Science
AAAS

Molecular and Evolutionary History of Melanism in North American Gray Wolves

Tovi M. Anderson, *et al.*

Science **323**, 1339 (2009);
DOI: 10.1126/science.1165448

answer, then she or he could short-sell this metal to profit from the expected reduction of the platinum price. Of course, the prices of catalysts are affected not only by advances in fuel cell technology but also by usage for other purposes. Still, financial engineers nowadays are capable of tailoring positions to isolate one specific source of risk. The fuel cell example illustrates how inventors at present may actually be able to profit from a discovery twice: through the monopoly granted by the patent and through positions in markets for catalysts. We propose to eliminate the former, leaving inventors only with the incentives provided by the latter. Our experiments show that this is sufficient to promote intellectual discovery.

In our experiments, subjects were paid as a function of the very best solution. We consider this to be an idealized situation afforded by the experimental setting, which allows for a clear interpretation of performance numbers. In practice, inventors will be compensated for second-best solutions as long as the best one has not been attained yet. This is precisely what happened in 19th century Cornwall: Engineers never really found the best way of building steam engines (we know this because many improvements were made afterward), but they were rewarded for providing better solutions.

The traditional patent system helps people with an idea even if they have no resources, because those who have the resources (usually venture capitalists) provide inventors with cash in return for a share in the intellectual property rights. In our markets system, resources would be generated in a similar way. The people with the ideas but no money could approach investors (such as fund managers) and inform them. These investors could then take positions to exploit expected changes in valuations from adoption of the new technology. They should be eager to pay for the information, and this payment will provide the inventors with the necessary cash for development.

We do not claim that our markets system will work under all circumstances. We envisage that it could replace the patent system whenever a technology builds on goods and services with economic rents, which means that their cost of provision is below market value. Such rents obtain for a variety of reasons. One is limited supply (volcanic ash in the concrete example, platinum in the fuel-cell case, artemisinin in the case of medication against drug-resistant malaria, or the claims to the items in the knapsacks in our experiments). Other important reasons are first-mover advantage (this seems to have been the case with steam engine technology in the mines of Cornwall in the mid-19th century) and lead in the learning curve [studied extensively in the economics literature; see (*7*)].

The success of our markets system relies on willingness to trade; without trading, those who make progress toward finding the optimal solution cannot exploit their acquired knowledge. In our setting, participants never know whether they have the optimal solution (because they would need much more time to check all possible solutions). Thus, there is always the possibility that one is trading with a counterparty who knows better. Why, then, would participants trade? We conjecture that they trade because they tend to be too confident that they are closer to solving the problem than others. Overconfidence is indeed an important human trait, best illustrated by the fact that more than 50% of people usually think they are better than the median (*12*). Other, non-pecuniary incentives for trade (such as a taste for being "right") may play a role. Further research is needed to identify the origin of the success of the markets system in promoting intellectual discovery.

Our proposal relies on anonymous, two-way markets. Until recently, setting up new markets required time. Modern technology, however, has enabled quick design, ready deployment, and low-cost management of markets. With the open-source software we developed for our experiments, jMarkets (*13*), setting up and launching of (online) markets can be done in a matter of hours.

Our experimental findings suggest that the patent system is not a universally superior way to incentivize intellectual discovery. We propose a markets-based system that we found to work better. Its main features are that the compensation for inventions is shared and that, because discovery remains in the public domain, it avoids both distortion in the provision of newly invented products and stifling of future discovery.

References and Notes

1. M. A. Heller, R. S. Eisenberg, *Science* **280**, 698 (1998).
2. A. B. Jaffe, *Res. Policy* **29**, 531 (2000).
3. R. J. Gilbert, D. M. G. Newbery, *Am. Econ. Rev.* **72**, 514 (1982).
4. M. Boldrin, D. Levine, *Am. Econ. Rev.* **92**, 209 (2002).
5. R. Siddall, *Geol. Soc. London Spec. Publ.* **171**, 339 (2000).
6. A. Nuvolari, *Camb. J. Econ.* **28**, 347 (2004).
7. K. Arrow, *Rev. Econ. Stud.* **29**, 155 (1962).
8. N. Gallini, S. Scotchmer, in *Innovation Policy and the Economy*, A. Jaffe, J. Lerner, S. Stern, Eds. (MIT Press, Cambridge, MA, 2002), vol. 2, pp. 51–78.
9. H. Kellerer, U. Pferschy, D. Pisinger, *Knapsack Problems* (Springer-Verlag, Heidelberg, 2004).
10. S. Sahni, *J. Assoc. Comput. Mach.* **22**, 115 (1975).
11. Materials and methods are available as supporting material on *Science* Online.
12. B. Biais, D. Hilton, K. Mazurier, S. Pouget, *Rev. Econ. Stud.* **72**, 287 (2005).
13. http://jmarkets.ssel.caltech.edu.

14. This study was partly funded by the U.S. NSF (grants SES-0616431 and SES-0317715) and the Swiss Finance Institute.

Supporting Online Material

www.sciencemag.org/cgi/content/full/323/5919/1335/DC1
Materials and Methods
SOM Text
Figs. S1 to S11
Tables S1 to S5
References

3 April 2008; accepted 5 January 2009
10.1126/science.1158624

Molecular and Evolutionary History of Melanism in North American Gray Wolves

Tovi M. Anderson,[1] Bridgett M. vonHoldt,[2] Sophie I. Candille,[1] Marco Musiani,[3] Claudia Greco,[4] Daniel R. Stahler,[2,5] Douglas W. Smith,[5] Badri Padhukasahasram,[6] Ettore Randi,[4] Jennifer A. Leonard,[7] Carlos D. Bustamante,[6] Elaine A. Ostrander,[8] Hua Tang,[1] Robert K. Wayne,[2] Gregory S. Barsh[1]*

Morphological diversity within closely related species is an essential aspect of evolution and adaptation. Mutations in the *Melanocortin 1 receptor* (*Mc1r*) gene contribute to pigmentary diversity in natural populations of fish, birds, and many mammals. However, melanism in the gray wolf, *Canis lupus*, is caused by a different melanocortin pathway component, the *K* locus, that encodes a beta-defensin protein that acts as an alternative ligand for Mc1r. We show that the melanistic *K* locus mutation in North American wolves derives from past hybridization with domestic dogs, has risen to high frequency in forested habitats, and exhibits a molecular signature of positive selection. The same mutation also causes melanism in the coyote, *Canis latrans*, and in Italian gray wolves, and hence our results demonstrate how traits selected in domesticated species can influence the morphological diversity of their wild relatives.

The correspondence between coat color and habitat is often attributed to natural selection, but rarely is supporting evidence provided at the molecular level. In North American gray wolves, coat color frequencies differ between wolves of forested and open habitats throughout western North America (*1*), including Denali National Park (*2*) and the Kenai Peninsula in Alaska (*3*), and much of the Canadian Arctic (*4*, *5*). These differences are especially dramatic between wolves of the high tundra that are migratory and follow barren-ground caribou to their breeding areas, and wolves that are year-round residents in the neighboring boreal forest and hunt nonmigratory prey.

Dark-colored wolves are extremely rare in the tundra but increase in frequency along a southwest cline toward forested areas (Fig. 1A). The potential selective value of dark versus light coat

[1]Departments of Genetics and Pediatrics, Stanford University, Stanford, CA 94305, USA. [2]Department of Ecology and Evolutionary Biology, University of California, Los Angeles, CA 91302, USA. [3]Faculty of Environmental Design, University of Calgary, Calgary, AB T2N 1N4, Canada. [4]Istituto Nazionale per la Fauna Selvatica, 40064 Ozzano Emilia (BO), Italy. [5]Yellowstone Center for Resources, National Park Service, Yellowstone National Park, WY 82190, USA. [6]Department of Biological Statistics and Computational Biology, Cornell University, Ithaca, NY 14853, USA. [7]Department of Evolutionary Biology, Uppsala University, 75236 Uppsala, Sweden. [8]National Human Genome Research Institute, Bethesda, MD 20892, USA.

*To whom correspondence should be addressed. E-mail: gbarsh@stanford.edu

Table 1. Distribution of *CBD103* alleles in wolves and coyotes. N/A, not applicable.

Animal and location		Phenotype†		
		White	Gray	Black
Forest wolves*	Total no.	12	2	7
	No. carrying K^B	0	1	7
Tundra/taiga wolves*	Total no.	10	8	2
	No. carrying K^B	0	5	2
Yellowstone wolves	Total no.	0	120	104
	No. carrying K^B	N/A	0	102
Coyotes‡	Total no.	0	61	6
	No. carrying K^B	N/A	0	6

*Forest and tundra/taiga wolves are from the Canadian Arctic (Fig. 1A). The overall frequency of dark (gray or black) wolves is 62 and 7% in the forest and tundra/taiga, respectively (*4*), and the genotype distributions shown do not represent population-based frequencies. All forest and tundra/taiga wolves carrying K^B were K^B/k^y; in the Yellowstone population, 10 were K^B/K^B and 92 were K^B/k^y. †This categorical designation of phenotypes, as defined at sample collection, does not fully capture the spectrum of normal coat color variation as indicated in Fig. 1B. ‡Gray coyotes surveyed were from Nebraska (30) or West Virginia (30); black coyotes were from Minnesota (2) or West Virginia (4).

