

无籽蜜柚蔗糖合成酶(SUS)和蔗糖转化酶(INV)基因家族生物信息学及表达分析

邓舒雅, 麦贻婷, 陈惠萍, 钮俊*

海南大学热带农林学院, 海口570228

摘要: 蔗糖是光合作用的最终产物, 通过韧皮部运输到植物的非光合组织中, 蔗糖的分解主要由蔗糖合成酶(SUS)和蔗糖转化酶(INV)催化。为了探究SUS和INV基因家族在无籽蜜柚汁胞发育中的功能, 本文通过HPLC研究, 结果表明阿拉伯糖和木糖在汁胞早期发育中含量较高, 而果糖、葡萄糖和蔗糖的含量随成熟过程显著增加。基于拟南芥数据库中SUS和INV蛋白序列, 与柚子的基因组数据进行同源比对, 分别筛选获得6个和12个具有完整开放阅读框(ORF)的Cg-SUS和Cg-INV基因。根据其在染色体上分布, 分别命名为Cg-SUS1~6和Cg-INV1~12, 并对其蛋白理化性质、内含子与外显子、系统进化树和蛋白功能结构域等方面加以分析。数字表达谱分析初步确定无籽蜜柚汁胞在发育过程中Cg-SUS2、Cg-SUS3和Cg-INV6转录较为活跃, 对其进行qRT-PCR分析, 结合各类糖含量的检测结果与基因表达做相关性分析, 结果显示Cg-SUS2表达量与木糖、阿拉伯糖含量之间存在极显著正相关性, 因此Cg-SUS2可能涉及阿拉伯糖和木糖的生物合成; Cg-SUS3和Cg-INV6表达量与葡萄糖、果糖、蔗糖含量之间存在显著正相关性, 说明Cg-SUS3可能参与到蔗糖的降解途径中, 从而促使转运过来的蔗糖生成UDP-葡萄糖和果糖, Cg-INV6可能是液泡己糖累积的关键酶。

关键词: 无籽蜜柚; 汁胞发育; 蔗糖合成酶; 蔗糖转化酶

柚子(*Citrus grandis*)是芸香科(Rutaceae)柑桔属(*Citrus*)植物, 主要分布于中国、泰国、马来西亚和印度等国, 而中国的种植面积居世界之首。柚子是一种具有多种营养和药用价值的树种, 富含维生素C和P、类黄酮、果胶和矿物质(包括钙、钾和铬)。它被医学界认为是治疗心血管和肾脏疾病的最有效的水果(Wu等2011)。无籽蜜柚(又名马来西亚蜜柚, *Citrus grandis* 'Seedless')于2002年从马来西亚引进在海南种植, 通常为常绿乔木, 成熟时果皮颜色为黄绿色、皮薄、无籽。目前国内有关无籽蜜柚研究主要集中在传统育种和丰产栽培技术来提高蜜柚产量等方面, 其汁胞发育过程中相关蔗糖代谢方面研究鲜见报道。

在高等植物中, 蔗糖(sucrose, Suc)是光合作用的最终产物, 是从光合作用源组织通过韧皮部运输到植物的非光合组织的主要碳源形式。蔗糖的分解由蔗糖合成酶(sucrose synthase, SUS, EC 2.4.1.13)或蔗糖转化酶(invertase, INV, EC 3.2.1.26)催化(Ruan 2014)。INV可以水解蔗糖为葡萄糖(glucose, Glu)和果糖(fructose, Fru), 而SUS可以催化蔗糖转化为UDP-葡萄糖(UDPG)和果糖(Coleman等2009)。根据亚细胞位置的不同, INV可分为细胞壁转化酶(CWIN)、液泡转化酶(VIN)和细胞质转化酶(CIN)

(Jin等2009), 根据其生化活性和pH的不同又可分为酸性转化酶和中性/碱性转化酶(Ruan等2010)。与INV相比, SUS是植物中广泛存在的一种UDP-糖基转移酶, 是催化蔗糖代谢的关键酶, 能更快地催化蔗糖的降解, 并在种子和果实的发育中发挥作用(Xu等2012)。

目前, 关于SUS和INV基因家族的研究主要集中在拟南芥(*Arabidopsis thaliana*), 已证实拟南芥中含有6个*At-SUS*和11个*At-INV*基因家族成员, 它们有不同的表达模式和功能(Baud等2004)。其中, *At-SUS1*和*At-SUS4*表现出不同表达的密切相关, *At-SUS2*是种子成熟的标志, *At-SUS3*是脱水的标志, *At-SUS5*和*At-SUS6*构成一个独立的群体, 其功能仍然未知(Baud等2004)。关于SUS基因家族的研究还涉及玉米(*Zea mays*) (Duncan等2006)、桃树(*Prunus persica*) (Zhang等2015)等; 关于INV基因家族的研究还涉及番木瓜(*Carica papaya*)、木薯(*Manihot esculenta*) (Yao等2015)等, 然而关于SUS和INV基因家族的研究在柚子中鲜见报道。

收稿 2018-07-16 修定 2018-09-29

资助 海南省自然科学基金面上项目(317045)和海南大学科研启动基金项目[KYQD(ZR)1701]。

* 通讯作者(niujun5555@163.com)。

基于拟南芥数据库中SUS和INV蛋白序列,与柚子基因组数据库进行同源比对,分别筛选获得6个和12个具有完整开放阅读框(ORF)的*Cg-SUS*和*Cg-INV*基因。根据其染色体分布位置,依次命名为*Cg-SUS1~6*及*Cg-INV1~12*。利用生物信息学方法,进行基因及编码的蛋白组成、进化树和蛋白功能结构域的分析。同时,基于先前无籽蜜柚汁胞转录组数据,开展*Cg-SUS*及*Cg-INV*基因的数字表达谱分析,并对其关键基因进行qRT-PCR检测。通过7个不同发育期无籽蜜柚汁胞可溶性糖含量的检测,对各类糖含量与无籽蜜柚*Cg-SUS2*、*Cg-SUS3*、*Cg-INV6*基因表达做相关性分析,为无籽蜜柚SUS和INV基因家族表达和功能提供关键信息。

1 材料与方法

1.1 植物材料

试验果实采摘于中国海南省海口市澄迈县无籽蜜柚(*Citrus grandis* (L.) Osbeck 'Seedless')基地(地理坐标约为19°73'N 109°99'E),基地土壤类型为砖红壤,砧木为本地酸柚。于2017年2~9月(落花后60、90、120、150、180、210和240 d)在园内选取树势与负载量相对一致、生长一致已进入结果期的无籽蜜柚为研究对象,第一次于60 d采果,之后每30 d摘取样果一次,共取样7次,选留无损伤、无虫斑、带有完好果柄的果实。每次采集样品迅速液氮冷冻后置于-80°C冰箱保存,供后续试验要求。

1.2 无籽蜜柚提取可溶性糖及测定方法

糖分提取:各时期取0.7 g汁胞样品放置在105°C烘箱中杀青处理30 min,然后在80°C中将其烘干至恒重,加入液氮研磨于离心管中保存,加入1.5 mL ddH₂O,放在超声波洗涤器中超声10 min,80°C水浴锅中水浴20 min,11 100×g离心10 min,收集1 mL上清液,向沉淀中再加入1.5 mL ddH₂O,重复超声步骤一次收集上清液1.5 mL,用真空浓缩仪60°C蒸干液体(约3 h),加入1.8 mL 80%乙醇放置于-20°C冰箱10 h,11 100×g离心10 min收集上清液1.5 mL,用真空浓缩仪60°C蒸干液体,加入0.7 mL蒸馏水溶解后反复吹打将离心管壁上黏附的沉淀溶解,然后在0.45 μm过滤微孔过滤器中过滤。

