

## 巴西固氮螺菌 Yu62draTG 基因及其下游区域的 克隆与核苷酸序列分析

马旅雁 李季伦

(北京农业大学农业生物技术国家重点实验室 北京 100094)

**摘 要** 以 *Azospirillum brasilense* Sp7 4.0kb draTG 片段为探针,自 *A. brasilense* Yu62 的基因文库中克隆了约 8kb 的 draTG 同源片段。通过对该片段的 Southern 杂交分析发现 *A. brasilense* Yu62 的 draTG 基因定位在 3.0kb EcoRI-Kpn I 片段上,其上游与 nifH 基因相邻。DNA 序列分析结果表明:该片段含有完整的 draTG,这两个基因下游还有两个开放阅读框架(ORF3 和 ORF4,其中 ORF4 是不完整的),draTG 及下游的 ORF3 推测以一个操纵元的方式转录;在 draG 及 ORF3 的上游区域均发现  $\sigma^{54}$  依赖型启动子的特征序列(DPE 及 UAS),推测它们与 draT 共转录外,还有可能单独转录。同源比较的结果表明 *Azospirillum* 的 DraTG 是非常保守的,它们在菌株及种间的差异都很小;紧接着 draG 的 ORF3 除与 *A. lipoferum* 和 *Rhodospirillum rubrum* 相应位置的 ORF 同源外,还与 *Azotobacter vinelandii* 的 ORF14 同源;ORF3 下游的 ORF4 与大肠杆菌的 yafJ 基因有较高的同源性。

**关键词** 巴西固氮螺菌, draTG, ORF, 核苷酸序列

固氮螺菌(*Azospirillum*)是一类可与许多禾本科作物及牧草根际联合固氮的微生物<sup>[1]</sup>,由于其具有分布广、数量多、适应性强等优点,成为最受重视的联合固氮菌。目前对其固氮遗传学研究的主要对象是巴西固氮螺菌(*A. brasilense*),该菌与深红红螺菌(*Rhodospirillum rubrum*)一样具有固氮酶翻译后的活性调控机制<sup>[2]</sup>。该调控是由 DRAT(Dinitrogenase reductase ADP-ribosyltransferase)和 DRAG(Dinitrogenase reductase activating glycohydrolase)催化的可逆 ADP-核糖修饰系统来完成<sup>[3]</sup>。编码这两个酶的 draT 和 draG 基因在 *A. brasilense* Sp7、*A. lipoferum*、*R. rubrum*、*Rhodobacter capsulatus* 这四个菌中都已得到序列分析。在前三个菌中 draTG 的上游均与 nifH 相邻,但转录方向相反。此外,在 *A. lipoferum*、*R. rubrum* 中 draTG 下游都发现一个功能未知的开放阅读框架(ORF)<sup>[4~7]</sup>。本文以 *A. brasilense* Sp7 draTG 片段为探针,从 *A. brasilense* Yu62 的基因文库中克隆了 8.0kb 的 draTG 同源片段。对该片段的 Southern 杂交分析表明 draTG 基因定位在 3.0kb EcoRI-Kpn I 片段上,其上游与 nifH 基因相邻。该片段及 draTG 上游区域(包括部分 nifH)被分别亚克隆并进行了全序列分析。由于篇幅有限,本文报道的只是 draTG 基因及其下游约 1kb 区域的序列分析结果。draTG 上游区域的核

863 计划资助项目。

本文部分内容曾在 1995 年“第十届国际固氮大会”上报告。

本文于 1996 年 11 月 5 日收到。

苷酸序列及 *draTG* 启动子的分析结果将在另一篇文章中报道。

## 1 材料和方法

### 1.1 菌株与质粒(表 1)

表 1 菌株与质粒

Table 1 Bacterial strains and plasmids

Strain and plasmid	Phenotype and/or Genotype	Reference or source
Strain <i>E. coli</i> K-12		
BHB2688	(N205 <i>recA</i> [ $\lambda$ <i>imm434</i> <i>cIts</i> <i>b2</i> <i>red</i> <i>Eam</i> <i>Sam</i> / $\lambda$ ])	[8]
BHB2690	(N205 <i>recA</i> [ $\lambda$ <i>imm434</i> <i>cIts</i> <i>b2</i> <i>red</i> <i>Dam</i> <i>Sam</i> / $\lambda$ ])	[8]
JM101	<i>supE</i> <i>thi</i> $\Delta$ ( <i>lac</i> <sup>-</sup> <i>proAB</i> )	[9]
	F' [ <i>traD36</i> <i>proAB</i> <sup>+</sup> <i>lacI</i> <sup>s</sup> <i>lacZ</i> $\Delta$ M15]	
LE392	<i>supE44</i> <i>supF58</i> <i>hsdR514</i> <i>galK2</i>	[10]
	<i>galT22</i> <i>metB1</i> <i>trpR55</i> <i>lacY1</i>	
<i>A. brasilense</i>		
Yu62	<i>Ap</i> <sup>r</sup> <i>Nx</i> <sup>r</sup> , wild type	[11]
Sp7	<i>Ap</i> <sup>r</sup> <i>Nx</i> <sup>r</sup> , wild type	[2]
Plasmid		
pUC19	<i>Ap</i> <sup>r</sup>	[12]
pYPZ103	<i>Ap</i> <sup>r</sup> , <i>A. brasilense</i> <i>draTG</i> (6.7kb <i>Sal</i> I fragment) in pUC19	[4]
pLYM106	<i>Ap</i> <sup>r</sup> , <i>A. brasilense</i> Yu62 <i>draTG</i> (8.0kb <i>Sal</i> I fragment) in pUC19	This study
pBSKS(-)	<i>Ap</i> <sup>r</sup>	Stratagene corp.