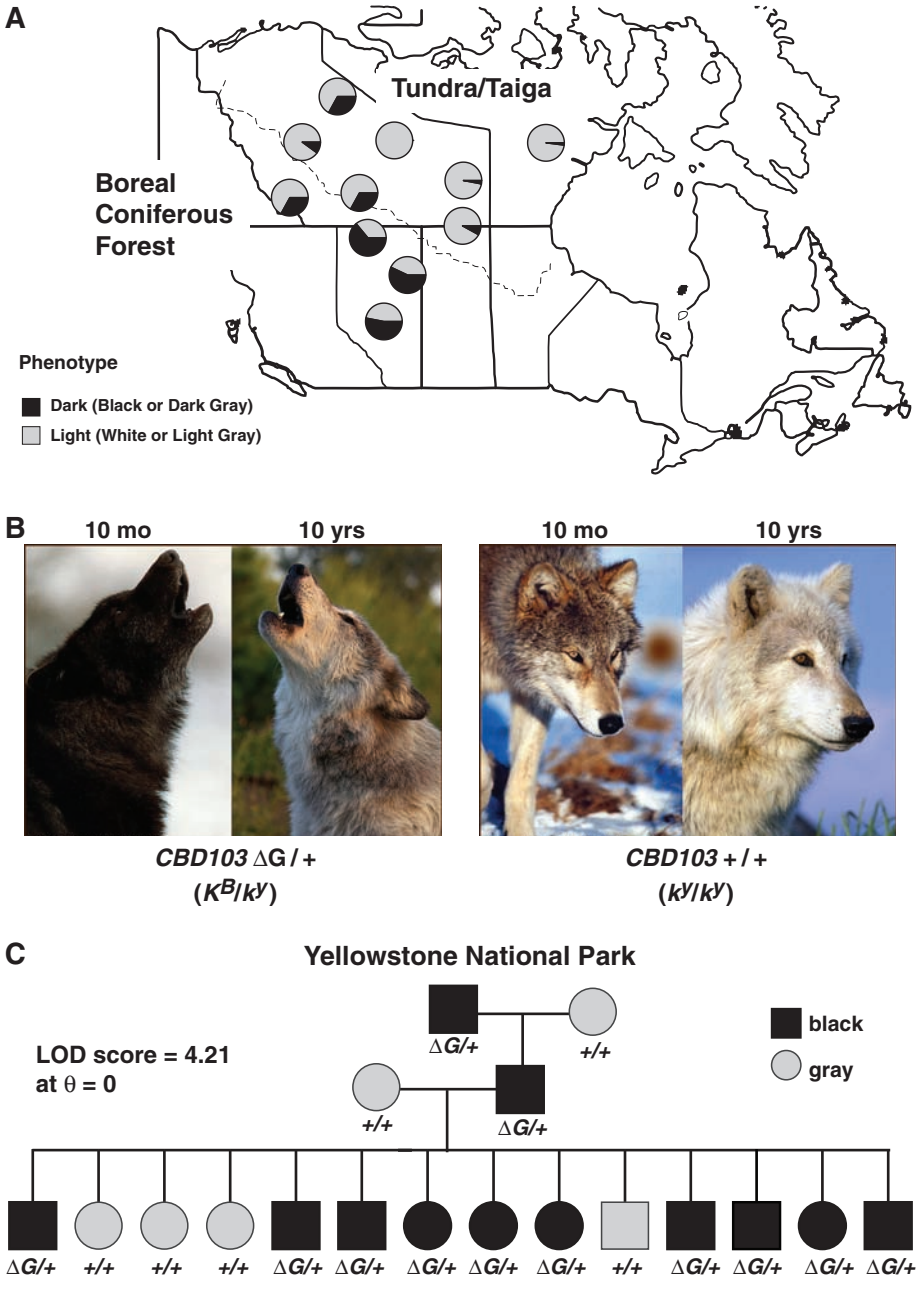

Fig. 1. Distribution of melanism and *K* locus genotypes in North American gray wolves. (**A**) Location and coat color phenotype of Canadian samples used here and as described (*4*). (**B**) Age-related graying and the associated difficulty of inferring genotype from phenotype in gray animals. Each pair of photos shows the same individual at different ages (10 months and 10 years) and documents an increasingly gray appearance at 10 years, reflecting the dilution of eumelanin in the K^B/k^y individual (left pair of images) and dilution of both eumelanin and pheomelanin in the k^y/k^y individual (right pair of images). [Images courtesy of Monty Sloan, Wolf Park, Battle Ground, Indiana] (**C**) Co-segregation of K^B and black coat color in a three-generation pedigree from the Leopold pack in Yellowstone National Park (*17*). ΔG indicates the dominant K^B allele, whereas + indicates the wild-type allele, k^y.

color has been suggested to include concealment during predation and/or indirect effects due to pleiotropy, but remains unresolved because the underlying gene(s) have not been identified (*5*–*7*).

In many vertebrates, natural pigmentary variation is controlled by the agouti– melanocortin 1 receptor (Mc1r) pathway, a ligand receptor pair that modulates the amount and type of pigment—red/yellow pheomelanin or brown/black eumelanin—produced by melanocytes in skin, hair, or feathers. Gain-of-function *Mc1r* mutations are well-recognized causes of melanism in many domestic and laboratory animal species (*8*, *9*), as well as in several natural populations of birds (*10*), rodents (*11*, *12*), and canids (*13*). Recently, we found that pigment type-switching in domestic dogs involves an additional component of the melanocortin pathway, the *K* locus, which encodes a beta-defensin protein, CBD103 (*14*, *15*).

Coat color in Canadian wolves is genetically complex, with phenotypes ranging from white to gray to black, and is also confounded by an independent effect of graying with age (Fig. 1B). However, in Yellowstone National Park, where a small number of founder animals from Canada were recently reintroduced (*16*, *17*), gray and black coat colors segregate as a Mendelian trait. We surveyed molecular variation in *Agouti*, *Mc1r*, and *CBD103* in wolves from North America and identified several *Mc1r* and *Agouti* polymorphisms. However, none of these were predicted to affect gene function and did not associate with black coat color (table S1). In contrast, in a 14-member, three-generation kindred from Yellowstone, we observed complete co-segregation between black coat color and markers at the *K* locus [logarithm of the odds ratio for linkage (lod) score = 4.21 at the maximum likelihood estimate of recombination fraction (θ) = 0, Fig. 1C], which is unlinked and lies on a different chromosome from *Agouti* and *Mc1r*.

In dogs, the ancestral *CBD103* allele (k^y) confers normal *Agouti* and *Mc1r* gene action, whereas a 3–base pair (bp) deletion ($CBD103^{\Delta G23}$ or K^B) suppresses *Agouti* gene action, leading to dominant inheritance of a black coat (*14*, *15*). We observed the same 3-bp deletion in 102 out of 104 black-colored wolves from Yellowstone and 9 out of 9 from the Canadian Arctic. Conversely, $CBD103^{\Delta G23}$ was absent from 120 of 120 gray-colored wolves from Yellowstone and from 22 of 22 white-colored wolves from the Canadian Arctic (Table 1). We also found $CBD103^{\Delta G23}$ in 6 of 10 gray-colored wolves from the Canadian Arctic, suggesting that gray coat color can result either from the absence of $CBD103^{\Delta G23}$ and a modified agouti phenotype (in which individual hairs contain both cream-colored pheomelanin and dark eumelanin) or from secondary factors such as age that dilute the pigmentation of hairs that contain only eumelanin. [Additional genealogy studies of the

