Uniwersytet Gdański

Faculty of Biology

University of Gdańsk

Roya Adavoudi Jolfaei, MSc

The Role of Hybridization in the Evolutionary Response to Environmental Change in the Genus *Canis*

Rola hybrydyzacji w ewolucyjnej odpowiedzi na zmiany środowiskowe u przedstawicieli rodzaju *Canis*

PhD thesis presented to The Scientific Council of Biological Sciences of the University of Gdańsk in order to obtain a doctoral degree in the field of exact and natural sciences in the discipline of biological sciences

This study was funded by the Polish National Science Centre (grant no. 2019/34/E/NZ8/ 00246 to Małgorzata Pilot)

Supervisor: dr hab. Małgorzata Pilot

Department of Evolutionary Genetics and Biosystematics

GDAŃSK 2025

Acknowledgment

I am deeply grateful for the journey that brought me to the completion of this thesis. What started as a distant dream slowly turned into reality, shaped by learning, challenges, and the incredible support of those around me. Along the way, I experienced unforgettable moments that I will carry with me for the rest of my life, moments of discovery, connection, and personal growth. This achievement is not mine alone; it reflects the encouragement, kindness, and wisdom of many people who walked beside me and made this path meaningful and possible.

First and foremost, I wish to express my heartfelt gratitude to my supervisor, Dr. Małgorzata Pilot. No words can truly express how grateful I am for her unwavering support, patience, and guidance throughout this journey. I have learned so much from her, not only through her scientific expertise but also through her quiet strength, resilience, and the way she leads with kindness and integrity. Dear Małgorzata, thank you for your belief in me, your endless patience, and the invaluable lessons you've taught me. I am truly honored to have been your student.

I am also deeply grateful to all the co-authors who contributed to this study by generously providing samples from across Eurasia. First, I would like to thank Prof. Wiesław Bogdanowicz for helping to organize sample collection across the region. I would also like to express my sincere appreciation to Dr. Milomir Stefanović, Dr. Karolina Doan, Dr. Alejandro Flores-Manzanero, Francelly Martinez-Sosa, Dr. Yellapu Srinivas, Dr. Elena Bykova, Dr. Ovidiu C. Banea, Dr. Duško Ćirović, Dr. Silviu Chiriac, Dr. Gianluca D'Amico, Dr. Mihajla Djan, Dr. Giorgos Giannatos, Dr. Sergey N Gashev, Dr. Jennifer Hatlauf, Dr. Vahram Hayrapetyan, Dr. Miklós Heltai, Dr. Kanstantsin Homel, Dr. Pavel Hulva, Dr. Angela Monica Ionică, Dr. Yadvendradev Vikramsinh Jhala, Dr. Jana Juránková, Dr. Mohammad Kaboli, Dr. Rasoul Khosravi, Dr. Natia Kopaliani, Dr. Rafał Kowalczyk, Dr. Miha Krofel, Dr. József Lanszki, Dr. Luca Lapini, Dr. Petros Lymberakis, Dr. Nikolay V Mamaev, Dr. Peep Männil, Dr. Georgi Markov, Dr. Sarah Marshall-Pescini, Dr. Andrei Daniel Mihalca, Dr. Anastasia Miliou, Dr. David Modrý, Dr. Vladislav Molchan, Dr. Carmela Musto, Dr. Mikhael Nikoforov, Prof. Henryk Okarma, Dr. Innokentiy M Okhlopkov, Prof. Massimo Scandura, Dr. Stéphane Ostrowski, Dr. Giedrė Pakeltytė, Dr. Dainis Edgars Ruņģis, Dr. Dragana Šnjegota, Dr.

László Szabó, Dr. George A. Tryfonopoulos, Dr. Elena Tsingarska, Dr. Anatoliy M. Volokh, Dr. Jan M. Wójcik, Dr. Maria Yakovleva, Dr. Alesya Lischyshyna, and all other people who generously provided samples and data. Your collaboration and contributions were vital in enabling me to conduct a comprehensive study. Thank you for being part of this work.

I gratefully acknowledge the financial support provided by the Polish National Science Centre (grant no. 2019/34/E/NZ8/ 00246 to Małgorzata Pilot).

I want to sincerely thank my husband, Rasoul, who truly believed in me and encouraged me to follow my dreams without ever letting me feel limited as a woman. His love, patience, and support have carried me through some of the toughest moments of this journey. Yet, no matter the distance, he was always just a call away, ready with words of encouragement, a listening ear, and quiet reassurances that gave me the courage to keep going when I felt overwhelmed or uncertain. Rasoul, your love means everything to me, and I could not have reached this without you by my side.

I am deeply thankful to my only sister, Roshanak, whose unwavering support and encouragement lifted me up during my darkest and most difficult moments. Your presence has been a bright light throughout this journey. I am truly and profoundly blessed to have you not only as my sister but as my closest friend and source of inspiration.

I would also like to thank my dear friends Morgane, Karolina, Francelly, Milomir, and Alejandro. Your friendship has been one of the most beautiful gifts during my time in Poland. Thank you for turning even the most difficult days into ones filled with laughter, comfort, and joy. Though our paths may now take us in different directions, the memories we created together will stay with me always. I feel incredibly lucky to have shared this journey with you.

Finally, I would like to sincerely thank everyone who has supported me, whether through advice, encouragement, or simply believing in me, your help has meant more than words can express.

Streszczenie

Niniejsza praca bada wzorce hybrydyzacji w obrębie gatunków z rodzaju *Canis*, obejmując wolno-żyjące psy (*Canis lupus familiaris*), wilki szare (*Canis lupus*) oraz szakale złociste (*Canis aureus*) na obszarze Eurazji. W oparciu o dane obejmujące polimorfizmy pojedynczych nukleotydów w skali całego genomu zbadano czynniki ekologiczne i ewolucyjne kształtujące wzorce hybrydyzacji pomiędzy tymi gatunkami. Choć hybrydyzacja pomiędzy psami domowymi a wilkami była już wcześniej badana, włączenie do analiz szakala złocistego pozwala uzyskać nowe informacje na temat wpływu odległości ewolucyjnej między krzyżującymi się gatunkami na częstość hybrydyzacji i introgresji.

Niniejsza praca przedstawia kompleksową analizę hybrydyzacji w obrębie rodzaju *Canis*, mającą na celu głębsze zrozumienie, w jaki sposób hybrydyzacja działa nie tylko jako źródło przepływu genów między gatunkami, lecz także jako mechanizm lokalnej adaptacji i zmian ewolucyjnych. W tym celu praca realizuje pięć głównych zadań: (1) zbadanie możliwych konsekwencji hybrydyzacji i jej wpływu na gatunki rodzicielskie, (2) znalezienie najlepszego sposobu szacowania proporcji wariantów pochodzących z hybrydyzacji u wilków, szakali i wolno-żyjących psów, (3) oszacowanie częstości hybrydyzacji między badanymi taksonami przy użyciu wybranej metody, (4) ocena roli hybrydyzacji w adaptacji gatunków poprzez identyfikację fragmentów chromosomów podlegających introgresji adaptacyjnej i analizę funkcji genów w nich zawartych, (5) zbadanie wpływu zmiennych środowiskowych na proporcje wariantów pochodzących z hybrydyzacji oraz częstość introgresji adaptacyjnej w populacjach psowatych. Cele te zostały zrealizowane w czterech rozdziałach pracy.

Rozdział 1 przedstawia systematyczny przegląd literatury dotyczącej skutków hybrydyzacji w różnych rzędach i rodzinach ssaków. Nasze wyniki pokazują, że negatywne konsekwencje hybrydyzacji, takie jak wypieranie rodzimego genotypu (występujące w 21% analizowanych prac) czy introgresja alleli pochodzących od zwierząt udomowionych (18%), są opisywane w literaturze znacznie częściej niż pozytywne skutki, takie jak nabycie nowej, adaptacyjnej zmienności genetycznej (8%). Przewaga negatywnych skutków w literaturze może wynikać z faktu, że wiele badań opiera się na neutralnych markerach genetycznych, które mają ograniczoną zdolność do wykrywania bardziej złożonych procesów, takich jak introgresja adaptacyjna czy specjacja hybrydowa. Z tego względu połączenie analiz loci neutralnych i markerów

zlokalizowanych w regionach kodujących może zapewnić pełniejszy i bardziej zrównoważony obraz zjawiska hybrydyzacji, uwzględniający zarówno jego negatywne skutki, jak i korzyści adaptacyjne oraz czynniki determinujące te efekty.

Rozdział 2 realizuje drugie zadanie poprzez ocenę metod rekonstrukcji pochodzenia osobników na poziomie globalnym (tzn. w całych genomach) i lokalnym (w obrębie chromosomów) w kontekście analizy introgresywnej hybrydyzacji pomiędzy trzema gatunkami psowatych, w oparciu o polimorfizmy pojedynczych nukleotydów w skali całego genomu. Wyniki pokazują, że metody analizy hybrydyzacji na poziomie globalnym (np. PCA, ADMIXTURE) zwykle zawyżają poziom hybrydyzacji w porównaniu z metodami lokalnymi (np. LAMP-LD, ELAI, Ghap). Różnice te mogą wynikać z odmiennych założeń metodologicznych, różnic w rodzaju analizowanych danych genetycznych oraz podejścia do brakujących danych. Zidentyfikowano dwa główne czynniki – niską jakość genotypów i genetyczną strukturę populacji – które mogą zwiększać niepewność i zmienność wyników między metodami. Wykazano, że metody globalne, takie jak ADMIXTURE, są bardziej podatne na te zakłócenia. Dlatego zalecamy łączne stosowanie metod lokalnych i globalnych, przy czym wynikom lokalnej analizy pochodzenia powinno się dawać priorytet w celu precyzyjnego oszacowania poziomu introgresji.

Rozdział 3 odnosi się do trzeciego i czwartego zadania, badając konsekwencje ewolucyjne hybrydyzacji u psowatych. Bazując na ustaleniach metodologicznych z poprzedniego rozdziału oraz wysokiej skuteczności metody ELAI w szacowaniu proporcji mieszanego pochodzenia osobników, oszacowano poziomy introgresji między gatunkami. Wyniki wykazały, że hybrydyzacja w obrębie rodzaju *Canis* jest powszechnym zjawiskiem na wielu obszarach ich występowania. Szczególnie wysoką częstość hybrydyzacji stwierdzono na Bałkanach, w Indiach, na Kaukazie oraz w północno-wschodniej Europie – co może być związane ze znacznymi zaburzeniami antropogenicznymi, dużą liczebnością wolno-żyjących psów oraz ekspansją zasięgu szakala złocistego. Zaobserwowaliśmy także wpływ odległości ewolucyjnej między gatunkami na częstość introgresji –udział wariantów pochodzących z introgresji od psów był większy u wilków niż u szakali (6,4% vs. 1,2%). Konsekwencje ewolucyjne hybrydyzacji wskazują, że zarówno dzikie psowate, jak i psy wolno żyjące mogą odnosić korzyści z tego procesu. Introgresja adaptacyjna może umożliwiać dzikim gatunkom nabywanie od psów alleli zwiększających przystosowanie, m.in. takich, które wzmacniają układ odpornościowy. Może to zwiększać odporność na nowe patogeny,

szczególnie w środowiskach, gdzie kontakt między dzikimi psowatymi a psami jest częsty, a patogeny dynamicznie się zmieniają. Z kolei wolno-żyjące psy nabywają więcej korzystnych wariantów genetycznych od wilków, co może wpływać na ich cechy morfologiczne, behawioralne i fizjologiczne. Obok sygnałów pozytywnej selekcji, stwierdzono również ślady selekcji negatywnej w blokach chromosomowych o obniżonym poziomie introgresji u psów i szakali. Wskazuje to, że niektóre warianty genetyczne mogą być dla nich szkodliwe, ale są skutecznie eliminowane z ich puli genowej. Podsumowując, ta praca podkreśla złożony charakter hybrydyzacji i introgresji jako procesów ewolucyjnych, które mogą wprowadzać zarówno korzystne, jak i szkodliwe warianty genetyczne.

Rozdział 4 realizuje ostatnie zadanie pracy poprzez zastosowanie analizy Random Forest (RF) oraz Redundancy Analysis (RDA) w celu identyfikacji kluczowych czynników środowiskowych wpływających na częstość występowania wariantów genetycznych pochodzących od psów w genomach dzikich psowatych. U wilków udział tych wariantów był dodatnio skorelowany z łagodniejszymi zimami, natomiast u szakali obserwowano odwrotną zależność. Obszary o łagodniejszym klimacie mogą sprzyjać większej liczebności wolno-żyjących psów, zwiększając prawdopodobieństwo kontaktu i krzyżowania się z wilkami. Z kolei niższe roczne temperatury mogą skłaniać szakale do zbliżania się do ludzkich osiedli w poszukiwaniu pożywienia, zwiększając prawdopodobieństwo interakcji z psami i tym samym częstość hybrydyzacji. Ponadto, stwierdzono pozytywną korelację między udziałem wariantów genetycznych pochodzących od psów u dzikich psowatych a śladem działalności człowieka (ang. *human footprint*). Obszary silnie przekształcone przez człowieka charakteryzują się większą liczebnością wolno-żyjących psów, co zwiększa szansę kontaktu z dzikimi psowatymi. Dodatkowo wykazano istotne powiązania między loci pochodzącymi od psów będacymi pod wpływem introgresji adaptacyjnej u wilków a zmiennymi środowiskowymi. Geny objęte introgresją adaptacyjną były związane z układem nerwowym, układem odpornościowym oraz metabolizmem. Wyniki te podkreślają znaczenie introgresji adaptacyjnej, która może pomóc wilkom lepiej przystosować się do środowisk zmodyfikowanych przez człowieka, gdzie mogą być narażone na nowe patogeny, stresory środowiskowe i zmiany w dostępności pokarmu. Podkreśla to również znaczenie hybrydyzacji jako aktywnego procesu ewolucyjnego, który może odgrywać większą rolę w adaptacji niż wcześniej sądzono.

Podsumowując, praca ta przedstawia kompleksową analizę czynników ekologicznych i konsekwencji ewolucyjnych hybrydyzacji między dzikimi psowatymi a psami wolno żyjącymi. Zastosowanie podejścia integrującego analizy genomowe, testy doboru naturalnego i analizy środowiskowe pozwoliło na pogłębione zrozumienie hybrydyzacji nie tylko jako źródła przepływu genów między gatunkami, ale również jako potencjalnego mechanizmu lokalnej adaptacji. Wyniki tej pracy uwypuklają złożony charakter hybrydyzacji i introgresji w procesach ewolucyjnych. Chociaż warianty pod wpływem introgresji adaptacyjnej mogą wspierać lokalne dostosowanie, hybrydyzacja może również wprowadzać warianty szkodliwe, zakłócając lokalnie dostosowane kompleksy genów lub zwiększając podatność na choroby i inne stresory. Wyniki te mają praktyczne znaczenie dla zarządzania populacjami dzikich zwierząt i ochrony przyrody. Kluczowe jest rozróżnianie przypadków, w których hybrydyzacja wspiera potencjał adaptacyjny gatunków, od tych, w których zagraża ich integralności genetycznej – szczególnie w odniesieniu do taksonów obejmujących formy udomowione, takich jak rodzaj *Canis*.

Summary

This thesis investigates hybridization patterns among *Canis* species, including free-ranging dogs (*Canis lupus familiaris*), Eurasian gray wolves (*Canis lupus*), and golden jackals (*Canis aureus*), across Eurasia. The genome-wide single nucleotide polymorphisms (SNP) data was applied to explore the ecological and evolutionary factors shaping hybridization patterns among these species. While hybridization between domestic dogs and grey wolves has been previously studied, including the golden jackal provides new insights into the effect of the evolutionary distance between the cross-breeding species on hybridization and introgression rates.

This thesis presents a comprehensive analysis of hybridization within the genus *Canis*, aiming to achieve a deeper understanding of how hybridization operates not only as a source of interspecific gene flow but also as a mechanism for local adaptation and evolutionary change. For this purpose, the thesis addresses five key objectives: (1) investigate possible consequences of hybridization and its impact on the parental species, (2) find the best method for estimating proportions of hybridization-derived variants in gray wolves, golden jackals, and free-ranging dogs, (3) estimate the rate of hybridization between the studied taxa using this method, (4) assess the role of hybridization in species adaptation by identifying adaptive introgressed chromosomal fragments and assessing the functions of genes included in these fragments, (5) test the effect of environmental variables on the proportions of hybridization-derived variants and adaptive introgression rates in canid populations. These objectives are addressed in four chapters of the thesis.

Chapter 1 presents a systematic literature review of studies that have reported the consequences of hybridization across various mammalian orders and families. Our results showed that negative consequences of hybridization, like genetic swamping (reported in 21% of studies) and introgression of variants from domestic animals (reported in 18% of studies), have been reported in the literature more frequently compared to the positive consequences, like gaining novel adaptive variation (reported in 8% of studies). The predominance of negative outcomes reported in the literature can be explained by the fact that many studies are based on neutral genetic markers, which are limited in detecting complex processes like adaptive introgression or hybrid speciation. Therefore, integrating both neutral loci and markers located in coding regions can provide a more comprehensive and balanced understanding of hybridization,

capturing not only its potential negative consequences but also its adaptive benefits and the underlying factors that shape these outcomes.

Chapter 2 addressed the second objective by evaluating the methods of individual ancestry reconstruction at the global (i.e. in the entire genomes) and local levels (within chromosomes) in the context of the analysis of introgressive hybridization among the three canids, based on genome-wide SNP data. The results revealed that global ancestry methods (e.g., PCA and ADMIXTURE) generally estimated higher hybridisation levels than local ancestry methods (e.g., LAMP-LD, ELAI, and Ghap). The inconsistency between the results may result from differences in their methodological frameworks, the types of genetic information they utilize, and their strategies for handling missing data. Two key factors, low-quality genotypes and subpopulation structure, were identified as major factors that can contribute to increasing uncertainty and variability between methods. We found that global ancestry analyses such as ADMIXTURE are more likely to be affected by these confounding factors. Therefore, we recommend a joint use of local and global methods, with results of local ancestry analysis being prioritized for precise estimation of introgression levels.

Chapter 3 addressed the third and fourth objectives by exploring the evolutionary consequences of hybridization in canids. Based on the methodological knowledge from the last chapter and the robust performance of ELAI in estimating individual ancestry proportions, the introgression rate among species was estimated using ELAI. Our results showed that hybridization in the genus *Canis* is common in their distribution range. In some regions, including the Balkans, India, the Caucasus, and northeastern Europe, a higher frequency of hybridization was found, which may have resulted from high anthropogenic disturbances, large population size of free-ranging dogs, and the range expansion of golden jackals. We also clearly show the effect of evolutionary distances between the species on introgression rates between them, since a higher frequency of dog introgression was found in wolves compared to golden jackals (6.4% vs. 1.2%). The evolutionary consequences of hybridization showed that both wild canids and free-ranging dogs may gain benefits from hybridization. Adaptive introgression may enable wild canids to acquire from dogs gene variants conferring adaptive advantage, including those that strengthen their immune systems. These beneficial genes may increase the resistance of wild canids to new pathogens, which would be particularly beneficial in environments where wild canids encounter dogs frequently and where pathogens are constantly evolving. Free-ranging dogs appear to have acquired a larger pool of

beneficial genetic variants from wolves, which may have contributed to some characteristics like morphological, behavioural, and physiological traits. Alongside detecting signals of positive selection, we also found signatures of negative selection in chromosomal blocks with reduced introgression levels in dogs and golden jackals. These results suggest that some introgressed gene variants may also have a deleterious effect on these species, but they can be efficiently removed from their gene pools. Overall, we highlight the complex nature of hybridization and introgression in the evolutionary process, showing that it can introduce both beneficial and maladaptive genetic variation.

Chapter 4 addresses the last objective by performing the Random Forest (RF) analysis and Redundancy analysis (RDA) to identify the key environmental factors that may contribute to the frequency of dog-derived genetic variants in wild canids. In wolves, the frequency of such variants showed positive association with milder winters, while in golden jackals, a reverse trend was observed. Regions with milder winters likely support larger free-ranging dog populations, increasing the likelihood of contact and interbreeding between dogs and wolves. Conversely, in golden jackals, lower annual temperatures may drive individuals toward human settlements in search of food, increasing interactions with domestic dogs and thereby hybridization rates. Additionally, our results showed that frequency of dog-derived genetic variants in wild canids is positively correlated with human footprint. Regions with high human disturbance often have a greater abundance of free-ranging dogs, increasing the likelihood of encounters with wild canids. Furthermore, we found a significant association between dog-derived loci under adaptive introgression in wolves and environmental factors. The genes under adaptive introgression were associated with the nervous system, immune system, and metabolism. These results emphasize the role of adaptive introgression, which can help wolves to better adapt to human-modified environments, where wolves may encounter new pathogens, environmental stressors, and dietary shifts. We highlighted the role of hybridization as an active evolutionary process, possibly being more important for adaptation than previously believed.

Overall, this thesis presents a comprehensive investigation into the ecological drivers and evolutionary consequences of hybridization between wild canids and free-ranging dogs. Using an integrative approach that combined genomic analyses, selection scans, and environmental associations, we achieved a deeper understanding of how hybridization operates not only as a source of interspecific gene flow but also as a mechanism for local adaptation. The results of this study highlight the complex nature of

hybridization and introgression in the evolutionary process. While variants under adaptive introgression may enhance local adaptation of species, hybridisation may also introduce deleterious variants, potentially disrupting locally adapted gene complexes or increasing vulnerability to disease and other stressors. The results of this study have practical implications for wildlife management and conservation. Recognizing when hybridization contributes to adaptive potential and when it threatens species integrity will be critical for informed decision-making in conservation genetics, particularly for taxa that include domesticated lineages, such as the genus *Canis.*

Table of Contents

List of Figures

List of Tables

General Introduction

Hybridization (the process of cross-breeding between two genetically distinct taxa) has a significant effect on many evolutionary processes such as speciation and adaptation (Harrison and Larson 2014; Chan et al., 2019; Elworth et al., 2019). Evolutionary biologists have been interested in interspecific hybridization since Darwin first described it in the context of speciation (Darwin, 1859). The term "introgression" (or "introgressive hybridization") was first introduced by Anderson and Hubricht in 1938. They described introgression as a widespread process of adaptive significance, in which alleles from one species are transferred into the gene pool of another species (Anderson and Hubricht 1938; Anderson, 1949). The theoretical framework explaining the evolutionary importance of introgression was later developed by Barton (2001). However, until 1990, less than 50 articles were published each year about interspecific hybridization (Schwenk et al., 2008). New molecular laboratory techniques such as Polymerase Chain Reaction (PCR) led to major advances in studies of hybridization (Dowling and Secor, 1997). Over the past decade, advances in genome sequencing resulted in genomic studies on a large number of non-model organisms, which demonstrated that hybridization is ubiquitous in nature and occurs in a broad range of taxa such as plants (Goulet et al., 2017), butterflies (Mallet et al., 2007), fishes (Porto-Foresti et al., 2008), birds (Ottenburghs, 2023), and mammals (Moroni et al., 2022). Many studies documented that ecosystem structure could be influenced by the introgression of genetic variants among different taxa (Payseur & Rieseberg, 2016; McFarlane and Pemberton, 2019; Edelman and Mallet, 2021).

The first step to answering questions regarding the impact of hybridization is accurately identifying hybrid individuals (McFarlane and Pemberton, 2019). Historically, introgression between species was inferred based on morphological characteristics, as intermediate phenotypes were often observed in hybrid individuals (Anderson and Hubricht, 1938; Anderson, 1948). However, since phenotypic traits can be influenced by environmental conditions, and in some cases the parental species display limited phenotypic differentiation, identifying admixed individuals using morphological traits can be problematic (Rieseberg et al., 1998; Randi, 2008). Hybrid individuals display distinct genome-wide patterns, which facilitates detection of first-generation hybrids and backcrosses using genetic data compared to morphological characteristics (Nason and Ellstrand, 1993; Anderson and Thompson, 2002; Mallet, 2005; Rogers & Bernatchez, 2007). In early 2000, some methods such as Multi Locus Sequence Typing (MLST), were the common frameworks to identify introgression, however, these methods were rapidly replaced by other approaches using a larger number of loci (e.g. AFLPs and microsatellites) (Dagilis et al., 2021). Although with the use of these genetic markers it was possible to identify the first-generation hybrids (F1) and early-generation backcrosses, detecting ancient backcrosses, multigenerational admixture, and fully understanding the extent of hybridization within populations was impossible. Because in later generations of backcrosses only a small proportion of loci retains the signature of hybridization, a large number of loci is needed to detect historical or ancient hybridization events (Taylor and Larson, 2019; Hibbins and Hahn, 2022). Over the past decade, the rapid advances in next-generation sequencing technologies

have resulted in the development of powerful methods to identify introgression between species. The whole genome sequencing and the genotyping of genome-wide single-nucleotide polymorphisms (SNPs) allowed researchers to study hybridization and introgression in more details (e.g., quantifying global admixture as well as identifying putatively introgressed regions in each chromosome).

During the past 20 years, many studies have documented the evolutionary consequences of hybridization (Grant, 1981; Arnold, 1997; Arnold, 2006; Arnold et al. 2008; Abbott et al., 2013; Arnold, 2015; Abbott et al., 2016; Sankararaman et al., 2016). Among different consequences, great attention has been paid to the possible negative impacts (e.g. genetic swamping and outbreeding depression), however, genome sequencing and genome-wide SNP markers provide an opportunity to assess both the negative and positive consequences of hybridization (e.g. adaptation and increase in genetic variation) (Edelman and Mallet, 2021). Due to rapid human-induced environmental changes, hybridization could potentially provide novel adaptive variations and reduce the risk of extinction through increasing genetic variation and creating new genetic combinations, especially in small and fragmented populations (Chan et al., 2019). Therefore, knowledge of various possible consequences of hybridization may lead to improved success in biodiversity conservation.

Hybridization between domestic animals and their wild relatives has been well-documented (McFarlane & Pemberton, 2019). In such cases, the chance of hybridization is higher due to the lack of fully developed reproductive barriers. An example of the negative consequences of hybridization between domestic and wild taxa can be observed in wildcats (*Felis silvestris silvestris*). The gene pool of Scottish wildcat is at serious risk of extinction due to the extensive genetic swamping caused by hybridization with domestic cat (*Felis catus*) (Howard-McCombe et al., 2023). The term 'genomically extinct' was used for this species as no pure individuals were identified in the study (Howard-McCombe et al., 2023).

Hybridization between species in the genus *Canis* is a notable example of this process in mammals (e.g. Randi, 2008; Galov et al., 2015; Pilot et al., 2018; Stronen et al., 2022; Stefanović et al., 2024). Although species from this genus are genetically distinct, they can produce fertile offspring, which can back-cross to their parental populations (Leonard et al., 2013). This natural hybridization provides an opportunity to study the ecological, evolutionary, and genetic/genomic consequences of hybridization in nature.

Genus *Canis* is an ideal model species for assessing different consequences of hybridization because of many reasons: (1) Since domestic dogs (*Canis lupus familiaris*) have adapted well to human-modified landscapes, investigating hybridization between this species and its wild relatives, such as the golden jackal (*Canis aureus*) and the grey wolf (*Canis lupus*), offers an ideal framework for studying anthropogenic introgression and evaluating the human influence on the genetic integrity of wild species; (2) species from the genus *Canis* occupy diverse ecological niches, ranging from anthropogenic to natural habitats, which makes them an ideal model to assess the influence of environmental conditions on the rate of hybridization and understand how changes in habitat may affect hybridization rates across different ecosystems; (3) Since hybridization can significantly influence the adaptive evolution of species (Barton,

2001), studying hybridization in various canid species offers a valuable opportunity to explore how it introduces new genetic variation that may facilitate adaptation to changing environments; (4) As hybridization between wild canids and domestic dogs can threaten the genetic integrity of wild species (Hindrikson et al., 2017; Donfrancesco et al., 2019), studying hybridization within this genus can help develop strategies to manage and conserve threatened species that are at risk of genetic swamping because of hybridization with other canids.

Therefore, studying genus *Canis* hybrids can help address broader questions in evolutionary biology, including the mechanisms of speciation, the impact of hybridization on genetic variation, and the effects of human disturbance on wildlife.

The aims of the study

This thesis presents a comprehensive study of hybridization within the genus *Canis*, aiming to answer key questions about its evolutionary consequences. The main aims of the study were addressed in four chapters by combining genetic (SNPs) and spatial data (environmental variables);

Investigating different possible consequences of hybridization

The first chapter establishes the context for the thesis by synthesizing existing research on hybridization in mammals and aims to broaden the theoretical understanding of its evolutionary consequences. While many studies have mostly focused on documenting whether hybridization occurs, much less is known about its consequences for the species. This chapter addresses that gap through a systematic review, with two main goals: (1) estimate the frequency of hybridization in different mammalian orders and families, and (2) evaluate its consequences (positive and negative) for parental species.

Identifying the best approach to estimating hybridization rate

In the second Chapter, local (i.e. chromosome-level) and global (genome-wide) ancestry inference methods were applied to estimate introgression rates in wolves, golden jackals, and free-ranging dogs. This chapter aims to: (1) assess the consistency between the results obtained from different methods, (2) compare the admixture proportions in the same individuals obtained from the analysis of the entire dataset versus populations from two different geographic regions (two local datasets), and (3) identify factors that can confound the results of local and global ancestry analyses.

Estimating the rate of hybridization between the studied taxa and assessing the role of hybridization in species adaptation

Building on the findings from Chapter 2, this chapter uses the most reliable method identified to investigate the evolutionary outcomes of hybridization in wild canids and free-ranging dogs. The central aim is to explore how hybridization contributes to adaptation through the introgression of beneficial gene variants. Specifically, this chapter addresses the following objectives: (1) identify genomic blocks showing overrepresentation of introgressed variants in wild canids and free-ranging dogs, (2) assess how the degree of genetic differentiation between species can influence the

proportion of introgressed gene variants under selection, (3) identify genes within adaptive introgressed blocks that are under positive selection, and (4) determine the functions of these positively selected gene variants and their potential benefits for admixed individuals.

Assessing the effect of environmental variables on introgression and adaptive introgression rates

Human disturbance is usually considered a main factor that may contribute to facilitating hybridization between wild canids and free-ranging dogs (e.g., Godinho et al., 2011; Lescureux and Linnell, 2014; Pilot et al., 2021), however, this relationship has never been proven, and there is still a lack of direct empirical evidence confirming this link. In this chapter, we address this gap by testing the hypothesis that introgression rate in admixed individuals are associated with environmental variables. The main aims of this chapter were to: (1) identify the key environmental factors that may be associated with hybrid ancestry in wild canids, (2) detect adaptive introgressed loci associated with environmental variables, and (3) identify genes within adaptive introgressed regions that are linked to environmental factors.

Bibliography

Abbott, R. J., Barton, N. H., & Good, J. M. (2016). Genomics of hybridization and its evolutionary consequences. Molecular Ecology.

Abbott, R., Albach, D., Ansell, S., Arntzen, J. W., Baird, S. J., Bierne, N., ... & Zinner, D. (2013). Hybridization and speciation. Journal of Evolutionary Biology, 26(2), 229-246.

Anderson, E. (1948). Hybridization of the habitat. Evolution, 1-9.

Anderson, E. C., & Dunham, K. K. (2008). The influence of family groups on inferences made with the program Structure. Molecular Ecology Resources, 8(6), 1219-1229.

Anderson, E. C., & Thompson, E. (2002). A model-based method for identifying species hybrids using multilocus genetic data. Genetics, 160(3), 1217-1229.

Anderson, E., & Hubricht, L. (1938). Hybridization in Tradescantia. III. The evidence for introgressive hybridization. American Journal of Botany, 396-402

Aria, M., Alterisio, A., Scandurra, A., Pinelli, C., & D'Aniello, B. (2021). The scholar's best friend: Research trends in dog cognitive and behavioral studies. Animal Cognition, 24(3), 541-553.

Arnold ML (1997) Natural Hybridization and Evolution. Oxford University Press, Oxford, UK.

Arnold ML (2006) Evolution Through Genetic Exchange. Oxford University Press, Oxford, UK.

Arnold ML (2015) Divergence with Genetic Exchange. Oxford University Press, Oxford, UK

Arnold, M. L., Sapir, Y., & Martin, N. H. (2008). Genetic exchange and the origin of adaptations: prokaryotes to primates. Philosophical Transactions of the Royal Society B: Biological Sciences, 363(1505), 2813-2820.

Barton, N. H. (2001). The role of hybridization in evolution. Molecular Ecology, 10(3), 551-568.

Chan, W. Y., Hoffmann, A. A., & van Oppen, M. J. (2019). Hybridization as a conservation management tool. Conservation Letters, 12(5), e12652.

Clay, L. et al. In defense of canine behavioral assessments in shelters: Outlining their positive applications. Journal of Veterinary Behavior. 38, 74–81 (2020).

Clay, L., Paterson, M. B., Bennett, P., Perry, G., & Phillips, C. C. (2020). Do behaviour assessments in a shelter predict the behaviour of dogs post-adoption?. Animals, 10(7), 1225.

Cordoni, G., & Palagi, E. (2019). Back to the future: A glance over wolf social behavior to understand dog–human relationship. Animals, 9(11), 991.

Dagilis, A. J., Peede, D., Coughlan, J. M., Jofre, G. I., D'Agostino, E. R., Mavengere, H., ... & Matute, D. R. (2021). 15 years of introgression studies: quantifying gene flow across Eukaryotes. BioRxiv, 2021-06.

Darwin, C. 1859 The origin of species by means of natural selection or the preservation of favoured races in the struggle for life. London, UK: John Murray.

Donfrancesco, V., Ciucci, P., Salvatori, V., Benson, D., Andersen, L. W., Bassi, E., ... & Mukherjee, N. (2019). Unravelling the scientific debate on how to address wolf-dog hybridization in Europe. Frontiers in Ecology and Evolution, 7, 175.

Dowling, T. E., & Secor, C. L. (1997). The role of hybridization and introgression in the diversification of animals. Annual review of Ecology and Systematics, 28(1), 593-619.

Dziech, A. (2021). Identification of wolf-dog hybrids in Europe–an overview of genetic studies. Frontiers in Ecology and Evolution, 9, 760160.

Edelman, N. B., & Mallet, J. (2021). Prevalence and adaptive impact of introgression. Annual Review of Genetics, 55, 265-283.

Elworth, R. L., Ogilvie, H. A., Zhu, J., & Nakhleh, L. (2019). Advances in computational methods for phylogenetic networks in the presence of hybridization. Bioinformatics and phylogenetics: seminal contributions of Bernard Moret, 317-360.

Farhat, N., Lazebnik, T., Monteny, J., Moons, C. P. H., Wydooghe, E., van der Linden, D., & Zamansky, A. (2023). Digitally-enhanced dog behavioral testing. Scientific Reports, 13(1), 21252.

Galov, A., Fabbri, E., Caniglia, R., Arbanasić, H., Lapalombella, S., Florijančić, T., ... & Randi, E. (2015). First evidence of hybridization between golden jackal (*Canis aureus*) and domestic dog (*Canis familiaris*) as revealed by genetic markers. Royal Society Open Science, 2(12), 150450.

Geza, E., Mugo, J., Mulder, N. J., Wonkam, A., Chimusa, E. R., & Mazandu, G. K. (2019). A comprehensive survey of models for dissecting local ancestry deconvolution in human genome. Briefings in bioinformatics, 20(5), 1709-1724.

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*(24), 5154-5166.

