

半滑舌鳎有效群体大小估计

王伟峰^{1,2}, 崔瑜^{2,3}, 蒋丽², 王悦玲^{2,3}, 王焕岭¹, 李恒德^{2*}

(1. 华中农业大学水产学院, 农业农村部淡水生物繁育重点实验室,
农业动物遗传育种与繁殖教育部重点实验室, 湖北 武汉 430070;

2. 中国水产科学研究院生物技术研究中心, 北京 100141;

3. 上海海洋大学, 水产科学国家级实验教学示范中心, 上海 201306)

摘要: 为了解半滑舌鳎当前的群体遗传结构, 探究半滑舌鳎有效群体大小(N_e)发展趋势及现状, 实验通过性别特异性分子标记随机鉴定了800尾半滑舌鳎的遗传性别, 选取297个遗传雌性半滑舌鳎样本进行简化基因组测序(2b-RAD), 获得了64 416个可用SNP标记, 利用这些标记进行全基因组范围的连锁不平衡分析得到各染色体上连锁不平衡分布; 根据标记间不同物理间距进行半滑舌鳎有效群体大小的估计, 初步了解各历史世代下有效群体规模; 通过选择0.1、0.2、0.5、1.0、2.0、5.0、10.0 Mb等7个不同染色体片段大小反映出半滑舌鳎群体遗传结构经过自然与人工共同选择的历史发展趋势。结果显示, 半滑舌鳎有效群体大小随其连锁不平衡程度衰减而呈连续下降趋势, 至2世代前其有效群体大小仅为29尾左右。

关键词: 半滑舌鳎; 简化基因组测序(2b-RAD); 连锁不平衡; 有效群体大小

中图分类号: Q 347; S 965.3

文献标志码: A

有效群体大小(effective population size, N_e)是指理想群体中与实际群体具有相同基因频率方差的繁殖个体数目, 它反映了群体平均近交系数增量的大小及群体遗传结构中基因的平均纯合度。近交不利于种质资源的发展, 最显著的结果是造成子代某些性状的表型均值降低, 即发生近交衰退^[1]。有效群体大小影响近交的程度, 较小时则表明, 群体倾向于近交水平的增加^[2]。因此, 开展有效群体大小的相关研究对遗传育种中的近交预测和控制, 以及了解遗传漂变对群体多样性的影响具有重要意义。

有效群体大小估计的具体方法最早由Wright^[3]提出, 但是由于不同研究背景下的群体结构、群体类型不同, 有效群体大小的估计方法也并不唯一^[4]。传统的有效群体估计方法是建立在群体内相互独立的个体随机交配基础上, 这在实际研究中往往很难实现。早期禽畜育种研究中多

以系谱信息辅助估计, 但由于禽畜家系往往都比较复杂, 很难避免由系谱信息发生错误而引起的估计误差^[5]。随着微卫星、单核苷酸多态性(single nucleotide polymorphism, SNP)分型方法的不断优化和广泛普及, 分子标记已经成为群体遗传学研究的重要辅助手段, 利用高密度分子标记间的连锁不平衡(linkage disequilibrium, LD)可以更准确地估计有效群体大小^[6]。水生动物育种研究中也有利用连锁不平衡估计有效群体大小的报道^[7], 但是由于其使用的分子标记仅为49个微卫星, 分布范围比较局限, 容易造成较大的估计误差。根据DNA序列的多态性估计有效群体大小的方法原理比较简单, 但是由于育种过程中有效群体规模不断变化, 导致其只能估计过去很长一段时间内有效群体大小的平均变化^[8]。

半滑舌鳎(*Cynoglossus semilaevis*)是我国特有的海水养殖珍品, 分布范围主要为东海、黄海

收稿日期: 2017-10-15 修回日期: 2018-04-03

资助项目: 中国水产科学研究院基本科研业务费专项(2016ZD0203, 2016B005)

