



# 水稻bZIP-5基因生物信息学分析

---

Group 4 : 李聪 何冠华 闫婷 龙海馨

水稻转录因子bZIP-5背景



水稻bZIP-5蛋白质生  
物信息分析

bZIP-5蛋白分析结果

# 一、背景

水稻转录因子bZIP-5主要调控水稻籽粒的长度和宽度

推测一：当bZIP-5与强启动子连接构建载体，并转入到水稻，可以增加转基因水稻籽粒长度、宽度、千粒重，从而增加水稻产量

推测二：将bZIP-5基因敲除，水稻千粒重可能会受到影响，然而目前对水稻转录因子bZIP的研究很少，很多机制还不了解

本文通过对水稻一号染色体上LOC\_0s01g59760进行生物信息学分析，了解该基因的基本性质，以及该基因所编码的蛋白质基本信息。



# 反式作用因子 (trans-acting factor)

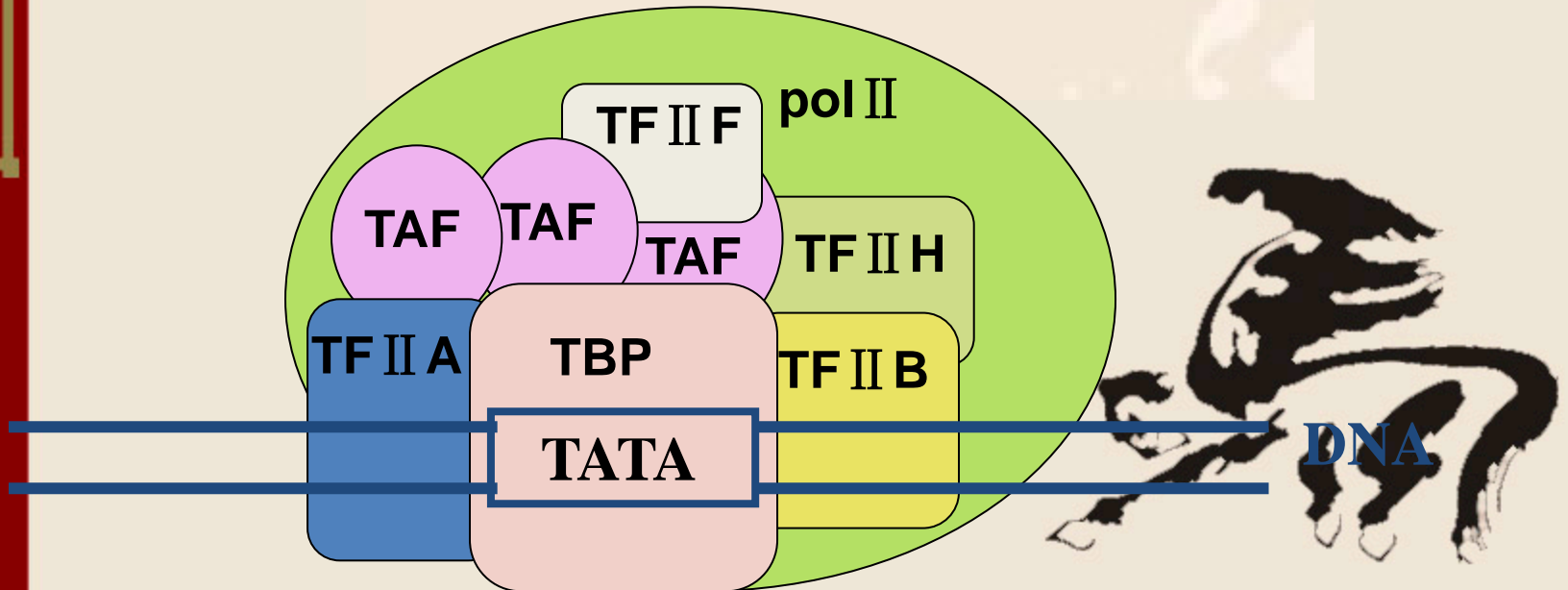
能直接或间接与**顺式作用元件**相互作用，进而调控基因转录的一类调节蛋白，统称为**反式作用因子**。

