

固氮调节蛋白 NifA 的生物信息学分析

Bioinformatics analysis of nitrogen fixation regulatory protein NifA

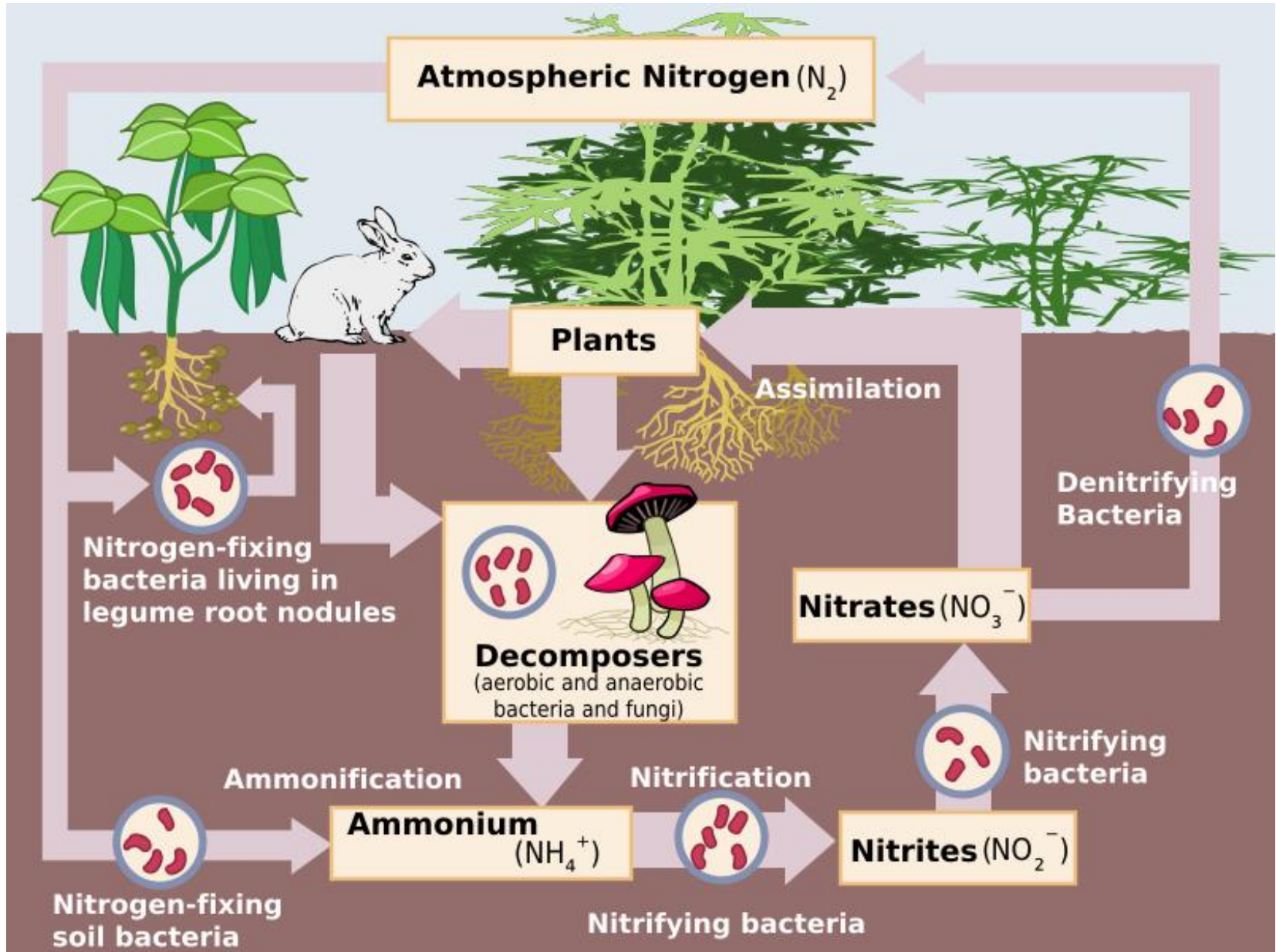
小组：G11

报告人：向鹏

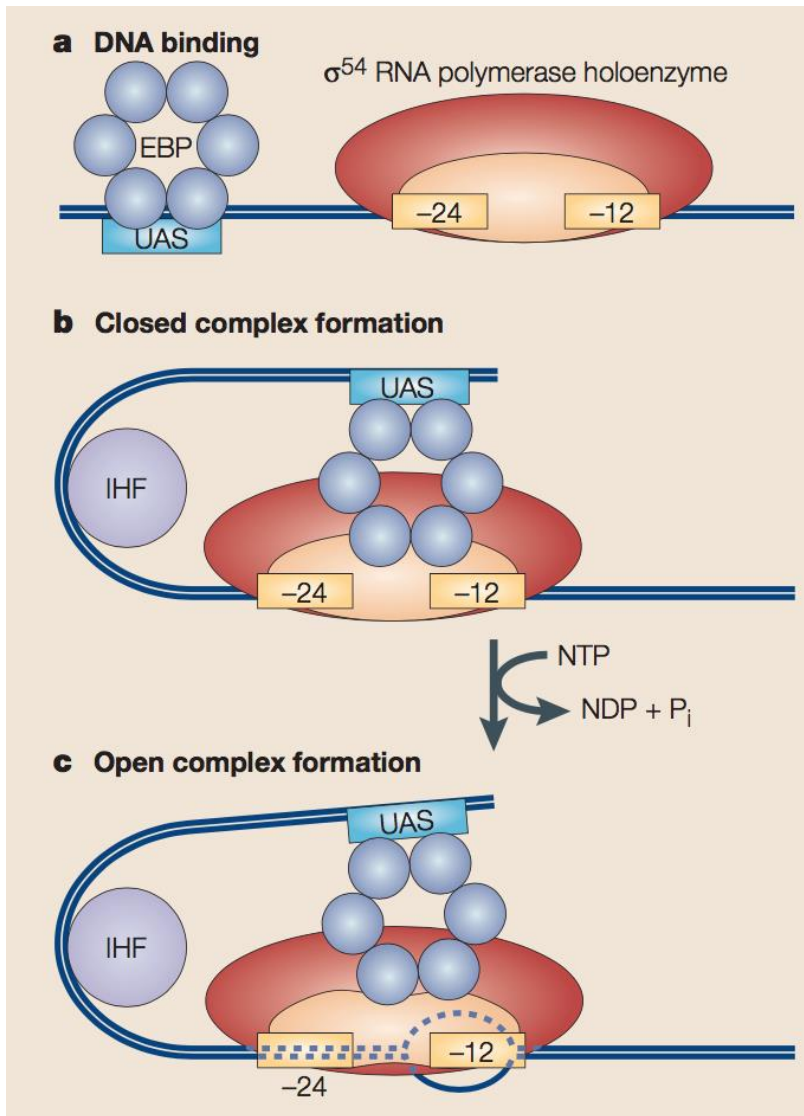
组员：向鹏，刘天旭，陈文茜，马原溪

时间：2019年1月12日

生物固氮

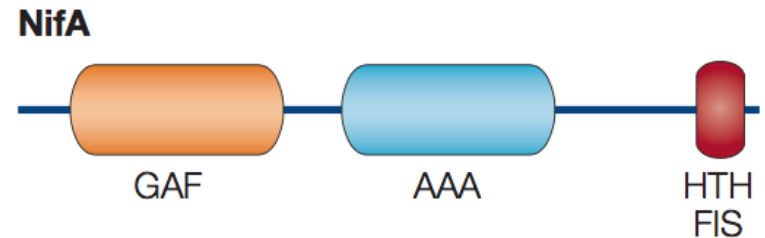


背景介绍



NifA

- EBP家族转录激活蛋白。
- 结合增强子发挥激活作用。



3个结构域
GAF、AAA及HTH FIS

常见固氮细菌中NifA及同源蛋白

>P03027 *Klebsiella pneumoniae* nifA

>A0A0B7G2N2 *Klebsiella variicola* nifA

>P30667 *Azospirillum brasilense* nifA

>P27713 *Herbaspirillum seropedicae* nifA

>P54929 *Azospirillum lipoferum* nifA

>P0CY94 *Rhodobacter capsulatus* nifA

>P09570 *Azotobacter vinelandii* nifA

>P56266 *Klebsiella oxytoca* nifA

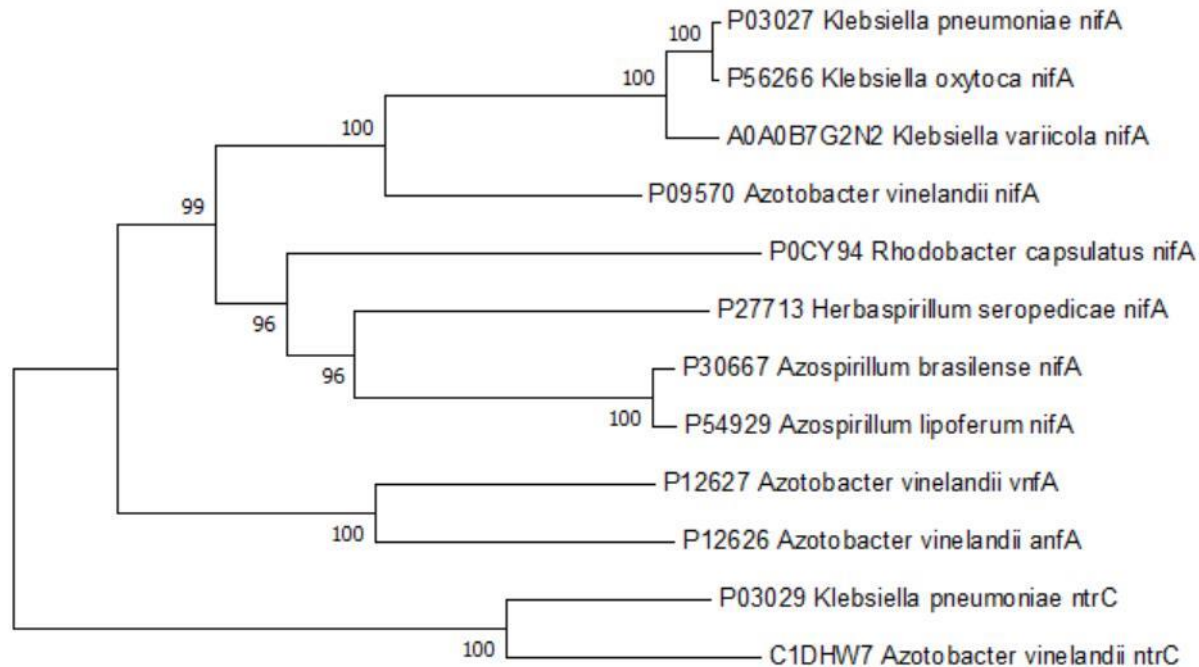
>P12627 *Azotobacter vinelandii* vnfA

>P12626 *Azotobacter vinelandii* anfA

>P03029 *Klebsiella pneumoniae* ntrC

>C1DHW7 *Azotobacter vinelandii* ntrC

常见固氮细菌中NifA同源蛋白全长的系统发育分析



将不同物种NifA及其同源蛋白序列用MEGA建立系统发育树：

- ①每类EBP都形成了独立于其他类群的集群。
- ②基因的出现早于物种。

常见固氮细菌中NifA同源蛋白的系统发育分析



不同物种的NiFA聚为一支，暗示其在不同物种中都具有保守的功能，而NiFA的同源蛋白单独聚为2支，暗示其发生了功能分化，参与不同的生物学途径。

NifA同源蛋白N-端多序列比对

N-terminal

```
NtrCKpn -----MQRGIAWIVDDSS-----IRWVLERALTGAGLSCT----TFESGNEVLDALTTKTPDV
AviNtrC -----MSRSETVWIVDDRS-----IRWVLEKALQQEGMTTQ----SFDSADGVLSRLSRQRPDV
VnfAAvi -----MSSLPQYCECGLGECRTDV---LPLLYEMSQIATESGDLSSIISILLRLMKRHMKVVRGMVTLYDRDSG----SIVLHESFGLSPEEAGK-----VYLLGEGIGRNVETGQSI
