SUPPORTING INFORMATION

Unregulated hunting and genetic recovery from a severe population decline: the cautionary case of Bulgarian wolves

Supplementary Tables

Table S1. List of samples used in this study, with detail on the number used on each analysis. Georeferenced samples were used in the spatial analysis. Note that the Jackal samples were identified molecularly during the study (see main text).

Dataset	Total	mtDNA	Microsatellites	Microsatellites without missing data	mtDNA & Microsatellites	Georeferenced
Wolf samples	109	105	102	92	98	92
Dog samples	14	0	14	14	0	0
Jackal samples	3	3	3	3	3	3

Table S2. *P*-values for Hardy-Weinberg exact tests carried out in GENEPOP (Rousset 2008) and diversity indices: F_{IS} - Inbreeding Coefficient, Ho- Observed Heterozygosity, He- Expected Heterozygosity, I- Shannon information index, all calculated in GENALEX (Peakall and Smouse 2006); All-Rich- Allelic richness, calculated in FSTAT (Goudet 2001).

Dataset	Locus	*P*-value	S.E.	Ho	He	F_{IS}	I	All-Rich
All individuals (N=102)	*2010*	0.0014	0.0001	0.576	0.719	0.199	1.348	4.97
	2017	0.0157	-	0.458	0.563	0.186	0.978	4
	2054	0.0163	0.0006	0.735	0.824	0.108	1.951	10.882
	2088	0	0	0.588	0.673	0.126	1.255	5.997
	2079	0.005	0.0002	0.414	0.526	0.213	0.867	4.969
	2096	0.0005	0	0.637	0.664	0.041	1.225	5.938
	VwF	0.1378	0.0016	0.733	0.753	0.027	1.754	9
	250	0.0022	0.0001	0.800	0.813	0.016	1.778	7.997
	253	0.1451	0.0013	0.765	0.804	0.049	1.804	7.941
	213	0.0033	0.0004	0.693	0.796	0.129	1.904	13.8
	2001	0.0236	0.0005	0.762	0.791	0.036	1.634	6.948
	130	0	0	0.624	0.770	0.190	1.684	8.901
	466	0.2788	0.003	0.696	0.750	0.071	1.598	9.765
	642	0	0	0.680	0.835	0.185	2.210	17.877
	All	NA	NA	0.654	0.734	0.113	1.571	8.499
Unrelated individuals only (N=62)	*2010*	0.0226	-	0.587	0.715	0.179	1.321	4.000
	2017	0.0128	-	0.403	0.523	0.228	0.900	4.000
	2054	0.0062	0.0003	0.703	0.823	0.146	1.958	10.937
	2088	0.0014	0.0001	0.531	0.690	0.230	1.287	5.000
	2079	0.0212	0.0004	0.444	0.559	0.205	0.954	4.984
	2096	0.0023	0.0001	0.641	0.664	0.035	1.248	5.968
	VwF	0.6578	0.0023	0.730	0.751	0.028	1.745	9.000
	250	0.3781	0.0018	0.778	0.813	0.043	1.771	7.968
	253	0.4288	0.0018	0.719	0.785	0.085	1.755	7.969
	213	0.0087	0.0007	0.683	0.786	0.132	1.807	12.905
	2001	0.2189	0.0014	0.746	0.791	0.057	1.639	6.968
	130	0.0006	0.0001	0.619	0.753	0.178	1.631	7.984
	466	0.0643	0.0012	0.594	0.753	0.211	1.617	8.906
	642	0.0001	0.0001	0.683	0.840	0.187	2.238	17.905
	All	NA	NA	0.633	0.732	0.139	1.562	8.178

Table S3. List of loci identified as having null alleles by the software MICROCHECKER (Van Oosterhout et al. 2004) for all samples pooled together, and separated by populations identified by the software STRUCTURE (Pritchard et al. 2000). Clusters A and B correspond to individuals with 60% ancestry to one of the two STRUCTURE clusters, while Cluster AB corresponds to individuals with less than 60% ancestry to both clusters.

Locus	All	Cluster A	Cluster B	Cluster AB
2010	Yes	No	Yes	No
2017	Yes	No	No	No
2054	Yes	No	No	No
2088	No	No	No	No
2079	Yes	No	Yes	No
2096	No	No	No	No
VwF	No	No	No	No
250	No	No	No	No
253	No	No	No	No
213	Yes	Yes	No	No
2001	No	No	No	No
130	Yes	Yes	No	No
466	No	No	No	No
642	Yes	No	No	No

Table S4. Results of the bottleneck tests from the software BOTTLENECK (Cornuet and Luikart 1996)

Mutation Model	Expected Het. Exc.	Observed Het. Exc.	p	Stand. Diff.	Wilcoxon H. Def.	Wilcoxon H. Exc.	Wilcoxon Two-tail
I.A.M.	8.23	14	0.00058	0.0003	1.0	0.00003	0.00006
T.P.M.	8.29	10	0.26	0.11298	0.94	0.067	0.13
S.M.M.	8.25	4	0.0215	0.00079	0.025	0.979	0.05

Table S5. Assignment probability for individuals with putative hybrid wolf-dog ancestry. These were defined as individuals having over 0.05 probability of assignment to any category other than pure wolf (marked with *; jackals are excluded) consistently between STRUCTURE (Pritchard et al. 2000) and NEWHYBRIDS (Anderson and Thompson 2002). Categories with assignment probabilities lower than 0.01 for all individuals are not shown. WBx - wolf backcross. See Figure S10 for the geographical distribution of these samples.

