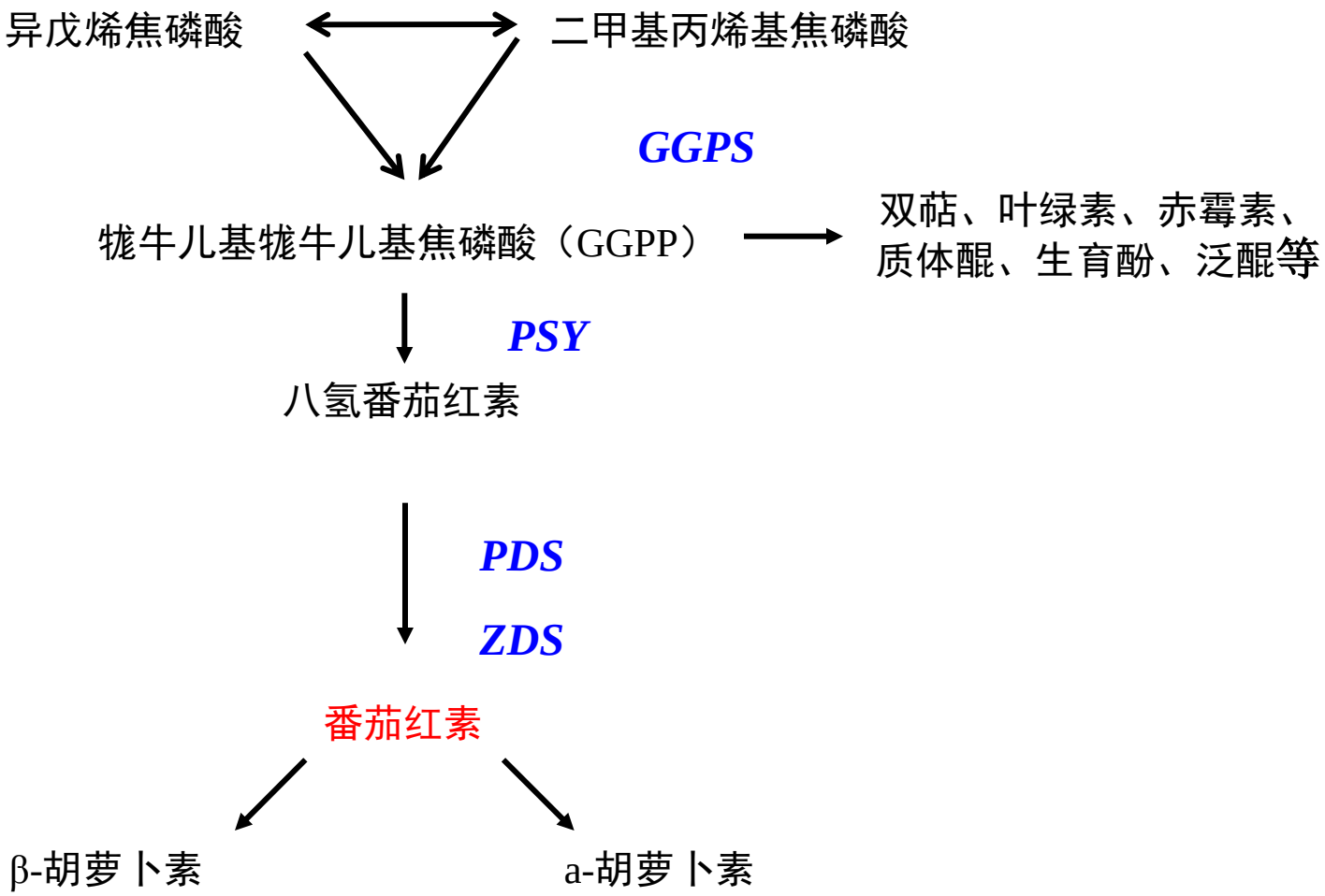
野生番茄和普通番茄 *GGPS*基因的差异比较

孙超 祝光涛 苏彦宾 白龙强

2013. 1. 11

研究背景

番茄红素是番茄果实类胡萝卜素代谢途径中的产物



研究背景

➢研究发现野生番茄LA1777的渐渗系TA517（该渐渗系以普通番茄E6203为遗传背景，含有来自LA1777第4染色体末端的片段），我们发现这两种番茄材料中的果实颜色具有有明显的差异。

E6203　　TA517

E6203　　TA517

研究背景

- 将野生番茄和普通番茄的*GGPS*基因并进行了克隆和比对，结果表明：

1. 两基因cDNA长度为1092bp， 编码363AA。

2. 两基因核苷酸序列存在23个碱基的差异。

3. 氨基酸水平存在13个差异。

生物信息分析策略

1、相关文献的获取和查阅
2、两基因DNA 序列、氨基酸序列的差异比较。
3、GGPS家族基因的系统发育
4、跨膜结构域、亚细胞定位的分析
5、蛋白质三维结构的预测和功能位点的分析

