



基于GBLUP和Bayes方法的“敏盖”绒山羊绒毛性状基因组预测准确性研究

张新¹, 边伟龙⁴, 尚瑞芳⁴, 宝勒尔⁴, 习海娇¹, 李金泉^{1,2}, 苏蕊¹, 吕琦¹, 王瑞军¹,
张燕军¹, 王志英^{1,3*}

(1. 内蒙古农业大学动物科学学院, 呼和浩特 010018; 2. 农业农村部肉羊遗传育种重点实验室, 呼和浩特 010018;
3. 内蒙古自治区羊遗传育种与繁殖重点实验室, 呼和浩特 010018;
4. 鄂尔多斯市云东依农种养殖有限责任公司, 鄂尔多斯 017200)

摘要: 旨在估计“敏盖”绒山羊绒毛性状的遗传参数和基因组育种值, 比较GBLUP和Bayes方法对“敏盖”绒山羊绒毛性状基因组育种值预测准确性的影响, 为优化“敏盖”绒山羊遗传选育策略提供参考依据。本研究以847只“敏盖”绒山羊的绒毛性状为研究对象, 利用GBLUP、Bayes A、Bayes B、Bayes C和Bayes LASSO五种方法进行绒毛性状(毛长、绒厚、绒细、产绒量)遗传参数及基因组育种值的估计, 进一步利用五倍交叉验证方法评价基因组育种值估计准确性。结果显示, 5种方法估计的毛长遗传力为0.11~0.32, 绒厚遗传力为0.14~0.33, 绒细遗传力为0.17~0.49, 产绒量遗传力为0.16~0.32。绒厚和毛长之间存在正的强遗传相关, 相关系数为0.48~0.84, 其它性状间的遗传相关相对较小, 为-0.06~0.46; GBLUP、Bayes A、Bayes B、Bayes C和Bayes LASSO方法估计的毛长、绒厚、绒细和产绒量的基因组育种值准确性分别为0.623 4~0.950 9、0.626 2~0.953 0、0.830 8~0.953 1、0.741 8~0.961 4。Bayes方法估计的各性状的基因组预测准确性均显著高于GBLUP方法, Bayes B方法的基因组预测准确性最高, 达0.950 9~0.961 4。综上所述, Bayes B方法被推荐用于“敏盖”绒山羊基因组选育, 可能会加快种群的遗传进展。

关键词: “敏盖”绒山羊; 绒毛性状; 遗传参数; 基因组选择

中图分类号: S827.2

文献标志码: A

文章编号: 0366-6964(2026)02-0704-12

Study on the Genomic Prediction Accuracy of Cashmere Traits in “Min Gai” Cashmere Goats Using GBLUP and Bayes Methods

ZHANG Xin¹, BIAN Weilong⁴, SHANG Ruifang⁴, BAO Leer⁴, XI Haijiao¹, LI Jinquan^{1,2}, SU Rui¹,
LÜ Qi¹, WANG Ruijun¹, ZHANG Yanjun¹, WANG Zhiying^{1,3*}

(1. College of Animal Science, Inner Mongolia Agricultural University, Hohhot 010018, China;
2. Key Laboratory of Meat Sheep Genetic, Breeding of Ministry of Agriculture and Rural Affairs, Hohhot 010018, China; 3. Key Laboratory of Sheep Genetic, Breeding and Reproduction in Inner Mongolia Autonomous Region, Hohhot 010018, China; 4. Ordos Yundong Yinong Breeding and Farming Co., Ltd., Ordos 017200, China)

收稿日期: 2025-07-29

基金项目: 2030科技创新-重大项目(2023ZD0405102); 羊遗传资源保护与创新利用(BR251201); 内蒙古自治区高等学校创新团队发展计划支持(NMGIRT2322); 财政部和农业农村部: 国家绒毛细羊产业技术体系(CARS-39)

作者简介: 张新, 硕士生, 主要从事羊遗传资源保护与育种方面的研究, E-mail: 1534185726@qq.com

* 通信作者: 王志英, 副教授, 主要从事羊遗传资源保护与育种方面的研究, E-mail: wzhy0321@126.com

Abstract: This study aimed to estimate the genetic parameters and genomic breeding values of “Min Gai” cashmere goat hair traits, compare the effects of GBLUP and Bayes methods on the accuracy of genomic breeding value prediction of hair traits of “Min Gai” cashmere goat, and provide a reference basis for optimising the genetic selection and breeding strategy of “Min Gai” cashmere goats. This study took the fleece traits of 847 “Min Gai” cashmere goats as the research object and used GBLUP, Bayes A, Bayes B, Bayes C, and Bayes LASSO to estimate the genetic parameters and genomic breeding values of fleece traits (fleece length, cashmere length, cashmere diameter, and cashmere yield). Furthermore, the accuracy of the genomic breeding value estimates was evaluated using five-fold cross-validation. The results showed that the heritability of fleece length estimated by the 5 methods was 0.11-0.32, the heritability of cashmere length was 0.14-0.33, the heritability of cashmere diameter was 0.17-0.49, and the heritability of cashmere yield was 0.16-0.32. There was a strong positive genetic correlation between cashmere length and fleece length, with a correlation coefficient of 0.48-0.84. The genetic correlation between other traits was relatively small, ranging from -0.06 to 0.46. The genomic breeding value accuracy estimated by the GBLUP, Bayes A, Bayes B, Bayes C, and Bayes LASSO methods for fleece length, cashmere length, cashmere diameter, and cashmere yield were 0.623 4-0.950 9, 0.626 2-0.953 0, 0.830 8-0.953 1, and 0.741 8-0.961 4, respectively. The genomic prediction accuracy of each trait estimated by the Bayes method was significantly higher than that estimated by the GBLUP method, and the Bayes B method had the highest genomic prediction accuracy, reaching 0.950 9-0.961 4. In summary, the Bayes B method is recommended for the genomic selection of “Min Gai” cashmere goats, which may accelerate the genetic progress of the population.

Keywords: “Min Gai” cashmere goat; cashmere traits; genetic parameter; genomic selection

* **Corresponding author:** WANG Zhiying, E-mail: wzhy0321@126.com

在中国,绒山羊不仅是一种重要的遗传资源,还承载着人们吉祥美好的文化寓意。“敏盖”绒山羊是为了改良内蒙古阿尔巴斯白绒山羊的产绒量、繁殖率等性状,同时响应黄河流域生态保护和绿色发展的精神,以内蒙古阿尔巴斯白绒山羊为母本,辽宁绒山羊为父本,经过级进杂交产生符合理想型的杂种后代,进一步通过横交固定、严格选育,形成的适合舍饲饲养模式,生产性能高,羊绒品质优,遗传性能稳定的新品种,目前尚未被命名^[1]。该品种的培育目标是在保持理想绒毛品质的前提下,进一步提高产绒量、繁殖率,以培育出新的高产质优的绒山羊种群。据统计,目前该种群的成年公母羊平均产绒量分别达1 200 g和1 000 g,平均屠宰率达48%以上,经产母羊的产羔率达170%,耐粗饲、抗逆性强,常年发情,适合一年两胎或两年三胎。

基因组选择技术(genomic selection, GS)自2001年由Meuwissen等^[2]首次提出,已成为畜牧遗传育种领域的重要工具。基因组育种值估计是基

