

An SNP within the Angiotensin-Converting Enzyme Distinguishes between Sprint and Distance Performing Alaskan Sled Dogs in a Candidate Gene Analysis

HEATHER J. HUSON, ALEXANDRA M. BYERS, JONATHAN RUNSTADLER, AND ELAINE A. OSTRANDER

From the Cancer Genetics Branch, National Human Genome Research Institute, National Institutes of Health, Bethesda, MD 20892 (Huson, Byers, and Ostrander); and the Institute of Arctic Biology, University of Alaska Fairbanks, Fairbanks, AK (Huson and Runstadler).

Address correspondence to Elaine A. Ostrander at the address above, or e-mail: eostrand@mail.nih.gov.

Abstract

The Alaskan sled dog offers a unique mechanism for studying the genetics of elite athletic performance. They are a group of mixed breed dogs, comprised of multiple common breeds, and a unique breed entity seen only as a part of the sled dog mix. Alaskan sled dogs are divided into 2 primary groups as determined by their racing skills. Distance dogs are capable of running over 1000 miles in 10 days, whereas sprint dogs run much shorter distances, approximately 30 miles, but in faster times, that is, 18–25 mph. Finding the genes that distinguish these 2 types of performers is likely to illuminate genetic contributors to human athletic performance. In this study, we tested for association between polymorphisms in 2 candidate genes; *angiotensin-converting enzyme (ACE)* and *myostatin (MSTN)* and enhanced speed and endurance performance in 174 Alaskan sled dogs. We observed 81 novel genetic variants within the *ACE* gene and 4 within the *MSTN* gene, including a polymorphism within the *ACE* gene that significantly (P value 2.38×10^{-5}) distinguished the sprint versus distance populations.

Key words: Alaskan sled dogs, angiotensin-converting enzyme, myostatin, performance genetics

Alaskan sled dogs are a population of dogs with Northern breed ancestry originally developed as a primary means of human transportation over snow-covered terrain (Rennick 1987; Collins 1991). Although they were an integral part of human habitation in northern climates in the late 1800s to early 1900s, sled dogs eventually experienced a decline in popularity as modern modes of transportation came into common use. However, in the 1930s, there was a resurgence in Alaskan sled dog popularity, with the launching of competitive sled dog racing and the transition of the Alaskan sled dog into a high performance athlete (Rennick 1987; Wendt 1999).

The sport has diverged over the past century into 2 distinct racing styles; sprint or short distance racing and the popular long-distance events (Attla 1972; Swenson 1987; Welch 1989) that cover several hundred miles over multiple days such as the Iditarod and Yukon Quest. Sprint racing is more analogous to short-distance track events with multiple

competition events defined, in this case, by the size of the dog team. The status of elite endurance and sprinting athlete has established the Alaskan sled dog as a new system for exploring the genetics of performance-enhancing polymorphisms (PEP) (Ostrander et al. 2009).

There are no limits regarding the size or appearance of Alaskan sled dogs as they are not an established breed recognized by any registering body such as the American Kennel Club (AKC) in the United States. Dogs are selected and bred solely on their athletic ability and have a unique physique that is capable of a quick, efficient gait, pulling strength, and endurance. Their weight, averaging 55 lbs, and the density of their coat varies, depending on racing style, location, and lineage.

We have shown previously, that as a population, Alaskan sled dogs are a mix of Northern breeds including the Alaskan Malamute and Siberian Husky, together with other AKC-recognized breeds which were introduced to enhance

different racing aspects of performance, such as speed or endurance (Huson et al. 2010). Our previous study showed, further, that the sprint versus distance populations could be genetically distinguished following analysis of a panel of 96 microsatellite-based markers. Although there are likely multiple genetic differences between the sprint and distance populations, in this study, we sought to look for allelic differences in the 2 populations with regard to just 2 candidate genes: *ACE* and *MSTN*.

Variants within the *ACE* gene were among the first PEPs found in humans (Gayagay et al. 1998), and *ACE* variants have been widely studied in the context of elite athletes, particularly high-altitude mountaineers (Montgomery et al. 1998). *ACE* is part of the rennin-angiotensin system and is responsible for degradation of the vasodilator bradykinin, regulation of inflammatory reactions in the lung, respiratory drive, erythropoiesis, tissue oxygenation, and the regulation of skeletal muscle efficiency (Thompson and Binder-Macleod 2006; Zhang et al. 2008). The most common PEP associated with the human *ACE* gene is the *I/D* polymorphism, a 287 bp intronic indel. The *I* allele is associated with lower serum and tissue *ACE* activity and improved performance in sports requiring high levels of endurance, such as marathon running (Rigat et al. 1990; Danser et al. 1995). The *I* allele is believed to facilitate the maximization of oxidative fuel for metabolism (Montgomery et al. 1999). Conversely, the *D* allele is associated with higher serum and tissue activity and superior performance in sports requiring short bursts of power (Thompson and Binder-Macleod 2006). The *D* allele is also associated with greater increase in left ventricular mass, higher VO_2max , and greater strength gain in response to training (Woods and Montgomery 2001).

We also investigated the role of the *MSTN* gene, for which we have previously demonstrated the presence of deletion mutations that are, in turn, associated with increased racing speed in whippet dogs (Mosher et al. 2007). Dogs heterozygous for the mutation exhibited a more muscular phenotype and consistently excelled in competition with faster race times than dogs homozygous for the wild genotype. Dogs carrying 2 copies of the deleterious mutation are heavily muscled (Mosher et al. 2007), and the resultant phenotype is colloquially termed “double muscling” (Girgenrath et al. 2005). Individuals homozygous for various *MSTN* mutations have been reported in mice (Szabo et al. 1998), cattle (Grobert et al. 1997; McPherron and Lee 1997), sheep (Clou et al. 2006), and humans (Schuelke et al. 2004), all of whom share similar phenotypes.

