

· 成果报告 ·

日本血吸虫(大陆株)成虫基因表达谱的研究

吴忠道, 余新炳, 徐 劲, 李 焱, 李晖婷, 周俊梅, 包俊英, 卞国武, 郑亦男, 邵 筱, 彭寨玉

(中山大学中山医学院寄生虫学教研室, 广东 广州 510080)

摘 要: 报告作者 1998 年至 2000 年有关日本血吸虫(大陆株)成虫基因表达谱研究的结果, 主要包括: 已测定 551 条 EST, 其中 519 条 EST 已送入国际基因数据库(GenBank/dbEST), 并获得了 GenBank 的登录号。经同源整合后, 519 条 EST 序列归类为 388 个基因序列, 其中, 24 条(6.19%) 为日本血吸虫已知基因序列; 18 条(4.64%) EST 与曼氏血吸虫已知基因序列同源, 90 条(23.19%) EST 与其它生物的已知基因序列同源, 同源基因包括精氨酸酶、酰胺酶、微纤维蛋白、肌动蛋白、细胞色素 C 氧化酶、二硫异构酶、加长因子、热激蛋白 70、Y 盒结合蛋白和抗凋亡因子等。其余的 256 条(65.98%) EST 在 GenBank 中未发现已知基因的同源序列。将非日本血吸虫同源序列 EST 或未检索到同源序列的 EST 初步定义为日本血吸虫亲基因/未知基因序列 EST, 共 364 条。已克隆日本血吸虫精氨酸酶基因、Y-盒结合蛋白基因和抗凋亡-1 因子 3 个新基因的全长 cDNA, 其 GenBank 的登录号分别为 AF402615、AF367371 和 AF333765。本研究结果丰富了血吸虫基因组数据, 并为进一步筛选日本血吸虫疫苗候选基因或药物靶位提供了科学依据。

关键词: 日本血吸虫/遗传学; 基因; 表达序列标签; 计算机生物学

中图分类号: R383.2⁺ **文献标识码:** A **文章编号:** 1000-257X(2002)06-0401-04

Schistosoma japonicum Mainland Strain: The Gene Expression Profile in Adult Worm WU Zhong-dao, YU Xin-bing, XU Jing, LI Yan, LI Hui-ting, ZHOU Jun-mei, BAO Jun-ying, BIAN Guo-wu, ZHENG Yi-nan, SHAO Xiao, PENG Zhai-yu. (Department of Parasitology, Zhongshan Medical College, Sun Yat-sen University, Guangzhou 510080, China)

Abstract: This paper summarized the results of the research on the gene expression profile in adult worm of *Schistosoma japonicum* (*S. j*) Mainland strain from 1998 to 2000. The main results are as follows: The EST approach has provided partial sequences of 551 clones. Among them, 519 had been deposited in dbEST, which were grouped as 388 different expressed genes of *S. j*. Homologous searching for significant similarities to sequences deposited in schistosoma database of EMBL, and GenBank with the basic local alignment search tool program (Blastn & Blastx) showed that 24 ESTs (6.19%) represented genes previously characterized in *S. j*, 18 ESTs (4.64%) similar to the sequences of *Schistosoma mansoni*, 90 ESTs (23.19%) similar to the known gene of other species except schistosome, these homologous genes included arginase gene, amidase gene, fibrillin gene, actin gene, cytochrome c oxidase gene, disulfide isomerase gene, elongin gene, heat shock protein 70 gene, Y-box binding protein gene and DAD1 (Defender Against Apoptotic Death-1) gene etc. 256 ESTs (65.98%) of ESTs had no significant matches with sequences of schistosoma or other species, which were considered to be the unidentified genes or newly characterized genes of this organism. For a further characterization of those clones, full-length sequencing of interesting clones were carried out. Some novel *S. j* genes such as arginase, *S. j* Y-box binding protein and DAD1 genes had been cloned and sequenced, their accession numbers were AF402615, AF367371 and AF333765 respectively. The study provides the information about the gene expression profile in *S. j* adult worm, and adds the sequence data for schistosoma genome database as well as screening of Schistosomiasis vaccine and drug development.

