

研究报告

Research Report

水稻 NRL 基因家族的全基因组鉴定与生物信息学分析

胡阳阳 孙琳琳 王多祥 张大兵 蔡文国*

上海交通大学生命科学技术学院, 上海, 200240

* 通信作者, wenguo.cai@sju.edu.cn

摘要 NRL (NPH3/RPT2-Like)基因家族是植物中特异存在的基因家族,根据拟南芥 NRL 基因家族中参与植物向光性响应的 NPH3 (NONPHOTOTROPIC HYPOCOTYL3)和 RPT2 (ROOT PHOTOTROPISM2)命名。本研究利用生物信息学方法鉴定到 27 个水稻 NRL 基因,分布在除 10 号染色体外的染色体上,同时基因复制分析结果表明,全基因组复制或片段复制比串联复制在水稻 NRL 基因家族扩张中起更大的作用。利用单、双子叶植物的七个物种进行的系统进化关系分析,将 192 个 NRL 基因家族成员分为 6 个不同的组。水稻 NRL 家族的结构特征分析显示,多数成员含有 NRL 基因家族的特征结构域,而基于 RNA-Seq 和微阵列数据的表达分析结果表明,水稻 NRL 家族基因在多个组织表达,同时部分基因在叶片、花序和花药中高表达。本研究对水稻 NRL 基因家族进行鉴定和分析,为进一步研究水稻 NRL 基因家族的功能提供了一定的参考和依据。

关键词 水稻, NRL, 基因家族, 生物信息学

Genome-wide Identification and Bioinformatics Analysis of NRL Gene Family in Rice

Hu Yangyang Sun Linlin Wang Duoxiang Zhang Dabing Cai Wenguo*

School of Life Sciences and Biotechnology, Shanghai Jiaotong University, Shanghai, 200240

* Corresponding author, wenguo.cai@sju.edu.cn

DOI: 10.5376/mpb.cn.2020.18.0034

Abstract NRL (NPH3/RPT2-Like) gene family is plant-specific and it is named after its two members in Arabidopsis, NPH3 (NONPHOTOTROPIC HYPOCOTYL 3) and RPT2 (ROOT PHOTOTROPISM 2), which are involved in phototropism. In this study, we identified 27 rice NRL genes distributed on 12 chromosomes except chromosome 10 with bioinformatic analysis. Gene duplication analysis revealed that whole genome duplication (WGD)/segmental duplication played a more vital role than tandem duplication in rice NRL gene family expansions. A total of 192 NRL gene family members from seven species were identified and subsequently they were divided into six groups in phylogenetic relationship analysis. Structural feature analysis showed that most members of rice NRL gene family contained specific domains of NRL family. Furthermore, based on RNA-Seq data and microarray data, expression analysis showed that rice NRL family genes were expressed in several tissues and some genes displayed higher expression in leaves, inflorescence and anther. Identification and analysis were conducted in NRL gene family from rice and these results provided some reference-able message for further functional research.

Keywords Rice, NRL, Gene family, Bioinformatics

本文首次发表在《分子植物育种》上, 现依据版权所有人授权的许可协议, 采用 Creative Commons Attribution License, 协议对其进行授权, 再次发表与传播

收稿日期: 2020 年 9 月 1 日; 接受日期: 2020 年 9 月 4 日; 发表日期: 2020 年 9 月 4 日

引用格式: 胡阳阳, 孙琳琳, 王多祥, 张大兵, 蔡文国, 2020, 水稻 NRL 基因家族的全基因组鉴定与生物信息学分析, 分子植物育种(网络版), 18(34): 1-12 (doi: 10.5376/mpb.cn.2020.18.0034) (Hu Y.Y., Sun L.L., Wang D.X., Zhang D.B., and Cai W.G., 2020, Genome-wide identification and bioinformatics analysis of NRL gene family in rice, Fenzi Zhiwu Yuzhong (Molecular Plant Breeding (online)), 18(34): 1-12 (doi: 10.5376/mpb.cn.2020.18.0034))

植物生长和形态学建成过程中会对外界光信号的波长、强度和方向等方面的变化进行响应应答,其中之一就是植物中存在的向光性,即植物器官产生朝向或远离特定方向光源弯曲的现象(Holland et al., 2009)。拟南芥向光素 PHOT1 (PHOTOTROPIN1)和 PHOT2 (PHOTOTROPIN2)参与调控向光性、叶绿体运动、气孔开放、叶片伸展和定位等生理活动(Christie et al., 2015)。拟南芥 NPH3 (NONPHOTOTROPIC HYPOCOTYL3)和 RPT2 (ROOT PHOTOTROPISM2)在向光性缺陷突变体的分子遗传分析中被鉴定,研究表明 NPH3 和 RPT2 参与向光素介导的植物向光性、叶片伸展和定位调控,同时 RPT2 也在叶绿体运动中发挥作用(Christie et al., 2015, Suetsugu et al., 2016, Christie et al., 2018,)。拟南芥 NPH3 与 RPT2 同属一个在植物中特异存在的 NRL (NPH3/RPH2-Like)家族,其家族成员主要包含三个特征结构域:N 端的 BTB (Broad complex, tramtrack, and bricà brac)结构域,C 端的卷曲螺旋结构域,以及中间的 NPH3 结构域(Motchoulski and Liscum, 1999, Sakai et al., 2000, Liscum et al., 2014)。NPH3 结构域是 NRL 基因家族成员共有的,而部分 NRL 基因家族成员缺少 N 端的 BTB 结构域或 C 端的卷曲螺旋结构域(Pedmale et al., 2010)。在 NRL 基因家族三个特征结构域的功能研究报告中,N 端的 BTB 结构域参与 E3 泛素连接酶 CRL3NPH3 介导的 phot1 的泛素化,C 端的卷曲结构参与 NPH3 与 phot1 间的蛋白互作,而 NPH3 结构域的功能尚未有研究报告 (Motchoulski and Liscum, 1999, Roberts et al., 2011)。

拟南芥中含有包括 NPH3、RPT2 在内的 33 个 NRL 基因家族成员,但功能并未局限在以 NPH3 和 RPT2 为代表的参与植物向光性响应等生理活动的范围内(Pedmale et al., 2010)。*AtNRL31* (*AtNCH1/AtSR1P1*) 参与不依赖生长素的叶绿体聚集运动,同时也是植物免疫中的正调控因子(Zhang et al., 2014, Suetsugu et al., 2016)。*AtNRL6* (*AtNYP2*)、*AtNRL7* (*AtNYP4*)、*AtNRL20* (*AtMAB4/ENP/NPY1*)、*AtNRL21* (*AtNYP5*) 和 *AtNRL30* (*AtNYP3*) 在与 AGC 激酶 (cAMP 依赖的蛋白激酶 A, cGMP 依赖的蛋白激酶 G 和磷脂依赖的蛋白激酶 C) 协同调控器官发生和根的向重力性应答中的生长素运输方面存在功能冗余 (Cheng et al., 2008, Li et al., 2011)。*AtNRL8* (*AtSETH6*) 参与调控花粉萌发和花粉管生长,而 *AtNRL23* (*AtD-OT3*) 与叶脉结构调节图案形成、维管系统发育以及生殖发育有关(Lalanne et al., 2004, Petricka et al.,

