



Sequence characterization of the melanocortin 1 receptor (*MC1R*) gene in sheep with different coat colours and identification of the putative *e* allele at the ovine *Extension* locus

L. Fontanesi^{a,*}, F. Beretti^{a,b,1}, V. Riggio^b, S. Dall'Olio^a, D. Calascibetta^b, V. Russo^a, B. Portolano^b

^a DIPROVAL, Sezione di Allevamenti Zootecnici, University of Bologna, Via F.lli Rosselli 107, 42100 Reggio Emilia, Italy

^b Dep. S.En.Fi.Mi.Zo., Sezione di Produzioni Animali, University of Palermo, Viale delle Scienze – Parco d'Orleans, 90128 Palermo, Italy

ARTICLE INFO

Article history:

Received 23 December 2009

Received in revised form 20 March 2010

Accepted 23 March 2010

Available online 20 April 2010

Keywords:

Coat colour

MC1R

Missense mutation

Ovine breeds

ABSTRACT

Sequence of the melanocortin 1 receptor (*MC1R*) gene (the *Extension* locus) was obtained from a panel of 73 animals belonging to 9 Italian sheep breeds or populations (Appenninica, Bergamasca, Comisana, Cornigliese-like, Delle Langhe, Massese, Merinizzata Italiana, Sarda and Valle del Belice) with different coat colours. Evaluation of the identified polymorphisms on this phenotype was reported with *in silico* predictions and comparative approaches within and across breeds and across species. Five novel single nucleotide polymorphisms (SNPs), organized in three haplotypes, were detected. Another haplotype, including the two missense mutations already described for the *E^D* allele, was identified in few Massese sheep. One SNP (c.199C > T) caused a predicted amino acid substitution (p.R67C) in a highly conserved position of the first intracellular loop of the *MC1R* protein. The same substitution causes recessive pheomelanism in other species. We propose that the p.67C allele represents the recessive *e* allele at the ovine *Extension* series that was, so far, not completely recognized in sheep by classical genetic studies. This polymorphism was analysed in a total of 388 sheep of the 9 investigated breeds. The p.67C allele was identified only in the Valle del Belice breed (allele frequency of 21.3% in 176 analysed animals of this breed) in which the presence of epistatic white-determining loci might mask, at least in part, its effects. Confirming the effect of this novel allele on coat colour will lead to new perspectives on the composition of specialized coloured sheep lines.

© 2010 Elsevier B.V. All rights reserved.

1. Introduction

Coat colour in farm animals is one of the most important traits considered for breed identification and attribution. Coat colour has other practical applications for fibre production in sheep, goats and other farm animals. Variability in this trait is due to the presence, distribution and biochemical activity of the melanocytes in which two types

of melanin (eumelanins and pheomelanins, determining black/brown and red/yellow colours, respectively) are synthesized. A large number of genes affects coat colour as reported in the laboratory mouse which is considered as the model species for these studies (Bennett and Lamoignon, 2003). However, two loci (*Agouti* and *Extension*) play a major role in determining coat colour controlling and regulating the relative amounts of eumelanins and pheomelanins in skin and hair (Searle, 1968). The *Agouti* locus encodes for the agouti signalling protein (ASIP; Bultman et al., 1992), a small paracrine signalling molecule that interacts with the product of the *Extension* locus. The *Extension* locus encodes for the melanocortin 1 receptor

* Corresponding author. Tel.: +39 0522 290516; fax: +39 0522 290523.
E-mail address: luca.fontanesi@unibo.it (L. Fontanesi).

¹ These authors contributed equally to this work.

(MC1R) that is a seven transmembrane domains protein belonging to the G-protein coupled receptor present on the surface of the melanocyte membrane (Robbins et al., 1993). Binding of the MC1R with the α melanocyte-stimulating hormone (α MSH) induces eumelanin synthesis, whereas alternative binding with ASIP causes a pigment-type switching from eumelanins to pheomelanins (Lu et al., 1994; Ollmann et al., 1998). The *Agouti* and *Extension* loci show epistatic interactions with few exceptions. Dominant *Extension* alleles induce black pigmentation, whereas recessive alleles extend the production of pheomelanins, determining red/yellow/pale pigmentation. The presence of wild type *Extension* alleles is usually needed for the expression of the mutated *Agouti* alleles that have, in general, opposite models of action, i.e. dominant *Agouti* alleles determine pheomelanin phenotypes, whereas recessive alleles cause black coat colour (Searle, 1968).

Gain of function and loss of function mutations of the *MC1R* gene determining dominant or partially dominant black/dark and recessive or partially recessive red/yellow/pale coat colour phenotypes, respectively, have been described in several mammals, such as mice (Robbins et al., 1993), humans (Valverde et al., 1995), cattle (Klungland et al., 1995; Joerg et al., 1996; Rouzaud et al., 2000), pigs (Kijas et al., 1998, 2001), horses (Marklund et al., 1996), foxes (Våge et al., 1997; Våge et al., 2005), beach and pocket mice (Hoekstra et al., 2006; Nachman et al., 2003), rabbits (Fontanesi et al., 2006), and goats (Fontanesi et al., 2009a).

In sheep, classical genetic studies have identified a two-allele series at the *Extension* locus: the dominant black (E^D) allele that may account for the black of a few coloured breeds; and the wild type (E^+) allele widely distributed in most breeds in which segregation of *Agouti* alleles (about 20 *Agouti* alleles have been described in sheep) should account for the majority of colour variation (Searle, 1968; Sponenberg, 1997). Contrary to other species, the recessive e allele of the *Extension* locus has not yet been clearly documented in sheep.

