

文章编号: 1674-5566(2023)06-1176-08

DOI: 10.12024/jsou.20221004005

3种鲃亚科鱼类全基因组微卫星分布特征分析

刘士力^{1,2,3}, 刘一诺³, 李飞³, 郑建波³, 程顺³, 蒋文枰³, 迟美丽³,
赵金良^{1,2}

(1. 上海海洋大学 农业农村部淡水水产种质资源重点实验室, 上海 203106; 2. 上海海洋大学 水产种质资源发掘与利用教育部重点实验室, 上海 203106; 3. 浙江省淡水水产研究所 浙江省淡水水产遗传育种重点实验室, 浙江 湖州 313001)

摘要: 利用MISA软件对3种鲃亚科鱼类全基因组中的完整型微卫星进行了搜索, 并对其分布特征及规律进行了分析。结果表明: 在翘嘴鲃(*Culter alburnus*, 1.055 Gb)、鳊浪白鱼(*Anabarilius grahami*, 1.006 Gb)和团头鲂(*Megalobrama amblycephala*, 1.116 Gb)全基因组中, 分别筛选出772 276、548 726和772 329个完整型微卫星。相对丰度分别为732.0、553.2和691.2个/Mb。微卫星总长度分别为14.29、9.42和15.02 Mb, 分别占基因组序列总长度的1.35%、0.94%和1.35%。在1~6个不同碱基重复类型完整型微卫星中, 3种鲃亚科鱼类的6种碱基类型数目排序是一致的。均是单碱基重复数目最多, 然后依次是二碱基、四碱基、三碱基、五碱基和六碱基。3种鱼三碱基和四碱基重复中数量最多的分别是AAT和AGAT, 但在翘嘴鲃二碱基重复中, AT重复微卫星的数量要多于AC重复。3种鱼中不同碱基类型微卫星分布频率随着重复单元拷贝数和长度的增加呈递减趋势, 其微卫星重复单元长度主要分布在10~30 bp, 分别占微卫星总数的90.54%、94.27%和89.17%。本研究结果可为进一步研究鲃亚科基因组特征提供参考, 并为今后进行这3种鱼的分子标记辅助育种及遗传信息评估等工作提供基础资料。

关键词: 鲃亚科; 微卫星; 基因组; 分布特征

中图分类号: S 917 **文献标志码:** A

微卫星(Microsatellite)在真核^[1]、原核^[2]和病毒^[3]基因组中均有存在, 由于其遗传呈共显性、分布广、数量多、多态性高且易于检测而被广泛用于群体遗传结构分析^[4]、遗传图谱的构建^[5]、亲缘关系鉴定^[6]以及种质鉴定^[7]等领域。目前开发微卫星的主要方法有磁珠富集法、数据库筛选法、近缘物种扩增法和高通量测序法。相比其他微卫星开发方法, 高通量测序法更加便利和高效。在鲤(*Cyprinus carpio*)^[8]、草鱼(*Ctenopharyngodon idella*)^[9]、瓦氏黄颡鱼(*Pelteobagrus vachelli*)^[10]和绒杜父鱼(*Hemitripterus villosus*)^[11]等鱼类上已经基于全基因组序列开展了微卫星序列的特征分析和标记开发的研究。

鲃亚科为东亚特有类群, 共计18属64种, 绝大部分属、种分布于中国^[12]。其中部分种类在我

国淡水养殖业及淡水捕捞业中占有重要的经济地位。其中翘嘴鲃(*Culter alburnus*)是鲃亚科中体型最大的一种鱼类, 具有重要的生态意义和经济价值; 团头鲂(*Megalobrama amblycephala*)又称武昌鱼, 是我国长江中下游湖泊主养的草食性经济鱼类之一, 2021年鳊鲂养殖产量达764 303 t^[13]。鳊浪白鱼(*Anabarilius grahami*)是云南抚仙湖特有土著鱼类, 又称为抗浪鱼。该鱼肉鲜味美, 是云南四大名鱼之一。近年来, 随着人工养殖技术日趋成熟, 鳊浪白鱼已成为当地名优养殖鱼类之一^[14]。

目前关于鱼类基因组微卫星方面的研究已有一些, 梁霞等^[8]对鲤全基因组微卫星的研究表明其微卫星随着重复单元的增加, 占比逐渐降低。黄纬杰等^[9]对草鱼全基因的研究表明其四碱基重复

收稿日期: 2022-10-26 修回日期: 2022-11-20

基金项目: 浙江省农业(水产)新品种选育重大科技专项(2016C02055-1, 2021C02069-3)

作者简介: 刘士力(1985—), 男, 博士研究生, 研究方向为遗传育种学。E-mail: liushili1212@126.com

通信作者: 赵金良, E-mail: jlzhao@shou.edu.cn

版权所有 ©《上海海洋大学学报》编辑部(CC BY-NC-ND 4.0)

Copyright © Editorial Office of Journal of Shanghai Ocean University (CC BY-NC-ND 4.0)

