Relaxation of selective constraint on dog mitochondrial DNA following domestication

Susanne Björnerfeldt, Matthew T. Webster and Carles Vilà

Genome Res. 2006 16: 990-994
Access the most recent version at doi:10.1101/gr.5117706

Letter

Relaxation of selective constraint on dog mitochondrial DNA following domestication

Susanne Björnerfeldt,[1] Matthew T. Webster,[1,2] and Carles Vilà[3]

Department of Evolutionary Biology, Uppsala University, SE-752 36 Uppsala, Sweden

The domestication of dogs caused a dramatic change in their way of life compared with that of their ancestor, the gray wolf. We hypothesize that this new life style changed the selective forces that acted upon the species, which in turn had an effect on the dog's genome. We sequenced the complete mitochondrial DNA genome in 14 dogs, six wolves, and three coyotes. Here we show that dogs have accumulated nonsynonymous changes in mitochondrial genes at a faster rate than wolves, leading to elevated levels of variation in their proteins. This suggests that a major consequence of domestication in dogs was a general relaxation of selective constraint on their mitochondrial genome. If this change also affected other parts of the dog genome, it could have facilitated the generation of novel functional genetic diversity. This diversity could thus have contributed raw material upon which artificial selection has shaped modern breeds and may therefore be an important source of the extreme phenotypic variation present in modern-day dogs.

[Supplemental material is available online at www.genome.org. The sequence data from this study have been submitted to GenBank under accession nos. DQ480489–DQ480511.]

In *The Origin of Species*, Darwin (1859) suggested that "several wild species of Canidae have been tamed and that their blood, in some cases mingled together, flows in the veins of our domestic [dog] breeds". We now know that dogs (*Canis familiaris*) are entirely derived from the domestication of wolves (*Canis lupus*) (Vilà et al. 1997); however, the origin of the huge morphological diversity that led Darwin to his speculation remains largely unknown (Sutter and Ostrander 2004). The domestic dog is the most phenotypically diverse mammal on earth. The large differences in size, conformation, behavior, and physiology between dog breeds exceed the differences among species in the dog family, Canidae (Coppinger and Coppinger 2001; Wayne 2001). Recent studies show that the origin of most dog breeds may derive from very recent selective breeding practices and are probably <200 yr old (Parker et al. 2004). However, selection acts upon existing variability. It is remarkable that the potential for such large diversification existed in the ancestral wolf population from where the domestication process was initiated. Furthermore, the time since domestication (at least 14,000 yr; Vilà et al. 1997; Sablin and Khlopachev 2002; Savolainen et al. 2002) seems insufficient to generate substantial additional genetic diversity. What is the origin of this diversity? We hypothesize that changes in the living conditions of dogs as a result of domestication resulted in the release of selective constraint allowing a faster accumulation of functional (non-silent) genetic diversity in a large array of genes.

Weakly deleterious mutations—those with selective effects close to the reciprocal of the effective population size—represent an important class of genetic variability (Ohta and Kimura 1971). Such mutations are expected to accumulate faster in populations with small effective sizes or in populations in which selection has been relaxed, resulting in a decline in fitness. Advantageous mutations, conversely, contribute little to patterns of genetic variation and are enriched in fixed differences between species. To examine whether the accumulation of deleterious mutations is increased in dogs compared with their wild ancestors, we have focused on mitochondrial DNA (mtDNA). The mitochondrial genome represents only a small part of the canine genome and has a unique mode of inheritance. However, while dog and wolf lineages are difficult to separate for nuclear genes (Parker et al. 2004; Vilà et al. 2005), mitochondrial lineages are clearly distinguishable for the two species (Vilà et al. 1997; Savolainen et al. 2002). This offers a good opportunity to evaluate the consequences of life with humans on a portion of the dog genome.

As is often seen in data sets where recombination is rare or absent, mtDNA commonly exhibits an excess of replacement (nonsynonymous) to silent (synonymous) changes in intraspecies polymorphism compared with inter-species divergence (Hasegawa et al. 1998). These findings likely reflect an increased preponderance of deleterious alleles segregating within populations. Among the dog mitochondrial genes, 13 encode different proteins of oxidative phosphorylation, which are important for the production of energy and heat. Additionally, the mtDNA includes two ribosomal RNA genes and 22 transfer RNA genes. In humans, large numbers of pathological mutations have been described in mtDNA genes (Chinnery et al. 2000). Studies on human mtDNA have shown that selection may have shaped the patterns of variability observed today (Ruiz-Pesini et al. 2004; Kivisild et al. 2006). In the case of the dog, since wolves still survive across the world, it is possible to compare the genome of the domestic species with that of its ancestor and thus identify both functional and silent nucleotide changes that appeared in the domestic lineages. In this study we compare patterns of molecular evolution in the mtDNA of dogs and wolves in order to understand the genetic consequences of the change in lifestyle associated with domestication.

[1]These authors contributed equally to this work.
[2]Present address: Smurfit Institute of Genetics, University of Dublin, Trinity College, Dublin 2, Ireland.
[3]Corresponding author.
E-mail carles.vila@ebc.uu.se; fax. +46-18-471-6310.
Article published online before print. Article and publication date are at http://www.genome.org/cgi/doi/10.1101/gr.5117706.

16:990–994 ©2006 by Cold Spring Harbor Laboratory Press; ISSN 1088-9051/06; www.genome.org

Results and Discussion

Previous studies have shown that domestic dog mtDNA sequences cluster in four main clades when compared with wolves, indicating different origination events (Vilà et al. 1997; Savolainen et al. 2002). In order to select samples of dogs representing several mtDNA lineages for the analysis, we sequenced the mitochondrial control region I for 88 dogs from 53 breeds. Among those individuals we selected 14 dogs, which included six from clade I (the clade that encompasses about 71% of today's dogs; Savolainen et al. 2002) and two or three from each one of the clades II, III, and IV (Vilà et al. 1997). Because we wished to characterize mutations that occurred on dog lineages since the emergence of each clade, the dogs in this study were selected to be representative of the full genetic diversity observed in each clade (Supplemental Fig. S1). Complete mtDNA sequences, excluding the tandem repeat located inside the control region (Hoelzel et al. 1994), were obtained through polymerase chain reaction (PCR) amplification and sequencing. The complete mitochondrial sequence was also obtained for six wolves from throughout the world trying to represent as much of the previously described wolf diversity (Vilà et al. 1999) as possible: Spain, Russia, Saudi Arabia (two individuals), North America, and Sweden. Three coyotes (*Canis latrans*) from Nebraska and Colorado (two individuals), USA, were also sequenced and used as outgroups.