Analysing the sheep *MC1R* gene, Våge et al. (1999) identified two missense mutations (p.M73K and p.D121N) determining the dominant black (E^D) allele in the Norwegian Dala breed. The presence of these two mutations was also observed in other sheep breeds: Corriedale, Damara, Black Merino, Black Castellana, and Karakul (Våge et al., 2003; Royo et al., 2008). Pharmacological characterization of these two amino acid substitutions revealed that the p.M73K substitution alone was able to constitutively activate the receptor, probably increasing the stability of the high affinity activated state. However, according to the phenotype observed in other species having the same two mutated positions (chicken p.M73K, Takeuchi et al., 1996; pig, p.D121N, Kijas et al., 1998), it seems reasonable to suppose that any of the two mutations could be sufficient to determine constitutive eumelanin synthesis. Another hypothesis could be that each of these mutations could produce a weaker activation of the receptor similar to that found in Alaska silver fox and only the presence of both mutated positions or homozygosity for one or the other substitution would be requested to obtain entirely black animals (Våge et al., 2003). The alternative form of

the *MC1R* gene identified by Våge et al. (1999) should correspond to the wild type *Extension* allele.

Recently, Norris and Whan (2008) characterized the sheep *ASIP* gene showing that a 190-kb tandem duplication encompassing this gene, the *AHCY* coding region (CDS) and the *ITCH* promoter region should be the cause of the white coat colour of dominant white and tan (A^{Wt}) *Agouti* sheep. In addition, a not yet characterized regulatory mutation as well as a deletion of 5 bp in exon 2 and a missense mutation in exon 4 were identified as the causes of the black recessive non-agouti (A^a) allele (Gratten et al., 2010; Norris and Whan, 2008; Royo et al., 2008). Analysis of the *ASIP* gene in a few Italian sheep breeds has been obtained by Fontanesi et al. (in press-a). Animals homozygous for the A^a allele would not be phenotypically distinguishable from animals carrying the E^D allele and the contemporary presence of both these alleles have been proposed in a few black breeds (Roberts and White, 1930). However, the role of the dominant A^{Wt} allele and of the duplicated *ASIP* gene in determining white coat colour in sheep has been disputed suggesting that other genetic factors, besides the *Agouti* alleles, are involved in determining the white phenotype in several ovine breeds (Renieri et al., 2008).

In this study we sequenced the *MC1R* gene in several Italian sheep breeds and populations identifying new mutations, one of which is possibly affecting coat colour. Putative interactions with other coat colour loci are discussed.

2. Materials and methods

2.1. Animals and samples

Blood, hair or milk samples were collected from a total of 388 sheep: 15 Appenninica (from 2 flocks), 14 Bergamasca (from 1 flock), 21 Comisana (from 3 flocks), 12 Cornigliese-like (from 1 flock), 8 Delle Langhe (from 1 flock), 91 Massese (from 3 flocks), 10 Merinizzata Italiana (from 1 flock), 41 Sarda (from 2 flocks) and 176 Valle del Belice (from 4 flocks) sheep. Coat colour descriptions were available for all animals. Coat colours of the sheep included in this study were: Appenninica, white hair and unpigmented skin; Bergamasca, white hair and unpigmented skin; Comisana, white hair and unpigmented skin with red cheeks, ears and part of the neck; Cornigliese-like, white hair and unpigmented skin with black spots in the head (8 sheep) or with white hair and unpigmented skin (2 sheep) or with black hair and pigmented skin (1 sheep) or with pigmented skin and white hair (1 sheep); Delle Langhe, white hair and unpigmented skin; Massese, black hair (78 sheep) or grey hair (13 sheep) and pigmented skin; Merinizzata Italiana, white wool and unpigmented skin; Sarda, white hair and unpigmented skin (40 sheep) or black hair and pigmented skin (1 sheep); Valle del Belice, white hair and unpigmented skin with sometimes red/brown or black spots in the head (cheeks and ears). Pictures of the head were available for 79 Valle del Belice sheep (Fig. 1).

2.2. DNA extraction and sequencing

The sequencing panel was constituted of 73 sheep (a subset of the animals reported in the previous section): 2 Appenninica, 3 Bergamasca, 10 Comisana, 5 Cornella (3 with white hair and with black spots in the face and in the body; 1 with black hair and pigmented skin; 1 with white hair and unpigmented skin), 2 Delle Langhe, 24 Massese (18 with black hair; 6 with grey hair), 4 Merinizzata Italiana, 13 Sarda (12 white hair and unpigmented skin, 1 with black hair) and 10 Valle del Belice (3 white hair and unpigmented skin and 7 with red or dark patches in the head) sheep. DNA was extracted using a standard phenol-chloroform protocol for blood, the Wizard® Genomic DNA Purification kit for blood and milk (Promega Corporation, Madison, WI, USA) or rapid extraction methods for milk and hair roots (Russo et al., 2007). Primers for *MC1R* amplification and sequencing (Table 1) were designed with Primer 3 (Whitehead



Fig. 1. Pictures of two Valle del Belice sheep: A) completely white; B) with pale red spots in the head. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of the article.)

Institute for Biomedical Research, Cambridge, MA, USA) from the published bovine, sheep and goat *MC1R* gene sequences (GenBank accession numbers: AF445641, Y13965, and FM212940, respectively; Rouzaud et al., 2000; Våge et al., 1999; Fontanesi et al., 2009a). PCR was performed using a TGradient thermal cycler (Biometra, Göttingen, Germany) or a PT-100 thermal cycler (MJ Research, Watertown, MA, USA) in a volume of 20 μ L containing 10–100 ng DNA template, 1 U DNA EuroTaq DNA polymerase (EuroClone Ltd., Paington, Devon, UK), 1 \times PCR Buffer, 2.5 mM dNTPs, 10 pmol of each primer and optimised $MgCl_2$ concentrations (from 1.0 to 2.0 mM). PCR profile was as follows: 5 min at 95 $^{\circ}$ C; 35 amplification cycles of 30 s at 95 $^{\circ}$ C, 30 s at the appropriate annealing temperature (Table 1), 30 s at 72 $^{\circ}$ C; 5 min at 72 $^{\circ}$ C. For the *MC1R* fragments sequencing 3–5 μ L of PCR product was treated with 2 μ L of ExoSAP-IT[®] (USB Corporation, Cleveland, Ohio, USA) following the manufacturer's protocol. Cycle sequencing of the PCR products was obtained with the Big Dye v3.1 kit (Applied Biosystems, Foster City, CA, USA) and sequencing reactions, after purification steps carried out with EDTA 0.125 M, Ethanol 100% and Ethanol 70%, were loaded on an ABI3100 Avant sequencer (Applied Biosystems). All sequences were visually inspected, edited, assembled, and aligned using the BioEdit v. 7.0.5.2 (<http://www.mbio.ncsu.edu/BioEdit/bioedit.html>) and the CodonCode Aligner (<http://www.codoncode.com/aligner>) soft-

ware. Nomenclature of the detected single nucleotide polymorphisms (SNPs) was obtained following the Nomenclature for the description of sequence variation (<http://www.hgvs.org/mutnomen/>).

