综述

鱼类基因数据库与生物信息学在鱼类基因开发上的应用

贝锦新　张　勇　李文笙　刘晓春　林浩然
(中山大学水生经济动物研究所,广州　510275)

APPLICATION OF GENOME DATABASES AND BIOINFORMATICS IN EXPLORATION FOR FISH GENES

BEI Jin-Xin, ZHANG Yong, LI Wen-Sheng, LIU Xiao-Chun and LIN Hao-Ran
(*Institute of Aquatic Economic Animals, Sun Yat-Sen (Zhongshan) University, Guangzhou* 510275)

关键词: 鱼类基因组; 同线性分析; 同源比对; 基因开发
Key words: Fish genome; Conservation of synteny; Homology search; Data mining
中图分类号: Q78　　文献标识码: A　　文章编号: 1000-3207(2008)03-0387-06

近十年来,随着分子生物学技术的发展,各国际组织间的合作研究使得资源基因组计划得以普遍实施。各种动物、植物基因组数据库均在互联网上陆续发布,生物信息量随之迅猛增长,然而如何利用并发掘这些数据蕴涵的宝藏,从中提取解释生命个体生长发育、免疫调节以及病害控制等机理的宝贵信息,已成为人类所面临的巨大挑战。因此,以获取、加工、储存、分配、分析和解读生物信息为手段,综合运用数学、计算机科学和生物学工具的交叉学科——生物信息学[1]由此诞生。也正是依靠生物信息学的辅助,人类基因组计划最终在 2003 年基本完成,基因组序列数据的拼接和组装也陆续完善。

各种鱼类大约占据了 50% 的脊椎动物种类,有着长远的进化史[2]。因此,利用这些丰富的鱼类资源,研究与系统发育和进化相关的种间保守或者特异的遗传和分子机制就显得相当重要。但相对于高等脊椎动物而言,鱼类的分子生物学研究比较滞后。在鱼类基因开发过程中,大部分是通过不同鱼类或者是其他高阶元物种的基因序列同源性,设计简并引物进行目的片段扩增;或者通过表达序列标签(Expressed sequence tags, EST)数据库发现并筛选鱼类功能基因。随着鱼类的基因组全序列测序拼接工作的不断完善,通过生物信息学和分子生物学的方法,新的鱼类功能基因逐渐被发现,这种技术路线采用了反向生物学的原理,即从基因到蛋白质再到功能研究的途径来发现新的生物活性分子[3]。国外科学家已经筛选到一批与发育、生殖及免疫相关的功能基因,而国内对于鱼类新功能基因的开发研究起步较晚,因此,我们应当运用生物信息学结合分子生物学的手段,发掘鱼类以及其他生物基因组数据库中的关键信息,开发并研究鱼类重要的功能基因,使得我国在这一领域能够有创新性的成果。为此,本文综述了目前鱼类基因组数据库的研究现状,并结合例子介绍利用这些数据库进行基因开发的几种方法。

1　鱼类基因公共数据库资源

1.1　斑马鱼 (*Danio rerio*, zebrafish)

斑马鱼隶属鲤形目(Cypriniformes)鲤科,由于个体小、周期产卵且产卵周期短、产卵量高、卵大、体外受精、胚胎透明、体外发育、胚胎早期发育快和易于大量获得样品等独有的特点而成为目前世界公认的模式脊椎动物之一。单倍体斑马鱼基因组由 25 条染色体组成,含有 1.7×10^3 Mb(1Mb= 10^6 个碱基对或 bp),大约为哺乳类基因组大小的二分之一[4]。2001 年 2 月开始,英国 Sanger 研究所开始了斑马鱼全基因组测序工作;主要通过 BAC(Bacterial artificial chromosome)、PAC(P1-derived artificial chromosome)文库的克隆测序和拼接,以及基因组鸟枪法测序(Whole genome shotgun sequencing)和组装两种策略。目前,该研究所和国际同行已经测序并拼接了 6653 条 DNA 片段,包含碱基约 1.6×10^3 Mb(参照: http://www.sanger.ac.uk/Projects/D-rerio/)。许多不同的服务器都提供了最新的斑马鱼基因组序列拼接数据,并能进行下载和比对(表 1)。同时,以美国华盛顿大学为代表的许多研究所也进行了不同规模的 EST 测序工作,这些丰富的斑马鱼 EST

