

家蚕 *Septin* 基因家族的全基因组鉴定及分析

王 磊, 朱保健, 魏国清, 钱 岑, 刘朝良*

(安徽农业大学生命科学学院, 合肥 230036)

摘 要: *Septin* 基因家族存在于除植物外的所有真核生物中, 是具有 GTPase 活性保守基因家族。*Septin* 蛋白功能多样, 能够参与细胞分裂、细胞内物质转运、细胞周期调控及细胞凋亡等生理反应, 并与肿瘤发生、神经功能障碍和病原物侵染的过程直接相关。利用果蝇的 5 个 *Septin* 基因家族蛋白序列检索家蚕基因组数据库, 发现家蚕基因组中存在 4 个 *septin* 基因, 其蛋白序列都具有完整的保守结构域; 通过与果蝇、酵母等 *Septin* 蛋白进行聚类分析, 发现该基因家族具有高度保守性, 并对家蚕 4 个 *septin* 进行命名为 *BmPnut*、*BmSeptin1*、*BmSeptin2* 和 *BmSeptin4*; 通过查找基因芯片数据库分析了家蚕 *septin* 基因家族在各组织的表达模式, 4 个 *septin* 基因在各组织表达量各不相同, 为下一步研究家蚕 *septin* 基因功能奠定了基础。

关键词: 家蚕; *Septin*; 基因家族; 基因组

中图分类号: S881.2

文献标识码: A

文章编号: 1672-352X (2014)03-0345-06

Genome-wide analysis and identification of *Septin* gene family in *Bombyx mori*

WANG Lei, ZHU Baojian, WEI Guoqing, QIAN Cen, LIU Chaoliang

(School of Life Science, Anhui Agricultural University, Hefei 230036)

Abstract: The *septin* gene family is broadly distributed in eukaryotes excluding plants, and conserved GTPase activity. Septins are of multifunctional proteins, participating in cytokinesis, vesicle trafficking, cell cycle regulation and apoptosis, and also involve in the pathogenesis of different diseases including neoplasia, neurodegeneration and pathogen infections. Based on *Drosophila melanogaster* septin amino acid sequences blasted with Silkworm Genome Database (SilkDB), we found 4 *septin* genes in *Bombyx mori* genome. These 4 septin protein's sequences have conserved domains, including G1, G2, G3, G4, G5 and coiled-coil domains. Phylogenesis tree and multiple sequence alignment results indicated that the gene family was highly conserved, *B. mori* septins were named *BmPnut*, *BmSeptin1*, *BmSeptin2* and *BmSeptin4* based on similarities with *D. melanogaster* septins. The expression profiles of *septins* in *B. mori* were analyzed by searching silkworm microarray data. The results in this study provide a good basis for further studies on functions of septins in *B. mori*.

Key words: *Bombyx mori*; *Septin*; gene family; genome

Septin 是具有 GTPase 活性的基因家族, 最早在酿酒酵母中发现其参与了芽生细胞周期过程^[1], 后来在线虫^[2]、果蝇^[3]、哺乳动物^[4]中也鉴定到直系同源基因, 但在原生物和植物基因组信息中未发现它的存在。*Septin* 家族早期在酵母中命名为 CDC 系列, 直到 2002 年 HUGO (human genome organization) 及 MGN (mouse genome nomenclature) 统一命名为 *septin* (缩写为 SEPT)^[5]。

Septin 属于含 P 环 (loop) 的 GTPase 超家族, 物种间及同种家族的序列同源性高, 都含有 GTP 结合部位, 其中在 GTP 结合结构域内至少含有 5 个保守的基序 (motif), 分别命名为 G1、G2、G3、G4 和 G5^[6]。G1 基序 (GxxxxGK [ST]) 形成一个弹性的环与核苷酸的磷酸群相作用。G3 基序是 DxxG 后面跟随一些疏水残基, 此区域能够结合 Mg^{2+} 和 GTP。G4 基序 (NKxD) 对于 GTP 结合的特异性很

收稿日期: 2013-11-22

基金项目: 安徽省自然科学基金青年项目 (No:1308085QC60) 和安徽农业大学稳定和引进人才项目 (引进 2013-8) 共同资助。

作者简介: 王 磊, 博士, 讲师。E-mail: wanglei20041225@163.com

* 通信作者: 刘朝良, 博士, 教授, 博士生导师。E-mail: cyschx@163.com

重要。在 septin 蛋白的 C 末端可以形成卷曲 (α helix coiled-coil) 结构域, 此结构域对于 septin 间相互作用是必不可少的^[7]。Septin 通常组成成低聚的复合体和高度有序的聚合物, 这些 septin 聚合物的组配方式、结构与功能在线虫、酵母和人都进行了大量的研究^[8-10]。

