

DOI: 10.11931/guihaia.gxzw202007007

巩元勇, 赵丽华, 闫飞, 等, 2023. 棉花 *DUR3* 基因的鉴定及进化分析 [J]. 广西植物, 43(2): 283–292.

GONG YY, ZHAO LH, YAN F, et al., 2023. Identification and evolutionary analysis of cotton *DUR3* gene [J]. *Guihaia*, 43(2): 283–292.



棉花 *DUR3* 基因的鉴定及进化分析

巩元勇¹, 赵丽华¹, 闫飞¹, 孙伊辰², 王慧², 刘来华^{2*}

(1. 攀枝花学院 生物与化学工程学院, 四川 攀枝花 617000; 2. 教育部植物与土壤互作重点实验室, 中国农业大学资源、环境及粮食安全中心, 北京 100193)

摘要: 植物 *DUR3* 同源蛋白属于钠离子/溶质共运蛋白家族的尿素高亲和力运输蛋白, 在植物体对外源尿素的主动吸收及内源尿素的再分配过程中具有重要作用。为明确棉花 *DUR3* 基因的结构和进化情况, 基于生物信息学的方法, 从全基因组水平鉴定陆地棉和雷蒙德氏棉的 *DUR3* 基因, 并对基因结构、跨膜结构域、基序分布、进化关系等进行分析。结果表明: (1) 从陆地棉 A 亚组和 D 亚组染色体各鉴定出 1 个 *DUR3* 基因, 从雷蒙德氏棉基因组鉴定出 1 个 *DUR3* 基因。这 3 个棉花 *DUR3* 同源蛋白同其他植物 *DUR3* 同源蛋白一样, 具有 15 个跨膜结构域, 具有 3 个位置一致、高度保守的基序。(2) 基因结构分析表明, 双子叶植物 *DUR3* 基因的外显子个数明显多于单子叶植物, 这 3 个棉花 *DUR3* 基因的外显子个数亦是如此。(3) 根据物种间种属亲缘关系, 对不同物种 *DUR3* 氨基酸序列构建的进化树显示, 棉花的同双子叶植物的聚在一起。(4) *DUR3* 直系同源基因和旁系同源基因的 K_a/K_s 比值普遍均大于 1, 说明这些基因在进化过程中主要受到正向选择的作用。该研究结果为深入研究棉花 *DUR3* 同源蛋白提供了理论基础。

关键词: 陆地棉, 雷蒙德氏棉, *DUR3* 基因, 生物信息, 进化

中图分类号: Q943 文献标识码: A 文章编号: 1000-3142(2023)02-0283-10

Identification and evolutionary analysis of cotton *DUR3* gene

GONG Yuanyong¹, ZHAO Lihua¹, YAN Fei¹, SUN Yichen²,
WANG Hui², LIU Laihua^{2*}

(1. *Biological and Chemical Engineering College, Panzhihua University, Panzhihua 617000, Sichuan, China*; 2. *Key Laboratory of Plant-Soil Interactions, Ministry of Education; Center for Resources, Environment and Food Security, China Agricultural University, Beijing 100193, China*)

Abstract: Plant *DUR3* homologous protein is a high affinity urea transporter which belongs to the family of sodium/solute symporter family, and plays an important role in the active absorption of exogenous urea and redistribution of endogenous urea by plants. The purpose of this study was to clarify structure and evolution situation of cotton *DUR3* gene. Based on bioinformatics methods, *DUR3* genes were identified from *Gossypium hirsutum* and *G. raimondii* genomic sequences, and then the gene structure, transmembrane domain, motif location, as well as phylogenetic relationship,

收稿日期: 2022-07-20

基金项目: 国家自然科学基金(31301682); 金沙江干热河谷生态修复与治理创新研究团队专项经费(035200179)。

第一作者: 巩元勇(1982-), 博士, 副教授, 主要从事植物生物技术研究, (E-mail) gyy2011qh@163.com。

*通信作者: 刘来华, 博士, 教授, 研究方向为植物营养分子生物学, (E-mail) ll1025@cau.edu.cn。

were systematically analyzed. The results were as follows: (1) Three *DUR3* genes were identified from the A and D subgroup chromosomes of upland cotton and *G. raimondii* genomic sequences. These three cotton *DUR3* homologous proteins, like other plant *DUR3* homologous proteins, had 15 transmembrane domains and three highly conserved motifs with consistent positions. (2) The gene structure analysis showed that the number of exons of *DUR3* genes in dicotyledons was significantly higher than that in monocotyledons, and so were the cotton *DUR3* genes. (3) Phylogenetic analysis revealed that the amino acid sequences of different species were classified according to species kinship, and cotton clustered in one branch with dicotyledons. (4) The K_a/K_s values of orthologous and paralogous *DUR3* genes were generally more than one, indicating that those genes mainly experienced positive selection among evolution. The results of this study provide a theoretical basis for further research on cotton *DUR3* homologous protein.

Key words: *Gossypium hirsutum*, *G. raimondii*, *DUR3* gene, bioinformatics, evolution

氮素在植物的生长过程中是一种非常重要的营养元素,因为它可以用来合成蛋白质、氨基酸、叶绿素等重要的氮化合物(Marschner, 1995)。尿素是全世界使用最普遍的化肥之一。因为脲酶——一种细菌分泌的可以降解尿素的酶,在土壤中几乎无处不在,大部分施入土壤中的尿素被脲酶分解为氨和二氧化碳,所以通常情况下土壤中尿素的浓度很低,不足以作为单一氮源满足植物生长需要(Kojima et al., 2006)。但是研究发现,施到土壤中的尿素存在 1~8 d 的半衰期(Liu et al., 2003a),在这期间土壤中会有一个较高浓度的尿素存在,为植物根系从土壤吸收尿素提供了宝贵的时间。至今为止,在植物中仅发现 2 类尿素转运蛋白,一类是具有低亲和力尿素运输能力、介导尿素被动迁移的 MIPs (Major Intrinsic Proteins),这是一类水通道蛋白,它们通常定位到细胞质膜及液泡膜上(Liu et al., 2003b);另一类是具有高亲和力尿素运输能力、介导尿素主动跨膜运输的属于钠离子/溶质共运蛋白家族(sodium/solute symporter family, SSS family)的 *DUR3* 同源蛋白,该类跨膜蛋白定位在质膜,且 *DUR3* 基因受到外界缺氮条件的诱导表达(Liu et al., 2003a; Liu et al., 2003b; Kojima et al., 2007)。 *DUR3* 的生理贡献不只局限于氮缺乏条件下的植物根部,在衰老的叶片中同样发挥一定的作用(Bohner et al., 2015)。植物体内的尿素不仅只能从外源土壤中获得,还能够从体内二次氮代谢的合成中获得,衰老叶片中尿素含量的增加就证明了这一观点(Bohner et al., 2015)。Bohner 等(2015)对拟南芥的研究显示, *AtDUR3* 基因在衰老叶片的维管组织中大量表达, *DUR3* 能够将叶片衰老引起的氮化合

