

# In Silicon Cloning and Bioinformatics Analysis of the *Raphanus Sativus* WUS Gene

Ma Guang

Department of Life Sciences, Hengshui College, Hengshui, China  
Email: maaohan@163.com

Received 2013

## ABSTRACT

The complete coding sequence of *Raphanus sativus* WUS gene was obtained by using *Arabidopsis thaliana* gene sequence (NM\_127349) as a probe in silico cloning. Then the hydrophilicity, secondary structure and advanced structure of WUS protein in *Raphanus sativus* are analyzed by using bioinformatics methods. The results show that the cDNA was 1298 bp, with no intron, contains an open reading frame of 936bp, encoding 312aa protein. The protein coded by *Raphanus sativus* gene showed 74% similarity to *Arabidopsis thaliana*.

**Keywords:** *Raphanus sativus*; WUS; Bioinformatics; Stem Cell

## 1. Introduction

Recent years, many genes regulating the plant stem cell development have been found in different tissues. WUS (WUSCHEL) gene is one of those, which encodes a transcription factor that specifies the adjacent cells to be stem cell [1]. A negative feedback loop that couples the antagonistic functions of the WUSCHEL and CLAVATA loci controls stem cell fate in the shoot apical meristem. In the embryonic meristem, the expression of CLV3 depends on WUS only. During post-embryo development, both WUS and STM are needed for CLV3 expression and the triggering of organogenesis [2]. In floral meristem, the expression of AG is activated by coexistence of WUS and LFY, AG acts as a negative regulator of WUS expression to down regulate the WUS level. The signal system established by WUS is also involved in ovule development [3]. The somatic embryogenesis can be promoted efficiently by WUS, especially in the presence of auxin [4]. The results of previous works indicated that cell competence to WUS activity is related to microenvironment and the combination of WUS signal with different environmental factors could activate different downstream genes [5].

Because the WUS gene is important in plant stem cell, it becomes necessary to clone the gene sequence and elucidate the gene structure. To date, the WUS gene in *Arabidopsis thaliana* was reported [6], but the cloning of *Raphanus sativus* WUS gene has not been reported. In this study, we cloned *Raphanus sativus* WUS based on in the silico cloning strategy. The sequence of *Raphanus sativus* WUS was analyzed by bioinformatics methods,

including open reading frame (ORF) analysis, BLAST, protein structure, physical and chemical properties prediction.

## 2. Materials and Methods

### 2.1. Bioinformatics Tools

<http://www.ncbi.nlm.nih.gov>  
<http://mobylye.pasteur.fr/cgi-bin/portal.py?form=cap3>  
<http://www.ncbi.nlm.nih.gov/gorf/gorf.html>  
<http://blast.ncbi.nlm.nih.gov/Blast.cgi>  
<http://frodo.wi.mit.edu/primer3>  
<http://www.expasy.org/tools/protparam.html>  
[http://npsa-pbil.ibcp.fr/cgi-bin/npsa\\_automat.pl?page=/NPSA/npsa\\_sopma.html](http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_sopma.html)  
<http://www.expasy.org/cgi-bin/protscale.pl>  
<http://swissmodel.expasy.org>

### 2.2. BLAST Searching of *Raphanus sativus* EST and Genomic Databases

The cDNA sequence of *Arabidopsis thaliana* protein WUSCHEL (WUS) mRNA, complete cds (GenBank, NM\_127349) was used as a probe to search the *Raphanus sativus* expressed sequence tag (EST) database (<http://www.ncbi.nlm.nih.gov/dbEST>) for a homologous clone, using the BLAST program. The EST sequence of score  $\geq 100$  and length  $\geq 100$  bps selected from the blast result, were generated contigs (<http://pbil.univ-lyon1.fr/cap3.php>). The longer contig was used as second probe. The above step was not repeated until the newly generated probe cannot be elongated. This approach led to a

sequence as a putative *Raphanus sativu* WUS cDNA.

