

• 药材与资源 •

西洋参赤霉素 20-氧化酶基因的克隆与序列分析

孙同玉, 祝娟, 孙鹏, 廖登群, 李先恩, 祁建军*

中国医学科学院 北京协和医学院药用植物研究所, 北京 100193

摘要: 目的 对西洋参 *Panax quinquefolium* 种子休眠与萌发过程中赤霉素 20-氧化酶 (GA20ox) 基因编码区进行克隆及生物信息学分析。方法 根据西洋参种子转录组的注释信息, 得到 9 条 GA20ox 相关序列, 通过比对分析确定了 GA20ox 的转录本。采用 qRT-PCR 方法得到西洋参 GA20ox (PqGA20ox) 的 cDNA 全长, 并对其编码蛋白进行生物信息学分析。采用实时荧光定量 PCR 方法检测 PqGA20ox 基因在根、茎、叶和种子休眠不同时期的表达水平。结果 生物信息学分析表明, PqGA20ox 基因全长 1 092 bp, 编码 363 个氨基酸残基, PqGA20ox 编码蛋白不含跨膜区, 不含信号肽。qRT-PCR 实验结果显示 PqGA20ox 基因在形态休眠起始期和生理休眠起始期表达量高于其他时期。结论 首次获得 PqGA20ox 基因的编码区序列, 为进一步研究西洋参种子解除休眠的分子机制奠定基础。

关键词: 西洋参; 种子休眠与萌发; 赤霉素 20-氧化酶; 克隆; 实时荧光定量 PCR

中图分类号: R282.12 文献标志码: A 文章编号: 0253-2670(2015)11-1656-05

DOI: 10.7501/j.issn.0253-2670.2015.11.018

Cloning and sequence analysis of PqGA20ox gene in *Panax quinquefolium*

SUN Tong-yu, ZHU Juan, SUN Peng, LIAO Deng-qun, LI Xian-en, QI Jian-jun

Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing 100193, China

Abstract: Objective To clone and analyze the coding region of gibberellin 20-oxidase (GA20ox) gene of *Panax quinquefolium* in seed dormancy and germination process. **Methods** According to *P. quinquefolium* seeds transcriptome annotation, one transcript coding of GA20ox was obtained by comparative analysis of nine related unigenes. The full-length cDNA of PqGA20ox was determined by using RT-PCR method. Then the bioinformatic analysis of this gene and encoded protein was performed. The expression level of PqGA20ox in the roots, stems, leaves, and several seed dormancy periods was detected by real time fluorescent quantitative PCR (RT-qPCR). **Results** Bioinformatic analysis showed that PqGA20ox contained 1 092 bp and encoded 363 amino acids. PqGA20ox encoded protein had neither transmembrane nor signal peptide. The expression level of PqGA20ox was higher in the inset periods of morphological and physiological dormancy than that in the other periods based on RT-qPCR analysis. **Conclusion** The PqGA20ox gene is cloned for the first time, and will provide a foundation for the dormancy release molecular mechanism of *P. quinquefolium*.

Key words: *Panax quinquefolium* L.; seed dormancy and germination; GA20ox; cloning; real time fluorescent quantitative PCR

西洋参 *Panax quinquefolium* L. 为五加科 (Araliaceae) 人参属 *Panax* L. 多年生草本植物, 以根入药, 是名贵药材。西洋参通过种子繁殖, 但其种子具有深度休眠的特性, 属于形态生理型休眠 (morphophysiological dormancy, MPD), 即种子既没有发育完全的胚又存在萌发抑制物质, 使种子不能正常萌发^[1], 需要长时间的处理解除休眠后才可萌发。种子休眠现象导致不能适时播种或出苗率

低, 这已成为制约药用植物人工栽培发展的关键因素。种子休眠与萌发是高等植物长期适应环境因素形成的复杂的适应性特征^[2], 是由多种植物激素调节的, 如脱落酸 (ABA)、赤霉素 (GA)、生长素、乙烯、油菜素内酯等。GA 能够促进休眠解除和发芽, 减弱 ABA 的影响^[3]。在生理休眠的种子中, 温度和 GA 可以同时解除休眠和促进发芽^[4-5]。赤霉素 20-氧化酶 (GA20ox) 和赤霉素 3-氧化酶 (GA3ox)

