

DOI: 10.3724/SP.J.1118.2016.15440

驯养、选育条件下尼罗罗非鱼群体的选择压力分析

唐首杰, 杨洁, 王成辉, 李思发, 赵金良

上海海洋大学, 农业部淡水水产种质资源重点实验室, 上海 201306

摘要: 家养动物是研究长期人工选择对动物基因组产生选择效应机制的独特对象, 尼罗罗非鱼(*Oreochromis niloticus*)是一种受人工干预(驯养、选育)历史较短的优良养殖对象, 可作为研究新近发生的人工干预对动物基因组产生影响的遗传机制的良好模型。本研究以 1 个尼罗罗非鱼埃及野生群体为对照组, 以 4 个“新吉富”罗非鱼选育系群体、2 个企业自主选育群体和 5 个驯养群体为实验组, 采用 3 种模型分析方法(岛屿模型、分级岛屿模型和贝叶斯似然法), 在 12 个微卫星位点上进行 F_{ST} -离群值点检测(F_{ST} -outlier test)。结果显示, 在本研究所分析的 12 个微卫星位点中, 4 个“新吉富”罗非鱼选育系群体在 2 个微卫星位点(OMO043, OMO114)受到了显著的正向选择压力($P<0.01$), 2 个企业自主选育群体在另外 2 个微卫星位点(OMO049, OMO100)受到显著的正向选择压力($P<0.01$), 而 5 个驯养群体只在 1 个微卫星位点(OMO013)受到了显著的正向选择压力($P<0.01$)。由此可见, 选育群体受到的正向选择位点数明显多于驯养群体, 选育群体与驯养群体受到正向选择的位点各异, 不同选育群体间受到正向选择的位点也各不相同。本研究结果表明, 不同的人工干预途径从不同的方向上对尼罗罗非鱼基因组产生了影响。

关键词: 尼罗罗非鱼; 野生群体; 驯养群体; 选育群体; 选择压力

中图分类号: S917

文献标志码: A

文章编号: 1005-8737-(2016)04-0900-14

遗传、变异、选择是推动生物进化的主要动力, 在生物的自然进化中, 自然选择致使生物朝着有利于物种生存和繁衍的方向发展, 而驯化和育种使动植物向着既适应特定生态环境又满足人类需要的方向发展。

研究表明, 许多家养动物(如狗^[1]、猪^[2]、马^[3]、山羊^[4]、牦牛^[5]、家蚕^[6]等)、农作物(如水稻^[7]、大豆^[8]、玉米^[9-10]等)的基因组深深留下了人工选择的烙印, 例如, 玉米基因组中有 2%~5%的基因在驯化和现代育种中受到了选择^[11]。家养动植物受到人工选择的遗传机制一直是进化研究的热点问题之一, 许多研究揭示了家养动植物在经历了漫长(几百甚至几千个世代)的人工选择后, 其基因组中产生了大量的受到选择作用的印记。因为在长期的人工选择过程中, 群体中的有利突变能

逐渐扩散并固定下来。尽管如此, 有关驯化初期动植物受到选择的遗传机制的研究很少, 仅在个别家养动物, 如大西洋鲑(*Salmo salar*)^[12]、狐狸^[13]等中有过报道。

尼罗罗非鱼(*Oreochromis niloticus*)是一种世界性的优良养殖对象, 自 1978 年被作为养殖品种引进到中国以后^[14], 经历了长期的人工驯养和人工选育, 产生了一系列的人工驯养群体(如被国内多家罗非鱼良种场及养殖企业作为保种和养殖的群体)和人工选育群体(如“吉富”品系^[15]、“新吉富”品系尼罗罗非鱼^[16]等)。而且, 经遗传改良的“新吉富”罗非鱼在短期内(17 年, 17 代)经历了高强度的人工选择, “新吉富”罗非鱼选育的目标性状包括生长速度、体型和尾鳍条纹整齐度等, 其中生长性状的选择响应高达每代 5%左右^[17]。因此, 尽

收稿日期: 2015-11-23; 修订日期: 2016-01-29.

基金项目: 国家现代农业产业技术体系建设专项(CARS-49-04B); 水产动物遗传育种中心上海市协同创新中心(ZF1206); 上海高校青年教师培养资助计划(ZZhy12004).

作者简介: 唐首杰(1981-), 男, 博士, 讲师, 从事水产动物种质资源与种苗工程研究. E-mail: sjtang@shou.edu.cn

管驯养和选育的时间较短, 尼罗罗非鱼选育群体和野生群体间具有明显的差别。与此同时, 在人工繁殖和养殖过程中, 包含着有意或无意的选择, 人工养殖过程中无意的选择对驯养群体来说可能是不利的, 通常会使驯养群体产生近亲繁殖, 导致其适合度下降。因此, 尼罗罗非鱼可用作研究当前的、正在进行的驯化的遗传机制的模型。尼罗罗非鱼在长期的人工驯养和人工选育过程中, 可能经历了近交、遗传瓶颈和选择等遗传学事件, 遗传瓶颈对基因组中各个基因位点的作用是相同的, 而选择作用只对人工选择的靶基因及其连锁区域产生影响, 例如, 选择作用可能在尼罗罗非鱼基因组中的基因编码区、基因调控区或其他非编码区留下特异性的“痕迹”, 致使驯养群体和选育群体在基因组层面相对于野生群体而言发生了改变。

目前, 许多方法被用来识别基因组中受到选择的区域^[18], 这些方法大体被分成两大类, 一类方法基于比较数据(来自不同物种), 能检测出古老的选择印记; 另一类方法基于群体基因组学数据, 能检测出新近发生的选择。后一类方法在最近 10 年来变得非常流行, 已被用于许多非模式生物的研究中^[19]。

利用群体基因组学数据来检测新近发生的选择作用的这些研究方法都是以 Lewontin 等^[20]提出的思想为前提的。其基本原理是: 受定向选择的位点比中性位点表现出更大的遗传分化, 受平衡选择的位点比中性位点表现出更低的遗传分化。因此, 这类方法旨在识别 F_{ST} 值显著偏离中性假设的位点, 这些位点被称为“离群值点”(outlier loci)^[21-24]。

随着近期基因组研究的进展, 许多高密度分布于基因组中的新型标记已被用于鉴定基因组中受到选择的靶位点。特别是许多生物体基因组全序列的测定, 基于群体基因组学数据的方法获得了更加广泛的应用。因为微卫星标记在基因组中均匀分布, 且多态信息含量丰富^[25], 尽管它们不可能成为人工选择的靶位点, 但与基因组中受选择位点相连锁的微卫星标记可能偏离中性预期,

因此, 微卫星标记常用来鉴定基因组中受到选择的区域^[26-32]。

本研究利用 12 个微卫星位点, 以尼罗罗非鱼埃及野生群体为对照, 对尼罗罗非鱼 5 个驯养群体和 6 个选育群体进行选择压力检测, 试图找出这些群体中最近因为人工干预(驯养或选育)而受到选择的印记。

1 材料与方法

1.1 实验材料

本研究采用的尼罗罗非鱼土著群体(以下简称“野生群体”)、引进后驯养群体(以下简称“驯养群体”)、选育良种群体(以下简称“选育群体”)的来源和采样情况如表 1 所示。其中, GD 群体是 2002 年广西国家级罗非鱼良种场从广东三高省级罗非鱼良种场引进; 2002 年广西国家级罗非鱼良种场转引的群体有 EGY 群体和 USA 群体, 其中 EGY 群体是 1999 年淡水渔业研究中心从埃及农业和农垦部水产研究中心实验室引进并经多代选育, USA 群体是 1992 年淡水渔业研究中心从美国奥本大学引进并繁育至今; GLD 群体为 2008 年广西国家级罗非鱼良种场从美国关岛大学引进。JNM、BL 群体均于 2009 年采自海南新吉水产科技有限公司, 分别为挪威吉诺玛有限公司和海南宝路吉富水产科技有限公司自主选育的品系, 它们的选育基础群体均为吉富(“GIFT”)品系尼罗罗非鱼。WY 群体于 2013 年采自国家级伟业罗非鱼良种场, 为该场 2008 年从苏丹境内白尼罗河与青尼罗河交汇处引入的尼罗罗非鱼群体。F₁₄、F₁₅、F₁₆、F₁₇ 群体为上海海洋大学的“新吉富”罗非鱼选育系 F₁₄(2010 年选育)、F₁₅(2011 年选育)、F₁₆(2012 年选育)、F₁₇(2013 年选育)。EW 群体为埃及野生群体, 采自埃及水产养殖中心。采集每个群体样本 30 尾, 剪取每尾样本的鳍条, 于 95%酒精中保存固定。运回实验室, 用于下一步的基因组 DNA 提取。