2.3. Genotyping

Three missense mutations were genotyped by PCR-RFLP: one was a novel mutation identified in this study (c.199C>T, p.R67C); two other mutations (c.218T>A, p.M73K; and c.361G>A, p.D121N) cause the E^D allele. PCR amplification was carried out with primer pairs used for sequencing parts of the *MC1R* gene (Table 1). PCR conditions were as already reported above. All digested products (5 μ L of PCR product in a total of 25 μ L of reaction volume including 5 U of restriction enzyme and 1 \times reaction buffer) were electrophoresed in 10% polyacrylamide:bisacrylamide 29:1 TBE 1 \times gels. DNA fragments were visualized with 1 \times GelRed Nucleic Acid Gel Stain (Biotium Inc., Hayward, CA, USA). The c.199C>T SNP creates/disrupts a *SsiI* endonuclease restriction site (CCGC). The 169 bp amplified fragment was digested at 37 $^{\circ}$ C overnight with *SsiI* restriction enzyme (Fermentas, Vilnius, Lithuania). Electrophoretic pattern for the c.199C allele consisted of three fragments of 82, 71, and 16 bp while allele c.199T consisted of only two fragments

Table 1

Primer sequences and PCR conditions for *MC1R* sequencing and PCR-RFLP analyses.

Primer pair	Forward and reverse primers (5'-3')	Fragment length (bp)	PCR conditions ^a	Use
1	agtgcctggaggtgtccatcc ctgagcgtcaccagcaagt	169 bp	62/2.0/PT	PCR-RFLP analysis with <i>SsiI</i> (c.199C>T; p.R67C) and <i>HinIII</i> (c.218T>A; p.M73K)
2	agccatgagttgagcaggac caggacaccagctccag	376 bp	64/1.0/PT	Sequencing
3	gtgagcgtcagcaacgtg acatagaggacggccatcag	366 bp	61/2.0/TG	Sequencing and PCR-RFLP analysis with <i>Tru1I</i> (c.361G>A; p.D121N)
4	gcctgggttgctcttcata tggctagcgatcctctttg	456 bp	58/2.0/TG	Sequencing

^a PCR conditions describe annealing temperature ($^{\circ}$ C)/ $MgCl_2$ concentration (mM)/Thermal cycler (TG = TGradient thermal cycler - Biometra, Göttingen, Germany; PT = PT-100 thermal cycler - MJ Research, Watertown, MA, USA).

of 87 and 82 bp respectively. For the c.218T>A mutation, the amplified fragment of 169 bp was digested with *Hin1I* endonuclease (Fermentas; recognition sequence: CATG) at 37 °C. Allele c.218A was characterized by two DNA fragments (146 + 23 bp), whereas allele c.218T consisted of three DNA fragments (119 + 27 + 23 bp). The analysis of the c.361G>A mutation was obtained digesting a 366 bp amplified *MC1R* gene region with *Tru1I* restriction enzyme (Fermentas; recognition sequence, TTAA) at 65 °C overnight. The resulting PCR-RFLP pattern was characterized by one fragment of 366 bp (allele c.361G) or two fragments of 265 + 101 bp (allele c.361A).

2.4. Data analysis

In silico functional analysis of the novel identified missense mutation was obtained using the SIFT (Sorting Intolerant From Tolerant) program (Ng and Henikoff, 2003) and the evolutionary analysis of coding SNPs tool of PANTHER (Protein ANALYSIS THrough Evolutionary Relationships; Thomas et al., 2003; Thomas and Kejariwal, 2004), whose predictions have been experimentally validated (Brunham et al., 2005). SIFT is a sequence homology-based tool that sorts intolerant from tolerant amino acid substitutions and predicts whether an amino acid substitution at a particular position in a protein will affect protein function and hence, potentially alter the phenotype. Positions with normalized probabilities <0.05 are predicted to be deleterious. To evaluate this prediction, SIFT calculates the median conservation value, which measures the diversity of the sequences in the alignment. Conservation is calculated for each position in the alignment and the median of these values is obtained. Conservation ranges from $\log_2 20 = 4.32$, when a position is completely conserved and only one amino acid is observed, to zero, when all 20 amino acids are observed at a position. Predictions based on sequence alignments with median conservation values higher than 3.25 are less diverse and will have a higher false positive error. Default options were used in our analysis. PANTHER estimates the likelihood of a particular non-synonymous (amino-acid changing) coding SNP to cause a functional impact on the protein. It calculates the substitution position-specific evolutionary conservation (subPSEC) score based on an alignment of evolutionarily related proteins (Thomas et al., 2003, 2006; Thomas and Kejariwal, 2004). The probability that a given variant will cause a deleterious effect on protein function is estimated by $P_{\text{deleterious}}$, such that a subPSEC score of -3 corresponds to a $P_{\text{deleterious}}$ of 0.5 (Brunham et al., 2005). The subPSEC score is the negative logarithm of the probability ratio of the wild type and mutant amino acids at a particular position. PANTHER subPSEC scores are continuous values from 0 (neutral) to about -10 (most likely to be deleterious).

RNAstructure software v. 4.6 (Mathews et al., 2004) was used to predict secondary structures and stability (expressed as minimum free energy) of the different sheep *MC1R* transcript alleles using default options. The 5'-untranslated region (5'-UTR) of the sheep *MC1R* gene (composed of only one exon as in all other mammals) was deduced aligning the obtained *MC1R* sequences with that of the cattle *MC1R* cDNA (GenBank accession number: AF445642; Rouzaud et al., 2000).

For the sequenced animals, haplotypes including the identified SNPs within the sheep *MC1R* gene were inferred using the PHASE program v. 2.1 (Stephens et al., 2001).