### 1.2 培养方法

巴西固氮螺菌用 LD 培养基培养<sup>[4]</sup>, 其它菌株均用 LB 培养基培养。巴西固氮螺菌 30℃ 培养, 大肠杆菌 (*Escherichia coli*) 37℃ 培养。

### 1.3 包装蛋白的制备

用两个溶源菌制备包装提取物。提取方法按文献[13]进行。

### 1.4 DNA 的制备

细菌总 DNA 的提取按文献[14]进行。 $\lambda$  噬菌体、质粒和单链 M13 噬菌体 DNA 的提取按文献[13]进行。

### 1.5 核酸印迹分子杂交

探针的生物素标记及杂交和检测均按 BioRad 公司生物素标记的说明书进行。

### 1.6 DNA 序列测定及计算机分析

据双脱氧末端终止法<sup>[15]</sup>, 以单链 M13 噬菌体 DNA 为模板, 按 Perkin Elmer Cetus 公司 Ampli Taq DNA 聚合酶测序试剂盒说明书进行。用 Beckman 公司的 MicroGenie 序列分析软件进行序列拼接、分析等。

### 1.7 酶及试剂

核酸限制酶、DNA 聚合酶 Klenow 片段购自 Promega 公司和华美公司, T4DNA 连接酶购自 New England Biolabs 公司, 生物素标记 DNA 检测盒购自 BioRad 公司, Ampli Taq DNA 聚合酶测序试剂盒购自 Perkin Elmer Cetus 公司,  $\alpha$ -<sup>32</sup>P-dATP 购自生辉公司, Erase-

$\alpha$ -Base-System 购自 Promega 公司。

## 2 结果与讨论

### 2.1 总 DNA 酶切产物的 Southern 杂交分析

为了确定 *A. brasilense* Yu62 中是否有 draTG 同源基因,同源片段的定位情况及与 *A. brasilense* Sp7 的同源程度等。首先用 *A. brasilense* Sp7 的 4.0kb draTG 片段做探针<sup>[4]</sup>与 *A. brasilense* Yu62 总 DNA 各酶切产物进行 Southern 杂交,同时以 *A. brasilense* Sp7 总 DNA 的 Sal I 酶切产物和 EcoR I 酶切产物做对照。杂交结果表明 *A. brasilense* Yu62 的 draTG 同源片段分别定位在 8.2kb SalI, 6.9kb Xho I 和 12kb EcoR I 片段上。BamHI 酶切产物的杂交片段较大,大约在 30kb 左右。而相应的 *A. brasilense* Sp7 的杂交片段分别定位在 6.7kb Sal I (与 Zhang<sup>[4]</sup>等报道的相符)和 11.5kb EcoR I 片段上(图 1)。从杂交结果可以看出:*A. brasilense* Yu62 中有 draTG 同源基因,而且与 *A. brasilense* Sp7 的同源性很高,但这两个菌株之间仍有差异。

### 2.2 基因文库的构建及筛选

提取高分子量的 *A. brasilense* Yu62 总 DNA(>50kb),用稀释法确定其 Sau3A 部分酶切最佳条件,据此进行大量的部分酶切,然后用 10%~40% 的蔗糖密度超离心分离酶切片段,并进行分部收集。根据分部收集样品的电泳检测回收 9~23kb 的片段。提取包装蛋白,用标准浓度的  $\lambda$ DNA 进行包装试验,效价可达  $2 \times 10^8$  pfu/ $\mu$ g  $\lambda$ DNA。将已制备好的  $\lambda$ EMBL3/EcoR I + BamH I 双酶切产物(0.5 $\mu$ g/ $\mu$ l)与回收片段以 0.5 $\mu$ g:1.2 $\mu$ g(6 $\mu$ l 连接体积)的比例和浓度进行连接。将全部连接液进行包装,建立了 *A. brasilense* Yu62 的基因文库,容量为  $2.5 \times 10^5$  个重组子,并将此文库进行扩增、保存备用。

