



西瓜 *PIP5K* 基因家族鉴定及其在雄性不育花蕾中的表达分析

张高原¹, 丁 谦², 魏兵强¹

(1. 甘肃农业大学 园艺学院, 兰州 730070; 2. 潍坊工程职业学院 花卉学院, 山东青州 262500)

摘 要 为了挖掘西瓜磷脂酰肌醇 4-磷酸 5-激酶(PIP5K)基因的功能, 利用生物信息学方法对西瓜 PIP5K 家族成员进行鉴定和特征分析, 并对其在多个组织的表达模式进行分析研究。结果显示: 西瓜基因组中含有 8 个 PIP5K 家族成员, 分布在 6 条染色体上; 多数基因定位在细胞质或细胞膜上; 共线性分析发现 4 个 *ClPIP5K* 基因与 5 个 *AtPIP5K* 基因存在共线关系, 并且 *ClPIP5K* 基因间存在片段复制事件; 进化树分析发现 *ClPIP5K* 基因和 *AtPIP5K* 基因被分成 4 组, 并且同一组内的基因具有相似的保守基序和基因结构; 启动子分析发现 *ClPIP5K* 基因中含有激素和胁迫响应相关顺式作用元件; 表达分析显示 *ClPIP5K* 基因在西瓜生长点、果柄、果肉、叶片、下胚轴、根以及花蕾等组织中表现出不同的表达水平。

关键词 西瓜; PIP5K; 基因家族; 雄性不育; 表达模式

中图分类号 S651

文献标志码 A

文章编号 1004-1389(2021)06-0883-11

磷脂酰肌醇 4-磷酸 5-激酶(PIP5K)是磷脂酰肌醇(phosphatidylinositol, PI)信号转导途径中的关键酶, 在多种细胞的功能发挥方面起着重要调控作用^[1]。大多数 PIP5K 蛋白的 N 端具有多个 MORN(Membrane Occupation and Recognition Nexus)结构域, C 端具有 1 个起催化功能的 PIPKc 结构域, 这些结构域在真核生物中高度保守^[2-3]。研究表明, *PIP5K* 基因对植物的多个组织器官的发育起调控作用, 如气孔^[4]、根^[5-7]、花^[8-11]等; 此外, 还参与调控植物逆境胁迫响应, 如盐害^[12]、水分胁迫^[13]等。目前, 在拟南芥(*Arabidopsis thaliana*)、水稻(*Oryza sativa*)、银杏(*Ginkgo biloba*)和大豆(*Glycine max*)中分别发现 11、11、7 和 22 个 *PIP5K* 基因^[14], 部分基因已被克隆。

虽然西瓜基因组数据已公布^[15], 但是关于西瓜 *PIP5K* 基因的鉴定及其在雄性不育花蕾中的表达分析未见相关报道。本试验拟利用生物信息学方法对西瓜 *PIP5K* 基因进行筛选和鉴定, 并对其蛋白理化性质、结构域、基因结构、motif 分布、染色体定位、进化树、共线性、顺式作用元件等特征以及组织表达模式进行分析, 以期为进一步研

究其功能奠定前期基础。

1 材料与方法

1.1 西瓜花蕾材料

西瓜雄性两用系中可育系‘MF-1’和不育系‘MS-1’种子均由甘肃农业大学瓜类研究所提供。种子于 2020 年 4 月 22 日播种于兰州市皋兰县忠和镇实试验田, 于 7 月 8 日分别采取雄性可育株和不育株的子蔓茎尖花蕾(长度为 3~5 mm)和茎尖下方 4~5 节处花蕾(长度为 7~9 mm), 每种花蕾至少采集 30 个, 储存在液氮中保存, 备用。

1.2 西瓜 PIP5K 家族成员的获取

首先从西瓜基因组数据库(ftp://cucurbit-genomics.org/pub/cucurbit/genome/watermelon/97103/v1/)^[15]下载蛋白组和基因组序列, 其次利用 Blast 本地化软件(Blast2. 2. 28)(https://blast.ncbi.nlm.nih.gov/Blast.cgi)构建西瓜蛋白组和基因组数据库, 然后以 Pfam 数据库(http://pfam.xfam.org)^[16]中的 PIP5K 保守结构域的隐马尔可夫模型(Hidden Markov Model; HMM)文件(序列号为 PF01504)为搜索序列, 利用 HMMER 3. 2 软件(http://www.hmmerr.

收稿日期:2020-09-17 修回日期:2020-11-05

基金项目:甘肃农业大学公开招聘博士科研启动基金(2017RCZX-30)。

第一作者:张高原,男,博士,讲师,研究方向为西瓜育种及分子生物学。E-mail:zhanggy@gsau.edu.cn

org/)^[17]在本地西瓜蛋白组数据库中进行 BlastP 搜索,阈值为 $E\text{-value} < 10^{-10}$,获取西瓜 PIP5K 候选成员。为确保候选成员中含有 PIPKc 结构域,利用 SMART (<http://smart.embl.de/>)在线软件鉴定西瓜 PIP5K 候选成员。最后从本地西瓜蛋白组和基因组数据库中下载鉴定后的西瓜 PIP5K 家族成员的蛋白和基因序列,作为后期分析的基础。拟南芥 PIP5K 成员^[14]的基因和蛋白序列来源于拟南芥数据库(<http://www.arabidopsis.org/>)。

1.3 西瓜 PIP5K 家族成员基本信息分析

蛋白序列长度、分子量及等电点等理化特征利用 ProtParam (<http://web.expasy.org/prot-param/>)在线软件进行分析。亚细胞定位分别利用 Euk-mPLoc 2.0 (<http://www.csbio.sjtu.edu.cn/bioinf/euk-multi-2/>)^[18]和 ProtComp version 9.0 server (<http://www.softberry.com>)进行分析。根据基因序列起始位点,利用 MapChart 2.2 软件绘制西瓜 PIP5K 基因染色体分布图。利用 MCScanX 软件 (<http://chibba.pgml.uga.edu/mcscan2/>)^[19]进行西瓜和拟南芥 PIP5K 基因的共线性分析,结果用 circos 软件 (<http://circos.ca/>)绘制^[20]。进化树由 MEGA-X 软件^[21]构建,序列比对方法均为 ClustalW,Bootstrap 的重复次数设置为 1 000,其中西瓜 PIP5K 家族成员的蛋白和编码序列进化树的构建方法均为邻接法(Neighbor-joining, NJ),执行参数分别为 p-distance 和 pairwise deletion^[22-24],剩余设置选择默认选项;拟南芥和西瓜 PIP5K 家族成员的蛋白和编码序列系统进化树的构建方法均为最大似然法(Maximum Likelihood, ML),执行参数分别为 Jones-Taylor-Thornton (JTT) model、Gamma Distributed (G) 和 Use all sites^[24],剩余设置选择默认选项。利用 OrthoFinder 2.3.11 软件^[25] (<http://www.stevekellylab.com/software/orthofinder>)分析拟南芥和西瓜 PIP5K 基因家族之间的直系同源基因和旁系同源基因。利用 MEME 5.1.1 软件^[26] (<http://meme-suite.org/>)对 PIP5K 蛋白序列进行 motif 分析,其中 motif 基序长度范围为 6~50,数量最大为 10,然后利用 WebLogo 3.7.4 (<http://weblogo.threeplusone.com/>)生成 motif 基序标签^[27]。利用 GSDS2.0 (Gene Structure Display Server) (<http://gsds.gao-lab.org/index.php>)进

