

薄壳山核桃 *SnRK2* 基因家族鉴定及表达分析

马文娟^{a,b,①}, 朱凯凯^{a,b,①}, 谌梦云^{a,b}, 赵娟^{a,b}, 谭鹏鹏^{a,b}, 彭方仁^{a,b,②}

(南京林业大学: a. 南方现代林业协同创新中心, b. 林学院, 江苏 南京 210037)

摘要: 通过生物信息学方法从薄壳山核桃 [*Carya illinoensis* (Wangenh.) K. Koch] 全基因组中鉴定出 *SnRK2* 基因家族成员, 并对该基因家族成员编码蛋白的理化性质及基因的结构、保守基序、共线性关系、顺式作用元件和表达模式进行了分析。结果表明: 从薄壳山核桃全基因组中鉴定出 13 个 *SnRK2* 基因, 命名为 *CiSnRK2.1* 至 *CiSnRK2.13*, 编码氨基酸残基数为 336~366, 理论等电点均小于 pI 7, 且所有编码蛋白为亲水蛋白。13 个 *CiSnRK2* 与 10 个拟南芥 [*Arabidopsis thaliana* (Linn.) Heynh.] *SnRK2* 的氨基酸序列高度保守, 包含相同的结构域, 如脱落酸 (ABA) 诱导调节结构域。13 个 *CiSnRK2* 分为 3 组, 薄壳山核桃与山核桃 (*C. cathayensis* Sarg.) 的 *SnRK2* 亲缘关系较近; 每组中的 *CiSnRK2* 具有相似的保守基序和基因结构。13 个 *CiSnRK2* 基因存在 12 对共线性关系, *Ka/Ks* 分析结果表明 *CiSnRK2* 基因进化过程主要受到正选择。*CiSnRK2* 基因的启动子包含丰富的环境胁迫类及激素类响应元件, 其中激素类响应元件以 ABA 响应元件最多。多数 *CiSnRK2* 基因在薄壳山核桃不同组织中有不同程度表达, 其中 *CiSnRK2.5*、*CiSnRK2.8*、*CiSnRK2.9* 和 *CiSnRK2.12* 的表达水平较低, *CiSnRK2.2*、*CiSnRK2.3* 和 *CiSnRK2.13* 的表达水平较高。经 100 $\mu\text{mol} \cdot \text{L}^{-1}$ ABA 处理 24 h 后, 除 *CiSnRK2.2*、*CiSnRK2.3*、*CiSnRK2.6* 外, 其余 10 个基因表达水平均上调。综合分析表明: 薄壳山核桃 *SnRK2* 基因家族成员具有 *SnRK2* 基因家族的基本特征; 该基因家族成员在不同组织中的表达特性存在差异, 部分 *CiSnRK2* 基因的表达呈现组织特异性; ABA 处理后部分 *CiSnRK2* 基因的表达水平显著上调, 由此推断这些基因可能参与薄壳山核桃的 ABA 诱导响应过程。

关键词: 薄壳山核桃; *SnRK2* 基因; 生物信息学分析; 组织特异性表达; ABA 诱导

中图分类号: Q946-33; S664.1 文献标志码: A 文章编号: 1674-7895(2023)02-0018-11

DOI: 10.3969/j.issn.1674-7895.2023.02.03

Identification and expression analysis of *SnRK2* gene family from *Carya illinoensis* MA Wenjuan^{a,b,①}, ZHU Kaikai^{a,b,①}, CHEN Mengyun^{a,b}, ZHAO Juan^{a,b}, TAN Pengpeng^{a,b}, PENG Fangren^{a,b,②} (Nanjing Forestry University; a. Co-Innovation Center for Sustainable Forestry in Southern China, b. College of Forestry, Nanjing 210037, China), *J. Plant Resour. & Environ.*, 2023, 32(2): 18-28, 46

Abstract: The *SnRK2* gene family members were identified from the whole genome of *Carya illinoensis* (Wangenh.) K. Koch by using bioinformatics methods, and the physicochemical properties of coding proteins of the gene family members, and their structures, conserved motifs, collinear relationships, *cis*-acting elements and expression patterns were analyzed. The results show that 13 *SnRK2* genes are identified from the whole genome of *C. illinoensis* and named *CiSnRK2.1* to *CiSnRK2.13*, the numbers of coding amino acid residues are 336-366, their theoretical isoelectric points are all smaller than pI 7, and all coding proteins are hydrophilic proteins. The amino acid sequences of 13 *CiSnRK2*s and 10 *SnRK2*s from *Arabidopsis thaliana* (Linn.) Heynh. are highly conserved, which contain the same

收稿日期: 2022-10-24

基金项目: 中央财政林业科技推广示范项目(苏[2022]TG04); 江苏省自然科学基金项目(BK20190749)

作者简介: 马文娟(1998—), 女, 安徽亳州人, 硕士研究生, 主要从事植物分子生物学研究。

朱凯凯(1988—), 男, 江苏南通人, 博士, 讲师, 主要从事植物分子生物学研究。

①共同第一作者

②通信作者 E-mail: frpeng@njfu.edu.cn

引用格式: 马文娟, 朱凯凯, 谌梦云, 等. 薄壳山核桃 *SnRK2* 基因家族鉴定及表达分析[J]. 植物资源与环境学报, 2023, 32(2): 18-28, 46.

domains such as abscisic acid (ABA)-inducible regulatory domain. Thirteen *CiSnRK2*s are divided into 3 groups, and the genetic relationships of *SnRK2*s between *C. illinoensis* and *C. cathayensis* Sarg. are relatively close; *CiSnRK2*s in each group have similar conserved motifs and gene structures. There are 12 pairs of collinear relationships among 13 *CiSnRK2* genes, and the *Ka/Ks* analysis result shows that the evolution of *CiSnRK2* genes is mainly affected by positive selection. The promoters of *CiSnRK2* genes contain abundant environmental stress and hormone response elements, and the ABA response elements are the most in hormone response elements. Most *CiSnRK2* genes are expressed at different levels in different tissues of *C. illinoensis*, in which, the expression levels of *CiSnRK2.5*, *CiSnRK2.8*, *CiSnRK2.9* and *CiSnRK2.12* are relatively low, while those of *CiSnRK2.2*, *CiSnRK2.3* and *CiSnRK2.13* are relatively high. After $100 \mu\text{mol} \cdot \text{L}^{-1}$ ABA treatment for 24 h, except for *CiSnRK2.2*, *CiSnRK2.3* and *CiSnRK2.6*, the expression levels of the other 10 genes are up-regulated. The comprehensive analysis shows that the *SnRK2* gene family members from *C. illinoensis* have the basic characteristics of *SnRK2* gene family; there are differences in expression characteristics of these genes in different tissues, and the expressions of some *CiSnRK2* genes show tissue specificity; the expression levels of *CiSnRK2* genes are significantly up-regulated after ABA treatment, therefore, it is concluded that these genes may be involved in the response process of *C. illinoensis* under ABA induction.

Key words: *Carya illinoensis* (Wangenh.) K. Koch; *SnRK2* gene; bioinformatics analysis; tissue specific expression; ABA induction

蔗糖非发酵蛋白激酶 (sucrose non-fermenting-1-related protein kinase, SnRK) 是一类植物特异性丝氨酸/苏氨酸蛋白激酶,广泛存在于高等植物中,参与调节植物生长发育及逆境胁迫反应^[1-3]。根据序列相似性及结构域特征可将 SnRK 分为 SnRK1、SnRK2 和 SnRK3 3 个亚家族^[4],其中,SnRK1 亚家族主要参与调节碳、氮代谢^[5],SnRK2 和 SnRK3 亚家族参与植物响应环境胁迫^[6-7]。

SnRK2 基因首次经脱落酸 (ABA) 处理的小麦 (*Triticum aestivum* Linn.) 中分离出来,受到 ABA 诱导表达^[8-9]。研究表明:将小麦的 *SnRK2.4* 基因转入拟南芥 [*Arabidopsis thaliana* (Linn.) Heynh.] 中能够提高植株对非生物胁迫的耐受性^[10];研究人员在拟南芥中陆续克隆到 10 个 *SnRK2* 基因,其中受 ABA 诱导的基因有 5 个,这些基因能激活 ABA 信号通路,从而应对干旱胁迫;将含有 *AtSnRK2.8* 基因的过表达载体转化到拟南芥中,其幼苗抗旱性显著提高^[11-12];水稻 (*Oryza sativa* Linn.) 中 *SAPK8* 和 *SAPK9* 基因能够被 ABA 诱导^[13],且 *SnRK2.9-5A* 基因的过表达提高了水稻对非生物胁迫的抗性^[14];欧洲甜樱桃 [*Prunus avium* (Linn.) Linn.] 经 ABA 处理及干旱胁迫后,其 *SnRK2* 和 *PP2C* 基因表达上调,从而促进果实中花青素及干旱胁迫信号分子的累积^[15]。

