



大豆NLP2蛋白的 序列、结构分析

汇报人：栗亚静

3G01：张竹亭、邵会敏、李颖睿

研究背景



刺槐



金花菜

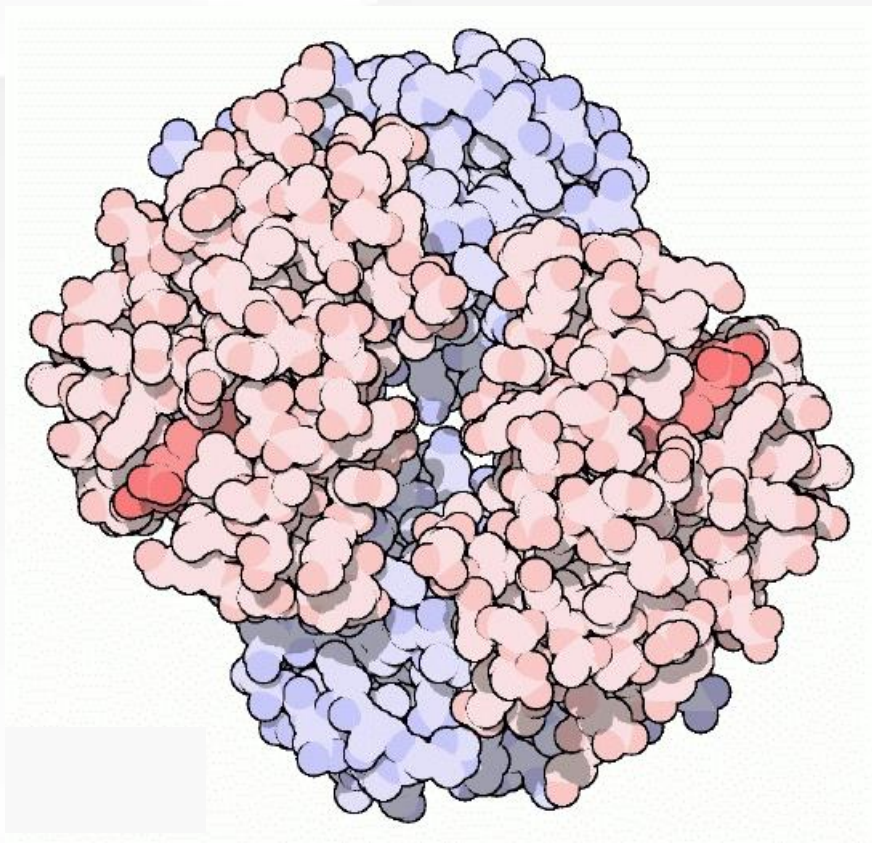


豌豆



豆科植物根瘤

研究背景



豆血红蛋白



根瘤

研究背景

PLANT SCIENCE

NIN-like protein transcription factors regulate leghemoglobin genes in legume nodules

Suyu Jiang¹, Marie-Françoise Jardinaud², Jinpeng Gao¹, Yann Pecrix^{2,3}, Jiangqi Wen⁴, Kirankumar Mysore⁴, Ping Xu⁵, Carmen Sanchez-Canizares⁶, Yiting Ruan¹, Qiujiu Li¹, Meijun Zhu¹, Fuyu Li¹, Ertao Wang¹, Phillip S. Poole⁶, Pascal Gamas², Jeremy D. Murray^{1,7*}

Leghemoglobins enable the endosymbiotic fixation of molecular nitrogen (N₂) in legume nodules by channeling O₂ for bacterial respiration while maintaining a micro-oxic environment to protect O₂-sensitive nitrogenase. We found that the NIN-like protein (NLP) transcription factors NLP2 and NIN directly activate the expression of leghemoglobins through a promoter motif, resembling a “double” version of the nitrate-responsive elements (NREs) targeted by other NLPs, that has conserved orientation and position across legumes. CRISPR knockout of the NRE-like element resulted in strongly decreased expression of the associated leghemoglobin. Our findings indicate that the origins of the NLP-leghemoglobin module for O₂ buffering in nodules can be traced to an ancient pairing of NLPs with nonsymbiotic hemoglobins that function in hypoxia.

研究背景

NLP家族 (Nin-like protein family)