Goulet, B. E., Roda, F., & Hopkins, R. (2017). Hybridization in plants: old ideas, new techniques. Plant Physiology, 173(1), 65-78.

Harrison, R. G., & Larson, E. L. (2014). Hybridization, introgression, and the nature of species boundaries. Journal of Heredity, 105(S1), 795-809.

Hibbins, M. S., & Hahn, M. W. (2022). Phylogenomic approaches to detecting and characterizing introgression. Genetics, 220(2), iyab173.

Hindrikson, M., Remm, J., Pilot, M., Godinho, R., Stronen, A. V., Baltrūnaité, L., ... & Saarma, U. (2017). Wolf population genetics in E urope: a systematic review, meta-analysis and suggestions for conservation and management. Biological Reviews, 92(3), 1601-1629.

Howard-McCombe, J., Jamieson, A., Carmagnini, A., Russo, I. R. M., Ghazali, M., Campbell, R., ... & Beaumont, M. A. (2023). Genetic swamping of the critically endangered Scottish wildcat was recent and accelerated by disease. Current Biology, 33(21), 4761-4769.

Leonard, J. A., Echegaray, J., Randi, E., Vilà, C., & Gompper, M. E. (2013). Impact of hybridization with domestic dogs on the conservation of wild canids. Free-ranging Dogs and Wildlife Conservation, 170.

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

Mallet, J. (2005). Hybridization as an invasion of the genome. Trends in Ecology & Evolution, 20(5), 229-237.

Mallet, J., Beltrán, M., Neukirchen, W., & Linares, M. (2007). Natural hybridization in heliconiine butterflies: the species boundary as a continuum. BMC Evolutionary Biology, 7, 1-16.

McFarlane, S. E., & Pemberton, J. M. (2019). Detecting the true extent of introgression during anthropogenic hybridization. Trends in Ecology & Evolution, 34(4), 315-326.

Moroni, B., Brambilla, A., Rossi, L., Meneguz, P. G., Bassano, B., & Tizzani, P. (2022). Hybridization between Alpine Ibex and Domestic Goat in the Alps: A Sporadic and Localized Phenomenon?. Animals, 12(6), 751.

Nason, J. D., & Ellstrand, N. C. (1993). Estimating the frequencies of genetically distinct classes of individuals in hybridized populations. Journal of Heredity, 84(1), 1-12.

Netto, W. J., & Planta, D. J. (1997). Behavioural testing for aggression in the domestic dog. Applied Animal Behaviour Science, 52(3-4), 243-263.

Ottenburghs, J. (2023). How common is hybridization in birds?. Journal of Ornithology, 164(4), 913-920.

Payseur, B. A., & Rieseberg, L. H. (2016). A genomic perspective on hybridization and speciation. Molecular Ecology, 25(11), 2337-2360.

Pilot, M., Greco, C., vonHoldt, B. M., Randi, E., Jędrzejewski, W., Sidorovich, V. E., ... & 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(5), 662-680.

Pilot, M., Moura, A. E., Okhlopkov, I. M., Mamaev, N. V., Manaseryan, N. H., Hayrapetyan, V., ... & Bogdanowicz, W. (2021). Human-modified canids in human-modified landscapes: The evolutionary consequences of hybridization for grey wolves and free-ranging domestic dogs. Evolutionary Applications, 14(10), 2433-2456.

Porto-Foresti, F., Hashimoto, D. T., Prado, F. D., Senhorini, J. A., & Foresti, F. (2013). Genetic markers for the identification of hybrids among catfish species of the family P imelodidae. Journal of Applied Ichthyology, 29(3), 643-647.

Randi, E. (2008). Detecting hybridization between wild species and their domesticated relatives. Molecular Ecology, 17(1), 285-293.

Randi, E., Hulva, P., Fabbri, E., Galaverni, M., Galov, A., Kusak, J., ... & Caniglia, R. (2014). Multilocus detection of wolf x dog hybridization in Italy, and guidelines for marker selection. PLoS One, 9(1), e86409.

Rieseberg, L. H., & Carney, S. E. (1998). Plant hybridization. The New Phytologist, 140(4), 599-624.

Rogers, S. M., Isabel, N., & Bernatchez, L. (2007). Linkage maps of the dwarf and normal lake whitefish (*Coregonus clupeaformis*) species complex and their hybrids reveal the genetic architecture of population divergence. Genetics, 175(1), 375-398.

Sankararaman, S., Mallick, S., Patterson, N., & Reich, D. (2016). The combined landscape of Denisovan and Neanderthal ancestry in present-day humans. Current Biology, 26(9), 1241-1247.

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: Biological Sciences, 363(1505), 2805-2811.

Stefanović, M., Bogdanowicz, W., Adavoudi, R., Martínez-Sosa, F., Doan, K., Flores-Manzanero, A., ... & Pilot, M. (2024). Range-wide phylogeography of the golden jackals (*Canis aureus*) reveals multiple sources of recent spatial expansion and admixture with dogs at the expansion front. Biological Conservation, 290, 110448.

Stronen, A. V., Mattucci, F., Fabbri, E., Galaverni, M., Cocchiararo, B., Nowak, C., ... & Caniglia, R. (2022). A reduced SNP panel to trace gene flow across southern European wolf populations and detect hybridization with other Canis taxa. Scientific Reports, 12(1), 4195.

Taylor, S. A., & Larson, E. L. (2019). Insights from genomes into the evolutionary importance and prevalence of hybridization in nature. Nature Ecology & Evolution, 3(2), 170-177.

1 Chapter 1

Consequences of Hybridization in Mammals: A Systematic Review

This chapter has been published as: Adavoudi, R., & Pilot, M. (2021). Consequences of hybridization in mammals: A systematic review. *Genes*, *13*(1), 50.

genes

MDPI

Review

Consequences of Hybridization in Mammals: A Systematic Review

Roya Adavoudi and Małgorzata Pilot *

Museum and Institute of Zoology, Polish Academy of Sciences, ul. Nadwiślańska 108, 80-680 Gdańsk, Poland; radavoudi@miiz.waw.pl
* Correspondence: mpilot@miiz.waw.pl

Abstract: Hybridization, defined as breeding between two distinct taxonomic units, can have an important effect on the evolutionary patterns in cross-breeding taxa. Although interspecific hybridization has frequently been considered as a maladaptive process, which threatens species genetic integrity and survival via genetic swamping and outbreeding depression, in some cases hybridization can introduce novel adaptive variation and increase fitness. Most studies to date focused on documenting hybridization events and analyzing their causes, while relatively little is known about the consequences of hybridization and its impact on the parental species. To address this knowledge gap, we conducted a systematic review of studies on hybridization in mammals published in 2010–2021, and identified 115 relevant studies. Of 13 categories of hybridization consequences described in these studies, the most common negative consequence (21% of studies) was genetic swamping and the most common positive consequence (8%) was the gain of novel adaptive variation. The total frequency of negative consequences (49%) was higher than positive (13%) and neutral (38%) consequences. These frequencies are biased by the detection possibilities of microsatellite loci, the most common genetic markers used in the papers assessed. As negative outcomes are typically easier to demonstrate than positive ones (e.g., extinction vs hybrid speciation), they may be over-represented in publications. Transition towards genomic studies involving both neutral and adaptive variation will provide a better insight into the real impacts of hybridization.

Keywords: adaptive introgression; genetic swamping; hybridization; hybrid speciation; hybrid zones; outbreeding depression; mammals

Citation: Adavoudi, R.; Pilot, M. Consequences of Hybridization in Mammals: A Systematic Review. *Genes* **2022**, *13*, 50. https://doi.org/10.3390/genes13010050

Academic Editor: Arne Ludwig

Received: 12 November 2021
Accepted: 20 December 2021
Published: 24 December 2021

1. Introduction

Until recently, hybridization was considered as a rare phenomenon in the animal kingdom and thus its role in animal evolution has been underappreciated [1]. Growing amount of whole genome sequence data produced in recent years made it possible to demonstrate that a broad range of animal species have experienced hybridization events [2]. Although hybridization frequency (i.e., the proportion of individuals interbreeding with another species and producing hybrid offspring) is low in most species [3], it has been estimated that around 25% of plant species and 10% of animal species have been affected by hybridization [4]. Hybridization is most frequent among recently diverged sister species, which are frequently characterized with incompletely developed prezygotic isolation (behavioural and mechanical) and postzygotic isolation (zygote mortality and hybrid inviability and sterility) [5]. Hybridization is now recognized as a widespread phenomenon with significant impact on species evolution [6,7] and potentially serious consequences for species conservation and management [8].

Cross-breeding between species with limited postzygotic isolation can often lead to an intake of genetic variation typical of one species into another species' gene pool—a process called introgression. Introgressive hybridization can affect creation, maintenance and loss of biodiversity [9]. In some cases, introgression may facilitate species evolutionary

responses to environmental changes by promoting rapid acquisition of new adaptive genetic variants [10,11], thus increasing adaptive potential of these species [12–15]. Moreover, introgressive hybridization can contribute to speciation by creating new genetic variation and functional novelty [6,16]. Introgression of alleles from one species into another one can create novel adaptive combinations of alleles and form a new admixed population, which is genetically distinct from both parental populations [17]. Such population may develop reproductive isolation from the parental populations and thus maintain distinctiveness even in a contact zone [18]. However, speciation through hybridization occurs less frequently in mammals compared to other vertebrates, because reproductive barriers between mammalian species are in general well established [19]. In species with low genetic variation, introgressive hybridization could increase genetic variation and decrease inbreeding, without any signs of outbreeding depression [20,21].

On the other hand, hybridization can be also considered as a potential threat to species survival [22–24]. Accumulation of deleterious variation [25], outbreeding depression [26–28] and genetic swamping [29–31] are among detrimental consequences of hybridization. In extreme cases, severe outbreeding depression and decline in the population growth rate below the replacement rate due to wasted reproductive effort in one or both parental lineages may lead to extinction [32]. High risk of extinction due to hybridization has been reported for rare or endangered species interbreeding with more common relatives [33,34]. However, hybridization was mentioned as a factor contributing to extinction in only 11 species out of 120,369 extinct species assessed [23]. The negative impact of hybridization should not be neglected, but the conservation policies should not be focused on the negative aspects of hybridization only [23]. From a conservation perspective, hybridization outcomes may range from considerable introgression with significant negative impacts, e.g., reduced survival or reproductive success of hybrids, through minimal introgression with negligible impact, to moderate introgression with significant positive impacts, e.g., increased fitness of admixed individuals [35]. Given than hybridization can represent either a threat to species survival or a potential pathway to evolutionary rescue, it is important to examine its impacts case-by-case [35].

Hybridization may be particularly common in widespread, abundant species and in non-indigenous species that were intentionally or unintentionally introduced into a new habitat by humans [32,36]. Hybridization has also been frequently reported for domesticated species and their wild relatives, e.g., wild boar and domestic pig [20,37], gray wolf and domestic dog [38–43], wild cat and domestic cat [44–47]. In such cases hybridization may lead to the introgression of gene variants typical for domestic animals into gene pools of wild species [48,49]. This may have a range of negative consequences, such as the loss of specific adaptations [41] and reduced viability [50]. These negative consequences are particularly pronounced in small, fragmented and isolated populations [51]. Moreover, introgressive hybridization can also affect feral populations of domesticated animals [52,53]. Cross-breeding between individuals originating from captive-bred populations and their wild conspecifics may have similar consequences as that between domesticated and wild populations [54].

In recent decades, several review papers on hybridization have been published. They were focused on specific aspects of hybridization and/or particular taxonomic groups, for example the evolutionary importance of natural hybridization [2], the role of hybridization in extinction [32], hybrid fitness [55], introgression during anthropogenic hybridization [56], mammalian hybrid zones [57], taxonomic problems associated with inter-specific gene flow [58], hybridization in European ungulates [59] and hybridization in New Zealand taxa [60]. Most studies to date focused on documenting hybridization events and analyzing their causes, while relatively little is known about the consequences of hybridization and its impact on the parental species [23]. To address this knowledge gap, we conducted a systematic review of studies on hybridization in mammals and assessed the frequency of different consequences of hybridization reported. In addition, we evaluated the contribution of different mammalian orders and families in published studies on hybridization. We selected

mammals as the focal class because of the large number of available studies, resulting in part from the profound role of species from this class in ecosystem functioning [61].

2. Materials and Methods

We focused on papers on hybridization in mammals published between 2010 and 2021. The database search for papers published in 2021 was completed on the 3rd of December of that year, so papers published after that date are excluded from the results. For finding relevant papers in the Web of Science, we employed the following string: ("hybridization*" OR "hybridisation*" OR "outbreeding*" OR "outcrossing" OR "admixture*" OR "admixed individual*" OR "hybrid zone" OR "hybrid individual$*" OR "backcrosse$") AND ("mammal$*" OR "vertebrate$" OR "consequence" OR "implication" OR "Extinction" OR "genetic swamping" OR "adaptive introgression" OR "hybrid speciation" OR "outbreeding depression") NOT ("protein$" OR "fish$" OR "plant$" OR "invertebrate$" OR "avian reptile$" OR "non avian reptile$" OR "fung$" OR "bird$" OR "Lizard$" OR "penguin$" OR "turtle$" OR "insect$" OR "frog$" OR "butterfl$" OR "homoploid" OR "moth$" OR "salamander$"). With these keywords we found limited numbers of relevant papers (49 papers), we therefore applied different sets of keywords: ("hybridization*" OR "hybridisation*" OR "hybrid$*" OR "outbreeding" OR "outcrossing" OR "admixture*" OR "introgression*" OR "admixed individual *"OR "hybrid zone" OR "hybrid individual$*" OR "backcrosse$") AND ("mammal$ *") NOT ("protein$" OR "cell$" OR "fish$" OR "plant$" OR "invertebrate$" OR "avian reptile$" OR "non avian reptile$" OR "fung$" OR "Cell$" OR "bird$" OR "Lizard$" OR "penguin$" OR "turtle$" OR "insect$" OR "frog$" OR "butterfl$" OR "homoploid" OR "moth$" OR "salamander$"). We combined the search results based on these two sets of keywords.

We excluded from the search books, review papers, theses, annuals or meeting reports. We also excluded papers published in journals that were outside of the following categories: Evolutionary Biology, Genetics and Heredity, Ecology, Biology, Zoology, Biodiversity Conservation, Multidisciplinary Science and Biochemistry and Molecular Biology. Abstracts of the papers that were identified after applying this automatic exclusion (793 and 641, respectively) were read and these papers that were not focused on the hybridization process in mammals were removed. We also removed one of the two copies of papers that overlapped between the paper sets resulting from the search with each set of keywords (Figure 1). We then combined the results from the two sets of papers, including only the papers meeting the following criteria: experimental studies, focused on mammalian species, that investigate hybridization among different species and subspecies. Studies that evaluated admixture among populations of the same species were removed from the analysis, with the exception of wild and domesticated forms that have been classified as the same species (e.g., wild cat *Felis silvestris silvestris* and domestic cat *Felis silvestris catus*), as well as wild vs captive/farmed populations of the same species. Although we designed the search terms to be as comprehensive as possible in the detection of studies on mammalian hybridization, some relevant studies could have been missed. Nevertheless, the resulting set of papers is free of human bias (except for the choice of keywords) and therefore provides a reliable overview of the current knowledge on the consequences of hybridization.

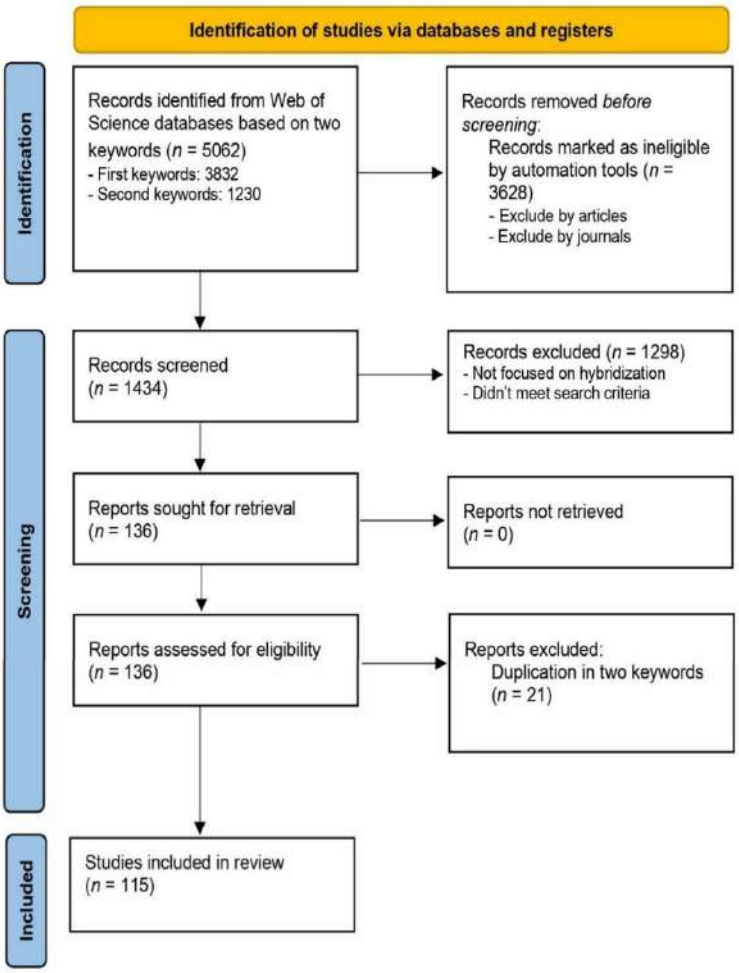


Figure 1. Flow diagram showing the selection stages of studies to be included in the review based on the two sets of keywords used.

3. Results

According to the inclusion criteria defined above, a total of 49 and 87 published papers were selected based on the first and second set of keywords. This was a small subset of papers initially identified with each set of keywords; we tried multiple versions of keywords and did not manage to improve the accuracy of results. After excluding overlapping papers between the two sets of keywords (15.4%), 115 papers were retained for the subsequent analysis (Figure 1). These papers are listed in Supplementary Table S1.

We calculated the frequency of mammalian orders and families involved in hybridization process in the surveyed studies (Figure 2). 39.13% of these studies focused on the order Carnivora (45 papers), of which Canidae (24.34%, 28 papers), Felidae (6.95%, eight papers), and Mustelidae (6.08%, seven papers) were the most-studied families. Order Cetartiodactyla was the second most-studied order (26.95%, 31 papers), in which Cervidae (7.82%, nine papers), Bovidae (6.95%, eight papers) and Suidae (5.21%, six papers), were the most-studied families. The third most represented order was Rodentia (15.65%, 18 papers), in which Cricetidae (8.69%, 10 papers), Sciuridae (5.21%, six papers) and Muridae (1.75%, two papers) had the largest contribution. Other orders were represented by only two families (Chiroptera, 5.21%, six papers and Diprotodontia 2.60%, three papers) or one family (Primates (3.40%, four papers), Lagomorpha (4.34%, five papers), Macroscelidea (0.86%, one study), Perissodactyla (0.86%, one study) and Soricomorpha (0.86%, one study).

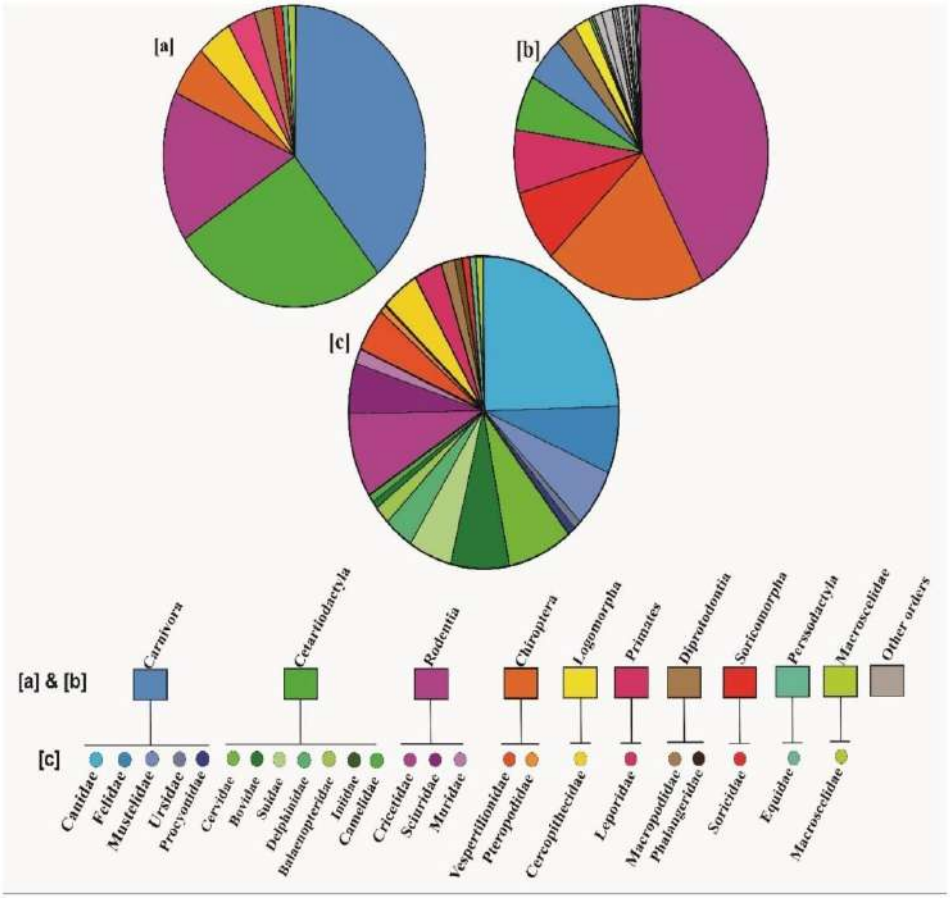


Figure 2. Relative frequencies of (**a**) orders represented in 115 papers included in the systematic review, (**b**) orders represented across all currently recognized mammalian species [62] and (**c**) families represented in 115 papers included in the systematic review.

We compared the frequencies of species representing different mammalian orders in the surveyed studies on hybridization with the frequencies of all currently recognized contemporary mammalian species [62] representing different mammalian orders. This comparison demonstrated a considerable bias in the number of studies on hybridization focused on representatives of different mammalian orders (Figure 2a,b).

We classified the hybridization outcomes described in the 115 papers included in the systematic review into 13 categories (Table 1). These categories were non-exclusive and in some cases the consequences of hybridization reported could be classified to more than one category. Among the 115 surveyed papers, 10 papers did not provide sufficient information to classify them in any category, e.g., [63,64]. Therefore, this classification is based on 105 papers. The reported frequencies of different hybridization outcomes in these papers should not be considered as reliable estimates of the real frequencies, because of the biases discussed below. We classified the impacts of each possible outcome as (1) positive (e.g., gaining novel adaptive variation), (2) neutral or unknown, (3) negative (e.g., extinction, loss of reproductive output) and (4) considered as negative. In this last category we included genetic swamping and introgression from a domesticated lineage, which are frequently described as negative in the literature [29,30,65–67]. Given that direct empirical evidence for their negative effects is limited, we did not classify them as unequivocally negative outcomes. It is important to note that extinction due to extreme genetic swamping is classified in a separate category, given its clearly negative impact.

Table 1. Outcomes of hybridization described in papers included in the systematic review. The outcomes are grouped by the character of their impact. The reported frequencies of different hybridization outcomes in the papers studied should not be considered as reliable estimates of the real frequencies.

	Results	Impacts	Number of Papers	Percentage	Description
1	Genetic swamping	Considered as negative	29	20.71	Genetic integrity of a species involved in hybridization being threatened by introgression from another species
2	Introgression from a domesticated lineage	Considered as negative	25	17.85	Genetic integrity of wild species being threatened by introgression from a domesticated lineage
3	Extinction due to extreme genetic swamping	Negative	3	2.14	Complete loss of genetic material of one of the species involved in hybridization
4	Outbreeding depression	Negative	7	5.0	Reduction or loss of specific adaptations and overall fitness
5	Morphological anomalies	Negative	2	1.4	Morphological anomalies with negative effects on fitness
6	Loss of reproductive output	Negative	3	2.14	Decrease in growth rate of parental species because of wasted reproductive effort
7	Increase in genetic diversity and reduction of inbreeding	Positive	3	2.14	Increase in genetic diversity via low rates of introgression, without any evidence of outbreeding depression; reduction of inbreeding levels
8	Gaining novel adaptive variation	Positive	11	7.85	Transferring of adaptive variants through hybridization
9	Hybrid speciation	Positive	4	2.85	Creation of a new species via hybridization
10	Intermediate phenotypic traits	Neutral or unknown	10	7.14	Intermediate morphological characteristics of hybrid individuals relative to the parental species
11	Hybrid zone	Neutral or unknown	14	10.0	Geographically restricted zones where genetically distinct species meet and mate
12	Hybridization without significant impacts	Neutral or unknown	5	3.57	Evidence of hybridization without substantial changes in the gene pools of each species
13	No or rare evidence of hybridization	Neutral or unknown	24	17.14	Hybridization is rare and does not result in introgression

Of 115 studies considered, 21 (18.26%) identified hybrids using microsatellite loci as the only genetic markers, 18 studies (15.65%) used mtDNA fragments, 35 (30.43%) studies used both microsatellite loci and mtDNA, 12 studies (10.43%) used genome-wide single nucleotide polymorphisms (SNPs) or whole genome sequencing, two studies (1.73%) used all these three types of markers, and the remaining studies used another method of hybrid detection. Altogether, 50.42% of studies used microsatellite loci as either the only type of genetic markers or together with other types.

To assess whether the analysis of genome-wide data may affect the type of hybridization outcomes observed, we calculated the frequencies of different outcomes based on 14 papers that used genome-wide SNPs or whole genome sequencing. As more than half of these papers (eight studies) focused on hybridization between domestic animals and their wild relatives, introgression from a domesticated lineage was the most common negative effect (36% of studies), followed by genetic swamping (18%). Novel adaptive variation was the only positive impact of hybridization and was reported in five studies (23%). In total, negative outcomes were reported more frequently than positive ones (54% and 23%, respectively). The frequencies of both negative and positive outcomes were higher than in the entire dataset (49% and 13%), while the frequency of neutral outcomes was smaller. However, these frequencies should be treated with caution due to the small number of studies and overrepresentation of hybridization with domestic animals among the studies considered.

4. Discussion

4.1. Hybridization in Mammalian Orders and Families

4.1.1. Mammalian Orders

The frequency of different mammalian orders in the studies included in this systematic review does not reflect the number of species within each order. Representatives of orders Carnivora and Cetartiodactyla prevail among the species studied, with the frequencies of 39% and 27%, respectively. The frequency of species from these orders among all mammalian species are 5% and 6%, respectively [62]. In contrast, two most species-rich mammalian orders, Rodentia (42% of species) and Chiroptera (21% of species) were represented in only 16% and 5% of studies, respectively. In studies of hybrid zones, rodents have been represented more frequently than other mammalian orders, but nevertheless only eight rodent genera have been studied, as reported in a review paper [57]. Therefore, the underrepresentation of rodents in hybridization studies may result from the focus on well-studied genera only, such as e.g., the genus *Mus*. Several studies published in the previous decade (i.e., not considered in this systematic review) detected signatures of hybridization in several bat species e.g., [68–71], but altogether hybridization was reported for less than 20 of over 1000 known bat species. This may be associated with a limited number of studies on hybridization in bats [72] or a stronger reproductive isolation in bat species compared with other mammals [70,73]. As bats can form mixed-species groups during mating seasons [74] and during maternal care [75], reproductive barriers are particularly important for the maintenance of species distinctiveness.

Accordingly, overrepresentation of orders Carnivora and Cetartiodactyla in hybridization studies may result from more relaxed reproductive barriers between congeneric species from these orders compared with other mammals, overrepresentation of studies focused on these orders, or a combination of both. High interest in studying species from these orders may result from their important roles in ecosystems and in some cases their high commercial value. Large species from the order Carnivora are keystone species in their ecosystems and are frequently subject to active management and conservation strategies [76–79]. Accordingly, many representatives of Artiodactyla are valuable game species. Moreover, species from Carnivora and Cetartiodactyla orders can compete with humans over resources by consuming game species, livestock depredation and fisheries depredation, as well as damaging crops and wild vegetation [78,80]. For these reasons, they are of particular interest to wildlife researchers, also in the context of hybridization.

In theory, the proportion of species within each mammalian order that are subject to introgressive hybridization could be used as a measure of the strength of reproductive barriers between species within each order. However, to achieve a reliable comparison between mammalian orders, several sources of bias would have to be accounted for, including the above-mentioned differences in intensity of research on different mammalian orders as well as differences in criteria used to define species. Therefore, in practice the frequency of detected hybridization cases is not a reliable measure of the strength of reproductive isolation.

4.1.2. Mammalian Families

The frequency of families within each mammalian order that are subject to hybridization studies is biased as well. For example, the Canidae family is represented in 62% of studies on the order Carnivora and 24% of all 115 studies on mammalian hybridization assessed in this systematic review. The second most frequently represented family within Carnivora is the Felidae family, represented in 18% of studies on this order. Canidae and Felidae are the only families within the order Carnivora that include domesticated species, the domestic dog (*Canis lupus familiaris*) and the domestic cat (*F. s. catus*), respectively. Among the studies included in this systematic review, most (68%) of studies on the Canidae family were focused on hybridization between gray wolves and domestic dogs [38,81–84]. Accordingly, seven out of eight studies (87%) on the Felidae family were focused on hybridization between domestic cats and wild cats [85,86]. Due to a recent

origin of domestic animals, their hybrids with the wild relatives are fertile and thus can back-cross into parental populations [87–89]. Moreover, global populations of domestic dogs and cats have been increasing with human population growth, and the majority of individuals globally are free-ranging and thus can breed freely [90]. Widespread occurrence of free-ranging domestic dogs and cats may have promoted their interactions with their wild relatives and as a result increased the rate of hybridization between them [81,91,92]. The presence of a domesticated species within a particular family and order may be thus an important factor increasing the hybridization rate.

Within Artiodactyla, Cervidae, Bovidae and Suidae were the most studied families. All hybridization cases described in the family Suidae were between the wild boar (*Sus scrofa*) and domestic pig (*Sus scrofa domesticus*). Although free-roaming domestic pigs are rare, hybridization with wild boars may occur in open domestic boar farms [93]. Cross-breeding with wild boars is also used intentionally by humans to obtain less aggressive and larger-sized animals, and to increase growth rate of offspring [93]. Climate change, low frequency of predators, supplementary feeding, reforestation of agricultural areas and intentional releases for hunting have led to the range expansion of the wild boar, which as a result has become one of the most widespread large mammals in the world and the second most frequent ungulate in Europe [94–97]. In many regions, the wild boar has been considered as a pest species for croplands [93,98]. One hypothesis for the vast distribution of wild boars is that introgression from domestic pigs could have led to their increased fitness and invasiveness [99,100]. Hybridization between domestic yaks (*Bos grunniens*) and wild yaks (*Bos mutus*) [101] (Bovidae) is spatially more restricted, given geographically restricted ranges of the wild species, but similarly as in the case of pig–wild boar hybridization, it occurs as both a spontaneous admixture and intentional cross-breeding by humans. Overall, 36% of studies (41 papers) considered in this review were focused on hybridization between domestic animals and their wild relatives, suggesting that the presence of domesticated forms within a family facilitates hybridization.

The effect of human activities on hybridization has long been known [102], and domestication is one of many anthropogenic factors that may increase the frequency of hybridization. The introduction of invasive species to distribution ranges of closely related species may have a similar effect [36,103]. Together with habitat fragmentation and destruction, introduced species are an important threat to global biodiversity [104–107]. Many wild ungulates are valuable game species, and therefore are strongly affected by humans by extensive translocations and introductions of non-native species, hunting, and artificial management; all these factors contribute to hybridization within ungulate families [108].

In particular, the Cervidae family includes multiple valuable game species. One of them, the sika deer (*Cervus nippon*), was deliberately introduced to many European countries for hunting [109], which has led to hybridization with native deer species in some regions [110–112]. Another cervid, European roe deer (*Capreolus capreolus*), is known to hybridize with Siberian roe deer (*Capreolus pygargus*) [15,113] and Italian roe deer (*Capreolus c. italicus*) [114]. Although natural processes (e.g., range expansion) could have caused hybridization in this genus, human-mediated introductions of Siberian roe deer, aimed at increasing body mass and trophy size of European roe deer, affected the rate of hybridization between these species [15]. In Bovidae family, hybridization was reported between Tatra chamois (*Rupicapra rupicapra tatrica*) and introduced Alpine chamois (*Rupicapra rupicapra rupicapra*) [21]. In that case, the introduction was carried out for conservation purposes.

In cetacean species, hybridization has been documented both in captive breeding sites and in the wild [115,116], with around 20% of species known to hybridize [117]. Cross-breeding was shown to be more common between species that have similar morphological and behavioral traits [117–119], and to be facilitated by population fragmentation [120–122]. Although until recently it was believed that hybridization in cetaceans is a dead-end process, as most known hybrids seemed to be infertile [120], a study on hybridization between fin whale (*Balaenoptera physalus*) and blue whale (*Balaenoptera musculus*) showed that the hybrid individuals can reproduce and survive to adulthood in specific circumstances [123].

4.2. Typical Outcomes of Hybridization between Mammalian Species

Depending on species and environmental conditions, hybridization may have either negative or positive impacts, and sometimes there may be very limited consequences. In some cases, hybridization can drive species toward extinction, while in others it provides an opportunity to create new species [124]. Genetic swamping, outbreeding depression, introgression of variants originating from domesticated lineages, and morphological anomalies are typically associated with negative consequences, such as loss of adaptive variation [26,125–130], high mortality rates [131], and even extinction [30]. However, hybridization can be considered as a beneficial process in some circumstances. Introgression from a closely related species may facilitate adaptation by providing novel adaptive variation; this may be particularly important when a population occupies a sub-optimal, poor-quality habitat, expands to a new habitat, or experiences rapid changes in local environmental conditions [15,20,132].

Of 13 categories of hybridization outcomes identified in the studies considered in this review (Table 1), genetic swamping and introgression of variants originating from domesticated lineages were two most common outcomes (21% and 18% of studies, respectively). Both these outcomes are commonly considered as negative. Another common outcome (17% of studies), "no or rare evidence of hybridization", can be classified as neutral. The most common positive outcome (8%) was the gain of novel adaptive variation. Graphical representation of the common outcomes of hybridization is presented on Figure 3.

