaWSH18* 蛋白质的磷酸化位点及信号肽预测分析

运用 NetPhos 对该蛋白质序列进行潜在磷酸化

位点预测分析,结果表明,*TaWSH18-1*、*TaWSH18-2* 和 *TaWSH18-3* 具有 5 个丝氨酸磷酸化位点、10 个苏氨酸磷酸化位点和 3 个酪氨酸磷酸化位点(图 4a)。*TaWSH18-4* 存在 4 个丝氨酸磷酸化位点、9 个苏氨酸磷酸化位点以及 2 个酪氨酸磷酸化位点(图 4b)。同时利用 SignalP 预测蛋白信号肽结果显示,克隆的这 4 个蛋白均无信号肽结构域及剪切位点。

2.5 小麦 *TaWSH18* 蛋白质亚细胞定位及跨膜结构域的预测分析

利用 TargetP 1.1 Server 和 PSORT 在线分析小麦 *TaWSH18* 蛋白,结果表明这 4 个蛋白可能定位于细胞质中,通过 TMHMM 工具预测结果表明,这 4 个蛋白均无跨膜螺旋区,为非膜蛋白。

2.6 小麦 *TaWSH18* 蛋白质的二级结构分析

神经网络 HNN 对 *TaWSH18-1*、*TaWSH18-2*/*TaWSH18-3* 和 *TaWSH18-4* 蛋白二级结构的预测表明,*TaWSH18-1* 和 *TaWSH18-2*/*TaWSH18-3* 蛋白的  $\alpha$  螺旋、延伸链和无规则卷曲的比例均为 56%、

7.11% 及 36.89% (图 5a), TaWSI18-4 蛋白  $\alpha$  螺旋的比例为 50.67%, 延伸链为 6.67%, 无规则卷曲为 42.67% (图 5b)。