2008)。对水稻中 *AtNPH3* 的同源基因 *CPT1* (COL-EOPTILE PHOTOTROPISM 1)研究表明,CPT1 介导的胚芽鞘向光性建成是通过生长素横向运输和随后的生长再分布实现的(Haga et al., 2005)。水稻 *WIN1* (*WINDING 1*)基因过表达后,幼苗出现螺旋表型和生长素不均匀分布,外源施加生长素极性运输抑制剂不影响螺旋表型,同时过表达 *WIN1* 不影响向光性建成(Cheng et al., 2017)。番茄栽培种 M82 中属于 NRL 基因家族的 *SIBTB5* 在根和花中高表达并受外界环境胁迫表达下调,而马铃薯中 NRL 基因家族成员 *StNRL1* 则是马铃薯晚疫病引起的植物免疫中的敏感因子,参与植物免疫调控(Yang et al., 2016, Li et al., 2018)。可见,NRL 基因家族成员广泛地参与植物生长发育和逆境响应。然而,对该基因家族成员尤其是单子叶植物中该基因家族成员的功能研究还有待深入。

目前对于 NRL 基因家族研究还比较少,仅在拟南芥中有相关报道,而水稻中 NRL 基因家族的全基因组分析未见相关报道。本研究筛选获得水稻 NRL 基因家族成员,利用生物信息学方法进行分析,为进一步研究水稻 NRL 基因家族提供参考和依据。

1 结果与分析

1.1 水稻 NRL 基因家族的鉴定

在本研究中,利用本地 BLAST 和 HMMER 分析方法,获得 30 个水稻 NRL 家族基因,并通过 NCBI 在线工具 Batch CD-Search (<https://www.ncbi.nlm.nih.gov/cdd/>), SMART (<http://smart.embl-heidelberg.de/>)以及 Pfam (<http://pfam.xfam.org/>)对获得的预测的水稻 NRL 家族基因进行筛选,鉴定到 27 个水稻 NRL 家族基因,并命名为 *OsNRL1*~*OsNRL27* (表 1)。水稻 NRL 家族基因的编码区长度在 924bp 到 2289bp 之间分布,氨基酸序列长度在 307~762 范围内,预测的蛋白质分子量大小在 31.37~80.62 kD 之间,理论等电点大小在 4.66~10.21 之间。水稻 NRL 基因家族成员被预测定位在细胞膜、细胞质、细胞核、质体、过氧化物酶体和内质网等细胞内位置,其中 11 个 NRL 基因家族成员被预测定位在细胞核中,6 个 NRL 基因家族成员被预测定位在细胞膜上。

1.2 水稻 NRL 基因家族的染色体分布和基因复制分析

从水稻基因组数据库 RGAP (Rice genome annotation project) (<http://rice.plantbiology.msu.edu/>) 获得水稻 NRL 家族基因的位置信息,利用 TBtools 软件绘

表 1 水稻 NRL 家族基因的基因信息

Table 1 Information of NRL family genes in rice

基因名	基因 ID	编码区长度	氨基酸序列长度	分子量(kD)	等电点	蛋白定位预测
Gene name	Gene ID	Length of CDS	Amino acid sequence length	MW (kD)	pI	Predicted protein location
<i>OsNRL1</i>	LOC_Os01g08130	924	307	31.37	7.58	质体 Plastid
<i>OsNRL2</i>	LOC_Os01g57230	1 980	659	70.63	5.33	过氧化物酶体 Peroxisome
<i>OsNRL3</i>	LOC_Os02g35970	2 289	762	80.62	7.73	细胞膜 Cytomembrane
<i>OsNRL4</i>	LOC_Os02g38120	1 962	653	69.62	7.04	细胞核 Nucleus
<i>OsNRL5</i>	LOC_Os03g10800	1 818	605	66.66	8.31	细胞核 Nucleus
<i>OsNRL6</i>	LOC_Os03g10880	1 881	626	68.03	8.2	细胞核 Nucleus
<i>OsNRL7</i>	LOC_Os03g22600	2 016	671	71.52	9.22	质体 Plastid
<i>OsNRL8</i>	LOC_Os03g41350	1 905	634	69.53	6.51	细胞核 Nucleus
<i>OsNRL9</i>	LOC_Os03g43990	1 875	624	69.53	7.07	质体 Plastid
<i>OsNRL10</i>	LOC_Os03g52880	2 004	667	72.93	6.17	细胞核 Nucleus
<i>OsNRL11</i>	LOC_Os03g55830	2 052	683	74.63	5.91	细胞核 Nucleus
<i>OsNRL12</i>	LOC_Os04g40100	1 884	627	66.34	9.57	细胞膜 Cytomembrane
<i>OsNRL13</i>	LOC_Os04g54400	1 935	644	70.04	4.66	细胞膜 Cytomembrane
<i>OsNRL14</i>	LOC_Os04g57800	1 701	566	61.74	8.42	细胞膜 Cytomembrane
<i>OsNRL15</i>	LOC_Os05g08530	1 506	501	51.33	10.21	细胞膜 Cytomembrane
<i>OsNRL16</i>	LOC_Os06g08550	1 935	644	70.02	5.97	内质网 Endoplasm
<i>OsNRL17</i>	LOC_Os07g36230	1 707	568	60.01	8.08	细胞膜 Cytomembrane
<i>OsNRL18</i>	LOC_Os07g39530	2 166	721	78.42	6.89	细胞核 Nucleus
<i>OsNRL19</i>	LOC_Os08g03650	1 926	641	71.31	6.38	细胞核 Nucleus
<i>OsNRL20</i>	LOC_Os09g09370	1 764	587	63.96	7.55	细胞质 Cytoplasm
<i>OsNRL21</i>	LOC_Os09g25330	1 857	618	68.33	8.29	细胞核 Nucleus
<i>OsNRL22</i>	LOC_Os11g02610	1 716	571	60.54	8.21	细胞核 Nucleus

基因名	基因 ID	编码区长度	氨基酸序列长度	分子量(kD)	等电点	蛋白定位预测
Gene name	Gene ID	Length of CDS	Amino acid sequence length	MW (kD)	pI	Predicted protein location
<i>OsNRL23</i>	LOC_Os11g02620	1 932	643	70.89	6.51	质体 Plastid
<i>OsNRL24</i>	LOC_Os12g02530	1 716	571	60.68	8.47	细胞核 Nucleus
<i>OsNRL25</i>	LOC_Os12g02540	1 956	651	71.85	6.74	质体 Plastid
<i>OsNRL26</i>	LOC_Os12g39380	1 917	638	69.63	5.26	内质网 Endoplasm
<i>OsNRL27</i>	LOC_Os12g41910	1 953	650	72.09	5.85	细胞质 - 细胞核 Cytoplasm-nucleus

制水稻 NRL 家族基因的染色体分布图(图 1)。27 个水稻 NRL 家族基因不均匀地分布在除第 10 号染色体外的水稻染色体上,其中 3 号染色体上分布的 NRL 基因最多,共有 7 个基因定位在 3 号染色体上,而 5、6 和 8 号染色体上分别只有一个 NRL 基因分布。

利用 MCScanX 软件对水稻串联复制事件和全基因组复制或片段复制事件进行分析,共有 7 个基因复制事件发生在水稻 NRL 基因家族中,包含 2 个串联复制事件和 5 个全基因组复制或片段复制事件(图 1;表 2)。基因对 *OsNRL22~OsNRL23* 和 *OsNRL24~OsNRL25* 是串联复制,而基因对 *OsNRL1~OsNRL15*、*OsNRL4~OsNRL12*、*OsNRL22~OsNRL24*、*OsNRL26~OsNRL8* 和 *OsNRL27~OsNRL9* 是全基因组复制或片段复制。使用 DnaSP 6.0 软件计算基因复制对的 Ka/Ks 值以评估水稻 NRL 基因进化的驱动力。Ka/Ks >1 即正选择, Ka/Ks <1 即纯化选择, Ka/Ks =1 即中性选择。水稻 NRL 家族基因中存在基因复制事件的基因对的 Ka/Ks 值在 0.107 0~0.546 5 之间,表明它们在进化过程中受到纯化选择(表 2)。

