

鲤科鱼类 *Angel* 相关元件的鉴定及其系统发育意义*

王莉^{1,2} 何舜平^{1*} 陈宜瑜³

1. 中国科学院水生生物研究所, 武汉 430072; 2. 中国科学院研究生院, 北京 100039;
3. 国家自然科学基金委员会, 北京 100085

摘要 *Angel* 相关元件属于微型反向重复转座元件(MITEs)的一个家族, 我们在鲤科雅罗鱼系鱼类的生长激素基因内含子 2 序列中分离并鉴定了这类元件, 而在鲃系的直系同源位点上没有发现, 从而为雅罗鱼系的单系性提供了新证据. 这类元件可能是在雅罗鱼系与鲃系分枝后插入到雅罗鱼系的共同祖先中的. 较高的序列变异及相对广泛的物种分布表明生长激素基因中的 *Angel* 相关元件是一类古老的转座子, 大约 26 Ma 前可能已经在雅罗鱼系共同祖先中出现.

关键词 *Angel* 相关元件 雅罗鱼系 单系性

微型反向重复转座元件(MITEs)是一类分布在真核生物基因组中的中度重复 DNA 片段, 它们大约 80—500 bp 长, 没有编码能力, 两端具有短末端反向重复序列(TIRs), 转座时在宿主基因组的靶位点产生正向重复序列(TSDs), 且有形成茎环结构的能力. 如果不同物种基因组的同一位点存在某一 MITE 插入, 这意味它们有一个共同离征, 能提供系统发育信息^[1]. 例如, 玉米中, *Heartbreaker* (*Hbr*) 家族已经成功地被用作分子标记^[2]. 但在脊椎动物中, 尚未有 MITEs 被用作系统发育标记的报道. Izsvák 等^[3]首次在斑马鱼(*Danio rerio*)的真核生物蛋白合成起始因子 4E(eIF-4E)基因中鉴定了一个 MITE—*Angel* 元件, 同时报道了在草鱼(*Ctenopharyngodon idellus*)和白鲢(*Hypophthalmichthys molitrix*)的生长激素(GH)基因内含子 2 中存在 *Angel* 相关元件(*Angel* 家族的成员)的插入. 然而, *Angel* 相关元件在其他鲤科鱼类生长激素基因中的分布情况却是未知的.

本研究中, 通过对目的片段的扩增, 探讨 *An-*

gel 相关元件在鲤科鱼类生长激素基因中的分布及其系统发育意义.

1 材料和方法

1.1 样品 DNA 的提取、PCR 扩增和测序

研究所用的样品包括了鲤科 11 个亚科^[4]的 27 个代表种, 保存于中国科学院水生生物研究所(表 1). 采用标准的苯酚/氯仿抽提法从 95%乙醇固定的肌肉组织中提取总基因组 DNA. 根据下载的 5 种鲤科鱼类 GH 基因序列(表 1), 分别在外显子 2 和 3 上设计引物 GF (5'-CCAGCGGCTYTTCAAYAAY-GCAGT-3')和 GR (5'-AGTCAGARTTGCAGAAAGACAGAGG-3')用来扩增长生长激素基因内含子 2 片段. PCR 反应总体积为 60 μ L, 其中含 100 ng 基因组 DNA, 引物 (10 μ mol/L) 各 1.5 μ L, 10 \times 缓冲液 6 μ L, dNTP (10 mmol/L) 0.75 μ L, Taq 聚合酶 3.0 U, 最后补足灭菌双蒸水至终体积. 反应条件为: 94 $^{\circ}$ C 预变性 3 min; 94 $^{\circ}$ C 变性 30 s, 62.5 $^{\circ}$ C 退火 45 s, 72 $^{\circ}$ C 延伸 1 min, 32 个循环; 72 $^{\circ}$ C 终延伸

2006-06-07 收稿, 2006-08-07 收修稿

* 国家杰出青年科学基金(批准号: 30225008)和国家自然科学基金(批准号: 30530120)资助项目

** 通信作者, E-mail: clad@ihb.ac.cn

5 min. PCR 扩增产物经 1.2% 溴化乙锭染色的低熔点琼脂糖凝胶电泳后, 用试剂盒纯化回收 (BioStar Glassmilk DNA Purification Kit). 为保证目的片段

测序的可靠性, 每一样品均用正反两个引物分别测序. 序列测定在自动测序仪 (Applied Biosystems 377 Stretch) 上进行, 其 GenBank 登录号见表 1.

