

DOI: 10.19663/j.issn2095-9869.20210126002

http://www.yykxjz.cn/

唐荣叶, 苏孟园, 杨汶珊, 徐杰杰, 王涛, 尹绍武. 斑点叉尾鲷全基因组微卫星分布特征分析. 渔业科学进展, 2022, 43(2): 89-97

TANG R Y, SU M Y, YANG W S, XU J J, WANG T, YIN S W. Analysis of microsatellite distribution characteristics in the channel catfish (*Ictalurus punctatus*) genome. Progress in Fishery Sciences, 2022, 43(2): 89-97

斑点叉尾鲷全基因组微卫星分布特征分析*

唐荣叶 苏孟园 杨汶珊 徐杰杰 王涛^① 尹绍武^①

(南京师范大学海洋科学与工程学院 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 为了解斑点叉尾鲷(*Ictalurus punctatus*)全基因组中完整型微卫星的分布特征,本研究使用生物信息学软件 MISA 对其全基因组微卫星进行了搜索并分析。结果显示,在斑点叉尾鲷29 条染色体中,筛选到完整型微卫星共 510 256 个,总长度达 11 036 941 bp。其中,微卫星数量最多的是 2 号染色体(25 284 个),其次分别是 3 号、1 号和 5 号染色体,29 号染色体的微卫星数量最少,只有 11 591 个。每条染色体长度与位于其上的微卫星的数量均显著相关(SPSS, $r=0.98$, $P<0.01$)。27 号染色体的微卫星相对丰度最大(785.03 个/Mb),11 号染色体的微卫星相对丰度最小(615.89 个/Mb)。6 种重复拷贝类型的微卫星中,单碱基的数量最多,占总数的 45.31%,随后依次是二碱基(38.53%)、三碱基(8.73%)、四碱基(6.93%)、五碱基(0.46%)和六碱基(0.04%)。斑点叉尾鲷全基因组微卫星中,前 10 种优势重复拷贝类别为 A、AC、AG、AT、AAT、AAAT、C、AAC、AAAC 和 AAG,表现出明显的 A、T 碱基优势。本研究结果可为进一步研究斑点叉尾鲷全基因组特征提供参考,并为今后进行斑点叉尾鲷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 斑点叉尾鲷; 全基因组; 微卫星; 分布特征

中图分类号 S917.4 **文献标识码** A **文章编号** 2095-9869(2022)02-0089-09

微卫星又称简单序列重复(simple sequence repeats, SSRs),广泛分布于真核、原核生物以及病毒中,是由 1~6 bp 碱基为单元重复串联而成的 DNA 序列。根据核心序列的排列差异,又分为完整型和不完整型(Morgante *et al.*, 2001)。其中,完整型指重复序列中不存在其他碱基或重复序列的情况,不完整型指重复序列中存在错配情况。目前,对微卫星的研究主要集中于完整型(郑燕等, 2012),因其具有多态性高、

共显性遗传等特点,目前被广泛应用于遗传图谱构建(Shen *et al.*, 2007; Xia *et al.*, 2010)、遗传多样性分析和种质资源保护(Narasimhamoorthy *et al.*, 2008)等研究中。

斑点叉尾鲷(*Ictalurus punctatus*),又名沟鲶,是一种原产于北美洲的大型淡水鱼类。由湖北省水产科学院在 1984 年引入我国后进行驯养,于 1987 年成功解决人工繁育的难题,成为我国重要的淡水经济鱼类之一(向建国等, 2004)。另外,较强的繁殖能力使其

* 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[CX(19)2034]和南京师范大学大学生创新创业训练计划项目共同资助 [This work was supported by Agricultural Major New Variety Creation Project of Jiangsu Province (PZCZ201742), Key Research and Development Program of Jiangsu Province (BE2017377), Jiangsu Agricultural Science and Technology Innovation Fund [CX(19)2034], and Innovation and Entrepreneurship Training Program for College Students in Nanjing Normal University].

唐荣叶, E-mail: 2508049280@qq.com

① 通讯作者: 王涛, 副教授, E-mail: seawater88@126.com; 尹绍武, 教授, E-mail: yinshaowu@163.com

收稿日期: 2021-01-26, 收修改稿日期: 2021-02-23

能形成庞杂的品系和群体,成为基因组学研究有代表性的生物群体和实验材料(张涛等,2019)。

近年来,由于远离原产地多年、亲本数量有限等原因,无可避免地对国内的斑点叉尾鲷群体进行近亲繁殖育种,造成了一定程度的遗传漂变和种质退化等问题。因此,斑点叉尾鲷种质资源的保护越来越受到关注。目前,已有随机扩增长度多态性 DNA(RAPD)标记(Liu *et al.*, 1998、1999)、扩增片段长度多态性(AFLP)标记(Mickett *et al.*, 2003; Simmons *et al.*, 2006)和表达序列微卫星(EST-SSR)标记(赫崇波等,2008)等多种分子标记技术应用于斑点叉尾鲷遗传育种和种质资源方面的研究,而关于斑点叉尾鲷全基因组微卫星分布规律的研究尚未见报道。本研究对斑点叉尾鲷29条染色体上微卫星序列的数量、频率、密度等分布情况进行统计分析,旨在为进一步系统地进行斑点叉尾鲷全基因组微卫星引物的开发、遗传多样性和遗传育种等研究提供基础资料。

