



日本七鳃鳗(*Lampetra japonica*)肝脏 ESTs 分析与比较转录组研究

朱丽娜^① 戴雅丽^② 马 飞^① 李庆伟^{①*}

(^① 辽宁师范大学生命科学学院, 大连 116029; ^② 安徽农业大学生命科学学院, 合肥 230036)

摘要 以日本七鳃鳗(*Lampetra japonica*)肝脏为材料构建 cDNA 文库, 在文库中随机挑选克隆子进行测序共得到 10077 条有效 ESTs(expressed sequence tags)序列. ESTs 序列分析显示, 8515 条 ESTs 拼成 648 条片段重叠群, 共得到 2210 条转录本, 其中 47.06% 的转录本预测为全长序列; 利用 BLAST 程序在 GenBank 数据库中进行同源性搜索发现 2053 条转录本有同源序列匹配, 占总转录本的 92.9%. 更进一步对这些基因产物进行 Gene Ontology 注释, 结果发现, 在日本七鳃鳗肝脏中与有颌类免疫、凝血和代谢相关的基因大量表达, 并预测了 8 个新基因. 通过对日本七鳃鳗与底鳗(*Fundulus heteroclitus*)、鼠(*Mus musculus*)、牛(*Bos taurus*)和人(*Homo sapiens*)肝脏转录组的比较分析, 发现日本七鳃鳗肝脏中比其他物种优势表达的是甲壳质酶和多糖代谢等相关的基因, 这些基因可能在日本七鳃鳗免疫中发挥重要作用. 此外, 也利用 TargetScan 软件对日本七鳃鳗肝脏转录组中 3'UTR 区进行 microRNA 靶标识别, 结果发现了与人类癌症基因调控同源的 microRNA 靶标, 这为研究人类癌症提供了有益的线索. 上述结果将为七鳃鳗功能基因和蛋白组学的研究以及脊椎动物的基因组进化提供重要的理论基础.

关键词 日本七鳃鳗(*Lampetra japonica*) 肝脏 EST 转录组比较

日本七鳃鳗(*Lampetra japonica*)属圆口纲, 七鳃鳗目, 七鳃鳗科, 七鳃鳗属, 是最原始的脊椎动物类群, 是联系无脊椎动物与脊椎动物之间的重要桥梁. 最近的考古发现, 现代七鳃鳗与 3.6 亿年前的祖先没有太大的变化, 是名副其实的“活化石”^[1], 因此七鳃鳗是研究脊椎动物起源和进化的首选材料. 目前, 对日本七鳃鳗的研究大多局限于形态结构、生态分布等方面^[2], 而关于基因组学和蛋白组学方面的研究则较少报道. 国内外现已有海七鳃鳗肠^[3]以及日本七鳃鳗口腔腺^[4]转录组的研究报道, 但尚未见最大的消化腺——肝脏的相关基因的报道. 本文对日本七鳃鳗

肝脏转录组的研究, 将为进一步研究脊椎动物肝脏的起源和进化提供理论基础.

由于七鳃鳗独特的生活方式和形态特点, 直到 19 世纪末人们对它的研究才兴起^[5], 而且主要集中在比较形态学和早期脊椎动物进化等方面, 例如实验证实了七鳃鳗与有颌类脊椎动物脑部基因、咽结构的同源性^[6,7], 探讨了 *Hox* 基因复制的机制^[8]. 肝脏作为机体重要的代谢器官日益引起人们的重视, Savina 等人比较了七鳃鳗肝脏线粒体不同时期生物能参数^[9], Rotchell 等人分析了 CYP1A 酶在七鳃鳗肝脏中的催化模式^[10], Gamper 等人研究了七鳃鳗产卵前肝脏

收稿日期: 2007-04-09; 接受日期: 2007-8-28

国家高技术研究发展计划(863 计划)项目(批准号: 2007AA09Z428)、国家自然科学基金项目(批准号: 60575005)、教育部科学技术研究重点项目(批准号: 206032)、大连市科技计划项目(批准号: 2006E11SF068)和大连市优秀青年科技人才基金项目资助

* 联系人, E-mail: liqw@263.net

可逆的代谢减缓^[11]。同时肝脏也肩负一线免疫防御的功能。最近英国*Nature*杂志报道,科学家在海七鳃鳗类淋巴细胞的细胞上发现了新型可变受体,利用一种不寻常的基因重排过程来产生受体多样性,证明海七鳃鳗具有一种特有的适应性免疫应答^[12]。七鳃鳗印记了从无脊椎向脊椎动物的进化历史,包含了各种功能基因的进化信息。因此研究七鳃鳗肝脏基因组及转录组具有十分重要的理论和实践意义。

本实验室率先构建了日本七鳃鳗口腔腺^[13]和肝脏的cDNA文库,目前已经完成了口腔腺 1280 条 ESTs 序列的生物信息学分析^[4]。本研究在构建日本七鳃鳗肝脏cDNA文库的基础上,随机挑选cDNA克隆进行单次测序构建EST库,对测得的 10077 条ESTs通过BLAST同源性比较、Gene Ontology(GO)蛋白功能分类等进行注释,并与底鳉(*Fundulus heteroclitus*)、鼠(*Mus musculus*)、牛(*Bos taurus*)、人(*Homo sapiens*)肝脏的转录组进行比较分析以及对 3'UTR区的microRNA靶标进行预测等,期望得到日本七鳃鳗肝脏中特异表达的基因,为下一步功能基因研究奠定基础。

1 材料和方法

1.1 材料

实验对象为日本七鳃鳗,于 12 月中下旬捕于黑龙江省松花江流域同江地区,取其新鲜的肝脏组织。PCR 引物、工具酶、RNA 提取试剂盒、RT-PCR 试剂盒、测序试剂盒、pBluescript II SK(+)克隆载体以及 *E. coli* JM109 Electro-cell 菌种均由宝生物工程(大连)有限公司提供。

