

一条大鼠睾丸新基因的生物信息学分析及真核表达 *

袁莉刚 刘 英 黄布敏 王 睿 曲亚玲

(甘肃农业大学动物医学院 甘肃 兰州 730070)

摘要 目的 探讨大鼠睾丸组织一条新基因的生物信息学特征和真核表达。方法 构建 pEGFP-N1 载体的融合质粒进行真核表达, 利用生物信息学手段分析基因和蛋白功能。结果 生物信息学分析表明 RSA14-44 的编码区序列与人类及鼠源 RAS 同源基因家族核酸序列达到 85%以上的同源性, RSA14-44 蛋白没有典型的跨膜结构域, 也没有典型的 N 末端信号肽; 与人类 RhoA 蛋白序列达到了 89%的同源且具有 Rho 家族成员的 GAAX 盒和 p-loop 结构的基序特征, RSA14-44 蛋白大部分氨基酸序列与 Rho 家族 7 个已知结构域高度同源, RSA14-44 基因真核表达定位于细胞质。结论: RSA14-44 基因真核表达定位于 CHO-K1 细胞质, 编码蛋白质与 Rho 家族同源性高, 为进一步研究其生物学功能提供参考。

关键词 克隆;真核表达;RSA-14-44 基因;生物信息学

中图分类号 Q811.4 Q786 **文献标识码** A **文章编号** :1673-6273(2011)05-819-05

Bioinformatic Analysis and Eukaryotic Expression of a New Rat Testicular cDNA Gene*

YUAN Li-gang, LIU Ying, HUANG Bu-min, WANG Rui, QU Ya-ling

(College of Veterinary Medicine, Gansu Agricultural University, Lanzhou, Gansu, 730070, China)

ABSTRACT Objective: To investigate the bioinformatic analysis and eukaryotic expression of a new rat testicular cDNA gene. **Methods:** The gene was amplified by PCR from genomic DNA of rat testis Genes and was eukaryotic expressed by a recombinant expression plasmid pEGFP-N1-14,the structure and the function of cloning genes were predicted by bioinformatics. **Results:** The homologue of the RSA14-44 gene between the rat and human chromosome RAS family DNA sequence was 85% with and the homologue of this gene coding protein was 89% with human family RhoA. There were typically GAAX box and p-loop sequence at the coding region;There were no typical membrane spaning domain and terminal signal peptide. However, the RSA14-44 amino acid sequence had high homologous with the seven known structural domain of family Rho.The localization for RSA14-44 in CHO cells was observed in cytoplasm. **Conclusion:** The coding protein of rat gene was localized in cytoplasm and it was high homologous with Rho family.

Key words: Cloning; Prokaryotic expression; RSA-14-44 gene; Bioinformatics

Chinese Library Classification (CLC): Q811.4 Q786 **Document code:** A

Article ID: 1673-6273(2011)05-819-05

前言

随着分子生物学和蛋白质组学的发展,利用分子生物学和生物信息学的手段研究新基因以及预测蛋白质的功能成为研究手段之一^[1,2]。现在越来越多编码精子相关蛋白的基因已经克隆、鉴定,生精细胞相关基因中有一类是既在生精细胞中表达又在体细胞中表达,如血管紧张素转换酶基因等^[3],由于基因的转录或转录后 mRNA 加工过程的差异,生精细胞中产生的 mRNA 在大小和顺序上与体细胞 mRNA 有所差异^[4,5];来源于大鼠睾丸组织初级精母细胞中的 RSA14-44 新基因(Genbank 登记号 AY149343),位于大鼠染色体 4q31,全长为 1124bp,可读编码框自 100bp_684bp,编码 194 个氨基酸,前期通过原核表达和酵母双杂交研究初步探索了基因功能^[6,7],本研究首先利用生物信息学技术分析克隆基因结构及蛋白功能,推测其分子

生物学特征,然后通过基因克隆构建融合质粒进行真核表达,为进一步研究其蛋白作用特点和调控机制提供了新思路。

1 材料与方法

1.1 实验材料

E.coli BL21(DE3)菌、E.coliDH5α 菌、pGEM-T 载体系统购自 Promega 公司 pEGFP-N1 购自 Clontech 公司,各种限制性内切酶、T4 DNA 连接酶和 DL2000 DNA Marker 购自大连宝生物有限公司, DNA 快速纯化回收试剂盒购自博大生物公司, 转染级质粒的提取采用 GIBCO 公司的 CONCERT 高纯度质粒纯化系统。