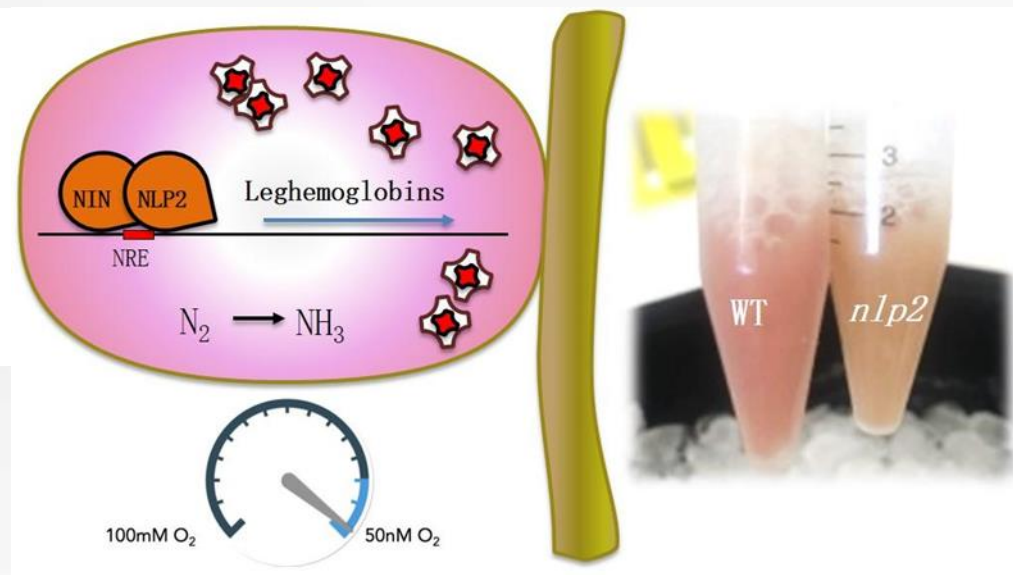
即NIN类似蛋白，为植物特有的一类转录因子，被命名为根节起源蛋白。

NIN全称: NODULE INCEPTION

nodule 指植物上的节结,小瘤

inception 开端、初期

功能: 参与豆科植物根瘤的形成, 调节植物氮代谢过程, 调控植物根部硝酸盐的吸收。



蛋白序列的获取与确定

物种名称	NCBI登录号	蛋白大小 AA	E值	序列覆盖率	相似性
大豆	KAH1251980.1	786	4e-131	75%	47.05%
	XP_003523488.1	785	4e-131	75%	47.05%
	XP_003518588.1	710	5e-112	71%	43.53%
	KAH1243973.1	716	2e-87	71%	39.01%
	XP_040865075.1	626	2e-82	51%	47.30%

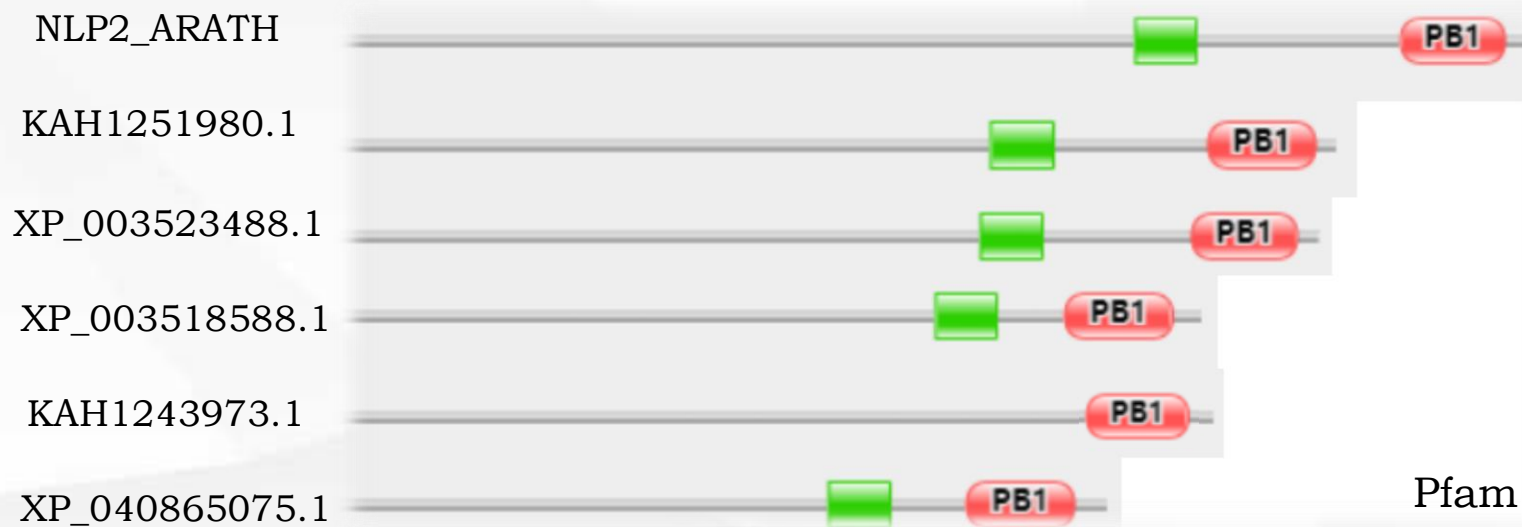
数据来源：NCBI

蛋白序列的获取与确定

Family & Domainsⁱ

Domains and Repeats

Feature key	Position(s)	Description	Actions	Graphical view	Length
Domain ⁱ	635 - 716	RWP-RK PROSITE-ProRule annotation	Add BLAST		82
Domain ⁱ	862 - 945	PB1 PROSITE-ProRule annotation	Add BLAST		84