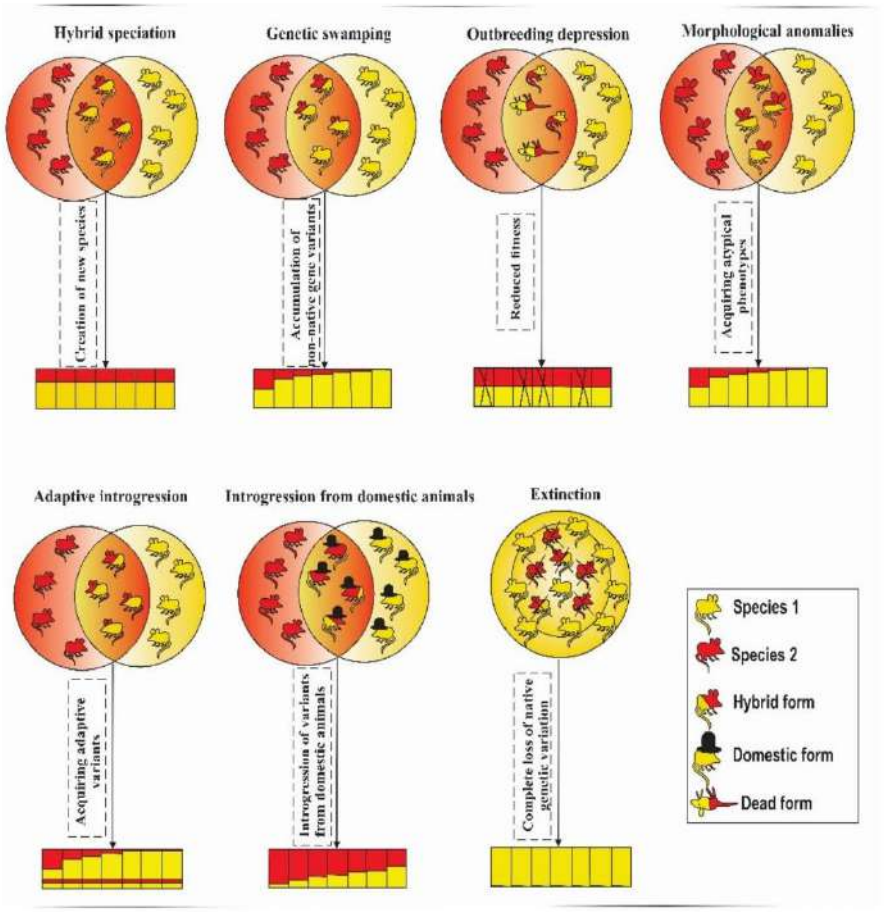


Figure 3. Graphical representation of the common outcomes of hybridization.

The frequencies of different outcomes of hybridization may be affected by the specific sets of keywords that were employed in the literature search. We included keywords such as "Genetic swamping", "Hybrid zone", "Hybrid speciation", "Extinction", and "Outbreeding depression", and therefore studies focusing on these topics may be overrepresented. Furthermore, information about the consequences of hybridization is missing from some studies, which could also affect the result. Nonetheless, we also identified outcomes that were not included in the keywords, and some which were included had very low frequencies among the selected papers (e.g., "Hybrid speciation").

4.2.1. Negative Outcomes

Genetic Swamping

Genetic swamping refers to the process where genotypes of one or both parental species are partially replaced by hybrid genotypes [32]. Genetic swamping is typically considered as a negative consequence of hybridization due to its disruptive effects on genetic integrity of species and potential to eliminate unique adaptations. Negative results of this process are well documented in some cases, e.g., when it leads to extinction (see below) or results in outbreeding depression e.g., [133]. However, many studies reporting genetic swamping do not assess its fitness consequences or long-term effects on the gene pool composition, and therefore it remains unclear whether the negative consequences of this process prevail among all the cases when it occurs. Genetic rescue, i.e., a reduction of extinction probability of a small, isolated population by restoring gene flow [134] is necessarily associated with genetic swamping, especially if the source of gene flow belongs to another species. Therefore, in some cases negative effects of genetic swamping on the species genetic integrity may be compensated by positive effects, such as reduction of inbreeding depression in isolated populations.

Nearly half (48%) of studies included in this systematic review that reported genetic swamping were focused on hybridization between domesticated mammals and their wild relatives, including wolf and domestic dog [38,81,135], wild boar and domestic pig [20,30,93] and wild cat and domestic cat [44]. More than a quarter (28%) of the studies reported genetic swamping of a native gene pool as the main result of hybridization between introduced species and native species, e.g., in Cervidae [36,136] and Mustelidae families [29,137,138]. Over 80% of the studies reporting genetic swamping focused on cases where hybridization was directly or indirectly caused by human actions, i.e., either domestication or species translocation (deliberate or unintentional). This implies that either such cases are considered as greater concern than introgression resulting from natural hybridization between pairs of wild native species and thus are studied more frequently, or genetic swamping is indeed more frequent when it involves a domesticated or introduced species cross-breeding with a native wild species.

Reproductive barriers between closely related species that evolved in geographic isolation may be weak, and therefore after the secondary contact due to translocation, cross-breeding and production of fertile offspring may be common. In such cases, continuous cross-breeding across generations may result in considerable genetic swamping [36]. Accordingly, reproductive isolation between domesticated mammals and their wild ancestors is frequently incomplete due to their recent divergence. For example, hybridization between wolves and domestic dogs results in introgression of hybridization-derived variants into gene pools of both canids [139]. In wolves, introgression of dog variants is mostly driven by drift, with only a small number of genes experiencing negative or positive selection due to this process [139]. In free-ranging domestic dogs, the observed proportion of candidate genes under positive selection was larger than those under negative selection, suggesting that introgression from wolves may provide dogs with an adaptive advantage [139]. This last case demonstrates that genetic swamping is not an unequivocally negative process, and its outcomes should be considered on an individual basis.

Extinction via Genetic Swamping

In extreme cases, extensive genetic swamping may lead a population or an entire species towards extinction [32,140]. Endemic species with patchy, isolated habitats are under a particularly high risk of extinction via hybridization with introduced or invasive closely related species. Of 115 papers considered in this systematic review, three papers mentioned the risk of extinction by genetic swamping [30,114,141]. One of these papers describes the case of the Java warty pig (*Sus verrucosus*), which is endangered with extinction via genetic swamping from the common Indonesian banded pig (*Sus scrofa vitattus*), because of high hybridization rates resulting from the breakdown of reproductive barriers and reduced fertility of hybrids [30]. The second paper reports the case of extinction of the endemic Italian roe deer (*C. c. italicus*) due to extreme genetic swamping from the introduced European roe deer (*Capreolus c. capreolus*) [114]. The third study presents a mathematical model on hybridization between the mountain hare (*Lepus timidus*) and European hare (*Lepus europaeus*), showing that under climate change scenarios, increased hybridization rate can lead to the mountain hare's extinction via genetic swamping [141]. These studies focused on endemic species, which are threatened by habitat fragmentation and overhunting, and are interbreeding with common, closely related species. Extinction of endemic species through genetic swamping has also been reported in other taxonomic groups, including plants e.g., [140].

Notably, identifying key factors involved in the extinction process is very challenging. In many cases, interaction of different forces such as environmental stress, low genetic diversity and small population size, may lead to the extinction vortex [142,143]. Therefore, while hybridization may play an important role in pushing a species towards extinction, it may not be the only contributing factor. Overall, hybridization contributed to extinction in only 11 documented cases [23].

Outbreeding Depression

Outbreeding depression occurs when cross-breeding between two species or populations that are adapted to different environmental conditions results in a loss of local adaptations and reduction of fitness in hybrid individuals [144]. In the studies reviewed here, outbreeding depression has been reported in native species that cross-bred with non-native or invasive species, such as the roe deer interbreeding with the introduced sika deer [36,145], and also in cases of admixture between wild and captive-bred populations e.g., wild versus captive-born American mink, [26]. In both cases, introgression may break-up co-adapted gene complexes, reducing fitness in wild populations and resulting in outbreeding depression [146,147]. Selection pressures in captive-bred populations, associated with adaptations to the captive environment and artificial selection on traits desirable for humans result in the presence of gene variants that are maladaptive in natural habitats [32,148]. Moreover, captive-bred populations have small sizes, which leads to low genetic diversity and inbreeding depression. Introgressive hybridization may introduce maladaptive gene variants present in such captive populations to natural populations, with negative effects on their fitness.

On the other hand, captive breeding programs for conservation purpose have become a conservation tool to prevent extinctions and support reintroductions in cases when remaining wild populations are small and have low genetic diversity [149,150]. For example, captive breeding program of European mink was lunched as a conservation tool for this critically endangered species [151,152].

Among the papers discussed in this systematic review, five studies were focused on admixture between wild and captive-bred populations of the same species. Two of these studies did not detect any signatures of hybridization, and the remaining three studies reported both outbreeding depression and genetic swamping. One of them reported genetic swamping for the native population and increase in genetic diversity for the captive-bred population studied.

Introgression from a Domesticated Lineage

Hybridization between wild species and their domesticated relatives frequently results in the introgression of gene variants typical of domesticated animals to wild populations. Although there is a considerable overlap between the studies included in this category with other categories, we consider it separately due to specific conservation problems resulting from this type of introgression [153]. Domesticated mammals are not separated from their wild ancestors by strong reproductive barriers, and therefore they are likely to cross-breed in regions where their ranges overlap. Introgression of domesticated species' alleles to wild species' gene pool may threaten the genetic integrity of wild species [154], and therefore it is typically considered as a negative process. However, such introgression may increase the genetic variability in wild species suffering the effects of a severe bottleneck, and/or accelerate the process of adaptation to changing environmental conditions by providing novel genetic variation [155]. For example, the Alpine ibex (*Capra ibex ibex*) acquired one of its two MHC DRB alleles from domestic goats (*Capra aegagrus hircus*), which critically increased diversity of this genetically impoverished species at the key component of the immune system [156].

Among 25 studies from our systematic review that fit in this class, 15 studies focused on hybridization between the grey wolf or the dingo and the domestic dog, four studies focused on the wild boar and the domestic pig, four studies on the wild cat and the domestic cat, one study on the wild sheep and the domestic sheep, and one study on the wild yak and the domestic yak. These studies show that the introgression from domestic animals into their wild relatives is more frequent than in the opposite direction. Population sizes of domestic animals are dependent on the human population size, and therefore human population growth combined with the fragmentation of natural habitats increases both the numbers of domestic animals and the probability of encounters with their wild relatives. Given that wild populations are typically considerably smaller than populations of domestic animals, a single hybridization and back-crossing event will have a larger effect on gene pools of wild populations compared with domestic ones.

Studies focused on hybridization between domestic animals and their wild relatives constituted 36% of studies considered in this systematic review, and thus they had significant impact on the overall proportions of different consequences of hybridization. By default, they were responsible for all the cases of introgression from domesticated lineages, which were reported in 44% of studies focused on hybridization with domestic animals. Further 24.5% of studies reported genetic swamping; in this case, the difference between these two consequences is only in wording, with the exception of introgression cases from wild to domestic populations. The most common positive consequence was gaining novel adaptive variation (7%). Overall, the frequency of hybridization outcomes considered as negative (70%) was considerably higher than the frequency of positive outcomes (9%). This suggests that studies on hybridization between domestic animals and their wild relatives have a disproportional contribution to the negative hybridization outcomes in the overall assessment. However, this is based on the assumption that introgression from domesticated lineages and genetic swamping are negative outcomes by default, which has rarely been tested. Given low divergence between domesticated animals and their wild relatives, it may be expected instead that introgression will rarely lead to increased mortality and infertility of admixed individuals, but the presence of atypical phenotypic traits may result in reduced fitness.

Morphological Anomalies

Interspecific hybridization influences phenotypic traits and may create novel or unusual traits [157,158]. Morphological anomalies and abnormal growth are common among hybrid individuals [159] and are sometimes used as a proxy to detect hybridization [160]. Morphological anomalies usually reduce fitness and in extreme cases may cause inviability, and therefore we classified them as a negative outcome of hybridization. In cases where morphological anomalies increase fitness of hybrid individuals, they are considered as

novel adaptive variation, which is a separate category of hybridization outcomes identified in this review (see below).

The divergence of phenotypic traits of admixed individuals from average traits within each of the cross-breeding species increases with their divergence time [157,161,162]. Depending on the species, the anomalies can occur in different body parts, including teeth, skull, horn shape, body size etc. Among the papers considered in this systematic review, we found only two papers that reported morphological anomalies in hybrid individuals, including abnormal placental growth in hybrids between two species of dwarf hamsters (*Phodopus campbelli* and *Phodopus sungorus*) [131] and skull, dental and horn anomalies in the wildebeest hybrids (*Connochaetes taurinus* and *Connochaetes gnu*) [157].

Loss of Reproductive Output

Hybridization can change reproductive output by leading to changes in mating behavior [163] or by wasting reproductive efforts. These changes typically involve reduction in reproductive success, and therefore are considered as a negative consequence of hybridization. For instance, unidirectional introgression from the fin whale to blue whale resulted in the reduction of reproductive rate of the blue whale, reducing its recovery [123]. In cases when most hybrid individuals are infertile and inviable, introgression does not happen or is rare, and therefore the consequences of hybridization are reduced to the production of F1 hybrids only. Moreover, in some cases the reproductive output differs between different generations of hybrids. For example, hybridization between *Microtus hartingi lydius* and *Microtus hartingi strandzensis* produces viable and prolific F1 hybrids, while in the F2 generation, males are sterile and the mortality rate is high [164]. Falling fertility rates and loss of reproductive outputs may lead to severe demographic declines in parental species and even a rapid extinction of local populations involved in cross-breeding.

4.2.2. Positive Outcomes

Increase in Genetic Diversity and Reduction of Inbreeding

In a specific case when genetic diversity of a population is very low and the rate of inbreeding is high, introgression from a non-native population or species can increase genetic diversity without any signs of outbreeding depression. This can be considered as a positive consequence of hybridization. Moreover, in small and fragmented populations that have low genetic diversity and experience inbreeding depression, hybridization can restore population viability [165,166]. Genetic rescue, i.e., restoration of genetic diversity and mitigation of inbreeding depression through gene flow can be a valuable tool in conservation of small, isolated populations [167]. For instance, introgressive hybridization with a non-native Alpine chamois (*R. r. rupicapra*) was shown to improve genetic diversity of Tatra chamois (*R. r. tatrica*), an endangered endemic population in the Tatra Mountains that suffered from a high level of inbreeding depression [21].

Although introgression from domesticated lineages can be considered as a threat for wild populations, in some circumstances it can increase genetic diversity and viability of wild populations. For example, increase in genetic diversity has been reported in European wild boars that cross-bred with domestic pigs [20]. Accordingly, admixture between feral and farmed populations of American mink (*Neovison vison*) increased genetic diversity of the invasive populations of these species in Europe, which could increase their adaptive potential and therefore compromise management efforts to control them [137]. Although this is a negative process from the conservation perspective, it can be considered as a positive outcome of hybridization in terms of increasing individual fitness in the invasive population.

Novel Adaptive Variation

In some cases, creation of novel genetic diversity via hybridization can facilitate species adaptation to variable or novel environmental conditions, without a loss of its genetic integrity [132,168]. Admixed individuals may acquire new adaptive traits, providing them

with selective advantages in comparison to their parental species [169,170]. For instance, in eastern Poland, introgression from the Siberian roe deer (*C. pygargus*) allowed the European roe deer (*C. capreolus*) to adapt better to severe winters, which are an important contributing factor of roe deer mortality [171]. Furthermore, hybridization between the coyote (*Canis latrans*) and the grey wolf (*Canis lupus*) in Canada has resulted in the introduction of novel adaptive variation to the coyote populations, allowing them to increase in body size, which in turn improved their success in hunting deer [172].

In several mammalian species, including humans, the presence of adaptive variation from their extinct relatives has been detected [173–175], implying that ancient hybridization events provided long-lasting positive fitness effects [176]. For example, Tibetan and Himalayan wolves experienced ancient introgression from an unknown canid lineage, which resulted in the introgression of an *EPAS1* haplotype that confers an adaptive advantage in high altitude environments [175]. In humans, ancient cross-breeding with Neanderthals and Denisovans in Eurasia resulted in introgression of novel adaptive variation, but also increased the genetic load compared with non-admixed African populations [25,177]. Altogether, among 14 papers considered in this systematic review that used genome-wide SNPs or whole genome sequencing, there were six papers reporting cases of ancient introgression. Three of these studies showed that ancient introgression was associated with gaining novel adaptive variation, and the remaining three papers reported ancient introgression without investigating its consequences.

Hybrid Speciation

Hybrid speciation refers to the process in which hybridization results in the creation of a new species, which is characterized by mixed ancestry and distinct genetic composition from its parental species [18]. Hybridization may act as a driving force in speciation by creating new hybrid phenotypes or providing necessary material for adaptive divergence [17]. Given that the creation of a new species increases biodiversity, it can be considered as a positive outcome of hybridization.

Three criteria should be met to demonstrate speciation via hybridization; first, confirmed evidence of a past hybridization event in the putative hybrid species, second, reproductive isolation between the hybrid species and its parental species, and finally the presence of isolating impacts of hybridization [178]. Hybrid speciation may allow the new species to colonize a new habitat [179].

In this review we found four studies that reported hybrid speciation [131,180–182]. These studies showed that the emergence of distinct phenotypic traits in hybrid individuals may play an important role in speciation by impeding gene flow between parental species and hybrid individuals [131]. For instance, differentiation in facial patterns in the primate genus *Cercopithecus* is one of the key mechanisms driving hybrid speciation in this genus [181,183]. Abnormal growth patterns in hybrids between two dwarf hamster species, *P. campbelli* and *P. sungorus*, were suggested to play an important role in speciation by contributing to reproductive isolation between these recently diverged species [131]. Despite its potentially important role in mammalian speciation, the genetic basis of growth-related developmental inviability is still unknown [131]. However, studies on a hybrid zone between subspecies the house mouse (*Mus musculus*), which is a model species in genetics, provided an insight into the molecular mechanisms underlying hybrid speciation. Dysfunction in the Mecp2 protein in the house mouse resulting from introgressive hybridization within the hybrid zone may induce changes in the expression of thousands of genes, which may initiate the speciation process [182].

4.2.3. Neutral or Unspecified Outcomes

Intermediate Phenotypic Traits

Although in some cases interspecific hybridization may create deleterious morphological anomalies or novel adaptive traits (see above), hybrid individuals frequently show intermediate phenotypic traits compared to their parents [184]. The additive effect, domi-

nance effect, and/or epistatic effect may create variation in polygenic traits [157,185]. In the additive model, F1 hybrid offspring shows intermediate phenotypes relative to their parents [185]. A classic example of the effect of hybridization on morphological traits are the Darwin finches in Galapagos, where most hybrid individuals have intermediate body size and beak shape compared with the parental species [186]. Several studies in this review reported intermediate phenotypes, e.g., in cetaceans [187], mustelids [188], camelids [189] and primates [181]. The presence of admixed individuals with intermediate phenotypes may impede species identification in the field. For instance, field identification of four chipmunk species (*Tamias* spp.) in the Sierra Nevada, USA, was associated with 14% error rate, which was in part attributed to sporadic hybridization among these species [190]. The fitness consequences of intermediate phenotypic traits have rarely been studied and therefore this outcome of hybridization could not be classified as either positive or negative.

Hybrid Zones

Hybrid zones are areas where two genetically distinct linages meet, mate and create viable offspring [191]. These geographic regions are usually narrow, with the width ranging from several meters to several kilometers [192]. Hybrid zones can be created through natural hybridization between parapatric or sympatric species [193,194]. Most hybrid zones are maintained by the balance between natural selection against hybrids and dispersal capabilities of the cross-breeding taxa [179,191]. If before the range expansion or removal of a geographic barrier, reproductive isolation between closely related species has not been complete, hybrid zone may be formed. For example, due to a recent divergence and weak reproductive isolation between the pine marten *Martes martes* and the sable *Martes zibellina* in Western Siberia, a vast hybrid zone has formed between these species after the Last Glacial Maximum [138]. Features of hybrid zones, such as fertility or sterility of hybrid individuals, directionality of mating, hybridization frequency, and geographic extent of introgression, vary considerably, and their examination can help understand the mechanisms of hybrid zone maintenance [195]. In woodrat species (*Neotoma bryanti* and *Neotoma lepida*), a hybrid zone has been maintained as a result of sporadic cross-breeding between these species and hybrid fertility [196]. Among studies included in this literature review, hybrid zones have been described in mice (*Mus musculus musculus* and *Mus musculus domesticus*) [197], different woodrat species [195,196,198–200], marmots [201], primates [181], artiodactyles [157,202], Diprotodontia [203,204] and carnivores [138,205].

Hybridization without Significant Impacts

Some studies included in this review showed that despite hybridization, populations maintained their genetic distinctiveness [206]. For instance, despite extensive rate of hybridization among different bat species in Poland, their gene pools have not been disrupted by introgression [74]. Furthermore, admixture between Italian wolves and domestic dogs did not affect the integrity of wolves' gene pool [207]. The lack of significant effects on the gene pool does not necessarily imply the lack of any effects, e.g., loss of reproductive effort. Given that these effects were not studied, we classified this type of outcome as neutral.

No or Rare Evidence of Hybridization

This category includes studies that have not found any signs of hybridization in the populations studied or found only very limited evidence e.g., [154,208–215]. We found 24 studies that fitted this category. This could include cases where hybridization was rare or did not occur, as well as cases where limitations in sampling and the use of low number of genetic makers could result in poor detection of hybridization [216–225]. The number of genetic markers is important to detect signatures of hybridization, especially if cross-breeding and/or back-crossing events are rare [56,124]. Moreover, small data sets may show only a preliminary assessment of hybridization [226], and comprehensive sampling is necessary to obtain reliable results.

In some cases, efficient conservation management may result in low rate of hybridization [212,227]. For example, because of careful monitoring and management, the Scandinavian wolf population shows a lower level of hybridization with dogs compared to other European wolf populations, which was demonstrated based on the comprehensive sampling and the analysis of whole genomes [82]. The rate of hybridization between the same pairs of species may differ regionally depending on environmental conditions [228]. For example, hybridization between bat species *Myotis myotis* and *Myotis blythii* has been reported in Europe, but in Turkey no signs of hybridization between these species have been detected [229].

Altogether, 38% of the studies assessed in this systematic review reported neutral or unspecified outcomes of hybridization. Within this group, "no or rare evidence of hybridization" was the most common hybridization outcome, which occurred in 17% of all the studies on hybridization. This result suggests that one of the five categories of hybridization outcomes delineated for the purpose of species conservation [35], "negligible impact and minimal introgression of genes into the species of concern", occurs relatively frequently. Therefore, a presumption that hybridization always constitutes a threat to biodiversity is incorrect, and instead the decision-making regarding the management and conservation of wild-living hybrids should be based on the examination of hybridization outcomes case-by-case [35].

4.2.4. Consequences of Hybridization for Threatened Species

Only 18% of studies considered in this review were focused on threatened species (having the IUCN Red List categories of Near Threatened (NT), Vulnerable (VU), Endangered (EN), Critically Endangered (CR), Extinct in the Wild (EW)). Negative consequences (e.g., genetic swamping, extinction via genetic swamping, introgression from domesticated lineages and loss of reproductive output) were reported in 38% of these studies and positive consequences (with only one category, gaining novel adaptive variation) were reported in 14% of them. In the remaining studies, the authors did not mention any positive or negative consequences or did not find any evidence of hybridization. The frequency of the positive consequences reported for all the studies assessed in this systematic review was very similar (13%), while the frequency of the negative consequences was higher (49%) compared to those observed in the studies on threatened species. This suggests that the negative consequences of hybridization are not intensified in endangered species, at least the mammalian species considered in this review.

5. Conclusions

Among the papers included in this systematic review, hybridization outcomes typically considered as negative had considerably higher frequency (49%) than those considered as positive (13%). However, these frequencies could have been biased by several factors and therefore should be treated with caution. For instance, two most frequent outcomes of hybridization, genetic swamping and introgression of variants from domestic animals are typically considered as negative, but this is not always the case. In some circumstances, moderate levels of genetic swamping or introgression of domestication-related variants may result in increased fitness and genetic rescue [155,230]. The cases where hybridization outcomes are unequivocally negative, leading to extinction or loss of reproductive output, are relatively rare. They were reported in 13% of the studies considered in this review—the same frequency as that of the positive outcomes. In cases when the hybridization outcome cannot be easily determined, e.g., when genetic swamping occurs at a low rate, long-term monitoring of admixed populations is required to conclude about advantages and disadvantages of introgressive hybridization. For this purpose, at least two consecutive generations should be monitored [12,13], but currently, many studies are based on a singular sampling effort, and fitness of sampled individuals is rarely assessed.

It is also important to stress that the detection of different outcomes of hybridization depends on the type of molecular markers applied. Microsatellite markers enable identifi-

cation of first-generation hybrids and recent back-crosses, but cannot reliably detect more distant hybridization events [56]. Since microsatellites are neutral genetic markers and are typically genotyped in small numbers (<100), they cannot be used to detect adaptive introgression or hybrid speciation, which are among the most frequently reported positive outcomes of hybridization. In contrast, some negative outcomes, such as genetic swamping, can be detected using a small number of neutral markers. Given that until recently microsatellite loci were the most frequently chosen markers in hybridization studies, and they were used in 50% of studies considered in this systematic review, the frequency of negative outcomes of hybridization may be overestimated.

A combination of neutral loci and those located within coding genes is better suited to provide an unbiased insight into the relative frequencies of positive and negative hybridization outcomes and identify factors that affect them. Single Nucleotide Polymorphisms (SNPs) can be genotyped in large numbers (hundreds of thousands to millions) using arrays or next generation sequencing, which makes them suitable for identification of adaptive loci [231,232] as well as detection of small proportions of hybrid ancestry and identification of F2-F4 backcrosses [233,234]. Therefore, the application of this type of genetic markers creates an opportunity to identify both positive and negative impacts of hybridization.

In some circumstances, hybridization can be used as a conservation tool to facilitate adaptation of populations to changing habitat conditions and to increase individual fitness in populations experiencing inbreeding depression [12,13]. To use hybridization this way, we need to achieve a better understanding on how to prevent negative effects of hybridization without eliminating the potential for the positive effects. This will require comprehensive studies focusing on the genetic effects of hybridization on both neutral and functional parts of the genome and fitness effects of cross-breeding on F1 hybrids and several generations of back-crosses. Experimental studies simulating different evolutionary scenarios may be the best way to achieve an unbiased assessment of the frequency of different hybridization outcomes.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/genes13010050/s1, Table S1: A list of papers considered in the systematic review.

Author Contributions: R.A. carried out the literature review and wrote the first draft of the manuscript. M.P. designed the study, supervised its implementation, and revised the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This study was funded by the Polish National Science Centre (grant no. 2019/34/E/NZ8/00246 to M.P.) and the Polish National Agency for Scientific Exchange—NAWA (Polish Returns Fellowship PPN/PPO/2018/1/00037 to M.P.).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The references of 115 papers used as data source in the systematic review are provided in the reference list. In addition, the list of these 115 papers is provided in the Supplementary Table S1. The summary data from these studies is provided in the Results section.

Acknowledgments: We thank three anonymous reviewers for their helpful comments on the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Mayr, E. *Animal Species and Evolution*; Belknap: Cambridge, MA, USA, 1963.
2. Taylor, S.A.; Larson, E. Insights from genomes into the evolutionary importance and prevalence of hybridization in nature. *Nat. Ecol. Evol.* **2019**, *3*, 170–177. [CrossRef] [PubMed]
3. Mallet, J. Hybridization as an invasion of the genome. *Trends Ecol. Evol.* **2005**, *20*, 229–237. [CrossRef] [PubMed]
4. Arnold, M.L. *Natural Hybridization and Evolution*; Oxford University Press on Demand: Oxford, UK, 1997.

5. Nesi, N.; Nakoune, E.; Cruaud, C.; Hassanin, A. DNA barcoding of African fruit bats (Mammalia, Pteropodidae). The mitochondrial genome does not provide a reliable discrimination between *Epomophorus Gamb* and *Micropteropus pusillus*. *C. R. Biol.* **2011**, *334*, 544–554. [CrossRef] [PubMed]
6. Abbott, R.; Barton, N.H.; Good, J.M. Genomics of hybridization and its evolutionary consequences. *Mol. Ecol.* **2016**, *25*, 2325–2332. [CrossRef]
7. Goulet, B.E.; Roda, F.; Hopkins, R. Hybridization in Plants: Old Ideas, New Techniques. *Plant Physiol.* **2017**, *173*, 65–78. [CrossRef] [PubMed]
8. Wayne, R.K.; Shaffer, H.B. Hybridization and endangered species protection in the molecular era. *Mol. Ecol.* **2016**, *25*, 2680–2689. [CrossRef]
9. Mota, M.R.; Pinheiro, F.; Leal, B.S.S.; Wendt, T.; Palma-Silva, C. The role of hybridization and introgression in maintaining species integrity and cohesion in naturally isolated inselberg bromeliad populations. *Plant Biol.* **2019**, *21*, 122–132. [CrossRef] [PubMed]
10. Brennan, A.C.; Woodward, G.; Seehausen, O.; Muñoz-Fuentes, V.; Moritz, C.; Guelmami, A.; Abbott, R.J.; Edelaar, P. Hybridization due to changing species distributions: Adding problems or solutions to conservation of biodiversity during global change? *Evol. Ecol. Res.* **2015**, *16*, 475–491.
11. Lavrenchenko, L.A.; Bulatova, N.S. The role of hybrid zones in speciation: A case study on chromosome races of the house mouse Mus domesticus and common shrew *Sorex araneus*. *Biol. Bull. Rev.* **2016**, *6*, 232–244. [CrossRef]
12. Chan, W.Y.; Peplow, L.M.; Menéndez, P.; Hoffmann, A.; van Oppen, M. Interspecific Hybridization May Provide Novel Opportunities for Coral Reef Restoration. *Front. Mar. Sci.* **2018**, *5*, 160. [CrossRef]
13. Hamilton, J.A.; Miller, J. Adaptive introgression as a resource for management and genetic conservation in a changing climate. *Conserv. Biol.* **2016**, *30*, 33–41. [CrossRef]
14. Hoffmann, A.A.; Sgro, C. Climate change and evolutionary adaptation. *Nature* **2011**, *470*, 479–485. [CrossRef]
15. Olano-Marin, J.; Plis, K.; Sönnichsen, L.; Borowik, T.; Niedziałkowska, M.; Jędrzejewska, B. Weak population structure in European roe deer (*Capreolus capreolus*) and evidence of introgressive hybridization with Siberian roe deer (*C. pygargus*) in northeastern Poland. *PLoS ONE* **2014**, *9*, e109147.
16. Seehausen, O. Hybridization and adaptive radiation. *Trends Ecol. Evol.* **2004**, *19*, 198–207. [CrossRef] [PubMed]
17. Abbott, R.; Albach, D.; Ansell, S.; Arntzen, J.W.; Baird, S.J.; Bierne, N.; Boughman, J.; Brelsford, A.; Buerkle, C.A.; Buggs, R. Hybridization and speciation. *J. Evol. Biol.* **2013**, *26*, 229–246. [CrossRef]
18. Mallet, J. Hybrid speciation. *Nature* **2007**, *446*, 279–283. [CrossRef]
19. Lavrenchenko, L. Hybrid speciation in mammals: Illusion or reality? *Biol. Bull. Rev.* **2014**, *4*, 198–209. [CrossRef]
20. Canu, A.; Vilaça, S.; Iacolina, L.; Apollonio, M.; Bertorelle, G.; Scandura, M. Lack of polymorphism at the MC1R wild-type allele and evidence of domestic allele introgression across European wild boar populations. *Mamm. Biol.* **2016**, *81*, 477–479. [CrossRef]
21. Zemanová, B.; Hájková, P.; Hájek, B.; Martínková, N.; Mikulíček, P.; Zima, J.; Bryja, J. Extremely low genetic variation in endangered *Tatra chamois* and evidence for hybridization with an introduced Alpine population. *Conserv. Genet.* **2015**, *16*, 729–741. [CrossRef]
22. Ruiz-García, M.; Pinedo-Castro, M.; Shostell, J.M. Small spotted bodies with multiple specific mitochondrial DNAs: Existence of diverse and differentiated tigrina lineages or species (*Leopardus* spp.: Felidae, Mammalia) throughout Latin America. *Mitochondrial DNA Part A* **2018**, *29*, 993–1014. [CrossRef]
23. Draper, D.; Laguna, E.; Marques, I. Demystifying Negative Connotations of Hybridization for Less Biased Conservation Policies. *Front. Ecol. Evol.* **2021**, *9*, 268. [CrossRef]
24. Cairns, K.M.; Nesbitt, B.J.; Laffan, S.W.; Letnic, M.; Crowther, M.S. Geographic hot spots of dingo genetic ancestry in southeastern Australia despite hybridisation with domestic dogs. *Conserv. Genet.* **2020**, *21*, 77–90. [CrossRef]
25. Harris, K.; Nielsen, R. The Genetic Cost of Neanderthal Introgression. *Genetics* **2016**, *203*, 881–891. [CrossRef] [PubMed]
26. Beauclerc, K.B.; Bowman, J.; Schulte-Hostedde, A.I. Assessing the cryptic invasion of a domestic conspecific: A merican mink in their native range. *Ecol. Evol.* **2013**, *3*, 2296–2309. [CrossRef]
27. Grobler, P.; van Wyk, A.M.; Dalton, D.L.; van Vuuren, B.J.; Kotzé, A. Assessing introgressive hybridization between blue wildebeest (*Connochaetes taurinus*) and black wildebeest (*Connochaetes gnou*) from South Africa. *Conserv. Genet.* **2018**, *19*, 981–993. [CrossRef]
28. Koutsogiannouli, E.A.; Moutou, K.A.; Sarafidou, T.; Stamatis, C.; Mamuris, Z. Detection of hybrids between wild boars (*Sus scrofa scrofa*) and domestic pigs (*Sus scrofa f. domestica*) in Greece, using the PCR-RFLP method on melanocortin-1 receptor (MC1R) mutations. *Mamm. Biol.* **2010**, *75*, 69–73. [CrossRef]
29. Colella, J.P.; Wilson, R.E.; Talbot, S.L.; Cook, J.A. Implications of introgression for wildlife translocations: The case of North American martens. *Conserv. Genet.* **2019**, *20*, 153–166. [CrossRef]
30. Drygala, F.; Rode-Margono, J.; Semiadi, G.; Frantz, A.C. Evidence of hybridisation between the common Indonesian banded pig (*Sus scrofa vittatus*) and the endangered Java warty pig (*Sus verrucosus*). *Conserv. Genet.* **2020**, *21*, 1073–1078. [CrossRef]
31. Haus, T.; Roos, C.; Zinner, D. Discordance between spatial distributions of Y-chromosomal and mitochondrial haplotypes in African green monkeys (*Chlorocebus* spp.): A result of introgressive hybridization or cryptic diversity? *Int. J. Primatol.* **2013**, *34*, 986–999. [CrossRef]
32. Todesco, M.; Pascual, M.A.; Owens, G.L.; Ostevik, K.L.; Moyers, B.T.; Hübner, S.; Heredia, S.M.; Hahn, M.A.; Caseys, C.; Bock, D.G. Hybridization and extinction. *Evol. Appl.* **2016**, *9*, 892–908. [CrossRef]