CHO 细胞系,DMEM 细胞培养液,胎牛血清、胰酶(GIBCO 公司),脂质体 LipofectamineTM2000(Invitrogen 公司);常规试剂均为分析纯。

* 基金项目:甘肃农业大学创新基金项目(GAU-CX0513)

作者简介:袁莉刚(1974-),女,副教授,博士,主要研究方向:动物发育生物学。

电话:0931-7631229 E-mail: yuan2918@126.com

(收稿日期:2010-12-03 接受日期:2010-12-28)

1.2 实验方法

1.2.1 大鼠基因组总 DNA 的提取 Wistar 大鼠睾丸基因组 DNA 提取 按《精编分子生物学实验指南》说明进行。

1.2.2 RSA14-44 真核表达质粒的构建 PCR 扩增 RSA14-44 的编码区序列,引物设计时在上下游引物的 5' 端和 3' 端分别加上 BamH 和 Xho 位点,引物设计 :1)5'-AAGCTT ATG AAGTCCCAGGTGCTTTGTACAGC - 3' 2)5' -GGATCC TTCACACGGTTCGTTTGCAAG - 3'。将扩增后的 PCR 产物首先连接到 pGEM-T 载体上以 BamH / Xho 双酶切鉴定后,经琼脂糖电泳和切胶回收后连接到经相同位点酶切的 pEGFP-N1 载体中,将连接产物转化 DH5α 感受态细胞,提取质粒经双酶切鉴定正确后测序。

1.2.3 CHO-K1 细胞的培养和转染 CHO 细胞的培养用含 10 % 胎牛血清的 DMEM 细胞培养液在 37 ℃ 5% CO₂ 及一定湿度条件下常规培养 2-3 天传代一次。转染前一天,以胰酶消化收获对数期生长的细胞,以 0.5~1.0× 10⁶ 重新接种于 6 孔板中,同时放入处理好的盖玻片,培养细胞至 90-95 % 的饱和度。采用脂质体介导的真核细胞转染 (LIPOFECTAMINETM 2000 Reagent) 用 250 μl 无抗生素无血清的 DMEM 分别稀释 5 μg 质粒和 5 μl 的 Lipofectamine2000,室温静置 2 min。将两种稀释液轻轻混合,室温放置 20 min。将混合液加至对应的孔板中,置于 37 ℃ 5 % CO₂ 条件下培养 6 小时后移去原来的培养液,更换成含 10 % 胎牛血清的培养液,转染后 48 小时在荧光显微镜下观察。

1.2.4 细胞固定、透化和复染 转染后 48h 用 PBS 清洗细胞两次,4 % 多聚甲醛固定液室温固定 10-15 min, PBS 漂洗 3 次,每次 5 min,细胞透化液于室温透化处理 10 min, PBS 漂洗 3 次,每次 5 min。透化处理好的细胞培养板中加入以 PBS 配制的 RNA 酶(100 μg/ml)溶液,于 37 ℃ 保温 30 min, PBS 漂洗 2 次,每次 1 min,碘化丙啶(2 μg/ml)染色 5-10min。PBS 漂洗 3 次,每次 5 min,去离子水漂洗去盐,封片后激光共聚焦显微镜

(LEICA, TCS NT)观察并照相。

1.3 生物信息学分析软件

瑞典生物信息研究所 ExPASy(Expert Protein Analysis System)蛋白分析系统和美国国家生物技术信息中心 Domain 分析工具 核酸蛋白序列分析软件 DNAMAN。

2 结果与分析

2.1 生物信息学分析

2.1.1 RSA14-44 的基因和蛋白序列 BLAST 结果 用 BLAST 搜索 GenBank 数据库,对 RSA14-44 进行核酸和蛋白质序列同源性比较, RSA14-44 的编码区序列与人类以及鼠源 RAS 同源基因家族成员 A (RHOA) 核酸序列达到 85% 以上同源; RSA14-44 蛋白与人类 GTP 连接蛋白 RhoA 蛋白序列达到了 89% 的同源。

