

• 专论与综述 •

# 药用动物基因组及 EST 研究

刘 颖,程克棣,朱 平\*

(中国医学科学院中国协和医科大学药物研究所 卫生部天然药物生物合成重点实验室,北京 100050)

**摘 要:**以生物信息学数据库 GenBank 的生物信息为基础,对包括《中国药典》2000 年版一部、《新编中药志》第 4 卷、《中国大百科全书》(中国传统医学分册)、《中药大辞典》及《天然药物化学》第 3 版中所涉及的 450 多种重要的药用动物的基因组、蛋白质及表达序列标签(EST)的注册情况进行了初步统计,就相关进展进行了综述。到 2004 年 10 月底截止的数据分析表明,所统计的药用动物中 42% 以上没有 DNA 报道;48% 以上没有蛋白序列的报道。除模式动物和一些与人们日常生活密切相关的动物外,在 GenBank 数据库中仅能找到 27 种药用动物的 EST 数据,而药用动物基因组方面的数据也十分有限。并对这一领域的研究现状与展望进行了讨论。

**关键词:**药用动物;基因组;表达序列标签

**中图分类号:**R282.12;R282.3

**文献标识码:**A

**文章编号:**0253-2670(2005)09-1412-04

## Current status of genome and EST research in medicinal animals

LIU Ying, CHENG Ke-di, ZHU Ping

(Key Laboratory of Biosynthesis of Natural Products, Ministry of Health, Institute of Materia Medica, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100050, China)

**Key words:** medicinal animals; genome; expressed sequence tag

世纪之交实施的人类基因组研究这项宏伟的科学计划以其鲜明的前沿性、先进性和带动性使生命科学成为 21 世纪的主流科学和新技术、新产业的基础与生长点。随着此项计划的深入,一个以蛋白质和药物基因组学为研究重点的后基因组时代已经拉开序幕,并对科学家们提出了更为严峻的挑战。动物是人类食物和药物的重要来源之一,其基因研究也备受关注,并在基因组水平或 cDNA 水平取得了一定的进展<sup>[1]</sup>。本文仅就与药用动物有关的基因研究作一介绍,旨在为促进我国药用动物的研究提供基本信息资料。

基因资源是一种有限的战略资源,从种质资源中分离出来的基因或者用种质资源育成的品种具有自主知识产权,往往具有很高的价值。因此,作为一种不同寻常的资源,基因资源已成为又一个可争夺的对象。基因序列是基因研究的基础,而基因专利则是基因产业的基础。随着注册基因的不断增多,不能不对药用动物基因资源及相关基因产业的未来予以特别的关注。药用动物是中药材的重要组成部分,是祖国医药学的重要内容。药用动物在我国有着悠久的药用历史,它具有活性强、疗效佳、见效快等特点。目前人工饲养的药用动物主要有鹿类和蛤蚧、蝎子、蚂蚁、林蛙、水蛭等。但也有不少药用动物来源于濒危动物,如犀角、虎骨、麝香、牛黄、羚羊角等。由于动物药功能广泛、药效显著,在中医临床上起着重要的作用,所以深受医学界和人民群众的喜爱,使其临床需

求量急剧增加。人类对野生药用动物生态环境的破坏,盲目和大量地捕猎、走私珍贵药用动物,使野生动物资源蕴藏量剧减,供需矛盾日益增加。综合市场价格、购买趋势和每年中医药厂家消耗量等因素考虑,由于药用而造成资源匮乏最为严重的物种有麝、穿山甲、蝮蛇等。重视和加强药用动物资源的保护和开发利用,寻找新的动物药资源是目前动物药研究工作的重大课题。随着中药现代化的发展,我国有关部门对涉及药用动物利用的国内法律、法规及国际公约的实施,尤其是《濒危野生动植物种国际贸易公约》的实施给予了高度重视,同时有计划、有控制、有指导地开展濒危及使用量大的药用动物种类的驯养繁殖工作,以确保减轻对野生种群的压力。作为源头的药用动物,其种质资源(包括基因资源)存在混杂情况,从众多种质资源中挑选优良种质是一项重要工作,同时必须重视和保护种质资源的多样性,因为物种的灭绝意味着它们所携带的遗传基因也随之消失。

目前世界各国均将研究重点放在种质资源的基因鉴定与分离方面,首先受到关注的是与人类关系密切的动物。如由多国科学家组成的两个小组于 2004 年 3 月 1 日宣布绘出鸡的基因组序列草图和遗传差异图谱,这两项成果在禽流感防治、改善家禽和人类健康等领域都将有应用价值。另外蜜蜂基因组草图已完成测序并向公众公布,使得对蜜蜂的研究产生重要的积极意义。领导蜜蜂基因组测序计划的美国休斯

