Mitochondrial DNA from Prehistoric Canids Highlights Relationships Between Dogs and South-East European Wolves

Fabio Verginelli,* Cristian Capelli,† Valentina Coia,‡ Marco Musiani,§ Mario Falchetti,|| Laura Ottini,||| Raffaele Palmirotta,*1 Antonio Tagliazzo,¶ Iacopo De Grossi Mazzorin,# and Renato Mariani-Costantini*8

*Department of Oncology and Neurosciences, University G. d’Annunzio, and Center of Excellence on Aging (Ce.S.I.), G. d’Annunzio Foundation, Chieti, Italy; †Institute of Legal Medicine, Catholic University of S. Cuore, Rome, Italy; ‡Department of Animal and Human Biology, University “La Sapienza,” Rome, Italy; §Faculty of Environmental Design, University of Calgary, Calgary, Alberta, Canada; ||Department of Experimental Medicine and Pathology (Section of History of Medicine), University La Sapienza, Rome, Italy; ¶Laboratory of Quaternary Paleontology and Archeozoology, Special Superintendence to the Luigi Pigorini National Prehistoric-Ethnographic Museum, Rome, Italy; and #Department of Cultural Heritage, University of Lecce, Lecce, Italy.

The question of the origins of the dog has been much debated. The dog is descended from the wolf that at the end of the last glaciation (the archaeologically hypothesized period of dog domestication) was one of the most widespread among Holarctic mammals. Scenarios provided by genetic studies range from multiple dog-founding events to a single origin in East Asia. The earliest fossil dogs, dated ≈17–12,000 radiocarbon (14C) years ago (YA), were found in Europe and in the Middle East. Ancient DNA (a-DNA) evidence could contribute to the identification of dog-founder wolf populations. To gain insight into the relationships between ancient European wolves and dogs we analyzed a 262-bp mitochondrial DNA control region fragment retrieved from five prehistoric Italian canids ranging in age from ≈15,000 to ≈3,000 14C YA. These canids were compared to a worldwide sample of 547 purebred dogs and 341 wolves. The ancient sequences were highly diverse and joined the three major clades of extant dog sequences. Phylogenetic investigations highlighted relationships between the ancient sequences and geographically widespread extant dog matrilines and between the ancient sequences and extant wolf matrilines of mainly East European origin. The results provide a-DNA support for the involvement of European wolves in the origins of the three major dog clades. Genetic data also suggest multiple independent domestication events. East European wolves may still reflect the genetic variation of ancient dog-founder populations.

Introduction

The dog (Canis familiaris L.), considered the first domestic animal, is descended from the grey wolf (Canis lupus L.), a highly mobile and adaptable carnivore that up to few centuries ago was widespread throughout the Holarctic (Kurtén 1968; Olsen 1985; Clutton-Brock 1995; Vila, Maldonado, and Wayne 1999). The places and dates of wolf domestication are much debated. No dog images have been identified in the Franco-Cantabric cave art, which strongly suggests that before ≈16,000 radiocarbon (14C) years ago (YA) the dog was still unknown in Western Europe (Delporte 1990; Nougier 1993; Clutton-Brock 1995). The earliest dog fossils consist in two crania from Eliseeyevichi 1 (Central Russian Plain, 13–17,000 14C YA) and in a mandible from Oberkassel (Germany, ≈14,000 14C YA) (Nobis 1979; Sablin and Khlopachev 2002). These represent robust wolf-sized dogs thought to have derived from the large northern Holarctic wolf form C. I. lupus. Remains of smaller dogs from Mesolithic (Natifuan) cave deposits in the Middle East (≈12,000 14C YA) were interpreted as evidence of independent domestication from a lighter Southwest Asian wolf form, Canis lupus arabs (Davis and Valla 1978; Dayan 1994; Clutton-Brock 1995; Tchernov and Valla 1997). Images in the Saharan and Iberic Epipaleolithic rock art and skeletal remains indicate that, by ≈10,000 14C YA, dogs were already present in areas as far apart as North Africa, Western Europe, and North America (Lindner 1950; Kurtén and Anderson 1980; Clutton-Brock 1995; Schwartz 1997).

