

水稻中结瘤素基因的同源基因研究

王彦章 俞冠翹 朱家璧*

(中国科学院上海生命科学研究院 植物生理生态研究所 植物分子遗传国家重点实验室, 上海 200032; * 通讯联系人, E-mail: jbzhu@iris.sipp.ac.cn)

Homologues of Nodulin Genes in *Oryza sativa*

WANG Yan-zhang, YU Guan-qiao, ZHU Jia-bi*

(National Laboratory of Plant Molecular Genetics, Institute of Plant Physiology & Ecology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200032, China; * Corresponding author, E-mail: jbzhu@iris.sipp.ac.cn)

Abstract: Seventy-five nodulin genes, found in nucleotide databases, were used as probes to scan the rice genome by bioinformatics methods. Thirty-one homologues of nodulin genes were detected in the rice genome which exhibited more than 35% identities in amino acid sequences in comparison with leguminous nodulin genes. This result indicated that leguminous nodulin homologous genes existed extensively in rice genome. In comparison with the leguminous nodulin gene *enod40*, sucrose synthase gene and *Rab* gene, their corresponding homologues in rice showed that they all belonged to the classification of orthologous genes, presumably originated from the common ancestral genes. However, leguminous nodulin genes mutated during the early evolution so as to meet the requirements of nodule organogenesis. The other 44 leguminous nodulin genes, which have no homologs in rice genome, play essential roles in establishing symbiosis between the *Rhizobium* and the host legumes. It was suggested that the lack of these leguminous nodulin genes in rice might result in its disability for nodulation and nitrogen fixation.

Key words: *Oryza sativa*; leguminous plants; nodulin genes; orthologous gene

摘 要: 以核酸数据库中检索到的 75 个豆科植物结瘤素基因作为探针,应用生物信息学方法对水稻基因组进行扫描分析。在水稻基因组中发现有 31 个与结瘤素基因具有同源性的基因,与相应的结瘤素基因比对,它们的氨基酸序列一致性至少在 35% 以上。这表明在水稻基因组中广泛存在豆科植物结瘤素基因的同源基因。豆科植物结瘤素基因 *enod40*、蔗糖合成酶基因和 *Rab* 基因与水稻中对应的同源基因比较分析表明,它们属于直向同源基因,可能来自于共同的祖先基因,在长期进化过程中豆科植物结瘤素基因受根瘤器官形成的需求而发生变异。然而,另有 44 个豆科结瘤素基因在水稻基因组中未显示有同源性基因的存在,这些结瘤素基因在豆科植物与根瘤菌建立共生过程中起着重要的作用。推测可能是由于水稻中缺少了这些豆科结瘤素基因,导致水稻不能结瘤固氮。

关键词: 水稻; 豆科植物; 结瘤素基因; 直向同源基因

中图分类号: Q943; S330; S511.03; S52.03

文献标识码: A

文章编号: 1001-7216(2005)03-0202-07

根瘤菌在氮源缺乏的条件下能够诱导豆科植物根部产生独特的器官——根瘤。根瘤的形成成为共生于根瘤中的根瘤菌(类菌体)创建了合适的生态位,类菌体将空气中的 N_2 还原为 NH_3 供宿主植物利用,形成了自然界中效率最高的共生固氮作用。豆科植物根瘤的形成需要共生伙伴间的特异性基因在时间和空间上的表达调控。根瘤菌的结瘤基因,负责根瘤菌信号分子——结瘤因子的合成和运输;宿主植物中特异性诱导表达的结瘤素基因,负责根瘤器官的发生和功能。在共生关系的建立过程中,根瘤菌启动植物基因的专一表达,促进侵染及辅助形成固氮根瘤的位置,而宿主植物提供根瘤发育的遗传信息。

然而,这种共生作用具有严格的宿主专一性,能够与根瘤菌建立共生关系的植物主要限于豆科植物,水稻等非豆科的重要农作物不能与根瘤菌形成

共生作用。Reddy 等^[1]利用豆科植物早期结瘤素基因作为探针,与不同水稻品种的基因组进行杂交,发现水稻中存在早期结瘤素基因 *enod2*、*enod5*、*enod12*、*enod14*、*enod40*、*enod55*、*enod70* 和 *enod93* 的同源基因。我们在核酸数据库中(www.ncbi.nlm.nih.gov/nucleotide)检索到豆科植物的 75 个结瘤素基因。本研究利用这 75 个豆科植物结瘤素基因作为探针,通过生物信息学的方法对水稻整个基因组进行扫描分析,在水稻基因组中找到了许多与豆科植物结瘤素基因具有高度同源性的基因,为水稻固氮潜能的研究提供了有价值的数据和信息。