### 2.3. Bioinformatics Analysis of the *Raphanus sativu* WUS Gene

The putative *Raphanus sativu* WUS cDNA was analyzed by on-line bioinformatics softwares. The open reading frame (ORF), protein structure, physical and chemical properties prediction *Raphanus sativu* WUS was studied.

## 3. Results

### 3.1. Identification of EST Sequence Containing Part of Putative *Raphanus sativu* WUS

With the cds of *Arabidopsis thaliana* protein WUSCHEL (WUS) mRNA (GenBank, NM\_127349) as probe, 16 EST sequences (score  $\geq 100$  and length  $\geq 100$ ) were found by Blastn searching the EST database, in NCBI (Figure 1).

The sequences were selected saved in a file with FASTA format. The file was submitted to on-line CAP software. Three contigs were obtained by assembling EST sequences. The longest contig of 1298 bp was selected as putative *Raphanus sativu* WUSCHEL (WUS) cDNA (Figure 2). Searching the cDNA sequence for potential coding regions by ORF finder (NCBI), an entire open reading frame (ORF) of 312 amino acids was detected with a potential start codon at the 132rd site and a stop codon at the 1069th site (Figure 2)

### 3.2. Gene Comparison between *Arabidopsis thaliana* and *Raphanus sativu* WUS cDNA by ClustalW2

The alignment of *Arabidopsis thaliana* and *Raphanus sativu* WUS amino acid sequences, constructed using ClustalW program, suggested that *Raphanus sativu* WUS amino acid sequences were very similar to *Arabidopsis thaliana* amino acid sequences. There were same 74.0% amino acid in *Arabidopsis thaliana* and *Raphanus sativu* WUS (Figure 3).

### 3.3. Secondary Structure of *Raphanus sativu* WUS

The secondary structure of *Raphanus sativu* WUS was

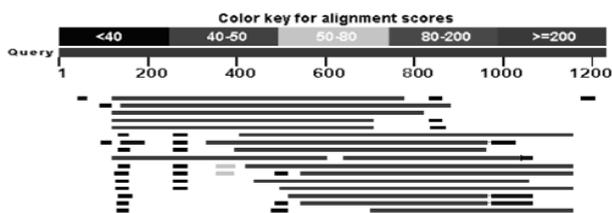


Figure 1. BLAST result of wus cDNA in the *Raphanus sativu* EST databank.

```

GGATATCTCCCCACATACAAAACCTAAAATCT
CTTTTGCTTAAAGTTCTTTTATGGTAAAGTAAA
AGTTCATTCTCTCACITTAATCACTGCTTCT
TCTTCCGATCATTITTTTCATCTTGATCTTGAA
atggagcaaccgcaatatcataacccatcacatcagcaaacgcat
M E Q P Q Y H N H H H Q Q T D
caagagagcggcacaacaacaagaatccggctctggtggttaccgg
Q E S G N N N K S G S G G Y P
tgtcgccaaacgagcacaagaatggacaccaacgacggagcagatc
C R Q T S T R W T P T T E Q I
agaatctcaaaagatctctactacaacaacggagttcggtcacca
R I L K D L Y Y N N G V R S P
acagccgaccagatccagaagaatctctgcacggctgagacagatc
T A D Q I Q K I S A R L R Q Y
ggaagatcggaggtaagaacgtcttttactggttccagaaccac
G K I E G K N V F Y W F Q N H
aaggctcgtgagcgacagaagaagagattcagcggcacaacctag
K A R E R Q K K R F S G T M
accacaccaacatcttcatcccccaactcggttatgatggccaac
T T P T S S S P N S V M M A N
gaccattatcatcatcaacatcatcctcttctgcatcatcatcat
D H Y H H Q H H P L L H H H H
catggtgttaccatgcataggcctgcttccggtcaatattaagctt
H G V T M H R P A S V N I K L
gaccagaaaaatcatctccttcatcagaacatcatatgccaac
D Q E N H L L H Q N I S Y A N
tttcataccgggaatttaaatcatgcaagttcaggcactgaaatgt
F H N G N L N H A S S G T E C
ggtgctcttaaatgcttctcccagtggtctacatgagtagccatctc
G A L N A S P S G Y M S H L
tatggatctatggaacaagattatcagtgagctacaacaactta
Y G S M E Q D Y S V S Y N N L
ggcggagatggacaacatgagatcataatcatcattactcagct
G G G W T N M D H N H H Y S A
ccagcttacaacttctctogtagaccaagcctctgctctggacta
P A Y N F F D R P K P L S G L
aaagatcacgaagaagaaggagaatattggtggcagatcttattctg
K D H E E E G E Y G G D A Y L
gaacatagacgtacacttctctctcctatgcacgggtgaggat
E H R R T L S L F P M H G E D
cacatcaacggcagtggtgatgcccttttgaagtacggacaatcg
H I N G S G D A F L K Y G Q S
gacgatcgtgattattatggtagaggcccttcttctctcaag
D D R D Y Y G R G P C A S L K
ttatgtctaaactcctacgcaggcgtctcaccggattaa
L C L N S Y A G V S P D *
TGATGCAGTAATCATAATAATAATAGTATCTCTC
GTACGTAGTTTATGTTATCTGTAATTACTATCGTCT
TTCGGGTTAGCTTTCATTGAGTGCCTTTTTCTCC
TTCTTTTAAATGTAAGACGTTGAACTTATATTAG
TATTACAAATAA GAACTCCAAAAAAAAAAAAAAAA
AAAAAAAA
    
```