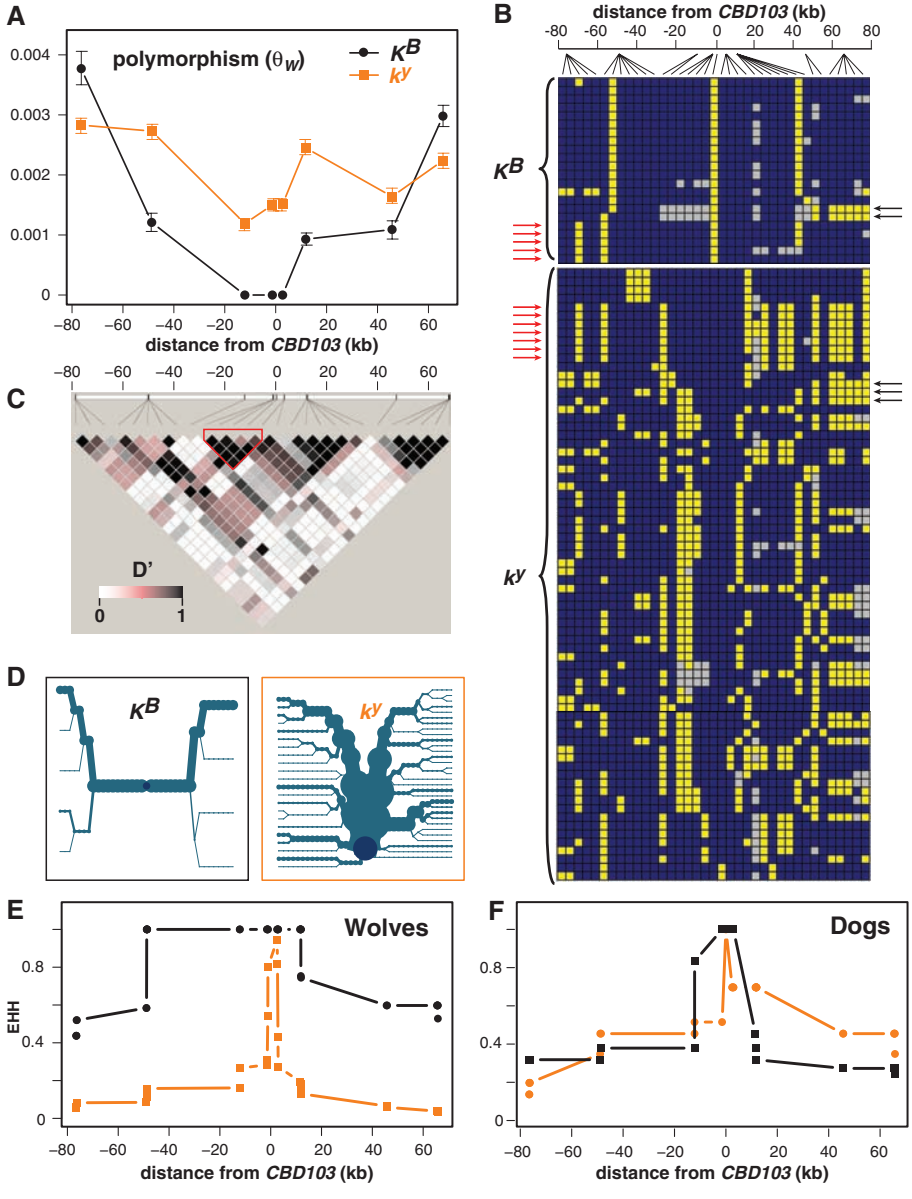

Fig. 2. Polymorphism and haplotype structure of the *K* locus in North American gray wolves [(A) to (E), 1 K^B/K^B, 20 K^B/k^y, and 26 k^y/k^y] and domestic dogs [(F), 6 K^B/K^B and 6 k^y/k^y]. (**A**) Polymorphism (θ_W, ±SD) as a function of distance from *CBD103*. (**B**) Wolf haplotype structure was inferred on the basis of 36 SNPs; each row represents a K^B- or k^y-bearing chromosome; blue and yellow squares represent the major and minor alleles, respectively; and the gray squares represent missing data. Red and black arrows indicate examples of haplotypes likely to represent historical recombination between K^B- and k^y-bearing chromosomes at the 5′ and 3′ ends of the locus, respectively. (**C**) Pairwise LD values (expressed as D') for all wolf chromosomes; the red outline indicates a core region (as in Fig. 3) unlikely to have undergone historical recombination. (**D**) Haplotype bifurcation diagrams for K^B- or k^y-bearing chromosomes, in which the central dark blue dot represents *CBD103*, branches represent haplotype divergence, and the thickness of the lines is proportional to the number of chromosomes. (**E** and **F**) EHH for K^B- or k^y-bearing chromosomes in wolves (E) and dogs (F) as a function of distance from $CBD103^{\Delta G23}$.

Yellowstone population (*17*) together with the paucity of *Mc1r* variation in wolves (table S1) suggest that black coat color reported for the two k^y/k^y Yellowstone wolves is likely to reflect phenotypic ambiguity or misclassification at the time of sampling.] Allele frequencies for $CBD103^{\Delta G23}$ in tundra and forest wolves overall were estimated at 0.02 and 0.19, corresponding to phenotype frequencies of 2 to 33% and 33 to 64% for dark wolves in tundra and forest populations, respectively (Fig. 1A) (*4*).

To investigate the evolutionary history of the melanistic *K* allele, we sequenced eight single-copy noncoding segments distributed across an ~150-kb region centered on *CBD103* in 32 Arctic and 15 unrelated Yellowstone wolves, as well as in 12 domestic dogs: 6 k^y/k^y (akita, basenji, boxer, bulldog, Doberman pinscher, and great dane) and 6 K^B/K^B (curly-coated retriever, Dalmatian, great dane, Labrador retriever, poodle, and Portuguese water dog). We identified 52 biallelic polymorphisms across all canids (36 in wolves) and estimated haplotype structure (tables S3 and S4, Fig. 2B, and fig. S2). The rate of polymorphism among all wolf amplicons was one single-nucleotide polymorphism (SNP) per 510 bp (Watterson's estimator, $\theta_W = 1.96 \times 10^{-3}$), which is similar to genome-wide measurements of polymorphism between the boxer and the gray wolf (1 out of 580 bp) and the coyote (1 out of 420 bp) (*18*). However, partitioning our data according to *K* locus genotype and proximity to *CBD103* revealed little or no polymorphism among K^B-bearing chromosomes close to *CBD103*, rising to levels at or above those observed in k^y-bearing chromosomes in the 75 kb spanning either side of the locus (Fig. 2A). This pattern, and the analogous one for nucleotide diversity (π, fig. S1), is also reflected in a significant difference in haplotype diversity between K^B (8 unique of 22 total) and k^y (59 unique of 72 total) chromosomes ($\chi^2 = 14.2$, $P < 0.001$). Together with the correlations between coat color and habitat (*2*–*5*), the combination of low diversity and high frequency suggests that K^B has been under positive selection in North American forest wolves.

Overall, the patterns of linkage disequilibrium (LD) across 150 kb surrounding the *K* locus were similar to comparisons between different breeds of domestic dogs (*18*), with relatively small haplotype blocks, including an ~4-kb *CBD103* core region within which there is no evidence for historical recombination (Fig. 2C). However, different evolutionary histories for the Arctic wolf K^B and k^y alleles were apparent when the SNP patterns (Fig. 2B) were depicted as haplotype bifurcation diagrams (Fig. 2D), which highlight a central region of ~60 kb devoid of polymorphism among wolf K^B haplotypes. This characteristic, and the corresponding difference between K^B and k^y chromosomes, were represented quantitatively by the extended haplotype homozygosity (EHH) statistic (*19*), which is the empirical probability that two chromosomes chosen at random remain identical at progressively increasing distances from *CBD103*. As depicted in Fig. 2, E and F, the distribution of EHH was considerably broader for K^B as compared to k^y chromosomes in wolves, whereas the

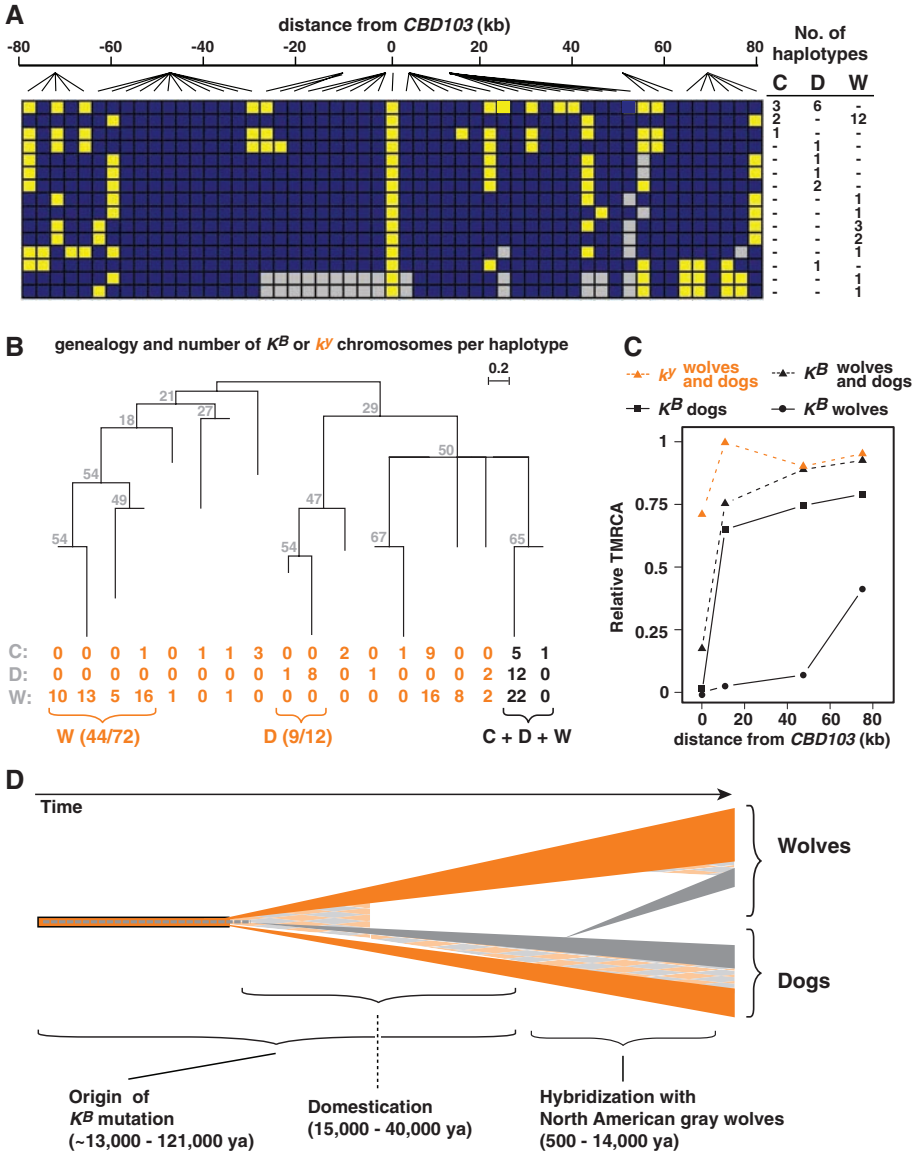

Fig. 3. Evolutionary relationships and history of the *K* locus in canids. (**A**) K^B haplotype structure in wolflike canids based on genotypes defined by 52 SNPs. Each row represents a K^B-bearing haplotype found in coyotes (C), dogs (D), or wolves (W) listed with their respective frequencies on the right and colored as in Fig. 2B. (**B**) Inferred genealogical relationships of the core region (Fig. 2C) haplotypes (with bootstrap values from 500 replicates shown next to branches). Each branch represents 1 of 18 different haplotypes, with the number of chromosomes for each haplotype indicated underneath according to species. (**C**) TMRCA estimates for indicated chromosome subsets calculated according to a molecular clock (*22*) and expressed as a fraction of the divergence time for all wolflike canids. Individual points represent sets of chromosome segments whose relative TMRCA increases as a function of distance from *CBD103*, presumably due to ancient hybridization and recombination. (**D**) Timeline scenario for *K* locus evolution in dogs and wolves, in which ancestral k^y chromosomes are indicated in orange, derivative K^B chromosomes in gray, and recombinant chromosomes as an orange-gray checkered pattern. The k^y-to-K^B mutation may have overlapped or even predated domestication, but the introgression of K^B into North American gray wolves is more recent.

distributions were nearly identical for K^B as compared to k^y chromosomes in dogs. Together with additional analyses of genome-wide SNP data [supporting online material (SOM) text and fig. S3], these observations suggest that K^B has risen to high frequency by a selective sweep.

