

DOI: 10.19663/j.issn2095-9869.20210206003

http://www.yykxjz.cn/

陈丽梅, 李莉, 石栩蔚, 秦艺铭, 刘利华, 郭永军. 基于转录组数据的毛蚶 SSR 分子标记开发与评价. 渔业科学进展, 2022, 43(3): 129-137

CHEN L M, LI L, SHI X W, QIN Y M, LIU L H, GUO Y J. Development and evaluation of SSR markers based on transcriptome sequencing in *Scapharca kagoshimensis*. Progress in Fishery Sciences, 2022, 43(3): 129-137

# 基于转录组数据的毛蚶 SSR 分子标记开发与评价\*

陈丽梅<sup>1</sup> 李莉<sup>2</sup> 石栩蔚<sup>1</sup> 秦艺铭<sup>1</sup> 刘利华<sup>1</sup> 郭永军<sup>1①</sup>

(1. 天津农学院水产学院 天津市水产生态及养殖重点实验室 天津 300384;

2. 山东省海洋生物研究院 山东 青岛 266104)

**摘要** 本研究基于毛蚶(*Scapharca subcrenata*)的转录组数据, 利用 MISA 软件对其中的微卫星位点进行挖掘。从 35 555 条 unigene 中共获得 3987 个 SSR, SSR 出现频率达 11.21%。SSR 重复类型主要以二核苷酸重复为主(58.06%), 其次为三核苷酸重复(19.04%)。共有 182 种重复基元, 不同类型 SSR 的重复基元分布特征不同, 其中, 二核苷酸重复基元中 AC/GT 类型比例最高, 为 45.70%。毛蚶转录组中 SSR 重复次数主要集中在 5~7 次, SSR 长度主要集中在 12~29 bp, 多态性均在中等以上。利用筛选出的 14 对 SSR 引物在山东潍坊毛蚶野生群体中进行遗传多样性分析, 结果显示, 平均有效等位基因数( $N_a$ )、平均观测杂合度( $H_o$ )、平均期望杂合度( $H_e$ )和多态性信息含量(PIC)分别为 15.4、0.682、0.852 和 0.817, 从 PIC 值来看, 本研究开发的 14 个微卫星标记均属高多态性标记( $PIC \geq 0.5$ )。此外, 有 7 个位点显著偏离哈迪-温伯格平衡(Hardy-Weinberg equilibrium, HWE) ( $P < 0.05$ )。结果表明, 基于毛蚶转录组数据开发微卫星标记是切实可行的, 研究结果丰富了毛蚶的分子标记数量, 对毛蚶的种群遗传学分析、遗传图谱构建及分子辅助育种等研究具有重要意义。

**关键词** 毛蚶; 转录组; SSR; 遗传多样性

**中图分类号** S917 **文献标识码** A **文章编号** 2095-9869(2022)03-0129-09

毛蚶(*Scapharca kagoshimensis*)俗称毛蛤、麻蛤、麻蚶等(王如才等, 2008), 属广温广盐性经济贝类。毛蚶广泛分布于中国、日本、朝鲜沿岸, 在中国以莱州湾、渤海湾、辽东湾、海州湾等浅海区资源尤为丰富(王庆志等, 2015)。对毛蚶的研究主要集中在形态学(陈蓉等, 2009; 宋菲菲等, 2012)、苗种繁育(翟林香等, 2010; 马云聪等, 2008)及养殖技术(王庆志等, 2015)等, 关于群体遗传学的研究较少(赵文等, 2011;

田吉腾等, 2016)。进入 20 世纪 80 年代, 受生境破坏及过度捕捞等因素的影响, 毛蚶自然资源逐年减少, 产量大幅度下降, 已远远不能满足市场需求, 毛蚶自然种群的遗传评估和资源恢复工作也亟待进行。

微卫星分子标记具有分布范围广、呈共显性遗传、多态性高和重复性好等特点, 在群体遗传多样性分析、种质资源评价、遗传图谱构建等研究中有明显优势(Nikolic *et al*, 2009; Bao *et al*, 2016; Kewwuwan

\*财政部和农业农村部: 国家现代农业产业技术体系(CARS-48)、山东省现代农业产业技术体系建设专项资金资助项目(SDAIT-14-04)和天津市企业特派员项目(19JCTPJC60100)共同资助 [This work was supported by China Agriculture Research System of MOF and MARA (CARS-48), Earmarked Fund for Modern Agro-Industry Technology Research System in Shandong Province (SDAIT-14-04), and Tianjin Technical Expert Project (19JCTPJC60100)]. 陈丽梅, E-mail: chenlimeicc@163.com

① 通讯作者: 郭永军, 研究员, E-mail: guoyongjun@tjau.edu.cn

收稿日期: 2021-02-06, 收修改稿日期: 2021-02-26

et al, 2016)。微卫星的开发方法主要有直接文库筛选法(战爱斌等, 2008)、磁珠富集法(宋丹丹等, 2019)、数据库检索法(魏大为等, 2015)、种间转移扩增法(Liu et al, 2007)等。目前, 毛蚶已开发出部分微卫星标记, 如 Feng 等(2009)利用磁珠富集法筛选了 14 对毛蚶微卫星引物, 陈辰(2015)利用同样的方法获得 41 对多态性较好的微卫星引物。另外, Li 等(2012)通过在魁蚶(*Scapharca broughtonii*)中构建 cDNA 文库, 得到了 25 条在魁蚶和毛蚶中均通用的 EST-SSR 引物。Dong 等(2012)利用泥蚶(*Tegillarca granosa*)的转录组数据, 开发了 62 个 EST-SSR 引物, 在毛蚶中的通用率为 25.81%。但毛蚶微卫星标记的数量还远不能满足其遗传分析需求, 需要开发更多可利用的微卫星标记。近年来, 利用转录组测序平台开发 SSR 分子标记已经得到越来越多的应用, 其相比于传统方法, 具有通量大、周期短、成本低且直接与某些功能基因相关联的优点(于爱清等, 2019)。