<http://www.shhydx.com>

类型要高于三碱基重复序列。近年来,已有3种鲃亚科鱼类的基因组已被公布出来,这为从基因组层面研究其微卫星提供了良好的素材。本研究从基因组水平上研究了3种鲃亚科鱼类微卫星的分布特征、微卫星分布频率以及微卫星的丰度,为进一步开发微卫星的标记提供数据支持。

1 材料与方法

1.1 基因组序列

用于比较分析的翘嘴鲃(*Culter alburnus*)基因组来自本课题组,基因组大小为1.055 Gb,原始数据已上传至NCBI Genome数据库,登录号PRJNA700093。鳊浪白鱼(*Anabarilius grahami*)基因组从NCBI Genome数据库下载,登录号为RJUV00000000,基因组大小为1.006 Gb^[15];团头鲂(*Megalobrama amblycephala*)基因组下载自GigaDB数据库(<http://dx.doi.org/10.5524/100305>),其基因组总大小约为1.116 Gb^[16]。所有序列均以FASTA格式保存。

1.2 微卫星序列的筛选与统计

利用微卫星筛选软件MISA (<http://pgrc.ipk-gatersleben.de/misa/>)对全基因组序列进行扫描,筛选符合条件的微卫星序列。本研究仅对1~6碱基重复完整型微卫星进行统计分析。搜索标准

参考MISA软件默认参数,即单核苷酸(Mono-)、二(Di-)、三(Tri-)、四(Tetr-)、五(Penta-)、六核苷酸(Hexa-nucleotides)重复序列的最小重复次数分别为软件默认的10、6、5、5、5和5次。考虑到碱基互补配对原则以及每个微卫星起始碱基的排列差异,将所有可循环的序列及其碱基互补序列归为同一类别。如AC重复拷贝类别,可以与之合并为一类的重复拷贝类别是CA、TG和GT。MISA分析后采用Excel对实验结果进行初步汇总和分析。

2 结果与分析

2.1 基因组微卫星数量分析

在翘嘴鲃、鳊浪白鱼和团头鲂全基因组中,分别筛选出772 276个、548 726个和772 329个完整型微卫星。微卫星总长度分别为14.29 Mb、9.42 Mb和15.02 Mb,分别占基因组序列总长度的1.35%、0.94%和1.35%。在1~6个不同碱基重复类型完整型微卫星中,3种鲃亚科鱼类的6种碱基类型数目排序是一致的。除了四核苷酸微卫星略多于三核苷酸微卫星外,随着核苷酸重复次数的增多,微卫星的比例逐渐降低。均是单碱基重复微卫星数量最多,然后依次是二碱基、四碱基、三碱基、五碱基和六碱基(表1)。

表1 不同重复类型的微卫星在3种鲃亚科鱼类基因组的分布情况

Tab. 1 Distribution of microsatellite for different repeat types in the three species of Culterinae fish genome

重复类型 Motif length	翘嘴鲃 <i>Culter alburnus</i>			鳊浪白鱼 <i>Anabarilius grahami</i>			团头鲂 <i>Megalobrama amblycephala</i>		
	总数量 Total number	丰度 Frequency/ Mb	占比 Proportion /%	总数量 Total number	丰度 Frequency/ Mb	占比 Proportion /%	总数量 Total number	丰度 Frequency/ Mb	占比 Proportion /%
单碱基 Mononucleotide	416 303	394.60	53.90	267 027	269.21	48.66	410 948	367.76	53.21
二碱基 Dinucleotide	241 289	228.70	31.24	188 686	190.23	34.39	230 169	205.98	29.80
三碱基 Trinucleotide	51 244	48.60	6.64	38 969	39.29	7.10	58 579	52.42	7.58
四碱基 Tetranucleotide	52 099	49.40	6.75	47 800	48.19	8.71	59 609	53.34	7.72
五碱基 Pentanucleotide	9 741	9.20	1.26	5 635	5.68	1.03	12 581	11.26	1.63
六碱基 Hexanucleotide	1 600	1.50	0.21	609	0.61	0.11	443	0.40	0.06
总计 Total	772 276	732.00	100.00	548 726	553.20	100.00	772 329	691.20	100.00

2.2 基因组微卫星中优势重复单元的组成分析

虽然3种鲃亚科鱼类基因组中的6种碱基类型微卫星数目排序是一致的,但各碱基重复类型中优势碱基类别有所不同。

单核苷酸重复类型中,A/T占有绝对优势,是主要的重复单元,A/T重复微卫星在翘嘴鲃、鳊浪

白鱼和团头鲂全基因组中的比例分别为96.59%、97.87%和97.63%(表2)。

翘嘴鲃二核苷酸重复类型中AT/TA重复单元最常见,占比43.97%,略高于AC/GT重复单元(41.29%),之后是AG/CT,共35 193个(14.59%),GC/CG数量最少,375个。鳊浪白鱼和团头鲂全

基因组中占比最高的二核苷酸重复类型是 AC/GT, 其比例分别为 49.59% 和 47.08%, AT/TA 在这 2 种鱼中的比例分别为 32.79% 和 36.48%。