1.2 实验方法

1.2.1 基因组 DNA 提取 基因组 DNA 提取采用常规的“酚-氯仿”方法进行^[33]。

表 1 尼罗罗非鱼土著群体、驯养群体和选育群体来源和采样情况表

Tab. 1 Sampling information and origination of native population, domesticated population and selective breeding population of *Oreochromis niloticus*

名称与代号 population code	采集地 collection location	样本数 sample size	性质 property
广东群体	GD	30	驯养群体 domesticated population
埃及群体	EGY	30	
美国群体	USA	30	
吉拉达群体	GLD	30	
伟业苏丹群体	WY	30	驯养群体 domesticated population
吉诺玛群体	JNM	30	
宝路群体	BL	30	
“新吉富”选育系 F ₁₄ 群体	F ₁₄	30	选育群体 selective breeding population
“新吉富”选育系 F ₁₅ 群体	F ₁₅	30	
“新吉富”选育系 F ₁₆ 群体	F ₁₆	30	
“新吉富”选育系 F ₁₇ 群体	F ₁₇	30	
埃及野生群体	EW	30	土著群体 native population

1.2.2 微卫星引物的筛选和多态性引物的 PCR 扩增 本实验所使用的微卫星位点均来自 NCBI (National Center for Biotechnology Information, <http://www.ncbi.nlm.nih.gov/>) 数据库, 利用 SSR Hunter 软件从 NCBI 数据库中搜索尼罗罗非鱼基因组中含有三核苷酸和四核苷酸核心序列的微卫星位点, 然后利用引物设计软件 Primer 5.0 软件设计引物, 共设计引物 24 对, 所有引物由上海生工生物工程有限公司合成。

对 24 对引物进行 PCR 扩增, PCR 产物经 8.0% 非变性聚丙烯酰胺凝胶电泳, 银染法检测^[33]。根据电泳结果中多态性的有无和扩增条带的清晰度来筛选引物, 筛选其中多态性高且扩增清晰稳定的引物作为后续实验的引物。

经过筛选得到 12 对高度多态的微卫星引物(表 2), 其中, 11 对引物的微卫星原始序列(OMO013、OMO040、OMO043、OMO049、OMO050、OMO092、OMO100、OMO114、OMO129、OMO232、OMO361) 分别来自罗非鱼的 10 个连锁群^[34], 另外一对引物的微卫星原始序列(UNH993)来自 Carleton 等^[35]。以所有群体的基因组 DNA 为模板, 对这 12 对微卫星引物进行 PCR 扩增, PCR 扩增反应体积为 10 μ L, 包括: 正向引物 0.5 μ L(5 μ mol/L), 反向引物 0.5 μ L(5 μ mol/L), 2 $\times$ PCR Reagent 5 μ L, DNA 模板 1 μ L, 无菌水 3 μ L。PCR 反应程序为: 94 $^{\circ}$ C

预变性 5 min, 接下来进行 35 个循环, 每个循环包括 94 $^{\circ}$ C 变性 30 s, 退火 45 s(各对引物退火温度见表 2), 72 $^{\circ}$ C 延伸 1 min; 最后一次循环结束后在 72 $^{\circ}$ C 延伸 10 min。

1.2.3 毛细管电泳分析 采用 QIAxcel 全自动毛细管核酸分析系统(QIAGEN 公司)对所有样本的 PCR 产物进行毛细管电泳分析, 分析参数设置参照仪器使用说明书。利用 2 种 DNA 分子量标准, 分别为 pBR322DNA/Msp I 和 Alignment Marker (15 bp 和 500 bp), 对 PCR 产物进行分子量估算。

1.2.4 数据统计 首先利用 MICRO-CHECKER version2.2.3 软件^[36]对得到的微卫星数据进行无效等位基因(null alleles)检测。随后将所有群体每个样本在 12 个微卫星位点上的基因型数据(等位基因片段)按 CREATE 软件^[37]可识别的格式输入 Microsoft Excel 文件中, 用 CREATE 软件将此数据格式文件分别转换为 ARLEQUIN version3.5.2.2 软件^[38]、FSTAT version2.9.3 软件^[39]和 GENEPOP version4.0 软件^[40]可识别的格式文件。利用 PGDSpider version 2.0.5.1 软件^[41]将 ARLEQUIN version3.5.2.2 软件^[38]可识别的格式文件转换成 BayeScan version2.1 软件^[42]可识别的格式文件。

采用 FSTAT version2.9.3 软件^[39]对所选用的 12 个微卫星位点进行连锁不平衡(linkage disequilibrium, LD)显著性检验, 以确定位点间是否为独立遗传。

表 2 本研究使用 12 个尼罗罗非鱼微卫星位点的特性
Tab. 2 Characteristics of the 12 microsatellite markers for *Oreochromis niloticus*

位点 locus	核心序列 repeat motif	引物序列(5'-3') primer sequence(5'-3')	退火温度/℃ annealing temperature	等位基因/bp allele size range	序列号 GenBank accession No.
OMO013	(TTTC) ₁₄	F: TCCCTGCGTATTATCTTT R: ATCCCTGGGCTAACTTCA	48	106~162	JX204845.1
OMO040	(CAT) ₁₃	F: GATGAACACTAAACTGTCCCTG R: TGTGCTGCTTCTCGCTCT	50	298~346	JX204855.1
OMO043	(TTTG) ₈ (AC) ₅	F: CATTCGGTTTATTGGTTATTGC R: TCATGCCTTCACCCCTCCC	50	298~340	JX204857.1
OMO049	(AGGA) ₁₀	F: TACTGCTGCCGTCCTTTA R: AACTGCCTGTGAGTGTCT	50	361~421	JX204860.1
OMO050	(TGA) ₁₃	F: TTAAGTGCCCAAGTTTGTT R: CATTCCCTCTGATTACGA	49	108~141	JX204861.1
OMO092	(TTC) ₈ TTT(TC) ₅	F: TATTGAGGCTTAGTTGGG R: GAAATGTAAGTGGCAGAT	50.4	135~174	JX204890.1
OMO100	(TAT) ₁₈	F: GGCAGTCTAAGCCCTAC R: ACATAAACAGCCGACGAG	52.8	206~248	JX204895.1
OMO114	(GTT) ₁₁	F: CAGTGGAAGACGCCRAGT R: TGRGACCAGCTCAGAGGC	51.7	238~280	JX204905.1
OMO129	(CCAT) ₉	F: GATCACTACTTTGTAGCTGGAT R: GAGCGAATGGTTGTCTGT	51.7	185~305	JX204914.1
OMO232	(AGAA) ₉	F: TTACATTCCCTGTCTTTAC R: GTTTGTTCTGGTGCTTAT	49	200~256	JX204976.1
OMO361	(TTC) ₉ TG(CTT) ₇	F: GTATGACAGCGAGCCAGAA R: AGCCCATCACTAAAGAAAGC	50	305~380	JX205069.1
UNH993	(CA) ₁₂ (CACG) ₅	F: TAAATGAAGGCAGGAGTA R: GGTGCGTAAGACGGTGTT	48	207~315	G68272.1

采用 3 种算法来进行 F_{ST} -离群值检验, 以检测受到正向选择的位点。第 1 种算法基于 Beaumont 等^[22]提出的岛屿模型(Island Model), 使用 LOSITAN version1.44 软件^[43]构建遗传分化指数(F_{ST})与杂合度(heterozygosity)的条件联合分布曲线图, 超出 99%置信区间上限的位点为受到强烈正向选择的位点。LOSITAN 软件参数设置为: 选择逐步突变模型(stepwise mutation model, SMM), 模拟 50000 次, 置信区间为 99%。

第 2 种算法基于 Excoffier 等^[44]提出的分级岛屿模型(Hierarchical Island Model), 在 ARLEQUIN version3.5.2.2 软件^[38]上进行运算, 由 R 函数(<https://www.R-project.org/>)构建遗传分化指数(F_{ST})与杂合度(heterozygosity)的条件联合分布曲线图, 超出 99%置信区间上限的位点为受到强烈正向选择的位点。参数设置为: 选择逐步突变模型(SMM), 模拟 20000 次, 置信区间为 90%、95%和 99%。

