

# 恩施黑猪基因组群体遗传学参数的估计与选择信号研究

吴林慧, 孙琦, 王荔茹, 张凯丽, 谢胜松, 李新云, 赵书红, 马云龙\*

(华中农业大学 农业动物遗传育种与繁殖教育部重点实验室, 武汉 430070)

**摘要:** 旨在调查恩施黑猪基因组群体遗传学参数与选择信号。本研究利用 Porcine 80K SNP 芯片, 通过计算亲缘系数、近交系数、连锁不平衡程度与有效群体大小等群体遗传学层面的参数评估恩施黑猪种群结构关系, 运用 CLR 和 iHS 方法检测恩施黑猪基因组选择信号, 并通过生物信息学分析揭示其潜在受选择基因。亲缘系数计算发现, 咸丰县恩施黑猪个体间平均亲缘系数为 0.12; 近交系数计算表明, 16% 的样本近交系数大于 0.125, 存在明显近交累积。此外, 本研究构建了恩施黑猪全基因组连锁不平衡图谱; 利用连锁不平衡信息, 恩施黑猪估计历史有效群体大小呈现逐代下降趋势, 其中 5 世代前有效群体大小约为 25 头。基于 CLR 方法共检测到 126 个显著选择信号候选区域, 总长约为 51.6 Mb, 占基因组总长的 2.1%。利用 iHS 方法, 共发现 248 个显著选择信号候选区域, 长度约为 78.78 Mb, 约占基因组总长的 3.2%。富集分析表明, 与选择信号区域重叠的 *LPAR2*、*NDUFA13*、*MEF2B*、*AHR* 基因分别与胴体长度(滴水损失)、精子形成、骨骼肌分化和总产仔数相关。研究表明, 现存恩施黑猪群体血统较窄, 近交累积严重, 有效群体大小较小, 并呈现继续缩小的趋势。选择信号分析揭示的一系列潜在受选择基因能够为未来恩施黑猪的遗传改良提供一定的参考依据。

**关键词:** 恩施黑猪; 亲缘系数; 近交系数; 连锁不平衡; 有效群体大小; 选择信号

中图分类号: S828.2

文献标志码: A

文章编号: 0366-6964(2019)03-0485-10

## A Study of the Population Genetics Parameters and Selection Signatures in Enshi Black Pig

WU Linhui, SUN Qi, WANG Liru, ZHANG Kaili,

XIE Shengsong, LI Xinyun, ZHAO Shuhong, MA Yunlong\*

(Key Laboratory of Agricultural Animal Genetics, Breeding and Reproduction of the Ministry of Education, Huazhong Agricultural University, Wuhan 430070, China)

**Abstract:** This study was designed to investigate the population genetics parameters and selection signatures of Enshi Black pig. Using the Porcine 80K SNP chip, a series of population genetics parameters, including the relationship coefficient, the inbreeding coefficient, the linkage disequilibrium and the effective population size, were calculated to evaluate the population structure relationship of Enshi Black pig. In addition, the CLR and iHS methods were used to identify the potential selection signatures and then the corresponding selected genes were revealed by bioinformatics analysis. The results indicated that the mean value of relationship coefficient was 0.12 for all individual-pairs in Xianfeng county, the inbreeding coefficient of 16% of all individuals was higher than 0.125. In addition, this study presented a linkage disequilibrium map for the whole

收稿日期: 2018-09-05

基金项目: 国家自然科学基金(31790414; 31601916); 国家生猪产业技术体系项目(CARS-35); 中央高校基本科研业务费专项基金(2662015QD018)

作者简介: 吴林慧(1997-), 女, 湖北黄冈人, 主要从事分子数量遗传研究, E-mail: 525835111@qq.com; 孙琦(1996-), 男, 内蒙古兴安盟人, 主要从事分子数量遗传研究, E-mail: 2386725428@qq.com。吴林慧和孙琦为同等贡献作者

\* 通信作者: 马云龙, 主要从事分子数量遗传研究, E-mail: Yunlong.Ma@mail.hzau.edu.cn

genome of Enshi Black pig. Analysis of the past effective population size based on the information of linkage disequilibrium suggested a decline in effective population size, and up to  $Ne=25$  before 5th generation. Based on CLR method, a total of 126 significant potential selection signature regions were detected, which were approximately 51.6 Mb in length, accounting for approximately 2.1% of the total length of the genome. Similarly, a total of 248 significant potential selection signature regions were detected using iHS method, which spanned approximately 78.78 Mb and accounted for approximately 3.2% of the total genome length. Enrichment analysis showed that genes such as *LPAR2*, *NDUFA13*, *MEF2B* and *AHR* overlapping with the selection signatures region were associated with carcass length (drip loss), sperm formation, skeletal muscle differentiation and total litter size. This study suggests that the genetic basis of Enshi black pig is narrow and inbreeding is obvious, and the effective population size is limited, and it continues to shrink. The candidate genes revealed by sweep analysis can provide some references for the future genetic improvement of Enshi Black pig.