本研究基于毛蚶鳃组织的转录组数据, 对其中的 SSR 位点进行搜索, 分析其分布特征及组成类型; 并利用所筛选的引物在潍坊毛蚶野生群体中进行了遗传多样性分析, 旨在为毛蚶微卫星分子标记的开发提供更为有效的方法, 为毛蚶种群遗传学分析、遗传图谱构建及分子标记辅助育种提供有力工具。

## 1 材料与方法

### 1.1 实验材料及转录组测序

实验所需毛蚶样品于 2017 年 10 月采自天津滨海新区大神堂海域, 壳长为(26.19±1.46) mm, 活体重为(20.01±2.89) g。选取活力好的毛蚶 3 只, 取其鳃组织, 构建高通量转录组测序文库, 并利用 Illumina HiSeq4000 测序平台(杭州联川生物技术股份有限公司)进行测序, 共获得 93.57 Gb 的数据量, 原始测序数据经 Trinity 软件(<http://trinityrnaseq.sourceforge.net/>)组装并去除冗余后, 获得了平均长度为 621.23 bp 的 unigene 35 555 条。

### 1.2 微卫星标记筛选和引物设计

微卫星序列的识别和定位使用 MISA 软件进行, MISA 搜索参数设置如下: 单核苷酸、二核苷酸、三核苷酸、四核苷酸、五核苷酸和六核苷酸的最少重复次数分别为 12、6、5、5、4 和 4, 复合 SSR 两个位点间最大间隔碱基数设置为 100, 结合 Primer 3 软件批量设计引物。主要参数设置: 引物长度为 18~27 bp, PCR 扩增产物在 100~280 bp 之间, 57℃ ≤ 退火温度

( $T_m$ ) ≤ 63℃, 20% ≤ GC 含量 ≤ 80%。同时, 利用 NCBI 数据库中的 Primer-BLAST 对结果进行进一步验证, 主要参数设置: Self complementarity < 6, Self 3 complementarity < 3, 且上下游引物  $T_m$  值差值不高于 2℃。引物由生工生物工程(上海)股份有限公司合成。

### 1.3 微卫星多态性分析

使用海洋动物组织基因组 DNA 提取试剂盒(天根生化科技(北京)有限公司)进行毛蚶肌肉组织 DNA 的提取, 委托生工生物工程(上海)股份有限公司合成荧光引物, 对毛蚶潍坊群体的 30 个个体进行 STR 检测(3730XL 测序分析仪, 美国 ABI 公司), 利用 GeneMapper 3.2 软件对结果进行基因分型。

等位基因数( $N_a$ )、期望杂合度( $H_o$ )和观测杂合度( $H_e$ )用 Microsatellite Analyser (MSA) 软件计算(Dieringer et al, 2003)。多态信息含量(polymorphism information content, PIC)用 PIC-CALC 0.6 软件计算。利用 Genepop 4.2 软件进行哈迪-温伯格平衡(Hardy-Weinberg equilibrium, HWE)分析(<http://genepop.curtin.edu.au/>), 并对  $P$  值进行 Bonferroni 校正。

## 2 结果与分析

### 2.1 毛蚶转录组中 SSR 的数量和分布特点

对毛蚶转录组中序列总长度 22 087 807 bp 的 35 555 条 unigene 序列进行 SSR 检测, 共得到 3987 个 SSR (完美型 SSR 为 3074 个), 分布在 3162 条 unigene 序列当中。SSR 发生频率(含 SSR 的 unigene 数/unigene 总数)为 8.89%, SSR 出现频率(检出 SSR 位点数/unigene 总数)为 11.21%, 平均每 5.54 kb 含有 1 个 SSR 位点(总 unigene 长度/搜索到的 SSR 数量)。其中, 612 条 unigene 含有 1 个以上 SSR 位点, 以复合微卫星形式出现的 SSR 数目为 913 个。

### 2.2 毛蚶转录组中 SSR 的重复基元类型和频率特征

从毛蚶转录组中筛查到的 3987 个 SSR 位点中, 二核苷酸重复 SSR 出现频率最高, 为 2315 个, 占总数的 58.06%; 其次是三核苷酸重复 SSR (759 个, 19.04%)和单核苷酸重复 SSR (696 个, 17.46%); 四核苷酸、五核苷酸和六核苷酸重复 SSR 所占比例最少, 分别占总数的 4.01%、1.15%和 0.28%。

3987 个 SSR 共有 182 种重复基元类型, 根据碱基互补配对原则和阅读起始碱基顺序的差异(栾生等, 2007), 对每种长度类型的重复基元进行同类兼并后, 单、二、三、四、五、六核苷酸重复基元类型分别为 2、4、10、18、17 和 8 种(表 1)。单核苷

表 1 毛蚶转录组 SSR 主要重复基元类型及分布比例

Tab.1 The type and percentage of main repeat motif of SSR in transcriptome of *S. kagoshimensis*

重复类型 Type of SSR	重复基元 Repeat motif	数量 Number	所占比例 Ratio/%	重复类型 Repeat type	重复基元 Repeat motif	数量 Number	所占比例 Ratio/%
单核苷酸 Mononucleotide	A/T	582	83.62	三核苷酸 Trinucleotide	AAC/GTT	119	15.68
	C/G	114	16.38		ACC/GGT	54	7.11
二核苷酸 Binucleotide	AC/GT	1058	45.70		AAG/CTT	39	5.14
	AG/CT	766	33.09		AGC/CTG	39	5.14
	AT/AT	480	20.73		AGG/CCT	29	3.82
	CG/CG	11	0.48		ACG/CGT	23	3.03
三核苷酸 Trinucleotide	AAT/ATT	264	34.78		ACT/AGT	12	1.58
	ATC/ATG	171	22.53		CCG/CGG	9	1.19

酸重复中出现最多的基元类型是 A/T, 占单核苷酸重复的比例为 83.62%。二核苷酸重复中出现较多的是 AC/GT 和 AG/CT 类型, 占二核苷酸重复的比例分别为 45.70%和 33.09%, AAT/ATT 基元在三核苷酸重复中出现的频率最高, 占三核苷酸重复的 34.78%。