从未经扩增的基因库中随机挑选了 6 个噬菌斑,提取各自的 DNA 并进行 Sal I 酶切分析,结果表明这 6 个均为重组子,插入片段大小在 9~22kb 之间,平均约为 16kb。根据公式:  $N = \ln(1 - P) / \ln(1 - f)$  ( $N$  - 重组子个数;  $P$  - 希望获得某目的基因的概率;  $f$  - 目的基因长度在单倍体基因组总长度中所占的比例)计算,该基因文库的容量适合实验要求。

以生物素标记的 *A. brasilense* Sp7 4.0kb draTG 片段做探针,杂交筛选 *A.*

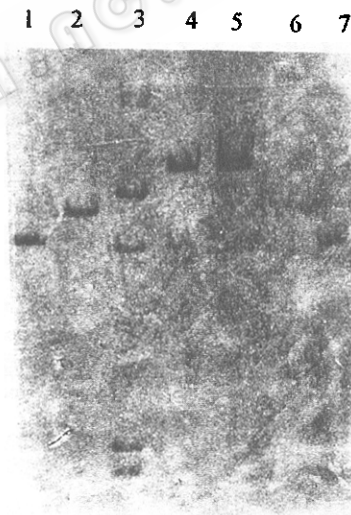


图 1 *A. brasilense* Yu62 总 DNA 酶切产物与生物素标记的 4.0kb draTG 片段的 Southern 的杂交  
Fig.1 Southern hybridization of several restriction endonucleases digested total DNA of *A. brasilense* Yu62 with Biotin-labelled 4.0kb draTG fragment of *A. brasilense*

Sp7.1. Sp7 DNA/Sal I ; 2. Yu62 DNA/Sal I ; 3. Biotin- $\lambda$ /Hind III marker(23.13, 9.42, 6.56, 4.36, 2.32, 2.03kb); 4. Sp7 DNA/EcoR I ; 5. Yu62 DNA/EcoR I ; 6. Yu62 DNA/BamH I ; 7. Yu62 DNA/Xho I

*brasilense* Yu62 的基因文库, 获得 10 个阳性噬菌斑, 分别称为 ETM1~3, ETM51~57。提取阳性斑的噬菌体 DNA, 进行酶切和杂交分析, 发现 *draTG* 定位在 8kb *Sal* I 片段上。

### 2.3 8kb *draTG* 同源片段的克隆与分析

将 8kb *Sal* I 片段克隆至 pUC19 上, 构成 pLYM106。用多种限制酶对 pLYM106 进行酶切分析, 建立了 8kb 插入片段的物理图谱(如图 2)。

用 *A. brasilense* Sp7 640bp *draT* 片段, 400bp *draTG* 下游片段<sup>[4]</sup> 及 500bp *A. brasilense* Yu62 *nifH* 片段<sup>[20]</sup> 做探针, 分别与该 8.0kb 片段的酶切产物做 Southern 杂交, 结果表明 3.0kb *EcoR* I - *Kpn* I 片段内含 *draTG* 基因, 而且 *draTG* 的上游与 *nifH* 基因相邻。

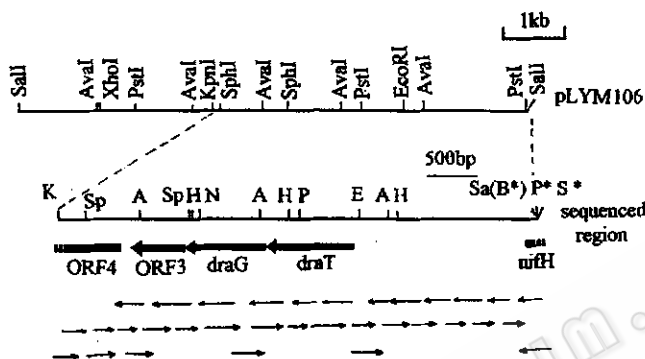


图 2 *A. brasilense* Yu62 *draTG* 基因区域(8.0kb *Sal* I 片段)的物理图谱

Fig. 2 Physical map of *draTG* region (8.0kb *Sal* I fragment) from *A. brasilense* Yu62

Heavy arrows indicate the coding regions and their transcribed direction. Light arrows indicate the sequencing direction and length. Restriction site: A - *Ava* I, B - *Bam* H I, E - *Eco* R I, H - *Hinc* II, N - *Nsi* I, P - *Pst* I, S - *Sal* I, Sa - *Sau* 3A I, Sp - *Sph* I (arsterisk indicate restriction sites from EMBL3 vector).