1.2 cDNA 文库构建与测序

剥离日本七鳃鳗肝脏组织,液氮研磨,使用 Trizol Reagent 试剂提取总 RNA。利用 Oligotex-dT₍₃₀₎ 进行 mRNA 的分离和纯化后,以带有 Not I 酶切位点的 Oligotex-dT Linker Primer 为反转录引物合成 cDNA 第一链,并进一步合成双链 cDNA。于 cDNA 片段两端添加 EcoR I 连接位点,酶切后克隆于带有 Not I /EcoR I 酶切位点的 pBluescript II SK(+)载体上构建成 cDNA 文库。随机挑选 cDNA 文库中的克隆,以载体两端通用引物进行 PCR 扩增。挑选插入片段长度大于 500 bp 克隆子提取质粒。进行片段 5'端单序列测定。测序试剂采用 TaKaRa TaqTM Cycle Sequencing Kit,序列测定采用 ABI377 型全自动分析。

1.3 EST 拼接及全长 cDNA 的检测

利用CAP3 软件^[14]进行EST拼接,将EST拼接成重叠群和单一序列,再将单一序列继续拼接直到不再产生新的重叠群为止,最小重叠区为 25 bp及一致性为 75%。利用在线工具 TargetIdentifier^[15](<https://fungalgenome.concordia.ca/tools/TargetIdentifier.html>)进行拼接后的全长检测,预测出全长的cDNA序列。参数设为默认值。

1.4 ORF 的预测和 UTR 区分析

根据NCBI上ORFfinder软件(<http://www.ncbi.nlm.nih.gov/projects/gorf/>)预测所有拼接后序列的可读框,取最大的可读框并分离出microRNA主要的调控作用区域 3' UTR区,从而进行microRNA靶标的预测识别。TargetScan 软件^[16](<http://www1.ncifcrf.gov/app/htdocs/appdb/drawpage.php?appname=miRNA>)将已知库中除植物外的所有 microRNA 序列与分离出的 3' UTR 区进行比对,得到对应的 microRNA 和靶标。

1.5 EST 注释、功能分类及新基因的寻找

采用本地下载的BLASTX程序,将拼接后的一致序列以可能的 6 个可读框进行翻译,然后与登录在NCBI非冗余数据库(nr)中的序列进行比对。对于有同源性的序列,取同源性最高的已知功能的基因注释这一基因功能。其余的为功能未知的蛋白质序列。如果有些序列在BLASTX和BLASTN程序的搜索中都没有比对到同源产物,再将这部分序列在整合了PROSITE, PRINTS, Pfam, ProDom等 13 个数据库内容的 InterProScan^[17](<http://www.ebi.ac.uk/InterProScan/>)搜索界面上搜索,如果也没有找到同源的蛋白结构域,就可以被认为是首次发现的EST序列。根据标准基因词汇体系(gene ontology, GO),我们把具有BLASTX注释信息的基因进行分类,应用Gopipe软件^[18](1.2 版)(<http://www.Fishgenome.org/bio-info/>)将其分为生物学途径、分子功能、细胞组件三类。

1.6 转录组的比较分析

为进行转录组比较,我们从NCBI的UniGene数据库下载其他物种肝脏的EST序列,分别下载底鳉 33369 条,鼠 123756 条,牛 48787 条,人 251224 条,然后利用软件 Godiff^[19](<http://www.fishgenome.org/>)

bioinfo/)寻找功能差异基因. Godiff是通过比较两个物种的基因表达量概率来确定同一基因在不同物种间表达量关系的软件. 将日本七鳃鳗的肝脏EST与上述 4 个物种的肝脏EST进行两两比较. 确定相关性并对结果进一步统计分析, 与底鳉、鼠、牛、人的肝脏转录组比较, 从进化上分析表达量低的基因, 再从同源性搜索、蛋白结构域、分子功能和参与的生物过程等方面分析表达量高的基因, 试图揭示进化上日本七鳃鳗肝脏与环境适应的特异表达模式.

2 结果

2.1 cDNA 文库的构建和 EST 测序

效价测定表明日本七鳃鳗肝脏 cDNA 文库的容量为 4.6×10^6 pfu/mL, 插入片段的平均大小为 1.2 kb, 符合建库标准. 从 cDNA 文库中随机挑取克隆从 5' 端单向测序, 去除低质量序列、载体来源和污染序列最后共得到 10077 条有效 ESTs 序列.

2.2 EST 拼接、片段重叠群分析和全长 cDNA 拼接检测

利用 CAP3 软件对 10077 条 ESTs 序列进行拼接

共得到 648 个重叠群(contig)和 1562 条单一序列. Contig 全长为 641486 bp, 平均长度 990 bp. 重叠群分布(表 1)表明, 大部分 contig 是由两个 ESTs 组成, 其中高重叠群所对应的同源蛋白见表 2. 从组成同源蛋白的 EST 数目可以大致估计该蛋白的表达量, 所以研究结果显示, 血浆阿朴脂蛋白 LAL2、纤维蛋白原、卵黄蛋白原、核糖体蛋白 L4、胸腺细胞 cDNA 及酸性哺乳动物甲壳质酶在肝脏中表达量较高.

运用在线工具 TargetIdentifier对拼接好的序列进行全长检测, 共预测出 1040 条全长 cDNA, 其中 65.12%的片段重叠群(422 条)和 39.56%的单一序列(618 条)预测为全长序列, 可以看出将近 50%的转录本预测为全长. 根据 ORF 的分布结果(表 3)可以看出编码框大多集中在 100~600 bp 之间, 大于 300 bp 的 ORF 有 627 条, 说明该序列大部分编码蛋白质. 大于 1000 bp 的 ORF 有 29 条, 按照同源比对结果对这 29 条序列进行初步分析(表 4), 其中几种较长的序列预测功能分别为核糖体蛋白、纤维蛋白原、热激蛋白、甲壳质酶和巨噬细胞 cDNA 等, 而且其分值和期望值(E)都说明序列的相似度极高, 证实了拼接之后序列的可信度.