因组选择的核心内容,依据所采用的统计模型,其计算方法主要分为直接法和间接法^[3-4]。直接法把个体视为随机效应,基于混合模型方程组求解基因组关系矩阵,进而获得待预测个体的基因组育种值,如GBLUP法^[5]、SSGBLUP法^[6]等;间接法利用参考群体估计标记效应值,在预测群体中通过累加标记效应值,进而获得个体的基因组育种值,常用的模型为贝叶斯类模型^[7],包括Bayes A法^[2]、Bayes B法^[2]、Bayes C法^[8]、Bayes D π 法^[9]、Bayes LASSO法^[10]等。GBLUP与Bayes方法估计结果的差异主要源于对标记效应分布的基本假设不同。GBLUP基于“无穷小模型”假设,即所有遗传标记对性状的效应均服从相同的正态分布,这使其在解析由大量微效基因控制的性状时表现较好。然而,对于受少数中大型效应基因影响的性状,该假设可能过于简化。与之相比,Bayes A、Bayes B等Bayes方法允许标记效应服从更具弹性的先验分布(如t分布或包含点质量混合分布),从而能更灵活地捕

捉不同大小的遗传效应。因此,当性状的遗传架构中存在效应较大的QTL时,Bayes方法通常能提供更准确的遗传参数估计和更高的预测准确性;反之,若性状由极为复杂的多基因控制,则GBLUP可能因其稳定性而更具优势。目前,GBLUP法和Bayes系列方法已在牛^[11]、猪^[12]、鸡^[13]等物种中被广泛应用。Kim等^[14]利用GBLUP和PBLUP法对Hanwoo牛的胴体重、眼肌面积、背膘厚度等性状开展基因组选择研究,发现GBLUP法的准确性为0.766~0.791,PBLUP的准确性为0.480~0.486,表明GBLUP法优于PBLUP法。Peters等^[15]使用BLUP、GBLUP和Bayes B方法对荷斯坦奶牛的产奶性状进行了基因组育种值估计准确性评价,结果发现在Bayes B方法下产乳量、乳脂率、乳蛋白率的基因组育种值估计准确性分别为0.899、0.915和0.897,其明显优于GBLUP法估计的准确性(0.667、0.715和0.635)。Liu等^[16]在三黄鸡的生长和屠宰性状的基因组选择研究中发现,GBLUP、Bayes LASSO和Bayes Mix4法用于以上性状的基因组育种值预测,准确性无显著差异,从数值上讲,Bayes Mix4法预测的基因组育种值准确性最高。Wolc等^[17]基于GBLUP和Bayes B方法对White Leghorn品系蛋壳质量性状进行基因组育种值预测,结果发现Bayes B(0.39~0.85)的基因组预测准确性高于GBLUP(0.28~0.67)法。综上所述,基因组选择的准确性具有群体、性状和方法等特异性,因此,有必要对不同群体开展系统的基因组选育研究,确定用于该种群的最优基因组选择方案,提升基因组预测的准确性,加快种群的世代遗传进展。

在基因组选择及遗传参数估计研究中,交叉验证是评估模型预测准确性与结果可靠性的核心方法,其中五倍交叉验证因兼顾计算效率与验证精度,已成为畜禽重要经济性性状遗传研究的常用手段。从研究进展来看,五倍交叉验证通过将全基因组样本随机划分为5个等规模子集,依次将其中4个子集作为训练集构建遗传参数估计模型,剩余1个子集作为测试集验证模型预测效果,经5次循环后取平均预测误差或相关系数作为最终评估指标,该方法已在牛、羊等畜禽的产奶量、胴体重等遗传研究中广泛应用^[14-15]。

本试验以“敏盖”绒山羊的绒毛性状为研究对象,利用GBLUP和Bayes系列方法对各性状进行方

差组分和遗传参数的估计,进一步计算性状之间的遗传相关,确定性状间是否可以间接选择。最后利用五倍交叉验证方法对育种值准确性进行评价,选择用于“敏盖”绒山羊基因组选择的最优方法。

1 材料与方法

1.1 试验动物及表型数据来源

本研究以内蒙古云东集团敏盖绒山羊原种繁育中心的“敏盖”绒山羊为研究对象,收集该群体2020—2024年共847个个体(包括139只公羊和708只母羊,年龄1~5岁)的7713条绒毛性状记录,所有个体均在标准化羊场环境下进行统一舍饲饲养管理。所研究性状包括毛长(fleece length)、绒厚(cashmere length)、绒细(cashmere diameter)、产绒量(cashmere yield)共4个性状。

表型数据的基本统计采用Microsoft Excel,对原始数据进行预处理,剔除“平均值±2.58倍标准差”范围外的数据用于后续分析,并利用R软件(V4.3.1)对各性状的数据进行正态分布检验。

1.2 基因型数据

本研究所用基因型数据是基于课题组研发的Illumina测序平台绒山羊10K育种芯片测序所得,使用PLINKv1.90软件对个体的基因型数据进行质量控制,质量控制的标准如下:1)删除染色体位置信息缺失的SNP标记位点;2)删除样本SNP位点缺失率大于10%的个体;3)删除样本个体缺失率高于10%的个体;4)删除最小等位基因频率小于5%的SNP位点;5)删除不符合哈代-温伯格平衡定律或 $P<10^{-5}$ 的SNP位点,最后保留9443个高质量SNPs用于后续分析。

1.3 基因组预测的统计模型

1.3.1 GBLUP模型 GBLUP法是VanRaden^[18]2008年在传统BLUP的基础上提出的,用基因组标记实测的亲缘关系 G 矩阵替代系谱推算的理论亲缘关系 A 矩阵,GBLUP法的模型为:

$$y = Xb + Zu + e$$

式中, y 为表型向量; b 为固定效应向量; u 为随机效应向量,即估计育种值,服从 $u \sim N(0, G\sigma_a^2)$ 的正态分布, σ_a^2 为加性遗传方差, e 为残差效应,服从 $e \sim N(0, I\sigma_e^2)$ 的正态分布, I 为单位矩阵, σ_e^2 为残差方差; X 和 Z 分别是固定效应和随机效应的关联矩阵; G 为基

因组关系矩阵:

$$G = \frac{ZZ'}{2 \sum_{i=1}^m P_i(1-p_i)}$$

其中, Z 为 $m \times n$ 的基因型矩阵(1和-1表示两种纯合子基因型,0表示杂合子基因型), m 为SNP个数, n 为个体数; p_i 为第 i 个个体的等位基因频率。

1.3.2 Bayes模型 Meuwissen等^[2]在2001年将Bayes方法运用到全基因选择当中,他提出的Bayes A和Bayes B方法解决了SNP标记数量远大于表型记录的问题。在Bayes A和Bayes B的基础上,研究人员又陆续提出了基于Bayes模型的新方法^[12-13,19],Bayes法的模型如下:

$$y = Xb + \sum_{i=1}^m Z_i g_i + e$$

式中, y 是性状的表型值向量; X 是 b 的关联矩阵; b 是固定效应向量; m 是标记的数量; Z_i 是第 i 个位点的基因型(0/1/2); g_i 是第 i 个位点的效应值; e 是随机残差效应向量。不同Bayes模型之间对标记效应分布的假设各不相同,具体假设如下:Bayes A:假设每个SNP位点都有效应,且效应方差服从相同的先验分布。SNP对表型的影响服从正态分布,意味只有少数标记具有大效应,而大部分标记只具有微小的效应。Bayes B:在Bayes A模型的基础上,引入概率常量 π ,并假定 π 个SNP标记没有效应,假设只有少数SNP有效应(比例为 $1-\pi$),若 $\pi=1$ 时,则所有SNP都具有效应。Bayes C:针对前两者做出了修正,将 π 作为未知参数,把 $1-\pi$ 个有效应SNP标记的方差 $\sigma_{g_i}^2$ 设置为常量,并服从正态分布。Bayesian LASSO:假定绝大部分的SNP效应接近0,服从双指数分布,与Bayes A相比允许标记效应极大值和极小值以更大的概率出现。

