A Frameshift Mutation in *MC1R* and a High Frequency of Somatic Reversions Cause Black Spotting in Pigs

J. M. H. Kijas,*,1,2 M. Moller,*,1 G. Plastow† and L. Andersson*

*Department of Animal Breeding and Genetics, Swedish University of Agricultural Sciences, S-751 24 Uppsala, Sweden and †PIC International Group, Fyfield Wick, Abingdon, Oxon, OX13 5NA, United Kingdom

ABSTRACT

Black spotting on a red or white background in pigs is determined by the *E* allele at the *MC1R/Extension* locus. A previous comparison of partial *MC1R* sequences revealed that *E* shares a missense mutation (D121N) with the *E*2 allele for dominant black color. Sequence analysis of the entire coding region now reveals a second mutation in the form of a 2-bp insertion at codon 23 (nt67insCC). This mutation expands a tract of six C nucleotides to eight and introduces a premature stop codon at position 56. This frameshift mutation is expected to cause a recessive red color, which was in fact observed in some breeds with the *E* allele present (Tamworth and Hereford). RT-PCR analyses were conducted using skin samples taken from both spotted and background areas of spotted pigs. The background red area had transcript only from the mutant nt67insCC *MC1R* allele, whereas the black spot also contained a transcript without the 2-bp insertion. This indicates that black spots are due to somatic reversion events that restore the frame and *MC1R* function. The phenotypic expression of the *E* allele is highly variable and the associated coat color ranges from red, red with black spots, white with black spots, to almost completely solid black. In several breeds of pigs the phenotypic manifestation of this allele has been modified by selection for or against black spots.

EXTENSION (E) is one of the classical mammalian coat color loci and allelic series have been proposed in a number of species (Searle 1968). The molecular basis for this locus is now well understood and *E* encodes the melanocortin receptor 1 (*MC1R*) expressed in melanocytes (Robbins et al. 1993). *MC1R* is a G-protein-coupled receptor and *MC1R* signaling determines whether the melanocyte produces black eumelanin or red/yellow pheomelanin. The binding of the ligand *et al.* whether the melanocyte produces black eumelanin or *tein-coupled receptor and MC1R signaling determines *in melanocytes (*Robbins et al.* 1993). The missense mutation D121N. The recessive red coat color of swine was studied in Duroc (*e/e*) and found to share a missense mutation L99P postulated to cause a constitutively active receptor. Another black breed, Hampshire (*EP2/EP2*), possesses *MC1R* 2, which contains a missense mutation L99P postulated to cause a constitutively active receptor. Another black breed, Hampshire (*EP2/EP2*), possesses *MC1R* 3, associated with the missense mutation D121N. The recessive red coat color of swine was studied in Duroc (*e/e*) and found to be associated with *MC1R* 4. This harbors two missense mutations, one of which (A240T) is a strong candidate to disrupt receptor function.

An unresolved issue in our previous study concerned the molecular basis for black-spotted pigs (Kijas et al. 1998). The presence of black spots for a long time has been attributed to an allele of *E* and named *E* for partial extension of black. Black spots may occur on a white or red background (Figure 1A). Two breeds fixed for *E* (Pietrain and Large White) were both found to carry *MC1R* 3, the allele associated with *EP2* and the black color of Hampshire. We postulated that the *E* allele
contains a second mutation either in a codon not included in our previous study (first 40 codons and last 25 codons) or in a flanking regulatory region.

The objective of the present study was to identify the causative mutation for black spotting in E/F pigs by determining the entire MCIR coding sequence and parts of the flanking sequences. The mutation is shown to be a 2-bp insertion in codon 23 leading to a frameshift and a premature stop codon. A diagnostic test was used to screen for the presence of this mutation among various breeds presumed to carry the E allele (Figure 1). We also show that the black spots observed in E/F homozygotes are due to somatic reversions restoring the reading frame.

