Patterns of gene flow in Eurasia do not explain LD decay in Central Asia

Gene flow clearly occurs in dogs, with population structure exhibiting isolation by distance everywhere, and admixture with European dogs in particular regions. Using f3 admixture statistics, we find that indeed all Old World populations geographically intermediate between two other populations appear to be genetically intermediate as well, consistent with this model of bidirectional gene flow (Table S1).

As Central Asia is a population intermediate between regions previously implicated in genetic theories of dog origins (East Asia and Middle East / Europe), we used ms simulations [1] (Table S2) to examine the effect of gene flow on central and peripheral populations to ensure that the intermediate position of Central Asia within Eurasia was not driving a spurious signature of a domestication origin in the LD decay data. We explored two alternate models (Fig S1). In the first, dogs arise in a single population; this population splits as dogs move westward (serial founder model). The non-founder populations each experience a bottleneck as they split off. Gene flow then occurs between the centrally located population (Population 2) and both the founder population (Population 1) and the most diverged population (Population 3). This is the pattern we would expect in Eurasia if dogs were from South East Asia and then had spread west (or were from Europe or the Middle East, and spread east). In the second model, dogs originate in a central population, two populations branch off from the central populations and after going through a bottleneck both the Eastern and Western population experience gene flow with the central population. This is the model we would expect if dogs originated in Central Asia and spread in both directions.

Under various levels of migration, in the first model Population 2 never exhibited lower LD than the founder population (Fig S2). As migration rates increased populations became more similar in LD curve. For model 2 the founding population was also the population with lowest LD. This is consistent with the interpretation that dogs arose in the Central Asia and spread outward, but not consistent with dogs arising in either East Asia or Europe / Middle East.

More complicated patterns of admixture may influence patterns of LD in more complex ways. Perhaps the most likely way for LD to be reduced in a region (apart from being a domestication origin) would be substantial ancient admixture from an archaic population. While there is evidence of localized admixture with wolves in our dataset (Fig S3) and previously published datasets [14], we did not see any D tests where Central Asian dogs were outliers, indicating outsized gene flow from another population. East Asian dogs showed clear evidence of admixture with Chinese wolves, and all other dog populations showed evidence of admixture with Middle Eastern wolves.

We note that Freedman et al 2014 demonstrated that admixture with Middle Eastern wolves was bidirectional, although TreeMix only identified admixture from Middle Eastern wolves into dogs in our dataset (Fig 4). We believe the lack of a reciprocal admixture edge biases the TreeMix topology, artificially putting Middle Eastern wolves as the nearest neighbor to dogs. Other TreeMix models run with different numbers of admixture edges are inconsistent about the grouping of Middle Eastern wolves with dogs. In some models, Chinese wolves are closest (but no admixture edge is seen between Chinese wolves and dogs despite highly significant D-tests supporting such admixture) and in others wolves are monophyletic. Thus, while we believe admixture edges shown in Figure 4 reflect real genetic patterns, some admixture events are not robustly detected by TreeMix, leading to inconsistent topologies at the most basal branches of dogs and wolves.

Table S1: Significant F3 tests. We tested all combinations of village dog populations from: Afghanistan, Africa, Europe, India, Middle East, Mongolia, Nepal, and Vietnam and we report the significant results.

Source1	Source2	Target	f3	Z	Lower	Upper	p-value
Europe	India	Middle East	-0.00919	-21.63	0.608	0.818	4.69E-104
Europe	Vietnam	Mongolia	-0.0160	-21.062	0.489	0.818	8.88E-99
Middle East	Vietnam	Mongolia	-0.0130	-20.119	0.623	0.772	2.52E-90
Vietnam	India	Nepal	-0.0124	-16.499	0.257	0.418	1.87E-61
Vietnam	Africa	Nepal	-0.0133	-16.192	0.303	0.58	2.87E-59
Middle East	Vietnam	Nepal	-0.0115	-15.497	0.531	0.767	1.82E-54
Vietnam	Africa	Mongolia	-0.0111	-13.385	0.299	0.553	3.70E-41
Europe	India	Afghanistan	-0.00849	-13.07	0.399	0.665	2.44E-39
Vietnam	Afghanistan	Mongolia	-0.00780	-11.134	0.191	0.273	4.29E-29
Europe	Vietnam	Nepal	-0.0102	-10.896	0.374	0.818	6.02E-28
Vietnam	Afghanistan	Nepal	-0.00787	-10.392	0.196	0.362	1.35E-25
Europe	Vietnam	Afghanistan	-0.00607	-7.276	0.614	0.872	1.72E-13
Middle East	Vietnam	Afghanistan	-0.00441	-6.266	0.812	0.955	1.85E-10
Europe	Nepal	Afghanistan	-0.00376	-6.046	0.383	0.711	7.42E-10
Europe	Nepal	Middle East	-0.00264	-6.035	0.598	0.856	7.95E-10
India	Mongolia	Nepal	-0.00310	-5.439	0.254	0.59	2.68E-08
Europe	Africa	Middle East	-0.00135	-4.247	0.26	0.558	1.08E-05
Africa	Mongolia	Middle East	-0.00145	-4.081	0.542	0.072	2.24E-05
Vietnam	Africa	Afghanistan	-0.00322	-3.665	0.133	0.403	0.000124
Europe	Vietnam	Middle East	-0.00142	-2.231	0.749	0.927	0.0128

---

[1] Hudson, R. R.(2002) Generating samples under a Wright-Fisher neutral model. Bioinformatics 18:337-8

Table S2: Code for reported ms simulations

Migration Rate	Model 1	Model 2
2.5 x 10-6	ms 36 100 -t 80 -r 80 100000 -I 3 12 12 12 -n 1 1 -n 2 1 -n 3 1 -m 2 1 0.2 -m 2 3 0.2 -m 1 2 0.2 -m 3 2 0.2 -eM 0.01 0 -en 0.025 3 0.01 -ej 0.0251 3 2 -en 0.0495 2 0.01 -ej 0.05 2 1	ms 36 100 -t 80 -r 80 100000 -I 3 12 12 12 -n 1 1 -n 2 1 -n 3 1 -m 2 1 0.2 -m 1 3 0.2 -m 1 2 0.2 -m 3 1 0.2 -eM 0.04 0 -en 0.049 3 0.01 -en 0.0495 2 0.01 -ej 0.0497 3 1 -ej 0.05 2 1
1.25 x 10-5	ms 36 100 -t 80 -r 80 100000 -I 3 12 12 12 -n 1 1 -n 2 1 -n 3 1 -m 2 1 1 -m 2 3 1 -m 1 2 1 -m 3 2 1 -eM 0.01 0 -en 0.025 3 0.01 -ej 0.0251 3 2 -en 0.0495 2 0.01 -ej 0.05 2 1	ms 36 100 -t 80 -r 80 100000 -I 3 12 12 12 -n 1 1 -n 2 1 -n 3 1 -m 2 1 1 -m 1 3 1 -m 1 2 1 -m 3 1 1 -eM 0.04 0 -en 0.049 3 0.01 -en 0.0495 2 0.01 -ej 0.0497 3 1 -ej 0.05 2 1
2.5 x 10-5	ms 36 100 -t 80 -r 80 100000 -I 3 12 12 12 -n 1 1 -n 2 1 -n 3 1 -m 2 1 2 -m 2 3 2 -m 1 2 2 -m 3 2 2 -eM 0.01 0 -en 0.025 3 0.01 -ej 0.0251 3 2 -en 0.0495 2 0.01 -ej 0.05 2 1	ms 36 100 -t 80 -r 80 100000 -I 3 12 12 12 -n 1 1 -n 2 1 -n 3 1 -m 2 1 2 -m 1 3 2 -m 1 2 2 -m 3 1 2 -eM 0.04 0 -en 0.049 3 0.01 -en 0.0495 2 0.01 -ej 0.0497 3 1 -ej 0.05 2 1
1 x 10-4	ms 36 100 -t 80 -r 80 100000 -I 3 12 12 12 -n 1 1 -n 2 1 -n 3 1 -m 2 1 8 -m 2 3 8 -m 1 2 8 -m 3 2 8 -eM 0.01 0 -en 0.025 3 0.01 -ej 0.0251 3 2 -en 0.0495 2 0.01 -ej 0.05 2 1	ms 36 100 -t 80 -r 80 100000 -I 3 12 12 12 -n 1 1 -n 2 1 -n 3 1 -m 2 1 8 -m 1 3 8 -m 1 2 8 -m 3 1 8 -eM 0.04 0 -en 0.049 3 0.01 -en 0.0495 2 0.01 -ej 0.0497 3 1 -ej 0.05 2 1
0.001	ms 36 100 -t 80 -r 80 100000 -I 3 12 12 12 -n 1 1 -n 2 1 -n 3 1 -m 2 1 80 -m 2 3 80 -m 1 2 80 -m 3 2 80 -eM 0.01 0 -en 0.025 3 0.01 -ej 0.0251 3 2 -en 0.0495 2 0.01 -ej 0.05 2 1	ms 36 100 -t 80 -r 80 100000 -I 3 12 12 12 -n 1 1 -n 2 1 -n 3 1 -m 2 1 80 -m 1 3 80 -m 1 2 80 -m 3 1 80 -eM 0.04 0 -en 0.049 3 0.01 -en 0.0495 2 0.01 -ej 0.0497 3 1 -ej 0.05 2 1
0.05	ms 36 100 -t 80 -r 80 100000 -I 3 12 12 12 -n 1 1 -n 2 1 -n 3 1 -m 2 1 4000 -m 2 3 4000 -m 1 2 4000 -m 3 2 4000 -eM 0.01 0 -en 0.025 3 0.01 -ej 0.0251 3 2 -en 0.0495 2 0.01 -ej 0.05 2 1	ms 36 100 -t 80 -r 80 100000 -I 3 12 12 12 -n 1 1 -n 2 1 -n 3 1 -m 2 1 4000 -m 1 3 4000 -m 1 2 4000 -m 3 1 4000 -eM 0.04 0 -en 0.049 3 0.01 -en 0.0495 2 0.01 -ej 0.0497 3 1 -ej 0.05 2 1

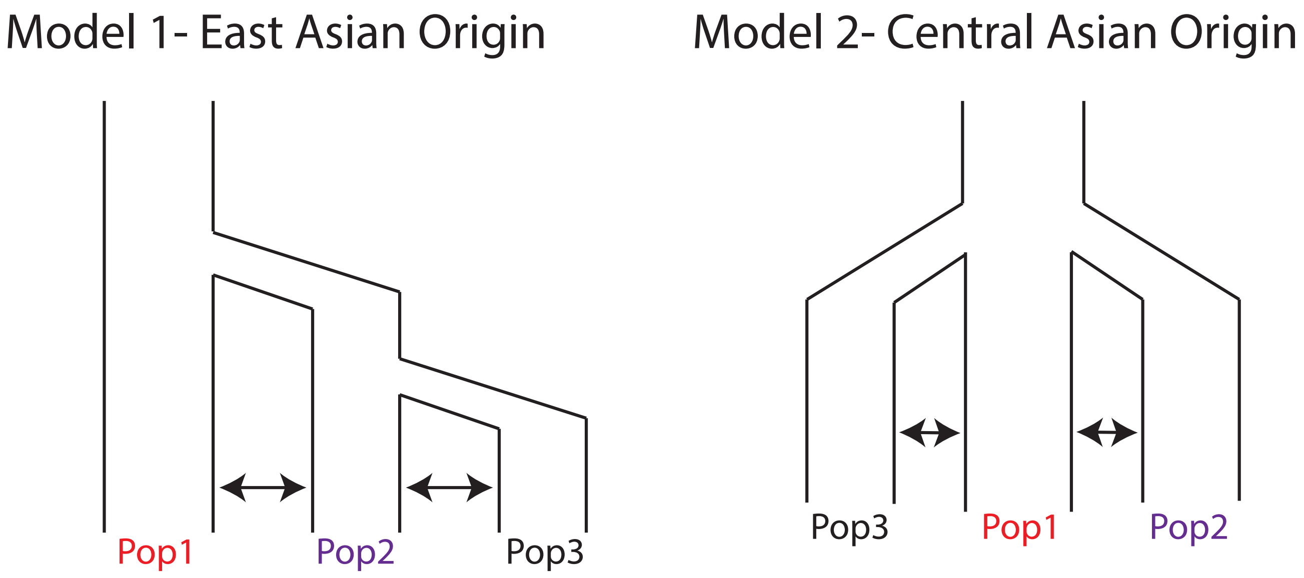

Fig. S1: Models for MS simulations.

