

## ORIGINAL ARTICLE

# Using multiple markers to elucidate the ancient, historical and modern relationships among North American Arctic dog breeds

SK Brown<sup>1,2</sup>, CM Darwent<sup>2</sup>, EJ Wictum<sup>1</sup> and BN Sacks<sup>1,3</sup>

Throughout most of the Americas, post-colonial dogs largely erased the genetic signatures of pre-historical dogs. However, the North American Arctic harbors dogs that are potentially descended from pre-historical ancestors, as well as those affected by post-colonial translocations and admixtures. In particular, Inuit dogs from Canada and Greenland are thought to descend from dogs associated with Thule peoples, who relied on them for transportation ca. 1000 years ago. Whether Thule dogs reflected an earlier colonization by Paleoeskimo dogs ca. 4500 years ago is unknown. During the Alaskan Gold Rush, additional sled dogs, possibly of post-colonial derivation, the Alaskan Husky, Malamute and Siberian Husky, were used in the Arctic. The genealogical relationships among and origins of these breeds are unknown. Here we use autosomal, paternal and maternal DNA markers to (1) test the hypothesis that Inuit dogs have retained their indigenous ancestry, (2) characterize their relationship to one another and to other Arctic breeds, and (3) estimate the age of North American indigenous matriline and patriline. On the basis of the agreement of all three markers we determined that Inuit dogs have maintained their indigenous nature, and that they likely derive from Thule dogs. In addition, we provide support for previous research that the Inuit dogs from Canada and Greenland dog should not be distinguished as two breeds. The Alaskan Husky displayed evidence of European introgression, in contrast to the Malamute and Siberian Husky, which appear to have maintained most of their ancient Siberian ancestry.

*Heredity* (2015) **115**, 488–495; doi:10.1038/hdy.2015.49; published online 24 June 2015

## INTRODUCTION

The origins of dogs in the Arctic are complex. Although the first dogs arrived the Americas ca. 12 000 years ago (Leonard *et al*, 2002), people and their dogs did not settle in the Arctic until ca. 4500 years ago (Paleoeskimo) and then again 1000 years ago (Thule) (Morey & Aaris-Sorensen 2002; Morey 2010; Gilbert *et al*, 2008; Raghavan *et al*, 2014). It is unclear whether the dogs associated with these two Arctic cultures arrived on the continent in two separate waves of colonization along with their human counter parts (see Raghavan *et al*, 2014), or if the original Paleoeskimo dogs served as founding stock for the later Thule dogs. It is also possible that Paleoeskimo dogs themselves arose from earlier Native American dogs initially brought to the continent at the end of the last ice age (Leonard *et al*, 2002). Thus, one of our most basic questions pertained to the roots of indigenous North American Arctic dogs.

Second, it is unclear to what extent today's Arctic dogs reflect that indigenous ancestry. Throughout most of the Americas, post-colonial dogs largely erased the genetic signatures of pre-historical dogs (Leonard *et al*, 2002). Today's Inuit dogs of Canada and Greenland seem to be an exception, most likely the descendants of pre-colonial indigenous dogs (Brown *et al*, 2013; van Asch *et al*, 2013). Since European colonization, however, inhabitants of the Arctic have adopted as sled dogs several additional breeds or types, including the Alaskan Husky (created in post-colonial North America), Alaskan

Malamute (hereafter, Malamute) and Siberian Husky. Although these types are of similar form and function to the Inuit dog, their genealogical relationships are unclear. The Alaskan Husky of today, which is defined on the basis of (and are bred for) function rather than ancestral 'purity', reflects a century or more of admixture (for example, Pointers, Shepherds, Salukis) and selection for performance (Coppinger and Coppinger, 2001; Huson *et al*, 2010). Malamutes were thought to be created by the Malemiut Inupiaq people of Alaska's Norton Sound region, and therefore presumably share a common ancestry with Inuit dogs (Handford 1998). The Malamute was used primarily for sledding but is heavier bodied than Inuit dogs, with different markings (Handford 1998). It is unknown to what extent native ancestry has been retained through breed development. The Siberian Husky was originally developed by the Chukchi people of the Chukotka Peninsula in eastern Siberia (Fiszdon & Czarkowska 2008) and was brought to Nome, Alaska, in 1908 for sled-dog sport racing.

Existing genetic work on these dogs provides an incomplete picture. Genomic studies have invariably excluded Inuit dogs (as well as other putative indigenous dogs), but suggest that other arctic spitz-type dogs (Malamutes, Siberian Huskies) form one of the most ancestral groups among the modern-defined breeds (Parker *et al*, 2004; vonHoldt *et al*, 2010; but see Larsen *et al*, 2012). Mitochondrial DNA studies have included a wider range of arctic dog breeds and types, including Inuit dogs, along with village dogs indigenous to a variety of non-Western

<sup>1</sup>Mammalian Ecology and Conservation Unit, Center for Veterinary Genetics, University of California, Davis, CA, USA; <sup>2</sup>Department of Anthropology, University of California, Davis, CA, USA and <sup>3</sup>Department of Population Health and Reproduction, School of Veterinary Medicine, University of California, Davis, CA, USA  
Correspondence: Dr SK Brown, Mammalian Ecology and Conservation Unit, Center for Veterinary Genetics, University of California, One Shields Avenue, Davis, CA 95616, USA.  
E-mail: skbrown@ucdavis.edu

Received 4 November 2014; revised 17 March 2015; accepted 28 April 2015; published online 24 June 2015

regions (for example, Pang *et al*, 2009; Brown *et al*, 2011, 2013; Sacks *et al*, 2013); these studies generally support the indigenous ancestry of Inuit dogs but are equivocal with respect to Malamute ancestry. In particular, genetic analysis of 582 base-pairs (bp) of mitochondrial DNA in modern (Savolainen *et al*, 2002; Pang *et al*, 2009; Klütsch *et al*, 2011; van Asch *et al*, 2013) and ancient (Brown *et al*, 2013) dogs demonstrated genetic continuity between Thule (pre-colonial) and modern Canadian and Greenland dogs by the continued presence and high prevalence of the North American haplotype A31, not found in modern Siberian dogs. The Malamute has been associated primarily with haplotype A29 (Klütsch *et al*, 2011; Brown *et al*, 2013; van Asch *et al*, 2013), which has been identified in prehistoric dog remains from sites in the Americas, in both Alaskan and Siberian huskies, and in Southeast Asian village dogs and Australian dingoes (Leonard *et al*, 2002; Savolainen *et al*, 2004; Brown *et al*, 2013; Sacks *et al*, 2013). The occurrence of this haplotype in indigenous populations on both continents renders it equivocal with respect to evidence for proximate origins in North America, but nevertheless supports the ancient origins of the spitz-type group in general.

