

Modelling population dynamics in presence of hybridization

Nina Santostasi

▶ **To cite this version:**

Nina Santostasi. Modelling population dynamics in presence of hybridization. Agricultural sciences. Université Montpellier; Università degli studi La Sapienza (Rome), 2020. English. NNT : 2020MONTG017 . tel-03144074

HAL Id: tel-03144074
https://theses.hal.science/tel-03144074

Submitted on 17 Feb 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



SAPIENZA
Università di Roma
Facoltà di Scienze Matematiche Fisiche e Naturali

PH. D. IN ENVIRONMENTAL AND EVOLUTIONARY BIOLOGY

Cycle: **XXII**
(A.Y. **2019/2020**)

Modelling population dynamics in presence of hybridization

Ph. D. Student
Nina Luisa Santostasi

Supervisors

Prof. Paolo Ciucci
Dept. Biology and Biotechnologies,
University of Rome La Sapienza, Rome, Italy

Prof. Olivier Gimenez
CEFE, CNRS, Univ Montpellier,
Univ Paul Valéry Montpellier 3, EPHE, IRD,
Montpellier, France

Coordinator
Prof. Gabriella Pasqua

THÈSE POUR OBTENIR LE GRADE DE DOCTEUR DE L'UNIVERSITÉ DE MONTPELLIER

En Ecologie et Biodiversité

École doctorale GAIA

Unité de recherche UMR 5175- CEFE- Centre d'Ecologie Fonctionnelle et Evolutive

En partenariat international avec Sapienza Université de Rome, Italie

Modélisation de la dynamique des populations en présence d'hybridation

Présentée par Nina Luisa Santostasi

20 Fèvrier 2020

Sous la direction de Olivier Gimenez et Paolo Ciucci

Devant le jury composé de

John F. Benson, Professeur, University of Nebraska	Rapporteur
Guillame Chapron, Professeur Associé, Swedish University of Agricultural Sciences	Rapporteur
Gentile Francesco Ficetola, Professeur Associé, Università degli Studi di Milano	Examinateur
Emmanuelle Cam, Professeur, Université de Bretagne Occidentale	Examinatrice
Olivier Gimenez, directeur de recherche, Université de Montpellier	Directeur
Paolo Ciucci, Professeur Associé, Sapienza Università di Roma	Directeur
Nolwenn Drouet-Hoguet, chargée de recherche, ONCFS	Invitèe
Diego Rubolini, Professeur Associé, Università degli Studi di Milano	Invitèe

UNIVERSITÉ DE MONTPELLIER

Alla professoressa Morizio, mia madre. Allenatore di maturandi, esaminandi, laureandi e dottorandi. Donna bellissima, gladiatore e generale romano. Le mie bibliografie non saranno mai più le stesse senza di te.

INDEX

1.GENERAL INTRODUCTION……5

2. CHAPTER 1: Development of a multievent capture–recapture model to estimate the prevalence of admixed individuals in wildlife populations.....……..……..22

3.APPENDIX CHAPTER 1…….……...……35

4. CHAPTER 2: Estimating the prevalence of admixed wolf x dog individuals in a study area in the northern Apennines, Italy…….……..55

5.APPENDIX CHAPTER 2.……....105

6. CHAPTER 3: Development of a population-based modelling approach to project hybridization dynamics……...……119

7.APPENDIX CHAPTER 3…….……...……..130

8. CHAPTER 4: Development of an individual-based model to evaluate the effect of different hybridization management strategies for the wolf x dog case study……...….175

9.APPENDIX CHAPTER 4…….…….……...…….215

10.GENERAL CONCLUSIONS……..……....235

1. GENERAL INTRODUCTION

1.1 Natural and anthropogenic hybridization

Hybridization, the interbreeding of individuals from genetically distinct populations, regardless of their taxonomic status (Allendorf et al. 2001) is acknowledged to be a relatively common evolutionary process both in the plant and animal kingdoms (Hewitt, 1988; Olden et al. 2004; Grabenstein & Taylor 2018). Hybridization is mainly observed between otherwise allopatric taxa that come into contact due to natural (natural hybridization) or anthropogenic causes (anthropogenic hybridization, e.g., changes in the abundance and distribution of species, the removal of barriers that cause isolated or restricted species to expand, and/or the uncontrolled diffusion of domestic species). Todesco et al. (2016) performed a literature review of hybridization studies to identify the ecological, evolutionary, and genetic factors that critically affect extinction risk through hybridization. They found that 72% of studies regarding anthropogenic hybridization reported an extinction threat and among those the most common anthropogenic causes of hybridization were husbandry or agriculture (55% of the studies), invasive species (54%), and habitat disturbance (36%). Most of the studies dealt with hybridization in plants, followed by fishes, birds, and mammals. Extinction risk was more common in hybridizing vertebrates than plants and appeared to be driven by fish (85% of the studies) and birds species (79% of the studies).

Anthropogenic hybridization is therefore considered as an increasing and significant threat to biodiversity (Rhymer & Simberloff 1996; Crispo et al. 2011) that has the potential of driving rare taxa to genomic extinction (the loss of combinations of genes and genotypes that have a unique evolutionary history) through two main mechanisms related to the fitness of the hybrids (Allendorf et al. 2001). If the fitness of hybrids is lower than that of parental individuals, hybridization can reinforce reproductive isolation between incompletely isolated species (Barton & Hewitt 1985), but it can also cause extinction through demographic swamping (where population growth rates are reduced due to the wasteful production of maladaptive hybrids; Allendorf et al. 2001; Wolf et al. 2001). If the fitness of the hybrids is greater than or equal to that of parental individuals,

hybridization can cause extinction through genetic swamping (where the rare form is replaced by hybrids; Allendorf et al. 2001). Following genomic extinction, the result is frequently a 'hybrid swarm' in which all the individuals composing a population are hybrids.

1.2 When hybridization is a threat: importance of quantifying the prevalence of admixed individuals in and its dynamics

As with any conservation threat, to evaluate potential management actions, it is fundamental to: 1) understand if hybridization poses a threat to the viability of a population that requires human intervention, 2) identify possible management actions and evaluate their effectiveness in an adaptive management loop (Williams et al. 2011). For the first point, Allendorf et al. (2001) proposed a classification framework to assess the severity of the hybridization threat. They divided anthropogenic hybridization cases in different types based mainly on two criteria: the fitness and relative abundance of admixed individuals (i.e., prevalence of admixed individuals). In the first anthropogenic hybridization type ("hybridization without introgression"), first-generation hybrids are sterile and the main threat for parental populations is the waste of reproductive effort (demographic swamping). In the second type ("widespread introgression"), hybrids are fertile but the parental populations are still prevalent and well distinguished. In the third type ("complete admixture"), widespread introgression can evolve into a situation in which only the admixed individuals remain (genetic swamping). Understanding the type of hybridization that is occuring is therefore fundamental to assess the threat posed by anthropogenic hybridization on the viability of the parental population and timely elaborate management strategies (Allendorf et al. 2001; Bohling 2016). To understand the hybridization type we need to quantify the proportion of admixed individuals in a population at a given time (hereafter prevalence) and understand if such quantity is predicted to increase toward complete admixture.

For the second point, depending on the hybridization type, a variety of management actions can be more effective or feasible to avoid genomic extinction: from removal and/or sterilization of

hybrids to the management of the human disturbances that facilitated hybridization in the first place (Allendorf et al. 2001; Bohling 2016).

1.2.1 Methodological challenges and solutions to estimate the prevalence of admixed individuals

The relative abundance (prevalence) of different categories of individuals within a population is an essential piece of information to understand processes in ecology, evolution, and conservation. For example, the prevalence of infected individuals is critical to understand disease dynamics (Jennelle, et al. 2007) and the prevalence of key demographic categories (e.g., mature females) is needed to predict the viability of populations (Caswell 2000). Estimating the prevalence of different population segments in wildlife populations is not trivial, therefore naive prevalence (the proportion of individuals of a given class in the sample) is often used as a proxy (Jennelle et al. 2007). However, this approach overlooks two main sources of bias: first, imperfect and/or heterogeneous probability of detection leads to biased abundance and prevalence estimates when it is ignored (Jennelle et al. 2007; Cubaynes et al. 2010), second, uncertainty in the classification of individuals is common when individuals are assigned to a specific category based on clues (i.e., the assignment of individuals to the category admixed and parental based on a limited number of genetic markers; Vähä & Primmer 2006).

Reliable estimates of prevalence can be obtained by correcting field counts by the proportion of undetected individuals (i.e., the ratio between the number of observed individuals and the probability of detection; Nichols 1992). The probability of detection is usually estimated by using capture–recapture models (CR) from a sample of individual encounter histories (Otis et al. 1978). In particular, multistate CR models estimate the probability of detection for different categories of individuals by assigning individuals to static or dynamic states. However, while classical CR models assume the correct assignment of all individuals to their state (Lebreton et al. 1992), multievent models relax this assumption by acknowledging the uncertainty in the classification

process in the model structure (Pradel 2005). Multievent models have been used to estimate a variety of population parameters in the presence of uncertainty in classification (e.g., rates of entry and exit from disease states: Conn & Cooch 2009; probability of skipping reproduction: Sanz-Aguilar et al. 2011, and the probability of survival of different age classes: Gervasi et al. 2017). However, multievent models have never been used to estimate the abundance and prevalence of different classes of individuals in a population.

1.2.2 Methods for projecting hybridization dynamics

Once hybridization has been estimated within a period of biological relevance (i.e., reproductive or generational), it is important to understand if such quantity is predicted to increase towards complete admixture. Moreover, if hybridization is recognized as a legitimate threat, the type and intensity of management action need to be chosen and their outcome need to be evaluated (Allendorf et al. 2001; Bohling 2016).

Population projection models are widely used in other contexts (e.g., harvest modelling Sutherland 2001) to simulate population dynamics under different biological/evolutionary scenarios and provide management recommendations (Cross & Beissinger 2001; Hradsky et al. 2019) but their application to assess hybridization as a threat are few, though recently increasing (e.g., Nathan et al. 2016). The first attempts to model hybridization-extinction dynamics had a genetic focus and were based on changes in allelic frequencies at one or more loci (Huxel 1999; Ferdy & Austerlitz 2002). On the other hand, ecological models explicitly examine the effects of life-history traits (e.g., survival and reproductive rates) on the hybridization outcome. Within this approach, two types of models have been used to model hybridization dynamics (Hall & Ayres 2008): 1) individual-based models that simulate the contribution of each individual to the hybridization dynamics of the entire population (e.g., Thompson et al. 2003; Hooftman et al. 2007), and 2) population-based models that can be used when only the mean fitness parameters of the main demographic stages are available (e.g., Wolf et al. 2001; Campbell et al. 2002). Both modelling approaches were applied to

105 hybridization in plant species (Hall & Ayres 2008; Todesco et al. 2016). However, few studies used
106 individual-based models (Fredrickson & Hedrick 2006; Nathan et al. 2019) to simulate
107 hybridization dynamics in animal species.
108
109 *1.3 General scope of the* Ph.D. project
110 The general scope of my Ph.D. was to develop models for the assessment and management of
111 extinction risk due to anthropogenic hybridization. To provide a practical demonstration of the use
112 of the developed models we used them on an emblematic anthropogenic hybridization case study:
113 the hybridization between wolves (*Canis lupus*) and dogs (*Canis lupus familiaris*) in human
114 dominated landscapes.
115
116 *1.3.1 Wolves x dog hybridization in Europe*
117 Wolves and dogs are interfertile and the first-generation hybrids can backcross with both parental
118 sub-species, generating gene flow between the two gene pools (Vilà & Wayne 1999). Hybridization
119 between wolves and dogs occurred repeatedly since dog domestication (15-10.000 years BP; Larson
120 et al. 2012; Pilot et al. 2018), but more recent (i.e., up to 2 generations since admixture)
121 introgressive hybridization has been detected in various wolf populations in Eurasia (Galaverni et
122 al. 2017; Pacheco et al. 2017; Pilot et al. 2018; Dufresnes et al. 2019). In Europe, the expansion of
123 recovering wolf populations through human-dominated landscapes (Chapron et al. 2014), where
124 free-ranging dogs have since long become the most abundant carnivore (Ritchie et al. 2014) creates
125 the conditions for increasing hybridization rates. Although few examples of positive selection of
126 introgressed domesticated alleles have been hypothesized (Anderson et al. 2009), consistent gene
127 flow (i.e., introgression) between domesticated forms and their wild ancestors is expected to have
128 deleterious consequences for wild species genomic integrity (e.g., reduction in fitness and adaptive
129 potential, loss of unique combinations of genes and genotypes that have a unique evolutionary
130 history; Allendorf et al. 2001; Bohling 2016). For these reasons, the assessment of this threat for the

131 genomic integrity of European wolf population is one the conservation priorities for wolves in
132 Europe (Hindrickson et al. 2017).
133 Studies on wolf-dog hybridization in several European countries have been mostly based on
134 non-invasive genetic sampling (reviewed in Dufresnses et al. 2019) and found revealed relatively
135 low proportions of admixed individuals in the collected samples (≤ 7% and ≤ 14% in Western and
136 Eastern Europe, respectively). However, these results may be flawed due to two main reasons. First
137 many of these samples have been collected opportunistically and at a coarse spatial (i.e., national)
138 and temporal (i.e., decades) scales (Dufresnes et al. 2019), whereas hybridization occurs at the local
139 scale and within one generation time. In support to that we highlight that, although large scale
140 assessments failed to detect high prevalence of admixture (Randi & Lucchini, 2002; Verardi et al.
141 2006; Lorenzini et al. 2014) more recent studies, based on intensive sampling at the local scale
142 (Caniglia et al. 2013; Bassi et al. 2017; Salvatori et al. 2019; this study), revealed higher admixture
143 rates even if still based on naive prevalence quantifications.
144 Second, naive prevalence quantifications may suffer from both the heterogeneous
145 detectability of parental and admixed individuals, and the uncertainty in their classification. The
146 systematic monitoring of prevalence in wild populations relies on non-invasive samples (e.g., hairs
147 or scats), therefore the relatively poor-quality DNA extracted from such samples only allows for the
148 amplification of a relatively low number of diagnostic loci (Randi et al. 2014), limiting the power to
149 discriminate between parental individuals and later (>2) generation backcrosses (Vähä & Primmer
150 2006). This generates uncertainty over the classification of some allegedly admixed genotypes (≥
151 3rd generation backcrosses; Vähä & Primmer 2006), that are generally arbitrarily assigned to wolves
152 (e.g., Pacheco et al. 2017; Dufrenses et al. 2019), a practice that avoids Type 1 error (i.e.,
153 misclassifying wolves as admixed individuals) but not necessarily Type 2 error (i.e., misclassifying
154 admixed individuals as wolves).
155 Based on the low proportions of admixed individuals detected in Western European studies
156 (Dufresnes et al. 2019) several authors concurred on two main hypotheses. The first is that wolf x

157 dog hybridization is currently rare and where it occurs, it most likely takes place in the peripheral
158 portion of the wolf distribution (Lorenzini et al. 2014) and/or during phases of range expansion
159 (Galaverni et al. 2017; Kusak et al. 2018). The second is that introgression of dog genes into wolf
160 populations is expected to be buffered by behavioral and selective constraints (e.g., the unsuccessful
161 integration of pregnant females in the natal packs, the reduced survival of first generation hybrid
162 litters due to limited paternal care, the lower success of admixed individuals in territorial or
163 predatory interactions; Vilà & Wayne 1999), and/or by the dilution of dog genes through
164 backcrossing into the parental wolf populations (Verardi et al. 2006). Based on these assumptions
165 and possibly due to the inherent biological and social complexity characterizing control of wolf-dog
166 hybridization (Donfrancesco et al. 2019), active management has lagged behind in Europe,
167 notwithstanding formal guidelines and recommendations provided by the European Union
168 (Salvatori & Ciucci 2018; Salvatori et al. 2019).
169
170 *1.3.2 Wolf x dog hybridization as a threat*
171 Although there was consensus among a panel of 42 wolf x dog hybridization experts across the
172 world regarding the fact that anthropogenic hybridization should be mitigated (Donfrancesco et al.
173 2019), clear evidence has not yet been found that the introgression of dog genes is a real
174 conservation issue.
175 Dogs diverged from the grey wolf *Canis lupus* between 11,000 and 35,000 years ago (Freedman &
176 Wayne, 2017). However, Larson et al. (2012) suggested that the current gene pool of most dog
177 breeds has been formed very recently, not more than 100–150 years ago and the breeds were subjected
178 to extensive selection and most likely passed through narrow bottlenecks. Domesticated mammals
179 in general exhibit a suite of behavioral, physiological, and morphological traits not observed in their
180 wild counterparts. These characteristics are known to include: increased tameness, coat color
181 changes, changes in craniofacial morphology, reductions in tooth size, alterations in ear and tail
182 form, more frequent and non-seasonal estrus cycles, alterations in adrenocorticotropic hormone

183 levels, changed concentrations of several neurotransmitters, prolongations in juvenile behavior, and
184 reductions in both total brain size and of particular brain regions (Wilkins et al. 2014). For dogs
185 several behavioural traits differentiated them from wild canids (Miklosi 2015) and the genetic basis
186 of these traits is being unveiled (vonHoldt et al. 2017). The consequence of the introgression of
187 such traits in a wild wolf population have not yet been addressed.
188 The only case where some of the long term consequences of the introgression of dog alleles
189 in wild wolf population gene pool have been explored is the introgression (and following positive
190 selection) of the melanistic mutation at the K-locus in the wolf gene pool in North American wolves
191 (Anderson et al. 2009). This mutation determines the black color in dogs (Candille et al. 2007)
192 and was likely introduced into North American wolves by ancient hybridization with dogs (12-
193 14,000 years ago; Anderson et al. 2009). Coulson et al. (2011) found that gray wolves had slightly
194 lower survival, recruitment, generation length, and lifetime reproductive success than black
195 heterozygotes for the melanistic mutation. Coulson et al. (2011) also observed that black
196 homozygotes had lower survival, recruitment, generation length, and lifetime reproductive success
197 than black heterozygotes and gray wolves. Stahler et al. (2013) carried a detailed analysis of the
198 life-history traits of the population for both different colors and sexes and found that gray females
199 had more surviving pups than black females, the latter of which were nearly all heterozygotes. It is
200 therefore possible that the beta-defensin gene that determines black color in wolves might have
201 pleiotropic effects on disease resistance or other immunologically related traits (Coulson et al.
202 2011) and result in fitness tradeoffs.
203 Although hybridization with dogs is not only a recent phenomenon and, in this particular case, it
204 seems to have led to potential adaptive variation (together with possible pleiotropic effects), it is
205 reasonable to hypothesize that the forces that shaped introgression of domesticated alleles in North
206 American wild wolf populations 12,000 years ago are much different than the forces that would act
207 on introgressed dog alleles in the Anthropocene. In addition to that, the relative abundance of
208 wolves and dogs changed: following humans evolutionary success, dogs have now become the most

abundant carnivore in the world (Ritchie et al. 2014) while wolves faced almost eradication (Chapron et al. 2014). As a result of conservation efforts, gray wolves are currently re-colonising parts of their former range in Western Europe across human-dominated landscapes (Chapron et al. 2014). Although gray wolves are able to survive and reproduce in anthropogenic habitats, the survival rates and the causes of mortality in these systems will likely differ from those in less human-dominated settings. These habitats are characterized, among other things, by high densities of anthropogenic foods (e.g., livestock, human food waste) that has been linked to changes in dietary habits of predators, which may lead to genetic, behavioural and demographic changes (Newsome et al. 2015). In this context the interaction of anthropogenic selection and introgressed domesticated alleles may contribute to the changes in ecological niche (and role) and evolutionary diversification. The possible outcomes of such process may lead to the evolution of a new dog-like wolf ecotype, or to the gradual replacement of wolves by admixed individuals in human-dominated landscapes (genetic swamping).

Given these considerations and following a precautionary approach, wolf x dog hybridization is largely perceived as a threat by experts of the field (Donfrancesco et al. 2019), a vision that is shared by the author of this manuscript.

1.3.3 Management of wolf x dog hybridization

The management of anthropogenic hybridization involves three types of actions (Donfrancesco et al. 2019): preventive (e.g., community engagement and education to decrease the number of free-ranging dogs), proactive (e.g., removal of free-ranging dogs, and reduction of factors facilitating hybridization such as poaching), and reactive interventions (e.g., the active management of admixed individuals as to reduce their prevalence in the parental wolf population with capture, captivity, sterilization and release, or with lethal removal; Donfrancesco et al. 2019). The three different types are characterized by differences in social acceptance (Donfrancesco et al. 2019) and applicability into the legal framework of European countries (Trouwborst 2014). Moreover preventive and

235 proactive measures are generally enforceable through the national and European legislation, agreed
236 upon and supported by experts and socially more acceptable. On the other hand reactive measures
237 are controversial due to the not-defined legal status of hybrids, the lack of consensus within both the
238 scientific community and the general public (Lorenzini et al. 2014; Donfrancesco et al. 2019). The
239 effectiveness of reactive management has also been questioned by several authors (Lorenzini et al.
240 2014, Pacheco et al. 2017).

241 It is difficult if not impossible to empirically assess the effectiveness of reactive
242 hybridization management strategies due to the general lack of long-term time data series and the
243 need to implement management timely. Demographic simulation models can help exploring
244 complex ecological and evolutionary processes that are not easily empirically measured (Nathan et
245 al. 2019) and can be instrumental to inform decision-making in a context of uncertainty (e.g.,
246 Gervasi & Ciucci 2018). To our knowledge, while population projections have been used to test the
247 effectiveness of the sterilization of admixed breeders is in the intensively monitored red wolf (*Canis*
248 *rufus*) x coyote (*Canis latrans*) system in North Carolina (Fredrickson & Hedrick 2006), they have
249 never been used to compare management strategies in the context of wolf x dog hybridization and
250 in general rarely for animal species (but see Nathan et al. 2019).

251

252 *1.4 Structure and specific objectives of the Ph.D manuscript*

253 This manuscript consists in four chapters, three of which (Chapter 1, 3 and 4) aiming at developing
254 methodological frameworks to quantify the prevalence of admixed individuals and its dynamics.
255 The second chapter is entirely dedicated to the application of the estimation model developed in
256 Chapter 1 to the case study on wolf x dog hybridization in the Northern Apennines, Italy.

257

258 *1.4.1 Chapter 1. Development of a prevalence multievent capture–recapture model to*
259 *estimate prevalence of admixed individuals in wildlife populations*

This chapter has been published: Santostasi, N.L., Ciucci, P., Caniglia, R., Fabbri, E., Molinari, L., Reggioni, W., Gimenez, O. (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.

In this chapter we: 1) developed a multievent capture–recapture model to estimate prevalence in free-ranging populations accounting for imperfect detectability and uncertainty in individual's classification, 2) carried out a simulation study to i) evaluate model performance, ii) compare it to naive quantifications of prevalence and iii) assess the accuracy of model-based estimates of prevalence under different sampling scenarios. The main results from this chapter were that i) the prevalence of hybrids could be estimated while accounting for both detectability and classification uncertainty ii) model-based prevalence consistently had better performance than naive prevalence in the presence of differential detectability and assignment probability and was unbiased for sampling scenarios with high detectability. Our results underline the importance of a model-based approach to obtain unbiased estimates of prevalence of different population segments.

1.4.2 Chapter 2. Estimating the prevalence of admixed wold x dog individuals in a study area in the northern Apennines, Italy

This chapter is currently submitted: Santostasi, N.L., Gimenez, O., Caniglia, R., Fabbri, E., Molinari, L., Reggioni, W. (Submitted). When bad gets worse: unmanaged anthropogenic hybridization between dogs and wolves may lead to widespread introgression. Animal Conservation.

In this chapter we adopted targeted non-invasive genetic sampling and the capture-recapture estimation model developed in Chapter 1 to estimate the prevalence of wolf-dog hybrids in a local, protected wolf population in the northern Apennines, Italy. We discuss the results in the light of previous assessment of prevalence of wolf x dog admixed individuals in Western Europe and we

illustrate the implications of the results for wolf conservation and for the management of wolf x dog hybridization in human-dominated landscapes. In particular, by taking into account imperfect detectability and the uncertainty associated with the assignment of backcrossed admixed individuals, we estimated 64-78% recent hybrids occurring in 6 out of the 7 surveyed packs. Ancestry analysis and genealogy reconstructions also confirmed multi-generational introgression and indicated that some admixed packs had one or both breeders of recent admixture. Our findings underline that in human-modified landscapes wolf-dog hybridization may raise to unexpected levels if left unmanaged, and that reproductive barriers or dilution of dog genes through backcrossing should not be expected, per se, to prevent occurrence and the spread of introgression.

1.4.3 Chapter 3. Development of a population-based modelling approach to project hybridization dynamics

This chapter is currently accepted pending minor revisions: Santostasi, N.L., Ciucci, P., Bearzi, G., Bonizzoni, S., Gimenez, O.(In press). Assessing the dynamics of hybridization through a matrix modelling approach. Ecological Modelling.

In this chapter we present a new matrix population model to project population dynamics of animal populations in presence of hybridization. We apply the model to two real-world case studies of terrestrial (wolf x dog) and marine mammal species (common dolphin *Delphinus delphis* x striped dolphin *Stenella coeruleoalba*). Specifically, we investigate 1) the possible outcomes of wolf x dog hybridization events for an expanding wolf population in Italy, under different reproductive isolation scenarios, 2) the genomic extinction probability of the two interbreeding dolphin species within a semi-enclosed gulf in Greece, under different hybrids' fitness scenarios, 3) the sensitivity of the probability of genomic extinction to the main demographic parameters in the two case studies. In particular our projections highlighted that i) hybridization leads to genomic extinction in the absence of reproductive isolation, ii) rare or depleted species are particularly vulnerable to genomic extinction, iii) genomic extinction depends mainly on demographic

parameters of parental species, iiii) maintaining healthy and abundant populations prevents genomic extinction.

1.4.4 Chapter 4. Development of an individual-based model to evaluate the effect of different management strategies for wolf x dog hybridization

In this chapter we built a detailed individual based model describing the life cycle of the gray wolf by contemplating social dynamics traits linked to hybridization rates. We applied this model to investigate the hybridization dynamics of wolves in a study population the Northern Apennines, Italy, to evaluate the effectiveness of different management scenarios aimed to reduce the abundance of admixed individuals during a ten-generation time. We showed that in presence of frequent immigration of admixed individuals any management action proved ineffective. In presence of immigration by pure wolves all management actions produced a decrease in prevalence, although their relative effectiveness changed depending on the mating choice scenario. In all the simulated scenarios, the impact of hybridization is predicted to extend at broad scales as large numbers of admixed dispersers are produced. Moreover, we identified demographic and social processes that need to be further investigated to more accurately project the outcomes of management alternatives.

LITERATURE CITED

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

Anderson, T.M., vonHoldt, B.M., Candille, S.I., Musiani, M., Greco, C., Stahler, D. R., Smith, D.M., Padhukasahasram,B., Randi,E., Leonard, J.A., Bustamante, C.D., Ostrander, E.A., Tang, H., Wayne, R.K. & Barsh, G.S. (2009). Molecular and evolutionary history of melanism in North American gray wolves. Science 323, 1339-1343.

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

Barton, N.H., Hewitt, G.M. (1989). Adaptation, speciation and hybrid zones. Nature 341, 497-503.

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

Campbell, D.R., Waser, N.M., Pederson, G.T. (2002). Predicting patterns of mating and potential hybridization from pollinator behavior. American Naturalist 159, 438-450.

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

Chapron, G., Kaczensky, P., Linnell, J.D., von Arx, M., Huber, D., Andrén, H., López-Bao, J.V., et al.. (2014).Recovery of large carnivores in Europe's modern human-dominated landscapes. Science 346, 1517- 1519.

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

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

Crispo, E., Moore, J., Lee-Yaw, J.A., Gray, S.M., Haller, B.C. (2011). Broken barriers: human-induced changes to gene flow and introgression in animals. BioEssays 33, 508-518.

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

Donfrancesco, V., Ciucci P., Salvatori V., Benson, D., Andersen, L.W., Bassi, E., 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., Remollino, N., Stoffel, C., Manz, R., Weber, J. M., Fumagalli, L. (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.

Ferdy, J., Austerlitz, F. (2002). Extinction and introgression in a community of partially cross-fertile plant species. American Naturalist 160, 74-86.

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

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

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

395 Gervasi, V., Boitani, L., Paetkau, D., Posillico, M., Randi, E., Ciucci, P. (2017). Estimating survival
396 in the Apennine brown bear accounting for uncertainty in age classification. Population
397 Ecology 59, 119–130.
398
399 Grabenstein, K.C., Taylor, S.A. (2018). Breaking barriers: causes, consequences, and experimental
400 utility of human-mediated hybridization. Trends in Ecology and Evolution 2342, 1-15.
401 Hall, R.J., Ayres D.R. (2008). What can mathematical modeling tell us about hybrid invasions?
402 Biological Invasions 11, 1217–1224.
403
404 Hindrikson, M., Remm, J., Pilot, M., Godinho, R., Stronen, A. V., Baltrūnaité, L., et al., 2017. Wolf
405 population genetics in Europe: A systematic review, meta-analysis and suggestions for
406 conservation and management. Biological Review 92, 1601-629.
407
408 Hooftman, D.A.P., De Jong, M.J., Oostermeijer, J.G.B., den Nijs, H.C.M. (2007). Modelling the
409 long-term consequences of crop-wild relative hybridization: a case study using four
410 generations of hybrids. Journal of Applied Ecology 44, 1035-1045.
411
412 Huxel, G.R. (1999). Rapid displacement of native species by invasive species: effects of
413 hybridization. Biological Conservation 89, 143-152.
414
415 Kusak, J., Caputo, V., Galov, A., Gomerčić, T., Arbanasić, H., Caniglia, R., Galaverni, M., Relji,
416 S., Huber, D., & Randi, E. (2018). Wolf-dog hybridization in Croatia. Veterinarski Arhiv 88,
417 375-395.
418
419 Larson, G., Karlsson, E.K, Perri, A, Webster, M.T., Ho, S.Y.H., Peters, J., Peter W. et al. (2012).
420 Rethinking dog domestication by integrating genetics, archeology, and biogeography.
421 Proceedings of the National Academy of Sciences 109, 8878-8883.
422
423 Lebreton, J., Burnham, K., Clobert, J., Anderson, D. (1992). Modelling survival and testing
424 biologcal hypothesis using marked animals: a unified approach with case studies. Ecological
425 Monographs 62, 67-118.
426
427 Lorenzini, R., Fanelli, R., Grifoni, G., Scholl, F. & Fico, R. (2014). Wolf-dog crossbreeding:
428 ‘Smelling’ a hybrid may not be easy. Mammalian Biology 79, 149-156.
429
430 Miklósi, Á. (2015). Dog behaviour, evolution, and cognition. Oxford University Press, Oxford, UK.
431
432 Nathan, L.R., Mamoozadehb, N., Tumasc, H.R., Gunselmand, S., Klasse, K., Metcalfef, A., Edgeg,
433 C., et al. (2019). A spatially-explicit, individual-based demogenetic simulation framework for
434 evaluating hybridization dynamics. Ecological Modelling 401, 40-51.
435
436 Newsome, T.M., Dellinger, J.A., Pavey, C.R., Ripple, W.J., Shores, C.R., Wirsing, A.J. Dickman,
437 C.R. (2015). The ecological effects of providing resource subsidies to predators. Global
438 Ecology and Biogeography 24, 1–11.
439
440 Nichols, J. D. (1992). Capture-recapture models: using marked animals to study population
441 dynamics. Bio Science, 42, 94–102.
442
443 Olden, J.D., LeRoy Poff, N., Douglas, M.R., Douglas, M.E., Fausch, K.D. (2004). Ecological and
444 evolutionary consequences of biotic homogenization. Trends in Ecology and Evolution 19,
445 18-24.

Otis, D. L., Burnham, K. P., White, G. C., & Anderson, D. R. (1978). Statistical inference from capture data on closed animal populations. Wildlife Monographs, 62, 3–135.

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

Pilot, M., Greco, C., vonHoldt, B.M., Randi, E., Jędrzejewski, W., Sidorovich, V.E., Konopiński, M.K., Ostrander, E.A. & Wayne, R.K. (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.

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

Randi, E., Hulva, P., Fabbri, E., Galaverni, M., Galov, A., Kusak, J., Bigi, D., Bolfikova, B.C., Smetanova, M. & Caniglia, R. (2014). Multilocus detection of wolf x dog hybridization in italy, and guidelines for marker selection. PLoS ONE 9. 1-13.

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

Rhymer, J.M., Simberloff, D., 1996. Extinction by hybridization and introgression. Annual Review of Ecology and Systematics 27, 83–109.

Ritchie. E.G., Letnic, C.R. & Vanak, A.T. (2014). Dogs as predators and trophic regulators, in: Gompper M. E (Ed.). In Free-Ranging Dogs and Wildlife Conservation: 55-68. Oxford: Oxford UnivPress.

Salvatori, V., Ciucci, P. (2018). Wolf-dog hybridization: issues on detection and management across Europe. In Wolf-human coexistence in the alps and in europe. Proceedings of the international final conference of the LIFEWOLFALPS project, Trento, 19-20 March 2018.

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

Sanz-Aguilar, A., Tavecchia, G., Genovart, M., Igual, J. M., Oro, D., Rouan, L., Pradel, R. (2011). Studying the reproductive skipping behavior in long-lived birds by adding nest inspection to individual-based data. Ecological Applications, 21, 555-564.

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

Thompson, C.J., Thompson, B.J.P., Ades, P.K., Cousens, R., Garnier-Gere, P., Landman, K., Newbigin, E., et al. (2003). Model-based analysis of the likelihood of gene introgression from genetically modified crops into wild relatives. Ecological Modelling 162, 199-209.

496 Sutherland, W.J. (2001). Sustainable exploitation: a review of principles and methods. Wildlife
497 Biology 7, 131-140.
498
499 Todesco, M., Pascual, M.A., Owens, G.L., Ostevik, K.L., Moyers, B.T., Hübner, S., Heredia, S.M.,
500 et al. (2016). Hybridization and extinction. Evolutionary Applications 9, 892-908.
501
502 Trouwborst, A. (2014). Exploring the legal status of Wolf-dog hybrids and other dubious animals:
503 International and EU law and the wildlife conservation problem of hybridization with
504 domestic and alien species. Review of European Comparative and International Environ.
505 Law. 23, 111-124.
506
507 Vähä, J.P. & Primmer, C.R. (2006). Efficiency of model-based Bayesian methods for detecting
508 hybrid individuals under different hybridization scenarios and with different numbers of loci.
509 Molecular Ecology 15, 63-72.
510
511 Verardi, A., Lucchini, V. & Randi, E. (2006). Detecting introgressive hybridization between free-
512 ranging domestic dogs and wild wolves (Canis lupus) by admixture linkage disequilibrium
513 analysis. Molecular Ecology 15, 2845-55.
514
515 Vilà, C. & Wayne, R.K. (1999). Hybridization between wolves and dogs. Conservation Biology 13,
516 195-198.
517
518 vonHoldt, B.M., Shuldiner, E., Koch, I.J., Kartzinel, R.Y., Hogan, A., Brubaker, L., et al., (2017).
519 Structural variants in genes associated with human Williams-Beuren syndrome underlie
520 stereotypical hypersociability in domestic dogs. Science Advances.
521
522 Wilkins, A.S., Wrangham R.W., Tecumseh Fitch, W. (2014). The “Domestication Syndrome” in
523 Mammals: A Unified Explanation Based on Neural Crest Cell Behavior and Genetics.
524 Genetics 197, 795-808.
525
526 Williams, B.K. (2011) Adaptive management of natural resources—framework and issues,
527 Journal of Environmental Management, 92, 1346-1353.
528
529 Wolf, D.E., Takebayashi, N., Rieseberg, L.H. (2001). Predicting the risk of extinction through
530 hybridization. Conservation Biology 15, 1039-1053.
531

532

CHAPTER 1

Received: 3 September 2018 | Accepted: 5 September 2018

DOI: 10.1002/ece3.4819

ORIGINAL RESEARCH

WILEY Ecology and Evolution Open Access

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

Nina Luisa Santostasi[1,2] | Paolo Ciucci[1] | Romolo Caniglia[3] | Elena Fabbri[3] | Luigi Molinari[4] | Willy Reggioni[4] | Olivier Gimenez[2]

[1]Department of Biology and Biotechnologies "Charles Darwin", University of Rome La Sapienza, Rome, Italy

[2]CEFE, CNRS, University of Montpellier, University Paul Valéry Montpellier 3, EPHE, IRD, Montpellier, France

[3]Italian Institute for Environmental Protection and Research (ISPRA), Unit for Conservation Genetics (BIO-CGE), Ozzano dell'Emilia, Bologna, Italy

[4]Wolf Apennine Center, Appennino Tosco-Emiliano National Park, Ligonchio, Italy

Correspondence
Nina Luisa Santostasi, Department of Biology and Biotechnologies "Charles Darwin", University of Rome La Sapienza, Rome, Italy.
Email: n.santostasi@gmail.com

Funding information
French National Research Agency, Grant/Award Number: ANR-16-CE02-0007.

Abstract

Estimating the relative abundance (prevalence) of different population segments is a key step in addressing fundamental research questions in ecology, evolution, and conservation. The raw percentage of individuals in the sample (naive prevalence) is generally used for this purpose, but it is likely to be subject to two main sources of bias. First, the detectability of individuals is ignored; second, classification errors may occur due to some inherent limits of the diagnostic methods. We developed a hidden Markov (also known as multievent) capture–recapture model to estimate prevalence in free-ranging populations accounting for imperfect detectability and uncertainty in individual's classification. We carried out a simulation study to compare naive and model-based estimates of prevalence and assess the performance of our model under different sampling scenarios. We then illustrate our method with a real-world case study of estimating the prevalence of wolf (*Canis lupus*) and dog (*Canis lupus familiaris*) hybrids in a wolf population in northern Italy. We showed that the prevalence of hybrids could be estimated while accounting for both detectability and classification uncertainty. Model-based prevalence consistently had better performance than naive prevalence in the presence of differential detectability and assignment probability and was unbiased for sampling scenarios with high detectability. We also showed that ignoring detectability and uncertainty in the wolf case study would lead to underestimating the prevalence of hybrids. Our results underline the importance of a model-based approach to obtain unbiased estimates of prevalence of different population segments. Our model can be adapted to any taxa, and it can be used to estimate absolute abundance and prevalence in a variety of cases involving imperfect detection and uncertainty in classification of individuals (e.g., sex ratio, proportion of breeders, and prevalence of infected individuals).

KEYWORDS

anthropogenic introgression, capture–recapture, hidden Markov models, hybridization, multievent models, prevalence, Viterbi algorithm

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2019 The Authors. *Ecology and Evolution* published by John Wiley & Sons Ltd.

1 | INTRODUCTION

The relative abundance (prevalence) of different population segments is a fundamental piece of information to understand processes in ecology, evolution, and conservation. For example, the prevalence of infected individuals is critical to understand the mechanisms driving disease dynamics (Jennelle, Cooch, Conroy, & Senar, 2007; Moreno-Torres, Wolfe, Saville, & Garabed, 2016); the prevalence of key demographic categories (e.g., mature females) is needed to assess the viability of endangered populations (Caswell, 2000) and when hybridization represents a threat, the prevalence of admixed individuals is needed to evaluate the appropriate management option (Allendorf, Leary, Spruell, & Wenburg, 2001).

However, estimating prevalence for wildlife populations is challenging and the raw percentage of individuals of a class in the sample (naive prevalence) is often used as a proxy (Jennelle et al., 2007). However, this approach overlooks two main sources of bias. First, imperfect and/or heterogeneous detection leads to biased abundance estimates when it is ignored (Cubaynes et al., 2010; Jennelle et al., 2007). Second, uncertainty in the classification of individuals (e.g., diseased/healthy, breeder/nonbreeder, male/female) is common in wildlife population studies where individuals are assigned to a specific status based on imperfect clues. Examples include determining sex or breeding status based on the behavior of individuals (Genovart, Pradel, & Oro, 2012) or establishing health status from the observation of outer symptoms only (Conn & Cooch, 2009). Another less explored but intriguing situation is assigning individuals to genetic classes (subpopulations) based on a limited number of genetic markers (Vähä & Primmer, 2006).

Reliable estimates of wildlife abundance can be obtained by correcting field counts by the proportion of undetected individuals (i.e., the ratio between the number of observed individuals and the probability of detection; Nichols, 1992). The probability of detection is usually estimated by using capture–recapture models (CR) from a sample of individual encounter histories (Otis, Burnham, White, & Anderson, 1978). In particular, multistate CR models estimate the probability of detection for different classes of individuals by assigning individuals to static or dynamic states. However, multistate CR models assume the correct assignment of all individuals to their state (Lebreton, Burnham, Clobert, & Anderson, 1992). Multievent models relax this assumption by acknowledging the uncertainty of the observation process in the model structure (Pradel, 2005). In these models, a hidden biological process (e.g., survival or dispersal) is modeled as a Markov chain of states (Pradel, 2005). The observation process (the data) arises from the underlying states through the probability of detection. To include uncertainty in state assignment, the observation process is further split into two steps: detection and state assignment conditional on detection (Gimenez, Lebreton, Gaillard, Choquet, & Pradel, 2012; Pradel, 2005). This formulation includes a probability of assignment (besides the probability of detection and the probabilities associated with the biological process), allowing for the inclusion of individuals classified with uncertainty (Pradel, 2005).

Multievent models have been used to estimate a variety of population parameters in the presence of uncertainty in state assignment. Examples of that include the rates of entry and exit from disease states (Conn & Cooch, 2009), the probability of skipping reproduction (Sanz-Aguilar et al., 2011), and the probability of survival of different age classes (Gervasi et al., 2017). However, multievent models have never been used to estimate the abundance of individuals in different states because the numerator of the abundance estimator (the number of observed individuals) is contaminated by uncertain observations.

Here, we develop a capture–recapture approach to estimate the prevalence of admixed individuals (hereafter "hybrids") in a population while simultaneously accounting for both imperfect detection and classification uncertainty. Specifically, we show how to use the multievent CR framework to estimate abundance of individuals in different states (i.e., "Parental," "Hybrid," "Dead") in the presence of uncertainty in state assignment. We first use multievent models to estimate survival and detection parameters; second, we use the Viterbi algorithm to assign the uncertain observed individuals to the most likely state (Rouan, Gaillard, Guédon, & Pradel, 2009; Zucchini, MacDonald, & Langrock, 2009), and lastly, we estimate prevalence via a Horvitz–Thompson-like estimator combined with a bootstrapping procedure to produce standard error and confidence intervals (Davison & Hinkley, 1997).

We assess the importance of incorporating detectability and uncertainty in state assignment by comparing the performance of model-based and naive prevalence under different scenarios. The accuracy of CR parameters' estimators depends on the recapture rate and on the number of capture occasions (Otis et al., 1978). Increasing the detectability and/or the number of occasions requires different sampling strategies and generates different costs in terms of financial and human resources (Lieury et al., 2017). Therefore, we also explore how different sampling strategies may affect the model performance.

Despite the increasing attention that researchers are devoting to hybridization cases (Schwenk, Brede, & Streit, 2008; Todesco et al., 2016), there have been only few attempts to estimate prevalence of hybrids in wild populations (Vaz Pinto, Beja, Ferrand, & Godinho, 2015). We illustrate our method with a case study by estimating the prevalence of wolf (*Canis lupus*) × dog (*Canis lupus familiaris*) hybrids in a wolf population in northern Italy. This is a case of anthropogenic hybridization (i.e., the interbreeding of individuals from genetically distinct populations due to human action; Allendorf et al., 2001) and is considered a major threat to wolf genomic integrity (Boitani, 2000). Therefore, accurately estimating the prevalence of hybrids in wolf populations is a priority to elaborate conservation strategies (Hindrickson et al., 2017). We show that in this case using naive prevalence as a proxy underestimates the prevalence of hybrids.

2 | METHODS

2.1 | Hidden Markov model

We assumed that animals are individually recognized at discrete encounter occasions, therefore obtaining an encounter history for each individual. Individuals can be in one of three possible nonobservable

states: alive and parental (P), alive and hybrid (H), or dead (D). We underline that hereafter the term "hybrid" refers to all categories of admixture and not only to first-generation hybrids. Upon its first encounter, an individual has a probability π_p to be a parental and the complementary probability $\pi_h = 1 - \pi_p$ to be a hybrid. The initial state probabilities describe the probability that an individual is in one or another state when first encountered. Then, the states change over time according to a first-order Markov process, with the state process being governed by apparent survival probabilities φ_p and φ_h. More specifically, the state process, which summarizes the underlying biological process, is represented by a transition probability matrix with states at time t in rows ("Parental," "Hybrid," "Dead") and $t + 1$ in columns ("Parental," "Hybrid," "Dead"):

$$\begin{array}{c} \\ \boldsymbol{P} \\ \boldsymbol{H} \\ \boldsymbol{D} \end{array} \begin{array}{c} \begin{array}{ccc} \boldsymbol{P} & \boldsymbol{H} & \boldsymbol{D} \end{array} \\ \begin{pmatrix} \varphi_p & 0 & 1-\varphi_p \\ 0 & \varphi_h & 1-\varphi_h \\ 0 & 0 & 1 \end{pmatrix} \end{array}$$

where parameter φ_p (resp. φ_h) is the probability that an individual alive and in state "Parental" (resp. "Hybrid") at time t is still alive in the study area and in state "Parental" (resp. "Hybrid") at time $t + 1$ and corresponds to the apparent survival probability of parental (resp. "Hybrid") individuals.

The second time series (or event process) is generated from the states at each occasion and describes the observation process. Individuals are detected at time t with probability of detection p_p for parental and p_h for hybrid individuals. Upon detection, an attempt is made of assigning the individuals to a hybridization state based on genetic and/or morphologic diagnostic features and there is a probability δ that an individual is assigned to the state "Parental" or "Hybrid". If the diagnostic clues are not sufficient to ascertain the hybridization state, the individual is classified as uncertain with the complementary probability $1-\delta$. The observation process is summarized by a matrix with states in rows ("Parental," "Hybrid," "Dead") and events in columns (0 = "Not detected," 1 = "Detected as parental," 2 = "Detected as hybrid," 3 = "Detected as uncertain"):

$$\begin{array}{c} \\ \boldsymbol{P} \\ \boldsymbol{H} \\ \boldsymbol{D} \end{array} \begin{array}{c} \begin{array}{cccc} \boldsymbol{0} & \boldsymbol{1} & \boldsymbol{2} & \boldsymbol{3} \end{array} \\ \begin{pmatrix} 1-p_p & p_p\delta_p & 0 & p_p(1-\delta_p) \\ 1-p_h & 0 & p_h\delta_h & p_h(1-\delta_h) \\ 1 & 0 & 0 & 0 \end{pmatrix} \end{array}$$

In the first row, the $(1 - p_p)$ term is the probability that an individual in state "Parental" is not detected, and $p_h\delta_h$ is the probability that an individual in state "Parental" is detected and assigned to the category "Parental" while $p_p(1 - \delta_p)$ is the probability that an individual in state "Parental" is detected (p_p) and not assigned to any category ($1 - \delta_p$). Note that an individual in state "Parental" cannot be detected as a hybrid; hence, the corresponding probability is 0. The second row is similar to the first one, except that it refers to the hybrid individuals. In the third row, all individuals are nondetected because they are dead. Equivalently, the event process can be decomposed as the product of a detection matrix by an assignment matrix which expresses the probability that an individual is assigned to a state given that it has been detected:

$$\begin{array}{c} \\ \boldsymbol{P} \\ \boldsymbol{H} \\ \boldsymbol{D} \end{array} \begin{array}{c} \begin{array}{ccc} \boldsymbol{0} & \boldsymbol{1} & \boldsymbol{2} \end{array} \\ \begin{pmatrix} 1-p_p & p_p & 0 \\ 1-p_h & 0 & p_h \\ 1 & 0 & 0 \end{pmatrix} \end{array} \begin{array}{c} \begin{array}{cccc} \boldsymbol{0} & \boldsymbol{1} & \boldsymbol{2} & \boldsymbol{3} \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \delta_p & 0 & 1-\delta_p \\ 0 & 0 & \delta_h & 1-\delta_h \end{pmatrix} \end{array}$$

when an animal is first encountered, the capture process is not modeled because an animal must be encountered at least once to enter the dataset, but the state ascertainment remains valid.

To illustrate the calculation of an encounter history, let us consider the case of a 3-year CR experiment. For instance, the encounter history "303" denotes an individual encountered at the first and third occasions but not at the second occasion. The state of this individual is never assigned. Assuming parameters are constant, we have

$$\begin{aligned} \Pr(303) = {} & \pi_h(1-\delta_h)\varphi_h(1-p_h)\varphi_h p_h(1-\delta_h) \\ & + \pi_p(1-\delta_p)\varphi_p(1-p_p)\varphi_p p_p(1-\delta_p) \end{aligned}$$

On the right side of the equation, the first element of the sum is the probability of the observed history if the underlying state is "Hybrid", while the second element is the probability of the observed history given that the underlying state is "Parental". The likelihood of the entire dataset is obtained as the product of the probabilities of all individual encounter histories assuming independence. In this paper, estimates of initial states (π_p and π_h), apparent survival (φ_p and φ_h), detection (p_p and p_h), and assignment probabilities (δ_p and δ_h) are obtained by maximizing the likelihood function. Different models can be built by allowing the parameters to vary according to, for example, state, time, or group of individuals classified based on discrete variables (e.g., age class or sex).

2.2 | Abundance and prevalence estimation

Naive prevalence (P_{naive}) at a given occasion is calculated as the proportion of observed hybrids in the sample:

$$P_{\text{naive}} = \frac{n_h}{n_p + n_h}$$

where n_h is the number of observed hybrids and n_p is the number of observed parentals.

Here, we propose a model-based prevalence estimate (P_{model}) as the ratio between hybrids' abundance in the population and total population abundance. Assuming that marked and unmarked individuals have the same probability of detection, total abundance at a given occasion ($\hat{N}_{\text{tot}}$) is then estimated with the Horvitz–Thompson estimator (McDonald & Amstrup, 2001) as:

$$\hat{N}_{\text{tot}} = \frac{n_p}{\hat{p}_p} + \frac{n_h}{\hat{p}_h} = \hat{N}_p + \hat{N}_h$$

where n_h is the number of observed hybrids, and n_p is the number of observed parental individuals, $\hat{p}_h$ is the recapture probability of hybrids, and $\hat{p}_p$ is the recapture probability of parental individuals. Model-based prevalence is then estimated as:

$$P_{\text{model}} = \frac{\hat{N}_h}{\hat{N}_p + \hat{N}_h}$$

The main difficulty is therefore in determining the number of observed parental individuals and hybrids, because the uncertain individuals have to be assigned to one of the two states ("Hybrid" or "Parental"). To do so, we use the Viterbi algorithm (Zucchini et al., 2009) which, given any observation sequence (here, the encounter histories) and the parameters estimated by the hidden Markov model, finds the most probable underlying sequence of states that has generated the observed data (Rouan et al., 2009). Once the uncertain observations are assigned to the most likely state, the number of observed hybrid and parental individuals can be reconstructed and the formula can be used to estimate prevalence. We obtain confidence intervals for the hybrids' prevalence by using a nonparametric bootstrap procedure (Davison & Hinkley, 1997).

2.3 | Evaluation of model performance and sampling strategy

To test the model performance, we generated encounter histories with known prevalence mimicking our case study on wolves and dogs in the Northern Apennines (see next section) and compared model-based prevalence and naive prevalence. Using the R (R Core Team, 2017) package HMM (Himmelmann, 2010), we simulated a cohort of 100 individuals that we split into 2 states "Hybrid" and "Parental". For all the scenarios, we set the initial proportion of wolves $\pi_p = 0.7$ and initial proportion of hybrids $\pi_h = 1 - \pi_p = 0.3$ as in our case study sample (see next section). Using values estimated by Caniglia et al. (2012) for the same wolf population, we considered state-dependent survival (constant over time) with parental survival ($\varphi_p = 0.8$) higher than hybrid survival ($\varphi_h = 0.7$). We then considered three hypothetical scenarios for detectability and assignment probability: (see Supporting information Table S1 for a complete list of parameters for the three scenarios) (1) state-dependent detectability ($p_p > p_h$) and homogeneous assignment probability, (2) homogeneous detectability and state-dependent assignment probability ($\delta_p > \delta_h$) and (3) homogeneous detectability and assignment probability. Within these three main scenarios, we evaluated the effect of lower and higher sampling intensities by comparing sampling strategies with low and high detectability and with 5 and 10 capture occasions. We simulated 100 datasets for each combination of parameters within the three main scenarios. We fitted constant and state-dependent models to the simulated data, and we calculated the relative bias and the root mean squared error (RMSE) for the model-based and the naive prevalence estimators. We also calculated the confidence interval coverage for the model-based estimator of prevalence.

2.4 | Case study

We collected fresh wolf scats in the Appennino Tosco-emiliano National Park from August 2016 to May 2017. We extracted, amplified and sequenced DNA from the scats following the procedures described in Caniglia, Fabbri, Galaverni, Milanesi, and Randi (2014). We identified wolves, dogs and putative hybrids based on the analysis of molecular markers listed in Randi et al. (2014) and using Bayesian genetic clustering procedures implemented in STRUCTURE 2.3.4 (Falush, Stephens, & Pritchard, 2003; Pritchard, Stephens, & Donnelly, 2000). We distinguished wolves, hybrids, and uncertain individuals based on their membership proportions to the wolf cluster (q_{wolf}) and 90% Bayesian credible intervals (BCI). We set the thresholds for the three categories based on the genetic clustering analyses performed on simulated genotypes by Pacheco et al. (2017), see Fig. S1 in the Supplementary Materials of Pacheco et al. (2017). We considered as pure wolves those individuals whose q_{wolf} was included in the range of q_{wolf} values of simulated pure wolves genotypes and did not overlap with that of simulated backcrosses (hybrid × parental). We classified as uncertain those individuals whose q_{wolf} was included in the range in which the q_{wolf} values of simulated pure wolves and backcrosses to wolves (first and second generation) overlapped. Finally, we considered as hybrids those individuals whose q_{wolf} overlapped with the range of q_{wolf} values of simulated backcrosses (first and second generation) and/or hybrids (first and second generation). We additionally considered as hybrids those individuals which presented a Y haplotype of canine origin or the deletion at the K-locus (Caniglia et al., 2013; Randi et al., 2014) regardless of their q_{wolf} values.

The CR data were pooled in 2-month capture occasions, with a total of five capture occasions. Thirty-nine individuals were sampled (19 wolves, 12 hybrids, 8 uncertains based on their q_{wolf} value). We applied the multievent CR model described above to test for differences in detectability and assignment probability between hybrid and parental individuals and to estimate prevalence of wolf × dog hybrids in the population. Since the hybridization assessment is performed only once for each genotype, we constrained the assignment probability to be estimated only upon first capture (see the models details in the Supporting information). We fitted models with state-dependent and constant parameters and a combination of the two. The models were fitted in the E-SURGE software (Choquet, Rouan, & Pradel, 2009) and the Viterbi algorithm was implemented using the R (R Core Team, 2017) package HMM (Himmelmann, 2010). We used the Akaike Information Criterion (AIC) for model selection, considering models within ΔAICc ≤ 2 as the most supported and used model averaging to account for uncertainty in model selection (Burnham & Anderson, 2002).

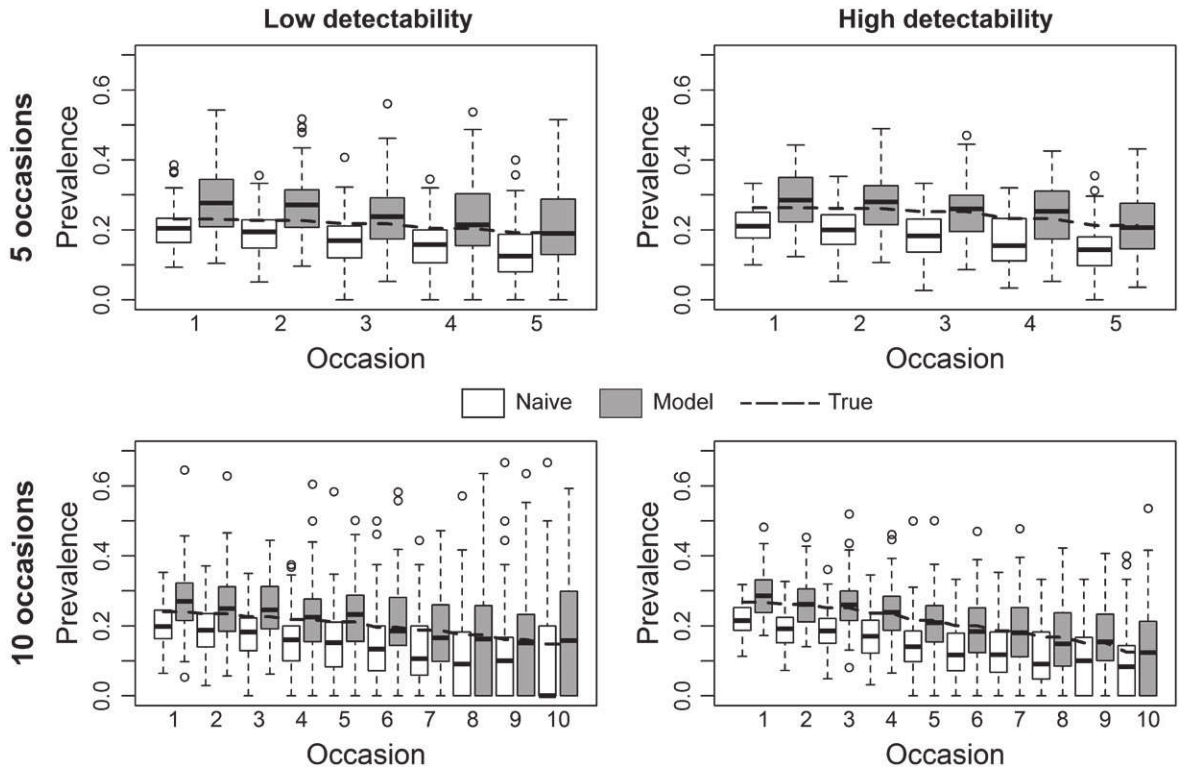


FIGURE 1 Scenario 1 of simulation (state-dependent detectability and homogeneous assignment probability). Sampling strategies with 5 (upper panels) versus 10 (lower panels) capture occasions and low (left-column panels) versus high (right-column panel) detectability. True prevalence is represented as a dashed line while the 100 values of naive and model-based prevalence are displayed in the white and gray boxplots, respectively

3 | RESULTS

3.1 | Evaluation of model performance

Naive prevalence had higher RMSE and percent relative bias than model-based prevalence in scenario 1 (state-dependent detectability and homogeneous assignment probability; Figure 1, Tables 1,2). The same occurred in scenario 2 (homogeneous detectability and state-dependent assignment probability; Supporting information Figure S1 and Tables S2, S3). Naive and model-based prevalence had similar RMSE and relative bias only in scenario 3 (homogeneous detectability and assignment probability; Supporting information Figure S1 and Tables S4, S5).

The bias associated with model-based prevalence tended to 0 in scenarios with high detectability (Tables 1,2; Supporting information Tables S2, S5). Interestingly, the bias associated with naive prevalence had the opposite behavior, as it increased at higher detectability (Tables 1,2; Supporting information Tables S2–S5). The negative bias in naive prevalence is due to the simulation settings for the true detection and assignment probabilities, and in particular to the fact that $p_p > p_h$ for scenario 1 and $\delta_p > \delta_h$ for scenario 2. Switching the true values for detectability in scenario 1 and for assignment probability in scenario 2 would cause naive prevalence to be positively-biased. The bias associated to the estimates of apparent survival, detectability, and probability of assignment was negligible for all scenarios (Figures 2–4; Supporting information Figures S3–S9), while the estimates of the initial state probability were slightly biased for scenario 1 (Figure 5). Estimates of parameters for the category with fewer individuals (in this case, the hybrids) were less precise, as

TABLE 1 Scenario 1 of simulation (state-dependent detectability and homogeneous assignment probability)

	Occ. 1	Occ. 2	Occ. 3	Occ. 4	Occ. 5
Root mean squared error					
Low detectability					
Naive	0.07	0.14	0.23	0.23	0.30
Model-based	0.30	0.15	0.06	0.07	0.02
High detectability					
Naive	0.26	0.36	0.48	0.41	0.48
Model-based	0.04	0.02	0.00	0.01	0.00
Percent relative bias					
Low detectability					
Naive	−0.03	−0.04	−0.05	−0.05	−0.05
Model-based	0.05	0.04	0.03	0.03	0.01
High detectability					
Naive	−0.05	−0.06	−0.07	−0.06	−0.07
Model-based	0.02	0.01	0.00	0.01	0.00

Note. Root mean squared error and relative bias of naive and model-based prevalence for sampling strategies with 5 capture occasions (Occ).

TABLE 2 Scenario 1 of simulation (state-dependent detectability and homogeneous assignment probability)

	Occ. 1	Occ. 2	Occ. 3	Occ. 4	Occ. 5	Occ. 6	Occ. 7	Occ. 8	Occ. 9	Occ. 10
Root mean squared error										
Low detectability										
Naive	0.14	0.22	0.21	0.33	0.33	0.30	0.40	0.36	0.26	0.78
Model-based	0.1	0.04	0.05	0.01	0.01	0.00	0.00	0.00	0.02	0.02
High detectability										
Naive	0.23	0.51	0.37	0.45	0.49	0.50	0.39	0.36	0.16	0.16
Model-based	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Percent relative bias										
Low detectability										
Naive	−0.04	−0.05	−0.05	−0.06	−0.06	−0.05	−0.06	−0.06	−0.05	−0.03
Model-based	−0.03	−0.02	−0.02	−0.01	−0.01	−0.00	−0.00	−0.01	−0.01	−0.01
High detectability										
Naive	−0.05	−0.07	−0.06	−0.07	−0.07	−0.07	−0.06	−0.06	−0.04	−0.04
Model-based	0.01	0.00	0.01	0.00	0.01	0.01	0.00	0.01	0.01	0.00

Note. Root mean squared error and relative bias of naive and model-based prevalence for sampling strategies with 10 capture occasions (Occ).

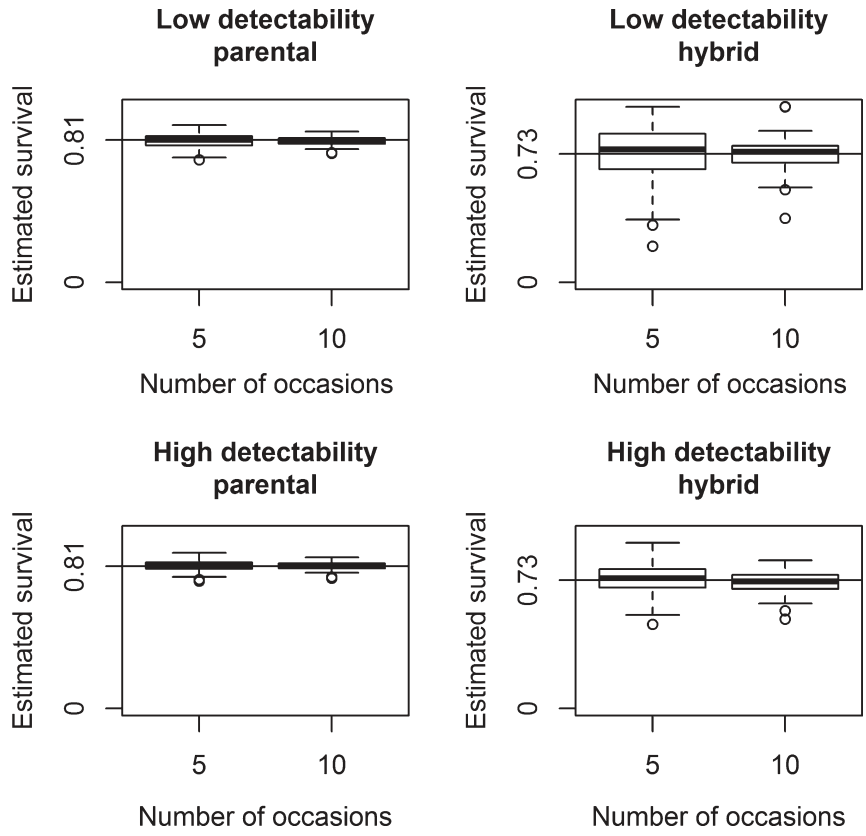


FIGURE 2 Scenario 1 of simulation (state-dependent detectability and homogeneous assignment probability). Boxplots of 100 simulated survival estimates for parentals (left two panels) and hybrids (right two panels) for each sampling strategy. Sampling strategies with low detectability are in the top row, and sampling strategies with high detectability are in the bottom row. Estimates obtained with sampling strategies with 5 and 10 capture occasions are compared in each panel

showed by the boxplot larger ranges (Figures 2–4; Supporting information Figures S2–S10).

