ANIMAL GENETICS Immunogenetics, Molecular Genetics and Functional Genomics

doi: 10.1111/age.12769

Characterization and expression analysis of *KIT* and *MITF-M* genes in llamas and their relation to white coat color

M. Anello, M. S. Daverio, M. B. Silbestro, L. Vidal-Rioja and F. Di Rocco

Laboratorio de Genética Molecular, Instituto Multidisciplinario de Biología Celular (IMBICE), CONICET-UNLP-CIC, Calle 526 e/10 y 11, La Plata 1900, Buenos Aires, Argentina.

Summary

The llama (*Lama glama*) is a fiber-producing species that presents a wide range of coat colors, among which white is one of the most important for the textile industry. However, there is little information about the molecular mechanisms that control the white phenotype in this species. In domestic mammals, a white coat is usually produced by mutations in the *KIT proto-oncogene receptor tyrosine kinase* (*KIT*) and *microphthalmia-associated transcription factor* (*MITF*) genes. In this work we have sequenced and described the coding regions of *KIT* and *MITF-M,* the melanocyte-specific isoform, and the two transcriptional variants *MITF*-M(−) and *MITF-M*(+). Moreover, we studied the expression of these genes in the skin of white and colored llamas. Although no variants were revealed to be associated with white coat color, significant differences between phenotypes were observed in the expression levels of *KIT* and *MITF-M*. Interestingly, white llamas expressed less *MITF-M*(+) than did colored ones, which is consistent with a consequent reduction in the synthesis of melanin. Even though our results indicate that downregulation of *KIT* and *MITF-M* expression is involved in white phenotype production in llamas, the causative gene of white coat color remains unknown.

Keywords camelids, coat color, melanogenesis genes, splice variants

Introduction

Domestic species usually present greater phenotypic variation than do wild species, and coat color is one of the most significant characteristics. Although wild species show a defined color or pattern, domestic animals are more diverse (Cieslak *et al.* 2011). The llama (*Lama glama*) is a domestic South American Camelid with a wide variety of coat colors, which makes the species valuable for fiber production. Among all possible color phenotypes, complete white is of interest because it is one of the most attractive for the textile industry. Thus, many producers are trying to shift the color of their herds by phenotypic selection. Although white inheritance in llamas is considered dominant over uniform colors (Frank *et al.* 2006), classical crossbreeding tends to produce a high frequency of non-uniform colors and indefinite patterns.

Pigmentation in mammals is a highly conserved process in which specialized cells of the skin, melanocytes, produce melanin through a biological pathway called melanogenesis. Basic coat colors are defined by the relative proportion of two types of melanin: eumelanin (black or brown) and pheomelanin (yellow or red), whereas the white coat color is determined by the absence of pigmentation. Additionally, dilution of melanin can lead to a pale-cream or almost white phenotype. According to their origin, white phenotypes can be divided into two main classes: albino white, caused by a failure in melanin production due to mutations at the albino locus (*tyrosinase* gene), and non-albino white, caused by disruptions in melanocyte development, differentiation or migration. In domestic animals, such as horse, cow and pig, non-albino white is due to mutations in either the *KIT* (*KIT proto-oncogene receptor tyrosine kinase*) or the *MITF*, (*microphthalmia-associated transcription factor*) genes.

KIT encodes a tyrosine kinase receptor located in the cell membrane and, together with its ligand (KITL), regulates proliferation, migration and survival of different cell types, including melanocytes (Roskoski 2005). KIT regulates melanogenesis via the activation of MITF, a basic helix–loop–helix leucine zipper (BHLH-LZ) protein that controls the differentiation and development of melanocytes and is responsible for upregulating transcription of pigmentation enzymes (TYR, TYRP1 and TYRP2). *MITF* genomic sequencing in vertebrates has revealed a complex intron–exon structure in which alternative splicing of exon 1 gives

Address for correspondence

F. Di Rocco, Laboratorio de Genética Molecular, Instituto Multidisciplinario de Biología Celular (IMBICE), CONICET-UNLP-CIC, Calle 526 e/10 y 11, La Plata 1900, Buenos Aires, Argentina.
E-mail: fdirocco@imbice.gov.ar

Accepted for publication 14 December 2018

rise to a family of transcription factors. Exon 1M, with its specific promoter, originates the melanocyte-specific isoform, MITF-M (Shibahara 2001; Hallsson *et al.* 2007).

KIT and MITF show complex interactions: MITF is needed for the maintenance of *KIT* expression in melanoblasts (melanocyte precursors), and KIT signaling modulates MITF activity and stability in melanocytes (Hou *et al.* 2000). Mutations in *KIT* and *MITF* have been reported to cause white and white-spotted phenotypes in several mammals. An example of this is the horse, which has more than 20 mutations in the *KIT* gene that produce dominant white (Haase *et al.* 2010; Holl *et al.* 2017) and spotted patterns like sabino, tobiano or roan (Dürig *et al.* 2017a; Negro *et al.* 2017). Another spotted phenotype in this species, 'splashed white', is caused by mutations in *MITF* (Hauswirth *et al.* 2012; Dürig *et al.* 2017b). In the dog, several authors have associated *MITF* and *KIT* with white spotting (Rothschild *et al.* 2006; Karlsson *et al.* 2007; Schmutz *et al.* 2009; Wong *et al.* 2013; Körberg *et al.* 2014), though some differences have been reported according to the breed.

