

水稻寡分蘖新突变基因 *TUB2* 的生物

信息分析

The biological information analysis of the rice fewer tillerings mutation *TUB2*

报告人: 邱海阳

郝陆洋

王杰

研究背景

基因克隆

蛋白序列分析与结构预测

未来研究计划

1.1 研究背景

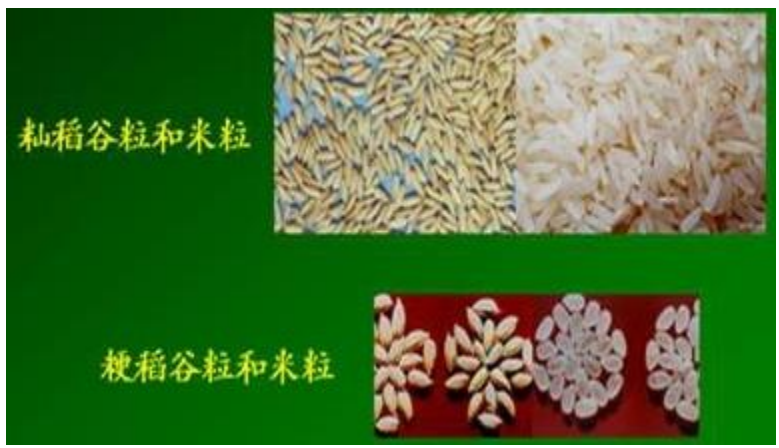
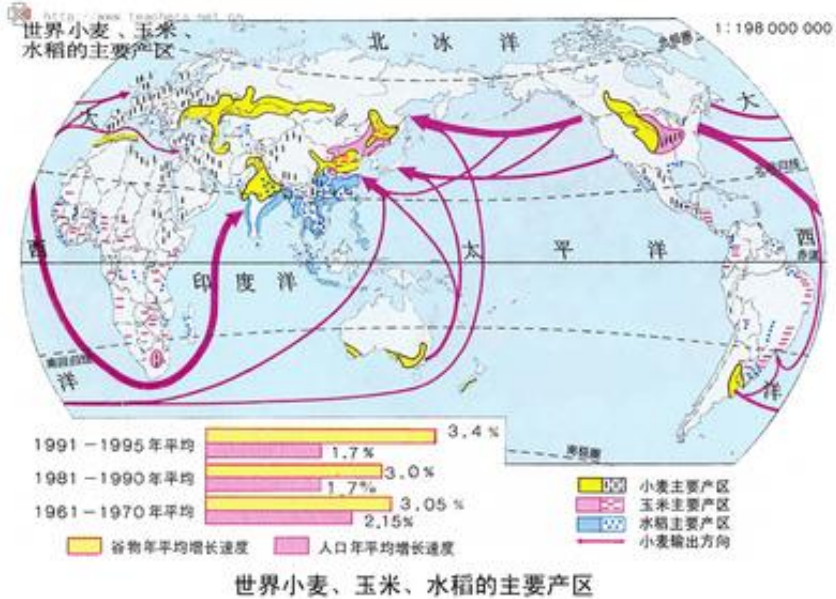
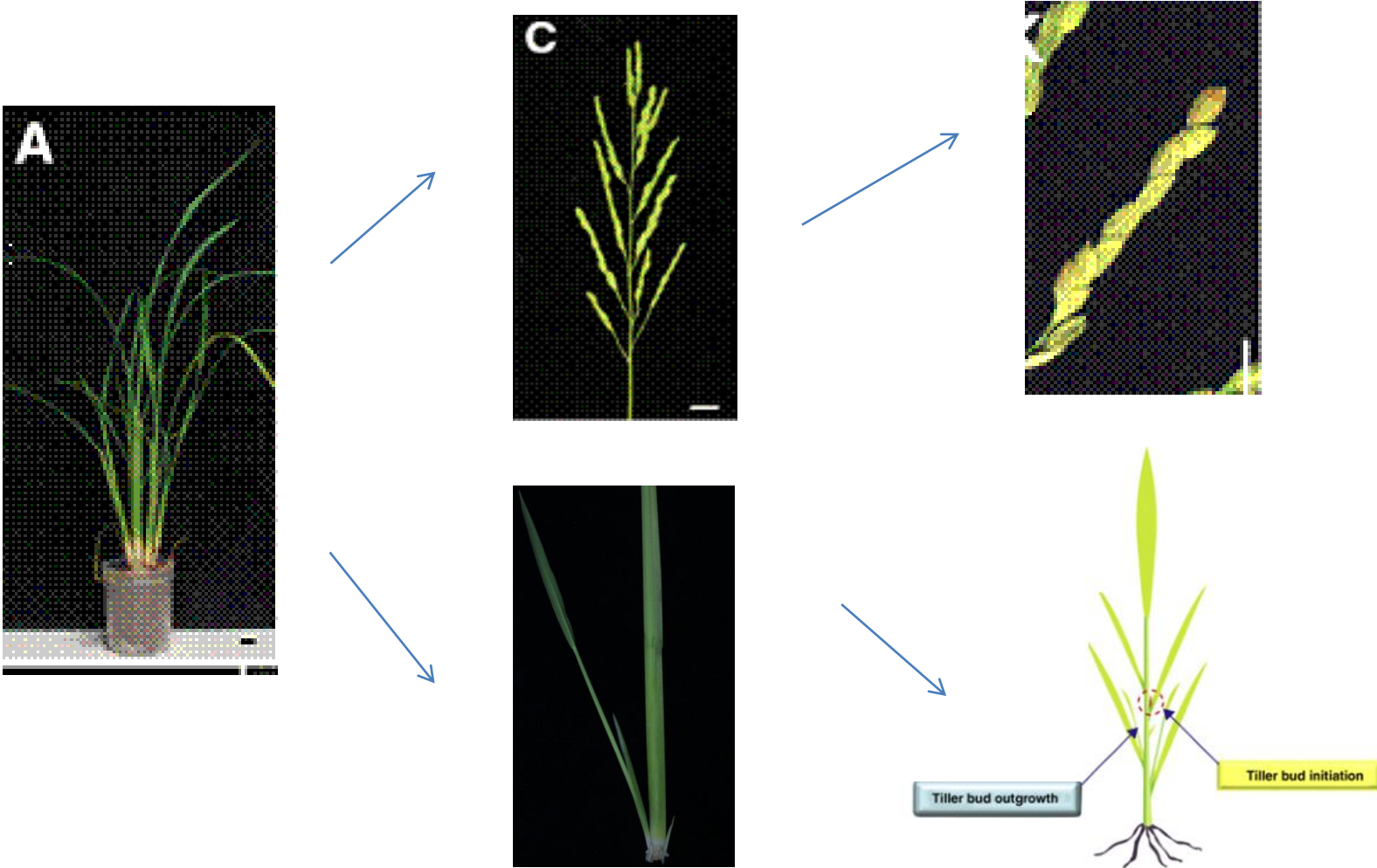