行基因结构分析^[28]。截取西瓜 PIP5K 编码基因上游 2 000 bp 序列,利用 PlantCARE 软件 (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/>)进行顺式作用元件分析^[29],结果用 Evolview v2 绘制^[30]。

1.4 西瓜 PIP5K 家族成员转录组分析

利用 NCBI-SRA 数据库 (<https://www.ncbi.nlm.nih.gov/sra/>)下载西瓜栽培品种 97103 的生长点、果柄、果肉、下胚轴、叶和根相关的 RNA-seq 数据 (SRP192188)^[31],然后利用 HISAT 2.2.0 软件^[32] (<https://daehwankimlab.github.io/hisat2/>)进行 RNA-seq 分析,R 语言的 Rsubread 包的 featureCounts 进行 TPM (Transcripts Per Million)表达定量分析^[33-34]。

1.5 西瓜花蕾材料的采集及表达分析

利用 TRNzol Universal 总 RNA 提取试剂 (TIANGEN,中国北京)提取西瓜雄性可育和不育花蕾总 RNA。利用 PrimeScript™ RT 试剂盒 (TaKaRa,中国大连)合成 cDNA。利用 Primer 5.0 设计西瓜 PIP5K 基因以及内参基因 ($\beta\text{-actin}$; Cla007792)的特异性引物(表 1)。采用罗氏 LightCycler96 实时荧光定量 PCR 仪 (Roche,Basel,Switzerland)进行基因表达量检测。Quantitative real-time PCR (qRT-PCR)反应体系如下:10 μL SYBR Green Master mix (2 $\times$),0.8 μL primer F (10 $\mu\text{mol/L}$),0.8 μL primer R (10 $\mu\text{mol/L}$),6.4 μL ddH₂O 和 2 μL cDNA 模板。PCR 反应程序如下:95 $^{\circ}\text{C}$ 预变性 30 s,然后 95 $^{\circ}\text{C}$ 5 s,60 $^{\circ}\text{C}$ 30 s 进行 50 个循环。以雄性可育花蕾(长度为 3~5 mm)为对照,计算其他花蕾中 ClaPIP5K 基因相对表达量,设置 3 次重复,计算方法采用 $2^{-\Delta\Delta\text{Ct}}$ 方法^[35]。采用 SPSS17 软件中的 t 检验对基因表达量进行统计分析。

2 结果与分析

2.1 ClaPIP5K 成员的染色体定位和共线性分析

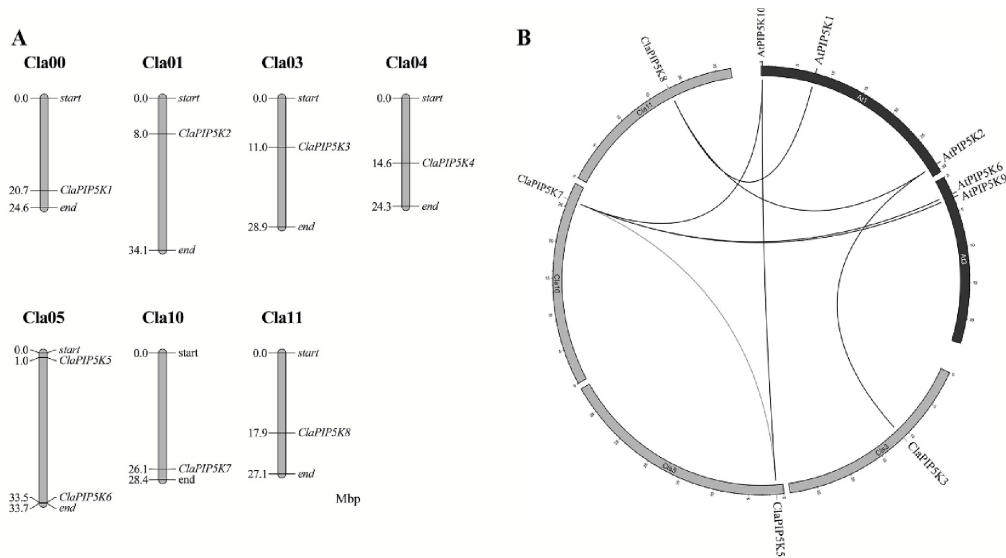
西瓜基因组中共鉴定出 8 个 PIP5K 家族成员,根据其在染色体上的位置进行编号 (ClaPIP5K1 ~ ClaPIP5K8)。结果如图 1-A 所示,西瓜数据库中将无法定位的染色体片段组合起来称为 0 号染色体,除了 ClaPIP5K1 定位在 0 号染色体外,其余 7 个 ClaPIP5K 基因分布在其他 6 条染色体上,其中 5 号染色体分布 2 个 ClaPIP5K 基因,其他 5 条染色体分别分布 1 个

表 1 西瓜 *PIP5K* 基因的 qRT-PCR 引物
Table 1 qRT-PCR primers of watermelon *PIP5K* genes

基因 Genes	上游引物 (5'→3') Forward primer	下游引物 (5'→3') Reverse primer
<i>β-actin</i>	GTCGTACAACAGGTATTGTG	AAGGTCCAGACGGAGGATAG
<i>ClaPIP5K1</i>	CGGTAAGCGAGCCAACGATTCC	ATGTGAGGACCTGGAGCAGTGG
<i>ClaPIP5K2</i>	TACCACGCCGAGAAGATCCTCTG	CTCCATTCGCCGACGTACATACAC
<i>ClaPIP5K3</i>	TCAAGTCGGGTCGGATGGAAGG	CCGTCGTAGAAATCGCCGTTGG
<i>ClaPIP5K4</i>	CAGAAGAAGTTGGCGGTGTGGAG	ACCCTACCATCAACCGACATCCTC
<i>ClaPIP5K5</i>	CAGGGTGCCAGAGTGCTACAATG	TCTCCTACCTGACCCTCAAATGCC
<i>ClaPIP5K6</i>	GCACACCCGAGAGCCATCATTC	CTGGCAGAGTCGTGTTTCCTTGTC
<i>ClaPIP5K7</i>	ATACAGAGGCGGGAAGTCAGGAG	AGGCGGAGTCAACTGGGAACC
<i>ClaPIP5K8</i>	CCACGCAATCTACAGAGACGACAG	TCCGTTGGGTAGGTGCTTCTCC

ClaPIP5K 基因。为了检测 *ClaPIP5K* 和 *At-PIP5K* 基因间的进化关系,利用 MScanX 软件进行共线性分析。结果如图 1-B 所示,4 个 *ClaPIP5K* 基因与 5 个拟南芥 *AtPIP5K* 基因存在共线关系,如 *ClaPIP5K7* 分别与 *AtPIP5K6*、

AtPIP5K9 和 *AtPIP5K10* 共线, *ClaPIP5K8* 分别与 *AtPIP5K1* 和 *AtPIP5K2* 共线, *ClaPIP5K3* 与 *AtPIP5K2* 共线;此外,1 个 *ClaPIP5K* 基因对 (*ClaPIP5K5* ~ *ClaPIP5K7*)发生片段复制事件。



