



小麦GS5-3B的蛋白序列 与结构分析

Protein Sequence and Structure analysis of
GS5 3B in Wheat

G08小组成员：刘俊钰、董安琪、赵雅芝、贾博宇

汇报人：董安琪

汇报时间：2026.1.10



目录

CONTENTS

- 1 研究背景与目的
- 2 序列分析
- 3 蛋白质性质分析
- 4 蛋白质结构预测
- 5 研究结论与展望



研究背景与目的

序列分析

蛋白质性质分析

蛋白质结构预测

研究结论与展望

研究背景



国家水稻数据中心
China Rice Data Center



EDGENE BIOT
艾迪晶生物

科研人自己的基因编辑平台
“载体设计+遗传转化+检测分析+种质种植”一体化功能基因研究解决方案

首页 | 优异种质 | 突变体 | 分子标记 | 基因&QTL | ONTOLOGY | 品种&系谱 | 技术&百科 | 文献资源

• 常规信息 最近更新: 2024年11月20日 6:40:00

基因(座)名称	丝氨酸羧肽酶; 籽粒大小 <i>regulator of grain size; serine carboxypeptidase 26; grain size 5</i>
基因符号	<i>GS5; OsSCP26</i>
所在染色体	5 (已克隆)

[GS5\(AK106800, Li et al. 2011; JN256055/JN256056, Li et al. 2016\)...](#)

GS5编码一个丝氨酸羧肽酶, 正向调控水稻籽粒大小。

【突变体表型】
GS5高表达具有更大的籽粒 (Li et al. 2011; Xu et al. 2015)。

【定位与克隆】
GS5位于第5染色体短臂的RM593和RM574之间。

【时空表达谱】
GS5等位基因具有节律性, 在白天转录逐步增加。GS5在绿色组织中的表达量通常高于非绿色组织如茎秆、根、幼穗以及胚乳; 非绿色组织中, GS5最高的表达量主要在长度为1-11cm的幼穗中; 开花阶段, 外稃/内稃中GS5表达量显著上调 (Xu et al. 2015)。

【亚细胞定位】

【生物学功能】
GS5编码一个丝氨酸羧肽酶, 是一个控制水稻粒宽、充实度和千粒重的数量性状基因 (Li et al. 2011)。
GS5启动子区域的两个关键单核苷酸多态性(SNPs)造成水稻幼穗中GS5的差异表达, 这决定了籽粒大小的差异。GS5表达增强会通过占据OsBAK1-7的胞外富亮氨酸重复(LRR)结构域, 竞争性地抑制OsBAK1-7和OsMSBP1间的互作, 进而阻止OsBAK1-7与OsMSBP1互作而产生的胞吞作用 (Xu et al. 2015)。