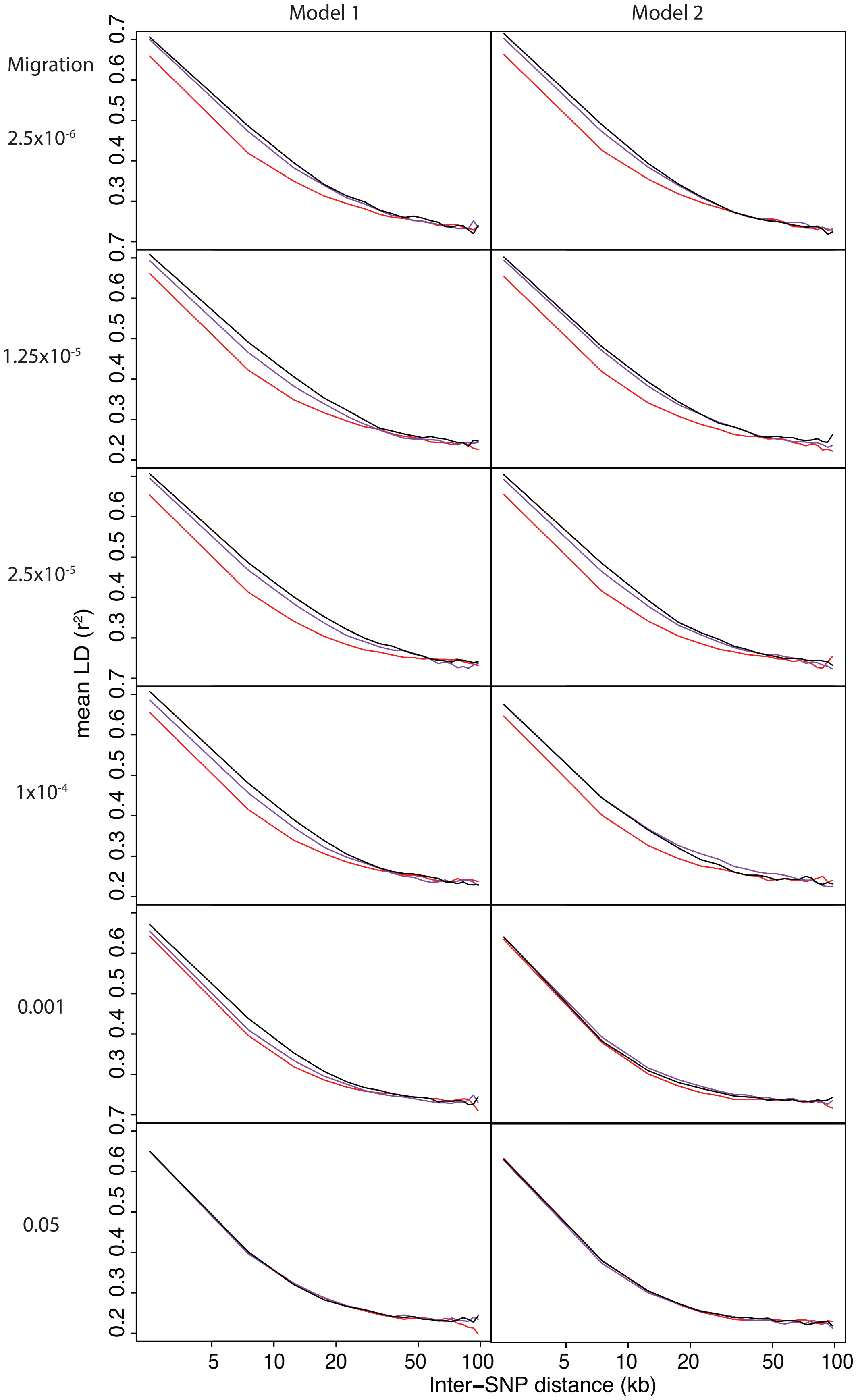

Fig. S2: LD curves for two contrasting models of dog history featuring various migration rates (percentage of sink population made up of source population per generation). In all cases the red population is the founding population.

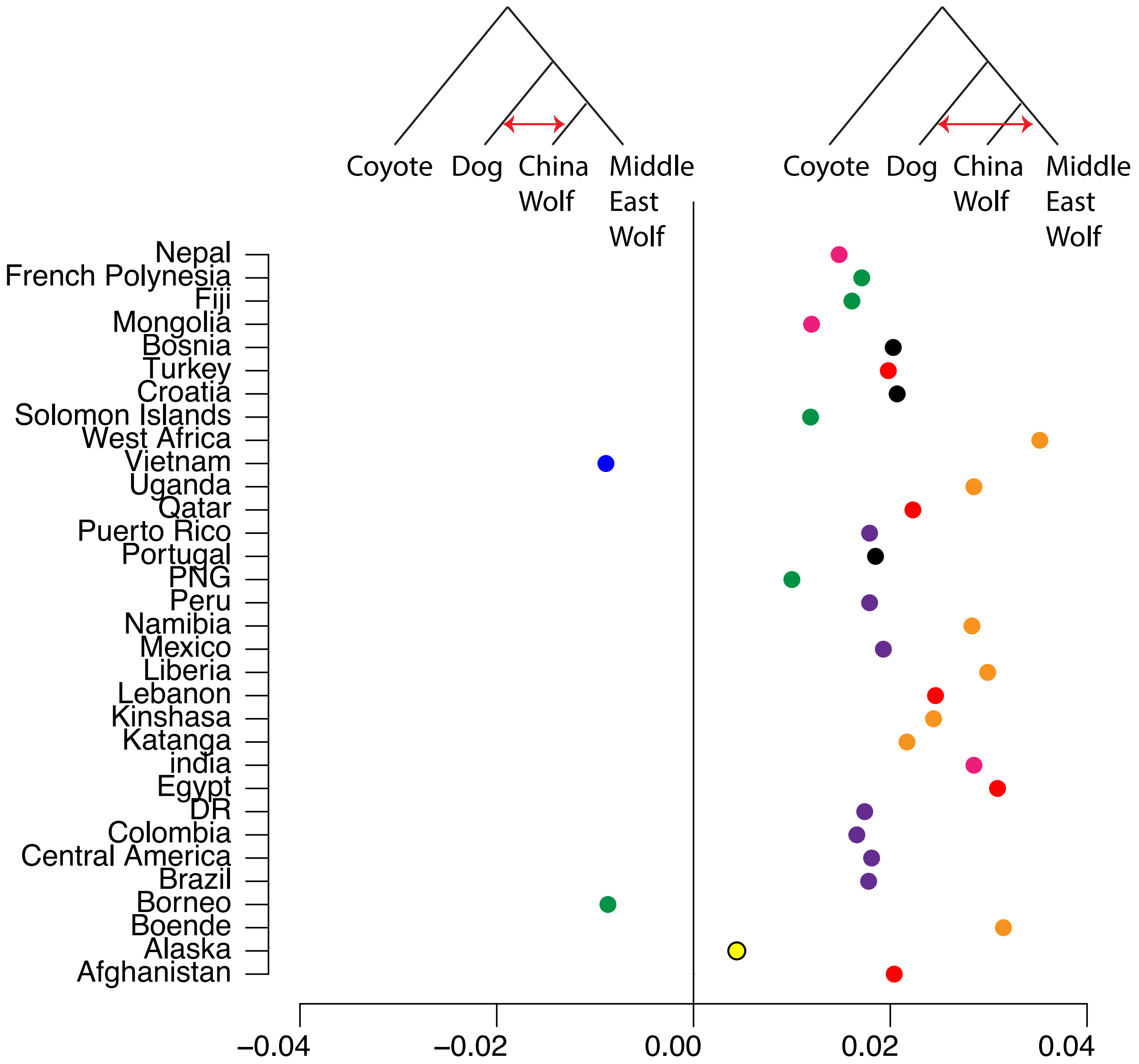

Fig. S3: D tests for all village dog populations in relation to Middle Eastern and Chinese wolves. Points are colored by region.

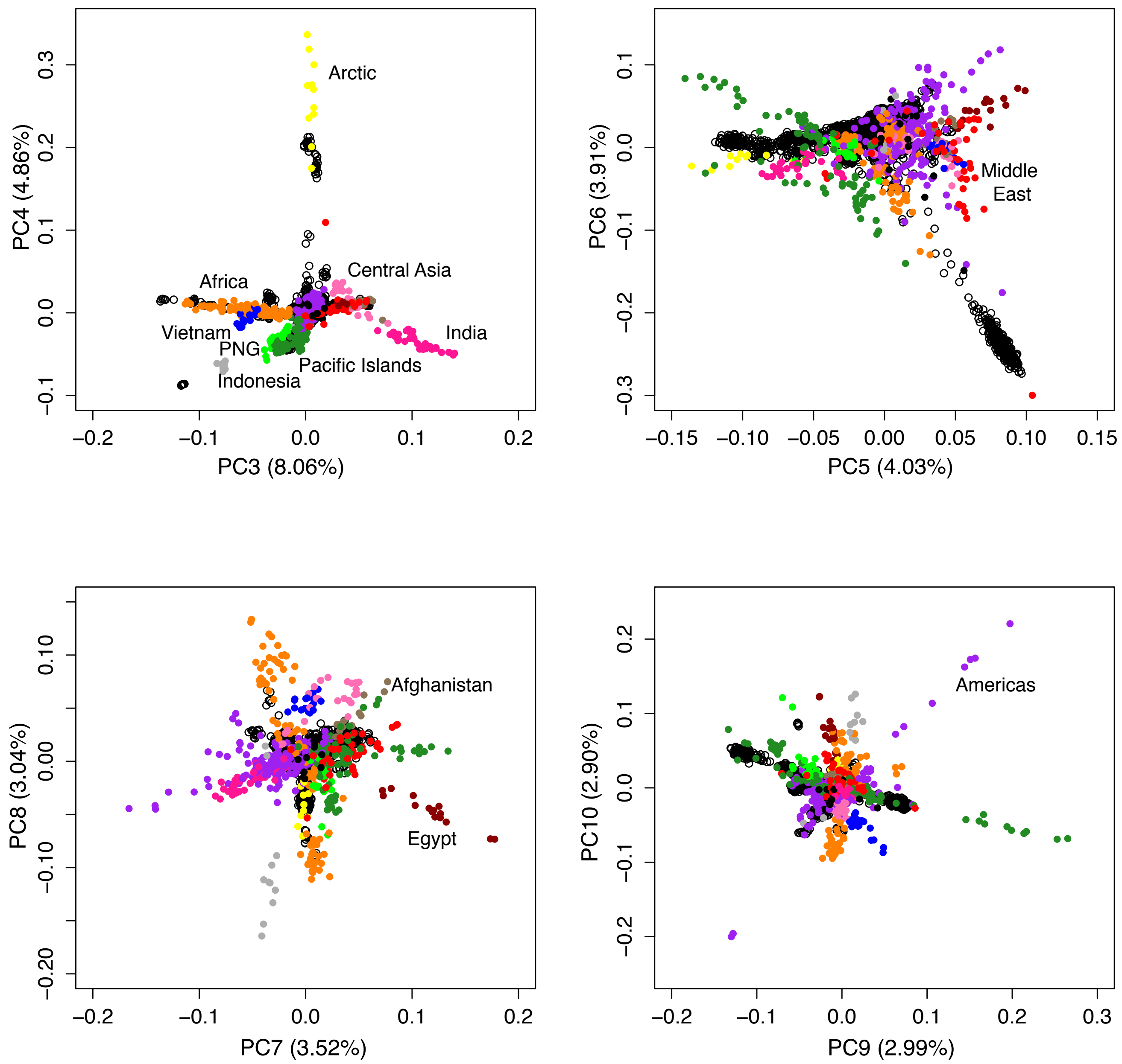

Fig. S4: Principal component analysis for village dogs (PC3–PC10). Village dogs are represented by colored dots while breed dogs, which have been projected onto the PCA space, are represented by open black circles.

