

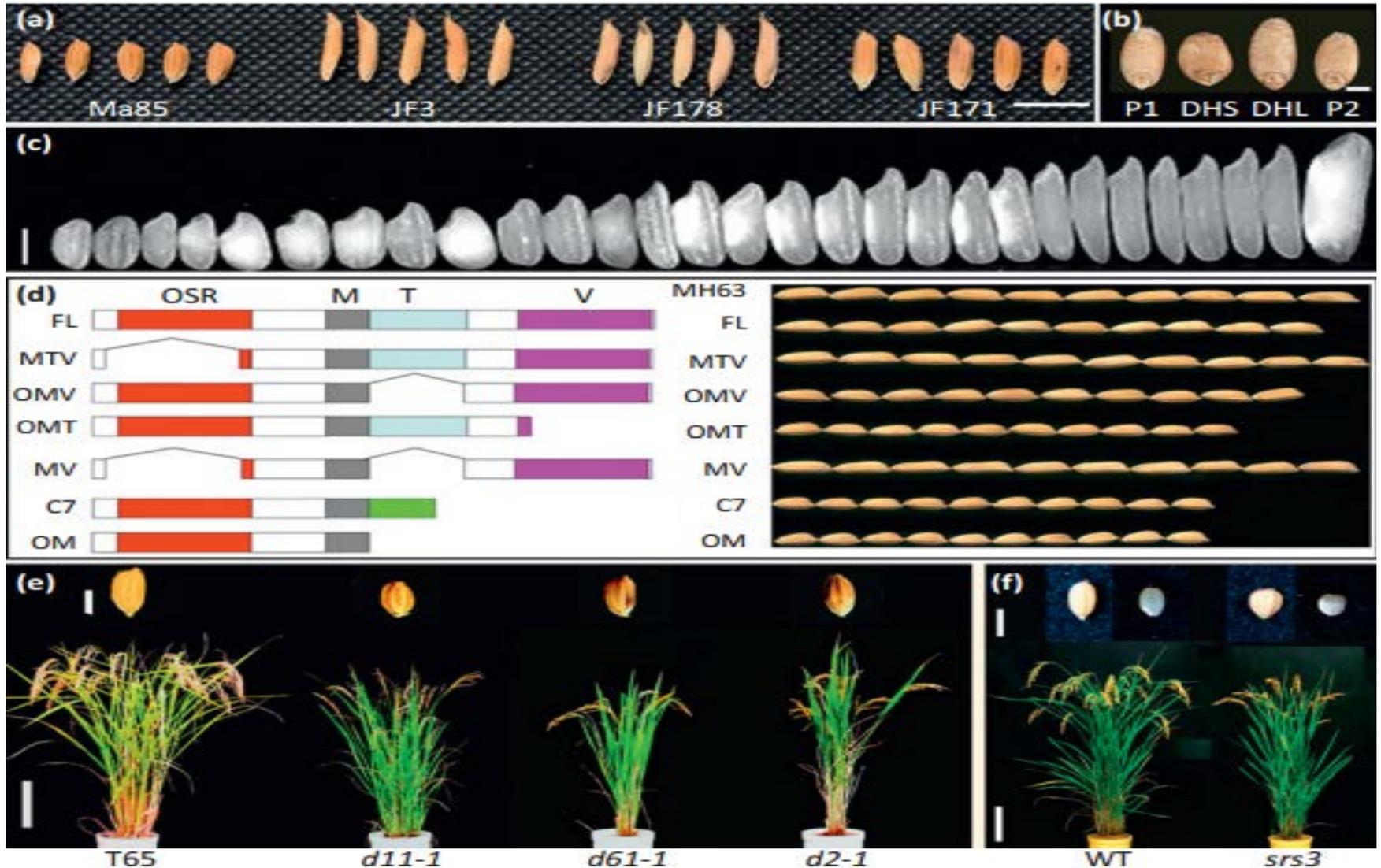
# Cloning Rice Grain Size Gene and Its Utilization