A median-joining network (Bandelt et al., 1999) for these haplotypes was constructed using Network v. 4.510 (<http://www.fluxus-engineering.com>), including the wild type goat *MC1R* coding sequence (GenBank accession number: FM212940; goat haplotype 4 in Fontanesi et al., 2009a) to orient the network.

Evaluation of Hardy-Weinberg equilibrium was carried out with chi-square analysis from the HWE software program (Linkage Utility Programs, Rockefeller University, New York, NY, USA).

3. Results

3.1. Identification of polymorphisms and *in silico* analysis of their effects

Sequencing 1045 bp of the *MC1R* gene (including the whole CDS of 954 bp, 38 bp of the 5'-UTR, and 53 bp of the 3'-UTR) in 73 sheep of different breeds revealed 5 novel SNPs that, together with the polymorphisms reported by

Table 2
SNPs identified in the sheep *MC1R* gene.

SNP	Gene region (amino acid change)	Reference
c.-31G>A	5'-UTR	This study
c.199C>T	CDS (p.R67C)	This study
c.218T>A	CDS (p.M73K)	Våge et al. (1999), this study
c.361G>A	CDS (p.D121N)	Våge et al. (1999), this study
c.429C>T	CDS (synonymous)	This study
c.600T>G	CDS (synonymous)	This study
c.735T>C	CDS (synonymous)	This study

Våge et al. (1999), increased substantially the number of markers of the *MC1R* gene in this species (Table 2). Three Massese sheep showed the c.218T>A and the c.361G>A substitutions causing the p.M73K and the p.D121N amino acid changes of the E^D allele (Våge et al., 1999). Of the novel SNPs, one (c.-31G>A) was located in the 5'-UTR region, whereas the others were situated in the coding region. Of the latter 4 polymorphisms, 3 were synonymous SNPs (c.429C>T; c.600T>G; c.735C>T) and one was a non-synonymous mutation (c.199C>T) causing an amino acid substitution (p.R67C) in the first intracellular loop of the *MC1R* protein (Fig. 2) in a highly conserved position across species (Fig. 3).

In silico prediction of the functional effect of the p.R67C amino acid substitution obtained with PANTHER indicated that it is highly probable that this amino acid change causes a deleterious effect on the normal functionality of the *MC1R* protein (subPSEC = -6.68827 ; $P_{\text{deleterious}}$ = 0.9756). The results obtained with SIFT indicated that the p.R67C substitution is predicted to affect protein function with a score of 0.01 (median sequence conservation = 3.22; with 396 other sequences represented at this position, data not shown), confirming the prediction of PANTHER.

Due to its location in the 5'-UTR, the c.-31G>A could have a functional effect altering the secondary structure of this region that, as consequence, can affect translation efficiency and mRNA stability. To evaluate this hypothesis, RNA secondary structure and stability of the two different 5'-UTR carrying the alternative alleles at this polymorphic

	51		▼		80
Sheep	VSLVENLWV	AAIAK	IVNLH	SPMYIFICCL	
Sheep_e	C...	
Goat	
Cattle	
Cat	..V.....	
Dog	..V.....G..	
Fox	..V.....G..	
Horse	T.....	
HumanA...	T.....C....	
ElephantM...	
Mouse	I..T.....	
RabbitL...	V.....C....	
ChickenL...	..L.....	..T.....	

Fig. 2. Alignment of the deduced sheep *MC1R* protein around the p.R67C substitution with the same protein regions of other species (GenBank accession numbers: goat, FM212940; cattle, Y19103; cat, FM180571; dog, AF064455; fox, AJ786717; horse, AF288357; human, AY363625; elephant, DQ648866; mouse, BC119294; rabbit, AM180878; chicken, AY220304). Dots indicate the same amino acid of the sheep protein.

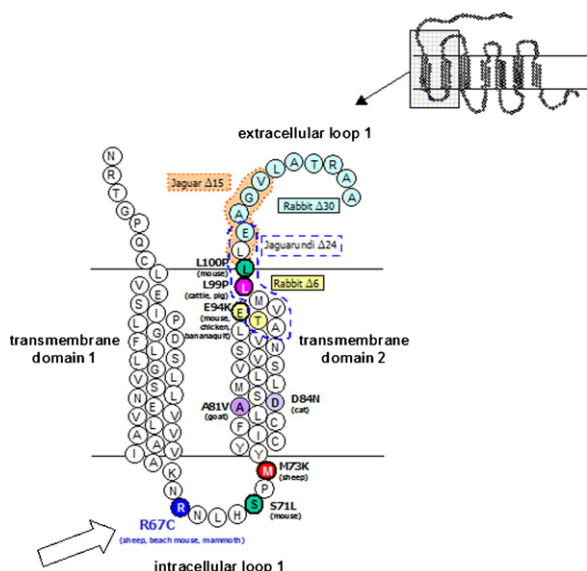


Fig. 3. Two dimensional structure of a portion of the sheep MC1R protein in which the p.R67C substitution occurs (indicated with an arrow). Amino acid substitutions and deletions identified in other species are indicated. Circles bordered by a thick black line represent the activating amino acid substitutions that have been previously reported to be associated to dominant eumelanin phenotype in mice (Robbins et al., 1993), cattle (Klungland et al., 1995), chicken (Takeuchi et al., 1996), pig (Kijas et al., 1998), sheep (Våge et al., 1999), and bananaquit (Theron et al., 2001). The deletions observed in jaguar and jaguarundi (Eizirik et al., 2003) have been surrounded by a dotted or a bold dotted line, respectively. The two deletions identified in rabbit have been reported in different colours (Fontanesi et al., 2006). A putative functional substitution identified in goat and the substitution causing the *Extension e* allele in cat are reported (Fontanesi et al., 2009a; Peterschmitt et al., 2009). The inset at the right top shows the secondary structure diagram of the whole protein with evidenced the region enlarged below it. The two dimensional-model of the MC1R protein was modified from Majerus and Mundy (2003). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of the article.)

site was analysed using the RNA structure software. However, no substantial differences were observed in minimum free energy and secondary structure considering the two alternative 5'-UTR (−129.6 and −129.7 kcal/mol for the c.-31G allele and for the c.-31A allele, respectively).