**Key words:** Enshi Black pig; relationship coefficient; inbreeding coefficient; linkage disequilibrium; the effective population size; selection signatures

随着社会经济的发展,人们对肉蛋奶的需求急剧增加。具有优良生产性能的商业品种在全世界范围得到广泛推广,并逐步取代了那些产能低下、集约化程度低的地方品种。据不完全统计,目前我国已有大约 55 个地方猪品种已经处于濒危状态<sup>[1]</sup>。但这些地方猪品种存在着许多优良的性状,是我国种质资源宝贵的财富。恩施黑猪便是该背景下中国地方猪品种的缩影。恩施黑猪因其脂肪储存能力和冷湿耐受性而闻名,肉质性状十分优良,脂肪含量高,被广泛用于干火腿和熏肉生产,加上中心产区的恩施黑猪与土司文化紧密相连,有着深厚的文化底蕴。根据中心产区咸丰县的调查,1981 年恩施黑猪母猪为 3 万头,2001 年为 4 万头,2005 年为 3.3 万头,2006 年为 2.2 万头,2007 年为 4.08 万头;恩施黑猪公猪 2001-2007 年,由于人工授精工作的开展,存栏始终保持在 3~4 个血统,共计 8 头(据《中国遗传资源志·猪志》,2011)。2013 年,从恩施市和咸丰县两地通过调查只收集到恩施黑猪公猪 14 头。目前,由于经费投入不足,恩施黑猪血统记录混乱,选配工作无序,纯种公猪逐年减少,品种保护形势不容乐观。

传统意义上,经典遗传参数的估计依靠系谱信息的详细记录,然而我国地方猪系谱信息普遍记录不全,这也使得我国地方猪三大经典遗传参数的估计鲜有报道。目前,关于猪相关性状遗传参数的研究主要见于西方瘦肉型猪品种及其商业杂交群体,主要包括肉质、生长等重要经济性状<sup>[2-3]</sup>。随着高密

度 SNP 芯片和重测序技术的应用,遗传参数的内涵被不断丰富,除三大经典遗传参数外,基因组杂合度、群体分化程度、基因组连锁不平衡程度、基因组近交系数与亲缘系数的计算已经被广泛应用于群体遗传学的研究<sup>[4-5]</sup>。同时,利用基因组数据进行种群结构分析、亲缘关系估计与基因组遗传参数的计算在畜禽种质资源的研究中也已得到广泛的应用<sup>[5-6]</sup>。一般而言,基因组亲缘系数与近交系数的估计有助于了解群体内亲缘关系状况,为合理制定群体选配计划提供重要的指导意见。基因组连锁不平衡程度与有效群体大小的估计,则有助于了解种群基因组结构特征,评估其历史种群规模。同时,畜禽长期的适应性进化使得不同畜禽品种间存在显著的表型差异,并逐渐形成了各自的典型品种特征。因此,通过选择信号的研究有助于揭示畜禽适应性进化的潜在遗传机制;有助于定位影响其优良性状的重要候选基因,为畜禽遗传改良提供重要参考<sup>[7-8]</sup>。

本研究基于 54 头恩施黑猪高密度 SNP 芯片数据,通过基因组群体遗传学参数的估计来评估恩施黑猪种群结构关系;利用选择信号分析鉴定恩施黑猪重要种质性状的潜在遗传基础,探索恩施黑猪适应性进化的遗传作用机制;旨在为恩施黑猪的保种与遗传改良工作提供科学的参考依据。

## 1 材料与方法

### 1.1 试验材料

本研究选取恩施州咸丰县全部血统的 54 头纯种

恩施黑猪个体构建资源群体,采集耳组织样,−20 °C 冷冻保存,用于提取全基因组 DNA。

## 1.2 试验方法

1.2.1 基因分型与质量控制 本研究利用 DNA 提取试剂盒(天根)提取耳组织全基因组 DNA 后,使用 Illumina Porcine SNP80K 芯片进行基因分型,共获得 68 516 个 SNPs 位点。经过严格质量控制,共有 33 429 个 SNPs 位点用于后续分析。其中,质量控制的标准为:1)剔除未知位置的标记位点;2)剔除性染色体上的标记位点;3)剔除最小等位基因频率低于 0.05 的标记位点。

1.2.2 利用基因组信息评估个体亲缘关系与近交情况 亲缘系数(relationship coefficient)是指两个有共同祖先的个体在某一基因座上具有相同等位基因的概率。本研究利用 PLINK 软件基于基因组同源相同(identity-by-descent)信息,评估个体间亲缘关系程度<sup>[9]</sup>。近交系数(inbreeding coefficient)是指某一个体任何基因座上的两个等位基因来自父母共同祖先同一基因的概率,也就是该个体由于父母近交而可能成为纯合子的概率。本研究基于常染色体基因组纯合片段(runs of homozygosity,ROH)占全部常染色体基因组总长度比例来评估个体近交系数,计算公式:

$$F_{ROH} = \frac{\sum L_{ROH_k}}{L_{genome}} \quad (1)$$

其中, $L_{ROH_k}$ 表示第  $k$  个 ROH 的长度, $L_{genome}$ 表示基因组总长,本研究为常染色体总长;基因组纯合片段利用 PLINK 软件完成,ROH 定义参数设置为:最小长度 200 kb,SNP 个数/ROH > 3,最小 SNP 密度 > 500 kb/SNP,最大间隔长度 < 1 000 kb<sup>[9]</sup>。

1.2.3 基因组连锁不平衡程度与有效群体大小估计 本研究运用 PLINK 软件计算各 SNP 对间的相关程度( $r^2$ ),用于评估恩施黑猪基因组连锁不平衡程度。参数设置为:—r2—ld-window-kb 1 000—ld-window 99 999—ld-window-r2 0。利用 Sved (1971)<sup>[10]</sup>构建有效群体大小与连锁不平衡程度间存在的近似关系,估计恩施黑猪历史有效群体大小,公式:

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

其中, $r^2$ 为不同分子标记间的连锁不平衡程度; $Ne$ 是  $1/(2c)$ 代之前的种群大小; $c$ 代表了摩尔根连锁遗传图谱距离; $\alpha$ 取值为 1 时,表示模型考虑突变,为 2 表示模型不考虑突变; $k$ 为 4 代表常染色体,2