表 1 重叠群分布情况表

每个重叠群含 EST 数目	重叠群数目	占全部重叠群百分数/%	EST 数目	占全部 EST 百分数/%
2	263	40.59	526	5.22
3	103	15.90	309	3.07
4	59	9.10	236	2.34
5	27	4.17	135	1.34
6	26	4.01	156	1.55
7	14	2.16	98	0.97
8	9	1.39	72	0.71
9	12	1.85	108	1.07
10	7	1.08	70	0.69
>10	128	19.75	8367	83.03
合计	648	100	10077	100

表 2 高重叠群的 EST 数目

重叠群序列	含有 EST 数目	长度/bp	同源蛋白功能
contig401	3255	1077	血浆阿朴脂蛋白 LAL2
contig35	777	5385	纤维蛋白原
contig547	669	6050	卵黄蛋白原
contig277	265	3425	核糖体蛋白 L4
contig168	202	2958	胸腺细胞 cDNA
contig582	191	3454	酸性哺乳动物甲壳质酶

表 3 ORF 分布

ORF 核苷酸数目/bp	全长序列数目	百分数/%
>1000	29	2.79
1000~600	87	8.37
600~300	511	49.13
300~100	405	38.94
<100	8	0.77
总计	1040	100

表 4 ORF 大于 1000 bp 基因的功能预测

序列	核苷酸数目	同源蛋白功能	E 值	分值
502	4955	核糖体 S27α蛋白	3.00×10^{-62}	239
65	2452	纤维蛋白原α-1 链前体	0	817
187	1791	40S 核糖体蛋白 S4	1.00×10^{-138}	493
119	1731	热激蛋白 90-α	0	827
582	1575	酸性哺乳动物甲壳质酶前体	1.00×10^{-109}	400
504	1508	甜菜碱-高半胱氨酸转甲基酶	0	638
515	1488	纤维蛋白原β链	0	914
252	1392	核糖体蛋白 S6	1.00×10^{-112}	411
34	1371	丝氨酸蛋白酶抑制剂	0	686
208	1347	骨髓巨噬细胞 cDNA	0	800
447	1323	丝氨酸蛋白酶抑制剂	4.00×10^{-87}	325
410	1318	α-2 微管蛋白	0	810
439	1302	S-腺苷高半胱氨酸水解酶	0	765
35	1299	纤维蛋白原γ链	0	796
325	1241	核糖体蛋白 L5a	1.00×10^{-129}	463
308	1224	丝氨酸蛋白酶抑制剂	6.00×10^{-37}	157
250	1218	真核翻译起始因子 3	1.00×10^{-84}	317
470	1215	染色体 3 SCAF14679	0	635
296	1185	控制翻译的肿瘤蛋白	1.00×10^{-64}	250
401	1164	血浆阿朴脂蛋白 LAL2 前体	5.00×10^{-78}	295
514	1155	眼科新糖蛋白	1.00×10^{-58}	138
218	1125	核糖体蛋白 L17	6.00×10^{-79}	298
360	1125	β肌动蛋白	0	553
518	1095	纤维蛋白原γ链多肽	9.00×10^{-58}	144
321	1059	血浆阿朴脂蛋白 LAL2 前体	5.00×10^{-6}	55
595	1050	ADP/ATP 移位酶	1.00×10^{-147}	523
280	1020	MGC81823 蛋白	1.00×10^{-145}	518
22	1014	细胞色素 P450, 4 家族	1.00×10^{-102}	373
21	1005	血清凝集素前体	0	661

2.3 3'非翻译区(3' UTR)的 microRNA 靶标分析

通过 TargetScan 软件将库中除植物外的所有

microRNA 通过不完全碱基配对的方式和分离得到的 791 条 3' UTR 区序列比对, 其中 611 条和 2568 种 microRNA 对应了 25500 个靶标位点. 配对到动物

microRNA 的靶标位点有 23107 个, 线虫 1895 个, 病毒 498 个. 不同物种的 microRNA 具有高度的保守性, 统计筛选出有 18 个物种的 miR-92 家族 microRNA 分别对应同一靶标位点. 其次 17 个物种的有 miR-7, miR-18 家族, 接下来 16 个物种的有 miR-19a 家族, 15 个物种的有 miR-16, miR-17 家族, 14 个物种的有 miR-15a, miR-20, miR-29, miR-100, miR-103, miR-181 和 miR-106 家族等.

2.4 EST 的注释和新基因的发现

利用 BLASTX 程序进行同源性比对后发现, 共有 2053 条转录本在数据库中可以找到同源物, 占总数的 92.9%, 其中 1686 条转录本推测的蛋白编码产物与数据库中功能已知的蛋白序列具有同源性, 367 条与推测的、假想的或功能未知的蛋白质序列具有同源性. 统计 BLASTX 比对后的结果, 发现有许多转录本的比对产物为同一蛋白质产物, 其中数目最多的是核糖体蛋白的各个亚型共 123 条, 其次是巨噬细胞 cDNA51 条, NADH 脱氢氧化酶 38 条, 胸腺细胞 cDNA32 条, 血浆阿朴脂蛋白 28 条, 与甲壳质代谢相关的酶 22 条等. 另外还发现 T 细胞受体(TCR)、B 细胞受体(BCR)、主要相容性复合物(MHC- I, MHC- II)、补体因子(C3)、补体因子 B(CFB)、转铁蛋白、热激蛋白以及与免疫相关的因子和蛋白等.

大规模测序产生 EST 和利用生物信息学方法对 EST 数据库进行检索, 可以筛选新的基因. 本研究在 BLASTX 和 BLASTN 程序的搜索中都没有比对到同源产物的有 18 条. 将这 18 条序列在 InterProScan 搜索界面上搜索, 又有 4 条能找到同源的蛋白结构域, 剩下的 14 条 ORF 编码框大于 300 bp 的有 8 条, 因此这部分序列可以初步被认为是首次发现的 EST 序列(分别为 contig68, contig600, LLLNSEM008FB09, LLLNSEM011FG11, LLLNSEM055FD08, LLLNSEM134FD08, LLLNSEM120FF10 和 LLLNSEM064FD-09), 为下一步新基因的克隆和功能研究提供了基础.