RWP-RK 和 PB1 结构域

蛋白序列的获取与确定

NLP2ARATH	ESASEFGVQTLGSRPGEKRRRTKTEKTIGLEVLRQYFAG	660
KAH1251980.1	ESYTAAFGGRRGARKSAGDKRRTKAEKTISLPVLRQYFAG	531
XP_003523488.1	ESYTAAFGGRRGARKSAGDKRRTKAEKTISLPVLRQYFAG	530
XP_003518588.1	QSQTHTFGVKRG..RKPGEKRRRTKAEKTISLPVLRQYFAG	511
KAH1243973.1	...TAAFCGRRVGRKSAGDKRRTKAEKTISLPVLRQYFAG	538
XP_040865075.1	ESHTHTFGGKRG..RKPGEKRRRTKAEKTISLPVLRQYFAG	418

g krrtk ekti l vlrqyfag

SLKDAAKSIGVCPTTLKRICRQHGITRWPSRKIKKVGHSL	700
SLKDAAKSIGVCPTTLKRICRQHGITRWPSRKIKKVGHSL	571
SLKDAAKSIGVCPTTLKRICRQHGITRWPSRKIKKVGHSL	570
SLKDAAKSIGVCPTTLKRICRQHGITRWPS.....	541
SLKDAAKSIG.....	548
SLKDAAKSIGVCPTTLKRICRQHGITRWPS.....	448

slkdaaksig

DNAMAN

蛋白理化性质分析

编号 NO.		等电点PI Isoelectric point	分子量Mw Molecular weight	不稳定指数 Instability index	脂肪族氨 基酸指数 Aliphatic index	消光系数 Extinction coefficients	亲水性 hydropathicity
1	NLP2_ARATH	5.65	107278.54Da	48.18	73.0	91300 90300	-0.476
2	KAH1251980.1	5.46	86992.04Da	53.45	77.2	92455 91330	-0.314
3	XP_003523488.1	5.51	86862.93Da	53.09	77.30	92455 91330	-0.310
4	XP_003518588.1	5.87	79084.29Da	51.92	77.17	86955 85830	-0.312
5	KAH1243973.1	5.83	79297.64Da	47.48	77.09	89935 88810	-0.277
6	XP_040865075.1	6.10	69678.77Da	52.04	78.55	75495 74370	-0.304