The striking phenotypic and genetic diversity of dogs clearly indicates that their founders were recruited from a large and varied wolf population (Clutton-Brock 1995; Wayne and Ostrander 1999; Parker et al. 2004). Mitochondrial DNA (mtDNA) analyses have been used to elucidate relationships between dog and wolf. Phylogenetic trees of dog and wolf hypervariable region 1 (HVR1) sequences show that dogs group into several clades (Tsuda et al. 1997; Vila et al. 1997; Leonard et al. 2002; Savolainen et al. 2002), a fact that clearly indicates independent origins from multiple wolf matrilines. Vila et al. (1997), using a fixed substitution rate mtDNA molecular clock calibrated on a 1 million years age for wolf-coyote divergence (Kurtén and Anderson 1980), estimated that dogs and wolves separated as early as 76–135,000 YA. Based on the same calibration and focusing on worldwide HVR1 sequence variation among dogs, Savolainen et al. (2002) proposed that dogs diverged from wolves 40–15,000 YA and were most probably domesticated in East Asia. This suggests that the available archaeozoological record could be biased toward West Eurasia because of the more extensive fieldwork conducted there. Concomitantly, Leonard et al. (2002) showed that pre-Columbian New World dog lineages derived from Eurasia and, more recently, Savolainen et al. (2004) provided substantial evidence that Australian dogs originated from a population of East Asian dogs. The molecular time estimates of dog origins cited above are challenged by newly uncovered evidence showing that clocks calibrated on phylogenetic substitution rates cannot be applied to the dating of geologically recent divergence events (Ho et al. 2005).
For several millennia the breeding and geographic distribution of dogs depended on man, while in the last few centuries human persecution caused the extinction or bottlenecking of most wolf populations (Vilà et al. 1999; Randi et al. 2000; Flagstad et al. 2003). Thus, the genetics of dogs and wolves bears the historic influence of humans, a strong confounding factor that can be circumvented only by ancient DNA (a-DNA) studies (Hofreiter et al. 2001b; Leonard et al. 2002). A-DNA has shed novel light on the evolutionary history of several species, notably including the Neanderthal, large Pleistocene carnivores, such as bears and hyenas, and domestic animals, such as horses, cattle, and native North American and Island Southeast Asian dogs (Greenwood et al. 1999; Leonard, Wayne, and Cooper 2000; Hofreiter et al. 2001b; 2004; Troy et al. 2001; Vilà et al. 2001; Jansen et al. 2002; Leonard et al. 2002; Savolainen et al. 2004).

Because of the geographic position and archaeozoological record, Italy is a country of potential relevance for dog evolutionary studies (Sauter 1948; Kurtén 1968; Gamble 1986). The present study is centered on novel a-DNA evidence from five prehistoric Italian canids (PICs) ranging in age from ~15,000 to ~3,000 14C YA. The ancient HVR1 sequences were pooled with the sequences of 341 extant wolves and 547 purebred dogs retrieved from databases. Our results provide novel insight into the relationships between dogs and wolf populations.

**Materials and Methods**

Specimens and Radiocarbon Dating

PIC specimens were provided by the Luigi Pigorini National Prehistoric-Ethnographic Museum (LPM), Rome, and by the Laboratory of Archaeozoology (LA), Archaeological Superintendence of Rome. A-DNA could be retrieved from five of eight tested specimens, including PIC-1 (LPM), a mandibular ramus of large canid (no associated teeth) from the Late Glacial levels of the Palidoro Upper Paleolithic rock shelter, Latium, central Italy (Bietti 1976–1977; Cassoli 1976–1977); PIC-2 (LPM), a fragment of large canid mandibial from the Early Holocene “terre brune” level of Romanelli Cave, a well-known Upper Paleolithic site in Apulia, southern Italy; PIC-3 (LPM), a fragment of large canid tibia also from the “terre brune” of Romanelli Cave (Blanc 1920, 1928; Cassoli, Fiore, and Tagliazzo 1994); PIC-4 (LPM), a tibia from a relatively complete dog skeleton found at the Casal del Dolce Eneolithic site, Latium (Fiore and Tagliazzo 1997); PIC-5 (LA), a dog femur from the Vejano Recent Bronze Age site, Latium (De Grossi Mazzorin and Tagliazzo 2000). PIC-1 through -3 were associated with cold final Pleistocene fauna and Epigravettian Upper Paleolithic industry. These three PIC specimens were within the range of morphometric variability of the Late Upper Pleistocene Italian wolves (Cassoli, Fiore, and Tagliazzo 1994) but lacked diagnostic features that could exclude attribution to large proto-dogs (Sablin and Khlopachev 2002). PIC-4 and PIC-5 were unambiguously attributable to medium-sized dogs, with no breed differentiation evidenced after analysis of the PIC-4 skeleton (Fiore and Tagliazzo 1997; De Grossi Mazzorin and Tagliazzo 2000). Specimens were radiocarbon dated from purified collagen by using accelerator mass spectrometry at BETA Analytic (Miami, Fla.). Reported dates are 14C years before A.D. 1950, corrected for 13C fractionation.

