

文章编号: 1000-0615(2017)07-1083-13

DOI: 10.11964/jfc.20160410360

罗氏沼虾种群SSR分析中样本量及标记量 对遗传多样性指标的影响

戴习林^{1*}, 刘洁¹, 李晶晶¹, 江宗冰¹, 明磊¹, 蒋飞²

(1. 上海海洋大学水产与生命学院, 上海 201306;

2. 上海申漕特种水产开发公司, 上海 201516)

摘要: 利用60对微卫星分子标记分别对罗氏沼虾3个养殖群体进行遗传多样性分析, 通过比较样本量及标记量对遗传多样性指标的影响, 探讨最适宜的样本量及标记量。结果显示, 样本量和标记量的大小均对遗传多样性指数有较大的影响, 其中样本量与平均等位基因数和平均有效等位基因数呈高度正相关, 与杂合度和Nei氏遗传多样性指数分别呈中度、高度相关, 样本量为10~30时, 遗传参数变化差异显著, 样本量大于30, 变化差异不显著; 标记量与遗传参数也存在不同程度的相关性, 标记量为5~25时, 各遗传多样性指数变化有明显差异, 标记量大于25时, 各遗传参数变化较小, 多态信息含量不同的标记直接影响到群体遗传参数。性别对群体遗传多样性指标无显著性影响。研究表明, 应尽可能选择多态信息含量高的微卫星分子标记开展罗氏沼虾群体遗传多样性分析, 可不考虑样本性别, 样本量不宜小于30, 标记量不宜少于25。

关键词: 罗氏沼虾; 遗传多样性; 微卫星; 样本量

中图分类号: Q 785; S 968.2

文献标志码: A

微卫星标记(microsatellite)又称为简单重复序列(simple sequence repeats, SSR)或短串联重复序列。作为第二代分子遗传标记, 因其具有分布广、多态信息含量高、检测快速方便和呈共显性遗传等优点, 而被广泛应用于评价生物遗传多样性、群体进化、构建遗传连锁图谱、制作系统发生树、血缘关系鉴定以及核基因组研究和基因定位等研究^[1-3]。联合国粮农组织在管理动物遗传资源和其持续发展战略中, 也将SSR作为推荐的分析工具^[4]。

对群体进行遗传结构的评估是现阶段水产动物育种的重点内容, 在进行遗传评价时, 样本量的确定和标记量的选取是正确评估群体遗传结构的首要步骤, 其直接影响评估结果的可靠性。在群体遗传学中, 合理的取样策略不仅

要有代表性, 还要尽可能的包含较多的遗传变异, 样本量太多, 人力物力花费较大, 而且不利于一些物种保护, 也易造成浪费, 而样本量少则可靠性又较差。因此, 确定满足分析所需的最低样本量和标记量是必须要考虑的问题。关于最低样本量和标记量的选取在水产动物方面的研究甚少, 目前仅涉及少数鱼类^[5], 有关虾类的研究迄今未见报道, 对于微卫星标记量的选取和样本量的选用常依据宏观生态学理论和参考其他水产物种的数据。本实验设置不同的样本量梯度和标记量梯度对3个罗氏沼虾(*Macrobrachium rosenbergii*)养殖群体进行分析, 探讨遗传多样性研究中所需的最适宜样本量和标记量, 以为罗氏沼虾及其他虾类的种质评估等研究提供参考依据。

收稿日期: 2016-04-14 修回日期: 2016-05-23

资助项目: 上海市现代农业产业技术体系(虾类)专项[沪农科产字(2015)第5号]

通信作者: 戴习林, E-mail: xldai@shou.edu.cn

1 材料与方 法

1.1 实验材料

分别从上海(SH, 体长4.2~8.7 cm, 体质量约1.4~16.4 g)、浙江(ZJ, 体长3.8~7.9 cm, 体质量约1.1~12.1 g)、马来西亚子三代(MF₃, 体长4.1~11.2 cm, 体质量约1.4~51.0 g)3个养殖群体中各随机选取60尾罗氏沼虾, 采集其步足样品置于无水乙醇中固定, -20 °C保存备用。SH群体与ZJ群体和MF₃群体间的遗传距离分别为0.1794、0.3531, ZJ群体与MF₃群体间遗传距离为0.3441。

1.2 微卫星引物

60对微卫星引物由上海生工有限公司合成, 其中7对G-SSR引物(位点为SUG-bp8-103b, SUGbp8-137, Y1, Y6, Y7, Y42, Y44)为实验室开发筛选, 6对G-SSR引物(位点为Mr7-88, SUGbp8-101b, SUGbp8-101c, SUGbp8-103a, MRMB11, W18)筛选自文献[6](表1), 其余47对微卫星分子标记均为实验室以罗氏沼虾马来西亚野生群体为基础开发的EST-SSR引物。

1.3 微卫星DNA多态性检测

采用标准的酚、氯仿抽提法提取基因组DNA。PCR反应液总体积为10 μL, 包括2×Master-

mix(0.1U Taq Polymerase/uL, 500 μmol/L dNTP each, 20 mmol/L Tris-HCl pH8.3, 100 mmol/L KCl, 3 mmol/L MgCl₂, 天根生化科技有限公司) 5 μL, 上、下游引物(10 μmol/L)各0.3 uL, 模板DNA0.4 μL, 用ddH₂O补足体积。PCR反应程序为94 °C预变性4 min, 94 °C变性30 s, 退火温度复性45 s, 72 °C延伸1 min, 35个循环, 最后72 °C延伸4 min。

PCR产物4 °C保存并用8%的非变性聚丙烯酰胺凝胶电泳分离, 银染法检测电泳结果。以pBR-322DNA/*MspI*(天根生化科技有限公司)作为分子量标准, 用Gelpro analyzer 4.5软件完成微卫星DNA数据的读取和校正。

1.4 数据处理

3个养殖群体样本量梯度设计为6个: 10、20、30、40、50、60尾, 随机数字表法抽样设置(SSR标记量为60个), 分析样本量对遗传多样性指标的影响, 同法设计3个养殖群体样本的SSR标记量梯度: 5、10、15、20、25、30、35、40、45、50、55、60个(样本量为60个), 分析标记量对遗传多样性指标的影响。按标记量梯度随机获得3个群体的剩余样本量梯度, 组成样本量与标记量的72个交叉组合, 分别计算各交叉组合遗传多样性指数。