Key words: *Schistosoma japonicum*/genetics; gene; expressed sequence tags; computer biology

血吸虫属多细胞脊椎动物且生活史复杂, 但基因组相对较小(270 Mb), 开展血吸虫基因组研究不但具有重要的医学意义如研制血吸虫病疫苗等, 而且还可以作为研究人类基因表达与调控, 以及开展基因结构与功能关系研究的理想模型(或称作模

式生物基因组)。随着分子生物学技术以及计算机技术和生物信息学技术的发展, 大规模测序已成为基因组计划的主要内容, 也是基因表达谱研究及发现新基因的有效手段。1991 年, 美国 NIH 的 MD Adams 开始运用表达序列标签(expressed sequence

收稿日期: 2002-08-01

基金项目: 广东“211 工程”重点建设基金资助项目(98169); 国家自然科学基金资助项目(30070683); 国家教育部博士点基金资助项目(200045); 广东省首批自然科学团队研究基金资助项目; 本项目获得 2001 年度中国高校自然科学二等奖

作者简介: 吴忠道(1962-), 男, 江西宜黄人, 医学博士, 博士后, 教授, 硕士生导师; 主要研究方向: 病原生物功能基因组研究。余新炳(1953-), 男, 安徽宿松人, 医学博士, 教授, 博士生导师, 主要研究方向: 病原生物功能基因组研究。

tags EST)法寻找新基因^[1]。1995年, Fanco 首先应用 EST 法对曼氏血吸虫(*S. m*)的基因进行了分离和序列测定^[2]。随着 WHO 血吸虫基因组计划的启动(1995年), *S. m* EST 的分离与测序进展迅速, 但日本血吸虫(*S. j*)EST 序列增长仍较缓慢, 至本研究启动时, 在 NCBI 的 GenBank/dbEST 中收录的血吸虫 EST 序列 9 737 条, 其中 *S. m* 有 8 216 条, 而 *S. j* 且 EST 仅有 834 条, 只占 *S. j* 表达基因数的 4.17%。因此, GenBank 中已有的 EST 序列资料还不足以对日本血吸虫基因谱进行全面的描述。所以, 开展日本血吸虫基因表达谱的研究对于进一步丰富国际血吸虫基因序列资料和深化对日本血吸虫基因组的认识, 对于寻找新的血吸虫病疫苗候选分子、诊断抗原及药物靶标均具有重要的意义。为此, 我们从 1998 年开始进行了日本血吸虫基因表达谱的研究, 现将阶段性研究工作总结如下。

1 EST 的获得及同源性分析^[3,4]

采用表达序列标签(EST)技术策略, 随机从已构建的日本血吸虫(大陆株)成虫 cDNA 文库(*S. j* cDNA 文库)中挑出单个 cDNA 克隆, 制备模板, 扩增出 cDNA 插入片段; 采用 PCR 产物直接序列测定技术, 对插入片段的 5' 端进行序列测定; 原始的碱基序列经人工处理, 去除模糊序列及载体序列后得到相应的 cDNA 克隆插入片段的 EST 序列, 并通过互联网与 EBI (European bioinformatics institute)中血吸虫核苷酸序列数据库和 NCBI GenBank 中的核苷酸序列及蛋白质序列进行同源性比较(BLASTn 或 BLASTx)。

本研究已完成 EST 测序 551 条, 其中 519 条为有分析价值的 EST 序列。根据 WHO 血吸虫基因组计划的有关规则并获得 NCBI GenBank 的认可, 建立了本实验室 cDNA 克隆和 EST 的命名方法^[3], 已在 GenBank/dbEST 中登录的 EST 序列有 519 条, 其 GenBank/dbEST 进入号见表 1。在 519 条 EST 序列中, 出现 1 次的 EST 序列为 313 条, 出现 2 次的有 53 条, 出现 3 次以上的有 22 条。出现次数最多的 EST 为 JAYL0340.GY, 共出现 14 次。通过聚类分析(cluster analysis)将所获得的 EST 整合归类为 388 个独立的 EST, 并采用 Blastn/Blastx 等程序, 对 EBI 血吸虫基因数据库及 GenBank 进行检索。结果发现与日本血吸虫已知序列同源的

EST 占 6.19%, 与曼氏血吸虫已知序列同源的 EST 占 4.64%, 与其它生物已知基因同源的 EST 占 23.19%。另外还有 65.98% EST (256 条)未发现有已知基因的同源序列, 其中, 214 条 EST 也未发现有同源 EST 序列存在, 因此, 这部分 EST 序列是 GenBank 中首次公布的日本血吸虫序列。