The only study involving arctic dog Y-chromosomes was restricted to low-resolution sequence variants and included two Inuit dogs, along with two Malamutes, a Siberian Husky, and a Chukotka dog from Siberia (Ding *et al*, 2011). Both Malamutes and an Inuit dog shared a haplotype otherwise not found in any other of the hundreds of dogs genotyped throughout Eurasia (Ding *et al*, 2011; Sacks *et al*, 2013), suggesting that use of a higher-resolution Y-chromosome marker and a larger sample of arctic dogs could help further elucidate their ancestry.

Here, we apply maternal, paternal and biparentally inherited markers to assess the genetic history of today's North American arctic sled dogs. Our objectives were to (1) test the hypothesis that today's Inuit dogs have retained indigenous North American ancestry, (2) characterize their relationship to one another and to the more recently imported or constructed arctic dog breeds and (3) estimate the age of North American indigenous matriline and patriline. For this last objective, we used a 6024-bp portion of the mitochondrial genome and Y-chromosome microsatellite haplotypes to test coalescent predictions corresponding to three alternative hypotheses about the timing of arrival from Asia: (a) ca.12000 years ago, corresponding to the initial introduction of dogs to the continent (Leonard *et al*, 2002), (b) ca. 4500 years ago, consistent with a Paleoeskimo dog origin, or (c) ca. 1000 years ago, consistent with origin from a subsequent Thule migration from eastern Siberia.

## MATERIALS AND METHODS

### Samples

We collected buccal swabs from arctic-breed owners and breeders between 1998 and 2010 (sampled by Veterinary Genetics Lab Dog Research Group) and 2012–2013 (sampled by SKB). Altogether, we collected a total of 178 individual buccal swabs from Inuit dogs (Canada and Greenland), Alaskan Husky, Malamute and Siberian Huskies (Table 1). Our sample sizes varied among the types of marker used owing to insufficient DNA quantities for some samples, processing of different batches among markers in variable order, restriction of Y-chromosome analyses to males (obviously) and removal of closely related individuals (based on autosomal microsatellites) for expanded sequencing (Appendix 1 online).

### Laboratory procedures

We extracted DNA from buccal swabs using a sodium hydroxide protocol (Brown *et al*, 2011) or Qiagen DNeasy kit, according to manufacturer's instructions. We genotyped and analyzed a total of 24 ISAG (International Society for Animal Genetics) canine autosomal microsatellite loci (AHT121,

AHT137, AHTH130, AHTH171, AHTH260, AHTk211, AHTk253, C22.279, FH2001, FH2054, FH2328, FH2848, INRA21, INU005, INU030, INU055, LEI004, REN105L03, REN162C04, REN169D01, REN169O18, REN247M23, REN54P11, REN64E19) as previously described (Pedersen *et al*, 2013). For comparison in the population tree analysis, we included 12 wolves from Yukon Territory, Canada (Appendix 1 online).

We first sequenced the 582-bp portion of the mitochondrial DNA (mtDNA) hypervariable region I (HVI) using primers H15422 and L16106 (Savolainen *et al*, 2002). Among individual Inuit dogs found to possess the A31 haplotype, we selected for expanded sequencing a subset ( $n = 10$ ; seven from Canada, three from Greenland) that were maximally unrelated ( $r < 0.22$ ) based on microsatellite analysis (see below) to ensure that we represented the greatest genealogical diversity in the sample set. We targeted 6024 bp of mitochondrial DNA, including 5582 bp identified previously as the most highly polymorphic regions of the dog mtDNA genome (Imes *et al*, 2012), using primers listed in Supplementary Table S1. The resulting sequence was annotated using the online web server MITOS (Bernt *et al*, 2013).

For all mitochondrial amplicons, we performed polymerase chain reactions in an ABI 2720 thermal cycler in 23- $\mu$ l volumes consisting of 3  $\mu$ l DNA, 1  $\times$  PCR buffer, 3 mM MgCl<sub>2</sub>, 0.2 mM dNTPs, 0.5  $\mu$ M of each primer and 1 U of AmpliTaq DNA polymerase (Applied Biosystems, Grand Island, NY, USA). Conditions were set at 94–96 °C for 10 min, followed by 40 cycles of 94 °C for 15–45 s, 50–56 °C for 30–45 s and 72 °C for 45–60 s, followed by 10 min of extension at 72 °C. Amplicons were purified with EXO-SAP-it (Affymetrix, Santa Clara, CA, USA) as per manufacturer's protocol and were then sequenced with Big Dye v 3.1 chemistry on an ABI 3730 capillary sequencer. Sequences were aligned to a reference mitogenome (Kim, 1998). Naming convention for HVI haplotypes followed Savolainen *et al*. (2002).

We genotyped male dogs using five dinucleotide-repeat microsatellites from the non-recombining region of the Y chromosome, including 650–79.2, 650–79.3, 990–35 (Bannasch *et al*, 2005), MS34A and MS41B (Sundqvist *et al*, 2001), in two separate multiplex reactions as previously described (Brown *et al*, 2011). We also genotyped these dogs at 29 Y-chromosome single-nucleotide polymorphisms reported by Ding *et al*. (2011) using a Sequenom platform with primers and conditions described by Sacks *et al*. (2013).

### Data analyses

We estimated autosomal microsatellite diversity statistics ( $H_O$ ,  $H_E$ ) and allelic richness using Microsatellite Toolkit (Park 2001). We used GENEPOP v. 4.2 (Raymond & Rousset 1995) to test for departures from Hardy–Weinberg equilibrium and linkage disequilibrium and to estimate  $F_{ST}$ . In tests for Hardy–Weinberg equilibrium and linkage disequilibrium, we adjusted  $P$ -values for multiple comparisons using the sequential Bonferroni correction (Rice 1989). We estimated the inbreeding coefficient ( $F_{IS}$ ) in Arlequin v. 3.5.1.3 (Excoffier *et al*, 2005), and estimated relatedness among individuals using a maximum likelihood approach implemented in MLrelate (Kalinowski *et al*, 2006).

**Table 1** Sample sizes<sup>a</sup> used for data analyses involving hypervariable region I (HVI) mtDNA, expanded mtDNA sequences, Y-chromosome microsatellites and autosomal microsatellites

Breed	HVI mtDNA	Expanded mtDNA	Y Chromosome	Autosomal microsatellites
Inuit	59	10	31	33
Canada	45	7	22	25
Greenland	14	3	9	8
Alaskan Husky	10	—	15	16
Malamute	14	—	14	15
Siberian Husky	4	—	11	10
Total	87	10	71	74

<sup>a</sup>Sample sizes vary among marker types for several reasons, including insufficient DNA quantities, processing of different batches among markers in variable order, restriction of Y-chromosome analyses to males, removal of closely related individuals (based on autosomal microsatellites) for expanded sequencing.