收稿日期: 2006-10-25; 修订日期: 2007-06-07
基金项目: 国家重点基础研究发展计划(编号: 2006CB101805)资助
作者简介: 贝锦新(1977—),男,广东揭西人;博士;主要从事鱼类分子免疫学研究。E-mail: beijinxin@hotmail.com
通讯作者: 林浩然, E-mail: lsslhr@mail.sysu.edu.cn; Tel: 020-84113791

序列资源均能在互联网数据库(表 1)中进行比对或提取。其中最重要的数据库资源,是由美国俄勒冈大学(University of Oregon)负责维护的斑马鱼信息网络平台(The Zebrafish Information Network, ZFIN, 表 1)。它整合了大量的有关斑马鱼研究的信息资源,包括基因组数据、EST 数据、研究参考文献、研究人员和单位等。此外,它还提供斑马鱼突变体, EST 或 cDNA,以及相关单克隆抗体等资源。

表 1 主要的鱼类分子生物信息数据库互联网链接地址

Tab. 1 The URLs of several major fish genomic databases

维护单位(Service provider)	简要说明(Comments)	数据库互联网链接地址 (Uniform Resource Locator)
美国, NCBI (United States of America, National Center for Biotechnology Information)	斑马鱼基因组、鸟枪法测序片段、EST 等	http://www.ncbi.nlm.nih.gov/genome/seq/BlastGen/BlastGen.cgi?taxid=7955
	包括罗非鱼、fugu、黑青斑河鲀、三刺鱼和斑马鱼等各种鱼类的原始基因组、EST 测序片段,提供 MEGABLAST 以及 discontiguous MEGABLAST 两种比对方式	http://www.ncbi.nlm.nih.gov/Traces/trace.cgi?cmd=show&f=blast&m=blast&s=blast
	支持以染色体位点的形式查找和显示基因信息,可以选择不同的物种	http://www.ncbi.nlm.nih.gov/mapview/
	包括 blastp 等比对蛋白数据库工具	http://www.ncbi.nlm.nih.gov/BLAST/
英国(United Kingdom) The Sanger Institute	斑马鱼基因组测序拼接序列数据库	http://www.sanger.ac.uk/cgibin/blast/submitblast/d_rerio/advanced
ENSEMBL (European Bioinformatics Institute and The Sanger Institute)	欧洲 EBI 和 SANGER 研究所提供包括斑马鱼、fugu、黑青斑河鲀、三刺鱼, medaka 等鱼类、哺乳类以及其他物种的基因组序列信息查询和下载	http://www.ensembl.org/index.html
	提供最新递交的斑马鱼、斑点叉尾鮰、虹鳟鱼、多鳍鱼、七鳃鳗、大西洋鲑鱼、fugu 和黑青斑河鲀等鱼类原始基因组、EST 等测序序列信息	http://trace.ensembl.org/cgibin/tracesearch
美国 TIGR (USA, The Institute for Genomic Research)	提供妊丽鱼、大西洋鲑鱼、鲶鱼、fugu、朴丽鱼、底鲹、medaka、虹鳟、斑马鱼等鱼类的 EST 序列比对	http://tigrblast.tigr.org/tgi/
英国 MRC(UK, Medical Research Council)	提供 fugu 的基因组各个拼接序列不同的版本,原始序列以及 EST 序列信息	http://fugu.biology.qmul.ac.uk/blast/
法国 GENOSCOPE (France)	提供黑青斑河鲀鱼基因组全序列的比对工具	http://www.genoscope.cns.fr/externe/tetraodon/
美国 JGI (USA, DOE Joint Genome Institute)	Fugu 基因组序列拼接第四版数据库	http://genome.jgi-psf.org/cgi bin/runAlignment?db=Takru4&advanced=1
日本东京大学 Mbase (medaka genome database, university of Toky)	Medaka 基因组序列、EST 序列以及 BAC 文库序列数据库	http://mbase.bioweb.ne.jp/~dclust/medaka_top.html
美国 NIH(USA, National Institutes of Health)	提供斑马鱼 EST 序列数据库	http://zgc.nci.nih.gov/
日本遗传学研究所 (Japan, National Institute of Genetics)	Medaka 全基因组鸟枪测序片段、拼接片段数据库	http://dolphin.lab.nig.ac.jp/medaka/#about_project
美国华盛顿大学 (Washington University)	含有约 20 万条来源于不同组织或发育阶段的斑马鱼 EST 序列	http://www.genetics.wustl.edu/fish_lab/frank/cgi bin/fish/
美国俄勒冈大学 (university of Oregon)	The Zebrafish Information Network; 整合了大量的有关斑马鱼研究的信息资源,包括基因组数据、EST 数据、研究参考文献、研究人员和单位等	http://zfin.org/

