

# 家养动物选择信号研究进展

潘章源<sup>1,2</sup>, 贺小云<sup>1</sup>, 王翔宇<sup>1</sup>, 郭晓飞<sup>1</sup>, 曹晓涵<sup>1,3</sup>, 胡文萍<sup>1</sup>, 狄冉<sup>1</sup>,  
刘秋月<sup>1</sup>, 储明星<sup>1</sup>

1. 中国农业科学院北京畜牧兽医研究所, 农业部畜禽遗传资源与种质创新重点实验室, 北京 100193;
2. 临沂大学生命科学学院, 临沂 276005;
3. 四川农业大学生命科学学院, 雅安 625014

**摘要:** 家养动物在人类生活中占有重要地位, 它们都经历驯化而来, 在自然和人工选择下, 适应了当地环境和人类需要, 形成了丰富多样的各类品种。驯化、自然和人工选择都会在基因组上留下选择信号。对这些选择信号进行研究, 可以直接挖掘到功能基因, 是目前最重要的功能基因筛选策略之一。当前已经对猪(*Sus scrofa*)、鸡(*Gallus gallus*)、牛(*Bos taurus*)、绵羊(*Ovis aries*)、山羊(*Capra hircus*)、犬(*Canis lupus familiaris*)及马(*Equus caballus*)等家养动物开展了选择信号研究, 并挖掘了大量功能基因。本文主要概述了选择信号的种类和检测方法及其在家养动物中的研究进展, 并讨论了选择信号分析的关键问题及其研究前景。

**关键词:** 家养动物; 选择信号; 驯化; 人工选择; 功能基因

## Selection signatures in domesticated animals

Zhangyuan Pan<sup>1,2</sup>, Xiaoyun He<sup>1</sup>, Xiangyu Wang<sup>1</sup>, Xiaofei Guo<sup>1</sup>, Xiaohan Cao<sup>1,3</sup>,  
Wenping Hu<sup>1</sup>, Ran Di<sup>1</sup>, Qiuyue Liu<sup>1</sup>, Mingxing Chu<sup>1</sup>

1. Key Laboratory of Farm Animal Genetic Resources and Germplasm Innovation of Ministry of Agriculture, Institute of Animal Science, Chinese Academy of Agricultural Sciences, Beijing 100193, China;
2. College of Life Science, Linyi University, Linyi 276005, China;
3. College of Life Science, Sichuan Agricultural University, Ya'an 625014, China

**Abstract:** Domesticated animals play an important role in human lives. All the animals were firstly domesticated from the wild, and then adapted to the local environment and human needs and formed various breeds. During the

收稿日期: 2016-05-30; 修回日期: 2016-08-12

基金项目: 国家自然科学基金项目(编号: 31472078, 31402041), 宁夏农林科学院科技创新先导资金项目(编号: DWJLC-2016001), 中国农业科学院科技创新工程(编号: ASTIP-IAS13), 国家肉羊产业技术体系专项(编号: CARS-39), 内蒙古自治区科技重大专项和内蒙古自治区战略性新兴产业发展专项资金资助[Supported by the National Natural Science Foundation of China (Nos.31472078, 31402041), the Science and Technology Innovation Guide Program of Ningxia Academy of Agriculture and Forestry Sciences (No. DWJLC-2016001), the Agricultural Science and Technology Innovation Program of China (No.ASTIP-IAS13), the Earmarked Fund for China Agriculture Research System (No.CARS-39), the Science and Technology Major Program of Inner Mongolia Autonomous Region of China, and the Special Fund for Development Plan for Strategic Emerging Industry of Inner Mongolia Autonomous Region of China]

作者简介: 潘章源, 博士, 研究方向: 动物遗传育种。E-mail: pzq170450077@163.com

通讯作者: 储明星, 研究员, 博士生导师, 研究方向: 动物遗传育种。E-mail: mxchu@263.net

刘秋月, 副研究员, 研究方向: 动物遗传育种。E-mail: qiuyue1983921@163.com

DOI: 10.16288/j.ycz.16-195

网络出版时间: 2016/10/20 9:04:26

URI: <http://www.cnki.net/kcms/detail/11.1913.R.20161020.0904.002.html>

long process, natural and artificial selection leave selection signals in the genomes. Seeking out these selection signals is one of the most important strategies in screening functional genes. Such studies have been carried out in pigs, chickens, cattle, sheep, goats, dogs, horses and other domestic animals. Lots of functional genes had been identified. Here we provide an overview of categories and detection methods of selection signals, outline progresses of selection signal studies in domestic animals, and discuss the key issues in selection signal analysis and its prospects.

**Keywords:** domesticated animals; selection signature; domestication; artificial selection; functional genes

家养动物在人类社会中扮演着重要的角色,是人类主要的生活物质来源。由野生动物变为家养动物,必须经历一个重要过程,即人工驯化,在该过程中,它们从野生状态变成了适应人类需要的状态,由野性十足变成性情温顺,或由低产变成高产<sup>[1]</sup>。随后人类根据自身的需要,采用人工选择,又培育出不同品种。这些过程会在基因组上表现为一些位点(Locus)受到选择,并留下选择特征的现象。在中性进化理论下,一个新的突变往往需要很长一段时间才能够在群体中达到一个较高的频率,并且这些突变周围的连锁不平衡程度会因重组率的影响而在这段时间内几乎完全衰减降解<sup>[2]</sup>。因此,基因组上绝大多数未受到选择作用的位点会始终处于随机漂变状态,彼此之间形成的连锁不平衡(Linkage disequilibrium, LD)容易衰减,单倍型(Haplotype)长度相对较短。然而在选择的作用下,群体有利等位基因频率则会在较短的时间内达到一个极值,重组的作用会受到一定程度的对冲而不能对长范围单倍型造成实质性的降解;同时,选择作用下的连锁不平衡会造成选择位点附近的中性位点的基因频率随之增加进而形成长范围的单倍型纯合<sup>[3]</sup>。群体遗传学中,将这种由选择作用造成的部分染色体片段的多态性降低称为选择性清除(Selective sweep)<sup>[4]</sup>;选择位点周围的中性位点得益于选择作用而表现出基因频率迅速增加,被称为“搭车”效应(Hitchhiking effect)<sup>[5,6]</sup>。因此选择性清除和“搭车”效应都是选择作用在基因组上留下的明显特征,被称为选择信号(Selection signature)。生物体选择作用往往针对某些性状,而内在机制是对某些基因进行选择,被选择的基因与该生物性状有直接关系,对选择信号的检测能揭示受选择的基因。因此找到这些选择信号,相当于找到行使相应功能的候选基因,对了解物种的进化和获得经济性状基因都具有重大意义。

## 1 选择信号种类

选择(Selection)定义为在人类或自然界的干预下,生物群体在世的传递过程中,某种基因型个体的比例发生变化的群体遗传学现象。在漫长的生物进化过程中,生物会受到来自外界环境不断的选择,这种选择在基因组上主要表现为遗传多样性的改变,在表型上表现为适合度的改变<sup>[7]</sup>。根据适合度变化的角度可以将选择分为以下几种形式:正选择(Positive selection)指有利突变在群体中得到固定的过程,表现为基因杂合度不断降低;平衡选择(Balancing selection)则表现为维持群体的多态性变异及杂合优势;负选择(也称净化选择, Purifying selection)指去除有害突变的作用,表现为突变型频率不断降低;还有一些比较少见的选择,包括最高生殖性状才可以产生子代的截断选择(Truncation selection);保持中间态位点的稳定选择(Stability selection);有利于一种性状极高值的分裂选择(Disruptive selection)等。选择种类对应选择信号种类,如正选择对应的选择信号为正选择信号。家养动物中最常见的是正选择信号,即会形成选择性清除的特征,所以大部分研究是针对该特征进行正选择信号筛选。