33. Balao, F.; Casimiro-Soriguer, R.; García-Castaño, J.L.; Terrab, A.; Talavera, S. Big thistle eats the little thistle: Does unidirectional introgressive hybridization endanger the conservation of *Onopordum hinojense*? *New Phytol.* **2015**, *206*, 448–458. [CrossRef]
34. Lepais, O.; Petit, R.; Guichoux, E.; Lavabre, J.; Alberto, F.; Kremer, A.; Gerber, S. Species relative abundance and direction of introgression in oaks. *Mol. Ecol.* **2009**, *18*, 2228–2242. [CrossRef] [PubMed]
35. Hirashiki, C.; Kareiva, P.; Marvier, M. Concern over hybridization risks should not preclude conservation interventions. *Conserv. Sci. Pract.* **2021**, *3*, e424. [CrossRef]
36. Eva, S.N.; Yamazaki, Y. Hybridization between native and introduced individuals of sika deer in the central part of Toyama Prefecture. *Mammal Study* **2018**, *43*, 269–274. [CrossRef]
37. Sagua, M.I.; Figueroa, C.; Acosta, D.; Fernández, G.; Carpinetti, B.; Birochio, D.; Merino, M.L. Inferring the origin and genetic diversity of the introduced wild boar (*Sus scrofa*) populations in Argentina: An approach from mitochondrial markers. *Mammal Res.* **2018**, *63*, 467–476. [CrossRef]
38. Pilot, M.; Greco, C.; vonHoldt, B.M.; Randi, E.; Jędrzejewski, W.; Sidorovich, V.E.; Konopiński, M.K.; Ostrander, E.A.; Wayne, R.K. Widespread, long-term admixture between grey wolves and domestic dogs across Eurasia and its implications for the conservation status of hybrids. *Evol. Appl.* **2018**, *11*, 662–680. [CrossRef]
39. Iacolina, L.; Scandura, M.; Gazzola, A.; Cappai, N.; Capitani, C.; Mattioli, L.; Vercillo, F.; Apollonio, M. Y-chromosome microsatellite variation in Italian wolves: A contribution to the study of wolf-dog hybridization patterns. *Mamm. Biol.* **2010**, *75*, 341–347. [CrossRef]
40. Hulva, P.; Černá Bolfíková, B.; Woznicová, V.; Jindřichová, M.; Benešová, M.; Mysłajek, R.W.; Nowak, S.; Szewczyk, M.; Niedźwiecka, N.; Figura, M. Wolves at the crossroad: Fission–fusion range biogeography in the Western Carpathians and Central Europe. *Divers. Distrib.* **2018**, *24*, 179–192. [CrossRef]
41. Munoz-Fuentes, V.; Darimont, C.T.; Paquet, P.C.; Leonard, J.A. The genetic legacy of extirpation and re-colonization in Vancouver Island wolves. *Conserv. Genet.* **2010**, *11*, 547–556. [CrossRef]
42. Santostasi, N.L.; Gimenez, O.; Caniglia, R.; Fabbri, E.; Molinari, L.; Reggioni, W.; Ciucci, P. Estimating Admixture at the Population Scale: Taking Imperfect Detectability and Uncertainty in Hybrid Classification Seriously. *J. Wildl. Manag.* **2021**. [CrossRef]
43. Cairns, K.M.; Newman, K.D.; Crowther, M.S.; Letnic, M. Pelage variation in dingoes across southeastern Australia: Implications for conservation and management. *J. Zool.* **2021**, *314*, 104–115. [CrossRef]
44. Mattucci, F.; Oliveira, R.; Lyons, L.A.; Alves, P.C.; Randi, E. European wildcat populations are subdivided into five main biogeographic groups: Consequences of Pleistocene climate changes or recent anthropogenic fragmentation? *Ecol. Evol.* **2016**, *6*, 3–22. [CrossRef] [PubMed]
45. Beutel, T.; Reineking, B.; Tiesmeyer, A.; Nowak, C.; Heurich, M. Spatial patterns of co-occurrence of the European wildcat *Felis silvestris silvestris* and domestic cats *Felis silvestris catus* in the Bavarian Forest National Park. *Wildl. Biol.* **2017**, *2017*. [CrossRef]
46. Oliveira, R.; Randi, E.; Mattucci, F.; Kurushima, J.; Lyons, L.A.; Alves, P. Toward a genome-wide approach for detecting hybrids: Informative SNPs to detect introgression between domestic cats and European wildcats (*Felis silvestris*). *Heredity* **2015**, *115*, 195–205. [CrossRef]
47. Le Roux, J.J.; Foxcroft, L.C.; Herbst, M.; MacFadyen, S. Genetic analysis shows low levels of hybridization between A frican wildcats (*Felis silvestris lybica*) and domestic cats (*F. s. catus*) in South Africa. *Ecol. Evol.* **2015**, *5*, 288–299. [CrossRef]
48. Boitani, L. Genetic considerations on wolf conservation in Italy. *Ital. J. Zool.* **1984**, *51*, 367–373. [CrossRef]
49. Gottelli, D.; Sillero-Zubiri, C.; Applebaum, G.D.; Roy, M.S.; Girman, D.J.; Garcia-Moreno, J.; Ostrander, E.A.; Wayne, R.K. Molecular genetics of the most endangered canid: The Ethiopian wolf *Canis simensis*. *Mol. Ecol.* **1994**, *3*, 301–312. [CrossRef] [PubMed]
50. Randi, E. Genetics and conservation of wolves *Canis lupus* in Europe. *Mammal Rev.* **2011**, *41*, 99–111. [CrossRef]
51. Torres, R.T.; Ferreira, E.; Rocha, R.G.; Fonseca, C. Hybridization between wolf and domestic dog: First evidence from an endangered population in central Portugal. *Mamm. Biol.* **2017**, *86*, 70–74. [CrossRef]
52. Popova, E.; Zlatanova, D. Living a dog's life: A putative gray wolf in a feral dog group. *Mammalia* **2020**, *84*, 115–120. [CrossRef]
53. Saetre, P.; Lindberg, J.; Leonard, J.A.; Olsson, K.; Pettersson, U.; Ellegren, H.; Bergström, T.F.; Vila, C.; Jazin, E. From wild wolf to domestic dog: Gene expression changes in the brain. *Mol. Brain Res.* **2004**, *126*, 198–206. [CrossRef]
54. Lounsberry, Z.T.; Quinn, C.B.; Statham, M.J.; Angulo, C.L.; Kalani, T.J.; Tiller, E.; Sacks, B.N. Investigating genetic introgression from farmed red foxes into the wild population in Newfoundland, Canada. *Conserv. Genet.* **2017**, *18*, 383–392. [CrossRef]
55. Arnold, M.L.; Martin, N.H. Hybrid fitness across time and habitats. *Trends Ecol. Evol.* **2010**, *25*, 530–536. [CrossRef] [PubMed]
56. McFarlane, S.E.; Pemberton, J.M. Detecting the true extent of introgression during anthropogenic hybridization. *Trends Ecol. Evol.* **2019**, *34*, 315–326. [CrossRef] [PubMed]
57. Shurtliff, Q.R. Mammalian hybrid zones: A review. *Mammal Rev.* **2013**, *43*, 1–21. [CrossRef]
58. Petit, R.J.; Excoffier, L. Gene flow and species delimitation. *Trends Ecol. Evol.* **2009**, *24*, 386–393. [CrossRef]
59. Iacolina, L.; Corlatti, L.; Buzan, E.; Safner, T.; Šprem, N. Hybridisation in European ungulates: An overview of the current status, causes, and consequences. *Mammal Rev.* **2019**, *49*, 45–59. [CrossRef]
60. Morgan-Richards, M.; Smissen, R.D.; Shepherd, L.D.; Wallis, G.P.; Hayward, J.J.; Chan, C.H.; Chambers, G.K.; Chapman, H.M. A review of genetic analyses of hybridisation in New Zealand. *J. R. Soc. N. Z.* **2009**, *39*, 15–34. [CrossRef]
61. Sinclair, A.R.E. Mammal population regulation, keystone processes and ecosystem dynamics. *Philos. Trans. R. Soc. B Biol. Sci.* **2003**, *358*, 1729–1740. [CrossRef]

62. Wilson, D.E.; Reeder, D.M. *Mammal Species of the World: A Taxonomic and Geographic Reference*; JHU Press: Baltimore, MD, USA, 2005; Volume 1.
63. Tolesa, Z.; Bekele, E.; Tesfaye, K.; Ben Slimen, H.; Valqui, J.; Getahun, A.; Hartl, G.B.; Suchentrunk, F. Mitochondrial and nuclear DNA reveals reticulate evolution in hares (*Lepus* spp., Lagomorpha, Mammalia) from Ethiopia. *PLoS ONE* **2017**, *12*, e0180137. [CrossRef]
64. Grobler, J.P.; Hayter, K.N.; Labuschagne, C.; Nel, E.; Coetzer, W.G. The genetic status of naturally occurring black-nosed impala from northern South Africa. *Mamm. Biol.* **2017**, *82*, 27–33. [CrossRef]
65. Moura, A.E.; Tsingarska, E.; Dąbrowski, M.J.; Czarnomska, S.D.; Jędrzejewska, B.; Pilot, M. Unregulated hunting and genetic recovery from a severe population decline: The cautionary case of Bulgarian wolves. *Conserv. Genet.* **2014**, *15*, 405–417. [CrossRef]
66. Cairns, K.M.; Brown, S.K.; Sacks, B.N.; Ballard, J.W.O. Conservation implications for dingoes from the maternal and paternal genome: Multiple populations, dog introgression, and demography. *Ecol. Evol.* **2017**, *7*, 9787–9807. [CrossRef]
67. Canu, A.; Apollonio, M.; Scandura, M. Unmasking the invader: Genetic identity of invasive wild boar from three minor islands off Sardinia (Italy). *Mamm. Biol.* **2018**, *93*, 29–37. [CrossRef]
68. Artyushin, I.; Bannikova, A.; Lebedev, V.; Kruskop, S. Mitochondrial DNA relationships among North Palaearctic Eptesicus (Vespertilionidae, Chiroptera) and past hybridization between common serotine and northern bat. *Zootaxa* **2009**, *2262*, 40–52.
69. Berthier, P.; Excoffier, L.; Ruedi, M. Recurrent replacement of mtDNA and cryptic hybridization between two sibling bat species *Myotis myotis* and *Myotis blythii*. *Proc. R. Soc. B Biol. Sci.* **2006**, *273*, 3101–3123. [CrossRef]
70. Bogdanowicz, W.; Van Den Bussche, R.A.; Gajewska, M.; Postawa, T.; Harutyunyan, M. Ancient and contemporary DNA sheds light on the history of mouse-eared bats in Europe and the Caucasus. *Acta Chiropterol.* **2009**, *11*, 289–305. [CrossRef]
71. Hoffmann, F.G.; Owen, J.G.; Baker, R.J. mtDNA perspective of chromosomal diversification and hybridization in Peters' tent-making bat (*Uroderma bilobatum*: Phyllostomidae). *Mol. Ecol.* **2003**, *12*, 2981–2993. [CrossRef]
72. Afonso, E.; Goydadin, A.-C.; Giraudoux, P.; Farny, G. Investigating hybridization between the two sibling bat species Myotis myotis and M. blythii from guano in a natural mixed maternity colony. *PLoS ONE* **2017**, *12*, e0170534. [CrossRef]
73. Vallo, P.; Benda, P.; Červený, J.; Koubek, P. Conflicting mitochondrial and nuclear paraphyly in small-sized West African house bats (Vespertilionidae). *Zool. Scr.* **2013**, *42*, 1–12. [CrossRef]
74. Bogdanowicz, W.; Piksa, K.; Tereba, A. Hybridization hotspots at bat swarming sites. *PLoS ONE* **2012**, *7*, e53334. [CrossRef]
75. Arlettaz, R.; Christe, P.; Lugon, A.; Perrin, N.; Vogel, P. Food availability dictates the timing of parturition in insectivorous mouse-eared bats. *Oikos* **2001**, *95*, 105–111. [CrossRef]
76. Linnell, J.D.; Swenson, J.E.; Andersen, R. Conservation of biodiversity in Scandinavian boreal forests: Large carnivores as flagships, umbrellas, indicators, or keystones? *Biodivers. Conserv.* **2000**, *9*, 857–868. [CrossRef]
77. Macdonald, E.; Burnham, D.; Hinks, A.; Dickman, A.; Malhi, Y.; Macdonald, D. Conservation inequality and the charismatic cat: Felis felicis. *Glob. Ecol. Conserv.* **2015**, *3*, 851–866. [CrossRef]
78. Tensen, L. Biases in wildlife and conservation research, using felids and canids as a case study. *Glob. Ecol. Conserv.* **2018**, *15*, e00423. [CrossRef]
79. Tisdell, C.; Nantha, H.S.; Wilson, C. Endangerment and likeability of wildlife species: How important are they for payments proposed for conservation? *Ecol. Econ.* **2007**, *60*, 627–633. [CrossRef]
80. Nyhus, P.J. Human–wildlife conflict and coexistence. *Annu. Rev. Environ. Resour.* **2016**, *41*, 143–171. [CrossRef]
81. Caniglia, R.; Fabbri, E.; Galaverni, M.; Milanesi, P.; Randi, E. Noninvasive sampling and genetic variability, pack structure, and dynamics in an expanding wolf population. *J. Mammal.* **2014**, *95*, 41–59. [CrossRef]
82. Smeds, L.; Aspi, J.; Berglund, J.; Kojola, I.; Tirronen, K.; Ellegren, H. Whole-genome analyses provide no evidence for dog introgression in Fennoscandian wolf populations. *Evol. Appl.* **2021**, *14*, 721–734. [CrossRef] [PubMed]
83. Boggiano, F.; Ciofi, C.; Boitani, L.; Formia, A.; Grottoli, L.; Natali, C.; Ciucci, P. Detection of an East European wolf haplotype puzzles mitochondrial DNA monomorphism of the Italian wolf population. *Mamm. Biol.* **2013**, *78*, 374–378. [CrossRef]
84. Pilot, M.; Moura, A.E.; Okhlopkov, I.M.; Mamaev, N.V.; Alagaili, A.N.; Mohammed, O.B.; Yavruyan, E.G.; Manaseryan, N.H.; Hayrapetyan, V.; Kopaliani, N. Global phylogeographic and admixture patterns in grey wolves and genetic legacy of an ancient Siberian lineage. *Sci. Rep.* **2019**, *9*, 1–13. [CrossRef]
85. Zwijacz-Kozica, T.; Ważna, A.; Muñoz-Fuentes, V.; Tiesmeyer, A.; Cichocki, J.; Nowak, C. Not European wildcats, but domestic cats inhabit Tatra National Park. *Pol. J. Ecol.* **2017**, *65*, 415–421. [CrossRef]
86. Nussberger, B.; Wandeler, P.; Weber, D.; Keller, L.F. Monitoring introgression in European wildcats in the Swiss Jura. *Conserv. Genet.* **2014**, *15*, 1219–1230. [CrossRef]
87. Leonard, J.A.; Echegaray, J.; Rand, E.; Vilà, C. Impact of hybridization with domestic dogs on the conservation of wild canids. *Free Ranging Dogs Wildl. Conserv.* **2013**, *170*, 170–184. [CrossRef]
88. Ottoni, C.; Van Neer, W.; De Cupere, B.; Daligault, J.; Guimaraes, S.; Peters, J.; Spassov, N.; Prendergast, M.E.; Boivin, N.; Morales-Muñiz, A. The palaeogenetics of cat dispersal in the ancient world. *Nat. Ecol. Evol.* **2017**, *1*, 1–7. [CrossRef]
89. Vigne, J.-D.; Guilaine, J.; Debue, K.; Haye, L.; Gérard, P. Early taming of the cat in Cyprus. *Science* **2004**, *304*, 259. [CrossRef]
90. Gompper, M.E. The dog-human-wildlife interface: Assessing the scope of the problem. *Free Ranging Dogs Wildl. Conserv.* **2014**, 9–54. [CrossRef]
91. Adams, J.R.; Leonard, J.A.; Waits, L.P. Widespread occurrence of a domestic dog mitochondrial DNA haplotype in southeastern US coyotes. *Mol. Ecol.* **2003**, *12*, 541–546. [CrossRef]

92. Vilà, C.; Maldonado, J.E.; Wayne, R.K. Phylogenetic relationships, evolution, and genetic diversity of the domestic dog. *J. Hered.* **1999**, *90*, 71–77. [CrossRef]
93. Dzialuk, A.; Zastempowska, E.; Skórzewski, R.; Twarużek, M.; Grajewski, J. High domestic pig contribution to the local gene pool of free-living European wild boar: A case study in Poland. *Mammal Res.* **2018**, *63*, 65–71. [CrossRef]
94. Apollonio, M.; Andersen, R.; Putman, R. *European Ungulates and Their Management in the Twenty-First Century*; Cambridge University Press: Cambridge, UK, 2010.
95. Herrero, J.; García-Serrano, A.; García-González, R. Reproductive and demographic parameters in two Iberian wild boar *Sus scrofa* populations. *Mammal Res.* **2008**, *53*, 355–364. [CrossRef]
96. Massei, G.; Kindberg, J.; Licoppe, A.; Gačić, D.; Šprem, N.; Kamler, J.; Baubet, E.; Hohmann, U.; Monaco, A.; Ozoliņš, J. Wild boar populations up, numbers of hunters down? A review of trends and implications for Europe. *Pest Manag. Sci.* **2015**, *71*, 492–500. [CrossRef]
97. Vetter, S.G.; Ruf, T.; Bieber, C.; Arnold, W. What is a mild winter? Regional differences in within-species responses to climate change. *PLoS ONE* **2015**, *10*, e0132178. [CrossRef]
98. Waithman, J.D.; Sweitzer, R.A.; Van Vuren, D.; Drew, J.D.; Brinkhaus, A.J.; Gardner, I.A.; Boyce, W.M. Range expansion, population sizes, and management of wild pigs in California. *J. Wildl. Manag.* **1999**, *63*, 298–308. [CrossRef]
99. Frantz, A.C.; Zachos, F.E.; Kirschning, J.; Cellina, S.; Bertouille, S.; Mamuris, Z.; Koutsogiannouli, E.A.; Burke, T. Genetic evidence for introgression between domestic pigs and wild boars (*Sus scrofa*) in Belgium and Luxembourg: A comparative approach with multiple marker systems. *Biol. J. Linn. Soc.* **2013**, *110*, 104–115. [CrossRef]
100. García, G.; Vergara, J.; Lombardi, R. Genetic characterization and phylogeography of the wild boar *Sus scrofa* introduced into Uruguay. *Genet. Mol. Biol.* **2011**, *34*, 329–337. [CrossRef]
101. Chai, Z.X.; Xin, J.W.; Zhang, C.F.; Zhang, Q.; Li, C.; Zhu, Y.; Cao, H.W.; Wang, H.; Han, J.L.; Ji, Q.M. Whole-genome resequencing provides insights into the evolution and divergence of the native domestic yaks of the Qinghai–Tibet Plateau. *BMC Evol. Biol.* **2020**, *20*, 1–10. [CrossRef]
102. Anderson, E.; Stebbins, G.L., Jr. Hybridization as an evolutionary stimulus. *Evolution* **1954**, *8*, 378–388. [CrossRef]
103. Flores-Manzanero, A.; Valenzuela-Galván, D.; Cuarón, A.D.; Vázquez-Domínguez, E. Conservation genetics of two critically endangered island dwarf carnivores. *Conserv. Genet.* **2021**, 1–15. [CrossRef]
104. Baskin, Y. *A Plague of Rats and Rubbervines: The Growing Threat of Species Invasions*; Island Press: Washington, DC, USA, 2013.
105. McNeely, J.A. *The Great Reshuffling: Human Dimensions of Invasive Alien Species*; IUCN: Gland, Switzerland, 2001.
106. Wittenberg, R.; Cock, M.J. *Invasive Alien Species: A Toolkit of Best Prevention and Management Practices*; CABI: Egham, UK, 2001.
107. Queirós, J.; Gortázar, C.; Alves, P.C. Deciphering anthropogenic effects on the genetic background of the Red deer in the Iberian Peninsula. *Front. Ecol. Evol.* **2020**, *8*, 147. [CrossRef]
108. Csányi, S.; Carranza, J.; Pokorny, B.; Putman, R.; Ryan, M. Valuing ungulates in Europe. In *Behaviour and Management of European Ungulates*; Whittles: Dunbeath, UK, 2014; pp. 13–45.
109. Lever, C. *Naturalized Mammals of the World*; Longman: Harlow, UK, 1985.
110. Kalb, D.M.; Bowman, J.L. A complete history of the establishment of Japanese sika deer on the Delmarva Peninsula: 100 years post-introduction. *Biol. Invasions* **2017**, *19*, 1705–1713. [CrossRef]
111. Krojerová-Prokešová, J.; Barančeková, M.; Kawata, Y.; Oshida, T.; Igota, H.; Koubek, P. Genetic differentiation between introduced Central European sika and source populations in Japan: Effects of isolation and demographic events. *Biol. Invasions* **2017**, *19*, 2125–2141. [CrossRef]
112. Takagi, T.; Matsumoto, Y.; Koda, R.; Tamate, H.B. Bi-directional movement of deer between Tomogashima islands and the western part of the Kii Peninsula, Japan, with special reference to hybridization between the Japanese sika deer (*Cervus nippon centralis*) and the introduced exotic deer. *Mammal Study* **2020**, *45*, 133–141. [CrossRef]
113. Świsłocka, M.; Czajkowska, M.; Matosiuk, M.; Saveljev, A.P.; Ratkiewicz, M.; Borkowska, A. No evidence for recent introgressive hybridization between the European and Siberian roe deer in Poland. *Mamm. Biol.* **2019**, *97*, 59–63. [CrossRef]
114. Mucci, N.; Mattucci, F.; Randi, E. Conservation of threatened local gene pools: Landscape genetics of the Italian roe deer (*Capreolus c. italicus*) populations. *Evol. Ecol. Res.* **2012**, *14*, 897–920.
115. Do Nascimento Schaurich, M.; Lopes, F.R.V.; de Oliveira, L.R. Hybridization phenomenon in cetacean and pinniped species. *Neotrop. Biol. Conserv.* **2012**, *7*, 199–209.
116. Glover, K.A.; Kanda, N.; Haug, T.; Pastene, L.A.; Øien, N.; Seliussen, B.B.; Sørvik, A.G.; Skaug, H.J. Hybrids between common and Antarctic minke whales are fertile and can back-cross. *BMC Genet.* **2013**, *14*, 1–11. [CrossRef]
117. Crossman, C.A.; Taylor, E.B.; Barrett-Lennard, L.G. Hybridization in the Cetacea: Widespread occurrence and associated morphological, behavioral, and ecological factors. *Ecol. Evol.* **2016**, *6*, 1293–1303. [CrossRef]
118. Miralles, L.; Lens, S.; Rodriguez-Folgar, A.; Carrillo, M.; Martin, V.; Mikkelsen, B.; Garcia-Vazquez, E. Interspecific introgression in cetaceans: DNA markers reveal post-F1 status of a pilot whale. *PLoS ONE* **2013**, *8*, e69511. [CrossRef] [PubMed]
119. Guo, W.; Sun, D.; Cao, Y.; Xiao, L.; Huang, X.; Ren, W.; Xu, S.; Yang, G. Extensive Interspecific Gene Flow Shaped Complex Evolutionary History and Underestimated Species Diversity in Rapidly Radiated Dolphins. *J. Mamm. Evol.* **2021**, 1–15. [CrossRef]
120. Bérubé, M.; Palsbøll, P.J. Hybridism. In *Encyclopedia of Marine Mammals*; Elsevier: Amsterdam, The Netherlands, 2018; pp. 496–501.
121. Willis, P.M.; Crespi, B.J.; Dill, L.M.; Baird, R.W.; Hanson, M.B. Natural hybridization between Dall's porpoises (*Phocoenoides dalli*) and harbour porpoises (*Phocoena phocoena*). *Can. J. Zool.* **2004**, *82*, 828–834. [CrossRef]

122. Brown, A.M.; Kopps, A.M.; Allen, S.J.; Bejder, L.; Littleford-Colquhoun, B.; Parra, G.J.; Cagnazzi, D.; Thiele, D.; Palmer, C.; Frere, C.H. Population differentiation and hybridisation of Australian snubfin (*Orcaella heinsohni*) and Indo-Pacific humpback (*Sousa chinensis*) dolphins in north-western Australia. *PLoS ONE* **2014**, *9*, e101427. [CrossRef]
123. Pampoulie, C.; Gíslason, D.; Ólafsdóttir, G.; Chosson, V.; Halldórsson, S.D.; Mariani, S.; Elvarsson, B.Þ.; Rasmussen, M.H.; Iversen, M.R.; Daníelsdóttir, A.K. Evidence of unidirectional hybridization and second-generation adult hybrid between the two largest animals on Earth, the fin and blue whales. *Evol. Appl.* **2021**, *14*, 314–321. [CrossRef]
124. Quilodrán, C.S.; Montoya-Burgos, J.I.; Currat, M. Harmonizing hybridization dissonance in conservation. *Commun. Biol.* **2020**, *3*, 1–10. [CrossRef] [PubMed]
125. Allendorf, F.W.; Leary, R.F.; Spruell, P.; Wenburg, J.K. The problems with hybrids: Setting conservation guidelines. *Trends Ecol. Evol.* **2001**, *16*, 613–622. [CrossRef]
126. Frankham, R.; Ballou, J.D.; Eldridge, M.D.; Lacy, R.C.; Ralls, K.; Dudash, M.R.; Fenster, C.B. Predicting the probability of outbreeding depression. *Conserv. Biol.* **2011**, *25*, 465–475. [CrossRef]
127. Brekke, T.D.; Henry, L.A.; Good, J.M. Genomic imprinting, disrupted placental expression, and speciation. *Evolution* **2016**, *70*, 2690–2703. [CrossRef]
128. van Wyk, A.M.; Kotzé, A.; Randi, E.; Dalton, D.L. A hybrid dilemma: A molecular investigation of South African bontebok (*Damaliscus pygargus pygargus*) and blesbok (*Damaliscus pygargus phillipsi*). *Conserv. Genet.* **2013**, *14*, 589–599. [CrossRef]
129. Bozarth, C.A.; Hailer, F.; Rockwood, L.L.; Edwards, C.W.; Maldonado, J.E. Coyote colonization of northern Virginia and admixture with Great Lakes wolves. *J. Mammal.* **2011**, *92*, 1070–1080. [CrossRef]
130. Garcia-Elfring, A.; Barrett, R.; Combs, M.; Davies, T.; Munshi-South, J.; Millien, V. Admixture on the northern front: Population genomics of range expansion in the white-footed mouse (*Peromyscus leucopus*) and secondary contact with the deer mouse (*Peromyscus maniculatus*). *Heredity* **2017**, *119*, 447–458. [CrossRef]
131. Brekke, T.D.; Good, J.M. Parent-of-origin growth effects and the evolution of hybrid inviability in dwarf hamsters. *Evolution* **2014**, *68*, 3134–3148. [CrossRef] [PubMed]
132. Neaves, L.E.; Zenger, K.; Cooper, D.W.; Eldridge, M. Molecular detection of hybridization between sympatric kangaroo species in south-eastern Australia. *Heredity* **2010**, *104*, 502–512. [CrossRef]
133. Templeton, A.R.; Hemmer, H.; Mace, G.; Seal, U.S.; Shields, W.M.; Woodruff, D.S. Local adaptation, coadaptation, and population boundaries. *Zoo Biol.* **1986**, *5*, 115–125. [CrossRef]
134. Bell, D.A.; Robinson, Z.L.; Funk, W.C.; Fitzpatrick, S.W.; Allendorf, F.W.; Tallmon, D.A.; Whiteley, A.R. The exciting potential and remaining uncertainties of genetic rescue. *Trends Ecol. Evol.* **2019**, *34*, 1070–1079. [CrossRef]
135. Mallil, K.; Justy, F.; Rueness, E.K.; Dufour, S.; Totis, T.; Bloch, C.; Baarman, J.; Amroun, M.; Gaubert, P. Population genetics of the African wolf (*Canis lupaster*) across its range: First evidence of hybridization with domestic dogs in Africa. *Mamm. Biol.* **2020**, *100*, 645–658. [CrossRef]
136. Smith, S.L.; Senn, H.V.; Pérez-Espona, S.; Wyman, M.T.; Heap, E.; Pemberton, J.M. Introgression of exotic *Cervus* (*nippon and canadensis*) into red deer (*Cervus elaphus*) populations in Scotland and the English Lake District. *Ecol. Evol.* **2018**, *8*, 2122–2134. [CrossRef]
137. Bifolchi, A.; Picard, D.; Lemaire, C.; Cormier, J.; Pagano, A. Evidence of admixture between differentiated genetic pools at a regional scale in an invasive carnivore. *Conserv. Genet.* **2010**, *11*, 1–9. [CrossRef]
138. Zhigileva, O.N.; Uslamina, I.M.; Gimranov, D.O.; Chernova, A.A. Mitochondrial DNA markers for the study of introgression between the sable and the pine marten. *Conserv. Genet. Resour.* **2020**, *12*, 329–336. [CrossRef]
139. Pilot, M.; Moura, A.E.; Okhlopkov, I.M.; Mamaev, N.V.; Manaseryan, N.H.; Hayrapetyan, V.; Kopaliani, N.; Tsingarska, E.; Alagaili, A.N.; Mohammed, O.B. Human-modified canids in human-modified landscapes: The evolutionary consequences of hybridisation for grey wolves and free-ranging domestic dogs. *Evol. Appl.* **2021**, *14*, 2433–2456. [CrossRef]
140. Gómez, J.M.; González-Megías, A.; Lorite, J.; Abdelaziz, M.; Perfectti, F. The silent extinction: Climate change and the potential hybridization-mediated extinction of endemic high-mountain plants. *Biodivers. Conserv.* **2015**, *24*, 1843–1857. [CrossRef]
141. La Morgia, V.; Venturino, E. Understanding hybridization and competition processes between hare species: Implications for conservation and management on the basis of a mathematical model. *Ecol. Model.* **2017**, *364*, 13–24. [CrossRef]
142. Godwin, J.L.; Lumley, A.J.; Michalczyk, Ł.; Martin, O.Y.; Gage, M.J. Mating patterns influence vulnerability to the extinction vortex. *Glob. Chang. Biol.* **2020**, *26*, 4226–4239. [CrossRef] [PubMed]
143. Soulé, M.; Gilpin, M.; Conway, W.; Foose, T. The millenium ark: How long a voyage, how many staterooms, how many passengers? *Zoo Biol.* **1986**, *5*, 101–113. [CrossRef]
144. Ralls, K.; Ballou, J.D.; Frankham, R. Inbreeding and outbreeding. In *Encyclopedia of Biodiversity*; Elsevier: Amsterdam, The Netherlands, 2001.
145. Biedrzycka, A.; Solarz, W.; Okarma, H. Hybridization between native and introduced species of deer in Eastern Europe. *J. Mammal.* **2012**, *93*, 1331–1341. [CrossRef]
146. Muhlfeld, C.C.; Kalinowski, S.T.; McMahon, T.E.; Taper, M.L.; Painter, S.; Leary, R.F.; Allendorf, F.W. Hybridization rapidly reduces fitness of a native trout in the wild. *Biol. Lett.* **2009**, *5*, 328–331. [CrossRef]
147. Templeton, A.R. Coadaptation and outbreeding depression. In *Conservation Biology: The Science of Scarcity and Diversity*; Sinauer Associates: Sunderland, MA, USA, 1986; pp. 105–116.

