



# 基于 RNA-Seq 分析 *RasI<sup>CA</sup>* 在家蚕后部丝腺过表达 对细胞周期通路基因的影响\*

马倩<sup>1,2\*\*</sup> 马俐<sup>2</sup> 李胜<sup>2</sup> 李恺<sup>1\*\*\*</sup>

(1. 华东师范大学生命科学学院, 上海 200241; 2. 中国科学院上海生命科学研究院植物生理生态研究所, 上海 200032)

**摘要** 【目的】 本文旨在挖掘野生型家蚕 *Bombyx mori* 和 *RasI<sup>CA</sup>* 过表达转基因家蚕后部丝腺中的转录本差异, 分析和验证细胞周期通路中的差异表达基因, 从而探讨 *RasI<sup>CA</sup>* 过表达转基因家蚕提高蚕丝产量的分子机制。【方法】 利用转录组学比较野生型和 *RasI<sup>CA</sup>* 过表达转基因家蚕后部丝腺中的基因转录本差异, 经实时荧光定量 PCR (qRT-PCR) 验证。【结果】 野生型家蚕和 *RasI<sup>CA</sup>* 过表达转基因家蚕后部丝腺组织中有 2 636 个差异表达基因, 其中细胞周期信号通路中有 42 个差异表达基因, 包括细胞周期依赖性激酶 (CDK)、细胞周期素 (Cyclin) 以及转录因子等。通过 qRT-PCR 检测 *cdk1*、*cyclinD1*、*cyclinD2*、*cdc7*、*cdh1*、*dp-1,2* 等 6 个基因在野生型和 *RasI<sup>CA</sup>* 过表达转基因家蚕后部丝腺组织中的相对表达量, 发现转基因家蚕中的表达量均显著高于野生型 ( $P < 0.05$ ), 其中 *cyclinD2* 的表达差异极显著 ( $P < 0.01$ )。qRT-PCR 结果与转录本差异一致, 表明 *RasI<sup>CA</sup>* 过表达后, 能够促进细胞周期通路基因的表达。【结论】 野生型家蚕和 *RasI<sup>CA</sup>* 过表达转基因家蚕后部丝腺中有大量差异表达基因, 且 *RasI<sup>CA</sup>* 能够在转录水平上调细胞周期通路, 影响后部丝腺组织的细胞分裂和器官生长, 从而促进蚕丝蛋白的合成。

**关键词** 转录组学, 转基因家蚕, *RasI<sup>CA</sup>* 过表达, 细胞周期基因

## RNA-Seq technology based transcriptomic analysis of differentially expressed genes in the cell cycle pathway of *RasI<sup>CA</sup>*-overexpressed and wild type posterior silk glands of *Bombyx mori*

MA Qian<sup>1,2\*\*</sup> MA Li<sup>2</sup> LI Sheng<sup>2</sup> LI Kai<sup>1\*\*\*</sup>

(1. School of Life Science, East China of Normal University, Shanghai 200241, China; 2. Institute of Plant Physiology and Ecology, Shanghai Institutes for Biological Science, Chinese Academy of Sciences, Shanghai 200032, China)

**Abstract** [Objectives] To further understand the molecular mechanisms underlying the overexpression of the *RasI<sup>CA</sup>* oncogene in the posterior silk gland (PSG) of *Bombyx mori* and how this improves silk yield at the transcriptional level. Transcriptomic data from differentially expressed genes (DEGs) in the cell cycle pathway of *RasI<sup>CA</sup>*-overexpressed and wild type PSG were analyzed and verified by quantitative real-time PCR (qRT-PCR). [Methods] Deep RNA sequencing of silkworm PSG was carried out and KEGG pathway enrichment analysis was conducted on the DEGs. The DEGs in the cell cycle pathway was selected to be verified by qRT-PCR. [Results] RNA-Seq revealed 2 636 DEGs in *RasI<sup>CA</sup>*-overexpressed PSG compared to wild type PSG. There were 42 DEGs distributed in the “cell cycle” pathway, including cyclin dependent kinases (CDKs), cyclins and transcription factors. Consistent with the transcriptomic data, qRT-PCR verification of the selected genes; *cdk1*, *cyclinD1*, *cyclinD2*, *cdc7*, *cdh1*, *dp-1,2*, indicated that all of these were significantly upregulated by *RasI<sup>CA</sup>*. [Conclusion] Transcriptomic analysis revealed that there are a number of DEGs in *RasI<sup>CA</sup>*-overexpressed and wild

\* 资助项目 Supported projects: 国家自然科学基金 (31201747); 上海市自然科学基金 (12ZR1452500); 上海生科院优秀青年人才领域前沿项目 (2012KIP305)

\*\*第一作者 First author, E-mail: maqian0927@126.com

\*\*\*通讯作者 Corresponding author, E-mail: kaili@bio.ecnu.edu.cn

收稿日期 Received: 2014-04-29, 接受日期 Accepted: 2014-05-23

type PSG. *RasI<sup>CA</sup>*-overexpression upregulates some DEGs in the cell cycle pathway that benefit silk gland growth and fibroin synthesis. These results advance our knowledge of the molecular mechanism underlying how *RasI<sup>CA</sup>*-overexpression in the PSG improves fibroin production and silk production.

