

RESEARCH

Open Access



# Molecular characterization of coat color gene in Sahiwal versus Karan Fries bovine

Talla Sridhar Goud<sup>1,2</sup>, Ramesh Chandra Upadhyay<sup>1</sup>, Vijaya Bhaskar Reddy Pichili<sup>3</sup>, Suneel Kumar Onteru<sup>4</sup> and Kiranmai Chadipiralla<sup>2\*</sup>

## Abstract

**Background:** *Melanocortin-1-receptor* gene (*MC1R*) plays a significant role in signaling cascade of melanin production. In cattle, the coat colors, such as red and black, are an outcome of eumelanin and pheomelanin pigments, respectively. The coat colors have become critical factors in the animal selection process. This study is therefore aimed at the molecular characterization of reddish-brown coat-colored Sahiwal cattle in comparison to the black and white-colored Karan Fries.

**Results:** The Sequence length of the *MC1R* gene was 954 base pairs in Sahiwal cattle. The sequences were examined and submitted to GenBank Acc.No. MG373575 to MG373605. Alignment of both (Sahiwal and Karan Fries) protein sequences by applying ClustalO multiple sequence alignment programs revealed 99.8–96.8% sequence similarity within the bovine. *MC1R* gene phylogenetic studies were analyzed by MEGA X. The gene *MC1R* tree, protein confines, and hereditary difference of cattle were derived from Ensemble Asia Cow Genome Browser 97. One unique single-nucleotide polymorphism (c.844C>A) (SNP) was distinguished. Single amino acid changes were detected in the seventh transmembrane structural helix region, with SNP at p.281 T>N of *MC1R* gene in Karan Fries cattle.

**Conclusions:** In this current research, we first distinguished the genomic sequence of the *MC1R* gene regions that showed evidence of coat variation between Indian indigenous Sahiwal cattle breed correlated with crossbred Karan Fries. These variations were found in the *Melanocortin 1 receptor* coding regions of the diverse SNPs. The conclusions of this research provide new insights into understanding the coat color variation in crossbred compared to the Indian Sahiwal cattle.

**Keywords:** Coat color, *Melanocortin 1 receptor* gene, Karan Fries cattle (*Bos taurus taurus*), Sahiwal (*Bos taurus indicus*), SNPs

## Background

Coat color is a part of the major essential features for identifying the modern breeds of cattle. Each breed has its own specific phenotypic, physiological, hormonal, and metabolic functions to sustain and adapt to diverse agro-geographical and tropical climatic conditions. Coloration is an effective part of variable phenotypical traits in a diversity of animals, and there are several hypotheses for its function, such as camouflage and signaling of

diverse detectable progression [1–3]. In mammals, coat color has been associated with their production and environmental adaptation. Darwin for the first time stated that a broad range of domestic animals shares multiple phenotypic characteristics, the most evident of which is a wide variety of coat colors [4, 5]. In vertebrates, the only character other than variable coat colors that occur The Sequence length of the *MC1R* gene was 954 base pairs. Coat color phenotypes are useful in the identification of diverse cattle breeds and other livestock. The coat color in animals is dependent on the percentage of eumelanin coating for black-brown pattern [6] to that of the pheomelanin coating for yellow-reddish pattern [7].

\* Correspondence: [cdpkiranmai@gmail.com](mailto:cdpkiranmai@gmail.com)

<sup>2</sup>Department of Biotechnology, Vikrama Simhapuri University, Andhrapradesh, Nellore 524320, India

Full list of author information is available at the end of the article

The  $\alpha$ -melanocyte-stimulating hormone ( $\alpha$ -MSH) and the melanocortin-1 receptor (*MC1R*) have been implicated to execute a pivotal function in regulating the pigment synthesis in several animal species for coat coloration [8, 9]. The hormone  $\alpha$ -MSH, upon binding with *MC1R* results in stimulating of adenylate cyclase through the G-protein coupled receptors, consequently leading to high intracellular cAMP levels. Tyrosinase is a well-known rate-limiting enzyme in the melanin pigment synthesis pathway, whose activity is dependent on the prominence of cyclic AMP [10–13]. Therefore, *MC1R* performs an essential role in the production of melanin in melanocytes. The variations within the *MC1R* gene must reveal an effect on the coat color of the diverse animal species like the cow, buffalo, Chinese yakow [14–18] sheep, goat [19, 20], pigs, horses [21, 22] foxes, dogs [23, 24], cats, and mice [25, 26].

The *MC1R* gene encoding for a 7-transmembrane domain [27] contains a single exon spanning over 954 base pairs. Its extension locus is positioned on chromosome 18 in cattle. It is well known that variation in the regulation and expression of the *MC1R* gene plays a vital role in causing variations in the pigmentation patterns [28]. For instance, differentiation in base coat color of the cattle is due to the hereditary alterations in the *MC1R* gene, traditionally termed as the extent locus, with alleles coding for black (ED), red (e), and wild-type (E+). Those alternative alleles were identified to observe a dominance model, in which ED > E+ > e. The native population allele (E+) encodes for a receptor that responds to both the  $\alpha$ -melanocyte-stimulating hormone ( $\alpha$ -MSH) ligand and its competitor, agouti-signaling protein (ASP). The effective ED allele causes a point mutation altering the amino acid from leucine to proline, and promoting the constitutive expression of an active receptor, which acts on most liable eumelanin because of its  $\alpha$ -MSH ligand binding mimicry [29].

Sahiwal cattle is a milch breed that mainly existed in the home track of Punjab, Haryana, Uttar Pradesh, and Madhya Pradesh states of India. Its milk yield ranges from 1140 to 3180 kg per lactation. It is remarkably resistant to tick-borne infections, thermo-tolerant, and perceived for its enhanced protection to internal and external parasites. The color of these animals ranges from lighter reddish-brown to a thicker predominant red occasionally with varying degrees of white patches over the neck and underline. Further, the intensity of the color normally shades in the extremities such as legs, tail, and head. In the current research study, the *MC1R* gene was amplified from the genomic DNA of Sahiwal cattle breed and compared its sequence with that of Karan Fries breed. In addition, a small cohort of Sahiwal cattle was genotyped for a SNP of the *MC1R* gene, c.844C>A.

## Methods

### Experimental environment and geographical location

All the experimental cows were housed in Climate Resilient Livestock Research Centre (CRLRC) located in Karnal district, Haryana state, India. The region is positioned at an altitude of 250 ms elevated than the mean sea level (MSL) geographically located at 29° 42" N 79° 54" E, latitude, and longitudinal view, respectively. The location experiences a huge variation in the temperatures measured between summers and winters. While the summers record a maximum temperature of about 46 °C, winter temperature drops to as low as 0 °C. The mean rainfall from July to August month is around 700 mm.

### Experimental design

A total of 30 clinically healthy Sahiwal cattle free from any physical abnormalities were selected in this study. All animals were closely scrutinized and maintained. Blood samples (5–7 ml) were collected from the animals by jugular vein perforation with the help of 18 G needle in K2-EDTA (1 mg/ml concentration) anticoagulant Vacutainer tubes® (BD Biosciences USA) under sterile conditions. The coat colors of the animals were determined by direct visual inspection and recorded at the time of blood specimen collection. Blood samples isolated were cautiously stored at – 20 °C for further molecular studies.

### Genomic DNA isolation from whole blood samples

The genomic DNA was isolated by standard proteinase K digestion process and phenol, chloroform, and isoamyl alcohol method [30]. To illustrate the quality of the isolated DNA, 200 ng of genomic DNA was loaded on 0.8% agarose gel electrophoresis along with 1 kb marker. Genomic DNA quality and intensity were also investigated by Nanodrop (Shimadzu BioSpec-nano, Japan). We achieved the DNA sample quality ranging between 1.80 and 1.85. An amount of 50–100 ng/ $\mu$ l of genomic DNA was applied as the template for the polymerase chain reaction.