To visualize patterns of genomic heterogeneity among individual Arctic dogs irrespective of prior categorization into a particular breed or type, we used a Bayesian model-based clustering approach implemented in program STRUCTURE v. 2.3.3 (Pritchard *et al*, 2000). We performed 100 independent runs using the correlated allele frequencies model (Falush *et al*, 2003), including 10 runs each for  $K=1-10$  genetic clusters. Each run was composed of 500 000 Markov Chain Monte Carlo (MCMC) repetitions from which the first 250 000 were discarded as burn-in. The final  $K$  (number of groups) was chosen that best fit the data (denoted by  $X$ ), by plotting the negative in  $\ln(-\ln(X/K))$  versus  $K$  ( $K$  ranging from 1 to 10). Finally, the  $K$  that was supported was then assessed to assure that this value did not sacrifice substantial explanatory power (Pritchard and Wen 2002; Waples and Gaggiotti 2006). To complement this analysis with one based on pre-categorized groups (that is, breeds/types), we also used program Populations (Langella 1999; <http://bioinformatics.org/~tryphon/populations/>) to estimate Nei's  $D_A$  (Takezaki & Nei 1996) and produce a neighbor-joining tree among populations with node support calculated bootstrapping on loci.

We estimated mitochondrial HVI haplotype diversity and mitochondrial pairwise  $F_{ST}$  in Arlequin. For expanded mitochondrial haplotypes, ideally, we would have measured the average (and s.d.) numbers of substitutions accumulating between individuals' haplotypes and the ancestral haplotype (Forster *et al*, 1996; Saillard *et al*, 2000). However, for recently derived lineages (for example, <1200 years), it was likely that we would find no mutations in the entire 6024 bp, which prevented use of this approach. Therefore, we turned this question around: instead of asking 'what is the coalescence time suggested by these data?' we asked 'what is the probability that these data (or more extreme values) would be observed if their true coalescence were <1000 years ago?, <5000 years ago?, <12 000 years ago?' Specifically, we treated the problem as one of decay, whereby the probability (or expected portion) of haplotypes having no mutations decreased with increasing coalescence time. We assumed a substitution rate of 2.5% per million years (Nabholtz *et al*, 2008), and determined the probability ( $P$ ) associated with the observed frequency of derived haplotypes (for example, 0) versus ancestral haplotypes (conservatively presumed to be the haplotype at highest frequency) corresponding to a range of the divergence times (1000–12 000 year), and used these values to identify the divergence time corresponding to the critical value at  $\alpha=0.05$ . Specifically, we used the binomial formula,

$$P = \binom{k}{x} * p^x * (1-p)^{(k-x)}$$

where  $k$  was the total number of dogs sequenced (that is, 10),  $x$  was the number of these dogs exhibiting the most frequent haplotype (for example, 10), and  $p$  varied according to the hypothetical divergence time (Figure 1).

We arranged Y-chromosome microsatellite alleles into five-locus haplotypes, which were nested within haplogroups defined according to single-nucleotide polymorphism haplotypes (hereafter, 'haplogroups' to distinguish from the more highly resolved, nested microsatellite haplotypes) and compared with published haplogroups and haplotypes (Brown *et al*, 2011; Ding *et al*, 2011; Sacks *et al*, 2013). We first constructed median-joining haplotype networks and then applied a reduced-median analysis to create the final networks, using program Networks (4.6.1.1.; Bandelt *et al*, 1995).

## RESULTS

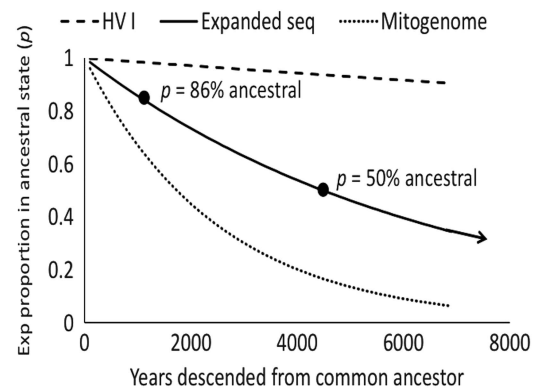
### Autosomal microsatellites

We observed an average of 4.91 alleles per locus among the 24 autosomal microsatellite loci, with breed-specific averages ranging from 3.38 to 6.58 (Table 2). Otherwise, all 24 loci were polymorphic in all breeds. Inuit dogs (both Canadian and Greenland) exhibited lower heterozygosity than the other breeds (Table 2). Tests of linkage disequilibrium revealed no locus-pairs in any of the breeds to be out of equilibrium. No loci showed significant deviations from Hardy–Weinberg equilibrium in Inuit dogs (both Canadian and Greenland), Alaskan Huskies, or Siberian Huskies. However, CC22.279 showed significant deviations from Hardy–Weinberg equilibrium in Malamutes. Four breeds showed heterozygote deficiency in FIS, when

calculated across loci, consistent with substructure within these breeds/types.

Our admixture analysis indicated the greatest support for  $K=4$  clusters, corresponding to each of the presumptive breeds (Figure 2). The structure appeared hierarchical. At  $K=3$ , the Alaskan and Siberian Huskies clustered together. At  $K=2$  all Inuit dogs grouped in one cluster versus Alaskan Huskies, Malamutes and Siberian Huskies in another.

Similarly to the admixture analysis, a population tree based on Nei's  $D_A$  indicated clustering of all Inuit dogs from Canada and Greenland with 100% bootstrap support and clustering the Malamute with 84% bootstrap support with this group, and the Yukon wolves and Siberian Husky with 64% bootstrap support (Supplementary Fig. S1). The pairwise  $F_{ST}$  estimates additionally indicated that Alaskan Huskies were more closely related to both Siberian Huskies and Malamutes than either of the latter breeds were to one another, consistent with high contributions from Siberian Huskies and Malamutes to the Alaskan Husky gene pool (Table 3).

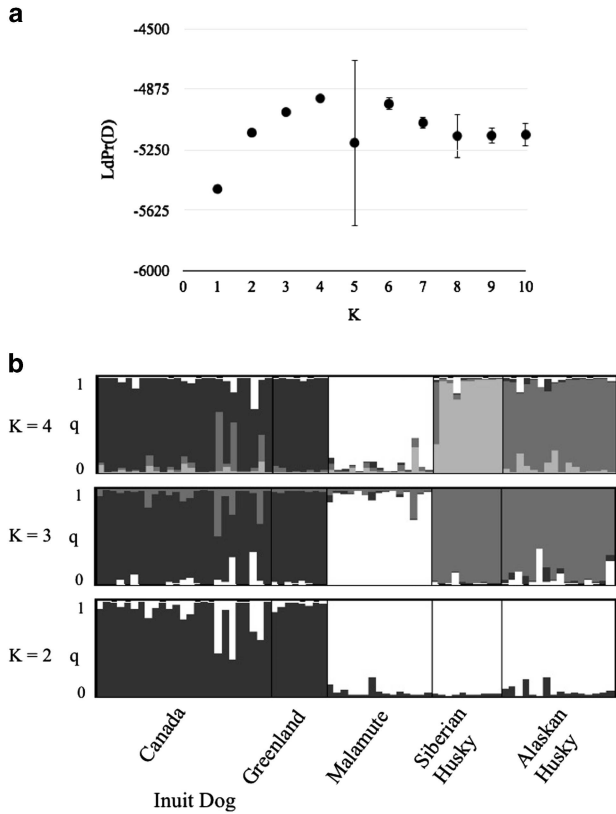


**Figure 1** Relationship between the expected proportion of mitochondrial haplotypes identical in state to the ancestral haplotype ( $p$ ) (as opposed to all mutant forms  $(1-p)$ ) as a function of time since divergence from a common ancestor for the expanded sequence (6024 bp). Filled circles and numeric values on the expanded-sequence curve correspond to two hypothetical time points, 1000 years, approximately when Thule culture and dogs spread east across North America, and 4400 years, corresponding to the Paleoekimo migration from Siberia. The expected proportion of ancestral haplotypes ( $p$ ) corresponding to 12 000 years ago (not shown) was 16% ( $P=0.16$ ). Curves for the hypervariable region I (HVI) fragment (582 bp) and the entire mitochondrial genome (16 000+bp) are shown for reference.