AnfAAvi MSDQATFEFDTDYFVEEFSHCFTGECRVKM---LPILYKISQIITGNADLADALSIVLGVMQQHLKMQRGIVTLYDMRAE----TIFIHDSFGLTEEEKRG-----IYAVGEGITGKVVETGKAI
NifAAvi -----MNATIPQRS AKQNPVELYDL--QLQALASIARTLSREQQIDELLEQVLAVLHNDLGLLHGLVTISDPEHG-----ALQIGAIHTDSEAVAQACEGVRYRSGEGVIGNVLKHGNSV
NifAKva -----MIP-ESDPD TT VRRFDLSQQFTAMQRISVLSRATEASKTLQEVLSVLHNDAFMQHGMICLYDSEQE-----ILSIEALQQTGGQPLPGSTQIRYRPGEGLVGTVLAQQGSL
NifAKpn -----MIHKSDSDTTVRRFDLSQQFTAMQRISVLSRATEASKTLQEVLSVLHNDAFMQHGMICLYDSQQE-----ILSIEALQQTEDQTLPGSTQIRYRPGEGLVGTVLAQQGSL
NifAKox -----MIHKSDSDTTVRRFDLSQQFTAMQRISVLSRATEASKTLQEVLSVLHNDAFMQHGMICLYDSQQE-----ILSIEALQQTENQTLPGSTQIRYRPGEGLVGTVLAQQGSL
NifARca -----MTDQQRSPASRRRSTQSIADRLAL---DALYEIAKTFAAAPDPVAEVPQIFNVLSSFLDLRHGVLALLAEPGEGAGVNPYVIAATAFQRSPEAPA----ADVLPDAVARIVFRSGVPF
NifAAbr -----MPGAMRQSTSN-----LEL---LTIYEVSKILGSSLDLQQLTREVLRALAYQLQMHGRVYLVGEDNV-----LRLVAANGLSNEAAAQI----EFRDGEGITGRILKTGMPA
NifAali -----MPGAMRQSTSN-----LEL---LTIYEVSKILGSSLDLQQLTREVLRALAYQLQMHGRVYLVGEDNV-----LRLVAANGLSNEAAAQI----EFRDGEGITGRILKTGMPA
NifAHse -----MATILDDRSVN-----LEL---VTIYEISKILGSSLDL SKTLREVLNVL SAHLETKRVL LSLMQDSGE-----LQLVSAIGLSYEEFQSG----RYRVGEGITGKIFQTETPI
: : :
NtrCKpn LLSDIR----MPGMDGLALLKQIKQRHPLMPIIIMT---AHSDLDAVSA--YQQGAFDYLPKPFIDEAVA--LVDRAISHYQEQQPRNAPIN-----SP--TADI-
AviNtrC IISDIR----MPGASGLELLAQIREHYRPLPIIIMT---AHSDLDSAVAS--YQGGAFEYLPKPFVDVEAVS--LIKRANLHAQEQQGLEVPITQ-----AN--TPEI-