Sample	NEWHYBRIDS software					STRUCTURE software			
	Wolf	F2	WBx 1	WBx 2	WBx 3	Wolf 1	Wolf 2	Dog	Jackal
BW9	0.69	0.24*	0.00	0.07*	0.00	0.01	0.94	0.05*	0.00
BW61	0.72	0.13*	0.02	0.12*	0.00	0.01	0.87	0.11*	0.02
BW67	0.75	0.16*	0.00	0.09*	0.00	0.04	0.83	0.10*	0.04
BW60	0.77	0.11*	0.01	0.10*	0.00	0.12	0.76	0.11*	0.01
BW78	0.63	0.01	0.02	0.31*	0.02	0.31	0.57	0.10*	0.03
BW58	0.91	0.02	0.01	0.06*	0.00	0.60	0.34	0.06*	0.00
BW114	0.52	0.25*	0.03	0.20*	0.00	0.46	0.30	0.23*	0.01
BW116	0.93	0.01	0.00	0.05*	0.00	0.70	0.24	0.05*	0.01
BW89	0.45	0.07*	0.07*	0.39*	0.02	0.88	0.02	0.09*	0.01
BW15	0.72	0.13*	0.02	0.13*	0.00	0.92	0.02	0.06*	0.00

Table S6. Comparison of genetic diversity estimates between the present study for Bulgarian wolves, and those obtained for other wolf populations. Standard deviation is included whenever described in the paper (after the ± symbol). Values from (Czarnomska et al. 2013) result from averaging estimates for all different regions reported. Values from (Vonholdt et al. 2008) result from averaging estimates for all years reported. Ne and Ne/N values from (Sastre et al. 2011) were calculated by averaging mean values from both programs used, with the most extreme values of 95% CI between the two programs used.

Population	Hd	Pi	Ho	F_{IS} (95% CI)	Ne (95% CI)	Ne/N (95% CI)
Bulgaria (This study)	0.75 ±0.019	0.022 ±0.0003	0.654 ±0.031	0.113 ±0.019	90.4 (63.7-214.8)	0.12 (0.08-0.29)
Iberia (Sastre et al. 2011)	0.555 ±0.041	0.010(Sastre et al. 2011)	0.525	0.177 (0.1-0.23)	43.2 (29.9-73.8)	0.024 (0.015-0.037)
Russia (Sastre et al. 2011)	0.672 ±0.043	0.016	0.663	0.147 (0.07-0.2)	225.3 (75.9-1741.6)	0.12 (004-0.94)
Finland (2004) (Aspi et al. 2006)	NA	NA	0.680	0.029 (-0.052-0.08)	39.5 (19.3-98.7) 51.5 (44.3-78.3) (Jansson et al. 2012)	0.42 (0.20-1.03) 028(0.24-0.42) (Jansson et al. 2012)
Finland (2009) (Jansson et al. 2012)	NA	NA	0.615	0.108 (NA)	30.4 (25.4-45.4)	0.097 (0.08-0.15)
Poland (Czarnomska et al. 2013)	0.35 ±0.2	0.009 ±0.005	0.552 ±0.034	0.222 ±0.06	NA	NA
Apennine Peninsula (Fabbri et al. 2007)	NA	NA	0.6 ±0.02	0.03 ±0.01	NA	NA
Italy only (Verardi et al. 2006)	NA	NA	0.46 ±0.18	0.127 ±0.03	NA	NA
Yellowstone (vonHoldt et al. 2008)	NA	NA	0.71 ±0.013	-0.0411 ±0.02	16.27 ±4.62	0.28 ±0.04

Supplementary Figures

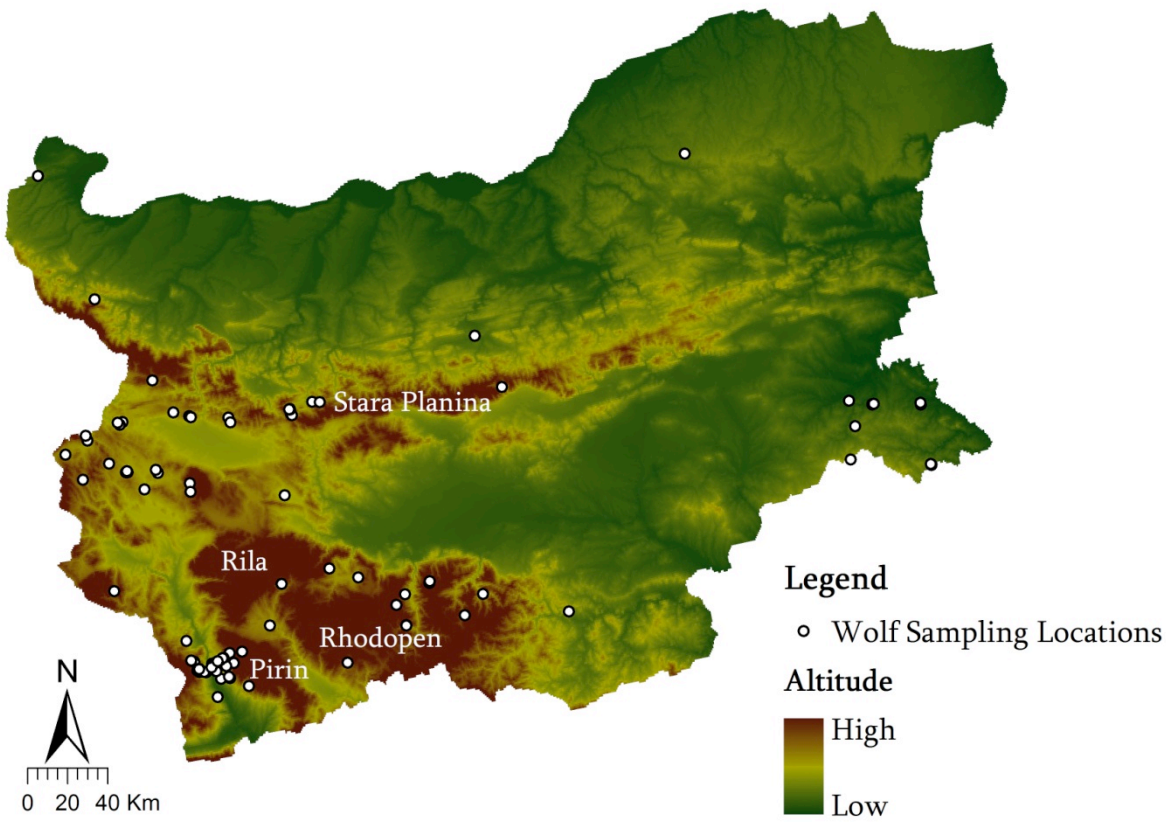

Figure S1. Altitude map of Bulgaria. The 4 main mountain systems are labelled, and the location of all samples used in the study indicate by white circles.