表 1 本文中使用的标本及其 GenBank 登录号

亚科	物种	采集地	标本号	GenBank 登录号	
鲇亚科	厚颌鲂 <i>Megalobrama pellegrini</i>	四川合江	IHB0405133	DQ219449	
	团头鲂 <i>Megalobrama amblycephala</i>			AF463498 ^{a)}	
	长春鳊 <i>Parabramis pekinensis</i>	湖南岳阳	IHB0411120	DQ219441	
	红鳍原鲃 <i>Cultrichthys erythropterus</i>	广西灵山	IHB0301134	DQ219450	
	飘鱼 <i>Pseudolaubuca sinensis</i>	湖南桃源	IHB0305194	DQ219440	
	南方拟鲃 <i>Pseudohemiculter dispar</i>	广西崇左	IHB1303127	DQ219439	
雅罗鱼亚科	鲢 <i>Elopichthys bambusa</i>	湖南桃源	IHB040511	DQ219444	
	赤眼鲮 <i>Squaliobarbus curriculus</i>	湖北金口	IHB0301132	DQ219437	
	草鱼 <i>Ctenopharyngodon idellus</i>			X60419 ^{a)}	
	鳊 <i>Ochetobius elongates</i>	广西腾县	IHB0817003	DQ219447	
鲴亚科	青鱼 <i>Mylopharyngodon piceus</i>	湖北武汉	IHB0401127	DQ219438	
	圆吻鲴 <i>Distoechodon tumirostris</i>	湖北金口	IHB020636	DQ219436	
鲢亚科	鳊 <i>Aristichthys nobilis</i>	湖北武汉	IHB040923	DQ219435	
鲃亚科	鲢 <i>Hypophthalmichthys molitrix</i>			M94348 ^{a)}	
鲃亚科	高体鳊 <i>Rhodeus ocellatus</i>	四川乐山	IHB0207101	DQ219442	
鲃亚科	蛇鲃 <i>Saurogobio dabryi</i>	重庆合川	IHB0405356	DQ219443	
鲃亚科	稀有鲃 <i>Gobiocypris rarus</i>	四川彭县	IHB0305124	DQ219448	
	马口鱼 <i>Opsariichthys bidens</i>	广西横县	IHB040687	DQ219445	
	宽鳍鱮 <i>Zacco platypus</i>	广西西林	IHB0303120	DQ219446	
	拟细鲫 <i>Nicholsicypris normalis</i>	海南吊罗山	IHB0301017	DQ219451	
	中华细鲫 <i>Aphyocypris chinensis</i>	四川彭县	IHB0305216	DQ219452	
	斑马鱼 <i>Danio rerio</i>			BX005440 ^{b)}	
	鲤亚科	鲤 <i>Cyprinus carpio</i>			X51969 ^{a)}
		鲫 <i>Carassius auratus</i>	湖北武汉	IHB040925	DQ219453
鲃亚科	中华倒刺鲃 <i>Spinibarbus sinensis</i>	四川南充	IHB031259	DQ219454	
野鲮亚科	东方墨头鱼 <i>Garra orientalis</i>	广西天峨	IHB0403442	DQ219456	
裂腹鱼亚科	南方裂腹鱼 <i>Schizothorax meridionalis</i>	云南腾冲	IHB03468	DQ219455	

a) 为从 GenBank 上下载的序列; b) 为斑马鱼基因组中与草鱼的 GH 序列相似性最高的序列 (e 值为 2×10^{-70})

1.2 数据分析

使用 Clustal X 软件^[5]和 SEAVIEW 程序^[6]对序列进行比对, 并加以手工校正. 5 种下载的鲤科鱼类 GH 基因内含子 2 序列和我们的数据一起分析. 采用 MEGA3.1^[7]计算 Angel 相关元件各碱基的含量、替代及序列变异情况. 基于最小自由能原理, 用 RNA structure 3.2^[8]软件折叠它们的二级结构.