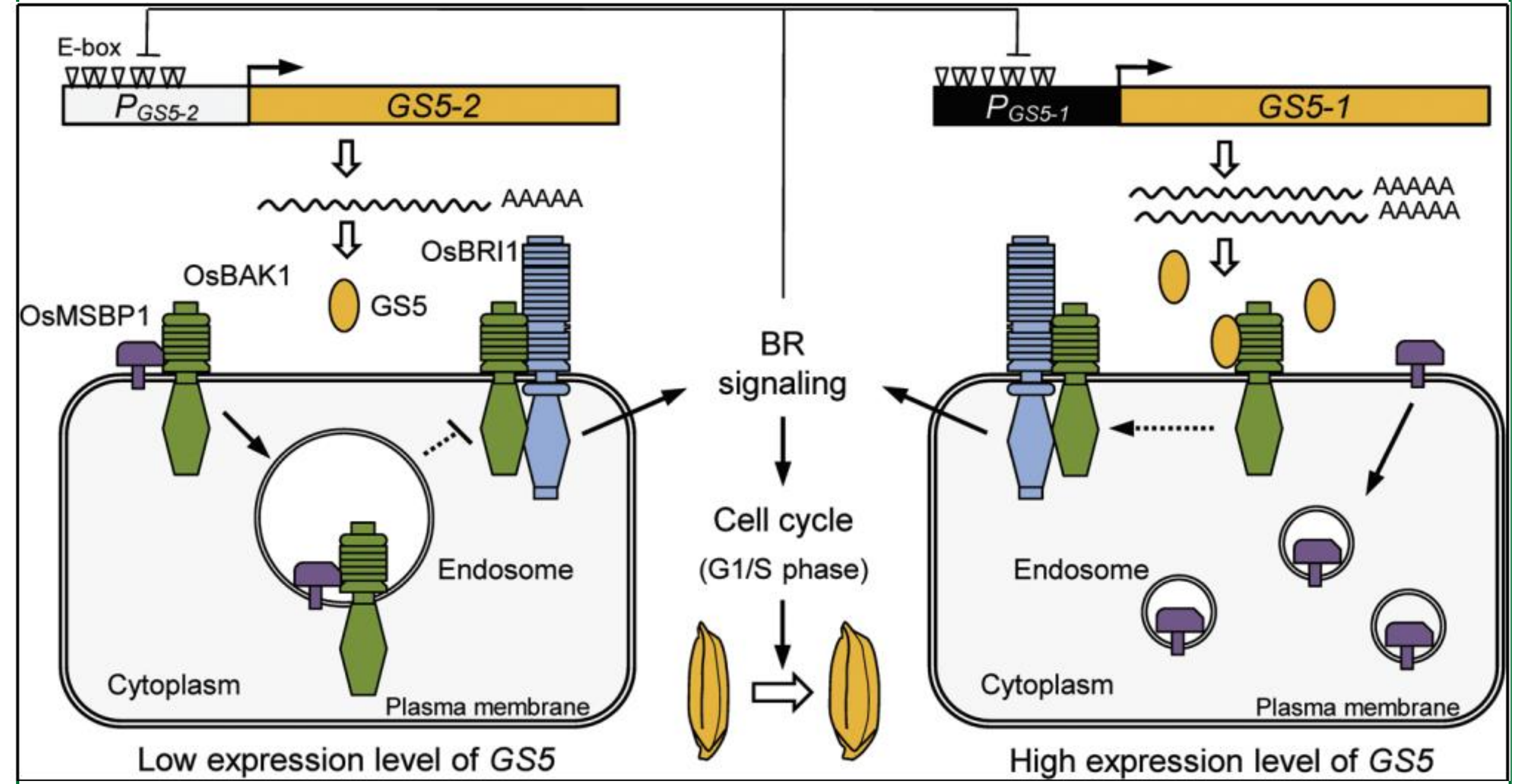
【相关登录号】

contigs及其产物: [AP014961](#) ↘ [BAS92383](#)
基因及产物ID号: [JN256057](#)→[AEO37082](#), [JN256058](#)→[AEO37083](#)
cDNAs及其产物: [JN256055](#)→[AEO37080](#), [JN256056](#)→[AEO37081](#), [AK106800](#)
参考基因组位点: [Os05g0158500 \(RAP-DB, PhytoAB公司抗体服务\)](#) ↔ [LOC_Os05g06660 \(本地、MSU-RGAP, 百格基因突变体服务\)](#) ↔ [LOC4337873 \(NCBI\)](#)



