Estimating Admixture at the Population Scale: Taking Imperfect Detectability and Uncertainty in Hybrid Classification Seriously

Nina Santostasi, Olivier Gimenez, Romolo Caniglia, Elena Fabbri, Luigi Molinari, Willy Reggioni, Paolo Ciucci

▶ To cite this version:

HAL Id: hal-03367111

https://hal.science/hal-03367111

Submitted on 7 Oct 2021

Research Article

Estimating Admixture at the Population Scale: Taking Imperfect Detectability and Uncertainty in Hybrid Classification Seriously

NINA L. SANTOSTASI (ID),[1] *Department of Biology and Biotechnologies "Charles Darwin", University of Rome La Sapienza, Rome, Italy*

OLIVIER GIMENEZ, *CEFE, CNRS, University of Montpellier, University Paul Valéry Montpellier 3, EPHE, IRD, Montpellier, France*

ROMOLO CANIGLIA, *Italian Institute for Environmental Protection and Research (ISPRA), Unit for Conservation Genetics (BIO–CGE), Ozzano dell'Emilia, Italy*

ELENA FABBRI, *Italian Institute for Environmental Protection and Research (ISPRA), Unit for Conservation Genetics (BIO–CGE), Ozzano dell'Emilia, Italy*

LUIGI MOLINARI, *Wolf Apennine Center, Appennino Tosco-Emiliano National Park, Ligonchio, Italy*

WILLY REGGIONI, *Wolf Apennine Center, Appennino Tosco-Emiliano National Park, Ligonchio, Italy*

PAOLO CIUCCI (ID),[2] *Department of Biology and Biotechnologies "Charles Darwin", University of Rome La Sapienza, Rome, Italy*

ABSTRACT Introgressive hybridization between domestic dogs and wolves (*Canis lupus*) represents an emblematic case of anthropogenic hybridization and is increasingly threatening the genomic integrity of wolf populations expanding into human-modified landscapes. But studies formally estimating prevalence and accounting for imperfect detectability and uncertainty in hybrid classification are lacking. Our goal was to present an approach to formally estimate the proportion of admixture by using a capture-recapture (CR) framework applied to individual multilocus genotypes detected from non-invasive samples collected from a protected wolf population in Italy. We scored individual multilocus genotypes using a panel of 12 microsatellites and assigned genotypes to reference wolf and dog populations through Bayesian clustering procedures. Based on 152 samples, our dataset comprised the capture histories of 39 individuals sampled in 7 wolf packs and was organized in bi-monthly sampling occasions (Aug 2015–May 2016). We fitted CR models using a multievent formulation to explicitly handle uncertainty in individual classification, and accordingly examined 2 model scenarios: one reflecting a traditional approach to classifying individuals (i.e., minimizing the misclassification of wolves as hybrids; Type 1 error), and the other using a more stringent criterion aimed to balance Type 1 and Type 2 error rates (i.e., the misclassification of hybrids as wolves). Compared to the sample proportion of admixed individuals in the dataset (43.6%), formally estimated prevalence was 50% under the first and 70% under the second scenario, with 71.4% and 85.7% of admixed packs, respectively. At the individual level, the proportion of dog ancestry in the wolf population averaged 7.8% (95% CI = 4.4–11%). Balancing between Type 1 and 2 error rates in assignment tests, our second scenario produced an estimate of prevalence 40% higher compared to the alternative scenario, corresponding to a 65% decrease in Type 2 and no increase in Type 1 error rates. Providing a formal and innovative estimation approach to assess prevalence in admixed wild populations, our study confirms previous population modeling indicating that reproductive barriers between wolves and dogs, or dilution of dog genes through backcrossing, should not be expected *per se* to prevent the spread of introgression. As anthropogenic hybridization is increasingly affecting animal species globally, our approach is of interest to a broader audience of wildlife conservationists and practitioners. © 2021 The Authors. *The Journal of Wildlife Management* published by Wiley Periodicals LLC on behalf of The Wildlife Society.

KEY WORDS anthropogenic introgression, Bayesian assignment, *Canis lupus*, capture-recapture, genetic swamping, multievent modeling, prevalence, wolf-dog hybridization.

[1]*E-mail: nina.santostasi@uniroma1.it*
[2]*E-mail: paolo.ciucci@uniroma1.it*

Hybridization between domesticated forms and their wild ancestors is considered an exemplary form of anthropogenic hybridization (Randi 2007). Although this phenomenon has repeatedly occurred since domestication (Frantz et al. 2016, Pilot et al. 2018), there is concern that, especially in human-dominated landscapes, the widespread occurrence of domesticated forms (Boivin et al. 2016) and their potential interactions with their wild ancestors may lead to an increased

risk of hybridization and gene flow (i.e., introgression; Randi 2007, Salvatori et al. 2020). This can eventually lead the rarer wild counterparts to massive introgression and eventually genomic extinction through swamping (Allendorf et al. 2001). Domesticated mammals in general exhibit a suite of morphological, physiological, and behavioral traits (e.g., coat color, craniofacial morphology, dentition, ears and tail shape and length, more frequent and non-seasonal estrus cycles, alterations in the adrenocorticotropic hormone, increased tameness and sociality) that are not observed in their wild counterparts (Wilkins et al. 2014). Although few examples of positive selection of introgressed domesticated traits in wild species have been hypothesized (Anderson et al. 2009, Coulson et al. 2011, Grossen et al. 2014), consistent gene flow between domesticated forms and their wild ancestors is expected to have deleterious consequences for genomic integrity and viability of wild species (e.g., reduction in fitness and adaptive potential, loss of unique combinations of genes and genotypes that have unique evolutionary history; Allendorf et al. 2001, Bohling 2016, Wayne and Shaffer 2016). For these reasons hybridization between wild and domesticated forms is considered a relevant threat to biodiversity (Todesco et al. 2016).

Admixture can be measured in terms of the degree of introgression (i.e., the proportion of alleles from a non-parental taxon averaged across individuals in the population; Miller et al. 2003) or the proportion of admixed individuals in a population (Allendorf et al. 2001); here, we refer to the latter (i.e., proportion of admixture, or prevalence) because it is fundamental to define appropriate management responses and to assess the dynamics of hybridization in a given population (Allendorf et al. 2001). In practical terms, assessing the proportion of admixture can be used in wildlife management decisions according to a 2-step process: first, at the population scale, where assessments are often conveniently based on the genotyping of non-invasive samples, it elucidates if and eventually where management measures are needed (Adams et al. 2003); and second, if population-wide surveys reveal admixture to an extent that necessitates reactive interventions (i.e., sterilization or removal; Gese and Terletzky 2015), more indepth genetic analyses are needed to identify hybrids and accordingly target management actions (vonHoldt et al. 2013). In turn, estimating the proportion of admixture at the population-wide scale requires reliably estimating the abundance of parental and admixed individuals in the population. These estimates, to be reliable, should ideally account for 3 critical aspects. First, they should be based on population samples that reflect biologically meaningful temporal (i.e., generational) and spatial scales. Second, they should derive from estimation methods that formally account for imperfect detectability and other potential sources of bias (Anderson 2001, Yoccoz et al. 2001); in particular, because prevalence is essentially a proportion measuring the relative abundance of admixed and parental individuals, the estimation process should account for a potentially different detectability of the 2 forms (i.e., admixed vs. parental). Third, the inherent uncertainty that generally afflicts the classification of individuals as parental or admixed, especially if based on poor-quality DNA samples, should be formally accounted for within the estimation framework (Santostasi et al. 2019). Specifically, even though genetic markers are considered at large more reliable than phenotypic cues of hybridization (Allendorf et al. 2001), uncertainty in detecting admixed individuals still remains and depends on 2 interacting factors: the number and type of genetic markers used, and the statistical methods and options adopted to assign sampled individuals to the parental or admixed reference populations (Vähä and Primmer 2006, Bohling et al. 2013). Because population-wise assessments are more efficiently conducted analyzing non-invasive samples, especially for elusive and threatened species, the relatively poor-quality DNA extracted from such samples allows for the amplification of a low number of diagnostic loci, therefore limiting the power to discriminate between parental and admixed individuals and their backcrosses (Vähä and Primmer 2006). One of the most commonly used methods to assign sampled individuals to the parental or the admixed category is the Bayesian clustering procedure implemented in programs such as NewHybrids, BAPS, and STRUCTURE (Pritchard et al. 2000, Anderson and Thompson 2002, Falush et al. 2003, Corander et al. 2008). Analysis using STRUCTURE probabilistically assigns individual genotypes to K populations (characterized by distinct allele frequencies) that are assumed to contribute to their gene pools. Specifically, each individual is assigned to a population based on the estimated membership proportion (q_i), which is the fraction of its genome that is inherited from ancestors in 1 of the 2 populations (Pritchard et al. 2000). Admixed individuals are then inferred when their estimated q_i value is intermediate between 2 clusters (a first-generation hybrid should theoretically have a $q_i = 0.5$). The choice of the threshold q_i value that discriminates parental from admixed individuals is traditionally fixed according to an arbitrarily defined standard (e.g., $q_i = 0.8$ for canids; Verardi et al. 2006, Wheeldon et al. 2010, Benson et al. 2012, Rutledge et al. 2012), or by referring to q_i values derived from simulated genotypes of known genealogy (Godinho et al. 2011, van Wyk et al. 2016, Caniglia et al. 2020). We argue that more formal methods should be used to account for the uncertainty in defining threshold q_i values and therefore to assign sampled individuals to the parental or the admixed categories. Moreover, in wolf (*Canis lupus*) × domestic dog hybridization studies, the choice of the threshold q_i value has been generally oriented at avoiding Type 1 error (i.e., misclassifying parental wolves as admixed individuals), with the consequence of underestimating Type 2 error (i.e., misclassifying backcrosses as wolves). Nevertheless, Type 2 error may bear relevant conservation implications, especially in small and expanding wolf populations (Donfrancesco et al. 2019). Especially if hybrids are detected using a limited number of genetic markers, calibrating threshold q_i values to obtain small Type 1 error rates means that a non-trivial proportion of recent backcrosses (Caniglia et al. 2020) could be erroneously classified as wolves. Nonetheless, backcrosses are an indication of ongoing introgression and they carry domesticated alleles that, even if in smaller proportions compared to first-generation hybrids, pose a risk for the genomic integrity of the wolf parental population (Allendorf et al. 2001).

