开花调控基因CONSTANS的生物信息学分析

汇报人　刘　薇

小　组　G04

组　员　简银巧

刘　薇

刘长友

张　哲

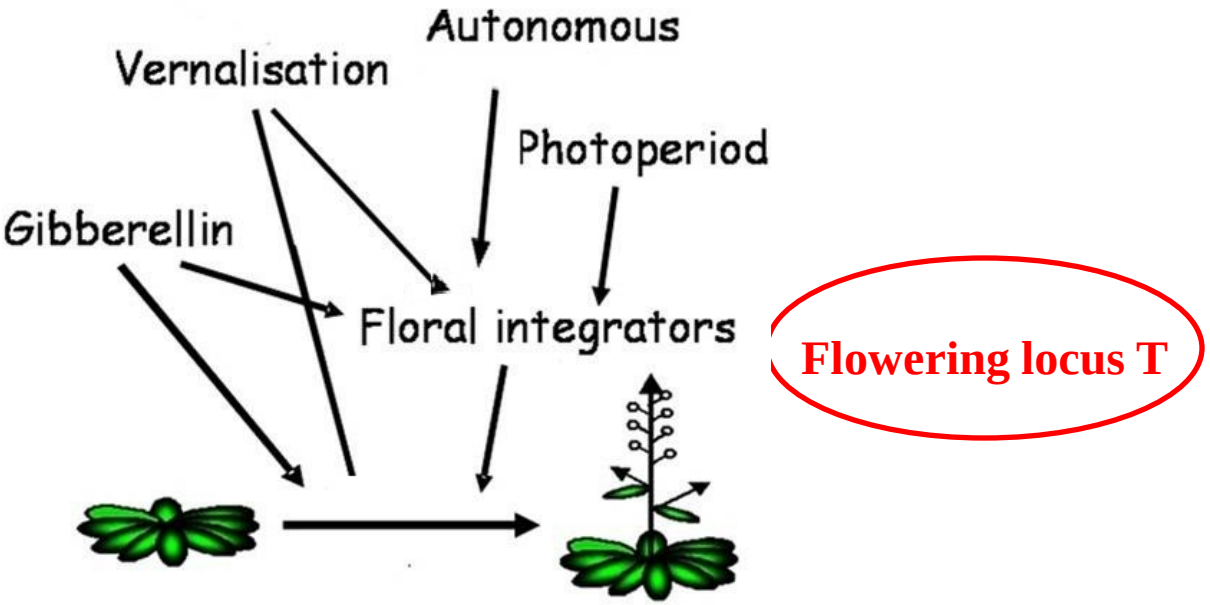
背景介绍

CO序列及结构分析

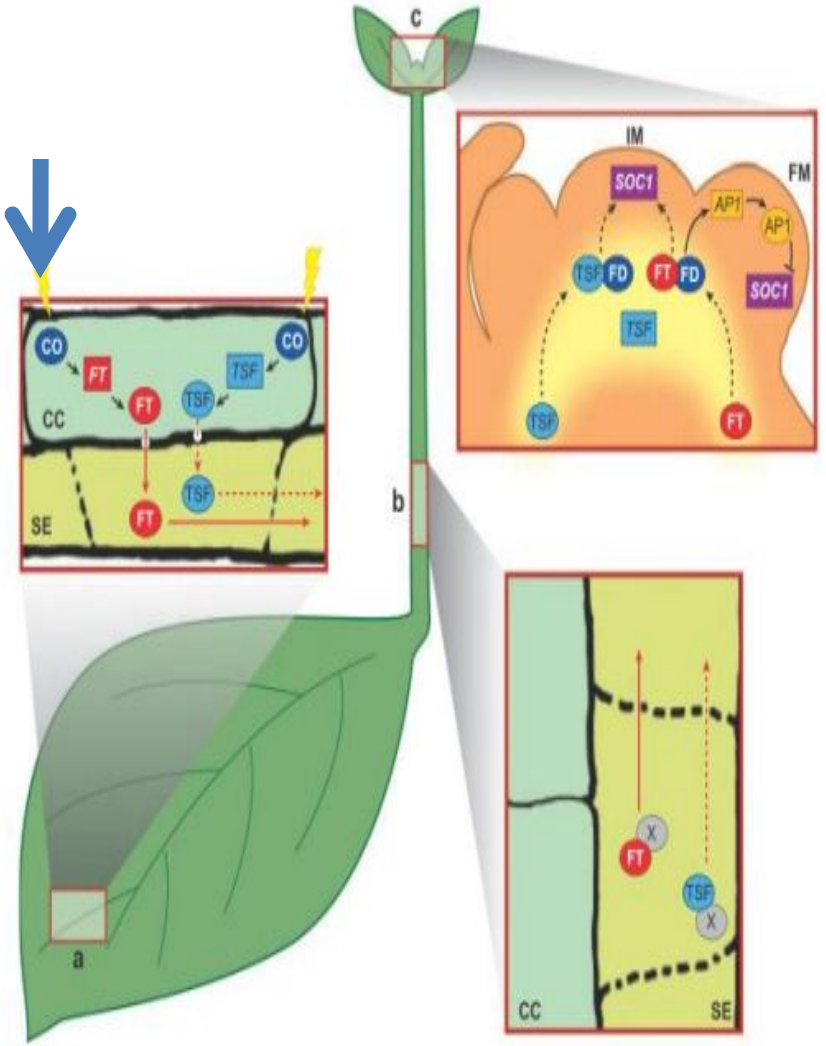
蛋白质结构预测

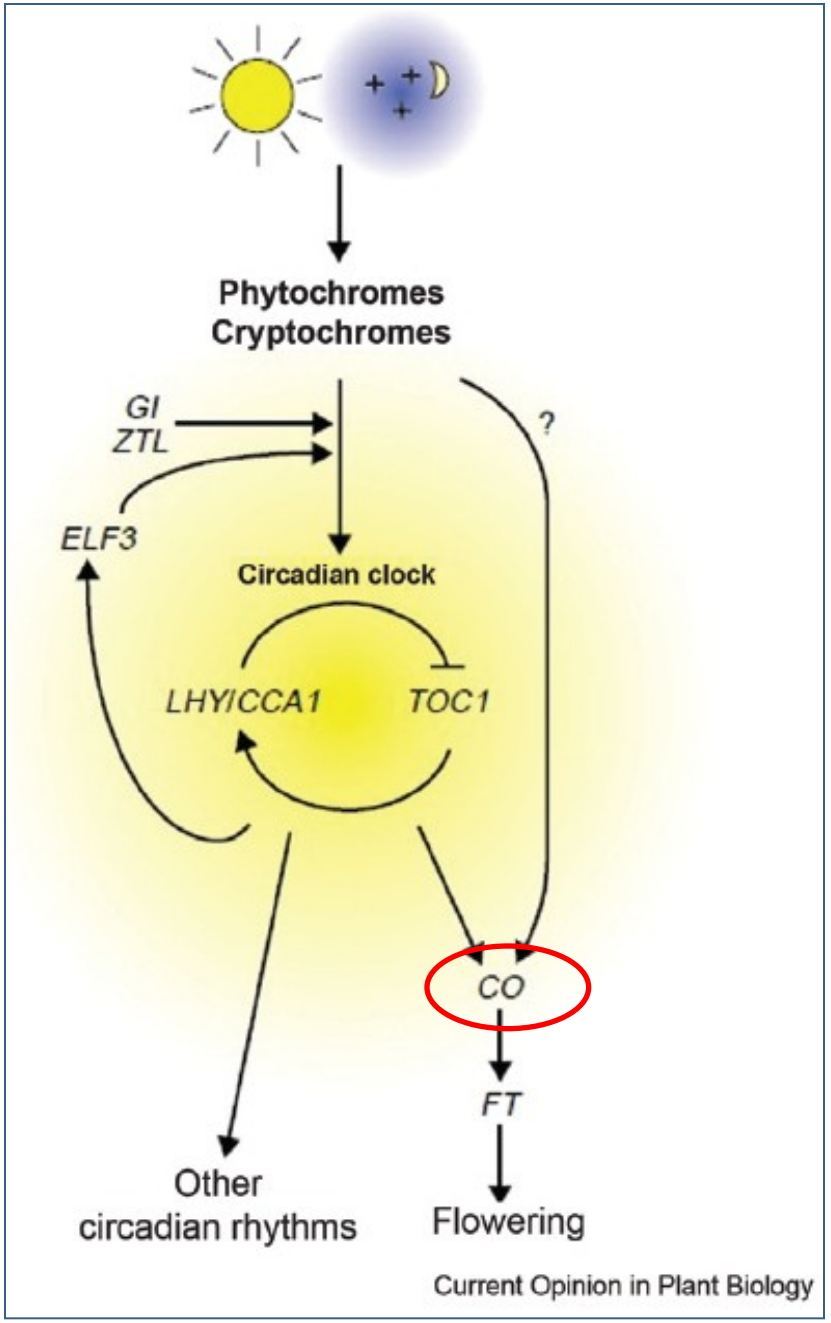
背景介绍

CONSTANS：1995年从拟南芥晚花突变体中发现。可以整合光信号和生物钟信号，节律性地激活 Flowering locus T 的表达，从而诱导开花。



调控开花的四种途径：

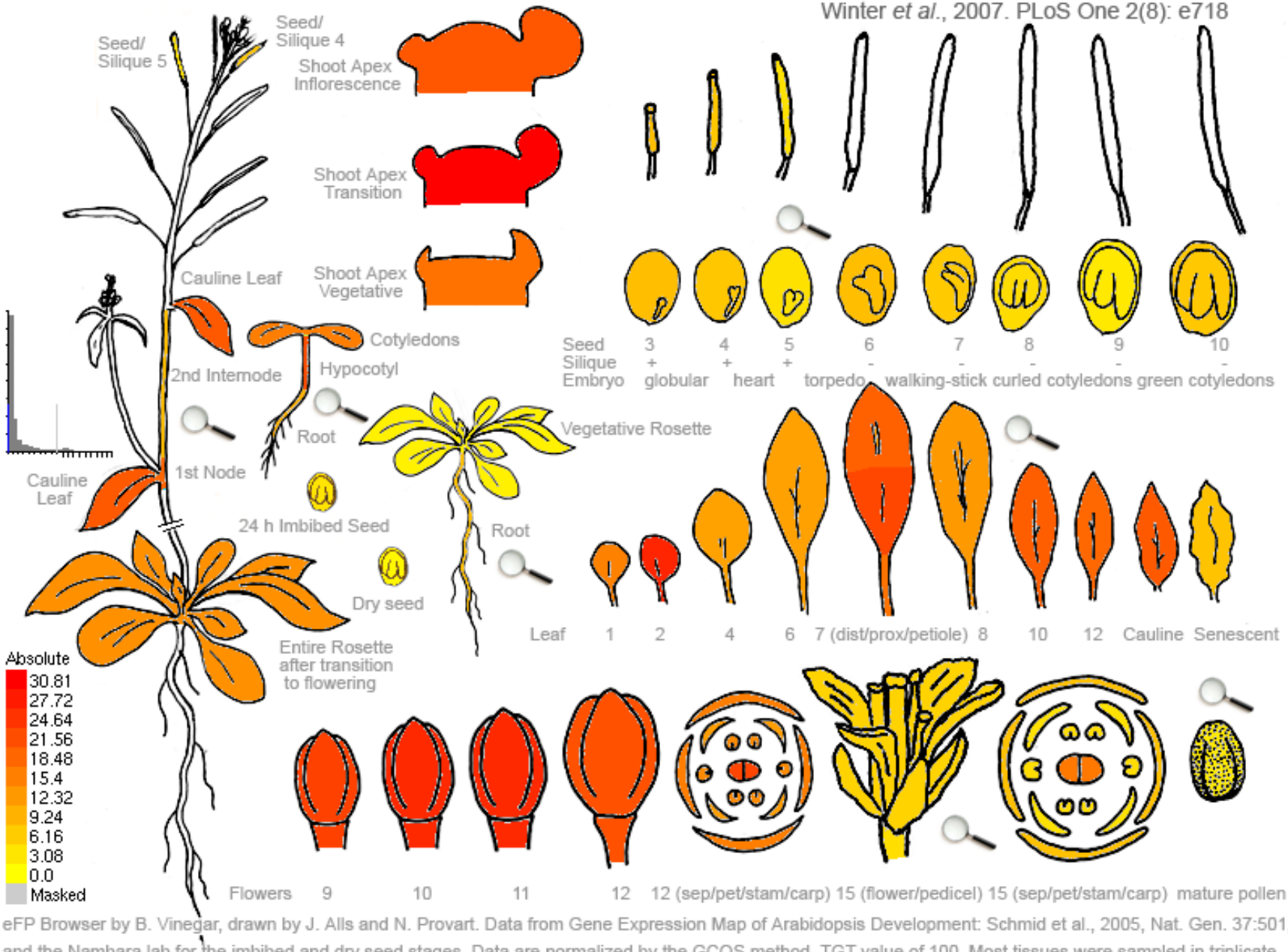

c
IM
SOC1
FM
AP1
AP1
SOC1
TSF
FD
FT
FD
TSF
TSF
FT
CO
FT
TSF
CO
CC
TSF
SE
FT
TSF
b
a
FT
X
TSF
X
CC
SE


Phytochromes
Cryptochromes
GI
ZTL
?
ELF3
Circadian clock
LHY/CCA1
TOC1
CO
FT
