

赵旭庭, 徐盼, 马政, 等. 枫泾猪、梅山猪和沙乌头猪全基因组 ROH 检测及候选基因鉴定[J]. 江苏农业学报, 2021, 37(5): 1234-1243.

doi: 10.3969/j.issn.1000-4440.2021.05.019

## 枫泾猪、梅山猪和沙乌头猪全基因组 ROH 检测及候选基因鉴定

赵旭庭<sup>1</sup>, 徐盼<sup>1</sup>, 马政<sup>1</sup>, 仲德<sup>1</sup>, 刘林雨<sup>1</sup>, 倪黎纲<sup>1</sup>, 吴井生<sup>2</sup>

(1. 江苏农牧科技职业学院, 江苏 泰州 225300; 2. 江苏农林职业技术学院, 江苏 句容 212400)

**摘要:** 通过检测枫泾猪、梅山猪和沙乌头猪的长纯合片段(Runs of homozygosity, ROH), 鉴定其与 3 个猪种经济性状相关的候选基因。利用全基因组单核苷酸多态性(Single Nucleotide Polymorphisms, SNP)芯片数据对这 3 个猪群体进行全基因组 ROH 检测, 统计 ROH 数量、长度及分布, 计算基于 ROH 的近交系数( $F_{ROH}$ ), 并在高频 ROH 区域鉴定候选功能基因。结果显示, 从枫泾猪、梅山猪和沙乌头猪群体中分别检测到 1 588 个、3 314 个和 4 419 个 ROH。枫泾猪的 ROH 平均长度为 12.84 Mb, 平均  $F_{ROH}$  为 0.084; 梅山猪的 ROH 平均长度为 10.55 Mb, 平均  $F_{ROH}$  为 0.164; 沙乌头猪的 ROH 平均长度为 10.52 Mb, 平均  $F_{ROH}$  为 0.141。在高频 ROH 区域共鉴定到 9 个枫泾猪候选功能基因和 8 个梅山猪候选功能基因。研究结果为枫泾猪、梅山猪和沙乌头猪的资源保护和分子标记辅助选择奠定了基础。

**关键词:** 猪; 长纯合片段; 近交系数; 候选基因

**中图分类号:** S813.1 **文献标识码:** A **文章编号:** 1000-4440(2021)05-1234-10

## Genome-wide scan for run of homozygosity and identification of corresponding candidate genes in Fengjing pigs, Meishan pigs and Shawutou pigs

ZHAO Xu-ting<sup>1</sup>, XU Pan<sup>1</sup>, MA Zheng<sup>1</sup>, ZHONG De<sup>1</sup>, LIU Lin-yu<sup>1</sup>, NI Li-gang<sup>1</sup>, WU Jing-sheng<sup>2</sup>

(1. Jiangsu Agri-animal Husbandry Vocational College, Taizhou 225300, China; 2. Jiangsu Vocational College of Agriculture and Forestry, Jurong 212400, China)

**Abstract:** By detecting runs of homozygosity (ROH) of Fengjing pig, Meishan pig and Shawutou pig, the candidate genes related to the economic traits of the three pig breeds were identified. The genome-wide single nucleotide polymorphisms (SNP) chip data were used to perform genome-wide ROH scan for the three pig breeds. The number, length and

distribution of ROH were counted, and the inbreeding coefficient ( $F_{ROH}$ ) was calculated based on ROH. In addition, the functional candidate genes were identified in the high frequency ROH region. The results showed that 1 588, 3 314 and 4 419 ROH were detected in the three pig populations, respectively. The average length of ROH was 12.84 Mb and the average  $F_{ROH}$  was 0.084 in Fengjing pigs. The average length of ROH was 10.55 Mb and the

收稿日期: 2021-02-07

基金项目: 国家自然科学基金(31702089); 江苏现代农业(生猪)产业技术体系[JATS(2020)346]; 江苏高校“青蓝工程”项目

作者简介: 赵旭庭(1966-), 男, 江苏靖江人, 硕士, 教授, 主要从事地方猪品种资源保护与遗传育种研究。(E-mail) jsfduc@163.com。徐盼为共同第一作者。

通讯作者: 吴井生, (E-mail) 17190607@qq.com

average  $F_{\text{ROH}}$  was 0.164 in Meishan pigs. The average length of ROH was 10.52 Mb and the average  $F_{\text{ROH}}$  was 0.141 in Shawutou pigs. Nine functional candidate genes of Fengjing pig and eight functional candidate genes of Meishan pig were identified in high frequency ROH region. The results lay a foundation for resource conservation and molecular marker-assisted selection of Fengjing pig, Meishan pig and Shawutou pig.

**Key words:** pig; runs of homozygosity (ROH); inbreeding coefficient; candidate gene

长纯合片段(Runs of homozygosity, ROH)是基因组上一定数量、一定密度的单核苷酸多态性(Single nucleotide polymorphisms, SNPs)表现为纯合的一段区域<sup>[1]</sup>。在人工选择的过程中,畜禽群体的基因组被重塑,所选区域中的 ROH 频率会提高,甚至产生“ROH 岛”。ROH 信息在畜禽亲缘关系鉴定、遗传多样性分析、近交衰退评估、选择信号鉴定和功能基因筛选等方面发挥着重要作用。Keller 等<sup>[2]</sup>在比较基于 ROH 的近交系数( $F_{\text{ROH}}$ )、基于基因组杂合度的近交系数( $F_{\text{HOM}}$ )、基于基因组关系矩阵的近交系数( $F_{\text{GRM}}$ )和基于系谱信息的近交系数( $F_{\text{PED}}$ )后发现, $F_{\text{ROH}}$ 是最有效的估计近交的方法。Mastrangelo 等<sup>[3]</sup>在 Pontremolese、Varzese-Ottone 和 Mucca Pisana 群体中检测到高水平的 ROH,针对这些群体,在实施配种计划中应尽量增加种公畜血统,减少其遗传多样性的损失,维持或者增加其群体有效含量。Saura 等<sup>[4]</sup>分析了近交对伊比利亚猪繁殖性状的影响,结果表明,近交系数每增加 0.1,其仔猪初生后的存活率、仔猪出生后的总数量有下降趋势。Zhong 等<sup>[5]</sup>利用简化基因组测序数据在金华猪群体的 ROH 岛中鉴别到与繁殖、肉质、食欲及抗病性状相关的基因。