As in black dogs and melanistic wolves, $CBD103^{\Delta G23}$ was associated with coat color in 67 coyotes (6 black and 61 gray, Table 1 and table S2). These findings suggest three possible evolutionary histories. First, the 3-bp deletion may be relatively old, having occurred in a canid ancestor more than 1 million years ago before the divergence of coyotes from wolves. Second, the 3-bp deletion may have occurred more recently in one of the species, followed by introgression into the others. Finally, the 3-bp deletion may represent a mutational hotspot, having recurred independently in coyotes, wolves, and dogs. To distinguish among these possibilities, we ascertained and compared coyote haplotypes (6 K^B and 18 k^y) with those from the North American wolf and dog.

The pattern of haplotype diversity for all three canids was similar to that observed in wolves alone and showed significantly less diversity among K^B (15 unique of 40 total) relative to k^y (66 unique of 102 total) chromosomes ($\chi^2 = 9.7$, $P = 0.003$). Of the 15 unique K^B haplotypes, 1 haplotype was observed in three coyotes and six dogs, and a second haplotype was observed in two coyotes and 12 wolves (Fig. 3A). However, none of the 66 unique k^y haplotypes were observed in more than one species (fig. S2).

Reconstruction of a phylogenetic network for the entire 150-kb region is complicated by historical recombination between extant K^B and k^y chromosomes (arrows in Fig. 2B) and the lack of a suitable approach for inferring accurate gene genealogies in the presence of recombination (*20*). However, by focusing on the 4-kb *CBD103* core region (Fig. 2C), a simple neighbor-joining tree was constructed for 18 core region haplotypes representing 142 (94 wolf, 24 dog, and 24 coyote) chromosomes (Fig. 3B). In this tree, all the K^B chromosomes define a 2-haplotype cluster, whereas the remaining 16 haplotypes (which represent all the k^y chromosomes) are more dispersed. Furthermore, many of the k^y chromosomes cluster by species (9 out of 12 of the dogs and 44 out of 72 of the wolves), unlike the K^B chromosomes. This contrasting phylogenetic pattern suggests that the K^B mutation occurred in a single species and was later distributed among dogs, wolves, and coyotes by interspecific hybridization. [The 24 k^y haplotypes from coyotes are no closer to each other than to k^y haplotypes from wolves or dogs (Fig. 3B), which is consistent with their history of hybridization with other canids (*21*)].

To gain additional insight into how *K* locus variation in dogs and wolves arose, we estimated coalescent time to the most recent common ancestor (TMRCA) as a function of cumulative distance from *CBD103* for k^y and K^B chromosomes from wolves, dogs, and both groups together. We applied a molecular clock approach to sequencing data from individual amplicons across the entire 150-kb region (Fig. 2), which assumes that mutations occur at the same constant rate at all sites in wolves and dogs and integrates the effects of both recombination and demography (*22*). Close to *CBD103*, TMRCA estimates were near zero for all K^B subsets (Fig. 3C) because there is little or no polymorphism in this region (Fig. 3A). However, at greater distances from *CBD103* (10 to 50 kb), estimates for dog chromosomes are similar to those of dog and wolf chromosomes considered together, regardless of genotype. This suggests that K^B in dogs is sufficiently old to have undergone extensive recombination with k^y chromosomes, and that the recombination history includes hybridization between dogs and wolves. However, in the same 10- to 50-kb range, TMRCA estimates for wolf K^B chromosomes were considerably less than those from dog K^B chromosomes (or from dog and wolf K^B chromosomes considered together), suggesting that K^B was introduced into North American wolves from dogs, not vice versa.

Introgression of K^B from dogs into North American wolves is also supported by geographical and ecological considerations. K^B is widely distributed among domestic dogs, including ancient breeds originating in Asia and Africa. In wolves, however, melanism has been reported outside North America only in Italy, where it is associated with molecular and/or morphologic evidence of recent hybridization with free-ranging dogs (*23*). Indeed, we also examined 22 samples from the Italian Apennines and observed K^B in six of seven black "wolves" (including one previously classified to be a dog-wolf hybrid) but 0 of 15 gray wolves. In contrast, genome-wide SNP analysis of 10 K^B/k^y and 10 k^y/k^y North American wolves showed no evidence for recent dog-wolf hybridization (SOM text and fig. S3B).

The dog was domesticated between 15,000 and 40,000 years ago in East Asia from gray wolves (*24*, *25*), and we estimate that K^B is at least 46,886 years old (95% confidence limit: 12,779 to 121,182 years); therefore, we cannot distinguish whether K^B arose before or after domestication. However, if K^B arose in Old World wolves before domestication, our data indicate that it must have been lost from the gene pool and reacquired in North America, perhaps from Native American dogs that accompanied humans across the Bering Strait 12,000 to 14,000 years ago (*26*) (Fig. 3D).

The wolf in the United States faces grave threats, in some cases by eradication, and in others by hybridization, such as in the Great Lakes region (*27*). However, apparent selection for the K^B locus in North American gray wolves shows how genetic diversity—preserved by humans in domestic dogs—may flourish in wild wolf populations. As the available tundra habitat declines because of development and/or global warming, the frequency of the K^B mutation may increase further in northern latitudes. Thus, the introduction of genetic diversity into a natural population from a mutation originally selected in domesticated animals may, ironically, provide a mechanism for that population to adapt to a changing environment. Interspecific hybridization has been widely observed between other domesticated species of animals and plants (*28*–*30*). Our results imply that variants that appear under domestication can be viable in the wild and enrich the genetic legacy of natural populations.

References and Notes

1. P. S. Gipson *et al.*, *Wildl. Soc. Bull.* **30**, 821 (2002).
2. L. D. Mech, L. G. Adams, T. J. Meier, J. W. Burch, B. W. Dale, *The Wolves of Denali* (Univ. of Minnesota Press, Minneapolis, MN, 1998).
3. R. O. Peterson, J. D. Wollington, T. N. Bailey, *Wildl. Monogr.* **88**, 3 (1984).
4. M. Musiani *et al.*, *Mol. Ecol.* **16**, 4149 (2007).
5. P. Jolicoeur, *Evolution* **13**, 283 (1959).
6. M. E. N. Majerus, *Melanism: Evolution in Action* (Oxford Univ. Press, Oxford, 1998).
7. A. L. Ducrest, L. Keller, A. Roulin, *Trends Ecol. Evol.* **23**, 502 (2008).
8. H. Klungland, D. I. Vage, *Ann. N. Y. Acad. Sci.* **994**, 331 (2003).
9. L. Andersson, *Ann. N. Y. Acad. Sci.* **994**, 313 (2003).
10. N. I. Mundy *et al.*, *Science* **303**, 1870 (2004).
11. H. E. Hoekstra, R. J. Hirschmann, R. A. Bundey, P. A. Insel, J. P. Crossland, *Science* **313**, 101 (2006).
12. M. W. Nachman, H. E. Hoekstra, S. L. D'Agostino, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 5268 (2003).
13. D. I. Vage *et al.*, *Nat. Genet.* **15**, 311 (1997).
14. J. A. Kerns *et al.*, *Genetics* **176**, 1679 (2007).
15. S. I. Candille *et al.*, *Science* **318**, 1418 (2007).
16. E. E. Bangs, S. Fritts, *Wildl. Soc. Bull.* **24**, 402 (1996).
17. B. M. vonHoldt *et al.*, *Mol. Ecol.* **17**, 252 (2008).
18. K. Lindblad-Toh *et al.*, *Nature* **438**, 803 (2005).
19. P. C. Sabeti *et al.*, *Nature* **419**, 832 (2002).
20. S. M. Woolley, D. Posada, K. A. Crandall, *PLoS ONE* **3**, e1913 (2008).
21. M. S. Roy, E. Geffen, D. Smith, E. A. Ostrander, R. K. Wayne, *Mol. Biol. Evol.* **11**, 553 (1994).
22. H. Tang, D. O. Siegmund, P. Shen, P. J. Oefner, M. W. Feldman, *Genetics* **161**, 447 (2002).
23. E. Randi, V. Lucchini, *Conserv. Genet.* **3**, 29 (2002).
24. C. Vila *et al.*, *Science* **276**, 1687 (1997).
25. P. Savolainen, Y. P. Zhang, J. Luo, J. Lundeberg, T. Leitner, *Science* **298**, 1610 (2002).
26. J. A. Leonard *et al.*, *Science* **298**, 1613 (2002).
27. J. A. Leonard, R. K. Wayne, *Biol. Lett.* **4**, 95 (2008).
28. R. Lecis *et al.*, *Mol. Ecol.* **15**, 119 (2006).
29. N. Halbert, J. Derr, *J. Hered.* **98**, 1 (2007).
30. N. Ellstrand, H. Prentice, J. Hancock, *Annu. Rev. Ecol. Syst.* **30**, 539 (1999).