A-DNA Retrieval, Analysis, and Authentication

Bone surfaces were scraped with sterile blades, UV irradiated for 1 h, and microtrephined to obtain at least 1 g of bone powder, stored in sealed 250 mg aliquots until use. Bone powder aliquots were washed overnight in 0.5 M ethylenediaminetetraacetic acid (EDTA) (pH 8.0), digested at 37°C for 48 h in 600 μl of protease K (20 mg/ml in 0.5 M EDTA [pH 8.0]/5% sodium dodecyl sulfate), and extracted as in Krings et al. (1997). Two nested partially overlapping polymerase chain reactions (PCRs) were used to amplify a 299-bp (with primers) canine HVR1 fragment encompassing the 262-bp segment (nucleotide positions [bps] 15431–15687) utilized for phylogenetic analyses. External (ext) and internal (int) primers, identified aligning 26 dog and 27 wolf haplotypes from Vilà et al. (1997), and sizes of PCR products were L15422extCTCTTGTGTTCCACATCAGC-3'/H15548extTTATATGCGTGTTGGGTTACC-3', 173 bp; L15426intCTCAGGTGGGTTACCC-3'/H15555intATGGGGCA-AACATTACAT-3', 157 bp; L15511extACTGTGCTAGTTGCTTGGTTACCC-3'/H15691extTTATGCGTGTTGGGTTACC-3', 220 bp; and L15529intCTCAGGTGGGTTACCC-3'/H15691intATGGGGCA-AACATTACAT-3', 199 bp. PCRs were carried out in 40-μl volumes with 10 mM Tris (pH 8.3), 1.5 mM MgCl2, 200 μM deoxynucleoside triphosphates, 1.5 units of AmpliTaq Gold (Applied Biosystems, Foster City, Calif.), 10 pM of each primer, and 4 μl of a-DNA extract. External PCR conditions were 50 cycles of 50°C for 45 s, 72°C for 45 s, and 94°C for 30 s; for internal PCRs (19 cycles) annealing was at 55°C. Primers produced no amplifications when tested on human DNA; conversely, primers designed for human mtDNA did not yield amplifications when tested on the a-DNAs. PCR products were directly sequenced on an ABI-PRISM 310 Sequencer (Applied Biosystems). Contamination was strictly monitored with extraction and PCR blanks.

Some of the main a-DNA authentication criteria, including the most stringent (i.e., independent replication), were followed (Hofreiter et al. 2001a). All PIC specimens were independently sampled, extracted, PCR amplified, and sequenced in Chieti (University G. d’Annunzio) and in Rome (Catholic University of S. Cuore). In both laboratories, a-DNA extractions and amplifications were performed in separate dedicated spaces where no modern canid samples had been manipulated (setup of PCR and sequencing conditions on modern dog DNA was performed in a third laboratory). Reagents, bench surfaces, and non-disposable equipment were routinely sterilized under 254-nm UV light. Chieti and Rome generated independent amplification products for the two partially overlapping HVR1 segments analyzed; consensus sequences were obtained by alignment of at least two independent direct sequences. Identical sequences were replicated from the PIC-3 a-DNA extracts; discrepancies in 3, 2, 1, and 1 nt were observed for the PIC-1, PIC-2, PIC-4, and PIC-5 sequences,
Phylogenetic Analysis

Frequencies, geographic origins, and GenBank accession numbers of the extant wolf and dog HVR1 sequences used in this study are given in Tables S1–S4 (see the Supplementary Material online). Extant wolf sequences were from 341 individuals (Europe, 256; Asia, 66; America, 20 [Tsuda et al. 1997; Vilà et al. 1997; Randi et al. 2000; Savolainen et al. 2002]). Extant dog sequences were from 547 purebred dogs, grouped according to continent as indicated by breed (Europe, 231; Asia, 255; Africa, 31; America, 30). Outbred dogs were not considered because of greater uncertainty about origins. PIC sequences (262 bp, nps 15431–15687) were pooled with unique GenBank HVR1 sequences (262–664 bp, nps 15431–15687/16039, Tables S1–S4, Supplementary Material online) of dogs (92 sequences), wolves (53 sequences), and coyotes (2 sequences). To avoid collapse of previously identified dog clades (Tsuda et al. 1997; Vilà et al. 1997; Leonard et al. 2002; Savolainen et al. 2002, 2004), sequences were evaluated along their full lengths, filling gaps with unknown nucleotides for shorter sequences. Arlequin 2000 (Schneider, Roessli, and Excoffier 2000) was used to identify haplotype sharing among ancient and extant sequences and to perform Fu’s Fs tests of selective neutrality for identified clades. The phylogenetic trees were constructed through Neighbor-Joining (NJ) algorithm using PHYLIP 3.6 Inferno (Felsenstein 2002). Distance matrices were based on the F84 model with a $\gamma$-distribution for substitution rate heterogeneity (Felsenstein 1984). The $\gamma$-shape and $\alpha$-transition/transversion parameters were set to 0.10 and 9.94, respectively, as estimated using Tree-Puzzle 5.0 (Schmidt et al. 2002). Significance was determined from bootstrap percentages obtained after replication of 100 trees. To obtain median-joining (MJ) networks of dog clades containing PICs, the DnaSP v.3 software (available at http://www.ub.es/dnasp/) was applied to sequence data to identify segregating variants in the 262-bp PIC segment. Tables of variants were used in the MJ algorithm option of NETWORK 4.000 (Bandelt, Forster, and Röhl 1999). The wolf MJ network was constructed as above, based on variants in the 262-bp PIC segment (44 sequences including 16 from Europe, 20 from Asia, 6 from America, and 2 shared between Europe and Asia [Tsuda et al. 1997; Vilà et al. 1997; Randi et al. 2000; Savolainen et al. 2002]). To avoid high-dimensional cubes, rapidly mutating variants at nps 15625 and 15643 were 10-fold down-weighted. Maximum parsimony (MP) algorithm was applied to purge the wolf MJ network of superfluous links. In all cases, default settings were $r = 2$ and $e = 0$ (most parsimonious pathway).