2.5 蛋白功能分类

将上面 BLASTX 的结果用 Gopipe 软件进行功能分类, 结果显示生物学途径 693 条, 分子功能 764 条, 细胞组件 565 条. 在生物学途径中(图 1)最多的是生理过程(664 条), 其次是新陈代谢(526 条)、细胞生长(220 条)、转运(158 条). 在分子功能中(图 2)最多的是

绑定功能(356 条). 细胞组件中最多的细胞内结构的组成(427 条)(图 3)(每个基因注释的不同功能可以在 3 种分类中交叉体现).

2.6 转录组比较分析

与其他物种肝脏转录组进行比较, 结果显示在日本七鳃鳗肝脏中低量表达的是与 RNA 剪接、DNA 修复、细胞调亡的负调节等相关的序列; 而高表达的则大都是多糖水解活性蛋白, 包括甲壳质代谢酶以及氨基葡萄糖代谢酶等, 它们的表达量远远大于其他物种(表 5). 表达上述功能的序列共有 22 条, 其中 12 条同源性搜索结果为酸性哺乳动物甲壳质酶, 其表达量分别是底鳉、鼠、牛和人的 123, 142, 70, 166 倍. 保守结构寻找发现, 日本七鳃鳗肝脏中的甲壳质酶与人类甲壳质酶有同样的蛋白结构域, 都属于多糖水解结构域 18 家族.

3 讨论

3.1 日本七鳃鳗 cDNA 文库构建和 EST 测序

日本七鳃鳗在分类学上属于脊椎动物门、圆口纲、七鳃鳗目, 是最原始的脊椎动物类群, 也是研究动物进化方面的重要材料之一. 目前日本七鳃鳗资源量相当匮乏, 处于易危状态, 且尚无有效的保护措施, 因此建立日本七鳃鳗不同组织的 cDNA 文库不仅为进化遗传学的研究提供依据, 也将为这一物种遗传资源的保存做出贡献.

3.2 日本七鳃鳗肝脏转录组的注释

根据重叠群中含有 EST 数目的多少和转录本的比对产物为同一蛋白质的重复数量可以初步估计该蛋白的表达量, 我们通过序列同源性比对, 发现核糖体蛋白, 血浆阿朴脂蛋白 LAL2, 纤维蛋白原, 卵黄蛋白原, NADH 脱氢氧化酶, 胸腺细胞 cDNA, 巨噬细胞 cDNA, 酸性哺乳动物甲壳质酶等这几种蛋白在日本七鳃鳗肝脏中表达量较高. 更进一步研究揭示日本七鳃鳗肝脏中也存在与有颌类脊椎动物免疫相关基因同源的表达片段, 包括上面高表达量的巨噬细胞 cDNA, 胸腺细胞 cDNA, 甲壳质代谢酶以及热激蛋白, TCR, BCR, MHC- I, MHC- II, 补体因子 C3, C2, CFB, 干扰素和转铁蛋白等(表 6). 肝脏内富有巨噬细胞, 利用它可以吞噬消化和清除血中及经肠道吸收的微生物异物等有害物质, 通过吞噬、隔离和消

