

华葵根瘤菌 *nifA* 基因的克隆和功能分析^{*}

高成江 沈思师 金润之

(中国科学院上海植物生理研究所 上海 200032)

摘 要 华葵根瘤菌(*Mesorhizobium huakuii* 即 *R. astragali*)159 的 *nifA* 基因的序列分析表明,该基因全长 1227bp,编码分子量为 44734D 的 NifA 蛋白。与其它 NifA 蛋白的序列比较发现,华葵根瘤菌 NifA 蛋白也存在保守的中间结构域和 C-末端 DNA 结合结构域,但其氨基端缺失。Tn5 定点突变得到的突变体是 Nif⁻表型。构建了 *nifA* 基因组成型表达的质粒,此质粒在大肠杆菌中对华葵根瘤菌 *nifH-lacZ* 有激活作用。

关键词 华葵根瘤菌, *nifA* 基因, DNA 序列, *nifH-lacZ* 融合基因

中图分类号: Q78 文献标识码: A 文章编号: 0001-6209(2000)03-0257-63

生物固氮是一些微生物将大气氮还原为氨的复杂过程。这个酶促过程需要许多固氮基因(*nif* 和 *fix*)的表达。大部分固氮基因启动子区都含有 TGGCAG-N₅-TTGCA 的保守序列。此类启动子的表达需要 *nifA* 基因编码的 NifA 蛋白作为正调节因子与 *ntrA* 基因编码的 σ^{54} 因子^[1]。多种固氮微生物的 *nifA* 基因已被克隆,并且它们的调控方式不同^[2]。在肺炎克氏杆菌(*Klebsiella pneumoniae*)中 *nifA* 基因由 *ntrC* 基因产物调节,而在苜蓿根瘤菌(*Rhizobium meliloti*)中 *nifA* 基因被 *fixJ* 调节^[3]。在慢生型大豆根瘤菌(*Bradyrhizobium japonicum*)中 *nifA* 基因与 *fixR* 基因组成一个操纵子,受到双重调控^[4]。

nifA 基因作为固氮基因的正调节基因,已经从 10 多种固氮微生物克隆出来并完成了序列测定。华葵根瘤菌是我国特有的温带快生型根瘤菌,与紫云英(*Astragali sinicus*)相互作用形成有效根瘤进行生物固氮。我们用苜蓿根瘤菌 *nifA* 基因作探针从华葵根瘤菌基因文库中筛选到 1 株阳性克隆 pRaG25,完成了 *nifA* 基因的序列测定,用 Tn5 对 *nifA* 基因进行了定点突变,并应用 *nifH-lacZ* 融合基因在大肠杆菌中测定了 *nifH* 基因对 *nifH* 基因的激活作用。

1 材料和方法

1.1 菌株和质粒(见表 1)

根瘤菌生长于 TY 培养基(28℃),大肠杆菌生长于 LB 培养基(37℃)。抗菌素及使用浓度:链霉素(Sm)500μg/mL,氨基青霉素(Ap)100μg/mL,四环素(Tc)15μg/mL。