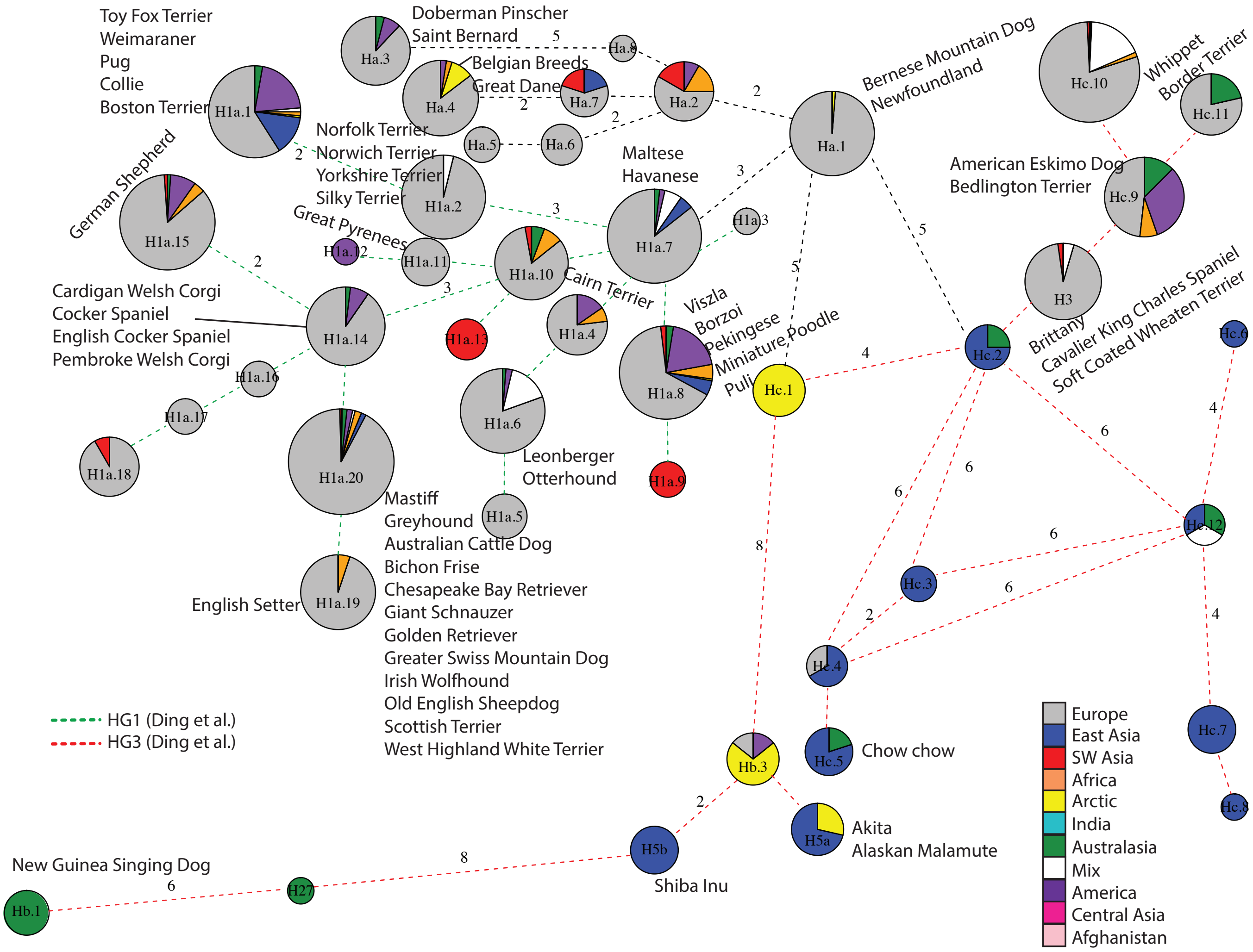

Fig. S5: Minimum spanning network for Y-haplogroup HG1-3. Circle size is proportional to haplotype frequency and line length is determined by number of mutations separating haplotypes (indicated by numerals when > 1 mutation). Previously reported haplotypes have maintained their published names. Some haplotypes are particularly common in individual breeds. Breed name is printed next to a haplotype if we had at least 5 individuals from that breed and more than 90% of those individuals carried the haplotype. HG1 and HG3 were separate haplogroups in Ding et al. [10]. Red dotted lines connect haplotypes previously grouped in HG3 and green dotted lines connect haplotypes previously grouped in HG1.

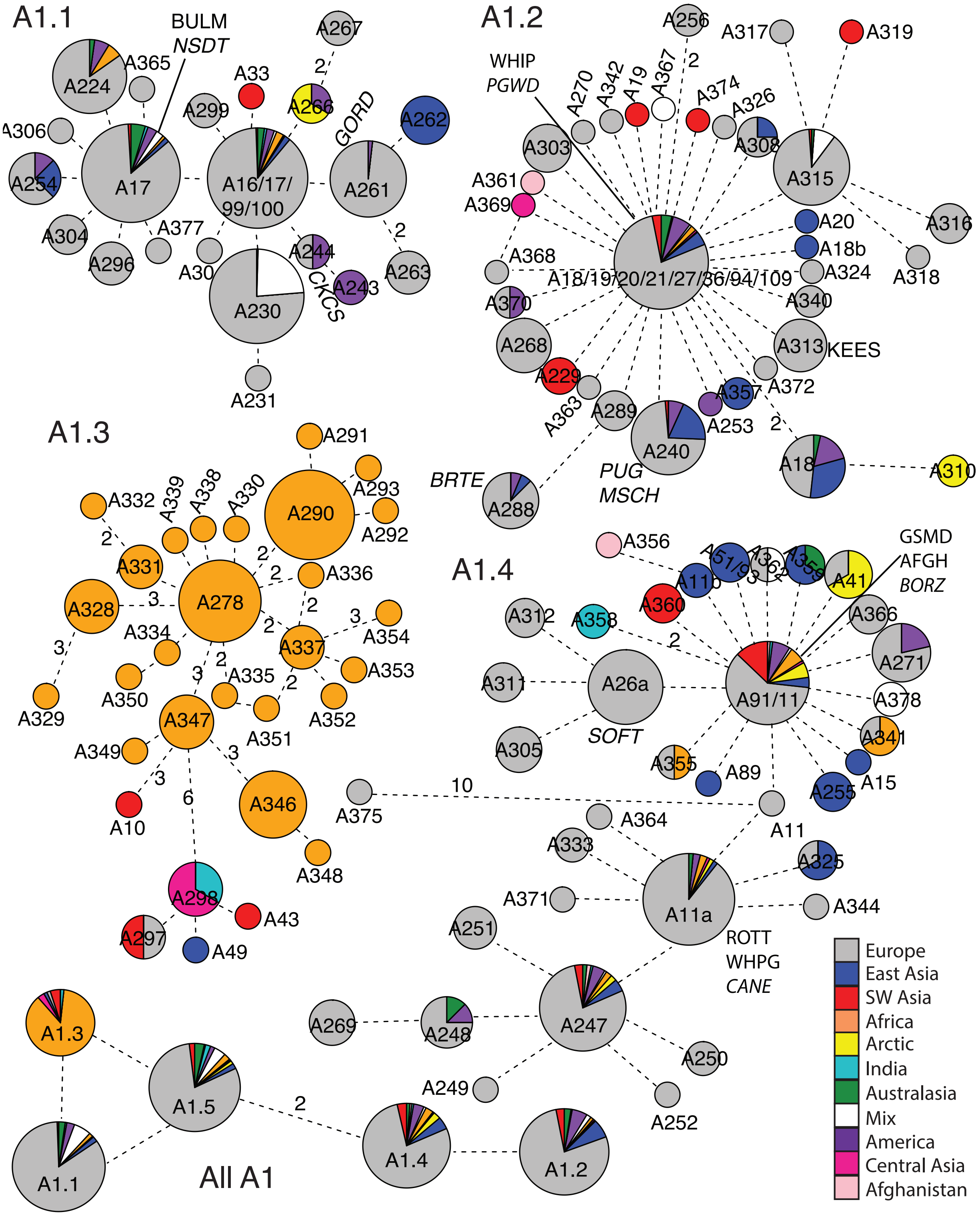

Fig. S6: Minimum spanning networks for haplogroup A1. A1 is the largest Mt haplogroup and we have subdivided it in the interest of readability. Circle size is proportional to haplotype frequency and line length is determined by number of mutations separating haplotypes (indicated by numerals when > 1 mutation). Previously reported haplotypes have maintained their published names. Novel haplotypes, or published haplotypes which did not follow the naming convention from Pang [8], have names larger than A200. Some haplotypes are particularly common in individual breeds. Breed name is printed next to a haplotype if we had at least 5 individuals from that breed and more than 80% (or 50% in italics) of those individuals carried the haplotype. The network in the bottom left corner indicates how the subhaplogroups fit together.

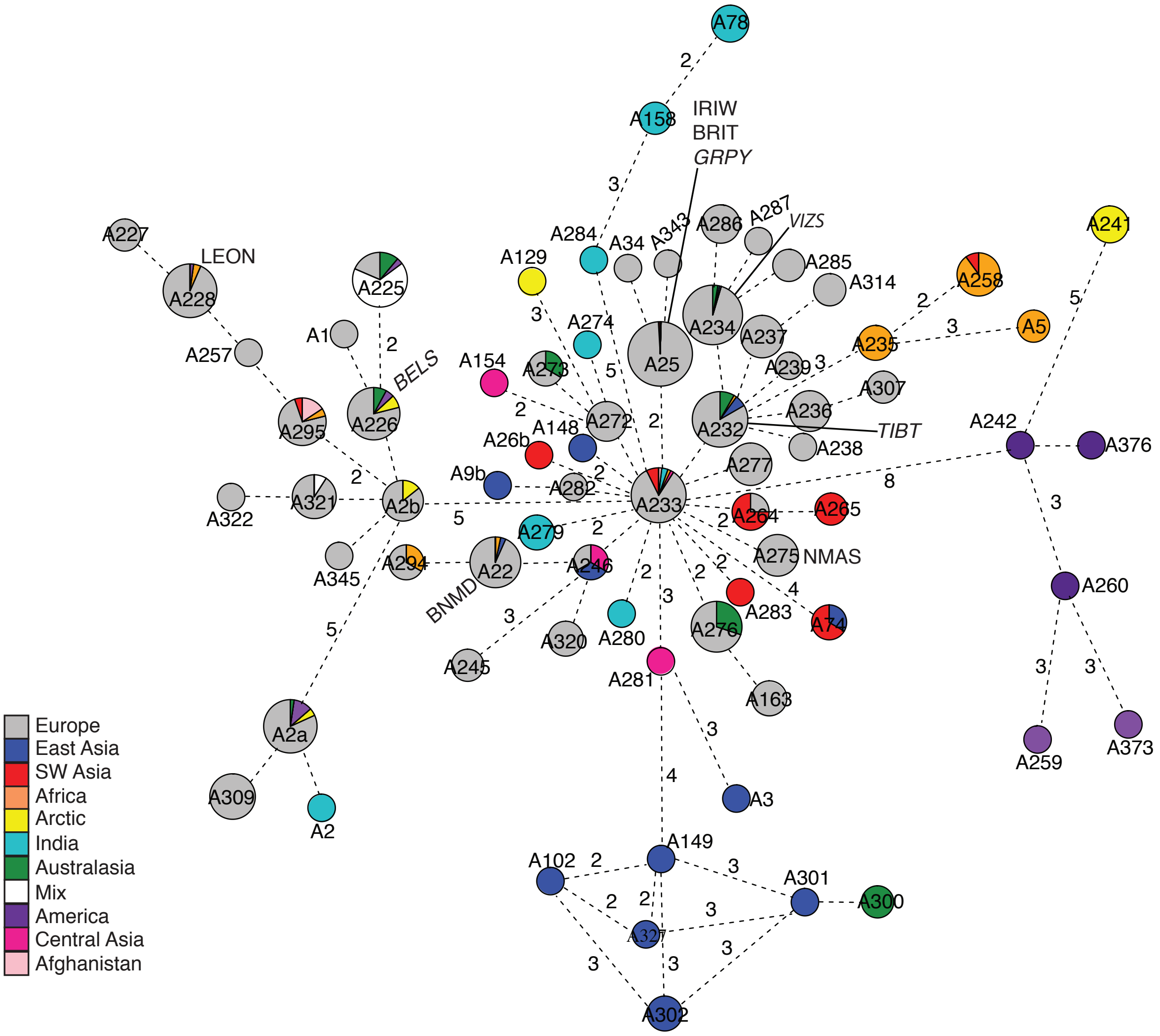

Fig. S7: A minimum spanning network for Mt subhaplogroup A1.5. Circle size is proportional to haplotype frequency and line length is determined by number of mutations separating haplotypes (indicated by numerals when > 1 mutation). Previously reported haplotypes have maintained their published names. Novel haplotypes, or published haplotypes which did not follow the naming convention from Pang [8], have names larger than A200. Some haplotypes are particularly common in individual breeds. Breed name is printed next to a haplotype if we had at least 5 individuals from that breed and more than 80% (or 50% in italics) of those individuals carried the haplotype. Ancient American haplotypes are from Thalmann et al. [11]. Modern American haplotypes clustering with the ancient haplotypes are from a village dog in Puerto Rico, a village dog in the Dominican Republican, and several Alaskan Village Huskies.