Hybridization between gray wolves and domestic dogs is an emblematic case of wild-domestic hybridization (Butler 1994). Because of intense artificial selection, dogs differ from wolves in several morphological, physiological, and behavioral traits, and many of these differences are genetically based (vonHoldt et al. 2017, Pendleton et al. 2018). Nonetheless, wolves and dogs are interfertile and first-generation hybrids can backcross into the wolf parental population, generating gene flow between the 2 forms (Vilà and Wayne 1999, Randi 2007). In Europe, several wolf populations are currently re-expanding their range across human-dominated landscapes (Chapron et al. 2014) where dogs have become the most abundant carnivore (Ritchie et al. 2014). Accordingly, recent introgressive hybridization (i.e., up to 3 generations in the past; Caniglia et al. 2020) has been recently detected in several wolf populations in Eurasia (Galaverni et al. 2017, Pacheco et al. 2017, Pilot et al. 2018, Salvatori et al. 2020). Detection and monitoring of wolf-dog hybridization within wolf populations is therefore considered a conservation priority at the European scale (Hindrikson et al. 2017, Donfrancesco et al. 2019, Salvatori et al. 2020). But there is no systematic and coordinated management of wolf-dog hybrids currently in place across Europe because of the inherent uncertainties in detecting hybrids and a lack of *ad hoc* planned monitoring programs to formally assess hybridization at the population scale (Salvatori et al. 2020). Previous assessments of wolf-dog hybridization in Europe (Dufresnes et al. 2019) estimated prevalence as the proportion of admixed individuals in a sample drawn from the population (i.e., naïve prevalence) and most of those estimates were based on convenience or opportunistic samples (e.g., incidentally found carcasses or opportunistically collected biological samples) pooled at the country scale and across time frames encompassing several wolf generations. Although these samples may be indicative of broad patterns of introgression, they overlook several sources of sampling and estimation bias and are hardly useful to inform management responses at the appropriate spatial and temporal resolution. More formal yet practical estimation approaches, based on appropriate sampling designs, are needed to assess and monitor wolf-dog hybridization, especially in light of the likely increasing phenomenon throughout Europe (Salvatori et al. 2020).
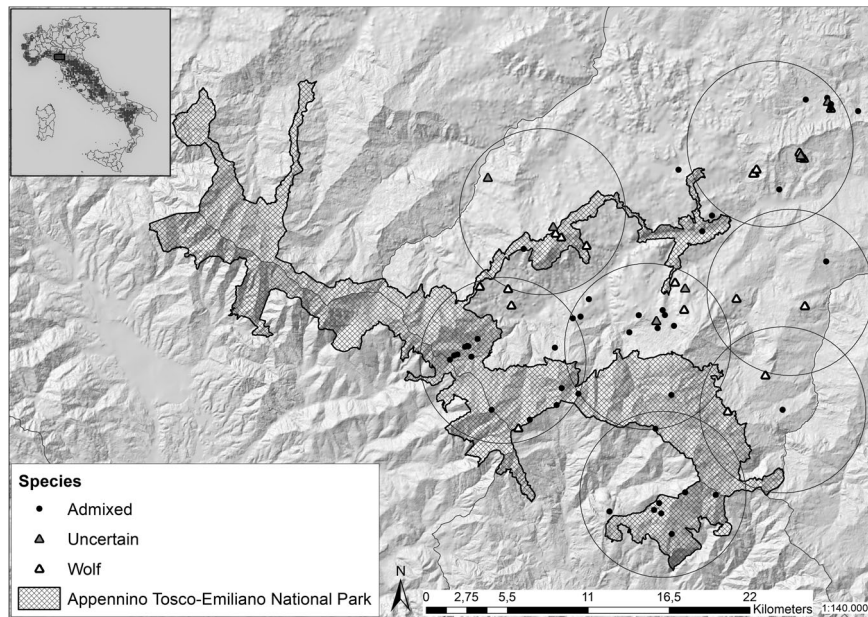
We applied a multievent capture-recapture (CR) model developed by Santostasi et al. (2019) to estimate prevalence of admixture in a wolf population in the northern Apennines, Italy, accounting for imperfect detectability and uncertainty in the probabilistic assignment of admixed individuals. Building on previous work aimed at optimizing the genetic procedures to detect wolf-dog hybrids (Caniglia et al. 2020), our first objective was to estimate the proportion of admixture accounting for the sampling and estimation problems that affect naïve estimates of prevalence; in doing so, we explored 2 alternative rationales to classify admixed individuals: the current practice to minimize Type 1 error versus a more precautionary approach to balance between Type 1 and 2 error rates. Our second objective was to investigate the reproductive status of admixed individuals within the studied population. Based on previous, anecdotal knowledge of admixture in this and adjacent wolf populations, we predicted we would detect a non-trivial extent of admixture and that formal estimates would be higher compared to naïve estimates, especially if false negatives were controlled for in the procedure used to classify admixed individuals (i.e., reducing Type 2 error rates). We also predicted that detection probability and survival would not necessarily be the same between wolves and admixed individuals.

STUDY AREA

Our 731-km^2 study area encompassed the Appennino Tosco–Emiliano National Park (PNATE), in the northern Apennines, Italy (Fig. 1) and the sudy period spanned 10 months (Aug 2015–May 2016). Elevation ranged from 400 to 2,100 m above sea level and the landscape was typically mountainous terrain at higher elevations, and rolling hills at lower altitudes (<1,000 m). The vegetation was mainly composed of temperate and sub-Mediterranean deciduous forests, predominated by beech (*Fagus sylvatica*) at higher elevations, alternated with prairies, meadows, pastures, and cultivated fields. Average annual temperatures were 0.6°C in winter, 10.6°C in spring, 16.2°C in summer and 4.2°C in autumn. Precipitations peak was in October and snow cover usually extended from December to March. Human presence (25 inhabitants/km^2) was limited throughout the year to lower elevations (<1,000 m), although it increased during summer because of tourism and livestock grazing at higher altitudes (Ciucci et al. 2003). In addition to wolves, large mammals in the study area included roe (*Capreolus capreolus*) and red (*Cervus elaphus*) deer and wild boar (*Sus scrofa*), which were hunted outside protected areas from October to December, and mouflon (*Ovis gmelini*).

Wolves naturally recolonized the area in the early 1980s from the central Apennines (Fabbri et al. 2007), and they are now locally established at high and stable density (Caniglia et al. 2014), thriving on wild and occasionally domestic ungulates (Ciucci et al. 1996). Specifically, the area lays within the core of the wolf range in the northern Apennines and comprises the territories of 7 wolf packs that have been intensively surveyed using a combination of field techniques (i.e., wolf-howling during summer, snow-tracking in winter, global positioning system (GPS)-telemetry, non-invasive genetic sampling, camera-trapping; Ciucci and Boitani 1999, Caniglia et al. 2014, Ciucci et al. 2018). Although our study wolf population cannot be considered closed, the tight territorial arrangement of the local wolf packs, and a marked environmental and anthropogenic gradient beyond the study area's borders, suggest our defined study population comprises a demographically and genetically cohesive wolf population. Stray dogs are scarce in the area, but uncontrolled working or hunting dogs may be occasionally present throughout the year, including the hunting season, which overlaps with the wolves' breeding period. A few admixed individuals have been previously reported in the study area (Caniglia et al. 2014).

Figure 1. Location of the study area along the northern Apennines, Italy (inset is the Italian wolf population range) and sampling locations of the wolf scats (dots and triangles) collected within each approximate wolf pack territory (circles) between August 2015 and May 2016. We classified sampled individuals based on multilocus genotypes (12 short tandem repeats [STRs]) and Bayesian clustering analysis (Program STRUCTURE). Based on simulated genotypes, individuals classified as uncertain had estimated membership proportion to the wolf cluster (q_w) values intermediate between those of wolves and admixed (wolf-dog) individuals.

METHODS

Sampling Methods

We applied non-invasive genetic sampling by systematically collecting fresh wolf scats every 2 weeks along fixed routes throughout each of 7 wolf pack territories. We used complementary collection methods (i.e., at scent posts, homesites, kill sites) to boost sample size and enhance sampling coverage. To ensure a demographically meaningful time frame over which to estimate prevalence of admixed individuals, we restricted sampling to a single breeding pulse (i.e., from pup rearing until the onset of the next breeding season) by analyzing wolf scats collected from August 2015 to May 2016, excluding those from pups born in spring 2016. Upon collection, we individually stored fecal samples at –20°C in 95% ethanol. We then extracted DNA from the scats using the Qiagen DNeasy Blood and Tissue Kit (Qiagen, Hilden, Germany) following the manufacturer's instructions and amplified fecal DNA using standard laboratory protocols as described in Caniglia et al. (2014) and Fabbri et al. (2018).

Detection and Classifications of Individuals

Based on a multiple-tube protocol (Taberlet et al. 1996) using procedures described in Fabbri et al. (2018), we genotyped fecal DNA samples amplifying them at 12 unlinked autosomal microsatellites (short tandem repeats [STRs]) selected for their polymorphism and reliable scorability for wolves and dogs (Caniglia et al. 2014) and routinely used for genotyping low-content DNA samples in non-invasive genetic monitoring projects (Caniglia et al. 2013, 2014; Fabbri et al. 2018), and a dominant 3-base pair (bp) deletion (named KB or CBD103DG23) of

the b-defensin CBD103 gene (the K-locus; Anderson et al. 2009), which represents a reliable indicator of dog introgression in some Italian wolf subpopulations (Caniglia et al. 2013). Additionally, we sexed samples by polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP) of the ZFX/ZFY (zinc-finger protein) sequences (Lucchini et al. 2002) and identified paternal haplotypes typing 4 STRs located on the Y chromosome (MS34A, MS34B, MS41A, and MS41B; Sundqvist et al. 2001) and maternal haplotypes analyzing 250 bp of the hypervariable domain of the mtDNA CR1 (Caniglia et al. 2013). We used the software Gimlet version 1.3.3 (Valière 2002) to reconstruct the consensus genotype from the results of the 4–8 replicated amplifications per locus, to estimate PCR success (the number of successful PCRs ÷ number of PCR runs across samples), allelic drop-out, and false allele rates and to match the detected genotypes to each other and to the Italian Institute for Environmental Protection and Research (ISPRA) *Canis* database for the identification of possible resamplings in the study area. We retained as reliable consensus genotypes those showing a reliability score $R \geq 0.95$ obtained by the software RELIOTYPE (Miller et al. 2002).

We performed Bayesian clustering procedures on the 12-loci multilocus reliable genotypes obtained from the 4–8 replicated amplifications per locus per sample using the R package parallel structure (Besnier and Glover 2013). Such software can produce more stable assignment coefficients that are not affected by sample sizes or samples with variable levels of admixture (Caniglia et al. 2020) because the software automatically subdivides a dataset of genotypes to be assigned

to predefined reference populations into multiple single projects (each project is composed of the reference populations and 1 of the genotypes to be assigned), which are independently run. For each individual genotype, we estimated the individual proportions of membership (q_i) and the 90% Bayesian credible intervals (BCI) to the 2 inferred clusters (K); details about the Bayesian assignment test models are reported in Caniglia et al. (2020). We used the admixture and the independent allele frequencies models, which are considered the most suitable approaches to trace gene flow between taxa with allele frequencies clearly different and that evolved independently (Falush et al. 2003). We performed 4 independent runs at K = 2 applying 500,000 Markov chain Monte Carlo iterations, discarding the first 50,000 as burn-in. We averaged results across different runs using the software CLUMPP 1.1.1 (Jakobsson and Rosenberg 2007). We used as reference populations a panel of 190 wolves and 89 dogs. We obtained wolf samples from areas with no documented cases of admixture and the wolves in this sample showed the typical wild coat color pattern and no apparent signal of other morphological dog-like traits (e.g., black coats, white claws, spur on the hind legs), did not share dog-derived Y and mtDNA haplotypes, and had a $q_w > 0.990$, a value obtained in previous Bayesian assignment procedures performed using 156,132 canine single-nucleotide polymorphisms (SNPs) and 39 canine STRs commonly used in recent studies on wolf-dog hybridization in Europe (Galaverni et al. 2017, Fabbri et al. 2018, Caniglia et al. 2020). Dog samples comprised 61 free-ranging dogs sampled in the same areas (Randi et al. 2014, Galaverni et al. 2017), plus 1 male and 1 female randomly chosen from 14 wolf-sized dog breeds available from the LUPA project data set (Lequarre et al. 2011, Vaysse et al. 2011, Caniglia et al. 2020). We used samples from the reference populations also in HybridLab (Nielsen et al. 2006) to simulate 12-STR genotypes for 100 individuals per each of the following parental and admixed ancestry classes: wild (PW) and domestic (PD) parentals, first (F1) and second (F2) generation hybrids, and 4 backcross generations originated either from F1s (BCW1–BCW4) or F2s (SBCW1–SBCW4) crossing with wild parentals (Caniglia et al. 2020), for 1,200 genotypes.
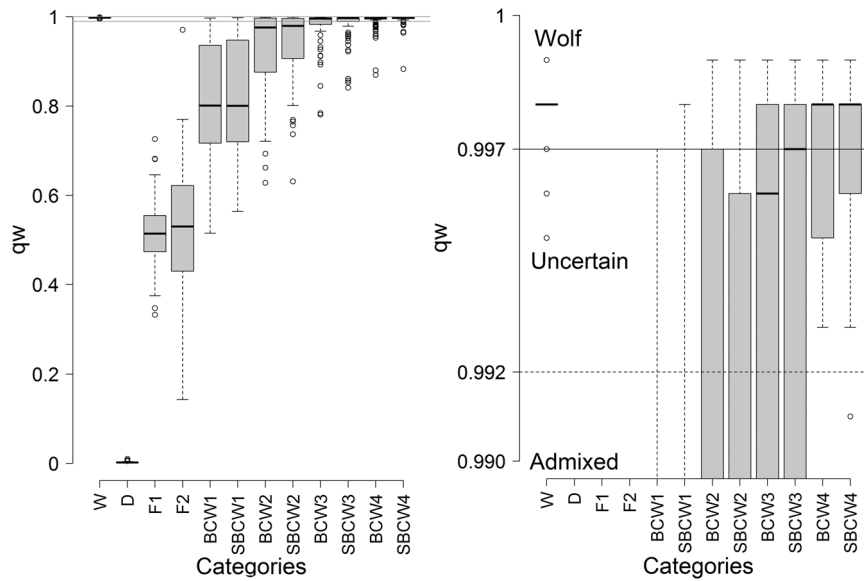
We classified the sampled genotypes as wolves or admixed individuals by comparing their individual q_i values with those of the reference populations and of simulated genotypes. In doing so, we examined 2 classification scenarios to gauge their effect on the final estimate of prevalence. The first classification (scenario A) reflected the widely adopted procedures essentially meant to reduce Type 1 error rates (Randi 2007, Godinho et al. 2011, Pacheco et al. 2017, Dufresnes et al. 2019). We accordingly used the threshold $q_w \leq 0.975$, as suggested by Caniglia et al. (2020) to identify recently admixed individuals in the Italian wolf population with a reduced 12-STR marker panel. The second classification (scenario B) reflected a more precautionary approach meant to balance between Type 1 and 2 error rates, thus reducing the number of simulated first- and second-generation backcrosses erroneously assigned to parental