Figure 2. Nucleotide sequence and deduced amino acid sequence of WUS of *Raphanus sativu*. The asterisk indicates the stop codon.

predicted utilizing online service ([http://npsa-pbil.ibcp.fr/cgi-bin/npsa\\_automat.pl?page=/NPSA/npsa\\_sopma.html](http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_sopma.html)). The results showed that *Raphanus sativu* WUS of 19.87% alpha helix, 10.58% extended strand, 66.03% random coil and 3.53% beta turn.

### 3.4. Hydrophilicity Prediction of *Raphanus Sativu* WUS

Hydrophilicity of *Raphanus sativu* WUS was predicted utilizing Program of ProtScale (Kyte and Doolittle). The results showed that most sites of *Raphanus sativu* WUS are in the hydrophilic region (Score: 0 ~ -0.35) (Figure 4). It was concluded that the *Raphanus sativu* WUS is a hydrophilic protein.

### 3.5. Advanced Structure of *Raphanus sativu* WUS

The structure prediction from primary to advanced

```

Arabidopsis MEPPQ-HQHSHHQADQESGNNMNNKSGSGGYTCRQTSTRWTPTEQIKILKELYNNNAIR 59
Raphanus   MEQPQYHNNHHQQTQESGNN--NKSGSGGYPCRQTSTRWTPTEQIRILKLDLYNMGVR 58
          ** ** *:*:*:*:***** *****.*****:***:*****.*
          *

Arabidopsis SPTADQIQKITARLRQFGKIEGKNVYWFQNHKARERQKKRFNGTMMTP-SSSPNSVMM 118
Raphanus   SPTADQIQKISARLRQYQKIEGKNVYWFQNHKARERQKKRFSGTMMTPSSSPNSVMM 118
          *****:*****:*****.*****.**,**** *****

Arabidopsis AAN---DHYHPLLHHHH-GVPMQRPANSVNVKLNQDHHLYHNNKPYPSFNNGNLNHASSG 174
Raphanus   ANDHYHHQHHP LLHHHHGVTMHRPA-SVNIKLDQENHLHQNISYANFHNGNLNHASSG 177
          * : .:***** **:*:* **:*:*:*:*:* *:* *.*.*:*****

Arabidopsis TECGVVNAS-NGYMSHVGYSMEQDCSMNYNVGGGWANMDH--HYSSAPYNFFDRAKPL 231
Raphanus   TECGALNASPSGYMSSHLYGSMEQDYSVSYNMLGGGWITMDHNHYSAPAYNFFDRPKPL 237
          ****.:*** .*****:***** *:.*:*:*:*:* ****:..*****.***

Arabidopsis FGLEGHQEEEECGDAYLEHRRTLPLFPMHGEDHINGSGAIWKYQSEVR-----PCA 285
Raphanus   SGLKDHEEEGEYGGDAYLEHRRTLSLFPMHGEDHINGSGDAFLKYQSDDRDYGRGPCA 297
          **:.*:*:* * *****.*****..*:* *****: * ***

Arabidopsis SLELRN----- 292
Raphanus   SLKLCLSYAGVSPD 312
          ***.***