物降解生成的尿素从叶肉细胞运送到质外体,这表明 *DUR3* 是植物体内氮素再分配的一个非常重要的转运通道蛋白。运用反向遗传学研究方法证实,在氮素缺乏的大田里, *OsDUR3* 基因插入系植株的水稻籽粒灌浆差,与正常的水稻相比,产量降低了 26.2%;在种子发育期, *OsDUR3* 基因插入系植株的水稻氮素主要积累在叶片中,花絮不能得到充分的发育;来自 *OsDUR3* 基因插入系植株老叶片的水稻尿素含量低于正常水稻。综上所述,在氮素缺乏的大田条件下 *DUR3* 参与了氮素转运和水稻产量(Beier et al., 2019)。

目前,只有拟南芥(*Arabidopsis thaliana*, *AtDUR3*) (Liu et al., 2003a)、水稻(*Oryza sativa*, *OsDUR3*) (Wang et al., 2012b)、玉米(*Zea mays*, *ZmDUR3*) (Zamin et al., 2014; Liu et al., 2015) 这 3 个高等植物的 *DUR3* 基因被克隆并验证了分子生理功能,这些研究都是围绕 *DUR3* 主动运输尿素开展。在酵母中的研究表明,酿酒酵母(*Saccharomyces cerevisiae*)的 *ScDUR3* 和白假丝酵母菌(*Candida albicans*) *CaDUR3* 除了具有转运尿素的功能以外,还能够转运多胺(Uemura et al., 2006; Kumar et al., 2011)。多胺是一类脂肪族小分子化合物,可广泛地参与到植物生长发育的各个阶段,同时对植物响应外界胁迫环境具有非常重要的作用(Groppa & Benavides, 2008),但是在植物中尚未发现转运多胺的蛋白。因此,植物 *DUR3* 同源蛋白是否具有转运多胺的功能值得进一步研究。

棉花是人类获取植物纤维的主要农作物,陆地棉(*Gossypium hirsutum*)是当前全球最重要的棉花栽培品种,全球 90% 以上棉田种植的都是陆地棉。陆地棉基因组是 AADD 型异源四倍体,使得

棉花基因组庞大而复杂,很大程度限制了棉花分子生物学的研究进度。随着二倍体 D 亚组雷蒙德氏棉(*G. raimondii*) (Wang et al., 2012a) 和 A 亚组亚洲棉(*G. arboreum*) (Li et al., 2014) 测序工作的完成,陆地棉遗传标准系 TM-1 全基因组序列也得以公布(Li et al., 2015; Zhang et al., 2015),在联合基因组研究所(Joint Genome Institute, JGI)上已经引入了陆地棉和雷蒙德氏棉基因组,这必将加速棉花功能基因的研究进度。本文从 JGI 上的陆地棉和雷蒙德氏棉基因组中鉴定棉花的 *DUR3* 基因,并进行相关生物信息学分析,可为该基因进一步的克隆和功能验证提供理论基础。

1 材料与方法

1.1 材料

从 NCBI(<https://www.ncbi.nlm.nih.gov/>) 搜索获得如下物种 *DUR3* 基因的序列:拟南芥(*Arabidopsis thaliana*, *AtDUR3*, NP_199351)、水稻(*Oryza sativa*, *OsDUR3*, NP_001065513)、玉米(*Zea mays*, *ZmDUR3*, KJ652242)、高粱(*Sorghum bicolor*, *SbDUR3*, XP_002438118.1)、谷子(*Setaria italica*, *SiDUR3*, XP_004965066.1)、番茄(*Solanum lycopersicum*, *SlDUR3*, XP_004245999.1)、葡萄(*Vitis vinifera*, *VvDUR3*, XP_002263043.1)、大豆(*Glycine max*, *GmDUR3*, XP_003523904.1)、大麦(*Hordeum vulgare*, *HvDUR3*, BAJ94433.1)、二穗短柄草(*Brachypodium distachyon*, *BdDUR3*, XP_003571687.1)、蒺藜苜蓿(*Medicago truncatula*, *MtDUR3*, XP_003612583.1)、毛果杨(*Populus trichocarpa*, *PtDUR3*, XP_002303472.2)、酿酒酵母(*Saccharomyces cerevisiae*, *ScDUR3*, L19875.1)、构巢曲霉(*Aspergillus nidulans*, *AnDUR3*, ACZ62639.1)、条斑紫菜(*Pyropia yezoensis*, *PyDUR3*, BAU04114.1)、长牡蛎(*Crassostrea gigas*, *CgDUR3*, XP_019929725.1)、砗磲(*Tridacna squamosa*, *TsDUR3*, MF073181.1), 从 JGI 的 Phytozome (<https://phytozome.jgi.doe.gov/pz/portal.html>) 获得上面所述植物 *DUR3* 基因的编码区及 CDS 序列信息。

1.2 棉花 *DUR3* 基因的鉴定及生物信息学分析

1.2.1 棉花 *DUR3* 基因的鉴定 以拟南芥 *AtDUR3* 基因的蛋白质序列(NCBI 登录号: NP_199351) 为探针,在 Phytozome 陆地棉(*Gossypium hirsutum*) 和雷蒙德氏棉(*Gossypium raimondii*) 基因组 BLAST

搜索获得 E 值小于 e^{-10} 的同源序列,将获得的序列(棉花和拟南芥)利用 InterProScan5 (<http://www.ebi.ac.uk/Tools/pfa/iprscan/>) 在线分析其功能结构域的存在情况,最终确定目标序列。

1.2.2 植物 *DUR3* 基因基本信息的获取 棉花及本研究所需植物 *DUR3* 基因的编码区长度、CDS 序列长度、外显子个数等信息均从 Phytozome 数据库中获得;利用 ProtParam tool (<http://web.expasy.org/protparam/>) 在线分析中所获得植物 *DUR3* 基因氨基酸序列的分子量、等电点等基本信息。