wolves. To formally account for the uncertainty in classification due to the expectedly large overlap in q_w values between wolves and backcrosses using a relatively limited number of loci (Vähä and Primmer 2006), we classified individuals according to 3 categories: wolves, including individuals whose q_w was higher than the maximum q_w of simulated first-generation backcrosses (BCW1) because this prevented erroneously assigning first-generation backcrosses to the wolf category (i.e., Type 2 error); admixed individuals, including those whose q_w was lower than the minimum q_w of reference wolves because this prevented erroneously assigning wolves to the admixed category (i.e., Type 1 error); and uncertains, including individuals whose q_w was between the minimum q_w of reference wolves and the maximum q_w of simulated BCW1 (Fig. 2; Table S1, available online in Supporting Information). To this aim, we used the minimum q_w of reference instead of simulated wolves because the former more realistically represents the expected genetic variability in the wolf population (Dufresnes et al. 2019). Whereas the first 2 categories are by definition those traditionally recognized in Bayesian-based assignment tests (i.e., scenario A), we considered the category uncertains only in the classification scenario B (see below).

For each of the 2 scenarios, we obtained a CR dataset containing the capture histories of the observed individuals organized in bi-monthly sampling occasions. We analyzed both datasets using CR models to estimate the abundance of wolves and admixed individuals.

Capture-Recapture Modeling and Prevalence Estimation

The multievent formulation of open population CR models explicitly handles uncertainty in individual classification by modeling the observed capture histories as 2 time series: the state process (i.e., the population dynamics during the study) and the event process (i.e., what we can observe through sampling; Pradel 2005). Following Pradel (2005) and Santostasi et al. (2019), we modeled the state process as a Markov chain of 3 partially hidden states: alive in the study area as wolf, alive in the study area as admixed, and dead or permanently emigrated. The state process was described by the initial state probability (π_w = the probability that an individual was in one or the other state when first encountered) and the apparent survival probability (φ = the probability that an individual survived and remained in the study area between sampling occasions). In our model, because individuals could not change their state between wolf and admixed, the only possible transition was between in the study area and permanently emigrated or dead; therefore, their state changed over time according to a first-order Markov process determined by the apparent survival probabilities only (Santostasi et al. 2019). We modeled the event process as conditional on the underlying state (to be observed, individuals had to be alive and in the study area) and represented by 2 consecutive steps: detection (p is the probability of finding and successfully

Figure 2. The left panel presents a boxplot of the q_w (estimated membership proportion to the wolf cluster) values obtained by Bayesian clustering analysis of the genotypes simulated from reference wolves ($n = 190$) and dogs ($n = 89$) in the northern Apennines, Italy, 2015–2016, grouped per genealogical class ($n = 100$ genotypes for each class). The grey horizontal lines define the area of uncertainty that is zoomed in the right panel. The right panel provides detail of the area of uncertainty between the minimum q_w of reference wolves (dashed line) and the maximum q_w of simulated first-generation backcrosses (solid line). We classified as wolves those sampled genotypes whose q_w was above the solid line, as admixed those whose q_w was below the dashed line, and as uncertains those in between. W = wolves; D = dogs; F1 and F2 = first- and second-generation hybrids, respectively; BCW1–4 = first- to fourth-generation backcrosses of F1 with wolves; SBCW1–4 = first- to fourth-generation backcrosses of F2 with wolves.

scoring an individual genotype) and state assignment (i.e., individual classification, δ is the probability of classifying an individual as wolf or admixed according to its q_w value). In classification scenario A, we did not consider uncertainty in the state assignment, reflecting the traditional adoption of a fixed threshold q_w value to discriminate between wolves and admixed individuals, and accordingly modeled the probability of assigning an individual to 1 of the 2 states (δ) as equal to 1, whereas the complementary probability (1 – δ: the probability of not classifying an individual) was equal to 0. Therefore, under scenario A, our model considered only 3 possible events, corresponding to individuals that could be 1) detected and classified as wolf, recorded as 1 in the capture history; 2) detected and classified as admixed, recorded as 2; and 3) not detected, recorded as 0. This model formulation corresponded to a multistate CR model, which assumed that each individual could be classified without ambiguity (Conn and Cooch 2009). In scenario B we accounted for uncertainty in the assignment procedure considering 4 possible events corresponding to individuals that could be 1) detected and classified as wolf, recorded as 1 in the capture history; 2) detected and classified as admixed, recorded as 2; 3) detected but not classified (i.e., classified as uncertain; recorded as 3); and 4) not detected, recorded as 0. Under this classification scenario, we did not assume δ was equal to 1, but it became a parameter to be estimated, and we used the complementary probability (1 – δ) to model the capture histories of the individuals classified as uncertain (Santostasi et al. 2019). For example, to illustrate the calculation of an encounter history of an

uncertain individual in a 3-session CR experiment, 303 would denote an individual detected on the first occasion, not detected on the second occasion, and detected again on the third occasion. The state of this individual (i.e., wolf or admixed) is not assigned in this phase. Assuming that the parameters do not change over time, the capture history can be written as (Santostasi et al. 2019):

$$\Pr(303) = \pi_h(1 - \delta_h)\varphi_h(1 - p_h)\varphi_h p_h(1 - \delta_h)$$
$$+ \pi_w(1 - \delta_w)\varphi_w(1 - p_w)\varphi_w p_w(1 - \delta_w),$$

where the subscripts h and w indicate parameters for admixed individuals and wolves, respectively.

Assuming independence between observations, we obtained the likelihood of the entire dataset as the product of the probabilities of all individual encounter histories and we obtained the parameter estimates by maximizing the likelihood function (Pradel 2005).

With scenario A, we estimated the population abundance at a given occasion ($\hat{N}$), given by the sum of the estimated abundances of parental wolves ($\hat{N}_w$) and of admixed individuals ($\hat{N}_h$), with the Horvitz–Thompson estimator (McDonald and Amstrup 2001) as:

$$\hat{N} = \frac{n_w}{\hat{p}_w} + \frac{n_h}{\hat{p}_h} = \hat{N}_w + \hat{N}_h,$$

where n_w was the number of parental individuals and n_h the number of hybrids detected and $\hat{p}_w$ and $\hat{p}_h$ were the estimated detectabilities of parental and admixed individuals,

respectively. We then estimated prevalence as (Santostasi et al. 2019):

$$Prevalence = \frac{\hat{N}_h}{\hat{N}_w + \hat{N}_h}.$$

With scenario B, the individuals classified as uncertain needed to be assigned to 1 of the 2 states (wolf or admixed). To solve the problem, we used the Viterbi algorithm, which, given the observed encounter histories of the uncertain individuals and the parameters estimated by the multievent model, reconstructed the most likely sequence of states that has generated the observed data (Rouan et al. 2009, Zucchini et al. 2016, Santostasi et al. 2019). We ran the Viterbi algorithm and calculated the probability of the 4 (the maximum possible with Program E–SURGE; Choquet et al. 2009) most likely capture history reconstructions for each uncertain individual. Then, once we reconstructed the number of observed wolves and admixed individuals, we used their respective abundance estimates to estimate prevalence. We estimated standard deviation and confidence intervals of the abundance and prevalence estimates via a non-parametric bootstrap (Davison and Hinkley 2002). We took a step further compared to Santostasi et al. (2019) and evaluated the extent to which the final prevalence estimate would be affected by taking into account the alternative and less likely Viterbi reconstructions. To do this, we calculated first the relative probability (weight) of each Viterbi reconstruction, and then a weighted estimate of prevalence (see Supporting Information for additional details).

For both scenarios A and B, to test for sources of variation in the probability of detection and apparent survival, we built a set of candidate models incorporating the effect of biologically relevant and time-dependent variables. We tested models with a state effect on detection and survival probability to test for possible differences that could arise from different behavior of parental and admixed individuals as reported in other species (Derégnaucourt et al. 2004, Battocchio et al. 2017), a time (capture occasion) effect on detection probability to test for heterogeneity due to variation in sampling effort or environmental conditions, and a pack effect on detection probability and survival to test for heterogeneity due to uneven spatial distribution of the sampling effort. We did not test for a time effect on apparent survival because our sampling was extended over a relatively short time frame. Based on several CR studies on wolf populations (Marucco et al. 2009; Cubaynes et al. 2010, 2014; Caniglia et al. 2014) that did not find a significant sex effect on apparent survival, we decided not to run models with sex-dependent survival. Because we performed the classification of genotypes only once for each genotype, we constrained the assignment probability to be estimated upon first capture only (Santostasi et al. 2019). We fitted the CR models and ran the Viterbi algorithm using Program E–SURGE (Choquet et al. 2009). We compared the models based on Akaike's Information Criterion corrected for small sample size (AIC_c). To account for uncertainty in model selection, we obtained model-averaged estimates considering models whose ΔAIC_c was ≤2 from the best-selected model (Burnham and Anderson 2002). In the absence of a goodness-of-fit test for multievent CR models (Pradel 2009), we used the R package R2ucare (Gimenez et al. 2018) to evaluate the fit of our data to the Cormack Jolly Seber model (Lebreton et al. 1992) that had the same structure but did not allow uncertainty in state assignment (Gimenez et al. 2012).