1 材料与方法

1.1 全基因组来源

斑点叉尾鲷全基因组数据从 NCBI Genome 数据库(<https://www.ncbi.nlm.nih.gov/genome/?term=Ictalurus+punctatus>)下载,登录号为 GCA_004006655.3,以 FASTA 格式文件保存,共包含 29 条染色体,总大小为 11 036 941 bp。

1.2 数据处理方法

利用微卫星搜索软件 MISA (<http://pgrc.ipk-gatersleben.de/misa/>)对斑点叉尾鲷全基因组中完整型微卫星序列进行搜索和统计。本研究统计标准设置为 MISA 默认参数,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(徐杰杰等,2020),即单碱基重复序列最小拷贝数设置为 10,二

碱基重复序列最小拷贝数设置为 6,三、四、五、六碱基重复序列最小拷贝数设置为 5。考虑碱基互补配对原则和碱基排列顺序的差异,对可聚为同一类型的微卫星进行兼并统计。

2 结果与分析

2.1 斑点叉尾鲷全基因组微卫星总体分布特征

从斑点叉尾鲷全基因组中共筛选出 510 256 个完整型微卫星,总长度为 11 036 941 bp,占全基因组 DNA 序列总长度的 1.45%。在斑点叉尾鲷 29 条染色体中,出现微卫星数量最多的是 2 号染色体,占微卫星总数的 4.96%,其次是 3 号染色体和 1 号染色体,分别占 4.72%和 4.64%。29 号染色体在所有染色体中所含微卫星数量最少,仅占 2.27%。线性回归分析表明,斑点叉尾鲷微卫星数量和其所在的染色体长度具有显著相关性(SPSS, $r=0.98$, $P<0.01$),微卫星频率与染色体长度也具有显著相关性(SPSS, $r=-0.67$, $P<0.01$) (图 1、图 2)。

2.2 斑点叉尾鲷 29 条染色体不同重复类型微卫星分布特征

通过统计分析可知,斑点叉尾鲷全基因组 1~6 碱基重复类型中,含量最高的是单碱基重复类型,其后依次是二碱基和三碱基。在斑点叉尾鲷各染色体碱基重复序列中,除 27 号和 29 号染色体二碱基数量大于单碱基外,其余 27 条染色体均是单碱基微卫星数量最多,其他依次是二碱基、三碱基、四碱基、五碱基和六碱基重复类型(表 1、图 3)。

2.3 斑点叉尾鲷全基因组微卫星各重复类别特征

在斑点叉尾鲷全基因组微卫星各重复类型中,同种碱基类型的不同类别分布有很大差异。在单碱基重复类型中,以 A 碱基重复类别最多, C 碱基重复类别

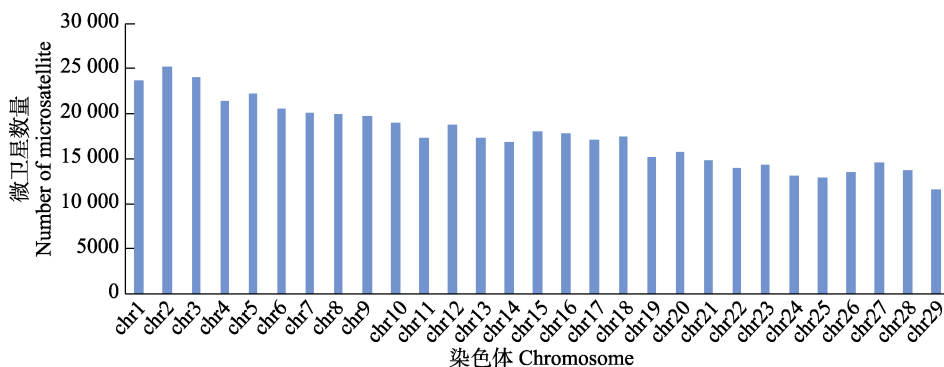


图 1 斑点叉尾鲷各染色体微卫星数量分布

Fig.1 Distribution of SSRs number in each chromosome of *I. punctatus*

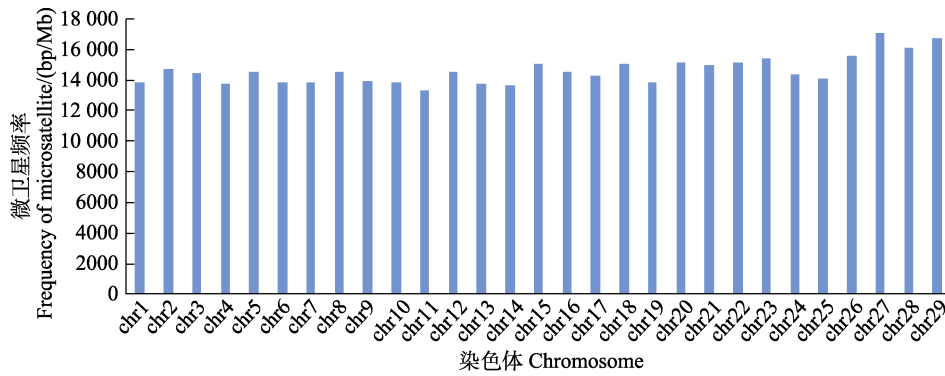


图 2 斑点叉尾鲷各染色体微卫星频率分布

Fig.2 Distribution of SSRs frequency in each chromosome of *I. punctatus*

表 1 斑点叉尾鲷各染色体上各种类型微卫星数量分布情况

Tab.1 Distribution of different types of SSRs in each chromosome of *I. punctatus*

