DOI:10.15886/j.cnki.rdswxb.2011.04.019

文章编号:1674 - 7054(2011)04 - 0349 - 06

水牛 *Cdc*42 cDNA 的克隆与生物信息学分析

张晓茹[1,2],匡文华[1,2],成 鹰[1],杜 丽[1],雷 明[1],
焦寒伟[1],张冬琳[1],郝永昌[1],祁 超[2],王凤阳[1]
(1. 海南大学 农学院/海南省热带动物繁育与疫病研究重点实验室(筹)/海口市动物基因工程重点实验室,
海南 海口 570228; 2. 华中师范大学 生命科学学院,湖北 武汉 430079)

摘 要: 通过 RT-PCR 扩增得到水牛细胞分裂周期蛋白 42(*Cdc*42) cDNA 序列,将扩增产物与 pMD20-T 载体连接,重组质粒经 PCR、酶切鉴定后测序并进行生物信息学分析。结果表明:克隆得到的序列全长 626 bp,含有 1 个大小为 576 bp 的开放阅读框,共编码 191 个氨基酸,相对分子质量为 21.3×10^3,理论等电点为 6.16;*Cdc*42 ORF 序列与 Genbank 上登载的黄牛、狗、小鼠和人的核苷酸序列同源性分别为 99.65%、96.70%、94.97% 和 94.10%,具有很高的相似性;该蛋白为非跨膜蛋白,存在 1 个 RHO 结构域。

关键词: *Cdc*42; 水牛; 基因克隆; 生物信息学分析
中图分类号: S 823.8+3 **文献标志码:** A

细胞分裂周期蛋白 42(Cell division cycle 42,*Cdc*42)是一种三磷酸鸟苷(GTP)结合蛋白,具有 GTP 酶活性,于酵母中首次发现[1],在功能上参与极性细胞的生长及萌芽位点的建立[2]。*Cdc*42 属于 Ras 家族内 Rho 亚家族成员,在活化状态下与 GTP 结合,去活化状态下则与 GDP 结合,在多种细胞功能活动中起分子开关作用。活化状态下的 *Cdc*42 参与肌动蛋白的重组[3]、细胞运动及极性的形成[4]、丝状伪足的形成[5-6]、胞膜运输和细胞内吞作用[7-8]、细胞周期及转录后表达调控[9-10]等生理进程。近年来的研究发现,*Cdc*42 的异常活动会导致许多疾病发生,如心血管病、糖尿病、神经元变性疾病[11];同时 *Cdc*42 还与肿瘤的发生发展有着密切的关系。多项研究发现 *Cdc*42 在肺癌、肝癌、肠癌、乳腺癌、食管癌、膀胱癌等组织及肿瘤细胞中过表达,推测该蛋白可能在肿瘤的增殖、转化、侵袭和转移中发挥作用[12-15]。另外,*Cdc*42 在布鲁氏菌的内化过程中发挥重要作用[16]。兴隆水牛是中国著名的地方水牛品种,产于中国的热带和亚热带地区,具有耐热、抗病力强等特性。迄今,尚未见到兴隆水牛 *Cdc*42 基因的研究报道。笔者采用 RT-PCR 技术,从兴隆水牛外周血白细胞中克隆得到 *Cdc*42 cDNA,并对该基因的序列结构和特征进行了生物信息学分析,同时通过软件对该蛋白质的空间结构进行了预测,旨在为进一步探讨 *Cdc*42 蛋白的生物学功能奠定基础。

1 材料与方法

1.1 材料 兴隆水牛新鲜血液采自海南定安。

1.2 试剂 血液总 RNA 提取试剂盒购自上海华舜生物工程有限公司;RNAiso Reagent、TaKaRa Ex *Taq* 酶、pMD20-T Vector 试剂盒以及 DNA 胶回收试剂盒购自宝生物工程(大连)有限公司;RevertAid™ H Minus First Strand cDNA Synthesis Kit 和限制性内切酶购自 Fermentas 公司;E. Z. N. A. Plasmid Mini Kit Ⅰ 试