收稿日期: 2015-02-06

作者简介: 孙同玉 (1989—), 女, 河北唐山人, 硕士研究生, 从事药用植物生物技术研究。E-mail: kao20yan12@163.com

*通信作者 祁建军, 男, 博士, 副研究员, 硕士生导师, 从事药用植物生物技术研究。E-mail: jjqi@implad.ac.cn

参与 GA 合成,而赤霉素 2-氧化酶 (GA2ox) 则参与 GA 的分解代谢^[6-8]。GA2ox 是一种多功能酶,不仅参与 GA 生物合成,还是 GA 生物合成过程中的关键限速酶,并控制植物体内有生物活性的 GA1 和 GA4 的合成数量^[9-10]。

本研究从前期高通量测序获得的西洋参种子 78 207 条 unigenes 序列注释信息中挖掘出 9 条 GA2ox 基因相关序列,通过 BlastP 比对和 ORF Finder 确定 GA2ox 的 unigenes 序列,设计引物,采用 qRT-PCR 方法扩增西洋参 GA2ox (PqGA2ox) 基因的 cDNA 全长。本研究对 PqGA2ox 的 cDNA 序列进行克隆及生物信息学分析,为阐明西洋参种子休眠解除与萌发过程中 GA 的作用奠定基础。

1 材料、试剂与方法

1.1 试验材料

样品取自于中国医学科学院北京协和医学院药用植物研究所试验田,经笔者鉴定为四年生西洋参 *Panax quinquefolium* L. 种子、根、茎、叶。对西洋参种子进行层积处理,在种子休眠的不同时期分别取样,用于检测 PqGA2ox 基因在不同组织中的表达差异。大肠杆菌 DH5 α 感受态细胞购自天根生物公司。

1.2 试剂

植物多糖多酚核糖核酸 (RNA) 提取试剂盒购自北京百泰克生物公司。PrimeScriptTM RT reagent

Kit with gDNA Eraser (Perfect Real Time) 反转录试剂盒、ex Taq 酶、pMD19-T vector、SYBR[®]Premix ex TaqTM (Tli RNase Plus) 试剂盒均购自日本宝生物公司。通用型 DNA 纯化回收试剂盒 (离心柱型) 购自天根生物公司。本研究所用引物均由生工生物工程 (上海) 股份有限公司北京分部合成。其他试剂均为国产分析纯。

1.3 方法

1.3.1 RNA 的提取和 cDNA 的合成 取西洋参不同部位 (根、茎、叶),不同休眠期的种子样品,按照 RNA 提取试剂盒操作步骤提取各样品总 RNA,利用 NanoDrop 2000 进行定量测定,1.2%的琼脂糖电泳检测其完整性。DNase I 去除提取物中所含的 DNA。按照反转录试剂盒操作说明将 RNA 反转录成第一链互补链 DNA (cDNA)。

1.3.2 PqGA2ox 基因的扩增及测序 通过分析西洋参种子转录组数据,得到 9 条与 GA2ox 相关的 unigene 基因,其中 unigene 61797 全长 1 289 bp 的转录本被注释为 GA2ox,详细信息见表 1。通过 NCBI 的 ORF Finder (<http://www.ncbi.nlm.nih.gov/gorf/gorf.html>) 分析,发现该转录本具有完整的开放阅读框 (open readingframe, ORF),其 ORF 为 1 092 bp。根据 ORF 序列设计引物 (F-5'-CTATCATTCAATGGCCTCATCT-3', R-5'-GCAGCAATTTGTGGGCTTAA-3')。引物由生工生物工程 (上海) 股份有限公司北京分部合成。

表 1 与 GA2ox 相关的 9 条 unigene 基因
Table 1 Nine unigenes related to GA2ox

基因	长度/bp	注释
unigene30205	411	gibberellin 20 oxidase 2 [<i>Solanum nigrum</i>] 2-oxoglutarate-dependent dioxygenase
unigene48771	155	gibberellin 20-oxidase-related protein [<i>Arabidopsis thaliana</i>]
unigene49112	282	gibberellin 20-oxidase [<i>Torenia fournieri</i>]
unigene55454	316	gibberellin 20-oxidase [<i>Populus trichocarpa</i>]
unigene60909	225	gibberellin 20-oxidase 2 [<i>Chrysanthemum x morifolium</i>]
unigene61797	1 289	gibberellin 20-oxidase [<i>Gossypium hirsutum</i>]
unigene67746	176	gibberellin 20-oxidase-like protein [<i>Arabidopsis thaliana</i>]
unigene7661	478	gibberellin 20-oxidase-like protein [<i>Selaginella moellendorffii</i>]
unigene26051	1 260	gibberellin 20-oxidase-like protein [<i>Selaginella moellendorffii</i>]