31. Supported by grants from NIH (G.S.B.), NSF (R.K.W., D.R.S., and D.W.S.), and the Swedish Research Council (J. A.L.). We are grateful to H. Chen and S. Schmutz for advice, to H. Manuel for technical assistance, and to members of the U.S. Department of Agriculture Wildlife Services and private citizens for assistance with sample collection. Sequences generated in this study are deposited in GenBank under accession numbers FJ609634 to FJ609641.

Supporting Online Material

www.sciencemag.org/cgi/content/full/1165448/DC1
Materials and Methods
SOM Text
Figs. S1 to S4
Tables S1 to S5
References

3 September 2008; accepted 15 January 2009
Published online 5 February 2009;
10.1126/science.1165448
Include this information when citing this paper.

Supporting Material

Materials and Methods

Sample collection

DNA samples from Canadian (*S1*), Yellowstone (*S2*), and Italian (*S3, S4*) wolves were extracted from blood, pelts, or tissues. DNA samples from domestic dogs were collected by cheek swab from owners and breeders according to a Stanford IACUC-approved protocol; all samples were prepared, as described (*S5*). DNA samples from 61 gray and 2 black coyotes were provided through a generous collaboration with the United States Department of Agriculture (John Pingley and Bill Bonwell, USDA Wildlife Services, West Virginia; Chad Fox, Harold McDaniel Jr., and Eric Wilhelm, USDA Wildlife Services, Virginia). Tissue samples from an additional 4 black coyotes were provided by Ed Smallwood, Mike Laska, and Matthew Boggs.

Amplification and sequencing of *Agouti*, *Mc1r*, and *CBD103*

The wolf and coyote *Agouti* genes were analyzed in 4 amplicons as follows: 5'-CACCCAACACACTTCTGCG-3' and 5'-TACCATACCAAAACATCTGC-3' (283 bp, exon 1); 5'-AGGGCACAGCCTCTTATCAA-3' and 5 '-CAGGGCTTTTCCAAACCATA-3' (617 bp, exon 2); 5'-CACCCTGAGACTTCCTGGAG-3' and 5'-GAGGCCAAGAAGCCTTTAGA-3' (295 bp, exon 3); 5'-AAGTCCAGCGGACAGTCG-3' and 5'-CACACCTTGGAGCAGCCTA-3' (636 bp, exon 4). [Previous results (*S6, S7*) indicate the potential for alternative 5' untranslated exons in canid *Agouti*; the amplicon containing exon 1 described here corresponds to a cDNA isolated from a Doberman Pinscher and lies in a genomic location homologous to what has been described as the ventral-specific exons from the mouse *Agouti* gene (*S8*)]. Additional SNPs from intronic regions of *Agouti* are listed in Table S2; primer sequences are available on request. The wolf and coyote *Mc1r* genes were analyzed in 3 amplicons covering a single large coding exon: 5'-CACTTGTACAGACCGGGAGAG-3' and 5'-ACGTCAATGATGTCGTCCAG-3' (493 bp, first third); 5'-GTGACGAATGTGCTGGAGAC-3' and 5'-AAATGCCCAGCAGGATAGTG-3' (484 bp, central third); 5'-TCTTTGTAGCCATGCTGGTG-3' and 5'-ATCCACCACACCACAGATCA-3' (486 bp, last third). The wolf and coyote *CBD103* mature coding regions (the location of the K^B ΔG23 mutation) were analyzed in a single amplicon, 5'-TGTCTTCATCCCTGTGAGGT-3' and 5 '-CCAGGAGGCATTTTCACACT-3' (396 bp).

For PCR, a touchdown protocol was used with the following conditions: (1) 94°/90 sec; (2) 94°/30 sec; 65°/30 sec (-0.5°/cycle); (4) 72°/60 sec; (5) repeat steps 2 – 4 x 20; (6) 94°/30 sec; (7) 55°/30 sec; (8) 72°/60 sec; (9) repeat steps 6 – 8 x 20; (10) 72°/5 min. Direct sequencing of amplified product after primer hydrolysis with ExoSapIt (USB, Cleveland, OH) was carried out with standard fluorescent dye terminator technology on a capillary instrument.

Genomic and bioinformatic analysis of *CBD103* and flanking regions

Genome sequence coordinates used throughout this paper refer to the dog genome assembly CanFam2.0 (*S9*) as displayed and annotated in the UCSC genome browser,

http://genome.cse.ucsc.edu/ (*S10*). In this assembly, *CBD103* G23 lies at chr16:61,902,782-61,902,785, very close to the end of chromosome 16 (62,570,175); however, our previous studies suggest the last 5 Mb of the chromosome are mistakenly inverted in the CanFam2.0 assembly (*S5*).

For population genetic analysis of *CBD103* and surrounding regions, 8 single-copy noncoding segments distributed across a ~150 kb region centered on *CBD103* were analyzed with nested primers, whose position, sequence, and associated polymorphisms are given in Tables S2 and S3. Amplification with outer primers was carried out with the following conditions: (1) 92°/60 sec; (2) 92°/20 sec; (3) 65°/20 sec (-0.6°/cycle); (4) 72°/30 sec; (5) repeat steps 2 – 4 x 11; (6) 92°/20 sec; (7) 58°/20 sec; (8) 72°/30 sec; (9) repeat steps 6 – 8 x 20; (10) 72°/3 min. Then, amplification with inner primers was carried out with the following conditions: (1) 92°/60 sec; (2) 92°/20 sec; (3) 58°/20 sec (4) 72°/30 sec; (5) repeat steps 2- 4 x 29; (6) 72°/3 min. Other conditions for amplification were carried out as described (*S11*), which permits direct sequencing after amplification without primer hydrolysis.

All sequencing results were analyzed with CodonCode Aligner (CodonCode Corporation, Dedham, MA) or Sequencher 4.2 (Gene Codes Corporation, Ann Arbor, MI), and for every polymorphism between or within a species, the trace files were inspected visually for confirmation.

Nucleotide diversity and polymorphism statistics were determined with DnaSP 4.0 (*S12*), haplotypes for the region surrounding *CBD103* were inferred with PHASE 2.1 (*S13*), extended haplotype homozygosity statistics and haplotype bifurcation diagrams were calculated and visualized with Sweep (*S14*), LD values were calculated and visualized with Haploview (*S15*), phylogenetic analysis of the *CBD103* core region was carried out with MEGA (*S16*), and the mtDNA network (Fig. S3) was constructed with TCS (*S17*). The TMRCA molecular clock analysis was carried out, as described (*S18*); estimates from 5' and 3' polymorphisms were combined, averaged, and presented as a function of absolute distance (independent of direction) from *CBD103*.

Supporting text

Population genetics and evolutionary history of the K^B mutation

We estimated allele frequencies of 0.19 and 0.02 for K^B in forest and tundra wolves, using coat color frequencies reported for 11 different Arctic populations; 4 forest populations representing a total of 68 wolves, and 7 tundra or taiga populations representing a total of 336 wolves (Fig. 1A, *S1*). These estimates assume that 50% of the gray wolves carry K^B but are non-penetrant due to graying with age.

Our estimate for the age of the k^y to K^B mutation derives from dog resequencing data. In previous studies, we identified 4 K^B/K^B dogs (2 Great Danes, 1 Large Munsterlander, 1 Poodle) in which haplotype analysis delineated a 9.1 kb interval containing *CBD103* that lacked evidence of historical recombination (*S5*). Among these 8 K^B chromosomes, we identified 3 mutations from 63984 bp of high-quality sequence. Assuming a mutation rate of 1×10^{-9}/yr and a star-shaped genealogy, these data predict that K^B occurred 46,886 years ago, with a 95% confidence limit of 12,779 - 121,182 years. This is a minimal estimate because the dogs we analyzed may sample only a subset of all K^B chromosomes.

However, regardless of whether the original k^y to K^B mutation occurred prior, during, or after domestication of dogs from wolves, the haplotype data (Fig. 3A, Fig. S3), the TMRCA analysis (Fig. 3C), and the geographical distribution of K^B (which is much broader in dogs than in wolves) suggest that North American gray wolves acquired K^B from dogs and not vice versa.

Genetic relatedness among wolves and dogs: effect of the K^B mutation

The higher frequencies of K^B in forest compared to tundra wolves that we (*S1*) and others (*S19 – S21*) have observed suggest that melanism and/or K^B provide a selective advantage in forested habitats. An alternative hypothesis is that K^B was introduced recently into North American gray wolves, and that the differences in habitat-specific allele and phenotype frequencies are due to genetic drift. If so, one would expect black wolves to carry additional regions of the dog genome besides the K^B-bearing segment on chromosome 16, as a signature of residual hybridization with dogs.

To explore this possibility, we analyzed large-scale genotype data obtained with the Affymetrix v2 platform, which represents 49,663 SNPs distributed widely across the canine genome. These data are part of a larger study that will be described in detail elsewhere (R.K.W., C.D.B., E.A.O., B.M.V.). We examined a subset of the data selected for geographic origin and *K* locus genotype to investigate if K^B was associated with unlinked markers due to population structure, and to determine if there was cryptic relatedness between dogs and wolves carrying K^B.