表 1 罗氏沼虾部分微卫星引物及引物信息

Tab. 1 Part of microsatellite primers and their information for *M. rosenbergii*

座位 locus	核心序列 core motif	引物序列 primer sequence(5'-3')	片段大小/bp size	退火温度/°C temperature
SUGbp8-103b	(CTG) ₇	F:ATTCGGTATCAGCTCTGC R:AGGTCATCACCCCTTTCCA	171~216	58
SUGbp8-137	(ATT) ₁₅	F:CGACTGGGTGGTATTTAT R:CGCTGACGTTTATTCTGT	300~349	55
Y1	(ATT) ₇	F:CATCAGCATTGGCAGTC R:CATTGGAGCCCTGAAC	130~205	55
Y6	(ATT) ₇	F:CATCAGCATTGGCAGTC R:AGCCCTTGAACCTGTTGTAT	124~191	52
Y7	(TG) ₁₁	F:ATGCCTGGAAAGAATGAG R:TTGTCTGAGCCTGAAACC	295~325	52
Y42	(AG) ₃₃	F:GGGCTCAAGAACGCTATG R:GACCCAATTACTGCTCAAAGG	240~267	50
Y44	(AG) ₁₀	F:GCGAGGAACACCCAAAG R:CGGAATCGTAATAGAGGC	158~211	54
Mr7-88	(CT) ₇	F:TATTACGATTCGTTGGCACA R:ATATTCTTTGTAGCGGCTGG	185~249	53
SUGbp8-101b	(CAT) ₂ (CTT) ₂ (TAAA) ₂	F:CCGTGGCTGAGAGTTTCATAC R:GGGGAATAAAGGACCTCTGG	207~238	60
SUGbp8-101c	(CAC) ₂	F:ATAGATGCCAACAGCACAGG R:GACAGTGAGGCACTTTACCG	252~276	60
SUGbp8-103a	(TCT) ₂ (ATA) ₂	F:AGGATTCTGCGAGGCTCTG R:CGTGTGTTCTTCATAGGCTTC	167~217	60
MRMB11	(CT) ₃₀	F:GACGCTGCCAAAAAGAAAAG R:ACCGTGCCATTAACCTCCAA	162~205	44
W18	(GTT) ₂ (TCT) ₂	F:GCAGCTGGAACGGATTTTAC R:AGATGCCGAAGTAGCTCCAG	240~260	50

使用PopGene(Version 3.2)软件统计微卫星基因座的等位基因数(observed number of alleles, N_a)、有效等位基因数(effective number of alleles, N_e)、观测杂合度(observed heterozygosity, H_o)、期望杂合度(expected heterozygosity, H_e)和Nei's遗传多样性指数(*Nei's genetic diversity*), 多态信息含量(polymorphism information content, PIC)的Botstein公式^[7]如下:

$$PIC = 1 - \left(\sum_{i=1}^n P_i^2 \right) - \left(\sum_{i=1}^n \sum_{j=i+1}^n 2P_i P_j \right)$$

式中, n 为某一位点上等位基因数; P_i 、 P_j 分别为第*i*和第*j*个等位基因在群体中的频率, $j=i+1$ 。

R语言3.2统计分析样本量、标记量与遗传多样性指标间的Pearson相关关系(correlation analysis)。方差分析样本量、标记量间的遗传多样性指数差异, Duncan氏法均值比较。*t*检验考察两性间遗传多样性指数差异。

2 结果

2.1 微卫星位点的扩增结果及遗传多样性

将60对微卫星引物在3个群体共180个个体中进行PCR扩增, 图1分别为标记36634在SH、

ZJ、MF₃ 3个群体中部分个体的扩增结果。各微卫星位点的遗传多样性见表2。

2.2 样本量对遗传多样性的影响

样本量对遗传多样性指标的影响 样本量越大, 检测到的 N_a 也越大(图2), N_a 与样本量之间呈高度正相关^[8](表3), 曲线拟合(curve estimation)等位基因数与样本量间的回归关系方程, 发现两者间呈显著的对数回归关系, 当样本量大于30时, N_a 、 N_e 逐渐趋于稳定。相关分析显示, 样本量与 H_o 无显著相关性, 与 H_e 和 H 呈正相关, 当样本量超过30, H_o 、 H_e 、 H 均逐步趋于稳定(图3)。以群体为重复数, 不同样本量下群体的遗传多样性均值比较结果, 当以 N_a 来衡量总体的遗传多样性指数时, 样本量40以上, 达到群体总体遗传多样性水平的95%; 以 N_e 、 H_e 和 H 来衡量时, 样本量30以上, 就已达到群体总体遗传水平的95%(表4)。

样本量对稀有等位基因的影响 合并3个群体统计分析稀有等位基因频率发现, 60对标记中有16对标记存在稀有等位基因(频率小于0.01)^[9]。当样本量达到30以上, 不同样本量的稀有等位基因频率差异不显著(表5)。

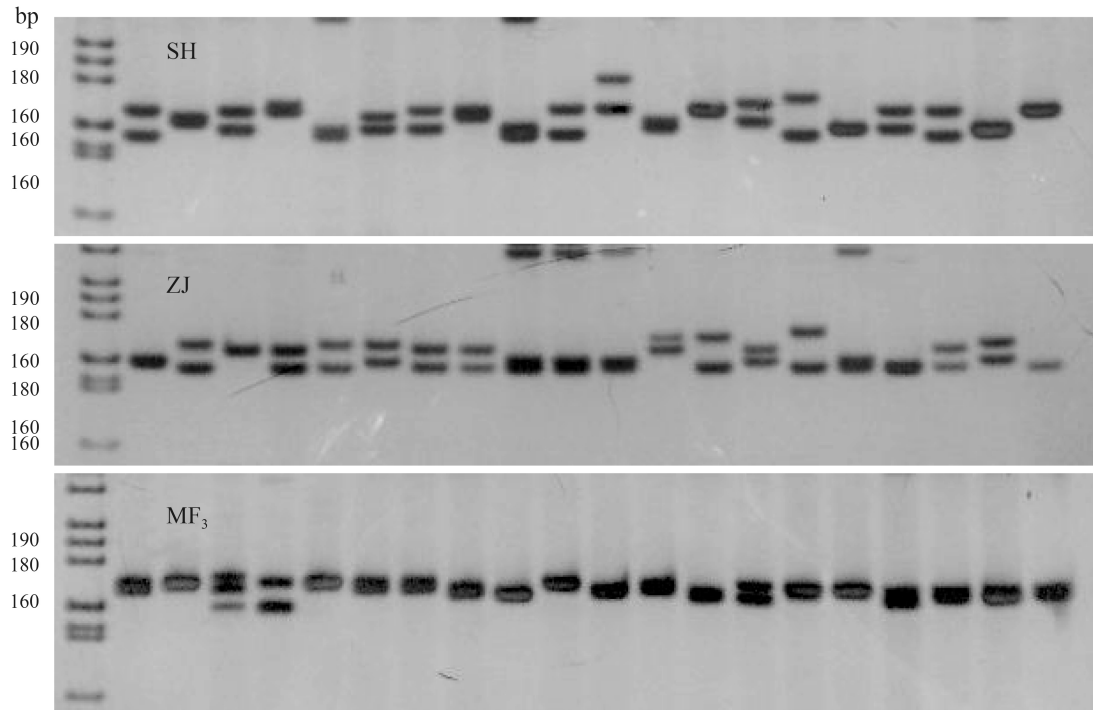


图1 微卫星标记36634在罗氏沼虾群体中部分个体扩增电泳图

Fig. 1 Amplification pattern of microsatellite locus 36634 in different populations of *M. rosenbergii*