1.3 水稻 NRL 基因家族系统进化关系分析

为进一步了解水稻 NRL 基因家族,从在线数据库 Phytozome (<https://phytozome.jgi.doe.gov/pz/portal.html>) 获得双子叶植物拟南芥、葡萄、番茄,单子叶植物玉米、大麦、二穗短柄草蛋白序列,分别鉴定到 33、24、27、31、26 和 24 个 NRL 基因家族成员,并进行系统进化关系分析(图 2)。结果显示,七个物种中 NRL 基因家族共有 192 个成员,可以分为六组(Group I~VI),分别有 8、23、42、39、54 和 26 个成员。Group VI 中的 *OsNRL3 (OsCPT1)* 与 *AtNPH3* 是同源基因,与之前的

报道一致(Haga et al., 2005)。在 Group VI 中,玉米和番茄中分别有两个与 *OsNRL3 (OsCPT1)* 和 *AtNPH3* 进化关系相近的基因,即 *Zm00001d004440*、*Zm00001d016898*、*Solyc01g105680* 和 *Solyc10g047530*,而大麦、二穗短柄草和葡萄分别只有 *HORVU6Hr1G015480*、*Bradi3g46480* 和 *VIT_203s0038g00270* 是 *OsNRL3 (OsCPT1)* 和 *AtNPH3* 的同源基因(图 2)。在 Group IV 中, *OsNRL22*、*OsNRL24* 以及 *Zm00001d005051*、*Zm00001d023314* 是 *AtRPT2* 的同源基因,而 *AtRPT2* 在大麦、二穗短柄草、番茄和葡萄中的同源基因分别是 *HORVU4Hr1G023260*、*Bradi4g25900*、*Solyc07g043130* 和 *VIT_206s0004g08230*。Group V 中含有 54 个 NRL 基因家族成员,是 NRL 基因家族成员最多的分组,而拟南芥和水稻均有 8 个 NRL 基因家族成员在该组中,其中 *AtNRL6*、*AtNRL7*、*AtNRL20*、*AtNRL21* 和 *AtNRL30* 存在功能冗余(Cheng et al., 2008, Li et al., 2011)。

1.4 水稻 NRL 家族基因结构和保守基序分析

对水稻 NRL 家族基因结构的分析结果显示,17 个水稻 NRL 家族基因包含 4 个外显子和 3 个内含子,如 *OsNRL18*、*OsNRL7*、*OsNRL11* 和 *OsNRL20*, 5 个水稻 NRL 家族基因包含 3 个外显子和 2 个内含子,即 *OsNRL10*、*OsNRL14*、*OsNRL17*、*OsNRL22* 和 *OsNRL24*, 2 个水稻 NRL 家族基因包含 5 个外显子和 4 个内含子,即 *OsNRL3* 和 *OsNRL6*, 而 *OsNRL15* 含有 2 个外显子和 1 个内含子, *OsNRL1* 只包含 1 个外显子, *OsNRL20* 包含的外显子/内含子数目最多,共有 6 个外显子和 5 个内含子(图 3B)。进一步的保守基序分析结果显示,多数水稻 NRL 家族成员包含 14~15 个保守基序(图 3C)。BTB 结构域被预测由 N 端的第 2、

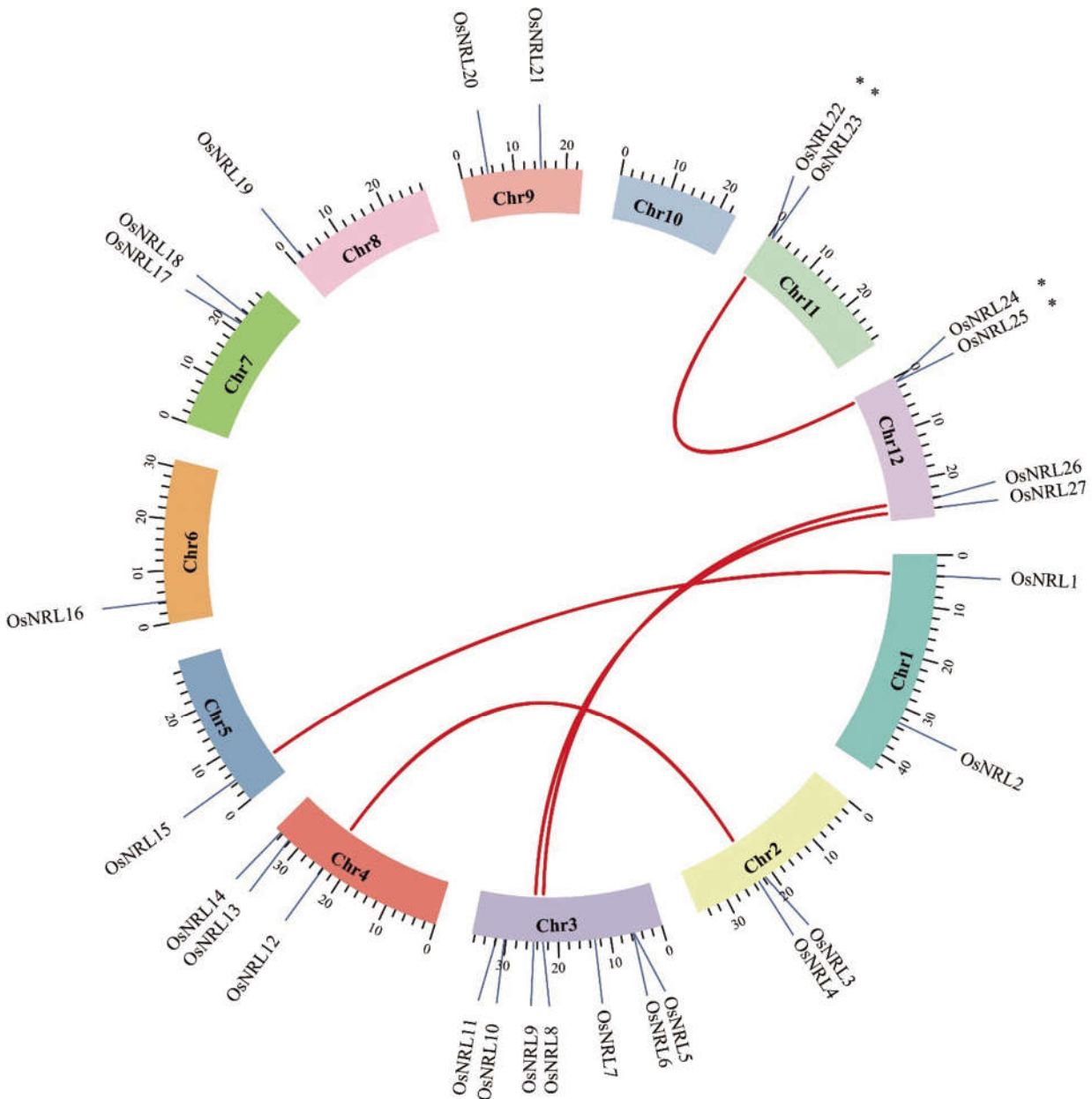


图 1 水稻 NRL 基因的染色体分布和基因复制

注: 刻度标尺以兆碱基对(Mb)划分染色体长度; 全基因组复制或片段复制基因对由红线连接; 串联复制基因对由星号标识

Figure 1 Chromosomal distribution and gene duplications of OsNRLs

Note: The length of chromosome is scaled in Megabase; Whole genome duplication (WGD)/segmental duplication gene pairs are linked by a red line; Tandem duplication gene pairs are marked with asterisks

4 和 6 基序组成, 第 1、3、5、7、10、11 和 14 基序被预测组成 NPH3 结构域, 而 C 端的第 13 基序被预测构成卷曲螺旋结构域。27 个水稻 NRL 家族成员都含有 NPH3 结构域, 2 个缺少 N 端的 BTB 结构域, 10 个缺少 C 端的卷曲螺旋结构域。BTB 结构域和卷曲螺旋结构域均可参与蛋白互作, 缺失任一结构域可能影响与潜在的特定蛋白的相互作用, 而 NPH3 结构域的具体功能还有待进一步研究 (Motchoulski and Liscum, 1999, Inada et al., 2004, Roberts et al., 2011)。