### Polymerase chain reaction (PCR)

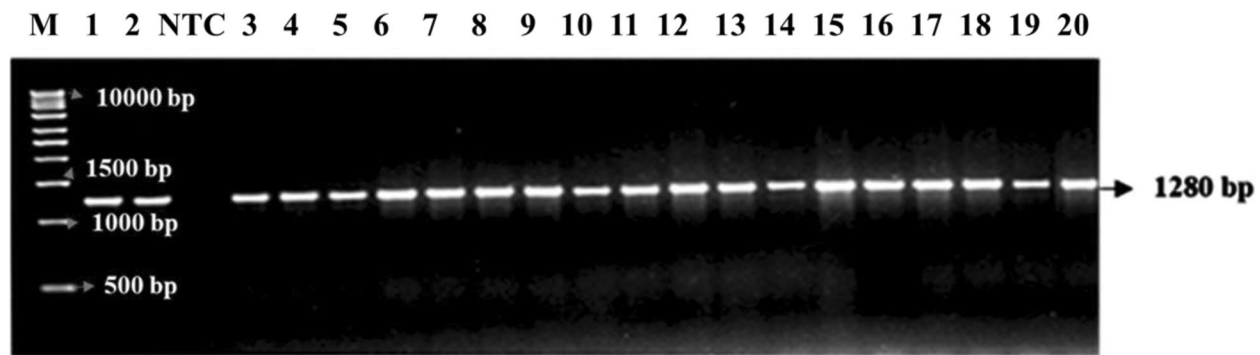
The primer sequences were designed using the reference gene sequence of *Bos taurus* cattle

(GenBank accession number: NM\_174108) and the Primer 3.0 software (Applied Biosystems, USA). To amplify the *MC1R* sequence of 1280 bp, we employed the following primers set.

Forward Primer: 5'-GGACCCTAGGAGAGCAAGCA C-3',

Reverse Primer: 5'-CTCACCTTCGAGGATGGTCT A-3'.

The PCR was performed in a reaction volume of 25  $\mu$ l, which contained 12.5  $\mu$ l 2X PCR Master mix buffer



**Fig. 1** Sahiwal Cattle *MC1R* gene PCR product amplification (M) marker 1kb (1-20) number of animals

(Sigma, USA), 1  $\mu$ l forward primer (10 pico-moles/ $\mu$ l), 1  $\mu$ l reverse primer (10 pico-moles/ $\mu$ l), 2  $\mu$ l of DNA (50–100 ng), and 8.5  $\mu$ l of nuclease-free water. The PCR protocol parameters and the procedure is as follows: 95  $^{\circ}$ C for 3 min; 35 cycles at 94  $^{\circ}$ C for 1 min, 59  $^{\circ}$ C for min, and 72  $^{\circ}$ C for 1 min; and final extension cycle at 72  $^{\circ}$ C for 10 min (Applied BioSystem Gene Amp<sup>®</sup> PCR 9700 USA). Each PCR product was separated by 1.4% agarose gel along with 1 kb molecular ladder (Sigma, USA). After cautious staining of the gel with ethidium bromide dye (0.5  $\mu$ g/ml), amplified bands were observed. Band pictures were clicked and stored by Gel documentation imaging equipment (Bio-Rad, USA) (Fig. 1).

#### *MC1R* gene PCR product sequencing and analysis

A complete quantity of 50  $\mu$ l PCR products was sequenced proceeding with 10  $\mu$ l of the forward and reverse primers of the *MC1R* gene. All the PCR products were washed in the spin column kit (Sigma, USA) with the maintenance of the ABI3730 DNA analyzer (Applied Biosystems) sequencing service done by Sci-Genom, Kerala, India, Pvt. Ltd. The sequence chromatogram ABI

files were visualized and analyzed by using Bio-Edit version 7.2. The *MC1R* gene Open reading frames were distinguished by using the ORF finder tool by NCBI (<http://www.ncbi.nlm.nih.gov/orffinder/>) and the nucleotide sequence was interpreted using the translate tool by ExPASy (<http://www.expasy.org>). The *MC1R* protein signal peptide prediction was recognized and exhibited by applying SignalP 4.0 server online (<http://www.cbs.dtu.dk/services/SignalP/>) based on the neural interface instructed on eukaryotes [31]. The bio-physicochemical properties of *MC1R* proteins were analyzed by applying ProtParam server ExPASy [32]. The integrated membrane proteins were analyzed by PROTTER online server (<http://wlab.ethz.ch/protter>) [33]. *MC1R* amino acid sequence, secondary structure, and hydrophobic surface accessibility (see Supplementary file) were predicted by NetSurfP-2.0 webserver (<http://www.cbs.dtu.dk/services/NetSurfP/>) [34]. The distinguished *MC1R* protein sequences were associated with the protein data sequence of *Bos frontalis*, *Bos grunniens*, *Bos taurus*, and *Bubalus bubalis* available at NCBI GenBank database using the pBLAST program.

**Table 1** *MC1R* protein amino acid frequency sites used in the mammalian species all frequencies are given in percent (%)

Name of the Species	Ala	Cys	Asp	Glu	Phe	Gly	His	Ile	Lys	Leu	Met	Asn	Pro	Gln	Arg	Ser	Thr	Val	Trp	Tyr	Total
<i>B. indicus</i>	9.78	4.42	1.89	1.89	5.36	5.05	2.52	7.57	1.89	17.35	2.21	3.15	4.42	4.10	4.73	5.99	4.42	9.15	1.26	2.84	317
<i>B. frontalis</i>	9.46	4.42	1.89	1.89	5.05	5.05	2.52	7.57	1.89	17.67	2.21	3.47	4.42	4.10	4.73	5.99	4.42	9.15	1.26	2.84	317
<i>B. grunniens</i>	9.78	4.42	1.89	1.89	5.05	5.05	2.52	7.57	1.89	17.67	2.21	3.47	4.42	4.10	4.73	5.99	4.10	9.15	1.26	2.84	317
<i>B. taurus</i>	9.46	4.42	1.58	1.89	4.73	5.05	2.52	7.57	1.89	17.35	2.21	3.79	4.73	4.10	5.05	5.99	4.42	9.15	1.26	2.84	317
<i>B. bubalis</i>	9.78	4.42	1.89	1.89	4.73	4.73	2.52	7.26	1.89	17.67	2.21	3.47	4.42	4.10	4.73	6.62	4.10	9.46	1.26	2.84	317
<i>C. hircus</i>	9.78	4.42	1.89	1.89	4.73	4.73	2.52	7.26	1.58	17.35	2.52	3.47	4.73	3.47	5.05	6.62	4.42	9.46	1.26	2.84	317
<i>O. aries</i>	9.46	4.42	1.89	1.89	4.73	4.73	2.52	7.26	1.58	17.35	2.52	3.47	4.42	3.79	5.05	6.62	4.42	9.78	1.26	2.84	317
<i>O. moschatus</i>	9.78	4.42	1.89	1.89	4.73	4.73	2.52	7.26	1.58	17.35	2.52	3.47	4.73	3.47	5.05	6.62	4.42	9.46	1.26	2.84	317
<i>E. caballus</i>	7.57	4.10	1.89	2.21	5.05	4.73	3.15	6.62	1.89	18.93	3.15	2.84	4.42	3.79	4.10	7.26	4.73	9.15	1.26	3.15	317
<b>Avg. %</b>	<b>9.43</b>	<b>4.38</b>	<b>1.86</b>	<b>1.93</b>	<b>4.91</b>	<b>4.87</b>	<b>2.59</b>	<b>7.33</b>	<b>1.79</b>	<b>17.63</b>	<b>2.42</b>	<b>3.40</b>	<b>4.52</b>	<b>3.89</b>	<b>4.80</b>	<b>6.41</b>	<b>4.38</b>	<b>9.32</b>	<b>1.26</b>	<b>2.87</b>	<b>317</b>

Ala alanine, Asp aspartic acid, Cys cysteine, Glu glutamic acid, Phe phenyl alanine, Gly glycine, His histidine, Ile iso-leucine, Lys lysine, Leu leucine, Met methionine, Asn asparagine, Pro proline, Gln glutamine, Arg arginine, Ser serine, Thr threonine, Val valine, Trp tryptophan, Tyr tyrosine

