

实用生物信息技术期末总结汇报

硒代半胱氨酸甲基转移酶（SMT）的序列、结构和功能分析

**structure and function analysis of selenocysteine methyltransferase
(SMT)**

2021年5月10日

报告人：张小雪

邮箱：375030749@qq.com

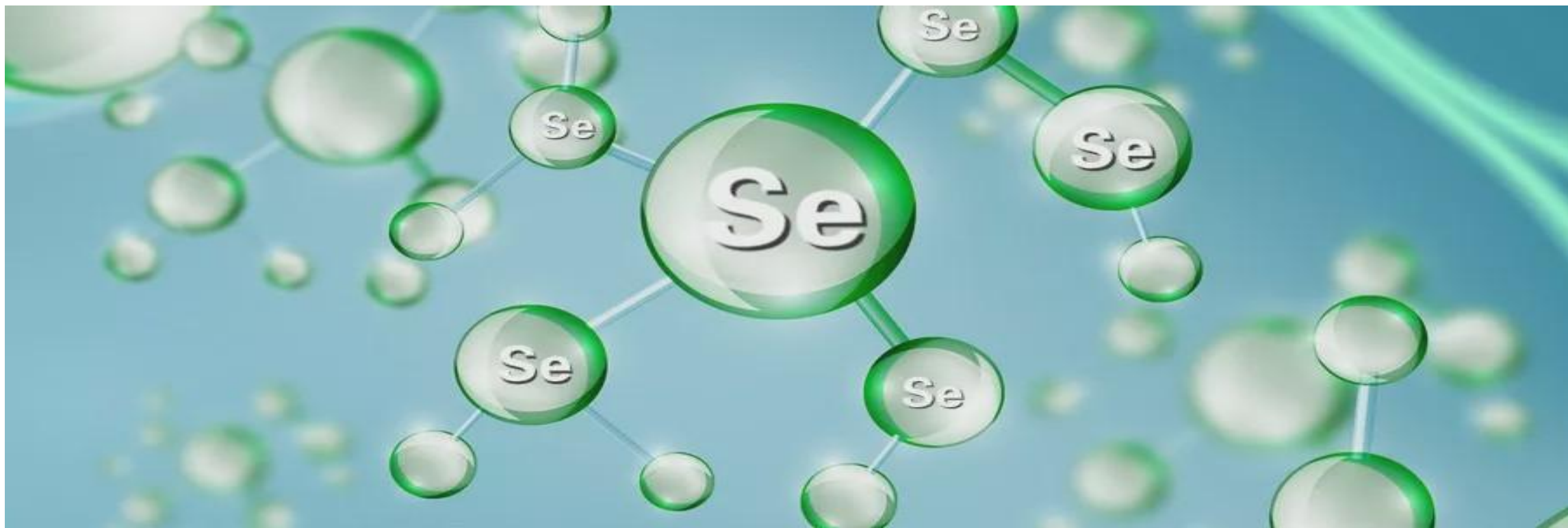
小组成员

编号	姓名	研究所	导师	课题
3G04A	于东洋	油料所	闫立英	花生白绢病致病机理
3G04B	张小雪	油料所	王新发	油菜硒高效理念相关QTL位点和功能基因的鉴定
3G04C	陶顺玉	油料所	张慧芳	花生含油量主效位点的图位克隆和候选基因的鉴定