通信作者: 李恒德, E-mail: hengde.li@cafs.ac.cn

和渤海海域^[9]。受繁殖技术操作难度大又相对封闭的局限,以及对养殖群体中亲本质量的严格要求,导致育种群体规模相对较小,极易发生近交衰退现象,进而降低优良性状表型均值。半滑舌鳎具有明显的性别二态性,其雌性生长速率约为雄性的4倍,并且自然状态下存在由雌性到雄性的性逆转现象^[10-11],这种性别差异生物学特性的存在导致生产中只关注提高雌性率,缺少严格科学系统的选育。因此,通过高密度SNP标记间连锁不平衡(LD)来估计半滑舌鳎有效群体大小具有一定的实际意义,将会为半滑舌鳎保种及选育工作提供合理的配种方案。国内外对有效群体大小的相关研究已经有诸多报道,且主要集中在禽畜育种中^[4],水生动物中主要以贝类为主^[12-14]。张国范^[1]通过设置理论上的有效群体大小梯度,对海湾扇贝(*Argopecten irradians*) F₁生长和存活的影响进行了深入探究。Hedgecock等^[13]根据不同时期中性基因突变频率,对人工养殖的长牡蛎(*Crassostrea gigas*)等多种贝类共16个群体进行了有效群体大小估计,结果显示其N_e都小于100甚至多数低于50(13个群体)。Rexroad等^[7]使用49个微卫星标记,分析了虹鳟(*Oncorhynchus mykiss*)的连锁不平衡程度并利用部分染色体进行了有效群体大小估计,在一定程度上反映了虹鳟的遗传结构受到了不同强度的选择。由于测序技术的迅速发展,禽畜育种中陆续出现利用高密度的SNP标记进行有效群体大小估计的报道^[15-17],水产动物育种研究中至今未见相关方法的应用。因此,通过高密度SNP标记进行水生动物有效群体大小估计具有重要意义。

1 材料与方法

1.1 实验对象

实验对象为采自山东潍坊和河北黄骅的800尾4月龄半滑舌鳎,均来自随机交配的养殖群体后代,通过性别特异性分子标记^[18]鉴定所有样本的遗传性别,其中遗传性别为雌性且符合质量标准的297尾半滑舌鳎DNA样本用于建库和分型。

1.2 基因型测定

提取DNA后,建立2b-RAD^[19]文库,使用Illumina HiSeq 2500平台对全部个体进行单末端测序。测序后共产生1 582 669 040个reads,平均每样本5 328 852个reads。最小等位基因频率小于0.05,

样本检出率(call rate)小于0.90及染色体位置未知的直接剔除。最后使用RADtyping v1.5进行2b-RAD分型,共获得64 416个SNP,并借助SOAP2^[20]将其定位于半滑舌鳎基因组^[21]。

1.3 有效群体大小估计

依据Sved^[22]提出的连锁不平衡与有效群体大小之间关系的公式:

$$E(r^2) \approx 1/(\alpha + KN_e c) \quad (1)$$

式中, r^2 为不同标记间的连锁不平衡程度, N_e 为有效群体大小, c 为标记间的遗传距离,单位为摩尔根(M), α 取值为1时表示不考虑突变, α 为2时则考虑突变^[23], K 的取值与染色体类型有关,常染色体为4,性染色体为2。

目前大多数研究都采用无突变模型进行高密度SNP估计有效群体大小^[6],本研究也选取无突变模型进行分析。根据公式(1),估计有效群体大小必须先获得标记间遗传距离(c)和连锁不平衡程度(r^2)。

1.4 连锁不平衡程度度量

LD的度量有多种指标,其中常用的为 D' 和 r^2 ,但 D' 受群体影响较大,中小群体容易出现偏差, r^2 则对群体大小敏感性不强^[24],本研究因此使用 r^2 作为LD的度量指标。对于2个SNP标记,其等位基因分别为A、a和B、b及相应的等位基因频率为 P_A 、 P_a 、 P_B 、 P_b ,等位基因间形成的4个单倍型频率分别为 P_{AB} 、 P_{Ab} 、 P_{aB} 、 P_{ab} ,2个标记间的 r^2 计算公式:

$$r^2 = \frac{(P_{AB}P_{ab} - P_{Ab}P_{aB})^2}{P_A P_a P_B P_b} \quad (2)$$

采用HaploView软件^[25]进行连锁不平衡分析并获得各SNP标记间的 r^2 值。

1.5 标记间遗传距离

对于高密度SNP标记,目前还无法直接通过连锁图谱获得各标记间的遗传距离,通常是根

$$\text{标记间遗传距离(cM)} = \frac{\text{染色体遗传距离(cM)}}{\text{染色体物理距离(Mb)}} \times \text{标记间物理距离(Mb)} \quad (3)$$