研究背景

OsGS5通过促进细胞分裂调控水稻籽粒大小和重量



研究背景与目的

序列分析

蛋白质性质分析

蛋白质结构预测

研究结论与展望



研究目的

研究背景与目的

序列分析

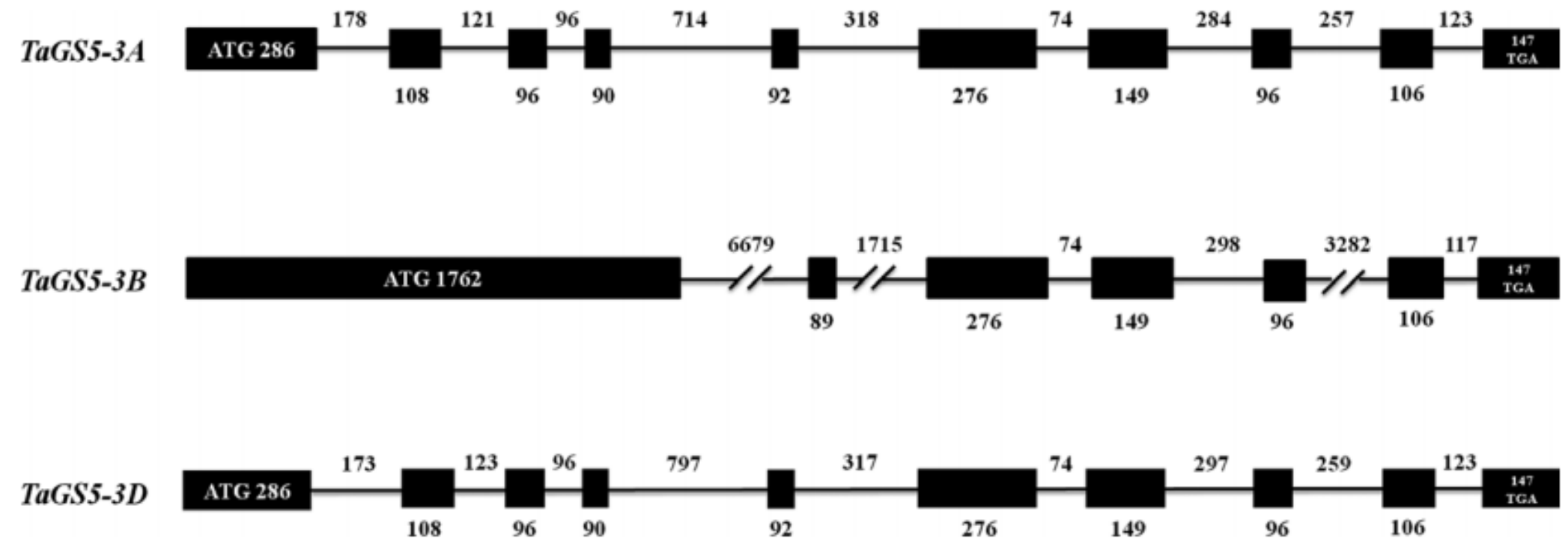
蛋白质性质分析

蛋白质结构预测

研究结论与展望



研究目的



已在小麦中分离出GS5同源基因，TaGS5同源基因优先在幼穗和发育中的籽粒中表达，并将其定位在3A、3B和3D染色体上。有研究表明，TaGS5-3A是籽粒大小的正向调控因子，但目前关于TaGS5-3B的功能研究较少，特此对小麦TaGS5-3B编码的丝氨酸羧肽酶样33的序列和结构进行分析。



研究背景与目的

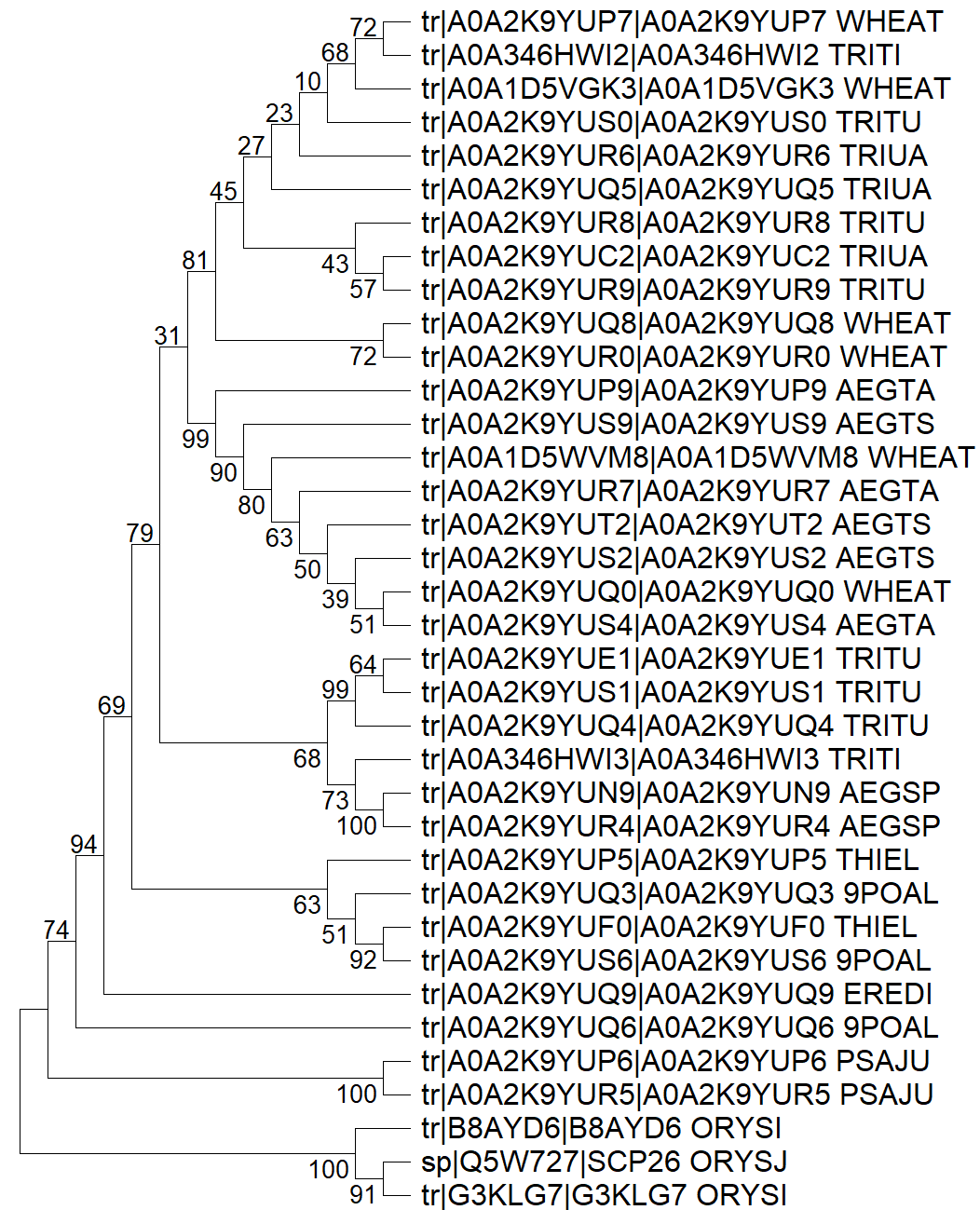
序列分析

蛋白质性质分析

蛋白质结构预测

研究结论与展望

系统发生树构建



MEGA12 邻接法(NJ)构建



UniProt构建系统发育树

**GS5基因在小麦属的进化过程中高度保守；
水稻序列独立成支，与小麦亲缘关系较远**



分析蛋白质的 pI、Mw、氨基酸组成、消光系数、稳定系数等

研究背景与目的

序列分析

蛋白质性质分析

蛋白质结构预测

研究结论与展望

Number of amino acids: 448
Theoretical pI: 8.28
Molecular weight: 50592.11

Amino acid composition:

Ala (A)	32	7.1%
Arg (R)	20	4.5%
Asn (N)	26	5.8%
Asp (D)	25	5.6%
Cys (C)	7	1.6%
Gln (Q)	21	4.7%
Glu (E)	18	4.0%
Gly (G)	37	8.3%
His (H)	8	1.8%
Ile (I)	19	4.2%
Leu (L)	39	8.7%
Lys (K)	26	5.8%
Met (M)	4	0.9%
Phe (F)	24	5.4%
Pro (P)	21	4.7%
Ser (S)	35	7.8%
Thr (T)	15	3.3%
Trp (W)	10	2.2%
Tyr (Y)	30	6.7%
Val (V)	31	6.9%
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): 44
Total number of positively charged residues (Arg + Lys): 50

Atomic composition:

Carbon	C	2463
Hydrogen	H	3722
Nitrogen	N	656
Oxygen	O	705
Sulfur	S	14

Formula: $C_{2463}H_{3722}N_{656}O_{705}S_{14}$
Total number of atoms: 7560

Extinction coefficients:

Extinction coefficients are in units of $M^{-1} cm^{-1}$, at 280 nm measured in water.
Ext. coefficient 105700
Abs 0.1% (=1 g/l) 1.948, assuming all pairs of Cys residues form cystines

Ext. coefficient 105200
Abs 0.1% (=1 g/l) 1.939, assuming all Cys residues are reduced

Estimated half-life:

The N-terminal of the sequence considered is M (Met).
The estimated half-life is: 30 hours (mammalian reticulocytes, in vitro).
>20 hours (yeast, in vivo).
>10 hours (Escherichia coli, in vivo).

Instability index:

The instability index (II) is computed to be 37.88
This classifies the protein as stable.