西瓜染色体编号为 Cla00~11,拟南芥染色体编号为 At1~5,染色体长度单位为 Mbp。黑色曲线表示西瓜和拟南芥 *PIP5K* 基因间的共线关系,灰色曲线表示西瓜 *PIP5K* 基因间的共线关系

Chromosome numbers of watermelon are Cla00~11,chromosome numbers of *Arabidopsis* are At1~5,and the length unit of chromosome is Mbp. Black curves indicate synteny relationship between watermelon and *Arabidopsis PIP5K* genes,and grey curves indicate synteny relationship among watermelon *PIP5K* genes

图 1 西瓜 *PIP5K* 家族成员的染色体分布(A)及共线性分析(B)

Fig. 1 Chromosome distribution (A) and syntenic analysis (B) of *PIP5K* family members in watermelon

2.2 *ClaPIP5K* 成员的蛋白和基因信息

ClaPIP5K 基因的 CDS 序列长度为 606~2 499 bp,蛋白序列长度为 201~832 aa,分子质量为 22.62~93.76 ku,等电点介于 6.6~8.87,其中 4 个蛋白呈碱性(表 2)。亚细胞定位分析显示多数 *ClaPIP5K* 蛋白位于细胞质或细胞膜上(表 2)。蛋白结构域分析显示,所有的 *ClaPIP5K* 成员在 C 端都具有 1 个起催化功能的 PIPKc 结构域,其中 6 个 *ClaPIP5K* 成员在 N 端还具有 7~8 个 MORN 结构域(图 2)。

表 2 西瓜 PIP5K 家族成员的基本信息

Table 2 Information of PIP5K family members in watermelon

基因名称 Gene name	基因号 Gene locus	编码序列 长度/bp CDS length	蛋白 Protein			亚细胞定位 PC Subcellular location by PC	亚细胞定位 EM Subcellular location by EM
			序列 长度/aa Length	分子 质量/ku MW	等电点 pI		
<i>ClaPIP5K1</i>	Cla000575	606	201	22.62	6.6	Ex	Cy
<i>ClaPIP5K2</i>	Cla005895	2 232	743	84.58	8.02	Pm	Cy,Ex
<i>ClaPIP5K3</i>	Cla002795	2 409	802	—	—	Pm	Cy
<i>ClaPIP5K4</i>	Cla003061	2 358	785	88.89	6.78	Pm	Cy,Ex
<i>ClaPIP5K5</i>	Cla021180	813	270	30.83	6.03	Pm	Cy
<i>ClaPIP5K6</i>	Cla009934	2 274	757	86.11	8.51	Pm	Cy,Ex
<i>ClaPIP5K7</i>	Cla017778	2 499	832	93.76	8.63	Pm	Cy,Ex
<i>ClaPIP5K8</i>	Cla023149	2 385	794	89.53	8.87	Pm	Cy

注:MW. 分子质量; pI. 等电点; Cy. 细胞质; Ex. 胞外; Pm. 质膜; PC. ProtComp version 9.0; EM. Euk-mPLoc 2.0。
Note:MW. Molecular mass; pI. isoelectric points; Cy. Cytoplasm; Ex. Extracell; Pm. Plasma membrane; PC. ProtComp version 9.0;
EM. Euk-mPLoc 2.0.

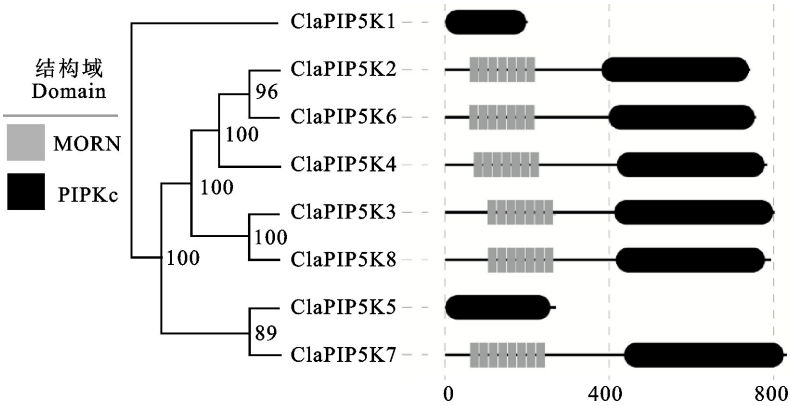


图 2 西瓜 PIP5K 家族成员的蛋白结构域分布

Fig. 2 Protein domain distribution of PIP5K family members in watermelon

2.3 *ClaPIP5K* 成员的蛋白序列进化树及保守基序分析

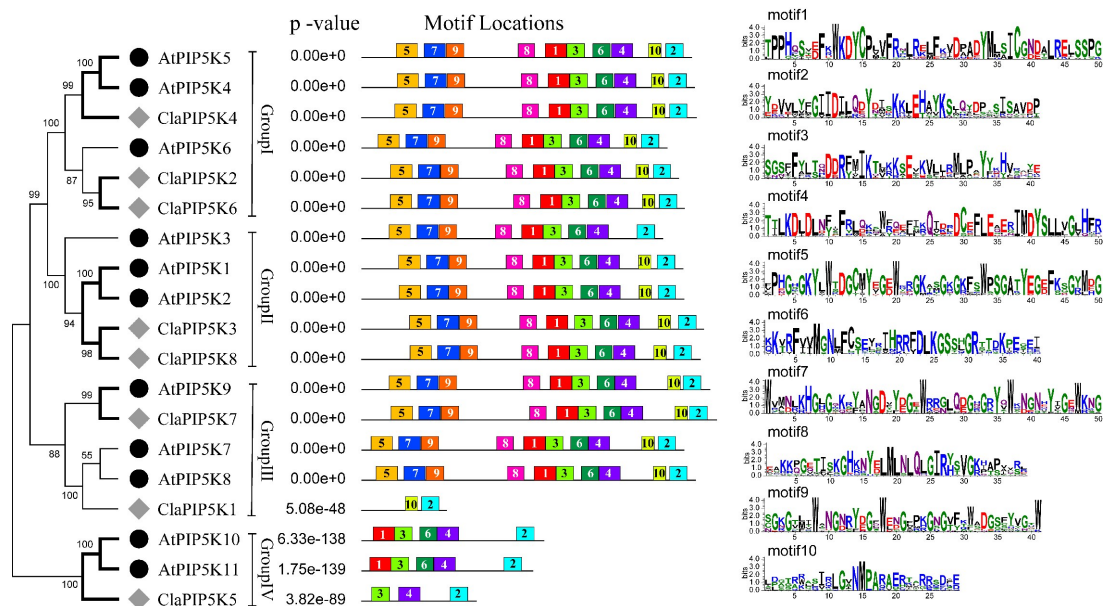
为了检测 *ClaPIP5K* 和 *AtPIP5K* 成员的蛋白序列进化情况以及同源序列对(组)所包含的保守基序数量及种类,利用 MEGA X 软件进行进化树分析,Orthofinder 软件进行同源序列分析, MEME 软件进行保守基序分析。结果如图 3 所示,蛋白序列进化树显示 *ClaPIP5K* 和 *AtPIP5K* 成员被分成 4 个组(Group I ~ IV),其中 Group I 所含成员最多,为 6 个,Group II 和 III 次之,分别为 5 个,Group IV 最少,为 3 个。蛋白保守基序分析显示 *ClaPIP5K* 蛋白中含有 10 个保守基序,每个保守基序含有 29~50 个氨基酸,其中 motif1、4、5 和 7 所含氨基酸数量最多,平均为 50 个, motif10 所含最少,为 29 个。此外, Group I ~ III 亚组中,除了 *ClaPIP5K1* 和 *AtPIP5K3* 外,其他 PIP5K 成员中均含有 10 个 motif。同源序列分析发现 *ClaPIP5K* 和 *AtPIP5K* 之间存在 5 个