色谱分析条件:采用高效液相色谱法(HPLC, Waters 2695, the United States),定性和定量都是用

标样法进行,检测到无籽蜜柚汁胞中阿拉伯糖、木糖、蔗糖、葡萄糖和果糖的含量,色谱柱使用碳水柱(carbon water column, 4.6 mm×250 mm),柱温:37°C;流动相:乙腈:水溶液(体积比70:30),流速为1.0 mL·min⁻¹。

1.3 *Cg-SUS*和*Cg-INV*基因的鉴定获取及染色体定位

从拟南芥基因组数据库网站(www.arabidopsis.org/)下载拟南芥SUS基因家族和INV基因家族的蛋白序列,在柚子基因组数据库(www.citrus.hzau.edu.cn/orange/index.php)中进行同源比对,阈值设置为1E⁻¹⁰,初步获得柚子*Cg-SUS*和*Cg-INV*基因家族的候选基因。接着将这些候选基因的蛋白序列提交到Pfam数据库(www.pfam.xfam.org/) (Finn等2016)中,进行进一步的蔗糖合成结构域(PF00862)和糖基转移结构域(PF00534)验证,最终鉴定出6个和12个具有完整ORF的*Cg-SUS*和*Cg-INV*基因。根据柚子基因组数据库中的信息,使用MapDraw V2软件绘制基因在染色体上的相对分布位置图,参照Jiang等(2013)染色体位置命名法对本研究中的*Cg-SUS*和*Cg-INV*基因家族进行系统命名。

1.4 *Cg-SUS*和*Cg-INV*基因理化性质、进化树及内含子外显子结构分析

采用BioEdit软件、ProtParam工具(<http://web.expasy.org/protparam/>),对无籽蜜柚*Cg-SUS*和*Cg-INV*基因序列及其编码的蛋白序列进行理化特性分析。

从拟南芥数据库(www.arabidopsis.org/)中调取拟南芥6个SUS、12个INV蛋白序列;同时,在NCBI基因组数据库中(www.ncbi.nlm.nih.gov/genome/gdv/)搜索获得毛果杨(*Populus trichocarpa*)、碧桃(*Prunus persica*)的SUS基因家族的蛋白序列,番茄(*Solanum lycopersicum*)、水稻(*Oryza sativa*)的INV基因家族的蛋白序列。利用进化树分析软件MEGA 7.0 (www.mega.com/),进行At-SUS、Cg-SUS、Pt-SUS、Pp-SUS和At-INV、Cg-INV、Sl-INV、Os-INV之间进化关系分析研究,构建系统发育树,其参数设置参照董金金等(2018)方法。

将Fasta格式的*Cg-SUS*和*Cg-INV*基因编码的核苷酸序列和相应的全核苷酸序列,通过GSDS软件(gsds1.cbi.pku.edu.cn/index.php)绘制*Cg-SUS*和*Cg-INV*基因的内含子和外显子的结构模式图(董洁静等2018)。

1.5 Cg-SUS和Cg-INV蛋白功能结构域分析

Cg-SUS和Cg-INV蛋白功能结构域的绘制分析使用SMART在线网站(<https://smart.embl.de/>) (Sigrist等2013), 利用Cg-SUS和Cg-INV蛋白序列, 选择与PFAM数据库进行筛选, 得出每个蛋白相应的蔗糖合成酶结构域(PF00862)、糖基转移酶结构域(PF00534)、糖基水解酶结构域(PF12899)的位置。

1.6 无籽蜜柚Cg-SUS和Cg-INV基因数字表达谱分析

从NCBI上下载无籽蜜柚各发育时期转录组原始数据(SAMN08640739、SAMN08640740、SAMN08640741和SAMN08640742), 去除低质量reads, 将clean reads映射到Cg-SUS和Cg-INV基因上, 通过RSEM v1.2.31 (Li和Dewey 2011)计算基因的表达水平(Fragment Per Kilobase of exon model per Million mapped reads, FPKM)。利用HemI软件绘制Cg-SUS和Cg-INV基因数字表达谱热图。

1.7 qRT-PCR及相关性分析

用RNeasy Plus Micro Kit试剂盒提取无籽蜜柚7个不同时期汁胞的总RNA, 具体操作步骤见产品说明。第一链cDNA的合成采用TaKaRa Reverse-Transcription Kit。依据之前的柑橘稳定内参研究, *UBC*和*ACTB*基因被选为内参基因(Yan等2012)。利用PrimerQuest (www.idtdna.com/PrimerQuest/Home/Index)软件设计qRT-PCR检测引物(表1)。采用TaKaRa公司SYBR *Premix Ex Taq*进行qRT-PCR分析(具体操作步骤按产品说明书), 每个反应设置3次重复, 用 $2^{-\Delta\Delta C_t}$ 方法计算相对表达量, 并利用SPSS软件进行无籽蜜柚Cg-SUS2、Cg-SUS3和Cg-INV6基因表达量及各糖含量相关性和显著性分析。

2 实验结果

2.1 无籽蜜柚可溶性糖含量分析

可溶性糖的成分是决定口味和感官品质的关键因素。本研究用HPLC法在无籽蜜柚汁胞中检

测到5种可溶性糖, 包括阿拉伯糖、木糖、果糖、葡萄糖和蔗糖(图1)。值得注意的是, 在汁胞发育早期, 主要的可溶性糖是阿拉伯糖和木糖, 但在成熟过程中逐渐降低到几乎为零。然而随着无籽蜜柚汁胞的成熟, 果糖、葡萄糖和蔗糖的含量显著增加, 这为后续研究无籽蜜柚SUS和INV基因家族及无籽蜜柚糖代谢提供基础性条件。

2.2 Cg-SUS和Cg-INV基因编码蛋白理化特性分析

通过Blastp分析, 从柚子基因组中共鉴定到6个Cg-SUS和12个Cg-INV基因。对Cg-SUS和Cg-INV基因家族cDNA和编码蛋白进行序列分析(表2), 结果显示该基因的双链分子量为957.20~2 042.28 kDa, ORF平均长度为1 578~3 366 bp, 共编码508~1 121个氨基酸。通过Protparam分析, 推测该蛋白家族的理化特性分析如表2所示, 相对分子量为56.24~126.27 kDa, 等电点为4.9~9.19。

2.3 Cg-SUS和Cg-INV基因的分析及染色体定位

对6个Cg-SUS和12个Cg-INV基因在染色体上的位置进行分析, 由图2可以看出其不均匀地分布在柚子9条染色体(Chr1~Chr9)上, 根据染色体排位, 将其依次命名为Cg-SUS1~6及Cg-INV1~12 (图2)。其中Chr5、Chr6、Chr8和Chr9各含1个Cg-SUS基因; Chr4含2个Cg-SUS基因; 其余各染色体不含Cg-SUS基因。对于Cg-INV而言, Chr2、Chr3、Chr4、Chr6、Chr7和Chr8各含1个Cg-INV基因; Chr1、Chr5和Chr9含2个Cg-INV基因。