**Table 2** Autosomal microsatellite statistics, including number of dogs genotyped ( $n$ ), expected (assuming HWE) heterozygosity ( $H_E$ ), expected heterozygosity s.d., observed (proportion of loci counted in heterozygous state) heterozygosity ( $H_O$ ), observed heterozygosity s.d., number (No.) of alleles, inbreeding coefficient ( $F_{IS}$ ),  $F_{IS}$   $P$ -value

Breed	$n$	$H_E$	$H_E SD$	$H_O$	$H_O SD$	No. alleles	$F_{IS}$	$F_{IS} P$
Inuit Dog	33	0.56	0.044	0.51	0.018	5.67	0.112	0.000*
Canada	25	0.55	0.048	0.51	0.021	5.29	0.088	0.005*
Greenland	8	0.53	0.045	0.51	0.037	3.38	0.024	0.431
Alaskan Husky	16	0.72	0.034	0.68	0.027	6.58	0.056	0.046*
Malamute	15	0.68	0.029	0.56	0.026	4.88	0.206	0.000*
Siberian Husky	10	0.67	0.027	0.64	0.031	4.46	0.071	0.086

\*Significant  $P$ -value.



**Figure 2** Admixture analysis based on 24 autosomal microsatellites of four Arctic dog breed/types, including (a) the average ( $\pm$ s.d.) estimated log probability of the data ( $\text{Pr}(X|K)$ ) versus the specified number of genetic clusters ( $K$ ) calculated from 10 replicate runs for each  $K$  value and (b) bar plot illustrating for each dog the ancestry ( $q$ ) apportioned to each of the  $K=2-4$  genetic clusters, arranged in breed groups. The most parsimonious number of clusters supported by the increase in  $\text{Pr}(X|K)$  per increase  $K(A)$  was  $K=2$  and the maximal partition supported by these criteria was  $K=4$ .

### Mitochondrial DNA

We obtained 88 HVI sequences (~582 bp) that represented 12 haplotypes from four arctic breeds (Table 4). Mitochondrial haplotype diversity was relatively high in Alaskan Huskies, Malamutes and Siberian Huskies, and lower in Inuit dogs (Table 5). The Inuit dogs were the only group to possess haplotype A31. Inuit dogs also possessed another private type, A181 ( $n=2$ ). Haplotype A29 occurred in all breeds. As with the microsatellite analyses, pairwise mitochondrial  $F_{ST}$  comparisons indicated considerable differentiation ( $F_{ST}>0.42$ ) between haplotype frequencies of Inuit dogs versus Alaskan Huskies, Malamutes and Siberian Huskies (Table 3). Inuit dogs had similarly low haplotype diversity (Table 5) relative to the higher diversity observed in the other three breeds.

Among individual Inuit dogs with A31 HVI haplotypes, we selected the 10 (seven from Canada, three from Greenland) least closely related individuals in our data set for expanded mtDNA sequencing (Supplementary Table S3). We detected no polymorphisms among these individuals in the 6024-bp fragment, which we deposited in Genbank (Accession No. KM821048). A Basic Local Alignment Search Tool (BLAST) search revealed no other matching haplotype and differed by the nearest completely overlapping haplotype by 11 substitutions (0.18%; Thalmann *et al*, 2013; Genbank No. KF661086). This haplotype was a 1000-year-old dog from Florida with the corresponding HVI haplotype matching A3, which is also the

**Table 3** Microsatellite pairwise  $F_{ST}$  (above diagonal) estimates calculated from 24 autosomal microsatellites among four Arctic breeds. Below the diagonal are the mitochondrial HVI DNA  $F_{ST}$  estimates among four Arctic breeds. Sample sizes (in italics) for microsatellite  $F_{ST}$  are on the right side of the ' $\backslash$ ' and the mitochondrial  $F_{ST}$  on the left hand side

	<i>Inuit dog</i>	<i>Canada</i>	<i>Greenland</i>	<i>Alaskan Husky</i>	<i>Malamute</i>	<i>Siberian Husky</i>
Inuit	59133	0.009	0.046*	0.106*	0.134*	0.163*
Canada	0.014	45125	0.102*	0.119*	0.149*	0.169*
Greenland	0.005	0.036	1418	0.110*	0.131*	0.190*
Alaskan Husky	0.533*	0.538*	0.428*	10116	0.067*	0.054*
Malamute	0.530*	0.542*	0.442*	0.006	14115	0.113*
Siberian Husky	0.460*	0.477*	0.428*	0.028	0.006	4110

\* $P < 0.05$ .

nearest known HVI haplotype basal to A31 (Pang *et al*, 2009). The divergence time corresponding to the observed frequency of haplotypes (that is, 10 ancestral, 0 mutant) at  $P=0.05$  was estimated at 1725 years, indicating that this was the oldest statistically supported time since divergence from a common ancestor. In particular, the  $P$ -values corresponding to 1000 years, 4400 years and 12 000 years were  $P=0.21$ ,  $P=0.001$  and  $P=9 \times 10^{-9}$ , respectively, providing support for coalescence on the Thule timeframe and evidence to reject coalescence on both the Paleoeskimo and late Pleistocene timeframes. Because these  $P$ -values assume a mutation rate of 2.5%, we also explored sensitivity to this assumption. We determined that the 4400-year coalescence timeframe would not have been statistically plausible (that is,  $P > 0.05$ ) for any mutation rate greater than 1% per site per million years, and the 1000-year timeframe would remain within statistical plausibility for mutation rates up to 4.8% per site per million years, indicating that our conclusions about the coalescence timeframe apply as long as the true substitution rate falls within a window of 1–4.8%.

### Y-chromosome DNA

We genotyped the 5 Y-chromosome microsatellites and 29 single-nucleotide polymorphisms of 71 male dogs and observed 14 Y-chromosome microsatellite haplotypes, a similar number to the number of mtDNA haplotypes. In contrast to mtDNA haplotypes, however, Y-chromosome haplotypes tended to be specific to a particular breed, with fewer haplotypes shared among breeds (Table 6). Sharing of Y haplotypes among breeds/types also differed from the mtDNA (Figure 3). For example, Inuit dogs from Canada and Greenland did not share Y-chromosome haplotypes, whereas Inuit dogs from Canada and Malamutes shared their most prevalent haplotype (H22-n33). Alaskan Huskies and Malamutes shared a single haplotype as did Malamutes and Siberian Huskies.