三核苷酸重复类型中, AAT/ATT 比例最高, 分别为 58.39%、56.46% 和 59.39%, 其次是 AAC/GTT, 占比分别为 12.16%、11.80% 和 12.27%。翘嘴鲌和鳊浪白鱼中占比第三的是 AAG, 比例分别为 7.99% 和 9.11%, 而团头鲂中占比第三的为 ACAG(7.15%)。

四核苷酸重复类型中 AGAT/ATCT 比例最高, 分别为 34.96%、36.79% 和 27.49%, 其次是 ATCC, 占比分别为 16.22%、18.63% 和 19.48%。

翘嘴鲌和团头鲂中占比第三的是 AAAT, 比例分别为 14.63% 和 17.72%, 而鳊浪白鱼中占比第三的为 ACAG(12.37%)。

五核苷酸重复类型中, 翘嘴鲌和团头鲂前 3 种优质碱基类型分别为 AATAT、AAAAT 和 AATAG。鳊浪白鱼中 AAGTG 数量要高于 AATAT 和 AAAAT; 六核苷酸重复类型中, 翘嘴鲌和团头鲂的 AACCT/AGGGT 占比最高, 在翘嘴鲌中占具明显优势, 数量为 955 个(59.69%)。鳊浪白鱼六核苷酸重复类型中数量最多的为 AATCAG, 数量为 63 个(10.34%)。

表 2 3 种鲌亚科鱼类基因组中各碱基重复类型中前 3 种优势碱基类别

Tab. 2 The top three dominant base classes in each base repeat type in the three species of Culterinae fish genome

重复类型 Motif length	翘嘴鲌 <i>Culter alburnus</i>	鳊浪白鱼 <i>Anabarilius grahmi</i>	团头鲂 <i>Megalobrama amblycephala</i>
单碱基 Mononucleotide	A(402 124) 96.59%	A (261 346) 97.87%	A (401 188) 97.63%
	C(14 179) 3.41%	C (5 681) 2.13%	C (9 760) 2.38%
	AT (106 084) 43.97%	AC (93 564) 49.59%	AC (108 370) 47.08%
二碱基 Dinucleotide	AC (99 637) 41.29%	AT (61 862) 32.79%	AT (83 976) 36.48%
	AG (35 193) 14.59%	AG (32 947) 17.46%	AG (37 417) 16.26%
	AAT (29 923) 58.39%	AAT (2 2001) 56.46%	AAT (34 790) 59.39%
三碱基 Trinucleotide	AAC (6 229) 12.16%	AAC (4 599) 11.80%	AAC (7 185) 12.27%
	AAG (4 096) 7.99%	AAG (3 553) 9.11%	ATC (4 186) 7.15%
	AGAT (18 212) 34.96%	AGAT (17 585) 36.79%	AGAT (16 385) 27.49%
四碱基 Tetranucleotide	ATCC (8 451) 16.22%	ATCC (8 903) 18.63%	ATCC (11 611) 19.48%
	AAAT (7620) 14.63%	ACAG (5 914) 12.37%	AAAT (10 561) 17.72%
	AATAT (3 462) 35.54%	AAGTG (1158) 20.55%	AATAT (4411) 35.06%
五碱基 Pentanucleotide	AAAAT (2 230) 22.89%	AATAT (837) 14.85%	AAAAT (3023) 24.03%
	AATAG (675) 6.93%	AAAAT (639) 11.34%	AATAG (835) 6.64%
	AACCCT (955) 59.69%	AATCAG (63) 10.34%	AACCCT (64) 14.45%
六碱基 Hexanucleotide	AACCCG (426) 26.63%	AAAGTG (62) 10.18%	ACATAT (35) 7.90%
	AGATAT (16) 1.00%	AAATGT (55) 9.03%	AGATAT (26) 5.87%

2.3 基因组微卫星重复单元数目和长度分析

3 种鲌亚科鱼类基因组中不同类型微卫星核心序列拷贝数变化趋势较为一致, 即随着微卫星

位点核心序列拷贝数的递增, 微卫星位点总体呈递减趋势。在翘嘴鲌、鳊浪白鱼和团头鲂基因组单碱基中重复拷贝数均主要集中在 10~24 次, 分

别占单碱基总数的99.47%、99.00%和94.53%;二碱基中重复拷贝数均主要集中在6~20次,分别占比93.22%、98.49%和93.66%;三碱基中核心重复拷贝数范围较为集中,主要集中在5~14次,分别占比99.63%、99.26%和98.26%;四碱基中翘嘴鲃和团头鲂微卫星核心重复拷贝数主要集中在5~19次,分别占比98.75%和96.46%。而鳊浪白鱼为90.98%,这是由于其有8.08%在20~24次;五碱基中3种鲃亚科鱼类微卫星核心重复拷贝数主要