Aliphatic index: 80.10

Grand average of hydropathicity (GRAVY): -0.293



研究背景与目的

序列分析

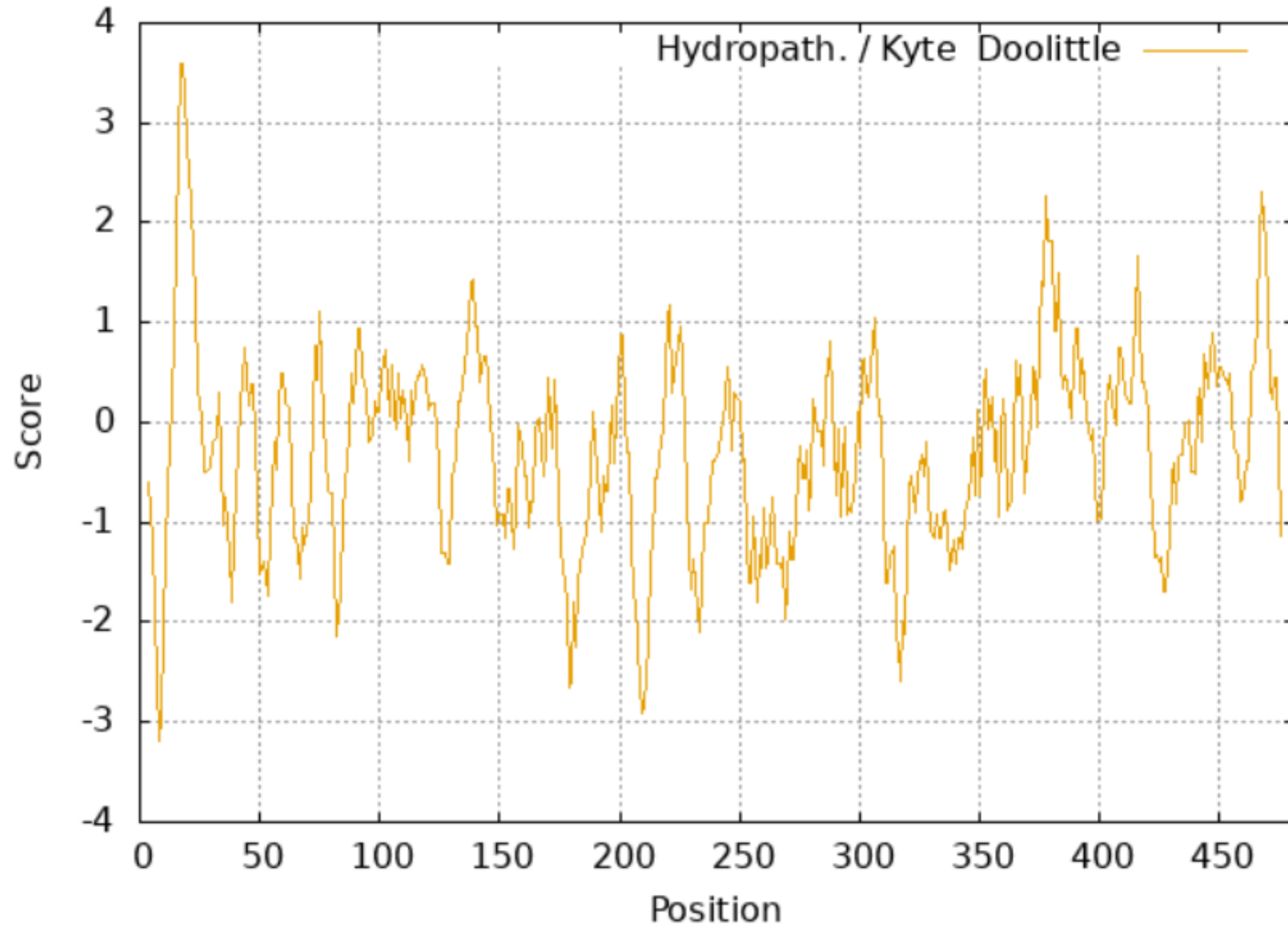
蛋白质性质分析

蛋白质结构预测

研究结论与展望

蛋白质的亲水性和疏水性分析

ProtScale output for user_sequences





蛋白质的跨膜区和信号肽分析

```
# AUW37688.1 Length: 481
# AUW37688.1 Number of predicted TMHs: 0
# AUW37688.1 Exp number of AAs in TMHs: 4.368839999999999
# AUW37688.1 Exp number, first 60 AAs: 4.19423
# AUW37688.1 Total prob of N-in: 0.24653
AUW37688.1 TMHMM2.0 outside 1 481
```