枫泾猪、梅山猪和沙乌头猪是江苏省优良的地方猪遗传资源,以产仔数高、肉质优良、耐粗饲和环境适应力强等特点而闻名。本研究利用中芯一号育种芯片数据对这 3 个猪群体进行全基因组 ROH 检测,统计 ROH 数量、长度及分布,计算基于 ROH 的近交系数,并在高频 ROH 区域鉴定候选功能基因,以期对枫泾猪、梅山猪和沙乌头猪的资源保护和分子标记辅助选择奠定基础。

## 1 材料与方法

### 1.1 试验动物

试验动物为来自镇江牧苑动物科技开发有限公司省级枫泾猪保种场的 107 头枫泾猪(FJ)、江苏农林职业技术学院国家级梅山猪保种场的 94 头梅山

猪(MS)和江苏兴旺农牧科技发展有限公司国家级沙乌头猪保种场的 146 头沙乌头猪(SWT)。采集猪耳组织样后保存于 $-20^{\circ}\text{C}$ 冰箱内待用。

### 1.2 SNP 芯片分型及质量控制

委托北京康普森生物技术有限公司使用中芯一号育种芯片对 107 头枫泾猪、94 头梅山猪和 146 头沙乌头猪进行基因分型。用 plink 软件分别对 3 个猪群体的芯片数据进行质量控制,质量控制标准:1)个体检出率 $\geq 90\%$ ;2)SNP 检出率 $\geq 90\%$ ;3)次等位基因频率(MAF) $\geq 0.01$ ;4)哈迪温伯格平衡 $P \geq 1 \times 10^{-5}$ ;5)剔除性染色体。

### 1.3 全基因组 ROH 检测

ROH 检测使用 Plink 软件,使用滑动窗口的方法对常染色体进行检测,具体的检测参数:1)ROH 长度大于 1 Mb;2)至少连续 50 个 SNPs;3)连续 SNP 间的距离小于 1 Mb;4)最多允许 ROH 中有 5 个缺失和 1 个杂合;5)窗口阈值为 0.01;6)密度为 0.01 SNP/kb。

### 1.4 基因组近交系数的计算

利用 ROH 计算近交系数( $F_{\text{ROH}}$ ),公式如下:

$$F_{\text{ROH}} = \sum L_{\text{ROH}} / L_{\text{auto}}$$

其中: $\sum L_{\text{ROH}}$ 为常染色体上 ROH 片段的长度之和, $L_{\text{auto}}$ 为常染色体的物理总长度。

根据上式计算  $F_{\text{ROH } 1.0 \sim 5.0 \text{ Mb}}$ 、 $F_{\text{ROH } 5.1 \sim 10.0 \text{ Mb}}$ 、 $F_{\text{ROH } > 10.0 \text{ Mb}}$ 和  $F_{\text{ROH total}}$ 4 种近交系数之间的 Pearson 相关系数,其中  $F_{\text{ROH total}}$ 为基于全部长度的 ROH 计算的近交系数。

### 1.5 高频 ROH 区域的检测及候选基因的鉴定

统计每个 SNP 在各自猪群体内参与组成 ROH 的次数占样本数的比例,并将高于 45% 的 SNP 区域作为高频 ROH 区域,利用 RCircos 包绘制 3 个猪群体 ROH 频率在染色体上分布的 Circos 图。基于高频 ROH 区段的物理位置,在猪基因组注释文件([ftp://ftp.ensembl.org/pub/release-100/gtf/sus\\_scrofa/Sus\\_scrofa.SsCrofa11.1.100.chr.gtf.gz](ftp://ftp.ensembl.org/pub/release-100/gtf/sus_scrofa/Sus_scrofa.SsCrofa11.1.100.chr.gtf.gz))中检索基因。

利用 DAVID 数据库 (<https://david.ncifcrf.gov/>) 对高频 ROH 区域注释到的基因进行 GO (Gene ontology, 基因本体) 功能富集和 KEGG (Kyoto Encyclopedia of Genes and Genomes, 京都基因与基因组百科全书) Pathway 富集分析, 当  $P \leq 0.05$  时, 则表示显著富集。将显著富集到生物学过程和信号通路的基因与猪数量性状基因座 (Quantitative Trait Locus, QTL) 数据库 (<https://www.animalgenome.org/cgi-bin/QTLdb/SS/index>) 进行比对, 获得候选功能基因。

## 2 结果与分析

### 2.1 基因型的质量控制

中芯一号育种芯片含有 51 315 个 SNP 位点, 经过质量控制后, 对剩余 107 头枫泾猪的 19 895 个 SNPs, 94 头梅山猪的 24 526 个 SNPs 和 146 头沙乌头猪的 32 275 个 SNPs 进行后续分析。

### 2.2 ROH 数量、长度及分布统计

由表 1、图 1、图 2 可以看出, 从枫泾猪、梅山猪

和沙乌头猪群体中分别检测到 1 588 个、3 314 个和 4 419 个 ROH, 其长度主要集中在 40.00 Mb 以内。枫泾猪 ROH 的平均长度为 12.84 Mb; ROH 的最长长度为 118.00 Mb, 含有 1 328 个 SNPs, 位于第 15 号染色体上; ROH 的最短长度为 2.70 Mb, 含有 50 个 SNPs, 位于第 13 号染色体上; 枫泾猪个体的平均 ROH 数量为 14.84 个, 以 15 号染色体上的 ROH 最多, 18 号染色体上的 ROH 最少。梅山猪 ROH 的平均长度为 10.55 Mb; ROH 的最长长度为 98.33 Mb, 含有 1 313 个 SNPs, 位于第 14 号染色体上; ROH 的最短长度为 1.95 Mb, 含有 59 个 SNPs, 位于第 6 号染色体上; 梅山猪个体 ROH 的平均数量为 35.26 个, 14 号染色体上的 ROH 最多, 12 号染色体上的 ROH 最少。沙乌头猪 ROH 的平均长度为 10.52 Mb; ROH 的最长长度为 130.29 Mb, 含有 1 727 个 SNPs, 位于第 1 号染色体上; ROH 的最短长度为 1.70 Mb, 含有 50 个 SNPs, 位于第 6 号染色体上; 沙乌头猪个体 ROH 的平均数量为 30.27 个, 1 号染色体上的 ROH 最多, 10 号染色体上的 ROH 最少。