**Key words** transcriptomics, transgenic silkworm, *RasI<sup>CA</sup>* overexpression, cell cycle genes

家蚕 *Bombyx mori* 是一种重要的经济昆虫, 20 世纪初, 其杂交育种技术促进了蚕丝业的发展; 但由于该技术存在缺陷, 近半个世纪来养蚕业已进入瓶颈期 (Talebi and Subramanya, 2009)。近年来, 包括转基因技术在内的分子育种新技术取得突破, 尤其是 *PiggyBac* 转座子介导的转基因家蚕技术于 2000 年建立后 (Tamura *et al.*, 2000), *GAL4/UAS* 技术的应用使目的基因能在特定时空进行表达 (Imamura *et al.*, 2003)。我们前期利用这一技术在家蚕后部丝腺特异性过表达 *RasI<sup>CA</sup>*, 提高了蚕丝产量 (Ma *et al.*, 2011)。

*RasI<sup>CA</sup>* 在家蚕后部丝腺过表达, 通过细胞体积的增大和核内复制的增加, 促进了后部丝腺组织的生长和蚕丝蛋白的合成 (Ma *et al.*, 2011)。家蚕后部丝腺细胞为核内复制方式, 只具有 DNA 扩增而不具有细胞分裂: 由连续的 S 期和 G 期组成, 而不存在有丝分裂过程中的 M 期 (Zimmet and Ravid, 2000)。5 龄期家蚕后部丝腺细胞核内 DNA 不断复制, 直至游走时期退出核内复制过程 (Zhang *et al.*, 2012), 这一阶段 DNA 复制的增加可使丝蛋白基因表达增强, 从而有利于蚕丝蛋白大量合成。细胞周期蛋白 (Cyclin) 和细胞周期依赖性蛋白激酶 (Cyclin dependent kinase, CDK) 协同作用, 调控细胞周期有序发生 (Pines, 1995)。在哺乳动物中研究发现, 原癌基因 *Ras* 能通过 Raf-MAPK 和 PI3K-AKT 等下游通路促使细胞周期通路基因表达和促进细胞分裂, 从而影响细胞的生命活动 (Orton *et al.*, 2005)。

近年来, 基因组、转录组和蛋白质组等各种功能基因组学研究的不断发展, 在昆虫学研究中的应用价值越来越重要 (张赞等, 2012)。这对阐释昆虫, 尤其是模式昆虫的生长代谢过程中的生理规律和分子机制起了重要的促进作用。其中, 转录组学旨在揭示生物功能基因的转录调控规

律 (Wang *et al.*, 2009)。我们通过对野生型家蚕和 *RasI<sup>CA</sup>* 过表达转基因家蚕进行转录组学分析, 发现了大量差异表达的功能基因 (Ma *et al.*, 2014), 有助于从转录水平上解释 *RasI* 过表达如何通过 Raf-MAPK 和 PI3K-AKT 通路来转录诱导下游基因的表达。

本研究以转基因家蚕作为实验材料, 进行转录组学分析, 比较野生型家蚕和 *RasI<sup>CA</sup>* 转基因家蚕后部丝腺的转录本差异, 进一步研究转基因家蚕提高蚕丝产量的分子调控机制。运用实时荧光定量 PCR (qRT-PCR) 方法对细胞周期通路差异基因进行分析, 确定 *RasI<sup>CA</sup>* 转基因家蚕中 *RasI<sup>CA</sup>* 对细胞周期通路基因表达的影响, 为今后综合杂交育种与转基因育种技术进行家蚕性状改良提供更可靠的理论基础和更佳候选基因。

## 1 材料与方法

### 1.1 材料和主要试剂

实验所用家蚕品种为大造 (中国农业科学院镇江蚕业研究所供种), 置于 25℃、80% 相对湿度环境中饲养, 光照时间为 14 h, 新鲜桑叶饲喂。转基因家蚕品系为 *Fil-GAL4/UAS-RasI<sup>CA</sup>*, 和野生型家蚕以同样条件饲养于实验室。

TRIzol Reagent 购自 Invitrogen 公司, M-MLV 反转录酶和 RRI 酶购自 TaKaRa 公司, 实时荧光定量 PCR 试剂 SYBR-Green Realtime PCR Master Mix 为日本 TOYOBO 公司产品, 引物合成由赛百盛公司完成, 其他化学试剂均为国产分析纯。

### 1.2 家蚕后部丝腺转录组数据分析

测序文库的建立参照 Illumina 公司试剂盒, 取野生型家蚕和转基因家蚕提取的 RNA 各 20 µg, 用带有 Oligo (dT) 的磁珠富集, 94℃ 反应 5 min 将其打断为 100~400 bp 的片段, 用六碱基

随机引物合成双链 cDNA, 对其进行末端修复及 3' 末端加 A 碱基后, 采用 paired-end 双向测序技术在 Illumina Hiseq 2000 测序仪上对文库进行 Solexa 测序。

### 1.3 数据分析

转录组数据分析, 对原始序列进行处理, 去除测序过程中的接头序列、低质量和不确定的序列。运用 TopHat (<http://tophat.cbcb.umd.edu/>) 和 Cufflinks (<http://cufflinks.cbcb.umd.edu/>) 软件对获得的高质量序列进行拼接和组装, 与家蚕基因组数据库和家蚕基因库进行分析比对, 通过 FRKM (Fragments per kilobase per million reads) 标准化公式计算得出基因表达丰度 (Unigene abundance)。对 unigene 进行 KEGG (Kyoto encyclopedia of genes and genomes) 注释和 GO (Geneontology) 注释。

实验数据运用 Microsoft Office Excel 2010 和统计分析软件 SPSS 17.0 进行统计分析。

### 1.4 家蚕后部丝腺 RNA 提取

饲喂野生型家蚕和 *RasI<sup>CA</sup>* 转基因家蚕至游走早期, 解剖取后部丝腺收集于 EP 管中, 迅速置于液氮中冷冻。采用 Trizol 提取总 RNA, 取适量材料于 1.5 mL 离心管中, 加入 1 mL Trizol, 用振荡器震荡破碎, 室温放置 5 min, 加入 200  $\mu$ L 体积的三氯甲烷, 充分混匀后静置 5 min, 4 条件下 12 000 r/min 离心 15 min; 小心吸取上清, 加入等体积异丙醇, 室温放置 10 min, 4 条件

下 12 000 r/min 离心 10 min, 弃上清, 用 1 mL 75% 乙醇洗涤 RNA 沉淀, 4 条件下 12 000 r/min 离心 5 min; 弃上清, 待 RNA 晾干后溶于适当体积 (1~4  $\mu$ g/ $\mu$ L) 的 DEPC 水中, -80 保存。取 RNA 做适当稀释, 测定  $A_{260}$  和  $A_{260}/A_{280}$  检测浓度和纯度。

