Preliminary analysis of the vascular-related NAC-domain 1 (VND 1) transcription factor

参与维管形成的转录因子VND1的初步分析

G02 ：孟　强
　　　解振诚
　　　石　磊
　　　黄鹏辉

2017年12月13号

中国农业科学院
Chinese Academy of Agricultural Sciences

CONTENT

中国农业科学院
Chinese Academy of Agricultural Sciences

　　VND1是植物中与维管形成相关的转录因子，属于NAC转录因子家族中的一员。NAC（NAM/ATAF/CUC）转录因子家族是植物特有的一类转录因子，在水稻、番茄、拟南芥中均有存在，也是目前发现的最大的转录因子家族之一。第一个NAC基因（NAM）是1996年Soure等在矮牵牛中克隆得到的，1997年Aida在拟南芥中发现具有类似功能的CUC2基因，NAM、ATAF、CUC编码的蛋白的N端都有一段保守的氨基酸序列，被命名为NAC结构域，包含该结构域的蛋白组成了NAC转录因子家族。该家族在植物的生长发育调控中发挥着重要作用，例如：调控顶端分生组织发育、细胞分化、子叶形成、侧根发生、花发育、激素信号传导、植物器官衰老、胚胎发育以及次生细胞壁形成和纤维的形成。此外，NAC转录因子还具有相应病原菌侵染和逆境胁迫应答的作用，当植物受到生物和非生物胁迫时，会诱导大量NAC转录因子成员的产生，从而调控植物进行胁迫响应

➤ Analyzed by NCBI

Analyzed by BLAST

Analyzed by BLAST

➤ Analyzed by Clustal Omega

Analyzed by MEGA

中国农业科学院
Chinese Academy of Agricultural Sciences

➢ Analyzed by NEBcutter

Linear Sequence: Vnd1

Help   Comments

Display: – All commercial single cutter restriction enzymes
         – Main non-overlapping, min. 100 aa ORFs

GC=39%, AT=61%

Cleavage code
◤ | blunt end cut
◣ | 5′ extension
◢ | 3′ extension
▼ | cuts 1 strand

Enzyme name code
Available from NEB
Has other supplier
Not commercially available
*: cleavage affected by CpG meth.
#: cleavage affected by other meth.
(enz.name): ambiguous site

PsiI   NspI
                        365 aa
1   1671

HindIII (AAGCTT) cuts at
942 (4 nt 5′ ext)

AvrII   MwoI   BmgBI   AvaII   HindIII   TatI   PacI   SphI
BsnI   BsmFI   HpaI   AseI
       BstYI   HincII   *FauI   HaeIII
       *#BspDI   *MspA1I
       *#ClaI   *BcgI   XmnI   PflFI
       *AatII   SacI   Tth111I
       *ZraI   Eco53kI   *BsaAI
       *BsaHI   BsaXI   *SnaBI
       *SfaNI   BsoBI   BsaBI
       *BtgZI   *PaeR7I   NdeI
       BspQI   *AvaI   SspI
       SapI   *XhoI   BtsIMutI
       *ArsI   *PspXI   EcoRI
       #PfoI   *HpaII   *BfuAI
       BpmI   *MspI   BspMI
                *BsrFI   BbsI
       PsiI   NsiI   BseRI

Main options   Availability   Display   Zoom   List
New DNA   NEB   2 cutters   Zoom in   0 cutters
Custom digest   All   3 cutters   More...   1 cutters
View sequence   All sites
ORF summary   Save all sites

HindIII

[Back to main display]

```
5'... A▼A  G  C  T  T ... 3'
3'... T  T  C  G  A▲A ... 5'
```

Available from NEB, Catalog # R0104

[View product page]

REBASE enzyme page | Methylation Sensitivity

HindIII-HF (High-Fidelity version) NEB Catalog # R3104

[View product page] 🕐 Time-Saver Supplied with **CutSmart** buffer.

Buffer name: NEBuffer 2.1
    Salt: 50 mM NaCl
    Main: 10 mM Tris-HCl
      pH: 7.9
      Mg: 10 mM MgCl2
     BSA: 100

Overlapping methylation:

NONE

Reaction temperature: 37 ° C

Neoschizomers:

NONE

Isoschizomers:

NONE

Sites in sequence: 1

Enzymes producing compatible ends:

End produced at 942:

Those cutting Vnd1:

➢ Analyzed by BioEdit

Create Restriction Map

Display:
- ☑ Display Map
- ☑ Alphabetical by Name
- ☑ Numeric by position
- ☐ List of unique sites
- ☐ Fasta list of fragments
- ☐ Create feature map

- ☑ Sites that cut five or fewer times
- ☐ Summary table of frequencies
- ☑ Summary of enzymes that do not cut
- ☐ 4-base cutters
- ☐ 5-base cutters
- ☑ 6-base cutters
- ☑ Enzymes with degenerate recognition

- ☑ large recognition sites (>6)
- ☐ All Isoschizomers (not recommended)

Manufacturer:

All Enzymes ▼ ➡

View by Manufacturer   Select from list

Title: Arabidopsis thaliana vascular related NAC-domain protein 1 (VND1), mRNA Restriction Map

Display Translations:

Top Strand
- ☐ Frame 1
- ☐ Frame 2
- ☐ Frame 3

Bottom Strand
- ☐ Frame -1
- ☐ Frame -2
- ☐ Frame -3

☐ Circular DNA (ends joined)

