

巨鲃(*Bagarius yarrelli*)全基因组 微卫星分布特征分析

杨汶珊^{1,2}, 唐荣叶^{1,2}, 苏孟园^{1,2}, 徐杰杰^{1,2}, 王 涛^{1,2}, 尹绍武^{1,2}

(1. 南京师范大学海洋科学与工程学院, 江苏 南京 210023)

(2. 江苏省特色水产育种与绿色高效养殖技术工程研究中心, 江苏 南京 210023)

[摘要] 对已公布在 NCBI 数据库中的巨鲃(*Bagarius yarrelli*)全基因组测序结果, 使用 MISA 软件对巨鲃全基因组中的微卫星进行筛选并分析其数量与分布特征. 在巨鲃基因组 570 806 968 bp 序列中, 共筛选出 360 235 个完整型微卫星, 其长度为 6 998 449 bp, 占基因序列总长度的 1.23%. 在 6 种完整型微卫星中, 微卫星数量最多的是单碱基类型, 约占总数的 44.65%, 其余碱基类型数量排序为二碱基(43.29%)、三碱基(6.12%)、四碱基(4.80%)、五碱基(1.02%)和六碱基(0.11%). 基因组中数量最多的前 10 种微卫星类别分别为: A、AC、AG、AT、AAT、C、ATAG、AAAT、ACT 和 ATC.

[关键词] 巨鲃, 全基因组, 微卫星, 分布特征

[中图分类号] S917 **[文献标志码]** A **[文章编号]** 1672-1292(2021)03-0062-07

Analysis of Microsatellite Distribution Characteristics in the Whole Genome of *Bagarius yarrelli*

Yang Wenshan^{1,2}, Tang Rongye^{1,2}, Su Mengyuan^{1,2}, Xu Jiejie^{1,2}, Wang Tao^{1,2}, Yin Shaowu^{1,2}

(1. School of Marine Science and Engineering, Nanjing Normal University, Nanjing 210023, China)

(2. Jiangsu Province Engineering Research Center for Aquatic Animals Breeding and
Green Efficient Aquacultural Technology, Nanjing 210023, China)

Abstract: The study uses the published genome-wide sequencing results of the *Bagarius yarrelli* on NCBI to screen and analyze the number and distribution of the microsatellites of the whole genome using MISA software. In the 570 806 968 bp longsequence of the *Bagarius yarrelli* genome, 360 235 perfect microsatellites are screened with a length of 6 998 449 bp, accounting for 1.23% of the total length of the genome sequence. Among the six types of microsatellites, mononucleotide is the most, accounting for 44.65% of the total, the other bases are dinucleotide(43.29%), trinucleotide(6.12%), tetranucleotide(4.80%), pentanucleotide(1.02%) and hexanucleotide(0.11%). The top 10 microsatellite copy types in the genome are: A, AC, AG, AT, AAT, C, ATAG, AAAT, ACT and ATC.

Key words: *Bagarius yarrelli*, genome, microsatellites, distribution characteristics

巨鲃(*Bagarius yarrelli*)属于脊椎动物亚门(Vertebrate), 硬骨鱼纲(Teleostei), 鲇形目(Siluriformes), 鲇科(Sisoridae), 鲇属(*Bagarius*). 巨鲃的皮肤为棕栗色, 身体呈流线形, 头部宽大、略扁. 因肌肉呈黄色, 俗称“面瓜鱼”“黄鱼”, 个体可超过 50 kg^[1]. 巨鲃在自然环境下, 以摄食鱼虾和水生昆虫等为主, 偶食植物碎屑; 在人工驯养条件下, 仅摄食动物饵料^[2]. 此外, 巨鲃肉质风味独特, 深受云南各地人民的喜爱. 目前, 云南省水产研究所等单位已成功地对巨鲃进行了人工繁殖^[3].

微卫星又称简单序列重复(simple sequence repeats, SSRs), 是指以 1~6 bp 核苷酸为基本重复单位组成的重复序列, 在自然界的各种生物基因组中均有分布^[4]. 微卫星因具有易于检测、多态信息含量高、高效稳

收稿日期: 2020-10-04.

基金项目: 江苏省农业重大新品种创制项目(PZCZ201742)、江苏省重点研发计划(现代农业)重点项目(BE2017377)、江苏省农业科技自主创新资金项目(CX(19)2034)、南京师范大学大学生创新创业训练计划项目.