表 1 枫泾猪、梅山猪和沙乌头猪的 ROH 长度和数量

Table 1 The length and number of runs of homozygosity (ROH) in Fengjing, Meishan and Shawutou pigs

种群	样本数 (个)	群体 ROH 数量	ROH 长度 (Mb)		猪个体 ROH 数量 (个)	
			均值±标准差	范围	均值±标准差	范围
枫泾猪	107	1 588	12.84±12.41	2.70~118.00	14.84±4.60	1~26
梅山猪	94	3 314	10.55±0.97	1.95~98.33	35.26±7.22	10~50
沙乌头猪	146	4 419	10.52±12.45	1.70~130.29	30.27±10.49	9~51

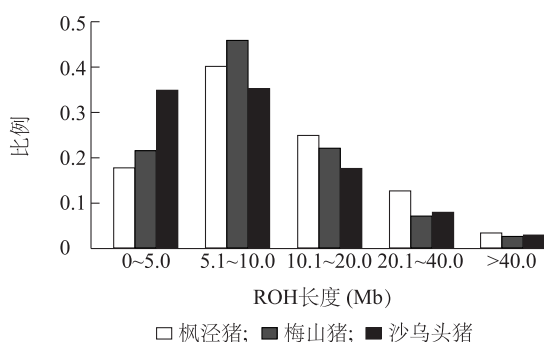


图 1 不同长度 ROH 占比

Fig.1 Proportion of ROH with different lengths

### 2.3 近交系数

由表 2 可以看出, 枫泾猪基因组  $F_{ROH\ total}$  的最大

值为 0.171, 最小值为 0.003, 平均值为 0.084; 梅山猪基因组  $F_{ROH\ total}$  的最大值为 0.305, 最小值为 0.057, 平均值为 0.164; 沙乌头猪基因组  $F_{ROH\ total}$  的最大值为 0.501, 最小值为 0.022, 平均值为 0.141; 3 个群体中梅山猪的平均  $F_{ROH\ total}$  最高。对比表 2 中 3 种猪群体的 4 种近交系数发现, 枫泾猪群体中  $F_{ROH\ 5.1\sim10.0\ Mb}$  与  $F_{ROH\ >10.0\ Mb}$ 、 $F_{ROH\ total}$  分别在 0.05、0.001 水平显著相关,  $F_{ROH\ >10.0\ Mb}$  与  $F_{ROH\ total}$  在 0.001 水平显著相关; 梅山猪群体中  $F_{ROH\ 5.1\sim10.0\ Mb}$  与  $F_{ROH\ 1.0\sim5.0\ Mb}$ 、 $F_{ROH\ total}$  及  $F_{ROH\ >10.0\ Mb}$  与  $F_{ROH\ total}$  均在 0.001 水平显著相关; 沙乌头猪群体中  $F_{ROH\ 1.0\sim5.0\ Mb}$  与  $F_{ROH\ total}$  在 0.01 水平显著相关,  $F_{ROH\ 1.0\sim5.0\ Mb}$  与  $F_{ROH\ 5.1\sim10.0\ Mb}$ 、 $F_{ROH\ 5.1\sim10.0\ Mb}$  与  $F_{ROH\ >10.0\ Mb}$ 、 $F_{ROH\ 5.1\sim10.0\ Mb}$  与  $F_{ROH\ total}$ 、 $F_{ROH\ >10.0\ Mb}$  与  $F_{ROH\ total}$  均在 0.001 水平显著相关 (表 3)。

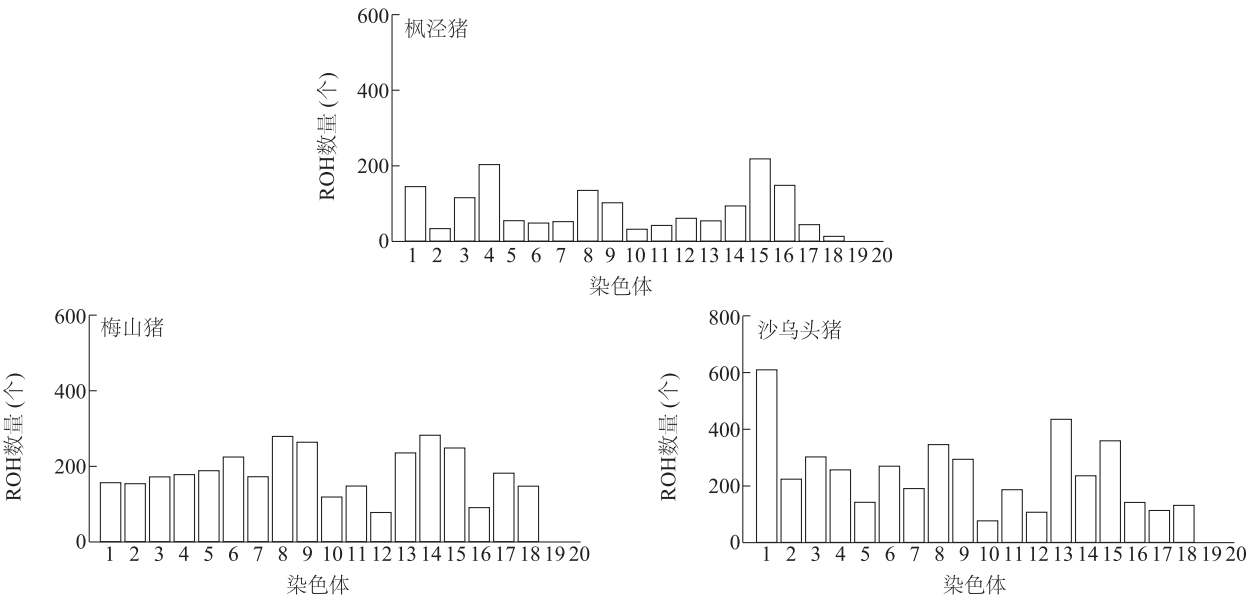


图 2 不同染色体的 ROH 数量  
Fig.2 The number of ROH on different chromosomes

表 2 基于 ROH 的枫泾猪、梅山猪和沙乌头猪的近交系数 ( $F_{\text{ROH}}$ )  
Table 2 Inbreeding coefficient ( $F_{\text{ROH}}$ ) of Fengjing, Meishan and Shawutou pigs based on ROH

