Research article

https://revistas.unal.edu.co/index.php/refame

Revista Facultad Nacional de Agronomía

Genome-wide association study on growth traits in Colombian Hair Sheep

Estudios de asociación en genoma completo para caracteres de crecimiento en Ovinos de Pelo Colombiano

UNIVERSIDAD NACIONAL DE COLOMBIA

https://doi.org/10.15446/rfnam.v77n1.105408

Yineth Alexandra Palacios Erazo[1]*, Manuel Fernando Ariza Botero[2], Moris de Jesús Bustamante Yánez[3], Óscar David Vergara Garay[3] and Luz Ángela Álvarez Franco[1]

Keywords:
Creole sheep
GWAS
Live weight
Microarray
SNPs
Weight gain

ABSTRACT

The Colombian hair sheep have characteristics of great interest, among the following: high capacity for adaptation, good fertility, high prolifically, and low presence of diseases, which have been little studied. Currently, genome-wide association studies (GWAS) have been widely used to detect and locate candidate genes. However, in sheep, there is a low number of investigations carried out in GWAS, because the available information is limited, compared to that of other species. This research aimed to conduct a genome-wide association study on muscle growth traits using the Illumina OvineSNPs50 BeadChip array. A GWAS using 54.241 single nucleotide polymorphisms (SNPs) was conducted in Ethiopian (44 individuals), Sudan (63), and Pelibuey (60) breeds of Creole hair sheep to evaluate eight growth traits. Quality control was performed using a linear regression model in PLINK. Moreover, a functional analysis was done in the KEGG database using the *Ovis aries* (sheep) genome v.3.1. In total, 44.396 SNPs that passed quality control were used for the analysis. The 10 most significant SNPs were identified for each trait. The functional analysis allowed the annotating of four candidate genes, namely CEP135, EMCN, PAM, and PIAS2, as the most relevant genes for the traits assessed. Additionally, 27 genes associated with phenotypic traits were considered promising and could also be influencing growth traits. This is the first GWAS on Colombian hair sheep to report genomic traits associated with muscle growth traits. Four candidate genes (CEP135, EMCN, PAM, and PIAS2) associated with eight growth traits were identified by genome-wide association in colombian hair sheep.

Palabras clave:
Ovino criollo
GWAS
Peso vivo
Microarreglo
SNPs
Ganancia de peso

RESUMEN

Los Ovinos de Pelo Colombiano tienen características de gran interés, entre ellas se resaltan: alta capacidad de adaptación, buena fertilidad, alta prolificidad y baja presencia de enfermedades que han sido poco estudiadas. Actualmente, los estudios de asociación de genoma completo (GWAS) han sido ampliamente usados para detectar y localizar genes candidatos. Sin embargo, en ovinos se cuenta con un bajo número de investigaciones al respecto, debido a que la información disponible se encuentra limitada, en comparación con la de otras especies. El objetivo de esta investigación fue realizar un análisis genómico asociado a caracteres de crecimiento muscular, mediante el uso del microarreglo OvineSNP50 BeadChip de Illumina. Se realizó un GWAS empleando 54,241 polimorfismos de nucleótido simple (SNPs), en ovinos de pelo criollo, en el cual se evaluaron las variedades raciales Etíope (44), Sudán (63) y Pelibuey (60), para ocho caracteres de crecimiento. Se llevó a cabo un control de calidad a través del programa PLINK, por medio de un modelo de regresión lineal. Posteriormente se hizo un análisis funcional, empleando el genoma ovino *Ovis aries* v.3.1, en la base de datos de KEGG. Después de haber realizado el control de calidad, se analizaron 44,396 SNP. Se identificaron los 10 SNPs más significativos para cada carácter. A través del análisis funcional se logró anotar cuatro posibles genes candidatos (CEP135, EMCN, PAM y PIAS2), como los más importantes en los caracteres evaluados. Adicionalmente, se encontraron 27 genes asociados a los caracteres fenotípicos, los cuales pueden ser prometedores y podrían estar influyendo en los caracteres de crecimiento. Este trabajo es el primer GWAS en ovinos de pelo colombiano en el país, que ha asociado caracteres fenotípicos de crecimiento muscular con caracteres genómicos. El GWAS permitió identificar cuatro posibles genes candidatos (CEP135, EMCN, PAM y PIAS2), asociados a ocho caracteres de crecimiento.

[1]Facultad de Ciencias Agropecuarias, Universidad Nacional de Colombia. Grupo de Investigación en Recursos Zoogenéticos. Palmira, Colombia. ypalaciose@unal.edu.co, laalvarezf@unal.edu.co

[2]Facultad de Medicina Veterinaria y Zootecnia, Universidad Nacional de Colombia. Grupo de Investigación en Genética Molecular Animal. Bogotá, Colombia. mfarizab@unal.edu.co

[3]Facultad de Medicina Veterinaria y Zootecnia, Universidad de Córdoba. Grupo de Investigación en Producción Animal Tropical. Montería, Colombia. morisbustamante@correo.unicordoba.edu.co, overgara@correo.unicordoba.edu.co

*Corresponding author

Received: September 10, 2023; Accepted: October 30, 2023

Rev. Fac. Nac. Agron. Medellín 77(1): 10625-10635. 2024

ISSN 0304-2847 / e-ISSN 2248-7026

Sheep are an important component of the environment based on economic and conservation perspectives (Knapik et al. 2017). These animals can adapt to landscapes that are difficult to cultivate and play an important role in food security, household economies, and cultural development of Colombian indigenous and farmer communities (Corpoica 2007). Traditionally, in the country, Colombian hair sheep have been used for meat production in extensive systems (Zuleta et al. 2009). These animals are highly diverse since crosses between different breeds date back to the colonization (Vergara et al. 2017). There are two varieties of sheep, which producers refer to as Ethiopian, or cherry red-haired sheep, and Sudan, or yellow-haired sheep; however, polychromy is common among sheep, with animals found in cherry red, yellow, white, black, and combinations of the above, due to the introduction of animals from West Africa, Ethiopia, Aruba, Curacao, and other Caribbean islands (Zuleta et al. 2009). Furthermore, in Colombia, exotic breeds are also used and crossed with Creole breeds to obtain hybrid vigor for meat production (Vergara-Garay et al. 2016).