Confidence interval coverage was always ≥0.89 for all scenarios with 5 sampling occasions (Table 3; Supporting information Tables S6, S7). For scenarios with 10 occasions (Table 4; Supporting information Tables S8, S9) coverage decreased after the 7th–8th occasion. This is because the low number of individuals present at the end of the study (due to a low apparent survival) affects the accuracy of the estimates (Figure 1; Supporting information Figures S1, S2).

3.2 | Evaluation of sampling strategies

The performance of model-based prevalence improved more at higher detectability than with an increasing number of capture

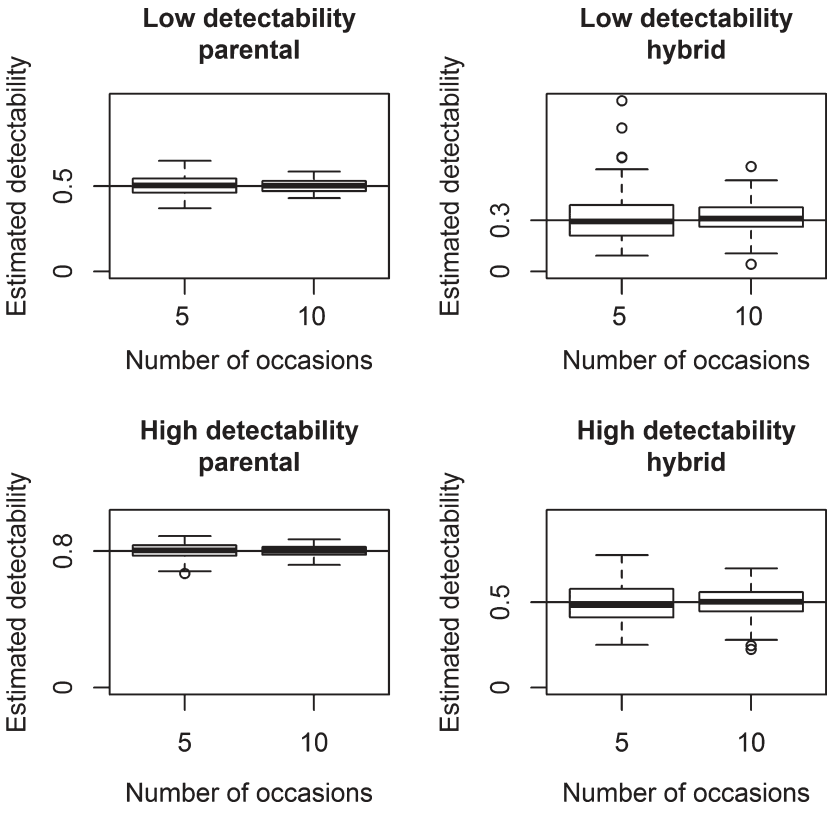


FIGURE 3 Scenario 1 of simulation (state-dependent detectability and homogeneous assignment probability). Boxplots of 100 simulated detectability estimates for parentals (left two panels) and hybrids (right two panels) for each sampling strategy. Sampling strategies with low detectability are in the top row, and sampling strategies with high detectability are in the bottom row. Estimates obtained with sampling strategies with 5 and 10 capture occasions are compared in each panel

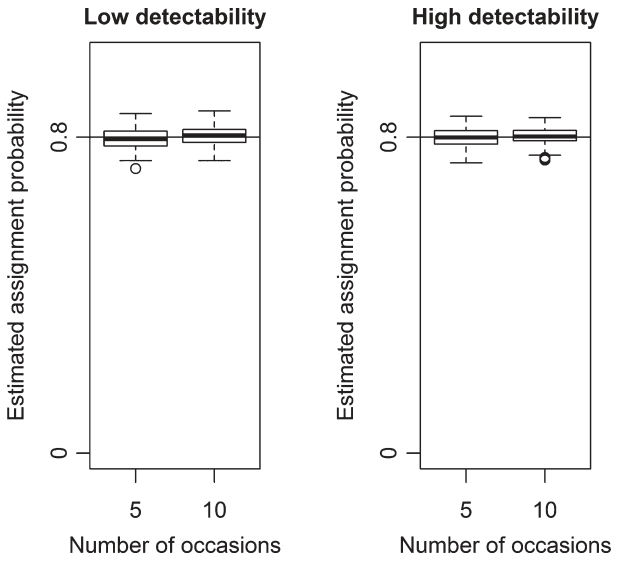


FIGURE 4 Scenario 1 of simulation (state-dependent detectability and homogeneous assignment probability). Boxplots of 100 simulated assignment probability estimates. Sampling strategies with low detectability are on the left panel, and sampling strategies with high detectability are on the right panel. Estimates obtained with sampling strategies with 5 and 10 capture occasions are compared in each panel

occasions. Specifically, in scenario 1 the RMSE and the relative bias approached 0 for both sampling strategies with 5 (Table 1) and 10 capture occasions (Table 2) when detectability changed from low to high. In contrast, in sampling strategies with low detectability, the RMSE and the relative bias of model-based prevalence decreased less rapidly from 5 to 10 capture occasions (Figure 1, Tables 1,2). We observed the same pattern both in scenario 2 (Supporting information Figures S1, S2, Tables S2, S3) and in scenario 3, although in the latter the bias was small regardless of the sampling strategy (Supporting information Tables S4, S5).

3.3 | Case study

The best-supported model had constant parameters (Table 5). Models with state-dependent apparent survival, detectability, and probability of assignment were also supported (ΔAICc < 2). The models with state-dependent probability of assignment were not identifiable due to the reduced sample size, and we discarded them (see Supporting information Table S10 for a complete list of fitted models). According to the model-averaged estimates of

TABLE 3 Scenario 1 of simulation (state-dependent detectability and homogeneous assignment probability)

Confidence interval coverage 5 occasions						
	Occ. 1	Occ. 2	Occ. 3	Occ. 4	Occ. 5	Average
Low p	0.94	0.97	0.96	0.94	0.89	0.94
High p	1.00	0.98	0.95	0.95	0.99	0.97

Note. Confidence interval coverage for sampling strategies with 5 capture occasions (Occ.).

TABLE 4 Scenario 1 of simulation (state-dependent detectability and homogeneous assignment probability)

Confidence interval coverage 10 occasions											
	Occ. 1	Occ. 2	Occ. 3	Occ. 4	Occ. 5	Occ. 6	Occ. 7	Occ. 8	Occ. 9	Occ. 10	Average
Low p	0.98	0.95	0.97	0.96	0.94	0.92	0.84	0.65	0.62	0.64	0.84
High p	1.00	0.99	0.97	0.97	0.95	0.91	0.92	0.85	0.91	0.75	0.92

Note. Confidence interval coverage for sampling strategies with 10 capture occasions (Occ.).

parameters (Table 6), model-based prevalence was consistently higher (range: 0.23–0.53) than naive prevalence (range: 0.20–0.46), with the latter always included in the confidence interval of the former (Table 7).

4 | DISCUSSION

We presented a hidden Markov model to estimate prevalence in wildlife population taking into account the imperfect detectability and uncertainty in individuals' classification. We compared model-based and naive prevalence showing that the latter can be severely biased when detectability is state-dependent, in agreement with Jennelle et al. (2007). In addition, we identified another source of bias in naive prevalence which was related to the state-dependent probability of assignment. We demonstrated that, if naive prevalence has to be used as a proxy for population prevalence, the burden of proof should be placed on demonstrating homogeneity in detection and assignment probabilities (Jennelle et al., 2007; MacKenzie & Kendall, 2002).

4.1 | Model assumptions

Our approach provides a framework to statistically test differences in detectability and probability of assignment and take them into account to produce unbiased estimates of prevalence. However, a series of assumptions must be met (Lebreton et al., 1992; Otis et al., 1978): (a) parameter and processes estimated for the marked individuals can be applied to the unmarked ones. In particular, because multievent CR models are conditional on first capture, we assume that capture probability is the same for unmarked and marked individuals in order to obtain abundance estimates, (b) marks do not affect the behavior of the individuals, (c) marks are not lost, and they are correctly recognized, (d) individuals alive in the population at time t have homogeneous detectability and apparent survival probability, (e) individuals are independent from each other, and (f) no births, deaths, emigration, or immigration occur during the capture occasions. The assumption of homogeneous detectability and survival can be relaxed by including different sources of heterogeneity in the model structure (Cubaynes et al., 2010; Pradel, 1993; Pradel, Hines, Lebreton, & Nichols, 1997). In particular when

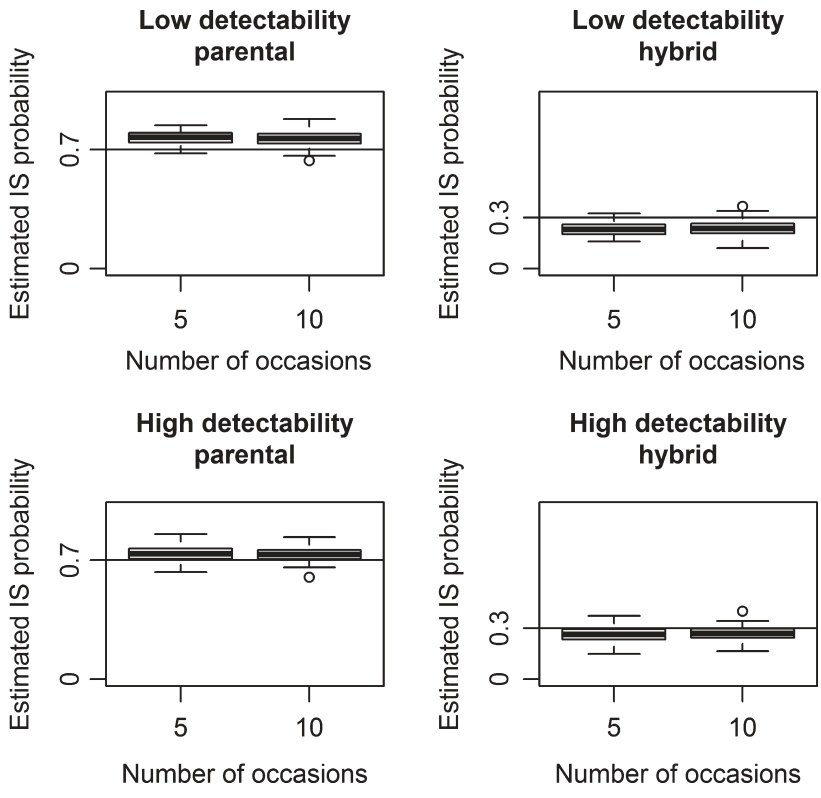


FIGURE 5 Scenario 1 of simulation (state-dependent detectability and homogeneous assignment probability). Boxplots of 100 simulated initial state probabilities estimates for parentals (left two panels) and hybrids (right two panels) for each sampling strategy. Sampling strategies with low detectability are in the top row, and sampling strategies with high detectability are in the bottom row. Estimates obtained with sampling strategies with 5 and 10 capture occasions are compared in each panel

TABLE 5 Model selection results for the case study on wolf × dog hybridization

Model	n_{par}	Deviance	QAICc	ΔQAICc
$\pi(i)\varphi(i)p(i)\delta(a1 + a2_fix)$	4	183.48	192.16	0
$\pi(i)\varphi(f)p(i)\delta(a1 + a2_fix)$	5	181.41	192.44	0.28
$\pi(i)\varphi(i)p(f)\delta(a1 + a2_fix)$	5	181.76	192.79	0.63

Note. The notation (.) indicates constant parameters while (state) indicates state-dependent parameters. π = initial state probability, φ = survival probability, p = detection probability, δ = assignment probability, n_{par} = number of parameters. The term ($a1 + a2_fix$) indicates that we constrained the model to have fixed assignment probabilities after first capture.

TABLE 6 Model-averaged parameter estimates for the case study on wolf × dog hybridization

Parameter	Estimate	95% C.I.
Initial proportion of individuals in state "Wolf" π_w	0.60	0.43–0.76
Detection probability of wolves p_w	0.42	0.17–0.67
Detection probability of hybrids p_h	0.50	0.26–0.77
Survival probability of wolves φ_w	0.72	0.39–0.91
Survival probability of hybrids φ_h	0.84	0.46–0.99
Assignment probability δ	0.79	0.64–0.89

estimating abundance attention must be paid to unaccounted for heterogeneity in capture probabilities, which is known to bias estimates of abundance (Pollock, Winterstein, Bunck, & Curtis, 1989).

4.2 | Limitations of the model

A potential drawback of our approach lies in convergence issues that might occur when there is a high proportion of uncertain individuals in the sample (Pradel et al., 2007). This problem can be overcome by confirming the state of just a handful of individuals with some error-free method. Specifically, Pradel et al. (2007) showed that in a case study in which the sex of about 80% of the individuals was uncertain, adding few genetic confirmations (24 individuals over 4,025) greatly improved the shape of the likelihood and hence the convergence of the optimization algorithm. This is particularly relevant because the model does not handle assignment errors, so having a high proportion of uncertain individuals is preferable to taking the risk of making assignment errors to reduce such proportion.

4.3 | Sampling guidelines

Through the simulations we showed that the precision of model-based prevalence increased proportionally more by enhancing detection probability than by increasing the number of sampling occasions. The precision of survival estimates increased by the same amount with increasing detectability and number of occasions, confirming that a cost-effective sampling strategy should maximize detectability within capture occasions instead of sampling more frequently. For a specific study, our simulation framework can be used to determine the best trade-off to obtain a cost-effective level of accuracy and precision of parameter estimates (Lieury et al., 2017).

4.4 | Case study

For the hybridization case study, potential differences in the detectability of parental and admixed individuals generate bias in naive prevalence and can originate from various reasons. For example, differences in vocalization behavior (Derégnaucourt, Guyomarc'h, & Spanò, 2005), migratory tendency (Derégnaucourt, Guyomarc'h, & Belhamra, 2004), and social status (Battocchio, Iacolina, Canu, & Mori, 2017) were documented between admixed parental individuals of different species and may cause differential detectability. In wolf packs, in particular, differences in detectability through scat sampling are related to social status and marking behavior (Cubaynes et al., 2010; Marucco et al., 2009).

Previous hybridization studies on wolves (Godinho et al., 2011; Pacheco et al., 2017) acknowledged that, due to uncertainty of classification, a proportion of backcrosses was assigned to the parental cluster (i.e., wolves), leading to an underestimation of their prevalence in the population. With our approach, a conservative q_{wolf} threshold can be used for the parental cluster, greatly reducing the chance of type II errors (erroneously classifying hybrids as parentals), as showed by the higher model-based prevalence values in the case study (Table 7). These results should be considered as a warning that relying on naive prevalence underestimates the hybridization-related risks for the conservation of the parental populations.

TABLE 7 Naive and model-based prevalence of hybrids for the case study on wolf × dog hybridization

Prevalence	Occ. 1	Occ. 2	Occ. 3	Occ. 4	Occ. 5
Naive	0.27	0.33	0.20	0.46	0.27
Model-based	0.32	0.42	0.23	0.53	0.43
95% CI	0.07–0.55	0.09–0.63	0.06–0.48	0.12–0.72	0.09–0.67

4.5 | Model extensions and broader applications of the model

The current model formulation contains two states ("Hybrid" and "Parental"). It might be desirable, however, to further split the "Hybrid" state into two distinct states "Hybrid" and "Backcross," or to include not just one but both parental species. This can be done by increasing the number of states and corresponding parameters depending on the available data and the power of the genetic tests. However, attention should be paid to avoid over-parameterization (Gimenez, Choquet, & Lebreton, 2003). Multievent models can quickly become parameter-rich and thus result in nonidentifiability in studies with small sample sizes.

The Viterbi algorithm has been previously applied in the multievent CR framework to reconstruct the reproductive life of individual roe deers (*Capreolus capreolus*; Rouan et al., 2009). We used the Viterbi algorithm to reconstruct the abundance of individuals in different states. This approach can be used in any case study that requires the estimation of abundance and prevalence of individuals in the presence of imperfect detection and uncertainty in state assignment, for example, the estimation of sex ratios in monomorphic species (Genovart et al., 2012; Pradel et al., 2007). Moreover, the model can be extended to include dynamic states by adding a transition probability matrix in the formulation. Such extension would expand its applications to other fields such as epidemiology (e.g., the estimation of the number of infected individuals in a population; Marescot et al., 2018) or reproductive biology studies (e.g., the number of breeders in a population; Desprez, McMahon, Hindell, Harcourt, & Gimenez, 2013).

ACKNOWLEDGMENTS

We are grateful to M. Canestrini and F. Moretti for sample collection and to M. Galaverni for assistance with genetic analyses. We thank Chris Grieves, Paul B. Conn, and one anonymous reviewer for their insightful suggestions during the editorial process. We also thank Rémi Choquet and Roger Pradel for their help in the model development. OG and NLS were supported by the French National Research Agency with a Grant ANR-16-CE02-0007. NLS was supported by a PhD grant by Sapienza University of Rome.

CONFLICT OF INTEREST

None declared.

AUTHOR CONTRIBUTIONS

NLS, OG, and PC conceived the ideas and designed methodology and led the writing of the manuscript; PC and WR organized and supervised the field data collection and logistics; LM collected the data; NLS and OG constructed the models, performed the simulations, and analyzed the data; RC and EF performed the genetic laboratory and cluster analyses. All authors contributed critically to the drafts and gave final approval for publication.

DATA ACCESSIBILITY

Codes of the models, simulations, and data are available from the Dryad Digital Repository (https://doi.org/10.5061/dryad.8g8r675).

ORCID

Nina Luisa Santostasi https://orcid.org/0000-0001-9148-7069

REFERENCES

Allendorf, F. W., Leary, R. F., Spruell, P., & Wenburg, J. K. (2001). The problems with hybrids: setting conservation guidelines. *Trends in Ecology & Evolution*, *16*, 613–622. https://doi.org/10.1016/S0169-5347(01)02290-X

Battocchio, D., Iacolina, L., Canu, A., & Mori, E. (2017). How much does it cost to look like a pig in a wild boar group? *Behavioural Processes*, *138*, 123–126. https://doi.org/10.1016/j.beproc.2017.03.001

Boitani, L. (2000). *Action plan for the conservation of Wolves in Europe (Canis lupus)*. Strasbourg, France: Council of Europe Publishing.

Burnham, K. P., & Anderson, D. R. (2002). *Model selection and multimodel inference: A practical information-theoretic approach*. New York, NY: Springer.

Caniglia, R., Fabbri, E., Cubaynes, S., Gimenez, O., Lebreton, J.-D., & Randi, E. (2012). An improved procedure to estimate wolf abundance using non-invasive genetic sampling and capture – recapture mixture models. *Conservation Genetics*, *13*, 53–64. https://doi.org/10.1007/s10592-011-0266-1

Caniglia, R., Fabbri, E., Galaverni, M., Milanesi, P., & Randi, E. (2014). Noninvasive sampling and genetic variability, pack structure, and dynamics in an expanding wolf population. *Journal of Mammalogy*, *95*, 41–59. https://doi.org/10.1644/13-MAMM-A-039

Caniglia, R., Fabbri, E., Greco, C., Galaverni, M., Manghi, L., Boitani, L., & Randi, E. (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. https://doi.org/10.1007/s10344-013-0703-1

Caswell, H. (2000). *Matrix population models: Construction, analysis and interpretation*. Sunderland, MA: Sinauer.

Choquet, R., Rouan, L., & Pradel, R. (2009). Program E-Surge: A software application for fitting multievent models. In D. L. Thompson, E. G. Cooch, & M. J. Conroy (Eds.), *Modeling demographic processes in marked populations*, volume 3 of Environmental and Ecological Statistics (pp. 845–865). New York, NY: Springer series.

Conn, P. B., & Cooch, E. G. (2009). Multistate capture – recapture analysis under imperfect state observation: an application to disease models. *Journal of Applied Ecology*, *46*, 486–492. https://doi.org/10.1111/j.1365-2664.2008.01597.x

Cubaynes, S., Pradel, R., Choquet, R., Duchamp, C., Gaillard, J.-M., Lebreton, J.-D., … Gimenez, O. (2010). Importance of accounting for detection heterogeneity when estimating abundance: the case of French wolves. *Conservation Biology*, *24*, 621–626. https://doi.org/10.1111/j.1523-1739.2009.01431.x

Davison, A. C., & Hinkley, D. V. (1997). *Bootstrap methods and their application*. Cambridge, UK: Cambridge University Press.

Derégnaucourt, S., Guyomarc'h, J.-C., & Belhamra, M. (2004). Comparison of migratory tendency in European Quail *Coturnix coturnix*, domestic Japanese Quail *Coturnix japonica* and their hybrids. *Ibis*, *147*, 25–36. https://doi.org/10.1111/j.1474-919x.2004.00313.x

Derégnaucourt, S., Guyomarc'h, J.-C., & Spanò, S. (2005). Behavioural evidence of hybridization (Japanese × European) in domestic quail released as game birds. *Applied Animal Behaviour Science*, *94*, 303–318. https://doi.org/10.1016/j.applanim.2005.03.002

Desprez, M., McMahon, C. R., Hindell, M. A., Harcourt, R., & Gimenez, O. (2013). Known unknowns in an imperfect world : incorporating uncertainty in recruitment estimates using multievent capture – recapture models. *Ecology and Evolution*, *3*, 4658–4668. https://doi.org/10.1002/ece3.846

Falush, D., Stephens, M., & Pritchard, J. K. (2003). Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics*, *164*, 1567–1587. https://doi.org/10.1111/j.1471-8286.2007.01758.x

Genovart, M., Pradel, R., & Oro, D. (2012). Exploiting uncertain ecological fieldwork data with multievent capture – recapture modelling: an example with bird sex assignment. *Journal of Animal Ecology*, *81*, 970–977. https://doi.org/10.1111/j.1365-2656.2012.01991.x

Gervasi, V., Boitani, L., Paetkau, D., Posillico, M., Randi, E., & Ciucci, P. (2017). Estimating survival in the Apennine brown bear accounting for uncertainty in age classification. *Population Ecology*, *59*, 119–130. https://doi.org/10.1007/s10144-017-0587-0

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

Gimenez, O., Lebreton, J.-D., Gaillard, J.-M., Choquet, R., & Pradel, R. (2012). Estimating demographic parameters using hidden process dynamic models. *Theoretical Population Biology*, *82*, 307–316. https://doi.org/10.1016/j.tpb.2012.02.001

Godinho, R., Llaneza, L., Blanco, J. C., Lopes, S., Álvares, F., García, E. J., & Ferrand, N. (2011). Genetic evidence for multiple events of hybridization between wolves and domestic dogs in the Iberian Peninsula. *Molecular Ecology*, *20*, 5154–5166. https://doi.org/10.1111/j.1365-294X.2011.05345.x

Himmelmann, L. (2010). HMM: HMM-hidden Markov Models. Scientific Software Development, 2010. R Package Version 1.0.

Hindrickson, M., Remm, J., Pilot, M., Godinho, R., Stronen, A. V., Baltrunaité, L., … Saarma, U. (2017). Wolf population genetics in Europe: A systematic review, meta-analysis and suggestions for conservation and managment. *Biological Reviews*, *92*, 1601–1629.

Jennelle, C. S., Cooch, E., Conroy, M. J., & Senar, J. C. (2007). State specific detection probability and disease prevalence. *Ecological Applications*, *17*, 154–167. https://doi.org/10.1890/1051-0761(2007)017[0154:SDPADP]2.0.CO;2

Lebreton, J., Burnham, K., Clobert, J., & Anderson, D. (1992). Modelling survival and testing biologcal hypothesis using marked animals: a unified approach with case studies. *Ecological Monographs*, *62*, 67–118. https://doi.org/10.2307/2937171

Lieury, N., Devillard, S., Besnard, A., Gimenez, O., Hameau, O., Ponchon, C., & Millon, A. (2017). Designing cost-effective capture – recapture surveys for improving the monitoring of survival in bird populations. *Biological Conservation*, *214*, 233–241. https://doi.org/10.1016/j.biocon.2017.08.011

MacKenzie, D. I., & Kendall, W. L. (2002). How should detection probability be incorporated into estimates of relative abundance? *Ecology*, *83*, 2387–2393. https://doi.org/10.2307/3071800

Marescot, L., Benhaiem, S., Gimenez, O., Hofer, H., Lebreton, J.-D., Olarte-Castillo, X. A., & East, M. L. (2018). Social status mediates the fitness costs of infection with canine distemper virus in Serengeti spotted hyenas. *Functional Ecology*, *32*(1237–1250), 1–14. https://doi.org/10.1111/1365-2435.13059/epdf

Marucco, F., Pletscher, D. H., Boitani, L., Schwartz, M. K., Pilgrim, K. L., & Lebreton, J. (2009). Wolf survival and population trend using non-invasive capture – recapture techniques in the Western Alps. *Journal of Applied Ecology*, *46*, 1003–1010. https://doi.org/10.1111/j.1365-2664.2009.01696.x

McDonald, T. L., & Amstrup, S. C. (2001). Estimation of population size using open capture-recapture models. *Journal of Agricultural, Biological and Environmental statistics*, *6*, 206–220. https://doi.org/10.1198/108571101750524553

Moreno-Torres, K., Wolfe, B., Saville, W., & Garabed, R. (2016). Estimating *Neospora caninum* prevalence in wildlife popuations using Bayesian inference. *Ecology and Evolution*, *6*, 2216–2225. https://doi.org/10.1002/ece3.2050

Nichols, J. D. (1992). Capture-recapture models: using marked animals to study population dynamics. *Bio Science*, *42*, 94–102. https://doi.org/10.2307/1311650

Otis, D. L., Burnham, K. P., White, G. C., & Anderson, D. R. (1978). Statistical inference from capture data on closed animal populations. *Wildlife Monographs*, *62*, 3–135. https://doi.org/10.2307/2287873

Pacheco, C., López-bao, J. V., García, E. J., Lema, F. J., Llaneza, L., Palacios, V., & Godinho, R. (2017). Spatial assessment of wolf-dog hybridization in a single breeding period. *Scientific Reports*, *7*, 1–10. https://doi.org/10.1038/srep42475

Pollock, K. H., Winterstein, S. R., Bunck, C. M., & Curtis, P. D. (1989). Survival analysis in telemetry studies: The stagged entry design. *Journal of Wildlife Management*, *53*, 7–15.

Pradel, R. (1993). Flexibility in survival analysis from recapture data: Handling trap–dependence. In J.-D. Lebreton, & P. M. North (Eds.), *Marked individuals in the study of bird populations* (pp. 29–37). Basel, Switzerland: Birkhaüser Verlag.

Pradel, R. (2005). Multievent: an extension of multistate capture – recapture models to uncertain states. *Biometrics*, *61*, 442–447. https://doi.org/10.1111/j.1541-0420.2005.00318.x

Pradel, R., Hines, J. E., Lebreton, J.-D., & Nichols, J. D. (1997). Capture – recapture survival models taking account of transients. *Biometrics*, *53*, 60–72. https://doi.org/10.2307/2533097

Pradel, R., Maurin-Bernier, L., Gimenez, O., Genovart, M., Choquet, R., & Oro, D. (2007). Estimation of sex-specific survival with uncertainty in sex assessment. *The Canadian Journal of Statistics*, *35*, 1–14. https://doi.org/10.1002/cjs.5550360105

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics*, *155*, 945–959. https://doi.org/10.1111/j.1471-8286.2007.01758.x

R Core Team. (2017). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Randi, E., Hulva, P., Fabbri, E., Galaverni, M., Galov, A., Kusak, J., … Caniglia, R. (2014). Multilocus detection of wolf × dog hybridization in Italy, and guidelines for marker selection. *PLoS ONE*, *9*, e86409. https://doi.org/10.1371/journal.pone.0086409

Rouan, L., Gaillard, J.-M., Guédon, Y., & Pradel, R. (2009). Estimation of lifetime reproductive success when reproductive status cannot always be assessed. In D. L. Thompson, E. G. Cooch, & M. J. Conroy (Eds.), *Modeling demographic processes in marked populations*, volume 3 of Environmental and Ecological Statistics (pp. 845–865). New York, NY: Springer series.

Sanz-Aguilar, A., Tavecchia, G., Genovart, M., Igual, J. M., Oro, D., Rouan, L., & Pradel, R. (2011). Studying the reproductive skipping behavior in long-lived birds by adding nest inspection to individual-based data. *Ecological Applications*, *21*, 555–564. https://doi.org/10.1890/09-2339.1

Schwenk, K., Brede, N., & Streit, B. (2008). Introduction. Extent, processes and evolutionary impact of interspecific hybridization in animals. *Philosophical Transactions of the Royal Society B*, *363*, 2805–2811. https://doi.org/10.1098/rstb.2008.0055

Todesco, M., Pascual, M. A., Owens, G. L., Ostevik, K. L., Moyers, B. T., Hübner, S., & Riesberg, L. H. (2016). Hybridization and extinction. *Evolutionary Applications*, *9*, 892–908. https://doi.org/10.1111/eva.12367

Vähä, J. P., & Primmer, C. R. (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. https://doi.org/10.1111/j.1365-294X.2005.02773.x

Vaz Pinto, P., Beja, P., Ferrand, N., & Godinho, R. (2015). Hybridization following population collapse in a critically endangered antelope. *Scientific Reports*, *6*, 1–9. https://doi.org/10.1038/srep18788

Zucchini, W., MacDonald, I. L., & Langrock, R. (2009). *Hidden Markov models for time series: an introduction using R*. Boca Raton, FL: Chapman & Hall/CRC Press.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Santostasi NL, Ciucci P, Caniglia R, et al. 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. *Ecol Evol*. 2019;9:744–755. https://doi.org/10.1002/ece3.4819

APPENDIX CHAPTER 1

SUPPLEMENTARY MATERIALS

1. MODEL DETAILS

1.1 Instruction to fit the model in E_SURGE:

INPUT DATA CODING:

We coded the data in the following way (individuals were assigned upon first capture and never again):

Example of coding of encounter histories.

Capture history	**State assignment**
10330	Sure wolf (W)
02330	Sure hybrid (H)
03303	Uncertain (H)

GEPAT:

Initial States

$$\begin{matrix} \boldsymbol{P} & \boldsymbol{H} \end{matrix}$$
$$\begin{pmatrix} \pi & 1-\pi \end{pmatrix}$$

Transitions

$$\begin{matrix} & \boldsymbol{P} & \boldsymbol{H} & \boldsymbol{D} \\ \boldsymbol{P} & \varphi_p & 0 & 1-\varphi_p \\ \boldsymbol{H} & 0 & \varphi_h & 1-\varphi_h \\ \boldsymbol{D} & 0 & 0 & 1 \end{matrix}$$

Events

$$\begin{matrix} & \mathbf{0} & \mathbf{1} & \mathbf{2} \\ \boldsymbol{P} & 1-p_p & p_p & 0 \\ \boldsymbol{H} & 1-p_h & 0 & p_h \\ \boldsymbol{D} & 1 & 0 & 0 \end{matrix} \qquad \begin{matrix} & \mathbf{0} & \mathbf{1} & \mathbf{2} & \mathbf{3} \\ \mathbf{0} & 1 & 0 & 0 & 0 \\ \mathbf{1} & 0 & \delta_p & 0 & 1-\delta_p \\ \mathbf{2} & 0 & 0 & \delta_h & 1-\delta_h \end{matrix}$$

GEMACO:

IS: to (for constant), t (for time)

TRANSITIONS: i (for constant), f (for state depndendent)

EVENTS:

- Step 1: firste+nexte (for constant), firste+nexte.f (for state depndendent), firste+nexte.t (for time)
- Step 2: a (age effect, where age is time since first capture, this allows for having assignment probability upon first capture ≠ assignment probability on recaptures), a.f (for having assignment probability upon first capture ≠ assignment probability on recaptures and a state effect)

IVFV:

Events

- **Step 1:** fixed first detection to 1
- **Step 2:** fix assignment probability at age 2 (meaning all assignment probabilities after first capture) to 0

2. TABLES

Table S1. Detectability (p) and assignment probabilities (δ) used for the three main scenarios of simulation for wolves (w) and hybrids (h). Scenario 1 implies state-dependent detectability and homogeneous assignment probability, Scenario 2 implies homogeneous detectability and state-dependent assignment probability, Scenario 3 implies homogeneous detectability and assignment probability.

	Scenario 1		Scenario 2		Scenario 3	
Detectability	High	Low	High	Low	High	Low
p_w	0.80	0.50	0.80	0.50	0.80	0.50
p_h	0.80	0.50	0.50	0.30	0.80	0.50
δ_w	0.80	0.80	0.80	0.80	0.80	0.80
δ_h	0.80	0.80	0.80	0.80	0.60	0.60

2.1 Simulation results for Scenarios 2 and 3

2.1.1 Root Mean Squared Error and Relative Bias

Table S2. Scenario 2 (homogeneous detectability and state-dependent assignment probability). Root mean squared error and relative bias of naive and model-based prevalence for sampling strategies with 5 capture occasions (Occ).

	Occ1	Occ2	Occ3	Occ4	Occ5
Root mean squared error					
	Low detectability				
Naive	0.21	0.25	0.21	0.41	0.29
Model-based	0.03	0.02	0.01	0.02	0.01
	High detectability				
Naive	0.31	0.26	0.26	0.23	0.15
Model-based	0.04	0.02	0.00	0.00	0.00
Percent relative bias					
	Low detectability				
	-0.05	-0.05	-0.05	-0.06	-0.05
	0.01	0.01	0.01	0.02	0.01
	High detectability				
Naive	-0.06	-0.05	-0.05	-0.05	-0.04
Model-based	0.02	0.01	0.00	0.00	0.00

Table S3. Scenario 2 (homogeneous detectability and state-dependent assignment probability). Root mean squared error and relative bias of naive and model-based prevalence for sampling strategies with 10 capture occasions (Occ).

	Occ1	Occ2	Occ3	Occ4	Occ5	Occ6	Occ7	Occ8	Occ9	Occ10
Root Mean Squared Error										
	Low detectability									
Naive	0.19	0.15	0.17	0.08	0.28	0.17	0.16	0.13	0.02	NaN
Model-based	0.10	0.06	0.05	0.06	0.01	0.00	0.00	0.00	0.01	0.01
	High detectability									
Naive	0.27	0.31	0.23	0.22	0.19	0.12	0.12	0.19	0.13	0.05
Model-based	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Percent relative bias										
	Low detectability									
Naive	-0.04	-0.04	-0.04	-0.03	-0.05	-0.04	-0.04	-0.04	-0.01	NaN
Model-based	0.03	0.02	0.02	0.02	0.01	0.00	0.00	0.00	0.01	0.01
	High detectability									
Naive	-0.05	-0.06	-0.05	-0.05	-0.04	-0.03	-0.04	-0.04	-0.04	-0.02
Model-based	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00

Table S4. Scenario 3 (homogeneous detectability and assignment probability). Root mean squared error and relative bias of naive and model-based prevalence for sampling strategies with 5 capture occasions (Occ).

	Occ1	Occ2	Occ3	Occ4	Occ5
Root mean squared error					
	Low detectability				
Naive	0.00	0.01	0.00	0.00	0.00
Model-based	0.05	0.00	0.01	0.00	0.00
	High detectability				
Naive	0.00	0.00	0.00	0.00	0.01
Model-based	0.03	0.01	0.00	0.00	0.00
Percent relative bias					
	Low detectability				
Naive	-0.00	-0.01	-0.00	-0.00	0.00
Model-based	0.00	-0.00	-0.01	-0.00	-0.00
	High detectability				
Naive	0.00	-0.00	0.00	0.00	0.01
Model-based	-0.02	-0.01	0.00	-0.00	0.00

Table S5. Scenario 3 (homogeneous detectability and assignment probability). Root mean squared error and relative bias of naive and model-based prevalence for sampling strategies with 10 capture occasions (Occ).

	Occ1	Occ2	Occ3	Occ4	Occ5	Occ6	Occ7	Occ8	Occ9	Occ10
Root Mean Squared Error										
	Low detectability									
Naive	0.05	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.00	0.01
Model-based	0.00	0.00	0.00	0.00	0.00	0.03	0.01	0.00	0.00	0.03
	High detectability									
Naive	0.00	0.00	0.001	0.00	0.01	0.00	0.00	0.01	0.01	0.00
Model-based	0.03	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Percent relative bias										
	Low detectability									
Naive	0.02	0.01	0.00	-0.00	-0.00	-0.01	-0.01	-0.00	0.00	0.03
Model-based	-0.00	-0.00	0.00	-0.01	-0.01	-0.02	-0.01	-0.01	-0.00	-0.02
	High detectability									
Naive	0.00	0.00	0.01	-0.00	0.00	-0.00	-0.01	-0.01	-0.01	0.00
Model-based	-0.02	-0.01	-0.00	-0.00	0.00	-0.00	-0.00	-0.01	0.00	0.00

71

72 2.1.2 Confidence intervals coverage

73 **Table S6.** Scenario 2 (homogeneous detectability and state-dependent assignment probability) for

74 sampling strategies with 5 capture occasions (Occ). Confidence interval coverage.

Confidence interval coverage 5 occasions						
	Occ1	Occ2	Occ3	Occ4	Occ5	Ave
Low p	0.97	1.00	0.99	0.95	0.98	0.98
High p	1.00	1.00	1.00	1.00	0.99	1.00

75

76 **Table S7.** Scenario 2 (homogeneous detectability and state-dependent assignment probability) for

77 sampling strategies with 5 capture occasions (Occ). Confidence interval coverage.

Confidence interval coverage 10 occasions											
	Occ1	Occ2	Occ3	Occ4	Occ5	Occ6	Occ7	Occ8	Occ9	Occ10	Ave
Low p	0.97	1.00	1.00	0.97	0.97	0.93	0.90	0.88	0.81	0.70	0.91
High p	0.99	1.00	1.00	1.00	1.00	0.99	0.98	0.94	0.97	0.94	0.98

78

79 **Table S8.** Scenario 3 (homogeneous detectability and assignment probability) for sampling

80 strategies with 5 capture occasions (Occ). Confidence interval coverage.

Confidence interval coverage 5 occasions						
	Occ1	Occ2	Occ3	Occ4	Occ5	Ave
Low p	0.98	0.98	0.99	0.98	0.99	0.98
High p	1	1	1	1	1	1

81

82

Table S9. Scenario 3 (homogeneous detectability and assignment probability) for sampling strategies with 10 capture occasions (Occ). Confidence interval coverage.

Confidence interval coverage 10 occasions											
	Occ1	Occ2	Occ3	Occ4	Occ5	Occ6	Occ7	Occ8	Occ9	Occ10	Ave
Low p	1.00	0.98	0.99	0.98	0.95	0.92	0.95	0.92	0.77	0.87	0.93
High p	1.00	1.00	1.00	0.98	1.00	1.00	0.99	0.98	0.97	0.95	0.98

2.2 Case study results

Table S10. Model selection results for the wolf x dog case study. The notation (.) indicates constant parameters, (state) indicates state-dependent parameters, (time) indicates time-dependent parameters. π = initial state probability, φ = survival probability, p = detection probability, δ = assignment probability, N. Par. = number of parameters. The term (a1+a2_fix) indicates that we constrained the model to have fixed assignment probabilities after first capture. The term (a1+a2_fix).state indicates that the assignment probability is state-dependent. Models with state dependent assignment probability have been dropped from the model averaging because they had non-identifiabe parameters.

Model	N. Par.	Deviance	QAICc	deltaAICc
π(i)phi(i)p(i)delta(a1+a2_fix).state	4	183.48	192.16	0.00
π(i)phi(i)p(i)delta(a1+a2_fix)	4	183.48	192.16	0.00
π(i)phi(state)p(i)delta(a1+a2_fix)	5	181.41	192.44	0.28
π(i)phi(i)p(state)delta(a1+a2_fix)	5	181.76	192.79	0.63
π(i)phi(state)p(i)delta(a1+a2).state	6	179.40	192.87	0.71
π(i)phi(i)p(state)delta(a1+a2).state	6	181.25	194.73	2.57
π(i)phi(state)p(state)delta(a1+a2_fix).state	6	181.25	194.73	2.57

π(i)phi(state)p(state)delta(a1+a2_fix)	6	181.40	194.88	2.72
π(i)phi(i)p(t)delta(a1+a2_fix)	7	180.63	196.63	4.47
π(i)phi(state)p(t)delta(a1+a2_fix)	8	178.62	197.24	5.08
π(i)phi(state)p(t)delta(a1+a2).state	9	176.27	197.60	5.44
π(t)phi(i)p(i)delta(a1+a2_fix)	8	181.48	200.10	7.94
π(t)phi(state)p(i)delta(a1+a2)	9	179.11	200.45	8.29
π(t)phi(i)p(state)delta(a1+a2_fix)	9	179.17	200.50	8.34
π(t)phi(state)p(i)delta(a1+a2).state	10	176.89	201.04	8.88
π(i)phi(i)p(t)delta(a1+a2).state	9	180.63	201.96	9.80
π(t)phi(i)p(i)delta(a1+a2).state	9	180.98	202.31	10.15
π(t)phi(i)p(state)delta(a1+a2).state	10	178.41	202.56	10.40
π(t)phi(state)p(state)delta(a1+a2_fix)	10	179.04	203.19	11.03
π(t)phi(state)p(state)delta(a1+a2).state	11	176.88	203.95	11.79
π(t)phi(i)p(t)delta(a1+a2_fix)	11	178.63	205.71	13.55
π(t)phi(state)p(t)delta(a1+a2_fix)	12	176.36	206.48	14.32
π(t)phi(state)p(t)delta(a1+a2).state	13	173.76	207.04	14.88
π(t)phi(i)p(t)delta(a1+a2).state	12	178.12	208.24	16.08

96

97

3. FIGURES

3.1 Simulated prevalence estimates for scenario 2: homogenous detectability and state-dependent assignment probability

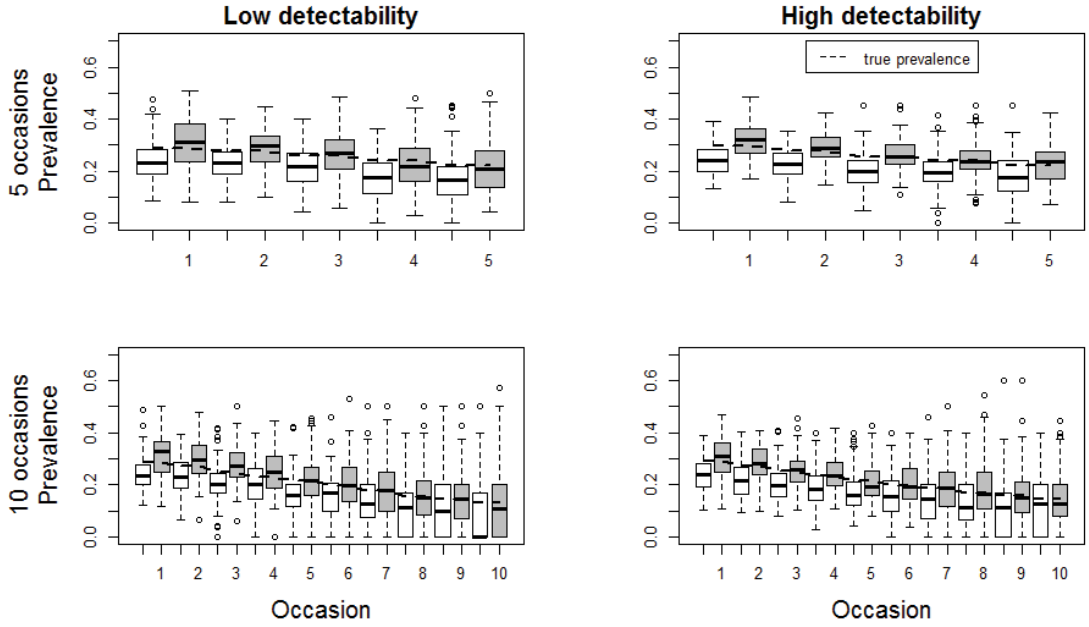


Figure S1. Scenario 2 (homogeneous detectability and state-dependent assignment probability). Sampling strategies with 5 (upper panels) vs 10 (lower panels) capture occasions and low (left-column panels) vs high (right-column panel) detectability. True prevalence is represented as a dashed line while the 100 values of naive and model-based prevalence are displayed in the white and grey boxplots, respectively.

3.2 Simulated prevalence estimates for scenario 3: homogenous detectability and assignment probability

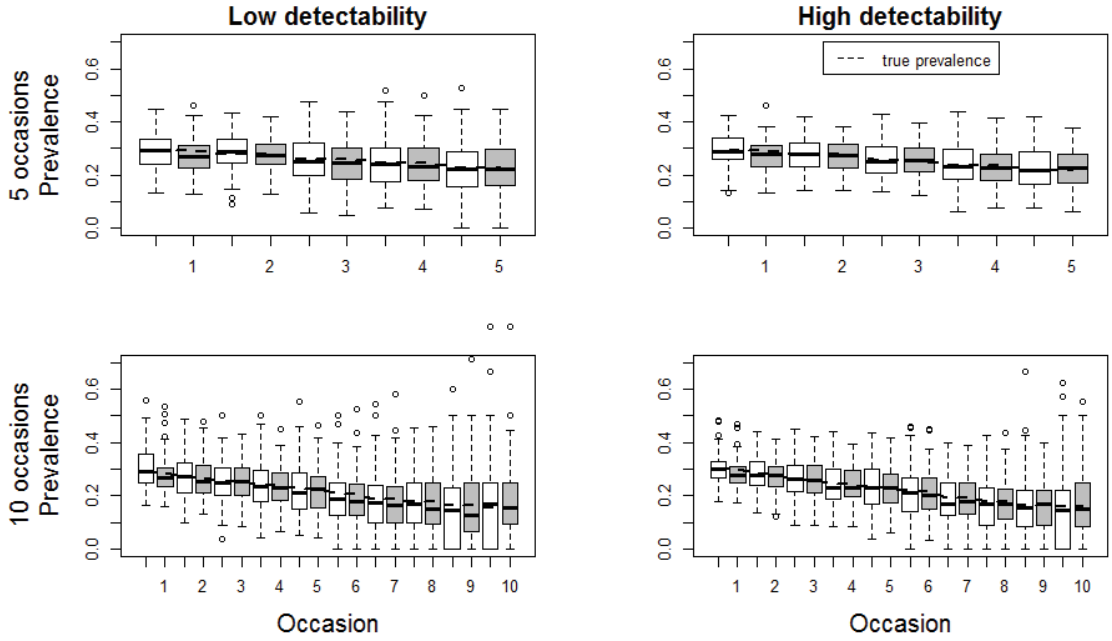


Figure S2. Scenario 3 (homogeneous detectability and assignment probability). Sampling strategies with 5 (upper panels) vs 10 (lower panels) capture occasions and low (left-column panels) vs high (right-column panel) detectability. True prevalence is represented as a dashed line while the 100 values of naive and model-based prevalence are displayed in the white and grey boxplots, respectively.

3.3 Simulated parameter estimates for simulations scenario 2:homogenous detectablity and state-dependent assignment probability

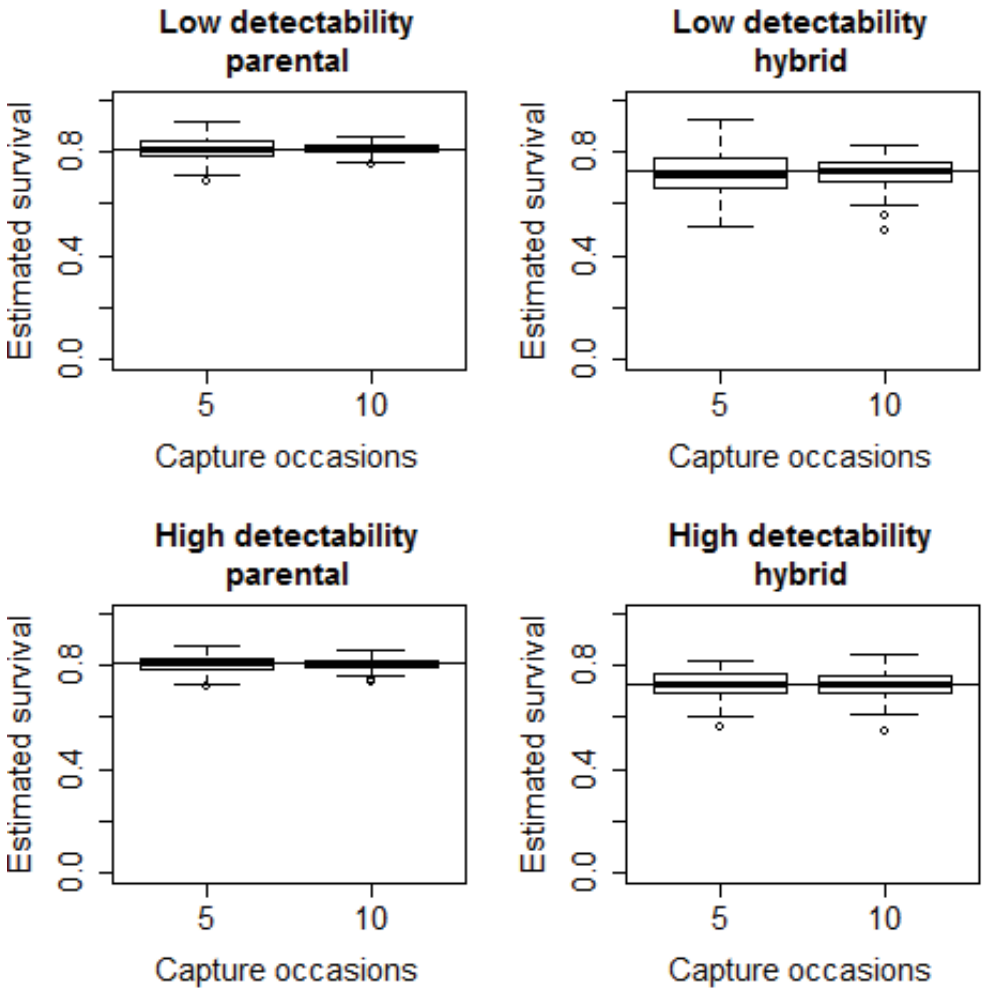


Figure S3. Scenario 2 (homogeneous detectability and state-dependent assignment probability). Boxplots of 100 simulated survival estimates for parentals (left two panels) and hybrids (right two panels) for each sampling strategy. Sampling strategies with low detectability are in the top row, sampling strategies with high detectability are in the bottom row. Estimates obtained with sampling strategies with 5 and 10 capture occasions are compared in each panel.

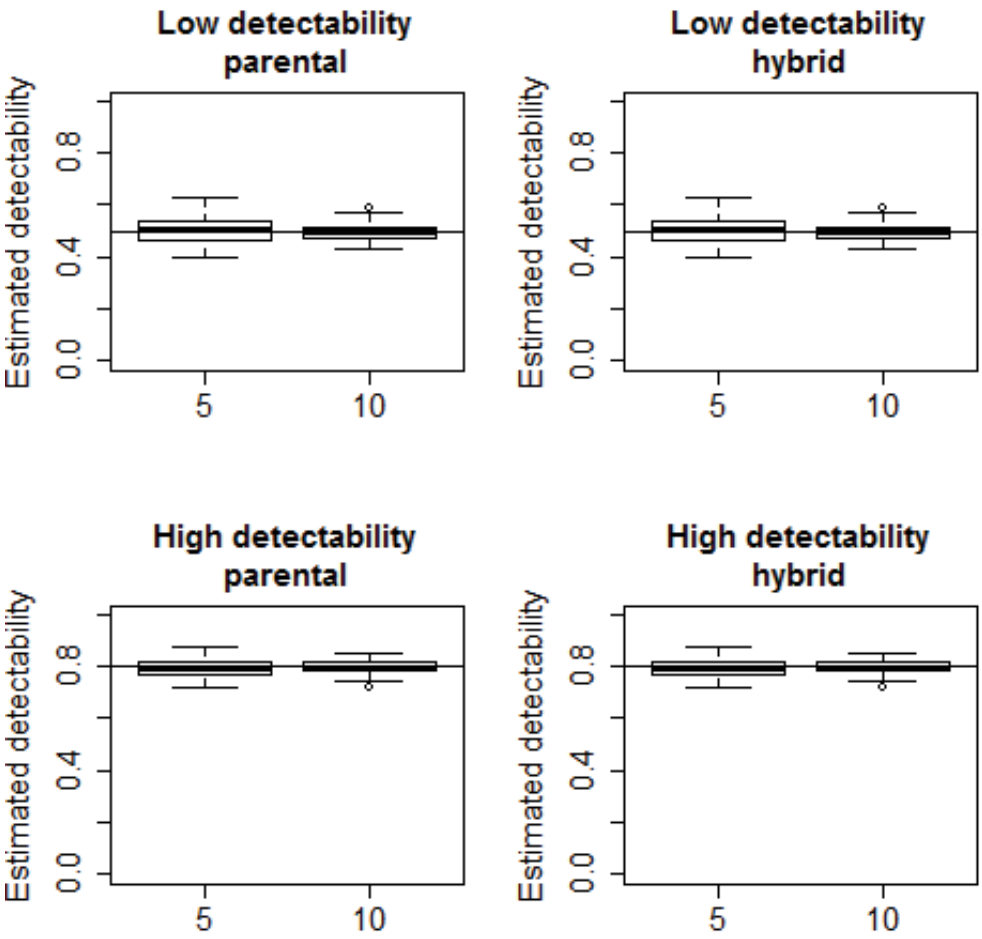


Figure S4. Scenario 2 (homogeneous detectability and state-dependent assignment probability). Boxplots of 100 simulated detectability estimates for parentals (left two panels) and hybrids (right two panels) for each sampling strategy. Sampling strategies with low detectability are in the top row, sampling strategies with high detectability are in the bottom row. Estimates obtained with sampling strategies with 5 and 10 capture occasions are compared in each panel.

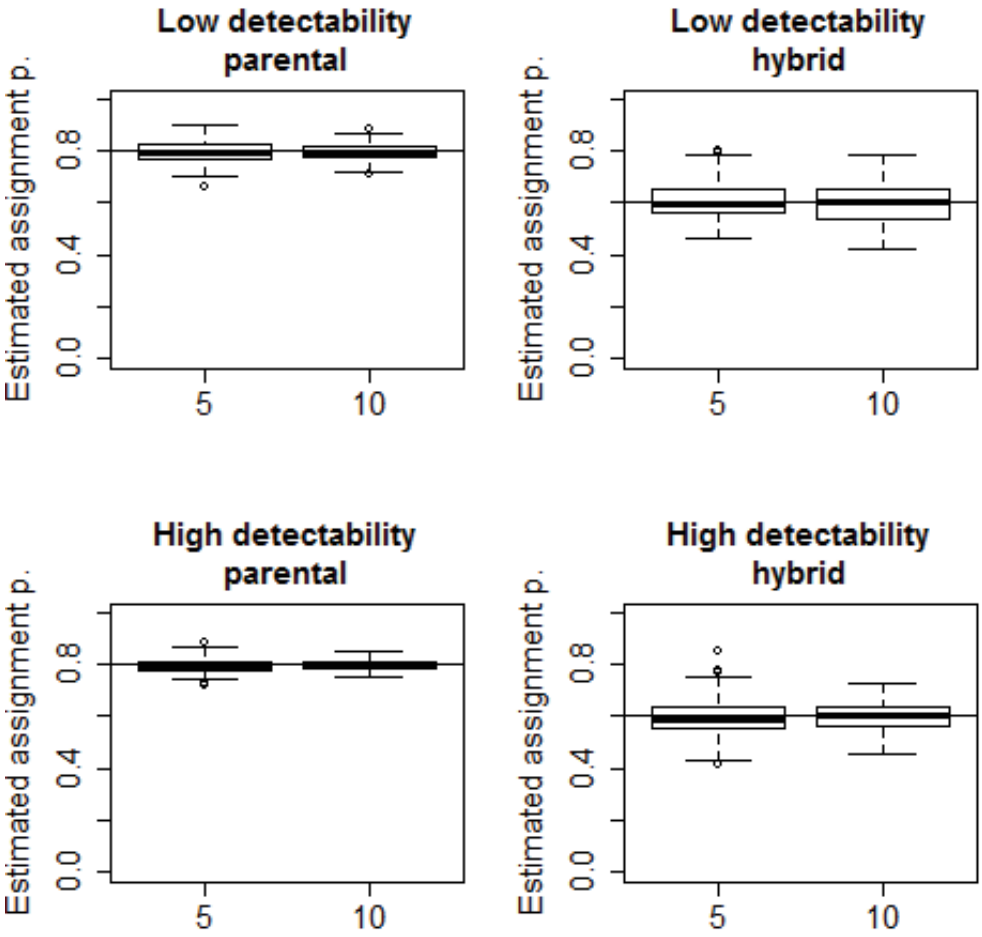


Figure S5. Scenario 2 (homogeneous detectability and state-dependent assignment probability). Boxplots of 100 simulated assignment probability estimates for parentals (left two panels) and hybrids (right two panels) for each sampling strategy. Sampling strategies with low detectability are in the top row, sampling strategies with high detectability are in the bottom row. Estimates obtained with sampling strategies with 5 and 10 capture occasions are compared in each panel.

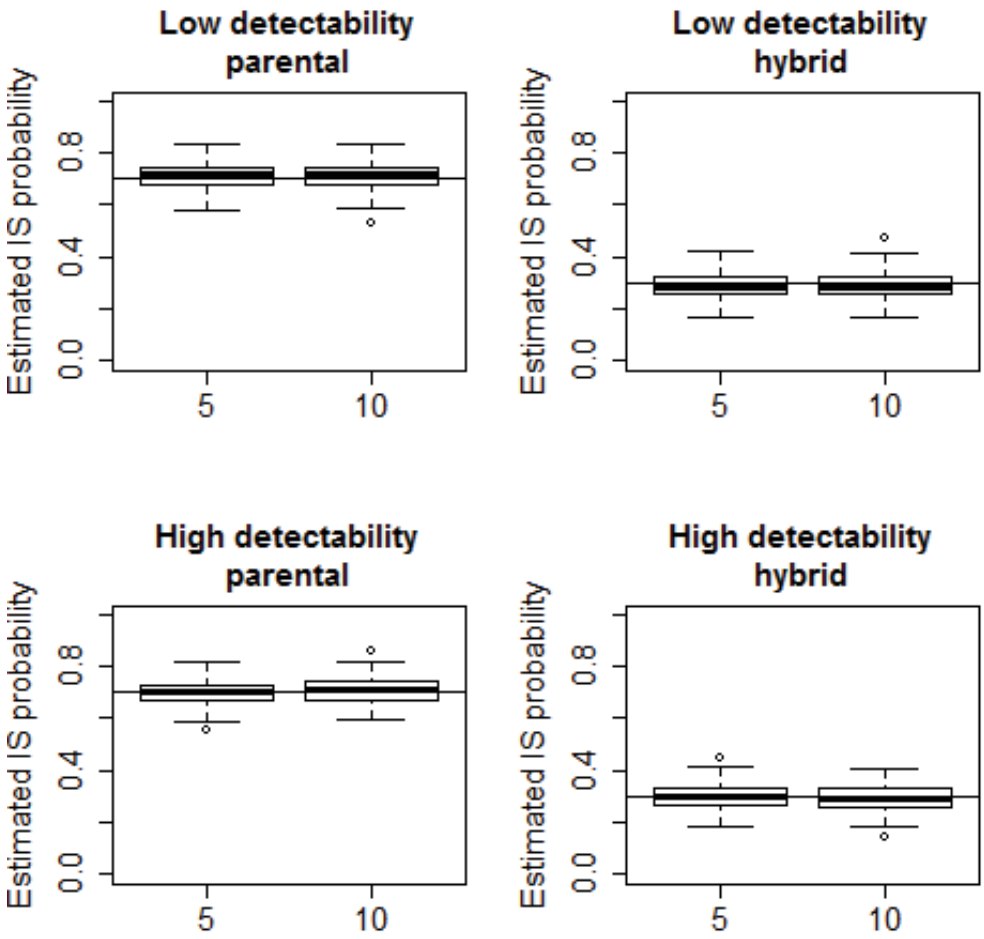


Figure S6. Scenario 2 (homogeneous detectability and state-dependent assignment probability). Boxplots of 100 simulated initial state (IS) probability estimates for parentals (left two panels) and hybrids (right two panels) for each sampling strategy. Sampling strategies with low detectability are in the top row, sampling strategies with high detectability are in the bottom row. Estimates obtained with sampling strategies with 5 and 10 capture occasions are compared in each panel.

3.4 Simulated parameter estimates for simulations scenario 3:homogenous detectablity and assignment probability

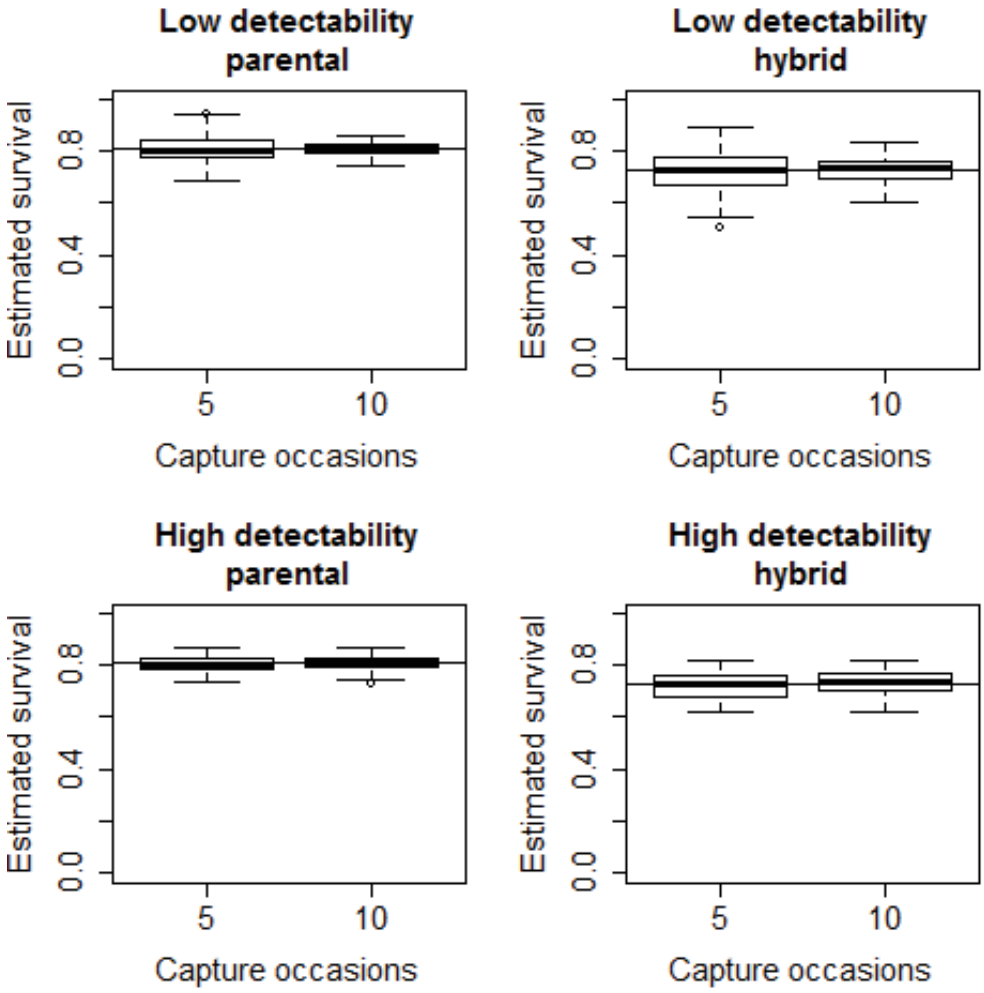


Figure S7. Scenario 3 (homogeneous detectability and assignment probability). Boxplots of 100 simulated survival estimates for parentals (left two panels) and hybrids (right two panels) for each sampling strategy. Sampling strategies with low detectability are in the top row, sampling strategies with high detectability are in the bottom row. Estimates obtained with sampling strategies with 5 and 10 capture occasions are compared in each panel.