Generate Map   Cancel

View All Currently Selected Enzymes

tcttcaccacttctctcttctttctaattaagttgttctcaccagtcaccaccattgatcttcctaggtttttcttttgtttccgtcacatact

```
1671 base pairs
Translations:  none


Restriction Enzyme Map:


1        TCTTCACCACTTCTCTCTTCTTTCTAATTAAGTTGTTCTCACCAGTCACCACCATTGATCTTCCTAGGTTTTTCTTTTTGT    80
1        AGAAGTGGTGAAGAGAGAAGAAAGATTAATTCAACAAGAGTGGTCAGTGGTGGTAACTAGAAGGATCCAAAAGAAAAACA    80
            MboII           EarI        HphI     HphI              MboII           AvrII        TspGWI
                                                BsrI                            BsaJI
                                                                                StyI


81       TTCCGTCACATACTCGATCAATTATAATTATATATTCGGCTGCGAAGATTTACAAGAAATAATGGAGCCAATGGAATCTT    160
81       AAGGCAGTGTATGAGCTAGTTAATATTAATATATAAGCCGACGCTTCTAAATGTTCTTTATTACCTCGGTTACCTTAGAA    160
                            PsiI                              MboII        NlaIV      BpmI
                            BbvI                                                      Eco57MI


161      GTAGCGTTCCTCCAGGATTTAGGTTCCATCCGACGGACGAAGAGCTTGTCGGGTACTATCTAAGGAAGAAAATCGCATCG    240
161      CATCGCAAGGAGGTCCTAAATCCAAGGTAGGCTGCCTGCTTCTCGAACAGCCCATGATAGATTCCTTCTTTTAGCGTAGC    240
                    FokI    MnlI     BstF5I    EarI          TspGWI                          MboII
                            NlaIV          BslI              MboII
                                   SapI                      MmeI


241      CAAAAGATTGATCTCGACGTCATCAGAGACATCGATCTCTACAGAATAGAACCATGGGATCTACAAGAACAATGTCGAAT    320
241      GTTTTCTAACTAGAGCTGCAGTAGTCTCTGTAGCTAGAGATGTCTTATCTTGGTACCCTAGATGTTCTTGTTACAGCTTA    320
            SfaNI       Hpy188III          ClaI     SfcI          BsaJI          AlwI
                        BsaHI                                     BtgI
                        ZraI                                      NcoI
                          AatII                                   StyI
                        BsmAI                                       BstYI
```

```
Restriction table:

Enzyme        Recognition                              frequency    Positions

AatII         G_ACGT'C                                 1            261
AlwI          GGATCnnnn'n_                              2            306, 539
ApaLI         G'TGCA_C                                  2            504, 1614
ApoI          r'AATT_y                                  2            1184, 1496
AseI          AT'TA_AT                                  1            1543
AvaI          C'yCGr_G                                  1            551
AvrII         C'CTAG_G                                  1            64
BanII         G_rGCy'C                                  2            656, 1628
BbsI          GAAGACnn'nnnn_                            1            1065
BbvI          GCAGCnnnnnnnn'nnnn_                       2            107, 742
BcgI          CGAnnnnnnTGCnnnnnnnnnn_nn'                1            693
BcgI          GCAnnnnnnTCGnnnnnnnnnn_nn'                1            659
BfrBI         ATG'CAT                                  1            538
Bme1580I      G_kGCm'C                                 2            508, 1618
BmgBI         CAC'GTC                                  1            503
BpmI          CTGGAGnnnnnnnnnnnnnnnn_nn'                1            156
BpuEI         CTTGAGnnnnnnnnnnnnnnnn_nn'                2            640, 1210
BsaAI         yAC'GTr                                  1            1417
BsaBI         GATnn'nnATC                              1            1399
BsaHI         Gr'CG_yC                                 1            258
BsaJI         C'CnnG_G                                  5            64, 293, 405, 426, 576
BsaXI         ACnnnnnCTCCnnnnnnn_nnn'                   1            643
BsaXI         GGAGnnnnnGTnnnnnnnnn_nnn'                 1            673
BseRI         GAGGAGnnnnnnnn_nn'                        1            978
BsiHKAI       G_wGCw'C                                 3            508, 656, 1618
BslI          CCnn_nnn'nnGG                             2            193, 931
BsmI          GAATG_Cn'                                2            571, 605
```

Primer Premier

File Edit View Function Translate Window Help

New	▶	DNA Sequence
Open	▶	Protein Sequence
Insert Sequence		
Save	Ctrl+S	
Save As		
Preferences	Ctrl+R	
Degenerate Bases	Ctrl+B	
Print		
Quit		

1 D:\Program Files (x86)\BioSoft\Primer Premier 5\data\DemoTNF.ALN

2 C:\Program Files\Primer Premier 5\data\DemoTNF.ALN

...ence

Translation:

Active Sequence: NewSequence ▼

Translations: Original DNA

DNA Protein

| 5' Seq No | Header | 3 | 10 | Find | Find Next | S | A | dsDNA |

1

Pos: 00001

File   Edit   View   Function   Report   Graph   Window   Help

Primer Premier

Primer:

[S][A]   Search   Results   Edit Primers

1   (1)   (1671)   25
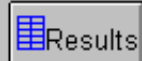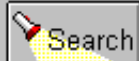
Direct Select:

```
3'AGAAGTGGTGAAGAGAGAAGAAAGA 5'
  |||||||||||||||||||||||||
5' TCTTCACCACTTCTCTCTTCTTTCTAATTAAGTTGTTCTCACCAGTCACCACCATTGATCTTCCTAGGTTTTCTTTTTGTTT 3'
        10        20        30        40        50
   S  S  P  L  L  S  S  F  -  L  S  C  S  H  Q  S  P  P  L  I
```