目的：推测两个酶是否有活性的差异
和影响酶活性的差异位点。

两序列氨基酸差异比较

1、两个氨基酸序列之间存在13个差异，差异的氨基酸主要集中在多肽链的N端，多肽链的C端相对保守。

2、氨基酸序列之间差异位点是：

6，24，34，38，48，52，60，

71，85，130，197，279，360

两序列氨基酸差异比较

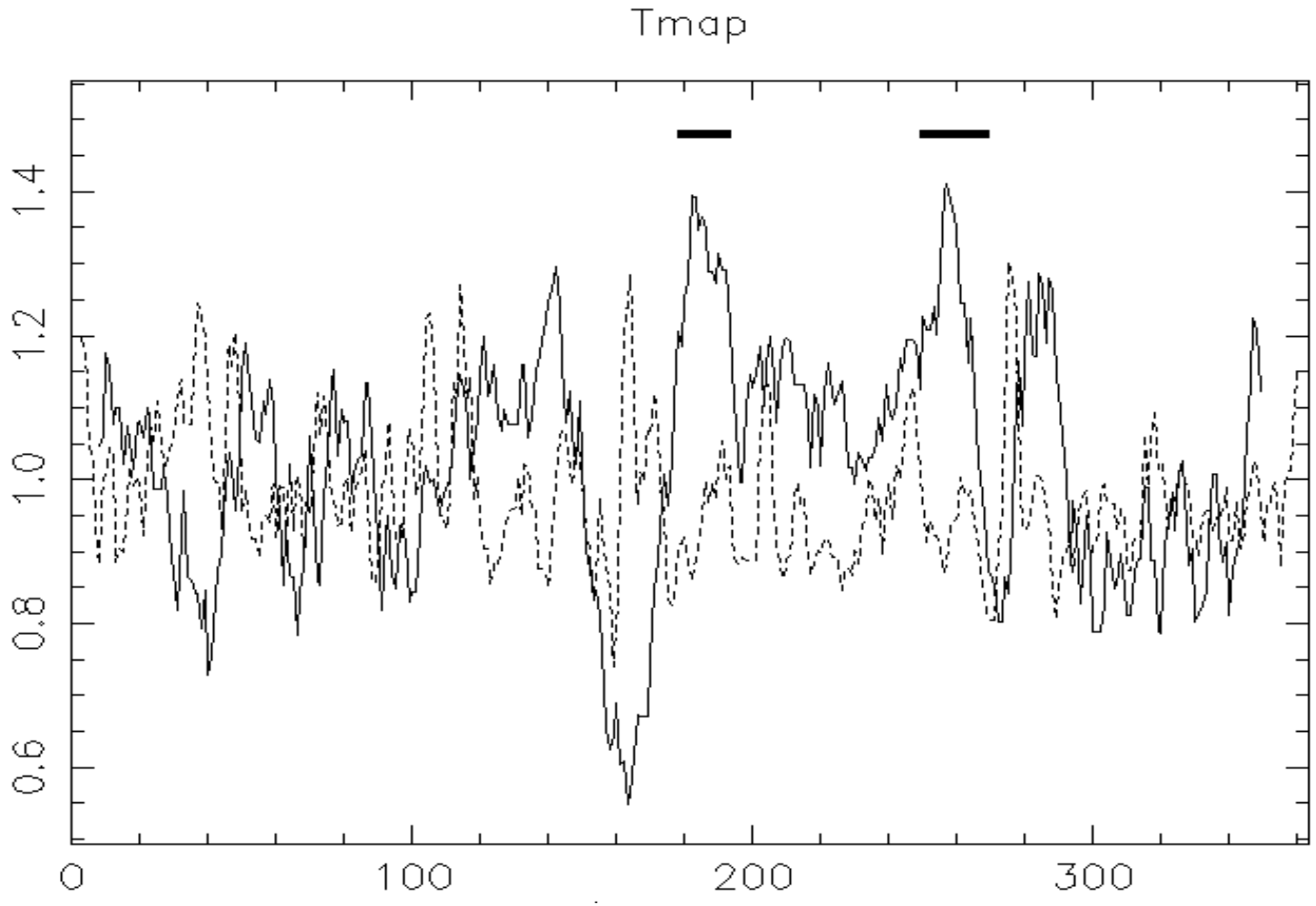
磷酸化位点预测（ NetPhos 2.0 ）和比较分析

野生	Pos	磷酸化位点区	Score	普通	Pos	磷酸化位点区	Score
Serine	9	NVVDSWGQA	0.975	Serine	9	NVVDSWGQA	0.975
	45	KQSLSYRPF	0.746		45	KQSLSYRPF	0.540
	50	YRPFSSVTV	0.809		214	ELAKSVGTE	0.665
	55	SVTVSAIAT	0.959		297	DVTKSSSEL	0.908
	214	ELA KSVGTE	0.665		345	AAFDSHKAA	0.658
	297	DVTKSSSEL	0.908				
	298	VTKSSSELG	0.973				
	345	AAFDSHKAA	0.658				
Tyrosine	23	QSLPYDSFN	0.596	Tyrosine	191	FAFEYLATA	0.507
	191	FAFEYLATA	0.507		316	DKTTYPKLL	0.665
	316	DKTTYPKLL	0.665				
				Threonine	197	ATATTGVPP	0.712

