Localization of White Spotting Locus in Boxer Dogs on CFA20 by Genome-Wide Linkage Analysis with 1500 SNPs

PETER A. LEEGWATER, MARJAN A. VAN HAGEN, AND BERNARD A. VAN OOST

From the Department of Veterinary Medicine of Companion Animals, Faculty of Veterinary Medicine, Utrecht University, PO Box 80154, 3508 TD Utrecht, The Netherlands. Marjan A. van Hagen is now at the Dutch Kennel Club, Amsterdam, The Netherlands. Bernard A. van Oost is now at the Department of Molecular Cell Biology, American University of the Caribbean, Sint Maarten, Netherlands Antilles.

Address correspondence to P. Leegwater at the address above, or e-mail: p.a.j.leegwater@vet.uu.nl.

Abstract

New techniques allow fast genotyping of large numbers of single-nucleotide polymorphisms (SNPs) of the genome. These techniques are used to map disorders with complex inheritance patterns and require large study groups. Linkage analysis of monogenetic traits exploits close family relationships between relatively small numbers of cases and controls. Linkage studies are typically performed with a set of microsatellite markers spaced at 10 cM. We were interested to test whether SNP typing could be applied in genome-wide linkage analysis because of the speed of the procedure. White spotting in Boxer dogs was chosen as a model because it is a semidominant trait, allowing the assignment of locus genotypes to each phenotyped dog. A set of just more than 1500 SNPs were typed in 5 families with heterozygous parents and offspring that included 11 white, 6 brown, and 19 spotted dogs. Multipoint linkage analysis was performed and a LOD score of 12.1 was obtained on canine chromosome 20. The CFA20 region was the only region with a positive LOD score. The gene MITF, coding for a transcription factor implicated in Waardenburg syndrome in humans, is located in the region close to a SNP that is in apparent linkage disequilibrium with the white spotting locus. Thus, MITF is a likely candidate for involvement in white spotting in boxers. We conclude that SNPs, spaced at an average distance of 1.6 Mb, are highly informative in linkage analysis of monogenic traits and are a powerful alternative to microsatellite markers.

Linkage analysis is based on the comparison of the segregation of a trait through a pedigree with the inheritance of alleles of polymorphic DNA markers. Typically, 300–400 microsatellite markers are used to map the inheritance of complete genomes. Even with the use of pipetting robots and automated detection of polymerase chain reaction (PCR) products, genome screens with microsatellite markers are laborious and time consuming. Presently, techniques are available to genotype many single-nucleotide polymorphisms (SNPs) in a single experiment. These techniques facilitate genetic association studies of complex traits in populations. Because of the speed of the procedure, we wondered whether SNP typing could be effective in family-based linkage analysis.

White spotting in Boxer dogs is a monogenic, semidominant trait (Van Hagen et al. 2004). The breed standard allows the partial white coloring of heterozygous dogs (flashy) but not the extreme white phenotype. As a consequence, approximately 4% of boxer puppies of a Dutch birth cohort were

euthanized due to white coloring (Nielen et al. 1998). The well-known boxer Tasha, of which the genomic DNA has been entirely sequenced, has the flashy phenotype. Earlier, the KIT and EDNRB genes have been excluded from involvement in white spotting in boxers (Van Hagen et al. 2004).