1.2 酶和试剂

限制性内切酶购自华美生物工程公司。DNA 测序试剂盒购自美国 Promega 公司。

作者简介:高成江(1971-)男,山东鄄城县人,中国科学院上海植物生理研究所博士研究生,研究方向是生物固氮的分子遗传学

收稿日期:1999-01-11,修回日期:1999-06-12

© 中国科学院微生物研究所期刊联合编辑部 <http://journals.im.ac.cn>

表 1 菌株和质粒
Table 1 Strains and Plasmids

Strains and plasmids	Genotype/Phenotype	Reference/Source
<i>E. coli</i>		
JM105	lac pro thi strA endA sbc15 hsdR4 (F' traD36 proAB lacI ^q lac △ ZM15)	Yanish-Perron 1985 [5]
YMC11	YMC9hutC _{K. aerogenes} glr(ALG) △ 2000	
<i>Mesorhizobium huakuii</i>		
159	Wild type , Sm ^r	This lab
159 – 1	<i>nifA</i> :Tn5 , Sm ^r Nm ^r	This study
159 – 2	<i>nifA</i> :Tn5 , Sm ^r Nm ^r	This study
Plasmids		
pRaG25	Clone isolated from <i>M. huakuii</i> 159 gene bank using <i>R. meliloti nifA</i> gene as probe , Tc ^r	This study
pSUP201	AP ^r ,Cm ^r ,Mob	[6]
pUC9	Cloning vector , Ap ^r	Yanish-Perron [7]
pGD499	Km ^r - <i>LacZ</i> in pPK290 , Tc ^r	
pRaG251	4.4 kb fragments homologous to <i>R. meliloti</i> <i>nifA</i> gene cloned in pUC9 ,Ap ^r	This study
pRaG252	2.3 kb fragments homologous to <i>R. meliloti</i> <i>nifA</i> gene cloned in pUC9 , Ap ^r	This study
pRaG253	0.9 kb fragments homologous to <i>R. meliloti</i> <i>nifA</i> gene cloned in pUC9 , Ap ^r	This study
pRaG254	<i>M. huakuii</i> 159 <i>nifA</i> ⁺ pUC9 derivative , Ap ^r	This study
pRaG255	<i>M. huakuii</i> 159 <i>nifH-lacZ</i> , pGD499 derivative , Tc ^r	This study
Phages		
M ₁₃ mp18/19	Sequencing vector	Messing 1983

1.3 DNA 制备、杂交实验

质粒 DNA 制备、Southern 转移、缺口翻译、DNA 杂交按 Sambrook 等人^[8]的方法进行。探针 DNA 用 α-³²P 标记 ,其放射性比强度为 7 × 10⁷ cpm/μg。

1.4 DNA 序列测定

将各酶解片段克隆到 M₁₃mp 和 pUC 载体 ,提取单链和双链 ,自动测序在 ABI model 373A DNA 分析仪进行 ,部分序列用双脱氧法由手工测定。DNA 序列分析用 Pcgene 程序进行。

1.5 半乳糖苷酶活性测定

将过夜细菌培养物按 1 :50 接入新鲜培养基 ,培养至对数期 ,离心收集菌体 ,悬浮于相同新鲜培养基中 ,在 37℃ 有氧或无氧(抽气后充氩气 3 次)条件下振荡培养 8 h ,按 Miller^[9]方法测定 β-半乳糖苷酶活力。

1.6 *nifA* 基因插入突变体的构建

把 2.3 kb *pst* I 片段克隆到 p_{sup}201 ,对此质粒进行 Tn5 插入诱变 ,使突变质粒重组进入华菱根瘤菌基因组 ,选择 Nm^rCm^s 双交换突变体 ,用 Southern 方法进一步鉴定。

1.7 结瘤实验

结瘤实验按金润之^[10]等人的方法进行。固氮酶活性测定按文献 [11] 进行。

2 结果和讨论

2.1 华葵根瘤菌 *nifA* 基因的克隆与序列测定

用苜蓿根瘤菌 *nifA* 基因作探针从华葵根瘤菌基因文库中筛选到 1 株阳性克隆 pRaG25 ,进一步的 DNA 杂交发现 2.3 kb *Pst* I / 4.4 kb *Hind* III 和 0.9 kb *Eco*R I 片段能与探针杂交 (结果未显示) ,从 pRaG251 的物理图谱显示 2.3 kb *Pst* I 片段包含在 4.4

kb *Hind*III 片段中 ,而 0.9 kb *Eco*R I 片段又包含在 2.3 kb *Pst* I 片段之中 ,如图 1 所示。将 0.9 kb *Eco*R I 片段克隆到 M₁₃mp 载体 ,采用双脱氧法完成了此序列测定。把此序列输入计算机分析发现该片段与不同来源的 *nifA* 基因存在同源性 ,并位于 *nifA* 基因的 3'端 ,此后我们又对 0.9 kb *Eco*R I 片段上游片段进行序列测定 ,从而完成了整个 *nifA* 基因的序列测定 ,测序策略如图 1。