不稳定指数分析可得
该五项蛋白均为不稳
定蛋白。

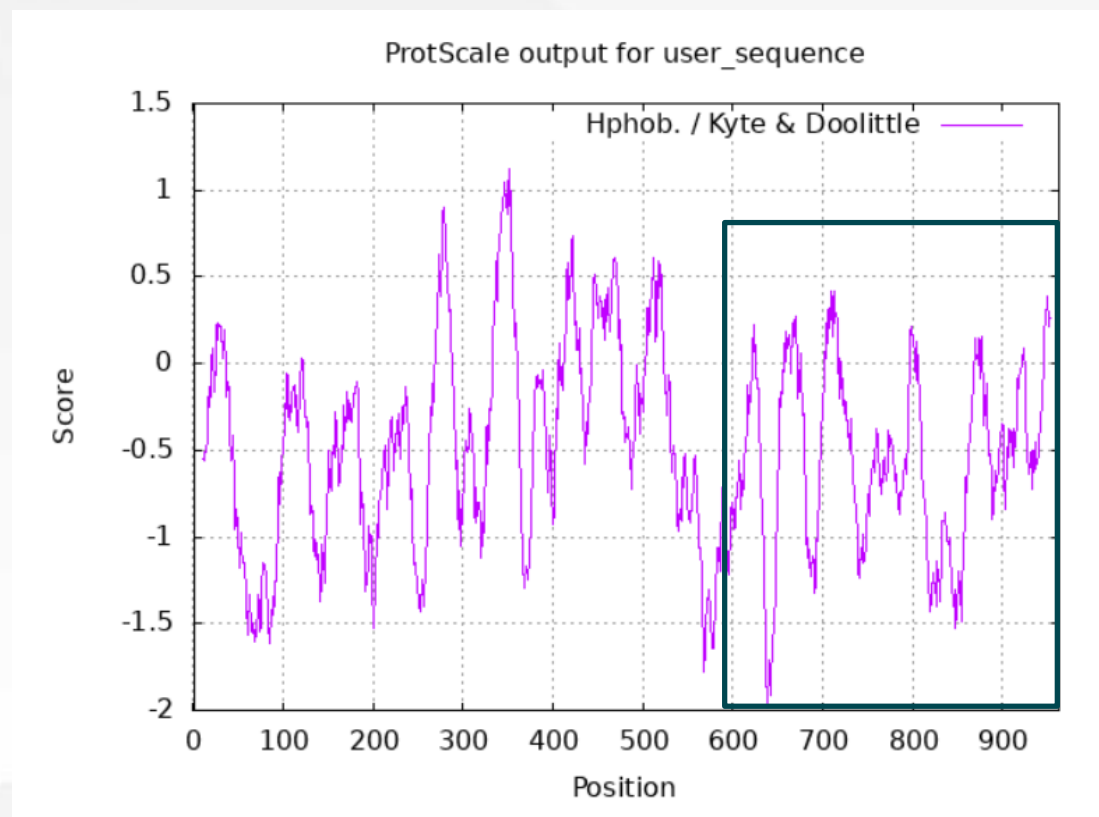
消光系数：

- ① 假设所有对Cys残
基形成胱氨酸；
- ② 假设所有半胱氨酸
残基都减少。

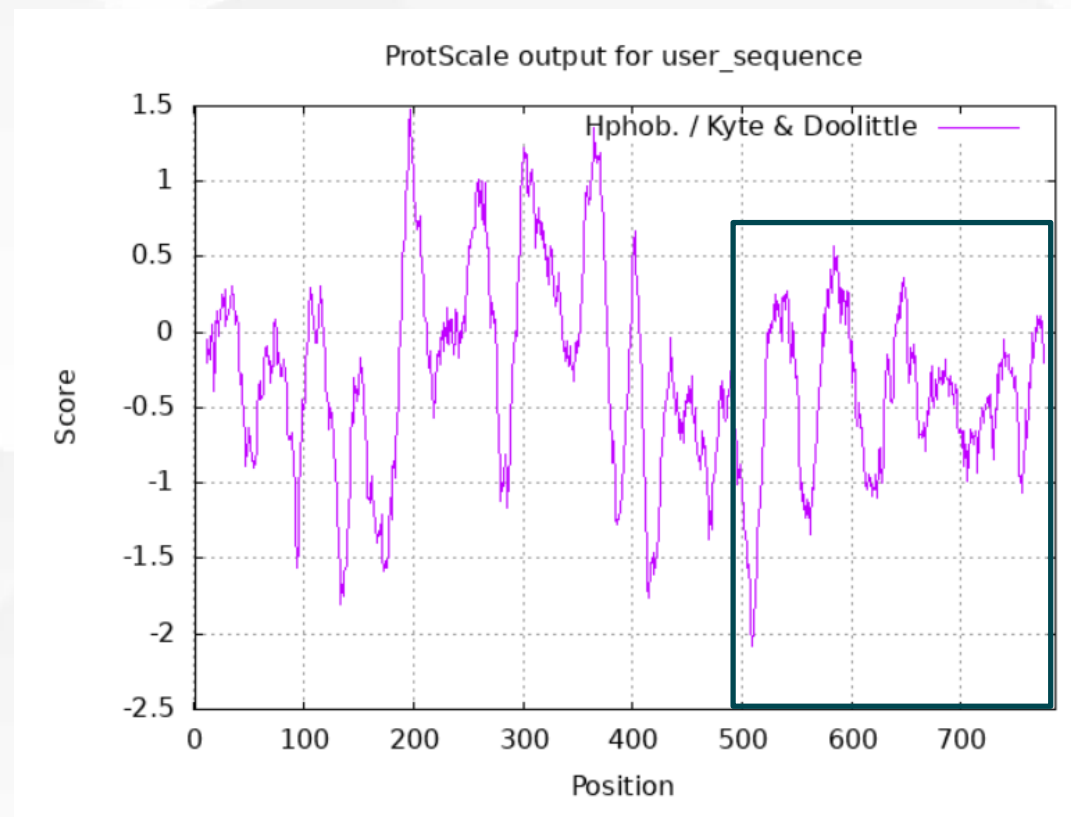
数据来源：Expasy

蛋白理化性质分析

亲水性分析



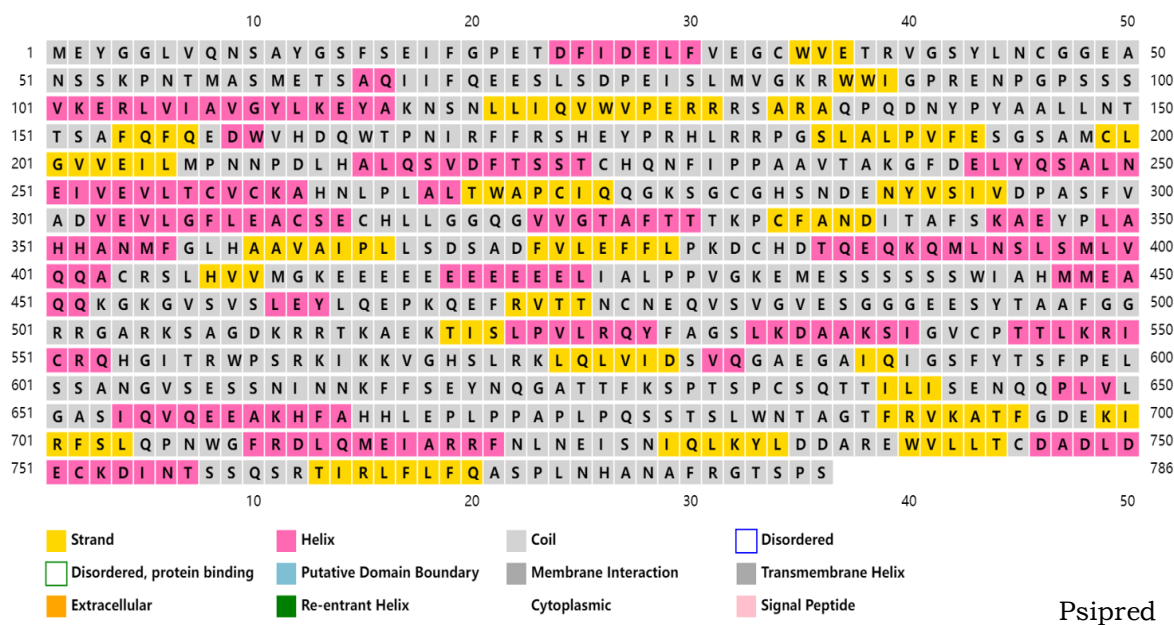
NLP2_ARATH



KAH1251980.1

蛋白结构分析

蛋白二级结构预测



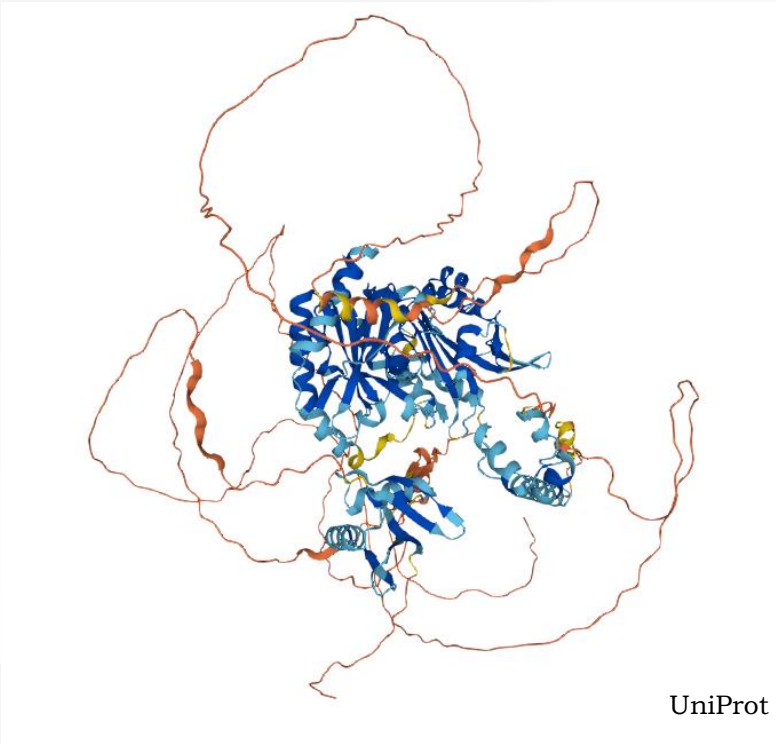
大豆KAH1251980.1

蛋白ID号 Entry name	α螺旋 (%) α-Helix	β折叠 (%) β-Strand	β转角 (%) β-Turn	无规卷曲 (%) Coil