Other
circadian rhythms
Flowering
Current Opinion in Plant Biology

At5g15840 246525_at *co*

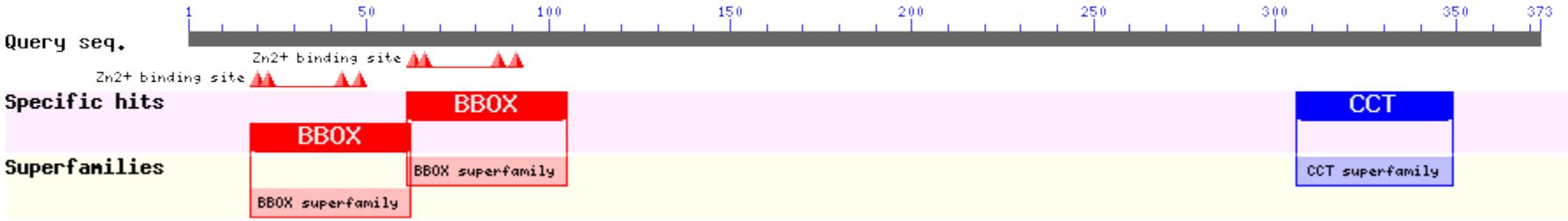
Arabidopsis eFP Browser at bar.utoronto.ca

Winter *et al.*, 2007. PLoS One 2(8): e718



eFP Browser by B. Vinegar, drawn by J. Alls and N. Provart. Data from Gene Expression Map of Arabidopsis Development: Schmid et al., 2005, Nat. Gen. 37:501, and the Nambara lab for the imbibed and dry seed stages. Data are normalized by the GCOS method, TGT value of 100. Most tissues were sampled in triplicate.