表 1 已送入 GenBank 登录的 EST 序列统计

Table 1 List of ESTs deposited in GenBank

Accession number	No of ESTs
AI725355-AI725361	7
AI740189-AI740204	16
AI816553-AI816554	2
AW160107-AW160132	26
AW231224-AW231255	32
AW282251-AW282275	25
AW329868-AW329884	17
AW736722-AW736785	64
AW738819-AW738828	10
BE123822-BE123844	23
BE128954-BE128960	7
BE187660-BE187701	42
BE217367-BE217399	33
BE601437-BE601442	6
BE607311-BE607321	11
BE819807-BE819830	24
BF059758-BF059768	11
BF169228-BF169272	45
BF186483-BF186482	24
BF219676-BF219681	6
BF348716-BF348730	15
BF713605-BF713630	26
BG797176-BG797186	11
BG886457-BG886485	29
BG891014-BG891020	7
Total	519

2 已知或同源基因的识别^[5~17]

同源性分析显示, 在已发现的同源基因中酶类基因占 27.46%、结构蛋白基因占 9.86%、表达调控相关基因占 9.86%、信号转导/受体相关基因占 3.52%、生理/代谢相关基因占 9.15%。其中, 卵壳蛋白基因序列是高表达的基因之一, 卵壳蛋白基

因EST有7条(7个拷贝),占1.35%。高拷贝的基因还包括一些酶类基因,如组织蛋白酶B、甘油醛-3-磷酸脱氢酶、 26×10^3 (kDa)谷胱甘肽-S-转移酶(GST)、胞质液氨肽酶及丝氨酸酶等。亲环素B及线粒体DNA也是高表达的基因之一。

2.1 酶类基因

发现的已知日本血吸虫酶类基因或曼氏血吸虫同源基因有: Sj26kDaGST、甘油醛-3-磷酸脱氢酶、丝氨酸酶、烯醇化酶、组织蛋白酶B、胞质液氨肽酶(cytosol aminopeptidase)、天冬氨酸蛋白酶、酰氨酶、超氧化物歧化酶(SOD)、天冬氨酸蛋白酶、果糖二磷酸醛缩酶、核苷酸激酶、细胞色素C氧化酶、二硫异构酶等。另有20个EST与其它生物的酶基因同源,这些基因编码的酶类包括精氨酸酶、酪氨酸蛋白酶、遍在蛋白缀合酶、假尿苷合成酶、氨基乙糖苷酶、钾泵-驱动ATP酶、甲硫氨酸腺苷转移酶、岩藻糖苷酶、羟化酶、二氢乳清酸氧化酶、酰磷酸酶、磷酸甘油酯变位酶、细胞色素C氧化酶、苯基tRNA合成酶、尿苷二磷酸葡萄糖-焦磷酸化酶、酪氨酸激酶、氨基酸转移酶、分泌性组织蛋白酶、反式异构酶和乳酸脱氢酶等。

2.2 结构蛋白基因

与结构蛋白基因同源的EST有6个。所谓结构蛋白是指构成血吸虫虫体实质性组织的蛋白,如卵壳蛋白、副肌球蛋白、动力蛋白、肌动蛋白、微纤蛋白和丝束蛋白等。已知卵壳蛋白可诱导宿主产生抗血吸虫生殖免疫的效果,是具有潜在研究价值的疫苗候选分子。微管是真核细胞的骨架之一,它对于支持细胞形态、参与细胞的各种运动、细胞内大分子物质的运输及细胞器的位移有重要作用。微管还是构成动物细胞中心体、纤毛和鞭毛、有丝分裂纺锤体等结构的主要元件。研究微管的形成和确定微管的物质成分,一直是细胞骨架研究中的一个重要方面。本研究得到与微管构成蛋白编码基因同源的EST,这些微管构成蛋白基因分别是肌动蛋白和微管蛋白等基因。血吸虫的体被为双层液态膜,富含大量晶体状的棘(crystalline spines)。光学实验提示,这些棘主要由上述几种蛋白构成的微管组成,它们对维持血吸虫成虫体被的形成和完整性的保持具有重要意义。对这些基因的功能研究将有助于寻找到新的血吸虫疫苗候选分子。

2.3 信号传导及表达调控相关基因

本研究发现的与信号传导及表达调控有关的

基因包括转录起始因子(translation initiation factors)、Y-盒结合蛋白(Y-box binding protein)、4-跨膜蛋白(tetra span transmember protein)、翻译可控肿瘤蛋白、翻译起始因子、伴侣蛋白、延长因子-1、延伸因子C、剪接因子、真核翻译起始因子、转录因子、成纤维细胞生长因子和抗凋亡-1因子等基因。Y-盒结合蛋白是与MHC II类分子的表达调控相关的重要功能蛋白。值得关注的是,体外合成的SmNF-YA能与同样包含CCAAT区域的Sm28GST的启动子相结合,这提示有类似结构的DNA结合蛋白能对Sm28GST起到转录调节的作用,而后者是血吸虫疫苗的重要候选分子,因此,研究日本血吸虫的Y盒结合蛋白可能成为血吸虫疫苗和新药研究一条新的思路。