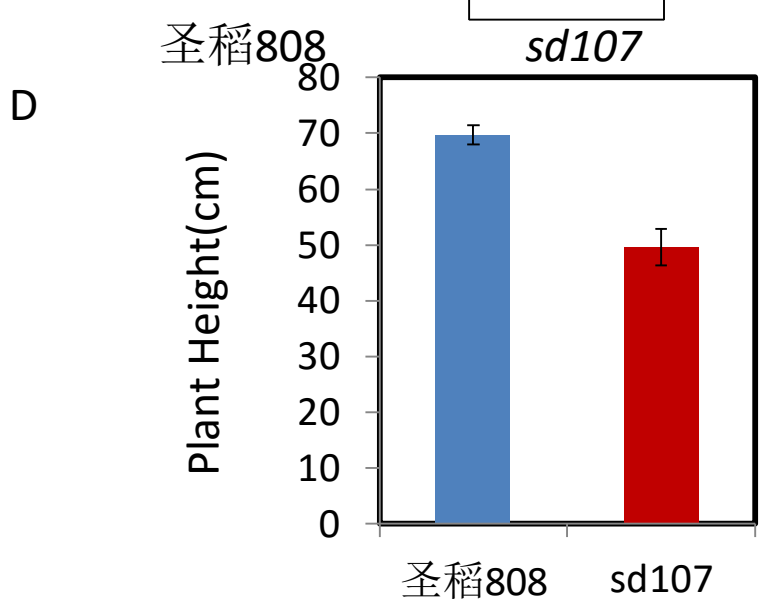
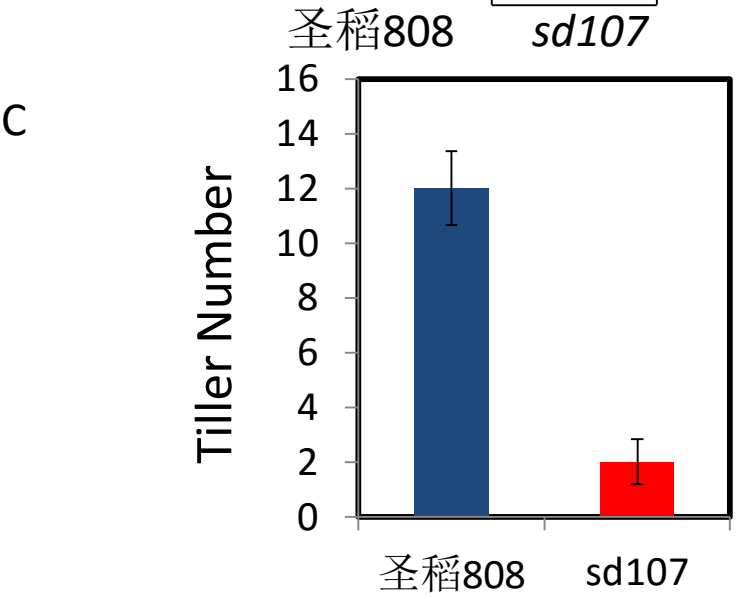
表1 籼稻和粳稻的形态比较

形态特征		籼 稻	粳 稻
叶	宽 度	较 宽	较 窄
	色 泽	淡 绿	深 绿
	剑叶开度	小	大
	茸 毛	较 多	较少或无茸毛
稈 毛	短而稀，散生在稈面上		长而密，集生在稈棱上
芒	多无芒，有芒时多为直立形短芒		从长芒-无芒，芒略呈弯曲状
粒 形	细而长，稍扁平		短而宽，较厚
脱粒性	易脱粒		难脱粒
谷粒对石炭酸反应	能为石炭酸染色且染色较深		不为石炭酸染色或染色较浅

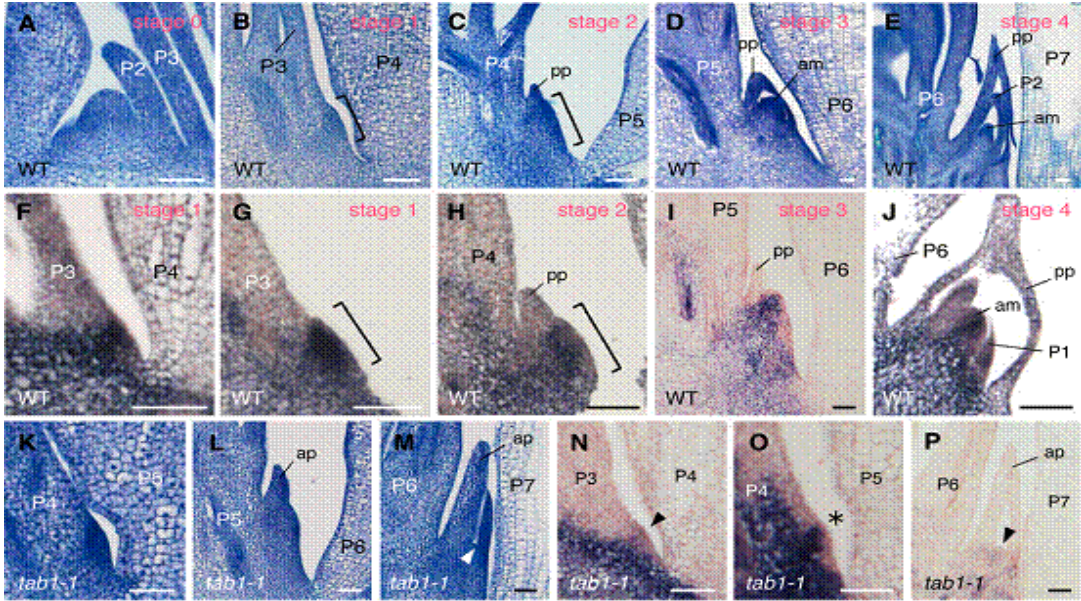
1.2 影响水稻产量的主要因素



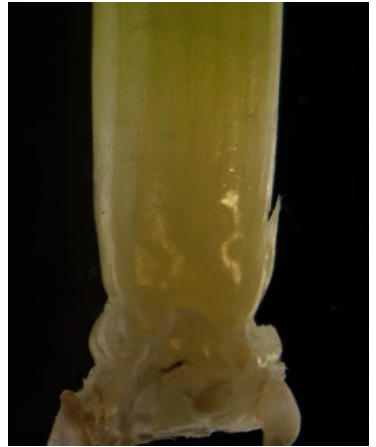
1.3 *sd107*植株，根部表型分析



1.4 *sd107* 植株分蘖芽表型分析



圣稻808(45days)



sd107(45days)

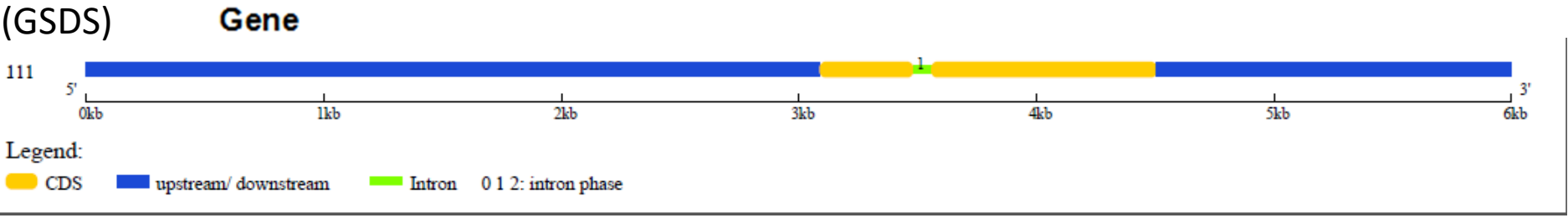
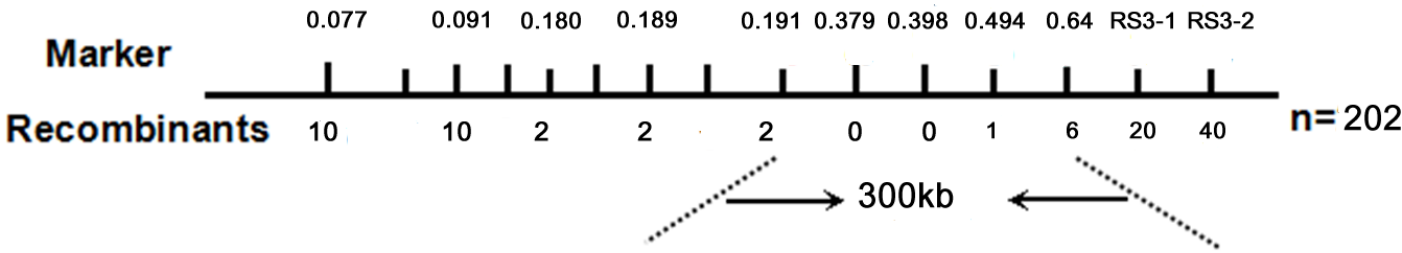
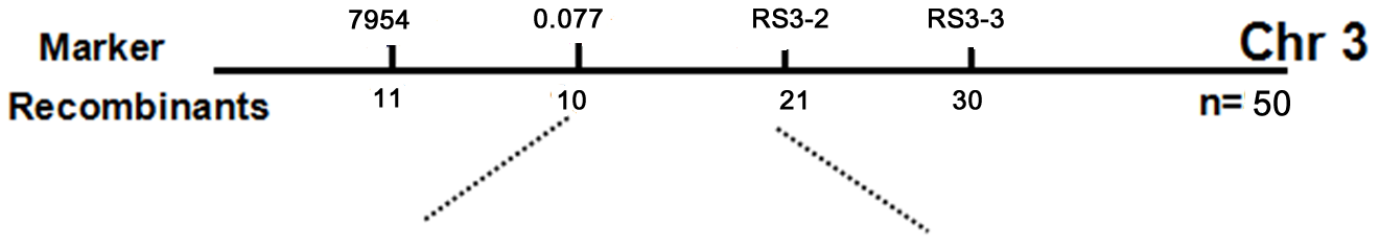