KAH1251980.1	29.13	14.12	4.58	52.16
XP_003523488.1	32.87	14.78	5.22	47.13
XP_003518588.1	31.55	15.92	5.35	47.18
KAH1243973.1	30.31	15.64	5.45	48.60
XP_040865075.1	32.91	17.57	5.43	44.09



蛋白结构分析

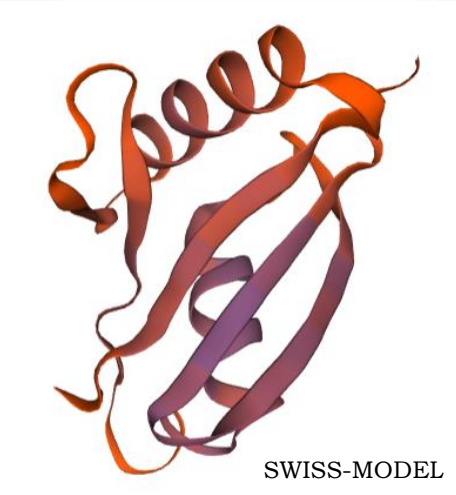
蛋白三级结构预测



NLP2_ARATH



NLP2_ARATH

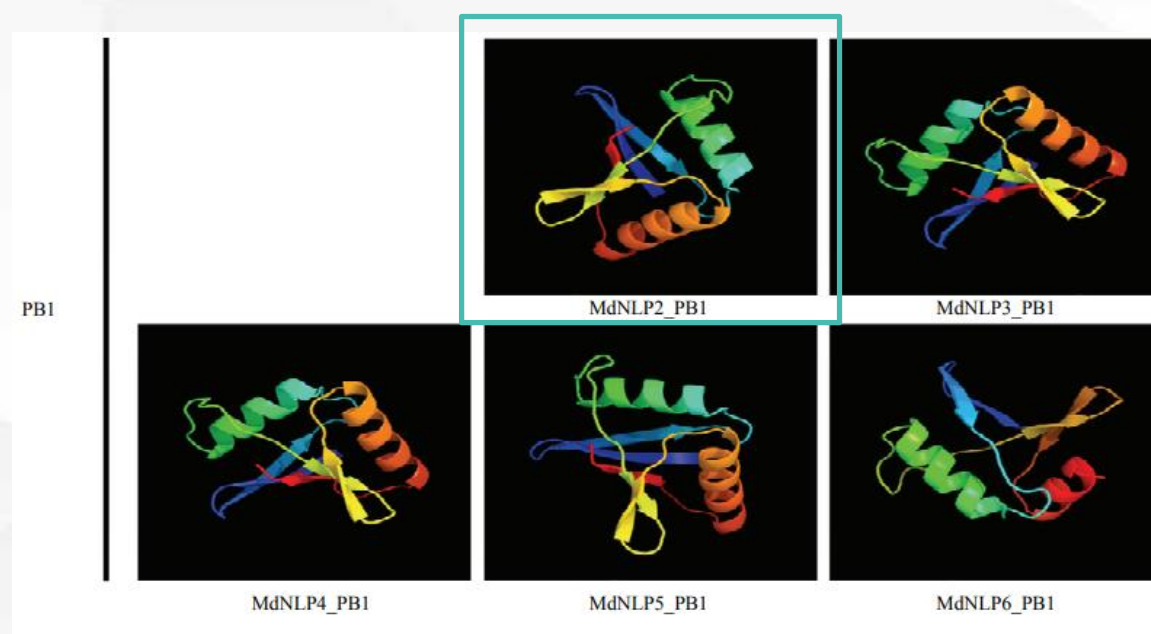
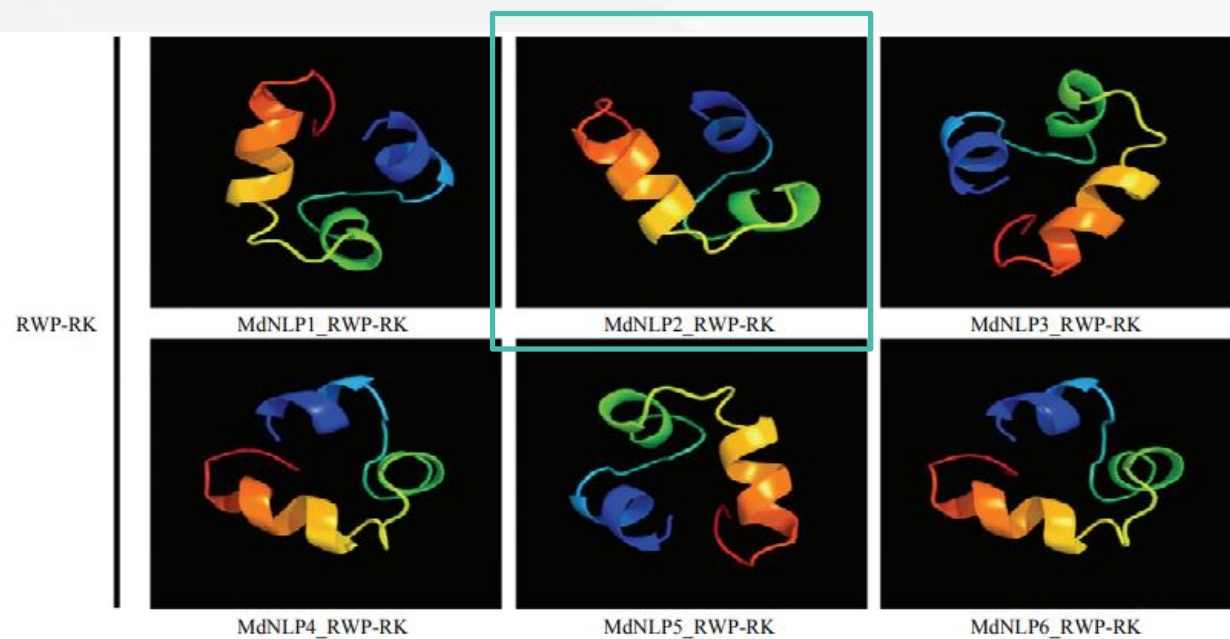


