

虾虎鱼科(Gobiidae)基因组微卫星 DNA 的分布特征*

白翠翠^{1,2,3} 柳淑芳^{1,2①} 庄志猛¹

(1. 农业部海洋渔业可持续发展重点实验室 中国水产科学研究院黄海水产研究所 青岛 266071;

2. 青岛海洋科学与技术国家实验室 海洋渔业科学与食物产出过程功能实验室 青岛 266200;

3. 山东农业大学 泰安 271018)

摘要 为了解虾虎鱼科(Gobiidae)鱼类基因组遗传结构特征,本研究自主开发矛尾复虾虎鱼(*Synechogobius hasta*)微卫星序列 153 条,结合从 GenBank 中筛选出的虾虎鱼科微卫星序列 535 条,合计 686 条微卫星序列,隶属于 19 种虾虎鱼,序列总长度为 295062 bp,包含 473 个微卫星位点,微卫星位点累计长度为 33370 bp。统计微卫星重复类型,发现在所有微卫星重复类型中以二碱基重复出现次数最多,位点数为 361 个,占总微卫星的 76.32%,其中,重复拷贝类别最多的是 AC (340 个),占全部微卫星的 71.88%,占二碱基重复微卫星的 94.18%,在二碱基重复类型中没有发现 AT 和 GC 两种重复类型。三碱基重复位点数为 35 个,占总微卫星的 7.4%,其中以 ACT 重复拷贝类别最多(12 个),占三碱基重复微卫星的 34.29%。四碱基重复位点数为 68 个,其中以 CTAT 重复最多(31 个),占总微卫星的 14.38%。五碱基重复类型中只有 TCTGG 和 ATCTA 两种类型,只占全部微卫星的 0.42%。六碱基重复 7 个,占总微卫星的 1.48%,六碱基重复中各重复类型出现次数相当。在所有微卫星重复类型中没有发现单核苷酸重复类型。

关键词 虾虎鱼科;微卫星;基因组

中图分类号 S917.4 **文献标识码** A **文章编号** 2095-9869(2016)05-0009-07

微卫星 DNA, 又称简单重复序列(Simple Sequence Repeat, SSR)或短串联重复序列(Short Tandem Repeats, STR),通常指基因组上 1-6 个核苷酸为重复单位组成的 DNA 序列(Toth *et al.*, 2000)。微卫星标记因多态性丰富、杂合度高、稳定性好、共显性遗传及符合孟德尔定律等优点,而被广泛用于种群遗传多样性研究、物种起源与进化分析、基因定型、指纹鉴定、遗传图谱构建与家系分析等。因此,分析物种的基因组微卫星的碱基组成、含量、分布特征,有助于从进化角度了解各物种内各重复类型的分布特点及其生物学功能(Gao *et al.*, 2005; Katti *et al.*, 2001; Prasad *et al.*, 2005;

Li *et al.*, 2005; 郭文久, 2004¹⁾)。迄今为止,已报道的基因组微卫星特征分析有 30 多个物种,包括人类和黑猩猩在内的灵长类、小鼠、家猪、鸡、蚊子、海鞘、斑马鱼及中国对虾等动物,也有拟南芥、水稻和小麦等植物和一些原核生物(郭文久, 2004¹⁾; Li *et al.*, 2004; Gao *et al.*, 2005)。研究表明,在真核生物基因组中,二核苷酸微卫星序列最为丰富(Ma *et al.*, 1996),但不同物种的微卫星序列含量各不相同(Wang *et al.*, 1994)。目前,水生生物基因组微卫星特征分析多局限于单一物种,即分类学的“种”级水平,如三疣梭子蟹(宋来鹏等, 2008)、日本囊对虾(栾生等, 2007)、

* 国家自然科学基金项目(31172411)、国际科技合作专项(2013DFA31410)和泰山学者建设工程专项共同资助。白翠翠, E-mail: baicuicui77@163.com