圣稻808(70days)



sd107(70days)

2.1 *SD107*基因的基因定位与克隆

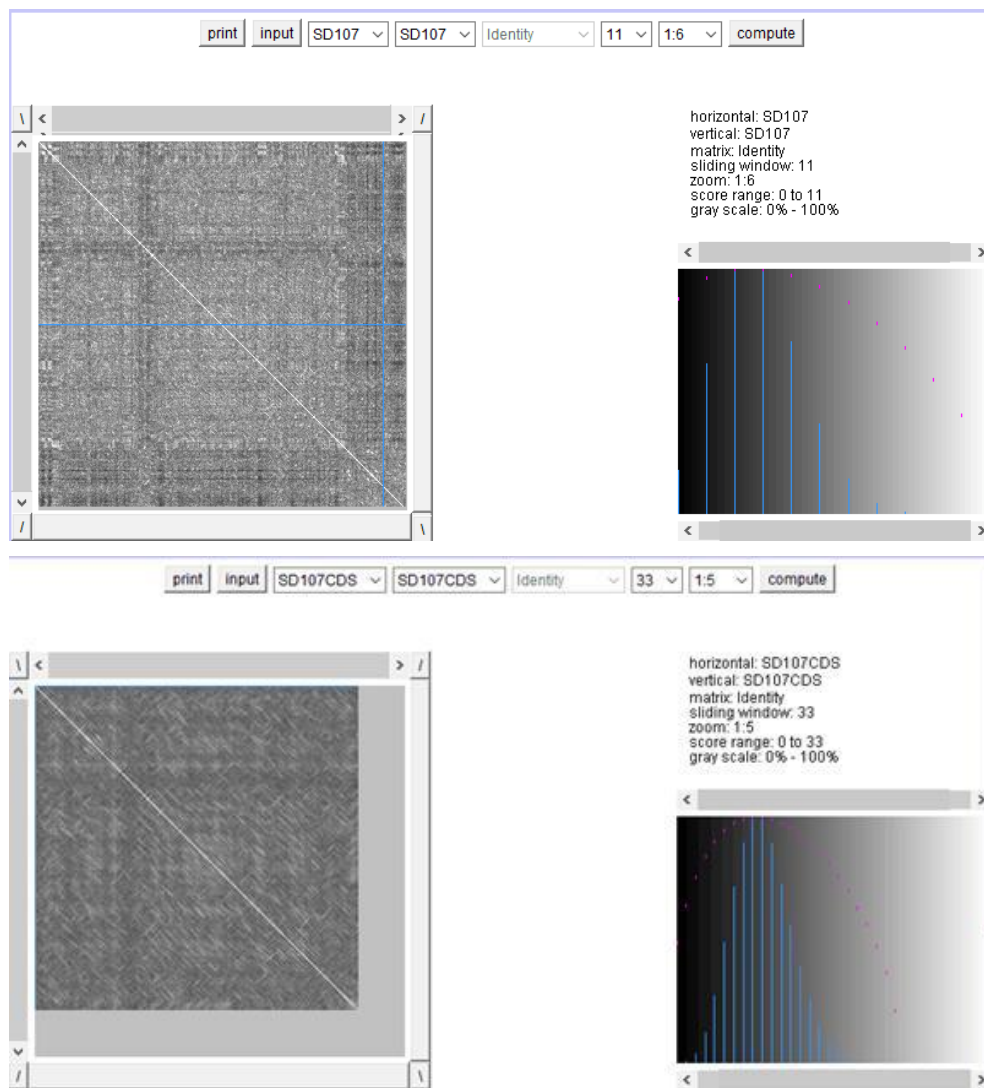


2.1 *SD107*基因的基因定位与克隆



LOC_Os03g01530 sequence information Genomic sequence length: 1811 nucleotides
CDS length: 1344 nucleotides
Protein length: 447 amino acids
Putative Function: tubulin/FtsZ domain containing protein, putative, expressed

2.2 dotlet分析



该基因全长和CDS长不存在重复区域

2.3 NCBI-Blastp

分别搜索Refseq和Swiss-Prot数据库，所选物种为Oryza Sativa, 其它参数如图上，经过尝试，发现BLOSUM值越高，PAM值越小，搜索同源基因相似度越高；Word Size越大，准确度越高，精度越差。

General Parameters

- Max target sequences: 250
- Short queries: Automatically adjust parameters for short input sequences
- Expect threshold: 0.01
- Word size: 3
- Max matches in a query range: 0

Scoring Parameters

- Matrix: BLOSUM90
- Gap Costs: Existence: 11 Extension: 1
- Compositional adjustments: Conditional compositional score matrix adjustment

Filters and Masking

- Filter: Low complexity regions

Putative conserved domains have been detected, click on the image below for detailed results.

Query seq. nucleoside binding site Taxol binding site alpha/beta domain interface beta_tubulin Tubulin_FtsZ_Cetz-like superfamily PLN00220

Distribution of 13 Blast Hits on the Query Sequence

Mouse-over to show define and scores, click to show alignments

Color key for alignment scores
<40
40-50
50-80
80-200
>=200

Query 1 80 160 240 320 400

2.4 同源比对

从NCBI Blast和Uniprot Blast 中分别找出*OsTUB2*的同源基因，并用MEGA中ClustalW进行序列比对，然后用Genedoc分析结果。

```

*          20          *          40          *          60          *          80          *          100         *          120
OsTUB2 : MREILHIQGGCCGNQIGAKFWEVVC AEHGIDATGRYVCD -SDLQLERVNYYNEASCGRFV PRAVLMDLEPGTMDSVRSRGPYGH IFRPDNFVFGQSGAGNNWAKGHYTEGAELIDAVLD : 118
AtTUB8 : MREILHIQGGCCGNQIGAKFWEVVC AEHGIDSTGRYVCE -NDLQLERVNYYNEASCGRFV PRAVLMDLEPGTMDSVRSRGPYGH IFRPDNFVFGQSGAGNNWAKGHYTEGAELIDAVLD : 118
AtTUB3 : MREILHIQGGCCGNQIGAKFWEVVC AEHGIDPTGRYVCT -SDLQLERVNYYNEASCGRFV PRAVLMDLEPGTMDSVRSRGPYGH IFRPDNFVFGQSGAGNNWAKGHYTEGAELIDAVLD : 118
AtTUB2 : MREILHIQGGCCGNQIGAKFWEVVC AEHGIDPTGRYVCT -SDLQLERVNYYNEASCGRFV PRAVLMDLEPGTMDSVRSRGPYGH IFRPDNFVFGQSGAGNNWAKGHYTEGAELIDAVLD : 118
OsTUB4 : MREILHIQGGCCGNQIGAKFWEVVC DEHGIDPTGRYVCTN -SDLQLERVNYYNEASCGRFV PRAVLMDLEPGTMDSVRTGPYGH IFRPDNFVFGQSGAGNNWAKGHYTEGAELIDAVLD : 118
OsTUB1 : MREILHIQGGCCGNQIGAKFWEVVC DEHGIDPTGRYVCT -SDLQLERVNYYNEASCGRFV PRAVLMDLEPGTMDSVRTGPYGH IFRPDNFVFGQSGAGNNWAKGHYTEGAELIDAVLD : 118
AtTUB6 : MREILHIQGGCCGNQIGAKFWEVVC DEHGIDPTGRYVCTN -SDLQLERVNYYNEASCGRFV PRAVLMDLEPGTMDSVRTGPYGH IFRPDNFVFGQSGAGNNWAKGHYTEGAELIDAVLD : 118
AtTUB7 : MREILHIQGGCCGNQIGAKFWEVVC DEHGIDPTGRYVCT -SDLQLERVNYYNEASCGRFV PRAVLMDLEPGTMDSVRSRGPYGH IFRPDNFVFGQSGAGNNWAKGHYTEGAELIDAVLD : 118
ZmTUB5 : MREILHIQGGCCGNQIGAKFWEVVC DEHGIDPTGRYVCT -SDLQLERVNYYNEASCGRFV PRAVLMDLEPGTMDSVRTGPYGH IFRPDNFVFGQSGAGNNWAKGHYTEGAELIDAVLD : 118
ZmTUB6 : MREILHIQGGCCGNQIGAKFWEVVC DEHGIDPTGRYVCTN -SDLQLERVNYYNEASCGRFV PRAVLMDLEPGTMDSVRTGPYGH IFRPDNFVFGQSGAGNNWAKGHYTEGAELIDAVLD : 118
ZmTUB1 : P----- --RAVLMDLEPGTMDSVRSRGPYGH IFRPDNFVFGQSGAGNNWAKGHYTEGAELIDAVLD : 58
AtTUB5 : MREILHIQGGCCGNQIGAKFWEVVC DEHGIDSTGRYVCTADLQLERVNYYNEASCGRFV PRAVLMDLEPGTMDSVRSRGPYGH IFRPDNFVFGQSGAGNNWAKGHYTEGAELIDAVLD : 119
mreilhigggccgnqig kfwev c ehgid tgr y g dlqler nvyneascgr vpra61mdlepgtmds6r3gp5gqifrdpndfvfgqsgagnnwakghytegaelid vld