## 2 选择信号的检测方法

随着生物技术的发展,对基因组的研究不断深入,已经从传统的微卫星、单基因 SNP 位点到现在的全基因组遗传多样性分析,尤其是 SNP 芯片和高通量测序技术的发展,为选择信号的研究提供了坚实的数据基础。在此基础上,发展出一系列选择信号检测方法,均基于不同的原理,各具优势。

### 2.1 基于群体分化的检测方法

自然界的生物群体由于地域隔离,环境差异产

生的选择压力会呈现出不同的群体结构, 目前有几种方法可检测这种群体间的选择压力, 包括  $F_{ST}$  检验<sup>[8]</sup>、LSBL(Lineage-specific-branch-length)<sup>[9]</sup>和  $di$ <sup>[10]</sup>。

$F_{ST}$  能通过群体间的差异解释遗传变异的比例。全基因组  $F_{ST}$  检验可以通过比较单个位点  $F_{ST}$  值找到局部受到选择的基因, 但由于一些物种的基因组比较大, 在研究中会使用滑框(大动物常用 50 k 的滑框)进行  $F_{ST}$  值的计算。在中性突变下,  $F_{ST}$  的大小主要受迁移和遗传漂变等因素的影响。但当一个突变受到人工选择或自然选择时, 其频率的升高就会增大种群的分化水平,  $F_{ST}$  值也越大( $0 < F_{ST} < 1$ , 1 表示该区域种群完全分化)。通过揭示变异的“离群”模式、多样性的缺失以及连锁标记受到选择的影响, 就可以实现对选择痕迹的检测。 $F_{ST}$  的计算具有多种方式, 目前最常用的方式是通过均方误差, 计算公式如下:

$$F_{ST} = \frac{MSP - MSG}{MSP + (n_c - 1)MSG}$$

其中  $MSG$  是群体内的均方误差,  $MSP$  是群体间的均方误差,  $n_c$  是校正后整个群体的平均样本大小, 此方法主要适用于不同群体之间选择信号的检测。

LSBL 计算公式是由 Shriver 等<sup>[9]</sup>首次提出。它通过成对的  $F_{ST}$  距离来计算(图 1): 首先需要定义 3 个群体 A、B、C, 3 个群体之间计算  $F_{ST}$  值, 得到彼此之间的  $F_{ST}$  距离( $d_{AB}$ 、 $d_{BC}$  和  $d_{AC}$ ), 此时 LSBL 计算公式为:  $x = (d_{AB} + d_{AC} - d_{BC})/2$ ;  $y = (d_{AB} + d_{BC} - d_{AC})/2$ ;  $z = (d_{AC} + d_{BC} - d_{AB})/2$ 。

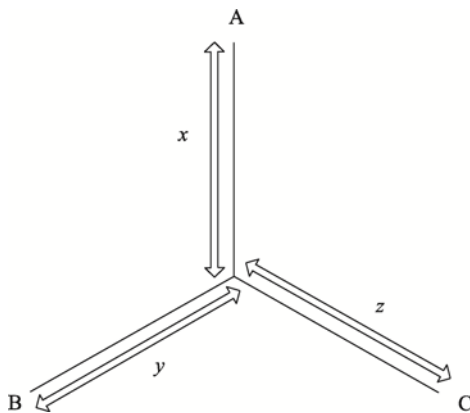


图 1 3 个群体间 LSBL 计算示意图

Fig. 1 A diagram demonstrating how branch lengths are estimated for a network with three populations

$di$  是由 Akey 等<sup>[10]</sup>首先提出来的计算方法, 同样也是基于  $F_{ST}$  计算多个群体间的差异, 公式如下:

$$di = \sum_{j \neq i} \frac{F_{ST}^{ij} - E[F_{ST}^{ij}]}{sd[F_{ST}^{ij}]}$$

$E[F_{ST}^{ij}]$  和  $sd[F_{ST}^{ij}]$  表示为  $i$  和  $j$  品种之间所有的 SNPs 的两两  $F_{ST}$  值的期望值和标准差。

## 2.2 基于位点频率谱的检测方法

基于基因型频率的计算方法主要有 Tajima's  $D$  检验<sup>[11]</sup>、 $H_p$  检验<sup>[5]</sup>等。这些方法只用考虑基因频谱的变化, 不需要考虑连锁因素。最早提出来的是 Tajima's  $D$  检验, 它通过比较群体突变率的参数估计值  $\theta_w$  和  $\pi$  的差异来检测正向选择。其公式如下:

$$D = \frac{\pi - \theta_w}{\sqrt{V(\pi - \theta_w)}}$$

其中  $\theta_w$  为群体遗传参数, 其理论值为  $4N_e\mu$ ,  $N_e$  为有效群体大小,  $\mu$  为突变频率。

理论上, 在中性进化条件下,  $\theta_w$  和  $\pi$  值近似相等, Tajima's  $D$  值为 0, 但是当等位基因频率偏离中性进化模型时, Tajima's  $D$  值就会显著偏离 0 值。当 Tajima's  $D > 0$  时, 说明可能存在群体瓶颈效应、平衡选择或突然的群体萎缩。当 Tajima's  $D < 0$  时, 说明存在近期选择作用、搭车效应或群体通过瓶颈效应后大量扩群。

$H_p$  检验也是目前常用的检测方法, 首先计算每个突变位点的最大等位基因数( $n_{MAJ}$ )和最小等位基因数( $n_{MIN}$ ), 然后按照滑框计算  $H_p$  值, 计算公式为:

$$H_p = \frac{2 \sum n_{MAJ} \sum n_{MIN}}{(\sum n_{MAJ} \sum n_{MIN})^2}$$

其中  $\sum n_{MAJ}$  和  $\sum n_{MIN}$  分别代表滑框内所有 SNP 的  $n_{MAJ}$  和  $n_{MIN}$  之和。当区域受到选择时,  $H_p$  值变低, 受到选择越强,  $H_p$  值越低。有些研究将其转换成标准化杂合度  $ZH_p$ , 转换公式为  $ZH_p = |H_p - \mu H_p| / \sigma H_p$ ,  $\mu H_p$  和  $\sigma H_p$  分别代表  $H_p$  的平均值和标准差<sup>[12]</sup>。

然而这种基于频谱的方法存在一定局限性, 无法判断选择类型, 正选择还是负选择; 无法排除连锁不平衡的效应, 可能是搭车效应, 无法确定功能位点。

### 2.3 基于连锁不平衡(LD)和单倍型的检测方法

基于 LD 和单倍型的方法主要有 EHH(Extended haplotype homozygosity)检验<sup>[13]</sup>、XP-EHH(Cross population extended haplotype homozygosity)检验<sup>[14]</sup>和 *iHS*(Integrated haplotype score)检验<sup>[15]</sup>。同一条染色体上的核苷酸位点存在连锁关系,多个位点连锁能形成“单倍型”。在中性条件下,基因组很难形成长范围的连锁不平衡的单倍型,因为新突变需要经历漫长的遗传漂变才能达到较高的频率,而这么漫长的时间内会发生大量的基因重组,使得这种连锁不断地被打破。而当群体处于正选择作用下时,致因突变及其连锁位点在正选择的作用下,在短时间内达到较高频率,形成大片段的纯合单倍型。扩展单倍型纯合度(EHH)检验正是基于这样的特征来筛选受选择基因,其计算公式如下:

$$EHH = \sum_{i=1}^s \frac{(2^{e_{ii}})}{(2^{c_i})}$$

其中  $e_{ii}$  表示第  $i$  个扩展单倍型在样本中的数目,  $c_i$  表示第  $i$  个核心单倍型在样本中的数目,  $S$  是独立单倍型的数目。

EHH 检验也存在局限性,即由于忽略了不同染色体片段之间重组率的差异,而并不适用于基因组范围内的选择信号强度的检测。认识到 EHH 检测的局限性后, Sabeti 等<sup>[14]</sup>又对其进行了改进,研发出了交叉群体扩展的单倍型纯合度(XP-EHH)检验。XP-EHH 计算基于 EHH 值,计算公式为:

$$XP-EHH = \ln \left( \frac{I_A}{I_B} \right)$$

其中  $I_A$  表示试验群体的 EHH 值关于遗传距离的积分,  $I_B$  表示参考群体 EHH 值关于遗传距离的积分,一般把群体  $B$  当作参考群体。