2.2 测序结果及序列分析

nifA 基因序列包括上游和下游序列如图 2 所示。利用 Pcgene 程序对该序列进行阅读框分析发现此序列包含两个大的 ORF(开放阅读框) ,两者相距 391 bp。ORF1 从 449 位核苷酸起始到 1417 位核苷酸终止 ,编码分子量为 37680D 的蛋白。ORF2 从 1807 位核苷酸起始 ,到 3033 位核苷酸终止 ,编码分子量为 44734D 的蛋白。ORF2 的 1227bp 核苷酸序列与其它 *nifA* 基因的核苷酸序列非常相似 ,并且此 ORF 编码的蛋白与其它 NifA 蛋白存在同源性 (图 3) ,因此我们将 ORF2 定为 *nifA* 基因。ORF1 的功能还有待进一步实验。在苜蓿根瘤菌中 *nifA* 基因与 *fix*ABCX 操纵子紧密连锁 ,在慢生型大豆根瘤菌中 *nifA* 基因与 *fixR* 基因连在一起构成一个操纵子 ,两者相距 187bp^[12]。华葵根瘤菌 ORF1 编码的蛋白与苜蓿根瘤菌 *fix*ABCX 基因及慢生型大豆根瘤菌 *fixR* 基因编码的蛋白没有同源性 ,因此华葵根瘤菌 *nifA* 基因的排列方式与苜蓿根瘤菌及慢生型大豆根瘤菌显著不同。ORF2 有两个可能的起始密码子(GTG 和 ATG) ,第一个起始密码子 GTG 位于 1807 核苷酸处 ,在其上游第 9 核苷酸处存在一个可能的 S.D. 序列(AGAAA)。第二个起始密码子 ATG 位于 1921 核苷酸处 ,在其上游第 7 核苷酸处有一个可能的 S.D. 序列(AGGAGA)。根据 NifA 蛋白的同源性比较我们将第一个起始密码子定为 *nifA* 基因的翻译起始位点。

由 ORF2 推测的氨基酸序列与苜蓿根瘤菌、慢生型大豆根瘤菌和肺炎克氏杆菌的 NifA 蛋白的同源性比较如图 3 所示。在 4 种 NifA 蛋白中都存在两个高度保守的结构域 :中间结构域以及羧基端 DNA 结合结构域。中间结构域是与 RNA 聚合酶的 σ^{54} 因子相

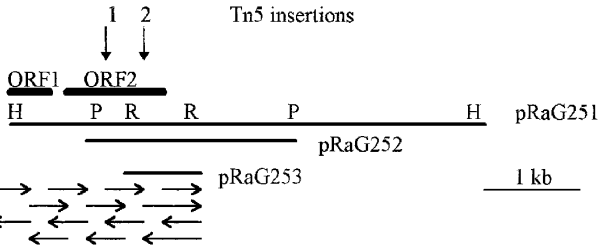


图 1 质粒 pRaG251 的物理图谱及 *nifA* 基因的测序策略
Fig.1 Physical map of plasmid pRaG251 and sequencing strategy for the *nifA* gene
The bold regions indicate ORF1 and ORF2 respectively. The thin arrows indicate the length and direction of sequenced fragments. The vertical arrows indicate Tn5 insertions in the *nifA* gene. Restriction enzyme sites : R , *Eco*R I ; P , *pst* I ; H , *Hind* III .