收稿日期: 2004-06-01; 修改稿收到日期: 2004-11-05。

基金项目: 国家 973 重点基础研究发展规划项目(2001CB108901); 上海生命科学院基金资助项目。

第一作者简介: 王彦章(1969—),女,在读博士研究生。

1 材料与方法

用 BLAST 分析序列和进行数据库搜索 (http://www.ncbi.nlm.nih.gov/BLAST)。采用 Clustal W(version 1.82) 软件进行多序列比较分析。

2 结果与分析

2.1 核酸数据库中检索到的豆科植物结瘤素基因

宿主豆科植物中特异性诱导表达的结瘤素基因,是根瘤形成及功能维持的关键因素。根据诱导表达时间的差异,结瘤素基因可分为早期结瘤素基因和晚期结瘤素基因两类^[2]。早期结瘤素基因是在根瘤发育的早期和固氮作用开始前表达的宿主基因;而晚期结瘤素基因是指在固氮作用开始后表达的宿主基因,此时根瘤结构已经形成。因此,早期结瘤素基因参与根瘤菌的侵染和根瘤发育过程,而晚期结瘤素基因主要负责根瘤的功能。我们在 GenBank 数据库中进行广泛的搜索,检索到 75 个有明确定义的豆科植物结瘤素基因(表 1 和表 3),包括早期结瘤素基因和晚期结瘤素基因。到目前为止除少数结瘤素基因通过遗传学方法确定了其功能外,大多数结瘤素基因的功能仍不清楚。

2.2 水稻基因组中存在结瘤素基因的同源基因

将检索到的 75 个豆科植物结瘤素基因的序列作为查询序列,在非冗余的 GenBank CDS translations、PDB、SwissProt、PIR 和 PRF 的水稻数据库中进行广泛的 BLAST 分析。经过筛选,在水稻基因组中发现有 31 个与结瘤素基因具有同源性的基因,其氨基酸序列比对的一致性达到 35% 以上(表 1)。这 31 个豆科植物结瘤素基因在结瘤过程中承担着重要的功能,其中某些基因突变后可导致明显的结瘤突变表型。我们从表 1 中选取豆科结瘤素基因 *enod40*、蔗糖合成酶基因和 *Rab* 基因与水稻中相应的同源基因进行比较分析,探讨这些基因在水稻与豆科植物中表达和功能方面的异同,以期通过这些基因找到水稻和豆科植物的特定联系,这对于探索水稻结瘤固氮潜能具有重要的意义。

2.2.1 早期结瘤素基因 *enod40* 与水稻中同源基因 *Osenod40* 的比较分析

我们以大豆早期结瘤素基因 *Gmenod40* (GenBank 数据库登录号 X69155) 作为查询序列,在水稻数据库中找到具有高度同源性的基因 *Osenod40* (GenBank 数据库登录号 AB024054)。将 *Osenod40*

与大豆早期结瘤素基因 *enod40* 进行比对,两者的氨基酸序列一致性达到 65%。将水稻 *OsENOD40* 与其他几种豆科植物的 *ENOD40* 进行多序列比较,结果显示 *OsENOD40* 与不同豆科植物中的 *ENOD40* 具有很高的同源性,其中酪氨酸、组氨酸、甘氨酸、丝氨酸、精氨酸和谷氨酸等序列是完全保守的(图 1)。

根据有关报道将豆科植物早期结瘤素基因 *enod40* 与水稻中对应的同源基因 *Osenod40* 功能进行比较分析,发现两者的功能存在一定的相似性,都在特定组织器官中专一表达。Kouchi 等^[3]曾报道豆科植物 *enod40* 在根瘤器官形成的起始过程中承担重要的功能。Yang 等^[4]的研究则表明 *enod40* 也参与根瘤后期发育和共生功能的形成。Kouchi 等^[5]后来从水稻中分离了 *Osenod40* 基因,发现此基因仅在水稻茎秆的维管束中表达,进一步研究表明转基因大豆根瘤中 *Osenod40-GUS* 的表达模式与大豆中 *Gmenod40-GUS* 的表达模式相似,主要在维管束中表达。因此,根据图 1 的多序列比较分析和综合上述已报道的研究结果,我们推测不同豆科植