1.2　青　(*Oryzias latipes*, Medaka)

青　是一种淡水颌针目(Beloniformes)鱼类; 体型小、生长周期短、繁殖力强、卵透明, 对于水中溶氧及温度的变化适应能力较强; 可作为环境、癌症以及发育等研究领域的一种实验对象[5]。青　的基因组由 24 对染色体组成, 含有碱基约(0.65–1) × 10^3 Mb; 其大小相当于人类基因组的三分之一、小于斑马鱼的二分之一[6]。日本科学家成功地获得许多自然的或者人为诱导的(Ethylnitrosourea, ENU)突变株; 最近, 他们成功地获得了透明青　突变株[7]。2002 年开始, 日本研究团体共同实施了青　基因组计划, 包括进行大规模 EST 测序、BAC 文库构建测序、全基因组鸟枪法测序以及基因遗传图谱制作等等。目前, 来源于不同青　cDNA 文库的将近 10 万条 EST 序列收录于日本东京大学管理的公共数据库(Mbase, 表 1)中, 为研究者提供比对分析工具。大约 90% 的基因组序列已经测定, 而拼接工作也在不断完善; 日本遗传学研究所(National Institute of Genetics)提供了有关的基因组序列数据库(表 1)。青　本身具有的特征以及目前取得的研究成果, 将逐渐使它成为一种新的模式脊椎动物。

1.3　红鳍东方　(*Takifugu rubripes*, fugu)

红鳍东方　, 俗称河　鱼, 为　形目(Tetraodontiformes)海水鱼类, 但对盐度变化具有较强地适应能力。河　体型较大, 繁殖周期比较长, 性成熟大约需要 3 年, 卵不透明[8]。其基因组大小为 400 Mb, 不到人类基因组的七分之一[9]; 虽然它的基因组比较小, 但作为一种脊椎动物, 它具有与哺乳类许多相同的生物学特征, 因此也具有许多相似的基因[10]。Brenner, *et al*.[9]研究发现河　鱼的基因结构比较紧凑(Compact), 也就是部分基因的内含子序列较短或少, 而且基因之间间隔较短; 并首次提出红鳍东方　可以作为脊椎动物基因研究的模式生物[9]。Elgar, *et al*.[11]以及 Aparicio, *et al*.[12]陆续报道了河　鱼的基因组测序和相关的拼接研究工作。在 Aparicio, *et al*. 的研究基础[12]上, 到 2004 年 8 月, JGI(The DOE Joint Genome Institute)以及其他研究组利用质粒、粘粒和 BAC 序列, 完成了大于 95% 覆盖率的基因组序列拼接工作。他们最新的数据库(表 1)提供了河　鱼基因组的 7213 条组装序列, 含有大约 393 Mb。伦敦大学玛丽皇后学院(Queen Mary, University of London)开发了利用河　鱼基因组序列信息以及大量的 EST 序列信息进行目的基因比对的生物信息学工具(表 1)。

1.4　黑青斑河　(*Tetraodon nigroviridis*, Spotted-green pufferfis)

2004 年, Jaillon, *et al*.[13]报道了黑青斑河　的基因组序列草图。黑青斑河　是　形目淡水鱼类, 体型小, 可在小型水族箱中饲养。它与红鳍东方　的亲缘关系很近, 由 18—30 百万年前的共同祖先进化而来[13]; 它的基因组大小为 350 Mb, 是已知基因组最小的脊椎动物。目前有 39 条超长拼接序列(Ultracontigs)已经定位到该种的 21 条染色体上, 覆盖了基因组的 64.2% (参照: http://www.genoscope.cns.fr/externe/tetraodon/)。此外, Jaillon, *et al*.[13]也提供了大量的 EST 序列信息, 可以通过 GENOSCOPE、ENSEMBL 或者 NCBI 数据库工具进行比对和提取(表 1)。

