

**CLONING, PROKARYOTIC EXPRESSION AND PROTEIN
PURIFICATION OF SUGARCANE ENDOGENOUS AZOTOBACTER
KLEBSIELLA VARIICOLA DX120E NIFK GENE**

**Kun-Kun Zhang^{1#}, Yong-Xiu Xing^{1,2#}, Min Shao¹, Chao-Xing Wu¹, Li-Tao Yang^{1,2*},
Yang-Rui Li^{1,2*}**

¹State Key Laboratory for Conservation and Utilization of Subtropical Agro-Bioresources, College of Agriculture, Guangxi University, Nanning 530004, China.

² Guangxi Key Laboratory of Sugarcane Genetic Improvement, Key Laboratory of Sugarcane Biotechnology and Genetic Improvement (Guangxi), Ministry of Agriculture, P.R. China, Sugarcane Research Center, Chinese Academy of Agricultural Sciences-Guangxi Academy of Agricultural Sciences, Nanning 530007, China.

ABSTRACT

NifK gene is a key nitrogenase and plays an important role in the nitrogen fixation process. It is necessary to obtain the nifK gene for exploring its relationship with nitrogen fixation. This study aimed to clone nifK gene, predict the secondary structure, the characters and functions of the nifK protein, and conduct prokaryotic expression. According to other nifK sequences registered in the NCBI, specific primers were designed. Klebsiella variicola DX120E nifK gene opening reading frame (ORF) was cloned by PCR amplification, and its nucleotide sequence, amino acid sequence and protein structure were analyzed by combining bioinformatics methods. Prokaryotic expression vector pET30a (+) was used to construct the recombinant expression vector pET30a-nifK. After PCR, double enzyme digestion and DNA sequencing, the recombinant plasmid was transformed into BL21 (DE3), and 1.0 mmol/L IPTG was used to induce the gene expression at 28 °C. The fusion protein expression was detected through SDS-PAGE electrophoresis. The nifK gene ORF cloned from Klebsiella

variicola DX120E in this study was 1563 bp, encoding 520 amino acids. The prokaryotic expression protein by induction was detected by mass spectrometry, which had an isoelectric point of 5.54, and molecular weight of 58.12 kDa. The gene has been registered in GenBank with accession number KF732647. Amino acid sequence comparison and phylogenetic tree analysis showed that the *nifK* gene had the highest similarity with *Klebsiella*. The *nifK* gene was cloned and the *nifK* protein was successfully expressed in a prokaryotic expression system, which laid the foundation for further studying the function of *DX120E nifK* gene in sugarcane.

Key words: Sugarcane; *Klebsiella variicola* DX120E; clone of *nifK* gene; prokaryotic expression; protein purification

Introduction

Sugarcane (*Saccharum officinarum* L.) is one of the most important sugar and energy crops. It has great economic importance due to its application in the food industry, and is also particularly valuable for its use in the production of ethanol (Menossi *et al.*, 2008). Originated in Asia, sugarcane is highly productive in the tropical and subtropical areas of the world. However, many factors affect the growth and development of sugarcane plants (Lisson *et al.*, 2005; Inman-Bamber *et al.*, 2005). For example, sugarcane production in China is facing many serious impediments, such as high nitrogen (N) fertilization and degeneration of the main cultivars (Wei *et al.*, 2014a). More than 60 % of the sugarcane fields in China are fertilized with over 500-700 kg N ha⁻¹ year⁻¹ (Li, 2010) whereas sugarcane fields in Brazil are fertilized with 60–70 kg N ha⁻¹ year⁻¹ (Urquiaga *et al.*, 2012). The much higher amount of nitrogen fertilizer input not only caused high sugarcane production costs but also produced serious adverse impact on the environment. Therefore, in China, it is imperative to reduce the N-fertilization. Biological nitrogen fixation is responsible for supplying more than 60% of the world's annual resources of new ammonia (Schlesinger *et al.*, 1991). This process is performed by nitrogenase,

which is composed of two metalloproteins: MoFe protein (ecoded by *nifD* and *nifK*) and Fe protein (ecoded by *nifH*) (Rubio *et al.*, 2005). The nucleotide sequence for coding regions of *nifHDK* genes and their respective genes are remarkably highly conserved among all nitrogen-fixing organisms (Dean *et al.*, 1992). These genes have been cloned and sequenced from many different N₂-fixing microorganisms. For example, The complete nucleotide sequence of the *nifK* open reading frame (ORF) from *Cyanothece* sp. ATCC 51142 is 1533 bp in length and encodes a polypeptide of 511 amino acids (Colón-López *et al.*, 1999). The β subunit of *Azotobacter vinelandii* MoFe protein has 522 amino acids (Brigle *et al.*, 1985). Strain DX120E was isolated from surface-sterilized roots of the main sugarcane cultivar ROC22 grown in Daxin, Guangxi, China. It is a Gram-negative bacterium and can able to fix N₂, produce siderophores and indole-3-acetic acids, and solubilize Ca₃(PO₄)₂ (Lin *et al.*, 2012). At present, there are only few reports about molecular aspect in this bacteria, more about inoculation with sugarcane. For example, under greenhouse condition, inoculation of strain DX120E in sugarcane cultivar GT21 for 7 days showed improved dry weight, and N, P and K contents obviously (Wei *et al.*, 2014a). Inoculation of bacteria DX120E in sugarcane cultivars B8 and GT21 effectively promoted the plant growth and nutrient uptake, significantly improved the nitrate reductase (NR) and glutamine synthetase (GS) activities, and the nitrate concentration in certain degree in leaves (Wei *et al.*, 2014b). Inoculation of strain DX120E in ROC22 and B8 showed increases in chlorophyll, N and P contents (Xing *et al.*, 2014). In this study, *nifK* gene was cloned and the protein was induced to lay foundation for further studying the molecular mechanism of strain DX120E promoting sugarcane growth, antibody preparation, enzyme activity in vitro and western blot analysis.