在酵母中鉴定了 7 种 septin (septin1~7), 主要参与了胞质分裂过程、膜分割区的界限设定 (compartmentalization)、囊泡运输和分泌过程、有丝分裂纺锤体的定向以及有丝分裂中细胞动力学的关卡检查^[11]。在线虫中发现了 2 个 septin (Unc-59 和 Unc-61), 其与卵裂沟相关, 在分裂末期定位于中间体, 并且为胚胎后期的胞质分裂所必需^[12]。在果蝇中, 5 个 septin 蛋白被鉴定, 参与了细胞质分裂以及分裂末期卵裂沟的迁移过程, septin 蛋白复合物具有结合和水解 GTP 的活性^[13]。在人类基因组中发现了 13 个 septin 基因, 分为 4 个亚家族, 并且有很多变异剪接本, 但同一基因不同变异剪接本在多数情况下编码相同的蛋白质。人 septin 蛋白不仅在细胞内物质的运输与分泌中有重要作用, 包括调节神经递质与血小板的释放反应等, 而且已明确了 septin 家族与 Parkinson 病、Alzheimer 病、Down's 综合征等神经系统疾病, 白血病及中枢神经系统肿瘤、结直肠癌、卵巢癌、乳腺癌、前列腺癌、膀胱癌、肾细胞癌等肿瘤, 志贺 (氏) 杆菌和立克次氏体之类的病毒和细菌在内的病原体感染, 男性不孕症以及遗传性神经肌肉萎缩 (HNA) 等疾病发生有关^[5, 11]。

家蚕 (*Bombyx mori*) 是我国重要的经济昆虫, 同时也是研究遗传学、发育生物学和行为学的理想模式动物。随着家蚕基因组计划的完成, 加速了对家蚕生理生化与分子生物学等方面的研究基础。本研究旨在利用已知的果蝇 septin 基因信息和家蚕基因组数据库和蛋白质数据库, 通过 Blast 比对分析家蚕中可能存在的 septin 基因, 并进行生物信息学分析, 为下一步研究家蚕 septin 基因的生物学功能以及利用家蚕 septin 来研究 septin 与人类疾病的关系奠定基础。

1 材料与方法

1.1 数据库的搜索

果蝇 (*Drosophila melanogaster*) septin 基因及其编码的蛋白质序列 (表 1) 来源于 NCBI 数据库。在西南大学家蚕基因组数据库 SilkDB (<http://silkworm.genomics.org.cn/>) 中分别以果蝇 septin 蛋

白序列 (表 1) 进行 tblastn 比对, E 值设定为 10^{-6} 。同时对家蚕 EST 数据库 (<http://silkbases.ab.a.u-tokyo.ac.jp/cgi-bin/index.cgi>)^[14] 进行比对, E 值设定为 0.001。候选序列再进行 Pfam (<http://pfam.sanger.ac.uk/>) 结构域分析, 若存在 septin 则认为该候选蛋白属于 septin 蛋白家族。利用 SignalP 3.0 (<http://www.cbs.dtu.dk/services/SignalP-3.0/>) 在线分析蛋白的信号肽序列。

1.2 多序列比对及系统发生树的构建

利用 Clustal X 程序 (<http://www.clustal.org/>)^[15] 对果蝇和家蚕 septin 蛋白序列进行多序列比对分析, 分析结果使用 MEGA4.0 (<http://megasoftware.net>)^[16] 程序采用 Neighbor Joining 法 (NJ) 进行系统发生树的构建, bootstrap 值设置为 500。家蚕 septin 蛋白利用 Clustal X 程序进行多序列比对, 比对结果利用 DNAMAN 软件进行输出结果的设置。

表 1 检索用果蝇 septin 基因家族信息

Table 1 The query information of septin gene family in *D. melanogaster*

基因名称 Gene name	氨基酸数 Amino acid number	登录号 Accession number
<i>Pnut</i>	539	NM_165597
<i>septin1</i>	361	NM_167747
<i>septin2</i>	419	NM_079693
<i>septin4</i>	427	NM_167530
<i>septin5</i>	422	NM_165578