148. Piorno, V.; Villafuerte, R.; Branco, M.; Carneiro, M.; Ferrand, N.; Alves, P. Low persistence in nature of captive reared rabbits after restocking operations. *Eur. J. Wildl. Res.* **2015**, *61*, 591–599. [CrossRef]
149. Berejikian, B.A.; Ford, M.J. Review of Relative Fitness of Hatchery and Natural Salmon. 2004. Available online: https://www.webapps.nwfsc.noaa.gov/assets/25/6429_02012005_154209_fitnesstm61final.pdf (accessed on 11 November 2021).
150. Fraser, D. Understanding animal welfare. *Acta Vet. Scand.* **2008**, *50*, 1–7. [CrossRef] [PubMed]
151. Kiik, K.; Maran, T.; Nagl, A.; Ashford, K.; Tammaru, T. The causes of the low breeding success of European mink (*Mustela lutreola*) in captivity. *Zoo Biol.* **2013**, *32*, 387–393. [CrossRef]
152. Maran, T.; Põdra, M.; Põlma, M.; Macdonald, D.W. The survival of captive-born animals in restoration programmes–Case study of the endangered European mink *Mustela lutreola*. *Biol. Conserv.* **2009**, *142*, 1685–1692. [CrossRef]
153. Randi, E. Detecting hybridization between wild species and their domesticated relatives. *Mol. Ecol.* **2008**, *17*, 285–293. [CrossRef]
154. Wierzbicki, H.; Zatoń-Dobrowolska, M.; Mucha, A.; Moska, M. Insight into the Genetic Population Structure of Wild Red Foxes in Poland Reveals Low Risk of Genetic Introgression from Escaped Farm Red Foxes. *Genes* **2021**, *12*, 637. [CrossRef] [PubMed]
155. Feulner, P.G.; Gratten, J.; Kijas, J.W.; Visscher, P.M.; Pemberton, J.M.; Slate, J. Introgression and the fate of domesticated genes in a wild mammal population. *Mol. Ecol.* **2013**, *22*, 4210–4221. [CrossRef] [PubMed]
156. Grossen, C.; Keller, L.; Biebach, I.; Consortium, I.G.G.; Croll, D. Introgression from domestic goat generated variation at the major histocompatibility complex of alpine ibex. *PLoS Genet.* **2014**, *10*, e1004438. [CrossRef]
157. Ackermann, R.R.; Brink, J.S.; Vrahimis, S.; De Klerk, B. Hybrid wildebeest (Artiodactyla: Bovidae) provide further evidence for shared signatures of admixture in mammalian crania. *S. Afr. J. Sci.* **2010**, *106*, 1–5. [CrossRef]
158. Tobler, M.; Carson, E.W. Environmental variation, hybridization, and phenotypic diversification in Cuatro *Ciénegas pupfishes*. *J. Evol. Biol.* **2010**, *23*, 1475–1489. [CrossRef] [PubMed]
159. Orr, H.A. Developmental anomalies in Drosophila hybrids are apparently caused by loss of microchromosome. *Heredity* **1990**, *64*, 255–262. [CrossRef]
160. Ackermann, R.R. Phenotypic traits of primate hybrids: Recognizing admixture in the fossil record. *Evol. Anthropol. Issues News Rev.* **2010**, *19*, 258–270. [CrossRef]
161. Falconer, D.S. *Introduction to Quantitative Genetics*; Pearson Education India: London, UK, 1996.
162. Stelkens, R.B.; Schmid, C.; Selz, O.; Seehausen, O. Phenotypic novelty in experimental hybrids is predicted by the genetic distance between species of cichlid fish. *BMC Evol. Biol.* **2009**, *9*, 1–13. [CrossRef]
163. Tung, J.; Charpentier, M.J.; Mukherjee, S.; Altmann, J.; Alberts, S.C. Genetic effects on mating success and partner choice in a social mammal. *Am. Nat.* **2012**, *180*, 113–129. [CrossRef]
164. Zorenko, T.A.; Atanasov, N.; Golenishchev, F.N. Behavioral differentiation and hybridization of the European and Asian forms of Harting'vole *Microtus hartingi* (Rodentia, Arvicolinae). *Russ. J. Theriol.* **2016**, *15*, 133–150. [CrossRef]
165. Chan, W.Y.; Hoffmann, A.A.; van Oppen, M.J. Hybridization as a conservation management tool. *Conserv. Lett.* **2019**, *12*, e12652. [CrossRef]
166. Whiteley, A.R.; Fitzpatrick, S.W.; Funk, W.C.; Tallmon, D.A. Genetic rescue to the rescue. *Trends Ecol. Evol.* **2015**, *30*, 42–49. [CrossRef]
167. Ralls, K.; Sunnucks, P.; Lacy, R.C.; Frankham, R. Genetic rescue: A critique of the evidence supports maximizing genetic diversity rather than minimizing the introduction of putatively harmful genetic variation. *Biol. Conserv.* **2020**, *251*, 108784. [CrossRef]
168. Salzburger, W.; Baric, S.; Sturmbauer, C. Speciation via introgressive hybridization in East African cichlids? *Mol. Ecol.* **2002**, *11*, 619–625. [CrossRef] [PubMed]
169. Mohammadi, Z.; Aliabadian, M.; Ghorbani, F.; Moghaddam, F.Y.; Lissovsky, A.A.; Obst, M.; Olsson, U. Unidirectional Introgression and Evidence of Hybrid Superiority over Parental Populations in Eastern Iranian Plateau Population of Hares (Mammalia: Lepus Linnaeus, 1758). *J. Mamm. Evol.* **2020**, *27*, 723–743. [CrossRef]
170. Chen, Z.-H.; Xu, Y.-X.; Xie, X.-L.; Wang, D.-F.; Aguilar-Gómez, D.; Liu, G.-J.; Li, X.; Esmailizadeh, A.; Rezaei, V.; Kantanen, J. Whole-genome sequence analysis unveils different origins of European and Asiatic mouflon and domestication-related genes in sheep. *Commun. Biol.* **2021**, *4*, 1–15. [CrossRef]
171. Okarma, H.; Jędrzejewska, B.; Jędrzejewski, W.; Krasiński, Z.A.; Miłkowski, L. The roles of predation, snow cover, acorn crop, and man-related factors on ungulate mortality in Białowieża Primeval Forest, Poland. *Acta Theriol.* **1995**, *40*, 197–217. [CrossRef]
172. Kays, R.; Curtis, A.; Kirchman, J.J. Rapid adaptive evolution of northeastern coyotes via hybridization with wolves. *Biol. Lett.* **2010**, *6*, 89–93. [CrossRef]
173. Barlow, A.; Cahill, J.A.; Hartmann, S.; Theunert, C.; Xenikoudakis, G.; Fortes, G.G.; Paijmans, J.L.; Rabeder, G.; Frischauf, C.; Grandal-d'Anglade, A. Partial genomic survival of cave bears in living brown bears. *Nat. Ecol. Evol.* **2018**, *2*, 1563–1570. [CrossRef] [PubMed]
174. Racimo, F.; Sankararaman, S.; Nielsen, R.; Huerta-Sánchez, E. Evidence for archaic adaptive introgression in humans. *Nat. Rev. Genet.* **2015**, *16*, 359–371. [CrossRef]
175. Wang, M.-S.; Wang, S.; Li, Y.; Jhala, Y.; Thakur, M.; Otecko, N.O.; Si, J.-F.; Chen, H.-M.; Shapiro, B.; Nielsen, R. Ancient hybridization with an unknown population facilitated high-altitude adaptation of canids. *Mol. Biol. Evol.* **2020**, *37*, 2616–2629. [CrossRef]
176. Ferreira, M.S.; Jones, M.R.; Callahan, C.M.; Farelo, L.; Tolesa, Z.; Suchentrunk, F.; Boursot, P.; Mills, L.S.; Alves, P.C.; Good, J.M. The legacy of recurrent introgression during the radiation of hares. *Syst. Biol.* **2021**, *70*, 593–607. [CrossRef]

177. Sankararaman, S.; Mallick, S.; Patterson, N.; Reich, D. The combined landscape of Denisovan and Neanderthal ancestry in present-day humans. *Curr. Biol.* **2016**, *26*, 1241–1247. [CrossRef] [PubMed]
178. Schumer, M.; Rosenthal, G.G.; Andolfatto, P. How common is homoploid hybrid speciation? *Evolution* **2014**, *68*, 1553–1560. [CrossRef]
179. Macholán, M. Hybrid Zone, Mouse. In *Brenner's Encyclopedia of Genetics*, 2nd ed.; Maloy, S., Hughes, K., Eds.; Academic Press: Cambridge, MA, USA, 2013; pp. 588–591.
180. Chang, S.W.; Oshida, T.; Endo, H.; Nguyen, S.; Dang, C.; Nguyen, D.; Jiang, X.; Li, Z.J.; Lin, L.K. Ancient hybridization and underestimated species diversity in Asian striped squirrels (genus *Tamiops*): Inference from paternal, maternal and biparental markers. *J. Zool.* **2011**, *285*, 128–138. [CrossRef]
181. Detwiler, K.M. Mitochondrial DNA analyses of Cercopithecus monkeys reveal a localized hybrid origin for *C. mitis doggetti* in Gombe National Park, Tanzania. *Int. J. Primatol.* **2019**, *40*, 28–52. [CrossRef]
182. Macholán, M.; Baird, S.J.; Dufková, P.; Munclinger, P.; Bímová, B.V.; Piálek, J. Assessing multilocus introgression patterns: A case study on the mouse X chromosome in central Europe. *Evol. Int. J. Org. Evol.* **2011**, *65*, 1428–1446. [CrossRef]
183. Allen, W.L.; Stevens, M.; Higham, J.P. Character displacement of Cercopithecini primate visual signals. *Nat. Commun.* **2014**, *5*, 1–10. [CrossRef] [PubMed]
184. Komárek, J.; Komárková-Legnerová, J. Phenotype diversity of the cyanoprokaryotic genus *Cylindrospermopsis* (Nostocales). *Czech Phycol.* **2003**, *3*, 1–30.
185. Ito, T.; Kawamoto, Y.; Hamada, Y.; Nishimura, T.D. Maxillary sinus variation in hybrid macaques: Implications for the genetic basis of craniofacial pneumatization. *Biol. J. Linn. Soc.* **2015**, *115*, 333–347. [CrossRef]
186. Grant, P.R.; Grant, B.R. Phenotypic and genetic effects of hybridization in Darwin's finches. *Evolution* **1994**, *48*, 297–316. [CrossRef]
187. Gridley, T.; Elwen, S.H.; Harris, G.; Moore, D.; Hoelzel, A.; Lampen, F. Hybridization in bottlenose dolphins—A case study of *Tursiops aduncus* × *T. truncatus* hybrids and successful backcross hybridization events. *PLoS ONE* **2018**, *13*, e0201722. [CrossRef] [PubMed]
188. Cserkész, T.; Kiss, C.; Barkaszi, Z.; Görföl, T.; Zagorodniuk, I.; Sramkó, G.; Csorba, G. Intra-and interspecific morphological variation in sympatric and allopatric populations of *Mustela putorius* and *M. eversmanii* (Carnivora: Mustelidae) and detection of potential hybrids. *Mammal Res.* **2021**, *66*, 103–114. [CrossRef]
189. Balcarcel, A.; Sánchez-Villagra, M.R.; Segura, V.; Evin, A. Singular patterns of skull shape and brain size change in the domestication of South American camelids. *J. Mammal.* **2021**, *102*, 220–235. [CrossRef]
190. Frare, C.F.; Matocq, M.D.; Feldman, C.R.; White, A.M.; Manley, P.N.; Jermstad, K.D.; Hekkala, E.R. Landscape disturbance and sporadic hybridization complicate field identification of chipmunks. *J. Wildl. Manag.* **2017**, *81*, 248–258. [CrossRef]
191. Barton, N.H.; Hewitt, G.M. Analysis of hybrid zones. *Annu. Rev. Ecol. Syst.* **1985**, *16*, 113–148. [CrossRef]
192. Barton, N.H.; Hewitt, G.M. Adaptation, speciation and hybrid zones. *Nature* **1989**, *341*, 497–503. [CrossRef]
193. Harrison, R. *Hybrid Zones and the Evolutionary Process*; Oxford University Press: New York, NY, USA, 1993.
194. Woodruff, D.S. Natural hybridization and hybrid zones. *Syst. Biol.* **1973**, *22*, 213–218. [CrossRef]
195. Mauldin, M.R.; Haynie, M.L.; Hanson, J.D.; Baker, R.J.; Bradley, R.D. Multilocus characterization of a woodrat (genus *Neotoma*) hybrid zone. *J. Hered.* **2014**, *105*, 466–476. [CrossRef]
196. Shurtliff, Q.R.; Murphy, P.J.; Matocq, M.D. Ecological segregation in a small mammal hybrid zone: Habitat-specific mating opportunities and selection against hybrids restrict gene flow on a fine spatial scale. *Evolution* **2014**, *68*, 729–742. [CrossRef]
197. Ďureje, Ľ.; Macholán, M.; Baird, S.J.; Piálek, J. The mouse hybrid zone in Central Europe: From morphology to molecules. *J. Vertebr. Biol.* **2012**, *61*, 308–318. [CrossRef]
198. Coyner, B.S.; Murphy, P.J.; Matocq, M.D. Hybridization and asymmetric introgression across a narrow zone of contact between *Neotoma fuscipes* and *N. macrotis* (Rodentia: Cricetidae). *Biol. J. Linn. Soc.* **2015**, *115*, 162–172. [CrossRef]
199. Jahner, J.P.; Parchman, T.L.; Matocq, M.D. Multigenerational backcrossing and introgression between two woodrat species at an abrupt ecological transition. *Mol. Ecol.* **2021**, *30*, 4245–4258. [CrossRef]
200. Mauldin, M.R.; Haynie, M.L.; Vrla, S.C.; Bradley, R.D. Temporal evaluation of a woodrat (genus *Neotoma*) hybrid zone based on genotypic and georeferenced data. *J. Mammal.* **2021**, *102*, 541–557. [CrossRef]
201. Brandler, O.; Kapustina, S.; Nikol'skii, A.; Kolesnikov, V.; Badmaev, B.; Adiya, Y. A study of hybridization between *Marmota baibacina* and *M. sibirica* in their secondary contact zone in Mongolian Altai. *Front. Ecol. Evol.* **2021**, *9*, 363. [CrossRef]
202. Gravena, W.; Da Silva, V.M.; Da Silva, M.N.; Farias, I.P.; Hrbek, T. Living between rapids: Genetic structure and hybridization in botos (Cetacea: Iniidae: *Inia* spp.) of the Madeira River, Brazil. *Biol. J. Linn. Soc.* **2015**, *114*, 764–777. [CrossRef]
203. Campbell, C.D.; Cowan, P.; Gruber, B.; MacDonald, A.J.; Holleley, C.E.; Sarre, S.D. Has the introduction of two subspecies generated dispersal barriers among invasive possums in New Zealand? *Biol. Invasions* **2021**, *23*, 3831–3845. [CrossRef]
204. Eldridge, M.D.; Pearson, D.J.; Potter, S. Identification of a novel hybrid zone within the black-footed rock-wallaby (*Petrogale lateralis*) in Western Australia. *Aust. J. Zool.* **2021**, *68*, 98–107. [CrossRef]
205. Kinoshita, E.; Abramov, A.V.; Soloviev, V.A.; Saveljev, A.P.; Nishita, Y.; Kaneko, Y.; Masuda, R. Hybridization between the European and Asian badgers (Meles, Carnivora) in the Volga-Kama region, revealed by analyses of maternally, paternally and biparentally inherited genes. *Mamm. Biol.* **2019**, *94*, 140–148. [CrossRef]

206. Baird, A.B.; Braun, J.K.; Engstrom, M.D.; Holbert, A.C.; Huerta, M.G.; Lim, B.K.; Mares, M.A.; Patton, J.C.; Bickham, J.W. Nuclear and mtDNA phylogenetic analyses clarify the evolutionary history of two species of native Hawaiian bats and the taxonomy of Lasiurini (Mammalia: Chiroptera). *PLoS ONE* **2017**, *12*, e0186085. [CrossRef]
207. Lorenzini, R.; Fanelli, R.; Grifoni, G.; Scholl, F.; Fico, R. Wolf-dog crossbreeding: "Smelling" a hybrid may not be easy. *Mamm. Biol.* **2014**, *79*, 149–156. [CrossRef]
208. Thompson, C.W.; Pfau, R.; Choate, J.R.; Genoways, H.H.; Finck, E.J. Identification and characterization of the contact zone between short-tailed shrews (Blarina) in Iowa and Missouri. *Can. J. Zool.* **2011**, *89*, 278–288. [CrossRef]
209. Andriollo, T.; Ashrafi, S.; Arlettaz, R.; Ruedi, M. Porous barriers? Assessment of gene flow within and among sympatric long-eared bat species. *Ecol. Evol.* **2018**, *8*, 12841–12854. [CrossRef]
210. Arbogast, B.S.; Schumacher, K.I.; Kerhoulas, N.J.; Bidlack, A.L.; Cook, J.A.; Kenagy, G. Genetic data reveal a cryptic species of New World flying squirrel: *Glaucomys oregonensis*. *J. Mammal.* **2017**, *98*, 1027–1041. [CrossRef]
211. Yannic, G.; Statham, M.J.; Denoyelle, L.; Szor, G.; Qulaut, G.Q.; Sacks, B.N.; Lecomte, N. Investigating the ancestry of putative hybrids: Are Arctic fox and red fox hybridizing? *Polar Biol.* **2017**, *40*, 2055–2062. [CrossRef]
212. Mengoni, C.; Mucci, N.; Randi, E. Genetic diversity and no evidences of recent hybridization in the endemic Italian hare (*Lepus corsicanus*). *Conserv. Genet.* **2015**, *16*, 477–489. [CrossRef]
213. Prevosti, F.J.; Ramírez, M.A.; Schiaffini, M.; Martin, F.; Udrizar Sauthier, D.E.; Carrera, M.; Sillero-Zubiri, C.; Pardiñas, U.F. Extinctions in near time: New radiocarbon dates point to a very recent disappearance of the South American fox *Dusicyon avus* (Carnivora: Canidae). *Biol. J. Linn. Soc.* **2015**, *116*, 704–720. [CrossRef]
214. Bell, K.C.; Van Gunst, J.; Teglas, M.B.; Hsueh, J.; Matocq, M.D. Lost in a sagebrush sea: Comparative genetic assessment of an isolated montane population of *Tamias amoenus*. *J. Mammal.* **2021**, *102*, 173–187. [CrossRef]
215. Zeng, L.; Liu, H.-Q.; Tu, X.-L.; Ji, C.-M.; Gou, X.; Esmailizadeh, A.; Wang, S.; Wang, M.-S.; Wang, M.-C.; Li, X.-L. Genomes reveal selective sweeps in kiang and donkey for high-altitude adaptation. *Zool. Res.* **2021**, *42*, 450. [CrossRef]
216. Fabbri, E.; Velli, E.; D'Amico, F.; Galaverni, M.; Mastrogiuseppe, L.; Mattucci, F.; Caniglia, R. From predation to management: Monitoring wolf distribution and understanding depredation patterns from attacks on livestock. *Hystrix Ital. J. Mammal.* **2018**, *29*, 101–110.
217. Inoue, T.; Murakami, T.; Abramov, A.V.; Masuda, R. Mitochondrial DNA control region variations in the sable *Martes zibellina* of Hokkaido Island and the Eurasian continent, compared with the Japanese marten *M. melampus*. *Mammal Study* **2010**, *35*, 145–155. [CrossRef]
218. Thomsen, C.L.; Andersen, L.W.; Stronen, A.V. Forensic DNA analyses suggest illegal trade of canid skins. *Mammal Res.* **2016**, *61*, 423–426. [CrossRef]
219. Eckert, I.; Suchentrunk, F.; Markov, G.; Hartl, G.B. Genetic diversity and integrity of German wildcat (*Felis silvestris*) populations as revealed by microsatellites, allozymes, and mitochondrial DNA sequences. *Mamm. Biol.* **2010**, *75*, 160–174. [CrossRef]
220. Leite, J.V.; Álvares, F.; Velo-Antón, G.; Brito, J.C.; Godinho, R. Differentiation of North African foxes and population genetic dynamics in the desert—Insights into the evolutionary history of two sister taxa, *Vulpes rueppellii* and *Vulpes vulpes*. *Org. Divers. Evol.* **2015**, *15*, 731–745. [CrossRef]
221. Sierra, A.B.A.; Castillo, E.R.; Labaroni, C.; Barrandeguy, M.E.; Martí, D.A.; Ojeda, R.; Lanzone, C. Genetic studies in the recently divergent *Eligmodontia puerulus* and *E. moreni* (Rodentia, Cricetidae, Sigmodontinae) from Puna and Monte deserts of South America. *Mamm. Biol.* **2017**, *87*, 93–100. [CrossRef]
222. Lawson, L.P.; Castruita, J.A.S.; Haile, J.S.; Vernesi, C.; Rovero, F.; Lorenzen, E.D. Unraveling elephant-shrews: Phylogenetic relationships and unexpected introgression among giant sengis. *Mol. Phylogenet. Evol.* **2021**, *154*, 107001. [CrossRef] [PubMed]
223. Morales, A.E.; Fenton, M.B.; Carstens, B.C.; Simmons, N.B. Comment on "Population genetics reveal *Myotis keenii* (Keen's myotis) and *Myotis evotis* (long-eared myotis) to be a single species". *Can. J. Zool.* **2021**, *99*, 415–422. [CrossRef]
224. Nagata, J.; Yasuda, M.; Yamashiro, A. Genetic analysis of a newly established deer population expanding in the Sasebo area in Nagasaki Prefecture, Japan reveals no evidence of genetic disturbance by Formosan sika deer. *Mamm. Study* **2021**, *46*, 1–13. [CrossRef]
225. Sarver, B.A.; Herrera, N.D.; Sneddon, D.; Hunter, S.S.; Settles, M.L.; Kronenberg, Z.; Demboski, J.R.; Good, J.M.; Sullivan, J. Diversification, introgression, and rampant cytonuclear discordance in rocky mountains chipmunks (sciuridae: *Tamias*). *Syst. Biol.* **2021**, *70*, 908–921. [CrossRef]
226. Korablev, M.P.; Korablev, N.P.; Korablev, P.N. Genetic diversity and population structure of the grey wolf (*Canis lupus* Linnaeus, 1758) and evidence of wolf× dog hybridisation in the centre of European Russia. *Mamm. Biol.* **2021**, *101*, 91–104. [CrossRef]
227. van Wyk, A.M.; Dalton, D.L.; Hoban, S.; Bruford, M.W.; Russo, I.R.M.; Birss, C.; Grobler, P.; van Vuuren, B.J.; Kotzé, A. Quantitative evaluation of hybridization and the impact on biodiversity conservation. *Ecol. Evol.* **2017**, *7*, 320–330. [CrossRef]
228. Furman, A.; Coraman, E.; Çelik, Y.E.; Postawa, T.; Bachanek, J.; Ruedi, M. Cytonuclear discordance and the species status of *Myotis myotis* and *Myotis blythii* (Chiroptera). *Zool. Scr.* **2014**, *43*, 549–561. [CrossRef]
229. Furman, A.; Çelik, Y.E.; Çoraman, E.; Bilgin, R. Reproductive isolation and morphological discrimination of *Myotis myotis macrocephalicus* and *M. blythii sl* (Chiroptera: Vespertilionidae) in Turkey. *Acta Chiropterol.* **2020**, *22*, 21–28. [CrossRef]
230. Burgarella, C.; Barnaud, A.; Kane, N.A.; Jankowski, F.; Scarcelli, N.; Billot, C.; Vigouroux, Y.; Berthouly-Salazar, C. Adaptive introgression: An untapped evolutionary mechanism for crop adaptation. *Front. Plant Sci.* **2019**, *10*, 4. [CrossRef] [PubMed]

231. Beaumont, M.A.; Balding, D.J. Identifying adaptive genetic divergence among populations from genome scans. *Mol. Ecol.* **2004**, *13*, 969–980. [CrossRef]
232. Excoffier, L.; Hofer, T.; Foll, M. Detecting loci under selection in a hierarchically structured population. *Heredity* **2009**, *103*, 285–298. [CrossRef]
233. Browett, S.S.; O'Meara, D.B.; McDevitt, A.D. Genetic tools in the management of invasive mammals: Recent trends and future perspectives. *Mamm. Rev.* **2020**, *50*, 200–210. [CrossRef]
234. Mattucci, F.; Galaverni, M.; Lyons, L.A.; Alves, P.C.; Randi, E.; Velli, E.; Pagani, L.; Caniglia, R. Genomic approaches to identify hybrids and estimate admixture times in European wildcat populations. *Sci. Rep.* **2019**, *9*, 1–15.

2 Chapter 2

Evaluation of Global and Local Ancestry Reconstruction Methods for Admixture Detection Using Genome-Wide SNP Data in Genus *Canis*

Abstract

Introgressive hybridization plays an important role in shaping gene pools of populations and species. Investigations of the hybridization process require accurate estimates of the introgression rates. This can be achieved using global or local ancestry estimates. While global ancestry represents the proportion of different ancestral populations across the entire genome, local ancestry identifies the alleles derived from different ancestries at each specific locus, which may be then averaged across loci to obtain the genome-wide estimates. In this study, we aimed to evaluate the methods of global and local ancestry reconstruction in the context of the analysis of introgressive hybridization. For this purpose, a dataset comprising 229120 SNPs from representatives of the genus *Canis* (gray wolves, golden jackals, and free-ranging dogs) was used to perform global ancestry analyses (PCA and ADMIXTURE) and local ancestry analyses (LAMP-LD, ELAI, and GHap). The results obtained from these methods were compared to assess their consistency. Furthermore, factors contributing to discrepancies among the results were evaluated, providing insights into the robustness and limitations of global and local ancestry analyses. We found that the global ancestry method (Admixture) estimated higher admixture proportions than local ancestry methods. The comparison of the three local ancestry methods showed that while the results of LAMP_LD were highly consistent with the results of ELAI in most samples, the ancestry proportions estimated by GHap were lower. The differences between local and global ancestry results can be attributed to their methodological approaches, the type of information they utilize, and their strategies for handling missing data. We emphasize two key factors, low-quality genotypes and the presence of subpopulation structures, as major contributors to inconsistencies between the methods, increasing the uncertainty and variability of estimates. We found that global ancestry analyses such as ADMIXTURE are more likely to be affected by these confounding factors. Therefore, global ancestry methods such as ADMIXTURE may not be suitable as standalone approaches for the precise inference of admixture proportions due to their susceptibility to confounding factors. We recommend a joint use of local and global methods of ancestry analysis, with local ancestry results being prioritized for precise inference of introgression rates.

Keywords: Introgressive hybridization, Global ancestry, Local ancestry, Genus *Canis*

2.1. Introduction

Rapid advances in genome sequencing and computational technologies have shown that introgressive hybridization, interbreeding between representatives of different taxa that results in the transfer of alleles between them, is a common process (Dagilis et al., 2021). Introgression can have a substantial impact on evolutionary patterns in the species affected, ranging from extinction via genetic swamping to hybrid speciation (Todesco et al., 2016; Adavoudi and Pilot, 2021). Given the broad range of potential consequences of introgressive hybridization, it is crucial to obtain accurate estimates of introgression rates, especially in the context of species management and conservation. In recent years, different methods have been developed to estimate introgression based on genetic and genomic data (reviewed in: Wangkumhang and Hellenthal, 2018; Elworth et al., 2019; Hibbins and Hahn, 2021; Dagilis et al., 2021; Thawornwattana et al., 2023; Sun et al., 2025). Methods for inferring introgression can be classified into six main categories based on the type of information they use, including sequence similarity (F_{ST}, d_{xy}), clinal changes along space or the genome, tree topology, clustering analyses, demographic models, and D-statistics and their extensions (Dagilis et al., 2021).

The methodology of inferring introgression and identification of hybrid individuals varies in different studies depending on the type of applied molecular markers. Nuclear microsatellite loci and SNP markers are the main molecular markers that have been used for the identification of hybrids (Goli et al., 2024). However, SNPs can provide more details about hybridization than microsatellite markers because of technologies that enable the simultaneously obtain data from a large number of loci (thousands to millions) spanning the entire genome, via the microarray genotyping or next-generation sequencing. The use of a large number of loci can improve the accuracy of ancestry estimates and enable the identification of older-generation backcrosses (Goli et al., 2024). Therefore, SNPs are currently most suitable markers for studying introgressive hybridization.

Global ancestry inference is one of the widely used approaches for estimating the proportion of ancestry across the entire genome. Global ancestry methods can identify introgression by assigning individuals to their major population or ancestry group and identify those individuals that carry genetic variants originating from different ancestry groups (Padhukasahasram, 2014). Methods for calculating global admixture are classified into model-based and algorithm-based approaches (Padhukasahasram, 2014; Tan and Atkinson, 2023). STRUCTURE (Pritchard et al., 2000), fastSTRUCTURE (Raj, et al., 2014), and ADMIXTURE (Alexander et al., 2009; Alexander and Lange, 2011) are model-based (Bayesian clustering) methods. These are the most commonly used software to infer global admixture proportions and identify the presence of introgression. Both STRUCTURE and ADMIXTURE are based on the probabilistic models and estimate an individual's ancestry and hybridization status using allele frequency differences. Probabilistic quantities of different genetic groups (ancestry coefficients, q-values) can be estimated without any prior taxonomic information (Anderson, 2008). The admixture score (membership coefficient) indicates the proportions of genetic material

that are inherited from parental populations. Non-admixed individuals are expected to have scores of 0 or 1, or closely approaching these values. However, both STRUCTURE and ADMIXTURE rely on arbitrary thresholds and there is no standard way to interpret the membership coefficient. In addition, the assignment membership coefficient might be affected by errors in genotyping (Ottenburghs, 2021). Therefore, arbitrary probability thresholds are usually used to classify individuals as non-admixed, first-generation hybrids, or backcrossed (Randi 2008; Ottenburghs, 2020). Moreover, results could be biased by the number of individuals in each population, as these models are based on allele frequencies. In addition, selecting an optimal value for the number of populations (K parameter) is important to avoid overinterpreting barplots of STRUCTURE and ADMIXTURE results (Lawson et al., 2018). This may be challenging in the case of complex datasets containing multiple species with nested structures within each species.

Algorithmic-based approaches rely on the characteristics of Principal Component Analysis (PCA) and are much faster and easier than model-based approaches (Price et al., 2006). PCA is widely applied for the inference and visualization of population structure (Patterson et al., 2006; Seldin et al., 2011), the analysis of the demographic history and admixture between populations (Elhaik, 2022). PCA is a multivariate analysis that projects samples into a PC space that captures the largest variation, instead of producing an intuitive admixture proportion output (Tan and Atkinson, 2023). In a PCA plot, admixed individuals are scattered between the clusters of the two unadmixed populations in the first two PCs (Ma and Amos, 2012). Also, the relative distance of each admixed individual from the centroids of parental population clusters can be used to estimate the proportions of admixture for each individual (Ma and Amos, 2012). Similar to determining the K parameter in model-based approaches, the choice of the number of PCs has a significant impact on PCA results (Waaij et al., 2023).

Global inference methods provide a broad overview of admixture proportions across the entire genome but lack the resolution to detect admixture at finer scales, such as along individual chromosomes or at specific loci. These methods operate on the assumption that each individual has identical genetic ancestry ratios at every genomic locus (Long, 1991). This constraint results in variability in admixture proportions across loci, leading to discrepancies between local ancestry and overall global ancestry in admixed individuals. The ability to infer admixture at specific chromosomal locations has been significantly advanced by developments in genome sequencing, SNP genotyping technologies, and computational tools (Kidd et al., 2012; Wu et al., 2021; Vi et al., 2023). The genomes of admixed individuals can be described as mosaics of chromosomal segments with different ancestries (Price et al., 2009; Gravel, 2012; Ma et al., 2015). Local ancestry can provide us with information about the proportions of introgression in admixed individuals by identifying alleles derived from different ancestries at a specific locus (Duan et al., 2017; Martin et al., 2018).

Many local ancestry estimation methods have been developed for human population studies (Padhukasahasram, 2014; Yuan et al., 2017; Mazandu et al., 2018; Geza et al., 2019; Wu et al., 2021; Tan and Atkinson, 2023). Hidden Markov Models (HMMs) and Conditional Random Fields (CRFs) are two main methods for inferring local ancestry (see Tan and Atkinson, 2023 for details). Curretly, more than 70% of local ancestry tools

are based on HMMs (Wu et al., 2021). HapMix (Price et al., 2009), ELAI (Guan, 2014), Saber (Tang et al., 2006), LAMP-LD (Baran et al., 2012), and MOSAIC (Salter - Townshend and Myers, 2019) are based on HMMs and utilize genotype data from ancestral populations (also known as source or reference populations) to predict the ancestry classification of admixed individuals by identifying the hidden ancestral states. In contrast, RFMix (Maples et al., 2013) and AICRF (Alizadeh et al., 2023) leverage a conditional random field (CRF) parameterized by random forests trained on reference panels. A CRF is a statistical model that can model the dependencies between multiple variables (Lafferty et al., 2001). For example, in RFMix, a conditional random field (CRF), parameterized by a random forest trained on reference panels, is used to infer local ancestry within each block (Maples et al., 2013). Additionally, some methods rely on different algorithms. For example, GHap (Utsunomiya et al., 2020)employs the K-means clustering method. Ghap can approximate ancestral lineages by grouping all observed haplotypes using the K-means algorithm (Hartigan and Wong, 1979).

The fundamental concepts, methods for estimating local and global ancestry, and the tools/software used for these analyses, along with their applications, have been discussed in detail in recent reviews (e.g., Yuan et al., 2017; Goli et al., 2024). Several studies compared the performance of various methods for inferring admixture proportions, primarily using human genome data. For example, the performance and computational speed of HAPMix and LAMP-LD were evaluated against a newly developed method, the two-layer HMM implemented in software ELAI, using simulated datasets, demonstrating that the two-layer HMM offers certain advantages, such as effectively handling missing data and making fewer errors in small regions spanning a few hundred SNPs (Guan, 2014). In another study, the mathematical and statistical approaches used by different local ancestry deconvolution methods were compared through simulations, providing guidance for researchers in selecting appropriate methods based on their data characteristics (Geza et al., 2019). This analysis showed that ELAI (Guan, 2014) as an effective tool for estimating admixture events.

Additionally, the accuracy, runtime, memory usage, and usability of five local ancestry methods, LAMP-LD, ELAI, RFMix, Loter, and MOSAIC, were compared using both simulated and real human genome data (Schubert et al., 2020). The study found that RFMix demonstrated the highest overall performance among the evaluated methods. However, depending on the specific application, other methods could provide comparable results while offering faster runtimes (Schubert et al., 2020).

While numerous studies evaluated the performance of various local and global ancestry methods, most assessments relied on simulation datasets based on the human genome. Simulated data often fail to capture the full complexity of real genomic data, which may involve varying levels of quality due to poor DNA yield, sequencing errors, and missing data. Additionally, real genomic data are shaped by complex demographic histories and intricate population structures, factors that significantly impact the outcomes of ancestry analysis.

In this study, we aimed to apply different methods to infer the proportions of introgression from both global and local perspectives and evaluate the consistency of results among these methods in a dataset composed of three canids, gray wolves (*Canis*

lupus), golden jackals (*Canis aureus)*, and domestic dogs (*C. lupus familiaris*), known from earlier studies to interbreed (Gopalakrishnan et al. 2018; Wang et al., 2020; Werhahn et al., 2020).). For this purpose, PCA and ADMIXTURE representing non-parametric and model-based approaches, respectively, were applied for global ancestry estimation. Additionally, multiple tools for local ancestry analysis, including LAMP-LD, ELAI, and GHap, each offering distinct advantages and methodological strengths, were applied. LAMP-LD is optimized for high-speed and efficient analysis, particularly suited for datasets with large sample sizes and dense SNP coverage (Baran et al., 2012). ELAI uses a probabilistic framework based on multi-layer hidden Markov models, making it highly effective for analyzing complex admixture histories (Guan, 2014). GHap, in contrast, focuses on haplotype-based analysis, leveraging phased genomic data to provide fine-scale insights into local ancestry patterns (Utsunomiya et al., 2020). These diverse methodological approaches enabled a comprehensive assessment and comparison of the introgression rate results. Moreover, the effects of the quality of genotype data, demographic histories such as population structure within the species studied, on the results of different methods were explored, and methods that have better performance under different conditions were assessed. The SNP genotype data from different populations of gray wolves, golden jackals, and domestic dogs were used to create multiple data sets. We aimed to (1) apply different methods and tools to infer the proportions of introgression in the canid dataset from both global and local perspectives, (2) assess the consistency between the results obtained from different methods, (3) compare the results of global ancestry in the entire dataset versus two different geographic populations (regional datasets), and (4) identify factors that can confound the results of local and global ancestry analyses.

2.2. Material and methods

Sample collection and laboratory procedures

Our samples were collected opportunistically and therefore represented various tissues. Muscle and skin tissue samples were collected from gray wolves, golden jackals, and domestic dogs across their Eurasian range. Additionally, saliva samples from domestic dogs were collected using the PERFORMAgene animal DNA collection kit (DNA Genotek), providing an alternative and non-invasive sampling method. We also incorporated golden jackal tissue samples from a prior study (Rutkowski et al., 2015) to expand the dataset and enhance comparability (Fig. 2.1). DNA was extracted from tissue samples using NucleoSpin Tissue Kit (Macherey Nagel, Duren, Germany), and DNA from saliva samples was extracted by PG-AC extraction kit, following the manufacturer's instructions with slight modifications (for more details see supplementary file).

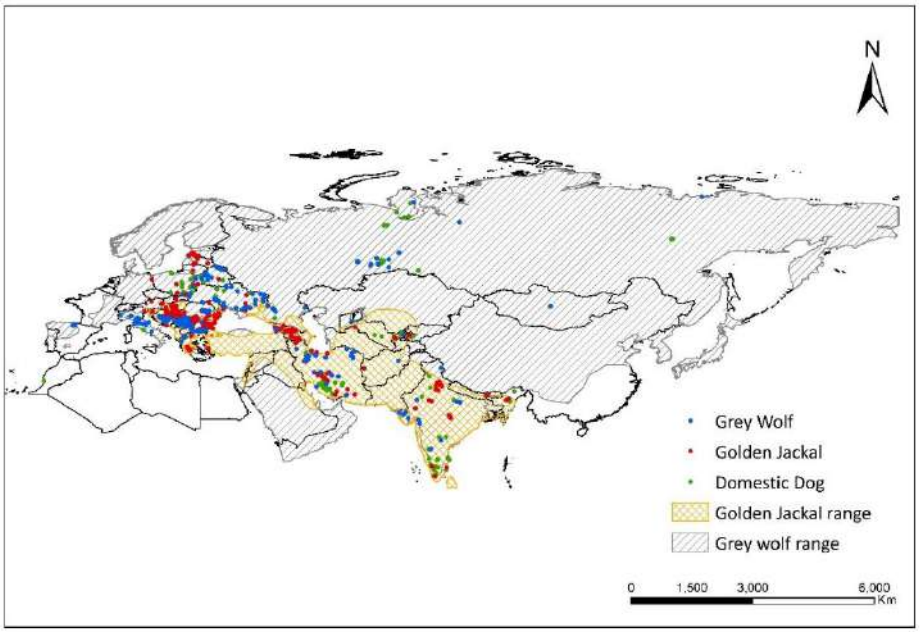


Fig. 2.1. Distribution of genetic samples and spatial range of grey wolves and golden jackals according to the IUCN (IUCN, 2024). Each marked sampling location can represent several samples.

SNP genotyping

The Axiom Canine HD Array and Axiom Canine Genotyping Array Set A (Thermo Fischer Scientific, USA) were used for the genotyping of 1378 and 249 samples, respectively. The Array Set A was used before the HD array was available. Axiom Analysis Suite version 5.1.1.1 (Thermo Fischer Scientific, USA) was applied to carry out SNP genotype calling. Identification of poor-quality samples was done using Dish QC (DQC), and sample QC call rate test. To achieve high-quality genotype data, the Axiom Best Practices Genotyping Analysis Workflow suggests very strict thresholds of 82 and 97% for DQC and QC call rates, respectively. While studies on model organisms can provide high quality data from all sampled individuals, this is not possible in studies on wild or feral animals, where samples are obtained opportunistically and individuals cannot be resampled. Therefore, more relaxed quality parameters were applied here to avoid discarding many individuals.