段内(G+C)%为 67.9%, 符合 *A. brasilense* 特点。序列分析结果表明该片段内有四个开放阅读框架(ORF)均具有较好的密码子利用率<sup>[21]</sup>。ORF1~3 的推测产物分别与其它菌的 *DraT*、*DraG* 及其下游 ORF 的推测产物同源。ORF4 的部分推测产物与大肠杆菌的 *YafJ* 同源。

在 *R. rubrum* 中, *draT* 基因以 TTG 作为起始密码子, 并且有两个终止密码子。 *A. brasilense* Yu62 *draT* 基因只有一个终止密码子, 推测可能与 *A. brasilense* Sp7 一样以 ATG 为起始密码子, 因为在该起始密码子上游 -9 至 -4 区域有 Shine-Dalgarno 序列 GAGAG, 推测即为核糖体结合位点(RBS)。 *A. brasilense* Yu62 的 *draT* 基因全长 888bp, 编码 296 个氨基酸, 其相应蛋白(DRAT)的分子量为 33.7kDa。而 894bp 的 *draG* 基因编码 298 个氨基酸, 相应 DRAG 酶的分子量为 32.4kDa。其起始密码子上游 -18 至 -12 区域的 ACGGGA 推测可能是 RBS, 但在 -75 ~ -70 之间也有一个典型的 SD 序列:

### 2.4 3.0kb *EcoR* I - *Kpn* I 片段的 DNA 序列测定

根据 8kb *draTG* 同源片段的酶切图谱, 酶切回收 3.0kb *EcoR* I - *Kpn* I 双酶切片段, 克隆至 M13mp19 上, 提取重组单链噬菌体的 RFDNA, 用 ExoIII 缺失法制备亚克隆, 提取各亚克隆的单链 DNA, 进行 DNA 序列测定。同时也酶切回收各种合适的片段分别亚克隆到 M13mp18/19 中, 进行 DNA 序列测定, 以确定一些缺失亚克隆没有测到的区域。最终获得了该插入片段共 2973bp 的核苷酸序列(如图 3), 其中 80% 以上为双向测序结果, 其余均是三次单向测序的结果。该片

GAGAG,该区域可能也是核糖体结合位点。*A. brasilense* Yu62 的 draT 和 draG 基因与 *A. brasilense* Sp7 及 *A. lipoferum* 的相应基因比较,有非常高的同源性(见后),说明这两个基因的保守性是很强的。紧接着 draG 的 ORF3,全长 450bp,其推测产物由 150 个氨基

```

EcoR I
TGTTTTCGAATTCGAAATGTATTGCGGGCTGCGCGACGTCGCCCTGCGGTGACGCGGACCCCTTGGCGGACGCGCGGTTTC 80
                                     RBS draT→ DPE1 160
GCGGTATCCATGAAGAATTTGTGGATTTGAGAGTGGCGCGGACGCGGTTCGGCGAGCGGTGGGTTCGGGATGCCGGG
                                     M A D G S A S G G L R D A R
GGCGGCTGATCGATCCGGATTGCGCGCTGGCGCTGAAGGCGCATCGCACCACTTGTCTGAACGTGACGGCGGAAGCGCTG 240
G R L I D P D S P L A L K A H R T N L L N V T A E A L
TGGCGGAGTCTCTTCAACGACGAGCGCGGCGCTGCGCATCGGCGGCACGCGCAGCGAAACGCGCTGCTGTTTCGAGGC 320
C A E S F N D E P R R L R I G G T R S E H A V L F E A
GCTCGACGAGAGCGCGAACAGCCTGATCGCGCGGACATCTTCAGCACTACATGGCCTTCACCTTCGGGCTGAACCCGG 400
L D E S P N S L I A A D I F Q H Y M A F T F G L N P
ACTTTTCCGGCGCGGAGGCGACGCGACGCGCGCGCTACCGCGCCAGCTACCTGCGCCTGCTGAAGGGGTGGCTGTTTC 480
D F S G A E G T D G R R R Y R A S Y L R L L K G W L F
GACAGCAACAACGCGGAGGGGGCTGTGCTGAAAGGCTGGGTGGAGAGCGCTTCGGTCTGCTGCCACCTATCACAAAGGA 560
D S N N A E G A V L K G W V E S R F G L L P T Y H K E
GCGGATCCTGCGCTTCGCTCCGACGCGTGGCGGACCTACGCGGAGGAGAAGGTTCGCCACCGCTTCACAAACAACAACA 640
P I L R F A S D A W R T Y G E E K V A T E F H N N N
Pst I
TCAACCTGCAGCTCGACCTGCTGTACGAATATTGCCAGTGGTGGATCGCGCGCGCTGGCCGGACGCGCGGATGCCGCTC 720
I N L Q L D L L Y E Y C Q W W M R R G W P D A P M P S
Hinc II
CGCGCCAGCCACATCCGCTCTACCGCGCGCTCAACGATTTGAGGAGCACCACATCGTCCGCGCGCCCGACAAGCGCAC 800
R A S H I R L Y R G V N D F E E H H I V A R P D K R T
GGCGGCTCTGCGCTCAACAACCTCAGCTCCTTTTCCATCGAGCGCGACATCGCCGCGCAGTTCGGGGACTACATCCTCG 880
A V L R L N N L S S F S I E R D I A G Q F G D Y I L
SD-like 960
AAACCTGGGTGCGCTCACCAAGGTGGTCTTCTCCGCGACATCCTGCCCGCTTACCCCTTCAAGCGCGGAGTAT
E T W V P L T K V V F F R D I L P R Y P F K G E G E Y
UAS RBS draG→
CTGGTCTGCGCGCGGACTACAGGTGGGGGTGTCGCTTCTCTGACAAACGATCACGGGAGGATCCGGCCCGGACTGA 1040
L V V G G D Y R V G V S L L M T D
Ava I
CCATTCCATCCGCTCGCGCGCGCTGGGCGCTATCTCGGGCTGGAATGCGGCGACGCGCTGGGCGCCACCGCTCGAGTTTC 1120
H S I R S R A L G A Y L G L D C G D A L G A T V E F
TGACAAAGGGCGAGATTGCCACCAATACGGCGTGCAACGACATCAAGGGCGCGGCTGGCTGAAGCTGCCCGCGGG 1200
L T K G E I A H Q Y G V H K H I K G G G W L K L P A G
CAGGTGACCGACGACACGAGATGTGATTCATCTCGGCGCGCCATCCTGGCGCGCCCGAGTGGGACGCGCGCGCGC 1280
Q V T D D T E M S I H L G R A I L A A P E W D A R R A
UAS
GGCGGAGGAGTTCGCGCTGTGCTGAAGGGCGTTCGGTGCATGTTCGGGACACGACAGCGCGCGCATCCGCGGCTCA 1360
A E E F A V W L K G V P V D V G D T T R R G I R R F
TCATGCACGGCACCTGTGCGGACCGGAGAGCGAGTATCACCGCGCAACGGGGCGCGGATGCGCAACCTGCGGTCGCC 1440
I M H G T L S E P E S E Y H A G N G A A M R N L P V A
CTGGCGACGCTGGGCGACGACGCGCGCTTCGAGCGCTGGACGCTGGAGCAGGCGCACATCACCCACTGCAACGCGATGTC 1520
L A T L G D D A A F E R W T V E Q A H I T H C N A M S
CGACGCGCGGACCTGACGCTCGGCCATATGGTGCGCGCTGCTGGTCTGGGCGGACGCTGCGCGACGTCGCGGACGAAT 1600
D A A T L T L G H M V R R L V L G G S V R D V R D E

