

文章编号: 1000-0615(2005)05-0599-07

表达序列标签数据库搜索克隆斑马鱼 QM 基因 及其数字化差异显示分析

金红建, 邵健忠, 项黎新
(浙江大学生命科学院, 浙江 杭州 310012)

摘要: QM 基因是一种与肿瘤抑制、细胞生长、分化、发育和凋亡等有关的重要基因,已在人类及许多物种中得到了克隆,但鱼类中尚无报道。我们利用表达序列标签数据库和电子克隆等技术,成功克隆了斑马鱼 QM 基因,并对该基因的结构进行了系统分析。结果显示,斑马鱼 QM 基因 cDNA 全长 769 bp,含 648 bp 开放阅读框架,编码 215 个氨基酸,5' UTR 25 bp,3' UTR122 bp,所编码的 QM 多肽分子量 24.6 kDa,等电点(pI) 10.6,含潜在的蛋白酶 C 磷酸化位点、N-乙酰化位点和酰胺化位点。比较斑马鱼 QM 和人等 13 个物种 QM 的同源性,发现其氨基酸序列相似性为 66%~92%。数字化差异显示分析结果表明,QM 基因在斑马鱼胚胎及成体多种组织中均有广泛表达。研究结果为今后利用生物信息学快速克隆新的鱼类功能基因打下了方法学基础,也为进一步研究鱼类 QM 基因功能提供了依据。

关键词: 斑马鱼; QM 基因; 表达序列标签; 电子克隆; 数字化差异显示分析

中图分类号: S917 文献标识码: A

In silico cloning and digital differential display analysis of *Danio rerio* QM gene by EST database searching

JIN Hong-jian, SHAO Jian-zhong, XIANG Li-xin
(College of Life Science, Zhejiang University, Hangzhou 310012, China)

Abstract: QM homologs, associated with tumor suppression, cell growth, differentiation, development and apoptosis, were isolated from human and many other organisms. However, QM derived from fish was not reported. Through integrated analysis of public database such as EST, TIGR Gene Index in silico, a full-length cDNA encoding QM homologue of human QM protein was identified from *Danio rerio*, which is 769 bp length and has a putative protein of 215 amino acids with a predicted mol. wt. of 24.6 kD and isoelectric point of 10.6. It was found that *Danio rerio* QM amino acid sequence includes 1 Ribosomal protein L10e signature, 2 Amidation sites, 1 Tyrosine sulfation site, 1 N-myristoylation site and 1 Protein kinase C phosphorylation site by scanning PROSITE motifs. The *Danio rerio* QM shares an 66% - 92% amino acid sequence identical to other proteins of the QM family, including human etc, respectively. Study showed that *Danio rerio* QM gene expressed in all *Danio rerio* embryo and adult tissues by digital differential display based on UinGene. The experiment provided us with a feasible bioinformatics method for cloning new functional genes, thereby preparing for further research on fish QM gene function.

Key words: *Danio rerio*; QM gene; expressed sequence tag (EST); in silico cloning; digital differential display

QM 基因是 Weissman 等首次从人维尔姆斯肿瘤(Wilms' tumor)相关的细胞系中分离的一种新基因^[1],随后发现该基因在自然界许多物种中

广泛存在,除了人类外,目前已在小鼠、鸡、果蝇、蚕、蛾、酵母、玉米和水稻等生物中克隆了 QM 同源基因,这些基因的核苷酸及其编码的氨基酸序

收稿日期: 2004-09-09

资助项目: 国家自然科学基金(30371096, 30571423); 浙江省自然科学基金(301307); 高等学校博士学科点专项基金资助(2003035040)

作者简介: 金红建(1976-),男,浙江永康人,博士研究生,主要从事鱼类分子免疫学研究。Tel: 0571-88273287, E-mail: jinjhj@zju.edu.cn