**Table 2** Physicochemical properties of MC1R protein sequence of the mammalian species

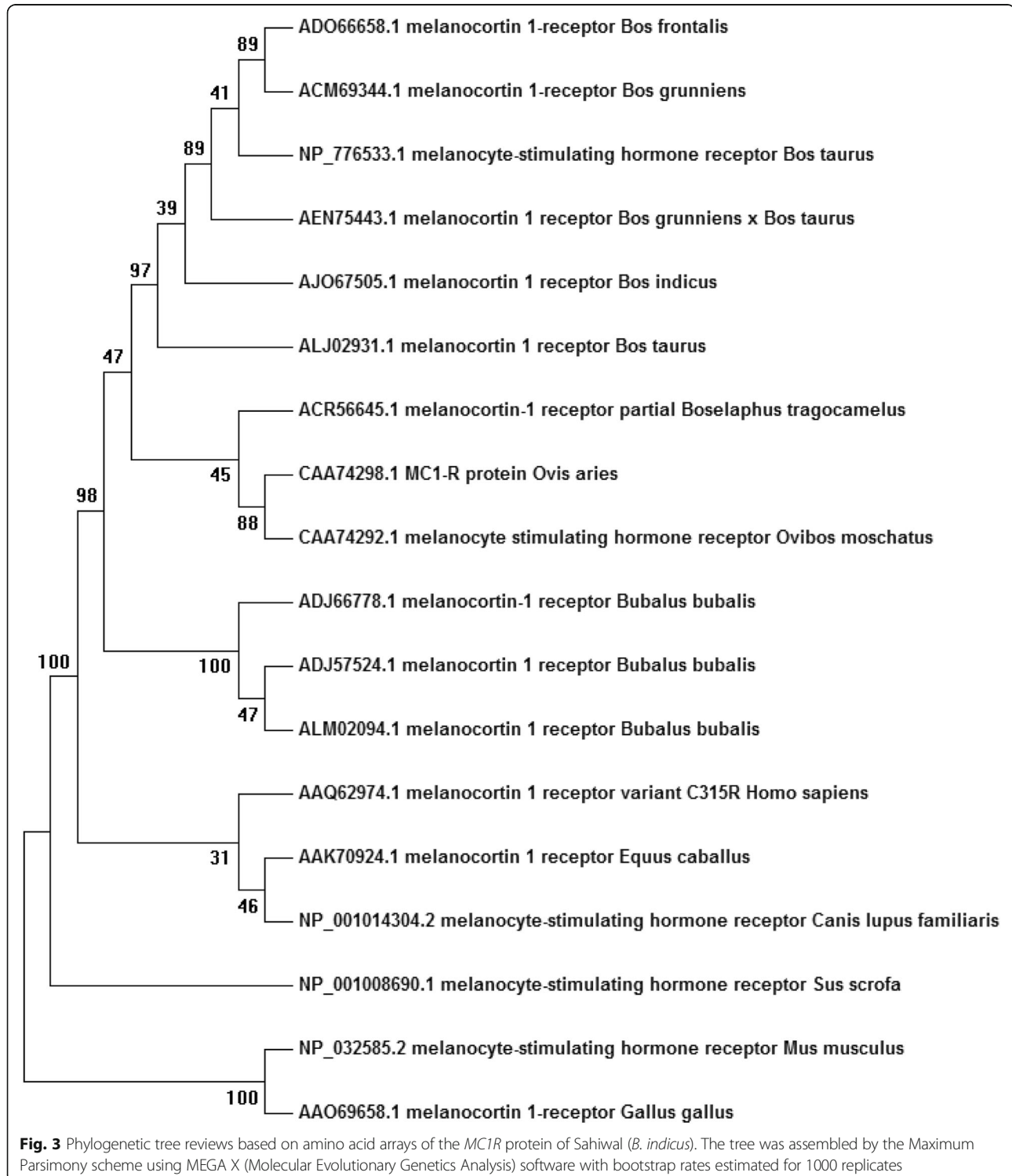
S.NO	Species	Open reading frame length in base pairs	Amino acids length	Chromosome number	NCBI accession number	% Identity with <i>Bos indicus</i>		Negatively charged residues (Asp + Glu)	Positively charged residues (Lys + Arg)	N- Glyc sites	O- Glyc sites	Phosphorylation sites
						Nucleotide	Amino acid					
1	<i>B. indicus</i>	954	317	18	MG373575	-	-	12	21	6	4	21
2	<i>B. frontalis</i>	954	317	18	HM488960	99.9	99.68	12	21	6	4	21
3	<i>B. grunnesis</i>	954	317	18	FJ624478	99.79	99.37	12	22	6	4	21
4	<i>B. taurus</i>	954	317	18	NC037345	99.2	99.37	11	22	6	2	21
5	<i>B. bubalis</i>	954	317	18	NW020228700	97.49	97.16	12	21	7	4	25
6	<i>C. hircus</i>	954	317	18	NW017189504	96.33	96.3	12	21	2	3	21
7	<i>O. aries</i>	954	317	14	NC019471	95.70	96.85	12	21	2	0	21
8	<i>O. moschats</i>	954	317	18	Y13958	96.33	96.85	12	21	6	3	21
9	<i>E. caballus</i>	954	317	3	NC009146	85.94	83.91	13	19	7	2	21



### Phylogenetic analysis

The *MC1R* protein sequences of diverse mammalian varieties were regained from the NCBI GenBank database. The phylogenetic tree study was done by MEGA X software by applying the Maximum Parsimony method. This method takes a minimum basic number of

mutations into consideration to analyze and elucidate any given series of sequences that are positioned together and does not allow any modification or amendment at every single site. Using this method, the evolutionary records were obtained from all the parsimony sites represented in the parenthesis, in which the





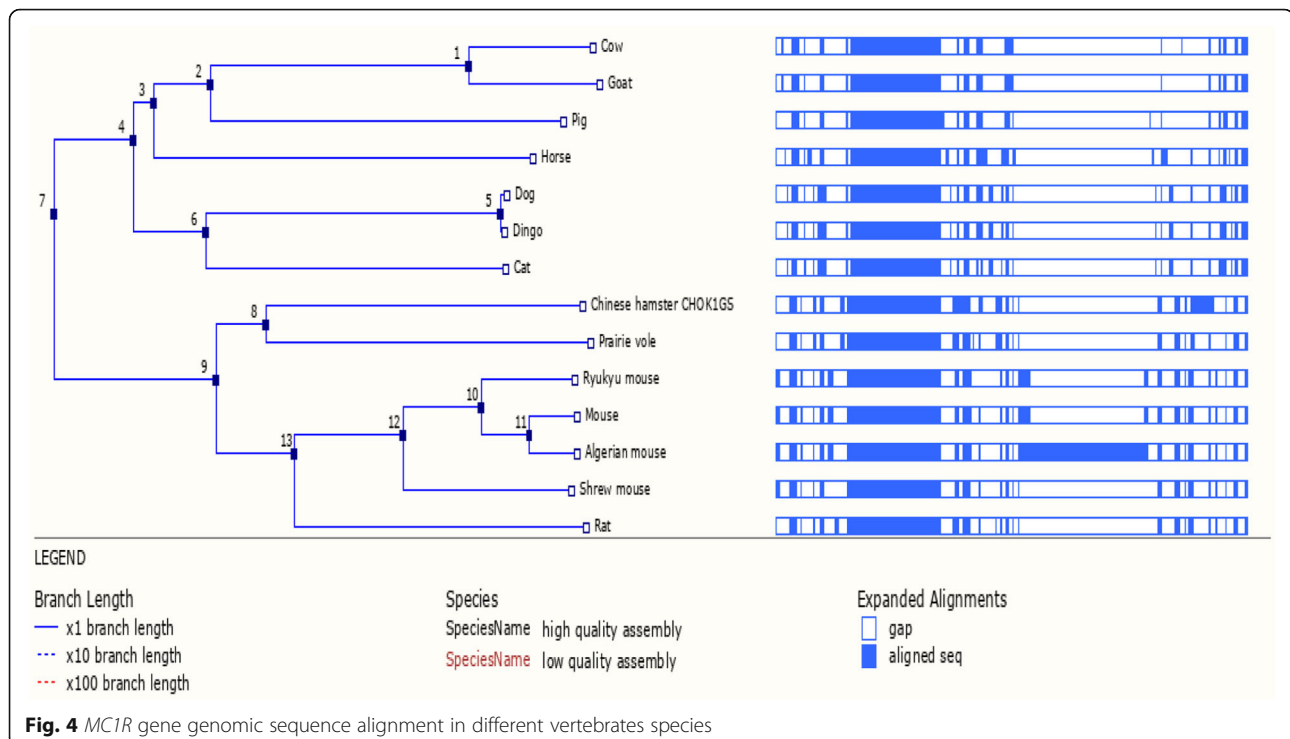
rate of duplicate trees, the associated grouped taxa from the bootstrap, distributed (1000 replicates) is presented near to the branches. Limited than fifty percentage of bootstrap replicates are deflated but the branches following to partitions were reproduced [35]. A total of nine amino acid sequences were carried in the study. In the last dataset, a total of 317 positions were covered. The molecular genetic analyses were carried in MEGA X [36].