SWISS-MODEL

KAH1251980.1

蛋白结构分析

蛋白三级结构预测

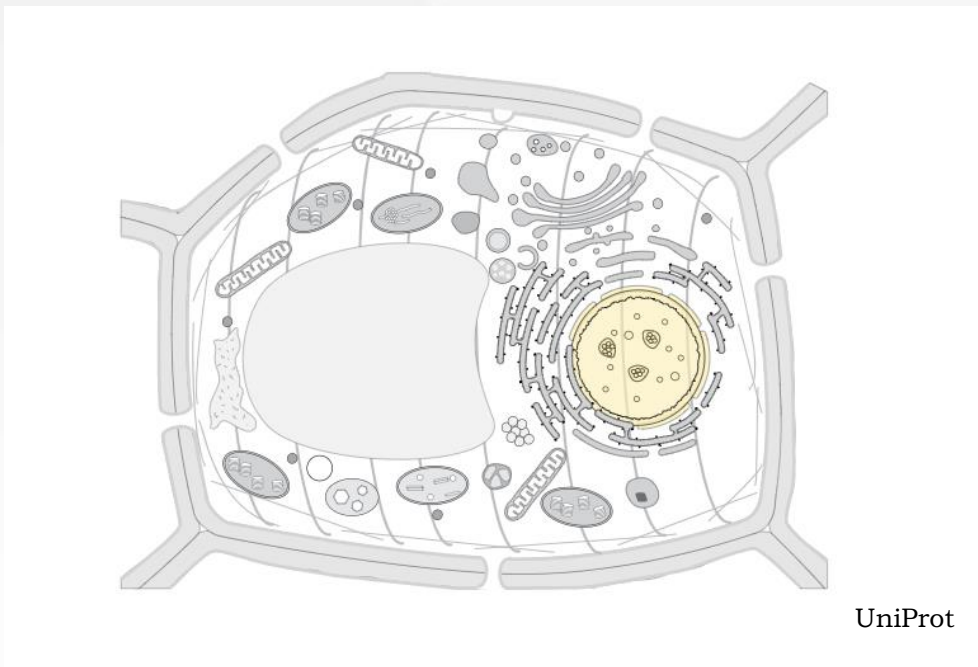


王寻等：苹果 NLP (Nin-Like Protein) 转录因子基因家族全基因组鉴定及表达模式分析



蛋白结构分析

亚细胞定位——细胞核



NLP7_ARATH



KAH1251980.1

进化关系分析

物种名称 Name	分类 Classification	物种名称 Name	分类 Classification
大豆 Glycine max	双子叶植物 Dicotyledons	粳稻 Oryza sativa Japonica Group	单子叶植物 Monocotyledons
拟南芥 Arabidopsis thaliana	双子叶植物 Dicotyledons	玉米 Zea mays	单子叶植物 Mnocotyledons
向日葵 Helianthus annuus	双子叶植物 Dicotyledons	共生菌属 Symbiodinium sp. CCMP2592	细菌 Bacterial
花生 Arachis hypogaea	双子叶植物 Dicotyledons	胶孢炭疽菌 Colletotrichum higginsianum	真菌 Fungi
芝麻 Sesamum indicum	双子叶植物 Dicotyledons		

