



中国农业科学院郑州果树研究所

Zhengzhou Fruit Research Institute, CAAS

# 葡萄 *SNAT* 基因的功能预测和原核表达分析

## Prediction of grape *VvSNAT* gene function and analysis of its expression in prokaryotic system

G02组成员: 吴艳迪 牛 娟

张学贤 肖水平

汇 报 人: 吴艳迪

指 导 老 师: 罗静初





中国农业科学院郑州果树研究所

Zhengzhou Fruit Research Institute, CAAS

# Outline

- 1. Topics of group member**
- 2. Background**
- 3. Gene cloning**
- 4. Sequence analysis and function prediction**
- 5. Combination with molecular experiment**



中国农学院郑州果树研究所

Zhengzhou Fruit Research Institute, CAAS

# 1. Topics of group members

**G02B 牛娟**

**题目：八月瓜种子硬度相关基因的挖掘及生物信息学分析**



**拟开展工作：**挖掘组学数据中细胞壁合成与降解相关相关的基因/蛋白，并对相关基因进行生物信息学分析和功能验证。

**目的及意义：**研究八月瓜种子硬度形成的分子机制，降低种皮硬度或使种皮变软，提高果实的可食率，为八月瓜的分子育种提供理论基础。

**涉及到的软件：**除本学期学习的软件外，后期可能还会涉及到热图构建，蛋白质功能，群体结构分析等相关软件等。



# 1. Topics of group members

## G02C 张学贤

### 题目：棉花育性恢复相关基因挖掘与功能验证

棉花虽已实现三系配套，但由于**恢复系来源狭窄**，**恢复力不强**等因素，在生产上仍没有真正实现大规模三系杂交种生产。因此，有必要针对棉花恢复基因及其作用机制开展相关研究。



恢复系

不育系

利用恢复基因等基因系材料，通过转录组测序研究其基因差异表达情况，筛选与育性恢复相关的差异表达基因，挖掘与棉花育性发育相关的候选基因并进行验证，探寻其对影响育性恢复的分子机制，为优良恢复系选育提供理论依据。



中国农学院郑州果树研究所

Zhengzhou Fruit Research Institute, CAAS

# 1. Topics of group members

G02D 肖水平

题目：棉花纤维发育相关基因的克隆、生物信息学分析及功能验证



**拟开展工作：**对棉花纤维发育相关基因进行克隆；利用生物信息学方法对克隆到的基因序列等进行分析；基因功能验证。

**目的及意义：**棉花纤维是重要的天然纺织原料。随着人们对高品质纤维需求的增加和纺织工艺的改进，挖掘控制纤维性状的基因，深入研究棉花纤维发育的分子机制，对于实现棉花纤维品质改良具有重要意义。



## 1. Topics of group members

### G02A 吴艳迪

#### 题目：葡萄褪黑素合成基因 *SNAT* 原核表达分析

**实验内容：**对葡萄中褪黑素合成基因*SNAT*进行初步序列分析及功能预测，设计引物克隆其cds序列，连接至表达载体，转入大肠杆菌BL(DE3)中诱导产生可溶性蛋白，加入相应前体物质从酶学功能方面初步鉴定克隆得到的葡萄*SNAT*基因的功能。

**目的及意义：**褪黑素在植物生长发育过程中发挥重要作用，葡萄作为我国重要的栽培果树，鉴定葡萄中褪黑素合成基因的功能有助于为采用分子手段调控葡萄中内源褪黑素含量奠定基础





## 2. Background

褪黑素（N-乙酰-5-甲氧基色胺）是一种吲哚胺类物质，它普遍存在于细菌、真菌、藻类、植物、动物、人类等绝大多数生物有机体中。

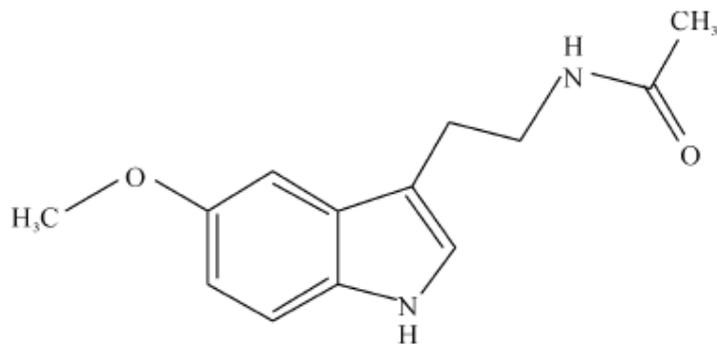
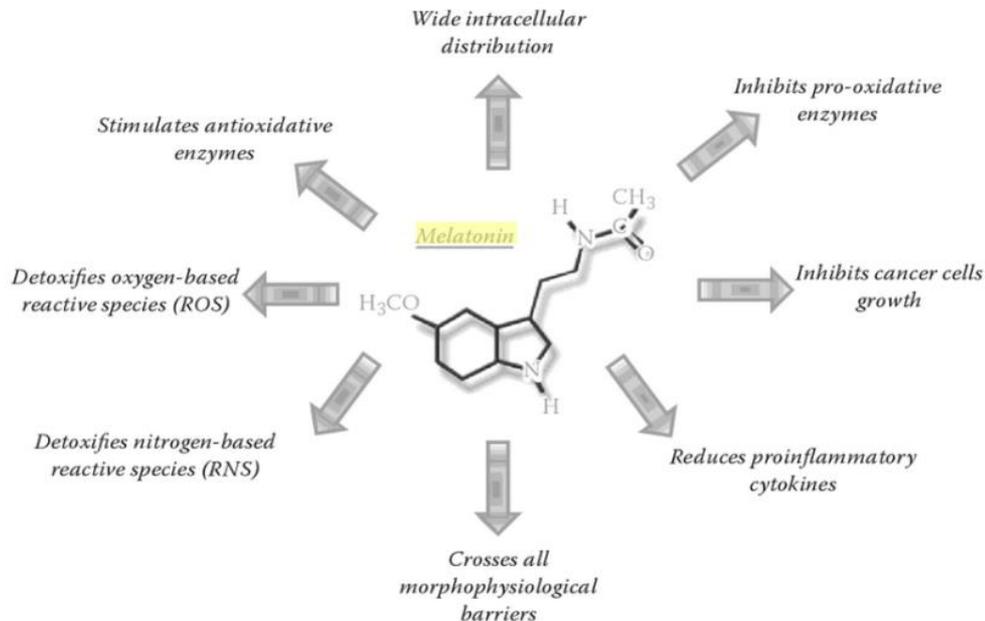


图1. 褪黑素的化学结构



## 2.1 Function of melatonin in animals

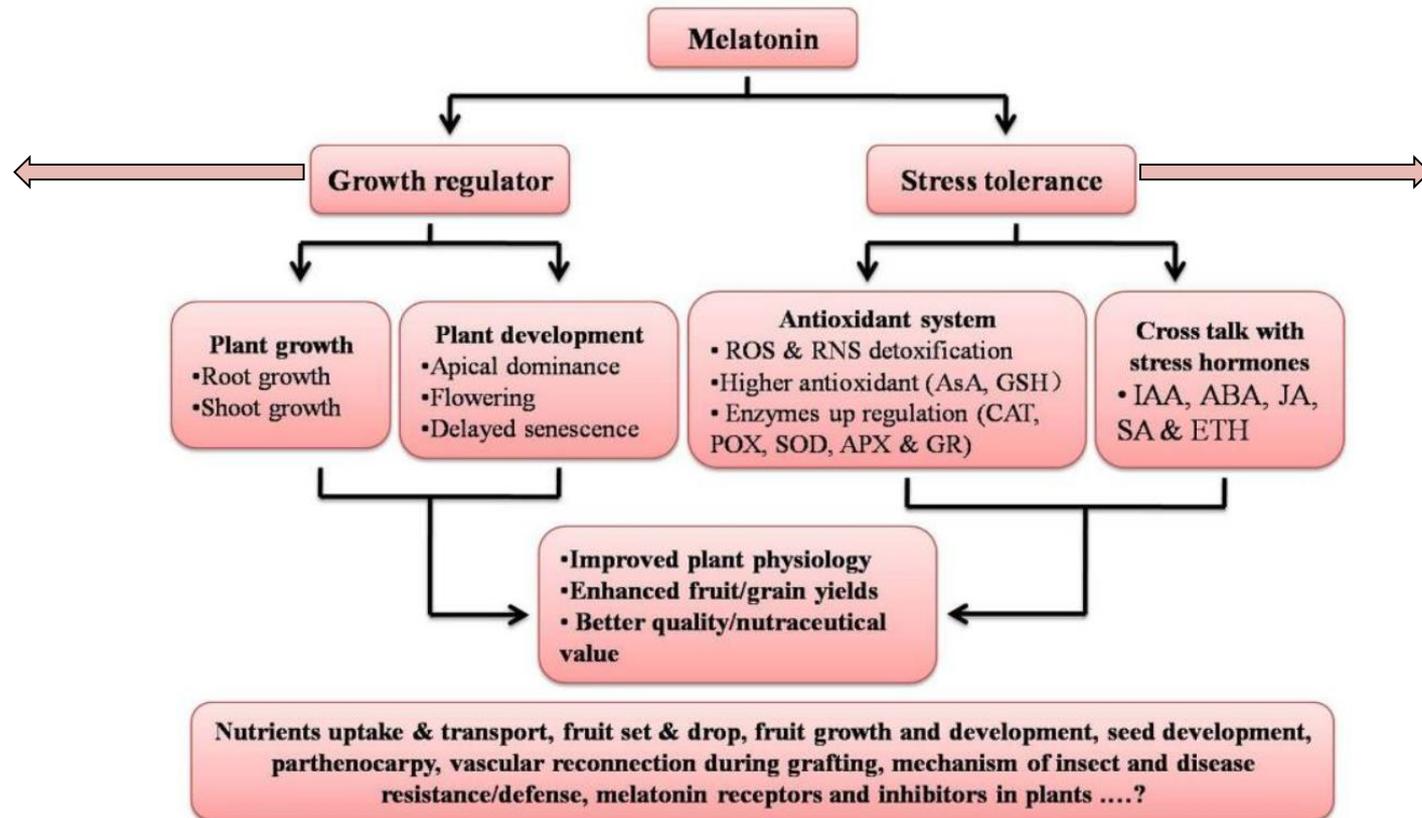
在动物体内，褪黑素具有调节昼夜节律和光周期反应、改善睡眠障碍、清除体内的自由基、减缓衰老、抗肿瘤、提高机体免疫力等多种功能。