Networks based on the 262-bp PIC segment were also constructed by MP method using TCS version 1.13 (Clement, Posada, and Crandall 2000), where the probability of parsimony, calculated for pairwise differences up to the 95% cutoff level, justifies the maximum number of mutational connections between pairs of sequences (indels considered as a fifth state). Rooting weights were calculated using the algorithm implemented in the TCS program, based on the frequency of each candidate root haplotype and on the number and frequencies of the haplotypes to which it is connected (Castelloe and Templeton 1994).

Spatial structure in haplotype distribution was examined with conventional $F$-statistics (Wright 1951) and hierarchical analysis of molecular variance (AMOVA; Excoffier, Smouse, and Quattro 1992). We calculated the matrices of pairwise Fst values between sampling locations. Locations of dog and wolf sequences were assigned according to area of origin for the dog breed and/or according to the references listed in Tables S1 and S2 (see the Supplementary Material online). Such locations were defined based on the latitude and longitude of the center of the country (or continental area) of origin. In AMOVA, we tested statistical significance after 1,000 permutations for the correlation of random genotypes within spatially clumped groups relative to the whole population. AMOVA utilizes both the frequency and sequence divergence between genotypes. We estimated the divergence between genotypes assuming a Tamura-Nei model of sequence evolution and a gamma distribution of the substitution rates with a value of $\alpha = 0.5$ (Tamura and Nei 1993; Wakeley 1993). A matrix of pairwise Euclidean distances between the central locations for each group was also calculated (mean latitude and longitude for relevant sampling locations). Finally, we obtained other matrices by calculating all pairwise geographic distances imposing a theoretical passage through the Bering Strait. Genetic and geographic matrices were tested for correlation using Mantel’s test (Liedloff 1999). Arlequin 2000 (Schneider, Roessli, and Excoffier 2000) and Genetix 4.05 (Belkhir et al. 2004) software were used for computations.

Results

We succeeded in retrieving 262-bp HVR1 sequences from five PICs that included PIC-1 (Late Glacial, 14,670 ± 130 14C YA), PIC-2 (Early Holocene, 9,860 ± 50 14C YA), PIC-3 (Early Holocene 9,670 ± 40 14C YA), PIC-4 (Eneolithic, 4,110 ± 40 14C YA), and PIC-5 (Recent Bronze Age, 3,040 ± 40 14C YA). These PICs yielded individual sequences that varied by 2–11 nt at 13 polymorphic sites (all variants occur in GenBank wolf and/or dog sequences). Independent replication of these sequences strongly supports their authenticity.

To highlight similarity relationships between PIC sequences and modern wolf and dog HVR1 sequences (262–664 bp), we constructed an NJ tree rooted using two coyote sequences. The NJ tree split dogs (92 haplotypes) and wolves (53 haplotypes) into nine clades, identified with Roman numerals in figure 1. Clades I, II, IV, VI, and IX corresponded to previously defined dog-dominated or dog-monophyletic clades (Tsuda et al. 1997; Vilà et al. 1997; Leonard et al. 2002; Savolainen et al. 2002, 2004). Clade I, the main dog clade, contained 56 dog haplotypes.
from Eurasia, the Americas, and Africa, 2 haplotypes shared between European wolves and dogs, and 3 Asian wolf haplotypes. Clade II comprised one West Asian and five European dog haplotypes. Clade IV encompassed eight dog haplotypes found in Eurasian, American, and African breeds. Clade VI, a dog-dominated clade comprising several wolves, included 15 dog haplotypes from Europe and Asia, 1 haplotype shared between European wolves and dogs, and 4 wolf haplotypes (three from Europe and one from Asia). Minor dog clade IX encompassed three East Asian haplotypes. Clades III, V, VII, and VIII included only wolves, with the exception of a dog haplotype embedded in clade VIII (Savolainen et al. 2002). Clade III included five European wolf haplotypes (among which the single haplotype of extant Italian/French wolves [Vila` et al. 1997; Randi et al. 2000; Lucchini et al. 2002]). Minor clades V and VII contained three wolf haplotypes each (two from Asia and one from Europe [clade V] and two from America and one from Europe [clade VI]). Clade VIII, the main wolf clade, comprised 29 haplotypes from the whole wolf geographic range and 1 East Asian dog haplotype. Wolf haplotypes W5 (Mongolia), W10 (China), and W13 (China) were outside defined clades. Dog clades II, IV, and IX had bootstrap support over 50%. Excluding wolf sequences in our bootstrap analysis gave over 50% support also for dog clade VI (data not shown). Because of extensive homoplasy, bootstrap support for clade I and for wolf clades could not be obtained (Savolainen et al. 2002).