染色体 Chromosome	单碱基 Mono-nucleotide	二碱基 Di-nucleotide	三碱基 Tri-nucleotide	四碱基 Tetra-nucleotide	五碱基 Penta-nucleotide	六碱基 Hexa-nucleotide	总数 Total number	占比 Proportion /%	丰度 Abundance /(ind./Mb)
chr1	10 706	9314	2026	1541	90	8	23 685	4.64	631.43
chr2	10 765	10 510	2210	1670	113	16	25 284	4.96	678.63
chr3	11 153	9133	1966	1708	106	8	24 074	4.72	671.46
chr4	9397	8655	1805	1477	94	15	21 443	4.20	619.81
chr5	9870	8857	1884	1503	87	14	22 215	4.35	666.09
chr6	9679	7564	1808	1440	94	7	20 592	4.04	640.71
chr7	9182	7664	1763	1344	78	11	20 042	3.93	645.94
chr8	9192	7383	1738	1506	134	9	19 962	3.91	656.43
chr9	9246	7323	1700	1400	95	5	19 769	3.87	654.32
chr10	9319	6579	1671	1309	81	4	18 963	3.72	652.86
chr11	8191	6362	1506	1216	84	14	17 373	3.40	615.89
chr12	8163	7613	1657	1241	94	8	18 776	3.68	667.10
chr13	8010	6474	1517	1187	89	2	17 279	3.39	630.03
chr14	7540	6573	1438	1167	72	9	16 799	3.29	629.18
chr15	7986	7062	1632	1301	81	8	18 070	3.54	692.71
chr16	8445	6545	1555	1214	67	9	17 835	3.50	686.91
chr17	7701	6527	1534	1209	85	7	17 063	3.34	658.37
chr18	8296	6242	1616	1272	75	6	17 507	3.43	714.20
chr19	6657	6175	1265	1031	70	5	15 203	2.98	630.12
chr20	6985	6069	1424	1140	85	9	15 712	3.08	699.23
chr21	6935	5508	1348	995	72	6	14 864	2.91	708.25
chr22	6194	5538	1237	918	65	3	13 955	2.73	702.17
chr23	6590	5309	1281	1049	71	5	14 305	2.80	725.33
chr24	6139	4763	1259	963	64	5	13 193	2.59	669.60
chr25	6063	4649	1209	956	65	4	12 946	2.54	665.34
chr26	5789	5457	1185	977	73	7	13 488	2.64	706.88
chr27	6132	6198	1174	981	69	4	14 558	2.85	785.03
chr28	6124	5438	1148	931	62	7	13 710	2.69	744.33
chr29	4675	5143	998	723	45	7	11 591	2.27	759.74

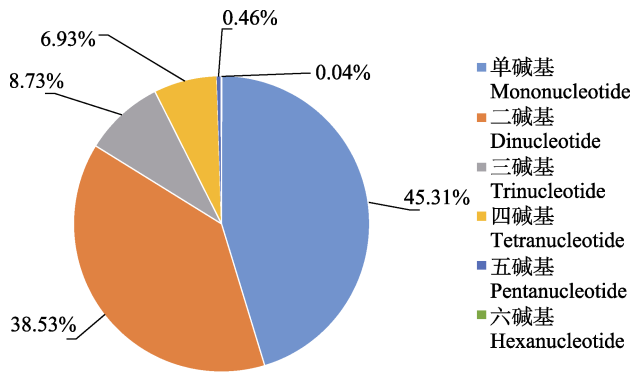


图3 斑点叉尾鲷全基因组微卫星各重复类型占比

Fig.3 Distribution of SSRs repeat types in genomes of *I. punctatus*

最少。二碱基重复类型的4种类别中,数量最多的是AC类别,其次是AG和AT,CG类别的数量最小,仅占二碱基类型的0.14%。三碱基类别中,数量最多的是AAT,其次是AAC和AAG。四碱基重复类型中,以AAAT和AAAC类别数量最多。五碱基中前3种类别分别是ATAAT(394个)、AAAAC(223个)和AAAAT(168个),分别占五碱基的16.69%、9.45%和7.12%。六碱基各类别数量相对于前5种重复类型数量较少且分布较分散,没有出现明显的优势类别。

2.4 斑点叉尾鲷微卫星各重复类型拷贝数分布特征

斑点叉尾鲷各重复类型微卫星的核心序列拷贝数范围差异很大,最小拷贝为5次,最大拷贝为127次,96.66%的微卫星核心序列拷贝数集中在5~30之间。单碱基重复拷贝数主要集中在10~30次,占单碱基所有微卫星数目的98.30%。二碱基重复拷贝数主要集中在5~30次,占二碱基总数目的93.47%。三碱基重

复拷贝数主要集中在5~20次,占三碱基总数目的98.28%。四碱基重复拷贝数主要集中在5~14次,占四碱基总数目的98.59%。五碱基重复拷贝数主要集中在5~12次,占五碱基总数目的96.53%。六碱基重复拷贝数主要集中在5~7次,占六碱基总数目的83.78%(图4)。