To construct a gene tree from the 23 complete mtDNA sequences, we first excluded the control region because of the high incidence of homoplasy (Ingman et al. 2000), resulting in a sequence length of 15,547–15,549 base pairs (bp). The average uncorrected pairwise sequence divergence between wolves and dogs was 0.47% (SE = 0.02), whereas average sequence divergence between coyotes and dogs plus wolves was 4.28% (SE = 0.11). A gene tree constructed with these sequences shows that all four clades of dogs are very well supported with bootstrap support values of 100% and Bayesian posterior probabilities of 1.00 (Fig. 1).

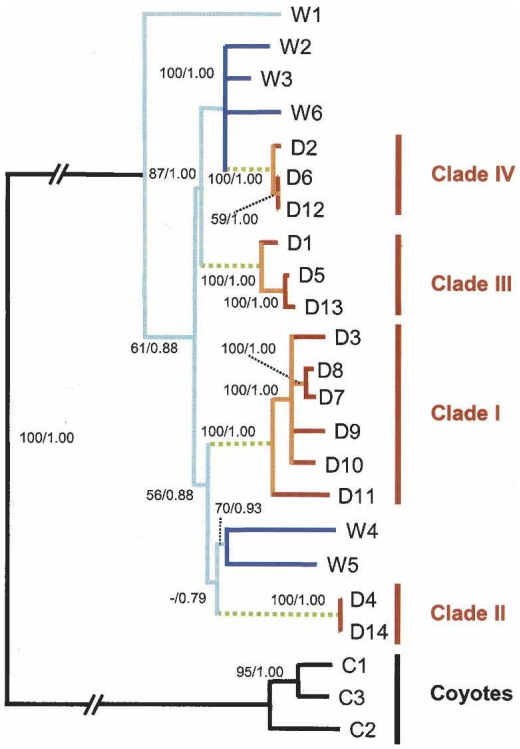

Figure 1. Phylogenetic tree of wolf (W), dog (D), and coyote (C) mtDNA sequences. The tree was constructed using a Bayesian approach. The same topology was obtained with a neighbor-joining approach. Support is indicated at the nodes as percent bootstrap support for 1000 neighbor-joining replicates and Bayesian posterior probabilities. Four clades of dog sequences (I to IV) are indicated as in Vilà et al. (1997). Internal dog branches are marked in orange, and internal wolf branches are marked in light blue. The branch leading to wolf haplotype W1 was basal to the rest of the tree and it was also considered internal. Internal branches that could not be conclusively associated to dogs or to wolves are indicated in discontinuous green.

We used a maximum-likelihood (ML) approach to estimate the rates of synonymous (d_S) and nonsynonymous (d_N) evolution in mtDNA genes along each individual branch of the gene tree in Figure 1. Branches leading only to one or more dog sequences were considered to be dog branches (in red and orange in Fig. 1) except for the four branches preceding the four clades (dotted lines). These branches were excluded from our analyses because they could not be uniquely assigned to dogs or to wolves. All other branches within the wolf/dog tree (in blue) were considered as wolf branches as they either represent evolution before dog domestication or recent wolf evolution. We first examined the tree for differences in the rate of silent nucleotide substitution between wolves and dogs. We used the ML estimates of d_S to perform relative rate tests at the four nodes of the tree associated with the origin of the four dog clades (see Methods). None of the tests showed significant differences between dog and wolf branches ($P > 0.05$ in all cases, data not shown), although d_S was on average 10.9% higher on lineages leading to dogs compared with wolves. We therefore have no evidence to suggest that the mutation rate in mtDNA differs between dogs and wolves. A similar test revealed that d_N was on average 40.3% higher on the dog lineages, although none of the individual tests were statistically significant.

We next examined ML estimates of d_N/d_S ratios in all branches of the gene tree. The average d_N/d_S ratio of divergence between a randomly chosen wolf (W5) and a randomly chosen coyote (C2) sequence (0.034; 95% confidence interval CI: 0.023–0.043) is significantly lower than the average ratio along wolf branches (0.091; 95% CI: 0.056–0.127; $P < 0.001$; confidence intervals and significance testing were calculated by bootstrap), which reflects intraspecific variation. Weakly deleterious mutations are expected to be more common in intra-specific variation than in divergence between species because purifying selection has had less time to act (Akashi 1999; Piganeau and Eyre-Walker 2003; Kivisild et al. 2006). This suggests that many weakly deleterious mutations are segregating within the wolf population. However, the average d_N/d_S ratio in dog branches (0.183; 95% CI: 0.096–0.279) is significantly higher than for wolves ($P = 0.040$). This result is especially surprising considering that selection is more effective in growing populations, increasing the probability that deleterious alleles are lost (Otto and Whitlock 1997), and that the domestic population has likely increased by more than six orders of magnitude since the time of domestication (Coppinger and Coppinger 2001). We believe that this could indicate a relaxation of constraint on dog mtDNA genes compared with wolves.

Within populations, the relative number of deleterious compared with neutral changes is expected to decline as a func-

tion of allele frequency as a result of purifying selection (Fay et al. 2001). This is because deleterious alleles tend to be removed from the population by selection and hence are unlikely to reach high frequencies. This implies that d_N/d_S is expected to be higher on terminal branches than internal branches of the wolf/dog portion of the gene tree, which all represent intra-specific variability. As a higher proportion of dog branches are terminal in the gene tree, this could potentially explain the higher d_N/d_S ratios in dog branches. In order to investigate this possibility, we divided dog and wolf branches into internal (orange for dogs and light blue for wolves, in Fig. 1) and terminal (red and dark blue) branches. The d_N/d_S ratios are surprisingly consistent within dog and within wolf branches (Fig. 2). This indicates that the effect of allele frequency on the proportion of nonsynonymous mutations has not been that great and that the differences between d_N/d_S ratios in wolves and dogs represent true differences in selective regimes, rather than simply being an effect of the accumulation of weakly deleterious mutations on terminal branches of the gene tree. It should however be noted that dog branches tend to be more distal in the tree compared with wolf branches, which could potentially contribute to the larger d_N/d_S in dogs.