Materials and methods

Materials

Experimental materials

The tested bacterial strain *Klebsiella variicola* DX120E was isolated from

surface-sterilized roots of ROC22 sugarcane planting in Daxin county, Guangxi, China.

Experiment reagents

EcoRI and XhoI restriction enzymes were bought from Fermentas Company (Shanghai); DNA Ligation Kit Ver. 2.0 and pMD18-T vector were from TaKaRa Company (Dalian); Biospin Gel Extraction Kit was purchased from BioFlux Company (Beijing); DH5 α and BL21 (DE3) were purchased from Beijing Gold Biotechnology Company; Prokaryotic expression vector pET-30a(+) were kept by the laboratory. Other conventional reagents were analytical grade reagents.

Experimental methods

Primer Design

According to *nifK* gene nucleotide sequences of other nitrogen-fixing species registered in GenBank, online software program Vector NTI Advance 11.0 was used to analyze the homology. The gene upstream primer 5'- ACGAATTCATGAGCCAAACGATTGATAA-3' (underline parts are restriction sites EcoRI), and downstream primer 5'- ATA CTCGAGTTAACGGACGAGATCGAAGCTG-3' (underline parts are restriction sites XhoI) were designed. The primers were synthesized by Sangon Corporation (Shanghai).

nifK gene cloning

Klebsiella variicola DX120E was used as the template for PCR amplification. The total volume of the PCR reaction system for amplification of *nifK* gene was 25 μ L, operation was done following the EsTaq polymerase instructions. The PCR reaction parameters included pre-denaturation at 95°C for 5 min; denaturation for 50 s at 94°C; annealing for 30 s at 68°C, and extending for 1 min at 72°C, 5 cycles; followed by 15 cycles of denaturation for 50 s at 94°C; annealing at 63°C for 30 s, and extension for 1 min at 72°C, then 15 cycles of denaturation for 50 s at 94°C; annealing for 30 s at 58°C, and extension for 1 min at 72°C, and the final extension for 10 min at 72°C. When the reaction finished, the PCR products were analyzed with 1.0% agarose gel electrophoresis. Then *nifK* gene was connected with pMD18-T vector at 16°C for 8

h, and then transformed into DH5 α competent cells by heat shock. Positive clones were screened out, and sequenced by Shanghai Sangon Company, China.

Bioinformatic analysis of nifK gene

The amino acid sequence of the gene was predicted by BioXM2.6, and basic physical and chemical properties of the speculated protein sequence of the gene were analyzed by ExPASy (<http://expasy.org/tools/>). The *nifK* gene and the homology with other species were analyzed by GenBank Blast (<http://blast.ncbi.nlm.nih.gov>) online analysis software. [Http://www.cbs.dtu.dk/services/SignalP/](http://www.cbs.dtu.dk/services/SignalP/) was used for protein signal peptide analysis. Online software <http://www.cbs.dtu.dk/services/TMHMM/> was used for membrane protein analysis. Protein blast program (<http://blast.ncbi.nlm.nih.gov/Blast.cgi>) and Motif Scan software were used to analyze the functional domains of the nifK protein. The phylogenetic tree of *nifK* gene amino acid sequences of DX120E with other species were constructed by MEGA6.0 software. The Vector NTI Advance 11.0 was used for amino acid sequence homology analysis.

Construction of nifK prokaryotic expression vector

The PCR products of *Klebsiella variicola* DX120E *nifK* gene were recovered and purified. The plasmids of expression vector pET-30a(+) were extracted, then double digestion was done by using EcoRI and XhoI, and the digestion products were recovered. The target gene and expression vector were connected by DNA Ligation Kit, and the ligation products were transformed into the expression strain DH5 α by heat shock. The correct recombinant plasmids bacteria were detected and sent to Shanghai Biological Engineering Company for sequencing, and the recombinant plasmids were verified by double digestion.

Prokaryotic expression of recombinant plasmid pET30a-nifK

The recombinant strain and the empty pET-30a strain were both transformed into the liquid LB medium with Kan 100 mg \cdot L⁻¹ to induce gene expression separately at 28°C 200 r/min. When OD₆₀₀ reached 0.6, the medium was added with 1.0 mmol/L IPTG for induction, and 2.0 mL medium were collected at 0, 1, 2, 3, 4 and 5 h, respectively. After completion of the induction, the product was centrifuged at 12000 \times g for 20 min, the supernatant and precipitate

were collected respectively, then 8 μ L 3 \times loading buffer were added and kept in boiling water bath for 5 min, and 25 μ L samples were taken for SDS-PAGE electrophoresis after 12000 \times centrifugation for 5 min. The concentrations of spacer gel and separation gel for SDS-PAGE was 4% and 12.5%, respectively. After electrophoresis, the gel was stained with Brilliant blue R-250 for imaging analysis.

Results

nifK gene cloning and sequence analysis

Klebsiella variicola DX120E was used as the template with specific primers for PCR amplification. After recycling, ligation and transformation of target gene, a 1563 bp sequence was obtained, which matches the expected size (Fig1).

Compared the obtained sequence with those in GenBank Blast, the result showed the 1563 bp fragment was the open reading frame (ORF) sequence of the *nifK* gene encoding 520 amino acids (Fig. 2), which was registered in GenBank with the accession number KF732647.