同源序列对(组),其中 4 个同源序列对(组) (*ClaPIP5K4/AtPIP5K4/AtPIP5K5*、 *ClaPIP5K3/ClaPIP5K8/AtPIP5K1/AtPIP5K2*、 *ClaPIP5K7/AtPIP5K9* 和 *ClaPIP5K2/ClaPIP5K6*)所含 motif 数量和种类均相似。而 *ClaPIP5K1* 和 *ClaPIP5K5* 由于部分氨基酸的缺失,导致其 motif 种类和数量上与同组的 *AtPIP5K* 成员有所差异。

2.4 *ClaPIP5K* 成员的基因序列进化树及基因结构分析

为了检测 *ClaPIP5K* 和 *AtPIP5K* 成员的基因序列进化情况以及同源序列的内含子和外显子分布情况,利用 MEGA X 软件进行进化树分析, Orthofinder 软件进行同源序列分析, GSDS 软件进行基因结构分析。结果如图 4 所示,基因序列的系统进化树和同源序列分析结果与蛋白序列的(图 3)基本一致;基因结构分析显示 75%(6/8) *ClaPIP5K* 基因含有 7 个内含子(图 4),并且 4 个同源序列对(组) (*ClaPIP5K4/AtPIP5K4/At-*

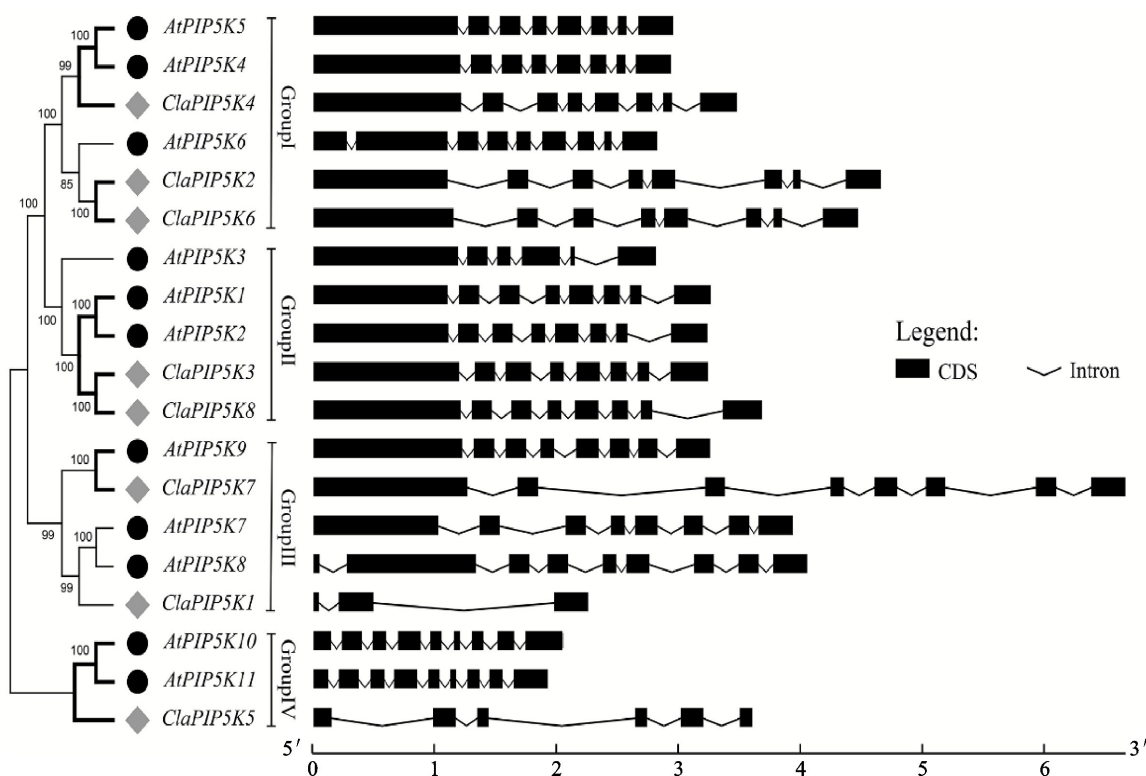
PIP5K5、*ClaPIP5K3* / *ClaPIP5K8* / *AtPIP5K1* / *ClaPIP5K6*) 所含内含子数量均为 7 个。另外,与 *AtPIP5K2*、*ClaPIP5K7* / *AtPIP5K9* 和 *ClaPIP5K2* / *AtPIP5K6* 相比, *ClaPIP5K2* 和 *ClaPIP5K6* 的第



加粗黑线代表同源序列对(组) Bold black lines indicate homologous pair (Group)

图 3 拟南芥和西瓜 *PIP5K* 成员的蛋白序列进化树及 motif 分布

Fig. 3 Protein sequence phylogenetic tree and motif distribution of PIP5K members in *Arabidopsis* and watermelon



加粗黑线代表同源序列对(组) The bold black lines indicated the homologous pair (Group)

图 4 拟南芥和西瓜 *PIP5K* 成员的基因序列进化树及内含子与外显子分布

Fig. 4 Phylogenetic tree of gene sequence and intron/exon distribution of PIP5K members in *Arabidopsis* and watermelon

一个内含子发生了缺失现象, *ClaPIP5K1* 和 *ClaPIP5K5* 出现了部分内含子和外显子的缺失现象,导致与同组的 *AtPIP5K* 基因在基因结构上有所差异。

2.5 *ClaPIP5K* 成员的顺式作用元件分析

为了明确*ClaPIP5K* 基因启动子所含的顺式作用元件种类及分布情况,利用 PlantCARE 软件进行分析。结果如图 5 所示,*ClaPIP5K* 基因启动子处主要包含 2 大类顺式作用元件:①激素响应元件,如响应生长素 IAA 的 AuxRR-core(1/8;含元件基因数/总基因数,下同),响应赤霉素 GA 的 TATC-box(1/8)、GARE-motif(2/8)和 P-box(4/8),响应水杨酸 SA 的 TCA-element(5/8),响应脱落酸 ABA 的 ABRE(4/8)和响应茉莉酸甲酯

