ANIMAL GENETICS Immunogenetics, Molecular Genetics and Functional Genomics

doi: 10.1111/age.12006

A *de novo* mutation in *KIT* causes white spotting in a subpopulation of German Shepherd dogs

A. K. Wong*¶, A. L. Ruhe*, K. R. Robertson*, E. R. Loew†, D. C. Williams‡ and M. W. Neff*§

*Veterinary Genetics Laboratory, University of California, Davis, CA, 95616, USA. †Department of Biomedical Sciences, Cornell University, Ithaca, NY, 14853, USA. ‡William R. Pritchard Veterinary Medical Teaching Hospital, University of California, Davis, CA, 95616, USA. §Van Andel Research Institute, Grand Rapids, MI, 49503, USA. ¶Department of Computer Science, Princeton University, 35 Olden St, Princeton, NJ, 08540, USA.

Summary

Although variation in the *KIT* gene is a common cause of white spotting among domesticated animals, *KIT* has not been implicated in the diverse white spotting observed in the dog. Here, we show that a loss-of-function mutation in *KIT* recapitulates the coat color phenotypes observed in other species. A spontaneous white spotting observed in a pedigree of German Shepherd dogs was mapped by linkage analysis to a single locus on CFA13 containing *KIT* (pairwise LOD = 15). DNA sequence analysis identified a novel 1-bp insertion in the second exon that co-segregated with the phenotype. The expected frameshift and resulting premature stop codons predicted a severely truncated c-Kit receptor with presumably abolished activity. No dogs homozygous for the mutation were recovered from multiple intercrosses ($P = 0.01$), suggesting the mutation is recessively embryonic lethal. These observations are consistent with the effects of null alleles of *KIT* in other species.

Keywords canine familiaris, coat color, melanocyte, piebald

Introduction

Piebald is a pattern of non-pigmentation and white spotting in mammals. The white spotting is caused by defects in the development and survival of melanoblasts, the neural-crest-derived precursors of pigment-producing melanocytes (reviewed in Jackson 1994; Steingrímsson *et al.* 2006). The genes responsible for white spotting typically influence multiple developmental processes and instruct various cell types; thus, they tend to be pleiotropic, often with deleterious effects. Across species, white spotting can be associated with anemia, deafness, mast cell deficiency, infertility, aganglionic megacolon and memory deficits (Jackson 1994). Artificial selection for white spotting presumably enriches for genes and alleles that avoid or resolve such adverse effects.

Artificial selection has produced many diverse coat colors in the dog that have been selectively bred for centuries, with distinctive patterns often serving as breed-defining hallmarks. Extensive variation in white spotting exists across breeds, with much of the phenotypic variation attributable to *MITF* (Karlsson *et al.* 2007), a master regulator of melanoblast migration and survival. *KIT*, encoding an essential component in the same melanogenesis pathway as MITF, is responsible for white spotting patterns in many mammals, including the pig (Marklund *et al.* 1998), horse (Brooks & Bailey 2005; Haase *et al.* 2007), human (Giebel & Spritz 1991), cow (Reinsch *et al.* 1999) and cat (Cooper *et al.* 2006). Interestingly, *KIT* has not been implicated in coat color phenotypes in the dog. Several genetic studies viewed *KIT* as a strong candidate for white spotting patterns but subsequently excluded the gene (Metallinos *et al.* 1998; van Hagen *et al.* 2004). The conspicuous absence of *KIT* variation in white spotting in the dog suggests that the biology of *KIT* may be modified in this species, such that thresholds for phenotypic effect have been altered.

Here, we show that *KIT* governs the dominant white spotting phenotype that spontaneously arose in a pedigree of German Shepherd dogs. The pattern includes white markings on the face, ventral abdomen, feet and tip of the tail, similar to the patterns observed in other species. We mapped this occurrence by linkage to a region encompassing canine *KIT*, and we identified an insertion that causes a frameshift early in the open reading frame.