2.4 Cg-SUS和Cg-INV基因进化树及内含子、外显子分析

为了探究Cg-SUS和Cg-INV基因家族在进化上的位置及亲缘关系, 本研究开展了Cg-SUS与其他3个物种(*A. thaliana*、*P. trichocarpa*、*P. persica*) SUS和(*A. thaliana*、*S. lycopersicum*、*O. sativa*) INV家族成员的系统进化树分析。从图3中能够看出, Cg-SUS基因家族可聚为3组, 分别为SUSI、SUSII和SUSIII,

表1 无籽蜜柚内参基因和关键基因qRT-PCR引物

Table 1 The qRT-PCR primer of reference genes and key genes in *C. grandis* 'Seedless'

| 基因 | 上游引物(5'→3') | 下游引物(5'→3') |
|----------------|------------------------|-------------------------|
| <i>UBC</i> | GGATTGGGATGGCAGAGAAA | CGAGCAACAACATGGGAAAC |
| <i>ACTB</i> | CGATTAGCCTTCGGATTGAGAG | GATCTGGCATCACACCTTCTAC |
| <i>Cg-SUS2</i> | CTATCACCACCCACTACGACTA | CCCTACACTGAGGAGAAGAGAA |
| <i>Cg-SUS3</i> | GTCCAGTCCTTGCCCTCTTTA | TTCCACCTGATACACCCTTT |
| <i>Cg-INV6</i> | TGGCTTGGATACATCACATAAT | GTCAATCGCAAACCTGTCTACTT |

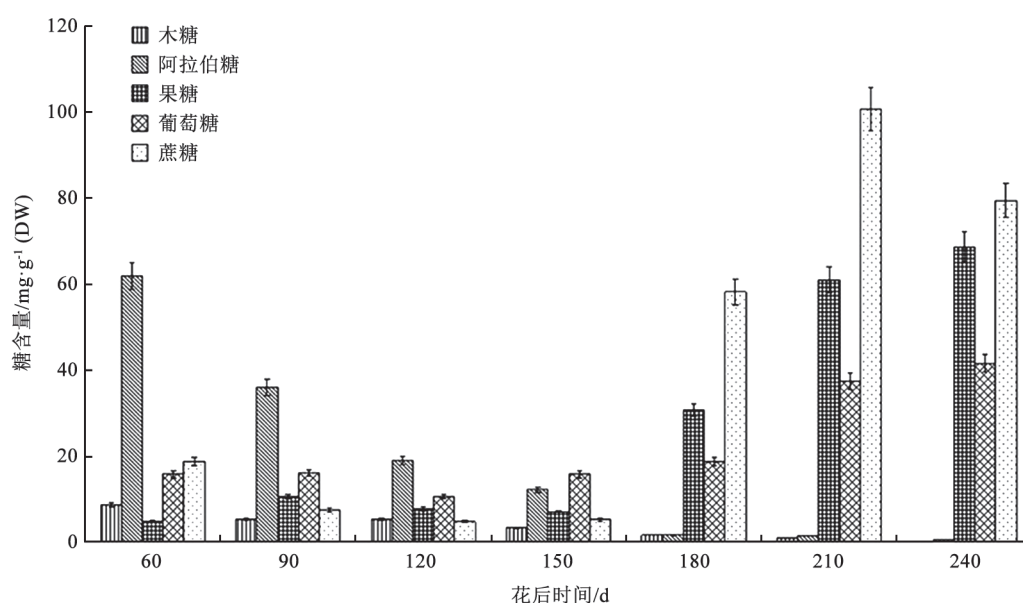


图1 无籽蜜柚汁胞发育过程中可溶性糖含量变化

Fig.1 Changes in soluble sugar contents in juice sacs during fruit development of *C. grandis* 'Seedless'

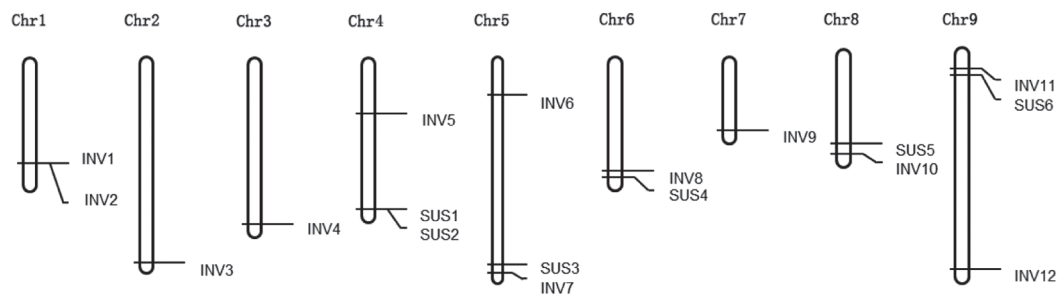
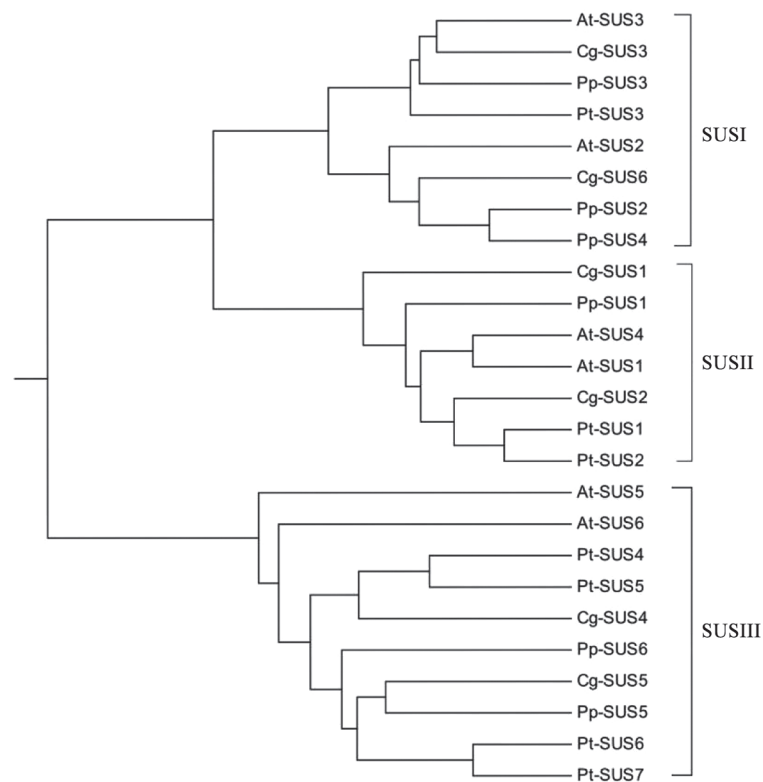
其中*Cg-SUS3*、*Cg-SUS6*位于SUSI组, *Cg-SUS1*和*Cg-SUS2*位于SUSII组, *Cg-SUS4*、*Cg-SUS5*位于SUSIII组。

在图4中, *Cg-INV*基因家族也可聚为3组, 分别是细胞质转化酶基因家族、液泡转化酶基因家族、细

胞壁转化酶基因家族。其中, *Cg-INV3*、*Cg-INV4*、*Cg-INV7*、*Cg-INV9*、*Cg-INV10*和*Cg-INV11*位于细胞质转化酶基因家族, *Cg-INV6*和*Cg-INV12*位于液泡转化酶基因家族, *Cg-INV1*、*Cg-INV2*、*Cg-INV5*和*Cg-INV8*位于细胞壁转化酶基因家族。