	Rating	Seq No	Length	Tm [°C]	GC%	Δ G [kcals/mol]	Activi [µg/OD
Sense	83	1	25	58.1	40.0	-41.6	37.0
Anti-sense	83	25	25	58.1	40.0	-41.6	27.9
Product	28	--	25	65.5	40.0	--	--

	Hairpin	Dimer	False Priming	Cross Dimer	No Hairpins Found
Sense	None	None	Found	Found	
Anti-sense	None	None	Found		

Search Criteria

Search For:

( ) PCR Primers   ( ) Sequencing Primers   ( ) Hybridization Probes

Search Type:

( ) Sense Primer   ( ) Compatible with Sense Primer
( ) Anti-sense Primer   ( ) Compatible with Anti-sense Primer
( ) Both   ( ) Pairs

Search Ranges:

Sense Primer:
1   to   1671

Anti-sense Primer:
1   to   1671

PCR Product Size:
100   bp   to   500   bp

Primer Length:
25   bp   ?   1   bp

Search Mode:
( ) Automatic   ( ) Manual

Search Parameters

✓ OK   ✗ Cancel   ? Help
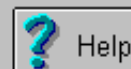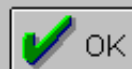
1321
1381
1441
1501
1561
1621

Primer Premier

File Edit View Function Report Graph Window Help

Primer Premier

Primer:

[S] [A] [Search] [Results] [Edit Primers]

31
(1)   (1671)

1670

Direct Select:

```
                    3' CAGGTGTTATGGTTGCTCTGAAAC 5'
                       |||||||||||||||||||||||||
5' GATGTCGAGCGTGCGGTGTGCACAACGGGGCTCCGTTGGGCATGCTTTGTCGTCCACAATACCAACGAGACTTTGG    3'
      1600      1610      1620      1630      1640      1650      1660      1670
    M   S   S   V   R   C   A   Q   R   G   S   V   G   H   A   L   S   S   T   I   P   T   R   L   W
```

	Rating	Seq No	Length	Tm [°C]	GC%	Δ G [kcals/mol]	Activity [µg/OD]	Degeneracy	Ta Opt [°C]
Sense	100	31	24	63.2	50.0	-42.4	32.5	1	--
Anti-sense	93	1670	24	60.6	45.8	-42.4	32.4	1	--
Product	82	--	1640	84.9	38.6	--	--	--	52.6

	Hairpin	Dimer	False Priming	Cross Dimer	No Hairpins Found	All
Sense	None	None	None	Found		
Anti-sense	Found	Found	None			
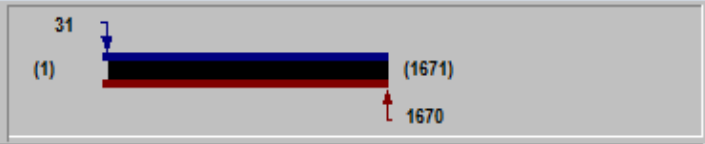
NewSequence

Original DNA

[DNA]   [Protein]

nd Next   [S] [A] [dsDNA]

```
        GAAAAC AGGTTCGTCA GTGAATTCTT GAGTACAAAC
        GTATTT GATAATTGAA TGAAACTACA TAGATGGCAT
        TACACT TATATGAAAA TATTAACAAT ATAACTAATA
1321    AAGAAAATTA AACATATGGT TTAGCCATGT CTAGTTTTAT GAAAAGAATG TGTTAGGGTT
1381    GTGAATGAGT GAGGATATAA ATCTATATTA CGTTACGTAC TATTTATATT CGAGTTTGAG
1441    TGACTAGGTC AGAGTTATTG TTACGTTATG AGGGTATGAA TTGTGAGGTT TTCTAAATTT
1501    TTTAATTATA GTTAGTTTTA TAGAGATGAT TTAAGGGGCC ATTAATTAAG TATAGTTAGA
1561    AATCAGAATA AGAAATCATT GTTGTGGGGC AAGGAGATGT CGAGCGTGCG GTGTGCACAA
1621    CGGGGCTCCG TTGGGCATGC TTTGTCGTCC ACAATACCAA CGAGACTTTG G
```

Pos: 01672

Search Results

◯ Sense   ⦿ Anti-sense   ◯ Pairs

824 anti-sense primers found.

#	Rating	Seq No	Length	Tm [°C]	GC%	Mark
1	100	58	24	63.5	50.0	☐
2	100	770	24	63.6	41.7	☐
3	100	858	24	69.3	58.3	☐
4	100	974	24	54.2	41.7	☐
5	100	1144	24	62.3	41.7	☐
6	100	1146	24	62.1	41.7	☐
7	93	1226	24	61.6	37.5	☐

Primer Premier

File　Edit　View　Function　Report　Graph　Window　Help

Primer Premier

Primer:

S　A

Search　Results　Edit Primers

(1)

835

(1671)
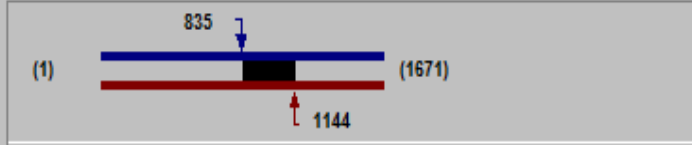
1144

Direct Select:

3' TACACCTCTACCTAAGCAGTTACA 5'