染色体物理距离和遗传距离可以分别从半滑舌鳎基因组^[21]和半滑舌鳎高密度连锁图谱^[26]中获得。

1.6 不同世代下有效群体大小

Hayes等^[27]提出遗传距离为 c 的染色体片段可以近似反映出前 $1/(2c)$ 世代的有效群体大小。因此, 根据公式(1)、(3), 可通过不同物理间距的SNP标记间的连锁不平衡, 估计不同世代下的有效群体大小。本实验选取0.1、0.2、0.5、1.0、2.0、5.0和10.0 Mb等7个不同SNP间距, 通过对这些不同间距SNP标记之间连锁不平衡程度分析, 估计半滑舌鳎在相应前128、64、26、13、7、3和2世代下的有效群体大小。

在本研究中, 除LD程度值 r^2 之外, 其余数据都使用R语言^[28]进行整理和分析。

2 结果

2.1 SNP在全基因组上的分布

参考半滑舌鳎基因组序列, 对获得的SNP标记进行定位, 确定其在全基因组中分布情况。质量控制前后SNP分布变化结果显示, 质量控制后, 样本个数保持不变, SNP则由64 416个减少为31 241个, 平均间距约为14 kb(图1)。

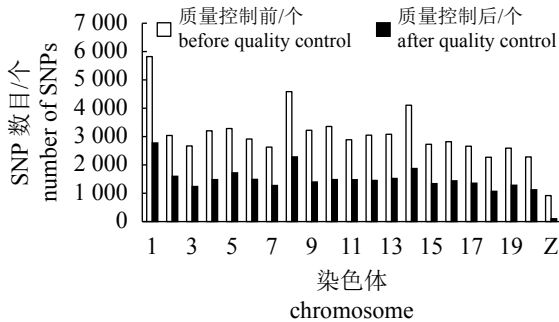


图1 质量控制前、后半滑舌鳎全基因组上SNP分布
Fig. 1 SNP distribution of whole genome before and after quality control

2.2 LD分布规律

从半滑舌鳎全基因组上不同SNP间距LD分布的整个趋势上看, 随着物理距离增大, LD逐渐减小, 但在某些距离较远的点之间也会出现比较强的LD(图2)。这表示基因组上的某些突变之间可能存在不受物理距离影响的连锁不平衡, 这与其他物种中出现的情况类似^[29]。同时, SNP间距较小时其LD变异程度较大, 而间距较大的SNP间LD的变异程度则相对较小(图2)。

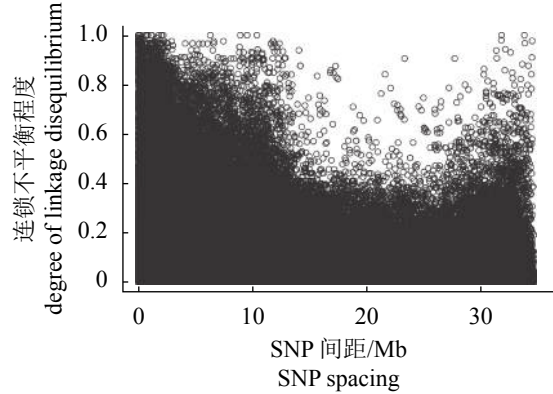


图2 半滑舌鳎全基因组连锁不平衡(r^2)分布
Fig. 2 Distribution of whole genome linkage disequilibrium (r^2) distribution of *C. semilaevis*

2.3 有效群体大小估计

半滑舌鳎有效群体大小呈逐渐下降的趋势, 从128世代前的2 869尾连续降至2世代前的29尾(表1, 图3)。并且, 即使相同的SNP间距, 不同SNP标记间连锁不平衡程度也有差异(图2), 因此有效群体大小估计也会有所差异。如在2世代时, 间距为10 Mb的不同SNP标记间估计的半滑舌鳎有效群体大小为17~40尾(表1)。

