Mammalian Genome 7, 895–899 (1996).

Mammalian Genome

A missense mutation in the gene for melanocyte-stimulating hormone receptor (*MC1R*) is associated with the chestnut coat color in horses

L. Marklund, M. Johansson Moller, K. Sandberg, L. Andersson

Department of Animal Breeding and Genetics, Swedish University of Agricultural Science, BMC, Box 597, S-751 24 Uppsala, Sweden

Received: 20 May 1996 / Accepted: 31 July 1996

Abstract. The melanocyte-stimulating hormone receptor gene (*MC1R*) is the major candidate gene for the chestnut coat color in horses since it is assumed to be controlled by an allele at the extension locus. *MC1R* sequences were PCR amplified from chestnut (*e/e*) and non-chestnut (*E/–*) horses. A single-strand conformation polymorphism was found that showed a complete association to the chestnut coat color among 144 horses representing 12 breeds. Sequence analysis revealed a single missense mutation (83Ser → Phe) in the *MC1R* allele associated with the chestnut color. The substitution occurs in the second transmembrane region, which apparently plays a key role in the molecule since substitutions associated with coat color variants in mice and cattle as well as red hair and fair skin in humans are found in this part of the molecule. We propose that the now reported mutation is likely to be the causative mutation for the chestnut coat color. The polymorphism can be detected with a simple PCR-RFLP test, since the mutation creates a *Taq*I restriction site in the chestnut allele.

Introduction

The inheritance of coat color has fascinated geneticists as well as animal breeders for many years. Chestnut is one of the major coat color variants in the horse and is characterized by an entire red coat lacking black or brown pigmentation. The chestnut coat color is in fact one of the first traits in a mammalian species that was found to show a simple Mendelian inheritance; Hurst reported as early as 1906 that chestnut was controlled by a recessive allele. Phenotypic similarities to recessive yellow (*e*) in the mouse and coat color variants in other species led Adalsteinsson (1974) to propose that chestnut is the bottom recessive allele (*e*) at the extension (*E*) coat color locus. This was later supported by comparative mapping data showing that the chestnut gene is located in a chromosomal region with homology to mouse Chromosome (Chr) 8 harboring the extension locus (Andersson and Sandberg 1982).

The extension locus controls, together with the agouti (*A*) locus, the relative amount of red/yellow pigment (phaeomelanin) and black pigment (eumelanin) in mammals (Jackson 1994). It is now well established that the extension locus encodes the melanocyte-stimulating hormone receptor (MSHR; Robbins et al. 1993), while agouti encodes an antagonist to this receptor (Lu et al. 1994). MSHR is member of a family of G protein-coupled receptors, each consisting of seven transmembrane domains (Mountjoy et al. 1992; Chhajlani and Wikberg 1992). (The designation of the gene encoding MSHR is *MC1R* for melanocortin receptor 1 since it belongs to a family of closely related genes encoding melanocortin receptors; the corresponding gene symbol in the mouse is *Mc1r*.)

In an agouti animal, the binding of αMSH to its receptor on the melanocyte induces a switch from the synthesis of phaeomelanin to synthesis of eumelanin (Jackson 1993). Animals homozygous for bottom recessive alleles at the extension locus do not express a functional MSH receptor and consequently show a red/yellow pigmentation. In contrast, animals carrying a top dominant allele at the extension locus express an overactive receptor which will cause a black pigmentation. Mutations in *MC1R* (formerly denoted *MSHR*) have been associated with coat color mutants assigned to the *E* locus in mice (Robbins et al. 1993) and in cattle (Klungland et al. 1995). Moreover, mutations in this gene have been found to be overrepresented in humans with red hair and fair skin (Valverde et al. 1995).

MC1R is thus the major candidate gene for chestnut, since this coat color is assumed to be controlled by an allele at the extension locus. We have previously reported an association between a restriction fragment length polymorphism (RFLP) in the *MC1R* gene and the chestnut color (Johansson et al. 1994; abstract presented at the XXIVth International Congress on Animal Genetics). Here we report the sequence of *MC1R* in chestnut and non-chestnut horses.

Materials and methods

Animals. A total of 114 chestnut (*e/e*) and 30 non-chestnut (*E/E* or *E/e*) horses previously subjected to paternity testing at the Blood Typing Laboratory in Uppsala were included in this study (Table 1); the material represented 12 different horse breeds. DNA was prepared from blood samples by standard protocols.