Reporter: Deng hongzhong  
Liu chaolei  
Cui yongtao  
Zhao ge

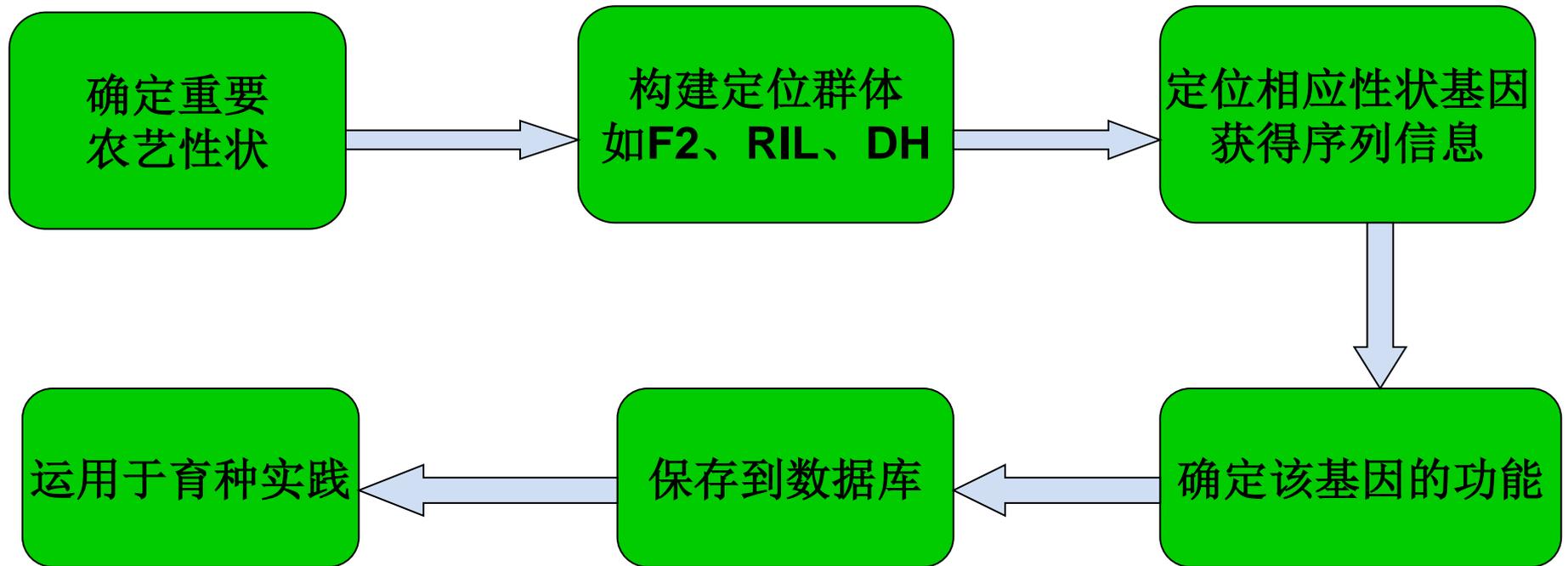
# 粒重——水稻产量的重要构成因素

- 产量构成三要素：每亩穗数、穗粒数和粒重
- 影响粒重因素：长度、宽度和厚度，属于数量性状（QTL）
- QTL: quantitative trait locus, 微效多基因控制的，也存在由少数主效基因控制

# The variations of rice grain shape



# QTL定位一般方法



# 文献搜索

利用NCBI中PubMed高级搜索引擎搜索到73篇相关文献，其中：

16 free full-text articles

9 reviews

5 years: 45

10 years: 68

最感兴趣两篇文章：

Genetic bases of rice grain shape:so many genes, so little known (review)

GS3, a major QTL for grain length and weight and minor QTL for grain width and thickness in rice, encodes a putative transmembrane protein