MsENOD40	MKLLCWQKSIHGSMANRQVTKRQWIPFWSL 30
MtENOD40	MKLLCWEKSIHGSMANRQVTKRQWIPFWSL 30
PsENOD40	MKFLCWQKSIHGSMANRQVTKRQWIPFWSL 30
SrENOD40	MK-LCWQKSIHGSLANRQVTKRQWTPGLVL 29
PvENOD40	MK-FCWQASIHGSQLANRQVTKRQWTPSGSL 29
GmENOD40a	ME-LCWQTSIHGSLANRQVTKRQWTPGLSL 29
GmENOD40b	ME-LCWLTTHGSWQTGKSQKRQWTPGLSL 29
OsENOD40	ME-DEWLEAHGSAPNRQVTKRQW----- 23
	*. * *** . : :***

图 1 豆科植物早期结瘤素基因 *enod40* 与水稻中对应同源基因 *enod40* 的多序列比较分析

Fig. 1. Multiple sequence alignment of leguminous early nodulin gene *enod40* and the corresponding homolog *enod40* in rice.

植物种类以及相应基因的登录号: Ms, *Medicago sativa* (X80263); Ps, *Pisum sativum* (X81064); Gm, *Glycine max* (a, D13504, X69155; b, D13504, X69154); Mt, *Medicago truncatula* (X80264); Sr, *Sesbania rostrata* (Y12714); Pv, *Phaseolus vulgaris* (X86441); Os, *Oryza sativa* (AB024054)。利用 Clustal W (version 1.82) 进行多序列比对。序列保守性用以下符号表示:“*”表示氨基酸残基相同;“:”表示保守的氨基酸;“.”表示部分保守的氨基酸。

Plant species and accession numbers of the corresponding genes: Ms, *Medicago sativa* (X80263); Ps, *Pisum sativum* (X81064); Gm, *Glycine max* (a, D13504, X69155; b, D13504, X69154); Mt, *Medicago truncatula* (X80264); Sr, *Sesbania rostrata* (Y12714); Pv, *Phaseolus vulgaris* (X86441); Os, *Oryza sativa* (AB024054). The Clustal W (version 1.82) was used for multiple sequence alignment. Sequence conservation is shown in the alignment according to the following keys: “*”, Invariant; “:”, Conserved; “.”, Partially conserved residue.

表 1 水稻基因组中存在的结瘤素基因的同源基因

Table 1. Homologous genes of the nodulin genes in rice genome.

