

# 拟南芥和水稻 *UXS* 基因家族生物信息学分析

潘玉欣, 王巍杰, 胡金山

(河北联合大学, 河北 唐山 063009)

**摘要:** 以 6 个拟南芥和 6 个水稻 *UXS* 基因家族序列为目标, 对其基因结构、保守结构域、基因表达、系统进化等方面进行了综合分析。结果显示, 12 个 *UXS* 基因均有内含子, 除 *OsUXS1* 基因仅在愈伤组织中表达外, 其余 11 个基因在根、叶以及愈伤组织均有表达。12 个 *UXS* 基因编码蛋白均存在较大范围的亲水区, 有较强的亲水性。12 个蛋白结构保守性较强, 含有该基因家族的保守域 3Beta-HSD 和 NAD-binding, 分成 2 个亚家族, 家族内结构相似的基因功能较为相似。综合分析表明, *UXS* 是一个多基因家族, 基因表达范围广, 结构保守性强。

**关键词:** 尿苷二磷酸葡萄糖醛酸脱羧酶; 拟南芥; 水稻; 生物信息学; 结构与功能预测

**中图分类号:** Q74      **文献标识码:** A      **文章编号:** 1004-3268(2011)11-0038-05

## Bioinformatic Analysis of *UXS* Gene Family in *Arabidopsis* and Rice

PAN Yu-xin, WANG Wei-jie, HU Jin-shan

(Hebei United University, Tangshan 063009, China)

**Abstract:** In this study, six UDP-glucuronate decarboxylase (*UXS*) genes separately from *Arabidopsis* and rice were analyzed from the gene structure, conservative motif, gene expression and phylogenesis. The results showed that 12 *UXS* genes had introns. All the *UXS* genes expressed in roots, leaves and calluses, except for *OsUXS1* expressing only in calluses. The 12 *UXS* proteins were hydrophilic proteins with a wide range of hydrophilic areas and had highly conserved structure with the family domain 3Beta-HSD and NAD-binding. The 12 *UXS* proteins belonged to two sub-families and their structures and functions were similar in a sub-family. The comprehensive analysis revealed that as a multi-gene family, the *UXS* gene expressed widely and had conservative structure.

**Key words:** UDP-glucuronate decarboxylase(*UXS*); *Arabidopsis*; Rice; Bioinformatics; Prediction of structure and function

木糖 (UDP-xylose, UDP-xyl) 是合成植物细胞壁多糖木糖、木聚糖、阿拉伯糖、果胶以及其他小分子代谢物必需的核苷糖<sup>[1]</sup>。在植物上, UDP-Xyl 主要由尿苷二磷酸葡萄糖醛酸脱羧酶 (UDP-glucuronate decarboxylase, *UXS*) 催化尿苷二磷酸葡萄糖醛酸 (UDP-glucuronate, UDP-GlcA) 形成, 由于 *UXS* 催化 UDP-GlcA 形成 UDP-xyl 反应为不可逆反应, 因此, *UXS* 在核苷糖转化过程中起重要作用,

参与多种细胞壁合成, 在植物界广泛存在<sup>[1-4]</sup>。

*UXS* 蛋白属于 NAD 异构酶/脱水酶基因家族 (NAD dependent epimerase/dehydratase, PF01370), 包含脱水酶、脱氢酶、表异构酶的保守域, 在 N 端存在该基因家族的典型模体序列, ADP-binding  $\beta\alpha\beta\alpha\beta$  折叠 (3Beta-HSD)、NAD(P)-binding 蛋白的一个特征序列 GXXGXXG 以及三位一体催化结构 Ser-Lys-Tyr, 其中 Lys 和 Tyr 存在于 YXXXXK 结构中。在