Address for correspondence

M. W. Neff, Van Andel Research Institute, Grand Rapids, MI 49503, USA.
E-mail: mark.neff@vai.org

Accepted for publication 02 September 2012

Materials and methods

DNA, pedigree and phenotyping

The German Shepherd pedigree segregating the phenotype (Fig. 1) was ascertained through the breeder of the

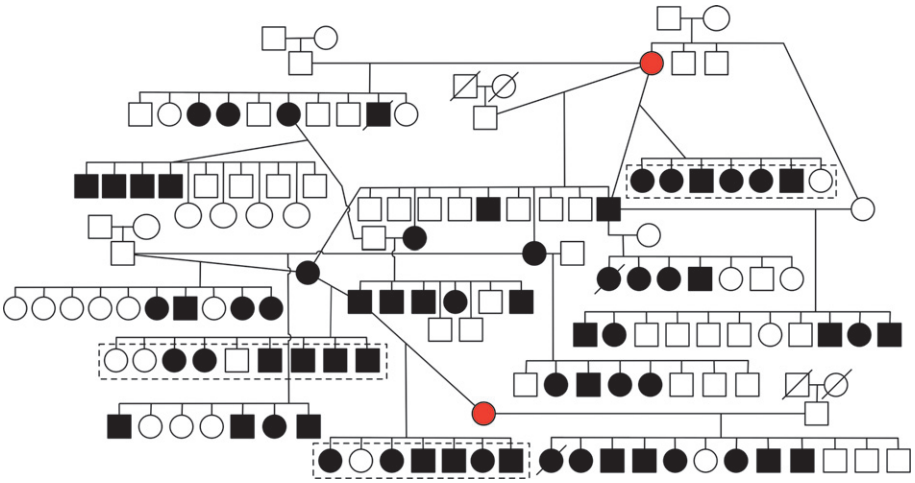

Figure 1 Pedigree segregating the spontaneous white spotting. Dogs with white spotting are denoted by solid black symbols. The proband is indicated in red and is represented twice for diagrammatic purposes. The German Shepherd bloodline was managed by private breeders interested in propagating the phenotype and hence the preferential breeding of spotted dogs. Litters from three intercrosses, informative for inferring a recessive lethal phenotype, are enclosed in dashed-line boxes.

proband. Blood and buccal swab samples from 135 pedigree members were converted to genomic DNA using standard methods (Bell *et al.* 1981; Oberbauer *et al.* 2003). The white spotting phenotype was qualitatively scored from photographs. Hearing was tested by electrophysiological recordings of the brainstem auditory evoked response (BAER) according to a previously described protocol (Strain *et al.* 1991).

Genotyping

DNA samples were genotyped using CA-repeat markers that were previously screened during assembly of a comprehensive canine linkage map (Wong *et al.* 2010). The set consisted of 143 markers selected to provide 20-cM sex-averaged spacing. If this first screen proved insufficient, additional markers were in abeyance to be applied to any gaps in coverage. The markers are described in Table S1, and their physical distribution is shown in Fig. S1. Markers for ensuring coverage, fine-mapping and haplotype analyses were available through DOGSET, an online database of *in silico* markers for the dog that computationally mines microsatellite loci with perfect CA repeats ($n \geq 12$ units) from the reference genome (Wong & Neff 2009). Because the Mendelian phenotype was successfully mapped in the primary screen, no additional markers were needed to address holes in genome coverage. Microsatellite markers were amplified by multiplex PCR using a previously described M13-tailed primer protocol (Oetting *et al.* 1995). Amplicons were separated by capillary electrophoresis and detected by fluorescence using ABI 3730 instruments (Applied Biosystems). Genotypes were automatically scored and manually verified with GENEMAPPER 4.0 (Applied Biosystems).

DNA sequencing

Annotated exons (www.ensembl.org) were interrogated for polymorphism by fluorescence-based DNA resequencing of PCR products. Site-specific primers for resequencing were downloaded from DOGSET (Wong & Neff 2009) and synthesized with T3 and T7 primer sequences appended to the 5′ ends of the forward and reverse oligonucleotides, respectively, to facilitate direct sequencing of amplicons. The BIGDYE TERMINATOR v3.1 Cycle Sequencing Kit (Applied Biosystems) was used to prepare product for analysis on ABI 3730 instruments. Sequence data were analyzed for variants using SNPDETECTOR (Zhang *et al.* 2005).