The 14 Y-chromosome microsatellite haplotypes were distributed among seven haplogroups, which facilitated comparison with previous studies, some of which only examined haplogroups (Ding *et al*, 2011). Greenland dogs exhibited haplotypes in the H23 and H23\* haplogroup, which was identified in dogs throughout the Middle East and East Asia, but corresponded to particular microsatellite haplotypes that had not been previously observed (Figure 3). Haplogroup H20, which we observed in Alaskan Huskies, Malamutes and Siberian Huskies, was previously reported only in a single Chukotka dog from Siberia



(Ding *et al*, 2011) and a mixed-breed dog in the USA (Sacks *et al*, 2013). Haplogroup H22, which we observed in Inuit dogs and Malamutes, was previously reported also only in North American Arctic dogs and in these two breeds/types (Ding *et al*, 2011). The Alaskan Huskies, which exhibited the highest Y-haplotype diversity, carried haplotypes in the H1/H2 haplogroup, which is closely associated with European breeds (Brown *et al*, 2011; Ding *et al*, 2011; Sacks *et al*, 2013). Four Siberian Huskies also possessed haplotypes in the H1/H2 haplogroup. Haplotype n65, found in an Inuit dog, differed from the basal n33 haplotype by a single mutation, suggesting an erroneous placement of this haplotype in the single-nucleotide polymorphism-structured network.

## DISCUSSION

We undertook the most comprehensive study to date of the genomic relationships among dogs of the North American Arctic. Ours was the first to assemble multilocus autosomal data on a substantial number of Inuit dogs, and consequently to enable the assessment of their contemporary relationships with one another and to more recently imported or created breeds/types. Moreover, our use of mtDNA and Y-chromosome markers provided an opportunity to assess ages of founder lineages on the basis of mutational accumulation.

Because previous ancient mtDNA analysis demonstrated genetic continuity between pre-colonial (~750 BP) Alaskan and Greenland dogs and modern Inuit dogs from Canada and Greenland, we knew at the outset of this study that the dominant maternal lineage composing these two arctic breeds was indigenous (Brown *et al*, 2013). However, the use of strictly maternally inherited DNA was insufficient to draw conclusions about the overall genomic constitution of these dogs and provided no information on paternal contributions, for example, from post-colonization European/Siberian dogs in these putative indigenous breeds. Ideally, we could have compared nuclear markers directly between ancient (pre-colonial) and modern samples. However, recovery of sufficient amounts of nuclear DNA from ancient samples was not feasible. Therefore, we instead relied on nuclear DNA solely from modern Inuit dogs, in reference to that of the known post-

colonial North American arctic dogs. Under the hypothesis that Inuit dogs maintained their original pre-colonial indigenous ancestry, we predicted that they would exhibit greater genetic distance from the post-colonial breeds/types than from one another, that the post-colonial breeds would cluster together relative to the putative indigenous breeds, that the putative indigenous breeds would exhibit distinct paternal (and maternal) ancestry and that the mutational accumulation in apparently North American-endemic maternal and paternal lineages would reflect a timeframe consistent with pre-colonial origins.

All markers supported these predictions, consistent with the indigenous ancestry of Inuit dogs. Both mtDNA and autosomal microsatellites indicated no detectable genomic differentiation between Inuit dogs from Canada and Greenland, and that they were distinct from Siberian Huskies, Alaskan Huskies and Malamutes, all of which clustered together.

The Y-chromosome findings were less straightforward, but nevertheless were consistent with indigenous patriline in the two putative indigenous breeds in that the microsatellite haplotypes clustered together but reflected mutational accumulation consistent with pre-colonial origins (see below). Haplogroups and/or haplotypes associated with these breeds also were distinctive and not previously reported in any other dog breed or type (except for Malamutes, but see below; Brown *et al*, 2011; Ding *et al*, 2011; Sacks *et al*, 2013).

In contrast to the other markers, however, Inuit dogs from Canada versus Greenland carried different Y chromosome in distinct haplogroups indicating that at least the modern representatives reflected distinct paternal lines. This observation could reflect recent breeding practices. Although ultimately derived from Canadian or Greenlandic Inuit-owned dogs, all of our samples were supplied by individuals who breed these animals with the intent of preserving their ancestry but who control mate-pairings in a way that is not necessarily reflective of their historical development in the context of Inuit (and Thule) culture. In Asian randomly breeding village dog populations, as with wild canid populations, paternal and maternal lines tend to be similarly heterogeneous (Brown *et al*, 2011; Sacks *et al*, 2008, 2013).

**Table 4** Numbers of individuals of four modern Arctic sled dog breeds/types exhibiting 13 mitochondrial HVI haplotypes (582 bp), with reference to previous identification in Arctic or Spitz type breeds

Haplotype	Inuit dog		Alaskan Husky	Malamute	Siberian Husky	Total	Comments
	Canada	Greenland					
A31	32	10				43	Inuit, Greenland dogs <sup>a, b, c</sup>
A181	2					2	Inuit dog <sup>d</sup>
A20	6					6	UT, Greenland, Inuit dogs <sup>a, d</sup>
A49		1				1	Finnish Laika; indel from A49 (Laikas), Karelian Bear dogs <sup>d</sup>
A18	3		1			4	UT <sup>a, b</sup> , but not identified in any other spitz-type northern dogs; probable ancient haplotype
A11	1 <sup>e</sup>		1	1 <sup>f</sup>	2 <sup>f</sup>	5	UT <sup>a, d</sup> , ancient N. American dogs <sup>c</sup>
A29	1	3	3	6	2 <sup>a</sup>	15	Widespread in East Asia and ancient North America <sup>a, b, c, d</sup>
A179			1	2		3	Alaskan Malamutes, Siberian Husky <sup>d</sup>
A2			2	1		3	UT <sup>a</sup> , Eurasian spitz-type dogs <sup>a, d</sup>
E2			2			2	Siberian Laika <sup>d</sup>
C1				1		1	UT, Siberian Huskies, Samoyed <sup>a</sup>
A17				3		3	UT, Siberian aboriginal sled dogs, Alaskan Malamute, Greenland dog, Samoyed <sup>a, b, d</sup>
Total	45	14	10	14	4	88	

UT refers to a universally occurring haplotype (Pang *et al*. 2009).

<sup>a</sup>Savolainen *et al*. 2002.

<sup>b</sup>Pang *et al*. 2009.

<sup>c</sup>Leonard *et al*. 2002.

<sup>d</sup>Klüttsch *et al*. 2011.

<sup>e</sup>Sequence also consistent with haplotypes A142 and A166 (Pang *et al*. 2009).

