DR. MALGORZATA  PILOT (Orcid ID : 0000-0002-6057-5768)

Article type      : Original Article

# Widespread, long-term admixture between grey wolves and domestic dogs across Eurasia and its implications for the conservation status of hybrids

Małgorzata Pilot[1], Claudia Greco[2], Bridgett M. vonHoldt[3], Ettore Randi[2,4], Włodzimierz Jędrzejewski[5,10], Vadim E. Sidorovich[6], Maciej K. Konopiński[7], Elaine A. Ostrander[8], Robert K. Wayne[9]

1. School of Life Sciences, University of Lincoln, Lincoln, UK; mpilot@lincoln.ac.uk

2. Department of Environmental Monitoring and Biodiversity Conservation, Italian National Institute for Environmental Protection and Research, Italy; claudia.greco@isprambiente.it and ettore.randi@unibo.it

3. Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ, USA; vonholdt@princeton.edu

4. Department 18/ Section of Environmental Engineering, Aalborg University, Aalborg, Denmark

5. Mammal Research Institute, Polish Academy of Sciences, Białowieża, Poland; wjedrzej1@gmail.com

6. Institute of Zoology, National Academy of Sciences of Belarus, Minsk, Belarus; vadim.sidorovich@gmail.com

7. Institute of Nature Conservation, Polish Academy of Sciences, Kraków, Poland; konopinski@iop.krakow.pl

8. Cancer Genetics Branch, National Human Genome Research Institute, National Institutes

of Health, Bethesda, MD, USA; eostrand@mail.nih.gov

9. Department of Ecology and Evolutionary Biology, University of California, Los Angeles,

CA, USA; rwayne@ucla.edu

10. Current address: Instituto Venezolano de Investigaciones Cientificas (IVIC), Centro de

Ecologia, Caracas, Venezuela

corresponding authors: Małgorzata Pilot and Robert K. Wayne

**Running head:** Wolf-dog admixture and management of hybrids

**Abstract**

Hybridisation between a domesticated species and its wild ancestor is an important

conservation problem, especially if it results in the introgression of domestic gene variants

into wild species. Nevertheless, the legal status of hybrids remains unregulated, partially

because of the limited understanding of the hybridisation process and its consequences. The

occurrence of hybridisation between grey wolves and domestic dogs is well-documented

from different parts of the wolf geographic range, but little is known about the frequency of

hybridisation events, their causes and the genetic impact on wolf populations. We analysed

61K SNPs spanning the canid genome in wolves from across Eurasia and North America and

compared that data to similar data from dogs to identify signatures of admixture. The

haplotype block analysis, which included 38 autosomes and the X chromosome, indicated the

presence of individuals of mixed wolf-dog ancestry in most Eurasian wolf populations, but

less admixture was present in North American populations. We found evidence for male-

biased introgression of dog alleles into wolf populations, but also identified a first-generation

hybrid resulting from mating between a female dog and a male wolf. We found small blocks

of dog ancestry in the genomes of 62% Eurasian wolves studied and melanistic individuals

with no signs of recent admixed ancestry, but with a dog-derived allele at a locus linked to melanism. Consequently, these results suggest that hybridisation has been occurring in different parts of Eurasia on multiple timescales and is not solely a recent phenomenon. Nevertheless, wolf populations have maintained genetic differentiation from dogs, suggesting that hybridisation at a low frequency does not diminish distinctiveness of the wolf gene pool. However, increased hybridisation frequency may be detrimental for wolf populations, stressing the need for genetic monitoring to assess the frequency and distribution of individuals resulting from recent admixture.

### INTRODUCTION

Hybridisation between defined taxonomic entities can be an important conservation problem when it involves an invasive and a native species, or a domesticated subspecies and its wild ancestor (Wayne & Shaffer, 2016). For example, hybridisation with introduced North American ruddy ducks (*Oxyura jamaicensis*) may endarger the genetic integrity of native Eurasian white-headed ducks (*Oxyura leucocephala*) (Munoz-Fuentes et al., 2007), and hybridisation with the domestic cat (*Felis silvestris catus*) is an important conservation threat to the European wildcat (*Felis silvestris silvestris*) (Lecis et al., 2006). Rapid human population growth and the spread of human-modified habitats can result in a parallel increase in domesticated species and decline of their wild relatives. Such changes in relative densities can increase the frequency of hybridisation, resulting in extensive introgression of derived "domesticated" gene variants into wild populations. Although such introgression is frequently considered maladaptive, it can also provide novel adaptations to a newly occupied or changing environment. For example, admixture between free-living Soay sheep and a modern sheep (*Ovis aries*) breed resulted in an introgression of a *TYRP1* gene variant associated with

light coat colour, which was favoured by natural selection in Soay sheep (Feulner et al., 2013). Another example comes from Alpine ibex (*Capra ibex ibex*), which was shown to acquire one of its two MHC *DRB* alleles from domestic goats (*Capra aegagrus hircus*) (Grossen et al., 2014). Our understanding of the hybridisation process and its consequences is still limited, and improving this knowledge has both theoretical importance (for understanding the role of hybridisation in speciation and adaptation) and practical applications in wildlife conservation and management of feral domestic populations.

The process of domestication is recent in an evolutionary time frame. The oldest domesticated species, the domestic dog *Canis lupus familiaris*, only diverged from the grey wolf *Canis lupus* between 11,000 and 35,000 years ago (Freedman & Wayne, 2017). Because the divergence between domesticated species and their wild relatives is recent, hybridisation between them is particularly frequent as reproductive isolation has not completely developed (Randi, 2008; Harrison & Larson, 2014). The case of wolf-dog hybridisation is particularly interesting due to extensive morphological, ecological and behavioural differences between the two subspecies, which may affect both hybridisation patterns and the fitness of admixed individuals (Anderson, et al. 2009; vonHoldt et al., 2010, 2017; Fan et al., 2016; Miao et al., 2017). Achieving a better understanding of wolf-dog hybridisation is also important from perspectives of conservation of the grey wolf, which is a keystone species in terrestrial Holarctic ecosystems. This knowledge may also contribute to better control of feral domestic dogs, which can pose a threat to both wildlife and humans (Gompper, 2014).

Domestic dogs coexist with grey wolves across the entire wolf range in the Holarctic. The relationship between the two subspecies is complex, and involves resource competition, predation, and disease transmission (Lescureux & Linnell, 2014). The two subspecies interbreed in the wild, and produce fertile offspring (Leonard et al., 2014). The context and relative frequency of different types of wolf-dog interactions is not well understood, partially

because the ecology of free-ranging dogs has not been extensively studied (but see Gompper, 2014). Therefore, although the occurrence of wolf-dog hybridisation is well documented (reviewed in Hindrikson et al., in press), little is known about its underlying ecological mechanisms. It is unknown whether hybridisation has occurred naturally at similar rate since the divergence of wolf and dog lineages, or if it has become more frequent recently as a result of the decline in wolf abundance and the parallel increase in dog numbers.

Occurrence of wolf-dog hybrids and/or back-crosses has been reported from most European populations, including Italy (Randi & Lucchini, 2002; Verardi et al., 2006; Caniglia et at., 2013; Lorenzini et al., 2014; Randi et al., 2014; Galaverni et al., 2017), the Iberian Peninsula (Godinho et al., 2011; Pacheco et al., 2017), North-Eastern Europe (Latvia and Estonia - Andersone et al., 2002; Hindrikson et al., 2012), the Balkans (Moura et al., 2014), and the Scandinavian Peninsula (Vilà et al., 2003). There are considerably fewer studies on Asian wolf populations, but recently the occurrence of wolf-dog hybridization has been reported from the Caucasus (Kopaliani et al., 2014; Pilot et al., 2014) and Iran (Aghbolaghi et al., 2014; Khosravi et al., 2013, 2015). All these studies focused on relatively small geographical areas, and therefore, little is known about geographical variation in the occurrence and frequency of admixed individuals which, if known, could shed a light on factors that favour hybridisation.

The extent of back-crossing of hybrids into wolf populations is also unknown. Studies based on microsatellite loci failed to reveal large-scale introgression of dog alleles into European wolf populations, despite the evidence that hybrids can be reintegrated into wolf populations (Andersone *et al*., 2002; Randi & Lucchini, 2002; Ciucci et al., 2003; Vilà et al., 2003; Godinho et al., 2011; Lorenzini et al., 2013). In contrast, genome re-sequencing data showed that Eurasian wolf genomes may have up to 25% of dog ancestry, and wolf populations with no signs of dog ancestry are rare in Eurasia (Fan et al., 2016).

Effective management of wolf populations that may be affected by hybridisation requires a clear understanding of how hybrids are defined and identified, and how their presence affects population viability. However, the presence of individuals with varying levels of dog ancestry in a population may make the distinction between pure and admixed individuals ambiguous. Individuals resulting from recent hybridisation are difficult to detect based on morphological features (Lorenzini et al., 2014), which may compromise efforts to eliminate them from wolf populations. Moreover, the introgression of dog alleles into wolf populations is not always maladaptive (Anderson et al., 2009; Coulson et al., 2011), and therefore it is unclear whether elimination of admixed individuals is always the most appropriate conservation strategy.

International, EU, US and national laws on endangered species conservation lack specific legislation on management of hybrids (Allendorf et al., 2001; Trouwborst, 2014; Wayne & Shaffer, 2016). Prevention and mitigation of wolf-dog hybridsation may be essential to comply with the Bern Convention on European Wildlife and Natural Habitats and the EU Habitats Directive. On the other hand, the prohibitions on the killing and capturing of the wolves introduced by these both legal frameworks also cover wild wolf-dog hybrids (Trouwborst, 2014). Therefore, better knowledge of the hybridisation process is needed to guide conservation legislation and practice (Wayne & Shaffer, 2016). In this study, we used genome-wide SNP data to analyse ancestry in Eurasian grey wolves in order to detect first generation dog-wolf hybrids, recent back-crosses, and signatures of more distant hybridisation events. This approach allowed us to assess the effect of hybridisation on grey wolves at a continental scale.

## MATERIALS AND METHODS

### Dataset

This study utilized previously reported genome-wide SNP data from wild canids and domestic dogs genotyped on an Affymetrix Canine SNP Genome Mapping Array at 60,584 high-quality autosomal SNP loci and 851 X chromosome SNP loci (vonHoldt et al., 2010). The original dataset of 225 grey wolves, 60 coyotes and 912 domestic dogs (vonHoldt et al., 2010), was previously used in studies focused on dog domestication (vonHoldt et al., 2010), the genetic architecture of morphological traits in the domestic dog (Boyko et al., 2010), wolf-coyote hybridisation in North America (vonHoldt et al., 2011) and signatures of selection in North American wolves (Schweizer et al., 2016). Here, we utilise this dataset in a novel way to study wolf-dog hybridisation in the wild.

From the original dataset, we selected 252 individuals: 54 Eastern European wolves, 20 Italian wolves, 6 Iberian wolves, 17 putative wolf-dog hybrids (9 from Eastern Europe and 8 from Italy), 28 Asian wolves (5 from Saudi Arabia, 7 from Israel, 2 from Oman, 1 from Iran, 3 from India, and 10 from China), 125 dogs of different breeds (1-2 individuals per breed), and two free-ranging non-breed dogs. Geographic distribution of the samples of wolves and admixed canids is presented in Figure 1. The dataset of pure-breed dogs besides modern breeds of European origin also included ancient breeds. The group of ancient breeds encompasses non-European breeds, largely of Asian origin, that are genetically distinct from breeds of European origin, as first proposed by Parker et al. (2004). In addition, we included 35 coyote (*Canis latrans*) genotypes from vonHoldt et al. (2010) dataset, representing most of the species range (from California to Vermont, and from Alabama to Manitoba). Coyotes distribution is limited to North America and therefore they do not interbreed with Eurasian wolves.