Computation and data analysis

Genotypes were screened for Mendelian inheritance using PEDCHECK (O'Connell & Weeks 1998). LOD scores were calculated using the twopoint option of CRI-MAP (Green 1990). SIMWALK was used to assess phase and haplotypic configurations (Sobel & Lange 1996). The chi-square test was used to assess statistical significance with the null hypothesis of expected genotypic ratios (1:2:1) for intercross progeny. Fisher's exact test was used to calculate the *P*-value for phenotype:genotype association.

Results

The proband, a female German Shepherd dog of purebred ancestry, was born to traditionally colored parents and was the only puppy with unusual white markings in a litter of four. The pattern of non-pigmentation was reminiscent of piebaldism – the dog's forelock and chest were white, and the feet and tail were white-tipped. The markings were

mostly symmetrical, but the boundaries of white patches were often irregular. The initial cross was repeated four times, but failed to produce any other offspring with white spotting (i.e., 38 normal sibs). This result was consistent with a *de novo* mutation event.

The proband produced mixed litters when crossed to unrelated normal sires, suggesting an autosomal dominant mode, which was supported by the outcomes of subsequent crosses (Fig. 1). Offspring with the pattern were similar in appearance to the proband, and although there was noticeable variation in the amount of white spotting and in the degree of symmetry of markings (Fig. 2), there was no evidence of segregating modifiers; no discernible systematic differences were observed in the white spotting phenotype of progeny from litters produced by outcrossing. Spotted offspring appeared to be healthy and exhibited normal hearing, as assayed by BAER (Fig. S2).

Genome-wide mapping and mutation discovery

A collection of 143 microsatellite markers spanning the genome (Table S1, Fig. S1) was genotyped on 95 pedigree members, and the resulting data were analyzed for linkage. A single linked locus was detected on CFA13 (two-point LOD = 15, recombination fraction = 0.03; Fig. 3). Twelve representative microsatellite loci spanning 10 Mb from the mapped region (CFA13:46.0–56.2 Mb, canFam2) were chosen to track the ancestral haplotype upon which the *de novo* 1-bp insertion occurred (Table S2). Genotyping localized discrete crossovers and refined the interval by haplotype analysis. All white-spotted dogs shared an identical-by-descent haplotype that could be traced to the sire of the proband (Fig. 4).

The fine-mapped region (CFA13:49.5–51.7 Mb; which included *KIT*) is the canine ortholog of the classic dominant white (W) locus in the mouse. DNA sequencing of exons revealed a 1-bp insertion of an adenine 70 bases downstream of the beginning of exon 2 (Fig. S3). Subsequent genotyping showed that the insertion was perfectly associated with the phenotype (64 spotted dogs and 31 normal littermates; $P < 0.001$). The mutation was a *de novo* event – observed in the proband but absent in both parents.

Functional inference

The insertion occurred in codon 47 (of 976 codons) and was expected to produce a frameshift with the first downstream premature stop at codon 57. Consequently, the allele was predicted to produce a severely truncated product (55/974 amino acid residues; Fig. S4). Given previous characterizations of c-Kit, the mutant protein was likely to be non-functional, implying that the dominant white spotting resulted from haploinsufficiency.

Because *KIT* is an essential gene, offspring having homozygous null alleles were predicted to be nonviable. Of 23 progeny produced from intercrosses, none were genotypically homozygous for the insertion. The absence of homozygotes in the F_2 was significant ($P = 0.01$): it suggested that the insertion and resulting frameshift were recessive lethal, causing early embryonic death.