收稿日期:2011 - 09 - 28
基金项目:国家转基因生物新品种培育重大专项(2009ZX08007 - 009B)
作者简介:张晓茹(1985 -),女,湖北武汉人,华中师范大学生命科学学院 2009 级硕士研究生.
通信作者:王凤阳(1970 -),男,教授,博导,主要从事动物功能基因组学研究. E-mail:fywang68@hotmail.com

剂盒购自 Omega 公司;大肠杆菌 DH5α 由本实验室保存;其他常规试剂为国产分析纯。

1.3 总 RNA 提取 从兴隆水牛静脉取血 1.5 mL,并注抗凝于离心管中,缓缓加入 5 倍体积冰冷的 w = 0.83% 的 NH_4Cl 溶液,上下翻转离心管数次后置于冰上,其间混匀 3 ~ 5 次,约 15 min 后,混合溶液呈透明状,红细胞已基本被裂解。4 000 r · min^{-1} 室温离心 5 min,收集白细胞沉淀,弃去上清,用 PBS 洗涤白细胞沉淀 2 次后,加入 RNAiso Reagent 后剧烈震荡裂解白细胞,按试剂盒说明提取总 RNA。

1.4 引物设计与合成 根据 Genbank 中登录的黄牛的 *Cdc*42 mRNA(NM_001046332)序列,设计 1 对引物。F1: AAAGTGGATACAAAACTGTTTC; R1: GCAGAAAGGGCTCTGGAG。引物合成及测序由上海生工生物工程技术有限公司完成。

1.5 cDNA 合成与基因克隆 以约 1 μg 的总 RNA 为模板,按试剂盒说明书方法利用 RevertAid™ H Minus First Strand cDNA Synthesis Kit 合成 cDNA 第 1 链。以 2 μL 上述反应液为模板,用引物 F1 和 R1 进行降落 PCR 反应。反应条件为:第 1 个循环 94 ℃预变性 3 min;第 2 个循环开始采用降落 PCR 的方法,退火温度从 66 ℃开始,每个循环减低 2 ℃,变性延伸温度不变,直到退火温度降到 44 ℃;再按 94 ℃ 1 min,44 ℃ 1 min,72 ℃ 1 min 进行 20 个循环的反应;最后,72 ℃延伸 5 min。反应结束后,将全部反应液进行琼脂糖凝胶电泳,回收目的条带。与 pMD20-T Vector 进行 TA 克隆后,转化大肠杆菌 DH5α,经蓝白斑筛选挑取白色单菌落,菌落 PCR 和酶切鉴定筛选出阳性重组克隆后送公司测序。

1.6 生物信息学分析 利用 NCBI 站点提供的相关信息完成基因同源性分析和开放阅读框的识别,核苷酸和氨基酸序列比对和进化树构建应用 DNAMAN 5.2.2 软件完成;利用 ProtParam 程序分析 *Cdc*42 蛋白的氨基酸组成、相对分子质量、等电点等理化性质;利用 SignalP,ProtScale,NetPhos,TMHMM,SMART,HNN,SWISS-MODEL 和 PSORT Ⅱ 网站在线程序对 *Cdc*42 蛋白信号肽、疏水性轮廓、磷酸化位点、跨膜信息、关键结构域、二级结构、三级结构及亚细胞定位进行预测。所用相关分析软件的生物信息学网址如下:

ExPASy-ProtParam: http://web.expasy.org/protparam/

ExPASy-ProtScale: http://web.expasy.org/protscale/

SignalP 3.0 Server: http://www.cbs.dtu.dk/services/SignalP/

NetPhos 2.0 Server: http://www.cbs.dtu.dk/services/NetPhos/

TMHMM Server v. 2.0: http://www.cbs.dtu.dk/services/TMHMM/

SMART: http://smart.embl-heidelberg.de/

HNN: http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_hnn.html

SWISS-MODEL: http://swissmodel.expasy.org/?pid=smh01&uid=&token=

PSORT Ⅱ Prediction: http://psort.hgc.jp/form2.html

2 结果与分析

2.1 水牛 *Cdc*42 cDNA 的克隆 以兴隆水牛外周血白细胞总 RNA 为模板,通过 RT-PCR 扩增出 1 条特异性条带。回收、克隆测序后,显示扩增的片段大小为 626 bp,与预期结果一致,说明所克隆的片段即为 *Cdc*42 目的基因。用限制性内切酶 *Hind*Ⅲ进行单酶切鉴定,结果如图 1 所示。将克隆片段与 pMD20-T 载体连接,挑取阳性克隆,提取重组质粒,用 *EcoR* Ⅰ和 *Xba* Ⅰ进行双酶切鉴定,结果如图 2 所示,表明目的基因 *Cdc*42 已成功插入到载体中。

2.2 测序结果及序列分析 核苷酸序列测序和分析结果表明,*Cdc*42 基因 cDNA 包含 1 个 576 bp 的开放阅读框,其中 A + T 含量为 54.9%,C + G 含量为 45.1%。*Cdc*42 基因核苷酸序列及其推导的氨基酸序列见图 3。ExPASy-ProtParam 软件在线分析结果表明,该基因共编码 191 个氨基酸,蛋白相对分子质量为 21.3×10^3,等电点(pI)为 6.16,酸性氨基酸残基总数(Asp + Glu)为 24,碱性氨基酸残基总数(Arg + Lys + His)为 23,疏水性氨基酸 66 个,极性氨基酸 50 个。将克隆得到的 *Cdc*42 mRNA 序列提交 Genbank(登录号:JN380079)。

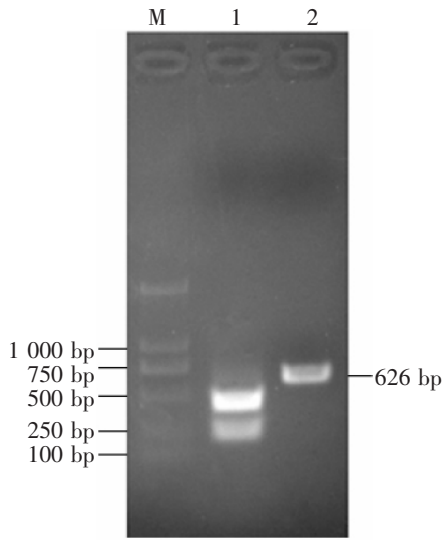

图 1　*Cdc*42 克隆片段及单酶切电泳

M: D2000 DNA marker;

1: *Cdc*42 经 *Hind* Ⅲ酶切产物;

2: *Cdc*42 克隆片段。

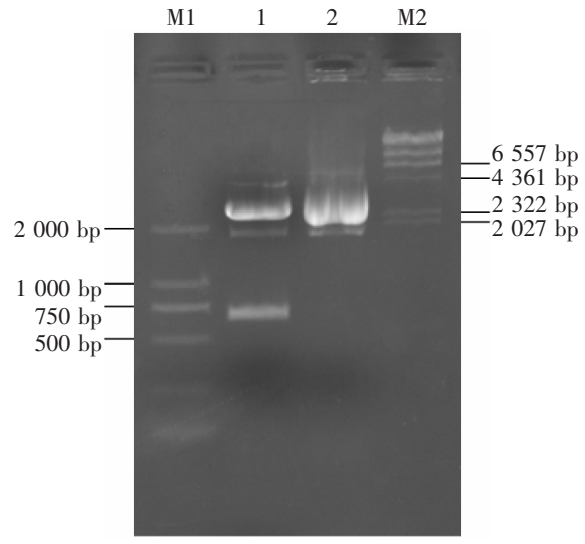

图 2　重组质粒 pMD20-*Cdc*42 双酶切鉴定电泳

M1: D2000 DNA marker;