薄壳山核桃 [*Carya illinoensis* (Wangenh.) K. Koch] 隶属于胡桃科 (Juglandaceae) 山核桃属 (*Carya* Nutt.), 原产于美国和墨西哥,是世界著名的坚果树

种之一^[16]。薄壳山核桃种仁富含不饱和脂肪酸、维生素及多种生物活性成分,可降低心脑血管疾病的发病率,同时为人体提供所需营养^[17]。近年来,研究者对薄壳山核桃抵御环境胁迫和品种改良方面取得一定进展^[18],但对于薄壳山核桃 *SnRK2* 基因家族的研究还未见报道。为了进一步研究薄壳山核桃 *SnRK2* 基因家族成员在 ABA 诱导下的分子响应机制,笔者在全基因组水平上对薄壳山核桃的 *SnRK2* 基因家族进行鉴定分析,并利用转录组学和 qRT-PCR 技术分别分析了 *CiSnRK2* 基因在不同组织中的表达特性及 ABA 诱导下的表达模式,以期为薄壳山核桃 *SnRK2* 基因功能研究及应用奠定理论基础。

1 材料和方法

1.1 材料

供试材料种植于南京林业大学美国山核桃试验基地 (东经 $119^{\circ}09'06''$ 、北纬 $31^{\circ}52'45''$), 为 9 年生健康薄壳山核桃品种 ‘Pawnee’ 的 1 年生实生苗,于 2022 年 6 月移植到气候室内,设置温度 (25 ± 2) $^{\circ}\text{C}$ 、空气相对湿度 70%~80%,每 2 天浇 1 次水,1 周后进行 ABA 处理。

1.2 方法

1.2.1 薄壳山核桃 *SnRK2* 家族成员鉴定及氨基酸序列分析 根据已报道的拟南芥 *SnRK2* 家族成员^[19],从拟南芥数据库 TAIR 中获得 10 个 *SnRK2* 氨基酸序

列,在 Pfam 数据库 (<http://pfam.xfam.org/>) 中下载含有 SnRK2 结构域的 HMM 文件 (PF00069),对氨基酸序列进行局部 BLASTp 比对,去除冗余序列^[20],筛选出包含已知保守结构域的薄壳山核桃 SnRK2 家族成员氨基酸序列,并采用相同的方法得到山核桃 (*C. cathayensis* Sarg.)、胡桃 (*Juglans regia* Linn.)、毛果杨 (*Populus trichocarpa* Torr. et Gray) 的 SnRK2 家族成员

氨基酸序列。对上述氨基酸序列的编码基因进行重命名,详见表 1。利用 ExPASy 在线网站 (<https://web.expasy.org/protparam/>) 分析薄壳山核桃 SnRK2 家族成员的理论等电点、理论相对分子量及疏水性,使用 CELLO V2.5 在线网站 (<http://cello.life.nctu.edu.tw/>) 对薄壳山核桃 SnRK2 家族成员进行亚细胞定位预测^[21]。

表 1 拟南芥、薄壳山核桃、山核桃、胡桃和毛果杨的 SnRK2 基因名和登录号

Table 1 Names and access numbers of SnRK2 genes from *Arabidopsis thaliana* (Linn.) Heynh., *Carya illinoensis* (Wangenh.) K. Koch, *C. cathayensis* Sarg., *Juglans regia* Linn., and *Populus trichocarpa* Torr. et Gray

基因名 Gene name	登录号 Access number	基因名 Gene name	登录号 Access number	基因名 Gene name	登录号 Access number
拟南芥 <i>Arabidopsis thaliana</i>		<i>CiSnRK2.10</i>	CiPaw.07G177700.1	<i>JrSnRK2.4</i>	JreChr06G11341
<i>AtSnRK2.1</i>	AT5g08590	<i>CiSnRK2.11</i>	CiPaw.08G048000.1	<i>JrSnRK2.5</i>	JreChr05G12310
<i>AtSnRK2.2</i>	AT3g50500	<i>CiSnRK2.12</i>	CiPaw.06G145800.1	<i>JrSnRK2.6</i>	JreChr05G12367
<i>AtSnRK2.3</i>	AT5g66880	<i>CiSnRK2.13</i>	CiPaw.05G050000.1	<i>JrSnRK2.7</i>	JreChr11G10571
<i>AtSnRK2.4</i>	AT1g10940	山核桃 <i>C. cathayensis</i>		<i>JrSnRK2.8</i>	JreChr03G10445
<i>AtSnRK2.5</i>	AT5g63650	<i>CcSnRK2.1</i>	CCA0638S0187	<i>JrSnRK2.9</i>	JreChr06G11295
<i>AtSnRK2.6</i>	AT4g33950	<i>CcSnRK2.2</i>	CCA1033S004	<i>JrSnRK2.10</i>	JreChr07G11743
<i>AtSnRK2.7</i>	AT4g40010	<i>CcSnRK2.3</i>	CCA0525S0140	<i>JrSnRK2.11</i>	JreChr04G10075
<i>AtSnRK2.8</i>	AT1g78290	<i>CcSnRK2.4</i>	CCA0507S0199	<i>JrSnRK2.12</i>	JreChr08G11848
<i>AtSnRK2.9</i>	AT2g23030	<i>CcSnRK2.5</i>	CCA1500S0025	毛果杨 <i>Populus trichocarpa</i>	
<i>AtSnRK2.10</i>	AT1g60940	<i>CcSnRK2.6</i>	CCA0507S0246	<i>PtSnRK2.1</i>	POPTR_0019s00790.1
薄壳山核桃 <i>Carya illinoensis</i>		<i>CcSnRK2.7</i>	CCA0696S0058	<i>PtSnRK2.2</i>	POPTR_0003s08240.1
<i>CiSnRK2.1</i>	CiPaw.11G072000.1	<i>CcSnRK2.8</i>	CCA1597S0007	<i>PtSnRK2.3</i>	POPTR_0004s15270.1
<i>CiSnRK2.2</i>	CiPaw.01G023500.1	<i>CcSnRK2.9</i>	CCA0514S0268	<i>PtSnRK2.4</i>	POPTR_0003s01110.1
<i>CiSnRK2.3</i>	CiPaw.02G010700.1	<i>CcSnRK2.10</i>	CCA0986S0064	<i>PtSnRK2.5</i>	POPTR_0004s14540.1
<i>CiSnRK2.4</i>	CiPaw.03G253100.1	<i>CcSnRK2.11</i>	CCA0659S0023	<i>PtSnRK2.6</i>	POPTR_0009s11010.1
<i>CiSnRK2.5</i>	CiPaw.04G178600.1	<i>CcSnRK2.12</i>	CCA0565S0054	<i>PtSnRK2.7</i>	POPTR_0007s05080.1
<i>CiSnRK2.6</i>	CiPaw.05G055400.1	胡桃 <i>Juglans regia</i>		<i>PtSnRK2.8</i>	POPTR_0002s09990.1
<i>CiSnRK2.7</i>	CiPaw.06G141000.1	<i>JrSnRK2.1</i>	JreChr04G11637	<i>PtSnRK2.9</i>	POPTR_0005s07370.1
<i>CiSnRK2.8</i>	CiPaw.03G145100.1	<i>JrSnRK2.2</i>	JreChr01G11053	<i>PtSnRK2.10</i>	POPTR_0004s22790.1
<i>CiSnRK2.9</i>	CiPaw.04G103100.1	<i>JrSnRK2.3</i>	JreChr02G10394	<i>PtSnRK2.11</i>	POPTR_0005s17290.1

1.2.2 氨基酸序列比对和系统发育树构建 使用 DNAMAN V7.0 软件对 *CiSnRK2* 及 *AtSnRK2* 家族成员氨基酸序列进行比对,使用 MEGA V7.0 软件中的邻接法根据拟南芥、薄壳山核桃、山核桃、胡桃和毛果杨 SnRK2 氨基酸序列的遗传距离构建系统发育树^[22],使用 Evolview 在线网站 (<https://evolgenius.info/>) 对结果进行可视化^[23]。

1.2.3 基因结构、保守基序及顺式作用元件分析 使用 TBtools 软件中的 Gene Structure 程序^[24] 对 *CiSnRK2* 进行基因结构分析;利用 MEME 在线网站 (<http://meme-suite.org/tools/meme>)^[25] 对 *CiSnRK2* 基因序列进行保守基序分析;从 Phytozome 数据库 (<https://phytozome-next.jgi.doe.gov/>) 中下载薄壳山