按其功能不同，常有以下三类：

**基本转录因子**：识别**promoter**元件

**转录调节因子**：识别**enhancer**或**silencer**

**共调节因子**：不能进行**DNA-蛋白质**相互作用



# 常见转录因子的结构域 (domain)

TF

**DNA结合域 (DNA binding domain)**

**Basic AA (K/R) rich, positively charged**

**转录激活域**

**(trans-activation domain)**

**蛋白质-蛋白质结合域**

**(dimerization, co-factors)**

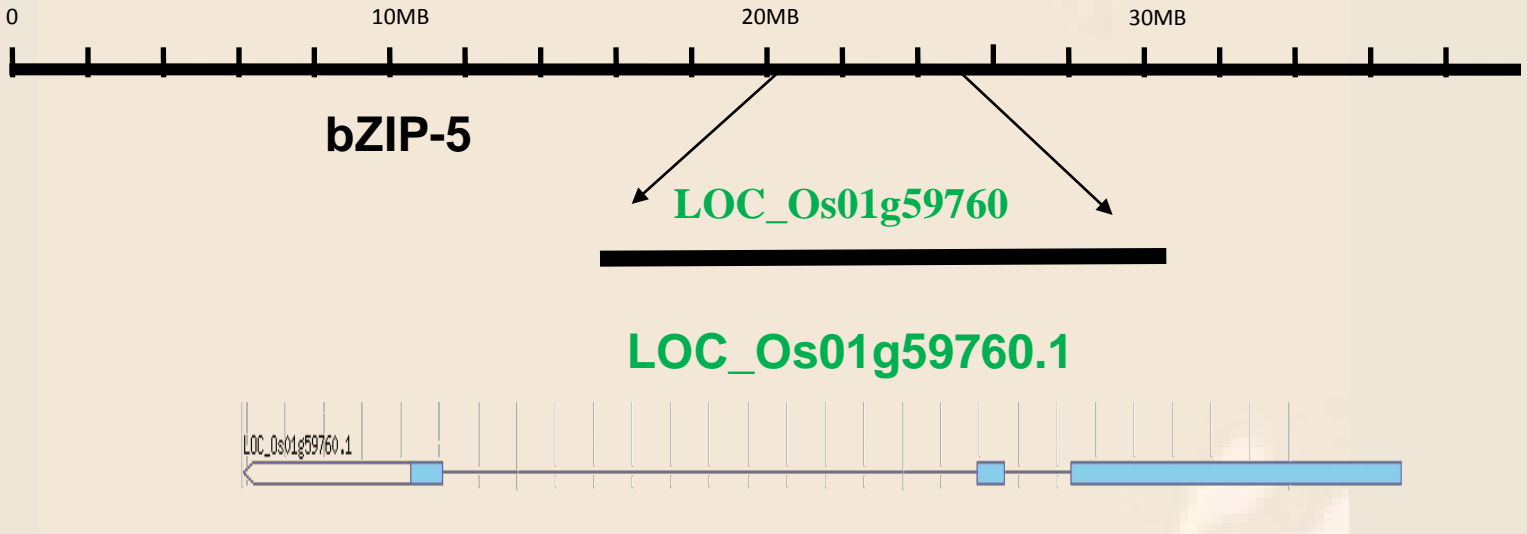
一、  
背景



- **bZIP-5基因性质，重复序列分析**
- **bZIP-5蛋白质性质、疏水性、信号肽、亚细胞定位、磷酸化、结构域预测！**
- **bZIP-5二级结构预测、三级结构预测**



# 水稻bZIP基因登陆号为: LOC\_Os01g59760



LOC\_Os01g59760 is a gene spanning chr1 from 34565292-34568290. LOC\_Os01g59760.1 is a mRNA spanning chr1 from 34565292-34568290



水稻  
bZIP  
基因

水稻  
bZIP  
基因

使用DNAMAN软件对核酸序列组分分析得到：bZIP-5  
核酸单链相对分子质量为313KDa；双链621.52KDa。  
A：23.7%；C：23.9%；G：34.5%；T：17.9%

SEQ DFXL.SEQ: 1008 bp;

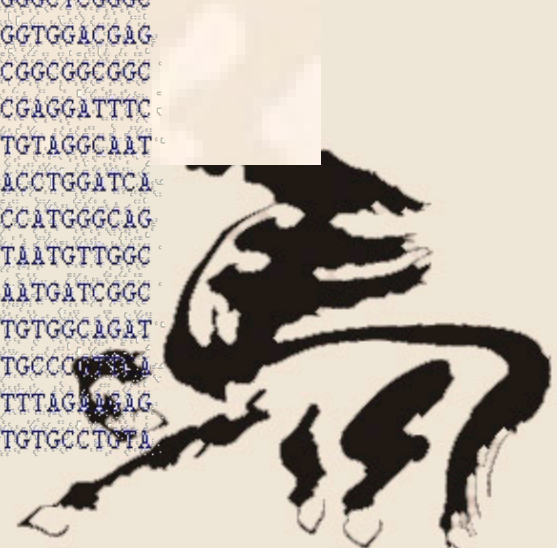
Composition: 239 A; 241 C; 348 G; 180 T; 0 OTHER