```

CGAACAAGCTCATCGGAAGCACCGTCCGTTCAAATTCCAGCCCTACCGGGCGCTGGCCACCGCCTACATCGTCGATACG 1680  
S N K L I A K H R P F K F Q P Y R G L A T A Y I V D T  
Nsi I Hinc II 1760  
ATGCAGACGGTATGCATTATTATTTCCAGACCGACTCCGTGGAATCCTGCGTCGTGAGACCGTCAACCAGGGCGGGGA  
M Q T V M H Y Y F Q T D S V E S C V V E T V N Q G G D  
Sph I 1840  
CGCCGACACACCGCGCCCATCGCCGGCATGCTGGCCGGCGCCACCTACGGGTGGAGACCATCCCGCCGCGCTGGCTGC  
A D T T G A I A G M L A G A T Y G V E T I P P R W L  
Hinc II DPE2 1920  
GCAAGCTCGACCGCAAAGTCTACGACGAAATCTGCCCGCAGGTGGACGGGCTGCTCGCCCGCTCCCGCGGCTCAAGCAG  
R K L D R K V Y D E I C A Q V D G L L A R S P A L K Q  
RBS ORF3 → 2000  
GGGTGACCGGACCGGCGACCGTGTCTTCTTGGAAAAGCCCGCTGCGCCGGCAACAACCGCCGAAGGCCCTGCTG  
G M A D V I F F E K P G C A G N N R Q K A L L  
2080  
GGGAGGGCGGGCACACCGTCCACGCCCGGACCTGCTGTCGAGCCCTGGACGGCGGACCGCTGCGCCCTTCTTCGG  
A E A G H T V H A R D L L S E P W T A D R L R P F F G  
2160  
CGACCGCCCGGTGGCGAGTGGTTCACCGCGCCCGCCCGCGGTGAAGAGCGGGAGGTGATCCGACGCGGTGGACG  
D R P V A E W F N R A A P A V K S G E V D P D A L D  
2240  
AGGCCCGCGCCCTGGCGCTGATGCTGAAAACCGCGCTGCTGATCCGCGCTCGCTGATGCAGGTGGGGACCGCTGTGAT  
E A A A L A L M L K T P L L I R R P L M Q V G D R C D  
Avr I 2320  
TCCGATTTCGAGGCGGAGCGGTGGACGCGCTGGATCGGCTGGCCCGCGGAGCGCGGAGGCAAGTTGGAAGTGGCGCC  
C G F E A E R V D A W I G L A A G A P E G K L E V R P  
RBS  
GCCCGGGATGCCCGCTGCCACCACCGCAAGGGGTGACGGGACAACCGGACTGTGCTAGGTTTGGCCCGTCCCTG  
R R D A A L P T T R K G V T G Q P G L S  
ORF4 → 2480  
TTTGAACCGACATTTACGCGATGGGGCCGATCCATGTGCGAACTCCTGGGCGGAGCGCAACGTCGCCACGGACAT  
M S A N V P T D I  
2560  
CTGCTTCAGCTTTCCGGGCTGATGCGCGCGCGGGGAGACCGGTCCGCACCGCGACGGTGGGGAATCGCCTTTTACG  
C F S F R G L M R R G G Q T G P H R D G W G I A F Y  
2640  
AAGGAAGGGCTCCCGACCTTCCACGACCGCGCGGAGCTCGGAGTCGGAGATCGCGCGGCTGGTCAGCAGCTATTCCG  
E G K G C R T F H D P A P S C E S E I A R L V S S Y S  
2720  
ATCAAGTCTCGCTGGTGTCTCCACATCCGCGCGGCCAACCGCGCGGGTGTGCTGGAGAACCCACCCCTTCAC  
I K S C V V I S H I R R A N R G R V S L E N T H P F T  
2800  
GCGGGAGCTGTGGGGCGCGTCTGGACCTTCGCCCACAACGGCCAGCTCAAGGGCATCAAGGAACGGACGCTGACCTTCT  
R E L W G R V W T F A H N G Q L K G I K E R T L T F  
Sph I 2880  
ACGAACCACTCGGACGACGGACAGCGAGCAGCCTTCTGCTGGCTGCTCGACCAGATCCGATGCAGTATCCGGAGCCG  
Y E P V G T T D S E H A F C W L L D Q I R M Q Y P E P  
2960  
CCCAAGAGGAGCGCGCCCTGATCGGGTTGATCCGCGACCTCGCCACCGACCTCGGCACGCTGGGCGTCTTCAACATGCT  
P K S S G A L M R L I R D L A T D L G T L G V F N M L  
Kpn I 2980  
GCTCAGCGACGGCCGTACC  
L S D G R Y