Putative wolf-dog hybrids were identified *a priori* based on genetic analyses using microsatellite loci (Randi & Lucchini, 2002) and/or morphological anomalies as compared to the typical grey wolf phenotype (e.g. dewclaws; Ciucci et al., 2003). Morphological anomalies alone do not allow for reliable identification of wolf-dog hybrids, and therefore we used them for indicative purposes only. We also analysed two individuals from Italy with black coat colouration, which is a dog-derived trait, but does not necessarily imply a recent admixed ancestry (Anderson et al., 2009; Caniglia et al., 2013; Galaverni et al. 2017). The inclusion of the pre-identified putative hybrids into our dataset allowed us to assess the accuracy of hybrid detection based on a small number of microsatellite loci, but it prevented us from estimating the frequency of hybridisation in the populations studied.

For comparative purposes, we also assessed the occurrence of dog admixture in North American wolves. For this analysis, we used genotypes of 48 individuals from vonHoldt et al. (2010) dataset, representing the following populations: Mexico (5 individuals), Yellowstone (18), Northern Quebec (6), forest (11), taiga (4) and tundra (3) habitats in Northern Canada, and Vancouver Island, British Columbia (1 individual). We excluded individuals from the wolf-coyote hybridisation zone in the Great Lakes Area (vonHoldt et al., 2011). For these populations, we would have to consider dog-wolf-coyote admixture and the outcome of such analysis would not be directly comparable to the analysis of dog-wolf admixture in Eurasia.

**Detection of admixed individuals based on global ancestry estimates**

To identify signatures of dog ancestry in Eurasian wolf populations and assess the accuracy of prior hybrid identification, we used the Bayesian clustering approach implemented in STRUCTURE (Pritchard et al., 2000) and ADMIXTURE (Alexander et al., 2009). We used both STRUCTURE and ADMIXTURE, as this allowed us to ensure that the inferred

admixture patterns are consistent and thus reliable. This analysis was based on the dataset consisting of European wolves, putative wolf-dog hybrids, dogs, and coyotes.

Prior to the analysis of population structure, we used PLINK (Purcell et al., 2007) to prune the dataset removing SNPs with genome-wide pairwise genotypic association coefficient $r^2 \geq 0.5$. Pruning was carried out with 50 SNP sliding windows, shifted and recalculated every 10 SNPs, yielding a set of loci that are not in strong linkage disequilibrium (LD). We also removed loci that were invariant in the analysed sample set, or had minor allele frequency (MAF) below 0.01, resulting in a dataset of 53,248 SNPs.

We ran STRUCTURE using 100,000 MCMC iterations preceded by 20,000 burn-in iterations with five replicates for K (the number of groups) from 1 to 10 on the pruned dataset. We used the correlated allele frequencies and admixture model, and checked whether the run parameters reach convergence within the burn-in period for each K. We used Structure Harvester (Earl & vonHoldt, 2012) to assess the optimal K value, based on likelihood values and Evanno et al. (2005) delta K values.

We ran ADMIXTURE analysis for K from 1 to 15, using the default termination criterion, which stops iterations when the log-likelihood increases by $< \varepsilon = 10^{-4}$ between iterations. The K value for which the model has best predictive accuracy was identified using a cross-validation procedure, where the runs are performed after removing 10% of the genotypes at random, with 10 repetitions. We assessed the optimal K as the value that resulted in the lowest cross-validation error.

To visualize the dominant components of variability in the dataset and the position of the putative hybrids relative to wolves and dogs, we carried out the principal component analysis (PCA), using SMARTPCA package from the software EIGENSOFT (Price et al., 2006).

**Ancestry block analysis in Eurasian wolves using LAMP software**

We used LAMP (Sankararaman et al., 2008) to carry out the ancestry block analysis, which infers blocks of wolf and dog ancestry along chromosomes in each individual. This analysis allowed us to assess the admixture status of the putative wolf-dog hybrids, identify additional admixed individuals, and assess the signatures of past admixture between wolves and dogs. LAMP's unique feature among software performing the ancestry block analysis is that it allows ancestry blocks estimation without defining *a priori* ancestral populations (wolves and dogs without signatures of past admixture in their ancestry). We used this feature in our analysis, because we were not able to identify *a priori* which individuals were pure wolves without past dog admixture. Instead, the identification of ancestral populations was an integrated part of the LAMP analysis. This was achieved in a similar way to the STRUCTURE analysis with K=2, which divides a dataset analysed into two genetic clusters without any prior information about population subdivision.

For the LAMP analysis we used the dataset consisting of wolves, putative hybrids, and pure-breed dogs. We assumed a mixture proportion of 0.50:0.50, which was the frequency of wolves and putative hybrids (125 individuals) *versus* dogs (125 individuals). The use of this ratio was based on a conservative assumption that the putative hybrids group with wolves rather than dogs. This assumption is supported by the fact that the set of the putative hybrids can include backcrosses besides F1 hybrids. Back-crossing into wolf populations is more likely than into dog populations, especially given that our dataset consisted of pure-breed dogs with breeding patterns controlled by humans.

All SNPs (61K) were included in the initial dataset, which was subsequently pruned for loci that were monomorphic for the analysed set of individuals, and for $r^2 > 0.1$. We used a recombination rate of 5e-10, and fraction of overlap between adjacent windows (offset) of

0.2. We assumed a recent admixture (10 generations since admixture), because otherwise the power to detect F1 hybrids and recent backcrosses was diminished: the assumption of 100 generations since admixture resulted in all individuals (wolves and dogs) being admixed.

**Ancestry block analysis in North American wolves using LAMP software**

To assess the signatures of past wolf-dog admixture in North America, we carried out the LAMP for North American grey wolves as described for Eurasian wolves. In this case we used 48 pure-breed dogs (1 individual per breed) to match with 48 wolves, in order to maintain the 0.50:0.50 mixture proportion. All other parameters were the same as described above.

**Ancestry block analysis in Eurasian wolves using PCADMIX software**

To assess the accuracy of the local ancestry inference in LAMP, we replicated the ancestry block analysis using PCADMIX software (Brisbin et al., 2012). This software does not carry out an unsupervised ancestry assignment, i. e. it requires prior information regarding allele frequencies in non-admixed populations. However, it has been shown to have better accuracy than the LAMP analysis using a supervised ancestry assignment mode (Brisbin et al. 2012). PCADMIX uses an algorithm based on the Principal Component Analysis to determine local ancestry along each chromosome for phased SNP genotype data. We phased the genotypes using FASTPHASE (Scheet & Stephens, 2006); the wolf and dog genotypes were phased together, with population information (wolf or dog) provided in an input file. Individuals previously identified as back-crosses were assigned to the wolf population for the purpose of phasing.

In the comparison between two methods of supervised ancestry assignment, the same set of non-admixed individuals from ancestral populations would be used. In LAMP, the

ancestral populations were only identified during the analysis, and therefore we used the results from LAMP to pre-define the set of non-admixed individuals to be used as an input for PCADMIX analysis. For wolves, the criteria for individuals to be included to a non-admixed set were: (1) an average proportion of autosomal SNP alleles of dog ancestry, as identified in LAMP, lower than 0.005; and (2) no more than one chromosome having over 10% of SNP alleles of dog ancestry. These criteria were met by 48 wolves. For dogs, we applied more strict criteria, as their overall level of admixture was lower than in wolves. The dogs included in the non-admixed set had: (1) an average proportion of autosomal SNP alleles of wolf ancestry lower than 0.003; and (2) each chromosome having less than 10% of SNP alleles of wolf ancestry. These criteria were met by 107 dogs. The admixture status of the remaining individuals was assessed using PCADMIX. The phased genotypes were pruned from loci in strong linkage disequilibrium (with $r^2 > 0.8$), and the analysis was carried out in windows of 20 SNPs. In contrast to LAMP, which assumed the markers to be nearly independent and therefore required heavy pruning to achieve $r^2 < 0.1$, PCADMIX accommodates nonindependent markers and therefore can use more relaxed criteria for LD pruning (Brisbin et al., 2012).

**Analysis of X chromosome data**

The X chromosome data were analysed for males and females separately. For males, we excluded SNPs from the Pseudo-Autosomal Region (PAR; first 6 Mb of the X chromosome). Outside the PAR, we found four loci for which 10 or more of 102 males had heterozygous calls. Because we could not explain this observation, we removed these loci from both males and females datasets. At remaining 508 SNP loci no more than 5 out of 102 males displayed heterozygous calls. These were most likely genotyping errors, and were treated as missing data. Although this implies a 5% error rate at some loci, most loci did not display any heterozygous calls, and the overall error rate was 0.075%. This is consistent with

the genotyping error rate for the entire microarray data, which was estimated for samples run in duplicates at less than 0.1% (Boyko et al. 2010). After the adjustments described above, we obtained X chromosome haplotypes for the males. We calculated genetic distances between these haplotypes as the proportion of SNP sites at which two haplotypes being compared are different, and constructed the neighbour-joining tree in MEGA (Tamura et al., 2011). The same set of 508 SNPs was analysed in LAMP.

For females, we used the same set of 508 loci as for males to phase the genotypes in FASTPHASE, using the homologous male haplotypes as additional input to enhance the results. We constructed the neighbour-joining tree for the inferred female haplotypes. We also used the entire set of 851 X chromosome SNPs for ancestry blocks analysis in LAMP and for population structure analysis in ADMIXTURE, carried out as described above for autosomes.

## Assessment of heterogeneity of dog ancestry proportions across the chromosomes in Eurasian wolves

In recently admixed populations, differences in ancestry proportions may arise among chromosomes as a result of non-random mating and selection. We applied the Chromosomal Ancestry Differences (CAnD) test (McHugh et al., 2016) to assess whether there are significant differences in dog ancestry contributions among the chromosomes in Eurasian wolves. Details of this analysis are described in Supporting Information, Part A.

## Analysis of a dataset of European wolves and European dog breeds

We carried out the analyses described above for a dataset consisting of European wolves and European dog breeds only, to assess whether the methods we applied provide

consistent identification of admixed individuals, independent of the context, i.e. the included wolf and dog datasets. Details are described in the Supporting Information, Part B.

**Estimation of heterozygosity, autozygosity and linkage disequilibrium**

We calculated observed and expected heterozygosity in wolf populations from different parts of Eurasia based on the 61K SNP set for autosomal chromosomes. To minimise the bias in heterozygosity estimates due to sample size, we included only the local populations with at least five individuals sampled, and selected a random subset of six individuals from each population where the total sample size was larger. We considered non-admixed Italian wolves and Italian admixed canids (the admixture status being confirmed/identified in this study) as separate groups. Admixed individuals were also excluded from calculations for other populations, except for Israel, where all individuals carried signatures of past admixture (see Results).

To assess the autozygosity level in admixed individuals and non-admixed wolves from different regions, we identified Runs Of Homozygosity (ROHs) in individual canids spanning at least 25 SNPs and longer than 100 kb. This analysis was carried out using the SNP set pruned for local LD (by removing SNPs with $r^2 > 0.5$) to minimize the detection of ROHs that result from strong LD and do not represent autozygosity.