结瘤素基因 Nodulin gene	GenBank 登录号 ¹⁾ Accession No.	植物来源 Plant species	与水稻中同源基因的氨基酸一致性 IR/%	根瘤发育中的作用或表型 FR	水稻中对应同源基因编码蛋白的 GenBank 登录号 ¹⁾ AR ¹⁾	水稻中对应同源基因功能的描述 DR
<i>Rab1p</i>	L14929	<i>Glycine max</i>	78	GTP 结合蛋白, 参与类菌体膜的生物发生	CAC39050	GTP 结合蛋白
SS	AF030231	<i>Glycine max</i>	75	蔗糖合成酶, 豌豆 <i>rug4</i> 突变导致固氮作用失效	S23543	蔗糖合成酶
<i>nodulin-26</i>	L12258	<i>Glycine max</i>	75	可能的通道蛋白	BAB63833	可能的 RAB7A 蛋白
<i>Rab7</i>	L14930	<i>Glycine max</i>	73	GTP 结合蛋白, 参与类菌体膜的生物发生	BAB93352	GTP 结合蛋白
<i>CCS2</i>	AF134835	<i>Medicago truncatula</i>	67	有丝分裂抑制因子	AAN74839	可能的细胞周期开关蛋白
<i>GmENOD93</i>	AB018378	<i>Glycine max</i>	66	早期结瘤素	BAA83559	OsENOD93a
<i>enod40a</i>	X69155	<i>Glycine max</i>	65	小肽, 在根瘤器官形成过程中起作用	AB024054	在维管束分化和功能方面承担作用
<i>nodulin-35</i>	M63743	<i>Glycine max</i>	64	根瘤内特异尿酸酶	BAB63819	可能的尿酸酶
<i>LjNOD16</i>	AF367434	<i>Lotus japonicus</i>	62	磷脂酰肌醇转移蛋白 IV	AAL83356	可能的磷脂酰肌醇磷脂酰乙醇胺转移酶
<i>GmN61</i>	AF434718	<i>Glycine max</i>	62	晚期结瘤素	AAP54087	可能的谷氨酸合成酶
<i>nod33</i>	AJ518838	<i>Phaseolus vulgaris</i>	58	可能的磷酸酶	AAL83638	可能的糖饥饿诱导蛋白
<i>SYMRK</i>	AF491998	<i>Lotus japonicus</i>	57	受体型激酶, 突变体的结瘤信号转导受阻, 根毛不发生卷曲, 仅顶端膨大	BAC83202	可能的结瘤因子受体激酶
<i>GS2</i>	AF207688	<i>Glycine soja</i>	56	腺苷三磷酸双磷酸酶	BAC83798	可能的核酸三磷酸酶
<i>MtN6</i>	AJ133118	<i>Medicago truncatula</i>	55	与前侵染线形成有关的蛋白	AAP54087	可能的谷氨酸合成酶
<i>enod18</i>	AJ271816	<i>Vicia faba</i>	53	细菌 MJ0577 家族相关的蛋白	BAC78561	未知功能的蛋白
<i>HAR1</i>	AB092810	<i>Lotus japonicus</i>	52	超级结瘤突变体	AAL79717	可能的受体蛋白激酶
<i>ripl</i>	U16727	<i>Medicago truncatula</i>	52	过氧化物酶	BAC79532	可能的过氧化物酶前体
<i>ENOD8</i>	AY166655	<i>Medicago sativa</i>	51	早期结瘤素	BAC80100	可能的 ENOD8 前体
<i>NARK</i>	AY166655	<i>Glycine max</i>	51	受体激酶, 突变体为超级结瘤突变体	AAL79717	可能的受体蛋白激酶
<i>Nfr1</i>	AJ575249	<i>Lotus japonicus</i>	51	可能的结瘤因子受体	BAD01244	类似 PERK1 的受体蛋白激酶
<i>NlJ70</i>	AF031243	<i>Lotus japonicus</i>	47	根瘤中特异性表达的蛋白	AAP68484	未知功能的蛋白
<i>Srglb</i>	X13505	<i>Sesbania rostrata</i>	45	豆血红蛋白	Q94FT7	非共生的血红蛋白
<i>SAT1</i>	AF069738	<i>Glycine max</i>	44	共生过程中氮运输因子	AAP03394	可能的氮运输因子
<i>LjNPP2C1</i>	AF092431	<i>Lotus japonicus</i>	42	根瘤中增强表达的蛋白磷酸酶 2C	BAB89568	可能的蛋白磷酸酶 2C
<i>nfr5</i>	AJ575254	<i>Lotus japonicus</i>	40	可能的结瘤因子受体	AAM19130	类似受体激酶的蛋白
<i>LjNOD65</i>	AF000392	<i>Lotus japonicus</i>	40	肽运输因子	AAG21906	可能的肽运输因子
<i>ENOD2B</i>	X16876	<i>Glycine max</i>	40	富含脯氨酸蛋白, 参与根瘤内皮层的形成	BAC15501	可能的 GmENOD2 前体
<i>nin</i>	LJA239041	<i>Lotus corniculatus</i>	40	根瘤起始调节因子, 突变体不能结瘤	AAM22710	根瘤起始蛋白
<i>Lel</i>	K00821	<i>Glycine max</i>	39	细胞粘连, 扩大宿主范围	BAC79695	可能的类似凝集素蛋白的激酶
<i>cppl</i>	AJ010165	<i>Glycine max</i>	38	参与调控豆血红蛋白的表达	BAC83433	可能的具有 CXC 结构域蛋白的转录因子
<i>psl</i>	X66368	<i>Pisum sativum</i>	36	有丝分裂增强子, 扩大宿主范围	BAC79695	类似凝集素蛋白的激酶

¹⁾ 根据基因的登录号, 可以查到有关此基因的文獻。¹⁾ The references of a gene could be obtained according to accession number of the gene.