We compared the d_N/d_S ratio along branches representing the divergence between coyotes and wolves, dog diversity, and wolf diversity for individual mitochondrial genes (Supplemental Fig. S2). We also estimated d_N/d_S ratios for three gene classes: ATPase genes, NADH dehydrogenase (complex I) genes, and cytochrome c oxidase (complex IV) genes (Supplemental Fig. S2). We did not note any trend for the observed differences in d_N/d_S ratios between branches to be biased toward any particular gene or gene class although the number of changes in individual genes is small and the confidence intervals are very large. Hence our data are compatible with a model whereby dogs are gradually accumulating weakly deleterious changes across all mtDNA genes.

In order to assign nucleotide changes at each gene to specific branches and investigate their biochemical properties, we compared the maximum-likelihood reconstructed ancestral sequences at each node with those at neighboring nodes. Table 1 shows the number of changes estimated for wolf and dog branches in the tree in Figure 1. These values corroborate the findings of the ML estimation of d_N/d_S ratios. There is a significant excess of nonsynonymous (NS) changes compared with synonymous (S) along both wolf and dog branches compared with divergence between wolves and coyotes ($P = 0.002$ and $P < 0.001$, respectively, G-test of independence). Additionally, a significant excess of nonsynonymous changes is also observed in dog compared with wolf lineages ($P = 0.033$). In line with our previous findings there is no difference between the ratio of synonymous and nonsynonymous changes between internal and terminal branches within either dogs or wolves ($P > 0.05$ in both cases). We used three different methods to examine the potential phenotypic effects of mutations: We divided changes into conservative or radical by charge and by polarity (Zhang 2000) and into benign, potentially damaging, or probably damaging using the PolyPhen database (see Methods; Table 1). The relative numbers of conservative to radical (or benign to damaging) changes are not significantly different between dog branches, wolf branches, or in coyote/wolf divergence (G-tests, $P > 0.05$ in all cases). Hence there appears to be no tendency for dogs to accumulate more radical or damaging changes than wolves, or for such changes to have been preferentially eliminated from the wolf or dog population. This is probably because the majority of radical or damaging changes are strongly deleterious and never reach detectable frequencies in any population. This is consistent with our hypothesis that a relaxation of selective constraint in dogs has resulted in an elevated accumulation of weakly deleterious variants.

Although the domestication process was likely initiated by just a few individuals (Vilà et al. 1997, 2005; Savolainen et al. 2002), the total world population of dogs is today estimated to be around 400 million (Coppinger and Coppinger 2001). As the initial dog population was small and was subsequently subdivided, deleterious mutations may have accumulated by genetic drift. In addition, as soon as the first dogs started to live with humans, it is likely that they were strongly selected for behavioral traits like tameness (Saetre et al. 2004). As the dogs' breeding program was controlled by humans, the intensity of stabilizing selection for other morphologic, behavioral, or physiological traits likely decreased. Individuals with lower metabolic efficiency were more likely to survive and reproduce than they were before. This relaxation of constraint may have allowed the accumulation of additional nonsynonymous mutations on the mitochondrial genome. It is therefore possible that this process led to an increase in functional genetic diversity throughout the entire dog genome, including both genes and elements affecting gene expression. For example, it has been suggested that variation at tandem repeats (Fondon III and Garner 2004) or the presence of short interspersed elements (SINEs; Wang and Kirkness 2005) closely associated with genes could have contributed to the phenotypic diversity in dogs, although there is no direct evidence of these elements being more frequent in dogs than in gray wolves. A relaxation of selective constraint could have contributed not only to the huge phenotypic diversity that exists in today's dogs but also to the appearance of

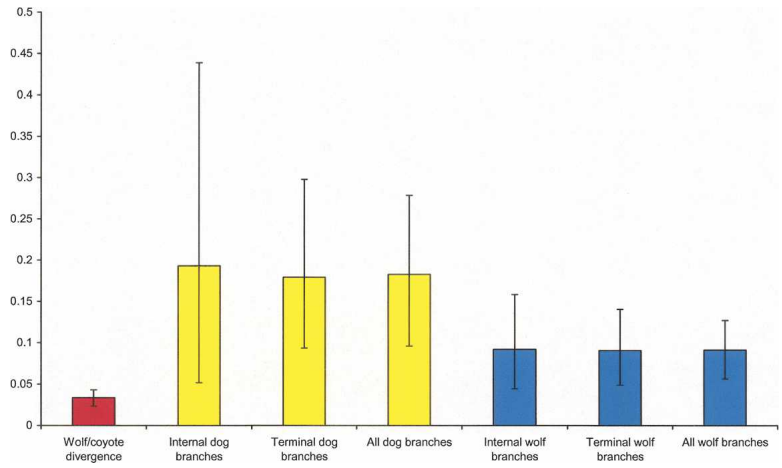

Figure 2. Maximum-likelihood estimates of d_N/d_S. Values for wolves are indicated in blue, for dogs in yellow, and for the divergence between a random wolf and coyote sequence in red. Ninety-five percent confidence intervals were bootstrap-derived.

Table 1. Number of nucleotide changes inferred from maximum-likelihood reconstructed ancestral nodes

	Wolf/coyote divergence	Internal wolf branches	Terminal wolf branches	All wolf branches	Internal dog branches	Terminal dog branches	All dog branches
RNA genes	90	9	14	23	2	10	12
Control region	62	11	30	41	4	14	18
Other noncoding	5	0	1	1	0	1	1
Synonymous (S)	546	63	83	146	14	41	55
Nonsynonymous (NS)	48	13	16	29	6	16	22
NS/S ratio	0.09	0.21	0.19	0.20	0.43	0.39	0.40
Nonsynonymous changes							
By charge							
Conservative	41	13	13	26	6	15	21
Radical	6	0	3	3	0	1	1
By polarity							
Conservative	33	9	9	18	3	10	13
Radical	14	4	7	11	3	6	9
PolyPhen							
Benign	42	11	16	27	4	16	20
Possibly damaging	2	1	0	1	1	0	1
Probably damaging	2	0	0	0	0	1	1

Wolf/coyote divergence refers to nucleotide changes between one random wolf and one random coyote sequence in Figure 1. Internal and terminal branches for wolves and for dogs are indicated in Figure 1.

the large variety of diseases that affect many dog breeds (Ostrander and Kruglyak 2000).