Ancestry Analysis and Genealogy

To reconstruct the ancestry of the sampled individual genotypes and to estimate their individual posterior probability of belonging to the assigned or other parental population, or of having a recent ancestor in either (Hubisz et al. 2009), we used the option population information to test for migrants implemented in STRUCTURE 2.3.4 (Falush et al. 2003). We a priori assigned individuals to the wolf or dog parental populations (2 genetic clusters) using the independent allele frequencies model and the POPFLAG = 1 to activate the POPINFO option. In this way we assumed that all reference wolves and dogs and the collected genotypes were a priori correctly identified and assigned to their own cluster. Because of a limited number of loci, we restricted the ancestry analysis to 2 generations backward (GENBACK = 2). Even in this analysis, we performed 4 independent runs at K = 2 and averaged results using the software CLUMPP 1.1.1 (Jakobsson and Rosenberg 2007). In addition, we estimated the genealogies of the pack members with a maximum likelihood approach implemented in COLONY 2.0 (Wang and Santure 2009) intending to investigate if admixed individuals produced litters (see the Genealogy Analyses section in the Supporting Information for further details).

RESULTS

Out of 152 collected scats, 65% were reliably genotyped an had an average number of positive amplifications per locus of 0.77 (range = 0.39–0.92), and average error rates of allelic drop-out of 0.18 ± 0.11 (SD) and false allele rates of 0.05 ± 0.03. We grouped consensus genotypes and assigned them to 39 individuals (15 females, 21 males, 3 of unknown sex; Table S2, available online in Supporting Information), sampled on average in 1.6 ± 0.9 capture occasions (Table S3, available online in Supporting Information). We sampled from 2−12 individuals in each of the 7 packs. All the 39 sampled individuals shared the typical Italian wolf mtDNA and none had the K^b melanistic deletion; 4 individuals shared a dog-derived Y haplotype, all assigned to the wolf category under scenario A, whereas under scenario B, 3 of them were classified as admixed and 1 as a wolf (Table S2).

Based on scenario A, out of the 39 sampled individuals, 22 were classified as wolf ($q_w \geq 0.975$) and 17 were classified as admixed ($q_w < 0.975$; Table S2), corresponding to a naïve prevalence of 43.6% and an average proportion of admixture of 7.8 (95% CI = 4.4−11%). Based on the simulated genotypes, this q_w threshold correctly identified 100% of simulated parental wolves ($n = 100$), and 100% of the F1
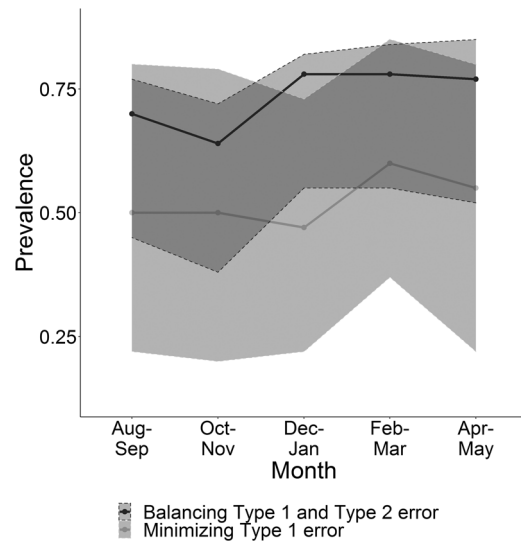
Table 1. Most supported capture-recapture models and model-averaged estimates of prevalence of wolf-dog admixed individuals in the northern Apennines, Italy, 2015–2016, for scenario A and B (p = detectability, π = initial state probability, φ = apparent survival probability, δ = assignment probability). The term (.) indicates constant parameters, the term (state) indicates state-dependent (i.e., wolf vs. admixed) parameters, the term (time) indicates occasion-dependent parameters. For each model, we provide the number of parameters (K), Akaike's Information Criterion corrected for small sample sizes (AIC$_c$), and difference in AIC$_c$ (ΔAIC$_c$). We considered models that had a ΔAIC$_c$ < 2 to be competitive and model averaged if multiple models were competitive.

Model	K	Deviance	AIC$_c$	ΔAIC$_c$
Scenario A				
$\pi(.)p(.)\varphi(state)$	4	152.37	161.03	0.00
$\pi(.)p(state)\varphi(state)$	5	152.34	163.35	2.32
$\pi(.)p(state)\varphi(.)$	4	155.10	163.76	2.73
$\pi(.)p(time)\varphi(state)$	7	148.41	164.38	3.35
$\pi(.)p(.)\varphi(.)$	3	158.03	164.43	3.40
$\pi(.)p(time)\varphi(.)$	6	154.15	167.60	6.57
$\pi(.)p(.)\varphi(pack)$	9	147.13	168.41	7.38
$\pi(.)p(state)\varphi(pack)$	10	146.57	170.65	9.62
$\pi(.)p(pack)\varphi(state)$	10	149.59	173.67	12.64
$\pi(.)p(time)\varphi(pack)$	12	144.11	174.11	13.08
$\pi(.)p(pack)\varphi(.)$	9	152.84	174.11	13.08
$\pi(.)p(pack)\varphi(pack)$	15	143.88	183.68	22.65
Scenario B				
$\pi(.)p(.)\varphi(.)\delta(.)$	4	176.49	185.17	0.00
$\pi(.)p(.)\varphi(state)\delta(.)$	5	175.46	186.49	1.32
$\pi(.)p(state)\varphi(.)\delta(.)$	5	175.62	186.65	1.48
$\pi(.)p(.)\varphi(state)\delta(.)$	6	175.39	188.86	3.67
$\pi(.)p(.)\varphi(pack)\delta(.)$	10	165.46	189.61	4.44
$\pi(.)p(time)\varphi(pack)\delta(.)$	7	173.64	189.64	4.47
$\pi(.)p(.)\varphi(state)\delta(.)$	8	172.62	191.23	6.06
$\pi(.)p(state)\varphi(pack)\delta(.)$	11	165.46	192.54	7.36
$\pi(.)p(pack)\varphi(.)\delta(.)$	10	171.45	195.60	10.42
$\pi(.)p(time)\varphi(pack)\delta(.)$	13	163.32	196.60	11.43
$\pi(.)p(pack)\varphi(state)\delta(.)$	11	171.41	198.48	13.31
$\pi(.)p(pack)\varphi(pack)\delta(.)$	16	162.74	206.31	21.14

Figure 3. Prevalence of admixed (wolf-dog) individuals in a protected wolf population in the northern Apennines, Italy, as assessed by non-invasive genetic sampling between August 2015−May 2016 and Bayesian clustering analysis based on multilocus genotypes (12 short tandem repeats [STRs]). Estimates of prevalence, obtained through an open population capture-recapture modeling approach, are shown for each 2-month sampling occasion. The 2 scenarios refer to 2 alternative rationales to cope with uncertainty in the classification of admixed individuals (i.e., the traditional practice of minimizing Type 1 error vs. a more precautionary approach balancing between Type 1 and 2 error rate).

(n = 100), 100% of the F2 (n = 100), 86% of the BCW1 (n = 86), 48% of the BCW2 (n = 48), 13% of the BCW3 (n = 13), and 8% of the BCW4 (n = 8) individuals (Table S4, available online in Supporting Information). The most supported model revealed different apparent survival between sampling occasions for wolves and admixed individuals and constant detection probability (Table 1). Upon first capture, individuals had a 0.56 probability (95% CI = 0.41−0.71) of being wolf (π_w) and a 0.44 (95% CI = 0.29−0.59) probability of being admixed. The probability of apparent survival between occasions was $\varphi_W = 0.56$

(95% CI = 0.31−0.79) for wolves and $\varphi_b = 0.92$ (95% CI = 0.38−0.99) for admixed individuals. Overall survival over the 10-month period (the product of the 4 bi-monthly estimates of apparent survival) was 0.10 (95% CI = 0.009−0.39) for wolves and 0.72 (95% CI = 0.02−0.96) for admixed individuals. We estimated a probability of detection $p = 0.48$ (95% CI = 0.29−0.68) for wolves and admixed individuals. Within each capture occasion, total population abundance ranged from 25 (95% CI = 10−44) to 31 (95% CI = 17−50) individuals, including 10 (95% CI = 3−26) to 17 (95% CI = 7−30) wolves and 13 (95% CI = 3−21) to 19 (95% CI = 8−31) admixed individuals (Table 2). Estimated prevalence under scenario A therefore ranged from 47% (95% CI = 21−79%) to 60% (95% CI = 37−84%) by sampling occasion and averaged 50% (95% CI = 22−80%) across the 10-month sampling period (Fig. 3; Table 2). Under scenario A, the proportion of admixed packs was 71.4%.

Table 2. Model-averaged estimates of prevalence of admixture (wolf-dog) in a protected wolf population in the northern Apennines, Italy, 2015–2016, estimated with multistate and multievent capture-recapture models. The 95% confidence intervals of the estimates are reported between parentheses.

	Sampling occasion				
	Aug–Sep	**Oct−Nov**	**Dec−Jan**	**Feb−Mar**	**Apr−May**
Scenario A					
Poulation abundance	25 (11–44)	25 (11–39)	31 (17–50)	31 (16–50)	23 (11–44)
Wolf abundance	13 (3–26)	13 (4–25)	17 (7–30)	13 (4–24)	10 (3–26)
Admixed abundance	13 (3–26)	13 (3–21)	15 (5–28)	19 (8–31)	13 (3–26)
Prevalence	0.50 (0.22–0.80)	0.50 (0.20–0.79)	0.47 (0.21–0.79)	0.60 (0.37–0.84)	0.55 (0.22–0.80)
Scenario B					
Poulation abundance	26 (13–78)	28 (14–85)	35 (17–109)	35 (17–109)	26 (13–80)
Wolf abundance	7 (3–26)	10 (4–35)	7 (3–26)	7 (3–26)	6 (2–20)
Admixed abundance	17 (8–48)	17 (8–48)	26 (12–72)	26 (12–72)	19 (8–51)
Prevalence	0.70 (0.45–0.77)	0.64 (0.38–0.72)	0.78 (0.55–0.82)	0.78 (0.55–0.84)	0.77 (0.52–0.85)

Conversely, based on scenario B, out of the 39 sampled individuals, 10 were classified as wolf ($q_w > 0.997$), 23 were classified as admixed ($q_w < 0.990$), and 6 were classified as uncertain ($0.990 \leq q_w \leq 0.997$; Table S2). Based on the simulated genotypes, but excluding those classified as uncertain ($n = 186$), all simulated wolves ($n = 76$), F1 ($n = 100$), F2 ($n = 100$), and BCW1 ($n = 94$) individuals were correctly classified. But 23% of BCW2 ($n = 79$) individuals were erroneously classified as wolves, a Type 2 error rate that increased to 43−83% for further generations of backcrosses (Table S5, available online in Supporting Information). The most supported models revealed constant detection and apparent survival probabilities, state-dependent survival, or constant apparent survival and state-dependent detection probability (Table 1). Upon first capture, individuals had a 0.30 (95% CI = 0.17−0.48) probability of being wolf (π_w) and a 0.70 (95% CI = 0.32−0.93) probability of being admixed. The probability of apparent survival between occasions was $\varphi_w = 0.73$ (95% CI = 0.40−0.91) and $\varphi_h = 0.77$ (95% CI = 0.50−0.93) for wolves and admixed individuals, respectively. Survival for the entire study period was 0.28 (95% CI = 0.03−0.75) for wolves and 0.35 (95% CI = 0.06–0.75) for admixed individuals. We estimated the probability of detecting a wolf as $p_w = 0.43$ (95% CI = 0.16−0.67) and the probability of detecting an admixed individual as $p_h = 0.47$ (95% CI = 0.26−0.69). Upon detection, the probability of being assigned either to the wolf or admixed categories was $\delta = 0.85$ (95% CI = 0.70−0.93). Based on the Viterbi algorithm, uncertain individuals had a higher probability of assignment to the category admixed (range = 0.31−0.80) than to the category wolf (0.09−0.31; Table S6, available online in Supporting Information). Within each capture occasion, total population abundance ranged from 26 (95% CI = 13−78) to 35 (95% CI = 17−109) individuals, including 6 (95% CI = 2−20) to 10 (95% CI = 4−35) wolves and 17 (95% CI = 8−48) to 26 (95% CI = 12−72) admixed individuals (Table 2). Estimated prevalence under scenario B therefore ranged from 64% (95% CI = 38−72%) to 78% (95% CI = 55−84%) by sampling occasion and averaged 70% (95% CI = 45−77%) across the 10-month sampling period (Fig. 3; Table 2). The second, third, and fourth most likely alternative Viterbi reconstructions had considerably lower weight than the most likely reconstruction (Table S7, available online in Supporting Information) and only marginally affected the estimated prevalence (range = 0.61–0.76 vs. 0.64–0.78; Table S8, available online in Supporting Information). Under scenario B, the proportion of admixed packs was 85.7%. The goodness-of-fit test did not detect signs of lack of fit (i.e., transience or trap-dependence) to the Cormack Jolly Seber model ($P > 0.05$ for all the tests; Table S9, available online in Supporting Information).