通讯作者: 王涛, 副教授, 研究方向: 水产动物遗传育种. E-mail: seawater88@126.com

定、分布广泛等优点,已被广泛应用于鱼类种群遗传多样性分析^[5-6]、遗传育种^[7]、亲子鉴定^[8]等研究中。

目前巨鲉全基因组已完成测序,Scaffold N50=3 129 371 bp,Contig N50=1 854 961 bp,并于2019年在线公布,但还未见巨鲉基因组水平上微卫星相关的研究报道。本研究通过对巨鲉全基因组中的微卫星进行筛选,统计其所含不同类型的微卫星的数量、丰度、密度等数据,分析其分布特征及组成情况,为今后开发巨鲉高质量的微卫星标记奠定一定的基础。

1 材料和方法

1.1 材料

巨鲉全基因组从 NCBI 数据库 (<https://www.ncbi.nlm.nih.gov/genome/?term=Bagarius+yarrelli>) 下载,该基因组序列总大小约为 570 Mb,下载的巨鲉全基因组所有序列均以 FASTA 文件格式保存。

1.2 方法

利用微卫星筛选软件 MISA (<http://pgrc.ipk-gatersleben.de/misa/>) 对巨鲉全基因组进行扫描,筛选符合条件的完整型微卫星序列,在筛选后的结果中统计出 1~6 个碱基重复的完整型微卫星,使用 EXCEL 统计分析出各碱基类型的数量和分布情况,分别列出其分布特征和占比情况。搜索标准参考 MISA 默认参数设置,即要求单碱基重复 ≥ 10 次,二碱基重复 ≥ 6 次,三、四、五、六碱基重复 ≥ 5 次。经过前期的大量研究可知,以此标准搜索全基因组微卫星得出的结果最优^[9]。根据起始碱基顺序的差异及碱基互补配对原则,对属于同一类别的微卫星进行同类合并,如三碱基 AAT,可以与之兼并的有 ATA、TAA、TTA、TAT 和 ATT。

2 结果与分析

2.1 巨鲉全基因组各重复类型微卫星分布特征

如表 1 所示,巨鲉全基因组完整型微卫星总数为 360 235 个,6 种完整型微卫星的总长度为 6 998 449 bp,占基因组的 1.23%,丰度为 631 个/Mb,密度为 12 261 bp/Mb。6 种微卫星丰度与微卫星数量变化的情况相同,即微卫星数量越多则丰度越大。每种微卫星的密度与微卫星的长度有关,即某一类型微卫星的长度越长,其密度越高。

如图 1 所示,数量最多的是单碱基类型,其次是二碱基类型,其余微卫星占比均小于 10%,依次为三碱基、四碱基、五碱基类型,占比最少的是六碱基类型。丰度随着 1~6 bp 碱基的基本重复单位的增加而降低。

如图 2 所示,在巨鲉 1~6 个碱基重复的 6 种完整型微卫星中,长度最长的是二碱基类型,其余根据降序排列依次为单碱基、四碱基、三碱基、五碱基。长度最短的为六碱基。

表 1 巨鲉全基因组中不同类型微卫星统计

Table 1 Different types of microsatellite statistics in the *B. yarrelli* genome

重复类型	数量	总长度/bp	丰度/(个/Mb)	密度/(bp/Mb)
单碱基	160 855	2 028 830	282	3 554
二碱基	155 942	3 789 386	273	6 639
三碱基	22 060	450 993	39	790
四碱基	17 279	576 300	30	1 010
五碱基	3 687	137 340	6	241
六碱基	412	15 600	1	27
总体	360 235	6 998 449	631	12 261

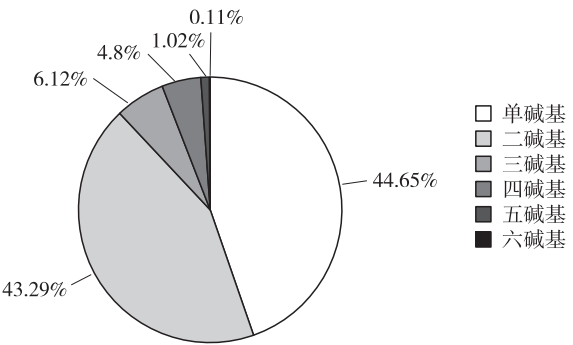


图 1 巨鲉不同类型微卫星数量占比

Fig. 1 The proportion of microsatellites of different types in *B. yarrelli*

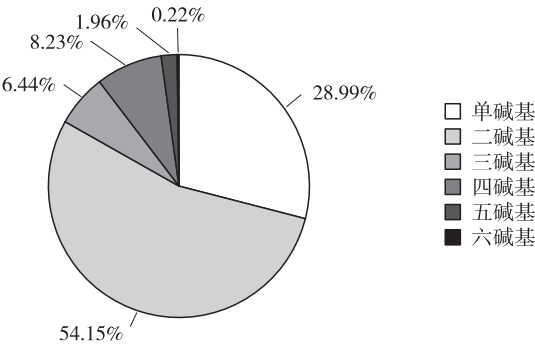


图 2 巨鲉不同类型微卫星总长度占比

Fig. 2 The proportion of different types of total length of microsatellites in *B. yarrelli*