表 2 罗氏沼虾60个微卫星位点的遗传多样性

Tab. 2 The genetic diversity for *M. rosenbergii* at 60 microsatellite loci

座位 locus	等位基因数 N_a	有效等位基因数 N_e	观测杂合度 H_o	期望杂合度 H_e	Nei's遗传多样性指数 H
27 463	2	1.9	0.31	0.46	0.46
15 331	3	1.9	0.49	0.49	0.48
16 813	6	1.7	0.40	0.43	0.43
28 780	3	1.3	0.20	0.22	0.22
6033	2	1.1	0.07	0.08	0.08
6221	5	2.1	0.49	0.53	0.53
245 505	2	1.1	0.09	0.11	0.10
270	5	2.3	0.34	0.58	0.57
20 823	5	2.6	0.28	0.62	0.62
55	4	2.9	0.17	0.66	0.65
713	7	3.4	0.27	0.71	0.71
25 718	5	4.0	0.54	0.75	0.75
25 992	5	3.6	0.51	0.73	0.72
9150	9	7.1	0.73	0.86	0.86
1773	4	2.8	0.09	0.64	0.64
2207	8	4.8	0.58	0.79	0.79
2343	4	2.7	0.32	0.63	0.63
11 375	4	2.1	0.16	0.53	0.53
13 540	5	2.6	0.40	0.62	0.62
17 493	4	1.7	0.31	0.40	0.40
20 398	2	1.4	0.29	0.31	0.31
Mr7-88	5	3.5	0.48	0.71	0.71
SUGbp8-103b	4	2.7	0.62	0.63	0.62
SUGbp8-101b	6	5.1	0.76	0.81	0.80
SUGbp8-101 c	5	3.5	0.46	0.72	0.72
SUGbp8-103a	11	2.9	0.13	0.65	0.65
MRMB11	6	3.8	0.46	0.74	0.74
SUGbp8-137	9	6.3	0.50	0.84	0.84
Y1	11	5.1	0.25	0.81	0.80
Y6	11	3.2	0.24	0.69	0.69
Y7	7	2.1	0.36	0.53	0.53
Y42	6	3.3	0.45	0.70	0.70
Y44	2	1.0	0.04	0.04	0.04
W18	3	1.6	0.15	0.38	0.38
2387	4	1.6	0.29	0.38	0.38

· 续表2 ·

座位 locus	等位基因数 N_a	有效等位基因数 N_e	观测杂合度 H_o	期望杂合度 H_e	Nei's遗传多样性指数 H
21 401	3	1.9	0.51	0.47	0.47
22 058	3	1.8	0.29	0.44	0.44
22 589	5	2.9	0.29	0.66	0.66
23 190	2	2.0	0.27	0.50	0.50
24 568	4	1.6	0.31	0.36	0.36
25 788	3	2.6	0.44	0.61	0.61
26 843	5	3.3	0.49	0.69	0.69
27 200	4	2.9	0.34	0.65	0.65
7352	5	3.3	0.24	0.70	0.70
32 714	3	1.7	0.07	0.40	0.40
34 797	3	1.2	0.07	0.20	0.20
36 634	11	6.0	0.49	0.83	0.83
25 484	2	1.5	0.24	0.32	0.32
25 791	5	3.1	0.62	0.68	0.68
25 218	2	1.6	0.08	0.37	0.37
22 375	6	2.0	0.20	0.50	0.50
22 860	8	3.2	0.31	0.69	0.69
24 694	4	2.7	0.59	0.63	0.62
750	3	2.1	0.32	0.53	0.53
7186	6	2.4	0.37	0.59	0.59
17 875	3	1.5	0.29	0.33	0.33
27 313	3	1.5	0.38	0.34	0.34
818	6	3.9	0.60	0.74	0.74
25 288	3	2.0	0.45	0.49	0.49
25 895	2	1.7	0.36	0.42	0.42

2.3 标记量对遗传多样性指标的影响

不同标记量与遗传多样性指数相关分析结果表明, 3个群体的遗传参数与标记量存在不同的相关关系(表6); 当标记数为5~25时, N_a 、 N_e 、 H_o 、 H_e 、 H 等遗传参数呈现上升趋势, 当标记数大于25时, 各遗传参数均趋于稳定(图4)。对3个群体不同标记量下群体遗传多样性指数均值比较发现, 当标记量大于20时, 各遗传多样性指数值已无显著性差异($P>0.05$)(表7)。

2.4 样本量与标记量的组合

分析样本量与标记量各梯度的交叉组合(横

坐标为样本量, 纵坐标为标记量)对3个群体遗传参数的影响, 从15个关系图发现, 3个群体间具有相同规律, 即样本量大于30, 标记量达25个以上, 各遗传指数逐渐趋于稳定(图5)。

2.5 多态信息含量不同的标记对遗传多样性指标的影响

SH群体中PIC大于0.5和小于0.5^[7]的标记数分别为20个和40个, ZJ群体中PIC大于0.5和小于0.5的标记数分别为25个和35个, MF₃群体中PIC大于0.5和小于0.5的标记数分别为13个和47个。由PIC>0.5标记获得的3个群体遗传多样性