### 1.5 cDNA 合成

选用 TaKaRa 公司的 M-MLV 逆转录酶, 进行第一链 cDNA 模板合成, 按照说明书进行操作, 反转录获得的 cDNA 模板稀释 10 倍后用于实时荧光定量 PCR。

### 1.6 实时荧光定量 PCR (Quantitative real-time PCR)

根据基因序列在 GenScript 网站设计 200 bp 左右的特异性引物 (表 1)。以反转录稀释后的 cDNA 为模板, 于 BioRad IQ5 PCR 仪 (USA) 进行操作。反应体系: Template 1.0  $\mu$ L (10 ng)、Realtime Master Mix 10  $\mu$ L、Primer1 (10  $\mu$ mol/L) 1  $\mu$ L、Primer2 (10  $\mu$ mol/L) 1  $\mu$ L、ddH<sub>2</sub>O 补充至 20  $\mu$ L。条件为: 95 预变性 1 min, 95 变性 15 s, 55 退火 10 s, 72 延伸 20 s, 40 个循环。以家蚕 *rp49* 基因作为内标参考, 实验重复 3 次。

## 2 结果与分析

### 2.1 转录组数据的 KEGG 通路分析

本研究采用 Illumina 平台的 paired-end 双向

表 1 实时荧光定量 PCR 检测基因和内参照基因的引物

Table 1 Primer pairs for qRT-PCR to detect transcriptions of target and internal reference genes

| 基因名称<br>Target genes | 引物序列<br>Primer sequences                                    |
|----------------------|-------------------------------------------------------------|
| <i>rp49</i>          | F:5'-CAGGCGGTTCAAGGGTCAATAC-3'; R:5'-TGCTGGGCTCTTCCACGA-3'  |
| <i>cdk1</i>          | F:5'-CAGAATGCTAAGAACGCCAA-3'; R:5'-TCAGTCCGGGAGGTAGTTTC-3'  |
| <i>cycD1</i>         | F:5'-CCGTAGATTTGCTGTGTGCT-3'; R:5'-AAGTTAAGGTGAGGGCGTGT-3'  |
| <i>cycD2</i>         | F:5'-TACGTGACCGTTTCGGATAAA-3'; R:5'-CTGCACAGTGCCGAAGTAGT-3' |
| <i>cdc7</i>          | F:5'-GCACTCGTTCGGTGTGCATAC-3'; R:5'-TGAGAGGTCAAGTGCGAGTC-3' |
| <i>cdh1</i>          | F:5'-AAGCTCTTCGTTTGGTCGAT-3'; R:5'-GGTTCTGTGAGTAGCCGTGA-3'  |
| <i>dp-1,2</i>        | F:5'-AATCAGGTGGCTGGGTCTAC-3'; R:5'-TTCCGCTTCCTTATTTCGTT-3'  |

测序技术对文库进行 Solexa 高通量测序。分析转录组数据发现, *RasI<sup>CA</sup>* 转基因和野生型家蚕分别表达了 6 962 个和 6 429 个基因, 其中有 2 636 个为差异表达基因 (Differentially expressed genes, DEGs)。利用 KEGG (Kyoto encyclopedia of genes and genomes) 数据库进行 Pathway 注释, 发现 *RasI<sup>CA</sup>* 转基因和野生型家蚕的 2 636 个差异表达基因存在于上百个信号通路中。表 2 列出了含有 20 个以上的差异表达基因的注释通路, 共计 51 个信号通路[含有 25 个及以上差异表达基因的 25 个信号通路已在 Ma 等 (2014) 中列出]。其中细胞周期是第 6 个差异表达基因

表 2 包含差异表达基因的 KEGG 通路列表  
Table 2 List of KEGG pathways involves mapped DEGs

| 信号通路<br>Pathway                                   | 类型<br>Pathway type                                              | 差异表达基因<br>DEGs |
|---------------------------------------------------|-----------------------------------------------------------------|----------------|
| ko00230, Purine metabolism                        | Metabolism/Nucleotide metabolism                                | 67             |
| ko03040, Spliceosome                              | Genetic information processing/Transcription                    | 52             |
| ko05200, Pathways in cancer                       | Human diseases/Cancers                                          | 49             |
| ko03013, RNA transport                            | Genetic information processing/Translation                      | 47             |
| ko05166, HTLV-I infection                         | Human diseases/Infectious diseases: Viral                       | 44             |
| ko00240, Pyrimidine metabolism                    | Metabolism/Nucleotide metabolism                                | 42             |
| <b>ko04110, Cell cycle</b>                        | <b>Cellular processes/Cell growth and death</b>                 | <b>42</b>      |
| ko04510, Focal adhesion                           | Cellular processes/Cell communication                           | 38             |
| ko04910, Insulin signaling pathway                | Organismal systems/Endocrine system                             | 35             |
| ko03008, Ribosome biogenesis in eukaryotes        | Genetic information processing/Translation                      | 35             |
| ko05016, Huntington's disease                     | Human diseases/Neurodegenerative diseases                       | 34             |
| ko04146, Peroxisome                               | Cellular processes/Transport and catabolism                     | 34             |
| ko00561, Glycerolipid metabolism                  | Metabolism/Lipid metabolism                                     | 32             |
| ko05010, Alzheimer's disease                      | Human diseases/Neurodegenerative diseases                       | 32             |
| ko04010, MAPK signaling pathway                   | Environmental information processing/Signal transduction        | 32             |
| ko04111, Cell cycle-yeast                         | Cellular processes/Cell growth and death                        | 32             |
| ko04144, Endocytosis                              | Cellular processes/Transport and catabolism                     | 31             |
| ko04810, Regulation of actin cytoskeleton         | Cellular processes/Cell motility                                | 30             |
| ko00564, Glycerophospholipid metabolism           | Metabolism/Lipid metabolism                                     | 29             |
| ko05169, Epstein-Barr virus infection             | Human diseases/Infectious diseases: Viral                       | 29             |
| ko00260, Glycine, serine and threonine metabolism | Metabolism/Amino acid metabolism                                | 28             |
| ko04120, Ubiquitin mediated proteolysis           | Genetic information processing/Folding, sorting and degradation | 28             |
| ko04113, Meiosis – yeast                          | Cellular processes/Cell growth and death                        | 28             |
| ko03010, Ribosome                                 | Genetic information processing/Translation                      | 25             |
| ko04145, Phagosome                                | Cellular processes/Transport and catabolism                     | 25             |