则代表 X 染色体。本研究选取 100、200、500、1 000、2 000、5 000、10 000 kb 等不同 SNPs 间距,基于上述关系估计恩施黑猪在相应前 500、250、100、50、25、10、5 代的有效群体大小,1 cm 近似等于 1 Mb。

1.2.4 选择信号分析 选择信号检测能够反映群体适应性进化的遗传机制。本研究采用两种互补的方法进行恩施黑猪正向选择信号检测。其中,基于位点频率谱差异原理的 CLR(composite likelihood ratio)方法对即将固定的选择信号具有较高选择效力<sup>[11-12]</sup>,基于扩展单倍型纯合原理的 iHS(integrated haplotype score)方法对正在进行的选择信号具有较高的检测效力<sup>[13]</sup>。本研究利用 sweepfinder 和 rehh 软件,采用无重叠滑动窗口方法,每个窗口 75 k,进行选择信号检测。使用 CLR 和 iHS 方法分别计算获得 32 642 与 32 449 个统计量。极值判断法(outlier)被用于选择信号显著性检验,各统计量进行秩排序,其中大于 0.01 分位数阈值的统计量被定义为选择信号位点<sup>[14]</sup>。

1.2.5 选择信号富集基因的功能注释 由于目前关于猪的基因注释信息较少,除了对猪相关基因注释数据库分析外,将找到的候选基因通过 BioMart(<http://www.biomart.org>)数据库同人基因组进行同源比对,获得人方面相对应的直系同源基因集,然后使用 DAVID6.7(<https://david.ncifcrf.gov/home.jsp>)软件对同源基因集进行功能富集分析,分析内容涵盖了细胞组分(cellular component)、分子功能(molecular function)、生物学过程(biological process)以及生物学通路(pathway)分析。此外,本研究参考 Ensembl genes 发布的最新猪基因数据信息(<http://www.ensembl.org/biomart/martview>),对恩施黑猪潜在选择信号区域与猪的注释基因进行比对,并结合 NCBI 基因数据库(<http://www.ncbi.nlm.nih.gov/gene/>)挖掘对应基因功能注释信息。

## 2 结果

### 2.1 恩施黑猪亲缘关系与近交程度评估

亲缘系数与近交系数是评估群体内个体与个体间关系的重要指标,是畜禽群体选配与遗传资源保护的重要参考依据。基于芯片数据的分析表明,54 个个体间可以形成 1 431 种关系对;其中个体间平均亲缘系数为 0.12,约 34.34% 的个体间亲缘

系数小于等于 0.062 5, 约 16.92% 的个体对间亲缘系数小于等于 0.125 (表 1)。在共同祖先无近交的情况下, 半同胞交配产生后代的近交系数为 0.125, 全同胞后代的近交系数为 0.25。基于 54 头恩施黑猪个体芯片数据计算发现, 群体平均近交系数为 0.069, 约 16.67% 的样本近交系数大于等于 0.125, 约一半的个体近交系数小于等于 0.062 5 (表 1)。基于上述分析推测, 恩施黑猪个体间亲缘关系普遍较近, 个体具有明显的近交累积现象。

表 1 恩施黑猪基因组亲缘系数与近交系数

Table 1 The summary of relationship coefficients and inbreeding coefficients in Enshi Black pig

项目 Item	亲缘系数 Relationship coefficient	近交系数 Inbreeding coefficient
均值±标准差 Mean±sd	0.12±0.11	0.069±0.06
众数 Mode	0	0.009
中位数 Median	0.11	0.062
≥0.5	0.53(0.63%)	0
0.25~0.5	0.30(13.99%)	0
0.125~0.25	0.18(32.94%)	0.17(16.67%)
0.062 5~0.125	0.10(16.92%)	0.10(33.33%)
≤0.062 5	0.01(34.34%)	0.016(50.00%)

## 2.2 恩施黑猪基因组连锁不平衡程度与有效群体大小估计

利用恩施黑猪 18 条常染色体基因组信息, 本研究通过计算各染色体 10 Mb 范围内全部 SNPs 对间的连锁不平衡程度来评估恩施黑猪基因组特征。共获得约 52 万对 SNPs 的  $r^2$  估计值。根据 SNP 间不同的标记间距, 将上述连锁不平衡程度  $r^2$  值进行分类分析发现, 随着 SNP 对间距的扩大, 平均  $r^2$  值呈现衰减趋势; 尽管各染色体 LD 衰减的速率存在一定差别, 但是整体衰减趋势相似, 且衰减速率较为平缓 (图 1)。整理各标记间距类别发现, 标记间距在 0.025 Mb 范围内的  $r^2$  均值为 0.28, 变异幅度最大, 其中 36.70% 的 SNPs 对的  $r^2 \geq 0.25$ ; 其后的间距类别,  $r^2$  均值逐渐减小, 变异幅度逐渐收窄, 且  $r^2 \geq 0.25$  的比例逐渐降低。

基于公式 (2), 本研究计算了 100、200、500 kb 等 SNP 间距的连锁不平衡程度, 分别估计了恩施黑猪区域重叠, 主要包括胴体长度、繁殖力、猪肉品质等。

猪群体在 500、250 和 100 代等世代的有效群体大小。总体趋势上, 在非变异模型下恩施黑猪有效群体规模从 500 代以前的  $N_e=891$  下降到 5 代以前  $N_e=28$ , 在变异模型下, 恩施黑猪有效群体规模从 500 代以前的  $N_e=640$  下降到 5 代以前的  $N_e=25$ , 说明恩施黑猪有效群体呈现逐代下降的趋势 (图 2)。利用 54 头恩施黑猪芯片数据, 估计恩施黑猪 20 年前 (50 世代) 有效群体大小约为 135 头, 其与目前报道的其他中国地方猪有效群体大小基本一致, 明显低于商品猪有效群体大小<sup>[15]</sup>。