To compare LD levels between admixed and non-admixed populations we calculated $r^2$ between all pairs of autosomal SNPs with a minor allele frequency >0.15 in each European population, based on 5-6 individuals each to minimize sample size effect. We estimated the distance at which $r^2$ coefficient decays below 0.5. All the above analyses were carried out in PLINK.

**RESULTS**

**Identification of admixed individuals using Bayesian clustering methods**

Results from STRUCTURE and ADMIXTURE were highly consistent for K-values between 2 and 4. At K=2 dogs were distinguished from wild canids, and at K=3 dogs, wolves and coyotes were identified as the three distinct groups. Italian wolves were indicated as the fourth group at K=4 (Figure 2, Supporting Figure S1). STRUCTURE and ADMIXTURE results were inconsistent regarding the optimum number of genetic clusters (K value). Using STRUCTURE, the highest delta K value was for K=2 and the second highest for K=3. In ADMIXTURE the lowest cross-validation error was obtained for K=6. Both in STRUCTURE and ADMIXTURE, the genetic clusters identified at K=6 included coyotes, two groups of dogs (see below), and clusters of Italian wolves, other European wolves, and Saudi Arabian wolves, with other Asian wolves having intermediate assignment values between the European and the Saudi Arabian clusters (see Supporting Information, Supporting Figure S1). The clustering patterns in dogs at K=6 differed between STRUCTURE and ADMIXTURE, but these patterns do not affect our inference regarding wolf-dog admixture. The allele frequency divergences among populations estimated in STRUCTURE are reported in Supporting Table S1).

Clusters identified at K=2 and K=3 corresponded with the three canid species/subspecies analysed, which allowed for identification of hybrids and backcrosses. Among individuals identified *a priori* as putative hybrids (nine from Eastern Europe and eight from Italy), we identified only one F1 hybrid, Italian canid #2757. This individual had about 45% assignment to the dog cluster in both STRUCTURE and ADMIXTURE at K=2. Four other putatively admixed canids from Italy and one from Eastern Europe had assignment probabilities the dog cluster of 10-17% (Figure 3, Supporting Table S2), suggesting that they were F2 or F3 back-crosses (offspring of F1 hybrids or F2 back-crosses breeding with pure

wolves). The remaining individuals identified *a priori* as putative hybrids had assignment probabilities to the dog cluster within the range for non-admixed wolves.

Two canids from European populations that were assumed *a priori* to be non-admixed wolves, had assignment probabilities to the wolf population ~83% and ~86%, respectively, which was outside the range for other European wolves (92-100%), but within the range for F2/F3 back-crosses (83-90%). Two other individuals from Eastern Europe had ambiguous admixture status, with assignment probabilities to the wolf population of about 90% inferred in both STRUCTURE and ADMIXTURE (Supporting Table S2).

Asian wolves had higher assignment probabilities to the dog cluster as compared with European wolves, ranging from 11-13% in Arabian Peninsula wolves, 9-15% in Chinese wolves, and 6-9% in Indian and Iranian wolves (Table 1). Four canids from the Arabian Peninsula (Israel and Oman) that were assumed *a priori* to be pure wolves, had assignment probabilities to the dog cluster of 15-21% (Figure 3; Supporting Table S2). This was outside the range for other Arabian Peninsula wolves, but within the range for European F2/F3 back-crosses, so these individuals could be back-crosses as well. However, both STRUCTURE and ADMIXTURE inferred some level of dog admixture in all Arabian Peninsula wolves, and therefore this inference is less robust compared with that for European canids.

To quantify uncertainty in ancestry estimates, we calculated 95% intervals for assignment probabilities from STRUCTURE, as well as standard errors for the cluster membership estimates from ADMIXTURE, which we used to calculate the 95% confidence intervals. For all individuals identified as F2/F3 back-crosses, the upper limits of both intervals were below the range of values for non-admixed individuals from the same geographic regions (Supporting Table S3). This supports our conclusion that these estimates reflect admixture rather than uncertainty of ancestry estimates.

In contrast, for most individuals with the admixture status classified as uncertain (see Supporting Table S3), the upper limits of both intervals were within the range of values for non-admixed individuals. This does not preclude these individuals as being further generation back-crosses, given the continuity of the assignment probability values from 0.75 to 1 in the wolf population (Supporting Figure S4). However, based on the existing genotyping data we cannot reliably distinguish F4 or further generation backcrosses from non-admixed individuals.

The dataset analysed also included coyotes, which at K=3 were identified as a distinct genetic cluster alongside dogs and wolves (Supporting Figure S1). Coyotes had assignment probabilities to the dog cluster between 0 and 15%, and to the wolf cluster between 0 and 7%. Coyotes hybridise with both dogs and North American grey wolves (vonHoldt et al., 2011), so this result could possibly reflect admixture. However, at K=3 most wolves (all coming from Eurasia) were also shown to have some share of coyote ancestry, and therefore it is more likely that these positive assignment probabilities reflect the common ancestry of wolves and coyotes. The inferred proportion of coyote ancestry in wolves declined with increasing K, but this was not the case for the inferred proportion of wolf and dog admixture in coyotes (Supporting Figure S1).

In the PCA plot (Figure 4; Supporting Figure S2), canids identified as the F1 hybrid and F2/F3 backcrosses based on STRUCTURE and ADMIXTURE analyses were distinct from their respective wolf populations and closer than other wolves to the dog cluster. The results of this analysis are described with more detail in Supporting Information, Part C.

**Ancestry block analysis in Eurasian wolves**

The ancestry block analysis carried out using LAMP identified two genetic clusters corresponding to wolves and dogs, and most individuals showed limited signs of admixed

ancestry (Figure 5; Supporting Figure S3). Across all autosomal chromosomes, the mean percentage of SNP alleles of dog ancestry was less than 5% for each wolf except for the few individuals discussed below (see also Supporting Table S2). However, only 41 out of 108 (38%) genotyped wolves had less than 10% of SNP alleles of dog ancestry on each chromosome, and only 25 of 108 (23%) wolves were completely free of small chromosomal blocks of assigned dog ancestry (i.e. had no SNP alleles of inferred dog ancestry). For European dog breeds, the mean percentage of SNP alleles of wolf ancestry was no higher than 0.3%. For ancient non-European breeds this percentage was between 0.2 and 11%.

LAMP results confirmed the admixture status of the F1 hybrid and all F2/F3 backcrosses from European populations identified based on STRUCTURE and ADMIXTURE analyses (Figure 3; Supporting Table S2). The F1 hybrid (individual #2757) revealed 50% of SNP alleles of dog ancestry spanning all autosomal chromosomes, with one copy of each chromosome having dog ancestry and the other wolf ancestry (recombination was not inferred in the LAMP analysis). Individuals classified as F2/F3 back-crosses based on STRUCTURE and ADMIXTURE results had 10-24% of SNP alleles of dog ancestry, consistent with the expected values (25% for F2 and 12.5% for F3 backcrosses). Additionally, two other individuals from Eastern Europe had 6-7% of SNP alleles of dog ancestry, which was marginally outside the range for other wolves (0-5%). These individuals could be back-crosses of further generations, but there is no strong support for this.

All individuals from Israel showed a relatively high percentage of SNP alleles of dog ancestry (5-14%), which is consistent with the STRUCTURE and ADMIXTURE analyses. Two individuals from Oman had only about 3% of SNP alleles of dog ancestry as inferred in LAMP, but had relatively high assignment probabilities to the dog cluster in STRUCTURE and ADMIXTURE analyses (Figure 3; Supporting Table S2).

We also counted the number of chromosomes for which the percentage of hybridisation-derived SNP alleles (i.e. alleles of wolf ancestry in dogs, and alleles of dog ancestry in wolves and putative hybrids/back-crosses) was higher than 10%. This number was low for European dog breeds (range: 0-1 chromosomes) and non-admixed wolves (0-8 chromosomes). For individuals identified as F2/F3 back-crosses, this number was considerably higher (12-23 chromosomes), while for individuals with uncertain admixture status we observed intermediate values of 3-11 (Supporting Table S2).

The distribution of individuals with varying levels of admixed ancestry differed between wolves and dogs (Supporting Figure S4). In dogs, the majority of individuals had a very low percentage of SNP alleles of wolf ancestry, and only a small proportion of individuals, largely from ancient non-European breeds, had a higher percentage of wolf-derived alleles. In contrast, the Eurasian wolf populations represented a continuous range of admixture levels from individuals with no detectable dog ancestry to an individual with 24% of dog ancestry, which is consistent with a F2 backcross.

Due to the continuity of admixture levels, we were not able to distinguish between F2 and F3 back-crosses, or between further generations of back-crosses and non-admixed wolves (see also vonHoldt et al., 2013). However, the F1 hybrid could be identified without unambiguity, as this individual had one copy of each autosomal chromosome originating from wolves and one from dogs.

**Ancestry block analysis for European wolves and European dog breeds**

The results from a dataset limited to European wolves and dog breeds of European origin only were in strong agreement with the results described above (Supporting Information Part B, Supporting Table S4). This demonstrates that the framework of the analysis (e.g. separate analysis for each wolf population versus joint analysis of different

populations) does not affect the ability to detect hybrids and back-crosses based on genome-wide SNP data.

**Ancestry blocks analysis in Eurasian wolves using PCADMIX software**

The inference of the dog admixture patterns in Eurasian wolves obtained from PCADMIX was consistent with the inference from LAMP. The average proportions of wolf ancestry in the assessed wolf dataset were 0.958 and 0.955 based on LAMP and PCAdmix analyses, respectively. There was a strong correlation between the dog ancestry proportions in individuals inferred using both methods (Spearman's rank correlation, rho = 0.875, $P = 2.62 \times 10^{-25}$; see Supporting Figure S5A). There was also a strong correlation in the number of chromosomes identified as admixed using the two methods (Spearman's rank correlation, rho = 0.854, $P = 6.07 \times 10^{-23}$; see Supporting Figure S5B). A chromosome was assumed to be admixed if it contained at least 10% of SNP alleles of dog ancestry assigned in LAMP, or at least 10% of windows of dog ancestry assigned in PCADMIX. The comparison of LAMP and PCADMIX further confirmed the continuous distribution of dog ancestry proportions in individuals from Eurasian wolf populations (Supporting Figure S5).

The individuals identified as F2/F3 hybrids based on the results from LAMP, STRUCTURE and ADMIXTURE analyses had also similarly high dog ancestry proportions inferred in PCADMIX, with the exception of individual #11254 from Spain (Figure 3; Supporting Table S2).

**Geographic distribution of admixed individuals in Eurasia**

Based on combined results from LAMP, STRUCTURE and ADMIXTURE analyses, we identified one F1 hybrid and nine F2/F3 back-crosses among 108 wolves and 17 putatively admixed individuals from Eurasia (Figure 3). Most of these individuals (one F1 hybrid and

seven F2/F3 back-crosses) were found in European populations, which could be because all 17 putatively admixed individuals identified *a priori* came from Europe. F2/F3 back-crosses were found in all European populations studied: two among 71 individuals from Eastern Europe, four among 20 individuals from Italy, and one among six individuals from the Iberian Peninsula. In the Arabian Peninsula, we identified two F2/F3 back-crosses among 14 individuals. These numbers cannot be used to reliably assess the frequency of recenly admixed individuals in the populations, because of small sample size and the presence of pre-selected putative hybrids in the sample. However, our results show that hybridisation is geographically widespread in Eurasian wolf populations.