Methods

Sampling strategy and mtDNA sequencing

Blood samples were obtained from the Swedish University of Agricultural Sciences at Uppsala, Sweden. Samples from 88 dogs corresponding to 53 different breeds were initially screened to identify dogs with mitochondrial DNA sequences pertaining to the four clades described by Vilà et al. (1997). Fourteen of these dogs from 13 different breeds were selected to represent all four clades (Supplemental Table S1). Additionally, tissue samples were obtained from six wolves and three coyotes. The wolves originated from Spain, Russia, Saudi Arabia, North America, and Sweden. Two samples were obtained from wolves from Saudi Arabia due to the large mtDNA diversity revealed by a previous study (Vilà et al. 1999). The coyote samples came from Nebraska and Colorado (two individuals), USA. DNA extraction, amplification, and sequencing are described in the Supplemental material.

Tree construction

We constructed a phylogenetic tree of all complete mtDNA sequences except the control region using a Bayesian approach as implemented in MrBayes 3 (Huelsenbeck and Ronquist 2001). The same tree topology was obtained using a neighbor-joining approach in PAUP 4.0 (Swofford 2002) using the model of sequence evolution that best fitted the data according to the program Modeltest 3.6 (Posada and Crandall 1998). Support of nodes was calculated using posterior probabilities for the Bayesian tree and with 1000 bootstrap pseudoreplicates for the neighbor-joining tree.

In order to distinguish between mutations along different branches of the gene tree, we classified wolf/dog branches as wolf internal, wolf terminal, dog internal, and dog terminal. The internal branches leading to each of the four dog clades were excluded from the analysis (Fig. 1) because they could not be conclusively assigned to dogs or to wolves. Internal wolf branches were defined as internal branches leading to only wolves or both wolves and dogs. Since the branch leading to wolf haplotype W1 was basal to the rest of the tree, it was also considered internal. Internal dog branches were defined as internal branches leading only to dog sequences. Terminal wolf and dog branches lead directly to a wolf or dog sequence, respectively.

Maximum-likelihood estimation of evolutionary rates

We used the PAML package (Yang 1997) to reconstruct the changes that occurred along each branch of the gene tree at all of the site classes within the mtDNA molecule. For sites outside of protein coding genes, we used the BaseML program to provide ML estimates of ancestral nucleotide states at each node of the tree and the number of changes along each branch using the Hasegawa et al. (1985) model of base substitution. These sites were classified as RNA genes, control region (excluding tandem repeat), and other noncoding regions. We used CodeML, implementing the codon substitution model of Goldman and Yang (1994) to provide ML estimates of the ancestral codons at each node in the tree and the nonsynonymous (d_N) and synonymous (d_S) substitution rates along each branch. The free-ratio model, where d_N/d_S is allowed to vary among branches, was used. Codon frequencies were calculated from the average nucleotide frequencies at the three codon positions.

We tested whether the rate of molecular evolution differs between dogs and wolves using an extension of the relative rate test (Sarich and Wilson 1973) based on lineage-specific ML estimates of d_N and d_S. This was done by first choosing the four nodes in the gene tree leading to the branches before each of the four dog clades (discontinuous green branches in Fig. 1). For each node, we then compared the average d_N and d_S values from the node to the tip of a dog branch in the clade in question with the average d_N and d_S values from the node to the tip of a wolf branch. We also estimated the average d_N/d_S values of the internal and external branch categories described in the previous section using ML estimates of d_N/d_S from each individual branch of the tree. For both procedures, generation of confidence intervals and estimation of the significance of differences between statistics on different branch categories were performed by nonparametric bootstrapping. We recalculated each statistic from data sets produced by resampling each codon with replacement with 1000 replicates (Felsenstein 1985). When comparing statistics for each individual gene, we resampled each gene separately.

We assigned each inferred amino acid change to a specific

branch in the gene tree by comparison of ML inferred sequences at each node. Nonsynonymous changes on each branch in the tree were categorized as radical or conservative by the criteria of both polarity and charge presented by Zhang (2000) and by using the PolyPhen application, which incorporates sequence, phylogenetic, and structural information to characterize the potential severity of the phenotypic effect produced by each mutation.

Acknowledgments

Åke Hedhammar, from the Department of Small Animal Clinical Sciences, Swedish University of Agricultural Sciences in Uppsala, Sweden, helped us to obtain samples from dogs. Ron Ruthstrom, from the Kansas Department of Wildlife and Parks, USA, and Jennifer Leonard, from the Department of Evolutionary Biology, Uppsala University, provided the coyote samples. Frank Hailer and Jennifer Leonard provided valuable comments and discussion. Financial support was provided by a grant of the Swedish Research Council for Environment, Agricultural Sciences and Spatial Planning to CV.

References

Akashi, H. 1999. Inferring the fitness effects of DNA mutations from polymorphism and divergence data: Statistical power to detect directional selection under stationarity and free recombination. *Genetics* **151:** 221–238.

Chinnery, P.F., Johnson, M.A., Wardell, T.M., Singh-Kler, R., Hayes, C., Brown, D.T., Taylor, R.W., Bindoff, L.A., and Turnbull, D.M. 2000. The epidemiology of pathogenic mitochondrial DNA mutations. *Ann. Neurol.* **48:** 188–193.

Coppinger, R. and Coppinger, L. 2001. *Dogs*. The University of Chicago Press, Chicago.

Darwin, C. 1859. *On the origin of species by means of natural selection*. Murray, London.

Fay, J.C., Wyckoff, G.J., and Wu, C.I. 2001. Positive and negative selection on the human genome. *Genetics* **158:** 1227–1234.

Felsenstein, J. 1985. Confidence limits on phylogenies: An approach using the bootstrap. *Evolution Int. J. Org. Evolution* **39:** 783–791.

Fondon III, J.W. and Garner, H.R. 2004. Molecular origins of rapid and continuous morphological evolution. *Proc. Natl. Acad. Sci.* **101:** 18058–18063.

Goldman, N. and Yang, Z. 1994. A codon-based model of nucleotide substitution for protein-coding DNA sequences. *Mol. Biol. Evol.* **11:** 725–736.

Hasegawa, M., Kishino, H., and Yano, T. 1985. Dating of the human–ape splitting by a molecular clock of mitochondrial DNA. *J. Mol. Evol.* **22:** 160–174.

Hasegawa, M., Cao, Y., and Yang, Z. 1998. Preponderance of slightly deleterious polymorphism in mitochondrial DNA: Nonsynonymous/synonymous rate ratio is much higher within species than between species. *Mol. Biol. Evol.* **15:** 1499–1505.