Not only do mutations in *MITF* and *KIT* have similar effects on the phenotype, there is also evidence that they interact genetically in an additive manner. For instance, double heterozygous mice for the semi-dominant alleles $KIT^{W\text{-}36H}$ and $MITF^{Mi\text{-}wh}$ have a more severe spotting phenotype than do mice heterozygous for either allele alone (Wen *et al.* 2010). Furthermore, Haase *et al.* (2013) reported that the accumulation of mutations at these loci progressively increases the extent of white markings in horse.

Llamas usually present the non-albino white phenotype, characterized by a completely white coat with pigmentation in the eyes, eye rims and snout. Association between pigmented phenotypes and mutations in other genes, such as the *melanocortin 1-receptor* (*MC1R*) and *agouti signaling protein* (*ASIP*) genes, was previously reported by Daverio *et al.* (2016). However, the molecular mechanisms responsible for the white phenotype are still unknown. In this work we addressed the molecular characterization of non-albino white llamas through the study of the *KIT* and *MITF-M* genes. The objectives of this work were to describe *KIT* and *MITF-M* coding regions, to analyze their variation and to study their expression levels in llamas.

Materials and methods

Samples

Samples were collected following the Argentine Ethical Guidelines for Biomedical Investigation in Animals from Laboratory, Farm, or Obtained from Nature (Resolution N° 1047/05 from CONICET, Argentina).

For gene description and variation analysis, skin biopsies from 31 llamas from four different breeding locations were collected. Biopsies were taken from the flank, after the last rib, using a disposable 3-mm-diameter biopsy punch and were conserved in RNAlater (SIGMA) until further extraction.

All animals were unrelated, and phenotypes included colored ($n = 16$) and non-albino white ($n = 15$). For the colored group, we selected pigmented llamas with uniformly colored fleece, including black and different shades of brown. For the non-albino white group, dark-eyed animals with a full white coat were included (Fig. S1). More information about the samples is provided in Table S1.

RNA extraction and cDNA synthesis

Total RNA was extracted by homogenization in TRIzol® and addition of chloroform to separate phases. The aqueous phase was used for the alcoholic precipitation of RNA. Then, it was washed and resuspended in RNase-free water. Reverse transcription to obtain cDNA was performed in a 20-μl reaction volume, using 1 μg/μl of RNA, RevertAid Reverse Transcriptase (Thermo Fisher Scientific) and random primers (Biodynamics), following the manufacturer's instructions.

Sequencing and description of *KIT* and *MITF-M*

To obtain *KIT* and *MITF-M* complete coding regions, PCR primers were designed using PRIMER-BLAST (Ye *et al.* 2012) over conserved regions of mRNA sequences from other mammals, such as sheep, goat and other ungulates, available from GenBank. A design of six and four overlapping fragments was used to fully cover the coding region of the *KIT* and *MITF* genes respectively. A list of all primers used, their sequences and the fragment amplified by each pair is shown in Table S2. Primer MITF1-F was specifically designed to anneal with exon1-M. As described in the mouse (Steingrímsson *et al.* 1994), *Mitf-M* has two alternative splicing products, named *Mitf-M*(−) and *Mitf-M*(+) depending on the addition of 18 bp between exons 5 and 6. To study these variants in llamas, we designed specific primers that flanked the insertion region (MITFvar F and MITFvar R; Table S2), separated the variants by 3% agarose electrophoresis, cut the two bands from the gel and sequenced them. For further analysis, we designed variant-specific primers that annealed within the 18-bp region or in the junction between exons 5 and 6. The primer design strategy for amplifying the *MITF-M* variants is shown in Fig. 1.

The amplification reaction for the *KIT* and *MITF-M* coding regions was carried out in a 25-μl PCR mix containing 1× PCR buffer (Invitrogen), 2.5 mM of $MgCl_2$, 0.2 mM of dNTPs, 60 μg of BSA, 1 U of Taq DNA polymerase (Invitrogen), 0.5 μM of each primer and 50 ng of cDNA. The cycling profile consisted of an initial denaturation step at 94 °C for 3 min; 30 cycles of 40 s at 94 °C, 50 s at 53–59 °C and 40 s at 72 °C; and a 5-min final extension at 72 °C. PCR products were checked on a 1%