1.3 蛋白结构分析

家蚕 Septin 蛋白序列提交 NCBI Conserved Domain (<http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>) 和 Pfam (<http://pfam.sanger.ac.uk/>) 分析; 卷曲结构利用在线 COILS (http://www.ch.embnet.org/software/COILS_form.html)^[17]。

1.4 表达模式分析

基因芯片数据来自 *Bombyx mori* Microarray Database (BmMDB)^[18], 利用基因芯片结果对 septin 基因的表达模式进行分析。

2 结果与分析

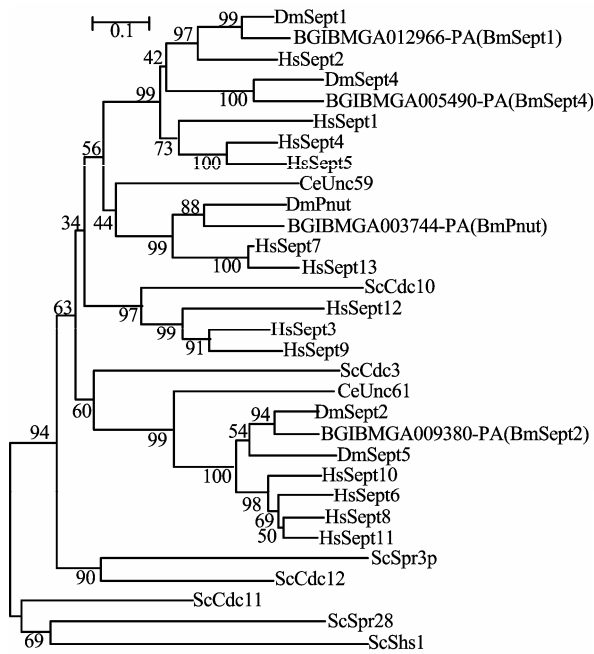
2.1 家蚕 septin 基因的鉴定

利用果蝇 septin 蛋白序列 BLAST 家蚕基因组数据库, 共鉴定到 4 个非冗余 septin 基因, 并且在 EST 库中都得到鉴定 (表 2)。家蚕 septin 基因家族数目 (4 个) 比酵母 (7 个)、果蝇 (5 个) 和人类 (13 个) 少, 可能在进化过程中某些 septin 蛋白丢失。为了便于分析和描述, 根据家蚕 septin 与果蝇、

表 2 家蚕 *septin* 基因家族主要信息
Table 2 Basic information of *septin* gene family in *B. mori*

基因名称 Gene name	基因序号 Gene ID (BGIBMGA)	氨基酸长度 AA Length	支架及位置 Scaffold & Position
<i>BmPnut</i>	003744-TA	383	nscf2674, 7185382...7191756
<i>BmSeptin1</i>	012966-TA	339	nscf3058, 2846860...2854832
<i>BmSeptin2</i>	009380-TA	419	nscf2948, 2387983...2394967
<i>BmSeptin4</i>	005490-TA	374	nscf2828, 4471984...4483075

外显子数目 No. of Exons	内含子数目 No. of Introns	染色体位置 Chromosomal Location	方向 Orientation	EST 比对 EST Blast
7	6	5	+	+
9	8	16	+	+
9	8	14	+	+
7	6	8	+	+



果蝇 5 个 *septin*, 线虫 2 个 *septin*, 酵母 7 个 *septin*, 人 13 个 *septin*

Drosophila melanogaster septin (DmPnut, gi | 730352; DmSept1, gi | 17647925; DmSept2, gi | 17738071; DmSept4, gi | 24642597; DmSept5, gi | 21356243), *Caenorhabditis elegans* septin (CeUnc59, gi | 17509405; CeUnc61, gi | 32566810), *Saccharomyces cerevisiae* septin (ScCdc3, gi | 6323346; ScCdc10, gi | 6319847; ScCdc11, gi | 6322536; ScCdc12, gi | 6321899; ScShs1, gi | 6319976; ScSpr28, gi | 6320424; ScSpr3, gi | 6321496), *Homo sapiens* septin (HsSept1, gi | 16604248; HsSept2, gi | 4758158; HsSept3, gi | 22035572; HsSept4, gi | 4758942; HsSept5, gi | 9945439; HsSept6, gi | 22035577; HsSept7, gi | 4502695; HsSept8, gi | 41147049; HsSept9, gi | 6683817; HsSept10, gi | 8088518; HsSept11, gi | 8922712; HsSept12, gi | 23242699; HsSept13, gi | 113418512)