收稿日期: 2011-05-21

基金项目: 河北省青年自然科学基金 (C2010000919); 唐山市科技局项目 (1014020C-4); 河北理工大学博士自然科学基金

作者简介: 潘玉欣 (1979-), 女, 河北沧州人, 讲师, 博士, 主要从事棉花遗传育种研究工作。E-mail: panyu-xin@163.com

三位一体催化结构 Ser-Lys-Tyr 的作用下,UDP-GlcA C<sub>4</sub> 位羟基的氢离子从糖的 C<sub>4</sub> 位置上转移到 NAD<sup>+</sup> 上,形成一个 C<sub>4</sub> 位酮的中间体,进而 NADH 的氢离子转移到中间体上,形成 UDP-xy<sup>[1,4]</sup>。现阶段已有多种植物 *UXS* 基因报道,但对拟南芥和水稻 *UXS* 家族的对比分析报道甚少,本研究选取拟南芥和水稻的 *UXS* 基因家族成员进行结构、表达模式以及系统进化等分析,旨在揭示植物 *UXS* 基因家族的结构特点、进化历史,为后续研究 *UXS* 家族的功能和进化上的多样性提供理论基础。

## 1 材料和方法

### 1.1 材料

从 TIGR 拟南芥和水稻数据库中获取已注册的拟南芥、水稻的 *UXS* 基因核酸及其对应蛋白质序列(表 1)。

表 1 拟南芥和水稻 *UXS* 家族成员

基因名称	位点	基因名称	位点
<i>AtUXS1</i>	AT3G53520	<i>OsUXS1</i>	LOC_Os03g17230
<i>AtUXS2</i>	AT3G62830	<i>OsUXS2</i>	LOC_Os01g21320
<i>AtUXS3</i>	AT5G59290	<i>OsUXS3</i>	LOC_Os03g16980
<i>AtUXS4</i>	AT2G47650	<i>OsUXS4</i>	LOC_Os01g62020
<i>AtUXS5</i>	AT3G46440	<i>OsUXS5</i>	LOC_Os05g29990
<i>AtUXS6</i>	AT2G28760	<i>OsUXS6</i>	LOC_Os07g47700

### 1.2 分析方法

1.2.1 *UXS* 基因结构分析 利用在线基因结构显示系统 GSDS (gene structure display server) (<http://gsds.cbi.pku.edu.cn/chinese.php>),对拟南芥、水稻 *UXS* 基因进行结构预测<sup>[5]</sup>。

1.2.2 编码 *UXS* 蛋白质的亲疏水性、跨膜结构以及亚细胞定位预测 通过在线 ProScale (<http://expasy.org/tools/protscale.html>) 分析系统,预测拟南芥和水稻 *UXS* 蛋白亲/疏水性,在线 TM-HMM Server v. 2.0 (<http://www.cbs.adt.dk/>

[service/TMHMM](http://www.cbs.adt.dk/service/TMHMM), 2.0) 分析系统预测跨膜结构,在线 PSORT 程序 (<http://www.psорт.org/>) 预测拟南芥和水稻 *UXS* 蛋白亚细胞定位。

1.2.3 *UXS* 保守结构域预测和多序列联配分析 分别利用 SMART 和 Clustal W 对拟南芥和水稻 *UXS* 蛋白结构域以及蛋白序列进行多序列联配分析。

1.2.4 *UXS* 蛋白多序列联配和系统发生树构建 以 Clustal W 序列联配结果为基础,用 MEGA 程序建立拟南芥和水稻的 *UXS* 蛋白系统进化树<sup>[6]</sup>。进化树生成采用邻接法(NJ method),校验参数为 Bootstrap 1000 次重复,对构建的系统发生树进行评估。

1.2.5 *UXS* 基因表达模式分析 对拟南芥与水稻 *UXS* 家族成员分别在 MPSS 相关物种网站<sup>[7]</sup>, <http://mpss.udel.edu/at/> 以及 <http://mpss.udel.edu/rice/> 中进行表达模式预测,分析愈伤组织、幼叶以及幼根中的表达量差异。

## 2 结果与分析

### 2.1 *UXS* 基因结构分析

比较 *UXS* 基因 CDS 与基因序列,12 条 *UXS* 序列均有内含子序列,内含子数量为 5~12 个不等(图 1)。其中 *OsUXS3* (LOC\_Os03g16980) 是内含子数目最多的成员,共有 12 个内含子,*OsUXS6* (LOC\_Os07g47700) 内含子数目最少,含有 5 个内含子。根据内含子数量,12 个 *UXS* 分成 A、B 2 个亚族,其中 A 亚族内含子数目较少,B 亚族内含子数量则较多,均在 9 个以上,包括 *AtUXS3*、*AtUXS5*、*AtUXS6* 以及 *OsUXS3*。A、B 亚族内成员,内含子、外显子分布具有一定程度上的相似性。*AtUXS2* (AT3G62830) 与 *AtUXS4* (AT2G47650),*OsUXS2* (LOC\_Os01g2132030) 与 *OsUXS5* (LOC\_Os05g29990) 内含子分布位置及数目都较为相似。