The increasing popularity of sheep meat production worldwide has prompted the sheep industry, producers, and geneticists to address the importance of growth and meat production in this species (Zuleta et al. 2009). Growth variables, such as body weight at different stages (Al-Mamun et al. 2015; Gholizadeh et al. 2015), have been evaluated as important growth indicators (Kemper et al. 2012). Given the low productive parameters in Colombian hair sheep (Romero et al. 2002), strategies to understand and improve these indicators must be implemented. In this regard, genomics-based approaches represent a valuable tool (Casas and White 2015).

The development of high-throughput SNP (single nucleotide polymorphism) genotyping has enabled the development of GWAS to identify candidate genes for traits of interest (Al-Mamun et al. 2015; Gholizadeh et al. 2015; Kominakis et al. 2017). Despite the discovery of various candidate genes through GWAS in recent years, a small number of these have been found in sheep due to the limited availability of genomic information on sheep (Zhang et al. 2013). Furthermore, most research has been conducted on exotic breeds, while studies on native sheep are scarce (Casas and White 2015). GWAS are expected to contribute to identifying SNPs associated with traits of interest; therefore, this study aimed to analyze associations between growth traits and genomic traits using the Illumina OvineSNPs50 BeadChip array.

MATERIALS AND METHODS

Study animals

The study population comprised 167 individuals of Colombian hair sheep (females and males) belonging to Sudan (OPC_S), Pelibuey (OPC_P), and Ethiopian (OPC_E) breeds. The number of individuals from each breed was 63, 60, and 44, respectively. The animals were in the departments of Cesar, Cordoba, and Valle del Cauca. The technical, scientific, and administrative procedures for research with animals, were adjusted to the regulations of Law 84 of 1989 and with resolution 8430 of 1993 (National Congress of Colombia).

Growth traits

The phenotypic growth traits assessed were birth weight (BW) measured during the first 24, adjusted weaning weight to 90 days (AWW) (equation 1), adjusted final weight adjusted to 1-year-old (AFW) (equation 2), pre-weaning daily weight gain (WGPRE) and post-weaning daily weight gain (WGPOS), fat thickness (FT), loin depth (LD), and loin eye area (LEA). The last three traits were measured at approximately 1 year of age, through a MyLab^{TM}One VET ultrasonograph (Esaote, Maastricht, Netherlands).

$$AWW = \left[\frac{(WW - BW}{WA}\right](100) + BW \qquad (1)$$

Where AWW is adjusted weaning weight to 90 days, WW is weaning weight, BW is birth weight, and WA is weaning age (days).

$$AFW = \left[\frac{(FW - WW}{AFWD}\right](365) + WW \qquad (2)$$

Where AFW is adjusted final weight adjusted to 1-year-old, FW is final weight, WW is weaning weight, and AFWD is the age at the final weight (days).

Samples and genotyping

Blood samples (10 mL from each animal) were obtained through jugular venipuncture and collected in vacutainer tubes with EDTA. Genomic DNA was extracted from the

blood samples using Gene JET Genomic DNA purification (Thermo Fisher Scientific). The DNA was quantified through spectrophotometry using NanoDrop® ND-1000 (Thermo Fisher Scientific). DNA genotyping was performed on an Illumina OvineSNPs50 BeadChip array, targeting 54,241 SNPs, according to the Infinium® Assay Super II Illumina® protocol for use on the HiScan®SQ System.

Quality control was performed in PLINK v1.9b5.2 (Purcell et al. 2007) by excluding individuals who were missing more than 5% of genotyping information, as well as SNPs with call rates below 95%, allele frequencies below 5%, and Hardy-Weinberg equilibrium $P<0.001$.

Genome association analysis

The association analysis between the genotype and eight phenotypic traits was based on a linear regression model (equation 3) using 44,396 filtered SNPs. The equation was adjusted according to the genomic location of the markers using Wald's standard statistic to calculate the *P* value for each SNP.

$$Y = X\beta + \in \quad (3)$$

Where Y is the vector of n phenotypic observations (BW, AWW, AFW, WGPRE, WGPOS, FT, LD, and LEA) for 167 individuals, X is a covariable matrix (n x p); where n are individuals and p are SNPs, β is a vector of p effects (SNPs), and ϵ is a vector of n residuals.

Post-GWAS analysis

Using Haploview v4.2 (Barrett et al. 2005), SNPs were filtered for each trait based on a probability value of $P<0.001$ to obtain a matrix of significantly associated SNPs, which was used in further analyses. Next, SNPs were classified according to *P*-value and correlation coefficient (R^2) using GWASTools (Gogarten et al. 2012) in R version 3.1 to identify the 10 most relevant SNPs associated with each phenotype studied.

Additionally, Genome Data Viewer was used to search the *Ovis aries* genome version 4 for growth genes, namely calpain (CAPN1), calpastatin (CAST), growth hormone (GH), growth hormone receptor (GHR), leptin (LEP), myostatin (MSTN), type 1 insulin-like growth factor (IGF-1), fatty acid-binding protein 9 (FABP9), and fatty acid-binding protein 4 (FABP4). SNPs located in these genes were identified for each trait assessed. On the other hand, a linkage disequilibrium value was established with a LD=0.001, under a window of 50,000 bp, using the SNPRelate 0.9.12 package (Zheng et al. 2012) from Bioconductor in the R program, to establish a set of filtered SNPs in equilibrium to avoid the strong influence of the groups of SNPs in the principal component analysis (PCA).

Functional analysis

The most significant genomic associations were searched in the KEGG database using *Ovis aries* genome v.3.1 in ReactomePA Bioconductor package (Yu and He 2016) in R to infer the biological processes of the genes involved. Furthermore, the DAVID database was used to obtain biological functions from additional databases, such as Gene Ontology, UniProt Keywords, InterPro, Kegg Pathway, Smart, and Pir SuperFamily. Finally, the KEGG identifiers retrieved were mapped to metabolic pathways in *Ovis aries*.

RESULTS AND DISCUSSION

In this research, a GWAS was conducted in 167 animals belonging to three breeds of Colombian hair sheep to determine associations between eight phenotypic traits and a panel of 54,241 SNPs using next-generation sequencing. The sex chromosome (OAR) and mitochondrial-associated contig were excluded from the analysis; furthermore, 44,396 SNPs were retained after discarding missing data for individuals and SNPs, low allele frequencies, and Hardy-Weinberg equilibrium (i.e., quality control).