2.2 巨鲉全基因组不同类型微卫星重复次数分布

巨鲉基因组中各类型微卫星重复次数集中在 5~30 次之间,数量占全部微卫星的 99.66%。在 6 种重复类型微卫星中,微卫星数量最多的拷贝次数均为 MISA 软件筛选出来的最小重复次数。如图 3 所示,单碱基重复的分布范围最广,最低重复次数为 10 次,最高重复次数为 249 次,其中微卫星主要重复次数为 10~20 次,其数量占单碱基 SSR 总数的 95.26%;二碱基重复的分布范围也较广,最低重复次数为 6 次,最高重复次数为 130 次,其中微卫星主要重复次数为 6~30 次,其数量占二碱基微卫星总数的 96.16%;三碱基微卫星重复拷贝数分布在 5~35 次之间,其中微卫星主要重复次数为 5~15 次,数量占三碱基微卫星总数的 99.18%;四碱基最低重复次数为 5 次,最高重复次数为 63 次,微卫星主要重复次数为 5~20 次,数量占四碱基微卫星总数的 96.81%;五碱基重复拷贝数分布在 5~40 次,微卫星主要重复次数为 5~15 次,数量占比为 96.53%;六碱基重复拷贝数分布范围最小,最低重复次数为 5 次,最高重复次数仅为 21 次,微卫星主要重复次数集中在 5~10 次,数量占比为 94.90%。

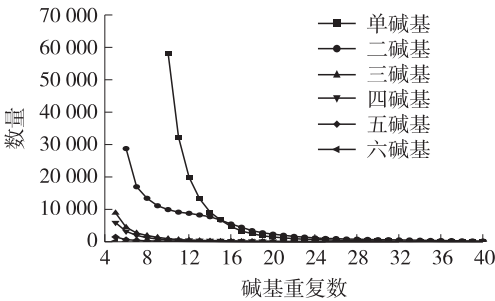


图 3 巨鲉各类型微卫星不同拷贝数分布
Fig. 3 Distribution of different copy numbers of various types of microsatellites in *B. yarrelli*

2.3 巨鲉全基因组微卫星各重复碱基类别特征

如图 4 所示,巨鲉全基因组微卫星中的优势拷贝类别按照数量从高到低排列分别为:A、AC、AG、AT、AAT、C、ATAG、AAAT、ACT 和 ATC,这 10 种拷贝类别总数为 339 090 个,占微卫星重复拷贝类别总数的 94.13%。

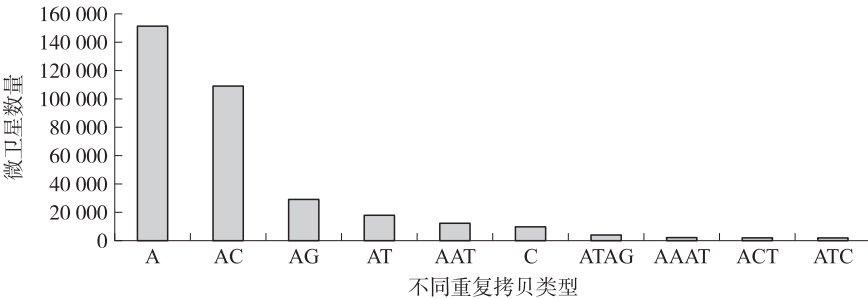


图 4 巨鲉全基因组中分布最多的不同重复拷贝类型微卫星
Fig. 4 Distribution of the most frequent microsatellite motifs of *B. yarrelli*

如表 2 所示,单碱基类别中,单碱基数量以 A 类别(A/T)为主,占有绝对的碱基优势,其次是 C 类别。二碱基类别中,AC 类别是主要的重复单元,其次分别是 AG 类别和 AT 类别;CG 类别数量最少,仅有 138 个,占比 0.09%。三碱基类型中,AAT 类别数量最多,其次分别是 ACT 类别和 ATC 类别;CCG 类别在三碱基 10 种类别中数量最少,只有 93 个,占比 0.42%。四碱基类型中,ATAG 类别重复数量最多,其次分别是 AAAT 类别和 AGAC 类别。五碱基类型中,AATCT 类别重复数量最多,其次分别是 AAGAG 类别和 AATAT 类别。六碱基类型中,AACCCT 类别重复数量最多,其次是 AGGGTT 类别,其余类别重复数量较少。

