

# Bioinformatics and Expression Analysis of GST Gene Encoding Glutathione S-Transferase in *Citrus maxima* cv. Shatian Yu

Qiong Luo, Jun Xiao, Shaomei Li, Xiaoxia Long, Huaying Liu, Xinmin Qin\*

College of Life Science, Guangxi Normal University, Guilin Guangxi  
Email: \*xm Qin@mailbox.gxnu.edu.cn

Received: Mar. 7<sup>th</sup>, 2017; accepted: Mar. 25<sup>th</sup>, 2017; published: Mar. 28<sup>th</sup>, 2017

## Abstract

In this paper, the transcriptome of the self-pollinated style and cross-pollinated style of *Citrus maxima* cv. Shatian Yu were sequenced by high-throughput sequencing technology. The glutathione S-transferase gene (*CmGST*) sequence of *Citrus maxima* cv. Shatian Yu was obtained through differential analysis method. The full length of *CmGST* gene was 872 bp (GenBank accession number: KY612456), and the ORF (Open Reading Frame) length was 657 bp, encoding a protein polypeptide of 219 amino acids with a predicted molecular weight of 25.00 kDa and pI of 5.48. Bioinformatics analysis showed that the *CmGST* protein contained the Tau GST-specific N-terminal domain (G site) and the C-terminal domain (H site). The homology analysis of amino acid sequence indicated that the glutathione S-transferase protein shared high homology with glutathione S-transferase protein of *Citrus clementina* (98%). This work provides a systemic sequence analysis of the glutathione S-transferase protein of *Citrus maxima* cv. Shatian Yu. It will provide the useful reference for further investigation of its structure and function.

## Keywords

*Citrus maxima* cv. Shatian Yu, Glutathione S-Transferase Gene, Bioinformatics

# 沙田柚谷胱甘肽转移酶基因(*CmGST*)的生物信息学及表达分析

罗琼, 肖君, 李少梅, 龙小霞, 刘华英, 秦新民\*

\*通讯作者。

文章引用: 罗琼, 肖君, 李少梅, 龙小霞, 刘华英, 秦新民. 沙田柚谷胱甘肽转移酶基因(*CmGST*)的生物信息学及表达分析[J]. 植物学研究, 2017, 6(2): 47-54. <https://doi.org/10.12677/br.2017.62008>

广西师范大学生命科学学院, 广西 桂林  
Email: xmqin@mailbox.gxnu.edu.cn

收稿日期: 2017年3月7日; 录用日期: 2017年3月25日; 发布日期: 2017年3月28日

## 摘要

利用高通量测序对沙田柚自交和异交花柱进行转录组测序, 通过差异分析得到了沙田柚谷胱甘肽转移酶基因(*CmGST*)序列。*CmGST*全长为872 bp (GenBank登录号为KY612456), 开放阅读框为657 bp, 编码219个氨基酸, 推测蛋白质分子量为25.00 kDa, 理论等电点pI为5.48。利用生物信息分析软件对*CmGST*基因进行同源性比对和系统进化分析表, 结果显示该基因编码的氨基酸具有谷胱甘肽S转移酶典型的C端与N端双结构域, 属于GST Tau家族蛋白; 与克莱门柚(*Citrus clementina*)的相似度约为98%。这些分析结果可为今后深入研究该蛋白的结构特征和功能提供参考。

## 关键词

沙田柚, 谷胱甘肽转移酶基因, 生物信息学

Copyright © 2017 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

植物自交不亲和性(Self-incompatibility, SI)是指能产生具有正常功能且同期成熟的雌雄配子的雌雄同体植物, 在自花授粉或相同基因型异花授粉时无法受精的现象[1]。据报道, 自然界中已在74科250属约3000种以上的显花植物中发现了自交不亲和现象[2] [3], 特别是十字花科、禾本科、豆科、蔷薇科、茄科、菊科、罂粟科、石蒜科等。

SI是目前所知的最主要的一种显花植物控制受精的机制, 是由雌蕊细胞和花粉之间相互作用产生的生理反应。这两种独特细胞间的接触是受空间和时间的授粉行为控制的, 这远不同于其他细胞间的信号转导[4]。因此, SI不仅为研究植物生殖细胞间信号识别和转导、细胞间相互作用和基因时空表达提供了一种理想模型, 而且在作物遗传改良和杂种优势利用上具有重要价值。理论和应用上的双重价值, 使得植物自交不亲和性成为近年来植物分子生物学和作物育种学研究的重要热点之一[5] [6]。