连锁不平衡程度的差异也反映在不同染色体上。不同染色体估计的有效群体大小具有很大差异, 如前2世代时, 半滑舌鳎有效群体大小差异为17~40尾(表1), 而在全部染色体中来看, 其数值应为18~192尾(表2)。同时, 根据性染色体上的连锁不平衡估计的有效群体大小已经达到其相同世代下常染色体的数倍以上, 以致无法在图中显示128、64世代的具体数值, 但从前5个世代的增长趋势不难发现这一规律(图3)。造成这种差异的原因是不同染色体在相同片段大小的情况下, 连锁不平衡程度较高的染色体估计的有效群体偏小, 反之, 染色体具有较小的连锁不平衡程度, 其有效群体估计值则偏大。如1和14染色体估计的有效群体较大, 其对应连锁不平衡程度(r^2)就比较小, 第3和19染色体上结果则恰好相反(表3)。对不同染色体上相同间距估计的有效群体大小进行方差分析可知, 各染色体间的差异显著($P < 0.05$)。同时, 有效群体大小由远世代至近世代的衰减趋势(表2), 与全基因组中连锁不平衡程度的衰减趋势基本一致(图2)。

表 1 半滑舌鲷根据常染色体SNP标记估计的各世代有效群体大小

Tab. 1 Effective population size for each generation estimated by the autosomal SNP markers

		SNP间距/kb SNP spacing						
		100	200	500	1 000	2 000	5 000	10 000
遗传距离	genetic distance	0.004	0.008	0.019	0.039	0.078	0.19	0.374
世代	generation	128	64	26	13	7	3	2
SNP对数	number of SNPs	7 589	14 685	34 286	64 381	116 923	211 314	286 691
N_e (mean±SD)		2 869.25±1 155.94	1 434.62±578.00	573.84±231.13	286.91±115.58	143.46±57.74	57.39±23.07	28.69±11.48

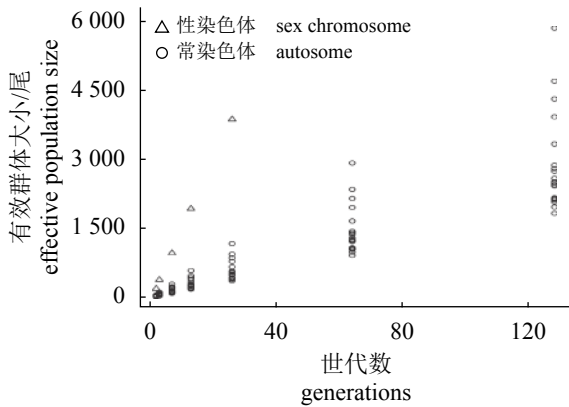


图 3 半滑舌鲷各历史世代(前128~前2世代)有效群体大小

Fig. 3 Effective population size of *C. semilaievis* in historical generations (ancestor generations from 128 to 2)

3 讨论

近年来,半滑舌鲷的市场需求不断扩大,导致繁殖方式愈发追求时效,模拟自然环境和人工授精等技术的不断成熟,使其选择进程大大超越了自然选择的历程。且其育种区域主要分布在渤海沿岸,分布范围相对集中。这些因素都会导致半滑舌鲷近交选择的程度逐渐加重,有效群体大小也很难维持在较高水平。根据本研究的估计,半滑舌鲷有效群体大小在2世代前仅为20~46尾。

根据标记间不同间距分析所得连锁不平衡程度有一定差异。在标记间距较小时,LD程度偏大,而随着标记间距增大LD程度逐渐衰减。从公式(1)、(3)能看出, r^2 与 N_e 应为正相关关系。

表 2 由不同染色体估计的半滑舌鲷有效群体大小

Tab. 2 Estimates of effective population size of *C. semilaievis* by different chromosomes

染色体号 chromosome	世代数 generations						
	128	64	26	13	7	3	2
1	5 838.483	2 919.410	1 167.720	583.800	291.952	116.778	58.391
2	2 114.623	1 056.534	422.583	211.281	105.665	42.263	21.130
3	1 968.275	984.166	393.615	196.779	98.405	39.361	19.682
4	4 303.107	2 151.578	860.414	430.339	215.146	86.059	43.030
5	2 588.190	1 293.850	517.656	258.822	129.404	51.767	25.883
6	2 447.223	1 224.058	489.715	244.780	122.386	48.962	24.481
7	2 138.462	1 068.990	427.474	213.766	106.867	42.751	21.379
8	3 917.491	1 958.167	783.355	391.725	195.852	78.341	39.170
9	2 074.835	1 037.932	415.060	207.540	103.781	41.511	20.755
10	3 328.829	1 664.410	665.848	332.884	166.446	66.582	33.289
11	2 485.219	1 242.911	497.405	248.609	124.338	49.731	24.865
12	2 513.340	1 256.754	502.718	251.365	125.668	50.272	25.134