5' TGAAGATAAAAACAACAATGATGTGGAGATGGATTCGTCAATGTT

1110　1120　1130　1140

E　D　K　N　N　N　D　V　E　M　D　S　S　M　F

	Rating	Seq No	Length	Tm [ºC]	GC%
Sense	100	835	24	69.5	58.3
Anti-sense	100	1144	24	62.3	41.7
Product	88	--	310	84.1	39.0

	Hairpin	Dimer	False Priming	Cross Dimer	N
Sense	None	None	None	None	
Anti-sense	None	None	None		

Edit Primer

HinfI

V　E　M　D　S　S　M

3' TACACCTCTACCTAAGCAGTTACA 5'

|||||||||||||||||||||||

5' ATGTGGAGATGGATTCGTCAATGTTTTTGAGCGAAAGAGAAGAAGAAAACAGGTTCGTCAGTGAATTCTTGAGTACAAACT 3'

V　E　M　D　S　S　M　F　L　S　E　R　E　E　N　R　F　V　S　E　F　L　S　T　N　S

1130　1140　1150　1160　1170　1180　1190　1

HinfI　EcoRI

Rating	Seq No	Length	Tm [ºC]	GC%	ΔG [kcals/mol]	Activity [μg/OD]	Degeneracy
100	1144	24	62.3	41.7	-42.8	31.7	1

Hairpin	Dimer	False Priming	No Hairpins Found
None	None	None	

Enzyme

Analyze

Prime

OK

Cancel

Help

Search Results

○ Sense　◉ Anti-sense　○ Pairs

824 anti-sense primers found.

#	Rating	Seq No	Length	Tm [ºC]	GC%	Mark
	100	58	24	63.5	50.0	
	100	770	24	63.6	41.7	
	100	858	24	69.3	58.3	
	100	974	24	54.2	41.7	
	100	1144	24	62.3	41.7	
	100	1146	24	62.1	41.7	
	93	1226	24	61.6	37.5	

Structure

Primary Structure

>NP_179397.1 vascular related NAC-domain protein 1 [Arabidopsis thaliana]
MEPMESCSVPPGFRFHPTDEELVGYYLRKKIASQKID
LDVIRDIDLYRIEPWDLQEQCRIGYEEQNEWYFFSHK
DKKYPTGTRTNRATMAGFWKATGRDKAVYDKTKLI
GMRKTLVFYKGRAPNGKKSDWIMHEYRLESDENAP
PQEEGWVVCRAFKKRATGQAKNTETWSSSYFYDEV
APNGVNSVMDPIDYISKQQHNIFGKGLMCKQELEG
MVDGINYIQSNQFIQLPQLQSPSLPLMKRPSSSMSI
TSMDNNYNYKLPLADEESFESFIRGEDRRKKKKQVM
MTGNWRELDKFVASQLMSQEDNGTSSFAGHHIVN
EDKNNNDVEMDSSMFLSEREEENRFVSEFLSTNSDY
DIGICVFDN

Amino acid composition: CSV format

Ala (A)	14	3.8%
Arg (R)	20	5.5%
Asn (N)	22	6.0%
Asp (D)	26	7.1%
Cys (C)	5	1.4%
Gln (Q)	17	4.7%
Glu (E)	31	8.5%
Gly (G)	21	5.8%
His (H)	6	1.6%
Ile (I)	19	5.2%
Leu (L)	20	5.5%
Lys (K)	28	7.7%
Met (M)	16	4.4%
Phe (F)	18	4.9%
Pro (P)	16	4.4%
Ser (S)	31	8.5%
Thr (T)	15	4.1%
Trp (W)	7	1.9%
Tyr (Y)	16	4.4%
Val (V)	17	4.7%
Pyl (O)	0	0.0%
Sec (U)	0	0.0%
(B)	0	0.0%
(Z)	0	0.0%
(X)	0	0.0%

Number of amino acids: 365

Molecular weight: 42476.67

Theoretical pI: 5.40

Total number of negatively charged residues (Asp + Glu): 57
Total number of positively charged residues (Arg + Lys): 48