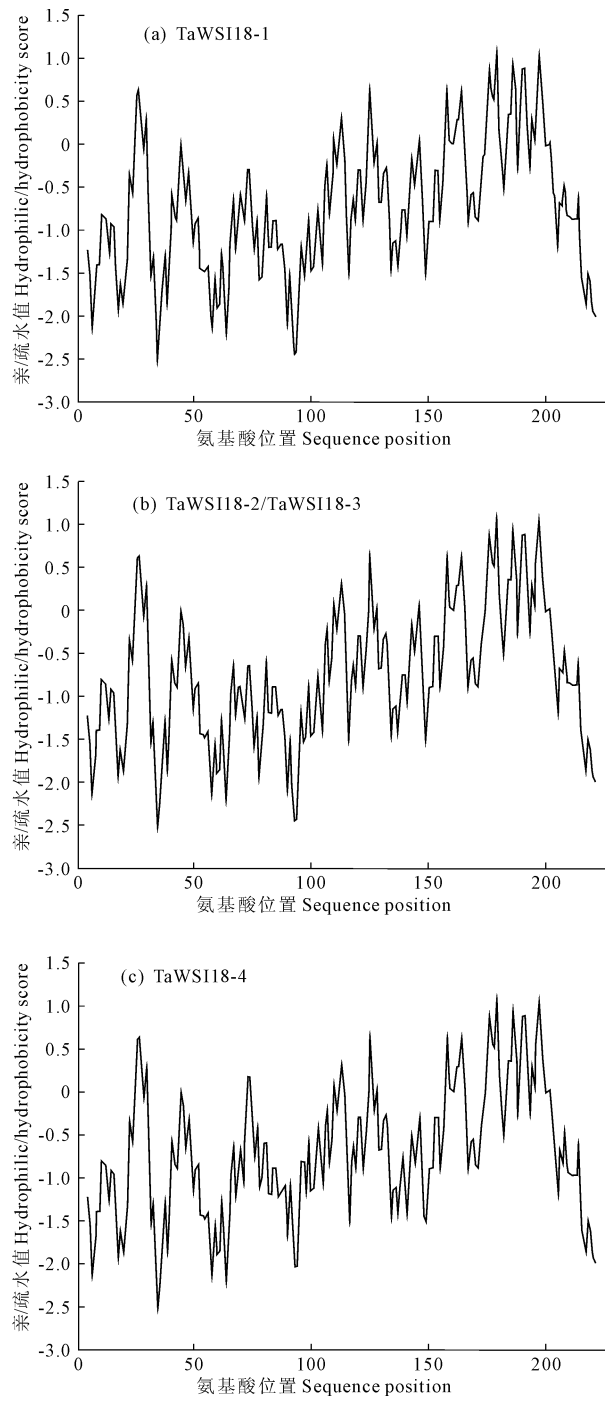


图 3 小麦 TaWSI18 蛋白的亲水/疏水性预测  
Fig.3 Hydrophilic/hydrophobicity analysis of TaWSI18 protein in *Triticum aestivum*

2.7 TaWSI18 蛋白序列比对及聚类分析

利用 NCBI 数据库中的 Blastp 程序对 TaWSI18-1、TaWSI18-2/TaWSI18-3 和 TaWSI18-4 的氨基酸序列进行分析,结果显示,小麦 TaWSI18 蛋白与无芒雀麦的 WSI18 蛋白(BAN15016)具有高度的同源性

(相似性为 94%);与二穗短柄草 LEA3 蛋白(XP\_003569641)的相似性为 70%,与其他物种的 WSI18 蛋白及 LEA 蛋白也有一定的相似性。其氨基酸序列的 C 端富含丙氨酸。ClutalX 比对分析表明,其 N 端存在与 LEA 蛋白的 N 端区域高度保守的序列,该保守性在禾本科植物的物种间,较其他亲缘关系较远的物种〔如双子叶的芜菁 (*Brassica rapa*, XP\_009115323) 和 拟 南 芥 (*Arabidopsis thaliana*, CAA63006)]更为保守(图 6)。