MeJA 的 CGTCA-motif(3/8)等元件;②胁迫响应元件,如响应干旱的 MYB 结合位点 MBS(5/8),响应低温的 LTR(4/8),响应防御与胁迫的 TC-richrepeats(5/8),厌氧诱导的 ARE(7/8)和响应损伤的 WUN-motif(6/8)等元件。其中 *ClaPIP5K4* 基因所含顺式作用元件种类和数量最多,共 9 种 16 个(1 个 AuxRR-core、1 个 GARE-motif、1 个 TCA-element、2 个 ABRE、1 个 CGTCA-motif、1 个 LTR、1 个 TC-richrepeats、7 个 ARE 和 1 个 WUN-motif), *ClaPIP5K3* 所含顺式作用元件种类和数量最少,共 3 种 6 个(1 个 MBS、4 个 ARE 和 1 个 WUN-motif)。另外, *ClaPIP5K2* 和 *ClaPIP5K6* 所含顺式作用元件种类和数量较相似。

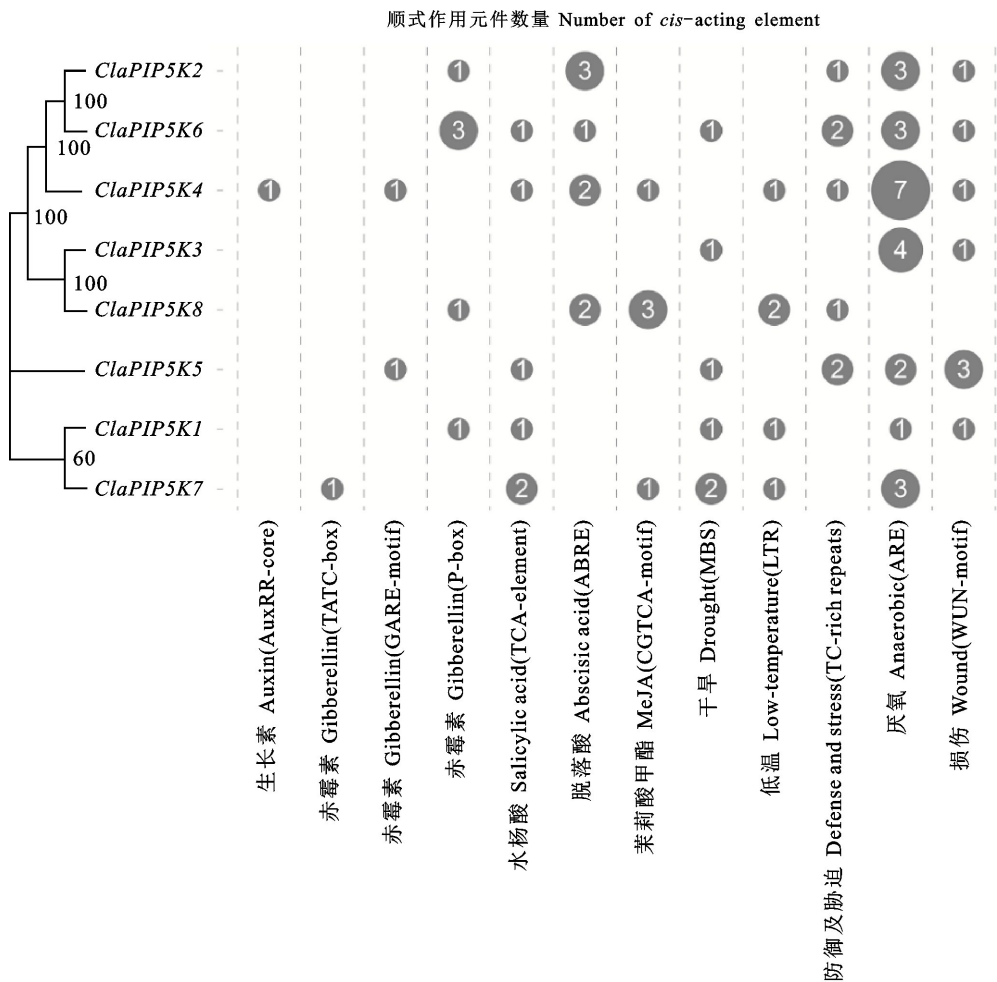


图 5 西瓜 PIP5K 成员的顺式作用元件预测

Fig. 5 Prediction of *cis*-acting elements of PIP5K members in watermelon

2.6 *ClaPIP5K* 基因在不同组织中的表达模式分析

为探索*ClaPIP5K* 基因在西瓜生长发育过程

中潜在的生物学功能,对已公布的西瓜栽培品种 97103 的生长点、果柄、果肉、下胚轴、叶和根的 RNA-seq 数据进行组织表达模式分析。结果如

表 3 所示,4 个*ClaPIP5K* 基因在 6 个组织的不同发育时期表现出较高的表达量(TPM>10),其中*ClaPIP5K1* 基因在授粉 34 d 后的生长点中表达量最高,*ClaPIP5K3* 基因在授粉 18 d 后的叶片中表达量最高, *ClaPIP5K8* 和 *ClaPIP5K7* 基因分别在授粉 10 d 和 26 d 后的果肉中表达量最高。此外,*ClaPIP5K5* 基因仅在生长点的 4 个时期表现出较高的表达量,具有明显的组织表达特异性。另外,本试验利用 qRT-PCR 技术检测了*ClaPIP5K* 基因在西瓜雄性可育和雄性不育花蕾中的表达分析。结果如图 6 所示,西瓜雄性可育植株花蕾和花药发育正常、饱满(图 6-A),而不育植株花蕾和花药发育不正常、干瘪(图 6-B);另

外,基因表达分析中(图 6-C),除了*ClaPIP5K5* 基因外,其他 7 个*ClaPIP5K* 基因在西瓜雄性可育和不育花蕾中均检测出不同程度的表达量。其中,与可育植株的 7~9 mm 长的花蕾对比后发现,4 个 *ClaPIP5K* 基因 (*ClaPIP5K1*、*ClaPIP5K3*、*ClaPIP5K7* 和 *ClaPIP5K8*) 在不育花蕾中具有明显的较高表达量,然而 *ClaPIP5K2* 表现出明显的较低表达量;此外,在 3~5 mm 和 7~9 mm 长的花蕾中,*ClaPIP5K4* 基因在不育花蕾的表达量均明显低于可育花蕾,然而 *ClaPIP5K6* 基因出现相反的表达趋势。以上结果表明*ClaPIP5K* 基因可能在西瓜生长点、叶片、果肉以及花蕾花药等组织生长发育方面具有重要作用。