表 2 巨鲉全基因组中各碱基重复类型中的优势碱基类别

Table 2 The dominant base classes in each base repeat type in the <i>B. yarrelli</i> genome							
SSR 类型	类别	数量	占各类型比例/%	SSR 类型	类别	数量	占各类型比例/%
单碱基	A	151 107	93.94	四碱基	ATAG	3997	23.13
	C	9 748	6.06		AAAT	2191	12.68
二碱基	AC	108 866	69.81		AGAC	1955	11.31
	AG	29 061	18.64	五碱基	AATCT	651	17.66
	AT	17 877	11.46		AAGAG	344	9.33
	AAT	12 288	55.70		AATAT	331	8.98
三碱基	ACT	1988	9.01	六碱基	AACCCT	102	24.76
	ATC	1967	8.92		AGGGTT	74	17.96
					AACCCG	13	3.16

3 讨论

3.1 巨鲈微卫星在全基因组中的含量分析

本研究以巨鲈全基因组为基础,利用生物信息学软件对巨鲈基因组中6种完整型微卫星进行搜索,共搜索到360 235个完整型微卫星,其长度为6 998 449 bp,占基因组序列总长度的1.23%。分析表明,巨鲈全基因组微卫星含量(1.23%)与已公布的人(*Homo sapiens*)(3%)^[10]、啮齿目的大鼠(*Rattus norvegicus*)(1.41%)^[11]和小鼠(*Mus musculus*)(2.85%)^[12]相比较低,但高于偶蹄目的牛(*Bos taurus*)(0.48%)、牦牛(*Bos mutus*)(0.58%)、藏羚羊(*Pantholops hodgsoni*)(0.54%)^[13-15]、食肉目的大熊猫(*Ailuropoda melanoleuca*)(0.64%)、北极熊(*Ursus maritimus*)(0.79%)^[16]、鲀形目的红鳍东方鲀(*Takifugu rubripes*)(0.73%)、菊黄东方鲀(*Takifugu flavidus*)(0.73%)、双斑东方鲀(*Takifugu bimaculatus*)(0.84%)、黑青斑河鲀(*Tetraodon nigroviridis*)(1.06%)^[9]。巨鲈全基因组微卫星含量与同为鲀形目的黄颡鱼(*Pelteobagrus fulvidraco*)(1.8%)^[17]有一定差异。这些结果表明,不同物种的微卫星在全基因组中含量具有较大的差异,这种差异一方面可能是不同物种基因组大小的区别造成的,另一方面可能还有物种进化差异的原因。

3.2 巨鲈全基因组各碱基类型微卫星特征分析

对比同属于鱼纲的物种,结果显示巨鲈基因组中的优势碱基类型排序与骨舌鱼目的美丽硬仆骨舌鱼(*Scleropages formosus*)^[18]以及鲀形目的红鳍东方鲀、菊黄东方鲀、双斑东方鲀^[9]的分布顺序一致,也与鲤形目的鲤鱼(*Cyprinus carpio*)^[19]的优势碱基类型分布顺序完全一致。但与同为鲀形目的黄颡鱼^[17]的优势碱基类型分布顺序不同,也与鲽形目的大菱鲆(*Scophthalmus maximus*)^[20]的优势碱基分布不同。

对比动物界各纲的物种,结果显示鱼纲(如巨鲈、美丽硬仆骨舌鱼、红鳍东方鲀、鲤鱼、黄颡鱼和大菱鲆)^[9,17-20]主要以单碱基和二碱基占优势。鸟纲(如白鹭(*Egretta garzetta*)、藏鸡(*Gallus gazette domesticus*)、绿尾虹雉(*Lophophorus lhuysii*)、红原鸡(*Gallus Lophophores*)和虎皮鹦鹉(*Melopsittacus undulatus*))^[21-25]和哺乳动物(如猪(*Sus scrofa*)、林麝(*Moschus berezovskii*)、大熊猫、北极熊、牦牛和水牛(*Bubalus bubalis*))^[14,16,26-27]的优势碱基类型排序均以单碱基微卫星占优势,而六碱基微卫星最少,但哺乳动物中的大鼠以二碱基微卫星占优势,六碱基微卫星最少^[11],说明大部分鸟纲动物和哺乳动物的微卫星分布顺序以单碱基微卫星占优势,六碱基微卫星占比最少,个别鸟纲和哺乳动物分布规律略有不同。而对比低等无脊椎动物中的一些昆虫纲动物,如草地贪夜蛾(*Bubalus bubalis*)、果蝇(*Drosophila melanogaster*)、蚊子(*Anopheles gamblae*)^[28-30],其优势碱基类型排序各有不同,优势碱基排序较为随机。在原核生物和酵母(*Saccharomyces cerevisiae*)的基因组中,处于优势的重序列类型是三碱基^[18]。