收稿日期:2005-01-20

作者简介:刘 颖(1979—),女,中国医学科学院 中国协和医科大学 2001 级硕士研究生,研究方向为生物转化。

Tel:(010)64711199-6045 E-mail:liuyliwd@sina.com.cn

\* 通讯作者 朱 平 Tel:(010)63165179 E-mail:zhuping@imm.ac.cn

顿贝勒医学院理查德-吉布斯说,如果没有蜜蜂和授粉,整个生态系统便会崩溃。另外,小鼠、大鼠、河豚、家蚕的基因测序也已基本完成。就整体而言药用动物有其特点:首先是种类很多;其次是不同动物的药用部位、药用价值及资源状况都各有千秋,而且大多数药用动物的遗传背景还不很清楚,因此在进行药用动物的基因研究时,其研究材料的取舍也是一个颇费斟酌的问题,了解世界范围内药用动物基因的研究情况是大有裨益的。

1 药用动物基因研究的总体情况

对包括《中国药典》2000 年版一部、《新编中药志》第 4 卷、《中国大百科全书》(中国传统医学分册)、《中药大辞典》及《天然药物化学》第 3 版中所涉及的 450 多种重要药用动物的 DNA 序列与蛋白质序列在 GenBank 数据库(<http://www.ncbi.nlm.nih.gov>)的注册情况进行了初步统计<sup>[2~6]</sup>。截至 2004 的 10 月底,455 种药用动物的 DNA 及蛋白质注册情况见表 1、2。总体看来,所统计的药用动物中 42% 以上没有 DNA 报道,48% 以上没有蛋白序列的报道,可见世界范围内对药用动物的研究都是十分有限的。

表 1 455 种药用动物 DNA 序列在 GenBank 注册数分布情况

Table 1 DNA sequence registration of 455 medicinal animals in GenBank

注册数/条	动物数	所占比例/%
0	192	42.2
1~9	133	29.2
10~99	92	20.2
100~1 000	24	5.3
大于 1 000	14	3.1

表 2 455 种药用动物蛋白质序列注册数分布情况

Table 2 Protein sequence registration of 455 medicinal animals in GenBank

注册数/条	动物数	所占比例/%
0	221	48.6
1~5	94	20.7
6~10	39	8.6
11~100	69	15.2
大于 100	32	7.0

从表 3 可看出,核酸序列注册较多的动物多为主要是药食两用、以食为主的动物。其中猕猴属于国家一级保护动物。蛋白质序列注册数较多的也是药食两用的动物,其中蛋白质序列注册最多的是褐家鼠和家鸡。蛋白质注册在千条以下、百条以上的药用动物钳蝎、马鹿、抹香鲸、东方鲑等。对其他药用动物的蛋白研究还较少。

对在 GenBank 注册的 DNA 序列信息的分析表明,在统计的药用动物中所注册的 DNA 序列主要涉及以下几方面:(1)线粒体相关基因:COX1、COX2、COX3、cytochrome B (cytB)、NADH dehydrogenase 等。(2)rRNA 基因:包括 18S、16S、12S 及其转录间隔区(internal transcribed spacer,ITS)

表 3 部分药用动物 DNA、蛋白质序列在 GenBank 注册情况

Table 3 DNA and protein sequence registration of some medicinal animals in GenBank

动物名称	注册蛋白数/条	注册核酸数/条
褐家鼠	47 665	1 148 264
家鸡	28 988	867 292
牛	13 140	1 820 717
家猪	8 633	450 677
家兔	6 778	8 564
犬	3 493	2 183 659
绵羊	2 976	21 125
猕猴	2 545	55 229
马	1 719	22 061
家蚕	1 274	397 554
鲤鱼	790	55 810
意大利蜂	334	11 224

等。(3)微卫星序列(microsatellite sequence)。所统计的药用动物在 GenBank 注册的绝大部分 DNA 序列属于以上 3 项。

根据已有的资料,药用动物基因的研究包括基因组 DNA、cDNA、表达序列标签(expressed sequence tag,EST)以及药用蛋白基因克隆与表达等不同层次的研究。下面仅就药用动物基因组 DNA 及 EST 的研究情况作一介绍。