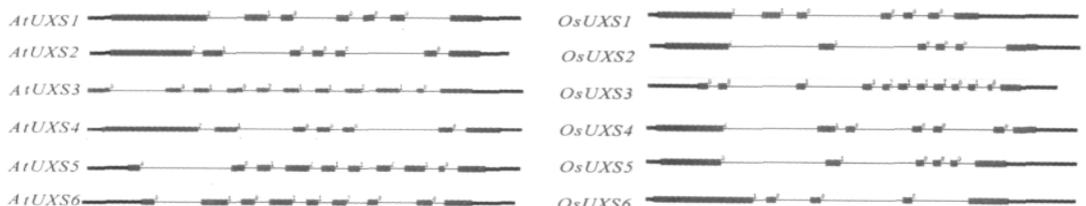


图 1 拟南芥和水稻 *UXS* 基因结构分布

### 2.2 编码蛋白质的跨膜结构、亲疏水性以及亚细胞定位分析

通过在线的 ProScal 对蛋白的亲/疏水性进行

分析,获得 12 个基因编码蛋白的亲/疏水性分布曲线(图 2)。12 个蛋白均存在较大范围的亲水区,有较强的亲水性,属于可溶性蛋白,N 端多含有一个明

显的疏水区。拟南芥和水稻 UXS 蛋白跨膜区预测显示(表 2)所示,50%的 UXS 氨基酸总值大于 18, N 端均有一个跨膜区存在,属于跨膜蛋白;50%的 UXS 氨基酸总值小于 18,不存在跨膜区。与亲疏

水性分析结果一致,含有跨膜区的 AtUXS1 等 6 个蛋白,N 端跨膜区位置均对应有一段明显的疏水区(图 3)。亚细胞定位如表 2,除 OsUXS2 外,不含跨膜区的 UXS 蛋白均定位于细胞膜上。

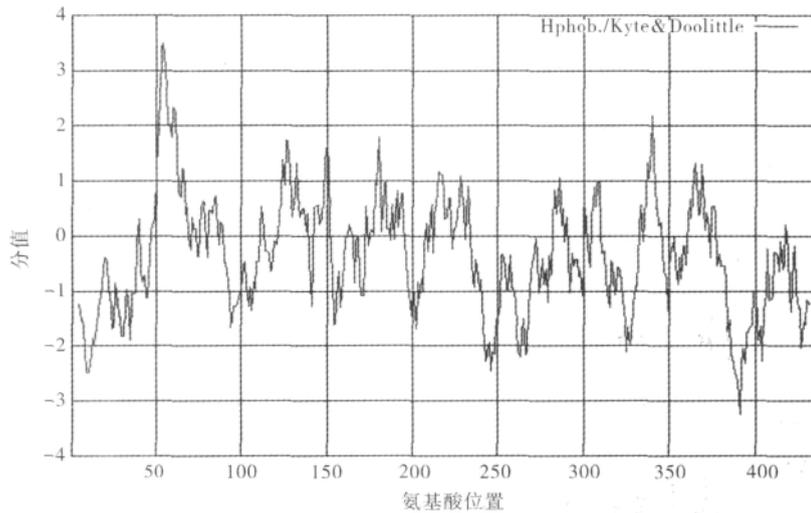
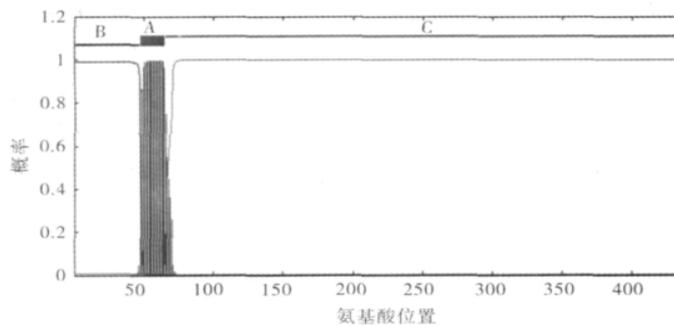