1: pMD20-TLR2 经 *EcoR* Ⅰ和 *Xba* Ⅰ酶切产物;

2: pMD20-TLR2 质粒; 3: λDNA/*Hind* Ⅲ DNA marker。

```
1          ATGCAGACAATTAAGTGCGTTGTTGTGGGTGATGGTGCCGTTGGTAAAACATGTCTCCTG
1           M  Q  T  I  K  C  V  V  V  G  D  G  A  V  G  K  T  C  L  L
61         ATATCCTACACAACAAACAAATTTCCATCGGAGTATGTACCGACTGTTTTTGACAACTAT
21          I  S  Y  T  T  N  K  F  P  S  E  Y  V  P  T  V  F  D  N  Y
121        GCTGTCACAGTTATGATTGGTGGGGAGCCATATACTCTTGGACTTTTTGATACTGCAGGG
41          A  V  T  V  M  I  G  G  E  P  Y  T  L  G  L  F  D  T  A  G
181        CAAGAGGATTATGACAGATTACGACCGCTGAGTTATCCACAAACAGATGTATTTCTAGTC
61          Q  E  D  Y  D  R  L  R  P  L  S  Y  P  Q  T  D  V  F  L  V
241        TGCTTTTCAGTGGTCTCTCCATCCTCATTCGAAAATGTGAAAGAAAAGTGGGTGCCTGAG
81          C  F  S  V  V  S  P  S  S  F  E  N  V  K  E  K  W  V  P  E
301        ATAACTCACCACTGTCCAAAGACTCCTTTCTTGCTTGTTGGGACCCAAATTGATCTCAGA
101         I  T  H  H  C  P  K  T  P  F  L  L  V  G  T  Q  I  D  L  R
361        GATGACCCCTCTACTATTGAGAAGCTTGCTAAGAATAAGCAGAAGCCTATCACTCCAGAG
121         D  D  P  S  T  I  E  K  L  A  K  N  K  Q  K  P  I  T  P  E
421        ACTGCTGAAAAACTGGCCCGTGACCTGAAGGCTGTCAAGTATGTGGAGTGCTCTGCACTC
141         T  A  E  K  L  A  R  D  L  K  A  V  K  Y  V  E  C  S  A  L
481        ACACAGAAAGGCCTAAAGAATGTATTTGATGAAGCAATATTGGCTGCCCTGGAGCCTCCA
161         T  Q  K  G  L  K  N  V  F  D  E  A  I  L  A  A  L  E  P  P
541        GAACCGAAGAAGAGCCGCAGGTGTGTGCTGCTATGA
181         E  P  K  K  S  R  R  C  V  L  L
```

图 3　*Cdc*42 基因核苷酸测序结果及预测蛋白

2.3　生物信息学分析

2.3.1　*Cdc*42 基因同源性分析　用 DNAMAN 软件将测序所得水牛 *Cdc*42 ORF 序列与 Genbank 中登陆的黄牛、人、狗、小鼠等动物的 ORF 序列进行比对，并利用其中的进化树分析程序进行了进化树分析(图 4)，发现兴隆水牛与黄牛(NM_001046332)的同源性最高，为 99.65%，与狗(NM_001003254)、小鼠(MMU37720)和褐鼠(NM_171994)的同源性较高，分别为 96.70%、94.97%和 94.10%；与人(AF498962)的同源性次之，为 91.15%；与原鸡(NM_205048)的同源性最低，为 88.19%。氨基酸同源性比对发现，水牛 *Cdc*42 蛋白与黄牛、狗、小鼠和褐鼠的同源性均为 100%，与人的为 95.29%，与原鸡的为 99.48%，说明 *Cdc*42 在进化过程中具有较高的保守性。