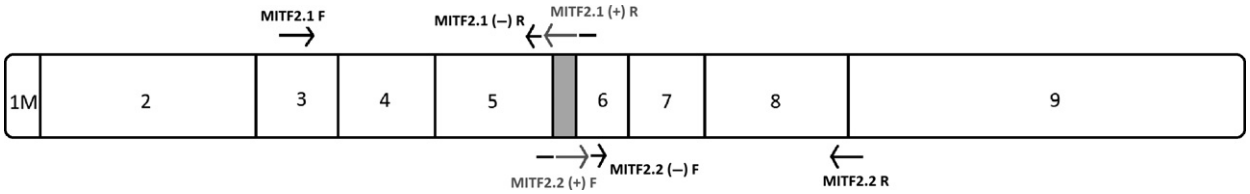

Figure 1 Schematic representation of primer design for the amplification of the transcriptional variants *MITF-M*(–) and *MITF-M*(+). Numbered rectangles correspond to exons in the cDNA; grey shaded denotes the additional 18 bp present in *MITF-M*(+). Right arrows represent forward primers, and left arrows reverse primers. Each variant was amplified into two fragments; one is represented above the cDNA scheme, and the other below. In each fragment amplification a common primer was used, represented by the complete black arrow, and it was combined either with a primer specific for *MITF-M*(+) (grey colored) or a primer specific for *MITF-M*(–) (fragmented black arrow).

agarose gel stained with GelRed™, purified by PEG precipitation and sequenced by Macrogen, Inc. The sequences obtained were aligned and analyzed using GENEIOUS (v.6.1.8; Biomatters). Exon boundaries were determined by comparison with the orthologous *KIT* and *MITF-M* genes of other species available from GenBank. Proteins topology and domains were predicted using CCTOP (Dobson *et al.* 2015) and CDD (Marchler-Bauer *et al.* 2017) software respectively and compared with UniProt reviewed entries (Bateman *et al.* 2017).

Gene expression profiles

For the expression analysis, skin biopsies from white and brown llamas were collected from the two breeding locations of the Buenos Aires province previously sampled. Biopsies were taken from young adult animals during the same week. Samples were kept in RNAlater (SIGMA) during their transportation to the laboratory, where they were immediately extracted to avoid RNA degradation.

Quantitative real-time PCR (qPCR) was carried out using specific primers designed to target the llama coding region for each gene and for each isoform in the case of *MITF-M*. Additionally, each primer of a pair annealed in a different exon or in the exon–exon junction to avoid genomic amplification (Table S2). *GAPDH* (*glyceraldehyde-3-phosphate dehydrogenase*) and ribosomal *18S* were tested as endogenous controls. Both genes have previously been used to study the expression of melanogenesis genes in skin of the sheep (Saravanaperumal *et al.* 2014). In our assays, only *18S* was stably expressed and did not show variation among groups. For that reason, it was subsequently used for data normalization.

Amplification reactions were carried out in a RotorGene Q (Qiagen) and consisted of 20 µl, including 4 µl of HOT FIREPol® EvaGreen® qPCR Mix Plus (ROX; Solis Biodyne), 0.5 mM of each primer and 1 ng of cDNA. The cycling parameters were: 15 min at 95 °C; 40 cycles of 15 s at 95 °C, 20 s at 60 °C and 20 s at 72 °C; and a final gradient from 95 to 72 °C. Each qPCR reaction was optimized; PCR efficiencies calculated from slope were within 90–110%, r^2 over 98%. Each gene was amplified in one event of qPCR that included three technical replicates for every sample and two no template controls. Melting curve analysis was performed following amplification to verify the absence of non-specific amplification or primer dimer. The cycle threshold (Ct) was determined by Rotor-Gene Q–Pure Detection software version 2.3.1. Quantification of transcript abundance was carried out using the comparative Ct method by Livak & Schmittgen (2001), and a *t*-test was used to assess if differences in expression were significant.

Results

KIT

The complete *KIT* coding region in llamas consists of 2919 bp and comprises 21 exons (GenBank accession no. MH062992).

One transcript variant was identified in the skin of llama. However, in sheep two transcript variants of *KIT*—*KIT*(+) and *KIT*(–)—have been reported (Saravanaperumal 2011). We used primers (KIT4-F and KIT4-R) that flanked the 12-bp insertion that differentiates variants (+) and (–), but a single product as well as a clear sequence from the electropherogram was always obtained.

The protein encoded by *KIT* is predicted to have 972 amino acids and well-conserved protein domains. Identity to orthologous proteins was high: 99% with other camelids and over 90% with other mammals. The KIT receptor presented a signal peptide of 25 amino acids, five extracellular immunoglobulin domains, a transmembrane segment from amino acid 525–545, an intracellular tyrosine kinase active domain from amino acid 553–928 and an intracellular C-terminal tail.