第 3 种算法为贝叶斯似然法(Bayesian Like-

lihood Method), 通过可逆跳跃马尔科夫链蒙特卡罗算法(Reversible Jump Markov chain Monte Carlo)获得后验分布(Posterior Distributions)^[45], 参数设置为 burn-in= 50000 次; chain length =500000 次。在 BayeScan version2.1 软件^[42]上完成运算。根据 Jeffrey^[46]的方法, 贝叶斯因子(Bayes factors, BF)大于 10(或 $\lg BF>1$)时, 表明某位点受到了强烈的正向选择。

以上 3 种算法所检测到的所有受正向选择的位点中, 至少被两种算法同时检测到的位点可判定为受到强烈正向选择的位点。

2 结果与分析

2.1 毛细管电泳及基因分型结果

毛细管电泳能直接从峰形图上判读出杂合子(图 1)和纯合子(图 2), 并通过与 DNA 标准分子量进行比对后得出每个峰对应的片段长度。引物 OMO129 在 BL 群体中的第 9 号样本的电泳结果

如图 1 所示, 引物 OMO129 在 BL 群体中的第 2 号样本的电泳结果如图 2 所示。引物 OMO129 在“新吉富”选育系 F₁₇ 群体第 1~12 号样本和第 13~24 号样本的电泳结果如图 3 和图 4 所示。

采用 MICRO-CHECKER version2.2.3 软件对 12 个群体的每个微卫星位点进行无效等位基因检测, 发现所有位点均不存在无效等位基因。经成对位点间的连锁不平衡检验, 发现各位点间连锁

不平衡现象不显著($P>0.05$)。说明本研究使用的 12 个微卫星位点均是独立遗传的, 因此保留所有位点进行后续分析。

2.2 选择压力分析结果

本研究在对 12 个微卫星位点进行选择压力分析前, 先将受试群体进行了分组, 把所有群体按其来源和性质分为以下 4 种分组模式: 第 1 种模式, 将 1 个野生群体单独归为一组, 5 个驯养群

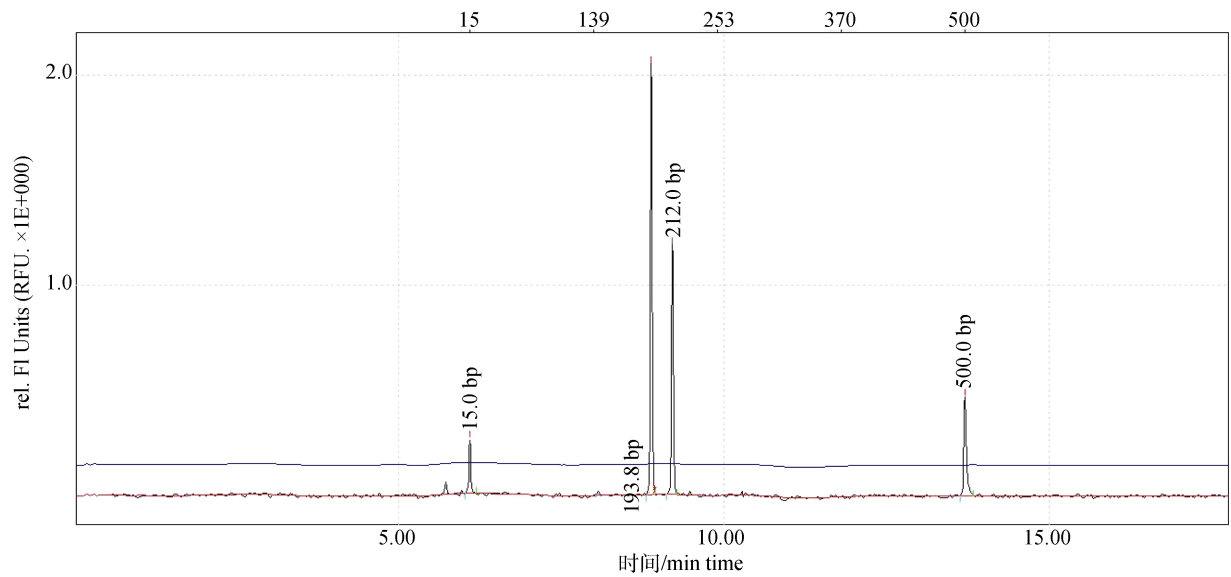


图 1 引物 OMO129 在 BL 群体中的第 9 号样本的电泳结果
Fig. 1 Electrophoresis results of locus OMO129 in the ninth sample of BL population

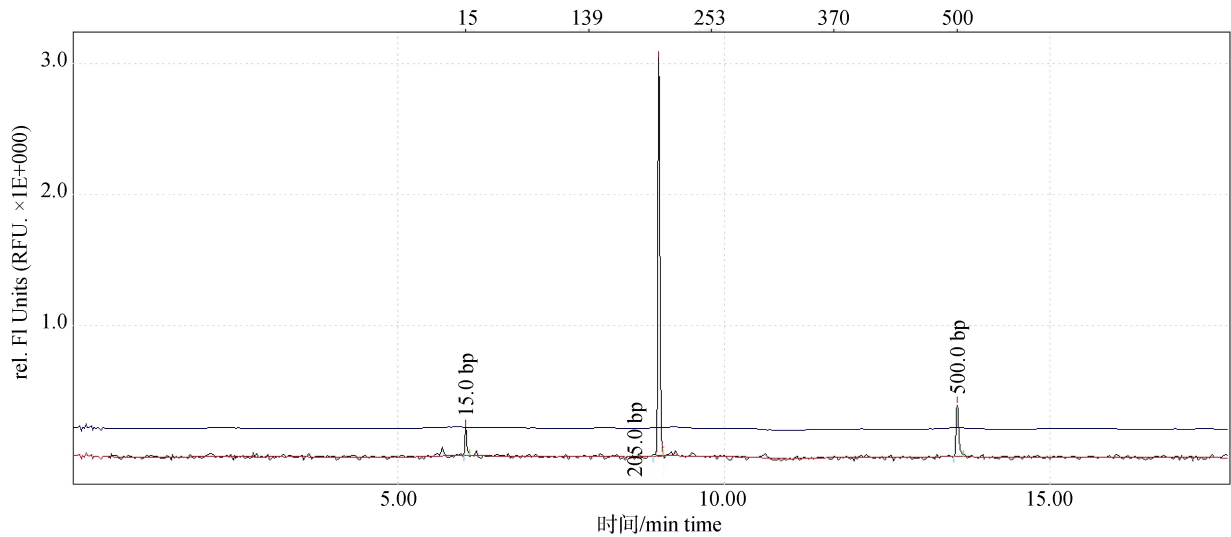


图 2 引物 OMO129 在 BL 群体中的第 2 号样本的电泳结果
Fig. 2 Electrophoresis results of locus OMO129 in the second sample of BL population