<sup>f</sup>Sequence also consistent with haplotypes A142, A226, A166 and A89 (Pang *et al*. 2009)

However, post-Victorian western breeding practices tend to use many fewer males than females in the creation of breeds, presumably because males can sire many more litters than females can produce in a lifetime (Sundqvist *et al*, 2006). This practice has resulted in a pattern whereby mtDNA haplotypes tend to be distributed widely across breeds but Y-chromosome haplotypes tend to be much more breed-specific (Bannasch *et al*, 2005; Brown *et al*, 2011; Pedersen *et al*, 2013). In addition, it is possible that the difference between a Canadian versus Greenland Inuit dog is defined solely on the basis of the sire, which would explain the contrasting patterns observed in the Y-chromosome and mtDNA between these breeds. In fact, the inability to differentiate them with the 24 autosomal microsatellites implies that they have not been systematically bred separately (at least not for many generations) despite the correspondence between nominal country of origin and paternal haplotype. Thus, we conclude

**Table 5 Mitochondrial HVI DNA diversity statistics, including number of individuals sequenced (*n*), number (No.) of haplotypes, haplotype diversity and nucleotide diversity**

Breed	<i>n</i>	No. haplotypes	Haplotype diversity ( <i>s.d.</i> )	Nucleotide diversity ( <i>s.d.</i> )
Inuit	59	7	0.47 (0.076)	0.0037 (0.0023)
Canada	45	6	0.47 (0.085)	0.0037 (0.0023)
Greenland	14	3	0.47 (0.136)	0.0036 (0.0024)
Alaskan Husky	10	6	0.88 (0.075)	0.0082 (0.0049)
Malamute	14	6	0.80 (0.090)	0.0056 (0.0035)
Siberian Husky	4	2	0.66 (0.204)	0.0034 (0.0028)

that the Inuit dog and Greenland dog are synonymous and simply contain two paternal haplogroups.

Our findings were also informative about the post-colonial North American breeds. Microsatellite, mitochondrial and Y-chromosome markers indicated a close relationship among Alaskan Huskies, Malamutes and Siberian Huskies. As expected on the basis of their documented breed history during and since the Alaskan Gold Rush (for example, Coppinger and Coppinger, 2001; Huson *et al*, 2010), Alaskan Huskies appeared genetically intermediate between Malamutes and Siberian Huskies, most likely because of these breeds contributing to their founding, and also exhibited numerous mitochondrial and Y-chromosome haplotypes previously identified in Siberian Huskies and among a variety of European breed dogs (for example, Pang *et al*, 2009; Klütsch *et al*, 2011; Brown *et al*, 2011, 2013; van Asch *et al*, 2013).

In contrast, Malamutes did not share any of their Y-chromosome haplotypes with western breeds, and, although generally less breed-specific, their mitochondrial haplotypes were all previously reported in other Arctic or Spitz type breeds, suggesting that they retained ancient Arctic dog ancestry. The Y-chromosome haplotype that was common to the Malamute, Alaskan Husky and Siberian Husky (7e) was subsequently found also in two Chukotka sled dogs from Siberia (unpublished data), confirming very recent (that is, post-colonial) shared ancestry of these breeds with dogs indigenous to Asia. The close clustering of Malamutes and Siberian Huskies (and Alaskan Huskies) based on microsatellites and mtDNA relative to Inuit dogs also was consistent with a primarily post-colonial derivation.

Lastly, the finding that both uniparentally inherited marker types from Inuit dogs were indigenous allowed us to use these markers to investigate the timing of their arrival from Siberia. On the basis of the expanded mitochondrial sequences corresponding to the HVI haplotype, A31, our statistical tests over a reasonable range of potential

**Table 6 Y-chromosome types, composed of haplogroups (defined by SNP haplotypes, beginning 'H') and microsatellite haplotypes (numerals and lower-case letters), in four Arctic breeds, and comments on whether microsatellite haplotypes were previously reported in Western breed dogs (WBD), Southeast Asia village dogs (SEAVD), other breeds, or were previously unreported (Novel)**

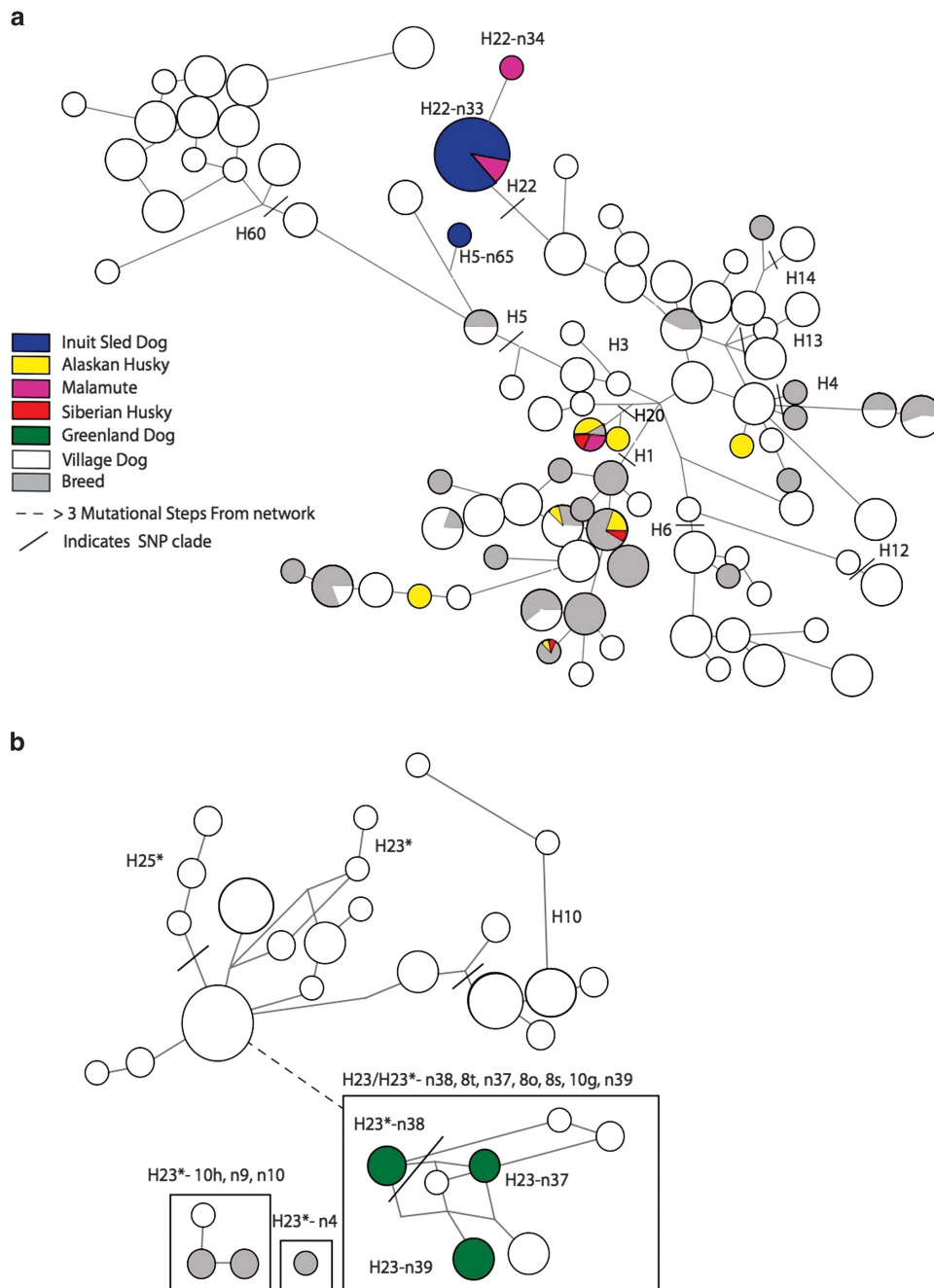
Y Type	Inuit dog		Alaskan Husky	Malamute	Siberian Husky	Total	Comments
	Canada	Greenland					
H1 <sup>a</sup>							
6q			5		1	6	5 WBDs <sup>b,d</sup> , 1 Dingo <sup>d</sup>
6zf			3		3	6	1 WBD <sup>b,d</sup>
7d			1			1	SEAVDs <sup>b,d</sup> , 1 WBD <sup>b,d</sup>
n32			2			2	Novel
H20 (7e)			2	1	7	10	Chukotka Sled dogs <sup>c</sup>
H22							
n33		21		9		30	Novel
n34				4		4	Novel
H3							
n36			1			1	Novel
H23							
n37		2				2	Novel
n39		4				4	Novel
H23* (n38)		3				3	Novel
H20 (n40)			1			1	Novel
H5 (n65)	1					1	