1.2.3 *DUR3* 基因的生物信息学分析 运用 DNAMAN 软件对不同植物 *DUR3* 基因的氨基酸序列进行多重序列比对;采用 MEGA6 软件和 Neighbor-Joining 法对不同物种的 *DUR3* 蛋白质序列构建进化树,校验参数 Bootstrap = 1 000;利用 Lasergene 软件的 MegAlign 分析不同植物 *DUR3* 基因氨基酸序列间的相似性;利用 MEME (<http://meme.sdsc.edu/meme/meme.html>) 鉴定分析植物 *DUR3* 蛋白质序列保守基序;利用 TMHMM (<https://services.healthtech.dtu.dk/service.php?TMHMM-2.0>) 在线预测植物 *DUR3* 蛋白序列的跨膜结构域;采用 GSDS9 (Gene Structure Dispely Server, <http://gsds.cbi.pku.edu.cn/>) (Hu et al., 2015) 在线绘制不同植物 *DUR3* 基因结构图;采用 DnaSPv5 计算获得棉花 *DUR3* 旁系同源基因和直系同源基因的非同义替换(K_a) 和同义替换(K_s), 计算出 K_a/K_s 比值。

2 结果与分析

2.1 棉花 *DUR3* 基因的鉴定及序列基本信息

从陆地棉异源四倍体标准系 TM-1 基因组中共鉴定出 2 个 *DUR3* 同源基因,一个是位于 A03 亚基因组的 *Gohir.A03g179800* 基因,另一个是位于 D08 亚基因组的 *Gohir.D08g001500* 基因,为表述方便,将 *Gohir.A03g179800* 基因命名为 *GhDUR3.1*,将 *Gohir.D08g001500* 基因命名为 *GhDUR3.2*。从雷蒙德氏棉基因组中鉴定出 1 个 *DUR3* 同源基因 *Gorai.005G228700*,命名为 *GrDUR3*。用 InterProScan5 分析这 3 个 *DUR3* 基因的氨基酸序列,它们都属于 Sodium/solute symporter (IPR001734) 蛋白家族下的 Urea active transporter (IPR031155) 家族,都具有 cd11476 (Na^+ /urea-polyamine cotransporter *DUR3*,

and related proteins; solute-binding domain) 这一功能结构域。

从表 1 可以看出,不同植物 *DUR3* 基因的编码区长度存在较大差异,从 *OsDUR3* 基因的 2 358 bp 到 *ZmDUR3* 基因的 5 567 bp,最长和最短相差 3 209 bp。然而,植物 *DUR3* 基因 CDS 序列长度的差别却不大,最短的是 *GhDUR3.1*,长度为 2 034 bp;最长的

是 *BdDUR3*,长度为 2 214 bp,相差仅为 180 bp。植物 *DUR3* 基因外显子的个数分为两类,双子叶植物外显子个数普遍较多,为 9 个或 10 个;单子叶植物外显子个数普遍较少,为 3 个或 4 个。植物 *DUR3* 基因推测的多肽的长度 (677 ~ 737 a. a) 和分子量 (72.905 ~ 77.608 kDa) 都相差不大,等电点只有 *SiDUR3* 的小于 7,其他的都高于 7 而处于碱性范围。

表 1 植物 *DUR3* 基因基本信息

Table 1 Basic information of plant *DUR3* genes

基因名 Gene name	编码区长度 Length of coding region (bp)	CDS 长度 Length of CDS (bp)	外显子个数 Number of exons	推测的多肽 Deduced polypeptide		
				长度 Length (a. a.)	分子量 MW (kDa)	等电点 pI
<i>AtDUR3</i>	4 360	2 115	10	704	75.913	8.72
<i>GmDUR3</i>	3 408	2 145	9	714	77.074	8.92
<i>MtDUR3</i>	4 789	2 136	9	711	77.006	8.75
<i>SiDUR3</i>	3 139	2 133	9	710	76.588	9.00
<i>PtDUR3</i>	2 922	2 124	9	707	76.480	8.21
<i>VvDUR3</i>	4 488	2 133	9	710	76.857	8.76
<i>SbDUR3</i>	4 784	2 190	4	729	77.029	8.19
<i>SiDUR3</i>	4 468	2 196	4	731	77.010	6.89
<i>BdDUR3</i>	2 419	2 214	3	737	77.608	7.81
<i>HvDUR3</i>	2 408	2 169	3	722	76.525	7.79
<i>OsDUR3</i>	2 358	2 166	3	721	76.416	8.67
<i>ZmDUR3</i>	5 567	2 196	4	731	77.425	7.17
<i>GhDUR3.1</i>	2 969	2 034	10	677	72.905	8.49
<i>GhDUR3.2</i>	3 023	2 094	9	697	75.190	8.43
<i>GrDUR3</i>	3 013	2 094	9	697	75.234	8.43

2.2 植物 *DUR3* 蛋白氨基酸序列比对

用 DNAMAN 对选取的拟南芥、水稻、玉米、高粱 (*Sorghum bicolor*)、谷子 (*Setaria italica*)、番茄 (*Solanum lycopersicum*)、葡萄 (*Vitis vinifera*)、大豆 (*Glycine max*)、大麦 (*Hordeum vulgare*)、二穗短柄草 (*Brachypodium distachyon*)、蒺藜苜蓿 (*Medicago truncatula*)、毛果杨 (*Populus trichocarpa*) 等不同层面模式植物和本研究鉴定出棉花的共 15 条 *DUR3* 蛋白氨基酸序列进行多重序列比对,如图 1 所示,这 15 条序列的一致性达到 79.81%,不一致区段主要存在于序列的两端,中间的功能区域部分一致性很高。*BdDUR3* 和 *MtDUR3* 序列一致性最低,为 71.8%;*GhDUR3.2* 和 *GrDUR3* 序列一致性最高,为 99.3%。

2.3 植物 *DUR3* 基因蛋白质保守基序分析

采用 MEME 在线分析获得这 14 个不同植物的 15 个 *DUR3* 蛋白基序信息,这 15 个蛋白序列都包含 3 个不同基序。如图 2 所示,这 3 个基序在氨基酸序列上的分布位置基本一致,出现的顺序都是 3 号基序在前,1 号基序紧随其后,2 号基序在最后面(靠近氨基酸序列的中间位置)。如图 3 所示,1 号基序和 2 号基序都有 50 个氨基酸组成,3 号基序有 33 个氨基酸组成,这 3 个基序的保守性都很高。

2.4 植物 *DUR3* 跨膜域预测

Cao 等 (2009) 研究表明,*DUR3* 属于跨膜蛋白,所以有必要分析不同植物 *DUR3* 蛋白质的跨膜域情况。利用 TMHMM Server v. 2.0 在线对这