The 262-bp PIC sequences clustered with major dog clades I, IV, and VI. PIC-1 (Late Glacial) and PIC-2 (Early Holocene) showed novel haplotypes. PIC-1 occupied a sister branch of clade IV, PIC-2 clustered in clade VI, and PIC-3 (Early Holocene) clustered in clade I and resulted identical with a reported a-DNA sequence retrieved from pre-European contact Alaskan dogs (Leonard et al. 2002; Savolainen et al. 2002). The PIC-3 sequence was shared by several Asian dog haplotypes and by W51-A11, an intriguing haplotype found in a Bulgarian wolf and in dog breeds with ancient Old and New World origins (Vila` et al. 1997; Randi et al. 2000; Leonard et al. 2002; Savolainen et al. 2002). PIC-4 (Eneolithic) was shared by four Old World dog haplotypes in clade IV and PIC-5 (Bronze Age) by five widespread dog haplotypes in clade I.

To define the evolutionary hierarchies of the haplotypes within major dog clades I, IV, and VI, we constructed...
262-bp networks using both MJ and MP approaches to obtain the most parsimonious pathways (ε = 0 for MJ and 95% confidence level for MP). The two independent approaches yielded identical network topologies for each clade. Clade I had a complex pattern (fig. 2). The nested design of this clade obtained by TCS resulted in four nested levels with two subclades, 3-1 and 3-2, at the third nesting level. Subclade 3-1 included three second nesting level groups, 2-1, 2-4, and 2-5; subclade 3-2 contained two closely related groups, 2-2 and 2-3. The algorithm implemented in the TCS program for calculating rooting weights (Castelloe and Templeton, 1994) located the root within subclade 2-1 (highest rooting probability 0.13 for A3), with a slightly lower rooting probability value for A11 in subclade 2-2 and A18 in subclade 2-3 (in both cases 0.11) and a markedly lower value for A29 in subclade 2-4 (0.06). The topologies of clades IV and VI were straightforward and without reticulations. Significant support for the expansion hypothesis was provided by the negative Fu’s Fs values obtained for clades I (−9.6, P = 0.01) and VI (−5.1, P = 0.01) and, within clade I, for subclades 3-1 (−6.3, P = 0.02) and 3-2 (−11.1, P < 0.01) at the third nesting level as well as subclades 2-2 (−8.4, P < 0.01) and 2-3 at the second nesting level (−5.2, P < 0.01).

The MJ networks of dog clades I, IV, and VI shown in figure 3 highlight the positions of PICs and of extant wolves relative to dogs. The three starlike subclades 2-1, 2-2, and 2-3 identified by nested clade analysis are outlined in the MJ network of clade I (fig. 3A-A1). Topological features suggest independent origins for 2-1 and 2-2, with 2-3 (monophyletic for dogs) clearly derived from 2-2. The central node of 2-1 included Mongolian/Chinese wolves (W2/W3) and a majority of Asian dogs, remaining dogs being from Africa. The central node of 2-2 also included mostly Asian dogs but comprised Bulgarian wolf W51 and PIC-3 (Early Holocene), while PIC-5 (Bronze Age) clustered to the central node of 2-3 with a majority of European dogs.

The networks of clades IV and VI suggested origins from single haplotype (fig. 3B-B1 and C-C1). Both clades contained a majority of Asian dogs; however, PIC-4 (Eneolithic) occupied the central node of clade IV (two mutational steps from the Late Upper Pleistocene sequence PIC-1) (fig. 3B-B1), and the central node of clade VI contained East European wolves (W6/W26, Romania-Russia/Greece), with PIC-2 (Early Holocene) and wolves from former Yugoslavia-Bulgaria-Greece (W8), Afghanistan (W7), and Bulgaria (W53) in derived nodes (fig. 3C-C1).

A 262-bp wolf MJ network was obtained to further analyze the relationships of PICs with extant wolves (fig. 4A). PICs grouped to two of at least four major wolf clusters identified in the network. The sequence shared between PIC-3 (Early Holocene) and the Bulgarian wolf W51 occupied the central node of one of the PIC-containing wolf clusters, with PIC-5 (Bronze Age), PIC-1 (Late Upper Pleistocene), and PIC-4 (Eneolithic) in derived nodes. These PICs were related to a branch including wolves from Romania/Greece (W20), Bulgaria (W35), Italy/France (W21), and former Yugoslavia/Bulgaria (W22/W34). An additional PIC-3/W51-derived branch included wolves with European and Asian origins. PIC-2 (Early Holocene)
was directly linked to the central node of the other PIC-containing wolf cluster, occupied by Romanian-Russian/Greek wolves (W6/W26). This cluster comprised wolves from Afghanistan (W7), former Yugoslavia-Bulgaria-Greece (W8), China (W10), and Bulgaria (W53). Notably, the wolf clusters that contained PICs included all the 262-bp sequences shared between wolf and dog, five of which (i.e., W6/W26, W8, W51, W52, W53) are reported in East European wolves and one (i.e., W2/W3 [Mongolia/China]) in Asian wolves.