① 通讯作者: 柳淑芳, 研究员, E-mail: liusf@ysfri.ac.cn

收稿日期: 2015-10-08, 收修改稿日期: 2015-12-02

1) 郭文久. 微卫星在基因组上的分布与功能及其计算方法初步研究. 四川农业大学博士研究生学位论文, 2004, 1-163

牙鲆(王蕾等, 2009)、红鳍东方鲀(崔建洲等, 2006)、中国对虾(高焕等, 2004)等, 但在分类学的“科”级水平上分析基因组微卫星的分布特征鲜见报道。在“科”级水平上研究其微卫星序列特征, 可加深了解科内其他物种基因组微卫星序列的分布规律, 提高微卫星标记的开发和应用效率。随着水生生物基因组大规模测序工作的开展, 微卫星标记的开发与应用研究将获得更新、更准确并且具有普遍意义的结果。

虾虎鱼科鱼类分布广泛、种类繁多、分类复杂, 共有 242 属, 2000 种之多, 是鱼类中最大的科(李思忠等, 1994)。据统计, 我国虾虎鱼类约 84 属 269 种(伍汉霖等, 2008)。虾虎鱼在水生种质资源中具有重要地位, 少数虾虎鱼类, 如矛尾复虾虎鱼(*Synechogobius hasta*)、大弹涂鱼(*Boleophthalmus pectinirostris*)等为经济鱼类, 其余大部分为大型鱼类饵料, 其中, 部分虾虎鱼属于观赏鱼类, 例如短虾虎属(*Brachygobius*)的种类等。

近年来, 由于水生生物种质资源的不断衰退, 对

虾虎鱼类的种质资源保护和开发利用也备受关注。在基因组水平上研究虾虎鱼科鱼类微卫星重复序列的分布特征, 能够深化认知虾虎鱼科鱼类基因组遗传结构特征, 将为虾虎鱼科鱼类微卫星标记的开发与应用提供参考。

1 材料与方法

1.1 样品和序列信息

2011 年 4 月在北黄海采集矛尾复虾虎鱼 48 尾, 剪鳍条, 用无水乙醇固定, 4℃保存备用。对上述样品进行微卫星自主开发, 获得矛尾复虾虎鱼微卫星序列 153 条, 含有 68 个微卫星位点和 85 个小卫星位点(Bai *et al.*, 2012)。同时, 从 GenBank 中筛选出虾虎鱼科微卫星序列 533 条, 含有 405 个微卫星位点和 128 个小卫星位点。本研究所分析的虾虎鱼科基因组微卫星序列合计 686 条, 共含有 473 个微卫星位点, 分别隶属于 16 属 19 种(表 1)。

表 1 虾虎鱼科微卫星序列信息
Tab.1 The SSR information of species of the Gobiidae family

种名 Species	序列数(微卫星数) Sequences (SSR loci)	GenBank 序列号 GenBank No.
矛尾复虾虎鱼 <i>Synechogobius hasta</i>	153(68)	本研究, in this study
矛尾复虾虎鱼 <i>Synechogobius hasta</i>	76(69)	EU714684-EU714759
双须多棘虾虎鱼 <i>Sufflogobius bibarbatuis</i>	15(14)	JN251023-JN251036
彼氏冰虾虎鱼 <i>Leucopsarion petersi</i>	97(31)	AB662996-AB663083 AB304787-AB304796
关岛阿胡虾虎鱼 <i>Awaous guamensis</i>	14(12)	HQ393961-HQ393975
大弹涂鱼 <i>Boleophthalmus pectinirostris</i>	150(130)	HQ852252-HQ852395 FJ598061-FJ598069
黑口新虾虎鱼 <i>Neogobius melanostomus</i>	16(11)	DQ9999744-DQ999981 EU732735-EU732744
大头新虾虎鱼 <i>Neogobius kessleri</i>	19(15)	EF629921-EF629939
尾纹裸头虾虎鱼 <i>Chaenogobius annularis</i>	16(13)	AB557720-AB557736
环肛麒麟虾虎鱼 <i>Coryphoterus peraonatus</i>	17(12)	GQ428991-GQ428998 AY738124-AY738132
白带高鳍虾虎鱼 <i>Pterogobius zonoleucus</i>	11 (6)	AB447471-AB447481
黄体尻虾虎鱼 <i>Gobius culus flavescens</i>	8(4)	EU295520-EU295527
青弹涂鱼 <i>Scartelaos histophorus</i>	14(21)	HQ128584-HQ128597
丝棘带眼虾虎鱼 <i>Ctenogobius feroculus</i>	18(10)	EF115293-EF115298 DQ402141-EF402152
高伦颌鳞虾虎鱼 <i>Gnatholepis callerensis</i>	7(6)	EF115288-EF115294
蛇头虾虎鱼 <i>Zostrisessor ophiocephalus</i>	12(11)	AY687320-AY687331
菊花吻虾虎鱼 <i>Rhinogobius sp.</i>	11(10)	AB159090-AB159099
云斑长臀虾虎鱼 <i>Pomatoschistus marmoratus</i>	3(3)	DQ334750-DQ334752
小长臀虾虎鱼 <i>Pomatoschistus minutus</i>	24(22)	AF516891-AF516903 DQ914817-DQ914827
小眼长臀虾虎鱼 <i>Pomatoschistus micropus</i>	5(5)	DQ334751-DQ334754