```

```

0          *          140         *          160         *          180         *          200         *          220         *          240
OsTUB2 : VVrKEAENCdCLQGfQVChSlGGGTGSgMGtLLlSkIReEYpDRmMLtFsvFpSPkVSDtVvEpyNatLsvHqLveNAdECmVLDNeAlYdICfRtLkLTPsFGdLNhLIATmSGVt : 237
AtTUB8 : VVrKEAENCdCLQGfQVChSlGGGTGSgMGtLLlSkIReEYpDRmMLtFsvFpSPkVSDtVvEpyNatLsvHqLveNAdECmVLDNeAlYdICfRtLkLTPsFGdLNhLIATmSGVt : 237
AtTUB3 : VVrKEAENCdCLQGfQVChSlGGGTGSgMGtLLlSkIReEYpDRmMLtFsvFpSPkVSDtVvEpyNatLsvHqLveNAdECmVLDNeAlYdICfRtLkLTPsFGdLNhLIATmSGVt : 237
AtTUB2 : VVrKEAENCdCLQGfQVChSlGGGTGSgMGtLLlSkIReEYpDRmMLtFsvFpSPkVSDtVvEpyNatLsvHqLveNAdECmVLDNeAlYdICfRtLkLTPsFGdLNhLIATmSGVt : 237
OsTUB4 : VVrKEAENCdCLQGfQVChSlGGGTGSgMGtLLlSkIReEYpDRmMLtFsvFpSPkVSDtVvEpyNatLsvHqLveNAdECmVLDNeAlYdICfRtLkLTPsFGdLNhLIATmSGVt : 237
OsTUB1 : VVrKEAENCdCLQGfQVChSlGGGTGSgMGtLLlSkIReEYpDRmMLtFsvFpSPkVSDtVvEpyNatLsvHqLveNAdECmVLDNeAlYdICfRtLkLTPsFGdLNhLIATmSGVt : 237
AtTUB6 : VVrKEAENCdCLQGfQVChSlGGGTGSgMGtLLlSkIReEYpDRmMLtFsvFpSPkVSDtVvEpyNatLsvHqLveNAdECmVLDNeAlYdICfRtLkLTPsFGdLNhLIATmSGVt : 237
AtTUB7 : VVrKEAENCdCLQGfQVChSlGGGTGSgMGtLLlSkIReEYpDRmMLtFsvFpSPkVSDtVvEpyNatLsvHqLveNAdECmVLDNeAlYdICfRtLkLTPsFGdLNhLIATmSGVt : 237
ZmTUB5 : VVrKEAENCdCLQGfQVChSlGGGTGSgMGtLLlSkIReEYpDRmMLtFsvFpSPkVSDtVvEpyNatLsvHqLveNAdECmVLDNeAlYdICfRtLkLTPsFGdLNhLIATmSGVt : 237
ZmTUB6 : VVrKEAENCdCLQGfQVChSlGGGTGSgMGtLLlSkIReEYpDRmMLtFsvFpSPkVSDtVvEpyNatLsvHqLveNAdECmVLDNeAlYdICfRtLkLTPsFGdLNhLIATmSGVt : 237
ZmTUB1 : VVrKEAENCdCLQGfQVChSlGGGTGSgMGtLLlSkIReEYpDRmMLtFsvFpSPkVSDtVvEpyNatLsvHqLveNAdECmVLDNeAlYdICfRtLkLTPsFGdLNhLIATmSGVt : 177
AtTUB5 : VVrKEAENCdCLQGfQVChSlGGGTGSgMGtLLlSkIReEYpDRmMLtFsvFpSPkVSDtVvEpyNatLsvHqLveNAdECmVLDNeAlYdICfRtLkLTPsFGdLNhLIATmSGVt : 238