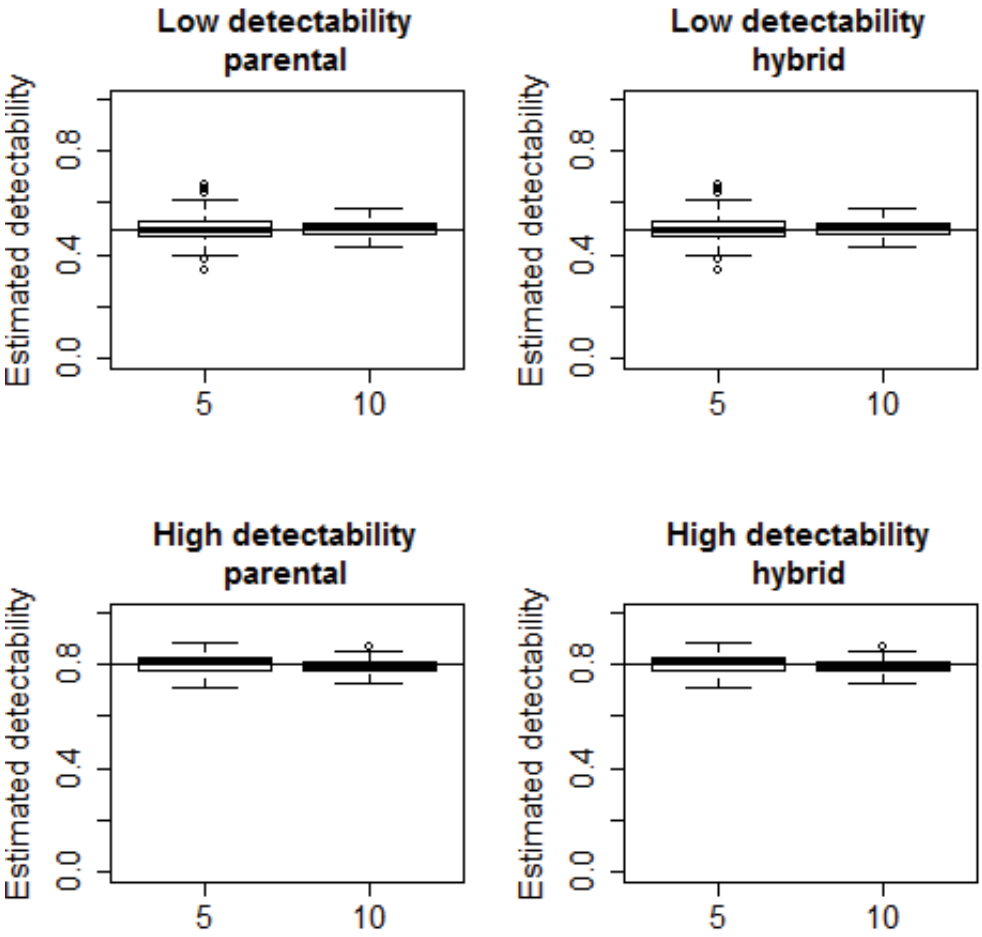


Figure S8. Scenario 3 (homogeneous detectability and assignment probability Boxplots of 100 simulated detectability estimates for parentals (left two panels) and hybrids (right two panels) for each sampling strategy. Sampling strategies with low detectability are in the top row, sampling strategies with high detectability are in the bottom row. Estimates obtained with sampling strategies with 5 and 10 capture occasions are compared in each panel.

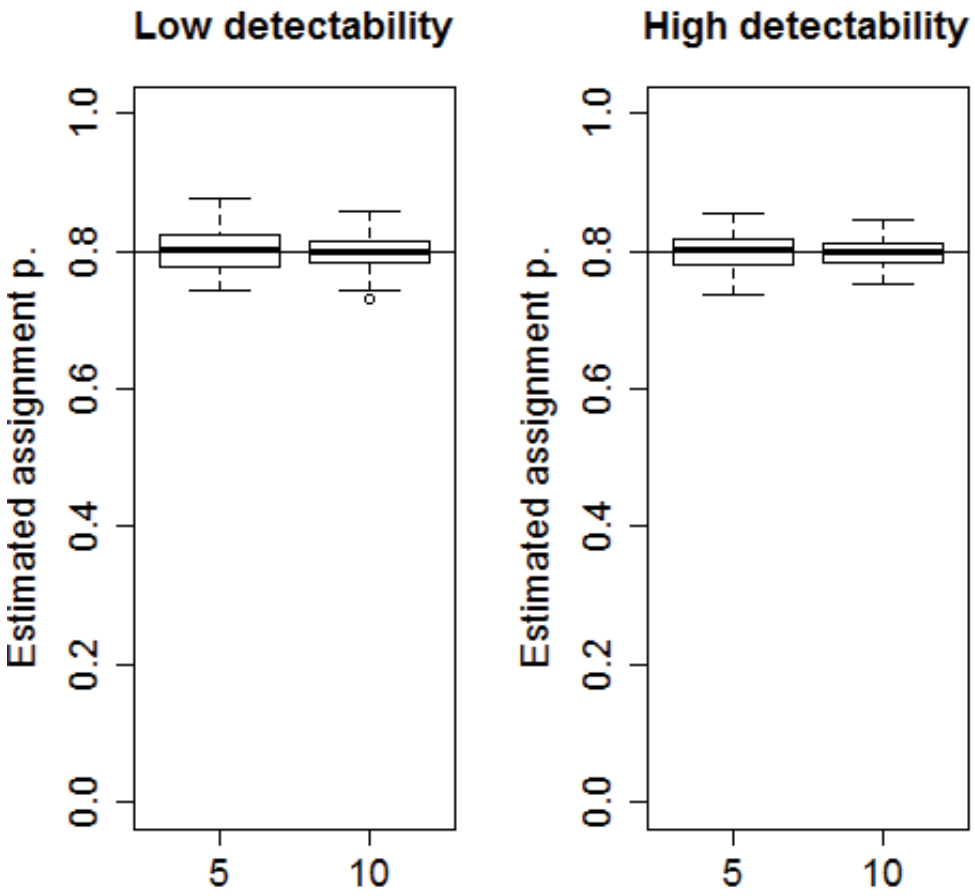


Figure S9. Scenario 3 (homogeneous detectability and assignment probability). Boxplots of 100 simulated assignment probability estimates. Sampling strategies with low detectability are on the left panel, sampling strategies with high detectability are on the right panel. Estimates obtained with sampling strategies with 5 and 10 capture occasions are compared in each panel.

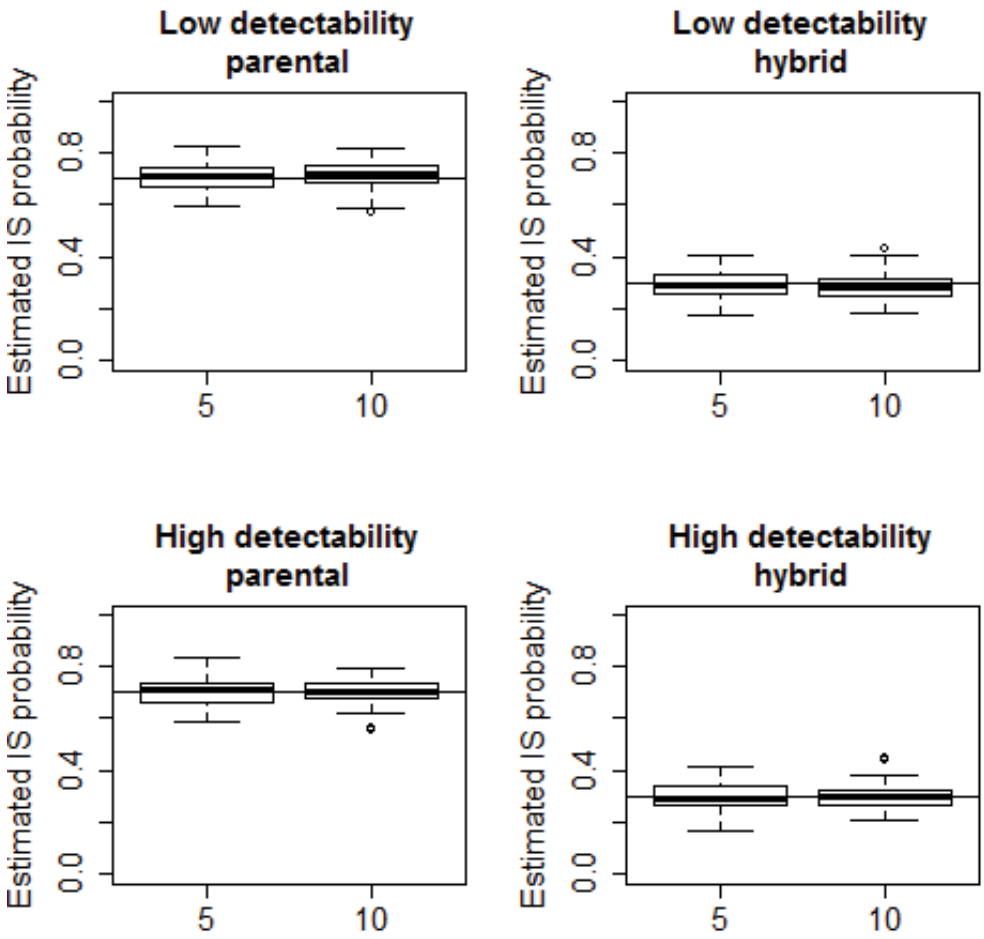


Figure S10. Scenario 3 (homogeneous detectability and assignment probability). Boxplots of 100 simulated initial state probability estimates for parentals (left two panels) and hybrids (right two panels) for each sampling strategy. Sampling strategies with low detectability are in the top row, sampling strategies with high detectability are in the bottom row. Estimates obtained with sampling strategies with 5 and 10 capture occasions are compared in each panel.

APPENDIX 2: Viterbi algorithm example

We observed the following capture histories of individuals that have been classified as Uncertains (coded as 3). Some are wolves, some are admixed individual. For example we will consider the following capture history:

	Occ1	Occ2	Occ3	Occ4	Occ5
Individual 1	3	3	0	3	3

From the estimates produced with the multievent model we know the following probabilities:

- **Initial state**: when capturing an individual for the first time there is a 0.60 (π_w) probability that it is a wolf and a 0.40 (π_h) probability that it is an admixed individual,

- **Transition probability:** wolves have a 0.72 (φ_w) probability of surviving and remaining in the study area, while admixed individuals have a 0.84 (φ_h) probability of surviving and remaining in the study area

- **Probabilty of detection (emission probabilities):** wolves have a 0.42 (p_w) probability of being detected, while admixed individuals have a 0.50 (p_h) probability of being detected.

The Viterbi algorithm establishes if the observed individual is more likely to be a wolf or an admixed individual by applying the estimated probabilities to the observed capture histories, for example:

The capture history of individual 1 (assuming that it is a wolf) is the following,

$$\pi_w \varphi_w p_w \varphi_w (1 - p_w) \varphi_w p_w \varphi_w p_w,$$

by applying the estimate probabilities we obtain:

(0.4)(0.72)(0.42)(0.72)(1-0.42)(0.72)(0.42)(0.72)(0.42) = 0.006.

The capture history of individual 1 (assuming that it is an admixed individual) is the following,

$$\pi_h \varphi_h p_h \varphi_h (1 - p_h) \varphi_h p_h \varphi_h p_h,$$

By applying the estimate probabilities we obtain:

(0.6)(0.84)(0.5)(0.84)(1-0.5)(0.84)(0.5)(0.84)(0.5) = 0.01.

Therefore Individual 1 is more likely to be an admixed individual

CHAPTER 2

Journal of Wildlife Management and Wildlife Monographs

Estimating admixture at the population scale: taking imperfect detectability and uncertainty in hybrid classification seriously

Journal:	*Journal of Wildlife Management and Wildlife Monographs*
Manuscript ID	JWM-20-0359
Wiley - Manuscript type:	Research Article
Date Submitted by the Author:	22-Oct-2020
Complete List of Authors:	Santostasi, Nina; University of Rome La Sapienza, Department of Biology and Biotechnology "Charles Darwin"; Centre d'Ecologie Fonctionnelle et Evolutive, Dynamique et Conservation de la Biodiversité Gimenez, Olivier; CEFE, CNRS, University of Montpellier, University Paul Valéry Montpellier 3, EPHE, IRD Camiglia, Romolo; Italian Institute for Environmental Protection and Research , Unit for Conservation Genetics (BIO-CGE) Fabbri, Elena; Italian Institute for Environmental Protection and Research , Unit for Conservation Genetics (BIO-CGE) Molinari, Luigi; Appennino Tosco-Emiliano National Park, Wolf Apennine Center Reggioni, Willy; Appenino Tosco Emiliano National Park, Wolf Apennine Center Ciucci, Paolo; University of Rome La Sapienza, Department of Biology and Biotechnologies "Charles Darwin"
Keywords:	Anthropogenic introgression, Bayesian assignment, Canis lupus, capture-recapture, genetic swamping, multi-event modelling, prevalence, wolf-dog hybridization
Abstract:	Introgressive hybridization between dogs and wolves represents an emblematic case of anthropogenic hybridization, and is increasingly threatening the genomic integrity of wolf populations expanding into human-modified landscapes. However, there is a lack of studies formally estimating prevalence and accounting for imperfect detectability and uncertainty in individual classification. 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. Individual multilocus genotypes were scored using a panel of 12 microsatellites and assigned to the reference wolf and dog populations through Bayesian clustering procedures. Based on 152 samples, our dataset comprised the capture histories of 39 individuals sampled in seven wolf packs and was organized in bi-monthly sampling occasions (Aug 2015□May 2016). We fitted CR models using a multi-event formulation to explicitly handle uncertainty in individual classification, and accordingly contemplated two model scenarios, one reflecting a traditional approach to classify individuals (i.e., minimizing

	the misclassification of wolves as hybrids; Type 1 error), and the other using a more stringent criterion aimed to balance between Type 1 and 2 error (i.e., the misclassification of hybrids as wolves). Compared to the naïve proportion of admixed individuals in the sample (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. The average dog introgression was 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 modelling 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 introgression. As anthropogenic hybridization is increasingly affecting animal species globally, our approach is of interest to a broader audience of wildlife conservationists and practitioners.

SCHOLARONE™
Manuscripts

22 Oct 2020

Nina L. Santostasi
University of Rome La Sapienza
Rome, Italy
nina.santostasi@uniroma1.it

RH: Santostasi et al. • Estimating proportion of admixture

Estimating admixture at the population scale: taking imperfect detectability and uncertainty in hybrid classification seriously

NINA L. SANTOSTASI[1] *Department of Biology and Biotechnologies "Charles Darwin", University of Rome La Sapienza, Rome, Italy; CEFE, CNRS, University of Montpellier, University Paul Valéry Montpellier 3, EPHE, IRD, Montpellier, France*
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[1], *Department of Biology and Biotechnologies "Charles Darwin", University of Rome La Sapienza, Rome, Italy*

[1] Email: nina.santostasi@uniroma1.it
[2] Email: paolo.ciucci@uniroma1.it

For Review Only

ABSTRACT

Introgressive hybridization between dogs and wolves represents an emblematic case of anthropogenic hybridization, and is increasingly threatening the genomic integrity of wolf populations expanding into human-modified landscapes. However, there is a lack of studies formally estimating prevalence and accounting for imperfect detectability and uncertainty in individual classification. 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. Individual multilocus genotypes were scored using a panel of 12 microsatellites and assigned to the reference wolf and dog populations through Bayesian clustering procedures. Based on 152 samples, our dataset comprised the capture histories of 39 individuals sampled in seven wolf packs and was organized in bi-monthly sampling occasions (Aug 2015–May 2016). We fitted CR models using a multi-event formulation to explicitly handle uncertainty in individual classification, and accordingly contemplated two model scenarios, one reflecting a traditional approach to classify individuals (i.e., minimizing the misclassification of wolves as hybrids; Type 1 error), and the other using a more stringent criterion aimed to balance between Type 1 and 2 error (i.e., the misclassification of hybrids as wolves). Compared to the naïve proportion of admixed individuals in the sample (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. The average dog introgression was 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 modelling 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 introgression. As anthropogenic hybridization is increasingly affecting

For Review Only

52 animal species globally, our approach is of interest to a broader audience of wildlife
53 conservationists and practitioners.
54
55 **KEYWORDS:** Anthropogenic introgression; Bayesian assignment**;** *Canis lupus*; capture-recapture;
56 genetic swamping; multi-event modelling; prevalence; wolf-dog hybridization.
57
58 Hybridization between domesticated forms and their wild ancestors is considered an exemplary
59 form of anthropogenic hybridization (Randi 2007). Although this phenomenon has repeatedly
60 occurred since domestication (Frantz et al. 2016; Pilot et al. 2018), there is concern that, especially
61 in human-dominated landscapes, the widespread occurrence of domesticated forms (Boivin et al.
62 2016) and their potential interactions with their wild ancestors may lead to an increased risk of
63 hybridization and gene flow (i.e., introgression; Randi 2007; Salvatori et al. 2020). This can
64 evetually lead the rarer wild counterparts to massive introgression and eventually genomic
65 extinction through swamping (Allendorf et al. 2001). Domesticated mammals in general exhibit a
66 suite of morphological, physiological and behavioral traits (e.g., coat color, craniofacial
67 morphology, dentition, ears and tail shape and length, more frequent and non-seasonal estrus cycles,
68 alterations in adrenocorticotropic hormone, increased tameness and sociality) that are not observed
69 in their wild counterparts (Wilkins et al. 2014). Although few examples of positive selection of
70 introgressed domesticated traits in wild species have been hypothesized (Anderson et al. 2009;
71 Coulson et al. 2011; Grossen et al. 2014), consistent gene flow beween domesticated forms and
72 their wild ancestors is expected to have deleterious consequences for wild species genomic integrity
73 and viability (e.g., reduction in fitness and adaptive potential, loss of unique combinations of genes
74 and genotypes that have a unique evolutionary history; Allendorf et al. 2001; Bohling 2016; Wayne
75 and Shaffer 2016). For these reasons hybridization between wild and domesticated forms is
76 considered a relevant threat to biodiversity (Todesco et al. 2016).

For Review Only

Admixture can be measured both in terms of degree of introgression (i.e., the proportion of alleles from a non-parental taxon averaged across individuals in the population; e.g., Miller et al. 2003) or proportion of admixed individuals in a population (*sensu* Allendorf et al. 2001); here, we refer to the latter (hereafter proportion of admixture, or prevalence) as 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 may inform wildlife management according to a two-step process: first, at the population scale, where assessments are often conveniently based on the genotyping of non-invasive samples, it informs if and eventually where management measures are needed (e.g., Adams et al. 2003); second, if population-wide surveys reveal admixture to an extent that necessitates interventions, more in-depth genetic analyses are needed to identify hybrids and accordingly target management actions (e.g., vonHoldt et al. 2013). In turn, estimating the proportion of admixture at the population-wide scale requires reliably estimating the abundance of both parental and admixed individuals in the population. These estimates, to be reliable, shoud ideally account for three 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 detectabilty 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 two 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 two interacting factors: the number and type of genetic markers used, and the statistical methods and options adopted to assign sampled

For Review Only

individuals to the parental or admixed reference populations (Bohling et al. 2013; Vähä and Primmer 2006). Because population-wise assessments are more efficiently conducted analyzing non-invasive samples, especially for elusive and treathened 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 method to assign sampled individuals to the parental or the admixed category is the Bayesian clustering procedures 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). In our study we carried out the assignment using STRUCTURE that 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 on the basis of the estimated membership proportion (q_i), that is the fraction of its genome that is inherited from ancestors in one of the two populations (Pritchard et al. 2000). Admixed individuals are then inferred when their estimated q_i value is intermediate between two clusters (e.g., 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 some rule of thumb of arbitrarily defined standard (e.g., q_i = 0.8 for Canids; Verardi et al. 2006; Rutledge et al. 2012; Wheeldon et al. 2010; Benson et al. 2012), or by making reference to q_i values derived from simulated genotypes of known genealogy (e.g., Godinho et al. 2011; van Wyk et al. 2016; Caniglia et al. 2020). We argue that more formal ways could be adopted to integrate the uncertainty in defining threshold q_i values, and therefore assign sampled individuals to the parental or the admixed categories, into the estimation process. Moreover, in wolf x 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

For Review Only

129 expanding wolf populations (Donfrancesco et al. 2019). Especially if hybrids are detected using a
130 limited number of genetic markers, calibrating threshod q_i values to obtain small Type 1 error rates
131 means that a non trivial proportion of recent backcrosses (*sensu* Caniglia et al. 2020) could be
132 erroneously classified as wolves. Nonetheless, backcrosses are an indication of ongoing
133 introgression and they carry domesticated alleles that, even if in smaller proportions compared to
134 F1s, pose a risk for the genomic integrity of the wolf parental population (Allendorf et al. 2001).
135 Here, we present a formal approach to estimate the proportion of admixed individuals by
136 making reference to a local wolf population affected by dog introgression in Italy. Hybridization
137 between gray wolves (*Canis lupus*) and dogs (*Canis lupus familiaris*) is an emblematic case of
138 wild-domestic hybridization (Butler 1994). Dog domestication originated about 35,000 – 11,000 ya
139 (Freedman and Wayne 2017). Due to intense artificial selection, dogs differentiated from wolves in
140 several morphological, physiological, and behavioral traits, and many of these differences are
141 genetically based (vonHoldt et al. 2017; Pendleton et al. 2018). Nonetheless, wolves and dogs are
142 interfertile and first generation hybrids can backcross into the wolf parental population, generating
143 gene flow between the two forms (Vilà and Wayne, 1999; Randi 2007). In Europe, several wolf
144 populations are currently re-expanding their range across human-dominated landscapes (Chapron et
145 al. 2014) where dogs have since long become the most abundant carnivore (Ritchie et al. 2014).
146 Accordingly, recent introgressive hybridization (i.e., up to three generations in the past; Caniglia et
147 al. 2020) has been lately detected in several wolf populations in Eurasia (Galaverni et al. 2017;
148 Pacheco et al. 2017; Pilot et al. 2018; Salvatori et al. 2020), suggesting the phenomenon is more
149 widespread than originally thought (Vilà and Wayne 1999; Randi et al. 2000). Detection and
150 monitoring of wolf-dog hybridization within wolf populations is therefore considered a
151 conservation prioritity at the European scale (Hindrickson et al. 2017; Donfrancesco et al. 2019;
152 Salvatori et al. 2020). However, no systematic and coordinated management of wolf-dog hybrids is
153 currently in place across Europe, also due to the inherent uncertainties in detecting hybrids and a
154 lack of *ad hoc* planned monitoring programs to formally assess hybridization at the population scale

For Review Only

(Salvatori et al. 2020). Previous assessments of wolf-dog hybridization in Europe (see Dufresnes et al. [2019] for a review) estimated prevalence as the sample proportion of individuals identified as admixed in the population (hereafter, naïve prevalence); in addition, 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 the light of the likely increasing phenomenon throughout Europe (Salvatori et al. 2020).

Based on previous methodological work (Santostasi et al. 2019), the goal of our study was to apply a formal approach to estimate the proportion of admixture that accounts for the sampling and estimation problems that affects naïve estimates of prevalence. By focussing on a wolf population in the northern Apennines (Italy), we applied a multi-event formulation of the CR model to formally address both imperfect detectability and the inherent uncertainty in the probabilistic assignment of admixed individuals (Pradel 2005; Jennelle et al. 2007; Santostasi et al. 2019). In doing so, we also explored more or less stringent rationales to classify admixed individuals (i.e., the current practice to minimize Type 1 error vs. a more precautionary approach to balance between Type 1 and 2 error rates) and the effect they had on the estimates of prevalence. We also used ancestry analysis and genealogy reconstruction to determine the reproductive status of admixed individuals within the studied population. Our findings, and the analytical approach we adopted, have broad implications for the assessment of the proportion of admixture in populations and species threatened by anthropogenic hybridization.

STUDY AREA

For Review Only

Our 731 km^2 study area is centered but extends beyond the Appennino Tosco-Emiliano national Park (PNATE), in the northern Apennines, Italy (Fig. 1). The area lays within the core of the wolf range in the northern Apennines and comprises the territories of seven wolf packs that we intensively surveyed using a combination of field techniques (i.e., wolf-howling during summer, snow-tracking in winter, GPS-telemetry, non-invasive genetic sampling, camera-trapping; Ciucci and Boitani, 1999; Caniglia et al. 2014; Ciucci et al. 2018). Although this wolf population can not 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 definition comprises a demographically and genetically cohesive wolf population. Rough topography with steep valleys and elevations characterize the area, and the vegetation is mainly composed by temperate and sub-Mediterranean deciduous forests, alternated with prairies, meadows, pastures and cultivated fields. Human presence (25 inhabitants/km^2) is limited throughout the year to lower elevations (below 1000 m), although it increases during summer due to tourism and livestock grazing at higher altitudes (Ciucci et al. 2003). 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 density (Caniglia et al. 2014) thriving on wild and occasionally domestic ungulates (Ciucci et al. 1996). Although stray dogs are scarce in the area, uncontrolled working or hunting dogs may be temporarily present. Occasional reports of admixed individuals have been previously reported in the study area (Caniglia et al. 2014) and led to the establishment of a EU-funded project (LIFE13 NAT/IT/000728) to further investigate the impact of recent hybridization on the local wolf population.

METHODS

Sampling methods

For Review Only

206 During 2015–16, we applied a non-invasive genetic sampling by systematically collecting wolf
207 scats every 2 weeks along fixed routes throughout each of 7 wolf pack territories, and used
208 complementary collection methods (i.e., at scent posts, homesites, kill sites) to boost sample size
209 and enhance sampling coverage. To ensure a demographically meaningful time frame over which to
210 estimate prevalence of admixed individuals, we restricted sampling to a single breeding pulse (i.e.,
211 from pup rearing until the onset of the next breeding season) by analysing wolf scats collected from
212 August 2015 to May 2016, excluding those from pups born in spring 2016. Upon collection, we
213 individually stored faecal samples at -20°C in 10 volumes of 95% ethanol. We then extracted and
214 amplified DNA from the scats using standard laboratory protocols.
215
216 **Detection and classifications of individuals**
217 Based on a multiple-tube protocol (Taberlet et al. 1996) using procedures described in Fabbri et al.
218 (2018), we genotyped faecal DNAs amplifying them at: a) 12 unlinked autosomal microsatellites
219 (STRs) selected for their polymorphism and reliable scorability for wolves and dogs (Caniglia et al.
220 2014) and routinely used for genotyping low-content DNA samples in non-invasive genetic
221 monitoring projects (Caniglia et al. 2013, Caniglia et al. 2014, Fabbri et al. 2018), and b) a
222 dominant 3-bp deletion (named KB or CBD103DG23) of the b-defensin CBD103 gene (the K-
223 locus; Anderson et al. 2009; Caniglia et al. 2013). Additionally, we sexed samples by PCR-RFLP of
224 the ZFX/ZFY (zinc-finger protein) sequences (Lucchini et al. 2002) and identified paternal
225 haplotypes typing 4 STRs located on the Y Chromosome (MS34A, MS34B, MSY41A and MS41B;
226 Sundqvist et al. 2001) and maternal haplotypes analyzing 250 bp of the hypervariable domain of the
227 mtDNA CR1 (Caniglia et al. 2013). We used the software Gimlet v.1.3.3 (Valière 2002) to
228 reconstruct the consensus genotype from the results of the 4-8 replicated amplifications per locus, to
229 estimate PCR success (the number of successful PCRs divided by the total number of PCR runs
230 across samples), allelic drop-out (ADO) and false alleles (FA) rates and to match the detected
231 genotypes to each other and to the ISPRA *Canis* database for the identification of possible

For Review Only

resemplings in he 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 (Besnier and Glover, 2013) on the 12-loci multilocus reliable genotypes obtained from the 4–8 replicated amplifications per locus per sample using the R package Parallel Structure (Caniglia et al. 2020). For each individual genotype we estimated the individual proportions of membership (q_i) and the 90% Bayesian Credible Intervals (BCI) to the two inferred clusters (details about the Bayesian assignment test models are reported in Caniglia et al. 2020). We used the "Admixture" and the "Indipendent Allele Frequencies" models, with 500,000 Markov Chain Monte Carlo iterations discarding the first 50,000 as burn-in. We used as reference populations a panel of 190 wolves and 89 dogs. Wolf samples were obtained from areas with no documented cases of admixture and: (*i*) showed the typical wild coat colour pattern and no apparent signal of other morphological or dog-like traits (e.g., black coats, white claws, spur on the hind legs), (*ii*) did not share dog derivated Y and mtDNA haplotypes, and (*iii*) had a q_i lower than 0.990, obtained in previous Bayesian assignment procedures performed using 156K canine SNPs and 39 canine STRs commonly used in recent studies on wolf x 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 one male and one female randomly chosen from 14 wolf-sized dog breeds available from LUPA project data set (Lequarre et al. 2011; Vaysse et al. 2011, Caniglia et al. 2020). The reference populations were used also in HybridLab (Nielsen et al. 2006) to simulate 100 genotypes for each of the following parental and admixed classes: wild (PW) and domestic (PD) parentals, first (F1) and second (F2) generation hybrids, and four backcross generations originated either from F1s (BCW1-BCW4) or F2 (SBCW1-SBCW4) crossing with wild parentals (Caniglia et al. 2020).

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 contemplated two classification scenarios to gauge their effect on the final estimate of

For Review Only

prevalence. The first classification (scenario A) reflected the widely adopted procedures essentially meant to reduce Type 1 error rates (e.g., Randi 2007; Godinho et al. 2011; Pacheco et al. 2017; Dufresnes et al. 2019). We acccordingly used the threshold $q_w \leq 0.975$, as suggested by Caniglia et al. (2020) to identify recent 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 three categories: (*i*) *Wolves*, including individuals whose q_w was higher than the maximum q_w of simulated first generation backcrosses (BCW1), as this prevented erroneously assigning first-generation backcrosses to the wolf category (i.e., Type 2 error); (*ii*) *Admixed individuals*, including those whose q_w was lower than the minimum q_w of reference wolves, as this prevented erroneously assigning wolves to the admixed category (i.e., Type 1 error); (*iii*) *Uncertains*, including individuals whose q_w was comprised between the minimum q_w of reference wolves and the maximum q_w of simulated BCW1 (Fig. 2; Table S1 Supplementary Materials); this reflected what suggested by Caniglia et al. (2020) who considered also a second *q*-threshold of 0.990 to identify admixed individuals of older backcross generations that share only a marginal dog ancestry. To this aim we used the minimum q_w of reference instead of simulated wolves as the former more realistically represents the expected genetic variability in the wolf population (Dufresnes et al. 2019). Whereas the first two categories are by definition those traditionally recognized in Bayesian-based assignment tests (i.e., scenario A), we contemplated the category *Uncertains* only in the classification scenario B (see below).

For each of the two 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 capture-recapture models to estimate the abundance of wolves and admixed individuals. With

For Review Only

scenario B we accounted for the uncertainty in the probabilistic assignment of admixed individuals (Santostasi et al. 2019).

Capture-recapture modelling and prevalence estimation

The multi-event formulation of open population CR models explicitly handles uncertainty in individuals classification by modelling the observed capture histories as two time series: the state process (i.e., the population dynamics during the study) and the event process (what we are able to observe through sampling; Pradel 2005). We modeled the state process as a Markov chain of 3 partially hidden states (Pradel 2005; Santostasi et al. 2019): alive in the study area as wolf, alive in the study area as admixed, and dead/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, since individuals could not change their state between parental and admixed, the only possible transition was between "in the study area" and "permanently emigrated/dead", therefore their state changed over time according to a first-order Markov process governed by only the apparent survival probabilities (Santostasi et al. 2019). We modelled 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 two consecutive steps: detection (p is the probability of finding and successfully 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 to assign an individual to one of the two 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 three possible events, corresponding to individuals that

For Review Only

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. Differently, in scenario B we accounted for uncertainty in the assignment procedure contemplating four 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 Admixed, recorded as 2; 3) detected but not classified (i.e., classified as Uncertain; recorded as 3); 4) not detected, recorded as 0. Under this classification scenario, δ was not assumed to be equal to 1 but became a parameter to be estimated, and the complementary probability (1-δ) was used 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 encountered at the first and third occasions but not at the second occasion. The state of this individual (i.e., Wolf or Admixed) is not assigned in this phase. Assuming parameters are constant, we have (Santostasi et al. 2019):

$$\Pr(303) = \pi_h(1-\delta_h)\varphi_h(1-p_h)\varphi_h p_h(1-\delta_h) + \pi_p(1-\delta_p)\varphi_p(1-p_p)\varphi_p p_p(1-\delta_p)$$

The likelihood of the entire dataset is obtained as the product of the probabilities of all individual encounter histories assuming independence, and the parameter estimates are obtained by maximum likelihood (Pradel 2005).

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

$$\hat{N} = \frac{n_p}{\hat{p}_p} + \frac{n_h}{\hat{p}_h} = \hat{N}_p + \hat{N}_h$$

where n_p was the number of parental individuals and n_h the number of hybrids detected, and $\hat{p}_p$ and $\hat{p}_h$ were the estimated detectabilities of parental and admixed individuals, respectively. Prevalence was then estimated as (Santostasi et al. 2019):

For Review Only

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

With scenario B, the individuals classified as uncertain needed to be assigned to one of the two states (admixed or parental). To do so, we used the Viterbi algorithm which, given any observation sequence (i.e., the encounter histories in this case) and the parameters estimated by the multi-event model, finds the most probable underlying sequence of states that has generated the observed data (Rouan et al. 2009; Zucchini et al. 2016; Santostasi et al. 2019). In our case the algorithm used the capture histories formulae of the uncertain individuals if they were wolves or admixed and caculated which of the two alternatives was the most probable, using the parameter estimates obtained with the multi-event model. We ran the Viterbi algorithm and calculated the probability of the four most likely capture histories reconstructions for each uncertain individual. Then, once the uncertain individuals were assigned to their most likely state and the number of observed parental and admixed individuals was reconstructed, their respective abundance estimates were used to estimate prevalence. Standard deviation and confidence intervals of the abundance and prevalence estimates were estimated via a non-parametric bootstrap (Davison and Hinkley, 2002).

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 (i) 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), (ii) time (capture occasion) effect on detection probability to test for heterogeneity due to variation in sampling effort and/or environmental conditions, (iii) 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 time effect on apparent survival because our sampling was extended over a relatively short time frame. Since the classification of genotypes was performed only once for each genotype, we constrained the assignment probability to be estimated upon first

For Review Only

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 their AICc (Akaike information criterion corrected for small sample size). To account for uncertainty in the selection of a single best model, we obtained model-averaged estimates considering models whose ΔAICc 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 has the same structure but does not allow uncertainty in state assignment (Gimenez et al. 2012).

Ancestry analysis and geneology

To reconstruct the ancestry of the sampled individual genotypes and to estimate their individual posterior probability to belong to the assigned or the other parental population, or to have 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; see Supplementary Materials for further details). In addition, to estimate the genealogies of the pack members and the reproductive status of the admixed individuals, we used a maximum-likelihood approach implemented in COLONY 2.0 (Wang and Santure 2009; see Supplementary Materials for further details).

RESULTS

Out of a total of 152 collected scats, 65% were reliably genotyped showing an average number of positive amplifications per locus of 0.77 (ranging from 0.39 to 0.92), and average error rates of ADO=0.18 (SD=0.11) and FA=0.05 (SD=0.03). Consensus genotypes were grouped and assigned to 39 individuals (15 females, 21 males, 3 of unknown sex; Table S2 Supplementary Materials), sampled on average (±SD) in 1.6 ± 0.9 capture occasions (Table S3 Supplementary Materials). We sampled from 2–12 individuals in each of the 7 packs. All the 39 sampled

For Review Only

individuals shared the typical Italian wolf mtDNA and none had the K^b melanistic deletion; four individuals shared a dog-derived Y haplotype, all assigned to the wolf category under scenario A while, under scenario B, three of them were classified as admixed and one as a wolf (Table S2 Supplementary Materials).

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 Supplementary Materials), corresponding to a naïve prevalence of 43.6% and an average degree of introgression of 7.8 (± 10.5 SD)% (95% CI: 4.4–11%). Based on the simulated genotypes, this q_w-threshold correctly identified 100% of simulated parental wolves, 100% of the F1, 100% of the F2, 86% of the BCW1, 48% of the BCW2, 13% of the BCW3, 8% of the BCW4 (Table S4 Supplementary Materials). The most supported model contemplated 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 (πw) and a 0.44 (95% CI: 0.59–0.29) 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 $\varphi_H = 0.92$ (95% CI: 0.38–0.99) for admixed individuals. Overall survival over the 10-months period (the product of the four bi-monthly estimates of apparent survival) was 0.10 (95% CI: 0.009–0.39) for wolves and 0.78 (95% CI: 0.02–0.96) for admixed individuals. We estimated a probability of detection p = 0.48 (95% CI: 0.29–0.68). Within each capture occasion, total population abundance ranged from 25 (95% CI: 10–44) to 31 (95% CI: 17–50) individuals, comprising 10 (95% CI: 3–26)–17 (95% CI: 7–30) wolves and 13 (95% CI: 3–21)–19 (95% CI: 9–32) admixed individuals (Table 2). Estimated prevalence under scenario A therefore ranged 47 (95% CI: 22–73)–60 (95% CI: 22–80)% 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%.

For Review Only

Conversely, according to classification of 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 Supplementary Materials). 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) were correctly classified. However, 23% of BCW2 (n = 79) were erroneously classified as wolves, a Type 2 error rate that increased to 43–83% for further generations of backcrosses (Table S5, Supplementary Materials). The most supported models contemplated constant detection and apparent survival probabilities, state dependent survival, or both 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 (πw) 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. Overall survival 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 probability of detecting a wolf as $p_W = 0.43$ (95% CI: 0.16–0.67) and 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 to be assigned to the category Admixed (range: 0.31–0.80) than to the category Wolf (0.09–0.31; Table S6 Supplementary Materials). Within each capture occasion, total population abundance ranged from 26 (95% CI: 13–78) to 35 (95% CI: 17–109) individuals, comprising 6 (95% CI: 2–20)–10 (95% CI: 4–35) wolves and 17 (95% CI: 8–48)–26 (95% CI: 13–78) admixed individuals (Table 2). Estimated prevalence under scenario B therefore ranged 64 (95% CI: 38–72)–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). Under scenario B, the proportion of admixed packs was 85.7%.

For Review Only

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 S7 Supplementary Materials).

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 ($0.120 \leq p \leq 0.999$) of having a grandparent in the dog population, hence of being first-generation backcrosses (Table S5, Appendix A). The remaining admixed individuals likely originated from backcrosses of further generations backwards (i.e., ≥ BCW2). Despite a relatively limited sample size, through the genealogy estimation we were able to identify ($p > 0.90$) likely admixed breeding pairs in two of the seven surveyed packs. Specifically, according to scenario A, the two admixed breeding pairs had one admixed breeder each, whereas, based on scenario B, one of the two pairs had two and the other one admixed breeders (Fig. 4). All of the admixed breeders were identified to be > first-generation backcrosses in the ancestry analysis (Table S8 Supplementary Materials). In 2015, these two admixed breeding pairs produced a minimum of 3 and 5 offspring each. Under scenario A 7 of these offspring were classified as Wolf and one as Admixed, whereas according to scenario B only 1 of them was classified as Wolf, 3 as Admixed, and 4 as Uncertain. Based on the Viterbi algorithm, the 4 individuals originally classified as Uncertain were classified as Admixed (Table S7 in Supplementary Materials).

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

For Review Only

460 detectability and uncertainty in individuals classification, both typically associated to the
461 genotyping of non-invasive samples. Differently, naïve estimates of prevalence do not take into
462 account detectability, and previous simulations demonstrated they are biased when the probability
463 of detection of parental and admixed individuals is heterogeneous (Santostasi et al. 2019).
464 Heterogeneity in detectability may arise because of expected differences in ecology and behavior
465 between parental and admixed individuals (e.g., social behavior, marking rate, spatial patterns,
466 resource selection), or because of sampling bias (e.g., spatially heterogeneous effort, edge effect;
467 Crespin et al. 2008). Although we did not find strong evidence of such heterogeneity in our
468 application, under scenario B high-ranking models did contemplate the effect of heterogeneous
469 detectability, suggesting that admixed individuals may have a higher probability of detection.
470 A critical point, common to both naïve and formal approaches to estimate prevalence,
471 concerns the reliability of the individual classification based on Bayesian clustering techniques. In
472 fact, in addition to the number and type of molecular markers, the classification is strongly affected
473 by the composition of the chosen reference samples and by the threshold q_i values chosen to
474 discriminate admixed from parental individuals (Vähä and Primmer 2006). To overcome these
475 problems, we relied on an improved procedure (Caniglia et al. 2020), based on (*i*) carefully
476 screened reference samples of pure Italian wolves, whose selection was aided by genomic tools
477 (Galaverni et al. 2017), and (*ii*) a back-end executable of the software STRUCTURE (i.e., R package
478 Parallel Structure) that produces more stable assignment coefficients which are not affected by
479 samples with variable levels of admixture as these are analysed one by one (Besnier and Glover
480 2013). Previous studies comparing the assignment and the detection power of admixed individuals
481 of several Bayesian softwares showed that the results obtained by STRUCTURE are comparable to
482 those obtained using other Bayesian approaches, such as those implemented in NewHybrids
483 (Caniglia et al. 2020) or BAPS (Randi et al. 2014), and better than partially Bayesian approaches
484 like those implemented in GeneClass (Piry et al. 2004, Sanz et al. 2009). We nevertheless used a
485 relatively limited number of microsatellite loci, though this reflects what is routinely done in

For Review Only

population-wide surveys adopting non-invasive genetic sampling (Caniglia et al. 2014; Fabbri et al. 2018). This implies a relatively reduced power to detect admixed individuals of second or further generations of backcross. In these circumstances, the traditional assignment approach is to define a threshold q_i value that comprises all wolves of the reference poulation, with the aim 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. However, similarly to other conservation contexts (McGarvey 2007; Saltz, 2011), it is also relevant to consider Type 2 error rates; specifically, an overlooked occurrence of admixed individuals, expecially 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 (*i*) adopting a multi-event formulation in CR modelling that more formally accounts for the uncertainty in detecting hybrids while balancing Type 1 and 2 error rates, and (*ii*) 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 (*sensu* 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 (i.e., using only the first q_w-threshold suggested by these authors and not taking into account the assignment interval including older admixed individuals) may still overlook a non trivial proportion of recent admixed individuals in the population, therefore understimating prevalence. Notably, the

For Review Only

512 approach suggested by Caniglia et al. (2020) had different aims than the formal estimation of
513 prevalence, and it was developed to provide a convenient and practical screening of samples while
514 standardizing assignment procedures across different genetic laboratories. Based on scenario B,
515 however, the consideration of individuals whose q_w value was comprised between the minimum q_i
516 of reference wolves and the maximum q_w of simulated first-generation backcrosses allowed us to
517 formally integrate the uncertainty in hybrid classification into the estimation process while
518 controlling for Type 2 error rates. In particular, the Viterbi algorithm offered a reproductible
519 procedure to assign the individuals whose classification was uncertain. In our study, 6 individuals
520 were classified as uncertain and, according to the algorithm, they all had a higher probability of
521 being admixed than wolves. Notably, because this assignment is based on the estimates of the
522 model parameters, it may be flawed if the estimates are biased. However, Santostasi et al. (2019)
523 assessed the performance of the multi-event CR model in estimating parameters (π, φ, p, δ) and
524 prevalence under different sample sizes and concluded that the model estimates are unbiased with
525 the value of detectability reported in our study. To further support the Viterbi algorithm-based
526 assignment, 4 of the 6 individuals originally classified as uncertain were confirmed to be the
527 progeny of two admixed individuals by the genealogical reconstruction. This leads us to suggest
528 that both the Viterbi algorithm and the genealogical reconstruction should be used as independent
529 methods to classify individuals that, based on their q_w value, are of uncertain assignment.
530 Expectedly, based on scenario B, the average prevalence of admixture in the wolf population raised
531 to 70%, no simulated wolf was erroneously assigned to the admixed category, no simulated first-
532 generation backcross, and only 23% of simulated second-generation backcrosses were erroneously
533 assigned to the the wolf cluster. Cumulatively, and limited to recent backcrosses, this corresponds
534 to a 65.2% decrease in Type 2 error rate, and no increase in Type 1 error rate, compared to scenario
535 A.

536 The above estimates of prevalence, ranging from 43.6% (naïve), to 50% (95% CI: 22–80%;
537 scenario A) and 70% (45–77%; scenario B), consistently reveal widespread admixture at the level

For Review Only

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 as, out of the 8 pups produced by the two admixed breeding pairs, 7 were identified as admixed under scenario B but only 1 under scenario A. Indeed, a 70% estimate of prevalence in a protected wolf population may cause some raised eyebrows, especially if this conclusion is based on a relatively small number of genetic markers. However, two considerations are in place here: first, being this assessment based on non-invasive samples genotypings, and hence a limited number of genetic markers, by no means our findings imply that 70% of the poulation has to be managed (e.g., captivated or sterilized) to contrast 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 works, 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. It could be argued that, regardless the proportion of admixture, individuals with an average 7.8% of dog introgression (95% CI: 4.3–11%) might not differ from the parentals in their ecological role, and that they can be considered as ecological surrogates (Wayne and Shaffer 2016; vonHoldt et al. 2017). However, we believe this argument should not interfere with the analytical capability to reliably assess proportion of admixture at the population scale. In absence of clear evidence on the ecological signature of introgressed individuals, a formal and precautionary approach should still be used to assess the proportion of admixture in wild populations.

Compared to previous estimates of wolf-dog hybridization (e.g., Randi and Lucchini 2002; Lorenzini et al. 2013; Kusak et al. 2018; Dufresnes et al. 2019), our estimate is based on a biologically more meaningful sample, as it was limited to a local wolf population and temporally

For Review Only

564 restricted to one breeding period. Hybridization is a highly dynamic phenomenon because
565 facilitating factors, such as lack of conspecifics, availability of heterospecific mates, anthropogenic
566 disturbance (Bohling and Waits, 2015), as well as movements of admixed individuals, may vary
567 considerably across time and space. Differently from ours, multi-generation samples pooled across
568 large areas and time frames, may impair a clear understanding of admixture dynamics in space and
569 time. On the other hand, our estimate has local significance and cannot be used to extrapolate
570 general conclusions about wolf-dog introgression at wider scales. This underlines the importance of
571 upscaling our sampling and estimation approach over larger areas and multiple timeframes, as
572 obtaining reliable assessment of the number of admixed populations and their prevalence over time
573 will inform the most adequate management strategies (Allendorf et al. 2001).
574 Regardless the scenario we adopted, the proportion of admixture we estimated is, to our
575 knowledge, the highest so far reported for a wolf population (summarized in Table 1 in Dufresenes
576 et al. [2019]; but see Salvatori et al. [2019]), and indicates that unmanaged wolf-dog hybridization
577 may lead to widespread introgression at local scale. Based on relatively low, naïve quantifications
578 of prevalence from previous hybridization studies (e.g., Verardi et al. 2006, Pacheco et al. 2017;
579 Kusak et al. 2018), several authors concurred on two main hypotheses: first, wolf-dog hybridization
580 is rare, and where it occurs most likely takes place in the peripheral portion of the wolf distribution
581 (Lorenzini et al. 2014) and/or during early phases of range expansion (Galaverni et al. 2017; Kusak
582 et al. 2018); second, introgression of dog alleles into wolf populations is expected to be buffered by
583 behavioural and selective constraints (e.g., the unsuccessfull integration of pregnant admixed
584 females in the natal packs, the reduced survival of F1 litters due to limited paternal care, the lower
585 success of admixed individuals in territorial or predatory interactions; Vilà and Wayne 1999), or by
586 dilution of dog genes through backcrossing into the parental wolf populations (Verardi et al. 2006).
587 However, the high proportion of admixture we reported contradicts both predictions and contrasts
588 with theoretical expectations on the functionality of reproductive barriers between wolves and dogs.
589 Although we did not detect any F1, we estimated a high prevalence of backcrosses of several

For Review Only

590 generations backwards, 30% of which had non-negligible probabilities of being first- or second-
591 generation backcrosses, indicating a time of original admixture dating as recently as 2006–2009
592 (i.e., 6–9 years before our sampling assuming a 3 years-wolf generation time; Skoglund et al. 2011).
593 Reports of F1 hybrids are rare both in the northern Apennines (Caniglia et al. 2014) and elsewhere
594 where wolf-dog hybridization has been detected (Godinho et al. 2011; Randi et al. 2014; Pilot et al.
595 2018; Salvatori et al. 2019). However absence of detection cannot be considered true absence when
596 detectability is not taken into account (MacKenzie 2005). Moreover, even if wolf x dog
597 hybridization events are rare, the production of admixed individuals by backcrossing may increase
598 exponentially when reproductive barriers are weak (one fertile F1 female can produce several
599 backcrossed litters; Fredrickson and Hedrick 2006; Santostasi et al. 2020). However, it is
600 noteworthy that in 2006 a male F1 hybrid, featuring a Y-haplotype typical of Czechoslovakian
601 wolfdogs, was sampled in one of our study packs. This hybrid successively sired two admixed
602 litters with a female wolf, as revealed by genealogical reconstruction of seven pups non invasively
603 sampled in 2010 and 2011 (R. Caniglia, pers. comm.). Therefore, crossmating between wolves and
604 dogs in the northern Apennines was not limited to the wolf re-colonization phase (i.e., late 1990s;
605 Galaverni et al. 2017), and our findings reveal that factors faciltating hybridization are currently
606 operating in a long-established wolf populations well after the re-colonization phase.
607 Anthropogenic disturbances may facilitate the disruption of pre-mating (e.g., agonistic behavior)
608 and post-mating (e.g., reduced fitness of F1) reproductive barriers in social canids (Rutledge et al.
609 2012; Bohling and Waits, 2015). In particular, human-caused wolf mortality may disrupt breeding
610 pairs, thereby increasing the chances of pack dissolution (Brainerd et al. 2008). Especially during
611 the breeding season, pack dissolution may release the social inhibition to mate of subordinate
612 wolves of both sexes, stimulating individuals to find mates and establish in vacant territories
613 (Bohling and Waits 2015). In human-dominated countries where free-ranging dogs are widespread
614 (Ritchie et al. 2014), or admixed individuals could be available as mates, the above mechanism may
615 contribute to greatly increase the chances of hybridization and backcrossing (Bohling and Waits

For Review Only

616 2015). In the same region of our study area, Caniglia et al. (2014) reported a high pack turnover,
617 possibly linked to the high mortality of wolf territory holders. The territories of the packs we
618 surveyed extend well beyond the boundaries of the PNATE, where hunting is allowed throughout
619 the year for wild boar population control, and instances of illegal wolf killings during drived hunts
620 with dogs are routinely reported among residents (L. Molinari, pers. comm.). This is in line with our
621 estimates of apparent survival that are considerably lower compared to other studies (e.g., Marucco
622 et al. 2009; Cubaynes et al. 2010; Caniglia et al. 2012). In addition, according to both scenarios A
623 and B apparent survival seems to be lower for wolves than for admixed individuals, a difference
624 that, if confirmed, would be key in promoting the spread of introgression (Rutledge et al. 2012;
625 Bohling and Waits 2015). We acknowledge that estimates of apparent survival can be also
626 negatively biased by transience generated by dispersing yearlings and subordinate adults (Jimenez
627 et al. 2017), but high disappearance rather than transience rates characterized the wolf population in
628 this region at a wider scale (Caniglia et al. 2014). We are therefore inclined to believe that human-
629 caused mortality is likely among the ultimate causes of the high introgression rates we detected.
630 Anthropogenic food sources (i.e., large livestock carcasses, butchery offals) are largely availabile in
631 some portions of our study area (L. Molinari, pers. comm.), and these may promote affiliative
632 interactions between solitary female wolves and dogs (Newsome et al. 2017) and also facilitate the
633 survival of solitary pregnant females and their admixed litters. The occurrence of at least two
634 admixed breeding pairs are further evidence of the reproductive success of admixed individuals. In
635 these conditions, hybrid-hybrid pairs can maintain dog genes at high frequency in the population
636 (see also Bassi et al. 2017; Salvatori et al. 2019), as also projected by simulation of hybridization
637 dynamics in social canids (Fredrickson and Hedrick, 2006; Santostasi et al. 2020).

638 We can not exclude that high admixture in our study area may have also originated through
639 dispersal of admixed individuals from other areas. Wolves, and expectedly so introgressed
640 individuals, can travel long dispersal distances from their natal territories also across human-
641 dominated countries (Ciucci et al. 2009). In an area located at the periphery of the wolf range in

For Review Only

central Italy, one admixed breeding pair produced 2–6 pups each year from 2005–2008, and some of the offsprings were successively 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 limited introgression has been so far currently detected (Dufresnes et al. 2019). Notably, one 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, pers. comm.).

MANAGEMENT IMPLICATIONS

In line with other wild-domesticated hybridization cases (e.g., Lecis et al. 2006; Stephens et al. 2015), our findings indicate that reproductive isolation between wild and domestic forms, and possibly dilution by backcrossing, may be not sufficient per se to prevent widespread introgression of domestic genes. While futher research is needed to better gauge the spatial and temporal dynamics of introgression over time, effective management of anthropogenic hybridization is of paramount importance (Allendorf et al. 2001). Depending on the extent of admixture, this could be achieved through preventive (e.g., education, information, communication) and proactive measures (e.g., effective control of facilitating factors such as free ranging dogs, poaching, anthropogenic food provisioning). The latter should be aimed in our case to enhance survival of wolf breeders and preserve the cohesiveness of wolf packs, as these are the two most influential factors contrasting the spread of introgression in social canids (Fredrickson and Hedrick, 2006; Bohlings and Waits 2015, Santostasi et al. 2020). However, in presence of widespread introgression as the one we revealed, reactive interventions should also be contemplated (e.g., Gese and Terletzky, 2015; van Wyk et al. 2016) to limit the spread domestic alleles by significantly reducing the reproductive contribution of the admixed individuals. In this perspective, population-based assessments such as ours ought to be

For Review Only

supported by a more in-depth assessment of introgression at the individual level in order to identify admixed individuals with greater confidence (i.e., a higher number of genetic markers). Population-wide, reliable assessments of admixture should be more frequently conducted to monitor the genetic status of populations threatened by hybridization before it is too late (Allendorf et al. 2001). They are useful to eventually indicate admixed propagules in due time and assess the effectiveness of preventive and proactive measures. In this perspective, new and sensitive techniques, such as microfluidic PCR, enable genotyping of reduced panels of ancestry-informative SNPs also on degradated DNA obtained from non-invasive samples, and this could be soon routinely applied for future assessment of admixture (Caniglia et al. 2020, von Thaden et al. 2017). In addition to this, we believe our modelling approach will serve for more frequent and reliable assessment and monitoring of anthropogenic hybridization where this is threatening the genomic integrity of wild parental populations.

AUTHOR CONTRIBUTIONS

OG, PC and NLS conceived the ideas, designed the methodology and led the writing of the manuscript; PC and WR organized and supervised the field data collection and logistics; LM collected the data; OG and NLS constructed the models and analyzed the data; EF and RC performed the genetic analyses. All authors contributed critically to the drafts and gave final approval for publication.

LITERATURE CITED

Adams, J. R., Kelly, B. T., Waits, L. P. 2003. Using faecal DNA sampling and GIS to monitor hybridization between red wolves (*Canis rufus*) and coyotes (*Canis latrans*). Molecular Ecology 12:2175-2186.

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

For Review Only

694 Anderson, E. C. and Thompson, E.A. 2002. A model-based method for identifying species hybrids
695 using multilocus genetic data. Genetics 160, 1217-1229.
696 Anderson, D. R. 2001. The need to get the basics right in wildlife field studies. Wildlife Society
697 Bulletin 29:1294-1297.
698 Anderson, T. M., vonHoldt, B. M., Candille, S. I., Musiani, M., Greco, C., Stahler, D. R., et al.,
699 2009. Molecular and evolutionary history of melanism in North American gray wolves.
700 Science 323:1339-1343.
701 Bassi, E., Canu, A., Firmo, I., Mattioli, L., Scandura, M. and Apollonio, M. 2017. Trophic overlap
702 between wolves and free-ranging wolf dog hybrids in the Apennine Mountains, Italy. Global
703 Ecology and Conservation 9:39-49.
704 Battocchio, D., Iacolina, L., Canu, A., and Mori, E. 2017. How much does it cost to look like a pig
705 in a wild boar group? Behavioural Processes, 138:123-126.
706 Benson, J. F., Patterson, B. R., and Wheeldon, T. J. 2012. Spatial genetic and morphologic
707 structure of wolves and coyotes in relation to environmental heterogeneity in a Canis hybrid
708 zone. Molecular Ecology 21:5934-5954.
709 Benson, J.F., Patterson, B.R. and Wheeldon, T.J. (2012), Spatial genetic and morphologic structure
710 of wolves and coyotes in relation to environmental heterogeneity in a Canis hybrid zone.
711 Mol Ecol, 21: 5934-5954. doi:10.1111/mec.12045
712 Besnier, F. and Glover, K. A. 2013. ParallelStructure: A R Package to distribute parallel runs of the
713 population genetics program STRUCTURE on multi-core computers. PLoS ONE 8:1-5.
714 Bohling, J. H. 2016. Strategies to address the conservation threats posed by hybridization and
715 genetic introgression. Biological Conservation, 203:321-327.
716 Bohling J. H., Adams, J,R,, Waits, L. P. 2013. Evaluating the ability of Bayesian clustering
717 methods to detect hybridization and introgression using an empirical red wolf data set.
718 Molecular Ecology 22:74-86.
719 Bohling, J. H. and Waits, L. P. 2015. Factors influencing red wolf–coyote hybridization in eastern
720 North Carolina, USA. Biological Conservation 184:108-116.
721 Boivin, N. L., Zeder, M. A., Fuller, D. Q., Crowther, A., Larson, G., Erlandson, J. M., Denham, T.
722 and Petraglia, M. D. 2016. Ecological consequences of human niche construction:
723 Examining long-term anthropogenic shaping of global species distributions. Proceedings of
724 the Naional Academy of Sciences of the United States of America 113:6388-6396.
725 Brainerd, S. M., Andrén, H., Bangs, E. E., Bradley, E. H., Fontaine, J. A., Hall, et al. 2008. The
726 effects of breeder loss on wolves. Journal of Wildlife Management 72:89-98.

For Review Only

727 Burnham, K. P., and D. R. Anderson. 1998. Model selection and inference: a practical information-
728 theoretic approach. Springer-Verlag, New York, New York, USA.
729 Butler, D. 1994. Bid to protect the wolf from genetic pollution. Nature 370:497.
730 Caniglia, R., Fabbri, E., Cubaynes, S., Gimenez, O., Lebreton, J.-D. and Randi, E. 2012. An
731 improved procedure to estimate wolf abundance using non-invasive genetic sampling and
732 capture–recapture mixture models. Conservation Genetics 13:53-64.
733 Caniglia, R., Fabbri, E., Galaverni, M., Milanesi, P. and Randi, E. 2014. Noninvasive sampling and
734 genetic variability, pack structure, and dynamics in an expanding wolf population. Journal of
735 Mammalogy 9:41-59.
736 Caniglia, R., Fabbri, E., Greco, C., Galaverni, M., Manghi, L., Boitani, L., Sforzi, A., and Randi, E.
737 2013. Black coats in an admixed wolf dog pack is melanism an indicator of hybridization in
738 wolves? European Journal od Wildlife Research 59:543-555.
739 Caniglia, R., Galaverni, M., Velli, E., Mattucci, F., Canu, A., Apollonio, M., et al. 2020. A
740 standardized approach to empirically define reliable assignment thresholds and appropriate
741 management categories in deeply introgressed populations. Scientific Reports 10:2862.
742 Chapron, G., Kaczensky, P., Linnell, J. D., von Arx, M., Huber, D., Andrén, H., et al. 2014.
743 Recovery of large carnivores in Europe's modern human-dominated landscapes. *Science*
744 346, 1517- 1519.
745 Ciucci P., and Boitani L. 1999. Nine-year dynamics of a wolf pack in the Northern Apennines
746 (Italy). Mammalia 63:377-384
747 Ciucci, P., Boitani, L., Falco, M. and Maiorano, L. 2018. Hierarchical, multi-grain rendezvous site
748 selection by wolves in southern Italy. Journal of Wildlife Management, 82:1049-1061.
749 Ciucci, P., Boitani, L., Raganella Pelliccioni, E., Rocco, M. and Guy, I. 1996. A comparison of
750 scat-analysis methods to assess the diet of the wolf *Canis lupus*. Wildlife Biology 2:37-48.
751 Ciucci, P., Masi, M. and Boitani, L. 2003. Winter habitat and travel route selection by wolves in the
752 northern Apennines, Italy. Ecography 26:223-235.
753 Ciucci, P., Reggioni, W., Maiorano, L. and Boitani, L. 2009. Long-distance dispersal of a rescued
754 wolf from the northern Apennines to the western Alps. Journal of Wildlife. Management.
755 73:1300-1306.
756 Corander, J., Marttinen, P., Sirén, J., Tang, J. 2008. Enhanced Bayesian modelling in BAPS
757 software for learning genetic structures of populations. BMC Bioinformatics 9, article
758 number 539.

For Review Only

759 Coulson, T., MacNulty, D. R., Stahler D. R., vonHoldt,B., Wayne, R. K., Smith, D. S. Modeling
760 effects of environmental change on wolf population dynamics, trait evolution, and life
761 history. Science 334:1275–1278.
762 Choquet, R., Rouan, L. and Pradel, R. 2009. Program E-Surge: a software application for fitting
763 multievent models. Pages 845-865 *in* Thompson, D. L., Cooch, E. G. and Conroy M. J.
764 editors. Modeling demographic processes in marked populations. Environmental and
765 ecological statistics series. Volume 3. Springer-Verlag, New York, New York, USA.
766 Crespin, L., Choquet, R., Lima, M., Merritt, J. and Pradel, R. 2008. Is heterogeneity of catchability
767 in capture–recapture studies a mere sampling artifact or a biologically relevant feature of the
768 population? Population Ecology 50:247-256.
769 Cubaynes, S., Pradel, R., Choquet, R., Duchamp, C., Gaillard, J. M., Lebreton, J. D., Marboutin, E.,
770 Miquel, C., Reboulet, A.M., Poillot, C., Taberlet, P. and Gimenez, O. (2010). Importance of
771 accounting for detection heterogeneity when estimating abundance: The case of French
772 wolves. Conservation Biology 24: 621-626.
773 Davison, A. C. and Hinkley, D. V. 2002. Bootstrap Methods and their Application. Cambridge
774 University Press, Cambridge, UK.
775 Derégnaucourt, S., Guyomarc'h, J. C., and Belhamra, M. 2004. Comparison of migratory tendency
776 in European Quail Coturnix coturnix, domestic Japanese Quail *Coturnix japonica* and their
777 hybrids. Ibis, 147:25-36.
778 Donfrancesco, V. Ciucci, P., Salvatori, V., Benson, D., Andersen L.W., Bassi E., et al. 2019.
779 Unravelling the scientific debate on how to address wolf-dog hybridization in europe .
780 Frontiers in Ecology and Evolution 7:1-13.
781 Dufresnes, C., Remollino, N., Stoffel, C., Manz, R., Weber, J. M., Fumagalli, L. 2019. Two decades
782 of non-invasive genetic monitoring of the grey wolves recolonizing the Alps support very
783 limited dog introgression. Scientific Reports 9:1-9.
784 Fabbri, E., Miquel, C., Lucchini, V., Santini, A., Caniglia, R., Duchamp, C., et al. Weber, J. M.,
785 Lequette, B., Marucco, F., Boitani,L., Fumagalli, L., Taberlet, P. and Randi, E. 2007. From
786 the Apennines to the Alps: colonization genetics of the naturally expanding Italian wolf
787 (*Canis lupus*) population. Molecular Ecology16:1661-1671.
788 Fabbri, E., Velli, E., D'Amico, F., Galaverni, M., Mastrogiuseppe, L., Mattucci, F. and Caniglia, R.
789 2018. From predation to management: monitoring wolf distribution and understanding
790 depredation patterns from attacks on livestock. HISTRIX 29:101-110.

For Review Only

791 Falush, D., Stephens, M. and Pritchard, J. K. 2003. Inference of population structure using
792 multilocus genotype data: Linked loci and correlated allele frequencies. *Genetics* 164, 1567-
793 1587.
794 Frantz, L. A. F., Mullin, V. E., Pionnier-Capitan, M., Lebrasseur, O., Ollivier, M., Perri. A., et al.
795 2016. Genomic and archaeological evidence suggest a dual origin of domestic dogs. Science
796 352:1228–1231.
797 Freedman, A. H., and Wayne R. K. 2017. Deciphering the origin of dogs: from fossils to genomes
798 Annual Review of Animal Biosciences 5:281-307.
799 Fredrickson, R. J. and Hedrick, P. W. 2006. Dynamics of hybridization and introgression in red
800 wolves and coyotes. Conservation Biology 20:1272-1283**.**
801 Galaverni, M., Caniglia, R., Pagani, L., Fabbri, E., Boattini, A., Randi, E. 2017. Disentangling
802 timing of admixture, patterns of introgression, and phenotypic indicators in a hybridizing
803 wolf population. Molecular Biology and Evoution 34:2324-2339.
804 Galaverni, M., Caniglia, R., Fabbri, E., Milanesi, P., Randi, E. 2016. One, no one, or one hundred
805 thousand: how many wolves are there currently in Italy? Mammal Research 61:13-24.
806 Genovesi, P. and Dupré, E. 2000. Strategia nazionale di conservazione del lupo (*Canis lupus*):
807 Indagine sulla presenza e la gestione di cani vaganti in Italia. Biologia e Conservazione della
808 Fauna 104:1-36.
809 Gese, E. M. and Terletzky, P. A. 2015. Using the 'placeholder' concept to reduce genetic
810 introgression of an endangered carnivore. Biological Conservation 192:11-19.
811 Gimenez, O., Lebreton, J.-D., Gaillard, J.-M., Choquet, R., and Pradel, R. 2012. Estimating
812 demographic parameters using hidden process dynamic models. Theoretical Population
813 Biology 82:307-316.
814 Gimenez, O., Lebreton, J. D., Choquet, R., Pradel, R. 2018. R2ucare: An r package to perform
815 goodness-of-fit tests for capture–recapture models. Methods in Ecology and Evolution 9:
816 1749-1754.
817 Godinho, R., Llaneza, L., Blanco, J. C., Lopes, S., Álvares, F., García, E. J., and Ferrand, N. 2011.
818 Genetic evidence for multiple events of hybridization between wolves and domestic dogs in
819 the Iberian Peninsula. Molecular Ecology 20:5154–5166.
820 Grossen, C., Keller, L., Biebach, I., The International Goat Genome Consortium and Croll D.
821 Introgression from domestic goat generated variation at the major histocompatibility
822 complex of Alpine ibex. PLoS Genetics 10: e1004438.