核桃全基因组序列,提取 13 个 *CiSnRK2* 基因上游 2 000 bp 区域序列,利用 PlantCARE 网站 (<https://phytozome-next.jgi.doe.gov/>)^[26] 预测顺式作用元件,用 TBtools 软件中的 Simple BioSequence Viewer 程序对结果进行可视化。

1.2.4 基因染色体定位和共线性分析 利用 TBtools 软件中的 Circos 程序对 13 个 *CiSnRK2* 基因的染色体定位进行可视化,使用 MCScanX 程序对 *CiSnRK2* 基因进行共线性分析。使用 MEGA V7.0 软件计算复制事件的同义替换率 (K_s) 和非同义替换率 (K_a),通过 K_a/K_s 分析选择压力, $K_a/K_s < 1$ 表示负选择, $K_a/K_s > 1$ 表示正选择, $K_a/K_s = 1$ 表示中性选择^[27]。

1.2.5 基因组织特异性表达分析 利用课题组获得

的薄壳山核桃转录组数据 (PRJNA799663) 从 SRA 数据库 (<https://submit.ncbi.nlm.nih.gov/subs/sra/>) 得到 *CiSnRK2* 在不同组织中的表达水平数据^[28]。利用 RSEM 软件^[29], 基于 FPKM 对所有 *CiSnRK2* 基因表达水平进行量化, 使用 TBtools 软件生成热图。

1.2.6 基因 ABA 诱导表达模式分析 向供试植株叶片喷施 $100 \mu\text{mol} \cdot \text{L}^{-1}$ ABA 溶液, 每组 3 株, 共 3 组。处理 0、3、6、12 和 24 h 后, 采集叶片, 于 $-80 \text{ }^{\circ}\text{C}$ 保存。

表 2 *CiSnRK2* 基因的 qRT-PCR 引物序列

Table 2 qRT-PCR primer sequences of *CiSnRK2* genes

基因名 Gene name	引物序列 (5'→3')		产物长度/bp Product length
	正向引物 Forward primer	反向引物 Reverse primer	
<i>CiSnRK2.1</i>	ACATTGCACCAGAGTTCTC	GGTATGTCGCCACCAACATTA	102
<i>CiSnRK2.2</i>	CAGCTCCTCGGTGAAGATATG	CAGTACTTCCGGAGCAATGTAA	110
<i>CiSnRK2.3</i>	AGTCCTCGGTTGAAGATTG	CAGTACTTCTGGAGCAATGTAGG	109
<i>CiSnRK2.4</i>	GGAGCTCGTGGCTATGAAAT	CCACCTCCTTGAAACGGATTA	119
<i>CiSnRK2.5</i>	GGAACCTCGTGGCCATGAAATA	CCACCTCCTTGAAACGGATTAT	119
<i>CiSnRK2.6</i>	CGGATAATTCTGAGGCTACCA	CAAGTCTCTCCATGTCTTC	102
<i>CiSnRK2.7</i>	CTCCAGAGTGCTGCCATTT	GATCTGCAGGGAGGTTCTTTAG	114
<i>CiSnRK2.8</i>	TGCTGGAGGAGAAGTATTTGAC	GCAGTAACTCACTCCGAAATA	103
<i>CiSnRK2.9</i>	GGAAGCTTATCACCACGCTTTA	CAAGACCTCCGGTGAATATAA	117
<i>CiSnRK2.10</i>	CAGTCTTCCATTACAAACCAAAG	GACACCACAGGACCAAACA	116
<i>CiSnRK2.11</i>	GTGTTCCATTACAAACCAAAGT	ACACCAGAGGACCAAACATC	113
<i>CiSnRK2.12</i>	GCGAAGATGAGGCGAGATATT	GGTTTCCATCCAGCAAGGTA	117
<i>CiSnRK2.13</i>	GAGCTCGTGGCCATGAAATA	CCACCTCCTTGAAACCGAATAA	118

以 cDNA 为模板进行 qRT-PCR 反应, 内参基因为 $\beta\text{-actin}$ ^[27]。反应体系总体积 $20.0 \mu\text{L}$, 包括 SYBR $10.0 \mu\text{L}$, $120 \text{ ng} \cdot \mu\text{L}^{-1}$ cDNA $1.0 \mu\text{L}$, $10 \mu\text{mol} \cdot \mu\text{L}^{-1}$ 上、下游引物各 $0.4 \mu\text{L}$, ddH_2O $8.2 \mu\text{L}$ 。反应程序: $95 \text{ }^{\circ}\text{C}$ 预变性 30 s ; $95 \text{ }^{\circ}\text{C}$ 变性 10 s ; $60 \text{ }^{\circ}\text{C}$ 退火 30 s ; $72 \text{ }^{\circ}\text{C}$ 延伸 30 s , 共循环 30 次。每个样品 3 次重复。根据扩增结果计算基因的相对表达量^[30]。

1.3 数据处理和统计分析

使用 IBM SPSS25 软件对数据进行统计分析; 采用单因素方差分析和多重比较进行显著性分析。

2 结果和分析

2.1 基因鉴定及其编码蛋白理化性质分析

从薄壳山核桃全基因组中筛选到 13 个 *SnRK2* 基因, 分别命名为 *CiSnRK2.1* 至 *CiSnRK2.13*, 并对其编码蛋白的理化性质进行分析, 结果见表 3。结果显示: 13 个 *CiSnRK2* 蛋白的氨基酸残基数为 $336 \sim 366$, 理论相对分子质量为 $38\ 007.50 \sim 41\ 354.85$, 理论等电

利用 RNA 提取试剂盒 (南京诺唯赞生物科技股份有限公司) 提取总 RNA, 利用 HiScript III All-in-one RT SuperMix Perfect for qPCR 反转录试剂盒 (南京诺唯赞生物科技股份有限公司) 合成 cDNA, 置于 $-20 \text{ }^{\circ}\text{C}$ 保存。使用 Primer Quest Tool 在线网站 (<https://sg.idtdna.com/PrimerQuest/Home/Index>) 设计 13 个 *CiSnRK2* 基因的引物 (表 2), 交由北京擎科生物科技股份有限公司合成。

点均小于 $\text{pI} 7$; 且 13 个 *CiSnRK2* 蛋白的 GRAVY 值均为负值, 说明所有的 *CiSnRK2* 蛋白均为亲水蛋白;

表 3 *CiSnRK2* 家族成员的理化性质¹⁾

Table 3 Physicochemical properties of *CiSnRK2* family members¹⁾

蛋白名 Protein name	N	<i>m</i>	pI	GRAVY 值 GRAVY value	SL
<i>CiSnRK2.1</i>	340	38 370.91	5.63	-0.35	Cy
<i>CiSnRK2.2</i>	366	41 246.92	4.75	-0.21	Nu
<i>CiSnRK2.3</i>	362	41 156.74	4.67	-0.27	Nu
<i>CiSnRK2.4</i>	355	40 771.18	5.94	-0.53	Nu
<i>CiSnRK2.5</i>	356	40 900.40	5.56	-0.51	Cy
<i>CiSnRK2.6</i>	363	41 354.85	4.87	-0.38	Nu
<i>CiSnRK2.7</i>	363	41 104.55	4.86	-0.34	Nu
<i>CiSnRK2.8</i>	341	38 691.85	5.26	-0.37	Cy
<i>CiSnRK2.9</i>	345	39 065.25	5.46	-0.42	Cy
<i>CiSnRK2.10</i>	338	38 212.81	5.45	-0.19	Cy
<i>CiSnRK2.11</i>	338	38 007.50	5.80	-0.21	Cy
<i>CiSnRK2.12</i>	336	38 232.76	6.19	-0.37	Cy
<i>CiSnRK2.13</i>	354	40 326.09	5.89	-0.45	Cy

¹⁾ N: 氨基酸残基数 Amino acid residue number; *m*: 理论相对分子质量 Theoretical relative molecular mass; pI: 理论等电点 Theoretical isoelectric point; SL: 亚细胞定位 Subcellular localization. Cy: 细胞质 Cytoplasm; Nu: 细胞核 Nucleus.