集中在5~19次,分别占比98.68%、98.19%和93.91%。翘嘴鲃和鳊浪白鱼核心重复拷贝数在24次以上的五碱基微卫星比例分别为0.22%和0.53%,但团头鲂中高达5.20%;鳊浪白鱼和团头鲂在六碱基中的重复拷贝数范围分布较窄,核心重复拷贝数均集中在5~9次,分别占比93.59%和93.68%。翘嘴鲃在六碱基中要分散一些,有32.56%的微卫星拷贝数在9次以上(图1),其中AACCT最大重复次数为75次。

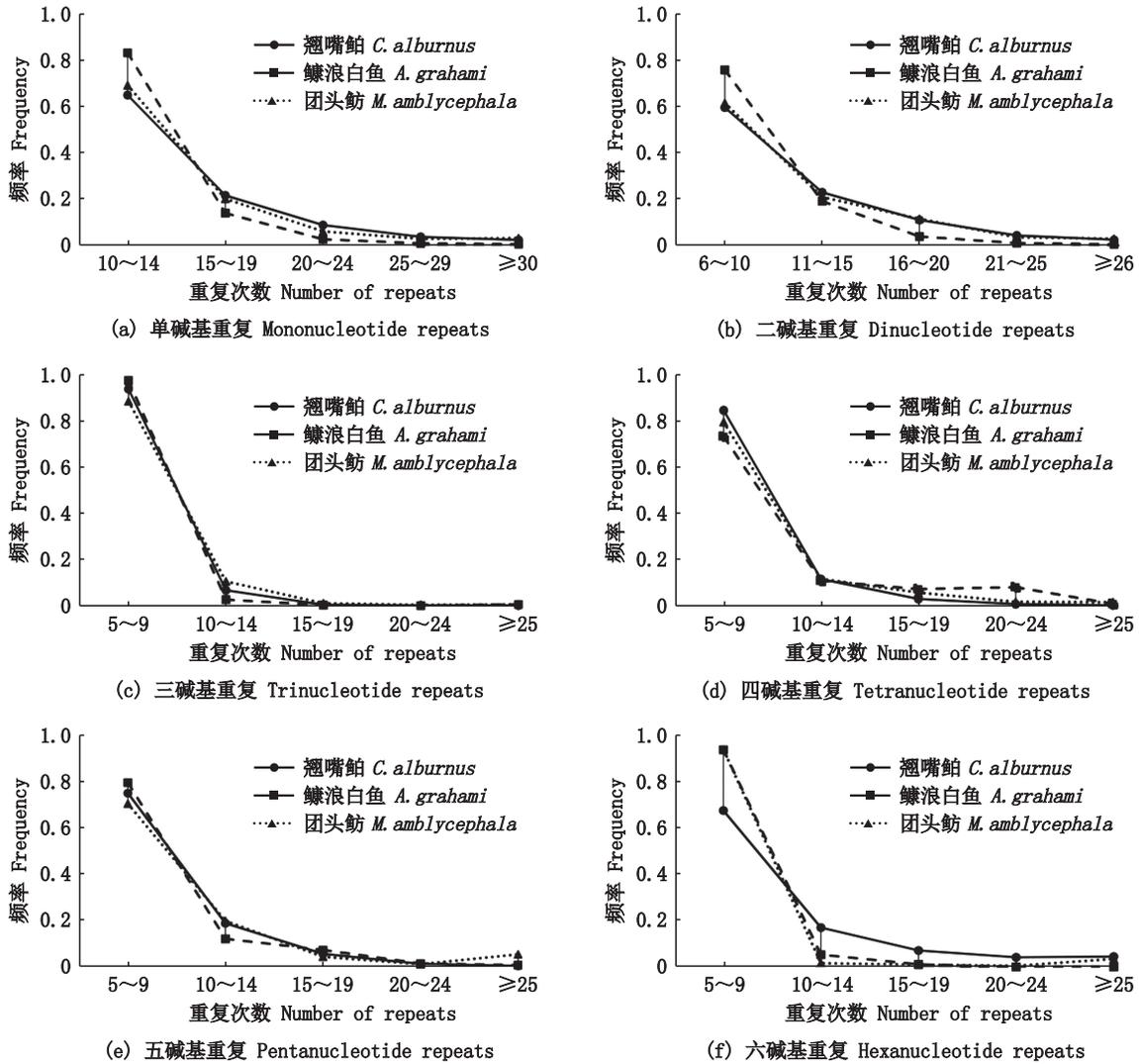


图1 3种鲃亚科鱼类6种重复类型微卫星的重复次数分布

Fig. 1 Distribution of microsatellite in six repeat types based on their copy numbers in three Culterinae species

翘嘴鲃、鳊浪白鱼和团头鲂基因组微卫星重复单元长度的分布见图2,其微卫星序列长度主要分布在10~30 bp,占比分别为90.54%、94.27%和89.17%。总体上,微卫星重复单元的长度分布表现出随长度增加频率急剧降低的趋势。