Ancestry Analysis and Genealogy
Ancestry analysis revealed that none of the 17 admixed individuals detected by Bayesian clustering procedures according to scenario A was a first- or second-generation hybrid; however, 41% ($n = 7$) of the admixed individuals had non-negligible posterior probabilities (0.120–0.999) of having a grandparent in the dog population, hence of being first-generation backcrosses (Table 3). The remaining admixed individuals likely originated from backcrosses of further generations backward (i.e., ≥BCW2). Despite a relatively limited sample size, through the genealogy estimation we were able to identify (posterior probability >0.90) likely admixed breeding pairs in 2 of the 7 surveyed packs. Specifically, according to scenario A, the 2 admixed breeding pairs had 1 admixed breeder each, whereas, based on scenario B, 1 of the 2 pairs had 2 and the other 1 admixed breeders (Fig. 4). All of the admixed breeders were identified to be more than first-generation backcrosses in the ancestry analysis (Table 3). In 2015, these 2 admixed breeding pairs produced a minimum of 3 and 5 offspring each. Under scenario A, 7 of these offspring were classified as wolf and 1 as admixed, whereas according to scenario B only 1 of them was classified as wolf, 3 as admixed, and 4 as uncertain; the latter, however, were successively classified as admixed based on the Viterbi algorithm (Table S6).

DISCUSSION
In this study, we presented an innovative approach to formally estimate population-wide prevalence of admixture in wild populations affected by anthropogenic hybridization. By recognizing that naïve estimates of the proportion of admixture generally suffer from various sources of bias, we applied a CR-based estimation approach to a local wolf population, accounting for imperfect detectability and uncertainty in hybrid classification, both typically associated with the genotyping of non-invasive samples. Contrary to our approach, naïve estimates of prevalence do not take into account detectability, and previous simulations demonstrated they are biased when the probability of detection of parental and admixed individuals is heterogeneous (Santostasi et al. 2019). Heterogeneity in detectability may arise because of expected differences in ecology and behavior between parental and admixed individuals (e.g., social behavior, marking rate, spatial patterns, resource selection), or because of sampling bias (e.g., spatially heterogeneous effort, edge effect; Crespin et al. 2008). Although we did not find strong evidence of such heterogeneity in our application, under scenario B high-ranking models did consider the effect of heterogeneous detectability, suggesting that admixed individuals may have a higher probability of detection.
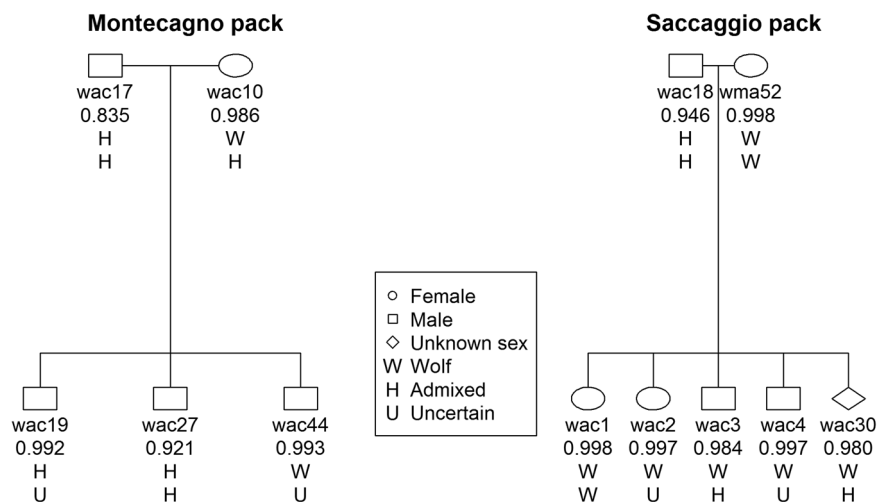
A critical point, common to both naïve and formal approaches to estimate prevalence, concerns the reliability of the individual classification based on Bayesian clustering techniques. In fact, in addition to the number and type of molecular markers, the classification is strongly affected by the composition of the chosen reference samples and by the threshold q_i values chosen to discriminate admixed from parental individuals (Vähä and Primmer 2006). To overcome these problems, we relied on an improved procedure (Caniglia et al. 2020), based on carefully screened reference samples of pure Italian wolves, whose selection was aided by genomic tools (Galaverni et al. 2017), and a back-end executable of the software STRUCTURE (i.e., R package parallel structure) that produces more stable assignment coefficients that are not affected by samples with variable

Table 3. Inferred ancestry of the 23 wolf-dog admixed and 6 uncertain individuals (according to scenario B) observed in northern Apennines, Italy, 2015–2016, using the population information model implemented in STRUCTURE. We show estimates of posterior probabilities for each individual to have ancestry in its *a priori* assigned population (q prior population), or in the other population in the present generation (dog or wolf), in the first past generation (parent) or the second past generation (grandparent).

Individual	Prior population	q prior population	Dog	Dog parent	Dog grandparent
HAC10	Wolf	0.995	0.000	0.000	0.005
HAC12	Wolf	0.951	0.000	0.000	0.049
HAC13	Wolf	0.983	0.000	0.000	0.017
HAC14	Wolf	0.993	0.000	0.000	0.007
HAC16	Wolf	0.663	0.000	0.000	0.337
HAC17	Wolf	0.801	0.000	0.000	0.199
HAC1	Wolf	0.819	0.000	0.000	0.181
HAC2	Wolf	0.858	0.000	0.000	0.142
HAC3	Wolf	0.936	0.000	0.000	0.064
HAC8	Wolf	0.985	0.000	0.000	0.015
HRE1	Wolf	0.429	0.000	0.008	0.563
W1967	Wolf	0.998	0.000	0.000	0.002
WAC10	Wolf	0.973	0.000	0.000	0.027
WAC17	Wolf	0.965	0.000	0.000	0.035
WAC18	Wolf	0.964	0.000	0.000	0.036
WAC19	Wolf	0.999	0.000	0.000	0.001
WAC22	Wolf	0.955	0.000	0.000	0.045
WAC23	Wolf	0.998	0.000	0.000	0.002
WAC25	Wolf	0.955	0.000	0.000	0.045
WAC27	Wolf	0.922	0.000	0.000	0.078
WAC29	Wolf	0.995	0.000	0.000	0.005
WAC2	Wolf	1.000	0.000	0.000	0.000
WAC30	Wolf	0.998	0.000	0.000	0.002
WAC33	Wolf	0.735	0.000	0.000	0.265
WAC3	Wolf	0.993	0.000	0.000	0.007
WAC42	Wolf	0.996	0.000	0.000	0.004
WAC44	Wolf	0.994	0.000	0.000	0.006
WAC46	Wolf	0.880	0.000	0.000	0.120
WAC4	Wolf	1.000	0.000	0.000	0.000

levels of admixture because these are analyzed one by one (Besnier and Glover 2013). Previous studies comparing the assignment and the detection power of admixed individuals of several Bayesian software options showed that the results obtained by STRUCTURE are comparable to those obtained using other Bayesian approaches, such as those implemented in NewHybrids (Caniglia et al. 2020) or BAPS (Randi et al. 2014), and better than partially Bayesian approaches like those implemented in GeneClass (Piry et al. 2004, Sanz et al. 2009). We nevertheless used a

Figure 4. Reconstructed genealogies for 2 of the 7 wolf packs non-invasively sampled in the northern Apennines, Italy, 2015–2016, to estimate admixture between wolves and dogs. For each pack, the 2 likely breeding individuals are on top of the diagram and are connected to their progeny through a vertical branch. For each symbol, the first line reports the individual identification, the second line the individual q_w (estimated membership proportion to the wolf cluster), the third line their classification based on a classification rationale aimed at minimizing Type 1 error (scenario A), and the fourth line their classification based on a classification rationale aimed at balancing between Type 1 and 2 error rates (scenario B). All individuals originally classified as uncertains (U) based on q_w were successively assigned to the admixed category (H) by a probabilistic, *a posteriori* procedure based on the Viterbi algorithm.

relatively limited number of microsatellite loci, though this reflects what is routinely done in population-wide surveys adopting non-invasive genetic sampling (Caniglia et al. 2014, Fabbri et al. 2018). This implies a relatively reduced power to detect second- or further generation backcrosses. In these circumstances, the traditional assignment approach is to define a threshold q_w value that comprises all wolves of the reference population to reduce Type 1 error rate in the assignment tests. According to this approach, for example, Godinho et al. (2011) and Randi (2007), using simulated genotypes, estimated that the proportions of first-generation backcrosses erroneously classified as wolves were 16% and 20%, respectively. Similar to other conservation contexts (McGarvey 2007, Saltz 2011), however, it is also relevant to consider Type 2 error rates: specifically, an overlooked occurrence of admixed individuals, especially in small and expanding populations, may increase the degree and spread of introgression compromising the genomic integrity of parental populations. To our knowledge, however, no approach has been currently developed to integrate the assessment of Type 2 errors within Bayesian-based clustering techniques and formal estimation of the proportion of admixture. We tackled this issue by adopting a multievent formulation in CR modeling that formally accounts for the uncertainty in detecting hybrids while balancing Type 1 and 2 error rates, and comparing estimates of prevalence obtained with the traditional approach to classify admixed individuals with those obtained by our more precautionary approach (i.e., scenarios A and B, respectively). Scenario A reflected assignment criteria (i.e., definition of threshold q_w values) indicated by Caniglia et al. (2020) to ensure the best performance in distinguishing between recent and older generations of admixture (Caniglia et al. 2020). According to this scenario ($q_w \geq 0.975$), prevalence in our wolf population was 50% and no simulated wolf genotypes were erroneously assigned to the admixed category. However, 14% and 52% of simulated first- and second-generation backcrosses, respectively, were erroneously assigned to the wolf category, indicating that such an approach may still overlook a non-trivial proportion of recently admixed individuals in the population, therefore underestimating prevalence. Also Caniglia et al. (2020) suggested a second q_w threshold of 0.990 to identify admixed individuals of older backcross generations; their approach, however, had different aims than the formal estimation of prevalence because it was developed to provide a convenient and practical screening of samples while standardizing assignment procedures across different genetic laboratories. In our application, and in particular under scenario B, we aimed at formally integrating the uncertainty in hybrid classification into the estimation process. In particular, the Viterbi algorithm offered a reproducible procedure to assign the individuals whose classification was uncertain. In our study, 6 individuals were classified as uncertain and, according to the algorithm, they all had a higher probability of being admixed than wolves. In comparison to Santostasi et al. (2019), who used classification thresholds derived from the literature to illustrate their

modeling procedure, here we used classification thresholds based on genotype simulations specifically calibrated on our study population (Caniglia et al. 2020).