## 2.2 Function of melatonin in plants

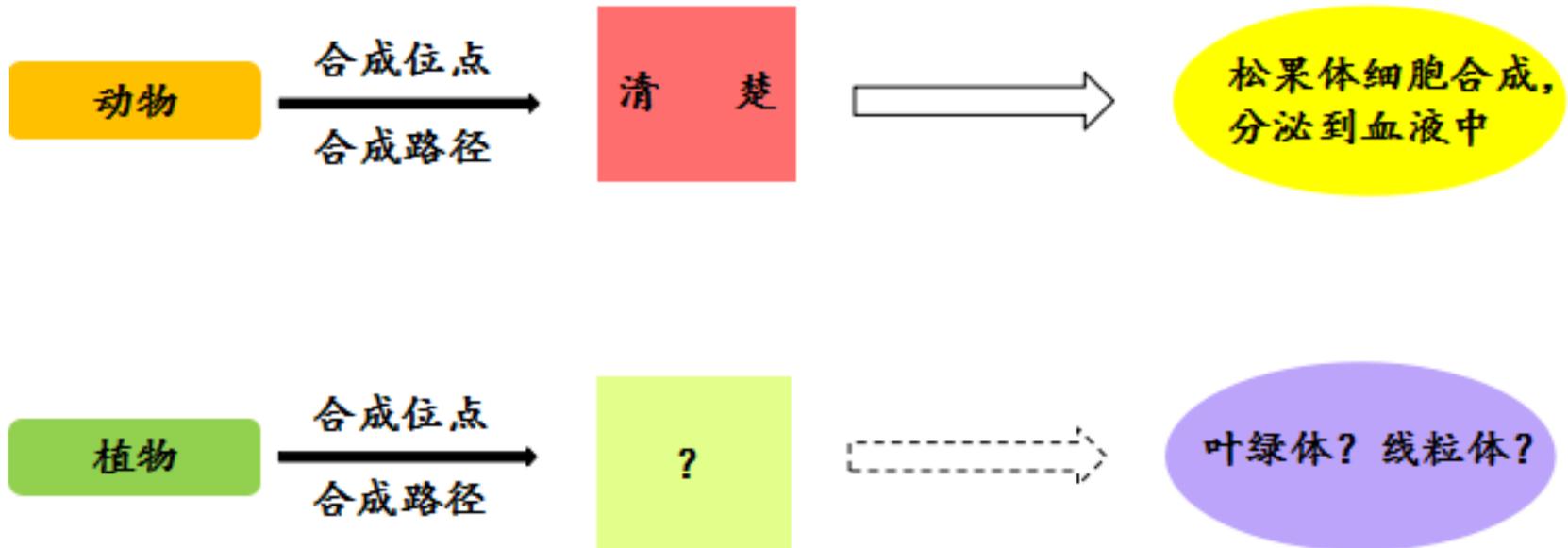
生长发育起到调节作用



缓解逆境胁迫对植物的损伤

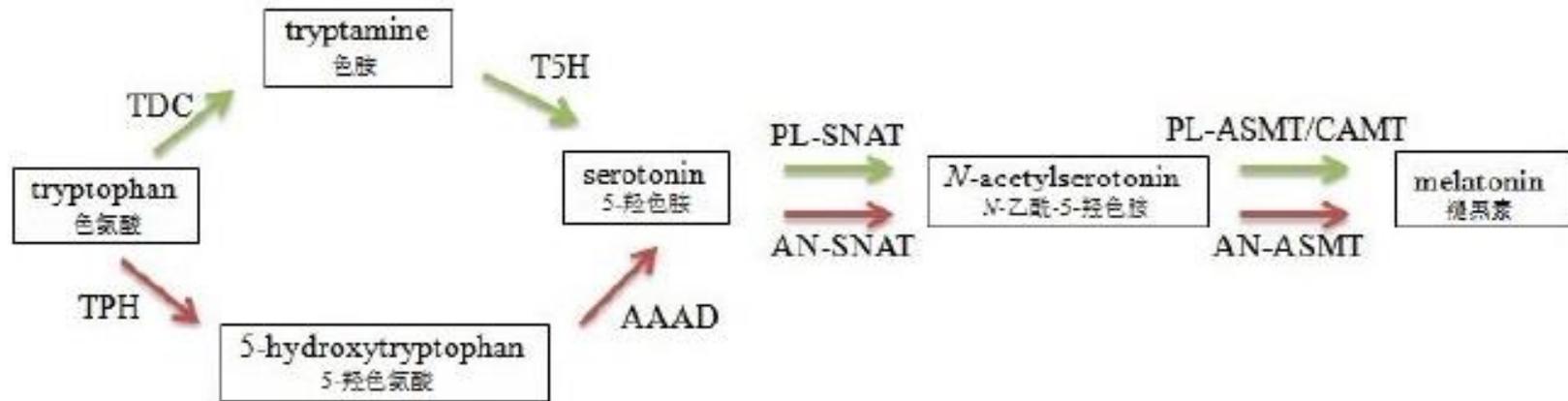


## 2.3 Biosynthesis of melatonin





## 2.3 Biosynthesis of melatonin



The green arrows identify the preferred pathway in plants while the red arrows identify the major pathway in animals. SNAT, serotonin N-acetyltransferase

褪黑素合成路径相关基因只在模式植物中被鉴定，其他高等植物（比如经济价值高的栽培果树——葡萄）中褪黑素合成基因的鉴定很有必要



## 3. Gene cloning

### 3.1 Screening of grape *SNAT* candidate genes

Run Blast: (blastp)

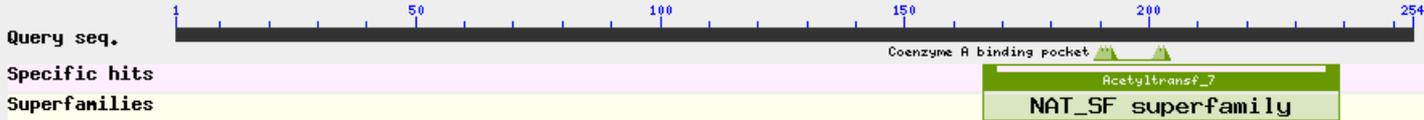
OsSNAT Genebank: BAG86973.1;

(Lei et al., 2013)

The screenshot shows the NCBI BLAST web interface. The 'Enter Query Sequence' section has 'BAG86973.1' entered in the text box. The 'Choose Search Set' section has 'Non-redundant protein sequences (nr)' selected for the database and 'Vitis vinifera (taxid:29760)' for the organism. The 'Program Selection' section has 'DELTA-BLAST (Domain Enhanced Lookup Time Accelerated BLAST)' selected. The 'BLAST' button is visible at the bottom left of the form.

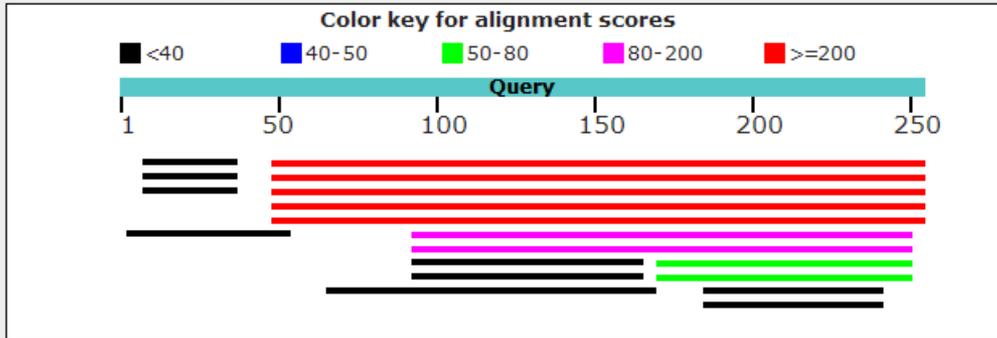


Putative conserved domains have been detected, click on the image below for detailed results.



Distribution of the top 18 Blast Hits on 18 subject sequences

Mouse over to see the title, click to show alignments



Alignments Download GenPept Graphics Distance tree of results Multiple alignment

Description	Max score	Total score	Query cover	E value	Ident	Accession	Select for PSI blast	Used to build PSSM
<input type="checkbox"/> unnamed protein product [Mtis vinifera]	299	299	81%	5e-103	72%	<a href="#">CBI31163.3</a>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
<input checked="" type="checkbox"/> PREDICTED: acetyltransferase NSI isoform X4 [Mtis vinifera]	299	299	81%	1e-102	72%	<a href="#">XP_002266361.2</a>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
<input checked="" type="checkbox"/> PREDICTED: acetyltransferase NSI isoform X3 [Mtis vinifera]	291	291	81%	1e-99	69%	<a href="#">XP_010653125.1</a>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
<input checked="" type="checkbox"/> PREDICTED: acetyltransferase NSI isoform X2 [Mtis vinifera]	283	283	81%	6e-96	60%	<a href="#">XP_019077452.1</a>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
<input checked="" type="checkbox"/> PREDICTED: acetyltransferase NSI isoform X1 [Mtis vinifera]	276	276	81%	7e-93	58%	<a href="#">XP_019077451.1</a>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
<input type="checkbox"/> hypothetical protein VITISV_037700 [Mtis vinifera]	113	113	62%	2e-30	40%	<a href="#">CAN70315.1</a>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
<input type="checkbox"/> PREDICTED: serotonin N-acetyltransferase 2, chloroplastic [Mtis vinifera]	113	113	62%	2e-30	40%	<a href="#">XP_002271276.1</a>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
<input type="checkbox"/> PREDICTED: uncharacterized N-acetyltransferase vcf52 [Mtis vinifera]	57.8	57.8	31%	2e-09	37%	<a href="#">XP_002270840.1</a>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>
<input type="checkbox"/> unnamed protein product [Mtis vinifera]	57.4	57.4	31%	3e-09	37%	<a href="#">CBI25520.3</a>	<input checked="" type="checkbox"/>	<input checked="" type="checkbox"/>



### 3. Gene cloning

#### Note: Principle for primer design

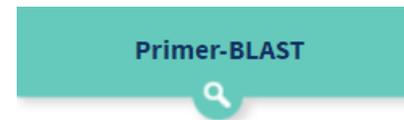
- (1) 引物长度控制在15-30bp以内（18-25bp）
- (2) GC含量一般设为40%-60%
- (3) T<sub>m</sub>一般设为55 °C-65 °C，退火温度一般要比解链温度低5 °C
- (4) 3'端不应超过4个连续的G或C
- (5) 引物自身不应存在互补序列，
- (6) 如扩增c<sub>ds</sub>序列，引物3'端不要终止于密码子的第3位，因密码子的第3位易发生简并，会影响扩增特异性与效率。



Primer Premier 6.0 Primer Premie...