## 目前研究成果:

# Cloned genes associated with grain shape and weight in rice

Gene	Trait	Gene difference <sup>a</sup>	Encoded protein	Refs
<i>D1</i>	Seed size	833-bp deletion	Heterotrimeric G protein $\alpha$	[21]
<i>D2</i>	Seed size	Nonsense mutation in <i>d2-1</i> ; missense mutation in <i>d2-2</i>	Cytochrome P450 (CYP90D) enzyme	[23]
<i>D11</i>	Seed size	1-bp deletion in <i>d11-1</i> ; 1-bp insertion in <i>d11-2</i> ; missense mutation in <i>d11-3</i> ; abnormal splicing in <i>d11-4</i>	Cytochrome P450 (CYP724B1) enzyme	[22]
<i>D61</i>	Seed size	Missense mutation in <i>d61-1</i> and <i>d61-2</i>	BR insensitive (BRI)-like leucine-rich repeat (LRR) receptor kinase	[25]
<i>GS3</i>	Grain length	Nonsense mutation	Membrane protein with multiple domains	[26,27]
<i>GW2</i>	Grain width	1-bp deletion	RING-type E3 ubiquitin ligase	[28]
<i>GW5/qGW5</i>	Grain width	1212-bp deletion	Arginine-rich protein of 144 amino acids	[29,42]
<i>GIF1</i>	Grain filling	1-bp deletion	Cell wall invertase	[30,43]
<i>GS5</i>	Grain shape	6-bp insertion	Serine carboxypeptidase	[32]
<i>GW8/SPL16</i>	Grain width	10-bp deletion in the promoter	SQUAMOSA promoter-binding protein-like 16	[31]
<i>SRS1/DEP2</i>	Seed size	Deletion in <i>srs1-1</i> and <i>srs1-4</i> ; nonsense mutation in <i>srs1-3</i> ; abnormal splicing in <i>srs1-2</i> and <i>srs1-5</i>	Protein of 1365 amino acids with unknown function	[33]
<i>SRS3</i>	Seed size	Missense mutation in <i>srs3</i> ; nonsense mutation in TCM768; abnormal splicing in TCM2092	Kinesin 13 protein	[34]
<i>SRS5</i>	Seed size	Missense mutation	Alpha-tubulin protein	[24]

<sup>a</sup>Genetic mutations and natural variations.

# GS3: a major QTL for grain length and weight

因为尚未开始实验，没有数据（原始序列信息），所以我们将基因序列已知、功能已鉴定的**GS3**，当作刚定位到的基因结构、功能未知序列，运用生物信息学软件来进行分析。主要是熟悉思路，便于我们以后科研工作。

**GS3: AB488612.1** (Genbank 登录号,原始序列)

# 运用Weblab中coderet软件分析获得CDS及 蛋白序列（编码232AA）

```
>ab488612_cds_1
```

```
atggcaatggcggcggcggccccggcccccaagtgcgcggcggcggccccgacccatgcggc  
cgccaccgcctccagctcgccgctcgaccgcctccaccggcgagatcggattcctcgagggt  
gaaataaattcaatcgaagggatccacgctgcctccagatgctgcagagaggttgaccgaa  
ttcatcggaagaactcctgatccattcataacgatttcacggagaagcgaagtcatgat  
cattctcaccacttcttgaagaagtttcgctgtttgtgcagagcaagtgcgctgctgcctc  
agctacctctcctggatctgctgctgcagcagcggccgggggctgctcctcctcctc  
tctcctccttcaacctcaagaggccgagctgctgctgcaactgcaactgcaactgctgc  
tctcctcctcctcctcctcatgtggggcggcggttaacgaagagtcggtgctgctgcgcgc  
cgcagctgctgctgcgctgctgctgctgcggcggcgtcggcgtccgcgcgtgcgcgagc  
tgcagctgctccccgcgctgcgcgtgctgcgcgcgcgcgctgcgcgggatgctcgtgccgc  
tgcacctgcccgtgcccgtgccccggcggtgctcctgcgcgtgccccggcgtgcaggtgc  
tctgcccggcgtccctcgttgctgccccccctgcttgtga
```

```
>ab488612_pro_1
```

```
MAMAAAPFPKSPFAPDFCGRHRLQLAVDALHREIGFLEGEINSIEGIHAASRCCREYDE  
FIGRTPDFITISSEKRSHDHSHFLLKFRCLCRASACCLSYLSWICCCSSAAGGCSSSS  
SSSFNLKRPSCCCNCCNCCSSSSSSSCGAALTKSPCRCRRSCCCRRCOCGGVGRACAS  
CSCSPPCACCAPPCAGCSCRCTCPCPCPGGCSCACFACRCCCGVPRCCPPCL
```