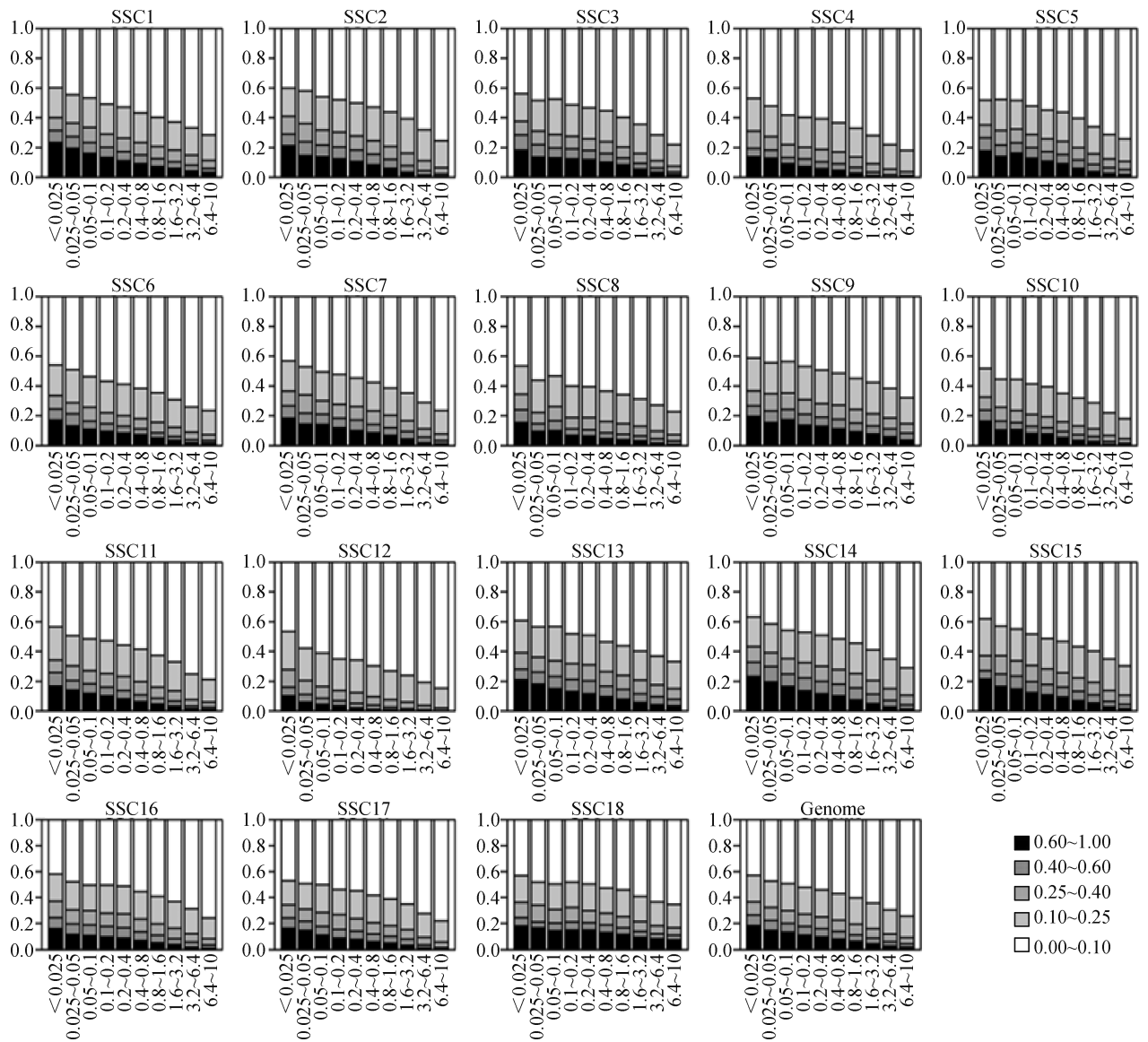
## 2.3 恩施黑猪选择信号分析

本研究使用 CLR 与 iHS 两种互补的检测方法进行选择信号分析。在 32 642 个 CLR 统计量中, 326 个统计量达到显著水平, 合并同一染色体上物理间距在 200 kb 范围内的选择信号, 共获得 126 个显著选择信号区域 (图 3)。全部区域的长度约为 51.6 Mb, 约占基因组总长的 2.1%。基于 iHS 方法计算每个核心 SNP 的检验统计量, 并将其绝对值平均到 75 kb 的窗口中, 共包含有 32 449 个 iHS 统计量, 其中, 324 个统计量达到显著水平, 同上合并相邻检验统计量, 共发现 248 个显著选择信号区域。全部区域的长度为 78.78 Mb, 约占基因组总长的 3.2%。研究发现, CLR 与 iHS 方法检测到的选择信号区域重叠较少 (图 3)。其主要原因可能是两种方法对不同类型的选择信号具有不同的检测灵敏度; CLR 方法对固定或者即将固定的选择信号具有较高的检测效力, 而 iHS 方法则对正在进行的选择信号具有较高的检测效力 (图 3)。

## 2.4 恩施黑猪基因富集分析

基于恩施黑猪正向选择信号区域, 利用 Ensembl 数据库, 通过 BioMart 抓取选择信号区域富集的基因集与人类基因组进行直系同源比对, 共获得 1 147 个人类直系同源基因。富集分析发现, 尽管绝大多数的生物学条目不能直观地反映具体的经济性状, 但是部分条目的生物学功能似乎与恩施黑猪的免疫能力、气味识别、感官能力相关。例如, 有 16 个基因被富集到“I 型干扰素介导信号通路调控”条目, 有 45 个基因被富集到“嗅觉传导”的生物学通路 (表 2)。