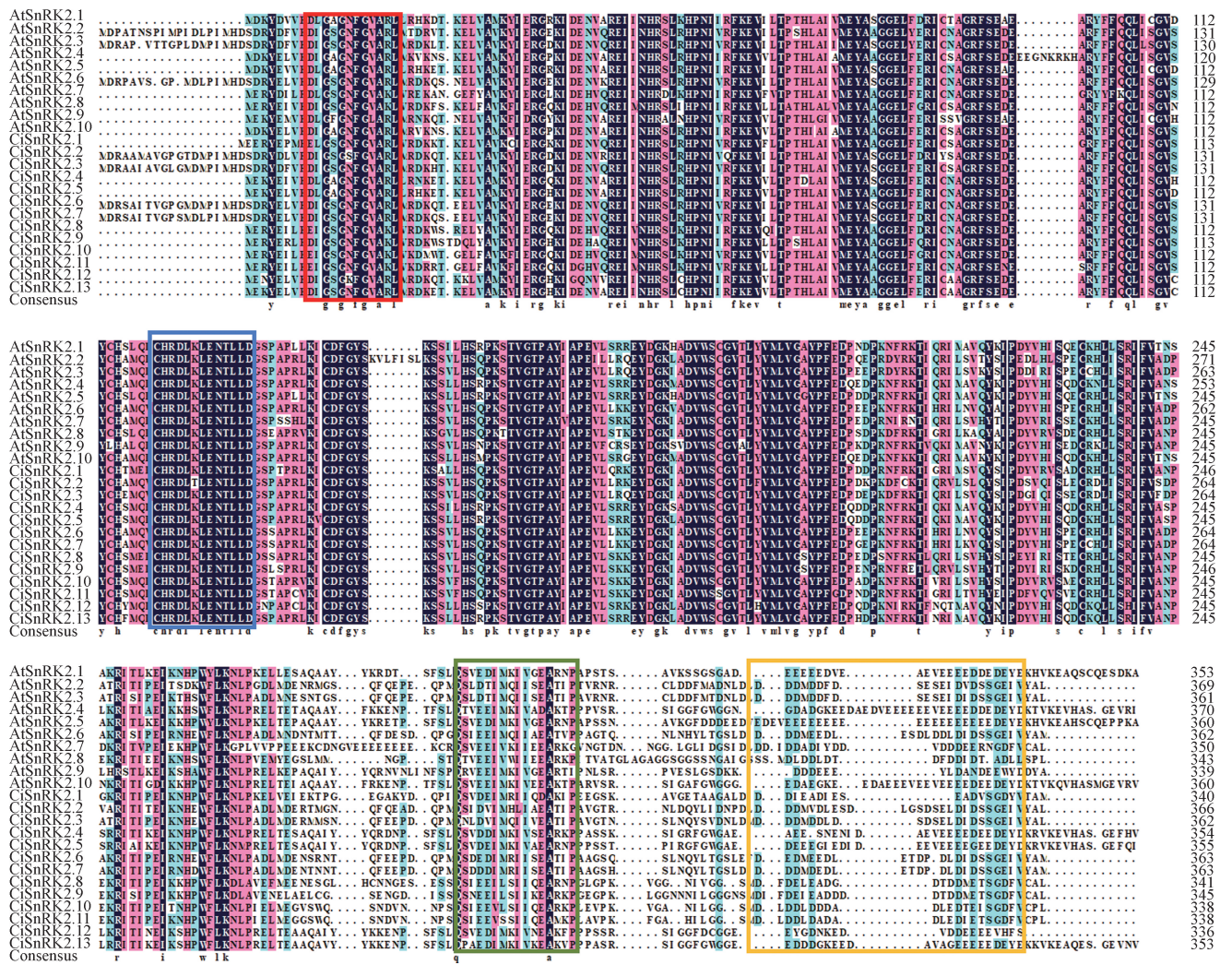
亚细胞定位预测 8 个 CiSnRK2 蛋白位于细胞质,5 个位于细胞核。

2.2 氨基酸序列比对

多重序列比对结果(图 1)显示:13 个 CiSnRK2 与 10 个拟南芥 SnRK2 的氨基酸序列高度保守, SnRK2 家族氨基酸序列包含 N 端和 C 端 2 个结构域,N 端为催化结构域,该结构域高度保守,包含 ATP 结合域和丝氨酸/苏氨酸蛋白激酶活性位点,是 SnRK2 的核心催化区域。C 端包含响应非生物胁迫结构域和 ABA 诱导调节结构域。

2.3 系统发育分析

根据薄壳山核桃、拟南芥、山核桃、胡桃和毛果杨的 SnRK2 氨基酸序列构建系统发育树,结果(图 2)显示:5 种植物的 SnRK2 家族成员构成 3 组, I 组包含 20 个 SnRK2,其中 4 个为 CiSnRK2, II 组包含 17 个 SnRK2,其中 4 个为 CiSnRK2, III 组包含 21 个 SnRK2,其中 5 个为 CiSnRK2。此外,薄壳山核桃 SnRK2 家族的每个成员基本上先与山核桃 SnRK2 家族成员聚在一起,表明薄壳山核桃与山核桃的 SnRK2 亲缘关系最近。

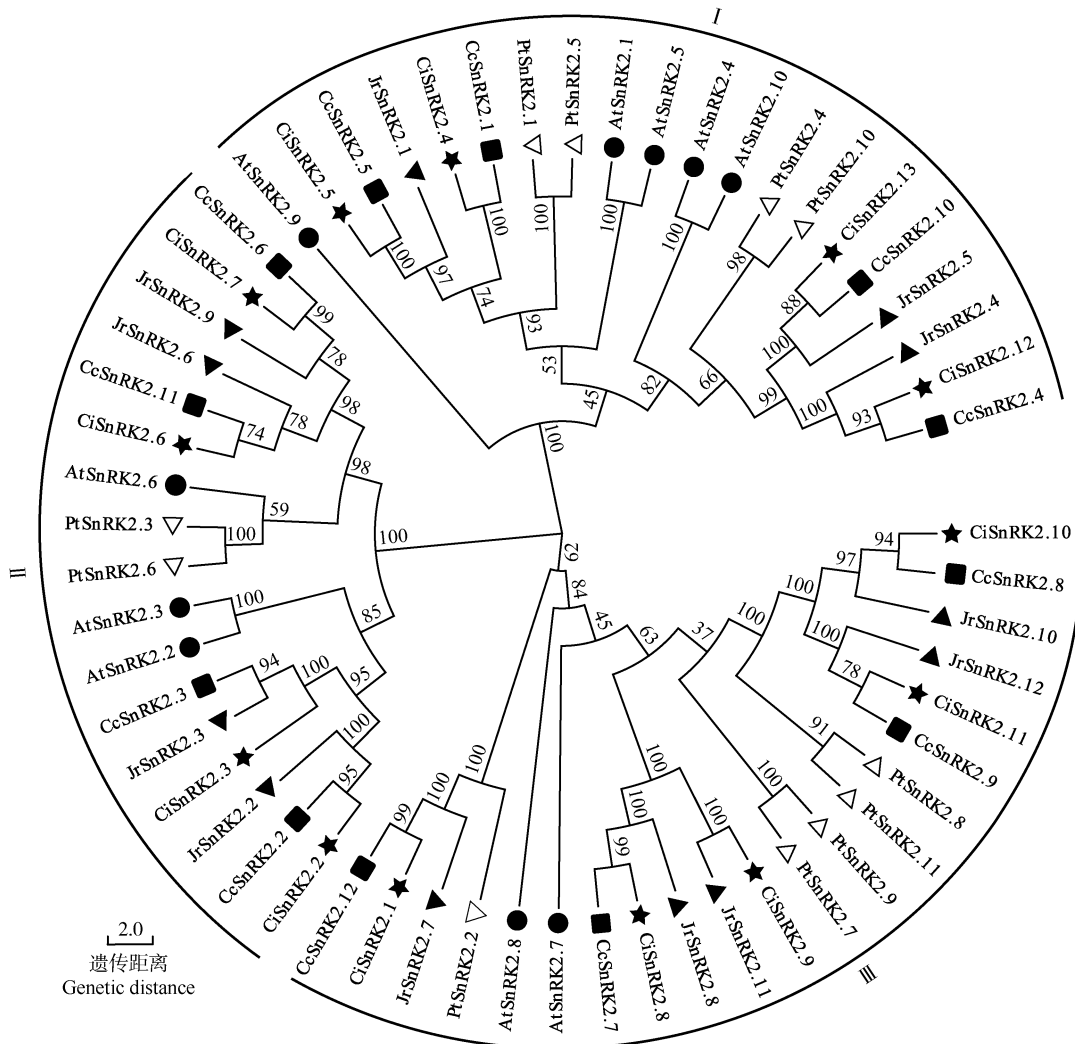


红色框表示 ATP 结合域 The red box represents ATP-binding domain; 蓝色框表示丝氨酸/苏氨酸蛋白激酶活性位点 The blue box represents the active site of serine/threonine protein kinase; 绿色框表示响应非生物胁迫结构域 The green box represents abiotic stress response domain; 黄色框表示 ABA 诱导调节结构域 The yellow box represents ABA-inducible regulatory domain.