表2 *Cg-SUS*和*Cg-INV*基因家族生物信息分析及其编码蛋白理化特性分析Table 2 Bioinformatics, coding protein, physical and chemical properties analyses of the *Cg-SUS* and *Cg-INV* families

| 基因ID | 基因名称 | 染色体位置 | 分子量(双链)/ kDa | 碱基数/nt | ORF/bp | 编码氨基酸个数 | 相对分子 量/kDa | 等电点 |
|------------|--------------|------------------------------|-----------------|--------|--------|---------|---------------|------|
| Cg4g018010 | <i>SUS1</i> | chr4 (23 888 364~23 891 860) | 1 449.24 | 2 388 | 2 388 | 795 | 90.83 | 5.81 |
| Cg4g018060 | <i>SUS2</i> | chr4 (23 933 632~23 928 856) | 1 736.42 | 2 860 | 2 421 | 806 | 92.29 | 5.85 |
| Cg5g037720 | <i>SUS3</i> | chr5 (44 425 171~44 418 874) | 1 841.64 | 3 036 | 2 130 | 709 | 81.56 | 6.38 |
| Cg6g016740 | <i>SUS4</i> | chr6 (18 501 251~18 505 366) | 1 739.14 | 2 867 | 2 514 | 837 | 95.05 | 6.13 |
| Cg8g016990 | <i>SUS5</i> | chr8 (14 094 631~14 099 139) | 1 765.82 | 2 911 | 2 718 | 905 | 102.79 | 7.54 |
| Cg9g002460 | <i>SUS6</i> | chr9 (2 507 989~2 513 760) | 1 630.00 | 2 687 | 2 427 | 808 | 92.66 | 6.01 |
| Cg1g009760 | <i>INV1</i> | chr1 (6 541 296~6 538 838) | 957.20 | 1 578 | 1 578 | 525 | 58.98 | 7.59 |
| Cg1g009770 | <i>INV2</i> | chr1 (6 554 408~6 547 807) | 2 042.28 | 3 366 | 3 366 | 1 121 | 126.27 | 8.69 |
| Cg2g032800 | <i>INV3</i> | chr2 (41 989 887~41 985 162) | 1 631.50 | 2 692 | 2 034 | 677 | 76.54 | 6.13 |
| Cg3g019340 | <i>INV4</i> | chr3 (26 561 313~26 556 579) | 1 364.73 | 2 250 | 1 677 | 558 | 63.66 | 6.07 |
| Cg4g006840 | <i>INV5</i> | chr4 (7 401 818~7 404 473) | 1 302.36 | 2 148 | 1 605 | 534 | 60.33 | 4.9 |
| Cg5g008630 | <i>INV6</i> | chr5 (6 290 979~6 295 067) | 1 754.29 | 2 891 | 1 527 | 508 | 56.24 | 5.23 |
| Cg5g040500 | <i>INV7</i> | chr5 (46 305 008~46 301 772) | 1 202.34 | 1 980 | 1 980 | 659 | 74.38 | 5.75 |
| Cg6g015210 | <i>INV8</i> | chr6 (17 403 417~17 398 694) | 1 184.63 | 1 952 | 1 725 | 574 | 65.25 | 9.19 |
| Cg7g009870 | <i>INV9</i> | chr7 (10 485 560~10 479 334) | 2 005.30 | 3 309 | 1 953 | 650 | 73.07 | 5.89 |
| Cg8g017930 | <i>INV10</i> | chr8 (15 928 129~15 924 290) | 1 514.23 | 2 497 | 1 932 | 643 | 72.42 | 6.72 |
| Cg9g000750 | <i>INV11</i> | chr9 (1 413 310~1 416 703) | 1 558.88 | 2 572 | 2 028 | 675 | 75.76 | 6.72 |
| Cg9g024960 | <i>INV12</i> | chr9 (36 156 923~36 161 625) | 1 361.28 | 2 242 | 1 929 | 642 | 71.30 | 5.61 |

图2 柚子 *Cg-SUS* 和 *Cg-INV* 基因染色体定位Fig.2 Chromosome localization of *Cg-SUS* and *Cg-INV* genes in *C. grandis*图3 *SUS* 基因系统进化树Fig.3 The evolutionary tree of *SUS* genes

图中各基因登录号分别为: *At-SUS1* (At5g-20830), *At-SUS2* (At5g49190), *At-SUS3* (At4g-02280), *At-SUS4* (At3g43190), *At-SUS5* (At5g-37180), *At-SUS6* (At1g73370), *Pt-SUS1* (ADR-81996.1), *Pt-SUS2* (ADR81997.1), *Pt-SUS3* (XP_002302727.1), *Pt-SUS4* (ADR81999.1), *Pt-SUS5* (ADR82000.1), *Pt-SUS6* (ADR82001.1), *Pt-SUS7* (ADR82002.1), *Pp-SUS1* (AF157908.1), *Pp-SUS2* (XP_007225257.2), *Pp-SUS3* (XP_007200961.2), *Pp-SUS4* (XP_007225235.1), *Pp-SUS5* (XP_007216050.2), *Pp-SUS6* (XP_007210378.1)。

通过将全长cDNA序列与相应的基因组DNA序列进行比较,揭示了每个*Cg-SUS*和*Cg-INV*的外显子-内含子结构。*Cg-SUS*内含子数目在12~14个之间,*Cg-INV*内含子数目在3~12个之间(图5),*Cg-SUS*和*Cg-INV*基因不仅具有密切的系统发育关系,而且也表现出复杂的内含子-外显子分布模式。

2.5 *Cg-SUS*和*Cg-INV*蛋白功能结构域分析

*Cg-SUS*和*Cg-INV*蛋白功能结构域分析表明,*Cg-SUS1*~*6*均含有蔗糖合成酶结构域(PF00862)和糖基转移酶结构域(PF00534)(图6),进一步说明该6个蛋白属于*Cg-SUS*蛋白家族成员。图6可以看出,*Cg-INV1*、*Cg-INV2*、*Cg-INV5*、*Cg-INV6*、*Cg-INV8*

