

文章编号:1000-0615(2008)05-674-10

样本容量对养殖群体内主要遗传结构分析参数的影响

鲁翠云¹, 金万昆², 孙效文¹, 李大宇³, 朱晓东⁴, 马海涛³,
于东梅³, 杨建新²

(1. 中国水产科学研究员黑龙江水产研究所, 黑龙江 哈尔滨 150070;

2. 国家级天津市换新水产良种场, 天津 301500;

3. 大连水产学院生命科学与技术学院, 辽宁 大连 116023;

4. 上海海洋大学水产与生命学院, 上海 200090)

摘要: 用微卫星标记评估养殖群体的遗传结构, 设置8个取样量梯度, 5个标记量梯度, 统计分析了各遗传参数的变化。结果: 群体等位基因数(A)随样本量的增加而呈现上升趋势, 但上升幅度由0.6667~0.9615($10 < n < 30$)下降到0.2308~0.3333($30 < n < 50$); 46.15%的位点出现高频率等位基因即等位基因纯合化趋势, 这可能与养殖群体长期受选择压力的影响有关。等位基因纯合化使多数位点的有效等位基因数(A_e)、期望杂合度(H_o)及多态信息含量(PIC)。在样本量大于40时, 在多个标记量梯度其变动范围均变小, 并出现平台现象。样本量、微卫星标记的数量和多态性水平对群体遗传结构均有较大的影响, 建议在用微卫星进行群体遗传评估时, 标记的数量在20个以上, 样本容量大小在40~50之间较为适合。

关键词: 样本容量; 养殖群体; 遗传结构; 参数

中图分类号: S917 **文献标识码:** A

对自然群体和养殖群体进行遗传结构的评估是现阶段水产动物遗传育种的重点, 在对群体进行遗传评价时, 样本的采集成为正确评估群体遗传结构的先决条件, 样本量的大小以及分子标记的选用直接影响评估的效率。样本量太少, 得出的结论缺乏说服力, 太多则容易造成浪费, 而且很多时候, 选择样本量过大为人为、物力所不能及, 尤其对于一些保护鱼类, 大样本的采集是不可能的, 因此确定满足实验需求的最低样本量是合理的实验设计所必需考虑的问题。

对于不同标记样本量的选择已经有一些报道, 刘丽等^[1]、周泽扬等^[2]、周艺彪等^[3-4]分别对RAPD、AFLP分析群体遗传结构的最佳采样量及位点数进行了研究。微卫星分子标记由于开发

费用较高, 在水产动物方面的开发及利用远落后于陆生生物, 所以分子标记的选用带有很大的随机性, 样本量大小的确定也没有一致的标准。使用基因座位的等位基因数与合理的样本量之间的关系是非线性关系, 计算也比较复杂。本文用26个微卫星分子标记对208个个体的镜鲤养殖群体的遗传结构进行评估, 分不同梯度的样本量、标记数量来探索样本量的大小和标记量的大小对群体遗传参数的影响, 以期在水产动物种质评估等研究中样本量和标记量的选择提供参考。

1 材料与方法

1.1 实验材料

镜鲤采自国家级天津市换新水产良种场, 随

收稿日期: 2007-08-16

资助项目: 农业部“引进国际先进农业科学技术”计划资助[2006-G55(A)]

作者简介: 鲁翠云(1980-), 女, 河北邢台人, 硕士, 研究实习员, 主要从事水产动物基因遗传育种研究。E-mail: xtlcy_123@163.com

通讯作者: 孙效文, Tel: 0451-84842646; E-mail: sunxw2002@163.com

机取样 200 尾,雌雄各 100 尾,年龄在 3~4 龄,体重 4~7 kg。随机抽样设置 8 个样本量梯度,依次为 10、20、30、40、50、100、150、200 个。

微卫星分子标记为本实验室通过磁珠富集的方法开发的鲤鱼基因组微卫星^[5],随机选用 26 个标记,由上海生工生物工程公司合成,采用随机数的方法设置 5 个分子标记梯度,依次为 3、8、14、20、26 个微卫星标记,检测不同标记量对遗传参数的影响。

1.2 DNA 提取

取镜鲤鳍条约 0.2 g,加入裂解液 400 μL (0.5%十二烷基肌氨酸钠,200 $\mu\text{g}\cdot\text{mL}^{-1}$ 蛋白酶 K,10 $\text{mmol}\cdot\text{L}^{-1}$ EDTA (pH 8.0)),50 $^{\circ}\text{C}$ 消化 2 h,用有机抽提法提取基因组 DNA,无水乙醇沉淀回收,溶于 200 μL TE,取少量稀释 20 倍,4 $^{\circ}\text{C}$ 保存备用。