The *ACE* and *MSTN* genes were chosen for investigation in Alaskan sled dogs due to their previous association with endurance or speed enhancement, respectively. The sprint and distance populations of Alaskan sled dogs have diverged over the past decades due to selection of dogs’ speed or endurance capabilities respective to the different performance requirements of the 2 racing styles. Therefore, we investigated whether any corresponding relationships existed between the candidate genes and population differentiation or population performance enhancement.

Materials and Methods

Sample Collection

A total of 174 Alaskan sled dogs were sampled from 8 “high performance” racing kennels. Four sprint kennels were deemed high performers by their points ranking which placed them within the top 25% of sprint or short-distance sled dog as recorded by the International Sled Dog Racing Association during sampling years (2005–2007) (ISDRA 2010). Ninety percent of the sprint racing Alaskan sled dogs were from open (10 or more dogs in a team) and 8-dog racing classes, with the remaining 10% competing in the 6-dog class. All dogs were conditioned at similar increasing mileage and speed throughout the training and racing season, which extends for approximately 7 months. This allowed for consistency in sample collection for the relative speeds and distance. The other 4 kennels were deemed “high performance” distance kennels because they finished in the top 15% of competitors for the Yukon Quest (Quest 2010) or Iditarod (Iditarod 2010) races during the 2 consecutive years (2007–2008) that sample collection was undertaken. In addition to the Alaskan sled dogs, 80 purebred dogs from 8 breeds including the Alaskan Malamute, Siberian Husky, Greyhound, Whippet, Mastiff, Staffordshire Bull Terrier, German Shorthaired Pointer, and English Pointer were included in the *ACE* gene study only. These breeds were selected based on either our previous study which demonstrated that they contributed to the makeup of the modern Alaskan sled dog (Alaskan Malamute, Siberian Husky, English and German Shorthaired Pointer) or because their athletic attributes made them reasonable candidates to consider for contribution (AKC 2010). To maximize diversity, we selected dogs from the same breed that were unrelated at the grandparent level or further removed.

Prior to blood collection, all owners signed an informed consent document, consistent with NHGRI Animal Care and Use Committee rules. Whole blood samples were collected from the cephalic vein in 3–5 ml ethylenediaminetetraacetic acid or ACD tubes. Sled dogs were sampled at their home kennels. Purebred dogs were sampled at AKC-sanctioned events. Samples were stored at 4 °C prior to extraction, and genomic DNA was isolated using standard proteinase K/phenol extraction methods by Health Gene (Toronto, Canada) or RX Bioscience (Rockville, MD). DNA samples were stripped of identifiers, numerically coded, and aliquoted for long-term storage at –70 °C. Detailed pedigrees were collected for each individual sampled and entered into an anonymous database.

Performance Ratings

Sled dogs were rated in terms of both speed and endurance with respect to the distinct styles of sprint and distance. Sprint dogs were reviewed for these criteria by competing at 18–25 mph for 10–30 miles, whereas distance dogs were rated at standards of 8–12 mph over 1000 miles. The performance phenotypes and rating criteria were defined by one of us (H.H.) and reviewed by 5 professional and independent dog mushers.

Speed was defined as an individual dog's ability to run at a specific rate of miles per hour that the team is traveling. A dog was ranked 1 if it was capable of maintaining the speed of the team during each run; 18–25mph for sprint dogs and 8–12mph for distance dogs; or ranked 2 if it was unable to maintain the required speed. Speed requirements were based on the performance levels of the kennels represented in the study.

Endurance was measured by assigning dogs to one of 3 ranks. Dogs who achieved Rank 1 covered the required mileage in good condition, whereas Rank 2 dogs completed the required mileage but struggled to do so. Rank 3 dogs were unable to finish the required mileage. Mileage requirements ranged from 8 to 30 miles for sprint dogs and 991–1150 miles for the distance dogs and were set according to race length requirements.

Sample Selection for *ACE* Gene Analysis

DNA from 20 Alaskan sled dogs (10 sprint and 10 distance) and 80 purebred dogs were initially sequenced using Sanger methodology for 99% of the *ACE* gene. Ten individuals unrelated within 3 generations and belonging to each group of sprint, distance, and the 8 domestic breeds were used to represent the distinct populations. The Alaskan sled dogs ranked elite for their speed and endurance within their respective racing populations. The region sequenced spanned approximately 20 kb on canine chromosome 9 (CFA9) and included 48 overlapping amplicons, averaging 700 bp in length. Amplicons covered all 28 exons, the associated introns, and putative flanking regions. Sequence data were analyzed for polymorphisms as described below and a total of 81 polymorphisms were found. The 10 distance and 10 sprint dogs were then compared at all polymorphisms to test for population-associated differences in the *ACE* gene sequence. In addition, each polymorphism was tested for a putative association with endurance and speed. In separate analysis, we compared allele distribution and frequency for each marker in the set of 80 purebred dogs versus the sprint and distance populations. Sixty-three additional Alaskan sled dogs, 24 elite distance and 39 elite sprint, were selected for subsequent genotyping of markers showing statistically significant differences in allele frequency between sprint and distance dogs.