3.2. Sheep MC1R haplotypes and genotyping results

The five novel identified SNPs and the two missense mutations of the E^D allele (Table 2) were organized in four haplotypes (submitted to EMBL/GenBank with the accession numbers FN600553, FN600554, FN600555, and FN600556, respectively) (Table 3). If the 5'-UTR polymorphic site is omitted (the 5'-UTR of the sheep MC1R gene was not reported in previous studies), haplotype 1 (c.-31G, c.199C, c.218T, c.361G, c.429C, c.600T, and c.735C nucleotides; [GCTGCTC]) corresponded to that identified by Våge et al. (1999), representing the wild type MC1R sequence (Table 3).

This haplotype was observed in all sequenced breeds except in Massese. Haplotype 2 [GTTGCTC] differs from haplotype 1 for the presence of the c.199T allele only, determining the p.R67C amino acid change. This haplotype was

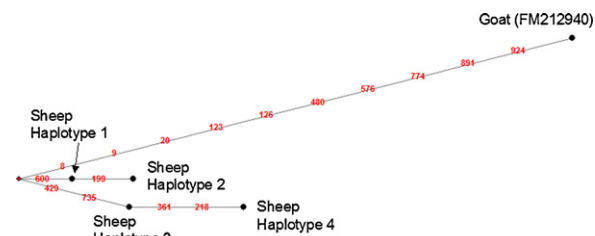


Table 3
MC1R haplotypes identified in the sequenced sheep panel.

Breed/Population ^a	No. of sequences ^b	Haplotype 1 [GCTGCTC]	Haplotype 2 [GTTGCTC]	Haplotype 3 [ACTGTGT]	Haplotype 4 [ACAATGT]
Appenninica	4	4	–	–	–
Bergamasca	6	5	–	1	–
Comisana	20	10	–	10	–
Cornigliese-like	10	8	–	2	–
Delle Langhe	4	3	–	1	–
Massese	48	–	–	45	3
Merinizzata Italiana	8	5	–	3	–
Sarda	26	21	–	5	–
Valle del Belice	20	10	8	2	–
Total	146	66	8	69	3

Haplotypes are indicated following the SNP positions in the *MC1R* gene: c.-31G>A, c.199C>T, c.218T>A, c.361G>A, c.429C>T, c.600T>G, and c.735C>T. Haplotype 2 is the putative *e* allele. Haplotype 4 represents the *E^D* allele.

^a The coat colour of the sequenced sheep was: white hair and unpigmented skin for all Appenninica, Bergamasca, Delle Langhe, and Merinizzata Italiana sheep, and 1 Cornigliese-like, 13 Sarda, and 3 Valle del Belice sheep; white hair with black spots in the face and in the body for 3 Cornigliese-like sheep; white hair and unpigmented skin with red or dark patches in the head for 7 Valle del Belice sheep; white hair with unpigmented skin with red cheeks, ears and part of the neck for all Comisana sheep; black hair with pigmented skin for 1 Cornigliese-like, 18 Massese and 1 Sarda sheep; grey hair and pigmented skin for 6 Massese sheep.

^b Two haplotype sequences for each animal.

segregation) or across breeds with different coat colour phenotypes.

Valle del Belice flocks usually include completely white animals together with animals having coloured patches in the head (around the eyes and/or in the ears or cheeks) (Fig. 1) probably due to incomplete fixation of different alleles at the *Spotting* locus (epistatic over other loci), as suggested for other breeds (Sponenberg, 1997; Renieri et al., 2008). The *Spotting* locus or other loci with similar phenotypic effects might act through inhibition or dysregulation of melanocyte migration from the neural crest at the embryonic level. This complicates the interpretation of the results as a complete characterization of the *Spotting* locus in sheep is lacking. Therefore, when spots are present (that means presence of melanocytes in these areas) it could be possible to evaluate if different mutations are associated with the presence of eumelanin or pheomelanin colours. To test this hypothesis we considered only the 79 Valle del Belice sheep for which we had pictures of the head. Of these animals, 17 were completely white, whereas 53 had red/brown spots and 9 had black/dark brown spots around the eyes, cheeks and/or ears. Of the three animals homozygous for allele c.199T (p.67C; we had pictures for only 3 out of 11 sheep with this genotype, see previous paragraph) two had red or brown spots and one was completely white. Red spots were also present in sheep heterozygous at this polymorphic site or homozygous for the alternative allele (data not shown). Therefore, there was no complete association between the genotypes at the p.R67C substitution and coat colour in this breed.

In the Massese breed, all 5 sheep carrying the *E^D* allele were black, indicating that this allele may have an effect in determining the black coat colour in this breed. However, other animals of this breed that had the black coat colour phenotype (no. 75) did not carry this allele. Sequencing and genotyping black and grey animals did not reveal any difference in the *MC1R* gene (apart from the presence of the *E^D* allele in five animals). In this breed the *ASIP* gene seems responsible for the black coat colour in most animals (Fontanesi et al., in press-a).

Coat colour heterogeneity was present in the Cornigliese-like population but no association with the presence of different *MC1R* haplotypes was observed in the sequenced animals (Table 3).

The only black Sarda sheep carried haplotype 1 and haplotype 3. Other four completely white sequenced sheep of this breed had the same haplotype combination.

Analysing the *MC1R* results across breeds, it was interesting to note that haplotype 3 was almost fixed in the Massese breed, but this haplotype was also present in almost all other breeds. The *in silico* analysis did not predict any functional effect of the SNPs included in haplotype 3. Therefore, the high frequency of this haplotype in the Massese breed might be interpreted as a possible result of different mechanisms, other than the direct selection on black/grey coat colour, that shaped the genetic pool of this sheep breed. Heterogeneity of *MC1R* haplotype combination was evident in several other breeds (Table 3) suggesting that haplotype 1 and haplotype 3 might not have functional effects on coat colour and could be considered wild type alleles at the *Extension* locus.