For the association analysis, the SNPs associated with each trait were filtered according to $P<0.001$; as a result, 7,000 significant SNPs were identified (Table 1). After ordering by lowest *P*-value and highest R^2, the 10 most significant SNPs showing the highest association with each of the eight traits were determined (Tables 2, 3, and 4) (Zhang et al. 2013), which resulted in a total of 80 SNPs for all traits. This study is the first approach in Colombia to identify candidate functional genes containing SNPs that are significantly associated with growth traits; furthermore, these SNPs may be involved in phenotypic determination. Despite the small sample size for a GWAS (Zhang et al. 2013; Peng et al. 2017), more than 10,000 SNPs were found to be significantly associated with eight growth traits (e.g., BW, AWW, WGPRE, AFW, and WGPOS), as well as traits that were

measured by ultrasonography (e.g., FT, LD, and LEA). Moreover, this research represents a pioneer study in Colombia, thereby, contributing essential information to the ovine sector in the country.

Table 1. Number of significant SNPs for each of the phenotypic traits assessed.

Phenotypic trait	SNPs*
BW	7437
AWW	1058
AFW	2872
WGPRE	1761
WGPOS	1076
FT	599
LD	147
LEA	1134

*Number of significant SNPs (*P*<0.001) based on the linear regression model.

The analysis indicated that chromosome OAR6 showed most significant associations (*P*<0.001), accounting for 12.5% of SNPs associated with AWW, WGPRE, AFW, WGPOS, FT, and LEA, followed by chromosome OAR1 with 11% of SNPs for AWW, WGPRE, AFW, FT, and LEA, and chromosome OAR5 with 9% of SNPs for BW, AWW, AFW, and WGPOS. Similarly, Zhang et al. (2013), suggested that chromosomes OAR1 and OAR3 are important for growth traits and meat production in sheep (i.e., for pre-weaning and post-weaning weight gain, daily weight gain, thoracic perimeter, shin circumference, weight at six months old, and weaning weight).

Association analysis for pre-weaning growth traits

Table 2 shows the 10 most significant SNPs associated with each pre-weaning phenotypic trait: BW, AWW, and WGPRE. For BW, annotation showed that 50% of SNPs were found in introns, 40% in the distal region, and 10% in promoters. Furthermore, six out of 10 SNPs were in five genes; however, functions were inferred for only three of these genes: LOC101107266, PAM, and GABRG2.

Association analysis for post-weaning growth traits

The 10 most significant SNPs associated with phenotypic traits AFW and WGPOS are found in Table 3. For AFW, 70% of SNPs were annotated in the distal region and 30% in introns. Five out of 10 SNPs were located in CEP135, EMCN, PRDM13, PIAS2, and SLC44A3 genes. Moreover, for WGPOS, 60% of SNPs were found in the distal region and 40% in introns. Six out of 10 SNPs were located in EFNA5, MAP3K7, CEP135, EMCN, PIAS2, and MTAP genes.

Association analysis for growth traits measured by ultrasonography

The 10 most significant SNPs associated with traits FT, LD, and LEA are shown in Table 4. For phenotypic trait FT, 80% of SNPs were annotated in the distal region, 10% in intron 2 of 20, and 10% were located downstream (1-2 Kb). Only one out of ten SNPs were located in a known gene, namely CEP135. Moreover, for LD, 80% of SNPs were annotated in the distal region and 20% in introns. Furthermore, four out of 10 SNPs were located in SUGP1, LOC101104424, PDYN, and TGFB3 genes. Finally, for LEA, 40% of SNPs were found in the distal region, 30% in introns, 20% in promoters, and 10% were located downstream. Five out of 10 SNPs were found in INPP4B, TBRG4, C17H4orf46, DDX24, and KIRREL genes.

Genomic location and importance of the most relevant SNPs

By annotating the 80 most significantly associated SNPs and mapping them to the chromosomes of the *Ovis aries* genome v3.1, eight relevant SNPs were identified on three chromosomes, namely OAR5, OAR6, and OAR23. In chromosome 5, SNPs OAR5_107968125.1 and OAR5_107977075.1 were associated with BW and located in the PAM gene (positions 99153391 bp and 99165665 bp, respectively). Furthermore, in chromosome 6, SNPs OAR6_77919148.1 (position 71481367 bp) and

Table 2. SNPs are significantly associated with BW, AWW, and WGPRE in Colombian hair sheep.