1.4 绒毛性状遗传相关及表型相关的估计

对于绒毛性状的育种值和表型值,利用R语言的COR包计算性状间的遗传相关和表型相关,计算公式如下^[20-21]:

$$\text{遗传相关: } r_g = \frac{\text{cov}(g_1, g_2)}{\sqrt{\sigma_{g_1}^2 \sigma_{g_2}^2}}$$

$$\text{表型相关: } r_p = \frac{\text{cov}(p_1, p_2)}{\sqrt{\sigma_{p_1}^2 \sigma_{p_2}^2}}$$

式中, $\text{cov}(g_1, g_2)$ 表示两个性状遗传协方差, $\sigma_{g_i}^2$

与 $\sigma_{g_2}^2$ 表示两个性状的遗传方差; $\text{cov}(p_1, p_2)$ 表示两个性状 p_1 、 p_2 之间的表型协方差, $\sigma_{p_1}^2$ 与 $\sigma_{p_2}^2$ 表示两个性状的表型方差。

1.5 育种值估计准确性评价

本研究采用五倍交叉验证的方法进行基因组育种值预测准确性的评价^[22],将研究群体分成5组,其中1组仅有基因型数据,被作为验证群体,其他4组的个体同时具有表型和基因型数据,作为参考群体,进行5次循环交叉验证。

$$r = \frac{\text{cov}(a, p)}{\sqrt{h^2}}$$

其中, $\text{cov}(a, p)$ 是性状估计育种值和表型值的协方差, h^2 是性状的估计遗传力。

对于每一种方法获得的5个基因组预测准确性值,利用R语言的stats包进行方差分析和多重比较,确定用于“敏盖”绒山羊绒毛性状的最佳基因组预测方法。

2 结 果

2.1 “敏盖”绒山羊绒毛性状描述性统计

“敏盖”绒山羊绒毛性状表型数据的基本描述性统计结果见表1,毛长、绒厚、绒细、产绒量的平均值分别为18.04 cm、8.82 cm、16.07 μm 和1 910 g,各性状的变异系数范围为8.34%~18.64%,说明经过质量控制后各性状个体之间的表型值差异相对较小。对选择的表型值进行正态性检验,由图1可知,各性状均服从正态分布。

2.2 基因型数据统计

将847个“敏盖”绒山羊个体通过绒山羊10K SNP芯片进行基因分型后,有10 000个原始SNPs,剔除不合格的557个SNPs,最终保留9 443个SNPs进行后续分析。质控前后各染色体上的SNP数目变化由图2显示,5号染色体被过滤掉的数目最多,25号染色体被过滤掉的数目最小。图3为质控后各染色体的SNP密度分布图,可以看出SNP均匀分布在“敏盖”绒山羊的各条常染色体上。

2.3 “敏盖”绒山羊绒毛性状方差组分和遗传参数的估计

利用GBLUP、Bayes A、Bayes B、Bayes C和Bayesian LASSO方法对“敏盖”绒山羊毛长进行方差组分和遗传参数估计,结果见表2。5种方法估计的加性遗传方差为0.27~0.81,遗传力为0.11~

表 1 “敏盖”绒山羊绒毛性状基本描述性统计

Table 1 Basic descriptive statistics of fleece traits in “Min Gai” cashmere goats

性状 Trait	个体数 Sample	最大值 Maximum	最小值 Minimum	平均值 Mean	标准差 Standard deviation	变异系数/% Variation coefficient
毛长/cm Fleece length	847	22.00	14.00	18.04	1.72	9.54
绒厚/cm Cashmere length	847	12.00	6.00	8.82	1.31	14.83
绒细/ μm Cashmere diameter	847	19.67	12.42	16.07	1.34	8.34
产绒量/g Cashmere yield	847	2 750	750	1 910	355.93	18.64

