Multiple and Ancient Origins of the Domestic Dog

Carles Viñà, Peter Savolainen, Jesús E. Maldonado, Isabel R. Amorim, John E. Rice, Rodney L. Honeycutt, Keith A. Crandall, Joakim Lundeberg, Robert K. Wayne*

Mitochondrial DNA control region sequences were analyzed from 162 wolves at 27 localities worldwide and from 140 domestic dogs representing 67 breeds. Sequences from both dogs and wolves showed considerable diversity and supported the hypothesis that wolves were the ancestors of dogs. Most dog sequences belonged to a divergent monophyletic clade sharing no sequences with wolves. The sequence divergence within this clade suggested that dogs originated more than 100,000 years before the present. Associations of dog haplotypes with other wolf lineages indicated episodes of admixture between wolves and dogs. Repeated genetic exchange between dog and wolf populations may have been an important source of variation for artificial selection.

The archaeological record cannot resolve whether domestic dogs originated from a single wolf population or arose from multiple populations at different times (1, 2). However, circumstantial evidence suggests that dogs may have diverse origins (3). During most of the Late Pleistocene, humans and wolves coexisted over a wide geographic area (1), providing ample opportunity for independent domestication events and continued genetic exchange between wolves and dogs. The extreme phenotypic diversity of dogs, even during the early stages of domestication (1, 3, 4), also suggests a varied genetic heritage. Consequently, the genetic diversity of dogs may have been enriched by multiple founding events, possibly followed by occasional interbreeding with wild wolves.

We sequenced portions of the mitochondrial DNA of wolves and domestic dogs. Initially, 261 base pairs (bp) of the left domain of the mitochondrial control region (5) were sequenced from 140 dogs representing 67 breeds and five cross-breeds and 162 wolves representing 27 populations from throughout Europe, Asia, and North America (Fig. 1) (6). Because all wild species of the genus Canis can interbreed (7) and thus are potential ancestors of the domestic dog, five coyotes (Canis latrans) and two golden, two black-backed, and eight Simenti jackals (C. aurucaes, C. mesomelas, and C. simensis, respectively) were also sequenced.

The control region of wolves and dogs was highly polymorphic (Fig. 1). We found 27 wolf haplotypes that differed on average by 5.31 ± 0.11 (±SE) substitutions (2.10 ± 0.04%), with a maximum of 10 substitutions (3.95%). The distribution of wolf haplotypes demonstrated geographic specificity, with most localities containing haplotypes unique to a particular region (Fig. 1). Four haplotypes (W2, W7, W14, and W22) had a widespread distribution. In dogs, 26 haplotypes were found. Only haplotype D6 also occurred in
some gray wolves from western Russia and Romania (W6). Sequence divergence among dogs was similar to that found among wolves. Dog haplotypes differed by an average of 5.30 ± 0.17 substitutions (2.06 ± 0.07%), with a maximum divergence of 12 substitutions (4.67%). Mitochondrial haplotype diversity in dogs could not be partitioned according to breeds. For example, in eight German shepherds examined, five distinct sequences were found, and in six golden retrievers, four sequences were detected. Moreover, many breeds shared sequences with other breeds. For instance, dog haplotypes D4, D3, D5, and D1 were found in 14, 14, 9, and 7 breeds, respectively. No dog sequence differed from any wolf sequence by more than 12 substitutions, whereas dogs differed from coyotes and jackals by at least 20 substitutions and two insertions. These results clearly support wolf ancestry for dogs. However, because mitochondrial DNA is maternally inherited, interbreeding between female dogs and male coyotes or jackals would not be detected. More limited studies of nuclear markers support the conclusion that the wolf was the ancestor of the domestic dog (8).

Several methods of phylogenetic analysis, including maximum likelihood (9), maximum parsimony (10), minimum spanning networks (11), and statistical parsimony (12), were used to investigate relationships among sequences. All analyses supported a grouping of dog haplotypes into four distinct clades, although the topology within and among clades differed among trees (13). As exemplified by the neighbor-joining analysis (Fig. 2A), three of the four monophyletic clades defined a larger clade containing all but three dog