以西洋参生理休眠种子 cDNA 为模板,PCR 反应体系: cDNA 2.0 μ L, ex Taq 酶 12.5 μ L, 上下游引物各 1 μ L, 超纯水 8.5 μ L, 终体积 25 μ L。PCR 反应程序: 94 $^{\circ}$ C 预变性 3 min, 以下进行 30 个循环: 98 $^{\circ}$ C 变性 10 s, 51 $^{\circ}$ C 退火 30 s, 72 $^{\circ}$ C 延伸 90 s, 循

环结束后 72 $^{\circ}$ C 延伸 5 min。PCR 产物用 1.2%琼脂糖凝胶电泳检测,对目的片段进行切胶回收。将回收产物连接到 pMD19-T 载体,转化大肠杆菌 (DH5 α 菌株),在含有氨苄的 LB 平板上进行培养,挑取单克隆经菌落 PCR 扩增和电泳检测后送上海生工生

物工程公司测序。

1.3.3 PqGA20ox 编码蛋白生物信息学分析 对 PqGA20ox 编码蛋白利用 ExPASy ProtParam (<http://www.expasy.ch/tools/protparam.html>) 进行理化性质分析, 利用 InterProScan (<http://www.ebi.ac.uk/Tools/pfa/iprscan/>)、NCBI Conserved Domains (<http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>) 和 ExPASy Prosite (<http://prosite.expasy.org/>) 进行蛋白质结构域和功能位点分析, 利用 SOPMA (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_sopma.html) 进行蛋白质二级结构分析。Signal P4.0 Server (<http://www.cbs.dtu.dk/services/SignalP/>) 进行分泌蛋白分析、WOLF PSORT (<http://wolfsort.org/>) 进行亚细胞定位、TMHMM Server (<http://www.cbs.dtu.dk/services/TMHMM/>) 进行跨膜区分析, ExPASy ProScale (<http://web.expasy.org/protscale/>) 进行疏水性预测。利用 SWISS-MODEL (<http://web.expasy.org/protscale/>) 建立该蛋白的三维模型。序列比对利用美国国立生物技术信息中心 (NCBI) 的蛋白质序列数据库进行, 再通过 MEGA 6.0 构建 Neighbor-joining (NJ) 系统进化树。

1.3.4 PqGA20ox 基因在不同组织的特异性表达分析 为了检测 PqGA20ox 基因的表达情况, 以本实验室前期筛选的肌动蛋白 (Actin) 基因为内参基因 (Actin-F: 5'-AACCAGCCAGGTGCGACA-3', Actin-R: 5'-CGAAGGGCCGGACTCAT-3'), 利用 qRT-PCR 技术进行扩增。反应体系: SYBR[®]Premix ex Taq[™] 10 μ L, 正反向引物各 0.8 μ L (F: 5'-CCAACTCAAAGTCCTTAGTCAGC-3', R: 5'-CCTTCTTGTCACCGAAAGC-3'), cDNA 模板 1.5 μ L, 超纯水 6.9 μ L, 终体积 20 μ L。扩增程序: 95 $^{\circ}$ C、30 s, 以下 40 个循环, 95 $^{\circ}$ C 变性 10 s, 58 $^{\circ}$ C 退火 20 s, 72 $^{\circ}$ C 延伸 20 s, 循环结束后 60 $^{\circ}$ C $\rightarrow$ 95 $^{\circ}$ C、30 s。

2 结果

2.1 西洋参 PqGA20ox 基因克隆

以西洋参生理休眠期种子 RNA 反转录得到的 cDNA 为模板, 利用 PCR 技术得到 1 条全长 1 092 bp 的序列, 对所扩增基因命名为 PqGA20ox, 将该核酸序列提交到 NCBI, 获得 GenBank 登录号为 KJ740524, 图 1 为扩增产物电泳图。