2.3.2　*Cdc*42 蛋白的特征分析　利用 TMHMM 程序预测 *Cdc*42 蛋白的跨膜结构，发现该蛋白无跨膜区，

说明该蛋白不是跨膜蛋白。利用 SignalP 3.0 Server-prediction 软件分析，发现该蛋白无信号肽出现。利用 NetPhos 2.0 Server 对蛋白序列的磷酸化位点进行预测，发现水牛 *Cdc*42 蛋白结构中磷酸化修饰位点共 13 个，其中丝氨酸蛋白激酶磷酸化位点 7 个，苏氨酸蛋白激酶位点 3 个，酪氨酸蛋白激酶磷酸化位点 3 个。利用网上在线工具 ExPASy-ProtScale 程序预测该蛋白的疏水性，结果如图 5 所示。

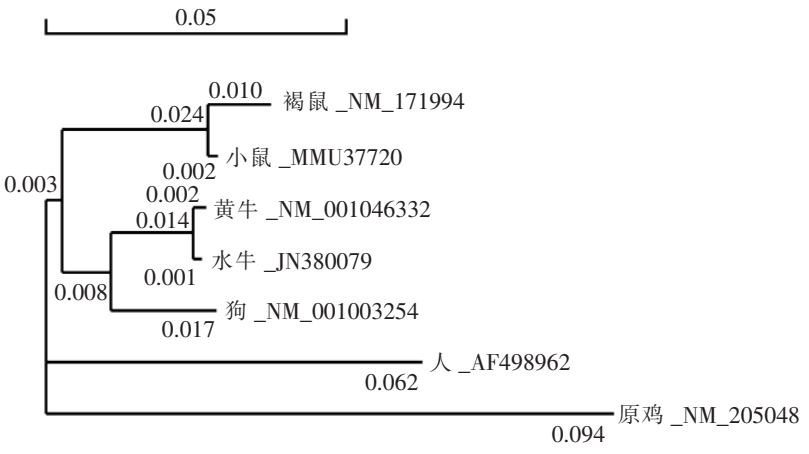

图 4 水牛 *Cdc*42 ORF 序列与其他物种间的遗传进化树

*Cdc*42 蛋白的疏水性最大值为 3.100，最小值为 −2.878，在氨基酸序列的 76 ~ 87 这个区域出现了较高的疏水性，表明这个区域富含疏水性氨基酸。