沙田柚为无患子目、芸香科、柑橘属植物, 为严格的配子体自交不亲和植物。关于沙田柚自交不亲和机理的研究取得一定得进展, 薛妙男等确定沙田柚为配子体自交不亲和类型, 花粉管在花柱1/2处生长受到抑制; 通过免疫胶体金技术, 确定自交1-3天花柱通道细胞S-蛋白产生部位及分布位置[7] [8] [9] [10]。杨继华等[11] [12]通过双向电泳鉴定沙田柚得到花柱特异性蛋白, 并对该蛋白的分子量、等电点以及N-末端氨基酸序列进行了测定。秦新民等分离和鉴定沙田柚花粉管特异性蛋白; 确定了特异蛋白在花粉管中的产生部位以及分布位置[13] [14] [15]。

为了进一步探讨沙田柚自交不亲和的机理, 我们对自交和异交花柱进行了转录组测序, 通过自交与异交花柱差异基因的比对, 获得了一个沙田柚谷胱甘肽转移酶基因, 并对该基因编码的蛋白质进行了生物化学特征分析, 旨在为沙田柚自交不亲和分子机理的深入研究提供帮助。

## 2. 材料与方法

### 2.1. 实验材料

沙田柚实验材料采自广西灵川县潮田乡大山口村果园十年生结果树。在盛花期对沙田柚样树进行人工自交(沙田柚 × 沙田柚)授粉和异交授粉(酸柚 ♂ × 沙田柚 ♀), 分别收集自交和异交 1~3 d 的授粉花柱以及当天未开花的花柱, 立即放入液氮中保存, 并放入-80℃超低温冰箱备用。

### 2.2. 方法

#### 2.2.1. RNA 的提取、建库和测序

总 RNA 的提取按改良 Trizol 法进行[16], 建库和测序同参考文献[17]。

#### 2.2.2. 序列分析和系统树构建

使用 DNAMAN、Finder、TMPRED、SWISS-MODEL、NetPhos2.0、SignalP4.1 等软件进行序列的开放阅读框(ORF)、编码氨基酸、编码蛋白的理化分析。将测定的基因序列与 GenBank 中的 13 种植物的同源序列进行比对, 基于氨基酸序列谷胱甘肽 S 转移酶基因的系统进化树用 DNAMAN 软件进行构树。

## 3. 结果

### 3.1. 基因的生物信息学分析

谷胱甘肽 S 转移酶基因(Unigene16804\_All)全长 872 bp (GenBank 登录号为 KY612456), 包含一个 657 bp 的开放阅读框(ORF)编码 219 个氨基酸(图 1)。

### 3.2. 编码蛋白质的分析及疏水性的预测

通过在线软件 protparam 分析(<http://www.expasy.org/tools/protparam.html>), 该基因基因编码的蛋白质分子质量为 25.00 kDa; 理论等电点为 5.48; 共由 3517 个原子组成, 分子式为  $C_{1149}H_{1753}N_{283}O_{322}S_{10}$ ; 不稳定指数为 33.42 (<40), 为稳定蛋白。

用 DNAMAN 预测该蛋白质的疏水性: 该基因编码的蛋白质疏水性最大值为 3.04, 亲水性最小值为 -3.11。综合表明该蛋白质疏水性平均值为 -1.93, 为亲水蛋白(图 2)。

### 3.3. 跨膜预测

跨膜预测的结果显示该基因编码的蛋白质存在着 2 个跨膜区。跨膜方向由内向外有 149~167 位氨基酸; 跨膜方向为由外向内的有 6~22 和 143~165 位氨基酸(图 3)。

### 3.4. 保守结构域分析

该蛋白质有两个保守结构域, 一是 GST\_C\_superfamily (谷胱甘肽转移酶超家族), 位于该蛋白质的第 85~215 位, 二是 Thioredoxin\_like superfamily (硫氧还蛋白超家族), 位于该蛋白质的第 1~75 位(图 4)。