MATERIALS AND METHODS

Animals: Genomic DNA samples from pigs representing 13 populations/breeds and all described alleles at the Extension locus were used (Table 1; Figure 1). The samples included European and Asian wild boars exhibiting the wild-type color, two breeds with the dominant black color (Meishan and Hampshire), and three breeds with recessive red color (Duroc, Tamworth, and Hereford). Animals from six breeds were all assumed to be homozygous E/E. Pietrain pigs are white with black spots while Linderöd pigs, a native Swedish breed, have black spots on a red or white background. Berkshire and Gloucester Old Spot pigs are both assumed to be homozygous E/E and to originate from black-spotted pigs in the United Kingdom. However, Berkshire has been selected for extension of black and is today almost entirely black but with the breed characteristic of six white points (feet, tail, and snout). In contrast, the Gloucester Old Spot has been selected in the other direction and most individuals exhibit a few black spots. Large White and Landrace pigs are homozygous for E but white because they carry the Dominant white allele at KIT (Marklund et al. 1998). Samples from a three-generation intercross between the European wild boar and Large White domestic pigs were also used. This pedigree has been used extensively for coat color research (Johansson et al. 1992; Johansson Moller et al. 1996; Mariani et al. 1996; Kijas et al. 1998; Marklund et al. 1998).

Sequence analysis and mutation detection: The previously described bacterial artificial chromosome (BAC) clone 978E4 (Kijas et al. 1998) was used to subclone the 5′ and 3′ portions of MCIR. The entire coding sequence, 827 bp upstream of the ATG translation start site, and the 3′-untranslated region (UTR) were sequenced (GenBank accession no. AF326520). Primers EPIG16 (5′-GGG AAG CTG GAC CCC CGA GAG CGA GCC GCC-3′) and EPIG24 (5′-CAC GTT CTC CAC GAG GCT CAC CAG C-3′) were used to amplify a fragment containing 43 bp of 5′-UTR, the ATG translation start codon, and the previously unreported section of the MCIR 5′ coding region. The product resulted in a fragment of 234 bp for the E allele as opposed to a 232-bp fragment for other alleles. The PCR profile was 94° for 10 min followed by a touchdown profile: 95° for 10 sec, 65°–55° for 30 sec with a decrease of 2° per cycle, and 72° for 60 sec. The touchdown was followed by 32 cycles with annealing at 55°. The PCR products were detected by fragment analysis on an ABI377 (Applied Biosystems, Foster City, CA) and analyzed using the Genotyper software version 2.0.

RNA isolation and RT-PCR analysis: Skin samples were collected from four pigs homozygous E/E and representing two breeds: Pietrain pigs (white with black spots) and Linderöd pigs (red with black spots). Total RNA was isolated from skin samples using the TRIZOL reagent (Life Technologies) according to the manufacturer’s recommendation. mRNA was extracted using the PolyATtract mRNA isolation system (Promega, Madison, WI) and cDNA was generated applying the First-Strand cDNA synthesis kit (Amersham Pharmacia Biotech), all according to the manufacturer’s instructions. The first-strand reaction was primed with the Nol(dT)18 primer. The EPIG16 and EPIG24 primers and the reaction conditions described above were used for the RT-PCR analysis.

RESULTS

Sequence analysis of the 5′ and 3′ regions of MCIR: We previously used primers designed against evolution-
arily conserved regions of \textit{MC1R} to amplify 758 bp comprising the major part of the porcine \textit{MC1R} coding sequence but the analysis did not include the 5' and 3' ends (Kijas et al. 1998). These missing parts have now been sequenced using subclones from BAC 978E4 containing \textit{MC1R}; the sequence has been deposited in GenBank with accession no. AF326520. The result revealed a gene encoding a receptor three amino acids longer than the corresponding mammalian homologs (Figure 2). The length difference is due to an apparent tandem duplication of codons 29–31. The amino acid level sequences identified within the first 40 residues were highest for pig and human (70.0%), followed by pig/cattle (65.0%) and pig/mouse (55.0%). These values are lower than those previously reported for the middle part of \textit{MC1R} (average 79%), indicating that the amino terminal region is less well conserved.