At: 拟南芥 *Arabidopsis thaliana* (Linn.) Heynh.; Ci: 薄壳山核桃 *Carya illinoensis* (Wangenh.) K. Koch.

图 1 薄壳山核桃和拟南芥 SnRK2 家族成员氨基酸序列比对

Fig. 1 Amino acid sequence alignment of SnRK2 family members from *Carya illinoensis* (Wangenh.) K. Koch and *Arabidopsis thaliana* (Linn.) Heynh.



★: 薄壳山核桃 *Carya illinoensis* (Wangenh.) K. Koch.; ●: 拟南芥 *Arabidopsis thaliana* (Linn.) Heynh.; ■: 山核桃 *Carya cathayensis* Sarg.; ▲: 胡桃 *Juglans regia* Linn.; △: 毛果杨 *Populus trichocarpa* Torr. et Gray. 分支上的数值表示自展支持率 The values on the branches represent the bootstrap support rates.

图 2 基于氨基酸序列的 *SnRK2* 家族成员的系统发育树
Fig. 2 Phylogenetic tree based on amino acid sequences of *SnRK2* family members

2.4 蛋白保守基序和基因结构分析

CiSnRK2 蛋白的保守基序分析结果(图 3-A)显示:每组的 motif 数量及类型都高度相似,且每个 *CiSnRK2* 蛋白都包含 motif1、motif2、motif3、motif4 和 motif5。II 组 *CiSnRK2* 蛋白的 motif 数量明显多于其他 2 组,且含有 motif6 和 motif10,而 I 组的 *CiSnRK2* 蛋白则有别于其他 2 组,均不含有 motif7 和 motif9,但含有 motif8,III 组中的 *CiSnRK2.9* 与同组中的其他蛋白不同,不含有 motif9。

基因结构分析结果(图 3-B)显示:III 组基因长

度略短,均小于 4 000 bp, I 和 II 组基因长度为 4 000~7 000 bp。所有 *CiSnRK2* 基因的外显子数均为 9,除 *CiSnRK2.4*、*CiSnRK2.5* 和 *CiSnRK2.7* 内含子数为 9 外,其他成员内含子数均为 8,且每个 *CiSnRK2* 基因的最后一个外显子长度均较长。每组中的 *CiSnRK2* 基因结构均十分相似,推测他们可能具有相似的功能。

2.5 基因共线性分析及 *Ka/Ks* 分析

共线性分析结果(图 4)显示:13 个 *CiSnRK2* 基因分布在薄壳山核桃的 9 个染色体上,其中,

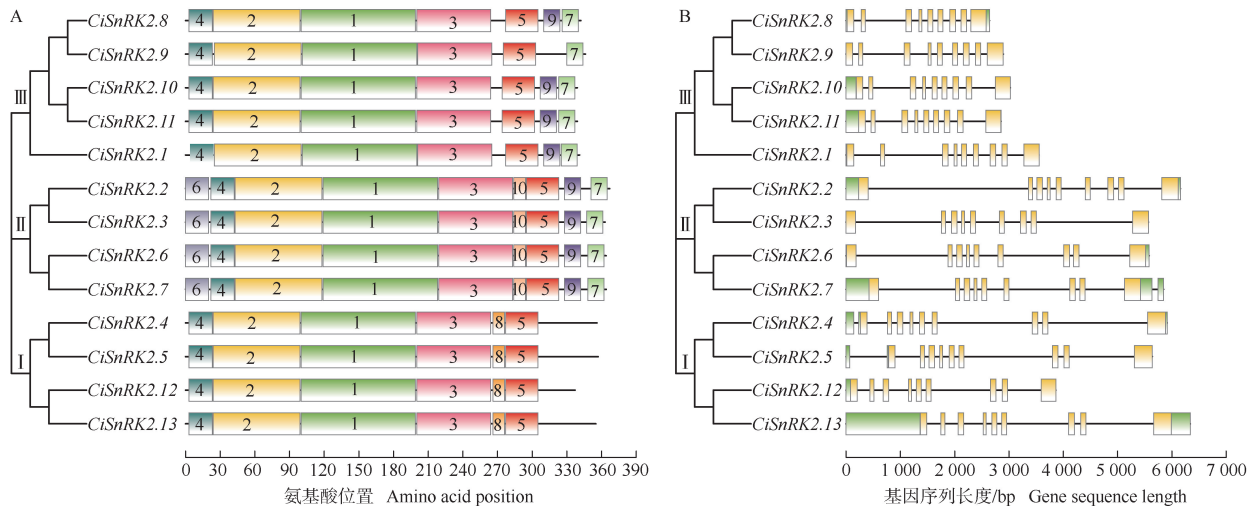


图 A 中不同颜色柱子表示不同保守基序 Different color pillars in figure A indicate different motifs; 图 B 中绿色柱子、黄色柱子和黑线分别表示非翻译区、外显子和内含子 The green pillars, yellow pillars, and black lines in figure B indicate untranslated regions, exons, and introns, respectively.

图 3 *CiSnRK2* 蛋白的保守基序 (A) 和 *CiSnRK2* 基因结构 (B)
 Fig. 3 Conserved motifs of *CiSnRK2* proteins (A) and *CiSnRK2* gene structures (B)

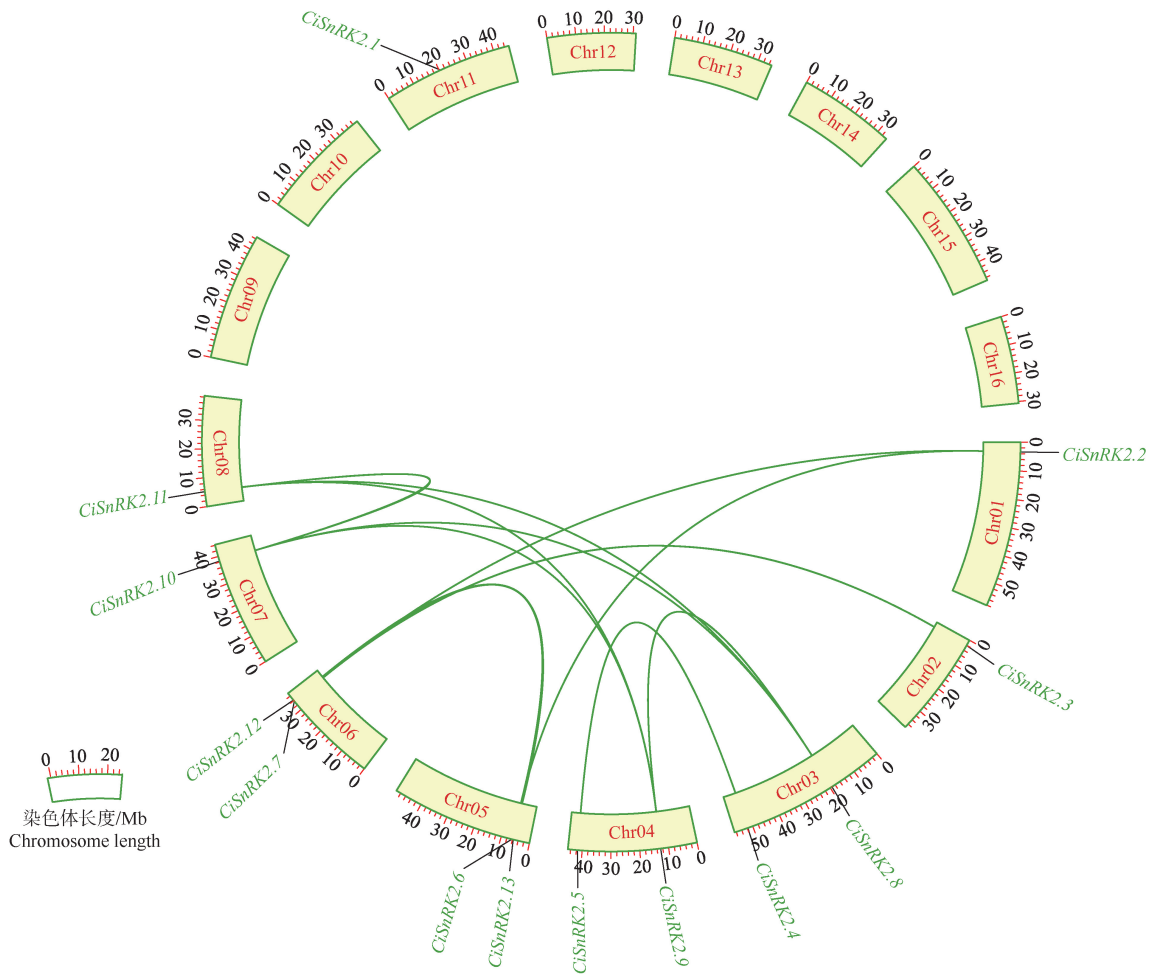


图 4 薄壳山核桃 *SnRK2* 基因家族成员的共线性分析
 Fig. 4 Collinearity analysis on *SnRK2* gene family members from *Carya illinoensis* (Wangenh.) K. Koch