### 3.5. 磷酸化位点预测

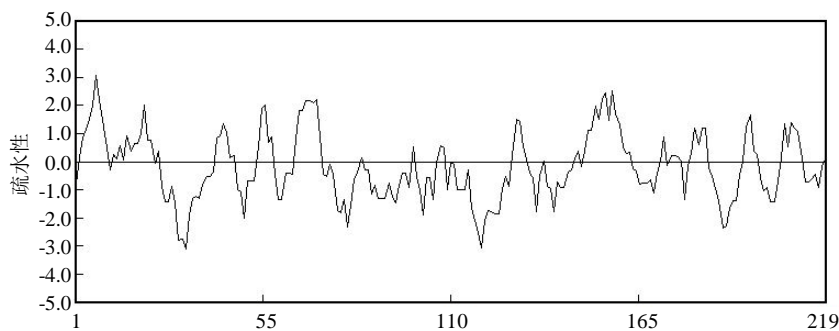
该基因编码的蛋白质可能的磷酸化位点共有 16 个, 其中 Ser 磷酸化位点 11 个, 分别位于肽链的 32、67、82、93、115、149、174、190、194、196、198 位; Thr 磷酸化位点 1 个, 位于肽链第 169 位; Tyr 磷酸化位点共有 5 个, 分别位于肽链 30、90、101、107 和 206 位。表明该蛋白的磷酸化以 Ser 磷酸化为主, 兼有 Thr 和 Tyr 磷酸化(图 5)。

```

1      ACCAAATTAAGCAACTAAATATTAATTAACCTAATTGTGCAATAAAAAATGGCTGAGAAT
                                         M A E N
61     GTGGTTTTATTGGGCTTTTGGCCAAGTCTTTTGGCAATGAGAGTGAAAAATGCACTGGCC
      V V L L G F W P S S F A M R V K I A L A
121    GAGAAGGGAATCAAATATGAATCAAAGAAGAGAACCTGTTCGATAAGACCCCTTACCTC
      E C G I K Y E S K E E N L F D K S P L L
181    CTTGAGATGAATCCGGTTGACAAGAAAATCCCTGTCTGATTCATGACGGGAAACCCATT
      L E M N P V D K K I P V L I H D G K P I
241    TGTGAATCCCTTATTATCATTCAATACATGTGATGAAGTTTGGCAGCACAAAATCCCTCTT
      C E S L I I I Q Y I D E V W Q H K S P L
301    ATGCCCTTCGATCCCTTACCTTCGATCTCAAGCCAGGTTTGGGCTGATTATGTGACAAA
      M P S D P Y L R S Q A R F W A D Y V D K
361    AATATATACGGTATTGGGAAGAGGATTTGGTCGGGAAAGGAGAAGATCAAGAGCCAGCA
      N I Y G I G K R I W S G K G E D Q E A A
421    AAGAAGGAGATGATAGGGTTTTTGAAGAACATGGAAGCAGAGCTCGGAGACAAACACTTC
      K K E M I G F L K N M E A E L G D K H F
481    TTCGGCGGAGAGATTTGGATTTGTAGATGTGGCGTTGGTCCATTCACAACATGGTTT
      F G G E S I G F V D V A L V P F T T W F
541    TACACGTACGAGACTTTTGGGAACCTTAGCATTGAAGCAGAGTGCCCAAGCTTATTCCG
      Y T Y E T F G N F S I E A E C P K L I A
601    TGGGCCAAGAGGTGCTCACAAGGAGAGTGTGTCTGCCCTCCCTGTGATCCTCACCGAC
      W A K R C S Q K E S V S A S L A D P H D
661    ATCTACGACTTGTCTTGGGCTTAGAATGAAATATGGGATTCAGTAGACAGGCAGTCCA
      I Y D F A L G L R M K Y G I Q *
721    GACAGTGTGATTTATATGGTTCCAATCCAGTGTGTGACGACATTTTATTATTA
781    GTTCCAAACTTCTGTCTTATTAGAGTTTAAAATCCCAAATAACAAAAACTATCTG
841    CATGTGTGTATGTAATGGCTGGTAAACAAAT
    
```