Sample Selection for the *MSTN* Gene Analysis

The *MSTN* gene was sequenced using Sanger methodology using DNA from 91 sprint Alaskan sled dogs and 2 whippets, the latter of which served as controls who carried a previously reported 2 bp deletion in exon 3 (Mosher et al. 2007). The canine *MSTN* gene spans approximately 5 kb on CFA37. All exons and noncoding regions were sequenced, as well as the flanking regions of the gene except for a 1039 bp GC-rich region in intron one. Sequencing was done using 12 overlapping amplicons averaging 700 bp in length. Dogs sequenced included 46 sprint dogs ranked as elite performers and 37 sprint dogs ranked as poor performers. An additional 8 sled dogs, without performance rating measurements, were also sequenced.

DNA Amplification and Sequencing

PCR amplification for both genes was performed in a 10 μ l volume containing 20 ng genomic DNA, 1.8 μ l GC melt, 1 μ l of 10 \times TaqGold buffer, 0.1 μ l of TaqGold (Applied Biosystems, www.appliedbiosystems.com), 1 μ l of 1 mM dNTPs, 0.3 μ l of 50 mM MgCl₂, 1 μ l of both forward and reverse 3 μ M primers, and 1.8 μ l water. Touchdown PCR was carried out as follows: 95 $^{\circ}$ C for 7 min, followed by 20 cycles of 94 $^{\circ}$ C for 30 s, then decreasing by 0.5 $^{\circ}$ C/cycle starting at 65 $^{\circ}$ C down to 55 $^{\circ}$ C for annealing for 30 s, followed by 20 cycles of 94 $^{\circ}$ C for 30 s, 55 $^{\circ}$ C for 30 s, 72 $^{\circ}$ C for 30 s, and a final extension of 72 $^{\circ}$ C for 7 min.

PCR products were sequenced using Big Dye version 3.1 on an ABI 3730x1 capillary electrophoresis unit (Applied Biosystems). Sequence reads were aligned and analyzed using Sequencher 4.8 software (Gene Codes, <http://www.genecodes.com/>). SNPs and indels were identified by manual comparison to the available canine reference sequences using UCSC Gene Browser (<http://ucsc.genome.edu>). Polymorphisms were numerically labeled as individual markers.

Statistical Analysis for the *ACE* Gene

Association tests using SNPs, indels, or haplotypes and permutation testing for association significance along with linkage disequilibrium (LD) plots were completed using Haploview 4.1 software (<http://www.broadinstitute.org/mpg/haploview>). All analyses used a case/control format and conducted pairwise comparisons between all markers with the exclusion of individuals with greater than 50% missing genotypes. Permutation tests were run for 10 000 cycles and performed separately for the *MSTN* and *ACE* genotypes with cases and controls set as described below.

All 81 *ACE* gene polymorphisms found by comparing DNA sequence from the sprint, distance, and purebred dogs were used in association testing in an attempt to identify markers that distinguished sprint versus distance dogs. Sixty-three additional Alaskan sled dogs (39 elite sprint, 24 elite distance) were genotyped for 3 markers that showed a statistically significant difference ($P \leq 0.005$; permutation $P \leq 0.05$) between the 10 sprint and 10 distance dogs. All the Alaskan sled dog genotype data were combined for a total of 49 elite sprint dogs and 34 elite distance dogs, which were reanalyzed for population association at the statistically significant SNPs.

To investigate whether the 3 *ACE* markers found to differentiate between the sprint and distance dog populations were also associated with endurance or sprinting/power, the same markers were also genotyped in the 80 domestic dogs. It was critical to first develop a scheme that would ensure that we were testing each marker for association to performance attributes rather than breed differences. Thus, we first compared allele frequency and distribution between the 49 sprint and 34 distance dogs and obtained P values. We then independently assigned the 8 pure breeds to 4 sets of 2 pairs, with each pair representing an athletic attribute. For speed, we paired Greyhound/

Whippet and German Shorthaired Pointer/English Pointer. For endurance/strength, we paired Alaskan Malamute/Siberian Husky and Mastiff/Staffordshire Bull Terrier. We then analyzed each of the 3 markers for each group of 20 dogs assigned to a pairing versus the other 60 and obtained *P* values. For example, allele frequencies and distribution were compared between the 20 Greyhound/Whippet (10 each breed), set as cases, in comparison to the 60 controls from the Alaskan Malamute, Siberian Husky, German Shorthaired Pointer, English Pointer, Mastiff, and Staffordshire Bull Terrier. Markers with raw *P* values of <0.005 in both sled dog only analysis and the purebred dog analysis were deemed to be potentially associated with an athletic attribute.

LD was investigated using the entire panel of polymorphic markers found in the *ACE* gene during the initial sequence analysis of the 20 Alaskan sled dogs and the 80 purebred dogs. LD plots were analyzed under the 4 gamete rule for the Alaskan sled dogs and the domestic breed pairs used in the association tests. This allowed us to test whether Alaskan sled dogs had LD patterns similar to any of the pure breeds.

Statistical Analysis for the *MSTN* Gene

To investigate *MSTN* marker association with speed performance, we were able to use a simpler scheme because we had samples from both elite and poor sprint performers. We thus compared genotypes at 4 markers from 46 elite performing sprint sled dogs (cases) versus 37 poor performers (controls) using Haploview 4.1 software. In addition, we genotyped sequence reads from 77 purebred dogs and a Golden Jackal sequenced previously by Mosher et al. (2007) for the 4 *MSTN* SNPs. This allowed us to determine if any variation seen between elite and poor performing sled dogs was also observed in domestic dog breeds. The 16 domestic dog breeds sequenced by Mosher et al. included an average of 4 dogs per breed (Afghan Hound, Akita, Australian Shepherd, Basenji, Beagle, Boxer, Bull Terrier, Flat-Coated Retriever, German Wirehaired Pointer, Greyhound, Ibizan Hound, Italian Greyhound, Mastiff, Pharaoh Hound, Saluki, Siberian Husky, and Whippet) and a Golden Jackal (Mosher et al. 2007).