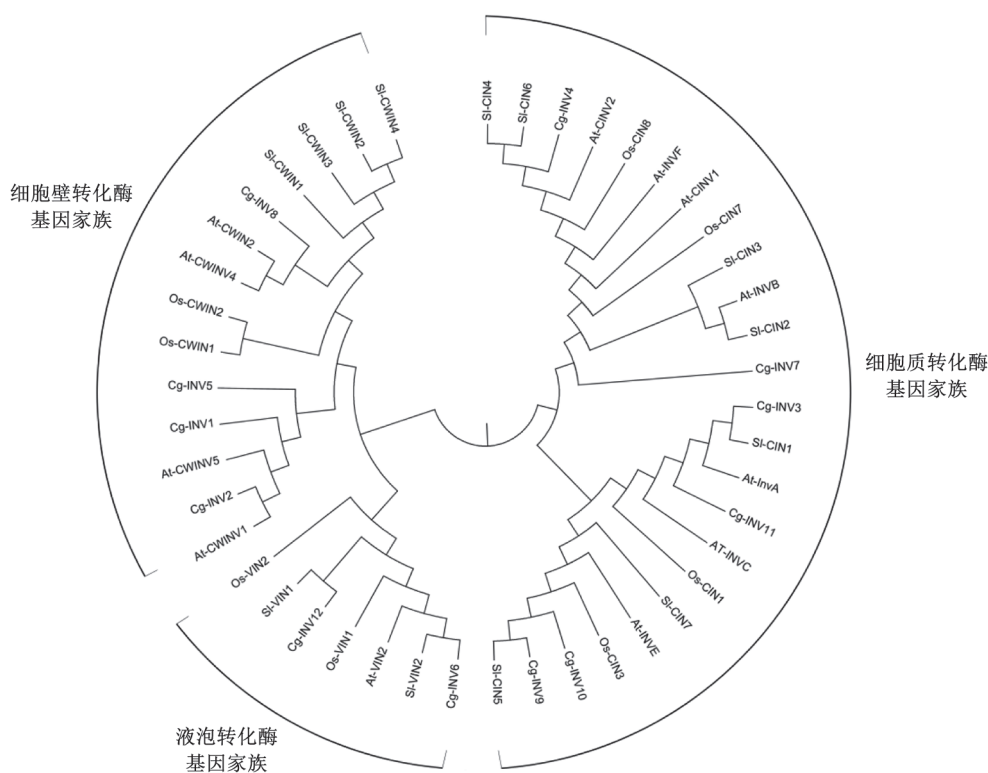


图4 INV基因系统进化树

Fig.4 The evolutionary tree of INV genes

图中各基因登录号分别为: *SI-CWIN1* (AJ2-72304), *SI-CWIN2* (AF506005), *SI-CWIN3* (AF506006), *SI-CWIN4* (AF506007), *SI-VIN1* (NM_001247689), *SI-VIN2* (NM_001247140), *SI-CIN1* (Solyc01g100810), *SI-CIN2* (Solyc01g111100), *SI-CIN3* (Solyc04g081440), *SI-CIN4* (Solyc06g065210), *SI-CIN5* (Solyc11g007270), *SI-CIN6* (Solyc11g020610), *SI-CIN7* (Solyc11g067050), *At-CWIN1* (At3g13790), *At-CWIN2* (At3g52600), *At-CWIN4* (At2g36190), *At-CWIN5* (At3g13784), *At-VIN1* (At1g62660), *At-VIN2* (At1g12240), *At-CIN1* (At1g56560), *At-CIN3* (At3g06500), *At-CIN5* (At5g22510), *At-CIN7* (At1g35580), *At-CIN9* (At4g09510), *Os-CWIN1* (AY578158), *Os-CWIN2* (AY578159), *Os-VIN1* (AF276703), *Os-VIN2* (AF276704), *Os-CIN1* (AK103334), *Os-CIN3* (AK121301), *Os-CIN7* (AK065562), *Os-CIN8* (AK102741)。

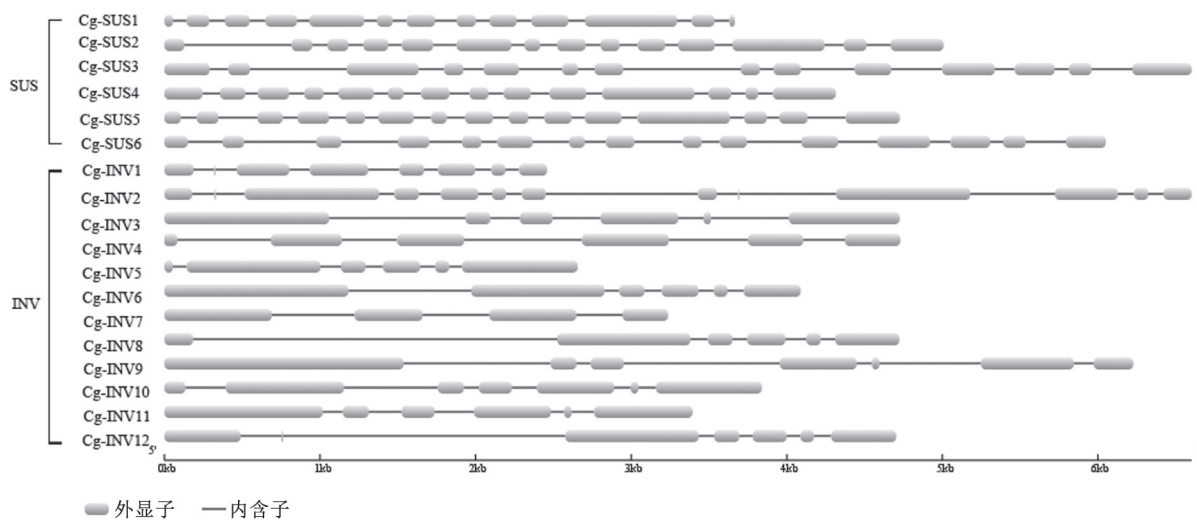


图5 Cg-SUS和Cg-INV基因内含子及外显子分析
Fig.5 Analysis of introns and exons of Cg-SUS and Cg-INV genes

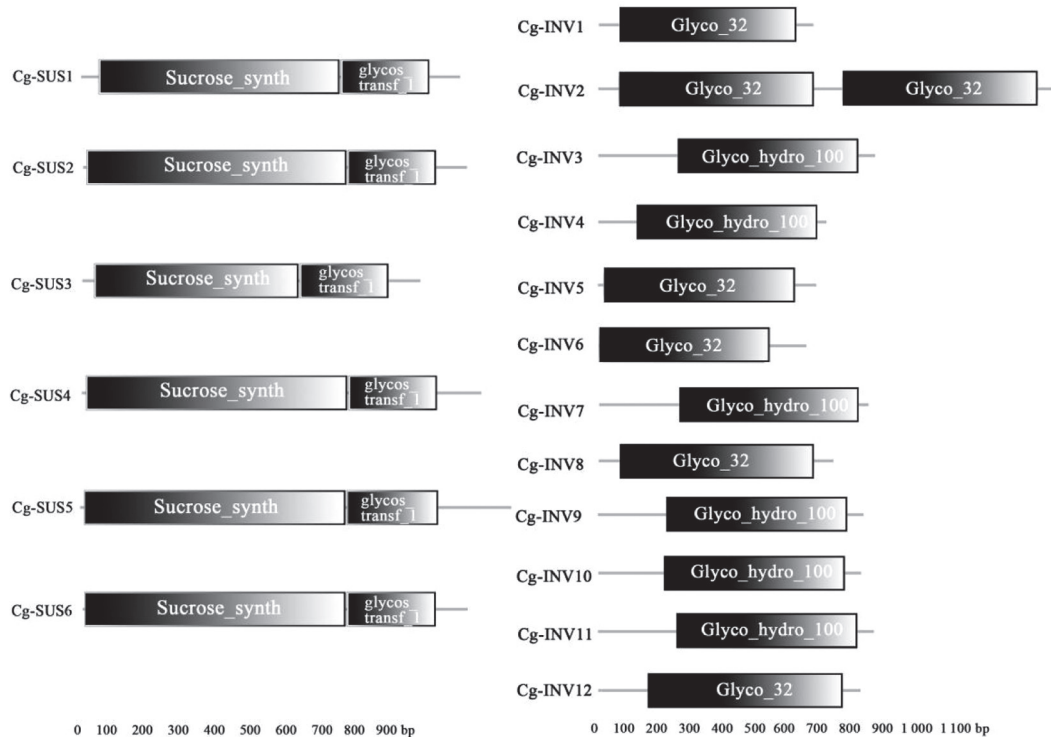


图6 Cg-SUS和Cg-INV蛋白功能结构域分析

Fig.6 Analysis of conserved protein domains of Cg-SUS and Cg-INV proteins

和Cg-INV12蛋白均含有Glyco_32 (SM000640)的糖基水解酶, 而Cg-INV3、Cg-INV4、Cg-INV7、Cg-INV9、Cg-INV10和Cg-INV11蛋白含有Glyco_hydro_100 (PF12899)的糖基水解酶, 这说明了Cg-INV蛋白家族能分为2种不同的水解酶类型。