2.2 PqGA20ox 基因编码蛋白特性分析

2.2.1 理化性质分析 利用在线软件 ExPASy ProtParam 预测 PqGA20ox 编码蛋白的理化性质。

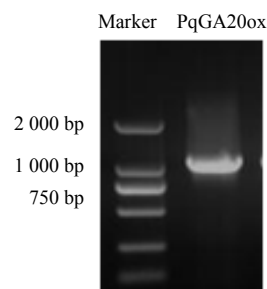


图 1 PqGA20ox 基因的电泳图

Fig. 1 Electrophoretogram of PqGA20ox gene

PqGA20ox 编码蛋白由 363 个氨基酸组成, 分子式 $C_{1857}H_{2855}N_{501}O_{542}S_{16}$, 总原子数 5 771。相对分子质量 41 384.0, 理论等电点 7.22, 带正电残基 (Asp+Glu) 44, 带负电残基 (Asp+Lys) 44。不稳定系数 34.51, 不稳定系数小于 40, 表示 PqGA20ox 编码蛋白稳定。脂肪系数 72.78, 总平均亲水性 -0.500, 负值表示亲水性, 正值表示疏水性, 总平均疏水性介于 -0.5~0.5 为两性氨基酸^[11], 所以 PqGA20ox 编码蛋白为两性氨基酸。

2.2.2 PqGA20ox 编码蛋白的二级结构分析及结构预测 采用 SOPMA 预测 PqGA20ox 编码蛋白的二级结构, 结果表明 PqGA20ox 编码蛋白二级结构中 α -螺旋 (alpha helix) 占 34.71%、延伸链 (extended strand) 占 17.91%、 β -折叠 (beta turn) 占 5.23%、无规则卷曲 (random coil) 占 42.15%。利用在线软件 InterProScan 和 Conserved Domains 进行结构域分析, 结果表明: 有 2 个比较保守的结构域, 分别是异青霉素 N 合成酶及非血红素双加氧酶 N 端的保守结构域和依赖于 2-酮戊二酸和 Fe^{2+} 依赖性双加氧酶保守结构域。以上结构域分别位于 15~352 位、42~144 位和 207~304 位氨基酸残基之间。利用 ExPASy 的 Prosite 分析 PqGA20ox 的功能位点, 结果表明, 保守功能域为 Fe^{2+} 结合域, 该保守域由第 229 位的 His, 231 位的 Asp 和 285 位的 His 残基组成 (图 2)。

2.2.3 PqGA20ox 编码蛋白三维结构建模 在 SWISS-MODEL 进行同源建模, PqGA20ox 编码蛋白三维结构模型见图 3。

2.2.4 PqGA20ox 编码蛋白亲疏水性分析 利用 ExPASy ProtScale 程序对西洋参 PqGA20ox 编码的蛋白做亲疏水性预测, 负值表示亲水性, 正值表示疏水性。结果显示该蛋白的疏水性最大值 2.511, 最小值 -2.556。从图 4 可以看出, PqGA20ox 编码蛋白的疏水性区域与亲水性区域交替出现, 没有典型



图 2 ExPASy 分析 PqGA20ox 编码蛋白的功能位点
Fig. 2 ExPASy analysis on functional site of PqGA20ox encoding protein

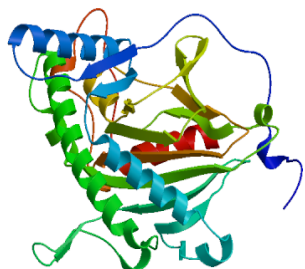


图 3 PqGA20ox 编码蛋白三维模型
Fig. 3 3D model for PqGA20ox encoding protein

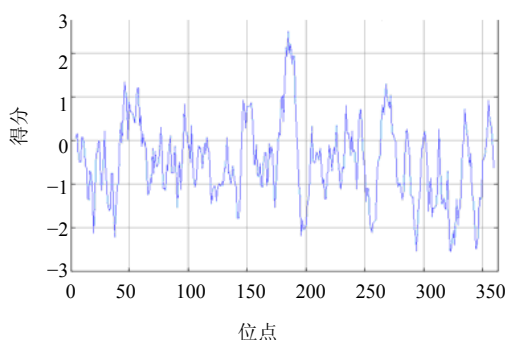


图 4 PqGA20ox 编码蛋白疏水性分析
Fig. 4 Hydrophobicity analysis of PqGA20ox encoding protein

的疏水性区域，属于两性氨基酸，与 Protparam 预测结果一致。

2.2.5 PqGA20ox 编码蛋白跨膜区、信号肽、亚细胞定位预测分析 利用 TMHMM 预测 PqGA20ox 编码蛋白的跨膜区，结果显示该蛋白不具有跨膜区，由此推断该蛋白不经过跨膜转运。信号肽是分析蛋白 N 端的一段氨基酸序列，是引导分泌性蛋白在糙面内质网上合成的最重要因素。利用 SignalP4.1 Server 预测 PqGA20ox 编码蛋白不具有信号肽，该蛋白不可能在细胞中发生移动，与跨膜区分析相吻合。利用在线工具 WOLF PSORT 预测该蛋白的亚细胞定位情况：细胞核的定位系数为 9.0 (nucl 9.0)，细胞质的定位系数为 3.0 (cyto 3.0)，叶绿体的定位系数为 1.0 (chlo 1.0)。因此，该蛋白最可能定位于细胞核。