```

Figure 3. Homology of *Arabidopsis thaliana* and *Raphanus sativu* WUS.

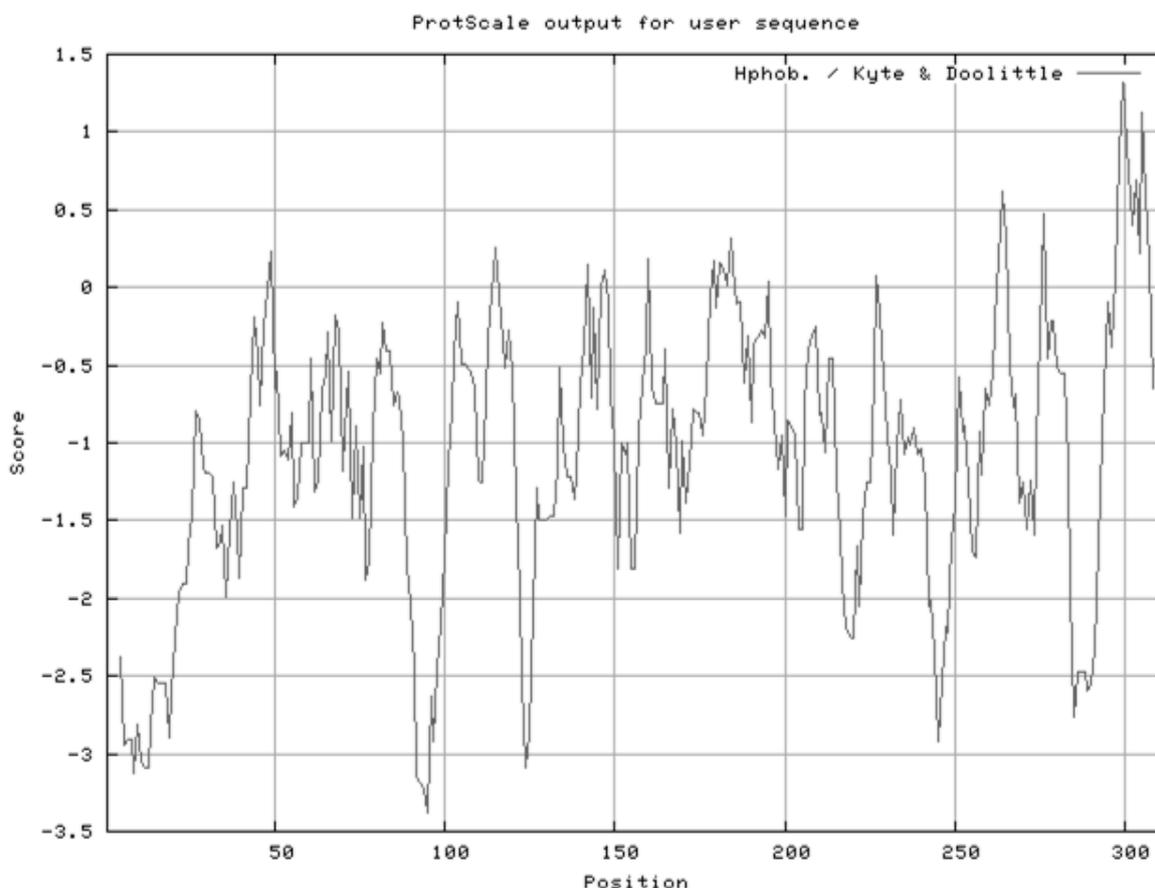


Figure 4. Hydrophilicity profile of *Raphanus sativu* WUS.

structure is an important task in field of the protein research. The three-dimensional structure model of *Raphanus sativu* WUS was predicted by the Swiss-Model

server, by homology modeling, based on the available structures. The result showed that there were 3  $\alpha$ -helices some irregular coiled peptides *Raphanus sativu* WUS

protein (Figure 5).