VnfAAvi VVPCIRDEPAFLNRTGSRDRSDANLSFICVPILR---GRQVMGTISAERLYDN--AELLKLDVEVLSILATTTAQAVELYLVENVENVALEA-ENRRLR----SALGERFK--PANI-
AnfAAvi VARRLQEHPDFLGR T--RVS RNGAKAKAAFFCVPIMR---AQKVLGTIAAERVYMN--PRLKQDVELLTMIAITMIAPLVELYLIENIERVRLN-ENRRLK----HALKERFK--PSNI-
NifAAvi VLGRISADPRFLDRLALYDLE-----MPFIAVPIKNP--EGNTIGVLAQ--PDCRADEHMPARTRFLEIVANLLAQT VRLVNI EDGREAADERDELRR-----EVRG-KYG--FENMV
NifAKva VLPRVADDQRFLDRLSLYDYD-----LPFIAVPLMGP--NARPIGVLAQ--PMARQEERLPACTRFLETVANLVAQTIRLMILPASPALSSRQPPKVERPPACSSSRG--VG--LDNM-
NifAKpn VLPRVADDQRFLDRLSLYDYD-----LPFIAVPLMGP--HSRPIGVLAQ--AMARQEERLPACTRFLETVANLIAQTIRLMILPASAQAPQSPRIERPRACTPSRG--FG--LENM-
NifAKox VLPRVADDQRFLDRLSLYDYD-----LPFIAVPLMGP--HSRPIGVLAQ--PMARQEERLPACTRFLETVANLIAQTIRLMILPASAQAPQSPRVERPRACTSSRG--FG--LENM-
NifARca VSFDLVAE--FGAEAVPKRLRDAG--QTLIAVPLRDPERSHFVLGVLAAYRSHDHNRSGFSADAVRVLTMVASLLEQALRFRRIARDRERALE-DTRRMLQTVTEQRGPAAPVSLDGI-
NifAAbr VVPNLAAEPLFLNRTGGREDLDEQV--ASLVGVPIKA---AGVVVGVLTIDRISDEGPQGHFGSDVRFLTMVANLIGQTVRLHRTVAEERRFMMR-ETFRMQK---ELRPVAAP--INDV-
NifAali VVPNLAAEPLFLNRTGGREDLDEQV--ASLVGVPIKA---AGVVVGVLTIDRISDEGPQGHFGSDVRFLTMVANLIGQTVRL---PHVAEEPLRDAETFRMQK---ELRPIAAP--INDV-
NifAHse VVRDLAQEPLFLARTSPRQSQDGEV--ISFVGVPIKA---AREMLGVLCVFR--DGQSPSRSDVHEVRL LTMVANLIGQTVRL YRSVAERQQLQE-EKRQLSR---QLQG-KYK--LDNV-
: : : . . * : : . : :
```