Oligo 7



Design primers specific to  
your PCR template



## 3.2 Primer design (PCR) (XM\_002266325.4)

Vector NTI - [NEWMOL]

File Edit View Primer Design Analyses Cloning Ge

### Find Primers in Selected Region of NEWMOL

Uniqueness		Qualities		Filter for Feature	
Primer	Amplicon	Structure	Pairs	Similarity	3' end

Region of Analysis  
From: 41 bp To: 810 bp Product Length  
Min: 750 bp Max: 770 bp

Maximum Number of Output Primers: 10

Analysis Conditions

Salt Conc. (mMol): 50.0 Probe Conc. (pMol): 250.0 dG Temperature (C): 25.0

Tm (C):  
>= 55.0 <= 65.0  
%GC: >= 40.0 <= 60.0  
Length: >= 18 <= 25

DNA

User-Defined Primers  
Sense:  Analyze  
Antisense:  Analyze

Attach to 5' terminus  
of Sense:  of Antisense:

Buttons: Load... Save... Default 确定 取消

#5: Product of length 745 (rating: 162)  
Contains region of the molecule from 72 to 816  
Tm: 77.0 C TaOpt: 55.6 C GC: 43.9  
Sense Primer:  
GCCCTCCTCTCCACTTCTCAGTA  
Similarity: 100.0%  
Length: 23 Tm: 55.5 C GC: 56.5  
dH: -169.6 kcal/mol dS: -437.5 cal/mol dG: -37.4 kcal/mol  
Antisense Primer:  
ATGCTTGCTCGAAGCGCA  
Similarity: 100.0%  
Length: 18 Tm: 55.3 C GC: 55.6  
dH: -147.0 kcal/mol dS: -373.2 cal/mol dG: -34.0 kcal/mol  
Tm Difference: 0.2  
GC Difference: 1.0



## 4. sequence analysis and function prediction

### 4.1 Sequence align (DNAMAN, Needle)

```

SNAT-NCBI ATGCAAACCCACAGCGCCCTCCCTCTCCA CTCTCA #-----
SNAT-测序 ATGCAAACCCACAGCGCCCTCCCTCTCCA CTCTCA #
Consensus atgcaaaccacagcgccctccctctcca cttctca # Aligned_sequences: 2
# 1: SNAT-测序
# 2: SNAT-NCBI
SNAT-NCBI ATCATTGTTCCCTTGC AATTTAAATCTTGGGGTTG # Matrix: EBLDSUM62
SNAT-测序 ATCATTGTTCCCTTGC AATTTAAATCTTGGGGTTG # Gap_penalty: 10.0
Consensus a cattgttcccttgc aatTTAAATCTTGGGGTTG # Extend_penalty: 0.5

SNAT-NCBI TTTGGGAATCCATCAGATCCGGGTTTTTGAAGAAT #
SNAT-测序 TTTGGGAATCCATCAGATCCGGGTTTTTGAAGAAT # Length: 246
Consensus tttgggaatccatcagatccgggtttttgaagaat # Identity: 246/246 (100.0%)
# Similarity: 246/246 (100.0%)
SNAT-NCBI GAAGAACCATTGCCCGAGGAGTTTGTCTTGTGTA # Gaps: 0/246 (0.0%)
SNAT-测序 GAAGAACCATTGCCCGAGGAGTTTGTCTTGTGTA # Score: 1290.0
Consensus gaagaaccattgcc cgaggagtttGTCTTGTGTA #

SNAT-NCBI AGGTGGAGATGTTGATGTCATGATCTCCAGGCCCT #-----
SNAT-测序 AGGTGGAGATGTTGATGTCATGATCTCCAGGCCCT #
Consensus aggtggagatgttgatgtctatgatctccaggccct #

SNAT-NCBI CTGCAGCTTTAAAAAATAGCTACATGGTAGCCACA #
SNAT-测序 CTGCAGCTTTAAAAAATAGCTACATGGTAGCCACA #
Consensus ctgcagctttaaaaaatagctacatggtagccaca #

SNAT-NCBI AAGAAGCTTATTGGCATGGCCCGTGCTACATCAGA #
SNAT-测序 AAGAAGCTTATTGGCATGGCCCGTGCTACATCAGA #
Consensus aagaagcttattggcatggcccggtgctacatcaga #

SNAT-NCBI ATATCAGGGCCAGGGACTTGGAAAGGCCCTCGTTG #
SNAT-测序 ATATCAGGGCCAGGGACTTGGAAAGGCCCTCGTTG #
Consensus atatcagggccagggacttggaaaggccctcgttg #

SNAT-NCBI CACTCTTTGCAGATAGCCAAGTGGTGGAGTTCTAT #
SNAT-测序 CACTCTTTGCAGATAGCCAAGTGGTGGAGTTCTAT #
Consensus cactctttgcagatagccaagtgggtggagtctat #

SNAT-NCBI TTCTGGCACCCAAGGTATTAG #
SNAT-测序 TTCTGGCACCCAAGGTATTAG #
Consensus ttctggcacccaaggtattag #

```

```

SNAT-测序      1 MQTHSALLSTSQYPIFSFGCNC DGSNLSLFPCLNMLGVGRARRKLVKVCQT      50
                |||
SNAT-NCBI      1 MQTHSALLSTSQYPIFSFGCNC DGSNLSLFPCLNMLGVGRARRKLVKVCQT      50

SNAT-测序     51 RASFWESIRSGLKNNTTQVIEFPSTNQE EEEPLPEEFVLVEKTLADGAI     100
                |||
SNAT-NCBI     51 RASFWESIRSGLKNNTTQVIEFPSTNQE EEEPLPEEFVLVEKTLADGAI     100

SNAT-测序    101 EQILFSSGGD VDVYDLQALCDKVGWFR RPLSKLSAALKNSYMWATLHSLR    150
                |||
SNAT-NCBI    101 EQILFSSGGD VDVYDLQALCDKVGWFR RPLSKLSAALKNSYMWATLHSLR    150

SNAT-测序    151 KSPGEERNEQK KLI GMARATSDHAFNATIWDV LVPDPSYQGGLGKALVEK    200
                |||
SNAT-NCBI    151 KSPGEERNEQK KLI GMARATSDHAFNATIWDV LVPDPSYQGGLGKALVEK    200

SNAT-测序    201 TIRALLQRDIGNITL FADSQVVEFYRNLGFEPDPEGIKGMFWHPRY     246
                |||
SNAT-NCBI    201 TIRALLQRDIGNITL FADSQVVEFYRNLGFEPDPEGIKGMFWHPRY     246

```

匹配率是99.33 %

但是其编码的氨基酸序列相同，而且也未出现移码现象



## 4.2 Search from the database

### 4.2.1 Search from UniProt datase

UniProtKB search results for "serotonin n acetyltransferase" AND reviewed:yes

UniProtKB results

UniProtKB consists of two sections:

- Reviewed (Swiss-Prot) - Manually annotated**  
Records with information extracted from literature and curator-evaluated computational analysis.
- Unreviewed (TrEMBL) - Computationally analyzed**  
Records that await full manual annotation.

The UniProt Knowledgebase (UniProtKB) is the central hub for the collection of functional information on proteins, with accurate, consistent and rich annotation. In addition to capturing the core data mandatory for each UniProtKB entry (mainly, the amino acid sequence, protein name or description, taxonomic data and citation information), as much annotation information as possible is added.

Filter by: Reviewed (12) Swiss-Prot

Popular organisms: Rice (2), Bovine (1), Human (1), Mouse (1), Rat (1), Other organisms

View by: Results table

Entry	Entry name	Protein names	Gene names	Organism	Length
Q16613	SNAT_HUMAN	Serotonin N-acetyltransferase	AANAT SNAT	Homo sapiens (Human)	207
O88816	SNAT_MOUSE	Serotonin N-acetyltransferase	Aanat Snat	Mus musculus (Mouse)	205
Q64666	SNAT_RAT	Serotonin N-acetyltransferase	Aanat Snat	Rattus norvegicus (Rat)	205
Q29495	SNAT_SHEEP	Serotonin N-acetyltransferase	AANAT SNAT	Ovis aries (Sheep)	207
Q5KQ16	SNAT1_ORYSJ	Serotonin N-acetyltransferase 1, ch...	SNAT1 GNAT5, NSI, SNAT, Os05g0481000, LOC_Os05g40260	Oryza sativa subsp. japonica (Rice)	254
O97756	SNAT_MACMU	Serotonin N-acetyltransferase	AANAT SNAT	Macaca mulatta (Rhesus macaque)	204
P79774	SNAT_CHICK	Serotonin N-acetyltransferase	AANAT SNAT	Gallus gallus (Chicken)	205
Q6Z1Y6	SNAT2_ORYSJ	Serotonin N-acetyltransferase 2, ch...	SNAT2 Os08g0102000, LOC_Os08g01170, B1147B12.22, OsJ_25726, P0015C07.5	Oryza sativa subsp. japonica (Rice)	200
O02785	SNAT_BOVIN	Serotonin N-acetyltransferase	AANAT SNAT	Bos taurus (Bovine)	207
Q51S55	SNAT_PANTR	Serotonin N-acetyltransferase	AANAT SNAT	Pan troglodytes (Chimpanzee)	207
Q9R0A9	SNAT_MESAU	Serotonin N-acetyltransferase	AANAT SNAT	Mesocricetus auratus (Golden hamster)	207