3 讨论

3.1 微卫星在全基因组中的含量

本研究以3种鲃亚科鱼类全基因组数据为基础,采用MISA软件对其中1~6碱基重复微卫星进行了统计分析。翘嘴鲃、鳊浪白鱼和团头鲂基因

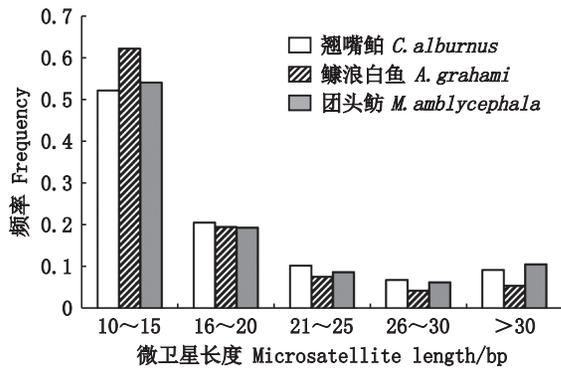


图2 3种鲃亚科鱼类不同微卫星长度的出现频率
Fig. 2 Frequency of microsatellites with different lengths in the three species of Culterinae fish

组中微卫星序列分别占基因组序列总长度的1.35%、0.94%和1.35%。与草鱼基因组微卫星含量(1.43%)^[9]、斑点叉尾鲷(*Ictalurus punctatus*, 1.45%)^[17]、瓦氏黄颡鱼(1.48%)^[10]和黄颡鱼(*Pelteobagrus fulvidraco*, 1.82%)^[18]等相比,这3种鲃亚科鱼类微卫星含量在全基因组中所占比例相对较低。但与鲤鱼(0.91%)、花斑无须鲢(*Ageneiosus marmoratus*, 0.91%)^[19]、红鳍东方鲀(*Takifugu rubripes*, 0.73%)、双斑东方鲀(*Takifugu bimaculatus*, 0.84%)和菊黄东方鲀(*Takifugu flavidus*, 0.84%)^[20]相比,这3种鲃亚科鱼类翘嘴鲌微卫星含量所占比例较高。巨鲃(*Bagarius yarrelli*, 1.23%)^[21]和黑青斑河鲀(*Tetraodon nigroviridis*, 1.06%)微卫星含量高于鳊浪白鱼,但低于翘嘴鲌和团头鲂。可见在同科甚至同属鱼类之间全基因组微卫星的含量也可能存在较大差异,这可能与物种差异、基因组大小以及组装质量有关。

3.2 微卫星各碱基类型特征

3种鲃亚科鱼类微卫星中单碱基的比例要高于其他碱基类型,与草鱼、鲤^[8]、巨鲃^[21]、斑点叉尾鲷^[17]、红鳍东方鲀、双斑东方鲀、菊黄东方鲀、黑青斑河鲀^[20]和美丽硬仆骨舌鱼(*Scleropages formosus*)^[22]等水生动物的优势碱基类型相同。而在黄颡鱼^[18]、瓦氏黄颡鱼^[10]、花斑无须鲢^[19]、金钱鱼(*Scatophagus argus*)^[23]和大菱鲆^[24]等鱼类中为二碱基占优势。尚未见三碱基为优势比例的鱼类,在原核生物和酵母的基因组中,三碱基重复序列类型处于优势地位,而比它们更高等的生物基因组中微卫星则倾向于两碱基和单碱基重复序列类型。巨鲃^[21]与同为鲃形目的黄颡鱼^[18]

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紧密联系,这是因为不同物种基因组的优势类型微卫星不同而导致的。在3种鲃亚科鱼类中,四碱基重复单元含量均高于三碱基重复单元,这在草鱼^[9]、黄颡鱼^[18]、瓦氏黄颡鱼^[10]和日本河鲀(*Fugu rubripes*)中也有发现^[25]。因此,我们推测物种基因组四碱基和三碱基重复单元相对丰度可能因种而异,并没有明显的规律性。

3.3 微卫星各碱基类型特征

在3种鲃亚科鱼类1~6碱基重复类型微卫星中,优势重复类型并不完全相同,但都具有明显的A/T优势。在单碱基重复中,A类别(A/T)数目最多,这与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,翘嘴鲌中AT类别微卫星数量最多,其次是AC类别,这与鳊浪白鱼和团头鲂等大多数硬骨鱼纲的二碱基优势重复类型均为AC类别不同,大部分鱼类二碱基类型中都是AC占优势。翘嘴鲌中CG类型微卫星最少,这与大多数生物中的情况一致^[26]。关于CG类别比AT类别少的原因,有研究认为,微卫星是在DNA复制的过程中产生滑移形成的^[27],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重复拷贝类别在数量上要多于CG重复拷贝类别。CG微卫星类型数量的稀少还被认为与胞嘧啶C在基因组中易被甲基化,容易脱氨基而转变成胸腺嘧啶T有关^[28]。也可以在一定程度上解释基因组中AC重复较多的现象。