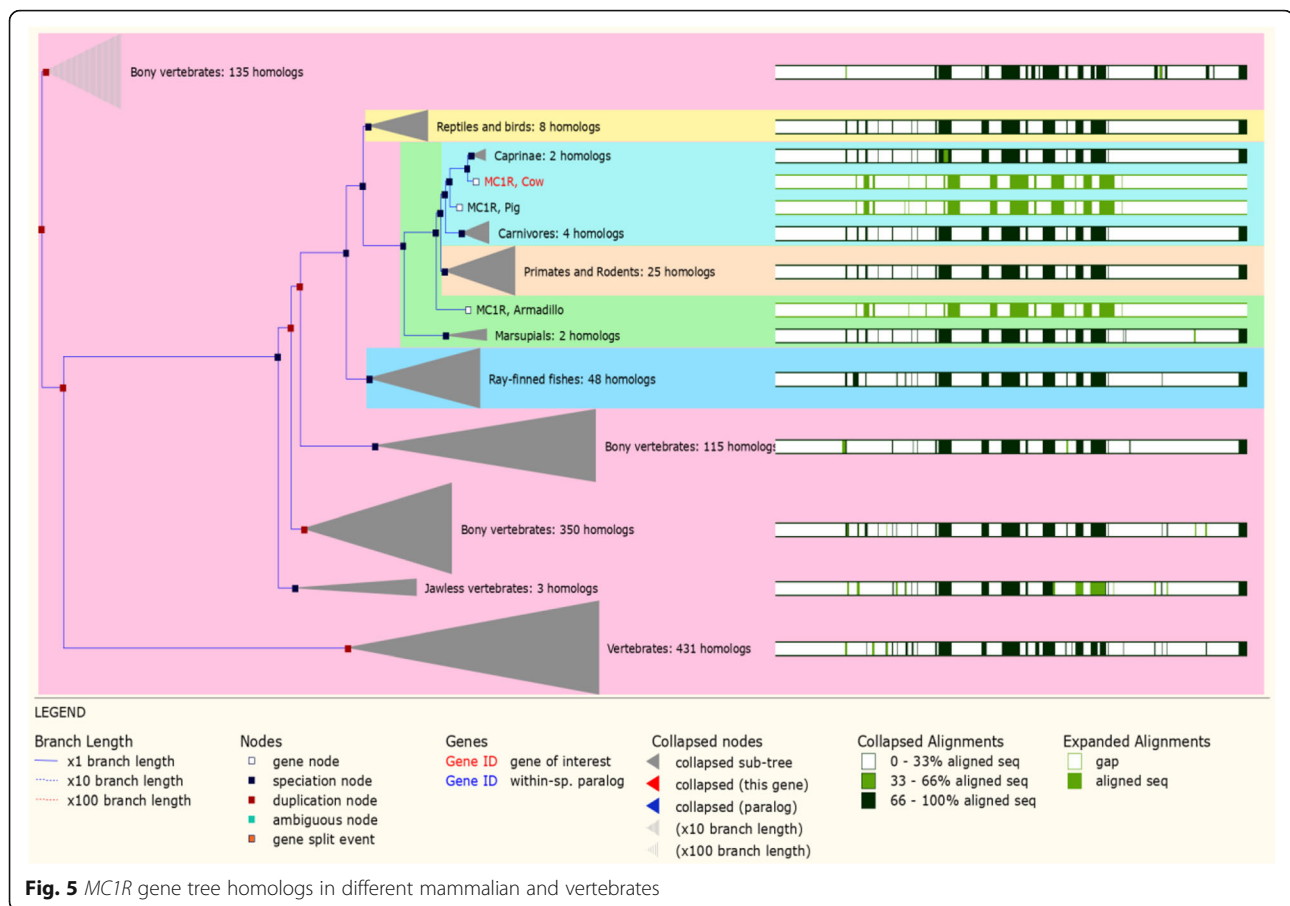
## Results

The *MC1R* gene was amplified and sequenced for the first time in Indian zebu Sahiwal cattle breed. The gene sequence fragment was achieved to be about 1280 base pairs that comprises coding regions and parts of the 5' and 3' noncoding regions (UTR) of the *MC1R* gene (185 and 141 bp sequentially). The complete coding domain of the Sahiwal cattle *MC1R* gene ORF length is 954 base pairs long and it encodes a protein of 317 amino acids. This *MC1R* gene is related with a specific G-protein linked receptor (GPCR). *MC1R* gene associated with the *B. indicus* reference gene sequences showed 100% similarity. The *MC1R* gene sequence was submitted to the NCBI GenBank database with accession number MG373575 to MG373605. Further, the *MC1R* gene sequence was translated by using Expasy protein translational tool. The *MC1R* protein sequence was further characterized as described below.

## *MC1R* protein sequence physicochemical characteristics in the ruminant species

The entire coding region of cattle *MC1R* protein (GenBank accession number NM\_174108) was compared with Sahiwal (*Bos indicus*), gayal (*Bos frontalis*), yak (*Bos grunniens*), Karan Fries (*Bos taurus*) cattle, buffalo (*Bubalus bubalis*), sheep (*Ovis aries*), goat (*Capra hircus*), and horse (*Equus caballus*). The *MC1R* (954 bp) conceal a protein of 317 amino acids with a predicted molecular mass of 34,937.00, 34,935.02, 34,946.00, 34,946.05, 34,817.88, 34,918.95, 34,859.89, and 35,213.36 daltons and iso-electric point of the protein was pI = 8.97, 8.97, 8.97, 9.05, 8.97, 8.97, 8.97, and 8.69 sequentially in the order of species mentioned above. The average number of amino acids in proportions for all the nine species is shown in Table 1. The other physicochemical characteristics of the *MC1R* protein sequence of the other mammals are presented in Table 2. Like protein, motif predictions were also made by employing the PROSITE motif exploration tool and other transmembrane helix predictions were investigated by using transmembrane prediction tools. Overall, we observed the 7-transmembrane helix region including inside and outside regions. A total of 317 amino acids with the 7-transmembrane structural helix region was identified in the *MC1R* protein. The *MC1R* protein sequence with other mammals was characterized by further studies.





**Fig. 5** *MC1R* gene tree homologs in different mammalian and vertebrates

### Comparative investigation of the *MC1R* protein sequence with other mammals

We compared the *MC1R* protein sequences of the Sahiwal, gayal, wild yak, Karan Fries, buffalo, goat, sheep, and horse. The results showed a high homology identity. Amino acid sequence alignment of the *MC1R* protein was done by using multiple sequence alignment Clustal Omega, (MUSCLE 3.8 version), European Bioinformatics Institute. The *MC1R* amino acid sequence showed very high homology with other mammals as compared with *Bos frontalis* 99.68%, *Bos grunniens* 99.37%, *Bos taurus* 99.37%, *Bubalus bubalis* 97.13%, *Capra hircus* 96.85%, *Ovis aries* 96.85%, and *Equus caballus* 83.91% (Fig. 2).

### Comparative molecular phylogenetic analysis of the Sahiwal *MC1R* with other mammals

After determining the amino acid alignment sequence of *MC1R*, we analyzed the *MC1R* protein sequence of the Sahiwal cattle breed. The phylogenetic tree was built based on the *MC1R* sequence of distinct mammalian ruminant species retrieved from GenBank to investigate the evolutionary relationship along with the Sahiwal cattle and *MC1Rs* of other ruminant species. The results indicated that *MC1R* of the Sahiwal cattle (*Bos indicus*)

have shown maximum close relations with yak (*Bos frontalis*), Gayal (*Bos grunniens*), Karan Fries (*Bos taurus*), and *Bubalus bubalis*, and a modest relation was observed with other ruminant species such as *Capra hircus*, *Ovis aries*. The sequence integrity was perceived to be least when related with *Homo sapiens* and *Canis lupus familiaris*, *Equus caballus*. The *MC1R* sequence of Sahiwal revealed 100% relation with that of *Mus musculus*, *Gallus gallus*, and *Sus scrofa* (Fig. 3).

The *MC1R* gene genomic region alignment was related with the other thirteen mammalian species (Fig. 4). Among them, the first internode represents genomic sequence alignment of the cattle and goat that are closely related to the same taxa. The second, third, fourth, fifth, and sixth branch internodes were connected with the pig, horse, dog, and dingo, followed by the seventh branch with cat, respectively. Eight and ninth branches were interconnected with the Chinese hamster and Prairie vole, followed by the ten, eleven, twelve, and thirteen internodes, linked with Ryukyu mouse, mouse, Algerian mouse, Shrew mouse, and rat.