1.2 数据统计

利用软件 Tandem Repeats Finder (Version 3.21) (Sugaya *et al.*, 2002)进行微卫星查找分析。微卫星序列的最小长度为 12 bp (郭文久, 2004)¹⁾, 即单碱基重复序列必须达到 12 个或 12 个拷贝以上, 而六碱基重复只需 2 个或 2 个拷贝以上即可。对查找到的微卫星序列进行统计分析, 根据碱基互补配对原则和阅读起始碱基顺序的差异, 对每种长度类型的重复单位基序进行同类兼并(表 2)。1-3 碱基重复可以兼并为表 2 中的几种类型, 4 碱基及其以上重复的兼并原则与表 2 中相同, 1-6 碱基重复兼并后的重复类型数目分别为 2、4、10、33、102、350。

为了衡量微卫星重复类型变异能力的大小, 根据栾生等(2007)引入变异系数概念, 变异系数的计算公式为:

$$CV = \frac{S}{\bar{x}} \times 100\%$$

式中, S 为某类型重复序列拷贝数的标准差, \bar{x} 为拷贝数的平均值。变异系数可以消除单位和(或)平均数不同对 2 个或多个资料变异程度比较的影响, 能够真实地反映重复类型变异水平的高低, 便于进行类型间的比较(栾生等, 2007)。

2 结果与分析

通过对 19 种虾虎鱼科鱼类已知微卫星序列进行

分析, 结果从中筛选出虾虎鱼科微卫星序列 686 条, 序列总长度为 295062 bp, 其中含有 473 个微卫星位点, 微卫星位点累计长度为 33370 bp。

2.1 微卫星各重复类型分布情况

由表 3 可以看出, 在 686 条微卫星重复序列中, 共 473 个微卫星位点。在各重复类型中以二碱基重复最多(361 个), 占总微卫星位点的 76.32%, 其中重复拷贝类别最多的是 AC 重复(340 个), 占二碱基重复微卫星的 94.18%, 没有发现 AT 和 GC 两种重复类型; 其次是四碱基重复(68 个), 占总微卫星的 14.38%, 在四碱基重复类型中以 CTAT 重复最多(31 个), 占四碱基重复微卫星的 45.56%; 第三类是三碱基重复类型(35 个), 占总微卫星的 7.40%, 在三碱基重复中以 ACT 占的比例最多, 占三碱基重复微卫星的 34.29%; 第四类是六碱基重复类型, 占全部微卫星的 1.48%; 第五类是五碱基重复类型, 其中只有 TCTGG 和 ATCTA 两种类型, 只占全部微卫星的 0.42%; 在所有碱基重复类型中没有发现单核苷酸重复类型。重复序列长度作为衡量标准得到的结果与此类似, 其中, 二碱基重复类型占微卫星序列累计长度的 79.60%, 其次, 四碱基重复占微卫星序列累计长度的 13.42%。

重复单位拷贝数变异能力分析显示(表 3), 二核苷酸重复中变异能力最大的是 AG 重复类型(76.51%), 其次是 AC 重复类型(67.59%); 三核苷酸重复中变异能力最大的是 ACT 重复类型(41.56%); 四核苷酸重复