AAT、AAC、AAG和ATG分别为三碱基中的前四类别,这表明AAN(N代表除A以外的任何碱基)类别在三碱基中具有绝对的碱基优势。该现象与鲤、斑点叉尾鲷^[17]和黄颡鱼^[18]等三碱基分布相一致。有研究表明三碱基微卫星不同于其它碱基类型,其与人类的遗传疾病具有一定的相关性^[29]。推测随着微卫星核心序列拷贝数的随机波动,有可能会改变功能基因的结构,从而影响到其相关性状的表达,引起部分遗传疾病的发生。因此对三碱基微卫星进行定位和研究分析,将有利于对其相关遗传疾病做出早期的预防与判断。在四碱基重复中,AGAT和ATCC类别的数量排在前二。微卫星(AGAT)_n在真核生物中广泛分布,可以用作结合转录因子家族GATA结合蛋白

的DNA共有序列^[30]。并认为与许多脊椎动物性染色体的分子分化有关^[31-32]。通过对微卫星作用的研究表明,在遗传标记上四、五和六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。本研究为开发这3种鱼类的四碱基微卫星标记提供了较大帮助。

3.4 微卫星各碱基类型拷贝数特征

在本实验的3种鲇亚科鱼类全基因组微卫星中,6种碱基类型总体变化特征就是随着重复拷贝数的增加,各类别微卫星的数目逐渐减少,这可能与微卫星位点长度的不断增加使得稳定性降低有关。WIERDL等^[33]在对微卫星序列的突变率进行研究时发现,对于GT序列,长度为105 bp时的突变率是长度为15 bp时突变率的500倍,据此得出结论认为微卫星的稳定性会随着长度的增加而降低,微卫星长度越长越不稳定。LEOPOLDINO等^[34]认为微卫星的突变率随着拷贝数的增加而增大,因此微卫星核心拷贝数越多微卫星发生突变的概率就越大,该微卫星数量就会减少。ELLEGREN^[35]研究认为微卫星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,一般情况下,核心序列拷贝数较多的微卫星数目较少。这些研究为更好地研究鲇亚科鱼类基因组特征及微卫星引物的筛选等提供了方向和依据。

参考文献:

- [1] TÓTH G, GÁSPÁRI Z, JURKA J, et al. Microsatellites in different eukaryotic genomes: survey and analysis [J]. *Genome Research*, 2000, 10(7): 967-981.
- [2] MRÁZEK J, GUO X, SHAH A. Simple sequence repeats in prokaryotic genomes [J]. *Proceedings of the National Academy of Sciences of the United States of America*, 2007, 104(20): 8472-8477.
- [3] ZHAO X Y, TAN Z Y, FENG H P, et al. Microsatellites in different Potyvirus genomes: survey and analysis [J]. *Gene*, 2011, 488(1/2): 52-56.
- [4] 崔文涛, 邹宇凡, 白志毅, 等. 安徽地区克氏原螯虾群体的遗传多样性和遗传结构分析[J]. *水产学报*, 2023, 47(9): 099607.
CUI W T, ZOU Y F, BAI Z Y, et al. Genetic diversity and structure analysis of *Procambarus clarkii* stocks in Anhui Province [J]. *Journal of Fisheries of China*, 2023, 47(9): 099607.
- [5] 郭香, 李琪, 孔令锋, 等. 基于微卫星标记整合长牡蛎遗传图谱[J]. *水产学报*, 2013, 37(6): 823-829.
- [6] GUO X, LI Q, KONG L, et al. Microsatellite-based consensus map of the Pacific oyster (*Crassostrea gigas*) [J]. *Journal of Fisheries of China*, 2013, 37(6): 823-829.
- [6] Bian Y, Liu S, Liu Y, et al. Development of a multiplex PCR assay for parentage assignment of the redclaw crayfish (*Cherax quadricarinatus*) [J]. *Aquaculture*, 2022, 550: 737813.
- [7] 费勤娇, 丁洪昌, 张鹏, 等. 基于转录组信息的坛紫菜 EST-SSR 标记的开发及新品系 SW-81 的种质鉴定[J]. *上海海洋大学学报*, 2021, 30(3): 464-474.
FEI Q J, DING H C, ZHANG P, et al. Development of EST-SSR marker based on transcriptome and the research of germplasm identification for the new strain (SW-81) of *Pyropia haitanensis*. *Journal of Shanghai Ocean University*, 2021, 30(3): 464-474.
- [8] 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]. *南京师大学报(自然科学版)*, 2021, 44(3): 103-111.
LIANG X, WANG H Q, MA Y X, et al. Distribution characteristics of microsatellites in the whole genome of *Cyprinus carpio*, Linnaeus [J]. *Journal of Nanjing Normal University (Natural Science Edition)*, 2021, 44(3): 103-111.
- [9] 黄纬杰, 郭向召, 张子豪, 等. 草鱼全基因组微卫星特征分析与亲子鉴定[J]. *水产学报*, 2022, 46(2): 161-172.
HUANG W J, GUO X Z, ZHANG Z H, et al. Analysis of microsatellite in the entire grass carp (*Ctenopharyngodon idella*) genome and the application in parentage identification [J]. *Journal of Fisheries of China*, 2022, 46(2): 161-172.
- [10] 彭洽, 李杰, 王涛, 等. 瓦氏黄颡鱼全基因组微卫星的分布特征及其定位的初步研究[J]. *南方水产科学*, 2022, 18(1): 90-98.
PENG Y, LI J, WANG T, et al. Preliminary study on distribution characteristics and positioning of microsatellites in whole genome of *Pelteobagrus vachelli* [J]. *South China Fisheries Science*, 2022, 18(1): 90-98.
- [11] 赵蕊蕊, 徐胜勇. 绒杜父鱼全基因组 survey 分析及微卫星分布特征[J]. *中国水产科学*, 2022, 29(7): 994-1001.
ZHAO R R, XU S Y. Whole-genome analysis and microsatellite distribution characteristics of *Hemibarbus villosus* [J]. *Journal of Fishery Sciences of China*, 2022, 29(7): 994-1001.
- [12] 陈宜瑜. 中国动物志·硬骨鱼纲·鲤形目(中卷)[M]. 北京: 科学出版社, 1998: 112-207.
- [13] CHEN YY. *Fauna Sinica, Osteichthyes, Cypriniformes II* (in Chinese). Beijing: Science Press, 1998: 112-207.
- [13] 农业农村部渔业渔政管理局, 全国水产技术推广总站,