Percentage: 23.7% A; 23.9% C; 34.5% G; 17.9% T; 0.0% OTHER

Molecular Weight (kDa): ssDNA: 313.67 dsDNA: 621.52

ORIGIN

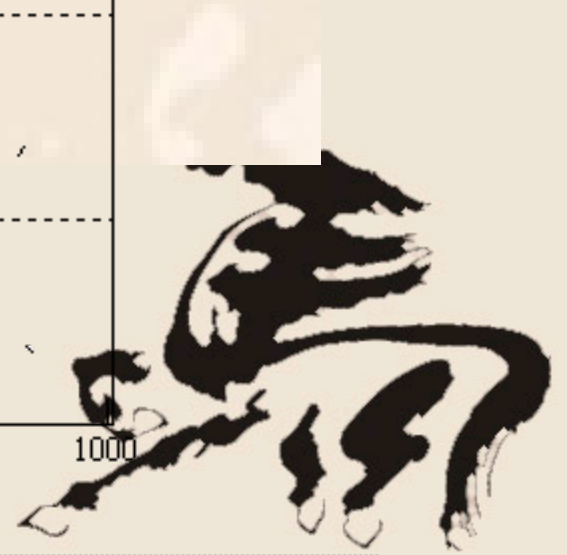
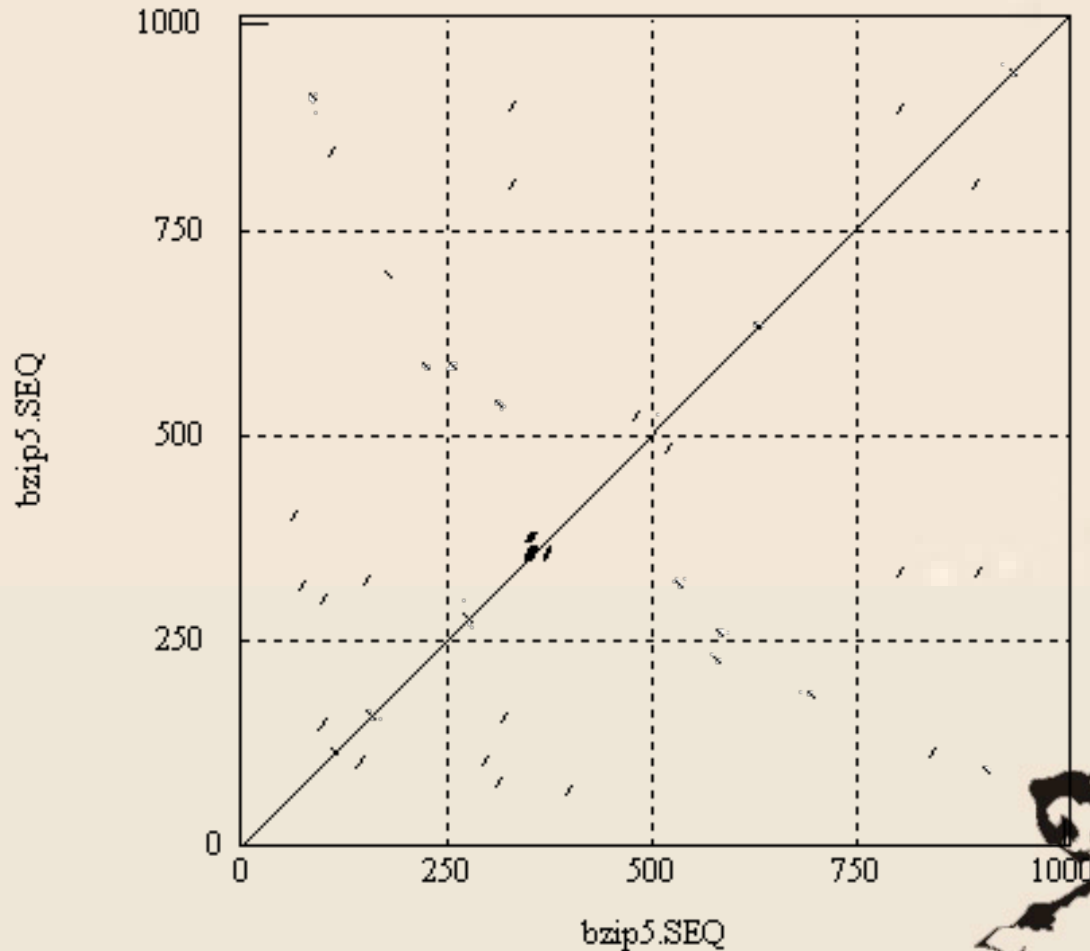
```
1      ATGGCGTCGA AGGCCGGAGG CGGGGGCGTC GCGCGGCGCG GAGGAGGGCG GATGCGGAGT
61     CTGGGGAGGC AGGGATCCAT GTACAGCCTC ACCCTCGACG AGGTGCAGAG CCAGCTGGGC
121    GAGCCGCTGC ACAGCATGAA CCTCGACGAG CTGCTCCGGA GCGTGTTCCT TGACGGCCTG
181    GCCATCGCCG ACGGTGCCGG TGCCACCACC AGTAGCCAGC AGCATCAGCC GGGCTCGGGC
241    CTCTGCGGCC AGGGGAGCAT CACGATGCCG CCTGAGCTCA GCAAGAAGAC GGTGGACGAG
301    GTGTGGAAGG GGATCCAGGC TGCTCCGAAG AGGAATGCCG AAACGGGCGG CGGCGGCGGC
361    GGAGGTCGAC GGCGGCGAGA GAGGCAGCCG ACGCTTGGGG AGGTGACGCT CGAGGATTTC
421    CTGGTCAAAG CTGGGGTTGT CACCCAAGGA TCTCTCAAGG AGCTTAGTGA TGTAGGCAAT
481    GTGGATCCGG TTGGAAGAGG TGTTACAGCA ACCGGGACTG TGGATCTGGC ACCTGGATCA
541    CACTGGATAG AGCAGTATAA GCAGCAGATT GCATCTACTG ATGCTCATCA CCATGGGCAG
601    CAAGGTGTGC AGGGTGCTTA TTTCCGAAT CGATTGGTCC CTCAGCCACT TAATGTTGGC
661    CCCGGTGCTA TTCTGGAGCC ATCCTACTCT GATGGCCAGA CTTCTTCAGG AATGATCGGC
721    GGAATGTCCG ATTCGCAGAC GCCTGGAAGG AAGCGAGGCA TGTCAGGGGA TGTGGCAGAT
781    AAGTTGATGG AGAGAAGGCA GAAGAGGATG ATCAAGAACA GGGAGTCAGC TGCCCGTCAA
841    AGAGCCAGGA AGCAGGCCTA CACTAATGAA CTTGAAAACA AGGTTTCTCG TTTAGGAGAG
901    GAGAATGTGA GGCTAAAGAG GCAAAAAGGAG TTGGATGAAT TAATTTGGCG TGTGCCTGTA
961    CCAGAACCCA AGTATCAACT CAGGAGAACA AGCTCCGCAG ATTTCTGA
```





使用DNAMAN软件对核酸重复序列分析：bZIP-5属于单拷贝序列！

水稻  
bZIP  
基因



水稻  
bZIP  
蛋白质

利用ProtParam工具分析可以得到：bZIP-5氨基酸序列  
The N-terminal of the sequence considered is M (Met).  
The estimated half-life is: 30 hours (mammalian  
reticulocytes, in vitro).  
只有1, 占0.3%, >20 hours (yeast, in vivo).  
58; 脂肪指数为:>10 hours (Escherichia coli, in vivo).  
蛋白质不稳定。