```

```

40          *          260         *          280         *          300         *          320         *          340         *
OsTUB2 : CcLrFpGqLNSdLrKlAVNlIppPrlHffMvGfApLtsRgSqQyYrAlTvPeLtQqMwDkNmMCAADpRHGRyLTASaMFRGkMStKEVDEqMlNvQnKNSsYfVewIpnNvKStVcDI : 356
AtTUB8 : CcLrFpGqLNSdLrKlAVNlIppPrlHffMvGfApLtsRgSqQyYrAlTvPeLtQqMwDkNmMCAADpRHGRyLTASaMFRGkMStKEVDEqMlNvQnKNSsYfVewIpnNvKStVcDI : 356
AtTUB3 : CcLrFpGqLNSdLrKlAVNlIppPrlHffMvGfApLtsRgSqQyYrAlTvPeLtQqMwDkNmMCAADpRHGRyLTASaMFRGkMStKEVDEqMlNvQnKNSsYfVewIpnNvKStVcDI : 356
AtTUB2 : CcLrFpGqLNSdLrKlAVNlIppPrlHffMvGfApLtsRgSqQyYrAlTvPeLtQqMwDkNmMCAADpRHGRyLTASaMFRGkMStKEVDEqMlNvQnKNSsYfVewIpnNvKStVcDI : 356
OsTUB4 : CcLrFpGqLNSdLrKlAVNlIppPrlHffMvGfApLtsRgSqQyYrAlTvPeLtQqMwDkNmMCAADpRHGRyLTASaMFRGkMStKEVDEqMlNvQnKNSsYfVewIpnNvKStVcDI : 356
OsTUB1 : CcLrFpGqLNSdLrKlAVNlIppPrlHffMvGfApLtsRgSqQyYrAlTvPeLtQqMwDkNmMCAADpRHGRyLTASaMFRGkMStKEVDEqMlNvQnKNSsYfVewIpnNvKStVcDI : 356
AtTUB6 : CcLrFpGqLNSdLrKlAVNlIppPrlHffMvGfApLtsRgSqQyYrAlTvPeLtQqMwDkNmMCAADpRHGRyLTASaMFRGkMStKEVDEqMlNvQnKNSsYfVewIpnNvKStVcDI : 356
AtTUB7 : CcLrFpGqLNSdLrKlAVNlIppPrlHffMvGfApLtsRgSqQyYrAlTvPeLtQqMwDkNmMCAADpRHGRyLTASaMFRGkMStKEVDEqMlNvQnKNSsYfVewIpnNvKStVcDI : 356
ZmTUB5 : CcLrFpGqLNSdLrKlAVNlIppPrlHffMvGfApLtsRgSqQyYrAlTvPeLtQqMwDkNmMCAADpRHGRyLTASaMFRGkMStKEVDEqMlNvQnKNSsYfVewIpnNvKStVcDI : 356
ZmTUB6 : CcLrFpGqLNSdLrKlAVNlIppPrlHffMvGfApLtsRgSqQyYrAlTvPeLtQqMwDkNmMCAADpRHGRyLTASaMFRGkMStKEVDEqMlNvQnKNSsYfVewIpnNvKStVcDI : 356
ZmTUB1 : CcLrFpGqLNSdLrKlAVNlIppPrlHffMvGfApLtsRgSqQyYrAlTvPeLtQqMwDkNmMCAADpRHGRyLTASaMFRGkMStKEVDEqMlNvQnKNSsYfVewIpnNvKStVcDI : 296
AtTUB5 : CcLrFpGqLNSdLrKlAVNlIppPrlHffMvGfApLtsRgSqQyYrAlTvPeLtQqMwDkNmMCAADpRHGRyLTASaMFRGkMStKEVDEqMlNvQnKNSsYfVewIpnNvKStVcDI : 357
CcLrFpGqLNSdLrKlAVNlIppPrlHffMvGfApLtsRgSqQyYrAlTvPeLtQqMwDkNmMCAADpRHGRyLTASaMFRGkMStKEVDEqMlNvQnKNSsYfVewIpnNvKStVcDI

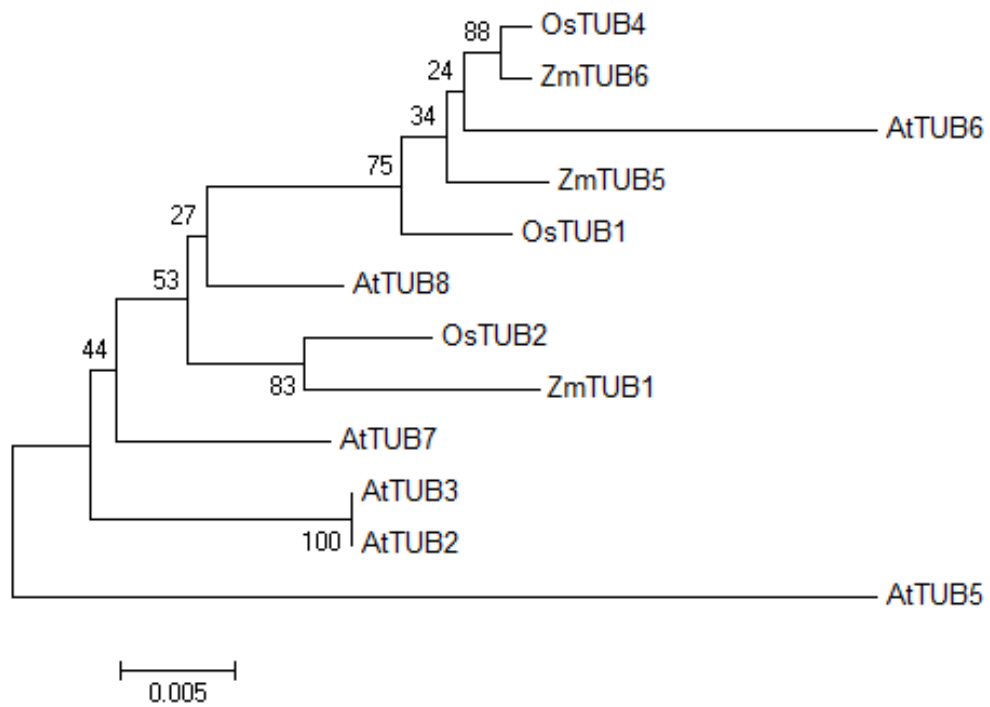
```

```

360          *          380          *          400          *          420          *          440          *
OsTUB2 : PpEgLRMAStFIgNStSIQEMFRVrVSeQFTAmFRrKAFLHWYtGEGMDemEftEaESNMNdLVsEYQqYQdAtAdEgEYdEeEeAdLQD--- : 447
AtTUB8 : PpEgLRMAStFIgNStSIQEMFRVrVSeQFTAmFRrKAFLHWYtGEGMDemEftEaESNMNdLVsEYQqYQdAtAdEgEYdEeEeAdLQD--- : 449
AtTUB3 : PpEgLRMAStFIgNStSIQEMFRVrVSeQFTAmFRrKAFLHWYtGEGMDemEftEaESNMNdLVsEYQqYQdAtAdEgEYdEeEeAdLQD--- : 450
AtTUB2 : PpEgLRMAStFIgNStSIQEMFRVrVSeQFTAmFRrKAFLHWYtGEGMDemEftEaESNMNdLVsEYQqYQdAtAdEgEYdEeEeAdLQD--- : 450
OsTUB4 : PpEgLRMAStFIgNStSIQEMFRVrVSeQFTAmFRrKAFLHWYtGEGMDemEftEaESNMNdLVsEYQqYQdAtAdEgEYdEeEeAdLQD--- : 447
OsTUB1 : PpEgLRMAStFIgNStSIQEMFRVrVSeQFTAmFRrKAFLHWYtGEGMDemEftEaESNMNdLVsEYQqYQdAtAdEgEYdEeEeAdLQD--- : 447
AtTUB6 : PpEgLRMAStFIgNStSIQEMFRVrVSeQFTAmFRrKAFLHWYtGEGMDemEftEaESNMNdLVsEYQqYQdAtAdEgEYdEeEeAdLQD--- : 449
AtTUB7 : PpEgLRMAStFIgNStSIQEMFRVrVSeQFTAmFRrKAFLHWYtGEGMDemEftEaESNMNdLVsEYQqYQdAtAdEgEYdEeEeAdLQD--- : 449
ZmTUB5 : PpEgLRMAStFIgNStSIQEMFRVrVSeQFTAmFRrKAFLHWYtGEGMDemEftEaESNMNdLVsEYQqYQdAtAdEgEYdEeEeAdLQD--- : 445
ZmTUB6 : PpEgLRMAStFIgNStSIQEMFRVrVSeQFTAmFRrKAFLHWYtGEGMDemEftEaESNMNdLVsEYQqYQdAtAdEgEYdEeEeAdLQD--- : 446
ZmTUB1 : PpEgLRMAStFIgNStSIQEMFRVrVSeQFTAmFRrKAFLHWYtGEGMDemEftEaESNMNdLVsEYQqYQdAtAdEgEYdEeEeAdLQD--- : 386
AtTUB5 : PpEgLRMAStFIgNStSIQEMFRVrVSeQFTAmFRrKAFLHWYtGEGMDemEftEaESNMNdLVsEYQqYQdAtAdEgEYdEeEeAdLQD--- : 449
pp GL MaStf6GNStSIQEMFRVrVSeQFTAmFRrKAFLHWYtGEGMDemEftEaESNMNdLVsEYQqYQdAtAdEgEYdEeEeAdLQD--- : 449

```

2.5 用MEGA6.0构建系统进化树



3.1 蛋白序列分析

- 3.1.1 由Uniprot中的Blast检索得到Os TUB2编码的为 β -tubulin蛋白。浏览注释信息，得其登录号。

Entry information¹

Entry name ⁱ	TBB2_ORYSJ
Accession ⁱ	Primary (citable) accession number: Q8H7U1 Secondary accession number(s): Q10T06
Entry history ⁱ	Integrated into UniProtKB/Swiss-Prot: August 30, 2005 Last sequence update: March 1, 2003 Last modified: June 8, 2016 This is version 90 of the entry and version 1 of the sequence. [Complete history]
Entry status ⁱ	Reviewed (UniProtKB/Swiss-Prot)
Annotation program	Plant Protein Annotation Program

1: - - - - - i

3.1.2 功能

微管蛋白是微管的主要成分，分为两种类型，

- (1) α -tubulin, 不具有GTP水解酶活性,
- (2) β -tubulin, 具有GTP水解酶活性.

β -tubulin的分子功能：参与GTP水解；参与GTP结合。构成细胞骨架的主要成分。

生物功能：参与微管的加工组装过程。

3.1.3 命名与分类

- Protein names : **Tubulin beta-2 chain**
- Gene names: **TUBB2**
- Ordered Locus Names: Os03g0105600, LOC_Os03g01530
- Organismⁱ [Oryza sativa subsp. japonica \(Rice\)](#)

3.1.4 亚细胞定位

用TargetP 程序分析 β -tubulin蛋白序列的亚细胞定位情况，Organism Group 选择植物，预测结果如下所示：

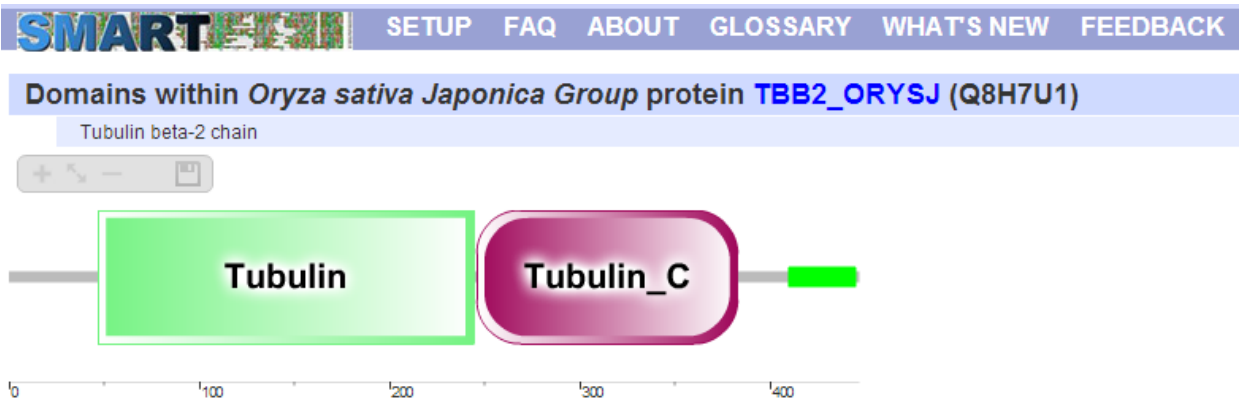
```
### targetp vl.1 prediction results #####
Number of query sequences: 1
Cleavage site predictions not included.
Using PLANT networks.

Name                Len      cTP      mTP      SP      other  Loc  RC
-----
Sequence            447     0.084    0.221    0.099    0.826  _    2
-----
cutoff                0.000    0.000    0.000    0.000

Explain the output. Go back.
```