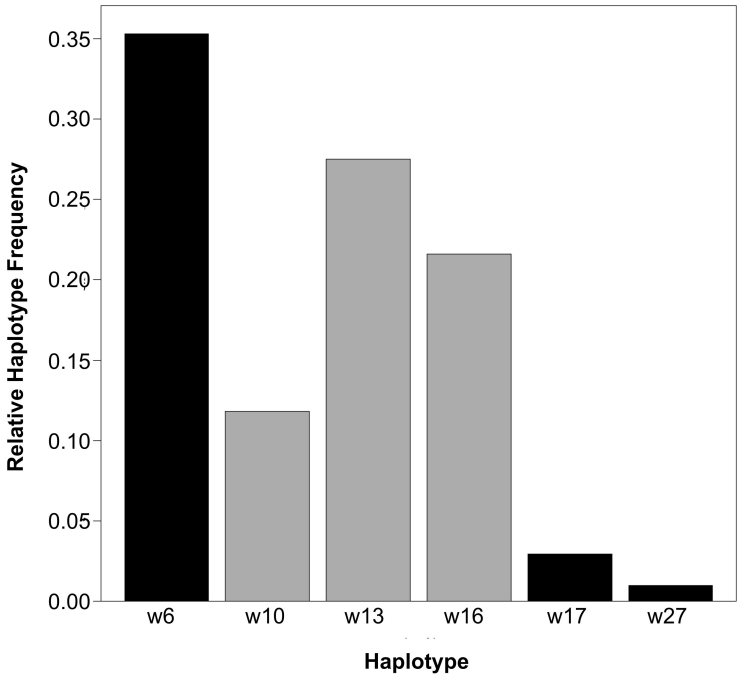

Figure S2. Frequency of mtDNA haplotypes. Haplotype designation follows the definition of (Pilot et al. 2010). Grey bars represent haplotypes belonging to haplogroup 1, while black bars represent haplotypes belonging to haplogroup 2 as defined in a previous study focusing on grey wolf phylogeography in Europe (Pilot et al. 2010). Plot drawn using ARLEQUIN (Excoffier et al. 2005)

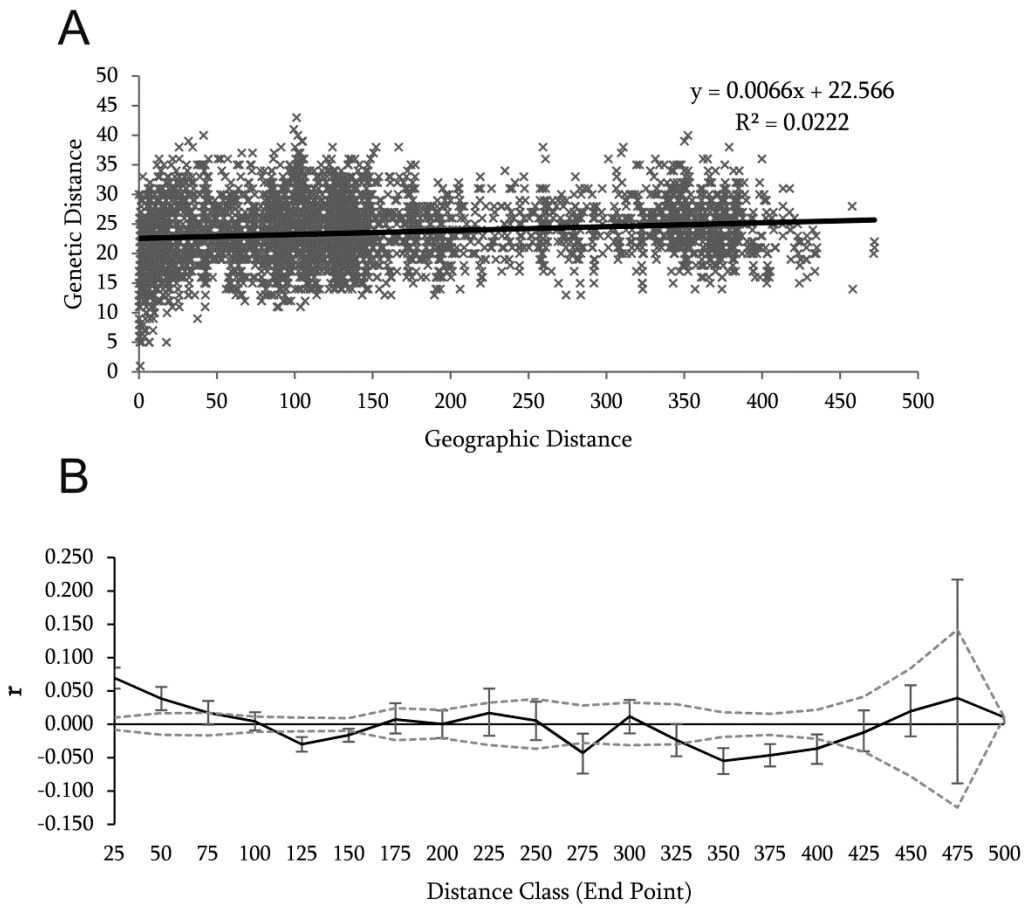

Figure S3. Results from the isolation by distance analysis as implemented in GENALEX (Peakall and Smouse 2006). A: mantel test results, with statistical significance assessed through 10,000 permutations. B: Spatial autocorrelation results, with 20 equal distance classes. r=correlation index; dashed lines represent 95% confidence limit as calculated by 10,000 permutations. Error bars in each r represent the 95% confidence limits as calculated by bootstrapping with 10,000 replicates.