Trait	Chromosome	SNPs	Position (bp)	R^2	P	Annotation	Ensembl Gene ID	NCBI Gene
BW	3	OAR3_61737307.1	58315054	0.441	$7.18x10^{-18}$	Intron (ENSOART00000022552/ENSOARG00000020700, intron 2 of 3)	ENSOARG00000020700	LOC 101107266
BW	5	OAR5_107968125.1	99153391	0.352	$3.00x10^{-14}$	Intron (ENSOART00000020038/ENSOARG00000018405, intron 3 of 25)	ENSOARG00000018405	PAM
BW	5	OAR5_107977075.1	99165665	0.352	$3.00x10^{-14}$	Intron (ENSOART00000020038/ENSOARG00000018405, intron 3 of 25)	ENSOARG00000018405	PAM
BW	5	s11274.1	71514994	0.340	$1.28x10^{-13}$	Intergenic distal	ENSOARG00000014001	GABRG2
BW	3	OAR3_89348294.1	84390370	0.322	$1.36x10^{-12}$	Intron (ENSOART00000008654/ENSOARG00000007947, intron 4 of 4)	ENSOARG00000007947	TMEM178A
BW	17	OAR17_29260298.1	26667946	0.322	$1.36x10^{-12}$	Intergenic distal	ENSOARG00000014791	-
BW	12	OAR12_20575087.1	17694724	0.309	$1.53x10^{-12}$	Intron (ENSOART00000011686/ENSOARG00000010714, intron 17 of 74)	ENSOARG00000010714	USH2A
BW	2	OAR2_51716542.1	48200678	0.298	$2.72x10^{-11}$	Intergenic distal	ENSOARG00000021824	-
BW	3	OAR3_74952313.1	70926536	0.294	$3.84x10^{-11}$	Intergenic distal	ENSOARG00000004203	-
BW	18	s03219.1	68137231	0.288	$7.96x10^{-11}$	Promotor (<=1Kb)	ENSOARG00000013958	-
AWW	14	s01263.1	58489098	0.190	$3.79x10^{-06}$	Intergenic distal	ENSOARG00000000628	LOC101113879
AWW	20	OAR20_49893668.1	45829455	0.159	$1.06x10^{-04}$	Intron (ENSOART00000028931/ENSOARG00000026887, intron 1 of 1)	ENSOARG00000026887	-
AWW	1	OAR1_23734999.1	23651580	0.159	$1.06x10^{-04}$	Intergenic distal	ENSOARG00000004178	LOC105605154
AWW	9	s33129.1	88576469	0.156	$1.26x10^{-04}$	Intergenic distal	ENSOARG00000011632	CNGB3
AWW	5	OAR5_49721052.1	45617625	0.146	$3.65x10^{-07}$	Intergenic distal	ENSOARG00000025310	-
AWW	1	OAR1_189179554.1	175482796	0.144	$4.11x10^{-04}$	Promotor (<=1Kb)	ENSOARG00000019298	CD200
AWW	6	OAR6_12248234.1	9785894	0.143	$5.16x10^{-04}$	Intergenic distal	ENSOARG00000017991	-
AWW	26	DU261801_281.1	37119640	0.139	$6.78x10^{-04}$	Intron (ENSOART00000004427/ENSOARG00000004073, intron 2 of 15)	ENSOARG00000004073	PSD3
AWW	1	s25125.1	96847814	0.139	$6.82x10^{-04}$	Intron (ENSOART00000022309/ENSOARG00000020482, intron 4 of 46)	ENSOARG00000020482	POF4DIP
AWW	10	OAR10_59207797.1	59207797	0.138	$8.01x10^{-07}$	Intergenic distal	ENSOARG00000017151	SLITRK1
WGPRE	1	s12060.1	75971485	0.236	$3.25x10^{-08}$	Intergenic distal	ENSOARG00000017670	PLPPR4
WGPRE	23	OAR23_39144674_X.1	39144675	0.225	$9.06x10^{-08}$	Promotor (<=1Kb)	ENSOARG00000026187	-
WGPRE	4	OAR4_24114290.1	24114290	0.205	$8.60x10^{-07}$	Intron (ENSOART00000009008/ENSOARG00000008280, Intron 7 of 7)	ENSOARG00000008280	-
WGPRE	10	OAR10_91128145.1	91128145	0.205	$8.86x10^{-10}$	Intergenic distal	ENSOARG00000009132	CHAMP1
WGPRE	7	OAR7_85269064.1	85269064	0.196	$2.07x10^{-06}$	Promotor (2-3Kb)	ENSOARG00000002286	CIPC
WGPRE	8	OAR8_1452721.1	1452721	0.197	$2.13x10^{-06}$	Intergenic distal	ENSOARG00000006317	-
WGPRE	19	s52415.1	11201714	0.193	$2.86x10^{-06}$	Intron (ENSOART00000000540/ENSOARG00000000498, Intron 27 of 27)	ENSOARG00000021896	-
WGPRE	6	OAR6_77919148.1	77919148	0.192	$3.26x10^{-06}$	Intergenic distal	ENSOARG00000026595	-
WGPRE	18	OAR18_58488706.1	58488706	0.180	$1.16x10^{-05}$	Intergenic distal	ENSOARG00000024407	-
WGPRE	3	s62226.1	26419161	0.180	$1.21x10^{-05}$	Promotor (<=1Kb)	-	LOC105601856

OAR6_27552838.1 (position 24157625 bp) were located in CEP135 and EMCN genes, respectively, and were associated with AFW and WGPOS. In particular, the first of these SNP (OAR6_77919148.1) was also found associated with FT. Finally, in chromosome 23, SNP OAR23_49635171_X.1 (position 46823961 bp), located in the PIAS2 gene, was associated with traits AFW and WGPOS. Additionally, SNPs OAR5_112451694.1 and OAR8_50320412.1 were found in chromosomes OAR5 and OAR8. These SNPs were located in functional genes EFNA5 and MAP3K7, which interact in a metabolic pathway associated with WGPOS and are involved in cell signaling processes, according to the KEGG annotation.

Previous studies that have used the Illumina OvineSNPs50 BeadChip array reported several SNPs, including OAR6_41003295.1 and OAR6_42945420.1, that were found to be significantly associated with traits such as body weight (Al-Mamun et al. 2015). In this study, both of these SNPs were associated with BW. Moreover, Gholizadeh et al. (2015), reported that SNP OAR16_46544413.1 was associated with weight at six months old; similarly, this SNP was associated with AWW in this study. Additionally, SNP s52984.1 has been associated with daily weight gain and weight at six months old, while SNPs s55067.1 and s34745.1 have been associated with weaning weight (Zhang et al. 2013). Furthermore, Ai-Mamun et al. (2015) reported that SNP OAR6_40409402.1 was associated with body weight. In this study, these four SNPs were associated with AFW. Moreover, SNPs s34745.1 and s55067.1, reported by Zhang et al. (2013), were associated with pre-weaning weight gain; similar to the results of this study that showed the association between these SNPs and WGPRE. Finally, OAR16_61248510.1 has been associated with post-weaning weight gain and s52984.1 with daily weight gain and weight at six months old (Zhang et al. 2013); in this case, both SNPs were associated with WGPOS.

Among the 80 most significant SNPs (Table 2, 3, and 4), those that were associated with more than one trait were identified as the most relevant SNPs. One SNP was located near two genes (CEP135 and EMCN), one in a single gene (PIAS2), and two SNPs in the same gene (PAM). Overall, four candidate genes were detected, which could be most influential on the growth traits assessed. However, 27 additional genes were also promising, including LOC101107266, GABRG2, TMEM178A, USH2A, LOC101113879, LOC105605154, CNGB3, CD200, PSD3, PDE4DIP, SLITRK1, PLPPR4, CHAMP1, CIPC, LOC105601856, PRDM13, SLC44A3, MTAP, SUGP1, LOC101104424, PDYN, TGFB3, INPP4B, TBRG4, C17H4orf46, DDX24, and KIRREL.