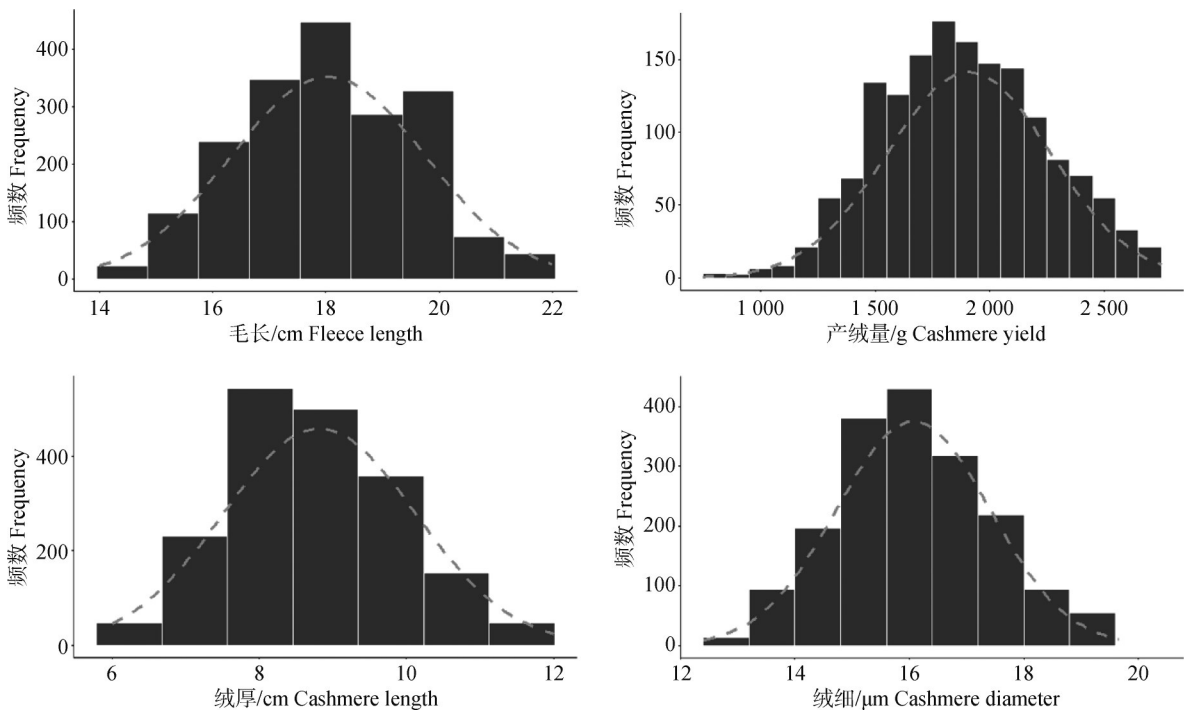


图 1 绒毛性状频率直方图

Fig. 1 Histogram of the frequency of cashmere traits

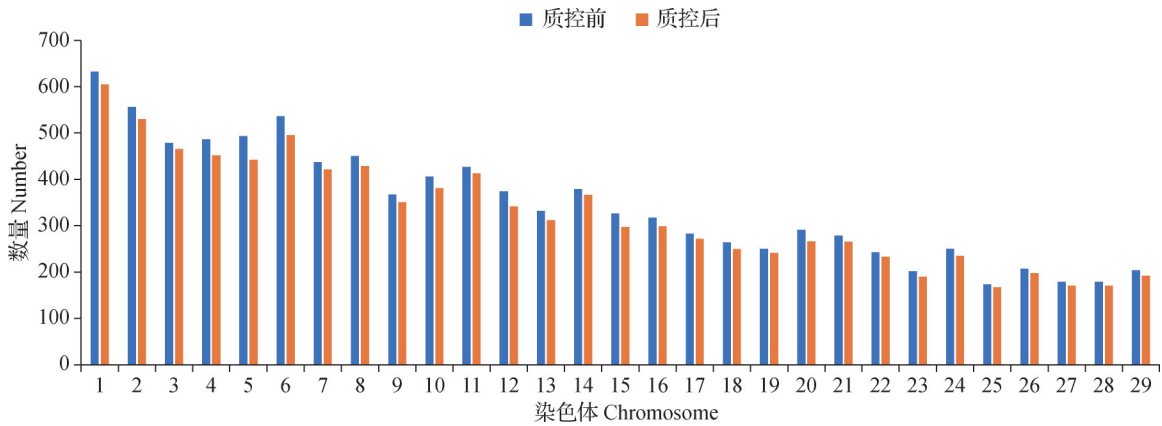


图 2 质控前后各染色体 SNPs 位点数目分布图

Fig. 2 The distribution of the SNPs number on each chromosome before and after quality control

0.32。其中 Bayes A 估计的毛长的加性遗传方差最高,对应的遗传力也最高,达 0.32,为高遗传力;其

他方法估计得遗传力均较小,为中等或者中等偏低遗传力。

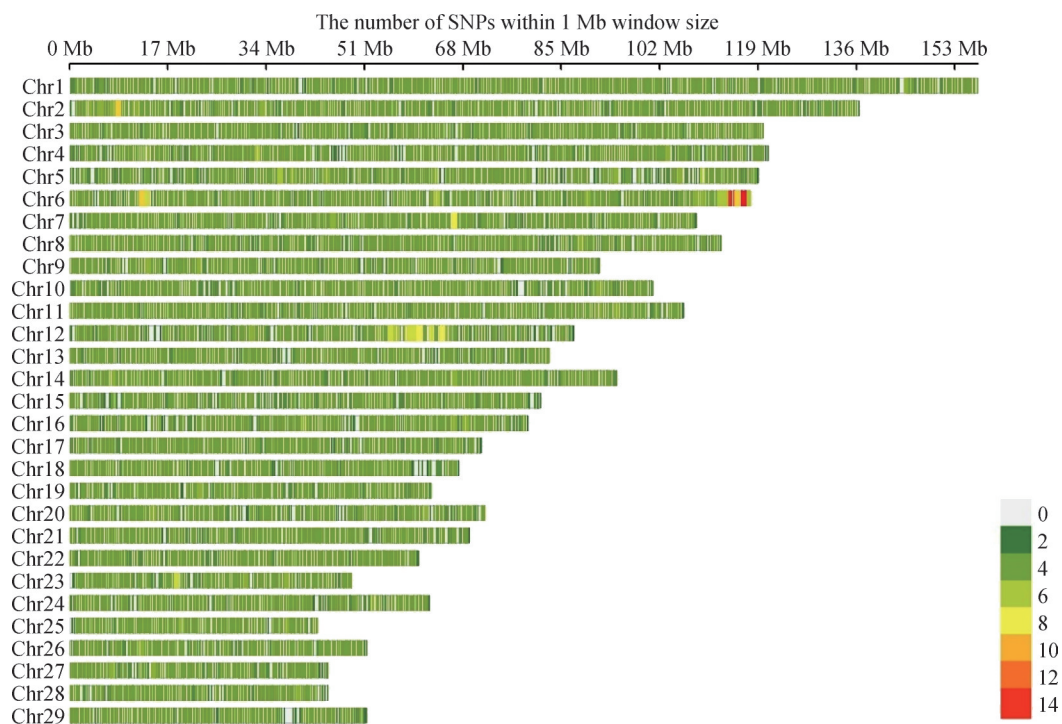


图3 质控后各染色体的SNPs密度分布图

Fig. 3 The distribution of the SNPs density on each chromosome after quality control

表2 毛长的方差组分和遗传参数的估计

Table 2 Estimation of variance components and genetic parameters for fleece length

方法 Method	σ_a^2	σ_e^2	σ_p^2	h_a^2
GBLUP	0.45	2.88	3.33	0.13
Bayes A	0.81	1.70	2.51	0.32
Bayes B	0.27	2.24	2.51	0.11
Bayes C	0.27	2.24	2.51	0.11
Bayesian LASSO	0.51	2.00	2.51	0.20

σ_a^2 为加性遗传方差; σ_e^2 为残差方差; σ_p^2 为表型方差; h_a^2 为遗传力。下同

σ_a^2 is additive genetic variance; σ_e^2 is residual variance; σ_p^2 is phenotypic variance; h_a^2 is heritability. The same as below

利用GBLUP、Bayes A、Bayes B、Bayes C和Bayesian LASSO方法对“敏盖”绒山羊产绒量进行方差组分和遗传参数估计,结果见表3。5种方法估计的加性遗传方差为0.08~0.19,遗传力为0.16~

0.32。其中GBLUP估计的产绒量的加性遗传方差最高,对应的遗传力也最高,达0.32,为高遗传力;其他方法估计得遗传力均较小,为中等或者中等偏低遗传力。

表3 产绒量方差组分和遗传参数的估计

Table 3 Estimation of variance components and genetic parameters for cashmere yield

方法 Method	σ_a^2	σ_e^2	σ_p^2	h_a^2
GBLUP	0.19	0.41	0.60	0.32
Bayes A	0.13	0.36	0.49	0.19
Bayes B	0.09	0.40	0.49	0.17
Bayes C	0.08	0.41	0.49	0.16
Bayesian LASSO	0.11	0.38	0.49	0.23

利用 GBLUP、Bayes A、Bayes B、Bayes C 和 Bayesian LASSO 方法对“敏盖”绒山羊绒厚进行方差组分和遗传参数估计,结果见表4。5种方法估计的加性遗传方差为 0.23~0.43,遗传力为 0.14~

0.33。其中 Bayes A 方法估计的产绒量的加性遗传方差最高,因此其计算得到的遗传力也最高(0.33),为高遗传力。其他方法估计得遗传力均较小,为中等或者中等偏低遗传力。

表4 绒厚方差组分和遗传参数的估计

Table 4 Estimation of variance components and genetic parameters for cashmere length

方法 Method	σ_a^2	σ_e^2	σ_p^2	h_a^2
GBLUP	0.25	1.58	1.83	0.14
Bayes A	0.43	0.89	1.32	0.33
Bayes B	0.26	1.06	1.32	0.19
Bayes C	0.23	1.09	1.32	0.17
Bayesian LASSO	0.32	1.00	1.32	0.24

利用 GBLUP、Bayes A、Bayes B、Bayes C 和 Bayesian LASSO 方法对“敏盖”绒山羊绒细进行方差组分和遗传参数估计,结果见表5。5种方法估计的加性遗传方差为 0.27~1.13,遗传力为

0.17~0.49。其中 GBLUP 估计的毛长的加性遗传方差最高,对应的遗传力也最高,达 0.49,为高遗传力。其他方法估计得遗传力相对较小,为 0.17~0.35。

表5 绒细方差组分和遗传参数的估计

Table 5 Estimation of variance components and genetic parameters for cashmere diameter