2.1.2 RSA14-44 基因编码蛋白的亲疏水性分析 利用瑞典生物信息研究所 ExPASy(Expert Protein Analysis System)蛋白分析系统中 Scanprosite 工具对 RSA14-44 进行亲疏水性分析, RSA14-44 基因编码的蛋白质没有典型的跨膜结构(图 1);用推测的 RSA14-44 克隆的氨基酸序列搜索 Prosite 数据库后,发现其具有 2 个 ASN 糖基化位点、3 个 PKC 磷酸化位点、5 个

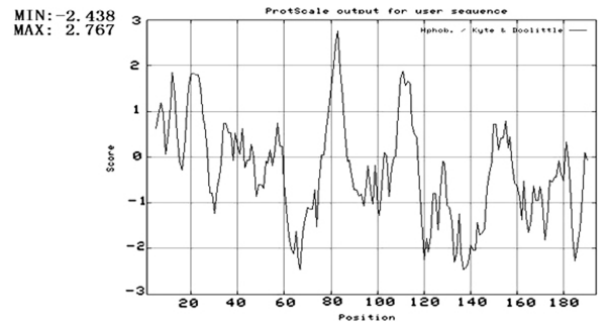


图 1 RSA14-44 基因编码蛋白的亲疏水性分析

Fig. 1 RSA14-44 gene encodes protein water-based affinities analysis

表 1 Prosite 分析 RSA14-44 基因编码蛋白所含可能的功能基序

Table 1 Analysis the functional motif of RSA14-44 gene encodes protein by Prosite tool

基序 Motif	匹配数 Numbers	匹配位点 Site
N- 糖基化位点 PDOC00001 PS00001 ASN_GLYCOSYLATION	2	69 - 72 NNST 70 - 73 NSTD
蛋白激酶 C 磷酸化位点 PDOC00005 PS00005 PKC_PHOSPHO_SITE	3	18 - 20 SIR 72 - 74 TdR 166 - 168 SeK
酪蛋白激酶 II 磷酸化位点 PDOC00006 PS00006 CK2_PHOSPHO_SITE	5	11 - 14 TmnE 55-58 SsiE 62 - 65 SymE 106 - 109 SIIE 137 - 140 ScdE
酪氨酸激酶磷酸化位点 PDOC00007 PS00007 TYR_PHOSPHO_SITE	1	20 - 26 Rsi.Eqk.Y
N- 烷基化位点 PDOC00008 PS00008MYRISTYL	2	100 - 105 GTpvTN 146 - 151 GGvvSL
异戊(间)二烯基位点 (CAAX box) PDOC00266 PS00294 PRENYLATION	1	184 - 187 CFLg

CK2 磷酸化位点、1 个 TYR 磷酸化位点、2 个十四烷基化位点和 1 个异戊(间)二烯基位点(表 1)。

2.1.3 RSA14-44 基因编码蛋白的结构域分析 利用美国国家生

物技术信息中心 Domain 分析工具进行结构域分析，发现 RSA14-44 基因编码蛋白大部分氨基酸序列与 7 个已知结构域高度同源(表 2 图 2)。

表 2 Prosite 分析 RSA14-44 基因编码蛋白所含结构域

Table 2 Analysis the functional domain of RSA14-44 gene encodes protein by Prosite tool