这些位点是否会影响酶分子的活性有待进一步研究

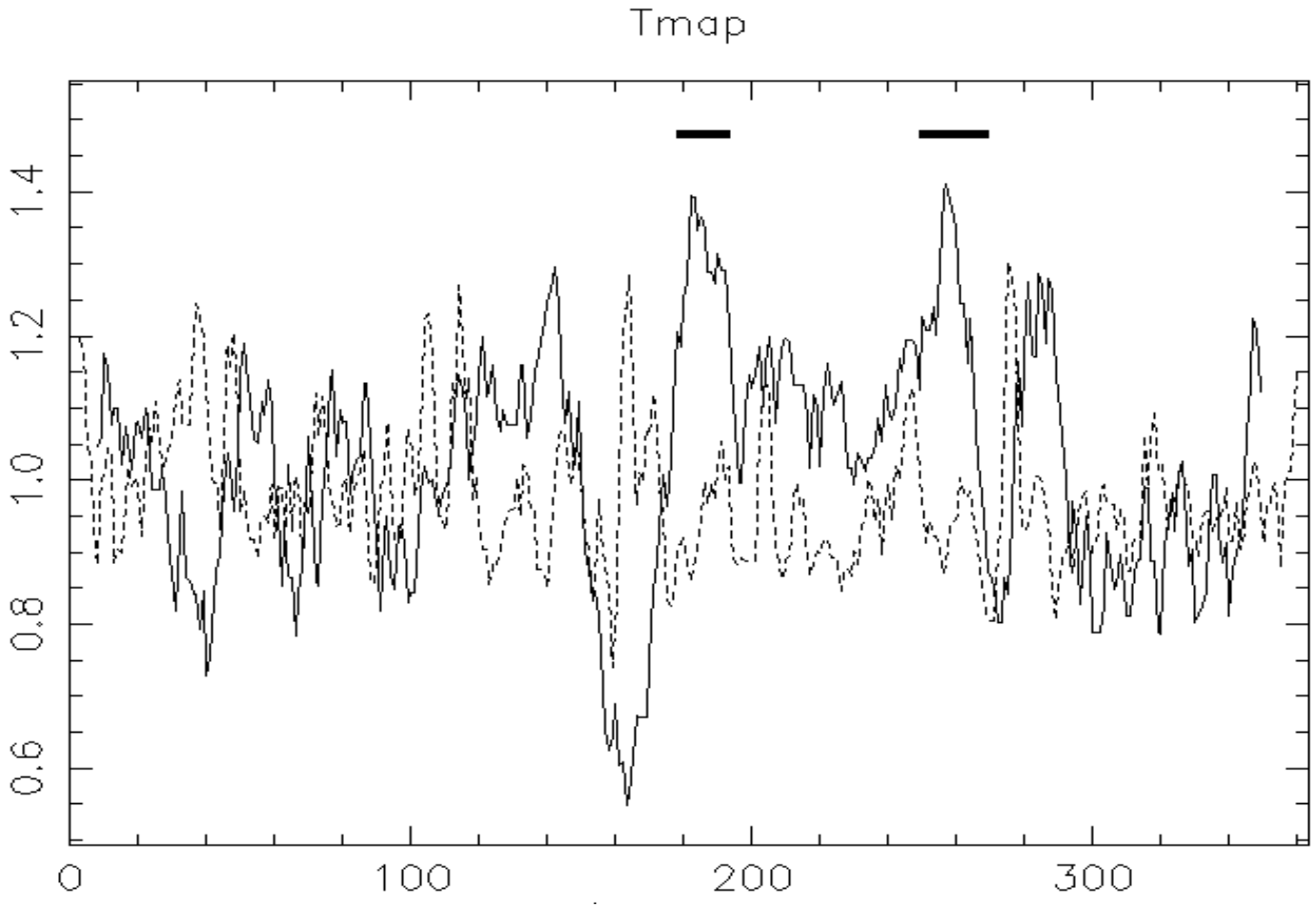
两序列氨基酸差异比较

疏水结构分析（Jemboss）



普通番茄：174-197　　1 EDVAVLAGDALLAFAFEYLATATT

245-273　　2 HIHKTAALLEASVVIGAILGGGADEEVDK

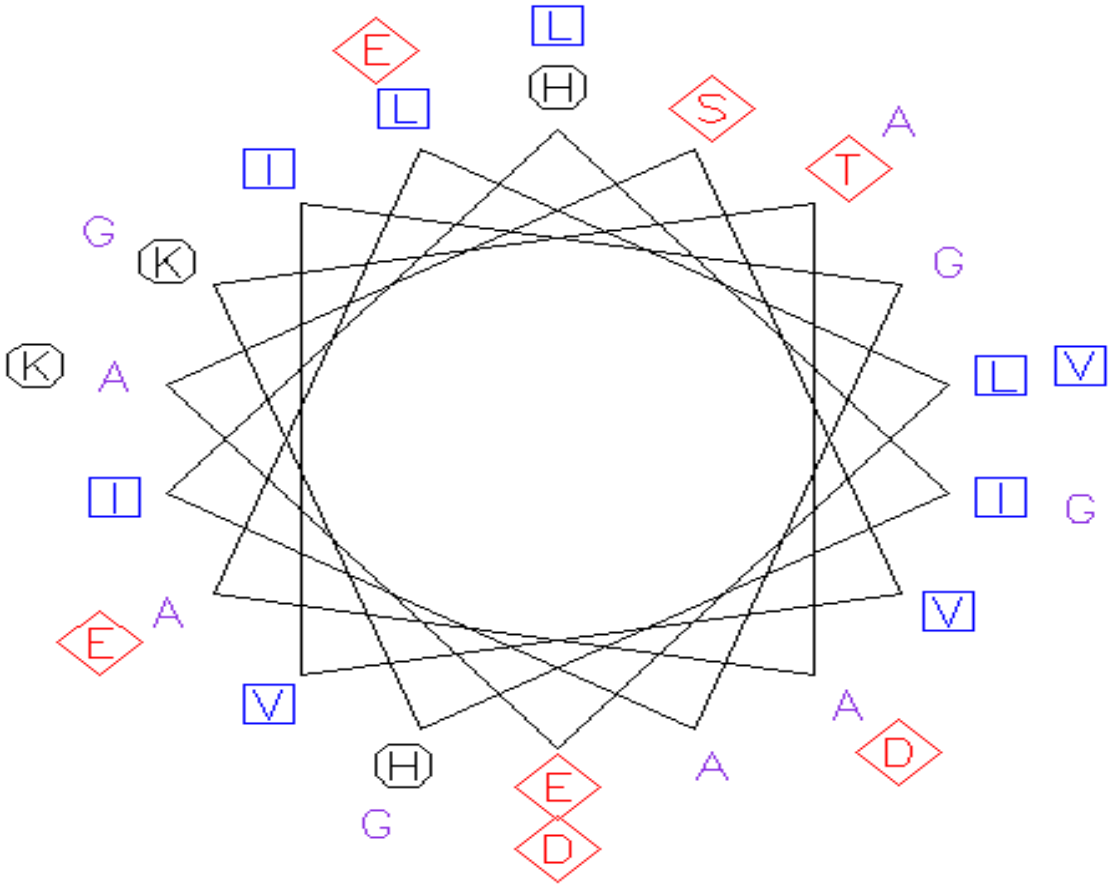


野生番茄 ： 174-198　　1 EDVAVLAGDALLAFAFEYLATATIG

245-273　　2 HIHKTAALLEASVVIGAILGGGADEEVDK

两序列氨基酸差异比较

Pepwheel分析



氨基酸BLAST

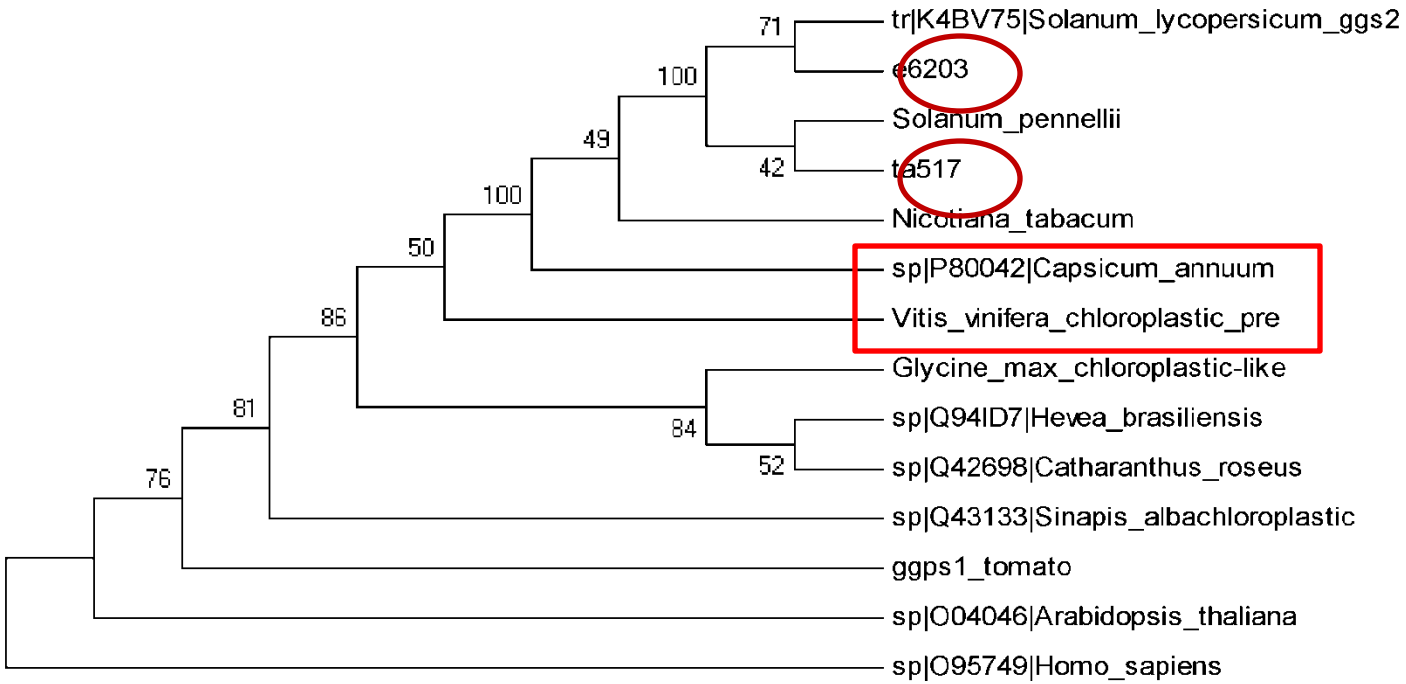
Uniprot protein Datebase

Color code for identity 0-100% =

Accession	Entry name	0Query hit363	0Match hit (sqrt scale)403	Name (Organism)
Query 2013010751T7MW552T				
Q1A7S9	Q1A7S9_SOLLC			Geranylgeranyl pyrophosphate synthase 2 (Solanum lycopersicum)
K4BV75	K4BV75_SOLLC			Uncharacterized protein (Solanum lycopersicum)
F1C933	F1C933_SOLPN			Geranylgeranyl pyrophosphate synthase 2 (Solanum pennellii)
B1N7F3	B1N7F3_NICAT			Geranylgeranyl pyrophosphate synthase (Nicotiana attenuata)