结论：N端保守程度很低，暗示不同物种的接受信号可能存在差异。

NifA同源蛋白中央域多序列比对

Central domain

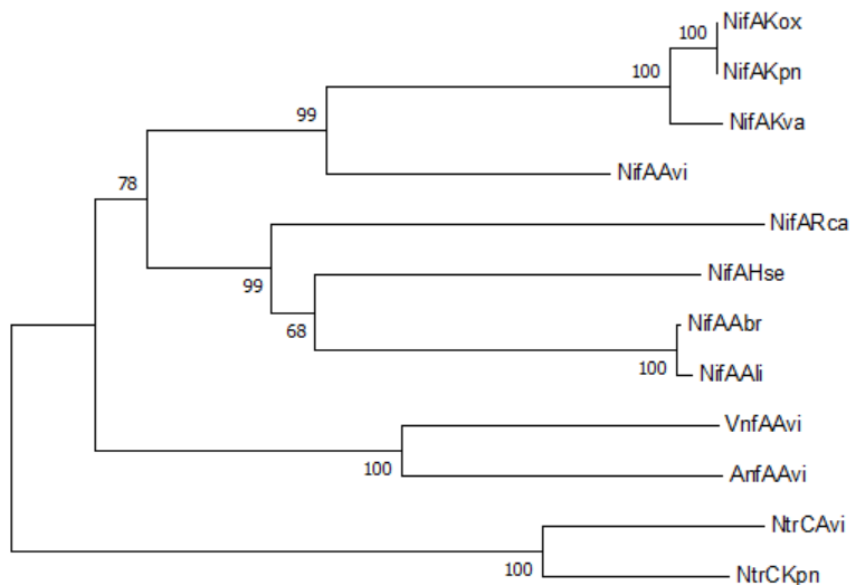
```
NtrCKpn IGEAPAMQDVFRIIGRLSRSSISVLINGESGTGKELVAHALHRHSPRAKAPFIALNMAAIPKDLIESELFGHEKGAFTGANTVVRQGRFEQADGGTLFLDEIGDMPDVDQTRLLRVLADGQ
AvintrC IGEAPAMQEVFRAIGRLSHSNITVLINGESGTGKELVAHALHRHSPRAKAPFIALNMAAIPKELMESELFGHEKGAFTGAASLRGRFEQADGGTLFLDEIGDMPADTQTRLLRVLADGE
VnfAAvi IGNSKPMLEVYQLIERVVRTRTTVLILGESGVGKELVAGAIHYNSPAAGKPFVKFNCAASLPESVIESELFGHEKGSFTGAIGLRKGRFEAAAGGTIFLDEVGEMSLTTQAKLLRVLQERS
AnfAAvi IGNSKPMQEVYELIHKVASTKATVILILGESGVGKELVANAIIHYNSPAEEAALVTSNCAPLPENLAESSELFGHEKGSFTGALTMHGKCFEQADGGTIFLDEVGELSPTVQAKLVRVLQNR
NifAAvi VGHPTMRRVFDQIRRVAKWNSTVLVLGESGTGKELIASAIHYKSPRAHRPFVRLNCAALPETLLESELFGHEKGAFTGAVKQRKGRFEQADGGTLFLDEIGEISPMFQAKLLRVLQEGE
NifAKva V GKSPAMRQIVEVIRQVSRWDTTTLVLRGESGTGKELIANAIHHHSPRAGAAAFVKFNCAALPDTLLESELFGHEKGAFTGAVRQRKGRFELADGGTLFLDEIGESSASFQAKLLRILQEGE
NifAKpn V GKSPAMRQIMDIIRQVSRWDTTTLVLRGESGTGKELIANAIHHHSPRAAAAFVKFNCAALPDNLLESELFGHEKGAFTGAVRQRKGRFELADGGTLFLDEIGESSASFQAKLLRILQEGE
NifAKox V GKSPAMRQIMDIIRQVSRWDTTTLVLRGESGTGKELIANAIHHHSPRAAAAFVKFNCAALPDNLLESELFGHEKGAFTGAVRQRKGRFELADGGTLFLDEIGESSASFQAKLLRILQEGE
NifARca V GSSPAIAEVVAQIKRVASTRMPVLLRGESGTGKELFARAVHAQSPRAKGFPIRVNCAALSETLLESELFGHEKGAFTGATALKKGRFELADGGTLFLDEIGEISPAFQSKLLRVLQEGE
NifAabr V CTSPNMLEVMAQVHRVAPFKSTVLIRGESGTGKELIARAIHNMSPRKDAPFIRVNCAALPESLLESELFGHEKGAFTGAQKDHKGRFELASGGTLFLDEIGDISPNFQAKLLRVLQEQE
NifAali V CTSPNMLEVMAQVHRVAPFKSTVLIRGESGTGKELIARAIHNMSPRKDAPFIRVNCAALPESLLESELFGHEKGAFTGAQKDHKGRFELASGGTLFLDEIGDISPNFQAKLLRVLQEQE
NifAHse I GISKAMQEVFAQVHQSPRSSTMLLRGESGTGKEVIARAIHYLSRKGDPFIKVNCAALSETLLESELFGHEKGAFTGAQGERKGRFELAHGGTLFLDEIGEISPAFQAKLLRVLQERE
: : : : : . .:: *****.* *:* ** .:: * *... : *****:*** ..* ** * ***:***:*. . *:.*:*:* :
```

```
NtrCKpn FYRVGGYAPVKVDVRIIAATHQNLELRVQEGKFREDFHRLNVIRVHLPPLRERREDIPRLARHFLQIAARELGVEAKQLHPETEMALTRLAWPGNVRQLENTCRWLTVMAAGQEV
AvintrC FYRVGGHTPVKVDVRIIAATHQNLENLQAGKFREDFHRLNVIIRIHIPRLSDRRADIPALTRHFLNRAAVELAVEPKTLKPETEYLSNLWPWPGNVRQLENTCRWITVMSAGREV
VnfAAvi FERVGGNTTIHVDLRVIAAATNRNLAEMVADGTFADLYYRLNVFPITIPPLRERGSDIITLADHFVSRFSREMGIEVNRISTPRLNMLQSYQWPGNVRELENVIERAMLLSEGDVI
AnfAAvi FERVGGSKPVKVDVRIIAATNRNLEEMVEQGTFRDLYYRLNVFPITVPLRERGSVIALADHFVSAFSRENGKNVKRISTPALNMLMSYHWPGNVRELENVMERAVILSDDVDI
NifAAvi FERVGGNQTVRVNRIVAATNRDLESEVEKGFREDLYYRLNVMAIRIPPLRERTADIPELAEFLLGKIGRQQGRPLT-VTDSAIRLLMSHRWPGNVRELENCLESAIMSEDGTI
NifAKva MERVGGDETLRVNRIIAATNRHLEEEVRLGHFREDLYYRLNVMPIALPPLRERQEDIAELAHFLVRKIGQHGRGTLR-ISEGAI RLLMEYSWPGNVRELENCLESAVMSSEGLI
NifAKpn MERVGGDETLRVNRIIAATNRHLEEEVRLGHFREDLYYRLNVMPIALPPLRERQEDIAELAHFLVRKIAHSQGRGTLR-ISDGAIRLLMEYSWPGNVRELENCLESAVLSSEGLI
NifAKox MERVGGDETLRVNRIIAATNRHLEEEVRLGHFREDLYYRLNVMPIALPPLRERQEDIAELAHFLVRKIAHSQGRGTLR-ISDGAIRLLMEYSWPGNVRELENCLESAVLSSEGLI
NifARca FERVGGAKTIKVDTRIVAATNRDLEDAVARGQFRADLYFRICVVPVIVLPLLRNRKSDIKPLAQLFLDRFNKQATNVK-FAADAFDQICRCQFPNGVRELENCVNRAAALS DGAIV
NifAabr FERVGSKTIKTDVRLICATNLNLEEAIGHGKFRADLYFRINVVTIHLPLRERRQDIGPLARHFVAKFAKDNMGLV-MEDEALEVLRCTWPGNVRELENCIERAATQSRDGI
NifAali FERVGSKTIKTDVRLICATNLNLEEAIGHGKFRADLYFRINVVTIHLPLRERRQDIGPLARHFVAKFAKDNMGLV-MEDEALEVLRCTWPGNVRELENCIERAATQSRDGI
NifAHse FERVGGSRSIKVDVRLVTATNRDLEKAVAKGEFRADLYFRINVVSI FIPPLRERREDIPYLVEHFLEKFRVENQRAMVAMSPQAMKVMNMCYWPNGVRELENCVERTATMMRGDLI
: **** .:..: *: **: * : * * **:*.*: .: : * * : * : * .: . : :*****:*** . . :
```