1.1 基因组研究:主要是进行基因组 DNA 核苷酸序列的测定,并绘制相关图谱。本文对药用动物的基因组 DNA 注册情况进行了初步统计,结果表明进行基因组核苷酸序列测定(注册 dbGSS)研究的药用动物非常少。主要集中在犬、牛、褐家鼠、家鸡上,对这 4 种动物的 dbGSS 研究较多,注册数目都超过了 10 万条。

基因组研究是一项耗资巨大的基础研究,但与药用动物“药用”的关系似乎不够直接。因此,要对有丰富物种多样性的药用动物进行广泛基因组研究几乎是不可能的,此类研究的必要性容易受到质疑。

1.2 EST 序列研究:表达序列标签(EST)是从 cDNA 克隆两端获得的短的 cDNA 部分序列。简单地说,一个 EST 就是对应于某一种 mRNA 的一个 cDNA 克隆的一段序列。它在新基因资源中扩展最为迅速。EST 技术的理论基础是,来自某一组织的足够数量的 EST 可代表特定组织中的基因的表达情况,可用于研究基因表达模式和基因组序列之间的复杂关系,已成为注释基因组序列的宝贵资源,在功能基因组学研究的各个方面发挥了重要的作用。通过 EST 技术,可了解细胞、组织或生命进程中基因表达模式的特征,鉴定出特异性基因,特别是与药用动物蛋白有关的基因,如蜂王浆蛋白前体基因等。对 EST 认真细致的生物信息学分析可以鉴定出组织特异性的 EST,是制备高质量 cDNA 微阵列的前提工作。基于对 EST 数据的分析,还可以挖掘出 mRNA 加工和成熟的分子机制,如 mRNA 加工信号元件,可变剪接和 3' 末端加工的多样性等。

从一个物种所得到的所有 cDNA(或称为 cDNA 文库)

代表着这一物种正在表达的有功能的基因。药用动物之所以为药,其基础是药用动物的药用蛋白或其他的活性成分,它们的产生都是由相应的基因所决定的,为了获得目的蛋白或化合物,研究者把与之合成相关的基因作为主要研究目标。与活性成分相关 cDNA 的克隆首先需要一定的生物信息,即一段核苷酸序列或蛋白质的氨基酸序列,以此为起点用各种生物技术的方法可以克隆到所需要的目的基因。虽然目前从动物中克隆的药用相关基因还非常有限,许多基因还没有被认识,但通过提供一定的生物信息如 EST,就有可能迅速克隆所需要的基因。利用 EST 可以避开大量的实验室工作而阐明一条基因的大部分序列。

表 4 为目前 EST 序列报道较多的药用动物。其中家鸡、牛、家猪、犬、家蚕注册的序列最多,均达到了 10 万条以上。注册序列数在千条以上的有猕猴、马、意大利蜂、鲤鱼、绵羊等。注册序列数在百条以上的有山羊、家鸽等。而其他药用动物 EST 序列数则较少或根本没有。

EST 研究常常需要建立 cDNA 文库并需要进行大量的 cDNA 克隆序列测定,没有足够的经费及信息技术的支撑及后续工作目标的抉择和延续,也只能是一项基础性研究,不容易看到其在药物研发中的实用价值。

表 4 在 GenBank 注册 EST 的部分药用动物  
Table 4 EST registration of some medicinal animals in GenBank

动物名称	注册 EST 数/条	动物名称	注册 EST 数/条
家鸡	527 686	鲤鱼	10 528
牛	432 891	鲈鱼	2 469
家猪	358 856	家兔	2 275
犬	154 722	意大利蜂	2 265
家蚕	116 541	山羊	637
猕猴	46 571	鳗鲡	196
马	15 290	家鸽	174
绵羊	10 952	中华蜜蜂	62

2 药用动物基因组、EST 及药用蛋白研究现状分析

在 GenBank 数据库中注册数较多的传统中药包括鹿、麝等动物。其中梅花鹿主要注册的 DNA 序列包括:微卫星序列(3 条)、核糖体基因(12 条)、线粒体相关基因(90 条)、MHC(21 条)、生长因子(2 条)等,这些工作主要是由日本的一些科研机构以及我国东北农业大学、吉林大学完成的;林麝注册的 DNA 序列包括:核糖体基因(1 条)、activin beta-A(1 条)、线粒体细胞色素 b 相关基因(5 条),目前报道的只有这 7 条,主要是由我国四川大学、东北林业大学完成的;大海马注册的 DNA 序列包括:线粒体细胞色素 b 相关基因(87 条)、核糖体基因(11 条)、MHC(1 条),主要是由加拿大的科学家完成的;中华蜜蜂注册的 DNA 序列主要是蜂王浆蛋白前体有关基因(4 条)、线粒体相关基因(39 条)、Hypopharyngeal gland cDNA 库(62 条)。