IR, Identities in amino acid aligned with the corresponding homologous genes in rice; FR, Function or phenotype during nodule development; AR, Accession No. of protein encoded by corresponding homologous genes in rice; DR, Description of the corresponding genes in rice.

物中高度同源的结瘤素基因 *enod40* 以及水稻对应同源基因 *Osenod40* 之间存在相似的调节机制。同时也暗示水稻中存在的某些高度同源的结瘤素基因与对应的豆科植物结瘤素基因很可能具有相似的表达机制。

2.2.2 结瘤素基因之一的蔗糖合成酶基因与水稻中对应同源基因蔗糖合成酶基因的比较分析

以大豆(*Glycine max*)蔗糖合成酶基因(GenBank 数据库登录号 AF030231)作为查询序列,在水稻基因组中发现 3 个与大豆蔗糖合成酶基因具有高度同源性的蔗糖合成酶基因 *RSus1* (GenBank 数据库登录号 X64770)、*RSus2* (GenBank 数据库登录号 X59046)、*RSus3* (GenBank 数据库登录号 L03366)。这 3 个基因与大豆蔗糖合成酶基因比对,其氨基酸序列一致性均达到 75%。将水稻中蔗糖合成酶基因 *RSus1* 与其他 5 种豆科植物中的蔗糖合成酶基因进行比对,结果显示水稻中的蔗糖合成酶基因与其他豆科植物中的蔗糖合成酶基因的氨基酸序列一致性在 70% 以上,具有高度的同源性(表 2)。

豆科植物蔗糖合成酶基因在根瘤内表达水平较高,与维持根瘤的正常发育和功能有关^[6,7]。在成熟根瘤中蔗糖合成酶参与蔗糖代谢,蔗糖合成酶基因除了负责提供根瘤固氮的碳源外,也是调控根瘤内碳代谢和氮固定的一种重要的方式,蔗糖合成酶基因突变导致不能有效固氮^[8,9]。Huang 等^[10]研究表明,*RSus1* 存在于叶肉细胞、根韧皮部以及种子糊粉层和胚乳细胞中,负责将蔗糖运输至胚乳细胞;*RSus2* 存在于叶组织的韧皮部、种子和胚乳中,可能是管家蛋白;*RSus3* 主要存在于胚乳细胞中,提供淀粉合成前体 UDPG。上述的研究揭示,尽管水稻中蔗糖合成酶基因和豆科植物蔗糖合成酶基因存在高度同源性,但是在豆科植物和水稻中呈现不同的表达模式,执行不同的功能。这一结果表明,高

度同源的豆科植物蔗糖合成酶基因和水稻蔗糖合成酶基因,随着进化和环境的改变,两者的功能也发生相应的改变。很可能豆科植物由于根瘤的形成,要求已存在的基因功能发生相应的改变,以执行新的功能。

2.2.3 结瘤素基因 *Rab1p* 与水稻中对应的同源基因 *OsRab* 比较分析

以大豆结瘤素基因 *Rab1p* (GenBank 数据库登录号 L14929)作为查询序列,对水稻数据库进行 BLAST 分析,在水稻基因组中找到一个高度同源的基因,称之为 *OsRab* 基因(GenBank 数据库登录号 AJ307662)。*OsRab* 与大豆的 *Rab1p* 基因进行比对,两者氨基酸序列一致性达到 78%。将水稻中 *OsRab* 基因与其他不同豆科植物中 *Rab* 基因进行多序列比较,表明水稻中 *OsRab* 基因与豆科植物中 *Rab* 基因存在高度的同源性(图 2)。

Rab 蛋白属于小 G 蛋白超级家族的成员,目前认为 *Rab* 蛋白调节植物细胞的基本生理过程^[11]。Cheon 等^[12]研究表明大豆结瘤素基因 *Rab1p* 的产物是一种 GTP 结合蛋白,属于小 G 蛋白超级家族,两者对于形成有效共生作用的类菌体膜区室的发育是必需的。对于水稻中 *OsRab* 基因编码的蛋白,Mayer 等^[13]推测可能是 GTP 结合蛋白。我们认为尽管目前无法确定水稻中 *Rab* 基因的功能,但是由于它与豆科植物中 *Rab* 基因存在很高的同源性,这暗示它们可能同属于一类直向同源基因,在表达和功能上存在一定的联系。

2.3 与水稻基因组比对不具有同源性的结瘤素基因

尽管在水稻基因组中发现了许多与结瘤素基因高度同源的基因。然而,我们通过对水稻基因组数据库和蛋白质数据库进行 BlastN 和 BlastP 搜索,经分析筛选,发现有 44 个结瘤素基因在水稻基因组

表 2 水稻蔗糖合成酶基因 *RSus1* 与豆科植物蔗糖合成酶基因的比对分析
Table 2. Comparative analysis of rice *RSus1* and sucrose synthase genes from various legumes.

