Check for updates

Genome-Wide Association Study Using Individual Single-Nucleotide Polymorphisms and Haplotypes for Erythrocyte Traits in Alpine Merino Sheep

Shaohua Zhu[1,2†], Tingting Guo[1,2†], Hongchang Zhao[1,2], Guoyan Qiao[1,2], Mei Han[1,2], Jianbin Liu[1,2], Chao Yuan[1,2], Tianxiang Wang[3], Fanwen Li[3], Yaojing Yue[2*] and Bohui Yang[1*]

[1] Animal Science Department, Lanzhou Institute of Husbandry and Pharmaceutical Sciences, Chinese Academy of Agricultural Sciences, Lanzhou, China, [2] Sheep Breeding Engineering Technology Center, Chinese Academy of Agricultural Sciences, Lanzhou, China, [3] Gansu Provincial Sheep Breeding Technology Extension Station, Sunan, China

Adaptation to high-altitude hypoxia is essential for domestic animals, such as yak, Tibetan chicken, and Tibetan sheep, living on high plateaus, as it ensures efficient oxygen absorption and utilization. Red blood cells are the primary medium for transporting oxygen in the blood. However, little is known about the genetic mechanism of erythrocyte traits. Genome-wide association studies (GWASs) based on single markers or haplotypes have identified potential mechanisms for genetic variation and quantitative traits. To identify loci associated with erythrocyte traits, we performed a GWAS based on the method of the single marker and haplotype in 498 Alpine Merino sheep for six erythrocyte traits: red blood cell count (RBC), hemoglobin (HGB), hematocrit (HCT), mean corpuscular hemoglobin (MCH), mean corpuscular hemoglobin concentration (MCHC), and RBC volume distribution width coefficient of variation (RWD_CV). Forty-two significant single-nucleotide polymorphisms (SNPs) associated with the six erythrocyte traits were detected by means of a single-marker GWAS, and 34 significant haplotypes associated with five erythrocyte traits were detected by means of haplotype analysis. We identified six genes (*DHCR24, SPATA9, FLI1, PLCB1, EFNB2*, and *SH2B3*) as potential genes of interest via gene function annotations, location, and expression variation. In particular, *FLI1* and *PLCB1* were associated with hematopoiesis and erythropoiesis, respectively. These results provide a theoretical basis for analyzing erythrocyte traits and high-altitude hypoxia adaptation in Alpine Merino sheep and will be a useful reference for future studies of plateau-dwelling livestock.

Keywords: genome-wide association study, single-nucleotide polymorphism, haplotype, high-altitude hypoxia adaptation, erythrocyte traits, alpine merino sheep

INTRODUCTION

The sheep (*Ovis aries*) was one of the first animals to be domesticated by humans (Zeder, 2008; Wang et al., 2014). Sheep provide a variety of resources, including wool, mutton, and milk, and thus play an indispensable role in the global agricultural economy. Merino and Merino-derived sheep breeds are widely distributed around the world; they represent vital genetic resources, both historically and economically, and have become the basis for breeding new varieties (Ciani et al., 2015). The Alpine Merino sheep is a new breed of Merino, developed by crossing Australian Merino sheep with Tibetan sheep, and lives in areas of the Qinghai-Tibet Plateau (average elevation > 4500 m above sea level) all year round. In addition to the quality of wool and meat they provide, they have the advantages of resistance to rough feeding and adaptation to high-altitude hypoxia. The latter adaptation refers to a process of morphological, physiological, and biochemical changes that occur in humans or animals that have survived and evolved in the hypoxic environment of the plateau for generations, and are thus well adapted to that environment. This adaptation is the result of the survival and evolution of populations (including humans and animals) in the hypoxic environment, and it can be stably inherited (Peng et al., 2017). It is an efficient way for the body to exist in a low-oxygen environment, allowing populations to use the lower level of oxygen (60% of that at sea level) on the plateau to complete normal physiological functions and survive in the harsh environment.

Hemocytes (blood cells) have a vital function in conferring resistance to diseases and in the transport of oxygen (Müller and Brem, 1991). They include three main cell types: erythrocytes (red blood cells, RBCs), leukocytes (white blood cells), and thrombocytes (platelets) (Kickler, 2005). Of these three components, RBCs are crucial for the transport of oxygen and the removal of respiratory byproducts in organisms. To be utilized by the body, oxygen from the external environment needs to be transported and exchanged via the lungs, transported through the blood, and exchanged between the blood and the various tissues in the body. The viscosity of blood and the deformability of RBCs play a crucial role in transporting and exchanging oxygen; several studies have shown that the viscosity of blood depends on existing shear forces, RBC aggregation, the hematocrit (HCT, vol% of RBCs in blood), and the biomechanical properties of RBCs (Chien, 1987; Tripette et al., 2009). Other erythrocyte traits, including mean corpuscular hemoglobin (MCH) and mean corpuscular hemoglobin concentration (MCHC), are associated with impaired deformability of RBCs (von tempelhoff et al., 2015). In addition, the oxygen-carrying capacity of the blood is affected by the HCT and HGB content (Eklom and Lill, 2006). Thus, in the processes noted above, erythrocytes have essential functions in oxygen storage, binding, transportation, and release (Balasubramanian et al., 2009).

The genome-wide association study (GWAS) was originally developed as a means to screen genetic loci associated with complex diseases, and it is gradually becoming an effective tool for locating genetic sites associated with quantitative traits (Lockett and Holloway, 2013; Verstockt et al., 2018). In recent years, GWAS and quantitative trait loci (QTL) mapping have been widely used in research in animals, including pig (Wang et al., 2017), cattle (McDaneld et al., 2014), sheep (Kominakis et al., 2017), horse (Dupuis et al., 2011), and chicken (Zhang et al., 2013), as well as in humans (Davenport et al., 2015). Moreover, many traits have been evaluated in these studies, such as wool (Wang et al., 2014) and milk production (Fang et al., 2017), growth (Crispim et al., 2015), reproduction (Hawken et al., 2011), immune response (Go et al., 2012), and also high-altitude hypoxia adaptation (Hendrickson, 2013).

Several previous studies have reported gene function screening and QTL mapping results for high-altitude hypoxia adaptation in humans, pig, yak, Tibetan chicken, and Tibetan sheep (Zhang et al., 2007; Qiu et al., 2012; Dunn et al., 2013; Wei et al., 2016). Simonson et al. (2010) found three genes, *EPAS1* (endothelial PAS domain protein 1), *EGLN1* (egl-9 family hypoxia inducible factor 1), and *PPARA* (peroxisome proliferator activated receptor alpha), associated with HGB concentration by whole-genome scan. Favier et al. (2001) found that most of the factors or enzyme changes caused by hypoxia are regulated by *EPAS1* and noted that the latter was a crucial transcription factor for stress hypoxia. Zhang et al. (2007) measured hematological characteristics in Tibetan chickens at different ages and found that the chickens retained a high level of RBCs and a stable level of hematocrit as they aged. Wei et al. (2016) studied the relationship between blood-related phenotypes and *EPAS1* genotypes in Tibetan sheep by genome-wide analysis and showed that *EPAS1* was associated with increased MCHC and mean corpuscular volume. Haase et al. found that, under low-oxygen conditions, the body can stimulate the synthesis and release of erythropoietin through cellular oxygen-sensing pathways such as HIF-1X, thus promoting hemoglobin production and erythrocyte proliferation to adapt to a harsh environment (Haase, 2013). Yang et al. (2016) found that *HMOX2* (also known as *HO-2*, heme oxygenase 2) as a modified gene could regulate the metabolism of downstream hemoglobin in the hypoxic metabolic pathway, indicating a mechanism for genetic adaptation of the Tibetan population to low-altitude hypoxia. Zou et al. (2008) identified 101 QTLs affecting RBCs traits on wild boar (*Sus scrofa*) chromosomes 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, and 13 . However, as mentioned in the above study, some of the candidate genes detected in the results were consistent, but some new candidate genes have also been discovered. This may reflect the complex genetic mechanisms of blood cells and the differences between species.