目录

- 1.背景
- 2.序列分析
- 3.结构分析
- 4.功能分析

1.背景



硒(Se) 是人体最重要的微量元素，具有抗氧化、改善免疫力、提高抵抗力等功能^[1]，被美国医学家称为至今发现最强烈也最具潜力的抗癌营养素^[2]。

中国土壤硒元素含量分布图



据地质学家考证，我国 72 % 的地区属于缺硒地区，2 / 3 的人口存在不同程度的硒摄入量不足。

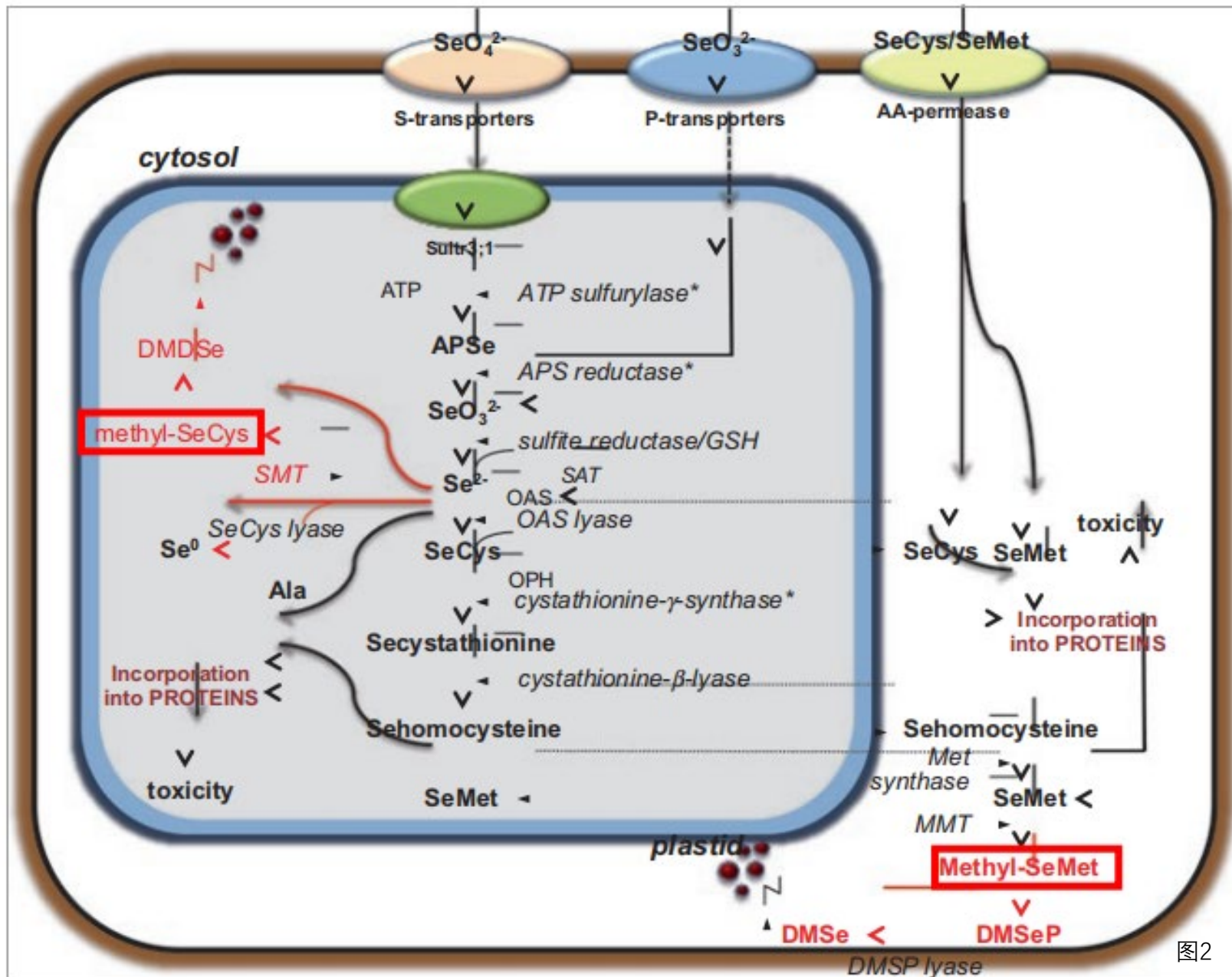


图2: 硒在植物中的代谢途径

硒代半胱氨酸甲基转移酶(SMT)是植物硒代谢途径中的一个关键酶。在植物硒代谢过程中, SMT催化硒代半胱氨酸甲基化形成甲基硒代半胱氨酸, 从而减少硒对蛋白质的掺入, 解除硒对植物的毒害^[3]。

图2



黄芪



拟南芥



水稻



烟草



青花菜



茶树

图3

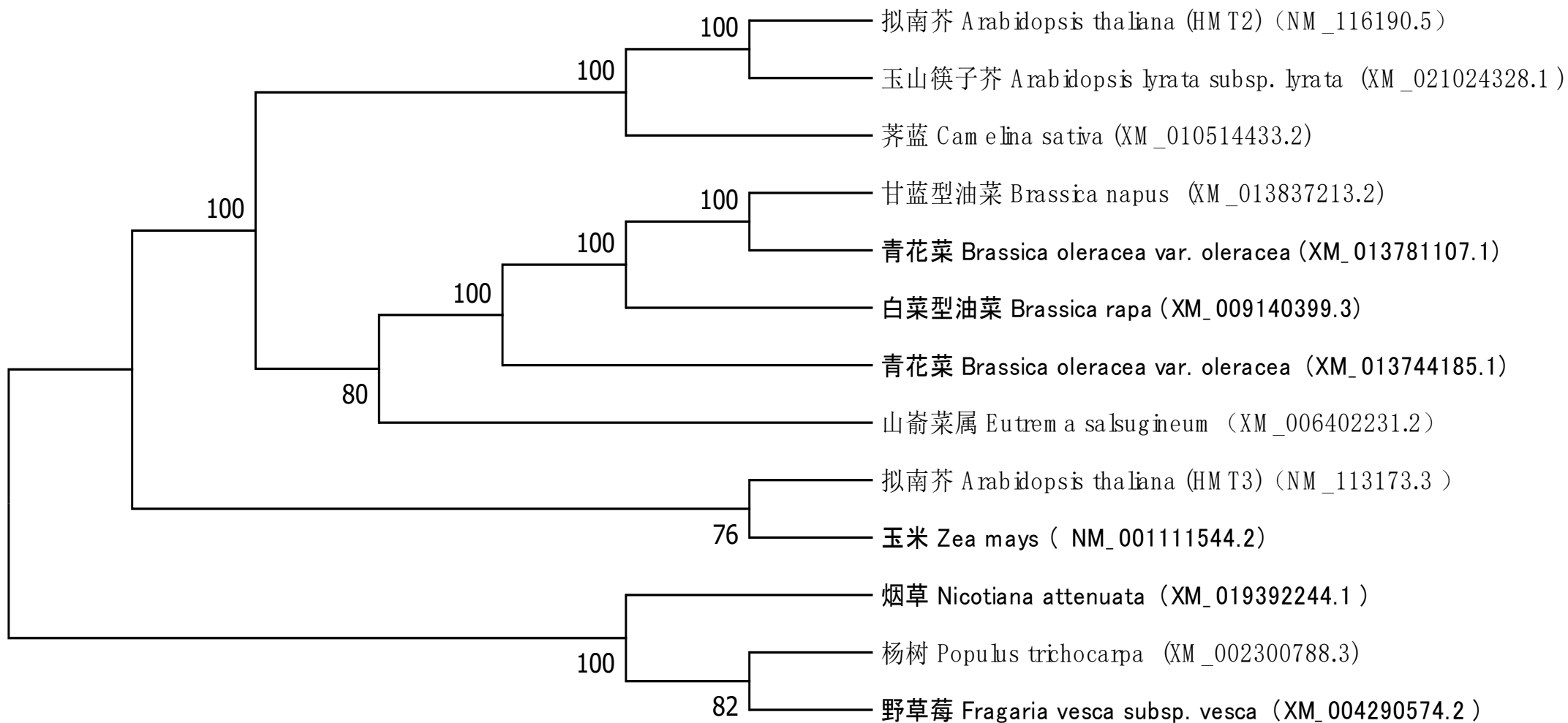
2.序列分析

SMT基因的系统发育

物种	拉丁文名	登录号
拟南芥	<i>Arabidopsis thaliana</i> (HMT2)	NM_116190.5
拟南芥	<i>Arabidopsis thaliana</i> (HMT3)	NM_113173.3
玉米	<i>Zea mays</i>	NM_001111544.2
山嵛菜属	<i>Eutrema salsugineum</i>	XM_006402231.2
青花菜	<i>Brassica oleracea</i> var. <i>oleracea</i>	XM_013744185.1
青花菜	<i>Brassica oleracea</i> var. <i>oleracea</i>	XM_013781107.1
白菜型油菜	<i>Brassica rapa</i>	XM_009140399.3
甘蓝型油菜	<i>Brassica napus</i>	XM_013837213.2
芥蓝	<i>Camelina sativa</i>	XM_010514433.2
杨树	<i>Populus trichocarpa</i>	XM_002300788.3
烟草	<i>Nicotiana attenuata</i>	XM_019392244.1
玉山筷子芥	<i>Arabidopsis lyrata</i> subsp. <i>lyrata</i>	XM_021024328.1
野草莓	<i>Fragaria vesca</i> subsp. <i>vesca</i>	XM_004290574.2