结构域 Structure field	特征 Characteristics
Rho	RAS 类小三磷酸鸟苷酸酶(GTPase)亚家族 ,包括 RhoA、cdc42、Rac 亚单位。参与肌动蛋白重组 ;在胞质分裂、粘附斑形成以及应力激活激酶受到刺激时参与调节细胞的转化。 GTPase sub-family, includ subunits of RhoA, cdc42 and Rac. Participate in actin recombination; participate in regulating cell transformation in cytokinesis adhere stain orming and when activate kinase stimulated.
RAB	小 GTPases 的 Rab 亚家族 ,在囊泡运输中发挥着重要作用。Rab 家族成员高度保守 ;主要定位于细胞内各细胞器的膜表面 ,在膜运输的各个环节中发挥作用。 Rab sub-family of little GTPases, plays vital function in juice bearing body transporting. Rab family member are high conservative; mostly locate at membrane surface in each organelle, play roles in every links of membrane transporting.
smart00174 RHO	Rho 属于 Ras 小 GTP 酶亚家族 ,包括 Cdc42 和 Rac 及 Rho 亚型。 Rho belongs to Ras little GTP enzyme sub-family, includ Cdc42, Rac and Rho subset.
smart00173 RAS	RAS 小 GTPases 的 Ras 亚家族成员 ,在结构和功能上与细菌 p21Ras 偶联受体酪氨酸激酶以及偶联到 G 蛋白受体上的蛋白激酶相似。RAS little GTPases Ras sub-family member, similitude with bacteria p21Ras couple receptor tyrosine kinase and protein kinase of couple G albumen receptor on structure and functions.
smart00176 RAN	Ran (Ras- 相关核蛋白) /TC4(Ras 亚家族成员) ;Ran 主要在蛋白通过核孔时的主动运输有关 Ras 家族 ,包括 Rab ,Rac ,Ral ,Ran ,Rap Ypt1 等亚家族 ,与 GTP_EFTU arf 以及 myosin 等都具有 P-loop 基序结构域。
pfam00071 Ras	Ras family, includ Rab, Rac, Ral, Ran, Rap Ypt1 sub-family and so on, has P - loop structure field, the same with GTP_EFTU, arf and myosin etc.
pfam00025	ADP- 核糖基化因子家族 ADP - ribosylation factor family

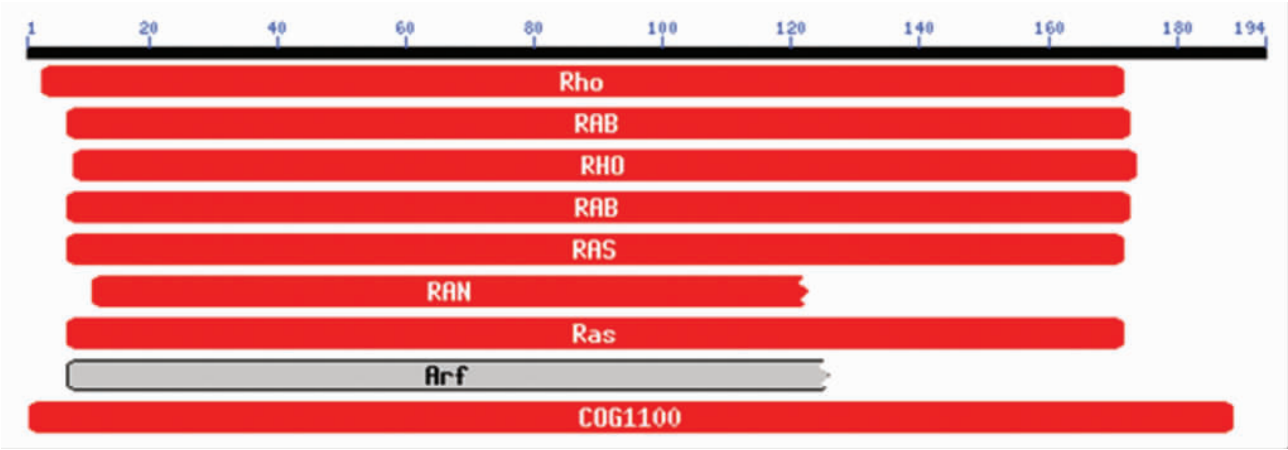


图 2 RSA14-44 基因编码蛋白的相关结构域分析

Fig.2 Analysis the domain of RSA14-44 gene encodes protein by Domain tool

2.2 RSA14-44 在真核细胞中的表达定位

2.2.1 pEGFP-N1-14 荧光表达载体的构建 PCR 扩增 RSA14-44

的编码区序列(584bp) ,引物设计时在上下游引物的 5' 端的分
别加上 Xho 和 BamH 位点进行扩增得到 PCR 产物(图 3) ,

首先克隆到 pGEM-T 载体上,然后分别以 BamH /Xho 双酶切,并经琼脂糖电泳和切胶回收后(图 4),连接到经相同酶切的增强绿色荧光表达载体 pEGFP-N1。将连接产物转化 DH5α