通讯作者: 邵健忠, Tel: 0571-88273287, E-mail: Shaojz@zju.edu.cn

列高度保守^[2]。在人类中,维尔姆斯肿瘤是一种在儿童期高发的肾母细胞瘤,QM基因被认为与该肿瘤的抑制有关^[3];在鸡中,QM同源基因Jif-1可与癌基因e-Jun的蛋白产物结合从而抑制e-Jun的反式激活转录调节活性,影响细胞生长与增殖^[4];在酵母中,QM同源基因GRC5/QSR1与线粒体呼吸作用有关,而且QM基因突变,可导致酵母蛋白质合成下降、生长和分裂阻滞、细胞骨架异常,QM基因缺失,则导致酵母死亡^[5];在小鼠胚胎发育过程中,QM参与基因转录和转译后的调节,与组织细胞分化关系密切^[6]。另外还有研究表明,QM是核糖体60S大亚基的一个外周蛋白,参与蛋白质特别是膜蛋白的合成与运输^[7];此外,QM还与e-Yes和其它Src家族成员相互作用,参与细胞信号的转导^[8]。目前的研究显示,QM是一种多功能蛋白,参与细胞生长、分化、发育和凋亡等基本生命活动^[9]。然而目前在鱼类中尚无QM基因的相关报道,我们利用生物信息学资源和技术,通过斑马鱼表达序列标签数据库和电子克隆方法,首次获得了斑马鱼(*Danio rerio*)QM基因(DrQM)全长cDNA序列,并且用数字化差异显示方法分析了DrQM在斑马鱼各种组织中表达情况,为今后利用生物信息学快速克隆新的鱼类功能基因打下方法学基础,也为进一步揭示鱼类QM基因的功能及其分子进化规律提供科学依据。

1 材料与方法

1.1 主要数据库和软件

主要数据库为美国国立生物技术信息中心(National Center for Biotechnology Information, NCBI) GenBank数据库(<http://www.ncbi.nlm.nih.gov/Database/index.html>)、EST数据库和UniGene数据库以及美国基因组研究所(The Institute of Genome Research, TIGR) TGI数据库(<http://www.tigr.org/tdb/tgi/>);核酸序列同源性分析软件使用BLASTN 2.2.2(<http://www.ncbi.nlm.nih.gov/BLAST/>), NCBI, USA)和WU-BLAST 2.0(<http://blast.wustl.edu/>), the Washington University BLAST);将所选EST组装为连续体软件为The Cap Sequence Assembler at IFOM(<http://bio.ifom-firc.it/ASSEMBLY/assemble.html>)或ESTmachine(<http://www.tigem.it/ESTmachine.html>);开放阅读框采用ORF Finder(NCBI, USA)核酸序列软件寻找;序列匹配分析软件用Clustal W 1.8(EMBL, Germany)和GeneDOC。蛋白质序列分析用DS Gene1.5、DNASTar5.0等软件包;信号肽预测用SignalP 3.0(<http://www.cbs.dtu.dk/services/SignalP/>)。

tigem.it/ESTmachine.html);开放阅读框采用ORF Finder(NCBI, USA)核酸序列软件寻找;序列匹配分析软件用Clustal W 1.8(EMBL, Germany)和GeneDOC。蛋白质序列分析用DS Gene1.5、DNASTar5.0等软件包;信号肽预测用SignalP 3.0(<http://www.cbs.dtu.dk/services/SignalP/>)。

1.2 DrQM基因的电子克隆与结构分析

在NCBI GenBank数据库中,以人的QM基因mRNA序列(accession# M64241)为查询序列对斑马鱼dbEST数据库进行Blastn分析,将检索到的匹配结果的E值小于10⁻²⁰的高度同源性的斑马鱼EST序列进行EST重叠群构建(计算机组装),在每个碱基至少经过两条以上EST序列验证后,用所得到的重叠群搜索TIGR斑马鱼基因索引(zebrafish gene index, ZGI)数据库,最后拼排得到基因的潜在全长cDNA。组装条件为Overlap Length > 35, Overlap Percent Identity > 80%。获得DrQM基因全长cDNA后,对其开放阅读框进行预测,并推导出DrQM氨基酸序列。然后比较DrQM和其它物种QM氨基酸序列相似性,构建系统进化树,并运用生物信息学软件,对DrQM蛋白的结构域、功能位点、氨基酸组成、等电点、分子量、电荷分布、亲水性、信号肽和二级结构等进行分析。

1.3 DrQM基因的数字化差异表达分析

以上述获得的DrQM基因全长cDNA为查询序列,对斑马鱼EST数据库进行BLASTN分析,将获得的一致性大于95%,E值小于10⁻²⁰的EST序列确认为是代表DrQM基因的EST,对每条EST来源的cDNA文库进行分析,可以得到文库构建的详细情况,其中包括组织来源、发育阶段、宿主菌、载体、克隆位点和提交者信息等。属于同一UniGene重叠群的EST,都可以通过数字化差异显示进行分析,计算出源于各种不同组织的cDNA文库中代表DrQM基因的EST数量,就可以反映出该基因在这几种组织中的表达水平。