Hoelzel, A.R., Lopez, J.V., Dover, G.A., and O'Brien, S.J. 1994. Rapid evolution of a heteroplasmic repetitive sequence in the mitochondrial DNA control region of carnivores. *J. Mol. Evol.* **39:** 191–199.

Huelsenbeck, J.P. and Ronquist, F. 2001. MRBAYES: Bayesian inference of phylogeny. *Bioinformatics* **17:** 754–755.

Ingman, M., Kaessmann, H., Pääbo, S., and Gyllensten, U. 2000. Mitochondrial genome variation and the origin of modern humans. *Nature* **408:** 708–713.

Kivisild, T., Shen, P., Wall, D.P., Do, B., Sung, R., Davis, K., Passarino, G., Underhill, P.A., Scharfe, C., Torroni, A., et al. 2006. The role of selection in the evolution of human mitochondrial genomes. *Genetics* **172:** 373–387.

Ohta, T. and Kimura, M. 1971. On the constancy of the evolutionary rate of cistrons. *J. Mol. Evol.* **1:** 18–25.

Ostrander, E.A. and Kruglyak, L. 2000. Unleashing the canine genome. *Genome Res.* **10:** 1271–1274.

Otto, S.P. and Whitlock, M.C. 1997. The probability of fixation in populations of changing size. *Genetics* **146:** 723–733.

Parker, H.G., Kim, L.V., Sutter, N.B., Carlson, S., Lorentzen, T.D., Malek, T.B., Johnson, G.S., DeFrance, H.B., Ostrander, E.A., and Kruglyak, L. 2004. Genetic structure of the purebred domestic dog. *Science* **304:** 1160–1166.

Piganeau, G. and Eyre-Walker, A. 2003. Estimating the distribution of fitness effects from DNA sequence data: Implications for the molecular clock. *Proc. Natl. Acad. Sci.* **100:** 10335–10340.

Posada, D. and Crandall, K.A. 1998. MODELTEST: Testing the model of DNA substitution. *Bioinformatics* **14:** 817–818.

Ruiz-Pesini, E., Mishmar, D., Brandon, M., Procaccio, V., and Wallace, D.C. 2004. Effects of purifying and adaptive selection on regional variation in human mtDNA. *Science* **303:** 223–226.

Sablin, M.V. and Khlopachev, G.A. 2002. The earliest Ice Age dogs: Evidence from Eliseevichi. *Curr. Anthropol.* **43:** 795–799.

Saetre, P., Lindberg, J., Leonard, J.A., Olsson, K., Pettersson, U., Ellegren, H., Bergstrom, T.F., Vilà, C., and Jazin, E. 2004. From wolf to domestic dog: Changes in brain gene expression. *Brain Res. Mol. Brain Res.* **126:** 198–206.

Sarich, V.M. and Wilson, A.C. 1973. Generation time and genomic evolution in primates. *Science* **179:** 1144–1147.

Savolainen, P., Zhang, Y.P., Luo, J., Lundeberg, J., and Leitner, T. 2002. Genetic evidence for an East Asian origin of domestic dogs. *Science* **298:** 1610–1613.

Sutter, N.B. and Ostrander, E.A. 2004. Dog star rising: The canine genetic system. *Nat. Rev. Genet.* **5:** 900–910.

Swofford, D.L. 2002. *PAUP*. Phylogenetic analysis using parsimony (*and other methods). Version 4*. Sinauer Associates, Sunderland, Massachusetts.

Vilà, C., Savolainen, P., Maldonado, J.E., Amorim, I.R., Rice, J.E., Honeycutt, R.L., Crandall, K.A., Lundeberg, J., and Wayne, R.K. 1997. Multiple and ancient origins of the domestic dog. *Science* **276:** 1687–1689.

Vilà, C., Amorim, I.R., Leonard, J.A., Posada, D., Castroviejo, J., Petrucci-Fonseca, F., Crandall, K.A., Ellegren, H., and Wayne, R.K. 1999. Mitochondrial DNA phylogeography and population history of the Gray Wolf *Canis lupus*. *Mol. Ecol.* **8:** 2089–2103.

Vilà, C., Seddon, J., and Ellegren, H. 2005. Genes of domestic mammals augmented by backcrossing with wild ancestors. *Trends Genet.* **21:** 214–218.

Wang, W. and Kirkness, E.F. 2005. Short interspersed elements (SINEs) are a major source of canine genomic diversity. *Genome Res.* **15:** 1798–1808.

Wayne, R.K. 2001. Consequences of domestication: Morphological diversity of the dog. In *The genetics of the dog* (eds. A. Ruvinsky and J. Sampson), pp. 43–60. CABI Publishing, Oxon, UK.

Yang, Z. 1997. PAML: A program package for phylogenetic analysis by maximum likelihood. *Comput. Appl. Biosci.* **13:** 555–556.

Zhang, J. 2000. Rates of conservative and radical nonsynonymous nucleotide substitutions in mammalian nuclear genes. *J. Mol. Evol.* **50:** 56–68.

Received January 4, 2006; accepted in revised form May 1, 2006.

ONLINE SUPPLEMENT

Mitochondrial DNA sequencing

DNA was extracted using a proteinase K digestion followed by a phenol- chloroform extraction (Sambrook *et al.* 1989) from either tissue or blood samples. Selection of dogs for the complete mtDNA study were done by sequencing the left domain control region with primers Thr-L 15926 5'-CAA TTC CCC GGT CTT GTA AAC C-3' and DL-H 16340 5'-CCT GAA GTA GGA ACC AGA TG-3' (Vilà *et al.* 1999). A neighbour-joining tree was built and 2-6 dogs from each clade were then sequenced for the complete mtDNA (Fig. S1). The entire mtDNA molecule was amplified and sequenced using 37 pairs of dog-specific plus three pairs of coyote-specific PCR primer pairs (Table S1). Primers were designed with the program Oligo (Molecular Biology Insights, Inc., Cascade, Colorado, USA) using a sequence from GeneBank (Accession Number: NC_002008).