1.5　其他种类

此外, 以美国、英国、日本等国为主的基因组中心分别开展了各种鱼类的基因组和 EST 测序工作。目前, 包括刺鱼形目的三刺鱼(*Gasterosteus aculeatus*, Three-spine stickleback)、鲶形目的斑点叉尾　(*Ictalurus punctatus*)、鲑形目虹鳟鱼(*Oncorhynchus mykiss*, Rainbown trout)和大西洋鲑鱼(*Salmo salar*)、多鳍鱼目多鳍鱼(*Polypterus bichir*)、鲈形目的妊丽鱼(*Astatotilapia burtoni*)和朴丽鱼(*Haplochromis* sp. red tail sheller)、　形目的底　(*Fundulus heteroditus*, killifish)以及圆口纲七鳃鳗(*Petromyzon marinus*)在内的许多鱼类基因序列信息, 都能在 TGI、ENSEMBL 或者 NCBI 等公共数据库服务器中得以查询提取(表 1)。随着这些种类基因序列信息的不断完善, 将为比较基因组学和比较生物学提供极其重要的平台。

2　利用基因数据库开发鱼类新基因

一般而言, 鱼类基因的克隆以及功能研究要比高等哺乳类的相关研究滞后许多。近几年, 得益于鱼类基因组数据库的不断更新完善, 从事鱼类研究工作的科学家开始重视这些数据库的功能, 并利用它们开发新型的鱼类基因。根据目前的报道, 利用基因数据库开发新基因的方法大致有两种, 包括基因编码蛋白序列相似性的同源比对法和基因组片段同线性进化现象的同线性分析法。以下分别列举一个例子, 介绍如何运用这两种方法。

2.1　同源比对

物种间基因的比较表明, 生物在进化过程中, 不同种属之间的同一基因的编码序列以及基因结构具有保守性, 表现为一定的同源性。根据这个特点, 可以利用已知物种基因的编码氨基酸序列在鱼类基因数据库中搜索候选基因, 然后运用生物信息学结合分子生物学的方法, 筛选并克隆鱼类的同源目的基因(Homologue)。主要过程(图 1), 以最近 Kaslin, *et al*.[14]报道的鱼类开胃素(Orexin)基因的克隆为例子。首先, 将同源的人 *orexin* 基因氨基酸序列输入斑马鱼基因数



图 1　同源比对方法开发鱼类新基因的策略

Fig. 1　The strategy of homology searching for fish genes in genomic database

据库(SANGER, 表 1)比对框, 选择TBLASTN 程序, E 值(Expect) = 1 进行查询(图 2)。图 3 结果显示有一个含有候选基因片段的拼接序列, 选择并下载含有该候选基因的邻接序列(Contig)。然后, 利用 GENSCAN 软件[15] (http://genes.mit.edu/GENSCAN.html)对该序列进行基因预测并将预测基因输入 NCBI GenBank 数据库进行 blastp 比对; 其中, 得到具有一段 orexin 蛋白结构(图 4)的预测基因。接着, 以这段预测的 cDNA 为基础, 设计特异引物进行全长 cDNA 克隆, 或者采取 3′和 5′- RACE(Rapid amplification of cDNA ends)的方法克隆目的基因。最后, 通过蛋白的氨基酸序列分析、同源基因的基因结构和表达模式相似性或者结合功能分析去验证该目的基因[14]。经比较, 预测的斑马鱼 *orexin* 基因包含所有编码成熟肽的 cDNA 序列, 说明这种方法的准确性比较高。目前, 有许多新的鱼类基因是通过这种方法发现的, 如 SOCS 3[16], II 型细胞因子及受体[17]等。