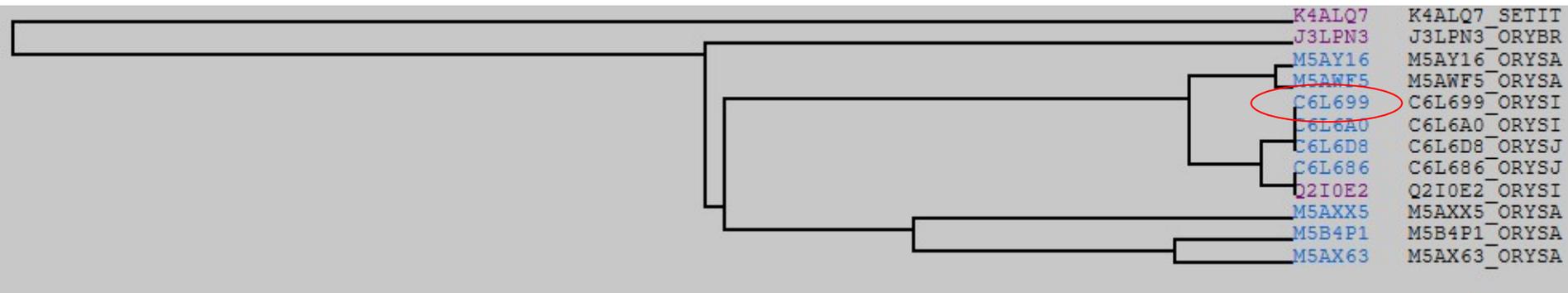
# 运用Blast进行序列比对

```
1 -----MAAAAAPRPKSPPAAPDPGCRHRL 24 J3LPN3 J3LPN3_ORYBR
1 -----MAMAAAPRPKSPPAPPDPGCRHRL 24 M5B4P1 M5B4P1_ORYSA
1 -----MAMAAAPRPKSPPAPPDPGCRHRL 24 M5AY16 M5AY16_ORYSA
1 -----MAMAAAPRPKSPPAPPDPGCRHRL 24 M5AWF5 M5AWF5_ORYSA
1 -----MAAAPRPKSPPAPPDPGCRHRL 22 C6L699 C6L699_ORYSI
1 -----MAMAAAPRPKSPPAPPDPGCRHRL 24 C6L6A0 C6L6A0_ORYSI
1 -----MAMAAAPRPKSPPAPPDPGCRHRL 24 C6L6D8 C6L6D8_ORYSJ
1 -----MAMAAAPRPKSPPAPPDPGCRHRL 24 C6L686 C6L686_ORYSJ
1 -----MAMAAAPRPKSPPAPPDPGCRHRL 24 Q2I0E2 Q2I0E2_ORYSI
1 -----MAMAAAPRPKSPPAPPDPGCRHRL 24 M5AXX5 M5AXX5_ORYSA
121 SSTEQLHRQSPAARLYRPAAVSQSPAESRNEKNNIMAAAPAPRPKSPPASPDPCGRHRL 180 K4ALQ7 K4ALQ7_SETIT
1 -----MAMAAAPRPKSPPAPPDPGCRHRL 24 M5AX63 M5AX63_ORYSA
*****

25 HLAVDALHREIGFLEGEINSIEGIHAASRCCREVDEFVGRTPDPFFITISSEKRNHSHH 84 J3LPN3 J3LPN3_ORYBR
25 QLAVDALHREIGFLEGEINSIEGIHAASRCCREVDEFIGRTPDPFFITISSEKRSHDHSHH 84 M5B4P1 M5B4P1_ORYSA
25 QLAVDALHREIGFLEGEINSIEGIHAASRCCREVDEFIGRTPDPCITISSEKRSHDHSHH 84 M5AY16 M5AY16_ORYSA
25 QLAVDALHREIGFLEGEINSIEGIHAASRCCREVDEFIGRTPDPFFITISSEKRSHDHSHH 84 M5AWF5 M5AWF5_ORYSA
23 QLAVDALHREIGFLEGEINSIEGIHAASRCCREVDEFIGRTPDPFFITISSEKRSHDHSHH 82 C6L699 C6L699_ORYSI