种群	ROH 长度	样本数(个)	近交系数 ( $F_{\text{ROH}}$ )			
			均值	标准差	最小值	最大值
枫泾猪	total	107	0.084	0.034	0.003	0.171
	1.0~5.0 Mb	98	0.005	0.003	0.001	0.011
	5.1~10.0 Mb	107	0.018	0.008	0.003	0.045
	>10.0 Mb	104	0.062	0.030	0.006	0.143
梅山猪	total	94	0.164	0.047	0.057	0.305
	1.0~5.0 Mb	94	0.015	0.005	0.002	0.029
	5.1~10.0 Mb	94	0.050	0.013	0.005	0.073
	>10.0 Mb	94	0.099	0.044	0.035	0.237
沙乌头猪	total	146	0.141	0.085	0.022	0.501
	1.0~5.0 Mb	146	0.018	0.006	0.005	0.040
	5.1~10.0 Mb	146	0.033	0.014	0.003	0.066
	>10.0 Mb	145	0.090	0.077	0.005	0.457

表 3 基于 ROH 的近交系数 ( $F_{\text{ROH}}$ ) 的相关系数  
Table 3 Correlation coefficients among the inbreeding coefficient ( $F_{\text{ROH}}$ ) based on ROH

近交系数	枫泾猪				梅山猪				沙乌头猪			
	A	B	C	D	A	B	C	D	A	B	C	D
$F_{\text{ROH } 1.0\sim 5.0 \text{ Mb}}$	1.000				1.000				1.000			
$F_{\text{ROH } 5.1\sim 10.0 \text{ Mb}}$	-0.038	1.000			0.480***	1.000			0.340***	1.000		
$F_{\text{ROH } >10.0 \text{ Mb}}$	-0.160	0.250*	1.000		-0.094	0.060	1.000		0.120	0.410***	1.000	
$F_{\text{ROH total}}$	-0.079	0.520***	0.970***	1.000	0.160	0.390***	0.940***	1.000	0.240**	0.560***	0.980***	1.000

A: $F_{\text{ROH } 1.0\sim 5.0 \text{ Mb}}$ ;B: $F_{\text{ROH } 5.1\sim 10.0 \text{ Mb}}$ ;C: $F_{\text{ROH } >10.0 \text{ Mb}}$ ;D: $F_{\text{ROH total}}$  \*、\*\*、\*\*\* 分别表示在 0.05、0.01、0.001 水平显著相关。

## 2.4 高频 ROH 区域及候选基因的鉴定

本研究选取 45% 作为高频 ROH 区域的阈值, 在枫泾猪、梅山猪和沙乌头猪群体中分别检测到 7 个、20 个和 9 个高频 ROH 区域, 分别注释到 414 个、547 个和 151 个基因(表 4)。枫泾猪、梅山猪和沙乌头猪没有共同的高频 ROH 区域, 枫泾猪和梅山猪在第 14 号染色体上有共同高频 ROH 区域, 梅山猪、沙乌头猪在第 8 号、11 号染色体上有共同高频 ROH 区域(图 3)。

利用 DAVID 数据库分别对枫泾猪、梅山猪和沙乌头猪群体中注释到的 414 个、547 个和 151 个基因进行 GO 功能富集分析和 KEGG Pathway 富集分析, 结果分别见表 5、表 6 和表 7。分析得出, 枫泾猪群体的 414 个基因中的 136 个基因显著富集于谷胱甘肽代谢过程、肝脏发育、炎症反应的负调控等 22 个 GO 条目和朊病毒病等 4 个信号通路; 梅山猪群体的 547 个基因中的 134 个基因显著富集于半胱氨酸型内肽酶活性的负调节、谷胱甘肽代谢过程、类黄酮葡萄糖醛酸化等 22 个 GO 条目和 11 个信号通路; 沙乌头猪群体的 151 个基因中的 24 个基因显著富集于类黄酮葡萄糖醛酸化、类黄酮生物合成过程、生物矿物组织发育等 11 个 GO 条目和抗坏血酸和醛糖酸盐代谢等 9 个信号通路。分别将枫泾猪群体的 136 个基因、梅山猪群体的 134 个基因和沙乌头猪群体的 24 个基因与猪 QTL 数据库进行比对, 得到枫泾猪候选功能基因 *HNF1A*、*CDH17*、*LIF*、*DECRI*、*ARL15*、*ACTN2*、*FABP5*、*FABP4* 和 *EWSRI* (表 5), 梅山猪候选功能基因 *CTSH*、*MSTN*、*GC*、*VEGFA*、*ZDHHC2*、*MTMR7*、*CDC5L* 和 *LIF* (表 6)。

## 3 讨论

ROH 可以指示亲缘的远近和近交的历史, ROH 片段越长, 家系内存在近世代近交的可能性越高<sup>[2, 6]</sup>。在 0~5.0 Mb 长度的 ROH 中, 枫泾猪所占比例最低, 沙乌头猪所占比例最高; 在大于 5.0 Mb 长度的 ROH 中, 枫泾猪所占比例最高, 沙乌头猪所占比例最低。该结果与 ROH 平均长度的结果一致(枫泾猪 ROH 的平均长度最长, 沙乌头猪 ROH 的平均长度最短)。3 个猪群中长度大于 5.0 Mb 的 ROH 所占比例都较高, 表明 3 个猪群经历了较高度近交, 可能与 3 个猪群的群体规模有关。在过去几十年里, 欧洲现代品种猪(主要是大白猪、长白