2.6 无籽蜜柚汁胞发育过程中Cg-SUS和Cg-INV基因分析

通过数字表达谱(图7)分析, Cg-SUS2~6和Cg-INV2~12在无籽蜜柚汁胞发育过程中均有表达, 而Cg-SUS1和Cg-INV1无表达, 推测Cg-SUS1和Cg-INV1可能在其他组织器官中特异性表达。图7表明, 在整个无籽蜜柚汁胞发育过程中, Cg-SUS2和Cg-SUS3有较高表达水平 $[\log_2(\text{FPKM}) > 6]$, 而Cg-SUS4~6呈现较低表达水平 $[\log_2(\text{FPKM}) < 2]$, 表明Cg-SUS2和Cg-SUS3是SUS介导蔗糖降解的主要酶。值得注意的是, Cg-SUS2在发育初期(120 d)具有较高表达水平, 而Cg-SUS3在发育后期(240 d)有较高表达水平, 表达模式的差异表明Cg-SUS2和Cg-SUS3可能具有不同的生物学功能。同样, 在INV基因中, 除了Cg-INV6 $[\log_2(\text{FPKM}) > 6]$, 其余INV基因在汁胞发育中呈现较低表达水平 $[\log_2(\text{FPKM}) < 2]$ 。

2.7 Cg-SUS2、Cg-SUS3和Cg-INV6的基因表达量与各糖含量的相关性分析

根据数字表达分析, Cg-SUS2、Cg-SUS3和Cg-INV6被筛选出来进行qRT-PCR验证。结果表明, Cg-SUS3和Cg-INV6呈上调表达趋势, Cg-SUS2表达量逐渐下调, 三者表达趋势基本与数字表达分析一致(图8)。相关性分析表明, Cg-SUS2表达量与木糖、阿拉伯糖含量之间存在极显著正相关性; Cg-SUS3和Cg-INV6表达量与果糖、葡萄糖、蔗糖含量之间存在显著正相关性(表3), 表明在无籽蜜柚汁胞发育过程中, Cg-SUS2可能参与阿拉伯糖和木糖的生物合成, Cg-SUS3可能参与蔗糖降解代谢, Cg-INV6可能是液泡己糖积累的直接原因。

3 讨论

果实发育涉及一系列的生理和生化过程, 包括细胞壁代谢和细胞膨压变化, 以及风味和营养成分有关的代谢变化(Gapper等2013)。可溶性糖的成分分析表明, 阿拉伯糖和木糖在无籽蜜柚汁胞发育早期中具有较高的含量, 而果糖、葡萄糖和蔗糖的含量随着果实发育成熟呈显著累积(图

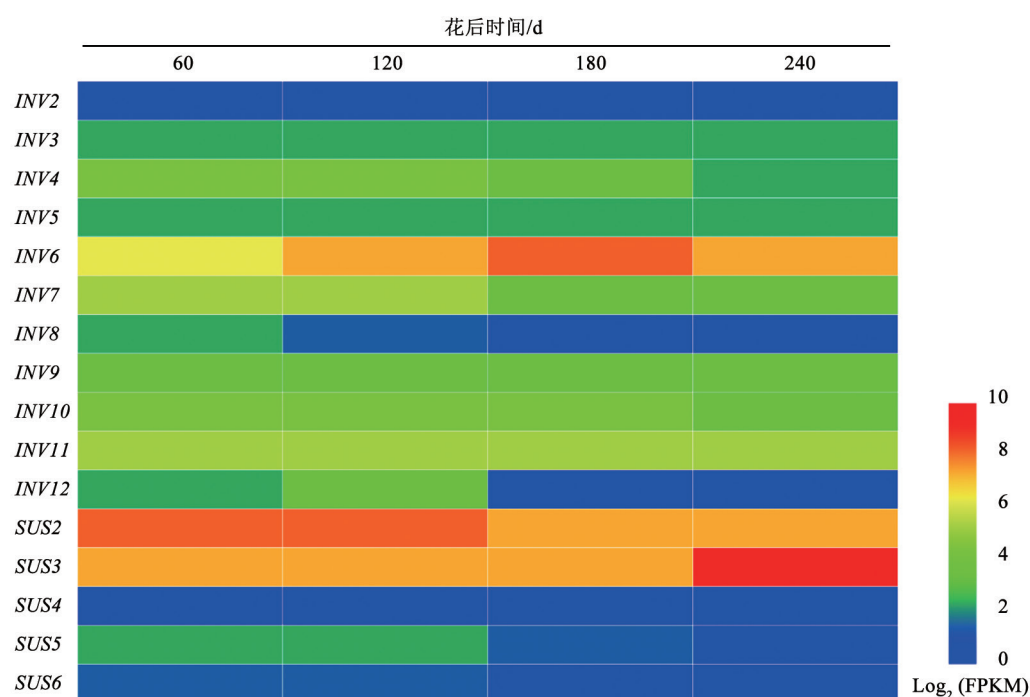


图7 无籽蜜柚汁胞Cg-SUS和Cg-INV基因数字表达谱分析

Fig.7 The digital expression analysis of Cg-SUS and Cg-INV genes in *C. grandis* 'Seedless'

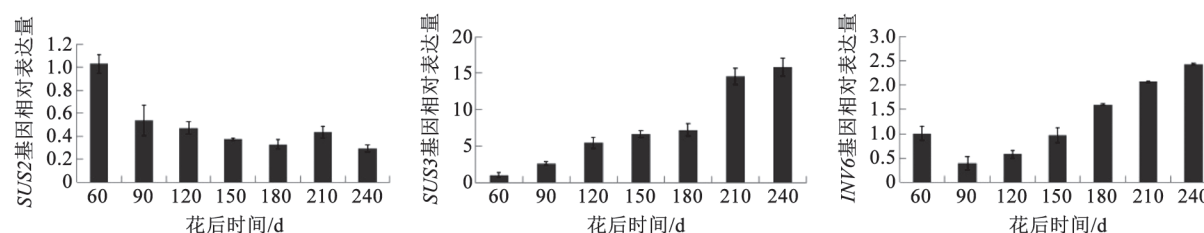


图8 无籽蜜柚发育过程中Cg-SUS2、Cg-SUS3和Cg-INV6基因表达谱分析

Fig.8 The gene expressions of Cg-SUS2, Cg-SUS3 and Cg-INV6 in *C. grandis* 'Seedless' during development

表3 无籽蜜柚Cg-SUS2、Cg-SUS3和Cg-INV6基因表达量与发育过程中可溶性糖含量的相关系数

Table 3 Relative correlation between the gene expressions of Cg-SUS2, Cg-SUS3, Cg-INV6 and soluble sugar contents in *C. grandis* 'Seedless' during development

| 基因 | 木糖含量 | 阿拉伯糖含量 | 果糖含量 | 葡萄糖含量 | 蔗糖含量 |
|---------|---------|---------|---------|--------|---------|
| Cg-SUS2 | 0.877** | 0.942** | -0.534 | -0.312 | -0.291 |
| Cg-SUS3 | -0.852* | -0.811* | 0.933** | 0.830* | 0.824* |
| Cg-INV6 | -0.718 | -0.616 | 0.914** | 0.810* | 0.876** |