续表 2 (Table 2 continued)

| 信号通路<br>Pathway                                       | 类型<br>Pathway type                                              | 差异表达基因<br>DEGs |
|-------------------------------------------------------|-----------------------------------------------------------------|----------------|
| ko04114, Oocyte meiosis                               | Cellular processes/Cell growth and death                        | 25             |
| ko04270, Vascular smooth muscle contraction           | Organismal systems/Circulatory system                           | 25             |
| ko04972, Pancreatic secretion                         | Organismal systems/Digestive system                             | 25             |
| ko00040, Pentose and glucuronate interconversions     | Metabolism/Carbohydrate metabolism                              | 25             |
| ko04142, Lysosome                                     | Cellular processes/Transport and catabolism                     | 23             |
| ko05164, Influenza A                                  | Human diseases/Infectious diseases: Viral                       | 23             |
| ko00980, Metabolism of xenobiotics by cytochrome P450 | Metabolism/Xenobiotics biodegradation and metabolism            | 23             |
| ko00190, Oxidative phosphorylation                    | Metabolism/Energy metabolism                                    | 23             |
| ko04914, Progesterone-mediated oocyte maturation      | Organismal systems/Endocrine system                             | 23             |
| ko05145, Toxoplasmosis                                | Human diseases/Infectious diseases: Parasitic                   | 23             |
| ko00620, Pyruvate metabolism                          | Metabolism/Carbohydrate metabolism                              | 22             |
| ko04141, Protein processing in endoplasmic reticulum  | Genetic information processing/Folding, sorting and degradation | 22             |
| ko04540, Gap junction                                 | Cellular processes/Cell communication                           | 22             |
| ko05146, Amoebiasis                                   | Human diseases/Infectious diseases: Parasitic                   | 22             |
| ko05202, Transcriptional misregulation in cancer      | Human diseases/Cancers                                          | 22             |
| ko04530, Tight junction                               | Cellular processes/Cell communication                           | 22             |
| ko03018, RNA degradation                              | Genetic information processing/Folding, sorting and degradation | 21             |
| ko03015, mRNA surveillance pathway                    | Genetic information processing/Translation                      | 21             |
| ko05222, Small cell lung cancer                       | Human diseases/Cancers                                          | 21             |
| ko04062, Chemokine signaling pathway                  | Organismal systems/Immune system                                | 21             |
| ko00860, Porphyrin and chlorophyll metabolism         | Metabolism/Metabolism of cofactors and vitamins                 | 21             |
| ko04724, Glutamatergic synapse                        | Organismal systems/Nervous system                               | 20             |
| ko05168, Herpes simplex infection                     | Human diseases/Infectious diseases: Viral                       | 20             |
| ko00071, Fatty acid metabolism                        | Metabolism/Lipid metabolism                                     | 20             |
| ko00500, Starch and sucrose metabolism                | Metabolism/Carbohydrate metabolism                              | 20             |
| ko04310, Wnt signaling pathway                        | Environmental information processing/Signal transduction        | 20             |

最多 (共 42 个差异表达基因) 的通路。

## 2.2 细胞周期通路分析

细胞周期信号通路 (Cell cycle [PATH : ko04110]) 含有 42 个 DEGs, 包括参与胰岛素信号通路的 CDK4/6、CDK2 等重要的细胞周期激

酶, 还有其他参与细胞周期调控的因子, 主要包括细胞周期蛋白 (Cyclin)、细胞周期依赖性蛋白激酶 (Cyclin dependent kinase, CDK)、细胞分裂周期基因 (Cell division cycle, Cdc) 和转录因子等。表 3 为细胞周期信号通路中的差异表达

表 3 细胞周期信号通路的差异表达基因  
Table 3 The differentially expressed genes in cell cycle pathway

| 基因 BGI 号<br>DEGs BGI No. | 名称<br>Map name | 注释<br>Definition                                                                                     |
|--------------------------|----------------|------------------------------------------------------------------------------------------------------|
| BGIBMGA010110-TA         | Smad4          | Mad 同源物<br>Mothers against decapentaplegic homolog 4                                                 |
| BGIBMGA001661-TA         | Miz1           | 锌指结构<br>Zinc finger and BTB domain-containing protein 17                                             |
| BGIBMGA008275-TA         | APC/C          | 后期促进复合物亚基 1<br>Anaphase-promoting complex subunit 1                                                  |
| <b>BGIBMGA013134-TA</b>  | <b>CDK1</b>    | <b>周期蛋白依赖性蛋白 1</b><br><b>Cyclin-dependent kinase 1</b>                                               |
| BGIBMGA005147-TA         | APC/C          | 后期促进复合物亚基 2<br>Anaphase-promoting complex subunit 2                                                  |
| BGIBMGA002070-TA         | Abl            | 原癌基因酪氨酸激酶<br>Proto-oncogene tyrosine-protein kinase ABL1                                             |
| BGIBMGA006899-TA         | E2F4,5         | 转录因子<br>E2F transcription factor 4/5                                                                 |
| BGIBMGA012496-TA         | ORC2           | 真核细胞 DNA 起始识别复合物亚基 2<br>Origin recognition complex subunit 2                                         |
| BGIBMGA001887-TA         | Mad1           | 有丝分裂纺锤体组装检验点蛋白 1<br>Mitotic spindle assembly checkpoint protein MAD1                                 |
| BGIBMGA008277-TA         | APC/C          | 后期促进复合物亚基 1<br>Anaphase-promoting complex subunit 1                                                  |
| BGIBMGA005370-TA         | Chk1,2         | 丝氨酸/苏氨酸激酶<br>Serine/threonine-protein kinase Chk2                                                    |
| BGIBMGA003747-TA         | CycB           | 细胞周期素 B<br>Cyclin B                                                                                  |
| BGIBMGA012734-TA         | Mad2           | 有丝分裂纺锤体组装检验点蛋白 2<br>Mitotic spindle assembly checkpoint protein MAD2                                 |
| BGIBMGA001344-TA         | CDK4,6         | 周期蛋白依赖性蛋白 4<br>Cyclin-dependent kinase 4                                                             |
| BGIBMGA011989-TA         | Miz1           | 锌指结构<br>Zinc finger and BTB domain-containing protein 17                                             |
| BGIBMGA002762-TA         | MCM4           | 微小染色体维持蛋白 4 (细胞分裂控制蛋白 54)<br>Minichromosome maintenance protein 4 (cell division control protein 54) |
| BGIBMGA007036-TA         | ATM/ATR        | 丝氨酸/苏氨酸激酶<br>Serine/threonine-protein kinase ATR                                                     |
| <b>BGIBMGA006232-TA</b>  | <b>DP-1,2</b>  | <b>转录因子</b><br><b>Transcription factor Dp-1</b>                                                      |
| BGIBMGA006210-TA         | APC/C          | 后期促进复合物亚基 5<br>Anaphase-promoting complex subunit 5                                                  |
| BGIBMGA013453-TA         | CDK2           | 周期蛋白依赖性蛋白 2<br>Cyclin-dependent kinase 2                                                             |
| BGIBMGA008625-TA         | MCM2           | 微小染色体维持蛋白 2<br>Minichromosome maintenance protein 2                                                  |
| BGIBMGA013084-TA         | Cdc45          | 细胞分裂控制蛋白 45<br>Cell division control protein 45                                                      |
| BGIBMGA005856-TA         | Chk1,2         | 丝氨酸/苏氨酸激酶<br>Serine/threonine-protein kinase Chk1                                                    |