从以上对比结果中可发现,巨鲈与鸟纲、昆虫纲以及硬骨鱼纲中的鲀形目与鲤形目中的一些物种的优势微卫星类型排序相同,但与同为鲀形目的黄颡鱼的优势碱基类型排序不同,这表明物种的亲缘关系与其优势碱基类型排序并无明显联系。Webster等^[31]研究了黑猩猩与人类基因组的微卫星分布,研究结果表明,即使物种的亲缘关系相近,其微卫星分布也存在很大差别。本研究中对鸟纲、哺乳动物、昆虫纲以及硬骨鱼纲的一些物种的比较分析也支持了这一说法。另外可发现,大多数物种为单碱基或二碱基占优势,少部分真菌及低等动物则为三碱基或六碱基数目最多,如酵母^[32]和二斑叶螨(*Tetranychus urticae*)^[33]以三碱基占优势,蚊子以六碱基占优势^[30]。黄杰等^[5]认为高等生物基因组更倾向于单碱基微卫星占优势,鱼纲和鸟纲以单碱基或二碱基占优势,未见三碱基占优势的报道。在原核生物和酵母的基因组中,处于优势的重序列类型是三碱基,而比其更高等的生物基因组则倾向于二碱基和单碱基重序列类型^[18]。这些结果一方面说明不同物种基因组的优势类型微卫星不同,另一方面说明物种在向高等动物的进化过程中倾向于选择二碱基和单碱基重序列类型。

3.3 巨鲈全基因组微卫星各碱基类型不同类别特征分析

在巨鲈6种碱基类型微卫星中,A、AC、AAT、ATAG、AATCT和AACCT类别分别为1~6碱基重复类型中数量最多的一类,可以看出明显的A/T优势。

在巨鲈单碱基重复中,A类别(A/T)数目最多,在单碱基微卫星中占比93.94%,这与已公布的人、黄颡鱼、红鳍东方鲀、中华按蚊(*Anopheles sinensis*)^[9-10,17,34]等大部分已公布的物种全基因组微卫星特征一致。

在二碱基重复中,AC 类别微卫星数量最多,其次是 AG 与 AT 类别,与人、黄颡鱼、美丽硬仆骨舌鱼和红鳍东方鲀^[9-10,17,21]等物种二碱基前三类别排序相一致. 与其他物种对比发现,大多数高等哺乳动物和硬骨鱼纲的二碱基优势重复类型均为 AC 类别,而鸟纲的二碱基优势重复类型为 AT 类别,推测高等哺乳动物、鸟纲和硬骨鱼纲的优势重复类型的差别可能与物种进化有关. 有研究表明,AT 类别重复为陆地生物的优势重复类型,而海洋生物海鞘(*Ciona intestinalis*)中 CG 和 AG 类别的重复次数较多,可看出微卫星优势碱基类型在陆地和海洋生物间并不相同^[35]. 同时对比显示,巨鲉、美丽硬仆骨舌鱼和黄颡鱼均属于淡水生物,其结果与海洋生物海鞘的分析有相同和不同之处,相同之处是 AG 类型均为排序第二的优势碱基类别,差异之处在于海鞘的 CG 类别重复次数较多,而在这三种鱼中 CG 类型均最少. 造成这种异同的原因可能是海洋生物和淡水生物生存环境的差别所致. 另外关于 CG 类别比 AT 类别少的原因,有研究认为,在 DNA 复制的过程中会产生滑移,从而形成了微卫星^[36],推测可能是由于 AT 之间的双键比 CG 之间的三键更容易断裂,因此更容易造成碱基滑动,从而导致 AT 重复拷贝类别比 CG 重复拷贝类别在数量上更具优势.

在三碱基重复中,AAT 类别占绝对优势,这与虎皮鹦鹉、红原鸡和黄颡鱼^[17,25-26]相同. 此外与其他高等动物、鸟纲和硬骨鱼纲的三碱基优势类别对比中并未发现明显规律. 有研究表明,某些遗传性病变与三碱基重复有关^[37],某些微卫星会由于其不稳定性改变基因的表达从而导致某些遗传性病变的发生^[38]. 因此,通过研究巨鲉的三碱基微卫星可以了解巨鲉某些疾病的发生原因,同时还可以帮助确定和预测一些人类疾病相关的基因.