方法 Method	σ_a^2	σ_e^2	σ_p^2	h_a^2
GBLUP	1.13	1.16	2.29	0.49
Bayes A	0.55	1.04	1.59	0.35
Bayes B	0.27	1.32	1.59	0.17
Bayes C	0.32	1.27	1.59	0.20
Bayesian LASSO	0.45	1.14	1.59	0.29

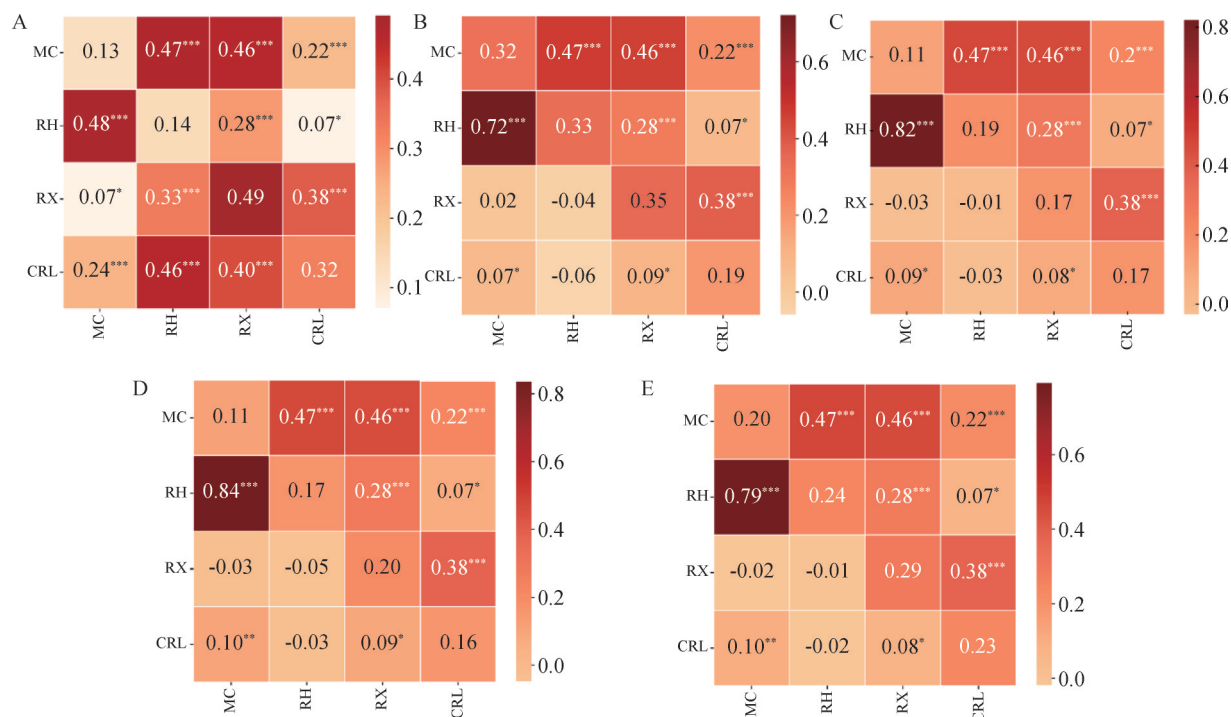
2.4 遗传相关及表型相关估计

利用 GBLUP、Bayes A、Bayes B、Bayes C 和 Bayesian LASSO 方法对“敏盖”绒山羊毛长、绒厚、绒细、产绒量进行遗传相关及表型相关的估计,结果如图4显示,GBLUP方法下绒毛性状间的遗传相关为 0.07~0.48,毛长和绒厚、绒厚和产绒量、绒细和产绒量均存在强的遗传正相关,相关系数分别为 0.48、0.46、0.40。Bayes A 方法下绒毛性状间的遗传相关为 -0.06~0.72,其中绒厚和毛长之间存在强的遗传正相关,相关系数为 0.72。绒细和绒厚、产绒量和绒厚之间遗传相关较低(-0.04、-0.06)。Bayes B 方法下绒毛性状间的遗传相关为 -0.03~0.82,绒厚与毛长之间存在正的强遗传相关,相关系数为 0.82。绒细和毛长、绒细和绒厚、产绒量和绒厚之间遗传相关较低(-0.03、-0.01、-0.03)。Bayes C 方法下绒毛性状间遗传相关为 -0.05~0.84,绒厚与毛长之间存在正的强遗传相

关,相关系数为 0.84。绒细和毛长、绒细和绒厚、产绒量和绒厚之间遗传相关较低(-0.03、-0.05、-0.03)。Bayes LASSO 方法下绒毛性状间遗传相关为 -0.02~0.79,绒厚与毛长之间存在正的强遗传相关,相关系数为 0.79。绒细和毛长、绒细和绒厚、产绒量和绒厚之间遗传相关均较低(-0.02、-0.01、-0.02)。各绒毛性状间的表型相关为 0.07~0.47,其中毛长和绒厚之间表型相关最高(0.47),其次为毛长和绒细(0.46)。在 GBLUP 和 Bayes 方法下均发现绒厚和毛长均成遗传正相关,其余性状相关性较低。

2.5 “敏盖”绒山羊绒毛性状基因组预测准确性比较

基于 GBLUP、Bayes A、Bayes B、Bayes C、Bayes LASSO 等 5 种方法对“敏盖”绒山羊绒毛性状进行基因组育种值估计,其准确性结果见表6和图5。结果显示,5种方法估计的毛长、绒厚、绒细和



A. GBLUP法的相关性热图; B. Bayes A法的相关性热图; C. Bayes B法的相关性热图; D. Bayes C法的相关性热图; E. Bayes LASSO法的相关性热图。上三角数据表示性状间的表型相关; 下三角数据表示性状间的遗传相关; 对角线表示遗传力。MC. 毛长; RH. 绒厚; RX. 绒细; CRL. 产绒量。数字上标*代表显著性水平: ***. $P < 0.001$ 代表极其显著, ** $0.001 < P < 0.01$ 代表极显著, * $0.01 < P < 0.05$ 代表显著

A. The correlation heat map of GBLUP method; B. The correlation heat map of Bayes A method; C. The correlation heat map of Bayes B method; D. The correlation heat map of Bayes C method; E. The correlation heat map of Bayes LASSO method. Upper triangular data indicate phenotypic correlations among traits; lower triangular data indicate genetic correlations among traits; diagonal lines indicate heritability. MC. Fleece length; RH. Cashmere length; RX. Cashmere diameter; CRL. Cashmere yield. Superscripts indicate significance levels: *** denotes $P < 0.001$ (highly significant), ** denotes $0.001 < P < 0.01$ (extremely significant), * denotes $0.01 < P < 0.05$ (significant)

图4 基于不同方法的“敏盖”绒山羊绒毛性状间的相关性热图

Fig. 4 Heat map of correlation between fleece traits of “Min Gai” cashmere goats based on different methodologies