Moreover, the gene *MC1R* tree of cow was compared with vertebrate species by using Ensemble Asia cow genome browser and we retrieved the *MC1R* reference sequence





(ENSBTAT00000032494) (Fig. 5) that was highly homologous with other vertebrates and mammalian species. The major portion of *MC1R* gene sequence alignment was highly homologous in 431, vertebrate species, 350, 115 bony vertebrates' and the gene sequence of the alignment was determined to be 100% similar. Among the primates and rodents, 25 different species revealed a very large homologous sequence. Moreover, 48 species of ray-finned fishes were remarkably homologous displaying 100% identity. The reptiles, birds, carnivores, jawless vertebrates, marsupials, and caprine species showing 100% identical homologous sequence was found to be 8, 4, 3, 2, 2. Cattle, pig, and armadillo sequence revealed 33–66% relevant homologs in their association. Subsequently, cattle and goat displayed the closely linked homologous sequence and connected to the identical branch node. Moreover, the *MC1R* protein region functions were defined below.

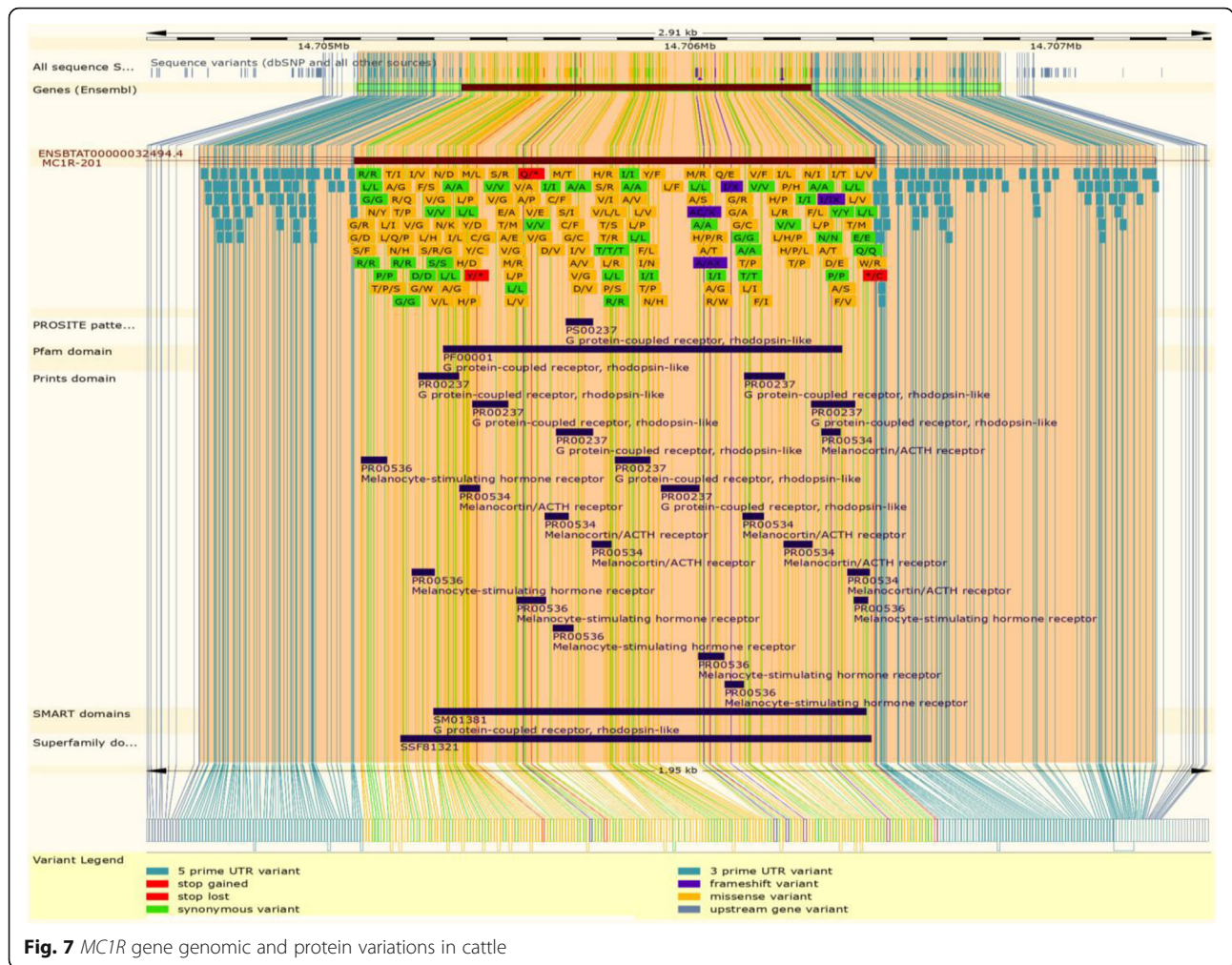
The *MC1R* protein region of the cattle is shown in Fig. 6. *MC1R* has seven transmembrane regions. *MC1R* gene codes the 317 amino acids; the *MC1R* protein domain was linked to a melanocyte-stimulating hormone receptor, adrenocorticotrophic hormone receptor, and G-protein linked receptor, rhodopsin-like receptors. Protein super family domain region of the *MC1R* is linked with the G-protein linked receptor, rhodopsin-like receptor, serpentine-like receptor, seven transmembrane G-protein receptor, and chemoreceptors. PROSITE family was similar to the GPCR, rhodopsin-like

7TM region. In Fig. 6, the *MC1R* coding domain (shown in the red-shaded area) predominantly represents missense variation and those shown in the green-shaded area represent synonymous variants. In the coding region, the amino acids present at 75th and 109th, position showed loss of a function of the stop codon. Similarly, the amino acids present at 104 and 231 positions of the coding region showed four frame shift variations.

The genetic variations in the genomic region of the *MC1R* gene are shown in Fig. 7. The full length of the *MC1R* gene which comprises the 1751 base pairs consists of 5', 3' flanking and un-translated regions; the range of the *MC1R* coding region was 954 base pairs. Most of the variants presented at the coding region and dominantly all were missense variants. Synonymous variants showed lightly in the coding domain. There were less frame shift variations found in the coding region of the *MC1R* gene in cattle.

#### Identification of mutation in the melanocortin 1 receptor 7-transmembrane region

Upon comparison of the Karan Fries with that of the zebu Sahiwal cattle, only one inter-breed SNP was identified (c.844C>A) in exonic region of the *MC1R* gene. One amino acid differences were found in the seven transmembrane helix region of the *MC1R* protein between Sahiwal and Karan Fries (Fig. 8). The amino-terminal



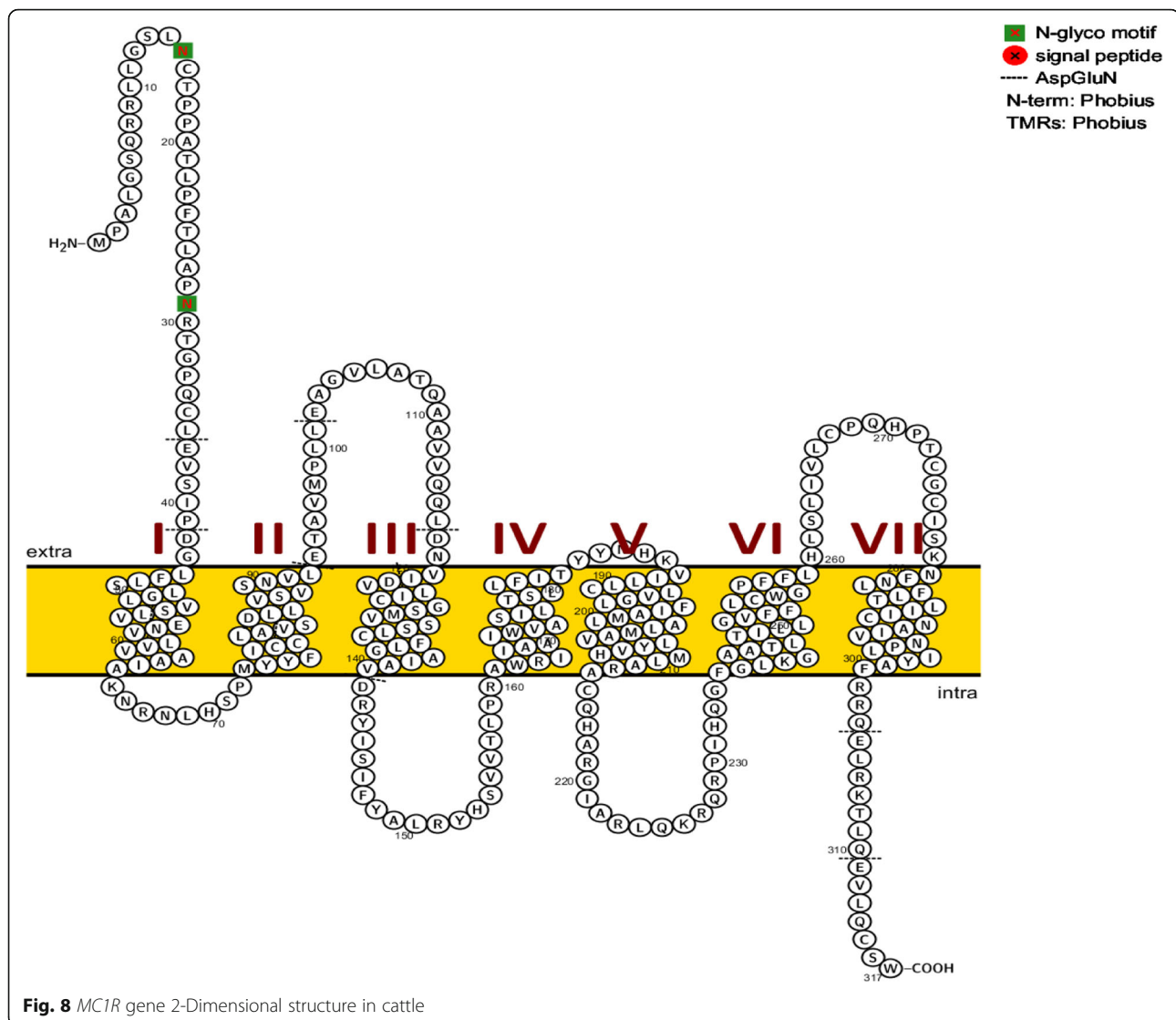
**Fig. 7** *MC1R* gene genomic and protein variations in cattle