2.3 毛蚶转录组中 SSR 长度及重复次数分布

本研究中, 毛蚶转录组中 SSR(完美型)长度在 12~84 bp 之间, 其中, 长度在 12~19 bp 之间的 SSR 占比最大, 为 SSR 总数的 61.06%。其次为长度 20~29 bp 的 SSR, 占总数的 22.38%。长度 ≥ 50 bp 的 SSR 仅占总数的 2.64%。SSR 的长度分布规律见图 1。

毛蚶转录组中不同类型 SSR 的重复次数分布见表 2。由表 2 可知, SSR 单核苷酸重复次数主要集中在 12~14 次, 二核苷酸重复次数主要集中在 6~8 次, 三核苷酸和四核苷酸重复次数主要集中在 5~6 次, 五核苷酸和六核苷酸重复次数主要为 4 次。总的来看, 重复 6 次的 SSR 占比最高, 为 778 个(19.51%), 其次为重复 5 次和 7 次, 分别为 488 个(12.24%)和 423 个(10.61%)。重复次数 ≥ 15 次的 SSR 位点共有 803 个, 占总 SSR 个数的 20.14%。

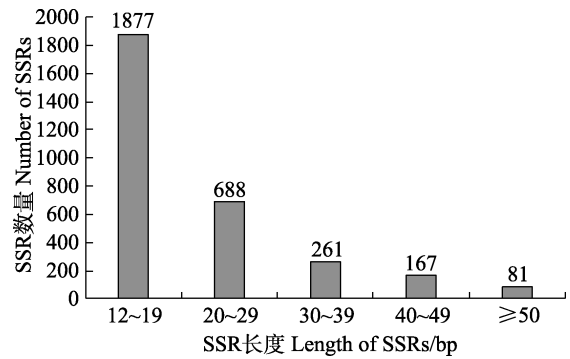


图 1 毛蚶转录组中的 SSR 长度分布  
Fig.1 Length distribution of SSRs in transcriptome of *S. kagoshimensis*

2.4 毛蚶 SSR 引物设计及群体遗传多样性分析

除部分微卫星序列侧翼序列过短或者本身结构不适合设计引物以外, 使用 Primer 3 软件对其中含有 SSR 的 3162 条序列进行引物设计, 最后从 60 对引物中筛选出扩增稳定、多态性较高的 14 对引物(表 3)。利用这 14 对引物对毛蚶潍坊野生群体的 30 个个体进行遗传多样性分析。等位基因数( $N_a$ )、观测杂合度( $H_o$ )、期望杂合度( $H_e$ )、多态信息含量(PIC)和哈迪-温伯格平衡检验( $P_{HWE}$ )等参数详见表 4。

表 2 毛蚶转录组中不同类型 SSR 重复次数分布

Tab.2 Repeat number of SSR in transcriptome of *S. kagoshimensis*

重复类型 Type of SSR	重复次数 Repeat number											
	4	5	6	7	8	9	10	11	12	13	14	≥15
单核苷酸 Mononucleotide	-	-	-	-	-	-	-	-	189	105	114	288
二核苷酸 Binucleotide	-	-	591	337	233	132	109	176	126	63	59	489
三核苷酸 Trinucleotide	-	411	147	78	47	14	8	12	8	9	4	21
四核苷酸 Tetranucleotide	-	72	38	8	11	6	3	5	5	3	4	5
五核苷酸 Pentanucleotide	34	5	2	-	3	-	1	-	1	-	-	0
六核苷酸 Hexanucleotide	11	-	-	-	-	-	-	-	-	-	-	0
总和 Total	45	488	778	423	294	152	121	193	329	180	181	803

表3 毛蚶14对SSR引物信息

Tab.3 Information of SSR primers in *S. kagoshimensis*

位点 Locus	重复基元 Repeat motif	引物序列 Primer sequences (5'~3')	产物平均长度 Mean product size /bp	退火温度 $T_m$ / $^{\circ}$ C
MH09	(AT) <sub>7</sub>	F: ATTCGGCACACTTTTGAGGT R: CCAATTTGGCAAGTGCTTTT	174	60
MH17	(TGA) <sub>6</sub>	F: ACAATGCCAAGGAAGAATCG R: CGGCTCATCAGCATACTCAA	272	60
BE02	(GA) <sub>12</sub>	F: CGTCTTTTGTGAAGATCCTTG R: ACGAAAATCCAAAAGAGCGA	258	56
BE04	(ATA) <sub>7</sub>	F: TTGCCGAGAAGGTTTGTTTT R: GCCCTACCTTGAGTCGTAA	213	58
BE10	(GGA) <sub>6</sub>	F: TGTAGTGCAAGGTGGAGCTG R: TGTCCAGTGCCATATCCAGA	254	61
BE16	(GT) <sub>12</sub>	F: GGTACAGTGTGGGCCAAAAG R: GAGTGGGGATGTTGGAGTGA	286	59
BE18	(TTC) <sub>6</sub>	F: CACGCTGTTTTACCTTTTG R: GAATCTGTCAATCGGTCGGT	267	61
RC02	(AC) <sub>12</sub>	F: GGTTAATGACGCACACATGC R: CTTTCAATGACCACTGGCAA	232	57
RC06	(TG) <sub>12</sub>	F: TTGCTGCCAGTGTGTGTGTA R: ACCAGTGAAAACCACCAAGC	247	60
RC15	(TAT) <sub>7</sub>	F: AACACAATGCCCTACTTGGC R: AACACCCAACAAACCTCAGAA	194	60
RC20	(AG) <sub>12</sub>	F: TGGTGACTGCTTACTTGGGA R: AGCGAAACGCTAAAAGCAAA	274	58
KT05	(AC) <sub>12</sub>	F: TTTATATGTGGCGGTCAGCA R: GATGGGCTATCCAGATGTCCT	159	60
KT09	(CAA) <sub>6</sub>	F: TGCGTCAACAACAAAAGGAG R: GGTACATTTCCAGCATTCCG	164	60
KT10	(TGA) <sub>6</sub>	F: CAGTACCAAAGACGGCCATT R: TGAGGTACTTGTTGGTCCAG	237	60