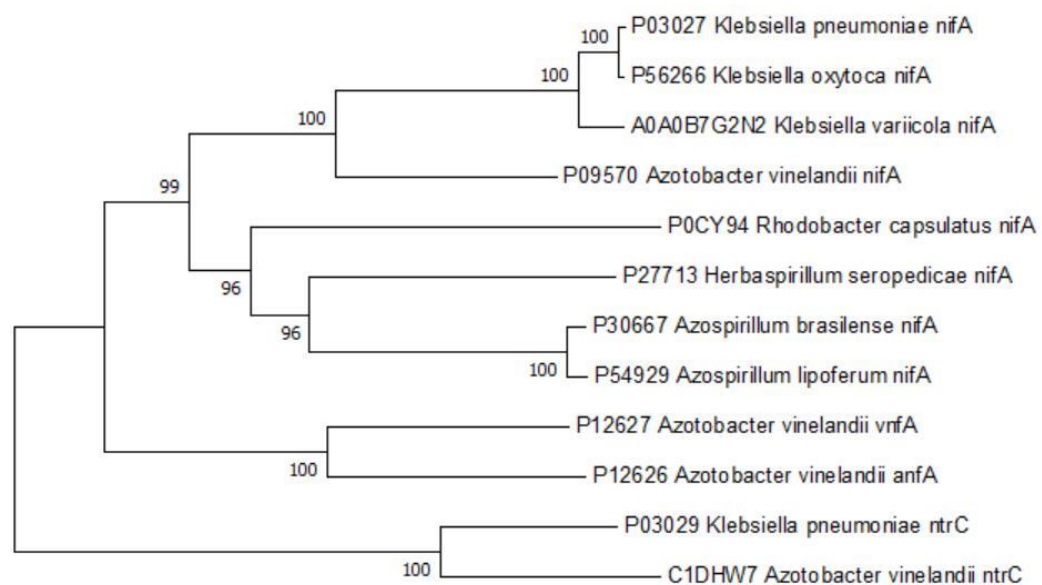
结论：中央结构域作为与RNA聚合酶结合的结构域，具有非常重要的功能，在不同物种非常保守。