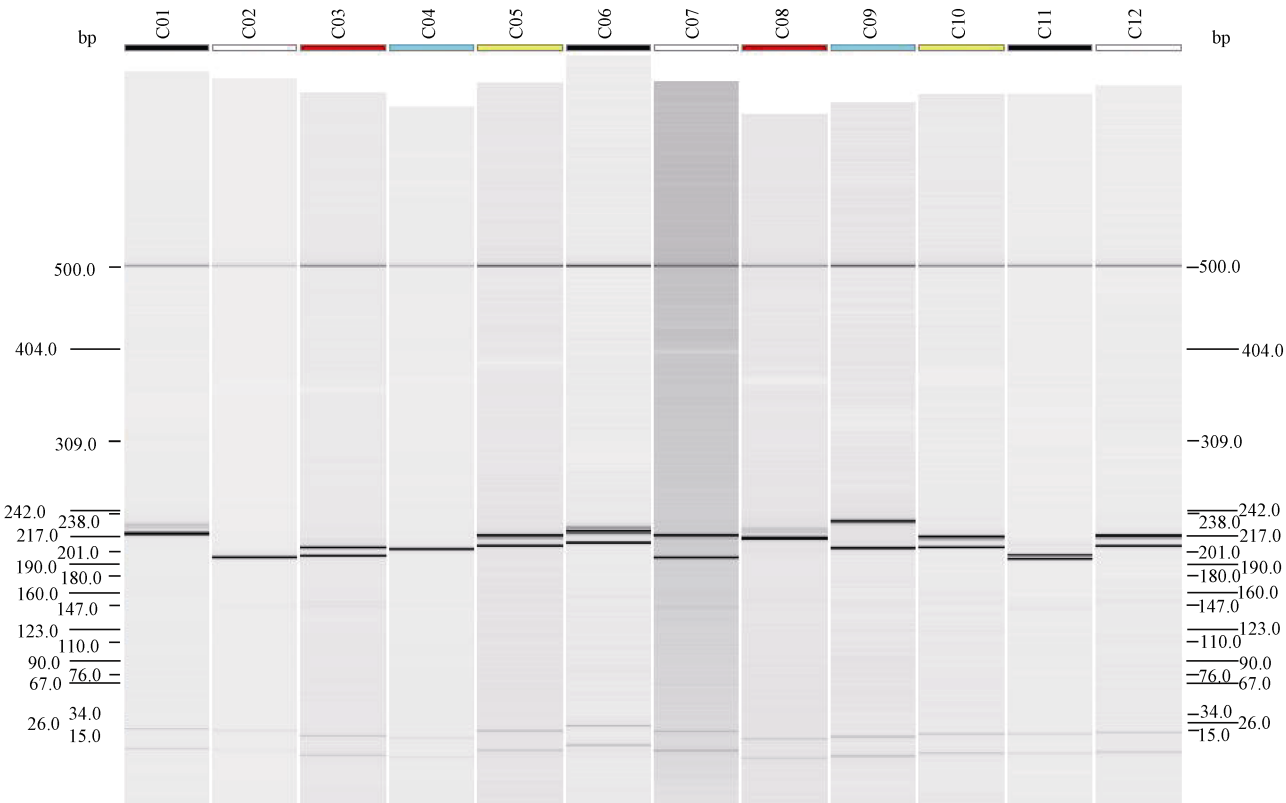


图 3 引物 OMO129 在“新吉富”选育系 F_{17} 群体第 1~12 号样本的电泳结果
Fig. 3 Electrophoresis results of locus OMO129 in the first to twelfth samples of F_{17} population

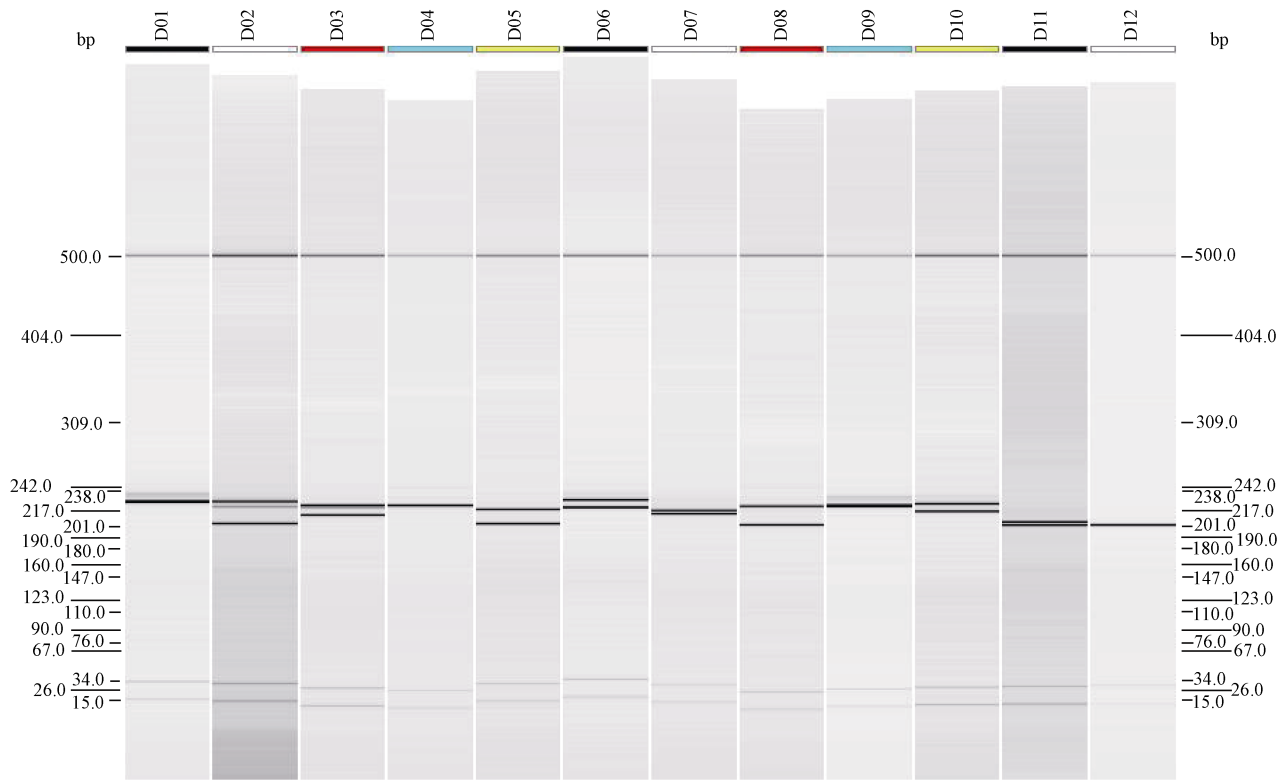


图 4 引物 OMO129 在“新吉富”选育系 F_{17} 群体第 13~24 号样本的电泳结果
Fig. 4 Electrophoresis results of locus OMO129 in the thirteenth to the twenty-fourth samples of F_{17} population

体归为一组, 4 个新吉富选育系群体(F_{14} 、 F_{15} 、 F_{16} 、 F_{17})归为一组, 2 个企业自主选育群体(BL、JNM)归为一组, 对所有群体进行分析; 第 2 种模式, 将 4 个新吉富选育系群体(F_{14} 、 F_{15} 、 F_{16} 、 F_{17})归为一组, 只对新吉富选育系群体进行分析; 第 3 种模式, 将 2 个企业自主选育群体(BL、JNM)归为一组, 只对企业自主选育群体进行分析; 第 4 种模式, 将 5 个驯养群体归为一组, 只对驯养群体进行分析。

分别应用 3 种软件(LOSITAN, ARLEQUIN, BAYESCAN)对 12 个微卫星位点进行选择压力分析的结果显示, 在第 1 种模式下, 3 种软件都在 OMO013 位点检测到了显著的正向选择压力, ARLEQUIN 软件和 LOSITAN 软件还同时在 OMO092 位点检测到了显著的正向选择压力(图 5a, 图 6a), 此外, LOSITAN 软件还在其他 4 个位点(OMO043, OMO100, OMO114, UNH993)检测到了显著的正向选择压力(图 5a), BayeScan 软件还在 OMO361 位点检测到了显著的正向选择压力(图 7a)。在第 2 种模式下, LOSITAN 软件和 ARLEQUIN 软件同时在 OMO043 位点和 OMO114 位点检测到了显著的正向选择压力(图 5b, 图 6b), 此外, BayeScan 软件分别在 OMO100 位点和 OMO129 位点检测到了显著的正向选择压力(图 7b)。在第 3 种模式下, LOSITAN 软件和 ARLEQUIN 软件同时在 OMO049 位点和 OMO100 位点检测到了显著的正向选择压力(图 5c, 图 6c), 此外, LOSITAN 软件还在 OMO013 位点和 UNH993 位点检测到了显著的正向选择压力(图 5c), BayeScan 软件在 OMO050 位点检测到了显著的正向选择压力(图 7c)。在第 4 种模式下, 3 种软件均在 OMO013 位点检测到了显著的正向选择压力(图 5d, 图 6d, 图 7d), 此外, LOSITAN 软件还在 OMO100 位点和 UNH993 位点检测到了显著的正向选择压力(图 5d)。综上所述, 在本研究所分析的 12 个微卫星位点中, 被两种软件同时检测到的位点判定为受到正向选择的位点(表 3), 因此, 新吉富选育系群体在 2 个微卫星位点(OMO043、OMO114)受到了显著的正向选择压力, 2

个企业自主选育群体在 2 个微卫星位点(OMO100、OMO049)受到了显著的正向选择压力, 而 5 个驯养群体只在一个微卫星位点(OMO013)受到了显著的正向选择压力, 选育群体受到的选择压力位点数多于驯养群体。

2.3 选择压力位点的生物信息学分析

为进一步了解选择压力位点在尼罗罗非鱼基因组中的位置和功能, 在 NCBI 中对 6 个受到选择压力的微卫星序列进行了核苷酸 BLAST 检索, 结果见表 4。6 条微卫星序列在 NCBI 中均检索到序列高度相似的 mRNA 片段, 相似度在 97%~99%, 这些 mRNA 在生物学过程中发挥着不同的功能, 如信号转导、转录调控等。