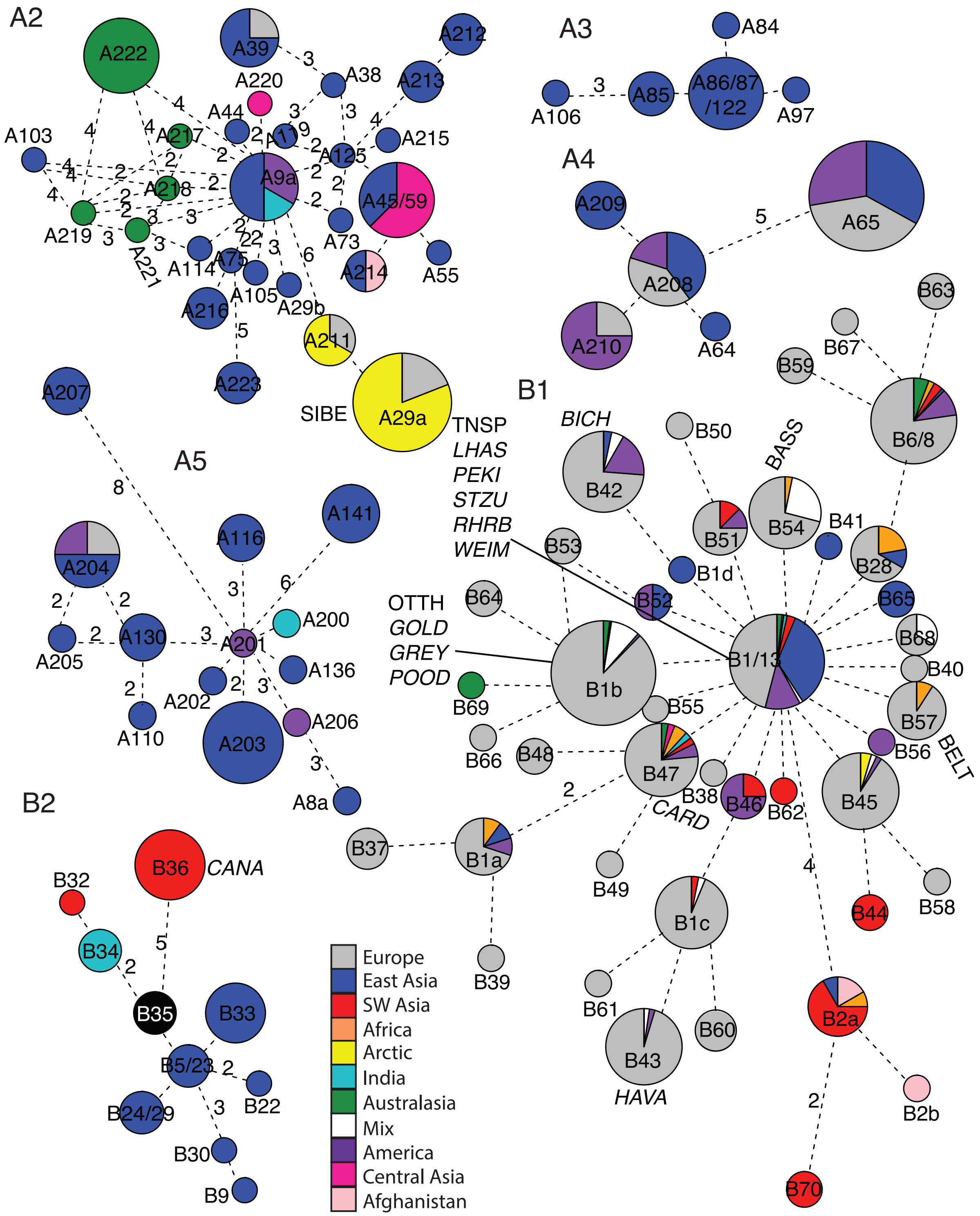

Fig. S8: Minimum spanning Mt haplotype networks for A and B haplogroups. Circle size is proportional to haplotype frequency and line length is determined by number of mutations separating haplotypes (indicated by numerals when > 1 mutation). Previously reported haplotypes have maintained their published names. Novel haplotypes, or published haplotypes which did not follow the naming convention from Pang et al [8], have names larger than A200 or B32. Some haplotypes are particularly common in individual breeds. Breed abbreviation is printed next to a haplotype if we had at least 5 individuals from that breed and more than 80% (or 50% in italics) of those individuals carried the haplotype. Ancient American haplotypes are from Thalmann et al. [11].

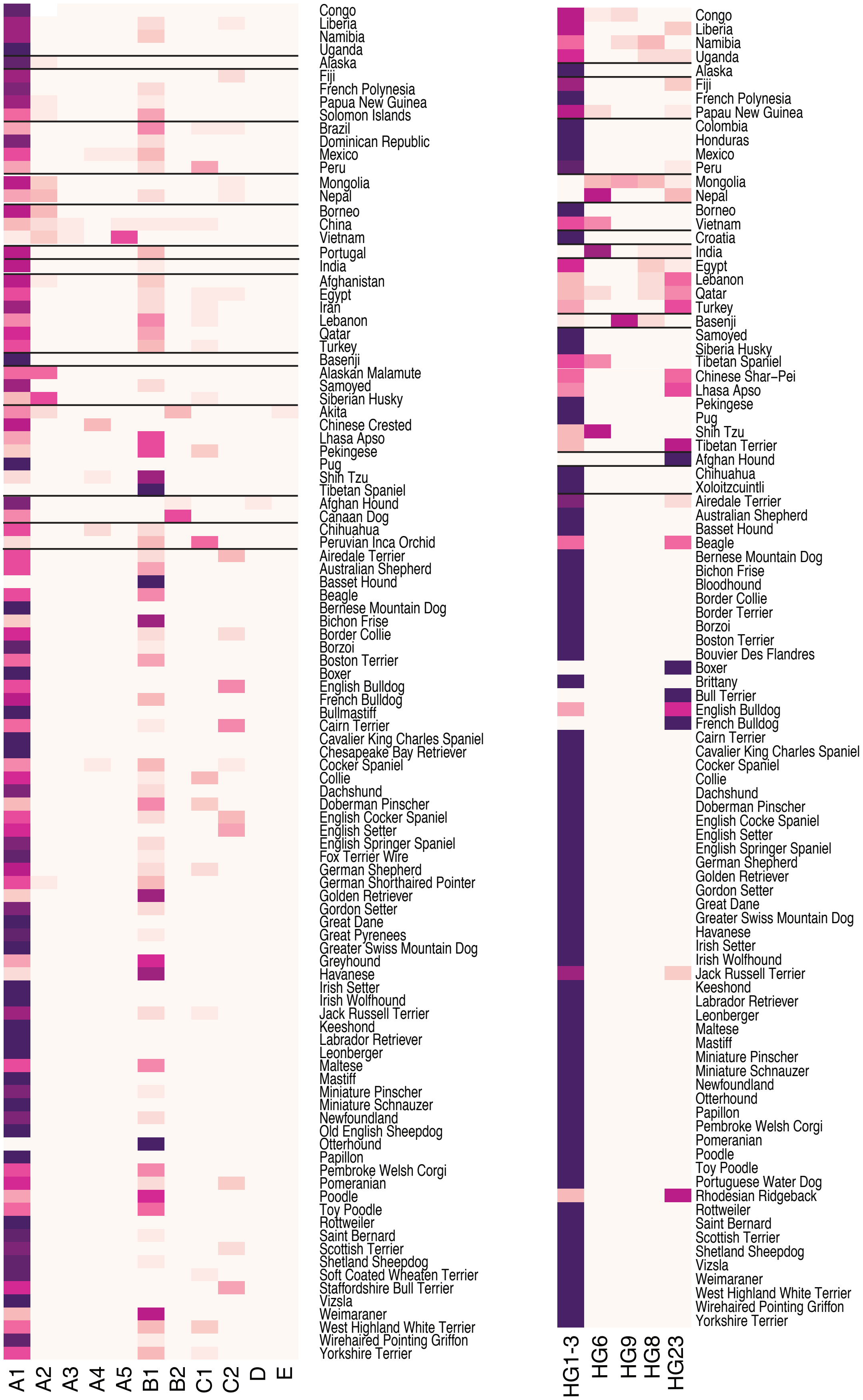

Fig. S9: Heat maps for haplogroups found within each population containing at least 10 dogs (at least 6 males). All Mt haplogroups (left) and Y haplogroups (right) are listed. Dark purple indicates that a group of dogs is fixed for a haplogroup while white indicates that the haplogroup is not found within that group of dogs.

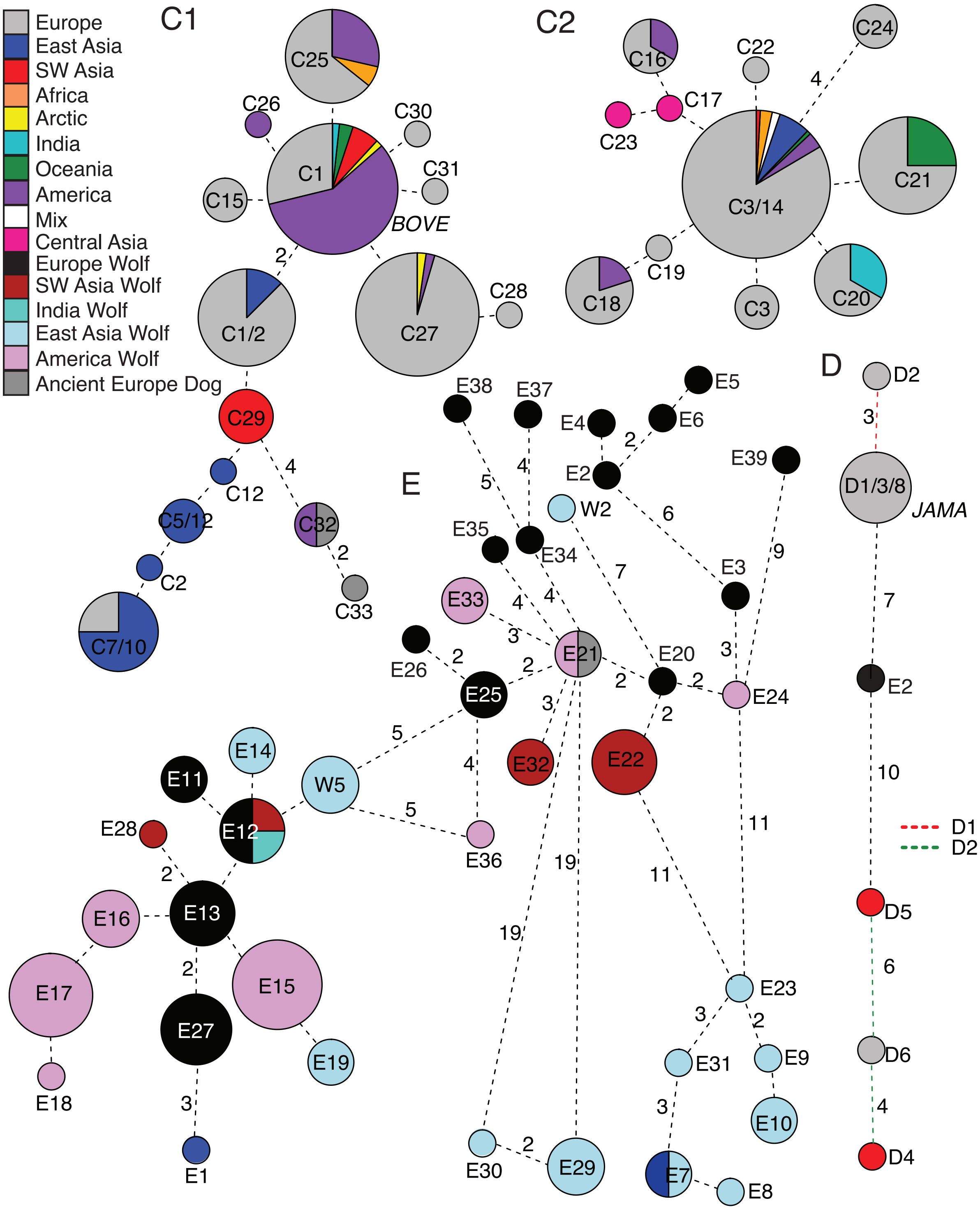

Fig. S10: Minimum spanning haplotype networks for C, D, and E haplogroups. Circle size is proportional to haplotype frequency and line length is determined by number of mutations separating haplotypes (indicated by numerals when > 1 mutation). Previously reported haplotypes have maintained their published names. Novel haplotypes, or published haplotypes which did not follow the naming convention from Pang et al. [8], have names larger than C15, or E1 (all D haplotypes have been previously reported). Some haplotypes are particularly common in individual breeds. Breed abbreviation is printed next to a haplotype if we had at least 5 individuals from that breed and more than 80% (or 50% in italics) of those individuals carried the haplotype. Ancient European wolf haplotypes are from Thalmann et al. [11].