*iHS* 也是基于 EHH 值改进的检验方法<sup>[15]</sup>,主要是针对一些家系或拥有祖先样品群体。计算公式如下:

$$iHS = \ln \left( \frac{iHH_A}{iHH_D} \right)$$

其中 *iHS* 表示整合单倍型值,  $iHH_A$  表示祖先等位基因的 EHH 整合得分值,  $iHH_D$  表示推断等位基因的 EHH 整合得分值。当 *iHS* 值为较大的正值时,

表示单倍型可能与祖先一致,当 *iHS* 值为较大的负值时,表示可能具有新衍生的单倍型<sup>[16]</sup>。*iHS* 检验是基于祖先等位基因的一种检验方法,其优势是通过祖先的信息能更准确地观察到进化中受选择的情况。而其劣势是现实中很难获得祖先的群体信息,因为很多古老的品种已经灭绝或野生品种很难获得。

EHH 检验、XP-EHH 检验、*iHS* 检验都是基于 LD,所以它们都对正选择信号特别敏感和对近期发生的选择敏感。但是 LD 很容易被时间打断,所以这 3 种方法很难检测到已经发生的古老的选择信号<sup>[17]</sup>。

### 2.4 新检测方法

XP-CLR (Cross-population composite likelihood ratio)<sup>[17]</sup>是一种相对比较新的选择信号检测方法,它的计算主要基于两个群体间多基因座(Multilocus)等位基因频率差异,因此该方法比基于位点频率谱的方法对选择造成的等位基因频率偏离更灵敏,检出率更高。目前已经用于家鼠(*Mus musculus*)<sup>[18]</sup>、玉米(*Zea mays ssp. mays*)<sup>[19]</sup>、黄瓜(*Cucumis sativus* L.)<sup>[20]</sup>、大豆(*Glycine max*)<sup>[21]</sup>等的选择信号挖掘。该方法计算过程比较复杂,需要原作者<sup>[17]</sup>提供软件。

HapFLK<sup>[22]</sup>检测方法是 2013 年提出的,主要是基于样品群体的分层结构,该方法同时应用了单倍型信息和群体分层结构使得受选择位点检出率更高,且在检测瓶颈效应和迁移方面比先前的方法更灵敏,可以检测温和清除(Soft sweeps)、未完成的清除(Incomplete sweeps)、发生在几个群体中的清除(Sweeps occurring in several populations)和直接作用于单倍型的清除(Selection acting directly on haplotypes)。该方法已经用于一些物种选择信号的筛选<sup>[23-25]</sup>。该方法计算过程已编成软件,下载地址 <https://forge-dga.jouy.inra.fr/projects/hapflk>。

各选择信号检测方法归纳见表 1。

## 3 家养动物选择信号研究

随着 SNP 芯片和高通量测序成本不断降低,家养动物的选择信号相关研究日益普遍。目前已经对猪(*Sus scrofa*)、鸡(*Gallus gallus*)、牛(*Bos taurus*)、绵羊(*Ovis aries*)、山羊(*Capra hircus*)、犬(*Canis lupus familiaris*)及马(*Equus caballus*)等家养动物的选择信号进行了筛选。



表 1 部分选择信号检测方法  
Table 1 A list of detection methods of selection signals

类别	方法	选择类型	群体	优点	缺点	参考文献
基于群体分化	$F_{ST}$	正向选择、平衡选择	两个群体间	群体间有差异的区域都能检测出。	无法判断受选择方向,种群历史可能影响。	[8]
	LSBL	正向选择、平衡选择	3 个群体以上		种群历史可能影响。	[9]
	$d_i$	正向选择	3 个群体以上		种群历史可能影响。	[10]
基于位点频率谱	Tajima's $D$	平衡选择、正向选择	群体内	简单,对一些人工选择信号特别有效。	无法排除连锁不平衡的效应,可能是搭车效应,无法确定功能位点。	[11]
	$H_p$	正向选择、负选择	群体内			[5]
基于连锁不平衡和单倍型	EHH	正向选择	群体内	对正选择信号特别敏感和准确,对近期发生的选择敏感。	不适用于基因组范围内的选择信号强度的检测。	[13]
	XP-EHH	正向选择	两个群体间		只能筛选正选择信号。	[14]
	$iHS$	正向选择	群体内		很难获得祖先的群体信息。	[15]
新方法	XP-CLR	正向选择	群体间	对等位基因频率偏离更灵敏,检出率高。	计算比较复杂	[17]
	HapFLK	正向选择、温和选择	群体间	受选择位点检出率高,能检出温和选择	应用率不高	[22]

3.1 猪选择信号检测

猪最早驯化于 9000 年前的欧亚地区,普遍认为家猪具有两个起源,欧洲野猪和亚洲野猪<sup>[26,27]</sup>。随后,科学家对这个驯化过程进行了解密,Leif Andersson 团队使用  $ZH_p$  方法对欧洲野猪和家猪进行分析,共检测到家猪 77 个受选择区域<sup>[28]</sup>,其中与体长相关的 3 个基因座表现出强选择信号,这 3 个候选基因分别为 1 号染色体上的 *NR6A1* 基因、4 号染色体上的 *PLAG1* 基因和 8 号染色体上的 *LCORL* 基因,通过关联性分析发现 *PLAG1* 和 *LCORL* 两个基因共同解释了群体中 18.4% 的体长变异,这表明猪在驯化过程中体长受到了选择。四川农业大学李学伟团队使用  $F_{ST}$  方法对藏野猪和家猪进行了选择信号的筛选,发现家猪 *KCNMA1* 和 *TRPC1* 基因受到选择,这可能与家猪大量分泌唾液有关<sup>[29]</sup>。驯化是一个复杂的过程,猪的驯化还伴随着一些与野猪基因交流事件<sup>[27,30]</sup>,这增加了对选择信号筛选的难度。

目前全世界有 300 多个各具特色的猪品种,在一定程度上这些特色都是由自然或人工选择作用而形成的。繁殖性状是猪最重要的性状之一,目前已经发现 *AREG*、*PRLP*、*ESR* 等多个基因在不同繁殖力群体中受到选择<sup>[31-33]</sup>,这些基因均与猪的繁殖性状密切相关,其中 *ESR* 直接影响产仔数,在猪标记辅助选择中被广泛应用<sup>[34]</sup>。肉质也是人类比较关注

的一种经济性性状,研究发现皮特兰、汉普夏与杜洛克猪群体的 *IGF2* 基因上存在选择信号,这与猪肉品质有关<sup>[35]</sup>。随着对猪的育种不断加强,越来越多的基因受到选择,然而检出率却不高,这就需要在育种过程中设计一些独特家系,加快这些主效基因的分离和检出。

3.2 鸡选择信号检测

鸡是最重要的家禽,目前认为家鸡起源于红色原鸡(*Gallus gallus*),在 1 万年前的印度和东南亚地区被驯化<sup>[36]</sup>。2010 年,Leif Andersson 团队对红色原鸡、蛋鸡及肉鸡的 9 个群体进行了全基因组测序(Whole genome sequencing, WGS)<sup>[12]</sup>,应用标准化杂合度( $ZH_p$ )方法筛选鸡驯化过程受选择的基因,在受选择基因过滤时,以已知黄色皮肤基因的  $ZH_p$  值为标准,发现仅剩 3 个区域,其中 *TSHR* 的  $ZH_p$  值最小,功能最为详细。在与红色原鸡对比,发现 *TSHR* 基因区域在驯养的鸡中受到强烈的选择(杂合度很低),对 36 个群体验证也证明了这一点。*TSHR* 在外显子区发生了一个非同义突变,该突变可能与鸡的季节性发情有关,通常野鸡为了节约能耗,只会在特定的季节发情繁衍,而人类为了让鸡常年发情,对该基因进行了强烈的选择。