Bioinformatic analysis of *nifK* gene

Expasy was used to analyze the *nifK* protein and it showed the molecular weight of *nifK* protein was 58.16 kDa with pI of 5.48, which was consistent with the mass spectrometric result. The *nifK* protein was an acidic protein. SOPMA was used to predict the secondary structure of the *nifK* protein, and the results showed that the protein contained α -helix, β strand, β folding and curl, and α -helix accounted for 42.88%, curl accounted for 38.46%, β folding accounted for 13.27%, while β strand accounted for 5.38% only (Fig. 3). SigalP 4.1 predictions showed that the protein had no signal peptide. The transmembrane helix of the protein was predicted using TMHMM online software and showed that it was an outer membrane protein (Tab. 1). NifK protein functional domains were analyzed through Motif Scan software, and the results showed that the sites 29-32, 47-50, 116-119, 194-197, 225-228, 260-263, 273-276 and 280-283 were casein kinase II phosphorylation sites; 73-78, 90-95,

125-130, 133-138, 190-195, 318-323, 384-389 and 441-446 were N-myristoylation sites; 298-300 and 333-335 were protein kinase C phosphorylation sites; 404-412 were Tyrosine kinase phosphorylation sites, 136-139 were asparagines glycosylation binding sites, 87-94 and 150-164 were nitrogenase component alpha/beta subunits (Fig. 3). The data in Fig. 4 showed conserved domains of *nifK* gene which revealed *nifK* was a highly conserved gene in the functional region. MEGA6.0 was used to construct a phylogenetic tree (Fig. 5), and the Vector NTI Advance 11.0 was used to make amino acid sequence homology analysis (Fig. 6). Both results showed that *nifK* protein has higher homology with *Klebsiella*, but relatively lower with non- *Klebsiella* species.

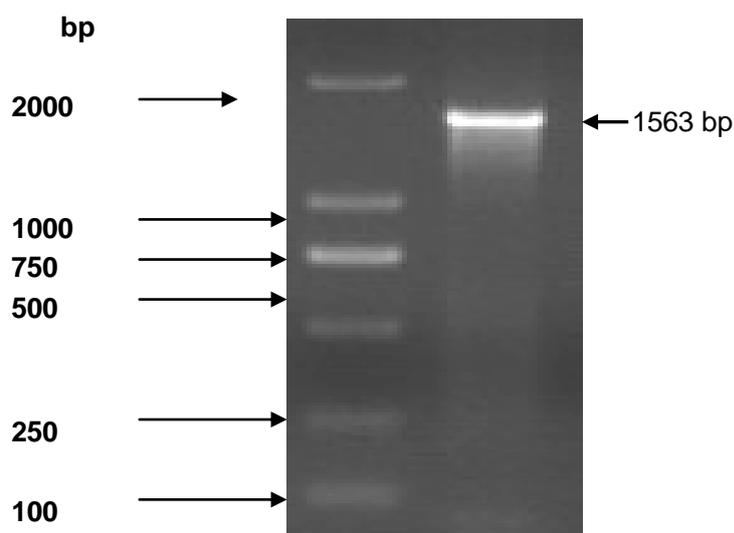


Fig. 1. PCR product of *nifK* gene. M: 2000bp; 1: *nifK* gene

```

1 ATGAGCCAAACGATTGATAAAATTCACAGCTGTTATCCGCTGTTTGAACAGGATGAATAC
M S Q T I D K I H S C Y P L F E Q D E Y
61 CAGACCCGTGCCAGAATAAAAAGACCCCTTGAAGAGGCGCAGCAGCCGCGTGTGCAG
Q T L F Q N K K T L E E A H D A Q R V Q
121 GAGGTTTTGCTGGACCACCACCGCGAGTATGAAGCGCTGAACTTCCAGCGCAAGCG
E V F A W T T T A E Y E A L N F Q R E A
181 CTGACTGTCGACCCGCCAAAGCTGCCAGCCGCTCGGCGCGTACTCTGCGCGTGGGG
L T V D P A K A C Q P L G A V L C A L G
241 TTCGCGGCAACCCCTGACCTGACGCGCTCCAGGGTGCCTGCGCTATTTTCGCACC
F A G T L P Y V H G S Q G C V A Y F R T
301 TACTTTAACCGCCATTTAAAGAGCCTGTCGCTGCGTCTCCGACTCCATGACCGAGGAC
Y F N R H F K E P V A C V S D S M T E D
361 GCGGCGGTGTTGCGCGCAACAACAACATGAATCTGGGCTGCAGAATGCCAGCGCGCTG
A A V F G G N N N M N L G L Q N A S A L
421 TATAAACCCGAGATTATCCCGCTCTCCACCACCTGTATGGCCGAGGTGATCGGTGACGAT
Y K P E I I A V S T T C M A E V I G D D
481 CTGCAGGCGTTTATGCCAACGCCAAAAAGAGGGATTGTTGACGACAGCATCGCCATT
L Q A F I A N A K K E G F V D D S I A I
541 CCTTACGCCCATACCCCGCTTATCGGACGCAATGTCACCGGCTGGGACAATATGTC
P Y A H T P S F I G S H V T G W D N M F
601 GAAGGGTTCGCGAAGACCTTTACTGCTGACTACCGCGGCGAGCCGGGCAACAGCAAAAAG
E G F A K T F T A D Y A G Q P G K Q K
661 CTCAATCTGGTGACCGGATTGAGACCTATCTCGGCAACTTCGCGTGTGAAAGCGGATG
L N L V T G F E T Y L G N F R V L K R M
721 ATGGCGCAGATGGATGCCGTCAGCCGCTCTCCGACCCATCAGAGGTGCTCGACACC
M A Q M D V P C S L L S D P S E V L D T
781 CCCGCGACGGCCATTACCGGATGACCGCGGCGCACCAGCCAGGAGGATCAAAACC
P A D G H Y R M Y A G G T S Q E E I K T
841 GCGCGGACCGCATTGACACCCGTGCTGCGACCGCTGGCAGCTGGTAAAAGCAAAAAG
A P D A I D T L L L L Q P W Q L V K S K K
901 GTGGTTCAGGAGATGTGGAACCCCGCCACCGAGGTGGCGTTCCGCTGGGCGTGGCC
V V Q E M W N Q P A T E V A V P L G L A
961 GCCACCGACCGCTGCTGATGACCGTCACTCAGCTGACCGGCAACCGATTGCGGATGCC
A T D A L L M T V S Q L T G K P I A D A
1021 CTGACCCCTGAGCGCGCGGCTGGTGCACATGATGCTGGATTCCCACACCTGGTGCAT
L T L E R G R L V D M M L D S H T W L H
1081 GGCAAAAATTCGCGCTCTACGGCGATCCGGATTTCGTGATGGGCTGACCGCTTCCCTG
G K K F G L Y G D P D F V M G L T R F L
1141 CTGGAGCTGGGCTGCGAGCCGCGGTGATCCTCAGCCATAACGCCAATAAACCGTGGCAA
L E L G C E P T V I L S H N A N K R W Q
1201 AAAGCGATGAAGAAAATGCTTGATGCCTCGCGTACGGTACGAAAGCGAAGTTCATC
K A M K K M L D A S P Y G Q E S E V F I
1261 AACTGCGACTGTGGCACTTCCGCTGCTGATGTTACCCGTCAGCCGACTTTATGATC
N C D L W H F R S L M F T R Q P D F M I
1321 GGTAACCTCCTACGGCAAGTTTATCCAGCGCGATACCCTGGCAAAAGGGCAAAAGCCTTCGAA
G N S Y G K F I Q R D T L A K G K A F E
1381 GTGCCGCTGATCCGCTGGGCTTTCCGCTGTTCCAGCCCATCATCTGCACCGCCAGACC
V P L I R L G F P L F D R H H L H R Q T
1441 ACCTGGGCTATGAAGGCGCAATGAACATCGTCAGCAGCTGGTGAACCCGCTGCTGGAA
T W G Y E G A M N I V T T L V N A V L E
1501 AAAGTGGACCGACACCGAGCCAGTTGGGCAAAACCGATTACAGCTTCGATCTCGTCCGT
K L D H D T S Q L G K T D Y S F D L V R
1561 TAA
*
```