东亚钳蝎是注册生理活性蛋白相关基因及蛋白较多的药用动物,所注册的 131 条核苷酸序列和 205 条蛋白序列除

个别外都是与生理活性蛋白相关的。已经注册的蛋白序列可大致分为以下几类:离子通道抑制剂或阻断剂:钾通道、钠通道与氯通道抑制剂、K<sup>+</sup>毒素样肽等;毒素类:哺乳动物神经毒素、酸性毒肽、昆虫毒素、防卫素样蛋白等;活性肽类:抗癫痫肽前体、肿瘤镇痛肽、抗凝血肽、类抗菌肽蛋白、抗神经激动肽等;生长因子类:成纤维细胞生长因子、缓激肽增强肽等。这方面的工作主要由我国武汉大学,中国科学院上海生物化学研究所、生理研究所、生化细胞生物化学研究所以及北京大学的科研工作者完成。

美国国立癌症研究所的基因组多样性研究室将穿山甲作为胎盘哺乳动物起源的研究对象之一,注册了穿山甲核苷酸序列,包括了活性蛋白基因并研究了重组蛋白、锌指蛋白(zinc finger protein)、神经营养生长因子、肾上腺素受体、腺嘌呤受体、磷脂酶 C 及 cAMP 相关基因等 15 项核苷酸序列和 11 条蛋白序列,相关结果在《Nature》杂志发表<sup>[7]</sup>。穿山甲注册的核苷酸序列中只有 1 项为线粒体基因,2 项为核糖体基因。因此,穿山甲也成为活性蛋白注册比例高的药用动物之一。

值得注意的是,如穿山甲,东亚钳蝎,梅花鹿和中华蜜蜂在最近几年注册的基因和蛋白数明显多于过去。从以上资料分析还可以看出一些趋势:(1)在药用动物基因组和 EST 研究方面所进行的工作还比较有限,研究的动物也相对集中。主要是牛、家猪、家鸡、犬、猕猴等动物。(2)现已进行或正在进行基因组 DNA 研究的主要目标动物是与人们基本生活相关的动物或者称之为广义上的药食同源动物,而重要药用动物基因组的研究十分有限。正在进行大规模 EST 研究的药用动物也为数有限,这可能与这些研究的深度及这类研究的投入较大而实际应用前景不十分明确有关。(3)随着人类功能基因组研究的深入,新的药物靶点不断出现,加之基因重组技术已十分成熟,研究和开发新药的需要将促进药用动物相关的药用蛋白的研究,重组动物药用蛋白的研发将成为新药研发中一个有希望的领域。(4)RNA 干扰技术在生物次生代谢过程研究中起到越来越重要的作用,相应的药用动物基因组学的研究是科学设计靶基因的 siRNA 的保证。

根据目前的研究趋势,对今后一段时间内药用动物基因组及 EST 研究可做出如下估计:由于投资较大而应用目标相对模糊,进行大规模基因组测序的药用动物种类近期不会有太大的增长,但与药用动物蛋白研究相关的 DNA 序列或 EST 的报道无论从动物种类或数量上都可能有较大的增长。目前 GenBank 数据库中不同生物的 dbEST 注册数据仍在不断更新。药用动物理应成为未来研究所关注的对象。但我国是否需要大规模投资进行药用动物 EST 的研究,则应论证研究对象的重要性和研究的目的性。但药用蛋白领域的研究应该有所加强。

3 结语

中国有丰富的药用动物资源,其蕴涵的基因资源急需保护和开发利用,这是我国药用生物资源可持续利用所必须的基础。中国政府对于药用生物可持续利用十分重视,我国科

学家在政府的支持下确实取得了不少的成果,但仍面临着不少严峻的问题。如:面对如此复杂和种类繁多的药用动物,该如何确定研究对象,采用哪种方法进行研究和保护,这需要广大科学工作者的共同努力。

#### References:

- [1] Rasin I, Ribnicky D M, Komarnytsky S, et al. Plants and human health in the twenty-first century [J]. *Trends Biotechnol*, 2002, 20(12): 522-531.  
[2] *Ch P* (中国药典) [S]. Vol 1. 2000.