2.3 PqGA20ox 编码蛋白序列比对及进化分析

将 PqGA20ox 的编码蛋白序列提交到 NCBI 上进行 Blast P 比对，搜索相似性序列，结果显示，PqGA20ox 与牡丹 *Paeonia suffruticosa* Andr. (ABQ52488.1) 的 GA20ox 序列相似性最高为 68%。

为了分析 PqGA20ox 的进化情况，从 NCBI 的 Nr 数据库中选取 11 个物种 11 条 GA20ox 序列，经过 Clustal X 比对并利用 MEGA6.0 中的 NJ 法，构建进化树。这些序列包括黑麦草 *Lolium perenne* L. (AAG43044.1)、小麦 *Triticum aestivum* Linn. (CAA74331.1)、玉米 *Zea mays* L. (ABG67710.1)、拟南芥 *Arabidopsis thaliana* (L.) Heynh. (CAB45519.1)、牡丹 (ABQ52488.1)、黄瓜 *Cucumis sativus* Linn. (CCD28480.1)、葡萄 *Vitis vinifera* Linn. (ABF61825.1)、茶树 *Camellia sinensis* L. (AGV40767.1)、向日葵 *Helianthus annuus* L. (CAQ43616.1)、笋瓜 *Cucurbita maxima* Duch. (CAC83626.1)、豌豆 *Pisum sativum* L. (CAA62846.1)。利用上述植物的 GA20ox 序列绘制系统进化树 (图 5)，PqGA20ox 与牡丹 (ABQ52488.1)、黄瓜 (CCD28480.1) 的 GA20ox 序列亲缘关系较近，这与序列 Blast P 分析结果一致。

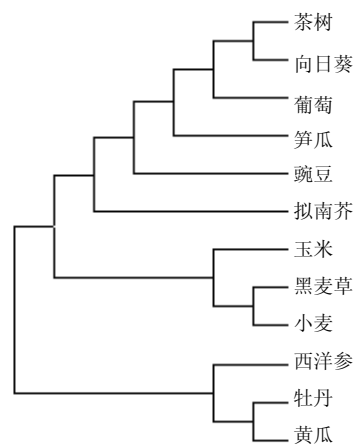


图 5 PqGA20ox 的进化分析
Fig. 5 Phylogenetic analysis of PqGA20ox

2.4 PqGA20ox 基因的组织特异性表达分析

利用 qRT-PCR 方法对 PqGA20ox 进行基因组织表达分析，材料包括根 (Ro)、茎 (St)、叶 (Le)、浸水种子即形态休眠起始期的种子 (S1)、形态休眠中期的种子 (S2)、裂口种子即解除形态休眠种子 (S3)、低温处理 24 h 种子即生理休眠起始期种

子 (S4)、生理休眠中期种子 (S5)、解除休眠的胚乳 (En) 和胚 (Em)。结果表明, PqGA2ox 在 S1 表达量高于 S2, S4 高于 S5, 而且在 S4 表达量最高, 且明显高于 S3 (图 6)。