ACGCTTCTTCCGCAAGGCCTTTGACCGCCATGGCCGGCCGGATCGCATCGTCAACGGCAGCCAAACAAATCACAAGGCAATCCTTTC 1230
CTGCAATGTGTAAAGCCGCTTTGCGGGATCAATCGCGCGCATCGCTGAAGCCAATCCGCAACCCGCCAAAGCCAGTACCTCAACAAATCGCCTT 1320
TGAGCAGGATCATCGACGGATCAAGTGGCCGATCAGCTCCATGCTCGGCTTCAAGCGGAAAGGCGAACCCGCGAAATTATCTGTCTGGC 1410
ATCGAGATGATCCACATGATGCGCAAAACGGCAGGCGAGGTATGCCCTACAATCCAGCCCCCTAATCGCGCAGCAGTTCCGCCATCTCACC 1500
GCATAGAAGCCGACCCGGATTCTGCCATCCTTTGTCCCGTTTCGATCTTTGCGACAGAACC GCCCTGGTTAAGGGCTCGCGCGCAAT 1590
GTGCGCAGCTTCGCCAGATGATCGACTGGACGAAGCACCGGATCGACTTCTCGACAGCGCCATCGCTCGGCGTCAGCGCCGGTAAAT 1680
TCTCGCACGCGGCTGGTAAACACGCTGCATCTGTGTATCAGCGCTTAAAGTGCTTGATGACTGCCTTCCACGCGAGCAAGCGCGACAAT 1770
GAATCGCAGGCTGCACGCGGTAGAAACTCTGGCTGTGGCGGCCAACGGTGGCTCATCCCTGAGCCAGACTAAGCCGTGTGCGCCGAGG 1860
M A A N G G S S L S Q T K P C A P R
TACGTTGCAAGCCAGATAACGGGGCGTTGTGTTTGTGCCCCGAAACCGAGGAGACGAATAATGGCCAATGGTGACAATAATTAGACCGT 1950
Y V A S Q I T G R C F C G P K P R R I M A N G D N N L D R
AATTTAGACCGTGGGGAACCGATAGCAAAAGCAGTACTTGTCTGACAGGGGATCATTGGCGACAGTTTGCCTTTGGCCAATTTGCTTCAA 2040
N L D R G E P I A K Q Y L L S E G I I G D S L P L R N L L Q
ACAGTTGCGCGCTCGCTAAGTACAACCTTTCGGCTTCTTTCGCGAGGCGAATCCGGTACCGGCAAGGAGCTGGTGCCTCAAGGCTATTAC 2130
T V A R V A K Y N F P V L L R G E S G T G K E L V A K A I H
GAGTTGTGCGCCCGGCCAAGCGGCCCTTCGTAAGGTCAACTGTGCGCGCTGACAGAAAGCGTCTCGAATCCGAATTATTTCGGTTCAT 2220
E L S P R A K R P F V K V N C A A L T E S V L E S E L F G H
GAGAAGGGCGATTCAACAGCGACTTCTACTCGCATAAGGACGCTTTGAGCTCGCTCACAAGGGGACGTTATTCTTGGACGAGATCGGC 2310
E K G A F T S A F Y S H K G R F E L A H K G T L F L D E I G
GAGATCTCACCCCTGGTTTCAGGCAAGGCTCTTACCGGTGCTGCAGGAGCAGGAGTTCGAGCGCGTGGCGGGCAACCAGACCATTAAGATC 2400
E I S P S F Q A K L L R V L Q E Q E F E R V G G N Q T I K V
GATGTTGCGATCATTGCGCGACGAACAAGAATTTGGAAGAGGCGGTGTAATAAATGAATTCGGTGGCGACCTCTATTACCGCTCAAT 2490
D V R I I A A T N K N L E E A V V K N E F R A D L Y Y R L N
GTGGTTCAATTGCTGGTGGCGCGCTACGCGAAAGCGGCACTGATATCCGCTCCTCGCAGCGGAGTTTCTCAAGAAGTTCAACGATGAG 2580
V V H L L V P P L R E R R T D I P L L A A E F L K K F N D E
AACGCGAATATGCTGACCTTCGATGCCAGTGGCATTGAAGTACTGACAAGCTGTGATTTTCCGGCAACATTGCGAGCTGCAAAACTGC 2670
N G N M L T F D A S A I E V L T S C D F P G N I R E L E N C
GTGTATCGGACAGCGGTTCTGCGCGAGGAGCATCAATCGGAGAGGCGACTTTGCATCGTGTGGGGCCAATGCTTCGCTGGCGGTTG 2760
V Y R T A V L A E G A S I G R G D F A C C R G Q C F A A A L
CATAAGCTCACATCTGGAGAAGTCCGCTGCGACCCCGGTGCAACATGCCGAGCAAGCGATTTCGGCCCCCTCGCGCGAGTCTCGTCA 2850
H K L T S G E L A L R P R S N M P Q Q A I S A P S A A V S S
GAGGCTGGAGSACAGCCCGCGCGCGGCTGCGCAGCAGTCCACTTGAACAATGCCGATCGGATCAGCGCGCCATGGACATATGTGGC 2940
E A G G Q P P A R P A D A V H L N N A D R I T A A M D I C G
TGGGTCCAAGCGAAGGCGCGCAGCCTGCTTGGACTGACACCGCGCAGATTGGTTACGCGTGAAAAACACGCCATCAATGTGAAGCGT 3030
W V Q A K A A R L L G L T P R Q I G Y A L K K H A I N V K R
TTTGAACAAACGACCGACGATGGCAGCTGGGGCAAATCCGACTTGGTGCTTTGGCGACTCGAGACGAAACGACGGAGTCTGCGGTCA 3120
F *
GTGGGCGTTGTACGTTGTTTGAATGTGGCCGCTGAGGCTGCTACCGCGCTTTTCAGGCAGGCAGTGCCTCACGGCTTCCCACGCGG 3210
CCATTGCCTGGCGTCGATGGGTTGCAATTTGTGCTGCCGAGCGGTTGGTCTGGGAGGG 3268

图2 *nifA* 基因的核苷酸序列和编码的氨基酸序列

Fig.2 Nucleotide and deduced amino acid sequence of *Mesorhizobium huakuii nifA* gene
Putative translation initiation codons are indicated by bold letters. Presumptive ribosome binding sites(S. D.)are underlined. The amino acid sequence of the ORF1 are not indicated in this figure.