2.4 生理功能相关蛋白基因

与生理功能相关蛋白基因同源的EST包括脂肪酸结合蛋白、热激蛋白、亲环素和伴侣蛋白基因等。本研究中,EST jayh0011是日本血吸虫的脂及酸结合蛋白(FABP)编码基因EST。在血吸虫生活中,脂类具有重要作用,它是虫体细胞膜的重要成分,又可改变宿主对血吸虫的体液免疫和细胞免疫,从而逃避宿主免疫攻击。但血吸虫不能合成长链脂肪酸和胆固醇,只能从宿主体内摄取有效成分,故作为脂肪酸运输载体的脂肪酸蛋白,在血吸虫疫苗和药物研究中受到广泛注意。1994年,WHO推荐了6种较好的血吸虫候选疫苗抗原分子中,其中FABP也在其中。EST jayh0094为日本血吸虫热激蛋白70(SjHSP/70)编码基因EST。HSP是细胞在一些应激条件,如热激、葡萄糖饥饿或受到病原感染时高效表达的一族蛋白。它实际上是一种被称为分子伴侣的进化上非常保守的蛋白质,这种分子能够结合和稳定另外一种蛋白质的不稳定构象,通过有控制的结合和释放,促进新生多肽的折叠、多聚体的装配或降解及细胞器蛋白的跨膜运输,防止蛋白质的热变性和聚集,从而避免细胞进一步受到伤害。在受到感染和发生自身免疫性疾病时,HSP可作为重要的抗原被免疫系统识别,因此在医学方面的作用也日益受到重视。已经证实当人受到细菌和寄生虫感染时,这些病原体来源的热激蛋白是一种非常重要的抗原。热激蛋白能作为免疫系统识别的重要抗原的原因有两个:一是当病原体被巨噬细胞吞噬时,它表达高水平的热激蛋白;二是在不同的病原体中,热激蛋白是高度保

守的,因此免疫系统可能很方便地识别这些高度保守的分子。HSP 作为疫苗候选抗原分子开发为疫苗研究提供了一种新的思路。

3 新基因识别及全长 cDNA 克隆^[18~22]

根据同源性检索结果,将非日本血吸虫已知序列包括与其它血吸虫同源的序列初步定义为日本血吸虫新基因/未知基因序列。本研究共发现 364 个日本血吸虫新基因序列 EST,其中 298 条 ESTs 在 GenBank 中未发现同源基因序列的 EST,即在数据库中未找到与之同源的血吸虫或其他生物基因或 EST 序列。因此,它们极有可能是日本血吸虫的特异基因,可能是更好的药物靶标和候选疫苗分子,具有重要的研究价值。对这类 EST 序列,目前正在通过对 dbEST、dbSTS 等数据库的重叠群(contigs)搜寻或定位信息(一致性>90%),然后通过电子拼接(eRACE)预测可能的全长 cDNA 序列,据此序列设计引物进行 PT-PCR 或筛选 cDNA 文库,克隆得到该基因的全编码序列。事实上,这项工作也是本研究后续工作的重点之一。

通过 cDNA 末端快速扩增(RACE)或 EST 序列的电子拼接等方法,已在国际上首先克隆了日本血吸虫精氨酸酶基因、Y-盒结合蛋白基因(Sj Y-box binding protein, SjYB-1)和抗凋亡-1 因子基因(Sj Defender against apoptotic death-1, SjDad-1)3 个新基因的全长 cDNA,其 GenBank 的登录号分别为 AF402615、AF367371 和 AF333765。

参考文献:

- [1] Adams M D, Kelley J M, Gocayne J D, *et al.* Complementary DNA sequencing: expressed sequence tags and human genome project[J]. *Sciences* 1991, 252(5013): 1651.
- [2] Franco G R, Adams M D, Tanaka M, *et al.* Identification of new *Schistosoma mansoni* genes by EST strategy using a directional cDNA library[J]. *Gene*, 1995, 152(2): 141.
- [3] 吴忠道,余新炳,李 焱,等. 日本血吸虫表达基因 EST 测定及同源性分析[J]. 中山医科大学学报, 2000, 21(4S): 25.
- [4] 吴忠道,余新炳,李 焱,等. 日本血吸虫(中国大陆株)表达

