

东方蝾螈 C-crkl 基因的生物信息学分析

孙菲阳

(深圳北理莫斯科大学, 广东 深圳 518172)

摘要 C-crkl 基因作为信号转导的关键基因, 参与了多种生理活动, 而目前对蝾螈 C-crkl 基因的研究十分匮乏。因此, 本研究采用 DNAMAN、SOPMA、MEGA 等软件拟对东方蝾螈 C-crkl 基因的基本生物学信息进行分析, 揭示了东方蝾螈 C-crkl 基因的基本生物学信息, 为进一步研究提供了依据。

关键词 东方蝾螈 C-crkl 基因 肢体再生 生物信息

中图分类号: Q3

文献标识码: A

文章编号: 1007-0745(2022)01-0058-03

近年来, 蝾螈肢体再生机制研究是再生医学研究的热点, 人们发现了大量生物分子参与肢体再生过程。随着基于转录组学、蛋白质组学技术的进步, 发现了大量与肢体再生相关的基因, 阐明这些基因在肢体再生过程中的作用, 将是肢体再生研究的重点。

信号接头蛋白 C-crkl 参与了受体整合蛋白和酪氨酸激酶等多种分子的信号转导, 与肿瘤的发生、发展密切相关, 在人体组织中不表达或少量表达, 而在乳腺癌等恶性胶质瘤中的 C-crkl 在 mRNA 与蛋白质水平均是高表达。然而目前对蝾螈 C-crkl 基因的研究十分匮乏, 因此本研究拟对东方蝾螈 C-crkl 基因的基本生物学信息进行分析, 为其在生物学等方面的功能提供理论依据。

1 文献综述

随着现代科学的进步, 人类在应对受伤的肢体中, 从依赖假肢到尝试新方法使肢体再生, 并在此基础上开拓了更专业的研究领域。研究发现, 哺乳动物的肢体无法持续再生主要是因为进化过程中已演变出优秀的人体免疫系统和愈合机制, 细胞的再生与伤口愈合中都处于相互抵制状态, 但蝾螈却具备强大的肢体再生能力。因此, 深入地研究蝾螈肢体再生的作用机理, 可为人类在再生医学方面指点迷津, 使断肢再生成为可能。

C-crkl 基因在癌细胞中过度表达, 与癌症进展密切相关。但 C-crkl 基因在肢体再生中是否存在重要作用仍然未知, 阻碍了研究。因此, 本文拟对东方蝾螈 C-crkl 基因的基本生物学信息进行分析, 为蝾螈 C-crkl 基因的进一步研究提供依据。

1.1 东方蝾螈概述

东方蝾螈 (*Cynops orientalis*), 也称中国火龙, 是国

家级珍稀保护动物, 属于两栖纲, 蝾螈属^[1]。主要分布在我国的江苏、浙江、湖北等省份, 背部与两侧分别呈现淡黑色, 腹部则为橙红色, 是有尾的两栖类脊椎动物^[2]。

蝾螈肢体再生研究进展: 德国科学家发现在蝾螈体内有一种特殊酶能够促进蝾螈肢体再生, 科学家猜想, 通过人为的合成这种酶, 或许可帮助肢体伤残及器官受伤的病人实现再生^[3]。蝾螈的肢体结构与人类肢体非常相似, 因此阐明蝾螈肢体再生的细胞机制对人类肢体再生具有重要意义。

目前研究揭示, 在蝾螈断肢前从肢体肘部游离肢体神经, 结果断肢后其肢体不能再生, 这是由于阻止了伤口表皮与肢体神经的正常相互作用, 从而导致肢体再生失败, 也表明肢体神经对蝾螈肢体成功再生至关重要^[4]。另外一些研究也揭示, 在断肢后人为、机械的去掉愈合的伤口表皮, 则肢体再生失败, 可见伤口表皮对蝾螈肢体成功再生十分重要^[5]。

1.2 C-crkl 基因的概述

C-crkl 是具有多功能的信号传导接头蛋白, 是原癌基因 C-crkl 编码的产物^[6], crkl 蛋白通过 mRNA 的选择性剪接, 产生从结构上划分的两种蛋白质: v-crkl 和 crkl i; crkl ii 和 crkl l, 高度同源的 SH2 和 SH3 结构域几乎可以构成所有的 C-crkl 蛋白^[7]。在信号分子传递的过程中, C-crkl 蛋白本身并不具备酶活性, 一般可与上下游的信号分子作用的是 SH2、SH3 结构域, 进行有序且特异性地信号传导, 参加肿瘤细胞的吸附、侵袭和增殖等信号的转导^[8]。

C-crkl 基因在癌症中的研究进展: C-crkl 广泛参与肿瘤细胞的迁移、生长、凋亡等。正常情况下, C-crkl 在人体中不表达或少量表达。而在乳腺癌、肺癌、卵巢癌等恶性的胶质瘤中 C-crkl 在 mRNA 与蛋白质水平