在四碱基重复中,ATAG 和 AAAT 类别的数量排在前三. 通过对微卫星作用的研究表明,在遗传标记上四碱基微卫星相比于二碱基和三碱基微卫星更为精确和可靠^[39]. 因此在开展巨鲉遗传多样性分析、物种鉴定以及亲子鉴定等研究中,巨鲉的四碱基多态性微卫星标记可能提供参照. 研究结果显示,巨鲉四碱基微卫星共有 17 279 个,同时其统计到的重复类别有 31 个,大量的微卫星和较多的重复类别在开发巨鲉四碱基微卫星标记中将提供较大帮助. 吴旭东等^[40]研究表明,鲇形目鱼类的微卫星侧翼序列在同目内种间具有保守性,同目鱼类的部分微卫星引物与其他的鱼类也存在一定的通用性. 这说明微卫星序列可能随着物种的进化而不断变化,由于选择压力等原因,使得亲缘关系相近的物种其微卫星侧翼序列也接近. 本研究可为同为鲇形目的其他鱼类的微卫星开发提供参照.

AATCT、AAGAG 和 AATAT 类别为五碱基中的前三类别,六碱基中的 AACCCT 和 AGGGTT 类别占主导地位. 基因组中重复最多的前 10 种微卫星拷贝类型表明,微卫星中存在明显的 A/T 优势,而 C/G 数量较少. 魏朝明等^[41]研究认为,由于不同碱基排列编码的基因其蛋白质拥有不同的功能,所以位于编码基因上的微卫星也行使着不同的功能. 因此 A/T 的碱基组合可能拥有某些功能,与物种的某些生物特性有关. 同时还有研究表明,物种基因组中微卫星标记数越多,A 和 T 碱基所占比例越高^[42]. 对于这一现象,Schlötterer 等^[36]认为,基因组 DNA 由于甲基化的发生,胞嘧啶 C 容易脱氨基从而转变成胸腺嘧啶 T. 基因组内 CG 含量少也是维持 DNA 热力学稳定性的必要条件之一^[43]. 另外,CG 碱基重复的测序工作较为困难也可能是数据中 CG 含量少的原因^[44].

3.4 巨鲉全基因组微卫星各碱基类型拷贝数特征分析

研究表明,在巨鲉基因组中,除了细微的波动外,6 种碱基类型均随着重复拷贝数的增加其微卫星数量逐渐递减. 同时数据显示,即使不同筛选软件参数设置不同,最低参数的微卫星数量也远超过倒数第二参数设定的微卫星重复拷贝数量. 这种现象同样出现在黄颡鱼、藏羚羊、虎皮鹦鹉和红原鸡等物种全基因组微卫星中. 对此,Wierdl 等^[45]认为,一方面与微卫星稳定性有关,即随微卫星长度的增加,其稳定性会下降,因此当重复拷贝数越高时越容易发生碱基突变而成为不完整型微卫星或非微卫星序列,这使得重复数越多的微卫星数量越少;另一方面,Harr 等^[46]认为,微卫星重复拷贝数越高,突变率就越高,使得越长的微卫星序列数目越少. 同时还有研究表明,长等位基因倾向于变得更短,从而抑制了微卫星长度变长的趋势,因此,微卫星长度越长,即重复拷贝数越多,该种微卫星越不易存在^[47].

4 结论

综上所述,在对巨鲉全基因组微卫星进行统计与分析,并对比其他物种的微卫星分布特征后得出如下结论:

不同物种的微卫星在全基因中含量具有较大差异,且不同物种的优势碱基类型与物种的亲缘关系无关;

原核动物以三碱基占优势,高等动物以单碱基和二碱基占优势;

各物种的重复类别微卫星都表现出明显的 A/T 碱基类别的优势;

6 种碱基类型基本随着重复拷贝数的增加其微卫星数量逐渐递减.

目前本研究已统计了微卫星在巨鲈全基因组中的数量、丰度和密度等数据,并分析了其特征和组成情况,为巨鲈种群遗传多样性分析、图谱构建和品系亲缘关系鉴定等提供了数据支持,期望能对巨鲈微卫星标记开发提供帮助.