CiSnRK2.1 分布在 11 号染色体上, 其余 12 个 *CiSnRK2* 基因均匀地分布在 1~8 号染色体上。13 个 *CiSnRK2* 基因存在 12 对共线关系, 其中 *CiSnRK2.6* 和 *CiSnRK2.13* 位于 5 号染色体, *CiSnRK2.7* 和 *CiSnRK2.12* 位于 6 号染色体, 属于串联重复, 其余 10 个基因对属于片段重复, 说明 *CiSnRK2* 基因家族的扩增方式可能是片段重复。

Ka/Ks 分析结果(表 4)显示: *CiSnRK2* 的 12 个共线基因对的 *Ka/Ks* 值为 0.67~1.90, 其中 8 个共线基因对的 *Ka/Ks* 值大于 1, 表明这些基因的进化过程主要受到正选择。

2.6 基因顺式作用元件分析

启动子顺式作用元件分析结果(图 5)显示: *CiSnRK2* 基因家族成员的启动子区主要包含响应环境胁迫和激素的顺式作用元件, 其中激素类响应元件

表 4 薄壳山核桃 *SnRK2* 基因家族成员中共线基因对的 *Ka/Ks* 分析
Table 4 *Ka/Ks* analysis on collinear gene pairs of *SnRK2* gene family members from *Carya illinoensis* (Wangenh.) K. Koch

共线基因对 Collinear gene pair	非同义替换率 Non-synonymous substitution rate (<i>Ka</i>)	同义替换率 Synonymous substitution rate (<i>Ks</i>)	<i>Ka/Ks</i>
<i>CiSnRK2.6/CiSnRK2.2</i>	1.78	1.57	1.13
<i>CiSnRK2.7/CiSnRK2.2</i>	2.38	1.96	1.22
<i>CiSnRK2.7/CiSnRK2.3</i>	1.96	1.58	1.24
<i>CiSnRK2.8/CiSnRK2.9</i>	1.50	2.24	0.67
<i>CiSnRK2.4/CiSnRK2.5</i>	1.26	1.61	0.78
<i>CiSnRK2.10/CiSnRK2.8</i>	1.82	1.65	1.11
<i>CiSnRK2.8/CiSnRK2.11</i>	2.00	1.82	1.10
<i>CiSnRK2.10/CiSnRK2.9</i>	2.22	1.39	1.59
<i>CiSnRK2.9/CiSnRK2.11</i>	1.84	1.90	0.97
<i>CiSnRK2.13/CiSnRK2.12</i>	1.82	2.10	0.87
<i>CiSnRK2.6/CiSnRK2.7</i>	2.23	1.17	1.90
<i>CiSnRK2.10/CiSnRK2.11</i>	1.73	0.98	1.77

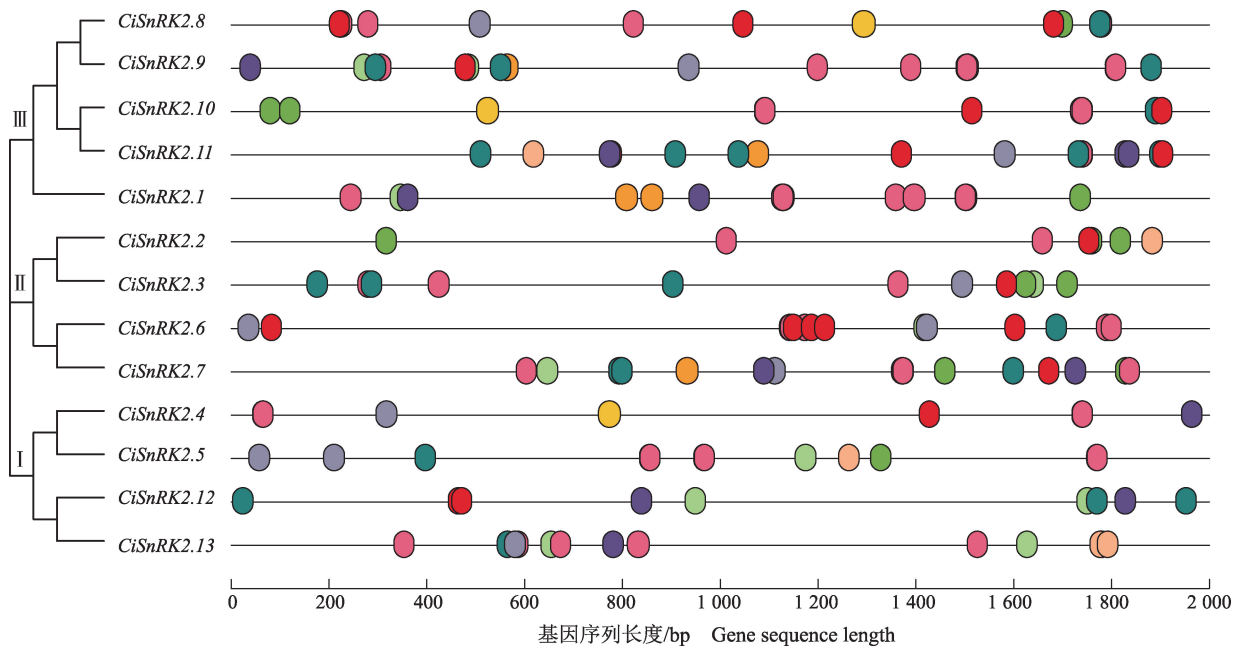


图 5 薄壳山核桃 *SnRK2* 基因家族成员顺式作用元件分析
Fig. 5 Analysis on *cis*-acting elements of *SnRK2* gene family members from *Carya illinoensis* (Wangenh.) K. Koch

主要包括 ABA 响应元件、茉莉酸甲酯响应元件、生长素响应元件、水杨酸响应元件和赤霉素响应元件。其中, ABA 响应元件的数量最多且分布最广, 其次为茉莉酸甲酯响应元件。

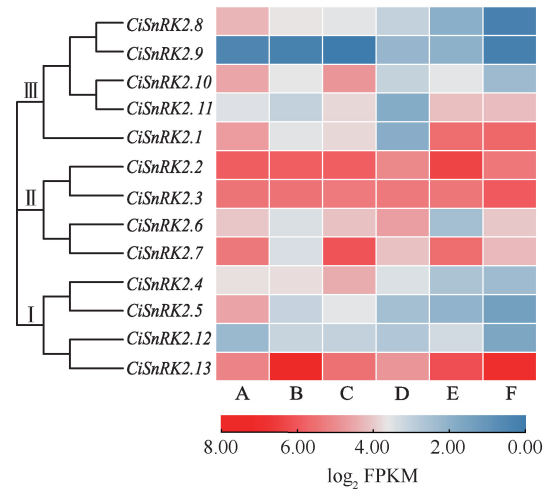
2.7 基因组织特异性表达分析

薄壳山核桃 *SnRK2* 基因家族成员的组织特异性表达分析结果(图 6)显示: *CiSnRK2.2*、*CiSnRK2.3* 和 *CiSnRK2.13* 在不同组织中均呈现较高的表达水平,

CiSnRK2.5 和 *CiSnRK2.8* 在雌花中高表达,在其他组织中的表达水平极低,*CiSnRK2.9* 和 *CiSnRK2.12* 在 6 个组织中的表达水平较低,*CiSnRK2.1* 和 *CiSnRK2.11* 在根和种子中表达量较高,*CiSnRK2.4* 和 *CiSnRK2.10* 在花和果实中表达量较高,*CiSnRK2.6* 在雌花、雄花、果实、叶片和种子中表达量均较高。

2.8 基因 ABA 诱导表达模式分析

100 $\mu\text{mol} \cdot \text{L}^{-1}$ ABA 处理下 *CiSnRK2* 基因在 24 h 内的相对表达量见表 5。结果显示:*CiSnRK2.1* 和 *CiSnRK2.11* 的相对表达量在 ABA 处理 3 h 显著升高;*CiSnRK2.1* 的相对表达量在 ABA 处理 6 h 达到峰值(较 0 h 升高了 3.38 倍)后逐渐下降;在 ABA 处理 24 h 后,除 *CiSnRK2.2*、*CiSnRK2.3*、*CiSnRK2.6* 外,其余 10 个基因的表达水平均上调,尤其是 *CiSnRK2.8* 和 *CiSnRK2.9* 的相对表达量较 0 h 分别升高了 6.18 和 9.65 倍,推测这些基因在响应 ABA 诱导中起正调控作用。



FPKM; 表达丰度 Expression abundance. A; 雌花 Female flower; B; 雄花 Male flower; C; 果实 Fruit; D; 叶片 Leaf; E; 根 Root; F; 种子 Seed.