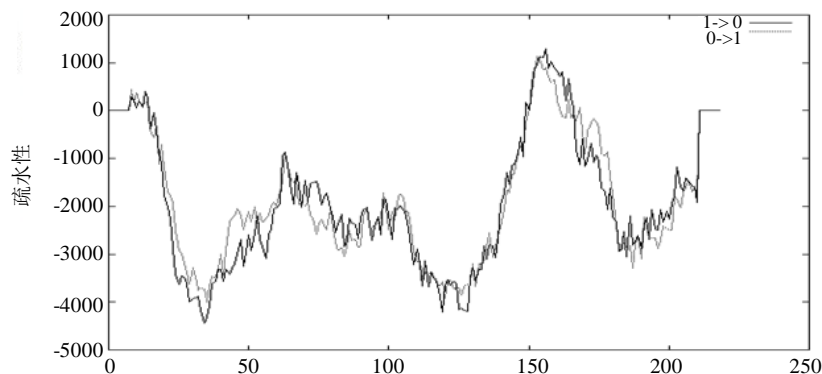
**Figure 1.** Nucleotide sequence and amino acid sequence of Uni-gene16804\_All gene

**图 1.** Unigene16804\_All 序列和推测氨基酸序列



**Figure 2.** Hydrophobicity analysing of Unigene16804\_All

**图 2.** Unigene16804\_All 蛋白疏水性分析



**Figure 3.** Result of TMpred Prediction of Unigene16804\_All

**图 3.** Unigene16804\_All 跨膜区预测

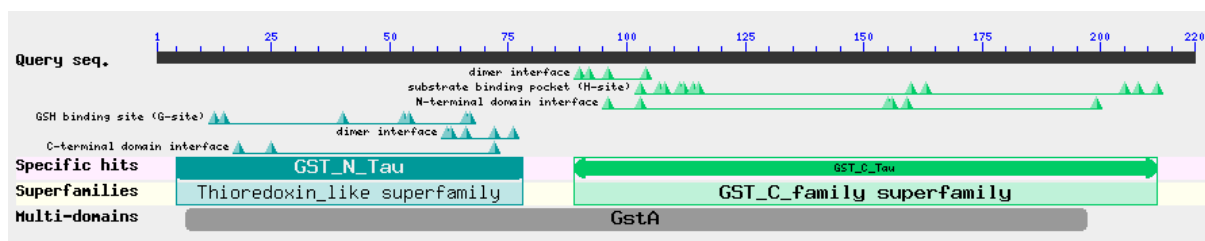


Figure 4. Conserved domains of porcine Unigene16804\_All

图 4. Unigene16804\_All 蛋白保守结构域

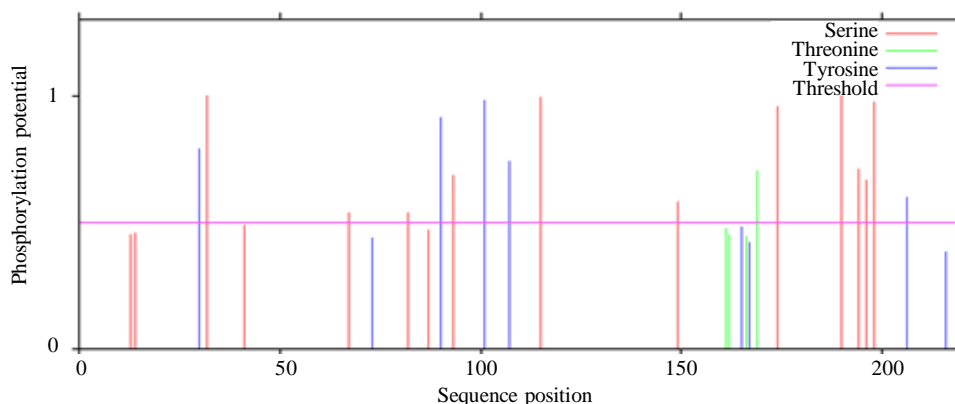


Figure 5. Phosphorylation site analysis of Unigene16804\_All

图 5. Unigene16804\_All 蛋白的磷酸化位点分析

### 3.6. 信号肽预测

利用 SignalP4.1 软件对沙田柚谷胱甘肽转移酶基因编码的氨基酸序列进行分析, 结果表明该蛋白质没有明显的信号肽。

### 3.7. 二级结构和三级结构预测

通过在线分析软件 Predict Protein (<http://www.predictprotein.org/>)对谷胱甘肽转移酶蛋白二级结构进行预测, 结果显示该蛋白含有螺旋、折叠和无规则卷曲, 所占比例分别是 54.34%、5.48%、40.18%。

利用 SWISS-Model 对沙田柚谷胱甘肽转移酶蛋白三级结构进行预测, 结果表明: 该蛋白三级结构含有 14 个  $\alpha$ -螺旋, 7 个  $\beta$ -折叠, 其间由无规则卷曲连接(图 6)。