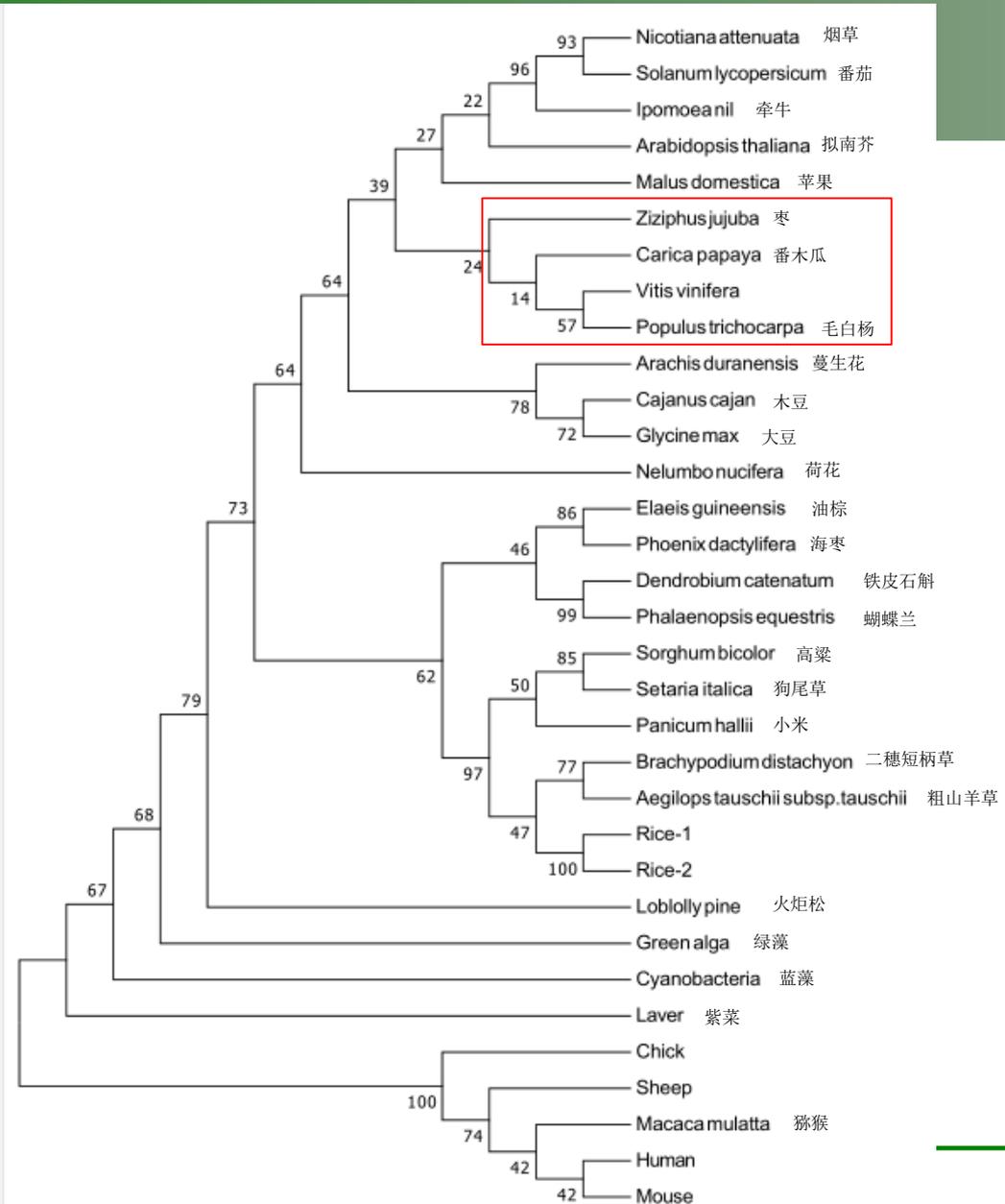
## 4.2.2 Search from NCBI database

Description	Max score	Total score	Query cover	E value	Ident	Accession	Select for PSI blast	Used to build PSSM
<input checked="" type="checkbox"/> <a href="#">PREDICTED: acetyltransferase NSI isoform X4 [Mitis vinifera]</a>	519	519	100%	0.0	100%	<a href="#">XP_002266361.2</a>	<input checked="" type="checkbox"/>	
<input type="checkbox"/> <a href="#">PREDICTED: acetyltransferase NSI isoform X3 [Mitis vinifera]</a>	510	510	100%	0.0	96%	<a href="#">XP_010653125.1</a>	<input checked="" type="checkbox"/>	
<input type="checkbox"/> <a href="#">PREDICTED: acetyltransferase NSI isoform X2 [Mitis vinifera]</a>	504	504	100%	8e-180	86%	<a href="#">XP_019077452.1</a>	<input checked="" type="checkbox"/>	
<input type="checkbox"/> <a href="#">PREDICTED: acetyltransferase NSI isoform X1 [Mitis vinifera]</a>	495	495	100%	3e-176	83%	<a href="#">XP_019077451.1</a>	<input checked="" type="checkbox"/>	
<input type="checkbox"/> <a href="#">unnamed protein product [Mitis vinifera]</a>	433	433	83%	6e-153	100%	<a href="#">CBI31163.3</a>	<input checked="" type="checkbox"/>	
<input type="checkbox"/> <a href="#">PREDICTED: acetyltransferase NSI isoform X1 [Juglans regia]</a>	394	394	100%	6e-137	77%	<a href="#">XP_018814362.1</a>	<input checked="" type="checkbox"/>	
<input type="checkbox"/> <a href="#">PREDICTED: acetyltransferase NSI isoform X1 [Theobroma cacao]</a>	384	384	100%	7e-133	78%	<a href="#">XP_007033470.2</a>	<input checked="" type="checkbox"/>	
<input type="checkbox"/> <a href="#">PREDICTED: acetyltransferase NSI isoform X2 [Juglans regia]</a>	382	382	88%	1e-132	83%	<a href="#">XP_018814363.1</a>	<input checked="" type="checkbox"/>	
<input checked="" type="checkbox"/> <a href="#">PREDICTED: serotonin N-acetyltransferase 1, chloroplastic isoform X3 [Nelumbo nucifera]</a>	382	382	100%	2e-132	75%	<a href="#">XP_010244575.1</a>	<input checked="" type="checkbox"/>	
<input type="checkbox"/> <a href="#">Nuclear shuttle interacting isoform 2 [Theobroma cacao]</a>	382	382	100%	4e-132	77%	<a href="#">EOY04396.1</a>	<input checked="" type="checkbox"/>	
<input type="checkbox"/> <a href="#">acetyltransferase NSI [Hevea brasiliensis]</a>	382	382	95%	9e-132	81%	<a href="#">XP_021684353.1</a>	<input checked="" type="checkbox"/>	
<input type="checkbox"/> <a href="#">acetyltransferase NSI isoform X2 [Lactuca sativa]</a>	379	379	96%	4e-131	76%	<a href="#">XP_023770682.1</a>	<input checked="" type="checkbox"/>	
<input type="checkbox"/> <a href="#">acetyltransferase NSI [Momordica charantia]</a>	378	378	99%	1e-130	75%	<a href="#">XP_022159397.1</a>	<input checked="" type="checkbox"/>	
<input type="checkbox"/> <a href="#">acetyltransferase NSI [Cynara cardunculus var. scolymus]</a>	376	376	94%	5e-130	80%	<a href="#">XP_024975398.1</a>	<input checked="" type="checkbox"/>	
<input checked="" type="checkbox"/> <a href="#">serotonin N-acetyltransferase 1, chloroplastic [Cajanus cajan]</a>	376	376	96%	1e-129	76%	<a href="#">XP_020203403.1</a>	<input checked="" type="checkbox"/>	
<input type="checkbox"/> <a href="#">acetyltransferase NSI [Citrus clementina]</a>	376	376	100%	2e-129	76%	<a href="#">XP_006429638.1</a>	<input checked="" type="checkbox"/>	
<input type="checkbox"/> <a href="#">hypothetical protein CUMW_001790 [Citrus unshiu]</a>	375	375	100%	7e-129	76%	<a href="#">GAY32321.1</a>	<input checked="" type="checkbox"/>	
<input checked="" type="checkbox"/> <a href="#">serotonin N-acetyltransferase 1, chloroplastic isoform X2 [Ziziphus jujuba]</a>	373	373	100%	7e-129	77%	<a href="#">XP_015897769.1</a>	<input checked="" type="checkbox"/>	



### 4.3 Phylogeny analysis

通过Blastp搜索NR数据库中同源序列，选取identity > 70%以上且注释为serotonin N-acetyltransferase的序列25条；搜索swiss-prot数据库中搜索到已审阅的identity > 60%的拟南芥、水稻serotonin N-acetyltransferase同源序列3条；另从Uniprot数据库找出5种已审阅的动物serotonin N-acetyltransferase序列加入系统进化树构建，共33个物种进行MEGA建树。







## 4.5 Primary Structure Analysis of Grape SNAT Amino Acid Sequence

### 4.5.1 Analysis of domain

**低复杂度区域**

**Pfam Acetyltransf\_7**

**乙酰转移酶结构域**

**Low complexity region**

This is a region of low compositional complexity, as detected by the [SEG](#) program. The region starts at position **79** and ends at position **92**.

**Low complexity region (14 aa):**

[Submit to BLAST](#) [Copy to clipboard](#)

EEEEPLPEEFVLVE

**Acetyltransf\_7 domain**

This is a [Pfam](#) domain. Please see the [Acetyltransf\\_7](#) entry in Pfam for full annotation.