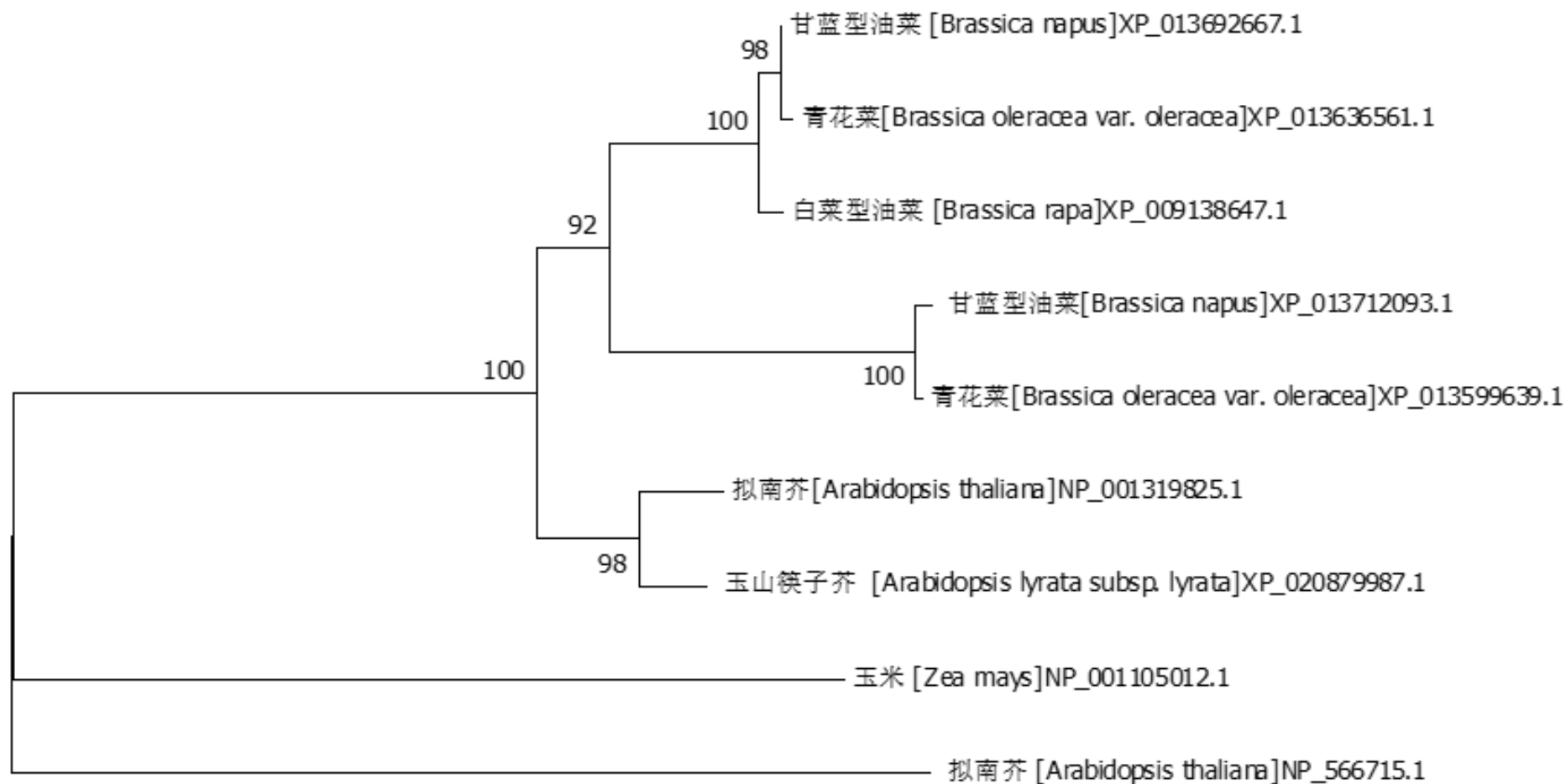
2.序列分析

SMT基因的系统发育



2.序列分析

SMT基因编码产物的系统发育



0.050

2. 序列分析 SMT编码产理化性质

Number of amino acids: 346

Molecular weight: 37862.93

Theoretical pI: 5.41

Carbon	C	1681
Hydrogen	H	2656
Nitrogen	N	448
Oxygen	O	523
Sulfur	S	11

Formula: C₁₆₈₁H₂₆₅₆N₄₄₈O₅₂₃S₁₁

Total number of atoms: 5319

Ala (A)	30	8.7%
Arg (R)	15	4.3%
Asn (N)	9	2.6%
Asp (D)	20	5.8%
Cys (C)	8	2.3%
Gln (Q)	4	1.2%
Glu (E)	29	8.4%
Gly (G)	27	7.8%
His (H)	8	2.3%
Ile (I)	25	7.2%
Leu (L)	30	8.7%
Lys (K)	24	6.9%
Met (M)	3	0.9%
Phe (F)	11	3.2%
Pro (P)	11	3.2%
Ser (S)	33	9.5%
Thr (T)	19	5.5%
Trp (W)	4	1.2%
Tyr (Y)	13	3.8%
Val (V)	23	6.6%

Ala (A) 8.25	Gln (Q) 3.93	Leu (L) 9.65	Ser (S) 6.63
Arg (R) 5.53	Glu (E) 6.72	Lys (K) 5.80	Thr (T) 5.35
Asn (N) 4.06	Gly (G) 7.07	Met (M) 2.41	Trp (W) 1.10
Asp (D) 5.46	His (H) 2.27	Phe (F) 3.86	Tyr (Y) 2.92
Cys (C) 1.38	Ile (I) 5.91	Pro (P) 4.73	Val (V) 6.86

半胱氨酸含量高

2.序列分析 SMT基因编码产理化性质

Instability index:

The instability index (II) is computed to be 38.77
This classifies the protein as stable.