### 3.8. 同源性分析

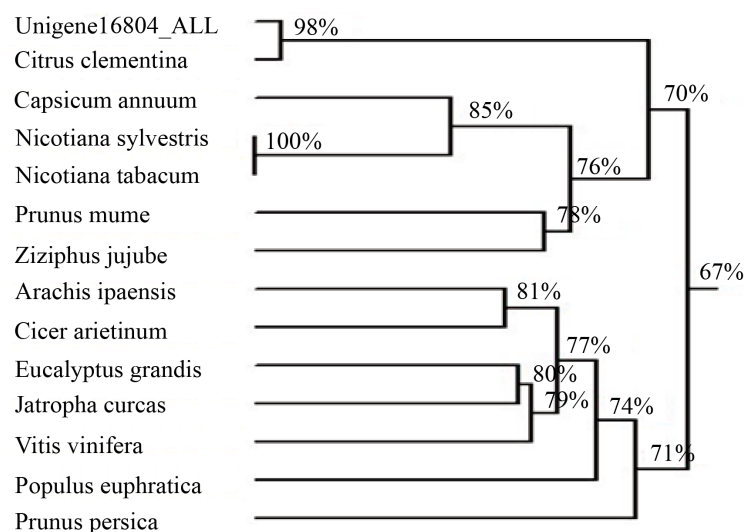
从 GenBank 数据库中下载 13 种植物谷胱甘肽转移酶基因编码的氨基酸序列, 序列进行比对结果表明沙田柚谷胱甘肽转移酶基因编码的氨基酸与芸香科的克莱门柚(*Citrus clementina*, XP\_006432070.1)谷胱甘肽转移酶蛋白的同源性为 98%。利用 DNAMAN 构建系统发育树, 结果表明沙田柚谷胱甘肽转移酶基因编码的蛋白质与芸香科的克莱门柚(*Citrus clementina*)有很近的亲缘关系, 属于同一个进化分支(图 7)。

## 4. 结论与讨论

氨基酸序列同源性分析表明, 本文所克隆的沙田柚基因(Unigene16804\_All)与克莱门柚谷胱甘肽转移酶的同源性均为 98%, 通过对该基因编码蛋白的功能结构域分析, 发现该蛋白具有谷胱甘肽转移酶保守结构域, 表明 Unigene16804\_All 属于谷胱甘肽转移酶超家族基因。



**Figure 6.** The tertiary structure of Unigene16804\_All  
**图 6.** Unigene16804\_All 蛋白的三级结构



Unigene16804\_ALL (沙田柚), *Citrus clementina* (克莱门柚, XP\_006432070.1), *Capsicum annuum* (辣椒, XP\_016556906.1), *Nicotiana sylvestris* (美花烟草, XP\_009776045.1), *Nicotiana tabacum* (烟草, XP\_016433645.1), *Prunus mume* (梅, XP\_008240792.1), *Ziziphus jujube* (枣, XP\_015874673.1), *Arachis ipaensis* (野生花生 XP\_016170182.1), *Cicer arietinum* (鹰嘴豆, XP\_004488707.1), *Eucalyptus grandis* (巨桉, XP\_010062724.1) *Jatropha curcas* (麻风树, XP\_012087687.1), *Vitis vinifera* (葡萄, XP\_002264437.1), *Populus euphratica* (胡杨, XP\_011036235.1), *Prunus persica* (桃, XP\_007213037.1)

**Figure 7.** Phylogenetic tree based on amino acid sequences of the *Glutathione S-transferase* protein  
**图 7.** 基于氨基酸序列的谷胱甘肽转移酶蛋白系统发育树

植物的谷胱甘肽转移酶(GST)是一个多基因家族,植物的GST分为 $\varphi$ 、 $\tau$ 、 $\zeta$ 、 $\theta$ 、 $\lambda$ 和脱氢抗坏血酸还原酶(DHARs)6类[18][19]。GST在植物的初级代谢、二级代谢、胁迫耐受和细胞信号转导中行使功能,从而影响植物的生长发育[20]。此外,GST家族蛋白还参与了植物黄酮类物质的积累和运转[21]。例如,牵牛花中An9基因编码了一个Phi类的GST,能与花青素形成复合物并参与其转运[22],而拟南芥中同属于Phi类的TF19基因则同时参与了花青素和浓缩单宁的转运[23]。类黄酮除了具有保护植物免受紫外线伤害、抵抗病原菌的侵害等功能外,类黄酮对植物生殖发育也具有重要的影响,如参与花粉-雌蕊的相互作用的过程[24]促进花粉萌发,及影响花粉管的极性生长中起作用[25][26]。