结合的部位,在其中还含有一可能的ATP结合位点(GESGTGKE)^[13],水解ATP并促进封闭型启动子复合物向开放型启动子复合物的转变。羧基端结构域是一个螺旋—转角—螺旋(helix-turn-helix)的结构,并与DNA结合。在华蔺根瘤菌、苜蓿根瘤菌和慢生型大豆根瘤菌NifA蛋白中均存在4个保守的C₂残基(Cys-X-Cys-X-Cys-X-Cys),而在肺炎

型大豆根瘤菌和肺炎克氏杆菌分别缺失 132、173 和 116 个氨基酸残基。从所列 4 种 NifA 蛋白的氨基酸同源性比较表明不同 NifA 蛋白的氨基端结构域存在很低的序列同源性,并且在三叶草根瘤菌(*Rhizobium leguminosarum* bv. *trifolii*)中 NifA 蛋白缺失氨基端结构域但仍保留其功能^[11]。

2.3 *M. huakuii* 159 *nifA* 的 Tn5 插入突变体的构建和功能

为了验证 *nifA* 基因的功能,用 Tn5 插入诱变的方法构建了两株 *nifA* 基因的插入突变体,Tn5 在 *nifA* 基因的大致插入位点如图 1。结瘤实验表明,突变体诱导的根瘤为白色,比野生型形成的根瘤小。固氮酶活性分别为野生型根瘤菌的 10%和 15%(表 2)。

表 2 华葵根瘤菌野生型及 *nifA* 基因 Tn5 插入突变体诱导的根瘤特性

Table 2 Characteristics of nodules induced by <i>M. huakuii</i> 159 and <i>M. huakuii</i> 159 <i>nifA</i> :Tn5 mutants			
Strains	<i>M. huakuii</i> 159	<i>M. huakuii</i> 159-1	<i>M. huakuii</i> 159-2
Acetylene reduction/%	100	15	10
Nodule phenotype	Red big	Light small	Light small

* Acetylene reduction was measured in 30 days after inoculating *Astragali sinicus*. All data were the averages of three independent assays.

2.4 *nifA* 基因对 *nifH-lacZ* 的激活

为了进一步验证 *nifA* 基因的功能,在大肠杆菌中试验了 *nifA* 基因对 *nifH* 基因的激活作用。将包含完整 *nifA* 基因的 4.4kb *Hind*Ⅲ片段克隆到 pUC9 上,这样 *nifA* 基因可在 *plac* 作用下由 IPTG 诱导组成型表达,从而构建了华葵根瘤菌 *nifA* 基因组成型表达质粒 pRaG254。由于华葵根瘤菌 159 的 *nifH* 基因已被克隆并完成了序列测定^[15],据此我们构建了华葵根瘤菌 *nifH-lacZ* 融合子。将包含 *nifH* 启动子区和部分编码区的 1.6kb *Sal* I 片段克隆到 pUC19,选择正确方向的克隆,再利用 pUC19 上的 *Bam*H I 和

*Hind*Ⅲ酶切位点,将 1.6kb *Bam*H I -*Hind*Ⅲ片段克隆到 pGD499 上,从而构建了华葵根瘤菌 *nifH-lacZ* 融合子。*nifH-lacZ* 的构建如图 4。将华葵根瘤菌 *nifA*^c 与 *nifH-lacZ* 共同引入到 YMC11 中,在有氧和微氧条件下测定 *nifA* 基因对 *nifH* 基因的激活作用(表 3),在有氧和微氧条件下 *nifH* 的活性分别提高了 19 倍和 27 倍,又由于 YMC11 是 *ntrC*⁻从而排除了 *ntrC* 对 *nifH* 的激活作用,所以可以断定我们克隆的基因是 *nifA* 基因。我们构建的 *nifA* 基因在 YMC11 中组成型表达其 NifA 蛋白的含量有可能比较高,使 *nifH-lacZ* 在有氧条件下具有很高的活性。*nifA* 基因更精确的功能和本身的调控研究正在进行中。