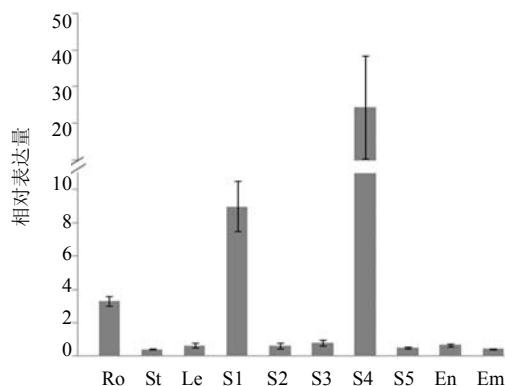


图 6 荧光定量检测 PqGA2ox 基因的表达

Fig. 6 Expression analysis of PqGA2ox using RT-qPCR

3 讨论

西洋参种子休眠属于 MPD 型休眠, 解除休眠的过程十分漫长。本研究针对西洋参种子休眠的研究现状, 首次克隆出西洋参种子休眠解除过程中 GA 的分解代谢关键基因 PqGA2ox, 并获得 GenBank 登陆号 KJ740524。通过生物信息学分析了 PqGA2ox 编码蛋白的二级结构、三级结构等基本特性, 既无信号肽, 又无跨膜结构域。GA2ox 基因编码氨基酸序列在不同的物种中一致性为 50%~60%^[12], PqGA2ox 基因编码氨基酸 Blast P 比对结果显示与牡丹 (ABQ52488.1) 的 GA2ox 序列相似性为 68%。通过 qRT-PCR 方法检测西洋参根、茎、叶和种子休眠解除过程中各时期的表达量, 发现 PqGA2ox 基因在形态休眠起始期 S1 表达量高于形态休眠其他时期, 生理休眠起始期 S4 高于生理休眠其他时期, 推测 PqGA2ox 基因在初期表达量高, GAs 在初期发挥作用。S4 表达量最高, 且是 S1 期的 3 倍多, 与 Ogawa 等^[13]的研究结果相同, 低温层积处理能够诱导 GA3ox、GA2ox 基因的表达, 促进 GA 生物合成, 同时加剧了 ABA 的分解代谢^[14], 打破种子休眠, 促进萌发。

本研究结果不仅为分析 PqGA2ox 基因的分子功能提供依据, 而且为西洋参种子休眠与萌发机制的研究提供基础。GA2ox 基因是 GAs 中的一员, 目前已发现一百多种 GA 分子, 但仅有少数具有生物学活性^[15], 为进一步研究 GAs 在种子休眠

与萌发中的作用, 在后续研究中需要克隆西洋参 GAs 基因家族的其他成员, 分析各成员间表达是否存在规律。

参考文献

- [1] Baskin J M, Baskin C C. A classification system for seed dormancy [J]. *Seed Sci Res*, 2004, 14(1): 1-16.
- [2] Koornneef M, Bentsink L, Hilhorst H. Seed dormancy and germination [J]. *Curr Opin Plant Biol*, 2002, 5(1): 33-36.
- [3] Birgit K, Marc A C, Gerhard L M. Plant hormone interactions during seed dormancy release and germination [J]. *Seed Sci Res*, 2005, 15(4): 281-307.
- [4] Baskin C C. Breaking physical dormancy in seeds-focussing on the lens [J]. *New Phytol*, 2003, 158(2): 227-238.
- [5] Kucerna B, Cohn M A, Leubner-Metzger G. Plant hormone interactions during seed dormancy release and germination [J]. *Seed Sci Res*, 2005, 15(4): 281-307.
- [6] Sakamoto T, Kobayashi M, Itoh H, et al. Expression of a gibberellin 2-oxidase gene around the shoot apex is related to phase transition in rice [J]. *Plant Physiol*, 2001, 125(3): 1508-1516.
- [7] Yamaguchi S, Annu R. Gibberellin metabolism and its regulation [J]. *Plant Biol*, 2008, 59: 225-251.
- [8] Chen Q F, Ya H Y, Li S M, et al. Isolation and analysis of homoeologous genes encoding gibberellins 2-oxidase 3 isozymes in common wheat [J]. *J Genet*, 2012, 91(3): 1-9.
- [9] Sakamoto T, Miura K, Itoh H, et al. An overview of gibberellin metabolism enzyme genes and their related mutants in rice [J]. *Plant Physiol*, 2004, 134(4): 1642-1653.
- [10] Susan E M, Rosemary L. Genetic regulation of seed dormancy in *Purshiat ridentata* (Rosaceae) [J]. *Ann Bot*, 2000, 85(4): 521-529.
- [11] 付海辉, 辛培尧, 许玉兰, 等. 几种经济植物 UFGT 基因的生物信息学分析 [J]. *基因组学与应用生物学*, 2010, 30(1): 92-102.
- [12] 石海燕, 张玉星. 高等植物赤霉素 20-氧化酶基因的克隆、表达及其调控技术的研究进展 [J]. *贵州农业科学*, 2011, 39(4): 19-21.
- [13] Ogawa M, Hanada A, Yamauchi Y, et al. Gibberellin biosynthesis and response during *Arabidopsis* seed germination [J]. *Plant Cell*, 2003, 15(7): 1591-1604.
- [14] Kermode A R. Role of abscisic acid in seed dormancy [J]. *J Plant Growth Regul*, 2005, 24(4): 319-344.
- [15] Gao X H, Xiao S L, Yao Q F, et al. An updated GA signaling 'relief of repression' regulatory model [J]. *Mol Plant*, 2011, 4(4): 601-606.