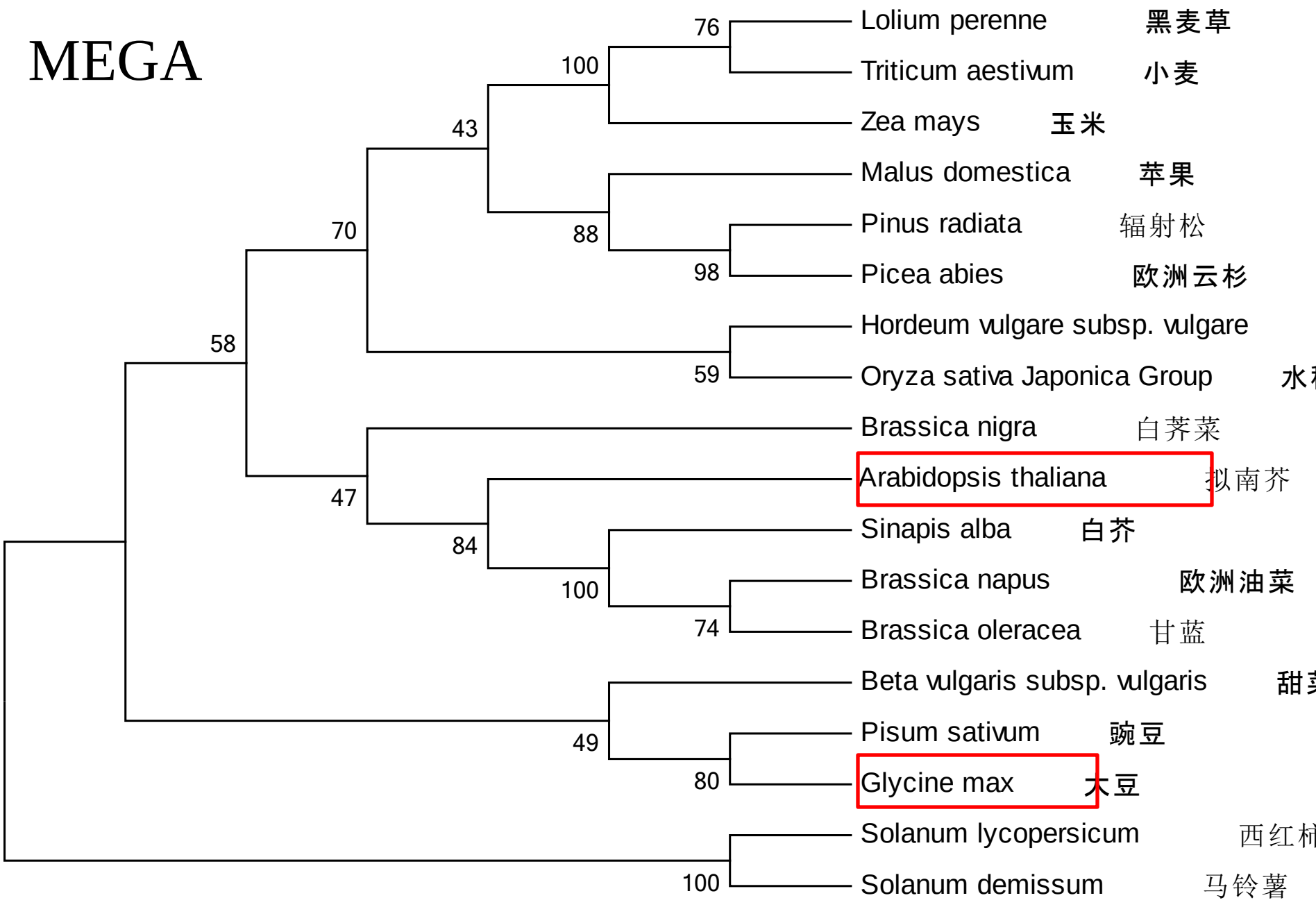
CDD



B-BOX：锌指蛋白。

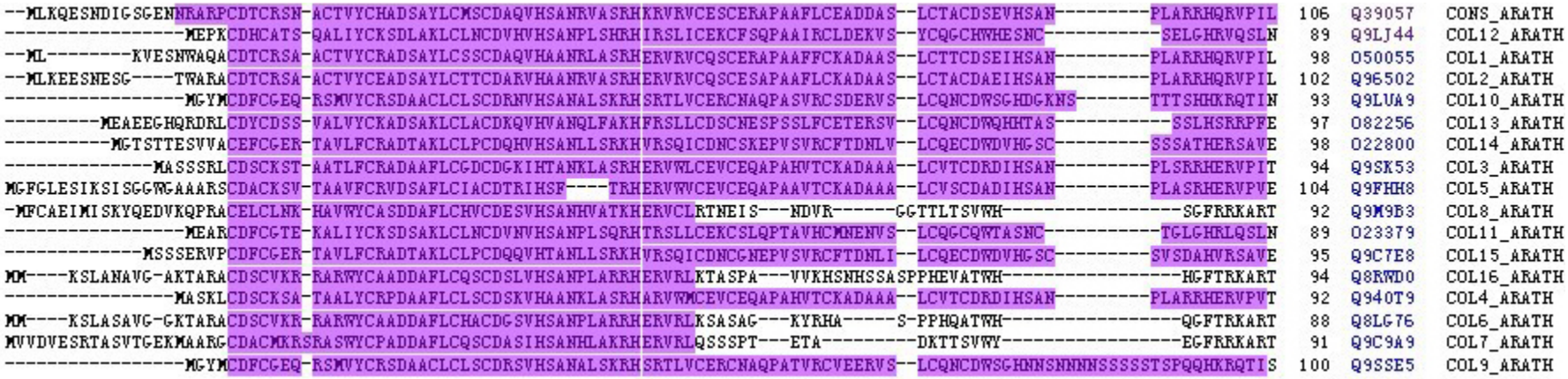
结合锌离子后能够稳定α-螺旋结构，使α-螺旋镶嵌于DNA的大沟中，因此含锌指结构的蛋白质都能与DNA或RNA结合。

CCT：具有核定位信号


MEGA
76
100
43
70
88
98
58
59
47
84
100
74
49
80
100
Lolium perenne 黑麦草
Triticum aestivum 小麦
Zea mays 玉米
Malus domestica 苹果
Pinus radiata 辐射松
Picea abies 欧洲云杉
Hordeum vulgare subsp. vulgare
Oryza sativa Japonica Group
Brassica nigra 白芥菜
Arabidopsis thaliana 拟南芥
Sinapis alba 白芥
Brassica napus 欧洲油菜
Brassica oleracea 甘蓝
Beta vulgaris subsp. vulgaris
Pisum sativum 豌豆
Glycine max 大豆
Solanum lycopersicum
Solanum demissum 马铃薯