表 2 虾虎鱼科鱼类微卫星碱基重复类型

Tab.2 The types of SSR repeats of species of the Gobiidae family

重复碱基类型 Types of repeat sequence	对应碱基种类 Types of nucleotide corresponding to repeat sequence	
二碱基重复 Dinucleotide repeat	AT	AT、TA
	AG	AG、GA、TC、CT
	AC	AC、CA、TG、GT
	GC	GC、CG
三碱基重复 Trinucleotide repeat	AAT	AAT、ATA、TAA、TTA、TAT、ATT
	AAC	AAC、ACA、CAA、TTG、TGT、GTT
	AAG	AAG、AGA、GAA、TTC、TCT、CTT
	ATC	ATC、TCA、CAT、TAG、AGT、GTA
	ACG	ACG、CGA、GAC、TGC、GCT、CTG
	ACT	ACT、CTA、TAC、TGA、GAT、ATG
	AGC	AGC、GCA、CAG、TCG、CGT、GTC
	GCC	GCC、CCG、CGC、CGG、GGC、GCG
	AGG	AGG、GGA、GAG、TCC、CCT、CTC
	ACC	ACC、CCA、CAC、TGG、GGT、GTG

1) 郭文久. 微卫星在基因组上的分布与功能及其计算方法初步研究. 四川农业大学博士研究生学位论文, 2004, 1-163

表 3 微卫星重复类型的频率和分布
Tab.3 The frequency and distribution of SSR of species of the Gobiidae family

重复类型 Repeat type	SSR 数 SSR number	占 SSR 总数比 Ratio(%)	累计 长度 Total length	占 SSR 序列 总长比 Length ratio of SSR(%)	占重复序列 总长比 Length ratio of repeat(%)	拷贝数 Copy number	平均 拷贝数 Mean copy	拷贝数 变异系数 Coefficient of variation (%)	平均 GC 含量 Mean GC content
AG	21	4.44	1212	3.63	0.41	14.0-104.0	28.86	76.51	0.50
AC	340	71.88	25351	75.97	8.59	13.0-234.0	37.28	67.59	0.50
Subtotal	361	76.32	26563	79.60	9.00	13.0-234.0	36.79		0.50
AAT	3	0.63	119	0.36	0.04	9.0-16.3	13.22	23.30	0.00
AAC	3	0.63	129	0.39	0.04	13.0-16.3	14.33	9.91	0.34
AAG	4	0.85	192	0.58	0.07	9.3-23.0	16.00	30.52	0.39
ATC	8	1.69	435	1.30	0.15	10.7-29.0	18.13	37.35	0.35
ACG	2	0.42	86	0.26	0.03	13.7-15.0	14.33	4.54	0.68
ACT	12	2.54	692	2.07	0.23	9.3-33.7	19.22	41.56	0.34
AGC	1	0.21	26	0.08	0.01	8.7-8.7	8.67	0	0.66
AGG	2	0.42	75	0.22	0.03	10.3-14.7	12.50	17.60	0.65
Subtotal	35	7.40	1754	5.26	0.59	9.0-33.7	25.06		0.52
CTAT	31	6.55	2462	7.38	0.83	3.25-60.0	19.85	62.15	0.27
ACAT	4	0.85	202	0.61	0.07	7.8-18.0	12.63	34.71	0.26
CATC	13	2.75	594	1.78	0.20	6.25-32.0	11.42	64.03	0.48
AAAC	3	0.63	142	0.43	0.05	7.0-20.0	11.83	49.09	0.26
CTAT	2	0.42	167	0.50	0.06	16.8-25.0	20.90	19.62	0.25
GTGC	3	0.63	142	0.43	0.05	10.8-13.0	11.83	7.62	0.77
GAAG	1	0.21	47	0.14	0.02	11.8-11.8	11.75	0	0.48
TCTG	6	1.27	322	0.96	0.11	7.0-22.0	13.42	43.08	0.49
ACAC	3	0.63	318	0.95	0.11	22.0-33.3	26.50	18.40	0.60
GTCA	1	0.21	35	0.10	0.01	8.75-8.75	8.75	0	0.52
TCTT	1	0.21	48	0.14	0.02	12.0-12.0	12.00	0	0.25
Subtotal	68	14.38	4479	13.42	1.52	3.25-33.3	16.47		0.37
TCTGG	1	0.21	53	0.16	0.02	10.6-10.6	10.60	0	0.55
ATCTA	1	0.21	25	0.07	0.01	5.0-5.0	5.00	0	0.20
Subtotal	2	0.42	78	0.23	0.03	5.0-10.6	7.80		0.38
TGCGTG	1	0.21	86	0.26	0.03	14.3-14.3	14.33	0	0.65
TAGGGC	1	0.21	51	0.15	0.02	8.5-8.5	8.50	0	0.55
GGTCCT	1	0.21	40	0.12	0.01	6.7-6.7	6.67	0	0.68
GGTCTT	2	0.42	197	0.59	0.07	5.8-27.0	16.42	68.30	0.55
AATACA	1	0.21	34	0.10	0.01	5.7-5.7	5.67	0	0.12
ACACAC	1	0.21	88	0.26	0.03	14.7-14.7	14.67	0	0.58
Subtotal	7	1.48	652	1.95	0.22	5.0-27.0	15.52		0.53
Total	473	1	33370	1	11.36				