NifA同源蛋白中央域系统发育进化树

中央结构域系统发育树



蛋白全长系统发育树



结论：利用中央结构域构建的系统发育树，与NifA全长蛋白构建的系统发育树类似，可反应NifA蛋白的演化情况。

NifA同源蛋白C-端多序列比对

C-terminal

NtrCKpn	LTQDLP-----SEL-----FETAIPDNPTQMLPDSWA
AviNtrC	HIDDLP-----PEL-----LSHAQENAPA----SNWE
VnfAAvi	HGYHLP-----PSL-----Q
AnfAAvi	HSYNLP-----PSL-----
NifAAvi	TRDVVSL-----TGV-----D
NifAKva	DRDVILF-----THQ-----D
NifAKpn	DRDVILF-----NHR-----D
NifAKox	DRDVILF-----NHR-----D
NifARca	LAEELACRQGACLSAELFRLQDGTSPIGGLA-----VGR-----VI-----TPTVRVSAPPPEP----A
NifAabr	RTESLSCSLNLCNSSVLFQYRTLGLASVGGGLAPSMGPGAINRVPPGPRGGPAAANAPKTPAMPAPVPEPAGAAAARGRPARRVV--PRPLAGLRRRPAGGSGPPDPACPCPSRAPLPPQA
NifAali	RTESLSCSLNLCNSSVLFQYRTLGLASVGGGLAPSMGPGSVNVRVPPGPRGVPAPANAPKAPAMPAPVPEPAGAGSAWPACASGCSAGPSPVCG-----AAQPAVPVPLIPLPLPEPSAPAA
NifAHse	TEVHFSCQQNKCLTKVLHE-----PGQ-----QQPVVV-----VPLERISAPYGAIFAEDW
.	
NtrCKpn	TLLGQW--ADRALRSQHQLLSEAQPEMERTLLTTALRHTQGHKQEAARLLGWGRNTLTRKLKELGME-----
AviNtrC	QALRQW--TDQALARGQSNLLDIVVPAFERVMIESVLKHTAGRRRDAALLLGWGRNTLTRKIKELGMNVGGSDEEDGDEG
VnfAAvi	APVVGD----SEAPPDG---LEARLGAIEYELIVEALKLHHGNMTEAATHLGLTARVLGLRMGKYNLNYKDYR-----
AnfAAvi	----QTS----KESGTAFLGLTLEEKIKAVECEMIVEALKNSSGHIGEAAKELGLARRMLGVRMERYGISYKSFSTRYA
NifAAvi	NESPPL----AAPLP--EVNLADETLD--DRERVIAALEQAGWVQAKAARLLGMITPRQIAYRIQTLNIHMRKI-----
NifAKva	RPAKAL----PASGPAEDSWLDNSLD--ERQLIAALEKAGWVQAKAARLLGMITPRQVAYRIQIMDITLPRL-----
NifAKpn	NPPKAL----ASSGPAEDGWLDNSLD--ERQLIAALEKAGWVQAKAARLLGMITPRQVAYRIQIMDITMPRL-----
NifAKox	NPPKAL----ASSGPAEDGWLDNSLD--ERQLIAALEKAGWVQAKAARLLGMITPRQVAYRIQIMDITMPRL-----
NifARca	PAPEPA----PEAPPREEVPLRKTQAQLSREELLFALLESAGWVQAKAARLLGMITPRQIAYALQKFEIELRKI-----
NifAabr	PPPSPA----AAPPPAAEVPLDEPESGSLRDRLLVAMERTGWVQAKAARLLGMITTRQVSYALRKYNIEIKRF-----
NifAali	AAPAPTSVTNAAPPPAAEVPLDEPESGSLRDRLLVAMERTGWVQAKAARLLGMITTRQVSYALRKYNIEIKRF-----
NifAHse	GQQGAT----GAAPPTSE-----RERLIWAMEQCQGWVQAKAARALNLSPRQMGYALQKFNIEVKKF-----
	: .: .** *. .: : :