SNP OAR6_77919148.1 was located on the distal region of the CEP135 (centrosomal protein 135) gene, which belongs to the family of centrosomal proteins that are an active component of the centrosome and are involved in the biogenesis of the centriole and control of cell cycle progression (Kumar et al. 2013). In humans, alterations in this gene cause reduced cell growth rates (Hussain et al. 2012). Furthermore, in *Chlamydomonas*, mutation of Bld10, an ortholog of CEP135, resulted in abnormal microtubules in interphase and the mitotic spindle; these defects occurred during cell division and significantly reduced cell growth rate (Matsuura et al. 2004). Based on this, the CEP135 gene is involved in cell growth rate and can influence AFW, WGPOS, and FT.

SNP OAR6_27552838.1 is located in the distal intergenic region of EMCN (Endomucin), which encodes a sialic-rich glycoprotein rich in type I O-glycoproteins and is specifically expressed in the venous and capillary endothelium. Recent studies suggest that this gene is a potent regulator of capillary formation from existing blood vessels, or angiogenesis (Park-Windhol et al. 2017). Furthermore, EMCN plays a key role in tissue damage and wound repair, the endometrial cycle, and the adaptation of the striated muscle to stress and exercise (Olfert et al. 2015). Therefore, there is a functional relationship between this gene and the traits assessed given the involvement of angiogenesis in the skeletal muscle and the association found here between EMCN and AFW and WGPOS.

SNPs OAR5_107968125.1 and OAR5_107977075.1 are located in introns of the PAM gene (Peptidylglycine Alpha-Amidating Monooxygenase), which catalyzes COOH-terminal amidation of peptide hormones (Traci et al. 2005). In mammals, amidation activity is presumed to be coded by a single gene product with two catalytic domains (PHM and PAL) that act sequentially to amidated peptides (Prigge et al. 2000). Amide peptides function as hormones, neuromodulators, and autocrine growth factors (Prigge et al. 2000; Merkler 1994). Therefore, this gene could be associated with BW.

Table 3. SNPs significantly associated with AFW and WGPOS in Colombian hair sheep.

Trait	Chromosome	SNPs	Position (pb)	R^2	*P*	Annotation	Ensembl Gene ID	NCBI Gene
AFW	6	OAR6_77919148.1	71481367	0.275	$3.66x10^{10}$	Intergenic distal	ENSOARG00000003002	CEP135
AFW	23	OAR23_39144674_X.1	37003177	0.273	$4.58x10^{10}$	Intergenic distal	ENSOARG00000022421	-
AFW	6	OAR6_27552838.1	24157625	0.256	$3.10x10^{-09}$	Intergenic distal	ENSOARG00000013694	EMCN
AFW	8	OAR8_39977285.1	37211967	0.224	$1.03x10^{-10}$	Intergenic distal	ENSOARG00000011775	PRDM13
AFW	5	OAR5_84496122_X.1	76876928	0.223	$1.24x10^{-07}$	Intergenic distal	ENSOARG00000014307	-
AFW	23	OAR23_49635171_X.1	46823961	0.223	$1.32x10^{-07}$	Intron (ENSOART00000003661/ENSOARG00000003371, intron 1 of 13)	ENSOARG00000003371	PIAS2
AFW	13	OAR13_26675629.1	24142711	0.113	$4.13x10^{-07}$	Intron (ENSOART00000003020/ENSOARG00000002785, intron 2 of 20)	ENSOARG00000021746	-
AFW	7	OAR7_104041756.1	95561909	0.211	$4.41x10^{-07}$	Intergenic distal	ENSOARG00000026731	-
AFW	7	OAR7_104050413.1	95569312	0.211	$4.41x10^{-07}$	Intron (ENSOART00000028761/ENSOARG00000026731, intron 2 of 2)	ENSOARG00000026731	-
AFW	1	s12060.1	71057388	0.210	$5.18x10^{-07}$	Intergenic distal	ENSOARG00000017408	SLC44A3
WGPOS	5	OAR5_112451694.1	103304973	0.180	$1.07x10^{-05}$	Intergenic distal	ENSOARG00000018761	EFNA5
WGPOS	20	s24831.1	11418394	0.176	$1.65x10^{-05}$	Intron (ENSOART00000016003/ENSOARG00000014698, Intron 17 of 22)	ENSOARG00000014906	-
WGPOS	5	OAR5_84496122_X.1	76876928	0.168	$3.70x10^{-05}$	Intergenic distal	ENSOARG00000014307	-
WGPOS	10	OAR10_1950751.1	4132402	0.167	$4.20x10^{-05}$	Intergenic distal	ENSOARG00000023450	-
WGPOS	8	OAR8_50320412.1	46840220	0.165	$5.09x10^{-05}$	Intron (ENSOART00000013401/ENSOARG00000012321, Intron 7 of 16)	ENSOARG00000012321	MAP3K7
WGPOS	6	OAR6_77919148.1	71481367	0.161	$7.30x10^{-05}$	Intergenic distal	ENSOARG00000003002	CEP135
WGPOS	6	OAR6_27552838.1	24157625	0.161	$7.52x10^{-05}$	Intergenic distal	ENSOARG00000013694	EMCN
WGPOS	23	OAR23_49635171_X.1	46823961	0.161	$8.41x10^{-05}$	Intron (ENSOART00000003661/ENSOARG00000003371, Intron 1 of 13)	ENSOARG00000003371	PIAS2
WGPOS	2	OAR2_96008804.1	89509922	0.159	$1.04x10^{-04}$	Intergenic distal	ENSOARG00000014411	MTAP
WGPOS	6	s48525.1	22902181	0.156	$1.30x10^{-04}$	Intron (ENSOART00000014572/ENSOARG00000013398, Intron 6 of 14)	ENSOARG00000023387	-

Table 4 . SNPs significantly associated with FT, LD, and LEA in Colombian hair sheep.