不稳定指数38.77
(小于40)

Aliphatic index: 89.94

脂肪指数89.94
总平均亲水性-0.170

Grand average of hydropathicity (GRAVY): -0.170

3. 结构分析

SMT基因编码产物信号肽分析

SignalP-5.0

Protein type	Signal Peptide (Sec/SPI)	Other
Likelihood	0.0013	0.9987

TargetP-2.0

Protein type	Other	Signal peptide	Mitochondrial transfer peptide	Chloroplast transfer peptide	Thylakoid luminal transfer peptide
Likelihood	0.9998	0.0002	0	0	0

3. 结构分析

Predictor: MultiLoc2-HighRes

Origin: plant

sp|Q4VVK0|SMTA_BRAOT cytoplasmic: 0.96 nuclear: 0.02 peroxisomal: 0.02 mitochondrial: 0.0 chloroplast: 0.0 Golgi apparatus: 0.0 extracellular: 0.0 plasma membrane: 0.0 ER: 0.0 vacuolar: 0.0

根据以上判断SMT定位于细胞质基质为亲水蛋白

3. 结构分析

SMT基因编码二级结构分析

HNNC :

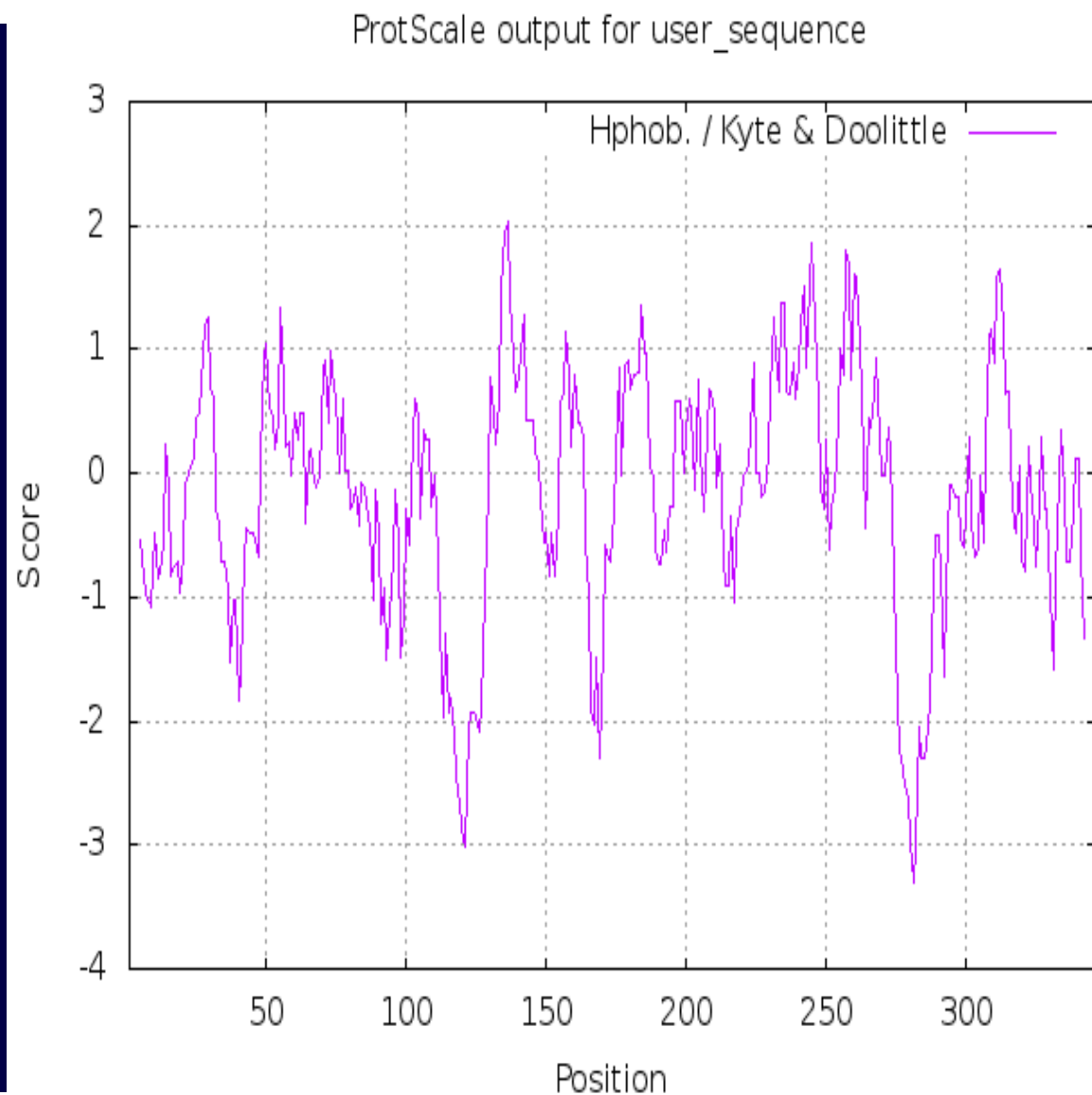
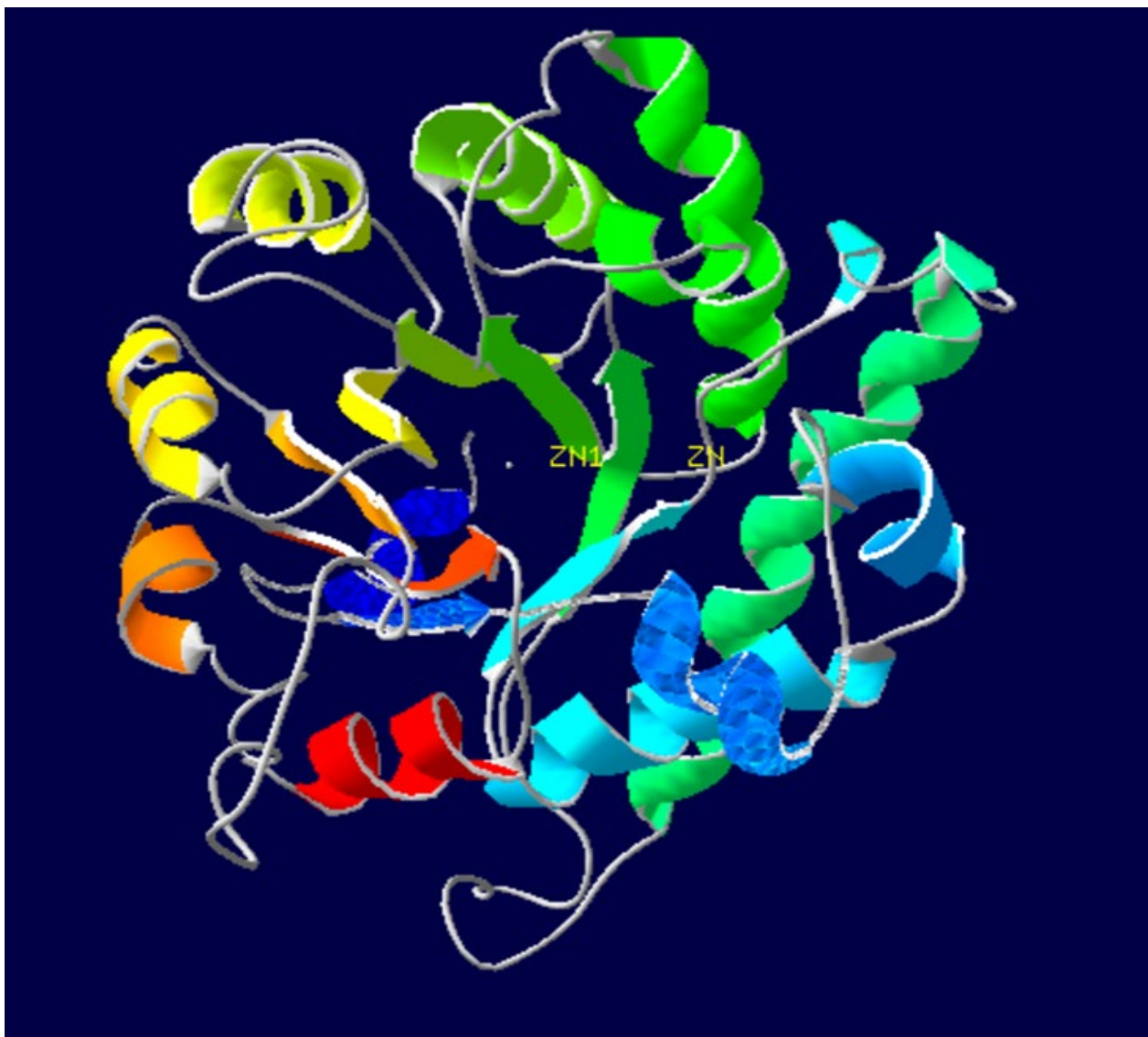
Alpha helix	(Hh)	:	160	is	46.24%
3 ₁₀ helix	(Gg)	:	0	is	0.00%
Pi helix	(Ii)	:	0	is	0.00%
Beta bridge	(Bb)	:	0	is	0.00%
Extended strand	(Ee)	:	59	is	17.05%
Beta turn	(Tt)	:	0	is	0.00%
Bend region	(Ss)	:	0	is	0.00%
Random coil	(Cc)	:	127	is	36.71%
Ambiguous states (?)		:	0	is	0.00%
Other states		:	0	is	0.00%