表4 基于14个SSR位点的潍坊毛蚶群体遗传多样性分析

Tab.4 Genetic diversity of Weifang population in *S. kagoshimensis* at 14 SSR loci

微卫星位点 Locus	重复基元 Repeat motif	等位基因数 $N_a$	观测杂合度 $H_o$	期望杂合度 $H_e$	多态信息含量 PIC	哈-温平衡检验 $P_{HWE}$
MH09	(AT) <sub>7</sub>	11	0.967	0.805	0.767	0.019 1
MH17	(TGA) <sub>6</sub>	6	0.533	0.771	0.719	0.009 9
BE02	(GA) <sub>12</sub>	19	0.933	0.942	0.921	0.294 7
BE04	(ATA) <sub>7</sub>	19	0.567	0.869	0.843	0.000 0*
BE10	(GGA) <sub>6</sub>	9	0.500	0.780	0.735	0.000 0*
BE16	(GT) <sub>12</sub>	25	0.759	0.962	0.943	0.000 0*
BE18	(TTC) <sub>6</sub>	23	0.733	0.953	0.933	0.003 2*
RC02	(AC) <sub>12</sub>	26	0.793	0.953	0.933	0.010 0
RC06	(TG) <sub>12</sub>	17	0.577	0.930	0.905	0.000 0*
RC15	(TAT) <sub>7</sub>	4	0.567	0.623	0.549	0.249 8
RC20	(AG) <sub>12</sub>	19	0.900	0.938	0.917	0.287 1
KT05	(AC) <sub>12</sub>	21	0.733	0.925	0.904	0.000 0*
KT09	(CAA) <sub>6</sub>	8	0.600	0.669	0.600	0.129 0
KT10	(TGA) <sub>6</sub>	9	0.379	0.802	0.769	0.000 0*
平均 Mean		15.4	0.682	0.852	0.817	-

\*表示 Bonferroni 校正后显著偏离哈迪-温伯格平衡( $P<0.05$ )\* means significant deviation from Hardy-Weinberg equilibrium after Bonferroni correction ( $P<0.05$ )

14 个微卫星位点总共检测到 216 个等位基因, 每个位点的等位基因个数从 4 个到 26 个不等, 平均等位基因数为 15.4 个。 $H_o$  为 0.379~0.967, 平均值为 0.682;  $H_e$  为 0.623~0.962, 平均值为 0.852, PIC 为 0.817。经 Bonferroni 校正后发现, 14 个位点中有 7 个偏离哈迪-温伯格平衡。

### 3 讨论

#### 3.1 毛蚶转录组 SSR 类型和分布规律

近年来, 越来越多的贝类中报道了利用转录组数据开发微卫星的方法, 在墨西哥湾扇贝(*Argopecten irradians concentricus*)的转录组测序数据中, SSR 的出现频率为 10%(谭杰等, 2018), 马氏珠母贝(*Pinctada martensii*)的 SSR 出现频率为 13.34%(王忠良等, 2015), 泥东风螺(*Babylonia lutosus*)的 SSR 出现频率为 13.62%(熊钢等, 2020)。本研究中, 毛蚶转录组中 SSR 的出现频率为 11.21%, 和上述报道相似, 但低于扁玉螺(*Neverita didyma*)的 86.53%(卢玮筱等, 2018)。在其他水生动物中, 转录组中 SSR 出现频率差异较大, 如牙鲆(*Paralichthys olivaceus*)为 43.24%(李超等, 2015), 大口黑鲈(*Micropterus salmoides*)为 11.30%(黄勇等, 2019), 金乌贼(*Sepia esculenta*)为 90.66%(张金勇等, 2020), 曼氏针乌贼(*Sepiella japonica*)为 48.70%(管奥等, 2018), 凡纳滨对虾(*Litopenaeus vannamei*)为 22.1%(杨铭等, 2017)。在 SSR 检索标准一致的情况下, SSR 出现频率主要与物种差异、转录组结构及测序数据大小有关。

除单核苷酸重复外, 大多数贝类如墨西哥湾扇贝、马氏珠母贝、泥东风螺及菲律宾蛤仔(*Ruditapes philippinarum*)(谭杰等, 2018; 王忠良等, 2015; 熊钢等, 2020; 闫路路, 2015)等转录组中 SSR 均以二核苷酸重复为主, 本研究中毛蚶转录组中 SSR 也以二核苷酸重复为主, 和上述研究结论一致。多数鱼类, 如牙鲆、刀鲚(*Coilia ectenes*)、密斑刺鲀(*Diodon hystrix*)(李超等, 2015; 于爱清等, 2019; 马军等, 2020)以及虾类, 如凡纳滨对虾、红螯螯虾(*Cherax quadricarinatus*)(杨铭等, 2017; 李喜莲等, 2020)等, 转录组中的 SSR 类型分布也符合这一规律。

本研究中二核苷酸重复基元类型以 AC/GT 最多, 占二核苷酸重复的 45.70%, 这与合浦珠母贝(*Pinctada fucata*)、泥东风螺的研究结果一致(曲妮妮, 2008; 熊钢等, 2020)。目前, 大多数鱼类的研究结果也验证了这一结论, 如牙鲆、刀鲚、团头鲂(*Megalobrama amblycephala*)和鳊鱼(*Miichthys miiuy*)

转录组中的二核苷酸重复类型中, AC/GT 重复基元分别占 75.25%、71.89%、65.39%和 68.2%(李超等, 2015; 于爱清等, 2019; 曾聪等, 2013; 车荣波等, 2015)。Toth 等(2000)对多个真核生物基因组的研究结果表明, 脊椎动物和节肢动物基因组中二核苷酸的类型多以 AC/GT 为主。与上述研究不同, 战爱斌等(2008)和李宏俊等(2009)通过富集文库法得到的栉孔扇贝(*Chlamys farreri*)和海湾扇贝(*Argopecten irradians*)基因组中 AG/CT 重复基元的比例远高于 AC/GT。不同贝类中重复单元类型占比不同的现象, 除了 SSR 筛选方法不同外, 还可能与不同编码基因的碱基组成偏好及体内甲基化酶活性有关(曾聪等, 2013)。