[参考文献] (References)

- [1] 田树魁,薛晨江,冷云,等. 巨鲈的生物学特性初步研究[J]. 水生态学杂志,2009,2(3):115-117.
- [2] 刘跃天,田树魁,冷云,等. 野生巨鲈生物学特性研究[J]. 现代农业科技,2010(18):302-303,307.
- [3] 薛晨江,张正雄,马建颜,等. 巨鲈人工繁殖初报与胚胎发育观察[J]. 水生态学杂志,2012,33(5):54-56.
- [4] KRISHNAN J,ATHAR F,RANI T S,et al. Simple sequence repeats showing 'length preference' have regulatory functions in humans[J]. Gene,2017,628:156-161.
- [5] 黄杰,周瑜,刘与之,等. 基于 454 GS FLX 高通量测序的四川山鹧鸪基因组微卫星特征分析[J]. 四川动物,2015,34(1):8-14.
- [6] PAGE R B,SANKAMETHAWEE W,PIERCE A J,et al. High throughput sequencing enables discovery of microsatellites from the puff-throated bulbul (*Alophoixus pallidus*) and assessment of genetic diversity in Khao Yai National Park, Thailand[J]. Biochemical Systematics and Ecology,2014,55:176-183.
- [7] 孙效文,张晓峰,赵莹莹,等. 水产生物微卫星标记技术研究进展及其应用[J]. 中国水产科学,2008,15(4):689-703.
- [8] 武耀,贾智英,李池陶,等. 筛选杂交鲤亲子鉴定的微卫星标记(英文)[J]. 农业生物技术学报,2012,20(5):549-559.
- [9] 徐杰杰,郑翔,张鑫宇,等. 4 种河鲈全基因组微卫星分布特征分析研究[J/OL]. 基因组学与应用生物学:1-11[2021-03-20]. <http://kns.cnki.net/kcms/detail/45.1369.q.20191126.1019.002.html>.
- [10] SUBRAMANIAN S,MISHRA R K,SINGH L. Genome-wide analysis of microsatellite repeats in humans;their abundance and density in specific genomic regions[J]. Genome Biology,2003,4(2):R13.
- [11] 涂飞云,李鹏,韩卫杰,等. 大鼠基因组微卫星分布特征分析[J]. 基因组学与应用生物学,2015,37(6):2374-2382.
- [12] 童晓玲,代方银,李斌,等. 小鼠基因组中的微卫星重复序列的数量、分布和密度(英文)[J]. 动物学报(英文版),2006,52(1):138-152.
- [13] 戚文华,蒋雪梅,肖国生,等. 牛和绵羊全基因组微卫星序列的搜索及其生物信息学分析[J]. 畜牧兽医学报,2013,44(11):1724-1733.
- [14] 戚文华,蒋雪梅,杜联明,等. 牦牛和水牛全基因组微卫星分布规律及其比较分析[J]. 基因组学与应用生物学,2015,34(7):1406-1412.
- [15] 戚文华,严超超,肖国生,等. 山羊和藏羚羊全基因组微卫星分布规律及其生物信息学分析[J]. 四川大学学报(自然科学版),2016,53(4):937-944.
- [16] 李午俊,李玉芝,杜联明,等. 大熊猫和北极熊基因组微卫星分布特征比较分析[J]. 四川动物,2014,33(6):874-878.
- [17] 徐杰杰,郑翔,李杰,等. 黄颡鱼(*Pelteobagrus fulvidraco*)全基因组微卫星分布特征分析[J]. 基因组学与应用生物学,2020,39(12):5488-5498.
- [18] 段永楠,刘奕,胡隐昌,等. 美丽硬仆骨舌鱼全基因组微卫星分布规律特征[J]. 中国农学通报,2019,35(23):152-158.
- [19] 梁霞,王慧琪,马宇璇,等. 鲤鱼(*Cyprinus carpio*)全基因组微卫星分布特征研究[J/OL]. 南京师大学报(自然科学版):1-14[2021-03-18]. <http://kns.cnki.net/kcms/detail/32.1239.N.20210113.1301.022.html>.
- [20] 阮晓红. 大菱鲆(*Turbot*)微卫星标记的筛选与应用[D]. 青岛:中国海洋大学,2009.
- [21] 黄杰,杨波,贾银平,等. 白鹭基因组微卫星分布规律研究[J]. 重庆师范大学学报(自然科学版),2019,36(5):66-71.
- [22] 郭新颖,张健,李梦柔,等. 藏鸡基因组微卫星特征分析[J]. 中国家禽,2020,42(1):116-120.
- [23] 崔凯,岳碧松. 绿尾虹雉全基因组微卫星分布规律研究[J]. 四川动物,2018,37(5):533-540.
- [24] 黄杰,杜联明,李玉芝,等. 红原鸡全基因组中微卫星分布规律研究[J]. 四川动物,2012,31(3):358-363.
- [25] 黄杰,原宝东,杨承忠. 虎皮鹦鹉全基因组中微卫星分布规律研究[J]. 野生动物学报,2017,38(3):422-426.