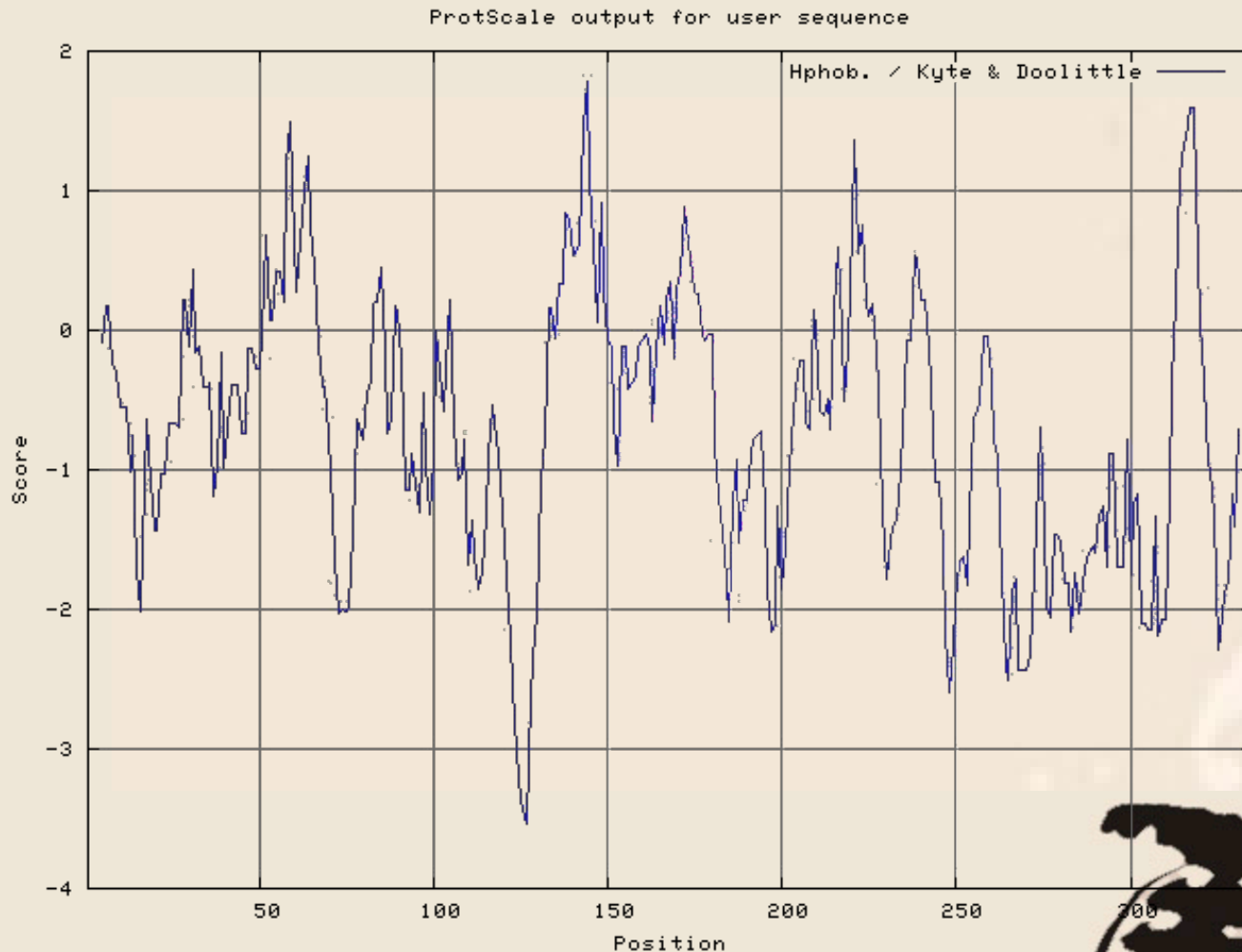
Amino acid composition:

CSV-format

Ala (A)	24	7.2%
Arg (R)	29	8.7%
Asn (N)	9	2.7%
Asp (D)	16	4.8%
Cys (C)	1	0.3%
Gln (Q)	23	6.9%
Glu (E)	22	6.6%
Gly (G)	44	13.1%
His (H)	6	1.8%
Ile (I)	9	2.7%
Leu (L)	28	8.4%
Lys (K)	17	5.1%
Met (M)	10	3.0%
Phe (F)	4	1.2%
Pro (P)	18	5.4%
Ser (S)	28	8.4%
Thr (T)	17	5.1%
Trp (W)	2	0.6%
Tyr (Y)	6	1.8%
Val (V)	22	6.6%
Pyl (O)	0	0.0%
Sec (U)	0	0.0%
(B)	0	0.0%
(Z)	0	0.0%
(X)	0	0.0%



# 水稻 bZIP 蛋白质



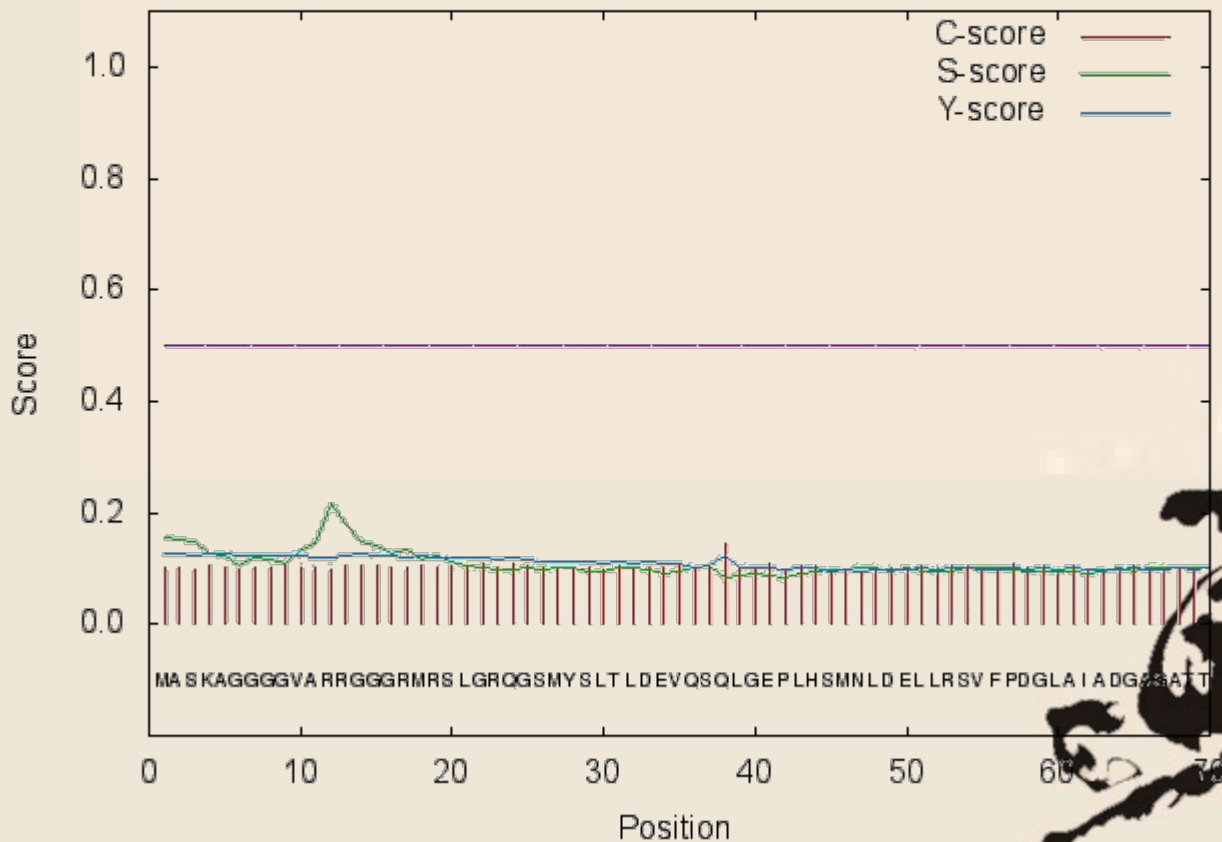
使用protScale疏水性分析：得到bZIP-5为亲水性蛋白，在蛋白质的C末端典型的疏水性区域，推测可能为信号肽！