The \textit{E'} allele is associated with a 2-bp insertion (nt67insCC) in \textit{MC1R}: Sequence data from the \textit{MC1R} 5' region were collected and compared between pig breeds known to carry different \textit{Extension} alleles: \textit{E}^{20} (Large Black), \textit{E}^{22} (Hampshire), \textit{E'} (Wild Boar), \textit{e} (Duroc), and \textit{E'} (Pietrain). The sequence comparison revealed the presence of an insertion of two C nucleotides at codon 23 (nt67insCC) in the \textit{MC1R} sequence associated with the \textit{E'} allele (Figure 3). The insertion causes a frameshift mutation that introduces a premature stop at codon 56. This defines a sixth allelic variant of \textit{MC1R} (*6) and distinguishes the \textit{E'} allele causing a black-spotted phenotype from \textit{ED2} for dominant black color, although they share the D121N missense mutation. The insertion of CC occurs in a GC-rich region and within a stretch of six Cs that is expanded to a mononucleotide repeat of eight Cs (Figure 3). The corresponding region in other mammalian species is interrupted by at least one T nucleotide. A nonsynonymous substitution was also identified at codon 17 in the \textit{ED1} allele (Figure 3).

The \textit{MCIR} nt67insCC mutation is associated with black spotting or red color across pig breeds: A PCR test was designed for typing the nt67insCC by a simple fragment size analysis. DNA samples representing 13 populations of pigs and all known \textit{Extension} alleles were analyzed (Table 1). A total of 45 animals homozygous for the alleles \textit{E}^{20}, \textit{E}^{22}, or \textit{E'} and thus solid colored were all found to be homozygous for the normal 232-bp fragment as expected. In contrast 65 animals representing six different breeds (Berkshire, Gloucester Old

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure1.png}
\caption{Illustration of pig coat color phenotypes associated with the \textit{E'} allele at the \textit{Extension/MC1R} locus. (A) F2 progeny from a wild boar/Large White intercross. The leftmost animal (\textit{E'}/\textit{E'}) is red with black spots; this is very similar to the phenotype observed in LinderoÈd pigs (not shown). The rightmost animal (\textit{E'}/\textit{E'}) is white with black spots; this is similar to the phenotype observed in Pietrain pigs (not shown). The second animal from the right is heterozygous \textit{E}^{1}/\textit{E'} and shows the wild-type color but with a black spot on the lower right side, indicating a somatic reversion. The second animal from the left is white due to the presence of the \textit{Dominant white} allele. (B) Tamworth. (C) Gloucester Old Spot. (D) Berkshire (photo copyright A. Christian and M. Rothschild, Iowa State University).}
\end{figure}
Figure 3.—Nucleotide and amino acid alignment of codons 1–40 of five different porcine MC1R alleles associated with different coat color phenotypes. The 2-bp insertion (nt67insCC) associated with the E<sup>p</sup> allele and a synonymous substitution at codon 17 in the E<sup>i</sup> allele are indicated. [MC1R<sup>*5</sup> is a wild-type allele found in the Japanese wild boar (Giuffra et al. 2000).] Identity to the master sequence is indicated by a dash.

Spot, Landrace, Large White, Linderöd, and Pietrain) somatic mutations restoring MC1R function. Genomic DNA were isolated from black spots of EP/EP homozygotes—presumed to be homozygous EP/EP—were all found to possess only the 234-bp fragment, indicating homozygosity for the 2-bp insertion. The screening also included three breeds that display a red phenotype and observed. This result is not surprising considering the fact that only a small fraction of the cells in the skin were therefore assumed to be fixed for the recessive e allele (Duroc, Hereford, and Tamworth). As expected, MC1R nt67insCC was not found in the sample of Duroc pigs, but surprisingly it was found to be present in both of the two MC1R copies. To further test the possibility of somatic reversions, skin samples from the red areas of the Hereford and Tamworth breeds (Table 1). Analysis of the entire coding sequence of MC1R in red animals homozygous for the 2-bp insertion confirmed that they were also homozygous for the D121N mutation associated with dominant black color, demonstrating that they are homozygous for the E<sup>p</sup>-MC1R<sup>*6</sup> allele. Moreover, these solid red E<sup>p</sup>/E<sup>p</sup> animals did not carry any additional mutation in the coding region that could explain the absence of black spots.