图 2 拟南芥 AtUXS1 蛋白的亲疏水性预测结果

表 2 拟南芥和水稻 UXS 蛋白跨膜区、亚细胞定位

蛋白名称	跨膜区位置	亚细胞定位	蛋白名称	跨膜区位置	亚细胞定位
AtUXS1	48-65	细胞膜	OsUXS1	37-59	细胞膜
AtUXS2	44-61	细胞膜	OsUXS2	无	细胞膜
AtUXS3	无	细胞质	OsUXS3	无	细胞质
AtUXS4	44-61	细胞膜	OsUXS4	44-66	细胞膜
AtUXS5	无	细胞质	OsUXS5	无	细胞质
AtUXS6	无	细胞质	OsUXS6	58-80	细胞膜



注:A 部分表示跨膜区,B 部分表示膜内,C 部分表示膜外

图 3 拟南芥 AtUXS1 蛋白跨膜区预测结果

### 2.3 基因结构域与多序列联配分析

采用 SMART 对蛋白质结构域功能分析显示,12 条 UXS 蛋白质序列均含有 Epimerase 保守域,属于 NAD dependent epimerase/dehydratase 家族,含有该基因家族的保守域 3Beta-HSD 和 NAD-binding,具体位置如表 3 所示。

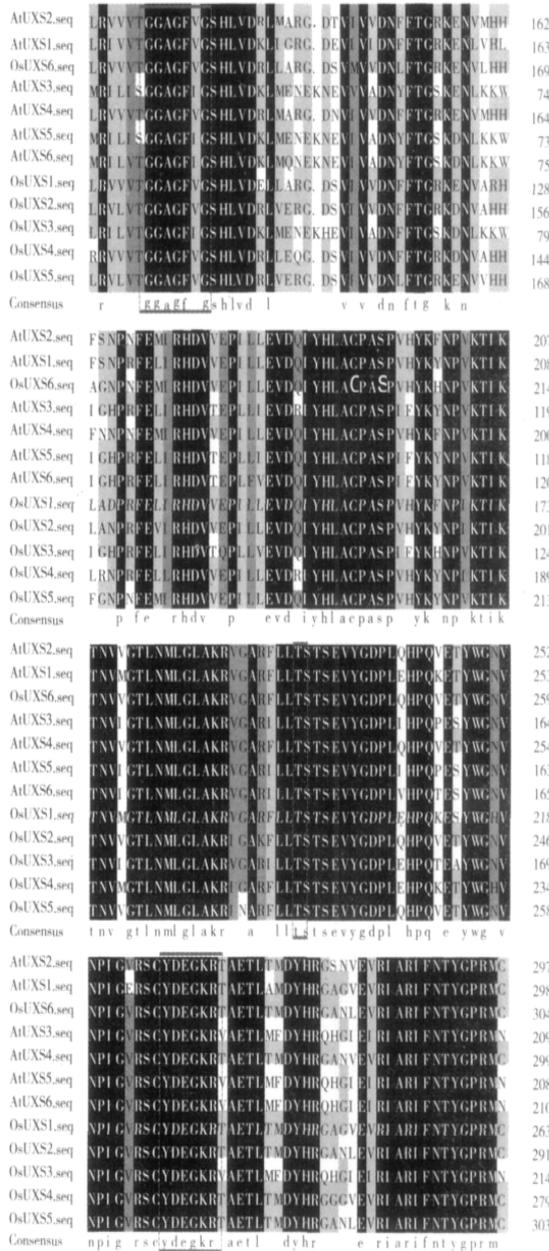
多重序列比较结果说明,12 条 UXS 氨基酸序列 N、C 末端差异较大。其中氨基酸最短序列 At-