系统发育树构建所需物种的信息

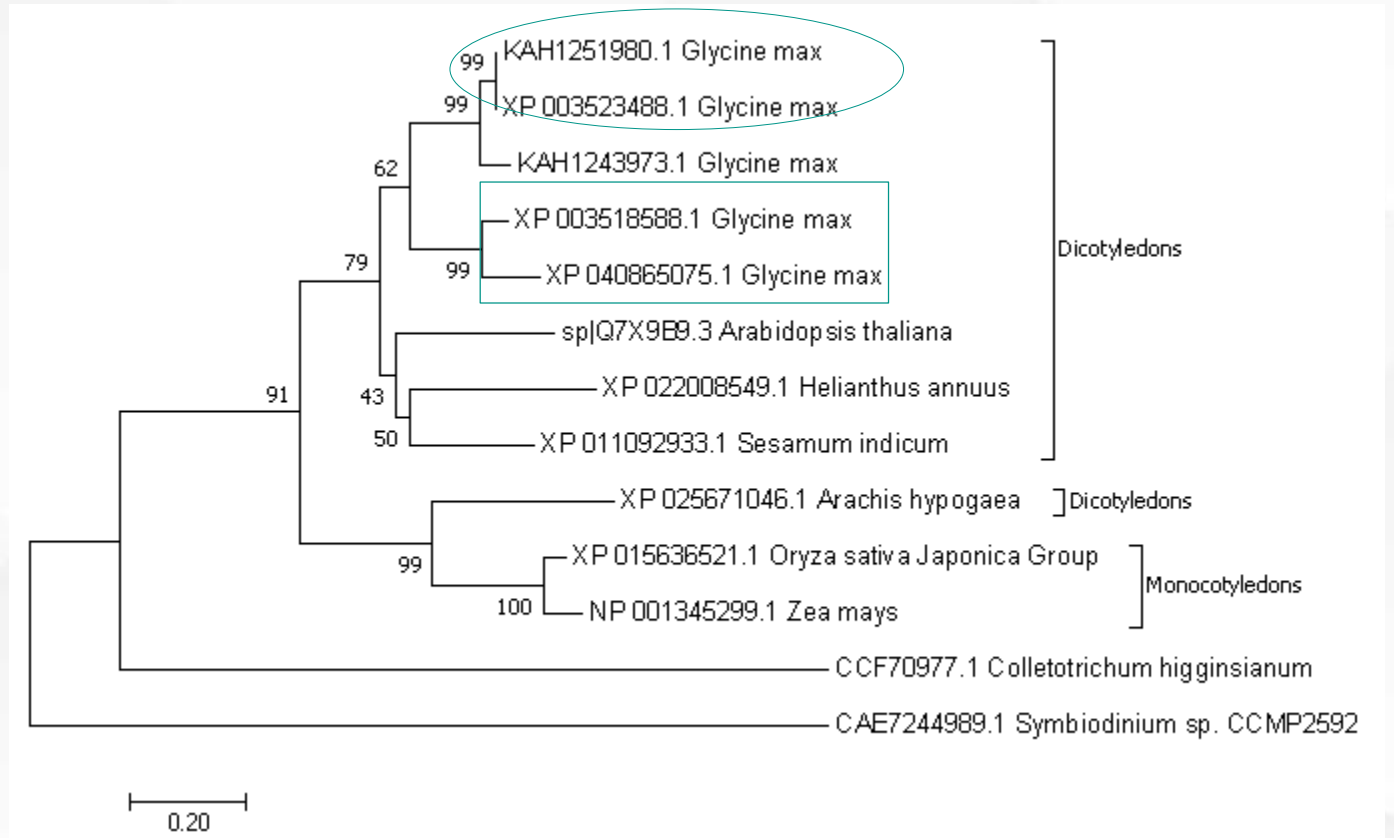
进化关系分析

构建系统发育树

大豆五条NLP2蛋白序列聚于一大分支，验证了这五个序列具有同源性。

单独分出的一个序列可能是由于该序列缺失一个重要的结构域RWP-RK。

双子叶植物序列和单子叶植物序列有明显区分，NLP2基因的进化关系与双子叶植物物种进化关系相吻合。



MEGA7

NLP2蛋白系统发育树
Phylogenetic tree of NLP2 protein

蛋白互作关系

Nitrate-responsive NIN-like protein transcription factors perform unique and redundant roles in Arabidopsis