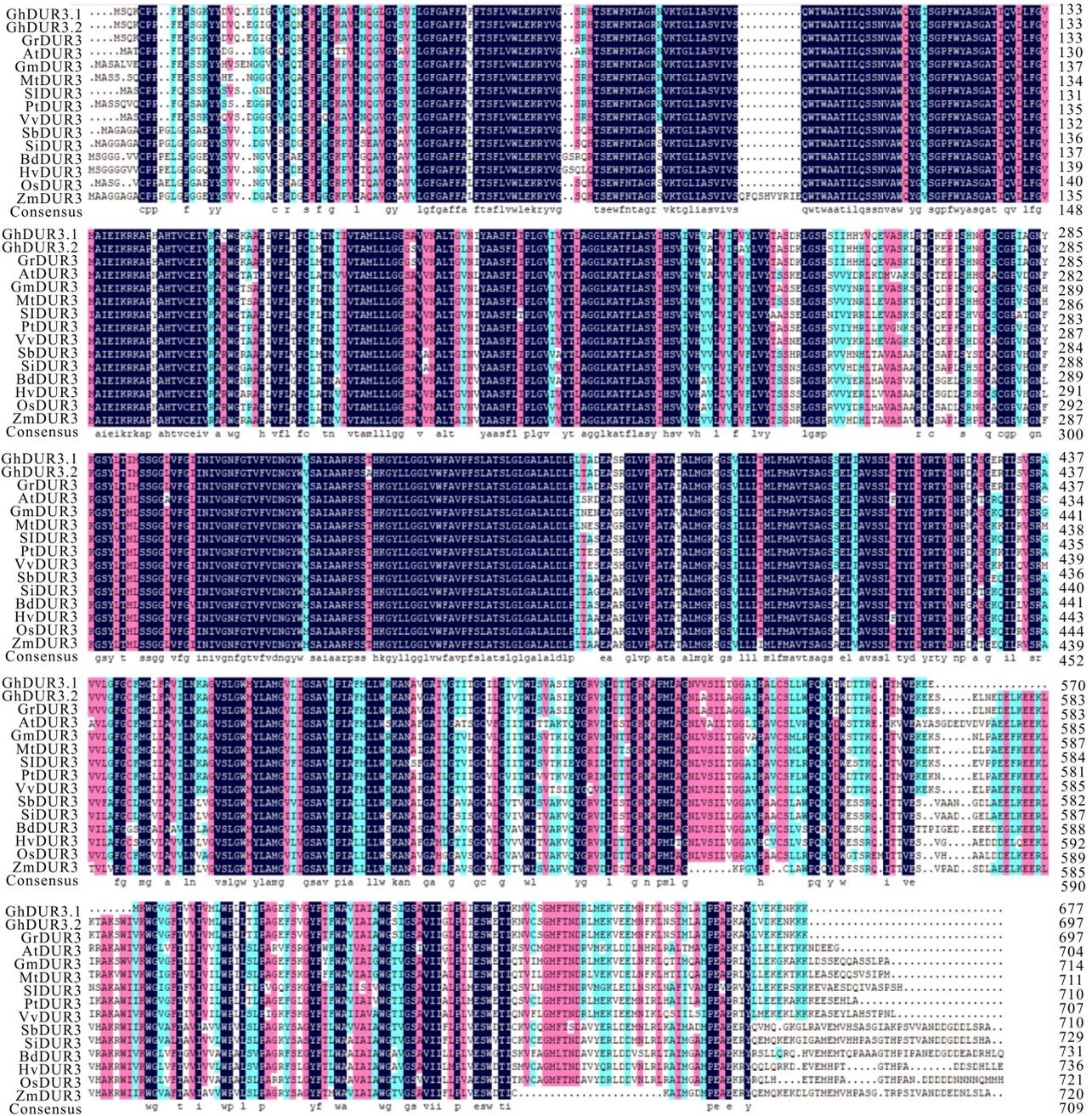


图 1 不同植物 *DUR3* 蛋白氨基酸多重序列比对

Fig. 1 Multiple sequence alignment of *DUR3* protein amino acid in different plants

15 条 *DUR3* 蛋白质的跨膜结构进行预测 (表 2), 15 条 *DUR3* 蛋白质都包含 15 个跨膜域 (图 4), 并且这些跨膜域在氨基酸序列上存在的位置也都非常相似。

2.5 植物 *DUR3* 基因结构分析

根据在 JPC 搜索获得的不同植物 15 个 *DUR3* 基因的编码区序列和 CDS 序列, 用 GSDS2.0 软件

在线绘制 *DUR3* 基因结构图, 以此来分析这 14 种植物 *DUR3* 基因的外显子和内含子情况。表 1 已经对植物 *DUR3* 基因外显子的个数进行了分析, 双子叶植物 *DUR3* 基因外显子个数平均为 9.2 个, 单子叶植物外显子个数平均为 3.5 个, 双子叶植物 *DUR3* 基因外显子的个数要明显多于单子叶植物 (图 5)。从图 5 可以看出, 具有相同外显子个数的基因所对应外