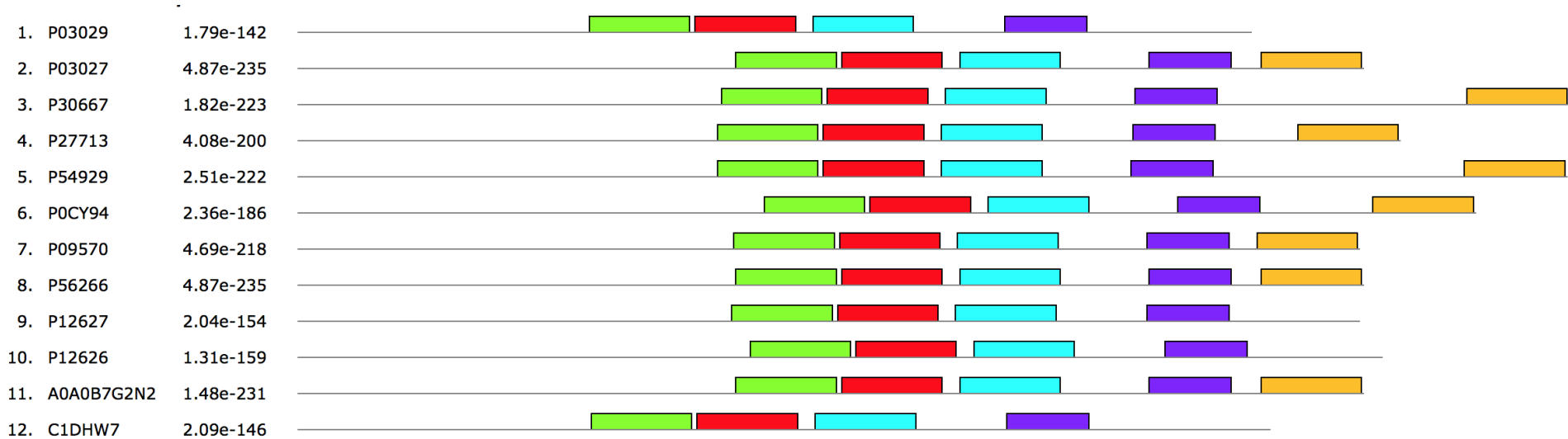
结论：C末端具有保守氨基酸残基，暗示其可能具有保守功能。

NifA及其同源蛋白motif分析



AL(M)-X9-AAX2-LG

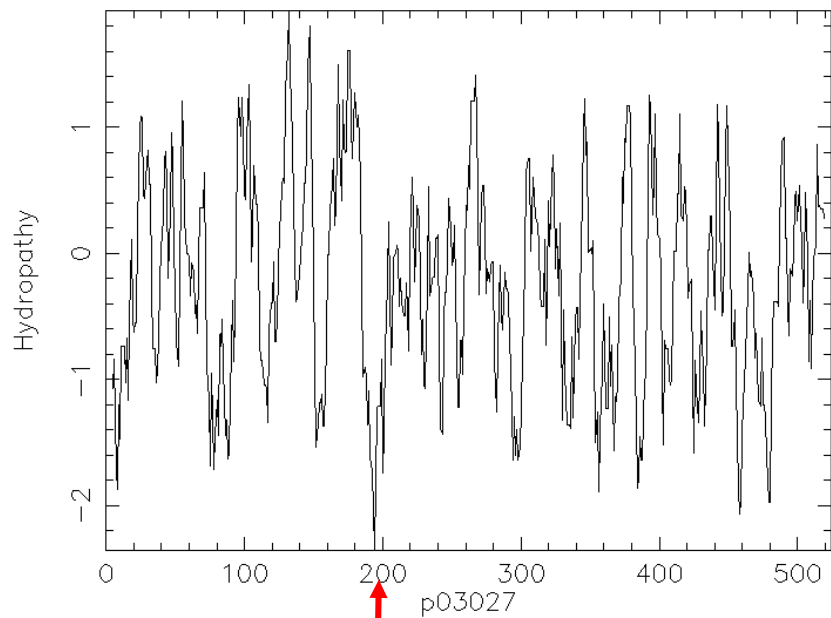
helix-turn-helix DNA-binding motif



Klebsiella pneumoniae NifA性质分析

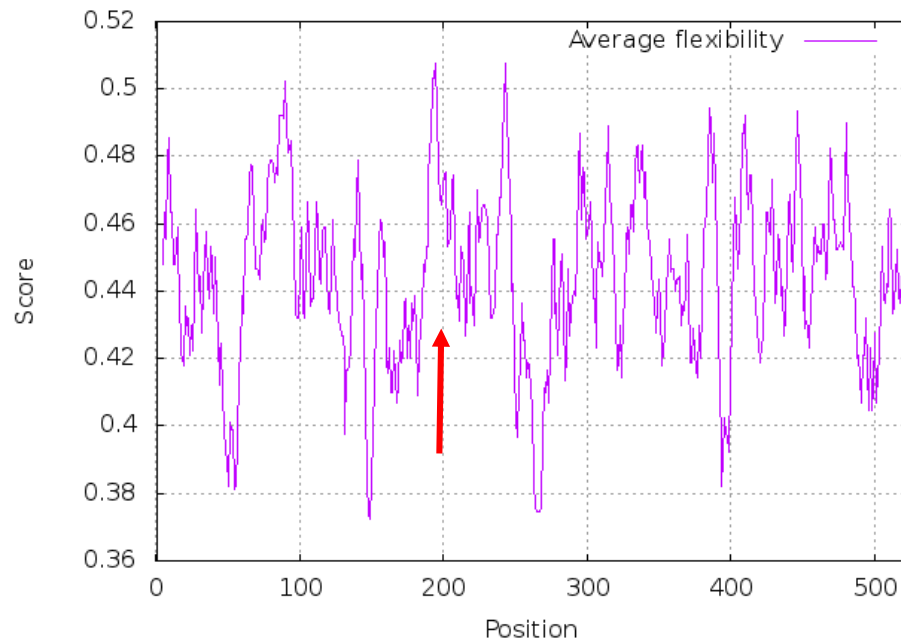
亲水性

Hydropathy Plot



柔韧性

ProtScale output for user_sequence



疏水柔韧的Q-linker作为功能域的接头。

不同分组的NifA同源蛋白结构域对比

肺炎克氏杆菌 (Kpn)



[Download](#) the data used to generate the domain

Source	Domain	Start	End
Pfam	GAF	17	176
Pfam	Sigma54_activat	212	379
Pfam	HTH_8	480	515

巴西固氮螺菌 (Abr)



[Download](#) the data used to generate the domain graph

Source	Domain	Start	End
Pfam	GAF	29	171
Pfam	Sigma54_activat	205	372
Pfam	HTH_8	583	618