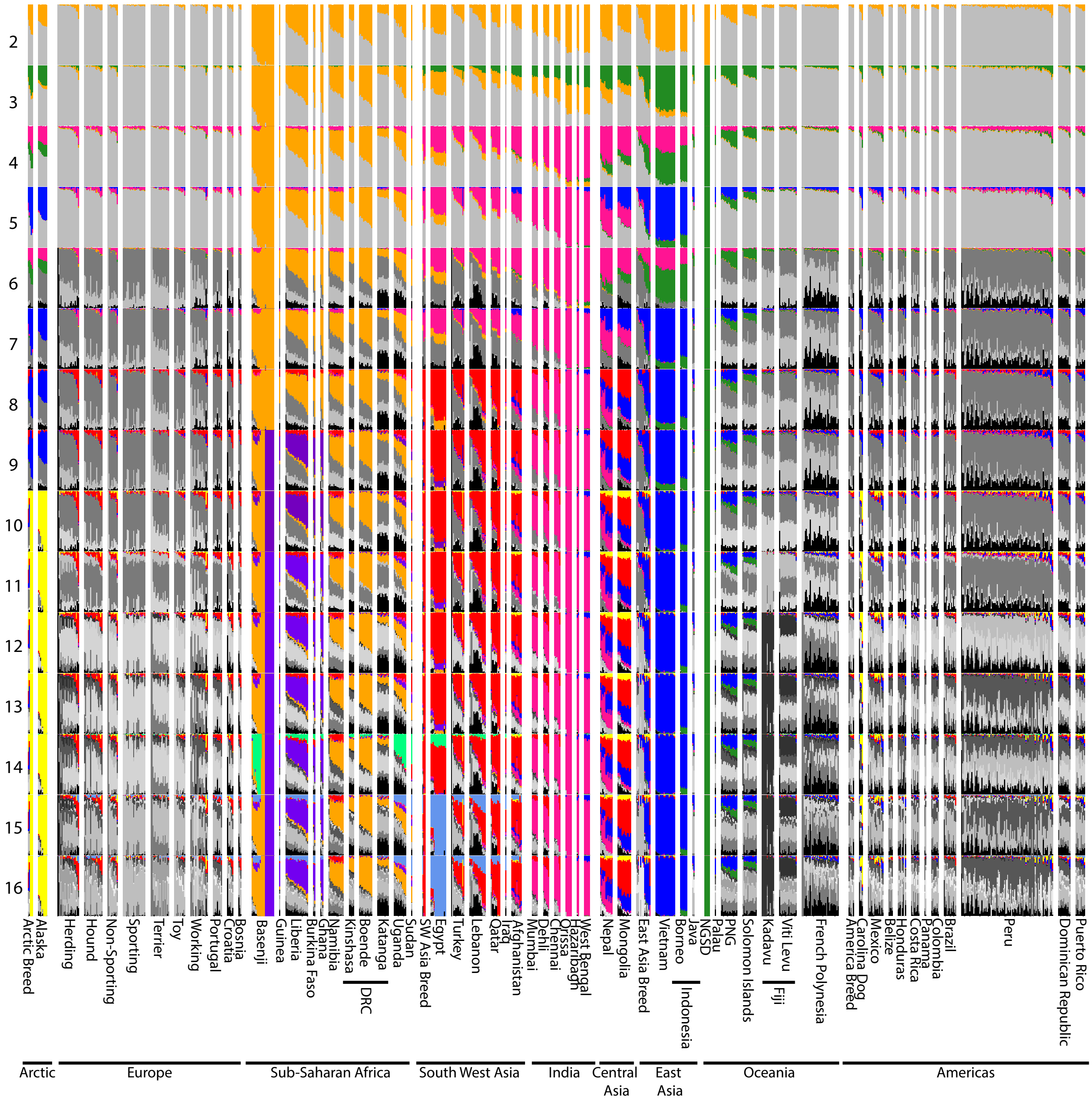

Fig. S11: ADMIXTURE analysis of village and purebred dogs at all K between 2 and 16. European components are in gray scale while regional components are in color. Each vertical line represents an individual, breeds are represented by a single individual with the exception of Basenjis, New Guinea Singing Dogs, and Carolina Dogs for which all unrelated individuals are included.

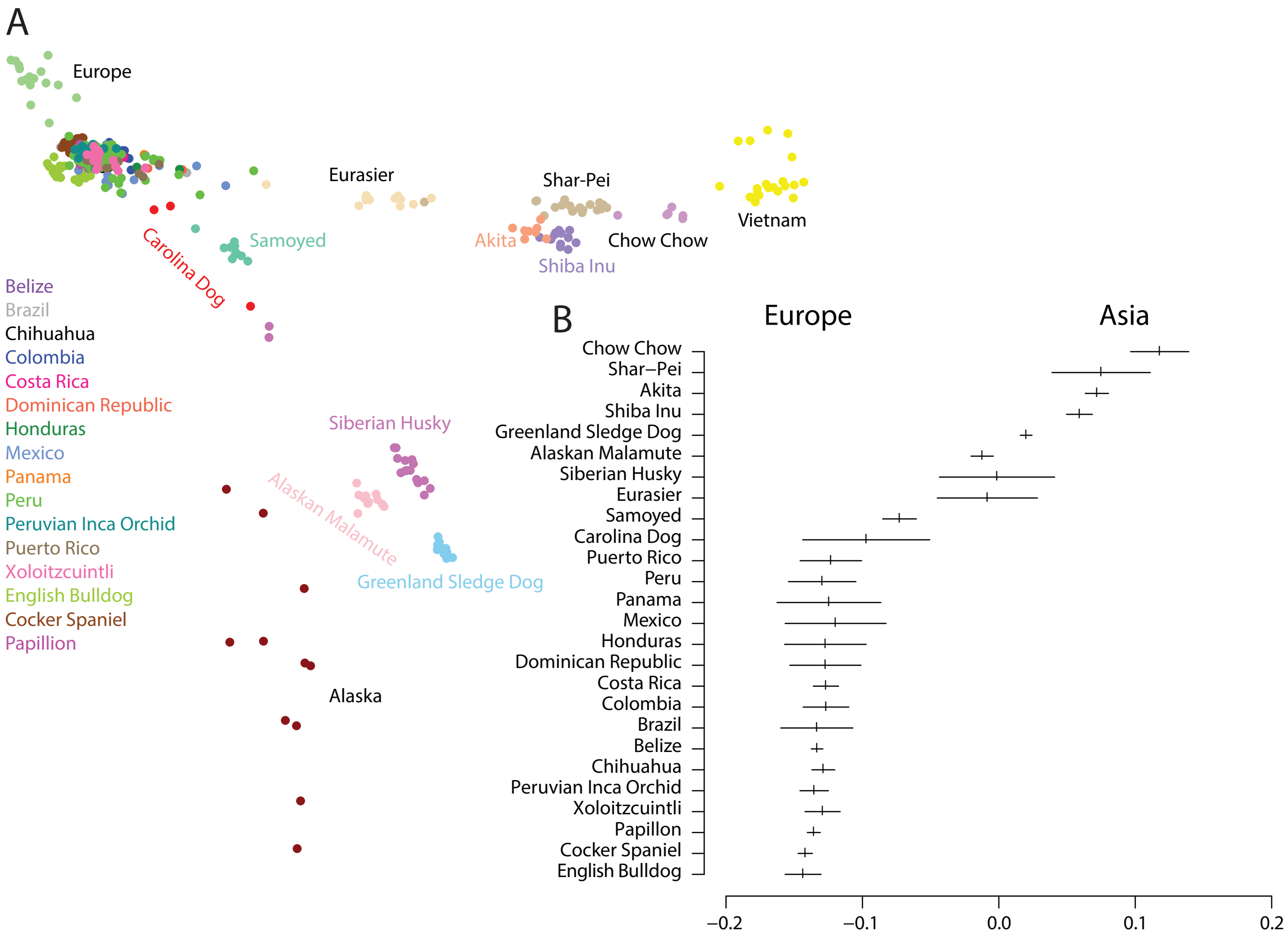

Fig. S12: A. Principal component analysis of European, Vietnamese, and Alaskan village dogs with American dogs (and representative European and Asian breed dogs) projected onto the PCA space. B. Mean value for PC1 of breed and village populations. Error bars indicate two standard deviations from the mean.

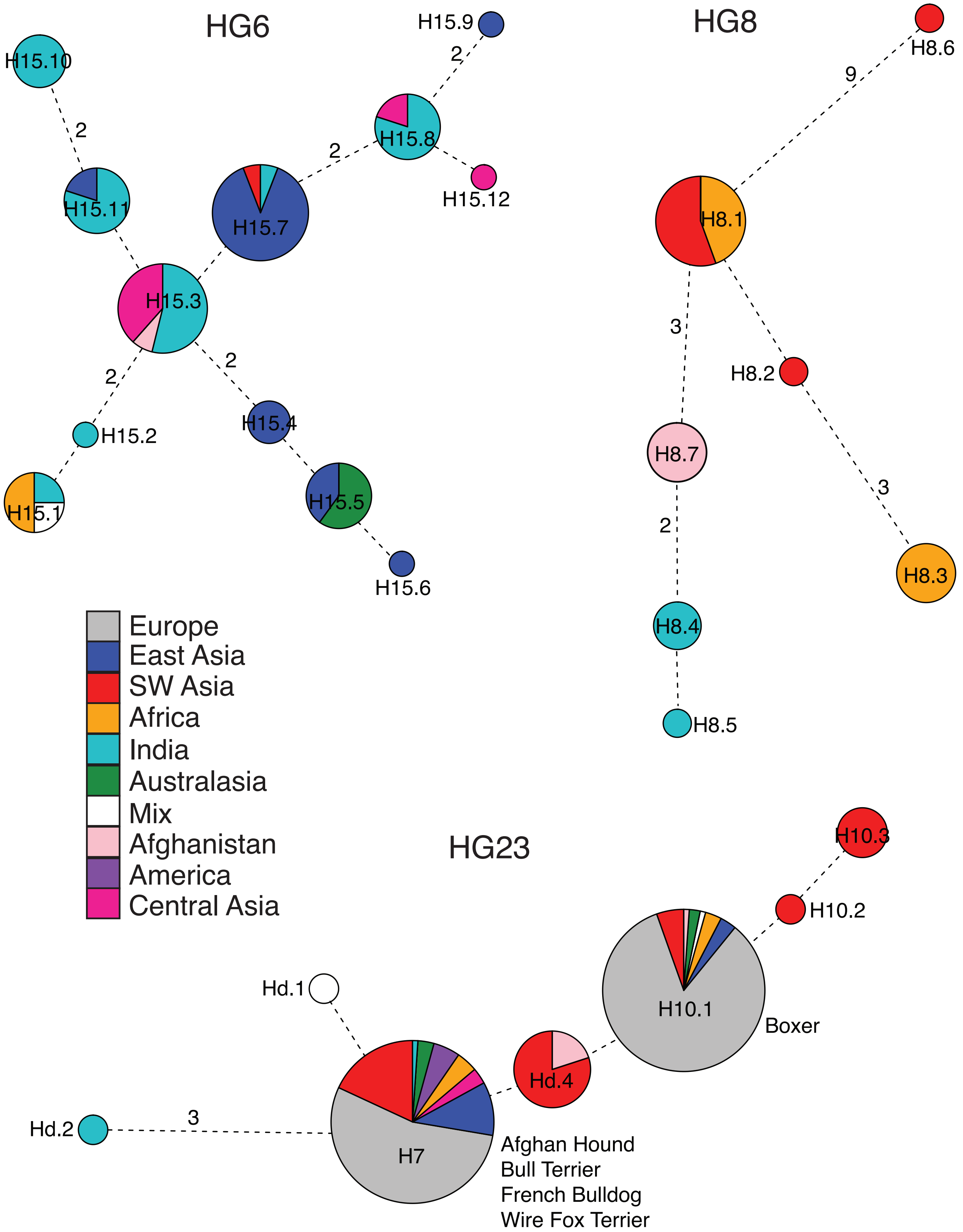

Fig. S13: Minimum spanning networks for the smaller Y haplotypes. Circle size is proportional to haplotype frequency and line length is determined by number of mutations separating haplotypes (indicated by numerals when > 1 mutation). Previously reported haplotypes have maintained their published names. Some haplotypes are particularly common in individual breeds. Breed name is printed next to a haplotype if we had at least 5 individuals from that breed and more than 90% of those individuals carried the haplotype.