中变异能力最大的是 CATC 重复类型(64.03%); 五核苷酸的各类型变异系数均为 0; 六核苷酸重复中变异能力最大的是 GGTCTT 重复类型。

2.2 微卫星不同重复类型的拷贝数

统计学发现, 在 6 种重复类型中, 拷贝数变化范

围从 3.3-234.0 不等, 主要分布在 15-60 之间, 占总微卫星位点数的 76.96%。与拷贝数的分布趋势相对应, 微卫星序列的长度主要分布在 30-80 个碱基的范围内。重复单位拷贝数范围以两核苷酸重复拷贝数最广, 在 13.0-234.0 之间, 其中, 主要分布在 10-50 之间(293 个), 占总微卫星位点数的 81.16%; 其次是四

碱基重复类型, 在 3.25–33.25 之间, 主要分布在 5.0–25.0 之间(56 个), 占总微卫星位点数的 82.35%; 三碱基重复类型分布在 9.0–33.7 之间; 六碱基重复类型分布在 5.0–27.0 之间; 以五碱基重复类型分布范围最窄, 在 5.0–10.6 之间。5 种长度重复单位类型微卫星平均拷贝数分别为 36.79、25.06、16.47、7.80、15.52, 其中, 以两核苷酸重复类型的平均拷贝数最高, 五核苷酸重复类型的平均拷贝数最低。

3 讨论

3.1 微卫星的结构与序列组成

单碱基重复是微卫星的一种重要类型, 不同物种间差异极大, 如中国对虾的单碱基重复占总微卫星数的 3.67%, 红鳍东方鲀的单碱基重复占总微卫星数的 0.81%, 而本研究所分析的 686 条虾虎鱼科微卫星序列中未发现单碱基重复类型, 除了分析过程软件设置的参数较严格以外, 还可能与虾虎鱼科的物种特异性有关。

本研究只发现了 AC 和 AG 两种二碱基重复类型, 未发现 AT 和 GC 重复类型。与其他海洋生物微卫星特征相比较, 虾虎鱼科与鳊(孙典巧等, 2011)和红鳍东方鲀(崔建洲等, 2006)的微卫星重复类型相似, 即 AC 含量最丰富; 而中国明对虾和日本囊对虾的 AT 重复拷贝最丰富(栾生等, 2007; 高焕等, 2004), 三疣梭子蟹的 AG 重复拷贝最丰富(宋来鹏等, 2008)。另外, 与这些物种的相同点为微卫星重复类别中均缺少 GC 重复。Schoreeret 等(1992)推测, GC 含量稀少的根本原因是: 基因组 DNA CpG 的甲基化使得胞苷酸(C)很容易经过脱氨基作用转变成胸腺嘧啶(T), 而少量的 GC 又是维持 DNA 热力学稳定性所必需的, 这样就使 GC 重复减少, 同时突变后的序列 TG(即 AC 类型)相应增加, 这同样可以在一定程度解释虾虎鱼科微卫星重复类型中 AC 含量最多的现象。

虾虎鱼科微卫星序列的三碱基重复类型共出现了 8 种, 以 ACT 重复类型所占比例最多, 其次是 ATC 类型; 没有发现 GCC 和 ACC 两种重复类型。与其他水生生物微卫星特征相比较发现, 三疣梭子蟹的三碱基重复也是以 ACT 重复类型出现次数最多(宋来鹏等, 2008); 而日本囊对虾(高焕等, 2004)和红鳍东方鲀(崔建洲等, 2006)则以 AAT 重复类型最多。Pearson 等(1998)报道, 在所有的微卫星重复类型中, 三碱基重复微卫星最为特殊, 一些三碱基重复与某些遗传性病变密切相关, 而通过改变相关基因的表达则有望控制某些遗传疾病的发生(Cummings *et al*, 2000)。