选择信号区域基因功能注释发现, 一些与重要经济性状相关的基因与恩施黑猪的选择信号潜在理论上, 家猪在适应性进化的过程中, 其繁殖性状和



横坐标表示各染色体(SSC1-SSC18)和全基因组 10 Mb 范围内不同的 SNP 间距类别,分别为<0.025、0.025~0.05、0.05~0.1、0.1~0.2、0.2~0.4、0.4~0.8、0.8~1.6、1.6~3.2、3.2~6.4、6.4~10 Mb;纵坐标表示不同  $r^2$  类别的占比

The x-axis indicates the different categories of SNP interval distance up to 10 Mb in each chromosome and the whole genome, including <0.025, 0.025-0.05, 0.05-0.1, 0.1-0.2, 0.2-0.4, 0.4-0.8, 0.8-1.6, 1.6-3.2, 3.2-6.4, 6.4-10 Mb; the y-axis indicates the proportion of different  $r^2$  categories

图 1 恩施黑猪基因组 10 Mb 范围内 SNPs 之间连锁不平衡程度随 SNP 间距增加的衰减趋势

Fig. 1 The tendency of linkage disequilibrium decay as the increasing distance between SNPs pairs ranged 10 Mb for the whole genome of Enshi Black pig

生长性状均受到不同程度的改良。研究发现,在 2 号染色体 58 和 144.8 Mb 处的选择信号区域分别与 *LPAR2* 和 *NR3C1* 基因重叠,前者与猪胴体长度相关,后者则影响猪生长调控相关生物信息学通路<sup>[16-18]</sup>。在 2 号染色体 58 Mb 处的选择信号区域与一系列繁殖相关的基因重叠,包括与精子形成相关的 *NDUFA13* 和 *TSSK6* 基因<sup>[19-20]</sup>。此外,9 号染

色体 86 Mb 处的选择信号区域与 *AHR* 基因相重叠,该基因能够显著影响大白猪的总产仔数,是亚洲地方猪渗入基因<sup>[21]</sup>(表 3)。据此可以推断,恩施黑猪在适应性驯化过程中,其繁殖性状受到一定程度的选择作用;尽管目前群体的繁殖性能较商品猪差,但上述选择信号区域的发现对恩施黑猪繁殖性能的遗传改良具有一定的指导作用。

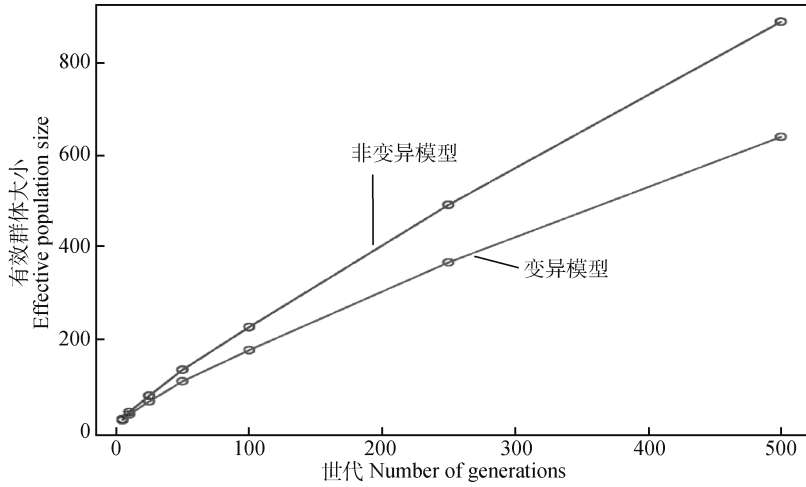
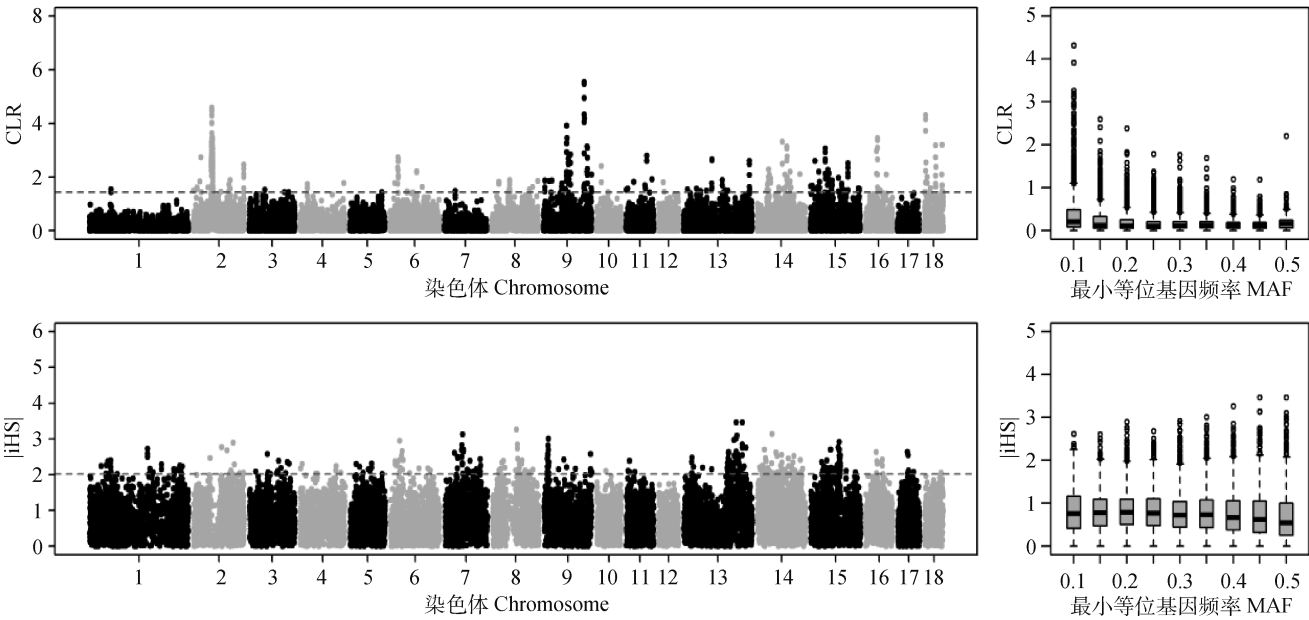