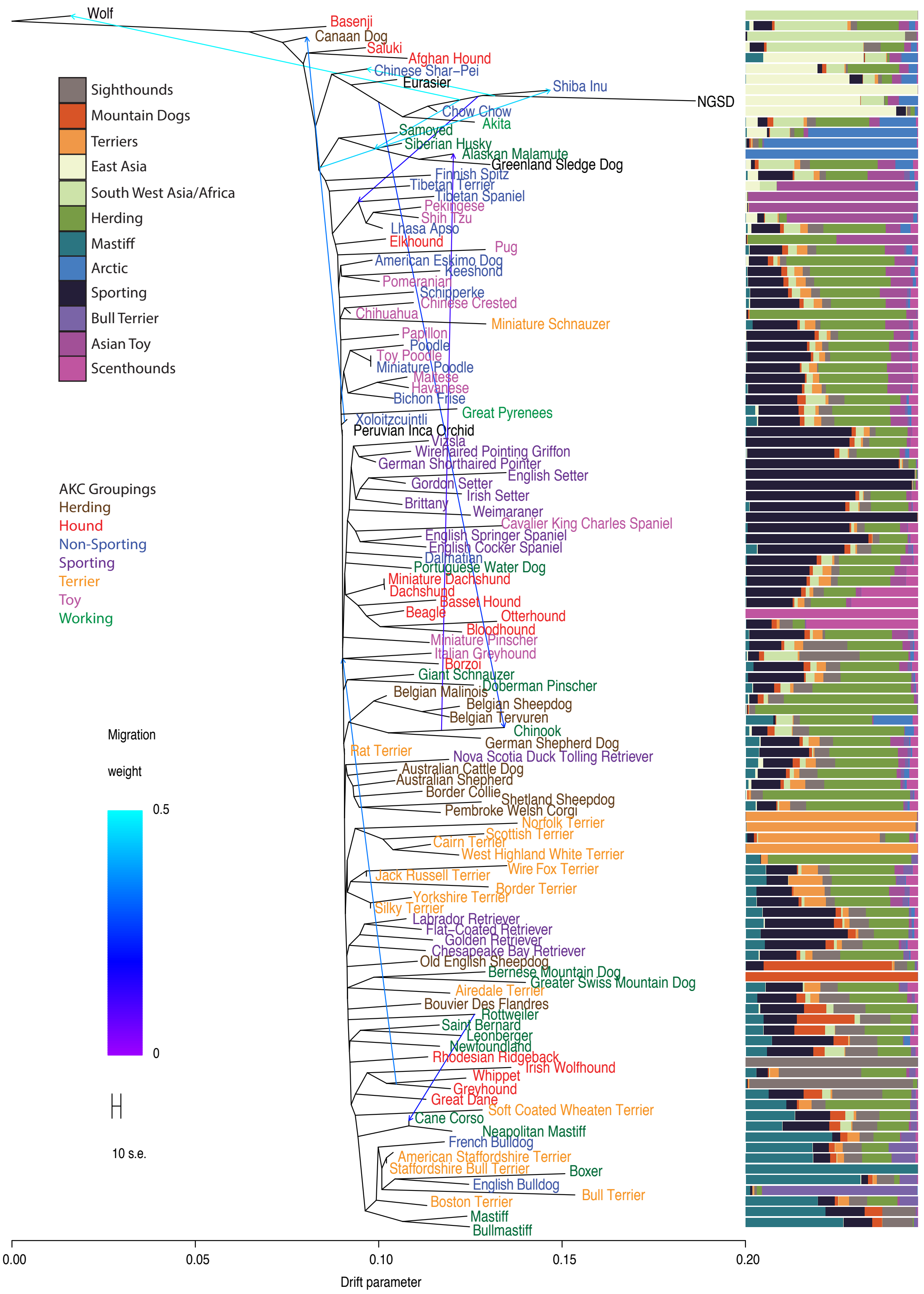

Fig. S14: Breeds form genetic clusters primarily according to geography in TreeMix (left) and ADMIXTURE (right). In both analyses, four basal clusters correspond to the Middle East, Africa, East Asia, and the Arctic with the fifth (European) cluster containing distinct subgroups corresponding to Sight Hounds, Scent Hounds, Sporting Dogs, Retrievers, Terriers, Mastiffs, and Herding Dogs, although many AKC groups (e.g. Toy, Working and Non-sporting) are spread throughout the tree. A few non-European breeds (Rhodesian Ridgeback, Chihuahua, Xoloitzcuintli and Peruvian Inca Orchid) group with European dogs, having been reconstituted over time from European stock. Four Asian breeds—Tibetan Spaniel, Lhasa Apso, Pekingese and Shih Tzu—also lie within the European cluster but retain between 18% and 44% Asian ancestry depending on breed (Fig. S15), consistent with their origins in East Asia followed by introgression from European dogs during modern times. Other admixture edges are consistent with known origins of modern breeds—Cane Corso as a mix of Rottweiler and Mastiff, Chinooks as a mix of German Shepherds and other dogs. The remaining admixture edges suggest gene flow between more ancient populations.

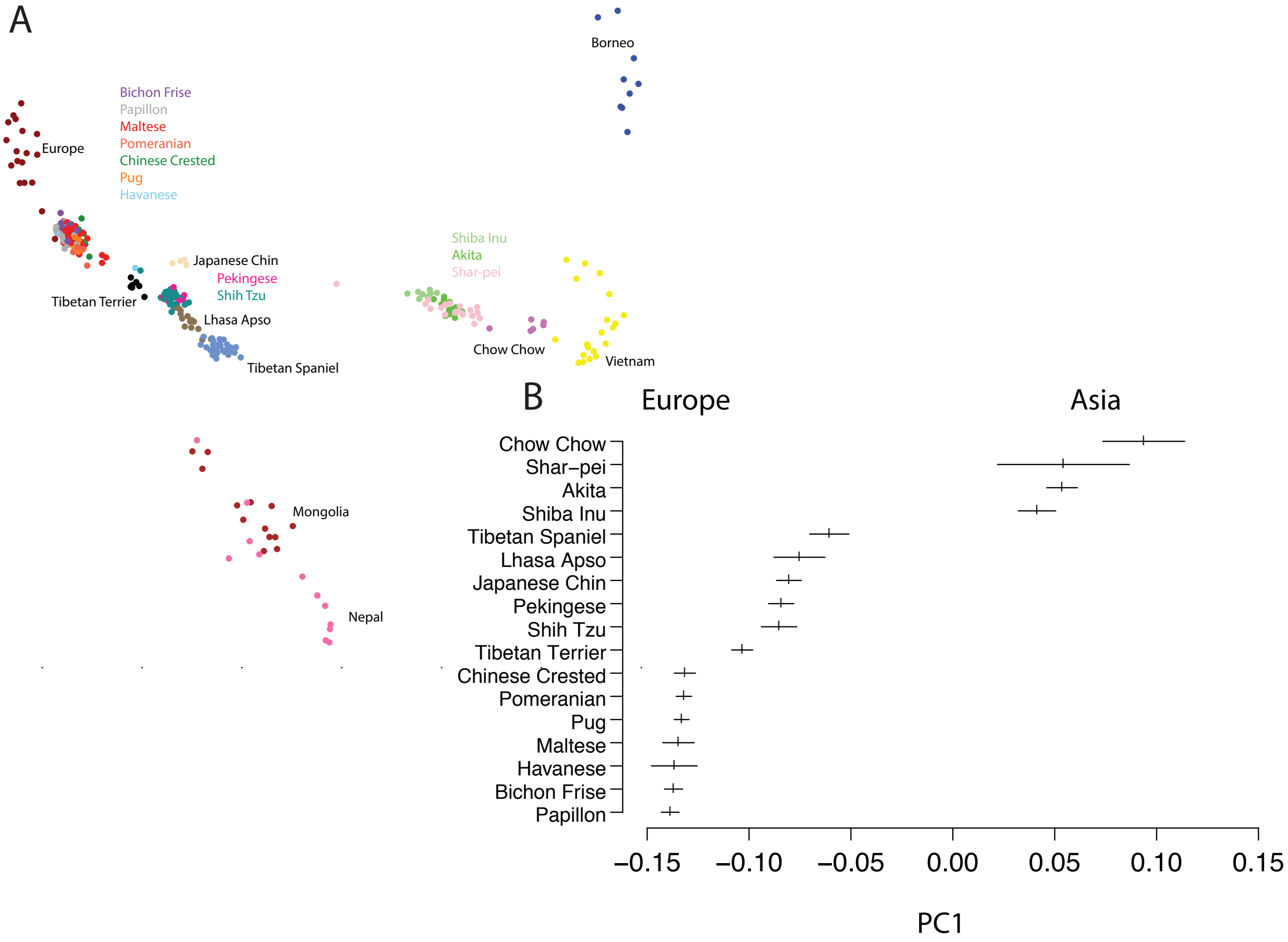

Fig. S15: A. Principal component analysis of European, Vietnamese, Bornese, Mongolian, and Nepalese village dogs with toy dogs (and representative Asian breed dogs) projected onto the PCA space. B. Mean value for PC1 of breed and village populations. Error bars indicate two standard deviations from the mean.

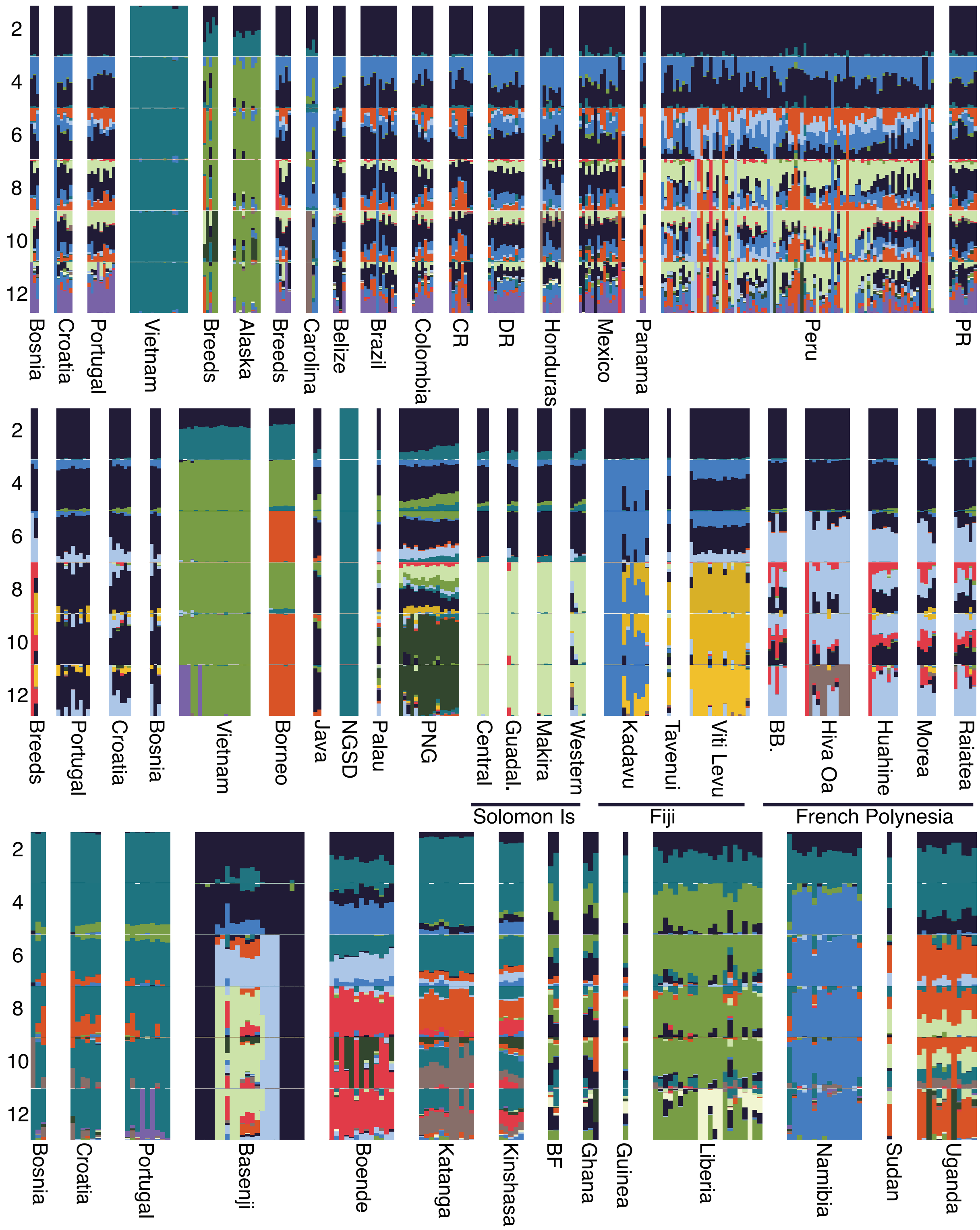

Fig. S16: Regional ADMIXTURE analyses. Top: New World Dogs, including Carolina dogs, American and Arctic breeds (Chihuahua, Peruvian Inca Orchid, Xoloitzcuintli, Greenland Sledge Dog, Alaskan Malamute, Siberian Husky, Eurasier, and Samoyed) with one individual per breed, and European and East Asian village dogs as outgroups. PR is Puerto Rico, DR is Dominican Republic, and CR is Costa Rica. Middle: Village dogs from Oceania and Island Southeast Asia, with select East Asian and European dogs. BB stands for Bora Bora. The breeds referenced here are Staffordshire Bull Terrier and Mastiff. Bottom: Village dogs from Africa with European village dogs as an out group. BF stands for Burkina Faso.