D3W9H8	D3W9H8_TOBAC			Geranylgeranyl diphosphate synthase (Nicotiana tabacum)
P80042	GGPPS_CAPAN			Geranylgeranyl pyrophosphate synthase... (Capsicum annuum)
D7F2D7	D7F2D7_SALMI			Geranylgeranyl diphosphate synthase (Salvia miltiorrhiza)
B7TCA9	B7TCA9_SALMI			Geranylgeranyl diphosphate synthase (Salvia miltiorrhiza)
D6BCR2	D6BCR2_CORAV			Geranylgeranyl diphosphate synthase (Corylus avellana)
O81099	O81099_HELAN			Geranylgeranyl pyrophosphate synthase (Helianthus annuus)
F6H042	F6H042_VITVI			Putative uncharacterized protein (Vitis vinifera)
C4NAM8	C4NAM8_HUMLU			Geranyl diphosphate synthase large subunit (Humulus lupulus)

GGPS 系统发育分析

Mega软件



多毛番茄和pennellii 聚在一块，GGPS并与定位于叶绿体中GGPS 聚在一起

Motif分析

Pfam：

PS00723 **POLYPRENYL_SYNTHASE_1** *Polyprenyl synthases signature 1* :

148 - 164: LIhDDlpcmDdddlRRG

[LIVM](2)-x-D-D-x(2,4)-D-x(4)-R-R-[GH]

PS00444 POLYPRENYL_SYNTHASE_2 Polyprenyl synthases signature 2 :