- 中国水产学会. 2022中国渔业统计年鉴[M]. 北京: 中国农业出版社, 2022: 25.
- Bureau of Fisheries - Ministry of Agriculture and Rural Affairs, National Fisheries Technology Extension Center, China Society of Fisheries. China fishery statistical yearbook-2022 [M]. Beijing: China Agriculture Press, 2022, 25.
- [14] 刘红艳, 熊飞, 董元火, 等. 鳊浪白鱼野生与养殖群体遗传多样性的ISSR分析[J]. 华中师范大学学报(自然科学版), 2012, 46(5): 591-595.
- LIU H Y, XIONG F, DONG Y H, et al. ISSR analysis on genetic diversity between the wild and cultured populations of *Anabarilius grahami* [J]. Journal of HuaZhong Normal University(Natural Sciences), 2012, 46(5): 591-595.
- [15] JIANG W, QIU Y, PAN X, et al. Genome assembly for a Yunnan-Guizhou Plateau “3E” fish, *Anabarilius grahami* (Regan), and its evolutionary and genetic applications [J]. Frontiers in Genetics, 2018, 9: 614.
- [16] LIU H, CHEN C, GAO Z, et al. The draft genome of blunt snout bream (*Megalobrama amblycephala*) reveals the development of *intermuscular* bone and adaptation to herbivorous diet[J]. *GigaScience*, 2017, 6(7): gix039.
- [17] 唐荣叶, 苏孟园, 杨汶珊, 等. 斑点叉尾鲷全基因组微卫星分布特征分析[J]. 渔业科学进展, 2022, 43(2): 89-97.
- TANG R Y, SU M Y, YANG W S, et al. Analysis of microsatellite distribution characteristics in the channel catfish (*Ictalurus punctatus*) genome [J]. Progress in Fishery Sciences, 2022, 43(2): 89-97.
- [18] 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.
- XU J J, ZHENG X, LI J, et al. Distribution characteristics of whole genome microsatellite of *Pelteobagrus fulvidraco* [J]. Genomics and Applied Biology, 2020, 39(12): 5488-5498.
- [19] 苏孟园, 杨汶珊, 唐荣叶, 等. 花斑无须鲢(*Ageneiosus marmoratus*)全基因组微卫星分布特征研究[J]. 南京师范大学学报(工程技术版), 2021, 21(2): 65-71.
- SU M Y, YANG W S, TANG R Y, et al. Microsatellite distribution in the whole genome of *Ageneiosus marmoratus* [J]. Journal of Nanjing Normal University (Engineering and Technology Edition), 2021, 21(2): 65-71.
- [20] 徐杰杰, 郑翔, 张鑫宇, 等. 4种河鲢全基因组微卫星分布特征分析研究[J]. 基因组学与应用生物学, 2021, 40(4): 1441-1451.
- XU J J, ZHENG X, ZHANG X Y, et al. Analysis of distribution characteristics of microsatellites in four genomes of puffer fish[J]. Genomics and Applied Biology, 2021, 40(4): 1441-1451.
- [21] 杨汶珊, 唐荣叶, 苏孟园, 等. 巨鲈(*Bagarius yarrelli*)全基因组微卫星分布特征分析[J]. 南京师范大学学报(工程技术版), 2021, 21(3): 62-68.
- YANG W S, TANG R Y, SU M Y, et al. Analysis of microsatellite distribution characteristics in the whole genome of *Bagarius yarrelli* [J]. Journal of Nanjing Normal University (Engineering and Technology Edition), 2021, 21(3): 62-68.
- [22] 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.
- DUAN Y N, LIU Y, HU Y C, et al. Distribution regularity of microsatellites in *Scleropages formosus* genome [J]. Chinese Agricultural Science Bulletin, 2019, 35(23): 152-158.
- [23] 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.
- WANG Y R, YANG W, REN X L, et al. Distribution patterns of microsatellites and development of polymorphic markers from *Scatophagus argus* genome [J]. Journal of Guangdong Ocean University, 2020, 40(4): 7-14.
- [24] 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.
- RUAN X H. Development, characterization and application of microsatellite markers in Turbot (*Scophthalmus maximus* L.) [D]. Qingdao: Ocean University of China, 2009.
- [25] EDWARDS Y J K, ELGAR G, CLARK M S, et al. The identification and characterization of microsatellites in the compact genome of the Japanese pufferfish, *Fugu rubripes*: perspectives in functional and comparative genomic analyses [J]. Journal of Molecular Biology, 1998, 278: 843-854.
- [26] STALLINGS R L. CpG suppression in vertebrate genomes does not account for the rarity of (CpG)_n microsatellite repeats[J]. Genomics, 1992, 13(3): 890-891.
- [27] SCHLÖTTERER C, TAUTZ D. Slippage synthesis of simple sequence DNA [J]. Nucleic Acids Research, 1992, 20(2): 211-215.
- [28] SCHOREERET F, GGARTLAR 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: 957-961.
- [29] 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.
- [30] PRIYADARSHINI P, MURTHY B S, NAGARAJU J, et al. A GATA-binding protein expressed predominantly in the pupal ovary of the silkworm, *Bombyx mori* [J]. Insect Biochemistry and Molecular Biology, 2003, 33: 185-195.