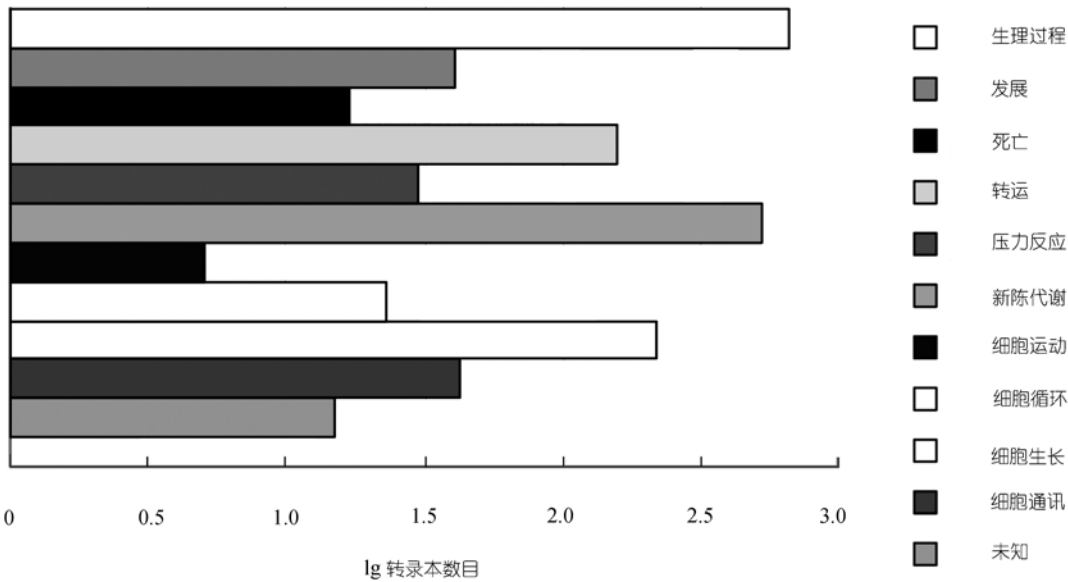


图 1 生物学途径

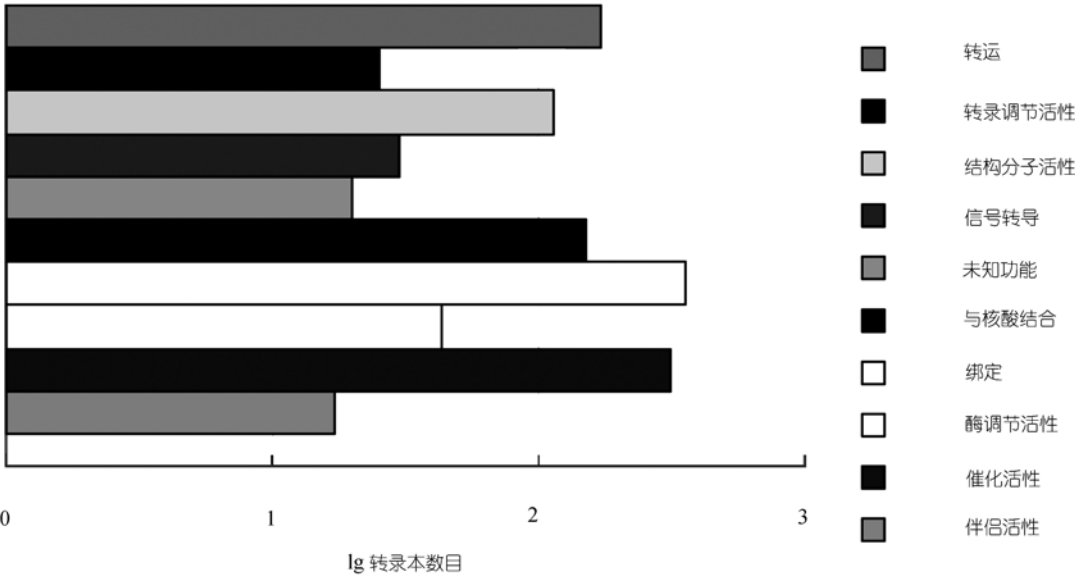


图 2 分子功能

除来改造入侵和内生的各种抗原；未成熟的T淋巴细胞在胸腺中培育，而且胸腺素还能提高淋巴细胞的杀伤能力诱导B细胞成熟^[20]。甲壳质是动物食用纤维，具有多种生物活性，有强化生物体免疫功能，提高生物体抗病能力，因此甲壳质代谢也与免疫相关^[21]。这3种基因表达量如此之高说明它们在日本七鳃鳗肝

脏免疫方面起着重要的防御作用。

补体系统是先天免疫系统中的中心辅助系统，其进化史可以追溯到棘皮动物(海胆)、非脊椎动物(棘皮动物、海鞘类)和圆口纲动物，它们拥有最原始的补体系统(凝集素途径)；干扰素系统是目前所知的机体防御反应中出现最早的细胞功能调节系统，也是

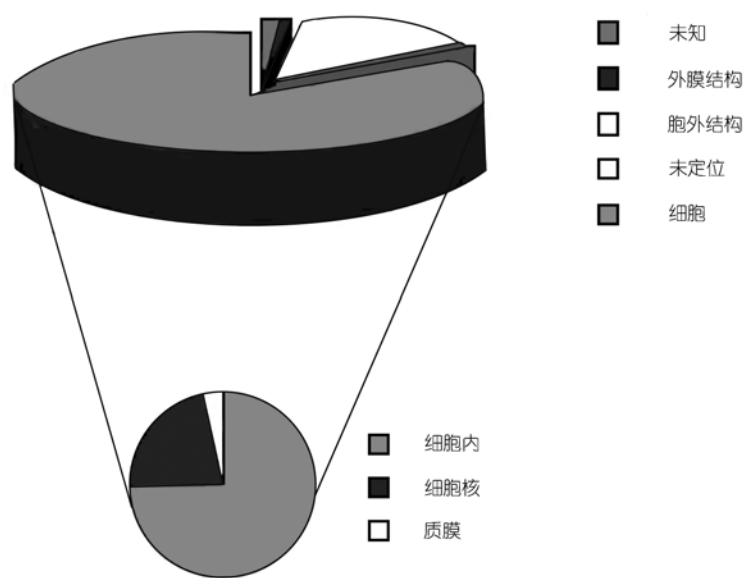


图 3 细胞组件

表 5 日本七鳃鳗肝脏转录组表达量高于其他物种的基因功能

GO 号	GO 功能	底鳉(倍数)	鼠(倍数)	牛(倍数)	人(倍数)
272	多糖分解代谢	62.74	72.45	44.91	79.49
44247	细胞内多糖分解代谢	124.53	71.90	65.14	123.96
4568	甲壳质酶激活	123.57	160.53	70.03	246.01
5529	糖结合	81.77	21.34	22.41	23.94
6030	甲壳质代谢	123.57	142.69	70.03	166.65
6032	甲壳质分解代谢	123.57	160.53	70.03	246.01
6040	氨糖代谢	41.19	36.69	28.01	35.14
6041	葡萄糖胺代谢	41.19	45.86	28.01	41.00
6043	葡萄糖胺分解代谢	123.57	160.53	70.03	172.20
6044	乙酰氨基葡萄糖胺代谢	41.19	45.86	28.01	44.10
6046	乙酰氨基葡萄糖胺分解代谢	123.57	160.53	70.03	246.01
46348	氨糖分解代谢	123.57	160.53	70.03	172.20
8061	甲壳质结合	123.57	1284.28	120.05	469.65

在生物中普遍存在的一个很保守的先天免疫系统;转铁蛋白具有抗菌杀菌、自我保护的抗病性能,是抑制细菌繁殖的重要因子^[22],因此可知日本七鳃鳗非特异性免疫系统在其自身抗病作用中发挥更大作用.日本七鳃鳗属于半寄生性动物,主要靠吸食其他鱼类血肉为食,非特异性防御机制在防止感染中扮演重要角色,潜在的非特异性防御机制可以在微生物入侵时发生作用,能更有效地清除、降解病原微生物和其他有害物质.

七鳃鳗已经表现出适应性免疫系统的一些特

征^[23~25],它们有一种适应性免疫响应,但不是我们在有颌类脊椎动物中看到的重组抗原受体,而是利用一种不寻常的基因重排过程来产生受体多样性^[12].Mayer等人^[3]在海七鳃鳗肠中发现有与小鼠类似的淋巴细胞,但没有发现B细胞、T细胞受体和MHC等关键因子,所以认为免疫系统还没进化到有颌类阶段.而我们的研究却发现上述各种因子在日本七鳃鳗肝脏中都存在,不过量很少,相似度也很低.在基因注释中发现热激蛋白的数目还是比较多的.越来越多的证据表明,热激蛋白能参与特异性和非特异