· 续表2 ·

染色体号 chromosome	世代数 generations						
	128	64	26	13	7	3	2
13	2 741.621	1 370.449	548.172	274.071	137.051	54.820	27.408
14	4 691.834	2 346.305	938.398	469.258	234.614	93.850	46.926
15	2 870.811	1 434.950	573.820	286.916	143.467	57.380	28.693
16	2 798.959	1 400.340	560.097	280.017	140.013	56.014	28.005
17	2 426.055	1 212.762	485.047	242.516	121.246	48.503	24.248
18	2 145.696	1 073.167	429.389	214.582	107.331	42.930	21.467
19	1 829.483	915.006	365.982	182.942	91.491	36.596	18.298
20	2 162.551	1 080.635	432.421	216.119	108.108	43.240	21.620
Z	19 255.65	9 642.976	3 864.520	1 930.672	963.903	385.633	192.952

表 3 各染色体不同物理间距的 r^2 分布情况Tab. 3 r^2 distribution of different physical distances of each chromosome

染色体号 chromosome	SNP间距/kb SNP spacing						
	100	200	500	1 000	2 000	5 000	10 000
1	0.032±0.091	0.029±0.092	0.029±0.085	0.027±0.078	0.026±0.072	0.020±0.054	0.017±0.045
2	0.035±0.114	0.040±0.111	0.031±0.085	0.031±0.082	0.027±0.072	0.024±0.064	0.020±0.052
3	0.033±0.096	0.042±0.110	0.031±0.087	0.032±0.089	0.027±0.078	0.022±0.064	0.018±0.050
4	0.041±0.115	0.035±0.102	0.035±0.096	0.031±0.086	0.032±0.088	0.025±0.070	0.019±0.048
5	0.038±0.115	0.036±0.102	0.030±0.084	0.034±0.089	0.027±0.073	0.023±0.061	0.018±0.051
6	0.033±0.095	0.037±0.108	0.033±0.096	0.027±0.071	0.030±0.083	0.024±0.064	0.021±0.063
7	0.038±0.103	0.033±0.087	0.033±0.093	0.031±0.088	0.028±0.080	0.027±0.081	0.016±0.046
8	0.035±0.090	0.038±0.110	0.034±0.094	0.030±0.080	0.028±0.079	0.020±0.052	0.017±0.043
9	0.032±0.087	0.027±0.075	0.032±0.095	0.026±0.071	0.025±0.071	0.023±0.064	0.020±0.054
10	0.033±0.091	0.035±0.097	0.032±0.090	0.03 ±0.083	0.026±0.069	0.023±0.066	0.022±0.061
11	0.033±0.088	0.044±0.119	0.033±0.089	0.029±0.079	0.025±0.073	0.023±0.065	0.019±0.051
12	0.045±0.130	0.033±0.089	0.029±0.078	0.029±0.083	0.027±0.079	0.025±0.070	0.018±0.050
13	0.030±0.092	0.033±0.094	0.029±0.083	0.025±0.065	0.023±0.070	0.024±0.070	0.021±0.059
14	0.030±0.074	0.030±0.084	0.031±0.088	0.025±0.069	0.029±0.084	0.022±0.062	0.019±0.052
15	0.029±0.072	0.035±0.097	0.029±0.073	0.027±0.084	0.026±0.074	0.024±0.066	0.021±0.057
16	0.037±0.102	0.035±0.102	0.031±0.086	0.030±0.077	0.030±0.081	0.025±0.065	0.021±0.021
17	0.032±0.080	0.031±0.087	0.033±0.090	0.028±0.074	0.026±0.071	0.024±0.062	0.018±0.046
18	0.044±0.119	0.024±0.062	0.034±0.099	0.030±0.084	0.025±0.068	0.026±0.072	0.020±0.053
19	0.041±0.124	0.039±0.113	0.030±0.085	0.025±0.069	0.025±0.072	0.023±0.066	0.021±0.057
20	0.028±0.080	0.039±0.114	0.034±0.084	0.027±0.080	0.026±0.073	0.023±0.064	0.021±0.064
Z	0.002±0.001	0.012±0.017	0.015±0.027	0.014±0.024	0.014±0.032	0.009±0.016	0.013±0.024