图3 *A. brasiliense* Yu62 *draTG* 基因区域的核苷酸序列

Fig.3 Nucleotide sequence of *draTG* genes region of *A. brasiliense* Yu62

DPE: Putative RpoN dependent promoter; RBS: putative ribosome binding site; UAS: Upstream activator sequence; SD: Shine-dalgarno sequence. The potential start condons and SD sequences are shaded. The DPE and UAS are showed by underlining and overlining, respectively. The arrow indicate the orientation of transcription.

酸组成,分子量为 16.1kDa,其起始密码子上游 -16~-11 区也有典型的 RBS:AGGGG。该 ORF3 与 *A. lipoferum* 和 *R. rubrum* 相应部位的 ORF 同源<sup>[5,6]</sup>。ORF4 与 ORF3 相隔 70bp,在这 70bp 内未发现保守的 SD 序列,但在 ORF4 起始密码子上游 -82~-87 区(ORF3 的编码区内)发现较典型的 SD 序列:AGGGG,推测是该 ORF 的核糖体结合位点。ORF4 是不完整的,其终止密码子应在 3.0kb 片段的下游。由该 ORF 推测的部分产物与大肠杆菌 YafJ 的前 181 个氨基酸的相似值为 77%,yafJ 的具体功能尚不清楚(Ohmori, H. 等未发表资料)。

在 *R. rubrum*, *A. lipoferum* 和 *A. brasilense* Sp7 中, draTG 及下游 ORF3 相隔都很近,因而推测它们是共转录的, *A. brasilense* Yu62 中的情况也如此。但 *R. rubrum* 中有迹象表明它们的转录方式并不是那么简单<sup>[19]</sup>。在 *A. brasilense* Yu62 的 draG 及其下游 ORF3 的上游区域均有  $\sigma^{54}$ -依赖型启动子的特征序列(DPE):GG-N<sub>10</sub>-GC(见图 3 的 DPE1 和 DPE2),而且在 ORF3 上游区域还发现了两个上游激活序列(UAS:NifA 结合部位的保守序列):TGT-N<sub>10</sub>-ACA(见图 3)。因而推测 draG 和 ORF3 与 draT 一起共转录之外,还可单独转录。

2.5 draT 和 draG 基因同源性分析

将 *A. brasilense* Yu62 draT 基因 DNA 序列分别与 *A. brasilense* Sp7<sup>[4]</sup>和 *A. lipoferum*<sup>[5]</sup>的相应基因比较,发现同源性非常高(分别为 97.4%和 96.7%)。将上述三个菌的 draT 基因推出的氨基酸序列进行比较,发现仅有几个氨基酸发生改变,而且均在 N-端(第 103 个氨基酸之前),中间区域的氨基酸很少有变化。此外,还将 *A. brasilense* Yu62 draT 基因推测的相应氨基酸序列与 *R. rubrum*<sup>[6]</sup>和 *Rhodobacter capsulatus*<sup>[7]</sup> DraT 氨基酸序列进行了比较,结果见表 2。