1.3 PCR 扩增及检测

建立 25 μL 反应体系,其中包括 PCR 混合缓冲液 18 μL ,微卫星上下游引物(10 $\mu\text{mol}\cdot\text{L}^{-1}$)各 0.5 μL ,*Taq* DNA 聚合酶 1 U,DNA 模板(50 ng) 1 μL ,补无菌水到 25 μL 。反应程序为 94 $^{\circ}\text{C}$ 预变性 3 min;94 $^{\circ}\text{C}$ 变性 30 s,50~54 $^{\circ}\text{C}$ 复性 30 s,72 $^{\circ}\text{C}$ 延伸 30 s,35 个循环;最后 72 $^{\circ}\text{C}$ 延伸 5 min。扩增产物用 2%的琼脂糖凝胶电泳检测,用 Gel-Pro Analyzer 4.5 软件分析电泳条带的大小。

1.4 数据处理

微卫星是共显性遗传,可从琼脂糖凝胶电泳图谱上直接判断出个体的基因型,用 Excel2003 统计不同取样量在不同标记水平的等位基因数(number

of alleles, *A*)、有效等位基因数(effective number of alleles, *A_e*)、观察杂合度(observed heterozygosity, *H_o*)、期望杂合度(expected heterozygosity, *H_e*)及各位点的多态信息含量(polymorphism information content, *PIC*),比较不同取样量对遗传多样性参数的影响。

2 结果

2.1 样本容量对平均等位基因数及有效等位基因数的影响

所选用的微卫星分子标记在群体内表现出不同程度的多态性,各位点检测到的等位基因数在 2~8 个,有效等位基因数 1.31~6.40 个,片段大小在 109~400 bp。采用随机抽样的方法设置 8 个样本容量梯度,统计各微卫星位点在不同样本容量所检测到的等位基因数及有效等位基因数。

等位基因的变化:各位点等位基因数受样本量的影响不同,5 个微卫星位点所检测到的等位基因数量在各样本量梯度不发生变化,占全部微卫星位点的 19.23%,16 个微卫星位点在样本量小于 30 时随样本量的增加等位基因数量明显增加,占 61.54%,而其中 87.5%(14 个)的标记在样本量大于 30 到达 40 时,等位基因数不再随着样本量的增加而增加。在 5 个标记量梯度(3,8,14,20,26),平均等位基因数均随着样本量的增加呈现上升趋势(图 1),但其上升幅度由 0.6667~0.9615(10 < n < 30)下降到 0.2308~0.3333(30 < n < 50)。

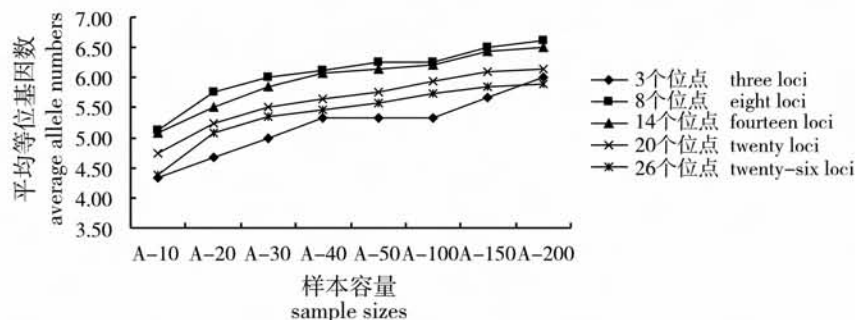


图 1 不同标记数量平均等位基因数随样本量的变化

Fig. 1 The change curve of average allele numbers along with sample sizes in different marker numbers

有效等位基因的变化:不同的微卫星位点有效等位基因数受样本容量的影响也不同,呈现 3 种变化规律,第一,有效等位基因数随样本容量的增大

变化很小,此类位点的特征是有有效等位基因数偏小,具有 1~2 个主效基因,在等位基因分布中占有 0.5 以上的基因频率,占所用标记的 46.15%(图 2-

HLJ041);第二,有效等位基因数随样本容量的增大而变化较大,主要发生在样本量为10~30阶段,并在样本量达到50以上波动变小,占所用标记的34.62%(图2-HLJ400);第三,有效等位基因数随样本容量的增大而变化较大,此类位点有效等位基因数较大,等位基因分布频率较均匀,随着样本容量的增大,稀有基因被检出,致使有效等位基因数随样本量的增大而增大或减小,必须在样本量达到

一个较大的水平才能趋于平衡,仅占所用标记的19.23%。而不同标记量的群体平均有效等位基因数则呈现较为一致的变化曲线,均随着样本量增加而减小,但在不同的取样量区间变化幅度有很大的差别,样本量在10~30之间平均等位基因数锐减,降幅为0.2685~0.0794,而样本量40~200的降幅仅为0.1077~0.0809,变化幅度相对较小,曲线逐渐趋于平稳(图3)。