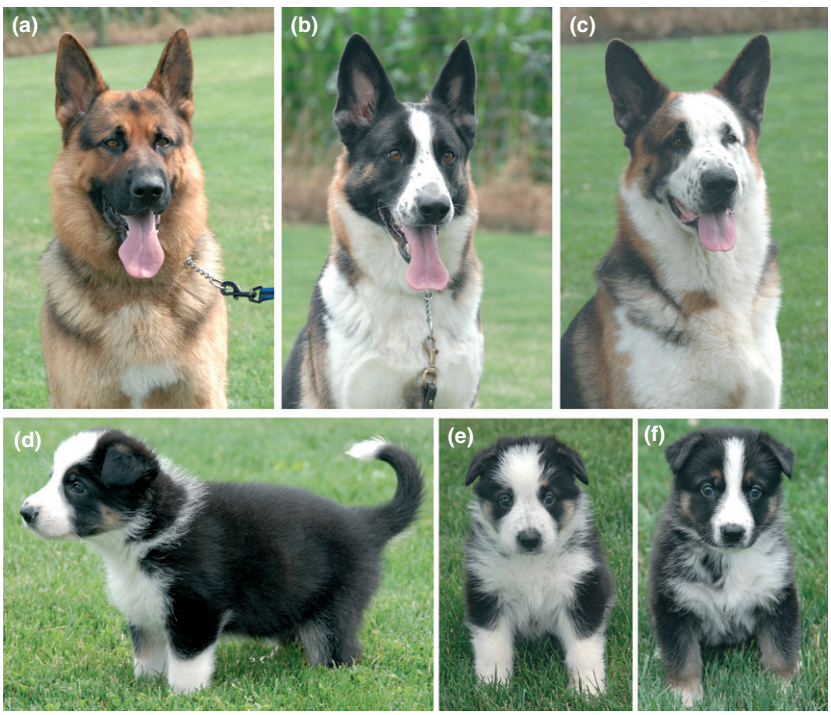

Figure 2 The white spotting phenotype. A normal (a) and two white-spotted German Shepherd dogs (b and c) are shown. The minor white patch on the chest of the dog shown in (a) is independent of the *KIT* mutation. (d) A white-spotted puppy shows the characteristic markings, with white forelock, muzzle, chest, ventral abdomen, collar and tip of the tail. The expressivity of the trait is reflected in the comparison of a pair of adult dogs (b and c) and a different pair of puppies (e and f).

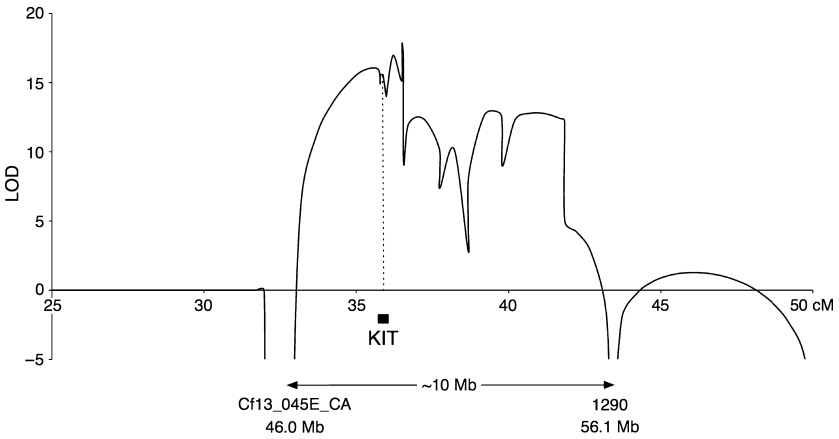

Figure 3 Multipoint LOD for region of CFA13. The LOD score is plotted on the *y*-axis; the *x*-axis plots the physical coordinates from canFam2.