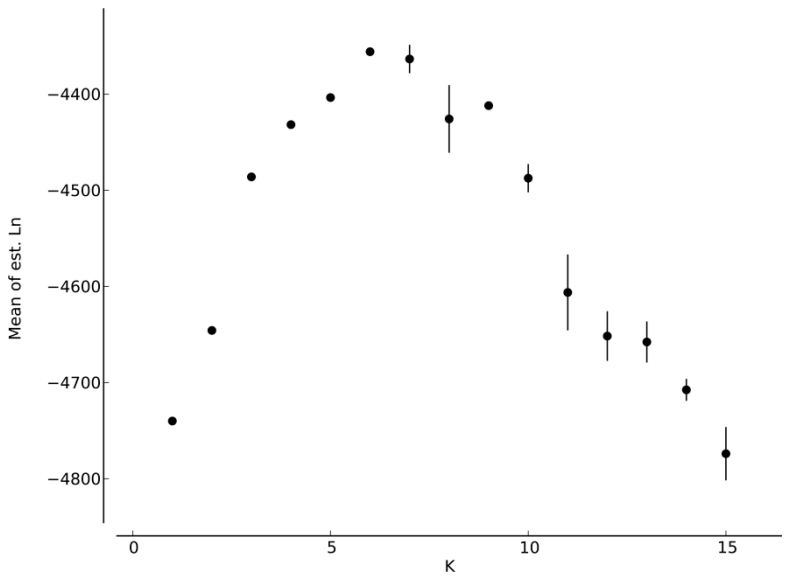

Figure S4. Posterior probability for each value of K tested using STRUCTURE for the dataset without dogs. Each point represents the mean probability of all runs, while the bars represent the range of values obtained. Graph made using STRUCTUREHARVESTER (Earl and vonHoldt 2012).

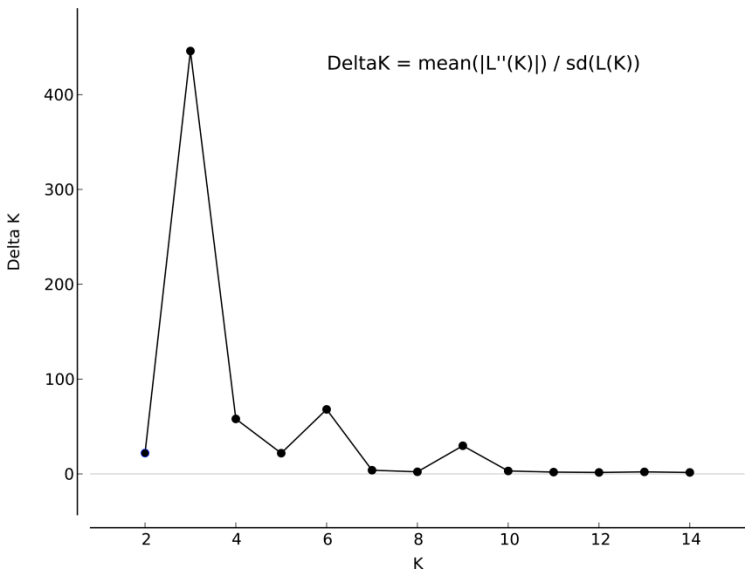

Figure S5. ΔK values for each K tested with STRUCTURE, using the method described in (Evanno et al. 2005). Graph produced in STRUCTUREHARVESTER (Earl and vonHoldt 2012).

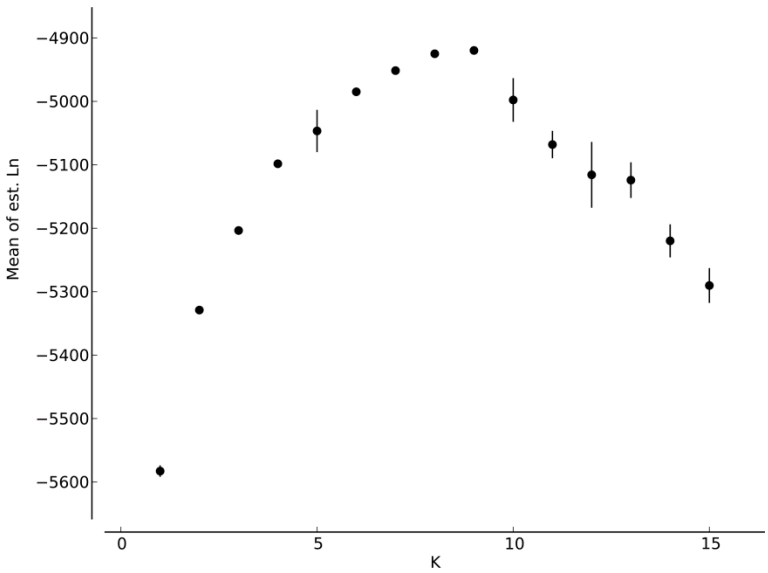

Figure S6. Posterior probability for each value of K tested using STRUCTURE for the dataset with feral dog samples added. Each point represents the mean probability of all runs, while the bars represent the range of values obtained. Due to posterior probability being virtually identical between different runs for several values of K=2 , we could not apply the (Evanno et al. 2005) method to this dataset. Graph made using STRUCTUREHARVESTER (Earl and vonHoldt 2012).