- [31] SUBRAMANIAN S, MISHRA R K, SINGH L. Genomewide analysis of Bkm sequences (GATA repeats): predominant association with sex chromosomes and potential role in higher order chromatin organization and function[J]. *Bioinformatics*, 2003, 19: 681-685.
- [32] ZIEMNICZAK K, TRALDI J B, NOGAROTO V, et al. In situ localization of (GATA)_n and (TTAGGG)_n repeated DNAs and W sex chromosome differentiation in Parodontidae (Actinopterygii: Characiformes) [J]. *Cytogenetic And Genome Research*, 2014, 144: 325-332.
- [33] 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.
- [34] LEOPOLDINO A M, PENA S D J. The mutational spectrum of human autosomal tetranucleotide microsatellites [J]. *Human Mutation*, 2003, 21(1): 71-79.
- [35] ELLEGREN H. Heterogeneous mutation processes in human microsatellite DNA sequences [J]. *Nature Genetics*, 2000, 24(4): 400-402.

Distribution characteristics of microsatellites in the whole genome of three Culterinae species

LIU Shili^{1,2,3}, LIU Yinuo³, LI Fei³, ZHENG Jianbo³, CHENG Shun³, JIANG Wenping³, CHI Meili³, ZHAO Jinliang^{1,2}

(1. Key Laboratory of Freshwater Aquatic Genetic Resources, Ministry of Agriculture and Rural Affairs, Shanghai Ocean University, Shanghai 201306, China; 2. Key Laboratory of Exploration and Utilization of Aquatic Genetic Resources, Ministry of Education, Shanghai Ocean University, Shanghai 201306, China; 3. Key Laboratory of Freshwater Aquatic Animal Genetic and Breeding of Zhejiang Province, Zhejiang Institute of Freshwater Fisheries, Huzhou 313001, Zhejiang, China)

Abstract: In this study, the microsatellite search software MISA was used to search for the perfect microsatellite in the whole genome of three Culterinae species, and its distribution pattern was analyzed. As a result, in the genome of (*Culter alburnus*, 1.055 Gb), (*Anabarilius grahami*, 1.006 Gb), and (*Megalobrama amblycephala*, 1.116 Gb), 772 276, 548 726, and 772 329 perfect microsatellites were screened out, respectively. The relative frequency were 732.0/Mb, 553.2/Mb, and 691.2/Mb, respectively. The total lengths of the microsatellites were 14.29 Mb, 9.42 Mb, and 15.02 Mb, respectively, accounting for 1.35%, 0.94%, and 1.35% of the total length of the genome sequence. Among the complete microsatellites with 1-6 different base repeat types, the orders of 6 base types from the three species of Culterinae were the same. Each had the largest number of mononucleotide repeats, followed by dinucleotide repeats, tetranucleotide repeats, trinucleotide repeats, pentanucleotide repeats, and hexanucleotide repeats. The highest frequency core repeat categories of trinucleotide and tetranucleotide in the three species of fish is AAT and AGAT, respectively. However, the number of AT is more than AC repeat in the dinucleotide of *Culter alburnus*. The distribution frequency of microsatellites of different base types in the three fish species decreased with the increase of the copy number and length of the repeat units. The length of the microsatellites repeat unit was mainly 10-30 bp, accounting for 90.54%, 94.27%, and 89.17% of the microsatellites, respectively. The results of this study may provide a reference for the further study of Culterinae genome characteristics and contribute basic data for future investigations into molecular marker-assisted breeding and genetic information assessment of 3 Culterinae species.

Key words: Culterinae; microsatellite; genome; distribution characteristics