- 基因的大规模随机测序体系的建立[J]. 广东寄生虫学会年报, 1999, 21: 44.
- [5] 吴忠道,余新炳,郑亦男,等. 日本血吸虫(中国大陆株)表达基因的分离和 EST 序列测定[J]. 中国人兽共患病杂志, 2000, 16(1): 3.
- [6] 吴忠道,余新炳,李 焱,等. 日本血吸虫特异性抗原基因的筛选和 EST 序列测定[J]. 中国血吸虫病防治杂志, 2000, 12(6): 330.
- [7] 李 焱,余新炳,吴忠道,等. EST 法筛选日本血吸虫表达基因序列的研究[J]. 广东寄生虫学会年报 1999, 21: 21.
- [8] 郑亦男,余新炳,吴忠道,等. 日本血吸虫卵壳蛋白基因的筛选及 EST 序列测定[J]. 广东寄生虫学会年报, 1999, 21: 23.
- [9] 李晖婷,余新炳,吴忠道,等. 日本血吸虫(中国大陆株)表达基因的 EST 序列测定及同源性分析[J]. 广东寄生虫学会年报, 1999, 21: 26.
- [10] 卞国武,余新炳,吴忠道,等. 日本血吸虫(中国大陆株)成虫表达序列标签的获取及分析[J]. 广东寄生虫学会年报, 2000, 22: 24.
- [11] 包俊英,余新炳,吴忠道,等. 日本血吸虫新基因的发现及其功能的初步预测[J]. 广东寄生虫学会年报, 2000, 22: 38.
- [12] 李 焱,余新炳,吴忠道,等. 日本血吸虫基因表达序列标签的获得和分析[J]. 地方病通报, 2001, 16(1): 6.
- [13] 李晖婷,余新炳,吴忠道,等. 日本血吸虫(中国大陆株)若干酶类基因的发现[J]. 中国人兽共患病杂志, 2001, 17(4): 46.
- [14] 刘兰英,余新炳. 日本血吸虫卵壳蛋白基因的克隆与表达[J]. 中国寄生虫学与寄生虫病杂志, 1998, 16(6): 415.
- [15] 周俊梅,余新炳,吴忠道,等. 日本血吸虫卵黄铁蛋白基因的体外扩增及逆转录病毒 PLXSN-FER1 重组质粒的构建. 中国人兽共患病杂志 2000, 16(1): 32.
- [16] 李 焱,余新炳,吴忠道,等. 日本血吸虫 26kDa 谷胱甘肽 S-转移酶基因真核表达载体的构建及序列测定. 中国人兽共患病杂志 2000, 16(3): 28.
- [17] 冯新港,余新炳,李 焱,刘彦文. 日本血吸虫大陆株脂肪酸结合蛋白 SjFABPc 编码区基因的体外扩增及克隆. 中国人兽共患病杂志, 1999, 15(3): 33.
- [18] 李 焱,余新炳,吴忠道,等. 日本血吸虫未知基因 JAYL0230 全长 cDNA 的克隆和测序分析[J]. 广东寄生虫学会年报, 2000, 22: 17.
- [19] 李 焱,吴忠道,郑焕钦,等. Sj Dad1 反义核酸抗日本血吸虫感染的实验研究. 热带医杂志, 2001, 1(1): 23.
- [20] 吴忠道,郑亦南,徐 劲,等. 日本血吸虫精氨酸酶编码基因 cDNA 的分离和序列分析[J]. 中山医科大学学报, 2001, 22(2): 85.
- [21] 吴忠道,李 焱,徐 劲,等. 日本血吸虫精氨酸酶的 cDNA 序列分析[J]. 中国人兽共患病杂志, 2001, 17(3): 22.

(编辑 张敏瑞)

《科学美国人》评出 2002 年度科研领袖 我国两科学家获此殊荣

中国青年报北京 2002 年 11 月 14 日电中国基因研究专家、中国科学院基因组信息学中心暨北京华大基因研究中心主任杨焕明,近日被《科学美国人》杂志评为“年度科研领袖人物”。与他一同分享该荣誉称号的还有:瑞士先正达公司的斯蒂芬·戈夫,美国华盛顿大学基因组中心的于军。他们因“在水稻基因组测序研究中发挥的领导作用”而获得这一荣誉。事实上,于军教授本人同时还是北京华大基因研究中心副主任、杨焕明的同事。而且,据北京华大基因研究中心办公室主任邓亚军透露,于军目前已辞去华盛顿大学教授职位,从本月起他已是华大中心的全职人员。因此,中国科学家在此次“年度科研领袖人物”评选中是“三占其二”。

(陈婷舒)