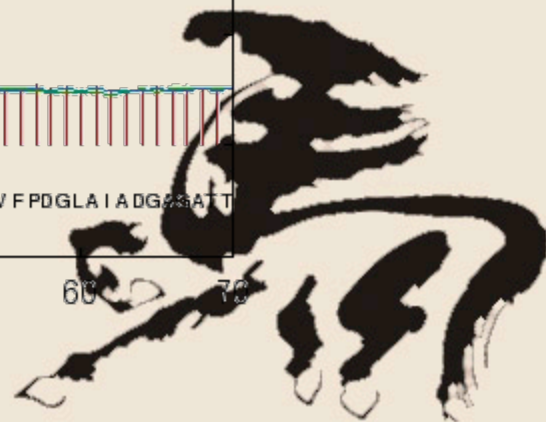
# 使用SignalP预测蛋白信号肽：该蛋白没有信号肽

# SignalP-4.1 euk predictions  
>LOC\_Os01g59760.1

SignalP-4.1 prediction (euk networks): LOC\_Os01g59760.1

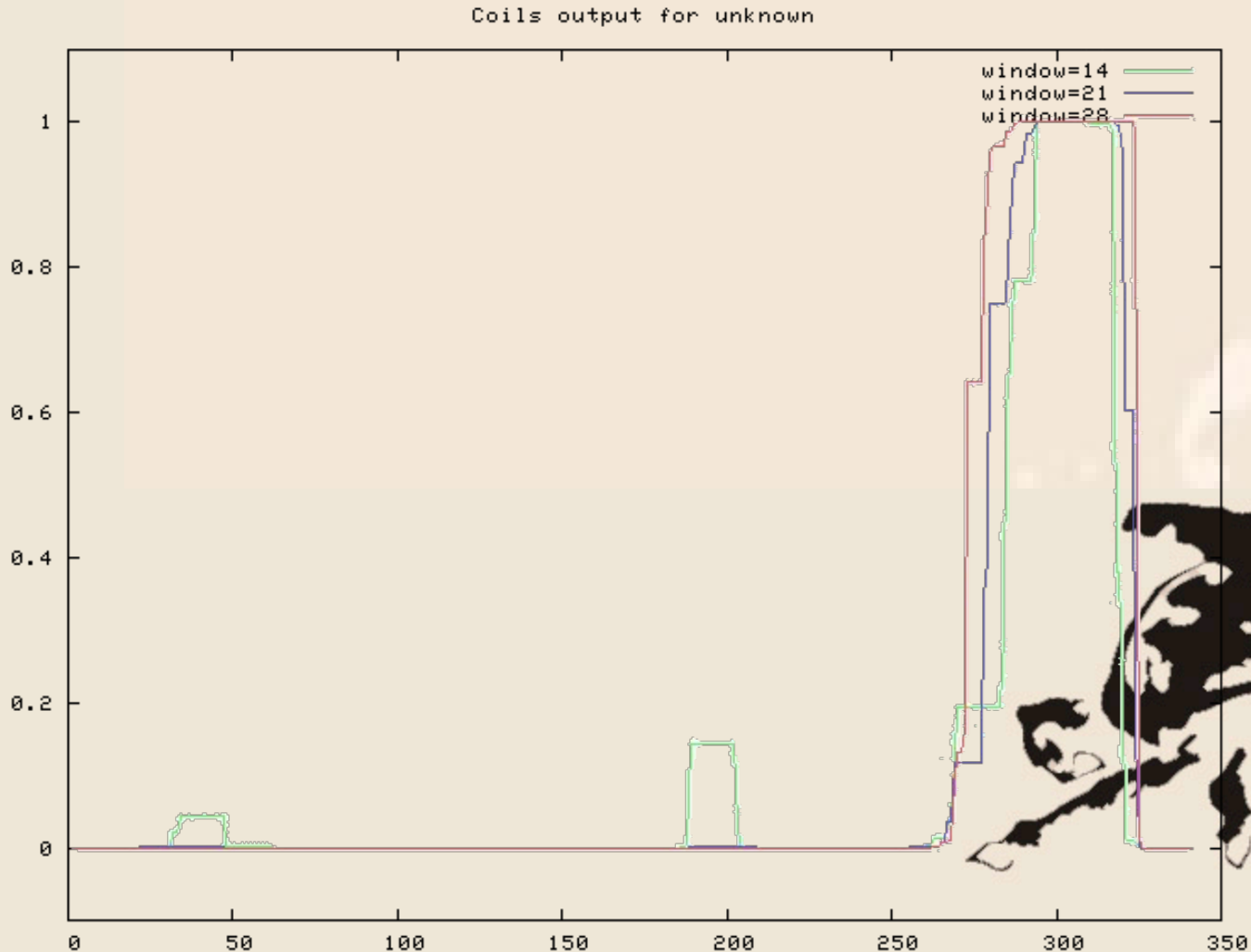


水稻  
bZIP  
蛋白质



使用MITOPROT预测蛋白质线粒体、叶绿体信号肽：该蛋白线粒体信号肽！推测这个蛋白质可能在线粒体调控基因！

水稻  
bZIP  
蛋白质

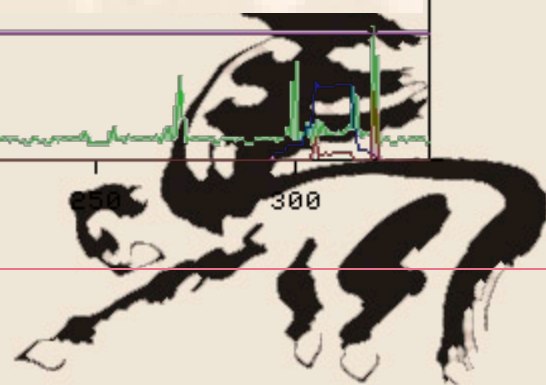
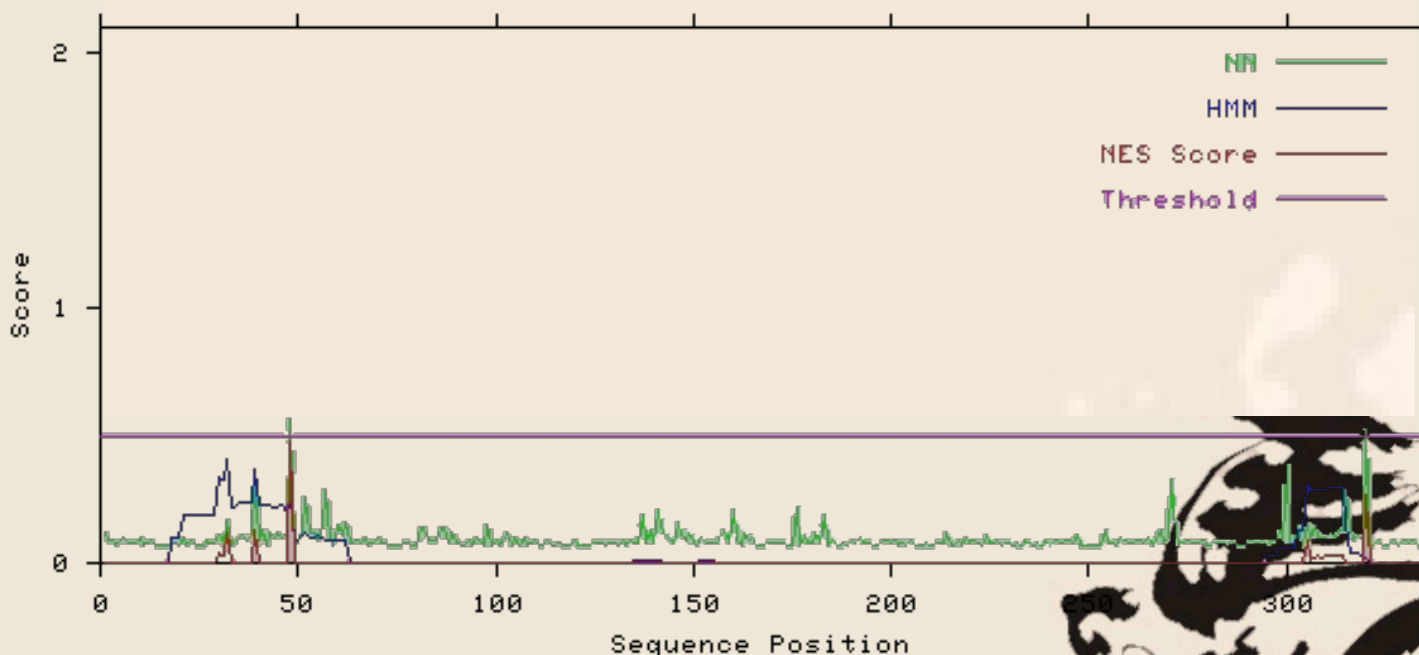