Fig. 2. cDNA sequence and deduced amino acid sequence of *Klebsiella variicola* DX120E *nifK* gene

Table 1. The analysis outer membrane helix of *nifK* gene

Facts	Number
Sequence length	520
Sequence Number of predicted TMHs	0
Sequence Exp number of AAs in TMHs	1.93885

Sequence Exp number, first 60 AAs	2e-05
Sequence Total prob of N-in	0.09190
Sequence TMHMM2.0 outside	1-520

10 20 30 40 50 60 70

```

MSQTIDKIHSCYPLFEQDEYQTLFQNKKTLEEAHDAQRVQEVFAWTTTAEYEALNFQREALTVDPKACQ
ccccchccccccccchhhhhhhhtccccccccchhhhhhhhhhhhhhhhhhhhhccheeeeeccccccc
PLGAVLCLAGFAGTLPYVHGSQGCVAAYFRITYFNRHFKEPVACVSDSMTEDAAVFGGNMNMNLGLQNASAL
chhhhhhhhtccccceeecccttchhhhhhhhhhhccccchhhhhhhhhhhheetcchhhhhhhhhhhhh
YKPEIIAVSTTCMAEVIGDDLQAFIANAKKEGFVDDSLAIPYAHTPSFIGSHVTGWDMNMFEGFAKFTAD
cccteeeeehhhhhhhhtcchhhhhhhhhhtccccccccceeeccccccccchhhhhhhhhhhhhhhcc
YAGQPGKQKLNLVITGFETYLGNFRVLRKRMMAQMDVPCSLSDPSEVLDTPADGHYRMYAGGT SQBEIKT
ccccccccceeeccccccccchhhhhhhhhhtccccccccccccccccccccceccccchhhhhhh
APDAIDTLLLPWQLVKSKKVVQEMWNPATEVAVPLGLAATDALLMTVSQLTGKPIADALTLERGRLVD
htchhhheeeccccchhhhhhhhtccccceccccccccchhhhhhhhhhtcccchhhhhhhhhhhhh
MMLDSHTWLHGKFKGLYGPDPFVMGLTRFLELGCPTVILSHNANKRWQKAMKMLDASPYGQSEVFI
hhhhhhhhettceeeeeecccthhhhhhhhhhhtccccceccccccccchhhhhhhhhhhccccccccceec
NCDLWHRSLMFTRQPDFMIGNSYKGFQRD TLAKGKA FEVPLIRLGFPLFDRHHLHRQTTWGYEGAMNI
cchhhhhhhhhhtccccceeeccctchhhhhhhhtcccchheeeeecccccccccccccccccecccthhh
VTTLVNAVLEKLDHDTSQLGKTDYSFDLVR
hhhhhhhhhhhtttccccccccceeeeh
    
```