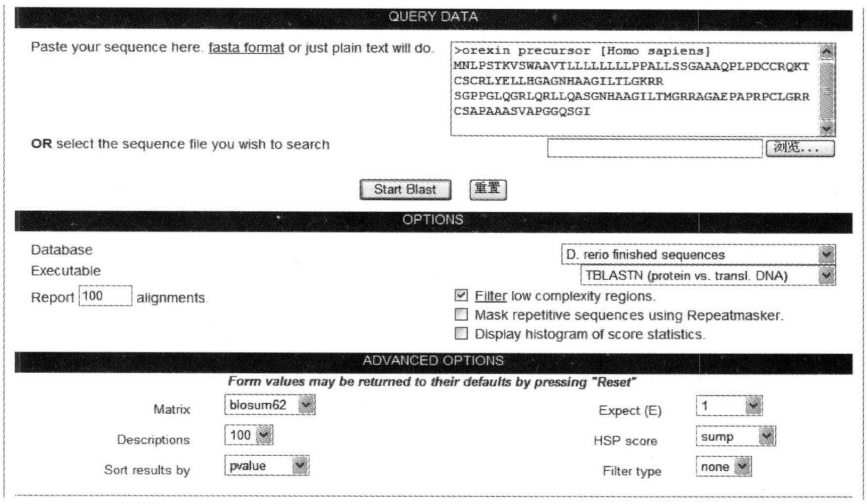


图 2 斑马鱼基因组数据库查询界面(互联网链接见表 1)

Fig. 2 The query interface of zebrafish genome database(See Tab. 1 for URL)

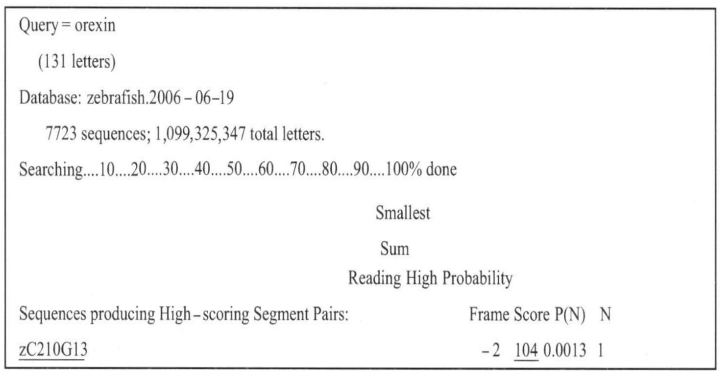


图 3 斑马鱼基因组数据库查询部分结果

Fig. 3 Results for the queried sequence. The underlined zC210G13 is the name and link to a contig

zC210G13 为候选拼接序列名称, 单击进入下载页面, P(N)值一般情况下取最低值。Frame 栏表示目的基因编码方向及起始位点, -2 即表示基因方向与拼接序列方向相反, 在 2 号位点起始编码

zC210G13 is the candidate contig, and could be downloaded via the link. The lowest P(N) value stands for highest probability. The -2 in Frame column indicated that the orientation of gene coding was reverse to the candidate contig sequence and the codon started at the second position

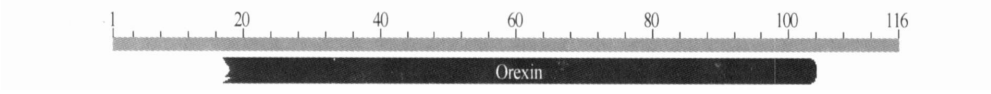


图 4 预测基因 BLASTP 比对 NCBI 蛋白数据库结果

Fig. 4 Results for the predicted protein using BLASTP and protein database in NCBI

上方数字标尺为预测基因编码蛋白长度, 下方为 orexin 蛋白结构及相对应区域

Ruler above showed the length of predicted protein and below showed the known protein motif