Structure

Secondary Structure
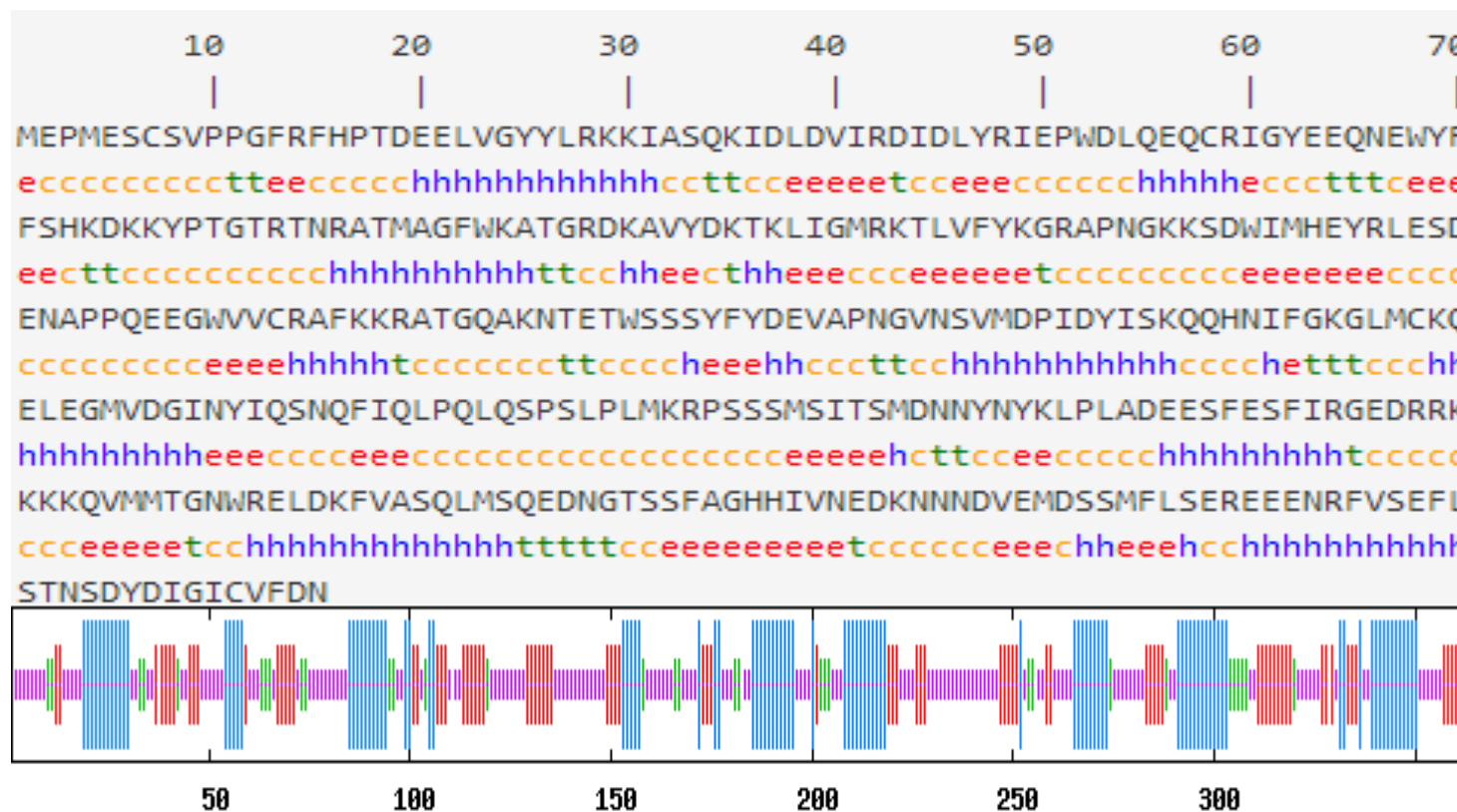
➢ Structure

Secondary Structure

```
# NP_179397.1 Length: 365
# NP_179397.1 Number of predicted TMHs:  0
# NP_179397.1 Exp number of AAs in TMHs: 0
# NP_179397.1 Exp number, first 60 AAs:  0
# NP_179397.1 Total prob of N-in:        0.00392
NP_179397.1      TMHMM2.0        outside      1    365
```


TMHMM posterior probabilities for NP_179397.1

Secondary Structure

SEQUENCE LENGTH: 365

Using the scale **Hphob. / Kyte & Doolittle**, the individual values for the 20 amino acids are:

Ala: 1.800   Arg: -4.500   Asn: -3.500   Asp: -3.500   Cys: 2.500   Gln: -3.500
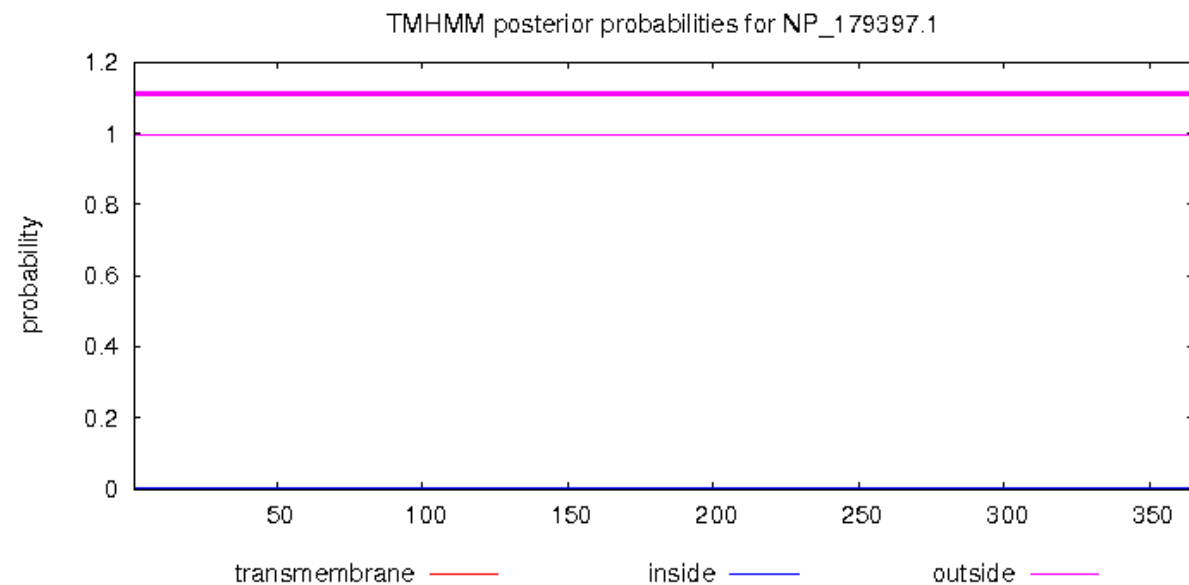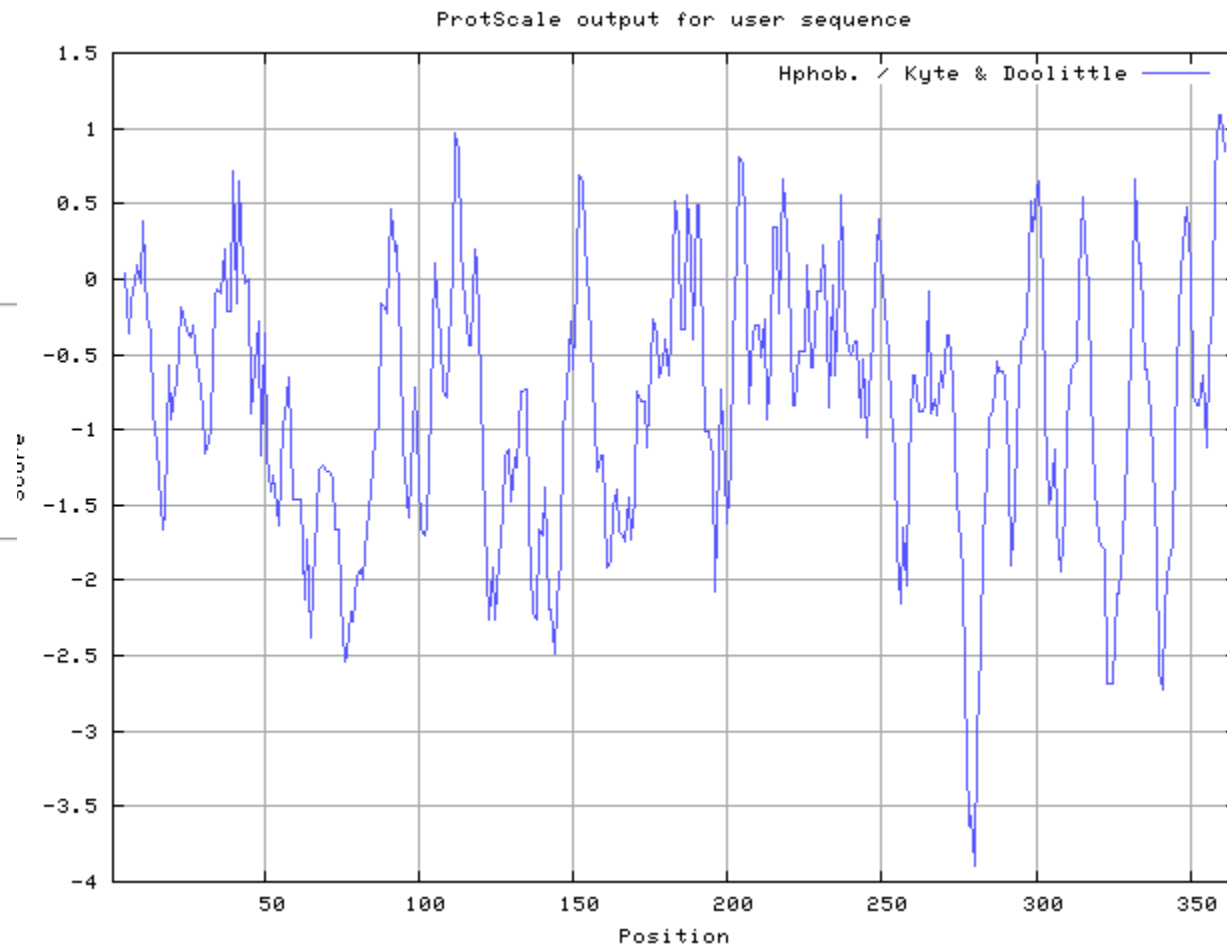Glu: -3.500   Gly: -0.400   His: -3.200   Ile: 4.500   Leu: 3.800   Lys: -3.900
Met: 1.900   Phe: 2.800   Pro: -1.600   Ser: -0.800   Thr: -0.700   Trp: -0.900
Tyr: -1.300   Val: 4.200   : -3.500   : -3.500   : -0.490

Weights for window positions 1,..,9, using **linear weight variation model**:

1	2	3	4	5	6	7	8	9
1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
edge				center				edge

中国农业科学院
Chinese Academy of Agricultural Sciences

Secondary Structure

Predicted localization for the Eukarya domain: Nucleus (GO term ID: GO:0005634) Prediction confidence 64
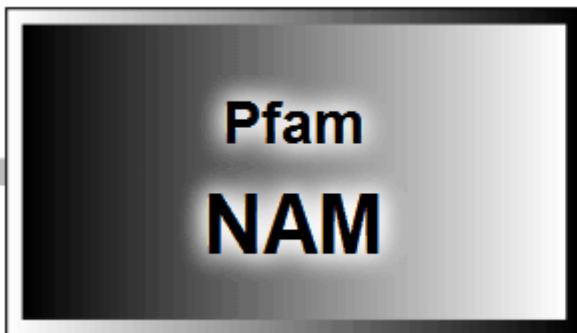
k = 9/23

91.3 %: nuclear
4.3 %: cytoplasmic
4.3 %: cytoskeletal

>> prediction for 151289405818780 is nuc (k=23)

中国农业科学院
Chinese Academy of Agricultural Sciences

➢ Structure

Secondary Structure

DNA-binding domain (DBD) and dimerization domain

Pfam domain sequence (128 aa):

Submit to BLAST **Copy to clipboard**

PPGFRFHPTDEELVGYYLRKKIASQKIDLDVIRDIDLYRIEPWDLQEQCRIGYEEQNEWY
FFSHKDKKYPTGTRTNRATMAGFWKATGRDKAVYDKTKLIGMRKTLVFYKGRAPNGKKSD
WIMHEYRL