<sup>a</sup>H1 = H1, H1\*, or H2\* because we could not consistently distinguish these haplogroups.

<sup>b</sup>Brown *et al*. 2011.

<sup>c</sup>Ding *et al*. 2011.

<sup>d</sup>Sacks *et al*. 2013.



**Figure 3** Y-chromosome haplotypes of 67 Arctic from four breeds/types in the present study organized in networks with previously published (Gray and white haplotypes-Brown *et al.* 2011; Sacks *et al.* 2013) haplotypes of East Asian village dogs ( $n=144$ ), New Guinea Singing Dogs ( $n=18$ ), Australian Dingoes ( $n=94$ ), Iranian village dogs ( $n=135$ ), African village dogs ( $n=3$ ) and western breed dogs ( $n=79$ ), illustrating (a) Arctic breed dogs that group into haplogroups H1-H22 (Southeast Asian haplogroup) and (b) Arctic breed dogs (Greenland dogs) that group into haplogroup H23.

mutation rates enabled us to strongly reject the hypotheses that these maternal lines traced their coalescence back to either the late Pleistocene or Paleoeskimo timeframes and invariably supported a timeframe corresponding to the arrival and expansion of Thule culture across the North American Arctic. An important caveat of this analysis is that coalescence need not necessarily imply origin. It is conceivable that a major bottleneck would have resulted in loss of several ancestral A31 matriline leaving only a single one corresponding to the contemporary population. However, we think this unlikely for a couple

reasons. First, although a historical culling of Canadian Inuit dogs substantially reduced their numbers across Canada in the late 1900s (Montcombroux, 2002), this was geographically localized and did not affect Greenland. That we found the same haplotype in the 10 most unrelated individuals across in our broad sample suggests it more likely that the lack of ancient coalescence reflects the bottleneck associated with the initial founding from dogs brought over a millennium ago by the Thule. Nevertheless, mitogenomic studies done directly on pre-colonial material are needed to confirm this conclusion.

We have demonstrated the relationship of contemporary Arctic breeds through the use of maternal, paternal and autosomal markers. Inuit dogs from Canada and Greenland appear to compose a single population despite reflecting different paternal lineages, and are indigenous to pre-colonial North America. Our findings further provide preliminary evidence that these indigenous Inuit dogs derived from ancestors brought to North America by Thule migrants from Siberia and not from the earlier dogs occasionally found with Paleo Eskimo cultures or those initially brought to North America initial movements in the late Pleistocene.

#### DATA ARCHIVING

Data available from the Dryad Digital Repository: <http://dx.doi.org/10.5061/dryad.46170>.

#### CONFLICT OF INTEREST

The authors declare no conflict of interest.

#### ACKNOWLEDGEMENTS

Funding was provided by a Polar Programs NSF-ARC-0330981 to Darwent and Sacks. For laboratory assistance we thank Z Lounsbury, S Reponen, D Nisan, T Kun and J Collins. We also thank the Arctic dog breeders who provided samples. In particular, we would like to thank S Hamilton, F Robinson, N Russell, D Budd and T Muldoon, who provided samples and history of their breeds. Finally, we would like to thank three anonymous reviewers for their comments.

Bandelt HJ, Forster P, Sykes BC, Richards MB (1995). Mitochondrial portraits of human populations using median networks. *Genetics* **141**: 743–753.

Bannasch D, Bannasch M, Ryun J, Famula T, Pedersen N (2005). Y chromosome haplotype analysis in purebred dogs. *Mamm Genome* **16**: 273–280.

Bernt M, Donath A, Jühling F, Externbrink F, Florentz C, Fritzsche G *et al.* (2013). MITOS: improved *de novo* metazoan mitochondrial genome annotation. *Mol Phylogenet Evol* **69**: 313–319.

Brooks BA, Wallis SE (1998). *Alaskan Malamute—Yesterday and Today*. Alpine Press: Crawford, CO, USA.

Brown SK, Darwent CM, Sacks BN (2013). Ancient DNA evidence for genetic continuity in Arctic dogs. *J Archaeol Sci* **40**: 1279–1288.

Brown SK, Pedersen NC, Jafarshorijeh S, Bannasch DL, Ahrens KD, Wu JT *et al.* (2011). Phylogenetic distinctiveness of Middle Eastern and Southeast Asian village dog Y chromosomes illuminates dog origins. *PLoS One* **6**: e28496.

Coppinger R, Coppinger L (2001). *Dogs—A Startling New Understanding of Canine Origin, Behavior, & Evolution*. Scribner: New York, NY, USA.

Ding ZL, Oskarsson M, Ardalan A, Angleby H, Dahlgren LG, Tepeli C *et al.* (2011). Origins of domestic dog in Southern East Asia is supported by analysis of Y-chromosome DNA. *Heredity* **108**: 507–514.

Excoffier L, Laval G, Schneider S (2005). Arlequin version 3.0: An integrated software package for population genetics data analysis. *Evol Bioinform Online* **1**: 47–50.

Falush D, Stephens M, Pritchard JK (2003). Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* **164**: 1567–1587.

Forster P, Harding R, Torroni A, Bandelt H-J (1996). Origin and evolution of Native American mtDNA variation: a reappraisal. *Am J Hum Genet* **59**: 935–945.