在虾虎鱼科微卫星重复序列的四碱基重复类型中, 共出现了 11 种微卫星重复类型。其中, CTAT 出现次数最多, 共 31 个, 其次是 CATC, 共 13 个, 这 2 种重复类型占四碱基重复类型的 64.7%。据报道, AAAX (AAAT、AAAC、AAAG)在灵长类和啮齿类中最丰富(Li *et al*, 2005), 在鳊鱼中, AAAC 在四核苷酸的重复序列中出现的频率也最高, 而在虾虎鱼科鱼类中该类型重复并不丰富, 仅出现 3 次。在红鳍东方鲀和日本囊对虾中, 均以 AGAT 重复出现的频率最高(崔建洲等, 2006; 高焕等, 2004)。

虾虎鱼科微卫星的五碱基重复类型出现的种类和数量均较少, 在整个微卫星重复类型中只出现了 2 个重复, 这与红鳍东方鲀基因组中五碱基重复类型的微卫星数量也较少的情况类似。五碱基重复类型的微卫星的生物学意义报道的不多, De Fonzo 等(1998)发现五碱基重复可能与一些人类疾病相关。

在虾虎鱼科微卫星重复序列中, 六碱基重复类型共发现 6 种重复类型, 其中只有 GGTCTT 出现 2 次, 其他每种重复类型均各出现 1 次, 并且各种重复拷贝类别的拷贝数均较少。对六碱基重复及其生物学意义的研究比较少, 还有待进一步研究。

3.2 微卫星重复单位拷贝数与变异能力分析

变异系数用来衡量观测值变异程度, 变异系数越大, 该重复类型的变异水平就越高(Larmuseau *et al*, 2007)。由表 3 可以看出, 变异系数最大的是 AG 重复类型(76.51%), 其次是 GGTCTT 重复类型(68.3%), 因此, 变异能力与碱基重复单位数量并无相关性, 而较长重复单位的类型有较大的变异能力。在所有的微卫星重复类型中, 二核苷酸重复类型的平均变异系数最高, 而五核苷酸重复类型的变异系数最小, 说明二核苷酸重复类型是基因组微卫星中最活跃的类型。分析结果表明, 拷贝数分布范围越广, 变异系数越大, 即变异系数与拷贝数分布范围正相关。

3.3 研究虾虎鱼科鱼类微卫星分布的意义

虾虎鱼种类多、分布广, 是经过长期进化和生存竞争保留下来的重要海洋鱼类, 与其他物种共同组成了有机生命链。刘群等(1997)认为, 渤海存在 2 条主要食物链, 其中一条是多毛类、虾、蟹通过虾虎鱼等小型鱼类转为石首鱼科等底层鱼类的食物。虾虎鱼类正处于底栖食物链的中继地位, 具有重要的生态学意义。如黄鳍刺虾虎鱼(*Acanthogobius flavimanus*)广泛分布在黄渤海沿岸软泥或砂泥底质的河口区, 是莱州湾的重要虾虎鱼类, 在食物链中, 它既是下一级经济

仔稚鱼和虾蟹类幼体的捕食者,又是上一级多种经济鱼类的重要饵料,在整个生态系统中占据重要地位(龚艳丽等,2000)。因此,虾虎鱼作为生态系统的重要组成部分,了解其物种多样性以及群体遗传学特性将会有重要的生态学意义。

虽然虾虎鱼种类繁多,但仅在少数几种虾虎鱼中有微卫星分离的报道(Mendonca *et al*, 2001; Martina *et al*, 2007; Dufour *et al*, 2007; Hepburn *et al*, 2005), 虾虎鱼“科”级水平微卫星的特征分布均不清楚,因此,对虾虎鱼类微卫星还有待进一步的研究。本研究发现,虾虎鱼科微卫星在总微卫星重复序列中占绝对优势的二碱基重复拷贝数主要分布在 10–50 之间,即虾虎鱼基因组中的微卫星长度主要分布在 20–100 个碱基的长度范围内,说明在长期进化中微卫星位点积累的变异相对较大,因而,预测这些微卫星序列具有丰富的多态信息,为下一步优化和筛选特异性微卫星标记奠定了基础。