Phylogeographic analyses were performed to test correlations between geographical locations and genetic relatedness of the wolf and/or dog HVR1 sequences. Straight-line distances ranging from 3,666 to 13,697 km separated central locations for the dog and wolf sequences from Europe, West Asia, East Asia, Africa, North America, and Central America that were analyzed in this study (Fig. S1A in the Supplementary Material online). Genetic relatedness was mediated by distance. The matrix of pairwise geographic distances calculated imposing a theoretical passage through the Bering Strait was correlated with the matrix of pairwise Fst values between these groups \((n = 6, r = 0.58, P < 0.05)\) (Fig. S1B in the Supplementary Material online). Correlations were not significant when plain Euclidean distances between locations were used. Our results did not support phylogeographic patterns within dog or wolf sequences analyzed independently, as well as within clades or subclades, because no significant relationships between genetic and geographic distances were found.

**Discussion**

At the end of the last glaciation, that coincides with the archaeologically hypothesized period of the origin of the dog, the wolf was one of the commonest among Holarctic mammals and could widely range across the Northern Hemisphere because low sea levels and prevalence of steppes/steppe-tundras facilitated its dispersal (Kurten 1968; Kurten and Anderson 1980; Williams 1998). Given a probable late date for the diffusion of man in North America (Dixon 2001), it is expected that wolves were initially domesticated in Eurasia, a view that agrees with genetic evidences (Leonard et al. 2002). Archaeozoological findings show that by 17–12,000 \(^{14}\)C YA proto-dogs, probably originating from different wolf populations, were already present in West Eurasia (Davis and Valla 1978; Nobis 1979; Dayan 1994; Clutton-Brock 1995; Tchernov and Valla 1997; Sablin and Khlopachev 2002). Molecular time estimates suggesting dog origins 76–135,000 YA or 40–15,000 YA (Vila` et al. 1997; Savolainen et al. 2002) need to be reevaluated based on recent findings showing that molecular clocks cannot be used to date geologically recent events, such as wolf-dog divergence (Ho et al. 2005).

To shed light on dog-founder wolf populations, we took advantage of the unique insight provided by PIC sequences. PIC-1 \((14,670 \pm 130 \, ^{14}\text{C} \, \text{YA})\) lived during the last Late Glacial Maximum \((\approx 22–14,000 \, ^{14}\text{C} \, \text{YA})\), PIC-2 \((9,860 \pm 50 \, ^{14}\text{C} \, \text{YA})\) and PIC-3 \((9,670 \pm 40 \, ^{14}\text{C} \, \text{YA})\), almost synchronous and sympatric, lived shortly after the Younger Dryas, a cold interval following the initial
FIG. 4.—Relationships between PICs and extant wolves. Panel A. Phylogenetic 262-bp network including 343 wolf HVRI sequences (44 haplotypes) and the 5 PIC sequences. Nodes of the two clusters containing PIC sequences are colored in green, when containing only extant wolf sequences, and in red, when containing PIC sequences. Panel B. PIC sites, known fossil dog sites predating 10,000 14C YA and approximate modern distribution of wolves in PIC-containing wolf clusters on the background of Western Eurasia at Late Glacial Maximum (~22–14,000 YA). Red dots with PIC designation point to PIC sites: PIC-1, Palidoro; PIC-5, Vejano; PIC-4, Casal del Dolce; PIC-2 and PIC-3, Romanelli Cave; green dots with sequence designation (prefix W followed by sequence number) indicate extant wolf locations. Extant haplotypes that share identical 262-bp sequences but differ in the rest of the sequence are indicated using their specific designations, haplotypes W1 and W2 (Mongolia), W3 and W10 (China), and W7 (Afghanistan), located outside the area shown, are not represented. Semicircles with black cores and specific designations indicate locations of the main fossil dog sites (Eliseyevich 1 and Oberkassel in Europe and Natufian sites in the Middle East). Light blue areas indicate glaciated regions; yellow bidirectional arrows suggest possible directions of migration.
postglacial warming (Björck et al. 1996; Williams 1998). PIC-1 through -3 were within the range of morphometric variability of the Late Upper Pleistocene Italian wolves (Cassoli, Fiore, and Tagliazuczo 1994) but could not be unambiguously assigned to wolves or wolf-sized proto-dogs because of the poor morphological resolving power of the fossil fragments (Sabin and Khlopachev 2002). However, despite intensive archaeological fieldwork, no proto-dogs have been identified in Italian Upper Paleolithic sites, while at the end of the last glaciation, wolves were common in Italy and are archaeologically well documented (Kurtén 1968; Radmilli 1974; Cassoli, Fiore, and Tagliazuczo 1994). PIC-4 (4,110 ± 40 14C YA) and PIC-5 (3,040 ± 40 14C YA), clearly identifiable as dogs, most likely reflect an indigenous Italian stock (De Grossi Mazzorin and Tagliazuczo 1998).