基于 GH 基因内含子 2 序列, 使用 PAUP* 4.0b10 软件^[9]构建最大简约树. 分析中选择启发式搜索 (heuristic searches), 树二等分再连接选项 (tree bisection-reconnection, TBR), 使用非参数自展法^[10]检验分支的可靠性, 参数设置为: 10 random stepwise additions sequences, 1000 次重复. 所

有位点均不加权. 当几个同等简约的无根树出现时, 最后结果将由 50% 一致树来表示.

2 结果与讨论

2.1 Angel 相关元件的特征和鉴定

通过序列测定和下载 (表 1), 共获得 27 种鲤科鱼类的 GH 基因内含子 2 序列. 多重比对后发现 21 个种的序列中都存在一个 DNA 片段的插入, 长度在 67—306 bp 之间, 其中 4 种鱼类的插入片段明显较短, 分别长 67 bp (长春鳊, *Parabramis pekinensis*), 148 bp (中华细鲫, *Aphyocypris chinensis*), 177 bp (鳊, *Aristichthys nobilis*), 224 bp (蛇鲃, *Saurogobio dabryi*). 插入片段中 A, T, G 和 C 的

平均含量为 34.3%, 32.2%, 16.4% 和 17.2%, A+T 的含量 (66.5%) 远高于 G+C 的含量 (33.6%); 平均的转换颠换比率是 1.073; 种间序列变异大约在 0.4% 到 22.4% 之间。

这 21 个插入片段都具有真核生物 MITEs 家族的共有特征(表 2), 包括能形成茎环二级结构(图未示出), 我们初步判定它们是 MITEs 的一个家族。而 MITEs 又是根据 TIR 和 TSD 序列被划分为不同家族的^[1]。Angel 元件具有 4 bp (TTAA) 的 TSD 序列和长 26 bp 的 TIR 序列^[3]。我们获得的 21 个插入片段大多具有与 Angel 元件完全一致的 TSD 序列, 它们的 TIRs 长度大约 26 bp, 与 Angel 元件的 TIRs 高度相似。因此, 我们认为它们属于 Angel 家族。草鱼和白鲢的该插入片段对应于已报道的 Angel 相关元件^[3]。将另外 19 个种的插入片段分别与 GenBank 中的序列相比较, 发现它们都和上述两个 Angel 相关元件的序列有一定的相似性。例如, 高体鲮鱼(*Rhodeus ocellatus*) 中长 282 bp 的插入片段与上述两个 Angel 相关元件的序列相似性很高(e 值分别为 4×10^{-15} , 4×10^{-9})。所以, 我们确定这 19 个插入片段为 Angel 相关元件。

表 2 Angel 相关元件(以高体鲮鱼 GH 基因中的此元件为代表)与 Angel 元件特征的比较

特征	Angel 相关元件	Angel 元件
长度 (bp)	282	312
靶位点正向	TTAA	TTAA
重复序列(TSD)		
末端反向重复	5'-TTAAAGGGATAGT	5'-TTAAAGGATAGTT
序列(TIR)	TCACTCAAAGTA-3'	CACCCAAAATGA-3'
能否形成茎环二级结构	能	能

2.2 Angel 相关元件在鲤科鱼类中的系统发育意义

我们将 GH 基因内含子 2 序列中 Angel 相关元件插入的有无标注于最大简约树上(图 1), 结果发现: 在雅罗鱼系 GH 基因中都有 Angel 相关元件的插入, 而鲃系和斑马鱼中则没有。在基于形态和分子的鲤科鱼类系统发育研究中, 尽管对这两个谱系所包括的具体内容还有争议, 大多数学者认为鲤科中存在两个谱系^{[11-13]1)}: 雅罗鱼系和鲃系。因此 Angel 相关元件在 GH 基因中的插入可以作为雅罗鱼系鱼类的一个共同特征, 从而为雅罗鱼系的单系性提供了新的证据。雅罗鱼系鱼类构成一个单系群, 这与以前的研究结果大体一致^{[12, 13]1)}。