**Fig. 1.** Substitutions and deletions (−) observed in 261 bp of control region sequence from wolves (W) and dogs (D). The dog sequence D13 had the same sequence as D4 except for an insertion of a 67-bp tandem repeat. The numerals I, II, III, and IV indicate assignments to the four clades of dog sequences. Wolf localities: Bulgaria (n = 1, W7); Croatia (n = 5, W2); Estonia (n = 1, W10; France (n = 2, W4); Finland (n = 2, W10); Greece (n = 7; W2, W5, W8, and W9); Italy (n = 12, W4); Poland (n = 1, W3); Portugal (n = 19; W1 and W2); Romania (n = 4; W5 and W6); Russia (n = 3; W6, W10, and W26); Spain (n = 46; W1 and W3); Sweden (n = 2; W2 and W10); Afghanistan (n = 3, W19); China (n = 3; W4, W19, and W27); India (n = 1, W12); Iran (n = 6; W16 and W17); Israel (n = 16, W11); Saudi Arabia (n = 7; W7, W12, W13, W14, and W15); Turkey (n = 2, W2; Alaska (n = 3, W20); Alberta (n = 1, W22); Labrador (n = 3, W22); Mexico (n = 5, W25); Montana (n = 1, W22); Northwest Territories (n = 3, W22); and Yukon (n = 3; W21, W23, and W24). Dog breeds: basenji (n = 1, D2); bassett (n = 1, D6); boxer (n = 1, D4); Norwegian buhund (n = 1, D11); bulldog (n = 1, D6); Chinese crested (n = 2; D2 and D25); chow chow (n = 3; D1, D2, and D3); collie (n = 1, D1); border collie (n = 3; D1 and D5); winelared dachshund (n = 3; D5 and D10); Australian dingo (n = 4, D18); grey Norwegian skihound (n = 9; D3 and D8); Eskimo dog (n = 1, D23); German shepherd (n = 8; D4, D5, D6, and D19); greyhound (n = 1, D9); groenendael (n = 1, D6); Mexican hairless (n = 6; D3, D6, D21, and D26); hamilton-stövare (n = 1, D5); Afghan hound (n = 3, D6); Alaskan husky (n = 2; D4 and D7); Siberian husky (n = 3, D3, D7, and D18); jähund (n = 3; D7 and D8); kaeshond (n = 1, D5); kuvasz (n = 1, D4); Leonberger (n = 2; D1 and D4); Norwegian lundehund (n = 1, D16); Maremman (n = 1, D6); Pyrenean mastiff (n = 1, D11); Newfoundland (n = 1, D4); otter hound (n = 1, D6); papillon (n = 2; D3 and D4); poodle (n = 1, D3); toy poodle (n = 1, D6); pug (n = 1, D26); Chesapeake Bay retriever (n = 1, D13); flat-coated retriever (n = 3; D4 and D10); golden retriever (n = 6; D4, D6, D15, and D24); labrador retriever (n = 6; D4 and D12); Rhodesian ridgeback (n = 1, D26); rottweiler (n = 2, D3); Samoyede (n = 3; D1, D4, and D5); St. Bernard (n = 1, D9); schipperke (n = 1, D4); giant schnauzer (n = 1, D9); English setter (n = 4; D3 and D5); Irish setter (n = 3; D1 and D9); New Guinea singing dog (n = 2, D18); shar pei (n = 1, D26); Icelandic sheepdog (n = 1, D3); Old English sheepdog (n = 1, D3); shiba inu (n = 1, D20); Cavalier King Charles spaniel (n = 1, D17); Irish water spaniel (n = 1, D6); springer spaniel (n = 1, D3); Tibetan spaniel (n = 1, D1); airedale terrier (n = 1, D7); border terrier (n = 2, D3); fox terrier (n = 2; D3 and D14); Norfolk terrier (n = 2, D4); West Highland terrier (n = 2, D7); Tibetan terrier (n = 2; D2 and D9); wachtelhund (n = 1, D5); whippet (n = 1, D3); Irish wolfhound (n = 2, D11); and crossbreeds (n = 5; D1, D3, D4, D5, and D18).
haplotypes and a subset of wolf haplotypes (W4 and W5). Clade I included 19 of the 26 dog haplotypes. This group contained representatives of many common breeds as well as ancient breeds such as the dingo, New Guinea singing dog, African basenji, and greyhound (14). Clade II included dog haplotype D8, from two Scandinavian breeds (elkhound and jämtlund), and was closely related to two wolf haplotypes found in Italy, France, Romania, and Greece (W4 and W5). Clade III contained three dog haplotypes (D7, D19, and D21) found in a variety of breeds such as the German shepherd, Siberian husky, and Mexican hairless. Finally, clade IV contained three haplotypes (D6, D10, and D24) that were identical or very similar to a wolf haplotype (W6) found in Romania and western Russia, which suggests recent hybridization between dogs and wolves. Many breeds contained representatives of more than one dog haplotype grouping (Fig. 1).

Because the overall bootstrap support for many of the internodes in Fig. 2A was low, 1030 bp of the control region were sequenced for 24 canids, including representatives of the four major clades (15). Although the association of clades was different, the analyses of the longer sequences provided stronger support for the four monophyletic groupings of dog haplotypes (Fig. 2B) (13). A Wilcoxon signed-rank test was used to assess the monophyly of dog clades (16). Monophyly of all dog haplotypes can be rejected, and monophyly of clades I, II, and III is marginally rejected (P = 0.0004 and P = 0.053, respectively). In both trees, dog clade II and IV are most closely related to wolf sequences from eastern Europe (Greece, Italy, Romania, and western Russia).