Regarding the variation of the *KIT* gene, 18 SNPs were found. One was located at the 5′-untranslated region, 15 bp before the translation starting site; 10 SNPs were found within the immunoglobulin-like (Ig-like) domains; five within the tyrosine kinase domains; one at the C-terminal tail; and one at the 3′-untranslated region, 17 bp after the stop codon (Table 1). Most of these substitutions were synonymous except for three SNPs located within the extracellular Ig-like domains of the protein. None of the

SNPs were associated with color phenotype, as no differences in the distribution of genotypes were observed when comparing colored with white animals (Table 2). Of the non-synonymous SNPs, the first one, c.866C>T in exon 5, changes threonine to methionine at position 289 of the KIT receptor. All llamas analyzed were homozygous for the C allele (which corresponds to threonine in the protein) except for two colored animals that were heterozygotes. The second non-synonymous SNP, c.967G>C, was found in exon 6, representing a change from valine to leucine in the protein. This SNP was very frequent, and we observed the three possible genotypes regardless of the color phenotype. The last non-synonymous SNP, c.1265C>T, produces a change in the protein from alanine to valine; however, this was observed just once in a pigmented llama.

MITF-M

The *MITF-M* coding region presented nine exons, including exon 1M, the melanocyte-specific first exon. Sequence identity of the complete coding region was 93–99% with other mammals, and exon 1M presented 100% identity with other species, including carnivores, even-toed ungulates, primates, rodents and placentals. The exon 1M sequence presented 33 bp (ATGCTGGAAATGCTAGAATA TAATCACTATCAG) that encoded the first 11 amino acids of the MITF-M protein (Met-Leu-Glu-Met-Leu-Glu-Tyr-Asn-His-Tyr-Gln).

Furthermore, *MITF-M* presented two alternative splice products: *MITF-M*(−) with 1242 bp and *MITF-M*(+) with the insertion of 18 bp between exons 5 and 6. These variants were first separated by electrophoresis and sequenced with a pair of flanking primers. An agarose gel with PCR amplification of both transcript variants is shown in Fig. 2a, and the corresponding sequences are shown in Fig. 2b (GenBank accession nos. MH062994 and MH062993). Once each sequence was obtained, new variant-specific primers were designed for the amplification of *MITF-M*(+) and *MITF-M*(−).

MITF-M(−) encoded a 413-amino acid protein, and the protein encoded by *MITF-M*(+) presented six extra amino acids (Ala-Cys-Ile-Phe-Pro-Thr) located before a BHLH-LZ domain. This domain is found in specific DNA-binding proteins that act as transcription factors and is required for DNA binding and dimerization. In llamas, it is located from amino acid 197–252. Besides the BHLH-LZ domain, the predicted protein presents a highly conserved domain from amino acids 11–87 that is usually found at the N-terminus of several transcription factors including MITF. It also presents a highly conserved transactivation domain that contains recruitment information for the transcription cofactor CBP/p300 (Sato *et al.* 1997); it comprises amino acids 114–132 of the llama predicted protein (Met-Gln-Met-Asp-Asp-Val-Ile-Asp-Asp-Ile-Ile-Ser-Leu-Glu-Ser-Ser-Tyr-Asn-Glu).

Only one SNP was found in the *MITF-M* coding region: a synonymous T to C substitution at nucleotide 532 (c.532T>C). We observed this polymorphism in both transcript variants. There was no relation between genotype and phenotype, as the three possible genotypes were observed in colored as well as white animals.

Expression profiles

All samples consistently amplified both the housekeeping gene and the gene of interest. The qPCR melting curves confirmed that only a single product was amplified in each PCR and no template controls showed no amplification. We compared expression levels of *KIT*, total expression of *MITF-M* and then each *MITF-M* variant separately for colored and white llamas. Results showed that expression levels of *KIT*, *MITF-M* and *MITF-M*(+) were significantly lower in white animals than in colored ones ($P < 0.05$). Expression difference for *MITF-M*(−) was not statistically significant between groups ($P = 0.1626$; Fig. 3).

Table 1 *KIT* polymorphisms.

Polymorphism (cDNA)	Gene location	Amino acid change	Protein domain
c.-15A>G	5′-UTR	–	–
c.324T>C	Exon 2	Syn	Ig-like domain
c.840A>G	Exon 5	Syn	Ig-like domain
c.866C>T	Exon 5	p.Thr289Met	Ig-like domain
c.967G>C	Exon 6	p.Val323Leu	Ig-like domain
c.1119C>T	Exon 7	Syn	Ig-like domain
c.1122A>G	Exon 7	Syn	Ig-like domain
c.1173G>A	Exon 7	Syn	Ig-like domain
c.1265C>T	Exon 8	p.Ala422Val	Ig-like domain
c.1377A>G	Exon 9	Syn	Ig-like domain
c.1485C>T	Exon 9	Syn	Ig-like domain
c.1830G>A	Exon 12	Syn	Tyrosine kinase domain
c.1845C>T	Exon 12	Syn	Tyrosine kinase domain
c.2715T>C	Exon 20	Syn	Tyrosine kinase domain
c.2721T>C	Exon 20	Syn	Tyrosine kinase domain
c.2745G>A	Exon 20	Syn	Tyrosine kinase domain
c.2808G>A	Exon 21	Syn	C-terminal tail
c.+17A>G	3′-UTR	–	–

Position +1 corresponds to ATG.
UTR, untranslated region; Syn, synonymous mutations.