3 讨论

与驯化历史悠久的家养动物相比, 尼罗罗非鱼选育群体、驯养群体受人工选育、人工驯养的时间非常短暂(小于 40 年), 尽管如此, 在本研究所分析的 12 个微卫星位点中, “新吉富”选育系群体在 2 个微卫星位点(OMO043、OMO114)受到了显著的正向选择压力, 2 个企业自主选育群体在 2 个微卫星位点(OMO100、OMO049)受到了显著的正向选择压力, 5 个驯养群体在 1 个微卫星位点(OMO013)受到了显著的正向选择压力, 说明选育群体和驯养群体在人工干预条件下, 可能经历了快速进化。这与 Vasemagi 等^[12]对大西洋鲑的研究结果类似。为研究驯养动物基因组在驯化早期受到选择的遗传机制提供了又一参考。

事实上, 微卫星位点一般不可能是受到选择作用的靶位点, 但它们与候选功能基因间是紧密连锁的。本研究所使用的微卫星位点位于罗非鱼的 10 个连锁群上^[34], 它们大多位于候选功能基因附近, 生物信息学分析也进一步确认, 6 个受到选择的位点均与已知候选功能基因间存在高度相似的区段, 推测这 6 个位点与已知候选功能基因是紧密连锁的。

同时, 选育群体与驯养群体基因组中受到选择的位点也有差异, 例如, “新吉富”选育系群体在 OMO043 和 OMO114 位点受到了正向选择, 企

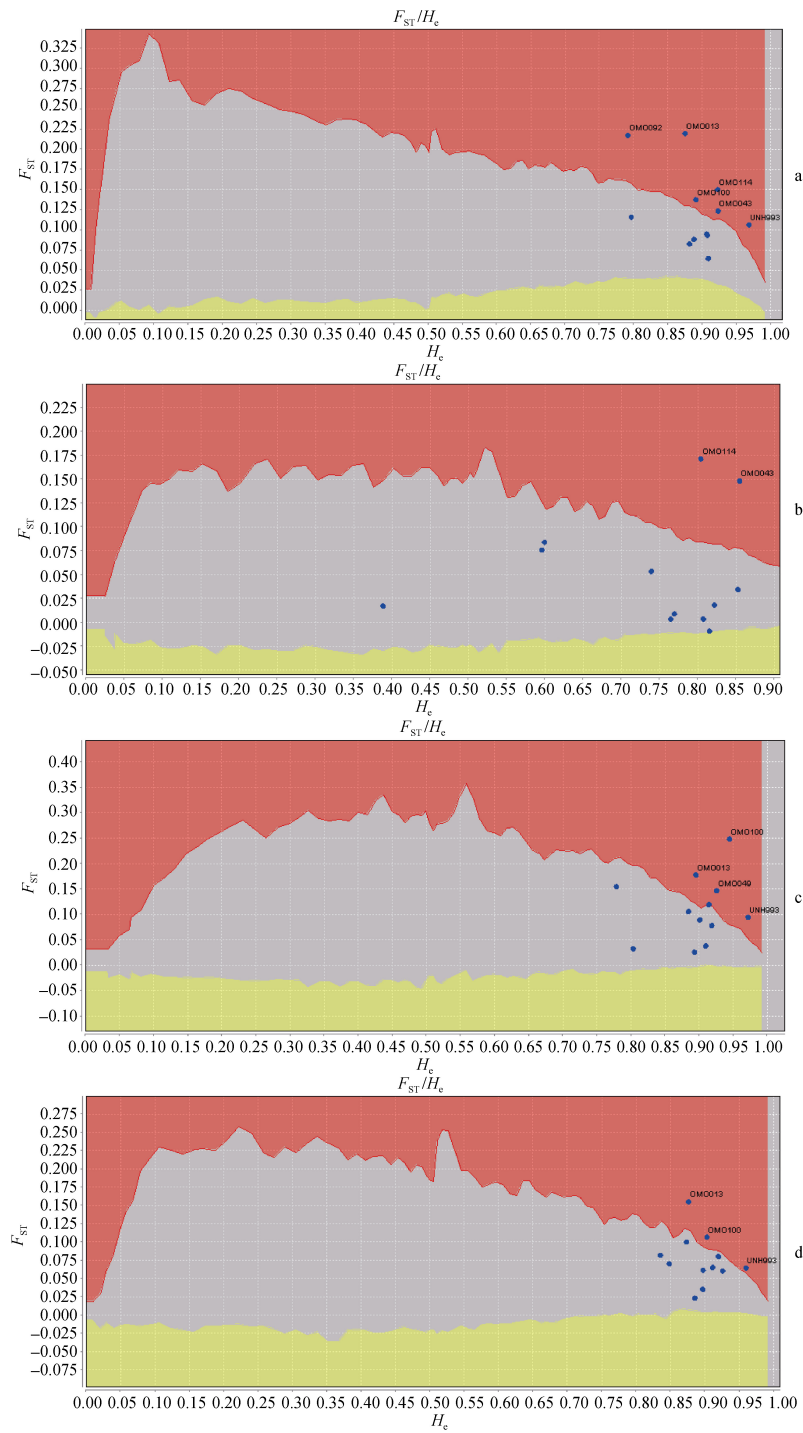


图 5 LOSITAN 软件在 4 种模式下的选择压力位点分析结果

a: 表示对所有群体的检测结果; b: 表示对 4 个“新吉富”选育系群体的检测结果; c: 表示对 2 个企业自主选育群体的检测结果; d: 表示对 5 个驯养群体的检测结果。红色区域为超出 99%置信区间上限的区域, 灰色区域为处于 99%置信区间以内的区域, 黄色区域为超出 99%置信区间下限的区域。红色区域内的点为受正向选择的位点, 灰色区域内的点为中性进化位点。

Fig. 5 Detection of loci under selection in four different modes by LOSITAN

a: Results for all populations; b: Results for four NEW GIFT strains of Nile tilapia; c: Results for two selective breeding populations developed by companies; d: Results for five domesticated populations. The area in red is the area beyond 99% upper confidence interval. The area in grey is the area within 99% confidence interval. The area in yellow is the area beyond 99% lower confidence interval. The loci under positive selection were in the red area. The neutral loci were in the grey area.