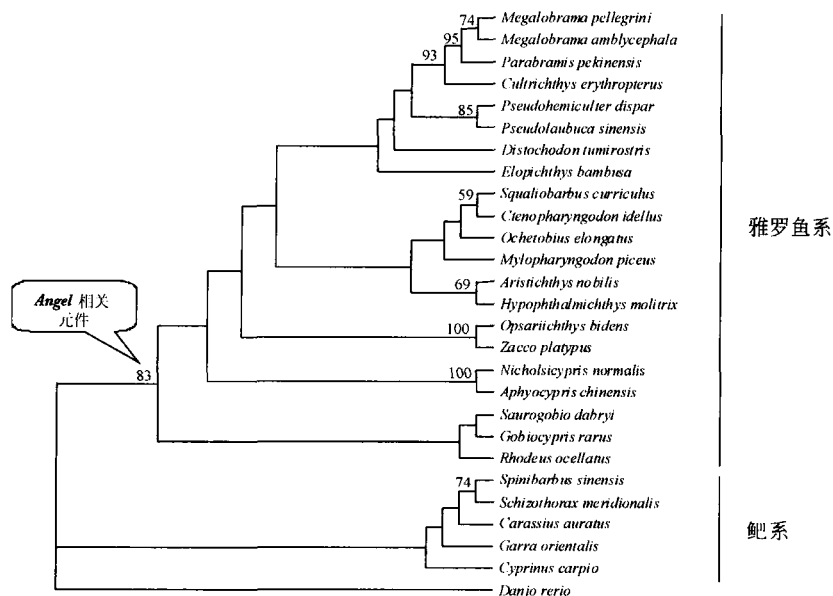


图 1 基于 GH 基因内含子 2 序列得到的 50% 最大简约无根一致树

树长 461 步, $CI=0.8091$, $RI=0.6879$ 。节点的数字为 1000 次自展支持率(仅显示 50% 以上)。

Angel 相关元件在雅罗鱼系的生长激素基因内含子 2 中存在却在鲃系和斑马鱼的位置缺失

1) Wang XZ, Li JB, He SP. Molecular evidence for the monophyly of East Asian groups of Cyprinidae (Teleostei: Cypriniformes) derived from the nuclear recombination activating gene 2 sequences. Mol Phylogenet Evol. 2006 (In press)

因为 MITEs 不编码转座酶或多肽, 它们被认为是非自主性 DNA 元件^[14], 能在相关自主元件提供的转座酶催化下进行插入和切除^[15]. 插入是正向而普遍的, 而切除则相对少见^[16]. 因此我们推测 *Angel* 相关元件是在雅罗鱼系和鲃系分歧之后插入到雅罗鱼系鱼类祖先的 GH 基因中的. 此外, MITEs 插入到某个位点后, 受到随机突变的影响其序列和长度的同质性会逐渐降低^[17]. 我们获得的 *Angel* 相关元件长度不一, 它们的 TSDs 序列和 TIRs 序列中包含较多的突变, 少数元件则缺失了一侧的 TSD 或 TIR. 较高的内部序列变异及相对广泛的物种分布表明它们属于一个比较古老的转座子家族. 根据已发表的雅罗鱼系与鲃系的分歧时间^[13], 我们估计大约 26 Ma 前 *Angel* 相关元件可能已经在雅罗鱼系鱼类祖先的 GH 基因中出现.