Materials and Methods

Dogs

Sires and dams of litters from the Dutch Boxer cohort that contained at least one extreme white boxer were phenotyped. A family was included in the study when the parents had the flashy heterozygous phenotype. The material included 11 extreme white, 19 flashy, and 6 brown offspring from 5 sires and 6 dams (Figure 1). Blood samples were obtained of all dogs except 2 sires. The DNA was isolated from the samples by the salt extraction method (Miller 1988).
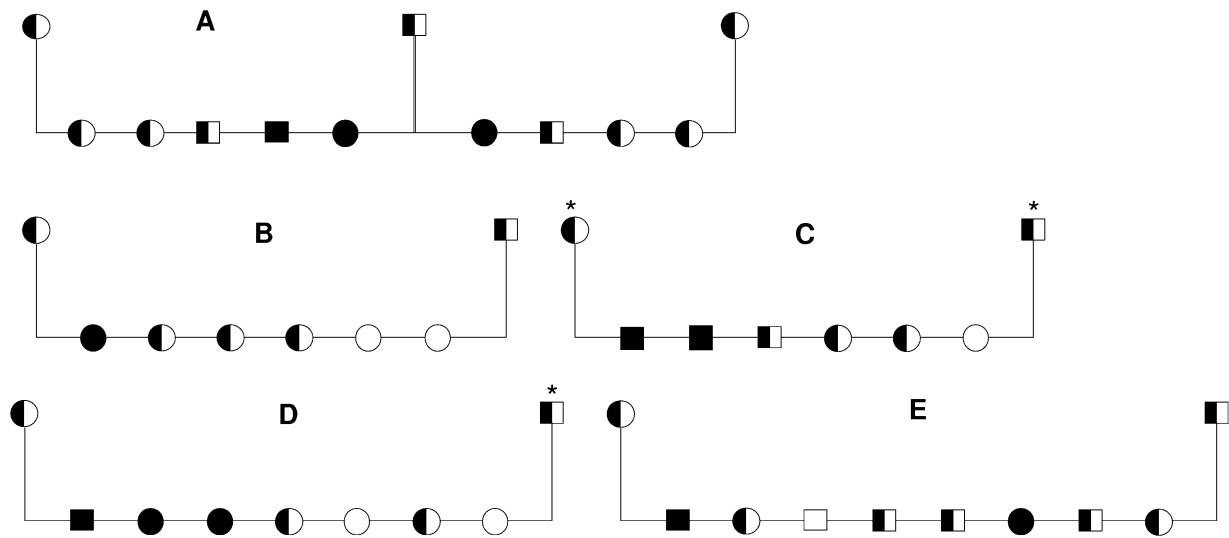
Figure 1. Pedigrees of boxer dogs segregating white coat color. White spotting is inherited as a semidominant trait in boxers, and heterozygous dogs can be recognized by partial white coloring of face and chest. Filled symbols represent extreme white dogs, open symbols brown dogs, and half-filled symbols partially white, flashy dogs. DNA samples of the sires of litters C and D were not available. The DNA of the dam of litter C failed to produce high-quality genotypes. The 3 dogs with missing data are indicated by an asterisk.

SNP Typing

A total of 1536 SNPs were selected from a set of 4500 SNPs that had been typed in various dog breeds by Waltham Center for Pet Nutrition. SNPs were chosen which had the highest minimal allele frequency in the golden retriever breed and were evenly spaced. Data for 1536 selected SNPs were developed by Illumina (San Diego, CA) for the GoldenGate assay of this company (Steemers and Gunderson 2005). This assay is based on ligation of allele-specific oligonucleotides wich are hybridized to genomic DNA of individual dogs and subsequent PCR with universal primers. Sentrix array matrices V.2 of beads with SNP-specific oligonucleotide addresses were used, and the genotypes were determined on a Beadstation 500 GX (Illumina). All SNPs that were not polymorphic in the Boxer families, had a low signal or had signal clusters of the three possible genotypes that were not clearly separated, were excluded. The SNPs that passed the quality control are listed in Table A1.

Linkage Analysis

The genotypes of adjacent SNPs were comprehensively tested for linkage with the white spotting phenotype. The estimation of the genetic distance between adjacent SNPs was based on the physical distance in build 1.1 of the canine genome. The rule of thumb was that 1 Mb of DNA corresponded to a genetic distance of 1 cM. The multipoint linkage calculations were performed with GENEHUNTER software (Kruglyak et al. 1996). GENEHUNTER is not suited to calculate linkage between a semidominant trait and polymorphic markers because it requires an affection status designated by a single digit, which represents affected,

unaffected, or unknown cases. There is no designation for heterozygous animals. To circumvent this problem, liability classes were used to define the underlying genotypes of the coat coloring. Brown boxers were placed in liability class 1, with penetrances 1, 0, and 0 for the genotypes 1 1, 1 2, and 2 2, respectively, in which allele 1 corresponds to brown and allele 2 to white. Flashy dogs were placed in liability class 2 with penetrances 0, 1, and 0 and white boxer dogs were placed in class 3 with penetrances 0, 0, and 1, for the respective genotypes. All dogs were defined as affected, and therefore, the liability class of each dog indicated the underlying genotype.

The maximal obtainable LOD score was calculated by assigning identical genotypes for 2 markers for each boxer dog of which we had DNA in the LINKAGE pedigree file (Lathrop et al. 1984). The first marker represented the coat color, whereas the second marker was an imaginative marker that was fully informative and displayed complete linkage. White boxers were genotyped 2 2 2 2, flashy boxers 1 2 1 2, and brown boxers 1 1 1 1. The flashy sires of which we did not have DNA were given genotypes 1 2 0 0. The 2-point LOD score between these markers in the available boxer families was calculated with MLINK (Lathrop et al. 1984).