2 结果

2.1 DrQM基因的电子克隆与结构分析

用人Wilm's tumor-related protein(QM)mRNA(accession# M64241)全序列BLASTN斑马鱼EST数据库,发现共有45条长度不一的斑马鱼EST与之高度同源(相似性80%以上,图1)。将

这些 EST 序列进行计算机拼排, 得到长度为 769 bp 的重叠群(记为 zebrafish QM contig1)。用它作为查询序列, 用 WU-BLAST 2.0. 程序搜索 TIGR 的斑马鱼基因索引, 发现与其中的 TC73541 索引有 99% 的序列相似性。TC73541 重叠群由 63 个 EST 片段组成, 它们来自于多种组织表达的基因片段。然后用 CAP3 程序对 zebrafish QM contig1 和 TC73541 进行组装, 得到含有一个最长为 648 bp 完整开放阅读框架和 AATAAA 加尾信号的斑马鱼全长 QM 基因 cDNA 序列, 称之为 DrQM (图 2, 图 3)。DrQM 基因全长 769 bp, 5' UTR 长度为

25 bp, 第 26~ 673 碱基是开放阅读框(ORF), 长 648 bp, 编码 215 个氨基酸, 3' UTR 长度为 122 bp, AATAAA 加尾信号出现在终止密码子下游 49 个核苷酸处。在加尾信号下游 11 个核苷酸处是多聚腺苷酸尾(图 3), 这与 AATAAA 加尾信号通常出现在多聚腺苷酸尾上游 11~ 30 个核苷酸处的现象相符合^[9]。DrQM 基因编码的多肽, 预测其分子量大小为 24.6 kDa, 等电点 pI 为 10.6, 其氨基酸组成特点是酸性氨基酸(Asp 和 Glu) 残基数为 7(7.9%), 碱性氨基酸(Arg 和 Lys) 残基数为 45(20.92%), 表明 DrQM 蛋白为碱性蛋白质。