利用 MEGA 5.0 软件进行聚类分析表明(图 7),小麦中的 WSI18 蛋白同源性最高,在不同物种间,与禾本科植物的相似性也颇高,其结果与小麦和这些物种同属于禾木科植物的传统分类相一致。其中与无芒雀麦相似性最高,与二穗短柄草次之。芜菁与拟南芥这两个双子叶十字花科则形成了另外的独立分支。聚类分析基本上代表了它们在经典分类上的地位。

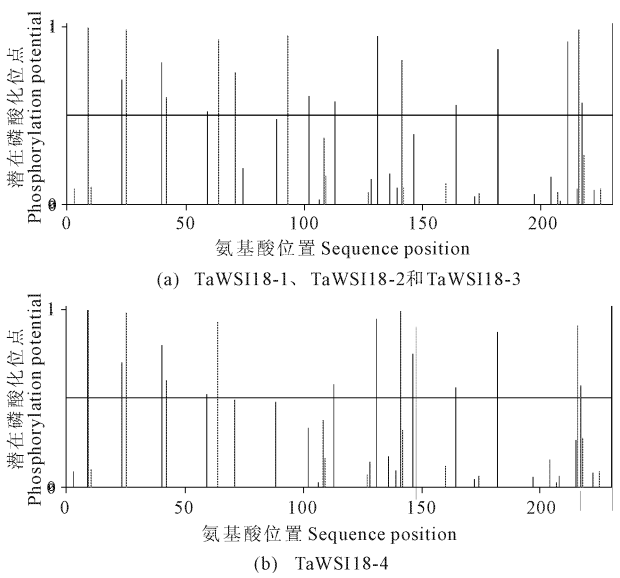


图 4 小麦 TaWSI18 蛋白的磷酸化位点预测  
Fig.4 Phosphorylation sites prediction of TaWSI18 protein in *Triticum aestivum*

3 讨 论

小麦是异源六倍体作物,含有 42 条染色体,其 DNA 含有大量的核苷酸重复,本实验以 1 条电子克隆序列为基础,设计 1 对特异引物,克隆得到 4 条高度同源的 WSI18 基因,这与小麦庞大的基因组及其异源六倍体有很多相关性。

从小麦中获得的 4 条水分胁迫诱导基因,其序列与 lea 基因的相似度很高,其中与二穗短柄草(XM\_003569593.1)和野生水稻(XM\_006644519)的 lea3 基因相似性达到 90%。WSI18 基因编码的蛋白