2.2　同线性分析

同源比对的方法适用于保守程度比较高的基因, 但对于一些高变异的基因, 如大部分的白细胞介素和一些激素类[18, 19], 就找不到候选片段。

比较基因组学研究发现, 处于不同进化地位的生物, 他们当中的许多单个基因不但在编码序列、调控序列以及基因结构上具有保守性, 而且基因间的位置关系也具有一定的保守性。也就是说, 一些连锁的基因在进化历程中, 总是组成相对稳定的片段, 在这些片段中的基因线性排列顺序以及结构高度相似, 这种现象被称为"保守同线性(Conserved synteny)"[10, 13, 20]。物种间分化时间越短, 同线性基因越多, 区域越长, 同线性程度越高。例如人和鼠之间含有 417 个大于 300 kb 的同线性区域[20], 而人和河　鱼享有 900 个含有 2—55 个基因的同线性区域[13]。生物体基因组进化的这种现象, 为克隆某些高变异基因如大多数的细胞因子提供了一种途径。最近几年, 这种方法被广泛用于开发鱼类的新基因, 如白细胞介素-2、-10、-15、-21, leptin 等[19, 21—23]。具体的策略(图 5), 以河　鱼白细胞介素(IL)-15 的基因克隆[21]为例子。首先, 通过 NCBI Map viewer(表 1)确定人类四号染色体上的 IL-15 基因以及其两侧基因的顺序和方向(图 6); 同样, 也能确定鸟类的同源基因, 发现 IL-15 基因区域存在同线性(图 6)。然后, 我们通过同源比对方法, 分别在河　鱼数据库中找到除了 IL-15 外的其他几个同源基因, 而且发现它们呈同线性分布在几个相邻的拼接序列中。第三步, 下载河　鱼 ZNF330 和 INPP4B 基因之间的 DNA 序列并运用 GENSCAN 进行基因预测; 经 BLASTP 比对 NCBI 蛋白数据库, 结果显示, 该序列编码一个与 IL-15 同源的基因。第四步, 根据预测的 cDNA 设计特异性引物, 分别进行 3′和 5′-RACE, 得到编码 IL-15 的全长 cDNA。最后, 根据氨基酸序列相似性、类似的蛋白高级结构、相同的基因结构以及相似的基因表达模式等特点, 我们认为该候选基因为人类 IL-15 的直向同源基因[21]。

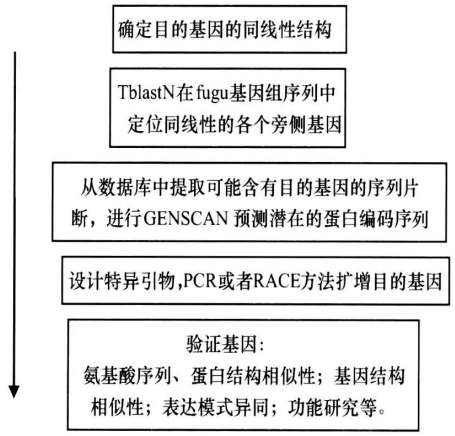


图 5　同线性分析法开发鱼类新基因的策略

Fig. 5　The strategy of conserved synteny analysis of fish genes in genomic database