The high genetic diversity of PIC-1, PIC-2, and PIC-3 is clearly demonstrated by their dispersal across the three major dog clades and suggests that these early PICs were connected with a large and varied population. This early diversity fits with the variation of the later PIC dogs, but the fact that two of the three Late Pleistocene/Early Holocene PIC sequences were novel, while both the later PIC dog sequences were found in modern dogs, suggests that much of the ancient variability has been lost (Vilà et al. 1999; Flagstad et al. 2003). Loss of genetic diversity is clearly evident when comparing the Late Pleistocene/Early Holocene PICs with extant Italian/French wolves, characterized by a single HVR1 haplotype (Vilà et al. 1997; Randi et al. 2000; Lucchini et al. 2002). This could reflect severe bottlenecking during the Holocene and suggests that caution is necessary in drawing historical inferences from modern population data, as previously noted for other Upper Pleistocene carnivores (Leonard, Wayne, and Cooper 2000; Hofreiter et al. 2004).

PIC HVR1 sequences were compared to a large sample of wolf and dog sequences worldwide. As observed by Savolainen et al. (2002), the interposition of wolf and dog clades, the presence of wolves within dog clades, and the lack of bootstrap support for clade I (the main dog clade) and for wolf clades indicated that wolf/dog taxonomic status was independent of HVR1 lineage. Significant bootstrap support was obtained for dog clades II, IV, and IX, as observed for corresponding clades by Savolainen et al. (2002), and, excluding wolves, for dog clade VI.

The complex pattern of the clade I network was compatible with a minimum of two independent founding events, for subclades 3-1 and 3-2 identified at the third nesting level by nested clade analysis. At the second nesting level, 3-1 encompassed subgroups 2-1, 2-4, and 2-5, while 3-2 included subgroup 2-2 and derived 2-3. The starlike shape of 2-1, 2-2, and 2-3 and the significantly negative Fu’s Fs test values obtained for subclades 3-1 and 3-2 at the third nesting level and subclades 2-2 and 2-3 at the second nesting level indicate sudden population expansion events. The presence of Asian but not European wolves and dogs in the starlike node of 2-1 suggests an Asian origin, as previously observed (Savolainen et al. 2002). However, the position of PIC-3 and W51/A11 (the haplotype identical between a Bulgarian wolf and ancient breeds with Old and New World origins [Randi et al. 2000]) suggests that 2-2 derived from European wolves. Furthermore, a European origin for dog-monophyletic subclade 2-3, derived from 2-2, is strongly supported by the central position of the Bronze Age dog PIC-5, together with a majority of European dogs. The networks of the two other major dog clades IV and VI are compatible with single origins that are linked to European wolves by the positions of PICs and of East European wolves. In fact, the Eneolithic dog PIC-4 joined the central node of clade IV, distantly related to a node occupied by the Late Upper Pleistocene PIC-1 sequence, while Romanian-Russian/Greek wolves clustered to the central node of clade VI, with the other Early Holocene sequence, PIC-2, in derived position. The apparent contrast between the central positions occupied by PICs and East European wolves and the Asian origins of the majority of dogs could reflect confounding effects of artificial breeding practices (there is evidence that the European dog breeds, stringently selected after the formalization of the breed concept in the 19th century, were derived from shared stock [Parker et al. 2004]). Although not containing PICs, it is worth mentioning that the composition of the remaining minor dog clades II and IX suggests recent geographically defined origins in Europe and Asia, respectively.

The above-discussed links between PICs and East European wolves are supported by the wolf phylogeographic network. In this respect, it is intriguing that the two wolf clusters containing PICs hold all the 262-bp HVR1 sequences shared between dog and wolf (five reported in East European wolves, one in wolves from Mongolia/China). Some of these shared sequences were previously attributed to dog gene flow into wolves (Randi et al. 2000). However, maternal introgressions of extant dogs into wolves following natural hybridization appear to be rare because of physiological, behavioral, and ecological barriers (Vilà and Wayne 1999). The interpretation of the dog-shared East European wolf sequences as representatives of true ancient European wolf genetic variation is supported by the close relationship between PIC-2 and the Romanian-Russian/Greek wolf sequence W6/W26 and by the fact that the Bulgarian wolf sequence W51 contains PIC-3.

We further investigated phylogeographic patterns among global wolf and/or dog HVR1 sequences. Geographic distances were not correlated with genetic distance values for dog or wolf sequences analyzed independently, even when analyses were restricted to clades and subclades. Lack of phylogeographic patterns in wolves, probably reflecting their high mobility, was also found by Vilà et al. (1999). On the other hand, we found that geographic distances calculated imposing a theoretical passage through the Bering Strait were correlated with genetic distances among European, West Asian, East Asian, African, North American, and Central American dogs and wolves analyzed together, suggesting genetic differentiation related to isolation by geographic distance (Mayr 1942, 1963; Wright 1943).