To the best of our knowledge, few GWASs have been conducted in sheep than in humans and pigs because of the limited information available for the sheep genome. Moreover, there is no specific parameter defined for high-altitude hypoxia adaptation, and it can be reflected by related characteristics such as erythrocyte traits, cardiopulmonary function, and oxygen circulation in cells and tissues. Therefore, studies of high-altitude hypoxia adaptation in sheep are rarely reported. The Alpine Merino sheep involved in this study had a mixture of Australian Merino sheep and Tibetan sheep ancestry, they quickly adapting to the Qinghai-Tibet Plateau and surviving in high altitude and hypoxia adaption, with the consequence that they present an ideal

subject for experiment in adaptation to high-altitude hypoxia. The main objectives of the present study were to detect significant SNPs for erythrocyte traits in Alpine Merino sheep by using the custom Affymetrix HD 630K microarrays and to explore the potential genes associated with high-altitude hypoxia adaptation.

MATERIALS AND METHODS

Ethics Statement
All the animal work involved in this study was conducted according to the guidelines for the care and use of laboratory animals promulgated by the State Council of the China. This research was approved by the Animal Management and Ethics Committee of Lanzhou Institute of Animal Husbandry and Veterinary Medicine, Chinese Academy of Agricultural Sciences (license number: 2019-008).

Animal Samples and Phenotypes
The sheep population used in this study was selected from Huangcheng pasture in Gansu Province, China, part of the Gansu Provincial Sheep Breeding Technology Extension Station. The station has a strictly standardized breeding and management system, which ensured the health of each individual sheep involved in the experiment. The flock consisted of 501 male Alpine Merino sheep, aged 18–24 months. Blood samples (5 mL) from each of the 501 sheep were collected from the jugular vein and transferred immediately to a vacuum blood collection tube (Yuli Medical Equipment Company Ltd., Jiangsu Province, China).

A standard set of erythrocyte data was recorded using a H-100iv diff whole blood analyzer (Sysmex, Kobe, Japan) within 24 h of specimen collection at Lanzhou Institute of Animal Husbandry and Veterinary Medicine, Chinese Academy of Agricultural Sciences. The remainder of each blood sample was stored at −20°C after the hematological data were recorded. Six erythrocyte traits closely associated with high-altitude hypoxia adaptation were then selected: red blood cell count (RBC), hemoglobin (HGB), hematocrit (HCT), mean corpuscular hemoglobin (MCH), mean corpuscular hemoglobin concentration (MCHC), and red blood cell volume distribution width coefficient of variation (RWD_CV). Then, the phenotype data were adjusted through remove the extreme values that exceeded three times the standard deviation before conducting the association study.

Genotyping and Quality Control
In total, 501 Alpine Merino sheep were genotyped using the custom Affymetrix HD 630K microarrays, and the genotyping platform adopted in this study was based on the Array Plate Processing Workflow for the Thermo Fisher Scientific (Affymetrix) GeneTitan System (Santa Clara, CA, United States). Before statistical analysis, SNPs were preprocessed using PLINK v1.9b4 (Purcell et al., 2007). SNPs were selected based on minor allele frequency >0.01, proportion of missing genotypes <0.05, and a Hardy–Weinberg equilibrium P-value > 10e−6. Samples with more than 10% missing genotypes were eliminated. According to the definition given by Gabriel, if the 95%

confidence upper boundary on one side of D' was lower than 0.98 (that is, there is no historical reorganization), and the lower boundary was higher than 0.7, then the pair was considered to be a "strong LD." If the D' of a pair of SNPs was lower than 0.7, the next haplotype block was initiated (Gabriel et al., 2002). The reference thresholds and division criteria for tightly linked haplotype internal SNPs were established based on the recommendations of the PLINK software user manual (Purcell et al., 2007). A total of 498 individuals with 439,398 autosomal SNPs and 89,676 haplotypes blocks with 325,062 haplotypes passed filters and quality controls, and these standard-compliant samples were collated for subsequent statistical analysis.

Statistical Analyses
The method of Bonferroni calibrated multiple tests (Yang et al., 2006) was used to determine significance thresholds. The threshold of the genome-wide significance level of each SNP and haplotype was 0.05, and the threshold of the suggestive or chromosomal significance level of each effective SNP site and haplotype was 1. After calibration of the significance thresholds, the corresponding thresholds for the SNP sites were set as 1.14e−7 (0.05/439,398) and 2.28e-6 (1/439,398); similarly, the corresponding thresholds for haplotypes were set as 1.54e-7 (0.05/325,062) and 3.08e-6 (1/325,062). Based on the whole genome sequence information, we used TASSEL 5.2.43 software (Bradbury et al., 2007) to perform principal component analysis on all the individuals involved in the study, then constructed and drew the principal component analysis plot. In addition, the R package "CMplot" was used to create Manhattan plots and quantile-quantile (Q-Q) plots for each GWAS result[1].

Studies of Single-Marker Association
The associated data for each SNP through regression analysis based on the generalized linear model (GLM) given below by Eq. (1) (Breslow and Clayton, 1993; McCulloch and Neuhaus, 2015):

$$y = Xb_i + Qv + e \qquad (1)$$

where y is the corrected phenotype, b_i is the regression coefficient of the phenotype on the SNPs, X represents the vector of the corresponding SNP indicators, v represents the effect of population structure, Q is the corresponding principal components matrix, e is the vector of residual errors with $e \sim N\left(0, \quad I\sigma_e^2\right)$, I is the identity matrix, and σ_e^2 is the variance of e. The single-marker analysis was conducted using TASSEL 5.2.43 software (Bradbury et al., 2007).

Studies of Haplotype-Based Association
Similarly to the single-marker analysis, GLM was adopted to perform the regression analysis in the haplotype-based association studies; y, v, Q, and e were as defined previously, b_i represents the regression coefficient of the phenotype on the haplotype blocks, and X represents the vector of the corresponding haplotype indicators. Furthermore, the standard expectation maximization (EM) algorithm was used to detect individual haplotype blocks and haplotype frequencies.

---

[1]https://cran.r-project.org/web/packages/CMplot/index.html

Specifically, EM iterations interact between performing the expectation step (ES) and establishing the maximized step (MS). The ES and MS establish an equation for the expected value of the log-likelihood, which is evaluated using the estimate of the haplotype and frequencies (Excoffier and Slatkin, 1995). According to recommendations of the PLINK software user manual (Purcell et al., 2007), we established the reference thresholds and division criteria for tightly connected haplotype internal SNPs. The haplotype-based analysis was also conducted using the TASSEL 5.2.43 software (Bradbury et al., 2007).

RESULTS

Phenotypic Statistics and Genotypic Characteristics

A total of 498 Alpine Merino sheep were used in this study, all from Huangcheng pasture in Gansu Province, China. Six phenotypic observations related to erythrocyte traits were measured and recorded: RBC, HGB, HCT, MCH, MCHC, and RWD_CV. Descriptive statistics of the phenotypic measurements (**Table 1**) include the mean, standard deviation, and coefficient of variation. The Q-Q plots for the test statistics from the GLM are shown in **Figures 1A,B** and indicate that there was no overall systemic bias in the analysis based on single-marker or haplotype analysis. In addition, the structure of this population is drawn based on the top three eigenvectors using principal component 1 (PC1), 2 (PC2), and 3 (PC3), the PCA plot (**Supplementary Figure S1**) suggests that some of the individuals have population stratification, therefore, the population stratification should be corrected, it could effectively control the group stratification. As shown in Eq. (1), v and Q represent the influence of the population structure and the corresponding principal component matrix, respectively.