与二穗短柄草(XP\_003569641)LEA3 蛋白的相似性为 70%,与其他物种的 LEA 蛋白也有一定的相似性,其中主要与 LEA3 蛋白的相似度较高。而且小麦 TaWSI18 蛋白的 N 端存在与 LEA 蛋白的 N 端区

域高度保守的序列。对小麦的 TaWSI18 蛋白序列进行 Blastp 分析及 ClustalX 比对和聚类分析(图 6 和图 7)表明,TaWSI18 蛋白序列可能属于 LEA3 - like 蛋白。

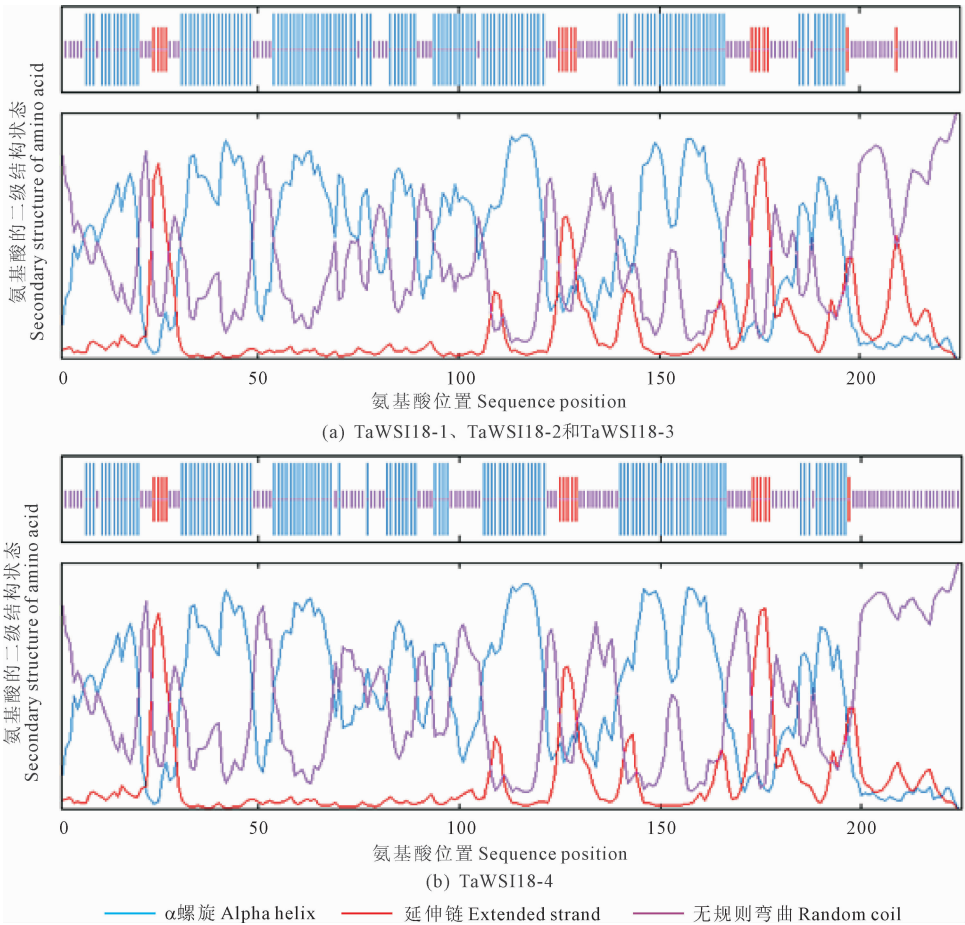


图 5 小麦 TaWSI18 蛋白质的二级结构分析

Fig.5 Secondary structure of TaWSI18 protein in *Triticum aestivum*

通过生物信息学分析,小麦 TaWSI18 蛋白相对分子量较小,蛋白序列缺少 Cys、Trp 和 Phe,富含 Ala、Gly、Lys 和 Thr,具有稳定性和亲水性,与已证实的 LEA 蛋白的序列特征极为相似<sup>[6,15]</sup>。Dure L 对 LEA3 蛋白的 11 个氨基酸残基进行生物信息学分析认为,该肽段以  $\alpha$  螺旋形式存在<sup>[16]</sup>,圆二色谱和红外光谱对 LEA3 蛋白的结构分析表明,该类蛋白在水溶液中以无序化结构存在,而在具有甲醇、蔗糖、甘油、乙烯、乙二醇或者脱水状态时,则形成  $\alpha$  螺旋结构<sup>[17-20]</sup>,TFE 和 SDS 也能够促使 LEA3 蛋白螺旋化<sup>[20]</sup>。许多研究还发现,缺水和复水条件下,LEA3 蛋白结构的螺旋化与无序化之间的转化有利于其在细胞中行使多种功能<sup>[21]</sup>。HNN 和 GOR 对 TaWSI18 蛋白的二级结构分析表明该类蛋白主要以  $\alpha$  螺旋和无规则卷曲形式存在。TaWSI18 蛋白的基本特点和二级结构与 LEA3 蛋白类似,推测其在细胞当中

与 LEA3 蛋白具有相同或相似的功能。

Takahashi R 等<sup>[12]</sup>发现,水稻的 *WSI18* 基因其蛋白序列的 N 端区域与 LEA3 蛋白高度同源。对该基因的启动子分析表明,其含有 ABA 响应元件,利用瞬时表达研究 *WSI18* 启动子和缺失功能元件启动子,结果表明,水稻的 *WSI18* 基因受干旱和 ABA 胁迫响应<sup>[13,22]</sup>。推测小麦 *TaWSI18* 基因的表达可能受到干旱和 ABA 等非生物胁迫的诱导表达,目前对小麦的 *TaWSI18* 基因的功能研究甚少,小麦基因组的庞大与复杂使重要功能基因的克隆存在困难。但随着测序技术的发展,新型大容量基因组 cDNA 文库的构建以及各种分子标记技术的发展,基因组文库和 EST 库的建立,为基因的克隆及其重要功能研究提供便利。EST 序列电子克隆基因为进一步深入研究 *WSI18* 基因提供参考和奠定基础。