使用NetNES 1.1 Server预测蛋白质核输出信号肽：没有核输出信号！不在细胞核中！

水稻  
bZIP  
蛋白质

>LOC\_Os01g59760.1 - NetNES 1.1 prediction

NetNES 1.1: Predicted NES signals in LOC Os01g59760.1



使用TargetP预测蛋白质亚细胞定位：该蛋白存在线粒体中，与前面的线粒体信号肽预测结果一致！

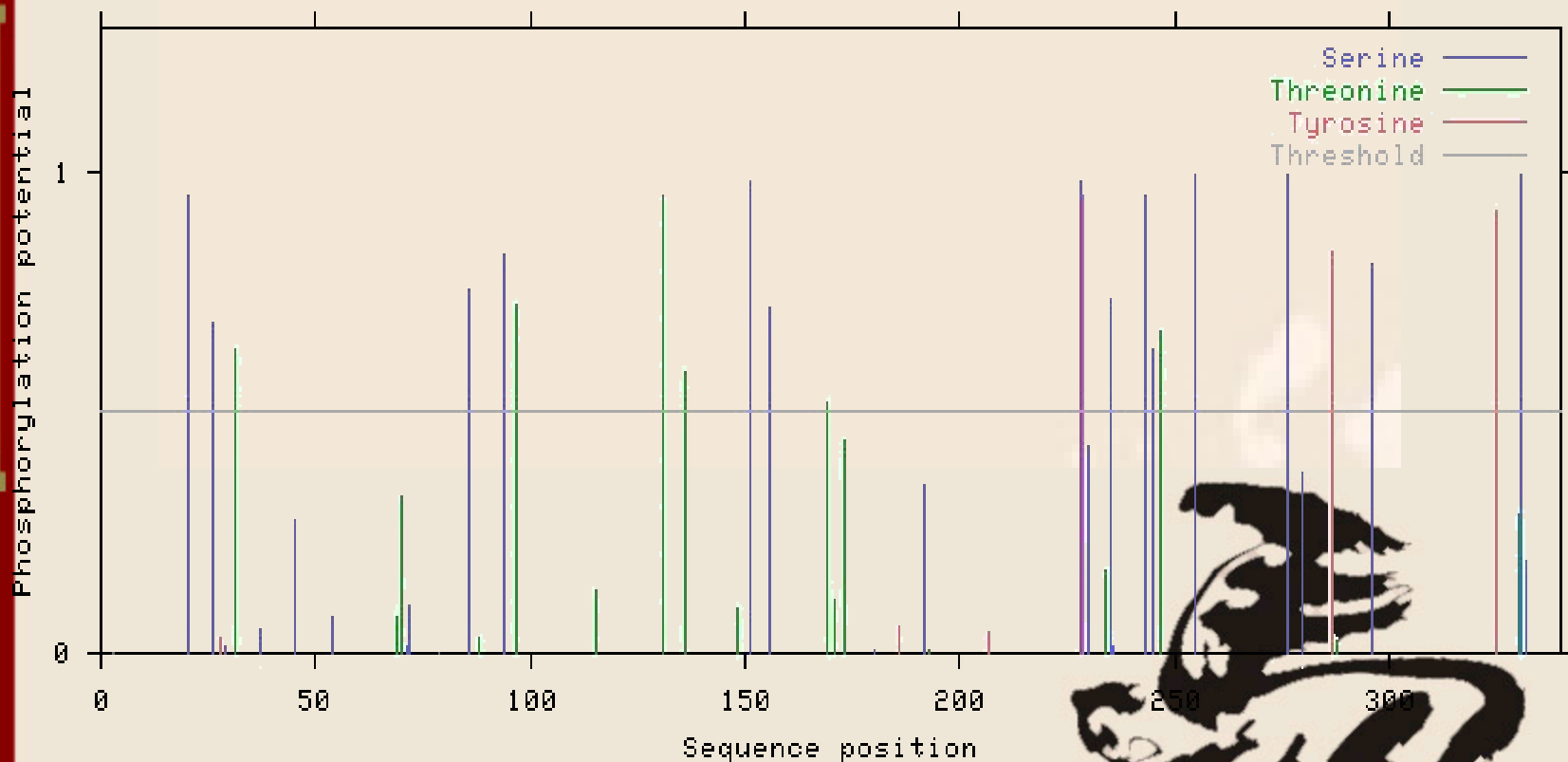
Name	Len	cIP	mIP	SP	other	Loc	RC
LOC_Os01g59760.1	335	0.097	0.691	0.007	0.300	M	4
cutoff		0.000	0.000	0.000	0.000		