图 6 薄壳山核桃 *SnRK2* 基因家族成员的组织特异性表达热图
Fig. 6 Heat map of tissue-specific expressions of *SnRK2* gene family members from *Carya illinoensis* (Wangenh.) K. Koch

表 5 100 $\mu\text{mol} \cdot \text{L}^{-1}$ ABA 处理下 *CiSnRK2* 基因的相对表达量 ($\bar{X} \pm SE$)¹⁾
Table 5 Relative expression levels of *CiSnRK2* genes under 100 $\mu\text{mol} \cdot \text{L}^{-1}$ ABA treatment ($\bar{X} \pm SE$)¹⁾

基因名 Gene name	在不同处理时间的相对表达量 Relative expression level at different treatment times				
	0 h	3 h	6 h	12 h	24 h
<i>CiSnRK2.1</i>	1.00±0.00c	2.89±0.56b	4.38±0.59a	3.55±0.18ab	2.31±0.20b
<i>CiSnRK2.2</i>	1.00±0.00a	0.76±0.08b	0.55±0.06b	0.72±0.03b	1.16±0.13a
<i>CiSnRK2.3</i>	1.00±0.00a	0.78±0.05bc	0.61±0.06c	0.75±0.01bc	0.81±0.04ab
<i>CiSnRK2.4</i>	1.00±0.00b	1.02±0.19b	1.55±0.44b	1.07±0.26b	3.36±0.24a
<i>CiSnRK2.5</i>	1.00±0.00b	0.69±0.10c	0.83±0.08bc	0.94±0.10bc	2.04±0.13a
<i>CiSnRK2.6</i>	1.00±0.00a	0.56±0.05b	0.61±0.04b	0.43±0.01b	1.00±0.12a
<i>CiSnRK2.7</i>	1.00±0.00b	0.77±0.02c	0.68±0.09c	0.77±0.04c	1.45±0.09a
<i>CiSnRK2.8</i>	1.00±0.00d	1.73±0.39d	3.68±0.64c	5.42±0.87b	7.18±0.34a
<i>CiSnRK2.9</i>	1.00±0.00b	0.98±0.29b	0.97±0.08b	1.10±0.09b	10.65±0.09a
<i>CiSnRK2.10</i>	1.00±0.00b	1.22±0.25b	1.33±0.30b	1.28±0.23b	2.01±0.14a
<i>CiSnRK2.11</i>	1.00±0.00c	1.58±0.05ab	1.55±0.10b	1.30±0.17bc	1.94±0.16a
<i>CiSnRK2.12</i>	1.00±0.00bc	0.62±0.17c	0.81±0.09bc	1.20±0.04b	2.15±0.20a
<i>CiSnRK2.13</i>	1.00±0.00bc	0.92±0.08c	1.07±0.08bc	1.22±0.13b	2.07±0.07a

¹⁾ 同行中不同小写字母表示同一基因的相对表达量在不同时间间差异显著 ($P < 0.05$) Different lowercases in the same row indicate the significant ($P < 0.05$) difference in relative expression level of the same gene at different treatment times.

3 讨论和结论

ABA 是植物响应胁迫及生长发育的重要植物激素^[31]。SnRK2 是植物特有的一类 Ser/Thr 蛋白激酶,在植物响应非生物胁迫以及 ABA 诱导过程中发挥重要作用^[32]。迄今为止,已有众多研究者在多种植物中鉴定出多个 *SnRK2* 基因,如拟南芥中 10 个^[19]、水

稻中 10 个^[13]、玉米 (*Zea mays* Linn.) 中 14 个^[32]、棉花 (*Gossypium hirsutum* Linn.) 中 20 个^[6] 等。本研究通过已知的 10 个拟南芥 SnRK2 氨基酸序列在薄壳山核桃全基因组中鉴定出 13 个 SnRK2,将其编码基因分别命名为 *CiSnRK2.1* 至 *CiSnRK2.13*。*CiSnRK2* 基因家族成员编码氨基酸残基数为 336~366,理论相对分子质量为 38 007.50~41 354.85,13 个 *CiSnRK2* 蛋白的 GRAVY 值均为负值,说明所有的 *CiSnRK2* 蛋

白均为亲水蛋白,这与陈娜娜等^[33]对苹果 (*Malus pumila* Mill.) *SnRK2* 基因家族成员蛋白质理化性质的研究结果类似。与拟南芥 *SnRK2* 进行多重序列比对发现, *CiSnRK2* 家族成员与拟南芥 *SnRK2* 家族成员的氨基酸序列高度保守,具有典型的 N 端和 C 端结构域,包含 ATP 结合域、经氨酸/苏氨酸蛋白激酶活性位点及 ABA 诱导调节结构域等,这与 Liu^[6]等在棉花中的研究结果一致。系统发育分析结果显示:5 种植物的 58 个 *SnRK2* 可分为 3 组,薄壳山核桃的 13 个 *SnRK2* 较为均匀地分布在 3 组中,并且与山核桃 *SnRK2* 家族成员有较近的亲缘关系。此外,保守基序和基因结构分析结果表明:同组的 *CiSnRK2* 蛋白具有相似的保守基序和基因结构,尤其是同组中 *CiSnRK* 蛋白的基序位置及排列顺序都十分相似。以上结果均表明植物中的 *SnRK2* 基因家族成员在进化过程中较为保守,没有发生太多变化。研究表明:大多数高等植物的 *SnRK2* 基因都包含 9 个外显子^[34],本文研究中 13 个 *CiSnRK2* 基因也包含 9 个外显子, *CiSnRK2.4* 和 *CiSnRK2.5* 含有 9 个内含子,其余 *CiSnRK2* 基因包含 8 个内含子,这与 Huang 等^[35]对小白菜 [*Brassica rapa* ssp. *chinensis* (Lour.) Hanelt] *SnRK2* 基因的研究结果较为一致。共线性分析发现,13 个 *CiSnRK2* 基因家族成员存在 12 对共线关系,有 8 对基因的 *Ka/Ks* 大于 1,4 对基因的 *Ka/Ks* 小于 1,基因复制会造成突变累积,这些基因在进化过程中主要受到正选择,即通过非同义突变导致的功能性改变来适应进化。

当植物受到外界环境如干旱、高盐、高温等非生物胁迫时,植物自身与抵抗这些胁迫的相关信号会通过信号转导途径转化为相应的抗性因子,从而激活与胁迫相关的转录因子,并与相应基因的顺式作用元件结合,启动相应基因的表达,对逆境胁迫做出应答^[36]。本研究结果表明:*CiSnRK2* 基因包含环境胁迫及激素类响应元件,所有 *CiSnRK2* 基因至少包含 1 种植物激素类响应元件,且激素类响应元件以 ABA 响应元件最多,由于植物激素类响应元件是参与植物生长以及应对各种胁迫反应的关键因子^[37],由此推断 *CiSnRK2* 基因家族成员可能在调节非生物胁迫中发挥重要作用。

组织特异性分析结果显示:13 个 *CiSnRK2* 基因中,Ⅱ组的 *CiSnRK2.2*、*CiSnRK2.3* 和 *CiSnRK2.13* 在各组织中表达水平较高,*CiSnRK2.5* 和 *CiSnRK2.8* 在雌

花中高表达,而在其他组织中的表达水平极低,*CiSnRK2.9* 和 *CiSnRK2.12* 在各组织中的表达水平较低,说明 *CiSnRK2* 基因具有多种组织表达模式。这 13 个 *CiSnRK2* 基因在 6 种组织中的表达结果表明:不同 *CiSnRK2* 基因具有不同的表达模式,在生物学功能上发挥不同作用。