Fiszdon K, Czarkowska K (2008). Social behaviours in Siberian huskies. *Annals of Warsaw University of Life Sciences – SGGW. Anim Sci* **45**: 19–28.

Handford JM (1998). Dog sledging in the eighteenth century: North America and Siberia. *Polar Record* **34**: 237–248.

Huson HJ, Parker HG, Runstadler J, Ostrander EA (2010). Genetic dissection of breed composition and performance enhancement in the Alaskan sled dog. *BMC Genet* **11**: 71.

Imes DL, Allard M, Wictum B, Penedo C, Sacks BN (2012). Identification of single nucleotide polymorphisms within the mtDNA genome of the domestic dog to discriminate individuals with common HVI haplotypes. *Forensic Sci Int Genet* **6**: 630–639.

Kalinowski ST, Wagner AP, Taper ML (2006). ML-RELATE: a computer program for maximum likelihood estimation of relatedness and relationship. *Mol Ecol Notes* **6**: 576–579.

Kim KS, Lee SE, Jeong HW, Ha JH (1998). The complete nucleotide sequence of the domestic dog (*Canis familiaris*) mitochondrial genome. *Mol Phylogenet Evol* **10**: 210–220.

Klütsch CFC, Seppälä EH, Fall T, Uhlén M, Hedhammar A, Lohi H *et al.* (2011). Regional occurrence, high frequency but low diversity of mitochondrial DNA haplogroup d1 suggests a recent dog-wolf hybridization in Scandinavia. *Anim Genet* **42**: 100–103.

Langella O (1999). Populations 1.2. 30. CNRS UPR9034.

Larsen G, Karlsson D, Perri A, Webster M, Ho SYW, Peters J *et al.* (2012). Rethinking dog domestication by integrating genetics, archeology, and biogeography. *Proc Natl Acad Sci* **109**: 8878–8883.

Leonard J, Wayne RK, Wheeler J, Valadez R, Guillén S, Vilà C (2002). Ancient DNA evidence for Old World origin of New World dogs. *Science* **298**: 1613–1616.

Montcombroux G (2002). *The Canadian Inuit Dog: Canada's Heritage*. 2nd edn Whippoorwill eBooks Publishing: Montreal.

Morey DF, Aaris-Sorensen K (2002). Paleo Eskimo dogs of the Eastern Arctic. *Arctic* **55**: 44–56.

Morey D (2010). *Dogs: Domestication and the Development of a Social Bond*. Cambridge University Press: New York, NY, USA.

Nabholz B, Glemin S, Galtier N (2008). Strong variations of mitochondrial mutation rate across mammals—the longevity hypothesis. *Mol Biol Evol* **25**: 120–130.

Pang J-F, Klütsch C, Zou X-J, Zhang A-B, Luo L-Y, Angelby H *et al.* (2009). mtDNA data indicate a single origin for dogs south of Yangtze River, less than 16,300 years ago, from numerous wolves. *Mol Biol Evol* **26**: 2849–2864.

Park S (2001). *Microsatellite toolkit*. Department of Genetics, Trinity College, Dublin: Ireland.

Parker HG, Kim LV, Sutter NB, Carlson S, Lorentzen TD, Malek TB *et al.* (2004). Genetic Structure of the Purebred Dog. *Science* **304**: 1160–1164.

Pedersen NC, Liu H, Theilén GH, Sacks BN (2013). The effects of dog breed development on genetic diversity and the relative influences of performance and conformation breeding. *J Anim Breed Genet* **130**: 236–248.

Pritchard JK, Stephens M, Donnelly P (2000). Inference of population structure using multilocus genotype data. *Genetics* **155**: 945–959.

Pritchard JK, Wen W (2002). Documentation for STRUCTURE Software Version 2. Available from URL <http://pritch.bsd.uchicago.edu>.

Raghavan M, DeGiorgio M, Albrechtsen A, Moltke I, Skoglund P, Korneliusson TS *et al.* (2014). The genetic prehistory of the New World Arctic. *Science* **345**: 1255832.

Raymond M, Rousset F (1995). GENEPOP (version 1.2): population genetics software for exact tests and ecumenicism. *J Heredity* **86**: 248–249.

Rice WR (1989). Analyzing tables of statistical tests. *Evolution* **223**–225.

Sacks BN, Bannasch DL, Chomel BB, Ernest HB (2008). Coyotes demonstrate how habitat specialization by individuals of a generalist species can diversify populations in a heterogeneous ecoregion. *Mol Biol Evol* **25**: 1384–1394.

Sacks BN, Brown SK, Stephens D, Pedersen NC, Wu J-T *et al.* (2013). Y chromosome analysis of Dingoes and Southeast Asian village dogs suggests a neolithic continental expansion from Southeast Asia followed by multiple Austronesian Dispersals. *Mol Biol Evol* **30**: 1103–1118.

Saillard J, Forster P, Lynnerup N, Bandelt HJ, Nørby S (2000). mtDNA variation among Greenland Eskimos: the edge of the Beringian expansion. *Am J Hum Genet* **67**: 718–726.

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

Savolainen P, Leitner T, Wilton AN, Matisoo-Smith E, Lundeberg J (2004). A detailed picture of the origin of the Australian dingo, obtained from the study of mitochondrial DNA. *Proc Natl Acad Sci USA* **101**: 12387–12390.

Sundqvist AK, Ellegren H, Olivier M, Vilà C (2001). Y chromosome haplotyping in Scandinavian wolves (*Canis lupus*) based on microsatellite markers. *Mol Ecol* **10**: 1959–1966.

Sundqvist A-K, Björnerfeldt S, Leonard JA, Hailer F, Hedhammar A *et al.* (2006). Unequal contribution of sexes in the origin of dog breeds. *Genetics* **172**: 1121–1128.

Takezaki N, Nei M (1996). Genetic distances and reconstruction of phylogenetic trees from microsatellite DNA. *Genetics* **144**: 389–399.

Thalmann O, Shapiro B, Cui P, Schuenemann VJ, Sawyer SK, Greenfield DL *et al.* (2013). Complete mitochondrial genomes of ancient canids suggest a European origin of domestic dogs. *Science* **342**: 871–874.

van Asch B, Zhang AB, Oskarsson MC, Klütsch CF, Amorim A, Savolainen P (2013). Pre-Columbian origins of Native American dog breeds, with only limited replacement by European dogs, confirmed by mtDNA analysis. *Proc R Soc Biol Sci* **280**: 20131142.

vonHoldt BM, Pollinger JP, Lohmueller KE, Han E, Parker HG, Quignon P *et al.* (2010). Genome-wide SNP and haplotype analyses reveal a rich history underlying dog domestication. *Nature* **464**: 898–902.

Waples RS, Gaggiotti O (2006). INVITED REVIEW: what is a population? An empirical evaluation of some genetic methods for identifying the number of gene pools and their degree of connectivity. *Mol Ecol* **15**: 1419–1439.

Supplemental Information accompanies this paper on Heredity website (<http://www.nature.com/hdy>)