Trait	Chromosome	SNPs	Position (bp)	R^2	*P*	Annotation	Ensemble Gene ID	NCBI Gene
FT	6	s61364.1	16067116	0. 256	$3.15x10^{-09}$	Intergenic distal	ENSOARG00000006816	-
FT	6	OAR6_77919148.1	71481367	0.227	$7.28x10^{-08}$	Intergenic distal	ENSOARG00000003002	CEP135
FT	13	OAR13_26675629.1	24142711	0.195	$2.33x10^{-09}$	Intron (ENSOART00000003020/ ENSOARG00000002785, Intron 2 of 20)	ENSOARG00000021746	-
FT	6	OAR6_81305080.1	74488523	0.186	$5.88x10^{-06}$	Intergenic distal	ENSOARG00000024398	-
FT	1	OAR1_248575929.1	230581424	0.186	$6.58x10^{-06}$	Río abajo (1-2kb)	ENSOARG00000021579	-
FT	2	s37719.1	32379564	0.171	$2.7x10^{-08}$	Intergenic distal	ENSOARG00000023484	-
FT	9	OAR9_40619495.1	38690895	0.166	$5.16x10^{-05}$	Intergenic distal	ENSOARG00000026523	-
FT	2	s58833.1	101443025	0.166	$5.36x10^{-05}$	Intergenic distal	ENSOARG00000014922	-
FT	4	OAR4_6864529.1	7128966	0.163	$6.83x10^{-05}$	Intergenic distal	ENSOARG00000021732	-
FT	1	OAR1_157215006.1	145677767	0.157	$1.10x10^{-04}$	Intergenic distal	ENSOARG00000025579	-
LD	10	OAR10_42329325.1	42329325	0.232	$6.36x10^{-08}$	Intergenic distal	ENSOARG00000015005	-
LD	11	OAR11_7879331.1	7879331	0.168	$4.17x10^{-05}$	Intron (ENSOART00000008542/ ENSOARG00000007848, Intron 4 of 10)	ENSOARG00000007910	-
LD	2	OAR2_20446382.1	20446382	0.151	$2.14x10^{-04}$	Intergenic distal	ENSOARG00000025744	-
LD	5	s53828.1	3679286	0.142	$5.08x10^{-04}$	Intergenic distal	ENSOARG00000008152	SUGP1
LD	15	s40457.1	47387694	0.142	$5.21x10^{-04}$	Intergenic distal	ENSOARG00000007342	LOC101104424
LD	5	OAR5_110584928.1	110584928	0.136	$8.92x10^{-04}$	Intergenic distal	ENSOARG00000025340	-
LD	10	OAR10_68443350.1	68443350	0.127	$2.37x10^{-03}$	Intergenic distal	ENSOARG00000000563	-
LD	13	s27419.1	52439172	0.126	$2.54x10^{-03}$	Intergenic distal	ENSOARG00000007354	PDYN
LD	22	OAR22_31810853.1	31810853	0.122	$3.70x10^{-03}$	Intergenic distal	ENSOARG00000011700	-
LD	7	s52907.1	84077554	0.118	$5.52x10^{-06}$	Intron (ENSOART00000002182/ ENSOARG00000001992, Intron 28 of 32)	ENSOARG00000002050	TGFB3
LEA	14	s01263.1	55308253	0.197	$1.89x10^{-06}$	Intron (ENSOART00000014897/ ENSOARG00000013693, Intron 3 of 5)	ENSOARG00000013698	-
LEA	1	OAR1_248575929.1	230581424	0.191	$3.96x10^{-06}$	Río abajo (1-2kb)	ENSOARG00000021579	-
LEA	17	OAR17_16676148.1	15020629	0.165	$4.96x10^{-05}$	Intergenic distal	ENSOARG00000011798	INPP4B
LEA	4	s12760.1	76582310	0.163	$6.63x10^{-05}$	Promotor (1-2kb)	ENSOARG00000013838	TBRG4
LEA	6	OAR6_50229073_X.1	45303722	0.163	$7.07x10^{-05}$	Intergenic distal	ENSOARG00000008049	-
LEA	6	OAR6_12297748.1	9841765	0.160	$9.17x10^{-05}$	Intergenic distal	ENSOARG00000017991	-
LEA	17	OAR17_43310387.1	40123170	0.155	$1.36x10^{-04}$	Intron (ENSOART00000005160/ ENSOARG00000004733, Intron 18 of 20)	ENSOARG00000004417	C17H4orf46
LEA	19	OAR19_8201989.1	7898831	0.147	$3.19x10^{-04}$	Promotor (<=1kb)	ENSOARG00000026642	-
LEA	18	s50461.1	57567836	0.147	$3.54x10^{-04}$	Intron (ENSOART00000015616/ ENSOARG00000014341, Intron 3 of 7)	ENSOARG00000014341	DDX24
LEA	1	OAR1_260939703.1	106932088	0.145	$4.08x10^{-04}$	Intergenic distal	ENSOARG00000007174	KIRREL

OAR23_49635171_X.1 was found in an intron of the PIAS2 gene (Protein Inhibitor of Activated STAT 2), which belongs to the family of STAT protein inhibitors that regulate the activity of many proteins and influence diverse processes such as the immune response, cancer formation, and cell cycle progression. Furthermore, a study on *Xenopus* indicated that this gene family could play an important role in embryogenesis regulation (Burn et al. 2011). Therefore, according to the literature, PIAS2 is involved in early growth processes, and in this study, it was found in association with AFW and WGPOS, which were measured during growth and development stages and in the adult stage. These findings suggest that this gene is not only involved in early growth stages but also later periods; however, this was not conclusive.

The APC showed that the three sheep breeds were separated, the OPC_P (yellow) from the OPC_E (blue) and OPC_S (red), however, some OPC_S individuals overlapped with the OPC_E (Figure 1). Which could be used in the model to check the population stratification.

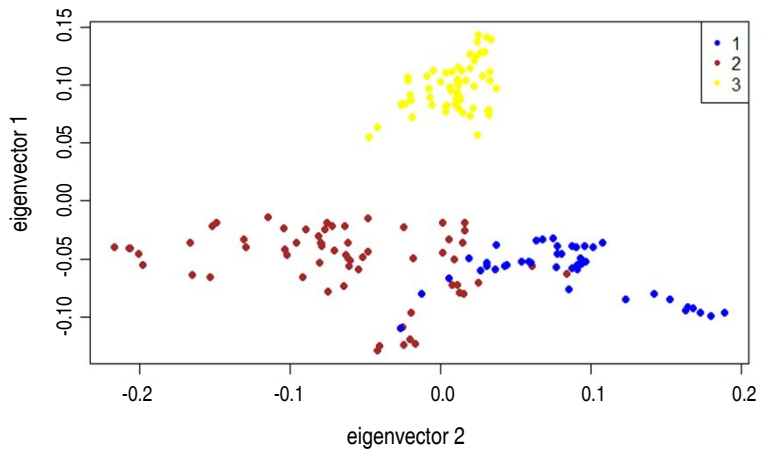

Figure 1. Principal component analysis for population stratification in three breeds of Colombian hair sheep.