**Ancestry block analysis in North American wolves**

Most North American wolves analysed here showed limited signs of dog admixture. With the exception of Mexican wolves and an individual from British Columbia (discussed below), each individual showed less than 0.6% of SNP alleles of dog ancestry (Table 1). Mexican wolves displayed 0.7 to 5.1% of SNP alleles of dog ancestry. An individual from Vancouver Island, British Columbia, had 21% of SNP alleles of dog ancestry and 16 chromosomes showing signs of admixture. This was the only individual among the North American wolves assessed that had an unambiguous signature of recent hybridisation. Among the wolves from North Canada there were two black individuals, and neither had detectable signs of dog ancestry.

European dog breeds showed low level of admixture with North American wolves (0.000-0.005), consistent with the corresponding result for Eurasian wolves. The level of North American wolf admixture detected in ancient non-European dog breeds (0.000-0.043) was lower than the level of Eurasian wolf admixture (0.000-0.306).

**Genetic differentiation between wolves and dogs at X chromosome**

The neighbour-joining tree of male X chromosome haplotypes clustered all wolves together with 67% bootstrap support and wolves from all regions, except China, with 97% support (Figure 6A). Chinese wolves were clustered together with 79% support. Italian wolves and Middle Eastern wolves (from Israel and Oman) formed two distinct subclades within the primary wolf clade with 99% and 87% bootstrap support, respectively. Haplotypes of Indian and Spanish wolves clustered with Eastern European wolves, but they were represented by only two individuals each. All F2/F3 backcrosses identified based on autosomal data (Figure 3) grouped with their respective populations. However, the F1 hybrid (individual #2757) demonstrated X chromosome haplotype clustering with dog haplotypes.

The neighbour-joining tree of female X chromosome haplotypes clustered all wolf haplotypes with 60% bootstrap support (Figure 6B). This tree is based on two haplotypes per female, which were reconstructed using FASTPHASE. Wolves from China, the Middle East (Saudi Arabia and Israel), Italy and Spain formed four distinct clusters with 99%, 58%, 99% and 95% bootstrap support, respectively. Both haplotypes of one female from China did not group with other Chinese haplotypes, but instead one of them grouped with West Asian haplotypes, and the second with Eastern European haplotypes. All F2/F3 backcrosses identified based on autosomal data (Figure 3) grouped with their respective populations. The X-chromosome haplotype trees based on a dataset limited to European wolves and dog breeds of European origin only were in strong agreement with the results described above, for both males and females (Supporting Information Part B, Supporting Figure S6).

LAMP analysis of ancestry blocks on the X chromosome in females revealed no signs of dog ancestry in European and West Asian wolves, with the exception of two individuals from Europe with 12% and 4% of SNP alleles of dog ancestry, respectively (Supporting Table S5). In contrast, only one Chinese wolf lacked a signature of dog ancestry, while the other seven

individuals analysed had 7-25% of SNP alleles of dog ancestry. Analysis of East Asian and Arctic dog breeds revealed 7-31% of SNP alleles of wolf ancestry, while this proportion was 2% in both Basenji and Dingo. European and West Asian breeds had the lowest proportion of 0-1.8% of SNP alleles originating from wolf admixture.

LAMP analysis using X chromosome in males showed that the F1 hybrid #2757 had the entire X chromosome of dog ancestry. In two males from Eastern Europe, the estimated percentage of dog ancestry was 15% and 9%, respectively, which was considerably higher than that observed using autosome data (Supporting Table S5). In contrast, none of the males identified as F2/F3 backcrosses based on autosomal chromosomes had detectable signature of dog ancestry in the X chromosome. No male dogs showed signature of wolf ancestry in the X chromosome (Table 1).

Analysis of population genetic structure at the X chromosome using ADMIXTURE (carried out for females only) distinguished wolves and dogs as distinct groups at K=2. Division into four groups (K=4) was identified as the most likely genetic structure with dogs, Italian wolves, and Arabian wolves forming separate clusters, while wolves from other regions (Eastern Europe, Spain, Iran, India and China) were grouped together. Only one European wolf (from Italy) showed signs of dog admixture at the X chromosome. No admixture was detected in West Asia, but Chinese wolves had assignment probabilities to the dog cluster ranging from 6% to 19%. In European dog breeds, the assignment probabilities to the dog cluster were in the range 81-99%, whereas ancient breeds of Asian origin had probabilities of 51-79%.

Standard errors for the cluster membership estimates from ADMIXTURE were higher for the X chromosome as compared with the autosomes. For all back-cross females, the confidence intervals for the wolf cluster membership at the X chromosome included the value of 1.000 (Supporting Table S3). Therefore, these individuals could have both copies of their

X chromosome originating from wolves, implying that they had male F1 hybrid ancestors originating from female wolf x male dog admixture.

**Heterogeneity of dog ancestry proportions across the chromosomes in Eurasian wolves**

The global CAnD test detected no significant heterogeneity in dog ancestry proportions across all autosomal chromosomes in the dataset of all wolves studied ($P$=0.072). For the datasets including both autosomal and X chromosome data for males and females, respectively, two autosomes had significantly lower proportions of dog ancestry as compared with the mean for all other autosomal chromosomes and the X chromosome (see Supporting Information, Part A, for details). We found no significant difference in dog ancestry proportions in the X chromosome as compared with the mean ancestry in autosomal chromosomes, in any of the datasets.

**Heterozygosity, autozygosity and linkage disequilibrium in admixed individuals**

Heterozygosity at autosomal chromosomes in admixed individuals from Italy ($H_O$=0.23, $H_E$=0.22) was considerably higher than in non-admixed Italian wolves ($H_O$=0.16, $H_E$=0.16), and was within the range of Eastern European wolf populations (Table 2). Wolf populations from south-western Europe and Saudi Arabia had lower heterozygosity levels compared to populations from Eastern Europe, Israel and China (Table 2).

Admixed individuals from Italy had a higher fraction of autozygous segments across all fragment sizes compared to Italian wolves and most other Eurasian wolf populations (Table 2). No autozygous segments were found in the F1 hybrid (individual #2757). Admixed individuals from Italy had lower LD levels ($r^2$ decayed below 0.5 at 387.5Kb) than non-admixed Italian wolves, where $r^2$ did not decay below 0.5 for the entire range of distances

considered (up to 1 Mb). However, these admixed wolves still had higher LD levels than Iberian (257 Kb) and Eastern European wolves (2.5-10 Kb).

## DISCUSSION

### Detection of wolf-dog admixture based on genome-wide SNP data

The application of genome-wide SNP data has substantially improved resolution to detect admixture between canid species, compared to data generated using 10-30 microsatellite loci (vonHoldt et al., 2011). However, even with improved resolution, detecting hybridisation between grey wolves and their domesticated subspecies remains challenging due to their recent shared common ancestry and the difficulty with *a priori* identification of non-admixed individuals that are required as a reference in most methods of admixture analysis.

We carried out ancestry block analyses in Eurasian wolves in comparison with domestic dogs, applying a method implemented in the LAMP software that did not require the use of non-admixed reference populations. Under the assumption of a recent admixture in the past 10 generations (corresponding to about 30 – 40 years; Mech & Seal, 1987; Mech et al., 2016) we were able to detect first generation wolf-dog hybrids, recent backcrosses, and assess the overall level of admixture. The results we obtained when comparing Eurasian wolves and dog breeds of diverse origin were highly consistent with the results for a reduced dataset consisting of European wolves and European dog breeds only. This finding suggests that the composition of wolf and dog datasets does not affect the ability to detect hybrids and recent back-crosses, based on genome-wide SNP data. High consistency of the results from LAMP with those from the PCADMIX software, which required the use of reference populations, shows that admixed individuals can be detected independent of the choice of a particular analytical approach.

On average, the frequency of dog-derived alleles in wolves was three times higher than the frequency of wolf-derived alleles in the pure-breed dogs we studied. If the alleles were incorrectly inferred as dog-derived due to recent common ancestry of wolves and dogs, reference bias in the dominantly dog panel used to design the array (vonHoldt et al. 2010), or imperfect resolution of our method, a similar frequency of inferred wolf-derived alleles should be expected in dogs (see also Supporting Information, Part D). The different proportions of alleles derived from hybridisation observed in the gene pools of dogs and wolves suggests more frequent introgression of dog alleles into the wolf gene pool than in the opposite direction. This is consistent with the expectations, given that we used pure-bred dogs, which are unlikely to interbreed with wolves except as a result of a deliberate human action. The estimated levels of wolf alleles introgression to free-ranging dog populations are likely to be higher.

**Advantages and limitations of the dataset used**

The dataset used in this study consisted of grey wolves sampled from across Eurasia, putative wolf-dog hybrids, and pure-bred dogs. Through the use of pure-bred dogs we ensured that we compare the wolf population (which admixture status was unknown prior to our analysis) with the non-admixed dog population. Pure-bred dogs can interbreed with wolves only via a deliberate human action, and we did not use breeds with a recent history of wolf admixture, such as Czechoslovakian wolf-dogs. In contrast, free-ranging dog populations in Eurasia show signatures of introgression from grey wolves (Kopaliani et al., 2014, Pilot et al., 2015, Fan et al., 2016). The comparison with pure-bred dogs allowed us to control the accuracy of our results, as we could expect limited levels of wolf admixture in pure-bred dogs. This would not be possible if we compared two populations with unknown admixture levels

(Eurasian wolves vs free-ranging dogs), and therefore pure-bred dogs were more appropriate for our purpose.

On the other hand, free-ranging dogs rather than pure-bred dogs are the source of the introgression of dog alleles into wolf gene pool, and Eurasian free-ranging dogs are a genetically distinct population instead of being an admixture of breeds (Pilot et al., 2015). Therefore, the comparison with free-ranging dogs may result in higher levels of estimated dog introgression into wolf populations.

Another important feature of this dataset was that it includes putative hybrids that were deliberately selected for genotyping from a larger dataset. This allowed us to assess the accuracy of hybrid detection in previous studies based on a small number of microsatellite loci, and provided us with a sufficient number of admixed individuals to make conclusions regarding the mechanisms of admixture (e.g. the sex-biased introgression). However, this dataset cannot be used to assess the frequency of hybrids and recent back-crosses in the wolf populations studied. For this purpose, a different sampling design will be required, without enrichment of the dataset for putatively admixed individuals.

Our dataset includes a relatively large sample of Italian wolves as compared with the sample sizes of other wolf populations. This could potentially affect the results of population structure analyses, showing that the Italian population is genetically distinct from other Eurasian populations (Figures 2 and 4). However, genetic distinctiveness of the Italian population was documented in a number of independent studies, and was shown to result from genetic drift during long-term isolation (Lucchini et al., 2004; Pilot et al., 2014; Montana et al., 2017). Our results are consistent with these earlier studies, and therefore we are confident that they are not an artefact of the uneven sample size.

**Wolf-dog hybridisation in Eurasia**

The ancestry block analysis unambiguously defined wolf and dog genetic clusters without any prior information about individuals' origin, which confirms the results based on microsatellite loci analyses showing that Eurasian wolf populations are not hybrid swarms (Randi & Lucchini, 2002; Godinho et al., 2011; Hindrikson et al., 2012; Lorenzini et al., 2014). On the other hand, 62% of genotyped wolves carried small chromosomal blocks that were inferred to originate from dogs (see Supporting Information, Part E). This is consistent with the inference from genome re-sequencing data, which suggests that most Eurasian grey wolves show some level of admixture with dogs (Fan et al., 2016). The presence of dog-derived chromosomal blocks of varying size in the wolf gene pool in different regions studied suggests that introgressive hybridisation has occurred in distinct regions of Eurasia on a variety of timescales and is not solely a recent phenomenon.