图 1 家蚕与果蝇、线虫、酵母及人 *septin* 蛋白进化分析
Figure 1 Phylogenetic relationships among the septin from *B. mori* and other species

线虫等 *septin* 的聚类分析结果 (图 1), 把家蚕 4 个 *septin* 分别命名为 *BmPnut*、*BmSeptin1*、*BmSeptin2* 和 *BmSeptin4* (表 2)。内含子分析表明, 家蚕 *septin* 基因都含有 6 个或 8 个内含子序列。4 个 *septin* 基因分别位于家蚕不同的染色体上, 表达的方向都是正向。通过信号肽软件预测发现, 家蚕 4 个 *septin* 都没有信号肽序列, 是非分泌蛋白, 可能直接在细胞内发挥作用。

2.2 家蚕 *septin* 结构域分析

通过结构域分析, 4 个家蚕 *septin* 蛋白都含有 GTP 结合结构域内的 5 个保守基序 G1、G2、G3、G4 和 G5, 一般 G1、G2 和 G3 位置紧密相连, 位于 N 末端; G4 和 G5 位于中部区域 (图 2)。通过 COILS 程序分析发现, 4 个家蚕 *septin* 在 C 末端都含有 coiled-coil 结构域, 但此结构域的长度不同, 其中 *BmSeptin2* 的 coiled-coil 结构域最长 (69 个氨基酸), *BmSeptin4* 最短 (41 个氨基酸) (图 2)。

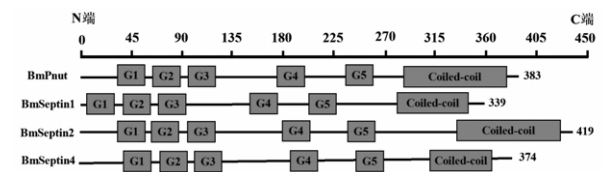


图 2 家蚕 *septin* 蛋白中保守基序的分配模式
Figure 2 The distribution mode of conserved domain in septin of *B. mori*

2.3 家蚕 *septin* 多序列联配

通过多序列联配发现, 家蚕各 *septin* 蛋白相似性不高, *BmPnut* 与 *BmSeptin1*、*BmSeptin2*、*BmSeptin4* 的相似性分别为 36.9%、32.9%和 34.8%, 其中 *BmSeptin1* 与 *BmSeptin4* 相似性最高, 为 52.8% (表 3)。通过图 3 可以看出, 家蚕 *septin* 在 N 末端和中部相似性较高, 而在 C 末端则相似性较低。其