281 - 293: IGllFQVvDDIlD

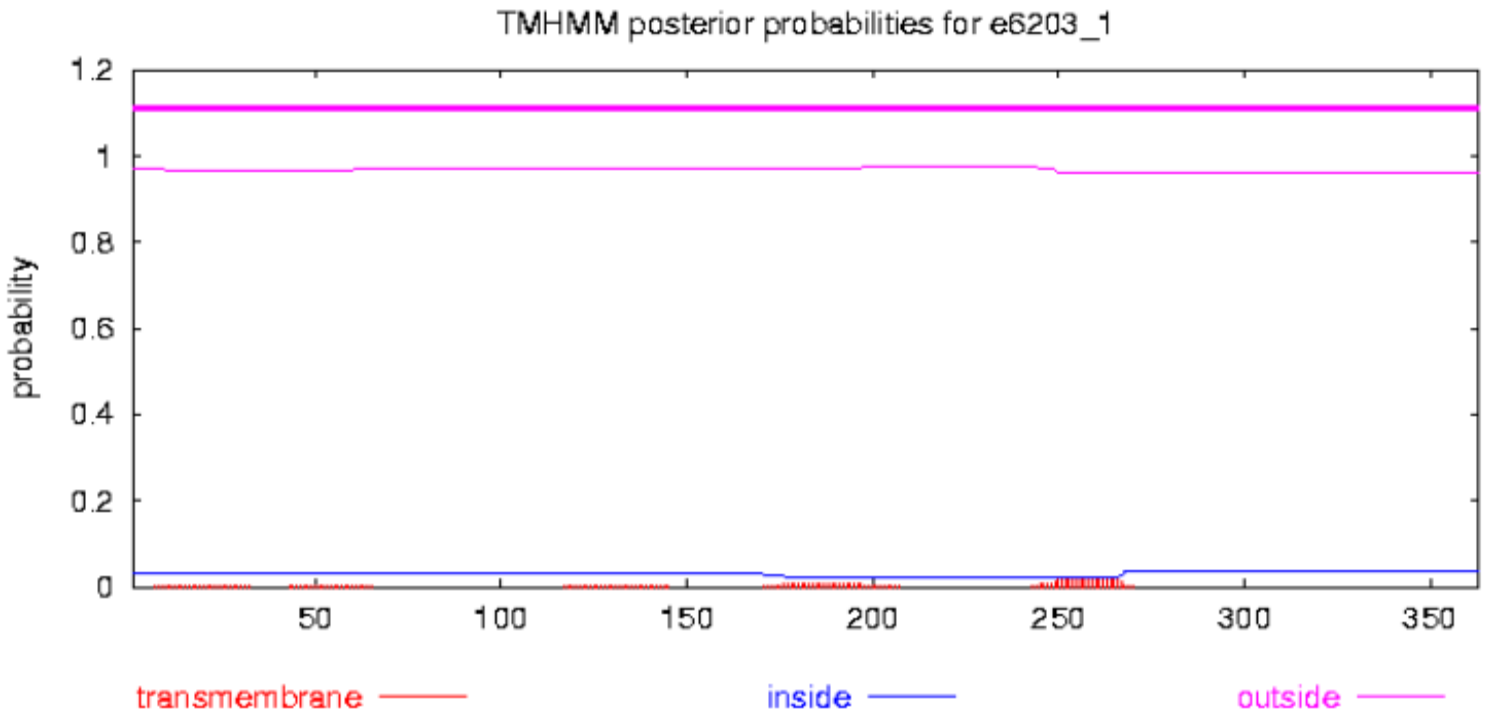
[LIVMFY]-G-x(2)-[FYL]-Q-[LIVM]-x-D-D-[LIVMFY]-x-[DNG]

95- 363：异戊烯基转移酶,

存在四个差异位点，分别是130、197、279、360

跨膜结构的预测

TMHMM Server v.2.0 prediction of transmembrance



野生番茄 **Number of predicted TMHs: 0**
普通番茄 **Number of predicted TMHs: 0**

信号肽的预测及亚细胞定位

TargetP 1.1 Server - prediction results

Name	Len	cTP	mTP	SP	other	Loc	RC	TPlen
ta517_1	363	0.166	0.081	0.062	0.299	_	5	-
cutoff		0.000	0.000	0.000	0.000			

SOSUI软件预测，为水溶性蛋白，不含信号肽

信号肽的预测及亚细胞定位

http://www.csbio.sjtu.edu.cn/bioinf/plant/

普通番茄

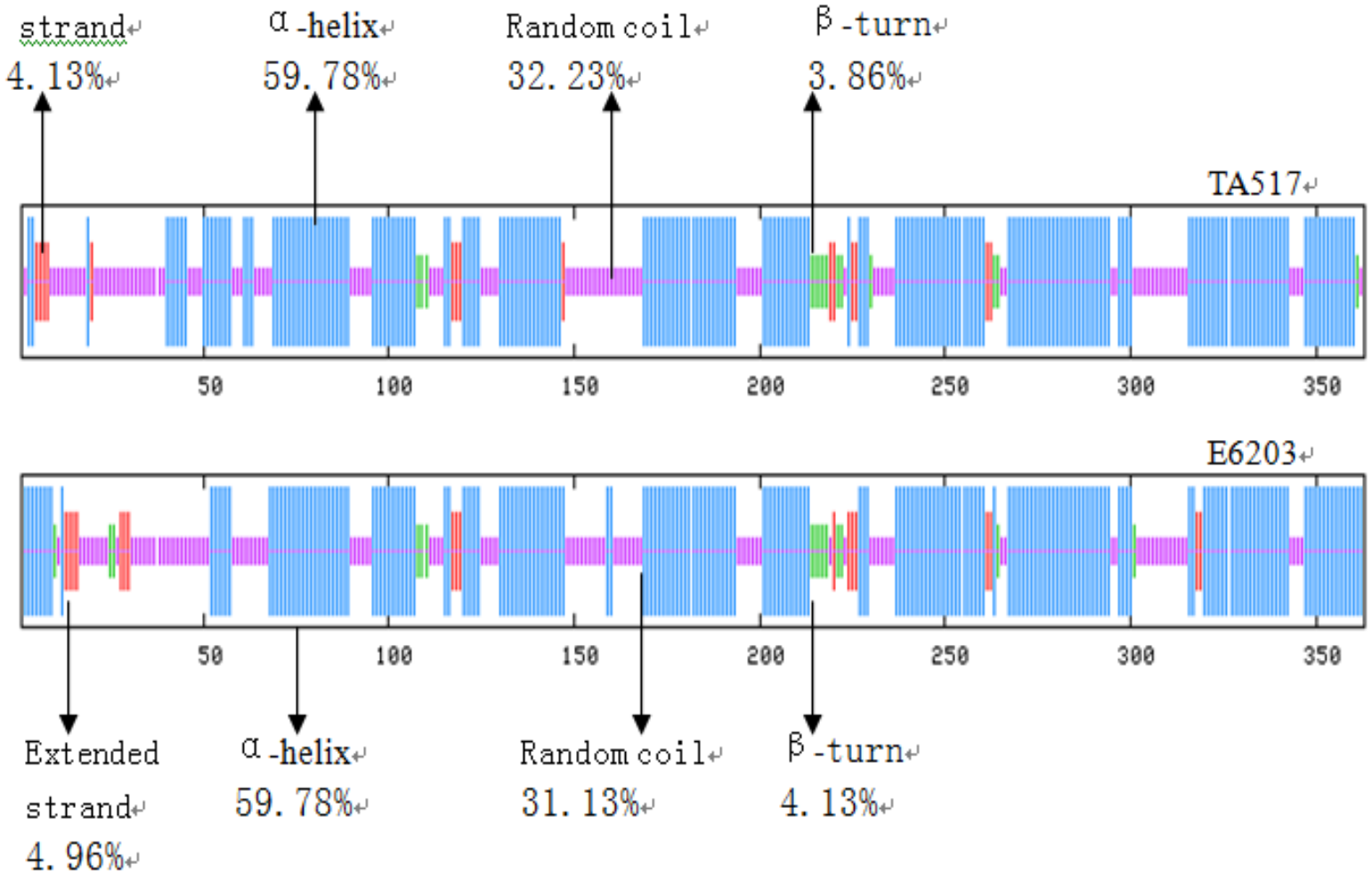
Accession number	Predicted location(s)	Prediction approach
P11111	Chloroplast	By fusing Gene Ontology (GO) information