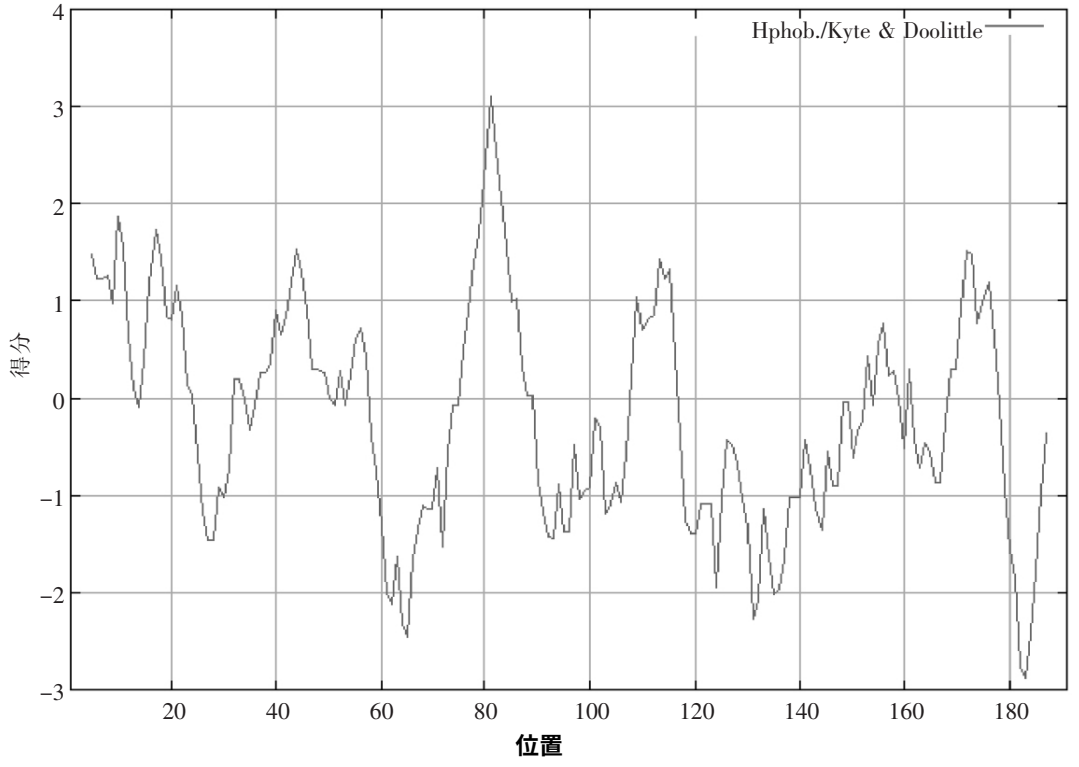

图 5 水牛 *Cdc*42 蛋白疏水性分析结果

2.3.3 *Cdc*42 蛋白的结构域分析 利用 SMART 程序预测 *Cdc*42 蛋白的关键结构域，发现该蛋白含有 1 个 RHO(Ras homology) 结构域，由 6 ~ 179 号氨基酸组成。利用 HNN 软件预测该蛋白的二级结构，结果如图 6 所示，该蛋白二级结构中 α 螺旋(Alpha helix) 占 27.23%，由 52 个氨基酸组成; 伸展结构(Extended strand) 占 25.65%，由 49 个氨基酸组成; 无规则卷曲(Random coil) 占 47.12%，由 90 个氨基酸组成。利用网上在线工具 SWISS-MODEL 分析了该蛋白的三维结构，结果如图 7 所示。

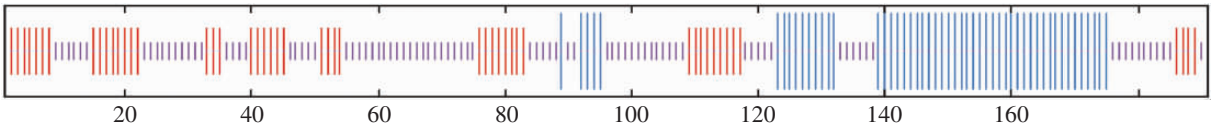

图 6 水牛 *Cdc*42 蛋白二级结构预测结果

注: 最长竖线表示 α 螺旋; 中长竖线表示伸展结构; 最短竖线表示无规则卷曲。

2.3.4　*Cdc*42 蛋白的亚细胞定位预测　利用网上在线工具 PSORT Ⅱ预测了水牛 *Cdc*42 蛋白的亚细胞定位情况,结果发现,该蛋白可能位于多个细胞器中,位于细胞质的概率为 39.1%,位于细胞核的概率为 17.4%,位于分泌系统小囊泡和细胞质膜的概率是 13.0%,位于高尔基体的概率为 8.7%,位于线粒体的概率为 4.3%。