Table 2 Distribution of genotypes for each *KIT* non-synonymous polymorphism between color groups.

Color groups	c.866C>T			c.967G>C			c.1265C>T		
	C/C	C/T	T/T	G/G	G/C	C/C	C/C	C/T	T/T
Non-albino white	15	–	–	6	5	4	15	–	–
Colored	14	2	–	6	8	2	15	1	–

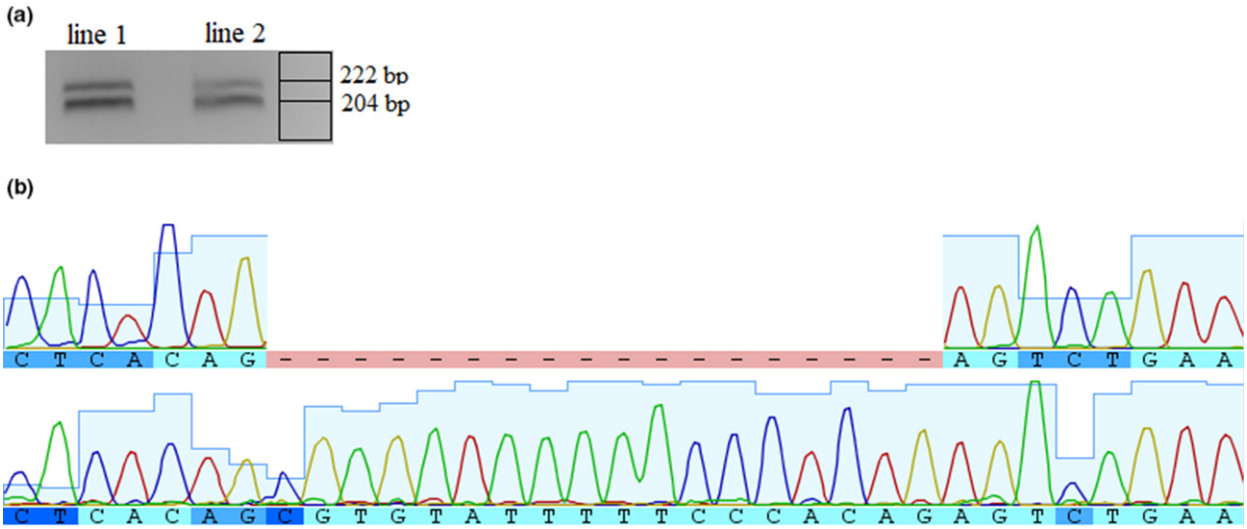

Figure 2 (a) Agarose gel showing bands corresponding to the variants of *MITF-M*(−) and *MITF-M*(+). Line 1: sample from a colored llama; line 2: sample from a white llama. (b) Aligned sequences of the two variants.

Discussion

This paper describes for the first time complete *KIT* and *MITF-M* coding regions and the expression level of these genes in the skin of llamas. Because the melanocyte-specific form of *MITF* was expressed and successfully amplified in both colored and white animals, our results show that white llamas do present melanocytes in the skin, in agreement with Frank's (2001) observations.

The structure of *KIT* and *MITF-M* genes was similar to that previously described in other vertebrates (Hallsson *et al.* 2007; Opatowsky *et al.* 2014). Although alternative splicing variants have been reported for both genes, in llamas we observed two transcript variants of *MITF-M* and only one of *KIT*. The presence or absence of 12 bp between exons 9 and 10 results in isoforms *KIT*(+) and *KIT*(−) (Reith *et al.* 1991; Gokkel *et al.* 1992). According to Reith *et al.* (1991), variants *KIT*(+) and *KIT*(−) are co-expressed in some cell types, but the *KIT*(−) form predominates. We found only the transcript variant that corresponds to *KIT*(−); *KIT*(+) may not be expressed in the skin of llama or may be expressed at such low levels that it could not be detected.

In many mammals, mutations in *KIT* and *MITF-M* are responsible for white phenotypes. We did not find mutations in those genes that could explain the white phenotype in llamas. Additionally, all the non-synonymous SNPs found in *KIT* were located in the Ig-like domain, a region where mutations usually result in hyperactivation of the receptor (Ray & Krishnamoorthy 2008). Instead, a causal mutation of white phenotype is expected to produce a decrease of KIT function.