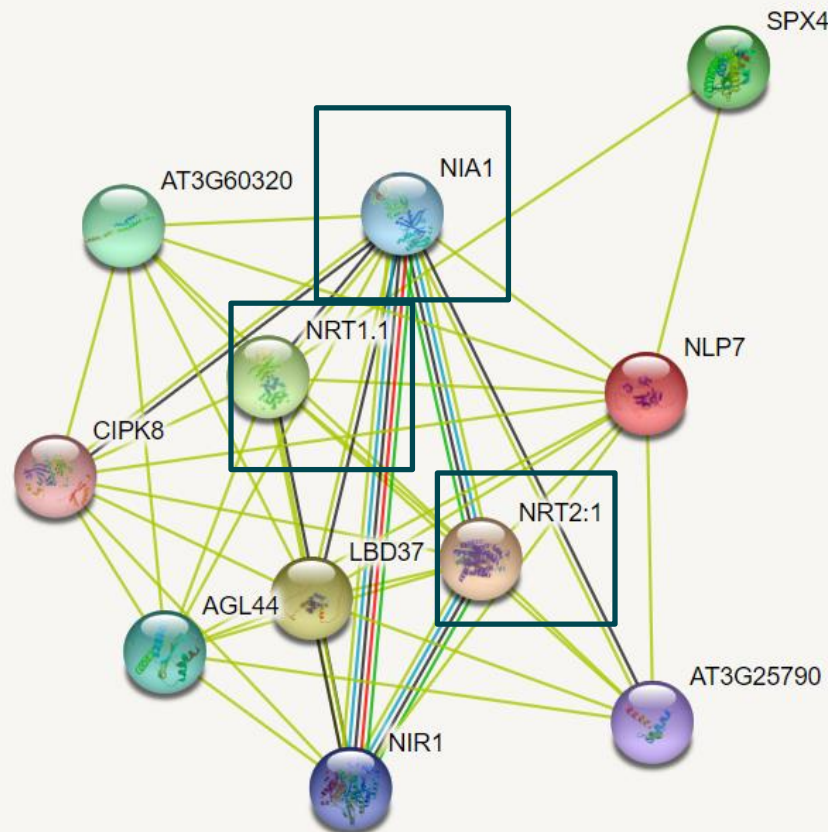
Mineko Konishi¹, Takayuki Okitsu¹, Shuichi Yanagisawa¹

Affiliations + expand

PMID: 34050740 DOI: 10.1093/jxb/erab246

Abstract

Upon sensing nitrate, NODULE INCEPTION (NIN)-like protein (NLP) transcription factors alter gene expression to promote nitrate uptake and utilization. Of the nine NLPs in Arabidopsis, the physiological roles of only three NLPs (NLP6-NLP8) have been characterized to date. To evaluate the unique and redundant roles of Arabidopsis NLPs, we assessed the phenotypes of single and higher order nlp mutants. Unlike other nlp single mutants, nlp2 and nlp7 single mutants showed a reduction in shoot fresh weight when grown in the presence of nitrate as the sole nitrogen source, indicating that NLP2, like NLP7, plays a major role in vegetative growth. Interestingly, the growth defect of nlp7 recovered upon the supply of ammonium or glutamine, whereas that of nlp2 did not. Furthermore, complementation assays using chimeric constructs revealed that the coding sequence, but not the promoter region, of NLP genes was responsible for the differences between nlp2 and nlp7 single mutant phenotypes, suggesting differences in protein function. Importantly, nitrate utilization was almost completely abolished in the nlp septuple mutant (nlp2 nlp4 nlp5 nlp6 nlp7 nlp8 nlp9), suggesting that NLPs other than NLP2 and NLP7 also assist in the regulation of nitrate-inducible gene expression and nitrate-dependent promotion of vegetative growth in Arabidopsis.



生物过程 (基因本体论)	
>	
♣ 围棋术语	▲ 描述
出发: 0071249	细胞对硝酸盐的反应
出发: 0051716	细胞对刺激的反应
出发: 0048527	侧根发育
出发: 0042128	硝酸盐同化
出发: 0015706	硝酸盐运输
分子功能 (基因本体论)	
♣ 围棋术语	▲ 描述
出发: 0015112	硝酸盐跨膜转运蛋白活性
出发: 0016661	氧化还原酶活性, 作用于其他含氮化合物作为供体

STRING

总结

- 1、通过PubMed查阅相关文献了解了NLN蛋白和NLP基因家族的演化背景和研究进展。
- 2、利用BLAST比对以拟南芥NLP2蛋白序列为参考序列获取了五条大豆NLP2转录因子的同源蛋白序列，并通过多序列比对结果和pfam结构域结果的对比发现存在PB1和RWP-RK这两个结构域以及其中一条序列缺失RWP-RK的特殊情况。
- 3、通过Expasy分析了大豆同源蛋白的基本理化性质，利用predictprotein和swiss model预测了大豆同源蛋白的二级结构和三级结构，亚细胞定位，并与拟南芥NLP2蛋白进行了对比分析，发现它们在以上这些方面具有高度的相似性。
- 4、利用MEGA对多个物种的NLP2同源序列构建系统发育树，发现在NLP2基因的进化关系与双子叶植物中物种进化关系基本吻合，猜测缺失RWP-RK结构域会干扰NLP2蛋白进化关系的判断。
- 5、通过STRING发现了拟南芥NLP7蛋白与硝酸盐转运相关蛋白存在互作，参考文献了解到NLP2和NLP7都与硝酸盐诱导拟南芥表达有关，进一步验证了NLP基因家族可以结合硝酸盐响应元件来激活下游基因的表达，参与调节植物氮代谢过程。
- 6、目前关于NLP2蛋白的研究非常有限，通过对大豆NLP2同源蛋白序列、结构的分析帮助更好地了解NLP2的调控固氮功能和与其他基因的互作关系，据报道当植物缺少NLP2后，豆血红蛋白基因的表达也受到了影响，且根瘤的固氮能力下降，因此猜想过表达NLP2是否会增强大豆根瘤固氮能力，需要相关实验进行后续验证。



请老师同学批评指正