PCR amplifications were done in 50µl PCR reactions. This included 1xPCR buffer containing 1.5mM $MgCl_2$, 0.25µM primer, 0.4mM dNTP, 0.75U AmpliTaq Gold (Applied Biosystems) and 30ng DNA. The PCR profile for the mtDNA control region included an initial denaturation step at 94°C for 7 min followed by 35 cycles of amplification (denaturation at 95°C for 1 min, annealing at 50°C for 2 min and extension at 72°C for 1 min and 30 s) and a final extension at 72°C for 10 min. The PCR profile for the complete mtDNA primers included an initial denaturation step at 95°C for 7 min followed by 14 touchdown-cycles (30 s of denaturation at 95°C, 30 s annealing starting at 58°C and decreasing 0.5°C each cycle, followed by extension at 72°C for 1 min), 20 cycles of amplification (denaturation at 95°C for 30 s, annealing at 51°C for 30 s and extension at 72°C for 1 min) and an additional extension step of 10 min at 72°C. Some of the fragments had to be run on a lower initial annealing temperature of 55°C in the touchdown cycles and 48°C in the amplification cycles. PCR products were purified using QIAquick PCR Purification Kit (Qiagen). The

purified PCR products sequenced on an ABI377 Prism DNA sequencer (Applied Biosystems) using chemistry and protocols suggested by the manufacturer.

References

Nei M, Kumar S (2000) Molecular Evolution and Phylogenetics. Oxford: Oxford University Press.

Sambrook J, Fritsch EF, Maniatis T (1989) Molecular Cloning: a Laboratory Manual. Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.

Vilà C, Savolainen P, Maldonado JE, Amorim IR, Rice JE et al. (1997) Multiple and ancient origins of the domestic dog. Science 276: 1687-1689.

Vilà C, Amorim IR, Leonard JA, Posada D, Castroviejo J et al. (1999) Mitochondrial DNA phylogeography and population history of the Gray Wolf Canis lupus. Mol. Ecol. 8: 2089-2103.

Figure S1. Neighbour-joining phylogenetic tree of 60 control region I sequences corresponding to 88 dogs. The tree was built using Jukes-Cantor model of sequence evolution (Nei & Kumar, 2000). Dog clades (I to IV) are indicated as in Vilà et al. (1997). Individuals chosen for the sequencing of the entire mitochondrial DNA molecule are indicated in colour inside each clade, with the code used in the text (see Table S2).

Figure S2. Maximum-likelihood estimates of dN/dS. Estimates for dog branches are indicated in yellow, for wolves in blue and for the wolf/coyote divergence in red. 95% confidence intervals were calculated by bootstrapping.

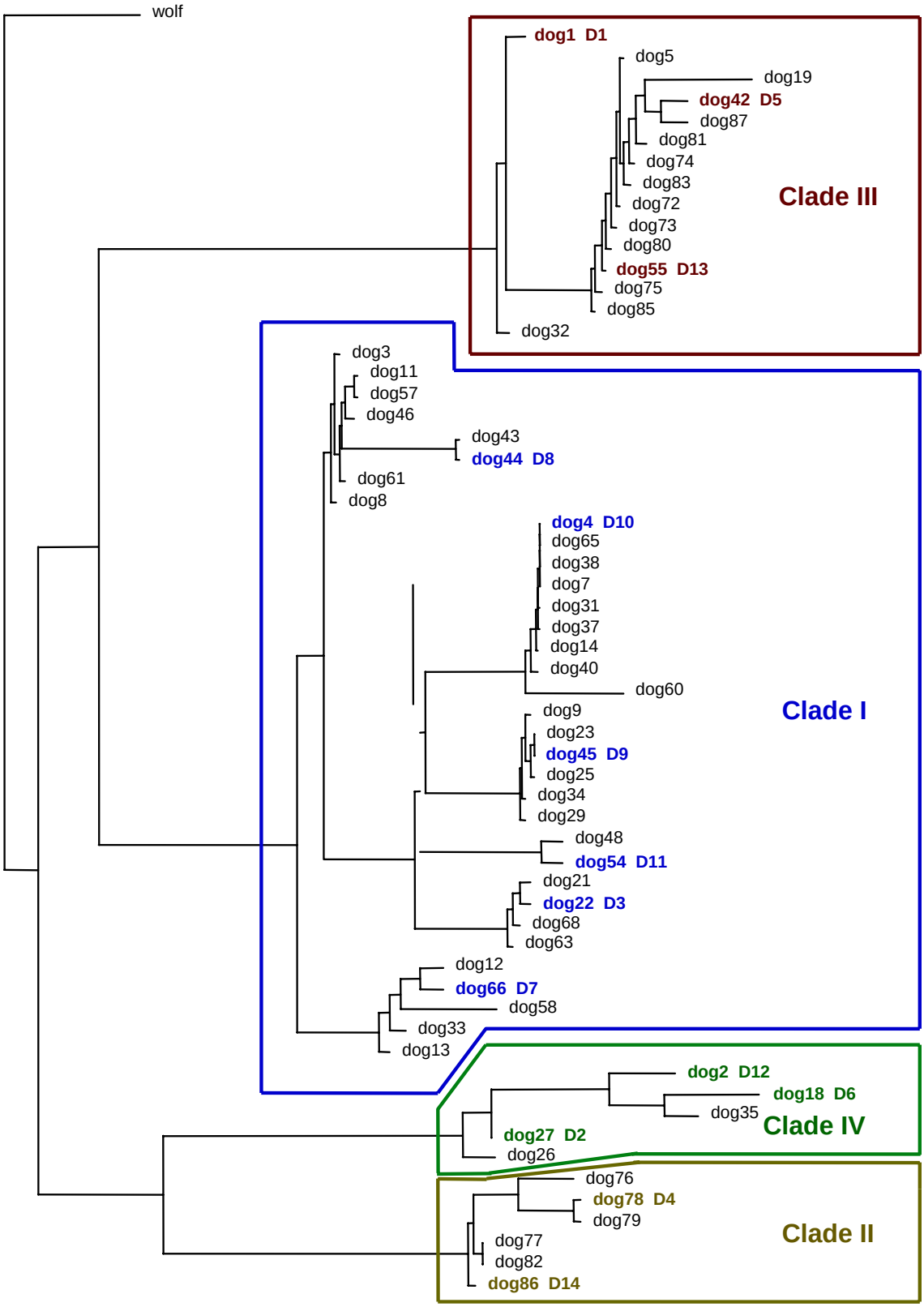
wolf
dog1 D1
dog5
dog19
dog42 D5
dog87
dog81
dog74
dog83
dog72
dog73
dog80
dog55 D13
dog75
dog85
dog32
Clade III
dog3
dog11
dog57
dog46
dog43
dog44 D8
dog61
dog8
dog4 D10
dog65
dog38
dog7
dog31
dog37
dog14
dog40
dog60
dog9
dog23
dog45 D9
dog25
dog34
dog29
dog48
dog54 D11
dog21
dog22 D3
dog68
dog63
dog12
dog66 D7
dog58
dog33
dog13
Clade I
dog2 D12
dog18 D6
dog35
Clade IV
dog27 D2
dog26
dog76
dog78 D4
dog79
dog77
dog82
dog86 D14
Clade II