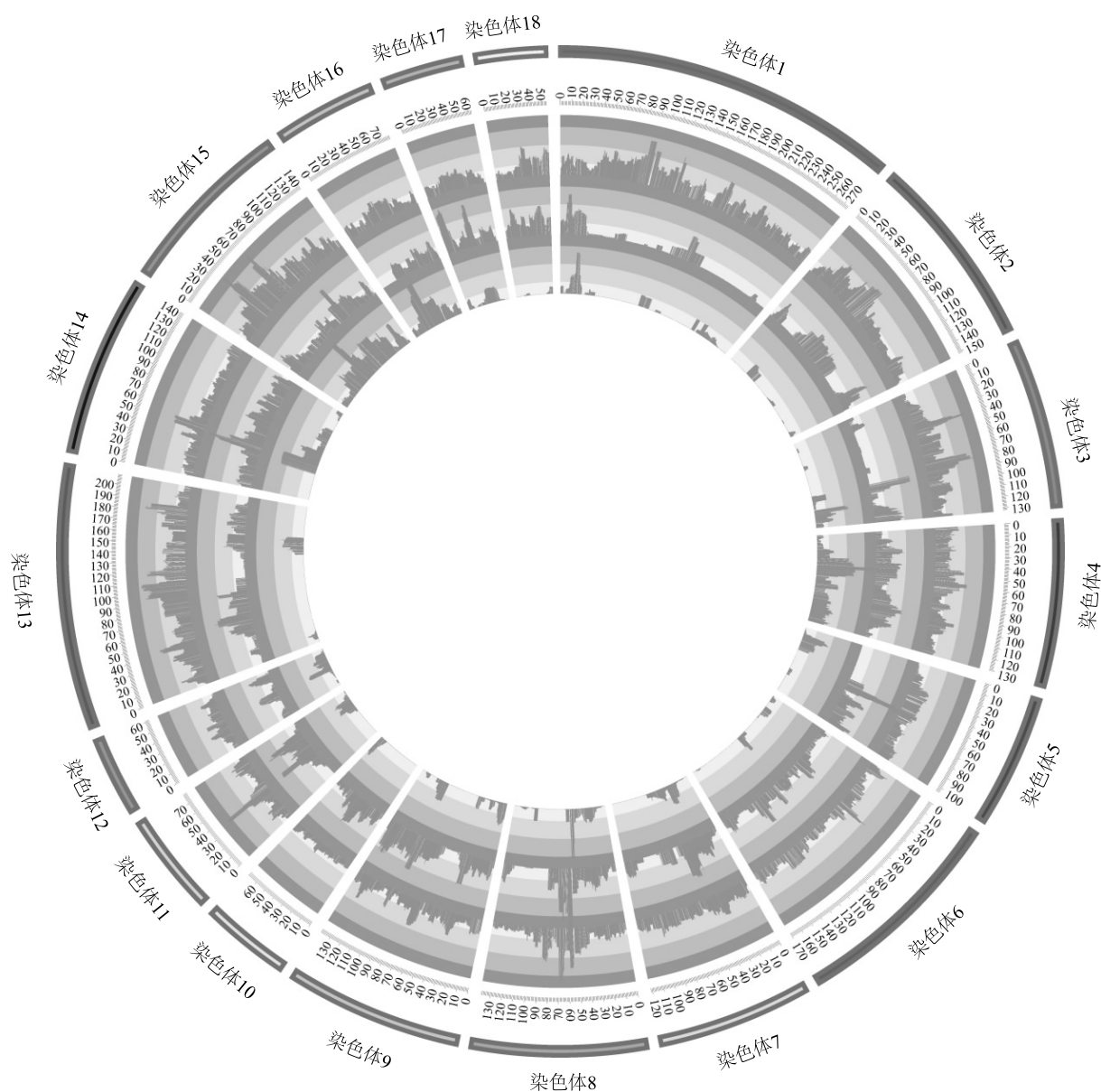
表 4 枫泾猪、梅山猪和沙乌头猪高频 ROH 区域

Table 4 High frequency ROH regions in Fengjing, Meishan and Shawutou pigs

群体	染色体	起始位置 (bp)	结束位置 (bp)	SNP 数量 (个)
枫泾猪	1	22 146 808	26 423 644	52
	3	11 0763 880	11 1051 475	5
	4	39 112 632	57 270 379	230
	8	49 420 599	53 587 541	56
	14	40 005 620	57 154 038	172
	15	53 340 080	60 677 580	83
	16	21 282 821	33 184 787	148
梅山猪	1	7 691 626	13 079 664	74
	2	138 401 227	139 202 583	10
	3	85 192 074	91 771 375	77
	4	95 557 016	95 601 325	2
	5	51 664 010	57 206 645	69
	7	34 502 131	34 562 337	3
	7	38 576 895	48 435 583	126
	8	55 063 034	69 866 924	198
	9	37 888 918	37 912 741	2
	9	92 955 449	98 824 983	88
	9	127 178 495	127 524 344	6
	9	129 091 764	131 787 116	41
	11	35 002 876	40 308 946	90
	11	58 175 011	63 590 389	59
	13	54 604 639	61 718 862	98
	13	63 166 998	63 556 910	7
	14	44 752 740	54 065 259	113
	15	90 255 657	94 627 948	85
	17	1 244 771	8 471 067	103
沙乌头猪	17	30 178 245	34 780 416	56
	3	39 035 953	40 561 777	41
	3	78 344 626	79 604 495	19
	3	80 895 719	81 276 729	9
	8	65 779 576	69 538 176	49
	11	34 538 535	38 031 870	80
	13	106 101 734	106 907 893	13
	15	33 196 684	38 253 783	85
	15	40 099 166	40 815 412	4
	15	62 628 159	65 057 504	33

猪和杜洛克猪)因瘦肉率高、生长速度快、饲料转化率高而不断被引进中国,目前占据着中国养猪业的统治地位,导致中国地方猪的养殖数量显著减少,枫泾猪、梅山猪和沙乌头猪的种群规模也因此受到影

响。在本研究中,枫泾猪、梅山猪和沙乌头猪样品采集于保种场,种群数量有限,从而导致近世代较高度度的近交。



每个环代表 1 个猪群体,从内到外的环分别为枫泾猪、梅山猪、沙乌头猪,柱子代表 ROH 频率。图中数据表示染色体的位置信息。

图3 ROH 分布频率的 Rcircos 图

Fig.3 Rcircos diagram of the ROH frequencies distribution

研究发现,基于 ROH 的近交系数( $F_{ROH}$ )与基于系谱信息的近交系数( $F_{PED}$ )间的相关系数随着 ROH 长度的增加而提高,在群体系谱信息不全的情

况下, $F_{ROH}$ 可以替代  $F_{PED}$ 来评估近交程度<sup>[7]</sup>。枫泾猪、梅山猪和沙乌头猪保种场早期的猪主要从主产区及周边长期从事地方猪饲养的农户收集而来,系