这表明,随着标记间距离不断增大,连锁不平衡程度逐渐衰减,估计所得有效群体大小理论上应随之降低。根据历史世代的 $1/2c$ ^[26]理论,即染色体在不同片段大小下能够反映不同世代的有效群体大小,短染色体片段可以估计远世代有效群体大小,而长染色体片段可以估计近世代有效群体大小。

相同标记间距时分析所得连锁不平衡程度同样表现出很大差异。这种差异可能是重组率差异、杂合度、遗传漂变和选择效应等因素综合反映的结果。本研究中也选择了标记距离更短的50 kb及距离更长的15 Mb时进行分析。前者所得LD程度值约为0.5,这表明标记间完全自由重组,此时的连锁为连锁平衡状态;15 Mb时连锁程度估计值与10 Mb时几乎相同,这表明标记间距离过大时LD程度估计很容易出现偏差,因而其对应的有效群体大小可能不够准确。鉴于此,本研究选取了0.1与10 Mb之间7个不同长度的染色体片段大小进行LD程度分析和有效群体大小估计。

半滑舌鳎为ZZ(♂)/ZW(♀)型^[30]性染色体系统,其中W染色体长度过小,测序所得的少量SNP标记难以对有效群体大小进行准确估计,故本实验不对W染色体进行分析。根据性染色体连锁不平衡估计的有效群体大小与常染色体之间具有明显差异,这可能是由于性染色体估计方法不同造成的,也可能与养殖群体中长期高强度人工选择高雌性性别比例群体有关。为了避免由性染色体带来的统计误差,实验在各世代有效群体大小估计时仅采用常染色体进行分析。

有效群体大小与连锁不平衡之间关系密切,而基因组范围内的连锁不平衡能够用于推断祖先的有效群体大小。本研究中半滑舌鳎的有效群体大小随世代进化呈逐渐减少的趋势,而有效群体的不断减少势必导致遗传变异也越来越小,最终将导致遗传进展受到极大影响。因此,通过连锁不平衡程度分析和有效群体大小的估计,进而合理控制半滑舌鳎有效群体规模具有重要意义。本研究中采用高密度SNP标记分析虽然能在一定程度上反映半滑舌鳎全基因组连锁不平衡程度,但是受研究对象的样本规模较小、分布范围比较集中及数据分析软件本身的局限等因素影响,最终结果可能不够准确,但从估计结果所反映出的有效群体大小变化趋

势来看,半滑舌鳎近年的选育工作已经显现了近交程度的加剧。

在今后的水产育种工作中,育种工作者应当借鉴相关研究成果,采取有效措施合理控制有效群体大小,避免因其有效群体规模过小导致遗传结构产生不利改变而影响遗传进展。同样,在进行其他水产养殖品种的保种和育种工作时,也应当采用合适的有效群体估计方法进行相关探究,从而掌握物种当前的群体遗传结构、估计当前或近代的有效群体规模,以避免近交等因素造成的不利影响。

参考文献:

- [1] 张国范. 海洋贝类遗传育种研究20年[J]. 厦门大学学报(自然科学版), 2006, 45(s2): 190-194.
Zhang G F. Advances in genetics and breeding of marine mollusk in 20 Years[J]. Journal of Xiamen University (Natural Science Edition), 2006, 45(S2): 190-194(in Chinese).
- [2] 赵力强, 梁健, 闫喜武, 等. 有效群体大小对蛤仔F₁生长和存活影响[J]. 海洋科学, 2015, 39(1): 64-69.
Zhao L Q, Liang J, Yan X W, et al. Effect of Effective Population Size on the F₁ growth and survival of the Manila clam *Ruditapes philippinarum*[J]. Marine sciences, 2015, 39(1): 64-69(in Chinese).
- [3] Wright S. Size of population and breeding structure in relation to evolution[J]. Science, 1938, 87(2263): 430-431.
- [4] Frankham R. Introduction to quantitative genetics (4th edn): by Douglas S. Falconer and Trudy F.C. Mackay Longman, 1996. £ 24.99 pbk (xv and 464 pages) ISBN 0582 24302 5[J]. Trends in Genetics, 1996, 12(7): 280-280.
- [5] Sørensen A C, Sørensen M K, Berg P. Inbreeding in Danish dairy cattle breeds[J]. Journal of Dairy Science, 2005, 88(5): 1865-1872.
- [6] 尼桂琰, 张哲, 姜力, 等. 利用全基因组连锁不平衡估计中国荷斯坦牛有效群体大小[J]. 遗传, 2012, 34(1): 50-58.
Ni G Y, Zhang Z, Jiang L, et al. Chinese Holstein Cattle effective population size estimated from whole genome linkage disequilibrium[J]. Hereditas, 2012, 34(1): 50-58(in Chinese).
- [7] Rexroad III C R, Vallejo R L. Estimates of linkage