现代的商品鸡主要分为两类:蛋鸡和肉鸡。其中蛋鸡的典型代表是白来航鸡,研究人员使用 60K

SNP 芯片对白来航母鸡遗传突变进行鉴定<sup>[37]</sup>, 随后应用 EHH 方法检测其中的选择信号, 发现了一系列与产蛋性状、代谢性状与免疫反应相关的基因, 如 *EYA2*、*NCKX1*、*LHX2* 等。Johansson 等<sup>[38]</sup>使用 60K SNP 芯片对经历过 50 个世代的歧化选择培育而成的 2 个体重相差极大的肉鸡品系进行遗传突变检测, 应用  $F_{ST}$  和群体杂合度分析了它们的选择信号, 发现 50 个潜在选择信号区域基本上与生长性状相关的 QTL 或基因重叠。同样使用 60K SNP 芯片对瘦肉系和肥胖系进行选择信号分析<sup>[39]</sup>, 分别在两个品系中发现 51 个和 57 个受选择区域, 其中包括多个与脂肪沉积有关的基因, 如 *PRPSAP1*、*BBS1*、*MAOA*、*MAOB*、*EHBPI* 和 *LRP2BP*。2014 年, Gholami 等<sup>[40]</sup>使用 600K SNP 芯片对 3 个蛋鸡品种和 14 个地方品种进行选择信号分析,  $ZH_p$  选择信号检测确认了之前文章报道的受选择候选基因, 包括 *TSHR*、*PRL*、*PRLHR*、*INSR*、*LEPR*、*IGF1* 和 *NRAMP1*, 同时发现与繁殖、产蛋等有关的候选基因 *NCOA1*、*SREBF2* 和 *RALGAP1*。2015 年, Örjan Carlborg 课题组通过选择信号定位方法鉴定了弗吉尼亚鸡的单性状双向选择系中多个主效基因在选择过程中产生的关键致因突变, 为长期单性状人工选择产生的多基因致因突变定位提供了新方法<sup>[41]</sup>。由于鸡世代间隔较短, 具有天然的优势, 可以快速培育多种家系, 通过高通量测序可以较快捷地筛选获得功能基因。

### 3.3 牛选择信号检测

牛群体比较复杂, 目前普遍认为包括普通牛 (*Bos taurus*)、水牛 (*Bos bubalus*)、牦牛 (*Bos grunniens*) 和瘤牛 (*Bos indicus*) 等不同的种, 普通牛和瘤牛是现代畜牧业最主要的两种牛, 它们均起源于原牛。然而, 原牛早在 1627 年灭绝, 无法直接比对原牛和家牛的基因组。2009 年, 牛基因组图谱研究团队完成了对海福特牛 (*Hereford*) 全基因组的测序<sup>[42]</sup>, 通过比对基因同义突变和非同义突变的比例 ( $Ka/Ks$ ), 发现在多个免疫基因上存在选择信号。除此之外, 该团队还重测了 6 个品种, 并利用这些 SNP 组成 SNP 芯片, 对 19 个品种 497 头牛进行 SNP 检测, 分析发现普通家牛正经历着由远古时代的大群体到当代群体迅速下降的一个过程, 这可能是由于驯化、选择和育

种形成的瓶颈效应导致的<sup>[43]</sup>。

根据牛的用途, 可将牛分为奶用型品种、肉用型品种及肉役兼用型品种。奶牛一直是倍受关注的群体, 其产奶量和奶品质存在高度选育。对奶牛选择信号进行分析, 发现在 6、14、20 号染色体上都存在强烈的选择信号<sup>[44,45]</sup>, 其中 6 号染色体选择信号区域包含 *WHITE*、*ABCG2* 和酪蛋白 3 个分别与产奶量、奶成分、奶蛋白量相关的基因<sup>[45~47]</sup>; 而在 14 号染色体上, 发现乳脂率性状相关的主效基因 *DGAT1* 受到强烈选择<sup>[45,48,49]</sup>; 20 号染色体的选择信号则包含有与产奶性状相关的 *GHR* 和 *PRLR* 基因<sup>[45,50]</sup>。饲料转化率和肌内脂肪含量是肉牛重要的两个指标, 研究发现与牛肉大理石花纹和脂肪含量相关的 14 号染色体区域在肉牛中存在明显的选择信号<sup>[43,48,51]</sup>。而与饲料转化率相关的 *ZRANB3* 和 *R3HDM1* 基因均在肉牛中受到选择<sup>[43,52]</sup>。牛的世代间隔长, 然而其家系最为完善, 甚至拥有相关疾病的详细记录, 因此, 对牛的选择信号筛选出的基因可以直接与一些性状进行关联, 能快速定位一些重要的功能基因。

### 3.4 羊选择信号检测

羊包括山羊和绵羊两个种, 它们可能均在 1 万年前驯化于新月沃土<sup>[53,54]</sup>。山羊可能是由两种野山羊 (*Capra aegagrus* 和 *Capra falconeri*) 驯化而来<sup>[53]</sup>, 而绵羊起源更为复杂, 目前仅知的欧洲绵羊可能由欧洲摩弗罗 (*O. musimon*) 驯化而来, 而亚洲绵羊的起源未知<sup>[55,56]</sup>。2014 年, Jiang 等<sup>[57]</sup>发表了绵羊基因组, 主要研究了基因扩张现象, 其中发现一个新基因 *LCE7A*, 该基因与毛生长相关, 通过分析对比绵羊和牛的同义突变和错义突变率, 发现该基因在绵羊中存在正选择信号。

当前全世界拥有 1400 多种绵羊。Kijas 等<sup>[58]</sup>应用 50K SNP 芯片对全球 75 个绵羊品种的 2819 个个体进行遗传突变检测, 使用  $F_{ST}$  方法筛选选择信号, 发现与羊角相关的 *RXP2* 基因受到强烈选择, 该基因与绵羊羊角性状 (有、无羊角) 密切相关。同时还发现多个基因存在选择信号, 包括: 与色素沉积有关的 *KIT*、*ASIP*、*MITF* 基因; 与骨骼形态和大小相关的 *NPR2*、*HMGA2*、*BMP2* 基因; 与繁殖相关的 *PRLR* 和 *TSHR* 基因。在使用 50K SNP 芯片对 2 个脂尾绵

羊品种和 13 个瘦尾绵羊品种进行选择信号进行分析,发现 *BMP2* 和 *VNRT* 基因区域受到选择<sup>[59]</sup>。在我国,绵羊拥有丰富的遗传多样性,杜立新课题组使用 50K SNP 芯片对我国 10 个绵羊品种选择信号进行分析,发现多个基因区域存在选择信号,包括与羊角相关的 *RXFP2* 基因;与脂肪可能有关的 *PDGFD* 基因;与生长性状可能相关的 *NPR2*、*SPAG8*、*HINT2* 基因;与绵羊产奶相关的 *PKD2*、*SPPI1*、*MEPE*、*IBSP* 等基因。野绵羊一直是探索自然选择的最佳模型,Kijas 课题组对来自美国本土的 3 个地方野生大角绵羊进行了全基因组测序<sup>[60]</sup>,分析杂合度发现公羊比母羊受到的选择更强, $ZH_p$  和  $\pi$  检验均表明 *RXFP2* 受到强烈选择,之前的研究表明 *RXFP2* 与羊角大小有关,作者认为大角绵羊的羊角进化显然与 *RXFP2* 的选择有关。虽然全世界有 1000 多种山羊,但是由于目前山羊规模化养殖程度较低,开展山羊育种工作较困难,对山羊选择信号的研究也较少。羊的研究工作进展整体不如猪、鸡、牛等家养动物,这正表明了羊具有广阔的研究空间。未来应加强羊资源家系的构建,提高选择信号的检出率。