Results

ACE Gene

We hypothesized that the canine *ACE* gene might play a role in differentiating between the sprint and distance populations of Alaskan sled dogs. We also wished to determine if any SNPs in the *ACE* gene were associated with speed versus endurance performance. Thus, the coding sequences, introns, and putative regulatory flanking regions of the *ACE* gene were initially sequenced in a set of 20 Alaskan sled dogs of which 10 were sprint and 10 were distance dogs. In addition, a set of 80 purebred dogs representing 10 dogs of each of 8 breeds (Alaskan Malamute, Siberian Husky, Staffordshire Bull Terrier, Mastiff, Whippet, Greyhound, German Shorthaired

Pointer, and English Pointer) were also fully sequenced. Purebred dogs were selected because they had either been shown to contribute to the genetic makeup of Alaskan sled dogs or because of their performance attributes. Analysis of the sequence reads from the combined data set of 100 dogs identified 81 polymorphisms. Two alleles, a *G* at Marker 60 and a *C* at Marker 75 were unique to Alaskan sled dogs. Seventeen polymorphisms were present only in the domestic breeds of which a *T* at Marker 28 was specific to the German Shorthaired Pointer breed (Supplementary Table 1).

Of the 81 variants, only 4 were located in exons. One was in an untranslated region of exon 28 (marker 80), 2 resulted in synonymous changes (markers 23 and 82) in exons 8 and 7, respectively, and a fourth, marker 46, caused a nonsynonymous change in exon 17. Both alleles of marker 80 were found in sprint and distance dogs and all the domestic breeds, with both homozygotes and heterozygotes present. Marker 23 caused an *A* to *G* change from the canonical sequence, which did not change the encoded amino acid glycine. Both alleles were found in the sled dog and purebred dog population. Marker 46 caused a *C* to *T* alteration that changed a threonine to methionine at position 14632203 bp. It was observed in both sprint and distance dogs, as well as purebreds. Finally, marker 82 was homozygous for the *T* allele in all Alaskan sled dogs and all purebred dogs tested, as opposed to the reported *C* allele in the boxer reference sequence (NW876331.1) (Supplementary Table 1).

Although a number of variants were found in noncoding regions, 2 (markers 24 and 56) were distinguished by the fact that they contained an allele distinct from that reported in the boxer reference sequence (NW876331.1). In the case of marker 24, we observed dogs that had homozygous and heterozygous deletions of the *C* allele at nucleotide 14627332 bp in the sprint, distance, and purebred populations. Frequencies were 0.400, 0.444, and 0.500, respectively (Supplementary Table 1). In the case of marker 56, the 20 sled dogs showed the *G* allele only present in the sprint dogs at a frequency of 0.35 (Table 1). Sequencing of an additional 63 sled dogs, described below for performance association, found the *G* allele in both the sprint and distance dogs with a frequency of 0.305 and 0.031, respectively (Table 2). Interestingly, of the 8 domestic breeds tested, only the Alaskan Malamute and Siberian Husky breeds carried the *G* allele with a frequency of 0.265 with the 20 dogs combined.

We next compared LD between the Alaskan sled dogs and the purebred dogs. Sprint and distance dogs were considered separately and LD plots produced (Figure 1A,B). We used the same groupings of the purebred dogs that we had developed for analysis of the *ACE* gene. Thus, closely related breeds were paired: Alaskan Malamute/Siberian Husky, Whippet/Greyhound, Mastiff/Staffordshire Bull Terrier, and German Shorthaired Pointer/English Pointer and LD plots were produced for those 4 combinations (Figure 1C–F). The LD plots demonstrate, first, that the sprint and distance dogs contain a substantial amount of LD in the *ACE* gene but that it differs between the 2 populations. There is considerably less LD in each of the

Table 1 Associated SNPs across the *ACE* gene comparing distance and sprint sled dogs

Marker	Chromosome	Location (bp)	Gene position	Allele	Distance ^a (case) MAF	Sprint ^b (control) MAF	P value	Permutation (10 000 cycles) P value
42	9	14631037	Intron 15	A:G	0	0.35	0.0036	0.0156
56	9	14635693	Intron 19	A:G	0	0.35	0.0036	0.0156
74	9	14637900	Intron 21	G:A	0	0.389	0.0051	0.0399

^a Ten elite distance dogs were set as cases.

^b Ten elite sprint dogs were set as controls.

4 purebred pairs. None of the 4 purebred pairs is reminiscent of the patterns observed in the Alaskan sled dogs, highlighting, again, the uniqueness of the breed.

We next wanted to determine if any of the markers could be used to distinguish between sprint and distance populations of the Alaskan sled dogs. Using Haploview (<http://www.broadinstitute.org/mpg/haploview>), both SNP and association tests were used to evaluate all 81 markers for association with either the sprint or distance population. In the initial analysis, 10 sprint and 10 distance dogs were compared at all markers. We observed that 3 markers, 42, 56, and 74, located at 14631037 bp, 14635693 bp, 14637900 bp, respectively, demonstrated *P* values lower than ≤ 0.005 (permutation *P* values ≤ 0.05) with variances in allele frequencies separating elite sprint and distance sled dogs (Table 1).