3 讨论

3.1 斑点叉尾鲷全基因组微卫星序列分布特征分析

微卫星标记是一种应用广泛的分子标记技术,斑点叉尾鲷是我国重要的水产养殖品种,将微卫星标记技术运用于斑点叉尾鲷的研究中,能很好地为斑点叉尾鲷分子标记辅助育种等方面的研究提供基础资料。本研究分析显示,在斑点叉尾鲷约762 Mb的全基因组序列中,微卫星含量为全基因组的1.45%。与人(*Homo sapiens*)(3%)(Subramanian *et al.*, 2003)、黄颡鱼(*Pelteobagrus fulvidraco*)(1.8%)(徐杰杰等, 2020)等物种相比,微卫星在斑点叉尾鲷全基因组中分布较少。而与猪(*Sus scrofa*)(0.85%)(戚文华等, 2014)、大熊猫(*Ailuropoda melanoleuca*)(0.64%)(李午佼等, 2014)、林麝(*Moschus berezovskii*)(0.42%)(卢婷等, 2017)等相比分布较多,其原因可能是不同物种的基因组大小有差异及基因组内的碱基组成和排列方式不同。

在本研究中,染色体DNA序列的长度与染色体上所含有的微卫星数量具有高度相关性(SPSS, $r=0.98, P<0.01$),这与黄杰等(2012)和戚文华等(2013)的研究结果一致。同时也支持了Hancock(1996)的假说:染色体序列越长,微卫星含量越高的趋势。统计

表2 斑点叉尾鲷全基因组微卫星数量最多的10种重复拷贝类别

Tab.2 Top 10 repeated copy categories of microsatellites in genomes of *I. punctatus*

重复类别 Motif	数量 Counts	长度 Length /bp	占比 Percentage /%	平均长度 Average length /bp	丰度 Abundance /(ind./Mb)	密度 Density /(bp/Mb)
A	215 753	2 958 299	93.35	13.71	283.14	3882.25
AC	123 484	3 065 578	62.80	24.83	162.05	4023.04
AG	39 884	1 025 944	20.28	25.72	52.34	1346.37
AT	32 978	1 490 536	16.77	45.20	43.28	1956.07
AAT	25 903	839 853	58.14	32.42	33.99	1102.16
AAAT	15 598	415 802	44.10	26.66	20.47	545.67
C	15 371	219 430	6.65	14.28	20.17	287.96
AAC	7320	151 989	16.43	20.76	9.61	199.46
AAAC	4908	123 278	13.88	25.12	6.44	161.78
AAG	3076	67 944	6.90	22.09	4.04	89.16

注:占比表示该类别在各自重复类型微卫星中的占比

Note: The percentage represents the proportion of different types of SSRs

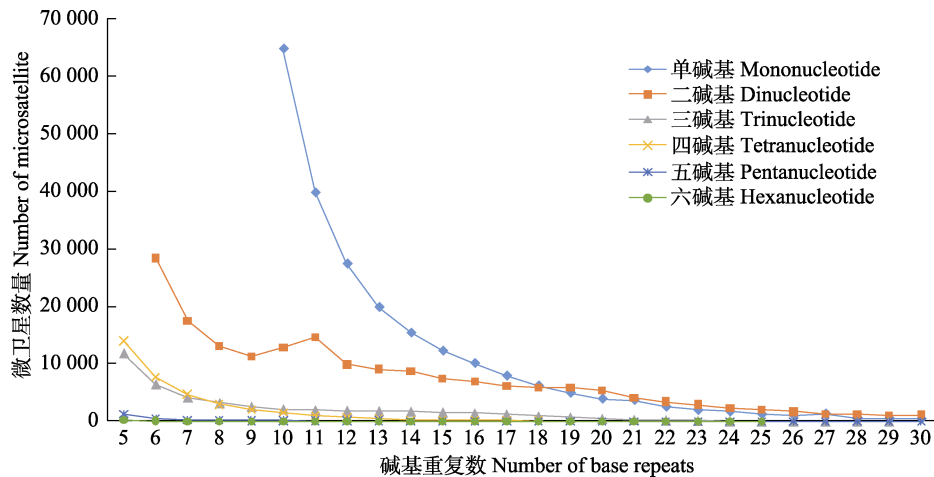


图 4 斑点叉尾鲷全基因组微卫星重复类型重复数分布

Fig.4 Distribution of number of base repetitions SSRs repeat types in *I. punctatus* genome

各染色体上的各种类型微卫星数量,发现各类型的比例除 27 号和 29 号染色体二碱基大于单碱基外,均符合 29 条染色体整体微卫星分布特征,即单碱基最多,其次是二碱基、三碱基、四碱基、五碱基和六碱基。这个规律与张琳琳等(2008)对赤拟谷盗(*Tribolium castaneum*)的研究结果相似,而与黄杰等(2012)对红原鸡(*Gallus gallus*)各染色体微卫星分布情况的研究结果相差甚远。高焕等(2005)认为,同一物种的不同染色体上,各种类型的重复序列分布有很大差异。而本研究的结果表明,各种类型的微卫星在各染色体上的分布符合整体规律,这可能和不同物种的不同类型微卫星的特定功能有关系。

斑点叉尾鲷全基因组微卫星中占比例最高的是单碱基类型,其次是二碱基类型,这与猪(戚文华等, 2014)、林麝(卢婷等, 2017)等哺乳动物以及菊黄东方鲀(*Takifugu flavidus*)(徐杰杰等, 2021)等水生动物的有关研究结果一致;而与一些其他物种的研究结果相异,如大鼠(*Rattus norvegicus*)(涂飞云等, 2015)、虾虎鱼科(Gobiidae)(白翠翠等, 2016)是以二碱基类型所占比例较高。酵母(*Saccharomyces cerevisiae*)(Katti *et al.*, 2001)、德国小蠊(*Blattella germanica*)(王晨等, 2015)等则以三碱基类型占比例最高。从大部分已公布的物种全基因组微卫星分布特征来看,大多数物种都是单碱基或者二碱基占优势,少部分低等动物则为三碱基或六碱基数目最多。黄杰等(2015)认为高等生物基因组更倾向于单碱基微卫星占优势。虽然以上推论不能概括出所有已公布物种的全基因组微卫星的分布特征,但在对比分析中,仍可以从整体上看到大致的规律。而对于不同物种的全基因组微卫星分布情况具有的差别。可能的原因有 2 种: 1)所用微卫星筛选软件