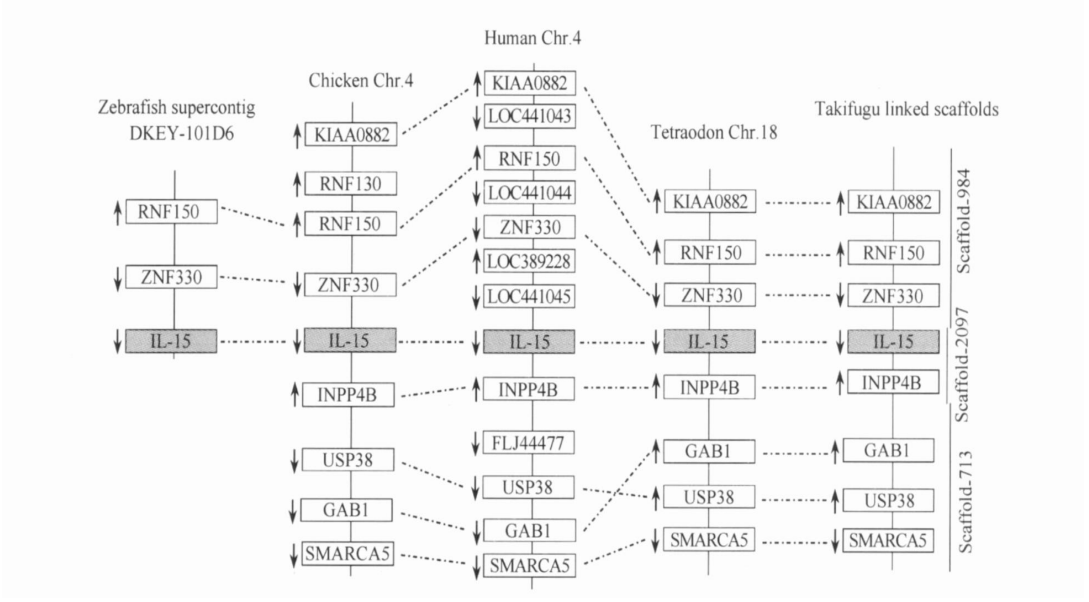


图 6　IL-15 及旁侧基因的同线性

Fig. 6　The conservation of synteny of IL-15 loci

方框内为各个基因缩写, 箭头表示基因在基因组片段中的方向, 虚线表示同源基因(摘自文献[21])

The genes were boxed, the orientation was shown by arrow, and the homologues were linked by dashes (Reference[21])

3　总结与展望

在进行鱼类新基因开发时, 首先可以通过查询各种鱼类 EST 数据库, 筛选同源的基因序列[1]。如果这种方法未能得到充分的数据, 可以结合本文中的两种方法, 找到模式鱼类的同源目标基因。对于那些没有充分的 EST 或者基因组序

列信息的实验鱼类,可以通过对以上几种模式鱼类同源基因的比对,确定这些基因是否存在保守的功能区域,据此降低引物或者探针设计的简并程度,提高特异性,然后进行 PCR 或者文库筛选克隆那些鱼类的同源基因。如果同源性不高以至于不能设计简并引物,那么可以利用基因组片段的同线性现象,选用该同源基因旁侧(最好是两侧)保守基因的编码序列:设计 PCR 引物,尝试反向 PCR 的方法对基因组进行步移[24];或者标记作探针,去筛选目的鱼类基因组文库中对应于该区域的基因组片段,将该片段测序,再进行基因预测,找到该区域的目的基因。

另外,在查找某些基因的时候,如果出现两个以上的候选序列,可以根据同源蛋白功能特征序列对每个预测基因进行初步筛选。一个同源基因,也可能在鱼类出现两个或以上拷贝,因为目前的研究说明了鱼类分支在进化过程中发生了一次全基因组复制(Whole genome duplication)[13]。例如,我们还发现一种鱼类特有的 IL-15 类似基因(IL-15L),但在基因表达模式上与 IL-15 不同,具体功能上的差异,仍有待于进一步的功能研究[21]。

总之,这里介绍的方法,为鱼类的基因开发提供了一种途径,但对于一些特殊的高变异基因(如白细胞介素-3,-4,-15等),通过同源性比对的方法未能发现;而且同线性分析也同样未能开发这些基因,这可能受限于基因组序列信息的不完整性,因此,有待鱼类基因组数据库的不断更新和完善。随着更多物种的 EST 和基因组测序计划的实施,鱼类分子生物学数据库将不断丰富起来,我们将可以更有效地利用生物信息学工具、比较基因组学等研究方法去开发与鱼类生殖、生长、免疫调控相关的功能基因。

参考文献:

[1] Zhao G P. Bioinformatics [M]. Beijing: Academic Press. 2002, 3—12 [赵国屏. 生物信息学. 北京: 科学出版社. 2002, 3—12]

[2] Nelson J S. Fishes of the world [M]. New York: John Wiley and Sons Inc. 1994, 2

[3] Liu H T. Cloning and the strategy of new cytokines [J]. *International Journal of Immunology*, 1995, **5**: 256—261 [刘红涛. 近年新细胞因子的克隆及其技术路线. 国外医学(免疫学分册), 1995, **5**: 256—261]

[4] Postlethwait J H. Evolution of the Zebrafish Genome [A]. In: Gong Z, Korzh V (Eds.), Fish Development and Genetics [C]. World Scientific, 2004, 581—611

[5] Mitani H, Shima A, Naruse K, *et al*. Medaka genome mapping to functional genomics [A]. In: Gong Z, Korzh V (Eds.), Fish Development and Genetics [C]. World Scientific, 2004, 612—636

[6] Wittbrodt J, Shima A, Schartl M. Medaka: a model organism from the far East [J]. *Nat Rev Genet*, 2002, **3**: 53—64