This conclusion is also supported by the result of ancestry block analysis for two black-coated individuals from Italy, showing no evidence of recent dog admixture. Both these individuals were heterozygous at the *CBD103* (beta-defensin) gene and carried a dog-derived allele linked to black colouration (Anderson et al., 2009; Caniglia et al., 2013), implying an ancient hybridisation event. This suggests that wolf-dog hybridisation in the Apennine Peninsula has occurred for many generations (in concordance with Randi et al. 2014), and that black wolves may be considered "pure" wolves with the exception of carrying the dog-derived *CBD103* allele (although particular black individuals can be hybrids or recent back-crosses). Our finding is consistent with the results of a recent study focused on wolf-dog hybridisation in Italy, which also detected a number of black-coated wolves that showed no detectable signs of dog ancestry (Galaverni et al., 2017). The assumption that all black wolves derive from recent hybridisation, providing the rationale for eliminating them (Salvatori, 2015), is therefore incorrect. In fact, Italian canids that we genetically identified as

recent back-crosses were not black-coated, hence, removing black wolves may not decrease the admixed ancestry of the population. This result shows that the elimination of individuals with atypical phenotypes is not always an appropriate management strategy for admixed populations.

Genetic introgression from a domesticated population into the wild ancestor is generally considered to be maladaptive as it compromises the genetic integrity of the wild species (Allendorf et al., 2001; Mallet, 2005), justifying management decisions to eliminate admixed individuals. However, in some cases, introgression of domestic gene variants may enhance adaptation. Anderson et al. (2009) showed that the mutation in *CBD103* gene linked to melanism exhibited a molecular signature of positive selection in North American grey wolves. Further studies showed that melanistic individuals which are heterozygous for the dog-derived *CBD103* variant have a selective advantage over grey individuals in forested habitats (Coulson et al., 2011; Stahler et al., 2013; Hedrick et al., 2014). This example shows that hybridisation may provide wolf populations a way of acquiring new adaptations to a rapidly changing environment. Elimination of individuals possessing a single dog-derived phenotypic trait may prevent such adaptations to be established in wolf populations. Therefore, management plans involving the lethal control of hybrids should consider both maladaptive and adaptive effects of admixture (see discussion in Wayne & Shaffer, 2016).

**Continuity of dog ancestry proportions in Eurasian wolf populations and detection accuracy of back-crosses**

If hybridisation has occurred regularly throughout generations and has been followed by back-crossing and gene introgression, we would expect that individuals with different proportions of dog ancestry, ranging between 0 and 0.25, would be present in the wolf population. This was indeed the case, as shown in Supporting Figure S4. Using similar logic,

if hybridisation has been infrequent enough for the wolf and dog populations to retain their genetic distinctiveness, few individuals should be expected to have a share of dog ancestry ranging from 0.25 (corresponding to F2 back-cross, i.e. offspring of a wolf and a F1 hybrid) and 0.5 (F1 hybrid), which can only be achieved from mating between recently admixed individuals. In our dataset, we did not observe any individuals having a proportion of dog ancestry within this range.

Although F1 hybrids could be unambiguously identified based on ancestry block analysis, it was impossible to distinguish between F2 and F3 back-crosses due to a lack of clear discontinuity between these two categories. There were also eight individuals in our dataset that could have been F4 back-crosses as they had > 5% of estimated dog ancestry (the expected value for this generation of back-crosses is 6.25%). However, some pure-breed dogs displayed similar levels of admixture, which likely reflects more distant hybridisation events. Therefore, the precision of this analysis was insufficient to unambiguously detect back-crossing at more distant levels.

The precision of back-crosses detection was improved, however, compared with microsatellite loci analysis (e.g. Randi et al. 2014). Eight individuals from Eastern Europe and three from Italy which were previously identified as admixed based on microsatellite analysis, did not present as genetic outliers from their wolf populations based on the genome-wide SNP data. This suggests that identification of admixed individuals based on a small number of microsatellite loci may be inaccurate beyond F1-F2 hybrids.

Genome-wide SNP genotyping is still too expensive to be used routinely for management decisions, and typically requires high-quality DNA extracts, precluding the use of non-invasive samples. We therefore suggest that both legal regulations and practical decisions regarding the management of admixed individuals clearly distinguish between F1 hybrids (which can be identified unambiguously based on a small number of genetic markers)

and back-crosses into wolf populations, which may be difficult to distinguish from non-admixed wolves without an extensive genetic analysis.

**Geographic patterns of admixture in Eurasian wolf populations**

Individuals with recent admixed ancestry were detected in each of the European populations studied. In contrast, we detected no hybrids or recent back-crosses in Iran, India and China, although the sample sizes from these countries were small. Contrasting patterns were found however in Chinese wolves in autosomes vs the X chromosome (see below). Wolves from the Arabian Peninsula showed signatures of dog admixture in eight out of 14 individuals, with two individuals identified as F2/F3 back-crosses. Given that no known admixed individuals from this region were included, this suggests that hybridisation has been particularly intense in this region. This finding is consistent with the inference of intense, bi-directional gene flow between Israeli wolves and dogs (Freedman et al., 2014), and the inference of gene flow from wolves to Saudi Arabian free-ranging dogs (Pilot et al., 2015).

Taken together, these results show that wolf-dog hybridisation is geographically widespread in Eurasia, but its frequency may vary considerably between regions. Earlier genetic studies on European wolves based on microsatellite loci, estimated the frequencies of admixed individuals at 5.6% in the Iberian Peninsula (Pacheco et al., 2017), 5% in Italy (Verardi et al., 2006) and 9.8% in Bulgaria (Moura et al. 2014). These varying estimates may suggest differences in hybridisation rate between regions, but could also result from differences in methodological approaches between the studies. A comparative assessment of hybridisation levels would require the use of the same genetic markers and analytical methods for different geographic regions, an even sample coverage and an unbiased sampling process, without preferential sampling of putative hybrids. The knowledge of large-scale geographic patterns of hybridisation may help understand whether different methods of wolf

management (regulated hunting, unregulated hunting, full protection) affect the frequency of hybridisation.

**Admixture patterns inferred from the X chromosome data**

The X-chromosome haplotypes of all back-crosses identified in this study grouped within the wolf cluster. This pattern suggests sex-biased introgression of dog alleles into wolf populations, with male dogs having a higher contribution than females. In Chinese wolves, all but one female had positive assignment probabilities (up to 25%) to the dog cluster at the X chromosome, but no admixture was detected in these individuals based on autosomal chromosomes. This result suggests an introgression of dog X chromosome haplotypes following an ancient hybridisation event, and possibly selection acting upon genes on the X chromosome. Representatives of East Asian dog breeds (Chow Chow and Akita) also had positive assignment probabilities to the wolf cluster, suggesting that the hybridisation resulted in a bi-directional introgression of X chromosome haplotypes between dogs and wolves.

The X and Y chromosome patterns imply that the only F1 hybrid identified in our sample set, Italian male #2757, was the offspring of a female dog and a male wolf. This individual was identified *a priori* as admixed based on its atypical phenotype (as reported in the ISPRA database "putative hybrid or dog-like"; the details of the phenotype or photos are not available) and its STRUCTURE-based assignment probability to the Italian wolf cluster was 0.44 based on 39 autosomal microsatellite loci (Randi et al. 2014). We found that the X chromosome haplotype of this male clustered with dogs, and Randi et al. (2014) found that he carried Y chromosome haplotype YH17 (inferred from four microsatellite loci data), which is commonly found in Italian wolves. This result indicates that this individual was the offspring of a male wolf. However, most previously described cases of natural wolf-dog hybridisation

involved female wolves that mated with male dogs (Andersone et al., 2002; Vila et al., 2003; Iacolina et al., 2010; Godinho et al., 2011). A review of wolf-dog hybridisation patterns worldwide concluded that mating between male wolves and female dogs is less frequent and/or it is rarely followed by back-crossing of the resulting hybrids into the wolf population (Leonard et al., 2014). Nevertheless, examples of hybrids having male wolf x female dog ancestry are known from earlier studies (Hindrikson et al., 2012), and the hybrid we have identified here unequivocally represents such case.

We also founds signs of female-mediated introgression of dog alleles in two males from Eastern Europe, which displayed relatively high estimated percentage of dog ancestry (15% and 9%, respectively) in the X chromosome, but lower percentage in the autosomes. On the other hand, all but one individuals identified as F2/F3 back-crosses based on the analysis of autosomes had 100% of X-chromosome SNP alleles matched to those defining wolf ancestry. This result suggests that mating of female wolves with male dogs may be more favourable for introgression of dog alleles into wolf populations. However, the test of heterogeneity in dog ancestry proportions between the X chromosome and autosomal chromosomes for the entire dataset of Eurasian wolves was non-significant, implying either the lack of sex bias in the introgression, or insufficient power. Futher studies are needed to clarify this, given important implications for the management of admixing populations. If the introgression of dog alleles is male-biased, it could be limited by sterilisation of free-ranging male dogs, but this won't be sufficient if the introgression is not sex-biased.

**Wolf-dog hybridisation in North America**

In contrast to Eurasian wolves, most North American grey wolves showed no signal of admixture with dogs. In Mexican wolves, we found SNP alleles matching dog ancestry, but their frequency (1-5%) was too small to make conclusions regarding the admixture status

of this population. This finding does not imply conservation concerns regarding the genetic integrity of the Mexican wolf population, but indicates the need for more extensive research into possible past hybridisation of Mexican wolves with other canids.

Among the North American wolves studied, only one individual, from Vancouver Island, British Columbia, was identified as a recent back-cross (probably F3). The Vancouver Island population has been shown previously to have experienced hybridisation with dogs, which likely occurred at early stages of recolonisation of the island in the 1970-1980s (Munoz-Fuentes et al., 2010). This result suggests that strong demographic fluctuations and range contractions/expansions promote cross-breeding with dogs. Eurasian wolf populations have experienced strong bottlenecks and range fluctuations, some of which are well documented either based on direct demographic inference (Boitani, 2003), or genetic analyses (e.g. Pilot et al., 2014; Fan et al., 2016; Montana et al., 2017). Moreover, Eurasian wolf populations have been sympatric with dogs for a longer period, given that the dog domestication occurred in Eurasia (Freedman & Wayne, 2017). Both factors may contribute to the higher frequency of alleles originating from dogs in gene pools of in Eurasian versus North American wolves.

The sympatric occurrence of coyotes in large parts of the North American wolf range may be of importance as well, as wolves may show preference towards mating with coyotes rather than dogs (vonHoldt et al., 2011, 2016). Although North American wolves show signatures of ancient hybridisation with dogs (Anderson et al., 2009), studies documenting recently admixed individuals are rare (Munoz-Fuentes et al., 2010). In consistence with our results, whole-genome sequence data showed signatures of recent admixture with dogs in Eurasian wolves, but not in North American wolves (Fan et al. 2016). Understanding the reason underlying this difference in hybridisation patterns may help develop effective strategies to manage admixing wolf populations.