表1 软件的名称及功能

| 软件名称 | 具体功能 | 网址 |
|-------------|---|--------------------------------------|
| DNAMAN | 序列比对, 进化树分析 (Sequence alignment and Analysis of the phylogenetic tree) | |
| ProtParam | 蛋白质理化特性分析 (Analysis of physicochemical properties of protein) | (http://web.expasy.org/protparam) |
| SOPMA | 蛋白二级结构的分析 (Analysis of protein secondary structure) | (https://npsa-prabi.ibcp.fr/) |
| Swiss Model | 蛋白质三维结构预测 (Three-dimension structure of the protein) | (https://www.swissmodel.expasy.org/) |
| NCBI | 基因编码蛋白的结构域 (The domain of a protein encoded by a gene) | (https://www.ncbi.nlm.nih.gov/) |

均是高表达^[9]。

冯润华等通过检测368例胃癌患者癌组织中的C-crkl表达水平发现,C-crkl在胃癌组织中表达明显升高^[10]。此外,李建芳等通过生物信息分析及荧光素酶报告系统发现miR-126特异性靶向调控C-crkl表达参与胃癌的进展^[11]。

朱进等研究发现,C-crkl蛋白在患者乳腺癌组织中表达程度明显高于良性乳腺瘤组织^[12]。李敏等发现miR-126可通过超声微泡介导来调节C-crkl表达抑制乳腺癌细胞生长,为乳腺癌的治疗提供了新的措施^[13]。李文燕等研究发现C-crkl在卵巢癌细胞高表达可能直接促进卵巢癌细胞的恶性增殖^[14]。

2 实验部分

基于前期的转录组测序数据,获得了东方蝾螈C-crkl(coC-crkl)的完整CDS区cDNA序列,本实验拟对coC-crkl基因的核酸序列、编码蛋白质的基本理化性质等进行分析。

2.1 实验工具

如表1所示。

2.2 实验方法

1. 基于前期转录组的测序结果,获得coC-crkl基因的cDNA序列。

2.coC-crkl基因的氨基酸序列及核苷酸序列分析:

打开DNAMAN软件,选择“File”“new”,将cDNA序列导入,点击“Sequence”“load sequence”,选择“From Selection”然后点击“protein”,选择“translation”,点击选项“Frame1”。

3.coC-crkl基因编码蛋白质结构域分析:

进入NCBI网页,选择“Domains & Structures”,

点击“Conserved Domain Database (CDD)”,选择“CD Search”。

4.coC-crkl基因编码的氨基酸序列的基本理化性质分析:

打开ProtParam网站,将氨基酸序列导入,选择“Compute parameters”。

5. 分析预测coC-crkl蛋白的二、三级结构:

打开SOPMA网站,从Secondary structure prediction项目栏中选择“SOPMA”,点击“SUBMIT”。

打开Swiss-Model网站,点击“Start Modelling”,将氨基酸序列导入,点击“Search For Templates”“Build Models”。

6. 获得不同物种C-crkl蛋白质的氨基酸序列:

打开NCBI网站,选择“Protein”,输入C-crkl、Search,点击FASTA,选择“Send to”,勾选“Show GI”,点击“Create file”,下载不同物种的氨基酸序列。

7.coC-crkl蛋白与其他物种C-crkl蛋白质氨基酸序列比对分析:

打开DNAMAN软件,选择“Sequence”“Alignment”,选择“Multiple Sequence Alignment”,点击“File”,选择“Full Alignment”“Input”点击“NO”,选择“Run in background”和“show progress”,点击“下一步”,选择默认数值,完成操作。