Significant SNPs

In single-marker analysis, we detected 42 significant SNPs, including 27 suggested and 15 genome-wide significant SNPs. These associated with six erythrocyte traits (**Supplementary Table S1**): 14 for RBC, 2 for HGB, 1 for HCT, 24 for MCH, 6 for MCHC, and 20 for RWD_CV, some of which were pleiotropic (12_39476100 were associated with both MCH and MCHC and 8_82239902 were associated with both RBC and RWD_CV). These SNPs were distributed on 20 autosomes, including Oar1, 2, 3, 4, 5, 7, 8, 9,10, 11, 12, 13, 14, 15, 17, 18, 21, 22, 24, and 26 (**Figure 2**). Of these 42 SNPs, 21 were located within genes and the others were located from 1,461 to 635,535 base pairs (bp) away from the nearest genes (**Supplementary Table S1**). The most significant SNP, 9_19840980, was associated with the RBC trait (p = 1.12e-10) (**Figure 2**). By means of genetic mapping and functional annotation, we identified five potential genes that were directly or indirectly related to high-altitude hypoxia adaptation (**Table 2**). Most of the SNPs were associated with MCH traits. Fourteen significant SNPs were associated with RBC and were located on Oar1, 2, 4, 8, 9, 13, 15, 17, and 21. Four SNPs were located within genes and others were located from 1,461 to 229,683 bp from the nearest gene. Two significant SNPs

were associated with the HGB trait, one (17_53452406) of them was located within *CAMKK2* (calcium/calmodulin-dependent protein kinase kinase 2), and the other was located on Oar18. Only one significant SNP was associated with the HCT trait; it was located 10,803 bp away from LRP1B (LDL receptor-related protein 1B). Twenty-four significant SNPs were associated with MCH and were distributed on Oar1, 2, 3, 4, 7, 8, 9, 10, 11, 12, 13, 15, 17, 21, and 22. Eleven SNPs were located within genes, and others were located from 1,461 to 635,535 bp from the nearest gene. Six significant SNPs were associated with MCHC, three of these were located within genes including *VPS13D* (vacuolar protein sorting 13 homolog D) on Oar12, OTOA (otoancorin) on Oar24, and RSU1 (Ras suppressor protein 1) on Oar13, and the others were located from 67,432 to 496,548 bp from the nearest gene. Twenty significant SNPs were associated with the RWD_CV trait, these were distributed on Oar1, 2, 3, 4, 5, 8, 9, 11, 14, 15, 17, and 21, 12 of them were located within genes, and the others were located from 1,461 to 207,033 bp from the nearest gene.

Significant Haplotypes

In haplotype analysis, we detected 32 significant haplotypes, including 24 suggested and eight genome-wide significant haplotypes. They were associated with five erythrocyte traits: 10 for RBC, 3 for HCT, 23 for MCH, 6 for MCHC, and 17 for RWD_CV, similarly to single-marker analysis, some of which were pleiotropic (haplotypes H1094_AA were associated with both RBC and MCH). These haplotypes were distributed on 15 autosomes, including Oar1, 2, 3, 4, 5, 7, 8, 9, 13, 16, 18, 19, 20, 21, and 26 (**Figure 3**). Of these 32 haplotypes, 13 were within or close to genes and the others were located from 1,668 to 1,176,920 bp away from the nearest genes (**Supplementary Table S2**). The most significant haplotype was associated with the RWD_CV trait (p = 1.21e-11) (**Figure 3**). By gene finding and functional annotation, we identified four potential genes that were directly or indirectly related to high-altitude hypoxia adaptation (**Table 3**). Most of the haplotypes were associated with MCH traits. Ten significant haplotypes were associated with RBC. Three significant haplotypes were associated with the HCT trait, which were located on Oar1, 9, and 13. Twenty-three significant haplotypes were associated with MCH and were distributed on Oar1, 2, 3, 4, 5, 7, 8, 13, 16, 20, and 21. Eight haplotypes were located close to or within genes, and others were located from 1,688 to 1,176,920 bp away from the nearest genes. Six significant haplotypes were associated with MCHC, two of them were located close to or within genes, and the others were located from 36,656 to 496,548 bp away from the nearest genes. Seventeen significant haplotypes were associated with the RWD_CV trait, 12 were distributed within genes, and the others were located from 6,529 to 229,076 bp away from the nearest genes.
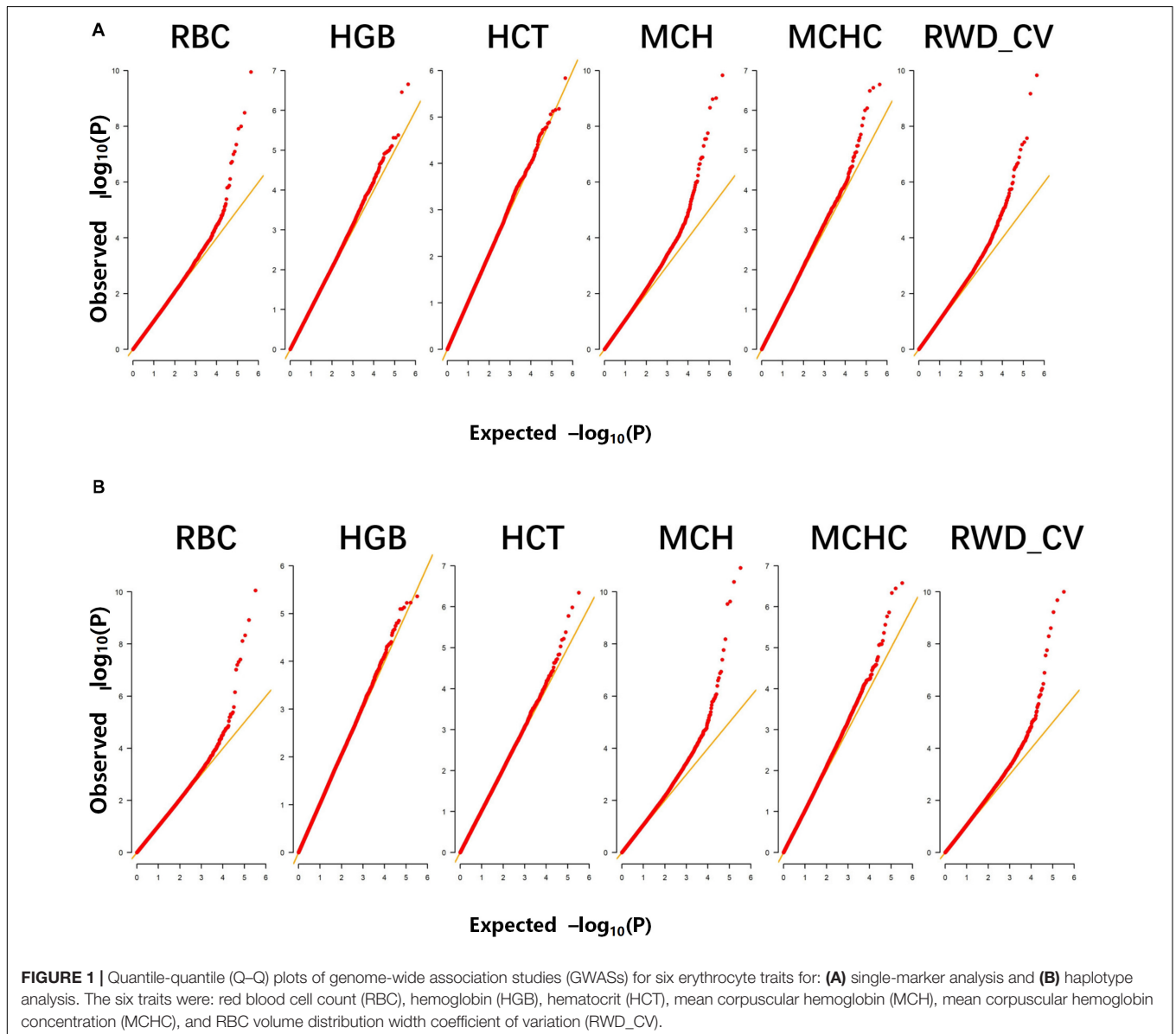
DISCUSSION