*A. brasilense* Yu62 draG 基因 DNA 序列与 *A. brasilense* Sp7、*A. lipoferum* 的 draG 基因同源性也很高,分别为 97.5%和 97%。这三个菌的 DraG 氨基酸序列之间也仅有几个氨基酸的差别,而且有意思的是, Yu62 菌株的 DraG 与 *A. lipoferum* 的同源性最高(表 3)。同样也将 *A. brasilense* Yu62 DraG 氨基酸序列与 *R. rubrum*、*Rhodobacter capsulatus* 相应蛋白的氨基酸序列进行了比较,结果见表 3。

以上的同源比较说明 *Azospirillum* 的 DraT 和 DraG 是非常保守的,不仅菌株间的差异小,而且种间差异也很小。

2.6 draG 下游 ORF3 的同源比较

表 2 几种固氮菌 DraT 相应氨基酸序列的同源百分率

Table 2 Percentage sequence matches calculated for the deduced amino acid sequences of draT from different nitrogen-fixing bacteria

	Ab(Yu62)	Ab(Sp7)	Al	Rr	Rc
Ab(Yu62)	*	99	98	49.5	31.8
Ab(Sp7)		*	97.3	49.8	31.8
Al			*	49.2	32.6
Rr				*	49.3
Rc					*

Ab: *A. brasilense*, Al: *A. lipoferum*, Rr: *R. rubrum*, Rc: *Rhodobacter capsulatus*

表 3 几种固氮菌 DraG 相应氨基酸序列的同源百分率

Table 3 Percentage sequence matches calculated for the deduced amino acid sequences of draG from different nitrogen-fixing bacteria

	Ab(Yu62)	Ab(Sp7)	Al	Rr	Rc
Ab(Yu62)	*	97.7	98.1	60.6	54
Ab(Sp7)		*	98.0	60.5	53.7
Al			*	60.8	53.8
Rr				*	61.1
Rc					*

Ab: *A. brasilense*, Al: *A. lipoferum*, Rr: *R. rubrum*, Rc: *Rhodobacter capsulatus*

C-端保守性较差,因为比较 *A. braslense* 和 *A. lipoferum* ORF3 的氨基酸序列,发现在前 126 个氨基酸的同源性为 97%,而整个编码区的同源百分率只有 83%。与 *R. rubrum* 的比较,也有同样的结果;前 121 个氨基酸的同源百分率大于 70%,但整个编码区的同源百分率只有 57.3%(表 4)。

该 ORF3 的功能尚不清楚,但在 *R. rubrum* 中,该 ORF 的突变株在黑暗条件下的 switch-off 速度较野生型更快,DRAG 蛋白量减少<sup>[17,18]</sup>。这说明它可能有抑制 DRAT 活性而稳定 DRAG 酶的功能。Zhang 等<sup>[19]</sup>对 DRAT 和 DRAG 酶活性调控的研究表明,在有光或氮饥饿时,有某种小分子量的信号组份可能参与这两个酶的活性调控过程。该

ORF3 与 draTG 紧密相联,且其上游区域有  $\sigma^{54}$  依赖启动子的特征:具有类似 DPE 的序列和 UAS(在 draG 基因编码区内,如图 3),如果这些序列确实是该 ORF 的启动子组份,那么在固氮条件下,该 ORF 显然可以单独转录。这使我们联想到,该 ORF3 编码的蛋白是否就是参与 DRAT 和 DRAG 活性调控的小分子信号感受蛋白? 这值得进一步研究。

本文研究的 *A. brasilense* Yu62 菌株是 1984 年杨洁彬等<sup>[11]</sup>从北京郊区玉米根际分离得到的。本文的研究结果不仅将有助于精确阐明巴西固氮螺菌固氮酶翻译后活性调控机制,同时也为应用于我国农业生产的抗铵菌株的构建打下了基础。

参 考 文 献

[1]Pedrosa F O. CRC Crit Rev Plant Sci. 1985, 6:345~384.