Also, we were interested in how the sprint and distance populations related to the purebred dog populations with regard to the 3 markers mentioned above. We hypothesize that because all the markers found were in the *ACE* gene and they easily distinguished the sprint and distance populations, they might highlight specific performance patterns associated with the purebred dogs. We first compared all sets of 2 purebred breeds against all other purebreds in order to determine if there were markers within the *ACE* gene that distinguished, specifically, purebred breeds associated with speed (Greyhound/Whippet) versus endurance (Malamute/Husky). Marker 56, which generates a single base pair change from A/G at 14635693 bp within intron 19 (A allele in reference sequence NW876331.1), was the only marker to have a significant *P* value (7.57×10^{-8}), for any purebred pair, in this case when we compared the Alaskan Malamute/Siberian Husky pair to all other purebreds (Table 2).

This was a particularly interesting marker as the previous analysis demonstrated that marker 56 was one of 3 which was useful for distinguishing sprint versus distance dogs

(*P* = 0.0036, Table 1). That analysis, however, only involved 10 dogs of each type. We expanded the analysis to include 63 additional for a total of 83 Alaskan sled dogs (49 sprint and 34 distance) and obtained a *P* value of 2.38×10^{-5} (Table 2).

We did, however, observe significant difference in minor allele frequencies (MAFs) between the Alaskan Malamute/Siberian Husky group compared with the Alaskan sled dog distance population. Specifically, we observed an MAF (G allele) of 0.265 in the Malamute/Husky group, whereas all other purebred dogs were homozygotes for the A allele. In addition, the G allele was found at a frequency of 0.305 in the sprint sled dogs and 0.031 for the distance dogs. Alleles were in Hardy–Weinberg equilibrium with respect to the sprint and distance populations. Had the marker been functional with respect to the performance aspects of the distance versus sprint populations, we would have expected that the distance dogs would carry the G allele much more frequently, more analogous to the Malamute/Husky group, from whom they presumably “inherit” a significant portion of their endurance, and a much lower frequency of the G allele, with respect to the sprint dogs. The fact that we observe the opposite suggests that while marker 56 is useful for distinguishing populations with the Alaskan sled dog sprint and distance groups as a population level, the marker is not a hallmark of any putative contribution the *ACE* gene may be making to performance.

MSTN Gene

We sequenced the coding region, introns, and putative regulatory regions flanking the *MSTN* gene in 91 sprint dogs. No obvious deleterious mutations were found including the 2 bp deletion at nucleotide 939–940, which we have previously reported in racing whippets (Mosher et al. 2007). Four polymorphisms were found in noncoding regions of the *MSTN* gene during the analysis of the sled dog sequence reads. One polymorphism in intron 2; a 4 bp

Table 2 *ACE* gene marker 56 shows *P* value scores ≤ 0.001 in separate association tests of Alaskan sled dogs and Purebred breeds

Group	Marker	Chromosome	Location (bp)	Gene position	Allele	Case MAF	Control MAF	<i>P</i> value
Sled dog ^a	56	9	14635693	Intron 19	A:G	0.031	0.305	2.38×10^{-05}
Purebreed ^b	56	9	14635693	Intron 19	A:G	0.265	0	7.57×10^{-08}

^a SNP association test comparing 34 elite distance dogs (cases) and 49 elite sprint dogs (controls).

^b SNP association test comparing 10 Alaskan Malamutes and 10 Siberian Huskies (20 cases) and 60 purebred dogs (10 dogs from each of the following breeds; Whippet, Greyhound, Mastiff, Staffordshire Bull Terrier, English Pointer, and German Shorthaired Pointer).

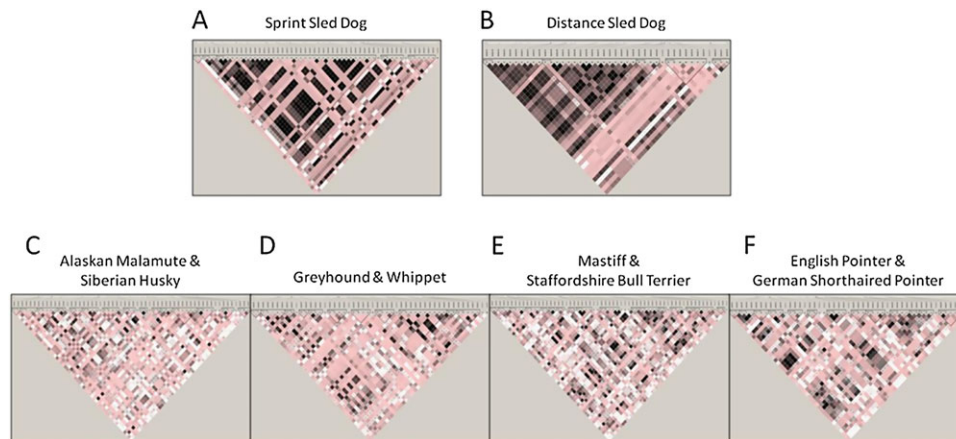


Figure 1. *ACE* gene LD plots were analyzed separately for sprint and distance dogs and closely related breed pairs to determine if sled dogs have similar LD patterns to the domestic breeds. LD was determined for each group based on the 81 polymorphisms found from manual sequence screening of the 20 Alaskan sled dogs (10 sprint, 10 distance) and 80 purebred dogs. The LD plots were analyzed with the 4 gamete rule and depicted using the alternate D'/LOD color scheme in Haploview 4.1. Low LOD and high LOD (logarithm (base 10) of odds) scores with low D' scores are in white, low LOD with high D' scores are in shades of pink, and high LOD with high D' scores are in black. (A) 10 sprint dogs; (B) 10 distance dogs; (C) 10 Alaskan Malamutes and 10 Siberian Huskies; (D) 10 Greyhounds and 10 Whippets (E) 10 Mastiffs and 10 Staffordshire Bull Terriers (F) 10 English Pointers and 10 German shorthaired Pointers. The Alaskan sled dogs demonstrate a substantial amount of LD in the *ACE* gene that differs between the sprint and distance populations. The purebred pairs show less LD in the *ACE* gene and no similarity in pattern to the Alaskan sled dog populations.

indel at nucleotide 3731257 (marker 01) was observed in 24 dogs. Also, a *T/A* SNP, downstream of the last coding exon at nucleotide 3720985 bp (marker 02), was found in 23 dogs. Two other polymorphisms were found upstream of the 5' end of the gene, an *A* to *G* change at 3736327 bp (marker 03) which was observed in 11 dogs and an *A* insertion at 3739468 bp (marker 04) which was observed in 73 dogs (Table 3).