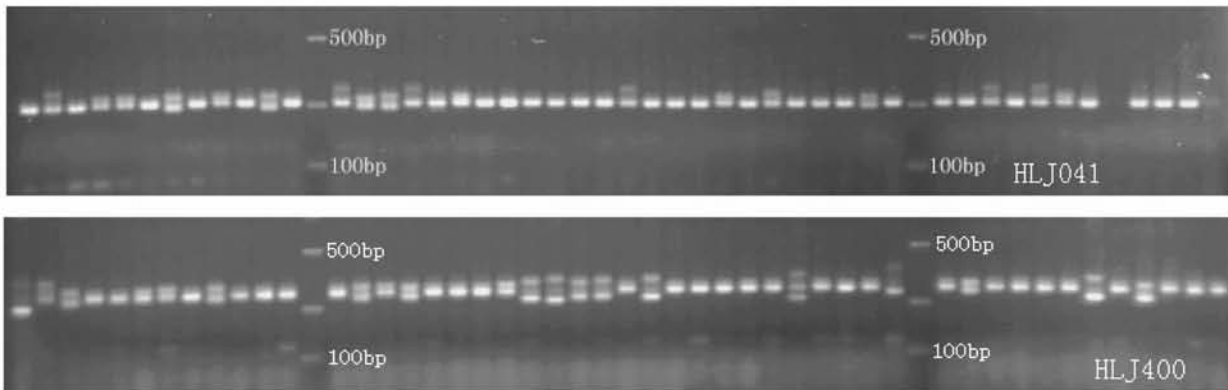


图2 微卫星位点 HLJ041、HLJ400 在部分样本的扩增结果

Fig. 2 Amplified result of HLJ041 and HLJ400 in partial samples

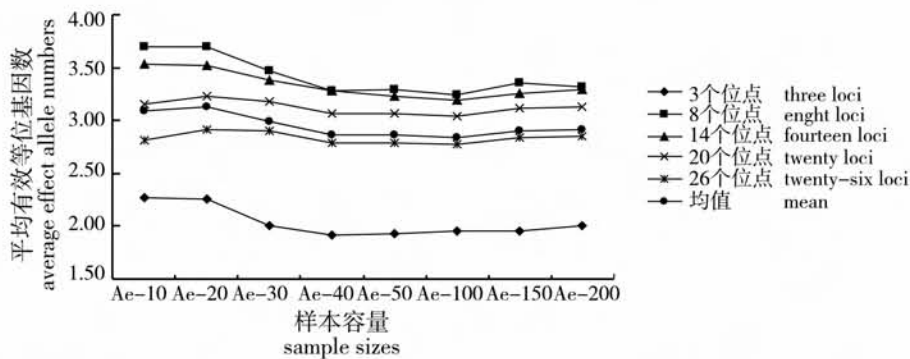


图3 不同标记数量平均有效等位基因数随样本量的变化

Fig. 3 The change curve of average effect allele numbers along with sample sizes in different marker numbers

平均有效等位基因的变化:相对于样本量对群体平均等位基因数的影响,标记的数量及其多态性对多样性参数的影响更为显著,如图3所示,不同标记及标记量的群体平均有效等位基因数相差达到1.2961~1.4743个,方差分析结果显示由于选用标记多态性及标记量所造成的平均等位基因数间差异极显著($P < 0.01$),而标记量较少的3个和8个标记,由于选用的标记多态性不同形成了统计量的两个极端,20个和26个标记量的变化曲线则接近于统计的平均值,尤其是当标记量

达到26个,样本量大于30时有效等位基因数接近于平均值(图3),能够较为客观的反映群体的实际有效等位基因数。

随机抽取相同标记量的不同标记的有效等位基因数变化:随机分5次抽取相同数量的标记,检验不同多态性水平的标记对遗传多样性统计参数的影响。随机5次抽取8个和20个标记统计的群体有效等位基因数,方差分析结果均差异极显著($P < 0.01$),如图4和图5所示,相对于5次随机抽取统计的平均值,8个标记的统计结果偏离较

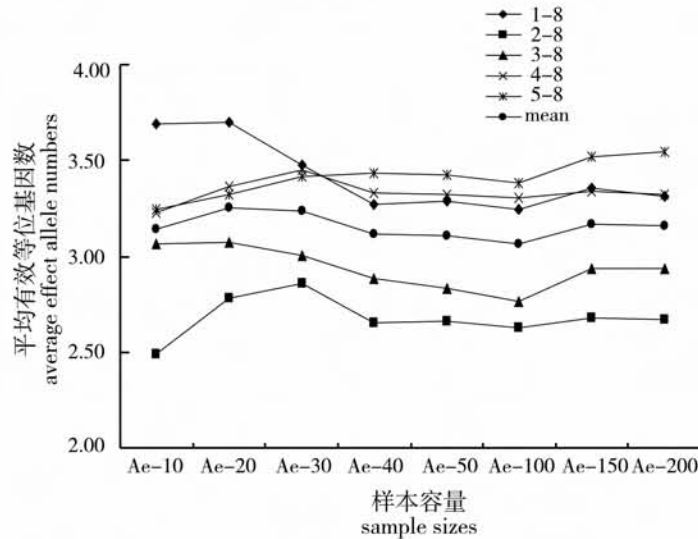


图4 5次随机抽取8个标记平均有效等位基因数随样本量的变化

Fig. 4 The change curve of average effect allele numbers along with sample sizes in 5 times randomly selected 8 markers