PCR amplification of the horse MC1R *gene.* A major portion of the single MC1R exon was amplified using two primer pairs. The first one EXT1 (5′-**GAT*GGATCC***TTCTGGGCTCCCTCAACTC, forward) and EXT4 (5′-GTAGTAGGCGATGAAGAGCGTGCT, reverse) amplified 519 bp of the 5′-part of *MC1R;* nucleotides in bold have been added to create a *Bam*HI restriction site for subcloning. EXT5 (5′-TGCGGTACCACAGCATCATGATGCTGC, forward) together with EXT2 (**CTT*GAATTC***GCATGTCAGCACCTCCTTGAG, reverse; an added *Eco*RI site is shown in bold/italics) amplified 496 bp of the 3′ part of *MC1R,* which partially overlapped with the product of EXT1+4 (Fig. 1). In total, 877 bp non-overlapping sequence (excluding the outer primer sites) was amplified out of the total 951 bp, which is the expected size of the horse MC1R coding sequence. The EXT1, 2, and 4 primers were designed on the basis of human and mouse sequences, derived from sequence databases, while the EXT5 primer was based on a partial horse sequence generated in this study.

All PCR reactions were carried out in 10 mM Tris-HCl, pH 8.3, 50 mM KCl, 1.5 mM $MgCl_2$, 0.2 mM of each dNTP, 5 pmol of each primer, and 0.25 U Taq polymerase (Advanced Biotechnology, London, UK) in a reaction volume of 10 μl. The PCR was performed with a PE9600 thermal cycler (Perkin Elmer, Foster City, Calif., USA) with 3 cycles at 95°C for 1 min, 55°C for 30 s, and 72°C for 1 min, followed by 27 cycles at 94°C for 30 s, 50°C for 30 s, and 72°C for 1 min, and a final extension at 72°C for 7 min.

The nucleotide sequence data reported in this paper have been submitted to GenBank and have been assigned the accession number X98012.

Correspondence to: L. Marklund

Table 1. Number of horses included in the study according to breed and coat color.

Breed	Non-chestnut			Chestnut
	E/E[a]	*E/e*[b]	*E/−*[c]	*e/e*
Standardbred		1	8	12
Thoroughbred		1		10
Swedish halfbred	4	3	3	16
North Swedish horse			1	3
Belgian			1	11
Arabian			2	12
American painthorse			1	3
Icelandic horse			2	13
Welsh pony				1
Shetland pony			1	13
New Forest pony			1	7
Gotland pony			1	13
Total:	4	5	21	114

[a] Deduced by having at least 10 non-chestnut offspring in matings to chestnut (*e/e*) mares.
[b] Deduced on the basis of having a chestnut (*e/e*) parent or offspring.
[c] *E/e* or *E/E*.

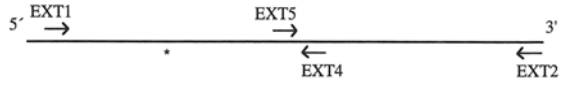

Fig. 1. The relative locations of the PCR primers used to amplify horse MC1R sequences. The location of the observed missense mutation is marked by an asterisk. PCR products obtained with EXT1+4 and EXT2+5 were 519 and 496 bp, respectively.

SSCP analysis. One microliter PCR product was mixed with 3 μl loading buffer (90% formamide, 20 mM NaOH, 0.05% xylene cyanole, 0.05% bromphenol blue) and heated to 95°C for 4 min. The denatured products were then separated in an 8% nondenaturing polyacrylamide gel (37.5:1; acrylamide:bisacrylamide) with a Hoefer SE600 unit (180 × 160 × 0.75 mm) with running conditions of +20°C, 4 W in 0.6 × TBE buffer (53 mM Tris, 53 mM boric acid, 1.5 mM EDTA-disodium) for 17 h. The SSCP patterns were visualized with silver staining as previously described (Marklund et al. 1995).