1.5 水稻 NRL 基因家族的表达分析

为了进一步分析水稻 NRL 家族基因在不同组织中的表达特征, 我们从水稻基因组数据库 RGAP (<http://rice.plantbiology.msu.edu/>) 获得水稻十个不同组织的 RNA-Seq 数据, 并对水稻 NRL 家族基因的表达数据作层次聚类分析(图 4)。数据表明, 水稻 NRL 家族基因在不同的水稻组织中有不同程度的表达, 根据聚类分析结果可以分为四组, 第一组到第四组分别有 7、13、4 和 3 个基因(图 4)。第一组中的基因,

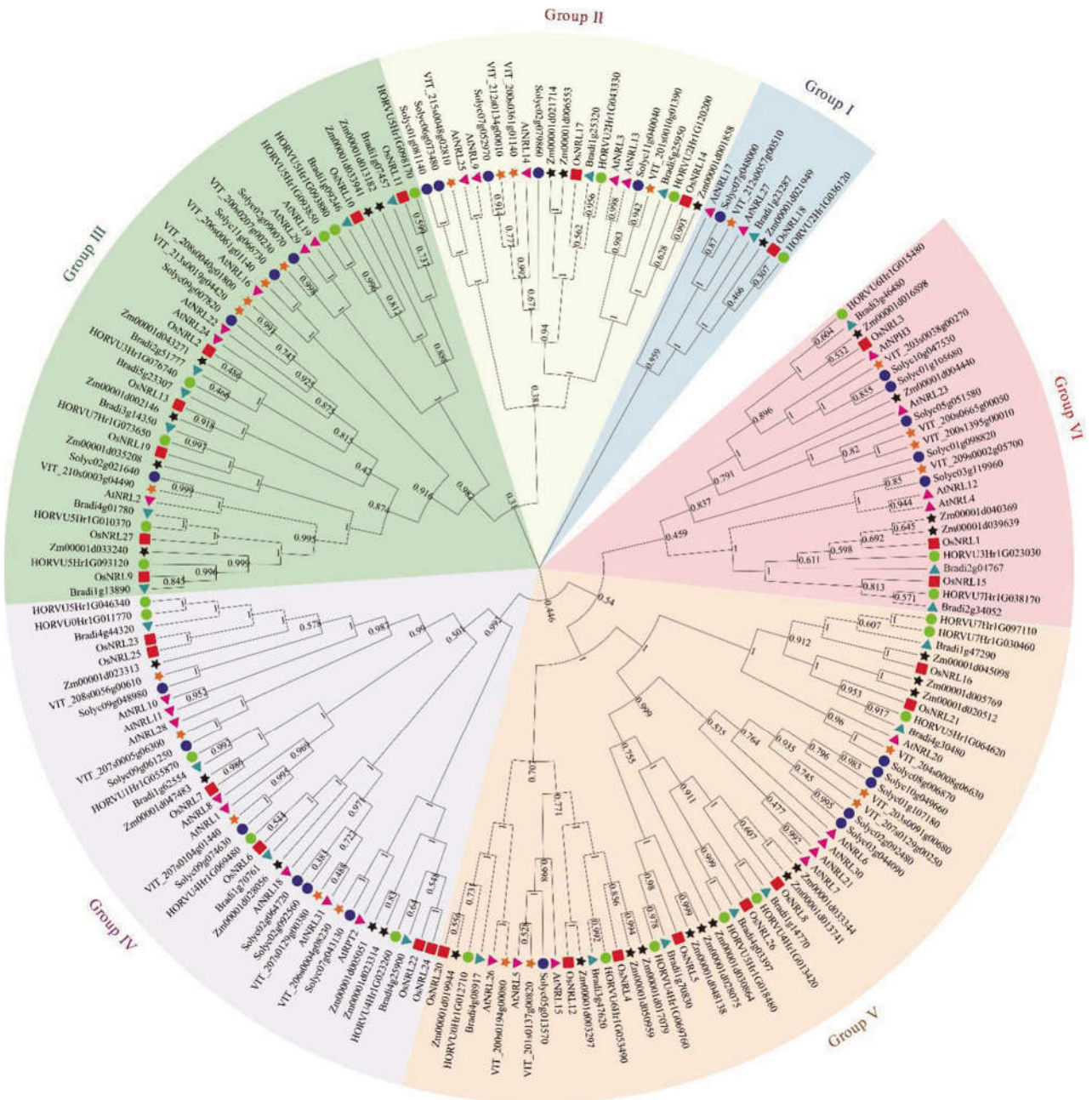


图2 水稻, 玉米, 大麦, 二穗短柄草, 拟南芥, 葡萄和番茄中 NRL 基因家族系统进化关系分析
 注: 利用 MEGA 7.0 使用邻接法(neighbor-joining, NJ)对 NRL 基因家族进行系统进化关系分析, Bootstrap 设置为 1000; 单子叶植物水稻, 玉米, 大麦和二穗短柄草的基因家族成员, 如 *OsNRL1*, *Zm00001d030864*, *HORVU0Hr1G011770* 和 *Bradi1g07457*, 分别以红色正方形, 黑色星形, 绿色圆形和青色三角形标识; 双子叶植物拟南芥, 葡萄和番茄的基因家族成员, 如 *AtNRL1*, *VIT_201s0137g00820* 和 *Solyc01g081140*, 分别以洋红色三角形, 橘黄色星形和蓝色圆形标识
 Figure 2 Phylogenetic relationship of NRL gene family in rice, maize, barley, Brachypodium Arabidopsis, grape and tomato
 Note: Neighbor-joining tree is constructed using MEGA 7.0 with 1000 bootstrap replicates; Rice, maize, barley and Brachypodium are marked with red squares, black stars, green circles and cyan triangles; Arabidopsis, grape and tomato are marked with magenta triangles, orange stars and blue circles

即 *OsNRL7*、*OsNRL8*、*OsNRL14*、*OsNRL16*、*OsNRL17*、*OsNRL19* 和 *OsNRL21* 在抽穗前的花序和雌蕊中高表达, 其中 *OsNRL7* 和 *OsNRL14* 在授粉后 5 天的种子中表达量也较高。第二组中的 13 个水稻 NRL 家族基因在抽穗前的花序高表达, 其中 *OsNRL23*、*OsNRL25*

和 *OsNRL26* 也在叶片有较高的表达量, *OsNRL20* 在抽穗后的花序中也呈现高表达。第三组中的 *OsNRL6*、*OsNRL12*、*OsNRL22* 和 *OsNRL24* 在叶片中高表达, 而第四组中的 *OsNRL2*、*OsNRL13* 和 *OsNRL18* 在抽穗后的花序和花药中高表达。值得注意的

表 2 水稻 NRL 家族复制基因对的 Ka/Ks 值

Table 2 Ka/Ks values for duplication gene pairs of NRL family in rice

基因 I	基因 II	Ka/Ks 值	复制类型	纯化选择
Gene I	Gene II	Value of Ka/Ks	Duplication type	Purifying selection
<i>OsNRL22</i>	<i>OsNRL23</i>	0.424 0	串联复制 Tandem duplication	是 Yes
<i>OsNRL24</i>	<i>OsNRL25</i>	0.460 2	串联复制 Tandem duplication	是 Yes
<i>OsNRL1</i>	<i>OsNRL15</i>	0.546 5	全基因组复制或片段复制 WGD or segmental duplication	是 Yes
<i>OsNRL4</i>	<i>OsNRL12</i>	0.504 2	全基因组复制或片段复制 WGD or segmental duplication	是 Yes
<i>OsNRL22</i>	<i>OsNRL24</i>	0.218 9	全基因组复制或片段复制 WGD or segmental duplication	是 Yes
<i>OsNRL26</i>	<i>OsNRL8</i>	0.107 0	全基因组复制或片段复制 WGD or segmental duplication	是 Yes
<i>OsNRL27</i>	<i>OsNRL9</i>	0.127 4	全基因组复制或片段复制 WGD or segmental duplication	是 Yes