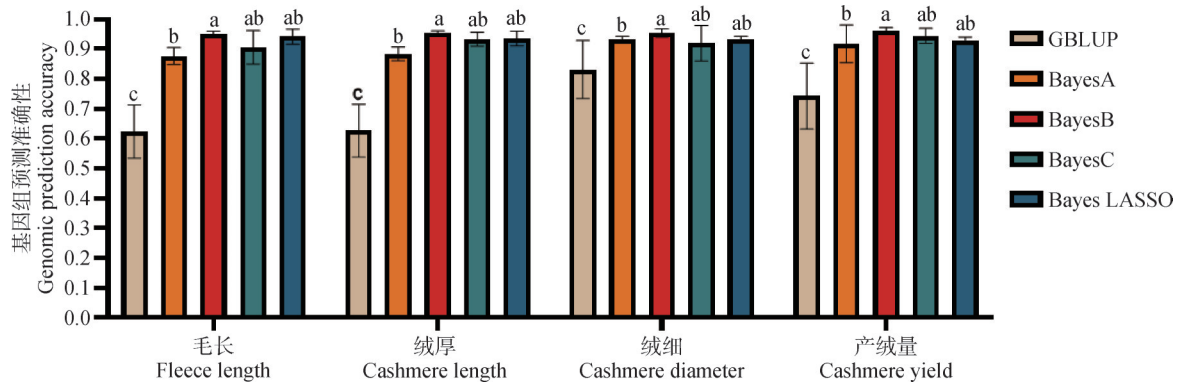
产绒量的基因组育种值准确性分别为 0.623 4~0.950 9; 0.626 2~0.953 0; 0.830 8~0.953 1; 0.741 8~0.961 4。对于每一个绒毛性状而言, Bayes B、Bayes C、Bayes LASSO 三种方法预测的 GEBV 准确性无显著差异, 其准确性显著高于 GBLUP 和 Bayes A 方法。从数值上讲, 相较于

GBLUP法, Bayes系列方法预测的毛长准确性提高了 40.57%~52.53%; 绒厚准确性提高了 41.06%~52.19%; 绒细准确性提高了 10.65%~14.72%; 产绒量提升了 23.60%~29.60%。综上所述, Bayes B方法用于“敏盖”绒山羊绒毛性状的基因组预测准确性最高, 达 0.950 9~0.961 4。

表6 不同方法下绒毛性状基因组预测的准确性

Table 6 Accuracy of genomic prediction of fleece traits under different methods

性状 Trait	Bayes A	Bayes B	Bayes C	Bayes LASSO	GBLUP
毛长 Fleece length	0.876 3±0.03 ^b	0.950 9±0.01 ^a	0.905 7±0.05 ^{ab}	0.940 9±0.02 ^{ab}	0.623 4±0.08 ^c
绒厚 Cashmere length	0.883 3±0.02 ^b	0.953 0±0.01 ^a	0.932 9±0.02 ^{ab}	0.935 3±0.02 ^{ab}	0.626 2±0.08 ^c
绒细 Cashmere diameter	0.930 3±0.01 ^b	0.953 1±0.01 ^a	0.919 3±0.05 ^{ab}	0.930 3±0.01 ^{ab}	0.830 8±0.09 ^c
产绒量 Cashmere yield	0.916 9±0.06 ^b	0.961 4±0.01 ^a	0.943 7±0.02 ^{ab}	0.927 9±0.01 ^{ab}	0.741 8±0.09 ^c



图表中字母不同表示组间差异显著 ($P < 0.05$), 字母相同表示组间差异不显著 ($P > 0.05$)

Different letters in the chart indicate significant differences between groups ($P < 0.05$), and the same letter indicates non-significant differences between groups ($P > 0.05$)

图5 不同方法估计的“敏盖”绒山羊绒毛性状基因组育种值准确性多重比较

Fig. 5 Multiple comparisons of accuracy of genomic breeding values for fleece traits in “Min Gai” cashmere goats with different methods

3 讨论

本研究采用GBLUP和Bayes系列方法(Bayes A、Bayes B、Bayes C、Bayes LASSO)对“敏盖”绒山羊绒毛性状进行遗传参数估计,结果发现毛长遗传力为0.11~0.32,绒厚遗传力为0.14~0.33,产绒量遗传力为0.17~0.32,绒细遗传力为0.17~0.49,不同方法估计的结果存在显著差异,Bayes A方法预测的毛长和绒厚的遗传力最高,GBLUP法预测的产绒量和绒细的遗传力最高。习海娇等^[23]基于加性-显性模型利用BLUP和GBLUP法对内蒙古绒山羊(阿尔巴斯型)进行遗传参数估计,其产绒量的遗传力为0.150~0.215。Rong等^[24]对内蒙古绒山羊(二狼山型)绒毛性状进行遗传参数估计,发现产绒量、绒细、毛长的遗传力分别为0.229、0.359、0.250。Zhao等^[25]使用ssGBLUP法对中国美利奴羊毛长的遗传参数进行估计,发现其遗传力为0.27。以上差异可能是由品种、数据量以及模型不同造成的。虽品种不同,但ssGBLUP与GBLUP同属基于基因组关系矩阵的方法,统计逻辑相近,因此结果具有一定可比性。王志英等^[26]对周岁内蒙古阿尔巴斯绒山羊开展遗传参数估计发现,毛长和绒细的遗传力分别为0.28和0.34,均属于中等偏高遗传力。陕北白绒山羊的产绒量和绒细遗传力分别为0.18、0.23^[27];内蒙古白绒山羊产绒量和绒细的遗传力分别为0.34和0.32^[28]。本研究中,“敏盖”绒山羊产绒量和绒细遗传力的结果与在辽宁绒

山羊报道的遗传力结果(0.357、0.421)^[29]基本一致。与国内部分地方绒山羊相比遗传力较高,因此,对其进行遗传选育可以快速实现遗传改良。因“敏盖”绒山羊育种过程是以内蒙古绒山羊(阿尔巴斯型)为母本,辽宁绒山羊为父本,经过级进杂交产生符合理想型的杂种后代。拥有内蒙古绒山羊和辽宁绒山羊的优质绒毛基因,三者遗传背景相似,使得核心绒毛性状的遗传变异水平接近。除遗传因素以外,各研究的群体结构差异、群体规模大小、评估方法及模型中考虑随机效应的差异,使得不同研究的遗传参数估计结果具有一定的异质性。

在畜禽的育种规划中,通常需统筹多个性状进行选择,性状间的遗传与表型相关是个体间接选择的标准^[30]。本研究对“敏盖”绒山羊绒毛性状进行遗传和表型相关的估计,发现性状间均存在不同程度的正相关,其中绒厚与毛长存在强的正遗传相关,遗传相关范围为0.66~0.84,这与李彦伯^[31]对内蒙古绒山羊(二狼山型)的研究结果相似,因此通过对绒厚的选择可以实现其他性状的间接选择。本研究中产绒量与毛长间遗传相关范围为0.07~0.61,李学武等^[32]研究结果表明内蒙古绒山羊产绒量与毛长的遗传相关为-0.11~0.37,这与本研究结果略有差异,造成差异的原因除不同品种间的遗传基础存在差异外,饲养条件、管理模式及环境因子(如温度、营养水平)均可能改变毛囊活性与性状表达,进而导致绒毛性状遗传参数估算出现偏差。本研究结果显示,基于GBLUP方法下性状间的表

型相关均低于遗传相关,说明环境因素对绒毛性状存在一定影响。具体而言,饲养管理、营养水平等因素可能是导致表型相关低于遗传相关的关键环境变量。当目标性状的遗传相关与表型相关均较高时,可依据其表型相关进行选种;反之,若两者差异较大,则不仅只依赖表型相关,而应基于遗传相关进行选种,通过性状间的间接选择,实现多个性状的遗传改良^[33]。