Cloning and sequencing. The EXT1+4 and EXT2+5 fragments were amplified from each of two chestnut (*e/e*) and two non-chestnut (*E/E*) homozygotes under the same conditions as described above with the exceptions that phosphorylated primers and the *Pwo* polymerase (Boehringer-Mannheim) were used. The PCR products were gel purified, with 1.5% agarose gel electrophoresis and the Qiaex kit (QIAGEN, Chatsworth, Calif.), before ligation into a *Sma*I-restricted pUC18 vector (Pharmacia Biotech, Uppsala, Sweden) according to the manufacturer's recommendations. Cycle sequencing was performed on at least three individual clones from each allele with dyeprimer chemistry (M13 forward and reverse) and the *Taq*FS enzyme (Perkin Elmer). The products were analyzed on an ABI377 sequencer (Applied Biosystem, Foster City, Calif., USA). PCR artifacts were avoided by using a proofreading enzyme for the PCR and by prescreening clones with SSCP analysis which is expected to detect most in vitro-generated artifacts.

Sequence alignments were done with the MEGA program (Kumar et al. 1993). Mouse (X65635), cattle (U39469), and human (X65634) *MC1R* sequences were from the GenBank database (accession numbers in parentheses).

PCR-RFLP analysis. Five microliter of the 519-bp EXT1-4 PCR product, amplified as described above, were transferred to microtiter plates, and restriction enzyme digestion was carried out in a total reaction volume of 25 μl with 2 U *Taq*I, 10 mM Tris-HCl (pH 8.5), 100 mM NaCl, 10 mM $MgCl_2$, 1 mM dithiothreitol, 0.2% (vol/vol) Triton X-100. The reactions were covered with mineral oil and incubated at 65°C overnight. The fragments were separated by 4% agarose gel electrophoresis and subsequently visualized by EtBr staining.

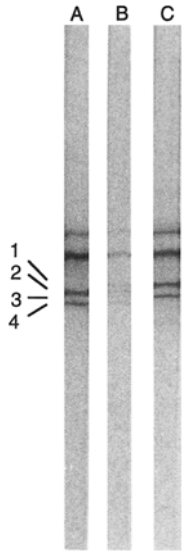

Fig. 2. Single-strand conformation polymorphism (SSCP) analysis of a non-chestnut *E/E* horse (A), a non-chestnut *E/e* horse (B), and a chestnut *e/e* horse (C). Four polymorphic bands (designated 1–4) were observed. Allele *1* showed bands 1 + 3, whereas allele *2* showed 2 + 4. The *MC1R* genotypes are as follows: *2/2* (A), *1/2* (B), and *1/1* (C).

```
          77  78  79  80  81  82  83  84  85  86  87  88
Horse-E   ATC TGC TGC CTG GCC GTG TCC GAC CTG CTG GTG AGC
   -e     --- --- --- --- --- --- -T- --- --- --- --- ---
                                   TaqI

Horse-E   Ile Cys Cys Leu Ala Val Ser Asp Leu Leu Val Ser
   -e     --- --- --- --- --- --- Phe --- --- --- --- ---
```

Fig. 3. Alignment of horse *MC1R* nucleotide sequences and corresponding amino acid sequences found in chestnut (*e*) and non-chestnut (*E*) horses. The codons are numbered according to the one in the corresponding human gene (Robbins et al. 1993). The presence of a *Taq*I restriction site in the *e*-associated sequence is indicated.

Table 2. Nucleotide and amino acid sequence identities in pairwise comparisons of horse, human, cattle, and mouse sequences for the melanocyte-stimulating hormone receptor (MC1R)[a,b].

Species	Horse	Human	Cattle	Mouse
Horse		0.8310	0.8245	0.714
Human	0.8007		0.8526	0.7519
Cattle	0.8272	0.8148		0.7482
Mouse	0.7796	0.7556	0.7562	

[a] Nucleotide and amino acid identities are given above and below the diagonal respectively. The comparison is based on 292 codons.
[b] GenBank accession numbers for human, cattle, and mouse sequences were X65634, U39469, and X65635 respectively.

Results

A two-allele SSCP was detected with the primer pair EXT1-4 (Fig. 2) while no polymorphism was detected for the EXT2+5 PCR product (data not shown). The observed polymorphism gave a classical SSCP pattern where each allele was defined by two bands each, most likely representing the two DNA strands. The MC1R SSCP showed a complete concordance with the chestnut coat color since all 114 chestnut (*e/e*) horses were homozygous *1/1*, whereas all non-chestnut horses (*E/e* or *E/E*) were *1/2* or *2/2*. Moreover, all non-chestnut horses that were deduced to be heterozygous (*E/e*) or homozygous (*E/E*) on the basis of pedigree information were typed as *1/2* or *2/2*, respectively. The association holds up among horses representing 12 breeds (Table 1).