α -螺旋 46.24%

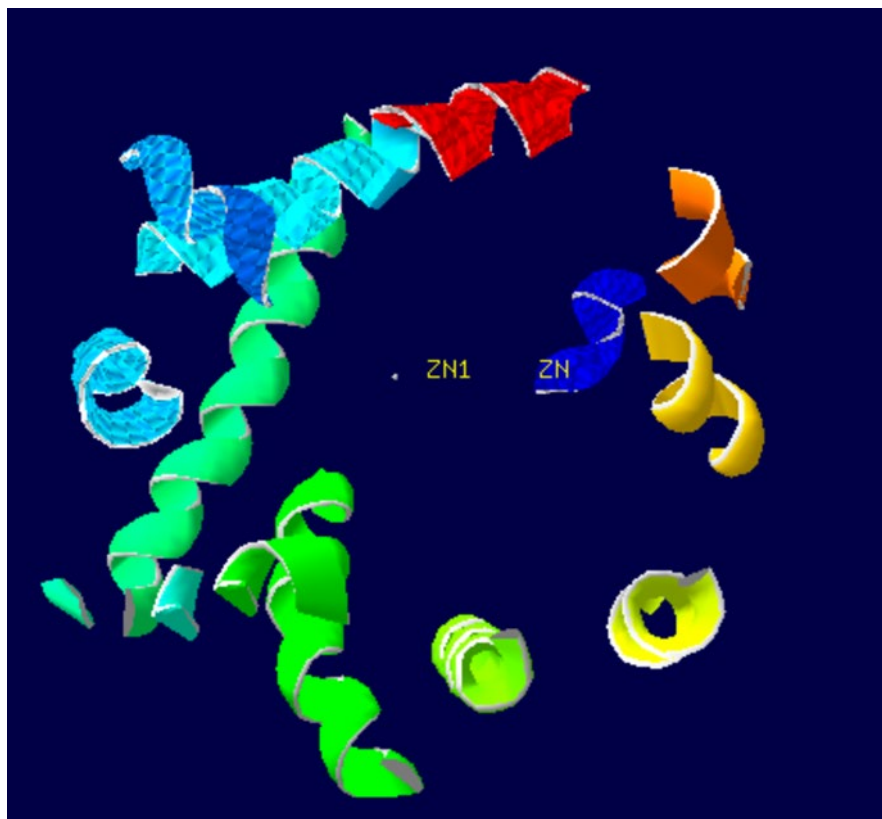
延伸链 17.05%

无规则卷曲 36.71%

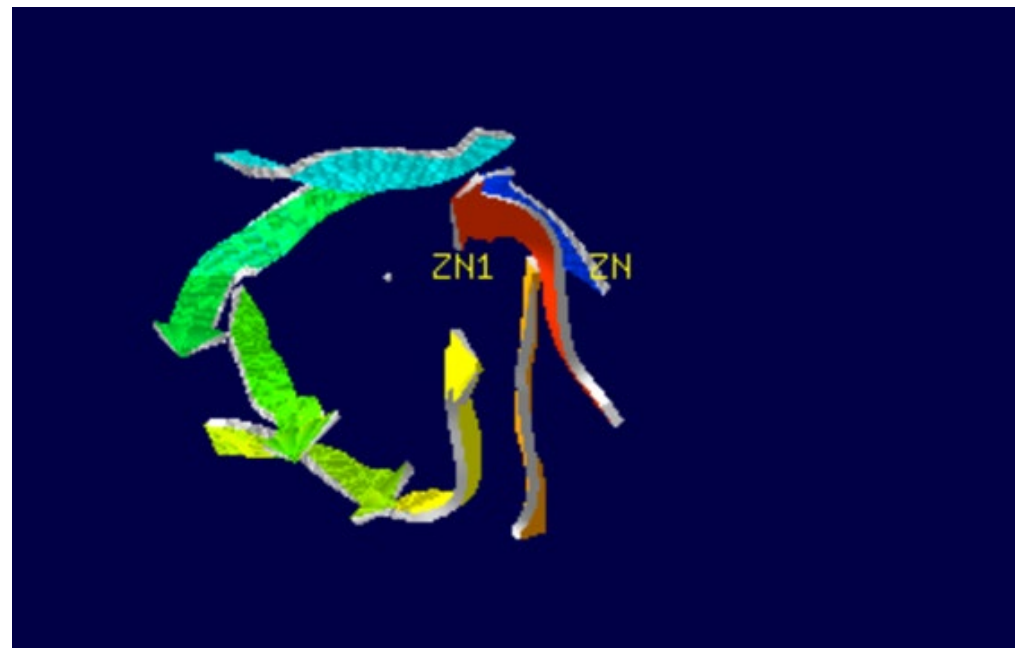
3. 结构分析 SMT基因编码三级结构分析



3. 结构分析 SMT基因编码三级结构分析

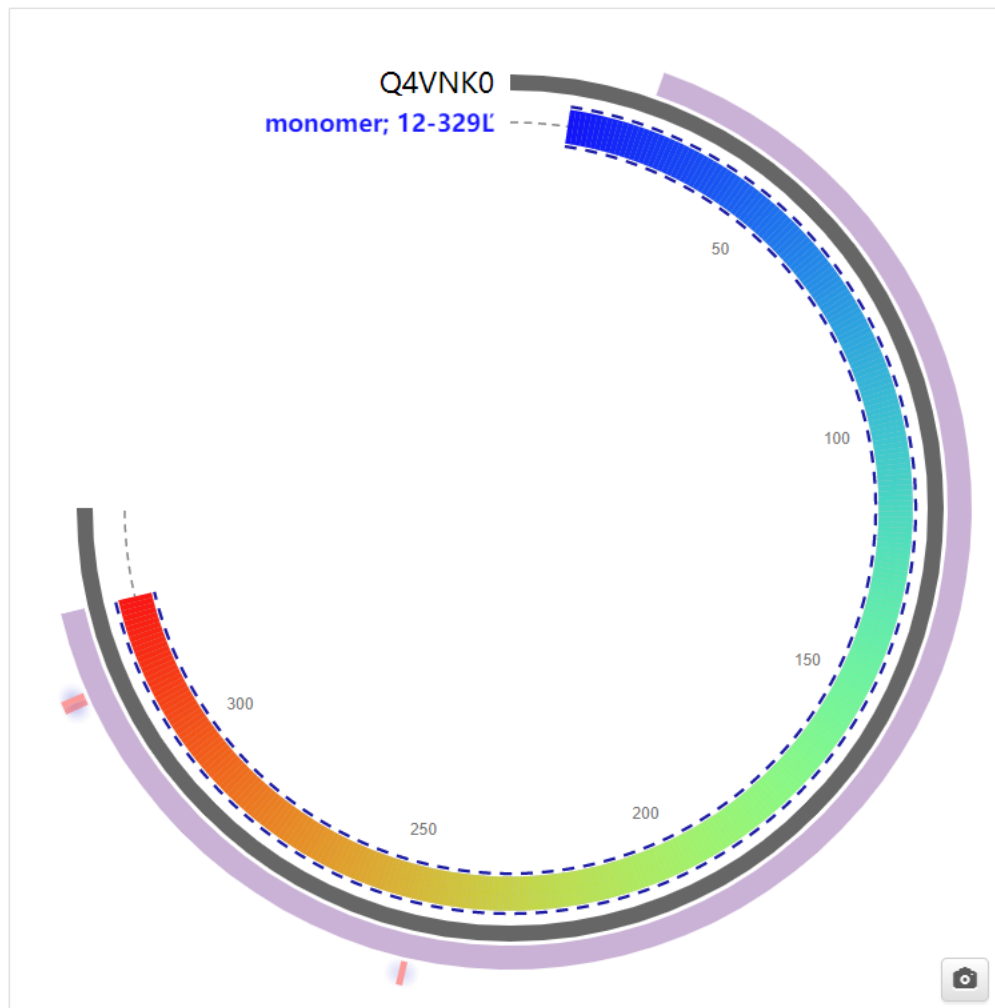


α -螺旋占49.8%
无规则卷曲占36.7%



延伸链占13.5%

3. 结构分析



Homology Model

Template: **5dmm.1.A** "Crystal Structure of the Homocysteine Methyltransferase MmuM from *Escherichia coli*, Metallated form" **Q47690**