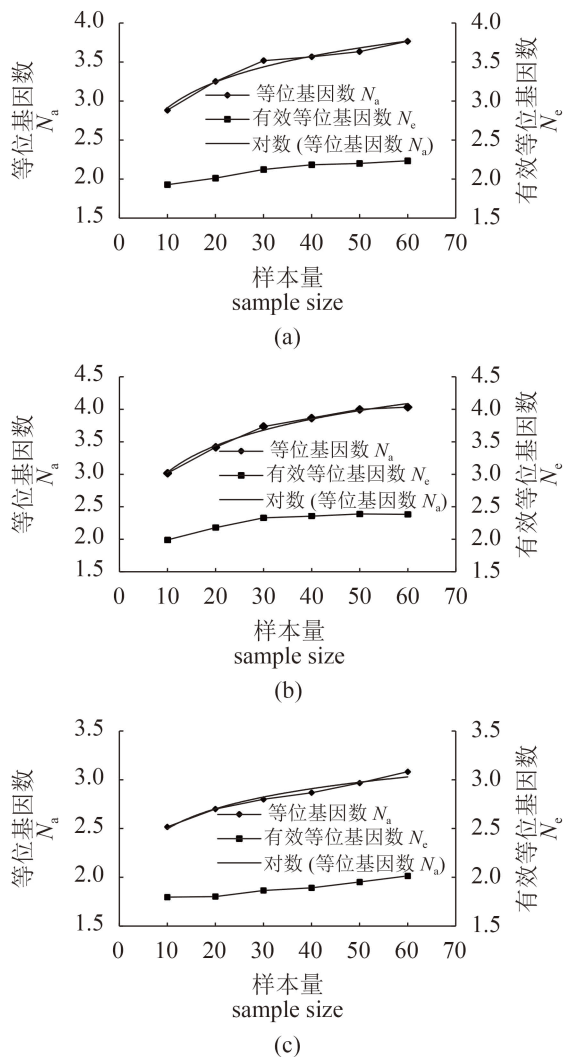


图2 样本量与平均等位基因数、平均有效等位基因数间的相关关系及拟合曲线

(a)(b)(c)分别为SH、ZJ、MF3群体，拟合曲线分别为： $\hat{y}=0.478 4\ln(x)+1.810 1$ ， $r^2=0.981 3$ ； $\hat{y}=0.591 \ln(x)+1.668 9$ ， $r^2=0.990 5$ ； $\hat{y}=0.297 3\ln(x)+1.811 6$ ， $r^2=0.971 4$

Fig. 2 Correlation between sample size and the average number of alleles, the average effective number of alleles and the effective alleles Curve Estimation

(a) (b) (c) represents the group of SH, ZJ, MF, respectively. Effective alleles Curve Estimation: $\hat{y}=0.478 4\ln(x)+1.810 1$, $r^2=0.981 3$; $\hat{y}=0.591 \ln(x)+1.668 9$, $r^2=0.990 5$; $\hat{y}=0.297 3\ln(x)+1.811 6$, $r^2=0.971 4$

指数均分别显著大于由PIC<0.5标记获得的遗传多样性指数(P<0.01)，可见选用不同PIC的标记对群体遗传多样性指数的评估结果存在极显著差异(P<0.01)(表8)。

2.6 性别与遗传多样性指数

MF₃群体雌雄虾遗传多样性指数之间无显著

表3 样本量与遗传多样性指数的相关性分析

Tab. 3 Correlation between sample size and genetic diversity index in different groups

群体 group	N_a	N_e	H_0	H_e	H
SH	0.94**	0.96**	-0.89*	0.94**	0.95**
ZJ	0.95**	0.89*	0.55	0.86*	0.87*
MF3	0.99**	0.98**	-0.45	0.84*	0.96**

注：“*”在0.05水平(双侧)上显著相关，“**”在0.01水平(双侧)上显著相关，下同

Notes: “*”means significant difference at the 0.05 level, “**”means extremely significant difference at the 0.01 level, the same below

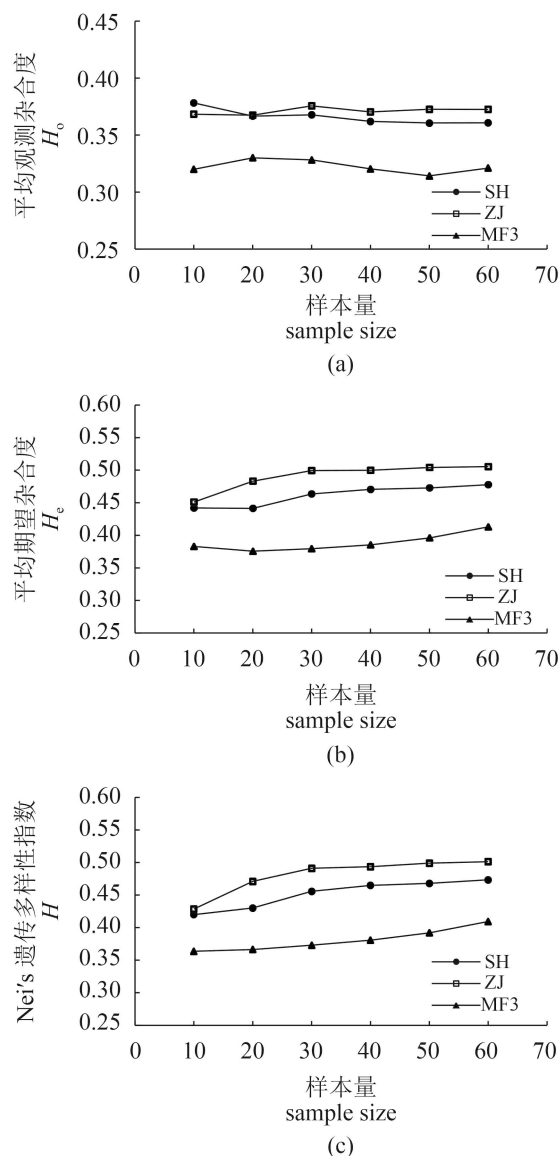


图3 三个群体的不同样本量与H₀、H_e和H的关系

Fig. 3 The relationship between H_0 , H_e and H in three groups with different sample sizes

表 4 不同样本量的群体遗传多样性指数

Tab. 4 Population genetic diversity indexes of different sample sizes

遗传指数 genetic index diversity index	不同样本量的差异显著性($P=0.05$) different sample size					
	10	20	30	40	50	60
N_a	2.81±0.26 ^a	3.12±0.38 ^b	3.35±0.49 ^c	3.43±0.51 ^{cd}	3.53±0.52 ^{cd}	3.63±0.49 ^d
占总体百分比/%	77.34	86.06	92.34	94.64	97.40	100
N_e	1.90±0.10 ^a	2.00±0.19 ^a	2.11±0.23 ^b	2.14±0.23 ^{bc}	2.18±0.22 ^{bc}	2.21±0.19 ^c
占总体百分比/%	86.15	90.33	95.32	98.35	98.60	100
H_e	0.42±0.04 ^a	0.43±0.05 ^{ab}	0.44±0.06 ^{bc}	0.45±0.06 ^{bc}	0.46±0.06 ^c	0.47±0.05 ^c
占总体百分比/%	91.41	93.12	96.15	97.14	98.34	100
H	0.40±0.04 ^a	0.42±0.05 ^{ab}	0.43±0.06 ^{bc}	0.44±0.06 ^{cd}	0.45±0.06 ^{cd}	0.46±0.05 ^d
占总体百分比/%	87.56	91.56	95.35	96.73	98.17	100

表 5 不同样本量稀有等位基因频率

Tab. 5 The rare allele frequency of difference sample sizes

样本 sample size	10	20	30	40	50	60	90	120	150	180
稀有等位基因频率 the frequency of rare alleles	0.063±0.056 ^a	0.075±0.067 ^{ab}	0.089±0.019 ^{bc}	0.095±0.016 ^c	0.096±0.012 ^c	0.098±0.014 ^c	0.099±0.014 ^c	0.102±0.007 ^e	0.101±0.005 ^e	0.103