Figure S1

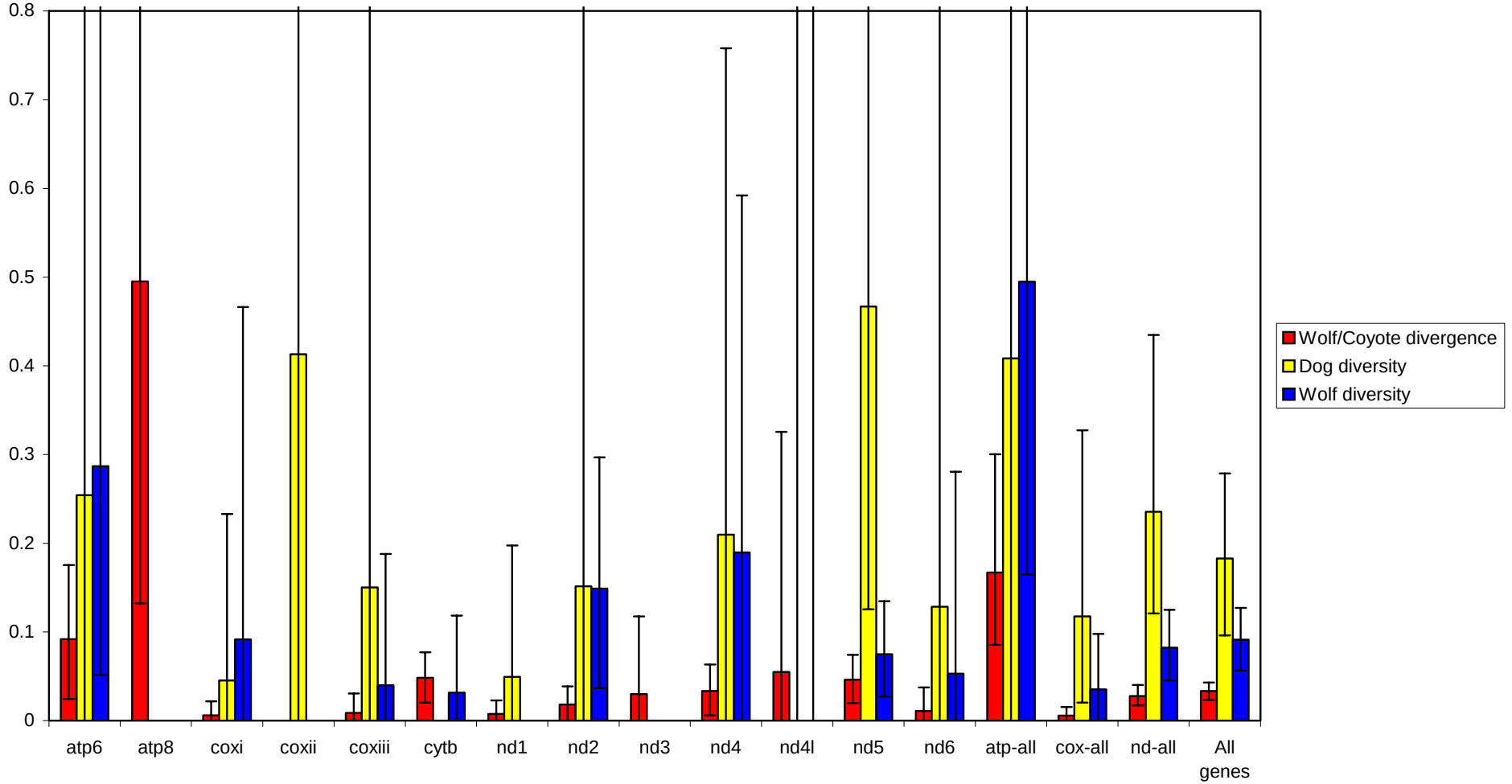
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
atp6
atp8
coxi
coxii
coxiii
cytb
nd1
nd2
nd3
nd4
nd4l
nd5
nd6
atp-all
cox-all
nd-all
All genes
Wolf/Coyote divergence
Dog diversity
Wolf diversity

Figure S2

Table S1. Samples included in the study. For the dogs, the clade including their mtDNA control region haplotype is indicated (see Figure S1).

Code	Species	Breed/Locality	Dog haplotype clade
D1	Dog	German Shepherd	III
D2	Dog	Flat Coated Retriever	IV
D3	Dog	Irish Setter	I
D4	Dog	Jämthund	II
D5	Dog	Black Russian Terrier	III
D6	Dog	Poodle	IV
D7	Dog	Cocker Spaniel	I
D8	Dog	Irish Soft Coated Wheaten Terrier	I
D9	Dog	West Highland White Terrier	I
D10	Dog	Miniature Schnauzer	I
D11	Dog	Siberian Husky	I
D12	Dog	Shetland Sheepdog	IV
D13	Dog	Swedish Elkhound	III
D14	Dog	Jämthund	II
W1	Wolf	W Russia	
W2	Wolf	Sweden	
W3	Wolf	Spain	
W4	Wolf	Saudi Arabia	
W5	Wolf	Saudi Arabia	
W6	Wolf	Canada	
C1	Coyote	Nebraska, USA	
C2	Coyote	Colorado, USA	
C3	Coyote	Colorado, USA	

Table S2. PCR primers used to amplify and sequence the complete mtDNA in dogs, wolves and coyotes.