4. Discussion

Classical genetic studies identified two alleles at the *Extension* locus in sheep: *E^D* and *E⁺* (Searle, 1968; Sponenberg, 1997). Allele *E^D*, that was subsequently characterized at the molecular level (Våge et al., 1999), is directly involved in affecting sheep pigmentation. This allele causes the dominant black coat colour. The polymorphisms we identified in the *MC1R* gene defined four haplotypes, two of which (haplotype 1 and haplotype 3) could be considered wild type alleles at the *Extension* locus. These two haplotypes are present in most of the analysed breeds. It seems unlikely that other regulatory mutations (present in non sequenced regions) in linkage disequilibrium with haplotype 3 might affect coat colour, as this haplotype was present in different breeds with different colours. However, it is interesting to note that the two missense mutations causing the *E^D* allele were inserted in a

haplotype 3-derived haplotype (haplotype 4). This information was not reported by Våge et al. (1999).

The p.R67C substitution (haplotype 2) may have important functional effects. This mutation causes an amino acid change in a highly conserved position of the first intracellular loop of the MC1R protein (Figs. 2 and 3). The same amino acid substitution has been identified in a beach mouse subspecies (*Peromyscus polionotus*; p.R65C position in mice due to the upstream lack of two amino acids) having adaptive pale coat colour (Hoekstra et al., 2006). *In vitro* functional tests (cyclic AMP and ligand-binding assays) of the two beach mouse alleles indicated a reduction in basal and stimulated cAMP formation caused by the p.65C allele (Hoekstra et al., 2006). As eumelanin production is associated with MC1R-mediated cAMP formation, reduced cAMP synthesis was the indirect cause of the lack of dark pigmentation in mice expressing the p.65C allele (Hoekstra et al., 2006). The p.R67C amino acid change was also identified in mammoth specimens (Römpler et al., 2006). *In vitro* functional analysis of MC1R constructed expression vectors revealed the mammoth p.R67C substitution as the main cause for MC1R dysfunction as also evidenced in beach mice (Römpler et al., 2006). Therefore, combining *in silico* prediction results and *in vitro* functional tests carried out in other species, it is highly probable that the sheep p.R67C amino acid change gives pheomelanin coat colour. This is confirmed in Valle del Belice sheep homozygous for the p.67C allele, in which no complete epistatic inhibition of coat colour could be attributed to other loci. The two sheep homozygous for this allele that presented coloured spots in the head showed a pheomelanin-like phenotype. Thus, we propose that the p.67C allele represents the *e* allele at the ovine *Extension* series that was, so far, not completely recognized in sheep by classical genetic studies (Searle, 1968; Sponenberg, 1997), probably due to its presence in few breeds (we have identified it only in the Valle del Belice breed), its low frequency (21.3% in this breed), and because it should be recessive. It will be interesting to genotype other sheep breeds to have a complete evaluation of its distribution and to infer its possible origin.

However, it is worthwhile to mention that the pheomelanin-like phenotype in Valle del Belice sheep does not seem to be caused only by the identified MC1R mutation. A similar complexity in determining the reddish coat colour has been identified in goats in which MC1R as well as other genes have been suggested to cause red/pale phenotypes (Fontanesi et al., 2009a).

Another issue that should be considered is the degree of redness or paleness that this new allele might give to the coat colour of the homozygous animals, depending on their genetic background. For example, in cattle several breeds carrying the *e* allele show a wide range of redness/paleness (from brown to white) (Klungland et al., 1995; Maudet and Taberlet, 2002; Rouzaud et al., 2000; Russo et al., 2007). The *A^{Wt}* allele has been suggested to determine, at least in part, the white pheomelanin coat colour (Sponenberg, 1997). The molecular basis of the *A^{Wt}* allele has been recently determined by Norris and Whan (2008) who identified a copy number variation affecting the *ASIP*, *AHCY*, and *ITCH* genes. However, it seems evident that other loci (e.g. *Spotting* locus) are involved in determining white coat colour

in sheep (Renieri et al., 2008), as also supposed in goats (Fontanesi et al., 2009b). Analysing the *ASIP* gene in several sheep breeds, we confirmed this hypothesis (Fontanesi et al., in press-a). Moreover the results we obtained for the MC1R gene can indirectly confirm that different loci other than the *Agouti* locus determine white hair. This derives from the fact that the expression of any mutated *Agouti* alleles would require a wild type *Extension* allele (Searle, 1968).

In silico and functional evaluations of the p.67C MC1R allele strongly indicate that this mutation inhibits the normal function of the MC1R protein, epistatic over the *Agouti* locus, independently to the presence of any *Agouti* allele. However, Valle del Belice sheep homozygous for the p.67C allele have white coat colour in the body except for a few small red or brown spots in the head (no. 2) or are completely white (no. 1). This could be an indirect proof of the fact that the *Agouti* locus would not have any direct effect on this coat colour and (an)other locus (loci) (epistatic over *Extension* and *Agouti*) should be involved in determining white coat colour in these animals. Epistatic effects of white caused by *KIT* gene alleles have been already described in pigs (Fontanesi et al., in press-b; Marklund et al., 1998).

5. Conclusion

The results we obtained confirmed, in general, that the MC1R gene is not the most important genetic factor that determines hair colour in sheep, at least in the analysed Italian breeds. Mutations in the *ASIP* gene seem to account for a larger number of coat colour phenotypes (Fontanesi et al., in press-a; Gratten et al., 2010; Norris and Whan, 2008; Royo et al., 2008). In addition, other loci should be involved in defining white hair with or without the combined effect of the *A^{Wt}* *Agouti* allele. However, the presence of a putative *e* allele at the *Extension* locus could open new perspectives for the constitution of specialized sheep lines for the production of coloured hair, as requested by wool market niches. Introgression of this new allele in black sheep homozygous for non-*agouti* alleles could make its expression possible without any confounding effects of other loci.

Acknowledgments

We thank Associazione Provinciale Allevatori of Bologna, Ferrara and Forlì and several people who helped in the sampling. This work was supported by the Assessorato Agricoltura e Foreste of the Regione Siciliana – U.O.B. 108, SOAT no. 69 Aragona (AG) and funded by the Italian MiPAF SELMOL project.