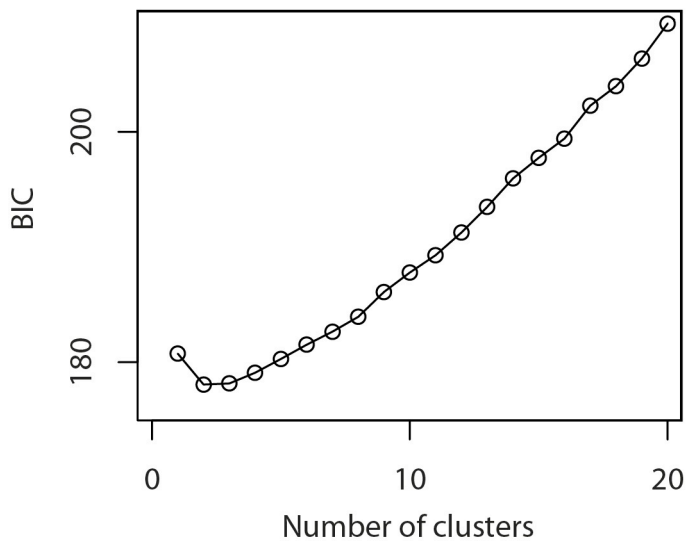

Figure S7. Values of BIC versus number of clusters using the DAPC (Jombart et al. 2010) method implemented in the R package ADEGENET (Jombart and Ahmed 2011). Support for two clusters is marginally better than support for three clusters.

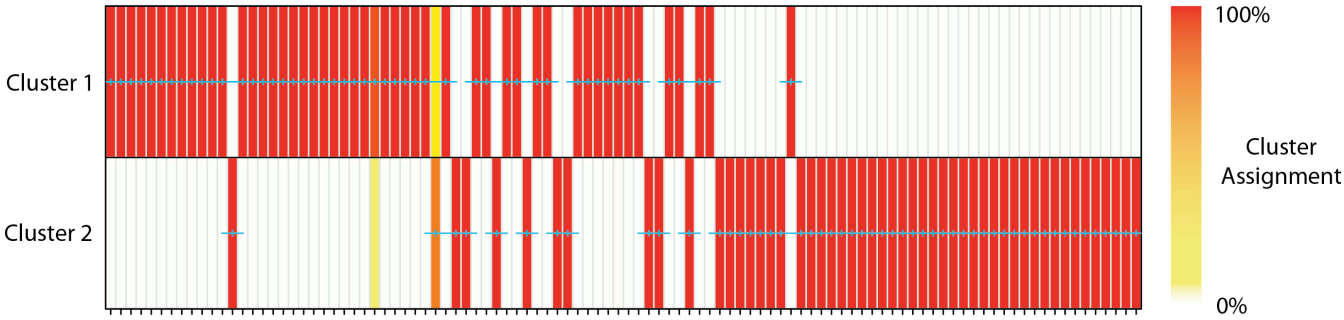

Figure S8. Assignment plot for two clusters obtained using the DAPC (Jombart et al. 2010) method implemented in the R package ADEGENET (Jombart and Ahmed 2011). Each vertical bar represents a single individual, colour coded according to assignment to both clusters, with red meaning full assignment and white meaning no assignment. Individuals are ordered as in the Structure plots in the main text.

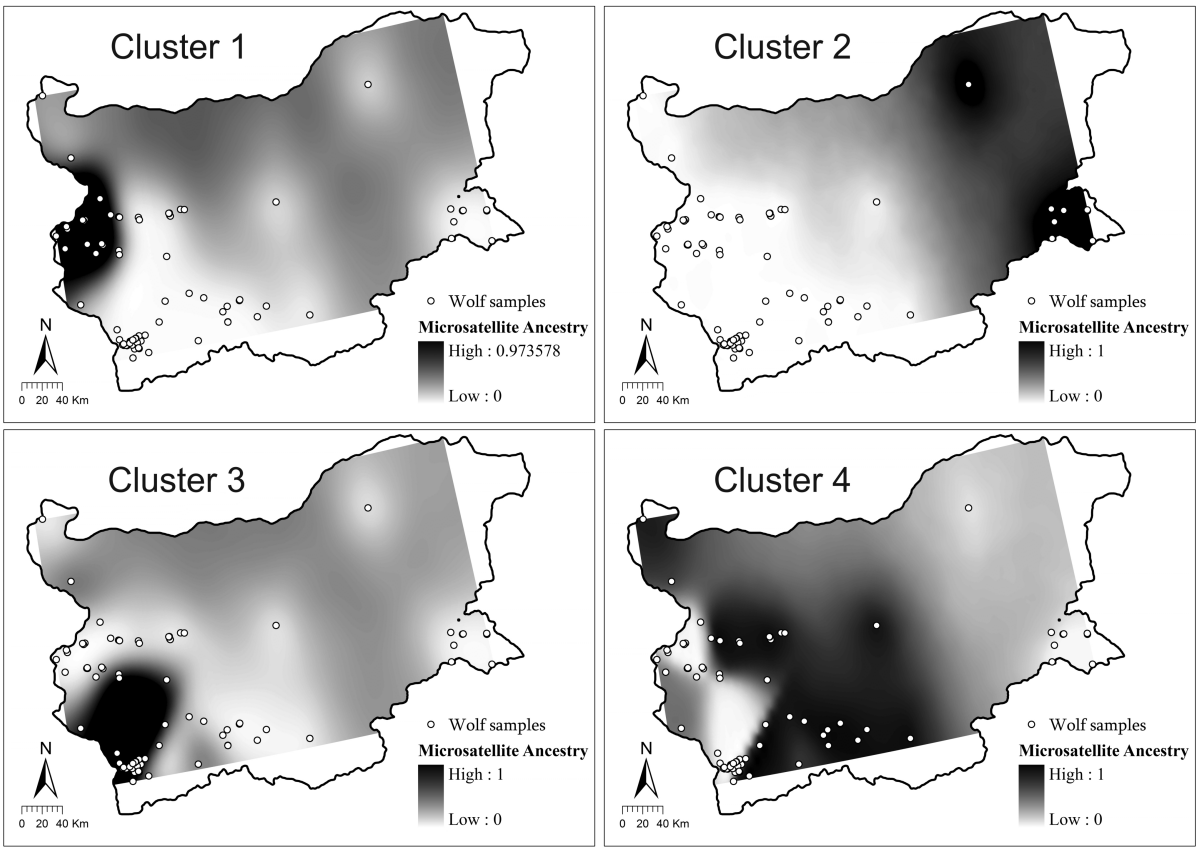

Figure S9. Microsatellite population structure as determined by GENELAND (Guillot et al. 2008) assuming four different population clusters. Individual samples used in this analysis are overlaid. Clusters are broadly coincident with those obtained for mtDNA. Cluster 4 has a higher degree of assignment than for mtDNA, but when GENELAND was run assuming three populations, Cluster 4 would alternate in grouping with any of the other 3 between runs.