- disequilibrium and effective population size in rainbow trout[J]. *BMC genetics*, 2009, 10: 83.
- [8] Tenesa A, Navarro P, Hayes B J, *et al.* Recent human effective population size estimated from linkage disequilibrium[J]. *Genome Research*, 2007, 17(4): 520-526.
- [9] Chen S L, Ji X S, Shao C W, *et al.* Induction of mitogenetic diploids and identification of WW super-female using sex-specific SSR markers in half-smooth tongue sole (*Cynoglossus semilaevis*)[J]. *Marine Biotechnology*, 2012, 14(1): 120-128.
- [10] Chen S L, Tian Y S, Yang J F, *et al.* Artificial gynogenesis and sex determination in half-smooth tongue sole (*Cynoglossus semilaevis*)[J]. *Marine Biotechnology*, 2009, 11(2): 243-251.
- [11] 季相山, 陈松林, 马洪雨, 等. 半滑舌鳎养殖群体中自然性逆转伪雄鱼的发现[J]. *水产学报*, 2010, 34(2): 322-327.
- Ji X S, Chen S L, Ma H Y, *et al.* Natural sex reversal of female *Cynoglossus semilaevis* in rearing populations[J]. *Journal of Fisheries of China*, 2010, 34(2): 322-327(in Chinese).
- [12] Hedgecock D, Sly F. Genetic drift and effective population sizes of hatchery-propagated stocks of the Pacific oyster, *Crassostrea gigas*[J]. *Aquaculture*, 1990, 88(1): 21-38.
- [13] Hedgecock D, Chow V, Waples R S. Effective population numbers of shellfish broodstocks estimated from temporal variance in allelic frequencies[J]. *Aquaculture*, 1992, 108(3-4): 215-232.
- [14] Zhang H B, Liu X, Zhang G F, *et al.* Effects of effective population size on the F₂ growth and survival of bay scallop *Argopecten irradians irradians* (Lamarck)[J]. *Marine fisheries*, 2005, 24(4): 114-120.
- [15] Kim E S, Kirkpatrick B W. Linkage disequilibrium in the North American Holstein population[J]. *Animal Genetics*, 2009, 40(3): 279.
- [16] Sargolzaei M, Schenkel F S, Jansen G B, *et al.* Extent of linkage disequilibrium in Holstein cattle in North America[J]. *Journal of Dairy Science*, 2008, 91(5): 2106-2117.
- [17] Qanbari S, Pimentel E C G, Tetens J, *et al.* The pattern of linkage disequilibrium in German Holstein cattle[J]. *Animal Genetics*, 2010, 41(4): 346-356.
- [18] 刘洋, 陈松林, 高峰涛, 等. 半滑舌鳎性别特异微卫星标记的SCAR转化及其应用[J]. *农业生物技术学报*, 2014, 22(6): 787-792.
- Liu Y, Chen S L, Gao F T, *et al.* SCAR-transformation of sex-specific SSR marker and its application in half-smooth tongue sole (*Cynoglossus semilaevis*)[J]. *Journal of Agricultural Biotechnology*, 2014, 22(6): 787-792(in Chinese).
- [19] Wang S, Meyer E, McKay J K, *et al.* 2b-RAD: a simple and flexible method for genome-wide genotyping[J]. *Nature Methods*, 2012, 9(8): 808-810.
- [20] Li R Q, Yu C, Li Y R, *et al.* SOAP2: an improved ultrafast tool for short read alignment[J]. *Bioinformatics*, 2009, 25(15): 1966-1967.
- [21] Chen S L, Zhang G J, Shao C W, *et al.* Whole-genome sequence of a flatfish provides insights into ZW sex chromosome evolution and adaptation to a benthic lifestyle[J]. *Nature Genetics*, 2014, 46(3): 253-260.
- [22] Sved J A. Linkage disequilibrium and homozygosity of chromosome segments in finite populations[J]. *Theoretical Population Biology*, 1971, 2(2): 125-141.
- [23] Weir B S, Hill W G. Effect of mating structure on variation in linkage disequilibrium[J]. *Genetics*, 1980, 95(2): 477-488.
- [24] Zhao H, Nettleton D, Soller M, *et al.* Evaluation of linkage disequilibrium measures between multi-allelic markers as predictors of linkage disequilibrium between markers and QTL[J]. *Genetical Research*, 2005, 86(1): 77-87.
- [25] Barrett J C, Fry B, Maller J, *et al.* Haploview: analysis and visualization of LD and haplotype maps[J]. *Bioinformatics*, 2005, 21(2): 263-265.
- [26] Song W T, Li Y Z, Zhao Y W, *et al.* Construction of a high-density microsatellite genetic linkage map and mapping of sexual and growth-related traits in half-smooth tongue sole (*Cynoglossus semilaevis*)[J]. *Plos One*, 2012, 7(12): e52097.
- [27] Hayes B J, Visscher P M, Mcpartlan H C, *et al.* Novel multilocus measure of linkage disequilibrium to estimate past effective population size[J]. *Genome Research*, 2003, 13(4): 635-643.
- [28] R Development Core Team. R: a language and environment for statistical computing[M]. Vienna, Austria: R Foundation for Statistical Computing, 2011.