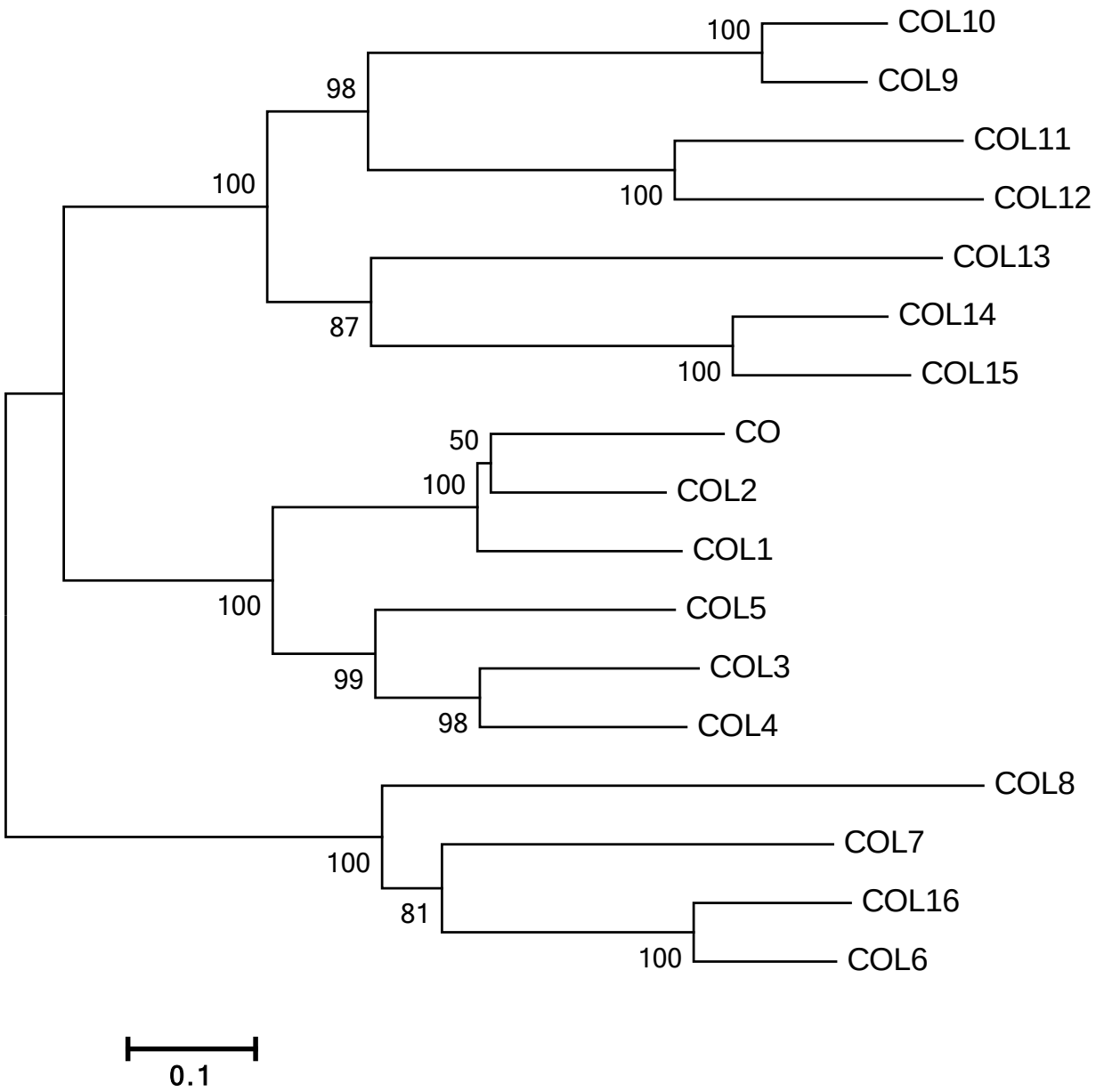
拟南芥CO家族成员序列比对

N 端



C 端

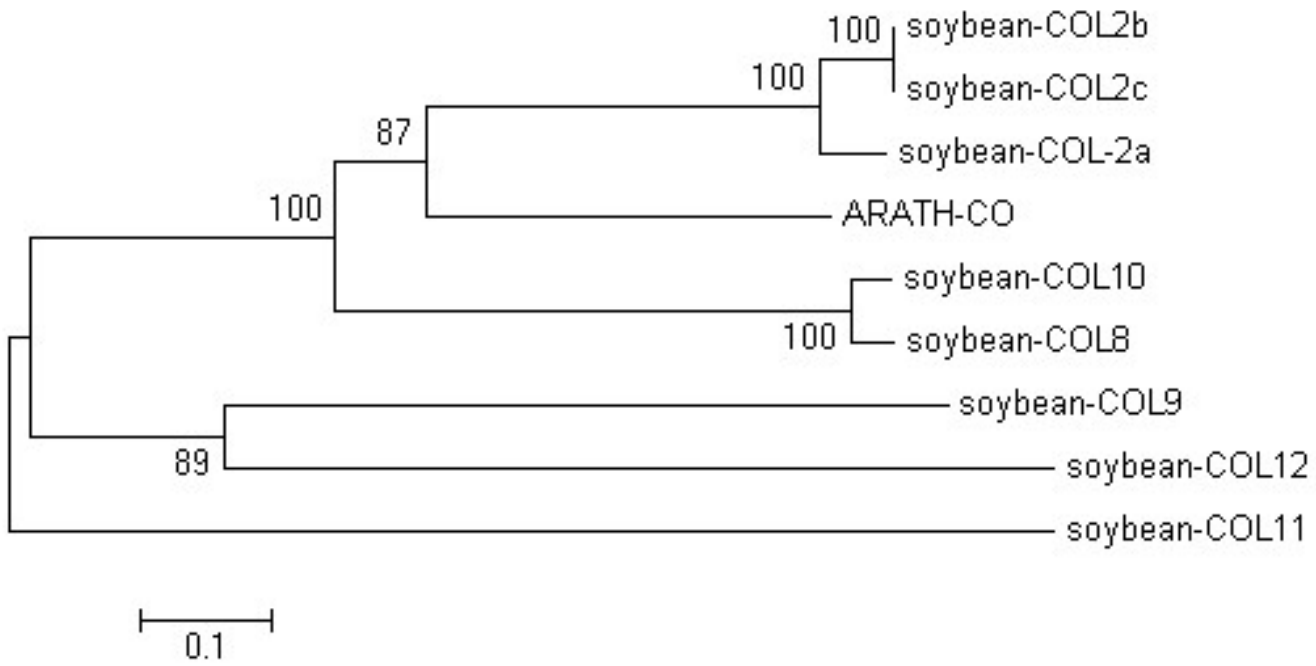
拟南芥CO家族成员进化树分析



- 大豆是典型的短日作物，不同品种早晚有较大的差异。
- 影响大豆开花候选基因的克隆，表达分析，转基因研究，鉴定其在成花转变中的作用，对探究基因与开花整合因子FT的相互关系以及调节大豆开花的分子网络的描绘有重要意义。