25 QLAVDALHREIGFLEGEINSIEGIHAASRCCREVDEFIGRTPDPFFITISSEKRSHDHSHH 84 C6L6A0 C6L6A0_ORYSI
25 QLAVDALHREIGFLEGEINSIEGIHAASRCCREVDEFIGRTPDPFFITISSEKRSHDHSHH 84 C6L6D8 C6L6D8_ORYSJ
25 QLAVDALHREIGFLEGEINSIEGIHAASRCCREVDEFIGRTPDPFFITISSEKRSHDHSHH 84 C6L686 C6L686_ORYSJ
25 QLAVDALHREIGFLEGEINSIEGIHAASRCCREVDEFIGRTPDPFFITISSEKRSHDHSHH 84 Q2I0E2 Q2I0E2_ORYSI
25 QLAVDALHREIGFLEGEINSIEGIHAASRCCREVDEFIGRTPDPFFITISSEKRSHDHSHH 84 M5AXX5 M5AXX5_ORYSA
181 QLAVDALHREIGFLEGEISSIDGVHAASRCCKEVDEFVGRNPDFFITIQPEKRSNEQSQQ 240 K4ALQ7 K4ALQ7_SETIT
25 QLAVDALHREIGFLEGEINSIEGIHAASRCCREVDEFIGRTPDPFFITISSEKRSHDHSHH 84 M5AX63 M5AX63_ORYSA
:***** **.*:*:*****:*****:*.*** ***. *****.:*.:

137 ----- 136 J3LPN3 J3LPN3_ORYBR
150 ----- 149 M5B4P1 M5B4P1_ORYSA
137 ----- 136 M5AY16 M5AY16_ORYSA
137 ----- 136 M5AWF5 M5AWF5_ORYSA
202 CPCPGGCSCACPACRCCCGVPRCCPPCL 229 C6L699 C6L699_ORYSI
204 CPCPGGCSCACPACRCCCGVPRCCPPCL 231 C6L6A0 C6L6A0_ORYSI
204 CPCPGGCSCACPACRCCCGVPRCCPPCL 231 C6L6D8 C6L6D8_ORYSJ
205 CPCPGGCSCACPACRCCCGVPRCCPPCL 232 C6L686 C6L686_ORYSJ
205 CPCPGGCSCACPACRCCCGVPRCCPPCL 232 Q2I0E2 Q2I0E2_ORYSI
150 ----- 149 M5AXX5 M5AXX5_ORYSA
246 ----- 245 K4ALQ7 K4ALQ7_SETIT
147 ----- 146 M5AX63 M5AX63_ORYSA
```