In the present study, both single-marker and haplotype-based analysis identified significant associations between erythrocyte

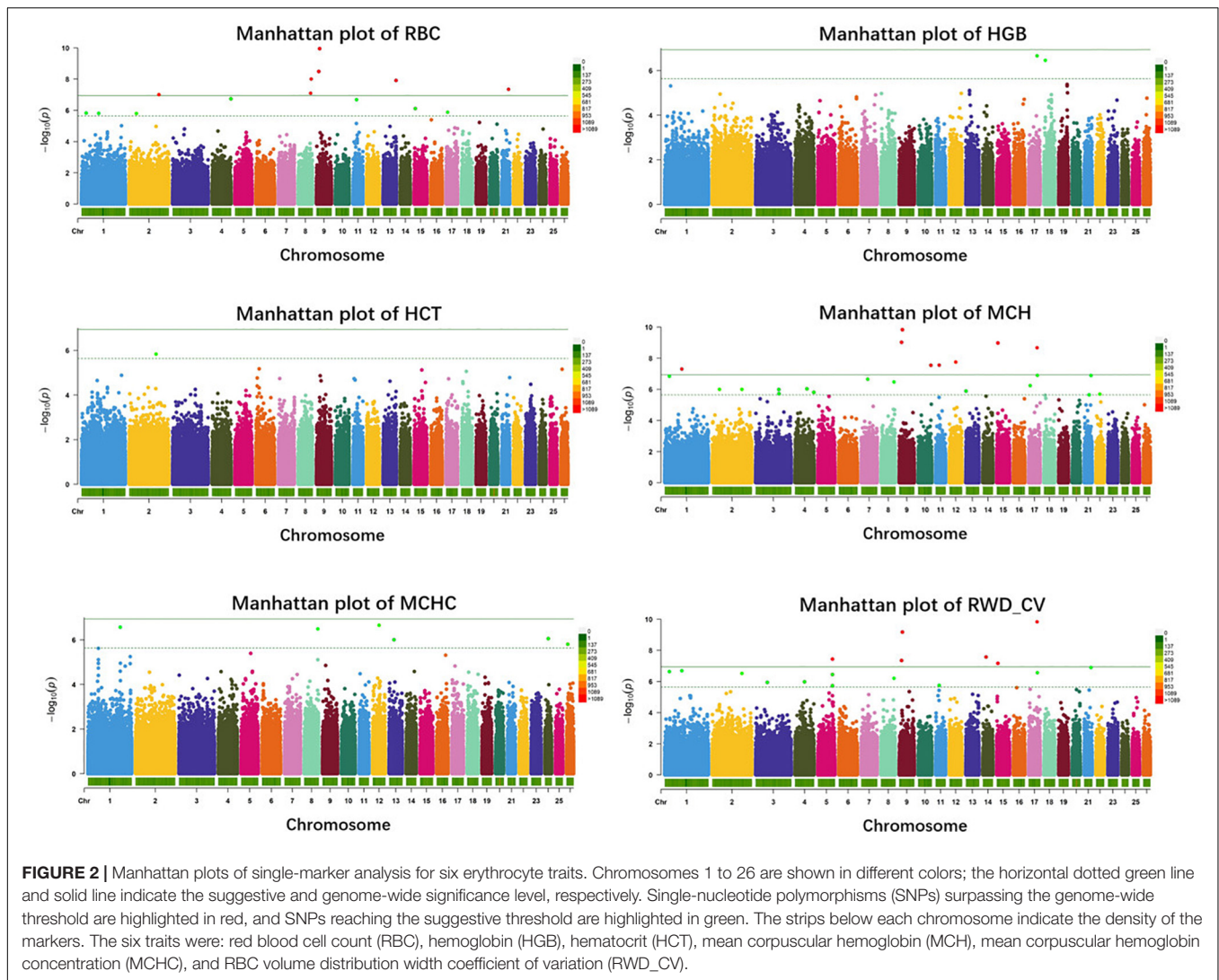**TABLE 1 |** Descriptive statistics of phenotype values of erythrocyte traits.

Trait	Abbreviation	Mean ± SD	CV[1] (%)	Number
Red blood cell count (10^{12}/L)	RBC	7.66 ± 1.69	22.08	494
Hemoglobin (g/L)	HGB	98.60 ± 11.64	11.80	489
Hematocrit (%)	HCT	0.27 ± 0.08	29.84	496
Mean corpuscular hemoglobin (pg)	MCH	13.24 ± 2.64	19.93	492
Mean corpuscular hemoglobin concentration (g/L)	MCHC	377.47 ± 103.77	27.49	488
RBC volume distribution width coefficient of variation	RWD_CV	0.39 ± 0.07	19.23	498

[1] CV, coefficient of variation.

FIGURE 1 | Quantile-quantile (Q–Q) plots of genome-wide association studies (GWASs) for six erythrocyte traits for: **(A)** single-marker analysis and **(B)** haplotype analysis. The six traits were: red blood cell count (RBC), hemoglobin (HGB), hematocrit (HCT), mean corpuscular hemoglobin (MCH), mean corpuscular hemoglobin concentration (MCHC), and RBC volume distribution width coefficient of variation (RWD_CV).

traits with comparable sheep genomic regions. The main reason for using GLM in haplotype-based association studies is that the process of obtaining the genetic relationship constructed by haplotypes is complex; elimination of rare haplotypes and low-frequency individuals followed by subsequent association studies could avoid adverse effects on significance testing effectively. The basic principle of single marker GWAS is to compare phenotypic differences in allelic groups, it is more concerned about the effect of each marker. Significant SNP may loss caused by low linkage disequilibrium (LD) and mutation between markers if the density

FIGURE 2 | Manhattan plots of single-marker analysis for six erythrocyte traits. Chromosomes 1 to 26 are shown in different colors; the horizontal dotted green line and solid line indicate the suggestive and genome-wide significance level, respectively. Single-nucleotide polymorphisms (SNPs) surpassing the genome-wide threshold are highlighted in red, and SNPs reaching the suggestive threshold are highlighted in green. The strips below each chromosome indicate the density of the markers. The six traits were: red blood cell count (RBC), hemoglobin (HGB), hematocrit (HCT), mean corpuscular hemoglobin (MCH), mean corpuscular hemoglobin concentration (MCHC), and RBC volume distribution width coefficient of variation (RWD_CV).

of markers is not high enough (Zhang et al., 2014), while the haplotype-based GWAS could effectively solve this problem, it is due to the fact that haplotype analysis could utilize both recent and ancestral reorganization information (Druet and Georges, 2010). However, haplotype analysis generally has greater degrees of freedom (df) than single-marker analysis, resulting in reduced detection capabilities (Zhang et al., 2012), it is difficult to balance between df and LD. In addition, the distribution of LDs in the whole genome is inconsistent, i.e., LD is higher in some districts and lower in other districts. Therefore, we conducted two different approaches to capture more associated SNPs. The differences in the numbers of markers (439,398 SNPs and 325,062 haplotypes) result in different *P*-values for the two methods. Although most of the genes involved are the same, the *P*-value distributions in the Manhattan plots from different analysis methods are slightly different. We detected some associations only via the SNP analysis, while other associations were detected only via the haplotype-based analysis. The two different methods detected that some of the associations were different. Therefore,

the overlap markers detected by two different methods could make the results more reliable, and the markers of non-overlap may represent new valuable associations detected by different methods. Our results show that the efficiency of the method depends on the type of genotype data. This is consistent with the findings of Xiaoming et al. (2019) and of Wu et al. (2014).