For Review Only

Hindrikson, M., Remm, J., Pilot, M., Godinho, R., Stronen, A. V., Baltrūnaité, L., 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., Falush, D., Stephens, M. and Pritchard, J. K. 2009. Inferring weak population structure with the assistance of sample group information. Molecular Ecology Resources 9:1322-1332.

Jennelle, C. S., Cooch, E., Conroy, M. J., and Senar, J. C. 2007. State specific detection probability and disease prevalence. Ecological Applications 17:154-167.

Jimenez, M. D., Bangs, E. E., Boyd, D. K., Smith, D. W., Becker, S. A., Ausband, D. E., Woodruff, S. P., Bradley, E. H., Holyan, J. and Laudon, K. 2017. Journal of Wildlife Management 81:581-592.

Kusak, J., Caputo, V., Galov, A., Gomerčić, T., Arbanasić, H., Caniglia, R., et al. 2018. Wolf-dog hybridization in Croatia. Veterinarski Arhiv 88:375-395.

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

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

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

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

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

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

MacKenzie, D. I. 2005. What Are the Issues with Presence-Absence Data for Wildlife Managers? The Journal of Wildlife Management 69:849–860.

For Review Only

McDonald, T. L., and Amstrup, S. C. 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., Adams, J. R., Waits, L. P. 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., Joyce, P., Waits, L. P. 2002. Assessing allelic dropout and genotype reliability using maximum likelihood. Genetics 160:357-66.

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

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

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

Pendleton, A. L., Shen, F., Taravella, A. M., Emery, S., Veeramah, K. R., Boyko, A. R., et al. 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., Greco, C., vonHoldt, B. M., Randi, E., Jędrzejewski, W., Sidorovich, V. E., Konopiński, M. K., Ostrander, E. A. and Wayne, R.K. 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., Alapetite,A., Cornuet, J.M., Paetkau, D., Baudouin, L., Estoup, A. 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.

For Review Only

Pritchard, J. K., Stephens, M., and Donnelly, P. 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., Hulva, P., Fabbri, E., Galaverni, M., Galov, A., Kusak, J., Bigi, D., Bolfikova, B. C., Smetanova, M. and Caniglia, R. 2014. Multilocus detection of wolf x dog hybridization in italy, and guidelines for marker selection. PLoS ONE 9:1-13.

Randi, E., Lucchini, V., Christensen, M. F., Mucci, N., Funk, S.M., Dolf, G., Loeschke, V. 2000. Mitochondrial DNA variability in Italian and east European wolves: Detecting the consequences of small population size and hybridization. Conservation Biology 14:464-473.

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

Ritchie. E. G., Letnic, C. R. and Vanak, A. T. Dogs as predators and trophic regulators. Pages 55-68 *in* M. E. Gompper, editor. Free-Ranging Dogs and Wildlife Conservation. Oxford UnivPress, Oxford, USA.

Rouan, L., Gaillard, J. M., Guédon, Y., and Pradel, R. 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.,White, B. N., Row, J. R., Patterson, B. R. 2012. Intense harvesting of Eastern wolves facilitated hybridization with coyotes. Ecology and Evolution 2:19-33.

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

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

Santostasi, N. L., Ciucci, P., Caniglia, R., Fabbri, E., Molinari L., Moretti F., Reggioni W. and Gimenez O. 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.

Santostasi, N. L., Ciucci, P., Bearzi, G., Bonizzoni, S. and Gimenez, O. (2020). Assessing the

For Review Only

924 dynamics of hybridization through a matrix modelling approach. Ecological Modelling
925 431:109120
926 Sanz, N., Araguas, R. M., Fernández, R., Vera, M., Garcìa Marìn, J. L. 2009. Efficiency of markers
927 and methods for detecting hybrids and introgression in stocked populations. Conservation
928 Genetics 10:225–236.
929 Saltz, D. 2011. Statistical inference and decision making in conservation biology. Israel Journal of
930 Ecology and Evolution 57:4, 309-317.
931 Skoglund, P., Gotherstrom, A., and Jakobsson, M. 2011. Estimation of population divergence times
932 from non-overlapping genomic sequences: Examples from dogs and wolves. Molecular
933 Biology and Evoution. 28:1505-1517.
934 Stephens, D., Wilton, A. N., Fleming, P. J.,Berry, O. 2015. Death by sex in an Australian icon: a
935 continent-wide survey reveals extensive hybridization between dingoes and domestic dogs.
936 Molecular Ecology 24:5643-5656.
937 Taberlet, P., Griffin, S., Goossens, B., Questiau, S., Manceau, V., Escaravage, N., Waits, L. P., and
938 Bouvet, J. 1996. Reliable genotyping of samples with very low DNA quantities using PCR.
939 Nucleic acids research, 24:3189-3194.
940 Todesco, M., Pascual, M. A., Owens, G. L., Ostevik, K.L., Moyers, B.T., Hübner, S., et al. 2016.
941 Hybridization and extinction. Evolutionary Applications 9:892-908.
942 Vähä, J. P., and Primmer, C. R. 2006. Efficiency of model-based Bayesian methods for detecting
943 hybrid individuals under different hybridization scenarios and with different numbers of
944 loci. Molecular Ecology 15:63-72.
945 Valière, N. 2002. GIMLET: a computer program for analysing genetic individual identification
946 data. Molecular Ecology Notes, 2:377-379.
947 van Wyk, A. M., Dalton, D. L., Hoban, S., Bruford, M. W., Russo, I. M., Birss, C., et al. 2016.
948 Quantitative evaluation of hybridization and the impact on biodiversity conservation.
949 Ecology and evolution, 7:320-330.
950 Vaysse, A., Ratnakumar, A., Derrien, T., Axelsson, E., Rosengren Pielberg, G., Sigurdsson, S., et
951 al. 2011. Identification of genomic regions associated with phenotypic variation between
952 dog breeds using selection mapping. PLoS Genetics 7:e1002316.
953 Verardi, A., Lucchini, V. and Randi, E. 2006. Detecting introgressive hybridization between free-
954 ranging domestic dogs and wild wolves (*Canis lupus*) by admixture linkage disequilibrium
955 analysis. Molecular Ecology 15:2845-55.
956 Vilà, C., and Wayne, R. K. 1999. Hybridization between wolves and dogs. Conservation Biology
957 13:195-198.

For Review Only

vonHoldt BM, Pollinger JP, Earl DA, Parker HG, Ostrander EA, Wayne RK. 2013. Identification of recent hybridization between gray wolves and domesticated dogs by SNP genotyping. Mammalian Genome 24:80-8.

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

von Thaden, A., Cocchiararo, B., Jarausch, A. Jüngling, H., Karamanlidis, A. A., Tiesmeyer, A., Nowak C., and Muñoz-Fuentes V. 2017. Assessing SNP genotyping of noninvasively collected wildlife samples using microfluidic arrays. Scientific Reports 7:10768.

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

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

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

Wilkins, A. S., Wrangham R. W., 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., Nichols, J. D., Boulinier, T. 2001. Monitoring of biological diversity in space and time. Trends in Ecology & Evolution: 16:446-453.

Zimen E, and Boitani L. 1975. Number and distribution of wolf in Italy. Zeitschrift Fur Saugetierkunde 40:102-112.

Zucchini, W., MacDonald, I. L., and Langrock, R. 2009. Hidden Markov models for time series: an introduction using R. Chapman & Hall/CRC Press, Boca Raton, Florida, USA.

ACKNOWLEDGMENTS

We are grateful to students, technicians and volunteers who assisted with field work. M. Canestrini and F. Moretti assisted with sample collection and administrative work. N. Mucci provided useful suggestions, M. Galaverni and E. Velli assisted the genetic analyses and supported the bayesian procedures. Data collection and genetic analyses were co-funded by the European Union (LIFE13/NAT/IT/000728) and the PNATE. NLS was supported by a PhD grant by The University

For Review Only

990 or Rome La Sapienza. OG and NLS were supported by the French National Research Agency grant
991 ANR-16-CE02-0007.

For Review Only

993 Figure captions

994

995 Figure 1. Location of the study area along the Northern Apennines in Italy (inlet: Italian wolf
996 population range), and sampling locations (dots) of the collected scats within each approximate
997 wolf pack territory (circles). Black dots refer to individuals genetically classified as wolves, while
998 white dots to individuals classified as wolf x dog admixed individuals (under classification scenario
999 B, aimed at balancing Type 1 and 2 classification errors, see text).

1000

1001 Figure 2. Left panel: boxplot of the q_w values of the genotypes simulated from reference wolves
1002 (n=190) and dogs (n=89) grouped per genealogical class (n=100 genotypes for each class). The grey
1003 horizontal lines define the area of uncertainty that is zoomed in the right panel. Right panel: detail
1004 of the area of uncertainty comprised between the minimum q_w of reference wolves (dashed line) and
1005 the maximum q_w of simulated first generation backcrosses (solid line). We classified as Wolves
1006 those sampled genotypes whose q_w was above the solid line, as Admixed those whose q_w was below
1007 the dashed line, and as Uncertains those in between. W=wolves; F1 and F2=first and second
1008 generation hybrids, respectively; BCW1-4=first to fourth generation backcrosses of F1 with wolves;
1009 SBCW1-4=first to fourth generation of backcrosses of F2 with wolves.

1010

1011 Figure 3. Prevalence of admixed individuals in a protected wolf population composed of seven wolf
1012 packs in the Northern Apennines, Italy (2015-2016), as assessed by non-invasive genetic sampling
1013 and Bayesian assigment procedures based on multilocus genoypes (12 STRs). Estimates of
1014 prevalence, obtained through an open population CR modelling approach, refer to an overall 10-
1015 month sampling period and are shown for each 2-month sampling occasion. The two scenarios refer
1016 to two alternative rationales to cope with uncertainty in the classification of admixed individuals
1017 (see text)

1018

For Review Only

Figure 4. Reconstructed genealogies for two of the seven wolf packs non-invasively sampled in the Northern Apennines, Italy (2015-2016), to estimate admixture between wolves and dogs. For each the two packs, the two 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 ID, the second line the individual q_w, and the third line their classification based on their q_w and a classification rationale aimed at minimizing type 1 error (scenario A) and the fourth line a classification rationale aimed at balancing between Type 1 and 2 error rates (scenario B). All individuals originally classified as Uncertains (U) based on the q_w were successively assigned to the admixed category (H) by our probabilistic, a posteriori procedure based on the Viterbi algorithm (see text).

For Review Only

1031 Table 1. Most supported capture recapture models and model averaged estimates of prevalence of
1032 admixed individuals in the PNATE for scenario A and B (p = detectability, π = initial state
1033 probability, φ = apparent survival probability, δ = assignment probability). The term (.) indicates
1034 constant parameters, the term (state) indicates state-dependent (i.e., Wolf vs. Admixed) parameters.
1035 We considered for the model averaging only the model that had a ΔQAICc <2 from the best model.

Model selection scenario A

Model	*# Id. Par.*	*Deviance*	*QAIC*	*QAICc*	*ΔQAICc*
π(.)p(.)φ(state)	4	152.37	160.37	161.03	0.00
π(.)p(state)φ(state)	5	152.34	162.34	163.35	2.32
π(.)p(state)φ(.)	4	155.10	163.10	163.76	2.73
π(.)p(time)φ(state)	7	148.41	162.41	164.38	3.35
π(.)p(.)φ(.)	3	158.03	164.03	164.43	3.40
π(.)p(time)φ(.)	6	154.15	166.15	167.60	6.57
π(.)p(.)φ(pack)	9	147.13	165.13	168.41	7.38
π(.)p(state)φ(pack)	10	146.57	166.57	170.65	9.62
π(.)p(pack)φ(state)	10	149.59	169.59	173.67	12.64
π(.)p(time)φ(pack)	12	144.11	168.11	174.11	13.08
π(.)p(pack)φ(.)	9	152.84	170.84	174.11	13.08
π(.)p(pack)φ(pack)	15	143.88	173.88	183.68	22.65

Model selection scenario B

Model	*# Id. Par.*	*Deviance*	*QAIC*	*QAICc*	*ΔQAICc*
π(.)p(.)φ(.)δ(.)	4	176.49	184.49	185.17	0
π(.)p(.)φ(state)δ(.)	5	175.46	185.46	186.49	1.32
π(.)p(state)φ(.)δ(.)	5	175.62	185.62	186.65	1.48
π(.)p(state)φ(state)δ(.)	6	175.39	187.39	188.86	3.67

For Review Only

π(.)p(.)φ(pack)δ(.)	10	165.46	185.46	189.61	4.44
π(.)p(t)φ(pack)δ(.)	7	173.64	187.64	189.64	4.47
π(.)p(t)φ(state)δ(.)	8	172.62	188.616	191.23	6.06
π(.)p(state)φ(pack)δ(.)	11	165.46	187.46	192.54	7.36
π(.)p(pack)φ(.)δ(.)	10	171.45	191.45	195.60	10.42
π(.)p(t)φ(pack)δ(.)	13	163.32	189.32	196.60	11.432
π(.)p(pack)φ(state)δ(.)	11	171.41	193.41	198.48	13.312
π(.)p(pack)φ(pack)δ(.)	16	162.74	194.74	206.31	21.14

Table 2. Model averaged estimates of prevalence of admixture in a protected wolf population in the northern Apennines, Italy (2015-2016) estimated with multistate and multievent capture-recapture models (p = detectability, π = initial state probability, φ = apparent survival probability). The term (.) indicates constant parameters, the term (state) indicates state dependent (i.e., Wolf vs. Admixed) parameters.

Estimates Scenario A					
Sampling occasion	*Aug–Sept*	*Oct–Nov*	*Dec–Jan*	*Feb–Mar*	*Apr–May*
Poulation abundance ($\hat{N}$)	25	25	31	31	23
	(11-44)	(11-39)	(17-50)	(16-50)	(11-44)
Wolf abundance ($\hat{N}_w$)	13	13	17	13	10
	(3-26)	(4-25)	(7-30)	(4-24)	(3-26)
Admixed individuals	13	13	15	19	13
abundance ($\hat{N}_h$)	(3-26)	(3-21)	(5-28)	(8-31)	(3-26)
Prevalence (95% CIs)	0.50	0.50	0.47	0.60	0.55
	(0.22-0-80)	(0.20-0.79)	(0.21-0.79)	(0.37-0.84))	(0.22-0.80)
Estimates Scenario B					

For Review Only

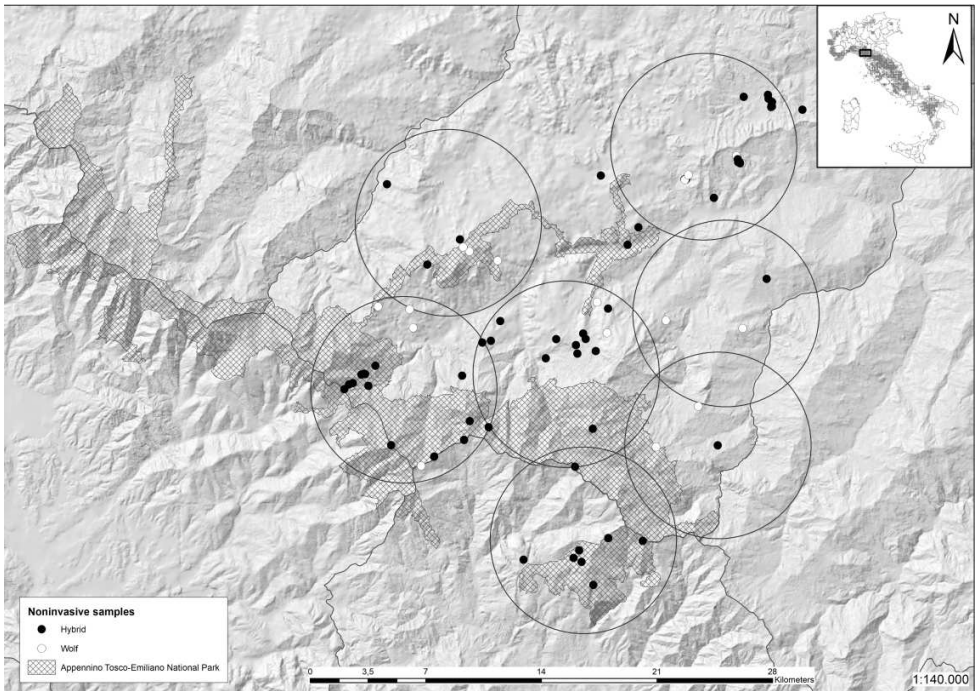
Poulation abundance ($\hat{N}$)	26	28	35	35	26
	(13-78)	(14-85)	(17-109)	(17-109)	(13-80)
Wolf abundance ($\hat{N}_w$)	7	10	7	7	6
	(3-26)	(4-35)	(3-26)	(3-26)	(2-20)
Admixed individuals	17	17	26	26	19
abundance ($\hat{N}_h$)	(8-48)	(8-48)	(12-72)	(12-72)	(8-51)
Prevalence (95% CIs)	0.70	0.64	0.78	0.78	0.77
	(0.45-0.77)	(0.38-0.72)	(0.55-0.82)	(0.55-0.84)	(0.52-0.85)

For

Summary for online Table of Contents. We applied a multi-event, capture-recapture model formulation to reliably estimating the proportion of admixed individuals in a wolf population affected by dog introgression, which is an emblematic case of anthropogenic hybridization. By accounting for generally neglected sources of bias and the uncertainty in classifying admixed individuals based on their multilocus genotypes, we detected a high proportion of introgressed individuals indicating that, under certain ecological conditions, reproductive barriers and dilution of dog genes into parental wolf populations can not be expected, per se, to prevent widespread introgression.

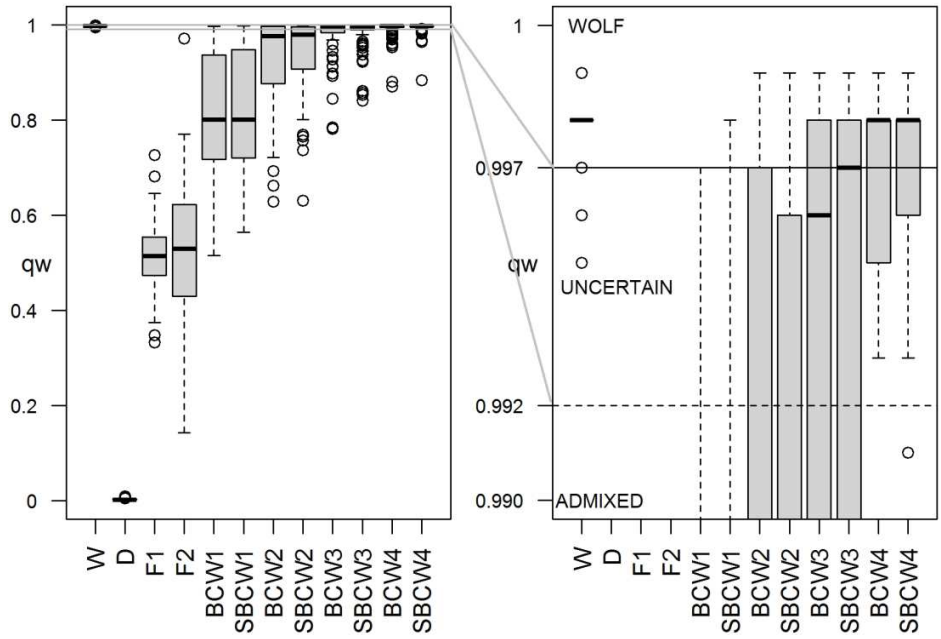
1052

For Review Only



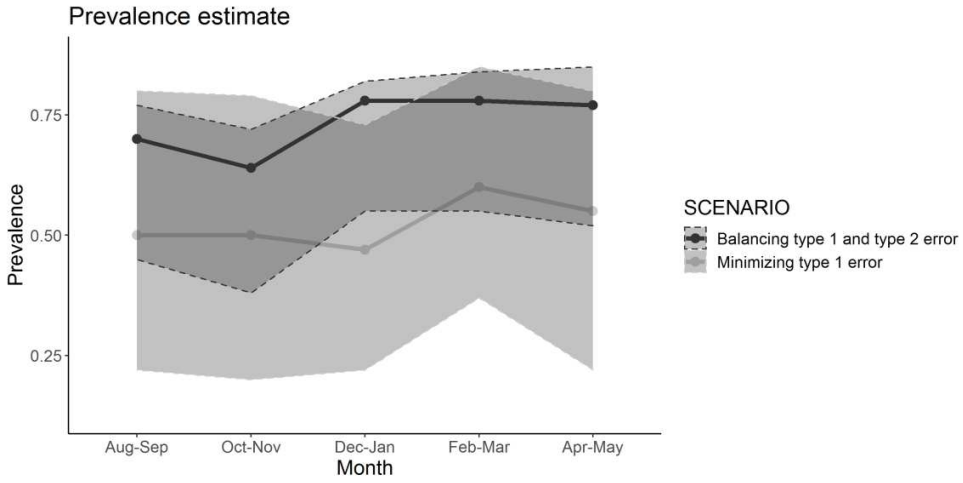
Location of the study area along the Northern Apennines in Italy (inlet: Italian wolf population range), and sampling locations (dots) of the collected scats within each approximate wolf pack territory (circles). Black dots refer to individuals genetically classified as wolves, while white dots to individuals classified as wolf x dog admixed individuals (under classification scenario B, aimed at balancing Type 1 and 2 classification errors, see text).

420x297mm (300 x 300 DPI)



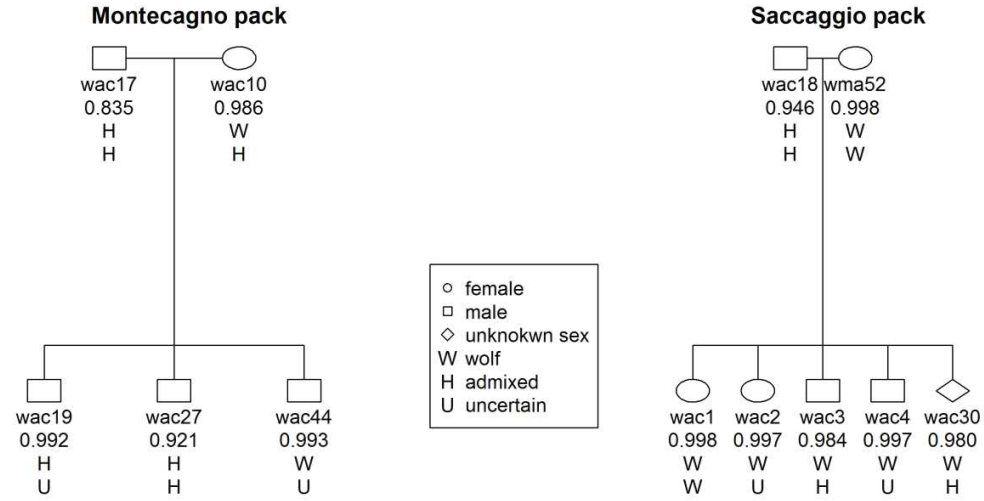
Left panel: boxplot of the qw values of the genotypes simulated from reference wolves (n=190) and dogs (n=89) grouped per genealogical class (n=100 genotypes for each class). The grey horizontal lines define the area of uncertainty that is zoomed in the right panel. Right panel: detail of the area of uncertainty comprised between the minimum qw of reference wolves (dashed line) and the maximum qw of simulated first generation backcrosses (solid line). We classified as Wolves those sampled genotypes whose qw was above the solid line, as Admixed those whose qw was below the dashed line, and as Uncertains those in between. W=wolves; 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 of backcrosses of F2 with wolves.

169x119mm (300 x 300 DPI)



Prevalence of admixed individuals in a protected wolf population composed of seven wolf packs in the Northern Apennines, Italy (2015-2016), as assessed by non-invasive genetic sampling and Bayesian assigment procedures based on multilocus genoypes (12 STRs). Estimates of prevalence, obtained through an open population CR modelling approach, refer to an overall 10-month sampling period and are shown for each 2-month sampling occasion. The two scenarios refer to two alternative rationales to cope with uncertainty in the classification of admixed individuals (see text).

254x127mm (300 x 300 DPI)



Reconstructed genealogies for two of the seven wolf packs non-invasively sampled in the Northern Apennines, Italy (2015-2016), to estimate admixture between wolves and dogs. For each the two packs, the two 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 ID, the second line the individual qw, and the third line their classification based on their qw and a classification rationale aimed at minimizing type 1 error (scenario A) and the fourth line a classification rationale aimed at balancing between Type 1 and 2 error rates (scenario B). All individuals originally classified as Uncertains (U) based on the qw were successively assigned to the admixed category (H) by our probabilistic, a posteriori procedure based on the Viterbi algorithm (see text).

209x159mm (300 x 300 DPI)

SUPPLEMENTARY MATERIALS

Estimating admixture at the population scale: taking imperfect detectability and uncerainty in hybrid classification seriously

NINA L. SANTOSTASI, Department of Biology and Biotechnologies "Charles Darwin", University of Rome La Sapienza, Rome, Italy; CEFE, CNRS, University of Montpellier, University Paul Valéry Montpellier 3, EPHE, IRD, Montpellier, France

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, Department of Biology and Biotechnologies "Charles Darwin", University of Rome La Sapienza, Rome, Italy

The following document contains:

- Details regarding ancestry and genealogy analyses;
- Table S1: the q_w range values of simulated and reference genotypes from Caniglia et al. (2020);
- Table S2: the q_w values, Bayesian Confidence Intervals, sex, y-haplotype pack and classification of the sampled individuals;
- Table S3: sampling results
- Table S4: the misclassification rates of the simulated genotypes according to Scenario A;
- Table S5: the misclassification rates of the simulated genotypes according to Scenario B;
- Table S6: Viterbi algorithm capture histories reconstruction for the 6 uncertain individuals (first row) according the three best models.
- Table S7: goodness of fit test results for the Cormack Jolly Seber model;
- Table S8: the probability of the admixed individuals of being first generation backcrosses estimated through ancestry analysis.

Ancestry and geneology analyses

We used the option "Population Information to test for migrants" implemented in software 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, as well as the collected genotypes, were a priori correctly identified and assigned to their own cluster. Due to a limited number of loci, we restricted ancestry analysis to two generations backwards (GENBACK = 2).

To estimate the genealogies of the packs and the reproductive role of admixed individuals, we used a maximum-likelihood approach implemented in COLONY 2.0 (Wang and Santure, 2009). For each area, we ran COLONY considering as candidate parents all the detected wolf or admixed

48 genotypes, allele frequencies estimated from the whole reference wolf population, amplification
49 error rates estimated by GIMLET 1.3.3 (Valiére, 2003), and probability of including fathers and
50 mothers in the candidate parental pair of 0.5 (Caniglia et al., 2014). We then verified the
51 genealogies using the approach based on the Mendelian inheritance of the alleles implemented in
52 GIMLET, accepting only parent-offspring combinations of individuals contemporarily sampled but
53 with at most one allele of incompatibilities on the 12 loci applied, and father-son combinations with
54 no incongruities at Y-STR haplotypes (Fabbri et al., 2018).
55

Table S1. Range of q_w values of simulated genotypes (n=100 for each category) and reference populations (n=190 for wolves and 89 for dogs) by Bayesian clustering analysis with 12 microsatellites loci (Caniglia et al., 2020). F1 and F2: first and second generation hybrids, respectively; BCW1–4: first to fourth generation backcrosses between F1 with wolves; SBCW1–SBCW4: first to fourth generation backcrosses between F2 with wolves.

Sample	**q_w range**
Reference wolves	0.992-0.999
Simulated wolves	0.995-0.999
Simulated F1	0.333-0.726
Simulated F2	0.143-0.770
Simulated BCW1	0.515-0.997
Simulated BCW2	0.628-0.999
Simulated BCW3	0.781-0.999
Simulated BCW4	0.870-0.999
Simulated SBCW1	0.564-0.999
Simulated SBCW2	0.631-0.999
Simulated SBCW3	0.841-0.999
Simulated SBCW4	0.883-0.999
Simulated dogs	0.001-0.010
Reference dogs	0.001-0.010

Table S2. Values of q_w of the 39 individual genotypes sampled in this study obtained by Bayesian clustering analysis and their corresponding categories according to Scenario A and B.

Individual code	Sex	Pack	q_w	95% BCI	Y Haplotype	Scenario A	Scenario B
HRE1	M	Cerreto	0.624	0.387-0.846	Wolf	Admixed	Admixed
HAC16	F	Orecchiella	0.655	0.427-0.846	Wolf	Admixed	Admixed
HAC17	M	Orecchiella	0.703	0.424-0.960	Wolf	Admixed	Admixed
HAC2	M	Saccaggio	0.745	0.467-1	Wolf	Admixed	Admixed
WAC33	F	Orecchiella	0.769	0.546-1	Wolf	Admixed	Admixed
HAC1	F	Saccaggio	0.800	0.568-1	Wolf	Admixed	Admixed
WAC46	M	Orecchiella	0.826	0.598-1	Wolf	Admixed	Admixed
HAC3	M	Saccaggio	0.831	0.538-1	Wolf	Admixed	Admixed
WAC17	F	Montecagno	0.835	0.589-1	Wolf	Admixed	Admixed
WAC25	F	Cerreto	0.88	0.648-1	Wolf	Admixed	Admixed
HAC12	M	Saccaggio	0.886	0.524-1	Wolf	Admixed	Admixed
HAC13	M	Orecchiella	0.900	0.67-1	Dog	Admixed	Admixed
HAC8	M	Cerreto	0.903	0.684-1	Wolf	Admixed	Admixed
WAC27	M	Montecagno	0.921	0.676-1	Wolf	Admixed	Admixed
WAC22	M	Cerreto	0.93	0.693-1	Wolf	Admixed	Admixed
WAC18	M	Saccaggio	0.946	0.7-1	Wolf	Admixed	Admixed
WAC29	NA	Saccaggio	0.961	0.811-1	Wolf	Admixed	Admixed
WAC30	NA	Saccaggio	0.980	0.852-1	Wolf	Wolf	Admixed
WAC3	M	Saccaggio	0.984	0.876-1	Wolf	Wolf	Admixed
HAC14	M	Gazzano	0.986	0.900-1	Dog	Wolf	Admixed
WAC10	M	Montecagno	0.986	0.896-1	Wolf	Wolf	Admixed

HAC10	M	Cerreto	0.988	0.912-1	Dog	Wolf	Admixed
WAC42	F	Campastrino	0.988	0.911-1	Wolf	Wolf	Admixed
WAC19	F	Montecagno	0.992	0.945-1	Wolf	Wolf	Uncertain
WAC23	M	Campastrino	0.992	0.948-1	Wolf	Wolf	Uncertain
WAC44	M	Montecagno	0.993	0.968-1	Wolf	Wolf	Uncertain
W1967	F	Campastrino	0.994	0.965-1	Wolf	Wolf	Uncertain
WAC2	F	Saccaggio	0.997	0.985-1	Wolf	Wolf	Uncertain
WAC4	M	Saccaggio	0.997	0.988-1	Wolf	Wolf	Uncertain
WAC12	F	Campastrino	0.998	0.993-1	Wolf	Wolf	Wolf
WAC13	M	Cerreto	0.998	0.993-1	Wolf	Wolf	Wolf
WAC16	F	Gazzano	0.998	0.994-1	Wolf	Wolf	Wolf
WAC1	F	Saccaggio	0.998	0.991-1	Wolf	Wolf	Wolf
WAC20	NA	Villa Minozzo	0.998	0.993-1	Wolf	Wolf	Wolf
WAC21	M	Campastrino	0.998	0.993-1	Wolf	Wolf	Wolf
WAC41	F	Cerreto	0.998	0.993-1	Wolf	Wolf	Wolf
WMA52	F	Saccaggio	0.998	0.994-1	Wolf	Wolf	Wolf
HAC9	M	Villa Minozzo	0.999	0.995-1	Dog	Wolf	Wolf
WAC43	F	Montecagno	0.999	0.995-1	Wolf	Wolf	Wolf

64

65

Table S3. Number of scats collected, successfully amplified and genotyped, and individual captures in the PNATE, Northern Apennines, Italy, from Aug 2015 to May 2016. Results are presented by 2-month capture occasions used in the capture-recapture analysis.

	Wolf-like scats		Individual genotypes	
Sampling occasions	**Collected**	**Genotyped**	**Sampled**	**New**
1 (Aug–Sept)	25	18	11	11
2 (Ott–Nov)	29	19	12	6
3 (Dec–Jan)	33	22	15	10
4 (Feb–Mar)	34	21	15	9
5 (Apr–May)	31	18	11	3
Total	152	98	64	**39**

Table S4. Classification of the simulated genotypes as Wolves or, Admixed according to Scenario A. W = wolves; 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 of backcrosses of F2 with wolves.

Simulated genotypes	Classified as Wolves	Classified as Admixed	TOT simulated
W	100	0	100
F1	0	100	100
F2	0	100	100
BCW1	14	86	100
SBCW1	14	94	100
BCW2	52	48	100
SBCW2	55	45	100
BCW3	87	13	100
SBCW3	86	14	100
BCW4	92	8	100
SBCW4	97	3	100
TOT	**597**	**503**	**1100**

77 **Table S5.** Classification of the simulated genotypes as Wolves, Admixed or Uncertains according
78 to Scenario B. W = wolves; F1 and F2 = first and second generation hybrids, respectively;
79 BCW1–4 = first to fourth generation backcrosses of F1 with wolves; SBCW1–4 = first to fourth
80 generation of backcrosses of F2 with wolves.

Simulated genotypes	**Classified as Wolves**	**Classified as Admixed**	**Classified as Uncertain**	**TOT simulated classified**	**TOT simulated**
W	76	0	24	76	100
F1	0	100	0	100	100
F2	0	100	0	100	100
BCW1	0	94	6	94	100
SBCW1	1	94	5	95	100
BCW2	18	61	21	79	100
SBCW2	12	65	23	77	100
BCW3	29	39	32	68	100
SBCW3	43	30	27	73	100
BCW4	58	23	19	81	100
SBCW4	59	12	29	71	100
TOT	**296**	**618**	**186**	**914**	**1100**

Table S6. Viterbi algorithm capture histories reconstruction for the 6 uncertain individuals (first row) according the three best models. Assign = reconstructed capture history, Prob = corresponding calculated probability for that reconstruction. In the reconstructed capture histories 1 means that the individual was attributed to the state "alive in the study area as a Wolf", 2 means that the individual was attributed to the state "alive in the study area as an Admixed" and D means that the individual was attributed to the state "Dead/Permanently emigrated." P parental is the cumulative probability across all the capture histories that the individual is a wolf and P admixed is the cumulative probability that an individual is an admixed.

	Capture histories											
	30000		**33300**		**03000**		**00300**		**00330**		**00003**	
	Assign.	*Prob.*	*Assign.*	*Prob.*	*Assign.*	*Prob.*	*Assign.*	*Prob.*	*Assign.*	*Prob.*	*Assign.*	*Prob.*
Best	2DDDD	0.40	222DD	0.34	02DDD	0.38	002DD	0.34	00222	0.43	00002	0.70
model	1DDDD	0.17	22222	0.23	01DDD	0.17	00222	0.23	0022D	0.26	00001	0.30
	22DDD	0.16	111DD	0.15	022DD	0.15	001DD	0.15	00111	0.19		
	11DDD	0.07	2222D	0.14	02222	0.10	0022D	0.14	0011D	0.11		
p admixed		**0.56**		**0.70**		**0.63**		**0.70**		**0.70**		**0.70**
p parental		**0.24**		**0.15**		**0.16**		**0.15**		**0.30**		**0.30**
	Assign.	*Prob.*	*Assign.*	*Prob.*	*Assign.*	*Prob.*	*Assign.*	*Prob.*	*Assign.*	*Prob.*	*Assign.*	*Prob.*
Second	2DDDD	0.31	22222	0.30	02DDD	0.29	00222	0.26	00222	0.50	00001	0.31
best	1DDDD	0.27	222DD	0.30	01DDD	0.25	002DD	0.26	0022D	0.22	00002	0.69

model	22DDD	0.14	111DD	0.16	02222	0.13	001DD	0.22	0011D	0.15		
	11DDD	0.09	2222D	0.13	022DD	0.19	0022D	0.11	00111	0.14		
p admixed		**0.44**		**0.74**		**0.55**		**0.63**		**0.72**		**0.69**
p parental		**0.36**		**0.13**		**0.25**		**0.11**		**0.28**		**0.31**
	Assign.	*Prob.*	*Assign.*	*Prob.*	*Assign.*	*Prob.*	*Assign.*	*Prob.*	*Assign.*	*Prob.*	*Assign.*	*Prob.*
Third	2DDDD	0.37	222DD	0.39	02DDD	0.35	002DD	0.31	00222	0.47	00002	0.70
best	1DDDD	0.16	22222	0.26	01DDD	0.15	00222	0.21	0022D	0.27	00001	0.30
model	22DDD	0.14	2222D	0.15	022DD	0.14	00111	0.16	00111	0.18		
	11DDD	0.08	11111	0.09	01111	0.1	001DD	0.13	0011D	0.08		
p admixed		**0.51**		**0.80**		**0.49**		**0.51**		**0.74**		**0.70**
p parental		**0.25**		**0.09**		**0.25**		**0.30**		**0.26**		**0.30**

Table S7. Overall goodness of fit for Cormack-Jolly-Seber model. TEST 2 evaluates the assumption of homogeneous detection probabilities and a A p-value < 0.05 indicates the presence of trap-dependence. TEST 3 evaluates the assumption of homogeneous survival probabilities and a p-value < 0.05 indicates the presence of transience.

Test	**Statistic**	***p*-value**	**Degrees of freedom**
TEST 2 Ct	0.5	0.7	2
TEST 3 Sr	3.3	0.3	3
Overall test	6.5	0.4	7

Table S8. Inferred ancestry of the 29 admixed and uncertain individuals observed in this study using the population information model implemented in STRUCTURE. We show estimates of posterior probabilities (q) for each individual to have ancestry in its a priori assigned population (q prior pop), or in the other population in the present generation (dog or wolf), in the first (parent) past generation or the second (grandparent) past generation.

Individual	Prior. pop	*q* prior pop	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

LITERATURE CITED

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

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

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

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

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

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

Ecological Modelling 431 (2020) 109120

Contents lists available at ScienceDirect

Ecological Modelling

journal homepage: www.elsevier.com/locate/ecolmodel

Assessing the dynamics of hybridization through a matrix modelling approach

Nina Luisa Santostasi[a,b,c,*], Paolo Ciucci[a], Giovanni Bearzi[c,d], Silvia Bonizzoni[c,d], Olivier Gimenez[b]

[a] Dept. Biology and Biotechnologies, University of Rome La Sapienza, Rome, Italy
[b] CEFE, Univ Montpellier, CNRS, Univ Paul Valéry Montpellier 3, EPHE, IRD, Montpellier, France
[c] Dolphin Biology and Conservation, Cordenons (PN), Italy
[d] OceanCare, Wädenswil, Switzerland

ARTICLE INFO

Keywords:
Genomic extinction
Hybrids
Hybridization
Matrix models
Mediterranean common dolphin
Population projections
Prevalence
Sensitivity analysis
Striped dolphin
Wolf

ABSTRACT

Hybridization affects the evolution and conservation status of species and populations. Because the dynamics of hybridization is driven by reproduction and survival of parental and admixed individuals, demographic modelling is a valuable tool to assess the effects of hybridization on population viability, e.g., under different management scenarios. While matrix models have been used to assess the long-term consequences of hybridization between crops and wild plants, to our knowledge they have not been developed for animal species. Here, we present a new matrix population model to project population dynamics in a system with two parental species or populations that interbreed. We consider the dynamics of males and females of the two parental groups as separate components, each described by species-specific vectors of initial abundance and projection matrices. Then we model hybridization as the production of hybrid fertile offspring due to the interaction of reproductive individuals of different parental species. Finally, we apply the model to two real-world case studies regarding a terrestrial and a marine mammal species in the presence of hybridization. Specifically, we investigate 1) the genomic extinction probability of two interbreeding dolphin species within a semi-enclosed gulf in Greece, under different hybrids' fitness scenarios, 2) the possible outcomes of wolf x dog hybridization events for an expanding wolf population in Italy, under different reproductive isolation scenarios, 3) the sensitivity of the probability of genomic extinction to the main demographic parameters in the two case studies.

1. Introduction

Hybridization, defined as the interbreeding of individuals from genetically distinct populations, regardless of their taxonomic status (Allendorf et al., 2001) is recognized as a relatively common phenomenon both in plants and animals (Hewitt, 1988; Olden et al., 2004; Grabenstein and Taylor, 2018). Hybridization is most commonly observed between otherwise allopatric taxa that come into contact due to natural (natural hybridization) or anthoropogenic causes (anthropogenic hybridization, e.g., human encroachment or the release of non-native taxa). The widespread occurrence of natural hybridization is raising attention due to its implications in evolutionary biology (Abbott et al., 2016). Additionally, the increasing occurrence of anthropogenic hybridization is considered a significant threat to biodiversity (Rhymer and Simberloff, 1996; Seehausen et al., 2008; Crispo et al., 2011).

Hybridization can have different consequences for the evolution and conservation of species. If the fitness of the admixed individuals is lower than that of parentals, hybridization can reinforce reproductive isolation between incompletely isolated species (Barton and Hewitt, 1989), but it can also cause extinction through demographic swamping (Allendorf et al., 2001; Wolf et al., 2001). If the fitness of the admixed individuals is greater than or equal to that of parental individuals, hybridization can cause fusion of species (Seehousen et al., 1997; Allendorf et al., 2001), genetic swamping (Allendorf et al., 2001), transfer of genetic material between species (potentially facilitating their adaptive evolution; Grant and Grant, 1992; Verhoeven et al., 2011), and the origin of new species (DeMarais et al., 1992). Understanding the potential consequences of hybrization is important to unveil evolutionary mechanisms such as how species integrity is maintained in the face of interspecific (and often intergeneric) gene flow (Crossman et al., 2016) and how new species can arise from the

* Corresponding author: Nina Luisa Santostasi
E-mail address: n.santostasi@gmail.com (N.L. Santostasi).

https://doi.org/10.1016/j.ecolmodel.2020.109120
Received 25 September 2018; Received in revised form 5 January 2020; Accepted 6 May 2020
0304-3800/ © 2020 Elsevier B.V. All rights reserved.

introduction of new allelic combinations generated by hybridization. Furthermore, understanding anthropogenic hybridization dynamics can help identify effective and timely management actions for threatened species. To this end there has been an effort to calibrate management actions to the severity of the threat by classifiying hybridization in different types (Allendorf et al., 2001; Bohling, 2016) based *inter alia* on admixed individuals fitness and relative abundance (i.e., prevalence, Santostasi et al., 2019). Depending on the hybridization type, a variety of management actions can be more effective or feasible to avoid genomic extinction: from admixed individuals removal and/or sterilization to the management of the human disturbances that cause hybridization in the first place (Allendorf et al., 2001; Bohling, 2016).

It is often difficult to define hybridization types, *inter alia* because of lack of information about admixed individuals prevalence and fitness. Projection models can tackle this uncertainty by simulating hybridization dynamics under different biological/evolutionary scenarios (Wolf et al., 2001; Fredrickson and Hedrick, 2006). Sensitivity analysis and/or the simulation of different possible management actions can also be used to provide management recommmendatios (Crouse et al., 1987; Cross and Beissinger, 2001).

The first attempts to model hybridization-extinction dynamics had a genetic focus and were based on changes in allelic frequencies at one or more loci (Huxel, 1999; Ferdy and Austerlitz, 2002). Compared to genetic models, ecological models place a greater emphasis on life-history traits, by explicitly examining the effects of fitness parameters (e.g., survival and reproductive rates) on the hybridization outcome. Within the ecological approach, two types of model have been used to model hybridization dynamics (Hall and Ayres, 2008): 1) individual-based models that simulate the contribution of each individual to the hybridization dynamics of the entire population (e.g., Thompson et al., 2003; Hooftman et al., 2007), and 2) population-based models that can be used when only the mean fitness parameters of the main demographic stages are available (e.g., Wolf et al., 2001; Campbell et al., 2002). Both modelling approaches were applied to hybridization in plant species (Hall and Ayres, 2008; Todesco et al., 2016). However, to our knowledge, few studies used individual-based models (Fredrickson and Hedrick, 2006; Nathan et al., 2019) and none used population-based models to simulate hybridization dynamics in animal species.

Here, we develop a population-based approach to project the dynamics of animal hybridization. By using a stage-based matrix model, and grouping individuals into genealogical categories, each described by their mean fitness parameters (i.e., age-specific survival, per capita fertility rate), our approach overcomes the need of realistic and accurate data at the individual level. To provide a practical example of this modeling approach, we illustrate its application to two case studies. The first refers to two delphinid species belonging to different genera, that interbreed in the Gulf of Corinth, Greece (Bearzi et al., 2016; Antoniou et al., 2019): the striped dolphin *Stenella coeruleoalba* and the common dolphin *Delphinus delphis*. The Mediterranean subpopulations of both species are classified as Vulnerable and as Endangered respectively in the IUCN Red List (Bearzi et al., 2003; Aguilar and Gaspari, 2012). Within the Gulf of Corinth, however, the subpopulation of common dolphin qualifies as Critically Endangered due to its small size (point estimate 22 individuals; Santostasi et al., 2016) and high ($\geq$50%) probability of extinction in three generations (Santostasi et al., 2018). Recently confirmed hybridization with the much more abundant striped dolphin population (point estimate 1331 individuals; Santostasi et al., 2016) threatens the persistence of common dolphins in the Gulf of Corinth (Santostasi et al., 2018; Antoniou et al., 2019). We regard this as a case of anthropogenic hybridization, considering that the steep decline of common dolphins due to anthropogenic impacts (Bearzi et al., 2003) locally leads to a lack of available mates and to mating with more abundant species (Frantzis and Herzing, 2002; Antoniou et al., 2019). The possible hybridization outcomes for the two isolated dolphin subpopulations in the Gulf of Corinth have not been previously explored (Antoniou et al., 2019).

In the second case study we evaluate the possible outcomes of hybridization between the wolf (*Canis lupus*) and its domestic counterpart, the dog (*Canis lupus familiaris*). Hybridization between wolves and dogs has been documented in several European countries and represents a well-known threat to wolf conservation (Boitani, 2000; 2003). Still, little is known about the possible outcomes of hybridization between wolves and dogs. Hybridization with dogs may represent a problem for recovering wolf populations expanding into human-dominated landscapes, where few potential wolf mates compete with free-ranging dogs that are abundant and widespread (Randi, 2008; Galaverni et al., 2017). Although reproductive isolation due to behavioral or physiological barriers has been often assumed to contrast introgressive hybridization (Vilà and Wayne, 1999; Randi and Lucchini, 2002; Galaverni et al., 2017), admixed wolf populations are increasingly being reported where wolves live in close contact with free-ranging dogs population (e.g., Italy: Caniglia et al., 2013; Galaverni et al., 2017; Salvatori et al., 2019).

The matrix model presented here allows to quantitatively assess the possibles outcomes of hybridization (i.e., genomic extinction vs. persistence) under different fitness scenarios. Providing management recommendations is beyond the scope of this study, but our model represents a valuable tool to inform management once appropriately customized and parametrized. While our focus is on mammalian species, the analytical approach described here is valid for other taxa, and it could be adopted to project the dynamics of admixed populations for situations entailing both natural and anthropogenic hybridization.

2. Methods

2.1. General model

We consider a system in which there are two parental groups (T1 and T2) that interbreed and produce an admixed progeny (H). We regard the admixed progeny as an absorbing state encompassing all offspring produced by pairs of different parental groups, where at least one of the parents is an admixed individual (Wolf et al., 2001). The possible crosses considered and the produced progeny are listed in Table 1. Based on the assumed dynamics of interbreeding and production of offspring, we project the future abundance of the three mixing groups over time. We present the projections step by step with linear equations and we introduce the equivalent matrix formulation to calculate the asymptotic growth rate and perform sensitivity analyses. Finally, we discuss the behavior of the model by applying it to our case studies.

2.1.1. Model equations

We assume that parental and admixed groups have similar life cycles, with three age classes: offspring (C individuals up to 1 year old), juveniles (J non-reproductive individuals up to 3 years old) and adults (A reproductive individuals $\geq$ 3 years old). The transitions among age

Table 1
Possible crosses in the system formed by two parental taxa (T1 and T2) and admixed individuals (H).

Female parent	Male parent	Offspring
T1	T1	T1
T2	T2	T2
T1	T2	H
T2	T1	H
H	T1	H
H	T2	H
T1	H	H
T2	H	H
H	H	H

classes are described by survival parameters (S) and the reproductive parameters are described by per capita fertility rates (f). At time t, the total number of individuals of the different groups (the parentals and the admixed) is:

$$N_{TOT} = N_1 + N_2 + N_h,$$

where N_1 is the total number of individuals in group 1, N_2 is the total number of individuals in group 2 and N_h is the total number of admixed individuals.

We model the dynamics of females and males separately. Below, we show the equations for females. Each group at time t is composed by females (f) and males (m) belonging to the three different age classes:

$$N_{1f}(t) = C_{1f}(t) + J_{1f}(t) + A_{1f}(t),$$
$$N_{2f}(t) = C_{2f}(t) + J_{2f}(t) + A_{2f}(t),$$
$$N_{hf}(t) = C_{hf}(t) + J_{hf}(t) + A_{hf}(t).$$

We assume that the rate of reproduction between adult females belonging to one group and adult males belonging to one of the three other groups, is proportional to the relative abundances of adult males $\alpha(t)$, $\beta(t)$ and $\gamma(t)$ which are given by the ratio between adult males (A_m) of one group divided by the total number of adult males $TOT_m(t)$:

$$\alpha(t) = A_{1m}(t)/TOT_m(t),$$
$$\beta(t) = A_{2m}(t)/TOT_m(t),$$
$$\gamma(t) = A_{hm}(t)/TOT_m(t).$$

Therefore, the number of offspring belonging to the parental group 1 produced at time t+1 is given by the number of females surviving to time t+1, $A_{1f}(t)Sa_1$, multiplied by their per capita fertility rate (f_1) multiplied by the relative abundance of adult males of group 1 $\alpha(t)$. The number of offspring belonging to the parental group 2 produced at time t+1 is therefore:

$$C_2(t+1) = A_{2f}(t)f_2Sa_2\beta(t).$$

The number of admixed offspring produced by e.g., the crossing between females of the group 1 and males of the group 2 is calculated as $A_{1f}(t)f_1Sa_1\beta(t)$ and the total number of admixed offspring at time t +1 is given by the sum of the contribution of all the possible crosses (Table 1):

$$\begin{aligned} C_h(t+1) &= A_1(t)f_1Sa_1\beta(t) + A_2(t)f_2Sa_2\alpha(t) + A_1(t)f_1Sa_1\gamma(t) \\ &\quad + A_2(t)f_2Sa_2\gamma(t) \\ &\quad + A_h(t)f_hSa_h\alpha(t) + A_h(t)f_hSa_h\beta(t) + A_h(t)f_hSa_h\gamma(t), \end{aligned}$$

where $A_{2f}(t)$ is the number of adult females of group 2 at time t, Sa_1 and Sa_2 are group-specific adult survival values, f_1 and f_2 are the group-specific per capita fertility rates. Because we model separately males and females, the number of offspring of each sex produced every year by each group is obtained by multiplying the total number of offspring by 0.5, assuming a 50:50 sex ratio at birth:

$$C_{1f}(t+1) = C_{1m}(t+1) = C_1(t+1)0.5,$$
$$C_{2f}(t+1) = C_{2m}(t+1) = C_2(t+1)0.5,$$
$$C_{hf}(t+1) = C_{hm}(t+1) = C_h(t+1)0.5.$$

Depending on the species mating system, hybridization can be modelled in different ways. For example, for species in which only the dominant individuals reproduce, it can be convenient to model hybridization at the level of the formation of the reproductive pairs. We consider this situation when dealing with the wolf x dog case study (section 3.2).

The number of female adults and juveniles at time t + 1 for the three groups are obtained as follows (the equations are showed only for group 1):

$$J_{1f}(t+1) = C_{1f}(t)Sc_1$$
$$A_{1f}(t+1) = J_{1f}(t)Sj_1 + A_{1f}(t)Sa_1$$

where Sc, Sj and Sa are respectively survival rates for offspring, juveniles and adults. The total number of females at time t + 1 is therefore:

$$\begin{aligned} N_{1f}(t+1) &= A_{1f}(t)f_1Sa_1\alpha(t)0.5 + C_{1f}(t)Sc_1 + J_{1f}(t)Sj_1 + A_{1f}(t)Sa_1 \\ N_{2f}(t+1) &= A_{2f}(t)f_2Sa_2\beta(t)0.5 + C_{2f}(t)Sc_2 + J_{2f}(t)Sj_2 + A_{2f}(t)Sa_2 \\ N_{hf}(t+1) &= [A_{1f}(t)f_1Sa_1\beta(t) + A_{2f}(t)f_2Sa_2\alpha(t) + A_{1f}(t)f_1Sa_1\gamma(t) \\ &\quad + A_{2f}(t)f_2Sa_2\gamma(t) + A_{hf}(t)f_hSa_h\alpha(t) + A_{hf}(t)f_hSa_h\beta(t) \\ &\quad + A_{hf}(t)f_hSa_h\gamma(t)]0.5 + C_{hf}Sc_h + J_{hf}(t)Sj_h + A_{hf}(t)Sa_h \end{aligned}$$

The number of male individuals in the three groups is obtained with the following equations:

$$\begin{aligned} N_{1m}(t+1) &= A_{1f}(t)f_1Sa_1\alpha(t)0.5 + C_{1m}Sc_1 + J_{1m}(t)Sj_1 + A_{1m}(t)Sa_1 \\ N_{2m}(t+1) &= A_{2f}(t)f_2Sa_2\beta(t)0.5 + C_{2m}Sc_2 + J_{2m}(t)Sj_2 + A_{2m}(t)Sa_2 \\ N_{hm}(t+1) &= [A_{1f}(t)f_1Sa_1\beta(t) + A_{2f}(t)f_2Sa_2\alpha(t) + A_{1f}(t)f_1Sa_1\gamma(t) \\ &\quad + A_{2f}(t)f_2Sa_2\gamma(t) \\ &\quad + A_{hf}(t)f_hSa_h\alpha(t) + A_{hf}(t)f_hSa_h\beta(t) + A_{hf}(t)f_hSa_h\gamma(t)] \\ &\quad 0.5 + C_{hm}Sc_h + J_{hm}(t)Sj_h + A_{hm}(t)Sa_h \end{aligned}$$

2.1.2. Matrix formulation

The model above can be conveniently formulated as $\boldsymbol{N}(t+1) = \boldsymbol{AN}(t)$ where:

$$\boldsymbol{N}(t) = [C_1(t)J_1(t)A_1(t)C_2(t)J_2(t)A_2(t)C_h(t)J_h(t)A_h(t)]^T$$

and

$$\boldsymbol{A} = \begin{pmatrix} 0 & 0 & f_1Sa_1\alpha(t)0.5 & 0 & 0 & 0 & 0 & 0 & 0 \\ Sc_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & Sj_1 & Sa_1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & f_2Sa_2\beta(t)0.5 & 0 & 0 & 0 \\ 0 & 0 & 0 & Sc_2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & Sj_2 & Sa_2 & 0 & 0 & 0 \\ 0 & 0 & f_1Sa_1(\beta(t)+\gamma(t))0.5 & 0 & 0 & f_2Sa_2(\alpha(t)+\gamma(t))0.5 & 0 & 0 & f_hSa_h(\alpha(t)+\beta(t)+\gamma(t))0.5 \\ 0 & 0 & 0 & 0 & 0 & 0 & Sc_h & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & Sj_h & Sa_h \end{pmatrix}$$

In the following section, we consider two real-world scenarios illustrating the relevance of the model described above. All the analyses are performed with the software R (R core team, 2018).

2.2. Applications

2.2.1. Dolphin model

We built stage-based matrices (Fig. 1; Taylor et al., 2007) for the two parental species and the admixed individuals with the following stages: calf (individuals up to 1 year old), juveniles (non-reproductive individuals up to species-specific age of first reproduction reported by Taylor et al., 2007) and adults (individuals that reached the age of first reproduction). We used the available stage-specific demographic parameters for the target populations in the Gulf of Corinth, Greece (Santostasi et al., 2016) and we used parameters estimated for other populations of the same species for the non available stage-specific demographic parameters (see Table 2 and Appendix A for details about model parametrization). We built three fitness scenarios for admixed individuals (Table 2): i) in the Null Model we assumed that admixed individuals had intermediate demographic traits between the two species, ii) in the Hybrid Vigour scenario we assumed that hybrids have higher survival (the upper 95% confidence limit estimated for the population by Santostasi et al., 2016) and annual per capita fertility rate (the highest annual pregnancy rate reported for Atlantic common dolphin subpopulations reported in Murphy et al., 2009) and that they

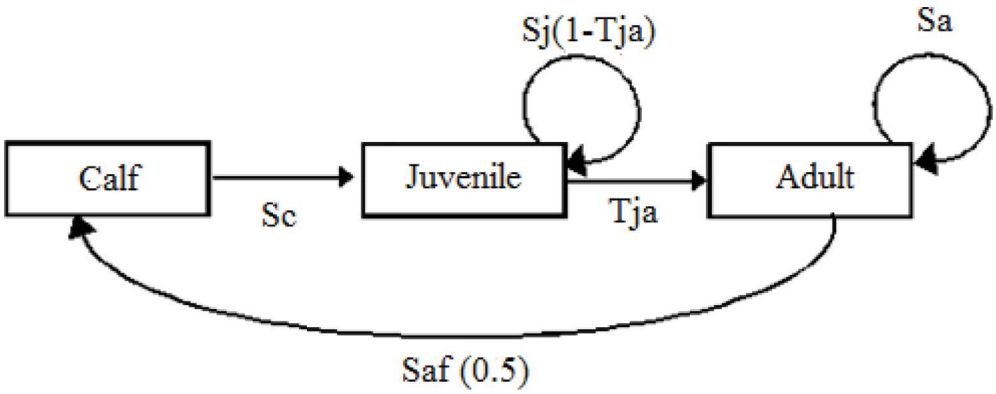


Figure 1. Life cycle used for the dolphin case study. Parameters are: Sc = survival of calves, Tja = transition rate from the juvenile to the adult stage modelled as the survival of juveniles (Sj) raised to the power of age of first reproduction - 1, Sa = survival of adults, f = per capita fertility rate, approximated as the annual pregnancy rate.

Table 2
Demographic parameters used for projecting the abundance of striped and common dolphins in the presence of hybridization using alternative parental and admixed individual fitness scenarios.

Scenario	All scenarios		Null	Outbreeding	Hybrid Vigour
	Striped dolphin	Common dolphin	Admixed dolphin	Admixed dolphin	Admixed dolphin
Initial abundance	1331[a]	22[a]	55[a]	55[a]	55[a]
Per capita fertility rate (f)	0.25[b]	0.26[c]	0.26	0.19[d]	0.33[e]
Age of first reproduction	11[b]	9[c,f]	10	11	9
Calf survival (Sc)	0.80[g]	0.80[g]	0.80	0.80[g]	0.80[g]
Juvenile survival (Sj)	0.94[a]	0.94[a]	0.94	0.92[a]	0.96[a]
Transition rate from the juvenile to the adult stage (Tja)	0.94[(10)]	0.94[(8)]	0.94[(9)]	0.92[(10)]	0.96[(8)]
Adult survival (Sa)	0.94[1]	0.94[1]	0.94	0.92[1]	0.96[1]

[a] Gulf of Corinth, Greece (Santostasi et al., 2016).
[b] Western Mediterranean (Calzada et al., 1997).
[c] Atlantic (Murphy et al., 2009)
[d] Iberian (Murphy et al., 2009).
[e] control group (Murphy et al., 2009).
[f] Eastern North Atlantic (Mannocci et al., 2012).
[g] Theoretical calculation by Taylor et al., (2007).

become reproductively mature earlier (having the youngest age of first reproduction between the two parental species), iii) in the Outbreeding Depression scenario we assumed that admixed individuals have lower survival (the lower 95% confidence limit estimated for this population by Santostasi et al., 2016) and annual per capita fertility rate (the lowest annual pregnancy rate reported for Atlantic common dolphin subpopulations reported in Murphy et al., 2009), and that they become reproductivley mature later (having the oldest age of first reproduction between the two parental species).

The mating systems of odontocete cetaceans (toothed whales) have been reported as either polygynous (some males with multiple partners) or polygynandrous (both males and females with multiple partners; Murphy et al., 2005). We assumed that mature females belonging to one species would reproduce with mature males belonging to the same species, the other species or the admixed individuals, proportionally to their relative abundance. We used deterministic projections (i.e., with constant parameters) to compare the predicted time of extinction for the two species with and without the hybridization effect. The matrix formulation corresponding to the dolphin model is presented below. The subscript Sc refers to striped dolphin, the subscript Dd refers to common dolphins the subscript h refers to admixed individuals. In the Appendix A we show the corresponding R code, the linear equations, and we describe how to incorporate demographic stochasticity.

$$\begin{pmatrix} 0 & 0 & f_{Sc}Sa_{Sc}\alpha(t)0.5 & 0 & 0 & 0 & 0 & 0 & 0 \\ Sc_{Sc} & (1-Tja_{Sc})Sj_{Sc} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & Tja_{Sc} & Sa_{Sc} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & f_{Dd}Sa_{Dd}\beta(t)0.5 & 0 & 0 & 0 \\ 0 & 0 & 0 & Sc_{Dd} & (1-Tja_{Dd})Sj_{Dd} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & Tja_{Dd} & Sa_{Dd} & 0 & 0 & 0 \\ 0 & 0 & f_{Sc}Sa_{Sc}(\beta(t)+\gamma(t))\,0.5 & 0 & 0 & f_{Dd}Sa_{Dd}\,(\alpha(t)+\gamma(t))0.5 & 0 & 0 & f_hSa_h(\alpha(t)+\beta(t)+\gamma(t))0.5 \\ 0 & 0 & 0 & 0 & 0 & 0 & Sc_h & (1-Tja_h)Sj_h & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & Tja_h & Sa_h \end{pmatrix}$$

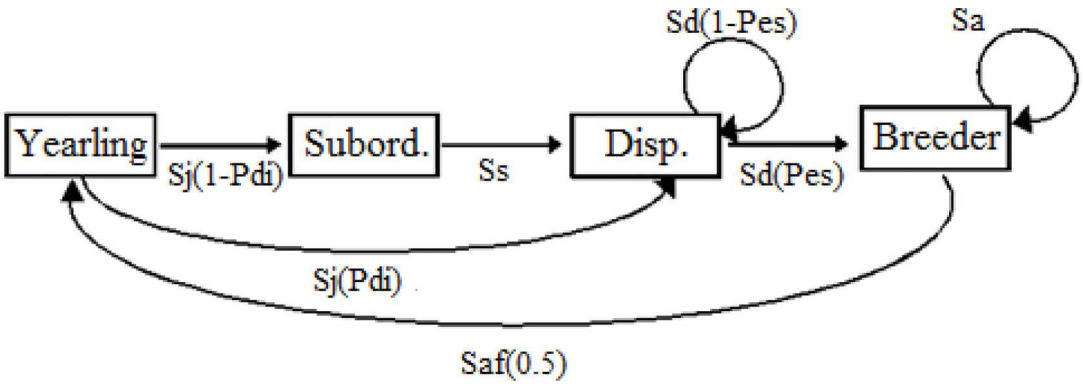


Figure 2. Life cycle for the wolf case study. Parameters are: Sj = survival rate of yearlings, Ss = survival rate of subordinates (yearlings and adults), Sd = survival rate of dispersers (yearlings and adults), Sa = survival rate of breeders, f = annual per capita fertility rate approximated as the litter size, Pdi = dispersal rate, Pes = transition rate to the breeder stage.

Table 3
Rates of formation of the reproductive pairs considered in the model.