注: 稀有等位基因频率为30次抽样平均值

Notes: The frequency of rare alleles is the mean value of 30 times' sampling

表 6 不同群体标记量与遗传多样性指数的相关性分析

Tab. 6 Correlation between loci numbers and genetic diversity indexes in different groups

种群 group	遗传多样性指数 genetic index				
	N_a	N_e	H_o	H_e	H
SH	0.66	0.71**	-0.12	0.67*	0.59*
ZJ	0.74*	0.84*	0.84*	0.87*	0.87*
MF ₃	0.67*	0.53	0.79**	0.75*	0.75*

性差异($P>0.05$), 表明性别对遗传多样性参数无显著性影响(表9)。

3 讨论

3.1 样本量对遗传多样性的影响

利用微卫星进行群体的遗传多样性分析时常采用两类多样性指标: 一个是 N_a 和 N_e ; 另一个是基于等位基因频率分布的参数, 包括 H_o 和 H_e ^[10], 在应用不同的遗传参数进行分析时所需的样本量和标记量有所不同。本实验中 N_a 随样本量的增大而增大, 没有出现平台现象, 且在3个养殖群体中, N_a 与样本量之间的相关系数均达到0.9以上, 表明两者之间呈现高度的正相关。只有当种群的采样量足够大时, 才能正确评估群

体的等位基因数。在微卫星中存在大量的稀有基因^[11], 随着样本量的不断增大, 稀有基因不断被检测出, 即样本量越大, 稀有基因被检测出的概率就越大。而 N_e 却不同, 虽然随着样本量的增大更多的稀有基因被检测出, 但是由于稀有基因的频率较低, 所以对 N_e 的贡献很低。有研究表明, N_a 的绝对数越接近检测到的 N_e , 就表明了等位基因在该群体中的分布越均匀^[5]。杂合度指的是随机抽取的样本中其两个等位基因不相同的可能性, 期望杂合度主要是根据群体内当前优势等位基因的分布频率来推算的, 稀有基因对其贡献则更小。当样本量大于30时, 不同样本量下稀有等位基因频率差异不显著($P>0.05$)。本实验当以 N_a 来衡量遗传多样性时, 样本量达到40接近群体的95%, 而用 H_e 和 H 来衡量时, 仅当