蔗糖合成酶 基因登录号 ¹⁾ Accession No. of sucrose synthase ¹⁾	来源植物 Plant species	与 <i>RSus1</i> 比对的氨基酸 序列一致性 Identities in amino acid aligned with <i>Rsus1</i> / %	与 <i>RSus1</i> 比对的氨基酸 序列相似性 Similarities in amino acid aligned with <i>Rsus1</i> / %	E 值 E-value
AF049487	<i>Medicago sativa</i>	72	83	0
AJ012080	<i>Pisum sativum</i>	72	83	0
M97551	<i>Vicia faba</i>	72	83	0
AJ131964	<i>Medicago truncatula</i>	72	83	0
AF315375	<i>Phaseolus vulgaris</i>	71	82	0

¹⁾ 根据基因的登录号,可以查到有关此基因文献。
¹⁾ The references of a gene are obtained according to accession number of the gene.

表3 与水稻基因组比对不具有同源性的结瘤素基因

Table 3. Nodulin genes showing no homologs aligned with rice genome.

结瘤素基因 Nodulin gene	功能或描述 Function or description	植物来源 Plant species	GenBank 数据库登录号 ¹⁾ Accession No. ¹⁾
<i>enod3</i>	金属结合蛋白	<i>Pisum sativum</i>	S45142
<i>enod5</i>	类似的阿拉伯糖苷蛋白	<i>Trifolium repens</i>	AJ310341
		<i>Vicia faba</i>	AJ250498
		<i>Pisum sativum</i>	S45139
		<i>Vicia sativa</i>	X83681
<i>enod6</i>	与前侵染线形有关的蛋白	<i>Medicago truncatula</i>	AJ133118
<i>enod7</i>	—	<i>Pisum sativum</i>	X93172
<i>enod10</i>	—	<i>Medicago sativa</i>	M91074, M91075, M91076
<i>enod11</i>	胞外基质蛋白	<i>Medicago truncatula</i>	AJ297721
<i>enod12</i>	富含脯氨酸的蛋白	<i>Vicia faba</i>	AJ277289, AJ277288
		<i>Medicago truncatula</i>	X68032
		<i>Pisum sativum</i>	M60585
		<i>Vicia sativa</i>	X83682
		<i>Medicago sativa</i>	X74355, X74354, X74356
<i>enod14</i>	金属结合蛋白	<i>Pisum sativum</i>	S45162
<i>enod16</i>	与光花青素苷相关的蛋白	<i>Medicago truncatula</i>	X99466
<i>enod20</i>	与光花青素苷相关的蛋白	<i>Medicago truncatula</i>	X99467
<i>enod55</i>	—	<i>Glycine max</i>	X691157, X69156
<i>VfENOD-GRP3</i>	富含甘氨酸的蛋白	<i>Vicia faba</i>	Z46780
<i>ddl7</i>	—	<i>Trifolium repens</i>	AJ310341
<i>nodulin45</i>	—	<i>Lupinus angustifolius</i>	L12388
	—	<i>Lupinus luteus</i>	X77044
<i>PsN6</i>	—	<i>Pisum sativum</i>	AB059550
<i>PsN466</i>	—	<i>Pisum sativum</i>	AB059553
<i>PsN335</i>	—	<i>Pisum sativum</i>	AB059552
<i>PsN314</i>	—	<i>Pisum sativum</i>	AB059551
<i>PsN1</i>	—	<i>Pisum sativum</i>	AB059549
<i>MtN1</i>	—	<i>Medicago truncatula</i>	Y10456
<i>nodulin 25</i>	—	<i>Medicago truncatula</i>	AJ277861
		<i>Vicia faba</i>	Z14046
		<i>Medicago sativa</i>	X13287
<i>nod-CCP1</i>	—	<i>Vicia faba</i>	AJ243462
<i>NGLb-2</i>	与半乳糖结合的蛋白	<i>Arachis hypogaea</i>	AF127559, AF127560
<i>nod28/32</i>	—	<i>Vicia faba</i>	AJ237842
<i>nod28/32-K40</i>	—	<i>Vicia hirsuta</i>	AJ237841
<i>nod28/32-A10</i>	—	<i>Vicia hirsuta</i>	AJ237840
<i>MtN13</i>	—	<i>Medicago truncatula</i>	Y100455
<i>PR10-1</i>	—	<i>Medicago truncatula</i>	Y08641
<i>ENBP1</i>	DNA 结合蛋白	<i>Vicia sativa</i>	X95995
<i>nod3</i>	—	<i>Glycine max</i>	X96792
<i>LjNOD21</i>	—	<i>Lotus japonicus</i>	AF000402
<i>LlPRP</i>	富含脯氨酸的 PRP2 蛋白前体	<i>Lupinus luteus</i>	U47661
<i>nodulin 30</i>	—	<i>Phaseolus vulgaris</i>	Z30347
<i>nodulin 44</i>	—	<i>Glycine max</i>	X05028
<i>Ngm-16</i>	—	<i>Glycine max</i>	X54307
<i>Nodulin-24</i>	类菌体周膜蛋白	<i>Glycine max</i>	M10595
<i>Nodulin-23</i>	—	<i>Glycine max</i>	X01704
<i>Nodulin-22</i>	—	<i>Glycine max</i>	X05024
<i>Nodulin-20</i>	—	<i>Glycine max</i>	X05020
<i>N-20t</i>	—	<i>Glycine max</i>	X60159
<i>PRP4</i>	富含脯氨酸的蛋白 4	<i>Medicago truncatula</i>	L23504
<i>VfNDSA2</i>	—	<i>Vicia faba</i>	Z46911
<i>Mszpt2-1 (Mscp17)</i>	细胞分化	<i>Medicago truncatula</i>	Y16131
<i>NORK</i>	不发生卷曲,仅根毛顶部膨大	<i>Medicago sativa</i>	AJ418377