Female parent	Male parent	Rate
W	W	$Df_w Sd_w Pes_w \alpha(t)$
W	D	$Df_w Sd_w Pes_h frd(t)$
W	H	$Df_w Sd_w Pes_h \gamma(t)$
H	W	$Df_w Sd_h Pes_h \gamma(t)$
H	H	$Df_w Sd_h Pes_h \gamma(t)$

Df_w = abundance of wolf females in dispersal, Pes_w = transition rate to the breeder stage for wolf females, Pes_h = transition rate to the breeder stage for admixed females, *frd* = constant annual frequency of reproductive events between female wolves and male dogs, α(t) = relative abundance of male wolves in dispersal, γ(t) = relative abundance of admixed males in dispersal.

2.2.2. Wolf x dog model

To model wolf life cycle we used the pre-breeding stage-structured model described in Marescot et al., (2012), that simplifies the complex wolf social structure in four age- and social-stages: yearlings, individuals in dispersal, subordinates, and breeders (Fig. 2). Only the dominant pair breeds, producing a single litter of pups every year. The pups that survived their first year (yearlings), may survive their second year (with survival rate Sj) and disperse (with dispersal rate Pdi), or they may remain in the natal pack (with non-dispersal rate 1 - Pdi) as subordinates (sexually mature individuals 1 to 2 years old, inhibited from breeding by the presence of breeders; Zimen, 1975). The model makes the following assumptions (Marescot et al., 2012): 1) if subordinates survive (with survival rate Ss) one year in the natal pack they will then leave and become individuals in dispersal by their third year of age; 2) the subordinates never directly transition to being breeders; 3) individuals in dispersal can either die or survive (with survival rate Sd), and gain access to reproduction by establishing a new pack (with transition rate to the breeder stage Pes) or remain dispersers (1-Pes); 4) breeders never lose their status remaining in the breeder stage and surviving with survival rate Sa. To run the projections we used the demographic parameters that were estimated for the expanding wolf population in the Italian Alps (Marucco et al., 2009; Marucco and McIntire, 2010 in which hybridization has not been detected yet (Fabbri et al., 2007). The details about model parametrization are shown in Appendix B.

We modeled hybridization at the level of the formation of the reproductive pairs. Hybridization was documented in the almost totality of cases to occur between female wolves and male dogs (Randi et al., 2008; Godinho et al., 2011, Pacheco et al., 2017), while the opposite case of hybridization between male wolves and female dogs appears to be rare (Hindrickson et al., 2012). Therefore, in the model we did not consider reproductive pairs formed by male wolves and female dogs. For simplicity, we did not model the occurrence of backcrossing to dogs (the reproduction of admixed individuals and dogs; Table 3). We assumed that a constant number of reproductive events happens every year between female wolves and male dogs (parameter *frd*). The rate of the formation of reproductive pairs between a female wolf and male wolf at time t +1 is therefore expressed as the product between the abundance of wolf females in dispersal (Df_w) by their survival (Sd_w) by the probability that a female reaches the breeder stage by establishing a new pack (Pes_w) by the relative abundance of wolf males in dispersal *α(t)*. The rates of formation of wolf-dog reproductive pairs, and of reproductive pairs in which at least one individual is admixed are formulated in Table 3.

For wolves, we used the probability of establishing a pack calculated by Marescot et al., (2012) as the transition rate to the breeder stage (Pes). Such annual rate varies uniformly between 0.3 to 0.7. For the Null Model, we assigned the the same average value (Pes = 0.5) to wolves and admixed individuals. We produced two alternative scenarios of reproductive isolation by simulating a reduced probability of establishing a pack for admixed females. In one scenario (Reproductive Isolation 1), we attributed the average value (0.5) to wolves and the minimum value (0.3) to the admixed individuals. In the second scenario (Reproductive Isolation 2) we assigned the maximum probability of establishing a pack to wolves (0.7) and the minimum probability (0.3) to the admixed individuals (Table 4). We used the prevalence of hybrids as a measure to define the type of hybridization reached at the end of the projection time-frame. The matrix formulation corresponding to the wolf x dog model is presented below:

$$\begin{pmatrix}
0 & 0 & 0 & f_w Sa_w 0.5 & 0 & 0 & 0 & 0 \\
Sj_w\, Pdi_w & 0 & Ss_w & 0 & 0 & 0 & 0 & 0 \\
Sj_w\,(1 - Pdi_w) & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & Sd_w Pes_w \alpha(t) & 0 & Sa_w & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & f_h Sa_h 0.5 \\
0 & 0 & 0 & 0 & Sj_h\, Pdi_h & 0 & Ss_h & 0 \\
0 & 0 & 0 & 0 & Sj_h\,(1 - Pdi_h) & 0 & 0 & 0 \\
0 & Sd_w Pes_h(frd + \gamma(t)) & 0 & 0 & 0 & Sd_h Pes_h(\alpha(t) + \gamma(t)) & 0 & Sa_h
\end{pmatrix}$$

Table 4
Demographic parameters used for projecting the abundance of wolves and admixed individuals in the presence of hybridization.

Scenario	Null Model		Repro. Isolation 1		Repro. Isolation 2	
Individual	Wolf	Admixed	Wolf	Admixed	Wolf	Admixed
Initial abundance (reproductive pairs)	6[a]	0	6[a]	0	6[a]	0
Average litter size (f)	3.387[a]	3.387[a]	3.387[a]	3.387[a]	3.387[a]	3.387[a]
Juvenile survival (Sj)	0.551[a]	0.551[a]	0.551[a]	0.551	0.551[a]	0.551
Subordinate survival (Ss)	0.82[a]	0.82[a]	0.82[a]	0.82[a]	0.82[a]	0.82[a]
Disperser Survival (Sd)	0.69[b]	0.69[b]	0.69[b]	0.69[b]	0.69[b]	0.69[b]
Breeder survival (Sa)	0.82[c]	0.82[c]	0.82[c]	0.82[c]	0.82[c]	0.82[c]
Dispersal rate (Pdi)	0.25[d]	0.25[d]	0.25[d]	0.25[d]	0.25[d]	0.25[d]
Transition rate to the breeder stage (Pes)	0.5[d,e]	0.5[d,e]	0.5[d,e]	0.3[d,e]	0.7[d,e]	0.3[d,e]

[a] Marucco and McIntire, 2010.
[b] Blanco and Cortés 2007.
[c] Marucco et al., 2009.
[d] Mech and Boitani, 2003.
[e] Marescot et al., 2012.

In the Appendix B we show the corresponding R code and the linear equations.

2.3. Sensitivity analysis

We performed a sensitiviy analysis to explore the relationship between the hybridization outcome (genomic extinction vs. persistence of parental and admixed populations after 50 years) and the demographic parameters of parental species and admixed individuals (Mc Carthy et al., 1995; Cross and Beissinger, 2001). Here, we use "genomic extinction" to refer to the disappearence of parental individuals from the population (Allendorf et al., 2001). We generated 15,000 parameter sets by drawing them from uniform distributions (Table 5) to emphasize the effects of variability in vital rates on model sensitivity (Cross and Beissinger, 2001). We projected population abundance with each parameter set and checked if the population went extinct or not using a quasi-extinction threshold of 5 parental mature females. We conducted logistic regressions to explore the relationship between the probability of genomic extinction of the two parental species as response variables and demographic parameters used in the projections as independent explanatory variables. For each regression we built full models including all the parameters and used a backward stepwise selection procedures to identify the sets of parameters of potential importance (Fredrickson and Hedrick, 2006). From those sets of parameters we ranked their relative importance in affecting the probability of genomic extinction of the parental species based on their standardized regression coefficients which are the regression coefficient divided by their standard error (Cross and Beissinger, 2001). In the Appendix A and B we show the code to perform the logistic regression sensitivity analysis and how to perform a sensitivity analysis of the asymptotic growth rate using the matrix formulation for the two case studies (R codes are also reported there).

3. Results

3.1. Striped x common dolphin model

Comparing the projections with and without hybridization, we see that when hybridization is ignored, the model predicts an exponential population growth (Fig. 3). When the impact of hybridization is taken into account, the least abundant species (the common dolphin) reaches the quasi-extinction threshold after 16 years, no matter the hybrid fitness scenario. In the Null Model and in the Hybrid Vigour scenario, the growth rate of the most abundant species (the striped dolphin) is also affected, becoming negative after about 40 years in the first scenario and after about 30 years in the second. In the absence of mechanisms that counter hybridization (e.g., assortative mating), the population will eventually be composed of a continuum of admixed classes (the "complete admixture" hybridization type described by Allendorf et al., 2001; Fig. 4). Conversely, under the Outbreeding Depression scenario, striped dolphin population abundance is expected to increase in the next 100 years, predicting a "widespread introgression" hybridization type (i.e., the coexistence of admixed and parental individuals; Fig. 4).

3.2. Wolf x dog model

Depending on the scenario, hybridization had different final outcomes. In the Null Model all the scenarios reached complete admixture (prevalence of admixed individuals = 1) after 50 years and prevalence increased more rapidly at the increasing of the recurrent gene flow from dogs (Fig. 5, upper-left panel). The reproductive isolation scenarios showed different outcomes. In the scenario with the weakest reproductive isolation (Reproductive Isolation 1) the prevalence increased less rapidly compared to the Null Model (Fig. 5, upper-right panel) but did not reach an asymptote, heading towards a final outcome of complete admixture (Allendorf et al., 2001). In the scenario of strongest reproductive isolation (Reproductive Isolation 2), prevalence reached an asymptote whose final value increased at the increasing of the intensity of the recurrent gene flow from dogs (Fig. 5, lower-left panel). In this last case the final outcome is the co-existence of admixed and parental individuals with a constant prevalence (assuming that all the demographic parameters remain constant over time). This last scenario falls into the definition of "widespread introgression" (Allendorf et al., 2001).

3.3. Sensitivity analysis

For the dolphin case study the probability of genomic extinction of both striped and common dolphins was most affected by the survival of mature individuals followed by juvenile survival with almost the same relative importance (Table 6). However, for common dolphins, the initial abundance was also important (Table 6).

For the wolf x dog case study, our sensitivity analysis showed that wolf genomic extinction probability was mostly affected by wolf breeder survival (Sa_w), followed by annual frequency of mating with dogs (*frd*), wolf access to reproduction (i.e., transition rate to the breeder stage, Pes_w) and annual wolf per capita fertility rate f_w, with the same relative importance (Table 7).

4. Discussion

Identifying the demographic factors affecting the outcome of hybridization helps both understanding evolutionary mechanisms and developing meaningful management and conservation measures when

Table 5
Demographic parameters range used for drawing the demographic parameters from uniform distributions to project the abundance of parental and admixed individuals in the sensitivity analyses.

Striped x common dolphin		
Taxon	Parameter	Range
Common dolphin	Initial abundance	16-32[a]
	Age of first reproduction	9-11[b,c,d]
	Calf survival	0.2-1
	Juvenile survival	0.2-1
	Adult survival	0.2-1
	Annual per capita fertility rate	0-0.33[b,c]
Admixed dolphins	Initial abundance	36-84[a]
	Age of first reproduction	9-11[b,c,d]
	Calf survival	0.2-1
	Juvenile survival	0.2-1
	Adult survival	0.2-1
	Annual per capita fertility rate	0-0.33[b,c]
Striped dolphin	Initial abundance	1331-1578[a]
	Age of first reproduction	9-11[b,c,d]
	Calf	0.2-1
	Juvenile survival	0.2-1
	Adult survival	0.2-1
	Annual per capita fertility rate	0-0.33[b,c]
Wolf x dog		
Taxon	Parameter	Range
Wolf	Annual per capita fertility rate	0-4.5[e]
	Juvenile survival	0.2-1
	Subordinate survival	0.2-1
	Disperser Survival	0.2-1
	Breeder survival	0.2-1
	Dispersal rate	0-0.4[e]
	Transition rate to the breeder stage	0-0.7[f]
Admixed	Annual per capita fertility rate	0-4.5[e]
	Juvenile survival	0.2-1
	Subordinate survival	0.2-1
	Disperser Survival	0.2-1
	Breeder survival	0.2-1
	Dispersal rate	0-0.4[e]
	Transition rate to the breeder stage	0-0-7[f]
Dog	Frequency of mating with dogs	0-1

These parameters are also used as explanatory variables in the logistic regression sensitivity analysis.
[a] Gulf of Corinth, Greece (Santostasi et al., 2016).
[b] Western Mediterranean (Calzada et al., 1997).
[c] Control group (Murphy et al., 2009).
[d] Eastern North Atlantic (Mannocci et al., 2012).
[e] Mech and Boitani, 2003.
[f] Marescot et al., 2012.

hybridization is a threat (Bohling, 2016). Compared to the genetic approach (Huxel, 1999; Ferdy and Austerlitz, 2002), our model has the advantage of making the link between hybridization and both demographic (e.g., demographic stochasticity) and ecological factors (e.g., environmental stochasticity). Moreover, our approach allows to relate the viability of the population to the contribution of different fitness components (such as survival and reproductive rates), ultimately providing the basis for sensitivity analyses.

On the other hand, ignoring genetic processes can be an important limitation (Hall and Ayres, 2008). In its present formulation, our model does not make a distinction between ancient (later generations backcrosses) and recent hybrids (first and second generation hybrids), which raises two issues. First, because later generations backcrosses have a limited mixed genomic content, in several cases they are indistiguishable from parentals with the current diagnostic techniques (Vähä and Primmer, 2006). Second, later generations backcrosses are often considered as parentals from a management perspective. For example, in the management of the anthropogenic hybridization between bontebok *Damaliscus pygargus pygargus* and blesbok *D. p. phillipsi* in South Africa, van Wyck et al., (2016) considered that backcrosses to parental bontebooks with $\geq$ 90% of bontebok genes are not to be removed from the population. In the management of the red wolf and coyote, canids with $>$ 87.5 % of red wolf genes are classified as red wolves and not sterilized or culled (Gese et al., 2015). To overcome this limitation, it may be desirable to split the "admixed" group into recent and ancient hybrids, and pool the ancient hybrids together with the parental populations. The threshold between the two groups could be set according to the defined acceptable threshold of admixture in the parentals, while also considering the power of the diagnostic system (Vähä and Primmer, 2006; van Wyck et al., 2016). Despite these limitations, our modelling framework provides a clear way to make explicit the assumptions underlying the definition of hybrids (Thompson et al., 2003).

Even though population-based approaches like the one described here require less-detailed information compared to individual-based approaches (Hall and Ayres, 2008), one still needs to know the survival and reproductive rates for all the population stages, and these may be unavailable for the target population. In this study, we overcame such limitation by resorting to parameters estimated for other populations, which may introduce an unknown level of uncertainty. Moreover, estimates of demographic rates of admixed individuals were not available and are likely to be unavailable in many studies. In such cases, one would need to make assumptions regarding admixed individuals survival and fertility rates. However, sensitivity analysis showed that genomic extinction probability was largely affected by the demographic parameters of parentals, whereas the parameters of admixed individuals had a lower relative importance. Nonetheless, the sensitivity of model predictions to unknown parameters should be tested through sensitivity analyses, and should be taken into account by considering different scenarios. Moreover, the comparison of alternative simulated scenarios with empirical data (i.e., the observed prevalence trends) can provide insight into the evolutionary mechanism shaping the observed patterns.

Another important limitation is that, for simplicity and for the absence of species-specific information, we assumed random mating and did not model more complex mating choice scenarios that could lead to assortative mating. If information about assortative mate choice is available, it can be modelled by changing the reproduction coefficients (α, β and γ in paragraph 2.1.1) from being proportional to the relative abundance of mature males to reflecting the observed percentage of conspecific and heterospecific matings.

At the moment, our projections cannot be validated due to the absence of long-term data on the quantitative evaluation of hybridization for the two case-studies. However, the predicted outcomes are consistent with previous studies showing that hybridization can lead to rapid genomic extinction in the absence of reproductive isolation (Huxel, 1999) or other forms of segregation (Wolf et al., 2001; Fredrickson and Hedrick, 2006). Two studies on wolf x dog hybridization (Salvatori et al., 2019; Santostasi et al., 2019) showed that the prevalence of hybrids can locally reach high levels (around 50%) where reproductive isolation between wolves and dogs is disrupted by anthropogenic disturbance (e.g., presence of free ranging dogs, food provisioning, high wolf anthropogenic mortality), supporting the results of our projections. We suggest that, given the importance of reproductive isolation mechanisms in determining the wolf x dog hybridization outcome, more research should be addressed to understand to what extent and in which circumstances those mehcanisms are effective in contrasting hybridization. In absence of such information, the lack of strong reproductive isolation should be assumed as a precautionary measure.

Studies on hybridization dynamics in cetaceans that could be used to validate our projections are not available. However, our finding that

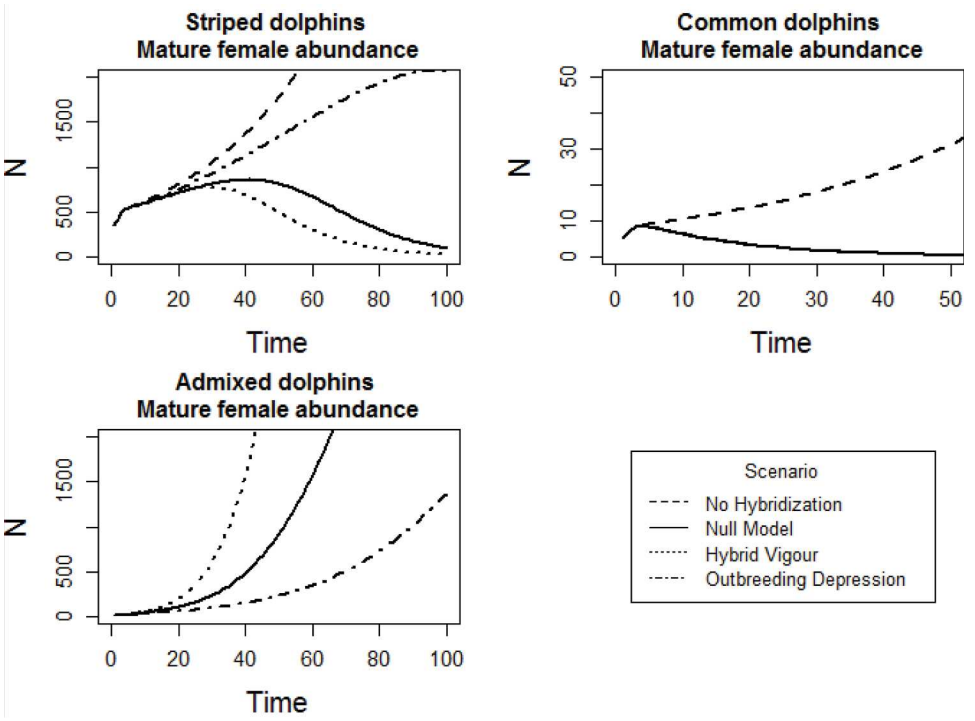


Figure 3. Population trajectories of striped, common and admixed dolphins under different parental and admixed individuals fitness scenarios (the continuos line represents the Null Model while the different types of dashed lines represent the model in absence of hybridization, the Outbreeding Depression Scenario and the Hybrid Vigour Scenario).

the genomic extinction risk for the parental population increases as their initial frequency decreases, making the least abundant species particularly vulnerable, is consistent with previous studies (Allendorf et al., 2001; Epifanio and Philipp, 2001).

We stress that the interest of our model is not the production of absolute predictions of population fate. Being long term hybridization dynamics otherwise difficult to test empirically, the practical value of our model is to illustrate the expected relative outcomes of alternative biological and managament scenarios. The comparison of those outcomes can be useful to highlight future research priorities and to inform

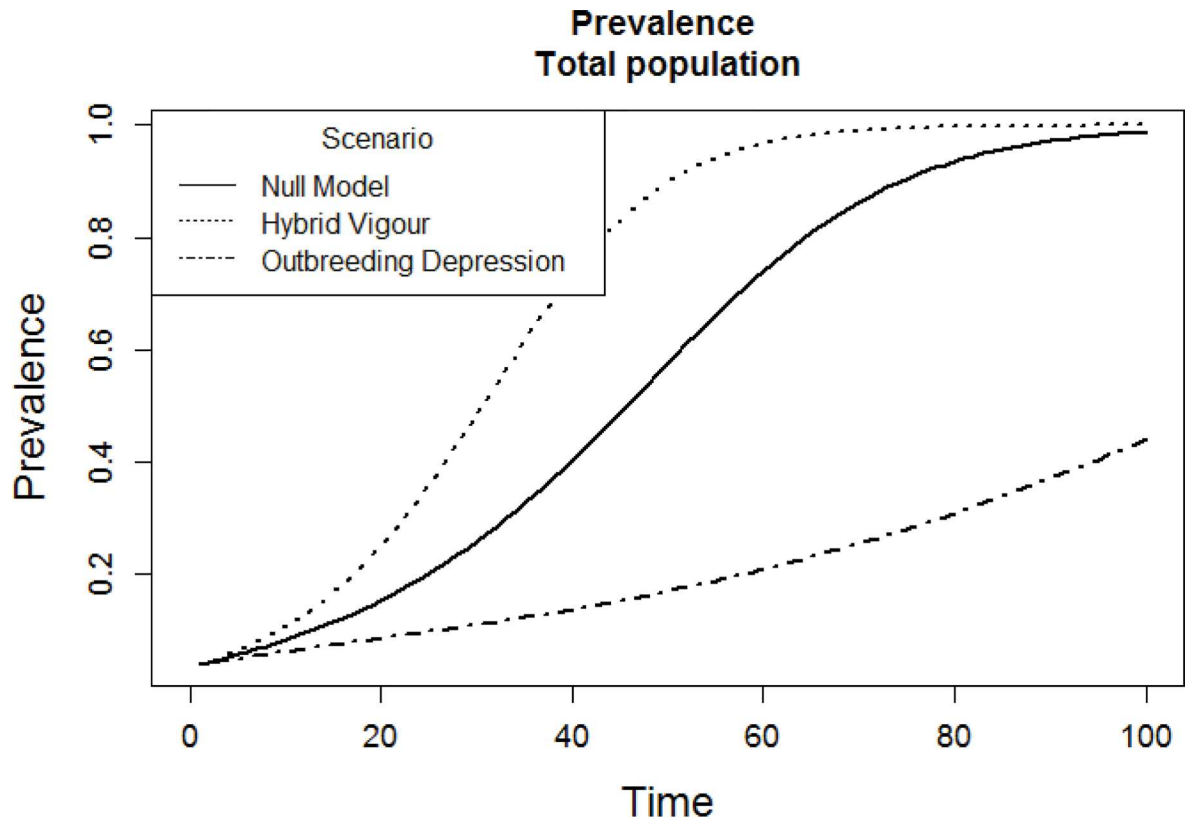


Figure 4. Projected prevalence of striped x common dolphin admixed individuals in the mixed dolphin population of the Gulf of Corinth, Greece under different parental and admixed individuals fitness scenarios (the continuos line represents the Null Model while the different types of dashed lines represent the Outbreeding Depression Scenario and the Hybrid Vigour Scenario).

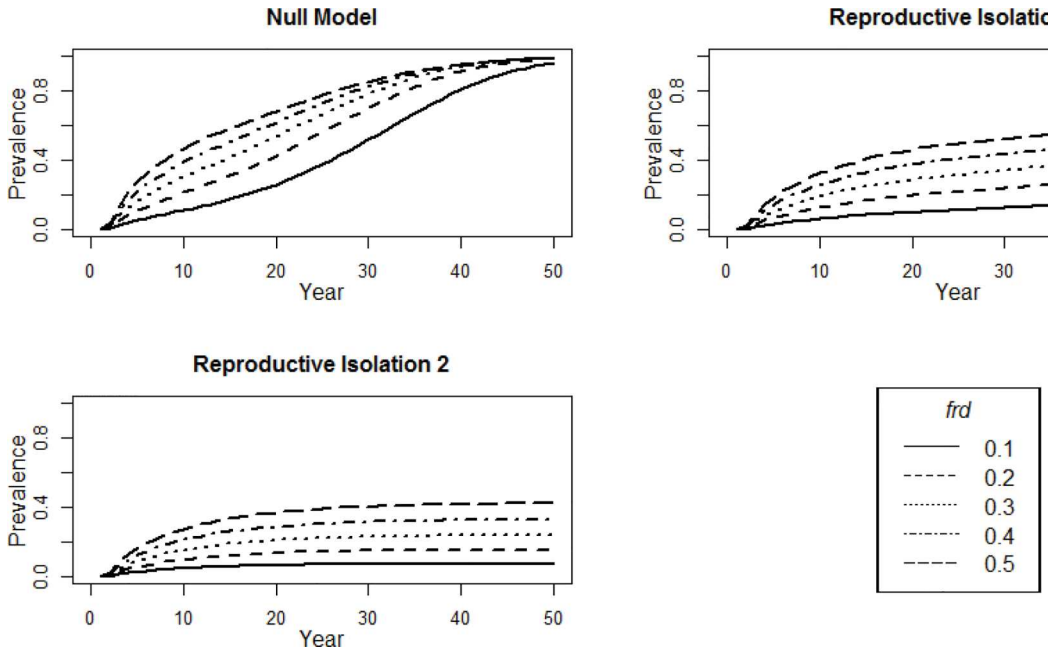


Figure 5. Projected prevalence of wolf x dog admixed individuals under increasing levels of recurrent gene flow from dogs (parameter *frd*) and under different reproductive isolation scenarios: the Null Model (upper-left panel), the Reproductive Isolation 1 scenario (upper-right panel) and the Reproductive Isolation 2 scenario (lower-left panel).

Table 6
Logistic regression sensitivity analysis results for striped and common dolphin population projections in the presence of hybridization.

Parameter	Coefficient	Standard Error	Standardized Coefficient	p value
Striped dolphin genomic extinction				
Sa_{Sc}	-92.90	3.65	-25.44	0.00
Sj_{Sc}	-10.28	0.49	-20.87	0.00
Sc_{Sc}	-2.99	0.33	-9.21	0.00
f_{Sc}	-7.52	1.76	-4.26	0.00
afr_{Sc}	0.24	0.10	2.40	0.02
f_{Dd}	-3.49	1.75	-1.99	0.05
Sj_{Dd}	-0.59	0.31	-1.92	0.05
ia_h	-0.01	0.01	-1.91	0.06
Sa_{Dd}	0.51	0.31	1.64	0.10
Common dolphin genomic extinction				
Sa_{Dd}	-91.40	6.30	-14.50	0.00
Sj_{Dd}	-9.13	0.69	-13.24	0.00
ia_{Dd}	-0.30	0.03	-10.78	0.00
Sc_{Dd}	-1.61	0.47	-3.43	0.00
Sa_{Sc}	1.46	0.43	3.38	0.00
Sa_h	1.17	0.44	2.63	0.01
afr_{Sc}	-0.23	0.15	-1.58	0.11

The parameters are sorted in descending order of relative importance, evaluated by looking at the standardized regression coefficients. Sa = adult survival, Sj = juvenile survival, Sc = calf survival, f = annual per capita fertility rate, ia = initial abundance. The subscript Sc refers to striped dolphin parameters, the subscript Dd refers to common dolphin parameters, the subscript h refers to admixed individuals parameters.

Table 7
Logistic regression sensitivity analysis results for the wolf population projections in the presence of hybridization with dogs.

Parameter	Coefficient	Standard Error	Standardized Coefficient	p value
Sa_w	-47.40	2.03	-23.32	0.00
f_w	-1.17	0.07	-16.89	0.00
Pes_w	-7.14	0.44	-16.05	0.00
frd	5.09	0.32	16.03	0.00
Sj_w	-4.54	0.35	-13.05	0.00
Sd_w	-3.83	0.34	-11.29	0.00
Ss_w	-2.62	0.32	-8.24	0.00
Sj_h	0.97	0.30	3.21	0.00
Pdi_w	-1.38	0.59	-2.34	0.02
Sa_h	0.52	0.31	1.70	0.09

The parameters are sorted in descending order of relative importance, evaluated by looking at the standardized regression coefficients. Sa = annual breeder survival, f = fertility rate (approximated as litter size), Pes (transition rate to the breeder stage), *frd* = annual frequency of mating with dogs, Sj = annual juvenile survival, Sd = annual disperser survival, Ss = annual subordinate survival, Pdi = annual dispersal rate. The subscript w indicates wolf demographic parameters and the subscript h indicates admixed individuals' demographic parameters.

decision-making in a context of uncertainty (Gervasi and Ciucci, 2018).

In alternative, not acknowledging hybridization in the projections may lead to underestimating the risk of genomic extinction. The common dolphin (the least abundant species) had a 100% probability of going extinct after a relatively short time (16 years) regardless the fitness scenario. A previous count-based projection (Santostasi et al., 2018) did not include the effect of hybridization (although it included the effect of demographic stochasticity) and estimated a considerably lower (50%) probability of demographic extinction after 15 years.

Although the two predictions are not entirely comparable due to the different methodologies, our results suggest that hybridization may pose an additional and important threat that must be considered when evaluating the conservation status of common dolphins. This is particularly relevant at larger scales, considering that, in the Mediterranean Sea common dolphins occurr in simpatry with the more abundant striped dolphins and mixed-species groups are not rare (Frantzis and Herzing, 2002; Giménez et al., 2017; Espada et al., 2019).

Interestingly, the population growth rate of the most abundant species, the striped dolphin, was also affected and hybridization could eventually lead this population to genomic extinction, under scenarios in which the fitness of admixed individuals is equal to or greater than that of parentals (i.e., the Null Model and the Hybrid Vigour scenario). However, striped dolphins are predicted to persist and even increase in

the next 100 years, if the fitness of admixed individuals is lower than that of parentals (i.e., the Outbreeding Depression scenario). More research on the fitness of admixed individuals would help understand which scenario represents the most likely alternative. Research on the genetic composition and ecology of admixed individuals would help understand if the predicted "hybrid swarm" may: 1) lead to a new species (Larsen et al., 2010; Lamichhaney et al., 2018); 2) preserve the integrity of the striped dolphin species by backcrossing and dilution of the common dolphin genetic contribution; or 3) allow for the persistence of adaptive common dolphin alleles in the striped dolphin, as suggested by other cases of adaptive introgression (Figueiró et al., 2017).

For the wolf x dog case study, our results confirm that hybridization with dogs should be expected to be a serious threat to the wolf genomic integrity, at least under weak reproductive isolation and/or frequent breeding with dogs. The sensitivity analysis pointed out that the most influential parameters decreasing the chances of genomic extinction are linked to social and reproductive integrity (i.e., survival of wolf breeders and the per capita fertility rate) and the annual frequency of mating with dogs. These results are in agreement with observations of hybridization in eastern wolves (*Canis lycaon*) and coyotes in Canada (Rutledge et al., 2012), and of red wolves and coyotes in North Carolina (Bohling and Waits, 2015). In both cases, high levels of mortality coupled with a large availability of coyotes have been identified as the main causes of hybridization. Our results offer further evidence that human-related factors contributing to hybridization (i.e., increased wolf mortality and dog presence) must be managed to avoid the risk of genomic extinction of wild wolves (Rutledge et al., 2012; Bohling and Waits, 2015).

In conclusion, our model provides estimates of genomic extinction risk in presence of hybridization by using data obtained during demographic monitoring programs of threatened populations. This model allows to link demographic parameters and environmental variables, therefore predicting hybridization dynamics under changing environments. Population projection models clearly represent a valuable tool to predict the outcome of hybridization, therefore contributing to management decisions (Kelly et al., 2010).

Declaration of Competing Interest

We wish to confirm that there are no known conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced its outcome.

Acknowledgements

OG and NLS were supported by the French National Research Agency with a Grant ANR-16-CE02-0007. NLS was supported by a PhD grant by Sapienza University of Rome. We thank the Editor and two anonimous reviewers for their insightful suggestions during the editorial process. We thank Dr. Federico Manna for his help with the matrix formulations.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.ecolmodel.2020.109120.

Literature cited

Abbott, R.J., Barton, N.H., Good, J.M., 2016. Genomics of hybridization and its evolutionary consequences. Mol. Ecol. 25, 2325–2332.
Antoniou, A., Frantzis, A., Alexiadou, P., Paskou, N., Poulakakis, N., 2019. Evidence of introgressive hybridization between *Stenella coeruleoalba* and *Delphinus delphis* in the Greek Seas. Mol. Phylogenet. Evol. 129, 325–337.
Allendorf, F.W., Leary, R.F., Spruell, P., Wenburg, J.K., 2001. The problems with hybrids: setting conservation guidelines. Trends Ecol. Evol. 16, 613–622.
Barton, N.H., Hewitt, G.M., 1989. Adaptation, speciation and hybrid zones. Nature 341, 497–503.
Aguilar, A., Gaspari, S., 2012. *Stenella coeruleoalba* (Mediterranean subpopulation). The IUCN Red List of Threatened Species.
Bearzi, G., Bonizzoni, S., Santostasi, N.L., Furey, N.B., Eddy, L., Valavanis, V.D., Gimenez, O., 2016. Dolphins in a scaled-down Mediterranean: the Gulf of Corinth's odontocetes. Mediterranean Marine Mammal Ecology and Conservation. Adv. Mar. Biol. 75, 297–331.
Bearzi, G., Reeves, R.R., Notarbartolo di Sciara, G., Politi, E., Cañadas, A., Frantzis, A., Mussi, B., 2003. Ecology, status and conservation of short-beaked common dolphins *Delphinus delphis* in the Mediterranean Sea. Mammal. Rev. 33, 224–252.
Bohling, J.H., 2016. Strategies to address the conservation threats posed by hybridization and genetic introgression. Biol. Conserv. 203, 321–327.
Bohling, J.H., Waits, L.P., 2015. Factors influencing red wolf–coyote hybridization in eastern North Carolina, USA. Biol. Conserv. 184, 108–116.
Boitani, L., 2000. Action plan for the conservation of Wolves in Europe (*Canis lupus*). Council of Europe Publishing, Strasbourg, France, pp. 1–84.
Boitani, L., 2003. Wolf conservation and recovery. In: Mech, L.D, Boitani, L. (Eds.), Wolves, Behaviour Ecology and Conservation. University of Chicago Press, Chicago, United States of America, pp. 317–340.
Calzada, N., Aguilar, A., Lockyer, C., Grau, E., 1997. Patterns of growth and physical maturity in the western Mediterranean striped dolphin, *Stenella coeruleoalba*. Can. J. Zool. 75, 632–637.
Campbell, D.R., Waser, N.M., Pederson, G.T., 2002. Predicting patterns of mating and potential hybridization from pollinator behavior. Am. Nat. 159, 438–450.
Caniglia, R., Fabbri, E., Greco, C., Galaverni, M., Manghi, L., Boitani, L., Sforzi, A., et al., 2013. Black coats in an admixed wolf dog pack is melanism an indicator of hybridization in wolves? Eur J Wildl Res 59, 543–555.
Crispo, E., Moore, J., Lee-Yaw, J.A., Gray, S.M., Haller, B.C., 2011. Broken barriers: human-induced changes to gene flow and introgression in animals. BioEssays 33, 508–518.
Cross, P.C., Beissinger, S.R., 2001. Using logistic regression to analyze the sensitivity of PVA models: a comparison of methods based on African wild dog models. Conserv. Biol. 15, 1335–1346.
Crossman, C.A., Taylor, E.B., Barrett-Lennard, L.G., 2016. Hybridization in the Cetacea: widespread occurrence and associated morphological, behavioral, and ecological factors. Ecol. Evol. 6, 1293–1303.
Crouse, D., Crowder, L., Caswell, H., 1987. A stage-based population model for loggerhead sea turtles and implications for conservation. Ecology 68, 1412–1423.
DeMarais, B.D., Dowling, T.E., Douglas, M.E., Minckley, W.L., Marsh, P.C., 1992. Origin of *Gila seminuda* (Teleostei: Cyprinidae) through introgressive hybridization: implications for evolution and conservation. Proc. Natl. Acad. Sci. U.S.A 89, 2747–2751.
Epifanio, J., Philipp, D., 2001. Simulating the extinction of parental lineages from introgressive hybridization: the effects of fitness, initial proportions of parental taxa, and mate choice. Rev. Fish Biol. Fish. 10, 339–354.
Espada, R., Olaya-Ponzone, L., Haasova, L., Martín, E., García-Gómez, J.C., 2019. Hybridization in the wild between *Tursiops truncatus* (Montagu 1821) and *Delphinus delphis* (Linnaeus 1758). PLoS ONE 14(4).
Fabbri, E., Miquel, C., Lucchini, V., Santini, A., Caniglia, R., Duchamp, C., Weber, J.M., et al., 2007. From the Apennines to the Alps: colonization genetics of the naturally expanding Italian wolf (*Canis lupus*) population. Mol. Ecol. 16, 1661–1671.
Ferdy, J., Austerlitz, F., 2002. Extinction and introgression in a community of partially cross-fertile plant species. Amer. Nat. 160, 74–86.
Figueiró, H.V., Li, G., Trindade, F.J., Assis, J., Pais, F., Fernandes, G., Santos, S.H.D., et al., 2017. Genome-wide signatures of complex introgression and adaptive evolution in the big cats. Sci. Adv. 3.
Frantzis, A., Herzing, D.L., 2002. Mixed-species associations of striped dolphins (*Stenella coeruleoalba*), short-beaked common dolphins (*Delphinus delphis*), and Risso's dolphins (*Grampus griseus*) in the Gulf of Corinth (Greece. Mediterranean Sea). Aquat. Mamm. 28, 188–197.
Fredrickson, R.J., Hedrick, P.W., 2006. Dynamics of hybridization and introgression in red wolves and coyotes. Conserv. Biol. 20, 1272–1283.
Galaverni, M., Caniglia, R., Pagani, L., Fabbri, E., Boattini, A., Randi, E., 2017. Disentangling timing of admixture, patterns of introgression, and phenotypic indicators in a hybridizing wolf population. Mol. Biol. Evol. 34, 2324–2339.
Gervasi, V., Ciucci, P., 2018. Demographic projections of the Apennine brown bear population *Ursus arctos marsicanus* (Mammalia: Ursidae) under alternative management scenarios. The European Zoological Journal 85, 242–252.
Gese, E.M., Knowlton, F.F., Adams, J.R., Beck, K., Fuller, T., Murray, D.L., Steury, T.D., et al., 2015. Managing hybridization of a recovering endangered species: the red wolf *Canis rufus* as a case study. Curr. Zool. 61, 191–205.
Giménez, J., Cañadas, A., Ramírez, F., Afán, I., García-Tiscar, S., Fernández-Maldonado, Castillo, J.J., et al., 2017. . Intra- and interspecific niche partitioning in striped and common dolphins inhabiting the southwestern Mediterranean Sea. Mar. Ecol. Progr. Ser. 567, 199–210.
Godinho, R., Llaneza, L., Blanco, J.C., Lopes, S., Álvares, F., García, E.J., Palacios, V., et al., 2011. . Genetic evidence for multiple events of hybridization between wolves and domestic dogs in the Iberian Peninsula. Mol. Ecol. 20, 5154–5166.
Grabenstein, K.C., Taylor, S.A., 2018. Breaking barriers: causes, consequences, and experimental utility of human-mediated hybridization. Trends Ecol. Evol. 2342, 1–15.
Grant, P.R., Grant, B.R., 1992. Hybridization of bird species. Science 256, 193–197.
Hall, R.J., Ayres, D.R., 2008. What can mathematical modeling tell us about hybrid invasions? Biol. Invasions 11, 1217–1224.
Hewitt, G.M., 1988. Hybrid zones: natural laboratories for evolutionary studies. Trends Ecol. Evol. 3, 158–167.
Hindrickson, M., Mannil, P., Ozolins, J., Krzywinsky, A., Saarma, U., 2012. Bucking the

trend in wolf-dog hybridization between female dogs and male wolves. PLoS ONE 7(10).
Hooftman, D.A.P., De Jong, M.J., Oostermeijer, J.G.B., den Nijs, H.C.M., 2007. Modelling the long-term consequences of crop-wild relative hybridization: a case study using four generations of hybrids. J. Appl. Ecol. 44, 1035–1045.
Huxel, G.R., 1999. Rapid displacement of native species by invasive species: effects of hybridization. Biol. Cons. 89, 143–152.
Kelly, B., Whiteley, A., Tallmon, D.A., 2010. The Arctic melting pot. Nature 468, 891.
Lamichhaney, S., Han, F., Webster, M.T., Andersson, L., Grant, B.R., Grant, P.R., 2018. Rapid hybrid speciation in Darwin's finches. Science 359, 224–228.
Larsen, P.A., Marchán-Rivadeneira, M.R., Baker, R.J., 2010. Natural hybridization generates mammalian lineage with species characteristics. Proc. Natl. Acad. Sci. USA. 107 11447–1152.
Mannocci, L., Dabin, W., Augeraud-Véron, E., Dupuy, J-F., Barbraud, C., Ridoux, V., 2012. Assessing the impact of bycatch on dolphin populations: the case of the common dolphin in the Eastern North Atlantic. PLoS ONE 7(2).
Marescot, L., Gimenez, O., Duchamp, C., Marboutinc, E., Chapron, G., 2012. Reducing matrix population models with application to social animal species. Ecol. Model. 232, 91–96.
Marucco, F., McIntire, E.J.B., 2010. Predicting spatio-temporal recolonization of large carnivore populations and livestock depredation risk: wolves in the Italian Alps. J. Appl. Ecol. 47, 789–798.
Marucco, F., Pletscher, D.H., Boitani, L., Schwartz, M.K., Pilgrim, K.L., Lebreton, J-D., 2009. Wolf survival and population trend using non-invasive capture–recapture techniques in the Western Alps. J. Appl. Ecol. 46, 1003–1010.
McCarthy, M.A., Burgman, M.A., Ferson, S., 1995. Sensitivity analysis for models of population viability. Biol. Cons. 73, 93–100.
Mech, L.D., Boitani, L., 2003. Wolves: Behavior, Ecology, and Conservation. University of Chicago Press, Chicago, United States of America, pp. 1–447.
Murphy, S., Collet, A., Rogan, E., 2005. Mating strategy in the male common dolphin (*Delphinus delphis*): what gonadal analysis tells us. J. Mammal. 86, 1247–1258.
Murphy, S., Winship, A., Dabin, W., Jepson, P., Deavill, R., Reid, J.R., Spurrie, C., et al., 2009. Importance of biological parameters in assessing the status of *Delphinus delphis*. Mar. Ecol-Prog. Ser. 388, 273–291.
Nathan, L.R., Mamoozadehb, N., Tumasc, H.R., Gunselmand, S., Klasse, K., Metcalfef, A., Edgeg, C., et al., 2019. A spatially-explicit, individual-based demogenetic simulation framework for evaluating hybridization dynamics. Ecol. Modell. 401, 40–51.
Olden, J.D., LeRoy Poff, N., Douglas, M.R., Douglas, M.E., Fausch, K.D., 2004. Ecological and evolutionary consequences of biotic homogenization. Trends Ecol. Evol. 19, 18–24.
Pacheco, C., López-bao, J.V., García, E.J., Lema, F.J., Llaneza, L., Palacios, V., Godinho, R., 2017. Spatial assessment of wolf-dog hybridization in a single breeding period. Sc. Rep. 7, 1–10.
R Core Team, 2018. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
Randi, E., 2008. Detecting hybridization between wild species and their domesticated relatives. Mol. Ecol. 17, 285–293.
Randi, E., Lucchini, V., 2002. Detecting rare introgression of domestic dog genes into wild wolf (*Canis lupus*) populations by Bayesian admixture analyses of microsatellite variation. Conserv. Genet 3, 31–45.
Rhymer, J.M., Simberloff, D., 1996. Extinction by hybridization and introgression. Annu. Rev. Ecol. Syst. 27, 83–109.
Rutledge Y., L., White N., B., Row R., J., Patterson R., B., 2012. Intense harvesting of eastern wolves facilitated hybridization with coyotes. Ecol. Evol. 2, 19–33.
Santostasi, N.L., Bonizzoni, S., Bearzi, G., Eddy, L., Gimenez, O., 2016. A robust design capture-recapture analysis of abundance, survival and temporary emigration of three odontocete species in the Gulf of Corinth, Greece. PLoS ONE 11(12).
Santostasi, N.L., Bonizzoni, S., Gimenez, O., Eddy, L., Bearzi, G., 2018. Common dolphins in the Gulf of Corinth are Critically Endangered. Aquat. Conserv. 1–9.
Santostasi, N.L., Ciucci, P., Caniglia, R., Fabbri, E., Molinari, L., Reggioni, W., Gimenez, O., 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. Ecol Evol 9, 744–755.
Salvatori V., Godinho, R., Braschi, C., Boitani, L., Ciucci, P., 2019. High levels of recent wolf × dog introgressive hybridization in agricultural landscapes of central Italy. Eur. J. Wildlife Res. 65, 73.
Taylor, B.L., Chivers, S.J., Larese, J., Perrin, W.F., 2007. Generation length and percent mature estimates for IUCN assessments of cetaceans. Administrative Report LJ-07-01 National Marine Fisheries Service, Southwest Fisheries Science Center, La Jolla, CA, United States of America, pp. 1–24.
Thompson, C.J., Thompson, B.J.P., Ades, P.K., Cousens, R., Garnier-Gere, P., Landman, K., Newbigin, E., et al., 2003. Model-based analysis of the likelihood of gene introgression from genetically modified crops into wild relatives. Ecol. Model. 162, 199–209.
Todesco, M., Pascual, M.A., Owens, G.L., Ostevik, K.L., Moyers, B.T., Hübner, S., Heredia, S.M., et al., 2016. Hybridization and extinction. Evol. Appl. 9, 892–908.
Vähä, J.P., Primmer, C.R., 2006. Efficiency of model-based Bayesian methods for detecting hybrid individuals under different hybridization scenarios and with different numbers of loci. Mol. Ecol. 15, 63–72.
van Wyk, A.M., Dalton, D.L., Hoban, S., Bruford, M.W., Russo, I.R.M., Birss, C., Grobler, B., et al., 2016. Quantitative evaluation of hybridization and the impact on biodiversity conservation. Ecol. Evol. 7, 1–11.
Verhoeven, K.J.F., Macel, M., Wolfe, L.M., Biere, A., 2011. Population admixture, biological invasions and the balance between local adaptation and inbreeding depression. Proc. R. Soc. Lond. [Biol] 278, 2–8.
Vilà, C., Wayne, R.K., 1999. Hybridization between wolves and dogs. Conservat. Biol. 13, 195–198.
Wolf, D.E., Takebayashi, N., Rieseberg, L.H., 2001. Predicting the risk of extinction through hybridization. Conservat. Biol. 15, 1039–1053.
Zimen, E., 1975. Social dynamics of the wolf pack. In: Fox, M.V. (Ed.), The Wild Canids: their Systematics, Behavioural Ecology and Evolution. Van Nostrand Reinhold, New York, United States of America, pp. 336–363.

APPENDIX 1: Striped and common dolphins hybridization in the Gulf of Corinth, Greece

1. Initial values and calculations

1.1 Annual pregnancy rate calculation for common dolphins

We used an annual pregnancy rate value derived from Table 5 in Murphy et al., (2009). We calculated the average fertility from different populations in United Kingdom, France, Ireland, Iberian Peninsula, NE Atlantic.

- uk=0.28 *United Kingdom*
- fr=0.29 *France*
- ir=0.20 *Ireland*
- ib=0.19 *Iberian Peninsula*
- ne=0.26 *North-East Atlantic*
- Control group=0.33 *Control group, healthy common dolphins according to Murphy et al., (2009)*

The average annual pregnancy rate is 0.26.

1.2 Calculation of percentage of adults, juveniles and calves and initial abundance for striped and common dolphins

We used the species-specific percentage of mature individuals calculated by Taylor et al., (2007). We divided it by two to obtain the percentage of mature males and females. We assumed an equal number of calves and juveniles among the non-mature individuals (1 - percent mature). To obtain the percentage of calves and juveniles we divided by two the percentage of non-mature individuals. Lastly, to obtain the number of female and male calves we divided by two the total percentages of calves and juveniles. We then multiplied the initial total abundance for the calculated percentages to obtain the initial abundance of adults juvenile and calves. The calculations are shown below.

Striped dolphins

- Percent mature (Taylor et al., 2007) = 0.53
  - Percent mature females = 0.53 / 2 = 0.265
  - Percent mature males = 0.53 / 2 = 0.265
- Percent non-mature individuals = 1 - 0.53=0.47
  - Percent calves = 0.47 / 2 = 0.235

- Percent female calves = 0.235 / 2 = 0.118
- Percent male calves = 0.235 / 2 = 0.118

- Percent juveniles = 0.47 / 2 = 0.235
  - Percent female juveniles = 0.235 / 2 = 0.118
  - Percent male juveniles = 0.235 / 2 = 0.118
- Initial total abundance (Santostasi et al., 2016) = 1,331 individuals
  - $C_{1f}(1)$= 1331 * 0.118 = 158 initial *abundance of female calves*
  - $J_{1f}(1)$ = 1331 * 0.118 = 158 *initial abundance of female juveniles*
  - $A_{1f}(1)$ = 1331 * 0.265 = 353 *initial abundance of female adults*
  - $C_{1m}(1)$ = 1331 * 0.118 = $C_{1f}(1)$ *initial abundance of male calves*
  - $J_{1m}(1)$ = 1331 * 0.118 = $J_{1f}(1)$ *initial abundance of male juveniles*
  - $A_{1fm}(1)$ = 1331 * 0.265 = $A_{1f}(1)$ *initial abundance of male adults*

Common dolphins

- Percent mature (Taylor et al., 2007) = 0.480
  - Percent mature females = 0.480 / 2 = 0.240
  - Percent mature males = 0.480 / 2 = 0.240
- Percent non-mature individuals =1 - 0.48 = 0.52
  - Percent calves = 0.52 / 2 = 0.260
    - Percent female calves = 0.260 / 2 = 0.130
    - Percent male calves = 0.260 / 2 = 0.130
  - Percent juveniles = 0.52 / 2 = 0.260
    - Percent female juveniles = 0.260 / 2 = 0.130
    - Percent male juveniles = 0.260 / 2 = 0.130
- Initial total abundance (Santostasi et al., 2016) = 22 individuals
  - $C_{2f}(1)$ = 22 * 0.130 = 3 *initial abundance of female calves*
  - $J_{2f}(1)$ = 22 * 0.130 = 3 *initial abundance of female juveniles*
  - $A_{2f}(1)$ = 22 * 0.240 = 5 *initial abundance of female adults*
  - $C_{2m}(1)$ = 22 * 0.130 = $C_{2f}(1)$ *initial abundance of male calves*
  - $J_{2m}(1)$ = 22 * 0.130 = $J_{2f}(1)$ *initial abundance of male juveniles*
  - $A_{2m}(1)$ = 22 * 0.240 = $A_{2f}(1)$ *initial abundance of male adults*

Admixed dolphins (striped x common)

- Percent mature (Taylor et al., 2007) = (0.480 + 0.530) / 2= 0.505 = average between striped and common dolphins
    - Percent mature females = 0.505 / 2 = 0.253
    - Percent mature males = 0.505 / 2 = 0.253
- Percent non-mature individuals = 1 - 0.505 = 0.495
    - Percent calves = 0.495 / 2 = 0.248
        - Percent female calves = 0.248 / 2 = 0.124
        - Percent male calves = = 0.260 / 2 = 0.124
    - Percent juveniles = 0.495 / 2 = 0.248
        - Percent female juveniles = 0.248 / 2 = 0.124
        - Percent male juveniles = 0.248 / 2 = 0.124
- Initial total abundance (Santostasi et al., 2016) = 55 individuals
    - $C_{hf}(1) = 55 * 0.124 = 6$ *initial abundance of female calves*
    - $J_{hf}(1) = 55 * 0.124 = 6$ *initial abundance of female juveniles*
    - $A_{hf}(1) = 55 * 0.253 = 14$ *initial abundance of female adults*
    - $C_{hm}(1) = 55 * 0.124 = C_{hf}(1)$ *initial abundance of male calves*
    - $J_{hm}(1) = 55 * 0.124 = J_{hf}(1)$ *initial abundance of male juveniles*
    - $A_{hm}(1) = 55 * 0.253 = A_{hf}(1)$ *initial abundance of male adults*

Initial proportions of males

- $TOT(1) = A_{1m}(1) + A_{2m}(1) + A_{hm}(1) = 371$ *number of total mature males*
    - $\alpha(1) = A_{1m}(1) / TOT_m(1) = 0.95$ *frequency of male striped dolphins*
    - $\beta(1) = A_{2m}(1) / TOT_m(1) = 0.01$ *frequency of male common dolphins*
    - $\gamma(1) = A_{hm}(1) / TOT_m(1) = 0.04$ *frequency of male admixed dolphins*

2. Model equations

Female striped dolphins projections

$C_{1f}(t+1) = A_{1f}(t)Sa_1f_1\alpha(t)0.5$ *female calves at time t+1*

$J_{1f}(t+1) = C_{1f}(t)Sj_1 + J_{1f}Sj_1(1-Sja_1)$ *female juveniles at time t+1*

$A_{1f}(t+1) = J_{1f}(t)Sja_1 + A_{1f}(t)Sa_1$ *female adults at time t+1*

Male striped dolphins projections

$C_{1m}(t+1) = A_{1f}(t)Sa_1f_1\alpha(t)0.5$ *male calves at time t+1*

$J_{1m}(t+1) = C_{1m}(t)Sj_1 + J_{1m}Sj_1(1\text{-}Sja_1)$ *male juveniles at time t+1*

$A_{1m}(t+1) = J_{1m}(t)Sja_1 + A_{1m}(t)Sa_1$ *male adults at time t+1*

Female common dolphins projections

$C_{2f}(t+1) = A_{2f}(t)Sa_2f_2\beta(t)0.5$ *female calves at time t+1*

$J_{2f}(t+1) = C_{2f}(t)Sj_2 + J_{2f}Sj_2(1\text{-}Sja_2)$ *female juveniles at time t+1*

$A_{2f}(t+1) = J_{2f}(t)Sja_2 + A_{2f}(t)Sa_2$ *female adults at time t+1*

Male common dolphins projections

$C_{2m}(t+1) = A_{2f}(t)Sa_2f_2\beta(t)0.5$ *female calves at time t+1*

$J_{2m}(t+1) = C_{2m}(t)Sj_2 + J_{2m}Sj_2(1\text{-}Sja_2)$ *female juveniles at time t+1*

$A_{2m}(t+1) = J_{2m}(t)Sja_2 + A_{2m}(t)Sa_2$ *female adults at time t+1*

Admixed females

$C_{2f}(t+1) =$

$[A_{1f}(t)Sa_1f_1\beta(t) +$ *(female striped x male common)*

$A_{2f}(t)Sa_2f_2\alpha(t) +$ *(female common x male striped)*

$A_{hf}(t)Sa_hf_h\gamma(t) +$ *(female admixed x male admixed)*

$A_{1f}(t)Sa_1f_1\gamma(t) +$ *(female striped x male admixed)*

$A_{2f}(t)Sa_2f_2\gamma(t) +$ *(female common x male admixed)*

$A_{hf}(t)Sa_hf_h\alpha(t) +$ *(female admixed x male striped)*

$A_{hf}(t)Sa_hf_h\beta(t)]0.5$ *(female admixed x male common)*

$J_{hf}(t+1) = C_{hf}(t)Sj_h + J_{hf}Sj_h(1\text{-}Sja_h)$ *female juveniles at time t+1*

$A_{hf}(t+1) = J_{hf}(t)Sja_h + A_{hf}(t)Sa_h$ *female adults at time t+1*

Admixed males

$C_2(t+1) =$

$[A_{1f}(t)Sa_1f_1\beta(t) +$ *(female striped x male common)*

$A_{2f}(t)Sa_2f_2\alpha(t) +$ *(female common x male striped)*

$A_{hf}(t)Sa_h f_h \gamma(t)$ + (*female admixed x male admixed)*

$A_{1f}(t)Sa_1 f_1 \gamma(t)$ + (*female striped x male admixed)*

$A_{2f}(t)Sa_2 f_2 \gamma(t)$ + (*female common x male admixed)*

$A_{hf}(t)Sa_h f_h \alpha(t)$ + (*female admixed x male striped)*

$A_{hf}(t)Sa_h f_h \beta(t)]0.5$ (*female admixed x male common)*

$J_{hm}(t+1) = C_{hm}(t)Sj_h + J_{hm}Sj_h(1\text{-}Sja_h)$ *male juveniles at time t+1*

$A_{hm}(t+1) = J_{hm}(t)Sja_h + A_{hm}(t)Sa_h$ *male adults at time t+1*

Proportion of males of each taxon at time t:

$Tot(t) = A_{1m}(t) + A_{2m}(t) + A_{hm}(t)$ *number of total mature males*

$\alpha(t) = A_{1m}(t) / TOT_m(t)$ *frequence of male striped dolphins*

$\beta(t) = A_{2m}(t) / TOT_m(t)$ *frequence of male common dolphins*

$\gamma(t) = A_{hm}(t) / TOT_m(t)$ frequence of male admixed individuals

The code to implement this model in a matrix formulation is available below:

```
##########################################
#    DOLPHIN CASE: MATRIX FORMULATION     #
##########################################

tspan=200
threshold1=5
threshold2=5

# Build some matrices for storing output
   nf <- matrix(0,9,tspan)      # storage of age-specific female abundance
   nm <- matrix(0,9,tspan)      # storage of age-specific male abundance
   Nf <- matrix(0,tspan,1)      # storage of total female abundance
   Nm <- matrix(0,tspan,1)      # storage of total male abundance
   tot <- matrix(0,tspan,1)     # storage of total abundance
   fr1 <- matrix(0,tspan,1)     # storage of relative abundance male species1
   fr2 <- matrix(0,tspan,1)     # storage of relative abundance male species2
```

```
  frh <-  matrix(0,tspan,1)     # storage of relative abundance admixed
  prevalence <- matrix(0,tspan,1)
  gr  <-  matrix(0,tspan-1,1)   # storage of time-specific  population  growth rates

# Initial  abundances and prevalence
  nf[,1]  <- c(1331*0.118,1331*0.118,1331*0.265,22*0.130,22*0.130,22
  *0.240,55*0.124,55*0.124,55*0.253)      # initial  population  abundance females in each age class

  nm[,1]  <- c(1331*0.118,1331*0.118,1331*0.265,22*0.130,22*0.130,22
  *0.240,55*0.124,55*0.124,55*0.253)      # initial  population  abundance males in each age class

  tot[1,1] <- nm[3,1]+nm[6,1]+nm[9,1]
  fr1[1,1] <- nm[3,1]/tot[1,1]
  fr2[1,1] <- nm[6,1]/tot[1,1]
  frh[1,1] <- nm[9,1]/tot[1,1]
  prevalence[1,1]=sum(nf[7:9,1])/(sum(nf[1:3,1])+sum(nf[4:6,1]))

##ONE PROJECTION##

# Define parameters

# Species 1
  Sc1  <- 0.80 # Calf survival (Taylor et al., 2007)
  Sj1  <- 0.94 # Juvenile  survival (Santostasi et al., 2016)
  Sja1 <- 0.94^10 # Transition  from juvenile  to adult = Annual juvenile  survival^number
  of years to maturity
  Sa1<- 0.94 # Adult survival (Santostasi et al., 2016)
  f1 <- 0.26 # Fertility  (Taylor et al., 2007)

# Species 2
  Sc2<- 0.80 # Calf survival (Taylor et al., 2007)
  Sj2<- 0.94 # Juvenile  survival (Santostasi et al., 2016)
```

```
    Sja2 <- 0.94^8 # Transition from juvenile to adult = Annual juvenile survival^number
    of years to maturity
    Sa2<- 0.94 # Adult survival (Santostasi et al., 2016)
    f2 <- 0.26 # Fertility (average from Table 5 of Murphy et al., 2007)

# Species h
    Sch<- 0.80      # Calf survival (same as parentals)
    Sjh<- 0.94      # Juvenile survival (same as parentals)
    Sjah<- 0.94^9 # Transition from juvenile to adult = Annual juvenile survival^number
    of years to maturity
    Sah<- 0.94      # Adult survival (same as parental)
    fh <- 0.26      # Fertility (average between parentals)

# Create the matrix population model

for (t in 1:(tspan-1)) {

  tot[t,] <- nm[3,t]+nm[6,t]+nm[9,t]   #number of total mature males

  fr1[t,] <- (nm[3,t]/tot[t,])          #Frequency of male S1
  fr2[t,] <- (nm [6,t]/tot[t,])         #Frequency of male S2
  frh[t,] <- (nm [9,t]/tot[t,])         #Frequency of male H

  A <- matrix(c( 0 , 0 , f1*Sa1*fr1[t,]*(0.5) , 0 , 0 , 0 , 0 , 0 , 0,
                 Sc1 , (1-Sja1)*Sj1 , 0 , 0 , 0 , 0 , 0 , 0 , 0,
                 0 , Sja1 , Sa1 , 0 , 0 , 0 , 0 , 0 , 0,
                 0 , 0 , 0 , 0 , 0 , f2*Sa2*fr2[t,]*0.5 , 0 , 0 , 0,
                 0 , 0 , 0 , Sc2 , (1-Sja2)*Sj2 , 0 , 0 , 0 , 0,
                 0 , 0 , 0 , 0 , Sja2 , Sa2 , 0 , 0 , 0,
                 0 , 0 , Sah*f1*(fr2[t,]+frh[t,])*0.5 , 0 , 0 , Sah*f2*(fr1[t,]+frh[t,])*0.5 , 0 , 0 ,
Sah*fh*(fr1[t,]+fr2[t,]+frh[t,])*0.5,
                 0 , 0 , 0 , 0 , 0 , 0 , Sch , (1-Sjah)*Sjh , 0,
```

```
            0 , 0 , 0 , 0 , 0 , 0 , Sjah , Sah),nrow = 9, byrow = TRUE)

 nf[,t+1] <- A%*%nf[,t]      # %*% = matrix multiplication in R
 nm[,t+1] <- A%*%nm[,t]
 Nf[t+1] <- sum(nf[,t+1])
 Nm[t+1] <- sum(nm[,t+1])
 gr[t] <- sum(nf[,t+1])/sum(nf[,t])   # per time step population growth rate
 prevalence[t+1,]=(sum(nf[7:9,t+1]))/(sum(nf[1:9,t+1]))

}


## Plots

### Striped dolphins trajectory
plot_freq=plot(1:(tspan), nf[3,], main="Striped dolphins \nMature female abundance",
           xlab = list("Time",cex=1.5),ylab = list("N",cex=1.5),
           type = "l",lwd = 2,cex.axis = 1.1, xlim=c(0,tspan),ylim=c(0,2000))

### Common dolphins trajectory
plot_freq=plot(1:(tspan), nf[6,], main="Common dolphins \nMature female abundance",
           xlab = list("Time",cex=1.5),ylab = list("N",cex=1.5),
           type = "l",lwd = 2,cex.axis = 1.1, xlim=c(0,tspan),ylim=c(0,50))

### Intermediate dolphins trajectory
plot_freq=plot(1:(tspan), nf[9,], main="Admixed dolphins \nMature female abundance",
           xlab = list("Time",cex=1.5),ylab = list("N",cex=1.5),
           type = "l",lwd = 2,cex.axis = 1.1, xlim=c(0,tspan),ylim=c(0,max(nf[9,])))

### Prevalence trajectory
plot_freq=plot(1:(tspan), prevalence, main="Prevalence \nTotal population",
           xlab = list("Time",cex=1.5),ylab = list("Nh/Ntot",cex=1.5),
```

```
type = "l",lwd  = 2,cex.axis = 1.1, xlim=c(0,tspan))
```

3. Adding demographic stochasticity

Given the reduced size of the common dolphin population we incorporated the effect of demographic stochasticity in the projections. For each time step the number of calves at time t+1 for each group is therefore extracted from a Poisson distribution (Gosselin and Lebreton, 2000) with mean rate = Afα(t) or Afβ(t) Afγ(t). The number of individuals in the other age classes at time t+1 are extracted from N(t) trials in a binomial distribution with probability of success equal to survival probability of that age class (Gosselin and Lebreton, 2000). The projections are repeated for 5000 runs and then the average probability of hitting the extinction threshold and the average time of extinction are calculated.