Fig. 3. Secondary structure prediction of nifK protein

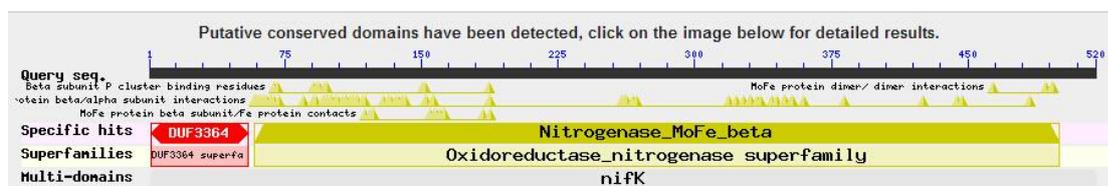


Fig. 4 Conserved domains of nifK

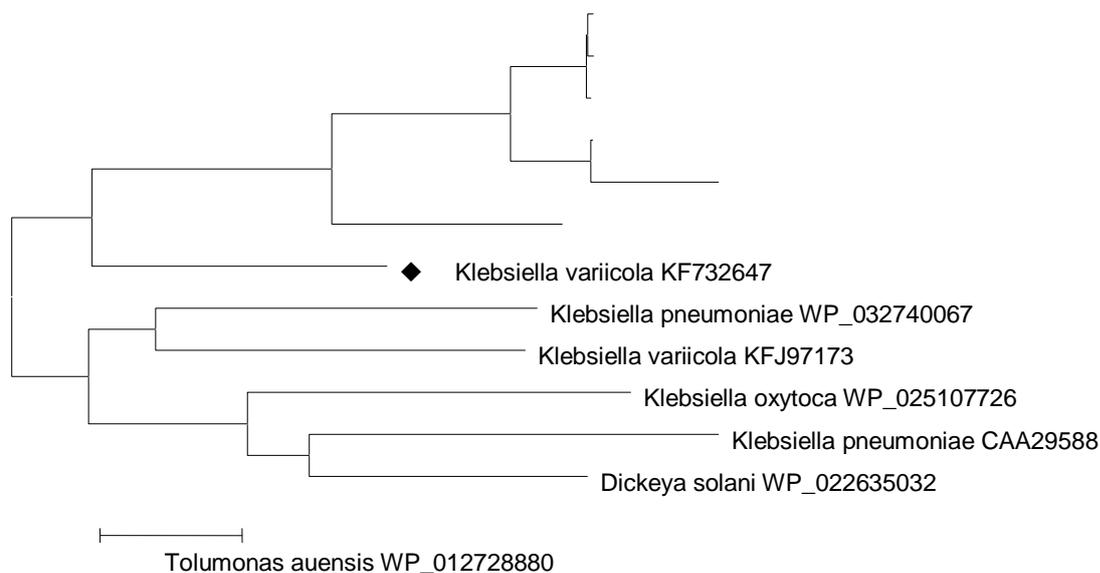
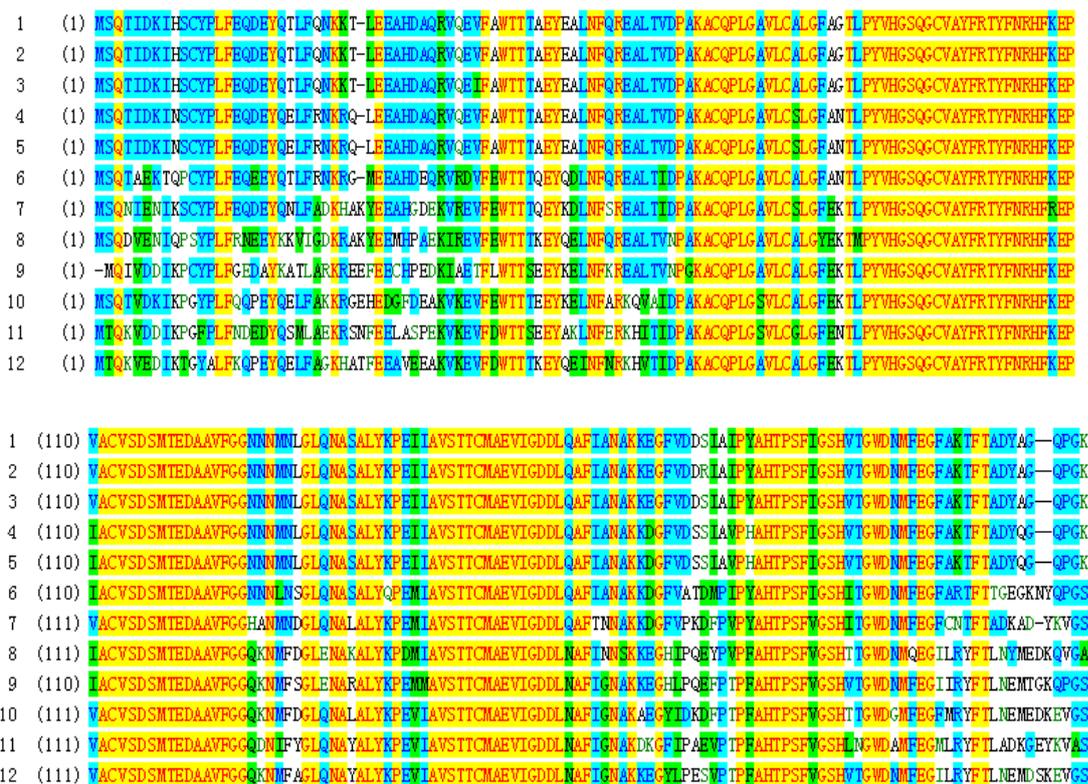


Fig. 5. Phylogenetic tree of *Klebsiellavariicola* DX120E nifK protein and other species nifK proteins