SMTL Version: 2020-07-22

Seq Identity: 42.39%

Seq Similarity: 0.40

1 x ZINC ION

Coordinates:

Model Quality Estimate

QMEAN		-1.92
C β		-1.65
All Atom		-1.41
solvation		-0.11
torsion		-1.57

根据大肠杆菌半胱氨酸甲基转移酶为模板构建的

3. 结构分析

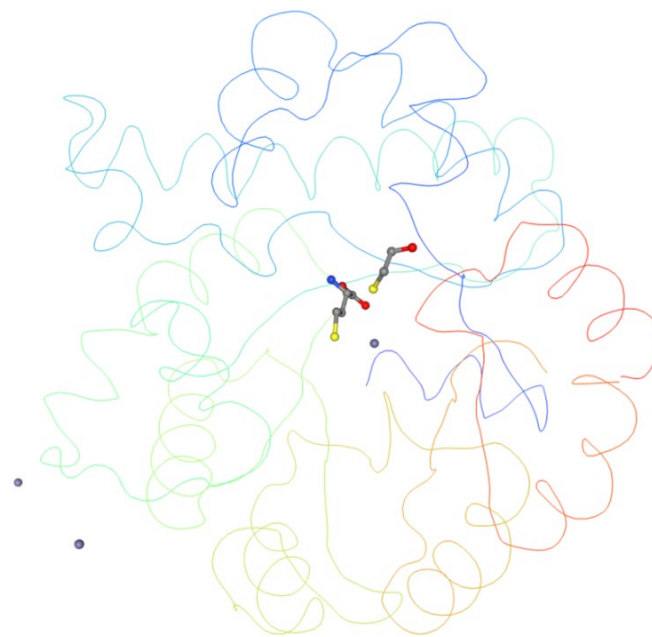
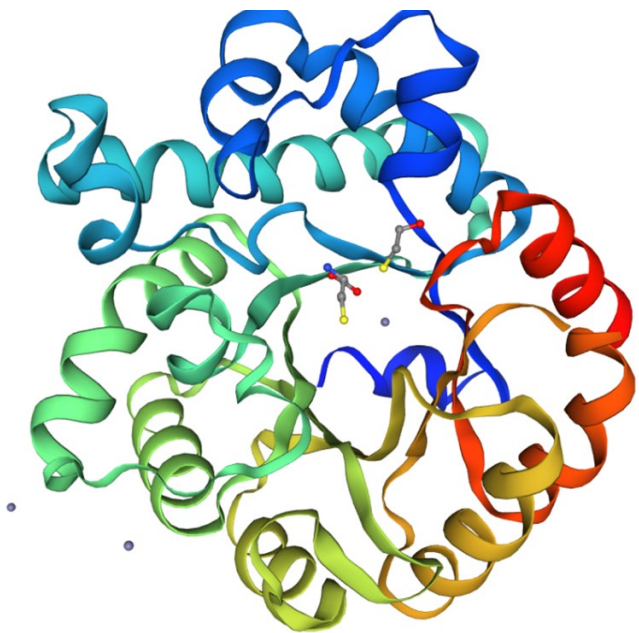
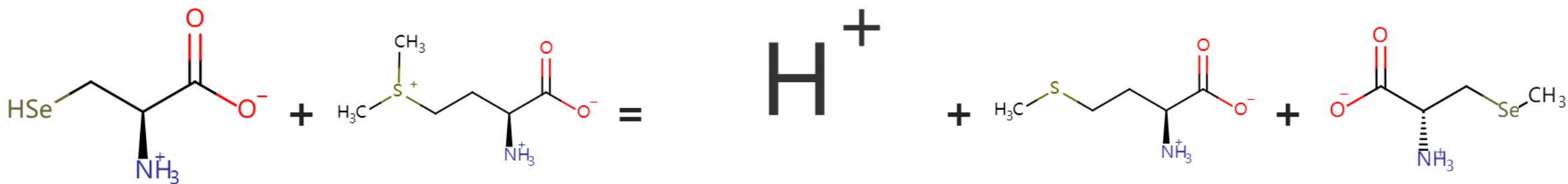
L-selenocysteine