UXS5 较最长氨基酸序列 OsUXS5 在 N 端大约缺少 69 个氨基酸,在 C 端大约缺少 11 个氨基酸。但 12 个蛋白中间部分氨基酸序列间同源性较高,蛋白保守性较强(图 4)。含有脱水酶、脱氢酶和表异构酶的保守区域,12 个 UXS 蛋白在 N 端存在着 UXS 保守域 GXXGXXG,该结构域是 ADP-binding  $\beta\alpha\beta\alpha\beta$  折叠(3Beta-HSD)和 NAD-binding 蛋白的一个特征序列。另外,该蛋白含有脱氢酶、表异构酶起

作用的三位一体催化结构 Ser-Lys-Tyr, 其中 Lys 和 Tyr 存在于 YXXXX 结构中。

表 3 拟南芥和水稻 *UXS* 蛋白保守域位置 SMART 预测结果

蛋白名称	NAD-binding 位置	3Beta-HSD 位置	蛋白名称	NAD-binding 位置	3Beta-HSD 位置
AtUXS1	124-337	123-383	OsUXS1	89-302	88-319
AtUXS2	123-339	122-353	OsUXS2	117-330	116-375
AtUXS3	34-251	33-269	OsUXS3	39-257	38-270
AtUXS4	125-341	124-355	OsUXS4	105-318	104-335
AtUXS5	33-250	32-268	OsUXS5	129-342	128-387
AtUXS6	35-252	34-270	OsUXS6	130-343	129-388



注:方框标志为 *UXS* 基因家族结构域 GXXGXXG、ST、YXXXX

图 4 拟南芥和水稻 *UXS* 蛋白中间序列 Clustal W 多序列联配

### 2.4 系统进化分析

根据进化树结果(图 5),拟南芥和水稻 12 条 *UXS* 蛋白家族分为 A、B 2 个亚族,在亚族分类上与

内含子分类结果完全一致,证实拟南芥和水稻 *UXS* 蛋白的相互交叉进化。而进化树 A、B 亚族均以拟南芥和水稻 *UXS* 旁系同源蛋白为主,AtUXS2 与 At-

UXS4, OsUXS2 与 OsUXS5 等这些旁系同源蛋白, 内含子分布也非常相似, 进一步说明单子叶和双子叶植物分化后又分别按物种特异性方式扩张, 产生新的基因, 具备新的功能。

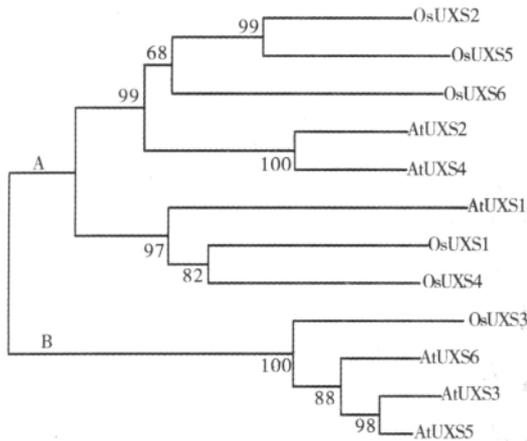


图 5 拟南芥和水稻 UXS 蛋白 MEGA 进化树分析

### 2.5 UXS 基因表达模式分析

结合 MPSS 网站分析结果, 选取根、叶、愈伤组织进行分析(表 4)。OsUXS1 仅在愈伤组织中表达, 其余各基因在 3 种组织中均有表达。不同基因存在组织表达差异。其中位于进化树同一分支的 AtUXS5 和 AtUXS6 基因, 内含子数目、疏水区、跨膜区、结构域等在位置结构特点上比较接近, 表达上均以愈伤组织较高。

表 4 拟南芥和水稻 UXS 蛋白表达分析

蛋白名称	根	叶	愈伤组织
AtUXS1	1	3	1
AtUXS2	1	1	3
AtUXS3	1	3	1
AtUXS4	3	1	1
AtUXS5	1	1	3
AtUXS6	1	1	3
OsUXS1	0	0	3
OsUXS2	1	3	1
OsUXS3	3	1	1
OsUXS4	3	1	1
OsUXS5	3	1	1
OsUXS6	1	3	1