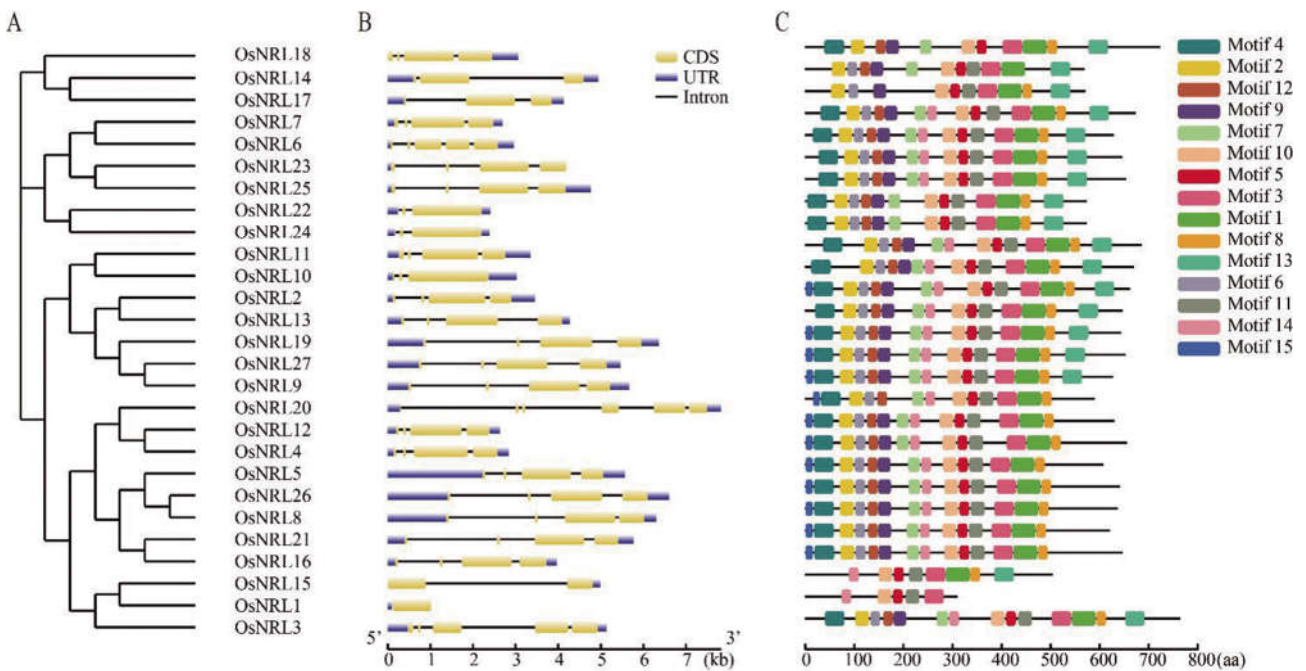


图 3 水稻 NRL 家族进化树, 基因结构和保守基序分析

注: A: 水稻 NRL 家族进化树; B: 水稻 NRL 家族基因结构分析; C: 水稻 NRL 家族保守基序分析

Figure 3 Phylogenetic tree, exon - intron structures and conserved motif analysis of NRL gene family in rice

Note: A: Phylogenetic tree of NRL family in rice; B: exon-intron structures analysis of NRL family in rice; C: Conserved motif analysis of NRL family in rice

是, 一些复制基因, 例如 *OsNRL22* 和 *OsNRL24*, 被分在同一个组中, 暗示一些全基因组复制或片段复制基因因有相同或相似的表达调控序列而具有相似的表达调控模式。因而, 进一步对复制基因的表达进行全局的分析。

利用微阵列数据对水稻 NRL 基因家族的全基因组复制或片段复制基因对和串联复制基因对的表达模式进行分析, 其中 *OsNRL1* 和 *OsNRL15*、*OsNRL27* 和 *OsNRL9* 两组基因对因缺少部分表达数据未纳入分析(图 5)。*OsNRL22* 和 *OsNRL23*、*OsNRL24* 和 *Os-*

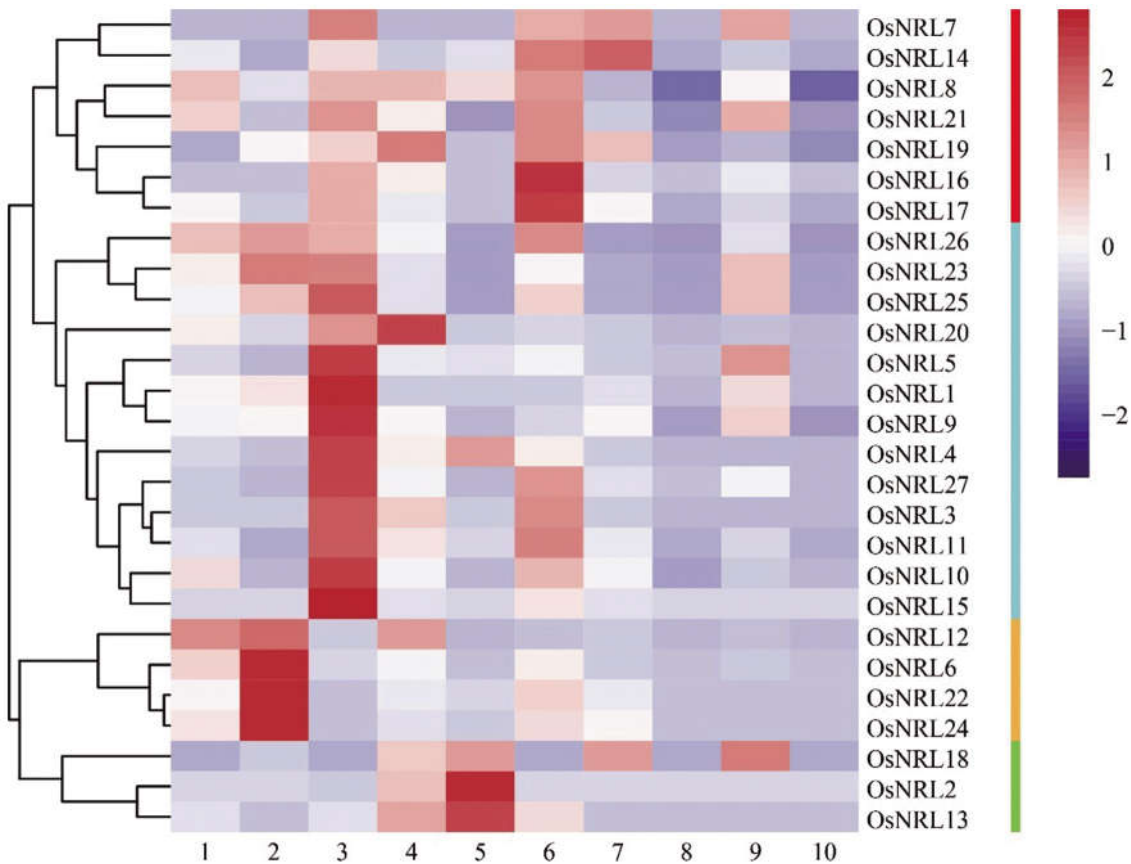


图4 水稻 NRL 基因在不同组织和生长发育阶段的表达

注: 1: 幼苗; 2: 叶片; 3: 抽穗前的花序; 4: 抽穗后的花序; 5: 花药; 6: 雌蕊; 7: 授粉后 5 d 的种子; 8: 授粉后 10 d 的种子; 9: 授粉后 25 d 的胚; 10: 授粉后 25 d 的胚乳