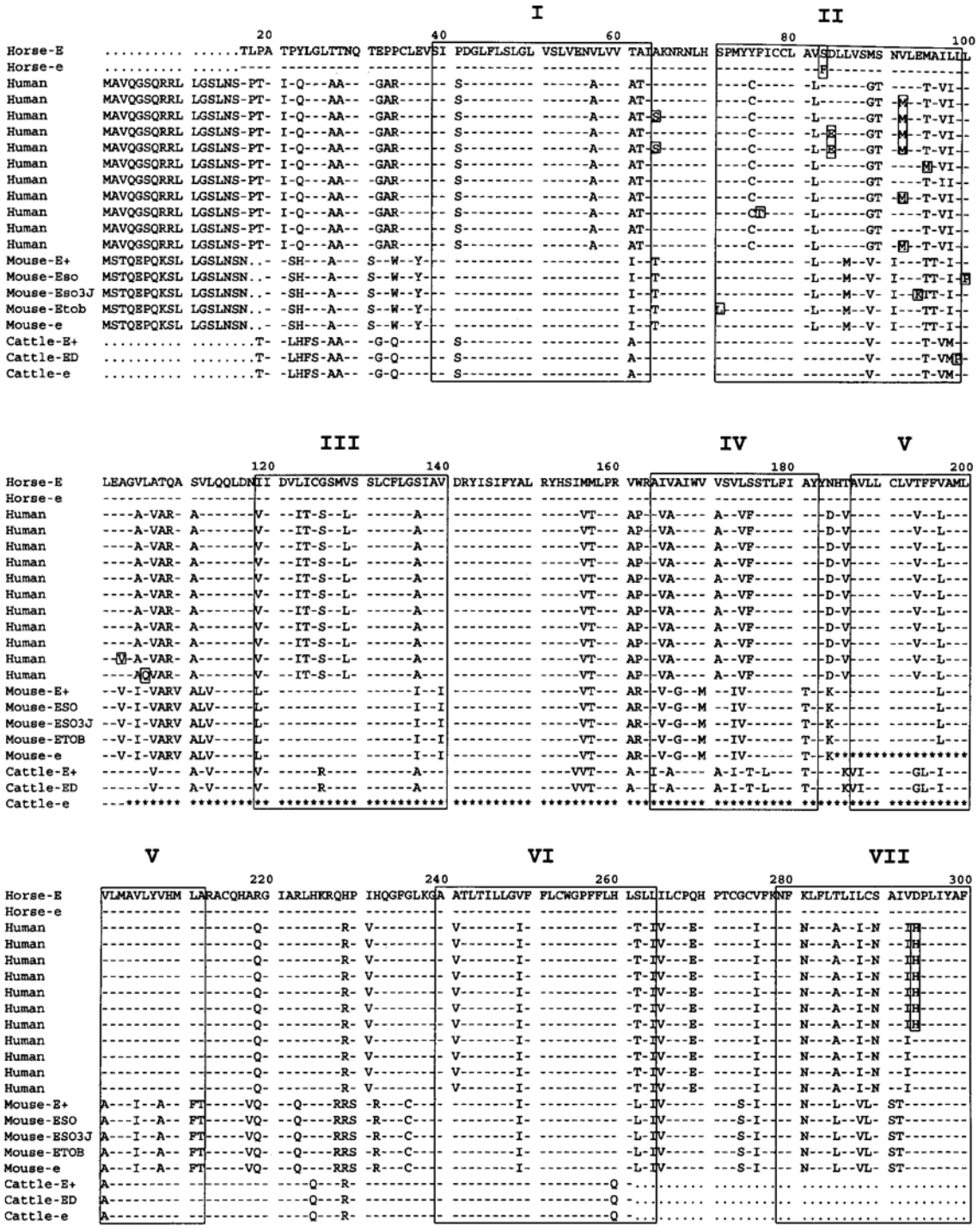

Fig. 4. Amino acid alignment of allelic MC1R sequences in horses (this study), humans (Valverde et al. 1995), mouse (Robbins et al. 1993), and cattle (Klungland et al. 1995; Joerg et al. 1996; accession number: U39469). Intraspecies substitutions are marked by small boxes, and the seven transmembrane domains are boxed and marked I-VII. * = sequences following the 1 = bp deletions found in mouse and cattle *e* alleles; - = identity with the horse *E* allele.; · = no sequence information available.