不同,参数设置不同; 2)可能与不同生物的基因表达及调控有关,体现了不同物种基因组的特征。

3.2 斑点叉尾鲷各碱基类型微卫星不同类别分布特征分析

不同物种的全基因组中重复序列不仅不同类型分布差异很大,同一类型不同类别也存在很大差异。斑点叉尾鲷全基因组微卫星中的前 5 种优势碱基类别分别为 A、AC、AG、AT 和 AAT,与同属于鲇形目的黄颡鱼(徐杰杰等, 2020)、花斑无须鲶(*Ageneiosus marmoratus*)(未发表)及巨鲰(*Bagarius yarrelli*)(未发表)的前 5 类优势碱基类别完全一致。由此可以猜测,物种亲缘关系越近,微卫星分布特征越接近。王月月等(2015)对 7 种家养动物全基因组微卫星分布的研究也得出了同样的结论。这可能说明微卫星序列是随着生物基因组的进化和发展而不断变化的,由于选择压力等缘故,使得亲缘关系相近的物种其微卫星序列分布规律也接近。

单碱基类型微卫星以 A 类别为主,这与已公布的绝大多数生物的全基因组微卫星单碱基类型特征相符合,如食蟹猴(*Macaca fascicularis*)(涂飞云等, 2018)、红尾蚺(*Boa constrictor*)(聂虎等, 2017)等。二碱基类型以 AC 所占比例最高,与大鼠(涂飞云等, 2015)、小鼠(*Mus musculus*)(童晓玲等, 2006)的二碱基类型的优势类别一致,而与德国小蠊(王晨等, 2015)、红原鸡(黄杰等, 2012)以 AT 碱基类别占优势的结果并不相同。Bachtrog 等(2000)曾用 6 个黑腹果蝇(*Drosophila melanogaster*)群体作为研究对象,得出二碱基重复类型中不同重复类别的微卫星突变率不同的结论,表明在黑腹果蝇基因组微卫星中 AC 类别相

对突变率比 AT 要大得多。而在斑点叉尾鲷基因组中,与黑腹果蝇不同,二碱基 AC 类别所占的比例最高。此现象可能由两个原因造成:1)在斑点叉尾鲷全基因组微卫星中,AC 类别重复序列的突变率明显低于其他类别,使得这个类别被较完整地保留下来;2) AC 类别重复序列容易由其他类别突变而成,且被不断地积累。猜想不同的物种在进化过程中总倾向于选择对自身发展有利的微卫星类别保留下来。对斑点叉尾鲷来说,可能微卫星 AC 重复类别对其进化和生长过程起着十分重要的作用。因此,无论是以上来源中的哪种途径,都使得 AC 类别在数量上占据优势地位。另外,CG 类别的含量最低,综合已有研究的物种的全基因组微卫星特征来看,二碱基 CG 类别重复序列的分布在绝大多数物种中都较低。

三碱基类型中,数目最多的是 AAT 类别,其次是 AAC 和 AAG,这与人的三碱基重复类型的前 3 种类别是一样的(Subramanian *et al*, 2003)。研究表明,某些三碱基重复类别可以形成特殊的结构,如发夹结构、三叶草结构和滑链结构等,这些结构非常的灵活,可能导致它们的遗传不稳定性(Pearson *et al*, 1996),此外,这些结构对基因的表达也起着十分重要的调控作用(戚文华等, 2016)。研究发现,人类的一些遗传疾病,如脆 X 综合征、肌强制性营养不良等与基因组中某些三碱基重复序列的增加有关(Schlötterer, 1998; Kovtun *et al*, 2001),而斑点叉尾鲷三碱基重复序列的前 3 种类别与人的一致,因此,斑点叉尾鲷的某些疾病可能也与三碱基重复序列相关(徐杰杰等, 2021)。

四碱基、五碱基的数量较多的前 2 种类别分别是 AAAT、AAAC 和 ATAAT、AAAAC,六碱基的数量相对于前 5 种重复类型比较少,数量最多的 2 种类别是 TGACTA 和 ATAGTC。与前 3 种碱基类型一样,四碱基、五碱基和六碱基也表现出明显的 A/T 碱基优势。有研究表明,DNA 序列中 G/C 含量越高,微卫星分布越少。对于这个现象,倪守胜等(2018)通过分析虾夷扇贝(*Patinopecten yessoensis*)基因组的微卫星分布特征,得出 DNA 复制滑动机制和重组机制使得 A/T 重复类型的产生几率更高的结论。Schorderet 等(1992)认为,由于基因组中 CpG 甲基化,胞嘧啶 C 容易在脱氨基作用下突变为胸腺嘧啶 T。这可能是二碱基中 CG 含量较少、TG (AC 类型)含量较多的原因。然而 Stallings (1992)的研究则表明,无论是否有 CpG 甲基化过程,基因组中的 CG 重复类别都是偏小的,他提出 CpG 结构是一种有害结构。