Figure 4 Expression of OsNRLs in different tissues and developmental stages

Note: 1: Shoots; 2: Leaves-20 days; 3: Pre-emergence inflorescence; 4: Post-emergence inflorescence; 5: Anther; 6: Pistil; 7: Seed-5 days after pollination; 8: Seed-10 days after pollination; 9: Embryo-25 days after pollination; 10: Endosperm-25 days after pollination

NRL25 两组串联复制基因呈现不同的表达模式, *OsNRL4* 和 *OsNRL12*、*OsNRL22* 和 *OsNRL24*、*OsNRL26* 和 *OsNRL8* 三组全基因组复制或片段复制基因分别呈现相似的表达模式。

2 讨论

NRL 基因家族是植物中特有的基因家族, 拟南芥 33 个基因家族成员中, 包括 *NPH3* 和 *RPT2* 在内共有 10 个成员的功能已经鉴定(Christie et al., 2018)。目前的研究观点认为, NRL 基因家族参与调控 AGC 激酶 VIII 亚家族成员(如向光素 *phot1* 和 *phot2*)介导的生长素依赖和非生长素依赖的生理反应(Suetsugu et al., 2016, Christie et al., 2018)。水稻中仅有两个 NRL 基因家族成员 *CPT1* 和 *WIN1* 被鉴定, *CPT1* 已被证明参与胚芽鞘向光性响应和生长素侧向运输, 而 *WIN1* 与生长素运输有关, 但不参与胚芽鞘和根的向性响应(Haga et al., 2005, Cheng et al., 2017)。

本研究利用生物信息学方法, 从水稻基因组中鉴定到 27 个 NRL 家族基因, 与先前的报道相比增加了两个 NRL 家族基因, 即 *OsNRL1* 和 *OsNRL15* (Gingerich et al., 2007)。与其它水稻 NRL 家族基因相比, *OsNRL1* 和 *OsNRL15* 的基因外显子/内含子结构数量较少, 同时两者编码的氨基酸序列长度也较短, 暗示它们可能是假基因, 但与一般假基因表达量低于同源基因的表达量不同, *OsNRL1* 和 *OsNRL15* 在多个组织中的表达量并未低于它们的同源基因 *OsNRL3* (*OsCPT1*) 的表达量(Zou et al., 2009)。因此, 判断 *OsNRL1* 和 *OsNRL15* 是否为假基因, 需要与今后的基因具体功能研究的结果相结合。

本研究中的 27 个水稻 NRL 家族基因不均匀地分布在 12 条染色体上, 并且存在一些复制基因对, 即 5 个全基因组复制或片段复制基因对和 2 个串联复制基因对。基因复制是进化的驱动力来源之一, 参与新基因的形成和功能分化, 而复制基因间可能存

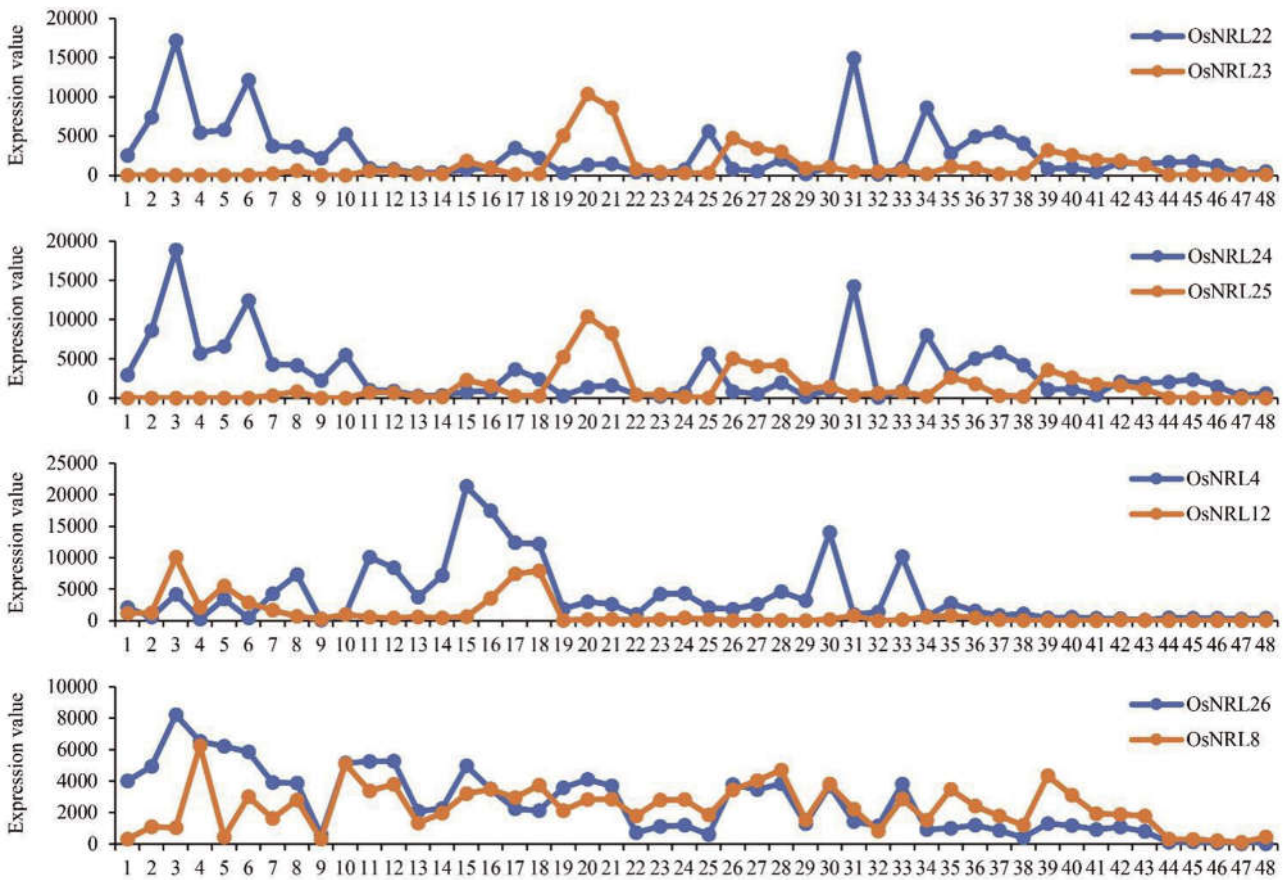


图 5 水稻 NRL 家族复制基因对的表达模式

注: 横轴的数字代表水稻生长阶段, 纵轴的 Expression value (表达值) 从微阵列数据获得 (GSE21396); 1~6: 叶片; 7~10: 叶鞘; 11~14: 根; 15~18: 茎; 19~21: 花序; 22~25: 花药; 26~28: 雌蕊; 29~31: 外稃; 32~34: 内稃; 35~38: 子房; 39~43: 胚; 44~48: 胚乳; 编号对应的水稻生长阶段的具体信息可由微阵列数据获得(GSE21396)

Figure 5 Expression patterns of duplication gene pairs of NRL family in rice

Note: The X-axis represents developmental stages of rice; The Y-axis represents the raw expression values obtained from the microarray data (GSE21396); 1~6: Leaf blade; 7~10: Leaf sheath; 11~14: Root; 15~18: Stem; 19~21: Inflorescence; 22~25: Anther; 26~28: Pistil; 29~31: Lemma; 32~34: Palea; 35~38: Ovary; 39~43: Embryo; 44~48: Endosperm; The matching Information between rice developmental stages and serial numbers obtains from the microarray data (GSE21396)