Results

Bead arrays, designed to type 1536 SNPs on the canine genome, were used for linkage analysis of white spotting in boxer dogs. Of these, 26 did not pass the data quality check and 445 SNPs were monomorphic in the Boxer pedigrees analyzed. The DNA of 1 dog, the dam of litter C, did not
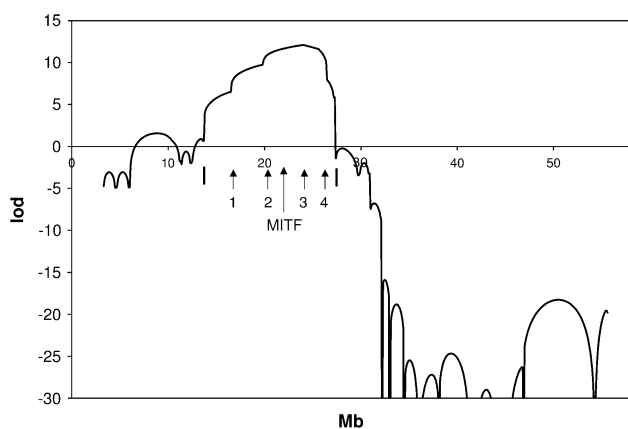
Figure 2. Multipoint linkage analysis of the white spotting locus and canine chromosome 20. The LOD score was calculated with GENEHUNTER software assuming a genetic distance of 1 cM for each Mb of DNA sequence. The critical region is bordered by the SNPs 235J11587 and 230J64832, indicated by bars. The positions of the informative SNPs in this region are indicated by numbered arrows. 1, G630J394570; 2, G630J398251; 3, G630J399481; and 4, 234J39413. The SNPs are from the Broad database CANFAM 1.0 (http://www.broad.mit.edu/mammals/dog/snp/). The position of the candidate gene *MITF* is indicated. Mb: position on BUILD 1.1 of CFA20 in mega base pairs.

Table 1. Haplotypes linked to brown and white alleles in 6 Boxer litters.

Allele	SNP number				N^b
	1^a	2^a	3^a	4^a	
Brown	A	C	G	T	7
	A	C	G	C	1
	G	C	G	C	2
	A	T	A	T	1
White	G	T	G	T	6
	A	T	A	T	3
	G	C	A	T	1
	A	C	A	T	1

[a] SNP1 = G630J394570; SNP2 = G630J398251; SNP3 = G630J399481; and SNP4 = 234J39413.

[b] Number of the haplotype observed in 11 heterozygous parents.

generate results, so both parents of this litter had unknown genotypes because the sire was not available (Figure 1). We calculated the multipoint LOD score for linkage of the semi-dominant white spotting phenotype to the genome. The LOD score had a peak value of 12.1 on chromosome 20 (Figure 2). This high score definitively places the white spotting locus on CFA20. It was the only region with a positive LOD score. The maximal obtainable LOD score with this material is 12.3 at theta 0, indicating that the SNP markers reached a very high level of informativeness. The location of the white spotting gene is bordered by recombination events that disturb the complete linkage between the inheritance of the phenotype and the chromosomal region. Recombinants which define these borders can be distinguished with SNPs 235J11587 and 230J64832 of the SNP database 1.0 of the Broad Institute (http://www.broad.mit.edu/mammals/dog/snp/). The respective positions of these SNPs are at 13.8 and 27.5 Mb of CFA20 of BUILD 2.1 of the DNA sequence map (http://www.ncbi.nlm.nih.gov/mapview/map_search.cgi?taxid=9615). Between these positions, there were only 4 SNPs that were typed and polymorphic in the boxer families. Comparison of haplotypes between brown and white boxers and between families indicate that the spotting locus is in linkage disequilibrium (LD) with the SNP G630J398251 (SNP2 of Figure 1 and Table 1). The LD entity D′ between the white spotting locus and this SNP was 0.8. This value was calculated from the deduced haplotypes in the 11 parents.

Discussion

The discovery of polymorphic microsatellites boosted molecular genetic research because of their abundance and the fact that there are often more than 2 alleles in the population. This property makes it easier to follow the descent of alleles through pedigrees and enhances the informativeness of the marker in linkage studies. Another advantage of microsatellites was the relative ease of genotyping when compared with other types of markers, for instance restriction fragment length polymorphisms. Compensating the disadvantage of having just 2 alleles, the big advantage of SNPs is the possibility for high-throughput genotyping. We questioned whether a set of approximately 1500 SNPs would be adequate for linkage analysis. Our results show that the set, which was selected from approximately 3000 SNPs that were shown to be polymorphic in golden retrievers, is highly efficient in mapping of the white spotting locus in boxer dogs. In general, the informativity of the genome-wide screen varied between 0.75 and 0.95 according to GENEHUNTER. This high level indicates that the SNP screen is at least as powerful as a screen with 300–400 microsatellites but is far more efficient because of the speed of the procedure. Apart from the design of the SNP set, the genotypes were generated in 1 week and the data analysis was done in 2 days.