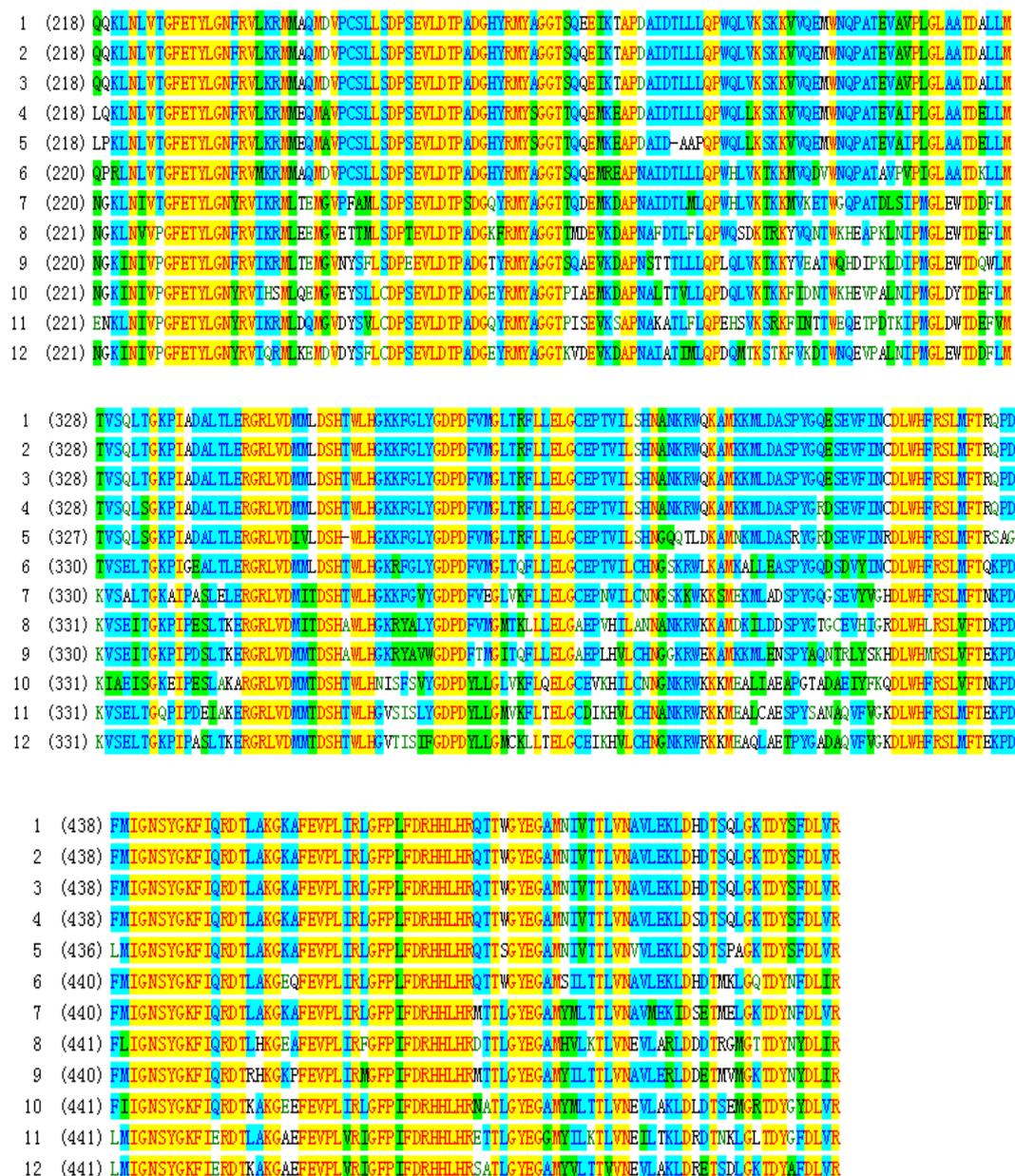


Fig. 6. Alignment of the deduced *nifK* amino acid sequence and its homologous amino acid sequences. 1. *Klebsiella variicola* DX120E; 2. *Klebsiella variicola*; 3. *Klebsiella pneumoniae*; 4. *Klebsiella oxytoca*; 5. *Klebsiella pneumoniae*; 6. *Dickeya solani*; 7. *Tolomonas auensis*; 8. *Thiobacillus prosperus*; 9. *Thioploca ingrica*; 10. *Aliagarivorans marinus*; 11. *Vibrio* sp.; 12. *Shewanella* sp.

Construction of prokaryotic expression plasmid and identification of recombinant

The *nifK* gene was recovered and purified, and pET-30a (+) plasmid was extracted at same time. EcoRI and XhoI restriction enzymes were applied for double digestion. The recovered target fragment was connected with expression vector pET-30a (+), and then transformed into DH5 α competent cells. The positive colonies were selected for PCR detection. The agarose gel electrophoresis showed the fragment was about 1563 bp. The obtained recombinant pET30a-nifK was doubly digested by EcoRI and XhoI and then sequencing. The sequencing results showed that *nifK* gene had inserted to vector plasmid. The DX120E *nifK* protein prokaryotic expression vector pET30a-nifK had successfully constructed. The correct recombinant was transformed into competent cells BL21 (DE3).

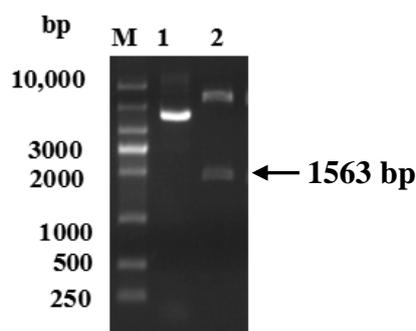


Fig. 7. Double enzyme digesting identification of restructured plasmid. M: 10 000 bp; 1: pET30a-nifK; 2: double digestion of pET30a-nifK

SDS-PAGE analysis

The empty pET-30a strain and recombinant strain were both transformed into the liquid LB medium with Kan 100 mg \cdot L⁻¹ to induce gene expression. When OD₆₀₀ reached 0.6, the medium was added with 1.0 mM IPTG for induction and 2.0 mL medium was collected at 0, 1, 2, 3, 4 and 5 h, respectively, for SDS-PAGE. The empty pET-30a strain and recombinant strain without IPTG as control (Fig. 8). As seen in Fig. 8, the target protein started expression at 1 h, reached the maximum expression level at 3 h, and then the expression

amount was not obvious.