A major portion of the *MC1R* coding sequence associated with the chestnut (*e*) and non-chestnut (*E*) allele was determined, and the two sequences were aligned. The only sequence difference found was a missense mutation at codon 83 (TCC → TTC), causing a non-conservative amino acid substitution (Ser → Phe) in the MC1R product associated with the *e* allele (Fig. 3). The point mutation creates a *Taq*I restriction site making it possible to identify this mutation by a simple PCR-RFLP analysis. Such an analysis verified that all the chestnut alleles included in this study carried the particular mutation.

The horse *MC1R* sequence shows a high sequence identity to the corresponding sequence in humans, cattle, and mouse both at the nucleotide and amino acid level (Table 2; Fig. 4). The horse, human, and cattle sequences were about equidistant, whereas the mouse sequence was slightly divergent, consistent with the presumed phylogenetic relationship of these four species (for example, Janke et al. 1994).

Discussion

We have here reported a complete concordance between the presence of a missense mutation in the *MC1R* gene and the chestnut coat color in horses based on the analysis of 144 horses representing 12 different breeds. This observation, together with the phenotypic similarity to extension coat color variants in other species (mouse and cattle) known to be controlled by *MC1R* lead us to conclude that the chestnut coat color is caused by a mutation in the *MC1R* gene. Moreover, we propose that the now reported missense mutation is the causative mutation. The observed mutation is non-conservative, replacing a polar residue (serine) with a hydrophobic one (phenylalanine). This is expected to alter the α-helix structure of the second transmembrane domain (Fig. 4), which may lead to a defect receptor. This is apparently a key part of the receptor, since a majority of the point mutations associated with extension coat color variants in mouse and cattle are located in this region. Point mutations in the same region have also been found to be overrepresented in humans with red hair and fair skin (Valverde et al. 1995). Serine at position 83 is conserved among all mammalian MC1R sequences reported so far except in the horse chestnut allele. However, we cannot formally exclude the possibility that the observed missense mutation is only very closely linked to another mutation (for example, in the promoter region), which abolishes *MC1R* expression. But until a non-chestnut horse homozygous for the now reported mutation has been found we will assume that it is the causative mutation. Functional studies of possible differences in ligand binding as well as signal transductions between the two allelic horse MSHR forms may resolve this question.

The finding that the same nucleotide substitution was associated with the chestnut allele among horses from 12 different breeds implies that the chestnut coat color is controlled by the same mutation that has been spread by migration between breeds or was present when the breeds were founded. The presence of a single mutation, controlling a phenotypic trait, appears to be a common phenomenon in farm animals. For instance, a single missense mutation in the calcium release channel gene is associated with malignant hyperthermia in different pig breeds (Fujii et al. 1991), as is the case for a duplication of the *KIT* gene, or a part of it, associated with dominant white color in pigs (Johansson Moller et al. 1996). This is in contrast to the situation in humans, where a plethora of mutations in the same gene is often found (for example, *CFTR*, the cystic fibrosis transmembrane conductance regulator gene; Cuppens et al. 1993).

Kriegesmann and associates (1996) recently reported a *Taq*I PCR-RFLP in the horse *MC1R* gene that most likely is identical to the polymorphism reported here. In contrast to our conclusion, they reported that this polymorphism is not informative for coat color in Arabian horses, a breed represented by 12 chestnut horses in our study. In the brief note, data were presented for 9 horses, and the *E* genotype could only be deduced for 3 horses, a chestnut (*e/e*) offspring and its two non-chestnut (*E/e*) parents. In perfect agreement with out data, the chestnut and non-chestnut horses were homozygous and heterozygous, respectively, for an *MC1R* allele with a *Taq*I restriction site (cf. Fig. 3). Their interpretation that this polymorphism is not associated with color is most likely based on the observation of two gray horses, which both have the same PCR-RFLP type as the chestnut horse. Gray is controlled by a dominant allele at a separate locus (*G*). A gray horse is born with its basic color (for example, black, bay, or chestnut) and gradually turns gray by age. The basic color of an adult gray horse is usually not visible. Hence, it cannot be excluded from the data present in the note that these two animals were gray on a chestnut background (*e/e, G/−*). Thus, we conclude that the data reported by Kriegesmann and colleagues (1996) do not refute our interpretation of the role of the *MC1R* gene in controlling the chestnut coat color.