This is a *Pfam* domain. Please see the **NAM** entry in *Pfam* for full annotation.

|	
Position:	10 *to* 137
E-value:	1.1e-33 (*HMMER3*)
Accession:	*PF02365*
Description:	
Interpro abstract (IPR003441):	*The NAC domain (for Petunia hybrida (Petunia) NAM and for Arabidopsis ATAF1, ATAF2, and ...(full abstract)*
GO process:	regulation of transcription, DNA-templated (*GO:0006355*)
GO function:	DNA binding (*GO:0003677*)

Pfam domain sequence (128 aa):

中国农业科学院
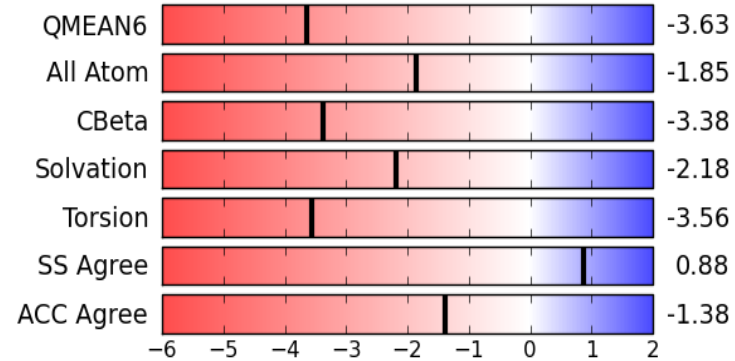**Chinese Academy of Agricultural Sciences**

Structure

3D structure (Swiss Model)

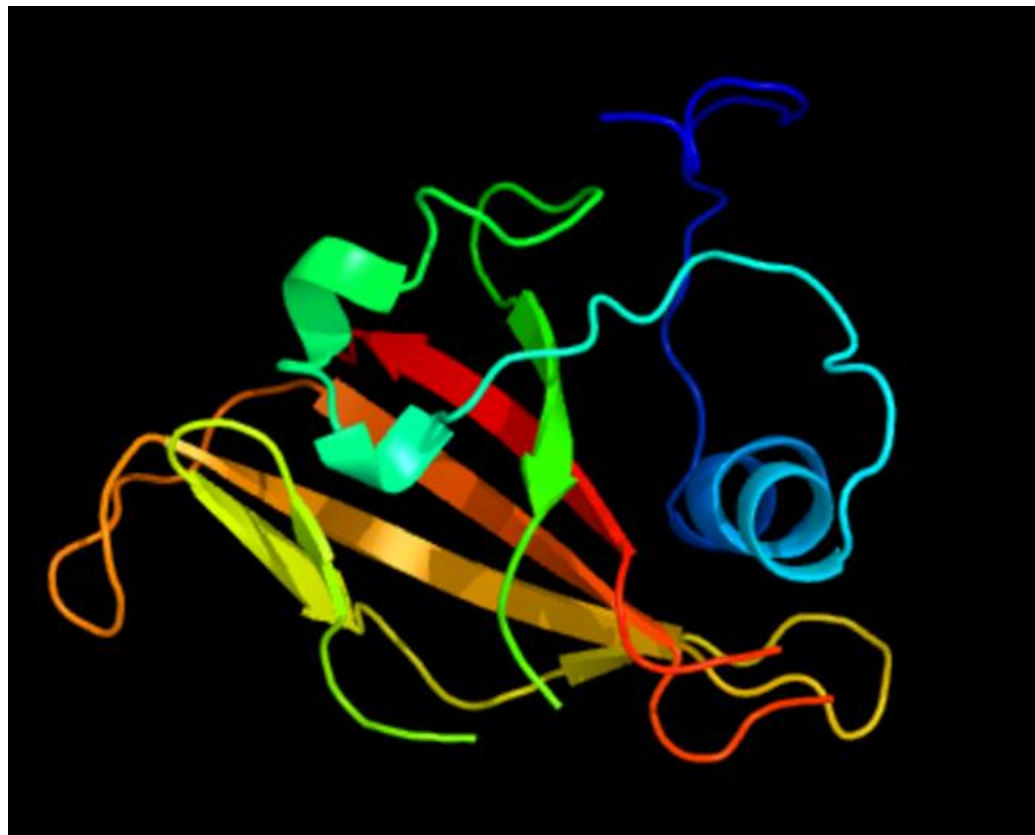中国农业科学院
Chinese Academy of Agricultural Sciences
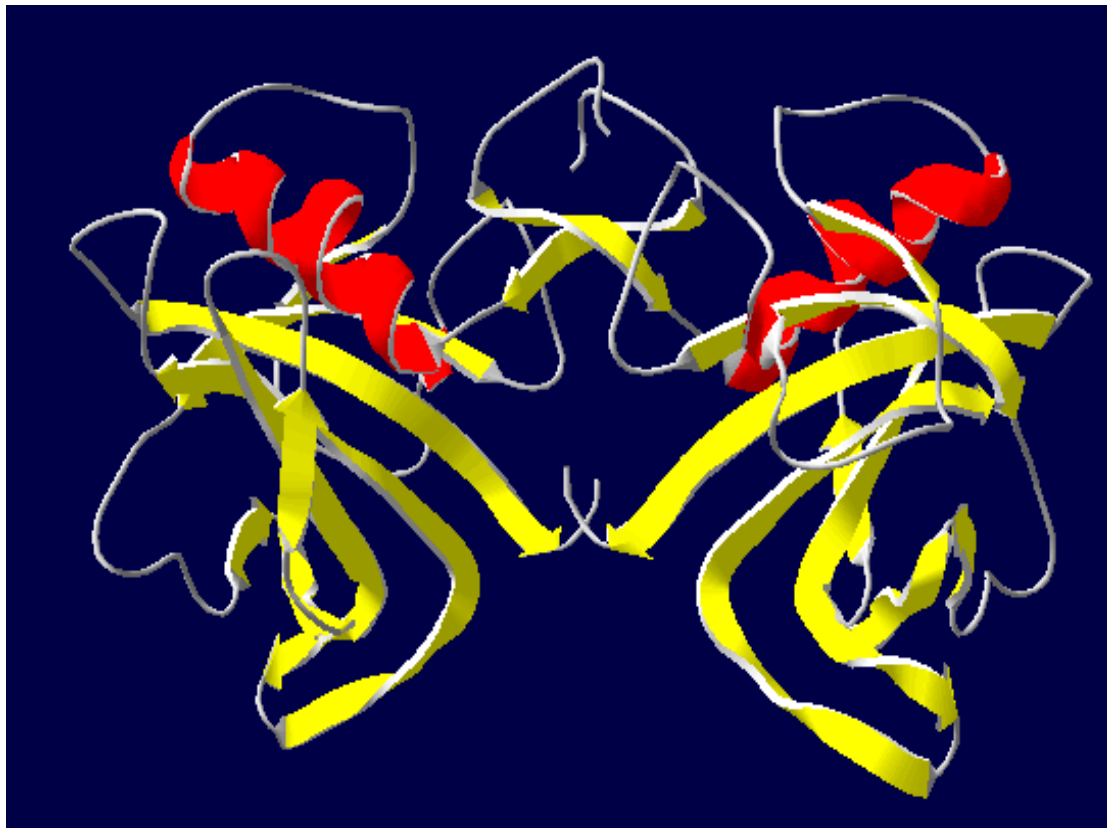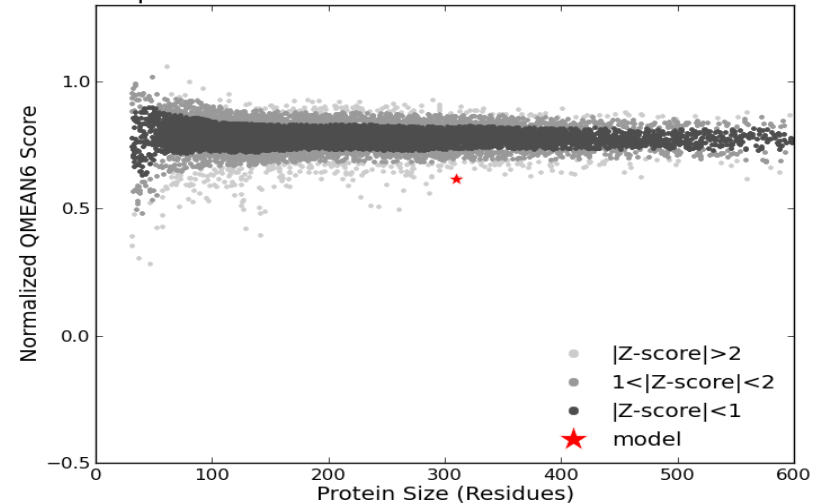
➤ Structure