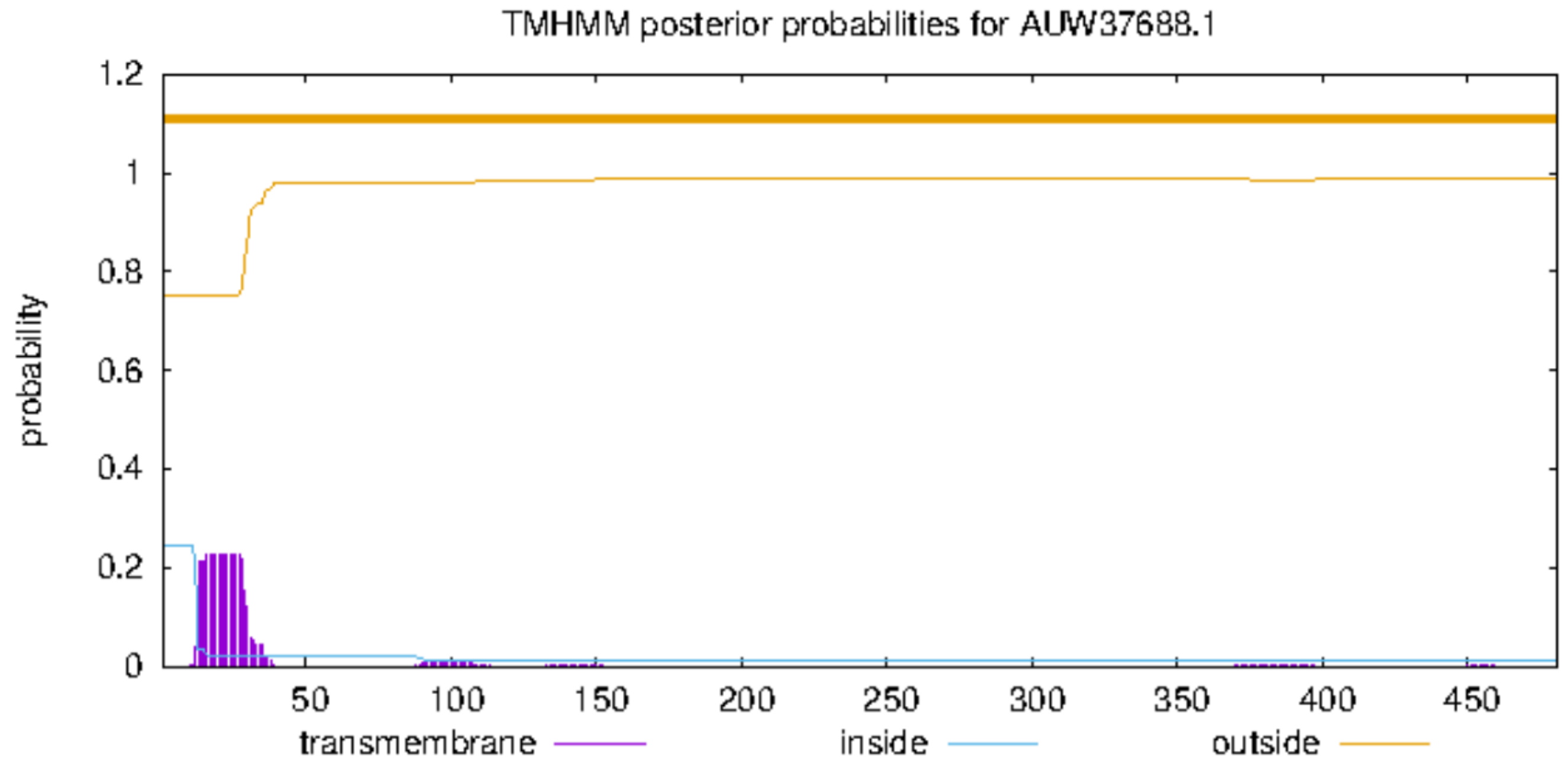
研究背景与目的

序列分析

蛋白质性质分析

蛋白质结构预测

研究结论与展望





信号肽预测

AUW37688.1 serine carboxypeptidase-like
33_ Triticum aestivum_

Prediction: Signal Peptide (Sec/SPI)

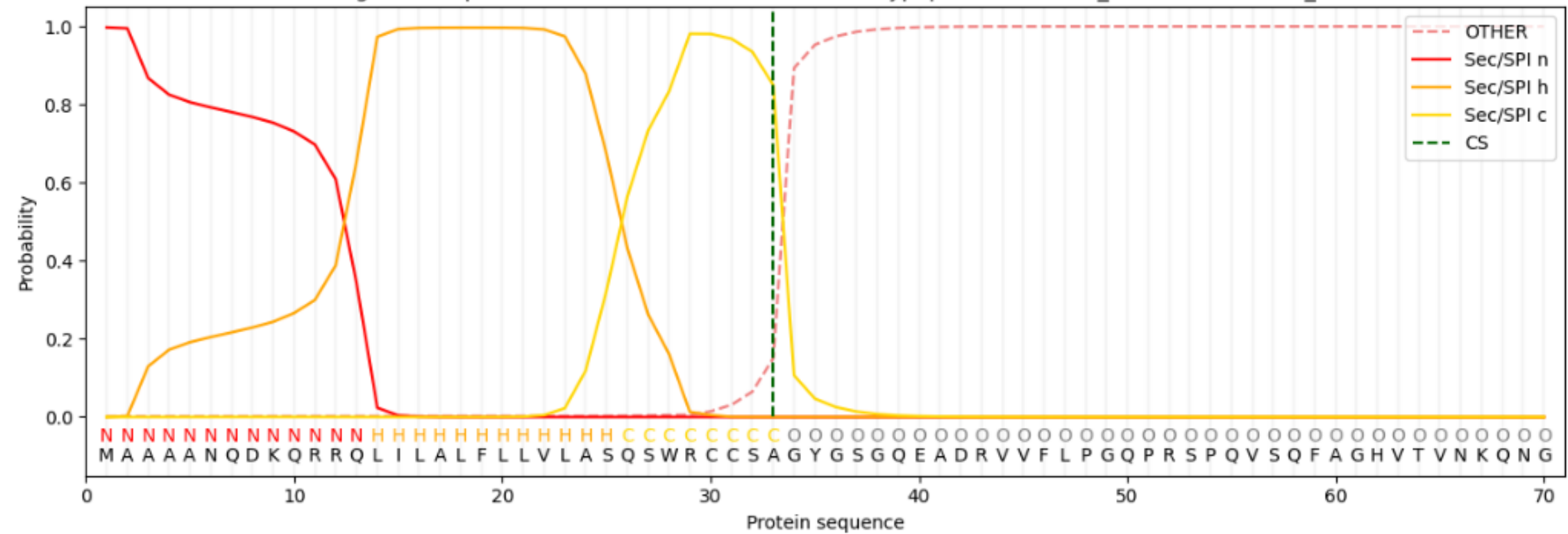
Cleavage site between pos. 33 and 34.