A two-stage approach in the Axiom Analysis Power Tools command line (version 2.11.4) was used for SNP genotyping of all samples. In the first step, Dish-QC threshold of 0.75 and sample call rate threshold of 80%were applied. In the next step, the genotype accuracy has been improved by using the Ps_CallAdjust SNPolisher function with the SNP confidence score threshold lowered to 0.01 and the SNP calling rate lowered to 90% to improve genotype accuracy as the number of missing calls increased. Samples with QC call rate threshold between 80 and 97% as well as those with DQC values between 0.75 and 0.82 were identified as low-quality samples and were distinguished by appending an "LQ" suffix to their identifiers. Genotypes generated using the Axiom Canine HD Array and the Axiom Canine Genotyping Array Set A, which were processed

separately using the two-stage approach described above, were merged as a single dataset (229,120 SNPs) and exported to the Plink file format.

Dataset creation

After merging individuals genotyped on Axiom arrays, a dataset of 1444 individuals including 338 samples of gray wolf, 483 samples of golden jackal, and 623 samples of free-ranging dog with 229,120 autosomal SNPs was created, hereafter referred to as the WJD dataset. Besides the WJD dataset, three different datasets have been generated for pairwise comparison between gray wolf and domestic dog (WD dataset), golden jackal and domestic dog (JD dataset), and gray wolf and golden jackal (WJ dataset), with the same number of SNPs (Table S 2.1).

Initial data processing

Plink software v 1.9 (Chang et al., 2015) was used for filtering genotypes of all datasets. SNPs with more than 10% missing data and a minor allele frequency (MAF) below 0.01 were removed. A relaxed 20% threshold for filtering out individuals with a high rate of missingness was applied. This allowed us to retain individuals with relatively high proportions of missing data (10-20%) to be kept in the dataset to test for the effect of missing data on the accuracy of admixture inference. The kinship coefficient between each pair of individuals was estimated using the KING (Manichaikul et al. 2010) algorithm in PLINK v 2.0 (Chang et al., 2015). To eliminate duplicate samples, we excluded one individual from each pair with a kinship coefficient exceeding 0.48, as duplicate samples typically exhibit a kinship coefficient of approximately 0.5 (Lee and Chen, 2016). The kinship coefficient is defined as the probability that two homologous alleles drawn from each of two individuals are identical by descent (IBD) (Jiang et al., 2022). Based on the kinship analyses, 58 duplicated individuals were identified and subsequently removed from the dataset. Only data from autosomal chromosomes 1–38 were used. Additionally, to meet the assumption of global ancestry models (such as PCA and Admixture), the WJD dataset was filtered to remove loci in strong linkage disequilibrium (LD; $r^2 > 0.1$), using sliding windows of 50 SNPs with a step size of 3 SNPs, and 172,357 SNPs were removed through linkage disequilibrium pruning. The rest of the datasets, including WD, JD, and WJ, were created after passing all filtering steps. The sample sizes for each dataset and the number of SNPs retained after quality control, categorized by whether global or local ancestry analyses were conducted, are provided in Table S 2.1.

Regional datasets

To analyze the population structure for each canid, Discriminant Analysis of Principal Components (DAPC) was carried out (Fig S 2.1, Fig S 2.2, Fig S 2.3). Based on the results from the population structure, we selected two regional populations to address potential confounding factors arising from population divergence, sample size, intrapopulation structure, and the composition of reference populations' influence on hybridization level estimation. To select regional datasets, we aimed to choose regions

with good sampling coverage from each of the three canids with different genotyping quality and sample sizes.

The first regional dataset included samples from India, consisting of 100 individuals: 21 gray wolves, 33 golden jackals, and 46 dogs (Fig 2.2, Table S 2.2). Nearly half of the samples from India (47 individuals) were low-quality samples (LQ). The second dataset, comprising 327 individuals: 110 gray wolves, 167 golden jackals, and 50 dogs from the Balkans (Bosnia and Herzegovina, Serbia, and Bulgaria; Fig 2.2, Table S 2.2), was selected based on high-quality genotyping data. The data from both datasets were extracted from the main filtered datasets. All subsequent analyses were conducted on both the main and regional datasets (Table S 2.2).

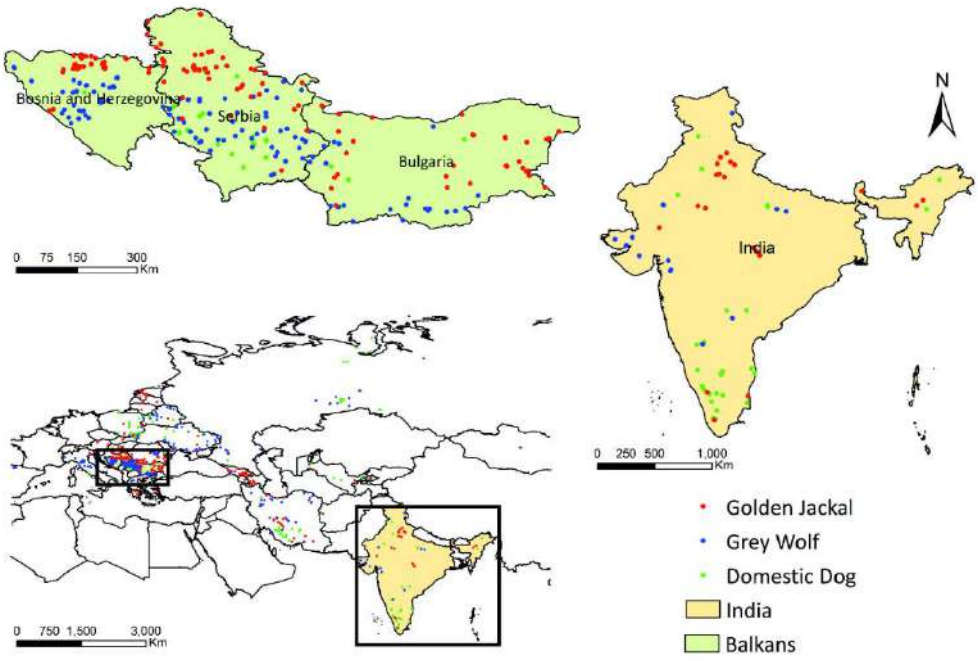


Fig. 2.2. Distribution of regional datasets, including India and the Balkans

Global Ancestry Analysis

For inferring global ancestry, both model-based and non-model-based approaches were used. Since one of the important assumptions of the global ancestry analysis is the absence of correlations between allele frequencies in different loci, datasets pruned according to LD were used for performing PCA and ADMIXTURE. To obtain three-way ancestry estimates, WJD dataset with 36K variants was used, while datasets WD, JD, and WJ were used for pairwise comparisons among gray wolves, golden jackals, and domestic dogs.

PCA

To compute eigenvalues and eigenvectors and visualize the predominant components of variability in all datasets, Plink software v. 1.9 (Chang et al., 2015) was used. PCA was performed independently for each dataset (WJD, WD, JD, and WJ), to obtain both

three-way and two-way admixture estimates. The values of the first and second principal components were used to visualize the results of the PCA and the position of putative hybrids in R v. 3.4.3. Samples located between the clusters in the first two PCs were identified as admixed samples based on the visual inspection of the plots.

ADMIXTURE

We applied the Bayesian clustering method implemented in ADMIXTURE software v. 1.3.0 (Alexander et al., 2009) for quantifying the proportions of different possible ancestral origins within the genome of each individual and determining potential signatures of dog ancestry in the wild canid populations and vice versa. Utilizing maximum likelihood methods like ADMIXTURE can significantly reduce the computational time required for estimating genetic ancestry compared to STRUCTURE (Goulet et al., 2017). Furthermore, compared to principal component analysis (PCA), ancestry coefficients (q-values) offer greater interpretability as they represent proportions of ancestry from multiple founding populations (Ko et al., 2023). We performed the ADMIXTURE analysis for K values ranging from 2 to 10, employing default termination criteria. For each K value, the analysis was repeated 10 times with different random seeds, and the cross-validation error was computed. The same as in the PCA analysis, the ADMIXTURE analysis was performed separately for each dataset using the LD-pruned data. The lowest cross-validation error was used to find the optimal ancestry clusters. Ancestry coefficients were then utilized to differentiate between purebred and admixed individuals. Bar plots illustrating the estimated ancestry proportions derived from q-values were generated using R v. 3.4.

Local ancestry inference (LAI)

In addition to global admixture, our study inferred the presence of chromosomal ancestry blocks resulting from hybridization through local ancestry analysis for all 38 autosomal chromosomes. We applied three distinct approaches (LAMP-LD, Ghap, and ELAI) with different algorithms to estimate local ancestry across all pairwise comparisons. All local ancestry analyses were performed using non-LD-filtered datasets.

LAMP-LD

Local Ancestry in Mixed Populations (LAMP) is based on the clustering algorithm (ICM) for the estimation of LA in recently admixed populations (Sankararaman et al., 2008). LAMP-LD (Baran et al., 2012) is an advancement of the LAMP algorithm designed to address multi-way admixtures, by integrating Hidden Markov Models (HMMs) with a window-based framework, resulting in highly accurate estimations (Padhukasahasram, 2014). LAMP-LD was carried out with two objectives: first, to estimate introgression rates between all pairwise comparisons (between domestic dogs, gray wolves, and golden jackals) and second, to identify non-admixed individuals. LAMP-LD estimates the admixture proportions without the need to define *a priori* ancestral non-admixed populations. Due to the presence of signatures of past admixture with dogs in most Eurasian wolf populations (Pilot et al., 2018), we used the results of

LAMP-LD analyses to identify non-admixed individuals to be included in reference panels in the ELAI analyses. In the LAMP-LD analysis, the mixture proportion (*alpha*) between two main clusters was set up so as to correspond to the frequency of the two taxa compared in each dataset (see Table 2.1). This was a conservative assumption that allowed for the scenario of no admixture. For all pairwise comparisons, the r^2 cutoff of 0.1 (*ldcutoff = 0.1*) was selected for pruning SNPs in linkage disequilibrium, and the fraction of overlap between adjacent windows (*offset*) was 0.2.

The recombination rate of 1×10^{-10} per base pair per generation was set (as in Pilot et al. 2021). We attempted to use a high-density recombination map from available sources (Campbell et al., 2016), but with its inclusion in the model, we were not able to obtain reliable admixture results. This was not caused by any problems with the recombination map, as we used it successfully in the ELAI analysis (see below). Therefore, we used a fixed recombination rate instead, an approach that was successfully applied in earlier studies (Pilot et al., 2018; Pilot et al., 2021; Sarabia et al., 2025). We considered the last 10 generations since admixture because the precision of the admixture estimates in LAMP-LD decreases with the increase in the number of generations since the admixture event (Suarez-Pajes et al., 2021).

Table 2.1. Mixture proportions were applied in the LAMP_LD analysis for the global, Indian, and Balkan datasets.

	(wolves *vs* dogs)	(jackals *vs* dogs)	(wolves *vs* jackals)
Global datasets	0.35:0.65	0.45:0.55	0.40: 0.60
India datasets	0.31:0.69	0.45:0.55	0.35:0.65
Balkan datasets	0.32:0.68	0.77:0.23	0.40:0.60

ELAI

In the next step, the inference of local ancestry was performed with ELAI v. 1.01 (Guan, 2014). ELAI uses a two-layer hidden Markov model to detect haplotype structure between unrelated individuals and assigns each local haplotype probabilistically to groups (Guan, 2014). Although LAMP-LD was used without defining any source populations, ELAI required the provision of reference populations. We defined reference populations for wild canid species and domestic dogs as sets of individuals that were not admixed or had very low levels of inferred admixture based on the LAMP-LD results. Different thresholds were set for wild canids and dogs for the classification of individuals as non-admixed to obtain at least 150 non-admixed individuals for each dataset representing all geographic regions studied (Table 2.2). For pairwise comparisons, ELAI was run with 30 steps in the expectation–maximization (EM) run between two main clusters (wolves vs dogs), (wolves vs jackals), and (jackals vs dogs) assuming 10 generations of admixture. The values for upper-layer clusters (the number of hybridizing taxa) and lower-layer clusters (representing population structure within each taxon) were set to 2 and 10, respectively. Moreover, three-way ELAI was applied to all admixed samples, including 568 admixed dogs, wolves, and golden jackals. The same parameters

described above were used, with a slight modification in the EM (20 steps). For the clustering structure, the upper-layer clusters (representing hybridizing taxa) were set to 3, while the lower-layer clusters (representing population structure within each taxon) were set to 15.

Table 2.2.Thresholds applied for the selection of pure and admixed individuals for the ELAI analysis in each dataset

Datasets		Threshold for pure wolves	Threshold for pure jackals	Threshold for pure dogs	Admixed samples (N)
Global datasets	Dataset WDJ	0.01	0.003	0.003	568
	Dataset WD	0.01	-	0.003	314
	Dataset JD	-	0.005	0.001	315
	Dataset JW	0.001	0.003	-	257
Indian datasets	Dataset IWD	0.005	-	0.001	26
	Dataset IJD	-	0.001	0.001	20
	Dataset IWJ	0*	0.01	-	10
Balkan datasets	Dataset BWD	0.005	-	0.002	91
	Dataset BJD	-	0.002	0.001	100
	Dataset BWJ	0*	0.005	-	59

*as no wolf samples in these populations had golden jackal admixture, ELAI was run only based on admixed jackal samples.

GHap

In addition to LAMP-LD and ELAI, which are both based on the Hidden Markov Model (HMM), we also applied another method with a different underlying approach. GHap can predict ancestries by constructing haplotype blocks, without the need to define non-admixed reference populations (Utsunomiya et al., 2020). We used an algorithm implemented as part of the GHap (Genome-wide Haplotyping) package v. 3.0.0 (Utsunomiya et al., 2016) in R v. 3.4.3, to identify the tracks of local ancestry based on the observed haplotypes (Utsunomiya et al., 2020). In contrast to LAMP-LD and ELAI, which analyze unphased genotypes, GHap requires phased and imputed datasets. Therefore, before the GHap analysis, genotype data for each chromosome were phased using Beagle v. 5.4 (Browning et al., 2021). To increase the accuracy of phasing, the values for the burn-in and iteration parameters were set up at 10 and 1000, respectively, and the genetic map file for the recombination rate was prepared based on the CanFam3 dog genome for each chromosome. After conducting phasing for the wolf, golden jackal, and dog genotypes separately, the phased data were merged to form phased pairwise datasets. The phased genotypes of each dataset (WD, JD, and WJ) were used as input

files for GHap. We used GHap to estimate haplotype ancestry without reference samples. The K-means algorithm (Hartigan & Wong, 1979) was employed to simultaneously infer haplotype structure and ancestry in an unsupervised manner (Utsunomiya et al., 2020). Subsequently, a random sample of seeding markers was used to group all haplotypes in groups with K from 2 to 10. After choosing the best K, the ancestry proportions of each K were estimated using 'ghap.anctest' function (Utsunomiya et al., 2020). For visualization of the estimated ancestry proportions for each dataset, 'ghap.karyoplot' function was used.

Comparisons of global ancestry and local ancestry inference methods

The results obtained from global and local ancestry analyses were compared using the squared pairwise Pearson correlation test (Pearson, 1895). The scatter plots were plotted using MedCalc software (Schoonjans, 2020). The degree of correlation and significance were used as indices of consistency between the results of different methods.

Comparison between the results from the regional datasets and the entire dataset

To find the effect of sample size and subpopulation structures on the results of global and local ancestry, we compared the results from both regional datasets with the entire dataset. The squared pairwise Pearson correlation test was used to estimate the degree of consistency. Moreover, Bland-Altman plots (Giavarina, 2015) were employed to detect discrepancies in ancestry proportion estimates between the datasets. Samples showing substantial differences in ancestry proportions between the regional and entire datasets were identified as outliers. Specifically, any samples falling outside the standard deviation lines from the mean difference on the Bland-Altman plots were flagged as outlier samples. This approach enabled us to assess the influence of dataset composition and structure on ancestry estimates.

2.3. Results

WJD dataset

Global ancestry analysis

The results of PCA identified wolves, dogs, and golden jackals as three distinct clusters (Fig 2.3). The first and second axes of the PCA explained approximately 31.4% and 17.7% of the total variance, respectively, and separated wolf, dog, and golden jackal as three different clusters. Nine putative hybrids, positioned between the species clusters, were identified (Table 2.3). Similar to the PCA results, the results of ADMIXTURE showed that at K=3, wolves, dogs, and golden jackals were identified as three distinct groups (Fig 2.3). However, compared to the PCA, more first-generation (F1) hybrids (25 putative hybrids) were identified based on the ADMIXTURE results. Six canids sampled as wolves were assigned to the dog cluster with 55-47% probability and were identified as F1 wolf-dog hybrids (Table 2.3). 18 canids sampled as golden jackals were identified as first-generation hybrids with either dogs or wolves (Table 2.3). Given the high genetic similarity and close evolutionary relationship between wolves and domestic dogs,

distinguishing between wolf and dog contributions to ancestry in admixed jackal samples is challenging (see Discussion).

Local ancestry analysis

The three-way ELAI analysis identified five possible F1 hybrids between wolves and dogs. However, only three of these individuals exhibited 50% dog ancestry across most of their chromosomes – a pattern expected from true hybrids (Table 2.3). In the case of admixed golden jackals, three first-generation hybrids with dogs were identified, all showing 50% dog ancestry across all chromosomes. No wolf-golden jackal hybrids were detected using the three-way ELAI analysis. A comparison of global and local ancestry results using the WJD dataset revealed inconsistencies, with more F1 hybrids detected based on global ancestry. These discrepancies were particularly evident in samples of lower quality (LQ) (Table 2.3).

Table 2.3. Putative F1 hybrids identified based on the results of different methods using the WJD dataset.

Sample_ID	PCA	ADMIXTURE	ELAI (three-way)
WSER466	✓	✓	✓
WSER483	✓	✓	✓
WBOS18	✓	✓	✓
WIRA616	✓	✓	✓
WIndD2449	✓	✓	✓
WIRA631_LQ	✓	✓	✗
WIRA1042_LQ	✓	✓	✗
JROM10658_LQ	✓	✓	✓
JHUN9531	✓	✓	✓
JBEL598_LQ	✓	✓	✓
JIndD471_LQ	✗	✓	✗
JIndD473_LQ	✗	✓	✗
JIndD474_LQ	✗	✓	✗
JIndD2629_LQ	✗	✓	✗
JIndDF4398_LQ	✗	✓	✗
JIndD2743_LQ	✗	✓	✗
JIndD2639_LQ	✗	✓	✗
JBUL241-19_LQ	✗	✓	✗
JWBOS38_LQ	✗	✓	✗
JKAU8321_LQ	✗	✓	✗
JWKAU5740_LQ	✗	✓	✗
JWKAU5741_LQ	✗	✓	✗
JGEO42_LQ	✗	✓	✗
JGRE9066_LQ	✗	✓	✗
JURK8926_LQ	✗	✓	✗

✓ Samples show 50% of ancestry from each of the two canids

✓ Samples show 50% of ancestry from each of the two canids in each chromosome.

✓ Samples show an average of 50% ancestry from each of the two canids but not within each chromosome

✗ Samples not identified as F1 hybrids

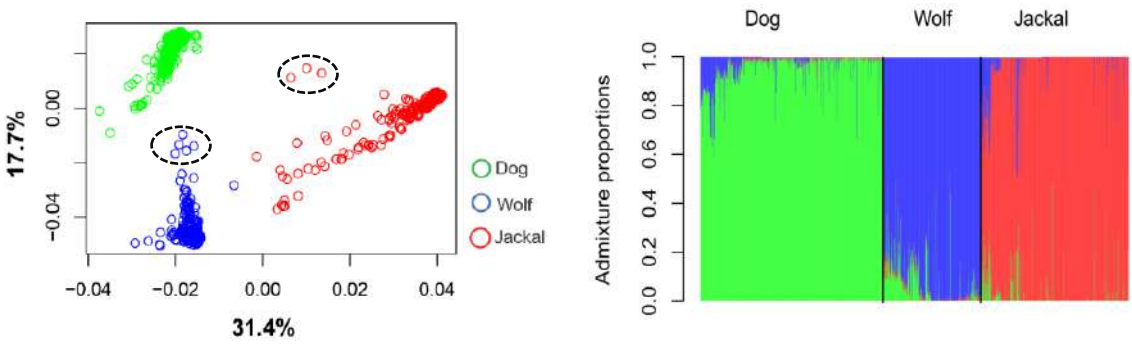


Fig. 2.3. PCA plot of first two principal components for the WJD and Admixture plots for K =3 in wolves, dogs, and golden jackals using the WJD dataset. In the PCA plot, the putative hybrids are marked with a black dashed circle.

WD dataset

Global ancestry analysis

The first and second axes of the PCA separate wolves from dogs and explain 30.47% and 11.03% of the total variance, respectively (Fig 2.4). Six putative hybrids, positioned between the species clusters, were identified as F1 hybrids based on the PCA result (Table 2.4). Based on the results of ADMIXTURE, seven canid samples were identified as F1 hybrids (Table 2.4, Fig 2.4). The average proportion of estimated dog ancestry in admixed wolves is reported in Table 2.5.

Local ancestry analysis

The proportion of chromosomal blocks of dog ancestry in wolf samples estimated in LAMP-LD varied between 0% and 55%. Six wolves displayed an average of 40%-55% dog ancestry and were identified as F1 hybrids. These six samples were also identified as F1 hybrid based on the ADMIXTURE results for the WJD and WD datasets. However, only four of these samples showed 50% dog ancestry in all their chromosomes and the last two samples (WBOS18 and WSER483) showed different proportions of dog ancestry in each chromosome (Table 2.4). Although all Indian wolf samples had varying proportions of dog ancestry based on the global analysis, LAMP_LD identified them as pure wolves, with more than 95% wolf ancestry, except two canids that were identified as an F1 and F2 hybrids. We found strong consistency between the results of ELAI and LAMP-LD, since the same six wolf samples were identified as F1 hybrids based on both methods. However, only four of them showed around 50% dog ancestry in all chromosomes based on the ElAI and LAMP_LD (Table 2.4). Although ELAI and LAMP-LD are independent methods that differ in their approaches, they produced consistent chromosomal patterns for these samples. In contrast, the results from GHap identified a greater number of apparently pure individuals compared to both LAMP-LD and ELAI (Fig. 2.5). The estimated proportion of dog ancestry in wolves based on the different methods presented below (Table 2.5).

Table 2.4. Putative F1 hybrids identified based on the results of different methods using the WD dataset.

Sample_ID	PCA	ADMIXTURE	LAMP-LD	ELAI	GHap
WSER466	✓	✓	✓	✓	✓
WSER483	✓	✓	✓	✓	✓
WBOS18	✓	✓	✓	✓	✓
WIRA616	✓	✓	✓	✓	✓
WIndD2449	✓	✓	✓	✓	✗
WIRA631_LQ	✓	✓	✓	✓	✗
WIndD3078_LQ	✗	✓	✗	✗	✗

✓ Samples show 50% of ancestry from each of the two canids

✓ Samples show 50% of ancestry from each of the two canids in each chromosome.

✓ Samples Show an average of 50% ancestry from each of the two canids but not within each chromosome

✗ Samples not identified as F1 hybrids

Table 2.5. The estimated dog ancestry proportions in wolves and jackals based on the different ancestry inference methods

Method	Dog ancestry in wolves	Dog ancestry in jackals	Wolf ancestry in jackals
ADMIXTURE	0.064	0.048	0.054
LAMP-LD	0.044	0.031	0.044
ELAI	0.042	0.020	0.036
GHap	0.034	0.006	0.006

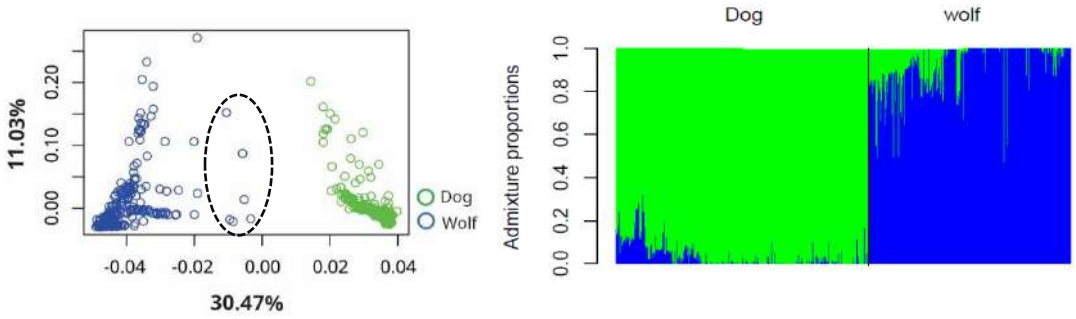


Fig. 2.4. PCA plot of two first principal components for the WD and Admixture plots for K =2 in wolves and dogs using the WD dataset. In the PCA plot, the putative hybrids are marked with black dashed line.

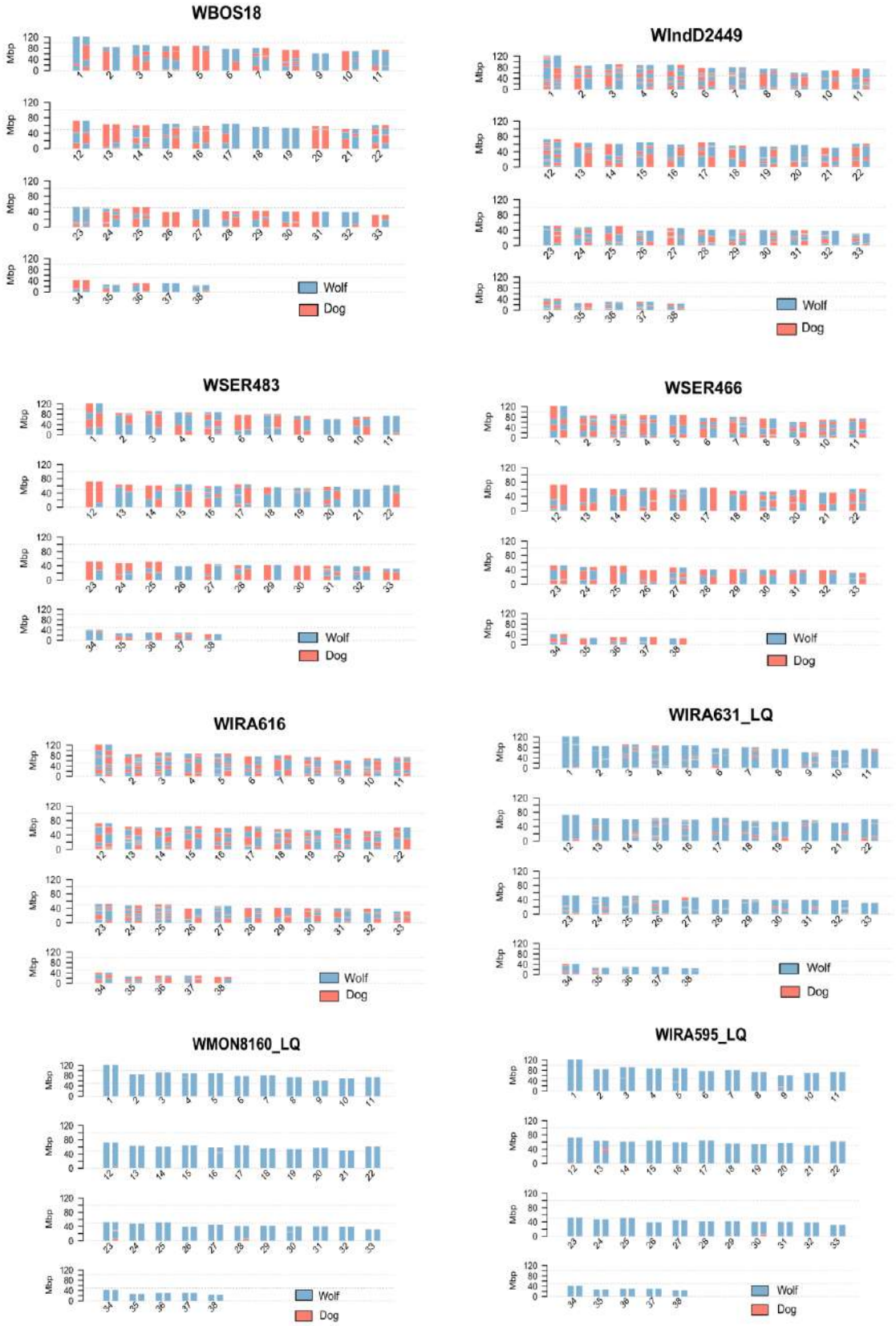


Fig. 2.5. Karyoplots of some wolf samples (including admixed and pure) in all 38 chromosomes based on the results of GHap using the WD dataset

JD dataset

Global ancestry analysis

The first and second axes of the PCA distinguished dog and golden jackal clusters (Fig 2.6) and explained approximately 40.6% and 14.9% of the total variance, respectively. Only three putative F1 hybrids, positioned between the species clusters, were identified based on the PCA (Table 2.6). Based on the results of ADMIXTURE, 17 more samples (20 samples altogether) were identified as F1 hybrids (Table 2.6, Fig 2.6). The proportion of estimated dog ancestry in jackals was 4.8% based on the ADMIXTURE results (Table 2.5).

Local ancestry analysis

The proportion of chromosomal blocks of dog ancestry in golden jackals varied between 0 and 50% based on the LAMP-LD. While dog samples showed limited signs of admixed ancestry (no more than 0.2%), 22 golden jackal samples showed 50% dog ancestry on average and were identified as putative F1 hybrids. However, only 12 samples showed 50% dog ancestry in all their chromosomes (Table 2.6). These results were mostly consistent with the results of ADMIXTURE, however, the estimated proportion of dog ancestries based on the LAMP_LD was smaller than ADMIXTURE in most cases. Based on the ELAI results, only four samples displayed 50% dog ancestry in all chromosomes and therefore were identified as F1 hybrids (Table 2.6). For other samples that were based on the LAMP-LD and identified as F1 hybrids, ELAI recognized them as F2 or even backcrosses. Therefore, compared to the results of the WD dataset, less consistency was found between the results of ELAI and LAMP-LD. GHap identified only three samples as F1 hybrids (Fig 2.7). In this case, the result of GHap was more consistent with the result of ELAI compared to other methods (Table 2.6). Among 24 golden jackal samples that were identified as F1 hybrids based on different methods, only one sample was not an LQ sample. The estimated proportion of dog ancestry in golden jackals based on the different methods presented below (Table 2.5).

Table 2.6. Putative F1 hybrids identified based on the results of different methods using the JD dataset.

Sample_ID	PCA	ADMIXTURE	LAMP-LD	ELAI	GHap
JROM10658_LQ	○	○	●	●	○
JHUN9531	○	○	●	●	○
JBEL598_LQ	○	○	●	●	○
JIndD471_LQ	✕	○	✕	✕	✕
JIndDF4398_LQ	✕	○	✕	✕	✕
JIndD2743_LQ	✕	○	✕	✕	✕
JIndD3274_LQ	✕	○	●	✕	✕
JIndD2175_LQ	✕	○	●	✕	✕
JIndD3018_LQ	✕	○	●	✕	✕
JIndD45_LQ	✕	○	◆	✕	✕
JIndDF417_LQ	✕	○	●	✕	✕
JIndD480_LQ	✕	○	●	✕	✕
JBUL241-19_LQ	✕	○	●	✕	✕
JBUL410-19_LQ	✕	✕	●	✕	✕
JBUL78-19_LQ	✕	✕	●	●	✕
JWBOS38_LQ	✕	○	●	✕	✕
JKAU8321_LQ	✕	○	◆	✕	✕
JWKAU5740_LQ	✕	○	◆	✕	✕
JWKAU5741_LQ	✕	✕	◆	✕	✕
JKAU8086_LQ	✕	○	◆	✕	✕
JKAU8341_LQ	✕	○	◆	✕	✕
JUKR8926_LQ	✕	○	◆	✕	✕
JGRE9066_LQ	✕	✕	◆	✕	✕
JUKR8600_LQ	✕	○	●	✕	✕

○ Samples show 50% of ancestry from each of the two canids

● Samples show 50% of ancestry from each of the two canids in each chromosome.

◆ Samples Show an average of 50% ancestry from each of the two canids but not within each chromosome

✕ Samples not identified as F1 hybrids

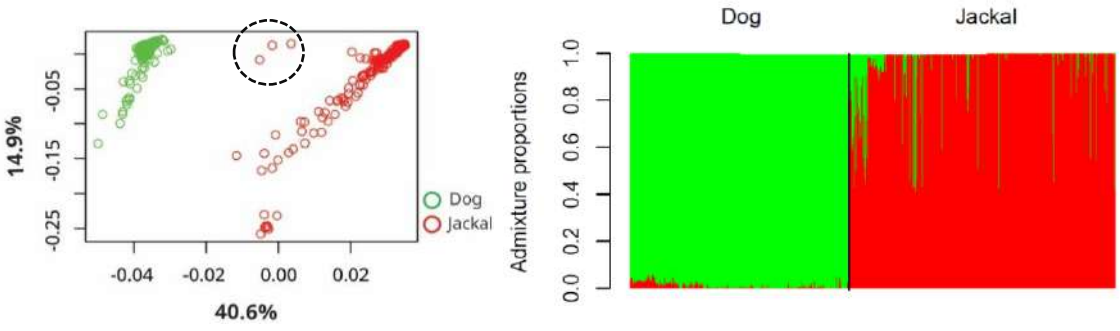


Fig. 2.6. PCA plot of two first principal components for the JD and Admixture plots for K =2 in golden jackals and dogs using the JD dataset. The putative hybrids are marked with black dashed circle.

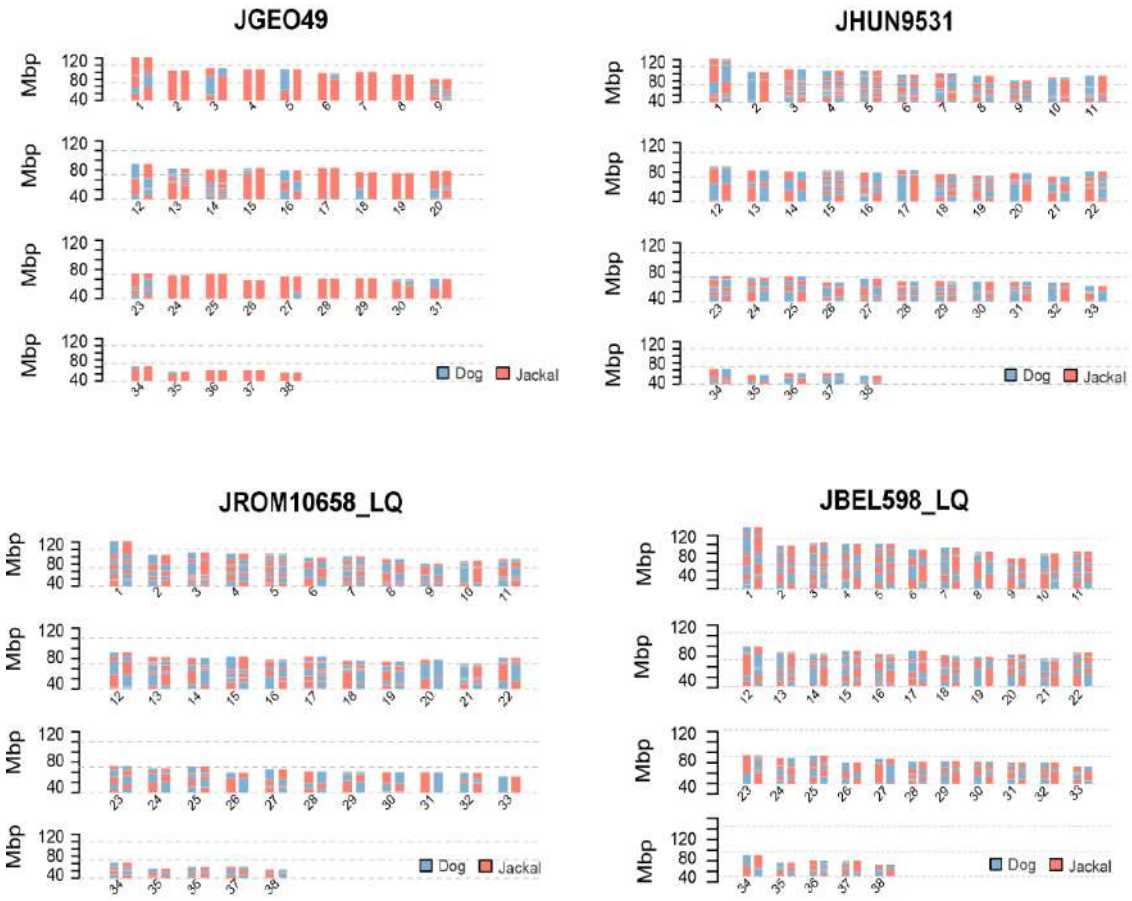


Fig. 2.7. Karyoplots of some admixed golden jackals (Dataset JD) in all 38 chromosomes based on the results of GHap.