图 2 根据全基因组连锁不平衡数据估计恩施黑猪不同历史世代有效群体大小

Fig. 2 The estimated effective population size over the past generations based on linkage disequilibrium data of genome



左侧为选择信号曼哈顿图,右侧为不同最小等位基因频率(MAF)下 CLR 与 iHS 统计量的分布图

The left Manhattan plots shows the significant selection signatures. The right plots are the distribution of CLR statistics and iHS statistics at different minimum allele frequencies

图 3 恩施黑猪选择信号检测结果示意图

Fig. 3 The visualization of selection signatures identification in Enshi Black pig

表 2 基因富集分析结果

Table 2 The result of terms revealed by enrichment analyses

类型 Type	条目 Term	基因数目 Gene number	P 值 P value
生物学过程 Biological process	GO:0060338:I 型干扰素介导信号通路调控 Regulation of type I interferon-mediated signaling pathway	16	<0.001
	GO:0002323:参与免疫应答的自然杀伤细胞激活 Natural killer cell activation involved in immune response	14	<0.001

(续表 2 Continued)

类型 Type	条目 Term	基因数目 Gene number	P 值 P value
	GO:0002286:T 细胞激活参与免疫应答 T cell activation involved in immune response	14	<0.001
	GO:0051607:防御病毒反应 Defense response to virus	24	<0.001
	GO:0007608:气味感官 Sensory perception of smell	21	<0.001
	GO:0045926:生长的负向调控 Negative regulation of growth	5	0.015
	GO:0002250:适应性免疫应答 Adaptive immune response	15	0.022
生物学通路 Biological pathway	hsa04140:调节自噬 Regulation of autophagy	14	<0.001
	hsa04650:自然杀伤细胞介导细胞毒性 Natural killer cell mediated cytotoxicity	21	<0.001
	hsa04740:嗅觉传导 Olfactory transduction	45	<0.001
	hsa04660:T 细胞受体信号通路 Cytokine-cytokine receptor interaction	14	0.007
	hsa00071:脂肪酸降解 Fatty acid degradation	8	0.013

表 3 部分恩施黑猪的选择信号区域富集重要候选基因

Table 3 Some candidate genes overlapping with the selection signature regions in Enshi Black pig

染色体 Chromosome	选择信号区域/Mb Selection signature region	P 值(方法) P value (method)	基因 Gene	基因功能 Function
1	183.275~184.000	<0.001(iHS)	<i>BMP4</i>	肾发育 <sup>[21]</sup> Kidney development
2	57.725~59.125	<0.001(CLR)	<i>LPAR2</i> ( <i>EDG4</i> )	胴体长度 <sup>[16]</sup> Carcass length 滴水损失 <sup>[16]</sup> Drip loss percentage
2	57.725~59.125	<0.001(CLR)	<i>NDUFA13</i>	精子形成 <sup>[19]</sup> Sperm formation
2	57.725~59.125	<0.001(CLR)	<i>TSSK6</i>	精子发生 <sup>[20]</sup> Spermatogenesis-related
2	57.725~59.125	<0.001(CLR)	<i>NR2C2AP</i>	抑制 TR4 介导的激活 <sup>[22]</sup> Suppress the TR4-mediated transactivation
2	57.725~59.125	<0.001(CLR)	<i>MEF2B</i>	骨骼肌分化 <sup>[23]</sup> Skeletal muscle differentiation

(转下页 Carried forward)

(续表 3 Continued)

染色体 Chromosome	选择信号区域/Mb Selection signature region	<i>P</i> 值(方法) <i>P</i> value (method)	基因 Gene	基因功能 Function
2	57.725~59.125	<0.001(CLR)	<i>TMEM161A</i>	氧化应激防护 <sup>[24]</sup> Protection against oxidative stress
2	144.95~145.225	0.009(iHS)	<i>NR3C1</i> ( <i>GR</i> )	生长 <sup>[17]</sup> Growth
9	86.225~86.800	0.003(CLR)	<i>AHR</i>	繁殖 <sup>[21]</sup> Reproductivity
13	176.075~176.350	0.004(iHS)	<i>ROBO1</i>	前脑神经元分化 <sup>[21]</sup> Forebrain neuron differentiation
14	79.700~80.125	0.008(CLR)	<i>KCNMA1</i>	唾液分泌 <sup>[25]</sup> Saliva secretion
17	50.075~50.350	0.008(iHS)	<i>CD40</i>	脂肪细胞分化 <sup>[26]</sup> Adipocyte differentiation