In this study, we discovered 42 significant SNPs via single-marker analyses; six erythrocyte traits were involved, including RBC, HGB, HCT, MCH, MCHC, and RWD_CV. We detected five potential genes (*DHCR24, EFNB2, SH2B3, PLCB1*, and *SPATA9*) that were directly or indirectly related to hypoxic adaptation based on gene functional annotations, positions, and reported expression variation. Heart ischemia is a significant threat to livestock living on the plateau, which is not conducive to oxygen intake and exchange. Marker 1_29018708 on Oar1, associated with three different traits including RBC, MCH, and RWD_CV, was located within *DHCR24* (24-dehydrocholesterol reductase). It has been reported that increased expression of this gene could activate phosphatidylinositol 3-kinase and induce

TABLE 2 | Associated SNPs and nearby potential genes for erythrocyte traits.

Trait [1]	SNP name	Oar [2]	Position (bp) [3]	*P*-value	Nearest gene [4]	
					Name	**Distance (bp)**
RBC	1_29018708	1	29018708	1.48e-06	*DHCR24*	Within
MCH	10_80901397	10	80901397	2.86e-08	*EFNB2*	635,535 bp
MCH	17_54731492	17	54731492	1.28e-07	*SH2B3*	Within
MCH	1_29018708	1	29018708	1.45e-07	*DHCR24*	Within
MCH	13_798029	13	798029	1.28e-06	*PLCB1*	Within
RWD_CV	1_29018708	1	29018708	2.30e-07	*DHCR24*	Within
RWD_CV	17_54731492	17	54731492	2.75e-07	*SH2B3*	Within
RWD_CV	5_92277630	5	92277630	3.67e-08	*SPATA9*	Within
RWD_CV	5_92265355	5	92265355	1.82e-06	*SPATA9*	Within
RWD_CV	5_92276610	5	92276610	1.82e-06	*SPATA9*	Within
RWD_CV	5_92256711	5	92256711	1.91e-06	*SPATA9*	Within

[1] HCT, hematocrit; MCH, mean corpuscular hemoglobin; MCHC, mean corpuscular hemoglobin concentration; RWD_CV, RBC volume distribution width coefficient of variation. [2] The Ovis aries chromosome number of each significant SNP. [3] The positions of each significant SNP associated with erythrocyte traits in Ovis aries Oar_4.0 assembly. [4] The nearest annotated gene to each significant SNP. The annotated gene database is from https://www.ncbi.nlm.nih.gov/assembly/GCF_000298735.2. SNP is designated as in a gene or as distance (bp) from a gene region in Ovis aries Oar_4.0 assembly.
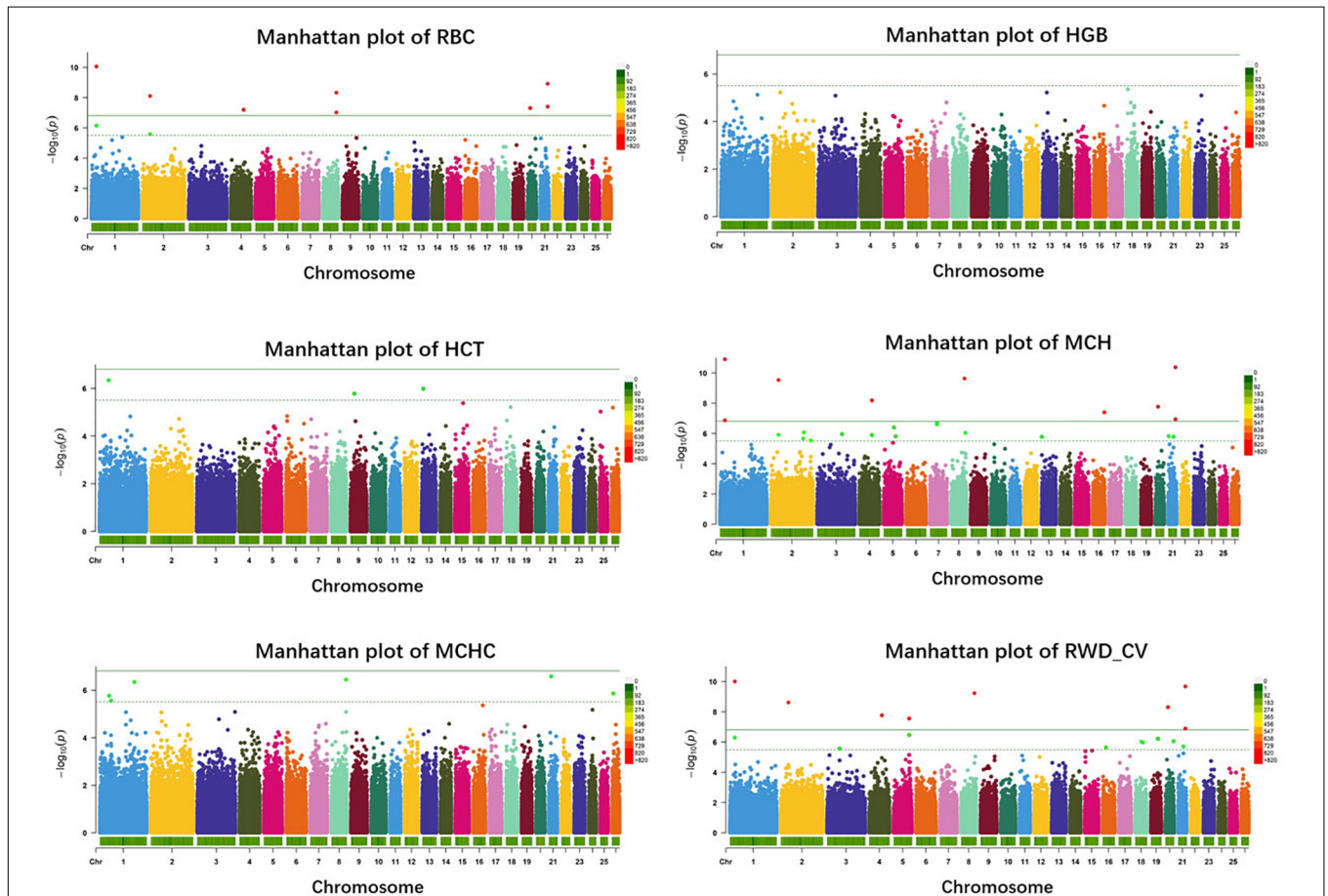

FIGURE 3 | Manhattan plots of haplotype-based analysis for six erythrocyte traits. Chromosomes 1 to 26 are shown in different colors; the horizontal dotted green line and solid line indicate the suggestive and genome-wide significance level, respectively. Haplotypes surpassing the genome-wide threshold are highlighted in red, and haplotypes reaching the suggestive threshold are highlighted in green. The strips below each chromosome indicate the density of the markers. The six traits were: red blood cell count (RBC), hemoglobin (HGB), hematocrit (HCT), mean corpuscular hemoglobin (MCH), mean corpuscular hemoglobin concentration (MCHC), and RBC volume distribution width coefficient of variation (RWD_CV).

TABLE 3 | Associated haplotype blocks and nearby potential genes for erythrocyte traits.