谱信息不全,因此需要通过基于 ROH 的近交系数 ( $F_{\text{ROH}}$ ) 来评估其近交程度。本研究基于 ROH 计算了每个猪群的 4 种近交系数 ( $F_{\text{ROH } 1.0 \sim 5.0 \text{ Mb}}$ 、 $F_{\text{ROH } 5.1 \sim 10.0 \text{ Mb}}$ 、 $F_{\text{ROH } > 10.0 \text{ Mb}}$  和  $F_{\text{ROH total}}$ ), 结果表明, 3 个猪群中枫泾猪的近交程度最低, 梅山猪的近交程度最高。与  $F_{\text{ROH } 1.0 \sim 5.0 \text{ Mb}}$ 、 $F_{\text{ROH } 5.1 \sim 10.0 \text{ Mb}}$  相比, 3 个猪群

的  $F_{\text{ROH } > 10.0 \text{ Mb}}$  与  $F_{\text{ROH total}}$  最接近, 与 4 种近交系数间的相关性分析结果一致,  $F_{\text{ROH } > 10.0 \text{ Mb}}$  和  $F_{\text{ROH total}}$  间的相关系数最高。Zhong 等<sup>[5]</sup> 利用简化基因组测序数据计算了金华猪的  $F_{\text{ROH } 1.0 \sim 5.0 \text{ Mb}}$ 、 $F_{\text{ROH } 5.1 \sim 10.0 \text{ Mb}}$ 、 $F_{\text{ROH } > 10.0 \text{ Mb}}$  和  $F_{\text{ROH all}}$ , 结果表明,  $F_{\text{ROH } > 10.0 \text{ Mb}}$  和  $F_{\text{ROH all}}$  间的相关系数最高, 与本研究结果一致。

表 5 枫泾猪注释基因的 GO 和 KEGG Pathway 富集分析

Table 5 GO and KEGG pathway enrichment analysis of annotated genes in Fengjing pigs

功能分类	功能注释号	功能描述	基因数 (个)	P 值	候选基因
GO-BP	GO:0006749	谷胱甘肽代谢过程	4	$1.05 \times 10^{-2}$	
GO-BP	GO:0001889	肝脏发育	4	$1.67 \times 10^{-2}$	<i>HNFI1A</i>
GO-BP	GO:0050728	炎症反应的负调控	5	$1.80 \times 10^{-2}$	
GO-BP	GO:0008299	异戊二烯生物合成过程	3	$2.14 \times 10^{-2}$	
GO-BP	GO:0048536	脾脏发育	4	$2.45 \times 10^{-2}$	<i>CDH17</i>
GO-BP	GO:2001235	凋亡信号通路的正调控	3	$2.53 \times 10^{-2}$	
GO-BP	GO:0050731	肽基酪氨酸磷酸化的正调控	4	$2.67 \times 10^{-2}$	
GO-BP	GO:0051726	细胞周期调节	5	$3.16 \times 10^{-2}$	
GO-BP	GO:0048538	胸腺发育	4	$3.40 \times 10^{-2}$	
GO-BP	GO:1902916	蛋白质多泛素化的正调控	2	$4.16 \times 10^{-2}$	
GO-BP	GO:0046888	荷尔蒙分泌负调控	2	$4.16 \times 10^{-2}$	<i>LIF</i>
GO-BP	GO:0015804	中性氨基酸转运	2	$4.16 \times 10^{-2}$	
GO-BP	GO:0042517	Stat3 蛋白酪氨酸磷酸化的正调控	3	$4.86 \times 10^{-2}$	<i>LIF</i>
GO-CC	GO:0005579	膜攻击复合物	3	$3.78 \times 10^{-3}$	
GO-CC	GO:0070062	胞外分泌体	44	$2.13 \times 10^{-2}$	<i>DECRI</i> 、 <i>ARL15</i> 、 <i>ACTN2</i> 、 <i>FABP5</i>
GO-CC	GO:0005737	细胞质	55	$3.55 \times 10^{-2}$	<i>HNFI1A</i> 、 <i>LIF</i> 、 <i>FABP4</i>
GO-MF	GO:0005212	晶状体的结构成分	4	$2.24 \times 10^{-3}$	
GO-MF	GO:0015179	L-氨基酸跨膜转运蛋白活性	3	$8.60 \times 10^{-3}$	
GO-MF	GO:0001158	增强子序列特异性 DNA 结合	3	$3.38 \times 10^{-2}$	
GO-MF	GO:0044822	poly(A) RNA 结合	22	$3.69 \times 10^{-2}$	<i>EWSR1</i>
GO-MF	GO:0005412	葡萄糖: 钠同向转运体活性	2	$4.16 \times 10^{-2}$	
GO-MF	GO:0015297	逆向转运蛋白活性	3	$4.34 \times 10^{-2}$	
KEGG Pathway	ssc05020	朊病毒病	4	$2.66 \times 10^{-2}$	
KEGG Pathway	ssc04151	PI3K-Akt 信号通路	13	$2.68 \times 10^{-2}$	
KEGG Pathway	ssc04810	肌动蛋白细胞骨架的调节	9	$3.94 \times 10^{-2}$	<i>ACTN2</i>
KEGG Pathway	ssc04120	泛素介导的蛋白质水解	7	$4.86 \times 10^{-2}$	

表 6 梅山猪注释基因的 GO 和 KEGG Pathway 富集分析

Table 6 GO and KEGG pathway enrichment analysis of annotated genes in Meishan pigs