MITF-M was much less variable than *KIT*. Only one synonymous mutation was found: c.532T>C. The same SNP was previously observed by Saravanaperumal *et al.*

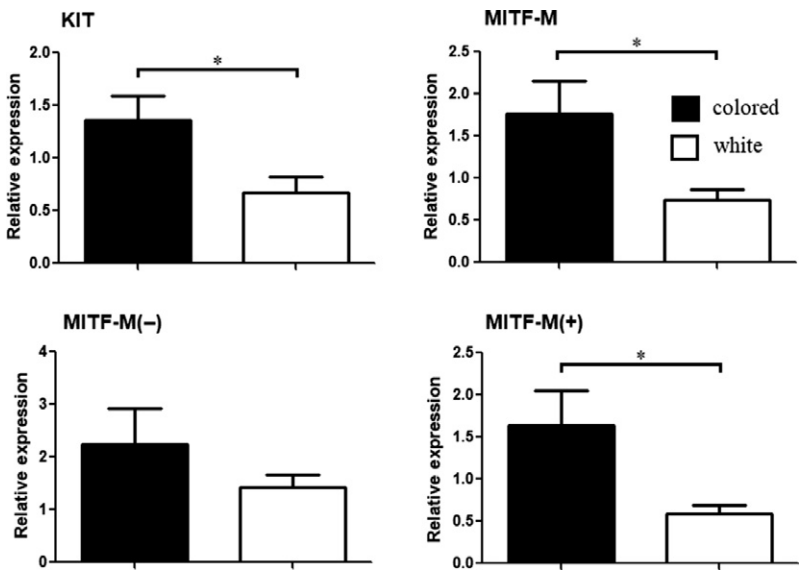

Figure 3 Relative expression of *KIT*, *MITF-M*, *MITF-M*(−) and *MITF-M*(+) in the skin of llamas. Bar graphs show mean values. Error bars represent standard error (SE). Six biological replicates were analyzed for *KIT* and *MITF-M*, and five for *MITF-M*(−) and *MITF-M*(+) isoforms; asterisk indicates statistically significant difference ($P < 0.05$).

(2014) in the sheep *MITF-M*(+) variant, and it was proposed to affect dimer formation or DNA binding of that variant due to its proximity to the basic region of MITF-M. On the contrary, in llama we observed this substitution in both transcript variants.

Although the coding regions of *MITF-M* and *KIT* did not show any mutations that could relate them to the white phenotype, the expression patterns observed indicate that these genes may have a role in the production of white coat color in llama. Our results reveal that the expression levels of *KIT* and *MITF-M* were significantly lower in white animals than in colored ones. Particularly, the difference found in *MITF-M* seems to be caused by the *MITF-M*(+) variant, given that *MITF-M*(−) did not show significant differences between white and colored animals. MITF-M(+) is known to bind target E-boxes in promoter DNA with higher affinity and transcribe tyrosinase more efficiently than MITF(−) does (Murakami *et al.* 2007). Consistent with that, $Mitf^{mi\text{-}sp}$ (microphthalmia spotted) mice, which express only the *Mitf-M*(−) variant, present a mild pigmentation (Bismuth *et al.* 2005). Similarly, white llamas express less *MITF-M*(+) than do colored ones, which could be explained by a decrease in the synthesis of melanin enzymes produced by the lower expression of this variant.

It is expected that the same expression pattern in both *KIT* and *MITF-M* genes would be found, given that they regulate each other. MITF-M is needed to maintain *KIT* expression in melanoblasts, and KIT is further required for the differentiation of these melanoblasts into melanocytes, which express *MITF-M* to produce melanin (Hou *et al.* 2000). Unpublished data from our laboratory indicate that other genes involved in the pigmentation process are also less highly expressed in white llamas compared to colored animals. Additionally, Munyard (2011) observed that eight color-related genes are down-regulated in white alpaca. The observed pattern in llamas could be explained by the action of another gene located upstream in the pathway that regulates the pigmentation process. One possible candidate gene is *ASIP*. Under normal conditions, binding of ASIP to MC1R inhibits signal transduction, causing the melanocytes to produce pheomelanin instead of eumelanin. *In-vitro* studies showed that *ASIP* overexpression inhibits melanin synthesis (Sakai *et al.* 1997; Aberdam *et al.* 1998; Le Pape *et al.* 2008). *In vivo*, *ASIP* overexpression has been shown to be associated with the white phenotype in beach mice (Manceau *et al.* 2011), sheep (Norris & Whan 2008) and alpaca (Chandramohan *et al.* 2013). Our laboratory has recently sequenced the coding region of *ASIP* in llamas, but no mutations were found to be associated with the white phenotype (Daverio *et al.* 2016). Therefore, it is necessary to study the expression and the transcript structure of this gene in llamas.

This work represents the first step to molecularly characterize non-albino white llamas. Moreover, the descriptions of *KIT* and *MITF-M* contribute to the genetic knowledge of a species whose genome is still not available.

Acknowledgements

We thank the llama owners who allowed us to take skin samples. We also appreciate the collaboration of E. Frank, L. Brodtmann, F. Veiga, V. Trasorras and J Martínez. This work was supported by grant PIP-00370 from the National Scientific and Technical Research Council (CONICET) and funds from the Commission of Scientific Research of the Province of Buenos Aires (CIC) F. Di Rocco is researcher from CIC.

Conflict of interest

The authors declare no conflict of interest.