# 运用Blast进行序列比对

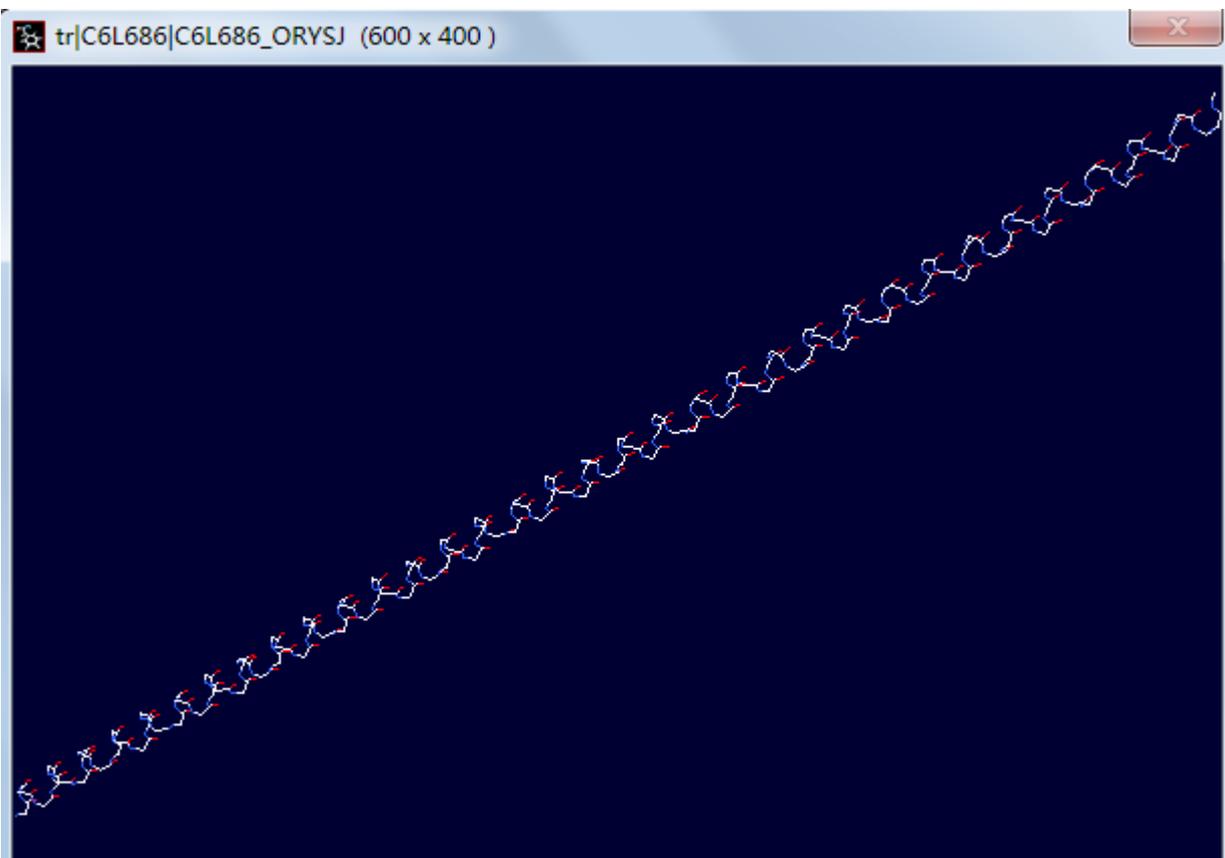


序列比对结果表明：

- 1) 前区段序列几乎完全相同（推测可能为保守序列），后区段序列差异较大；
- 2) 阅读相似性较高的序列发现：所有蛋白亚细胞定位于细胞质中；它们的功能作为信号转换器而存在。如相似度为82%的K4ALQ7，作为信号途经G-蛋白协同受体。