中,与其他 septin 相比, BmSeptin2 在 308~323 和 366~388 氨基酸处多了 2 段序列。

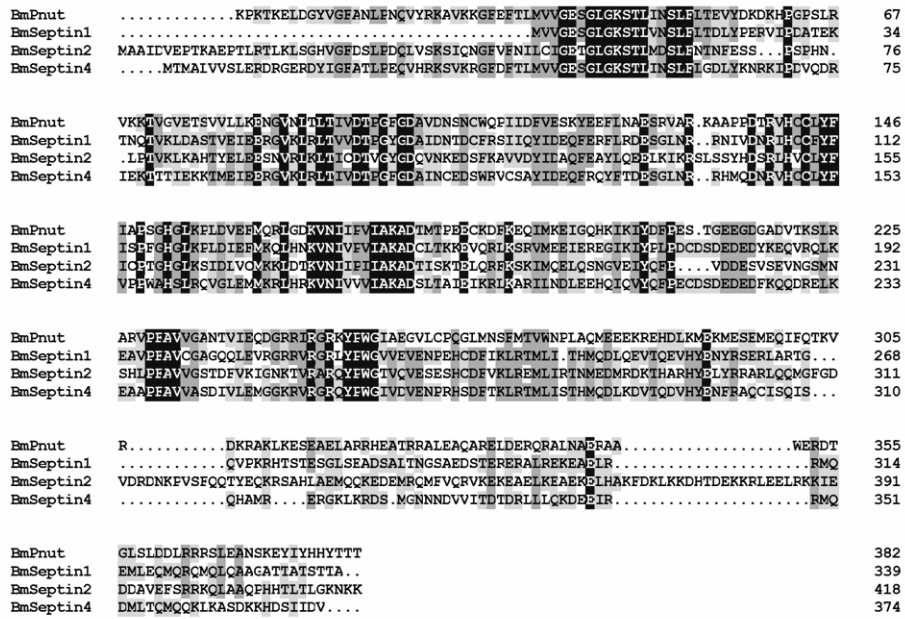
2.4 家蚕 septin 表达模式

利用家蚕基因芯片数据库对家蚕 septin 在各组织的表达情况进行分析,检测的家蚕组织有睾丸 (testis)、卵巢 (ovary)、头部 (head)、脂肪体 (fat body)、表皮 (integument)、中肠 (midgut)、血细胞 (hemocyte)、马氏管 (malpighian tubule)、前中部丝腺 (A/MSG) 和后部丝腺 (PSG) (图 4)。4 个家蚕 septin 基因在各个组织中表达有很大差异。 BmPnut (芯片号为 sw11225) 在睾丸 (testis) 中表达量最高; BmSeptin1 (芯片号为 sw14493) 在睾丸、卵巢 (ovary) 和血细胞 (hemocyte) 中表达量较高;

BmSeptin2 (芯片号为 sw05096) 与 BmSeptin1 表达模式相似,除在睾丸、卵巢和血细胞中表达量较高外,在头部 (head) 表达量也较高; BmSeptin4 (芯片号为 sw13656) 在中肠 (midgut) 中表达量最高 (图 4)。

表 3 家蚕 septin 间相似性比较 Table 3 Percent identity of B. mori septin %

Table with 5 columns: Gene, BmPnut, BmSeptin1, BmSeptin2, BmSeptin4. Rows show percent identity values for each gene pair.



黑色阴影表示完全一致,深灰色阴影代表相似性大于 75%,浅灰色阴影代表相似性大于 50%

Shading type: black=100%, dark grey ≥75%, light grey ≥50%

图 3 家蚕 4 个 septin 蛋白的多序列比对

Figure 3 Multiple sequence alignment of four septins in B. mori

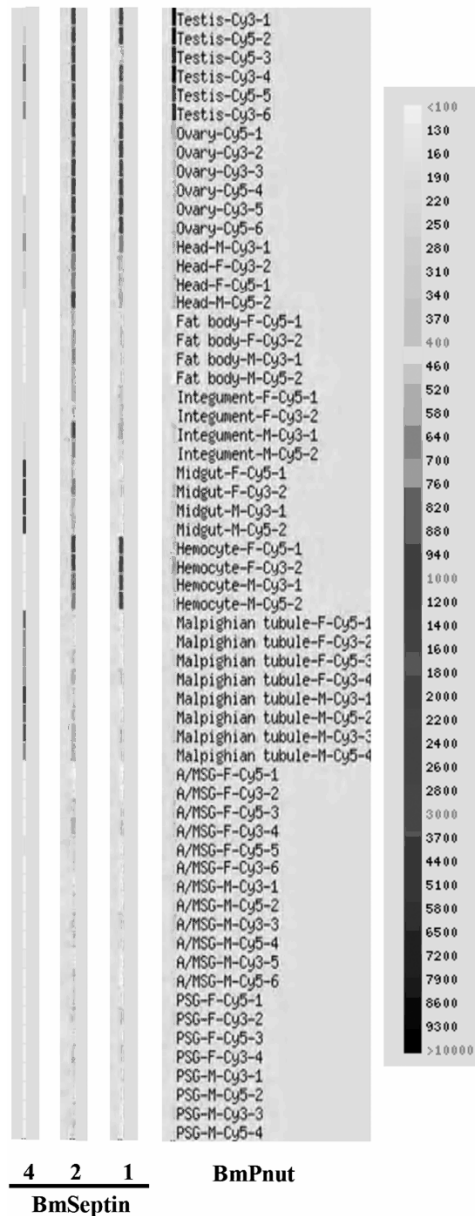
3 讨论

Septin 基因家族成员众多,在不同物种中数目不尽相同。本研究在家蚕基因组中鉴定到 4 个 septin 基因家族成员,比已报道的线虫基因组中最少 (2 个),而人类基因组中发现有 13 个成员。家蚕 septin 基因家族数目比果蝇少 1 个,在家蚕中 septin 的功能可能多样化,某个 septin 具有缺失 septin 基因的功能。另外,家蚕与果蝇的生活环境、生活习性等都有很大的不同,这也可能是导致 septin 基因家族数目不同的原因。通过进化基因组学的方法对信息