The present report is an advance in horse coat color genetics and provides the first molecular characterization of a coat color locus in this species. The finding leads to the development of a simple diagnostic DNA test that can be used by horse breeders to identify carriers of the recessive chestnut allele. The chestnut coat color is favored in some breeds but not in others. Moreover, it provides a tool to easily resolve uncertainties whether different shades of chestnut, such as the very dark, liver chestnut, are controlled by the same locus as the more common types of chestnut (Adalsteinsson 1974).

Acknowledgments. We thank the staff at the blood group laboratory for providing all horse samples.

Note added in proof: The basic color (chestnut or non-chestnut) of the two gray horses included in the study by Kriegesmann and coworkers (1996) was not known (B. Brenig, personal communication). Thus, their data are not in conflict with those presented here.

References

Adalsteinsson, S. (1974). Inheritance of the palomino color in Icelandic horses. J. Hered. 65, 15–20.

Andersson, L., Sandberg, K. (1982). A linkage group composed of three coat color genes and three serum protein loci in horses. J. Hered. 73, 91–94.

Chhajlani, V., Wikberg, J.E.S. (1992). Molecular cloning and expression of the human melanocyte stimulating hormone receptor cDNA. FEBS Lett. 309, 417–420.

Cuppens, H., Marynen, P., De Boeck, C., Cassiman, J.-J. (1993). Detection of 98.5% of the mutations in 200 Belgian cystic fibrosis alleles by reverse dot-blot and sequencing of the compete coding region and exon/intron junctions of the CFTR gene. Genomics 18, 693–697.

Fujii, J., Otsu, K., Zorzato, F., De Leon, S., Khanna, V.K., Weiler, J.E., O'Brien, P.J., MacLennan, D.H. (1991). Identification of a mutation in porcine ryanodine receptor associated with malignant hyperthermia. Science 253, 448–451.

Hurst, C.C. (1906). On the inheritance of coat colour in horses. Proc. Roy. Soc. 77, 388–394.

Jackson, I.J. (1993). Colour-coded switches. Nature 362, 587–588.

Jackson, I.J. (1994). Molecular and developmental genetics of mouse coat color. Annu. Rev. Gent. 28, 189–217.

Janke, A., Feldmaier-Fuchs, G., Kelley Thomas, W., von Haesler, A., Pääbo, S. (1994). The marsupial mitochondrial genome and the evolution of placental mammals. Genetics 137, 243–256.

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.

Johansson, M., Marklund, L., Sandberg, K., Andersson, L. (1994). Cosegregation between the chestnut coat colour in horses and polymorphisms at the melanocyte stimulating hormone (MSH) receptor locus. Anim. Genet. 25 (suppl):35.

Johansson Moller, M., Chaudhary, R., Hellmén, E., Höyheim, B., Chowdhary, B., Andersson, L. (1996). Pigs with the dominant white coat color phenotype carry a duplication of the *KIT* gene encoding the mast/stem cell growth factor receptor. Mamm. Genome, in press.

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.

Kriegesmann, B., Jansen, S., Bishop, M., Brenig, B. (1996). The equine MSH-R *Taq*I RFLP is not informative for hair colour in Arabian horses. Anim. Genet. 27, 64.

Kumar, S., Tamura, K., Nei, M. (1993). MEGA: Molecular Evolutionary Genetics Analysis, version 1.0. (University Park, Pa.: The Pennsylvania State University.)

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.

Marklund, S., Chaudhary, R., Marklund, L., Sandberg, K., Andersson, L. (1995). Extensive mtDNA diversity in horses revealed by PCR-SSCP analysis. Anim. Genet. 26, 193–196.

Mountjoy, K.G., Robbins, L.S., Mortrud, M.T., Cone, R.D. (1992). The cloning of a family of genes that encode the melanocortin receptors. Science 257, 1248–1251.

Robbins, L.S., Nadeau, J.H., Johnson, K.R., Kelly, M.A., Roselli-Rehfuss, 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.

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. Nature Genet. 11, 328–330.