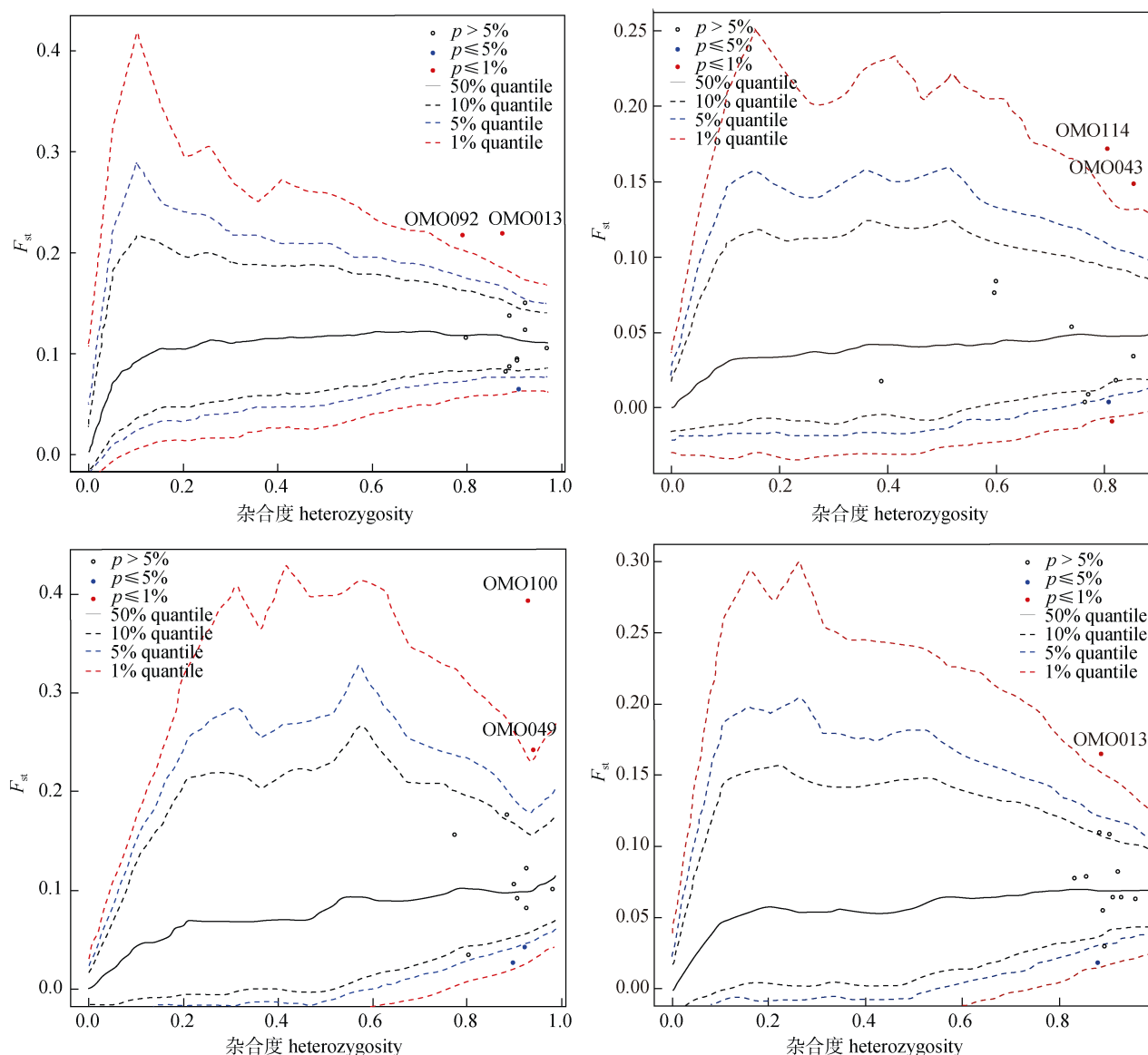


图 6 ARLEQUIN 软件在 4 种模式下的选择压力位点分析结果

a: 表示对所有群体的检测结果; b: 表示对 4 个“新吉富”选育系群体的检测结果; c: 表示对 2 个企业自主选育群体的检测结果; d: 表示对 5 个驯养群体的检测结果。红色虚线表示 99%置信度水平, 蓝色虚线表示 95%置信度水平, 黑色虚线表示 90%置信度水平, 黑色实线表示 50%置信度水平, 红色实心点表示超出 99%置信度水平的位点, 蓝色实心点表示处在 95%~99%置信度内的位点, 空心点表示处于 95%置信度以内的位点。

Fig. 6 Detection of loci under selection in four different modes by ARLEQUIN

a: Results for all populations; b: Results for four NEW GIFT strains of Nile tilapia; c: Results for two selective breeding populations developed by companies; d: Results for five domesticated populations. The dotted line in red represents 99% confidence interval. The dotted line in blue represented 95% confidence interval. The dotted line in black represents 90% confidence interval. The solid line in black represented 50% confidence interval. The solid point in red represents the locus beyond 99% confidence interval. The solid point in blue represented the locus within 95%~99% confidence interval. The hollow point represents the locus within 95% confidence interval.

业自主选育群体在 OMO100 和 OMO049 位点受到了正向选择, 而驯养群体只在 OMO013 位点受到了正向选择, 说明不同的人工干预途径(选育、驯

养)从不同的方向上对尼罗罗非鱼基因组产生影响。人工选育对选育群体的干预是有目的、有计划的, 而人工驯养对驯养群体的干预是随机的、无

意识的, 其基因组中受到选择的区域也可能是随机的, 这可能使驯养群体与选育群体在基因组层次上发生显著的遗传分化。这一现象在大西洋鲑选育群体和驯养群体基因组中也有类似的报道^[12]。

表 3 3 种算法检测到的受正向选择位点统计表
Tab. 3 Outlier loci detected under positive selection by three algorithms

算法 algorithms	所有群体 all populations	4 个“新吉富”选育系群体 four NEW GIFT strains of Nile tilapia	2 个企业自主选育群体 two selective breeding populations developed by companies	5 个驯养群体 five domesticated populations
分级岛屿模型 Hierarchical Island Model	<i>OMO013, OMO092</i>	<i>OMO043, OMO114</i>	<i>OMO049, OMO100</i>	<i>OMO013</i>
岛屿模型 Island Model	<i>OMO013, OMO043, OMO092, OMO100, OMO114, UNH993</i>	<i>OMO043, OMO114</i>	OMO013, <i>OMO049, OMO100, UNH993</i>	<i>OMO013, OMO100, UNH993</i>
贝叶斯似然法 Bayesian Likelihood Method	<i>OMO013, OMO361</i>	OMO100, OMO129	OMO050	<i>OMO013</i>

注: 加粗且倾斜的位点为至少被两种软件同时检测到受正向选择的位点。
Note: Loci in bold and italic indicated they were detected under positive selection at least twice by two kinds of softwares.

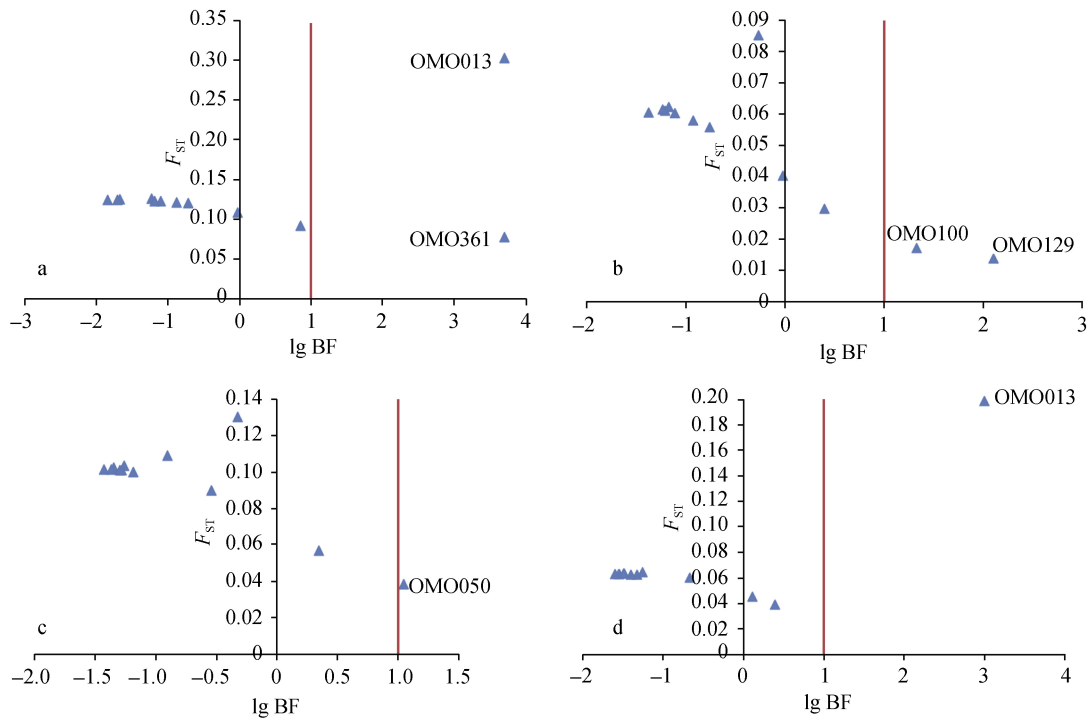


图 7 BayeScan 软件在 4 种模式下的选择压力位点分析结果

a: 表示对所有群体的检测结果; b: 表示对 4 个新吉富选育系群体的检测结果; c: 表示对 2 个企业自主选育群体的检测结果; d: 表示对 5 个驯养群体的检测结果。红色实线右侧的三角形图标为受到正向选择的位点。

Fig. 7 Detection of loci under selection in four different modes by BayeScan
a: Results for all populations; b: Results for four NEW GIFT strains of Nile tilapia; c: Results for two selective breeding populations developed by companies; d: Results for five domesticated populations. The triangle on the right side of red solid line represents the loci under positive selection.