#### 3.2 毛蚶转录组 SSR 的长度和重复次数

SSR 标记多态性的高低是判断其物种可用性的重要依据。根据 Temnykh 等(2001)的研究, SSR 的长度大小是影响其多态性高低的关键因素, 当 SSR 长度  $\geq 20$  bp 时, 呈高度多态性,  $12 \text{ bp} \leq \text{长度} \leq 20$  bp 时, 呈中等多态性; 前者更容易发生变异, 原因是在较长的模板上滑动错配发生的几率更高(毛俐慧等, 2018)。本研究在设置筛选参数时, 已剔除了长度在 12 bp 以下的 SSR 序列, 12~19 bp 和长度  $\geq 20$  bp 的序列分别占 SSR 总数的 66.23%和 33.77%, 在此基础上设计引物, 保证了引物具有较高的多态性。

本研究筛选出的 14 对多态性较好的 SSR 引物中, 二核苷酸重复主要为 6 次, 三核苷酸重复主要为 7 次, 随着重复次数的增加, 各类型 SSR 的数量均逐渐降低。重复次数范围和 SSR 数量变化规律和大多数同类研究一致。相比之下, 陈辰(2015)和 Feng 等(2009)用磁珠富集方法开发的毛蚶 SSR 中, 高重复序列所占比例较大, 其二核苷酸重复次数一般集中在 13~27 次。造成这种差异的原因, 一方面是由于转录组中开发的 SSR 重复次数往往要低于基因组中的 SSR (Xia *et al.*, 2018), 另一方面是磁珠富集的杂交洗脱等程序能够将大部分低重复序列除去(孙效文等, 2005)。Schlötterer 等(2000)认为, 重复次数与 SSR 位点突变率存在一定程度正相关。虽然本研究从转录组中开发的 SSR 重复次数相对较低, 但从毛蚶群体遗传学分析的结果来看, 依然显示了较高的多态性。

#### 3.3 潍坊毛蚶群体遗传多样性分析

本研究对基于转录组数据得到的微卫星引物进行筛选, 并在潍坊毛蚶野生群体中进行了遗传多样性分析, 其平均等位基因数( $N_a$ )、观测杂合度( $H_o$ )和期望杂合度( $H_e$ )分别为 15.4、0.682 和 0.852, 与陈辰(2015)

( $N_a$ : 15.2;  $H_o$ : 0.719;  $H_e$ : 0.856)的研究结果类似, 相比 Feng 等(2009)( $N_a$ : 21.64;  $H_o$ : 0.798;  $H_e$ : 0.931)多态性略低。14 个微卫星位点的 PIC 值范围为 0.549~0.943, 平均值为 0.817。根据 Botstein 等(1980)的标准,  $PIC \geq 0.5$ 、 $0.25 \leq PIC < 0.5$  和  $PIC < 0.25$  分别代表高度多态性、中度多态性和低度多态性。从结果来看, 本研究开发的 14 个微卫星标记均属高多态性, 能够提供丰富的遗传信息, 在毛蚶种群遗传学研究、遗传图谱构建及家系分析鉴定中具有较高的实用性。另外, 本研究采集的样本为山东潍坊野生群体, 也说明目前群体依然保持着较高的遗传多样性。

经 Bonferroni 校正后发现, 14 个微卫星位点中有 7 个显著偏离哈迪-温伯格平衡( $P < 0.05$ ), 微卫星位点偏离哈迪-温伯格平衡的现象在相关报道中并不少见, 如在菲律宾蛤仔(闫路路, 2015)、青蛤(方军等, 2020)、里式拟石磺(*Paraoncidium reevesii*)(吴欣等, 2016)和熊本牡蛎(*Crassostrea sikamea*)(黄飘逸等, 2020)的群体遗传学研究中, 微卫星位点偏离哈迪-温伯格的比例分别为 75%、70.37%、35.71%和 20%。一般来说, 群体偏离平衡的原因主要是近交或存在无效等位基因而导致的纯合子过剩或取样样本量过小。其中, 微卫星标记中存在无效等位基因是一个较为普遍的现象, 尤其是在海洋贝类中(Reece *et al.*, 2004)。本研究采集的群体为野生群体, 从平均等位基因数、杂合度等遗传参数来看, 均体现出较高的多态性, 故推测微卫星中无效等位基因的存在或样本量较小是造成偏离哈迪-温伯格平衡的主要原因。

本研究结果表明, 利用高通量转录组测序筛选毛蚶微卫星是一种高效可行的方法。研究结果增加了毛蚶的可用微卫星标记, 对毛蚶的种群遗传学分析、遗传图谱构建及分子标记辅助育种都具有重要的意义。