- [29] Hill W G. Linkage disequilibrium among multiple neutral alleles produced by mutation in finite population[J]. *Theoretical Population Biology*, 1975, 8(2): 117-126.
- [30] 周丽青, 杨爱国, 柳学周, 等. 半滑舌鲷染色体核型分析[J]. 水产学报, 2005, 29(3): 417-419.
- Zhou L Q, Yang A G, Liu X Z, *et al.* The karyotype of the tonguefish *Cynoglossus semilaevis*[J]. *Journal of Fisheries of China*, 2005, 29(3): 417-419(in Chinese).

Effective population size estimation of half-smooth tongue sole (*Cynoglossus semilaevis*)

WANG Weifeng^{1,2}, CUI Yu^{2,3}, JIANG Li², WANG Yueling^{2,3}, WANG Huanling¹, LI Hengde^{2*}

(1. Key Laboratory of Agricultural Animal Genetics, Breeding and Reproduction of Ministry of Education, Key Laboratory of Freshwater Animal Breeding of Ministry of Agriculture and Rural Affairs, College of Fisheries, Huazhong Agricultural University, Wuhan 430070, China;

2. Centre for Applied Aquatic Genomics, Chinese Academy of Fishery Sciences, Beijing 100141, China;

3. National Demonstration Center for Experimental Fisheries Science Education, Shanghai Ocean University, Shanghai 201306, China)

Abstract: To understand the current population genetic structure of half-smooth tongue sole and explore the trend and status of effective population size (N_e) in half-smooth tongue sole, in this paper, 800 4-month-old half-smooth tongue sole were identified by gender-specific genetic markers, and 297 genetic females were selected. Then 64 416 SNP markers were obtained by simplified genome sequencing method (2b-RAD). The genome-wide linkage disequilibrium (LD) was used to obtain the linkage disequilibrium distribution on each chromosome. The effective population size was calculated based on the different physical distances between the markers for each historical generation. In order to analyze the genetic structure of the population of half-smooth tongue sole selected by natural and artificial means, the sizes of 7 different chromosomal fragment lengths were detected. And the results showed that the effective population size two generations ago was only about 29. The high-throughput 2b-RAD sequencing method was first applied to the estimation of the effective population size of aquatic animals, which provided a reference for the breeding work of other aquatic species. The effective population sizes decreased when the historical generation was closer. It is important for half-smooth tongue sole breeding and conservation to estimate the effective population size for current generation.

Key words: *Cynoglossus semilaevis*; simplified genome sequencing (2b-RAD); linkage disequilibrium; effective population size

Corresponding author: LI Hengde. E-mail: hengde.li@cafs.ac.cn

Funding projects: Special Funds for Chinese Academy of Fishery Sciences Basic Research Business (2016ZD0203, 2016B005)