本研究系统探讨了GBLUP、Bayes A、Bayes B、Bayes C和Bayes LASSO方法对“敏盖”绒山羊毛长、绒厚、绒细、产绒量性状基因组预测的准确性。结果表明,Bayes方法估计的基因组育种值预测准确性均高于GBLUP方法,Bayes B方法的基因组预测准确性比Bayes A、Bayes C、Bayes LASSO方法有更好的预测能力。多重比较结果显示,GBLUP法与Bayes法预测的绒毛性状基因组准确性存在显著差异。可能是由于Bayes模型通过引入不同的先验分布和模型结构来适应遗传模型和数据特征,从而在性状的遗传结构未知时,能够比GBLUP法更准确地预测加性遗传效应值^[34]。Kjetså等^[35]利用GBLUP、Bayes C、Bayes GC三种方法对长白母猪的产仔猪数、仔猪死亡率等6个繁殖性状进行基因组育种值估计,发现Bayes GC的预测准确性比GBLUP预测准确性高,这说明Bayes GC更适合用于长白母猪产仔猪数等母性性状基因组育种值估计。张嘉欣^[36]利用Bayes A、Bayes B、Bayesian LASSO、Bayesian RR、GBLUP方法对内蒙古绒山羊产绒量、毛长、绒长、绒细性状进行基因组准确性预测,发现GBLUP预测的绒毛性状的基因组育种值准确性显著优于Bayes方法。葛菲^[37]在阿什旦牦牛早期生长性状基因组选择研究中发现,Bayes B模型对大部分性状的基因组预测准确性都优于Bayes A模型,与本研究结果相似,说明Bayes B方法的假设比其他贝叶斯方法更符合早期生长性状的遗传效应分布。刘勃攻等^[38]利用PBLUP、SSGBLUP、GBLUP、Bayes A、Bayes B、Bayes C等6种方法对湘佳黑凤鸡S1系产蛋性状进行基因组育种值的估计,结果表明,Bayes B方法在6种方法中获得最高的育种值估计准确性,与本研究结果相似。Baby等^[39]在韩国伯克希尔种猪群中使用GBLUP和Bayes B方法对其肉品质性状进行了基因组选择的研究,结果显示Bayes B方法下肉品质性状的基因组选择准确性均低于GBLUP方

法,与本试验结果相反,这可能是由于试验品种、群体规模大小、性状差异等因素造成的。在本研究比较的5种方法中,Bayes B方法是对“敏盖”绒山羊绒毛性状进行基因组选择的最优方法,在实际育种过程中,在运算效率允许的情况下,优先考虑运用Bayes B方法进行基因组育种值估计。

4 结 论

本研究使用GBLUP、Bayes A、Bayes B、Bayes C和Bayes LASSO五种方法对“敏盖”绒山羊绒毛性状进行遗传评估及育种值准确性评价,研究结果表明,GBLUP方法下产绒量和绒细的遗传力最高(0.32、0.49),Bayes A方法下毛长和绒厚的遗传力最高(0.32、0.33)。5种方法对其绒毛性状遗传相关性的估计表明,绒厚与毛长之间存在着强的遗传正相关,产绒量与毛长、产绒量与绒细之间呈相对较低的正遗传相关。5种方法对其绒毛性状的基因组育种值预测的准确性表明,Bayes类方法的基因组预测准确性高于GBLUP方法,Bayes类方法中Bayes B的准确性最高。尽管Bayes B的计算过程相对复杂,但其分析流程易于标准化,一旦建立便可稳定、批量运行,技术门槛可控,能实现更早、更精准的种羊筛选。因此,Bayes B方法可以有效应用于“敏盖”绒山羊绒毛性状的基因组选择,加快种群遗传进展,缩短世代间隔。

参考文献(References):

- [1] 王泽平,张丽芬,胡日查,等. 伊金霍洛旗绒山羊产业现状及发展思路[J]. 畜牧业环境, 2023, (15): 114-115.
WANG Z P, ZHANG L F, HU R Z, et al. Status quo and development idea of cashmere goat industry in Yijinhuo Banner[J]. *Livestock Environment*, 2023 (15): 114-115. (in Chinese)
- [2] MEUWISSEN T H, HAYES B J, GODDARD M E. Prediction of total genetic value using genome-wide dense marker maps [J]. *Genetics*, 2001, 157(4): 1819-1829.
- [3] ZHANG Z, ZHANG Q, DING X. Advances in genomic selection in domestic animals [J]. *Chin Sci Bull*, 2011, 56: 2655-2663.
- [4] MISZTAL I, LEGARRA A. INVITED REVIEW: Efficient computation strategies in genomic selection [J]. *Animal*, 2017, 11(5): 731-736.

- [5] VANRADEN PM. Efficient methods to compute genomic predictions [J]. *Dairy Sci*, 2008, 91 (11) : 4414-4423.
- [6] CHRISTENSEN O F, LUND M S. Genomic prediction when some animals are not genotyped [J]. *Genet Sel Evol*, 2010, 42(1):2.
- [7] ZHAO H, XIA J, ZHANG X, et al. Diet affects muscle quality and growth traits of grass carp (*Ctenopharyngodon idellus*): A comparison between grass and artificial feed [J]. *Front Physiol*, 2018, 9:283.
- [8] VERBYLA K L, HAYES B J, BOWMAN P J, et al. Accuracy of genomic selection using stochastic search variable selection in Australian Holstein Friesian dairy cattle [J]. *Genet Res (CAMB)*, 2009, 91(5):307-311.
- [9] HABIER D, FERNANDO R L, KIZILKAYA K, et al. Extension of the bayesian alphabet for genomic selection [J]. *BMC Bioinformatics*, 2011, 12:186.
- [10] PARK T, CASELLA G. The Bayesian Lasso [J]. *J Am Stat Assoc*, 2008, 103(482):681-686.
- [11] YU H, ZHANG K, CHENG G, et al. Genome-wide analysis reveals genomic diversity and signatures of selection in Qinchuan beef cattle [J]. *BMC Genomics*, 2024, 25(1):558.
- [12] WU X, ZHANG H, LONG H, et al. Genome-wide selection signal analysis to investigate wide genomic heredity divergence between Eurasian Wild Boar and Domestic Pig [J]. *Animals (Basel)*, 2023, 13 (13) : 2158.
- [13] GUO X, XING C H, WEI W, et al. Genome-wide scan for selection signatures and genes related to heat tolerance in domestic chickens in the tropical and temperate regions in Asia [J]. *Poult Sci*, 2022, 101(7): 101821.
- [14] KIM E H, KANG H C, SUN D W, et al. Estimation of breeding value and accuracy using pedigree and genotype of Hanwoo cows (Korean cattle) [J]. *Anim Breed Genet*, 2022, 139(3):281-291.
- [15] PETERS S O, KIZILKAYA K, IBEAGHA-AWEMU E M, et al. Comparative accuracies of genetic values predicted for economically important milk traits, genome-wide association, and linkage disequilibrium patterns of Canadian Holstein cows [J]. *Dairy Sci*, 2021, 104(2):1900-1916.
- [16] LIU T, QU H, LUO C, et al. Accuracy of genomic prediction for growth and carcass traits in Chinese triple-yellow chickens [J]. *BMC Genet*, 2014, 15:110.
- [17] WOLC A, DROBIK-CZWARNO W, JANKOWSKI T, et al. Accuracy of genomic prediction of shell quality in a White Leghorn line [J]. *Poult Sci*, 2020, 99 (6):2833-2840.
- [18] VANRADEN P M. Efficient methods to compute genomic predictions [J]. *Dairy Sci*, 2008, 91 (11) : 4414-4423.
- [19] YI N, XU S. Bayesian LASSO for quantitative trait loci mapping [J]. *Genetics*, 2008, 179(2):1045-1055.
- [20] MUCHA S, TORTEREAU F, DOESCHL-WILSON A, et al. Animal board invited review: Meta-analysis of genetic parameters for resilience and efficiency traits in goats and sheep [J]. *Animal*, 2022, 16(3):100456.
- [21] LI W H, LI G Y, ZHANG J, et al. Estimates of (co) variance components and phenotypic and genetic parameters of growth traits and wool traits in Alpine Merino sheep [J]. *Anim Breed Genet*, 2022, 139(3) : 351-365.
- [22] GAO L Y, XU Q, HE Y X, et al. Comparison of genomic prediction methods for early growth traits of Inner Mongolia cashmere goats based on multi trait models [J]. *Yi Chuan*. 2024, 46(5):421-430.
- [23] 习海娇, 李金泉, 张燕军, 等. 显性效应对内蒙古绒山羊产绒量和绒纤维直径育种值估计准确性的影响 [J]. *畜牧兽医学报*, 2025, 56(2):571-581.
- XI H J, LI J Q, ZHANG Y J, et al. Influence of dominance effects on the accuracy of breeding value estimation of cashmere production and cashmere diameter in Inner Mongolia Cashmere Goats [J]. *Acta Veterinaria et Zootechnica Sinica*, 2025, 56(2) : 571-581. (in Chinese)
- [24] RONG Y, WANG X, NA Q, et al. Genome-wide association study for cashmere traits in Inner Mongolia cashmere goat population reveals new candidate genes and haplotypes [J]. *BMC Genomics*, 2024, 25(1):658.
- [25] ZHAO B, LUO H, HUANG X, et al. Integration of a single-step genome-wide association study with a multi-tissue transcriptome analysis provides novel insights into the genetic basis of wool and weight traits in sheep [J]. *Genet Sel Evol*, 2021, 53(1):56.
- [26] 王志英, 张燕军, 王瑞军, 等. 内蒙古阿尔巴斯周岁绒山羊主要经济性状遗传参数的研究 [J]. *中国草食动物科学*, 2012, (S1):199-202.
- WANG Z Y, ZHANG Y J, WANG R J, et al. Study on genetic parameters of major economic traits in Albas week-old cashmere goats in Inner Mongolia [J]. *Chinese Herbivore Science*, 2012, (S1) : 199-202. (in