表 3 *ClaPIP5K* 基因在西瓜栽培品种‘97103’的不同组织中的表达分析

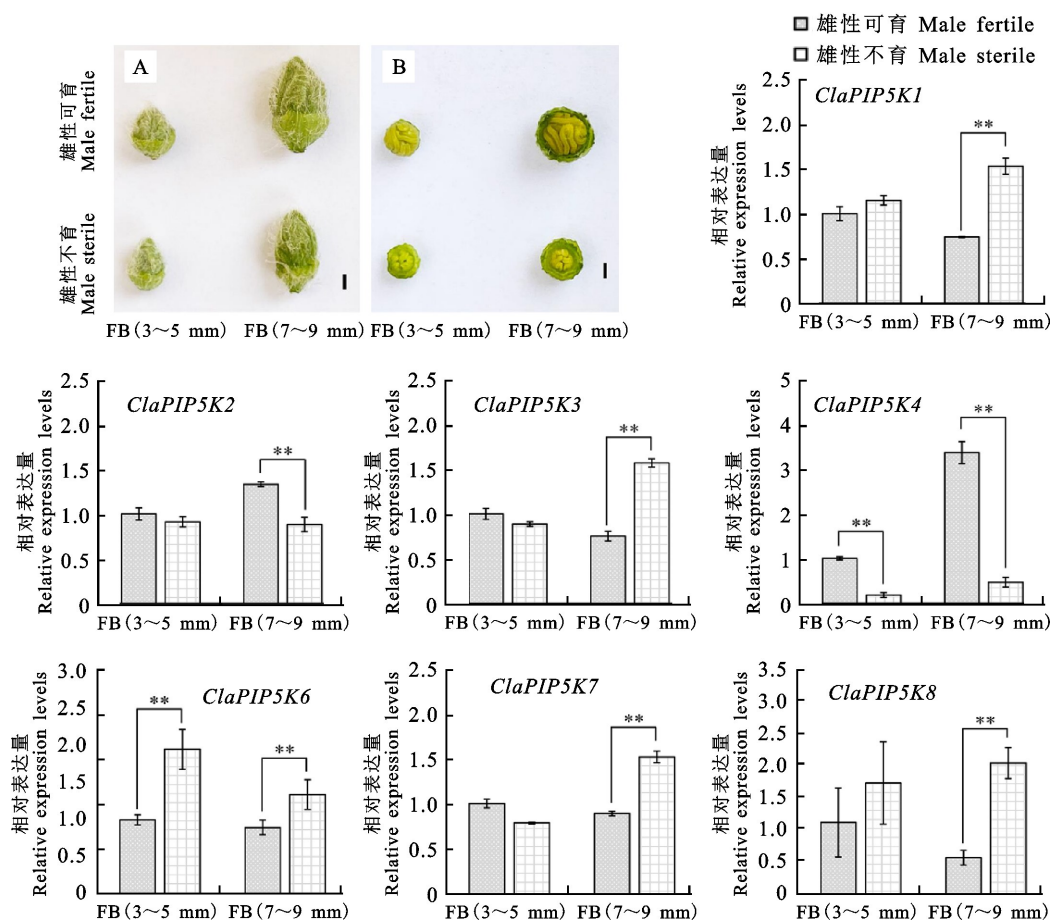
Table 3 Expression analysis of *ClaPIP5K* genes in different tissues of watermelon cultivar ‘97103’

组织 Tissue	授粉后时间/d Time after pollination	基因 Genes							
		<i>ClaPIP5K1</i>	<i>ClaPIP5K2</i>	<i>ClaPIP5K3</i>	<i>ClaPIP5K4</i>	<i>ClaPIP5K5</i>	<i>ClaPIP5K6</i>	<i>ClaPIP5K7</i>	<i>ClaPIP5K8</i>
生长点 Apical point	10	24.74	2.62	12.63	1.60	9.52	3.83	9.46	9.75
	18	19.08	3.12	13.36	2.61	21.21	5.58	13.24	16.52
	26	33.06	2.47	14.26	2.21	7.64	4.24	10.73	11.80
	34	43.93	3.50	12.99	1.94	24.78	4.95	15.88	7.68
果柄 Carpodium	10	22.05	0.23	13.86	0.14	1.08	2.11	23.68	7.82
	18	17.23	0.14	10.36	0.10	0.52	3.10	38.39	14.71
	26	19.66	0.15	10.49	0.18	0.76	1.55	28.57	11.10
	34	26.14	0.50	11.69	0.47	0.90	1.65	34.85	7.17
果肉 Fruit flesh	10	19.36	0.04	11.08	0.03	0.07	2.44	19.12	48.27
	18	23.92	0.02	13.11	0.02	0.29	1.93	58.19	20.45
	26	21.70	0.00	10.57	0.00	0.14	1.25	93.93	15.54
	34	19.26	0.03	13.81	0.00	0.19	1.76	67.01	10.09
下胚轴 Hypocotyl	10	23.47	0.16	13.19	0.12	0.61	1.32	25.07	11.52
	18	20.89	0.05	13.46	0.33	0.32	1.75	37.99	16.28
	26	17.24	0.07	11.82	0.17	0.26	1.92	29.33	19.96
	34	23.90	0.17	13.85	0.16	0.50	1.66	26.22	12.16
叶 Leaves	10	9.90	0.16	21.83	0.65	0.05	0.75	20.72	8.68
	18	10.59	0.16	34.29	2.07	0.00	2.17	28.12	7.43
	26	11.78	0.17	14.86	1.04	0.07	0.47	14.00	8.47
	34	12.80	0.07	11.28	1.42	0.18	0.50	15.09	15.69
根 Roots	10	9.54	0.25	25.69	0.12	0.12	1.70	27.32	24.34
	18	10.51	0.15	27.88	0.18	0.12	1.72	29.03	23.62
	26	9.02	0.12	25.16	0.05	0.13	1.13	29.88	24.99
	34	9.56	0.12	24.72	0.11	0.11	1.00	22.12	31.95

3 讨论与结论

磷脂酰肌醇 4-磷酸 5-激酶(PIP5K)是植物内一种磷酸类激酶,在不同植物中的成员数量有所不同。拟南芥中发现含有 11 个 PIP5K 成员,大

致分成 4 组(Ⅰ~Ⅳ)。除了第Ⅳ组含有 2 个 At-PIP5K 成员外,其他组均含有 3 个。西瓜含有 8 个 *ClaPIP5K* 成员,除了第Ⅰ组所含 *ClaPIP5K* 成员数量与拟南芥一样外,其他组的数量均比拟南芥的少 1 个,说明这 3 组(Ⅱ~Ⅳ)的 *ClaPIP5K*



A. 花蕾形态特征; B. 花药形态特征; C. *ClaPIP5K* 基因相对表达量; FB (3~5 mm): 3~5 mm 长的花蕾; FB (7~9 mm): 7~9 mm 长的花蕾; 标尺长度为 1 mm; * *: $P < 0.01$

A. phenotypic characterization of floral buds; B. phenotypic characterization of anthers; C. relative expression levels of *ClaPIP5K* genes; FB (3~5 mm): 3~5 mm length of floral buds; FB (7~9 mm): 7~9 mm length of floral buds; Scalebars = 1 mm; * *: $P < 0.01$

图 6 *ClaPIP5K* 基因在雄性可育系和不育系花蕾中的表达分析

Fig. 6 Expression analysis of *ClaPIP5K* genes in floral buds of watermelon male fertile and sterile lines

成员在进化过程中出现了缺失现象。除了 C 端具有起催化功能的 PIPKc 结构域外, N 端多 MORN 结构域也是植物大多数 PIP5K 家族成员所特有的特征^[2-3]。例如, 拟南芥中有 9 个 AtPIP5Ks 具有 7~8 个 MORN 结构域, 水稻中有 5 个 OsPIP5Ks 具有 7~8 个 MORN 结构域^[2], 而西瓜中也发现 6 个 *ClaPIP5K* 成员具有 7~8 个 MORN 结构域(图 1)。亚细胞定位预测显示多数 *ClaPIP5K* 定位在细胞质或细胞膜上, 说明这些基因可能在细胞质或细胞膜上行使其功能。以前的研究也证实了 *PIP5K* 基因作为一种细胞膜传感器, 在磷脂类代谢中发挥着重要作用^[36]。