MAP3K7 (Mitogen-Activated Protein Kinase Kinase Kinase 7) and EFNA5 (Efrin-A5) genes

MAP3K7 and EFNA5 genes are found in shared metabolic pathways and are involved in signaling processes. MAPK (Mitogen-Activated Protein Kinases) are members of the serine/threonine kinase protein family activated by growth and stress factors. These proteins play an important role in intracellular signal transduction, which enables the cell to integrate different extracellular signals. Therefore, MAPK participates in signaling cascades that regulate cell growth, differentiation, proliferation, and cell death (Roberts and Der 2007).

EFNA5 gene belongs to Eph receptors and their ligands (i.e., efrins), which constitute the largest family of tyrosine kinase receptors (Pasquale 2005; Kullander and Klein 2002). Additionally, this gene behaves as a chemotactic molecule that participates in the correct positioning and formation of the neuromuscular junction (Li and Johnson 2013). Accordingly, satellite cells are a resident population of stem cells of the adult skeletal muscle tissue, which is responsible for growth and regeneration. These cells generally aggregate near the ends of muscle fibers, as well as the neuromuscular junction. Li and Johnson (2013) examined the effects of EFNA5 signaling on satellite cells in bovines and found that chemokines and growth factors participate in the localization of these cells. MAP3K7 and EFNA5 genes participate in metabolic pathways involved in cellular processes; furthermore, both genes were associated with WGPOS, which was measured during the growth and development stage until the adult stage. Therefore, a coherent relationship between these genes and the growth trait was established.

Search for growth-associated genes reported in the literature

The search for CAPN1, CAST, GH, GHR, LEP, MSTN, IGF-1, FABP9, and FABP4 genes in the *Ovis aries* genome v.4.0

allowed identifying SNPs located in CAPN1, CAST, GHR, and FABP9. CAPN1 gene contained SNPs s30026.1, which showed association with traits BW, AFW, and FT (P<0.05). CAST gene contained two SNPs: OAR5_101792466.1 was significantly associated with traits BW, AFW, and FT, while s59216 was associated with AWW and LEA. Furthermore, there were five SNPs located in GHR, including three significantly associated SNPs, namely OAR16_34620156.1 for WGPOS, OAR16_34694443.1 for LD and LEA, and OAR16_34857607.1 for BW. Moreover, SNP OAR9_60512150.1, located in the FABP9 gene, was significantly associated with WGPRE and FT.

The annotation of the candidate genes indicated that several of these show some degree of association with cell growth, apoptosis, angiogenesis, and metabolic pathways that are directly or indirectly involved in muscle growth in different species. Therefore, several of these genes could play a similar role in sheep. The sheep genome has been poorly studied compared with other species; therefore, these results contribute to an exploratory analysis of candidate novel genes that can be used in future studies to elucidate the participation of these genes in productive traits.

CONCLUSIONS

This study is one of the first to conduct a genomic analysis of Creole hair sheep in Colombia to evaluate genomic traits associated with muscle growth traits, although the sample size is small, this is pioneering research, and its information represents a fundamental input for the country's sheep sector. In conclusion, four candidate genes, namely CEP135, EMCN, PAM, and PIAS2, were associated with eight growth traits. The CEP135 gene was the most relevant because it was associated with three of the eight traits evaluated post-weaning (AFW, WGPOS, and FT). The EMCN and PIAS2 genes were associated with the same traits (AFW and WGPOS), while the PAM gene was associated only with one preweaning growth trait (BW). It is important to highlight that the functions of these genes indicate their involvement mainly in cellular growth, apoptosis, and angiogenesis. These results can be used as a basis for future exploratory research, metabolic network analysis, and functional validation of the candidate genes.

ACKNOWLEDGMENTS

To the Ministerio de Ciencia, Tecnología, e innovación for funding this research, Universidad Nacional de Colombia - Palmira and Bogotá Headquarters and Universidad de Cordoba for the collaborative work. To Asovinos, el Hatico Natural Reserve, and small-scale producers of the departments of Cesar and Cordoba for their collaboration.

REFERENCES

Al-Mamun HA, Kwan P, Clark SA, Ferdosi MH et al (2015) Genome-wide association study of body weight in Australian Merino sheep reveals an orthologous region on OAR6 to human and bovine genomic regions affecting height and weight. Genetics Selection Evolution 47: 66. https://doi.org/10.1186/s12711-015-0142-4

Romero H, Vanegas R, Riveros E and Arcos D (2002) Ovinos colombianos de pelo: Alternativa productiva para el sur del departamento del Tolima. PE: Corporación Colombiana de Investigación Agropecuaria, Corpoica. Ibagué, Colombia. 78 p. http://hdl.handle.net/20.500.12324/12831

Barrett JC, Fry B, Maller J and Daly MJ (2005) Haploview: analysis and visualization of LD and haplotype maps. Bioinformatics 21(2): 263-265. https://doi.org/10.1093/bioinformatics/bth457

Burn B, Brown S and Chang C (2011) Regulation of Early Xenopus Development by the PIAS genes. Developmental Dynamics: an official publication of the American Association of Anatomists 240(9): 2120–2126. https://doi.org/10.1002/dvdy.22701

Casas E and White SN (2015) Aplicación de la genómica en características de importancia económica en poblaciones ovinas. pp 35. In: memorias del X seminario internacional de producción de ovinos en el trópico; november 23-26: Villa Hermosa, Tabasco, México.