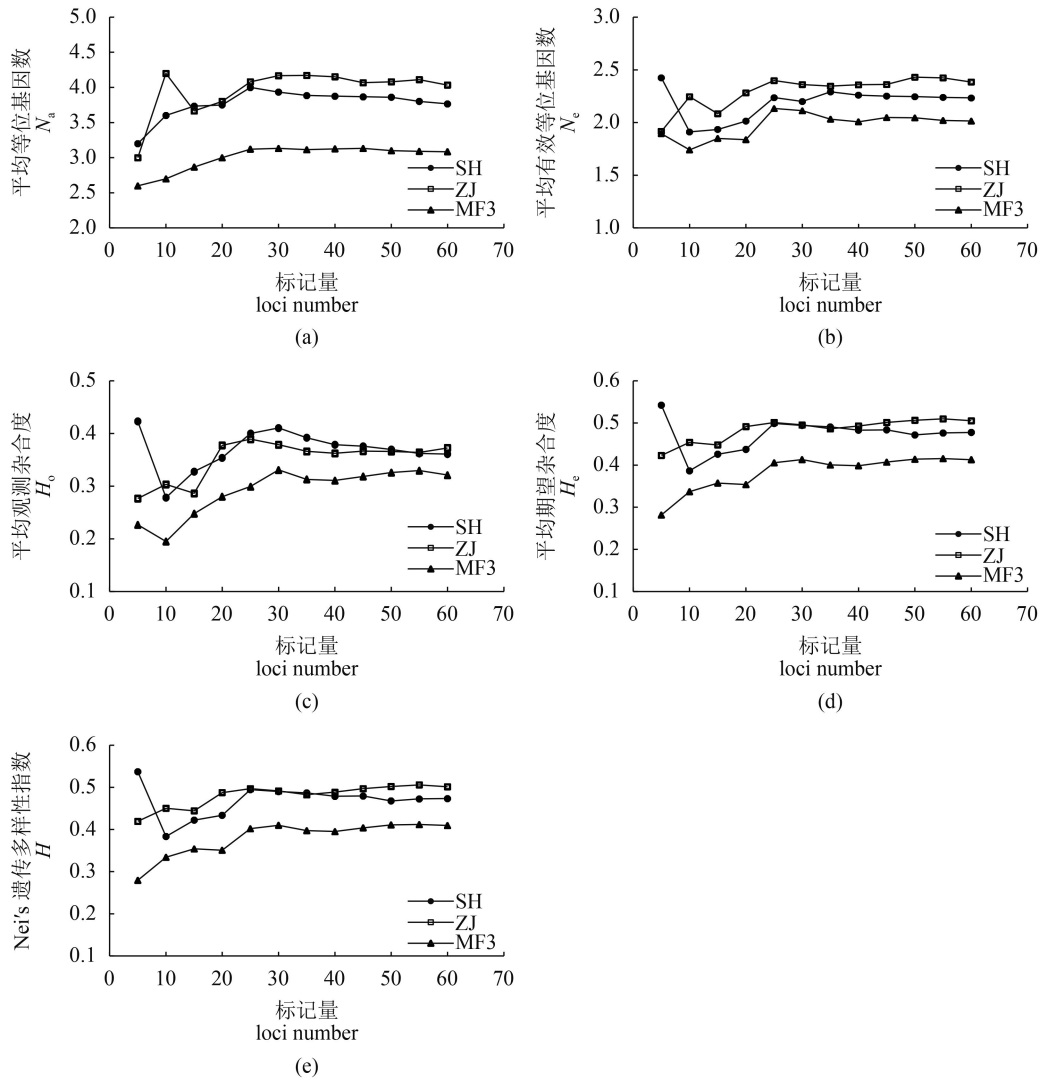


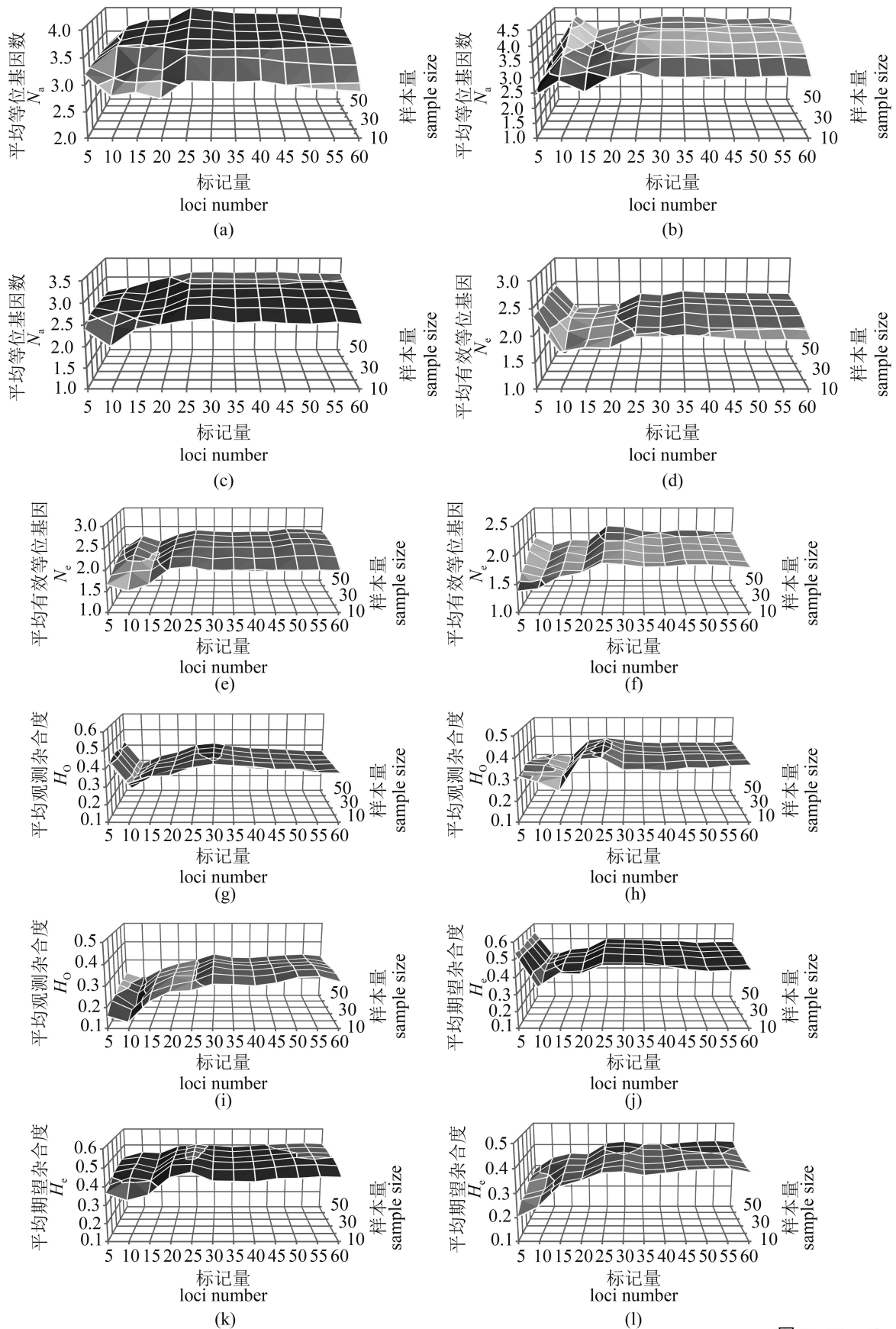
图 4 三个群体不同标记量对 N_a , N_e , H_o , H_e 和 H 的影响

Fig. 4 The relationship between N_a , N_e , H_o , H_e , H and loci number for SH, ZJ, MF3 groups

表 7 不同标记量的群体遗传多样性指数

Tab. 7 Population genetic diversity index of different loci numbers

	遗传多样性指数 genetic index				
	N_a	N_e	H_o	H_e	H
5	2.93±0.31 ^a	2.08±0.30 ^{ab}	0.31±0.10 ^{bc}	0.42±0.13 ^{ab}	0.41±0.13 ^{ab}
10	3.50±0.75 ^{bc}	1.97±0.26 ^a	0.26±0.06 ^a	0.39±0.06 ^a	0.39±0.06 ^a
15	3.42±0.48 ^b	1.96±0.12 ^a	0.29±0.04 ^{ab}	0.41±0.05 ^{ab}	0.41±0.05 ^{ab}
20	3.52±0.45 ^{bc}	2.04±0.22 ^{ab}	0.34±0.05 ^{cd}	0.43±0.07 ^{ab}	0.42±0.07 ^{ab}
25	3.73±0.53 ^c	2.26±0.13 ^b	0.36±0.06 ^{cd}	0.47±0.05 ^b	0.46±0.05 ^b
30	3.74±0.54 ^c	2.22±0.13 ^b	0.37±0.04 ^d	0.47±0.05 ^b	0.46±0.05 ^b
35	3.72±0.55 ^c	2.22±0.17 ^b	0.36±0.04 ^{cd}	0.46±0.05 ^b	0.46±0.05 ^b
40	3.72±0.53 ^c	2.21±0.18 ^b	0.35±0.04 ^{cd}	0.46±0.05 ^b	0.45±0.05 ^b
45	3.69±0.49 ^c	2.22±0.16 ^b	0.35±0.03 ^{cd}	0.46±0.05 ^b	0.46±0.05 ^b
50	3.68±0.51 ^{bc}	2.24±0.19 ^b	0.35±0.02 ^{cd}	0.46±0.05 ^b	0.46±0.05 ^b
55	3.67±0.52 ^{bc}	2.23±0.20 ^b	0.35±0.02 ^{cd}	0.47±0.05 ^b	0.46±0.05 ^b
60	3.63±0.49 ^{bc}	2.21±0.19 ^b	0.35±0.03 ^{cd}	0.47±0.05 ^b	0.46±0.05 ^b



(图5 Fig. 5)