References

- Bandelt, H.J., Forster, P., Röhl, A., 1999. Median-joining networks for inferring intraspecific phylogenies. *Mol. Biol. Evol.* 16, 37–48.
- Bennett, D.C., Lamoreux, M.L., 2003. The color loci of mice—a genetic century. *Pigment Cell Res.* 16, 333–344.
- Bultman, S.J., Michaud, E.J., Woychik, R.P., 1992. Molecular characterization of the mouse *agouti* locus. *Cell* 71, 1195–1204.
- Brunham, L.R., Singaraja, R.R., Pape, T.D., Kejariwal, A., Thomas, P.D., Hayden, M.R., 2005. Accurate prediction of the functional significance of single nucleotide polymorphisms and mutations in the *ABCA1* gene. *PLoS Genet.* 1, e83.

- Eizirik, E., Yuhki, N., Johnson, W.E., Menotti-Raymond, M., Hannah, S.S., O'Brien, S.J., 2003. Molecular genetics and evolution of melanism in the cat family. *Curr. Biol.* 13, 448–453.
- Fontanesi, L., Beretti, F., Riggio, V., Dall'Olio, S., Gómez González, G., Finocchiaro, R., Davoli, R., Russo, V., Portolano, B., 2009a. Missense and nonsense mutations in melanocortin 1 receptor (*MC1R*) gene of different goat breeds: association with red and black coat colour phenotypes but with unexpected evidences. *BMC Genet.* 10, 47.
- Fontanesi, L., Beretti, F., Riggio, V., Gómez González, E., Dall'Olio, S., Davoli, R., Russo, V., Portolano, B., 2009b. Copy number variation and missense mutations of the agouti signaling protein (*ASIP*) gene in goat breeds with different coat colours. *Cytogenet. Genome Res.* 126, 333–347.
- Fontanesi, L., Dall'Olio, S., Beretti, F., Portolano, B., Russo, V., in press-a. Coat colours in the Massese sheep breed are associated with mutations in the agouti signaling protein (*ASIP*) and melanocortin 1 receptor (*MC1R*) genes. *Anim.*
- Fontanesi, L., D'Alessandro, E., Scotti, E., Liotta, L., Crovetto, A., Chiofalo, V., Russo, V., in press-b. Genetic heterogeneity and selection signature at the *KIT* gene in pigs showing different coat colors and patterns. *Anim. Genet.*
- Fontanesi, L., Tazzoli, M., Beretti, F., Russo, V., 2006. Mutations in the melanocortin 1 receptor (*MC1R*) gene are associated with coat colours in the domestic rabbit (*Oryctolagus cuniculus*). *Anim. Genet.* 37, 489–493.
- Gratten, J., Pilkington, J.G., Brown, E.A., Beraldi, D., Pemberton, J.M., Slate, J., 2010. The genetic basis of recessive self-colour pattern in a wild sheep population. *Heredity* 104, 206–214.
- Hoekstra, H.E., Hirschmann, R.J., Bundey, R.A., Insel, P.A., Crossland, J.P., 2006. A single amino acid mutation contributes to adaptive beach mouse color pattern. *Science* 313, 101–104.
- Joerg, H., Fries, H.R., Meijerink, E., Stranzinger, G.F., 1996. Red coat color in Holstein cattle is associated with a deletion in the *MSHR* gene. *Mamm. Genome* 7, 317–318.
- Kijas, J.M.H., Moller, M., Plastow, G., Andersson, L., 2001. A frameshift mutation in *MC1R* and a high frequency of somatic reversions cause black spotting in pigs. *Genetics* 158, 779–785.
- Kijas, J.M.H., Wales, R., Törnsten, A., Chardon, P., Moller, M., Andersson, L., 1998. Melanocortin receptor 1 (*MC1R*) mutations and coat color in pigs. *Genetics* 150, 1177–1185.
- Klungland, H., Våge, D.I., Gomez-Raya, L., Adalsteinsson, S., Lien, S., 1995. The role of melanocyte-stimulating hormone (MSH) receptor in bovine coat color determination. *Mamm. Genome* 6, 636–639.
- Lu, D., Willard, D., Patel, I.R., Kadwell, S., Overton, L., Kost, T., Luther, M., Chen, W., Woychik, R.P., Wilkison, W.O., Cone, R.D., 1994. Agouti protein is an antagonist of the melanocyte-stimulating-hormone receptor. *Nature* 371, 799–802.
- Majerus, M.E.N., Mundy, N.I., 2003. Mammalian melanism: natural selection in black and white. *Trends Genet.* 19, 585–588.
- Marklund, L., Johansson Moller, M., Sandberg, K., Andersson, L., 1996. A missense mutation in the gene for melanocyte-stimulating hormone receptor (*MC1R*) is associated with the chestnut coat color in horses. *Mamm. Genome* 7, 895–899.
- Marklund, S., Kijas, J., Rodriguez-Martinez, H., Rönnstrand, L., Funa, K., Moller, M., Lange, D., Edfors-Lilja, I., Andersson, L., 1998. Molecular basis for the dominant white phenotype in the domestic pig. *Genome Res.* 8, 826–833.
- Maudet, C., Taberlet, P., 2002. Holstein's milk detection in cheeses inferred from melanocortin receptor 1 (*MC1R*) gene polymorphism. *J. Dairy Sci.* 85, 707–715.
- Mathews, D.H., Disney, M.D., Childs, J.L., Schroeder, S.J., Zuker, M., Turner, D.H., 2004. Incorporating chemical modification constraints into a dynamic programming algorithm for prediction of RNA secondary structure. *Proc. Natl. Acad. Sci. U.S.A.* 101, 7287–7292.
- Nachman, M.W., Hoekstra, H.E., D'Agostino, S.L., 2003. The genetic basis of adaptive melanism in pocket mice. *Proc. Natl. Acad. Sci. U.S.A.* 100, 5268–5273.
- Ng, P.C., Henikoff, S., 2003. SIFT: predicting amino acid changes that affect protein function. *Nucleic Acids Res.* 31, 3812–3814.
- 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 Res.* 18, 1282–1293.