值得注意的是, “新吉富”罗非鱼选育系群体和企业自主选育群体同为受人工选育的群体, 但他们受到的选择压力位点不同, 究其原因, 可能与选育基础群体来源及选育手段有关。“新吉富”罗非鱼选育群体的基础群体为 1994 年从国际水生生物资源管理中心(ICLARM)引进的尼罗罗非鱼吉

表 4 与已知基因显著相关的选择压力位点
Tab. 4 Candidate loci under selection with significant similarity to known genes

位点名称 locus name	BLAST 最佳匹配 BLAST hit	<i>E</i>	相似度/% identity	基因功能 gene function
OMO043	<i>Oreochromis niloticus</i> adaptor-related protein complex 1, gamma 1 subunit (<i>ap1g1</i>), transcript variant X2, mRNA XM_005467055.2	0.0	97	细胞间蛋白转运等
OMO049	<i>Oreochromis niloticus</i> epoxide hydrolase 2, cytoplasmic (<i>ephx2</i>), transcript variant X1, mRNA, XM_003454680.3	0.0	98	催化内源性和外源性的环氧化物水解
OMO114	<i>Oreochromis niloticus</i> actin filament-associated protein 1-like 2 (LOC100697549), transcript variant X1, mRNA XM_005460066.2	0.0	99	信号转导等
OMO100	<i>Oreochromis niloticus</i> small ubiquitin-related modifier 3 (LOC100690637), transcript variant X2, mRNA XM_005477195.2	0.0	98	转录调控等
OMO013	<i>Oreochromis niloticus</i> transmembrane and coiled-coil domain family 1 (<i>tmcc1</i>), transcript variant X13, mRNA XM_013275949.1	0.0	98	基因表达调控等
OMO092	<i>Oreochromis niloticus</i> cdc42 effector protein 4-like (LOC100697350), transcript variant X3, mRNA XM_005454977.2	0.0	99	肌动蛋白细胞骨架组装等

富(GIFT)品系^[16],自1997年至今,采用群体选育方法,以生长速度、形态特征和遗传纯度为选择指标,每年选育1代,从鱼苗培养到亲鱼,共进行4次选择,总选择率为6%^[47]。在长期的人工选育过程中,选育群体基因组中与生长、体型等相关功能基因相连锁的区域可能受到了选择。BL(宝路)罗非鱼选育群体是以挪威皇家极品水产有限公司从挪威国家水产遗传研究中心引进的吉富罗非鱼第14代家系为基础群体^[48],由海南宝路吉富水产科技有限公司通过累代的家系选育技术获得的新品系。JNM(吉诺玛)罗非鱼选育群体的基础群体为吉富品系尼罗罗非鱼,是国际水生生物资源管理中心(ICLARM)与菲律宾、挪威等国科研单位合作,通过对4个非洲原产地(埃及、加纳、肯尼亚、塞内加尔)直接引进的尼罗罗非鱼品系与4个亚洲国家和地区(以色列、新加坡、泰国、中国台湾)养殖比较广泛的尼罗罗非鱼品系进行杂交并经综合逐代选育获得的优良品系。该品系由挪威吉诺玛(海南)有限公司从ICLARM引进,经过连续世代综合选育(群体、家系等)后,其生长速度、抗逆性和出肉率等指标得到了显著提高。不同人工选育条件下,尼罗罗非鱼各选育群体基因组中受到正向选择的位点各异,这一结果为研究人工控制条件下鱼类基因组的进化机制提供了参考。

参考文献:

[1] Akeya J M, Ruheb A L, Akeya D T, et al. Tracking footprints of artificial selection in the dog genome[J]. Proc Natl Acad Sci USA, 2010, 107(3): 1160–1165.

[2] Amaral A J, Ferretti L, Megens H J, et al. Genome-wide footprints of pig domestication and selection revealed through massive parallel sequencing of pooled DNA[J]. PLoS ONE, 2011, 6(4): e14782.

[3] Petersen J L, Mickelson J R, Rendahl A K, et al. Genome-wide analysis reveals selection for important traits in domestic horse breeds[J]. PLoS Genetics, 2013, 9(1): e1003211.

[4] Nomura K, Yonezawa T, Mano S, et al. Domestication process of the goat revealed by an analysis of the nearly complete mitochondrial protein-encoding genes[J]. PLoS ONE, 2013, 8(8): e67775.

[5] Wang Z F, Yonezawa T, Liu B, et al. Domestication relaxed selective constraints on the yak mitochondrial genome[J]. Mol Biol Evol, 2011, 28(5): 1553–1556.

[6] Guo Y, Shen Y H, Sun W, et al. Nucleotide diversity and selection signature in the domesticated silkworm, *Bombyx mori*, and wild silkworm, *Bombyx mandarina*[J]. J Insect Sci, 2011, 11: 155.

[7] Liu Q P, Wang H, Hu H C. Genome-wide identification and evolutionary analysis of positively selected miRNA genes in domesticated rice[J]. Mol Gen Genom, 2015, 290(2): 593–602.

[8] Li Y H, Zhao S C, Ma J X, et al. Molecular footprints of domestication and improvement in soybean revealed by whole genome re-sequencing[J]. BMC Genom, 2013, 14:

- 579.
- [9] Wright S I, Vroh B I, Schroeder S G, et al. The effects of artificial selection on the maize genome[J]. *Science*, 2005, 308(5726): 1310–1314.
- [10] Vigouroux Y, McMullen M, Hittinger C T, et al. Identifying genes of agronomic importance in maize by screening microsatellites for evidence of selection during domestication[J]. *Proc Natl Acad Sci USA*, 2002, 99(15): 9650–9655.
- [11] Yamasaki M, Wright S I, McMullen M D. Genomic screening for artificial selection during domestication and improvement in maize[J]. *Ann Bot*, 2007, 100: 967–973.
- [12] Vasemagi A, Nilsson J, McGinnity P, et al. Screen for footprints of selection during domestication/captive breeding of Atlantic salmon[J]. *Comp Funct Genom*, 2012, 2012 Article ID 628204.
- [13] Trut L, Oskina I, Kharlamova A. Animal evolution during domestication: the domesticated fox as a model[J]. *BioEssays*, 2009, 31: 349–360.
- [14] Li J L, Li S F. Introduction and research advances of *Oreochromis niloticus* in China mainland[J]. *Journal of Fisheries of China*, 2001, 25(1): 90–95. [李家乐, 李思发. 中国大陆尼罗罗非鱼引进及其研究进展[J]. *水产学报*, 2001, 25(1): 90–95.]
- [15] Li S F, Li J L. Introduction to a new breed-GIFT Nile tilapia (*Oreochromis niloticus* L.)[J]. *China Fisheries*, 1998(4): 27, 36. [李思发, 李家乐. 养殖新品种简介——吉富品系尼罗罗非鱼[J]. *中国水产*, 1998(4): 27, 36.]
- [16] Li S F, Cai W Q. Characteristics and feeding techniques of “NEW GIFT” Nile tilapia (*Oreochromis niloticus* L.)[J]. *Scientific Fish Farming*, 2008(5): 21–22. [李思发, 蔡完其. 全国水产原良种审定委员会审定品种——“新吉富”罗非鱼品种特点和养成技术要点[J]. *科学养鱼*, 2008(5): 21–22.]
- [17] Tang S J, He A Y, Li S F, et al. Comparison of growth performance among the late selected generations of F_{13} - F_{15} of NEW GIFT strain of Nile tilapia (*Oreochromis niloticus*) [J]. *Journal of Shanghai Ocean University*, 2013, 22(1): 1–6. [唐首杰, 何安元, 李思发, 等. “新吉富”罗非鱼选育后期世代 F_{13} - F_{15} 的生长性能比较研究[J]. *上海海洋大学学报*, 2013, 22(1): 1–6.]
- [18] Nielsen R C, Bustamante A G, Clark S, et al. A scan for positively selected genes in the genomes of humans and chimpanzees[J]. *PLoS Biol*, 2005(3): 976–985.
- [19] Wilding C S, Butlin R K, Grahame J. Differential gene exchange between parapatric morphs of *Littorina saxatilis* detected using AFLP markers[J]. *J Evol Biol*, 2001, 14: 611–619.
- [20] Lewontin R, Krakauer J. Distribution of gene frequency as a test of theory of selective neutrality of polymorphisms[J]. *Genetics*, 1973, 74: 175–195.
- [21] Bowcock A, Kidd J, Mountain J, et al. Drift, admixture, and selection in human evolution: a study with DNA polymorphisms[J]. *Proc Natl Acad Sci USA*, 1991, 88: 839–843.
- [22] Beaumont M A, Nichols R A. Evaluating loci for use in the genetic analysis of population structure[J]. *Proc Roy Soc London B-Biol Sci*, 1996, 263: 1619–1626.
- [23] Vitalis R, Dawson K, Boursot P. Interpretation of variation across marker loci as evidence of selection[J]. *Genetics*, 2001, 158: 1811–1823.
- [24] Beaumont M A, Balding D J. Identifying adaptive genetic divergence among populations from genome scans[J]. *Mol Ecol*, 2004, 13(4): 969–980.
- [25] Guichoux E, Lagache L, Wagner S, et al. Current trends in microsatellite genotyping[J]. *Mol Ecol Res*, 2011, 11: 591–611.
- [26] Vasemagi A, Nilsson J, Primmer C R. Express Sequence Tag-linked microsatellite as a source for gene-associated polymorphisms for detecting signatures of selection in Atlantic salmon (*Salmo salar* L.)[J]. *Mol Biol Evol*, 2005, 22: 1067–1076.
- [27] Li M H, Iso-Touru T, Laurén H, et al. A microsatellite-based analysis for the detection of selection on BTA1 and BTA20 in northern Eurasian cattle (*Bos taurus*) populations[J]. *Gen Select Evol*, 2010, 42: 32.
- [28] Kasapidis P, Silva A, Zampicini G, et al. Evidence for microsatellite hitchhiking selection in European sardine (*Sardina pilchardus*) and implications in inferring stock structure[J]. *Sci Mar*, 2012, 76(1): 123–132.
- [29] Xu L H, Wang C H, Wang J, et al. Selection pressures have driven population differentiation of domesticated and wild common carp (*Cyprinus carpio* L.)[J]. *Genet Mol Res*, 2012, 11(3): 3222–3235.
- [30] Wang L, Liu S, Zhuang Z, et al. Population genetic studies revealed local adaptation in a high gene-flow marine fish, the small yellow croaker (*Larimichthys polyactis*) [J]. *PLoS ONE*, 2013, 8(12): e83493.
- [31] Rhode C, Vervalle J, Merwe A B, et al. Detection of molecular signatures of selection at microsatellite loci in the South African abalone (*Haliotis midae*) using a population genomic approach[J]. *Mar Genom*, 2013, 10: 27–36.
- [32] Wu L, Ci Y J, Huang S, et al. Genetic comparison and selection pressure analysis in complete setlines breeding and wild populations of Chinese mitten crab (*Eriocheir sinensis*) [J]. *Journal of Fishery Sciences of China*, 2015, 22(2): 204–213.