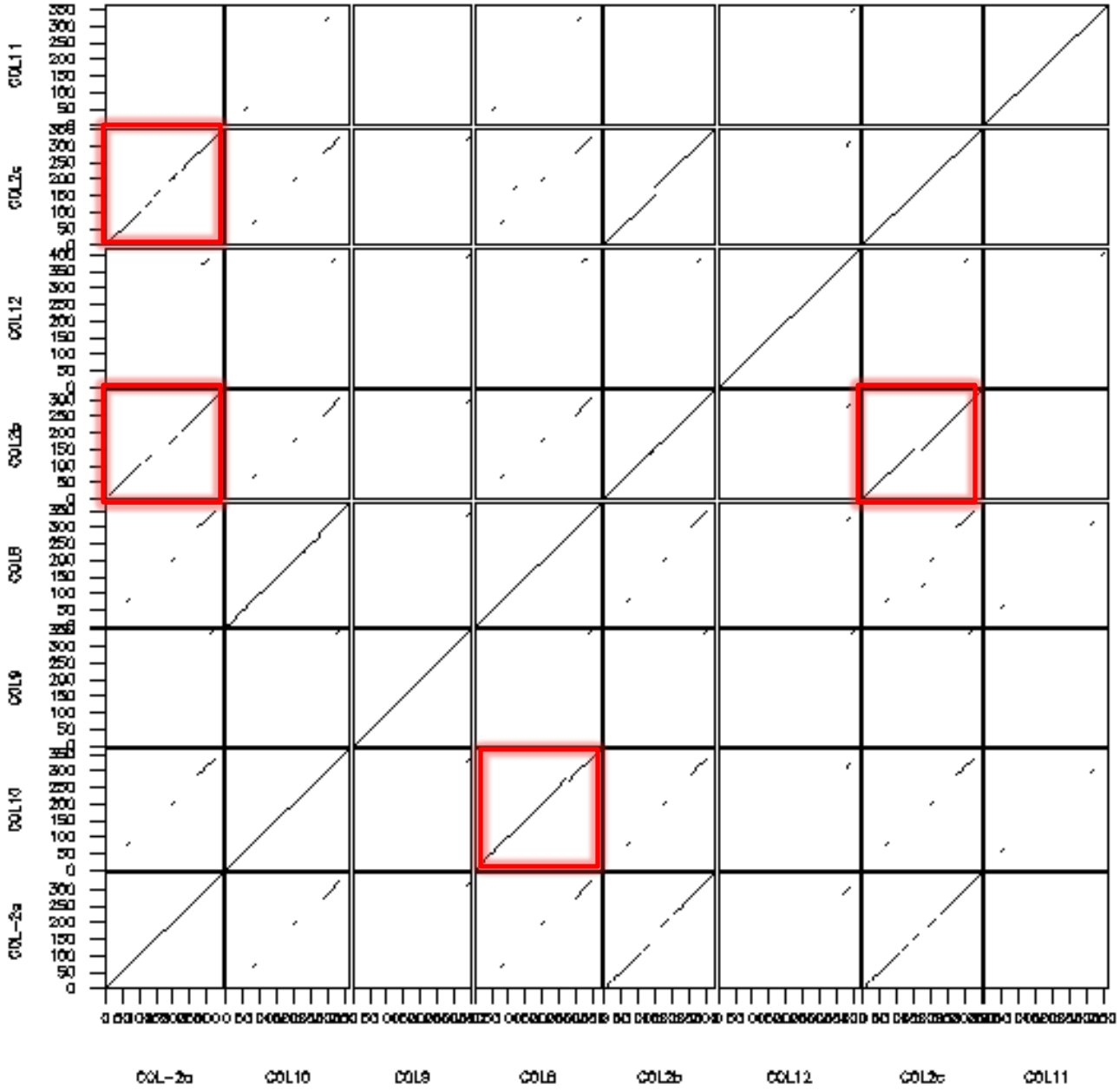
大豆CO-Like基因分析

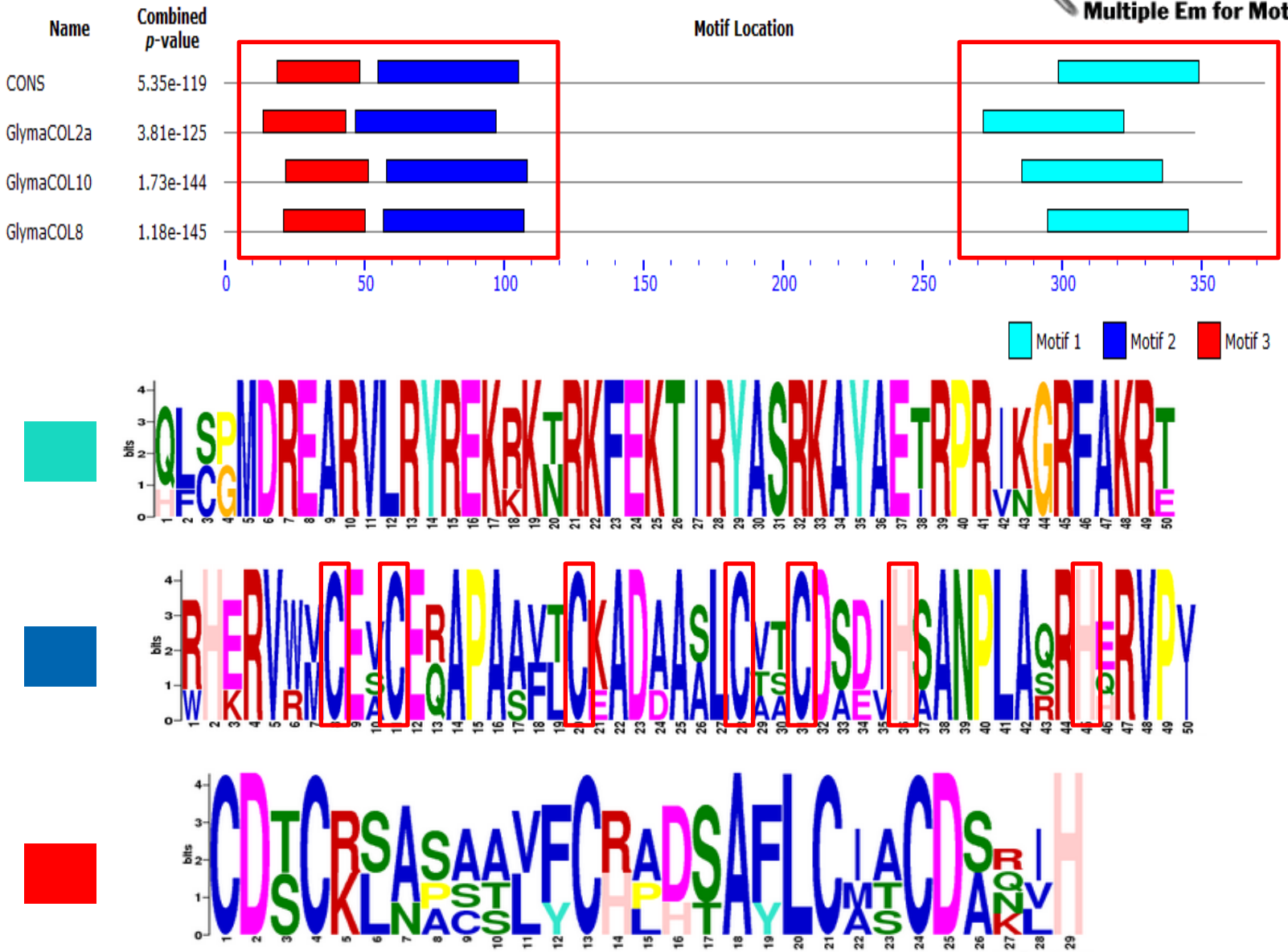
	Entry	Entry name	Status	Protein names	Gene names	Organism	Length
☑	B9MSS5	B9MSS5_SOYBN	☆	CONSTANS-like 2a	COL2a	Glycine max (Soybean) (Glycine hispida)	348
☑	D0EP05	D0EP05_SOYBN	☆	CONSTANS-like zinc finger protein	COL10	Glycine max (Soybean) (Glycine hispida)	365
☑	D0EP04	D0EP04_SOYBN	☆	CONSTANS-like zinc finger protein	COL9	Glycine max (Soybean) (Glycine hispida)	351
☑	D0EP03	D0EP03_SOYBN	☆	CONSTANS-like zinc finger protein	COL8	Glycine max (Soybean) (Glycine hispida)	374
☑	B9MSS6	B9MSS6_SOYBN	☆	CONSTANS-like 2b	COL2b	Glycine max (Soybean) (Glycine hispida)	328
☑	D0EP07	D0EP07_SOYBN	☆	CONSTANS-like zinc finger protein	COL12	Glycine max (Soybean) (Glycine hispida)	418
☑	B9MSS7	B9MSS7_SOYBN	☆	Zinc finger protein CONSTANS-LIKE 2	COL2c	Glycine max (Soybean) (Glycine hispida)	352
☑	D0EP06	D0EP06_SOYBN	☆	CONSTANS-like zinc finger protein	COL11	Glycine max (Soybean) (Glycine hispida)	366