野生番茄

Accession number	Predicted location(s)	Prediction approach
P11112	Chloroplast	By fusing PseAA composition

蛋白质二级结构分析

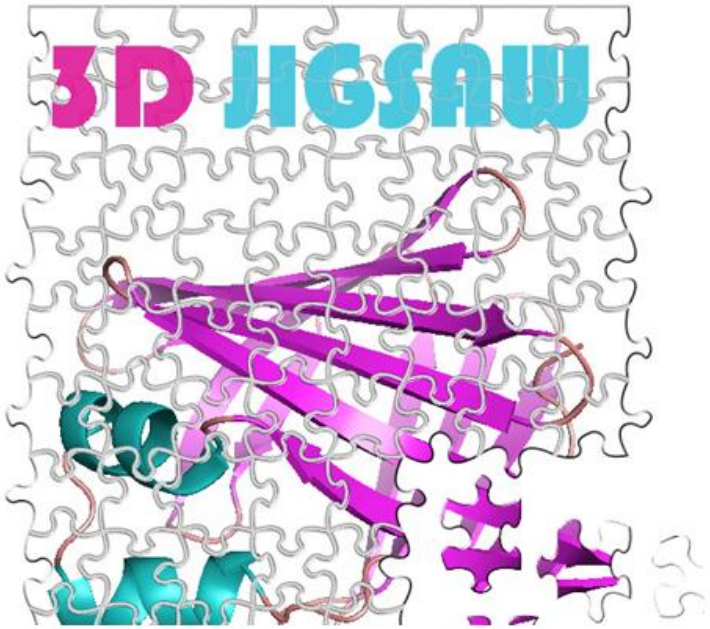
SMOPA程序，对推导氨基酸序列进行二级结构预测



蛋白质三级结构分析

3D-JIGSAW

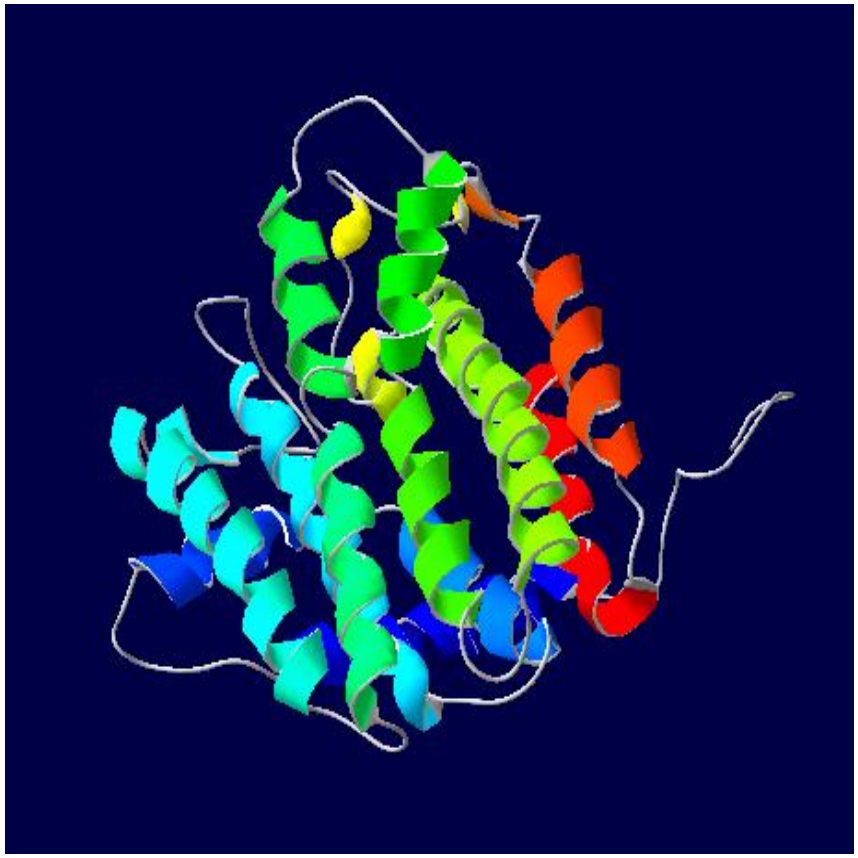
Comparative Modelling



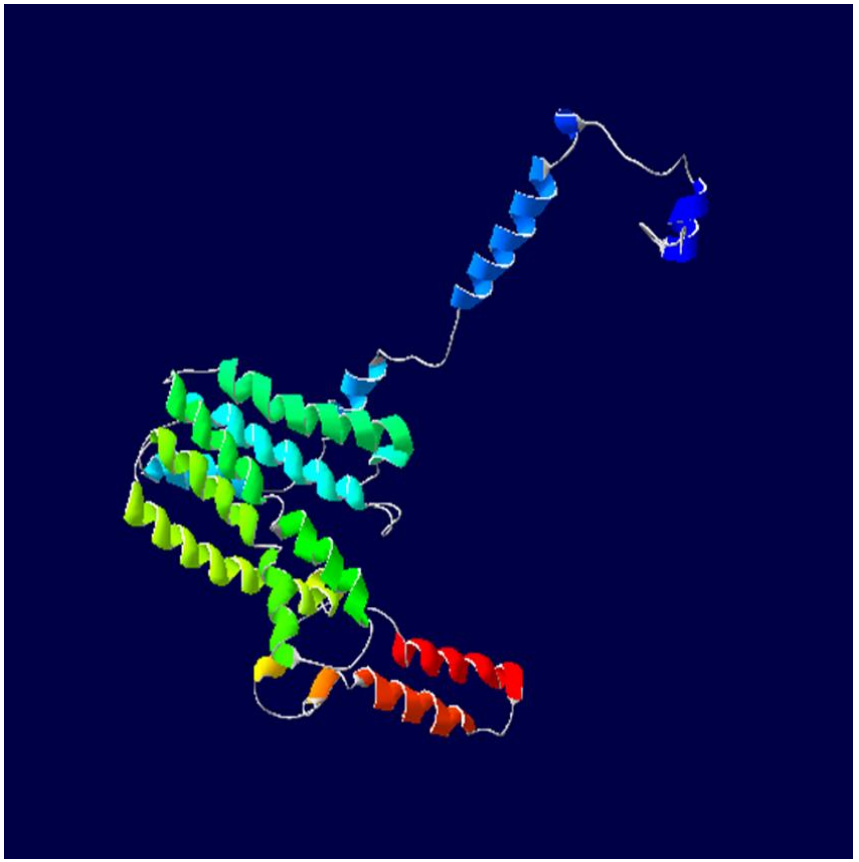
Database updates: 3D-JIGSAW (31 Dec) Domain Fishing (31 Dec) **Server updates**