Our estimation process is a 2-step process: first, we estimated parameters through a multievent model, and then we assigned uncertain individuals to a state based on the results of the Viterbi algorithm. In case there is substantial support (i.e., weight or relative probability) for the most likely Viterbi reconstruction (as in our case, Table S7), the uncertainty in individual identification can be dealt with in the first modeling step (i.e., parameter estimation). Should the Viterbi reconstructions reveal higher uncertainty, this can be accounted for also by weighting the prevalence estimate by the probability of each reconstruction (as described in the Supporting Information). Another caveat concerning the assignment of uncertain individuals is that, because this assignment is based on the estimates of the model parameters, it may be flawed if the estimates are themselves biased. Santostasi et al. (2019) assessed the performance of the multievent CR model in estimating parameters (π, φ, p, δ) and prevalence under different sample sizes and concluded that the model estimates are unbiased with the value of detectability reported in our study. To further support the Viterbi algorithm-based assignment, 4 of the 6 individuals originally classified as uncertain were confirmed to be the progeny of 2 admixed individuals by the genealogical reconstruction. This leads us to suggest that, where a large enough sampling coverage would allow, the Viterbi algorithm and the genealogical reconstruction should be used as independent methods to classify individuals that, based on their q_w value, are of uncertain assignment. Expectedly, based on scenario B, the average prevalence of admixture in the wolf population raised to 70%, no simulated wolf was erroneously assigned to the admixed category, no simulated first-generation backcross were erroneously assigned to the wolf cluster, and only 23% of simulated second-generation backcrosses were erroneously assigned to the wolf cluster. Cumulatively, and limited to recent backcrosses, this corresponds to a 65.2% decrease in Type 2 error rate, and no increase in Type 1 error rate, compared to scenario A.

The above estimates of prevalence, ranging from 43.6% (naïve) to 50% (95% CI = 22−80%; scenario A) and 70% (45−77%; scenario B) consistently reveal widespread admixture at the level of nuclear markers of this wolf population, even though the difference in absolute terms reveals that alternative analytical approaches may strongly affect the estimates. Based on theoretical grounds, we maintain that scenario B corresponds to the most reliable estimate in our context. Support is also provided by the genealogy reconstruction; out of the 8 pups produced by the 2 admixed breeding pairs, 7 were identified as admixed under scenario B but only 1 under scenario A. A 70% estimate of prevalence in a protected wolf population may cause concern, especially if this conclusion is based on a relatively small number of genetic markers. Two considerations are in place here. First, because this assessment is based on non-invasive samples genotyping, and hence a limited number of genetic markers, our findings no not imply that 70% of the population has to

be managed (e.g., captured or sterilized) to control the spread of admixture. Ours is a population assessment and by definition (see Introduction) is not meant to provide practical management indications at the individual level. Second, our estimate of prevalence is based on the same genetic protocols and markers currently used to detect hybrids using non-invasive samples (Caniglia et al. 2020). Compared to previous work, it simply extends the estimation approach to correct for imperfect detectability and advances a more formal way to address the uncertainty in assignment of admixed individuals. Our approach can be also applied to estimating the prevalence of admixture in wild canids where K = 3 (e.g., wolf-coyote [*Canis latrans*]-dog in northeastern North America, or wolf-jackal [*Canis aureus*]-dog in central Europe) provided parental reference populations are well-sampled, represented, and reliably distinguished on the basis of diagnostic molecular markers; synthetic genotypes can then be generated through simulations to establish pairwise (e.g., wolf × coyote, coyote × dog, wolf × dog) q_i thresholds using the procedure we illustrated. In the CR framework, the presence of 3 parental populations can be modeled by increasing the number of possible states for parental and admixed individuals. This would quickly increase the number of corresponding parameters leading to the risk of overparametrization (Gimenez et al. 2003); therefore, attention should be paid to adequate sample sizes (Santostasi et al. 2019) or possible ways of simplifying the system (e.g., pooling admixed categories together).

Compared to previous estimates of wolf-dog hybridization (Randi and Lucchini 2002, Lorenzini et al. 2014, Kusak et al. 2018, Dufresnes et al. 2019), our estimate is based on a biologically more meaningful sample because it was limited to a local wolf population and temporally restricted to 1 breeding period. Hybridization is a highly dynamic phenomenon because facilitating factors, such as lack of conspecifics, availability of heterospecific mates, anthropogenic disturbance (Bohling and Waits 2015), and movements of admixed individuals, may vary considerably across time and space. Multi-generation samples pooled across large areas and time frames may impair a clear understanding of admixture dynamics in space and time. On the other hand, our estimate has local significance and cannot be used to extrapolate general conclusions about wolf-dog introgression at wider scales. This underlines the importance of upscaling our sampling and estimation approach over larger areas and multiple timeframes; obtaining a reliable assessment of the number of admixed populations and their prevalence over time will inform the most adequate management strategies (Allendorf et al. 2001).

Regardless of the adopted scenario, the proportion of admixture we estimated is, to our knowledge, the highest so far reported for a wolf population (summarized in Table 1 in Dufresnes et al. [2019]) and indicates that unmanaged wolf-dog hybridization may lead to widespread introgression at the local scale. Based on relatively low, naïve quantifications of prevalence from previous hybridization studies (Verardi et al. 2006, Pacheco et al. 2017, Kusak et al. 2018), several authors concurred on 2 main hypotheses. First, wolf-dog hybridization is rare, and where it occurs it most likely takes place in the peripheral portion of the wolf distribution (Lorenzini et al. 2014) and during early phases of range expansion (Galaverni et al. 2017, Kusak et al. 2018). Second, introgression of dog alleles into wolf populations is expected to be buffered by behavioral and selective constraints (e.g., the unsuccessful integration of pregnant admixed females in the natal packs, the reduced survival of F1 litters due to limited paternal care, the lower success of admixed individuals in territorial or predatory interactions; Vilà and Wayne 1999), or by dilution of dog genes through backcrossing into the parental wolf populations (Verardi et al. 2006). The high proportion of admixture we reported, however, contradicts both predictions and contrasts with theoretical expectations on the functionality of reproductive barriers between wolves and dogs. Despite the high extent of admixture we reported, the absence of dog mtDNA haplotypes and the low occurrence of dog Y-haplotypes in our sample should be of no surprise. In fact, wolf-dog hybridization is highly unidirectional being most often induced by the initial crossing between a female wolf and a male dog, as confirmed by several studies that never reported private dog mtDNA haplotypes within admixed wolf populations, except in a few cases (Hindrikson et al. 2017). Although we did not detect any F1 hybrids, we estimated a high prevalence of backcrosses of several generations backward, 30% of which had non-negligible probabilities of being first- or second-generation backcrosses, indicating a time of original admixture dating as recently as 2006−2009 (i.e., 6−9 years before our sampling assuming a 3-yr wolf generation time; Skoglund et al. 2011). Reports of F1 hybrids are rare both in the northern Apennines (Caniglia et al. 2014) and elsewhere where wolf-dog hybridization has been detected (Godinho et al. 2011, Randi et al. 2014, Pilot et al. 2018, Salvatori et al. 2019). Absence of detection, however, cannot be considered true absence when detectability is not taken into account (MacKenzie 2005). Moreover, even if wolf × dog hybridization events are rare, the production of admixed individuals by backcrossing may increase exponentially when reproductive barriers are weak (1 fertile F1 female can produce several backcrossed litters; Fredrickson and Hedrick 2006, Santostasi et al. 2020). In 2006 a male F1 hybrid, featuring a Y-haplotype typical of Czechoslovakian wolfdogs, was sampled in one of our study packs. This hybrid successively sired 2 admixed litters with a female wolf, as revealed by genealogical reconstruction of 7 pups non-invasively sampled in 2010 and 2011 (R. Caniglia, ISPRA, unpublished data). Therefore, cross-mating between wolves and dogs in the northern Apennines was not limited to the wolf recolonization phase (i.e., the late 1990s; Galaverni et al. 2017), and our findings reveal that factors facilitating hybridization are currently operating in a long-established wolf population well after the recolonization phase.

Anthropogenic disturbances may facilitate the disruption of pre-mating (e.g., agonistic behavior) and post-mating (e.g., reduced fitness of F1) reproductive barriers in social canids (Rutledge et al. 2012, Bohling and Waits 2015).

In particular, human-caused wolf mortality may disrupt breeding pairs, thereby increasing the chances of pack dissolution (Brainerd et al. 2008). Especially during the breeding season, pack dissolution may release the social inhibition to mate of subordinate wolves of both sexes, stimulating individuals to find mates and establish in vacant territories (Bohling and Waits 2015). In human-dominated countries where free-ranging dogs are widespread (Ritchie et al. 2014), or admixed individuals could be available as mates, the above mechanism may contribute to greatly increasing the chances of hybridization and backcrossing (Bohling and Waits 2015). In the same region of our study area, Caniglia et al. (2014) reported a high pack turnover, possibly linked to the high mortality of wolf territory holders. The territories of the packs we surveyed extend well beyond the boundaries of the PNATE, where hunting is allowed throughout the year for wild boar population control, and instances of illegal wolf killings during hunts with dogs are routinely reported among residents (L. Molinari, Appennino Tosco-Emiliano National Park, personal observation). This is in line with our estimates of apparent survival that are considerably lower compared to other studies (Marucco et al. 2009, Cubaynes et al. 2010, Caniglia et al. 2012). According to both scenarios A and B, apparent survival seems to be lower for wolves than for admixed individuals, even though the large variability we obtained about these estimates does not support firm conclusions in our case. Nevertheless, should this difference be confirmed, it would be key in promoting the spread of introgression (Rutledge et al. 2012, Bohling and Waits 2015). We acknowledge that estimates of apparent survival can be also negatively biased by transience generated by dispersing yearlings and subordinate adults (Jimenez et al. 2017), but high disappearance rather than transience rates characterized the wolf population in this region at a wider scale (Caniglia et al. 2014). We are therefore inclined to believe that human-caused mortality is likely among the ultimate causes of the high introgression rates we detected. Anthropogenic food sources (i.e., large livestock carcasses, butchery offals) are largely available in some portions of our study area (L. Molinari, personal observation), and these may promote affiliative interactions between solitary female wolves and dogs (Newsome et al. 2017) and also facilitate the survival of solitary pregnant females and their admixed litters. The occurrence of ≥2 admixed breeding pairs is further evidence of the reproductive success of admixed individuals. In these conditions, hybrid-hybrid pairs can maintain dog genes at high frequency in the population (Bassi et al. 2017, Salvatori et al. 2019), as also projected by simulation of hybridization dynamics in social canids (Fredrickson and Hedrick 2006, Santostasi et al. 2020).

We cannot exclude that high admixture in our study area may have also originated through dispersal of admixed individuals from other areas. Wolves, and expectedly so introgressed individuals, can travel long dispersal distances from their natal territories also across human-dominated countries (Ciucci et al. 2009). In an area located at the periphery of the wolf range in central Italy, 1 admixed breeding pair produced 2−6 pups each year from 2005−2008, and some of the offspring were later sampled at about 40 km from their natal territory (Caniglia et al. 2013). Given that the northern Apennines play an important role as a functional corridor for wolves dispersing to the Alps (Fabbri et al. 2007), a high proportion of admixture in local wolf populations in the northern Apennines may increase the risk of introgression of dog ancestry spreading into the Alpine wolf population, where detection of introgression has been so far been limited (Dufresnes et al. 2019). Notably, 1 of the admixed pups sampled in our study area during our survey was retrieved dead 2 years later in the Italian Prealps, at a linear distance of 237 km along the main dispersal route to the Western Alps (L. Molinari, unpublished data).