Fig. 8 showed that, compared with the empty vector pET-30a (+), pET30a- nifK existed a fusion protein band in the molecular weight of approximately 58 kDa position, which was consistent with the expected size and the mass spectrometric *nifK* gene.

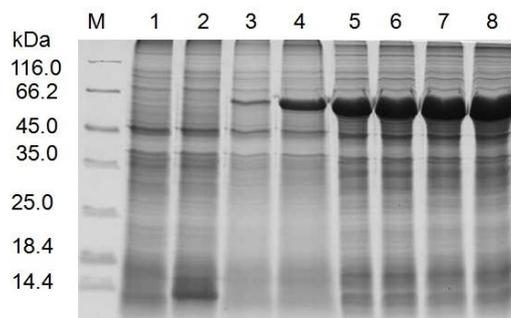


Fig. 8. SDS-PAGE analysis of restructured nifK protein. M: protein marker; 1: pET-30a without induction; 2: 1.0 mmol/L IPTG induced pET-30a (+) for 3 h; 3: pET-nifK without induction; 4-8: 1.0 mmol/L IPTG induced pET-nifK for 1h-5h

Solubility analysis of fusion protein

The recombinant bacteria pET30a-nifK after IPTG induction was broken by ultrasonic fragmentation. After analysis, the recombinant protein could express both in the supernatant and precipitate (Fig. 9).

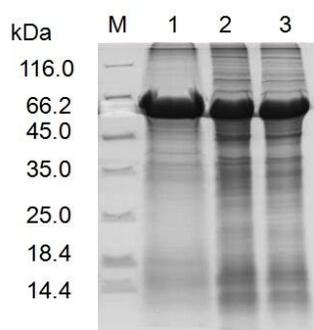


Fig. 9. Soluble analysis of pET-nifK protein. M: protein marker; 1: IPTG induced 3h precipitation of pET-nifK; 2: IPTG induced 3h supernatant of pET-nifK; 3: IPTG induced 3h bacteria of pET-nifK

Discussion

The abundant but inert N₂ in the atmosphere is converted to useful NH₃ by the metalloenzyme nitrogenase. The Fe-protein is a dimer of identical subunits encoded by the *nifH* gene, with a molecular weight in the range of 58-72 kDa, both subunits are bridged by one [4Fe-4S] metal center and contain two nucleotide [MgATP or MgADP] binding sites (Peters *et al.*, 1995; Kim *et al.*, 1994; Schindelin *et al.*, 1997). The MoFe-protein is a tetramer and is composed of two identical halves, containing α subunit and β subunit encoded by *nifD* and *nifK* genes, respectively, with a molecular weight in the range of 200-240 kDa (Howard *et al.*, 1996; Burgess *et al.*, 1996). The MoFe-protein contains metal clusters: the P-cluster and the iron-molybdenum cofactor or FeMo-co which participates in electron transfer from the Fe protein to the FeMo-cofactor (Chan *et al.*, 1993). NifH and NifDK perform the reduction of N₂ into NH₃ by coupling the reduction process to the free energy liberated from the hydrolysis of MgATP (Bulen *et al.*, 1996). In this study, the *nifK* gene was cloned from the nitrogen fixation strain *Klebsiella variicola* DX120E. The ORF is 1563 bp and has 520 amino acids. Comparisons of the amino acid sequence with other known *nifK* gene sequences suggested that *Klebsiella variicola* DX120E was most similar to *Klebsiella* spp. Phylogenetic tree also implied that nifK protein had the highest level of conservation with different species, which met the evolutionary relationships. It was predicted that nifK protein molecular weight of *Klebsiella variicola* DX120E is 58.12 kDa, and the pI is 5.54. The prokaryotic expression system using pET30a (+) with T7 bacteriophage promoter could integrated poly-histidine-tagged into the protein. This method has many advantages such as overexpression, being easy to operate, and purification (Shan *et al.*, 2008). T7 promoter on pET30a expression vector can specifically bind with T7 RNA polymerase within *Escherichia coli* BL21 (DE3), and then start downstream target gene expression of the T7 promoter. This study construct prokaryotic expression vector pET30a-nifK, after PCR, restriction enzyme digestion and sequencing, the correct recombinant strain was induced by IPTG out of a treated

58.12 kDa fusion protein. The obtained recombinant protein laid foundation for further studying the *nifK* gene function, antibody preparation on nitrogen fixation.

Conclusions

This study cloned *Klebsiella variicola* DX120E *nifK* gene from sugarcane and registered in GenBank with accession number KF732647. The ORF of *nifK* gene is 1563 bp, which encodes 520 amino acids. The protein molecular weight is 51.1 kDa, and the pI is 5.54. Bioinformatics analysis showed that the gene is highly conserved in the functional region. Phylogenetic analysis showed that the gene had the highest homology with that of *Klebsiella*. Prokaryotic expression results showed that the gene was expressed in fusion protein which relative molecular weight is 58.12 kDa approximately. The results laid molecular mechanism foundation for further studying *Klebsiella variicola* DX120E *nifK* gene interaction with sugarcane.