References

Aberdam E., Bertolotto C., Sviderskaya E.V., De Thillott V., Hemesath T.J., Fisher D.E., Bennett D.C., Ortonne J.P. & Ballotti R. (1998) Involvement of microphthalmia in the inhibition of melanocyte lineage differentiation and of melanogenesis by agouti signal protein. *Journal of Biological Chemistry* **273**, 19560–5.

Bateman A., Martin M.J., O'Donovan C. *et al.* (2017) UniProt: the universal protein knowledgebase. *Nucleic Acids Research* **45**, D158–69.

Bismuth K., Maric D. & Arnheiter H. (2005) MITF and cell proliferation: the role of alternative splice forms. *Pigment Cell Research*, **18**, 349–59.

Chandramohan B., Renieri C., La Manna V. & La Terza A. (2013) The alpaca agouti gene: genomic locus, transcripts and causative mutations of eumelanic and pheomelanic coat color. *Gene* **521**, 303–10.

Cieslak M., Reissmann M., Hofreiter M. & Ludwig A. (2011) Colours of domestication. *Biological Reviews* **86**, 885–99.

Daverio M.S., Rigalt F., Romero S., Vidal-Rioja L. & Di Rocco F. (2016) Polymorphisms in *MC1R* and *ASIP* genes and their association with coat color phenotypes in llamas (*Lama glama*). *Small Ruminant Research* **144**, 83–9.

Dobson L., Reményi I. & Tusnády G.E. (2015) CCTOP: a Consensus Constrained TOPology prediction web server. *Nucleic Acids Research* **43**, W408–12.

Dürig N., Jude R., Holl H., Brooks S.A., Lafayette C., Jagannathan V. & Leeb T. (2017a) Whole genome sequencing reveals a novel deletion variant in the *KIT* gene in horses with white spotted coat colour phenotypes. *Animal Genetics* **48**, 483–5.

Dürig N., Jude R., Jagannathan V. & Leeb T. (2017b) A novel *MITF* variant in a white American Standardbred foal. *Animal Genetics* **48**, 123–4.

Frank E.N. (2001) *Descripción y Análisis de Segregación de Fenotipos de Color y Tipos de Vellón en Llamas Argentina*. Doctoral Thesis. University of Buenos Aires, Argentina.

Frank E.N., Hick M.V.H., Gauna C.D., Lamas H.E., Renieri C. & Antonini M. (2006) Phenotypic and genetic description of fibre traits in South American domestic camelids (llamas and alpacas). *Small Ruminant Research* **61** (Special Issue 2–3), 113–29.

Gokkel E., Grossman Z., Ramot B., Yarden Y., Rechavi G. & Givol D. (1992) Structural organization of the murine *c-kit proto-oncogene*. *Oncogene* **7**, 1423–9.

 doi: 10.1111/age.12769

Haase B., Rieder S., Tozaki T., Hasegawa T., Penedo M.C., Jude R. & Leeb T. (2010) Five novel *KIT* mutations in horses with white coat colour phenotypes. *Animal Genetics* **42**, 337–40.

Haase B., Signer-Hasler H., Binns M.M., Obexer-Ruff G., Hauswirth R., Bellone R.R., Burger D., Rieder S., Wade C.M. & Leeb T. (2013) Accumulating mutations in series of haplotypes at the *KIT* and *MITF* loci are major determinants of white markings in Franches-Montagnes horses. *PLoS ONE* **8**, 1–10.

Hallsson J.H., Haflidadóttir B.S., Schepsky A., Arnheiter H. & Steingrímsson E. (2007) Evolutionary sequence comparison of the *Mitf* gene reveals novel conserved domains. *Pigment Cell Research* **20**, 185–200.

Hauswirth R., Haase B., Blatter M. *et al.* (2012) Mutations in *MITF* and *PAX3* cause 'splashed white' and other white spotting phenotypes in horses. *PLoS Genetics* **8**, e1002653.

Holl H.M., Brooks S.A., Carpenter M.L., Bustamante C.D. & Lafayette C. (2017) A novel splice mutation within equine *KIT* and the *W15* allele in the homozygous state lead to all white coat color phenotypes. *Animal Genetics* **48**, 497–8.

Hou L., Panthier J.J. & Arnheiter H. (2000) Signaling and transcriptional regulation in the neural crest-derived melanocyte lineage: interactions between *KIT* and *MITF*. *Development (Cambridge)* **127**, 5379–89.

Karlsson E.K., Baranowska I., Wade C.M. *et al.* (2007) Efficient mapping of mendelian traits in dogs through genome-wide association. *Nature Genetics* **39**, 1321–8.

Körberg I.B., Sundström E., Meadows J.R.S. *et al.* (2014) A simple repeat polymorphism in the MITF-M promoter is a key regulator of white spotting in dogs. *PLoS ONE* **9**, e104363.

Le Pape E., Wakamatsu K., Ito S., Wolber R. & Hearing V.J. (2008) Regulation of eumelanin/pheomelanin synthesis and visible pigmentation in melanocytes by ligands of the melanocortin 1 receptor. *Pigment Cell and Melanoma Research* **21**, 477–86.