感受态细胞;重组质粒进行 BamH /Xho 的双酶切鉴定,将克隆质粒提交上海博亚生物公司测序,测序结果证实成功构建了 pEGFP-N1-14 质粒。

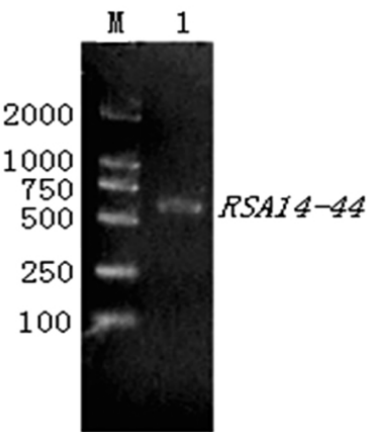


图 3 RSA14-44cDNA 扩增的电泳结果 :M :DL2000 Marker ;1 : RSA14-44 PCR 扩增产物

Fig.3 RT-PCR products of RSA14-44 cDNA :M: DL2000 Marker; 1: Products of RSA14-44 cDNA

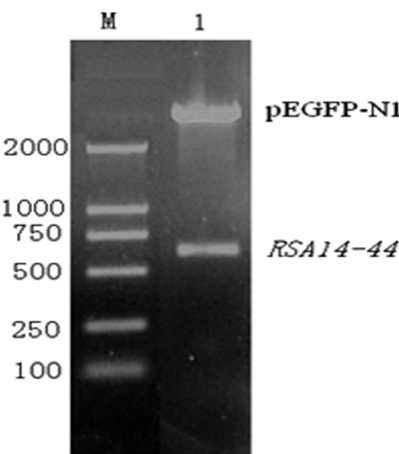


图 4 pEGFP-N1-14 cDNA 双酶切电泳结果 M :DL2000 Marker ;1 : pEGFP-N1-14 双酶切结果

Fig.4 Endonuclease digestion identification of the recombinant plasmid pEGFP-N1-14 cDNA :M: DL2000 Marker; 1: Product from pEGFP-N1-14 cDNA digested by BamH and Xho

2.2.2 pEGFP-N1-14 在 CHO-K1 细胞中的定位 结果显示 , pEGFP-N1-14(图 5-A)在细胞质中相对富集 ,与经碘化丙啶染

色的图像(图 5-B)进行叠加后(图 5-C)表明 pEGFP-N1-14 主要定位于胞质。

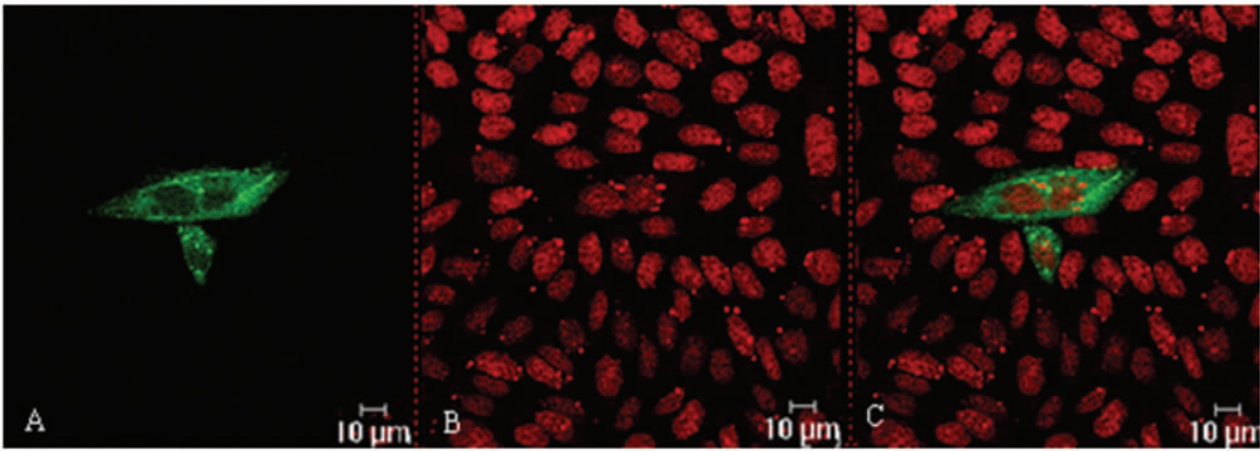


图 5 pEGFP-N1-14 瞬时表达蛋白在 CHO 细胞中的定位(Bar = 10μ m) :A :pEGFP-N1-14 蛋白在 CHO 细胞中的瞬时表达 ;B :碘化丙啶染色的 CHO 细胞 ;C :两张图片的叠加结果