S-methyl-L-methionine

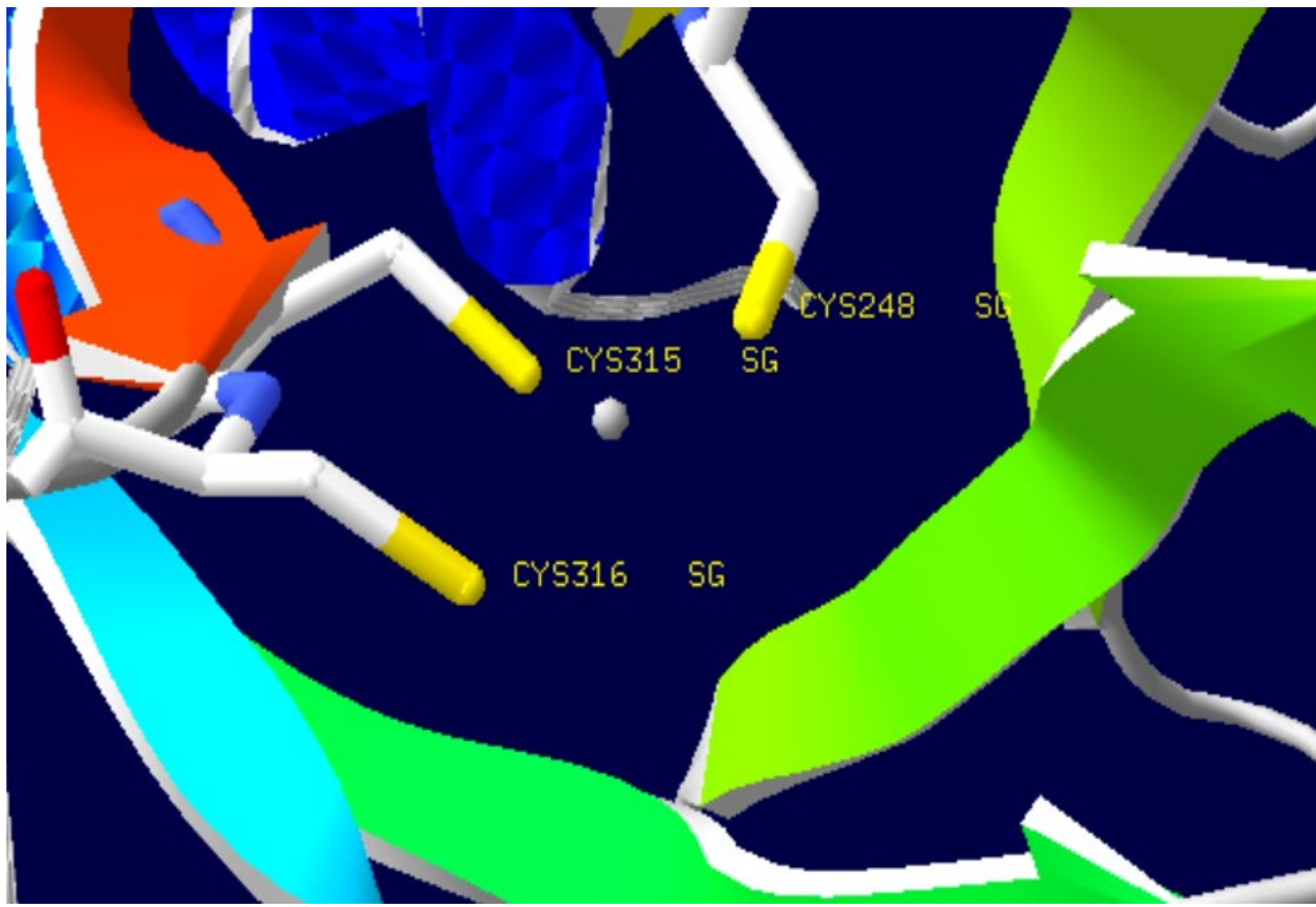
H⁺

L-methionine

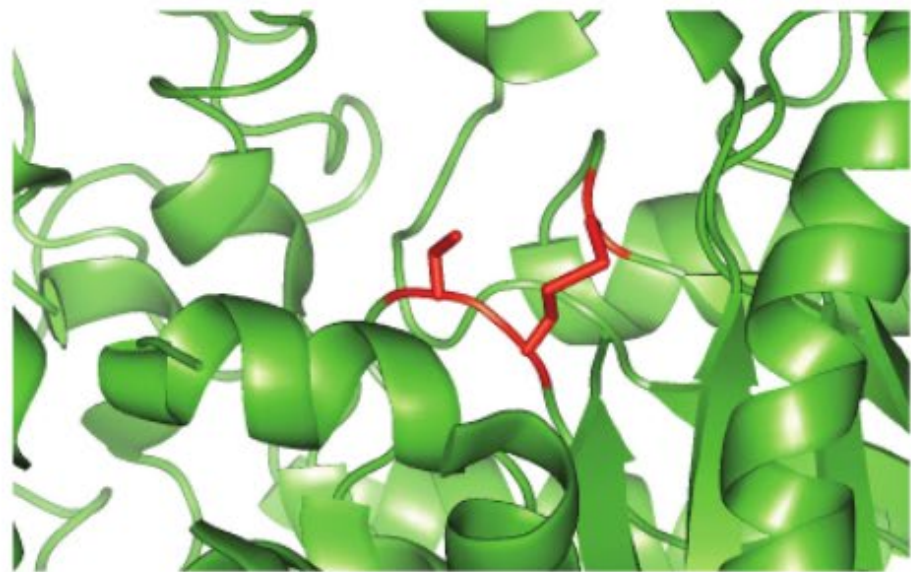
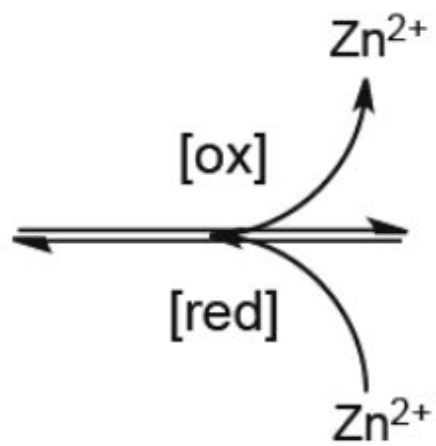
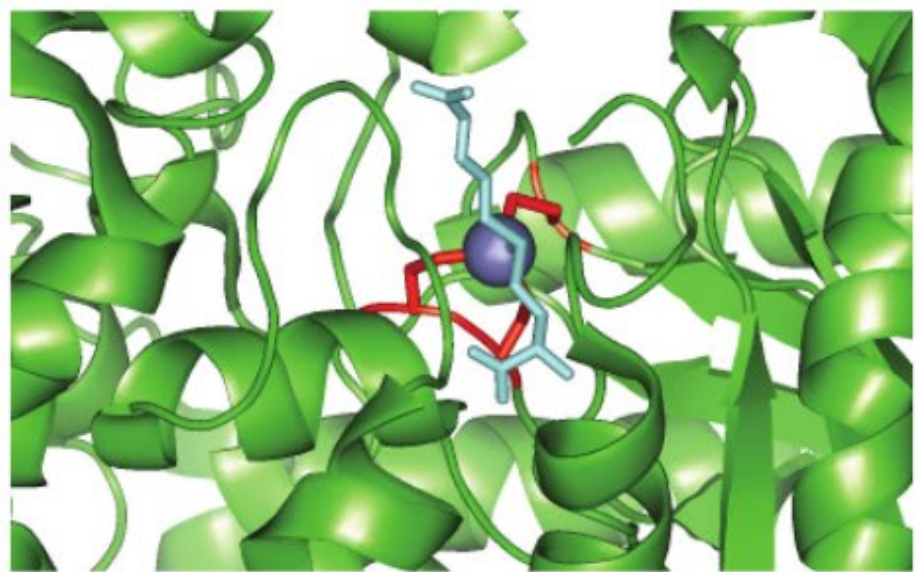
Se-methyl-L-selenocys



3. 结构分析



4功能分析



此图来自参考文献[5]

总结与展望

- 1) 在西兰花中SMT 基因编码产物结构稳定，为亲水蛋白，无跨膜结构，定位于细胞基质
- 2) 甘蓝型油菜SMT和青花菜SMT分为亲缘关系较近，可能都是由白菜型油菜SMT进化而来。另外表明甘蓝型有才SMT基因可能是由HMT 基因演化而来。植物具有硒、硫代谢的部分共用途径也表明HMT逐渐演化为SMT。其次可以看到在亲缘关系近的其稳定值很高，可能功能都相似。
- 3) 以大肠杆菌的同型半胱氨酸甲基转移酶为模板构建蛋白酶的3D结构。在还原条件下，Cys248，Cys315和Cys316与 Zn^{2+} 螯合，形成酶的活性形式，氧化条件下，Cys248和Cys315形成二硫键，导致 Zn^{2+} 的释放和酶的失活，且此中心有催化活性。
- 4) 油菜和西兰花同为芸薹属亲缘关系非常接近，希望通过研究硒代半胱氨酸甲基转移酶在西兰花中的一些功能，从而去研究其是否与油菜富硒有关，从而定位到基因上研究。

SMT 基因生物信息学分析的主要工具

程序或软件	说明	网址
ORF Finder	图形的序列分析工具，分析并找到序列的ORF区	http://www.ncbi.nlm.nih.gov/gorf/orfig.cgi
Protparam	分析未知蛋白质理化性质	http://web.expasy.org/protparam/
Prot Scale	蛋白质疏水性/亲水性	http://web.expasy.org/protscale/
TMHMM2.0	判定某蛋白是否为膜蛋白	http://www.cbs.dtu.dk/services/TMHMM/
Signal P5.0	预测真核生物信号肽	http://www.cbs.dtu.dk/services/SignalP/
Target P2.0	亚细胞定位	http://www.cbs.dtu.dk/services/TargetP/
MultiLoc	亚细胞定位	http://abi.inf.uni-tuebingen.de/Services/MultiLoc
HNN	蛋白质二级结构预测	http://npsa-pbil.ibcp.fr/
SWISS-MODEL	同源建模	http://swissmodel.expasy.org/workspace/