MANAGEMENT IMPLICATIONS

In line with other wild-domesticated hybridization cases, our findings indicate that reproductive isolation between wild and domestic forms and dilution by backcrossing may not be sufficient *per se* to prevent widespread introgression of domestic genes. Depending on the extent of admixture, effective management of anthropogenic hybridization could be achieved through preventive (e.g., education, information, communication), proactive (e.g., effective control of hybridization facilitating factors such as free-ranging dogs, poaching, anthropogenic food provisioning), and reactive measures. In our case, proactive interventions should be aimed to enhance the survival of wolf breeders and preserve the social cohesiveness of wolf packs. In presence of widespread admixture, however, reactive interventions should also be considered, aiming at reducing the reproductive contribution of the admixed individuals. In this perspective, population-based assessments such as ours should be supported by a more in-depth assessment of introgression at the individual level. Monitoring the genetic status of populations threatened by anthropogenic hybridization should be conducted more frequently with the aim to detect demes where admixture is originating before it spreads further and jeopardizes the genomic integrity of wild parental populations beyond the reach of practicable management interventions.

ACKNOWLEDGMENTS

LITERATURE CITED

Adams, J. R., B. T. Kelly, and L. P. Waits. 2003. Using faecal DNA sampling and GIS to monitor hybridization between red wolves (*Canis rufus*) and coyotes (*Canis latrans*). Molecular Ecology 12:21752186.

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

Anderson, D. R. 2001. The need to get the basics right in wildlife field studies. Wildlife Society Bulletin 29:1294–1297.

Anderson, E. C., and E. A. Thompson. 2002. A model-based method for identifying species hybrids using multilocus genetic data. Genetics 160:1217–1229.

Anderson, T. M., B. M. vonHoldt, S. L. Candille, M. Musiani, C. Greco, D. R. Stahler, D. W. Smith, B. Padhukasahasram, E. Randi, J. A. Leonard, et al. 2009. Molecular and evolutionary history of melanism in North American gray wolves. Science 323:1339–1343.

Bassi, E., A. Canu, I. Firmo, L. Mattioli, M. Scandura, and M. Apollonio. 2017. Trophic overlap between wolves and free-ranging wolf dog hybrids in the Apennine Mountains, Italy. Global Ecology and Conservation 9:39–49.

Battocchio, D., L. Iacolina, A. Canu, and E. Mori. 2017. How much does it cost to look like a pig in a wild boar group? Behavioural Processes 138:123–126.

Benson, J. F., B. R. Patterson, and T. J. Wheeldon. 2012. Spatial genetic and morphologic structure of wolves and coyotes in relation to environmental heterogeneity in a Canis hybrid zone. Molecular Ecology 21:5934–5954.

Besnier, F., and K. A. Glover. 2013. ParallelStructure: a R package to distribute parallel runs of the population genetics program STRUCTURE on multi-core computers. PLoS ONE 8:1–5.

Bohling, J. H. 2016. Strategies to address the conservation threats posed by hybridization and genetic introgression. Biological Conservation 203:321–327.

Bohling, J. H., J. R. Adams, and L. P. Waits. 2013. Evaluating the ability of Bayesian clustering methods to detect hybridization and introgression using an empirical red wolf data set. Molecular Ecology 22:74–86.

Bohling, J. H., and L. P. Waits. 2015. Factors influencing red wolf-coyote hybridization in eastern North Carolina, USA. Biological Conservation 184:108–116.

Boivin, N. L., M. A. Zeder, D. Q. Fuller, A. Crowther, G. Larson, J. M. Erlandson, T. Denham, and M. D. Petraglia. 2016. Ecological consequences of human niche construction: examining long-term anthropogenic shaping of global species distributions. Proceedings of the National Academy of Sciences of the United States of America 113:6388–6396.

Brainerd, S. M., H. Andrén, E. E. Bangs, E. H. Bradley, J. A. Fontaine, W. Hall, Y. Iliopoulos, M. D. Jimenez, E. A. Jozwiak, O. Liberg, et al. 2008. The effects of breeder loss on wolves. Journal of Wildlife Management 72:89–98.

Burnham, K. P., and D. R. Anderson. 2002. Model selection and inference: a practical information–theoretic approach. Springer–Verlag, New York, New York, USA.

Butler, D. 1994. Bid to protect the wolf from genetic pollution. Nature 370:497.

Caniglia, R., E. Fabbri, S. Cubaynes, O. Gimenez, J.-D. Lebreton, and E. Randi. 2012. An improved procedure to estimate wolf abundance using non-invasive genetic sampling and capture-recapture mixture models. Conservation Genetics 13:53–64.

Caniglia, R., E. Fabbri, M. Galaverni, P. Milanesi, and E. Randi. 2014. Noninvasive sampling and genetic variability, pack structure, and dynamics in an expanding wolf population. Journal of Mammalogy 9:41–59.

Caniglia, R., E. Fabbri, C. Greco, M. Galaverni, L. Manghi, L. Boitani, A. Sforzi, and E. Randi. 2013. Black coats in an admixed wolf dog pack is melanism an indicator of hybridization in wolves? European Journal of Wildlife Research 59:543–555.

Caniglia, R., M. Galaverni, E. Velli, F. Mattucci, A. Canu, M. Apollonio, N. Mucci, M. Scandura, and E. Fabbri. 2020. A standardized approach to empirically define reliable assignment thresholds and appropriate management categories in deeply introgressed populations. Scientific Reports 10:2862.

Chapron, G., P. Kaczensky, J. D. Linnell, M. von Arx, D. Huber, H. Andrén, J. Vicente López-Bao, M. Adamec, F. Álvares, O. Anders, et al. 2014. Recovery of large carnivores in Europe's modern human-dominated landscapes. Science 346:1517–1519.

Choquet, R., L. Rouan, and R. Pradel. 2009. Program E-Surge: a software application for fitting multievent models. Pages 845–865 in D. L. Thomson, E. G. Cooch, and M. J. Conroy, editors. Modeling demographic processes in marked populations. Environmental and ecological statistics series. Volume 3. Springer-Verlag, New York, New York, USA.

Ciucci, P., and L. Boitani. 1999. Nine-year dynamics of a wolf pack in the northern Apennines (Italy). Mammalia 63:377–384.

Ciucci, P., L. Boitani, M. Falco, and L. Maiorano. 2018. Hierarchical, multi-grain rendezvous site selection by wolves in southern Italy. Journal of Wildlife Management 82:1049–1061.

Ciucci, P., L. Boitani, E. Raganella Pelliccioni, M. Rocco, and I. Guy. 1996. A comparison of scat-analysis methods to assess the diet of the wolf Canis lupus. Wildlife Biology 2:37–48.

Ciucci, P., M. Masi, and L. Boitani. 2003. Winter habitat and travel route selection by wolves in the northern Apennines, Italy. Ecography 26:223–235.

Ciucci, P., W. Reggioni, L. Maiorano, and L. Boitani. 2009. Long-distance dispersal of a rescued wolf from the northern Apennines to the western Alps. Journal of Wildlife. Management 73:1300–1306.

Conn, P. B., and E. G. Cooch. 2009. Multistate capture-recapture analysis under imperfect state observation: an application to disease models. Journal of Applied Ecology 46:486–492.

Corander, J., P. Marttinen, J. Sirén, and J. Tang. 2008. Enhanced Bayesian modeling in BAPS software for learning genetic structures of populations. BMC Bioinformatics 9:article number 539.

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

Crespin, L., R. Choquet, M. Lima, J. Merritt, and R. Pradel. 2008. Is heterogeneity of catchability in capture-recapture studies a mere sampling artifact or a biologically relevant feature of the population? Population Ecology 50:247–256.

Cubaynes, S., D. R. MacNulty, D. R. Stahler, K. A. Quimby, D. W. Smith, and T. Coulson. 2014. Density-dependent intraspecific aggression regulates survival in northern Yellowstone wolves (*Canis lupus*). Journal of Animal Ecology 83:1344–1356.

Cubaynes, S., R. Pradel, R. Choquet, C. Duchamp, J.-M. Gaillard, J.-D. Lebreton, E. Marboutin, C. Miquel, A.-M. Reboulet, C. Poillot, et al. 2010. Importance of accounting for detection heterogeneity when estimating abundance: the case of French wolves. Conservation Biology 24:621–626.

Davison, A. C., and D. V. Hinkley. 2002. Bootstrap methods and their application. Cambridge University Press, Cambridge, United Kingdom.

Derégnaucourt, S., J. C. Guyomarc'h, and M. Belhamra. 2004. Comparison of migratory tendency in European Quail Coturnix coturnix, domestic Japanese Quail Coturnix japonica and their hybrids. Ibis 147:25–36.

Donfrancesco, V., P. Ciucci, V. Salvatori, D. Benson, L. W. Andersen, E. Bassi, J. C. Blanco, L. Boitani, R. Caniglia, A. Canu, et al. 2019. Unravelling the scientific debate on how to address wolf-dog hybridization in Europe. Frontiers in Ecology and Evolution 7:1–13.

Dufresnes, C., N. Remollino, C. Stoffel, R. Manz, J. M. Weber, and L. Fumagalli. 2019. Two decades of non-invasive genetic monitoring of the grey wolves recolonizing the Alps support very limited dog introgression. Scientific Reports 9:1–9.

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

Fabbri, E., E. Velli, F. D'Amico, M. Galaverni, L. Mastrogiuseppe, F. Mattucci, and R. Caniglia. 2018. From predation to management: monitoring wolf distribution and understanding depredation patterns from attacks on livestock. HISTRIX 29:101–110.

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

Frantz, L. A. F., V. E. Mullin, M. Pionnier-Capitan, O. Lebrasseur, M. Ollivier, A. Perri, A. Linderholm, V. Mattiangeli, M. D. Teasdale, E. A. Dimopoulos, et al. 2016. Genomic and archaeological evidence suggest a dual origin of domestic dogs. Science 352:1228–1231.

Fredrickson, R. J., and P. W. Hedrick. 2006. Dynamics of hybridization and introgression in red wolves and coyotes. Conservation Biology 20:1272–1283.

Galaverni, M., R. Caniglia, L. Pagani, E. Fabbri, A. Boattini, and E. Randi. 2017. Disentangling timing of admixture, patterns of introgression, and phenotypic indicators in a hybridizing wolf population. Molecular Biology and Evoution 34:2324–2339.

Gese, E. M., and P. A. Terletzky. 2015. Using the 'placeholder' concept to reduce genetic introgression of an endangered carnivore. Biological Conservation 192:11–19.

Gimenez, O., R. Choquet, and J.-D. Lebreton. 2003. Parameter redundancy in multistate capture-recapture models. Biometric Journal 45:704–722.

Gimenez, O., J.-D. Lebreton, R. Choquet, and R. Pradel. 2018. R2ucare: an R package to perform goodness-of-fit tests for capture-recapture models. Methods in Ecology and Evolution 9:1749–1754.

Gimenez, O., J.-D. Lebreton, J. M. Gaillard, R. Choquet, and R. Pradel. 2012. Estimating demographic parameters using hidden process dynamic models. Theoretical Population Biology 82:307–316.

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

Grossen, C., L. Keller, I. Biebac, The International Goat Genome Consortium, and D. Croll. 2014. Introgression from domestic goat generated variation at the major histocompatibility complex of Alpine ibex. PLoS Genetics 10:e1004438.

Hindrikson, M., J. Remm, M. Pilot, R. Godinho, A. V. Stronen, L. Baltrūnaité, S. D. Czarnomska, J. A. Leonard, E. Randi, C. Nowak, et al. 2017. Wolf population genetics in Europe: a systematic review, meta-analysis and suggestions for conservation and management. Biological Reviews 92:1601–629.

Hubisz, M. J., D. Falush, M. Stephens, and J. K. Pritchard. 2009. Inferring weak population structure with the assistance of sample group information. Molecular Ecology Resources 9:1322–1332.

Jakobsson, M., and N. A. Rosenberg. 2007. CLUMPP: a cluster matching and permutation program for dealing with label switching and multimodality in analysis of population structure. Bioinformatics 23:1801–1806.