蛋白质三级结构分析

普通番茄



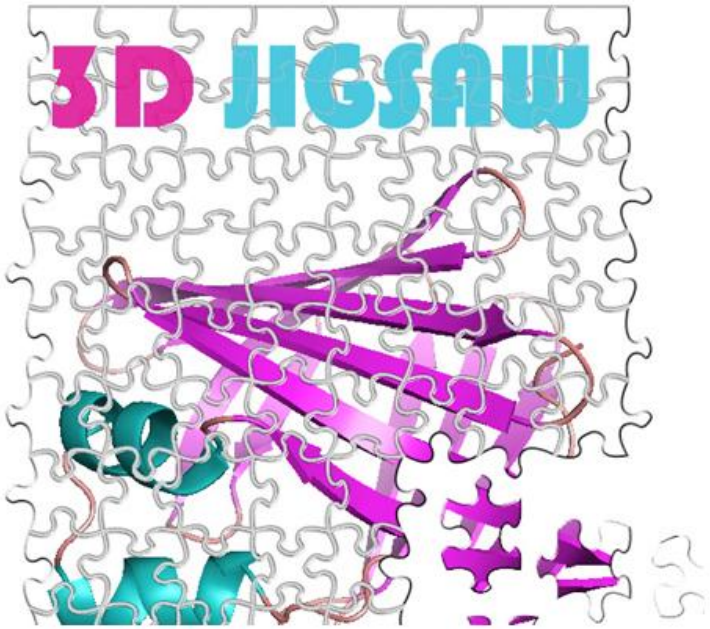
野生番茄



蛋白质三级结构分析

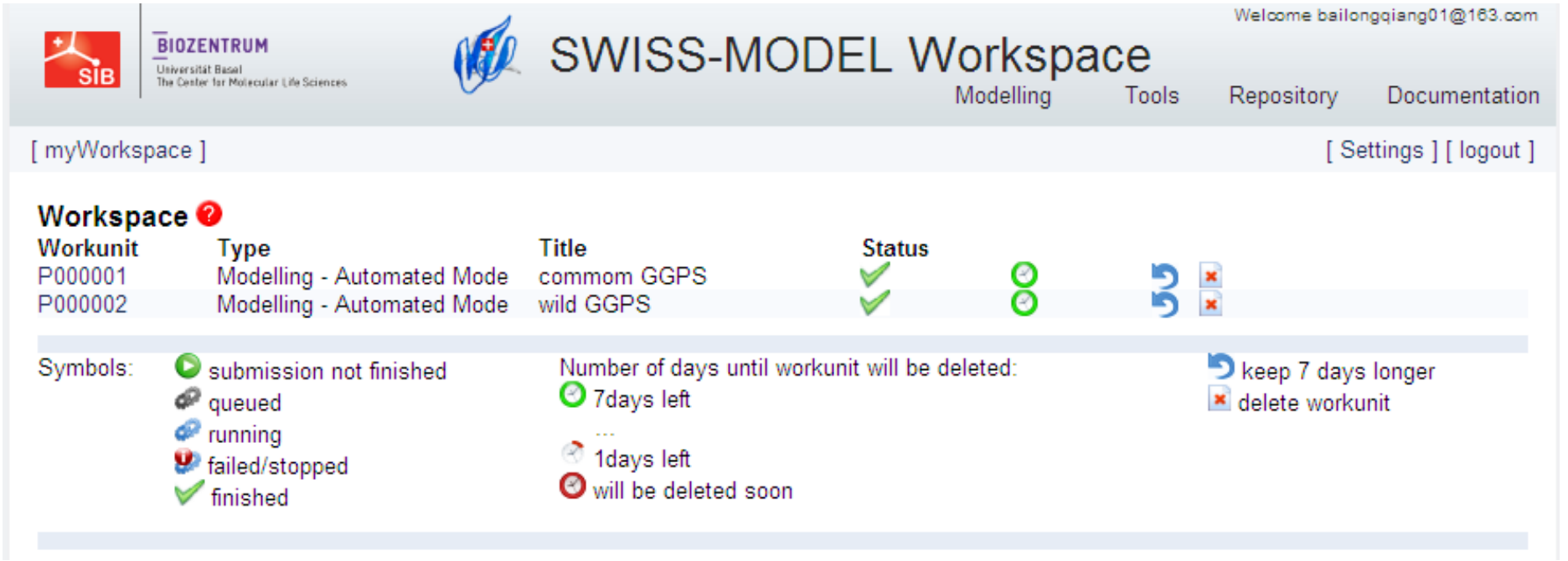
3D-JIGSAW

Comparative Modelling



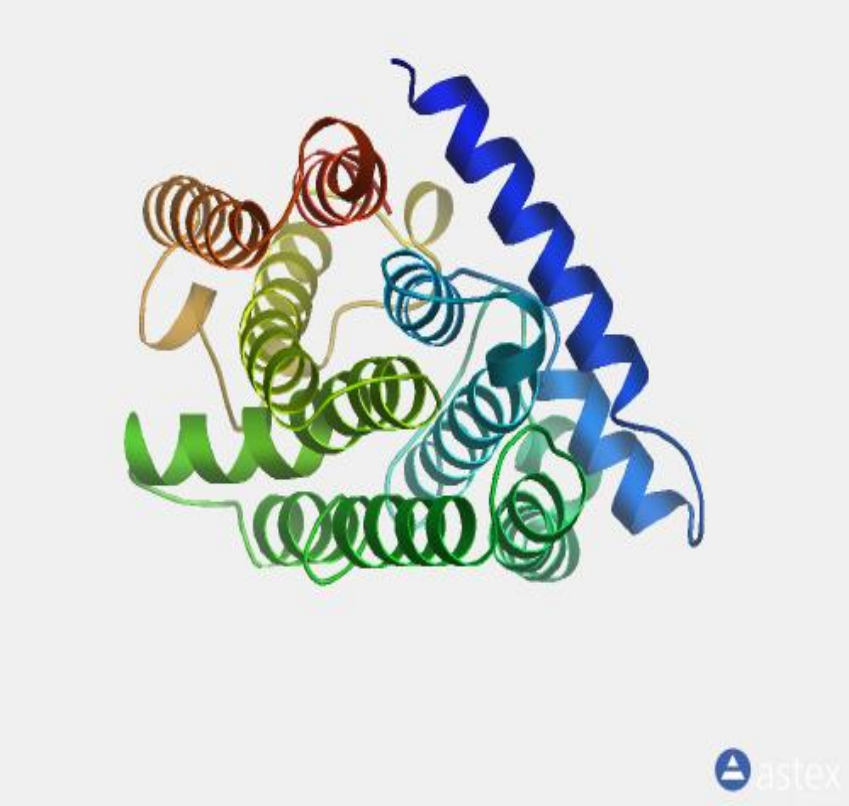
Database updates: 3D-JIGSAW (31 Dec) Domain Fishing (31 Dec) **Server updates**

蛋白质三级结构分析

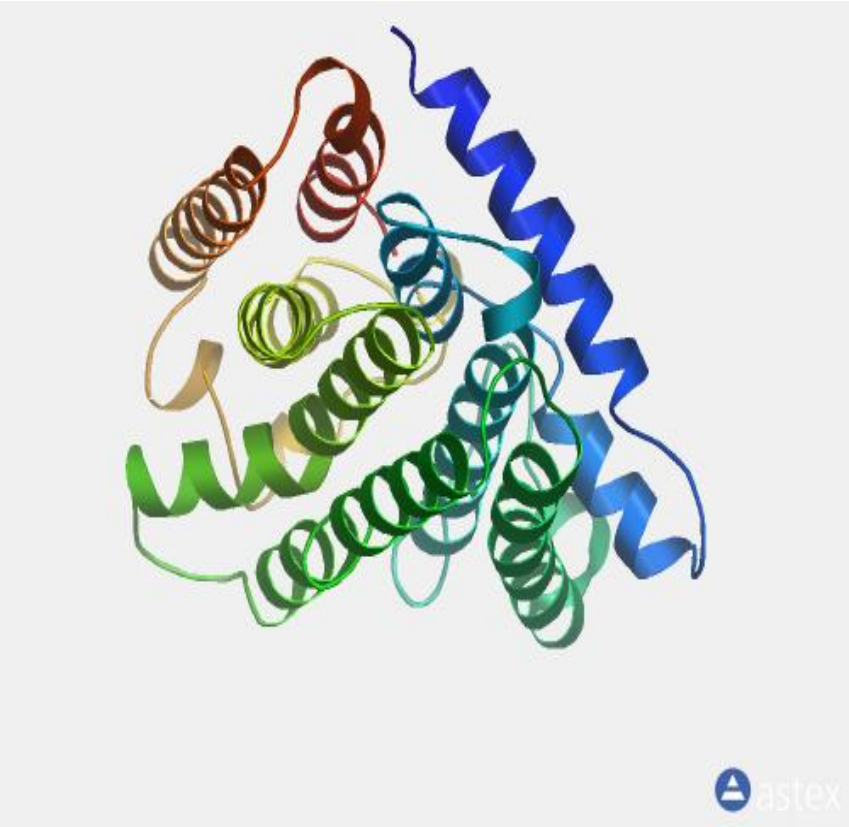

BIOZENTRUM
Universität Basel
The Center for Molecular Life Sciences
SWISS-MODEL Workspace
Welcome bailongqiang01@163.com
Modelling
Tools
Repository
Documentation
[myWorkspace]
[Settings] [logout]
Workspace
Workunit
Type
Title
Status
P000001
Modelling - Automated Mode
commom GGPS
P000002
Modelling - Automated Mode
wild GGPS
Symbols:
submission not finished
queued
running
failed/stopped
finished
Number of days until workunit will be deleted:
7days left
...
1days left
will be deleted soon
keep 7 days longer
delete workunit

蛋白质三级结构分析

普通番茄



野生番茄



影响功能位点的预测

亚细胞定位：结合生物信息学信息，下一步利用分子生物学手段进行验证

酶活性比较：酶分子活性的体外检测，和转基因验证

功能位点分析：结合生物信息学预测的功能位点，进行基因的定点突变。

谢谢大家！

希望大家批评指正！