素一型受体基因超家族进行了系统地研究,提出动物适应环境可以反映在基因数目变异上的假说,这对于理解基因家族数目的差异有很大帮助 [19-20]。

通过进化分析发现家蚕 septin 与果蝇 septin 的同源性很高 (图 1),并且家蚕 septin 都含有保守的结构域,这可能与它的结构和功能是密切相关的。Septin 基因家族最保守的核心结构域是 GTP 结合部位,具有 GTPase 活性,在这结构域内部的 G1~G5 基序都对应对应的功能,这有助于解析家蚕 septin 基因的功能。另外,在家蚕 septin 蛋白 C 末端,都有保守的卷曲结构,这可能与形成 septin 复合体相

关。在酵母及人类 *septin* 研究中, 发现了很多 *septin* 复合体的现象, 以复合体的形式来执行功能, 而在复合体形成的过程中, 卷曲结构起到了连接的作用。因此, 推测在家蚕中 *septin* 也形成复合体, 这需要进一步的研究证实。



BmPnut 芯片号为 sw11225, BmSeptin1 芯片号为 sw14493, BmSeptin2 芯片号为 sw05096, BmSeptin4 芯片号为 sw13656

(BmPnut probe ID: sw11225, BmSeptin1 probe ID: sw14493, BmSeptin2 probe ID: sw05096, BmSeptin4 probe ID: sw13656)

图 4 家蚕 *septin* 基因在各组织分布的基因芯片结果
Figure 4 Tissue distributions of *B. mori* septins from microarray database

基因芯片结果显示 3 个 *septin* 基因 (*BmPnut*, *BmSeptin1* 和 *BmSeptin2*) 在家蚕睾丸中表达量高,

由此可以预测 *septin* 在家蚕生殖方面起作用。在人类医学的研究过程中发现 *septin* 家族可能与男性不孕症有关, 免疫组织化学表明 *Septin4* 定位于精子环并且在特发性弱精子症患者精子中的表达显著低于正常男性, 人类部分精子无力症患者的精子有类似的 *septin* 环结构紊乱现象^[21]。*Septin* 基因敲除后的雄性老鼠精子形态及运动情况受到影响, 产生了不育症^[22]。另外, BmPnut 和 BmSeptin1 在感染 BmNPV 的细胞系中检测到, 而在正常细胞系中没有发现 (结果未给出), 这也预测家蚕 *septin* 可能参与了家蚕的免疫。在人类感染的发生以及病原微生物与人类相互作用的过程中, *septin* 蛋白可能参与了人类感染过程^[11]。外来病原微生物进入细胞及在细胞内的运动存在一种 *septin* 蛋白依赖机制, 同时细胞吞噬反应及免疫产生的细胞骨架重组也与 *septin* 蛋白的功能相关。但在感染微生物的家蚕脂肪体 EST 文库中, 没有发现家蚕 *septin* 基因的表达 (表未列出), 这也提示 *septin* 可能对不同病原物的应激组织和应激反应有所不同。

Septin 基因家族与人类疾病的关系已有大量的研究, 业已证实 *septin* 家族与神经系统疾病、肿瘤、感染及男性不育症等疾病发生相关。家蚕是研究人类生理、疾病、遗传和开发新型药物的优良模式动物, 已经建立了家蚕一些人类疾病和药物筛选等方面的研究^[23]。本研究从全基因组对家蚕 *septin* 基因家族进行了分析, 也探讨了 *septin* 蛋白可能存在的生物学功能, 为今后分析家蚕 *septin* 功能, 探索以 *septin* 家族为切入点建立研究人类疾病与药物筛选的家蚕模型提供理论基础。

参考文献:

- [1] Hartwell L H, Mortimer R K, Culotti J, et al. Genetic control of the cell division cycle in yeast: v. genetic analysis of CDC mutants [J]. *Genetics*, 1973, 74 (2): 267-286.
- [2] Amir S, Wang R, Matzkin H, et al. MSF-A interacts with hypoxia-inducible factor-1 α and augments hypoxia-inducible factor transcriptional activation to affect tumorigenicity and angiogenesis [J]. *Cancer Res*, 2006, 66(2): 856-866.
- [3] Munoz-Soriano V, Nieto-Arellano R., Paricio N. Septin 4, the *Drosophila* ortholog of human CDCrel-1, accumulates in parkin mutant brains and is functionally related to the Nedd4 E3 ubiquitin ligase[J]. *J Mol Neurosci*, 2012,48(1): 136-143.
- [4] Kinoshita M, Field C M, Coughlin M L, et al. The mammalian septin complexes: composition, ultrastructure and interaction [J]. *Mol Biol Cell*, 2000, 11(Suppl): 102
- [5] 吕炳建, 来茂德. *Septin* 基因家族与人类疾病[J]. 国际遗传学杂志, 2006, 29(06): 441-445.