<b>Position:</b>	<b>156 to 232</b>
<b>E-value:</b>	<b>1.4e-13 (HMMER3)</b>

**Accession:**

**Pfam domain sequence (77 aa):**

[Submit to BLAST](#) [Copy to clipboard](#)

ERNEQKLLIGMARATSDHAFNATIWDVLVDPYQQQLGKALVEKTIRALLQRDIGNITL  
FADSQVVEFYRNLGFEP



## 4.5.2 Analysis of ProtParam

### Amino acid composition:

CSV format

Ala (A)	16	6.5%
Arg (R)	14	5.7%
Asn (N)	12	4.9%
Asp (D)	12	4.9%
Cys (C)	5	2.0%
Gln (Q)	12	4.9%
Glu (E)	17	6.9%
Gly (G)	18	7.3%
His (H)	4	1.6%
Ile (I)	12	4.9%
Leu (L)	26	10.6%
Lys (K)	13	5.3%
Met (M)	4	1.6%
Phe (F)	12	4.9%
Pro (P)	13	5.3%
Ser (S)	20	8.1%
Thr (T)	12	4.9%
Trp (W)	4	1.6%
Tyr (Y)	6	2.4%
Val (V)	14	5.7%
Pyl (O)	0	0.0%
Sec (U)	0	0.0%
(B)	0	0.0%
(Z)	0	0.0%
(X)	0	0.0%

Number of amino acids: 246

Molecular weight: 27502.25

Theoretical pI: 6.12

Total number of negatively charged residues (Asp + Glu): 29

Total number of positively charged residues (Arg + Lys): 27

### Instability index:

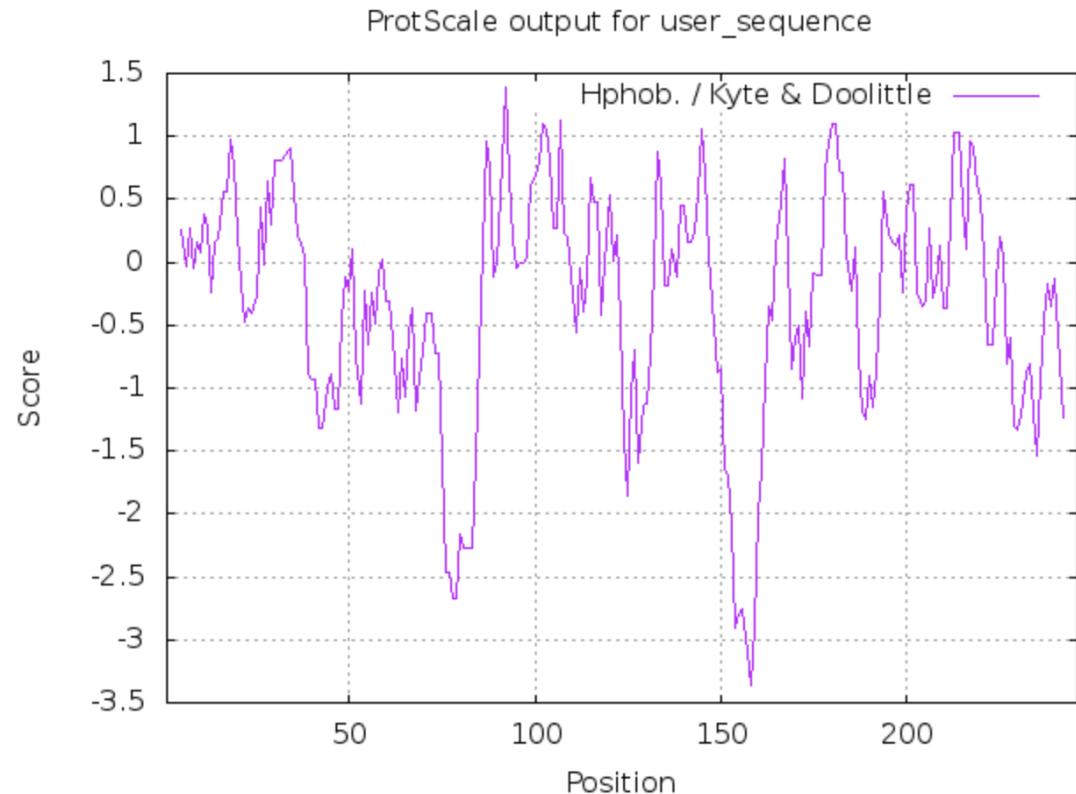
The instability index (II) is computed to be 43.99  
This classifies the protein as unstable.



## 4.6 Secondary Structure Analysis of Grape SNAT Amino Acid Sequence

### 4.6.1 Hydrophilic and hydrophobic analysis (ProtScale)

数值越高表明该区域多肽链的疏水性越强, 该蛋白存在数段亲水性较强的区域, 蛋白总体亲水性较强。

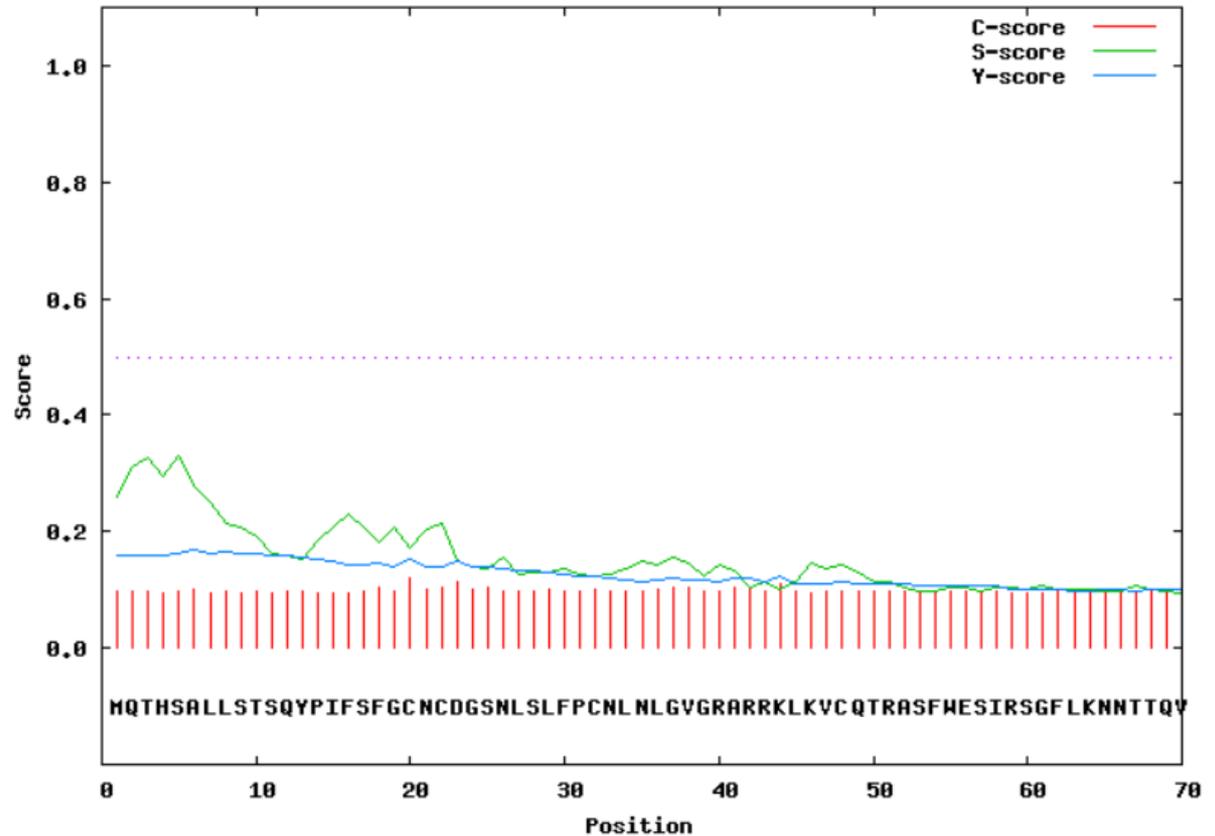




## 4.6.2 Signal peptide prediction

```
# SignalP-4.0 gram- predictions
# Measure Position Value Cutoff signal peptide?
max. C 20 0.120
max. Y 6 0.168
max. S 5 0.330
mean S 1-5 0.305
0 1-5 0.233 0.570 NO
Name=tmp_signalp_seq SP=NO D=0.233 D-cutoff=0.570 Networks=SignalP-noTM
```

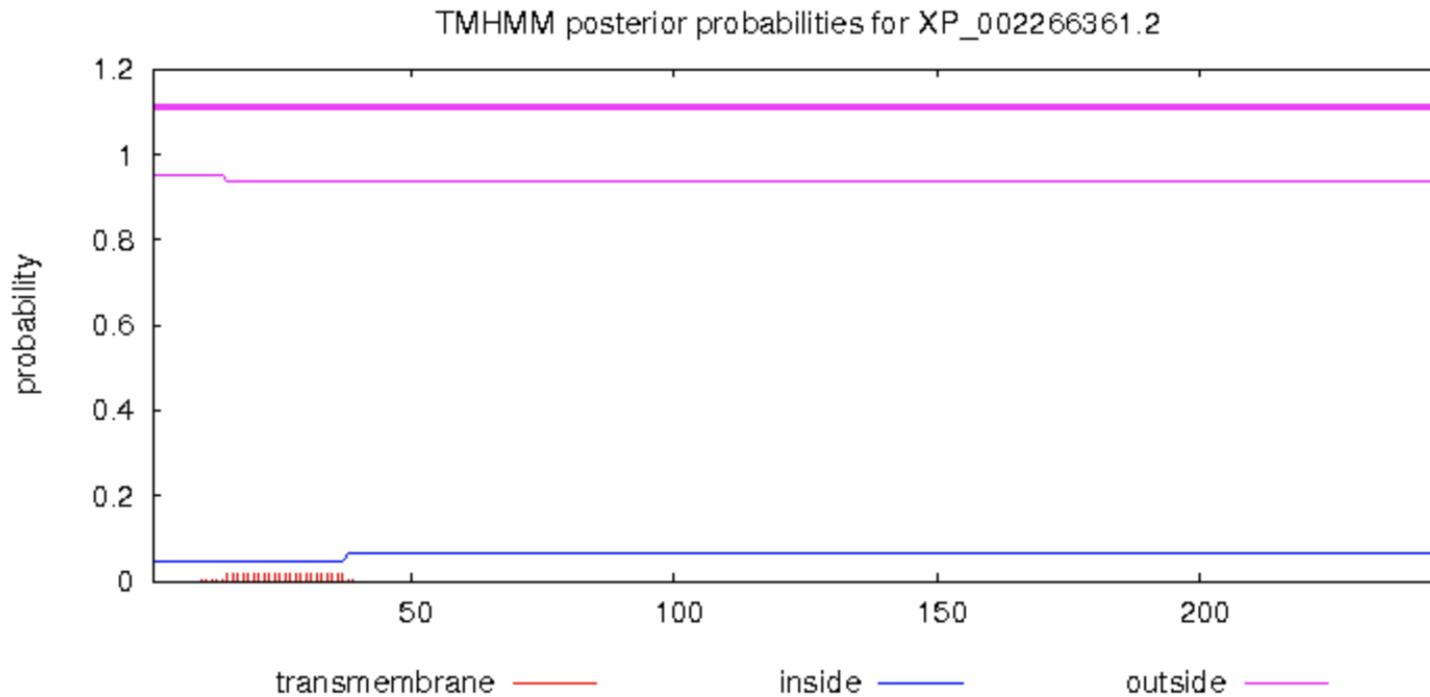
SignalP-4.0 prediction (gram- networks)