Acknowledgements: The present study was supported by the National High Technology Research and Development Program ("863" Program) of China (2013AA102604), National Natural Science Foundation of China (31101122, 31171504, 31471449), International Scientific Exchange Program projects (2013DFA31600), Guangxi Special Funds for Bagui Scholars and Distinguished Experts, Guangxi Natural Science Foundation (2013GXNSFAA019082), and Guangxi R & D Research Program projects (Gui Ke Chan 1123008-1, Gui Ke Gong 1222009-1B, Gui Ke He 1347004-2).

References:

- Brigle, K.E., Newton, W.E., and Dean, D.R. (1985). Complete nucleotide sequence of the *Azotobacter vinelandii* nitrogenase structural gene cluster. *Gene*, 37(1): 37–44.
- Bulen, W.A., and Lecomte, J.R. (1966). The nitrogenase system from *Azotobacter*: two enzyme requirements for N₂ reduction, ATP dependent H₂ evolution and ATP hydrolysis.

- Proceedings of the National Academy of Sciences of the United States of America, 56(3): 979–986.
- Burgess, B.K., and Lowe, D.J. (1996). Mechanism of molybdenum nitrogenase. *Chemical Reviews*, 96(7): 2983–3012
- Chan, M.K., Kim, J., and Rees, D.C. (1993). The nitrogenase FeMo-cofactor and P-cluster pair: 2.2A resolution structures. *Science*, 260(2): 792–794.
- Colón-López, M.S., Tang, H.Y., Tucker, D.L., and Sherman, L.A. (1999). Analysis of the *nifHDK* operon and structure of the NifH protein from the unicellular, diazotrophic cyanobacterium, *Cyanothece* strain sp. ATCC 51142. *Biochimica et Biophysica Acta*, 1473: 363–375.
- Dean, D.R., Jacobson, M.R., Stacey, G.S., Burris, R.H., Evans, H.J. (1992). *Biological Nitrogen Fixation*. Chapman and Hall, New York.
- Howard, J.B., and Rees, D.C. (1996). Structural basis of biological nitrogen fixation. *Chemical Reviews*, 96(7): 2965–2982.
- Inman-Bamber, N.G., and Smith, D.M. (2005). Water relations in sugarcane and response to waters deficits. *Field Crops Research*, 92 (2): 185–202.
- Kim, J., and Rees, D.C. (1994). Nitrogenase and biological nitrogen-fixation. *Biochemistry*, 33(2): 389–397.
- Li, Y.R. (2010). *Modern sugarcane Science*. China Agriculture Press, Beijing.
- Lin, L., Li, Z.Y., Hu, C.J., Zhang, X.C., Chang, S.P., Yang, L.T., Li, Y.R., and An, Q.L. (2012). Plant growth-promoting nitrogen-fixing enterobacteria are in association with sugarcane plants growing in Guangxi, China. *Microbes in the Environments*, 27(4): 391–398.
- Lisson, S.N., Inman-Bamber, N.G., Robertson, M.J., and Keating, B.A. (2005). The historical and future contribution of crop physiology and modeling research to sugarcane production systems. *Field Crops Research*, 92(2): 321–335.
- Menossi, M., Siva-Filho, M.C., Vincentz, M., Van-Sluys, M.A., and Souza, G.M. (2008).

- Sugarcane functional genomics: gene discovery for agronomic trait development. *International Journal of Plant Genomics*, 2008: 1–11.
- Peters, J.W., Fisher, K., and Dean, D.R. (1995). Nitrogenase structure and function: a biochemical-genetic perspective. *Annual Review of Microbiology*, 49(2): 335–366.
- Rubio, L.M., and Ludden, P.W. (2005). Maturation of nitrogenase: a biochemical puzzle. *Journal of Bacteriology*, 118(2): 405–414.
- Schindelin, N., Kisker, C., Schlessman, J.L., Howard, J.B., and Rees, D.C. (1997). Structure of ADP × AIF(4)(-)-stabilized nitrogenase complex and its implications for signal transduction. *Nature*, 387: 370–376.
- Schlesinger, W.H. 1991. *Biogeochemistry: An Analysis of Global Change*. Academic Press Inc.
- Shan, L.R., Yang, Z., An, N., Song, Z., and Guo, Z.K. (2008). Cloning, prokaryotic expression, purification and identification of human small ubiquitin-related modifier 4 (SUMO4). *Journal of Northwest A & F University: Natural Science Edition*, 36(9): 11-16.
- Urquiaga, S., Xavier, R.P., de Morais, R.F., Batista, R.B., Schultz, N., Leite, J.M., Maia, e Sá J.M., Barbosa, K.P., de Resende, A.S., Alves, B.J.R., and Boddey, R.M. (2012). Evidence from field nitrogen balance and ¹⁵N natural abundance data for the contribution of biological N₂ fixation to Brazilian sugarcane varieties. *Plant and Soil*, 356: 5–21.
- Wei, C.Y., Li, L., Luo, L.J., Xing, Y.X., Hu, C.J., Yang, L.T., Li, Y.R., and An, Q.L. (2014a). Endophytic nitrogen-fixing *Klebsiella variicola* strain DX120E promotes sugarcane growth. *Biology and Fertility of Soils*, 50(4):657-666.
- Wei, C.Y., Xing, Y.X., Lin, L., Yang, L.T., Li, Y.R., and Hu, C.J. (2014b). Growth-promoting effect of inoculating *Klebsiella variicola* DX120E on different sugarcane cultivars. *Chinese Journal of Applied Ecology*, 25(7): 2085-2092.
- Xing, Y.X., Mo, Y., Luo, L.J., Liu, L., Wei, C.Y., Yang, L.T., and Li, Y.R. (2014). The effect of inoculating several azotobacter with Sugarcane varieties on element NPK accumulation. *Journal of Guangxi Normal University (Natural Science Edition)*, 32(3):102-108.