Jimenez, M. D., E. E. Bangs, D. K. Boyd, D. W. Smith, S. A. Becker, D. E. Ausband, S. P. Woodruff, E. H. Bradley, J. Holyan, and K. Laudon. 2017. Wolf dispersal in the Rocky Mountains, Western United States: 1993–2008. Journal of Wildlife Management 81:581–592.

Kusak, J., V. Caputo, A. Galov, T. Gomerčić, H. Arbanasić, R. Caniglia, M. Galaverni, S. Reljić, D. Huber, and E. Randi. 2018. Wolf-dog hybridization in Croatia. Veterinarski Arhiv 88:375–395.

Lebreton, J.-D., K. Burnham, J. Clobert, and D. Anderson. 1992. Modeling survival and testing biological hypothesis using marked animals: a unified approach with case studies. Ecological Monographs 62:67–118.

Lequarre, A. S., L. Andersson, C. Andre, C. M. Fredholm, C. Hitte, T. Leeb, H. Lohi, K. Lindblad-Toh, and M. Georges. 2011. LUPA: a European initiative taking advantage of the canine genome architecture for unravelling complex disorders in both human and dogs. Veterinary Journal 189:155–159.

Lorenzini, R., R. Fanelli, G. Grifoni, F. Scholl, and R. Fico. 2014. Wolf-dog crossbreeding: 'smelling' a hybrid may not be easy. Mammalian Biology 79:149–156.

Lucchini, V., E. Fabbri, F. Marucco, S. Ricci, L. Boitani, and E. Randi. 2002. Noninvasive molecular tracking of colonizing wolf (Canis lupus) packs in the western Italian Alps. Molecular Ecology 11:857–868.

MacKenzie, D. I. 2005. What are the issues with presence-absence data for wildlife managers? Journal of Wildlife Management 69:849–860.

Marucco, F., D. H. Pletscher, L. Boitani, M. K. Schwartz, K. L. Pilgrim, and J.-D. Lebreton. 2009. Wolf survival and population trend using noninvasive capture-recapture techniques in the Western Alps. Journal of Applied Ecology 46:1003–1010.

McDonald, T. L., and S. C. Amstrup. 2001. Estimation of population size using open capture-recapture models. Journal of Agricultural, Biological and Environmental Statistics 6:206–220.

McGarvey, D. J. 2007. Merging precaution with sound science under the Endangered Species Act. BioScience 57:65–70.

Miller, C. R., J. R. Adams, and L. P. Waits. 2003. Pedigree-based assignment tests for reversing coyote Canis latrans introgression into the wild red wolf Canis rufus population. Molecular Ecology 12:3287–3301.

Miller, C. R., P. Joyce, and L. P. Waits. 2002. Assessing allelic dropout and genotype reliability using maximum likelihood. Genetics 160:357–66.

Newsome, T. M., P. J. S. Fleming, C. R. Dickman, T. S. Doherty, W. J. K. Ripple, E. G. Ritchie, and A. J. Wirsing. 2017. Making a new dog? BioScience 67:374–381.

Nielsen, E. E. G., L. A. Bach, and P. Kotlicki. 2006. Hybridlab (version 1.0): a program for generating simulated hybrids from population samples. Molecular Ecology Notes 6:971–973.

Pacheco, C., J. V. López-Bao, E. J. García, F. J. Lema, L. Llaneza, V. Palacios, and R. Godinho. 2017. Spatial assessment of wolf-dog hybridization in a single breeding period. Scientific Reports 7:1–10.

Pendleton, A.L., F. Shen, A.M. Taravella, S. Emery, K.R. Veeramah, A.R. Boyko, and J.M. Kidd. 2018. Comparison of village dog and wolf genomes highlights the role of the neural crest in dog domestication. BMC Biology 16:article number 64.

Pilot, M., C. Greco, B. M. vonHoldt, E. Randi, W. Jędrzejewski, V. E. Sidorovich, M. K. Konopiński, E. A. Ostrander, and R. K. Wayne. 2018. Widespread, long-term admixture between grey wolves and domestic dogs across Eurasia and its implications for the conservation status of hybrids. Evolutionary Applications 11:662–680.

Piry, S., A. Alapetite, J. M. Cornuet, D. Paetkau, L. Baudouin, and A. Estoup. 2004. GENECLASS 2: a software for genetic assignment and first-generation migrant detection. Journal of Heredity 95:536–539.

Pradel, R. 2005. Multievent: an extension of multistate capture-recapture models to uncertain states. Biometrics 61:442–447.

Pradel, R. 2009. The stakes of capture-recapture models with state uncertainty. Pages 781–795 in D. L. Thomson, E. G. Cooch, and M. J. Conroy, editors. Modeling demographic processes in marked populations. Environmental and ecological statistics series. Volume 3. Springer-Verlag, New York, New York, USA.

Pritchard, J. K., M. Stephens, and P. Donnelly. 2000. Inference of population structure using multilocus genotype data. Genetics 155:945–959.

Randi, E. 2007. Detecting hybridization between wild species and their domesticated relatives. Molecular Ecology 17:285–293.

Randi, E., P. Hulva, E. Fabbri, M. Galaverni, A. Galov, J. Kusak, D. Bigi, B. Černá Bolfíková, M. Smetanová, and R. Caniglia. 2014. Multilocus detection of wolf x dog hybridization in Italy, and guidelines for marker selection. PLoS ONE 9:1–13.

Randi, E., and V. Lucchini. 2002. Detecting rare introgression of domestic dog genes into wild wolf (Canis lupus) populations by Bayesian admixture analyses of microsatellite variation. Conservation Genetics 3:29–43.

Ritchie, E. G., C. R. Letnic, and A. T. Vanak. 2014. Dogs as predators and trophic regulators. Pages 55–68 in M. E. Gompper, editor. Free-ranging dogs and wildlife conservation. Oxford University Press, Oxford, United Kingdom.

Rouan, L., J. M. Gaillard, Y. Guédon, and R. Pradel. 2009. Estimation of lifetime reproductive success when reproductive status cannot always be assessed. Pages 845–865 in D. L. Thompson, E. G. Cooch, and M. J. Conroy, editors. Modeling demographic processes in marked populations. Environmental and ecological statistics series. Volume 3. Springer-Verlag, New York, New York, USA.

Rutledge, L. Y., B. N. White, J. R. Row, and B. R. Patterson. 2012. Intense harvesting of eastern wolves facilitated hybridization with coyotes. Ecology and Evolution 2:19–33.

Saltz, D. 2011. Statistical inference and decision making in conservation biology. Israel Journal of Ecology and Evolution 57:309–317.

Salvatori, V., V. Donfrancesco, A. Trouwborst, L. Boitani, J. D. C. Linnell, F. Alvares, M. Åkesson, V. Balys, J. C. Blanco, S. Chiriac, et al. 2020. European agreements for nature conservation need to explicitly address wolf-dog hybridisation. Biological Conservation 248:108525.

Salvatori, V., R. Godinho, C. Braschi, L. Boitani, and P. Ciucci. 2019. High levels of recent wolf × dog introgressive hybridization in agricultural landscapes of central Italy. European Journal of Wildlife Research 65:73–88.

Santostasi, N. L., P. Ciucci, G. Bearzi, S. Bonizzoni, and O. Gimenez. 2020. Assessing the dynamics of hybridization through a matrix modeling approach. Ecological Modeling 431:109120.

Santostasi, N. L., P. Ciucci, R. Caniglia, E. Fabbri, L. Molinari, W. Reggioni, and O. Gimenez. 2019. Use of hidden Markov capture-recapture models to estimate abundance in the presence of uncertainty: application to the estimation of prevalence of hybrids in animal populations. Ecology and Evolution 9:744–755.

Sanz, N., R. M. Araguas, R. Fernández, M. Vera, and J. L. Garcìa Marìn. 2009. Efficiency of markers and methods for detecting hybrids and introgression in stocked populations. Conservation Genetics 10:225–236.

Skoglund, P., A. Gotherstrom, and M. Jakobsson. 2011. Estimation of population divergence times from non-overlapping genomic sequences: examples from dogs and wolves. Molecular Biology and Evolution 28:1505–1517.

Sundqvist, A.-K., H. Ellegren, M. Olivier, and C. Vilà. 2001. Y chromosome haplotyping in Scandinavian wolves (*Canis lupus*) based on microsatellite markers. Molecular Ecology 10:1959–1966.

Taberlet, P., S. Griffin, B. Goossens, S. Questiau, V. Manceau, N. Escaravage, L. P. Waits, and J. Bouvet. 1996. Reliable genotyping of samples with very low DNA quantities using PCR. Nucleic Acids Research 24:3189–3194.

Todesco, M., M. A. Pascual, G. L. Owens, K. L. Ostevik, B. T. Moyers, S. Hübner, S. M. Heredia, M. A. Hahn, C. Caseys, D. G. Bock, and L. H. Rieseberg. 2016. Hybridization and extinction. Evolutionary Applications 9:892–908.

Vähä, J. P., and C. R. Primmer. 2006. Efficiency of model-based Bayesian methods for detecting hybrid individuals under different hybridization scenarios and with different numbers of loci. Molecular Ecology 15:63–72.

Valière, N. 2002. GIMLET: a computer program for analysing genetic individual identification data. Molecular Ecology Notes 2:377–379.

van Wyk, A. M., D. L. Dalton, S. Hoban, M. W. Bruford, I.-R. M. Russo, C. Birss, P. Grobler, B. Janse van Vuuren, and A. Kotzé. 2016. Quantitative evaluation of hybridization and the impact on biodiversity conservation. Ecology and Evolution 7:320–330.

Vaysse, A., A. Ratnakumar, T. Derrien, E. Axelsson, G. Rosengren Pielberg, S. Sigurdsson, T. Fall, E. H. Seppälä, M. S. T. Hansen, C. T. Lawley, et al. 2011. Identification of genomic regions associated with phenotypic variation between dog breeds using selection mapping. PLoS Genetics 7:e1002316.

Verardi, A., V. Lucchini, and E. Randi. 2006. Detecting introgressive hybridization between free–ranging domestic dogs and wild wolves (*Canis lupus*) by admixture linkage disequilibrium analysis. Molecular Ecology 15:2845–55.

Vilà, C., and R. K. Wayne. 1999. Hybridization between wolves and dogs. Conservation Biology 13:195–198.

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

vonHoldt, B. M., E. Shuldiner, I. J. Koch, R. Y. Kartzinel, A. Hogan, L. Brubaker, S. Wanser, D. Stahler, C. D. L. Wynne, E. A. Ostrander, et al. 2017. Structural variants in genes associated with human Williams-Beuren syndrome underlie stereotypical hypersociability in domestic dogs. Science Advances 3:e1700398.

Wang, J., and A. W. Santure. 2009. Parentage and sibship inference from multi-locus genotype data under polygamy. Genetics 181:1579–1594.

Wayne, R. K., and H. B. Shaffer. 2016. Hybridization and endangered species protection in the molecular era. Molecoular Ecology 25:2680–2689.

Wheeldon, T. J., B. R. Patterson, and B. N. White. 2010. Sympatric wolf and coyote populations of the western Great Lakes region are reproductively isolated. Molecular Ecology 19:4428–4440.

Wilkins, A. S., R. W. Wrangham, and W. Tecumseh Fitch. 2014. The "domestication syndrome" in mammals: a unified explanation based on neural crest cell behavior and genetics. Genetics 197:795–808.

Yoccoz, N. G., J. D. Nichols, and T. Boulinier. 2001. Monitoring of biological diversity in space and time. Trends in Ecology & Evolution 16:446–453.

Zucchini, W., I. L. MacDonald, and R. Langrock. 2016. Hidden Markov models for time series: an introduction using R. Chapman & Hall/CRC Press, Boca Raton, Florida, USA.

Associate Editor: Robin Russell.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at the publisher's website.