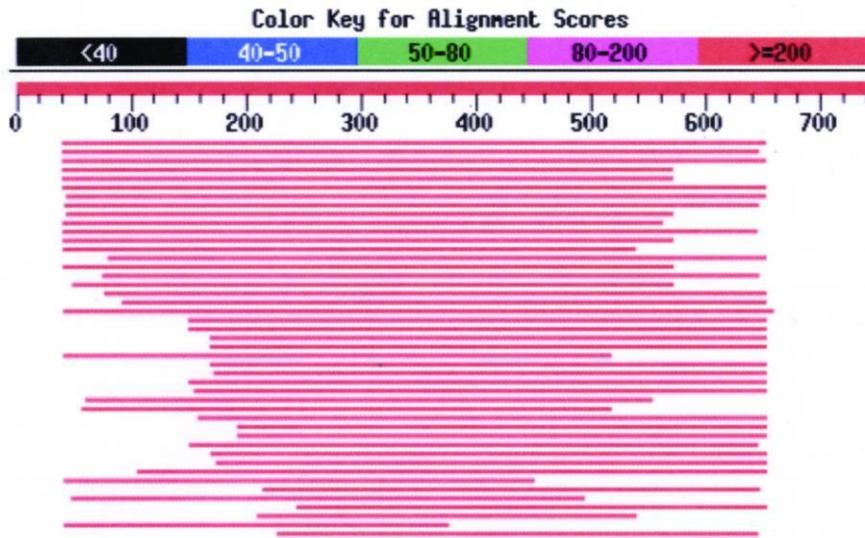


图 1 用人类 QM 基因搜索斑马鱼 EST 数据库结果

Fig. 1 Alignment of human QM mRNA in *Danio rerio* EST database

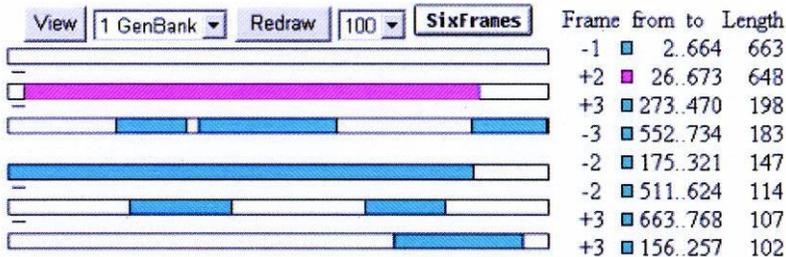


图 2 DrQM 基因 cDNA 全序列 ORF 预测

Fig. 2 ORF prediction of the assembled DrQM cDNA

从 GenBank 中检索到 13 种生物 QM 同源基因序列, 并用 Clustal X 软件进行比对, 这些序列和 DrQM 氨基酸序列相似性分别为: 人 *Homo sapiens* (M64241) 92%, 家鼠 *Mus musculus* (X75312)

91%, 牛 *Bos taurus* (AF143815) 91%, 鸡 *Gallus gallus* (Q08200) 90%, 珍珠贝 *Pinctada fucata* (AY156926) 76%, 菱形斑纹蛾 *Plutella xylostella* (AB180439) 77%, 家蚕 *Bombyx mori* (AF395838)

77%, 果蝇 *Drosophila melanogaster* (U49721) 77%, 酵母 *Schizosaccharomyces pombe* (U33214) 66%, 火炬松 *Pinus taeda* (AF013804) 68%, 水稻 *Oryza sativa*

(X81691) 67%, 玉米 *Zea mas* (U06108) 67% 和非洲油棕榈 *Elaeis guineensi* (AF295636) 69% (图4), 可见, QM 基因在各物种间比较保守。

```

1 ACCACCGCCGTCGAGATCGCAACCATGGGCCGCCGTCACGCCGTTGTTACAGATACTG
      26 M G R R P A R C Y R Y C
61 TAAGAACAAGCCTTACCCCAAGTCCCGTTTCTGTCGAGGTTCCTGATCCCAAAATCAG
      K N K P Y P K S R F C R G V P D P K I R
121 GATCTTCGACTTGGGCAGGAAGAAGGCTAAGGTGGATGAGTTCCCTGTGTGTCACAT
      I F D L G R K K A K V D E F P L C A H M
181 GGTGTCGTATGAGTATGAGCAGCTGTCCAGTGAAGCTCTTGGAGGCTGCCGTATCTGCGC
      V S D E Y E Q L S S E A L E A A R I C A
241 CAACAAGTATAATGGTGAAGACTTCCGGTAAAGATGGCTTCCACATCCCGTGGCCCTGCA
      N K Y M V K T C G K D G F H I R V R L H
301 CCCCTTCCACGTCATCCGCATCAACAAAATGTTGTCTGTGCCGGAGCTGATAGGCTCCA
      P F H V I R I N K M L S C A G A D R L Q
361 GACTGGTATGCGTGGTGGTTCGGGAAGCCTCAGGGCACAGTGGCCCGTGTGCACATTGG
      T G M R G A F G K P Q G T V A R V H I G
421 CCAGGTCATCATGTCGTGCGCACAAAGGCCAGAACAAAGGAGCATGTGATGAGGCTCT
      Q V I M S V R T K A Q N K E H V I E A L
481 GAGGAGAGCCAAGTCAAGTTCCTGGCCGTCAGAAGATCCACGTCCTCAAGAAGTATGG
      R R A K F K F P G R Q K I H V S K K Y G
541 CTTACAAAAGTCAACACGTTGACTTTGATAACATGCTGGCGGAGAGCGTCTGATTCC
      F T K F N T C D F D N M L A E K R L I P
601 TGATGGCTGCGGGGTGAATAACATCCCCAGCCGTGGACCCCTGAGCCGCTGGAAGGCCCT
      D G C G V K Y I P S R G P L S R W K A L
661 TCATGCCAACTTAAGCACCTCACACTTGCTTTTCGTTTGTGAAAAAGGCCCAAGGACCA
      H A N * 673
721 CAAATAAAGGCCATTCTTTGAAAAAAAAAAAAAAAAAAGGCCGCGCGT