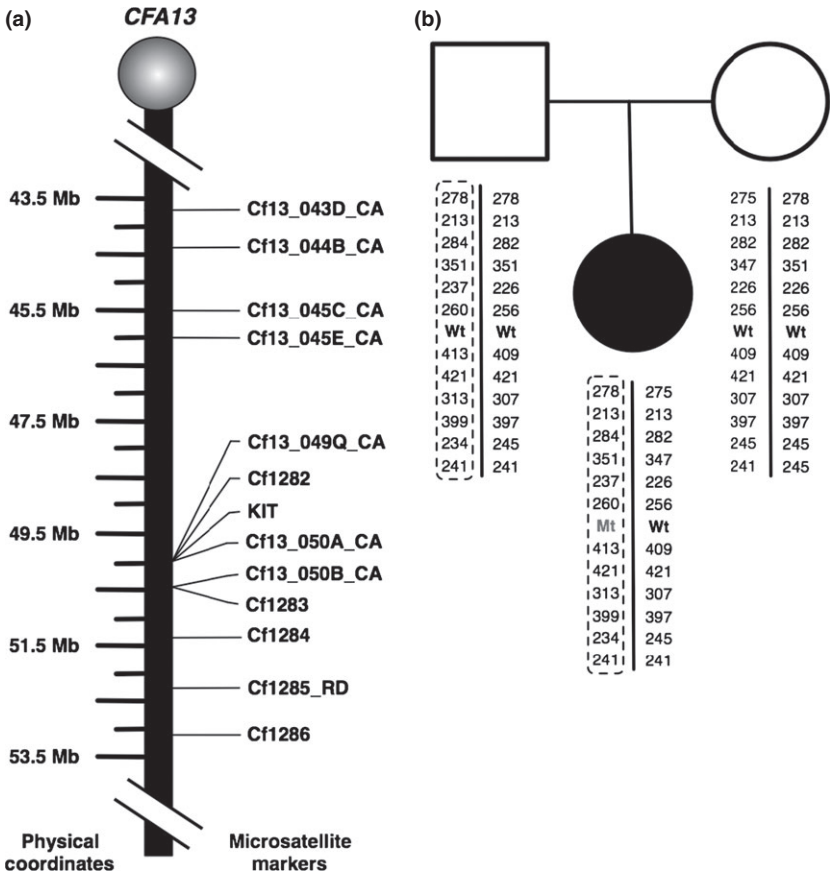

Figure 4 The *KIT* frameshift mutation haplotype analysis. (a) Fine-mapped region of CFA13 with physical coordinates (canFam2.0) of marker loci. (b) Haplotypic configuration of marker alleles for the parents and proband.

Discussion

We have found that a dominant white spotting pattern in the dog stemmed from a putative null allele of *KIT*. The molecular basis of the phenotype can be leveraged from earlier studies of white spotting in other mammals, most notably from the classical genetics of the mouse (reviewed in Jackson 1994). Historically, the *W* locus (i.e., *KIT*) was among several genes (e.g., *MITF*, *SOX10* and *PAX3*) that produced dominant white spotting patterns. These genes govern multiple developmental processes, with melanogenesis consistently being the most sensitive to dosage effects. *KIT*-based spotting in the German Shepherd dogs parallels the phenotypic outcomes of null alleles of *W*, where haploinsufficiency reduces c-Kit signaling, causing defects in the proliferation, migration and survival of neural-crest-derived melanoblasts. This deficit results in areas of the skin being devoid of melanocytes and thus non-pigmented. The *de novo* mutation in *KIT* has the same threshold effects on melanogenesis that are observed in other species.

Pleiotropism

Studies of white spotting in the mouse have shown that the genes responsible can have deleterious effects. Murine *KIT* mutations, for instance, can cause anemia, deafness, mast cell deficiency, infertility and memory deficits. White spotting patterns in the dog historically have been associated with a low-penetrance syndromic deafness, similar to the deafness observed in the mouse. This defect is now understood to result from a stochastic failure of melanoblasts to colonize the stria vascularis of the inner ear (Steel & Barkway 1989; Price & Fisher 2001). *MITF*, the major white spotting locus in the dog (Karlsson *et al.* 2007), may contribute to the risk of deafness in mostly white dogs. We did not observe any deleterious effects of the mutation in the German Shepherd dogs.

Interestingly, null alleles of *KIT* cause milder phenotypes than do hypomorphic mutations (e.g., within the juxtamembrane or tyrosine kinase domains depicted in Fig. S4), at least in the mouse (Jackson 1994; Steingrímsson *et al.* 2006). This observation may be explained by the fact that c-Kit dimerizes upon activation by ligand (i.e., *KITLG*, *KIT ligand*; also called stem cell factor or SCF) and thus affords a dominant-negative effect. The 1-bp insertion in *KIT* in the German Shepherd dogs predicts that <6% of the native peptide sequence is produced. The truncated product is unlikely to interact with wild-type product and thus may confer milder phenotypes. The German Shepherd *KIT* allele has the same effect as deletion alleles in the mouse, both in the hemizygous (i.e., dominant white spotting) and in the homozygous (i.e., recessive lethal) states. This suggests that in the dog (as in the mouse), as little as 40% of signaling activity from the receptor tyrosine kinase may be sufficient for *KIT*-dependent processes other than melanogenesis (Pielberg *et al.* 2002).