Our analysis also found low levels of admixture from North American wolves in domestic dogs breeds of both European and non-European origin, with the highest level estimated at 4%. By comparison, the admixture analysis between Eurasian wolves and domestic dogs indicated considerable input (up to 11%) of wolf-derived variants into ancient breeds, particularly breeds of in East Asian and Arctic origin. This implies that the past hybridisation event(s) resulting in wolf admixture in ancient breeds occurred in Eurasia rather than North America (see Supporting Information, Part D, for further discussion of wolf admixture in dogs).

The large differences in the frequency of dog-derived alleles in the Eurasian versus North American wolf populations provide evidence that dog admixture inferred in LAMP does not represent background noise produced by the method. With the exception of Mexican wolves and one individual from Vancouver Island discussed above, the maximum share of dog ancestry detected in North American wolves was 0.006. This value may be considered as the maximum rate of erroneous assignment of dog origin to small chromosomal segments using LAMP. In the Eurasian populations studied, 54% of individuals had an estimated proportion of dog ancestry exceeding this value. If small chromosomal segments attributable to dogs in Eurasian wolves were false positives produced by LAMP, they should have been detected in North American wolves with a similar frequency as in Eurasian wolves, which was not the case.

**Heterozygosity, autozygosity and linkage disequilibrium in admixed individuals**

Italian wolves have low heterozygosity and high LD as a result of long-term isolation and a bottleneck (Pilot et al., 2014; Montana et al. 2017). Italian canids identified as wolf-dog hybrids or back-crosses had considerably higher heterozygosity and lower LD than pure Italian wolves. Although hybridisation is generally expected to increase LD, in this case it

had an opposite effect, due to Italian wolves displaying particularly long stretches of LD (Pilot et al., 2014).

A study of the Scandinavian wolf population showed that the most heterozygous individuals establish themselves as breeders (Bensch et al., 2006). If this is a general rule, back-crossed individuals may have a selective advantage over pure wolves in populations with low heterozygosity levels, such as the Italian population. Interestingly, back-crossed Italian individuals carried a higher fraction of autozygous segments across all fragment sizes than pure wolves, suggesting that hybridisation was followed by mating with related individuals in subsequent generations. This conclusion is supported by the work of Caniglia et al. (2013) who chronicled a recent hybridisation event followed by breeding between close relatives in a single Italian pack of wolves.

**Conclusions and management implications**

We detected the presence of small blocks of dog ancestry in the genomes of 62% wolves sampled from all Eurasian populations analysed, suggesting that hybridisation has occurred in different parts of Eurasia, throughout multiple generations, and is not solely a recent phenomenon. Nevertheless, the wolf populations have maintained a distinct genetic profile from dogs, suggesting that hybridisation and backcrossing has occured at a low frequency.

We found that two melanistic wolves who carried a dog-derived allele at a beta-defensin locus, displayed no signs of recent admixed ancestry. In contrast, some individuals identified *a priori* as "pure" wolves were shown to be F2 or F3 back-crosses. This result implies that phenotype alone cannot be reliably used to distinguish between back-crosses and non-admixed individuals. Our data also suggests that Eurasian wolf populations represent a continuum of genotypes from "pure" wolves to F2 back-crosses. This makes the definition of

genetically "pure" wolves ambiguous, and raises questions about appropriate management of back-crossed individuals, as they may be too difficult to identify and too numerous to be removed from wolf populations (see Wayne & Shaffer, 2016).

Back-crossed individuals are typically integrated into wolf packs, and disruption of pack structure due to culling may enhance hybridisation (Moura et al. 2014). Therefore, even if admixed individuals could be unambiguously identified, their removal may be ineffective and could eventually generate more hybrids. The efficient management of admixed populations should be focused, instead, on reducing the factors which cause hybridisation, such as small population size, the presence of free-ranging dogs and unregulated hunting (Moura et al., 2014). Also, increasing the proportion of natural wolf habitats and their natural prey may enhance retention of wolf genomic elements by natural selection (Wayne and Shaffer, 2016). We also recommend that any documents regulating legal status of admixed canids should distinguish between F1 hybrids and back-crosses into wild populations. Although this study was specifically focused on grey wolves and domestic dogs, our conclusions are applicable to any case where hybridisation with a domesticated species may affect the genetic integrity of a closely related wild species.

Protection and Research (ER and CG), University of Bologna, Italy (CG), Institute of Nature Conservation, Polish Academy of Sciences (MKK), the US National Science Foundation (RKW), and the Intramural Program of the National Human Genome Research Institute (EAO).

**Data archiving**

This project was based on the analysis of data published in vonHoldt et al. (2010), which can be accessed at http://genome-mirror.bscb.cornell.edu/cgi-bin/hgGateway (see 'SNPs' track under the Variations and Repeats heading).

**References**

Aghbolaghi M.A., Rezaei H.R., Scandura M., Kaboli M. (2014) Low gene flow between Iranian grey wolves (*Canis lupus*) and dogs documented using uniparental genetic markers. Zoology in the Middle East, 60, 95-106.

Alexander, D.H., Novembre J., Lange K. (2009) Fast model-based estimation of ancestry in unrelated individuals. Genome Research, 19, 1655–1664.

Allendorf, F. W., Leary, R. F., Spruell P., Wenburg, J. K. (2001). The problems with hybrids: setting conservation guidelines. Trends in Ecology and Evolution, 16, 613–622.

Anderson, T., vonHoldt, B., Candille, S., Musiani, M., Greco, C., Stahler, D., …, Barsh, G. (2009). Molecular and evolutionary history of melanism in North American gray wolves. Science, 323, 1339-1343.

Andersone, Z., Lucchini, V., Randi, E., Ozolins, J. (2002). Relationships between wolves and dogs in Latvia as documented using mitochondrial and microsatellite DNA markers. Mammalian Biology, 67, 79–90.

Bensch, S., Andren, H., Hansson, B., Pedersen, H. C., Sand, H., Sejberg, D., … Liberg, O.

(2006). Selection for Heterozygosity Gives Hope to a Wild Population of Inbred Wolves. PLoS ONE, 1, e72.

Boitani, L. (2003). Wolf conservation and recovery. In: Mech LD, Boitani L (eds) Wolves. Behavior, ecology and conservation. University of Chicago Press, Chicago, pp 317–340.

Boyko, A.R., Quignon, P., Li, L., Schoenebeck, J., Degenhardt, J. D., Lohmueller, K. E., …, Ostrander E. A. (2010). A simple genetic architecture underlies quantitative traits in dogs. PLoS Biology, 8, e1000451.

Brisbin A., Bryc, K., Byrnes J., Zakharia F., Omberg L., Degenhardt J., … Bustamante, C.D. (2012). PCAdmix: Principal Components-Based Assignment of Ancestry along Each Chromosome in Individuals with Admixed Ancestry from Two or More Populations. Human Biology, 84, 343–364.

Caniglia, R., Fabbri, E., Greco, C., Galaverni, M., Manghi, L., Boitani, L., … Randi, E. (2013). Black coats in an admixed wolf x dog pack is melanism an indicator of hybridization in wolves? European Journal of Wildlife Research, 59, 543-555.

Ciucci, P., Lucchini, V., Boitani, L., Randi, E. (2003). Dewclaws in wolves as evidence of admixed ancestry with dogs. Canadian Journal of Zoology, 81, 2077-2081.

Coulson, T., MacNulty, D. R., Stahler, D. R., vonHoldt, B., Wayne, R. K., Smith, D. W. (2011). Modeling effects of environmental change on wolf population dynamics, trait evolution and life history. Science, 334, 1275-1278.

Earl, D. A., vonHoldt, B. M. (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.

Fan, Z., Silva, P., Gronau, I., Wang, S., Armero, A. S., Schweizer, R. M., …Wayne, R. K. (2016). Worldwide patterns of genomic variation and admixture in gray wolves. Genome Research, 26, 163-173.

Feulner, P. G. D., Gratten, J., Kijas, J. W., Visscher, P. M., Pemberton, J. M., Slate J. (2013) Introgression and the fate of domesticated genes in awild mammal population. Molecular Ecology 22, 4210-4221.

Freedman, A. H., Gronau, I., Schweizer, R. M., Ortega-Del Vecchyo, D., Han, E., Silva, P. M.,…Novembre J. (2014). Genome sequencing highlights the dynamic early history of dogs. PLoS Genetics, 10, e1004016.

Freedman, A. H., Wayne, R. K. (2017). Deciphering the origin of dogs: from fossils to genomes. Annual Review of Animal Biosciences, 5, 281-307.

Galaverni, M., Caniglia, R., Pagani, L., Fabbri, E., Boattini, A., Randi E. (2017). Disentangling Timing of Admixture, Patterns of Introgression, and Phenotypic Indicators in a Hybridizing Wolf Population. Molecular Biology and Evolution. Advance Online Publication, doi:10.1093/molbev/msx169.

Godinho, R., Llaneza, L., Blanco, J.C., Lopes, S., Álvares, F., García, E.J., ... Ferrand N. (2011). Genetic evidence for multiple events of hybridization between wolves and domestic dogs in the Iberian Peninsula. Molecular Ecology, 20, 5154–5166.

Gompper, M. E. (2014). The dog–human–wildlife interface: assessing the scope of the problem. In Free-ranging dogs and wildlife conservation (ed. M.E. Gompper), pp. 9–54. Oxford, UK: Oxford University Press.

Grossen, C., Keller, L., Biebach, I., The International Goat Genome Consortium, Croll, D. (2014) Introgression from Domestic Goat Generated Variation at the Major Histocompatibility Complex of Alpine Ibex. PLoS Genet 10(6), e1004438.

Hedrick, P. W., Stahler, D. R., Dekker D. (2014). Heterozygote Advantage in a Finite
Population: Black Color in Wolves. Journal of Heredity, 105, 457-465.

Hindrikson, M., Mannil, P., Ozolins, J., Krzywinski, A., Saarma, U. (2012). Bucking the
Trend in Wolf-Dog Hybridization: First Evidence from Europe of Hybridization
between Female Dogs and Male Wolves. PLoS ONE, 7(10), e46465.

Hindrikson, M., Remm, J., Pilot, M., Godinho, R., Stronen, A. V., Baltrūnaité, L., … Saarma
U. (in press). Wolf population genetics in Europe: a systematic review, meta-analysis
and suggestions for conservation and management. Biological Reviews, doi:
10.1111/brv.12298.

Harrison, R. G., Larson, E. L. (2014). Hybridization, Introgression, and the Nature of Species
Boundaries. Journal of Heredity: 105(Special Issue): 795–809.

Iacolina, L., Scandura, M., Gazzola, A., Cappai, N., Capitani, C., Mattioli, L., … Apollonio,
M. (2010). Y-chromosome microsatellite variation in Italian wolves: A contribution to
the study of wolf-dog hybridization patterns. Mammalian Biology, 75, 341–347.

Khosravi, R., Rezaej, H. R., Kaboli, M. (2013). Detecting Hybridization Between Iranian
Wild Wolf (*Canis lupus pallipes*) and Free-ranging Domestic Dog (Canis familiaris) by
Analysis of Microsatellite Markers. Zoological Science, 30, 27-34.

Khosravi, R., Aghbolaghi, M. A., Rezaej, H. R., Norani, E., Kaboli, M. (2015). Is black coat
colour in wolves of Iran an evidence of admixed ancestry with dogs? Journal of
Applied Genetics, 56, 97-105.