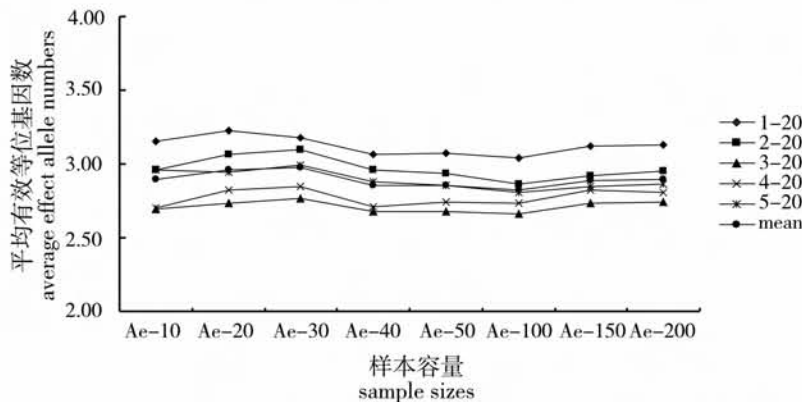


图5 5次随机抽取20个标记平均有效等位基因数随样本容量的变化

Fig. 5 The change curve of average effect allele numbers along with sample sizes in 5 times randomly selected 20 markers

大,5次随机抽取的平均方差为0.0974,20个标记的平均方差为0.0221。在标记量较少(8个)的情况下,采用不同的分子标记在样本容量较小的情况下(10~30)致使统计结果出现非常大的偏差,而在样本量达到40以后,偏差减小,变化曲线也较为一致;在标记量达到20个时,不同标记量抽样间的差异较小,而且随着样本量的增加变化曲线一致,即使在样本量较小的情况下(10~30),也能够反映群体的遗传结构。各标记量水平不同取样量区间的平均等位基因数及有效等位基因数的变化幅度见表1。

2.2 样本容量对杂合度的影响

基因杂合度也称为基因多样性,一般认为它

是度量群体遗传变异的一个最适参数。26个位点的观察杂合度差异较大,在0.2050~1.0000之间,其中61.54%(16个)的位点观察杂合度小于0.5,而5个位点的观察杂合度大于0.8。各微卫星位点随样本量的变化观察杂合度的变化并不十分明显,只是在位点非常少的情况下(3个位点),由于标记的选用所引起的偏差较大(图6)。期望杂合度在样本量在10~30之间变动幅度较大,达到0.0305。不同标记数量的平均杂合度随样本量变化趋势基本一致,尤其在标记数大于8时(图7);期望杂合度的变化与有效等位基因的变化趋势相似:样本量少于30时,杂合度变动较大,最小

变幅为 0.0036, 最大变幅为 0.0305; 样本量超过 40 时期望杂合度变化变小, 变幅仅为 0.0061~0.0155(表 2); 不同标记量由于选择标记的多态性不同, 杂合度变化较大(图 7)。

表 1 不同取样量范围在不同标记量水平平均等位基因数及有效等位基因数的变动幅度

Tab. 1 The change extent of average allele numbers and average effect alleles numbers in different sample range of different markers numbers

标记量 loci numbers	A				Ae			
	10~30	30~50	50~100	100~200	10~30	30~50	50~100	100~200
3	0.6667	0.3333	0.0000	0.6667	0.2658	0.0752	0.0265	0.0527
8	0.8750	0.2500	0.0000	0.3750	0.2279	0.1827	0.0441	0.0703
14	0.7857	0.2857	0.0714	0.2857	0.1537	0.1455	0.0482	0.1013
20	0.7500	0.2500	0.2000	0.2000	0.0794	0.1032	0.0322	0.0905
26	0.9615	0.2308	0.1538	0.1538	0.0908	0.1136	0.0191	0.0809
mean	0.8078	0.2700	0.0851	0.3362	0.1635	0.1240	0.0234	0.0791

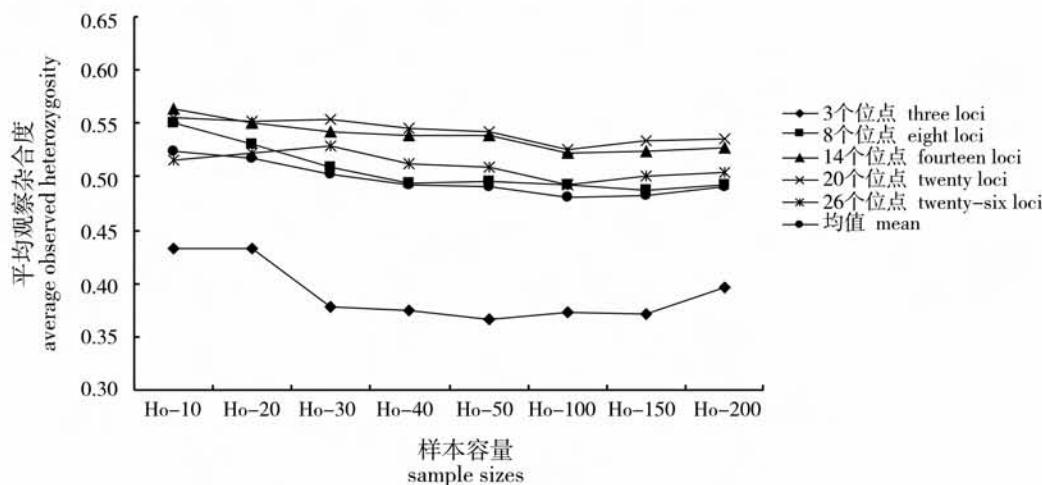


图 6 不同标记数量平均观察杂合度随样本量的变化

Fig. 6 The change curve of average observed heterozygosity along with sample sizes in different marker numbers