## 参 考 文 献

- BAO M M, ZHOU Q, XIE C X, *et al.* Ten polymorphic microsatellite markers characterized for *Schizothorax pseudaksaiensis* and applied for population genetic analysis. *Journal of Genetics*, 2016, 95: e5-e8
- BOTSTEIN D, WHITE R L, SKOLNICK M, *et al.* Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *American Journal of Human Genetics*, 1980, 32(3): 314-331
- CHE R B. Development of molecule markers and gene differential expression analysis of *Miichthys miiuy* based on transcriptome sequencing data. Master's Thesis of Zhejiang Ocean University, 2015 [车荣波. 基于转录组数据的鲢鱼分子标记筛选及基因差异表达分析. 浙江海洋学院硕士研究生学位论文, 2015]
- CHEN C. Population genetics of the ark shell. Doctoral Dissertation of Ocean University of China, 2015 [陈辰. 毛蚶群体遗传学研究. 中国海洋大学博士研究生学位论文, 2015]
- CHEN R, LIU J Y, TANG L J, *et al.* Morphological variations analysis of five different populations of *Scapharca subcrenata* in China. *Marine Sciences*, 2009, 33(7): 64-69 [陈蓉, 刘建勇, 唐连俊, 等. 中国 5 个海区不同群体毛蚶形态差异分析. *海洋科学*, 2009, 33(7): 64-69]
- DIERINGER D, SCHLÖTTERER C. Microsatellite analyser (MSA): A platform independent analysis tool for large microsatellite datasets. *Molecular Ecology Notes*, 2003, 3(1): 167-169
- DONG Y H, YAO H H, LIN Z H, *et al.* Characterization of 62 polymorphic EST-SSR markers in the blood clam (*Tegillarca ganosa*) and their cross-amplification in *Scapharca subcrenata*. *Conservation Genetics Resources*, 2012, 4(4): 991-997
- FANG J, SHEN Y P, ZHANG L L, *et al.* Development and verification of SSR markers in *Cyclina sinensis*. *Journal of Applied Oceanography*, 2020, 39(2): 214-220 [方军, 沈彦鹏, 张雷雷, 等. 基于转录组数据的青蛤微卫星标记开发与验证. *应用海洋学学报*, 2020, 39(2): 214-220]
- FENG Y W, LI Q, KONG L F. Isolation and characterization of 14 polymorphic microsatellite loci in the ark shell *Scapharca subcrenata* (Bivalvia: Arcidae). *Conservation Genetics*, 2009, 10: 1125-1127
- GUAN A, WU Y T, CHEN Y, *et al.* Deep sequence-based transcriptome analysis of microsatellites in the cuttlefish (*Sepiella japonica*). *Progress in Fishery Sciences*, 2018, 39(3): 144-151 [管奥, 毋玉婷, 陈宇, 等. 曼氏无针乌贼转录组微卫星特征分析. *渔业科学进展*, 2018, 39(3): 144-151]
- HUANG P Y, MAO H T, YU Z N, *et al.* Evaluation of genetic diversity and versatility of wild and cultured populations based on high-throughput sequencing of *Crassostrea sikamea* microsatellite markers. *Journal of Fisheries of China*, 2020, 44(3): 368-377 [黄飘逸, 马海涛, 喻子牛, 等. 利用高通量测序开发的熊本牡蛎微卫星标记评价野生和养殖群体的遗传多样性及通用性. *水产学报*, 2020, 44(3): 368-377]
- HUANG Y, GONG W B, CHEN H G, *et al.* Sequencing and bioinformatic analysis for transcriptome of *Micropterus salmoides* based on RNA-seq. *South China Fisheries Science*, 2019, 15(1): 106-112 [黄勇, 龚望宝, 陈海刚, 等. 基于 RNA-Seq 高通量测序技术的大口黑鲈转录组分析. *南方水产科学*, 2019, 15(1): 106-112]
- KEWWUWAN K, KUBOTA S, LIU Q, *et al.* Detection of growth-related quantitative trait loci and high-resolution

- genetic linkage maps using simple sequence repeat markers in the kelp grouper (*Epinephelus bruneus*). *Marine Biotechnology*, 2016, 18(1): 57–84
- LI C, HOU J L, WANG G X, *et al.* Bioinformatic analysis of SSR markers in transcriptomic sequencing *Paralichthys olivaceus*. *Marine Fisheries*, 2015, 37(2): 122–127 [李超, 侯吉伦, 王桂兴, 等. 基于牙鲆RNA-seq数据中SSR标记的信息分析. *海洋渔业*, 2015, 37(2): 122–127]
- LI H J, LIU X, DU X D, *et al.* Development of microsatellite markers in bay scallop and their inheritance patterns in an F<sub>1</sub> hybrid family. *Marine Sciences*, 2009, 33(12): 4–8 [李宏俊, 刘晓, 杜雪地, 等. 海湾扇贝微卫星标记开发及其分离方式分析. *海洋科学*, 2009, 33(12): 4–8]
- LI M, ZHU L, ZHOU C Y, *et al.* Development and characterization of EST-SSR markers from *Scapharca broughtonii* and their transferability in *Scapharca subcrenata* and *Tegillarca granosa*. *Molecules*, 2012, 17(9): 10716–10723
- LI X L, GUO J L, LI Q, *et al.* The information analysis of SSR loci in the red claw crayfish (*Cherax quadricarinatus*) transcriptome. *Hubei Agricultural Sciences*, 2020, 59(7): 207–211 [李喜莲, 郭建林, 李倩, 等. 红螯螯虾转录组中的SSR位点信息分析. *湖北农业科学*, 2020, 59(7): 207–211]
- LIU B Q, ZENG Q G, WANG Y J, *et al.* The cross-species amplification and validation of EST-SSR loci in *Porphyra haitanensis*. *Acta Hydrobiologica Sinica*, 2007, 31(2): 149–154
- LU W X, CHEN Y X, QI P Z. Analysis of SSR Information in the transcriptional group of *Neverita didyma*. *Journal of Zhejiang Ocean University (Natural Science)*, 2018, 37(3): 215–220 [卢玮筱, 陈永霞, 祁鹏志. 扁玉螺转录组SSR信息分析. *浙江海洋大学学报(自然科学版)*, 2018, 37(3): 215–220]
- LUAN S, KONG J, WANG Q Y, *et al.* Analysis of microsatellites in the genome of Kuruma prawn *Marsupenaeus japonicus*. *Progress in Natural Science*, 2007, 17(6): 731–740 [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]
- MA J, LIU J X, JIANG Z J, *et al.* Development and identification of SSR markers based on RNA-seq data of *Diodon hystris*. *South China Fisheries Science*, 2020, 16(1): 127–136 [马军, 刘嘉鑫, 江智景, 等. 基于RNA-seq数据的密斑刺鲀SSR分子标记开发与鉴定. *南方水产科学*, 2020, 16(1): 127–136]
- MA Y C, YAN S H, SU W Q, *et al.* Practical techniques of breeding and culture of *Scapharca subcrenata*. Beijing: China Ocean Press, 2008 [马云聪, 闫升华, 苏文清, 等. 毛蚶育苗和养殖实用技术. 北京: 海洋出版社, 2008]
- MAO L H, LIU J X, DING H Q, *et al.* Microsatellite characterization analysis and primers design of the whole transcriptome of *Curcuma alismatifolia*. *Molecular Plant Breeding*, 2018, 16(22): 7407–7414 [毛俐慧, 刘建新, 丁华侨, 等. 姜荷花全长转录组微卫星特征分析和引物设计. *分子植物育种*, 2018, 16(22): 7407–7414]
- NIKOLIC N, FEVE K, CHEVALET C, *et al.* A set of 37 microsatellite DNA markers for genetic diversity and structure analysis of Atlantic salmon *Salmo salar* populations. *Journal of Fish Biology*, 2009, 74(2): 458–466
- QU N N. Microsatellite DNA isolation and genetic diversity of the pearl oyster *Pinctada fucata*. Master's thesis of Huazhong Agricultural University, 2008 [曲妮妮. 合浦珠母贝微卫星DNA标记分离与遗传多样性研究. 华中农业大学硕士研究生学位论文, 2008]
- REECE K S, RIBEIRO W L, GAFFNEY P M, *et al.* Microsatellite marker development and analysis in the eastern oyster (*Crassostrea virginica*): Confirmation of null alleles and non-mendelian segregation ratios. *Journal of Heredity*, 2004, 95(4): 346–352
- SCHLÖTTERER C. Evolutionary dynamics of microsatellite DNA. *Chromosoma*, 2000, 109(6): 365–371
- SONG D D, WANG L, WEI H J, *et al.* Construction and polymorphism analysis of a microsatellite library in red swamp crayfish *Procambarus clarkii*. *Fisheries Science*, 2019, 38(6): 819–826 [宋丹丹, 王龙, 魏红静, 等. 克氏原螯虾微卫星文库构建及多态性分析. *水产科学*, 2019, 38(6): 819–826]
- SONG F F, WANG Q X, LIU M, *et al.* Morphological variation analysis of *Scapharca subcrenata* in China. *Periodical of Ocean University of China (Natural Science)*, 2012, 42(11): 40–45 [宋菲菲, 王其翔, 刘名, 等. 不同地区毛蚶群体的形态差异的比较研究. *中国海洋大学学报(自然科学版)*, 2012, 42(11): 40–45]
- SUN X W, JIA Z Y, WEI D W, *et al.* Comparison between magnetic beads enriched and small inserted fragment library for microsatellite sequences of common carp. *Journal of Fishery Sciences of China*, 2005, 12(2): 126–132 [孙效文, 贾智英, 魏东旺, 等. 磁珠富集法与小片段克隆法筛选鲤微卫星的比较研究. *中国水产科学*, 2005, 12(2): 126–132]
- TAN J, YAO G Y, LIU Z G, *et al.* SSR information analysis and unigene functional annotation of transcriptome from *Argopecten irradians* “Bohaihong” and *Argopecten irradians concentricus*. *Genomics and Applied Biology*, 2018, 37(12): 5242–5250 [谭杰, 姚高友, 刘志刚, 等. 扇贝“渤海红”与墨西哥湾扇贝转录组SSR信息分析及unigene功能注释. *基因组学与应用生物学*, 2018, 37(12): 5242–5250]
- TEMNYKH S, DECLERCK G, LUKASHOVA A, *et al.* Computational and experimental analysis of microsatellites in rice (*Oryza sativa* L.): Frequency, length variation, transposon associations, and genetic marker potential. *Genome Research*, 2001, 11(8): 1441–1452
- TIAN J T, HOU Y, LIU Z H, *et al.* Genetic diversity of different