在功能冗余(Panchy et al., 2016)。已有研究报道, 串联复制和全基因组复制或片段复制在单个基因家族的扩张中起主要作用(Kong et al., 2019)。本研究中水稻 NRL 基因家族中存在 7 个基因复制事件, 表明基因复制在水稻 NRL 基因家族的基因扩张中扮演比较重要的角色, 并且全基因组复制或片段复制相比串联复制发挥更重要的作用。在系统进化关系分析中, 包括水稻在内的七个物种的 NRL 基因家族成员可以分为 6 组(Christie et al., 2018)。AtNPH3 和 AtRPT2 分别在 Group VI 和 Group IV 中, 与 AtNPH3 同组的水稻同源基因 *OsNRL3 (OsCPT1)* 也参与向光性响应, 与 AtRPT2 同组的水稻同源基因 *OsNRL22* 和 *OsNRL24* 未见相关报道(Haga et al., 2005)。AtNPH3 和 AtRPT2 在其它五个物种中分别有 7 个和 6 个同源基因, 而

这些同源基因可能与 *OsNRL3 (OsCPT1)* 功能相似, 也参与向光素介导的植物向光性等生理响应。

本研究发现, 水稻 NRL 家族基因在多个组织中均有表达, 同时部分基因在部分组织中表现为高表达, 如叶片、抽穗前后的花序、花药和雌蕊。*OsNRL3 (OsCPT1)* 是第一个报道的水稻 NRL 家族基因, 突变后使得水稻胚芽鞘向光性接近完全消失(Haga et al., 2005)。*OsNRL3 (OsCPT1)* 在抽穗前的花絮中高表达, 同时在雌蕊中也有较高表达量, 表明 *OsNRL3 (OsCPT1)* 可能也在水稻生殖发育发挥一定的功能。对 *OsNRL7 (OsWIN1)* 而言, 一方面, 其过表达株系 WIN1-OX 的幼苗出现螺旋表型, 同时该株系与 T-DNA 插入突变体株系 WIN1Act 的胚芽鞘均表现为正常的向光性响应, 说明水稻 NRL 家族基因中以

OsNRL7 (*OsWIN1*)为代表的部分基因可能存在功能的特异性,另一方面,两类株系的种子大小较野生型种子小,而 *OsNRL7* (*OsWIN1*)包括种子和胚在内的多个水稻组织均有一定的表达,表明 *OsNRL7* (*OsWIN1*)可能参与水稻种子大小的调控 (Cheng et al., 2017)。结合组织表达分析和进化关系分析,水稻 NRL 家族基因中与 *OsNRL3* (*OsCPT1*)相近的基因,即 *OsNRL1* 和 *OsNRL15*,与 *OsNRL7* (*OsWIN1*)相近的基因,即 *OsNRL6*、*OsNRL7*、*OsNRL22*、*OsNRL23*、*OsNRL24* 和 *OsNRL25*,可能与两者相似,在植物向光性和生长素运输等生理活动中发挥功能。此外,由 RNA-Seq 和微阵列表达数据,*OsNRL4* 和 *OsNRL12*、*OsNRL22* 和 *OsNRL24*、*OsNRL26* 和 *OsNRL8* 三组全基因组复制或片段复制基因组内表达模式相似,尤其是 *OsNRL22* 和 *OsNRL24* 复制基因对,表达模式基本一致,并且两者均在叶片中高表达,表明两者的启动子区的顺式作用元件构成可能较为相似,所发挥的功能可能相近,对它们的研究需要考虑可能存在的功能冗余。

3 材料与方法

3.1 水稻 NRL 基因家族成员鉴定

水稻基因组信息 MSU 7.0 版本通过水稻基因组数据库 RGAP (<http://rice.plantbiology.msu.edu/>) 下载获得。NPH3 结构域的隐马尔科夫模型(hidden markov model, HMM) (PF03000)和 BTB 结构域的隐马尔科夫模型 (PF00651) 从 Pfam (<http://pfam.xfam.org/>) 获取。利用隐马尔科夫模型,使用 HMMER 3.0 软件对水稻蛋白序列进行本地 HMM SEARCH 预测 NRL 家族基因,E 值小于 1×10^{-10} 。同时使用 BLAST 2.9.0 软件进行本地 BLASTP 检索,E 值小于 1×10^{-10} 。结合两者结果,共鉴定到 30 个可能的水稻 NRL 基因家族成员。使用 NCBI 在线工具 Batch CD-Search (<https://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi>), SMART (<http://smart.embl-heidelberg.de/>) 以及 Pfam 对获得的预测的水稻 NRL 家族基因进行筛选确定,去除预测结构域长度小于 NPH3 结构域或 BTB 结构域长度 50%的基因,27 个 NRL 基因被确定属于水稻 NRL 基因家族。利用水稻基因组数据库 RGAP (<http://rice.plantbiology.msu.edu/>) 获得水稻 NRL 家族基因的基因信息 (黄岳等, 2015; 陈旭等, 2017)。利用在线网站 Plant-mSubP (<http://bioinfo.usu.edu/Plant-mSubP/>) 进行蛋白亚细胞定位预测。

3.2 水稻 NRL 基因家族的染色体分布和基因复制分析

通过水稻基因组数据库 RGAP (<http://rice.plantbiology.msu.edu/>) 获取水稻中基因在染色体上的位置信息并筛选出水稻 NRL 基因家族的染色体分布信息。利用 MCScanX 软件对水稻中基因复制事件进行分析并筛选提取水稻 NRL 基因家族中的基因复制对信息,并利用 TBtools 软件绘制水稻 NRL 基因的染色体分布和基因复制示意图(Chen et al., 2020)。使用 DnaSP 6.0 软件计算 Ka/Ks (非同义替换率/同义替换率),进行复制基因对选择压力分析。

3.3 NRL 基因家族系统进化分析

利用通过在线数据库 Phytozome (<https://phytozome.jgi.doe.gov/pz/portal.html>) 获得的双子叶植物(拟南芥,葡萄,番茄)和单子叶植物(玉米,大麦,二穗短柄草)蛋白序列进行 NRL 基因家族成员鉴定,并利用 MEGA 7.0 使用邻接法对 NRL 基因家族成员进行系统进化关系分析,其中模型设置为泊松模型,Bootstrap 设置为 1000,同时使用在线工具 Evolvview v2 (<http://nar.oxfordjournals.org/content/44/W1/W236>) 绘制进化树。使用 TBtools 软件中 IQ-tree 组件,利用最大似然法构建水稻 NRL 基因家族进化树,其中 Bootstrap 设置为 1 000 (Chen et al., 2020)。

3.4 水稻 NRL 家族基因结构、保守基序分析

利用在线网站 CSDS 2.0 (<http://gsds.cbi.pku.edu.cn/>) 对水稻 NRL 基因家族成员进行结构分析,利用在线网站 MEME (<http://meme-suite.org/tools/meme>) 对水稻 NRL 基因家族成员进行保守基序分析,使用默认参数,其中基序数为 15 个,并使用软件 TBtools 绘制保守基序图示意图(Chen et al., 2020)。

3.5 水稻 NRL 基因家族表达分析

从水稻基因组数据库 RGAP (<http://rice.plantbiology.msu.edu/>) 下载获得水稻不同组织表达数据。使用 R 语言 pheatmap 包对水稻 NRL 家族基因表达数据使用 Z-score 方法进行标准化,并进行表达层次聚类分析。利用由 NCBI Gene Expression Omnibus (<https://www.ncbi.nlm.nih.gov/geo/>) 获得的田间整个生长过程中水稻各种组织/器官的时空基因表达微阵列数据(GSE21396),分析水稻 NRL 家族复制基因对的表达模式。