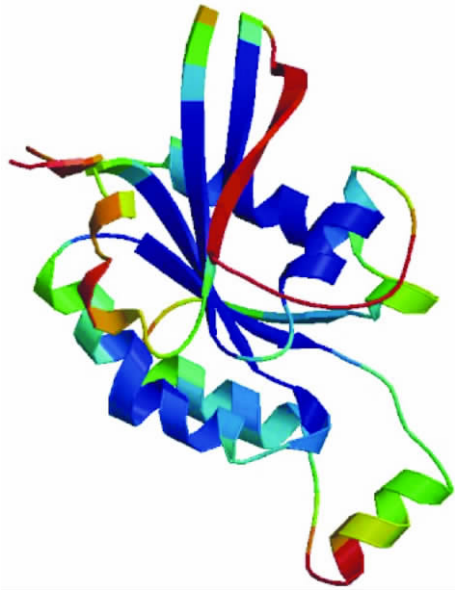

图 7　水牛 *Cdc*42 蛋白的三维结构预测

3　讨　论

*Cdc*42 作为多种细胞信号通路中的开关分子,其生物活性十分广泛,在细胞生长、分化、程序性死亡、细胞极性的建立等方面起重要作用。在 *Cdc*42 细胞信号通路中,*Cdc*42 与鸟嘌呤核苷酸交换因子(guanine nucleotide exchange factor, GEF)结合形成蛋白复合物,通过复合物上的 DH(Db1 homology)结构域或 DHR2(DOCK180 homology region)结构域将 *Cdc*42 定位于不同组织的细胞骨架,细胞外膜(细胞核外膜或液泡外膜),高尔基体复合物等中,从而形成不同的细胞信号通路,行使不同的细胞功能。近年来 *Cdc*42 与各种疾病和肿瘤易感性之间的关系更是深入研究的热点,其生物学活性与肿瘤发生、侵袭和转移、细胞周期进展及转录的调节、细胞的分泌和凋亡、细菌的内化等的关系正在逐渐明确,可见 *Cdc*42 在各种生命活动中的重要性。

笔者克隆的水牛 *Cdc*42 cDNA 大小为 626 bp,含 1 个大小为 576 bp 的开放阅读框,共编码 191 个氨基酸,相对分子质量为 21.3×10^3。序列分析发现,*Cdc*42 ORF 序列与 Genbank 上登载的黄牛核苷酸序列同源性为 99.65%,说明本实验成功克隆了水牛 *Cdc*42 cDNA。其 ORF 序列与狗、小鼠、人等物种的核苷酸序列同源性在 90% 以上,具有很高的相似性,这说明 *Cdc*42 在生物进化过程中具有较强的保守性。笔者还对兴隆水牛 *Cdc*42 基因进行了蛋白质理化性质分析及结构预测,为进一步研究 *Cdc*42 基因在水牛细菌性感染的作用机理研究提供了基础材料。

参考文献:

[1] JOHNSON D I, PRINGLE J R. Molecular characterization of *Cdc*42, a Saccharomyces cerevisiae gene involved in the development of cell polarity [J]. J Cell Biol,1990,111(1):143 – 152.

[2] JOHNSON D I. *Cdc*42: An essential Rho-type GTPase controlling eukaryotic cell polarity [J]. Microbiol Mol Biol Rev,1999,63(1):54 – 105.

[3] RIDLEY A J, HALL A. The small GTP-binding protein rho regulates the assembly of focal adhesions and actin stress fibers in response to growth factors [J]. Cell,1992,70(3):389 – 399.

[4] ATWOOD S X, CHABU C, PENKERT R R, et al. *Cdc*42 acts downstream of Bazooka to regulate neuroblast polarity through Par-6 aPKC. [J]. J Cell Sci,2007,120(18):3200 – 3206.

[5] NOBES C D, HALL A. Rho,rac ,and *Cdc*42 GTPases regulate the assembly of multimolecular focal complexes associated with actin stress fibers, lamellipodia, and filopodia [J]. Cell,1995,81(1):53 – 62.

[6] KOZMA R, AHMED S, BEST A, et al. The Ras-related protein *Cdc*42Hs and bradykinin promote formation of peripheral actin microspikes and filopodia in Swiss 3T3 fibroblasts [J]. Mol Cell Biol,1995,15(4):1942 – 1952.

[7] LUNA A, MATAS O B, MARTINER-Menárguez J A, et al. Regulation of protein transport from the Golgi complex to the endoplasmic reticulum by *Cdc*42 and N-WASP. [J]. Mol Biol Cell,2002,13(3):866 – 879.