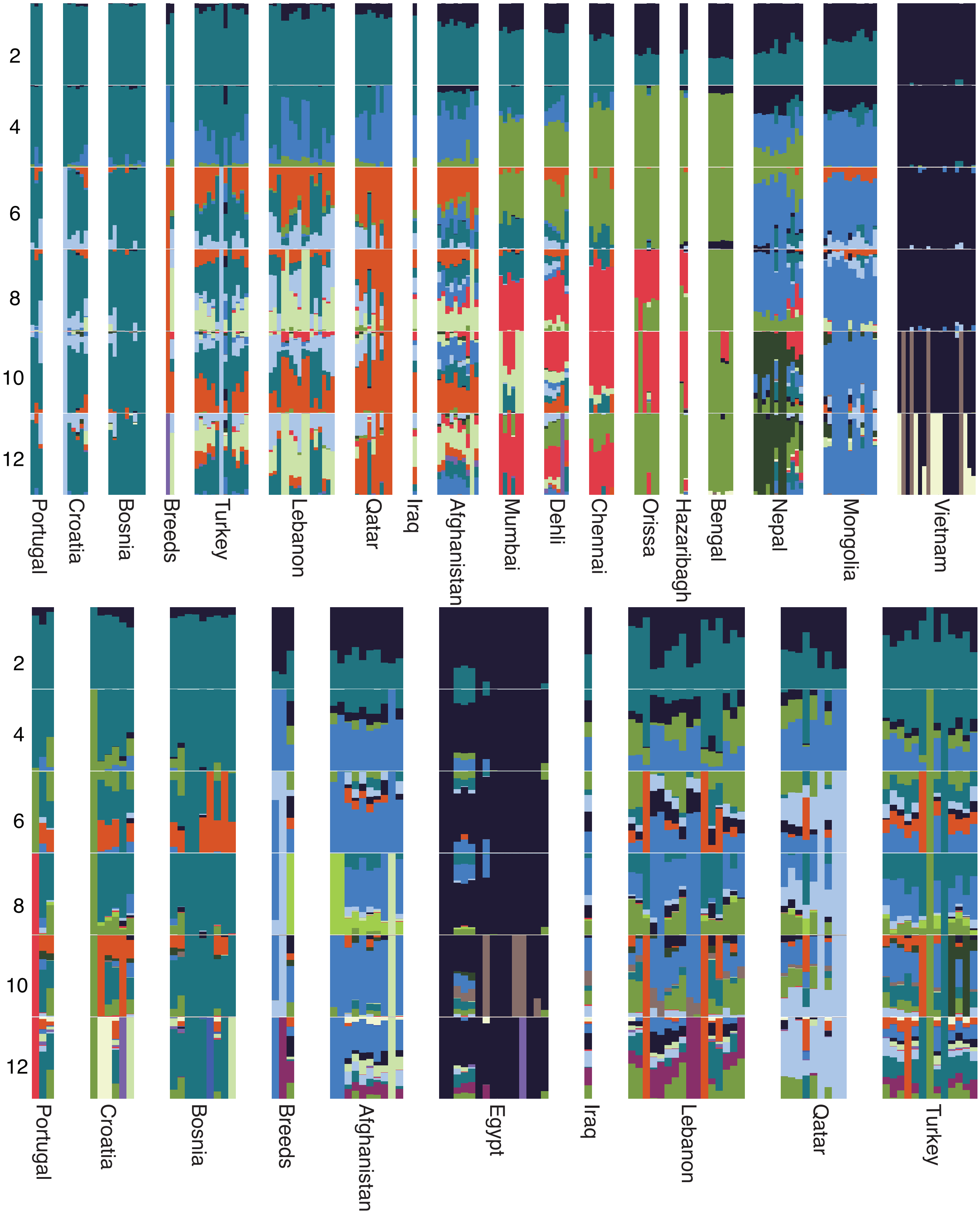

Fig. S17: Regional ADMIXTURE analyses, with two subsets of Eurasian village dogs–spanning the entire continent (top) and focused on Europe and South West Asia(bottom). In both cases an Afghan Hound and a Saluki are included as breeds.).

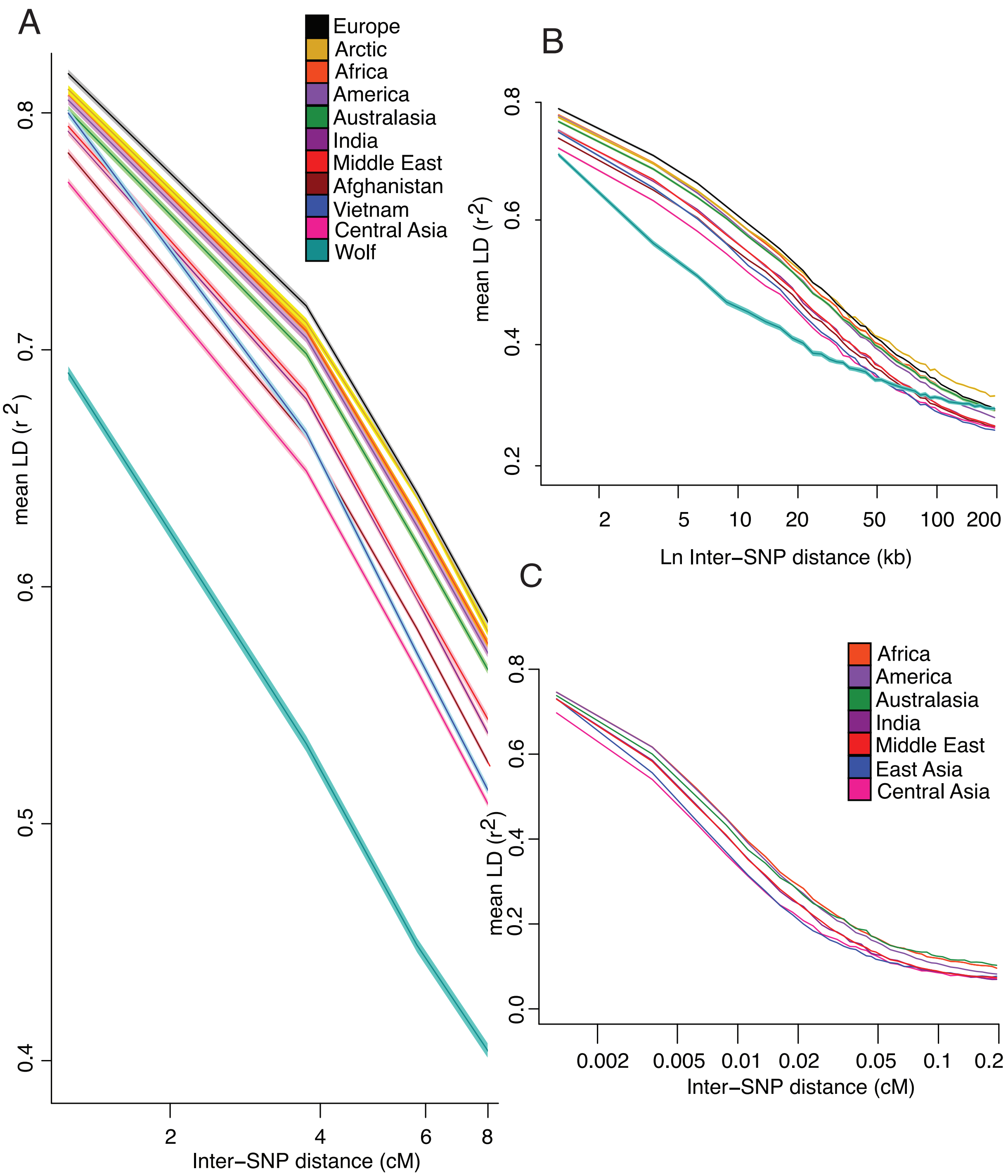

Fig. S18: LD decay for village dog populations worldwide. (A) LD at small inter-SNP distances. Dark lines represent means and lighter shading represent 1 s.e. interval calculated from 100 replicates. For each replicate, $N = 6$ individuals were selected randomly. This is a subsection of Fig 5 enlarged to make the 1 s.e. intervals visible. (B) LD curves for the same populations calculated in the same way as in A but with inter-SNP distance measured in kb. (C) LD curves calculated in the same way as A but for $N = 20$, a smaller number of populations were considered based which populations contained more than 20 individuals. Although the mean LD is lower, the relative ordering of populations stays the same.

Table S3: Shared vs. regional haplotypes in 651 village dogs

		Mt		Y	
region	population	both	local	both	local
Africa	Burkina Faso	0	2	0	1
	Congo	18	15	16	6
	Ghana	4	1	1	1
	Guinea	1	0	1	0
	Liberia	10	11	15	1
	Namibia	13	2	6	5
	Nigeria	0	1	-	-
	Sudan	0	1	-	-
	Uganda	7	5	9	2
America	Belize	4	0	3	0
	Brazil	11	2	4	0
	Colombia	9	0	6	0
	Costa Rica	7	1	2	0
	Dominican Republic	10	2	2	0
	Honduras	8	0	6	0
	Mexico	15	0	10	0
	Panama	2	0	-	-
	Peru	83	8	61	0
	Puerto Rico	7	2	5	0
Arctic	Alaska	7	4	4	2
Oceania	Fiji	29	0	14	0
	French Polynesia	35	1	15	0
	Palau	1	0	-	-
	Papua New Guinea	16	4	9	6
	Solomon Island	14	1	3	0
Central Asia	Mongolia	5	9	1	10
	Nepal	6	6	2	4
East Asia	Borneo	2	7	0	6
	China	29	51	-	-
	Java	2	0	2	0
	Thailand	4	3	-	-
	Vietnam	0	20	6	10
Europe	Bosnia	3	0	3	0
	Croatia	6	0	6	0
	Portugal	6	3	5	0
India	India	15	20	2	25
South West Asia	Afghanistan	8	5	1	2
	Egypt	7	8	2	6
	Iran	12	1	-	-
	Iraq	1	0	1	0
	Lebanon	9	7	9	6
	Qatar	5	4	2	5
	Turkey	11	2	9	1

Table S4: Mt haplotype diversity in novel and combined Mt data sets

	Semi-Custom CanineHD Array				Combined			
Population	N	H	HT(HG)	HT(SD)	N	H	HT(HG)	HT(SD)
Africa Breed	30	0.93	15(1)	11.06(3.36)	33	0.93	15(1)	11.92(1.10)
Africa Village	91	0.96	41(4)	14.05(3.81)	91	0.96	41(4)	14.91(1.66)
America Breed	46	0.92	20(6)	12.06(1.53)	48	0.93	21(6)	12.38(1.67)
America Village	171	0.94	49(7)	13.45(1.92)	171	0.94	49(7)	13.82(1.66)
Arctic Breed	39	0.84	11(4)	8.77(1.07)	42	0.83	12(4)	8.92(1.24)
Arctic Village	11	0.82	5(2)	-	11	0.82	5(2)	-
Oceania Breed	5	0	1(1)	-	6	0.33	2(1)	-
Oceania Village	101	0.93	29(5)	12.44(1.64)	105	0.93	29(5)	12.48(1.79)
Cental Asia Village	25	0.95	16(5)	13.83(0.95)	26	0.95	17(5)	13.92(1.04)
East Asia Breed	151	0.85	32(10)	10.68(1.76)	166	0.86	38(10)	11.67(2.05)
East Asia Village	31	0.96	19(5)	14.36(1.42)	118	0.99	74(10)	17.95(1.11)
Europe Breed	3414	0.95	151(7)	13.99(1.82)	3534	0.95	164(8)	14.20(1.92)
Europe Village	18	0.96	15(3)	-	18	0.96	15(3)	-
India Village	31	0.98	23(7)	16.41(1.12)	35	0.97	24(7)	15.44(1.44)
South West Asia Breed	31	0.85	10(4)	8.60(0.88)	37	0.90	16(4)	11.00(1.36)
South West Asia Village	67	0.97	34(6)	15.26(1.47)	80	0.96	36(6)	14.74(1.67)

Sample size (N) and haplotype heterozygosity (H) by region. HT(HG) is the number of unique haplotypes (haplogroups) and HT(SD) is the average number (standard deviation) of haplotypes from a sub-sample of 20 dogs.

Table S5: Y haplotype diversity in breed dogs (top) and village dogs (bottom)

Population	N	H	HT(HG)	HT(SD)
Africa	10	0.51	3(3)	-
America	26	0.67	7(1)	5.91(0.78)
Arctic	15	0.79	5(1)	-
East Asia	83	0.88	12(3)	9.17(1.35)
Europe	1708	0.91	34(3)	10.65(1.64)
South West Asia	14	0.59	5(3)	-
Africa	64	0.93	16(5)	11.58(1.49)
America	99	0.86	15(2)	8.09(1.33)
Arctic	6	0.87	4(1)	-
Oceania	47	0.94	19(3)	12.36(1.34)
Central Asia	17	0.84	6(4)	-
East Asia	24	0.92	13(2)	11.68(0.85)
Europe	14	0.93	9(1)	
India	27	0.89	12(4)	10.19(1.01)
South West Asia	46	0.94	20(4)	12.41(1.38)

Sample size (N) and haplotype heterozygosity (H) by region. HT(HG) is the number of unique haplotypes (haplogroups) and HT(SD) is the average number (standard deviation) of haplotypes from a sub-sample of 20 dogs.