表 6 日本七鳃鳗肝脏中免疫相关基因

数目	基因	数目	基因
51	骨髓巨噬细胞 cDNA	1	杀伤细胞免疫球蛋白受体
32	胸腺细胞 cDNA	1	干扰素- γ 诱导蛋白
18	CD11c +ve 抗原	1	人类白细胞抗原 B 关联性转录物 1
12	酸性哺乳动物甲壳质酶	1	MHC I 类抗原
7	热激蛋白	11	MHC II 类抗原
6	蛋白酶体 26S ATP 酶	1	淋巴毒素
5	ABC 转运子	11	杀伤细胞凝集素受体
4	C3 补体前体	11	B 细胞移动基因蛋白
2	巨噬细胞激动蛋白	11	B 细胞受体
2	七鳃鳗蛋白酶体 LMPX(7)	11	CD3 信号传感器
2	补体因子 B	1	树突状细胞蛋白(GA17 protein)
1	CD9-蛋白	1	补体调节 GPI-锚定蛋白
1	CD63-蛋白	1	补体受体 2
1	胸腺素	1	补体受体 1
1	T 细胞受体	1	补体因子 H-related 5
1	蛋白酶体 β 4 亚型(PSMB4)	1	C2 补体前体
1	蛋白酶体 β 7 亚型(PSMB7)	1	CD81 抗原
1	蛋白酶体 α 2 亚型 (PSMA2)	1	CD68 抗原
1	蛋白酶体 α 3 亚型(PSMA3)	1	转铁蛋白

性免疫反应，认为热激蛋白之所以能促进特异性免疫反应，是因为它能作为分子伴侣携带抗原信息，激发机体的免疫反应，但有时候热激蛋白也可以直接激发免疫反应^[26,27]。由于七鳃鳗中存在极少的B细胞、T细胞受体和MHC等关键抗原呈递分子，推测日本七鳃鳗在适应性免疫的早期阶段尝试用热激蛋白作为呈递抗原或直接激活适应性免疫反应。以上证据表明日本七鳃鳗已经具备了但不同于有颌类适应性免疫的一些免疫应答机制，这仍需更进一步的实验验证。

表达量高的另一类为参与物质代谢的血浆阿朴脂蛋白LAL1 和LAL2，核糖体蛋白以及能量代谢的NADH脱氢氧化酶、细胞色素氧化酶和ATP合成酶等。以前的研究发现血浆阿朴脂蛋白LAL1 和LAL2 在七鳃鳗的血液中大量表达^[28]，我们的结果也发现在日本七鳃鳗肝脏中它们也是高表达。阿朴脂蛋白是由阿朴E基因控制，在肝脏中分泌，主要起着转运、溶解胆固醇和磷脂的作用。而且脂蛋白还诱导巨噬细胞胆固醇酯蓄积,显著促进巨噬细胞增殖，增强非特异性免疫机制。小鼠肝脏SAGE数据分析也显示阿朴脂蛋白mRNA的表达占总表达的百分比>3.8%。日本

七鳃鳗肝脏中高表达的基因也编码血浆蛋白。此外，我们的研究结果也显示共有 123 条序列与核糖体蛋白的 73 个亚型匹配。

最后一类为凝血因子，包括纤维蛋白原的 α , β , γ 链，凝血因子VIIa 和IXa 以及凝血素等。纤维蛋白原的高表达是因为肝脏是体内合成凝血因子的场所，其中最重要的就是纤维蛋白原^[29]，所以本研究也说明了日本七鳃鳗的凝血机制可能已较为完善。

还有一些低表达的基因，比如仅有一条 EST 序列注释的基因有 634 个，两条的有 108 个，表明肝脏中表达的基因种类繁多。大量注释基因的获得为进一步了解七鳃鳗肝脏的机制及获取相关基因提供了重要的序列信息。

3.3 microRNA 靶标分析

microRNA是最近发现的一类约有 18~24 个核苷酸的非编码的小RNA分子，这些小RNA分子起着负调控基因表达作用，但是绝大多数动物中的microRNA 并不导致靶mRNA的降解，仅降低其靶基因的蛋白质表达水平，其mRNA水平几乎没有受到什么影响^[30]，所以在EST水平我们就可以通过预测

出microRNA相应的靶标, 来寻找可能的负调控表达模式。

俄亥俄州立大学的Volinia等人^[31]通过分析来自肺部、胸部、胃部、前列腺、结肠和胰腺等处的癌细胞样品 540 份, 发现了由过量表达的大部分microRNA组成的实体癌症信号, 在这些microRNA中包括 miR-17-5p, miR-20a, miR-21, miR-92, miR-106a和miR-155 家族。而在我们的结果中找到 miR-92, miR-17, miR-20, miR-106 家族的microRNA分别对应的靶标位置。成年型B细胞慢性淋巴型白血病(CLL)病人常有miR-15a 和miR-16-1 基因簇的缺失或是表达下调^[30]。我们也发现miR-15a和miR-16 家族的靶标位置。日本七鳃鳗和人类癌症基因调控的microRNA家族有如此的相似, 因此可以成为一种新的与人类癌症疾病相关的动物研究模型。

3.4 转录组的比较

通过不同基因组之间的比较可以获得大量的遗传进化信息, 对系统发育中代表性物种之间全方位表达基因和基因家族的比较分析更可以揭示转录组和表达基因在进化过程中的演变, 为探讨脊椎动物起源提供一些重要的启示^[32]。七鳃鳗代表着从无脊椎动物到脊椎动物进化过程中的一个重要过渡阶段, 通过日本七鳃鳗和现有的完整肝脏转录组的底鳃、鼠、牛和人比较可以追踪进化过程中某些基因的变化, 从而提供了有关脊椎动物起源和进化的线索。Godiff是寻找不同转录组之间功能基因差异表达的软件, 我们通过比较发现在日本七鳃鳗肝脏中比其他物种表达量少的基因有RNA剪接、DNA修复、细胞调亡的负调节等, 而且随着进化地位的上升其表达量也逐级升高, 这说明肝脏在这些方面的功能进化机制越来越完善。而筛选出日本七鳃鳗肝脏中高表达的则大都是多糖水解活性, 包括甲壳质代谢以及氨基葡萄糖代谢等, 它们的表达量大于其他物种 100 倍左右。在前面BLASTX的同源注释中发现它本身的表达量也较高, 所以推测甲壳质酶在日本七鳃鳗肝脏中的大量表达定有它重要的作用机制。