### 3.5 犬选择信号检测

犬是最早被驯化的家养动物,化石证据表明犬的驯化史大概始于 30 000~33 000 年前<sup>[61,62]</sup>,目前普遍认为犬是由灰狼驯化而来。2013 年,科研人员通过对灰狼和犬进行全基因组测序<sup>[63]</sup>探究了这一驯化过程,使用  $F_{ST}$  和  $H_p$  方法共筛选到 36 个受选择区域,包含 122 个基因,其中一半的区域与脑功能和神经发育有关,而脑和神经与动物行为密切相关,与狼相比,犬较为温顺、忠诚、合群,因此这些基因可能与狼和犬性情差异有关;同时还发现多个与脂肪和淀粉消化相关的基因,犬携带了额外拷贝的淀粉酶基因 *AMY2B*,其表达量是狼的 28 倍之多。此外,犬生成的麦芽糖酶-葡糖淀粉酶含量比狼高 12 倍,这是由于其编码基因 *MGAM* 发生了一些突变。藏獒作为犬中之王倍受关注,对藏獒进行重测序发现,藏獒 *EPAS1* 基因存在强烈的选择信号,这与人类适应高原受选择基因一致,表明了高原适应性方面藏獒与人类存在趋同进化的过程<sup>[64,65]</sup>。犬是人类最好的伙伴,在人类的喜好和选择作用下,

形成了各种形态各异的品种,如茶杯大小的茶杯犬,这为选择信号的筛选提供了遗传基础。

### 3.6 马选择信号检测

马的驯化时间大概是 5000 年前<sup>[66,67]</sup>,早期马被用于骑行和获得肉、奶,后来马被广泛用于运输、通讯和战争<sup>[68]</sup>,而近年用于赛跑。伴随这些用途,马被不断选育。2013 年,科研人员使用 50K SNP 芯片对 33 个品种的 744 匹马的遗传突变进行检测<sup>[69]</sup>,随后使用  $F_{ST}$  方法筛选选择信号,发现 *MSTN* 基因受到强烈选择,并且通过关联分析发现启动子区的一个插入突变(SINE insertion)和内含子区的一个 SNP 位点(g.66493737C/T)与肌纤维比例有关,此外还发现可能与步态相关的基因 *DMRT3* 和可能与体型大小相关的基因 *ECA11* 受到选择。目前马的基因组已经公布<sup>[70]</sup>,马具有优良的纯血血统,尤其是赛马,家系完整,这为通过全基因组测序挖掘选择信号提供了良好的遗传基础。

## 4 家养动物选择信号应考虑的关键问题

家养动物在受到选择过程中,在基因组上会表现为突变位点以及与其连锁的位点构成的单倍型被保留,在分化程度上表现为两组显著差异的遗传突变类型,在杂合度上表现为杂合度降低。但这是理想状态,现实中选择信号复杂多变,选择信号受到多个因素的影响。因此分析选择信号时,需要考虑多个关键问题。(1)搭车效应。一些基因由于相邻位点或连锁位点受到选择而产生假阳性信号,为了排除这个效应,在对选择信号分析过程中,通常需要使用基于单倍体连锁不平衡的检验方法来校验受选择的位点。(2)种群历史(Demographic history)。种群历史对整个基因组都有很大的影响,有时表现出的变化与选择信号一致或类似,如种群快速扩繁时会出现大量的低频位点,这与净化选择效果一致;此外,种群经过瓶颈效应后,个体数量比较少,这时产生的遗传漂变可能大于选择作用。所以在处理选择信号时,还需要把该区域与整个 DNA 序列特征进行比较,判断其是选择作用或受到种群历史影响。(3)选择类型。不同的选择类型在基因组上展示不同的现象,正选择表现为基因杂合度不断降低;平衡



选择表现为维持群体的多态性变异及杂合优势；净化选择表现为突变型频率不断降低和低频位点增多，所以很多时候需要使用多种方法筛选选择信号，避免漏掉一些关键基因，同时也需要综合分析这些选择信号，判断选择信号类型，避免假阳性。(4)选择强度。一个基因的进化速度受选择强度影响，人工或自然选择越强，选择经历的时间越短，选择信号就越强，选择痕迹越明显。(5)选择时间。对于自然选择而言，选择发生的时间在很大程度上影响到信号的强弱。如果属于较为古老的选择，即可能随着时间的推移选择痕迹慢慢消失，被新出现的中性突变位点所掩盖；而近期的正选择相对容易发现，不同的时间尺度要用不同的方法检测。

检测选择信号的方法有很多，但是每种方法都有自己的局限性，基于遗传分化的算法无法避免种群历史的干扰，基于单位点的方法无法排除连锁因素的影响，基于连锁的方法只能判断近期的选择信号，这就要求人们应综合利用这些方法。Grossman 等<sup>[71]</sup>构建了一个综合选择方法 CMS (Composite of multiple signals)，综合了  $iHS$ 、 $XP-EHH$ 、 $iHH$ 、 $F_{ST}$  和  $DAF$  5 种选择方法产生的信号，对人类基因组进行筛选，最后获得了真正受选择的致因突变位点。由此可见只有整合多种选择方法，结合更多的遗传模式，提高数据的可靠性，才能筛选获得正确的选择信号，确定功能基因。

## 5 展 望

探索选择信号已经成为筛选功能基因最有效的方法之一。家养动物受选择的基因通常影响着一些重要的经济性状。通过选择信号筛选出的基因，结合单性状关联分析或全基因组关联分析 (Genome-wide association study, GWAS) 以及后续的功能验证实验等方式有望确定影响性状的功能位点，后续可将这些功能位点作为分子标记应用到育种中。

随着高通量测序成本不断降低，全基因组测序得到越来越广泛的应用，选择信号检测是从全基因组测序结果中挖掘功能基因的最佳方法，因此未来这两者的结合将更加密切。同时选择信号检测方法也一直在不断改进，检出率和准确性进一步提高，结合测序产生的海量高深度测序数据，这些方法在

未来基因组结构变异 (Structural variation, SV) 的鉴定方面也将得到应用。

## 参考文献 (References):

- [1] Diamond J. Evolution, consequences and future of plant and animal domestication. *Nature*, 2002, 418(6898): 700–707. [\[DOI\]](#)
- [2] King JL, Jukes TH. Non-darwinian evolution. *Science*, 1969, 164(3881): 788–798. [\[DOI\]](#)
- [3] Sabeti PC, Schaffner SF, Fry B, Lohmueller J, Varilly P, Shamovsky O, Palma A, Mikkelsen TS, Altshuler D, Lander ES. Positive natural selection in the human lineage. *Science*, 2006, 312(5780): 1614–1620. [\[DOI\]](#)
- [4] Smith JM, Haigh J. The hitch-hiking effect of a favourable gene. *Genet Res*, 1974, 23(1): 23–35. [\[DOI\]](#)
- [5] Fay JC, Wu CI. Hitchhiking under positive Darwinian selection. *Genetics*, 2000, 155(3): 1405–1413. [\[DOI\]](#)
- [6] McVean G. The structure of linkage disequilibrium around a selective sweep. *Genetics*, 2007, 175(3): 1395–1406. [\[DOI\]](#)
- [7] Frank SA. Natural selection. II. Developmental variability and evolutionary rate. *J Evol Biol*, 2011, 24(11): 2310–2320. [\[DOI\]](#)
- [8] Lewontin R, Krakauer J. Distribution of gene frequency as a test of the theory of the selective neutrality of polymorphisms. *Genetics*, 1973, 74(1): 175–195. [\[DOI\]](#)
- [9] Shriver MD, Kennedy GC, Parra EJ, Lawson HA, Sonpar V, Huang J, Akey JM, Jones KW. The genomic distribution of population substructure in four populations using 8, 525 autosomal SNPs. *Hum Genomics*, 2004, 1(4): 274–286. [\[DOI\]](#)
- [10] Akey JM, Ruhe AL, Akey DT, Wong AK, Connelly CF, Madeoy J, Nicholas TJ, Neff MW. Tracking footprints of artificial selection in the dog genome. *Proc Natl Acad Sci USA*, 2010, 107(3): 1160–1165. [\[DOI\]](#)
- [11] Tajima F. Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, 1989, 123(3): 585–595. [\[DOI\]](#)
- [12] Rubin CJ, Zody MC, Eriksson J, Meadows JRS, Sherwood E, Webster MT, Jiang L, Ingman M, Sharpe T, Ka S, Hallböök F, Besnier F, Carlborg Ö, Bed'hom B, Tixier-Boichard M, Jensen P, Siegel P, Lindblad-Toh K, Andersson L. Whole-genome resequencing reveals loci under selection during chicken domestication. *Nature*, 2010, 464(7288): 587–591. [\[DOI\]](#)
- [13] Sabeti PC, Reich DE, Higgins JM, Levine HZP, Richter DJ, Schaffner SF, Gabriel SB, Platko JV, Patterson NJ,