推测：GS3编码蛋白位于细胞质，可能是一种信号转换器

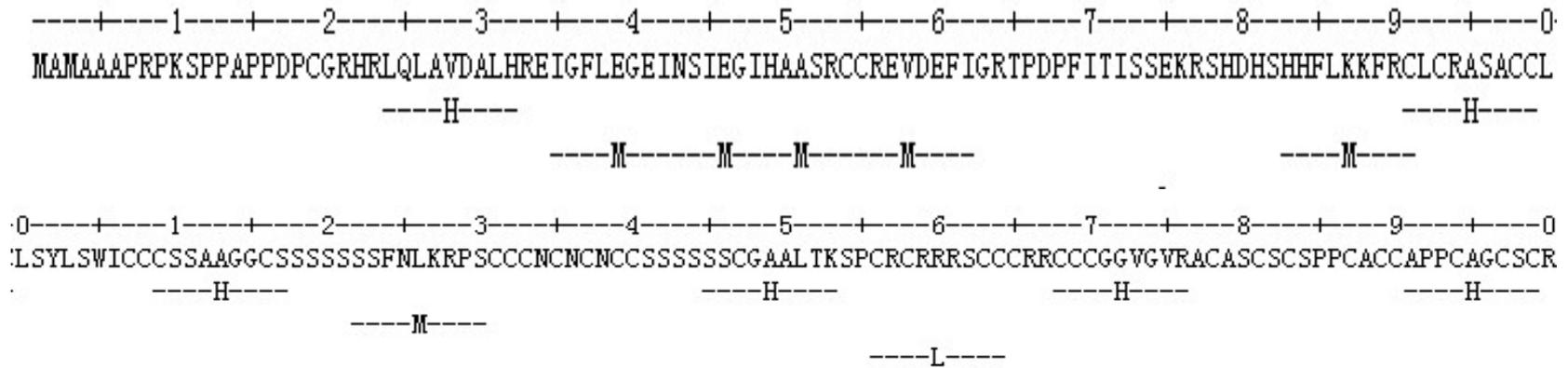
# Swiss-Pdbviewer结果



将蛋白序列导入Swiss-Pdbviewer中发现：**GS3**编码的蛋白结构单一，为无折叠、无回旋的棒状结构

# SOSUlauf Result

sequence length 232 a. a



Hydrophobic auf  
28, 95, 113, 149, 172, 195, 209

- 24-32 LQLAVDALH
- 91-99 CLCRASACC
- 109-117 CSSAAGGCS
- 145-153 SSCGAALTK
- 168-176 CCCGGVGR
- 191-199 APPCAGCSC
- 205-213 CPCPGGCSC

该蛋白多肽链上间隔存在亲水区段；疏水区段主要位于中间偏后位置。预测跨膜区也存在后半段。

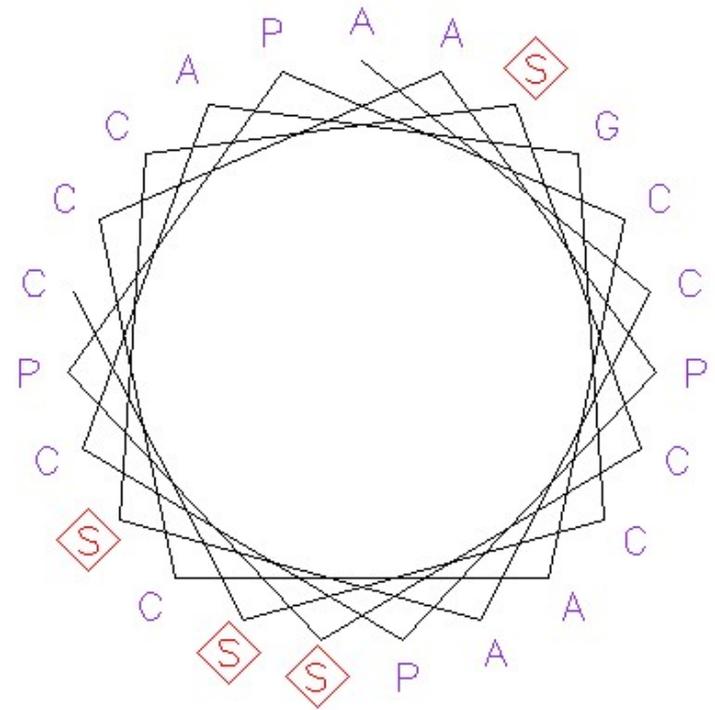
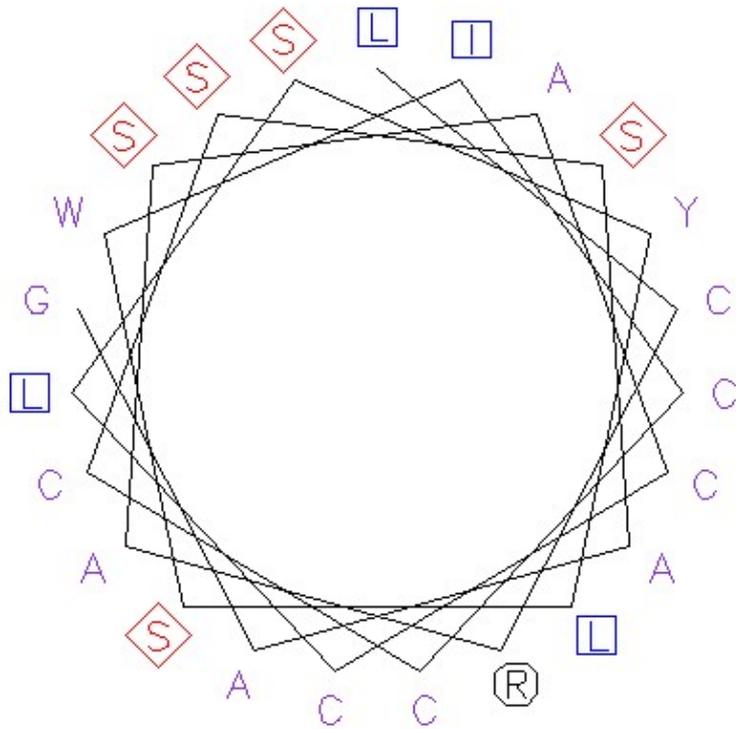
# SOSUI预测跨膜区段