本文克隆的谷胱甘肽转移酶基因在沙田柚未授粉, 以及自交花柱和异交花柱中的表达存在较为明显的差异: 未授粉花柱中该基因的表达量(FPKM)为 0.106, 自花授粉 1, 2, 3 d 花柱中的表达量(FPKM)分别为 3.32、5.98、2.45, 异花授粉 1, 2, 3 d 花柱中基因的表达量(FPKM)则分别为 2.26、3.51 和 14.16。从上述结果可以看出无论自交授粉还是异交授粉, 第 1 d 花柱中谷胱甘肽转移酶基因的表达都表现为迅速升高, 在自交授粉状态下, 基因的表达第 2 d 稍许升高, 第 3 d 则下降。而异交授粉过程中, 该基因的表达则表现为持续升高。

沙田柚谷胱甘肽转移酶基因的这种表达模式与其参与花粉-雌蕊的相互作用和影响花粉管的极性生长的功能相符。薛妙男等[7]观察到沙田柚自、异花授粉的 1 d, 其花粉的萌发和花粉管的生长行为相同。花粉在柱头上经水合作用开始萌发, 4~8 h 花粉管通过柱头乳突细胞间隙进入柱头, 16~24 h 即可进入花柱道。随后, 自交与异交授粉花粉管的生长出现明显差异, 自花授粉的花粉管 2~3 d 时在花柱的 1/2 处生长受阻, 而异交授粉花粉管则能正常生长, 进入子房完成受精。谷胱甘肽转移酶基因在自花和异花授粉的 1 d 均高表达, 与其参与花粉-雌蕊的相互作用功能相关。随后该基因的表达与花粉管在花柱中的生长密切相关, 异花授粉 2~3 d 内基因表达活性急剧上升, 而同期自花授粉状态下, 该基因的表达呈现下降, 与之对应的是花粉管生长受阻, 不能完成受精。因此可以推断谷胱甘肽转移酶可能参与了沙田柚授粉和自交不亲和性过程。但其机理有待进一步收入研究。

## 基金项目

国家自然科学基金(31360477); 广西教育厅项目(2013YB036)。

## 参考文献 (References)

- [1] Takayama, S. and Isogai, A. (2005) Self-Incompatibility in Plants. *Annual Review of Plant Biology*, **56**, 467-489. <https://doi.org/10.1146/annurev.arplant.56.032604.144249>
- [2] Lewis, D. (1979) *Sexual Incompatibility in Plants*. Edward Arnold Limited, London.
- [3] Brewbaker, J.L. (1959) Biology of the Angiosperm Pollen Grain. *The Indian Journal of Genetics & Plant Breeding*, **19**, 121-133.
- [4] 蓝兴国, 于晓敏, 李玉花. 配子体自交不亲和信号转导的研究进展[J]. 遗传, 2005, 27(4): 677-685.
- [5] 安雯, 曹阳, 李俞涛, 等. 植物自交不亲和性的研究进展与展望[J]. 信阳师范学院学报, 2009, 22(4): 630-634.
- [6] 胡彬, 蒋建雄, 易自力. 植物配子体自交不亲和和机制研究进展[J]. 中国农学通报, 2012, 28(18): 168-173.
- [7] 薛妙男, 陈腾士, 杨继华. 沙田柚自交和异交亲和性观察[J]. 园艺学报, 1995, 22(2): 127-132.
- [8] 薛妙男, 杨继华. 沙田柚花粉管在花柱中的生长途径及其识别[J]. 广西师范大学学报, 2001, 19(2): 60-66.
- [9] 薛妙男, 李义平, 张杏辉, 等. 沙田柚自交花柱 S1-RNase 的免疫胶体金定位[J]. 广西师范大学学报, 2000, 18(1): 81-84.
- [10] 薛妙男, 李义平, 张杏辉, 等. 沙田柚自交花柱中识别蛋白的免疫金定位[J]. 园艺学报, 2001, 28(1): 59-61.
- [11] 杨继华, 李红艳, 薛妙男. 沙田柚花柱 S-糖蛋白的分离与鉴定[J]. 广西师范大学学报, 2000, 18(4): 66-70.
- [12] 杨继华, 尧桂荣, 薛妙男. 沙田柚花柱 S-糖蛋白的纯化和 N-端序列测定[J]. 广西师范大学学报, 2001, 19(1): 72-79.
- [13] 秦新民, 李惠敏, 薛妙男, 等. 沙田柚自交、异交花粉管蛋白的双向电泳分析[J]. 广西植物, 2004, 24(6): 566-569.
- [14] 秦新民, 莫花浓, 万珊, 等. 沙田柚花粉管特异蛋白的免疫细胞化学研究[J]. 广西师范大学学报, 2008, 26(4): 113-115.
- [15] 秦新民, 莫花浓, 石菁萍, 等. 沙田柚花粉管 S1-RNase 免疫胶体金定位研究[J]. 广西农业科学, 2009, 40(5): 483-485.
- [16] Bariola, P.A., Howard, C.J., Taylor, C.B., et al. (1994) The *Arabidopsis thaliana* Ribonuclease Gene RNS1 Is Tightly Controlled in Response to Phosphate Limitation. *The Plant Journal*, **6**, 673-685.