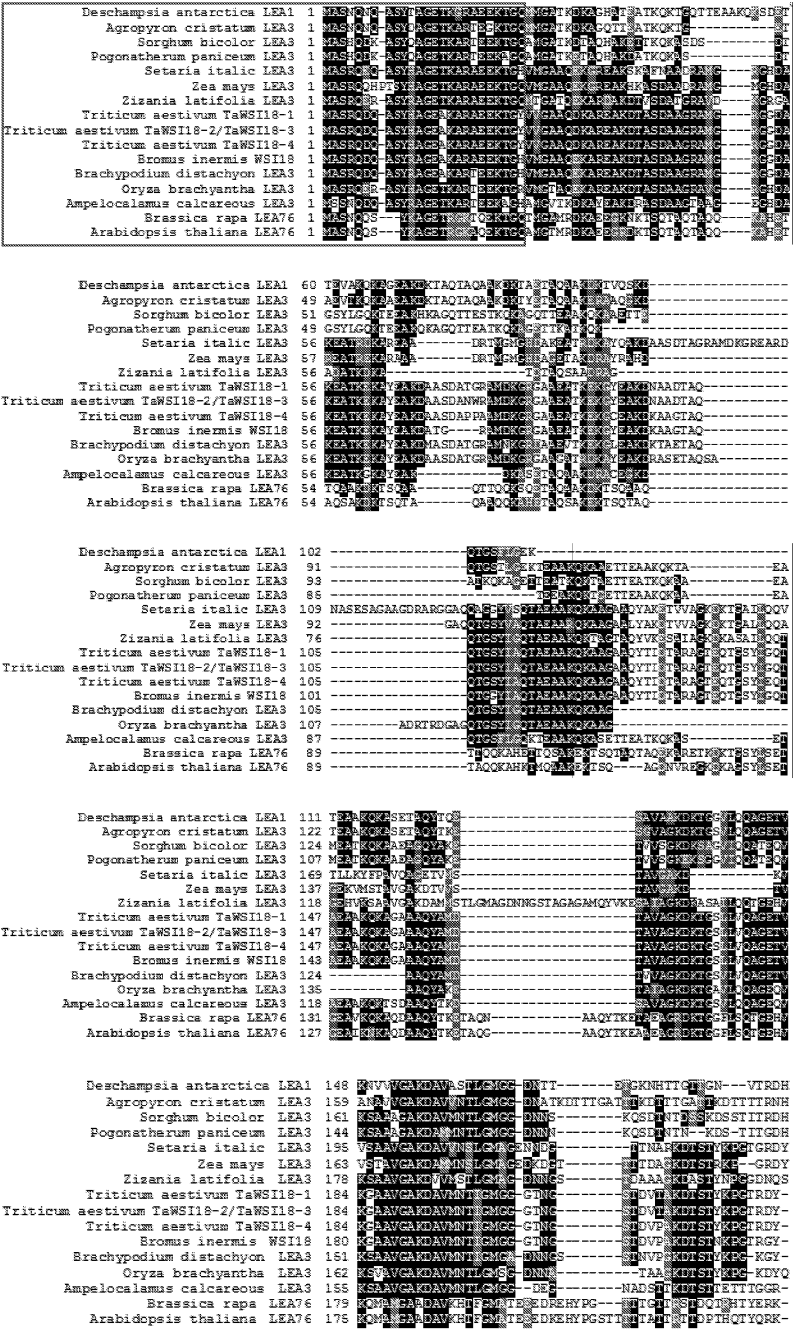


图 6 ClustalX 比对分析小麦 *TaWSI18* 蛋白与 LEA 蛋白氨基酸序列(黑色框内为 N 端保守区)  
Fig.6 Protein sequence multi-alignment of *TaWSI18* proteins and LEA protein of *Triticum aestivum* using ClustalX (The black box for N – terminal conservative region)

参考文献:

[1] 杨献光,梁卫红,齐志广,等.植物非生物胁迫应答的分子机制[J].麦类作物学报,2006,26(6):158-161.  
[2] Kaur N, Gupta A K. Signal transduction pathways under abiotic stresses in plants[J]. Current Science, 2005,88(11):1771-1780.  
[3] Shinozaki K, Yamaguchi Shinozaki K. Gene expression and signal transduction in water-stress response[J]. Plant Physiology, 1997,115(2):327-334.  
[4] Dure L, Chlan C. Developmental biochemistry of cottonseed embryoge-

nesis and germination: XII. Purification and properties of principal storage proteins[J]. Plant physiology, 1981,68(1):180-186.  
[5] Dure L, Galau G A. Developmental biochemistry of cottonseed embryogenesis and germination: XIII. Regulation of biosynthesis of principal storage proteins[J]. Plant physiology, 1981,68(1):187-194.  
[6] Battaglia M, Olvera-Carrillo Y, Garcarrubio A, et al. The enigmatic LEA proteins and other hydrophilins[J]. Plant Physiology, 2008,148(1):6-24.  
[7] Dalal M, Kumar G S, Mayandi K. Identification and expression analysis of group 3 LEA family genes in *Sorghum bicolor* (L.)[J]. Moench Acta Physiologiae Plantarum, 2013,35(3):979-984.