将 draG 下游的 ORF3 的 DNA 序列与 Genbank 进行比较,发现该 ORF 不仅与 *A. li-*  
*opferum* 及 *R. rubrum* 相应位置的 ORF 同源,而且与 *Azo-*  
*tobacter vinelandii* nifB 和 nifQ 基因间的 ORF14 同源,该 ORF14 的功能尚不清楚<sup>[17]</sup>。

由 ORF3 推测出的氨基酸序列与上述三菌的相应 ORF 推出的氨基酸序列的同源百分率见表 4。氨基酸序列的比较发现该 ORF 编码的蛋白,其

表 4 *A. brasilense* Yu62 draTG 下游 ORF 相应氨基酸序列与其它固氮菌相应 ORF 氨基酸序列的同源百分率

Table 4 Percentage sequence matches calculated for the deduced amino acid sequences of draTG downstream ORF from *A. brasilense* Yu62 and ORF from other nitrogen-fixing bacteria

	Ab	Al	Rr	Av(ORF14)
Ab	*	83.3	57.3	44.8
Al		*	56.6	46.3
Rr			*	38.0
Av				*

Ab: *A. brasilense*, Al: *A. lipoferum*, Rr: *R. rubrum*, Av: *A. vinelandii*



- [2]Hartmann A, Fu H A, Burris R H. J Bacteriol, 1986, 165:864~870.
- [3]Ludden P W, Roberts G P. Curr Top Cell Regul, 1989, 30:23~55.
- [4]Zhang Y P, Burris R H, Roberts G. J Bacteriol, 1992, 174, 3364~3369.
- [5]Inoue A, Shigematsu T, Hidaka M, et al. Gene, 1996, 170:101~106.
- [6]Fitzmaurice W P, Saari L L, Lowery R G, et al. Mol Gen Genet, 1989, 218:340~347.
- [7]Masepohl B, Krey R, Klipp W. J Gen Microbiol, 139:2667~2675.
- [8]Hohn B, Murray K. Proc Natl Acad Sci USA, 1977, 82:7525~7529.
- [9]Messing J. Recomb DNA Tech Bull, 1979, 2(2):43.
- [10]Borck K. Mol Gen Genet, 1976, 146:199~207.
- [11]杨洁彬, 曹增良, 李季伦. 北京农业大学学报, 1984, 10(3):321~329.
- [12]Yanisch-Perron C. Gene, 1985, 33:103~119.
- [13]Sambrook J, Maniatis T, Fritsch E F. Molecular Cloning: A Laboratory Manual, 2nd ed. Cold Spring Harbor Laboratory Press, 1989.
- [14]Rodriguez R L, Tait R C. Recombinant DNA Techniques: An Introduction, The Benjamin/Cummings Publishing Company, 1983.
- [15]Sanger F. Proc Natl Acad Sci USA, 1977, 74:5463~5467.
- [16]Joerges R D, Bishop P E. J Bacteriol, 1988, 170:1475~1487.
- [17]Lies D P. In 9th International Congress on Nitrogen Fixation: Program and Abstracts, 1992, EXS476
- [18]Liang J H, Nielsen G M, Lies D P, et al. J Bacteriol, 173:6903~6909.
- [19]Zhang Y P, Grunwald S K, Lies D P, et al. In Nitrogen Fixation: Fundamentals and Applications, Tikhonovich IA & Provorov NA (eds), Kluwer Academic Publishers, Dordrecht, Boston, London, 1995, pp177~182.
- [20]阎大来, 何路红, 马旅雁等. 生物工程学报, 1995, 11, 205~210.
- [21]Fani R, Allotta G, Bazzicalupo M. Mol Gen Genet, 1989, 220:81~87.

## Cloning and Sequencing of draTG Genes and Their Downstream Region of *Azospirillum brasilense* Yu62

Ma Lüyan Li Jilun

(National Laboratories for Agrobiotechnology, Beijing Agricultural University, Beijing 100094)

**Abstract** A 8kb fragment was cloned by probing the gene library of *Azospirillum brasilense* Yu62 with the 4.0kb draTG fragment of *A. brasilense* Sp7. DNA hybridization of this fragment demonstrated that draTG genes were located in a 3.0kb EcoR I - Kpn I fragment, and contiguous to nifH gene. This 3.0kb fragment was completely sequenced on both strands. Sequence analysis of the fragment revealed that it included the full length draTG genes and two ORF downstream of draG (ORF3 and ORF4, ORF4 is incomplete). The draTG and downstream ORF3 are assumed to be co-transcribed as a single operon. Promoter element analysis of the sequenced region showed that there were some elements of  $\sigma^{54}$ -dependent promoter (DPEs and UASs) in upstream region of draG and ORF3. This suggests that the draG and ORF3 might be transcribed independently except for being cotranscribed with draT. Both the DNA and amino acid sequences of the draTG genes from *A. brasilense* Yu62 were compared with those of other nitrogen-fixing bacteria. The results showed the DraTG were highly conservative. There were only few changes among strains and/or species. Homology analysis of ORF revealed that the ORF3 immediately downstream of draG was homologous not only with the ORF at the same position in *Rhodospirillum rubrum* and *A. lipoferum*, but also with the ORF14 of *Azotobacter vinelandii*, and the ORF4 showed extensive similarity to yafJ of *Escherichia coli*.

**Key words** *Azospirillum brasilense*, draTG, ORF, nucleic acid sequence