*表示差异达到显著水平($P < 0.05$), **表示差异达到极显著水平($P < 0.01$)。

1)。在植物中,阿拉伯糖和木糖是组成细胞壁多糖的基本组成成分,柑橘类成熟过程中的贮藏化合物由果糖、葡萄糖和蔗糖组成(Wu等2011)。由此可见,阿拉伯糖和木糖可能有助于早期无籽蜜柚汁胞胞间层和初级细胞壁的形成(Cosgrove 1997),

而果糖、葡萄糖和蔗糖是无籽蜜柚主要的贮藏化合物。众所周知,糖不仅是生长发育所需能量和碳骨架的丰富来源,而且是调节基因表达的信号分子(Ruan 2014)。在高等植物中,蔗糖是光合作用的最终产物,能够通过韧皮部运输到植物的非

光合组织中,蔗糖的分解反应由SUS或INV催化。INV可分为细胞壁转化酶(CWIN)、液泡转化酶(VIN)和细胞质转化酶(CIN)。与INV相比,SUS是植物中广泛存在的一种UDP-糖基转移酶,是催化蔗糖代谢的关键酶,能更快地催化蔗糖的降解,并在种子和果实的发育中发挥作用(Xu等2012),同时它能够降解蔗糖生成UDP-葡萄糖和果糖,参与胍胍质、纤维素、淀粉、糖脂和糖蛋白的生物合成(Coleman等2009)。本研究基于拟南芥数据库中SUS和INV蛋白序列与柚子的基因组数据同源序列比对,筛选分别获得6个和12个具有完整ORF的*Cg-SUS*和*Cg-INV*。根据之前Jiang等(2013)的染色体命名法,分别命名为*Cg-SUS1~6*和*Cg-INV1~12*。

先前在碧桃(Zhang等2015)等植物的研究中,通过系统发育和分子结构的比较,SUS蛋白家族在进化上可聚为三类(SUSI、SUSII和SUSIII)。在本研究中,*Cg-SUS*基因与其他植物同源基因的系统发育分析结果证实了这一分类(图3)。同样,本研究证实了INV基因家族也可聚为三类,分别是细胞质转化酶基因家族、液泡转化酶基因家族、细胞壁转化酶基因家族(图4)(Ruan 2014)。这些研究结果表明,SUS、INV家族基因在不同植物物种的进化或分离过程中具有高度保守性。研究发现,*Cg-SUS*基因家族均具有2个结构域(图6),即蔗糖合酶结构域和糖基转移酶结构域,这两个结构域分别对应蛋白质结构中的N端“调控”结构域和C端GT-B结构域。从图6中能够看出,*Cg-INV3*、*Cg-INV4*、*Cg-INV7*、*Cg-INV9*、*Cg-INV10*和*Cg-INV11*基因都含有Glyco_hydro_100(PF12899)的糖基水解酶,结合进化树分析能够看出(图4),*Cg-INV3*、*Cg-INV4*、*Cg-INV7*、*Cg-INV9*、*Cg-INV10*和*Cg-INV11*均属于细胞质转化酶基因家族,即碱性转化酶基因家族,在先前的研究中也证实Glyco_hydro_100结构域与碱性转化酶相关(Gao等2016)。

*Cg-SUS*和*Cg-INV*基因数字表达谱分析发现5个*Cg-SUS*基因均有表达(图7),但只有*Cg-SUS2*和*Cg-SUS3*在汁胞发育过程中呈现出活跃转录。这些结果表明,*Cg-SUS2*和*Cg-SUS3*是形成无籽蜜柚汁胞的主要蔗糖降解酶。不同SUS基因的表达可能受到糖的影响,在分离的马铃薯(*Solanum tuberosum*)叶片中,蔗糖诱导了*St-SUS4*基因的表达,

而*St-SUS3*的表达不受影响(Fu和Park 1995)。在拟南芥种子的6个*At-SUS*家族成员中,转录因子*At-LEC2*能够使*At-SUS2*和*At-SUS3*呈现出高表达水平(Angeles-Núñez和Tiessen 2012)。Bieniawska等(2007)报道称*At-SUS1*和*At-SUS4*的双突变体显著影响拟南芥种子可溶性糖含量。在玉米中,*Zm-SUS1*参与了种子淀粉的生物合成,*Zm-SUS2*参与了胚乳细胞壁的完整性,而*Zm-SUS3*参与了基底层胚乳转移细胞的形成(Chourey等1998)。从图7中发现*Cg-SUS2*和*Cg-SUS3*之间完全相反的表达模式,这暗示了*Cg-SUS*在无籽蜜柚汁胞发育过程中起着不同的作用。从表3中发现,*Cg-SUS2*与木糖、阿拉伯糖之间存在极显著正相关性,表明在无籽蜜柚汁胞发育过程中,*Cg-SUS2*可能涉及阿拉伯糖和木糖的生物合成。UDP-糖基转移酶催化糖基转移反应,是植物中较为普遍的修饰反应,从而调节受体分子在细胞和机体内的活性(Coleman等2009)。结合*Cg-SUS3*表达模式与葡萄糖、果糖、蔗糖含量之间存在显著正相关性,表明*Cg-SUS3*可能在胍胍质、糖脂和糖蛋白等生物合成过程及可溶性糖贮存方面发挥重要作用。

同时,在本研究中,11个*Cg-INV*基因表达,*Cg-INV6*在无籽蜜柚汁胞发育过程中显示出了活跃的转录水平,在进化树(图4)中发现*Cg-INV6*属于液泡转化酶基因家族。液泡成分似乎与细胞膨胀压力和内部糖分的调节有关,液泡转化酶活性在成熟期增加,酶活性水平与己糖水平相关,它似乎是在成熟组织中存储和控制糖通量的一部分(Gayler和Glasziou 2010)。在之前Klann等(1996)对番茄研究中,*Sl-VIN*基因的沉默能够导致蔗糖的累积及己糖含量的减少。在本研究中,我们确定*Cg-INV6*在细胞发育后期表达水平较高,此外相关性分析表明*Cg-INV6*表达水平与果糖、蔗糖、葡萄糖含量之间存在显著正相关性。这些结果说明*Cg-INV6*可能是液泡己糖累积的关键酶,涉及到液泡介导的渗透作用调控汁胞细胞的扩张。

参考文献(References)

- Angeles-Núñez JG, Tiessen A (2012). Regulation of *At-SUS2*, and *At-SUS3*, by glucose and the transcription factor *Lec2* in different tissues and at different stages of *Arabidopsis* seed development. *Plant Mol Biol*, 78 (4): 377