参 考 文 献

- 崔建洲, 申雪艳, 杨官品, 等. 红鳍东方鲀基因组微卫星特征分析. 中国海洋大学学报(自然科学版), 2006, 36(2): 249–254
- 高焕, 刘萍, 孟宪红, 等. 中国对虾(*Fenneropenaeus chinensis*)基因组微卫星特征分析. 海洋与湖沼, 2004, 35(5): 424–431
- 龚艳丽, 孟田湘, 万瑞景. 黄鳍刺虾虎鱼(*Acanthogobius flavimanus*)前期幼鱼摄食习性和摄食率的初步研究. 青岛海洋大学学报(自然科学版), 2000, 30(4): 609–613
- 刘群, 苏振明, 曾晓起, 等. 渤海虾虎鱼类的生物学调查——关于区系分布与生态学基础的研究. 青岛海洋大学学报(自然科学版), 1997, 27(2): 157–165
- 李思忠, 陈星玉, 陈小平, 等译. 世界鱼类. 台湾: 水产出版社, 1994, 355–357
- 栾生, 孔杰, 王清印, 等. 日本囊对虾(*Marsupenaeus japonicus*)基因组微卫星特征分析. 自然科学进展, 2007, 17(6): 731–740
- 宋来鹏, 刘萍, 李健, 等. 三疣梭子蟹基因组微卫星特征分析. 中国水产科学, 2008, 15(5): 738–744
- 孙典巧, 孙悦娜, 王日昕, 等. 鲩鱼 EST 序列中微卫星标记的初步筛选及特征分析. 水生生物学报, 2011, 35(5): 753–760
- 王蕾, 刘继红, 张立冬, 等. 牙鲆基因组(CAG)_n 微卫星 DNA 特征分析. 中国水产科学, 2009, 16(6): 807–815
- 伍汉霖, 钟俊生. 中国动物志 硬骨鱼纲鲈形目(五)虾虎鱼亚目. 北京: 科学出版社, 2008, 1–988
- Bai CC, Liu SF, Zhuang ZM, *et al*. Seventeen polymorphic microsatellite markers for the Javelin goby, *Synechogobius hasta* (Gobiidae). Genet Mol Res, 2012, 11(2): 1465–1468
- Cummings CJ, Zoghbi HY. Fourteen and counting; unraveling trinucleotide repeat diseases. Hum Mol Gen, 2000, 9(6): 909–916
- De Fonzo V, Bersani E, Aluffi-Pentini F, *et al*. Are only repeated triplets guilty? J Theor Biol, 1998, 194(1): 125–142
- Dufour BA, Hogan TM, Heath DD. Ten polymorphic microsatellite markers in the invasive round goby (*Neogobius melanostomus*) and cross-species amplification. Mol Ecol Notes, 2007, 7(6): 1205–1207
- Gao H, Kong J. The microsatellite and minisatellites in the genome of *Fenneropenaeus chinensis*. DNA Sequence, 2005, 16(6): 426–436
- Hepburn RI, Mottillo EP, Bentzen P, *et al*. Polymorphic microsatellite loci for the masked goby, *Coryphopterus personatus* (Gobiidae). Conserv Genet, 2005, 6(6): 1059–1062
- Katti MV, Ranjekar PK, Gupta VS. Differential distribution of simple sequence repeats in eukaryotic genome sequences. Mol Biol Evol, 2001, 18(7): 1161–1167
- Larmuseau MHD, Hellemans B, Van Houdt JKJ, *et al*. Development and characterization of nine polymorphic microsatellite markers in the sand goby *Pomatoschistus minutus* (Gobiidae). Mol Ecol Notes, 2007, 7(1): 147–149
- Li C, Han B. Diversity of simple sequence repeats in *Arabidopsis thaliana* and rice. Acta Bot Sin, 2004, 46(5): 603–609
- Li CY, Li JB, Zhou XG, *et al*. Frequency and distribution of microsatellites in the genome of filamentous fungus, *Neurospora crassa*. Agr Sci China, 2005, 4(2): 118–124
- Ma ZQ, Roder M, Sorrells ME. Frequencies and sequence characteristics of di-, tri-, and tetranucleotide microsatellite in wheat. Genome, 1996, 39(1): 123–130
- Martina V, Marketa O, Andrea S, *et al*. Isolation and characterization of microsatellites in *Neogobius kessleri* (Perciformes, Gobiidae) and cross-species amplification within the family Gobiidae. Mol Ecol Notes, 2007, 7(4): 701–704
- Mendonca H, Smith J, Brinegar C. Isolation and characterization of four microsatellite loci in the tidewater goby (*Eucyclogobius newberryi*). Mar Biotechnol, 2001, 3(2): 91–95
- Pearson CE, Sinden RR. Trinucleotide repeat DNA structures: dynamic mutations from dynamic DNA. Curt Opin Struct Biol, 1998, 8(3): 321–330
- Prasad MD, Muthulakshmi M, Madhu M, *et al*. Survey and analysis of microsatellites in the silkworm, *Bombyx mori*: frequency, distribution, mutations marker potential and their conservation in heterologous species. Genetics, 2005, 169(1): 197–214
- Schoreer DF, Gartlar SM. Analysis of CpG suppression in methylated and nonmethylated species. Proc Natl Acad Sci USA, 1992, 89(3): 957–961
- Sugaya T, Ikeda M, Taniguchi N. Relatedness structure estimated by microsatellites DNA markers and mitochondrial DNA