The coyote and wolf have a sequence divergence of 0.075 ± 0.002 (17) and diverged about one million years ago, as estimated from the fossil record (18). Consequently, because the sequence divergence between the most different genotypes in clade I (the most diverse group of dog sequences) is no more than 0.010, this implies that dogs could have originated as much as 135,000 years ago (19). Although such estimates may be inflated by unobserved multiple substitutions at hypervariable sites (20), the sequence divergence within clade I clearly implies an origin more ancient than the 14,000 years before the present suggested by the archaeological record (21). Nevertheless, bones of wolves have been found in association with those of hominids from as early as the middle Pleistocene, up to 400,000 years ago (1, 22). The ancient dates for domestication based on the control region sequences cannot be explained by the retention of ancestral wolf lineages, because clade I is exclusively monophyletic with respect to dog sequences and thus the separation between dogs and wolves has been long enough for coalescence to have occurred. To explain the discrepancy in dates, we hypothesize that early domestic dogs may not have been morphologically distinct from their wild relatives. Conceivably, the change around 10,000 to 15,000 years ago from nomadic hunter-gatherer societies to more sedentary agricultural population centers may have imposed new selective regimes on dogs that resulted in marked phenotypic divergence from wild wolves (23).

Although individual breeds show uniformity with respect to behavior and morphology, most breeds show evidence of a genetically diverse heritage because they contain different haplotypes. Moreover, dog sequences cluster with different groups of wolf haplotypes. Therefore, after the origin of dogs from a wolf ancestor, dogs and wolves may have continued to exchange genes. Backcrossing events could have provided part of the raw material for artificial selection and for the extraordinary degree of phenotypic diversity in the domestic dog. Domestic species of plants and animals whose wild progenitors are extinct cannot be enriched through periodic interbreeding, and change under artificial selection may be more limited. Consequently, the preservation of wild progenitors may be a critical issue in the continued evolution of domestic plants and animals.

REFERENCES AND NOTES
7. A. P. Gray, Mammalian Hybrids, a Check-list with Bibliography (Farnham Royal, Bucks, Commonwealth Agricultural Bureaux, UK, 1954).
13. The coyote was used as an outgroup on the basis of morphological and mitochondrial DNA sequence analyses (9). M. Nowak, North American Quaternary Canis (Monograph of the Museum of Natural History, University of Kansas, N. Lawrence, KS, 1979); R. K. Wayne, Trends Genet. 6, 218 (1990); C. Villà et al., unpublished data. Trees retaining the four indicated clades resulted from neighbor-joining searches with Tamura-Nei, Kimura 2–parameter, Kimura 3–parameter, and HK85 models of evolution; parsimony searches with equal, 2:1, and 10:1 weights of transversions to transitions; and maximum likelihood analyses with the HK85 model of evolution, taking into account differences in nucleotide frequencies and transition and transversion biases.
15. We amplified the entire canine control region with primers in (6) and the following primers: L16462 (5’-CATACATAACCTGCGGCGTTAC-3’), H222 (5’-AACTATATGTCTGTCAGGAAC-3‘), L296 (5’-ATA- CAAACCCCCCTAACCT-3‘), and H652 (5’-AACGATGGACCAAACCCTG-3‘). The primer numbering refers to their approximate location in the human mitochondrial genome (9). The 1030-bp sequence that we report includes the entire control region except for a hypervariable region close to the 3’ end that encompasses an imperfect 20-bp repeat having a variable number of units [A. R. Hoelzel, J. V. Lopez, G. A. Dover, S. J. O’Brien, J. Mol. Evol. 39, 191 (1995)].
17. Sequence divergence measures and tests were based on a gamma correction [K. Tamura and M. Nei, Mol. Biol. Evol. 10, 513 (1993)] with the empirically determined parameter α = 0.18 for the 261-bp sequences and α = 0.22 for the 1030-bp sequences [S. Kumar, K. Tamura, M. Nei, MEGA: Molecular Evolutionary Genetic Analysis, Version 1.07 (Pennsylvania State University, University Park, PA, 1993)].
19. A relative rate test indicated that substitution rates were similar for all wolf and dog sequences (9). T. A. Parson et al., Nature Genet. 15, 369 (1997).
23. We thank P. Taberlet for providing wolf sequences from France and Poland and J. Castrénio, Grupo Lobo-Portugal, E. Geffen, and many others for their help in obtaining samples. C.V. was supported by a postdoctoral fellowship from the Spanish Ministerio de Educación y Ciencia, and I.R.A. was supported in part by NSF grants to R.L.H. (DEB-9306022) and R.K.W. (DEB-9309293) and by a Sloan Young Investigator Award to K.A.C.