BLASTN	比较核酸序列和核算数据库	http://blast.ncbi.nlm.nih.gov/Blast.cgi
MegaX	分子进化与系统发育分析	http://www.megasoftware.net/

图片文献

- [1]Weeks B S, Hanna M S, Cooperstein D. Dietary selenium and selenoprotein function [J] . Medical science monitor: international medical journal of experimental and clinical research, 2012, 18 (8): 127— 132.
- [2] Rayman M P, Combs J G F, Waters D J. Selenium and vitamin E supplementation for cancer prevention [J] . JAMA: the journal of the American Medical Association, 2009, 301(18): 1876.
- [3] Gupta M, Gupta S. An Overview of Selenium Uptake, Metabolism, and Toxicity in Plants. Front Plant Sci. 2017 Jan 11;7:2074. doi: 10.3389/fpls.2016.02074. PMID: 28123395; PMCID: PMC5225104.
- [4]Evans JC, Huddler DP, Jiracek J, Castro C, Millian NS, Garrow TA, Ludwig ML. Betaine-homocysteine methyltransferase: zinc in a distorted barrel. Structure. 2002 Sep;10(9):1159-71. doi: 10.1016/s0969-2126(02)00796-7. PMID: 12220488.
- [5]Pace NJ, Weerapana E. Zinc-binding cysteines: diverse functions and structural motifs. Biomolecules. 2014 Apr 17;4(2):419-34. doi: 10.3390/biom4020419. PMID: 24970223; PMCID: PMC4101490.

图1、3来自百度文库

图2来自书籍Selenium in plants

Thank you for your attention!