与线粒体或核基因分子标记相比, MITEs 的插入用作系统发育标记的优点在于: 插入通常是不可逆的, 因为没有已知的机制能从基因组中特异地移走 MITEs; 可以避免趋同进化和平行进化的影响; 结果的推断主要依靠 MITEs 插入的有无, 而不是获得的 DNA 序列数据, 可以减少由于突变饱和、碱基组成的偏倚性或测序带来的错误. 但是解决一个类群的系统发育关系通常需要多个 MITEs 插入位点. 与 SINEs^[18] 类似, 大量的 MITEs 插入位点可以成为分子系统学研究的有力工具.

参 考 文 献

- 1 Petersen G, Seberg O. Phylogenetic evidence for excision of *Stowaway* miniature inverted-repeat transposable elements in Triticeae (Poaceae). *Mol Biol Evol*, 2000, 17(11): 1589—1596
- 2 Alexandra MC, Cory B, Alexander N, et al. The MITE family *Heartbreaker* (*Hbr*): Molecular markers in maize. *Proc Natl Acad Sci USA*, 2000, 97(18): 10083—10089
- 3 Izsvák Z, Ivics Z, Shimoda N, et al. Short inverted-repeat transposable elements in teleost fish and implications for a mechanism of their amplification. *J Mol Evol*, 1999, 48(1): 13—21
- 4 陈宜瑜. 中国动物志, 硬骨鱼纲, 鲤形目. 中卷. 北京: 科学出版社, 1998

- 5 Thompson JD, Gibson TJ, Plewniak F, et al. The Clustal_X windows interface: Flexible strategies for multiple sequences alignment aided by quality analysis tools. *Nucleic Acids Res*, 1997, 25(24): 4876—4882
- 6 Galtier N, Gouy M, Gautier C. SEAVIEW and PHYLO_WIN: Two graphic tools for sequence alignment and molecular phylogeny. *Comput Appl Biosci*, 1996, 12(6): 543—548
- 7 Kumar S, Tamura K, Nei M. MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. *Briefings in Bioinformatics*, 2004, 5: 150—163
- 8 Mathews DH, Disney MD, Childs JL, et al. Incorporating chemical modification constraints into a dynamic programming algorithm for prediction of RNA secondary structure. *Proc Natl Acad Sci USA*, 2004, 101: 7287—7292
- 9 Swofford DL. PAUP*. Phylogenetic Analysis Using Parsimony (* and other methods). Version 4. 0b10, Sunderland, Massachusetts: Sinauer Associates, 2002
- 10 Felsenstein JP. Confidence limits on phylogenies: An approach using the bootstrap. *Evolution*, 1985, 39: 783—791
- 11 陈湘舜, 乐佩琦, 林人端. 鲤科鱼类的科下类群及其宗系发生关系. *动物分类学报*, 1984, 9(4): 424—440
- 12 Liu HZ, Chen YY. Phylogeny of the East Asian Cyprinids inferred from sequences of the mitochondrial DNA control region. *Can J Zool*, 2003, 81(12): 1938—1946
- 13 何舜平, 刘焕章, 陈宜瑜, 等. 基于细胞色素 b 基因序列的东亚鲤科鱼类系统发育研究(鱼纲: 鲤形目). *中国科学, C 辑*, 2004, 34(1): 96—104
- 14 Feschotte C, Jiang N, Wessler SR. Plant transposable elements: Where genetics meets genomics. *Nat Rev Genet*, 2002, 3(5): 329—341
- 15 Feschotte C, Zhang X, Wessler SR. Miniature inverted-repeat transposable elements (MITEs) and their relationship with established DNA transposons. In: *Mobile DNA II*. Washington: American Society of Microbiology Press, 2002, 1147—1158
- 16 Wessler SR. Transposable elements associated with normal plant genes. *Physiol Plant*, 1998, 103(4): 581—586
- 17 Santiago N, Herráiz C, Goñi JR, et al. Genome-wide analysis of the Emigrant family of MITEs of *Arabidopsis thaliana*. *Mol Biol Evol*, 2002, 19(12): 2285—2293
- 18 Shedlock AM, Okada N. SINE insertions: powerful tools for molecular systematics. *BioEssays*, 2000, 22(2): 148—160