作者贡献

胡阳阳是本研究的实验设计和执行人; 胡阳阳

完成数据分析、论文初稿的写作;孙琳琳和王多祥参与实验设计、试验结果分析,论文写作与修改;张大兵和蔡文国是项目的构思者及负责人,指导实验设计,数据分析,论文写作与修改。全体作者都阅读并同意最终的文本。

致谢

本研究由国家重点研发计划(2016YFD0100804)、江苏省农业科技自主创新资金(JASTIF, CX(18)1001)和中国博士后科学基金(Grant 2017M621451)共同资助。

参考文献

- Chen C.J., Chen H., Zhang Y., Thomas H.R., Frank M.H., He Y. H., and Xia R., 2020, TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data, *Mol. Plant*, 13(8): 1194-1202
- Chen X., Shi L., Zhu L., Shi J.X., and Xu J., 2017, Molecular evolution characteristics and expression pattern analysis of the heat shock protein 70(HSP70) gene superfamily in plant, *Jiyinzuxue Yu Yingyong Shengwuxue (Genomics and Applied Biology)*, 36(10): 4282-4294 (陈旭, 石垒, 朱璐, 石建新, 许杰, 2017, 植物 HSP70 蛋白家族分子进化特征及其表达模式分析, *基因组学与应用生物学*, 36(10): 370-382)
- Cheng M.L., Lo S.F., Hsiao A.S., Hong Y.F., Yu S.M., and Ho T. D., 2017, Ectopic expression of WINDING 1 leads to asymmetrical distribution of auxin and a spiral phenotype in rice, *Plant & Cell Physiology*, 58(9): 1494-1506
- Cheng Y.F., Qin G.J., Dai X.H., and Zhao Y.D., 2008, NPY genes and AGC kinases define two key steps in auxin-mediated organogenesis in Arabidopsis, *Proc. Natl. Acad. Sci. USA*, 105(52): 21017-21022
- Christie J.M., Blackwood L., Petersen J., and Sullivan S., 2015, Plant flavoprotein photoreceptors, *Plant & Cell Physiology*, 56(3): 401-413
- Christie J.M., Suetsugu N., Sullivan S., and Wada M., 2018, Shining light on the function of NPH3/RPT2-Like proteins in phototropin signaling, *Plant Physiology*, 176 (2): 1015-1024
- Gingerich D.J., Hanada K., Shiu S.H., and Vierstra R.D., 2007, Large-scale, lineage-specific expansion of a bric-a-brac/tramtrack/broad complex ubiquitin-ligase gene family in rice, *Plant Cell*, 19(8): 2329-2348
- Haga K., Takano M., Neumann R., and Iino M., 2005, The Rice COLEOPTILE PHOTOTROPISM1 gene encoding an ortholog of Arabidopsis NPH3 is required for phototropism of coleoptiles and lateral translocation of auxin, *Plant Cell*, 17 (1): 103-115
- Holland J.J., Roberts D., and Liscum E., 2009, Understanding phototropism: from Darwin to today, *J. Exp. Bot.*, 60 (7): 1969-1978
- Huang Y., Zhu L., Luo R.H., Zhang D.B., and Xu J., 2015, Evolutionary and expression profile analysis of the non-specific lipid transfer proteins genes superfamily in Arabidopsis thaliana and rice (*Oryza sativa*), *Jiyinzuxue Yu Yingyong Shengwuxue (Genomics and Applied Biology)*, 34 (7): 1510-1521 (黄岳, 朱璐, 罗人和, 张大兵, 许杰, 2015, 拟南芥和水稻非特异脂质转运蛋白的分子进化、表达模式以及在花药发育中的功能分析, *基因组学与应用生物学*, 34 (7): 1510-1521)
- Inada S., Ohgishi M., Mayama T., Okada K., and Sakai T., 2004, RPT2 is a signal transducer involved in phototropic response and stomatal opening by association with phototropin 1 in Arabidopsis thaliana, *Plant Cell*, 16(4): 887-896
- Kong W.L., An B.G., Zhang Y., Yang J., Li S.M., Sun T., and Li Y.S., 2019, Sugar transporter proteins (STPs) in gramineae crops: comparative analysis, phylogeny, evolution, and expression profiling, *Cells*, 8(6): 560
- Lalanne E., Michaelidis C., Moore J.M., Gagliano W., Johnson A., Patel R., Howden R., Vielle-Calzada J.P., Grossniklaus U., and Twell D., 2004, Analysis of transposon insertion mutants highlights the diversity of mechanisms underlying male progamic development in Arabidopsis, *Genetics*, 167 (4): 1975-1986
- Li J.H., Su X.X., Wang Y.L., Yang W., Pan Y., Su C.G., and Zhang X.G., 2018, Genome-wide identification and expression analysis of the BTB domain-containing protein gene family in tomato, *Genes & Genomics*, 40(1): 1-15
- Li Y.T., Dai X.H., Cheng Y.F., and Zhao Y.D., 2011, NPY genes play an essential role in root gravitropic responses in Arabidopsis, *Mol. Plant*, 4(1): 171-179
- Liscum E., Askinosie S.K., Leuchtman D.L., Morrow J., Willenburg K.T., and Coats D.R., 2014, Phototropism: growing towards an understanding of plant movement, *Plant Cell*, 26 (1): 38-55
- Motchoulski A., and Liscum E., 1999, Arabidopsis NPH3: A NPH1 photoreceptor-interacting protein essential for phototropism, *Science*, 286(5441): 961-964
- Panchy N., Lehti-Shiu M., and Shiu S.H., 2016, Evolution of gene duplication in plants, *Plant Physiology*, 171 (4): 2294-2316
- Pedmale U.V., Celaya R.B., and Liscum E., 2010, Phototropism: mechanism and outcomes, *Arabidopsis Book*, 8: e0125
- Petricka J.J., Clay N.K., and Nelson T.M., 2008, Vein patterning screens and the defectively organized tributaries mutants in Arabidopsis thaliana, *Plant Journal*, 56(2): 251-263
- Roberts D., Pedmale U.V., Morrow J., Sachdev S., Lechner E., Tang X.B., Zheng N., Hannink M., Genschik P., and Liscum

- E., 2011, Modulation of phototropic responsiveness in Arabidopsis through ubiquitination of phototropin 1 by the CUL3-Ring E3 ubiquitin ligase CRL3NPH3, *Plant Cell*, 23 (10): 3627-3640
- Sakai T., Wada T., Ishiguro S., and Okada K., 2000, RPT2. A signal transducer of the phototropic response in Arabidopsis, *Plant Cell*, 12(2): 225-236
- Suetsugu N., Takemiya A., Kong S.G., Higa T., Komatsu A., Shimazaki K., Kohchi T., and Wada M., 2016, RPT2/NCH1 subfamily of NPH3-like proteins is essential for the chloroplast accumulation response in land plants, *Proc. Natl. Acad. Sci. USA*, 113(37): 10424-10429
- Yang L., McLellan H., Naqvi S., He Q., Boevink P.C., Armstrong M., Giuliani L.M., Zhang W., Tian Z.D., Zhan J.S., Gilroy E.M., and Birch P.R., 2016, Potato NPH3/RPT2-Like protein StNRL1, targeted by a *Phytophthora infestans* RXLR effector, is a susceptibility factor, *Plant Physiology*, 171(1): 645-657
- Zhang L., Du L.Q., Shen C.J., Yang Y.J., and Poovaiah B.W., 2014, Regulation of plant immunity through ubiquitin-mediated modulation of Ca²⁺-calmodulin-AtSR1/CAMTA3 signaling, *Plant Journal*, 78(2): 269-281
- Zou C., Lehti-Shiu M.D., Thibaud-Nissen F., Prakash T., Buell C. R., and Shiu S.H., 2009, Evolutionary and expression signatures of pseudogenes in Arabidopsis and rice, *Plant Physiology*, 151(1): 3-15