Primer pair:	Primer names:	Primer sequence:
1a[1]	For 29	5′-GCA CTG AAA AAT GCC AAG ATG-3′
	Rev 670	5′-TAG CGA AAG GTG GTG AGG TT-3′
1b[1]	For 36	5′-AAA ATG CCA AGA TGA GTC-3′
	Rev 708	5′-TTG AGG GTT TGC TGA AGA-3′
2	For 549	5′-ATT CGC CAG AGG ACT ACT AG-3′
	Rev 1212	5′-GGTACT ATC TCT ATC GCT CC-3′
3	For 1057	5′-CAC CCA GAA AGA TTT CAT TAC-3′
	Rev 1680	5′-GGA GTT GAT GTA GTA TGG TT-3′
4	For 1639	5′-CAA ACA ATA TAA CTT AAT CCC-3′
	Rev 2141	5′-GTG GTA TTC CCG CCT CTT CA-3′
5	For 1950	5′-GCA TTT CTA GTA TTG GAG GCA C-3′
	Rev 2579	5′-CAT CCC TTG TCC TTT CGT ACT-3′
6	For 2476	5′-GTT CAA CGA TTA AAG TCC TAC-3′
	Rev 3083	5′-GAG GCT TGA TAT TGC TAG TA-3′
7	For 2927	5′-CCT CTA CGA CCA CTT ACA TC-3′
	Rev 3423	5′-GAA AGA ATA GGG CGA AAG GA-3′
8	For 3204	5′-CAG TCC TCC TAA TAA ACG GG-3′
	Rev 3858	5′-CTG ACC TTA CTG TAG AAT ATA G-3′
9	For 3673	5′-CAT TAT CAC CGC AAG TAT CC-3′
	Rev 4258	5′-CAG AAG TGG AAT GGA GAT AG-3′
10	For 4140	5′-CTA TCA ACC TCC TTT ACT CC-3′
	Rev 4825	5′-GTA AGT GCG GTG CTA TAT GT-3′
11	For 4683	5′-CAT TAT CTG GCT TCA TCC CC-3′
	Rev 5388	5′-GTG ATT AGT GGA GAA CAG TC-3′
12	For 5230	5′-TCT AGG CTG CTT CTT TGA-3′
	Rev 6149	5′-TCT TTT TTC CCT GAG TAG-3′

13	For 5812	5′-GAG TCT CTT CTA TTT TAG GGG-3′
	Rev 6543	5′-GAA TCA GTG GGC AAA TCC TC-3′
14	For 6388	5′-TCT TAT TTA CAG TAG GCG GG-3′
	Rev 6949	5′-GGT TAT GAC ATT GGC TTG AAA-3′
15	For 6892	5′-AAT AAG AAA GGA AGG AAT CG-3′
	Rev 7416	5′-CTT GTG TTG GGA TTA TGT A-3′
16	For 7320	5′-TAA CCG TGA AAA CAA TAG GC-3′
	Rev 7882	5′-GTT GAA ATA GGA TGA AGA GG-3′
17	For 7651	5′-CTT TAT ACC CAT TGT TCT TG-3′
	Rev 8248	5′-GGC GTA AAT GAG TGA GGT AAT-3′
18	For 8049	5′-CCA TTT TAT TCC CAA CAC CC-3′
	Rev 8501	5′-GGT AGC CCC TCC ATT CAA A-3′
19	For 8255	5′-CAA CTC TCT ATA AAC CTC GG-3′
	Rev 8891	5′-CGT ATC GTA GTC CTT TTT GTA-3′
20	For 8808	5′-ATG CCA GTG ATG ACG AGA TG-3′
	Rev 9347	5′-CGA AGT GGT GGT TTG ATG TG-3′
21	For 9246	5′-CAC TGG ATT TCA CGG ATT AC-3′
	Rev 9847	5′-CAT ATT CGG TTC ATT CTA GCC-3′
22	For 9729	5′-AGC GTC ACA AAC CAA CAA G-3′
	Rev 10241	5′-GGG ATT AGT ATG ATA GTA GGG-3′
23	For 10104	5′-AGT ATT TGC TGC CTG CGA AG-3′
	Rev 10647	5′-GTA GAG TCC TGC GTT TAG TC-3′
24	For 10561	5′-TCC TAT TTG AAG CAA CAC TG-3′
	Rev 11222	5′-CGG CTA TGG ATT CGT TCG TA-3′
25	For 11093	5′-TAT CGT AGC GGT TCT TAT TC-3′
	Rev 11741	5′-GAC CAA CGG ATT ACT TCT AT-3′
26	For 11566	5′-TCG GTC CCA TCT ACT GTA AG-3′
	Rev 12224	5′-GCC AAC TCC TTC TCA ACC AA-3′
27	For 12050	5′-CCT GTA GCC CTT TTT GTC AC-3′
	Rev 12704	5′-GCT TGA GGT AGA GAA CGC T-3′
28	For 12588	5′-ACC AAA CTA TTC AAA CCC TC-3′

	Rev 13284	5′-GTG CCA GGA TGA AAC CCA AG-3′
29	For 13085	5′-CGC TTC TCC CCT ATA ATC C-3′
	Rev 13712	5′-GGT TTT TTA GTG AAG AGG-3′
30	For 13564	5′-AAT CCT CAG TCT ACT AAT CC-3′
	Rev 14036	5′-CTA TTT ATG GTG GGC TTG TG-3′
31	For 13881	5′-CTC AGT AGC CAT AGC AGT TG-3′
	Rev 14636	5′-GAG AGA AGA TTA GTG ATT ACA G-3′
32	For 14371	5′-TTT CAT CAG TCA CCC ACA TC-3′
	Rev 15025	5′-GGA TAG CAT AGG CGA ATA GA-3′
33	For 14838	5′-CCA TTT CAC CCT TAC TAC AC-3′
	Rev 15382	5′-GGC GGT TAC TCT CCA TTT TT-3′
34a[2]	For 15232	5′-TCG GAC AAG TCG CTT CAA TC-3′
	Rev 15801	5′-GTA AGA ACC AGA TGC CAG G-3′
34b[2]	For 15216	5′-CAC CCT TTC ATC ATC ATC-3′
	Rev 15789	5′-TGC CAG GTA TAG TTT CAT-3′
35a[2]	For 15516	5′-TAA ACC CTT CTC CCC TCC C-3′
	Rev 16116	5′-CCT GAA ACC ATT GAC TGA A-3′
35b[2]	For 15375	5′-AAT CAC CCT CCC TAA GAC TC-3′
	Rev 16117	5′-TCC TGA AAC CAT TGA CTG AA-3′
36	For 15965	5′-GCA ACG GCA CTA ACT CTA AC-3′
	Rev 16640	5′-CGA GAC CAA ATG CGT GTA AG-3′
37	For 16521	5′-CCA AAC CCC AAA AAC AGG AC-3′
	Rev 355	5′-CTT GAA CAC GCT TTA CGC CG-3′

1. Primer pair 1a was used for dogs and wolves, did not work on coyotes or dog D4. Primer pair 1b was used for these.

2. Primer pairs 34a and 35a were used for dogs and wolves, did not work on coyotes. Primer pairs 34b and 35b were used for these.