```

图3 DrQM 基因序列及其推导的氨基酸序列

Fig. 3 Nucleotide and deduced amino acid sequence of the assembled DrQM

用 DNASTar 软件建立了包括斑马鱼在内的 14 个物种的 QM 分子进化树, 蛋白质多序列比对采用基于氨基酸相对突变率矩阵 PAM250 的 Jotun Hein 方法。来源相同的物种在进化树上往往聚集成簇, 序列间的遗传距离对应于进化树中连接两个物种水平线的长度。结果显示, QM 基因可归为 4 大类, 包括脊椎动物、无脊椎动物、酵母和植物, 其中 DrQM 和鸡 QM 在进化上的距离最短(图5)。

通过搜索 PROSITE 数据库, 获得了 DrQM 多肽序列中可能包含的结构域和功能位点, 包括 2 个蛋白酶 C 磷酸化位点 SVR(137~139) 和 SKK(168~170), 1 个 N-乙酰化位点 GMRGAF(114~119), 2 个酰胺化位点 MGRR(1~4) 和 LGRK(36~39), 此外还存在 1 个核糖体蛋白 L10 特征序列 RGAFGKPGTVARV(116~129)。

根据 Kyte 和 Doolittle 的方法^[10], 对 DrQM 多肽的疏水/亲水性进行了分析, 结果显示, 其疏水性平均值(average hydrophobicity)为 -0.53, 和其它已知蛋白比较, DrQM 可能是一种可溶性蛋白。

用 SignalP 3.0 程序, 对 DrQM 多肽序列中信号肽的剪切位点进行了预测, 分析结果表明, DrQM 多肽 N-端不存在信号肽, 说明它不是一种分泌蛋白。此外, 用 Zhou 和 Fasman 的算法^[11], 预测了 DrQM 蛋白的二级结构, DrQM 有 6 个 α 螺旋区, 富含 10 个 β 折叠片以及 6 个转角区。

2.2 DrQM 基因的数字化差异表达分析

以 DrQM 基因的全长 cDNA 对斑马鱼 EST 数据库进行 Blastn 搜索, 获得 178 条高度同源的 EST。其中 168 条 EST 来自于 9 种不同组织的 cDNA 文库, 31 条 EST 来自于胚胎时期组织, 其余 137 条 EST 源于成体斑马鱼组织。各种组织中每百万转录本中的 DrQM 基因 EST 数分别为: 脑 1868, 鳍 1982, 胚胎 541, 睾丸 1083, 卵巢 307, 心脏 1537, 嗅觉器官 168, 肾 228, 肝 1773, 未知组织 928。数字化差异表达分析显示, QM 基因在脑和鳍中有较强表达, 在其余 7 种组织中也均有表达(图6)。说明 QM 基因在斑马鱼成体组织和胚胎中有广泛表达。