sequence of the *MC1R* protein contains a total of forty-three amino acids (1–43) (MPALGSQRRLGSLNCTPPA TLPFTLAPNRTGPQCLEVSIPDG). The (TM1) first trans-membrane structural helix precinct of the 7-TM comprises twenty-one amino acids (44–63) (LFLSLGLV SLVENLVVAAIA) followed by the first intracellular domain region that contains eleven amino acids (64–75) (KNRNLHSPMY). The (TM2) second transmembrane helix region contains twenty-three amino acids (76–98), (FICCLAVSDLLVSVSNVLETAVM) followed by the second extracellular domain region that contains nineteen amino acids (99–117) (PLLEAGVLATQAAVVQQLD), with an amino acid mutation occurring at (p.99 L>P) position. The (TM3) third trans-membrane structural helix region contains twenty-three amino acids (118–140) (NVIDVLICGSMVSSLCFLGAIIV) followed by the second intracellular domain regions containing nineteen amino acids (141–160) (DRYISIFYALRYHSVVTLP). The (TM4) fourth transmembrane structural helix contains twenty-two amino acids (161–183) (RAWRIIAIIVVASI LTSLLFIT). Third extracellular transmembrane domain

region contains two amino acids (184–186) (YY). Fifth (TM5) transmembrane structural helix region contains twenty-two amino acids presents (187–209), (NHKVILLC LVGLFIAMLALMAV). The third intracellular transmembrane domain region contains thirty-four amino (210–244) (LYVHMLARACQHARGIARLQKRQRPIHQGFGL KG) followed by the sixth (TM6) trans-membrane structural helix region that carries twenty-two amino acids (245–267) (AATLTILLGVFFLCWGPFFLHL). The fourth extracellular transmembrane domain region contains thirteen amino acids (268–281) (SLIVLCPQHPTCG), in which the p.281 T>N is the variation between Sahiwal and Karan Fries. Seventh (TM7) transmembrane structural helix region contains the eighteen amino acids (282–300) (CISKNFNLFLLIICNAI) in position. The fourth intracellular transmembrane domain region contains sixteen amino acids (301–317) (RQELRKTLLQEVLCQSW).

### Discussion

Coat color shows various types of shades in cattle and other mammalian species. *MC1R* gene plays a key role



in regulating the eumelanin and pheomelanin biosynthesis pathways in mammals. Prior to the current study, there was no information available regarding coat color genes in Sahiwal (*Bos indicus*). Due to an inherited variation in the *MC1R* gene, which codes for the melanocyte-stimulating hormone ( $\alpha$ -MSH) receptor, has been proved to be associated with the coat color of cows [9, 37, 38]. The *Bos* species *MC1R* has been broadly studied at three important loci such as the ED (dominant, dominant black),  $E^+$  (intermediate, recessive black), and  $e$  (recessive, red). The interchange of a T base (ED) with that of the C base ( $E^+$ ) at the 296th coding region leads to the variations of the amino acid from base T296C: p. Leu 99 Pro. The  $e$  allele derived from the deletion of a G at region 310/311 of the coding sequence. With regard to the skin color phenotype of Norwegian and Icelandic cattle herds, those with the ED allele present a dominant black color. Cattle possessing the  $e$ /

$e$  genotype always exhibited red, while animals homozygous or heterozygous for the wild-type allele  $E^+/E^+$  and  $E^+/e$  could produce either black, brownish, or red [14]. Red coat coloration is connected with the  $e$  locus of Holstein cow [15]. In Korean and Japanese beef cattle ED, *MC1R* gene has been briefly described in the  $e$  loci [39]. Black coat color Japanese cow had no homozygous allele's  $e/e$ , and consequently obtained consistent, expressing their black coat color as ED and  $E^+$  alleles, which resulted in black pigmentation. In addition, the majority of Korean Hanwoo cattle varieties, which had a coat color changing from yellowish-brown to dark brown, including red skin color, were all homozygous for  $e/e$  [40]. Despite, Korean cattle owning a yellowish-brown color were perceived to have gene regularities of the  $E^+/e$  (0.05) and  $e/e$  (0.95) [41]. In (*Bubalus bubalis*) black coat color of 49 buffaloes, white and grey coat color of 136 swamp buffaloes, 31 hybrid off-springs of



river buffalo, and swamp buffaloes, *MC1R* gene was sequenced. Among these, a total of three SNPs were found. First SNP, c.618 G>C was the same sense mutation and among other SNPs, two were missense mutation c.310G>A and c.384C>T due to differences in the amino acid array p.S104G and p.I128M mentioned already. The black coat color was linked with the SNP104 in buffaloes [17]. In Thai white zebu (*Bos indicus*) cattle four SNPs c.296 T>C, c.416C >T, c.663A>C, and c.725A>C were reported earlier by Mekchay S [42]. Two novel alleles were already reported earlier in the Ed1 (c.T667C) (p.W223R) and Ed2 (c.651In12: p.218InARGI) in Brown Swiss cattle. In Simmental cattle, one of the allele ef was found (c.C890T), (p.T297I) by [43]. Recent studies in gayal (*Bos frontalis*) revealed nine SNPs including five SNPs in the coding domain (C201T, C583T, T663C, A871G, and T876C) and four SNPs (G-1A, C-106 T, A-127C, T-129C) in specific 5' noncoding region published by previously [44]. Eight SNPs in Karan Fries crossbreed cattle were identified (c.296 T>C, c.583C>T, c.663 T>C, c.830C>A, c853G>A, c.880G>A, c.906C>G, and c.927C>T) recently in *MC1R* gene while compared with Tharparkar [16] along with we found the novel SNP (c.844C>A and p.281 T>N) in the present study. Hanna et al. reported both SNPs in *Bos indicus* source cattle at c.583C>T and c.663 T>C of *MC1R* [45]. Those were the only extra coding modifications that we recognized in this intron-less gene and they were in the complete phase with the *Bos indicus* E+ allele (c.296C>T). Discovery of this unique haplotype (TTC) in *Bos indicus* breed and crossbreed cattle prescribed that nucleotide divergence in *MC1R* obtained from *Bos indicus* did not provide to variation in the degree of black in EDE+ heterozygotes [45]. The dominant allele ED, owning a particular amino acid substitution from leucine to proline is reported in bovine E locus [14]. Three leucine amino acid residues were changed to leucine-proline-leucine motifs found in mouse ESO allele. In bovine, proline-leucine-leucine read motifs were found in the ED allele region [14]. We also observed the same kind of results in Karan Fries cross-breed cattle having dominant black color due to the amino acid substitution at 99th position. Moreover, we recognized extensive haplotypes (due to linkage) that stretched for more than 1 Mb either side of *MC1R*. Each EDE+ having a couple of distinct alleles from the *MC1R* gene were attributed from the haplotypes of their progenitors to be heterozygous for the CCT/TTC *MC1R* haplotypes. Inside families, heterozygotes that received the related set of long-range haplotypes revealed a broad change in their degree of black coat color pigmentation; consequently, we hypothesize that it is unlikely that a regulatory modification associated with *MC1R* could explain the observed black and white patch effect.