甲壳质酶是普遍存在的以甲壳质为底物的糖苷水解酶, 为单基因所编码。在日本七鳃鳗肝脏中优势表达的甲壳质酶根据同源性分析多为酸性哺乳动物甲壳质酶(AMCase), 该酶2001年首次在人体中发现^[33], 推测在胃肠消化管和肺脏中起消化和防卫作用。同

年Suzuki等人^[34]报道AMCase基因在牛的肝脏中表达, 推测牛可以防御甲壳类病原微生物。在小鼠中克隆和表达酸性哺乳动物甲壳质酶可以水解红色毛癣菌和白念珠菌细胞壁, 具有潜在的抗真菌功能^[35]。Yang等人^[36]发现了一种在感染丝虫机体中特异高表达的甲壳质酶CV-CHI-I, 在正常情况下数据显示该酶表达可以达到 22%, 但感染后可达到 42%~62%。同时IgG3 表达量增高, 外周单个核细胞(PBMC)也增多, 巨噬细胞在七鳃鳗肝脏中的优势表达就是一个有利证据, 而且被酶分解的甲壳质的免疫功能已被国际所公认。Nishimura等人^[37]往小鼠腹腔内注射经醋酸处理的脱乙酰基几丁质的多孔珠, 结果发现该物质可刺激小鼠产生白介素 1 及增强体内巨噬细胞的活性。几丁聚糖的分解产物氨基葡萄糖也能活化自然杀伤细胞(NK)、淋巴因子激活的杀伤细胞(LAK)。壳多糖还能诱发局部巨噬细胞增生, 并使其活性增强, 有免疫功能^[38], 所以甲壳质酶与免疫相关。

综合以上结果可推测日本七鳃鳗肝脏中甲壳质酶免疫的代谢途径: 首先由甲壳质酶完成甲壳类病原微生物细胞壁的“瓦解任务”, 接下来壳多糖刺激巨噬细胞增生, 使肝脏中大量巨噬细胞被激活, 再利用它吞噬消化和清除血中及经肠道吸收的微生物异物等有害物质来进行非特异性免疫。日本七鳃鳗肝脏内分解该活性物质的甲壳质酶高表达可能也与日本七鳃鳗特异的生活方式有关, 甲壳质广泛存在于自然界低等菌类、虾、蟹、昆虫等甲壳动物的外壳及高等植物的细胞壁中。我们知道仔七鳃鳗营泥砂中生活主要以浮游类为食物, 待到成体时大部分日本七鳃鳗为半寄生性动物, 既营寄生生活, 又能独立生活, 寄生时经常用吸盘附在其他鱼体上, 用吸盘内和舌上的角质齿锉破鱼体, 吸食其血与肉。营独立生活时, 则以浮游生物为食, 所以推测吸取血液中的病原微生物和浮游类生物可能是甲壳质的来源。

联系以上 MHC- I, MHC- II, C3, CFB, 干扰素、转铁蛋白、热激蛋白以及与免疫相关的蛋白和因子参与免疫应答的证据, 特别是转录组比较得出的甲壳质代谢、氨基葡萄糖代谢以及巨噬细胞激活的大量表达的结果, 可以初步推断日本七鳃鳗免疫中占有重要地位的可能是非特异性机制, 特有的免疫应答应与甲壳质代谢有关。此外根据日本七鳃鳗具有一定的适应性免疫能力以及存在适应免疫所必备的因子, 推测日本七鳃鳗可能已经进化到一个接近但不同于

高等有颌类适应性免疫的原始阶段.

4 结论

目前,除人类及少数的模式动植物、微生物之外,大多数生物的基因组测序工作尚未完成,日本七鳃鳗在进化上处于特殊的地位,所以构建 cDNA 文库具有相当重要的意义.本研究成功构建了日本七鳃鳗肝脏的 cDNA 文库,测序得到 10077 条 ESTs,这些序列基本能反映日本七鳃鳗肝脏组织表达的信息.通过对日本七鳃鳗肝脏 EST 的序列分析有三方面的发现:一是通过拼接、注释和功能分类揭示了日本七鳃鳗肝脏中蛋白的表达特点,并分析了蛋白高效表达的原因和机制;二是通过对 3' UTR 区的分析发现了与人类癌症基因调控相关的 microRNA 家族的靶位点,希望由此能为治疗人类各种癌症带来福音;三是通过与其他物种肝脏转录组的比较发现了一些日本七鳃鳗特异表达的功能基因,并着重分析了在日本七鳃鳗肝脏中比其他物种优势表达的甲壳质酶,进一步根据功能注释的信息推测日本七鳃鳗在免疫方面进化的阶段和特有的应答模式.我们实验室对日本七鳃鳗不同组织转录组做了系统的研究,有着重要的学术意义和应用前景,为下一步功能基因的克隆筛选和蛋白组学的研究奠定了基础.