续表 3 (Table 3 continued)

| 基因 BGI 号<br>DEGs BGI No. | 名称<br>Map name | 注释<br>Definition                                                                                       |
|--------------------------|----------------|--------------------------------------------------------------------------------------------------------|
| BGIBMGA001250-TA         | CycD1          | <b>细胞周期素 D</b><br><b>Cyclin D</b>                                                                      |
| BGIBMGA009588-TA         | DNA-PK         | DNA 依赖蛋白激酶催化亚单位<br>DNA-dependent protein kinase catalytic subunit                                      |
| BGIBMGA011028-TA         | Cdc20          | 细胞分裂周期蛋白 20<br>Cell division cycle 20, cofactor of APC complex                                         |
| BGIBMGA004270-TA         | Orc6           | 真核细胞 DNA 起始识别复合物亚基 6<br>Origin recognition complex subunit 6                                           |
| BGIBMGA003351-TA         | MCM7           | 微小染色体维持蛋白 7 ( 细胞分裂控制蛋白 47 )<br>Minichromosome maintenance protein 7 (cell division control protein 47) |
| BGIBMGA001198-TA         | CycD2          | <b>细胞周期素 D2</b><br><b>Cyclin D2</b>                                                                    |
| BGIBMGA000828-TA         | MCM6           | 微小染色体维持蛋白 6<br>Minichromosome maintenance protein 6                                                    |
| BGIBMGA013609-TA         | APC/C          | 后期促进复合物亚基 6<br>Anaphase-promoting complex subunit 6                                                    |
| BGIBMGA004646-TA         | Bub3           | 细胞周期阻滞蛋白<br>Cell cycle arrest protein BUB3                                                             |
| BGIBMGA001251-TA         | CycD           | 细胞周期素 D2<br>Cyclin D2                                                                                  |
| BGIBMGA007979-TA         | Cdc7           | <b>细胞分裂控制蛋白 7</b><br><b>Cell division control protein 7</b>                                            |
| BGIBMGA003403-TA         | E2F1,2,3       | 转录因子<br>E2F transcription factor 1/3                                                                   |
| BGIBMGA002754-TA         | Orc1           | 真核细胞 DNA 起始识别复合物亚基 1<br>Origin recognition complex subunit 1                                           |
| BGIBMGA001550-TA         | CDK7           | 周期蛋白依赖性蛋白 7<br>Cyclin-dependent kinase 7                                                               |
| BGIBMGA014215-TA         | Orc5           | 真核细胞 DNA 起始识别复合物亚基 5<br>Origin recognition complex subunit 5                                           |
| BGIBMGA013227-TA         | Esp1           | 姐妹染色体分离酶<br>Separase                                                                                   |
| BGIBMGA007634-TA         | Mps1           | 丝氨酸/苏氨酸激酶<br>Serine/threonine-protein kinase TTK/MPS1                                                  |
| BGIBMGA007924-TA         | Cdh1           | <b>细胞分裂周期类 20 蛋白</b><br><b>Cell division cycle 20-like protein 1, cofactor of APC complex</b>          |
| BGIBMGA004493-TA         | Cdc14          | 细胞分裂周期蛋白 14<br>Cell division cycle 14                                                                  |

基因及其功能注释。在这些差异表达基因中，CDK 和 Cyclin 复合物以及 Cdc、MCM 调控细胞周期过程，Smad 介导细胞表面受体向细胞核的转导过程，Mad 和 Bub 为检验点蛋白，Miz1、E2F 和 DP-1, 2 为转录因子，APC、Abl、Chk、ATM/ATR、Esp、Mps 和 DNA-PK 是调控细胞周期过程的激酶，ORC 是复制的启动子。如图 1