The 4 SNPs were then analyzed in a panel of 77 purebred dogs and a wild canid previously genotyped by Mosher et al. to determine if the allele was truly unique to the sprint dog population. There was an average of 4 dogs representing each of 16 breeds (Afghan Hound, Akita, Australian Shepherd, Basenji, Beagle, Boxer, Bull Terrier, Flat-Coated Retriever, German Wirehaired Pointer, Greyhound, Ibizan Hound, Italian Greyhound, Mastiff, Pharaoh Hound, Saluki, Siberian Husky, and Whippet) and a Golden Jackal (Mosher et al. 2007). We observed all polymorphisms in the domestic breeds, thus demonstrating that they are not unique to the sprint sled dogs. Marker 03 was deemed poor quality due to less than 50% genotype call rate in both the sled dogs and the purebred dogs. MAF showed a difference of ≤ 0.095 between the sled dogs and the purebred panel for the 3 remaining markers (Table 3).

SNP association tests were performed to identify whether any of the 4 markers were in association to sprint sled dog performance. Haploview 4.1 software found that neither the *MSTN* gene nor surrounding markers had a significant *P* value (raw $P \leq 0.005$) that would have suggested an association with either sled dogs performing poorly in speed or being ranked elite in their speed performance.

Discussion

The Alaskan sled dog provides researchers with a unique system in which to study the genetics of athletic performance. In this study, we focus on understanding specific genes that are candidates for distinguishing the sprint and distance populations of Alaskan sled dogs along with being potentially influential in athletic performance. We targeted 2 genes; *angiotensin-converting enzyme (ACE)* and *myostatin (MSTN)* and tested for association between gene polymorphisms and both sled dog population differentiation and athletic attributes such as endurance, speed, and power. Novel genetic variants were found within both genes. Four *MSTN* gene polymorphisms were found in the screening of 91 sprint dogs and confirmed in a panel of 77 purebred dogs from 16 breeds; Afghan Hound, Akita, Australian Shepherd, Basenji, Beagle, Boxer, Bull Terrier, Flat-Coated Retriever, German Wirehaired Pointer, Greyhound, Ibizan Hound, Italian Greyhound, Mastiff, Pharaoh Hound, Saluki, Siberian Husky, and Whippet and a Golden Jackal. However, none of the *MSTN* variants showed association with any traits of interest.

Eighty-one polymorphic markers were identified in the *ACE* gene through sequence analysis of 100 dogs, 10 each from the sprint sled dogs, distance sled dogs, Alaskan Malamute, Siberian Husky, Whippet, Greyhound, German Shorthaired Pointer, English Pointer, Mastiff, and Staffordshire Bull Terrier. Variation in *ACE* gene markers 60 and 75 was only found within the Alaskan sled dogs.

We previously demonstrated that a panel of 96 genome wide microsatellite-based markers successfully differentiates

Table 3 MAFs of 3 *MSTN* gene polymorphisms found in sprint Alaskan sled dogs, 16 domestic breeds, and a Golden Jackal

Gene	Group	Marker	Chromosome	Location (bp)	Gene position	Allele	MAF
MSTN	Sprint sled dog ^a	Marker 01	37	3731257	Intron 2	C:A	0.281
MSTN	Sprint sled dog ^a	Marker 02	37	3720985	Downstream 3'	T:A	0.125
MSTN	Sprint sled dog ^a	Marker 04	37	3739468	Upstream 5'	A:C	0.151
MSTN	Pure breeds ^b	Marker 01	37	3731257	Intron 2	C:A	0.297
MSTN	Pure breeds ^b	Marker 02	37	3720985	Downstream 3'	T:A	0.22
MSTN	Pure breeds ^b	Marker 04	37	3739468	Upstream 5'	A:C	0.161

^a Ninety-one sprint Alaskan sled dogs.

^b Seventy-seven total purebred dogs averaging 4 dogs per 16 breeds (Afghan Hound, Akita, Australian Shepherd, Basenji, Beagle, Boxer, Bull Terrier, Flat-Coated Retriever, German Wirehaired Pointer, Greyhound, Ibizan Hound, Italian Greyhound, Mastiff, Pharaoh Hound, Saluki, Siberian Husky, and Whippet) and a Golden Jackal.

Alaskan sled dogs into 2 populations based on their racing style of sprint (30 miles at 18–25 mph) or distance (1000 miles at 8–12 mph) (Huson et al. 2010). Here, we successfully identified 3 individual *ACE* gene markers (markers 42, 56, and 74) with a raw *P* value of ≤ 0.005 (permutation *P* value < 0.05) that differentiated between the sprint and distance sled dog populations.