<https://doi.org/10.1046/j.1365-313X.1994.6050673.x>

- [17] 秦新民, 张渝, 刘玉洁, 等. 沙田柚 S-RNase 基因的克隆及序列分析[J]. 广西师范大学学报(自然科学版), 2015, 33(1): 139-145.
- [18] Dixon, D.P., Davis, B.G. and Edwards, R. (2002) Functional Divergence in the Glutathione Transferase Superfamily in Plants. *Journal of Biological Chemistry*, **277**, 30859-30869. <https://doi.org/10.1074/jbc.M202919200>
- [19] Moons, A. (2005) Regulatory and Functional Interactions of Plant Growth Regulators and Plant Glutathione S-Transferases (GSTs). *Vitamins & Hormones*, **72**, 155-202.
- [20] 胡廷章, 周大祥, 罗凯. 植物谷胱甘肽转移酶的结构与功能及其基因表达[J]. 植物生理学通讯, 2007, 43(1): 195-200.
- [21] Grotewold, E. (2001) Subcellular Trafficking of Phytochemicals. *Recent Research Developments in Plant Physiology*, **2**, 31-48.
- [22] Alfenito, M.R., Souer, E., Goodman, C.D., et al. (1998) Functional Complementation of Anthocyanin Sequestration in the Vacuole by Widely Divergent Glutathione Transferases. *Plant Cell*, **10**, 1135-1150. <https://doi.org/10.1105/tpc.10.7.1135>
- [23] Kitamura, S., Shikazono, N. and Tanaka, A. (2004) Transparent Testa 19 Is Involved in the Accumulation of Both Anthocyanins and Proanthocyanidins in *Arabidopsis*. *Plant Journal*, **37**, 104-114. <https://doi.org/10.1046/j.1365-313X.2003.01943.x>
- [24] Hahlbrock, K. and Scheel, D. (1989) Physiology and Molecular Biology of Phenylpropanoid Metabolism. *Annual Review of Plant Biology*, **40**, 347-369. <https://doi.org/10.1146/annurev.pp.40.060189.002023>
- [25] Mo, Y., Nagel, C. and Taylor, L.P. (1992) Biochemical Complementation of Chalcone Synthase Mutants Defines a Role for Flavonols in Functional Pollen. *Proceedings of the National Academy of Sciences of the United States of America*, **89**, 7213-7217. <https://doi.org/10.1073/pnas.89.15.7213>
- [26] Taylor, L.P. and Grotewold, E. (2005) Flavonoids as Developmental Regulators. *Current Opinion in Plant Biology*, **8**, 317-323.

期刊投稿者将享受如下服务:

1. 投稿前咨询服务 (QQ、微信、邮箱皆可)
2. 为您匹配最合适的期刊
3. 24 小时以内解答您的所有疑问
4. 友好的在线投稿界面
5. 专业的同行评审
6. 知网检索
7. 全网络覆盖式推广您的研究

投稿请点击: <http://www.hanspub.org/Submission.aspx>

期刊邮箱: [br@hanspub.org](mailto:br@hanspub.org)