3.1 Results

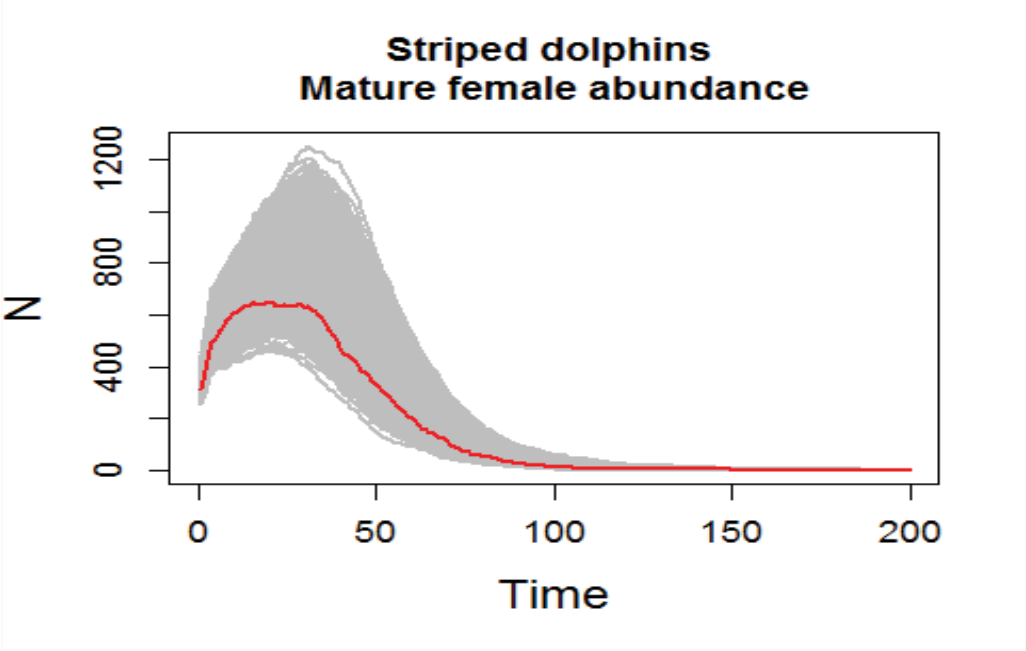


Figure 1. Stochastic projections for mature female striped dolphins. Each grey line represents the result of one simulation run. The red line is one random projection chosen to visualize one trajectory.

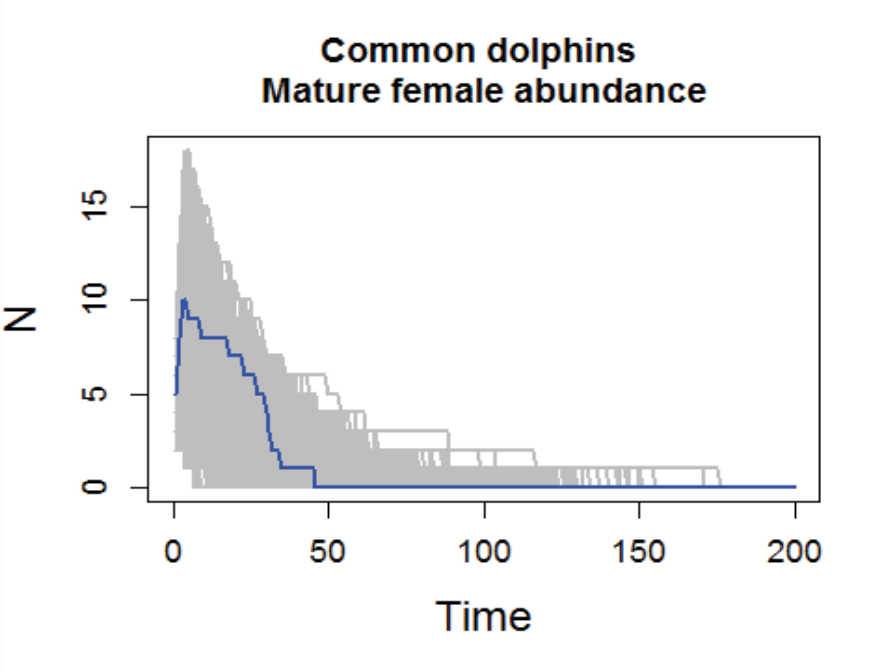


Figure 2. Stochastic projections for mature female common dolphins. Each grey line represents the result of one simulation run. The blue line is one random projection chosen to visualize one trajectory.

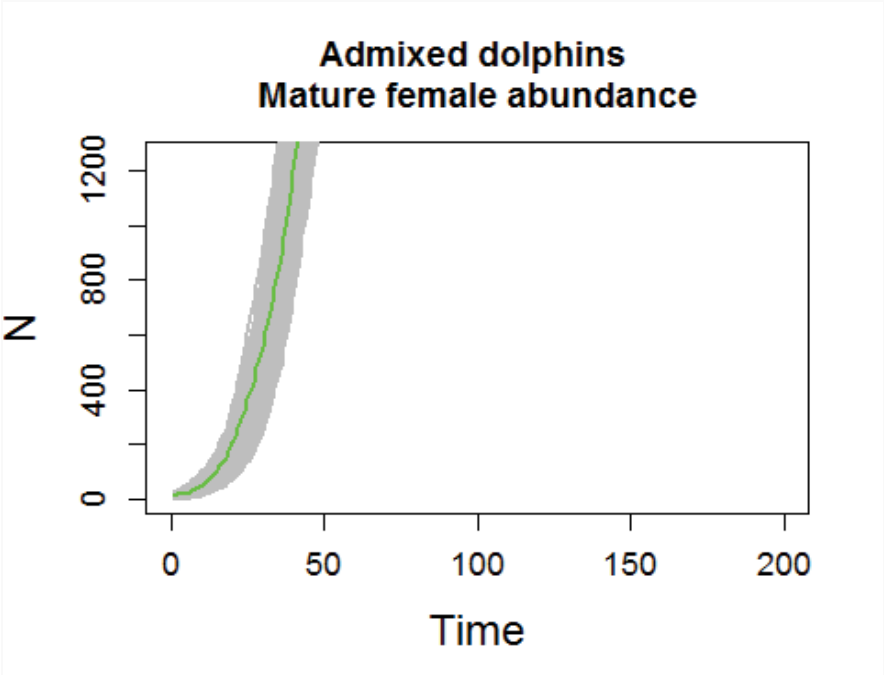


Figure 3. Stochastic projections for admixed mature female dolphins. Each grey line represents the result of one simulation run. The green line is one random projection chosen to visualize one trajectory. Mean time of extinction is 130 years for striped dolphins and 15 years for common dolphins.

The code implement demographic stochasticity is available below.

```
##########################################
#         DEMOGRAPHIC STOCHASTICITY        #
##########################################
tspan=200 # time span for projections
S=5000 # number of simulations
threshold1=5
threshold2=5

#Arrays

# Striped dolphins
   N1f <- array(0,c(3,tspan,S)) # matrix for female projected abundances over time
   N1m <- array(0,c(3,tspan,S)) # matrix for male projected abundances over time
   ini1_striped=1331            # initial abundance (Santostasi et al., 2016)
   SE_Ninit1=113

## Parameters
   sc1 <- 0.80   # Calf survival (Taylor et al., 2007)
   sj1 <- 0.94   # Juvenile survival (Santostasi et al., 2016)
   sja1 <- 0.94^10 # Transition from juvenile to adult = Annual juvenile survival ^ number
   of years to maturity
   sa1<- 0.94     # Adult survival (Santostasi et al., 2016)
   f1 <- 0.25     # Fecundity (Taylor et al., 2007)

# Common
   N2f <- array(0,c(3,tspan,S)) # matrix for female projected abundances over time
   N2m <- array(0,c(3,tspan,S)) # matrix for male projected abundances over time
   ini2_common=22               # initial abundance (Santostasi et al., 2016)
   SE_Ninit2=4
```

```
## Parameters
    sc2<- 0.80     # Calf survival (Taylor et al., 2007)
    sj2<- 0.94     # Juvenile survival (Santostasi et al., 2016)
    sja2 <- 0.94^7 # Transition from juvenile to adult = Annual juvenile survival ^ number
    of years to maturity
    sa2<- 0.94     # Adult survival (Santostasi et al., 2016)
    f2 <- 0.26     # Fertility average from Table 5 of Murphy et al., (2007)

# Hybrid
    Nhf <- array(0,c(3,tspan,S)) # matrix for female projected abundances over time
    Nhm <- array(0,c(3,tspan,S)) # matrix for male projected abundances over time
    inih_hybrid=55               # initial abundance (Santostasi et al., 2016)
    SE_Ninith=12

## Parameters
    sch<- 0.80        # Calf survival (same as parentals)
    sjh<- 0.94        # Juvenile survival (same as parentals)
    sjah<- 0.94^(8.5) # Transition from juvenile to adult = Annual juvenile survival ^
    number of years to maturity
    sah<- 0.94        # Adult survival (same as parentals)
    fh <- 0.28        # Fecundity (average between parentals)

    tot <- matrix(0,S,tspan)  # tot mature male abundance assuming a 1:1 sex ratio
    fr1 <- matrix(0,S,tspan)  # storage of time-specific male S1 frequency
    fr2 <- matrix(0,S,tspan)  # storage of time-specific male S2 frequency
    frh <- matrix(0,S,tspan)  # storage of time-specific male H frequency

    time1=rep(0,S) # time of extinction striped
    time2=rep(0,S) # time of extinction commons
    min1=rep(0,S)  # how many sims hit the extinction threshold for striped dolphins
    min2=rep(0,S)  # how many sims hit the extinction threshold for common dolphins
```

```
for (s in 1:S){

 ### Matrices and Initial abundances for species 1 = striped dolphins

    ini1=round(rnorm(1,mean=ini1_striped,sd= SE_Ninit1))
    N1f[1,1,s] <- round(ini1*0.118) # abundance of female calves at t = 1
    N1f[2,1,s] <- round(ini1*0.118) # abundance of female juveniles individuals at t = 1
    N1f[3,1,s] <- round(ini1*0.265) # abundance of female adult individuals at t = 1

    N1m[1,1,s] <- round(ini1*0.118) # abundance of male calves individuals at t = 1
    N1m[2,1,s] <- round(ini1*0.118) # abundance of male juveniles individuals at t = 1
    N1m[3,1,s] <- round(ini1*0.265) # abundance of male adult individuals at t = 1

 ### Matrices and Initial abundances for species 2 = common dolphins

    ini2=round(rnorm(1,mean=ini2_common,sd= SE_Ninit2))
    ini2

    N2f[1,1,s] <- round(ini2*0.130) # abundance of female calves at t = 1
    N2f[2,1,s] <- round(ini2*0.130) # abundance of female juveniles individuals at t = 1
    N2f[3,1,s] <- round(ini2*0.240) # abundance of female adult individuals at t = 1

    N2m[1,1,s] <- round(ini2*0.130) # abundance of male calves individuals at t = 1
    N2m[2,1,s] <- round(ini2*0.130) # abundance of male juveniles individuals at t = 1
    N2m[3,1,s] <- round(ini2*0.240) # abundance of male adult individuals at t = 1

 ### Matrices and Initial abundances for hybrids = common x striped dolphins

    inih=round(rnorm(1,mean=inih_hybrid,sd= SE_Ninith))
    inih

    Nhf[1,1,s] <- round(inih*0.124) # abundance of female calves at t = 1
```

```
  Nhf[2,1,s] <- round(inih*0.124)  # abundance of female juveniles  individuals  at t = 1
  Nhf[3,1,s] <- round(inih*0.253)  # abundance of female adult individuals  at t = 1

  Nhm[1,1,s] <- round(inih*0.124)  # abundance of male calves individuals  at t = 1
  Nhm[2,1,s] <- round(inih*0.124)  # abundance of male juveniles  individuals  at t = 1
  Nhm[3,1,s] <- round(inih*0.253)  # abundance of male adult individuals  at t = 1

### Initial total abundances of males and frequencies

  tot[s,1]=N1m[3,1,s]+N2m[3,1,s]+Nhm[3,1,s]
  fr1[s,1]=N1m[3,1,s]/tot[s,1]
  fr2[s,1]=N2m[3,1,s]/tot[s,1]
  frh[s,1]=Nhm[3,1,s]/tot[s,1]

for(t in 2:tspan){

 N1f[1,t,s] <- rpois(1,round(N1f[3,t-1,s]*f1*sa1*fr1[s,t-1]*0.5))  #calves
 N1f[2,t,s] <- rbinom(1,round(N1f[1,t-1,s]*sj1+
               (N1f[2,t-1,s]*(1-sja1))),sj1)  #juveniles
 N1f[3,t,s] <- rbinom(1,round(N1f[2,t-1,s]),sja1)+
           rbinom(1,round(N1f[3,t-1,s]),sa1)  #adults

 N1m[1,t,s] <-  rpois(1,round(N1f[3,t-1,s]*f1*sa1*fr1[s,t-1]*0.5))  #calves
 N1m[2,t,s] <-  rbinom(1,round(N1m[1,t-1,s]*sj1+
                (N1m[2,t-1,s]-N1m[2,t-1,s]*sja1)),sj1)  #juveniles
 N1m[3,t,s] <-  rbinom(1,round(N1m[2,t-1,s]*(1-sja1)),sja1)+
           rbinom(1,round(N1m[3,t-1,s]),sa1)  #adults

 N2f[1,t,s] <- rpois(1,round(N2f[3,t-1,s]*f2*sa2*fr2[s,t-1]*0.5))  #calves
 N2f[2,t,s] <- rbinom(1,round(N2f[1,t-1,s]*sj2+
           (N2f[2,t-1,s]*(1-sja2))),sj2)  #juveniles
 N2f[3,t,s] <- rbinom(1,round(N2f[2,t-1,s]),sja2)+
```

```
            rbinom(1,round(N2f[3,t-1,s]),sa2)  #adults

  N2m[1,t,s] <- rpois(1,round(N2f[3,t-1,s]*f2*sa2*fr2[s,t-1]*0.5))  #calves
  N2m[2,t,s] <- rbinom(1,round(N2m[1,t-1,s]*sj2+
            (N2m[2,t-1,s]*(1-sja2))),sj2)  #juveniles
  N2m[3,t,s] <- rbinom(1,round(N2m[2,t-1,s]),sja2)+
            rbinom(1,round(N2m[3,t-1,s]),sa2)  #adults

  Nhf[1,t,s] <- round((rpois(1,round(N1f[3,t-1,s]*sa1*f1*fr2[s,t-1]))+
rpois(1,round(N2f[3,t-1,s]*sa2*f2*fr1[s,t-1]))+
            rpois(1,round(Nhf[3,t-1,s]*sah*fh*frh[s,t-1]))+
            rpois(1,round(N1f[3,t-1,s]*sa1*f1*frh[s,t-1]))+
            rpois(1,round(N2f[3,t-1,s]*sa2*f2*frh[s,t-1]))+
            rpois(1,round(Nhf[3,t-1,s]*sah*fh*fr1[s,t-1]))+
            rpois(1,round(Nhf[3,t-1,s]*sah*fh*fr2[s,t-1])))*0.5)  #calves
  Nhf[2,t,s] <- rbinom(1,round(Nhf[1,t-1,s]*sjh+
            (Nhf[2,t-1,s]*(1-sjah))),sjh)  #juveniles
  Nhf[3,t,s] <- rbinom(1,round(Nhf[2,t-1,s]),sjah)+
            rbinom(1,round(Nhf[3,t-1,s]),sah)  #adults

  Nhm[1,t,s] <-  round((rpois(1,round(N1f[3,t-1,s]*sch*f1*fr2[s,t-1]))+
            rpois(1,round(N2f[3,t-1,s]*sa1*f2*fr1[s,t-1]))+
            rpois(1,round(Nhf[3,t-1,s]*sah*fh*frh[s,t-1]))+
            rpois(1,round(N1f[3,t-1,s]*sa1*f1*frh[s,t-1]))+
            rpois(1,round(N2f[3,t-1,s]*sa2*f2*frh[s,t-1]))+
            rpois(1,round(Nhf[3,t-1,s]*sah*fh*fr1[s,t-1]))+
            rpois(1,round(Nhf[3,t-1,s]*sah*fh*fr2[s,t-1])))*0.5)  #calves
  Nhm[2,t,s] <- rbinom(1,round(Nhm[1,t-1,s]*sjh+
            (Nhm[2,t-1,s]*(1-sjah))),sjh)  #juveniles
  Nhm[3,t,s] <- rbinom(1,round(Nhm[2,t-1,s]),sjah)+
            rbinom(1,round(Nhm[3,t-1,s]),sah)  #adults
```

```
  tot[s,t] <- round((N1m[3,t,s]+N2m[3,t,s]+Nhm[3,t,s]))   #number of total mature males at time t

  fr1[s,t] <- (N1m[3,t,s]/tot[s,t])          #frequency of male S1 at time t
  fr2[s,t] <- (N2m [3,t,s]/tot[s,t])         #frequency of male S2 at time t
  frh[s,t] <- (Nhm [3,t,s]/tot[s,t])         #frequency of male H at time t


  #Extinction probability and time of extinction for striped dolphins

  min1[s]= min(N1f[3,,s])  #vector with minimum value for each projection
  time_temp1=N1f[3,,s]<threshold1  #vector saying if, in the current sim, mature females are below the
threshold
  time1[s]=min(which(time_temp1 == TRUE)) #vector saying when mature females first hit the
threshold

  #Extinction probability and time of extinction for common dolphins

  min2[s]= min(N2f[3,,s])  #vector with minimum value for projection

  time_temp2=N2f[3,,s]<threshold2  #vector saying if, in the current sim, mature females are below the
threshold

  time2[s]=min(which(time_temp2 == TRUE)) #vector saying when mature females first hit the
threshold
 }
}
time_ext1=time1[time1<Inf]  #exclude Inf values
Mean_text1=mean(time_ext1)  #average time of extinction
```

3. **Logistic regression sensitivity analysis script**

The code to reproduce the logistic regression sensitivity analysis presented in the main text is provided below.

```
#####################################################################
#       DOLPHIN CASE: LOGISTIC REGRESSION SENSITIVITY ANALYSIS      #
#####################################################################
npar<-21
sets<-15000 #parameters set for sensitivity analysis
tspan<-50 #projections time
threshold1<-5 #extinction threshold for striped
threshold2<-5 #extinction threshold for common

# Build some matrices for storing output
nf <- array(NA,dim=c(9,tspan,sets))      # storage of age-specific female abundance
nm <- array(NA,dim=c(9,tspan,sets))      # storage of age-specific male abundance

Nf <- matrix(0,tspan,sets)       # storage of total female abundance
Nm <- matrix(0,tspan,sets)       # storage of total male abundance
tot <- matrix(0,tspan,sets)      # storage of total abundance
fr1 <- matrix(0,tspan,sets)      # storage of realtive abundance male species1
fr2 <- matrix(0,tspan,sets)      # storage of realtive abundance male species2
frh <- matrix(0,tspan,sets)      # storage of realtive abundance admixed
prevalence <- matrix(0,tspan,sets)
gr <- matrix(0,tspan-1,sets)     # storage of time-specific population growth rates
regr_data=matrix(0,sets,npar+3)
colnames(regr_data)=c("ia1","Sc1","Sj1","Sja1","Sa1","f1","afr1","ia2","Sc2","Sj2","Sja2","Sa
2","f2","afr2",
                      "iah","Sch","Sjh","Sjah","Sah","fh","afrh","Ext_Sc","Ext_Dd","Ext_H")

##PERFORM N=SETS MODEL# RUNS AND STORE THE RESULTS#

for (s in 1:(sets)) {
```

```
# Define parameters

# Species 1
ia1 =round(runif(1,min=1331, max=1578))
afr1=round(runif(1,min=8, max=10))
Sc1 <- runif(1,min = 0.2, max = 1) # Calf survival (Taylor et al., 2007)
Sj1 <- runif(1,min = 0.2, max = 1)# Juvenile survival (Santostasi et et al., 2016)
Sja1 <- Sj1^afr1# Transition from juvenile to adult = Annual juvenile survival^number of years to maturity
Sa1 <-runif(1,min = 0.2, max = 1)# Adult survival (Santostasi et al., 2016)
f1 <- runif(1,min = 0.19, max = 0.33) # Fertility (Taylor et al., 2007)

# Species 2

ia2=round(runif(1,min=16, max=32))
afr2=round(runif(1,min=8, max=10))
Sc2<- runif(1,min = 0.2, max = 1) # Calf survival (Taylor et al., 2007)
Sj2<- runif(1,min = 0.2, max = 1)# Juvenile survival (Santostasi et a.,l 2016)
Sja2 <- Sj2^afr2 # Transition from juvenile to adult = Annual juvenile survival^number of years to maturity
Sa2<- runif(1,min = 0.2, max = 1) # Adult survival (Santostasi et al., 2016)
f2 <- runif(1,min = 0.19, max = 0.33) # Fertility (average from Table 5 of Murphy et al., 2007)

# Species h
iah =round(runif(1,min=36, max=84))
afrh=round(runif(1,min=8, max=10))
Sch <- runif(1,min = 0.2, max = 1)      # Calf survival (same as parentals)
Sjh <- runif(1,min = 0.2, max = 1)      # Juvenile survival (same as parentals)
Sjah <- Sjh^afrh                        #Transition from juvenile to adult = Annual juvenile survival^number of years to maturity
Sah <- runif(1,min = 0.2, max = 1)      # Adult survival (same as parental)
fh <- runif(1,min = 0.19, max = 0.33)   # Fertility (average between parentals)
```

```
Ext_Sc <- 0
Ext_Dd <- 0
Ext_H  <- 0


# Initial abundances and prevalence
nf[,1,s]  <-
c(ia1*0.118,ia1*0.118,ia1*0.265,ia2*0.130,ia2*0.130,ia2*0.240,iah*0.124,iah*0.124,iah*0.253
)    # initial population abundance females in each age class
nm[,1,s]  <-
c(ia1*0.118,ia1*0.118,ia1*0.265,ia2*0.130,ia2*0.130,ia2*0.240,iah*0.124,iah*0.124,iah*0.253
)    # initial population abundance males in each age class
tot[1,s]  <- nm[3,1,s]+nm[6,1,s]+nm[9,1,s]
fr1[1,s]  <- nm[3,1,s]/tot[1,s]
fr2[1,s]  <- nm[6,1,s]/tot[1,s]
frh[1,s]  <- nm[9,1,s]/tot[1,s]
prevalence[1,s]=sum(nf[7:9,1,s])/(sum(nf[1:3,1,s])+sum(nf[4:6,1,s]))

# Create the matrix population model

for (t in 1:(tspan-1)) {

  tot[t,s] <- nm[3,t,s]+nm[6,t,s]+nm[9,t,s]   #number of total mature males

  fr1[t,s] <- (nm[3,t,s]/tot[t,s])          #Frequency of male S1
  fr2[t,s] <- (nm [6,t,s]/tot[t,s])         #Frequency of male S2
  frh[t,s] <- (nm [9,t,s]/tot[t,s])         #Frequency of male H

  A <- matrix(c( 0 , 0 , f1*Sa1*fr1[t,s]*(0.5)  , 0 , 0 , 0 , 0 , 0 , 0,
                 Sc1 , (1-Sja1)*Sj1  , 0 , 0 , 0 , 0 , 0 , 0 , 0,
                 0 , Sja1 , Sa1 , 0 , 0 , 0 , 0 , 0 , 0,
```

```
            0 , 0 , 0 , 0 , 0 , f2*Sa2*fr2[t,s]*0.5  , 0 , 0 , 0,
            0 , 0 , 0 , Sc2 , (1-Sja2)*Sj2  , 0 , 0 , 0 , 0,
            0 , 0 , 0 , 0 , Sja2 , Sa2 , 0 , 0 , 0,
            0 , 0 , Sah*f1*(fr2[t,s]+frh[t,s])*0.5  , 0 , 0 , Sah*f2*(fr1[t,s]+frh[t,s])*0.5  , 0 , 0 ,
Sah*fh*(fr1[t,s]+fr2[t,s]+frh[t,s])*0.5,
            0 , 0 , 0 , 0 , 0 , 0 , Sch , (1-Sjah)*Sjh  , 0,
            0 , 0 , 0 , 0 , 0 , 0 , 0 , Sjah , Sah),nrow = 9, byrow = TRUE)

  nf[,t+1,s] <- A%*%nf[,t,s]      # %*% = matrix multiplication in R
  nm[,t+1,s] <- A%*%nm[,t,s]
  Nf[t+1,s] <- sum(nf[,t+1,s])
  Nm[t+1,s] <- sum(nm[,t+1,s])
  gr[t,s] <- sum(nf[,t+1,s])/sum(nf[,t,s])   # per time step population growth rate
  prevalence[t+1,s]=(sum(nf[7:9,t+1,s]))/(sum(nf[1:9,t+1,s]))
}

if(any(nf[3,,s] < threshold1)) Ext_Sc=1
if(any(nf[6,,s] < threshold1)) Ext_Dd=1
if(any(nf[9,,s] < threshold1)) Ext_H=1
regr_data[s,]=c(ia1,Sc1,Sj1,Sja1,Sa1,f1,afr1,ia2,Sc2,Sj2,Sja2,Sa2,f2,afr2,iah,Sch,Sjh,Sjah,Sah,f
h,afrh,Ext_Sc,Ext_Dd,Ext_H)

}

#### logistic regression ####

  ###Influence of demographic parameters on striped dolphins genomic extinction

  mylogit1 <- glm(Ext_Sc ~ ia1+Sc1+Sj1+Sa1+f1+ia2+Sc2+Sj2+Sa2+
          f2+Sch+Sjh+iah+Sah+fh+afr1+afr2+afrh, data = as.data.frame(regr_data), family =
"binomial")
  #Removed the sja1,sja2 and sjah bc correlated with sj
```

```
modsel1=stepAIC(mylogit1,trace=F)  #backward stepwise selection function
summary(modsel1)
summary(modsel1)$coefficients  #Coefficients table
Ext_Sc_coeff <- summary(modsel1)$coefficients  #Standardized coefficients table
Ext_Sc_coeff_ord= Ext_Sc_coeff[order(abs(Ext_Sc_coeff[,3]),decreasing=T),]  #Sort the standardized coeff by abs. value
Ext_Sc_coeff_ord


###INFLUENCE OF DEMOGRAPHIC PARAMETERS ON Dd EXTINCTION

    mylogit3  <- glm(Ext_Dd ~ ia1+Sc1+Sj1+Sa1+f1+ia2+Sc2+Sj2+Sa2+
                f2+iah+Sch+Sjh+Sah+fh+afr1+afr2+afrh,  data = as.data.frame(regr_data), family = "binomial")
    modsel3=stepAIC(mylogit3,trace=F)  #backward stepwise selection function
    summary(modsel3)
    summary(modsel3)$coefficients  #Coefficients table
    Ext_Dd_coeff <- summary(modsel3)$coefficients#Standardized coefficients table
    Ext_Dd_coeff_ord= Ext_Dd_coeff[order(abs(Ext_Dd_coeff[,3]),decreasing=T),]  #sort absolute values standardized coeff
    Ext_Dd_coeff_ord

    ###Influence of demographic parameters on common dolphins genomic extinction

    mylogit5  <- glm(Ext_H ~ ia1+Sc1+Sj1+Sa1+f1+Sc2+Sj2+ia2+Sa2+
                f2+iah+Sch+Sjh+Sah+fh+afr1+afr2+afrh,  data = as.data.frame(regr_data), family = "binomial")
    modsel5=stepAIC(mylogit5,trace=F)  #backward stepwise selection function
    summary(modsel5)
    summary(modsel5)$coefficients  #Coefficients table
    Ext_H_coeff <- summary(modsel5)$coefficients  #Standardized coefficients table
```

```
Ext_H_coeff_ord= Ext_H_coeff[order(abs(Ext_H_coeff[,3]),decreasing=T),] #sort absolute values standardized coeff
Ext_H_coeff_ord
```

4. Analytical sensitivity analysis (

The finite annual rate of increase can be calculated as the dominant eigenvalue of the transition matrix **A**. The sensitivity of the growth rate to a proportionate change in each parameter (elasticity) can be obtained using calculus (based on Caswell 1978). The parameters with the higher elasticity and sensitivity was the survival of admixed individuals (Table S1).

Table S1.

Parameter	**Sensitivity**	**Elasticitiy**
Sc1	0.00E+00	0.00E+00
Sj1	0.00E+00	0.00E+00
Sja1	4.12E-12	2.14E-12
Sa1	0.00E+00	0.00E+00
f1	-8.88E-12	-2.14E-12
Sc2	0.00E+00	0.00E+00
Sj2	0.00E+00	0.00E+00
Sja2	0.00E+00	0.00E+00
Sa2	0.00E+00	0.00E+00
f2	0.00E+00	0.00E+00
Sch	0.00E+00	0.00E+00
Sjh	**1.38E-01**	**1.24E-01**
Sjah	2.16E-02	1.19E-02
Sah	8.84E-01	7.99E-01

fh	3.05E-01	7.64E-02

The code to implement the sensitivity analysis is available below:

```
##############################################
#                SENSITIVITY ANALYSIS        #
##############################################

# Create the dolphin population model function with generic parameters

source('build_matrix_dolphin.r')

# assign values to model param

source('input_param_dolphin.r')

# vector of parameters
theta <- c(Sc1,Sj1,Sja1,Sa1,f1,Sc2,Sj2,Sja2,Sa2,f2,Sch,Sjh,Sjah,Sah,fh)
theta

# Apply function build_matrix to obtain the dolphin matrix
A <- build_matrix_dolphin(theta)
A

# compute asymptotic growth rate
max(Re(eigen(A)$values))

# Create function to calculate sensitivity wrt parameter in position pos
# in vector of parameters that follows and matches the input parameters
# of function build_matrix
# delta is the perturbation parameter (set to 1e-4 by default)
```

```
sens_elas_num <- function(pos,Sc1,Sj1,Sja1,Sa1,f1,Sc2,Sj2,Sja2,Sa2,f2,Sch,Sjh,Sjah,Sah,fh,delta=1e-4){

  # param character format
  param_char <- c('Sc1','Sj1','Sja1','Sa1','f1','Sc2','Sj2','Sja2','Sa2','f2','Sch','Sjh','Sjah','Sah','fh')

  # get parameters
  theta <- c(Sc1,Sj1,Sja1,Sa1,f1,Sc2,Sj2,Sja2,Sa2,f2,Sch,Sjh,Sjah,Sah,fh)

  # build  matrix
  A <- build_matrix_dolphin(theta)

  # calculate growth rate
  lambda  <- max(Re(eigen(A)$values))

  # get focal parameter
  c <- theta[pos]

  # modify  the focal parameter c by a very small amount
  c_new <- c * (1 + delta)
  theta_new <- theta
  theta_new[pos] <- c_new

  # build  A_new with perturbed  focal parameter
  A_new <- build_matrix_dolphin(theta_new)

  # calculate growth rate
  lambda_new <- max(Re(eigen(A_new)$values))

  # calculate sensitivity  [sens = df(x)/dx = (lam.new-lam)/(c*delta)]
  sens <- (lambda_new-lambda)  / ( c * delta)
```

```
  # calculate elasticity [elas = sens*c/lam = (lam.new-lam)/(lam*delta)]
  elas <- (lambda_new - lambda)/(lambda*delta)

  res <- list(param = param_char[pos], sens = sens, elas = elas)
  return(res)
}

# compute elasticity growth rate wrt Sc1 (pos=1 dans theta) for param in input_param_dolphin.r
sens_elas_num(1,Sc1,Sj1,Sja1,Sa1,f1,Sc2,Sj2,Sja2,Sa2,f2,Sch,Sjh,Sjah,Sah,fh)  # Sc1

# compute elasticity growth rate wrt Sc2 (pos=6 dans theta) for param in input_param_dolphin.r
sens_elas_num(6,Sc1,Sj1,Sja1,Sa1,f1,Sc2,Sj2,Sja2,Sa2,f2,Sch,Sjh,Sjah,Sah,fh)  # Sc2

# compute them all
par_char = NULL
par_sens = NULL
par_elas = NULL

for (i in 1:length(theta)){
  res = sens_elas_num(i,Sc1,Sj1,Sja1,Sa1,f1,Sc2,Sj2,Sja2,Sa2,f2,Sch,Sjh,Sjah,Sah,fh)
  par_char = c(par_char,res$param)
  par_sens = c(par_sens,res$sens)
  par_elas = c(par_elas,res$elas)
}

# display results
data.frame(param = par_char, sensitivity = par_sens, elasticitiy = par_elas)
```

The script containing the build_matrix_dolphin function (build matrix dolphins.R) is:

```
##########################################
#                BUILD MATRIX DOLPHINS            #
##########################################
```

```
build_matrix_dolphin <- function(theta){

Sc1 <- theta[1] # Calf survival (Taylor et al., 2007)
Sj1 <- theta[2] # Juvenile survival (Santostasi et al., 2016)
Sja1 <- theta[3] # Transition from juvenile to adult = Annual juvenile survival^number
of years to maturity
Sa1 <- theta[4] # Adult survival (Santostasi et al., 2016)
f1 <- theta[5] # Fertility (Taylor et al., 2007)

Sc2 <- theta[6] # Calf survival (Taylor et al., 2007)
Sj2 <- theta[7] # Juvenile survival (Santostasi et al., 2016)
Sja2 <- theta[8] # Transition from juvenile to adult = Annual juvenile survival^number of years to
maturity
Sa2 <- theta[9] # Adult survival (Santostasi et al., 2016)
f2 <- theta[10] # Fertility (average from Table 5 of Murphy et al., 2007)

Sch <- theta[11] # Calf survival (same as parentals)
Sjh <- theta[12] # Juvenile survival (same as parentals)
Sjah <- theta[13] # Transition from juvenile to adult = Annual juvenile survival^number of years to
maturity
Sah <- theta[14] # Adult survival (same as parental)
fh <- theta[15] # Fertility (average between parentals)

# init abundance
nf <-
c(1331*0.118,1331*0.118,1331*0.265,22*0.130,22*0.130,22*0.240,55*0.124,55*0.124,55*0.253) #
females
nm <-
c(1331*0.118,1331*0.118,1331*0.265,22*0.130,22*0.130,22*0.240,55*0.124,55*0.124,55*0.253) #
males
```

```
tot <- nm[3] + nm[6] + nm[9] # number of total mature males
fr1 <- nm[3]/tot # Frequency of male S1
fr2 <- nm[6]/tot # frequency of male S2
frh <- nm[9]/tot # frequency of male H

A <- matrix(c(0,0,f1*Sa1*fr1*0.5,0,0,0,0,0,0,
Sj1 , (1-Sja1)*Sj1 , 0 , 0 , 0 , 0 , 0 , 0 , 0,
0 , Sja1 , Sa1 , 0 , 0 , 0 , 0 , 0 , 0,
0 , 0 , 0 , 0 , 0 , f2*Sa2*fr2*0.5 , 0 , 0 , 0,
0 , 0 , 0 , Sj2 , (1-Sja2)*Sj2 , 0 , 0 , 0 , 0,
0 , 0 , 0 , 0 , Sja2 , Sa2 , 0 , 0 , 0,
0 , 0 , Sah*f1*(fr2+frh)*0.5 , 0 , 0 , Sah*f2*(fr1+frh)*0.5 , 0 , 0 , Sah*fh*(fr1+fr2+frh)*0.5,
0 , 0 , 0 , 0 , 0 , 0 , Sjh , (1-Sjah)*Sjh , 0,
0 , 0 , 0 , 0 , 0 , 0 , 0 , Sjah , Sah),nrow = 9, byrow = TRUE)

  return(A)

}
```

The script containing the parameters (input_param_dolphin.r) is:

```
##############################################
#                INPUT PARAM DOLPHINS                #
##############################################

Sc1 <- 0.80 # Calf survival (Taylor et al., 2007)
Sj1 <- 0.94 # Juvenile survival (Santostasi et al., 2016)
Sja1 <- 0.94^10 # Transition from juvenile to adult = Annual juvenile survival^number of years to
maturity
Sa1<- 0.94 # Adult survival (Santostasi et al., 2016)
f1 <- 0.25 # Fertility (Taylor et al., 2007)

Sc2<- 0.80 # Calf survival (Taylor et al., 2007)
```

```
Sj2<- 0.94 # Juvenile survival (Santostasi et al., 2016)
Sja2 <- 0.94^8 # Transition from juvenile to adult = Annual juvenile survival^number of years to maturity
Sa2<- 0.94 # Adult survival (Santostasi et al., 2016)
f2 <- 0.26 # Fertility (average from Table 5 of Murphy et al., 2007)

uk <- 0.28 #United kingdom
fr <- 0.29 #France
ir <- 0.20 #Ireland
ib <- 0.19 #Iberian peninsula
ne <- 0.26 #Noth-East Atlantic
cont <- 0.33 #Control group, healthy common dolphins according to Murphy et al., 2007)

fert <- c(uk,fr,ir,ib,ne,cont)

Sch <- 0.80 # Calf survival (same as parentals)
Sjh <- 0.94 # Juvenile survival (same as parentals)
Sjah <- 0.94^9 # Transition from juvenile to adult = Annual juvenile survival^number of years to maturity
Sah <- 0.94 # Adult survival (same as parental)
fh <- 0.26 # Fertility (average between parentals)
```

CITED LITERATURE

Caswell, H., 1978. A general formula for the sensitivity of population growth rate to changes in life history parameters. Theor. Popul. Biol. 14, 215–230.

Gosselin, F., Lebreton, J-D., 2000. Potential of branching processes as a modeling tool for conservation biology. In: Quantitative Methods for Conservation Biology. Springer, New York, NY.

Murphy, S., Collet, A., Rogan, E., 2005. Mating strategy in the male common dolphin (*Delphinus delphis*): what gonadal analysis tells us. J. Mammal. 86, 1247–1258.

Murphy, S., Winship, A., Dabin, W., Jepson, P., Deavill, R., Reid, J.R., Spurrie, C., et al., 2009. Importance of biological parameters in assessing the status of *Delphinus delphis*. Mar. Ecol-Prog.Ser. 388, 273–291.

Santostasi, N.L., Bonizzoni, S., Bearzi, G., Eddy, L., Gimenez, O., 2016. A robust design capture-recapture analysis of abundance, survival and temporary emigration of three odontocete species in the Gulf of Corinth, Greece. PLoS ONE.

Taylor, B.L., Chivers, S.J., Larese, J., Perrin, W.F., 2007. Generation length and percent mature estimates for IUCN assessments of cetaceans. Administrative Report LJ-07-01 National Marine Fisheries Service, Southwest Fisheries Science Center, La Jolla, CA, United States of America, pp. 808 pp. 1–24. 809
810

APPENDIX 2. Wolves and dog hybridization

1. Initial values and calculations

Wolves

Females

- Initial number of female breeders = $A_{wf}(1) = 6$ = number of packs in the Italian Alps in year 1999 (Marucco and McIntire 2010)
- Initial number of female subordinates = $S_{wf}(1) = 6$
- Initial number of females in dispersal = $D_{wf}(1) = 6$
- Initial number of female juveniles = $J_{wf}(1)$ = [average litter size (3.389, Marucco and McIntire 2010) * number of packs * 0.5] = 10

Males

- Initial number of male breeders = $A_{wm}(1) = A_{wf}(1)$
- Initial number of male subordinates = $S_{wm}(1) = S_{wf}(1)$
- Initial number of male dispersers = $D_{wf}(1)$
- Initial number of male dispersers = $D_{wm}(1) = D_{wf}(1)$
- Initial number of male juveniles = $J_{wm}(1) = J_{wm}(1)$

Admixed individuals

Females

- Initial number of female breeders = $A_{hf}(1) = 0$
- Initial number of female subordinates = $S_{hf}(1) = 0$
- Initial number of females in dispersal = $D_{hf}(1) = 0$
- Initial number of female juveniles = $J_{hf}(1) = 0$

Males

- Initial number of male breeders = $A_{hm}(1) = 0$
- Initial number of male subordinates = $S_{hm}(1) = 0$
- Initial number of males in dispersal = $D_{hm}(1) = 0$
- Initial number of male juveniles = $J_{hfm}(1) = 0$

Dogs

- Hybridization with dogs is simulated as a constant annual frequency of reproductive events between female wolves and male dogs (parameter *frd*). The increasing values of *frd* tested are 0.1, 0.2, 0.3, 0.4, 0.5.

Initial proportion of males

- $Tot_m(1) = D_{wm}(1) + D_{hm}(1) = 6$ = initial total number of male wolves and admixed individuals in dispersal
    - $\gamma(1) = D_{hm}(1) / Tot_m(1)$ - *frd*(1) = 0 = initial proportion of admixed individuals in dispersal
    - $\alpha(1) = (1$-*frd*$(1) - \gamma(1)$ = initial proportion of wolves in dispersal, depends on the *frd*(1) parameter

2. Model equations

Female wolves

$J_{wf}(t+1) = A_{wf}(t)Sa_w f_w 0.5$ *juvenile wolf females*

$S_{wf}(t+1) = J_{wf}(t)Sj_w(1 - Pdi_w)$ *subordinates wolf females*

$D_{wf}(t+1) = J_{wf}(t)Sj_w Pdi_w + S_{wf}(t)Ss_w$ *dispersers wolf females*

$A_{wf}(t+1) = D_{wf}(t)Sd_w Pes_w \alpha(t) + A_{hf}(t)Sa_w$ *female wolf breeders*

Male wolves

$J_{wm}(t+1) = A_{wf}(t)Sa_w f_w 0.5$ *juvenile wolf males*

$S_{wm}(t+1) = J_{wm}(t)Sj_w(1 - Pdi_w)$ *subordinates wolf males*

$D_{wm}(t+1) = J_{wm}(t)Sj_w Pdi_w + S_{wm}(t)Ss_w$ *dispersers wolf males*

$A_{wm}(t+1) = D_{wm}(t)Sd_w Pes_w \alpha(t) + A_{hm}(t)Sa_w$ *male wolf breeders*

Female admixed

$J_{hf}(t+1) = A_{hf}(t)Sa_h f_h 0.5$ *juvenile admixed females*

$S_{hf}(t+1) = J_{hf}(t)Sj_h(1-Pdi_h)$ *subordinate admixed females*

$D_{hf}(t+1) = J_{hf}(t)Sj_{hf} Pdi_h + S_{hf}(t+1)Ss_h$ *dispersers admixed females*

$A_{hf}(t+1) = D_{wf}(t)Sd_w Pes_h$*frd*$(t)$ + *female wolf x male dog*

$D_{wf}(t)Sd_w Pes_h \gamma(t)$ + *female wolf x male admixed*

$D_{hf}(t)Sd_h Pes_h \alpha(t)$ + *female admixed x male wolf*

$D_{hf}(t)Sd_hPes_h\gamma(t)$ + *female admixed x male admixed*

$A_{hf}(t)Sa_h$)

Male admixed

$J_{hm}(t+1) = A_{hf}(t)Sa_hf_h0.5$ *juvenile admixed males*

$S_{hm}(t+1) = J_{hm}(t)Sj_h(1-pd_h)$ *subordinate admixed males*

$D_{hm}(t+1) = J_{1hm}(t)Sj_hPdi_h + S_{hm}(t+1)Ss_h$ *dispersers admixed males*

$A_{hf}(t+1) = D_{wm}(t)Sd_wPes_h$*frd*$(t)$+ *female wolf x male dog*

$D_{wm}(t)Sd_wPes_h\gamma(t)$ + *female wolf x male admixed*

$D_{hm}(t)Sd_hPes_h\alpha(t)$ + *female admixed x male wolf*

$D_{hm}(t)Sd_hPes_h\gamma(t))$+ *female admixed x male admixed*

$A_{hm}(t)Sa_h$)

Proportion of males of each taxon at time t:

$Tot_m(t) = D_{wm}(t) + D_{hm}(t)$ *total of males in dispersal*

$\gamma(t) = D_{hm}(t) / Tot_m(t)$ - *frd*(t) *proportion of admixed individuals in dispersal*

$\alpha(t) = 1$-*frd*$(t) - \gamma(t)$ *proportion of wolves in dispersal*

The code to run the model in a matrix formulation is available below:

```
##########################################
#         WOLF CASE: MATRIX FORMULATION        #
##########################################

### Parameters
#Wolves
tspan=50
Sjw=0.55   # Marucco and McIntire 2010
Ssw=0.82  # Marucco & McIntire  2010
Sdw=0.65  # Marucco & McIntire  2010
Saw=0.82  # Marucco & McIntire  2010
fw=3.39   #average litter size Marucco and McIntire 2010
pdiw=0.25  # average dispersal rate 0.10 and 0.40 (Mech and Boitani, 2003)
```

```
pesw=0.50 # average annual p establish pack (0.3-0.7) Marescot et al., 2012.

#Admixed
Sjh=Sjw  # same as wolf
Ssh=Ssw # same as wolf
Sdh=Sdw  # same as wolf
Sah=Saw  # same as wolf
fh=fw # same as wolf
pdih=pdiw  # same as wolf
pesh=pesw # same as wolf

# Build some matrices for storing output
nf <- matrix(0,8,tspan)  # storage of age-specific female abundance
nm <- matrix(0,8,tspan)  # storage of age-specific male abundance
Nf <- matrix(0,tspan,1)  # storage of total female abundance
Nm <- matrix(0,tspan,1)  # storage of total male abundance

tot <- matrix(0, tspan,1)
frd <- matrix(0.1, tspan,1) # Constant annual rate of formation of wolf-dog mixed pairs
frw <- matrix(0,tspan,1)
frh <- matrix(0, tspan,1)
prevalence <- matrix(0, tspan,1)
grw <- matrix(0,tspan-1,1)  # storage of time-specific population growth rates
grh <- matrix(0,tspan-1,1)  # storage of time-specific population growth rates

# Initial abundances and relative abundances of males

nf[,1] <- c(10,6,6,6,0,0,0,0)  # initial population abundance females in each age class
nm[,1] <- c(10,6,6,6,0,0,0,0)  # initial population abundance females in each age class

tot[1,] <- nm[2,1]+nm[6,1]
frh[1,] <- ((nm[6,1]/tot[1,]-frd[1,]))  #Frequency of hybrid males in dispersal
```

```
if (nm[6,1]/tot[1,]-frd[1,]<0)  {
  frh[1,]=0
}

frw[1] <- (1-frd[1,]-frh[1,])  #Frequency of male wolves in disperal
if (1-frd[1,]-frh[1,]<0)  {
  frw[1,]=0
}

prevalence[1]=sum(nf[5:8,1])/(sum(nf[5:8,1])+sum(nf[1:4,1]))

# Create the pre matrix population  model

for (t in 1:(tspan-1))  {
  tot[t,] <- nm[2,t]+nm[6,t]
  frh[t,] <- ((nm[6,t]/tot[t,]-frd[t,]))  #Frequency of hybrid males in dispersal
  if (nm[6,t]/tot[t,]-frd[t,]<0)  {
    frh[t,]=0
  }

  frw[t] <- (1-frd[t,]-frh[t,])  #Frequency of male wolves in disperal
  if (1-frd[t,]-frh[t,]<0)  {
    frw[t,]=0
  }

    B =  matrix(c( 0 , 0 , 0 , fw*Saw*0.5 , 0 , 0 , 0 , 0 ,
                Sjw*pdiw  , 0 , Ssw , 0 , 0 , 0 , 0 , 0  ,
                Sjw*(1-pdiw)  , 0 , 0 , 0 , 0 , 0 , 0 , 0 ,
                0 , Sdw*pesw*frw[t,] , 0 , Saw , 0 , 0 , 0 , 0 ,
                0 , 0 , 0 , 0 , 0 , 0 , 0 , fh*Sah*0.5  ,
                0 , 0 , 0 , 0 , Sjh*pdih  , 0 , Ssh , 0  ,
                0 , 0 , 0 , 0 , Sjh*(1-pdih)  , 0 , 0 , 0 ,
```

```
        0 , Sdw*pesh*(frd[t,]+frh[t,])  , 0 , 0 , 0 , Sdh*pesh*(frw[t,]+frh[t,])  , 0 , Sah  ),nrow = 8,
byrow = TRUE)

  nf[,t+1] <- B%*%nf[,t]      # %*% = matrix multiplication in R
  nm[,t+1] <- B%*%nm[,t]
  prevalence[t+1,]=sum(nf[5:8,t+1])/(sum(nf[5:8,t+1])+sum(nf[1:4,t+1]))
  Nf[t+1] <- sum(nf[,t+1])
  Nm[t+1] <- sum(nm[,t+1])
  grw[t] <- sum(nf[1:4,t+1])/sum(nf[1:4,t])   # per time step population growth rate wolves
  grh[t] <- sum(nf[5:8,t+1])/sum(nf[5:8,t])   # per time step population growth rate admixed
}
### Prevalence
plot_freq=plot(1:(tspan), prevalence, main="Reproductive isolation 2",
        ylab="",xlab="",type = "l",lty=1,lwd = 2,xlim=c(0,tspan))
```

3. Logistic regression sensitivity analysis

The code to reproduce the logistic regression sensitivity analysis presented in the main text is provided below.

```
  ######################################################################
  #      WOLF CASE CASE: LOGISTIC REGRESSION SENSITIVITY ANALYSIS  #
  ######################################################################

   library(MASS)
   library(tidyverse)
   library(broom)

   ### Parameters

   #Wolves
     npar<-15
     sets<-15000 #parameters set for sensitivity analysis
```

```
  tspan=50
  threshold1<-3  #extinction  threshold

# Build  some arrays (nsets matrices(ageclasses x tspan))
# and matrices (tspan x nsets) for storing output  for each parm. set

  nf<-array(NA,dim=c(8,tspan,sets))      # storage of projected age-specific female abundance
  nm<-array(NA,dim=c(8,tspan,sets))      # storage of projected age-specific male abundance

  Nf<-matrix(0,tspan,sets)              # storage of projected total female abundance
  Nm<-matrix(0,tspan,sets)              # storage of projected total male abundance
  tot <- matrix(0,tspan,sets)           # storage of projected total abundance
  frw <- matrix(0,tspan,sets)           # storage of projected frequency male wolves
  frh <- matrix(0,  tspan,sets)         # storage of projected frequency admixed
  prevalence <- matrix(0,  tspan,sets)  # storage of projected prevalence
  grw  <-  matrix(0,tspan-1,sets)       # storage of time-specific  population  growth rates
  grh  <-  matrix(0,tspan-1,sets)       # storage of time-specific  population  growth rates
  frd<-matrix(NA,tspan,sets)

# Build  matrix  to store demographic  parameters and extinction
# outcome  for logistic  regression

  regr_data <- matrix(0,sets,npar+2)
  colnames(regr_data)=c("Sjw","Ssw","Sdw","Saw","fw","pdiw",
                "pesw","Sjh","Ssh","Sdh","Sah","fh","pdih","pesh","frdog","Ext_W","Ext_H")

##Loop over 15000  parameter sets drawn from uniform  distributions
  ### Draw parameters -> Run projection  over tspan years -> save parameters and extinction
outcome

  for (s in  1:(sets)) {
```

```
# Define parameters
  #Wolves
  Sjw=runif(1,min = 0.2, max = 1)   # min-max Juvenile survival (from Marescot et al., 2012)
  Ssw=runif(1,min = 0.2, max = 1)   # min-max Subordinate survival (from Marescot et al., 2012)
  Sdw=runif(1,min = 0.2, max = 1)   # min-max Dispersers survival (from Marescot et al., 2012)
  Saw=runif(1,min = 0.2, max = 1)   # min-max Alpha survival from (from Marescot et al., 2012)
  fw=runif(1,min = 0, max = 4.5)    # min-max litter size (from Mech and Boitani 2003)
  pdiw= runif(1,min = 0, max = 0.40) # min-max dispersal rate 0.10 and 0.40 (from Mech and
Boitani, 2003)
  pesw= runif(1,min = 0, max = 0.70) # min-max p. establish pack (0.3-0.7) (from Marescot et al.,
2012)

  #Admixed (same parameters ranges as wolves)
  Sjh=runif(1,min = 0.2, max = 1)
  Ssh=runif(1,min = 0.2, max = 1)
  Sdh=runif(1,min = 0.2, max = 1)
  Sah=runif(1,min = 0.2, max = 1)
  fh=runif(1,min = 0, max = 4.5)
  pdih=runif(1,min = 0, max = 0.40)
  pesh=runif(1,min = 0, max = 0.70)

  #Dogs (modelled as a constant annual rate of formation of wolf-dog mixed pairs)
  frdog=runif(1,min = 0, max = 1) # Constant annual rate of formation of wolf-dog mixed pairs
  frd[,s] <- (round(frdog,1))     # Vectorized to be used in the following projection

  # Initialize Ext_W and Ext_H objects (extinction outcomes of W and H)
  Ext_W <- 0
  Ext_H <- 0

  # Initial values (t=1)
  nf[,1,s] <- c(10,6,6,6,0,0,0,0)     # initial population abundance females in each age class
  nm[,1,s] <- c(10,6,6,6,0,0,0,0)     # initial population abundance females in each age class
```

```
tot[1,s] <- nm[2,1,s]+nm[6,1,s]       # Total males in dispersal (h+w)
frh[1,s] <- ((nm[6,1,s]/tot[1,s]-frd[1,s]))  #Frequency of hybrid males in dispersal
if (nm[6,1,s]/tot[1,s]-frd[1,s]<0)  {       #Avoid negative numbers
  frh[1,s]=0
}
frw[1,s] <- (1-frd[1,s]-frh[1,s])    #Frequency of male wolves in disperal
if (1-frd[1,s]-frh[1,s]<0)  {        #Avoid negative numbers
frw[1,s]=0
}
prevalence[1,s]=sum(nf[5:8,1,s])/(sum(nf[5:8,1,s])+sum(nf[1:4,1,s]))   #Prevalence at t=1

# Create the matrix population  model and loop over time

for (t in 1:(tspan-1)) {

  tot[t,s] <- nm[2,t,s]+nm[6,t,s]
  frh[t,s] <- ((nm[6,t,s]/tot[t,s]-frd[t,s]))  #Frequency of hybrid males in dispersal
  if (nm[6,t,s]/tot[t,s]-frd[t,s]<0)  {
    frh[t,s]=0
  }

  frw[t,s] <- (1-frd[t,s]-frh[t,s])    #Frequency of male wolves in disperal
  if (1-frd[t,s]-frh[t,s]<0)  {
    frw[t,s]=0
  }

  B =  matrix(c( 0 , 0 , 0 , fw*Saw*0.5 , 0 , 0 , 0 , 0 ,
                 Sjw*pdiw  , 0 , Ssw , 0 , 0 , 0 , 0 , 0  ,
                 Sjw*(1-pdiw)  , 0 , 0 , 0 , 0 , 0 , 0 , 0 ,
                 0 , Sdw*pesw*frw[t,s] , 0 , Saw , 0 , 0 , 0 , 0 ,
                 0 , 0 , 0 , 0 , 0 , 0 , 0 , fh*Sah*0.5 ,
                 0 , 0 , 0 , 0 , Sjh*pdih  , 0 , Ssh , 0  ,
```

```
                0 , 0 , 0 , 0 , Sjh*(1-pdih)  , 0 , 0 , 0 ,
                0 , Sdw*pesh*(frd[t,s]+frh[t,s])  , 0 , 0 , 0 , Sdh*pesh*(frw[t,s]+frh[t,s])  , 0 , Sah
),nrow = 8, byrow = TRUE)

    nf[,t+1,s] <- B%*%nf[,t,s]      # %*% = matrix multiplication in R
    nm[,t+1,s] <- B%*%nm[,t,s]
    prevalence[t+1,s]=sum(nf[5:8,t+1,s])/(sum(nf[5:8,t+1,s])+sum(nf[1:4,t+1,s]))
    Nf[t+1,s] <- sum(nf[,t+1,s])
    Nm[t+1,s] <- sum(nm[,t+1,s])
    grw[t,s] <- sum(nf[1:4,t+1,s])/sum(nf[1:4,t,s])   # per time step population growth rate wolves
    grh[t,s] <- sum(nf[5:8,t+1,s])/sum(nf[5:8,t,s])   # per time step population growth rate admixed
    }

  if(any(sum(nf[3:4,t+1,s]) < threshold1)) Ext_W=1
  if(any(sum(nf[7:8,t+1,s]) < threshold1)) Ext_H=1

regr_data[s,]=c(Sjw,Ssw,Sdw,Saw,fw,pdiw,pesw,Sjh,Ssh,Sdh,Sah,fh,pdih,pesh,frdog,Ext_W,Ext_
H)
  }
#########################################################
#       Logistic regression sensitivity analysis        #
#########################################################

###INFLUENCE OF DEMOGRAPHIC PARAMETERS ON WOLF EXTINCTION
 mylogit1 <- glm(Ext_W ~ Sjw+Ssw+Sdw+Saw+fw+pdiw+
            pesw+Sjh+Ssh+Sdh+Sah+fh+
            pdih+pesh+frdog,  data = as.data.frame(regr_data), family = "binomial")

 modsel1=stepAIC(mylogit1)  #backward stepwise selection function
 summary(modsel1)
 summary(modsel1)$coefficients  #Coefficients table of selected model
 Ext_W_coeff <- summary(modsel1)$coefficients  #Standardized coefficients
```

```
Ext_W_coeff_ord= Ext_W_coeff[order(abs(Ext_W_coeff[,3]),decreasing=T),] #Sort the standardized coeff by abs. value
Ext_W_coeff_ord
```

4. Analytical sensitivity analysis

The finite annual rate of increase can be calculated as the dominant eigenvalue of the transition matrix **A**. The sensitivity of the growth rate to a proportionate change in each parameter (elasticity) can be obtained using calculus (based on Caswell,1978). The parameters with the higher elasticity and sensitivity were the survival of both wolves and admixed individuals (Table S2).

Table S2.

Parameter	**Sensitivity**	**Elasticitiy**
sjw	0.228	0.124
ssw	0.108	0.088
sdw	0.193	0.124
saw	**0.816**	**0.664**
fw	0.037	0.124
pdiw	0.027	0.007
pesw	0.251	0.124
sjh	0.228	0.124
ssh	0.108	0.088
sdh	0.193	0.124
sah	**0.816**	**0.664**
fh	0.037	0.124
pdih	0.027	0.007
pesh	0.251	0.124
frd	-0.139	-0.014

The code to reproduce the sensitivity analysis is available below:

```
##########################################
#                SENSITIVITY ANALYSIS                #
##########################################

# Create the wolf population model function
# with generic parameters
source('build_matrix_wolf.r')

# assign values to model param
source('input_param_wolf.r')

# vector of parameters
theta <- c(sjw,ssw,sdw,saw,fw,pdiw,pesw,sjh,ssh,sdh,sah,fh,pdih,pesh,frd)
theta

# Apply function build_matrix to obtain the dolphin matrix
A <- build_matrix_wolf(theta)
A
# compute asymptotic growth rate
max(Re(eigen(A)$values))
# Create function to calculate sensitivity wrt parameter in position pos
# in vector of parameters that follows and matches the input parameters
# of function build_matrix
# delta is the perturbation parameter (set to 1e-4 by default)
sens_elas_num <-
function(pos,sjw,ssw,sdw,saw,fw,pdiw,pesw,sjh,ssh,sdh,sah,fh,pdih,pesh,frd,delta=1e-4){

        # param char format
        param_char <- c('sjw','ssw','sdw','saw','fw','pdiw','pesw','sjh','ssh','sdh','sah','fh','pdih','pesh','frd')

        # get parameters
```

```
    theta <- c(sjw,ssw,sdw,saw,fw,pdiw,pesw,sjh,ssh,sdh,sah,fh,pdih,pesh,frd)

    # build  matrix
    A <- build_matrix_wolf(theta)

    # calculate growth rate
    lambda <- max(Re(eigen(A)$values))

    # get focal parameter
    c <- theta[pos]

    # modify  the focal parameter c by a very small amount
    c_new <- c * (1 + delta)
    theta_new <- theta
    theta_new[pos] <- c_new

    # build  A_new with perturbed  focal parameter
    A_new <- build_matrix_wolf(theta_new)

    # calculate growth rate
    lambda_new <- max(Re(eigen(A_new)$values))

    # calculate sensitivity  [sens = df(x)/dx = (lam.new-lam)/(c*delta)]
    sens <- (lambda_new-lambda)  / ( c * delta)

    # calculate elasticity  [elas = sens*c/lam = (lam.new-lam)/(lam*delta)]
    elas <- (lambda_new  - lambda)/(lambda*delta)

    res <- list(param  = param_char[pos], sens = sens, elas = elas)
    return(res)
}
```

```
# compute elasticity growth rate wrt sjw (pos=1 dans theta) for param in input_param_wolf.r
sens_elas_num(1,sjw,ssw,sdw,saw,fw,pdiw,pesw,sjh,ssh,sdh,sah,fh,pdih,pesh,frd)

# compute elasticity growth rate wrt pdiw (pos=6 dans theta) for param in input_param_wolf.r
sens_elas_num(6,sjw,ssw,sdw,saw,fw,pdiw,pesw,sjh,ssh,sdh,sah,fh,pdih,pesh,frd)

# compute them all
par_char = NULL
par_sens = NULL
par_elas  = NULL
for (i in 1:length(theta)){
        res = sens_elas_num(i,sjw,ssw,sdw,saw,fw,pdiw,pesw,sjh,ssh,sdh,sah,fh,pdih,pesh,frd)
        par_char = c(par_char,res$param)
        par_sens = c(par_sens,res$sens)
        par_elas = c(par_elas,res$elas)
}
# display results
data.frame(param = par_char, sensitivity  = par_sens, elasticitiy  = par_elas)
```

The script containing the build_matrix_wolf function (build matrix wolf.r) is:

```
##########################################
#                BUILD MATRIX  WOLF                #
##########################################

build_matrix_wolf  <- function(theta){

sjw <- theta[1] # Marucco & McIntire  2010
ssw <- theta[2] # Marucco & McIntire  2010 assuming that alpha and subordinates have same survival
sdw <- theta[3] # Marucco & McIntire  2010
saw <- theta[4] # Marucco & McIntire  2010
fw <- theta[5] #average litter size 4 to 9 pups (Mech and Boitani, 2003)
```

```
pdiw <- theta[6] # average dispersal rate 0.10 and 0.40 (Mech and Boitani, 2003)
pesw <- theta[7] # average p annual p establish pack (0.3-0.7) Marescot et al., 2012.
sjh <- theta[8] # Marucco & McIntire 2010
ssh <- theta[9] # Marucco & McIntire 2010
sdh <- theta[10] # Marucco & McIntire 2010
sah <- theta[11] # Marucco & McIntire 2010
fh <- theta[12] #average litter size 4 to 9 pups (Mech and Boitani, 2003)
pdih <- theta[13] # average dispersal rate 0.10 and 0.40 (Mech and Boitani, 2003)
pesh <- theta[14] # average p annual p establish pack (0.3-0.7) Marescot et al., 2012.
frd <- theta[15]

# init abundance
nf <- c(10,6,6,6,0,0,0,0) # females
nm <- c(10,6,6,6,0,0,0,0) # males

tot <- nm[2] + nm[6] # number of total mature males
frh <- ifelse((nm[6]/tot - frd) > 0, nm[6]/tot - frd, 0)
frw <- ifelse((1 - frd - frh) > 0, 1 - frd - frh, 0)

A <- matrix(c( 0 , 0 , 0 , fw*saw*0.5 , 0 , 0 , 0 , 0 ,
sjw*pdiw , 0 , ssw , 0 , 0 , 0 , 0 , 0 ,
sjw*(1-pdiw) , 0 , 0 , 0 , 0 , 0 , 0 , 0 ,
0 , sdw*pesw*frw, 0 , saw , 0 , 0 , 0 , 0 ,
0 , 0 , 0 , 0 , 0 , 0 , 0 , fh*sah*0.5 ,
0 , 0 , 0 , 0 , sjh*pdih , 0 , ssh , 0 ,
0 , 0 , 0 , 0 , sjh*(1-pdih) , 0 , 0 , 0 ,
0 , sdw*pesh*(frd+frh) , 0 , 0 , 0 , sdh*pesh*(frw+frh), 0 ,sah),nrow = 8, byrow = TRUE)
return(A)
  }
```

The script containing the parameters (input_param_wolf.r) is:

```
##########################################
#              PUT PARAMETERS WOLF       #
##########################################
sjw <- 0.55 # Marucco & McIntire 2010
ssw <- 0.82 # Marucco & McIntire 2010 assuming that alpha and subordinates have same survival
sdw <- 0.65 # Marucco & McIntire 2010
saw <- 0.82 # Marucco & McIntire 2010
fw <- 3.39 #average litter size Marucco & McIntire 2010
pdiw <- 0.25 # average dispersal rate 0.10 and 0.40 (Mech and Boitani, 2003)
pesw <- 0.5 # average p annual p establish pack (0.3-0.7) Marescot et al., 2012.
sjh <- 0.55 # Marucco & McIntire 2010
ssh <- 0.82 # Marucco & McIntire 2010
sdh <- 0.65 # Marucco & McIntire 2010
sah <- 0.82 # Marucco & McIntire 2010
fh <- 3.39 #average litter size Marucco & McIntire 2010
pdih <- 0.25 # average dispersal rate 0.10 and 0.40 (Mech and Boitani, 2003)
pesh <- 0.5 # average annual p establish pack (0.3-0.7) Marescot et al., 2012.
frd <- 0.1
```

LITERATURE CITED

Caswell, H., 1978. A general formula for the sensitivity of population growth rate to changes in life history parameters. Theor. Popul. Biol. 14, 215–230.

Marescot, L., Gimenez, O., Duchamp, C., Marboutinc, E., Chapron, G., 2012. Reducing matrix population models with application to social animal species. Ecol. Model. 232, 91–96.

Marucco, F., McIntire, E.J.B., 2010. Predicting spatio-temporal recolonization of large carnivore populations and livestock depredation risk: wolves in the Italian Alps. J. Appl. Ecol. 47, 789–798.