功能分类	功能注释号	功能描述	基因数 (个)	P 值	候选基因
GO-BP	GO:2000117	半胱氨酸型内肽酶活性的负调节	7	$5.16 \times 10^{-7}$	
GO-BP	GO:0006749	谷胱甘肽代谢过程	7	$3.24 \times 10^{-5}$	
GO-BP	GO:0052696	类黄酮葡萄糖醛酸化	4	$1.90 \times 10^{-3}$	
GO-BP	GO:0009813	类黄酮生物合成过程	4	$1.90 \times 10^{-3}$	
GO-BP	GO:0007626	运动行为	6	$5.59 \times 10^{-3}$	
GO-BP	GO:0070371	ERK1 和 ERK2 级联	4	$9.01 \times 10^{-3}$	CTSH
GO-BP	GO:0072520	生精小管发育	3	$3.37 \times 10^{-2}$	
GO-BP	GO:0031214	生物矿物组织发育	3	$4.04 \times 10^{-2}$	
GO-BP	GO:2001235	凋亡信号通路的正调控	3	$4.75 \times 10^{-2}$	CTSH
GO-BP	GO:0016197	内体转运	3	$4.75 \times 10^{-2}$	
GO-CC	GO:0005615	细胞外间隙	36	$3.31 \times 10^{-3}$	MSTN、CTSH、GC、VEGFA
GO-CC	GO:0072562	血液微粒	6	$2.40 \times 10^{-2}$	GC
GO-CC	GO:0005887	质膜的组成部分	29	$3.59 \times 10^{-2}$	ZDHHC2
GO-CC	GO:0014069	突触后密度	4	$3.63 \times 10^{-2}$	
GO-CC	GO:0016020	膜	29	$4.63 \times 10^{-2}$	MTMR7、CDC5L、VEGFA
GO-MF	GO:0004869	半胱氨酸型内肽酶抑制剂活性	7	$2.84 \times 10^{-5}$	
GO-MF	GO:0002020	蛋白酶结合	7	$2.84 \times 10^{-5}$	
GO-MF	GO:0004364	谷胱甘肽转移酶活性	6	$4.33 \times 10^{-4}$	
GO-MF	GO:0015020	葡萄糖醛酸转移酶活性	4	$3.34 \times 10^{-3}$	
GO-MF	GO:0001135	转录因子活性, RNA 聚合酶 II 转录因子募集	3	$7.89 \times 10^{-3}$	LIF、CDC5L
GO-MF	GO:0005126	细胞因子受体结合	3	$2.63 \times 10^{-2}$	
GO-MF	GO:0005245	电压门控钙通道活性	4	$2.78 \times 10^{-2}$	
KEGG Pathway	ssc00980	细胞色素 P450 对异生物物质的代谢	10	$4.87 \times 10^{-6}$	
KEGG Pathway	ssc00982	药物代谢-细胞色素 P450	10	$4.87 \times 10^{-6}$	
KEGG Pathway	ssc05204	化学致癌	10	$2.72 \times 10^{-5}$	
KEGG Pathway	ssc00983	药物代谢-其他酶	6	$1.09 \times 10^{-3}$	
KEGG Pathway	ssc00053	抗坏血酸和醛糖酸盐代谢	5	$2.11 \times 10^{-3}$	
KEGG Pathway	ssc00480	谷胱甘肽代谢	7	$2.28 \times 10^{-3}$	
KEGG Pathway	ssc00140	类固醇激素生物合成	7	$2.52 \times 10^{-3}$	
KEGG Pathway	ssc00040	戊糖和葡萄糖醛酸的相互转化	5	$2.99 \times 10^{-3}$	
KEGG Pathway	ssc00860	卟啉和叶绿素代谢	5	$1.24 \times 10^{-2}$	
KEGG Pathway	ssc00562	磷酸肌醇代谢	6	$3.34 \times 10^{-2}$	MTMR7
KEGG Pathway	ssc04070	磷脂酰肌醇信号系统	7	$4.39 \times 10^{-2}$	MTMR7

表 7 沙乌头猪注释基因的 GO 和 KEGG Pathway 富集分析

Table 7 GO and KEGG pathway enrichment analysis of annotated genes in Shawutou pigs

功能分类	功能注释号	功能描述	基因数 (个)	P 值
GO-BP	GO:0052696	类黄酮葡萄糖醛酸化	4	$3.40 \times 10^{-5}$
GO-BP	GO:0009813	类黄酮生物合成过程	4	$3.40 \times 10^{-5}$
GO-BP	GO:0031214	生物矿物组织发育	3	$2.98 \times 10^{-3}$
GO-BP	GO:0070175	釉质矿化的正调控	2	$1.51 \times 10^{-2}$
GO-BP	GO:0042475	含牙本质牙齿的形成	3	$2.28 \times 10^{-2}$
GO-CC	GO:0043231	细胞内膜结合的细胞器	5	$7.13 \times 10^{-3}$
GO-CC	GO:0048179	激活素受体复合物	2	$2.22 \times 10^{-2}$
GO-MF	GO:0015020	葡萄糖醛酸转移酶活性	4	$4.04 \times 10^{-5}$
GO-MF	GO:0046872	金属离子结合	12	$1.10 \times 10^{-2}$
GO-MF	GO:0030345	牙釉质的结构成分	2	$1.92 \times 10^{-2}$
GO-MF	GO:0016361	激活素受体活性, I 型	2	$2.55 \times 10^{-2}$
KEGG Pathway	ssc00053	抗坏血酸和醛糖酸盐代谢	5	$9.00 \times 10^{-6}$
KEGG Pathway	ssc00040	戊糖和葡萄糖醛酸的相互转化	5	$1.32 \times 10^{-5}$
KEGG Pathway	ssc00140	类固醇激素生物合成	6	$2.02 \times 10^{-5}$
KEGG Pathway	ssc00983	药物代谢-其他酶	5	$3.95 \times 10^{-5}$
KEGG Pathway	ssc00860	卟啉和叶绿素代谢	5	$6.55 \times 10^{-5}$
KEGG Pathway	ssc00980	细胞色素 P450 对异生物质的代谢	5	$2.78 \times 10^{-4}$
KEGG Pathway	ssc00982	药物代谢-细胞色素 P450	5	$2.78 \times 10^{-4}$
KEGG Pathway	ssc00830	视黄醇代谢	5	$4.06 \times 10^{-4}$
KEGG Pathway	ssc05204	化学致癌	5	$6.08 \times 10^{-4}$

本研究通过将富集到生物学过程和信号通路的基因与猪 QTL 数据库进行比对,鉴定到 9 个枫泾猪候选功能基因和 8 个梅山猪候选功能基因,其中 *LIF* 基因是共同的候选功能基因,*HNF1A* 基因是猪肉质和胴体性状的候选基因<sup>[8]</sup>,*CDH17* 基因影响公猪睾丸中雄烯酮水平<sup>[9]</sup>,*LIF* 基因影响猪的产活仔数和 21 日龄仔猪数<sup>[10]</sup>,*DECRI* 基因与杜洛克猪脂质组成性状显著相关<sup>[11]</sup>,*ARL15* 基因是猪的眼肌面积的候选基因<sup>[12]</sup>,*ACTN2* 基因与猪肉的烹饪损失有关<sup>[13]</sup>,*FABP5* 和 *FABP4* 基因与猪的脂肪性状相关<sup>[14-15]</sup>,*EWSRI* 基因影响猪肉肉色<sup>[13]</sup>,*CTSH* 基因与饲料增质量比相关<sup>[16]</sup>,*MSTN* 基因与猪日增质量和体质量等性状显著相关<sup>[17]</sup>,*GC* 基因与皮特兰猪的 pH<sub>24h</sub> 相关<sup>[18]</sup>,*VEGFA* 基因影响猪后腿步态评分<sup>[19]</sup>,*ZDHHC2* 基因是鲁莱黑猪肌内脂肪含量的候选基因<sup>[20]</sup>,*MTMR7* 基因与猪血液中平均红细胞血红蛋白含量有关<sup>[21]</sup>,*CDC5L* 基因与猪血细胞比容和