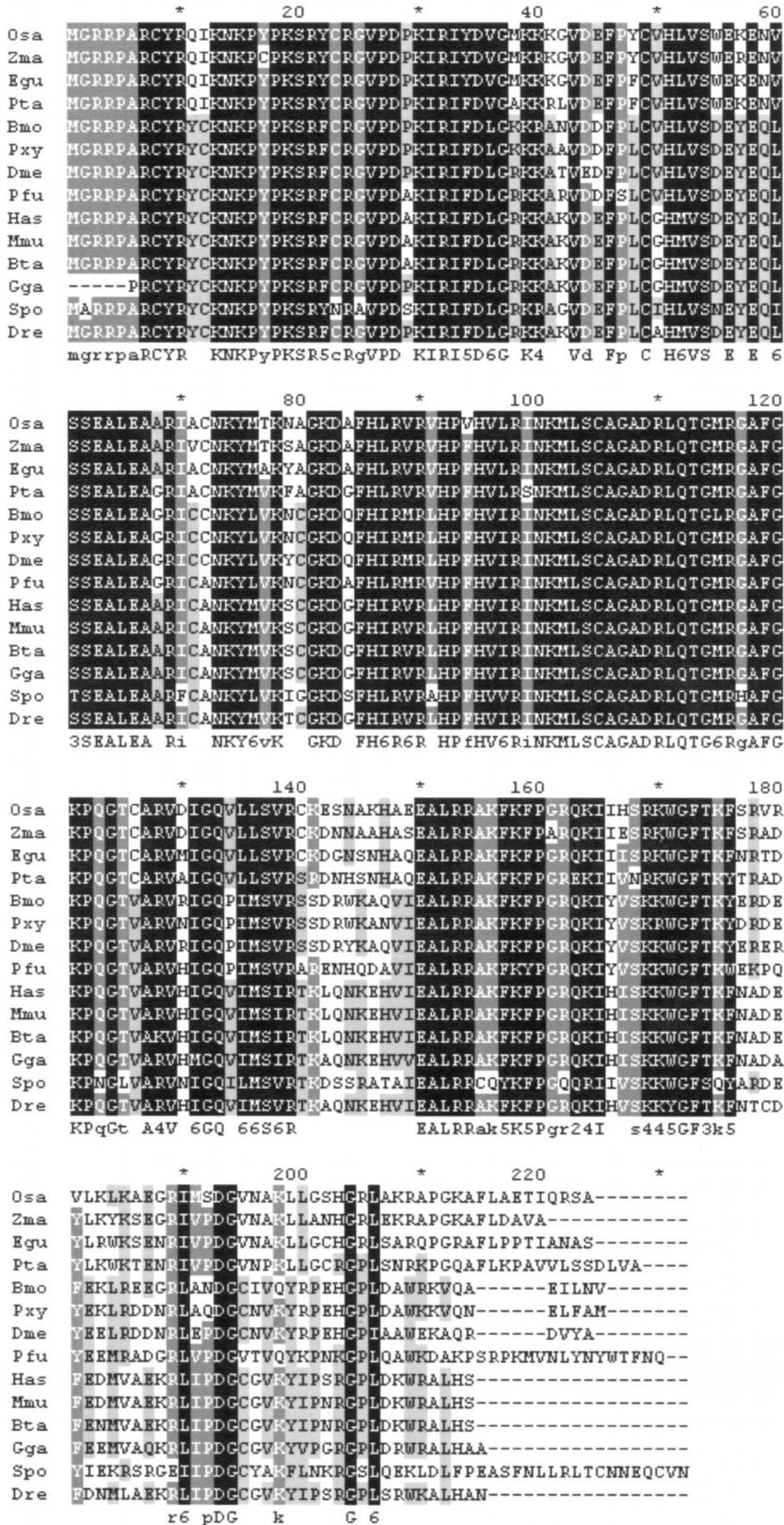


图 4 DrQM 与其它 13 种 QM 基因的氨基酸序列同源性比较

Fig. 4 Alignment of the predicted amino acid sequence of

the DrQM gene with those of corresponding proteins in other organisms

Osa 水稻 (*O. sativa*); Zma 玉米 (*Z. mays*); Egu 非洲油棕榈 (*E. guineensis*); Pta 火炬松 (*P. taeda*); Mmu 家鼠 (*M. musculus*); Bmo 家蚕 (*B. mori*); Pxy 菱形斑纹蛾 (*P. xylostella*); Dme 果蝇 (*D. melanogaster*); Pfu 珍珠贝 (*P. fucata*);

Has 人 (*H. sapiens*); Bta 牛 (*B. taurus*); Gga 鸡 (*G. gallus*); Spo 酵母 (*S. pombe*); Dre 斑马鱼 (*D. rerio*)

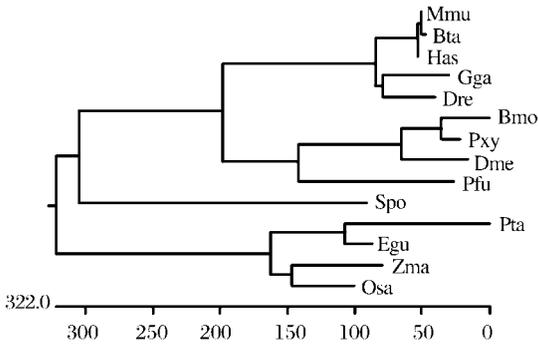


图5 DrQM 的系统发育分析

Fig. 5 Phylogenetic tree showing the evolutionary relationship of DrQM with other proteins of the QM family. 进化树下方的标尺代表核苷酸替代数的 $10 \times$ 值。Units on the scale underneath the tree represents $10 \times$ number of nucleotide substitution events.