This amino acid sequence is of a MEMBRANE PROTEIN  
which have 2 transmembrane helices.

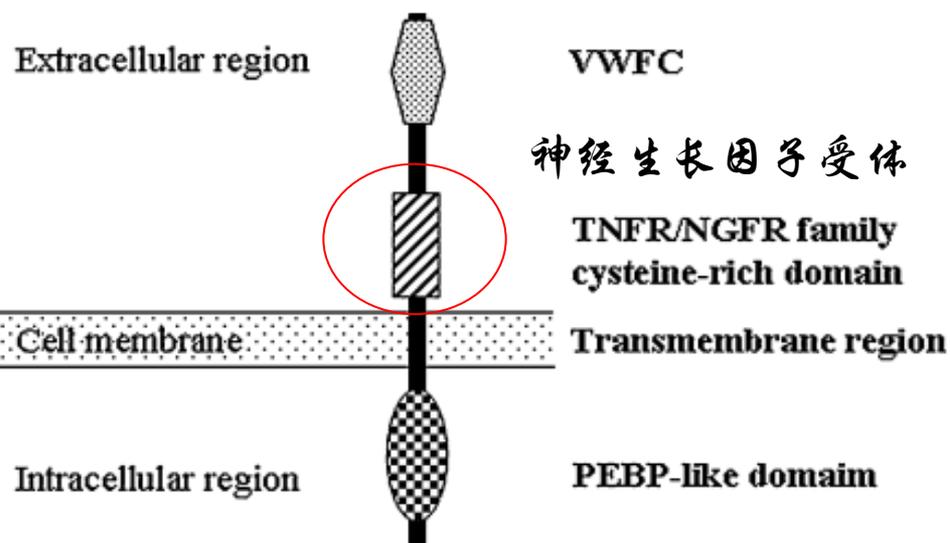
No.	N terminal	transmembrane region	C terminal	type	length
1	92	LCRASACCLSYLSWICCCSSAAG	114	PRIMARY	23
2	177	ACASCSCSPPCACCAPPCAGCSC	199	SECONDARY	23

存在两个跨膜区段，分别为氨基端残基第92-114；  
177-199，长度均为23AA。

# Pepwheel:



# 该蛋白在细胞膜上模型



结论：**GS3**基因控制着水稻粒型性状，其编码长232AA的棒状蛋白。该蛋白位于细胞膜上，可能作为信号转换器参与控制粒型性状。

**Thank you for your attention**