该蛋白无信号肽序列



## 4.6.3 Transmembrane helix analysis (TMHMM)



该蛋白无跨膜结构



## 4.6.4 Transit peptide prediction



### chlorop v1.1  
Number of query s

Name

Sequence

---Detailed output

Residue	--MN-sc	Raw	D
---------	---------	-----	---

```

1      ATGCAAACCCACAGCGCCCTCCTCTCCACTTCTCAGTACCCCATTTTCTCTTTTGGTTGCAATTGTGACGGATCAAATTTATCATTGTTC
    M Q T H S A L L S T S Q Y P I F S F G C N C D G S N L S L F
91     CCTTGCAATTTAAATCTTGGGGTTGGGAGAGCAAGAAGAAAGCTTAAGGTTTGTCAAACGAGAGCTAGCTTTTGGGAATCCATCAGATCC
    P C N L N L G V G R A R R K L K V C Q T R A S F W E S I R S
181    GGGTTTTTGAAGAATAATACAACACAAGTTATTGAACCACCCTCCACAAACCAAGAAGAGGAAGAACCATTGCCCGAGGAGTTTGTCTT
    G F L K N N T T Q V I E P P S T N Q E E E E P L P E E F V L
271    GTTGAAAAGACTCTAGCTGACGGAGCAATTGAACAGATAATATTTTCTTCAGGTGGAGATGTTGATGTCTATGATCTCCAGGCCTTATGT
    V E K T L A D G A I E Q I I F S S G G D V D V Y D L Q A L C
361    GATAAGGTGGGCTGGCCCCGAAGGCTCTATCAAACTATCTGCAGCTTTAAAAAATAGCTACATGGTAGCCACATTGCATTCTTAAGG
    D K V G W P R R P L S K L S A A L K N S Y M V A T L H S L R
451    AAATCACCTGGAGAAGAGAGGAATGAGCAAAAGAAGCTTATTGGCATGGCCCGTCTACATCAGATCATGCCTTCAATGCCACAATTGG
    K S P G E E R N E Q K K L I G M A R A T S D H A F N A T I W
541    GATGTCCTTGTGATCCTTCATATCAGGGCCAGGGACTTGGAAAGGCCCTCGTTGAAAAGACTATAAGAGCTCTTCTGCAAAGGGACATT
    D V L V D P S Y Q G Q G L G K A L V E K T I R A L L Q R D I
631    GGCAATATAACACTCTTTGCAGATAGCCAAGTGGTGGAGTTCTATCGAAATTTAGGTTTTGAACCTGACCAGAGGGCATCAAAGGTATG
    G N I T L F A D S Q V V E F Y R N L G F E P D P E G I K G M
721    TTCTGGCACCCAAGGTATTAG
    F W H P R Y *

```



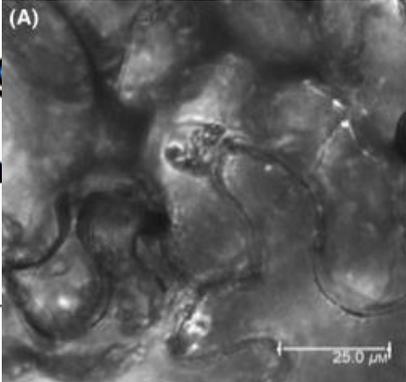
## 4.6.4 Subcellular localization (Target P)

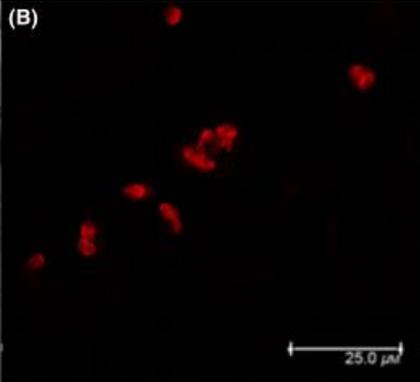
CENTERFOR  
RBIOLGI  
CALSEQU  
ENCEANA  
LYSIS CBS

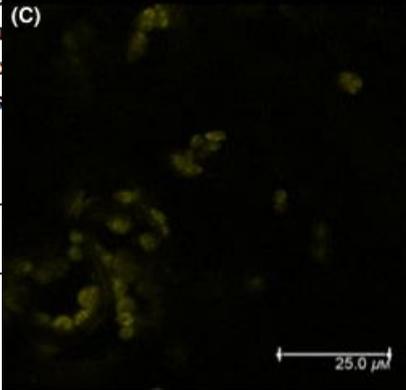
Target  
Tech

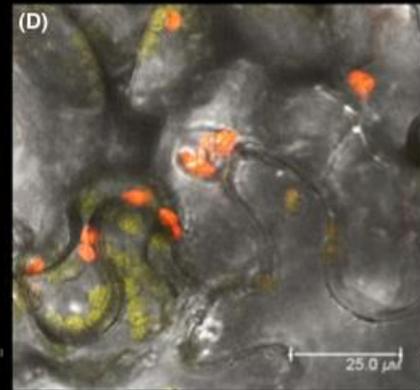
```
### targetp v1.1 pre
Number of query sequ
Cleavage site predic
Using PLANT network

Name
-----
Sequence
-----
cutoff
```









###

####

-----

-----

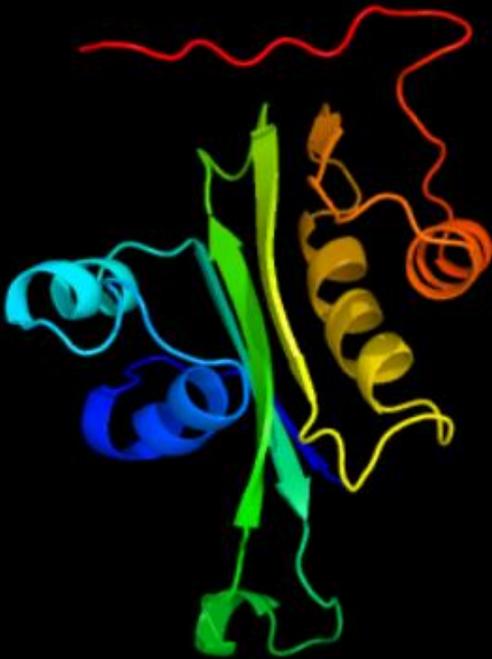
-----

**Explain the output.**

Localization of OsSNAT. (A) Bright-field image of Agrobacterium-infiltrated tobacco. (B) The red fluorescence of OsSNAT-mCherry and (C) the yellow fluorescence of chlorophyll (Chl) were monitored 12 hr after XVE induction, and (D) the two fluorescence images were merged. *N. benthamiana* leaves infiltrated with Agrobacterium with XVE-inducible OsSNAT-mCherry were grown for 2 days in a growth room before XVE induction and visualization by confocal microscopy. Bars = 25 μm.



## 4.3 3D structure prediction



Model (left) based on template [d2atra1](#)

Top template information

**Fold:** Acyl-CoA N-acyltransferases (Nat)

**Superfamily:** Acyl-CoA N-acyltransferases (Nat)

**Family:** N-acetyl transferase, NAT

Confidence and coverage

Confidence:

99.9%

Coverage: 52%

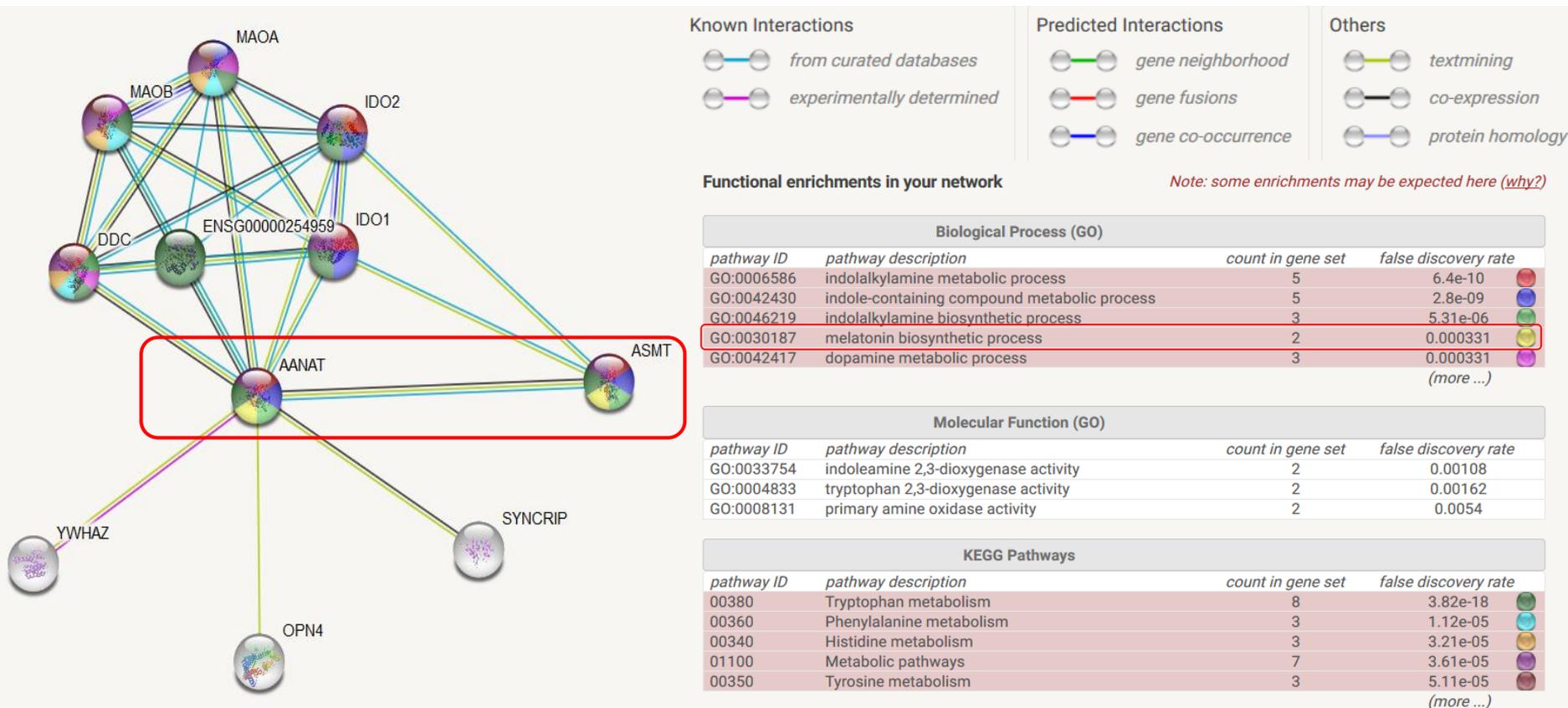
129 residues ( 52% of your sequence) have been modelled with 99.9% confidence by the single highest scoring template.

(Phary2)



## 4.7 Function prediction of SNAT protein

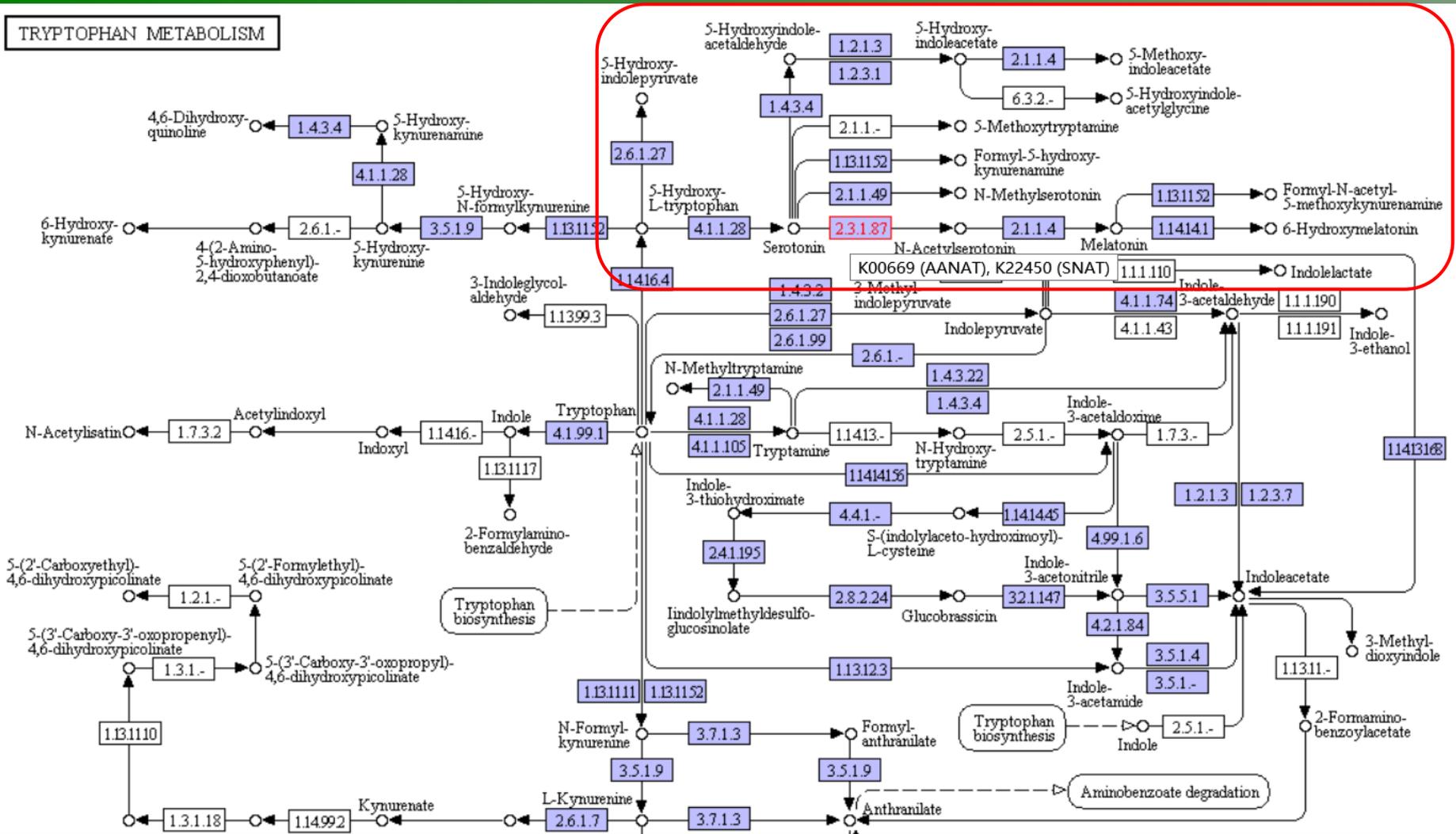
### 4.7.1 The protein interaction network of Human SNAT