- [吴廉, 慈元吉, 黄姝, 等. 中华绒螯蟹配套系育种群体与野生群体的遗传比较与选择压力分析[J]. 中国水产科学, 2015, 22(2): 204–213.]
- [33] Sambrook J, Russell D W. Molecular Cloning: A Laboratory Manual. 3rd edition[M]. New York: Cold Spring Harbor Laboratory Press, 2001.
- [34] Liu F, Sun F, Li J, et al. A microsatellite-based linkage map of salt tolerant tilapia (*Oreochromis mossambicus* × *Oreochromis* spp.) and mapping of sex-determining loci[J]. BMC Genom, 2013, 14: 58.
- [35] Carleton K L, Streelman J T, Lee B Y, et al. Rapid isolation of CA microsatellites from the tilapia genome[J]. Anim Genet, 2002, 33: 140–144.
- [36] Van Oosterhout C, Hutchinson W F, Wills D P M, et al. MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data[J]. Mol Ecol Notes, 2004, 4(3): 535–538.
- [37] Coombs J A, Letcher B H, Nislow K H. CREATE: a software to create input files from diploid genotypic data for 52 genetic software programs[J]. Mol Ecol Res, 2008, 8(3): 578–580.
- [38] Excoffier L, Lischer H E L. Arlequin suite ver 3.5: A new series of programs to perform population genetics analyses under Linux and Windows[J]. Mol Ecol Res, 2010, 10: 564–567.
- [39] Goudet J. FSTAT: A Program to Estimate and Test Gene Diversities and Fixation Indices (version 2. 9. 3). 2001. <http://www.unil.ch/izea/software/fstat.html>.
- [40] Raymond M, Rousset F. GENETPOP (version 1.2): population genetics software for exact tests and ecumenicism[J]. J Hered, 1995, 86: 248–249.
- [41] Lischer H E L, Excoffier L. PGDSpider: An automated data conversion tool for connecting population genetics and genomics programs[J]. Bioinformatics, 2012, 28: 298–299.
- [42] Foll M, Gaggiotti O. A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective[J]. Genetics, 2008, 180(2): 977–993.
- [43] Antao T, Lopes A, Lopes R J, et al. LOSITAN: A workbench to detect molecular adaptation based on a F_{st} -outlier method[J]. BMC Bioinform, 2008, 9: 323.
- [44] Excoffier L, Hofer T, Foll M. Detecting loci under selection in a hierarchically structured population[J]. Heredity, 2009, 103(4): 285–298.
- [45] Green P J. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination[J]. Biometrika, 1995, 82: 711–732.
- [46] Jeffrey H. Some tests of significance, treated by the theory of probability[J]. Proc Camb Philos Soc, 1935, 31: 203–222.
- [47] Li S F, He X J, Hu G C, et al. Improving growth performance and caudal fin stripe pattern in selected F_6 – F_8 generations of GIFT Nile tilapia (*Oreochromis niloticus* L.) using mass selection[J]. Aquac Res, 2006, 37: 1165–1171.
- [48] Ma D Y, Rye M, Wang Y X, et al. Genetic improvement of tilapias in China: Genetic parameters and selection responses in growth of Nile tilapia (*Oreochromis niloticus*) after six generations of multi-trait selection for growth and fillet yield[J]. Aquaculture, 2011, 322–323: 51–64.

Analysis of selective pressure on Nile tilapia (*Oreochromis niloticus*) populations during domestication/selective breeding

TANG Shoujie, YANG Jie, WANG Chenghui, LI Sifa, ZHAO Jinliang

Key Laboratory of Freshwater Fishery Germplasm Resources, Ministry of Agriculture; Shanghai Ocean University, Shanghai 201306, China

Abstract: Domesticated animals provide a unique opportunity to identify genomic targets for artificial selection to a captive environment. Nile tilapia (*Oreochromis niloticus*) is a useful model species for studying the genetic basis of recent, ongoing domestication, as reared Nile tilapia strains have experienced intense artificial selection for only a short period of time. In this study, we screened 12 microsatellite loci from 11 independently reared strains (6 used for selective breeding and 5 for ranching) and their wild progenitor population of Nile tilapia to identify recent selection footprints related to domestication and selective breeding. F_{ST} -outlier tests were implemented using three different genetic software programs (LOSITAN, ARLEQUIN, and BAYESCAN) to identify loci under positive selection. All approaches assumed that directional selection increases genetic differentiation between populations and reduces variability at linked loci. However, because all of the tests are based on different assumptions, identifying outlier loci simultaneously using all approaches provides additional support for the candidate status of a particular locus. The LOSITAN program uses coalescent simulations to generate a neutral joint distribution of F_{ST} and heterozygosity; loci with the highest F_{ST} : heterozygosity ratios are candidates for having experienced selection. Coalescent simulations were performed using 12 samples and a sample size of 30, assuming island and stepwise mutation models. The mean F_{ST} value was used with other values close to the mean to obtain half of the data points above and half of those below the median, as suggested in the software manual. The second method to detect selection footprints was based on the hierarchical island model and was implemented in ARLEQUIN software. Coalescent simulations were performed using 12 samples and a sample size of 30, assuming a stepwise mutation model. R-project software was used to generate a neutral joint distribution of F_{ST} and heterozygosity. The third method to detect selection footprints also identified loci that exhibit extreme differentiation compared with the remainder of the genome using the Bayesian likelihood method implemented via the reversible jump Markov chain Monte Carlo method. This approach was implemented in BayeScan software. F_{ST} was modeled using a logistic regression model and locus and population effects by relaxing the assumption of the symmetrical island model and allowing for asymmetries in population structure. The estimates of the model parameters were automatically adjusted during short pilot runs (10 pilot runs with 5000 iterations each). We used a burn-in of 50,000 iterations and a total chain length of 500,000 iterations to identify loci under selection. BayeScan estimates the posterior probability that a locus is under selection by calculating the Bayes factor, which is given by the ratio of the posterior model probabilities of two models (selected/neutral), given the data. According to Jeffreys' interpretation, if the Bayes factor BF is >10 [$\lg(\text{BF}) > 1$], "strong evidence" favors one model over the other and corresponds to a posterior probability >0.91 . The results showed that two loci (OMO043 and OMO114) provided significant evidence of positive artificial selection in four genetically improved NEW GIFT strain populations. Another two loci (OMO049 and OMO100) showed significant evidence of positive artificial selection in two selective breeding stocks produced by three companies. However, only one loci (OMO013) experienced positive artificial selection in five domesticated populations. The number of loci detected under selection in selective breeding populations was larger than that in domesticated populations. The loci detected under selection varied among the selective breeding and domesticated populations. These results indicate that the Nile tilapia genome has been affected by artificial selection in different directions.

Key words: *Oreochromis niloticus*; wild population; domesticated population; genetically improved population; selective pressure

Corresponding author: TANG Shoujie. E-mail: sjtang@shou.edu.cn