- Baud S, Vaultier MN, Rochat C (2004). Structure and expression profile of the sucrose synthase multigene family in *Arabidopsis*. *J Exp Bot*, 55 (396): 397–409
- Bieniawska Z, Paul DB, Garlick AP, et al (2007). Analysis of the sucrose synthase gene family in *Arabidopsis*. *Plant J*, 49 (5): 810–828
- Chourey PS, Taliercio EW, Carlson SJ, et al (1998). Genetic evidence that the two isozymes of sucrose synthase present in developing maize endosperm are critical, one for cell wall integrity and the other for starch biosynthesis. *Mol Gen Genet*, 259 (1): 88–96
- Coleman HD, Yan J, Mansfield SD (2009). Sucrose synthase affects carbon partitioning to increase cellulose production and altered cell wall ultrastructure. *Proc Natl Acad Sci USA*, 106: 13118
- Cosgrove DJ (1997). Assembly and enlargement of the primary cell wall in plants. *Annu Rev Cell Dev Biol*, 13 (13): 171–201
- Dong JJ, Liu W, Li M, et al (2018). Gene expression and phylogenetic analysis of MADS-box family genes in *Ginkgo biloba*. *Plant Physiol J*, 54 (6): 1055–1063 (in Chinese with English abstract) [董金金, 刘伟, 李萌等(2018). 银杏MADS-box基因家族的表达及系统发育分析. *植物生理学报*, 54 (6): 1055–1063]
- Dong JJ, Xu JY, Lin JJ et al (2018). Genome-wide identification and expression analysis of *PLCs* gene family in maize (*Zea mays*). *Plant Physiol J*, 54 (6): 1018–1028 (in Chinese with English abstract) [董洁静, 徐晶宇, 林俊俊等(2018). 玉米*PLCs*基因家族鉴定及表达谱分析. *植物生理学报*, 54 (6): 1018–1028]
- Duncan KA, Hardin SC, Huber SC (2006). The three maize sucrose synthase isoforms differ in distribution, localization, and phosphorylation. *Plant Cell Physiol*, 47 (7): 959–971
- Finn RD, Coggill P, Eberhardt RY, et al (2016). The pfam protein families database: towards a more sustainable future. *Nucleic Acids Res*, 44: 279–285
- Fu H, Park WD (1995). Sink- and vascular-associated sucrose synthase functions are encoded by different gene classes in potato. *Plant Cell*, 7 (9): 1369–1385
- Gapper NE, Mcquinn RP, Giovannoni JJ (2013). Molecular and genetic regulation of fruit ripening. *Plant Mol Biol*, 82 (6): 575
- Gao F, Cao XF, Si JP, et al (2016). Characterization of the alkaline/neutral invertase gene in *dendrobium officinale* and its relationship with polysaccharide accumulation. *Genet Mol Res*, 15 (2): 15027647
- Gayler KR, Glasziou KT (2010). Physiological functions of acid and neutral invertases in growth and sugar storage in sugar cane. *Physiol Plant*, 27 (1): 25–31
- Jiang H, Wu Q, Jing J, et al (2013). Genome-wide identification and expression profiling of ankyrin-repeat gene family in maize. *Dev Genes Evol*, 223 (5): 303–318
- Jin Y, Ni D, Ruan Y (2009). Post translational elevation of cell wall invertase activity by silencing its inhibitor in tomato delays leaf senescence and increases seed weight and fruit hexose level. *Plant Cell*, 21: 2072–2089
- Klann EM, Hall B, Bennett AB (1996). *Antisense acid invertase (TIV1)* gene alters soluble sugar composition and size in transgenic tomato fruit. *Plant Physiol*, 112 (3): 1321–1330
- Li B, Dewey CN (2011). Rsem: accurate transcript quantification from rna-seq data with or without a reference genome. *BMC Bioinformatics*, 12 (1): 323–323
- Ruan YL (2014). Sucrose metabolism: gateway to diverse carbon use and sugar signaling. *Annu Rev Plant Biol*, 65 (1): 33–67
- Ruan YL, Jin Y, Yang YJ, et al (2010). Sugar input, metabolism, and signaling mediated by invertase: roles in development, yield potential, and response to drought and heat. *Mol Plant*, 3 (6): 942–955
- Sigrist CJ, De CE, Cerutti L, et al (2013). New and continuing developments at prosite. *Nucleic Acids Res*, 41 (D1): 344–347
- Wu SJ, Ng CC, Tzeng WS, et al (2011). Functional antioxidant and tyrosinase inhibitory properties of extracts of taiwanese pummelo (*Citrus grandis* osbeck). *Afr J Biotechnol*, 10 (39): 7668–7674
- Xu SM, Brill E, Llewellyn DJ, et al (2012). Overexpression of a potato sucrose synthase gene in cotton accelerates leaf expansion, reduces seed abortion, and enhances fiber production. *Mol Plant*, 5: 430–441
- Yan J, Yuan F, Long G, et al (2012). Selection of reference genes for quantitative real-time RT-PCR analysis in *Citrus*. *Mol Bio Rep*, 39 (2): 1831
- Yao Y, Geng MT, Wu XH, et al (2015). Genome-wide identification, expression, and activity analysis of alkaline/neutral invertase gene family from cassava (*Manihot esculenta* crantz). *Plant Mol Biol Rep*, 33 (2): 304–315
- Zhang C, Yu M, Ma R, et al (2015). Structure, expression profile, and evolution of the sucrose synthase gene family in peach (*Prunus persica*). *Acta Physiol Plant*, 37 (4): 81

Bioinformatics and expression analysis of the sucrose synthase (SUS) and invertase (INV) gene families in *Citrus grandis* ‘Seedless’

DENG Shu-Ya, MAI Yi-Ting, CHEN Hui-Ping, NIU Jun*

Institute of Tropical Agriculture and Forestry, Hainan University, Haikou 570228, China

Abstract: Sucrose is the end product of photosynthesis. Upon translocation through the phloem to sinks, sucrose is degraded by either sucrose synthase (SUS) or invertase (INV). To study the functions of *SUS* and *INV* genes in *Citrus grandis* ‘Seedless’, we used HPLC analysis, and the results indicate that the contents of arabinose and xylose were higher in early development of *C. grandis* ‘Seedless’ juice sacs, while the contents of fructose, glucose and sucrose significantly increased during the ripening process. Based on *Arabidopsis* *SUS* and *INV* proteins, we used BLASTP to search the homologous protein in pomelo genome. A total of 6 *Cg-SUS* and 12 *Cg-INV* with complete open reading frame (ORF) were identified. Based on their positions on *C. grandis* chromosomes, genes were named as *Cg-SUS1–6* and *Cg-INV1–12*, respectively. Additionally, the analysis on their physicochemical properties, intron and exon, phylogenetic tree, and amino acid functional domain were performed. The digital expression analysis indicated that *Cg-SUS2*, *Cg-SUS3* and *Cg-INV6* showed an active transcript in juice sacs during fruit development of *C. grandis* ‘Seedless’, and also their expression profiles were detected by qRT-PCR. Combined with the correlation analysis between the soluble sugar contents and the gene expressions, the results showed that the expression of *Cg-SUS2* had established a close positive correlation between the contents of xylose and arabinose. So *Cg-SUS2* might be involved in biosyntheses of arabinose and xylose. The expressions of *Cg-SUS3* and *Cg-INV6* had a close positive correlation between the contents of glucose, fructose and sucrose, suggesting that *Cg-SUS3* may be involved in the degradation of sucrose, which is converted into UDP-glucose and fructose. And *Cg-INV6* may be a key enzyme in vacuolar hexose accumulation.

Key words: *Citrus grandis* ‘Seedless’; juice sacs development; sucrose synthase; invertase

Received 2018-07-16 Accepted 2018-09-29

This work was supported by Hainan Provincial Natural Science Foundation (317045) and Hainan University Research Funded Projects [KY-QD(ZR)1701].

*Corresponding author (niujun5555@163.com).