Table S6: Ancestry percentages for African and island populations estimated by F3 tests

Source population 1	Source population 2	Target population	F3	Z score	Ancestry proportion from source 1
Europe	Vietnam	French Polynesia	0.01	14.82	0.958-0.991
Europe	Borneo	French Polynesia	0.01	13.02	0.958-0.981
Europe	Vietnam	Fiji	0.01	15.34	0.952-0.966
Europe	Borneo	Fiji	0.01	12.57	0.951-0.959
Europe	Vietnam	Solomon Islands	-0.00	-2.28	0.785-0.855
Europe	Borneo	Solomon Islands	-0.01	-6.48	0.783-0.863
Europe	Vietnam	PNG	-0.02	-32.26	0.680-0.715
Europe	Borneo	PNG	-0.02	-30.68	0.678-0.784
Europe	Basenji	Katanga	-0.01	-5.02	0.884-0.856
Europe	Basenji	Kinshasa	-0.03	-27.66	0.749-0.730
Europe	Basenji	Uganda	-0.03	-40.38	0.580-0.638
Europe	Basenji	Namibia	-0.03	-32.83	0.525-0.609
Europe	Basenji	Liberia	-0.03	-39.95	0.522-0.599
Europe	Basenji	Boende	-0.031	-33.72	0.394-0.355

Table S7: F_{st} between American populations

	Alaska	ArcB	Bos	Cro	EAB	EurB	Por	Vie	Mon	AmB	Car	Mex	Hon	CR	Col	Bra	Peru	DR
Arctic breed	0.02																	
Bosnia	0.08	0.06																
Croatia	0.07	0.06	0.00															
East Asia breed	0.04	0.03	0.02	0.02														
Europe breed	0.05	0.03	0.00	0.00	0.03													
Portugal	0.07	0.07	0.01	0.00	0.03	0.00												
Vietnam	0.11	0.09	0.15	0.15	0.06	0.13	0.15											
Mongolia	0.05	0.04	0.05	0.04	0.02	0.04	0.05	0.08										
America breed	0.09	0.06	0.00	0.01	0.02	0.00	0.02	0.14	0.06									
Carolina Dog	0.06	0.05	0.02	0.02	0.02	0.00	0.03	0.13	0.04	0.04								
Mexico	0.06	0.06	0.01	0.00	0.02	0.00	0.00	0.13	0.04	0.01	0.02							
Honduras	0.07	0.07	0.02	0.01	0.03	0.00	0.01	0.15	0.05	0.02	0.02	0.01						
Costa Rica	0.07	0.06	0.01	0.00	0.02	0.00	0.01	0.14	0.04	0.01	0.02	0.00	0.01					
Colombia	0.07	0.07	0.01	0.01	0.02	0.00	0.01	0.14	0.04	0.02	0.02	0.00	0.01	0.01				
Brazil	0.07	0.07	0.01	0.00	0.03	0.00	0.00	0.14	0.04	0.02	0.02	0.00	0.01	0.00	0.01			
Peru	0.06	0.04	0.00	0.00	0.03	0.00	0.01	0.13	0.04	0.00	0.01	0.00	0.01	0.01	0.00	0.00		
Dominican Rep	0.06	0.07	0.01	0.01	0.03	0.00	0.01	0.14	0.04	0.02	0.02	0.00	0.01	0.01	0.01	0.00	0.00	
Puerto Rico	0.06	0.06	0.01	0.01	0.02	0.00	0.00	0.13	0.04	0.02	0.02	0.00	0.01	0.00	0.00	0.00	0.00	0.00

Table S8: F_{st} between Pacific Island populations

	Kadavu	Europe	Huahine	Moorea	Raiatea	PNG	NGSD	Borneo	Vietnam	Hiva Oa	Solomon Islands	Bora Bora
Europe	0.04											
Huahine	0.04	0.01										
Moorea	0.04	0.01	0.00									
Raiatea	0.04	0.01	0.00	0.00								
PNG	0.04	0.03	0.02	0.02	0.02							
NGSD	0.39	0.14	0.41	0.46	0.44	0.29						
Boreno	0.17	0.09	0.16	0.17	0.17	0.10	0.32					
Vietnam	0.16	0.14	0.15	0.15	0.15	0.08	0.26	0.05				
Hiva Oa	0.04	0.02	0.01	0.01	0.01	0.03	0.39	0.17	0.15			
Solomon Islands	0.04	0.03	0.03	0.02	0.03	0.01	0.32	0.12	0.11	0.032		
Bora Bora	0.04	0.01	0.00	0.00	0.00	0.02	0.46	0.17	0.15	0.01	0.03	
Viti Levu	0.02	0.02	0.02	0.01	0.02	0.02	0.37	0.16	0.14	0.02	0.03	0.02

Table S9: F_{st} between African populations

	Basenji	Boende	Bosnia	Croatia	Egypt	Europe Breeds	Ghana	Katanga	Kinshasa	Liberia	Namibia	Portugal
Boende	0.06											
Bosnia	0.21	0.09										
Croatia	0.21	0.08	0.00									
Egypt	0.13	0.06	0.06	0.05								
Europe Breeds	0.17	0.07	0.00	0.00	0.05							
Ghana	0.06	0.04	0.08	0.08	0.06	0.02						
Katanga	0.17	0.06	0.02	0.01	0.05	0.01	0.07					
Kinshasa	0.14	0.04	0.02	0.01	0.05	0.00	0.05	0.01				
Liberia	0.08	0.03	0.05	0.05	0.04	0.05	0.02	0.04	0.03			
Namibia	0.09	0.03	0.06	0.05	0.05	0.04	0.04	0.04	0.02	0.03		
Portugal	0.20	0.08	0.01	0.00	0.05	0.00	0.08	0.02	0.02	0.05	0.05	
Uganda	0.10	0.04	0.04	0.04	0.04	0.03	0.04	0.03	0.02	0.02	0.03	0.04

Table S10: F_{st} between Asian populations

	Vie[a]	Bas[b]	Ben[c]	Bos[d]	Che[e]	Cro[f]	Dehli	EAB[g]	Egypt	EB[h]	Leb[i]	Mon[j]	Mum[k]	Nepal	Ori[l]	Por[m]	Qatar	SWAB[n]	Tur[o]
Bas	0.25																		
Ben	0.13	0.17																	
Bos	0.15	0.21	0.13																
Che	0.14	0.16	0.03	0.08															
Cro	0.15	0.21	0.12	0.00	0.07														
Dehli	0.12	0.16	0.03	0.05	0.01	0.04													
EAB	0.06	0.13	0.07	0.02	0.05	0.02	0.03												
Egypt	0.14	0.13	0.07	0.06	0.05	0.05	0.03	0.05											
EB	0.13	0.17	0.08	0.00	0.05	0.00	0.03	0.03	0.05										
Leb	0.13	0.16	0.08	0.02	0.05	0.01	0.03	0.03	0.03	0.01									
Mon	0.08	0.18	0.07	0.05	0.05	0.04	0.03	0.02	0.05	0.04	0.03								
Mum	0.13	0.16	0.04	0.05	0.01	0.05	0.00	0.03	0.04	0.03	0.03	0.04							
Nepal	0.07	0.16	0.04	0.06	0.04	0.05	0.02	0.02	0.04	0.05	0.04	0.02	0.03						
Ori	0.15	0.17	0.01	0.13	0.02	0.12	0.03	0.08	0.07	0.07	0.09	0.08	0.03	0.05					
Por	0.15	0.20	0.12	0.01	0.08	0.00	0.05	0.03	0.05	0.00	0.02	0.05	0.05	0.06	0.12				
Qatar	0.13	0.16	0.08	0.03	0.05	0.03	0.03	0.03	0.03	0.02	0.01	0.03	0.03	0.04	0.08	0.03			
SWAB	0.14	0.18	0.09	0.02	0.05	0.04	0.03	0.03	0.04	0.01	0.02	0.04	0.04	0.04	0.09	0.04	0.01		
Tur	0.13	0.17	0.09	0.01	0.06	0.00	0.03	0.02	0.04	0.01	0.00	0.03	0.03	0.04	0.10	0.01	0.01	0.02	
Afg	0.11	0.16	0.06	0.03	0.04	0.03	0.02	0.02	0.03	0.02	0.01	0.02	0.02	0.02	0.07	0.03	0.01	0.02	0.01

[a] Vietnam

[b] Basenji

[c] Bengal

[d] Bosnia

[e] Chennai

[f] Croatia

[g] East Asia Breed

[h] Europe Breed

[i] Lebanon

[j] Mongolia

[k] Mumbai

[l] Orissa

[m] Portugal

[n] South West Asian Breed

[o] Turkey

Table S11: Sampling locations for village dogs

region	country	location	N_{mal}	N_{fem}	notes
Africa	Guinea	Kankan	1	0	US import
Africa	South Sudan	Sudan	0	1	US import
Africa	Nigeria	unknown	1	0	sampled in Monrovia
Africa	Namibia	northern	10	4	see Boyko et al 2009
Africa	Namibia	central	1	0	see Boyko et al 2009
Africa	Uganda	mainland	11	1	see Boyko et al 2009
Africa	Liberia	Lofa	13	6	various districts
Africa	Liberia	Monrovia	3	1	
Africa	DR Congo	Boende	8	7	
Africa	DR Congo	Kinshasa	5	1	
Africa	DR Congo	Katanga	9	3	
Africa	Ghana	Cape Coast	2	3	
Africa	Burkina Faso	Bobo-Dioulasso	1	1	
America	Peru	Arequipa	11	5	various communities
America	Peru	Cusco	18	7	various communities
America	Peru	Ica	7	1	Nazca
America	Peru	Loreto	13	12	various communities
America	Peru	Puno	12	5	various communities
America	Panama	various	0	2	US imports
America	Mexico	Mexico City	5	1	
America	Mexico	Morelia	5	4	
America	Brazil	various	4	9	
America	Dominican Republic	Sosua	2	10	
America	Honduras	Roatán	6	2	
America	Colombia	Colombia	6	3	
America	United States	Puerto Rico	5	4	see Boyko et al 2009
America	Belize	Belize	3	1	
America	Costa Rica	Costa Rica	2	6	
Arctic	United States	Alaska	6	5	coastal and interior Alaskan Village Huskies
Central Asia	Mongolia	various	11	2	Bankhar
Central Asia	Nepal	various	6	6	
East Asia	Indonesia	Borneo	6	3	
East Asia	Vietnam	Lào Cai	4	0	
East Asia	Vietnam	Hà Giang	4	1	
East Asia	Vietnam	Cao Bang	4	2	
East Asia	Vietnam	Lang Son	4	1	
East Asia	Indonesia	Jakarta	2	0	
Europe	Portugal	various	5	4	
Europe	Croatia	various	6	0	
Europe	Bosnia	various	3	0	
India	India	Orissa	6	0	
India	India	Chennai	4	2	
India	India	Dehli	6	0	
India	India	Hazaribagh	1	1	
India	India	Mumbai	6	0	
India	India	West Bengal	5	1	
Australasia	Papua New Guinea	Port Moresby	8	1	
Australasia	Papua New Guinea	Eastern Highlands	7	4	
Australasia	Palau	Palau	0	1	US import
Australasia	French Polynesia	Society Islands	10	14	Bora Bora, Huahine, Moorea, Raiatea
Australasia	French Polynesia	Marquesas	5	7	Hiva Oa
Australasia	Fiji	Viti Levu	5	11	various villages
Australasia	Fiji	Kadavu	8	4	
Australasia	Fiji	Tavenui	1	0	sampled in Kadavu
Australasia	Solomon Islands	Guadalcanal	1	2	
Australasia	Solomon Islands	Makira	1	3	
Australasia	Solomon Islands	Central	1	3	
Australasia	Solomon Islands	Western	0	4	
Middle East	Qatar	various	7	2	
Middle East	Egypt	Luxor	8	4	see Boyko et al 2009
Middle East	Egypt	Giza	1	2	see Boyko et al 2009
Middle East	Egypt	Kharga	1	0	see Boyko et al 2009
Middle East	Lebanon	Beruit	8	1	
Middle East	Lebanon	Bekaa	7	0	
Middle East	Iraq	various	1	0	US import
Middle East	Afghanistan	various	3	10	US imports
Middle East	Turkey	Istanbul	8	2	
Middle East	Turkey	Giresun	2	1	