Trait [1]	NSNP [3]	Oar [4]	Start [5]	End [5]	SNP1 [6]	SNP2 [6]	*P*-value	Nearest gene [7]	
								Name	**Distance (bp)**
RBC	2	1	29018708	29023326	1_29018708	1_29023326	8.77e-11	*DHCR24*	Within
RBC	2	1	29018708	29023326	1_29018708	1_29023326	7.00e-07	*DHCR24*	Within
MCH	2	1	29018708	29023326	1_29018708	1_29023326	1.21e-11	*DHCR24*	Within
MCH	2	1	29018708	29023326	1_29018708	1_29023326	1.33e-07	*DHCR24*	Within
MCH	3	21	31336799	31343543	21_31336799	21_31343543	1.64e-06	*FLI1*	Within
MCH	2	13	798029	800471	13_798029	13_800471	1.66 e-06	*PLCB1*	Within
MCH	3	21	31336799	31343543	21_31336799	21_31343543	1.97e-06	*FLI1*	Within
RWD_CV	2	5	92277625	92277630	5_92277625	5_92277630	2.75e-08	*SPATA9*	Within
MCH	2	1	29018708	29023326	1_29018708	1_29023326	9.93e-11	*DHCR24*	Within
MCH	2	1	29018708	29023326	1_29018708	1_29023326	5.17e-07	*DHCR24*	Within

[1] MCH, mean corpuscular hemoglobin; RWD_CV, red blood cell volume distribution width coefficient of variation; HCT, hematocrit. [2] The name of significant haplotypes for each erythrocyte trait. [3] Number of SNPs included in haplotype block. [4] The Ovis aries chromosome number of each significant haplotype. [5] The start and end positions of each significant haplotype block associated with erythrocyte traits in Ovis aries Oar_4.0 assembly. [6] The first and second SNPs in the haplotype block. [7] The nearest annotated gene to each haplotype block. The annotated gene database is from https://www.ncbi.nlm.nih.gov/assembly/GCF_000298735.2. SNP is designated as a gene or as distance (bp) from a gene region in Ovis aries Oar_4.0 assembly.

HO-1 (heme oxygenase-1) (Wu et al., 2012). *HO-1* induction prevents heart ischemia and reperfusion injury (Otterbein et al., 2003), while the normal function of the heart ensures blood circulation and oxygen transport, which is critical for livestock to survival in hypoxic conditions. *EFNB2* (ephrin B2) encodes a member of the ephrin (EPH) family, and the EPH receptors are the largest subfamily of receptor tyrosine kinases (*RTK*s), which have been implicated in mediating developmental events, especially in the nervous system and in erythropoiesis (Bennett et al., 1995; Suenobu et al., 2002). Duffy et al. (2010) pointed out that mesenchymal stem cells (MSCs) expressing *EFNB2* could promote cardiovascular formation in ischemic tissue; a very significant marker (10_80901397, p = 2.86e-08) found in our study was located 635,535 bp away from this gene. In addition, another gene, *SH2B3* (SH2B adaptor protein 3), closely related to hematopoietic function, was revealed in this study. Marker 17_54731492 on Oar17, associated with MCH and RWD_CV, was located within this gene that encodes a member of the *SH2B* adaptor family of proteins, which is an adaptor protein regulating the production of B cells (Li et al., 2000). Satoshi et al. found that *SH2B3* is expressed in early hematopoietic progenitor cells and that the hematopoietic capacity of hematopoietic stem cells (HSCs) increases significantly when the expression of SH2B decreases. They speculated that this gene plays a key role in HSC amplification and function (Takaki et al., 2000). All red blood cells are formed by differentiation of HSCs, and HSCs would self-renew throughout the life cycle of the body to continuously generate new blood cells (Till et al., 1964; Reya, 2003). Marker 13_798029 on Oar13, associated with MCH, was located within PLCB1 (phospholipase C beta 1). Several studies have proved that *PLCB1* participates in the body's hematopoietic system and regulates the differentiation and production of red blood cells (Cocco et al., 2015). Follo et al. (2012) determined that *PLCB1* inhibited the differentiation process of red blood cell lines, while another isoenzyme, PI-PLC gamma, increased during the differentiation of red blood

cells, which regulate the production of normal red blood cells. A total of four SNPs were located on Oar5 (**Supplementary Table S1**), associated with the RWD_CV trait, and were located within *SPATA9* (markers 5_92277630, 5_92265355, 5_92276610, and 5_92256711). Artigas et al. identified 16 new loci that affect lung function through genome-wide association and large-scale follow-up (46,411 individuals), including *SPATA9, TGFB2, C10orf11,* and *LRP1*; they found that SPATA9 was expressed in the lungs, airway smooth muscle, bronchial epithelial cells, and peripheral blood mononuclear cells. Moreover, they also found that this gene was associated with the forced expiratory volume in 1 s (FEV1) and the ratio of FEV1 to forced vital capacity (FVC) (Soler Artigas et al., 2011). Interestingly, a reduced ratio of FEV1 to FVC is used to define airway obstruction, and a reduced FEV1 is used to grade the severity of airway obstruction (Jakeways et al., 2003). These two parameters directly affect the absorption and utilization of oxygen by livestock in low-oxygen environments.

We identified 34 significant haplotypes associated with the five erythrocyte traits. By genetic mapping and functional annotation of these haplotypes, three important genes, including *DHCR24* on Oar1, *PLCB1* on Oar13, and *SPATA9* on Oar5, were overlapped with single-marker GWAS, which will not be discussed further in this section. In addition to the above three genes, we identified another important gene in the haplotype analysis, *FLI1* (Fli-1 proto-oncogene, ETS transcription factor). One significant haplotype block associated with MCH was found within this gene. *FLI1* is a member of the ETS transcription factor family, and analysis of the expression of members of this family in hematopoietic tissues, as well as studies of targeted mutations in ETS family members, suggest that they are critical for regulation of normal hematopoietic development (Graves and Petersen, 1998). Spyropoulos et al. (2000) demonstrated that hematopoiesis is severely impaired in *FLI1* mutant mice at mid-gestation, and they speculated that the loss of function of *FLI1* may lead to disruption of normal hematopoietic function in mice. Based on our results from single-marker GWAS and haplotype analysis, we

speculate that the expression of these six genes may be related to high-altitude hypoxia adaptation, and we will continue to investigate the biological functions of the above potential genes.

CONCLUSION

We identified 42 SNPs and 34 haplotypes, some of which were pleiotropic, such as 8_82239902, and were associated with both RBC and RWD_CV. Six genes were selected as potential for genes involved in high-altitude hypoxia adaptation based on functional annotations and gene mapping. In particular, *PLCB1* and *FLI1* are associated with the production of RBCs and normal hematopoietic function, respectively. The current results provide a valuable reference for studies of high-altitude hypoxia adaptation in domestic animals, such as the Alpine Merino sheep, living on high-altitude plateaus. Further research on these genes will be conducted, and the identified SNPs and haplotypes will be investigated in other populations and functionally verified to identify causal mutations.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://figshare.com/articles/dataset/Phenotype_dataset_blood_498_txt/12072261, https://figshare.com/articles/dataset/Genotype_dataset_GWASQC_498_vcf/12072249.

ETHICS STATEMENT

The animal study was reviewed and approved by the Animal Management and Ethics Committee of Lanzhou Institute of Animal Husbandry and Veterinary Medicine, Chinese Academy of Agricultural Sciences. Written informed consent was obtained from the owners for the participation of their animals in this study.

AUTHOR CONTRIBUTIONS

SZ and TG conceptualized the study. SZ and CY contributed to the data curation. SZ and TG contributed to the formal analysis. BY contributed to the funding acquisition. HZ and MH investigated the study. TG performed the methodology. YY and BY contributed to the project administration. TW and FL contributed to the resources. JL and CY contributed to the software. HZ contributed to the validation. GQ visualized the study. SZ and TG wrote the original draft. YY and BY wrote, reviewed and edited the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2020.00848/full#supplementary-material

FIGURE S1 | Population structure map drawn from the first three principal components.

TABLE S1 | All associated SNPs and nearby genes for erythrocyte traits.

TABLE S2 | All associated haplotype blocks and nearby genes for erythrocyte traits.