## 4.7.2 Metabolic pathway of rice

### TRYPHOPHAN METABOLISM

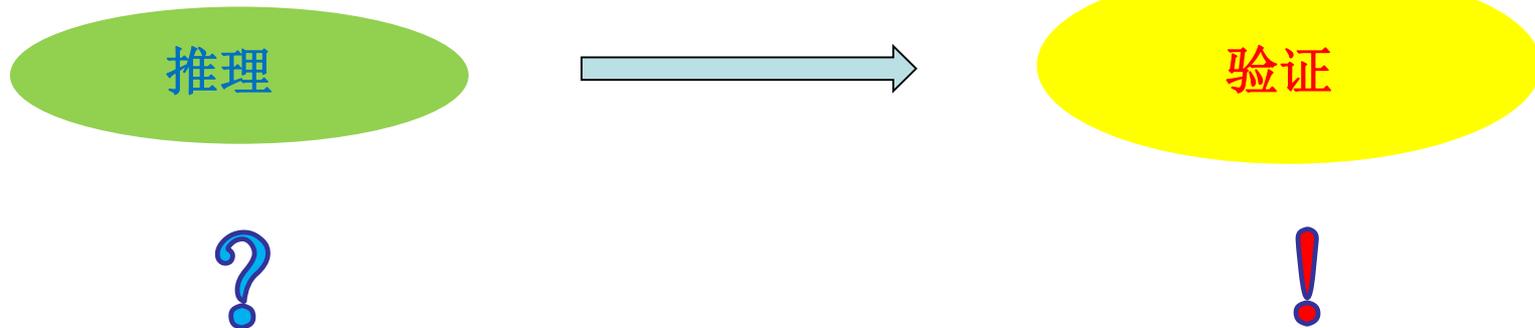




## 5. Combine with molecular experiment

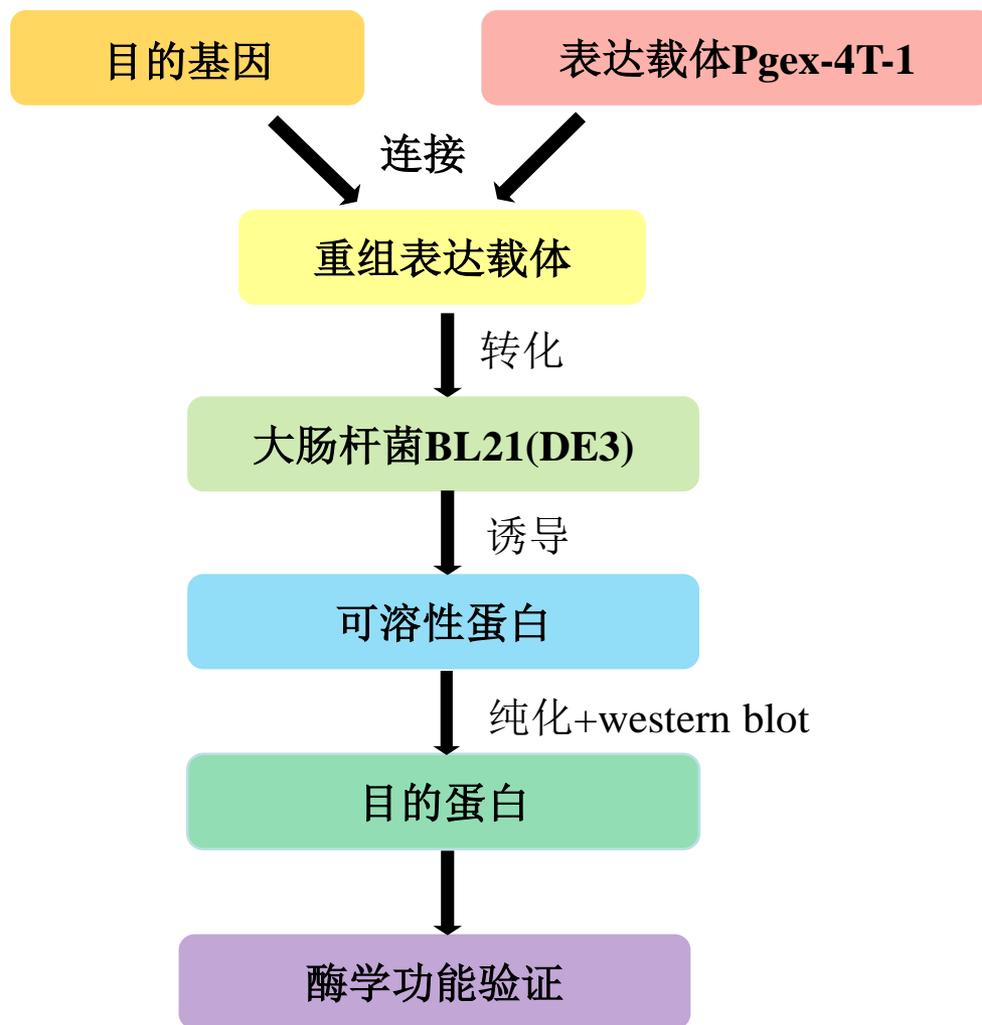
### Problem

葡萄中SNAT（5-羟色胺-N-乙酰基转移酶）是不是可能也具有与人，拟南芥中SNAT相似的功能，可以将5-羟色胺转化为N-乙酰-5-羟色胺？





## 5.1 Experiment process





## 5.2 Rare codon analysis of amino acid

*E. coli* Codon Usage Analyzer 2.1

### *SNAT*

ATG CAA ACC CAC AGC GCC CTC CTC TCC ACT TCT CAG TAC **CCC** ATT TTC TCT TTT GGT TGC  
AAT TGT GAC **GGA** TCA AAT TTA TCA TTG TTC CCT TGC AAT TTA AAT CTT  
**GGG** GTT **GGG** **AGA** GCA **AGA** **AGA** AAG CTT AAG GTT TGT CAA **ACG** **AGA** GCT AGC TTT TGG  
GAA TCC ATC **AGA** TCC **GGG** TTT TTG AAG AAT AAT ACA ACA CAA GTT ATT GAA CCA **CCC** TCC  
ACA AAC CAA GAA GAG GAA GAA CCA TTG **CCC** GAG GAG TTT GTT CTT GTT GAA AAG  
ACT **CUA** GCT GAC **GGA** GCA ATT GAA CAG **AUA** **AUA** TTT TCT TCA GGT **GGA** GAT GTT GAT GTC  
TAT GAT CTC CAG GCC TTA TGT GAT AAG GTG GGC TGG **CCC** **CGA** **AGG** CCT **CUA** TCA  
AAA **CUA** TCT GCA GCT TTA AAA AAT AGC TAC ATG GTA GCC ACA TTG CAT TCC TTA **AGG** AAA  
TCA CCT **GGA** GAA GAG **AGG** AAT GAG CAA AAG AAG CTT ATT GGC ATG GCC CGT GCT ACA  
TCA GAT CAT GCC TTC AAT GCC ACA ATT TGG GAT GTC CTT GTT GAT CCT TCA TAT CAG GGC  
CAG **GGA** CTT **GGA** AAG GCC CTC GTT GAA AAG ACT **AUA** **AGA** GCT CTT CTG CAA **AGG** GAC  
ATT GGC AAT **AUA** ACA CTC TTT GCA GAT AGC CAA GTG GTG GAG TTC TAT **CGA** AAT TTA GGT  
TTT GAA CCT GAC CCA GAG GGC ATC AAA GGT ATG TTC TGG CAC CCA **AGG** TAT TAG

稀有密码子均匀分布，而且没有连续三个的稀有密码子排列，对翻译蛋白的可溶性影响不大

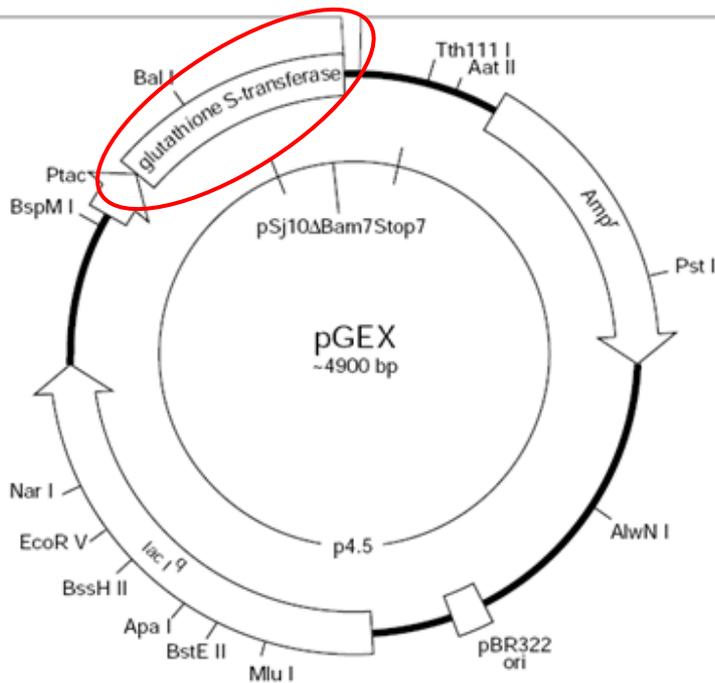


## 5.3 Plasmid profile of pGEX-4T-1

### pGEX-4T-1 (27-4580-01)

Thrombin

Leu	Val	Pro	Arg	Gly	Ser	Pro	Glu	Phe	Pro	Gly	Arg	Leu	Glu	Arg	Pro	His	Arg	Asp	
CTG	GTT	CCG	CGT	GGA	TCC	CCG	GAA	TTC	CCG	GGT	CGA	CTC	GAG	CGG	CCG	CAT	CGT	GAC	TGA
				BamH I		EcoR I		Sma I		Sal I		Xho I		Not I					Stop codons



含有GST促溶标签，有助于重组蛋白的可溶性表达



## 5.4 Analysis of restriction enzyme sites of target genes

Vector NTI - [NEWMOL]

File Edit View Primer Design Analyses Cloning Gel Analysis List Window Align Assemble Tools Help

ATC TRG ATC TRG B I U Courier New 13

- NEWMOL
- General Description
- Standard Fields
- Comments
- Annotations
- Feature Map
- Restriction/Methylation Map

NEWMOL  
741 bp

1	ATGCAAAACC ACAGCGCCCT CCTCTCCACT TCTCAGTACC CCATTITTC TTTTGGTITG AATTGTGAGC GATCAAATTT ATCATTGTTC CTTTGCAATT TACGTTTGGG TGTCGCGGGA GGAGAGGTGA AGAGTCATGG GGTAAAAGAG AAAACCAACG TTAACACTGC CTAGTTTAAA TAGTAACRAG GSAACGTTAA
101	TAAATCTTGG GGTGGGGAGA GCAAGAAGAA AGCTTAAAGT TTGTCAAACG AGAGCTAGCT TTTGGGAATC CATCAGATCC GGGTTTTTGA AGAATAATAC ATTTAGAACC CCAACCCTCT CGTTCITCTT TCGAATTCOA AACAGTITGC TCTCGATCGA AAACCCCTAG GTAGTCTAGG CCAAAAAACT TCTTATTATG
201	AACACRAGTT ATTGAACCAC CCTCCACAAA CCAAGAAGAG GRAGAACCAT TGCCCGRAGA GTTTGTTCTT GTTGAAAGA CTCTAGCTGA CGSAGCAATT TTGTGTTCAA TAACITGGTG GGAGGTGTTT GGTTCITCTC CTCTCTGGTA ACGGGCTCCT CAACCAAGAA CACTTTTCTC GAGATCGACT GCCTCGTTAA
301	GAACAGATAA TATTTTCTTC AGGTGGAGAT GTTGATGTCT ATGATCTCCA GGCCTTATGT GATRAGGTGG GCTGGCCCCG AAGGCCTCTA TCAAAACTAT CTTGCTAATT ATAAAGAAG TCCACCTCTA CCACTACAGA TACTAGAGGT CCGGAATACA CTATTCCACC CGACCGGGGC TTCCGGAGAT AGTTTTGATA
401	CTGCAGCTTT AAAAAATAGC TACATGTTAG CCACATTGCA TTCTTAAAGG AAATCACCTG GAGAAGAGAG GAATGAGCAA AAGAGCTTAA TTGGCATGGC GACGTCGAAA TTTTTATCG ATGTACCAIC GGTGTAACGT AAGSAAITCC TTAGTGGAC CICTTCTCTC CITACTCGTI TTCTTCGARI AACCGTACCG
501	CCGTGCTACA TCAGATCATG CCTTCAATGC CACAATTGG GATGTCCTTG TTGATCCTTC ATATCAGGGC CAGGGACTTG GAAAGGCCCT CGTTGAAAAG GGCACGATGT AGTCTAGTAC GSAAGTITACG GTGTAAACC CTACAGGAAC AACTAGGAAG TATAGTCCCG GTCCCTGAAC CTTTCCGGGA GCAACTTTTC
601	ACTATAAGAG CTCTCTGCA AAGSGACATT GSCAATATAA CACTCTTTCG AGATAGCCAA GTGGTGGAGT TCTATCGAAA TTTAGTTTTT GAACCTGACC TGATATTTTC GAGAAGACGT TTCCCTGTAA CCGTATATTT GTGAGAAGCG TCTATCGGTT CACCACCTCA AGATAGCTTT AAATCCAAA CTITGGACTGG
701	CAGAGGGCAT CAAAGGTATG TTCTGGCACC CAAGTATTA G GTCTCCGTA GTTCCATAC AAGACCGTGG GTTCCATAAT C



中国农业科学院郑州果树研究所

Zhengzhou Fruit Research Institute, CAAS

请各位老师、同学批评指导

谢谢!