表 2 不同取样量范围在不同标记量水平平均观察杂合度及期望杂合度的变动幅度

Tab. 2 The change extent of average allele numbers and average effect alleles numbers in different sample range of different markers numbers

标记量 loci numbers	Ho				He			
	10~30	30~50	50~100	100~200	10~30	30~50	50~100	100~200
3	0.0556	0.0111	0.0067	0.0233	0.0252	0.0061	0.0067	0.0216
8	0.0417	0.0133	0.0025	0.0006	0.0305	0.0074	0.0068	0.0131
14	0.0214	0.0043	0.0164	0.0057	0.0235	0.0061	0.0068	0.0136
20	0.0017	0.0113	0.0160	0.0098	0.0036	0.0111	0.0049	0.0127
26	0.0128	0.0197	0.0158	0.0110	0.0260	0.0155	0.0050	0.0124
mean	0.0215	0.0120	0.0088	0.0098	0.0099	0.0092	0.0061	0.0147

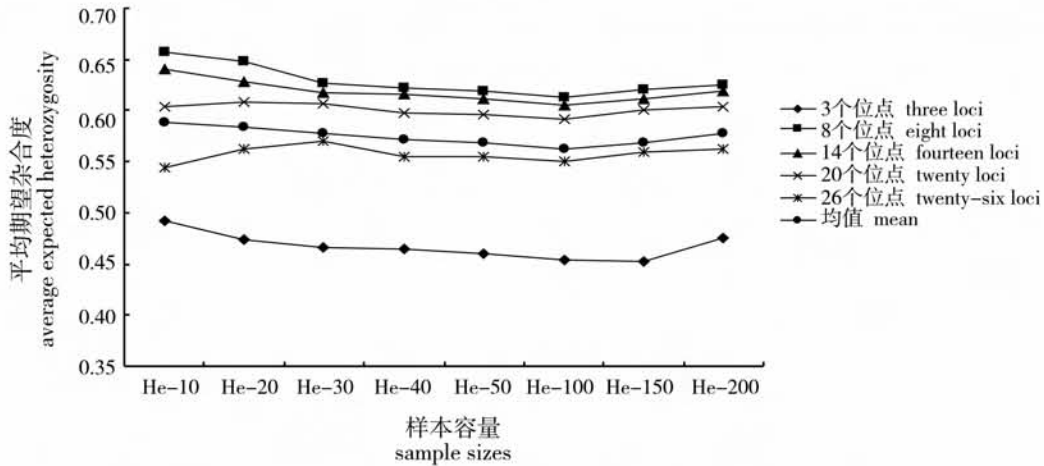


图7 不同标记数量平均期望杂合度随样本量的变化

Fig. 7 The change curve of average expected heterozygosity along with sample sizes in different marker numbers

2.3 样本容量对位点多态信息含量的影响

所检测的26个微卫星位点在镜鲤养殖群体的多态信息含量为0.1901~0.8238,依据Bostein^[6]提出的评价标准,15个位点为高度多态标记,占所用标记的57.69%;9个为中度多态标记,占34.62%,各位点多态信息含量受样本量影响的程度不同,69.23%的分子标记在样本量小于30时变化较剧烈,而73.77%的标记在样本量大于40时,多态信息含量的变化趋于平稳。不同标记数量的平均多态信息含量随样本量的变化也有差异,但是在样本量大于30的情况下变化趋于平稳。同样的样本量不同标记数量的平均多态信息含量差异较大,可能跟所选用的标记的多态性有关,如图8可见,在标记数量少(8个)的情况下,

可能多态性标记占的比例较高,得出的平均多态信息含量在一个较高的水平;而随着标记数量的增加,一些较低度多态性标记的引入,使平均多态信息含量有所降低。

随机分5次抽取相同数量标记的群体多态信息含量统计结果显示,在总体标记数量固定的情况下,标记量较少(8个),样本容量也较少时(10~30),PIC值出现较大波动(图9);而在标记量达到20个时,不同标记量抽样间的PIC值差异较小,随着样本量的增加变化曲线趋于一致,即使在样本量较小的情况下(10~30),也能够反映群体的遗传结构,但是在样本量达到40个以上时,PIC值基本保持不变(图10)。