所示，它们参与了细胞周期 G1-S-G2-M 期的重要过程，在家蚕后部丝腺细胞的核内复制过程中具有重要作用。

2.3 细胞周期通路差异表达基因的 qRT-PCR 验证

以家蚕 *rp49* 基因作为内参，选取细胞周期通路中的 6 个差异表达基因 *cdk1*、*cycD1*、

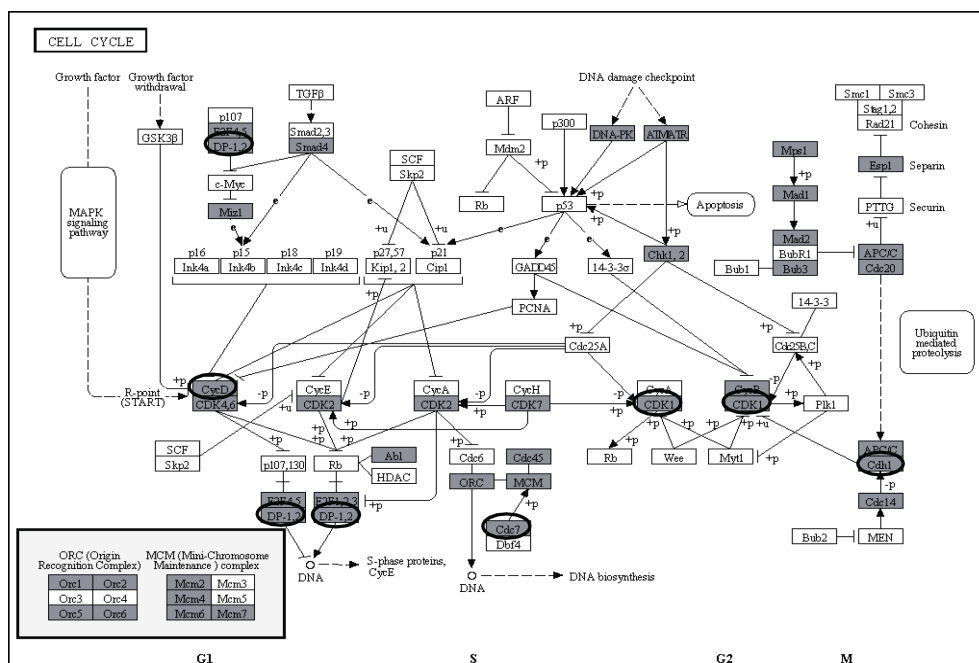


图 1 细胞周期信号通路 (来自 KEGG 数据库)

Fig. 1 The cell cycle pathway (from KEGG)

阴影部分为细胞周期信号通路中的差异表达基因, 圆圈圈示部分为进行 qRT-PCR 的基因。

The shaded parts are differentially expressed genes in cell cycle pathway, genes in the circle are chosen to be verified by qRT-PCR.

*cycD2*、*cdc7*、*cdh1* 和 *dp-1,2* 进行 qRT-PCR 验证, 在图 1 中以黑色圆圈标出, 这些基因编码的蛋白在细胞周期的不同阶段发挥着重要功能。如图 2 所示, 所选基因在野生型家蚕和 *Ras1<sup>CA</sup>* 转基因家蚕后部丝腺中的相对表达量有显著差异 ( $P < 0.05$ ), *cycD2* 基因表达的差异极显著 ( $P < 0.01$ )。这一结果表明, 在转录水平上 *Ras1<sup>CA</sup>* 过表达可以明显促进细胞周期通路基因的表达, 与转录本的差异是一致的。*Ras1<sup>CA</sup>* 转基因家蚕后部丝腺组织中, *Ras1<sup>CA</sup>* 通过上调 CDK 激酶、cyclin 以及 DP-1,2 转录因子的编码基因, 从转录水平上调控细胞周期信号通路, 对于后部丝腺中细胞 DNA 的复制和最终蛋白的合成有着重要的促进作用。

### 3 讨论

RNA-Seq 研究方法促进了昆虫学的研究, 通过该技术获得了飞蝗 *Locusta migratoria* 群居型和散居型及稻纵卷叶螟 *Cnaphalocrosis medinalis*

多种昆虫不同发育阶段的转录组数据, 发现了多个调控生长发育的基因和代谢通路 (Chen *et al.*, 2010; Li *et al.*, 2012)。通过分析埃及伊蚊 *Aedes aegypti* 吸血后不同时间段的转录组数据, 发现了调控其生理生化功能的重要基因 (Bonizzoni *et al.*, 2011)。RNA-Seq 技术可以推动昆虫学领域的研究, 为害虫的防治和资源昆虫的利用提供基础数据资源 (刘金定等, 2013)。利用高通量测序技术, 获得 *Ras1<sup>CA</sup>* 转基因家蚕后部丝腺的数据信息, 同时将表达差异显著的基因与生理现象相联系, 为深入研究转基因家蚕蚕丝高产的分子机制奠定基础。家蚕后部丝腺在 5 龄时大量合成并分泌蚕丝蛋白的主要成分——丝素蛋白, 本研究取家蚕上簇早期后部丝腺进行 RNA-Seq 分析, 发现野生型家蚕和 *Ras1<sup>CA</sup>* 转基因家蚕的后部丝腺中有 2 636 个差异表达基因, 分布在 277 个信号通路中, 涉及能量代谢和细胞运动等。*Ras1<sup>CA</sup>* 促进了 Raf-MAPK 和 PI3K-Akt 信号通路中的基因表达, 包括 *mek*、*jnk* 和 *erk* 等, 可能影响下游多种效应