Polydot

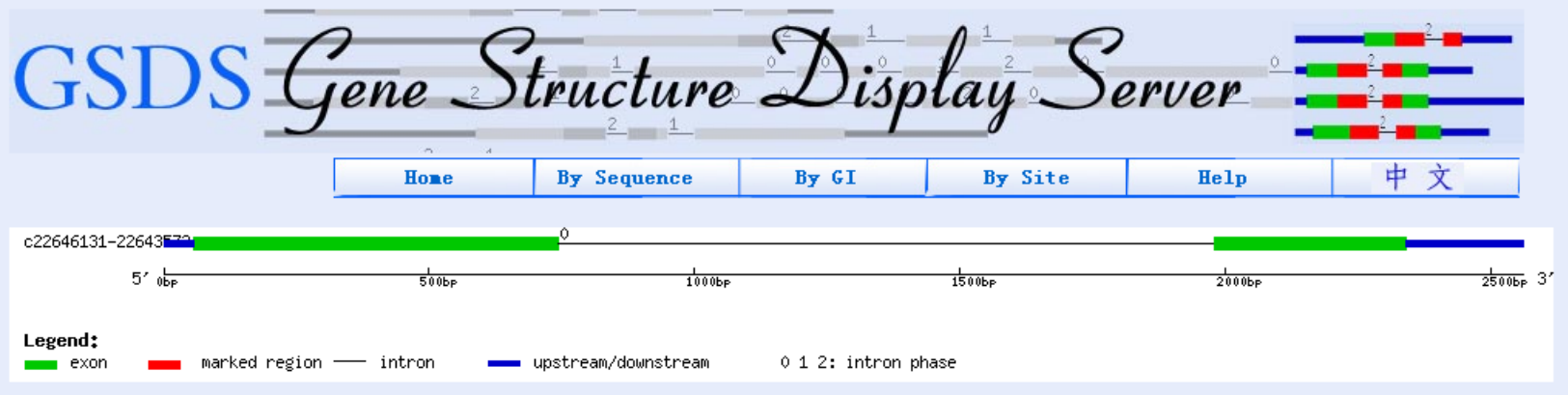
COL8和COL10之间，COL2a和COL2b和COL2c之间相似度很高

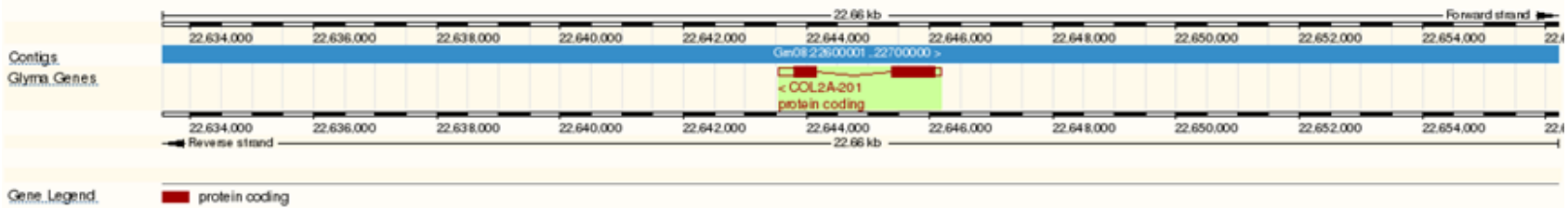


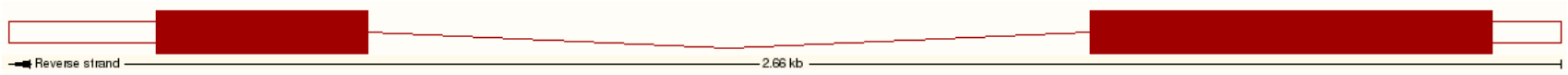


B-box： $CX_2CX_8CX_7CX_2CX_4HX_8H$

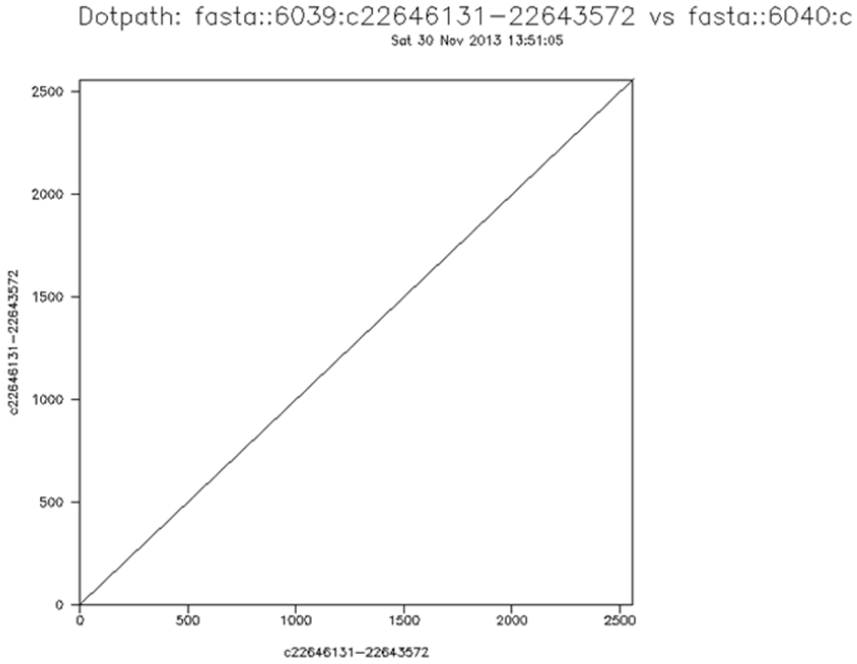
GmCOL2a的基因结构



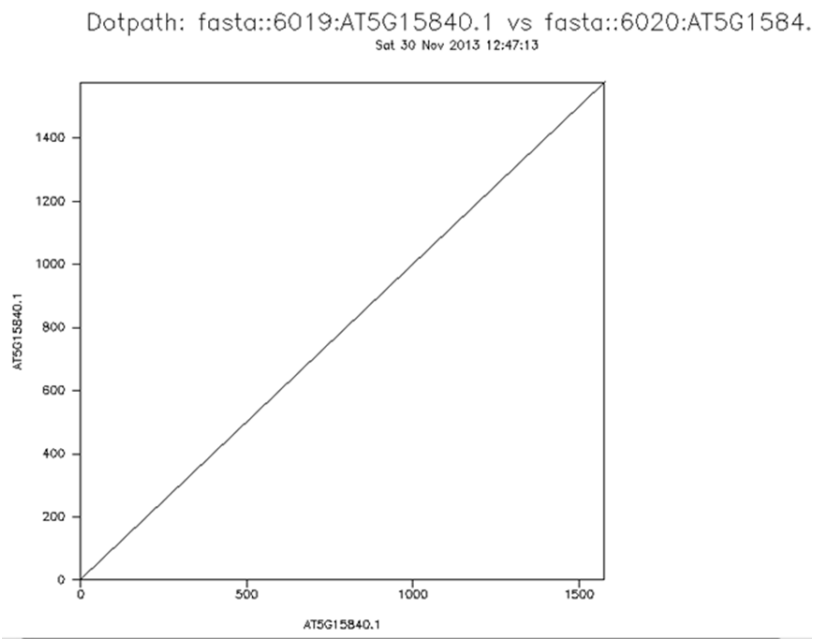




GmCOL2a的重复序列分析

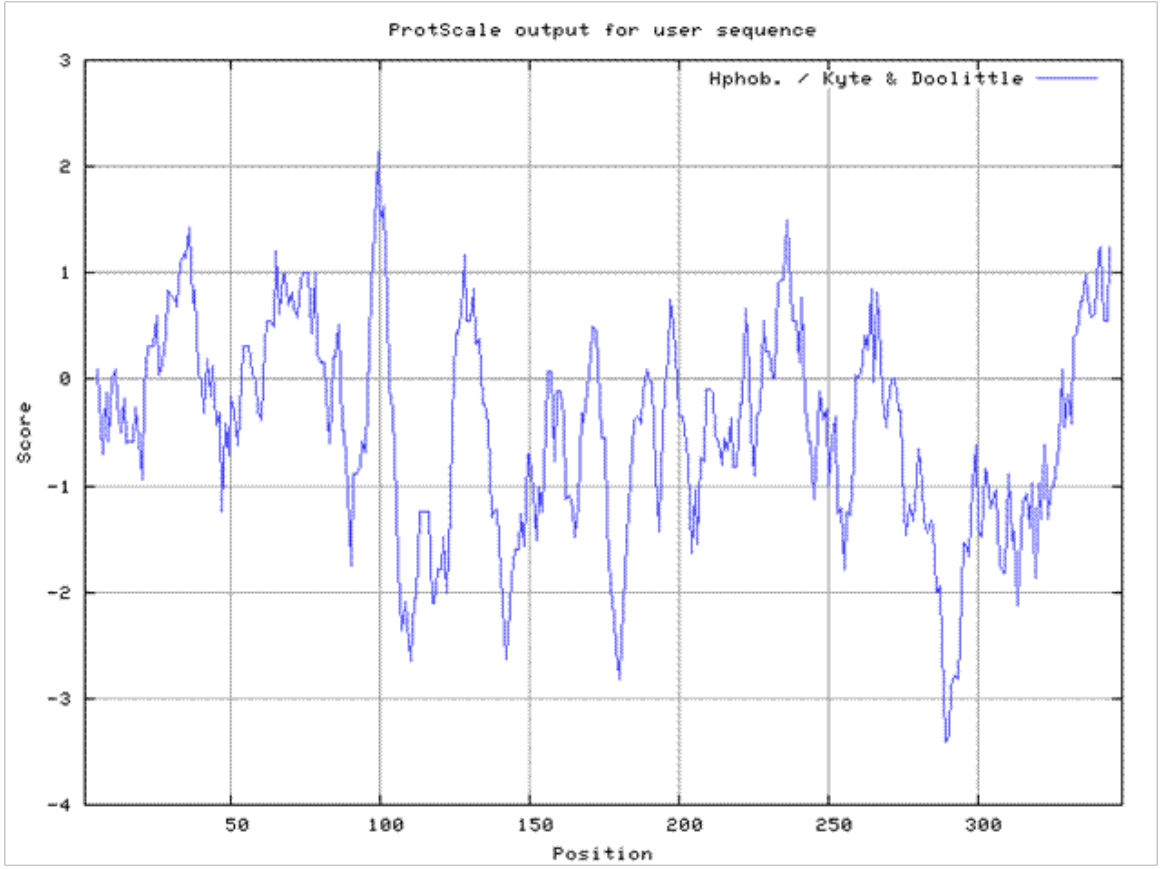


大豆

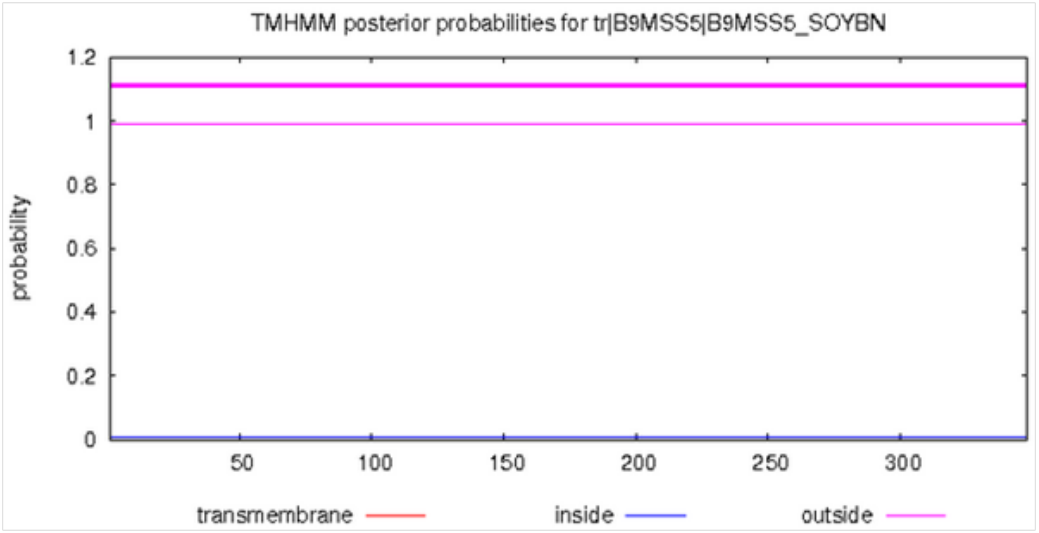


拟南芥

ProScale



TMHMM



GmCOL2a的亚细胞定位预测

CENTERFO RBIOLOGI CALSEQU ENCEANA LYSIS CBS

TargetP 1.1 Server - prediction results

Technical University of Denmark

```
### targetp v1.1 prediction results ######################################
Number of query sequences:  1
Cleavage site predictions not included.
Using PLANT networks.

Name                    Len      cTP     mTP      SP   other  Loc   RC
-------------------------------------------------------------------------
Soybean_COL_2a          348    0.144   0.262   0.018   0.680     _     3
-------------------------------------------------------------------------
cutoff                         0.000   0.000   0.000   0.000
```