- McDonald GJ, Ackerman HC, Campbell SJ, Altshuler D, Cooper R, Kwiatkowski D, Ward R, Lander ES. Detecting recent positive selection in the human genome from haplotype structure. *Nature*, 2002, 419(6909): 832–837. [DOI]
- [14] Sabeti PC, Varilly P, Fry B, Lohmueller J, Hostetter E, Cotsapas C, Xie XH, Byrne EH, McCarroll SA, Gaudet R, Schaffner SF, Lander ES, International HapMap Consortium. Genome-wide detection and characterization of positive selection in human populations. *Nature*, 2007, 449(7164): 913–918. [DOI]
- [15] Voight BF, Kudaravalli S, Wen XQ, Pritchard JK. A map of recent positive selection in the human genome. *PLoS Biol*, 2006, 4(3): e72. [DOI]
- [16] Biswas S, Akey JM. Genomic insights into positive selection. *Trends Genet*, 2006, 22(8): 437–446. [DOI]
- [17] Chen H, Patterson N, Reich D. Population differentiation as a test for selective sweeps. *Genome Res*, 2010, 20(3): 393–402. [DOI]
- [18] Staubach F, Lorenc A, Messer PW, Tang K, Petrov DA, Tautz D. Genome patterns of selection and introgression of haplotypes in natural populations of the house mouse (*Mus musculus*). *PLoS Genet*, 2012, 8(8): e1002891. [DOI]
- [19] Hufford MB, Xu X, van Heerwaarden J, Pyhäjärvi T, Chia JM, Cartwright RA, Elshire RJ, Glaubitz JC, Guill KE, Kaeppler SM, Lai JS, Morrell PL, Shannon LM, Song C, Springer NM, Swanson-Wagner RA, Tiffin P, Wang J, Zhang GY, Doebley J, McMullen MD, Ware D, Buckler ES, Yang S, Ross-Ibarra J. Comparative population genomics of maize domestication and improvement. *Nat Genet*, 2012, 44(7): 808–811. [DOI]
- [20] Qi JJ, Liu X, Shen D, Miao H, Xie BY, Li XX, Zeng P, Wang SH, Shang Y, Gu XF, Du YC, Li Y, Lin T, Yuan JH, Yang XY, Chen JF, Chen HM, Xiong XY, Huang K, Fei ZJ, Mao LY, Tian L, Städler T, Renner SS, Kamoun S, Lucas WJ, Zhang ZH, Huang SW. A genomic variation map provides insights into the genetic basis of cucumber domestication and diversity. *Nat Genet*, 2013, 45(12): 1510–1515. [DOI]
- [21] Zhou ZK, Jiang Y, Wang Z, Gou ZH, Lyu J, Li WY, Yu YJ, Shu LP, Zhao YJ, Ma YM, Fang C, Shen YT, Liu TF, Li CC, Li Q, Wu M, Wang M, Wu YS, Dong Y, Wan WT, Wang X, Ding ZL, Gao YD, Xiang H, Zhu BG, Lee SH, Wang W, Tian ZX. Resequencing 302 wild and cultivated accessions identifies genes related to domestication and improvement in soybean. *Nat Biotechnol*, 2015, 33(4): 408–414. [DOI]
- [22] Fariello MI, Boitard S, Naya H, SanCristobal M, Servin B. Detecting signatures of selection through haplotype differentiation among hierarchically structured populations. *Genetics*, 2013, 193(3): 929–941. [DOI]
- [23] Fariello MI, Servin B, Tosser-Klopp G, Rupp R, Moreno C, International Sheep Genomics Consortium, San Cristobal M, Boitard S. Selection signatures in worldwide sheep populations. *PLoS One*, 2014, 9(8): e103813. [DOI]
- [24] Kijas JW. Haplotype-based analysis of selective sweeps in sheep. *Genome*, 2014, 57(8): 433–437. [DOI]
- [25] Manunza A, Cardoso TF, Noce A, Martínez A, Pons A, Bermejo LA, Landi V, Sánchez A, Jordana J, Delgado JV, Adán S, Capote J, Vidal O, Ugarte E, Arranz JJ, Calvo JH, Casellas J, Amills M. Population structure of eleven Spanish ovine breeds and detection of selective sweeps with BayeScan and hapFLK. *Sci Rep*, 2016, 6: 27296. [DOI]
- [26] Larson G, Dobney K, Albarella U, Fang MY, Matisoo-Smith E, Robins J, Lowden S, Finlayson H, Brand T, Willerslev E, Rowley-Conwy P, Andersson L, Cooper A. Worldwide phylogeography of wild boar reveals multiple centers of pig domestication. *Science*, 2005, 307(5715): 1618–1621. [DOI]
- [27] Groenen MAM, Archibald AL, Uenishi H, Tuggle CK, Takeuchi Y, Rothschild MF, Rogel-Gaillard C, Park C, Milan D, Megens HJ, Li ST, Larkin DM, Kim H, Frantz LA, Caccamo M, Ahn H, Aken BL, Anselmo A, Anthon C, Auvin L, Badaoui B, Beattie CW, Bendixen C, Berman D, Blecha F, Blomberg J, Bolund L, Bosse M, Botti S, Bujie Z, Bystrom M, Capitanu B, Carvalho-Silva D, Chardon P, Chen C, Cheng R, Choi SH, Chow W, Clark RC, Clee C, Crooijmans RPMA, Dawson HD, Dehais P, De Sapio F, Dibbits B, Drou N, Du ZQ, Eversole K, Fadista J, Fairley S, Faraut T, Faulkner GJ, Fowler KE, Fredholm M, Fritz E, Gilbert JG, Giuffra E, Gorodkin J, Griffin DK, Harrow JL, Hayward A, Howe K, Hu ZL, Humphray SJ, Hunt T, Hornshøj H, Jeon JT, Jern P, Jones M, Jurka J, Kanamori H, Kapetanovic R, Kim J, Kim JH, Kim KW, Kim TH, Larson G, Lee K, Lee KT, Leggett R, Lewin HA, Li YR, Liu WS, Loveland JE, Lu Y, Lunney JK, Ma J, Madsen O, Mann K, Matthews L, McLaren S, Morozumi T, Murtaugh MP, Narayan J, Nguyen DT, Ni P, Oh SJ, Onteru S, Panitz F, Park EW, Park HS, Pascal G, Paudel Y, Perez-Enciso M, Ramirez-Gonzalez R, Reecy JM, Rodriguez-Zas S, Rohrer GA, Rund L, Sang YM, Schachtschneider K, Schraiber JG, Schwartz J, Scobie L, Scott C, Searle S, Servin B, Southey