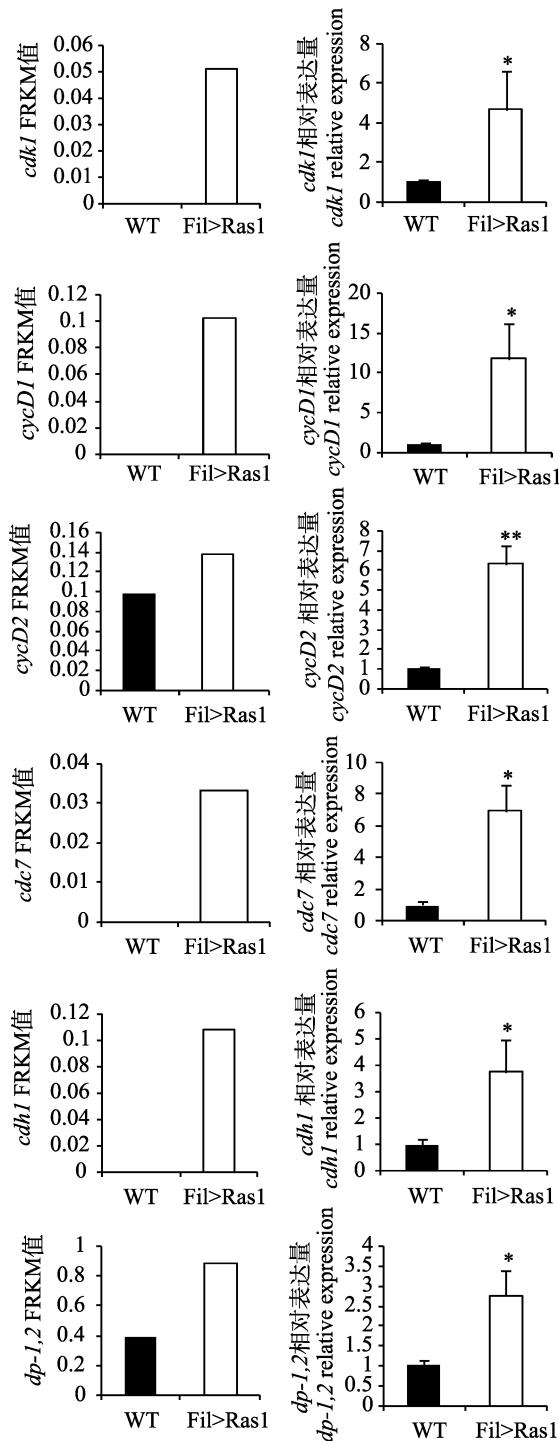


图2 实时荧光定量 PCR 验证细胞周期通路相关基因的变化  
Fig. 2 Verification of differential expression of DEGs in cell cycle pathway by qRT-PCR

WT 表示野生型家蚕, Fil>Ras1 表示转基因家蚕。图中\*表示野生型和 *Ras1<sup>CA</sup>* 转基因家蚕细胞周期通路相关基因表达有显著差异 ( $P < 0.05$ ), \*\*表示有极显著差异 ( $P < 0.01$ )。

WT: Wild type silkworms, Fil>Ras1: Transgenic silkworms.

\* and \*\* indicate the difference of DEGs expression in cell cycle pathway between WT and Fil>Ras1 transgenic silkworms is significant at 0.05 and 0.01 level, respectively.

分子,参与细胞增殖的调节(Zhang and Liu 2002)。

转录组测序结果表明,野生型家蚕和 *Ras1<sup>CA</sup>* 转基因家蚕的细胞周期信号通路中有 42 个差异表达基因,涉及细胞周期整个过程。在细胞周期的调控过程中 CDK-cyclin 为主要调控复合物,细胞周期蛋白依赖性激酶抑制剂 (CKI) 起负调控作用,转录因子在不同阶段促进或抑制 DNA 的合成,上文结果表明差异表达基因包括编码 CDKs、cyclins、Cdc 和转录因子的基因,本研究 qRT-PCR 验证所选取的 6 个基因涵盖了以上几类,具有一定的代表性。在哺乳动物中研究表明 CDKs 和 cyclins 在调节细胞周期从 G1 到 S 期的转化过程中具有重要作用,CDK2 可与 cyclinE 相结合,CDK4/6 可与 cyclinD 相结合,激活下游因子,调节细胞周期 G 期到 S 期的过程(King *et al.*, 1996; Evan and Vousden, 2001), CDK1 主要负责 M 期的调控(Weigel and Moore, 2007),BmCyclinB 和 BmCyclinB3 能够抑制细胞增殖,完成细胞周期过程(Pan *et al.*, 2013)。通过对表 3 列出的细胞周期信号通路差异表达基因进行分析,发现 *Ras1<sup>CA</sup>* 过表达影响家蚕后部丝腺细胞周期相关基因的表达,从而影响细胞周期这一生理过程,促进 DNA 的复制和丝蛋白的合成。

在胚胎发育早期,家蚕丝腺的细胞数目即已决定,其核内复制方式决定了在整个幼虫期丝腺组织只进行细胞体积的增大和重量的增加,细胞数目不再改变,因此细胞的大小决定了后部丝腺器官的大小(Tashiro *et al.*, 1968; Edgar and Orr-Weaver, 2001)。家蚕后部丝腺合成分泌丝蛋白的主要组成部分丝素蛋白,丝素重链蛋白(Fib-H)、丝素轻链蛋白(Fib-L)和 P25 蛋白(Fibrohexamerin, Fhx)在后部丝腺大量表达(Tanaka *et al.*, 1993)。随着细胞周期相关基因的表达,细胞核内 DNA 不断复制,丝素蛋白基因的拷贝数也大量增加,在转录水平和翻译水平上都有利于丝蛋白的合成,丝腺的生长和丝蛋白的合成对于提高蚕丝产量都有着重要的意义(Ma *et al.*, 2011)。

qRT-PCR 验证结果表明, *Ras1<sup>CA</sup>* 过表达在转录水平能够促进 *cdk1*、*cycD1*、*cycD2*、*cdc7*、*cdh1* 和 *dp-1,2* 基因的表达,与转录本分析结果一致,

表明转录组数据可信,也进一步证实细胞周期相关基因的表达受到 *RasI<sup>CA</sup>* 的转录调控,解释了 *RasI<sup>CA</sup>* 转基因家蚕蚕丝高产的分子机制。但是, *RasI<sup>CA</sup>* 在后部丝腺中过表达后,是通过哪一条或几条下游通路影响细胞周期信号通路的,还需要进一步研究。

本研究中的实验材料 Fil-GAL4/UAS-*RasI<sup>CA</sup>* 转基因家蚕品系,将杂交育种与转基因育种优势有机结合,提高了蚕丝产量,带来新的蚕丝产业革命。运用转录组学技术进行分析并结合 qRT-PCR 验证,可进一步的了解 *RasI<sup>CA</sup>* 在家蚕中过表达后,对后部丝腺细胞中表达基因的转录调控,通过促进细胞周期相关基因的表达,调控细胞周期这一生理过程,增加 DNA 的复制,为蚕丝蛋白的合成和分泌提供了大量的遗传物质 (Shigematsu *et al.*, 1978)。对 *RasI<sup>CA</sup>* 转基因家蚕后部丝腺的转录组学研究,阐释了可能影响蚕丝高产的分子调控机制,为蚕丝业的发展和分子育种技术提供了理论支持。