预测结果显示， β -tubulin蛋白定位在细胞中除叶绿体、线粒体等其他未知部位较高，RC (Reliability class) 值为2（分1-5，1的可靠性最高），表明预测的可靠性较高。其构成细胞质中细胞骨架的主要成分，所以推测应该主要在细胞质中。

3.1.6 TUB2的结构域



Confidently predicted domains, repeats, motifs and features:

Name	Start	End	E-value
Tubulin	47	244	3.61e-67
Tubulin_C	246	383	3.69e-41
coiled coil	410	444	N/A

Click on a row to highlight the feature in the diagram above. Click the feature name for more information.

Features NOT shown in the diagram: ?

Name	Start ▲	End	E-value	Reason
low complexity	138	151	N/A	overlap
low complexity	427	444	N/A	overlap

Click on a row to highlight the feature in the diagram above. Click the feature name for more information.

3.1.7 查找motif



MEME discovers novel, **ungapped** motifs (recurring, fixed-length patterns) in your sequences (sample output from sequences). MEME splits variable-length patterns into two or more separate motifs. See this Manual for more information.

Data Submission Form

Perform motif discovery on DNA, RNA or protein datasets.

Select the motif discovery mode
 Normal mode Discriminative mode [?](#)

Select the sequence alphabet
 Use sequences with a standard alphabet or specify a custom alphabet. [?](#)
 DNA, RNA or Protein Custom

Input the primary sequences
 Enter sequences in which you want to find motifs. [?](#)
 [?](#)

Select the site distribution
 How do you expect motif sites to be distributed in sequences? [?](#)

Select the number of motifs
 How many motifs should MEME find? [?](#)

DISCOVERED MOTIFS

	Logo	E-value ?	Sites ?	Width ?	More ?	Submit/Download ?
1.		1.4e-027	2	50	I	→
2.		5.1e-027	2	50	I	→
3.		1.9e-024	2	50	I	→

Stopped because requested number of motifs (3) found.

MOTIF LOCATIONS

Only Motif Sites [?](#) Motif Sites+Scanned Sites [?](#) All Sequences [?](#)

Name ?	p-value ?	Motif Location ?
1. sp Q8H7U1 TBB2_ORYSJ	1.47e-181	
2. sp P18025 TBB1_MAIZE	2.44e-182	

3.2 组成蛋白的氨基酸理化性质分析

3.2.1 从Rice Genome Annotation 中得到TUB2蛋白的序列长度及预测分子量等。

Gene Attributes	
Chromosome:	Chr3
CDS Coordinates (5'-3'):	339986 - 341796
Nucleotide length:	1344
Predicted protein length:	448
Predicted molecular weight:	50182.4
Predicted pI:	4.4575

3.2.2 利用ExPASy网站提供ProtParam的氨基酸组成分析程序，统计TUB2微管蛋白20种不同氨基酸的组成

ProtParam

User-provided sequence:

```

10      20      30      40      50      60
MREILHIQGG QCGNQIGAKF WEVWCAEHGI DATGRYDGD S DLQLERVNVY YNEASCGRFV

70      80      90      100     110     120
PRAVLMDLEP GTMDSVRS GP YGHIFRPDNF VFGQSGAGNN WAKGHYTEGA ELIDAVLDVV

130     140     150     160     170     180
RKEAENCDC L QGFQVCHSLG GGTGSGMGTL LISKIREEYP DRMMLTFSVF PSPKVS DTVV

190     200     210     220     230     240
EPYNATLSVH QLVENADECM VLDNEALYDI CFRILKLTP SFGDLNHLIS ATMSGVTCCL

250     260     270     280     290     300
RFPGQLNSDL RKLAVNLI PF PRLHFFMVGF APLTSRGSQQ YRALTVPELT QQMWDAKNMM

310     320     330     340     350     360
CAADPRHGRY LTASAMFRGK MSTKEVDEQM LNVQNKNS SY FVEWIPNNVK STVCDIPTG

370     380     390     400     410     420
LKMASTFIGN STSIQEMFRR VSEQFTAMFR RKAFLHWYTG EGMDEMEFTE AESNMNDLVS

430     440
EYQQYQDATA DDEGEYEDEE EADLQD
    
```

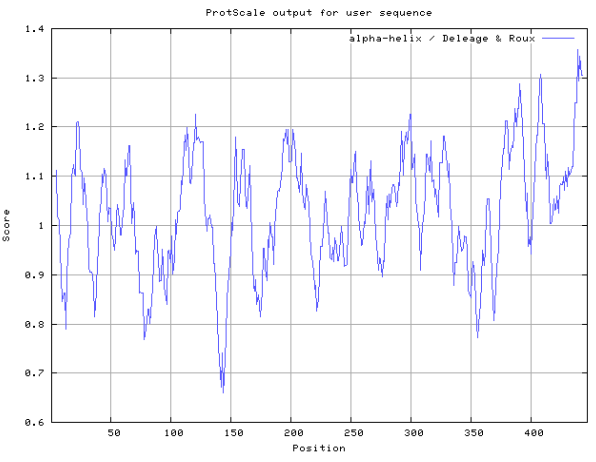
Amino acid composition: [CSV format](#)

Ala (A)	29	6.5%
Arg (R)	23	5.1%
Asn (N)	22	4.9%
Asp (D)	29	6.5%
Cys (C)	12	2.7%
Gln (Q)	21	4.7%
Glu (E)	35	7.8%
Gly (G)	35	7.8%
His (H)	10	2.2%
Ile (I)	15	3.4%
Leu (L)	35	7.8%
Lys (K)	14	3.1%
Met (M)	21	4.7%
Phe (F)	23	5.1%
Pro (P)	18	4.0%
Ser (S)	29	6.5%
Thr (T)	26	5.8%
Trp (W)	5	1.1%
Tyr (Y)	15	3.4%
Val (V)	30	6.7%
Pyl (O)	0	0.0%
Sec (U)	0	0.0%
(B)	0	0.0%
(Z)	0	0.0%
(X)	0	0.0%

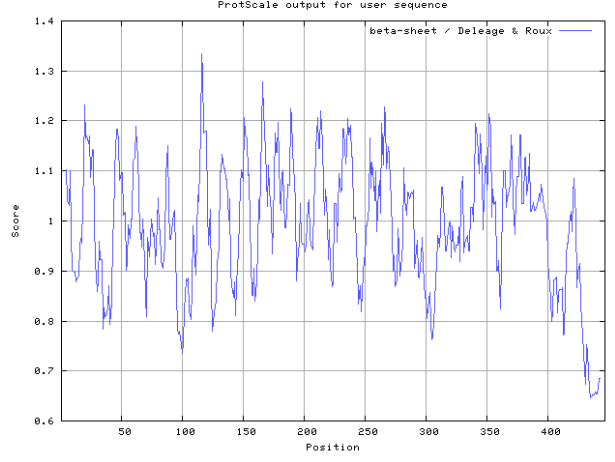
Total number of negatively charged residues (Asp + Glu): 64
 Total number of positively charged residues (Arg + Lys): 37