We first analyzed genotypes from 10 Canadian wolves (5 K^B/k^y and 5 k^y/k^y) and 10 Yellowstone wolves (5 K^B/k^y and 5 k^y/k^y), and asked whether any SNPs other than those closely linked to *CBD103* were associated with K^B. From 28,739 SNPs with call rates > 80% and a minor allele frequency > 0.05, 32 lie within 2 Mb of *CBD103*, but outside the core region previously used to define *K* locus haplotypes (Figs. 2B, 3A, S2). In wolves, none of these 32 SNPs were strongly associated with K^B (Table S5). As a more sensitive test, we carried out a principal component analysis of the 20 wolves using progressively larger genomic regions: SNPs on the distal 10 Mb of dog chromosome 16 (144), all SNPs on the same chromosome as *CBD103* (739), and all SNPs (28,739). For each set of SNPs, one or more of the major principal components separated Canadian from Yellowstone wolves but did not distinguish K^B/k^y from k^y/k^y individuals (Fig.S3A, S3B), which indicates that geography is the major source of genetic differentiation, and suggests that the K^B-associated segment introgressed from dogs into black wolves was not accompanied by other large regions of the dog genome.

To probe potential relatedness between North American wolves and dogs, we considered data for the 20 wolves together with 20 dogs from different breeds and 5 wolf-dog hybrids, and applied principal component analysis to the full set of ~45,000 SNPs. For this dataset, the first principal component separated dogs and wolves, leaving the Italian dog-wolf hybrids at an intermediate position, while the second principal component separated the hybrids from both the dogs and the North American wolves (Fig. S3C). As with the previous analysis of wolves alone, K^B/k^y individuals were not separated from k^y/k^y individuals.

As a complementary approach, we used mitochondrial haplotypes previously obtained for 37 Canadian wolves (*S1*) to construct a phylogenetic network, and then projected

information about *K* locus genotype onto that network. As shown in Fig. S4, mitochondrial DNA haplotypes for K^B-carrying wolves are broadly distributed, and no more closely related to each other or to dog haplotypes than to haplotypes for non-K^B wolves. Taken together, these observations suggest that the geographic differences in K^B allele frequencies are explained by a positive selection for K^B in forested compared to open habitats, and not by genetic drift following dog-wolf hybridization.

Supplemental references

S1. M. Musiani *et al.*, *Mol Ecol* 16, 4149 (2007).

S2. B. M. Vonholdt *et al.*, *Mol Ecol* 17, 252 (2008).

S3. E. Randi, V. Lucchini, *Conservation Genetics* 3, 29 (2002).

S4. E. Fabbri *et al.*, *Mol Ecol* 16, 1661 (2007).

S5. S. I. Candille *et al.*, *Science* 318, 1418 (2007).

S6. D. I. Vage *et al.*, *Nat Genet* 15, 311 (1997).

S7. J. A. Kerns *et al.*, *Mamm Genome* 15, 798 (2004).

S8. H. Vrieling, D. M. Duhl, S. E. Millar, K. A. Miller, G. S. Barsh, *Proc Natl Acad Sci U S A* 91, 5667 (1994).

S9. K. Lindblad-Toh *et al.*, *Nature* 438, 803 (2005).

S10. D. Karolchik *et al.*, *Nucleic Acids Res* 31, 51 (2003).

S11. B. M. Smits *et al.*, *Pharmacogenet Genomics* 16, 159 (2006).

S12. J. Rozas, J. C. Sanchez-DelBarrio, X. Messeguer, R. Rozas, *Bioinformatics* 19, 2496 (2003).

S13. M. Stephens, N. J. Smith, P. Donnelly, *Am J Hum Genet* 68, 978 (2001).

S14. P. C. Sabeti *et al.*, *Nature* 419, 832 (2002).

S15. J. C. Barrett, B. Fry, J. Maller, M. J. Daly, *Bioinformatics* 21, 263 (2005).

S16. K. Tamura, J. Dudley, M. Nei, S. Kumar, *Mol Biol Evol* 24, 1596 (2007).

S17. M. Clement, D. Posada, K. A. Crandall, *Mol Ecol* 9, 1657 (2000).

S18. H. Tang, D. O. Siegmund, P. Shen, P. J. Oefner, M. W. Feldman, *Genetics* 161, 447 (2002).

S19. P. S. Gipson *et al.*, *Wildlife Society Bulletin* **30**, 821 (2007).

S20. L. D. Mech, L. G. Adams, T. J. Meier, J. W. Burch, B. W. Dale, *The Wolves of Denali*. (University of Minnesota Press, 1998).

S21. R. O. Peterson, J. D. Wollington, T. N. Bailey, *Wildlife Monographs* **88**, 3 (1984).

S22. C. Vila *et al.*, *Science* 276, 1687 (1997).

Table S1. *Agouti* and *Mc1r* variation in black and non-black wolves

Gene	Position[1]	Location[1]	SNP[2]	Effect[2]	No. black[3]			No. non-black[3]		
					R/R	R/V	V/V	R/R	R/V	V/V
Agouti	chr24:26,351,138	5' flank	G/A	No	4	0	0	5	1	0
Agouti	chr24:26,351,277	5' flank	G/A	No	0	0	4	0	1	5
Agouti	chr24:26,357,184	Intron 1	A/G	No	0	1	3	0	1	5
Agouti	chr24:26,357,895	Intron 1	T/G	No	0	0	3	0	1	5
Agouti	chr24:26,358,085	Intron 1	C/G	No	0	0	3	0	1	5
Agouti	chr24:26,359,694	Intron 1	G/A	No	3	0	0	4	1	0
Agouti	chr24:26,359,752	Intron 1	T/C	No	0	1	2	0	1	4
Agouti	chr24:26,360,917	Intron 1	G/A	No	0	1	3	0	1	5
Agouti	chr24:26,360,980	Intron 1	G/A	No	0	1	3	0	1	5
Agouti	chr24:26,361,818	Intron 2	T/A	No	0	1	3	0	1	2
Agouti	chr24:26,363,092	Intron 3	G/A	No	3	1	0	5	1	0
Agouti	missing	Coding	G/T	A82S	4	0	0	5	1	0
Agouti	chr24:26,365,984	Coding	G/A	No	4	0	0	5	1	0
Agouti	chr24:26,366,507	3' flank	C/T	No	4	0	0	4	2	0
Agouti	chr24:26,366,958	3' flank	G/C	No	3	1	0	3	3	0
Mc1r	chr5:66,693,084	Coding	A/G	S90G	10	0	0	16	1	0
Mc1r	chr5:66,692,869	Coding	G/A	No	10	0	0	16	1	0

[1]Location in genome from CanFam 2.0 (*S9*); position given relative to *Agouti* or *Mc1r* transcript. The region of the *Agouti* gene coding for codon 82 is missing from the dog assembly.

[2]SNP information given as reference sequence/variant sequence, with predicted effect, if any, on protein.

[3]No. of individuals according to genotype of Reference and Variant alleles. Wolves surveyed for variation in *Agouti* were from Canada. Wolves surveyed for variation in *Mc1r* were from Canada (21) or Yellowstone National Park (6); we also sequenced *Mc1r* in 2 Swedish wolves (not included in the Table) and did not observe any variation.

Table S2. *Agouti* and *Mc1r* variation in black and gray coyotes

Gene	Location[1]	Position[1]	SNP[2]	Effect[2]	No. black[3]			No. gray[3]		
					R/R	R/V	V/V	R/R	R/V	V/V
Agouti	chr24:26,326,144	5' flank	C/T	No	0	0	2	1	0	3
Agouti	chr24:26,327,291	5' flank	T/C	No	0	0	2	0	1	3
Agouti	chr24:26,327,318	5' flank	A/G	No	0	0	2	0	1	3
Agouti	chr24:26,327,427	5' UTR	T/G	No	2	0	0	3	1	0
Agouti	chr24:26,366,050	Coding	C/G	No	1	0	5	2	2	8
Agouti	chr24:26,366,076	3' UTR	G/C	No	3	3	0	6	0	6
Agouti	chr24:26,366,591	3' flank	G/C	No	2	0	0	2	1	0
Agouti	chr24:26,366,623	3' flank	C/A	No	2	0	0	2	1	0
Agouti	chr24:26,366,627	3' flank	G/T	No	2	0	0	2	1	0
Agouti	chr24:26,366,778	3' flank	T/C	No	2	0	0	2	0	1
Agouti	chr24:26,366,958	3' flank	G/C	No	2	0	0	2	1	0
Mc1r	chr5:66,693,444	5' flank	C/T	No	6	0	0	59	2	0
Mc1r	chr5:66,693,226	Coding	C/T	No	3	3	0	22	29	10
Mc1r	chr5:66,693,084	Coding	A/G	S90G	6	0	0	60	1	0
Mc1r	chr5:66,693,068	Coding	C/T	T95M	6	0	0	58	3	0
Mc1r	chr5:66,693,047	Coding	C/G	A102G	1	4	1	27	22	12
Mc1r	chr5:66,693,036	Coding	T/C	No	5	1	0	50	8	3
Mc1r	chr5:66,692,869	Coding	G/A	No	5	1	0	57	4	0
Mc1r	chr5:66,692,684	Coding	G/A	R223Q	6	0	0	60	1	0
Mc1r	chr5:66,692,562	Coding	G/A	M264V	0	0	6	0	4	53

[1]Location in genome from CanFam 2.0 (*S9*); position given relative to *Agouti* or *Mc1r* transcript.

[2]SNP information given as reference sequence/variant sequence, with predicted effect, if any, on protein.

[3]No. of individuals according to genotype of Reference and Variant alleles.