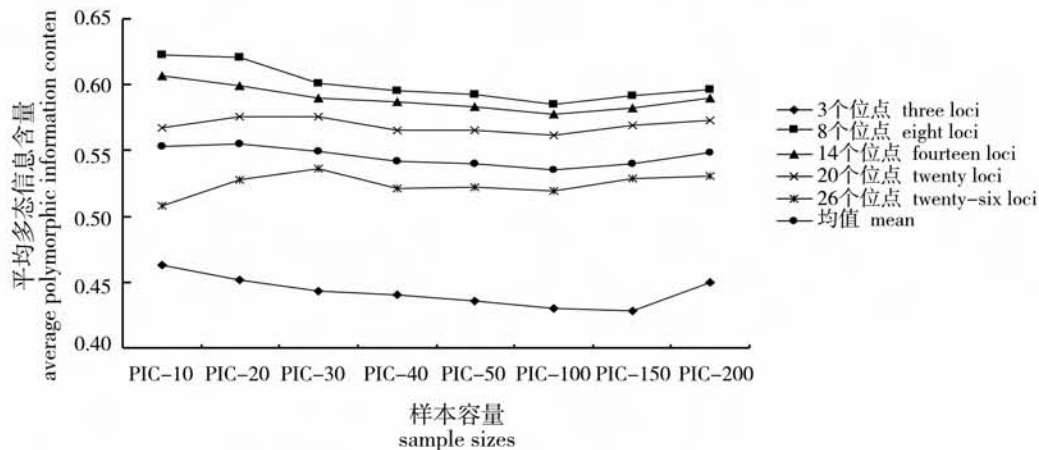


图8 不同标记数量平均多态信息含量随样本量的变化

Fig. 8 The change curve of population average polymorphic information content along with sample sizes in different marker numbers

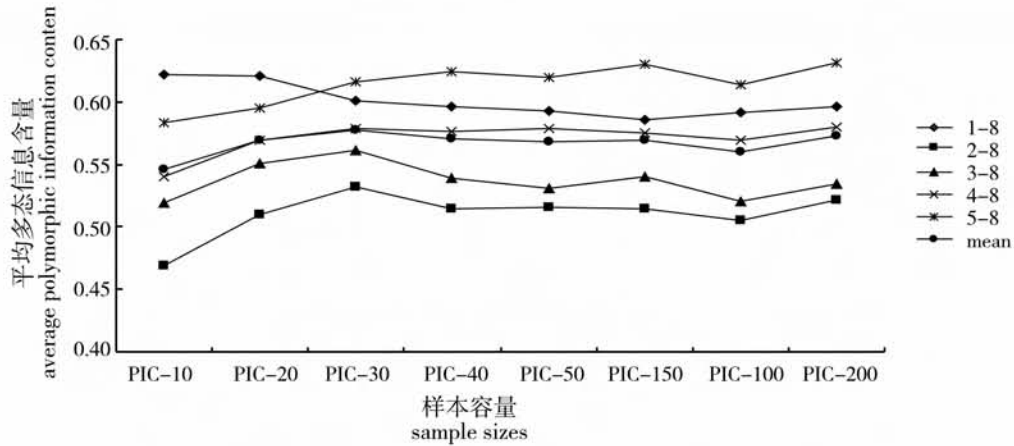


图 9 5次随机抽取 8 个标记群体平均多态信息含量随样本容量的变化

Fig. 9 The change curve of population average polymorphic information content along with sample sizes in five times randomly selected 8 markers

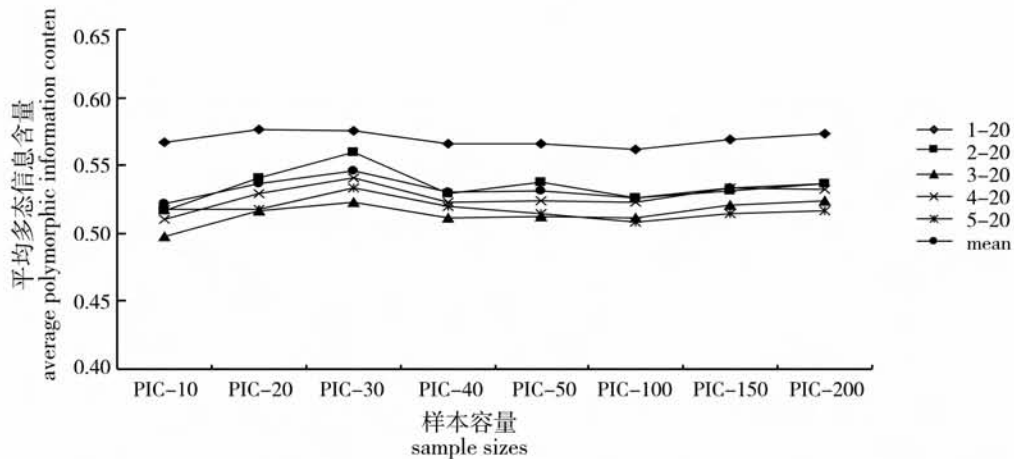


图 10 5次随机抽取 20 个标记群体平均多态信息含量随样本容量的变化

Fig. 10 The change curve of population average polymorphic information content along with sample sizes in five times randomly selected 20 markers