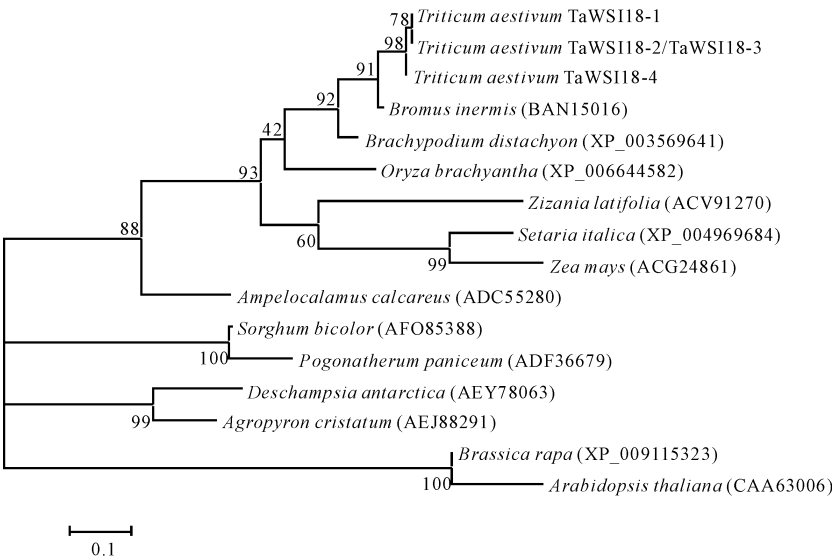


图 7 系统进化分析

Fig.7 Phylogenetic tree of TaWSI18 and other homologous sequences

[8] Vaseva I I, Anders I, Feller U. Identification and expression of different dehydrin subclasses involved in the drought response of *Trifolium repens*[J]. *Journal of Plant Physiology*, 2014,171(3-4):213-224.

[9] Amara I, Odena A, Oliveira E, et al. Insights into maize LEA proteins: from proteomics to functional approaches[J]. *Plant and Cell Physiology*, 2012,53(2):312-329.

[10] Hatanaka R, Hagiwara-Komoda Y, Furuki T, et al. An abundant LEA protein in the anhydrobiotic midge, PvLEA4, acts as a molecular shield by limiting growth of aggregating protein particles[J]. *Insect Biochemistry and Molecular Biology*, 2013,43(11):1055-1067.

[11] 张瑞越,徐兆师,李连城,等.一个新的小麦非生物胁迫诱导基因的克隆及表达特性[J]. *中国农业科学*,2007,40(5):875-881.

[12] Takahashi R, Joshee N, Kitagawa Y. Induction of chilling resistance by water-stress, and cdna sequence-analysis and expression of water stress-regulated genes in rice[J]. *Plant Molecular Biology*, 1994,26(1):339-352.

[13] Joshee N, Kisaka H, Kitagawa Y. Isolation and characterization of a water stress-specific genomic gene, pws18, from rice[J]. *Plant and Cell Physiology*, 1998,39(1):64-72.

[14] 张德礼,孙晓静,凌伦奖,等.人类SR蛋白超家族新成员—SFRS12(SRrp508)的基因克隆和特性分析[J].*遗传学报*,2002,29(5):377-383.

[15] Garay-Arroyo A, Colmenero-Flores J. M, Garcarrubio A, et al. Highly hydrophilic proteins in prokaryotes and eukaryotes are common during conditions of water deficit[J]. *Journal of Biological Chemistry*, 2000,275(8):5668-5674.

[16] Dure L. A repeating 11 – MER Amino – acid motif and plant desiccation[J]. *Plant Journal*, 1993,3(3):363-369.

[17] Dure L. Occurrence of a repeating 11 – mer amino acid sequence motif in diverse organisms[J]. *Protein and Peptide Letters*, 2001,8(2):115-122.

[18] Goyal K, Tisi L, Basran A, et al. Transition from natively unfolded to folded state induced by desiccation in an anhydrobiotic nematode protein[J]. *Journal of Biological Chemistry*, 2003,278(15):12977-12984.

[19] Toller D, Jaquinod M, Mangavel C, et al. Structure and function of a mitochondrial late embryogenesis abundant protein are revealed by desiccation[J]. *Plant Cell*, 2007,19(5):1580-1589.

[20] Wolkers W F, McCready S, Brandt W F, et al. Isolation and characterization of a D – 7 LEA protein from pollen that stabilizes glasses in vitro[J]. *Biochimica Et Biophysica Acta – Protein Structure and Molecular Enzymology*, 2001,1544(1-2):196-206.

[21] Sun X, Rikkerink E H A, Jones W T. Multifarious roles of intrinsic disorder in proteins illustrate its broad impact on plant biology[J]. *Plant Cell*, 2013,25(1):38-55.

[22] Zhu W N, Zhang L S, Lv H, et al. The dehydrin wzy2 promoter from wheat defines its contribution to stress tolerance[J]. *Functional & Integrative Genomics*, 2014,14(1):111-125.