- [3] Xiao P G. *Modern Chinese Materia Medica* (新编中药志) [M]. Vol IV. Beijing: Chemical Industry Press, 2000.  
[4] Yao X S. *Natural Products Chemistry* (天然药物化学) [M]. 3rd ed. Beijing: Peoples' Medical Publishing House, 2001.  
[5] Shi D B. *Chinese Traditional Medicine* (中国传统医学) [M]. Beijing: Encyclopedia of China Publishing House, 1992.  
[6] Jiangsu New Medical College. *Dictionary of Chinese Materia Medica* (中药大辞典) [M]. Vol 1. Shanghai: Shanghai Scientific and Technical Publishers, 1996.  
[7] Murphy W J, Eizirik E, Johnson W E, et al. Molecular phylogenetics and the origins of placental mammals [J]. *Nature*, 2001, 409(6820): 614-618.

## 植物药材次生代谢产物的积累与环境的关系

苏文华<sup>1,2</sup>, 张光飞<sup>1</sup>, 李秀华<sup>1</sup>, 欧晓昆<sup>1</sup>

(1. 云南大学生态学与地植物学研究所, 云南 昆明 650091; 2. 云南省生物资源保护与利用重点实验室, 云南 昆明 650091)

**摘要:**许多植物药的活性成分是其所含的次生代谢物质,人工种植时其产量取决于初生产物的积累,质量取决于次生产物的积累。药材质量及有效性的基础是植物的次生代谢。次生代谢物质的产生是发育程度、组织分化及外界刺激因素通过影响生物合成基因表达而控制。对次生代谢产物在植株体内形成和积累的诱导有不同的诱导机制假说。但都认为环境胁迫时,次生代谢产物数量增加。次生代谢产物积累与生长环境条件需求方面存在矛盾,是药用植物种植的难题。

**关键词:**药用植物;环境条件;质量;产量;次生代谢

**中图分类号:**R282.21

**文献标识码:**A

**文章编号:**0253-2670(2005)09-1415-04

### Relationship between accumulation of secondary metabolism in medicinal plant and environmental condition

SU Wen-hua<sup>1,2</sup>, ZHANG Guang-fei<sup>1</sup>, LI Xiu-hua<sup>1</sup>, OU Xiao-kun<sup>1</sup>

(1. Institute of Ecology and Geobotany, Yunnan University, Kunming 650091, China; 2. Laboratory for Conservation and Utilization of Bio-resources, Yunnan University, Kunming 650091, China)

**Key words:** medicinal plant; environmental condition; quality; yield; secondary metabolism

现代药学研究表明,许多植物药的药理作用与其所含的次生代谢物质有关<sup>[1]</sup>。这类药用植物人工种植时其产量取决于初生代谢产物的积累,而其质量就取决于次生代谢产物的积累,保持其药材质量及有效性的基础是植物的次生代谢。

次生代谢与初生代谢一样是植物体内重要的生理代谢,次生物质在协调与环境的的关系上充当着重要角色。植物面临环境胁迫时,次生代谢产物能提高植物自身保护和生存竞争能力。国内外对于植物的次生代谢已开展了大量的研究。在次生代谢与环境关系、次生代谢产物积累的内在规律等方面积累了大量的基础知识,形成了一些基础理论。这些成果是认识药用植物次生代谢有效成分变化,以及药用植物种植时有效成分控制的理论基础。

#### 1 次生代谢产物合成与积累的环境诱导作用

一些植物不论环境如何,都会合成、积累次生代谢物质。然而,对大多数植物而言,次生代谢产物的合成与积累往往受制于所处环境的变化。它们根据所处环境的变化来决定合成次生代谢产物的种类和数量,只有在特定的环境下才合成特定的次生代谢产物,或者显著地增加特定的次生代谢产物在体内的产量。碳同位素技术研究已证明,棉花中的驱虫挥发性次生代谢产物是在受到昆虫刺激后才合成的,并不事先得合成储存这些挥发性物质<sup>[2]</sup>。植物体内次生代谢产物的合成与积累需要环境条件的诱导。

可诱导或影响次生代谢产物合成积累的环境因子很多,有生物因子、非生物因子。生物因子如病虫害的侵袭、个体密度等;非生物因子如水分、温度、养分、紫外线、土壤理化性质、空气污染等<sup>[2-4]</sup>。环境因子对植物体内次生代谢的影响和作

收稿日期:2005-01-13

基金项目:国家自然科学基金资助项目(30360009)

作者简介:苏文华(1962—),男,云南昆明人,副教授,主要从事生理生态及野生药用植物驯化栽培研究。