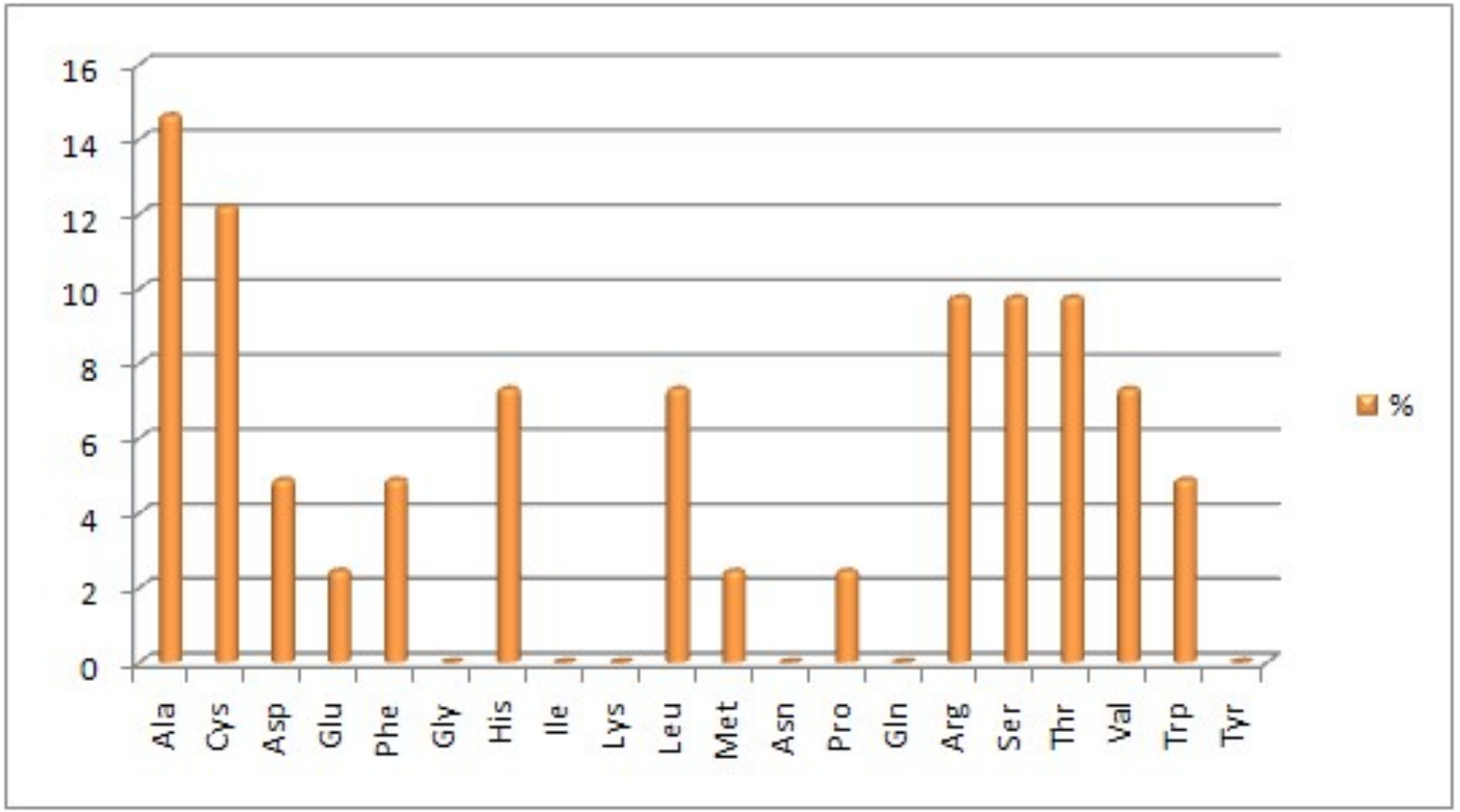
GmCOL2a的氨基酸组分分析

Pepstats

AA	SwissProt*	Glymax-COL2a	
	%	number	%
Ala	8.25	38	10.92
Cys	1.37	11	3.161
Asp	5.45	29	8.333
Met	2.42	7	2.011
Phe	3.86	15	4.31
Gln	3.93	8	2.299
His	2.27	18	5.172
Tyr	2.92	9	2.586
Ile	5.95	11	3.161
Lys	5.84	13	3.736
Gly	7.07	15	4.31
Asn	4.06	17	4.885
Pro	4.7	18	5.172
Glu	6.75	16	4.598
Arg	5.53	21	6.034
Ser	6.57	31	8.908
Thr	5.34	21	6.034
Val	6.87	27	7.759
Trp	1.08	4	1.149
Leu	9.66	19	5.46