血红蛋白相关<sup>[22]</sup>。本研究鉴定到的候选功能基因与猪的生长发育、胴体性状、肉质性状和繁殖性状相关,下一步试验将利用候选基因法研究候选功能基因与相关性状的关系,从而为解析猪生长发育、胴体性状、肉质性状和繁殖性状的分子机制提供参考。

#### 参考文献:

- [1] GIBSON J, MORTON N E, COLLINS A. Extended tracts of homozygosity in outbred human populations[J]. Human Molecular Genetics, 2006, 15(5): 789-795.
- [2] KELLER M C, VISSCHER P M, GODDARD M E. Quantification of inbreeding due to distant ancestors and its detection using dense single nucleotide polymorphism data[J]. Genetics, 2011, 189(1): 237-249.
- [3] MASTRANGELO S, SARDINA M T, TOLONE M, et al. Genome-wide identification of runs of homozygosity islands and associated genes in local dairy cattle breeds[J]. Animal, 2018, 12(12): 2480-2488.
- [4] SAURA M, FERNÁNDEZ A, VARONA L, et al. Detecting in-

- breeding depression for reproductive traits in Iberian pigs using genome-wide data[J]. *Genetics Selection Evolution*, 2015, 47(1): 1.
- [5] ZHONG X, SUN H, ZHANG Z, et al. Assessment of autozygosity derived from runs of homozygosity in Jinhua pigs disclosed by sequencing data[J]. *Frontiers in Genetics*, 2019, 10: 274.
- [6] HOWRIGAN D P, SIMONSON M A, KELLER M C. Detecting autozygosity through runs of homozygosity: a comparison of three autozygosity detection algorithms[J]. *BMC Genomics*, 2011, 12(1): 460.
- [7] PERIPOLLI E, STAFUZZA N B, MUNARI D P, et al. Assessment of runs of homozygosity islands and estimates of genomic inbreeding in Gyr (*Bos indicus*) dairy cattle[J]. *BMC Genomics*, 2018, 19(1): 34.
- [8] KAYAN A, UDDIN M J, KOCAMIS H, et al. Association and expression analysis of porcine HNF1A gene related to meat and carcass quality traits[J]. *Meat Science*, 2013, 94(4): 474-479.
- [9] DRAG M H, KOGELMAN L J, MARIBO H, et al. Characterization of eQTLs associated with androstenedione by RNA sequencing in porcine testis[J]. *Physiological Genomics*, 2019, 51(10): 488-499.
- [10] MUCHA A, ROPKA-MOLIK K, PIÓRKOWSKA K, et al. Effect of *EGF*, *AREG* and *LIF* genes polymorphisms on reproductive traits in pigs[J]. *Animal Reproduction Science*, 2013, 137(1/2): 88-92.
- [11] RAMÍREZ O, QUINTANILLA R, VARONA L, et al. DECR1 and ME1 genotypes are associated with lipid composition traits in Duroc pigs[J]. *Journal of Animal Breeding and Genetics*, 2014, 131(1): 46-52.
- [12] ZHUANG Z W, LI S Y, DING R R, et al. Meta-analysis of genome-wide association studies for loin muscle area and loin muscle depth in two Duroc pig populations[J]. *PLoS One*, 2019, 14(6): e0218263.
- [13] LI X P, KIM S W, DO K T, et al. Analyses of porcine public SNPs in coding-gene regions by re-sequencing and phenotypic association studies[J]. *Molecular Biology Reports*, 2011, 38(6): 3805-3820.
- [14] ESTELLÉ J, PÉREZ-ENCISO M, MERCADÉ A, et al. Characterization of the porcine *FABP5* gene and its association with the *FAT1* QTL in an Iberian by Landrace cross[J]. *Animal Genetics*, 2006, 37(6): 589-591.
- [15] MERCADÉ A, PÉREZ-ENCISO M, VARONA L, et al. Adipocyte fatty-acid binding protein is closely associated to the porcine locus on chromosome 4[J]. *Journal of Animal Science*, 2006, 84(11): 2907-2913.
- [16] RUSSO V, FONTANESI L, SCOTTI E, et al. Single nucleotide polymorphisms in several porcine cathepsin genes are associated with growth, carcass, and production traits in Italian Large White pigs[J]. *Journal of Animal Science*, 2008, 86(12): 3300-3314.
- [17] TU P A, SHIAU J W, DING S T, et al. The association of genetic variations in the promoter region of myostatin gene with growth traits in Duroc pigs[J]. *Animal Biotechnology*, 2012, 23(4): 291-298.
- [18] CINAR M U, KAYAN A, UDDIN M J, et al. Association and expression quantitative trait loci (eQTL) analysis of porcine *AMBP*, *GC* and *PPP1R3B* genes with meat quality traits[J]. *Molecular Biology Reports*, 2012, 39(4): 4809-4821.
- [19] GUO Y M, ZHANG X F, REN J, et al. A joint genomewide association analysis of pig leg weakness and its related traits in an F<sub>2</sub> population and a Sutai population[J]. *Journal of Animal Science*, 2013, 91(9): 4060-4068.
- [20] WANG Y P, NING C, WANG C, et al. Genome-wide association study for intramuscular fat content in Chinese Lulai black pigs[J]. *Asian Australasian Journal of Animal Sciences*, 2019, 32(5): 607-613.
- [21] WANG J Y, LUO Y R, FU W X, et al. Genome-wide association studies for hematological traits in swine[J]. *Animal Genetics*, 2013, 44(1): 34-43.
- [22] ZHANG Z Y, HONG Y, GAO J, et al. Genome-wide association study reveals constant and specific loci for hematological traits at three time stages in a White Duroc×Erhualian F<sub>2</sub> resource population[J]. *PLoS One*, 2013, 8(5): e63665.

(责任编辑:徐 艳)