- BR, Sperber G, Stadler P, Sweedler JV, Tafer H, Thomsen B, Wali R, Wang J, Wang J, White S, Xu X, Yerle M, Zhang GJ, Zhang J, Zhang J, Zhao SH, Rogers J, Churcher C, Schook LB. Analyses of pig genomes provide insight into porcine demography and evolution. *Nature*, 2012, 491(7424): 393–398. [DOI]
- [28] Rubin CJ, Megens HJ, Barrio AM, Maqbool K, Sayyab S, Schwochow D, Wang C, Carlborg Ö, Jern P, Jørgensen CB, Archibald AL, Fredholm M, Groenen MAM, Andersson L. Strong signatures of selection in the domestic pig genome. *Proc Natl Acad Sci USA*, 2012, 109(48): 19529–19536. [DOI]
- [29] Li MZ, Tian SL, Jin L, Zhou GY, Li Y, Zhang Y, Wang T, Yeung CKL, Chen L, Ma JD, Zhang JB, Jiang A, Li J, Zhou CW, Liu YK, Sun XQ, Zhao HW, Niu ZX, Lou PE, Xian JL, Shen XY, Liu SQ, Zhang SH, Zhang MW, Zhu L, Shuai SR, Bai L, Tang GQ, Liu HF, Jiang YZ, Mai MM, Xiao J, Wang X, Zhou Q, Wang ZQ, Stothard P, Xue M, Gao XL, Luo ZG, Gu YR, Zhu HM, Hu XX, Zhao YF, Plastow GS, Wang JY, Jiang Z, Li K, Li N, Li XW, Li RQ. Genomic analyses identify distinct patterns of selection in domesticated pigs and Tibetan wild boars. *Nat Genet*, 2013, 45(12): 1431–1438. [DOI]
- [30] Frantz LA, Schraiber JG, Madsen O, Megens HJ, Cagan A, Bosse M, Paudel Y, Crooijmans RPMA, Larson G, Groenen MAM. Evidence of long-term gene flow and selection during domestication from analyses of Eurasian wild and domestic pig genomes. *Nat Genet*, 2015, 47(10): 1141–1148. [DOI]
- [31] Wilkinson S, Lu ZH, Megens HJ, Archibald AL, Haley C, Jackson IJ, Groenen MAM, Crooijmans RPMA, Ogden R, Wiener P. Signatures of diversifying selection in European pig breeds. *PLoS Genet*, 2013, 9(4): e1003453. [DOI]
- [32] Ma YL, Zhang HH, Zhang Q, Ding XD. Identification of selection footprints on the X chromosome in pig. *PLoS One*, 2014, 9(4): e94911. [DOI]
- [33] Bonhomme M, Chevalet C, Servin B, Boitard S, Abdallah J, Blott S, SanCristobal M. Detecting selection in population trees: the Lewontin and Krakauer test extended. *Genetics*, 2010, 186(1): 241–262. [DOI]
- [34] Rothschild M, Jacobson C, Vaske D, Tuggle C, Wang LZ, Short T, Eckardt G, Sasaki S, Vincent A, McLaren D, Southwood O, van der Steen H, Mileham A, Plastow G. The estrogen receptor locus is associated with a major gene influencing litter size in pigs. *Proc Natl Acad Sci USA*, 1996, 93(1): 201–205. [DOI]
- [35] Ojeda A, Huang LS, Ren J, Angiolillo A, Cho IC, Soto H, Lemús-Flores C, Makuza SM, Folch JM, Pérez-Enciso M. Selection in the making: A worldwide survey of haplotypic diversity around a causative mutation in porcine *IGF2*. *Genetics*, 2008, 178(3): 1639–1652. [DOI]
- [36] Tixier-Boichard M, Bed'hom B, Rognon X. Chicken domestication: from archeology to genomics. *C R Biol*, 2011, 334(3): 197–204. [DOI]
- [37] Li DF, Liu WB, Liu JF, Yi GQ, Lian L, Qu LJ, Li JY, Xu GY, Yang N. Whole-genome scan for signatures of recent selection reveals loci associated with important traits in White Leghorn chickens. *Poultry Sci*, 2012, 91(8): 1804–1812. [DOI]
- [38] Johansson AM, Pettersson ME, Siegel PB, Carlborg Ö. Genome-wide effects of long-term divergent selection. *PLoS Genet*, 2010, 6(11): e1001188. [DOI]
- [39] Zhang H, Wang SZ, Wang ZP, Da Y, Wang N, Hu XX, Zhang YD, Wang YX, Leng L, Tang ZQ, Li H. A genome-wide scan of selective sweeps in two broiler chicken lines divergently selected for abdominal fat content. *BMC Genomics*, 2012, 13: 704. [DOI]
- [40] Gholami M, Erbe M, Gärke C, Preisinger R, Weigend A, Weigend S, Simianer H. Population genomic analyses based on 1 million SNPs in commercial egg layers. *PLoS One*, 2014, 9(4): e94509. [DOI]
- [41] Sheng ZY, Pettersson ME, Honaker CF, Siegel PB, Carlborg Ö. Standing genetic variation as a major contributor to adaptation in the Virginia chicken lines selection experiment. *Genome Biol*, 2015, 16: 219. [DOI]
- [42] The Bovine Genome Sequencing and Analysis Consortium, Elsik CG, Tellam RL, Worley KC. The genome sequence of taurine cattle: a window to ruminant biology and evolution. *Science*, 2009, 324(5926): 522–528. [DOI]
- [43] The Bovine HapMap Consortium. Genome-wide survey of SNP variation uncovers the genetic structure of cattle breeds. *Science*, 2009, 324(5926): 528–532. [DOI]
- [44] Khatkar MS, Thomson PC, Tammen I, Raadsma HW. Quantitative trait loci mapping in dairy cattle: review and meta-analysis. *Genet Sel Evol*, 2004, 36(2): 163–190. [DOI]
- [45] Qanbari S, Pimentel ECG, Tetens J, Thaller G, Lichtner P, Sharifi AR, Simianer H. A genome-wide scan for signatures of recent selection in Holstein cattle. *Anim Genet*, 2010, 41(4): 377–389. [DOI]
- [46] Weikard R, Widmann P, Buitkamp J, Emmerling R, Kuehn C. Revisiting the quantitative trait loci for milk production traits on BTA6. *Anim Genet*, 2012, 43(3): 318–323. [DOI]
- [47] Nilsen H, Olsen HG, Hayes B, Sehested E, Svendsen M,