MC1R function is restored in E<sup>p</sup> homozygotes with black spots: nt67insCC leads to a frameshift at codon 23 and a premature stop codon. Frameshift mutations cause recessive red color in mice and cattle (Robbins et al. 1993; Klungland et al. 1995) and nt67insCC is thus expected to cause a recessive red phenotype, as observed for some of the Tamworth and Hereford pigs included in this study (animals either heterozygous 232/234 or homozygous 234/234). However, the presence of black spots on the red or white background observed in many E<sup>p</sup> homozygotes (Figure 1) is very unexpected given the nature of this mutation. To the best of our knowledge, spots of black pigment are not present in other mammals homozygous for MC1R loss-of-function mutations.

We postulated that the black spots may be due to somatic mutations restoring MC1R function. Genomic DNA were isolated from black spots of E<sup>p</sup>/E<sup>p</sup> homozygotes and analyzed with the diagnostic DNA test, but only the 234-bp fragment containing the insertion was observed. This result is not surprising considering the fact that only a small fraction of the cells in the skin are melanocytes (Junqueira et al. 1998) and we expect that a somatic mutation should be present in only one MC1R copy. To further test the possibility of somatic reversions, skin samples from the red areas and black spots of Linderöd pigs and from the white areas and black spots of Pietrain pigs were collected for RT-PCR analysis. mRNA was prepared and RT-PCR amplification was carried out across the MC1R region. The results clearly showed the presence of MC1R transcripts lacking the 2-bp insertion in mRNA from black spots despite the fact that PCR analysis of genomic DNA showed that the animals were homozygous for the insertion (Figure 4); the loss of the 2-bp insertion in the 232-bp cDNA fragment was confirmed by direct DNA sequencing.

The observation of transcripts both with and without the insertion from the black spots was not unexpected since partial expression of the normal transcript should be sufficient to restore the dominant black phenotype. Furthermore, it is possible that the data partly reflect MC1R expression in nonpigmentary cells present in skin (Adachi et al. 2000). Only the 234-bp fragment containing the CC insertion was observed using skin samples from the red areas of Linderöd pigs, as expected from the red coat color. Surprisingly, both the 232- and 234-bp fragments were obtained from the white areas of skin from Pietrain despite the lack of pigmentation (Figure 4). This may reflect some migration of
MC1R and Black Spotting in Pigs

The insertion clearly inactivates MC1R function and is thus expected to give a uniform red coat color as observed in the Tamworth and Hereford pigs carrying this allele. However, we also show that black spots on $E^v/E^v$ homozygotes express transcripts in which the normal reading frame has been restored. The black color of these spots is due to the presence of the D121N substitution. The wild-type $E^v$ allele is dominant to $E^s$, and $E^v/E^s$ heterozygotes from our Wild Boar intercross showed the wild-type color. However, a few black spots were generally observed in $E^v/E^v$ heterozygotes but not in $E^v/E^s$ homozygotes (Figure 1A).

Somatic DNA reversion is the most likely explanation for the presence of MC1R transcripts without the 2-bp insertion in mRNA from black spots. An alternative explanation is transcriptional slippage. In fact, transcriptional slippage has been reported to occur at mononucleotide repeats both in mammalian and bacterial cells (Linton et al. 1997; Larsen et al. 2000). However, the observation of distinct black spots strongly suggests that these reflect the expansion of clones with somatic DNA reversions. Furthermore, transcriptional slippage is expected to generate a distribution of transcripts with variable length whereas we observed only two fragment sizes (normal and reverted). It should be possible to distinguish these two possibilities by analyzing genomic DNA from primary melanocyte cell lines established from black spots.

As the insertion mutation represents an expansion of a mononucleotide tract, 6C to 8C, a possible mechanism of these mutations is replicational slippage, (nt67insCC) while the presence of the 232-bp fragment indicates somatic reversions. The scale for the detected fluorescence signal from PCR fragments is shown to the right.