Kopaliani, N., Shakarashvili, M., Gurielidze, Z., Qurkhuli, T., Tarkhnishvili, D. (2014). Gene
Flow between Wolf and Shepherd Dog Populations in Georgia (Caucasus). Journal of
Heredity, 105, 345-353.

Leonard, J. A., Echegaray, J., Randi, E., Vila, C. (2014). Impact of hybridization with domestic dogs on the conservation of wild canids. In: Free-ranging dogs and wildlife conservation (ed. ME Gompper), pp. 170-184. Oxford, UK: Oxford University Press.

Lecis, R., Pierpaoli, M., Birò, Z.S., Szemethy, L., Ragni, B., Vercillo, F., Randi, E. (2006). Bayesian analyses of admixture in wild and domestic cats (*Felis silvestris*) using linked microsatellite loci. Molecular Ecology, 15, 119-131.

Lescureux, N., Linnell, J.D.C. (2014). Warring brothers: The complex interactions between wolves (*Canis lupus*) and dogs (*Canis familiaris*) in a conservation context. Biological Conservation, 171, 232–245.

Lorenzini, R., Fanelli, R., Grifoni, G., Scholl, F., Fico, R. (2014). Wolf-dog crossbreeding: "Smelling" a hybrid may not be easy. Mammalian Biology, 79, 149-156.

Lucchini, V., Galov, A., Randi, E. (2004). Evidence of genetic distinction and long-term population decline in wolves (*Canis lupus*) in the Italian Apennines. Molecular Ecology, 13, 523–536.

Mallet, J. (2005). Hybridization as an invasion of the genome. Trends in Ecology & Evolution, 20, 229–237.

McHugh, C., Brown, L., Thornton, T. (2016). Detecting heterogeneity in population structure across the genome in admixed populations. Genetics, 204, 43–56.

Mech, L. D., Barber-Meyer, S. M., Erb, J. (2016). Wolf (*Canis lupus*) Generation Time and Proportion of Current Breeding Females by Age. PLoS ONE 11(6), e0156682.

Mech, L.D., Seal, U. S. (1987). Premature reproductive activity in wild wolves. Journal of Mammalogy, 68, 871–873.

Miao, B., Wang, Z., Li, Y. (2017). Genomic Analysis Reveals Hypoxia Adaptation in the Tibetan Mastiff by Introgression of the Gray Wolf from the Tibetan Plateau. Molecular Biology and Evolution, 34, 734–743.

Montana, L., Caniglia, R., Galaverni, M., Fabbri, E., Ahmed A., Černá Bolfíková B, … Randi, E. (2017). Combining phylogenetic and demographic inferences to assess the origin of the genetic diversity in an isolated wolf population. PLoS ONE 12(5), e0176560.

Moura, A. E., Tsingarska, E., Dąbrowski, M. J., Czarnomska, S. D., Jędrzejewska, B., Pilot M. (2014). Unregulated hunting and genetic recovery from a severe population decline: the cautionary case of Bulgarian wolves. Conservation Genetics, 15, 407–415.

Munoz-Fuentes, V., Darimont, C.T., Paquet, P.C., Leonard, J.A. (2010). The genetic legacy of extirpation and re-colonization in Vancouver Island wolves. Conservation Genetics, 11, 547–556.

Munoz-Fuentes, V., Vila, C., Green, A. J., Negro, J. J., Sorenson, M. (2007). Hybridization between white-headed ducks and introduced ruddy ducks in Spain. Molecular Ecology, 16, 629–638.

Pacheco, C., López-Bao, J. V., García, E. J., Lema, F. J., Llaneza, L., Palacios, V., Godinho, R. (2017). Spatial assessment of wolf-dog hybridization in a single breeding period. Scientific Reports, 7, 42475.

Parker, H.G., Kim, L.V., Sutter, N. B., Carlson, S., Lorentzen, T. D., … Ostrander E. A. (2004). Genetic structure of the purebred domestic dog. Science, 304, 1160–1164.

Pilot, M., Greco, C., vonHoldt, B.M., Jędrzejewska, B., Randi, E., Jędrzejewski, W., … Wayne, R. K. (2014). Genome-wide signatures of population bottlenecks and diversifying selection in European wolves. Heredity, 112, 428–442.

Pilot, M., Malewski, T., Moura, A. E., Grzybowski, T., Oleński, K., Ruść, A., … Bogdanowicz, W. (2015). On the origin of mongrels: evolutionary history of free-breeding dogs in Eurasia. Proceedings of the Royal Society B: Biological Sciences, 282, 20152189.

Price, A. L., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N. A., Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. Nature Genetics, 38, 904 – 909.

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

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D ., … Sham, P. C. (2007). PLINK: a tool set for whole-genome association and population-based linkage analyses. American Journal of Human Genetics, 81, 559-575.

Randi, E. (2008). Detecting hybridization between wild species and their domesticated relatives. Molecular Ecology, 17, 285-293.

Randi, E., Lucchini, V. (2002). Detecting rare introgression of domestic dog genes into wild wolf (*Canis lupus*) populations by Bayesian admixture analyses of microsatellite variation. Conservation Genetics, 3, 31-45.

Randi, E., Hulva, P., Fabbri, E., Galaverni, M., Galov, A., Kusak, J., …Caniglia, R. (2014). Multilocus Detection of Wolf x Dog Hybridization in Italy, and Guidelines for Marker Selection. PLoS ONE, 9(1), e86409.

Salvatori, V. (2015). Reducing the loss of wolf gene pool in Province of Grosseto, Italy: action plan for the after LIFE IBRIWOLF project period. Project LIFE10NAT/IT/265 "Pilot actions for the reduction of wolf gene pool loss in central Italy" – IBRIWOLF.

Sankararaman, S., Sridhar, S., Kimmel, G., Halperin, E. (2008). Estimating local ancestry in admixed populations. American Journal of Human Genetics, 82, 290–303.

Scheet, P. and Stephens, M. (2006). A fast and flexible statistical model for large-scale population genotype data: applications to inferring missing genotypes and haplotypic phase. American Journal of Human Genetics, 78, 629–644.

Schweizer, R.M., vonHoldt, B.M., Harrigan, R., Knowles, J. C., Musiani, M., Coltman, D., Novembre, J., Wayne, R. K. (2016). Genetic subdivision and candidate genes under selection in North American gray wolves. Molecular Ecology, 25, 257–279.

Stahler, D. R., MacNulty, D. R., Wayne, R. K., vonHoldt, B., Smith, D. W. (2013). The adaptive value of morphological, behavioural and life-history traits in reproductive female wolves. Journal of Animal Ecology, 82, 222-234.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S. (2011). MEGA5: Molecular Evolutionary Genetics Analysis using Likelihood, Distance, and Parsimony methods. Molecular Biology and Evolution, 28, 2731-2739.

Trouwborst, A. (2014). Exploring the legal status of wolf-dog hybrids and other dubious animals: international and EU law and the wildlife conservation problem of hybridization with domestic and alien species. Review of European Community and International Environmental Law, 23,111–124.

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.

Vilà, C., Walker, C., Sundqvist, A. K., Flagstad, Ø., Andersone, Z., Casulli, A., … Ellegren, H. (2003). Combined use of maternal, paternal and bi-parental genetic markers for the identification of wolf-dog hybrids. Heredity, 90, 17–24.

vonHoldt, B. M., Fan, Z., Ortega-Del Vecchyo, D., Wayne R. K. (2017). EPAS1 variants in high altitude Tibetan wolves were selectively introgressed into highland dogs. PeerJ, 5:e3522.

vonHoldt, B., Pollinger, J. P., Lohmueller, K. E., Han, E., Parker, H. G., Quignon, P., … Wayne, R. K. (2010). Genome-wide SNP and haplotype analyses reveal a rich history underlying dog domestication. Nature, 464, 898-903.

vonHoldt, B. M., Pollinger, J. P., Earl, D. A., Knowles, J. C., Boyko, A. R., Parker, H., …
Wayne, R. K. (2011). A genome-wide perspective on the evolutionary history of
enigmatic wolf-like canids. Genome Research, 21, 1294-305.

vonHoldt, B. M., Kays, R., Pollinger, J. P., Wayne, R. K. (2016). Admixture mapping
identifies introgressed genomic regions in North American canids. Molecular Ecology,
25, 2443–2453.

vonHoldt, B. M., Pollinger, J. P., Earl, D. A., Parker, H., Ostrander E. A., Wayne, R. K.
(2013). Identification of recent hybridization between gray wolves and domesticated
dogs by SNP genotyping. Mammalian Genome, 24, 80–88.

Wayne, R. K., Shaffer, H. B. (2016). Hybridization and Endangered Species Protection in the
Molecular Era. Molecular Ecology, 25, 2680–2689.

Table 1. Summary of the results of the ancestry blocks analyses in LAMP for Eurasian wolf populations, North American wolves populations and pure-bred dogs, in comparison with the results of tests based on the analysis of population genetic structure (using STRUCTURE and ADMIXTURE, assuming K=2).

| | LAMP results | | | | | |
|---|---|---|---|---|---|---|
| Canid group | autosomes | No. admixed autosomes | chr X females | chr X males | STRUCTURE | ADMIXTURE |
| Chinese wolves (10) | 0.981-1.000 | 0-3 | 0.753-1.000 | 1.000 | 0.853-0.900 | 0.861-0.908 |
| Indian and Iranian wolves (4) | 0.959-0.996 | 0 | 1.000 | 1.000 | 0.913-0.930 | 0.922-0.938 |
| Arabian Peninsula wolves (6) | 0.999-1.000 | 0-6 | 1.000 | 1.000 | 0.866-0.886 | 0.875-0.893 |
| Arabian Peninsula F2/F3 back-crosses (2) | 0.861-0.914 | 12-16 | 1.000 | 1.000 | 0.790-0.823 | 0.797-0.831 |
| Arabian Peninsula F4+ back-crosses (uncertain) | 0.934-0.968 | 3-10 | 1.000 | 1.000 | 0.833-0.872 | 0.841-0.879 |
| European wolves (85) | 0.949-1.000 | 0-8 | 0.964-1.000 | 1.000 | 0.922-1.000 | 0.930-1.000 |
| European F2/F3 back-crosses (7) | 0.760-0.904 | 12-23 | 0.877-1.000 | 1.000 | 0.825-0.893 | 0.832-0.896 |
| European F4+ back-crosses (uncertain) (4) | 0.929-0.946 | 8-11 | - | 0.849-1.000 | 0.901-0.977 | 0.907-0.977 |
| North American wolves (42) | 0.994-1.000 | 0-1 | - | - | - | - |
| Mexican wolves (5) | 0.949-0.993 | 1-7 | - | - | - | - |
| European dog breeds (105) | 0.000-0.003 | 37-38 | 0.000-0.018 | 0.000 | 0.001-0.083 | 0.000-0.091 |
| non-European dog breeds (20) | 0.002-0.112 | 20-38 | 0.000-0.306 | 0.000 | 0.116-0.348 | 0.121-0.354 |

Table note: Number of samples is provided in brackets after the name of each population.

LAMP results are presented as the percentage of SNP alleles of wolf ancestry in autosomal chromosomes (at average) and in X chromosome (assessed only for individuals with sex known *a priori*, and separately for males and females). We also report the number of admixed autosomal chromosomes, i.e. having less than 90% of SNP alleles of wolf ancestry. The results of STRUCTURE and ADMIXTURE analyses are presented as the assignment probability of a given individual to the wolf cluster. "North American wolves" denote all North American wolf populations except Mexican wolves, which are presented separately. North American wolves were analysed in a separate LAMP run rather than with Eurasian wolves.