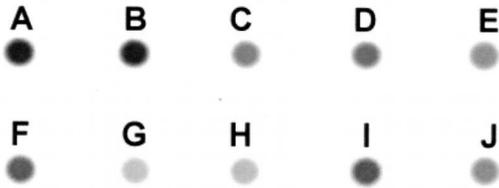


图6 DrQM 的数字化差异表达分析

Fig. 6 Tissue expression analysis of DrQM by digital differential display

点的浓度基于每百万转录本中的基因 EST 数; A. 脑; B. 鳍; C. 胚胎; D. 睾丸; E. 卵巢; F. 心脏; G. 嗅觉器官; H. 肾; I. 肝; J. 未知组织
spot intensity based on gene ESTs per million transcripts; A. brain; B. fin; C. embryo; D. testis; E. ovary; F. heart; G. olfactory; H. kidney; I. liver; J. unreported tissue type

3 讨论

随着人类和多种模式生物基因组计划的实施,表达序列标签(expressed sequence tag, EST)在基因结构与功能研究中的应用日益广泛^[12],各种生物信息学技术也应运而生并不断完善,充分利用生物信息学资源和技术研究基因组结构与功能、发现新的功能基因、揭示基因起源与分子进化规律等已成为后基因组时代的核心任务。随着各种物种 EST 数目的剧增以及 IMAGE 协定(THE I. M. A. G. E. CONSORTIUM, <http://mage.lnl.gov/>)的扩展,越来越多的新基因不仅可以通过对

EST 数据库的系统分析而加以鉴定,而且还可以通过索取包含该基因的相应 EST 的 IMAGE 克隆而直接获得,从而避免或减轻了采用传统 RACE、文库筛选等实验室方法克隆全长基因的繁琐工作^[3]。大量研究表明,利用 EST 数据库鉴定人类等基因在各种生物中的同源序列是目前基因克隆研究中最快捷和有效的方法,当前国际上相当数量的基因克隆研究都是从同源 EST 分析开始的,而利用 EST 数据库中丰富的组织来源、发育和疾病状态等信息进行基因的大规模表达分析,则是近年兴起的基因表达研究强有力的工具和方法,已经成为与基因表达系列分析(serial analysis of gene expression, SAGE)、基因芯片等并驾齐驱的高通量(high-throughput)表达分析的方法之一^[13-15]。

我们应用日益增长的斑马鱼 EST 数据库以及电子克隆(in silico cloning)等生物信息学技术,通过将互为重叠的斑马鱼 EST 序列进行排列(alignment)和组装,首次成功获得了 DrQM 基因全长 cDNA 序列。在此基础上,我们在 DrQM 基因开放阅读框架(ORF)内设计了特异引物,通过 RACE-PCR 扩增了 DrQM 基因片段,测序结果显示,其基因序列与电子克隆所获得的结果完全一致,由此证明了利用 EST 数据库克隆的 QM 基因是正确和可靠的。此外,我们还通过 DrQM 基因与其它物种 QM 基因的同源性分析,找出了不同物种间 QM 基因的同源保守序列,根据该保守序列设计引物,在草鱼头肾 cDNA 文库中克隆了草鱼 QM 基因,基因结构分析表明,两种鱼类 QM 基因的核苷酸序列同源性达到 89%,而氨基酸序列同源性达到 96%(另文报道),说明 QM 是一种在鱼类中比较保守的基因。另外我们还通过电子克隆方法,在斑马鱼和草鱼等鱼类中鉴定了一批免疫功能基因如 ILF、NKEF 等在鱼类中的同源序列^[16]。上述研究验证了利用表达序列标签数据库搜索鉴定鱼类新基因的可行性,这也是我们应用生物信息学技术作为基因克隆辅助手段,克隆鱼类功能基因的一次有意义的尝试。可以预示,今后几年随着多种模式生物基因组测序工作的完成以及鱼类基因数据库的不断完善,EST 分析必将作为指导鱼类和其它物种新基因克隆的重要辅助手段而得到广泛应用,从而成为利用模式生物基因组研究成果的重要手段。

数字化差异显示分析 (digital differential display, DDD) 是运用计算机技术对代表某一基因的、源于不同组织、不同生理状态或在特定发育阶段表达的 EST 序列数据进行分析, 从而绘制该基因在不同时空差异表达的图谱^[12]。我们利用该方法, 分析了 DrQM 的组织表达规律。由于在数字化差异表达分析时, 分析结果取决于同源基因 EST 数量和组织中 EST 总量的比值, 对于 EST 总量不在同一数量级的组织进行比较, 有可能产生误差, 因此, 我们在研究中只选择了 EST 总量在同一数量级 (10 000~ 50 000) 的几种组织进行差异表达分析, 从而避免了误差的产生。