3D structure (Phary2)

Model (left) based on template c3ulxA

Top template information

PDB header: dna binding protein
Chain: A: PDB Molecule: stress-induced transcription factor nac1;
PDBTitle: crystal structural of the conserved domain of rice stress-responsive2 nac1

Confidence and coverage

Confidence: **100.0%** Coverage: **40%**

146 residues (40% of your sequence) have been modelled with 100.0% confidence by the single highest scoring template.

中国农业科学院
Chinese Academy of Agricultural Sciences

➢ Predicted Functional Partners:

- VND1是植物中参与维管形成的转录因子，具　有一段保守的DNA结合结构域（NAC-domain）,属于NAC转录因子家族

- VND1总共包含365个氨基酸、相对分子量为42476.67、等电点为5.4

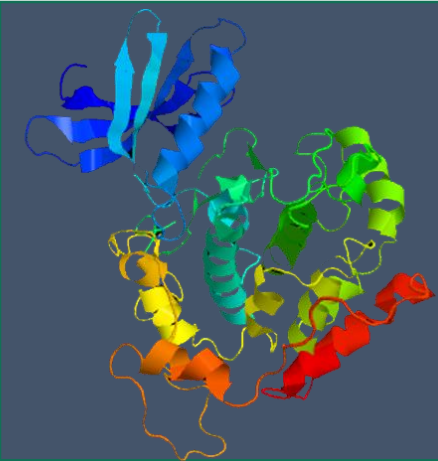- VND1没有跨膜结构，所以其不是跨膜蛋白

- VND1属于亲水蛋白，并且具有核定位信号，　将其位于细胞核中

中国农业科学院
Chinese Academy of Agricultural Sciences

衷心感谢罗老师一学期以来的耐心指导！祝愿罗老师身体健康，工作愉快！

感谢我组所有成员对学习、讨论和实习的配合，感谢班级同学对我组的帮助和指导！

中国农业科学院
Chinese Academy of Agricultural Sciences

请老师批评指正，谢谢！

中国农业科学院
Chinese Academy of Agricultural Sciences