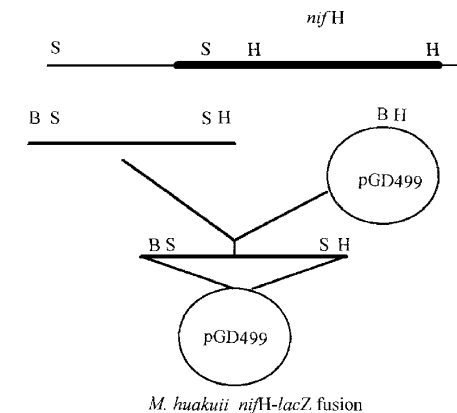


图 4 华葵根瘤菌 *nifH-lacZ* 的构建
Fig.4 Construction of *M. huakuii nifH-lacZ* fusion
S : *Sal* I , B : *Bam*H I , H : *Hind*Ⅲ .

表 3 华葵根瘤菌 *nifA* 基因在 YMC11 中对 *nifH-lacZ* 的激活作用

Table 3 Expression of *M. huakuii* 159 *nifH-lacZ* in the presence of *nifA* gene in YMC11(Averages of three independant assays)

Strains/Plasmids	Relevant characteristics	β -galactosidase activity(miller units)	
		Aerobically	Microaerobically
YMC11(pRaG254)	<i>nifA^c</i>	16	45
YMC11(pRaG255)	<i>nifH-lacZ</i>	120	138
YMC11(pRaG254 + pRaG255)	<i>nifA^c + nifH-lacZ</i>	2285	3768

致谢 感谢沈善炯先生对本工作的指导。

参 考 文 献

[1] Gonzalez V , Olvera L , Soberon X , *et al.* *Mol Microbiol* , 1998 , **28** (1) :55~67.

[2] Fisher H M. *Trends in Microbiol* , 1996 , **4** (8) :317~320.

[3] David M , Caveran M , Batut J , *et al.* *Cell* , 1988 , **54** :671~683.

[4] Thony B , Anthamatten D , Hennecke H. *J Bacteriol* , 1989 , **171** :4162~4169.

[5] Backman K , Chen Y M , Magasanik B. *Proc Natl Acad Sci USA* , 1981 , **78** :3743~3747.

[6] Simon R , Priefer U , Puhler A. *Biotechnology* , 1983 , **1** :784~791.

[7] Ditta G , Schmidhauser T , Yakobson E , *et al.* *Plasmid* , 1985 , **13** :149~153.

[8] Sambrook J , Fritsch E F , Maniatis T. *Molecular Cloning , A Laboratory Manual*. Second ed. , New York : Cold Spring Harbor Laboratory Press , 1989.

[9] Miller J H *Experimnts in Molecular Genetics* , Cold Spring Harbor. New York : Cold Spring Harbor Labarotory Press , 1972. 352~355.

[10] 金润之 朱劲松 江群益 等.微生物学报 ,1993 ,**33**(3) :170~176.

[11] Lismaa S E , Waston J M. *Mol Microbiol* , 1989 , **3** (7) :943~955.

[12] Thony B , Fisher H M , Anthamatten D , *et al.* *Nucleic Acids Res* , 1987 , **15** :8479~8499.

[13] Berger D , Narberhaus F , Lee H S , *et al.* *J Bacteriol* , 1995 , **177** :191~199.

[14] Michiels J , D 'hooghe I , Verreth C , *et al.* *Arch of Microbiol* , 1994 , **161** :404~408.

[15] 金润之 朱劲松 沈思师 等.科学通报 ,1993 ,**38**(21) :1990~1994.

IDENTIFICATION AND FUNCTIONAL ANALYSIS OF *NIFA* FROM
MESORHIZOBIUM HUAKUII 159

Gao Chengjiang Shen Sishi Jin Runzhi

(Shanghai Institute of Plant Physiology , the Chinese Academy of Sciences , Shanghai 200032)

Abstract : The *nifA* gene from *Mesorhizobium huakuii* 159 was identified and sequenced. An open reading frame coding for 44734D. NifA protein would detected. Comparison of the deduced amino acid sequence with other NifA proteins revealed the presence of the conserved central domin and the C-termainal DNA-binding domains. But the predicted *M. huakuii* NifA protein lacked a N-terminal domain that is present in the homologous proteins. Mutants with Nif⁻ phenotype were obtained by site-directed Tn5 mutagenesis. A *nifA* constitutively expressed plasmid was contruted , and this plasmid could activate the *nifH* expression in YMC11.

Key words : *Mesorhizobium huakuii* *nifA* gene DNA sequence *nifH-lacZ* fusion