保守三模体结构

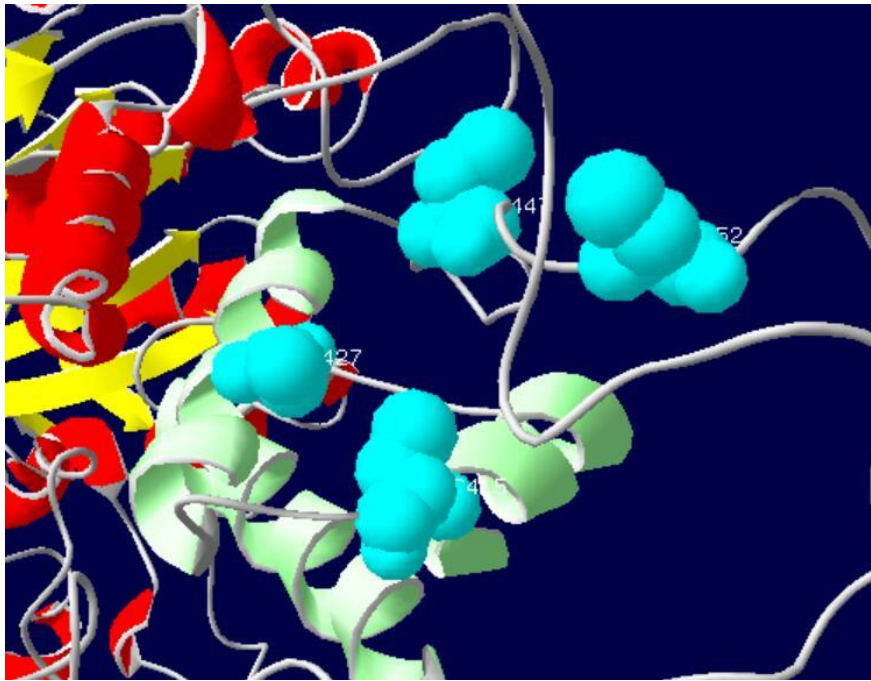
N端结构域用于接收外界信号。

中央结构域用于ATP的水解及和RNA聚合酶全酶相互作用。

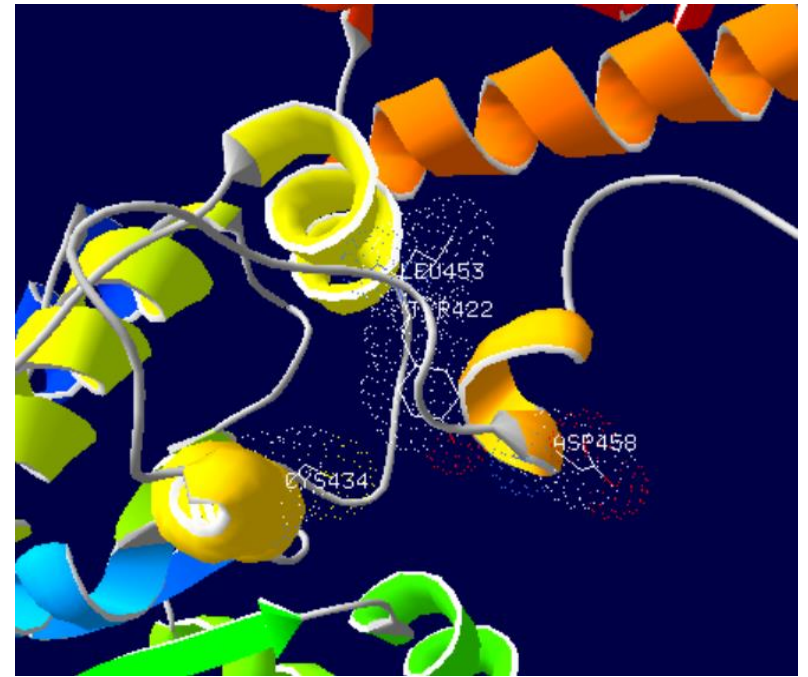
C端helix-turn-helix结构域可以和含有TGT-N₆₋₁₀-ACA序列增强子结合。

预测NifA结构与功能分析

巴西固氮螺菌
四个半胱氨酸微环境



肺炎克氏杆菌
该位点不保守

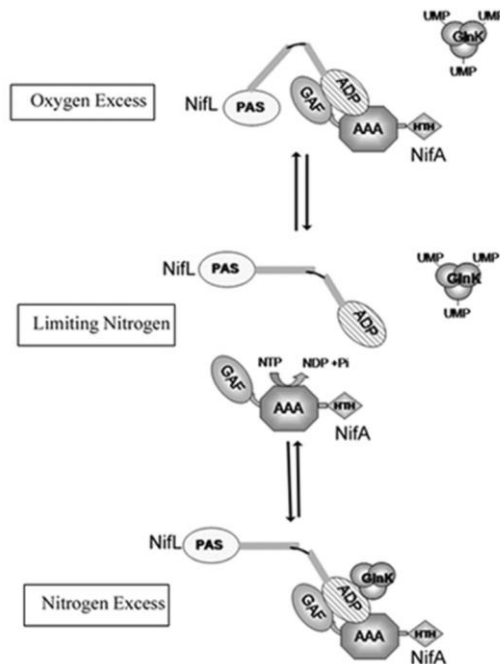


巴西固氮螺菌的NifA具有4个半胱氨酸，而在肺炎克氏杆菌该位点不保守，可能是造成NifA蛋白功能发生分化的原因。

常见固氮细菌中NifA调控模式分组

Klebsiella pneumoniae
Azotobacter vinelandii
Klebsiella oxytoca

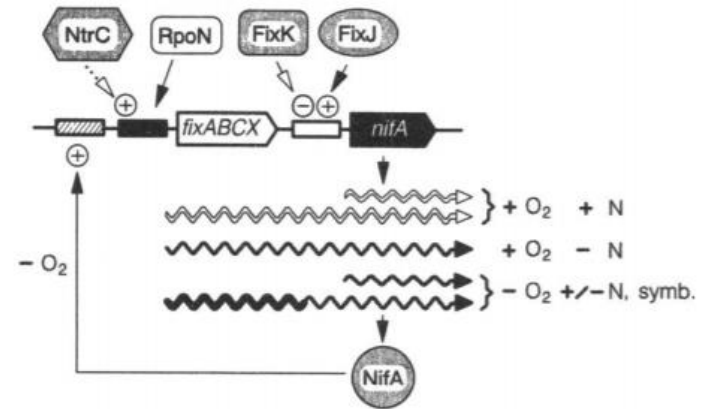
与nifL, GlnK互作
 GAF域接受化学信号 (2-OX)
 对氧不敏感



Azospirillum brasilense
Herbaspirillum seropedicae
Azospirillum lipoferum
Rhodobacter capsulatu

无nifL
 氮sensor(GlnB)对N端翻译后修饰
 感受氧敏感性

根瘤菌的调控模式



Thanks for listening! ~