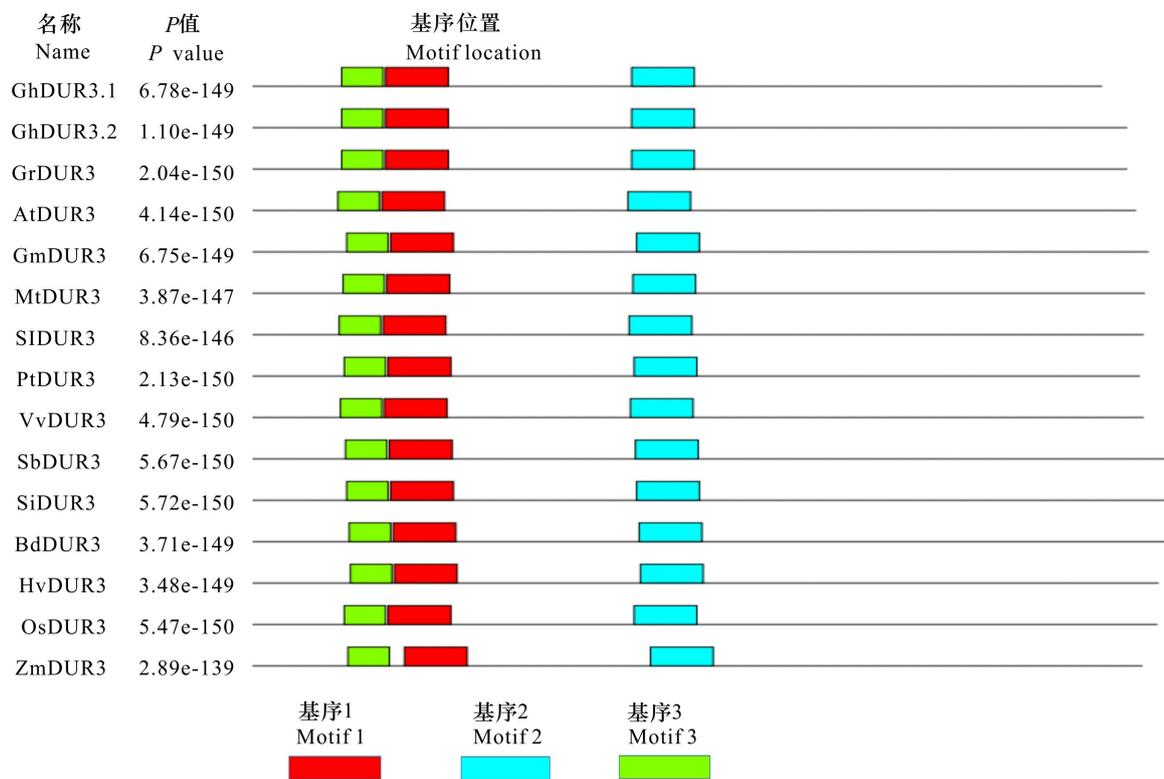


图 2 不同植物 DUR3 蛋白基序分布

Fig. 2 Motif locations of DUR3 proteins in different plants

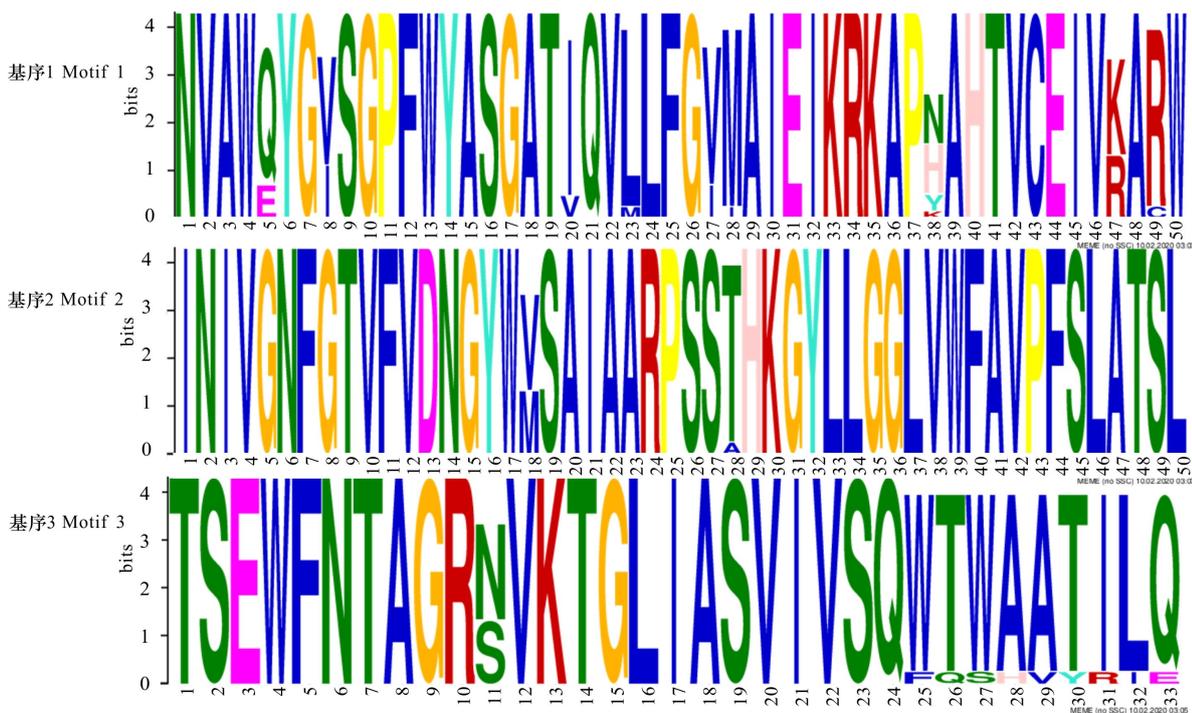


图 3 不同植物 DUR3 蛋白基序序列

Fig. 3 Motif sequences of DUR3 proteins in different plants

表 2 植物 *DUR3* 蛋白跨膜域预测Table 2 Prediction of the transmembrane regions of *DUR3* proteins in plants

蛋白名 Protein name	跨膜位置 Transmembrane position	跨膜数 Transmembrane number
At <i>DUR3</i>	37~59, 80~99, 114~133, 160~182, 187~209, 216~238, 289~311, 332~354, 388~410, 431~450, 460~482, 487~506, 526~548, 595~617, 621~643	15
Gm <i>DUR3</i>	44~66, 87~106, 121~143, 164~186, 196~218, 223~245, 296~318, 339~361, 395~417, 438~456, 466~488, 493~515, 530~552, 593~615, 625~647	15
Mt <i>DUR3</i>	41~63, 84~103, 118~140, 161~183, 193~215, 220~242, 293~315, 336~358, 389~411, 432~454, 459~481, 488~510, 530~552, 589~611, 621~643	15
Sl <i>DUR3</i>	38~60, 81~100, 115~134, 159~181, 191~213, 220~242, 290~312, 333~355, 389~411, 432~450, 460~482, 489~511, 526~548, 586~608, 618~640	15
Pt <i>DUR3</i>	42~64, 85~104, 119~141, 162~184, 194~216, 221~243, 294~316, 337~359, 390~412, 433~455, 460~482, 489~511, 531~553, 590~612, 622~644	15
Vv <i>DUR3</i>	39~61, 82~101, 116~138, 159~181, 191~213, 218~240, 291~313, 334~356, 387~409, 430~452, 457~479, 486~508, 528~550, 589~611, 621~643	15
Sb <i>DUR3</i>	43~65, 86~105, 120~139, 166~188, 193~215, 222~244, 295~317, 338~360, 394~416, 437~459, 463~485, 492~514, 532~554, 594~616, 626~648	15
Si <i>DUR3</i>	43~65, 86~105, 120~142, 162~184, 194~216, 223~245, 296~318, 339~361, 395~417, 438~460, 464~486, 493~515, 533~555, 594~612, 627~649	15
Bd <i>DUR3</i>	43~65, 86~108, 123~145, 166~188, 198~220, 225~247, 298~320, 341~363, 394~416, 437~459, 469~491, 498~520, 535~557, 598~616, 631~653	15
Hv <i>DUR3</i>	44~66, 87~109, 124~146, 167~189, 199~221, 226~248, 299~321, 342~364, 395~417, 438~460, 470~492, 494~516, 536~558, 596~618, 628~650	15
Os <i>DUR3</i>	42~64, 85~104, 119~141, 162~184, 194~216, 221~243, 294~316, 337~359, 393~415, 436~458, 462~484, 491~513, 528~550, 591~613, 623~645	15
Zm <i>DUR3</i>	44~66, 87~106, 121~143, 164~186, 196~218, 225~247, 297~319, 340~362, 393~415, 444~463, 467~489, 494~516, 534~556, 596~618, 628~650	15
Gh <i>DUR3</i> .1	40~62, 83~102, 117~139, 160~182, 192~214, 219~241, 292~314, 335~357, 390~412, 433~452, 462~484, 491~510, 530~552, 573~592, 602~624	15
Gh <i>DUR3</i> .2	40~62, 83~102, 117~139, 160~182, 192~214, 219~241, 292~314, 335~357, 388~410, 431~453, 458~480, 487~509, 529~551, 589~611, 621~643	15
Gr <i>DUR3</i>	40~62, 83~102, 117~139, 160~182, 192~214, 219~241, 292~314, 335~357, 388~410, 431~453, 458~480, 487~509, 529~551, 589~611, 621~643	15