- Ollmann, M.M., Lamoreux, M.L., Wilson, B.D., Barsh, G.S., 1998. Interaction of Agouti protein with the melanocortin 1 receptor in vitro and in vivo. *Genes Dev.* 12, 316–330.
- Peterschmitt, M., Grain, F., Arnaud, B., Deléage, G., Lambert, V., 2009. Mutation in the melanocortin 1 receptor is associated with amber colour in the Norwegian Forest Cat. *Anim. Genet.* 40, 547–552.
- Renieri, C., Valbonesi, A., La Manna, V., Antonini, M., Lauvergne, J.J., 2008. Inheritance of coat colour in Merino sheep. *Small Rumin. Res.* 74, 23–29.
- Roberts, J.A.F., White, R.G., 1930. Colour inheritance in sheep. IV. White colour, recessive black colour, recessive brown colour, badger-face pattern and reverse badger-face pattern. *J. Genet.* 22, 165–180.
- Robbins, L.S., Nadeau, J.H., Johnson, K.R., Kelly, M.A., Roselli-Rehfuess, L., Baack, E., Mountjoy, K.G., Cone, R.D., 1993. Pigmentation phenotypes of variant extension locus alleles result from point mutations that alter MSH receptor function. *Cell* 72, 827–834.
- Römpler, H., Rohland, N., Lalueza-Fox, C., Willerslev, E., Kuznetsova, T., Rabeder, G., Bertranpetit, J., Schöneberg, T., Hofreiter, M., 2006. Nuclear gene indicates coat-color polymorphism in mammoths. *Science* 313, 62.
- Rouzaud, F., Martin, J., Gallet, P.F., Delourme, D., Goulemot-Leger, V., Amigues, Y., Ménissier, F., Levéziel, H., Julien, R., Oulmouden, A., 2000. A first genotyping assay of French cattle breeds based on a new allele of the extension gene encoding the melanocortin-1 receptor (*MC1r*). *Genet. Sel. Evol.* 32, 511–520.
- Royo, L.J., Alvarez, I., Arranz, J.J., Fernández, I., Rodríguez, A., Pérez-Pardal, L., Goyache, F., 2008. Differences in the expression of the *ASIP* gene are involved in the recessive black coat colour pattern in sheep: evidence from the rare Xalda sheep breed. *Anim. Genet.* 39, 290–293.
- Russo, V., Fontanesi, L., Scotti, E., Tazzoli, M., Dall'Olio, S., Davoli, R., 2007. Analysis of melanocortin 1 receptor (*MC1R*) gene polymorphisms in some cattle breeds: their usefulness and application for breed traceability and authentication of Parmigiano Reggiano cheese. *Ital. J. Anim. Sci.* 6, 257–272.
- Searle, A.G., 1968. Comparative Genetics of Coat Colour in Mammals. Logos Press, London.
- Sponenberg, D.P., 1997. Genetics of colour and hair texture. In: Piper, L., Ruvinsky, A. (Eds.), *The Genetics of Sheep*. CAB International, New York, pp. 51–86.
- Stephens, M., Smith, N.J., Donnelly, P., 2001. A new statistical method for haplotype reconstruction from population data. *Am. J. Hum. Genet.* 68, 978–989.
- Takeuchi, S., Suzuki, H., Yabuuchi, M., Takahashi, S., 1996. A possible involvement of melanocortin 1-receptor in regulating feather color pigmentation in the chicken. *Biochim. Biophys. Acta* 1308, 164–168.
- Theron, E., Hawkins, K., Bermingham, E., Ricklefs, R.E., Mundy, N.I., 2001. The molecular basis of an avian plumage polymorphism in the wild: a melanocortin-1-receptor point mutation is perfectly associated with the melanin plumage morph in the bananaquit, *Coereba flaveola*. *Curr. Biol.* 11, 550–557.
- Thomas, P.D., Campbell, M.J., Kejariwal, A., Mi, H., Karlak, B., Daverman, R., Diemer, K., Muruganujan, A., Narechania, A., 2003. PANTHER: a library of protein families and subfamilies indexed by function. *Genome Res.* 13, 2129–2141.
- Thomas, P.D., Kejariwal, A., 2004. Coding single-nucleotide polymorphisms associated with complex vs. Mendelian disease: evolutionary evidence for differences in molecular effects. *Proc. Natl. Acad. Sci. U.S.A.* 101, 15398–15403.
- Thomas, P.D., Kejariwal, A., Guo, N., Mi, H., Campbell, M.J., Muruganujan, A., Lazareva-Ulitsky, B., 2006. Applications for protein sequence-function evolution data: mRNA/protein expression analysis and coding SNP scoring tools. *Nucleic Acids Res.* 34, W645–W650.
- Våge, D.I., Fleet, M.R., Ponz, R., Olsen, R.T., Monteagudo, L.V., Tejedor, M.T., Arruga, M.V., Gagliardi, R., Postiglioni, A., Nattrass, G.S., Klungland, H., 2003. Mapping and characterization of the dominant black colour locus in sheep. *Pigment Cell Res.* 16, 693–697.
- Våge, D.I., Fuglei, E., Snipstad, K., Beheim, J., Landsem, V.M., Klungland, H., 2005. Two cysteine substitutions in the *MC1R* generate the blue variant of the Arctic fox (*Alopex lagopus*) and prevent expression of the white winter coat. *Peptides* 26, 1814–1817.
- Våge, D.I., Klungland, H., Lu, D., Cone, R.D., 1999. Molecular and pharmacological characterization of dominant black coat color in sheep. *Mamm. Genome* 10, 39–43.
- Våge, D.I., Lu, D., Klungland, H., Lien, S., Adalsteinsson, S., Cone, R.D., 1997. A non-epistatic interaction of agouti and extension in the fox, *Vulpes vulpes*. *Nat. Genet.* 15, 311–315.
- Valverde, P., Healy, E., Jackson, I., Rees, J.L., Thody, A.J., 1995. Variants of the melanocyte-stimulating hormone receptor gene are associated with red hair and fair skin in humans. *Nat. Genet.* 11, 328–330.