3.2.3 二级结构

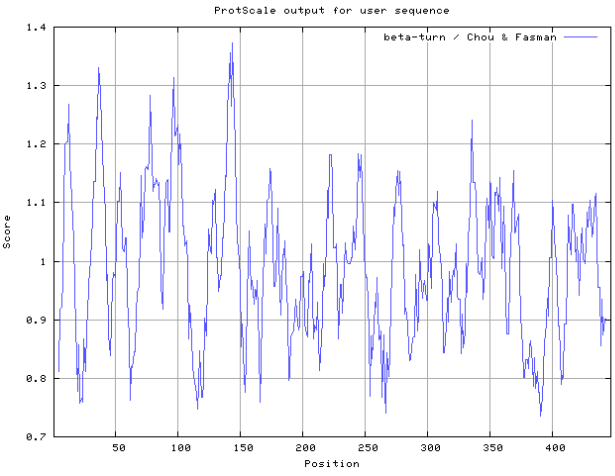
α-Helix (alpha-helix / Deleage & Roux)



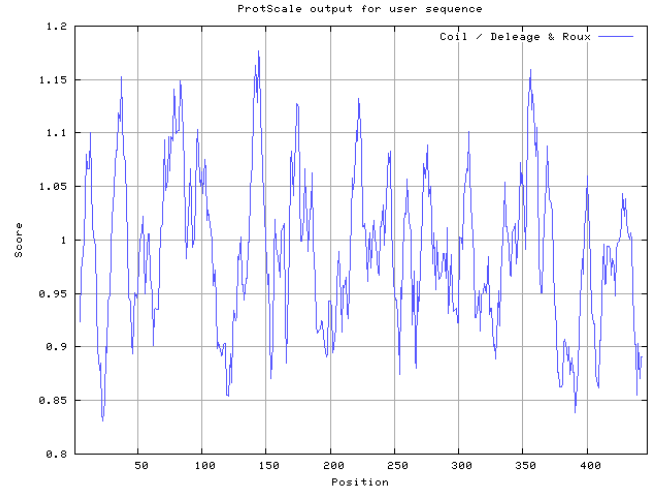
beta-sheet / Deleage & Roux



beta-turn / Chou & Fasman

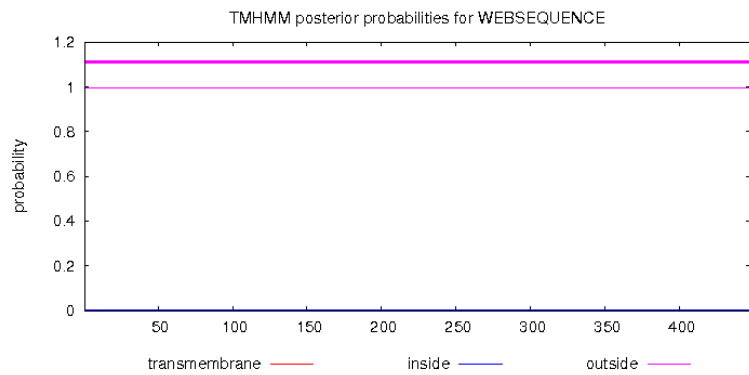


Coil (Coil / Deleage & Roux) :



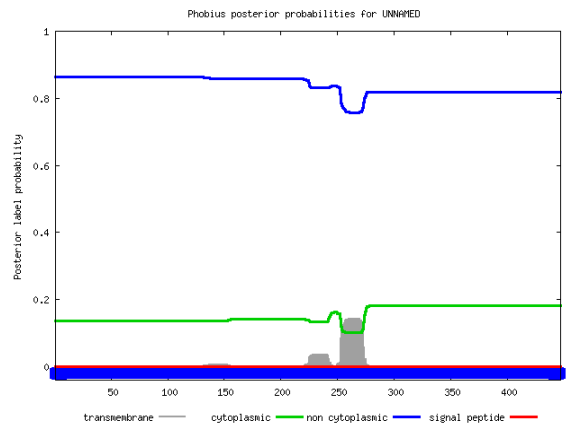
3.2.4 跨膜螺旋预测

- 用**TMHMM**对 *TUB2* 进行分析，其预测的跨膜螺旋如下图所示：



TMHMM 预测没有跨膜螺旋

- 用**Phobius**程序对 *TUB2* 进行分析，其预测的跨膜螺旋如下图所示：



- 与TMHMM 预测结果基本相符。

- 用DAS-TMfilter 对 TUB2进行分析，如下表所示：

```
Calculating prediction for the following proteins  
with reference library 08:
```

```
>
```

```
... Done.
```

```
*** List of predicted non-TM-protein codes ***
```

```
>
```

```
*** List of predicted TM-protein codes ***
```

```
none
```

```
=== Result of the prediction ===
```

```
>
```

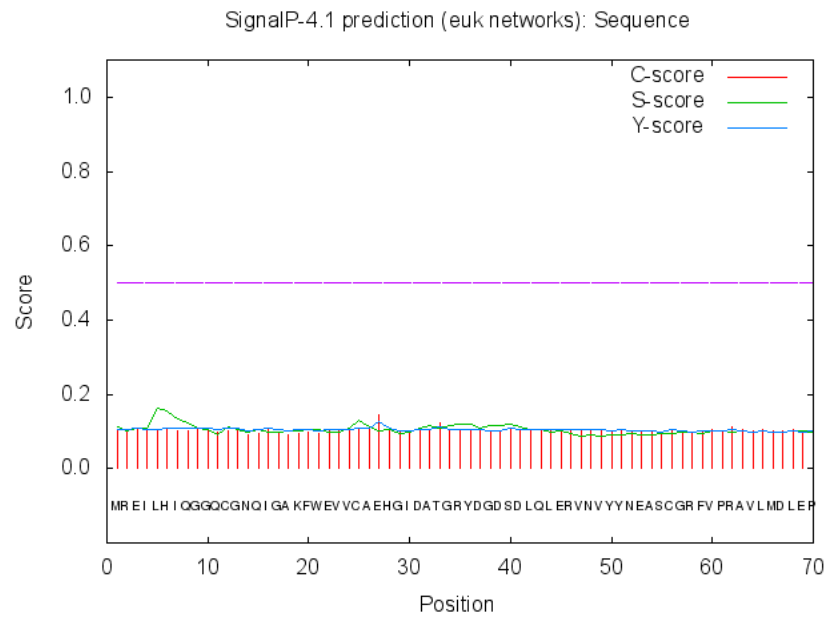
```
# TMH: 0 Q: trusted !!! Warning! Non-TM protein!
```

结果为无跨膜蛋白

- **总结：**各跨膜螺旋预测程序所得结果大同小异， β -tubbulin蛋白无跨膜蛋白，在各程序中TMHMM 和Phobius 的显示方式较为直观，DAS-TMfilter 给出了螺旋的具体位置和预测得分，比较清晰明了。

3.2.5 信号肽预测

利用 **CBS** 网站提供的程序**SignalP**

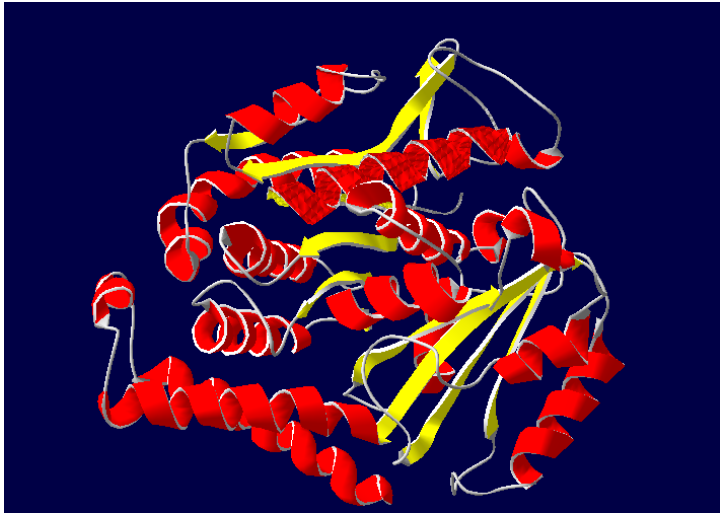


#	Measure	Position	Value	Cutoff	signal peptide?
	max. C	27	0.146		
	max. Y	27	0.128		
	max. S	5	0.162		
	mean S	1-26	0.111		
	D	1-26	0.119	0.450	NO