Coat color might be a complex trait, which is linked to many genes, pathways, or networks. *MC1R*, a known key determinant of color phenotype in bovine [46]. The transmembrane domain is composed of several polar and non-polar amino acid changes in the transmembrane domain have been described to regulate gene function due to variation of the protein localization and interaction with different molecules [47, 48]. Due to SNP changes (c.844C>A, p281T>N) might alter the function and the structure of *MC1R* protein. A novel Single SNP in Karan Fries cattle (black and white color) breed compared to indigenous Sahiwal cattle (reddish yellow color) due to amino acid changes threonine to asparagine in A 7-TM inner region of the *MC1R* gene. Because of the protein sequence changes due to the hydroxylation of the amino acid threonine which will involve in the pathway of the pheomelanin synthesis it can cause reddish coat color. Due to the lack of the eumelanin synthesis in the production of black and white color. Several studies reported that in mammals, the pigmentation of the coat color can be affected by 150 to 300 genes reported by [49, 50]. Various coat color phenotypes distinguished by the Reggiana and Holstein cattle are recognized in the *MC1R* gene region, on the eighteenth chromosome. These two breeds showed the intense allele frequency divergence in the *MC1R* region was observed by [51], which recently reported that the *MC1R* gene (c.871G>A) SNP can cause the mutation in brown Japanese cattle [52]. Senczuk et al. reported that different grey and non-grey coat color cattle breeds [53]. The *MC1R* gene variant is associated with black color in Swiss Holstein cattle breed [54]. Recently, Kasprzak et al. described *MC1R* polymorphism in different coat color cattle breeds of Central Europe [55].

## Conclusions

Analysis of Sahiwal *MC1R* gene revealed that the SNP c.844C>A or p.281 T>N might be the possible reason for its coat color variation from Karan Fries cattle. Further molecular studies need to be conducted to prove this association with the synthesis of different melanins.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s43141-021-00117-2>.

**Additional file 1: Table S1.** Web tools and function for bioinformatics analysis. **Figure S1.** Hydrophilicity and hydrophobicity analyses. **Figure S2.** Signal peptide prediction. **Figure S3.** Transmembrane domain prediction. **Figure S4.** Phosphorylation site prediction. **Figure S5.** Secondary structure prediction.

## Abbreviations

*MC1R*: Melanocortin 1 receptor gene; SNPs: Single nucleotide polymorphisms; 7-TM: Seven transmembrane domains; GPCR: G-protein coupled receptor; PCR: Polymerase chain reaction; gDNA: Genomic

deoxyribonucleic acid;  $\alpha$ -MSH: Alpha melanocyte-stimulating hormone; ORF: Open reading frame; CDS: Coding sequence

### Acknowledgements

We declare our thanks to the Director-General, Indian Council of Agricultural Research (ICAR) and Director, National Dairy Research Institute (NDRI), (NICR A) Karnal, required facilities to carry out this research study.

### Authors' contributions

TSG performed the experiments; RCU, VBR, SKO, CK analyzed the data and prepared the manuscript. RCU, TSG procured and prepared the samples and wrote a part of the manuscript. RCU, TSG conceptualized the idea, planned and organized the experiments, improved the manuscript. All authors read and approved the final manuscript.

### Funding

Financial support for this research study provided by the National Initiative on Climate Resilient Agriculture (NICRA Grant No-2049-3033) project the Indian Council of Agricultural Research (ICAR), New Delhi. The funder had no role in study design, interpretation of data and data collection and analysis, decision to publish, or in writing the manuscript.

### Availability of data and materials

GenBank sequence files are available from the NCBI database (Accession ID: MG373575 to MG373605) other supporting data is provided in the [supplementary material](#).

### Competing interest

The authors declare that there are no conflicts of interest.

### Ethics approval and consent to participate

All the test animals were maintained from the Institutional Animals Ethics Committee (IEAC) of ICAR-NDRI, Karnal as per the CPCSEA (13th article) rules followed by the Indian Government. Standards regarding the ethical practice of animals during the complete procedure were strictly followed (note, there is no need for an ethical number for routine blood sample collection).

### Consent for publication

Not applicable.

### Author details

<sup>1</sup>Climate Resilient Live Stock Research Centre, ICAR-National Dairy Research Institute, Karnal, Haryana 132001, India. <sup>2</sup>Department of Biotechnology, Vikrama Simhapuri University, Andhrapradesh, Nellore 524320, India. <sup>3</sup>Department of Life Science & Bioinformatics, Assam University Diphu Campus, Diphu, Assam 782462, India. <sup>4</sup>Molecular Endocrinology, Functional Genomics and Structural Biology, Animal Biochemistry Division, ICAR-National Dairy Research Institute, Karnal, Haryana 132001, India.

Received: 14 November 2019 Accepted: 6 January 2021

Published online: 29 January 2021

### References

- Caro T (2005) The adaptive significance of coloration in mammals. *Bioscience* 55:125–136
- Takahashi A (2013) Pigmentation and behaviour: potential association through pleiotropic genes in *Drosophila*. *Gene* 518(3):165–174
- Miyagi R, Terai Y (2013) The diversity of male nuptial coloration leads to species diversity in Lake Victoria. *Gene* 518(3):145–153
- Darwin C (1859) On the origin of species by means of natural selection or the preservation of favored races in the struggle for life. John Murray, London
- Darwin C (1868) The variation of animals and plants under domestication. John Murray, London
- Adalsteinsson S, Bjarnadottir S, Vage DI, Jonmundsson JV (1995) Brown coat color in Icelandic cattle produced by the loci extension and agouti. *J Hered* 86:395–398
- Wang N, Hebert DN (2006) Tyrosinase maturation through the mammalian secretory pathway: bringing color to life. *Pig Cell Res* 19:3–18
- Robbins LS, Nadeau JH, Johnson KR, Kelly MA, Roselli-Rehffuss L, Baack E, Mountjoy KG, Cone RD (1993) Pigmentation phenotypes of variant extension locus alleles result from point mutations that alter MSH receptor function. *Cell* 72:827–834
- Selz Y, Braasch I, Hoffmann C, Schmidt C, Schultheis C, Scharl M, Volff JN (2007) Evolution of melanocortin receptors in teleost fish: the melanocortin type 1 receptor. *Gene* 401:114–122
- Deng WD, Xi DM, Gou X, Yang SL, Shi XW, Mao HM (2008) Pigmentation in black-boned sheep (*Ovis aries*): association with polymorphism of the tyrosinase gene. *Mol Biol Rep* 35:379–385
- Deng W, Tan Y, Wang X, Xi D, He Y, Yang S, Mao H, Gao S (2009) Molecular cloning, sequence characteristics and polymorphism analyses of tyrosinase-related protein 2 gene with black traits from Blackboned sheep (*Ovis aries*). *Genome* 52:1001–1011
- Zhang JQ, Chen H, Sun ZJ, Liu XL, Qiang-Ba YZ, Gu YL (2010) Flesh color association with polymorphism of the tyrosinase gene in different Chinese chicken breeds. *Mol Biol Rep* 37:165–169
- Mountjoy KG, Robbins LS, Mortrud MT, Cone RD (1992) The cloning of a family of genes that encode the melanocortin receptors. *Science* 257:1248–1251
- Klungland H, Vage DI, 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
- Joerg H, Fries HR, Meijerink E, Stranzinger GF (1996) Red coat color in Holstein cattle is associated with a deletion in the *MSHR* gene. *Mamm Genome* 7:317–318
- Goud TS, Upadhyay RC, Onteru SK, Pichili VBR, Chadipiralla K (2020) Identification and sequence characterization of melanocortin 1 receptor gene (*MC1R*) in *Bos indicus* versus (*Bos taurus X Bos indicus*). *Anim Biotechnol* Aug 31(4):283–294
- Miao YW, Wu GS, Wang L, Li DL, Tang SK, Liang JP, Mao HM, Luo HR, Zhang YP (2010) The role of *MC1R* gene in buffalo coat color. *Sci China Life Sci* 53: 267–272
- Xi D, Wu M, Fan Y, Huo Y, Leng J, Gou X, Mao H, Deng W (2012) Isolation and characteristics of the melanocortin 1 receptor gene (*MC1R*) in the Chinese yak (*Bos grunniens* × *Bos taurus*). *Gene* 498:259–263
- Fontanesi L, Beretti F, Valentina R, Stefania DO, Elena GG, Raffaella F, Roberta Russo DV, Baldassare P (2009) 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
- Wu ZL, Li XL, Liu YQ, Gong YF, Liu ZZ, Wang XJ, Xin TR, Ji Q (2006) The red head and neck of Boer goats may be controlled by the recessive allele of the *MC1R* gene. *Ani Res* 55:313–322
- Kijas JMH, Wales A, Törnsten Chardon P, Moller M, Andersson L (1998) Melanocortin receptor 1 (*MC1R*) mutations and coat color in pigs. *Genetics* 150:1177–1185
- Marklund L, Johansson MM, 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
- Vage DI, Lu D, Klungland H, Lien S, Adalsteinsson S, Cone RD (1997) A non-epistatic interaction of agouti and extension in the fox, *Vulpes vulpes*. *Nat Genet* 15(3):311–315
- Newton JM, Wilkie AL, He L, Jordan SA, Metallinos DL, Holmes NG, Jackson IJ, Barsh GS (2000) Melanocortin 1 receptor variation in the domestic dog. *Mamm Genome* 11:24–30
- Eizirik E, Yuhki N, Johnson WE, Menotti MR, Hannah SS, O'Brien SJ (2003) Molecular genetics and evolution of melanism in the cat family. *Curr Biol* 13:448–453
- Nachman MW, Hoekstra HE, D'Agostino SL (2003) The genetic basis of adaptive melanism in pocket mice. *Proc Natl Acad Sci U S A* 100:5268–5273
- Werthe LA, Hawkins GA, Eggen Petit AEC, Kreigesmann BE, Bishop MD (1996) Rapid communication: melanocyte stimulating hormone receptor (*MC1R*) maps to bovine chromosome 18. *J Anim Sci* 74:262–262
- McRobie HR, King LM, Fanutti C, Coussons PJ, Moncrief ND and Thomas APM (2014) Melanocortin 1 receptor (*MC1R*) gene sequence variation and melanism in the gray (*Sciurus carolinensis*), fox (*Sciurus niger*), and red (*Sciurus vulgaris*) squirrel. *J Hered* 105(3):423–428
- Lu D, Vage DI, Cone RD (1998) A ligand-mimetic model for constitutive activation of the melanocortin-1 receptor. *Mol Endo* 12:592–604
- Goud TS, Upadhyay RC, Kumar A, Karri S, Choudhary R, Ashraf S, Singh SV, Kumar OS, Kiranamai CH (2018) Novel extraction of high-quality genomic DNA from frozen bovine blood samples by using detergent method. *Open Vet J* 8:415–422