polymerase chain reaction restriction fragment length polymorphism analyses in the wild population of kuruma prawn *Paenaeus japonicus*. *Fisheries Science*, 2002, 68(4): 793–802

Toth G, Gaspari ZJ. Microsatellites in different eukaryotic genomes survey and analysis. *Genome Res*, 2000, 10(7): 967–981

Wang Z, Weber JL, Zhong G, *et al.* Survey of plant short tandem DNA repeats. *Theor Appl Genet*, 1994, 88(1): 1–6

(编辑 冯小花)

Characteristic Analysis of Microsatellite DNA in the Genome of Gobiidae

BAI Cuicui^{1,2,3}, LIU Shufang^{1,2①}, ZHUANG Zhimeng¹

(1. Key Laboratory of Sustainable Development of Marine Fisheries, Ministry of Agriculture, Yellow Sea Fisheries Research Institute, Chinese Academy of Fishery Sciences, Qingdao 266071;

2. Laboratory for Marine Fisheries Science and Food Production Processes, Qingdao National Laboratory for Marine Science and Technology, Qingdao 266200; 3. Shandong Agricultural University, Taian 271018)

Abstract The family Gobiidae is one of the largest fish families consisting of more than 200 genera and 2000 species. This family was found in near shore seawater, brackish water, and freshwater in the temperate and tropical areas all over the world. In this study we applied microsatellite DNA (short sequence repeat, SSR) to characterize the genetic structure of Gobiidae. We analyzed 153 SSR sequences amplified in our laboratory and 535 SSR sequences downloaded and screened from GenBank. The 686 sequences originated from 19 different species of Gobiidae. The total length of these sequences was about 295062 bp, in which 473 microsatellite repeat loci were found by using the bio-software Tandem Repeats Finder. The length of the repeat sequences was 33370 bp. We found that 361 dinucleotide repeat sequences accounted for the majority (76.32%) of the 473 repeat loci. Moreover, the AC repeat sequences were the most common (94.18) in the dinucleotide repeat sequences, whereas AT and GC repeat sequences were not observed. The number of the trinucleotide repeat sequences was 35 (7.4%), including 12 ACT repeat sequences that were the majority (34.29%). There were 68 tetranucleotide repeat sequences including 31 CTAT that was the dominant (14.38%). The numbers of both tetranucleotide and hexanucleotide repeat sequences were 2 (0.42%) and 7 (1.48%) respectively. Intriguingly repeats of single nucleotide sequences were absent. Our results about the distribution of microsatellite contributed to a better understanding of the genomic structure of goby fish, and might provide important information for the development and application of goby fish microsatellite markers.

Key words Gobiidae; Microsatellite; Genome

① Corresponding author: LIU Shufang, E-mail: liusf@ysfri.ac.cn