Table S3. Oligonucleotide primers for *CBD103* and surrounding region[1]

Name	Size	Location	Forward	Reverse
1O	731	61826079	GGACCATTGGAGGTGTCTGT	TGGTTTGTTTTTGCTGCTGT
1I	545	61826079	GGACCATTGGAGGTGTCTGT	AGGCTTTCCCAGATCCTCAT
2O	754	61853654	GGCTCCAGAAAGTGCAGAGT	AGATGCTCTTCGTCCACACA
2I	564	61853710	CCCAGCTTCAGGAAGTAGCA	GCTCCTTTTGGCATCACTGT
3O	910	61890274	TTCTCATTTTTATGAGATAGAGAGTCA	TTCTGGTTGTGTGGTGTTCC
3I	650	61890389	GTCTGAGACCCTGGTGGTGT	TTCCTGGGTAACAGGTGAATG
4O	653	61901112	TCCTGTCTGGACACATGCAC	TGCACTTCTGTGGAAACTGC
4I	612	61901134	CCTCCCTCAAGATCCATATCC	CCCTCCAACATGATGCAACT
5O	703	61905135	GGATGCAAGAGGGTGAGATG	TGCCTTTAAAATGCCTTCCA
5I	570	61905227	AAGGGGCATTTTGACAAGTG	ACCAAGAGTTCACCGTGGAG
6O	726	61914204	TTGAGCGAGCATCACAAAAC	ACTCATCAGAGCCAGGCATT
6I	601	61914263	TGGCTAACAGCTTGACCAGA	GACTTAGGGGGCTTTGGTTC
7O	407	61948285	GGTTGCTGCTCAATGGGTAT	TCTCATGCACACACACACAAA
7I	333	61948320	CAACATGGCTGTGTGTCAGA	TCTCATTGTCCACTTGCTCCT
8O	739	61968022	CGGGAAGCTCTTCAAGGATA	GCCAGCCTCAGAGTTTGTGT
8I	560	61968143	GGGAGGGGTTAATGGTATGC	CAGGGATTGCCCTGTAAGAA

[1]Each primer pair is numbered according to its relative position (from centromere-proximal to centromere-distal) and designated as Outer or Inner. Size of the amplicon is given along with the location on chromosome 16 of the forward primer from CanFam 2.0 (*S9*).

Table S4. Polymorphisms surrounding *CBD103*[1]

Location	Reference	Variant	W poly.	D poly.	C poly.
61826211	A	G	Yes	Yes	Yes
61826328	T	C	Yes	Yes	Yes
61826329	C	T	Yes	Yes	Yes
61826373	G	A	Yes	No	No
61826511	C	T	Yes	Yes	Yes
61853853	C	T	Yes	Yes	No
61853923	T	C	Yes	Yes	Yes
61853997	T	C	No	No	Yes
61854003	C	T	Yes	No	No
61854007	G	A	No	No	Yes
61854037	C	T	Yes	No	Yes
61854114	T	A	No	No	Yes
61854131	C	A	Yes	No	Yes
61854138	A	C	Yes	No	Yes
61854157	C	T	No	No	Yes
61854170	G	A	No	No	Yes
61854203	C	A	Yes	Yes	Yes
61890602	G	A	Yes	Yes	Yes
61890615	A	G	No	Yes	No
61890676	C	T	Yes	No	Yes
61901288	T	A	Yes	No	Yes
61901326	C	T	No	Yes	Yes
61901423	G	C	Yes	No	Yes
61901539	C	T	No	No	Yes
61901650	C	T	Yes	No	No
61901689	G	A	Yes	No	No
61905249	C	A	No	Yes	No
61905310	G	C	Yes	Yes	No
61905325	T	A	No	No	Yes
61905458	A	G	Yes	No	No
61905476	C	T	No	No	Yes
61905647	C	T	Yes	Yes	Yes
61905725	C	T	Yes	Yes	Yes
61914287	C	T	Yes	Yes	Yes
61914408	G	A	No	No	Yes
61914533	G	A	Yes	Yes	Yes
61914602	C	T	Yes	Yes	Yes
61914627	C	G	Yes	Yes	Yes
61914628	T	C	Yes	Yes	Yes
61914660	G	C	Yes	Yes	Yes
61914730	A	G	Yes	No	No
61914806	G	A	No	No	Yes
61914836	T	C	No	No	Yes
61948392	A	G	Yes	Yes	Yes
61948488	A	G	No	Yes	Yes
61948582	A	G	Yes	No	No
61968262	A	T	Yes	Yes	Yes
61968342	A	G	Yes	Yes	Yes
61968379	C	T	No	No	Yes
61968477	T	C	Yes	Yes	No
61968530	T	A	Yes	No	No
61968603	C	T	Yes	Yes	Yes

[1]Location in genome from CanFam 2.0 (*S9*) together with the reference and variant sequence, and an indication of whether a polymorphism is observed in Wolves, Dogs, and Coyotes.

Table S5. Genotypes for North American wolves on distal chromosome 16

	No. of genotypes in k^y/k^y animals[1]			No. of genotypes in K^B/k^y animals[1]		
	A/A	A/B	B/B	A/A	A/B	B/B
Coordinate[2]						
59903711	8	2	0	8	1	0
59961368	1	1	8	1	1	7
59968089	2	1	7	1	0	7
60009775	8	0	0	7	2	0
60009816	2	4	3	5	0	4
60084613	2	1	5	0	4	6
60296878	2	3	4	1	3	5
60344356	4	1	4	0	3	6
60472608	2	6	2	2	6	1
60483301	1	1	8	0	2	7
60583700	2	6	1	2	6	1
60805247	2	6	2	2	5	3
60820685	6	2	2	7	3	0
61164580	2	2	6	2	2	6
61165580	2	6	2	3	4	3
61166730	1	1	8	0	4	6
61195694	1	3	5	0	2	6
61210684	1	3	5	0	5	5
61265861	2	2	6	1	5	4
61298611	0	1	8	0	1	8
61323096	0	2	8	0	5	4
61352882	0	1	9	0	2	8
61434067	1	3	5	0	4	5
61436659	0	3	7	0	1	8
61504548	2	4	4	2	7	1
61620850	2	4	4	1	9	0
61643373	1	1	7	0	3	7
61690807	0	1	9	1	6	2
61718721	1	4	4	5	4	1
61735384	2	6	2	3	4	2
61836356	6	2	2	5	2	0
62422368	0	3	7	1	3	3

[1]Data are presented as No. of homozygotes (A/A or B/B) and heterozygotes (A/B) observed among 10 k^y/k^y and 10 K^B/k^y animals for 32 SNPs on distal chromosome 16 as described in Supporting Text. None of these SNPs exhibit a distribution similar to that observed for the *CBD103* core region (Figs. 2B, 2C) in which all K^B/k^y animals are heterozygotes and all k^y/k^y animals are homozygotes.

[2]Location in genome from CanFam 2.0 (*S9*); in this assembly, the K^B mutation is located at 61,902,782-61902785.

Supplemental Figures

Fig. S1. Nucleotide diversity (π, +/- sd) as a function of distance from *CBD103*. Dataset is analogous to that presented in Fig. 2A, and represents 22 K^B and 72 k^y haplotypes from 32 Canadian wolves and 15 Yellowstone National Park founder wolves.

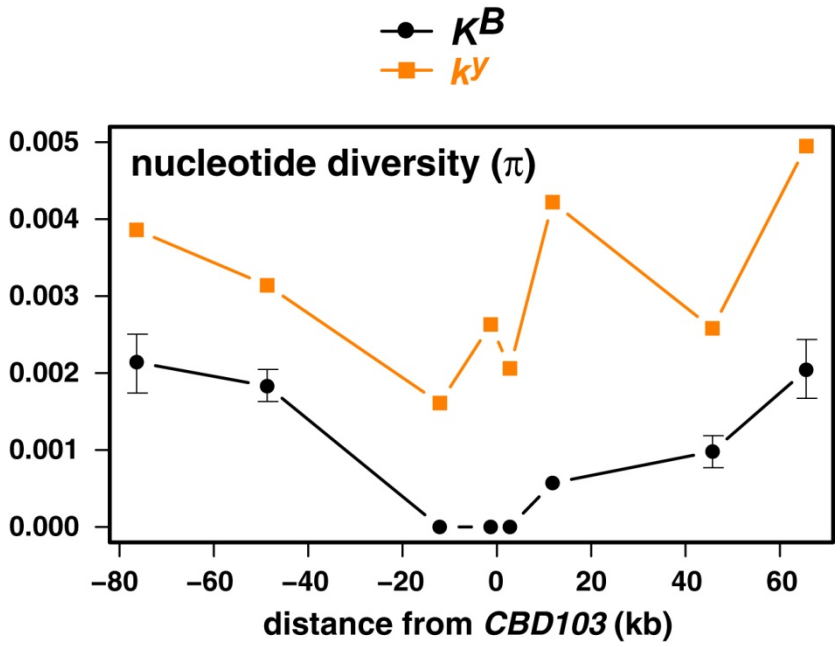

Fig. S2. Haplotypes surrounding *CBD103* in coyotes, dogs, and wolves. As in Fig. 2B (but with all canids rather than wolves alone), each row represents a K^B- or k^y-bearing chromosome, blue and yellow squares represent the major and minor alleles, respectively, gray squares represent missing data. Much greater diversity is apparent among k^y than K^B haplotypes.