通过进化树、同源序列、保守基序和基因结构等分析, 发现位于同一组的拟南芥和西瓜 *PIP5K* 成员在保守基序种类和基因结构上具有高度的相

似性。例如 4 个同源序列对(组) (*ClaPIP5K4*/AtPIP5K4/ AtPIP5K5、*ClaPIP5K3*/ *ClaPIP5K8*/ AtPIP5K1/ AtPIP5K2、 *ClaPIP5K7*/AtPIP5K9 和 *ClaPIP5K2*/ *ClaPIP5K6*)都具有 10 个 motifs 和 7 个内含子。但是也发现了一些差异, 如 AtPIP5K6 与 *ClaPIP5K2*/ *ClaPIP5K6* 的 motif 种类分布上相似, 但由于 AtPIP5K6 在 N 端多了 1 个内含子, 其基因结构与 *ClaPIP5K2*/ *ClaPIP5K6* 有所不同; *ClaPIP5K1* 和 *ClaPIP5K5* 由于大量序列的丢失, 导致与同组的 AtPIP5K 在保守基序分布和基因结构上差异较大。以上结果表明多数 PIP5K 成员在进化过程中相对保守。

基因复制事件在基因家族成员数量扩增、生物多样性和物种形成等方面具有重要作用^[14]。本研究发现, 4 个 *ClaPIP5K* 基因与 5 个拟南芥

AtPIP5K 基因存在复制情况,1 个 *ClaPIP5K* 基因对 (*ClaPIP5K5*-*ClaPIP5K7*) 存在片段复制现象,说明全基因组复制或者片段复制在 *ClaPIP5K* 基因复制中扮演着重要作用。

启动子分析显示 *ClaPIP5K* 基因含有响应 IAA、GA、ABA、SA、MeJA 等激素以及干旱、低温等非生物胁迫等方面的顺式作用元件,这与玉米 *PIP5K* 基因启动子生信分析结果相似^[37]。基于已公布的西瓜 RNA-seq 数据,分析了 *ClaPIP5K* 基因在不同组织的不同发育时期的表达模式。结果显示 4 个 *ClaPIP5K* 基因 (*ClaPIP5K1*、*ClaPIP5K3*、*ClaPIP5K7* 和 *ClaPIP5K8*) 在生长点、果柄、果肉、下胚轴、叶和根的不同发育时期均表现出较高的表达量,而 *ClaPIP5K5* 只在生长点表现出较高的表达量,说明 *ClaPIP5K* 基因可能协同或者单独地参与调控西瓜不同组织形态建成相关途径。此外,以前研究表明拟南芥 *AtPIP5K4* 或 *AtPIP5K5* 缺失突变株的花药和花粉发育受到严重抑制^[9-10]。本研究发现 *ClaPIP5K4* 基因在雄性不育花蕾发育的 2 个阶段中一直维持着显著的低表达量状态,并且 *ClaPIP5K4* 与 *AtPIP5K4* 和 *AtPIP5K5* 为同源基因,表明 *ClaPIP5K4* 基因可能参与调控西瓜花药或花粉的生长发育。总之,本研究首次利用生物信息学方法对西瓜 *PIP5K* 基因家族进行了鉴定,包括蛋白理化性质、结构域、保守基序、基因结构、进化情况、启动子以及组织表达模式等方面的分析,为后期研究其功能奠定了一定基础。

参考文献 Reference:

- [1] 梁塔娜,张艳欣,李丽丽,等. 植物中 *PIP5K* 基因家族调控功能的研究进展[J]. 贵州农业科学,2018,46(7):23-26.
LIANG T N,ZHANG Y X,LI L L,*et al.* Research progress of phosphatidylinositol 4-phosphate 5-kinase gene family[J]. *Guizhou Agricultural Sciences*,2018,46(7):23-26.
- [2] 马 晖. 水稻磷脂酰肌醇信号传导途径中-磷酸磷脂酰肌醇激酶(*OsPIP1*)的分子遗传分析[D]. 北京:中国科学院研究生院(上海生命科学研究院),2004.
MA H. Molecular genetic analysis of phosphatidylinositol phosphate kinases (*OsPIP1*)-phosphatidylinositol signaling pathway in *Oryza sativa* [D]. Beijing: Graduate University of Chinese Academy of Sciences (Shanghai Institutes for Biological Sciences),2004.
- [3] 孙 冲,姚 远,耿梦婷,等. 木薯磷脂酰肌醇磷酸 5-激酶 *PIP5K9* 基因的克隆及表达分析[J]. 分子植物育种,2016,14(9):2290-2296.
SUN CH,YAO Y,GENG M T,*et al.* Cloning and expression analysis of phosphatidylinositol phosphate 5-kinase *PIP5K9* from *Cassava* [J]. *Molecular Plant Breeding*,2016,14(9):2290-2296.
- [4] LEE Y,KIM Y W,JEON B W,*et al.* Phosphatidylinositol 4,5-bisphosphate is important for stomatal opening[J]. *Plant Journal*,2007,52(5):803-816.
- [5] LOU Y,GOU J Y,XUE H W. *PIP5K9*, an *Arabidopsis* phosphatidylinositol monophosphate kinase, interacts with a cytosolic invertase to negatively regulate sugar-mediated root growth[J]. *Plant Cell*,2007,19:163-181.
- [6] KUSANO H,TESTERINK C,VERMEER J E M,*et al.* The *Arabidopsis* phosphatidylinositol phosphate 5-Kinase *PIP5K3* is a key regulator of root hair tip growth[J]. *Plant Cell*,2008,20(2):367-380.
- [7] MEI Y,JIA W J,CHU Y J,*et al.* *Arabidopsis* phosphatidylinositol monophosphate 5-kinase 2 is involved in root gravitropism through regulation of polar auxin transport by affecting the cycling of PIN proteins[J]. *Cell Research*,2012,22(3):581-597.
- [8] MA H,XU S P,LUO D,*et al.* *OsPIP1*, a rice phosphatidylinositol monophosphate kinase, regulates rice heading by modifying the expression of floral induction genes[J]. *Plant Molecular Biology*,2004,54(2):295-310.
- [9] ISCHEBECK T,STENZEL I,HEILMANN I. Type B phosphatidylinositol-4-phosphate 5-kinases mediate *Arabidopsis* and *Nicotiana tabacum* pollen tube growth by regulating apical pectin secretion[J]. *Plant Cell*,2008,20:3312-3330.
- [10] SOUSA E,KOST B,MALHO R. *Arabidopsis* phosphatidylinositol-4-monophosphate 5-kinase 4 regulates pollen tube growth and polarity by modulating membrane recycling[J]. *Plant Cell*,2008,20:3050-3064.
- [11] UGALDE J M,RODRIGUEZ-FURLAN C,RYCKE R D,*et al.* Phosphatidylinositol 4-phosphate 5-kinases 1 and 2 are involved in the regulation of vacuole morphology during *Arabidopsis thaliana* pollen development[J]. *Plant Science*,2016,250:10-19.
- [12] HIRAYAMA T,OHITO C,MIZOGUCHI T,*et al.* A gene encoding a phosphatidylinositol-specific phospholipase C is induced by dehydration and salt stress in *Arabidopsis thaliana* [J]. *Proceedings of the National Academy of Sciences of the United States of America*,1995,92(9):3903-3907.
- [13] MIKAMI K,KATAGIRI T,IUCHI S,*et al.* A gene encoding phosphatidylinositol-4-phosphate 5-kinase is induced by water stress and abscisic acid in *Arabidopsis thaliana* [J]. *Plant Journal*,1998,15(4):563-568.
- [14] ZHANG Z B,LI Y T,HUANG K,*et al.* Genome-wide systematic characterization and expression analysis of the phosphatidylinositol 4-phosphate 5-kinases in plants[J]. *Gene*,2020,756:144915.
- [15] GUO SH G,ZHANG J G,SUN H H,*et al.* The draft genome of watermelon (*Citrullus lanatus*) and resequencing