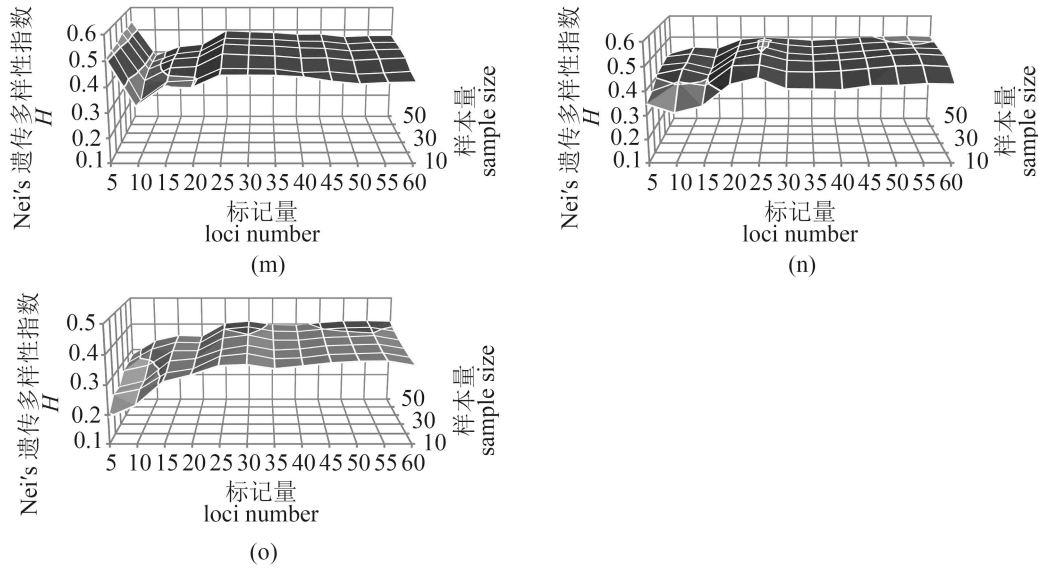


图 5 样本量与标记量的交叉组合与遗传指数间关系

(a)(b)(c)分别为SH、ZJ、MF₃群体N_a与样本量和标记量的关系；(d)(e)(f)分别为SH、ZJ、MF₃群体N_e与样本量和标记量的关系；(g)(h)(i)分别为SH、ZJ、MF₃群体H_o与样本量和标记量的关系；(j)(k)(l)分别为SH、ZJ、MF₃群体H_e与样本量和标记量的关系；(m)(n)(o)分别为SH、ZJ、MF₃群体H与样本量和标记量的关系

Fig. 5 The relationship between the cross combinations of sample size and loci number and genetic index

(a)(b)(c) stand for the relationship between N_a and sample size and loci number for SH, ZJ, MF₃ groups; (d)(e)(f) stand for the relationship between N_e and sample size and loci number for SH, ZJ, MF₃ groups; (g)(h)(i) stand for the relationship between H_o and sample size and loci number for SH, ZJ, MF₃ groups; (j)(k)(l) stand for the relationship between H_e and sample size and loci number for SH, ZJ, MF₃ groups; (m)(n)(o) stand for the relationship between H and sample size and loci number for SH, ZJ, MF₃ groups

表 8 三个群体中多态信息含量不同的标记对群体遗传多样性指数的影响

Tab. 8 Evaluation of different genetic parameters with the markers of different polymorphism information contents in three groups

遗传指数 genetic index		群体 populations		
		SH	ZJ	MF ₃
PIC<0.5	N _a	2.8±1.2	2.9±1.1	2.4±0.9
	N _e	1.7±0.4	1.7±0.4	1.6±0.4
	H _o	0.30±0.17	0.30±0.17	0.27±0.20
	H _e	0.37±0.15	0.37±0.16	0.31±0.19
	H	0.37±0.15	0.37±0.16	0.31±0.19
PIC>0.5	N _a	5.7±2.4	5.6±1.9	5.4±1.7
	N _e	3.4±1.0	3.4±1.0	3.6±1.0
	H _o	0.49±0.22	0.47±0.17	0.46±0.19
	H _e	0.69±0.07	0.69±0.07	0.70±0.08
	H	0.68±0.07	0.68±0.07	0.70±0.08

样本量为30，就已达到群体总水平的95%以上，这与朱维岳等^[12]在野生大豆居群采样策略的研究中所得到的结果类似。

已有一些学者从统计学和实验的角度得出了用于微卫星遗传分析时最适宜的样本量。包文斌等^[4]在利用29对微卫星标记对4个中国地方

表9 雄性组和雌性组的遗传多样性指数

Tab. 9 Genetic diversity index of male and female groups

	平均等位基因数 N_a	平均有效等位基因数 N_e	观测杂合度 H_o	期望杂合度 H_e	Nei's遗传多样性指数 H
雄性♂	2.9±1.6	2.0±1.1	0.32±0.22	0.41±0.24	0.40±0.23
雌性♀	2.8±1.5	1.9±0.9	0.30±0.22	0.38±0.24	0.38±0.23

鸡品种遗传多样性的微卫星分析中发现, 样本量在20~25时较为合适。张根喜等^[13]通过研究证明, 当样本量超过30时, 遗传多样性指数趋于稳定。闫路娜等^[11]以中国飞蝗种群的微卫星遗传分析为例, 评估了样本量对群体遗传多样性指标的影响, 结果显示, 30~50个个体是微卫星DNA分析中所需要的最小样本量。牛东红等^[14]采用微卫星标记进行缙蛭(*Sinonovacula constricta*)群体遗传研究时指出, 最小样本量应控制在30以上。

3.2 标记的选取对遗传多样性的影响

研究所用微卫星标记的多态性对群体的遗传多样性水平有较大影响。*PIC*是衡量一个标记多态性好坏的重要指标, 实验通过利用不同多态信息含量的标记对SH、ZJ、MF₃ 3个群体进行遗传多样性参数分析比较发现, 选用不同多态信息含量的微卫星分子标记对群体遗传参数的评估有着显著性差异。本实验中, 随着标记量的增大, 各遗传多样性指数均呈现出两个变化阶段: 当标记数小于25时, 各指数出现明显的变化, 得出的结果可靠性较低, 但是当标记数大于25时, 各遗传参数逐渐趋于稳定。该结论与李鸥等^[5]研究观察到的结果相似, 即在标记数大于25时, 各遗传参数呈现稳定趋势。

3.3 最适宜样本量和标记量的组合

Nei^[15]在早期的研究中通过对墨西哥脂鲤两个群体蛋白位点的期望杂合度的比较发现, 采用较多的标记数比采用较大的样本量更易得出可靠的结论。在标记数为67, 样本量为20时得到的 H_e 比采用标记数为17, 样本量为79时得到的数值更加可靠。张继全等^[16-17]在用实例分析和计算机模拟中研究了Nei's标准遗传距离估测精度的影响发现, 位点数偏少时估测出的遗传距离精度也较低, 但随着样本量的增加, 遗传距离估测精度也在增加。本研究比较了3个群体中样本量及标记量对群体遗传参数的影响, 证实了在取样量大于30, 标记量大于25时群体的遗传多样

性参数变化较小, 可靠度更高。

3.4 性别对遗传多样性指数的影响

在进行样本采集的过程中, 样本的性别因素也是研究者经常要考虑的问题。包文斌等^[4]研究认为, 雌雄两组间的期望杂合度和有效等位基因数没有显著性差异, 并且近交系数也无显著差异。张根喜等^[13]以微卫星标记进行边鸡遗传多样性分析数据为例, 证明在微卫星分析中性别对遗传多样性指标无显著差异。这与本实验结果相同, 即在使用微卫星分析罗氏沼虾群体遗传多样性时不需要考虑个体的性别问题。

参考文献:

- [1] 陈国宏, 季从亮, 王敏强, 等. 12个地方鸡种群遗传结构及遗传多样性分析[J]. 畜牧兽医学报, 2006, 37(2): 105-111.
Chen G R, Ji C L, Wang M Q, et al. Analysis of genetic structure and diversity of 12 Chinese indigenous chicken breeds[J]. Acta Veterinaria et Zootechnica Sinica, 2006, 37(2): 105-111(in Chinese).
- [2] 孙桂荣, 朱庆, 李亮, 等. 丝羽乌骨鸡群体遗传结构的微卫星标记分析[J]. 中国畜牧杂志, 2004, 40(9): 3-5.
Sun G R, Zhu Q, Li L, et al. Genetic variation of nine microsatellite loci in silky fowl[J]. Chinese Journal of Animal Science, 2004, 40(9): 3-5(in Chinese).
- [3] Weigend S, Romanov M N. Current strategies for the assessment and evaluation of genetic diversity in chicken resources[J]. World Poult Sci J, 2001, 57(3): 275-288.
- [4] 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.
Bao W B, Shu J T, Xu S H, et al. Effects of sample size and sex ratio on various genetic diversity measures with microsatellite markers[J]. Chinese Journal of Animal Science, 2007, 43(1): 6-9(in Chinese).
- [5] 李鸥, 赵莹莹, 郭娜, 等. 草鱼种群SSR分析中样本量及标记数量对遗传多样性的影响[J]. 动物学研究, 2009,

- 30(2): 121-130.
- Li O, Zhao Y Y, Guo N, *et al.* Effects of sample size and loci number on genetic diversity in wild population of grass carp revealed by SSR[J]. Zoological Research, 2009, 30(2): 121-130(in Chinese).
- [6] 戴习林, 邓平平, 施永海, 等. 罗氏沼虾SSR标记再开发及其影响因素初探[J]. 生物技术通报, 2012(10): 142-149.
- Dai X L, Deng P P, Shi Y H, *et al.* A preliminary study on re-developing of *Macrobrachium rosenbergii* microsatellite markers and influence factors[J]. Biotechnology Bulletin, 2012(10): 142-149(in Chinese).
- [7] Botstein D, White R L, Skolnick M, *et al.* Construction of a genetic linkage map in man using restriction fragment length polymorphisms[J]. American Journal of Human Genetics, 1980, 32(3): 314-331.
- [8] 丁敏. 实时荧光PCR技术检测猪肺炎支原体研究[D]. 中国农业科学院, 2008.
- Ding M. The development of real time quantitative PCR in the detection of mycoplasma hyopneumoniae[D]. Chinese Academy of Agricultural Sciences, 2008(in Chinese).
- [9] Asimit J, Zeggini E. Rare variant association analysis methods for complex traits[J]. Annual Review of Genetics, 2010, 44: 293-308.
- [10] Leberg P L. Estimating allelic richness: effects of sample size and bottlenecks[J]. Molecular Ecology, 2002, 11(11): 2445-9.
- [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]. Current Zoology, 2004, 50(2): 279-290.
- Yan L N, Zhang D X. Effects of sample size on various genetic diversity measures in population genetic study with microsatellite DNA markers[J]. Acta Zoologica Sinica, 2004, 50(2): 279-290(in Chinese).
- [12] 朱维岳, 周桃英, 钟明, 等. 基于遗传多样性和空间遗传结构的野生大豆居群采样策略[J]. 复旦学报: 自然科学版, 2006, 45(3): 321-327.
- Zhu W Y, Zhou T Y, *et al.* Zhong M. Sampling strategy for wild soybean (*Glycine soja*) populations based on their genetic diversity and fine-scale spatial genetic structure[J]. Journal of Fudan University(Natural Science), 2006, 45(3): 321-327(in Chinese).
- [13] 张跟喜, 丁馥香, 赵秀华, 等. 微卫星分析中样本量和性别对群体遗传多样性的影响[J]. 畜牧与兽医, 2011, 43(6): 36-39.
- Zhang G X, Ding F X, Zhao X H, *et al.* Effect of micro satellite analysis of sample size and sex in genetic diversity[J]. Animal Husbandry & Veterinary Medicine, 2011, 43(6): 36-39(in Chinese).
- [14] 牛东红, 陈慧, 林国文, 等. 缙蛭群体微卫星分析中样本量对遗传多样性指标的影响[J]. 海洋科学进展, 2010, 28(2): 203-208.
- Niu D H, Chen H, Lin G W, *et al.* Effects of sample size on genetic diversity index for population of *Sinonovacula constricta* in use of microsatellite DNA marker[J]. Advances in Marine Science, 2010, 28(2): 203-208(in Chinese).
- [15] Nei M, Roychoudhury A K. Sampling variances of heterozygosity and genetic distance[J]. Genetics, 1974, 76(2): 379-390.
- [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.
- Zhang J Q, Shao C R, Wang M Y, *et al.* Accuracy of the estimation of multi-loci genetic distance[J]. Chinese Journal of Animal and Veterinary Sciences, 1998, 29(2): 128-131(in Chinese).
- [17] 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.
- Zhang J Q, Shao C R, Wang M Y, *et al.* Accuracy of the estimation of Nei's standard genetic distance[J]. Chinese Journal of Animal and Veterinary Sciences, 1998, 29(1): 27-32(in Chinese).

Effects of sample size and loci number on genetic diversity index in *Macrobrachium rosenbergii* with microsatellite markers

DAI Xilin^{1*}, LIU Jie¹, LI Jingjing¹, JIANG Zongbing¹, MING Lei¹, JIANG Fei²

(1. College of Fisheries and Life Sciences, Shanghai Ocean University, Shanghai 201306, China;

2. Shencao Special Aquatic Product Development Company, Shanghai 201506, China)

Abstract: As assessment of genetic structure of population is the key point of aquatic animal breeding, the selection of sample size and loci number is the first step of correct evaluation of population genetic structure. Here, we solved the problem from an analysis of genetic diversity of three breeding groups of *Macrobrachium rosenbergii* at 60 polymorphic microsatellite loci. The results show that selection of different sample sizes and loci numbers both had a great effect on the genetic diversity index. The sample size was positively correlated with the average number of alleles and the average effective number of alleles, and the sample size was moderate and highly correlated with heterozygosity and *Nei's* genetic diversity index, respectively. Only when the sample size exceeds 30, they tend to be stable. When we select the loci, we found the different polymorphism information contents also directly affect the accuracy of the genetic parameters. When the loci number was 5–25, there was a significant difference in the genetic diversity index ($P < 0.05$). When the loci number was more than 25, the genetic parameters changed little ($P > 0.05$). So, we suggested that when microsatellite markers were used to estimate genetic diversity of population as far as possible to choose high polymorphism information content and we can't consider sample gender, meanwhile the minimum loci number and sample size are needed to exceed 25 and 30, respectively.

Key words: *Macrobrachium rosenbergii*; genetic diversity; microsatellite DNA marker; sample size

Corresponding author: DAI Xilin. E-mail: xldai@shou.edu.cn

Funding projects: Agriculture Research System of Shanghai (201505)