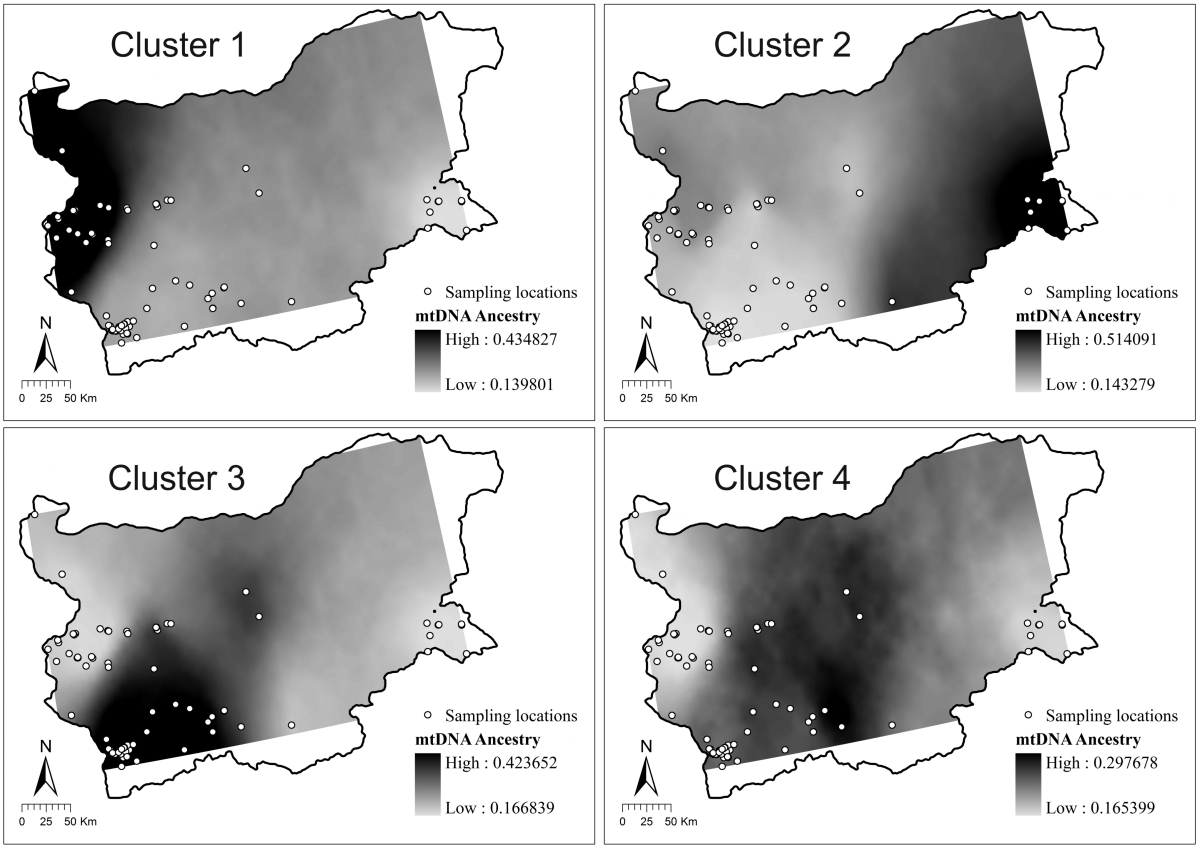

Figure S10. mtDNA population structure as determined by GENELAND (Guillot et al. 2008). Individual samples used in this analysis are overlaid, classified by their corresponding haplogroup as defined in (Pilot et al. 2010). Note how different core mtDNA clusters occupy regions with differing topography profiles shown in Figure S1.