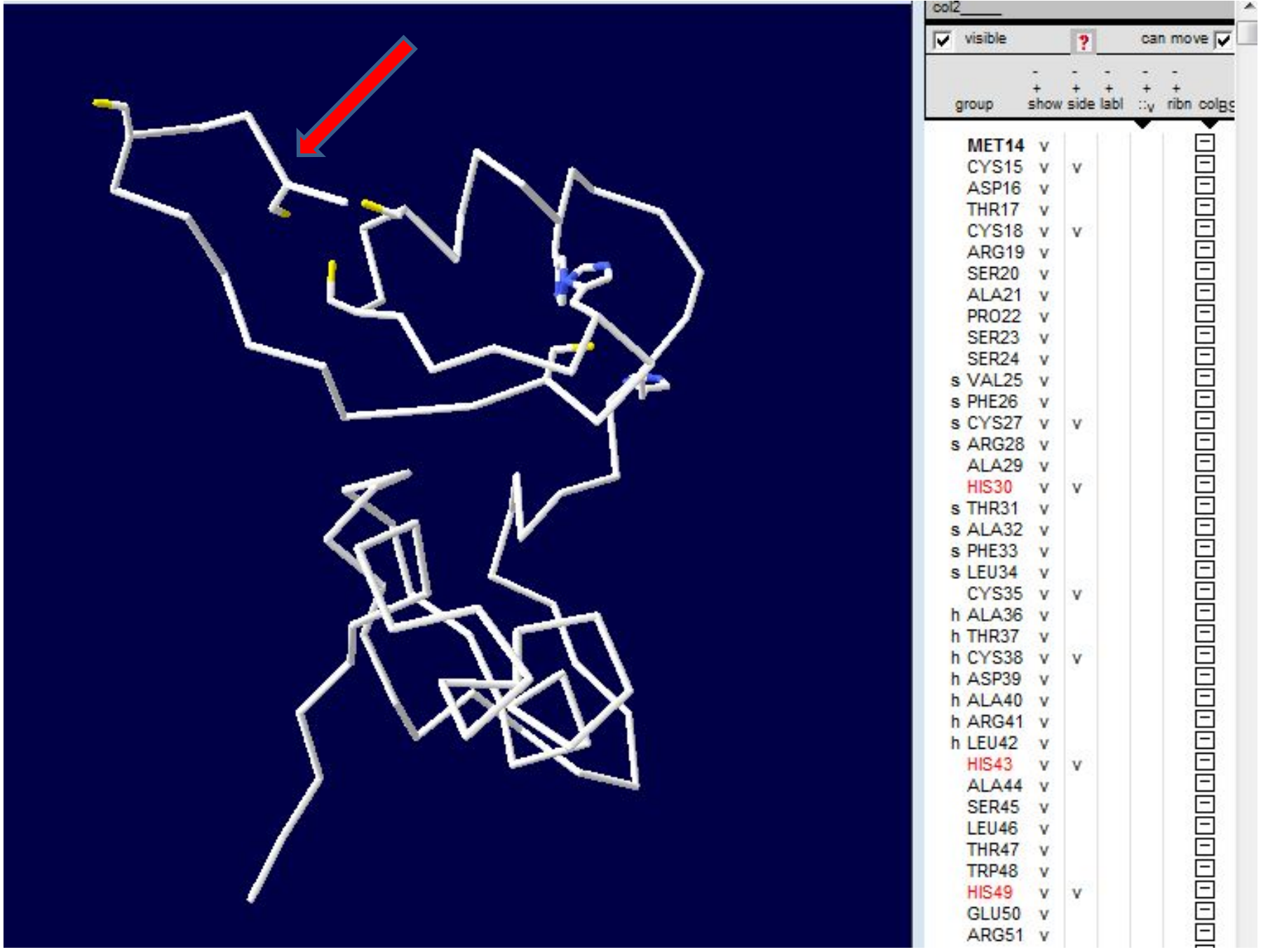
Pepstats

GmCOL2a B-box区氨基酸组分分析

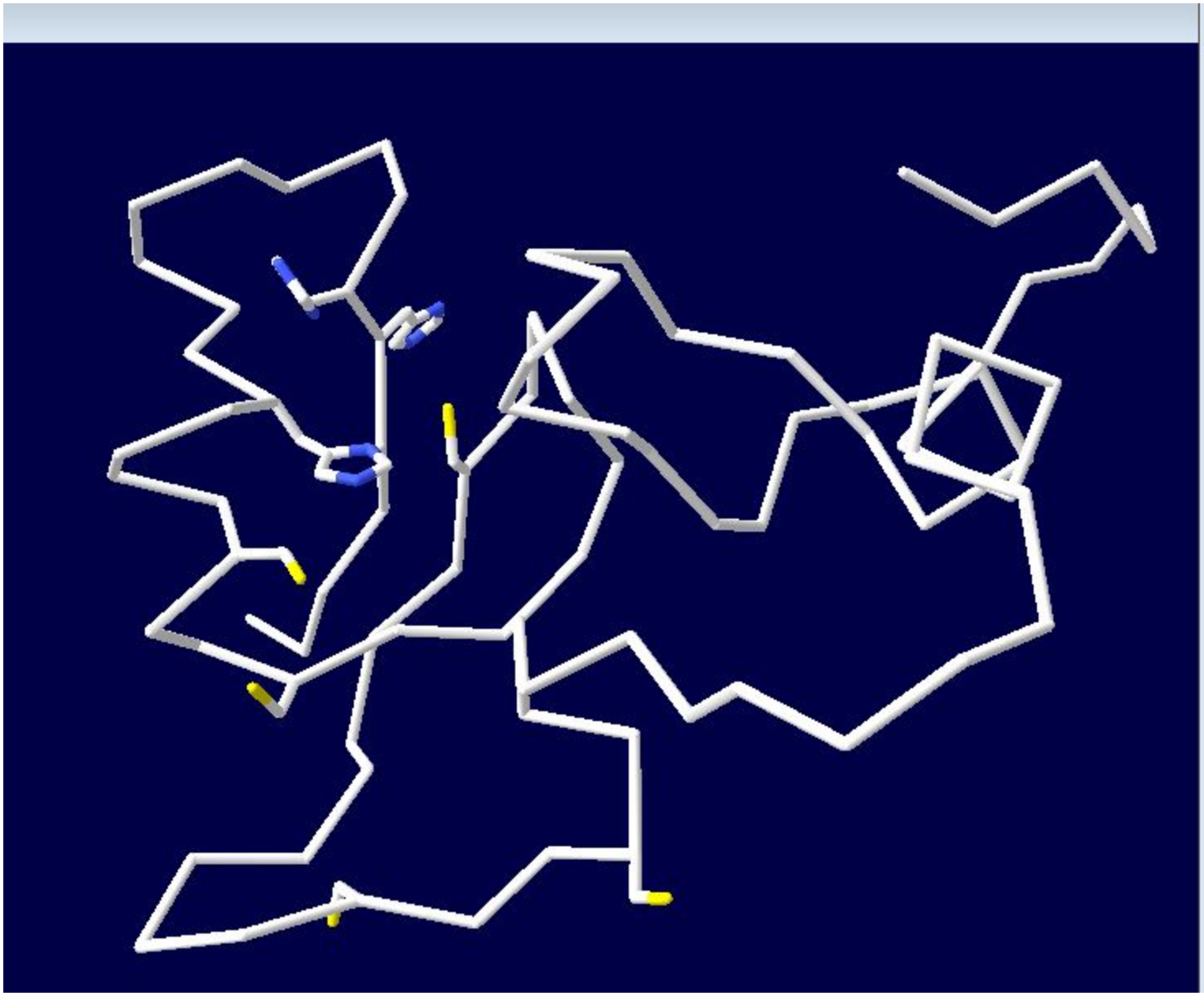



Sequence
Secondary structure
SS confidence
Disorder
Disorder confidence

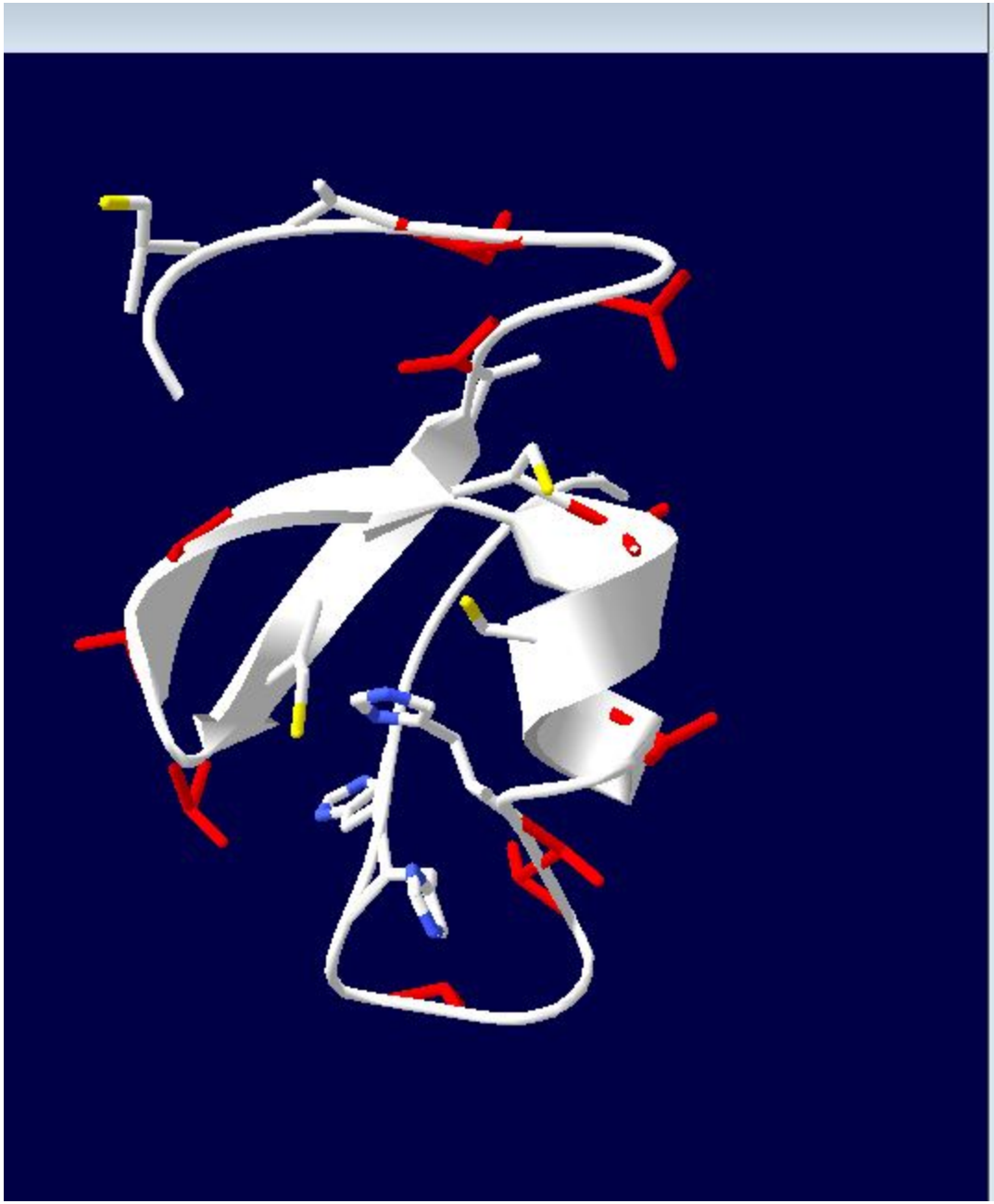
$CX_2CX_8CX_7CX_2CX_4HX_8H$



$$CX_2CX_8CX_7CX_2CX_4HX_8H$$



group	show	side	labl	ribn	col
VAL52	v				
TRP53	v				
VAL54	v				
CYS55	v	v			
GLU56	v				
ALA57	v				
CYS58	v	v			
GLU59	v				
ARG60	v				
ALA61	v				
PRO62	v				
ALA63	v				
ALA64	v				
s PHE65	v				
s LEU66	v				
s CYS67	v	v			
s LYS68	v				
ALA69	v				
ASP70	v				
s ALA71	v				
s ALA72	v				
s SER73	v				
s LEU74	v				
CYS75	v	v			
h ALA76	v				
h SER77	v				
h CYS78	v	v			
h ASP79	v				
h ALA80	v				
ASP81	v				
ILE82	v				
HIS83	v	v			
ALA84	v				
ALA85	v				
ASN86	v				
PRO87	v				
LEU88	v				
ALA89	v				
SER90	v				
ARG91	v				
HIS92	v	v			
HIS93	v	v			
ARG94	v				



visible ? can move

group	show	side	labl	ribn	colgs
CYS55	v	v		v	
GLU56				v	
ALA57	v	v		v	
CYS58	v	v		v	
GLU59				v	
ARG60				v	
ALA61	v	v		v	
PRO62				v	
ALA63	v	v		v	
ALA64	v	v		v	
s PHE65				v	
s LEU66				v	
s CYS67	v	v		v	
s LYS68				v	
ALA69	v	v		v	
ASP70				v	
s ALA71	v	v		v	
s ALA72	v	v		v	
s SER73				v	
s LEU74				v	
CYS75	v	v		v	
h ALA76	v	v		v	
h SER77				v	
h CYS78	v	v		v	
h ASP79				v	
h ALA80	v	v		v	
ASP81				v	
ILE82				v	
HIS83	v	v		v	
ALA84	v	v		v	
ALA85	v	v		v	
ASN86				v	
PRO87				v	
LEU88				v	
ALA89	v	v		v	
SER90				v	
ARG91				v	
HIS92	v	v		v	
HIS93	v	v		v	
ARG94				v	
VAL95				v	
PRO96				v	
ILE97				v	

Thank you !