预测结果表明 *TUB2* 存在信号肽段的可能性很小。

3.3 结构预测

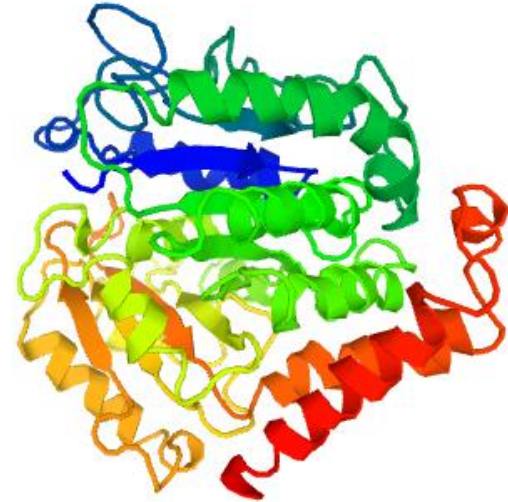
(1) SWISS MODEL 推测其 *TUB2* 的3D结构



Sequence Identity 84.27%

15个 α 螺旋和11个 β 折叠。

5jco. 1. F



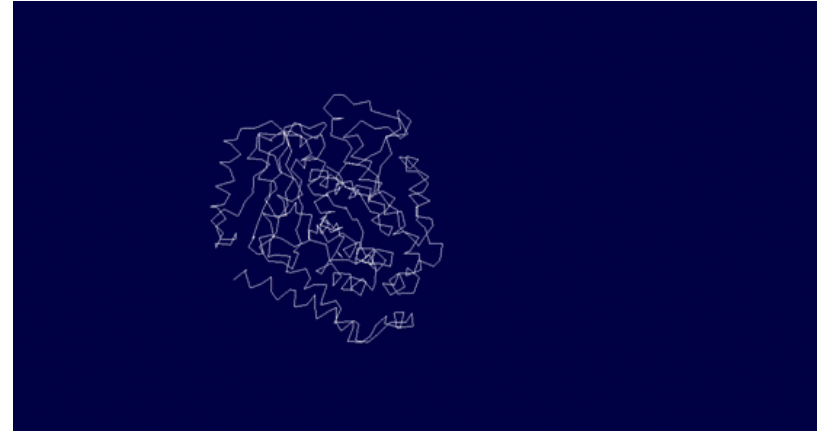
Sequence identity: 87%

13个 α 螺旋, 9个 β 折叠。

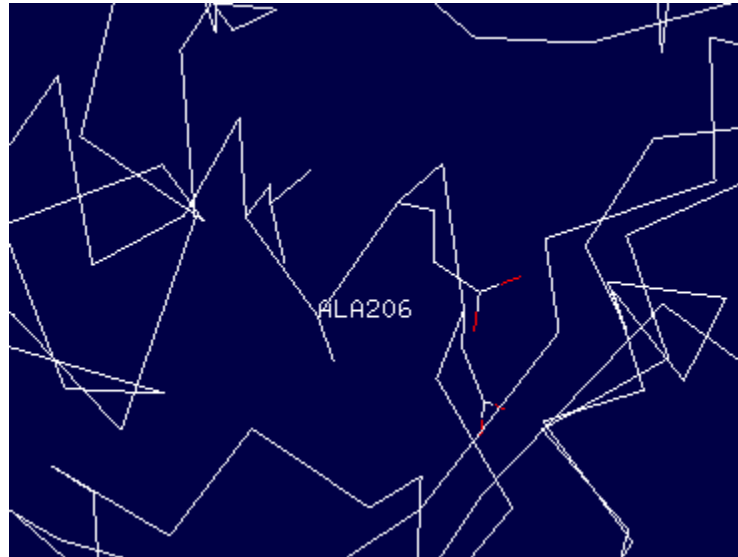
模板(template): 1tubB

[Q8H7U1.](#)

(2) **Phyre2** 预测结构, **Swiss-PdbViewer**观察:



13个 α 螺旋, 9个 β 折叠



4. 未来研究计划

- 通过 *TUB2* 基因的克隆，研究分析 β -tubulin 蛋白的功能和在调控水稻分蘖上的作用机理。
- 经文献阅读和生物信息搜索，发现在 *TUB2* 基因上有与 *OsTB1* 相关的 motif, *OsTB1* 基因负调控水稻分蘖，属于 TCP 家族，*TB1* 是一种最早起源于玉米的转录因子。所以可以通过 EMSA 实验验证其 *OsTUB2* 与 *OsTB1* 在水稻分蘖上的调控关系。



Plant Transcription Factor Database

v3.0

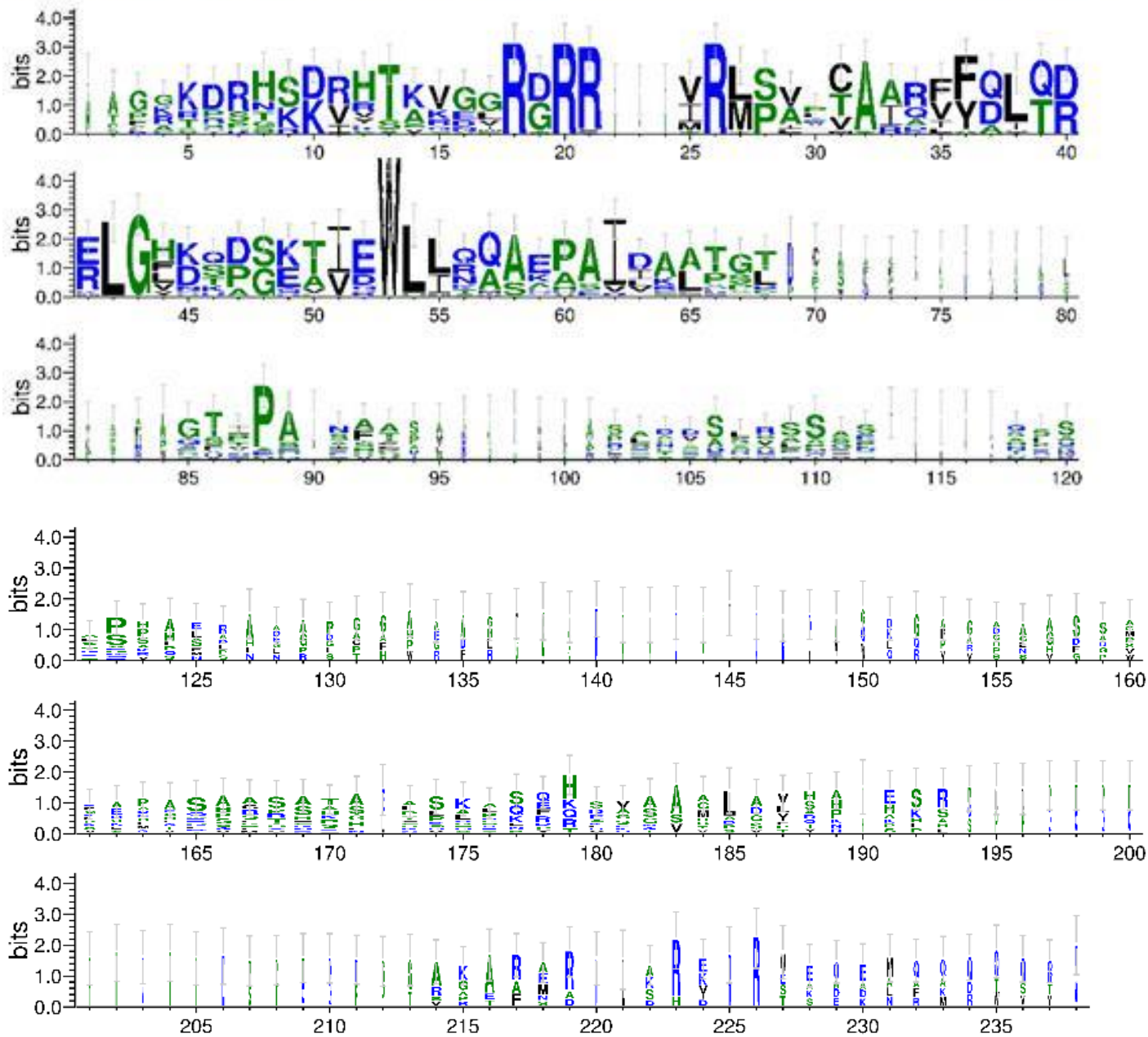
Center for Bioinformatics, Peking University, China

Previous versions: v1.0 v2.0

[Home](#) [Blast](#) [Search](#) [Download](#) [Prediction](#) [Help](#) [About](#) [ATRM](#) [Links](#)

(e.g., LFY)

Multiple Sequence Alignment for *Oryza sativa subsp. japonica* TCP family



谢谢大家！