注: 0 表示没有表达, 1 表示表达, 3 表示表达量最高

### 3 结论与讨论

本研究对拟南芥和水稻 UXS 基因结构、表达谱、结构域、功能等方面进行了综合的预测分析。12 个蛋白同源性比较分析表明, 利用 UXS 基因 cDNA 序列推导的氨基酸序列间具有高度的相似性, 说明该基因家族在初级结构上是保守的。不同植物中的 UDP-葡糖醛酸脱羧酶家族一般都存在于细胞质或细胞膜上。PSORT 分析也显示 12 个蛋白位于细胞质或细胞膜上。而对于多数蛋白来说, UXS 蛋白在 N 端存在一个明显的疏水区, 有跨膜区存在, 是跨膜蛋白, 这种蛋白

通常位于高尔基体上。与此类似, Hayashi 等以充足的证据证明大豆的 UDP-葡糖醛酸脱羧酶存在于高尔基体上, 需要在高尔基体上才有活性<sup>[8]</sup>。

对拟南芥水稻的 UXS 家族成员进行进化树分析, 结果显示, 12 个 UXS 分成 A、B 2 个亚族, B 类亚家族中的成员内含子多, 大部分的成员外显子都呈极不连续状。A 类亚家族成员内含子较少, 功能可能更为特化, 说明 A 类亚家族产生的较晚。B 类亚家族基因进化比 A 类亚家族更保守, 但在基因复制中插入不少内含子。A、B 亚家族中存在大量的旁系同源, 表明了拟南芥、水稻进化分离后 UXS 基因按照物种特异性方式进行了扩张。UXS 同一基因家族不同基因的活性是不同的<sup>[8]</sup>。

结合 MPSS 网站, 对拟南芥和水稻 UXS 家族成员表达模式分析, 发现不同 UXS 基因存在组织表达差异, 说明 UXS 基因作为细胞壁多糖合成的重要影响因素, 主要是在中间代谢中起作用, 在碳代谢中与糖蛋白、多糖合成等多种代谢过程相关, 广泛参与多种细胞壁形成过程。部分同源基因表达量相似, 说明 UXS 基因家族可能存在功能上的保守性。UXS 基因参与多糖代谢途径, 影响细胞壁形成, 通过生物信息学预测比较得到的信息, 将进一步为后续该基因家族功能研究奠定基础。

#### 参考文献:

- [1] April D H, Maor B P. Biosynthesis of UDP-xylose cloning and characterization of a novel *Arabidopsis* gene family, UXS, encoding soluble and putative membrane-bound UDP-Glucuronic acid decarboxylase isoforms [J]. *Plant Physiology*, 2002, 130: 2188-2198.
- [2] Zhang Q S, Shirley N, Lahnstein J, et al. Characterization and expression patterns of UDP-D-Glucuronate decarboxylase genes in barley [J]. *Plant Physiology*, 2005, 138: 131-141.
- [3] Maor B P, Griffith C L, Doering T L. Functional cloning and characterization of a UDP glucuronic acid decarboxylase: The pathogenic fungus *Cryptococcus neoformans* elucidates UDP-xylose synthesis [J]. *Proc Natl Acad Sci*, 2001, 21: 12003-12008.
- [4] Feingold D S. Aldo (and keto) hexoses and uronic acids [M]// *Plant carbohydrates I. Intracellular carbohydrates, new series, encyclopedia of plant physiology*. Berlin: Springer-Verlag, 1982: 3-76.
- [5] Guo A Y, Zhu Q H, Chen X, et al. GSDS: a gene structure display server [J]. *Hereditas*, 2007, 29(8): 1023-1026.
- [6] Kumar S, Dudley J, Nei M, et al. MEGA: A biologist-centric software for evolutionary analysis of DNA and protein sequences [J]. *Briefings in Bioinformatics*, 2008, 9: 299-306.
- [7] Nakano M, Nobuta K, Vemaraju K, et al. Plant MPSS databases; Signature-based transcriptional resources for analyses of mRNA and small RNA [J]. *Nucleic Acids Research*, 2006, 34: 731-735.
- [8] Hayashi T, Koyama T, Matsuda K. Formation of UDP-xylose and xyloglucan in soybean Gog1 membranes [J]. *Plant Physiol*, 1998, 87: 341-345.