3.3 斑点叉尾鲷各碱基类型拷贝数特征分析

斑点叉尾鲷 6 种重复类型中,除二碱基在 10~11

拷贝数处有微小变动之外,基本上呈现随重复次数递增、微卫星数量逐渐递减的规律。这个规律与绝大部分物种的全基因组微卫星分布特征相符。Ellegren (2000)认为,在基因座上,微卫星序列的长度会维持在一定的范围内,因为长等位基因倾向于变短,从而抑制了微卫星长度的无限增长。另外,该现象可能和微卫星的突变率增大有关。有研究表明,同种类型的微卫星中拷贝次数多的要比次数少的具有更高的突变率,往往重复序列越长,突变的概率越大(Leopoldino *et al*, 2003)。因此,微卫星长度越长,其稳定性就越低,越容易发生碱基突变而成为不完整型微卫星或者非微卫星序列,使得重复数越多的完整型微卫星数量越少。长序列表现出更高的多态性,在进行遗传多样性研究时,可适当选择长度较大的微卫星位点作为遗传标记(王月月等, 2015)。

综上所述,本研究使用软件 MISA 对斑点叉尾鲷全基因组中完整型微卫星进行了搜索并分析,得出斑点叉尾鲷全基因组微卫星的分布特征如下:斑点叉尾鲷的染色体长度越长,所含微卫星的数量越多;各染色体上微卫星数量大致符合单碱基>二碱基>三碱基>四碱基>五碱基>六碱基的整体分布规律。各类型微卫星均表现出明显的 A/T 碱基优势。微卫星序列的拷贝数越大,其数量就越少。本研究为斑点叉尾鲷微卫星标记的开发、遗传多样性分析和分子标记育种等提供了数据支持。

参 考 文 献

- BACHTROG D, AGIS M, IMHOF M, *et al*. Microsatellite variability differs between dinucleotide repeat motifs: Evidence from *Drosophila melanogaster*. *Molecular Biology and Evolution*, 2000, 17(9): 1277-1285
- BAI C C, LIU S F, ZHANG Z M. Characteristic analysis of microsatellite DNA in the genome of Gobiidae. *Progress in Fishery Sciences*, 2016, 37(5): 9-15 [白翠翠, 柳淑芳, 庄志猛. 虾虎鱼科(Gobiidae)基因组微卫星 DNA 的分布特征. *渔业科学进展*, 2016, 37(5): 9-15]
- ELLEGREN H. Heterogeneous mutation processes in human microsatellite DNA sequences. *Nature Genetics*, 2000, 24(4): 400-402
- GAO H, KONG J. Distribution characteristics and biological function of tandem repeat sequences in the genomes of different organisms. *Zoological Research*, 2005, 26(5): 555-564 [高焕, 孔杰. 串联重复序列的物种差异及其生物功能. *动物学研究*, 2005, 26(5): 555-564]
- HANCOCK J M. Simple sequences and the expanding genome. *BioEssays*, 1996, 18(5): 421-425
- HE C B, CHEN S J, GAO X G, *et al*. Genetic diversity analysis