### 3 讨论

#### 3.1 利用基因组信息估计恩施黑猪群体遗传学参数

长期以来,中国乃至世界地方畜禽品种由于商业化品种的冲击,饲养规模不断缩小,能够提供基本遗传信息的系谱记录缺失,部分品种处于濒危状态,甚至濒临灭绝。恩施黑猪种群调查表明,近年来恩施黑猪群体规模不断缩小,现存血统数目有限,系谱记录缺失,配种关系混乱。如何保证恩施黑猪的遗传多态,做好保种和遗传改良工作显得尤为关键。随着测序技术的发展,畜禽基因组数据呈现爆发式增长,这为在基因组层面进行群体遗传学参数的估计提供了可能。基于亲缘系数和近交系数的计算,发现恩施黑猪群体长期处于高度近交状态,存在明显的近交累积。对恩施黑猪有效群体大小的估计结果同样不容乐观,研究表明,其有效群体规模从 500 代以前的约 1 000 头,逐代下降至 5 代前的 25 头。估计恩施黑猪 20 年前(50 世代)有效群体大小约为 135 头,和目前其它中国地方猪报道的有效群体大小基本一致,明显低于商品猪有效群体大小<sup>[15]</sup>。基于以上对于恩施黑猪历史有效群体大小和趋势的估计可以推断,恩施黑猪群体在近代有效群体的规模将进一步缩小,这对于恩施黑猪群体内的基因多样性和异质性的保护提出了挑战。

在这种情况下,利用基因组数据进行畜群个体间亲缘关系的评估,有助于填补系谱记录缺失造成

的选配混乱;通过基因组近交系数的估计,可以为恩施黑猪种猪的选留提供参考。利用基因组同源相同信息估计个体间亲缘关系,在全基因组关联分析等研究中常被用于无关个体的挑选,部分研究使用的标准为个体间  $\pi_{\text{hat}} < 0.125$ <sup>[27-28]</sup>。传统上,利用系谱信息估计亲缘系数为 0.125 时,反映的是至少 3 代以上个体间的亲缘关系。与此同时,有研究认为亲缘关系在 3 代以外的个体可以近似为无关个体<sup>[4]</sup>。本研究表明,约 50% 的个体对间亲缘系数大于等于 0.125,即个体间存在较近的亲缘关系(表 1)。同样,利用基因组信息通过计算基因组纯合片段占比发现,约 16.67% 的样本近交系数大于等于 0.125。上述分析表明,恩施黑猪个体间亲缘关系较近,具有明显的近交累积现象。理论上,相对于使用传统系谱信息估计的亲缘系数与近交系数的期望值,充分利用基因组信息,通过恰当的统计方法确实能够更加准确地估计亲缘系数与个体近交系数。但是,如何构建恰当的亲缘系数与近交系数的估计方法仍然需要开展更深入的研究。

#### 3.2 恩施黑猪人工选择作用

在恩施黑猪品种形成的过程中,由于选择的作用会在基因组上留下选择印迹即所谓的选择信号。通过检测识别选择信号,便有可能找到影响恩施黑猪优良性状如耐寒和高储脂能力的潜在遗传变异,为相关区域基因的功能研究提供重要参考。

本研究基于高密度 SNP 芯片数据,运用了两种方法对恩施黑猪进行全基因组选择信号检测,研究



发现,共约有 250 个基因组区段受到了强烈的正向选择作用。进一步通过基因富集分析和候选基因分析发现,这些基因与恩施黑猪的生长发育、生产繁殖以及储脂能力有很大的关系,据此可以推测,恩施黑猪驯化过程中人为的选择往往针对某些对人类有利的性状,如胴体长度、繁殖力、猪肉品质等。

从内在机制来看,选择是针对功能相对应的基因,在恩施黑猪中检测到以下几个具有代表性的受到选择作用的基因:*LPAR2(EDG4)*、*NDUFA13*、*MEF2B*、*AHR*。它们分别对应的功能为胴体长度(滴水损失)、精子形成、骨骼肌分化和繁殖力<sup>[16, 19, 22-23, 29]</sup>。它们之所以受到强烈选择,可能是因为恩施黑猪在人工驯化和育种过程中,外部形态和行为特征发生改变:繁殖力增强、四肢发育良好、肉质性状优良、储脂能力增强,也就是说恩施黑猪在适应性进化过程中似乎正是向着人类选择方向进行的。上述潜在选择信号区域为揭示恩施黑猪适应性进化的基因,了解恩施黑猪的进化历程并获得其优良的经济性状具有非常深远的意义。

## 4 结 论

本研究结果表明,现存恩施黑猪群体血统较窄,近交累积严重,有效血统数目有限,且有效群体规模有不断缩小的趋势,目前保种形势严峻,恩施黑猪特色种质特性随时存在丢失风险。使用 CLR 与 iHS 两种互补的检测方法分别发现大约 2.1% 和 3.2% 的恩施黑猪基因组受到选择,影响生长与繁殖等重要经济性状的 *LPAR2*、*NR3C1*、*NDUFA13*、*TSSK6* 与 *AHR* 等基因与上述选择信号区域重叠,上述基因的发现为未来恩施黑猪的遗传改良提供了一定的参考依据。