QM 基因作为一种与肿瘤抑制、细胞生长、分化、发育和凋亡等有关的重要基因, 已在人类和许多物种中得到了克隆并进行了较深入的生物学性质与功能研究, 但鱼类中尚无 QM 基因克隆和功能等方面的研究报道, 因此我们对鱼类 QM 基因的克隆, 丰富了 QM 基因的研究内容, 为进一步开展 QM 在鱼类中的生物学功能研究打下了基础。同时鱼类作为生物进化中一个具有重要地位的物种, 对鱼类 QM 基因的研究, 对揭示 QM 基因分子起源与进化等也将具有重要的理论意义。

参考文献:

- [1] Weissman B, Stanbridge E, Farmer A, *et al.* Molecular characterization of QM, a novel gene with properties consistent with tumor suppressor function [J]. *Cold Spring Harb Symp Quant Biol*, 1994, 59: 573- 576.
- [2] Farmer A A, Loftus T M, Mills A A, *et al.* Extreme evolutionary conservation of QM, a novel e-Jun associated transcription factor [J]. *Hum Mol Genet*, 1994, 3: 723- 728.
- [3] Weissman B E, Saxon P J, Pasquale S R, *et al.* Introduction of a normal human chromosome 11 into a Wilms' tumor cell line controls its tumorigenic expression [J]. *Science*, 1987, 236: 175 - 180.
- [4] Monteclaro F S, Vogt P K. A Jun-binding protein related to a putative tumor suppressor [J]. *Proc Natl Acad Sci USA*, 1993, 90: 6726- 6730.
- [5] Tron T, Yang M, Dick F A, *et al.* QSR1, an essential yeast gene with a genetic relationship to a subunit of the mitochondrial cytochrome bcl complex, is homologous to a gene implicated in eukaryotic cell differentiation [J]. *J Biol Chem*, 1995, 270: 9961- 9970.
- [6] Eisinger D P, Jiang H P, Serrero G. A novel mouse gene highly conserved throughout evolution: regulation in adipocyte differentiation and in tumorigenic cell lines [J]. *Biochem Biophys Res Comm*, 1993, 196: 1227- 1232.
- [7] Masson JY, Vadnai J, Ramotar D. The Schizosaccharomyces pombe spqM gene is a new member of the Qm transcription factor family [J]. *Gene*, 1996, 170: 153- 154.
- [8] Oh H S, Kwon H, Sun S K, *et al.* QM, a putative tumor suppressor, regulates proto-oncogene *c-myc* [J]. *The Journal of Biological Chemistry*. 2002, 277(39): 36489- 36498.
- [9] Wellbum S C, Lillico S, Murphy N B. Programmed cell death in procyclic form trypanosoma brucei rhodesiense identification of differentially expressed genes during con A induced death [J]. *Mem Inst Oswaldo Cruz*, 1999, 94(2): 229- 234.
- [10] Kyte J, Doolittle R F. A simple method for displaying the hydrophobic character of a protein [J]. *J Mol Biol*, 1982, 157: 105- 132.
- [11] Chou P Y, Fasman G D. Empirical predictions of protein conformation [J]. *Annual Review of Biochemistry*, 1978, 47: 251- 276.
- [12] Adams M D, Kelley J M, Gocayne J D, *et al.* Complementary DNA sequencing: expressed sequence tags and human genome project [J]. *Science*, 1991, 252: 1651- 1656.
- [13] Muhle R A, Pavlidis P, Grundy W N, *et al.* A high-throughput study of gene expression in preterm labor with a subtractive microarray approach [J]. *Am J Obstet Gynecol*, 2001, 185(3): 716- 724.
- [14] Caulli J P, Artinger M, Swain P M, *et al.* High throughput analysis of differential gene expression [J]. *J Cell Biochem Suppl*, 1998, 30- 31: 286- 296.
- [15] Gill R W, Sansau P. Rapid in silico cloning of genes using expressed sequence tags (ESTs) [J]. *Biotechnol Annu Rev*, 2000, 5: 25- 44.
- [16] 金红建, 王兰州. 斑马鱼 ILF2 基因的电子克隆分析 [J]. *中国计量学院学报*, 2003, 14(1): 70- 73.