Probability 0.849979

Protein type	Other	Signal Peptide (Sec/SPI)	Lipoprotein signal peptide (Sec/SPII)	TAT signal peptide (Tat/SPI)	TAT Lipoprotein signal peptide (Tat/SPII)	Pilin-like signal peptide (Sec/SPIII)
Likelihood	0.0027	0.9954	0.0013	0.0002	0.0002	0.0002

Download: [PNG](#) / [EPS](#) / [Tabular](#)

SignalP 6.0 prediction: AUW37688.1 serine carboxypeptidase-like 33_ Triticum aestivum_



研究背景与目的

序列分析

蛋白质性质分析

蛋白质结构预测

研究结论与展望

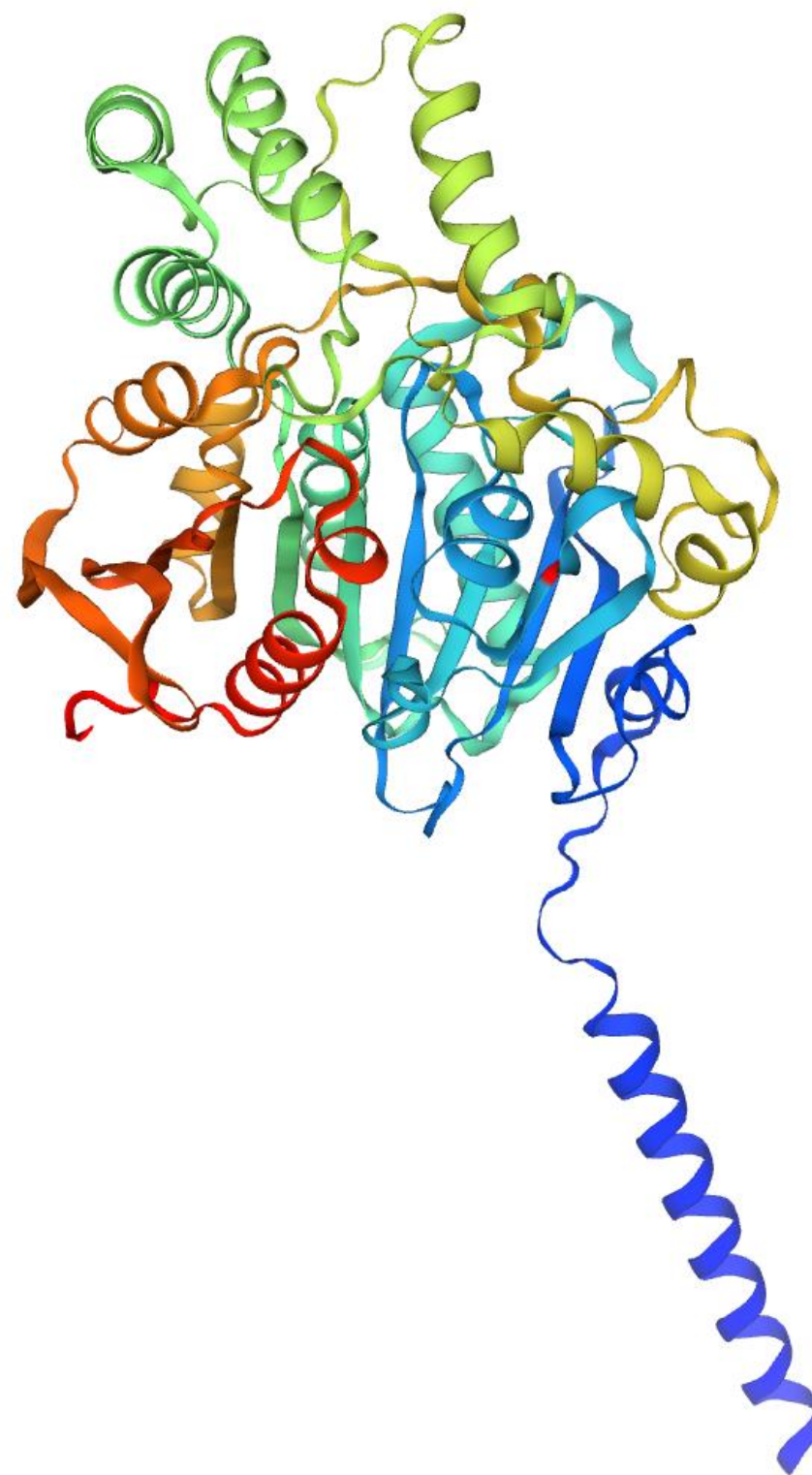
蛋白质保守基序分析



水稻、小麦及其近缘种的S10肽酶家族motif1-motif8高度保守，且在氨基酸序列上的分布及其位置基本一致。



蛋白质三级结构预测



同源建模 (SWISS-MODEL)