Mech, L.D., Boitani, L., 2003. Wolves: Behavior, Ecology, and Conservation. University of Chicago Press, Chicago, United States of America, pp. pp. 1–447.

CHAPTER 4

CHAPTER 4

TITLE

Individual-based modeling to assess management efficacy to mitigate wolf x dog hybridization

LIST OF AUTHORS AND AFFILIATIONS

Nina Luisa Santostasi[a,b]
Sarah Bauduin[b]
Oksana Grente[b]
Olivier Gimenez[b*]
Paolo Ciucci[a*]

*Co-last authors

a) Department of Biology and Biotechnologies, Sapienza University of Rome, Viale Dell'Università 32, 00185, Roma, Italy
b) CEFE, CNRS, Univ. Montpellier, Univ. Paul Valéry Montpellier 3, EPHE, IRD, Montpellier, France

TARGET JOURNAL

Journal of Applied Ecology

ABSTRACT

Hybridization between domesticated forms and their wild ancestors is a particular case of anthropogenic hybridization caused by the widespread occurrence of domesticated species in conjunction with increasing anthropogenic pressures (e.g., overexploitation, habitat destruction, impact from invasive species). Although hybridization between wolves and dogs occurred repeatedly during the course of dog domestication, there is concern that this phenomenon is increasing, particularly where wolves are expanding their range through human-dominated landscapes (e.g., Europe) where free-ranging dogs have become the most abundant carnivore. To contain this threat, European conservation legislation (i.e., the EU Council Directive 92/43/EEC, and Recommendation nr 173, 2014 of the Bern Convention) requires that wolf x dog hybridization be mitigated through effective management. In this context, the use of demographic simulation models can be instrumental to inform decision-making. We built a detailed individual based model (IBM) simulating the life cycle of gray wolves that incorporates aspects of wolf sociality

35 influencing hybridization. We applied this model to investigate hybridization dynamics of wolves in
36 the Northern Apennines, Italy, to evaluate the effectiveness of different management scenarios
37 aimed to reduce the abundance of admixed individuals during a ten-generation timeframe. We
38 showed that in the presence of continual immigration of admixed individuals, any management
39 action proved ineffective. In the presence of immigration by pure wolves, all management actions
40 produced a decrease in the prevalence of admixture, although their relative efficacy varied relative
41 to different mate choice scenarios. In all simulations, the impact of hybridization was predicted to
42 extend across broad spatial scales, driven by the dispersal of large numbers of admixed animals.
43 Moreover, we identified demographic and social processes that need to be investigated further to
44 more accurately predict the outcomes of alternative management strategies.
45
46 **1. INTRODUCTION**
47 Hybridization is defined as the intermixing of two distinct related taxa (Gompert & Buerkle 2016).
48 Although natural hybridization is acknowledged to be a positive force capable of introducing
49 beneficial adaptive genetic variation (Abbott et al. 2013), anthropogenic hybridization is considered
50 to be potential threat to the persistence of populations and species (Rhymer & Simberloff, 1996;
51 Todesco et al. 2016).
52 Hybridization between domesticated forms and their wild ancestors is a specific case of
53 anthropogenic hybridization that can be exacerbated by human pressures (e.g., overexploitation,
54 habitat destruction) and, in particular, the widespread occurrence of domesticated species (Boivin et
55 al. 2016). Multiple cases of hybridization between wild and domestic species have been identified
56 as threats for wildlife populations around the world. These include hybridization between wildcat
57 (*Felis silvestris*) and domestic cat (*Felis catus*) in Europe (Beaumont et al. 2001), American bison
58 (*Bison bison*) and cattle (*Bos taurus*) in North America (Ward et al. 1999), dingo (*Canis lupus*
59 *dingo*) and domestic dog (*Canis lupus familiaris*) in Australia (Stephens et al. 2015), and Ethiopian
60 wolf (*Canis simiensis*) and domestic dog in Africa (Gottelli et al. 1994). A notable example of
61 anthropogenic hybridization occurs between gray wolves (*Canis lupus*) and dogs in human-
62 dominated landscapes. As in several other cases of wild-domestic hybridization, the two *Canis*
63 subspecies are interfertile and their progeny can backcross with both parental populations, leading
64 to introgression of domesticated genes into the gene pool of wild wolf populations (Vilà & Wayne
65 1999). Although hybridization between wolves and dogs occurred repeatedly during the course of
66 dog domestication, there is concern that this phenomenon is increasing, particularly where wolves
67 are recolonizing human-dominated landscapes (e.g., Europe) and free-ranging dogs have become
68 the most abundant carnivore (Ritchie et al. 2014). Indeed, largescale studies have detected recent

hybridization events in several wolf populations in Eurasia (Dufresnes et al. 2019; Salvatori et al. Submitted) and intensive studies at finer-scales have found that that the proportion of admixed individuals can represent more than half of the total population of hybridizing canids (Salvatori et al. 2019; Santostasi et al. Submitted). These results raise concern for the potentially deleterious consequences of introgression of domesticated genes in the wild wolf gene pool. To contain this threat, European conservation legislation (i.e., the EU Council Directive 92/43/EEC, and Recommendation nr 173, 2014 of the Bern Convention) mandates that wolf × dog hybridization should be mitigated through effective management. However, there is little evidence that management has been effective, especially in the countries where hybridization is becoming increasingly prevalent, such as Italy (Salvatori and Ciucci, 2018; Salvatori et al. Submitted).

The management of anthropogenic hybridization involves three types of actions (Donfrancesco et al. 2019): preventive (e.g., community engagement and education to decrease the number of free-ranging dogs or to prevent the illegal killing of wolves), proactive (removal of free-ranging dogs, and reduction of factors facilitating hybridization such as poaching), and reactive (e.g., the active management of admixed individuals to reduce their prevalence and breeding in the parental wolf population through three alternative approaches: 1) sterilization and release, 2) removal to captivity, or 3) lethal removal). Decisions about which approach is appropriate to manage wolf x dog hybridization are controversial, and differ both with regard to acceptance within the scientific community (Donfrancesco et al. 2019), and with respect to applicability to the legal mandate of different European countries (Trouwborst 2014). Preventive and proactive measures can generally be enforced through national and European legislation, are agreed upon and supported by experts, and are often socially acceptable. However, reactive measures are controversial due to the undefined legal status of hybrids and the lack of consensus within both the scientific community (Donfrancesco et al. 2019) and the general public. The practical and economic feasibility of reactive management strategies has also been questioned (Lorenzini et al. 2014, Pacheco et al. 2017). Nevertheless, when introgression is widespread, reactive interventions might be the only tool to reduce introgression (Salvatori et al. 2019, Santostasi et al. Submitted) and have been used to mitigate other instances of hybridization (e.g., red wolf, *Canis rufus* and coyote *Canis latrans* in North America, Stoskopf et al. 2005; bontebock *Damaliscus pygargus pygargus* and blesbok *D. p. phillipsi* in South Africa; vanWyk et al. 2016; wildcat and domestic cat in Scotland, Senn et al. 2019).

It is difficult to assess the effectiveness of reactive management strategies within an experimental framework, as long time series on hybridization dynamics and responses to management are generally unavailable. Therefore, management decisions about hybridization are

often difficult to evaluate rigorously, which can lead to confusion and debate within the scientific community regarding strategies for effective management (Donfrancesco et al. 2019). In this context, demographic models are useful to simulate complex ecological and evolutionary processes that are not easily empirically measured (Nathan et al. 2019) and such modeling results can be instrumental to inform decision-making amidst uncertainty (e.g., Gervasi & Ciucci 2018). One notable example is red wolf × coyote hybridization in North Carolina, USA, where the efficacy of sterilizing admixed breeders was first evaluated with a simulation model (Fredrickson & Hedrick 2006) and later supported by a long-term observational study (Gese & Terletzky 2016).

Two approaches have been used to model hybridization dynamics (Hall & Ayres 2008). Population-based models require only that the mean demographic parameters of the main demographic classes are available (e.g., Wolf et al. 2001), whereas individual-based models (IBMs) simulate the contribution of each individual to the hybridization dynamics of the entire population (e.g., Thompson et al. 2003; Hooftman et al. 2007). However, these two approaches have only rarely been used to model anthropogenic hybridization. Specifically, we recently applied stage-structured matrix models to project hybridization dynamics between wolf × dog hybridization dynamics under different reproductive isolation scenarios (Santostasi et al. in Press). Earlier, Fredrickson & Hedrick (2006) applied individual-based projections to predict red wolf×coyote hybridization under different mate choice and management scenarios. A population-based approach is sufficient to make predictions at the population level, but it does not account for the inherent complexity of wolf social structure and social dynamics. IBMs are more suitable to include detailed social structure components because they are bottom-up models simulating the fate of individuals according to behavioral rules depending on the individuals' characteristics. Results at the population scale emerge from the individual-level simulations. Due to their flexibility, IBMs have been increasingly used for the projection of population dynamics in complex systems such as social species (Hrasdky et al. 2019) or hybridization (Nathan et al. 2019). IBMs including some aspects of gray wolves' social ecology (e.g., age dependent dispersal, reproduction limited to one breeding pair per social group) have been published (e.g., Marucco & McIntire 2010, Chapron et al. 2016). However, some components of wolf social ecology that are linked to hybridization rates in canids were omitted.

All species in the Canidae family are cooperative breeders and the perturbation of social dynamics is hypothesized to affect hybridization rates (Bohling & Waits 2015). In detail, dynamics such as pack disruption after mortality of breeding wolves (Bohling & Waits 2015), asymmetric replacement of breeders (the fact that vacant female breeder position are prioritarily filled by subordinates while the vacant male breeder positions are prioritarily filled by dispersers; vonHoldt et al. 2008), mate choice (Hinton et al. 2018), or inbreeding avoidance (vonHoldt et al. 2008) have

been hypothesized to affect hybridization rates in canid species. For this reason, the impact of different management strategies on wolf behavior and social dynamics should be considered when evaluating hybridization dynamics and management options.

We built an IBM for gray wolves that accounts for important aspects of wolf social dynamics that may influence hybridization. To present a flexible individual modelling framework to project hybridization dynamics and to explore how different management strategies interact with social dynamics, we applied this model to a local wolf population such as the one in the Northern Apennines, Italy, in which we recently assessed prevalence of admixture (Santostasi et al. Submitted).

We chose this population because we locally detected a high and unexpected extent of introgression following years of wolf conservation efforts in the Northern Apennines that has not included any attempt to mitigate hybridization. Therefore, the observed extent of admixture in this population represents an empirical outcome of what may occur locally in other wolf populations without management, which can be contrasted with our model predictions under alternative management scenarios.

We tested he following hypotheses: 1) management strategies aimed at removing admixed individuals (i.e., culling and captivation) vs. management strategies aimed at neutralize reproduction in admixed individuals (i.e., sterilization) have different outcomes depending on the underlying natural reproductive barriers (i.e., assortative vs. disassortative mating); 2) management strategies that target reproductive individuals only are more effective than strategies that target all age classes.

2. METHODS

2.1 General description and sub-model structure

The model was an individual-based model programmed in R using package NetLogoR (Bauduin et al. 2019), in which wolves were agents whose behavior was determined by their individual characteristics: age, sex, social status, wolf genomic content, pedigree, breeding experience (Table 1). Specifically, wolf genomic content was an index reporting the percentage of wolf genome in each individual and it was calculated as the sum of half the wolf genomic content of the parents (Nathan et al. 2019; Table 1). The model was not spatially explicit and individuals were not assigned to a specific location. However, we introduced a carrying capacity parameter (the maximum estimated number of wolf packs that can be sustained in the population). The time step was one year. Each year, all individuals went through the same processes, organized in sub-models, each one affecting every individual differently depending on their characteristics. The first 6 sub-models simulated the life cycle of wolves: 1) reproduction, 2) aging, 3) mortality, 4) change of social status, 5) immigration, 6) permanent emigration, and the last one (7) simulates management (Fig. 1). In turn, the change of status sub-model contained several additional sub-models: 4*i*) pack dissolution , 4*ii*) replacement of breeding females by subordinates, 4*iii*) dispersal, 4*iiii*) establishment of non.residents (by adoption, by replacing a missing breeder, or by occupying new territories, Fig. 1). The management sub-model includes the following sub-models (Fig. 1): 7*i*) removal and 7*ii*) sterilization. Below, we provide the detailed descriptions of the sub-models, and discuss how density dependence, inbreeding avoidance, and hybridization were included in the model. The rationale behind parameters choice has been described in Appendix 1. The chosen parameters are detailed in the Parametrization section.

2.2 Sub-models description

2.2.1 Reproduction

Each year, any wolf pack (in this model corresponding to at least one territorial breeder) with a breeding pair produced a litter. Only one breeding pair in a pack reproduced and no other adult members were allowed to reproduce (Packard 2003). We simulated the number of pups produced in each litter by randomly drawing it from a Poisson distribution with rate equal to the average litter size at one month after birth (Table 2). We did not model reproductive senescence.

2.2.2 Aging

In this sub-model, the age of the individuals was updated. Therefore, pups born in the current time step were 1 year old (i.e., in their first year), yearlings born in the previous time step were 2 years

old (i.e., in their second year), and adults were 3+ years old and are considered reproductively mature. We did not model senescence or any increase of the probability of mortality with age. To represent realistic age distribution in the population, all wolves reaching 14 years old were forced to die (Marucco & McIntire, 2010).

2.2.3 Mortality

Mortality in wolves depends strongly on age (in particular pups have lower survival than adults) and social status (non-residents usually have lower survival than resident wolves, Blanco & Cortés 2007; Smith et al. 2010). In this model, non-resident wolves were broadly defined as non-territorial individuals who left their established territory (e.g., dispersers and floaters; Smith et al. 2010). Accordingly, we assigned mortality rates depending on the age class (different values for pups vs. adults and yearlings in packs) and residency status (different values for non-residents vs. residents; Table 2).

2.2.3.1 *Density dependence in adult mortality*

We modeled density dependence by considering adult survival as constant until all the available territories (i.e., a proxy of carrying capacity) were occupied and we modeled adult survival as density-dependent after carrying capacity is reached. We modeled density-dependence as a linear decline in the adult survival at the increase of pre-reproduction density (density of adults and yearlings; Cubaynes et al. 2014).

2.2.4 Change of social status

2.2.4.1 Pack dissolution

By pooling data from multiple wolf populations across Europe and North America Brainerd et al. (2008) found that mortality of breeding wolves may disrupt the pack cohesion and cause its dissolution and the consequent abandonment of territory, depending on the pack size and the number of missing breeders. We modeled this process by performing Bernoulli trials to determine the status dissolution vs. non-dissolution of packs following the death of 1 or bothbreeders. Specifically, large packs (> 4 members) did not dissolve (Brainerd et al. 2008), whereas packs comprising only pups always dissolved. Additionally, small packs (≤ 4 members) with one missing breeder had a higher probability of dissolution (Table 2) than small packs with no missing breeders (Brainerd et al 2008). When a pack dissolved, all the former pack members became non-residents.

2.2.4.2 Change of individual social status

2.2.4.2.a <u>Replacement of breeding females by subordinates</u>

225 Previous studies have reported that females within gray wolf packs are more likely to replace
226 breeders than males (Jędrzejewski et al. 2005; vonHoldt et al. 2008; Caniglia et al. 2014). Thus, in
227 the model, breeding females who died were replaced by an eligible female in the pack if one was
228 present (before the replacement by a disperser, Fig. 1). We modeled this process by randomly
229 drawing a subordinate, mature (≥3 years old) female to become the breeding female for all packs
230 within which a female breeder had died. To avoid inbreeding, (Caniglia et al. 2014; vonHoldt et al.
231 2008) if the breeding male was present and was closely related with the chosen female, he will lose
232 its status, and the pack will remain without a male breeder until the male breeder replacement sub-
233 model that occurs later during the same year (see below; See Appendix 3 for further explanation).
234 *2.2.4.2.a.1 Implementation of inbreeding avoidance*
235 Given the choice, wolves avoid mating with close relatives (vonHoldt et al. 2008). In all the sub-
236 models including the formation of a breeding pair, we modeled inbreeding avoidance by imposing a
237 relatedness threshold above which individuals would choose alternative mates. We set this
238 threshold at $r \geq 0.125$, a relatedness coefficient corresponding to first cousins (vonHoldt et al. 2008;
239 Caniglia et al. 2014). We allowed inbreeding to occur only under "extenuating circumstances" (i.e.,
240 absence of alternative mates in the breeding season; vonHoldt et al. 2008) as the last pairing option
241 (see Male breeder replacement by subordinate sub-model).
242
243 *2.2.4.2.b* Dispersal
244 Wolves routinely disperse in response to competition and aggression related to food availability and
245 breeding opportunities within their pack (Packard 2003). Wolves disperse at different ages because
246 of limited food resources (Ballard et al. 1987, Gese & Mech 1991), and often postpone dispersal in
247 areas of high prey availability (Ballard et al. 1987; Blanco and Cortés 2007; Jimenez et al. 2017).
248 However, dispersal rates increase rapidly with the onset of sexual maturity (Gese & Mech 1991;
249 Packard 2003) such that most wolves disperse from their natal pack by the age of 3 years (Gese &
250 Mech 1991, Jimenez et al. 2017). We simulated the maximum number of individuals that could be
251 sustained in a pack using a normal distribution with a mean corresponding to the mean annual pack
252 size characterizing the study area (Table 2). If a pack was composed by more wolves than the
253 maximum simulated pack size, subordinates will leave the pack until the number of wolves in the
254 pack is equal to the maximum. We simulated the dispersal due to the onset of sexual maturity by
255 giving the highest dispersal probability to adult subordinates, followed by yearlings and pups (Table
256 2; Haight & Mech 1997). Individuals selected to leave the pack became non-residents.
257
258 *2.2.4.2.c* Establishment of non-residents

2.2.4.2.c.1 Establishment by adoption in an existing pack

Non-residents sometimes join packs already containing a breeding pair and become adoptees (Ballard et al. 1987, Mech and Boitani 2003). Most adoptees are males and are 1-3 years old (Mech and Boitani 2003). We modeled this process by selecting the packs which were not at their maximum size (determined in the dispersal sub-model) and by randomly drawing from a pool of young non-residents as many potential adoptees as the number of individuals needed to reach the maximum pack size. The adoption of males was favored by giving them priority in the selection (first selecting randomly among young male non-residents and then selecting randomly among young female non-residents if needed). Then, as the interaction between a single individual and an unknown pack can be negative due to competition and territorial aggression, we applied a Bernoulli trial with probability 0.5 of success to each possible adoption. If the adoption is successful, the young individual will join the pack, otherwise it will remain in the non-residents pool. Individuals that leave a pack in the Dispersal sub-model could re-join their former pack with the adoption sub-model. In this way we accounted for individuals that leave the pack temporarily.

2.2.4.2.c.2 Establishment by replacing a missing breeder

2.2.4.2.c.2.a Replacement of breeding females

If the missing breeding female breeder has not been replaced by a subordinate mature female (due to the absence of available mature subordinates), it can be replaced by a disperser. We modeled this process by randomly drawing a female replacer from the mature non-residents pool by first excluding all the females which are closely related to the breeding male (if present). See Appendix 3 for further explanation.

2.2.4.2.c.2.b Replacement of breeding males

Previous studies on gray and eastern wolves reported male-biased dispersal to packs where dispersers filled vacant male breeding positions (Jędrzejewski et al. 2005; vonHoldt et al. 2008; Rutledge, et al. 2011). Accordingly, replacement of male breeders was simulated first with male dispersing wolves joining the pack(before the Male breeder replacement b a subordinate sub-model). We randomly drew a mature male replacer from the non-residents pool by first excluding all the males which were closely related to the breeder female (if present).

2.2.4.2.c.3 Establishment by occupying a new territory

2.2.4.2.c.3.a Density dependence

We imposed a maximum number of available territories that can be sustained by the environment (carrying capacity). To include density dependence in the occupation of new territories we:

1. Calculated the number of occupied territories
2. Calculated a probability of establishment as:

$$Probability\ of\ establishment = \frac{\text{number of available territories}}{\text{carrying capacity}}$$

This probability reaches its maximum (1) when the number of available territories is low and it reaches its minimum (0) when the number of occupied territories is equal to the carrying capacity

3. Calculated the number of mature non-residents potentially available to establish in a vacant territory
4. Selected the mature non-residents that were going to successfully establish by applying a Bernoulli trial to each mature disperser with the probability of success being the probability of establishment.

Because there were several sub-models involving the occupation of available territories (establishment by pair bonding, by budding, and alone), the probability of establishment was updated after each one of them.

2.2.4.2.c.4. Establishment by pair bonding

One of the main mechanisms of pair formation is for dispersing wolves to find each other (Mech & Boitani 2003). If there were vacant territories available, two non-residents that were not closely related can form a breeding pair and establish a new pack. We modeled this process by randomly pairing non-closely related mature males and females in the pool of the non-residents that were selected to successfully establish.

2.2.4.2.c.5. Establishment by budding

Another breeding strategy for dispersing wolves is to pair with a subordinate from an existing pack and establish a new pack (Mech & Boitani 2003; Brainerd et al. 2008). In this situation, the mate for each disperser destined to establish was randomly drawn from the pool of mature subordinates from the existing packs (excluding the closely related ones). Once the subordinate is selected we applied a Bernoulli trial with a 0.5 probability of success to simulate the possibility that the interaction between a single disperser and a member of an unknown pack could not necessarily end up with a successful budding. If the budding is successful the disperser and the subordinate become breeders of a new pack and establish in an available territory.

2.2.4.2.c.6. Establishment alone

Non-residents that were selected to successfully establish but do not find any available mate may establish alone in a vacant territory. Once they establish alone, wolves acquire resident status, obtain a new and unique pack identity, and become breeders of the future pack.

2.2.4.2.d Replacement of breeding males by subordinates

When no mature male disperser could replace the missing breeding male in a pack, one of the subordinates could take over the breeding position. If there were several mature male subordinates in the pack, one is randomly selected among the least genetically related to the current breeding female, otherwise if there is not any breeding female, the choice among the potential breeders is random. In the particular case when among these subordinates there was the male previously dismissed from its breeding position (during the sub-model Replacement of breeding females by subordinates), because he was too related to the new female breeder, this individual gets his breeding position back. Once the new breeding male was chosen, and in the case it was not a dismissed individual, the current breeding female (if there is any) could be replaced if the newly formed breeding pair is too related. If the breeding female is too related to the breeding male and there is a mature female subordinate less related to the breeding male, she became the breeding female and the current breeding female was dismissed. If there were several mature female subordinates least related to the breeding male, the choice among them is random. If the current breeding female is less or equally related to the breeding male than the mature female subordinates, she maintained her breeder status. Once new breeding individuals were chosen, they were be able to mate the next year. See Appendix 3 for further explanation.

2.2.5 Immigration

Wolves can disperse for long distances, therefore populations are rarely completely isolated (Ciucci et al. 2009). We simulated the arrival of a variable number of immigrants each year (1 to 5) into the population by first randomly selecting the number of immigrants from a uniform distribution between 1 and 5. Those individuals entered the population as non-residents and are considered to be unrelated to the individuals in the population.

2.2.6 Permanent emigration

Individuals could exit the population by dispersing at long distance and we selected a variable number (1 to 5) of individuals that permanently emigrate from the population each year. Those individuals were removed from the population.

2.3. Hybridization

Studies quantifying the factors influencing hybridization rates between gray wolves and dogs are missing. However, several studies investigated the factors affecting hybridization between two North American wolf species (red wolf *Canis rufus* and Eastern wolf *Canis lycaon*) and coyotes (*Canis latrans*). We relied on this literature to model wolf x dog hybridization mechanisms.

Wolf density is the first common element among studies of *Canis* hybridization, with low density of wolves caused by anthropogenic mortality facilitating hybridization due to the lack of conspecific mates and the disruption of social structure (Hinton et al. 2018). This has been shown for eastern wolves interbreeding with coyotes (Rutledge et al. 2011; Benson et al. 2012; 2014) and red wolves interbreeding with coyotes (Bohling & Waits, 2015; Hinton et al. 2018); it has also been hypothesized for gray wolves and dogs in Europe (Godinho et al. 2011; Salvatori et al. 2019). More precisely, Bohling & Waits (2015) showed that red wolf-coyote admixed pairs were more likely to form after the death of one breeder in presence of high availability of heterospecific mates, and that young, first-time female breeders were responsible for a significant proportion of hybridization events. These authors hypothesized that the death of red wolf breeders during the breeding season forces reproductively active red wolves to quickly locate another mate (Bohling & Waits 2015). As stable red wolf pairs dissolved, the inhibition to breed is removed for young wolves, providing them with the opportunity to reproduce (Brainerd et al. 2008). The inexperience of these individuals, coupled with the high availability of heterospecific mates, may facilitate hybridization.

2.3.1 Hybridization implementation in the IBM

Hybridization with dogs in our model could occur in two sub-models, the "pack dissolution" and the "establishment alone" sub-models (Fig. 1). After pack dissolution, young (3-4 years old) first time breeders were attributed a probability of mating with dogs that is assumed to decline as the number of wolves and hybrids increases. Also, young first-time breeder females that were selected to establish alone in the establish alone sub-model will have the same probability of mating with dogs. Following Fredrickson & Hedrick (2006), we calculated this probability (P_{WD}) as:

$$P_{WD} = P_{max} e^{-Nrw}$$

where N is the number of wolves and hybrids in the population, P_{max} is the maximum value for P_{WD} (when $N = 0$, $P_{WD} = P_{max}$), $r_w = \frac{ln(P_{min}/P_{max})}{N_{thresh}}$ is a constant affecting the rate of change in P_{WD} with N_{thresh} the threshold N for which P_{min} is reached.

To calculate the probability of a hybrid pairing with a dog (P_{HD}), we first calculated the probability of a dog pairing with a dog (P_{DD}):

$$P_{DD} = (1 - P_{max})e^{Nrd}$$

where $r_d = \frac{[(1-P_{min})/(1-P_{max})]}{N_{thresh}}$

and then the probability of a hybrid pairing with a dog (P_{HD}) was

$$P_{HC} = P_{WD} + (1 - A_H) \times (P_{DD} - P_{WD})$$

where A_H was the proportion of red wolf ancestry of the hybrid (Fredrickson & Hedrick 2006). Therefore, the increased probability of a hybrid pairing with a dog relative to that of a wolf was proportional to the ancestry difference between a pure wolf and the hybrid.

The females that will mate with dogs are selected based on Bernoulli trials with probability of mating with dog being P_{HD}. Those females would: i) not be available for any following pairing option or for adopting individuals, ii) establish alone in a vacant territory, iii) produce a hybrid litter in the "reproduction" sub-model; the wolf genomic content of the litter will be calculated as ½ of the mother's wolf genomic content.

2.3.2 Random and assortative mate choice scenarios

We simulated two mate choice scenarios: random and assortative. In the latter, in the sub-models involving the formation of breeding pairs, individuals chose the mate that is the most similar in wolf genomic content among the available mates.

2.4. Management sub-models

We simulated two main reactive management strategies: removal and sterilization. For the sterilization scenarios, we assumed that surgical sterilization (i.e., vasectomy and tubal ligation) did not affect social or territorial behavior. We also simulated two approaches for the selection of the individuals to be managed: non-targeted (individuals selected randomly) and targeted selection (aimed at selecting breeders and mature non-breeders). We selected a wolf genomic content threshold below which the individuals were subject to management (Table 2), this threshold corresponds to the wolf genomic content of a first generation backcross (75%). We chose this threshold because first generation backcrosses can be unequivocally distinguished from wolves with few molecular markers (Vähä & Primmer 2006). This threshold represent a conservative choice to eliminate the possibility of erroneously sterilizing or removing a pure wolf (i.e., admixed individuals of further generations back are increasingly difficult to detect with certainty; Santostasi

et al. Submitted). We arbitrarily fixed the percentage of the admixed population to be managed (removed or sterilized) at 20% (Table 2). For the non-targeted selection, we selected the individuals to be managed in the following way: first we calculated the corresponding number of admixed individuals to be managed by applying the pre-determined percentage to the number of admixed individuals in the population, second we removed or sterilized the selected number of admixed individuals from the population in the Removal sub-model or in the Sterilization sub-model. The status of the sterilized breeders individuals did not change, but they would produce litters size equal to 0.

For the targeted selection we aimed at selecting the breeders (60% of the total managed) and the mature non-breeders (40% of the total managed). Therefore, after calculating the number of admixed individuals to be removed/sterilized, we calculated the number of admixed breeders to be removed/sterilized and the number of admixed non-breeders to be removed/sterilized and drew them randomly from the admixed breeders and mature non-breeders in the population.

2.5. Parametrization

2.5.1 Case study and initial population

As a case study, we applied our model to simulate population dynamics in presence of hybridization for a wolf population in the Northern Apennines, Italy. The prevalence of hybrids in a sub-area (delineated as the area occupied by seven wolf packs) in the Northern Apennines has been estimated to be 64–78% by Santostasi et al (Submitted). We simulated hybridization dynamics starting from an initial population based upon the abundance of wolves and admixed individuals estimated by Santostasi et al (Submitted). We then simulated the dynamics of this wolf population over 10 generations (30 years; Skoglund et al 2011) by setting a carrying capacity for the overall projection area corresponding to 50 packs (cf. Caniglia et al. 2014).

2.5.1 Parameters

We chose the parameters based on the following criteria: i) when population-specific parameters were available, we used parameters that were specific for the Northern Apennines wolf population (i.e., pack carrying capacity, territory size, relatedness threshold for breeding; see Table 2), ii) when population-specific parameters were not available, we used parameters from the near and connected population in the Italian Alps (i.e., adult and yearling mortality; see Table 2), iii) if parameters were non-available for the northern Apennines and non-available for near and connected populations we browsed the available literature on gray wolf (*Canis lupus*) demographic parameters. We explain

457 the rationale for choice for each parameter in Appendix 1. We simulated large availability of dogs
458 by setting the maximum probability of mating with dogs as 0.9 and the maximum probability of
459 mating with dogs as 0.5 and a threshold wolves and admixed individuals density for the probability
460 of mating with dogs to be minimum equal to 100 individuals (wolves + admixed; Fig. 2).
461
462 **2.6 Simulation scenarios**
463 We projected hybridization dynamics using a set of hierarchical scenarios in which we first
464 accounted for immigration options (i.e., pure wolves vs admixed), then for mate-choice (i.e.,
465 random vs. assortative mating), and lastly for management alternatives (i.e., no management vs.
466 removal or sterilization, each considering targeted vs. non-targeted removal), for a total of 20
467 simulated scenarios (5 management scenarios for each of the 2 mate-choice options, for each of the
468 2 immigration options). In particular, we began by simulating two immigration scenarios: one
469 scenario in which all the immigrants are pure wolves (wolf genomic content = 100%), and one
470 scenario in which all the immigrants are admixed individuals whose wolf genomic content ranges
471 from that of a first generation hybrid (50%) to that of a third generation backcross (93.75%). Then
472 for each immigration scenario, we simulated two mate-choice scenarios: random and assortative.
473 Lastly, for each mate-choice scenario, we simulated 5 different management scenarios, comprising
474 one scenario with no management and 2 management scenarios (i.e., sterilization vs. removal)
475 considering both targeted and non-targeted selection of managed individuals. For each scenario, we
476 performed 50 replicates (population trajectories) and analyzed the following outputs: 1) prevalence
477 of admixed individuals, 2) wolf genomic content of the individuals in the population, 3) number of
478 wolf and admixed residents, and 4) number of wolf and admixed non-residents. In the output, we
479 considered all the individuals whose wolf genomic content was equal or greater to that of a third
480 generation backcross (93.75%) to be wolves, a category that is difficult to distinguish from wolves
481 when relying on molecular markers as currently used to detect hybridization through non-invasive
482 genetic surveys (Caniglia et al. 2018).
483
484 **3. RESULTS**
485 **3.1 Wolf immigration scenarios**
486 3.1.2 No management scenarios
487 The results showed that, in scenarios simulating immigration of pure wolves, the prevalence of
488 admixed individuals was projected to increase towards complete admixture in absence of
489 management in the random mating scenario, while in the assortative mating scenario it was
490 projected to decrease towards a plateau in which admixed individuals accounted for about 50% of

491 the population (for population here and thereafter we refer to all canids in the hybrid zone; Fig. 3).
492 The random mating scenario led to a general decrease of wolf ancestry (Fig. 4). In 10 generations,
493 50% of the individuals were second or third generation backcrosses (wolf ancestry 0.91–0.94), 15%
494 were from first or second generation backcrosses (wolf ancestry 0.71–0.91), and only 15% were
495 third generation backcrosses to pure wolves (wolf ancestry 0.94–1; Fig. 4). However, in the
496 assortative mating scenario, the ancestry of the individuals was more polarized between individuals
497 with high wolf content and individuals with low wolf genomic content (Fig. 4). In 10 generations,
498 the population was composed by 50% of individuals being from third to fifth generation
499 backcrosses (wolf genomic content ranging 0.93–0.98), 15% from first generation backcross to dog
500 to third generation backcrosses to wolf (wolf genomic content ranging 0.30–0.90), and 15% from
501 fifth generation backcrosses to pure wolves (wolf genomic content ranging 0.98–1; Fig. 4).
502 Starting from the first year, the population produced an increasing number of non-residents
503 that accounted for about 0.25–24% of the total population in both the random and assortative
504 mating scenarios (Fig. 5). In the random mating scenario, an average of 61% of the total non-
505 residents were admixed (Fig. 6), while in the assortative mating scenario an average of 49% of the
506 non-residents were admixed (Fig. 6).
507
508 3.1.3 Alternative management approaches
509 The two different management approaches we simulated affected prevalence of admixed
510 individuals differently depending on whether mate choice was random or assortative. For the
511 random mating scenario, removal through targeted selection of the individuals appeared to be the
512 most effective approach, obtaining the highest decrease in prevalence of admixed individuals,
513 followed by removal through non-targeted removal (Fig. 7). Non-targeted sterilization was the least
514 effective strategy, while targeted sterilization led to a substantial improvement in the efficacy of this
515 approach (Fig 7). For the assortative mating scenario, the most effective management approach was
516 targeted sterilization followed by targeted removal, non-targeted removal, and non-targeted
517 sterilization (Fig. 7). Compared to the random mating scenario, the outcome of the different
518 management scenarios varied less markedly. For the random mating scenario, average wolf ancestry
519 had the highest increase with targeted and non-targeted removal followed by targeted and non-
520 targeted sterilization (Fig. 8). The same pattern characterized the assortative mating scenario (Fig.
521 8).
522
523 **3.2 Admixed individuals immigration scenarios**

524 The results showed that, in the all the scenarios simulating the immigration of admixed individuals,
525 the final outcome was complete admixture, regardless of the mate choice and the management
526 scenario (Figures 4-8). In the scenarios with no management, the prevalence of admixed individuals
527 increased regardless of the mating choice scenario. In the scenarios with management, the
528 prevalence of admixed individuals increased to 100% within the random mating scenario. In the
529 assortative mating scenario, the prevalence of admixed individuals initially decreased but began
530 increasing once all the available territories were occupied.
531

4. DISCUSSION

Our modelling framework provides a new tool to inform decision making in the management of a globally-significant threat: anthropogenic hybridization. We were able to 1) investigate the impact of mate choice in driving hybridization dynamics in the absence of management, and 2) compare the effectiveness of different management strategies. By comparing alternative scenarios informed by the available literature, we were able to produce results informative for management despite considerable uncertainty regarding the strength of the underlying reproductive barriers. Moreover, we identified key processes that need to be investigated to effectively plan the management of hybridization. The results reported in this study should not be considered as generic because they are calibrated on a specific population characterized by a high prevalence of admixed individuals, but our model has the necessary flexibility to be customized for other *Canis* populations.

One main result is that in presence of frequent immigration of admixed individuals any management action seems ineffective. In particular, in the Assortative mating scenario, the effectiveness of management was contrasted by the fact that the majority of the territories were occupied by admixed pairs. Although the formation on wolf-wolf pairs is more likely than the formation of wolf-hybrid pairs in this scenario, the saturated landscape contrasts their establishment favoring the maintenance of high prevalence. This mechanism seems to reproduce what was observed for eastern wolves in Ontario, Canada where a landscape saturated with territories of coyote-eastern wolf hybrids reduced the success of wolves in expanding their range outside of the Algonquin Provincial Park (Benson & Patterson 2013). Understanding dispersal behavior of wolves and admixed individuals and source-sink dynamics of adjacent populations by integrating different data sources is fundamental to understand the mechanisms that structures hybridization patterns and accordingly allocate management efforts (Benson et al. 2012). The sensitivity of the results of the model to the different mate immigration scenarios highlights that this topic should represent a research priority. In the scenarios assuming immigration by pure wolves and in absence of management, the hybridization outcome after 10 generations diverged depending on the mate choice scenario. While in the random mating scenario the prevalence of admixed individuals was predicted to increase, in the assortative mating scenario it was predicted to decrease towards a plateau in which admixed individuals accounted for about half of the population. Moreover, in both scenarios without management, the impact of hybridization was predicted to extend across broader scales as large numbers of admixed non-residents were produced. Those admixed non-residents managing to establish a territory and reproduce would increase the spatial extent of hybridization geographically, potentially leading to introgression in other populations (Salvatori et al. 2019). Therefore, we recommend reactive management aimed at reducing prevalence of admixed

individuals under either scenarios, such management should be a joint effort among regional administrations responsible for wolf management across a wider scale (Salvatori et al. 2019). In the scenarios involving management, the simulated reactive management strategies reduced the prevalence of wolf x dog admixed individuals to different extents and the relative effectiveness of the strategies depended on the mate choice scenario. In the random mating scenario, removal of admixed individuals led to a greater decrease in prevalence compared to sterilization. In this case, pure wolves were injected each year in the non-residents pool increasing the probability that the removed breeders were replaced by wolves. Given that the mating was random, the mates of the removed admixed individuals were wolves or admixed individuals. Thus, if the breeders that remained in the population were wolves, the removal would in part contribute to the formation of wolf-wolf pairs. If the breeders that remain in the population were admixed, the removal would in part contribute to the formation of wolf-admixed pairs and to the dilution of the dog ancestry in the future litters. On the other hand, if mate choice is random and there is immigration of pure wolves, sterilization of admixed individuals may be counter-productive. Due to the long-term monogamy of wolves, the sterilization of admixed individuals interferes with the reproduction of those wolves that are paired with sterile admixed individuals, limiting wolf reproduction. Moreover, the presence of sterile mates would also prevent the potential replacement of the admixed individual by a wolf disperser. In the red wolf coyote-hybrid system admixed red wolf-coyote individuals in heterospecific pairs are likely to be displaced by pure red wolves in search for mates (Fredrickson & Hedrick 2006, Gese & Terletzky 2016). However, the same process has never been documented in wolf-wolf x dog admixed pairs and should be investigated.

When the mate choice is assortative, on the other hand, targeted sterilization produces the highest decrease in the prevalence of admixed individuals. In this case, there is a high chance that the sterilization would affect admixed-admixed pairs while not affecting the reproductive rate of wolf-wolf pairs. Evidence in support of assortative mating have been found in the red wolf-coyote hybrid system (Bohling & Waits 2015) and in the eastern-wolf coyote system (Rutledge et al. 2010) but have not directly been investigated in the wolf-dog system. A recent study reconstructed the pedigree of admixed individuals in the Northern Apennines, Italy reporting the occurrence of wolf-admixed pairs (Santostasi et al. Submitted) indicating that at least in some circumstances, the mate choice is not assortative. However, the anecdotal nature of such report does not allow to draw general conclusions. The sensitivity of the results of the model to the different mate choice scenarios highlights that this topic should represent a research priority to properly understand and manage wolf x dog hybridization.

599 Although the removal of admixed individuals is practiced in some countries (e.g., Spain;
600 Pacheco et al. 2017), sterilization of admixed individuals may be the only applicable reactive
601 management strategy that issocially acceptable, given concerns regarding animal welfare and the
602 economic costs of management. One useful result emerging from our simulations is that the
603 targeted approach dramatically improves the efficacy of the sterilization strategy. Such
604 improvement was likely due to the fact that sterilized admixed breeders act as 'placeholders'
605 preventing other admixed individuals from reaching a breeding position (Gese & Terletzky 2015).
606 A random approach, leading to the sterilization of non-breeding individuals, is not as effective in
607 reducing admixturebecause many of those individuals would have not have produced litters..
608 Interestingly, the same difference did not hold for the comparison between targeted and non-
609 targeted removal. The effectiveness of targeted removal may be partially counteracted by the fact
610 that the removal of breeders and mature individuals creates breeding opportunities that may be
611 filled by other admixed individuals. Moreover, sterilization involves the capture and the
612 manipulation of individuals, ensuring a higher selection of admixed individuals compared with
613 removal by culling. Removal by culling, if not done selectively, should be highly discouraged as it
614 has been demonstrated both by demographic projections (Fredrickson & Hedrick 2006: Santostasi
615 et al. in Press) and by observation studies (Rutledge et al. 2011; Benson et al. 2014) that reducing
616 the survival of mature wolves increases hybridization rates. To provide effective management
617 recommendations, this model should be carefully calibrated with parameters estimated from the
618 target population and the predicted decrease in prevalence should be one element to be considered
619 in the framework of a cost-benefit analysis.
620 Compared to other models (i.e., Fredrickson & Hedrick 2006, Santostasi et al. in Press) our
621 individual-based approach allowed us to distinguish between different hybrid categories both for the
622 estimation of the prevalence of admixed individuals and for the simulation of management. This
623 improves the realism of the model and mitigates the limited power of distinguishing > 3 generation
624 backcrosses with few loci. It also allowed us to explicitly set an acceptable threshold of dog
625 ancestry and to consider the individuals below that threshold as parental individuals.
626 Our model relies on a high number of parameters and assumptions, however, parameters
627 such as litter size, survival and dispersal rates, territory size and pack size may greatly vary
628 depending on population density and environmental conditions (Fuller et al. 2003) and will strongly
629 influence hybridization dynamics. Our model provides the flexibility to set these parameters to fit
630 any study population by customizing the demographic parameters. Where population-specific
631 parameters are not available the model allows to draw parameter values from a range of plausible

values and to test the sensitivity of hybridization rates to different demographic and environmental factors (McCarthy et al. 1995; Cross & Beissinger 2001; Ovenden et al. 2019).

In conclusion, our modeling approach showed that even with widespread introgression, implementation of reactive strategies can be a solution to reduce introgression, provided that the target population is not subject to frequent immigration of admixed individuals. On the other hand, the absence of management leads to complete admixture locally and to the geographic expansion of anthropogenic hybridization. Although further research is needed to validate the models' assumptions and predictions, all management scenarios we simulated produced a decrease in prevalence of admixed individuals. Thus, implementation of these management strategies should be considered for the mitigation of the introgression of domesticated genes into wild wolf populations..

Acknowledgements.

Nina Luisa Santostasi was supported by a PhD grant by Sapienza University of Rome. Olivier Gimenez and Nina Luisa Santostasi were supported by the French National Research Agency grant ANR-16-CE02-0007. We are grateful to Dr. Richard J. Fredrickson for his valuable advice in the early phase of the model construction.

CITED LITERATURE

Abbott, R.J., Barton, N.H., Good, J.M. (2016). Genomics of hybridization and its evolutionary consequences. Molecular Ecology 25, 2325–2332.

Ballard, W.B., Withman, J.S., Gardrner C.L. (1987). Ecology of an exploited population in south-central Alaska. Wildlife Monographs 98. The wildlife society, Bethesda, MD, USA.

Bauduin, S. McIntire, E.J., Chubaty, A.M. (2019). NetLogoR: a package to build and run spatially explicit agent-based models in R. Ecography, 42, 1841-1849.

Beaumont, M., Barratt, E.M., Gottelli, D., Kitchener, A.C., Daniels, M.J., Pritchard, J. Bruford, M.W. (2001). Genetic diversity and introgression in the Scottish wildcat. Molecular Ecology 10, 319-336.

Benson, J.F., Patterson, B.R. (2013). Inter-specific territoriality in a Canis hybrid zone: spatial segregation between wolves, coyotes, and hybrids. Oecologia 173, 1539-1550.

Benson, J. F., Patterson, B.R., Mahoney, P.J. (2014). A protected area influences genotype-specific survival and the structure of a Canis hybrid zone. Ecology, 95, 254-264.

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

Blanco, J.C. and Cortés, Y. (2007), Dispersal patterns, social structure and mortality of wolves living in agricultural habitats in Spain. Journal of Zoology, 273, 114-124.

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

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

Brainerd, S.M., Andrén, H., Bangs, E.E., Bradley, E.H., Fontaine, J.A., Hall, W., Iliopoulos, Y., et al. (2008). The effects of breeder loss on wolves. Journal of Wildlife Management 72, 89-98.

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

Caniglia, R., Mattucci, F., Velli, E., Galaverni, M., Canu, A. & Scandura, M. (2018). How much hybrid should a hybrid be? Empirical evaluation of proportions of membershipto define reliable assignment thresholds. Proceedings of the XI Congresso Italiano di Teriologia, Firenze, Italy.

Cariappa, C., Oakleaf, J.K., Ballard, W.B., Breck, S.W. (2011) A reappraisal of the evidence for regulation of wolf populations. The Journal of Wildlife Management, 75, 726-730

698 Chapron, G., Wikenros, C., Liberg, O., Wabakken, P., Flagstad, Ø, Milleret, C., J Månsson, J. et al.
699 (2016). Estimating wolf (*Canis lupus*) population size from number of packs and an individual
700 based model. Ecological Modelling 339, 33-44.
701
702 Ciucci, P., Reggioni, W., Maiorano, L. & Boitani, L. (2009). Long-distance dispersal of a rescued
703 wolf from the northern Apennines to the western Alps. Journal of Wildlife Management 73,
704 1300-1306.
705
706 Cross, P.C., Beissinger, S.R. (2001). Using logistic regression to analyze the sensitivity of PVA
707 models: a comparison of methods based on African wild dog models. Conservation Biology 15,
708 1335-1346.
709
710 Cubaynes, S. , MacNulty, D. R., Stahler, D. R., Quimby, K. A., Smith, D. W. Coulson, T. (2014).
711 Density-dependent intraspecific aggression regulates survival in northern Yellowstone wolves
712 (*Canis lupus*). Journal of Animal Ecology 83, 1344-1356.
713
714 Cubaynes, S., Pradel, R., Choquet, R., Duchamp, C., Gaillard, J.-M., Lebreton, J.-D., Marboutin, E.,
715 et al. (2010). Importance of accounting for detection heterogeneity when estimating abundance:
716 the case of French wolves. Conservation Biology 24, 621-626.
717
718 Donfrancesco, V., Ciucci P., Salvatori V., Benson, D., Andersen, L.W., Bassi, E., et al. (2019).
719 Unravelling the scientific debate on how to address wolf-dog hybridization in Europe. Frontiers
720 in Ecology and Evolution 7, 1-13.
721
722 Dufresnes, C., Remollino, N., Stoffel, C., Manz, R., Weber, J. M., Fumagalli, L. (2019). Two
723 decades of non-invasive genetic monitoring of the grey wolves recolonizing the Alps support
724 very limited dog introgression. Scientific Reports 9,1-9.
725
726 Fredrickson, R.J., Hedrick, P.W. (2006). Dynamics of hybridization and introgression in red wolves
727 and coyotes. Conservation Biology 20, 1272-1283.
728
729 Gese E.M., Mech, D. (1991). Dispersal of wolves (*Canis lupus*) in northeastern Minnesota, 1969-
730 1989. Canadian Journal of Zoology 69, 2946-2955.
731

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

Gervasi, V., Ciucci P. (2018). Demographic projections of the Apennine brown bear population *Ursus arctos marsicanus* (Mammalia: Ursidae) under alternative management scenarios. The European Zoological Journal 85, 243-253.

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

Gompert, Z., Buerkle, C. A. (2016). What, if anything, are hybrids: enduring truths and challenges associated with population structure and gene flow. Evolutionary Applications 9, 909-923

Gottelli, D., Sillero-Zubiri, C., Applebaum, G.D., Roy, M.S., Girman, D.J., Garcia-Moreno, J., Ostrander, E.A., Wayne, R.K. (1994). Molecular genetics of the most endangered canid: the Ethiopian wolf *Canis simensis*. Molecular Ecology 3, 301-312.

Hall, R.J., Ayres D.R. (2008). What can mathematical modeling tell us about hybrid invasions? Biological Invasions 11, 1217-1224.

Hooftman, D.A.P., De Jong, M.J., Oostermeijer, J.G.B., den Nijs, H.C.M. (2007). Modelling the long-term consequences of crop-wild relative hybridization: a case study using four generations of hybrids. Journal of Applied Ecology 44, 1035-1045.

Hinton, J.W., Gittleman, J.L., van Manen, F.T., Chamberlain, M.J. (2018). Size-assortative choice and mate availability influences hybridization between red wolves (*Canis rufus*) and coyotes (*Canis latrans*). Ecology and Evolution 8, 3927-3940

Hradsky, BA, Kelly, LT, Robley,j A, Wintle, BA. (2019). FoxNet: an individual-based model framework to support management of an invasive predator, the red fox. Journal of Applied Ecology 56, 1460–1470.

766 Jędrzejewska, B. , Jędrzejewski, W. , Bunevich, A. N., Minkowski, L. And Okarma, H. (1996),
767 Population dynamics of Wolves *Canis lupus* in Bialowieża Primeval Forest (Poland and Belarus)
768 in relation to hunting by humans, 1847-1993. Mammal Review, 26, 103-126.
769
770 Jimenez, M.D., Bangs, E.E., Boyd, D.K., Smith, D.W., Becker, S.A., Ausband, D.E., Woodruff, et
771 al. (2017), Journal of Wildlife Management 81, 581-592.
772
773 Lorenzini, R., Fanelli, R., Grifoni, G., Scholl, F. & Fico, R. (2014). Wolf-dog crossbreeding:
774 'Smelling' a hybrid may not be easy. Mammalian Biology. 79, 149-156.
775
776 Mancinelli, S., Boitani, L., Ciucci P. (2018). Determinants of home range size and space use
777 patterns in a protected wolf (*Canis lupus*) population in the central Apennines, Italy. Canadian
778 Journal of Zoology 96, 828-838.
779
780 Marucco, F., McIntire, E.J.B. (2010). Predicting spatio-temporal recolonization of large carnivore
781 populations and livestock depredation risk: wolves in the Italian Alps. Journal of Applied
782 Ecology 47, 789-798.
783
784 McCarthy, M. A., Burgman, M. A.,Ferson, S. (1995). Sensitivity analysis for models of population
785 viability. Biological Conservation 73, 93-100.
786
787 Fuller, L., Mech, D., Cochrane J.F. (2003). Wolf population dynamics. In: Mech, L.D., Boitani, L.
788 (Eds.). Wolves: Behavior, Ecology, and Conservation. University of Chicago Press, Chicago,
789 USA, pp. 1–35.
790
791 Mech, L.D., Boitani, L. (2003). Wolf social ecology. In: Mech, L.D., Boitani, L. (Eds.). Wolves:
792 Behavior, Ecology, and Conservation. University of Chicago Press, Chicago, USA, pp. 161–192.
793
794 Nathan, L.R., Mamoozadehb, N., Tumasc, H.R., Gunselmand, S., Klasse, K., Metcalfef, A., Edgeg,
795 C., et al. (2019). A spatially-explicit, individual-based demogenetic simulation framework for
796 evaluating hybridization dynamics. Ecological Modelling 401, 40-51.
797
798 Pacheco, C., López-bao, J. V., García, E. J., Lema, F. J., Llaneza, L., Palacios, V. & Godinho, R.
799 (2017). Spatial assessment of wolf-dog hybridization in a single breeding period. Scientific

800 Reports 7, 1-10.
801
802 Packard, J. (2003). Wolf behavior: reproductive, social and intelligent. In: Mech, L.D., Boitani, L.
803 Wolf social ecology. In: Mech, L.D., Boitani, L. (Eds.). Wolves: Behavior, Ecology, and
804 Conservation. University of Chicago Press, Chicago, USA, pp. 161-192.
805
806 Rhymer, J.M., Simberloff, D. (1996). Extinction by hybridization and introgression. Annual Review
807 of Ecology and Systematics 27, 83-109.
808
809 Ritchie. E.G., Letnic, C.R. & Vanak, A.T. (2014). Dogs as predators and trophic regulators, in:
810 Gompper M.E (Ed.). Free-Ranging Dogs and Wildlife Conservation: 55-68. Oxford University
811 Press, Oxford, UK, pp. 55-68.
812
813 Rutledge, L.Y., Garroway, C.J., Loveless, K.M., Patterson, B.R. (2010). Genetic differentiation of
814 eastern wolves in Algonquin Park despite bridging gene flow between coyotes and grey wolves.
815 Heredity 105, 520-531.
816
817 Rutledge, L.Y., White, B.N., Row, J.R., Patterson, B.R. (2011). Intense harvesting of eastern
818 wolves facilitated hybridization with coyotes. Ecology and Evolution 2, 19-33.
819
820 Santostasi, N.L., Ciucci, P., Bearzi, G., Bonizzoni, S., Gimenez, O. (In press). Assessing the
821 dynamics of hybridization through a matrix modelling approach. Ecological Modelling.
822
823 Santostasi, N.L., Gimenez, O., Caniglia, R., Fabbri, E., Molinari, L., Reggioni, W. (Submitted).
824 When bad gets worse: unmanaged anthropogenic hybridization between dogs and wolves may
825 lead to widespread introgression. Animal Conservation
826
827 Salvatori, V., Donfrancesco, V., Trouwborst, A., Boitani, L., Linnell, J., Alvares, F., Åkesson, M.,
828 et al. (Submitted). European agreements for nature conservation need to explicitly address
829 wolfdog hybridisation. Biological Conservation.
830
831 Salvatori, V., Godinho, R., Braschi, C., Boitani, L. Ciucci, P. (2019). Record-high levels of
832 wolf x dog introgressive hybridization in agricultural landscapes of central Italy. European Journal
833 of Wildlife Research 65, 73-88.

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

Smith, D. W., Bangs, E. E., Oakleaf, J. K., Mack, C. , Fontaine, J. , Boyd, D. , Jimenez, M., et al. (2010). Survival of Colonizing Wolves in the Northern Rocky Mountains of the United States, 1982-2004. The Journal of Wildlife Management, 74, 620-634.

Stephens, D., Wilton, A.N., Fleming, P.J. Berry, O. (2015). Death by sex in an Australian icon: a continent-wide survey reveals extensive hybridization between dingoes and domestic dogs. Molecular Ecology, 24, 5643-5656.

Stoskopf, M., Beck, K., Fazio, B., Fuller, T., Gese, E., Kelly, B., Knowlton, F., et al. (2005). Implementing recovery of the red wolf integrating research scientists and managers. Wildlife Society Bulletin 33, 1145-1152.

Thompson, C.J., Thompson, B.J.P., Ades, P.K., Cousens, R., Garnier-Gere, P., Landman, K., et al. (2003). Model-based analysis of the likelihood of gene introgression from genetically modified crops into wild relatives. Ecological Modelling 162, 199-209.

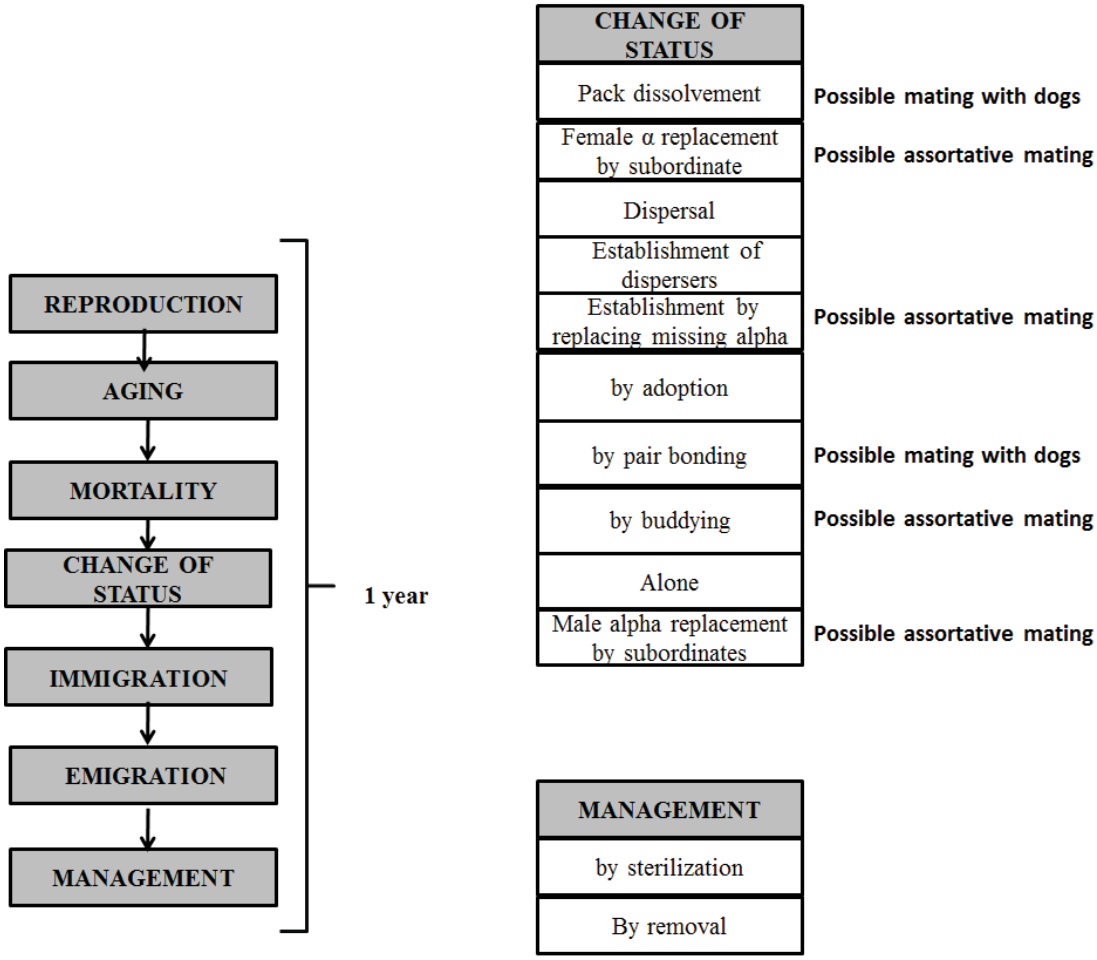
Todesco, M., Pascual, M.A., Owens, G.L., Ostevik, K.L., Moyers, B.T., Hübner, S., Heredia, S.M., et al. (2016). Hybridization and extinction. Evolutionary Applications 9, 892-908.

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

Vilà, C., Wayne, R.K. (1999). Hybridization between wolves and dogs. Conservation Biology 13, 195-198.

867 Ward, T.J., Bielawski, J.P., Davis, S.K., Templeton, J.W., Derr, J.N. (1999). Identification of
868 domestic cattle hybrids in wild cattle and bison species: a general approach using mtDNA
869 markers and the parametric bootstrap. Animal Conservation, 2, 51-57.
870
871 Wolf, D.E., Takebayashi, N., Rieseberg, L.H. (2001). Predicting the risk of extinction through
872 hybridization. Conservation Biology 15, 1039-1053.
873

874 **FIGURES**



875

876 **Figure 1**. General model structure. The seven main sub-models representing the wolf life cycle are
877 listed on the left. The further structure of the Change of social status and of the Management sub-
878 model are shown on the right.

879

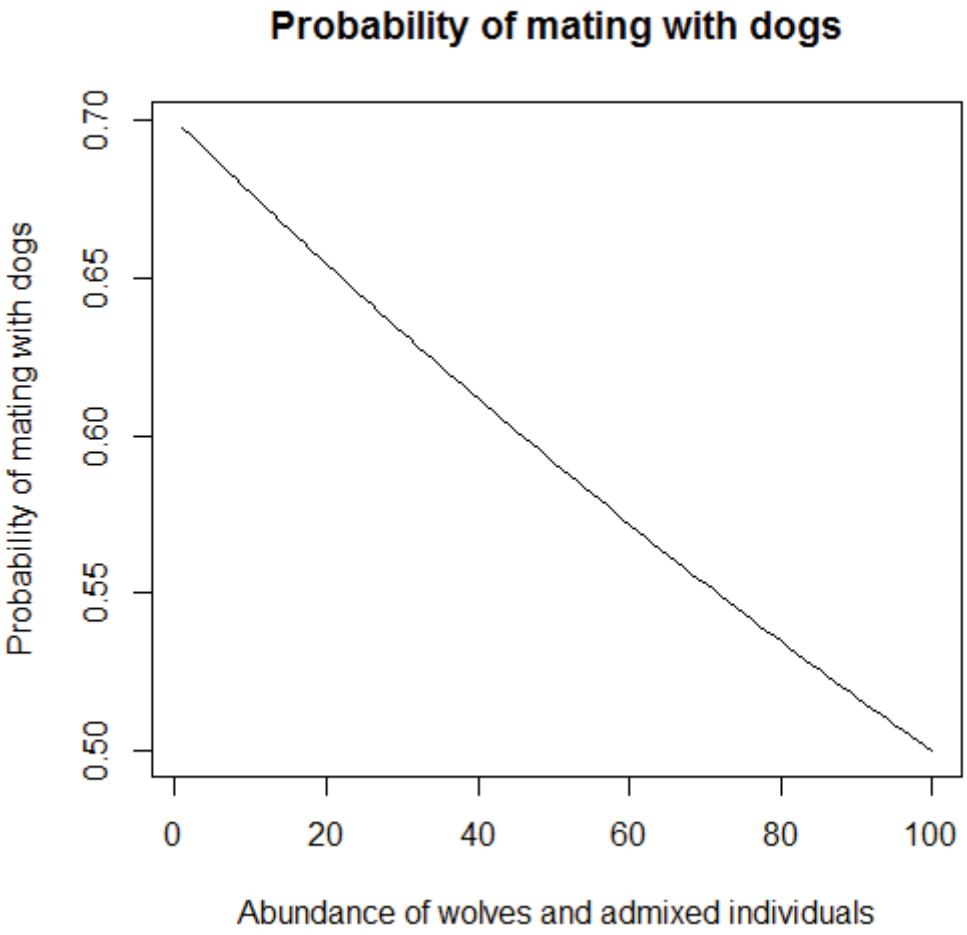


Figure 2. Probability of mating with dogs (i.e., the original hybridization events producing F_1s). The probability is maximum (0.9) when the abundance of wolves and admixed individuals is small and decreases exponentially reaching its smallest value (0.5) when the abundance of wolves and admixed individuals reaches a threshold of 100, after the threshold is hit the probability becomes 0.

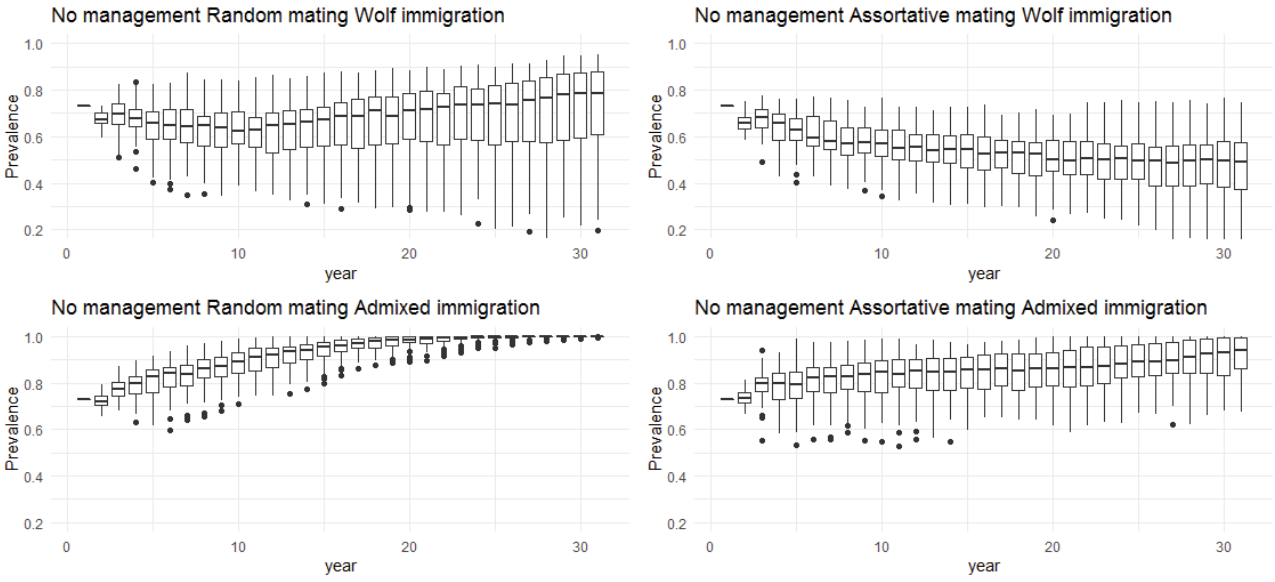


Figure 3. Prevalence of admixed individuals projected for 30 years (10 generations) under the Random mating (left column) and the Assortative mating scenarios (right column) and under the Wolf immigration (top row) and Admixed immigration (bottom row) scenarios. Each boxplot summarizes the prevalence values obtained for each of the simulated year and for each simulation replicate (n=50).