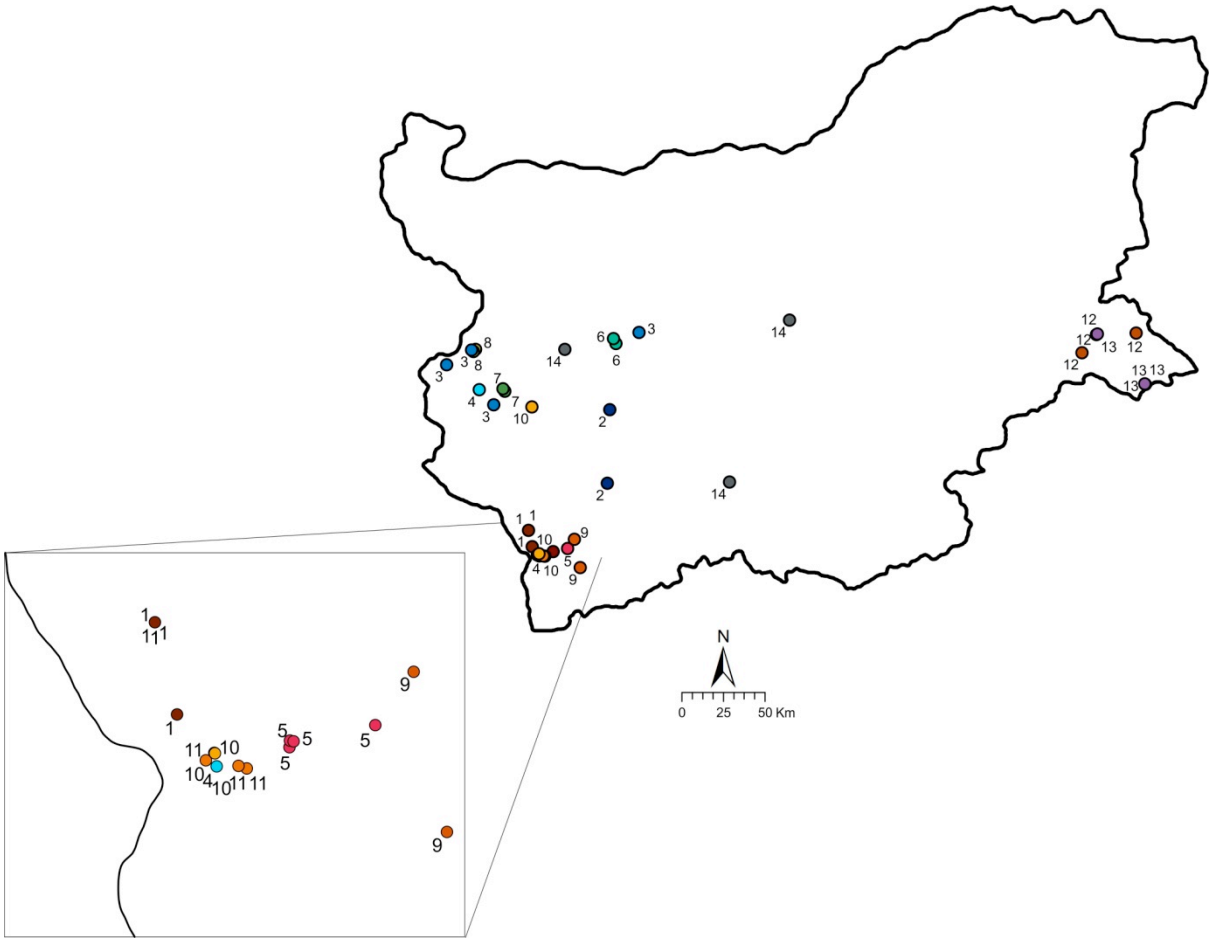

Figure S11. Distribution of core family groups. Numbers indicate the particular family group to which an individual belongs.

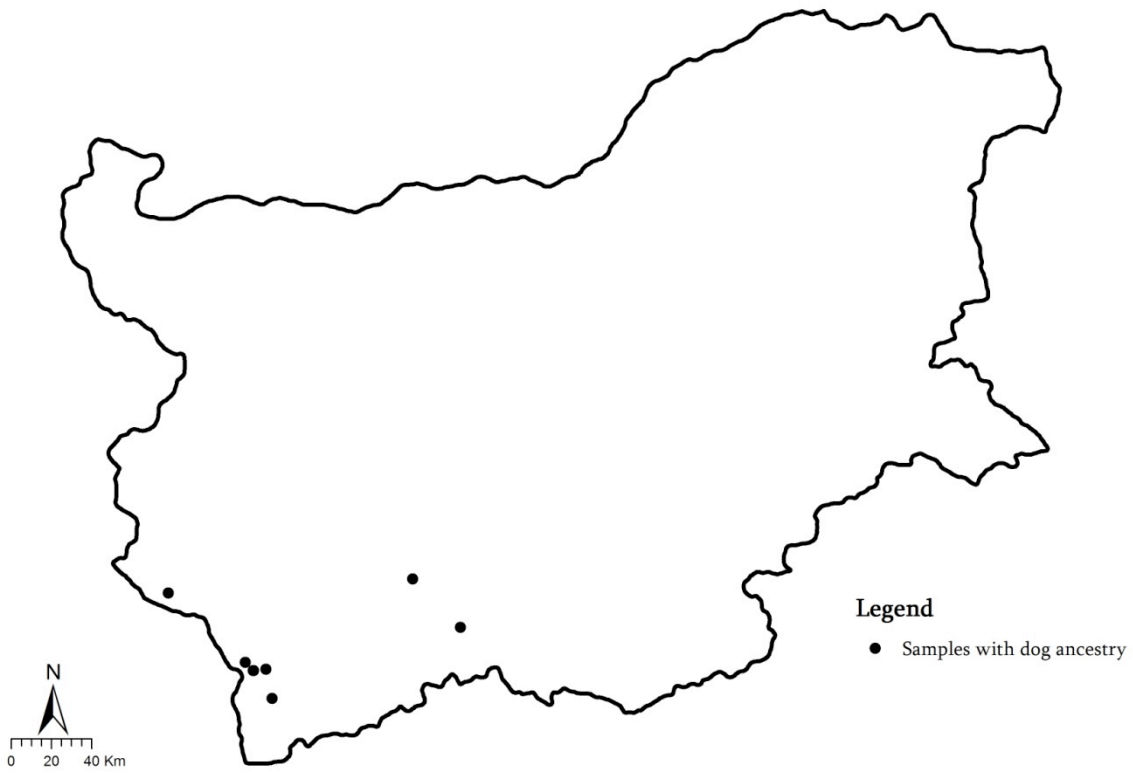

Figure S12. Distribution of individuals with hybrid dog ancestry as identified in Table S3.