[7] Wakamatsu Y, Pristyazhnyuk S, Kinoshita M, *et al*. The see-through medaka: a fish model that is transparent throughout life [J]. *Proc Natl Acad Sci U S A*, 2001, **98**: 10046—10050

[8] Sun Z Z. Biological characters and artificial breeding of pufferfish [J]. *Shandong Fisheries*, 2002, **19**: 44—46 [孙中之. 红鳍东方 的生物学特性及人工育苗技术. 齐鲁渔业, 2002, **19**: 44—46]

[9] Brenner S, Elgar G, Sandford R, *et al*. Characterization of the pufferfish (Fugu) genome as a compact model vertebrate genome [J]. *Nature*, 1993, **366**: 265—268

[10] Liu S M, Yang L. Pufferfish: An ideal model for genetic studies of vertebrate [J]. *Int J Genet*, 1998, **21**: 1—4 [刘山鸣, 杨立. 脊椎动物基因研究的理想模式生物-河豚鱼. 国外医学遗传学分册, 1998, **21**: 1—4]

[11] Elgar G, Clark M S, Meek S, *et al*. Generation and analysis of 25 Mb of genomic DNA from the Pufferfish *Fugu rubripes* by sequence scanning [J]. *Genome Research*, 1999, **9**: 960—971

[12] Aparicio S, Chapman J, Stupka E, *et al*. Whole-Genome Shotgun Assembly and Analysis of the Genome of *Fugu rubripes* [J]. *Science*, 2002, **297**: 1301—1310

[13] Jaillon O, Aury J M, Brunet F, *et al*. Genome duplication in the teleost fish *Tetraodon nigroviridis* reveals the early vertebrate proto-karyotype [J]. *Nature*, 2004, **431**: 946—957

[14] Kaslin J, Nystedt J M, Ostergard M, *et al*. The orexin/hypocretin system in zebrafish is connected to the aminergic and cholinergic systems [J]. *J Neurosci*, 2004, **24**: 2678—2689

[15] Burge C, Karlin S. Prediction of complete gene structures in human genomic DNA [J]. *J Mol Biol*, 1997, **268**: 78—94

[16] Jin H J, Shao J Z, Xiang L X. Identification and characterization of suppressor of cytokine signaling 3 (SOCS-3) homologues in teleost fish [J]. *Mol Immunol*, 2007, **44**: 1042—1051

[17] Lutfalla G, Crollius H R, Stange-Thomann N, *et al*. Comparative genomic analysis reveals independent expansion of a lineage-specific gene family in vertebrates: the class II cytokine receptors and their ligands in mammals and fish [J]. *BMC Genomics*, 2003, **4**: 29

[18] Harrison G, Wedlock D. Marsupial cytokines: structure, function and evolution [J]. *Dev Comp Immunol*, 2000, **24**: 473—484

[19] Kurokawa T, Uji S, Suzuki T. Identification of cDNA coding for a homologue to mammalian leptin from pufferfish, *Takifugu rubripes* [J]. *Peptides*, 2005, **26**: 745—750

[20] Bourque G, Pevzner P A, Tesler G. Reconstructing the Genomic Architecture of Ancestral Mammals: Lessons From Human, Mouse, and Rat Genomes [J]. *Genome Res*, 2004, **14**: 507—516

[21] Bei J X, Suetake H, Araki K, *et al*. Two interleukin (IL)-15 homologues in fish from two distinct origins [J]. *Mol Immunol*, 2006, **43**: 860—869

[22] Bird S, Zou J, Kono T, *et al*. Characterization and expression analysis of interleukin 2 (IL-2) and IL-21 homologues in the Japanese pufferfish, *Fugu rubripes*, following their discovery by synteny [J]. *Immunogenetics*, 2004, **56**: 909—923

[23] Zou J, Clark M S, Secombes C J. Characterisation, expression and promoter analysis of an interleukin 10 homologue in the puffer fish, *Fugu rubripes* [J]. *Immunogenetics*, 2003, **55**: 325—335

[24] Jong A Y, Tang A, Liu D P, *et al*. Inverse PCR: Genomic DNA Cloning [A]. In: Chen B Y, James H W (Eds.), PCR Cloning Protocols (2nd) [C]. New Jersey: Humana Press. 2002, 301—307