WJ dataset

Global ancestry analysis

The results of PCA identified golden jackals and wolves as two separate clusters (Fig 2.8). The first and second axes of the PCA explained approximately 34.2% and 13% of the total variance, respectively. The identified F1 hybrids based on the results of PCA for all datasets are provided in Table 2.7. The result of ADMIXTURE analysis at K=2 including golden jackals and wolves showed that 65 golden jackal samples had lower than 90% of golden jackal ancestry proportions including all Indian golden jackals (Fig 2.8). Of these samples, 15 samples showed between 41% and 58 % golden jackal ancestry (Table 2.7). We also found that 20 wolves from Iran, India, and the Balkans had golden jackal admixture proportions between 10-23%, which corresponds to F2 or F3 backcrosses.

Local ancestry analysis

In the wolf-golden jackal admixture analysis, all wolves carried less than 0.1 of golden jackal ancestry in their chromosomes (except for two samples from Iran that were identified as F1 and backcrosses).

The proportion of chromosomal blocks of wolf ancestry in golden jackal samples was estimated at between 0 and 50%. The average proportions of wolf ancestry in 21 golden jackals were estimated between 43 to 50%, however, only 17 golden jackals displayed around 50% of chromosomal blocks originating from recent admixture with wolves in most of their chromosomes (Table 2.7). Most of these samples were recognized as backcrosses based on the results of ADMIXTURE, therefore inconsistency was found between the results of global ancestry and LAMP-LD. Based on the results of ELAI, 37 jackal samples carried less than 90% golden jackal ancestry, including nine samples where the average proportion of wolf ancestry was estimated to be between 40% and 50%. However, of these samples, only four individuals displayed 50% wolf ancestry in all chromosomes (Table 2.7). These results were consistent with the results of LAMP-LD. The mean percentage of different ancestries has been reported in Table S 2.9. In the wolf-golden jackal admixture analysis in GHap, all wolf samples had more than 98% wolf ancestry and two golden jackal samples were identified as F1 hybrids (2.7, Fig 2.9). The estimated proportion of wolf ancestry in golden jackals based on the different methods presented below (Table 2.5).

Table 2.7. Putative F1 hybrids identified based on the results of different methods using the WJ dataset.

Sample_ID	PCA	ADMIXTURE	LAMP-LD	ELAI	GHap
JROM10658_LQ	○	✗	✓	✓	○
JHUN9531	○	✗	✓	✓	○
JBEL598_LQ	○	✗	✓	✓	✗
JIndD471_LQ	✗	✗	✓	✗	✗
JIndDF417_LQ	✗	✗	✓	✓	✗
JIndDF4398_LQ	✗	✗	✓	✗	✗
JIndD45_LQ	✗	✗	✓	✓	✗
JIndD473_LQ	✗	✗	✓	✗	✗
JIndD480_LQ	✗	✗	✓	✓	✗
JIndD2175_LQ	✗	✗	✓	✗	✗
JIndD2743_LQ	✗	✗	✓	✗	✗
JDIndD3018	✗	✗	✓	✓	✗
JIndD3274_LQ	✗	✗	✓	✗	✗
JIndDF4398rep_LQ	✗	✗	✓	✗	✗
JUKR8600_LQ	✗	✗	✓	✗	✗
JIndD2628_LQ	✗	○	✗	✗	✗
JIndD2202_LQ	✗	○	✗	✗	✗
JIndD2219_LQ	✗	○	✗	✗	✗
JIndD2637_LQ	✗	○	✗	✗	✗
JIndD2630_LQ	✗	○	✗	✗	✗
JIndD3014_LQ	✗	○	✗	✗	✗
JIndD476_LQ	✗	○	✗	✗	✗
JIndD467_LQ	✗	○	✗	✗	✗
JIndD479_LQ	✗	○	✗	✗	✗
JIndD554_LQ	✗	○	✗	✗	✗
JIndDF2250_LQ	✗	○	✗	✗	✗
JBUL241-19_LQ	✗	✗	✓	✗	✗
JBUL410-19_LQ	✗	○	✓	✗	✗
JBUL78-19_LQ	✗	✗	✓	✓	✗
JGEO42_LQ	✗	○	✓	✗	✗
JHUN9025_LQ	✗	○	✗	✗	✗
JROM3500_LQ	✗	○	✗	✗	✗
JUKR8600_LQ	✗	✗	✓	✓	✗
WIRA631-LQ	○	✗	✓	✗	✗

○ Samples show 50% of ancestry from each of the two canids

✓ Samples show 50% of ancestry from each of the two canids in each chromosome.

✓ Samples Show an average of 50% ancestry from each of the two canids but not within each chromosome

✗ Samples not identified as F1 hybrids

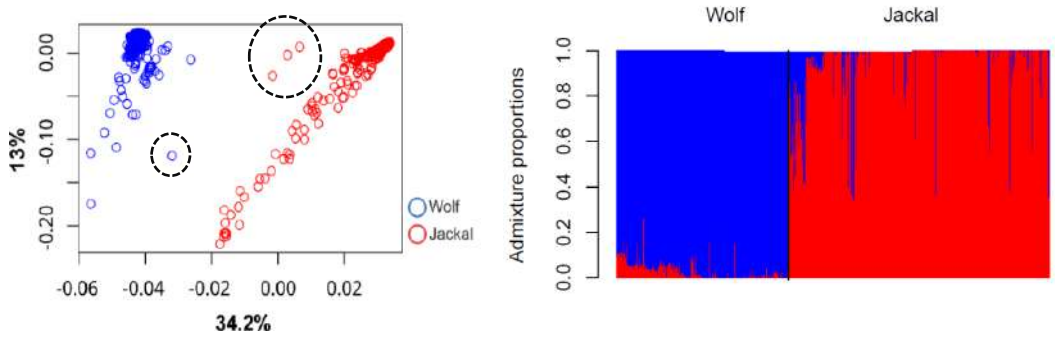


Fig. 2.8. PCA plot of two first principal components for the WJ and Admixture plots for K =2 in golden jackals and wolves using the WJ dataset. The putative hybrids are marked with black dashed circle.

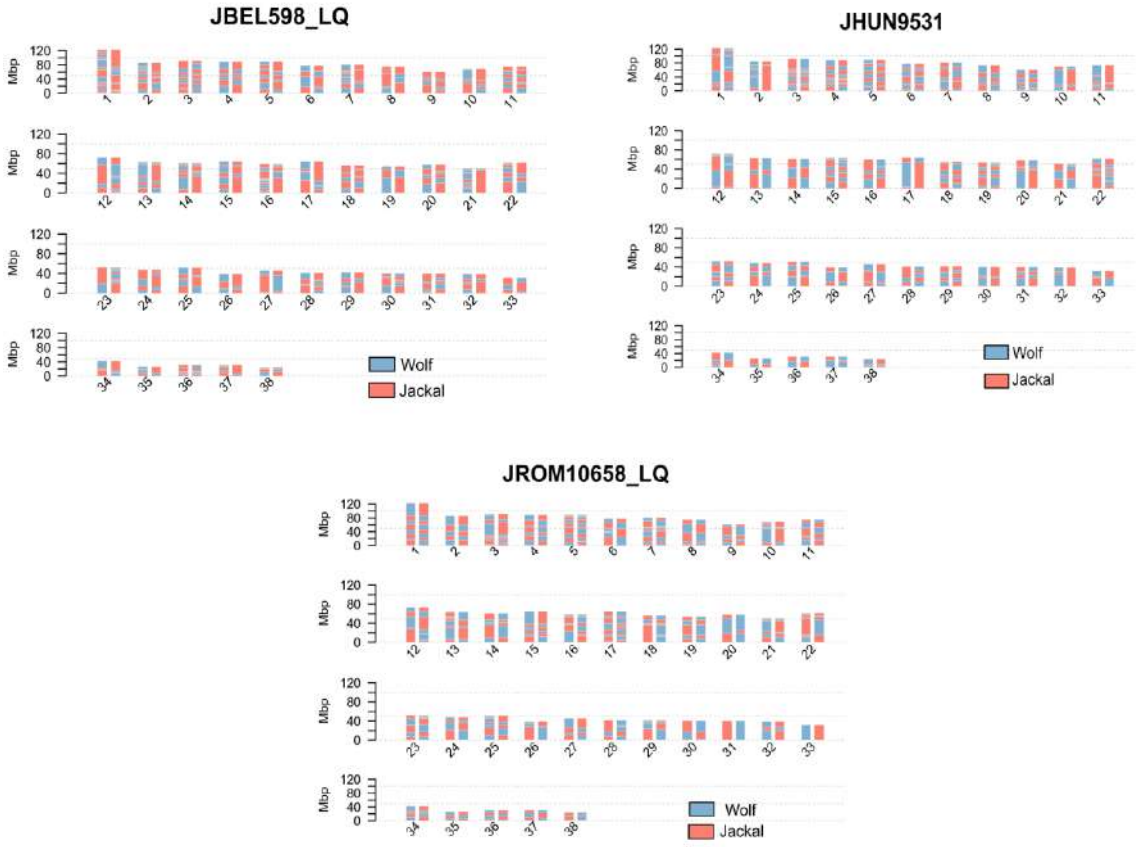


Fig. 2.9. Karyoplots of some admixed golden jackals (Dataset WJ) in all 38 chromosomes based on the results of GHap.

Regional dataset

Global ancestry analysis

For the India dataset (IWJD), the first and second axes of the PCA explained approximately 21.5% and 14.1% of the total variance, respectively (Fig 2.10). For the Balkan dataset (BWJD), the first and second axes of the PCA explained approximately 34.8% and 14.8% of the total variance (Fig 2.10). In both datasets the first and second components separate clusters of wolves, dogs, and golden jackals from each other. While seven F1 hybrids were identified using the BWJD dataset, only three individuals were classified as F1 hybrids based on the entire dataset (WJD) (Table S 2.3).

In the ADMIXTURE analysis for the Indian dataset (IWJD) at K=3, Indian wolves, dogs, and golden jackals were identified as three distinct groups (Fig 2.10). The single F1 hybrid identified using the India dataset was also identified as an F1 hybrid when analyzed with the entire dataset (Table 2.8). However, there were some inconsistencies between the results of the entire dataset and the India dataset. For example, none of the Indian samples identified as F1 jackal-dog hybrids based on the entire dataset were identified as hybrids when analyzed using the Indian datasets. Furthermore, all Himalayan wolves were identified as admixed individuals (Fig 2.10, Table 2.8) using the Indian dataset. In contrast, the results of ADMIXTURE for the Balkan dataset (BWJD) were mostly consistent with the results of the entire dataset (Table 2.8, Fig 2.10).

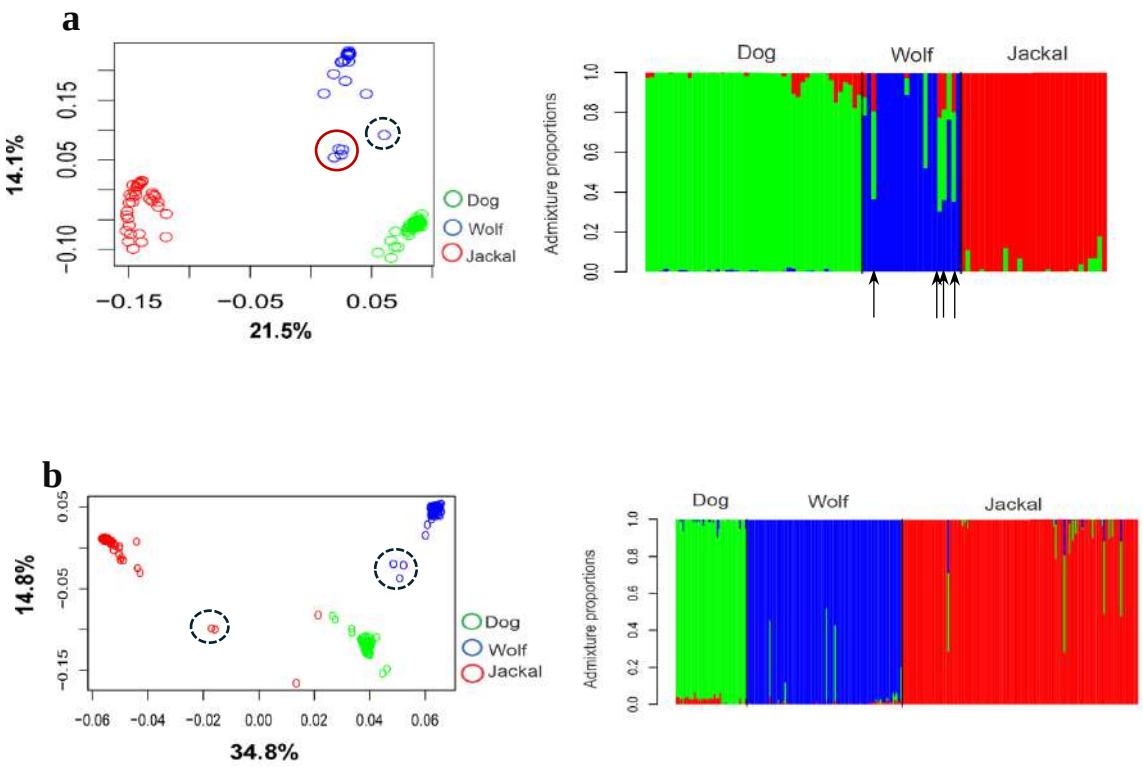


Fig. 2.10. PCA plot of two first principal components and Admixture plots for K =3 in wolves, golden jackals, and dogs using (a) the IWJD dataset and (b) the BWJD dataset. The putative hybrids are marked with black dashed circles. Himalayan wolves are marked with a red circle and black arrows in PCA and Admixture plots.

Table 2.8. Putative F1 hybrids identified based on ADMIXTURE results using the entire dataset and both regional datasets. Samples identified as F1 hybrids in both the regional and entire datasets are highlighted in bold.

Putative hybrids		**Entire datasets**			**India**			**Balkan**		
India	Balkan	WJD	WD	JD	IWJD	IWD	IJD	BWJD	BWD	BJD
WIndD2449		✓	✓		✓	✓				
WIndD3078_LQ			✓			✗				
WIndD3011		✗		✗	✓		✗			
WIndD3012		✗		✗	✓		✗			
WIndDLDK_LQ		✗		✗	✓		✗			
WIndD48		✗		✗	✓		✗			
JIndD471_LQ		✓		✓	✗		✗			
JIndD473_LQ		✓		✗	✗		✗			
JIndD474_LQ		✓		✗	✗		✗			
JIndD2629_LQ		✓		✗	✗		✗			
JIndDF4398_LQ		✓		✓	✗		✗			
JIndD2743_LQ		✓		✓	✗		✗			
JIndD2639_LQ		✓		✗	✗		✗			
JIndD3274_LQ		✗		✓	✗		✗			
JIndD2175_LQ		✗		✓	✗		✗			
JIndD3018_LQ		✗		✓	✗		✗			
JIndD45_LQ		✗		✓	✗		✗			
JIndDF417_LQ		✗		✓	✗		✗			
JIndD480_LQ		✗		✓	✗		✗			
	WSER466	✓	✓					✓	✓	
	WSER483	✓	✓					✓	✓	
	WBOS18	✓	✓					✓	✓	
	JBUL241-19_LQ	✓		✓				✓		✓
	JBUL410-19_LQ	✗		✓				✓		✓
	JBUL78-19_LQ	✗		✗				✓		✓
	JWBOS38_LQ	✓		✓				✓		✓

✓ Samples show 50% of ancestry from each of the two canids
✗ Samples not identified as F1 hybrids

Local ancestry analysis

We found consistent results between the results of LAMP-LD and ELAI using the Indian wolf-dog dataset (IWD) and the entire dataset (WD) (Table 2.9). In contrast, inconsistencies were found between the results of LAMP-LD and ELAI using the dog-jackal dataset from India (IJD) and the entire dataset (JD) (Table 2.9, Table S 2.8). The results of GHap in both Indian wolf-dog and jackal-dog datasets were consistent with the entire dataset (Table S 2.10, Fig S 2.8). The results of LAMP-LD and ELAI using the Balkan dataset (wolf-dog (BWD) and jackal-dog (BJD datasets) were completely consistent with these results from the entire dataset (WD, JD) (Table 2.9, Table S 2.9). Although consistent results were found between the results of Ghap using BWD and the

entire dataset (WD), inconsistent results were found between the Ghap results using BJD and the entire dataset (JD) (Table S 2.10, Fig S 2.9). The proportion of estimated dog ancestry in wolves and golden jackals based on different methods in both datasets (India and Balkan) is provided below (Table 2.10).

Table 2.9. Putative F1 hybrids identified based on LAMP-LD results using the entire and both regional datasets. Samples identified as F1 hybrids in both the regional and entire datasets are highlighted in bold.

Putative hybrids		**Entire datasets**		**India**		**Balkan**	
India	Balkan	WD	JD	IWD	IJD	BWD	BJD
WIndD2449		✓		✓			
JIndD471_LQ							
JIndD2743_LQ			✓		✗		
JIndD3274_LQ			✓		✗		
JIndD2175_LQ			✓		✗		
JIndD3018_LQ			✓		✗		
JIndD45_LQ			✓		✗		
JIndDF417_LQ			✓		✗		
JIndD480_LQ			✓		✗		
	WSER466	✓				✓	
	WSER483	✓				✓	
	WBOS18	✓				✓	
	JBUL241-19_LQ		✓				✓
	JBUL410-19_LQ		✓				✓
	JBUL78-19_LQ		✓				✓
	JWBOS38_LQ		✓				✓

✓ Samples show 50% of ancestry from each of the two canids in each chromosome.

✓ Samples Show an average of 50% ancestry from each of the two canids but not within each chromosome

✗ Samples not identified as F1 hybrids

Table 2.10. The estimated dog proportions in wolves and jackals in the Indian and Balkan datasets based on the different methods

Method	Dog ancestry in wolves (India)	Dog ancestry in jackals (India)	Dog ancestry in wolves (Balkan)	Dog ancestry in jackals (Balkan)
ADMIXTURE	0.150	0.023	0.019	0.021
LAMP-LD	0.082	0.000	0.023	0.016
ELAI	0.097	0.000	0.022	0.010
GHap	0.034	0.000	0.021	0.003

Comparisons of global and local ancestry inference methods

The estimated ancestry proportions for each species using global ancestry were compared to the local ancestry methods (Table 2.5). We found that the estimated global ancestry proportions by ADMIXTURE for each individual were higher than those estimated by local ancestry methods in all datasets. In most samples, the local ancestry estimated by GHap was lower than that estimated by other methods. For instance, while

many samples were identified as backcrosses or F3/F4 by ELAI and LAMP-LD, GHap classified these samples as pure individuals.

Comparing the estimated dog ancestry in wolf samples based on the local and global ancestry methods

The squared Pearson correlation was significant for each pairwise comparison between different methods of inferring ancestry proportions (Table 2.11, Fig 2.11). The dog ancestry proportions in wolves estimated by ADMIXTURE showed a higher correlation with LAMP-LD and ELAI results (Pearson correlation= 0.87 and 0.86, respectively), compared to GHap. In local ancestry analyses, a particularly strong correlation was detected between LAMP-LD and ELAI, with a Pearson correlation coefficient of 0.97. The differences in estimated admixture proportions between ELAI and LAMP-LD were generally minor, with most discrepancies being below 0.1. However, two samples, WIRA595_LQ and WMON8160_LQ, exhibited differences greater than 0.1. These samples were classified as F1 hybrids by LAMP-LD, but as backcrosses by ELAI, showing 22% and 15% dog ancestry, respectively, based on ELAI estimates. After excluding all low-quality samples from the WD dataset, the correlation coefficients increased consistently across all methods, exceeding 90% (Table 2.11).

Comparing the estimated dog ancestry in jackal samples based on the local and global ancestry methods

Based on the ADMIXTURE results, among 33 golden jackal samples from India, 30 samples were identified as F2/F3, which indicates the sensitivity of ADMIXTURE analysis to the global population structure within the golden jackal samples. ADMIXTURE estimated higher rates of dog ancestry in golden jackal samples compared to the local ancestry analyses. Moreover, in both pairwise comparisons, ADMIXTURE/LAMP-LD and ADMIXTURE/ELAI, samples with differences exceeding 0.1 in estimated admixture proportions were consistently low-quality samples. This variability in admixture estimates among different methods indicates that low-quality genotype data introduces inconsistencies in the estimation process. Although all pairwise comparisons among methods showed statistically significant correlations, weaker correlations were observed between GHap and the other analyses, particularly between ADMIXTURE and GHap (Table 2.11, Fig. 2.12). According to GHap results, most jackal samples had more than 97% jackal ancestry, with the exception of four samples identified as F1 hybrids or backcrosses. This likely explains the weaker correlation between GHap and the other methods. After excluding low-quality samples from the JD dataset, the correlations between GHap and the other methods improved substantially, reflecting the impact of data quality on consistency across analytical approaches (Table 2.11). Comparing the estimated wolf ancestry in dog and jackal samples are provided in the supplement (Fig S 2.10, Fig S 2.11).

Table 2.11. The squared pairwise Pearson correlation of dog ancestry in wolves and golden jackals was estimated by global and local ancestry.

	Dog ancestry in wolves		**Dog ancestry in wolves (No-LQ samples)**		**Dog ancestry in golden jackals**		**Dog ancestry in golden jackals (No-LQ samples)**	
Method	*r*	P-value	*r*	P-value	*r*	P-value	*r*	P-value
Admixture-LAMP-LD	0.87	<0.001	0.92	<0.001	0.91	<0.001	0.84	<0.001
Admixture-GHap	0.78	<0.001	0.90	<0.001	0.28	<0.001	0.81	<0.001
Admixture-ELAI	0.86	<0.001	0.97	<0.001	0.84	<0.001	0.91	<0.001
LAMP-LD-ELAI	0.97	<0.001	1	<0.001	0.93	<0.001	0.96	<0.001
LAMP-LD-GHap	0.87	<0.001	0.99	<0.001	0.37	<0.001	0.99	<0.001
GHap-ELAI	0.93	<0.001	0.99	<0.001	0.52	<0.001	0.95	<0.001

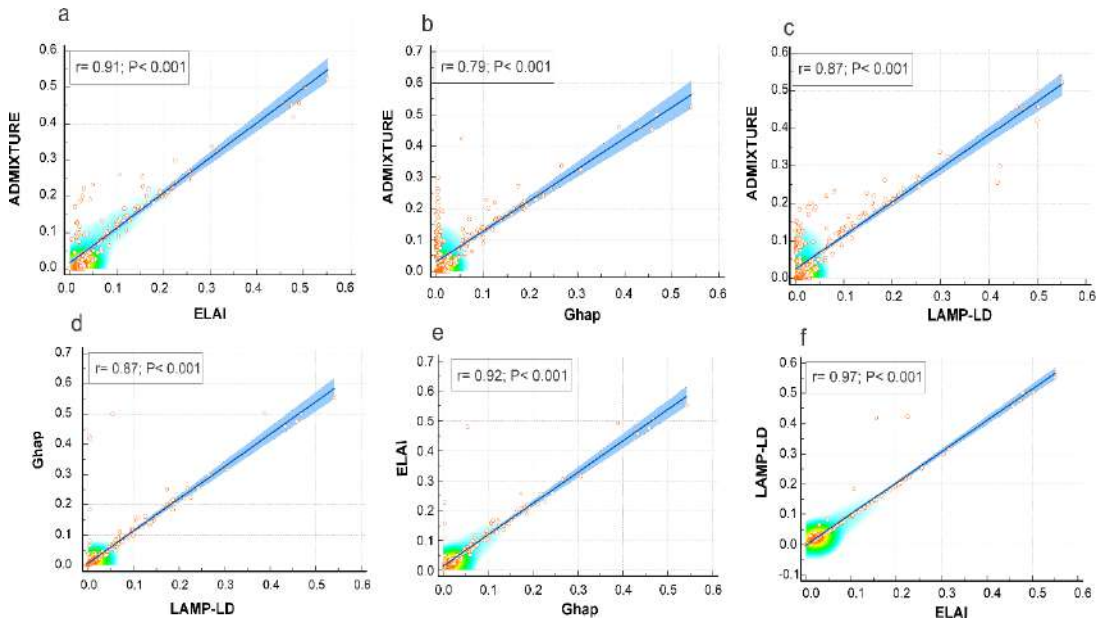


Fig. 2.11. The scatter plots between the estimated dog ancestry in wolves between ADMIXTURE and ELAI (a), ADMIXTURE and GHap (b), ADMIXTURE and LAMP-LD (c), LAMP-LD and GHap (d), ELAI and GHap (e), and LAMP-LD and ELAI (f).

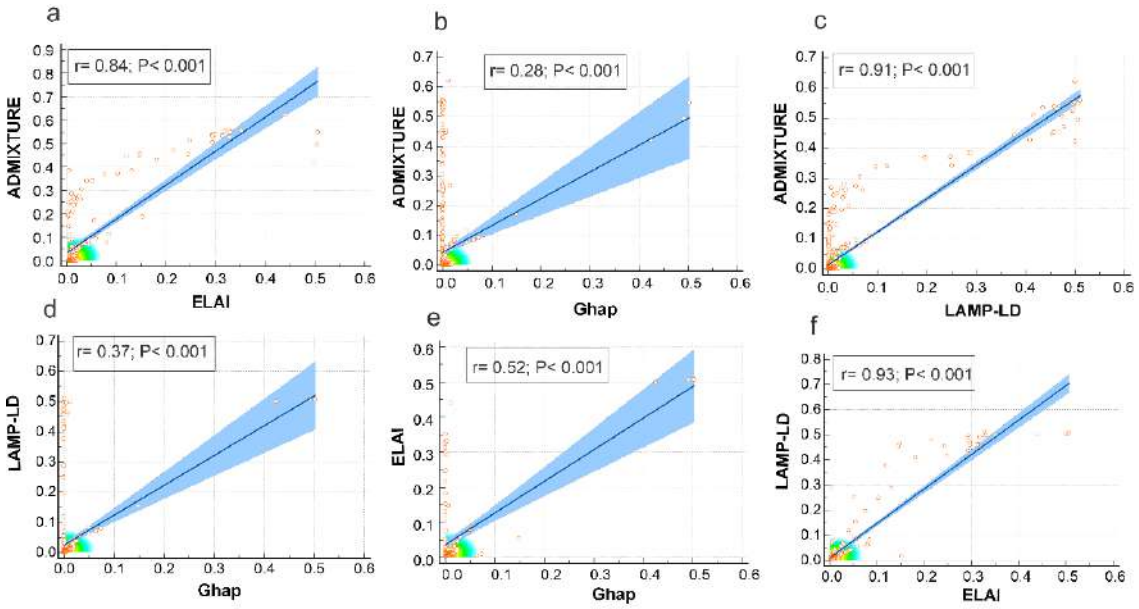


Fig. 2.12. The scatter plots between the estimated dog ancestry in jackals between ADMIXTURE and ELAI (a), ADMIXTURE and GHap (b), ADMIXTURE and LAMP-LD (c), LAMP-LD and GHap (d), ELAI and GHap (e), and LAMP-LD and ELAI (f).

Comparison of the results of the Indian dataset with the entire dataset

The average estimated dog ancestry in Indian wolves was higher than in other pairwise comparisons (Table 2.10). The estimated global ancestry proportions by ADMIXTURE for each individual were higher than those estimated by local ancestry methods in all datasets. We compared the results obtained from the Indian datasets with those from the entire dataset to investigate their consistency, using the squared Pearson correlation coefficient. For dog ancestry in Indian wolves, significant correlations were observed between the results from the entire dataset and the Indian dataset (Table 2.12). The Bland-Altman plot identified Himalayan wolves as outliers in the analyses of dog ancestry in wolves based on the ADMIXTURE result (Fig. 2.13, Table S 2.12, S 2.14). These samples exhibited differences of more than 0.38 in the estimated dog ancestry proportions when comparing results from the entire dataset to those derived from the India-specific dataset. The estimates of dog admixture in Indian jackals derived from the entire dataset were notably higher than those obtained from the Indian dataset alone. Significant correlations between the Indian and entire datasets were observed only for the Admixture and GHap analyses (Table 2.12). These findings suggest substantial differences in dog ancestry estimations when comparing the two datasets. Bland-Altman plots revealed outlier samples with notable discrepancies in dog ancestry among the jackals (Figure 2.13, Table S 2.14, S 2.18). Tables S 2.12 – S 2.18 and Figs S 2.12 – S 2.15 show the list of outlier samples.

Table 2.12. The average and squared pairwise Pearson correlation of dog ancestry in wolves and golden jackals using both the entire dataset and the Indian dataset.

	Dog ancestry in wolves				Dog ancestry in golden jackals			
Method	Average (Entire dataset)	Average (India)	*r*	P-value	Average (Entire dataset)	Average (India)	*r*	P-value
Admixture	0.150	0.159	0.45	0.039	0.287	0.023	0.79	<0.001
LAMP-LD	0.048	0.082	0.73	<0.001	0.149	0.000	0.07	0.7
GHap	0.031	0.034	1.00	<0.001	0.000	0.000	0.37	0.035
ELAI	0.041	0.097	0.86	0.01	0.078	0.003	0.44	0.45

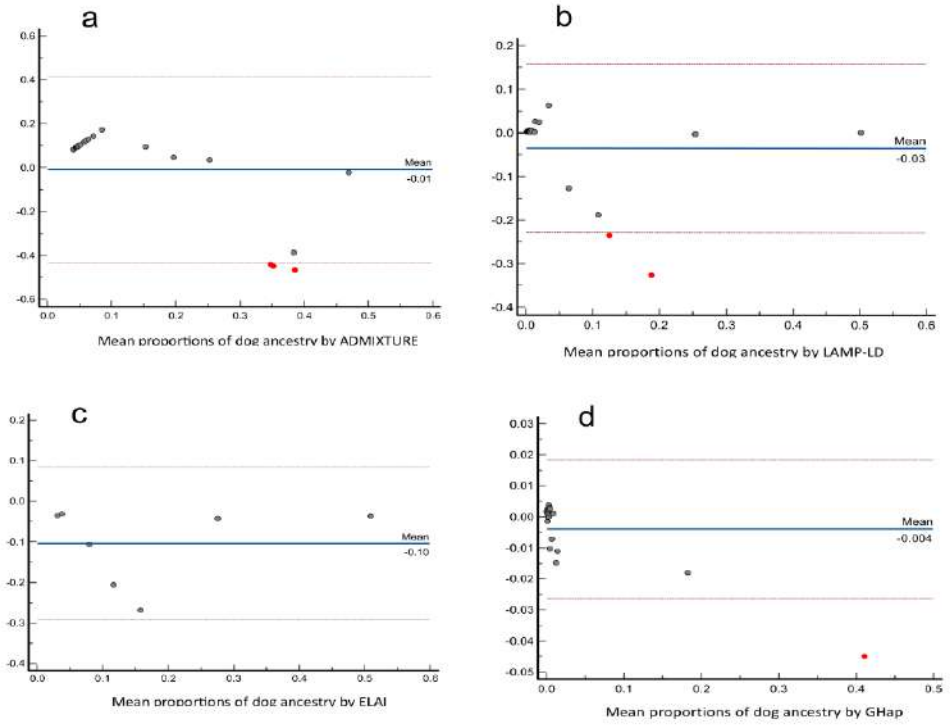


Fig. 2.13. Bland-Altman plots of average estimated dog ancestry in wolf samples from India by ADMIXTURE (a), LAMP-LD (b), ELAI (c), and GHap (d). The x-axis shows the average proportions of dog ancestry using the entire and Indian datasets and the y-axis shows the difference between estimated ancestry using the entire and Indian datasets. The outlier samples are displayed in yellow color.

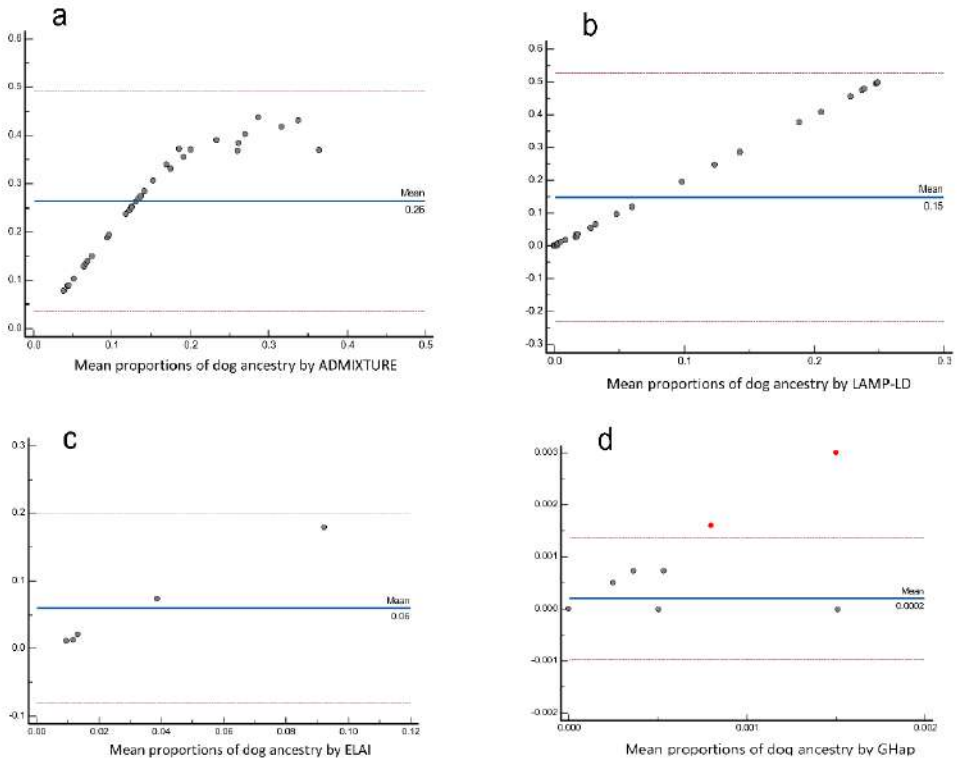


Fig. 2.14. Bland-Altman plots of average estimated dog ancestry in jackals samples from India by ADMIXTURE (a), LAMP-LD (b), ELAI (c), and GHap (d). The x-axis shows the average proportions of dog ancestry using the entire and Indian datasets and the y-axis shows the difference between estimated ancestry using the entire and Indian datasets. The outlier samples are displayed in yellow color.