3 实验结果

3.1 coC-crkl基因序列分析

通过DNAMAN软件分析发现,coC-crkl完整CDS区基因大小为567 bp,编码188个氨基酸。

3.2 coC-crkl结构域分析

使用NCBI Conserved Domain在线软件分析,发现

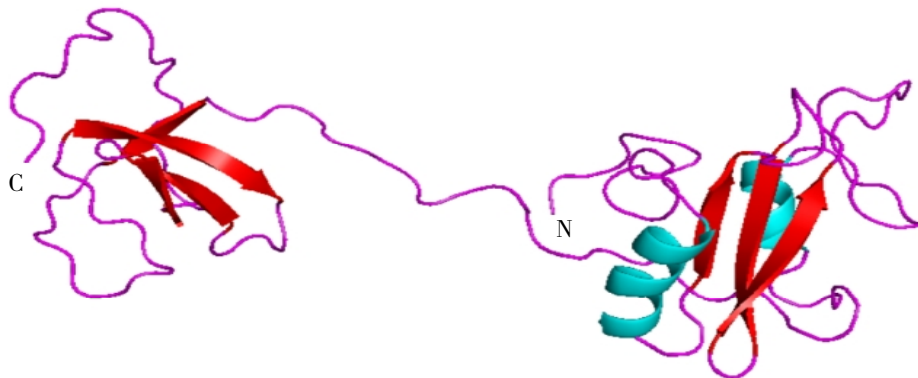


图1 东方蝾螈 C-crak 蛋白的三级结构预测

coC-crak 包含两个结构域, 分别是 SH2 结构域 (5-107 氨基酸) 和 SH3 结构域 (119-173 氨基酸)。

3.3 coC-crak 的理化性质分析

通过 ProtParam 分析发现, coC-crak 的分子量为 215 09.99 Da, 等电点为 5.16。

3.4 coC-crak 二级及三级结构预测

分析发现, coC-crak 的 β -折叠占 6.38%, 延伸链占 19.68%, α -螺旋占 18.62%, 55.32% 为无规则卷曲。以 2eyy.1A (同源性为 81.72%) 为模板, 对 coC-crak 进行同源建模, PyMOL Viewer 显示 coC-crak 蛋白质三级结构, 结果如图 1 所示。

3.5 coC-crak 氨基酸序列比对

利用 NCBI 下载不同物种的 C-crak 序列, 通过 DNAMAN 软件进行多序列比对。发现除 *Lepeophtheirus salmonis* 和 *Mytilus coruscus* 外, 其余物种 C-crak 的 SH2 结构域氨基酸序列高度保守, 本次所选物种的 c-crak 蛋白质序列 SH3 结构域高度保守。

4 结论

C-crak 是细胞信号传导的重要基因, 研究发现在多种恶性肿瘤中 C-crak 的 mRNA 及其蛋白表达程度很高, 表明其与肿瘤发展密切相关。本研究基于前期转录组测序结果获得 coC-crak 基因的 CDS 区 cDNA 序列, 通过生物信息学软件对其分析。

本研究采用多种软件对 coC-crak 基因 CDS 区核酸序列、编码蛋白质的基本理化性质、多序列比对进行分析, 阐明了 coC-crak 基因的基本生物学信息, 为其生物学功能研究提供理论依据。

参考文献:

[1] 王帅. 东方蝾螈的攀附爬行能力及其表面形态研究

[D]. 南京航空航天大学, 2016.

[2] 李顺才, 刘晓丽, 郝晓. 东方蝾螈的形态观察和解剖学研究 [J]. 中国兽医科学, 2005, 35(01): 60-63.

[3] 陆洪良, 耿军, 徐卫, 等. 东方蝾螈幼体热耐受性和游泳表现的热驯化响应 [J]. 生态学报, 2017, 37(05): 1603-1610.

[4] 唐婕. 东方蝾螈肢体再生的多组学分析 [D]. 西北大学, 2020.

[5] 耿文鑫, 严兴荣, 陈煜娴, 等. 东方蝾螈肢体异位再生的研究 [J]. 中国组织工程研究, 2009, 13(37): 7301-7304.

[6] 阴传敏, 姚紫薇. 信号接头蛋白 c-Crk 与恶性肿瘤关系的研究进展 [J]. 重庆医学, 2008, 37(02): 191-192.

[7] 翟波, 廖侃. 接头蛋白 c-Crk 的结构和功能 [J]. 生命的化学, 2003, 23(03): 198-201.

[8] 阴传敏. 信号接头蛋白 Crk 对宫颈癌恶性潜能的影响 [D]. 重庆医科大学, 2008.

[9] 冯润华, 毕任达, 陆晟, 等. Crk 在胃癌中的表达及其临床意义 [J]. 外科理论与实践, 2018, 23(03): 258-262.

[10] 李敏, 王健宝, 杨大艳. 超声微泡介导 miR-126 通过调节 CRK 抑制乳腺癌细胞生长 [J]. 中国免疫学杂志, 2020, 36(16): 1967-1972.

[11] 朱进, 胡孝渠, 郭贵龙, 等. 乳腺癌组织中 Crk 蛋白的表达及其临床意义 [J]. 实用肿瘤杂志, 2009(05): 446-449.

[12] Feller SM, Knudsen B, Hanafusa H. C-Abl kinase regulates the protein binding activity of c-Crk. EMBO J. 1994, 13(10): 2341-2351.

[13] Beitner-Johnson D, LeRoith D. Insulin-like growth factor-I stimulates tyrosine phosphorylation of endogenous c-Crk. J Biol Chem. 1995, 270(10): 5187-5190.

[14] 耿文鑫, 严兴荣. 东方蝾螈肢体异位再生的研究 [J]. 中国组织工程研究与临床康复, 2009, 10(13): 267.