大量研究表明:*SnRK2* 基因参与植物多种非生物胁迫反应,盐和干旱环境诱导水稻中的 *SAPK3*、*SAPK8* 和 *SAPK9* 的表达水平升高^[38],ABA 和聚乙二醇(PEG)处理后丹参 (*Salvia miltiorrhiza* Bunge) *SnRK2.7* 可能在干旱等导致的渗透胁迫过程中发挥重要作用^[39]。本研究中多数 *CiSnRK2* 基因对 ABA 诱导有响应。在 ABA 处理 24 h 后, *CiSnRK2.2*、*CiSnRK2.3*、*CiSnRK2.6* 的表达水平未发生显著变化,其余 9 个 *CiSnRK2* 基因的相对表达量均上调;*CiSnRK2.1*、*CiSnRK2.4*、*CiSnRK2.8* 和 *CiSnRK2.9* 的组织表达水平与其余 *CiSnRK2* 基因相比较低,但在 ABA 处理后的相对表达量显著升高;共线基因 *CiSnRK2.8* 和 *CiSnRK2.9* 的表达水平高于其余 *CiSnRK2* 基因,即在 ABA 诱导后显著上调表达。这些结果表明不同 *CiSnRK2* 基因在 ABA 诱导下的响应程度不同,可能发挥不同作用。

本研究共鉴定出 13 个薄壳山核桃 *SnRK2* 基因家族成员,可分为 3 组,具有 *SnRK2* 家族的基本特征,且同组成员间的基因结构与保守基序较为相似,并且在基因复制及进化过程中受到正选择。13 个 *CiSnRK2* 基因在不同组织中具有不同的表达模式,其中 *CiSnRK2.2*、*CiSnRK2.3* 和 *CiSnRK2.13* 在 6 个组织中高表达;在 ABA 诱导下,13 个 *CiSnRK2* 基因中大多数基因表达水平高于对照,共线基因 *CiSnRK2.8* 和 *CiSnRK2.9* 可能在应对 ABA 诱导时具有强烈的调控作用。

参考文献:

- [1] 刘涛,王萍萍,何红红,等. 草莓 *SnRK2* 基因家族的鉴定与表达分析[J]. 农业生物技术学报, 2019, 27(12): 2150-2163.
- [2] NAKASHIMA K, FUJITA Y, KANAMORI N, et al. Three *Arabidopsis* *SnRK2* protein kinases, SRK2D/*SnRK2.2*, SRK2E/*SnRK2.6*/OST1 and SRK2I/*SnRK2.3*, involved in ABA signaling are essential for the control of seed development and dormancy[J]. *Plant and Cell Physiology*, 2009, 50(7): 1345-1363.
- [3] BOUDSOCQ M, LAURIÈRE C. Osmotic signaling in plants. Multiple pathways mediated by emerging kinase families[J]. *Plant Physiology*, 2005, 138(3): 1185-1194.

- [4] HRABAK E M, CHAN C W M, GRIBSKOV M, et al. The arabidopsis CDPK-SnRK superfamily of protein kinases [J]. *Plant Physiology*, 2003, 132(2): 666-680.
- [5] CELENZA J L, CARLSON M. A yeast gene that is essential for release from glucose repression encodes a protein kinase [J]. *Science*, 1986, 233(4769): 1175-1180.
- [6] LIU Z, GE X Y, YANG Z R, et al. Genome-wide identification and characterization of *SnRK2* gene family in cotton (*Gossypium hirsutum* L.) [J]. *BMC Genetics*, 2017, 18: 54.
- [7] ZHANG Y H, WAN S Q, WANG W D, et al. Genome-wide identification and characterization of the *CsSnRK2* family in *Camellia sinensis* [J]. *Plant Physiology and Biochemistry*, 2018, 132: 287-296.
- [8] ANDERBERG R J, WALKER M K. Isolation of a wheat cDNA clone for an abscisic acid-inducible transcript with homology to protein kinases [J]. *Proceedings of the National Academy of Sciences of the United States of America*, 1992, 89: 10183-10187.
- [9] HOLAPPA L D, WALKER-SIMMONS M K. The wheat abscisic acid-responsive protein kinase mRNA, PKABA1, is up-regulated by dehydration, cold temperature, and osmotic stress [J]. *Plant Physiology*, 1995, 108(3): 1203-1210.
- [10] MAO X G, ZHANG H Y, TIAN S J, et al. TaSnRK2.4, an SNF1-type serine/threonine protein kinase of wheat (*Triticum aestivum* L.), confers enhanced multistress tolerance in *Arabidopsis* [J]. *Journal of Experimental Botany*, 2010, 61(3): 683-696.
- [11] BOUDSOCQ M, BARBIER-BRYGOO H, LAURIÈRE C. Identification of nine sucrose nonfermenting 1-related protein kinases 2 activated by hyperosmotic and saline stresses in *Arabidopsis thaliana* [J]. *The Journal of Biological Chemistry*, 2004, 279(40): 41758-41766.
- [12] FUJITA Y, NAKASHIMA K, YOSHIDA T, et al. Three SnRK2 protein kinases are the main positive regulators of abscisic acid signaling in response to water stress in *Arabidopsis* [J]. *Plant and Cell Physiology*, 2009, 50(12): 2123-2132.
- [13] KOBAYASHI Y, YAMAMOTO S, MINAMI H, et al. Differential activation of the rice sucrose nonfermenting1-related protein kinase2 family by hyperosmotic stress and abscisic acid [J]. *The Plant Cell*, 2004, 16: 1163-1177.
- [14] UR REHMAN S, WANG J Y, CHANG X P, et al. A wheat protein kinase gene *TaSnRK2.9-5A* associated with yield contributing traits [J]. *Theoretical and Applied Genetics*, 2019, 132(4): 907-919.
- [15] SHEN X J, GUO X, ZHAO D, et al. Cloning and expression profiling of the *PacSnRK2* and *PacPP2C* gene families during fruit development, ABA treatment, and dehydration stress in sweet cherry [J]. *Plant Physiology and Biochemistry*, 2017, 119: 275-285.
- [16] 彭方仁, 李永荣, 郝明灼, 等. 我国薄壳山核桃生产现状与产业化发展策略 [J]. *林业科技开发*, 2012, 26(4): 1-4.
- [17] REYES-VÁZQUEZ N D C, DE LA ROSA L A, MORALES-LANDA J L, et al. Phytochemical content and potential health applications of pecan [*Carya illinoensis* (Wangenh) K. Koch] nutshell [J]. *Current Topics in Medicinal Chemistry*, 2022, 22(2): 150-167.
- [18] 李小飞, 邓秋菊, 曹凡, 等. 激素浸种对薄壳山核桃种子萌发的影响及其种子吸胀过程中 SOD 和 POD 活性变化 [J]. *植物资源与环境学报*, 2016, 25(2): 111-113.
- [19] SAHA J, CHATTERJEE C, SENGUPTA A, et al. Genome-wide analysis and evolutionary study of sucrose non-fermenting 1-related protein kinase 2 (*SnRK2*) gene family members in *Arabidopsis* and *Oryza* [J]. *Computational Biology and Chemistry*, 2014, 49: 59-70.
- [20] FINN R D, COGGILL P, EBERHARDT R Y, et al. The Pfam protein families database: towards a more sustainable future [J]. *Nucleic Acids Research*, 2016, 44(D1): D279-D285.
- [21] YU C S, LIN C J, HWANG J K. Predicting subcellular localization of proteins for Gram-negative bacteria by support vector machines based on *n*-peptide compositions [J]. *Protein Science*, 2004, 13: 1402-1406.
- [22] KUMAR S, STECHER G, TAMURA K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets [J]. *Molecular Biology and Evolution*, 2016, 33(7): 1870-1874.
- [23] ZHANG H K, GAO S H, LERCHER M J, et al. EvolView, an online tool for visualizing, annotating and managing phylogenetic trees [J]. *Nucleic Acids Research*, 2012, 40(W1): W569-W572.
- [24] CHEN C J, CHEN H, ZHANG Y, et al. TBtools: an integrative toolkit developed for interactive analyses of big biological data [J]. *Molecular Plant*, 2020, 13(8): 1194-1202.
- [25] BAILEY T L, JOHNSON J, GRANT C E, et al. The MEME suite [J]. *Nucleic Acids Research*, 2015, 43(W1): W39-W49.
- [26] LESCOT M, DÉHAIS 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.
- [27] ZHU K K, FAN P H, MO Z H, et al. Identification, expression and co-expression analysis of R2R3-MYB family genes involved in graft union formation in pecan (*Carya illinoensis*) [J]. *Forests*, 2020, 11(9): 917.
- [28] ZHU K K, FAN P H, LIU H, et al. Insight into the *CBL* and *CIPK* gene families in pecan (*Carya illinoensis*): identification, evolution and expression patterns in drought response [J]. *BMC Plant Biology*, 2022, 22: 221.
- [29] LI B, DEWEY C N. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome [J]. *BMC Bioinformatics*, 2011, 12: 323.
- [30] LIVAK K J, SCHMITTGEN T D. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta Ct}$ method [J]. *Methods*, 2001, 25(4): 402-408.