显子的长度也基本一致, 如 *GhDUR3. 2*、*GmDUR3*、*GrDUR3*、*MtDUR3*、*PtDUR3*、*SlDUR3*、*VvDUR3* 等具有 9 个外显子的基因, *BdDUR3*、*HvDUR3*、*OsDUR3* 这 3 个具有 3 个外显子的基因, *SbDUR3*、*SiDUR3*、*ZmDUR3* 这 3 个具有 4 个外显子的基因。

2.6 *DUR3* 基因系统发育分析

为了分析 *DUR3* 基因在不同物种间的进化关系, 选取了植物、真菌、藻类、软体动物等 4 类共 20 个 *DUR3* 蛋白多肽序列通过 MEGA6 软件构建进化树。结果(图 6)表明, 所有的植物聚集在一个分支, 两个真菌 [酿酒酵母 (*Saccharomyces cerevisiae*, *ScDUR3*) 和构巢曲霉 (*Aspergillus nidulans*, *AnDUR3*)] 聚在一个分支, 两个软体动物 [长牡蛎 (*Crassostrea gigas*, *CgDUR3*) 和砗磲 (*Tridacna squamosa*, *TsDUR3*)] 聚在一个分支, 藻

类 [条斑紫菜 (*Pyropia yezoensis*, *PyDUR3*)] 单独一个分支, 所有的单子叶植物和双子叶植物又分别聚集在两个不同的分支。

2.7 植物 *DUR3* 同源基因非同义突变率与同义突变率的比值分析

先利用 MEGA6 的 ClustalW 对所有植物的 *DUR3* 同源基因的 CDS 序列进行比对, 然后用 DnaSPv5 计算获得旁系同源基因和直系同源基因的非同义替换 (K_a) 和同义替换 (K_s), 计算出 K_a/K_s 比值。从表 3 可以看出, 除了 *GhDUR3. 2* 和 *GrDUR3* 以外, 其他的 *DUR3* 直系同源基因间的 K_a/K_s 比值均大于 1, 表明 *DUR3* 基因在不同植物间的进化过程中受到正向选择的作用。陆地棉 *DUR3* 两个旁系同源基因的 K_a/K_s 比值为 4.12, 也大于 1, 表明陆地棉 *DUR3* 基因在自身内的进化过程中同样受到正向选择的作用。

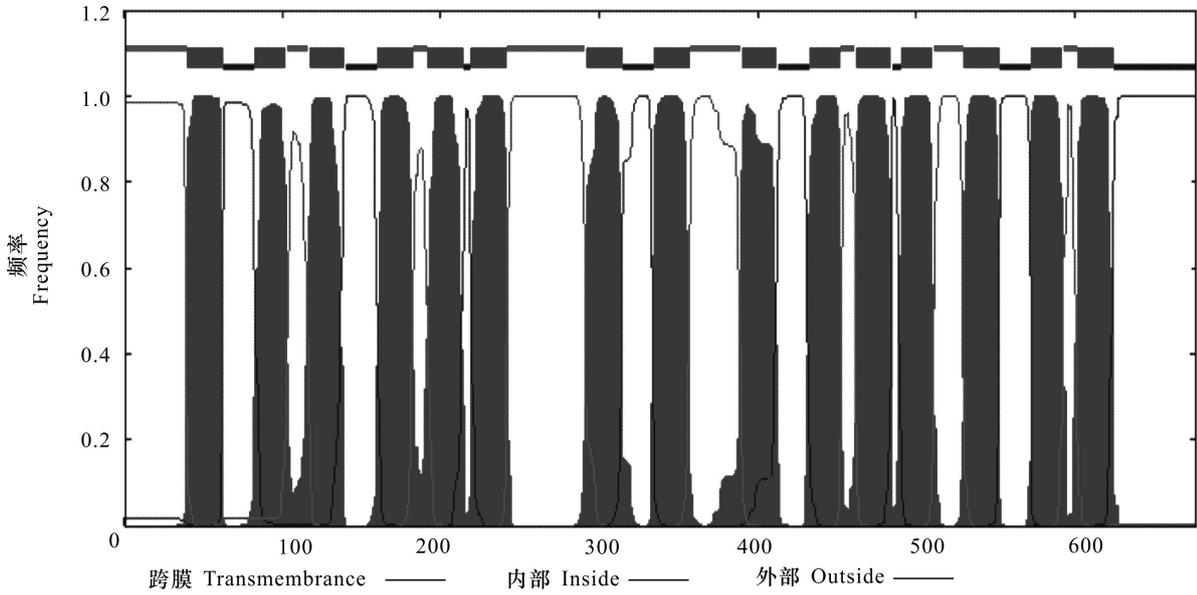


图 4 陆地棉 GhDUR3.1 蛋白跨膜结构域预测分析

Fig.4 Transmembrane structure prediction analysis of the GhDUR3.1 protein in *Gossypium hirsutum*