- of 20 diverse accessions[J]. *Nature Genetics*, 2013, 45(1): 51-58.
- [16] EL-GEHALI S, MISTRY J, BATEMAN A, *et al.* The Pfam protein families database in 2019[J]. *Nucleic Acids Research*, 2019, 47(D1): D427-D432.
- [17] FINN R D, CLEMENTS J, EDDY S R. HMMER web server: interactive sequence similarity searching[J]. *Nucleic Acids Research*, 2011, 39 (Web Server issue): W29-W37.
- [18] CHOU K CH, SHEN H B. A new method for predicting the subcellular localization of eukaryotic proteins with both single and multiple sites: Euk-mPLoc 2.0[J]. *PLoS One*, 2010, 5(4): e9931.
- [19] WANG Y P, TANG H B, DEBARRY J D, *et al.* MCS-canX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity[J]. *Nucleic Acids Research*, 2012, 40(7): e49.
- [20] KRZYWINSKI M, SCHEIN J, BIROL I, *et al.* Circos: an information aesthetic for comparative genomics[J]. *Genome Research*, 2009, 19(9): 1639-1645.
- [21] KUMAR S, STECHER G, LI M, *et al.* MEGA X: Molecular evolutionary genetics analysis across computing platforms[J]. *Molecular Biology and Evolution*, 2018, 35(6): 1547-1549.
- [22] FELSENSTEIN J. Confidence limits on phylogenies: An approach using the bootstrap[J]. *Evolution*, 1985, 39(4): 783-791.
- [23] SAITOU N, NEI M. The neighbor-joining method: A new method for reconstructing phylogenetic trees[J]. *Molecular Biology and Evolution*, 1987, 4(4): 406-425.
- [24] NEI M, KUMAR S. *Molecular Evolution and Phylogenetics*[M]. New York: Oxford University Press, 2000: 237-296.
- [25] EMMS D M, KELLY S. OrthoFinder: phylogenetic orthology inference for comparative genomics[J]. *Genome Biology*, 2019, 20(1): 238-252.
- [26] BAILEY T L, WILLIAMS N, MISLEH C, *et al.* MEME: discovering and analyzing DNA and protein sequence motifs[J]. *Nucleic Acids Research*, 2006, 34 (Web Server issue): 369-373.
- [27] CROOKS G E, HON G, CHANDONIA J M, *et al.* WebLogo: A sequence logo generator[J]. *Genome Research*, 2004, 14(6): 1188-1190.
- [28] HU B, JIN J P, GUO A Y, *et al.* GSDB 2.0: an upgraded gene feature visualization server[J]. *Bioinformatics*, 2015, 31(8): 1296-1297.
- [29] LESCOT M, DEHAIS P, THIJS G, *et al.* PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences[J]. *Nucleic Acids Research*, 2002, 30(1): 325-327.
- [30] HE Z L, ZHANG H K, GAO SH H, *et al.* Evolview v2: an online visualization and management tool for customized and annotated phylogenetic trees[J]. *Nucleic Acids Research*, 2016, 44 (W1): W236-W241.
- [31] GUO SH G, ZHAO SH J, SUN H H, *et al.* Resequencing of 414 cultivated and wild watermelon accessions identifies selection for fruit quality traits[J]. *Nature Genetics*, 2019, 51(11): 1616-1623.
- [32] KIM D, LANGMEAD B, SALZBERG S L. HISAT: a fast spliced aligner with low memory requirements[J]. *Nature Methods*, 2015, 12(4): 357-360.
- [33] LIAO Y, SMYTH G K, SHI W. featureCounts: an efficient general-purpose program for assigning sequence reads to genomic features[J]. *Bioinformatics*, 2014, 30(7): 923-930.
- [34] LIAO Y, SMYTH G K, SHI W. The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads[J]. *Nucleic Acids Research*, 2019, 47(8): e47.
- [35] LIVAK K J, SCHMITTGEN T D. Analysis of relative gene expression data using real-time quantitative PCR and the 2^{-ΔΔC_T} method[J]. *Methods*, 2001, 25(4): 402-408.
- [36] LIU A ZH, SUI D X, WU D Q, *et al.* The activation loop of PIP5K functions as a membrane sensor essential for lipid substrate processing[J]. *Science Advances*, 2016, 2(11): e1600925.
- [37] 周玉琼. 玉米 *PIP5K* 基因启动子的克隆与功能分析[D]. 合肥: 安徽农业大学, 2012.
- ZHOU Y Q. Cloning and functional analysis of *PIP5K* gene promoter in *Zea mays* [D]. Hefei: Anhui Agricultural University, 2012.

Identification of *PIP5K* Gene Family in Watermelon (*Citrullus lanatus*) and Its Expression Analysis in Male Sterile Floral Buds

ZHANG Gaoyuan¹, DING Qian² and WEI Bingqiang¹

(1. College of Horticulture, Gansu Agricultural University, Lanzhou 730070, China;

2. College of Floriculture, Weifang Engineering Vocational College, Qingzhou Shandong 262500, China)

Abstract In order to explore the function of phosphatidylinositol 4-phosphate 5-kinase (*PIP5K*) genes in watermelon, the watermelon *PIP5K* gene family was identified and their characteristics and expression patterns were analyzed by using bioinformatics method in this study. The results showed that there were 8 *PIP5K* members in watermelon, which were distributed on six chromosomes. Most genes were located in cytoplasm or cell membrane. Syntenic analysis showed that there were synteny relationships between 4 *ClPIP5K* and 5 *AtPIP5K* genes, and one *ClPIP5K* gene pair was identified as segmental duplication. The phylogenetic tree analysis showed that *ClPIP5K* and *AtPIP5K* genes were divided into 4 groups, and the genes had similar conserved motifs and gene structure in the same group. The promoter analysis showed that the *ClPIP5K* genes contained *cis*-acting elements related to hormone and stress response. The expression pattern analysis showed that *ClPIP5K* genes displayed different expression levels in the apical point, carpel, fruit flesh, leaf, hypocotyl, root and floral bud of watermelon.

Key words Watermelon; *PIP5K*; Gene family; Male sterile; Expression pattern

Received 2020-09-17

Returned 2020-11-05

Foundation item Open Scientific Research Startup Funds for Ph. D Degree Holders of Gansu Agricultural University(No. 2017RCZX-30).

First author ZHANG Gaoyuan, male, Ph. D, lecturer. Research area: watermelon breeding and molecular biology. E-mail: zhanggy@gsau.edu.cn

(责任编辑:潘学燕 Responsible editor: PAN Xueyan)