Our data are in agreement with the notion that wolf and dog populations have been connected through Bering in the past (Clutton-Brock 1995). The nonrandom geographic distribution of wolf and dog HVR1 sequences could reflect phylogeographic patterns preceding the reproductive isolation of dogs from wolves. Our findings would also support
the hypothesis of multiple independent origins of dogs and/or of frequent interbreeding between early proto-dogs and wolves throughout a vast geographic range (Sablin and Khlopachev 2002). In fact, genetic separation between dogs and wolves is likely to have occurred only after the Neolithic agropastoral revolution (≈8,000 YA) that resulted in incompatibility between wolves and humans because of the presence of livestock (Lindner 1950; Clutton-Brock 1995; Sablin and Khlopachev 2002).

The last glaciation could have restricted Upper Pleistocene carnivores such as bears and hyenas to refugia (Hofreiter et al. 2004) but could have facilitated the dispersal of a highly mobile and adaptable predator such as the wolf, which can thrive in open steppe environments (Kurtén 1968). The mtDNA haplotypes of the extant East European wolves may represent a relic of the last glaciation. In fact, the relationships between PICs and East European wolves are consistent with the paleogeographic background (Fig. 4B). At Last Glacial Maximum (≈22–14,000 14C YA), because of the extensive regression of the Eastern Mediterranean Sea, Italy participated in a vast temperate steppe province extending eastward into the Danubian basin and southeastward, through a land bridge across the Bosphorus, into West Asia (Williams 1998). Due to the presence of ice sheets to the north, this southern steppe, that linked the distant locations of the earliest known dog sites in Europe and in the Middle East (Davis and Valla 1978; Nobis 1979; Benecke 1994; Dayan 1994; Tchernov and Valla 1997; Sablin and Khlopachev 2002), was the main route of faunal and human migration into Europe and West Asia (Sauter 1948; Kurtén 1968; Gamble 1986). As shown in figure 4B, the distribution of the PIC-related wolf haplotypes bridges the geographic gap between the earliest fossil dog sites.

In conclusion, the data show that the prehistoric canids of the Italian peninsula were genetically diverse and not closely related to the extant Italian wolves. Genetic data obtained comparing the ancient sequences with extant dog and wolf sequences and early archaeozoological evidences concur in suggesting that Late Glacial/Early Holocene wolf populations of the West Eurasian steppes (that stretched over South-Eastern Europe and West Asia) contributed to the origins of the dog (Davis and Valla 1978; Nobis 1979; Benecke 1994; Dayan 1994; Tchernov and Valla 1997; Sablin and Khlopachev 2002). Genetic data also suggest multiple independent Asian and European domestication events. In spite of the severe global decline of wolves, extant East European wolf populations may still carry genetic signatures of dog-founder populations and deserve dedicated study and conservation efforts.

The question of dog origins is geographically, genetically, and archaeologically complex and clearly requires additional multidisciplinary studies. In this respect, further a-DNA evidence from relevant areas of the world should allow to better understand the evolution of wolves and dogs.

Supplementary Material

The following tables are available online: Table S1 (frequencies, geographic origins, GenBank accession numbers, and references of the extant wolf and coyote HVRI sequences [262–664 bp, nps 15431–15687/16039]); Table S2 (frequencies, geographic origins, GenBank accession numbers, and references of the purebred dog HVRI sequences [262–664 bp, nps 15431–15687/16039]); Table S3 (frequencies and geographic origins of unique 262-bp HVRI sequences [nps 15431–15687] of extant wolves [excluding sequences shared with dogs]); and Table S4 (frequencies and geographic origins of unique 262-bp HVRI sequences [nps 15431–15687] of purebred dogs [including sequences shared with wolves]).

The following figures are available online: Fig. S1 (A, global map with sampling locations for the dog and wolf HVRI sequences studied; B, plot of the genetic distance vs. the geographic distance between dog and wolf populations sampled in Europe [EU], West Asia [WA], East Asia [EA], Africa [AF], North America [NA], and Central America [CA]).

The sequences reported in this paper have been deposited in the GenBank database (accession numbers AY741666–AY741670).

Acknowledgments

This work was supported by the Italian National Research Council (CNR, Project Beni Culturali-Archivio Biologico coordinated by A. Guarino) and by the Ministero dell’Istruzione, dell’Università e della Ricerca (MIUR), Progetti di Ricerca di Interesse Nazionale (PRIN) grants 2000 and 2003, coordinated by L. R. Angeletti. We thank Carlés Vilà, University of Uppsala, Sweden; Giovanni Destro-Bisol, Department of Animal and Human Biology, University La Sapienza, Rome; and two anonymous reviewers for critical suggestions. We acknowledge the help of Associazione per lo Studio degli Animali Estinti (ASAE) (www.asae-onlus.org).

Literature Cited

Cassoli, P. F., I. Fiore, and A. Tagliacozzo. 1994. Butchering and differentiated introduction of main mammals in the


Liedloff, A. 1999. Mantel version 2.0 Mantel non-parametric test calculator. School of Natural Resource Sciences, Queensland University of Technology, Brisbane, Australia.


Schneider, S., D. Roessli, and L. Excoffier. 2000: a software for population genetics data analysis. Genetics and Biometry Laboratory, University of Geneva, Switzerland.


William Martin, Associate Editor

Accepted August 16, 2005