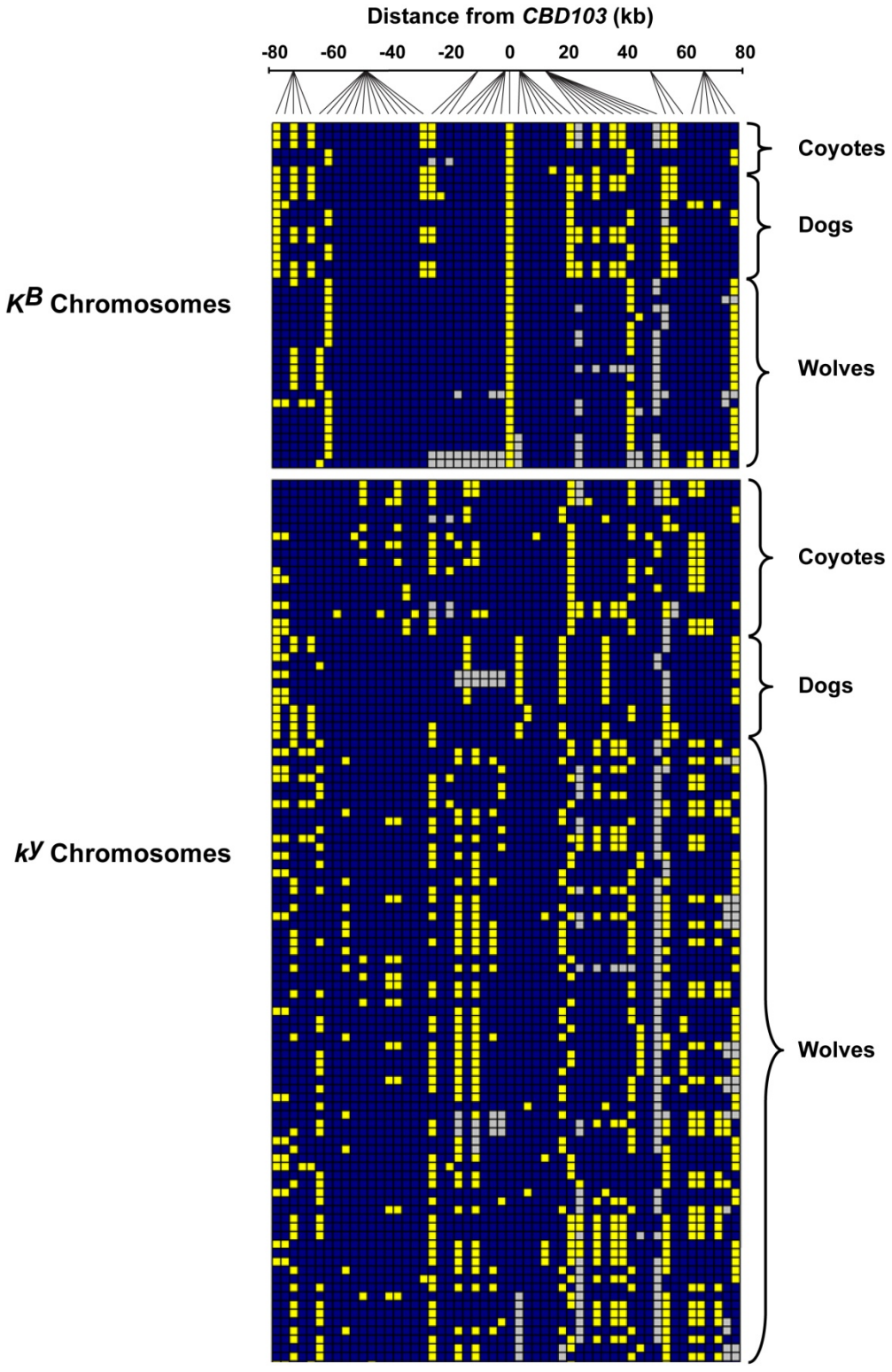
Distance from CBD103 (kb)
-80
-60
-40
-20
0
20
40
60
80
Coyotes
Dogs
Wolves
K^B Chromosomes
Coyotes
Dogs
Wolves
k^y Chromosomes

Fig. S3. Principal component analysis of wolf and dog genotypes. Genotypes for 20 wolves (5 K^B/k^y and 5 k^y/k^y from Canada, and 5 K^B/k^y and 5 k^y/k^y from Yellowstone) were analyzed by principal component analysis based on 3 progressively larger genomic regions; color (K^B/k^y vs. 5 k^y/k^y) and origin (Canada vs. Yellowstone) were then tested separately as distinguishing factors for each of the 5 first principal components. (A) p values, displayed on the ordinate as the negative logarithm, based on ANOVA; horizontal line corresponds to p = 0.05. (B) Scatterplot of individual wolves according to values for PC1 and PC2. (C) A second principal component analysis carried out on genotypes from the 20 wolves together with 20 dogs of different breeds and 5 wolf-dog hybrids from Italy; scatterplot of individuals according to values for PC1 and PC2.

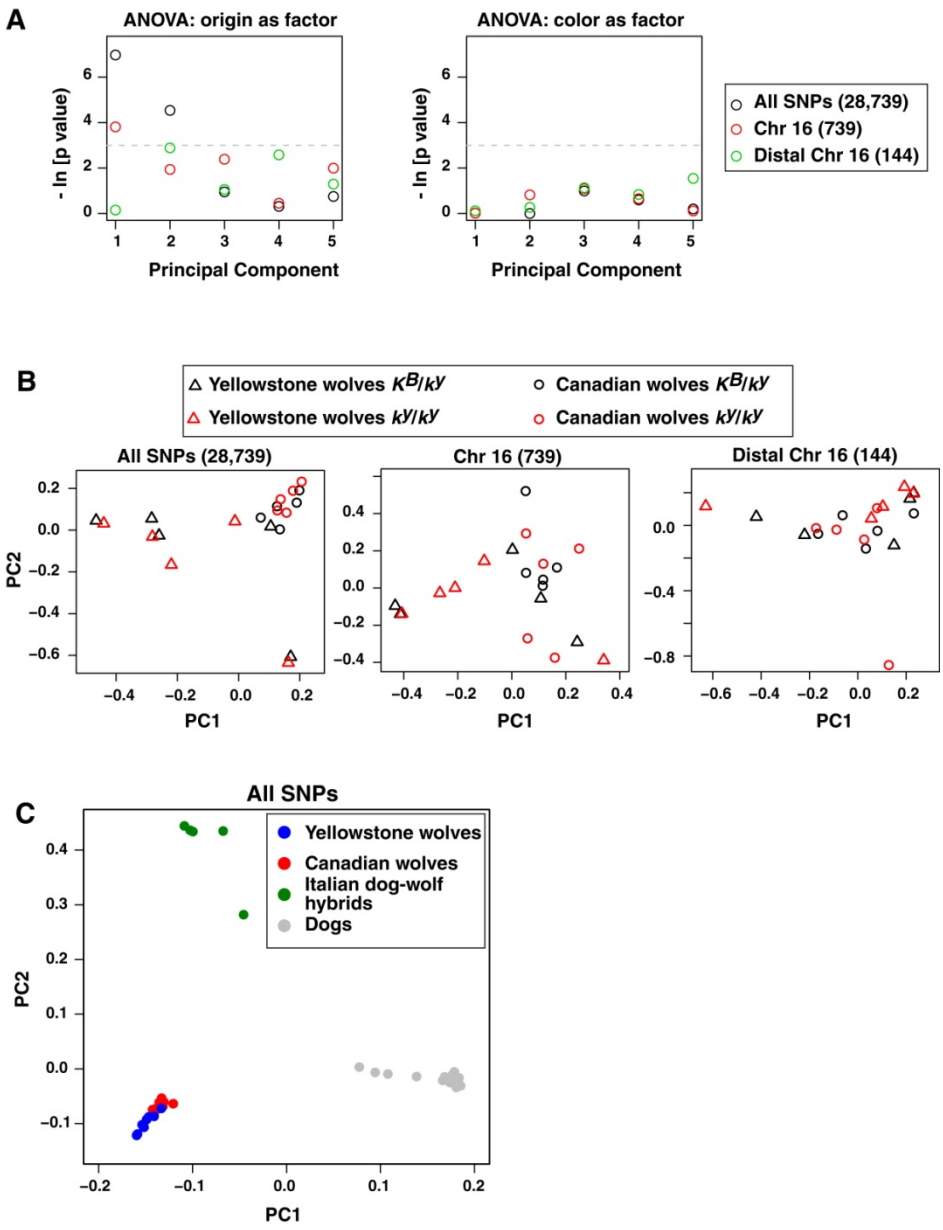
A
ANOVA: origin as factor
ANOVA: color as factor
- ln [p value)
Principal Component
All SNPs (28,739)
Chr 16 (739)
Distal Chr 16 (144)
B
Yellowstone wolves KB/kY
Yellowstone wolves kY/kY
Canadian wolves KB/kY
Canadian wolves kY/kY
All SNPs (28,739)
Chr 16 (739)
Distal Chr 16 (144)
PC1
PC2
C
All SNPs
Yellowstone wolves
Canadian wolves
Italian dog-wolf hybrids
Dogs
PC1
PC2

Fig. S4. Haplotype network for mitochondrial DNA control region from Canadian wolves and dogs. Wolf haplotypes are from (*S1*); dog haplotypes are from sequences representing the five major dog clades (Genbank AF531664, AF531729, AF531720, AF531738, and AF531741) (*S22*). Each polygon represents a haplotype whose area is proportional to its frequency; small circles are inferred haplotypes that represent individual "steps" connecting the network. The network was constructed from 37 haplotypes (14 black wolves, 18 non-black wolves, 5 dogs) with the number of black wolves indicated for each haplotype. Wolves carrying the K^B allele are not more closely related to each other than they are to wolves carrying only the k^y allele.