[8] RODAL A A, MOTOLA-BARNES R N, LITTLETON J T. Nervous wreck and *Cdc*42 cooperate to regulate endocytic actin assembly during synaptic growth [J]. J Neurosci,2008,28(33):8316 – 8325.

[9] SPURING C C, GODMAN C A, NOONAN E J, et al. HDAC3 overexpression and colon cancer cell proliferation and differen-

tiation [J]. Mol Carcinog,2008,47(2):137 – 147.

[10] COSO O A, CHIARIELLO M, YU J C, et al. The small GTP-binding proteins Rac1 and *Cdc*42 regulate the activity of the JNK/SAPK signaling Pathway [J]. Cell,1995,81(7):1137 – 1146.

[11] ROSSMAN K L, DER C J, SONDEK J. GEF means go: turning on RHO GTPases with guanine nucleotide-exchange factors [J]. Nat Rev Mol Cell Biol,2005,6(2):167 – 180.

[12] CUIYAN Z, JIE H, FANG Z, et al. Overexpression of RhoE in Non-small Cell Lung Cancer (NSCLC) is associated with smoking and correlates with DNA copy number changes [J]. Cancer Biol Ther,2007,6(3):335 – 342.

[13] GRISE F, BIDAUD A, MOREAU V. Rho GTPases in hepatocellular carcinoma [J]. Biochim Biophys Acta,2009,1795: 137 – 151.

[14] FENG J G, LIU Q, QIN X, et al. Clinicopathological pattern and Annexin A2 and *Cdc*42 status in patients presenting with differentiation and lymphnode metastasis of esophageal squamous cell carcinomas [J/OL]. Mol Biol Rep,2011,21: DOI: 10.1007/s11033 – 011 – 0859 – 2.

[15] STENGEL K, ZHENG Y. *Cdc*42 in oncogenic transformation, invasion, and tumorigenesis [J]. Cell Signal,2011,23(9): 1415 – 1423.

[16] GUZMÁN-VERRI C, CHAVES-OLARTE E, VONEICHEL-STREIHER C, et al. GTPases of the Rho subfamily are required for Brucella abortus internalization in non-professional phagocytes: direct activation of *Cdc*42 [J]. J Biol Chem,2001,276: 44435 – 44443.

Cloning and Sequence Analysis of *Cdc*42 cDNA from Buffalo

ZHANG Xiao-ru[1,2], KUANG Wen-hua[1,2], CHENG Ying[1], DU Li[1], LEI Ming[1], JIAO Han-wei[1], ZHANG Dong-lin[1], HAO Yong-chang[1], QI Chao[2], WANG Feng-yang[1]

(1. College of Agriculture, Hainan University, Hainan Key Lab of Tropical Animal Reproduction & Breeding and Epidemic Disease Research (Construction Period), Animal Genetic Engineering Key Lab of Haikou, Haikou 570228, China; 2. College of Life Sciences, Huazhong Normal University, Wuhan 430079, China)

Abstract: The objective of the study was to clone full-length cDNA of *Cdc*42 (Cell division cycle 42) gene of buffalo and analyze this gene. RT-PCR were performed to clone *Cdc*42 cDNA from buffalo, and the cloned *Cdc*42 cDNA sequence had been sent to GeneBank and got a accession number: JN380079. The bioinformatics analysis results showed that the *Cdc*42 cDNA sequence was 626 bp length, containing an open reading frame of 576 bp, encoding 191 amino acids. Relative molecular weight was 21.3 kD, pI was 6.16. Further analysis results indicated that the homology of *Cdc*42 gene of buffalo are 99.65%, 96.70%, 94.97% and 94.10%, compared to cattle, dog, mouse and human, respectively; the protein was non-transmembrane protein, which include a RHO domain. These results set a strong basis for further research of the function of *Cdc*42 of buffalo.

Key words: *Cdc*42; Buffalo; cloning; bioinformatics analysis