Fig 5 Protein localization of transient expression of pEGFP-N1-14 in CHO-K1 cells(Bar = 10μ m) :A:Transient expression of pEGFP-N1-14 in CHO-K1 cells; B: stained by PI; C: overlay of A and B picture

3 讨论

本研究选择大鼠睾丸组织初级精母细胞中的一条名为 RSA14-44 的新基因(Genbank 登记号 AY149343)进行研究 ;由于该基因的功能未知,因此首先通过生物信息学的手段对 RSA14-44 功能进行了推测。RSA14-44 基因核苷酸及其编码的蛋白质的同源性分析表明其与人类的 Rho-GTP (Ras homolog gene)基因家族有高度的相关性。功能基序分析 RSA14-44 蛋白显示有 2 个 ASN 糖基化位点、3 个 PKC 磷酸化位点、5 个 CK2

磷酸化位点、1 个 TYR 磷酸化位点、2 个十四烷基化位点和 1 个异戊(间)二烯基位点。这些位点与 Rho 家族成员基序特征相似,尤其是氨基酸序列末端的异戊(间)二烯基位点是一个特殊有 GAAX 盒,它是 Rho 家族蛋白在 GTP 与 GDP 交换中的主要识别信号^[8,9]。本研究中 RSA14-44 的 C 末端不仅具有 CAAX 基序,而且还有一个在 Rho 家族中相当于保守的 p-loop 结构,该结构域能够结合 GTP 或 ATP,并具有相应的 GTPase 或 ATPase 活性^[10,11]。利用美国国家生物技术信息中心 Domain 分析工具进行结

构域分析发现 RSA14-44 基因编码蛋白大部分氨基酸序列与 7 个已知结构域高度同源 Rho 结构域、Rab 结构域、Smart00174、smart00173、smart00176、pfam00071、pfam00025。Rho 家族蛋白是 Ras 超家族的成员,具有 GTP 酶活性,在真核细胞信号传递过程中发挥分子开关的作用^[12]。RhoGTP 酶是从酵母到人类的所有真核细胞的肌动蛋白形成细胞骨架的主要调节者^[13,14];并且调节神经元发育和重塑^[15-17],影响基因表达、细胞增殖和凋亡^[18,19]等。蛋白的结构域决定其功能,在本研究中,RSA14-44 是从初级精母细胞得到的一条基因,生物信息学提示其与 Rho 基因家族高度相关,可以此为起点可对其功能进行深入探索,初级精母细胞向精子细胞发育过程中,也伴随细胞凋亡和细胞骨架的重组变化^[20];由此可揭示这个基因在精子细胞形态变化过程,甚至在神经细胞发育及其可塑性等方面的潜在作用。

亲疏水性分析结果表明 RSA14-44 蛋白没有典型的跨膜结构域,也没有典型的 N 末端信号肽,这表明这个蛋白可能分布于细胞质。本研究中 RSA14-44 在真核细胞中的荧光载体表达定位于细胞质中,即是对这种推测的验证,也证实 RSA14-44 蛋白既在生精细胞中表达又在体细胞中表达。研究发现 Rab 家族成员高度保守,主要定位于细胞内各细胞器的膜表面,在膜运输各个环节中发挥作用^[21],提示可进一步研究 RSA14-44 蛋白在不同细胞器的定位以探讨其功能。

本研究通过生物信息学分析和蛋白的真核表达,初步认为 RSA14-44 蛋白与 Rho 蛋白的功能相似,为进一步研究其功能提供了参考。

参考文献(References)