- [6] Cao L H, Yu W B, Wu Y H, et al. The evolution, complex structures and function of septin proteins [J]. *Cell Mol Life Sci*, 2009, 66 (20): 3309-3323.
- [7] Casamayor A, Snyder M. Molecular dissection of a yeast septin: distinct domains are required for septin interaction, localization, and function [J]. *Mol Cell Biol*, 2003, 23 (8): 2762-2777.
- [8] John C M, Hite R K, Weirich C S, et al. The *Caenorhabditis elegans* septin complex is nonpolar [J]. *EMBO J*, 2007, 26(14): 3296-3307.
- [9] Bertin A, McMurray M A, Grob P, et al. *Saccharomyces cerevisiae* septins: Supramolecular organization of heterooligomers and the mechanism of filament assembly [J]. *Proc Natl Acad Sci USA*, 2008, 105 (24): 8274-8279.
- [10] Low C, Macara I G. Structural analysis of septin 2, 6, and 7 complexes [J]. *J Biol Chem*, 2006, 281(41): 30697-30706.
- [11] 余文博, 江松敏, 余龙. *septin* 基因家族的研究进展[J]. *遗传*, 2008, 30(09): 1097-1107.
- [12] Nguyen T Q, Sawa H, Okano H, et al. The *C. elegans* septin genes, *unc-59* and *unc-61*, are required for normal postembryonic cytokinesis and morphogenesis but have no essential function in embryogenesis [J]. *J Cell Sci*, 2000, 113(21): 3825-3837.
- [13] Field C M, Alawar O, Rosenblatt J, et al. A purified *Drosophila* septin complex forms filaments and exhibits GTPase activity [J]. *J Cell Biol*, 1996, 133 (3): 605-616.
- [14] Mita K, Morimyo M, Okano K, et al. The construction of an EST database for *Bombyx mori* and its application [J]. *Proc Natl Acad Sci USA*, 2003, 100 (24): 14121-14126.
- [15] Larkin M A, Blackshields G, Brown N P, et al. Clustal W and Clustal X version 2.0 [J]. *Bioinformatics*, 2007, 23 (21): 2947-2948.
- [16] Tamura K, Dudley J, Nei M, et al. MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) Software Version 4.0 [J]. *Mol Biol Evol*, 2007, 24 (8): 1596-1599.
- [17] Lupas A, Van Dyke M, Stock J. Predicting coiled coils from protein sequences [J]. *Science*, 1991, 252 (5009): 1162-1164.
- [18] Xia Q, Cheng D, Duan J, et al. Microarray-based gene expression profiles in multiple tissues of the domesticated silkworm, *Bombyx mori* [J]. *Genome Biol*, 2007, 8(8): R162.
- [19] Pfister P, Rodriguez I. Olfactory expression of a single and highly variable V1r pheromone receptor-like gene in fish species [J]. *Proc Natl Acad Sci USA*, 2005, 102 (15): 5489-5494.
- [20] Wang G D, Shi P, Zhu Z H, et al. More functional V1R genes occur in nest-living and nocturnal terricolous mammals [J]. *Genome Biol Evol*, 2010, 2: 277-283.
- [21] 李玉山, 冯晓霞, 吉晓菲, 等. SEPT4 蛋白在特发性弱精子症患者精子中的表达[J]. *中华男科学杂志*, 2011, 17(08): 699-702.
- [22] Ihara M, Kinoshita A, Yamada S, et al. Cortical organization by the septin cytoskeleton is essential for structural and mechanical integrity of mammalian spermatozoa [J]. *Dev Cell*, 2005, 8 (3): 343-352.
- [23] 张桂征, 张雨丽, 费美华, 等. 家蚕模式在人类疾病与医药研究中的应用[J]. *广西蚕业*, 2011, 48(4): 35-42.