## 参考文献(References):

- [1] 郑雪君,杨婷婷. 中国地方猪品种的保护与利用分析[J]. 中国畜牧杂志, 2015, 51(16): 24-27.  
ZHENG X J, YANG T T. Analysis on the conservation and utilization of Chinese local pig breeds[J]. *Chinese Journal of Animal Science*, 2015, 51(16): 24-27. (in Chinese)
- [2] MIAR Y, PLASTOW G, BRUCE H, et al. Genetic and phenotypic correlations between performance traits with meat quality and carcass characteristics in commercial crossbred pigs [J]. *PLoS One*, 2014, 9(10): e110105.
- [3] LEE J H, SONG K D, LEE H K, et al. Genetic parameters of reproductive and meat quality traits in Korean Berkshire pigs[J]. *Asian-Australas J Anim Sci*, 2015, 28(10): 1388-1393.
- [4] AI H S, HUANG L S, REN J. Genetic diversity, linkage disequilibrium and selection signatures in Chinese and Western pigs revealed by genome-wide SNP markers[J]. *PLoS One*, 2013, 8(2): e56001.
- [5] BOSSE M, MEGENS H J, MADSEN O, et al. Using genome-wide measures of coancestry to maintain diversity and fitness in endangered and domestic pig populations[J]. *Genome Res*, 2015, 25(7): 970-981.
- [6] QANBARI S, PIMENTEL E C G, TETENS J, et al. The pattern of linkage disequilibrium in German Holstein cattle[J]. *Anim Genet*, 2010, 41(4): 346-356.
- [7] 薛周敏源, 宋显威, 吴林慧, 等. 畜禽选择信号检测方法及其统计学问题[J]. 畜牧兽医学报, 2018, 49(6): 1099-1107.  
XUE Z Y Y, SONG X W, WU L H, et al. The identification methods of selection signatures in livestock and its statistical problems[J]. *Acta Veterinaria et Zootechnica Sinica*, 2018, 49(6): 1099-1107. (in Chinese)
- [8] QANBARI S, SIMIANER H. Mapping signatures of positive selection in the genome of livestock[J]. *Livest Sci*, 2014, 166: 133-143.
- [9] PURCELL S, NEALE B, TODD-BROWN K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses [J]. *Am J Hum Genet*, 2007, 81(3): 559-575.
- [10] SVED J A. Linkage disequilibrium and homozygosity of chromosome segments in finite populations [J]. *Theor Popul Biol*, 1971, 2(2): 125-141.
- [11] NIELSEN R, WILLIAMSON S, KIM Y, et al. Genomic scans for selective sweeps using SNP data [J]. *Genome Res*, 2005, 15(11): 1566-1575.
- [12] MA Y, DING X, QANBARI S, et al. Properties of different selection signature statistics and a new strategy for combining them [J]. *Heredity*, 2015, 115(5): 426-436.
- [13] VOIGHT B F, KUDARAVALLI S, WEN X Q, et al. A map of recent positive selection in the human genome [J]. *PLoS Biol*, 2006, 4(3): e72.
- [14] QANBARI S, PAUSCH H, JANSEN S, et al. Classic selective sweeps revealed by massive sequencing in cattle [J]. *PLoS Genet*, 2014, 10(2): e1004148.
- [15] TRASPOV A, DENG W J, KOSTYUNINA O, et al.

- Population structure and genome characterization of local pig breeds in Russia, Belorussia, Kazakhstan and Ukraine[J]. *Genet Sel Evol*, 2016, 48:16.
- [16] SHAN T L, LI K, TANG Z L, et al. Full-length coding sequences, polymorphism and chromosomal localizations of the porcine *EDG4* and *EDG7* genes[J]. *Mol Biol Rep*, 2009, 36(4):751-756.
- [17] EDEA Z, HONG J K, JUNG J H, et al. Detecting selection signatures between Duroc and Duroc synthetic pig populations using high - density SNP chip[J]. *Anim Genet*, 2017, 48(4):473-477.
- [18] JIA Y M, GAO G C, SONG H G, et al. Low-protein diet fed to crossbred sows during pregnancy and lactation enhances myostatin gene expression through epigenetic regulation in skeletal muscle of weaning piglets[J]. *Eur J Nutr*, 2016, 55(3):1307-1314.
- [19] YANG Y, CHENG L Y, WANG Y, et al. Expression of *NDUFA13* in asthenozoospermia and possible pathogenesis [J]. *Reprod BioMed Online*, 2017, 34(1):66-74.
- [20] WANG P, HUO H L, WANG S Y, et al. Cloning, sequence characterization, and expression patterns of members of the porcine *TSSK* family[J]. *Genet Mol Res*, 2015, 14(4):14908-14919.
- [21] AI H S, FANG X D, YANG B, et al. Adaptation and possible ancient interspecies introgression in pigs identified by whole-genome sequencing[J]. *Nat Genet*, 2015, 47(3):217-225.
- [22] YANG Y, WANG X, DONG T F, et al. Identification of a novel testicular orphan receptor-4 (TR4)-associated protein as repressor for the selective suppression of TR4-mediated transactivation [J]. *J Biol Chem*, 2003, 278(9):7709-7717.
- [23] ESTRELLA N L, DESJARDINS C A, NOCCO S E, et al. MEF2 transcription factors regulate distinct gene programs in mammalian skeletal muscle differentiation[J]. *J Biol Chem*, 2015, 290(2):1256-1268.
- [24] MONTESANO GESUALDI N, CHIRICO G, CATTANESE M T, et al. AROS-29 is involved in adaptive response to oxidative stress [J]. *Free Radic Res*, 2006, 40(5):467-476.
- [25] LI M Z, TIAN S L, JIN L, et al. Genomic analyses identify distinct patterns of selection in domesticated pigs and Tibetan wild boars [J]. *Nat Genet*, 2013, 45(12):1431-1438.
- [26] WON S, JUNG J, PARK E, et al. Identification of genes related to intramuscular fat content of pigs using genome-wide association study [J]. *Asian-Australas J Anim Sci*, 2018, 31(2):157-162.
- [27] WATERWORTH D M, LI L, SCOTT R, et al. A low-frequency variant in *MAPK14* provides mechanistic evidence of a link with myeloperoxidase: a prognostic cardiovascular risk marker [J]. *J Am Heart Assoc*, 2014, 3(4):e001074.
- [28] SIEDLINSKI M, CHO M H, BAKKE P, et al. Genome-wide association study of smoking behaviours in patients with COPD [J]. *Thorax*, 2011, 66(10):894-902.
- [29] BOSSE M, MEGENS H J, FRANTZ L A F, et al. Genomic analysis reveals selection for Asian genes in European pigs following human-mediated introgression [J]. *Nat Commun*, 2014, 5:4392.

(编辑 郭云雁)