We hypothesized that these 3 *ACE* gene markers had the potential to be associated with athletic attributes such as endurance or speed/power exhibited, respectively, by the distance and sprint populations. Separate analysis of allele frequency and distribution within the pure breed pairings established marker 56 as having significant G allele association (*P* value of 7.57×10^{-8}) between the Alaskan Malamute/Siberian Husky pairing and the other 6 domestic breeds. An expanded analysis of the Alaskan sled dogs to include 49 sprint and 34 distance dogs improved the *P* value from 0.0036 (10 sprint and 10 distance) to 2.38×10^{-5} for marker 56 (Table 2). However, we observed a significant difference in the MAF (G allele) between the Alaskan Malamute/Siberian Husky group at 0.265 compared with the Alaskan sled dog distance population at 0.031. In contrast, the MAF (G allele) of the sprint dogs at 0.305 was more analogous to the Alaskan Malamute/Siberian Husky breed pairing. In a previous study by our group, the distance dogs showed a 25% higher degree of Alaskan Malamute and Siberian Husky in their total breed composition than the sprint dogs. We also found an 11% increase in these 2 breeds when comparing high and low endurance performance distance sled dogs. We therefore expected the distance sled dogs to be similar in allele frequency to the Alaskan Malamute and Siberian Husky based on their common athletic attribute of endurance and previous identification of these 2 purebred breeds being higher component breeds within the distance sled dog population (Huson et al. 2010). The fact that we observe the opposite suggests that while marker 56 is useful for distinguishing between sprint and distance populations of Alaskan sled dogs, the marker is not a hallmark of any putative contribution the *ACE* gene may be making to sprint versus distance performance.

One explanation for these results may relate to the founder populations of the Northern breeds that created the sled dogs over a century ago are genotypically different,

especially for performance genes, from the registered AKC Alaskan Malamutes and Siberian Huskies we sampled at conformation events. Dogs shown in conformation events are selected based on AKC standards for body structure, not performance abilities (Club 1997). Another explanation is that the G allele may have arisen separately in the sprint dog population and hence is not in the domestic dogs or it may have come from a lineage that was not investigated.

The A allele of marker 56 appears to be near fixation in the distance sled dogs, whereas the G allele has been selected for in the sprint dogs (Table 2). This suggests that the G allele may be under selection in the sprint dogs for a trait other than endurance that was not in evidence when we compared domestic breeds such as the greyhounds and whippets, who share the attribute of speed. The association of marker 56 in the sled dogs may also reflect that this SNP is in LD with another other variants that are more biologically relevant. It would be interesting to sample and genotype working Alaskan Malamutes and Siberian Huskies to determine if they demonstrate selection for marker 56 in the working dogs, as opposed to the dogs bred for conformation.

The inclusion of purebred dogs in the analysis was important for several reasons. First, it was necessary for carrying out tests of performance association. Second, and more importantly, our knowledgebase regarding the composition of sprint and distance dogs is built on our previous clustering analyses, which identified distinct contributions of multiple pure breeds, uniquely, to both sprint and distance populations (Huson et al. 2010). There are limitations to the use of domestic breeds as a control for determining whether markers were associated with performance attributes selected for in sprint and distance sled dogs as opposed to being a population identifier. Although performance is a genetically complex trait, this approach required a specific marker to demonstrate association to both the sled dog population attributed with speed (sprint) or endurance (distance) along with the respective domestic breeds displaying the same athletic attribute, therefore assuming the same genetic mechanism effecting performance in the different populations. However, the genetic components integral to the speed exhibited by greyhounds and whippets may be significantly different from those selected for in sprint sled dogs. Although we hypothesized that the genetic

basis for endurance is more likely to be similar between distance sled dogs and their component breeds of Alaskan Malamute and Siberian Husky, those similarities may not lie within the *ACE* gene. Nevertheless, understanding allele frequencies of critical SNPs in the context of both the sled dog and the contributing purebred dog populations was thus important. Comparison of LD across the *ACE* gene in the Alaskan sled dogs and the 4 pairs of domestic breeds analyzed corroborated our earlier findings, highlighting the uniqueness of the Alaskan sled dog breed. A substantial amount of LD in the *ACE* gene was found in both the sprint and distance dogs, but with noticeable differences in pattern between the 2 populations. By comparison, purebred pairs showed considerably less LD and no pattern similarity to either Alaskan sled dog population (Figure 1).

Identifying genes and their subsequent markers that distinguish between elite endurance performing distance sled dogs and elite sprinting sled dogs has the potential to illuminate contributors in the complex genetic arena of human athletic performance. Although none of the variants identified in the *ACE* or *MSTN* genes were significantly associated with any behavioral traits, the finding of markers within the *ACE* gene which distinguishes these 2 populations of Alaskan sled dogs, and the developed understanding of how the populations relate to one another as well as various purebred breeds, sets the stage for genome-wide association studies aimed at finding performance-associated genes.

Supplementary Material

Supplementary material can be found at <http://www.jhered.oxfordjournals.org/>.

Funding

IDEA Networks of Biomedical Research Excellence (INBRE). INBRE grant 5P20RR016466 from NCRR (J.R.) and the Intramural program of the National Human Genome Research Institute for their support (H.H., H.P. E.A.O.).

Acknowledgments

We thank Kenneth and Lori Chezik, Dr Dawn Brown, Greg Sellentin, and Deborah McGrath for sharing their expertise and review of the performance phenotype rating criteria. We appreciate collection assistance from Danielle Dillon, Keiko Herrick, Lori Gildehaus, Ian Herriott, and the kennel owners and handlers, and we gratefully acknowledge the sled dog owners who have generously contributed information and DNA to our studies.

References

American Kennel Club. 2010. Conformation events. Raliegh (NC): American Kennel Club Publishing. [cited 2010 Aug]. Available from: <http://www.akc.org/events/conformation/>.
ADMA, Alaska Dog Musher's Association. 2010. Fairbanks (AK): Alaska Dog Musher's Association Publishing. [cited 2010 Aug]. Available from: <http://www.sleddog.org/2010/03/21/open-north-american-results-day-3-final/>.