## 参考文献 (References)

- Bonizzoni M, Dunn WA, Campbell CL, Olson KE, Dimon MT, Marinotti O, James AA, 2011. RNA-seq analyses of blood-induced changes in gene expression in the mosquito vector species, *Aedes aegypti*. *BMC Genomics*, 12(1): 82.
- Chen S, Yang PC, Jiang F, Wei YY, Ma ZY, Kang L, 2010. *De novo* analysis of transcriptome dynamics in the migratory locust during the development of phase traits. *PLoS ONE*, 5(12): e15633.
- Edgar BA, Orr-Weaver TL, 2001. Endoreplication cell cycles: more for less. *Cell*, 105(3): 297–306.
- Evan GI, Vousden KH, 2001. Proliferation, cell cycle and apoptosis in cancer. *Nature*, 411(6835): 342–348.
- Imamura M, Naka J, Inoue S, Quan GX, Kanda T, Tamura T, 2003. Targeted gene expression using the GAL4/UAS system in the silkworm *Bombyx mori*. *Genetics*, 165(3): 1329–1340.
- King RW, Deshaies RJ, Peters JM, Kirschner MW, 1996. How proteolysis drives the cell cycle. *Science*, 274(5293): 1652–1659.
- Li SW, Yang H, Liu YF, Liao QR, Du J, Jin DC, 2012. Transcriptome and gene expression analysis of the rice leaf folder, *Cnaphalocrocis medinalis*. *PLoS ONE*, 7 (11): e47401.
- Liu JD, Zhang Z, Huang SQ, Li F, 2013. Insect RNA-Seq data analysis pipeline. *Chinese Journal of Applied Entomology*, 50(5): 1458–1468. [刘金定, 张赞, 黄水清, 李飞, 2013. 昆虫 RNA-Seq 数据的分析流程. 应用昆虫学报, 50(5): 1458–1468.]
- Ma L, Ma Q, Li X, Cheng LL, Li K, Li S, 2014. Transcriptomic analysis of differentially expressed genes in the *RasI<sup>CA</sup>*-overexpressed and wild type posterior silk glands. *BMC Genomics*, 15(1): 182.
- Ma L, Xu HF, Zhu JQ, Ma SY, Liu Y, Jiang RJ, Xia QY, Li S, 2011. *RasI<sup>CA</sup>* overexpression in the posterior silk gland improves silk yield. *Cell Research*, 21(6): 934–943.
- Orton RJ, Sturm OE, Vyshemirsky V, Calder M, Gilbert DR, Kolch W, 2005. Computational modelling of the receptor-tyrosine-kinase-activated MAPK pathway. *The Biochemical Journal*, 392(2): 249–261.
- Pan MH, Hong KL, Chen XY, Pan C, Chen XM, Kuang XX, Lu C, 2013. *BmCyclinB* and *BmCyclinB3* are required for cell cycle progression in the silkworm, *Bombyx mori*. *Science China Life Sciences*, 56(4): 360–365.
- Pines J, 1995. Cyclins and cyclin-dependent kinases: a biochemical view. *Biochemical Journal*, 308(3): 697–711.
- Shigematsu H, Kurata K, Takeshita H, 1978. Nucleic acids accumulation of silk gland of *Bombyx mori* in relation to silk protein. *Comparative Biochemistry and Physiology Part B: Comparative Biochemistry*, 61(2): 237–242.
- Talebi E, Subramanya G, 2009. Genetic distance and heterosis through evaluation index in the silkworm, *Bombyx mori* (L.). *American Journal of Applied Sciences*, 6(12): 1981–1987.
- Tamura T, Thibert C, Royer C, Kanda T, Abraham E, Kamba M, Komoto N, Thomas JL, Mauchamp B, Chavancy G, Shirk P, Fraser M, Prudhomme JC, Couble P, 2000. Germline transformation of the silkworm *Bombyx mori* L. using a *PiggyBac* transposon-derived vector. *Nature Biotechnology*, 18(1): 81–84.
- Tanaka K, Mori K, Mizuno S, 1993. Immunological identification of the major disulfide-linked light component of silk fibroin. *Journal of Biochemistry*, 114(1): 1–4.
- Tashiro Y, Morimoto T, Matsuura S, Nagata S, 1968. Studies on the posterior silk gland of the silkworm *Bombyx mori*. I. Growth of posterior silk gland cells and biosynthesis of fibroin during the fifth larval instar. *The Journal of Cell Biology*, 38(3): 574–588.
- Wang Z, Gerstein M, Snyder M, 2009. RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics*, 10(1): 57–63.
- Weigel NL, Moore NL, 2007. Cyclins, cyclin dependent kinases, and regulation of steroid receptor action. *Molecular and Cellular Endocrinology*, 265-266(1): 157–161.
- Zhang CD, Li FF, Chen XY, Huang MH, Zhang J, Cui HJ, Pan MH, Lu C, 2012. DNA replication events during larval silk gland development in the silkworm, *Bombyx mori*. *Journal of Insect Physiology*, 58(7): 974–978.
- Zhang W, Liu HT, 2002. MAPK signal pathways in the regulation of cell proliferation in mammalian cells. *Cell Research*, 12(1): 9–18.
- Zhang Z, Liu JD, Huang SQ, Li F, 2012. Application of bioinformatics in entomology. *Chinese Journal of Applied Entomology*, 49(1): 1–11. [张赞, 刘金定, 黄水清, 李飞, 2012. 生物信息学在昆虫学研究中的应用. 应用昆虫学报, 49(1): 1–11.]
- Zimmet J, Ravid K, 2000. Polyploidy: Occurrence in nature, mechanisms, and significance for the megakaryocyte-platelet system. *Experimental Hematology*, 28(1): 3–16.