- populations of *Scapharca subcrenata* based on mitochondrial CO I gene. *Marine Sciences*, 2016, 40(1): 1-9 [田吉腾, 侯丫, 刘志鸿, 等. 基于线粒体 COI 基因的毛蚶群体遗传多样性. *海洋科学*, 2016, 40(1): 1-9]
- TOTH G, GASPARI Z, JURKA J. Microsatellites in different eukaryotic genomes: Survey and analysis. *Genome Research*, 2000, 10(7): 967-981
- WANG R C, WANG Z P. Science of marine shellfish culture. Qingdao: China Ocean University Press, 2008 [王如才, 王昭萍. 海水贝类养殖学. 青岛: 中国海洋大学出版社, 2008]
- WANG Q Z, ZHANG M, TENG W M, *et al.* A review: Research progress of biology in ark shell *Scapharca subcrenata* insight from aquaculture. *Journal of Dalian Ocean University*, 2015, 30(4): 437-443 [王庆志, 张明, 滕炜鸣, 等. 毛蚶养殖生物学研究进展. *大连海洋大学学报*, 2015, 30(4): 437-443]
- WANG Z L, DING Y, XU Y H, *et al.* Polymorphism of EST-SSRs from *Pinctada martensii* based on transcriptome datasets. *Oceanologia et Limnologia Sinica*, 2015, 46(3): 687-693 [王忠良, 丁燊, 许尤厚, 等. 基于转录组数据的马氏珠母贝 EST-SSR 位点的信息分析及其多态性检测. *海洋与湖沼*, 2015, 46(3): 687-693]
- WEI D W, LIAN Z Q, WU X D, *et al.* Preliminary screening of *silurus lanzhouensis* EST-SSR markers based on Clariidae and Siluridae ESTs. *Freshwater Fisheries*, 2015, 45(5): 19-23 [魏大为, 连总强, 吴旭东, 等. 基于胡鲇科和鲇科 ESTs 序列初步筛选兰州鲇 EST-SSR 分子标记. *淡水渔业*, 2015, 45(5): 19-23]
- WU X. Development of SSR and SNP for Onchidiidae based on transcriptome sequencing. Master's Thesis of Shanghai Ocean University, 2016 [吴欣. 基于转录组测序的石磺科贝类 SSR 和 SNP 两种分子标记的开发. 上海海洋大学硕士学位论文, 2016]
- XIA Y, LUO W, YUAN S Q, *et al.* Microsatellite development from genome skimming and transcriptome sequencing: Comparison of strategies and lessons from frog species. *BMC Genomics*, 2018, 19: 886
- XIONG G, WANG X Q, WANG P, *et al.* Development and genetic diversity analysis of *Babylonia lutosa* with EST-SSR markers. *Progress in Fishery Sciences*, 2020, 41(4): 117-124 [熊钢, 王晓清, 王佩, 等. 泥东风螺 EST-SSR 开发及其群体遗传多样性分析. *渔业科学进展*, 2020, 41(4): 117-124]
- YAN L L, QIN Y J, YAN X W, *et al.* Development of microsatellite markers in *Ruditapes philippinarum* using next-generation sequencing. *Acta Ecologica Sinica*, 2015, 35(5): 1573-1580 [闫路路, 秦艳杰, 闫喜武, 等. 基于转录组平台的蛤仔微卫星标记筛选. *生态学报*, 2015, 35(5): 1573-1580]
- YANG M, YU Y, ZHANG X J, *et al.* Development of microsatellite markers from the transcriptome sequences of Pacific white shrimp (*Litopenaus vannamei*). *Marine Sciences*, 2017, 41(2): 96-102 [杨铭, 于洋, 张晓军, 等. 基于转录组数据的凡纳滨对虾微卫星标记开发. *海洋科学*, 2017, 41(2): 96-102]
- YU A Q, SHI Y H, XU J B, *et al.* Characteristic analysis of microsatellites in selected *Coilia ectenes* using a transcriptome dataset. *Progress in Fishery Sciences*, 2019, 40(5): 101-109 [于爱清, 施永海, 徐嘉波, 等. 长江刀鲚选育群体转录组 EST-SSR 的分布特征分析. *渔业科学进展*, 2019, 40(5): 101-109]
- ZENG C, GAO Z X, LUO W, *et al.* Characteristics of microsatellites in blunt snout bream (*Megalobrama amblycephala*) EST sequences using 454 FLX. *Acta Hydrobiologica Sinica*, 2013, 37(5): 982-988 [曾聪, 高泽霞, 罗伟, 等. 基于 454 GS FLX 高通量测序的团头鲂 ESTs 中微卫星特征分析. *水生生物学报*, 2013, 37(5): 982-988]
- ZHAI L X, MAO H Q. Biological characteristics and artificial breeding techniques of *Scapharca subcrenata*. *Modern Agricultural Science and Technology*, 2010(4): 345-346 [翟林香, 毛红茜. 毛蚶的生物学特性及人工育苗技术. *现代农业科技*, 2010(4): 345-346]
- ZHAN A B, HU J J, HU X L, *et al.* Isolation and characterization of microsatellite markers for Zhikong scallop by screening SSR-enriched library. *Journal of Fisheries of China*, 2008, 32(3): 353-361 [战爱斌, 胡景杰, 胡晓丽, 等. 富集文库-菌落原位杂交法筛选栉孔扇贝的微卫星标记. *水产学报*, 2008, 32(3): 353-361]
- ZHANG J Y, HE M C, XIANG Z L, *et al.* Microsatellite screening and characterization analysis of *Sepia esculenta* based on full-length transcriptome sequencing. *Progress in Fishery Sciences*, 2020, 41(6): 149-155 [张金勇, 何暮春, 项子龙, 等. 基于全长转录组测序的金乌贼微卫星位点筛选与特征分析. *渔业科学进展*, 2020, 41(6): 149-155]
- ZHAO W, ZHANG L, BI J H. ISSR analyses of genetic diversity within and among five stocks of *Scapharca subcrenata* along Liaoning coast. *Journal of Fisheries of China*, 2011, 35(6): 854-862 [赵文, 张雷, 毕进红. 辽宁沿海 5 个毛蚶群体遗传多样性的 ISSR 分析. *水产学报*, 2011, 35(6): 854-862]