Atla GBL. 1972. Everything I know about training and racing sled dogs. Hayward (CA): Typart Publishing Service.
Clop A, Marcq F, Takeda H, Pirottin D, Tordoir X, Bibe B, Bouix J, Caiment F, Elsen JM, Eychenne F, et al. 2006. A mutation creating a potential illegitimate microRNA target site in the myostatin gene affects muscularity in sheep. *Nat Genet.* 38:813–818.
Club AK. 1997. The Complete dog book. Official publication of the American Kennel Club. New York: Howell Book House.
Collins MJC. 1991. Dog driver: a guide for the serious musher. Crawford (CO): Alpine Publications.
Danser AH, Schalekamp MA, Bax WA, van den Brink AM, Saxena PR, Riegger GA, Schunkert H. 1995. Angiotensin-converting enzyme in the human heart. Effect of the deletion/insertion polymorphism. *Circulation.* 92:1387–1388.
Gayagay G, Yu B, Hambly B, Boston T, Hahn A, Celermajer DS, Trent RJ. 1998. Elite endurance athletes and the ACE I allele—the role of genes in athletic performance. *Hum Genet.* 103:48–50.
Girgenrath S, Song K, Whittemore LA. 2005. Loss of myostatin expression alters fiber-type distribution and expression of myosin heavy chain isoforms in slow- and fast-type skeletal muscle. *Muscle Nerve.* 31:34–40.
Grobert L, Martin LJ, Poncelet D, Pirottin D, Brouwers B, Riquet J, Schoeberlein A, Dunner S, Menissier F, Massabanda J, et al. 1997. A deletion in the bovine myostatin gene causes the double-muscling phenotype in cattle. *Nat Genet.* 17:71–74.
Huson HJ, Parker HG, Runstadler J, Ostrander EA. 2010. A genetic dissection of breed composition and performance enhancement in the Alaskan sled dog. *BMC Genet.* 11:71.
Iditarod. 2010. The official site of the Iditarod. Wasilla (AK): Iditarod Trail Committee, Inc. [cited 2010 Aug]. Available from: <http://www.iditarod.com/>.
ISDRA, 2010. ISDRA; Merrifield (MN): International Sled Dog Racing Association. International Sled Dog Association Publishing. [cited 2010 Aug]. Available from: <http://www.isdra.org/>.
McPherron AC, Lee SJ. 1997. Double muscling in cattle due to mutations in the myostatin gene. *Proc Natl Acad Sci U S A.* 94:12457–12461.
Montgomery H, Clarkson P, Barnard M, Bell J, Brynes A, Dollery C, Hajnal J, Hemingway H, Mercer D, Jarman P, et al. 1999. Angiotensin-converting-enzyme gene insertion/deletion polymorphism and response to physical training. *Lancet.* 353:541–545.
Montgomery HE, Marshall R, Hemingway H, Myerson S, Clarkson P, Dollery C, Hayward M, Holliman DE, Jubb M, World M, et al. 1998. Human gene for physical performance. *Nature.* 393:221–222.
Mosher DS, Quignon P, Bustamante CD, Sutter NB, Mellersh CS, Parker HG, Ostrander EA. 2007. A mutation in the myostatin gene increases muscle mass and enhances racing performance in heterozygote dogs. *PLoS Genet.* 3:e79.
Ostrander EA, Huson HJ, Ostrander GK. 2009. Genetics of athletic performance. *Annu Rev Genomics Hum Genet.* 10:407–429.
Quest Y. 2010. Yukon Quest; Sled Dog Race; Alaska, Yukon. Fairbanks (AK): Zee Creative Publishing. [cited 2010 Aug]. Available from: <http://www.yukonquest.com/>.
Rennick P. 1987. Dogs of the North. Anchorage (AK): Alaska Geographic Society p. 117.
Rigat B, Hubert C, Alhenc-Gelas F, Cambien F, Corvol P, Soubrier F. 1990. An insertion/deletion polymorphism in the angiotensin I-converting enzyme gene accounting for half the variance of serum enzyme levels. *J Clin Invest.* 86:1343–1346.

- Schuelke M, Wagner KR, Stolz LE, Hubner C, Riebel T, Komen W, Braun T, Tobin JF, Lee SJ. 2004. Myostatin mutation associated with gross muscle hypertrophy in a child. *N Engl J Med.* 350:2682–2688.
- Swenson R. 1987. The secrets of long distance training and racing. Wasilla (AK): L&B Color Printing.
- Szabo G, Dallmann G, Muller G, Patthy L, Soller M, Varga L. 1998. A deletion in the myostatin gene causes the compact (Cmpt) hypermuscular mutation in mice. *Mamm Genome.* 9:671–672.
- Thompson WR, Binder-Macleod SA. 2006. Association of genetic factors with selected measures of physical performance. *Phys Ther.* 86:585–591.
- Welch J. 1989. The speed mushing manual: how to train racing sled dogs. Eagle River (AK): Sirius Publishing.
- Wendt R. 1999. Alaska dog mushing guide: facts and legends. Fairbanks (AK): Goldstream Publications.
- Woods DR, Montgomery HE. 2001. Angiotensin-converting enzyme and genetics at high altitude. *High Alt Med Biol.* 2:201–210.
- Zhang X, Wang C, Dai H, Lin Y, Zhang J. 2008. Association between angiotensin-converting enzyme gene polymorphisms and exercise performance in patients with COPD. *Respirology.* 13:683–688.

**Received October 19, 2010; Revised March 10, 2011;
Accepted March 10, 2011**

Corresponding Editor: John Fyfe