研究背景与目的

序列分析

蛋白质性质分析

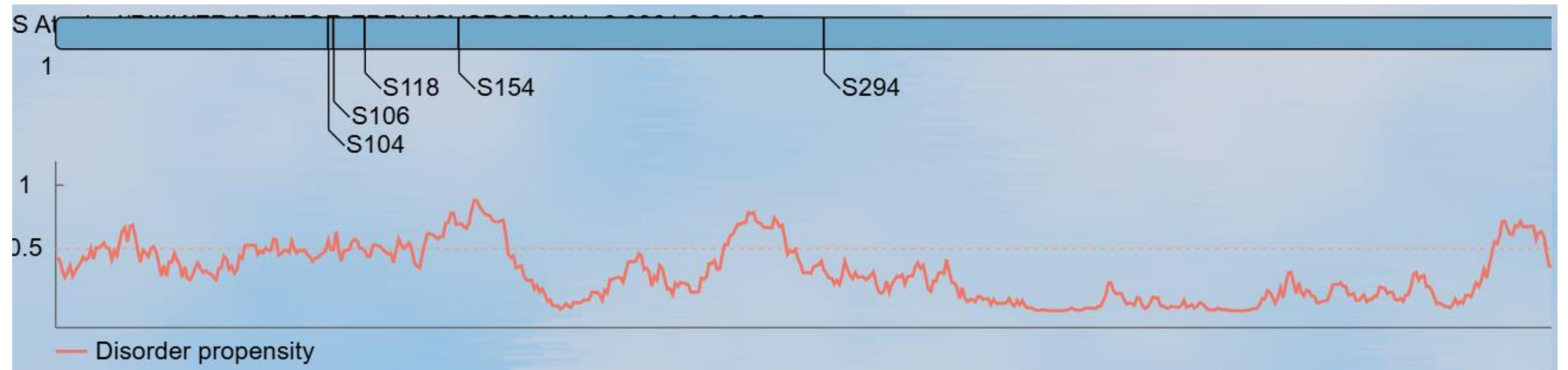
蛋白质结构预测

研究结论与展望



蛋白质的翻译后修饰位点预测

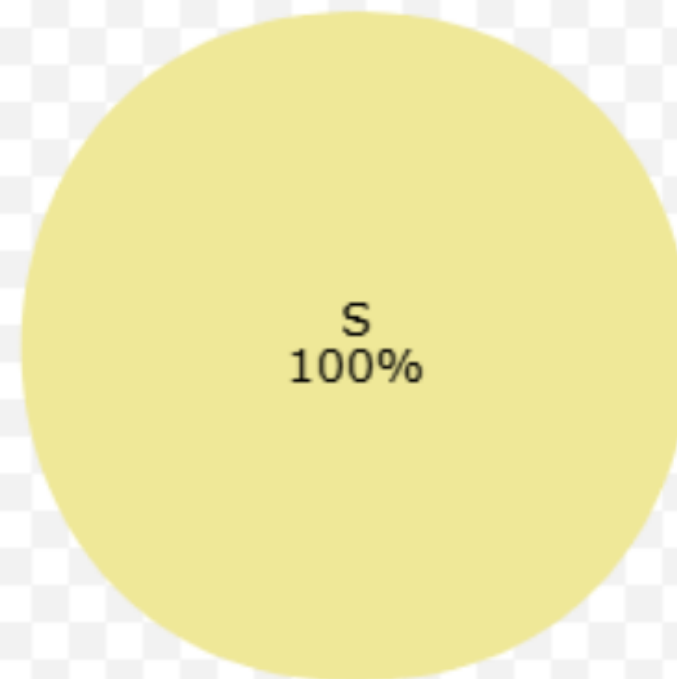
GPS 6.0 - Kinase-specific Phosphorylation Site Prediction:



Distribution of kinase families



Distribution of S/T/Y p-sites



研究背景与目的

序列分析

蛋白质性质分析

蛋白质结构预测

研究结论与展望



研究背景与目的

序列分析

蛋白质性质分析

蛋白质结构预测

研究结论与展望

研究结论与展望



研究结论

- **结论：**由小麦TaGS5-3B编码的丝氨酸羧肽酶样33序列在水稻、小麦及其近缘种中高度保守，具有8个保守基序。为亲水性蛋白有利于参与籽粒发育的水溶功能，结构预测其具有PIKK激酶特异性磷酸化位点，可在籽粒细胞分裂的信号转导中发挥作用。预测定位在细胞膜外，但也有研究表明可能是膜蛋白，因此要考虑生物信息学工具预测的不确定性。
- **不足：**翻译后修饰位点的预测未结合具体激酶家族的作用机制，修饰位点的真实性及功能关联性尚未验证；研究中无法预测到互作蛋白，且对下游通路未深入探究。
- **展望：**目前丝氨酸羧肽酶样33在uniprot中信息未经人工注释，且变异种、亚细胞定位、互作蛋白仍不清楚，后续除完善原有生物信息学分析内容外，进一步通过免疫沉淀、质谱分析等技术筛选其互作蛋白，结合磷酸化实验验证修饰位点功能，解析其参与的调控通路；运用染色、荧光蛋白、Western blot结合生物信息学工具多次进行亚细胞定位分析。

小组成员与分工



从左往右依次为：

 **赵雅芝：背景信息整理 + 蛋白序列比对**

 **刘俊钰：蛋白理化性质分析**

 **贾博宇：系统发育树构建**

 **董安琪：蛋白结构预测 + PPT制作**



感谢观看

敬请老师和各位同学批评指正

汇报人：董安琪

汇报时间：2026.1.10