REFERENCES

Balasubramanian, M., Moorthy, P. S., Neelagandan, K., and Ponnuswamy, M. N. (2009). Purification, crystallization and preliminary crystallographic study of haemoglobin from camel (*Camelus dromedarius*): a high oxygen-affinity lowland species. *Acta Crystallogr.* 65(Pt 8), 773–775. doi: 10.1107/s1744309109024154

Bennett, B., Zeigler, F., Gu, Q. M., Fendly, B., Goddard, A., Gillett, N., et al. (1995). Molecular cloning of a ligand for the EPH-related receptor protein-tyrosine kinase HtK. *Proc. Natl. Acad. Sci. U.S.A.* 92, 1866–1870. doi: 10.1073/pnas.92.6.1866

Bradbury, P., Zhang, Z., Kroon, D., Casstevens, T., Ramdoss, Y., and Buckler, E. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* 23, 2633–2635. doi: 10.1093/bioinformatics/btm308

Breslow, N. E., and Clayton, D. G. (1993). Approximate inference in generalized linear mixed models. *J. Am. Stat. Assoc.* 88, 9–25. doi: 10.2307/2290687

Chien, S. (1987). Red cell deformability and its relevance to blood flow. *Annu. Rev. Physiol.* 49, 177–192. doi: 10.1146/annurev.ph.49.030187.001141

Ciani, E., Emiliano, L., D'Andrea, M., Alloggio, I., Marroni, F., Ceccobelli, S., et al. (2015). Merino and Merino-derived sheep breeds: a genome-wide intercontinental study. *Genet. Select. Evol. GSE* 47:64. doi: 10.1186/s12711-015-0139-z

Cocco, L., Manzoli, L., Faenza, I., Ramazzotti, G., Yang, Y., McCubrey, J., et al. (2015). Modulation of nuclear PI-PLCbeta1 during cell differentiation. *Adv. Biol. Regul.* 60, 1–5. doi: 10.1016/j.jbior.2015.10.008

Crispim, A., Kelly, M., Guimarães, S., Silva, F., Fortes, M., Wenceslau, R., et al. (2015). Multi-Trait GWAS and new candidate genes annotation for growth curve parameters in brahman cattle. *PLoS One* 10:e0139906. doi: 10.1371/journal.pone.0139906

Davenport, E., Cusanovich, D., Michelini, K., Barreiro, L., Ober, C., and Gilad, Y. (2015). Genome-wide association studies of the human gut microbiota. *PLoS One* 10:e0140301. doi: 10.1371/journal.pone.0140301

Druet, T., and Georges, M. (2010). A hidden markov model combining linkage and linkage disequilibrium information for haplotype reconstruction and quantitative trait locus fine mapping. *Genetics* 184, 789–798. doi: 10.1534/genetics.109.108431

Duffy, G., D'Arcy, S., Ahsan, T., Nerem, R., O'Brien, T., and Barry, F. (2010). Mesenchymal stem cells overexpressing ephrin-B2 rapidly adopt an early endothelial phenotype with simultaneous reduction of osteogenic potential. *Tissue Eng. A* 16, 2755–2768. doi: 10.1089/ten.tea.2009.0623

Dunn, L., Midwinter, R., Ni, J., Hamid, H., Parish, C., and Stocker, R. (2013). New insights into intracellular locations and functions of heme oxygenase-1. *Antioxid. Redox Signal.* 20, 1723–1742. doi: 10.1089/ars.2013.5675

Dupuis, M.-C., Zhang, Z., Druet, T., Denoix, J.-M., Charlier, C., Lekeux, P., et al. (2011). Results of a haplotype-based GWAS for recurrent laryngeal neuropathy in the horse. *Mamm. Genome* 22, 613–620. doi: 10.1007/s00335-011-9337-3

Eklom, K., and Lill, A. (2006). Development of parameters influencing blood oxygen-carrying capacity in nestling doves. *EMU* 106, 283–288. doi: 10.1071/MU05061

Excoffier, L., and Slatkin, M. (1995). Maximum-likelihood estimation of molecular haplotype frequencies in a diploid population. *Mol. Biol. Evol.* 12, 921–927. doi: 10.1093/oxfordjournals.molbev.a040269

Fang, L., Sahana, G., Su, G., Yu, Y., Zhang, S., Lund, M., et al. (2017). Integrating Sequence-based GWAS and RNA-Seq provides novel insights into the genetic basis of mastitis and milk production in dairy cattle. *Sci. Rep.* 7:45560. doi: 10.1038/srep45560

Favier, J., Kempf, H., Corvol, P., and Gasc, J. M. (2001). Coexpression of endothelial PAS protein 1 with essential angiogenic factors suggests its involvement in human vascular development. *Dev. Dyn.* 222, 377–388. doi: 10.1002/dvdy.1207

Follo, M., Mongiorgi, S., Clissa, C., Paolini, S., Giovanni, M., Martelli, A., et al. (2012). Activation of nuclear inositide signalling pathways during erythropoietin therapy in low-risk MDS patients. *Leukemia* 26, 2474–2482. doi: 10.1038/leu.2012.133

Gabriel, S., Schaffner, S., Nguyen, H., Moore, J., Roy, J., Blumenstiel, B., et al. (2002). The Structure of haplotype blocks in the human genome. *Science* 296, 2225–2229. doi: 10.1126/science.1069424

Go, Y., Bailey, E., Cook, F., Timoney, P., Coleman, S., MacLeod, J., et al. (2012). Genome-wide association study (GWAS) among four horse breeds identifies a common haplotype in equine chromosome 11 (ECA11) associated with the in vitro CD3+ T cell susceptibility/resistance to equine arteritis virus infection. *J. Equine Vet. Sci.* 32, S74–S75. doi: 10.1016/j.jevs.2012.08.159

Graves, B. J., and Petersen, J. M. (1998). "Specificity within the ets family of transcription factors," in *Advances in Cancer Research*, eds G. F. Vande Woude, and G. Klein, (Cambridge, MA: Academic Press), 1–57. doi: 10.1016/s0065-230x(08)60738-1

Haase, V. (2013). Regulation of erythropoiesis by hypoxia-inducible factors. *Blood Rev.* 27, 41–53. doi: 10.1016/j.blre.2012.12.003

Hawken, R., Zhang, Y., Fortes, M., Collis, E., Barris, W., Corbet, N., et al. (2011). Genome-wide association studies of female reproduction in tropically adapted cattle. *J. Anim. Sci.* 90, 1398–1410. doi: 10.2527/jas.2011-4410

Hendrickson, S. (2013). A genome wide study of genetic adaptation to high altitude in feral Andean Horses of the páramo. *BMC Evol. Biol.* 13:273. doi: 10.1186/1471-2148-13-273

Jakeways, N., McKeever, T., Lewis, S. A., Weiss, S., and Britton, J. (2003). Relationship between FEV1 reduction and respiratory symptoms in the general population. *Eur. Respir. J.* 21, 658–663. doi: 10.1183/09031936.03.00069603

Kickler, T. (2005). Hematology: basic principles and practice, 4th edition. *Transfusion* 45, 1233–1233. doi: 10.1111/j.1537-2995.2005.00192.x

Kominakis, A., Hager-Theodorides, A., Zoidis, E., Saridaki, A., Antonakos, G., and Tsiamis, G. (2017). Combined GWAS and 'guilt by association'-based prioritization analysis identifies functional candidate genes for body size in sheep. *Genet. Select. Evol.* 49:41. doi: 10.1186/s12711-017-0316-3

Li, Y. J., He, X. Q., Schembri-King, J., Jakes, S., and Hayashi, J. (2000). Cloning and characterization of human Lnk, an adaptor protein with pleckstrin homology and src homology 2 domains that can inhibit T cell activation. *J. Immunol.* 164, 5199–5206. doi: 10.4049/jimmunol.164.10.5199

Lockett, G., and Holloway, J. (2013). Genome-wide association studies in asthma; perhaps, the end of the beginning. *Curr. Opin. Allergy Clin. Immunol.* 13, 463–469. doi: 10.1097/ACI.0b013e328364ea5f

McCulloch, C. E., and Neuhaus, J. M. (2015). *Generalized Linear Mixed Models.* Hoboken, NJ: Wiley, 845–852.

McDaneld, T., Kuehn, L., Thomas, M., Snelling, W., Smith, T., Pollak, E., et al. (2014). Genomewide association study of reproductive efficiency in female cattle. *J. Anim. Sci.* 92, 1945–1957. doi: 10.2527/jas.2012-6807

Müller, M., and Brem, G. (1991). Disease resistance in farm animals. *Experientia* 47, 923–934.