Table 2. Heterozygosity and autozygosity in admixed individuals from Italy as compared with local populations of non-admixed wolves from different parts of Eurasia.

| Local populations | $H_O$ | $H_E$ | No. of homozygous segments | Average length of homozygous segments (Kb) |
|---|---|---|---|---|
| Italian hybrids/ backcrosses | 0.234 | 0.220 | 6.8 | 3603 |
| Italian wolves | 0.161 | 0.155 | 1.6 | 2449 |
| Iberian wolves | 0.173 | 0.169 | 1.5 | 1902 |
| East European wolves | 0.214-0.235 | 0.219-0.263 | 2.2-6.5 | 1771-5142 |
| Saudi Arabian wolves | 0.179 | 0.156 | 7.4 | 6445 |
| Israeli wolves | 0.215 | 0.222 | 6.1 | 6695 |
| Chinese wolves | 0.221 | 0.235 | 2.8 | 4154 |

Table note: $H_O$ – observed heterozygosity, $H_E$ – expected heterozygosity. Autozygosity is measured as average number of homozygous segments per individual, and their average length. East European wolves are represented by several local populations and therefore the range of values is provided.

**Figure legends**

Figure 1. Geographic distribution of the Eurasian wolf samples analysed (red dots), and the

F1 hybrid and F2/F3 back-crosses (black stars) identified in the study. Sample

locations in Europe, including Russia, are precise; sample locations in Asia are

approximate. The area highlighted in pink represents the wolf distribution range in

Eurasia.

Figure 2. Genetic differentiation between regional populations of wolves and domestic dogs

inferred using the program ADMIXTURE, assuming two, three and four genetic

clusters (K). Plots for a broader range of K values are shown in Supporting Figure

S1. Coyote – *Canis latrans*; Dog – *Canis lupus familiaris*: **EU**ropean breeds, **NON**

**EU**ropean breeds; Wolf – *Canis lupus*: from left to right **SP**ain, **IT**aly, **E**astern

**E**urope (Bulgaria, Croatia, Greece, Turkey – European part, Slovakia, Poland,

Lithuania, Belarus, Ukraine, Russia), **AS**ia (Israel, Arabia, Oman, Iran, India, China)

Figure 3. Dog ancestry proportions in the F1 hybrid, F2/F3 back-crosses and putative further

generation backcrosses, estimated from the ancestry blocks analyses in LAMP and

PCADMIX, and the analysis of population genetic structure in STRUCTURE and

ADMIXTURE. Individuals' admixture status was inferred based on the results from

LAMP, STRUCTURE and ADMIXTURE, while the PCADMIX was carried out as a

follow-up analysis. The LAMP and PCADMIX results are presented as the mean

percentage of SNP alleles of dog ancestry in autosomal chromosomes. The

STRUCTURE and ADMIXTURE results are presented as the assignment probability of

an individual to the dog cluster (assuming K=2). The ancestry proportions were

calculated for autosomal chromosome data.

Figure 4. Principal component analysis (PCA) illustrating the extent of genetic diversification

between Eurasian wolf populations and domestic dogs, and showing the position of

the inferred wolf-dog F1 hybrid and recent back-crosses relative to wolf and dog populations. Individuals labelled as "possibly admixed" are individuals with uncertain admixture status reported in Figure 3 and Supporting Table S2. The coyotes are included as outgroup. The PCA plot constructed without the coyotes is shown in Supporting Figure S2.

Figure 5. Results of ancestry block analysis in LAMP for two chromosomes, shown as an example. Dog ancestry is marked in red and wolf ancestry in yellow. Each row represents one individual, with dogs followed by wolves and admixed canids. A part of the graph is enlarged to show individual hybrids and backcrosses. The ancestry plots for all 38 chromosomes are available in the Supporting Figure S3.

Figure 6. Evolutionary relationships of X chromosome haplotypes in (A) females and (B) males (right) inferred using the neighbour-joining method. The distances were calculated using the p-distance method. Bootstrap support is shown if higher than 50%.
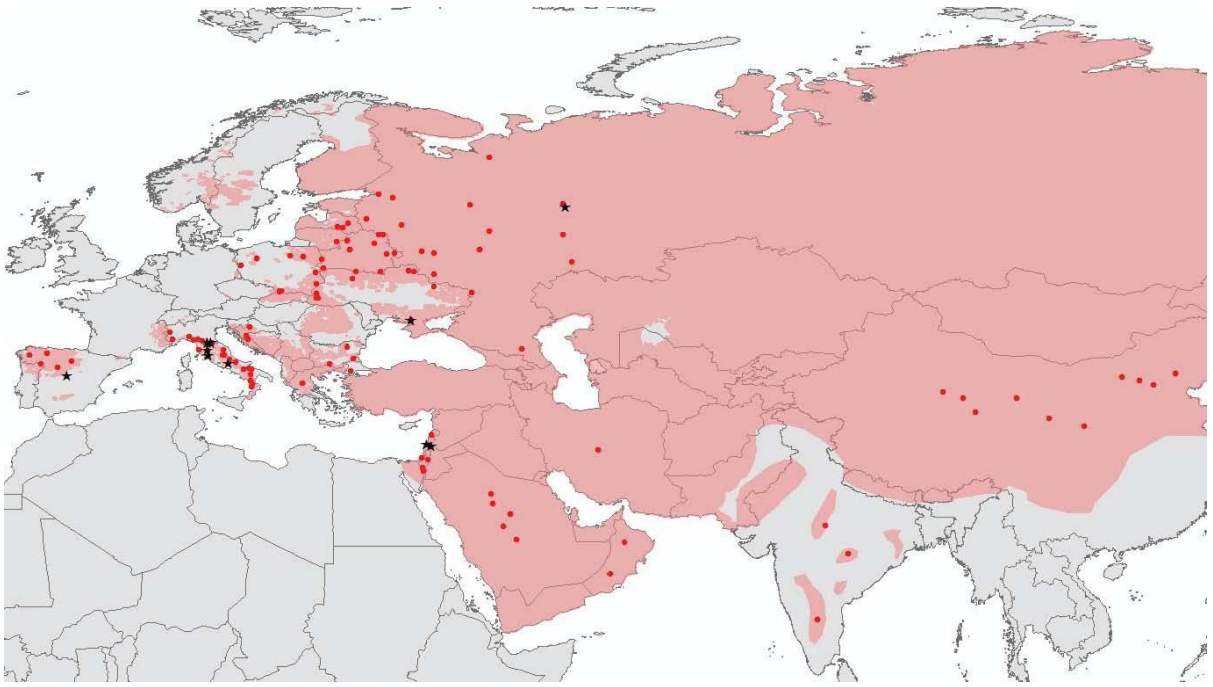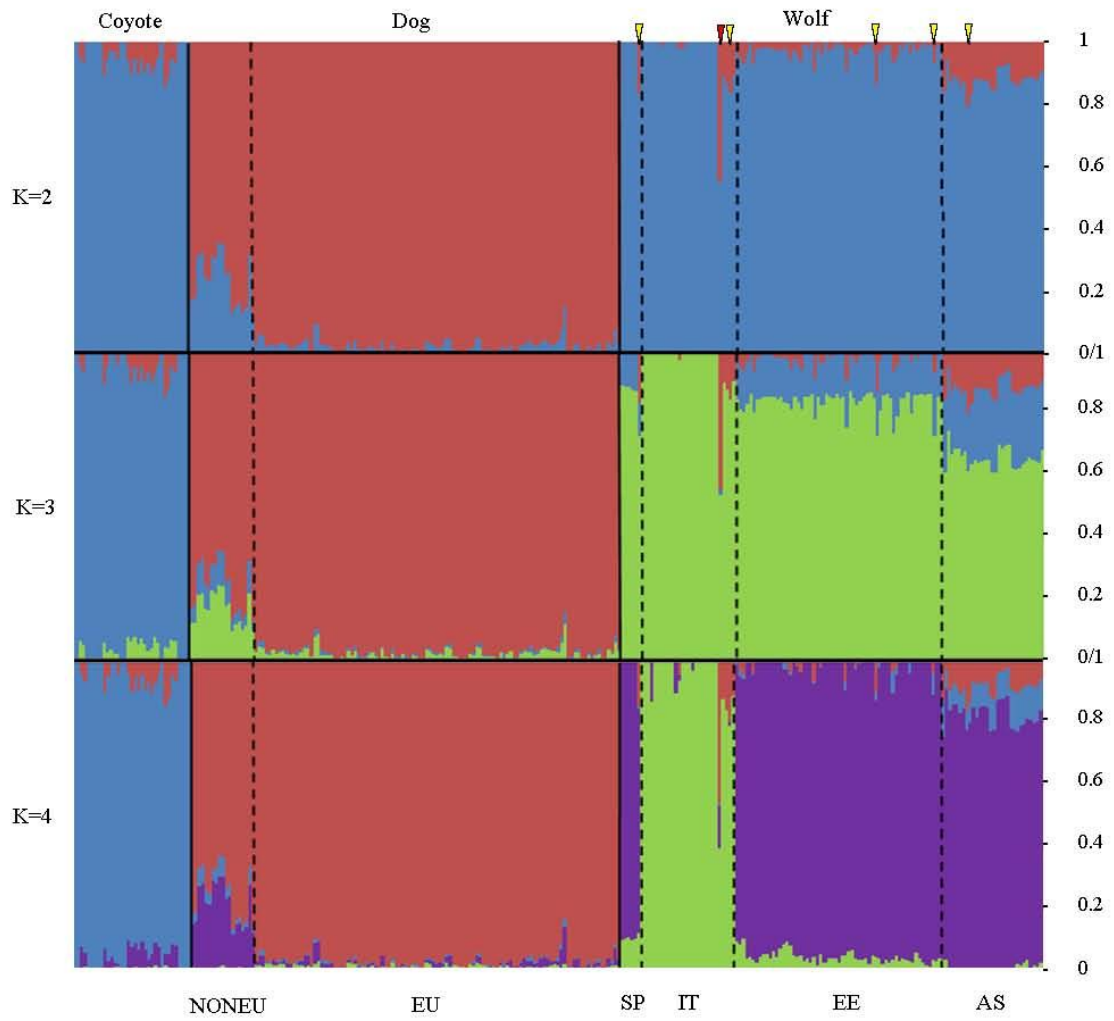
LEGEND

▼ F1 hybrid

▼ Backcross: from left to right in geographical order 1 individual from Spain, 4 from Italy, 1 from Ukraine, 1 from Russia and 2 from Israel.

Chromosome 1

Chromosome 3

dogs

wolves and backcrosses

Chromosome 1

Chromosome 3

Backcross #11254
Backcross #11348

F1 hybrid #2757
Backcross #2770
Backcross #2772
Backcross #2796
Backcross #2803

Backcross #11254
Backcross #11348

F1 hybrid #2757
Backcross #2770
Backcross #2772
Backcross #2796
Backcross #2803

LEGEND
- East European wolves
- Italian wolves
- Spanish wolves
- Iranian and Indian wolves
- Arabian Peninsula wolves
- Chinese wolves
- Domestic dog
- East European F2/F3 backcrosses
- Italian F1 hybrid
- Italian F2/F3 backcrosses
- Arabian Peninsula F2/F3 backcrosses
- Bootstrap support > 50%