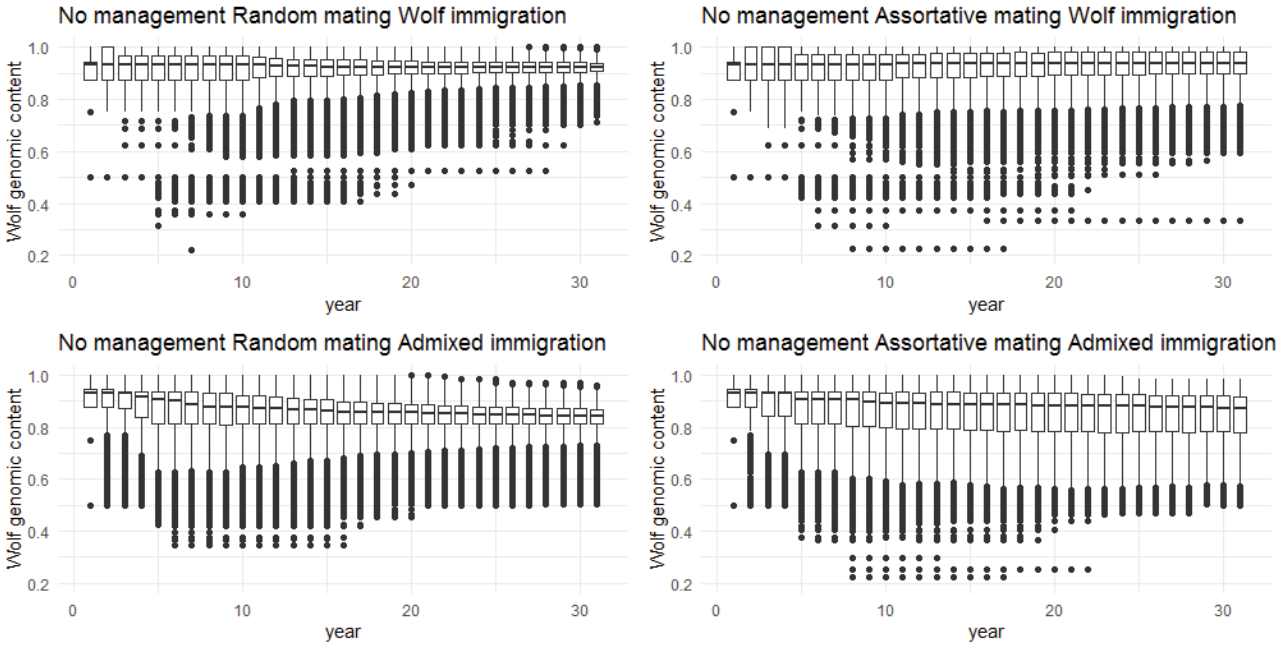


Figure 4. Average wolf genomic content of the simulated individuals under the Random mating (left column) and the Assortative mating scenarios (right column) and under the Wolf immigration (top row) and Admixed immigration (bottom row) scenarios. Each boxplot summarizes the average wolf genomic content across the individuals in the population obtained for each of the simulated year and for each simulation replicate (n=50).

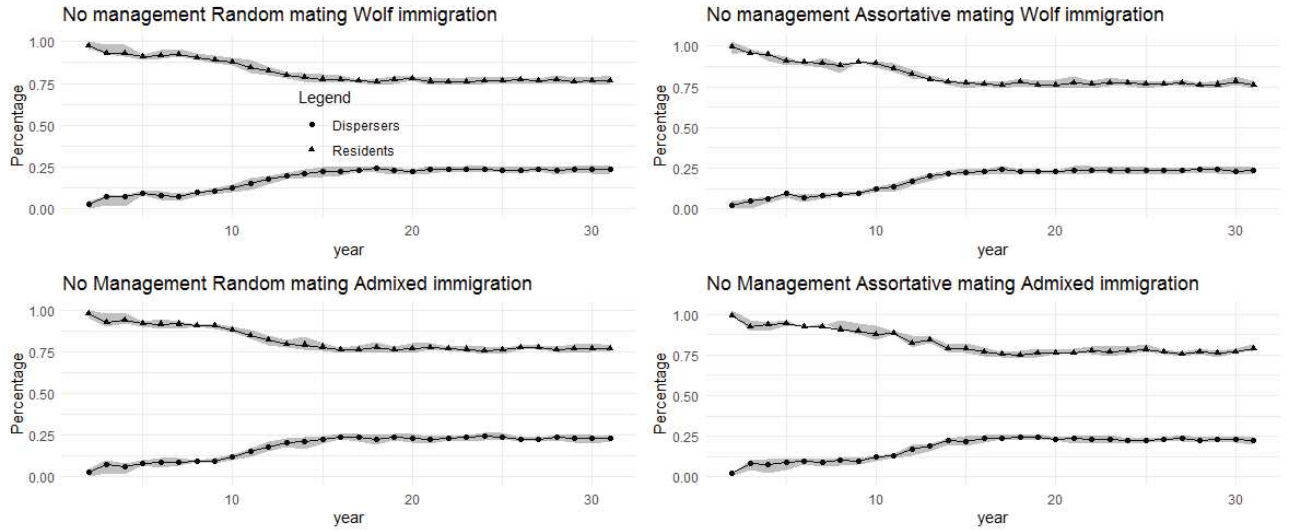


Figure 5. Percentage of disperser and resident individuals in the simulated population under the Random mating (left column) and the Assortative mating scenarios (right column) and under the Wolf immigration (top row) and Admixed immigration (bottom row) scenarios. Each point represents the average value for each year over the simulation replicates (n=50). The shaded area is the interquartile range of the values.

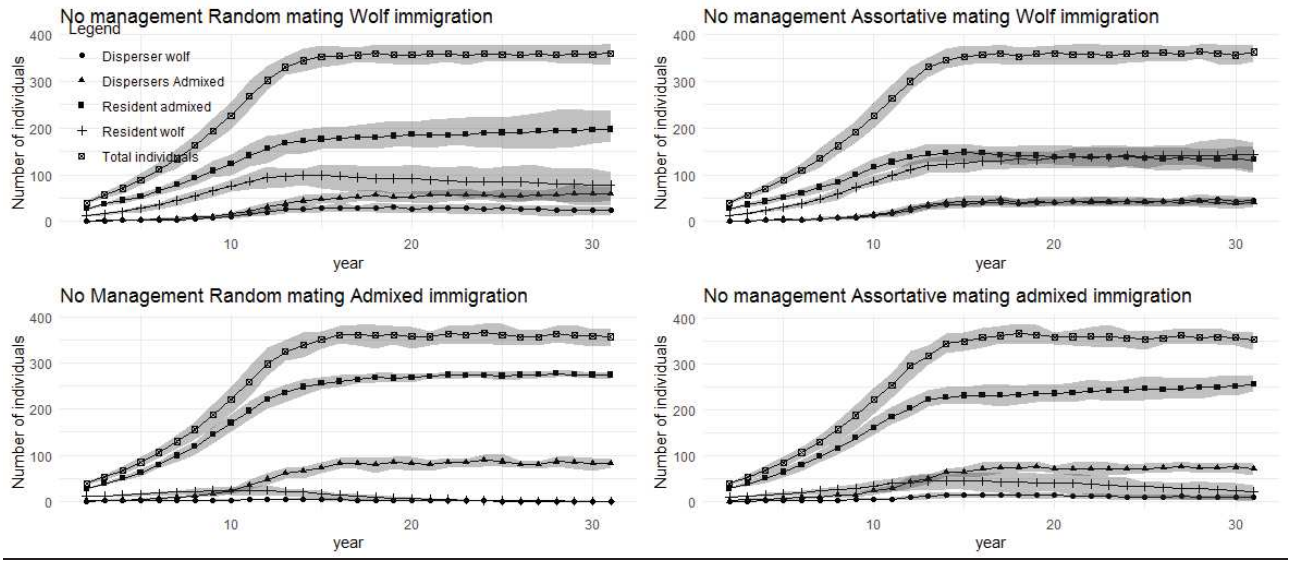


Figure 6. Projected number of resident and non-resident wolves and admixed individuals under the Random mating (left column) and the Assortative mating scenarios (right column) and under the Wolf immigration (top row) and Admixed immigration (bottom row) scenarios. Each point represents the average value for each year over the simulation replicates (n=50). The shaded area is the interquartile range of the values. Right panel: number of total non-resident and resident individuals per simulated year divided in the admixed and wolf categories. Each point represents the average value for each year over the simulation replicates (n=50). The shaded area is the range between the first and third quartiles of the values for each scenario.

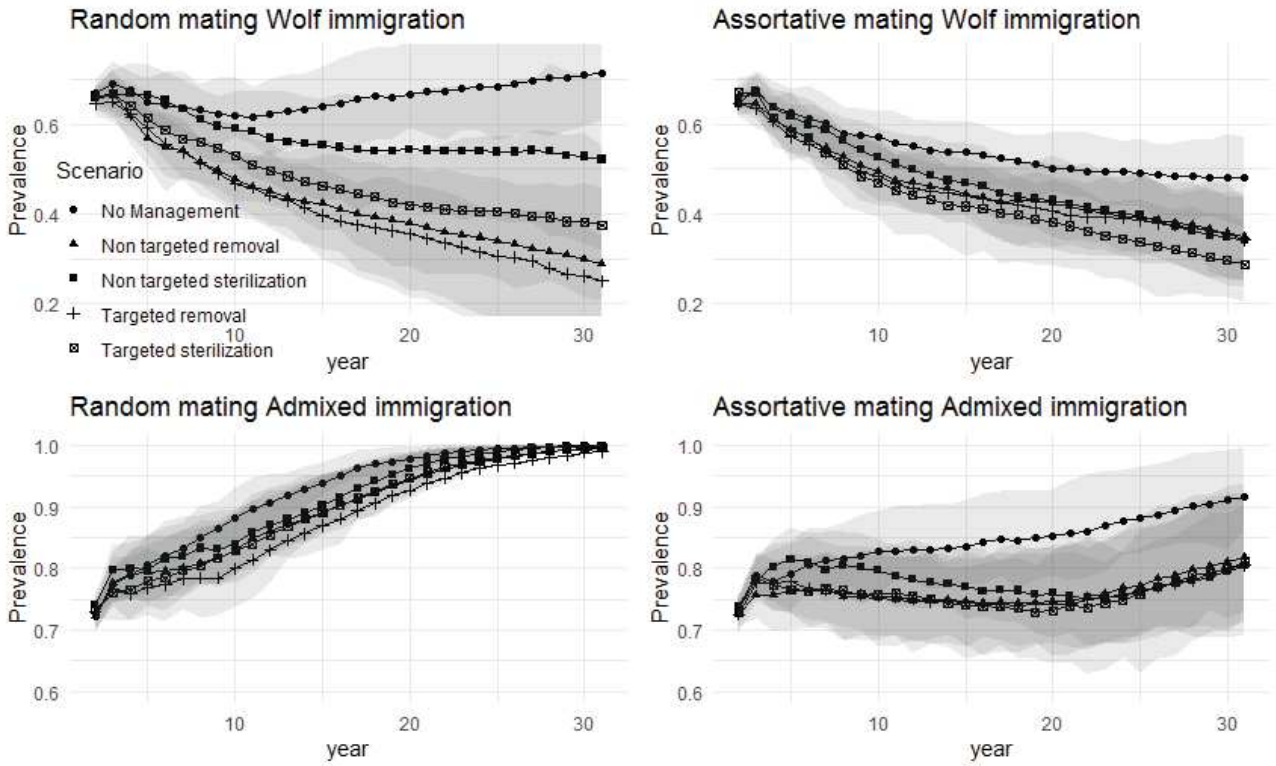


Figure 7. Simulated prevalence dynamics under different management strategies (different point shapes) under the Random mating (left column) and the Assortative mating scenarios (right column) and under the Wolf immigration (top row) and Admixed immigration (bottom row) scenarios. Each point represents the average value for each year over the simulation replicates (n=50). The shaded area is the interquartile range of the values for each scenario.

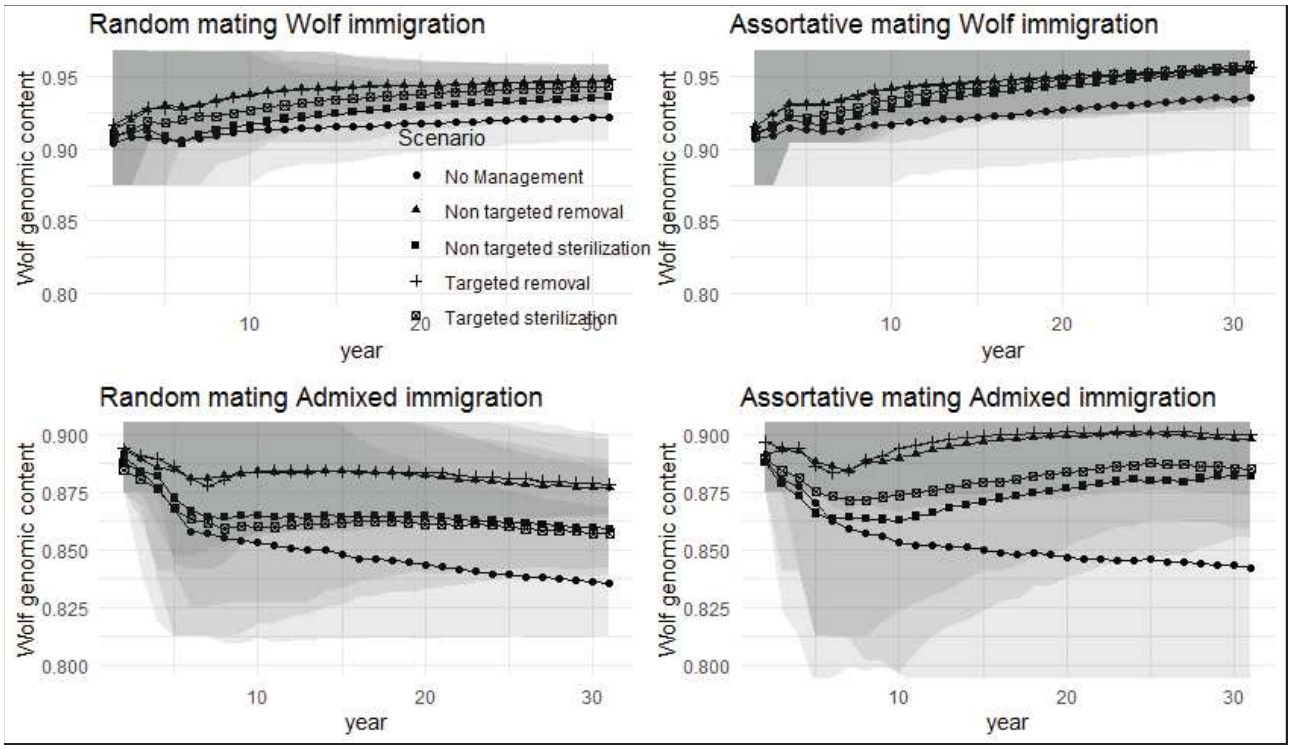


Figure 8. Scenario simulating the immigration of pure wolves. Simulated average individual wolf genomic content dynamics under different management strategies (different point shapes) and under the Random mating (left column) and the Assortative mating scenarios (right column) and the Wolf immigration (top row) and Admixed immigration (bottom row) scenarios. Each point represents the average wolf genomic content value for the individuals in the population for each year over the simulation replicates (n=50). The shaded area is the interquartile range of the values for each scenario.

940 **TABLES**

941

942 **Table 1.** Individual characteristics used in the Individual-Based Model.

Age class	- pup (0-12 months old; age =1) - yearling (12-24 months old; age =2) - adults (25+ months old; 3+)
Mother ID and father ID	Identity of the parents
Residency status	non-resident vs. resident in a pack
If resident: • **Pack ID**	pack ID number
• **Social status**	Breeder vs. subordinate
Wolf genomic content	Percentage of wolf genomic content calculated as ½ wolf genomic content of the mother + ½ wolf genomic content of the father
Sex	Female vs. male
If female • **Breeder experience**	has reproduced before vs. never reproduced
Cohort	year of birth

943

944

945 **Table 2.** Parameters used for the projections.

Submodel	Parameter	Value	Reference
Reproduction	Mean litter size	6.1	Sidorovich et al. 2007
	pups sex ratio	50:50	
Mortality	Annual pup mortality	0.398 (95% CI: 0.273, 0.579)	Smith et al 2010
	Yearling true mortality	0.18 (+/- 0.04)	Marucco and McIntire, 2010
	Adult true mortality when vacant territories are available	0.18 (+/- 0.04)	Marucco and McIntire, 2010
	Adult mortality at carrying capacity	$\text{logit}(\text{survival}_{\text{adult}}) = 1.196-0.505)(\text{Density}_{\text{adult\&yearlings}})$	Cubaynes et al 2014
	Territory size for calculating density	104 km^2	Mancinelli et al 2019
	Dispersing pup survival	0	
	Disperser survival (yearlings and adults)	0.7	Blanco and Cortés, 2007
Pack dissolution	Probability of dissolution for small packs with 1 breeder	0.258	Brainerd et al. 2008
	Probability of dissolution for packs with 0 breeder	0.846	Brainerd et al. 2008
	Pack size threshold for potential dissolution	4.055	Brainerd et al 2008
Dispersal	Mean pack size	5.6 (SD = 1.251, Min = 3, Max = 8)	Caniglia et al 2014
	Pup dispersal probability	0.25	Haight and Mech, 1997
	Yearling dispersal probability	0.5	Haight and Mech, 1997
	Adult dispersal probability	0.9	Haight and Mech,

			1997
	Probability of successful adoption	0.5	Arbitrary
Establishment	Carrying capacity	50 packs	Caniglia et al 2014
	Relatedness threshold above which mating is avoided	0.125	Caniglia et al 2014
	Probability of successful budding	0.5	Arbitrary
Hybridization	Pmin	0.5	
	Pmax	0.9	
	Nthresh	100	
Migration	Number of permanent emigrants	1 to 5	Arbitrary
	Number of immigrants	1 to 5	Arbitrary
Management	Percentage to be managed	20%	
	Wolf ancestry threshold to be managed	0.75% first generation backcrosses	
	Percentage of breeders for targeted management	60% of the managed individuals	
	Percentage of mature non-breeders for targeted management	40% of the managed individuals	

946

947

APPENDIX 1: rationale behind parameters choice

We chose the parameters based on the following criteria:

- when population- specific parameters were available, we used parameters that were specific for the Northern Apennines wolf population (i.e., pack carrying capacity, territory size, relatedness threshold for breeding; see Table 2)
- when population-specific parameters were not available, we used parameters from the near and connected population in the Italian Alps (i.e., adult and yearling mortality; see Table 2)
- If parameters were non-available for the northern Apennines and non-available for near and connected populations we browsed the available literature on gray wolf (*Canis lupus)* demographic parameters. Below we explain the rationale for choice of those parameters.

Litter size

We relied on the litter size listed in Sidorvitch et al 2006. This study relies on data from 101 litters found by experienced hunters between 1985 and 2003 and refers to pups that are up to one month of age. Such number is not far from the average litter size calculated from 16 subpopulations in Europe: 4.4 - 7.7 (average 5.9) pups per year by Sidorvitch et al 2006.

Pup mortality

To derive pup mortality we used the annual survival value estimated by Smith et al 2010 for gray wolf pups in the Northern Rocky Mountains. This mortality includes includes autumn and winter mortalities and is not far from pup survival estimated by Jedrzewksa et al. (1996) in Europe (i.e., Belarus and Poland) which is 0.35.

Dispersers mortality

We used the value estimated by Blanco and Cortés (2007) for Iberian wolves (*Canis lupus signatus*) dispersing across human-dominated landscapes.

Adult mortality at carrying capacity

We used the linear relationship estimated by Cubaynes et al. (2014) for gray wolves in Yellowstone in conditions of availability of food and saturated landscape, which resemble the ecological conditions of the Northern Apennines.

Probability of pack dissolution

We relied on the probabilities estimated by Brainerd et al. (2008) based on a meta-analysis from multiple wolf populations across Europe and North-America.

Dispersal probabilities

Wolves routinely disperse in response to competition and aggression related to food availability and breeding opportunity within their pack (Mech and Boitani 2003). Wolves disperse at different ages because of limited food resources (Fritts and Mech 1981, Ballard et al. 1987, Mech 1987, Fuller 1989, Gese and Mech 1991), and often postpone dispersal in areas of high prey availability (Ballard et al. 1987; Blanco and Cortes 2007; Jimenez et al 2017). The proportion of yearlings that disperse may increase when food is scarce (Peterson and Page 1983, Messier 1985). As an example in a Spanish agricultural area (with high human-population density, a dense network of roads and a shortage of wild ungulates), 5 of the 7 known-age wolves in the study dispersed when between 2 and 3 years old (adult class in our model), and the other 2 as yearlings (subadult class in our model) (Blanco and Cortes, 2007). The dispersal of the 2 yearlings happened when the food surplus had disappeared. As another example in a 15 year study in the the Norther Rocky Mountains Jimenez et al. (2017) found an average age of dispersal of 32 months (about 3 years old, adult class in our model) and attributed this older age at dispersal to abundant wild prey in the area. However, dispersal rates increase rapidly with the onset of sexual maturity (Packard and Mech 1980; Messier 1985b; Gese and Mech 1991) so that most wolves disperse from their natal pack by the age of 3 years (Gese and Mech 1991, Mech et al. 1998, Jimenez et al 2017).

Based on literature evidence we modelled the following biological mechanism: regardless of the food availability (that affects the proportion of yearling dispersers), dispersal probability increases with age due to mate competition. Individuals that did not disperse as yearlings would almost surely leave as adults at the onset of sexual maturity. We used a 90% probability for adults' dispersal to account for the fact that, although rare, there are observation of adults that remain in their natal pack for serveral years (e.g. Gese and Mech 1991). In this way we make sure that most wolves disperse from their natal pack by the age of 3 years which is largely supported by the literature. We preferred to simplify the model by directly coding the dispersal mechanism (chances of dispersing increasing with sexual maturity) instead of using the proportion of dispersers observed in different age classes in different studies because they can suffer from observed bias and they can be ecological-context dependent (i.e., related to food availability).

Probability of successful establishment and probability of successfull budding

1016 These processes have been observed in free-ranging wolves, but their actual incidence at the
1017 population level are unkown. In particular the presence of unrelated wolves (adoptees) has been
1018 observed in the Northern Apennines by Caniglia et al. (2014) and also the fundation of new packs
1019 by subordinates and dispersers (budding). Since we did not have available parameter to quantify
1020 such processes at the population lelve we used arbitrary values (probability of success 0.5) and then
1021 explored the sensitivity of model outputs to the variation in those parameters.
1022

APPENDIX 2: Individual Based Model Sesitivity analysis

from Bauduin, S., Grente, O., Santostasi, N.L., Ciucci, P., Duchamp, C., Gimenez, O. (In preparation). From individual and pack dynamics to population responses: An individual-based approach to model wolf life cycle. Target Journal: Ecological Modelling.

METHODS:

We run the model modifying one parameter value (Table 1) at a time. We increase and decrease the focused parameter value by 5% and run 200 replicates of 25 years simulations (Ovenden et al., 2019). The model is considered sensitive to a parameter if a model output (i.e., mean value over the 200 replicates) with the one modified parameter varies more than 20% from the original results (Kramer-Schadt et al., 2005; Ovenden et al., 2019). We examine the following model outputs:

- Number of packs with both breeders
- Number of new packs founded each year
- Number of individuals
- Proportion of residents in the population
- Relatedness between breeding pairs

Results

The parameter affecting the most model outputs is the carrying capacity: the number of new packs created and the relatedness between breeders in packs are sensitive to this parameter. The number of new packs created is more sensitive to the pack size threshold involved in the pack dissolvement process and also to the probability of succesfull adoption. The relatedness between breeders in packs is sensitive to the relatedness threshold. Among the 21 parameters tested, the increase or the decrease of the used values for 17 of them did not modify the oucome. The complete table with the value tested for the parameters and the results of the five outputs for each simulation run is available Table S1.

TABLE S1. Complete results of the sensitivity analysis. The first line of the table is the name of the simulation runs: for the original complete version of the model and the runs S1 to S42 are the runs where one parameter of the model was modified by decreasing or increasing its value by 5%. The second line informs which parameter was modified in the run and the following line gives the value used for this parameter. Then, the five following line are the five selected model outputs: the number of packs with both breeders, the number of new packs created each year, the number of

individuals, the proportion of resident individuals and the relatedness between the breeders. The result values are the mean values over the 200 simulation replicates for each run. The column "[- 20% ; + 20%]" presents the results for the run with M0 with the range - 20% and + 20% of the result values. Then, table cells are the mean values of the model outputs obtained with the runs S1 to S42. Grey cells are model results outside of the reference range of M0 results [- 20% ; + 20%].

Simulation run	[- 20%; + 20%]	S1	S2	S3	S4	S5	S6
Parameter modified		**Mean litter size (-5%)**	**Mean litter size (+5%)**	**Pup mortality (-5%)**	**Pup mortality (+5%)**	**Yearling mortality (-5%)**	**Yearling mortality (+5%)**
Parameter value used		5.795	6.405	0.572	0.632	0.171	0.189
Number of packs with both a male and a female alpha	29.3 [23.4;35.2]	29.5	29.5	29.4	29.3	29.5	29.4
Number of new packs created	2.1 [1.7;2.5]	2.3	2.2	2.4	2.3	2.3	2.3
Number of individuals	186.7 [149.4;224]	182.1	191.7	193.9	179.2	187.8	186.3
Proportion of resident individuals	0.71 [0.57;0.85]	0.72	0.69	0.68	0.72	0.7	0.7
Relatedness between the male alpha and the female alpha in a pack	0.06 [0.05;0.07]	0.06	0.06	0.06	0.06	0.06	0.06
Simulation run		S7	S8	S9	S10	S11	S12

Parameter modified		**Non density-dependent adult mortality (-5%)**	**Non density-dependent adult mortality**	**Slope of density dependent adult mortality (-5%)**	**Slope of density dependent adult mortality (+5%)**	**Dispersing pup mortality (-5%)**	**Dispersing pup mortality (+5%)**
Parameter value used		0.171	0.189	-0.48	-0.53	0.95	NA
Number of packs with both a male and a female alpha	29.3 [23.4;35.2]	29.4	29.4	29.5	29.5	29.5	
Number of new packs created	2.1 [1.7;2.5]	2.3	2.5	2.2	2.3	2.3	
Number of individuals	186.7 [149.4;224]	187.4	186.1	187.5	186.4	189.2	
Proportion of resident individuals	0.71 [0.57;0.85]	0.7	0.71	0.71	0.7	0.7	
Relatedness between the male alpha and the female alpha in a pack	0.06 [0.05;0.07]	0.06	0.06	0.06	0.06	0.06	
Simulation run		S13	S14	S15	S16	S17	S18
Parameter modified		**Disperser**	**Disperser**	**Dissolvement**	**Dissolvement**	**Dissolvement**	**Dissolvement**

		mortality (-5%)	mortality	probability for pack with 1 breeder (-5%)	probability for pack with 1 breeder (+5%)	probability for pack with 0 breeder (-5%)	probability for pack with 0 breeders (+5%)
Parameter value used		0.295	0.326	0.245	0.271	0.804	0.888
Number of packs with both a male and a female alpha	29.3 [23.4;35.2]	29.5	29.4	29.4	29.6	29.5	29.6
Number of new packs created	2.1 [1.7;2.5]	2.1	2.2	2.2	2.3	2.1	2.3
Number of individuals	186.7 [149.4;224]	189	186.2	186	186.5	186.8	186.9
Proportion of resident individuals	0.71 [0.57;0.85]	0.7	0.71	0.71	0.7	0.71	0.7
Relatedness between the male alpha and the female alpha in a pack	0.06 [0.05;0.07]	0.06	0.06	0.06	0.06	0.06	0.06
Simulation run		S19	S20	S21	S22	S23	S24
Parameter modified		**Pack size threshold for diss. (-5%)**	**Pack size threshold for diss. (+5%)**	**Mean pack size (-5%)**	**Mean pack size (+5%)**	**Pup dispersal probability (-5%)**	**Pup dispersal probability (+5%)**

Parameter value used		3.852	4.258	4.185	4.625	0.238	0.263
Number of packs with both a male and a female alpha	29.3 [23.4;35.2]	29.6	29.4	29.3	29.5	29.4	29.5
Number of new packs created	2.1 [1.7;2.5]	**1.3**	**2.5**	2.4	2.1	2.4	2.4
Number of individuals	186.7 [149.4;224]	188.6	187.8	180.3	192.8	186.7	185.9
Proportion of resident individuals	0.71 [0.57;0.85]	0.72	0.7	0.69	0.72	0.7	0.71
Relatedness between the male alpha and the female alpha in a pack	0.06 [0.05;0.07]	0.06	0.06	0.06	0.06	0.06	0.06
Simulation run		S25	S26	S27	S28	S29	S30
Parameter modified		**Yearling dispersal probabilit y (-5%)**	**Yearling dispersal probability (+5%)**	**Adult dispersal probability (-5%)**	**Adult dispersal probability (+5%)**	**Probability of adopting (-5%)**	**Probability of adoption (+5%)**
Parameter value used		0.475	0.525	0.855	0.945	0.475	0.525
Number of packs with both a male and	29.3 [23.4;35.2]	29.4	29.5	29.5	29.5	29.6	29.5

a female alpha							
Number of new packs created	2.1 [1.7;2.5]	2.4	2.4	2.2	2.4	**2.7**	2.4
Number of individuals	186.7 [149.4;224]	187.5	188.5	186.7	188.1	188.4	187.8
Proportion of resident individuals	0.71 [0.57;0.85]	0.7	0.7	0.7	0.7	0.7	0.7
Relatedness between the male alpha and the female alpha in a pack	0.06 [0.05;0.07]	0.06	0.06	0.06	0.06	0.06	0.06
Simulation run		S31	S32	S33	S34	S35	S36
Parameter modified		Carrying capacity (-5%)	Carrying capacity (+5%)	Territory size (-5%)	Territory size (+5%)	Relatedness threshold (-5%)	Relatedness threshold
Parameter value used		28.5	31.5	98.8	109.2	0.119	0.131
Number of packs with both a male and a female alpha	29.3 [23.4;35.2]	27.9	30.9	29.4	29.5	29.5	29.5
Number of new packs created	2.1 [1.7;2.5]	**1.4**	**1.6**	2.3	1.9	2.2	2.2

Number of individuals	186.7 [149.4;224]	189.1	211.4	184.5	189.2	187.2	187.1
Proportion of resident individuals	0.71 [0.57;0.85]	0.66	0.66	0.72	0.69	0.7	0.71
Relatedness between the male alpha and the female alpha in a pack	0.06 [0.05;0.07]	**0.05**	**0.05**	0.06	0.06	**0.05**	0.06
Simulation run		S37	S38	S39	S40	S41	S42
Parameter modified		**Probability of budding (-5%)**	**Probability of budding (+5%)**	**Number of immigrants (-5%)**	**Number of immigrants (+5%)**	**Proportion of emigrants (-5%)**	**Proportion of emigrants**
Parameter value used		0.475	0.525	c(0.317, 0.333, 0.35)	c(0.35, 0.333, 0.317)	0.095	0.105
Number of packs with both a male and a female alpha	29.3 [23.4;35.2]	29.5	29.5	29.5	29.4	29.4	29.4
Number of new packs created	2.1 [1.7;2.5]	2.2	2.3	2.4	2.3	2.1	2.4
Number of individuals	186.7 [149.4;224	186.9	187.5	186	188.4	187.7	186.8

	]						
Proportion of resident individuals	0.71 [0.57;0.85]	0.7	0.7	0.71	0.7	0.7	0.7
Relatedness between the male alpha and the female alpha in a pack	0.06 [0.05;0.07]	0.06	0.06	0.06	0.06	0.06	0.06

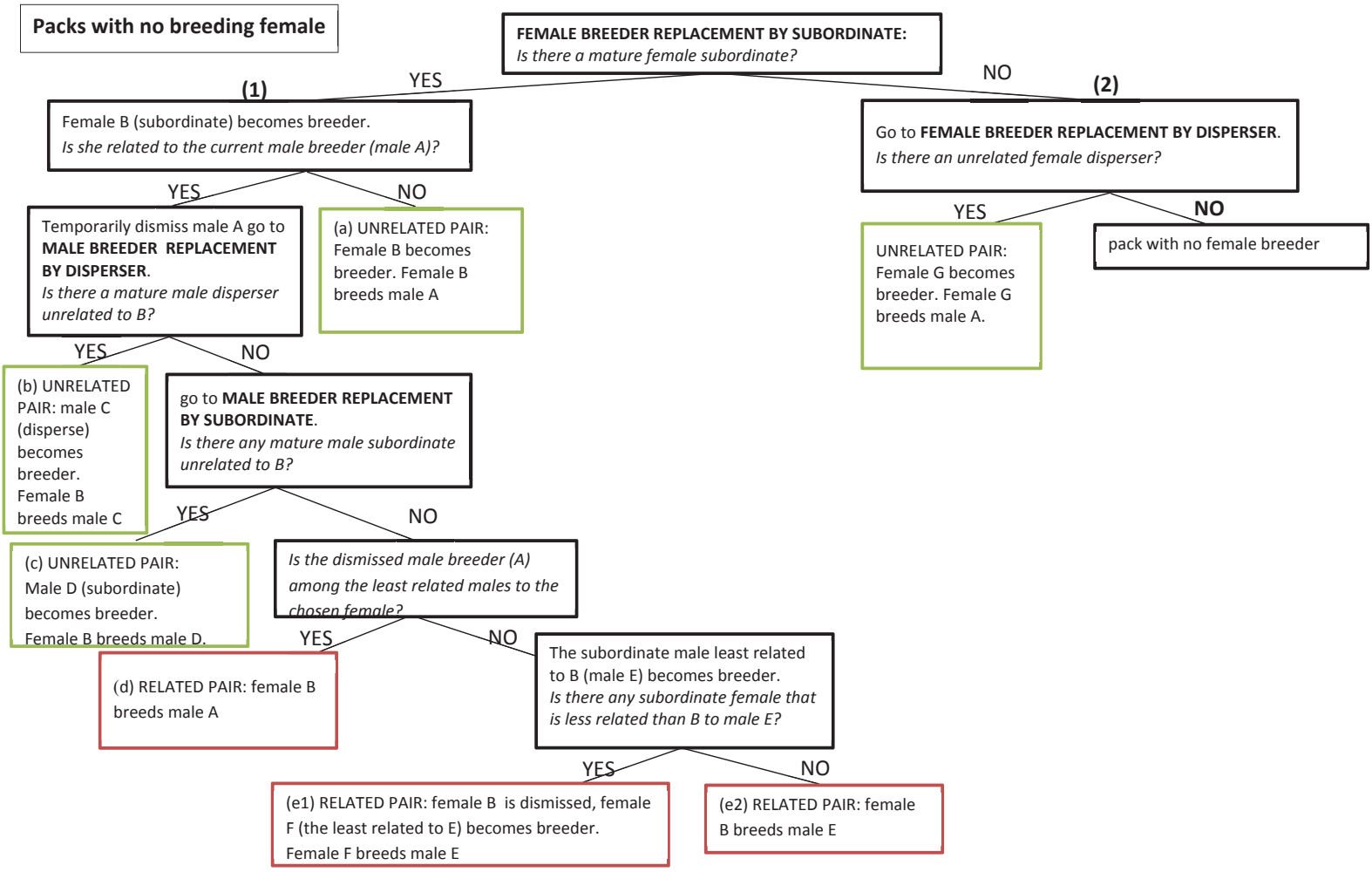

Packs with no breeding female
FEMALE BREEDER REPLACEMENT BY SUBORDINATE:
Is there a mature female subordinate?
YES
NO
(1)
(2)
Female B (subordinate) becomes breeder.
Is she related to the current male breeder (male A)?
Go to FEMALE BREEDER REPLACEMENT BY DISPERSER.
Is there an unrelated female disperser?
YES
NO
YES
NO
Temporarily dismiss male A go to MALE BREEDER REPLACEMENT BY DISPERSER.
Is there a mature male disperser unrelated to B?
(a) UNRELATED PAIR: Female B becomes breeder. Female B breeds male A
UNRELATED PAIR: Female G becomes breeder. Female G breeds male A.
pack with no female breeder
YES
NO
(b) UNRELATED PAIR: male C (disperse) becomes breeder. Female B breeds male C
go to MALE BREEDER REPLACEMENT BY SUBORDINATE.
Is there any mature male subordinate unrelated to B?
YES
NO
(c) UNRELATED PAIR: Male D (subordinate) becomes breeder. Female B breeds male D.
Is the dismissed male breeder (A) among the least related males to the chosen female?
YES
NO
(d) RELATED PAIR: female B breeds male A
The subordinate male least related to B (male E) becomes breeder.
Is there any subordinate female that is less related than B to male E?
YES
NO
(e1) RELATED PAIR: female B is dismissed, female F (the least related to E) becomes breeder. Female F breeds male E
(e2) RELATED PAIR: female B breeds male E

Figure S1. Decision tree for breeders replacement for packs that have only the breeding male (no breeding female). The black squares contain the steps and the criteria that lead to the formation a pair. Green squares indicate the formation of an unrelated pair, red squares indicate the formation of a related pair.

Description of the paths leading to the different ouctcomes for Fig. S1.

If fhe female breeder is missing, the female breeding position is taken over by:

- (1) a subordinate:
    - (a) if the current male breeder is unrelated, the new breeding female and the current breeding male form a pair.
    - (b) if the current male breeder is related, the new breeding female pairs with an unrelated disperser that takes over the breeding position. The former breeding male becomes a subordinate (being an adult subordinate he will have high probabiliy of dispersing in the next year).
    - (c) if the current male breeder is related and mature male dispersers unrelated to the new breeding female are not available, the new breeding female pairs with an unrelated subordinate that takes over the breeding position. The former breeding male becomes a subordinate.
    - (d) If the current breeding male is related, there are not mature dispersers nor mature subordinates that are unrelated to the new breeding female and the current breeding male is among the least related pack members to the new breeding female, the new breeding female and the current breeding male form a related pair.
    - (e) If the current breeding male is related, there are not mature dispersers nor mature subordinates that are unrelated to the new breeding female and the current breeding male is not among the least related pack members to the new breeding female, the least related subordinate takes over the breeding position and:
        - e1. If the new breeding female is among the least related pack members to the new breeding male, the new breeding female and the new breeding male form a pair
        - e2. If the new breeding female is not among among the least related pack members to the new breeding male, the least related female to the new breeding male takes over the breeding position and forms a related pair with the new alpha male.

- (2) If no mature subordinate is available the female breeding position is taken over by
    - (f) an unrelated mature female disperser, the current breeding male and the new breeding female form an unrealted pair.
    - (g) If no unrelated mature disperser is available, the pack remains without a breeding female.

Packs with no breeding male

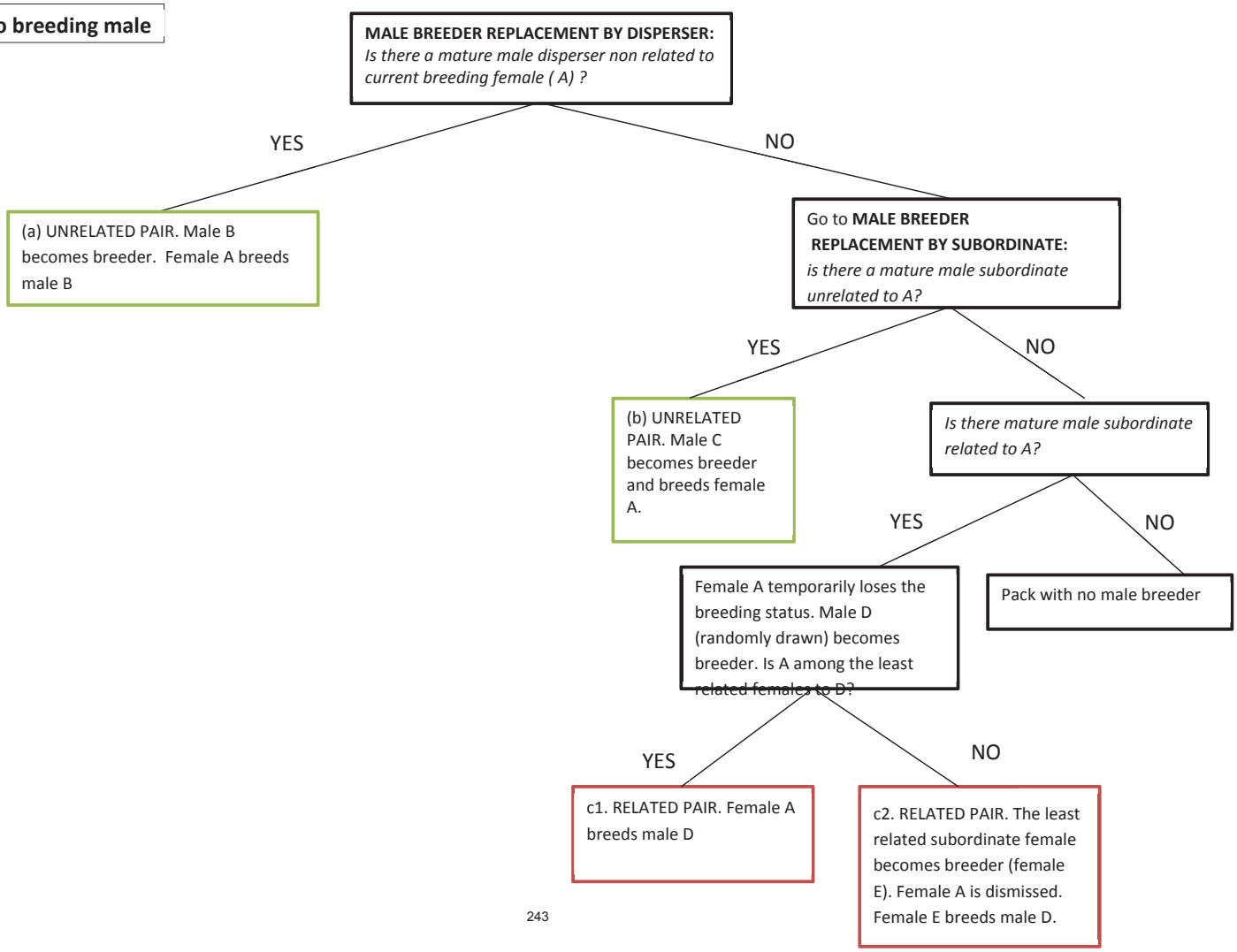


Figure S2. Decision tree for breeders replacement for packs that have only the breeding female (no breeding male). The black squares contain the steps and the criteria that lead to the formation a pair. Green squares indicate the formation of an unrelated pair, red squares indicate the formation of a related pair.

Description of the paths leading to the different ouctcomes for Fig. s2

The male breeder dies. The male breeding position is taken by:

- (1) (a) a disperser unrelated to the current breeding female.
- (2) If no unrelated male disperser is available, by a mature male subordinate
    - (b) unrelated to the breeding female.
    - (c) If no unrelated male subordinate is available The male breeding position is taken by a related subordinate male (randomly chosen)
        - (c1) if the current female breeder is among the least related females to the new breeding male the current breeding females breeds the new breeding male
        - (c2) if the current female breeder is not among the least related females to the new breeding male she looses the breeding status and the female least related to the new breesing male becomes breeder.

Packs with no breeders

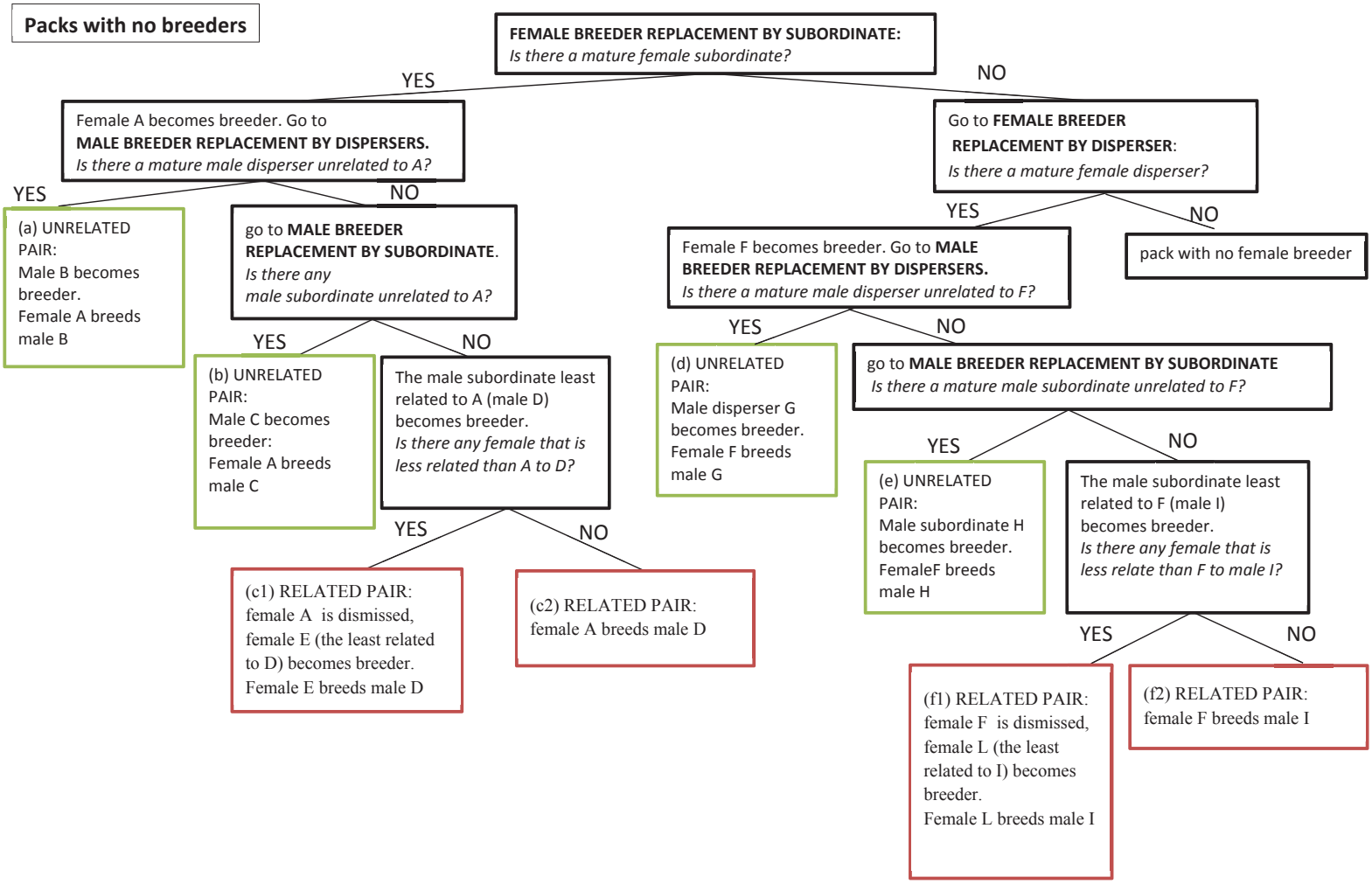


Figure S3. Decision tree for breeders replacement for packs that have no breeders. The black squares contain the steps and the criteria that lead to the formation a pair. Green squares indicate the formation of an unrelated pair, red squares indicate the formation of a related pair.

(1) If the female breeding position is taken over by a subordinate:

- (a) the new breeding female pairs with an unrelated disperser that takes over the breeding position. The former breeding male becomes a subordinate (being an adult subordinate he will have high probabiliy of dispersing in the next year).
- (b) If there are no unrelated dispersers available the new breeding female pairs with an unrelated subordinate that takes over the breeding position. The former breeding male becomes a subordinate.
- (c) If there are no unrelated dispersers available not unrelated subordinates, the least related to the new breeding female will take over the breeding position and:
    - (d1) If the new breeding female is among the least related pack members to the new breeding male, the new breeding female and the new breeding male form a pair
    - (d2). If the new breeding female is not among among the least related pack members to the new breeding male, the least related female to the new breeding male takes over the breeding position and forms a related pair with the new alpha male.

(2) If the female breeding position is taken over by anunrelated mature male disperser:

- (d) the new breeding female and an unrelated male disperser form a pair
- (e) if no male unrelated male disperser is available, the new breeding female and an unrelated male subordinate form a pair
- if no male unrelated male disperser, nor subordinate male are available the male subordinate the least related to the new breeding female takes over the breeding posisions and
    - (f1) If the new breeding female is among the least related pack members to the new breeding male, the new breeding female and the new breeding male form a pair
    - (f2) If the new breeding female is not among among the least related pack members to the new breeding male, the least related female to the new breeding male takes over the breeding position and forms a related pair with the new alpha male.

36

37

38

39

40

6. GENERAL CONCLUSIONS

We developed three complementary modeling approaches that can be used to 1) estimate the abundance of admixed individuals in a population, 2) project hybridization dynamics and 3) evaluate the effectiveness of different management strategies. By applying these models on the case study on wolf x dog hybridization in Italy, we highlighted the important implications of underestimating anthropogenic hybridization in wildlife populations. Moreover, we provided indications and quantitative support for the implementation of proactive and reactive management strategies to mitigate wolf x dog hybridization. These models should be regarded as new tools that are now available for the evaluation and management planning of anthropogenic hybridization cases.

6.1 Improvement in prevalence estimation methods and application to the case study

In the first chapter we demonstrated that the approach that is currently used to assess the extent of hybridization in many case studies (not limited to the wolf x dog case) relies on a biased estimator, naive prevalence, that leads to underestimating the hybridization-related risks for conservation. We provided an alternative and rigorous estimation framework that accounts for the main sources of bias (imperfect detectability and uncertainty in individuals' classification). The comparison of simulated naive and model-based prevalence estimates showed that the first can be severely biased when the probability of detection is state-dependent (admixed vs. parental). In the wolf x dog hybridization case study (Chapter 2) we found support for the presence of state-dependent detectability, although also models with constant parameters ranked high in the model selection. Given the potential bias introduced by differential detectability we recommend that it should be tested for and considered when estimating prevalence in agreement with MacKenzie & Kendall (2002) and Jennelle et al. (2007). In addition, we identified another source of bias in naive prevalence which was related to the state-dependent probability of assignment. However, in our case study (Chapter 2), the models with state-dependent probability of assignment were not

identifiable due to the reduced sample size and we had to discard them, but the consequences of this issue deserve further investigation.

By explicitly considering the uncertainty of classification of the individuals we propose a classification framework that overcomes the subjectivity of choosing between classifying all the uncertain individuals as admixed (i.e., overestimating prevalence) or classifying all the uncertain individuals as parentals (i.e., underestimating prevalence). The use of the Viterbi Algorithm allows to objectively re-classify the uncertain individuals due to the most likely category, given the observed data (Rouan et al. 2009). In our case study (Chapter 2) although our sample was limited, our independent reconstruction of pack genealogy largely supported the performance of the probabilistic attribution of the uncertain individuals.

Lastly, we relied on dataset simulations to provide practical indications for improving the monitoring of prevalence in the case study with the objective of indicating the sample size needed to obtain precise and accurate estimates of prevalence. We concluded that the most accurate estimates are obtained by maximizing the detectability within a minimum of 5 sampling occasions (Chapter 1) and we provided a minimum number of scats to be collected at each occasion and for each wolf pack to obtain a substantial improvement in the precision of the estimates (Chapter 2). The same simulation framework can be adapted to produce sampling indications in other case studies.

6.1.1.Implications for wolf conservation

In the second chapter we applied targeted small-scale sampling and the estimation framework developed in Chapter 1 to the wolf population occupying a study area in the Northern Apennines, Italy and estimated the highest estimate of admixture ever reported, to our knowledge, for a gray wolf population. Surprisingly such high prevalence occurred in what is considered as a core area of the Italian wolf distribution range (Corsi et al. 1999). By detecting first and second generation backcrosses we showed that hybridization between wolves and dogs in the northern Apennines was

not limited to the wolf re-colonization phase (i.e., late 1990s; Galaverni et al. 2017) but it may be recent and recurrent in this study area. Alternatively, high admixture in the study area may have originated also through dispersal of admixed individuals from other areas.

In addition to that, our low estimate of apparent survival raises the concern, supported by previous studies in the Northern Apennines (Caniglia et al. 2014) that human-caused mortality may be impairing the stability of wolf packs, enhancing the reproductive turnover and, in this way, facilitating hybridization and introgression as documented for other hybridization cases involving social canids. Bohling & Waits (2015) investigated hybridization events between red wolves (*Canis rufus*) and coyotes (*Canis latrans*) in eastern North Carolina, USA from 2001 to 2013 and observed that more than half of the hybridization events followed the disruption of social units due to mortality of one or both breeders. Rutledge et al. (2011) analyzed genetic data from eastern wolves (*Canis lycaon*) sampled in Algonquin Provincial Park (APP), Ontario, Canada from 1964 to 2007 showing that eastern wolf culling campaign were followed by a significant increase in coyote mitochondrial and nuclear DNA introgression into the eastern wolf population. Moreover, Benson et al. (2014) showed that higher admixed eastern wolf x coyote survival in non-protected areas adjacent to APP favored high prevalence of admixed individuals. The capture-recapture model with differential survival probability was the second best the model selection in Chapter 2 and although we cannot draw conclusions due to uncertainty in model selection, this aspect deserves further investigation.

Lastly, the occurrence of at least two admixed breeding pairs in the studied wolf population in the Northern Apennines, Italy (Chapter 2), is further evidence of persistent lack of reproductive isolation highlighting that, in conditions of high admixture, hybrid-hybrid pairs can maintain dog genes at high frequency in the wolf population (Bassi et al. 2017; Salvatori et al. 2019) instead of diluting them with backcrossing.

Contrarily to previous knowledge, our findings underline that wolf-dog hybridization is indeed a serious threat for wolf genomic integrity, affecting also core areas of the wolf range and

that, in absence of active management actions, no reproductive barriers nor dilution through backcrossing are sufficient to contrast the spread of the phenomenon.

6.2 Improvement in hybridization projection methods and application to the case study

In this work we produced two approaches for the projection of hybridization dynamics. The population-based approach has the advantage of requiring less parameters. However, the biggest disadvantage is that it cannot easily consider the fine-scale social dynamics that are known to affect hybridization rates in social canids. Moreover, it does not allow to easily distinguish between different hybrid categories (i.e., recent vs. ancient). The individual based model, on the other hand requires a large number of parameters and assumptions, but it allows to explicitly model the interaction between social dynamics and management. We are aware that these models cannot be validated due to the absence of long-term data series on the quantitative evaluation of wolf x dog hybridization. However, the comparison of predicted and observed dynamics in other more intensively monitored hybridization cases (e.g., red wolves-coyote and eastern wolves-coyote) could provide interesting insights. Moreover, rather than being useful for providing absolute predictions, projections are of practical value by illustrating the expected, relative outcome of alternative biological and management scenarios (e.g., Gervasi & Ciucci 2018) that would be otherwise difficult to test empirically on the wolf x dog hybridization case.

6.2.1 Implications for the management of wolf x dog hybridization

Both population-based and individual-based projections results showed that hybridization with dogs should be considered as a serious threat to the wolf genomic integrity, at least under the listed circumstances which are characteristics of human-dominated landscapes (disruption of reproductive isolation barriers and availability of dogs and/or admixed individuals as mates). Both population-based and individual-based projections results confirmed that the outcome of wolf x dog

104 hybridization in absence of reproductive isolation mechanisms (i.e., assortative mating, outbreeding
105 depression) and in presence of recurrent mating with dogs is complete admixture.
106 The sensitivity of the results produced by both models to the different reproductive isolation
107 scenarios highlights that the further investigation of the mechanisms promoting wolf x dog should
108 be addressed as research priority. However, the high prevalence of admixed individuals observed in
109 local studies (i.e., the case study presented in Chapter 2, but also the case study of Salvatori et al.
110 2019) show that at least in some areas characterized by high human pressure natural reproductive
111 isolation and/or dilution through backcrossing are not operating. These populations will therefore
112 head to a complete admixture scenario in the absence of management, as exemplified by the
113 management status of wolf x dog hybrids in Italy. In addition to that, the sensitivity analysis in our
114 population-based model pointed out that the most influential parameters decreasing the chances of
115 complete admixture are linked to social and reproductive integrity (i.e., survival of wolf alpha
116 individuals and the per capita fertility rate) and the annual frequency of mating with dogs. Our
117 results offer further evidence that human-related factors contributing to hybridization (i.e., increased
118 wolf mortality and dog presence) should be urgently managed through preventive and proactive
119 management actions to avoid the risk of genomic extinction of wild wolves (Rutledge et al. 2011;
120 Bohling & Waits 2015).
121 In further support to the need of implementing management, we showed that also in
122 scenarios in which hybridization is predicted to decrease thanks to the action of reproductive
123 barriers (i.e., assortative mating) the impact of hybridization is going to extend at broader scales, as
124 large numbers of admixed dispersers are produced. Those admixed dispersers that manage to breed
125 and reproduce will spread the hybridization geographically and affect other populations at the
126 landscape scale (Salvatori et al. 2019). Reactive management aimed at reducing prevalence of
127 admixed individuals is therefore recommended and should be a joint effort among regional
128 administrations responsible for wolf management across a wider scale (Salvatori et al. 2019).

When introgression is widespread to the extent recorded in some areas (e.g., Chapter 2) we argue that reactive interventions should be considered (Allendorf et al. 2001). Although the removal of admixed individuals is practiced in some countries (e.g., Spain; Pacheco et al. 2017), the sterilization may be the only applicable reactive management strategy when the protection status of admixed individuals is not clear and when considering the social acceptance and animal welfare costs of management. Our modelling approach helped elucidating that a targeted sterilization approach dramatically improves the effectiveness of the sterilization strategy. Such improvement relates to the fact that sterilized admixed breeders act as 'placeholders' preventing other admixed individuals from reaching a breeding position (Gese & Terletzky 2015).

6.3 Future perspectives

Human activities involved in anthropogenic hybridization are predicted to increase (Bohling 2016) and management will be needed to conserve current biodiversity as well as the ecological circumstances and evolutionary processes that support it (Todesco et al. 2016). The analytical approaches developed in this project can be extended for the evaluation and management planning of anthropogenic hybridization involving other species. Non-invasive genetic monitoring is already used for the monitoring of several hybridizing species (e.g., cutthroat trout, *Oncorhynchus clarki*, European wildcats *Felis sylvestris sylvestris*, Greater Spotted Eagles *Aquila clanga*, Iberian wolves *Canis lupus signatus*, red wolves *C. rufus;* Bohling 2016) and it could provide the data for such assessments.

One possible development for the individual based approach is to use it to investigate the relationship between prevalence indexes (e.g., naive prevalence, number of admixed pairs, number of admixed packs) and population prevalence with the approach used by Chapron et al. (2016) to estimate a conversion factor from the number of packs to the total population size. Such approach could help in elaborating less expensive and more feasible prevalence monitoring techniques.

One important development of the wolf x dog hybridization assessment would be to investigate the spatial genetic structure of the distribution of different Canis types (i.e., wolves, dogs and wolf x dog hybrids) and to model genetic ancestry in response to environmental variables (e.g.. prey availability or human disturbance). The combined use of using genetic, morphologic, demographic and behavioral data, would provide valuable information both for understanding biological and evolutionary mechanisms involved in hybridization and for producing sound management strategies (Benson et al. 2012).

In this context, one important extension to the produced projection models would be to make them spatially explicit (Nathan et al. 2019). Such extension would also allow to investigate the consequences of hybridization for niche differentiation and space use (Ellington and Murray 2015; Otis et al 2017). Understanding such processes would help determining if admixed individuals can act as ecological surrogates for the parental individual and if they are therefore worthy of protection (Wayne & Shaffer 2016). Particularly interesting are the recent development in genomics that are increasing our power of investigating the genetic basis of behavioral traits (vonHoldt et al. 2017). Such information could soon produce interesting results regarding the differential behavior between parental and admixed individuals that could be included in individual-based projections to scale up the consequences of individuals' differential behavior to population level processes such as hybridization dynamics.

LITERATURE CITED

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

Benson, J.F., Patterson, B.R. (2013). Inter-specific territoriality in a Canis hybrid zone: spatial segregation between wolves, coyotes, and hybrids. Oecologia 173, 1539-1550.

Benson, J. F., Patterson, B.R., Mahoney, P.J. (2014). A protected area influences genotype-specific survival and the structure of a Canis hybrid zone. Ecology, 95, 254-264.

185 Benson, J.F., Patterson, B.R., Wheeldon, T.J. (2012). Spatial genetic and morphologic structure of
186 wolves and coyotes in relation to environment heterogeneity in a Canis hybrid zone.
187 Molecular Ecology, 21, 5934-5954.
188
189 Bohling, J.H. (2016). Strategies to address the conservation threats posed by hybridization and
190 genetic introgression. Biological Conservation 203, 321-327.
191
192 Bohling, J.H. & Waits, L.P. (2015). Factors influencing red wolf–coyote hybridization in eastern
193 North Carolina, USA. Biological Conservation 184, 108-116.
194
195 Caniglia, R., Fabbri, E., Galaverni, M., Milanesi, P. & Randi, E. (2014). Noninvasive sampling and
196 genetic variability, pack structure, and dynamics in an expanding wolf population. Journal of
197 Mammalogy 9, 41-59.
198
199 Corsi, F., Duprè, E., Boitani, L. (1999). A large-scale model of wolf distribution in Italy for
200 conservation planning. Conservation Biology 13, 150-159.
201
202 Gervasi, V. & Ciucci P. (2018). Demographic projections of the Apennine brown bear population
203 *Ursus arctos marsicanus* (Mammalia: Ursidae) under alternative management scenarios. The
204 European Zoological Journal 85, 242-252.
205 Ellington, E.H. & Murray, D.L. (2015), Influence of hybridization on animal space use: a case study
206 using coyote range expansion. Oikos, 124, 535-542.
207
208 Galaverni, M., Caniglia, R., Fabbri, E., Milanesi, P., Randi, E. (2016). One, no one, or one hundred
209 thousand: how many wolves are there currently in Italy? Mammalian Resrearch 61, 13-24.
210
211 Gese, E. M. & Terletzky, P.A. (2015). Using the 'placeholder' concept to reduce genetic
212 introgression of an endangered carnivore. Biological Conservation 192, 11-19.
213
214 Jennelle, C.S., Cooch, E., Conroy, M.J. & Senar, J.C. (2007). State specific detection probability
215 and disease prevalence. Ecological Applications 17, 154-167.
216
217 MacKenzie, D.I. & Kendall, W.L. (2002). How should detection probability be incorporated into
218 estimates of relative abundance? Ecology 83, 2387-2393.
219
220 Nathan, L.R., Mamoozadehb, N., Tumasc, H.R., Gunselmand, S., Klasse, K., Metcalfef, A., Edgeg,
221 C., et al. (2019). A spatially-explicit, individual-based demogenetic simulation framework for
222 evaluating hybridization dynamics. Ecological Modelling 401, 40-51.
223
224 Otis, J.A., Thornton, D., Rutledge, L., Murray, D.L. (2017). Ecological niche differentiation across
225 a wolf-coyote hybrid zone in eastern North America. Diversity Distribution. 2017; 23: 529-
226 539.
227
228 Pacheco, C., López-bao, J. V., García, E. J., Lema, F. J., Llaneza, L., Palacios, V. & Godinho, R.
229 (2017). Spatial assessment of wolf-dog hybridization in a single breeding period. Scientific
230 Reports 7, 1-10.
231
232 Rouan, L., Gaillard, J.-M., Guédon, Y., Pradel, R. (2009). Estimation of lifetime reproductive
233 success when reproductive status cannot always be assessed. In D. L. Thompson, E. G.
234 Cooch, & M. J. Conroy (Eds.), Modeling demographic processes in marked populations,
235 volume 3 of Environmental and Ecological Statistics (pp. 845–865). New York, NY: Springer

236 series.
237
238 Rutledge, L.Y.,White, B.N., Row, J.R., Patterson, B.R. (2011). Intense harvesting of eastern wolves
239 facilitated hybridization with coyotes. Ecology and Evolution 2, 19-33
240
241 Salvatori, V., Godinho, R., Braschi, C., Boitani, L. & Ciucci, P. (2019). High levels of recent wolf ×
242 dog introgressive hybridization in agricultural landscapes of central Italy. European Journal of
243 Wildlife Research 65, 73-88.
244
245 Wayne, R.K. & Shaffer, H.B. (2016), Hybridization and endangered species protection in the
246 molecular era. Molecular Ecology 25, 2680-2689.
247

248