3 讨论

微卫星分子标记是一类由 1~6 个碱基组成的串联重复 DNA 序列,随机分布于绝大多数真核生物基因组中,数量庞大,每隔 10~50 kb 就存在 1 个微卫星^[7]。由于其重复性好、呈共显性遗传等优点已广泛应用于遗传图谱构建、群体遗传结构分析及种质资源鉴定等多个领域。微卫星分子标记的应用受限于开发费用高,有足够多的标记的物种还不是很多,尤其是水产动物,多数没有足够的微卫星标记可供选择。因此,在微卫星标记及其位点数量的选用上具有很大的局限性和随机性,样本数量的取用也没有统一的标准,影响了

研究结果的可靠性和可比性^[8-10]。本研究利用本室克隆的大量鲤微卫星标记^[5]从其中选取 26 个扩增效果好并呈多态性的标记,设计不同样本量及标记量对遗传参数值的影响。

等位基因数随着样本量的增大而增加,没有出现常见的平台现象,样本量 200 时等位基因数达到最大值,虽然取样量大于 30 时较低于 30 时等位基因增加的趋势下降,但仍没有出现平台的迹象,说明选育群体也还保留一定的稀有基因,也说明这个群体奠基群体较大,经长期较大压力选择,仍保留奠基群体遗传多样性的特点。

有效等位基因数随着样本量的增大而增加,但超过 40 个样本时基本不再变化,出现常见的平

台现象。随着样本量的增加虽然稀有等位基因仍然增加,但这些稀有等位基因频率过低,对有效等位基因数增减基本没有影响。

其他遗传参数如观察杂合度、期望杂合度、多态信息含量在样本量大于 40 时,不管标记是多少其变动范围都变小且都出现平台现象,原因是这些参数都是来源于等位基因频率的变动结果,由于等位基因频率在样本超过 40 时的变动变小,因此这些参数也趋于平衡;另外,即使样本量超过 40 时,上述参数在不同标记时的变化比较大,如 8 个标记与 20 个标记时有显著差异。另外,在 8 个标记和 20 个标记下的 5 次随机取样证实,8 个标记时取样批次间遗传参数变动较大,20 个标记时变动较小,这两项分析结果都支持微卫星标记进行群体遗传分析时,使用 20 个以上标记更能代表群体的真实遗传结构。

很多学者从统计学及实验分析的角度得出了多种标记用于群体遗传分析的最适样本量和位点数。刘丽等^[1]通过分析认为 RAPD 标记在样本数达到 20 且位点数达到 70 才能保证结果的可靠性;周艺彪等^[3]认为在用 AFLP 进行群体遗传多样性分析时,每个种群内的样本量最好不低于 30 个,位点数最好不低于 338 个;闫路娜等^[11]分析了 8 个微卫星位点样本量对各种遗传多样性指标的影响,认为样本大小与所观测到的等位基因数呈正相关,而与期望杂合度无显著相关,对于种群遗传和分子生态学研究,30~50 个个体是微卫星 DNA 分析所需的最小样本量;而包文斌等^[12]用 29 个微卫星标记的研究结果则认为样本量在 20~25 较为适宜。Nei^[13-14]早期的研究也指出,在进行群体内平均期望杂合度及遗传距离的评价时,位点数量的增加比样本数量的增加更能够得到可靠的结果,当样本容量偏小时,须观测 50 个以上的基因位点;而当样本容量达到 50,检测的基因位点又在 25 个以上时,抽样样本偏差较小,检测值就越接近真实值。张继全等^[15-16]从计算机模拟和实例分析两种方法研究了 Nei 氏标准遗传距离估测精度的影响因素,指出位点数较少时遗传距离估测精度较低,可靠性不高,并且随个体数增加,遗传距离估测精度增高。本研究比较了样本容量及标记数量对群体有效等位基因、多态信息含量等参数的影响,证实在取样量大于 40 时,标记量在 20 个以上时群体的遗传参数变化

较小。

主效等位基因对以基因分布为基础的统计参数如有效等位基因、期望杂合度等的影响较大,而微效基因的影响则小得多,经过选择的养殖群体尤其明显。所以在评估养殖群体的遗传结构时可以采用统计学最小样本量 30 个,在条件允许的情况下达到 40 个以上为宜;而野生群体理论上多数位点的等位基因分布较平均,取样数量应适当增加以便更为全面的了解等位基因的分布,但是考虑到野生群体取样相对困难,在条件允许的情况下应当取样数量达到 30 尾。另外在取样时应当考虑所取样品的生活习性和分布,例如是否具有洄游、群居等生活习性,并选取合适的取样时间和取样点,以便取样能够较全面的代表群体。在分子标记的选用方面,我们的统计结果与 Nei^[14]的结论相符,证实较多分子标记对群体的评估更为全面,偏差也较小,在分子标记数量小于 10 时,会因为所选用标记的多态性过高或过低,对群体的评价出现人为选择的偏差,而在标记数量大于 20 时这种偏差则小得多,而 20 个标记量相对于整个基因组来说也是很有限的。所以在用微卫星标记进行水产动物群体分析时,应根据所研究种类的特点采样 30~50 尾,并尽可能多的选用分子标记(20 以上),对群体的遗传结构进行较为全面的评价。