参 考 文 献

- Janvier P. Palaeontology modern look for ancient lamprey. *Nature*, 2006, 443: 921—924[DOI]
- 张 凯, 郭文场, 李训德, 等. 日本七鳃鳗肌肉脂肪酸的组成分析. *中国水产科学*, 2000, 7(1): 116—117
- Mayer W E, Uinukool T, Tichy H, et al. Isolation and characterization of lymphocyte-like cells from a lamprey. *Proc Natl Acad Sci USA*, 2002, 99(22): 14350—14355[DOI]
- 高 琪, 逢 越, 吴 毓, 等. 日本七鳃鳗(*Lampetra japonica*)口腔腺表达序列标签(EST)分析. *遗传学报*, 2005, 32(10): 1045—1052
- Hardisty M W. The skeleton. In: Hardisty M W, Potter I C, eds. *The Biology of the Lampreys*. London: Academic Press, 1981, 333
- Murakami Y, Ogasawara M, Satoh N, et al. Compartments in the lamprey embryonic brain as revealed by regulatory gene expression and the distribution of reticulospinal neurons. *Brain Res Bull*, 2002, 57: 271—275[DOI]
- McCauley D W, Bronner-Fraser M. Importance of Sox E in neural crest development and the evolution of the pharynx. *Nature*, 2006, 441: 750—752[DOI]
- Takio Y, Pasqualetti M, Kuraku S, et al. Evolutionary biology: Lamprey Hox genes and the evolution of jaws. *Nature*, 2004, 416: 386—387
- Savina M V, Emelyanova L V, Belyaeva E A. Bioenergetic parameters of lamprey and frog liver mitochondria during metabolic depression and activity. *Comp Biochem Physiol B*, 2006, 145: 296—305[DOI]
- Rotchell J M, Steventon G B, Bird D J. Catalytic properties of CYP1A isoforms in the liver of an agnathan (*Lampetra fluviatilis*) and two species of teleost. *Comp Biochem Physiol C*, 2000, 125: 203—214
- Gamper N L, Savina M V. Reversible metabolic depression in hepatocytes of lamprey (*Lampetra fluviatilis*) during pre-spawning: Regulation by substrate availability. *Comp Biochem Physiol B*, 2000, 127: 147—154[DOI]
- Pancer Z, Amemiya C T, Ehrhardt G A, et al. Somatic diversification of variable lymphocyte receptors in the agnathan sea lamprey. *Nature*, 2004, 430: 174—180[DOI]
- 李旭霞, 逢 越, 肖 蓉, 等. 日本七鳃鳗(*Lampetra japonica*)口腔腺 cDNA 文库的构建. *辽宁师范大学学报*, 2004, 27(1): 73—75
- Huang X, Madan A. CAP3: A DNA sequence assembly program. *Genome Research*, 1999, 9: 868—877[DOI]
- Min X J, Butler G, Storms R, et al. TargetIdentifier: A webserver for identifying full-length cDNAs from EST sequences. *Nucleic Acids Res*, 2005, 33: 669—672[DOI]
- Lewis B P, Burge C B, Bartel D P. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell*, 2005, 120(1): 15—20[DOI]
- Zdobnov E M, Apweiler R. InterProScan—an integration platform for the signature-recognition methods in InterPro. *Bioinformatics*, 2001, 17(9): 847—848[DOI]
- Chen Z Z, Xue C H, Zhu S, et al. GoPipe: streamlined gene ontology annotation for batch anonymous sequences with statistics. *Biophys Prog Biochem*, 2005, 32(2): 187—191
- Chen Z Z, Wang W, Liu J J, et al. GODiff: mining functional differentiation between EST-based transcriptomes. *BMC Bioinformatics*, 2006, 7(1): 72[DOI]
- 王 玢, 左明雪. 人体及动物生理学. 第二版. 北京: 高等教育出版社, 2001
- 陈 萍, 高未平. 壳聚糖的保健功能. *中国食品卫生杂志*, 2001, 13(3): 38—39
- 龙 华. 水生动物 10 种非特异性免疫分子的研究进展. *长江大学学报*, 2005, 2(11): 67—72
- Good R A, Finstad J, Litman G W. The biology of lampreys II. *Immunology*, 1972, 1: 405—432
- Hagen M, Filosa M F, Youson J H. The immune response in adult sea lamprey (*Petromyzon marinus* L.): The effect of temperature. *Comp Biochem Physiol A*, 1985, 1: 207—210[DOI]
- Perey D Y, Finstad J, Pollara B, et al. Evolution of the immune response. VI. First and second set skin homograft rejections in primitive fishes. *Lab Invest*, 1968, 1: 591—597

- 26 Srivastava P K, Menoret A, Basu S. Heat shock proteins come of age: Primitive functions acquire new roles in an adaptive world. *Immunity*, 1998, 8: 657—665[DOI]
- 27 Srivastava P. Interaction of heat shock proteins with peptides and antigen presenting cells: chaperoning of the innate and adaptive immune responses. *Annu Rev of Immunol*, 2002, 20: 395—425[DOI]
- 28 Pontes M, Xu X, Graham D, et al. cDNA sequences of apolipoproteins from lamprey. *Biochemistry*, 1987, 26(6): 1611—1617[DOI]
- 29 陈廷贵. 22 周孕龄人胎肝转录组及 SARS-CoV(BJ-01)基因组的生物信息学研究. 中国人民解放军军事医学科学院. 博士论文, 北京: 军事医学科学院, 2005, 91—100
- 30 Urora E K. Oncomirs-microRNAs with a role in cancer. *Nat Rev Cancer*, 2006, 6(4): 259—269[DOI]
- 31 Volinia S, Calin G A, Liu C G, et al. A microRNA expression signature of human solid tumors defines cancer gene targets. *Proc Natl Acad Sci USA*, 2006, 103(7): 2257—2261[DOI]
- 32 吴松峰, 朱云平, 贺福初. 转录组与蛋白质组比较研究进展. *生物化学与生物物理进展*, 2005, 32(2): 99—105
- 33 Boot R G, Blommaert E F C, Swart E, et al. Identification of a novel acidic mammalian chitinase distinct from chitotriosidase. *J Biol Chem*, 2001, 276(9): 6770—6778[DOI]
- 34 Suzuki M, Morimatsu M, Yamashita T, et al. A novel serum chitinase that is expressed in bovine liver. *FEBS Lett*, 2001, 506(2): 127—130[DOI]
- 35 陈凌, 沈柱, 刘玉峰, 等. 小鼠 AMCase 基因突变体的 cDNA 克隆及序列分析. *解放军医学杂志*, 2006, 31(3): 231—233
- 36 Yang W, Gillian E, James M, et al. Human immune responses to infective stage larval-specific chitinase of filarial parasite, *Onchocerca volvulus*, Ov-CHI-1. *Filaria J*, 2003, 2(1): 6—20[DOI]
- 37 Nishimura K, Nishimura S, Seo H, et al. Macrophage activation with multi-porous beads prepared from partially deacetylated chitin. *J Biomed Mater Res*, 1986, 20(9): 1359—1372[DOI]
- 38 Xiao H. Immunology and antitumor activities of chitin/chitosan and its derivatives. *Mar Sci*, 1999, 3(3): 30—32