## Development and Evaluation of SSR Markers Based on Transcriptome Sequencing in *Scapharca kagoshimensis*

CHEN Limei<sup>1</sup>, LI Li<sup>2</sup>, SHI Xuwei<sup>1</sup>, QIN Yiming<sup>1</sup>, LIU Lihua<sup>1</sup>, GUO Yongjun<sup>1</sup><sup>①</sup>

(1. Tianjin Key Laboratory of Aqua-Ecology and Aquaculture, College of Fishery Science, Tianjin Agricultural University, Tianjin 300384, China; 2. Marine Biology Institute of Shandong Province, Qingdao, Shandong 266104, China)

**Abstract** *Scapharca kagoshimensis* is a marine shellfish of great economic value. In recent decades, populations of *S. kagoshimensis* have declined due to environmental destruction and overfishing. To enhance our understanding of the genetic diversity and population-level genetic structure of *S. kagoshimensis*, microsatellite loci were detected based on the transcriptome data of *S. kagoshimensis* using MISA software. A total of 3987 Single Sequence Repeats (SSRs) were identified from 35,555 unigenes and the frequency of their occurrence was 11.21%. The main types of repeats were dinucleotides and trinucleotides, which accounted for 58.06% and 19.04%, respectively. A total of 182 types of repeat motifs were classified in all SSRs, and AC/GT was the most abundant among dinucleotide repeats (45.70%). The repeat numbers of SSRs primarily ranged between five and seven, and the number of SSRs gradually decreased as repeat number increased. Motif length was predominantly between 12 and 29 bp, and the SSR polymorphism level was above moderate. Among the 60 designed primer pairs, 14 pairs proved to be polymorphic microsatellite markers and were amplified in 30 wild individuals sampled from Weifang in Shandong Province. The results showed that the average number of alleles ( $N_a$ ), average observed heterozygosity ( $H_o$ ), average expected heterozygosity ( $H_e$ ), and polymorphism information content (PIC) were 15.4, 0.682, 0.852, and 0.817, respectively. All 14 loci were highly polymorphic (PIC  $\geq 0.5$ ). After Bonferroni correction, seven of the 14 loci deviated significantly from the Hardy-Weinberg equilibrium ( $P < 0.05$ ). These results indicate that it is feasible to develop microsatellite markers based on the *S. kagoshimensis* transcriptome. The polymorphic microsatellite loci obtained in this study will facilitate further studies on population genetic management, genetic mapping, and molecular assisted breeding of *S. kagoshimensis*.

**Key words** *Scapharca kagoshimensis*; Transcriptome; SSR; Genetic diversity

① Corresponding author: GUO Yongjun, E-mail: guoyongjun@tjau.edu.cn