¹⁾ 根据基因的登录号,可以查到有关此基因的文献;“—”表示此基因的功能不清楚。

¹⁾ The references of a gene are obtained according to accession number of the gene; “—” means function of a gene is unknown.

对于水稻中广泛存在豆科植物结瘤素基因的同源基因,可以从分子系统发生学和进化的角度进行分析和解释:(1)在水稻中这些同源基因可能是保守的,它们可能是豆科植物中结瘤素基因的直向同源基因。豆科植物中的结瘤素基因在形成根瘤器官的过程中发生变异和改变。另外,在其他的非豆科植物中(如烟草、拟南芥等)也发现类似的豆科植物结瘤素基因^[17]。这表明同时存在于单子叶和双子

叶植物中的结瘤素同源基因很可能来自共同的祖先基因,单子叶植物和双子叶植物分开后,与结瘤功能有关的结瘤素基因发生了进化。(2)某些早期结瘤素基因不仅在根瘤器官形成时诱导表达,在其他的组织器官中也有表达。这表明结瘤素基因的生物功能可能是多样的,不仅限于根瘤器官发生,还在植物生长和发育方面执行更普遍的功能。

然而,也有许多豆科植物结瘤素基因在水稻基

因组中找不到同源性的序列,这些基因在共生过程中承担着重要的功能。根据广泛的杂交和图谱比较结果,认为植物中决定形态和生理功能差异的基因可能是由于少数遗传基因差异造成的,这些基因可能是调节基因^[18,19]。由此,推测水稻中缺少的这些豆科结瘤素基因的同源基因,很可能是豆科植物结瘤固氮的关键性基因,正是缺少这些基因导致水稻不能结瘤固氮。

利用生物信息学的方法,我们在水稻基因组中发现的31个与结瘤素基因具有高度同源性的基因,其中大多数只是注释的未知基因,优先分离和克隆水稻中这些基因,确定这些基因的功能将为水稻结瘤固氮潜能的研究提供更多的线索。随着豆科模式植物 *Medicago truncatula* 和 *Lotus japonicus* 基因组测序数据的公布,将提供更多关于豆科植物共生作用的遗传信息。将来可以利用比较遗传学,在整个基因组、转录组和蛋白质组的水平上,对水稻和豆科植物进行全面比较分析,寻求两者之间的差异,进而为探索水稻结瘤固氮的研究提供更多更全面的信息。