- of channel catfish (*Ictalurus punctatus*) germplasms using EST-SSR markers. *Journal of Agricultural Biotechnology*, 2008, 16(5): 815–823 [赫崇波, 陈姝君, 高祥刚, 等. EST-SSR 标记对我国斑点叉尾鲷种质资源的研究. *农业生物技术学报*, 2008, 16(5): 815–823]
- HUANG J, DU L M, LI Y Z, *et al.* Distribution regularities of microsatellites in the *Gallus gallus* genome. *Sichuan Journal of Zoology*, 2012, 31(3): 358–363 [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. *四川动物*, 2012, 31(3): 358–363]
- HUANG J, ZHOU Y, LIU Y Z, *et al.* Characteristics of microsatellites in *Arborophila rufipectus* genome sequences using 454 GS FLX. *Sichuan Journal of Zoology*, 2015, 34(1): 8–14 [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. *四川动物*, 2015, 34(1): 8–14]
- KATTI M V, RANJEKAR P K, GUPTA V S. Differential distribution of simple sequence repeats in eukaryotic genome sequences. *Molecular Biology and Evolution*, 2001, 18(7): 1161–1167
- KOVTUN I V, GOELLNER G, MCMURRAY C T. Structural features of trinucleotide repeats associated with DNA expansion. *Biochemistry and Cell Biology*, 2001, 79(3): 325–336
- LEOPOLDINO A M, PENSA S D. The mutational spectrum of human autosomal tetranucleotide microsatellites. *Human Mutation*, 2003, 21(1): 71–79
- LI W J, LI Y Z, DU L M, *et al.* Comparative analysis of microsatellite sequences distribution in the genome of giant panda and polar bear. *Sichuan Journal of Zoology*, 2014, 33(6): 874–878 [李午佼, 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. *四川动物*, 2014, 33(6): 874–878]
- LIU Z J, LI P, ARGUE B J, *et al.* Inheritance of RAPD markers in channel catfish (*Ictalurus punctatus*) × blue catfish (*I. furcatus*) and their F₁, F₂ and backcross hybrids. *Animal Genetics*, 1998, 29: 58–62
- LIU Z J, LI P, ARGUE B, *et al.* Random amplified polymorphic DNA markers: Usefulness for gene mapping and analysis of genetic variation of catfish. *Aquaculture*, 1999, 174(1): 59–68
- LU T, WANG C, DU C, *et al.* Distribution regularity of microsatellites in *Moschus berezovskii* genome. *Sichuan Journal of Zoology*, 2017, 36(4): 420–424 [卢婷, 王晨, 杜超, 等. 林麝全基因组微卫星分布规律研究. *四川动物*, 2017, 36(4): 420–424]
- MICKETT K, MORTON C, FENG J, *et al.* Assessing genetic diversity of domestic populations of channel catfish (*Ictalurus punctatus*) in Alabama using AFLP markers. *Aquaculture*, 2003, 228(1/2/3/4): 91–105
- MORGANTE M, HANAFEY M, POWELL W. Microsatellites are preferentially associated with non-repetitive DNA in plant genomes. *Nature Genetics*, 2001, 30: 194–200
- NARASIMHAMOORTHY B, SAHA M C, SWALLER T, *et al.* Genetic diversity in switch grass collections assessed by EST-SSR markers. *Bioenergy Research*, 2008, 1(2): 136–146
- NI S S, YANG Y, LIU S F, *et al.* Microsatellite analysis of *Patinopecten yessoensis* using next-generation sequencing method. *Progress in Fishery Sciences*, 2018, 39(1): 107–113 [倪守胜, 杨钰, 柳淑芳, 等. 基于高通量测序的虾夷扇贝基因组微卫星特征分析. *渔业科学进展*, 2018, 39(1): 107–113]
- NIE H, CAO S S, ZHAO M L, *et al.* Comparative analysis of microsatellite distributions in genomes of *Boa constrictor* and *Protobothrops mucrosquamatus*. *Sichuan Journal of Zoology*, 2017, 36(6): 639–648 [聂虎, 曹莎莎, 赵明朗, 等. 红尾蚺和原矛头蝮基因组微卫星分布特征比较分析. *四川动物*, 2017, 36(6): 639–648]
- PEARSON C E, SINDEN R R. Alternative structures in duplex DNA formed within the trinucleotide repeats of the myotonic dystrophy and fragile X loci. *Biochemistry*, 1996, 35(15): 5041–5053
- QI W H, JIANG X M, XIAO G S, *et al.* Distribution regularities of microsatellites in the pig genome. *Animal Husbandry and Veterinary Medicine*, 2014, 46(8): 9–13 [戚文华, 蒋雪梅, 肖国生, 等. 猪全基因组中微卫星分布规律. *畜牧与兽医*, 2014, 46(8): 9–13]
- QI W H, JIANG X M, XIAO G S, *et al.* Seeking and bioinformatics analysis of microsatellite sequence in the genomes of cow and sheep. *Acta Veterinaria et Zootechnica Sinica*, 2013, 44(11): 1724–1733 [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. *畜牧兽医学报*, 2013, 44(11): 1724–1733]
- QI W H, YAN C C, XIAO G S, *et al.* Distribution regularities and bioinformatics analysis of microsatellite in the whole genomes of goat and Tibetan antelope. *Journal of Sichuan University (Natural Science)*, 2016, 53(4): 937–944 [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. *四川大学学报(自然科学版)*, 2016, 53(4): 937–944]
- SCHLÖTTERER C. Genome evolution: Are microsatellites really simple sequences. *Current Biology*, 1998, 8(4): 132–134
- SCHORDERET D F, GARTER S M. Analysis of CpG suppression in methylated and nonmethylated species. *Proceedings of the National Academy of Sciences of USA*, 1992, 89(3): 957–961
- SHEN X, YANG G, LIU Y, *et al.* Construction of genetic linkage maps of guppy (*Poecilia reticulata*) based on AFLP and microsatellite DNA markers. *Aquaculture*, 2007, 271(1):