#### 4. Discussion

In silico cloning is a method developed in recent years for functional gene identification by using genome and EST database. Compared to the traditional methods, such as molecular hybridization, screening of genomic or cDNA library, it is advanced for low cost, high efficiency, easy operation, etc [7]. With more and more EST and genome sequencing data were reported, it would become possible and feasible to isolate and identify the functional genes by in silico cloning. Many successful examples strongly support the fact that in silico cloning is absolutely a feasible tool for gene cloning and presents some advantages, compared to the traditional methods [8]. In this study, the full cDNA of *Raphanus sativu* WUS was obtained primarily by searching and splicing the EST sequences. The structure and function were analyzed and predicted using bioinformatics methods successfully. The results revealed that it is a convenient technique for cloning novel gene by searching EST database with homologous gene of model living things. To our knowledge, it was the first report about cloning of *Raphanus sativu* WUS cDNA with in silico cloning. This research



Figure 5. The predicted three-dimensional structure of *Raphanus sativu* WUS.

achievement will provide theory and reference for plant stem research in *Raphanus sativu*.

#### 5. Acknowledgements

The research is supported by fund of Hebei province colleges and universities (No. Z2010159) and fund of Hengshui city science and technology research (No. 09055Z).

#### REFERENCES

- [1] I. Baurle and T. Laux, "Regulation of WUSCHEL Transcription in the Stem Cell Niche of the *Arabidopsis* Shoot Meristem," *Plant Cell*, Vol. 17, 2005, pp. 2271-2280. <http://dx.doi.org/10.1105/tpc.105.032623>
- [2] Y. Y. Xu and K. Chong, "Progress in Research on Plant Stem Cell Organizer Gene WUSCHEL," (in Chinese) *J. Plant Phyol. Mol. Bio.*, Vol. 31, 2005, pp. 461-468.
- [3] G. Sena, X. Wang, H. Y. Liu, H. Hofhuis and K. D. Birnbaum, "Organ Regeneration Does Not Require a Functional Stem Cell Niche in Plants," *Nature*, Vol. 457, 2009, pp. 1150-1153. <http://dx.doi.org/10.1038/nature07597>
- [4] T. Kondo, S. Sawa A. Kinoshita, S. Mizuno, T. Kakimoto, H. Fukuda and Y. Sakagami, "A Plant Peptide Encoded by CLV3 Identified by *in Situ* MALDI-TOF MS Analysis," *Science*, Vol. 313, 2006, pp. 845-848. <http://dx.doi.org/10.1126/science.1128439>
- [5] U. Brand, M. Grunewald, M. Hobe and R. Simon, "Regulation of CLV3 Expression by Two Homeobox Genes in *Arabidopsis*," *Plant Physiology*, Vol. 129, 2002, pp. 565-575. <http://dx.doi.org/10.1104/pp.001867>
- [6] T. Laux, F. X. Klaus, Mayer, B. Berger and G. Juren, "The WUSCHEL Gene Is Required for Shoot and Floral Meristem Integrity in *Arabidopsis*," *Development*, Vol. 122, 1996, pp. 87-96.
- [7] H. M. Zhang, M. G. Jiang and Y. J. Feng, "In Silico Cloning of MgEno-1 cDNA from *Magnaporthe grisea*," *China J. Bioinformatics*, 2005, pp. 57-61.
- [8] H. Li, G. Y. Zhou, H. Y. Zhang, L. Lin and J. Liu, "In Silico Cloning and Bioinformatic Analysis of PEPCK Gene in *Fusarium oxysporum*," *African Journal of Biotechnology*, Vol. 9, 2010, pp. 1864-1870.