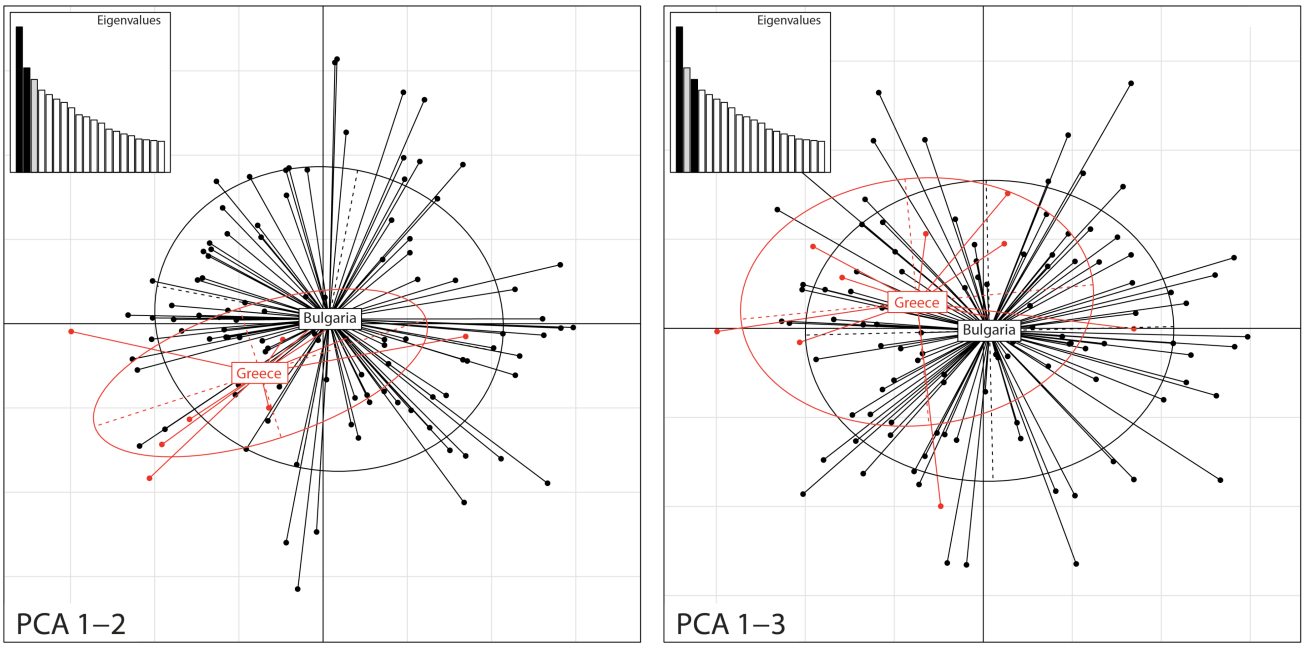

Figure S13. Principal Component Analyses comparing 10 wolf samples obtained from Greece with the Bulgarian samples. The 3 most informative components do not clearly separate Greek and Bulgarian samples. Analysis carried out using the ADEGENET software (Jombart, 2008)

References

Anderson EC, Thompson EA (2002) A model-based method for identifying species hybrids using multilocus genetic data. Genetics 160:1217-1229

Aspi J, Roininen E, Ruokonen M, Kojola I, VilÀ C (2006) Genetic diversity, population structure, effective population size and demographic history of the Finnish wolf population. Molecular Ecology 15:1561-1576.

Cornuet JM, Luikart G (1996) Description and power analysis of two tests for detecting recent population bottlenecks from allele frequency data. Genetics 144 (4):2001-2014

Czarnomska SD, Jędrzejewska B, Borowik T, Niedziałkowska M, Stronen AV, Nowak S, Mysłajek RW, Okarma H, Konopiński M, Pilot M, Śmietana W, Caniglia R, Fabbri E, Randi E, Pertoldi C, Jędrzejewski W (2013) Concordant mitochondrial and microsatellite DNA structuring between Polish lowland and Carpathian Mountain wolves. Conservation Genetics 14 :573-588.

Earl DA, vonHoldt BM (2012) STRUCTURE HARVESTER: a website and program for visualizing STRUCTURE output and implementing the Evanno method. Conservation Genetics Resources 4:359-361.

Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. Molecular Ecology 14:2611-2620

Excoffier L, Laval G, Schneider S (2005) Arlequin ver. 3.0: An integrated software package for population genetics data analysis. Evolutionary Bioinformatics Online 1:47-50

Fabbri E, Miquel C, Lucchini V, Santini A, Caniglia R, Duchamp C, Weber J-M, Lequette B, Marucco F, Boitani L, Fumagalli L, Taberlet P, Randi E (2007) From the Apennines to the Alps: colonization genetics of the naturally expanding Italian wolf (*Canis lupus*) population. Molecular Ecology 16:1661-1671.

Goudet J (2001) FSTAT, a program to estimate and test gene diversities and fixation indices (version 2.9.3). Available from http://wwwunilch/izea/softwares/fstathtml

Guillot G, Santos F, Estoup A (2008) Analysing georeferenced population genetics data with Geneland: a new algorithm to deal with null alleles and a friendly graphical user interface. Bioinformatics 24:1406-1407.

Jansson E, Ruokonen M, Kojola I, Aspi J (2012) Rise and fall of a wolf population: genetic diversity and structure during recovery, rapid expansion and drastic decline. Molecular Ecology 21:5178-5193.

Jombart T, Ahmed I (2011) Adegenet 1.3-1: new tools for the analysis of genome-wide SNP data. Bioinformatics 27:3070-3071.

Jombart T, Devillard S, Balloux F (2010) Discriminant analysis of principal components: a new method for the analysis of genetically structured populations. BMC Genetics 11:94.

Peakall ROD, Smouse PE (2006) Genalex 6: genetic analysis in Excel. Population genetic software for teaching and research. Molecular Ecology Notes 6:288-295.

Pilot M, Branicki W, Jedrzejewski W, Goszczynski J, Jedrzejewska B, Dykyy I, Shkvyrya M, Tsingarska E (2010) Phylogeographic history of grey wolves in Europe. BMC Evolutionary Biology 10:104.

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

Rousset F (2008) genepop'007: a complete re-implementation of the genepop software for Windows and Linux. Molecular Ecology Resources 8:103-106. Sastre N, Vilà C, Salinas M, Bologov VV, Urios V, Sánchez A, Francino O, Ramírez O (2011) Signatures of demographic bottlenecks in European wolf populations. Conservation Genetics 12:701-712.

Van Oosterhout C, Hutchinson WF, Wills DPM, Shipley P (2004) MICRO-CHECKER: software for identifying and correcting genotyping errors in microsatellite data. Molecular Ecology Notes 4:535-538

Verardi A, Lucchini V, Randi E (2006) Detecting introgressive hybridization between free-ranging domestic dogs and wild wolves (*Canis lupus*) by admixture linkage disequilibrium analysis. Molecular Ecology 15:2845-2855.

Vonholdt BM, Stahler DR, Smith DW, Earl DA, Pollinger JP, Wayne RK (2008) The genealogy and genetic viability of reintroduced Yellowstone grey wolves. Molecular Ecology 17:252-274.