- Chinese)
- [27] 马彩英,陈玉林,冯喜波,等. 舍饲后陕北白绒山羊主要经济性状遗传力估计[J]. 家畜生态学报,2019,40(7):75-77.
MA C Y, CHEN Y L, FENG X B, et al. Estimation on Heritability Capacity of Major Economic Traits of Shaanbei White Cashmere Goat after Indoor Feeding [J]. *Journal of Livestock Ecology*, 2019, 40(7): 75-77. (in Chinese)
- [28] 李学武,王瑞军,王志英,等. 内蒙古绒山羊抓绒性状的遗传参数及遗传进展估计的研究[J]. 黑龙江畜牧兽医,2017,(1):117-120+124.
LI X W, WANG R J, WANG Z Y, et al. Study on the estimation of genetic parameters and genetic progress for fleece traits of Inner Mongolian cashmere goats [J]. *Heilongjiang Animal Husbandry and Veterinary Medicine*, 2017,(1):117-120+124. (in Chinese)
- [29] 马 宁,李永军,宋雅芹,等. 辽宁绒山羊主要经济性状遗传参数的估计[J]. 吉林农业大学学报,2005,(4):446-448+452.
MA N, LI Y J, SONG Y Q, et al. Estimates of genetic parameters of main economic traits in Liaoning Cashmere Goat [J]. *Journal of Jilin Agricultural University*, 2005,(4):446-448+452. (in Chinese)
- [30] 张 沅. 家畜育种学[M]. 北京:中国农业出版社,2001.
ZHANG Y. *Livestock Breeding* [M]. Beijing: China Agricultural Press, 2001. (in Chinese)
- [31] 李彦伯. 内蒙古绒山羊(二狼山型)重要经济性状遗传参数估计[D]. 呼和浩特:内蒙古农业大学,2023.
LI Y B. Estimation of genetic parameters of important economic traits in Inner Mongolia cashmere goat (Erlang Mountain type) [D]. Hohhot: Inner Mongolia Agricultural University, 2023. (in Chinese)
- [32] 李学武,刘 燕,王瑞军,等. 内蒙古绒山羊不同毛被类型产绒量和体重的遗传参数估计[J]. 中国农业科学,2018,51(12):2410-2417.
LI X W, LIU Y, WANG R J, et al. Genetic Parameter Estimation of Cashmere Yield and Body Weight at Different Staple Types of Inner Mongolian Cashmere Goats [J]. *Chinese Agricultural Science*, 2018, 51(12):2410-2417. (in Chinese)
- [33] 王大广,赵志辉,赵玉民,等. 萨福克羊表型测定值遗传参数的估计[J]. 中国草食动物,2003,(S1):67-68.
WANG D G, ZHAO Z H, ZHAO Y M, et al. Estimation of genetic parameters for phenotypic determination of Suffolk sheep [J]. *Chinese Herbivores*, 2003,(S1):67-68. (in Chinese)
- [34] ZENG J, PSZCZOLA M, WOLC A, et al. Genomic breeding value prediction and QTL mapping of QTLMAS2011 data using Bayesian and GBLUP methods[J]. *BMC Proc*, 2012, 6(Suppl 2):S7.
- [35] KJETSÅ M V, GJUVSLAND A B, NORDBØ Ø, et al. Accuracy of genomic prediction of maternal traits in pigs using Bayesian variable selection methods[J]. *Anim Breed Genet*, 2022, 139(6):654-665.
- [36] 张嘉欣. 基于GBLUP和BAYES方法对内蒙古绒山羊绒毛性状基因组预测准确性研究[D]. 呼和浩特:内蒙古农业大学,2023.
ZHANG J X. Studies on the accuracy of genome prediction of cashmere traits in Inner Mongolia Cashmere Goats based on GBLUP and Bayes methods [D]. Hohhot: Inner Mongolia Agricultural University, 2023. (in Chinese)
- [37] 葛 菲. 阿什旦牦牛早期生长性状的全基因组选择与关联分析[D]. 北京:中国农业科学院,2021.
GE F. Genomic selection and genome-wide association analysis for early growth traits in Ashidan yak [D]. Beijing: Chinese Academy of Agricultural Sciences, 2021. (in Chinese)
- [38] 刘勃攻,杨 杰,肖 健,等. 利用BLUP和BAYES方法对湘佳黑凤鸡S1系产蛋性状基因组育种值预测效果的比较研究[J]. 中国家禽,2025,47(1):1-7.
LIU B G, YANG J, XIAO J, et al. Comparative study on the effect of predicting genomic breeding values for laying traits of Xiangjia Heifeng S1 line chicken using BLUP and Bayes methods [J]. *Chinese Poultry*, 2025, 47(1):1-7. (in Chinese)
- [39] BABY S, HYEONG K, LEE Y M, et al. Evaluation of genome based estimated breeding values for meat quality in a berkshire population using high density single nucleotide polymorphism chips [J]. *Asian-Australas J Anim Sci*, 2014, 27(11):1540-1547.

(编辑 郭云雁)