DISCUSSION

Coat color phenotypes comprising numerous black spots occur in several mammalian species such as Dalmatian dogs, Leopard-spotted horses, and the pig. Previous studies have clearly indicated that porcine black spotting is controlled by an allele at the Extension/MC1R locus, $E^v$. However, it remained unresolved how an $MC1R$ mutation could cause a red coat color with distinct black spots since mutations at this locus usually give uniform dominant black color or recessive red color. This is now explained by our observation of two functionally significant $MC1R$ mutations in the $E^v$ allele, a 2-bp insertion (nt67insCC) causing a frameshift unique to this allele and a missense mutation (D121N), shared with the $E^{v2}$ allele associated with dominant black color. The insertion clearly inactivates MC1R function and is thus expected to give a uniform red coat color as observed in the Tamworth and Hereford pigs carrying this allele. However, we also show that black spots on $E^v/E^v$ homozygotes express transcripts in which the normal reading frame has been restored. The black color of these spots is due to the presence of the D121N substitution. The wild-type $E^v$ allele is dominant to $E^s$, and $E^v/E^s$ heterozygotes from our Wild Boar intercross showed the wild-type color. However, a few black spots were generally observed in $E^v/E^v$ heterozygotes but not in $E^v/E^s$ homozygotes (Figure 1A).

Somatic DNA reversion is the most likely explanation for the presence of $MC1R$ transcripts without the 2-bp insertion in mRNA from black spots. An alternative explanation is transcriptional slippage. In fact, transcriptional slippage has been reported to occur at mononucleotide repeats both in mammalian and bacterial cells (Linton et al. 1997; Larsen et al. 2000). However, the observation of distinct black spots strongly suggests that these reflect the expansion of clones with somatic DNA reversions. Furthermore, transcriptional slippage is expected to generate a distribution of transcripts with variable length whereas we observed only two fragment sizes (normal and reverted). It should be possible to distinguish these two possibilities by analyzing genomic DNA from primary melanocyte cell lines established from black spots.

As the insertion mutation represents an expansion of a mononucleotide tract, 6C to 8C, a possible mechanism of these mutations is replicational slippage, (nt67insCC) while the presence of the 232-bp fragment indicates somatic reversions. The scale for the detected fluorescence signal from PCR fragments is shown to the right.
There are more than 10 coat color mutations in the mouse that show phenotypic reversion spots (Silvers 1979). These include the recessive *pink-eyed unstable* allele (Gondo et al. 1993) and some dominant *W/Kit* alleles (De Sepulveda et al. 1995). However, the *E* allele in pigs is unique as regards the high frequency of reversion spots and because it is a small duplication within a mononucleotide repeat. For instance, the genetically unstable *pearl* allele in the mouse is caused by a tandem duplication of 739 bp in the *Ap3bl* gene (Feng et al. 1999).

The coat color phenotype associated with the *E*/*E* genotype is influenced by other loci and it has apparently been modified by selection in some breeds. White pigs of the Large White and Landrace breeds do not show black spots because of epistatic interaction of the *Dominant white/KIT* alleles causing a defect in melanocyte migration (Marklund et al. 1998). In the absence of *Dominant white* alleles, black spots are observed in some breeds (Pietrain, Linderöd) but not in others (Tamworth, Hereford). It is noteworthy that Tamworth was originally (late 18th, early 19th century) red with black spots but nowadays pigs with black spots are excluded from breed books so that there is selection against spots in this breed (Porter 1993). We saw both *e* and *E* alleles in the Tamworth samples, both of which resulted in red coat color. The description of the development of the Tamworth breed suggests that the red coat in this breed resulted from *E*; it may be that the presence of the *e* allele is the result of a more recent introgression of Duroc into this breed in some regions. The presence of black spots in some pig breeds but not in others is probably explained by genetic differences in the underlying mechanism controlling the rate of somatic mutations or the expansion of revertant clones. A similar effect of the genetic background on the presence/absence of reversion spots has been reported for the *pearl* coat color mutation in mouse (Russel and Major 1956). There has also been selection for modifying the extension of black spotting in the Berkshire breed as well as the Gloucester Old Spot (Figure 1). In Berkshire, this selection has resulted in animals that are entirely black but with six white points (four feet, nose, and tail). An opposite selection pressure has created an almost completely white phenotype with a few black spots in the Gloucester Old Spot. Interestingly, it was already suggested by Sewall Wright in 1918 that the black color of Berkshire was an extended form of black spotting (Wright 1918). This was later confirmed by Hetzer (1954), who demonstrated that the degree of black spotting could be efficiently manipulated via selection in both directions, resulting in a range of phenotypes from only two or three black spots to being completely black. Such selection may influence both the frequency of somatic mutations and the expansion of revertant clones, as well as the occurrence of transcriptional slippage. We have not had access to skin samples of Berkshire pigs to test the latter possibility.