Livak K.J. & Schmittgen T.D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods* **25**, 402–8.

Manceau M., Domingues V.S., Mallarino R. & Hoekstra H.E. (2011) The developmental role of agouti in color pattern evolution. *Science* **331**, 1062–5.

Marchler-Bauer A., Bo Y., Han L. *et al.* (2017) CDD/SPARCLE: functional classification of proteins via subfamily domain architectures. *Nucleic Acids Research* **45**, D200–3.

Munyard K. (2011) *Inheritance of White Colour in Alpacas* (RIRDC Publication No. 11/074). Rural Industries Research and Development Corporation, Barton, Australia.

Murakami M., Iwata Y. & Funaba M. (2007) Expression and transcriptional activity of alternative splice variants of *Mitf* exon 6. *Molecular and Cellular Biochemistry* **303**, 251–7.

Negro S., Imsland F., Valera M., Molina A., Solé M. & Andersson L. (2017) Association analysis of *KIT, MITF* and *PAX3* variants with white markings in Spanish horses. *Animal Genetics* **48**, 349–52.

Norris B.J. & Whan V.A. (2008) A gene duplication affecting expression of the ovine *ASIP* gene is responsible for white and black sheep. *Genome Research* **18**, 1282–93.

Opatowsky Y., Lax I., Tomé F., Bleichert F., Unger V.M. & Schlessinger J. (2014) Structure, domain organization, and different conformational states of stem cell factor-induced intact KIT dimers. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 1772–7.

Ray P. & Krishnamoorthy N.A.R. (2008) Emerging functions of c-kit and its ligand stem cell factor in dendritic cells. *Cell Cycle* **15**, 32826–32.

Reith A.D., Ellis C., Lyman S.D., Anderson D.M., Williams D.E., Bernstein A. & Pawson T. (1991) Signal transduction by normal isoforms and W mutant variants of the Kit receptor tyrosine kinase. *EMBO Journal* **10**, 2451–9.

Roskoski R. (2005) Structure and regulation of Kit protein-tyrosine kinase—the stem cell factor receptor. *Biochemical and Biophysical Research Communication* **338**, 1307–15.

Rothschild M.F., Van Cleave P.S., Glenn K.L., Carlstrom L.P. & Ellinwood N.M. (2006) Association of *MITF* with white spotting in Beagle crosses and Newfoundland dogs. *Animal Genetics* **37**, 606–7.

Sakai C., Ollmann M., Kobayashi T., Abdel-Malek Z., Muller J., Vieira W.D., Imokawa G., Barsh G.S. & Hearing V.J. (1997) Modulation of murine melanocyte function *in vitro* by agouti signal protein. *EMBO Journal* **16**, 3544–52.

Saravanaperumal S.A. (2011) *Sheep Coat Colour Genes: Characterization of Multiple Splice Variants and Differential Gene Expression Analysis of* SCF, c-KIT *and* MITF *in White Merino Sheep*. Ph.D. Thesis. University of Camerino, Italy.

Saravanaperumal S.A., Pediconi D., Renieri C. & La Terza A. (2014) Alternative splicing of the sheep *MITF* gene: novel transcripts detectable in skin. *Gene* **552**, 165–75.

Sato S., Roberts K., Gambino G., Cook A., Kouzarides T. & Goding C.R. (1997) CBP/p300 as a co-factor for the microphthalmia transcription factor. *Oncogene* **14**, 3083–92.

Schmutz S.M., Berryere T.G. & Dreger D.L. (2009) *MITF* and white spotting in dogs: a population study. *Journal of Heredity*, **100** (Suppl 1), S66–74.

Shibahara S. (2001) Microphthalmia-associated transcription factor (MITF): multiplicity in structure, function, and regulation. *Journal of Investigative Dermatology Symposium Proceedings* **6**, 99–104.

Steingrímsson E., Moore K.J., Lamoreux M.L. *et al.* (1994) Molecular basis of mouse microphthalmia (mi) mutations helps explain their developmental and phenotypic consequences. *Nature Genetics* **8**, 256–63.

Wen B., Chen Y., Li H. *et al.* (2010) Allele-specific genetic interactions between *Mitf* and *Kit* affect melanocyte development. *Pigment Cell and Melanoma Research* **23**, 441–7.

Wong A.K., Ruhe A.L., Robertson K.R., Loew E.R., Williams D.C. & Neff M.W. (2013) A *de novo* mutation in *KIT* causes white spotting in a subpopulation of German Shepherd dogs. *Animal Genetics* **44**, 305–10.

Ye J., Coulouris G., Zaretskaya I., Cutcutache I., Rozen S. & Madden T.L. (2012) PRIMER-BLAST: a tool to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics* **13**, 134.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Figure S1 Phenotypes.

Table S1 Samples.

Table S2 Complete list of primers used in this work.