The white spotting locus is situated on CFA20 between positions 13.8 and 27.5 Mb of BUILD 2.1 of the canine genome. According to the most recent annotation, 39 genes are located in this region. One of these genes, *MITF*, coding for the microphthalmia-associated transcription factor, is known to be involved in coat color. Mutations of this factor cause Waardenburg syndrome type 2 in humans that is characterized by hypopigmentation and deafness (Tassabehji et al. 1994). Mutations in *MITF* are correlated with white spotting in mouse and hamster (Steingrimsson et al. 1994; Hodgkinson et al. 1998). The mode of inheritance of these traits is variable and semidominant inheritance of white spotting caused by *MITF* mutations has been observed (Steingrimsson et al. 1994; Hodgkinson et al. 1998). In the mouse, one *MITF* mutation is known which does not cause deafness or microphthalmia in the homozygous state (Steingrimsson

et al. 2003). One of the typed SNPs contained in the 13.7 Mb critical region on chromosome 20 is in LD with the white spotting locus. This SNP is located at 0.7 Mb distance from *MITF*. It is not unusual to observe LD over large distances in the dog. Due to the population structure of dog breeds, with recent bottlenecks and popular sires, LD extends over much larger regions than in human populations (Sutter et al. 2004). The LD strengthens *MITF* as a candidate gene for white spotting in boxer dogs.

Supplementary Material

Supplementary Appendix can be found at http://www.jhered.oxfordjournals.org/.

Acknowledgments

References

Hodgkinson CA, Nakayama A, Li H, Swenson L-B, Opdecamp K, Asher JH Jr, Arnheiter H, Glaser T. 1998. Mutation at the anophthalmic white locus in Syrian hamsters: haploinsufficiency in the *Mitf* gene mimics human Waardenburg syndrome type 2. Hum Mol Genet 7:703–708.

Kruglyak L, Daly MJ, Reeve-Daly MP, Lander ES 1996. Parametric and nonparametric linkage analysis: a unified multipoint approach. Am J Hum Genet. 58:1347–1363.

Lathrop GM, Lalouel JM, Julier C, Ott J. 1984. Strategies for multilocus linkage analysis in humans. Proc Natl Acad Sci USA. 81:3443–3446.

Miller SA, Dykes DD, Polesky HF. 1988. A simple salting out procedure for extracting DNA from human nucleated cells. Nucleic Acids Res. 16:1215.

Nielen AL, Van der Gaag I, Knol BW, Schukken YH. 1998. Investigation of mortality and pathological changes in a 14-month birth cohort of boxer puppies. Vet Rec. 142:602–606.

Steemers FJ, Gunderson K. 2005. Illumina profile: technology and assays. Pharmacogenomics. 6:777–782.

Steingrimsson E, Arnheiter H, Hallsson JH, Lamoreux ML, Copeland NG, Jenkins NA. 2003. Interallelic complementation at the mouse *mitf* locus. Genetics. 163:267–276.

Steingrimsson E, Moore KJ, Lamoreux ML, Ferre-D'Amare AR, Burley SK, Sanders Zimring DC, Skow LC, Hodgkinson CA, Arnheiter H, Copeland NG, et al. 1994. Molecular basis of mouse microphthalmia (mi) mutations helps explain their developmental and phenotypic consequences. Nat Genet 8:256–263.

Sutter NB, Eberle MA, Parker HG, Pullar BJ, Kirkness EF, Kruglyak L, Ostrander EA. 2004. Extensive and breed-specific linkage disequilibrium in *Canis familiaris*. Genome Res. 14:2388–2396.

Tassabehji M, Newton VE, Read AP. 1994. Waardenburg syndrome type 2 caused by mutations in the human microphthalmia (*MITF*) gene. Nat Genet 8:251–255.

Van Hagen MA, Van der Kolk J, Barendse MA, Imholz S, Leegwater PA, Knol BW, Van Oost BA. 2004. Analysis of the inheritance of white spotting and the evaluation of *KIT* and *EDNRB* as spotting loci in Dutch boxer dogs. J Hered. 95:526–531.