31. Petersen TN, Brunak S, Von Hejine G, Nielsen H (2011) SignalP 4.0 discriminating signal peptides from transmembrane regions. *Nat Methods* 8: 785–786
32. Sigrist CJA, De Castro E, Cerutti L, Cuče BA, Hulo N et al (2012) New and continuing developments at PROSITE. *Nucleic Acids Res* 41:1–4
33. Omasits U, Ahrens CH, Müller S, Wollscheid B (2014) Protter: interactive protein feature visualization and integration with experimental proteomic data. *Bioinformatics* 30(6):884–886
34. Michael SK, Martin CJ, Henrik N, Kamilla KJ, Vanessa IJ, Casper KS, Morten OAS, Ole W, Morten N, Bent P, Paolo M (2019) NetSurfP-2.0: improved prediction of protein structural features by integrated deep learning. *Protocols: structure, function and bioinformatics*. <https://doi.org/10.1002/prot.25674>
35. Felsenstein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution* 39:783–791
36. Kumar S, Stecher G, Li M, Niyaz C, Tamura K (2018) MEGA X: molecular evolutionary genetics analysis across computing platforms. *Mol Biol and Evol* 35:1547–1549
37. Simon JD, Peles DN (2010) The red and the black. *Acc Chem Res* 43:1452–1460
38. Slominski A, Tobin DJ, Shibahara SJ (2004) Melanin pigmentation in mammalian skin and its hormonal regulation. *Phy Rev* 84:1155–1228
39. Sasazaki S, Usui M, Mannen H, Hiura TSuji S (2005) Allele frequencies of the extension locus encoding the melanocortin-1 receptor in Japanese and Korean cattle. *J Anim Sci* 76:129–132
40. Kim SH, Yoon JT (2013) Changes in melanogenesis-related genes by the variation of *MC1R* in dark-muzzled Korean native cattle and Korean brindle cattle. *J Anim Vet Adv* 12(12):1101–1108
41. Hyeah G, Doori L, Okwan, Keunchang SK, Dong C, and JY, Seong HK (2009) Fast identification of Korean cattle based on simultaneous detection of various single nucleotide polymorphisms markers by capillary electrophoresis. *Bull Korean Chem Soc* 30:2141–2144.
42. Mekchay S (2008) Identification of single nucleotide polymorphism on *MC1R* gene in Thai native cattle. In: *Competition for Resources in a Changing World: New Drive for Rural Development*. Tropentag, October 7-9, Hohenheim
43. Graphodatskaya D, Joerg H, Stranzinger G (2002) Molecular and pharmacological characterization of the *MSH-R* alleles in Swiss cattle breeds. *J Recept Signal Transduct Res* 22:421–430
44. Xi D, Wu M, Fan Y, Huo Y, Leng J, Gou X, Mao H, Deng W (2012) Nucleotide diversity of the melanocortin 1 receptor gene (*MC1R*) in the gayal (*Bos frontalis*). *Mol Biol Rep* 39:7293–7301
45. Hulsman Hanna LL, Sanders JO, Riley DG, Abbey CA, Gill CA (2014) Identification of a major locus interacting with *MC1R* and modifying black coat color in an F2 Nellore-Angus population. *Genet Sele Evol* 46:4
46. Zhang Y, Li Q, Ye S, Faruque MO, Yu Y, Sun D, Zhang S, Wang Y (2014) New variants in the melanocortin 1 receptor gene (*MC1R*) in Asian cattle. *Anim Genet* Aug 45(4):609–610
47. Cooper CJ, Dutta NT, Martin CE, Piscione TD, Thomer PS, Jones N (2018) Characterization of a novel disease-associated mutation within *NPHS1* and its effects on nephrin phosphorylation and signaling. *PLoS One*. Sep 13;13(9)
48. Bocharov EV, Nadezhdin KD, Urban AS, Volynsky PE, Pavlov KV, Efremov RG, Arseniev AS, Bocharova OV (2019) Familial L723P mutation can shift the distribution between the alternative APP transmembrane domain cleavage cascades by local unfolding of the E-cleavage site suggesting a straightforward mechanism of Alzheimer's disease pathogenesis. *ACS Chem Biol* Jul 19;14(7):1573–1582
49. Hofreiter M, Schöneberg T (2010) The genetic and evolutionary basis of color variation in vertebrates. *Cell Mol Life Sci* 67:2591–2603
50. Cieslak M, Reissmann M, Hofreiter M, Ludwig A (2011) Colors of domestication. *Biol Rev Camb Philos Soc* 86:885–899
51. Bertolini F, Schiavo G, Bovo S, Sardina MT, Mastrangelo S, Dall'Olio S, Portolano B, Fontanesi L (2020) Comparative selection signature analyses identify genomic footprints in Reggiana cattle, the traditional breed of the Parmigiano-Reggiano cheese production system. *Animal* 14(5):921–932
52. Matsumoto H, Kojima M, Takamuku H, Kimura S, Kashimura A, Imai S, Yamauchi K, Ito S (2020) *MC1R* c.310G>- and c.871G > A determine the coat color of Kumamoto sub-breed of Japanese Brown cattle. *Anim Sci J Jan-Dec;91(1): e13367* doi: <https://doi.org/10.1111/asj.13367>
53. Senczuk G, Guerra L, Mastrangelo S, Campobasso C, Zoubeyda K, Imane M, Marletta D, Kusza S, Karsli T, Gaouar SBS, Pilla F, Ciani E (2020) The Bovita consortium. Fifteen shades of Grey: combined analysis of genome-wide SNP data in steppe and Mediterranean Grey cattle sheds new light on the molecular basis of coat color. *Genes (Basel)* 11(8):932
54. Hauser M, Wolf-Hofstetter S, Acklin-Menzi F, Studer E, Rediger D, Seefried F, Drögemüller C (2020) Grey, curly and short-haired Swiss Holstein cattle show genetic traces of the Simmental breed. *Schweiz Arch Tierheilkd*. 162(9):551-559
55. Kasprzak-Filipek K, Sawicka-Zugaj W, Litwińczuk Z, Chabuz W, Šveistienė R, Bulla J (2020) Polymorphism of the Melanocortin 1 receptor (*MC1R*) gene and its role in determining the coat colour of central European cattle breeds. *Animals (Basel)* 10(10):E1878

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Submit your manuscript to a SpringerOpen® journal and benefit from:**

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

---

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)

---