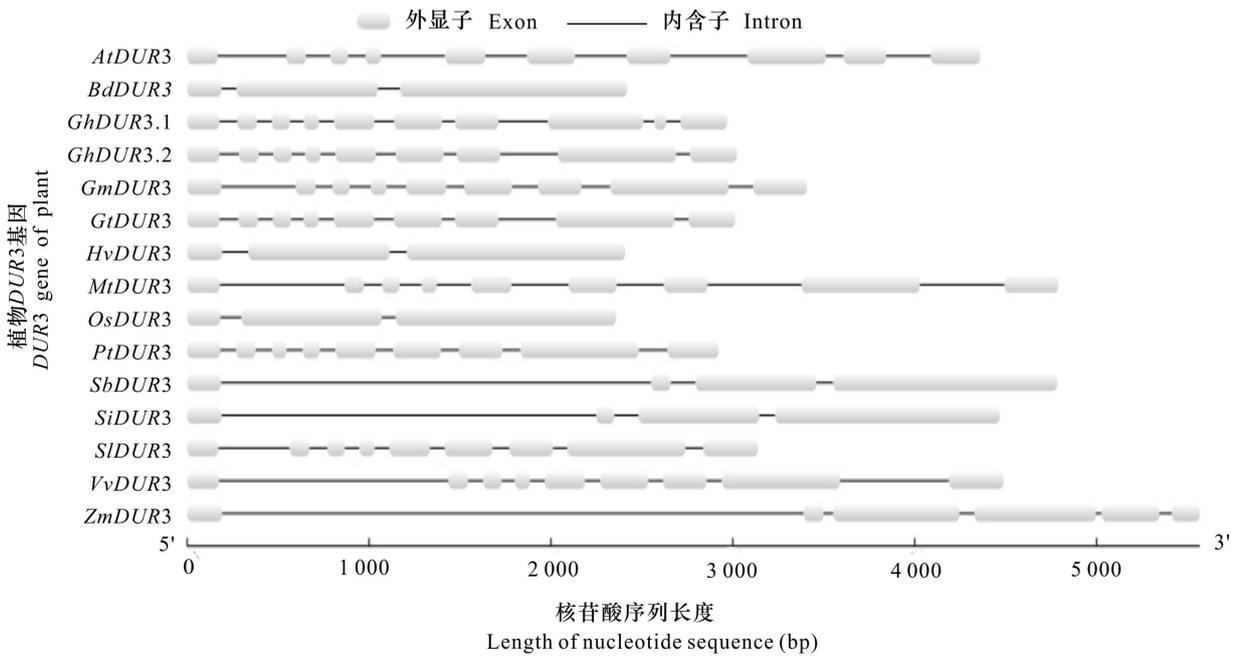


图 5 不同植物 DUR3 基因外显子和内含子结构

Fig. 5 Exon and intron structures of *DUR3* genes in different plants

3 讨论与结论

在二倍体基因组植物中, *DUR3* 同源基因以单一的基因模式 (singleton) 存在 (曹凤秋等, 2009)。

本研究结果也证实, 文中所述植物除了陆地棉是异源四倍体 (AADD), 其他植物皆属于二倍体, 它们的 *DUR3* 基因在各自的基因组中都是单一的基因模式。因为陆地棉是异源四倍体 (AADD), 其基因组有 A 亚组和 D 亚组组成, 所以鉴定出的两

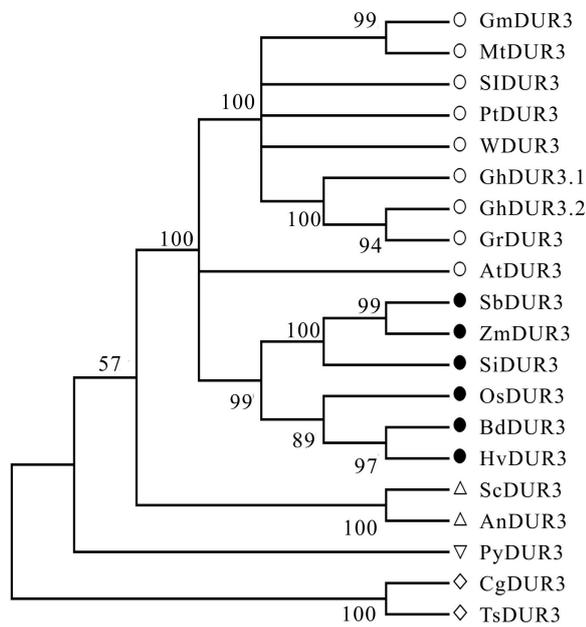
图 6 不同物种 *DUR3* 氨基酸序列进化树分析

Fig. 6 Phylogenetic tree analysis of *DUR3* proteins in different species

GhDUR3.1 和 *GhDUR3.2* 分属于 A 亚组和 D 亚组的基因。雷蒙德氏棉是 D 亚组二倍体棉花,其基因组的 *GrDUR3* 基因与同位于 D 亚组的 *GhDUR3.2* 基因的 CDS 序列和多肽序列相似度均高于 99%,这两个基因的基本信息除了编码区序列长度和氨基酸多肽分子量存在微小差异以外,其他都完全一致,就连多肽序列预测的跨膜结构域位置也完全一致,可见这两个基因在分子生理功能上也必将高度吻合。

整体而言,本文所列举的植物 *DUR3* 氨基酸序列的一致性接近 80%,对这些蛋白序列的等电点、跨膜结构域、基序等的分析都表明植物 *DUR3* 具有比较高的保守性。植物 *DUR3* 多肽序列只有 *SiDUR3* 的等电点略低于 7,为 6.89,其他的等电点基本上都高于 7,其中 *SiDUR3* 的等电点最高,达到 9,表明植物 *DUR3* 通常情况下属于碱性蛋白(表 1)。植物 *DUR3* 都具有 3 个基序,这 3 个基序高度保守,而且在氨基酸序列的排布上也高度一致。植物 *DUR3* 的跨膜结构域数量相同,都包含 15 个跨膜结构域,并且这些跨膜结构出现在氨基酸序

表 3 植物 *DUR3* 直系同源基因的 K_a/K_s 比值Table 3 K_a/K_s ratio of orthologous genes from plant

直系同源基因 Orthologous gene	K_a	K_s	K_a/K_s	直系同源基因 Orthologous gene	K_a	K_s	K_a/K_s
<i>AtDUR3</i>	0.426	0.166	2.566	<i>AtDUR3</i>	0.426	0.176	2.420
<i>GmDUR3</i>	0.295	0.110	2.682	<i>GmDUR3</i>	0.298	0.105	2.838
<i>MtDUR3</i>	0.328	0.101	3.248	<i>MtDUR3</i>	0.330	0.097	3.402
<i>SiDUR3</i>	0.346	0.118	2.932	<i>SiDUR3</i>	0.349	0.125	2.792
<i>PtDUR3</i>	0.273	0.105	2.600	<i>PtDUR3</i>	0.271	0.110	2.464
<i>VvDUR3</i>	0.305	0.109	2.798	<i>VvDUR3</i>	0.302	0.117	2.581
<i>GhDUR3.1</i>				<i>GhDUR3.2</i>			
<i>SbDUR3</i>	0.565	0.163	3.466	<i>SbDUR3</i>	0.576	0.166	3.470
<i>SiDUR3</i>	0.559	0.162	3.451	<i>SiDUR3</i>	0.570	0.163	3.497
<i>BdDUR3</i>	0.593	0.179	3.313	<i>BdDUR3</i>	0.614	0.181	3.392
<i>HvDUR3</i>	0.556	0.169	3.290	<i>HvDUR3</i>	0.571	0.168	3.399
<i>OsDUR3</i>	0.574	0.177	3.243	<i>OsDUR3</i>	0.585	0.179	3.268
<i>ZmDUR3</i>	0.559	0.160	3.494	<i>ZmDUR3</i>	0.575	0.162	3.549
<i>GrDUR3</i>	0.020	0.003	7.960	<i>GrDUR3</i>	0.002	0.003	0.800