The black spots in *E*/*E* pigs may occur on a red or white background (Figure 1). We have previously shown that the background color is controlled by another locus (loci) since full-sibs sharing the same *E* alleles identical by descent exhibited a red or white background (Martani et al. 1996). The mode of inheritance for this phenotypic variation has not yet been established. The white background may be caused by mutations in one or more genes required for pheomelanin but not eumelanin synthesis, resembling the coat color variants *fading yellow* in guinea pigs (Wright 1917), *gray-lethal*, and *grizzled* in the mouse (Silvers 1979). An alternative explanation for the white background is that it is a defect in melanocyte migration/survival in the absence of functional *MCIR* expression. Data in the mouse indicate that the *e/Mc1r* locus is a modifier for *piebald* and affects the amount of white spotting (Lamoreux and Russell 1979; Payan et al. 1995). Furthermore, white leg markings are more extensive in chestnut (*e/e*) horses than in horses with the dominant *E* allele (Woolf 1995), also suggesting that *MCIR* function contributes to the survival/migration of melanoblasts.

A model involving a defect in melanocyte migration/survival gains support by the interesting observation that the black spots are consistently larger on a white background than on a red background (Figure 1), suggesting that revertant clones expand more extensively in the absence of melanocytes in white areas. This would be consistent with the observation of a more extensive coat color pigmentation when murine neural crest-derived cells of pigmented C57BL/6J origin are injected *in utero* into *Kit* mutant embryos lacking melanocytes, compared with the situation when the same cells are injected into BALB/c embryos containing unpigmented melanocytes (Huszár et al. 1991). A general absence of melanocytes from white areas may also suggest an explanation for the observation that we detected mutant *MCIR* transcripts only from red areas in which melanocytes evidently are present, but detected both normal and mutant transcripts from white areas (Figure 4). RT-PCR may have amplified transcripts from a few revertant melanocytes that migrated into white areas from the surrounding black spots. The two possible explanations for the white background could be distinguished by histological examination of the white areas of *E*/*E* pigs with black spots since the model with a mutation affecting pheomelanin synthesis predicts the presence of melanocytes without pigment while the latter model implies the absence of melanocytes.

Mammalian coat color genetics has served as a model for studying gene action and interaction since the beginning of the last century (Searle 1968; Silvers 1979). This is well illustrated by this study where we describe a complex *MCIR* allele containing two functionally significant mutations, one frameshift and one missense
mutation. The $E^a$ allele is assumed to determine a red coat color but the phenotype ranges from red, red with black spots, white with few black spots, white with many black spots, to almost completely solid black due to somatic reversions and the action of modifying loci.

Sincere thanks to Leena Sahlström and Dr. Erik Bongcam-Rudloff for valuable assistance, to Dr. Tetsuo Hori, Ms. Guy Kiddy, Dr. Max Rothschild, and Mrs. S. Bowser for providing DNA samples from Japanese wild boar, Gloucester Old Spot, Hereford, and Tamworth pigs, and to Dr. Richard Wales for valuable comments on the article.

The study was supported by the Swedish Research Council for Forestry and Agriculture.

LITERATURE CITED


Johansson Moller, M., R. Chaudhary, E. Hellmen, B. Hoyheim, B. Chowdhary et al., 1996 Pigs with the dominant white coat color phenotype carry a duplication of the KIT gene encoding the mast/stem cell growth factor receptor. Mammm. Genome 7: 822–830.


Vage, D. I., D. Lu, H. Klungland, S. Lien, S. Adalsteinsson et al., 1997 A non-epistatic interaction of agouti and extension in the coat color but the phenotype ranges from red, red with black spots, white with few black spots, white with many black spots, to almost completely solid black due to somatic reversions and the action of modifying loci.


Communicating editor: C. Haley