Expressivity of white spotting

The variable expressivity of the white spotting pattern was considerable (Fig. 2), and it could result either from stochastic developmental processes or from heritable variation. Pavan *et al.* (1995) exploited the variable expressivity of piebald mice to identify modifiers that were attributable to strain background differences (Dunn & Charles 1937). A similar approach in the dog could identify additional variants segregating in the German Shepherd dog, which interact genetically with the *KIT* allele.

Efficient scanning set for canine linkage analysis

Mendelian traits and diseases are effectively mapped in the dog using linkage analysis. This approach plays to the strengths of canine pedigrees, which typically include multiple generations of extant dogs, relatively large litter (sibship) sizes and ascertainment by the breeder (Neff & Rine 2006). These familial attributes conferred mapping power to the study reported here. The comprehensive canine linkage map (Wong *et al.* 2010) enabled multipoint analysis. The observed negative LOD scores (Table S1), which are common in multipoint linkage, reflect a well-powered experiment. A negative LOD score indicates greater support for a marker being unlinked (null hypothesis) to the phenotype than for showing evidence of linkage (thus, the ratio is <1 and the LOD is <0). By convention, a LOD <−2 represents statistical significance for excluding regions in candidate gene studies. Thus, if sufficient mapping power were obtained, most of the genome would yield a negative LOD score when mapping a Mendelian trait. It should be noted, however, that the observed number of recombination events between adjacent markers and a causal mutation may not reflect actual genetic distance owing to insufficient genotype data or genotyping errors. In these cases, likelihood of the multipoint 'model' may be less than the null model and also produce a negative LOD. Nonetheless, the results reported here reflect a well-powered linkage study.

The marker set assembled for this study enabled an efficient linkage analysis of a pedigree of 96 dogs in <1 week. The cost of the entire analysis was less than that of assaying a single fixed SNP array on a minimal sample set of 12 dogs. The microsatellite-based linkage scanning set is thus a reasonable approach for studying a rare mutation that segregates within a pedigree and for mapping *de novo* mutations that have not yet become broadly distributed in a breed gene pool.

Acknowledgements

We thank breeders C. McCann, M. Lopizzo and S. Estell, as well as numerous owners, for participating in the study and submitting DNA and information from their dogs. We thank S. Okamura, C. Penedo and D. Nadziejka for helpful comments. This work was financially supported by the Veterinary Genetics Laboratory of the University of California, Davis, and by the Van Andel Research Institute in Grand Rapids, Michigan.

References

Bell G.I., Karam J.H. & Rutter W.J. (1981) Polymorphic DNA region adjacent to the 5′ end of the human *insulin* gene. *Proceedings of the National Academy of Sciences of the United States of America*, **78**, 5759–63.

Brooks S.A. & Bailey E. (2005) Exon skipping in the *KIT* gene causes a Sabino spotting pattern in horses. *Mammalian Genome*, **16**, 893–902.

Cooper M.P., Fretwell N., Bailey S.J. & Lyons L.A. (2006) White spotting in the domestic cat (*Felis catus*) maps near *KIT* on feline chromosome B1. *Animal Genetics*, **37**, 163–5.

Dunn L.C. & Charles D.R. (1937) Studies on spotting patterns I. Analysis of quantitative variations in the pied spotting of the house mouse. *Genetics*, **22**, 14–42.

Giebel L.B. & Spritz R.A. (1991) Mutation of the *KIT* (mast/stem cell growth factor receptor) protooncogene in human piebaldism. *Proceedings of the National Academy of Sciences of the United States of America*, **88**, 8696–9.

Green P., Fall K., Crooks S. (1990) Documentation for CRI-MAP, version 2.4.

Haase B., Brooks S.A., Schlumbaum A. *et al.* (2007) Allelic heterogeneity at the equine *KIT* locus in dominant white (W) horses. *PLoS Genetics*, **3**, e195.

van Hagen M.A., van der Kolk J., Barendse M.A., Imholz S., Leegwater P.A., Knol B.W. & van Oost B.A. (2004) Analysis of the inheritance of white spotting and the evaluation of *KIT* and *EDNRB* as spotting loci in Dutch boxer dogs. *Journal of Heredity*, **95**, 526–31.