- [26] 戚文华,蒋雪梅,肖国生,等. 猪全基因组中微卫星分布规律[J]. 畜牧与兽医,2014,46(8):9-13.
- [27] 卢婷,王晨,杜超,等. 林麝全基因组微卫星分布规律研究[J]. 四川动物,2017,36(4):420-424.
- [28] 张雪莲,王红梅,王磊,等. 草地贪夜蛾基因组微卫星的分布规律[J]. 应用昆虫学报,2020,57(6):1287-1298.
- [29] SCHUG M D, WETTERSTRAND K A, GAUDETTE M S, et al. The distribution and frequency of microsatellite loci in *Drosophila melanogaster*[J]. Molecular Ecology, 1998, 7(1):57-70.
- [30] 余泉友,李斌,李关荣,等. 蚊子全基因组中微卫星的丰度及其分布(英文)[J]. 生物化学与生物物理进展,2005,32(5):435-441.
- [31] WEBSTER M T, SMITH N G C, ELLEGREN H. Microsatellite evolution inferred from human-chimpanzee genomic sequence alignments[J]. Proceedings of the National Academy of Sciences of the United States of America, 2002, 99(13):8748-8753.
- [32] KATTI M V, RANJEKAR P K, GUPTA V S. Differential distribution of simple sequence repeats in eukaryotic genome sequences[J]. Molecular Biology and Evolution, 2001, 18(7):1161-1167.
- [33] 汪自立,黄杰,杜联明,等. 二斑叶螨和肩突硬蜱基因组微卫星分布规律研究[J]. 四川动物,2013,32(4):481-486.
- [34] 王小婷,张玉娟,何秀,等. 中华按蚊全基因组微卫星的鉴定、特征及分布规律[J]. 昆虫学报,2016,59(10):1058-1068.
- [35] 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[D]. 成都:四川农业大学,2004.
- [36] SCHLÖTTERER C, TAUTZ D. Slippage synthesis of simple sequence DNA[J]. Nucleic Acids Research, 1992, 20(2):211-215.
- [37] PEARSON C E, SINDEN R R. Trinucleotide repeat DNA structures: dynamic mutations from dynamic DNA[J]. Current Opinion in Structural Biology, 1998, 8(3):321-330.
- [38] CUMMINGS C J, ZOGHBI H Y. Fourteen and counting: unraveling trinucleotide repeat diseases[J]. Human Molecular Genetics, 2000, 9(6):909-916.
- [39] ARCHIE E A, MOSS C J, ALBERTS S C. Characterization of tetranucleotide microsatellite loci in the African Savannah Elephant (*Loxodonta africana africana*)[J]. Molecular Ecology Notes, 2003, 3(2):244-246.
- [40] 吴旭东,连总强,侯玉霞,等. 大口鲇微卫星标记在三个鲇形目鱼类种群间适用性研究[J]. 水生生物学报,2011,35(4):638-645.
- [41] 魏朝明,孔光耀,廉振民,等. 蜜蜂全基因组中微卫星的丰度及其分布[J]. 昆虫知识,2007,44(4):501-504.
- [42] 王耀嵘,杨尉,任席林,等. 金钱鱼基因组微卫星分布特征分析及多态性标记开发[J]. 广东海洋大学学报,2020,40(4):7-14.
- [43] SCHORDERET D F, GARTLER S M. Analysis of CpG suppression in methylated and nonmethylated species[J]. Proceedings of the National Academy of Sciences of the United States of America, 1992, 89(3):957-961.
- [44] 高焕,刘萍,孟宪红,等. 中国对虾(*Fenneropenaeus chinensis*)基因组微卫星特征分析[J]. 海洋与湖沼,2004,35(5):424-431.
- [45] WIERDL M, DOMINSKA M, PETES T D. Microsatellite instability in yeast: dependence on the length of the microsatellite[J]. Genetics, 1997, 146(3):769-779.
- [46] HARR B, SCHLÖTTERGER C. Long microsatellite alleles in drosophila melanogaster have a downward mutation bias and short persistence times, which cause their genome-wide under representation[J]. Genetics, 2000, 155(3):1213-1220.
- [47] ELLEGREN H. Heterogeneous mutation processes in human microsatellite DNA sequences[J]. Nature Genetics, 2000, 24(4):400-402.

[责任编辑:严海琳]