- Nome T, Meuwissen T, Lien S. Casein haplotypes and their association with milk production traits in Norwegian Red cattle. *Genet Sel Evol*, 2009, 41: 24. [DOI]
- [48] Hayes BJ, Chamberlain AJ, Maceachern S, Savin K, McPartlan H, MacLeod I, Sethuraman L, Goddard ME. A genome map of divergent artificial selection between *Bos taurus* dairy cattle and *Bos taurus* beef cattle. *Anim Genet*, 2009, 40(2): 176–184. [DOI]
- [49] Cole JB, VanRaden PM, O'Connell JR, Van Tassell CP, Sonstegard TS, Schnabel RD, Taylor JF, Wiggans GR. Distribution and location of genetic effects for dairy traits. *J Dairy Sci*, 2009, 92(6): 2931–2946. [DOI]
- [50] Viitala S, Szyda J, Blott S, Schulman N, Lidauer M, Mäki-Tanila A, Georges M, Vilkki J. The role of the bovine growth hormone receptor and prolactin receptor genes in milk, fat and protein production in Finnish Ayrshire dairy cattle. *Genetics*, 2006, 173(4): 2151–2164. [DOI]
- [51] Wiener P, Wilkinson S. Deciphering the genetic basis of animal domestication. *Proc Royal Soc B Biol Sci*, 2011, 278(1722): 3161–3170. [DOI]
- [52] Barendse W, Harrison BE, Bunch RJ, Thomas MB, Turner LB. Genome wide signatures of positive selection: The comparison of independent samples and the identification of regions associated to traits. *BMC Genomics*, 2009, 10: 178. [DOI]
- [53] Luikart G, Gielly L, Excoffier L, Vigne JD, Bouvet J, Taberlet P. Multiple maternal origins and weak phylogeographic structure in domestic goats. *Proc Natl Acad Sci USA*, 2001, 98(10): 5927–5932. [DOI]
- [54] Ryder ML. Sheep. In: Mason IL, ed. *Evolution of Domesticated Animals*. London, New York: Longman, 1984: 63–85. [DOI]
- [55] Hiendleder S, Kaupé B, Wassmuth R, Janke A. Molecular analysis of wild and domestic sheep questions current nomenclature and provides evidence for domestication from two different subspecies. *Proc Royal Soc B Biol Sci*, 2002, 269(1494): 893–904. [DOI]
- [56] Meadows JRS, Hiendleder S, Kijas JW. Haplogroup relationships between domestic and wild sheep resolved using a mitogenome panel. *Heredity*, 2011, 106(4): 700–706. [DOI]
- [57] Jiang Y, Xie M, Chen W, Talbot R, Maddox JF, Faraut T, Wu CH, Muzny DM, Li YX, Zhang WG, Stanton JA, Brauning R, Barris WC, Hourlier T, Aken BL, Searle SMJ, Adelson DL, Bian C, Cam GR, Chen YL, Cheng SF, DeSilva U, Dixon K, Dong Y, Fan GY, Franklin IR, Fu SY, Fuentes-Utrilla P, Guan R, Highland MA, Holder ME, Huang GD, Ingham AB, Jhangiani SN, Kalra D, Kovar CL, Lee SL, Liu WQ, Liu X, Lu CX, Lv T, Mathew T, McWilliam S, Menzies M, Pan SK, Robelin D, Servin B, Townley D, Wang WL, Wei B, White SN, Yang XH, Ye C, Yue YJ, Zeng P, Zhou Q, Hansen JB, Kristiansen K, Gibbs RA, Flicek P, Warkup CC, Jones HE, Oddy VH, Nicholas FW, McEwan JC, Kijas JW, Wang J, Worley KC, Archibald AL, Cockett N, Xu X, Wang W, Dalrymple BP. The sheep genome illuminates biology of the rumen and lipid metabolism. *Science*, 2014, 344(6188): 1168–1173. [DOI]
- [58] Kijas JW, Lenstra JA, Hayes B, Boitard S, Neto LRP, San Cristobal M, Servin B, McCulloch R, Whan V, Gietzen K, Paiva S, Barendse W, Ciani E, Raadsma H, McEwan J, Dalrymple B, International Sheep Genomics Consortium Members. Genome-wide analysis of the world's sheep breeds reveals high levels of historic mixture and strong recent selection. *PLoS Biol*, 2012, 10(2): e1001258. [DOI]
- [59] Moioli B, Pilla F, Ciani E. Signatures of selection identify loci associated with fat tail in sheep. *J Anim Sci*, 2015, 93(10): 4660–4669. [DOI]
- [60] Kardos M, Luikart G, Bunch R, Dewey S, Edwards W, McWilliam S, Stephenson J, Allendorf FW, Hogg JT, Kijas J. Whole-genome resequencing uncovers molecular signatures of natural and sexual selection in wild bighorn sheep. *Mol Ecol*, 2015, 24(22): 5616–5632. [DOI]
- [61] Ovodov ND, Crockford SJ, Kuzmin YV, Higham TFG, Hodgins GWL, van der Plicht J. A 33, 000-year-old incipient dog from the Altai Mountains of Siberia: evidence of the earliest domestication disrupted by the Last Glacial Maximum. *PLoS One*, 2011, 6(7): e22821. [DOI]
- [62] Davis SJM, Valla FR. Evidence for domestication of the dog 12, 000 years ago in the Natufian of Israel. *Nature*, 1978, 276(5688): 608–610. [DOI]
- [63] Axelsson E, Ratnakumar A, Arendt ML, Maqbool K, Webster MT, Perloski M, Liberg O, Arnemo JM, Hedhammar Å, Lindblad-Toh K. The genomic signature of dog domestication reveals adaptation to a starch-rich diet. *Nature*, 2013, 495(7441): 360–364. [DOI]
- [64] Gou X, Wang Z, Li N, Qiu F, Xu Z, Yan DW, Yang SL, Jia J, Kong XY, Wei ZH, Lu SX, Lian LS, Wu CX, Wang XY, Li GZ, Ma T, Jiang Q, Zhao X, Yang JQ, Liu BH, Wei DK, Li H, Yang JF, Yan YL, Zhao GY, Dong XX, Li ML, Deng WD, Leng J, Wei CC, Wang C, Mao HM, Zhang H, Ding

- GH, Li YX. Whole-genome sequencing of six dog breeds from continuous altitudes reveals adaptation to high-altitude hypoxia. *Genome Res*, 2014, 24(8): 1308–1315. [DOI]
- [65] Li Y, Wu DD, Boyko AR, Wang GD, Wu SF, Irwin DM, Zhang YP. Population variation revealed high-altitude adaptation of Tibetan Mastiffs. *Mol Biol Evol*, 2014, 31(5): 1200–1205. [DOI]
- [66] Lippold S, Matzke NJ, Reissmann M, Hofreiter M. Whole mitochondrial genome sequencing of domestic horses reveals incorporation of extensive wild horse diversity during domestication. *BMC Evol Biol*, 2011, 11: 328. [DOI]
- [67] Ludwig A, Pruvost M, Reissmann M, Benecke N, Brockmann GA, Castaños P, Cieslak M, Lippold S, Llorente L, Malaspinas AS, Slatkin M, Hofreiter M. Coat color variation at the beginning of horse domestication. *Science*, 2009, 324(5926): 485. [DOI]
- [68] Outram AK, Stear NA, Bendrey R, Olsen S, Kasparov A, Zaibert V, Thorpe N, Evershed RP. The earliest horse harnessing and milking. *Science*, 2009, 323(5919): 1332–1335. [DOI]
- [69] Petersen JL, Mickelson JR, Rendahl AK, Valberg SJ, Andersson LS, Axelsson J, Bailey E, Bannasch D, Binns MM, Borges AS, Brama P, da Câmara Machado A, Capomaccio S, Cappelli K, Cothran EG, Distl O, Fox-Clipsham L, Graves KT, Guérin G, Haase B, Hasegawa T, Hemmann K, Hill EW, Leeb T, Lindgren G, Lohi H, Lopes MS, McGivney BA, Mikko S, Orr N, Penedo MC, Piercy RJ, Raekallio M, Rieder S, Røed KH, Swinburne J, Tozaki T, Vaudin M, Wade CM, McCue ME. Genome-wide analysis reveals selection for important traits in domestic horse breeds. *PLoS Genet*, 2013, 9(1): e1003211. [DOI]
- [70] Wade CM, Giulotto E, Sigurdsson S, Zoli M, Gnerre S, Imsland F, Lear TL, Adelson DL, Bailey E, Bellone RR, Blöcker H, Distl O, Edgar RC, Garber M, Leeb T, Mauceli E, MacLeod JN, Penedo MCT, Raison JM, Sharpe T, Vogel J, Andersson L, Antczak DF, Biagi T, Binns MM, Chowdhary BP, Coleman SJ, Della Valle G, Fryc S, Guérin G, Hasegawa T, Hill EW, Jurka J, Kiialainen A, Lindgren G, Liu J, Magnani E, Mickelson JR, Murray J, Nergadze SG, Onofrio R, Pedroni S, Piras MF, Raudsepp T, Rocchi M, Røed KH, Ryder OA, Searle S, Skow L, Swinburne JE, Syvänen AC, Tozaki T, Valberg SJ, Vaudin M, White JR, Zody MC, Broad Institute Genome Sequencing Platform, Broad Institute Whole Genome Assembly Team, Lander ES, Lindblad-Toh K. Genome sequence, comparative analysis, and population genetics of the domestic horse. *Science*, 2009, 326(5954): 865–867. [DOI]
- [71] Grossman SR, Shylakhter I, Karlsson EK, Byrne EH, Morales S, Frieden G, Hostetter E, Angelino E, Garber M, Zuk O, Lander ES, Schaffner SF, Sabeti PC. A composite of multiple signals distinguishes causal variants in regions of positive selection. *Science*, 2010, 327(5967): 883–886. [DOI]

(责任编辑: 赵要风)