Jackson I.J. (1994) Molecular and developmental genetics of mouse coat color. *Annual Review Genetics*, **28**, 189–217.

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

Marklund S., Kijas J., Rodriguez-Martinez Heriberto. et al. (1998) Molecular basis for the dominant white phenotype in the domestic pig. *Genome Research*, **8**, 826–33.

Metallinos D.L., Bowling A.T. & Rine J. (1998) A missense mutation in the *endothelin-B receptor* gene is associated with lethal white foal syndrome: an equine version of Hirschsprung disease. *Mammalian Genome*, **9**, 426–31.

Neff M.W. & Rine J. (2006) A fetching model organism. *Cell* **124**(2), 229–31.

Oberbauer A.M., Grossman D.I., Eggleston M.L., Irion D.N., Schaffer A.L., Pedersen N.C. & Belanger J.M. (2003) Alternatives to blood as a source of DNA for large-scale scanning studies of canine genome linkages. *Veterinary Research Communications*, **27**, 27–38.

O'Connell J.R. & Weeks D.E. (1998) PEDCHECK: a program for identification of genotype incompatibilities in linkage analysis. *American Journal of Human Genetics*, **63**, 259–66.

Oetting W.S., Lee H.K., Flanders D.J., Wiesner G.L., Sellers T.A. & King R.A. (1995) Linkage analysis with multiplexed short tandem repeat polymorphisms using infrared fluorescence and M13 tailed primers. *Genomics*, **30**, 450–8.

Pavan W.J., Mac S., Cheng M. & Tilghman S.M. (1995) Quantitative trait loci that modify the severity of spotting in piebald mice. *Genome Research*, **5**, 29–41.

Pielberg G., Olsson C., Syvänen A.C. & Andersson L. (2002) Unexpectedly high allelic diversity at the *KIT* locus causing dominant white color in the domestic pig. *Genetics*, **160**, 305–11.

Price E.R. & Fisher D.E. (2001) Sensorineural deafness and pigmentation genes: melanocytes and the *MITF* transcriptional network. *Neuron*, **30**, 15–8.

Reinsch N., Thomsen H., Xu N. *et al.* (1999) A QTL for the degree of spotting in cattle shows synteny with the *KIT* locus on chromosome 6. *Journal of Heredity*, **90**, 629–34.

Sobel E. & Lange K. (1996) Descent graphs in pedigree analysis: applications to haplotyping, location scores, and marker-sharing statistics. *American Journal of Human Genetics*, **58**, 1323–37.

Steel K.P. & Barkway C. (1989) Another role for melanocytes: their importance for normal stria vascularis development in the mammalian inner ear. *Development*, **107**, 453–63.

Steingrímsson E., Copeland N.G. & Jenkins N.A. (2006) Mouse coat color mutations: from fancy mice to functional genomics. *Developmental Dynamics*, **235**, 2401–11.

Strain G.M., Jackson R.M. & Tedford B.L. (1991) Postnatal development of the visual-evoked potential in dogs. *American Journal of Veterinary Research*, **52**, 231–5.

Wong A.K. & Neff M.W. (2009) *DOGSET*: pre-designed primer sets for fine-scale mapping and DNA sequence interrogation in the dog. *Animal Genetics* **40**, 569–71.

Wong A.K., Ruhe A.L., Dumont B.L. *et al.* (2010) A comprehensive linkage map of the dog genome. *Genetics* **184**, 595–605.

Zhang J., Wheeler D.A., Yakub I., Wei S., Sood R., Rowe W., Liu P.P., Gibbs R.A. & Buetow K.H. (2005) SNPDETECTOR: a software tool for sensitive and accurate SNP detection. *PLoS Computational Biology*, **1**, e53.

Supporting information

Additional supporting information may be found in the online version of this article.

Figure S1 Genome-wide distribution of microsatellite markers included in the genome scan for linkage analysis.

Figure S2 Graphs of brainstem auditory evoked response (BAER) results.

Figure S3 DNA sequence polymorphism and mutation detection.

Figure S4 Structure of the c-Kit receptor.

Table S1 Microsatellite markers in the genome scanning set for linkage analysis.

Table S2 Microsatellite markers for fine-mapping of the *KIT* locus.