- [1] Storf S, Pfeiffer F, Dilks K, et al. Mutational and bioinformatic analysis of haloarchaeal lipobox-containing proteins. [J] *Bioinformatics*, 2010, 4(7):331-7
- [2] Biswas S, Raoult D, Rolain JM. A bioinformatic approach to understanding antibiotic resistance in intracellular bacteria through whole genome analysis. [J] *Int J Antimicrob Agents*, 2008; 32 (3): 207-20
- [3] Howard TE, Shai SY, Langford KG, et al. Transcription of testicular angiotensin-converting enzyme (ACE) is initiated within the 12th intron of the somatic ACE gene [J]. *Mol Cell Biol*, 1990, 10 (8): 4294-302
- [4] Lalancette C, Miller D, Li Y, et al. Paternal contributions: new functional insights for spermatozoal RNA [J]. *J Cell Biochem*, 2008, 104(5):1570-9
- [5] Kleene KC. Patterns, mechanisms, and functions of translation regulation in mammalian spermatogenic cells. [J] *Cytogenet Genome Res*, 2003; 103(3-4):217-24
- [6] 袁莉刚, 田永强, 刘英, 等. 新基因 RSA14-44 在大肠杆菌中的表达及抗体制备[J] *甘肃农业大学学报*, 2004, 5(39): 499-502
YUAN Li-gang, TIAN Yong-qiang, LIU Ying, et al. The Fusion

- Expression, Purification of RSA14-44 and Preparation of Its Polyclonal Antibody [J]. *Journal of GANSU Agricultural University*, 2004, 5(39):499-502
- [7] 袁莉刚, 田永强, 刘英, 等. 酵母双杂交筛选与 RSA14-44 基因相互作用的蛋白[J] *甘肃农业大学学报*, 2004, 6(39): 679-682
YUAN Li-gang, TIAN Yong-qiang, LIU Ying, et al. Screening of the protein interacting with RSA-14-44 using yeast two hybrid method [J]. *Journal of GANSU Agricultural University*, 2004, 6(39):679-682
 - [8] Roberts PJ, Mitin N, Keller PJ, et al. Rho Family GTPase modification and dependence on CAAX motif-signaled posttranslational modification. [J] *J Biol Chem*, 2008, 283(37):25150-63
 - [9] Michaelson D, Ali W, Chiu VK, et al. Postprenylation CAAX processing is required for proper localization of Ras but not Rho GTPases[J]. *Mol Biol Cell*, 2005, 16(4):1606-16
 - [10] Saraste M, Sibbald PR, Wittinghofer A. The P-loop-a common motif in ATP-and GTP-binding proteins [J]. *Trends Biochem Sci*, 1990, 15 (11):430-4
 - [11] Blouin C, Butt D, Roger AJ. Rapid evolution in conformational space: a study of loop regions in a ubiquitous GTP binding domain [J]. *Protein Sci*, 2004, 13(3):608-16
 - [12] Hall A, Nobes CD. Rho GTPases: molecular switches that control the organization and dynamics of the actin cytoskeleton [J]. *Philos Trans R Soc Lond B Biol Sci*, 2000, 355(1399):965-70
 - [13] Caron E, Hall A. Identification of two distinct mechanisms of phagocytosis controlled by different Rho GTPases [J]. *Science*, 1998, 282(5394):1717-21
 - [14] Fukata Y, Amano M, Kaibuchi K. Rho-Rho-kinase pathway in smooth muscle contraction and cytoskeletal reorganization of non-muscle cells[J]. *Trends Pharmacol Sci*, 2001, 22(1):32-9
 - [15] Koh CG. Rho GTPases and their regulators in neuronal functions and development[J]. *Neurosignals*, 2007, 15(5):228-37
 - [16] Curtis I de. Functions of Rac GTPases during neuronal development [J]. *Neurosci*, 2008, 30(1-3):47-58
 - [17] Yoshihara Y, De Roo M, Muller D. Dendritic spine formation and stabilization[J]. *Curr Opin Neurobiol*, 2009, 19(2):146-53
 - [18] Kardassis D, Murphy C, Fotsis T, et al. Control of transforming growth factor beta signal transduction by small GTPases [J]. *FEBS J*, 2009, 276(11):2947-65
 - [19] Hall A. Rho GTPases and the actin cytoskeleton[J]. *Science*, 1998, 279 (5350):509-14
 - [20] Amlani S, Vogl AW. Changes in the distribution of microtubules and intermediate filaments in mammalian Sertoli cells during spermatogenesis[J]. *Anat Rec*, 1988, 220(2):143-60
 - [21] Gurkan C, Koulov AV, Balch WE. An evolutionary perspective on eukaryotic membrane trafficking [J]. *Adv Exp Med Biol*, 2007, 607: 73-83