PROBABILITY  
of export to mitochondria: 0.6930

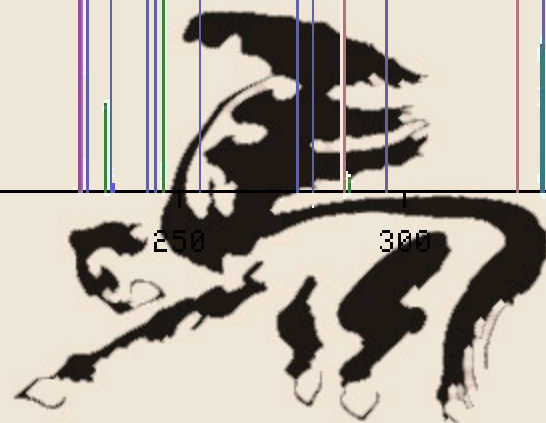


使用NetPhos 2.0 Serve预测蛋白质磷酸化位点  
预测：发现可磷酸化的位点一共有23个：  
Ser:14;Thr:6; Tyr:3

NetPhos 2.0: predicted phosphorylation sites in LOC 0s01g59

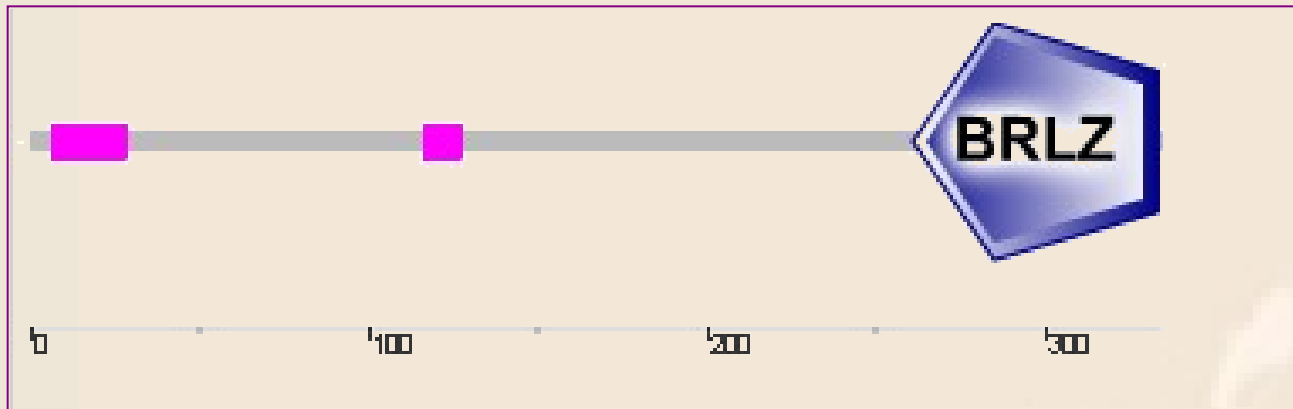


水稻  
bZIP  
蛋白质





使用SMART预测蛋白质结构域：有BRLZ结构域，具有亮氨酸拉链结构，可以与DNA大沟直接作用！

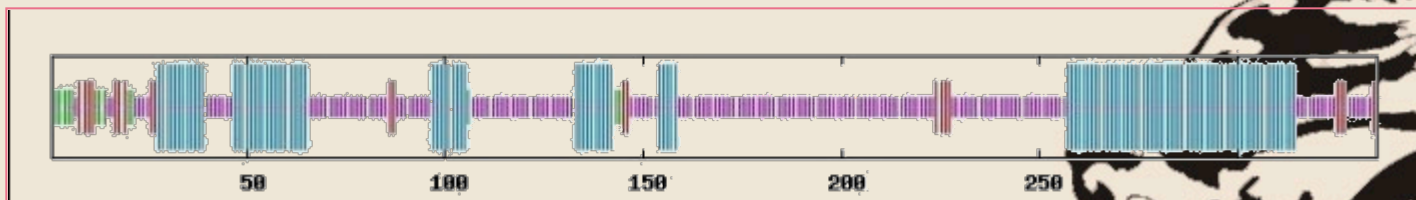


The basic-leucine zipper (bZIP) domain transcription factors of eukaryotic are proteins that contain a basic region mediating sequence-specific DNA-binding followed by a leucine zipper region required for dimerisation. sequence-specific DNA binding, sequence-specific DNA binding transcription factor activity.

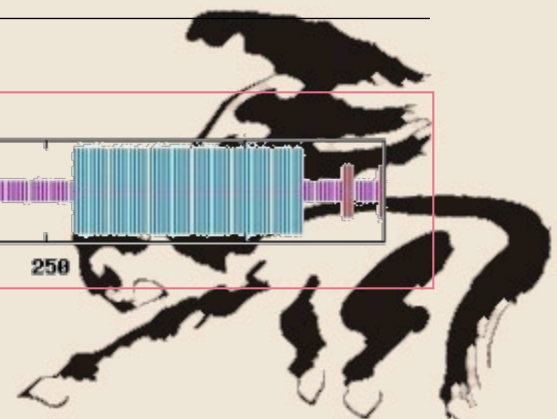


使用SOPMA预测蛋白质二级结构：无规卷曲占55.52%！蛋白氨基酸序列260-310有连续 $\alpha$ -螺旋，推测为亮氨酸拉链结构所在！

二级结构	二级结构简称	个数	百分比
Alpha helix $\alpha$ -螺旋	Hh	254	34.33%
Extended strand伸展链	Ee	54	6.27%
Beta turn $\beta$ -转角	Tt	23	3.88%
Random coil无规卷曲	Cc	54	55.52%

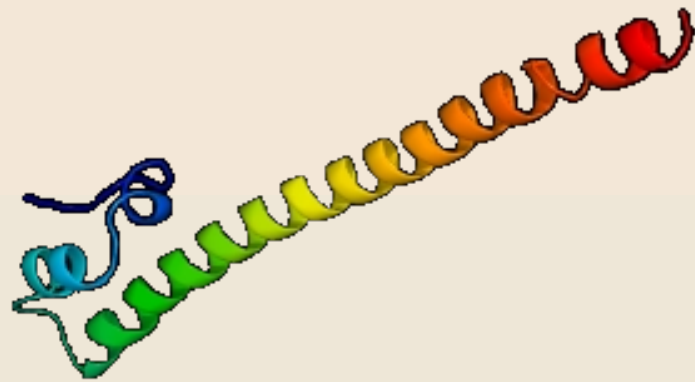


水稻  
bZIP  
蛋白质



使用SWISS-MODEL预测蛋白质三级结构：采用同源建模的方式预测，由于PDB库中只有一段螺旋，因此没有预测出该蛋白质的全部三级结构！

使用PHYRE折叠识别预测蛋白质三级结构：也只是得到其中的一段螺旋！这个蛋白质三级预测需采用从头合成的方法预测！



经过生物信息分析可以得到：

- 1、bZIP-5核酸单链相对分子质量为313KDa；双链621.52KDa.
- 2、bZIP-5基因属于单拷贝序列。
- 3、bZIP-5氨基酸序列个数为335个；分子质量约为36KDa；PI为9.56；氨基酸序列中Gly(G)含量最高，Cys含量最低。蛋白质不稳定。
- 4、bZIP-5为疏水性蛋白，具有线粒体信号肽，位于线粒体中！
- 5、bZIP-5可磷酸化的位点一共有23个：Ser:14;Thr:6; Tyr:3
- 6、bZIP-5具有亮氨酸拉链，可以直接与DNA大沟相互作用，具有转录激活域
- 7、bZIP-5蛋白二级结构无规则卷曲占主要部分！



## bZIP-5生物信息分析不足之处:



- 1、只是分析了其中的部分信息，还有很多没有分析出来！
- 2、使用BLAST同源分析，多序列比对后，没有得到明确的信息！已分析过的的bZIP家族与bZIP-5蛋白质没有特别近的同源关系！
- 3、没有构建出bZIP-5蛋白质三级结构，需从头开始预测！



接下来实验计划:

- 1、构建bZIP-5基因强启动子的载体，并转到烟草，拟南芥，水稻种观察表型！并做转录表达分析
- 2、构建构建bZIP-5 RNAi基因强启动子的载体，并转到烟草，拟南芥，水稻种观察表型！并作转录表达分析
- 3、对bZIP-5蛋白亚细胞定位
- 4、寻找bZIP-5转录因子与DNA结合的片段
- 5、寻找bZIP-5蛋白质与其它蛋白质之间的互作



非常感谢罗老师这段时间的教导  
非常感谢第四组其它组员：龙海馨、闫婷、  
何冠华对这次报告的支持与贡献  
非常感谢这段时间给予第四组帮助的小伙伴们

thank you