参考文献:

- Reddy P M, Aggarwal R K, Ramos M C, *et al.* Widespread occurrence of the homologue of the early nodulin (ENOD) genes in *Oryza* species and related grasses. *Biochem Biophys Res Comm*, 1999, 258: 148—154.
- Verma D P S, Fortin M G, Stanley J, *et al.* Nodulins and nudulins of *Glycine max*. *Plant Mol Biol*, 1986, 7: 51—61.
- Kouchi H, Hata S. Isolation and characterization of novel nodulin cDNAs representing genes expressed at early stages of soybean nodule development. *Mol Gen Genet*, 1993, 238: 106—119.
- Yang W C, Katinakis P, Hendriks P, *et al.* Characterization of *GmENOD40*, a gene showing novel patterns of cell-specific expression during soybean nodule development. *Plant J*, 1993, 3: 573—585.
- Kouchi H, Takane K, So R B, *et al.* Rice *ENOD40*: isolation and expression analysis in rice and transgenic soybean root nodules. *Plant J*, 1999, 18(2): 121—129.
- Gordon A J, Minchin F R, James C L, *et al.* Sucrose synthase in legume nodule is essential for nitrogen fixation. *Plant Physiol*, 1999, 120: 867—877.
- Thummler F, Vernon D P S. Nodulin-100 of soybean is the submit of sucrose synthase regulated by the availability of free heme in nodules. *J Biol Chem*, 1987, 262: 14730—14736.
- Gordon A J, James C L. Enzymes of carbohydrate and amino acid metabolism in developing and mature nodules of white clover. *J Exp Bot*, 1997, 48: 895—903.
- Craig J, Barratt P, Tatge H, *et al.* Mutations at the *rug4* locus alter the carbon and nitrogen metabolism of pea plants through an effect on sucrose synthase. *Plant J*, 1999, 17: 353—362.
- Huang J W, Chen J T, Yu W P, *et al.* Complete structure of three rice sucrose synthase isogenes and differential regulation of their expression. *Biosci Biotech Biochem*, 1996, 60(2): 233—239.
- Yang Z B. Small GTPases: versatile signaling switches in plants. *Plant Cell*, 2002, (suppl): 375—388.
- Cheon C I, Lee N G, Siddique A B, *et al.* Roles of plant homologs of *Rab1p* and *Rab7p* in the biogenesis of the peribacteroid membrane, a subcellular compartment formed de novo during root nodule symbiosis. *EMBO J*, 1993, 12(11): 4125—4135.
- Mayer K, Murphy G, Tarchini R, *et al.* Conservation of microstructure between a sequenced region of the genome of rice and multiple segments of the genome of *Arabidopsis thaliana*. *Genome Res*, 2001, 11(7): 1167—1174.
- Christiansen A, Hansen A C, Vijn I, *et al.* A novel type of DNA binding protein interacts with a conserved sequence in an early nodulin ENOD12 promoter. *Plant Mol Biol*, 1996, 32(5): 809—821.
- Kosuta S, Chabaud M, Loughon G, *et al.* A diffusible factor from arbuscular mycorrhizal fungi induces symbiosis-specific *MtENOD11* expression in roots of *Medicago truncatula*. *Plant Physiol*, 2003, 131(3): 952—962.
- Pichon M, Journet E P, Dedieu A, *et al.* *Rhizobium meliloti* elicits transient expression of the early nodulin gene *ENOD12* in the differentiating root epidermis of transgenic alfalfa. *Plant Cell*, 1992, 4(10): 1199—1211.
- Miklashevichs E, Röhrig H, Schell J, *et al.* Perception and signal transduction of rhizobial NOD factors. *Crit Rev Plant Sci*, 2001, 20: 373—394.
- Doebley J F, Stec A, Gustus C. Teosinte branched1 and the origin of maize: evidence for epistasis and the evolution of dominance. *Genetics*, 1995, 141(1): 333—346.
- Paterson A H, Lin Y R, Li Z, *et al.* Convergent domestication of cereal crops by independent mutations at corresponding genetic loci. *Science*, 1995, 269: 1714—1717.