Corpoica - Corporación Colombiana de Investigación Agropecuaria (2007) Manual para el manejo de las razas criollas ovino de pelo y caprino dentro del plan fomento. PE: Produmedios. Cundinamarca, Colombia. 52 p. http://hdl.handle.net/20.500.12324/13251

Gholizadeh M, Rahimi-Mianji G and Nejati-Javaremi A (2015) Genomewide association study of body weight traits in Baluchi sheep. Journal of Genetics 94(1): 143–146. https://doi.org/10.1007/s12041-015-0469-1

Gogarten SM, Bhangale T, Conomos MP, Laurie CA, McHugh CP et al (2012) GWASTools: an R/Bioconductor package for quality control and analysis of Genome-Wide Association Studies. Bioinformatics 28: 3329–3331. https://doi.org/10.1093/bioinformatics/bts610

Hussain MS, Baig SM, Neumann S, Nürnberg G, Farooq M et al (2012) A truncating mutation of CEP135 causes primary microcephaly and disturbed centrosomal function. American Journal of Human Genetics 90(5): 871–878. https://doi.org/10.1016/j.ajhg.2012.03.016

Kemper KE, Visscher PM and Goddard ME (2012) Genetic architecture of body size in mammals. Genome Biology 13: 244. https://doi.org/10.1186/gb4016

Knapik J, Ropka-Molik K and Pieszka M (2017) Genetic and nutritional factors determining the production and quality of sheep meat – a review. Annals of Animal Science 17(1): 23–40. https://doi.org/10.1515/aoas-2016-0036

Kominakis A, Hager-Theodorides AL, Zoidis E, Saridaki A et al (2017) Combined GWAS and 'guilt by association' based prioritization analysis identifies functional candidate genes for body size in sheep. Genetics Selection Evolution 49: 41. https://doi.org/10.1186/s12711-017-0316-3

Kullander K and Klein R (2002) Mechanisms and functions of Eph

and ephrin signalling. Nature Reviews Molecular Cell Biology 3(7): 475-486. https://doi.org/10.1038/nrm856

Kumar A, Rajendran V, Sethumadhavan R and Purohit R (2013) CEP proteins: the knights of centrosome dynasty. Protoplasma 250(5): 965-983. https://doi.org/10.1007/s00709-013-0488-9

Li J and Johnson SE (2013) Ephrin-A5 promotes bovine muscle progenitor cell migration before mitotic activation. Journal of Animal Science 91(3): 1086-1093. https://doi.org/10.2527/jas.2012-5728

Matsuura K, Lefebvre PA, Kamiya R and Hirono M (2004) Bld10p, a novel protein essential for basal body assembly in Chlamydomonas: localization to the cartwheel, the first ninefold symmetrical structure appearing during assembly. The Journal of Cell Biology 165(5): 663–671. https://doi.org/10.1083/jcb.200402022

Merkler DJ (1994) C-terminal amidated peptides: production by the in vitro enzymatic amidation of glycine-extended peptides and the importance of the amide to bioactivity. Enzyme and Microbial Technology 16(6): 450–456. https://doi.org/10.1016/0141-0229(94)90014-0

Olfert IM, Baum O, Hellsten Y and Egginton S (2015) Advances and challenges in skeletal muscle angiogenesis. American Journal of Physiology Heart and Circulatory Physiology 310(3): H326–H336. https://doi.org/10.1152/ajpheart.00635.2015

Park-Windhol C, Ng YS, Yang J, Primo V, Saint-Geniez M and D'Amore PA (2017) Endomucin inhibits VEGF-induced endotelial cell migration, growth, and morphogenesis by modulating VEGFR2 signaling. Scientific Reports 7(1): 17138. https://doi.org/10.1038/s41598-017-16852-x

Pasquale EB (2005) Eph receptor signalling casts a wide net on cell behaviour. Nature Reviews Molecular Cell Biology 6: 462-475. https://doi.org/10.1038/nrm1662

Peng WF, Xu SS, Ren X, Lv FH, Xie XL et al (2017) A genome-wide association study reveals candidate genes for the supernumerary nipple phenotype in sheep (*Ovis aries*). Animal Genetics 48(5): 570-579. https://doi.org/10.1111/age.12575

Prigge ST, Mains RE, Eipper BA and Amzel LM (2000) New insights into copper monooxygenases and peptide amidation: structure, mechanism and function. Cellular and Molecular Life Sciences 57(8): 1236– 1259. https://doi.org/10.1007/pl00000763

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MA et al (2007) PLINK: a tool set for whole-genome association and population-based linkage analyses. American Journal of Human Genetics 81(3): 559-575. https://doi.org/10.1086/519795

Roberts PJ and Der CJ (2007) Targeting the Raf-MEK-ERK mitogen-activated protein kinase cascade for the treatment of cancer. Oncogene 26(22): 3291-3310. https://doi.org/10.1038/sj.onc.1210422

Traci AC, Ning Y, Hsu M, Peng B et al (2005) Deletion of peptide amidation enzymatic activity leads to edema and embryonic lethality in the mouse. Developmental Biology 287(2): 301-313. https://doi.org/10.1016/j.ydbio.2005.09.001

Vergara GO, Medina H, Sierra RC, Simanca SJ and Bustamante YM (2017) Determinación de la curva de crecimiento en ovinos criollos de pelo, mediante la utilización del modelo Gompertz, en el trópico bajo colombiano. Revista UDCA. Actualidad & Divulgación Científica 20(2): 385-391. https://doi.org/10.31910/rudca.v20.n2.2017.396

Vergara-Garay O, Llorente-Martínez E, Ramos-Caro L, Bustamante-Yánez M and Simanca-Sotelo JC (2016) Descripción del crecimiento en ovinos criollos utilizando el modelo Brody. Orinoquía 20(2): 34-9. https://doi.org/10.22579/20112629.351

Yu G and He Q (2016) ReactomePA: an R/Bioconductor package for reactome pathway analysis and visualization. Molecular Biosystems 12(2): 477-479. https://doi.org/10.1039/C5MB00663E

Zhang L, Liu J, Zhao F, Ren H, Xu L et al (2013) Genome-Wide Association Studies for Growth and Meat Production Traits in Sheep. PLoS One 8(6): e66569. https://doi.org/10.1371/journal.pone.0066569

Zheng X, Levine D, Shen J, Gogarten SM, Laurie C and Weir BS (2012) A high-performance computing toolset for relatedness and principal component analysis of SNP data. Bioinformatics. 28: 3326–3328. https://doi.org/10.1093/bioinformatics/bts606

Zuleta SM, Roncallo FB, Ávila ME, Mojica RJE, Castro RE and Rodríguez FG (2009) Manejo de ovinos de pelo en el trópico. PE: Produmedios. Bogotá, Colombia. 59 p.