Otterbein, L., Soares, M., Yamashita, K., and Bach, F. (2003). Heme oxygenase-1: unleashing the protective properties of heme. *Trends Immunol.* 24, 449–455. doi: 10.1016/S1471-4906(03)00181-9

Peng, Y., Cui, C., He, Y., Ouzhuluobu, Zhang, H., Yang, D., et al. (2017). Down-Regulation of EPAS1 transcription and genetic adaptation of tibetans to high-altitude hypoxia. *Mol. Biol. Evol.* 34, 818–830.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M., Bender, D., et al. (2007). Plink: a tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81, 559–575. doi: 10.1086/519795

Qiu, Q., Zhang, G., Ma, T., Qian, W., Wang, J., Ye, Z., et al. (2012). The yak genome and adaptation to life at high altitude. *Nat. Genet.* 44, 946–949. doi: 10.1038/ng.2343

Reya, T. (2003). Regulation of hematopoietic stem cell self-renewal. *Recent Prog. Hormone Res.* 58, 283–295. doi: 10.1210/rp.58.1.283

Simonson, T., Yang, Y., Huff, C., Yun, H., Qin, G., Witherspoon, D., et al. (2010). Genetic evidence for high-altitude adaptation in tibet. *Science* 329, 72–75. doi: 10.1126/science.1189406

Soler Artigas, M., Loth, D., Wain, L., Gharib, S., Obeidat, M. E., Tang, W., et al. (2011). Genome-wide association and large-scale follow up identifies 16 new loci influencing lung function. *Nat. Genet.* 43, 1082–1090.

Spyropoulos, D., Pharr, P., Lavenburg, K., Jackers, P., Papas, T., Ogawa, M., et al. (2000). Hemorrhage, impaired hematopoiesis, and lethality in mouse embryos carrying a targeted disruption of the fli1 transcription factor. *Mol. Cell. Biol.* 20, 5643–5652. doi: 10.1128/MCB.20.15.5643-5652.2000

Suenobu, S., Takakura, N., Inada, T., Yamada, Y., Yuasa, H., Zhang, X., et al. (2002). A role of EphB4 receptor and its ligand, ephrin-B2, in erythropoiesis. *Biochem. Biophys. Res. Commun.* 293, 1124–1131. doi: 10.1016/S0006-291X(02)00330-3

Takaki, S., Sauer, K., Iritani, B., Chien, S., Ebihara, Y., Tsuji, K.-I., et al. (2000). Control of B cell production by the adaptor protein Lnk: definition of a conserved family of signal-modulating proteins. *Immunity* 13, 599–609. doi: 10.1016/S1074-7613(00)00060-1

Till, J. E., McCulloch, E. A., and La, S. (1964). A stochastic model of stem cell proliferation, based on the growth of spleen colony-forming cells. *Proc. Natl. Acad. Sci. U.S.A.* 51, 29–36. doi: 10.1073/pnas.51.1.29

Tripette, J., Alexy, T., Hardy-Dessources, M. D., Mougenel, D., and Connes, P. (2009). Red blood cell aggregation, aggregate strength and oxygen transport potential of blood are abnormal in both homozygous sickle cell anemia and sickle-hemoglobin C disease. *Haematologica* 94, 1060–1065. doi: 10.3324/haematol.2008.005371

Verstockt, B., Smith, K., and Lee, J. (2018). Genome-wide association studies in Crohn's disease: past, present and future. *Clin. Transl. Immunol.* 7:e1001. doi: 10.1002/cti2.1001

von tempelhoff, G.-F., Schelkunov, O., Demirhan, A., Tsikouras, P., Rath, W., Velten, E., et al. (2015). Correlation between blood rheological properties and red blood cell Indices(MCH, MCV, MCHC) in healthy women. *Clin. Hemorheol. Microcircul.* 62, 45–54. doi: 10.3233/CH-151944

Wang, K., Liu, Y., Xu, Q., Liu, C., and Fang, M. (2017). A post-GWAS confirming GPAT3 gene associated with pig growth and a significant SNP influencing its promoter activity. *Anim. Genet.* 48, 478–482. doi: 10.1111/age.12567

Wang, Z., Zhang, H., Yang, H., Wang, S., Rong, E., Pei, W., et al. (2014). Genome-wide association study for wool production traits in a chinese merino sheep population. *PLoS One* 9:e107101. doi: 10.1371/journal.pone.0107101

Wei, C., Huihua, W., Liu, G., Zhao, F., Kijas, J., Ma, Y., et al. (2016). Genome-wide analysis reveals adaptation to high altitudes in Tibetan sheep. *Sci. Rep.* 6:26770. doi: 10.1038/srep26770

Wu, B., Chen, K., Shrestha, S., Ong, K. L., Barter, P., and Rye, K.-A. (2012). High density lipoproteins inhibit vascular endothelial inflammation by increasing 3β-hydroxysterol-{delta}24 reductase expression and inducing heme oxygenase-1. *Circ. Res.* 112, 278–288. doi: 10.1161/CIRCRESAHA.111.300104

Wu, Y., Fan, H., Wang, Y., Zhang, L., Gao, X., Chen, Y., et al. (2014). Genome-wide association studies using haplotypes and individual SNPs in simmental cattle. *PLoS One* 9:e109330. doi: 10.1371/journal.pone.0109330

Xiaoming, M., Jia, C., Fu, D., Chu, M., Ding, X., Wu, X., et al. (2019). Analysis of hematological traits in polled yak by genome-wide association studies using individual SNPs and haplotypes. *Genes* 10:463. doi: 10.3390/genes10060463

Yang, D., Peng, Y., Ouzhuluobu, Bianbazhuoma, Cui, C., Bianba, et al. (2016). HMOX2 functions as a modifier gene for high-altitude adaptation in tibetans. *Hum. Mutat.* 37, 216–223. doi: 10.1002/humu.22935

Yang, Q., Cui, J., Chazaro, I., Cupples, L., and Demissie, S. (2006). Power and type I error rate of false discovery rate approaches in genome-wide associate studies. *BMC Genet.* 6(Suppl. 1):S134. doi: 10.1186/1471-2156-6-S1-S134

Zeder, M. A. (2008). Domestication and early agriculture in the Mediterranean Basin: origins, diffusion, and impact. *PNAS* 105, 11597–11604. doi: 10.1073/pnas.0801317105

Zhang, F., Zhang, Z., Yan, X., Chen, H., Zhang, W., Hong, Y., et al. (2014). Genome-wide association studies for hematological traits in Chinese Sutai pigs. *BMC Genet.* 15:41. doi: 10.1186/1471-2156-15-41

Zhang, H., Wu, C., Chamba, Y., and Ling, Y. (2007). Blood characteristics for high altitude adaptation in tibetan chickens. *Poultry Sci.* 86, 1384–1389. doi: 10.1093/ps/86.7.1384

Zhang, X., Nie, Q., and Xu, Z. (2013). Overview of genomic insights into chicken growth traits based on genome- wide association study and microrna regulation. *Curr. Genomics* 14, 137–146. doi: 10.2174/1389202913140201402 0006

Zhang, Z., Guillaume, F., Sartelet, A., Charlier, C., Georges, M., Farnir, F., et al. (2012). Ancestral haplotype-based association mapping with generalized linear mixed models accounting for stratification. *Bioinformatics* 28, 2467–2473. doi: 10.1093/bioinformatics/bts348

Zou, Z., Ren, J., Yan, X., Huang, X., Yang, S., Zhang, Z., et al. (2008). Quantitative trait loci for porcine baseline erythroid traits at three growth ages in a White Duroc × Erhualian F2 resource population. *Mamm. Genome* 19, 640–646. doi: 10.1007/s00335-008-9142-9