根据本研究结果,建议在用微卫星进行群体遗传结构分析时,取样量大于 40、标记量大于 20。

参考文献:

- [1] 刘丽,刘楚吾,曾健民. RAPD 分析中最适样本量和位点数的研究[J]. 湛江海洋大学学报,2005,25(4):1-4
- [2] 周泽扬,夏庆友,鲁成. 分子系统学研究中分子位点数与遗传差异信息可靠性的关系[J]. 遗传,1998,20(5):12-15.
- [3] 周艺彪,姜庆五,赵根明,等. 湖北钉螺扩增片段长度多态性分子标记遗传多样性研究的合理样本量与分子位点数[J]. 中华流行病学杂志,2005,26(12):951-954
- [4] 周艺彪,姜庆五,赵根明,等. 湖北钉螺数量性状研究 I. 遗传变异信息的可靠性与样本量的关系[J]. 中国血吸虫病防治杂志,2005,17(2):81-85.
- [5] 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[J]. 中国水产科

- 学, 2005, 12(2): 126-132.
- [6] Bostein D, White R L, Skolnick M, *et al.* Construction of a genetic linkage map in man using restriction fragment length polymorphisms[J]. *Am J Hum Genet*, 1980, 32: 317-331.
- [7] 何平. 真核动物中微卫星及其应用[J]. *遗传*, 1998, 20(4): 42~47.
- [8] 常玉梅, 孙效文, 梁利群. 中国鲤几个代表种群基因组 DNA 遗传多样性分析[J]. *水产学报*, 2004, 28(5): 481-486.
- [9] 胡雪松, 李池陶, 马波, 等. 3 个德国镜鲤养殖群体遗传变异的微卫星分析[J]. *水产学报*, 2007, 31(5): 575-582.
- [10] 全迎春, 李大宇, 曹鼎辰, 等. 微卫星 DNA 标记探讨镜鲤的种群结构与遗传变异[J]. *遗传*, 2006, 28(12): 1541-1548.
- [11] 闫路娜, 张德兴. 种群微卫星 DNA 分析中样本量对各种遗传多样性度量指标的影响[J]. *动物学报*, 2004, 50(2): 279-290.
- [12] 包文斌, 束婧婷, 许盛海, 等. 样本量和性比对微卫星分析中群体遗传多样性指标的影响[J]. *中国畜牧杂志*, 2007, 43(1): 6-9.
- [13] Nei M, Roychoudhury A K. Sampling variances of heterozygosity and genetic distance[J]. *Genetics*, 1974, 76: 379-390.
- [14] Nei M. Estimation of average heterozygosity and genetic distance from a small number of individuals[J]. *Genetics*, 1978, 89: 583-590.
- [15] 张继全, 邵春荣, 王毓英, 等. Nei 氏标准遗传距离的估测精度[J]. *畜牧兽医学报*, 1998, 29(1): 27-32.
- [16] 张继全, 邵春荣, 王毓英, 等. 多位点基因型遗传距离的估测精度[J]. *畜牧兽医学报*, 1998, 29(2): 128-131.

Effects of sample size on various genetic structure parameter in cultured population genetic study

LU Cui-yun¹, JIN Wan-kun², SUN Xiao-wen¹, LI Da-yu³, ZHU Xiao-dong⁴,
MA Hai-Tao³, YU Dong-mei³, YANG Jian-xin²

(1. Heilongjiang Fisheries Research Institute, Chinese Academy of Fishery Sciences, Harbin 150070, China;

2. Tianjin Huanxin Excellent Fisheries Seed Farm, Tianjin 301500, China;

3. College of Life Science and Technology, Dalian Fisheries University, Dalian 116023, China;

4. College of Fisheries and Life, Shanghai Ocean University, Shanghai 200090, China)

Abstract: This study examines the effects of sample size on various genetic structure parameters. The population genetic structure of cultured mirror carp (*Cyprinus carpio*) was estimated using microsatellite markers. 8 different sample sizes and 5 marker numbers were set up to analyze most genetic parameters including population alleles distribution, allele numbers (A), effect allele numbers (A_e), observed heterozygosity (H_o), expected heterozygosity (H_e) and polymorphism information content (PIC). As a result, population allele numbers rise dramatically along with sample sizes from 10 to 40 and then held a steady increase trends. Meanwhile 46.15% of loci showed high frequency allele. The results indicate that cultured population show predominant alleles in most loci because of selection pressure. Due to many homozygosity, the change of effect allele numbers, expected heterozygosity and polymorphism information content are little when the sample size is more than 40. In addition, the number of microsatellite markers and polymorphism level also have a little effect on the estimation of population genetic structure. So, we suggested that the numbers of markers is more than 20 and sample size is 40 to 50 when using microsatellite markers to estimate population genetic structure.

Key words: sample size; culture population; genetic structure; parameter