- 178–187
- SIMMONS M, MICKETT K, KUCUKTAS H, *et al.* Comparison of domestic and wild channel catfish (*Ictalurus punctatus*) populations provides no evidence for genetic impact. *Aquaculture*, 2006, 252(2/3/4): 133–146
- STALLINGS R L. CpG suppression in vertebrate genomes does not account for the rarity of (CpG)_n microsatellite repeats. *Genomics*, 1992, 13(3): 890–891
- SUBRAMANIAN S, MISHRA R K, SINGH L. Genome-wide analysis of microsatellite repeats in humans: Their abundance and density in specific genomic regions. *Genome Biology*, 2003, 4(2): 1–10
- TONG X L, DAI F Y, LI B, *et al.* Microsatellite repeats in mouse: Abundance, distribution and density. *Current Zoology*, 2006, 52(1): 138–152 [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. *动物学报*, 2006, 52(1): 138–152]
- TU F Y, LIU J, HAN W J, *et al.* Analysis of microsatellite distribution characteristics in the entire genome of *Macaca fascicularis*. *Chinese Journal of Wildlife*, 2018, 39(2): 400–404 [涂飞云, 刘俊, 韩卫杰, 等. 食蟹猴全基因组微卫星分布特征分析. *野生动物学报*, 2018, 39(2): 400–404]
- TU F Y, LIU X H, DU L M, *et al.* Distribution characteristics of microsatellites in the rat genome. *Acta Agriculturae Universitatis Jiangxiensis*, 2015, 37(4): 708–711 [涂飞云, 刘晓华, 杜联明, 等. 大鼠全基因组微卫星分布特征研究. *江西农业大学学报*, 2015, 37(4): 708–711]
- WANG C, DU L M, LI P, *et al.* Distribution patterns of microsatellites in the genome of the German cockroach (*Blattella germanica*). *Acta Entomologica Sinica*, 2015, 58(10): 1037–1045 [王晨, 杜联明, 李鹏, 等. 德国小蠊全基因组中微卫星分布规律. *昆虫学报*, 2015, 58(10): 1037–1045]
- WANG Y Y, LIU X X, DONG K Z, *et al.* Distribution difference of microsatellite in 7 domestic animals genomes. *China Animal Husbandry and Veterinary Medicine*, 2015, 42(9): 2418–2426 [王月月, 刘雪雪, 董坤哲, 等. 7种家养动物全基因组微卫星分布的差异研究. *中国畜牧兽医*, 2015, 42(9): 2418–2426]
- XIA J, LIU F, ZHU Z, *et al.* A consensus linkage map of the grass carp (*Ctenopharyngodon idella*) based on microsatellites and SNPs. *BMC Genomics*, 2010, 11(1): 135–151
- XIANG J G, ZHOU J, JIN H. A study on the biological and biochemical characteristics of the *Ictalurus punctatus*. *Hunan Journal of Hunan Agricultural University (Natural Science)*, 2004, 30(4): 355–358 [向建国, 周进, 金宏. 斑点叉尾鲷的生物学与生理生化特性研究. *湖南农业大学学报(自然科学版)*, 2004, 30(4): 355–358]
- XU J J, ZHENG X, LI J, *et al.* Distribution characteristics of whole genome microsatellite of *Pelteobagrus fulvidraco*. *Genomics and Applied Biology*, 2020, 39(12): 5488–5498 [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. *基因组学与应用生物学*, 2020, 39(12): 5488–5498]
- XU J J, ZHENG X, ZHANG X Y, *et al.* Analysis of distribution characteristics of microsatellites in the four genomes of puffer fish. *Genomics and Applied Biology*, 2021, 40(4): 1441–1451 [徐杰杰, 郑翔, 张鑫宇, 等. 4种河鲀全基因组微卫星分布特征分析研究. *基因组学与应用生物学*, 2021, 40(4): 1441–1451]
- ZHANG L L, WEI Z M, LIAN Z M, *et al.* Abundance of microsatellites in the entire genome and EST of *Tribolium castaneum*. *Chinese Bulletin of Entomology*, 2008, 45(1): 38–42 [张琳琳, 魏朝明, 廉振民, 等. 赤拟谷盗全基因组和 EST 中微卫星的丰度. *昆虫知识*, 2008, 45(1): 38–42]
- ZHANG T, GAN J, CHEN J W, *et al.* Morphological characteristics and electrophoretic analysis of two kinds of isozymes from *Ictalurus punctatus*. *Chinese Fishery Quality and Standards*, 2019, 9(6): 57–64 [张涛, 甘金, 陈建武, 等. 斑点叉尾鲷形态观察及其两种同工酶的电泳分析. *中国渔业质量与标准*, 2019, 9(6): 57–64]
- ZHENG Y, JIN G L, WU W R. Relationship between sequence completeness and polymorphism of microsatellites. *Genomics and Applied Biology*, 2012, 31(6): 587–591 [郑燕, 金谷雷, 吴为人. 微卫星序列完整性与多态性的关系. *基因组学与应用生物学*, 2012, 31(6): 587–591]

(编辑 冯小花)

Analysis of Microsatellite Distribution Characteristics in the Channel Catfish (*Ictalurus punctatus*) Genome

TANG Rongye, SU Mengyuan, YANG Wenshan, XU Jiejie, WANG Tao^①, YIN Shaowu^①

(College of Marine Science and Engineering, Nanjing Normal University Nanjing, Jiangsu Province Engineering Research Center for Aquatic Animals Breeding and Green Efficient Aquacultural Technology, Nanjing, Jiangsu 210023, China)

Abstract To understand the distribution of perfect microsatellites in the genome of the channel catfish (*Ictalurus punctatus*), we used MISA, a bioinformatics software package, to search and analyze the microsatellites. A total of 510 256 perfect microsatellites were isolated from 29 chromosomes of *I. punctatus*, with a total length of 11 036 941 bp. The chromosome containing the largest number of microsatellites was chromosome 2 (25 284), followed by chromosomes 3, 1, and 5. Chromosome 29 had the lowest number of microsatellites (11 591). The length of each chromosome was significantly correlated with the number of microsatellites it contained (SPSS, $r = 0.98$, $P < 0.01$). The highest relative abundance of microsatellites was found on chromosome 27 (785.03 ind./Mb), and the lowest was on chromosome 11 (615.89 ind./Mb). Among the six repeat types, mononucleotides were the most frequent, accounting for 45.31% of the total, followed by dinucleotides (38.53%), trinucleotides (8.73%), tetranucleotides (6.93%), pentanucleotides (0.46%), and hexanucleotides (0.04%). The predominantly repeated microsatellite sequences in the *I. punctatus* genome were A, AC, AG, AT, AAT, AAAT, C, AAC, AAAC, and AAG, showing an obvious inclination towards A and T bases. The results of this study provide a reference for the further study of *I. punctatus* genome characteristics and contribute basic data for future investigations into molecular marker-assisted breeding and genetic information assessment of *I. punctatus*.

Key words *Ictalurus punctatus*; Whole genome; Microsatellites; Distribution characteristics

① Corresponding author: WANG Tao, E-mail: seawater88@126.com; YIN Shaowu, E-mail: yinshaowu@163.com