注: K_a . 非同义替换; K_s . 同义替换。

Note: K_a . Non-synonymous substitution; K_s . Synonymous substitution.

个列的位置也相似。陆地棉 *GhDUR3.1* 和 *GhDUR3.2*、雷蒙德氏棉 *GrDUR3* 都符合上面的共性。

在进化关系上,不同物种 *DUR3* 氨基酸序列进化树的结果表明,这些基因还是根据物种间种属亲缘关系的远近而进行了聚类,植物的 *DUR3* 氨

核酸序列聚集在一起,真菌的聚集在一起,软体动物的聚集在一起,藻类单独在一个分支,植物又根据单子叶植物纲和双子叶植物纲又分属各自的分支(图6)。棉花属于双子叶植物,所以它聚在双子叶植物的大分支上,其中棉花的3个DUR3亲缘关系最近聚在一个分支。

*DUR3*直系同源基因和旁系同源基因的 K_a/K_s 比值普遍均大于1,说明这些基因在进化过程中主要受到正向选择的作用。此外,植物*DUR3*的基因结构具有一定的规律可以遵循,*DUR3*基因的外显子个数的分析结果表明,双子叶植物的外显子个数普遍较多,单子叶植物的外显子个数普遍较少。本研究结果为研究植物*DUR3*基因的结构进化提供了理论参考。

参考文献:

- BEIER MP, FUJITA T, SASAKI K, et al., 2019. The urea transporter DUR3 contributes to rice production under nitrogen-deficient and field conditions [J]. *Physiol Plantarum*, 167(1): 75-89.
- BOHNER A, KOJIMA S, HAJIREZAEI M, et al., 2015. Urea retranslocation from senescing *Arabidopsis* leaves is promoted by DUR3-mediated urea retrieval from leaf apoplast [J]. *Plant J*, 81(3): 377-387.
- CAO FQ, LIU GW, WANG WH, et al., 2009. Molecular processes of urea metabolism and transport in higher plants [J]. *Chin Bull Bot*, 44(3): 273-282. [曹凤秋, 刘国伟, 王伟红, 等, 2009. 高等植物尿素代谢及转运的分子机理 [J]. *植物学报*, 44(3): 273-282.]
- GROPPA MD, BENAVIDES MP, 2008. Polyamines and abiotic stress: recent advances [J]. *Amino Acids*, 34(1): 35-45.
- HU B, JIN JP, GUO AY, et al., 2015. GSDS 2.0: an upgraded gene feature visualization server [J]. *Bioinformatics*, 31(8): 1296-1297.
- KOJIMA S, BOHNER A, GASSERT B, et al., 2007. AtDUR3 represents the major transporter for high-affinity urea transport across the plasma membrane of nitrogen-deficient *Arabidopsis* roots [J]. *Plant J*, 52(1): 30-40.
- KOJIMA S, BOHNER A, NICOLAUS VON WIRÉN. 2006. Molecular mechanisms of urea transport in plants [J]. *J Membr Biol*, 212(2): 83-91.
- KUMAR R, CHADHA S, SARASWAT D, et al., 2011. Histatin 5 uptake by *Candida albicans* utilizes polyamine transporters Dur3 and Dur31 proteins [J]. *J Biol Chem*, 286(51): 43748-43758.
- LI FG, FAN GY, WANG KB, et al., 2014. Genome sequence of the cultivated cotton *Gossypium arboreum* [J]. *Nat Genet*, 46(6): 562-574.
- LI FG, FAN ZY, LU GH, et al., 2015. Genome sequence of cultivated Upland cotton (*Gossypium hirsutum* TM-1) provides insights into genome evolution [J]. *Nat Biotechnol*, 33(5): 524-530.
- LIU GW, SUN AL, LI DQ, et al., 2015. Molecular identification and functional analysis of a maize (*Zea mays*) DUR3 homolog that transports urea with high affinity [J]. *Planta*, 241(4): 861-874.
- LIU LH, LUDEWIG U, WOLF B FROMMER, et al., 2003a. AtDUR3 encodes a new type of high-affinity urea/H⁺ symporter in *Arabidopsis* [J]. *Plant Cell*, 15(3): 790-800.
- LIU LH, LUDEWIG U, GASSERT B, et al., 2003b. Urea transport by nitrogen-regulated tonoplast intrinsic proteins in *Arabidopsis* [J]. *Plant Physiol*, 133(3): 1220-1228.
- MARSCHNER H, 1995. Mineral nutrition of higher plants [M]. London: Academic Press: 889.
- UEMURA T, KASHIWAGI K, IGARASHI K, 2006. Polyamine uptake by DUR3 and SAM3 in *Saccharomyces cerevisiae* [J]. *J Biol Chem*, 282(10): 7733-7741.
- WANG KB, WANG ZW, LI FG, et al., 2012a. The draft genome of a diploid cotton *Gossypium raimondii* [J]. *Nat Genet*, 44(10): 1098-1103.
- WANG WH, KÖHLER B, CAO FQ, et al., 2012b. Rice DUR3 mediates high-affinity urea transport and plays an effective role in improvement of urea acquisition and utilization when expressed in *Arabidopsis* [J]. *New Phytol*, 193(2): 432-444.
- ZAMIN L, TOMASI N, WIRDNAM C, et al., 2014. Isolation and functional characterization of a high affinity urea transporter from roots of *Zea mays* [J]. *BMC Plant Biol*, 14(1): 222.
- ZHANG TZ, HU Y, JIANG WK, et al., 2015. Sequencing of allotetraploid cotton (*Gossypium hirsutum* L. acc. TM-1) provides a resource for fiber improvement [J]. *Nat Biotechnol*, 33(5): 531-537.

(责任编辑 李 莉)