Comparison of the results of the Balkan dataset with the entire dataset

Table 2.10 shows the average ancestral proportions in the Balkan wolf and golden jackal samples using global and local ancestry. The estimated global ancestry proportions by ADMIXTURE for each individual were higher than those estimated by local ancestry methods in all datasets.

High consistency was observed between the entire and Balkan datasets in estimating dog ancestry in Balkan wolves (Table 2.13). Bland-Altman plots revealed outlier samples with discrepancies in dog ancestry estimates among Balkan wolves (Fig. 2.15, Table S 2.19). However, these discrepancies were negligible, with differences lower than 0.03 between the two datasets. Significant correlations were found between global and local ancestry estimates for dog ancestry in jackals (Table 2.13), but outliers appeared in ADMIXTURE, LAMP-LD, ELAI, and GHap results (Fig 2.15, Table S 2.21). Most of these outlier samples were low-quality samples. Tables S 2.19 – S 2.25 and Figs S 2.16 – S 2.21 show the list of outlier samples.

Table 2.13. The average and squared pairwise Pearson correlation of dog ancestry in wolves and golden jackals using both the entire dataset and the Balkans dataset.

	Dog ancestry in wolves				Dog ancestry in golden jackals			
Method	Average (Entire dataset)	Average (Balkan)	*r*	P-value	Average (Entire dataset)	Average (Balkan)	*r*	P-value
Admixture	0.017	0.019	0.99	<0.001	0.017	0.021	1	<0.001
LAMP-LD	0.017	0.023	1	<0.001	0.015	0.016	1	<0.001
GHap	0.021	0.021	1	<0.001	0.003	0.003	0.87	<0.001
ELAI	0.021	0.022	1	<0.001	0.009	0.010	1	<0.001

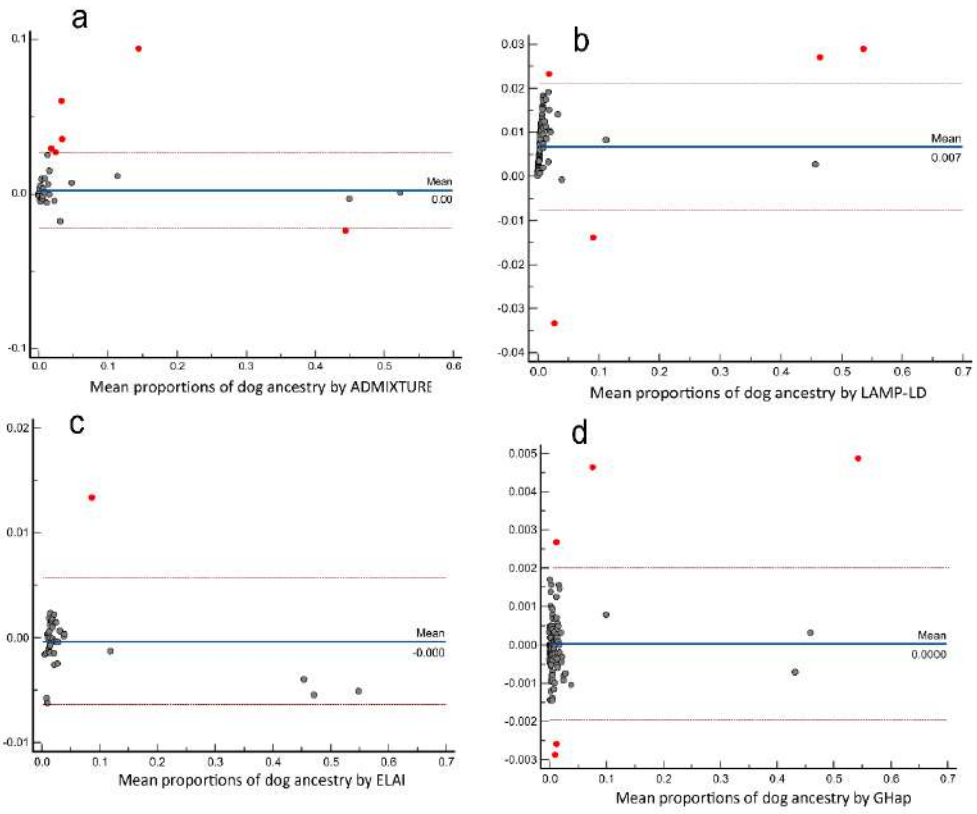


Fig. 2.15. Bland-Altman plots of average estimated dog ancestry in wolf samples from the Balkans by ADMIXTURE (a), LAMP-LD (b), ELAI (c), and GHap (d). The x-axis shows the average proportions of dog ancestry using the entire and Balkan dataset and the y-axis shows the difference between estimated ancestry using the entire and Balkan datasets. The outlier samples are displayed in yellow color.

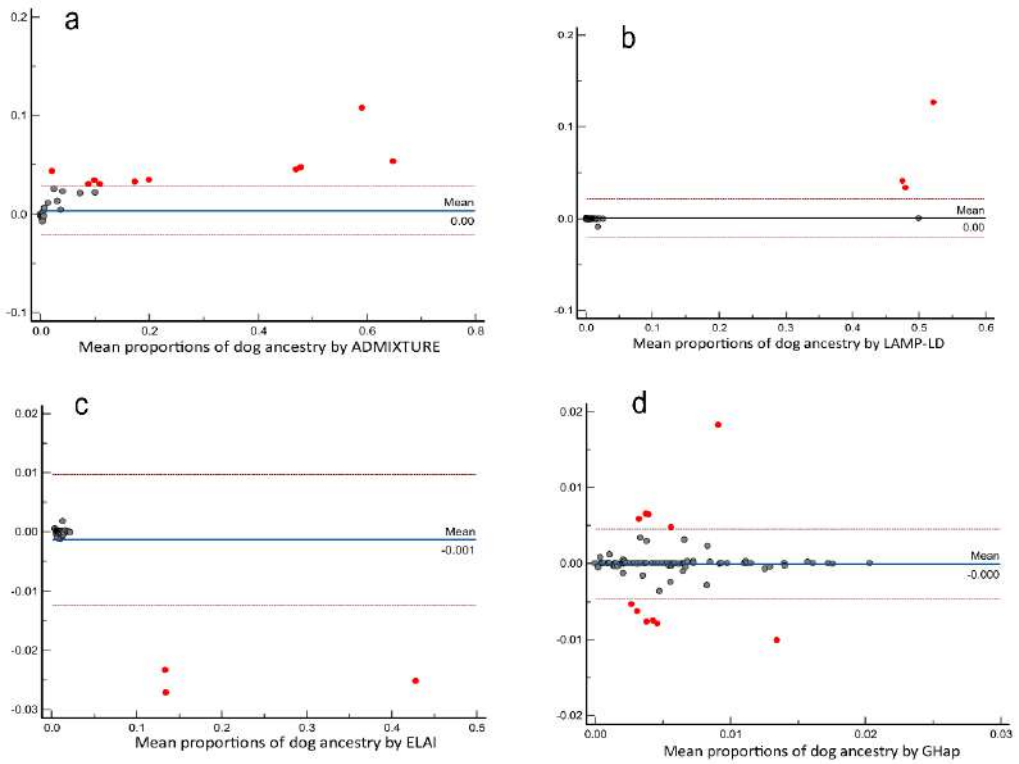


Fig. 2.16. Bland-Altman plots of average estimated dog ancestry in jackal samples from the Balkans by ADMIXTURE (a), LAMP-LD (b), ELAI (c), and GHap (d). The x-axis shows the average proportions of dog ancestry using the entire and Balkan datasets, and the y-axis shows the difference between estimated ancestry using the entire and Balkan datasets.

2.4. Discussion

Global ancestry analysis

Based on our results of PCA, the two first components explained more than 40% of the variance. However, some samples with low-quality SNP genotype data, especially jackal samples, could not be properly clustered into their groups using PCA. Although all datasets were filtered based on missing genotypes, minor allele frequency, and linkage disequilibrium, the presence of individuals with low-quality SNP genotypes may potentially confound the results of PCA. Himalayan wolves (*Canis lupus chanco*), a highly distinct wolf lineage (Werhahn et al., 2017) did not cluster with the other wolf samples in the PCA plots, which might have led to an incorrect conclusion of these individuals being admixed if the evolutionary history of the Himalayan population was unknown. This indicates that underlying intraspecific population structure can complicate the interpretation of principal components for the purpose of interspecific admixture detection. Although PCA is commonly used as an initial analysis in many population genetics studies, it has faced criticism (Chari et al., 2021; Björklund, 2019; Elhaik et al., 2014; McVean, 2009; Novembre et al., 2008). We recommend using PCA as one of the simplest and fastest initial methods for inferring admixture, which can

effectively identify outliers and potential hybrid samples. However, while it's a great starting point, PCA should not be used on its own to draw conclusions.

Unlike PCA, which captures the largest variation and projects samples into a PC space (Tan and Atkinson, 2023), model-based approaches such as ADMIXTURE (Alexander et al., 2009), estimate admixture proportions between 0 and 1, enabling the quantification of individual ancestries. Based on the results of ADMIXTURE at K=3, wolf, dog, and golden jackal were identified as three clusters, and putative hybrids were determined based on the ancestry coefficients. The results from the pairwise comparisons (datasets WD, JD, and WJ) were mostly consistent with those from dataset WDJ. However, we found inconsistencies in samples from India (Himalayan wolves), which demonstrates how genetic substructures may influence ADMIXTURE results. Although increasing the number of clusters (K) in ADMIXTURE can help to account for genetic substructures, it may also introduce several problems. Higher K values can lead to overfitting and the appearance of false admixture signals (Liu et al., 2022; Wang, 2022). Besides this challenge, ADMIXTURE overestimated admixture proportions in comparison with other methods. Therefore, we recommend using ADMIXTURE in conjunction with additional methods to ensure the robustness and accuracy of the findings.

Local ancestry analysis

The comparisons of the three methods of local ancestry analysis, showed high consistency between the results of LAMP-LD and ELAI. It may arise from their shared methodological principles; both of them are based on Hidden Markov Models (HMMs) and similar assumptions (e.g. using genotype information of source populations to predict the classification of ancestries of admixed individuals) (Baran et al., 2012). Moreover, since the results of LAMP-LD were used to define the source population in the ELAI analysis, we expected a high level of consistency between the outcomes of these methods. Significant differences were observed between the results of GHap and those of the two other methods, LAMP-LD and ELAI, with GHap producing the lowest proportions of introgressed ancestries. Utsunomiya et al. (2020) identified three factors that can significantly reduce the accuracy of local ancestry estimation by GHap: (1) an increase in the number of divergent lineages, (2) a decrease in the degree of divergence among lineages, and (3) the presence of shorter ancestry tracts. When source populations are insufficiently divergent and the number of generations since admixture increases, ancestral chromosome segments become shorter and harder to identify (Cottin et al., 2020). These shorter tracts contain fewer SNPs, reducing their informativeness and potentially introducing biases in ancestry inference (Avadhanam and Williams, 2023; Gravel, 2012). Therefore, the relatively low estimates of introgressed ancestry obtained using GHap in our datasets may be attributed to the varying degrees of evolutionary divergence and differing admixture histories among wolves, dogs, and golden jackals. Due to ongoing gene flow among these species, recombination gradually breaks down ancestral chromosomal blocks into shorter tracts. These shorter tracts contain fewer informative SNPs, making it more difficult for GHap to accurately infer local ancestry.

Although local ancestry analyses are powerful, they have some limitations. Most of these analyses require specific and complex information, such as the number of

generations since the last admixture event. This becomes particularly challenging in cases where admixture occurs continuously over time, as is likely in canids, rather than as a single, well-defined event. In such scenarios, accurately determining the number of generations since admixture is difficult, which complicates the application of local ancestry inference methods. Consequently, uncertainty or the lack of accurate population information can lead to biased estimates in local ancestry methods (Suarez-Pajes et al., 2021). Our findings highlight that the selection of appropriate methods for local ancestry analysis depends on the characteristics and quality of the dataset, as well as the specific objectives of the study. GHap showed lower consistency with other local ancestry inference methods, such as LAMP-LD and ELAI. In contrast, LAMP-LD and ELAI showed strong consistency and were effective for fine-scale ancestry analysis. However, the accuracy of LAMP-LD and ELAI can be affected by genotype quality (see below). These results also emphasize the need for careful consideration of methodological assumptions and input data quality when interpreting local ancestry estimates.

Comparisons of global ancestry and local ancestry inference methods

Although the estimated ancestry proportions by global and local ancestry typically should be consistent (Tan and Atkinson, 2023), some individuals with the same proportions of global ancestry can display differences in the distribution of local ancestry blocks due to the differences in population substructure (Secolin et al., 2019). The differences in accuracy between local and global ancestry results can be attributed to their methodological approaches and the type of information they utilize. For example, while the three-way ELAI model successfully distinguishes between wolf and dog ancestry in jackals, ADMIXTURE lacks the resolution to differentiate between these closely related sources. Local ancestry methods incorporate additional information such as linkage disequilibrium, haplotype structure, and relatedness which may make them more accurate than global ancestry methods. These features enable a more detailed analysis of ancestry at specific genomic regions. In contrast, ADMIXTURE, particularly its widely used unsupervised algorithm, relies on estimations that are not based on haplotypes, which limits its performance in identifying fine-scale population structure in admixed populations (Uren et al., 2020). In this study, ADMIXTURE estimated higher admixture proportions than local ancestry methods. The differences in the estimated admixture proportions between ADMIXTURE and local ancestry analysis methods can be attributed to several factors.

First of all, higher admixture proportions may be estimated due to ADMIXTURE's lack of limitations for the number of generations since hybridization was considered, while in local ancestry analysis a maximum of 10 generations since admixture was assumed, given that assuming a larger number of generations may lead to a significant overestimation of admixture in local ancestry methods (Pilot et al., 2018). Second, a significant limitation of ADMIXTURE is its assumption of independence for alleles present at neighboring SNPs, requiring the removal of SNPs in linkage disequilibrium (LD). This often leads to the exclusion of valuable information and usually retains many SNPs still in weak LD (Ko et al., 2023), which potentially leads to reduced estimation accuracy and increased false positives in ancestry inference (Padhukasahasram, 2014;

Geza et al., 2019; Uren et al., 2020). In this study, LD filtering removed 172,104 variants from all datasets, while all these SNPswere kept in local ancestry analyses. Third, since ADMIXTURE uses all markers across the genome to obtain global ancestry estimates, the obtained results may not reflect accurately the amount of ancestry variation in local chromosomal regions (Wang et al., 2011). Fourth, as the allele frequencies between subpopulations differ and ADMIXTURE assumes that individuals are sampled from the same homogeneous population, the presence of population stratification (subpopulations) may lead to false positive findings in the ADMIXTURE results (Skotte et al., 2013). For instance, the genetic substructuring within Indian samples (e.g. Himalayan wolves) caused most of these samples to be identified as F2/F3 or backcrosses based on the ADMIXTURE results. However, this problem can be fixed by increasing the number of clusters (K) and choosing the optimal K value. Moreover, various evolutionary processes could affect ancestry coefficients calculated in ADMIXTURE (Lawson et al. 2018; Anderson and Dunham 2008; Barilani et al. 2007). For example, the presence of close relatives violates the assumptions of the admixture model. This can result in biased outcomes by creating subclusters of related individuals, making it difficult to distinguish between ancestral groups and clusters of relatives (Thornton and Bermejo, 2014; Anderson & Dunham, 2008). Therefore, detecting related individuals is useful in identifying potential biases in admixture model results (Garcia-Erill and Albrechtsen, 2020).

Besides revealing higher proportions of introgressed variants through ADMIXTURE compared to local ancestry analysis, we identified some samples that, despite showing approximately 50% ancestry from each of the two canids at the global level, did not display this pattern consistently across individual chromosomes based on the local ancestry analyses. One possible explanation for this discrepancy is that these samples may result from F1 × F1 crosses, although such events are considered rare in nature. Another hypothesis for the inconsistency between local and global ancestry results is that these samples might indeed be true F1 hybrids, where local ancestry analysis failed to correctly assign ancestry for certain chromosomes. Methodological factors, such as low genotype quality, could significantly impact local ancestry inference. For instance, in the JD dataset, all 18 samples that did not show 50% ancestry across all chromosomes were low-quality (LQ) samples. This highlights the role of poor genotype quality and technical issues in producing inaccurate results. However, in the WD dataset, we identified two samples (WSER 483 and WBOS 18) that were not LQ samples but still failed to show 50% ancestry across all chromosomes, and showed the same chromosomal pattern based on the ELAI and LAMP-LD results. In this case, it is possible that these samples are F1 × F1 crosses, though further investigation is needed to confirm this hypothesis.

Although the Pearson correlation between the estimated dog ancestry by GHap and the other local and global ancestry methods was highly significant, a low correlation was found between them. This low correlation may be explained by the fact that GHap requires phased genotype input to infer individual ancestry proportions and local ancestry (Utsunomiya et al., 2020), while LAMP-LD and ELAI can operate with unphased genotype inputs. The quality of phased data depends on sample size (Avadhanam and Williams, 2023; Browning et al., 2021) and is also closely related to the density of genetic

markers (Browning and Browning, 2011). Therefore, a large datasets increase the accuracy of statistical phasing (Browning and Browning, 2022). In local ancestry analyses, the use of phased genotype datasets as input files can influence results due to phasing uncertainties in admixed samples (Thornton and Lorenzo Bermejo, 2014). For example, Chen et al. (2013) demonstrated that local ancestry estimation is more reliable when using unphased datasets, as phasing errors in admixed populations can compromise accuracy. Similarly, Avadhanam and Williams (2023) found that phase-free approaches outperform phase-based methods in local ancestry analysis, largely because the quality of phased data can introduce biases. These findings highlight the advantages of using tools like LAMP-LD and ELAI, which do not require pre-phased genotypes. Their ability to work with unphased input data makes them particularly valuable when high-quality phased genotypes are unavailable or uncertain. However, when high-quality, well-phased genotypes are available, GHap can use this information to provide precise estimates of ancestry proportions.

Factors confounding the results of local and global ancestry analysis

Low-quality genotype data

Our findings emphasize the impact of genotype quality on the consistency of global and local ancestry estimates across different methods. For example, using ADMIXTURE at K=3, we identified 25 first-generation hybrids, but only eight of these were classified as F1 hybrids by the three-way ELAI analysis. The majority of LQ samples were not recognized as F1 at the chromosomal level using ELAI. This suggests that inconsistencies between the methods are more likely to occur in low-quality (LQ) samples, as all high-quality F1 samples yielded consistent results in both local and global analyses. This pattern is further supported by results from the WD and JD datasets. In the WD dataset, 30 of 315 wolf samples were classified as LQ. Among these, global ancestry analysis identified 21 samples (70%) with less than 90% wolf ancestry, whereas LAMP-LD and ELAI classified only five samples (17%) below this threshold. However, in these LQ samples, GHap identified just one sample with approximately 89% wolf ancestry. In the JD dataset, LQ samples were more prevalent, comprising 102 out of 478 golden jackal samples. Of these, global ancestry analysis using Admixture identified 53 samples (52%) with less than 90% jackal ancestry. However, LAMP-LD and ELAI identified 28 and 25 samples, respectively, in this category, while GHap identified only two. GHap uses an imputed dataset, where all missing data are imputed based on the pure samples. This may explain why it detected considerably fewer hybrids.

These findings suggest that global ancestry analyses are particularly sensitive to the effects of low-quality (LQ) genotypes compared to local ancestry methods, leading to a higher frequency of inconsistencies among methods. This variability stems from the fact that each method uses different algorithms to handle missing or erroneous data, which can cause discrepancies in admixture estimates. After removing low-quality samples, the Pearson correlations between methods improved significantly, indicating that data quality plays a crucial role in ensuring reliable and consistent ancestry results. This result was consistent with the result from the earlier study (Kendall et al., 2024), which found a high correlation between global ancestry (ADMIXTURE) and local ancestry (RFMix)

when the quality of input genotypes was improved after phasing and imputation (Kendall et al., 2024).

In addition, it is possible that the high admixture proportions in genotyped individuals may result in an increased proportion of missing data, contributing to low-quality genotyping. If the genotyping array is not well-suited or specifically designed for more than one species, the inclusion of admixed samples may contribute to a higher number of LQ samples. Therefore, we do not recommend the exclusion of all low-quality samples, as this may lead to the unintended removal of individuals with admixed ancestry. For instance, samples such as WIRA631_LQ, JBEL598_LQ, and JROM10658_LQ, despite their lower quality, were consistently identified as F1 hybrids across all chromosomes. Instead, we recommend the selective removal of LQ samples that produce highly inconsistent results across different analytical approaches.

Subpopulation structures

Population structure is a potential confounding factor in global and local ancestry analysis (Martin et al., 2018; Schubert et al., 2020). Using a regional dataset from India, we observed that population structure significantly impacted the ancestry inference. For instance, all Himalayan wolves (*Canis lupus chanco*) were identified as backcrosses or F1 hybrids with dogs based on the ADMIXTURE analysis. However, the ADMIXTURE analysis based on the entire dataset with sampling range covering the entire Eurasia, Himalayan wolves were found to have lower admixture proportions. In the local ancestry analysis (three-way-ELAI), Himalayan wolves showed more than 98% wolf ancestry. This discrepancy demonstrates that limited datasets containing regional subpopulations can lead to misclassification in global ancestry analyses, while local ancestry methods may also be inaccurate if the reference panels do not adequately represent subpopulation diversity. Therefore, when analyzing geographically widespread datasets, it is important to compare results from regional and global datasets to assess consistency. Additionally, conducting a population structure analysis prior to hybridization analysis is crucial to identify any underlying substructure that could confound the results.

2.5. Conclusion

Despite numerous attempts to identify admixture rates, accurately detecting hybridization using genomic data remains challenging (Kong and Kubatko, 2021). Global and local ancestry methods are the most popular techniques for inferring ancestries. While the individual admixture proportions inferred using different methods were generally consistent, some inconsistencies were observed. These discrepancies can be attributed to the effects of demographic histories and intra-specific population structure, the quality of the input dataset, as well as the varying assumptions of the methods, and the limitations of the methods themselves. Although global ancestry methods are widely used, they may not be suitable as the sole methods of inferring admixture as they are more prone to confounding factors like the presence of population structure. We recommend using local ancestry analysis alongside global ancestry

methods, with local ancestry results being prioritized for precise interpretation of admixture rates.

2.6. Bibliography

Alexander, D. H., & Lange, K. (2011). Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. BMC Bioinformatics, 12, 1-6.

Alexander, D. H., Novembre, J., & Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. Genome Research, 19(9), 1655-1664.

Alizadeh, F., Jazayeriy, H., Jazayeri, O., & Vafaee, F. (2023). AICRF: ancestry inference of admixed population with deep conditional random field. Journal of Genetics, 102(2), 49.

Anderson, E. C., & Dunham, K. K. (2008). The influence of family groups on inferences made with the program Structure. Molecular Ecology Resources, *8*(6), 1219-1229.

Anderson, E. C., & Thompson, E. (2002). A model-based method for identifying species hybrids using multilocus genetic data. Genetics, 160(3), 1217-1229.

Avadhanam, S., & Williams, A. L. (2023). Phase-free local ancestry inference mitigates the impact of switch errors on phase-based methods. BioRxiv, 2023-12.

B L Browning, X Tian, Y Zhou, and S R Browning (2021). Fast two-stage phasing of large-scale sequence data. The American Journal of Human Genetics, 108(10):1880-1890. doi:10.1016/j.ajhg.2021.08.005

Baran, Y., Pasaniuc, B., Sankararaman, S., Torgerson, D. G., Gignoux, C., Eng, C., ... & Halperin, E. (2012). Fast and accurate inference of local ancestry in Latino populations. Bioinformatics, 28(10), 1359-1367.

Barilani, M., Sfougaris, A., Giannakopoulos, A., Mucci, N., Tabarroni, C., & Randi, E. (2007). Detecting introgressive hybridisation in rock partridge populations (*Alectoris graeca*) in Greece through Bayesian admixture analyses of multilocus genotypes. Conservation Genetics, 8, 343-354.

Björklund, M. (2019). Be careful with your principal components. Evolution, 73(10), 2151-2158.

Browning, B. L., & Browning, S. R. (2011). A fast, powerful method for detecting identity by descent. The American Journal of Human Genetics, 88(2), 173-182.

Browning, B. L., Tian, X., Zhou, Y., & Browning, S. R. (2021). Fast two-stage phasing of large-scale sequence data. The American Journal of Human Genetics, 108(10), 1880-1890.

Caliebe, A., Tekola-Ayele, F., Darst, B. F., Wang, X., Song, Y. E., Gui, J., ... & IGES ELSI Committee. (2022). Including diverse and admixed populations in genetic epidemiology research. Genetic Epidemiology, 46(7), 347-371.

Campbell, C. L., Bhérer, C., Morrow, B. E., Boyko, A. R., & Auton, A. (2016). A pedigree-based map of recombination in the domestic dog genome. G3: Genes, Genomes, Genetics, 6(11), 3517-3524.

Chang, C.C., Chow, C.C., Tellier, L.C., Vattikuti, S., Purcell, S.M., Lee, J.J., 2015. Second-generation PLINK: rising to the challenge of larger and richer datasets. Gigascience 2015 (4), 7.

Chari, T., Banerjee, J. & Pachter, L. The specious art of single-cell genomics. BioRxiv.

Chen, M., Yang, C., Li, C., Hou, L., Chen, X., & Zhao, H. (2014, June). Admixture mapping analysis in the context of GWAS with GAW18 data. In BMC Proceedings (Vol. 8, pp. 1-5). BioMed Central.

Clay, L., Paterson, M. B., Bennett, P., Perry, G., & Phillips, C. C. (2020). Do behaviour assessments in a shelter predict the behaviour of dogs post-adoption?. Animals, 10(7), 1225.

Conomos, M. P., Miller, M. B., & Thornton, T. A. (2015). Robust inference of population structure for ancestry prediction and correction of stratification in the presence of relatedness. Genetic Epidemiology, 39(4), 276-293.

Cordoni, G., & Palagi, E. (2019). Back to the future: A glance over wolf social behavior to understand dog–human relationship. Animals, 9(11), 991.

Cottin, A., Penaud, B., Glaszmann, J. C., Yahiaoui, N., & Gautier, M. (2020). Simulation-based evaluation of three methods for local ancestry deconvolution of non-model crop species genomes. G3: Genes, Genomes, Genetics, 10(2), 569-579.

Dagilis, A. J., Peede, D., Coughlan, J. M., Jofre, G. I., D'Agostino, E. R., Mavengere, H., ... & Matute, D. R. (2021). 15 years of introgression studies: quantifying gene flow across Eukaryotes. BioRxiv, 2021-06.

Dattani, S., Howard, D. M., Lewis, C. M., & Sham, P. C. (2022). Clarifying the causes of consistent and inconsistent findings in genetics. Genetic Epidemiology, 46(7), 372-389.

Dias-Alves, T., Mairal, J., & Blum, M. G. (2018). Loter: a software package to infer local ancestry for a wide range of species. Molecular Biology and Evolution, 35(9), 2318-2326.

Dowling, T. E., & Secor, C. L. (1997). The role of hybridization and introgression in the diversification of animals. Annual review of Ecology and Systematics, 28(1), 593-619.

Duan, Q., Xu, Z., Raffield, L. M., Chang, S., Wu, D., Lange, E. M., ... & Li, Y. (2018). A robust and powerful two-step testing procedure for local ancestry adjusted allelic association analysis in admixed populations. Genetic Epidemiology, 42(3), 288-302.

Dziech, A. (2021). Identification of wolf-dog hybrids in Europe–an overview of genetic studies. Frontiers in Ecology and Evolution, 9, 760160.

Edelman, N. B., & Mallet, J. (2021). Prevalence and adaptive impact of introgression. Annual Review of Genetics, 55, 265-283.

Elhaik, E. (2022). Principal Component Analyses (PCA)-based findings in population genetic studies are highly biased and must be reevaluated. Scientific Reports, 12(1), 14683.

Elhaik, E., Tatarinova, T., Chebotarev, D., Piras, I. S., Maria Calò, C., De Montis, A., ... & Wells, R. S. (2014). Geographic population structure analysis of worldwide human populations infers their biogeographical origins. Nature Communications, 5(1), 3513.

Elworth, R. L., Ogilvie, H. A., Zhu, J., & Nakhleh, L. (2019). Advances in computational methods for phylogenetic networks in the presence of hybridization. Bioinformatics and phylogenetics: Seminal Contributions of Bernard Moret, 317-360.

Farhat, N., Lazebnik, T., Monteny, J., Moons, C. P. H., Wydooghe, E., van der Linden, D., & Zamansky, A. (2023). Digitally-enhanced dog behavioral testing. Scientific Reports, 13(1), 21252.

Galov, A., Fabbri, E., Caniglia, R., Arbanasić, H., Lapalombella, S., Florijančić, T., ... & Randi, E. (2015). First evidence of hybridization between golden jackal (*Canis aureus*) and domestic dog (*Canis familiaris*) as revealed by genetic markers. Royal Society Open Science, 2(12), 150450.

Garcia-Erill, G., & Albrechtsen, A. (2020). Evaluation of model fit of inferred admixture proportions. Molecular Ecology Resources, 20(4), 936-949.

Geza, E., Mugo, J., Mulder, N. J., Wonkam, A., Chimusa, E. R., & Mazandu, G. K. (2019). A comprehensive survey of models for dissecting local ancestry deconvolution in human genome. Briefings in Bioinformatics, 20(5), 1709-1724.

Giavarina, D. (2015). Understanding bland altman analysis. Biochemia Medica, 25(2), 141-151.

Gopalakrishnan, S., Sinding, M. H. S., Ramos-Madrigal, J., Niemann, J., Castruita, J. A. S., Vieira, F. G., ... & Gilbert, M. T. P. (2018). Interspecific gene flow shaped the evolution of the genus *Canis*. Current Biology, 28(21), 3441-3449.

Goulet, B. E., Roda, F., & Hopkins, R. (2017). Hybridization in plants: old ideas, new techniques. Plant Physiology, 173(1), 65-78.

Gravel, S. (2012). Population genetics models of local ancestry. Genetics, 191(2), 607-619.

Guan, Y. (2014). Detecting structure of haplotypes and local ancestry. Genetics, 196(3), 625-642

Harmoinen, J., von Thaden, A., Aspi, J., Kvist, L., Cocchiararo, B., Jarausch, A., ... & Nowak, C. (2021). Reliable wolf-dog hybrid detection in Europe using a reduced SNP panel developed for non-invasively collected samples. BMC Genomics, 22(1), 473.

Hartigan, J. A., & Wong, M. A. (1979). A k-means clustering algorithm. Applied statistics, 28(1), 100-108.

Hibbins, M. S., & Hahn, M. W. (2022). Phylogenomic approaches to detecting and characterizing introgression. Genetics, 220(2), iyab173.

Kendall, C., Robinson, J., Debortoli, G., Nooranikhojasteh, A., Christian, D., Newman, D., ... & Viola, B. (2024). Global and local ancestry estimation in a captive baboon colony. Plos One, 19(7), e0305157.

Kidd, J. M., Gravel, S., Byrnes, J., Moreno-Estrada, A., Musharoff, S., Bryc, K., ... & Bustamante, C. D. (2012). Population genetic inference from personal genome data: impact of ancestry and admixture on human genomic variation. The American Journal of Human Genetics, 91(4), 660-671.

Ko, S., Chu, B. B., Peterson, D., Okenwa, C., Papp, J. C., Alexander, D. H., ... & Lange, K. L. (2023). Unsupervised discovery of ancestry-informative markers and genetic admixture proportions in biobank-scale datasets. The American Journal of Human Genetics, 110(2), 314-325.

Kong, S., & Kubatko, L. S. (2021). Comparative performance of popular methods for hybrid detection using genomic data. Systematic Biology, 70(5), 891-907.

Lafferty, J., McCallum, A., & Pereira, F. (2001, June). Conditional random fields: Probabilistic models for segmenting and labeling sequence data. In International Conference on Machine Learning (Vol. 1, No. 2, p. 3).

Lawson, D. J., Davies, N. M., Haworth, S., Ashraf, B., Howe, L., Crawford, A., ... & Timpson, N. J. (2020). Is population structure in the genetic biobank era irrelevant, a challenge, or an opportunity?. Human Genetics, 139, 23-41.

Lawson, D. J., Van Dorp, L., & Falush, D. (2018). A tutorial on how not to over-interpret STRUCTURE and ADMIXTURE bar plots. Nature Communications, 9(1), 3258.

Lazaridis, I., Patterson, N., Mittnik, A., Renaud, G., Mallick, S., Kirsanow, K., ... & Krause, J. (2014). Ancient human genomes suggest three ancestral populations for present-day Europeans. Nature, 513(7518), 409-413.

Liu, C. C., Shringarpure, S., Lange, K., & Novembre, J. (2020). Exploring population structure with admixture models and principal component analysis. Methods in Molecular Biology, 2090, 67-86.

Ma, J., & Amos, C. I. (2012). Principal components analysis of population admixture. PloS one, 7(7), e40115.

Ma, X., Fu, Y., Zhao, X., Jiang, L., Zhu, Z., Gu, P., ... & Tan, L. (2016). Genomic structure analysis of a set of Oryza nivara introgression lines and identification of yield-associated QTLs using whole-genome resequencing. Scientific Reports, 6(1), 27425.

Mallet, J. (2005). Hybridization as an invasion of the genome. Trends in ecology & evolution, 20(5), 229-237.

Maples, B. K., Gravel, S., Kenny, E. E., & Bustamante, C. D. (2013). RFMix: a discriminative modeling approach for rapid and robust local-ancestry inference. The American Journal of Human Genetics, 93(2), 278-288.

Martin, E. R., Tunc, I., Liu, Z., Slifer, S. H., Beecham, A. H., & Beecham, G. W. (2018). Properties of global-and local-ancestry adjustments in genetic association tests in admixed populations. Genetic Epidemiology, 42(2), 214-229.

Mazandu, G. K., Geza, E., Seuneu, M., & Chimusa, E. R. (2019). Orienting future trends in local ancestry deconvolution models to optimally decipher admixed individual genome variations. Bioinformatics Tools for Detection and Clinical Interpretation of Genomic Variations, 35.

McFarlane, S. E., & Pemberton, J. M. (2019). Detecting the true extent of introgression during anthropogenic hybridization. Trends in Ecology & Evolution, 34(4), 315-326.

McVean, G. (2009). A genealogical interpretation of principal components analysis. PLoS Genetics, 5(10), e1000686.

Moroni, B., Brambilla, A., Rossi, L., Meneguz, P. G., Bassano, B., & Tizzani, P. (2022). Hybridization between Alpine Ibex and Domestic Goat in the Alps: A Sporadic and Localized Phenomenon?. Animals, 12(6), 751.

Nason, J. D., & Ellstrand, N. C. (1993). Estimating the frequencies of genetically distinct classes of individuals in hybridized populations. Journal of Heredity, 84(1), 1-12.

Netto, W. J., & Planta, D. J. (1997). Behavioural testing for aggression in the domestic dog. Applied Animal Behaviour Science, 52(3-4), 243-263.

Novembre, J., & Stephens, M. (2008). Interpreting principal component analyses of spatial population genetic variation. Nature Genetics, 40(5), 646-649.

Oliveira, S., Marchi, N., & Excoffier, L. (2024). Assessing the limits of local ancestry inference from small reference panels. Molecular Ecology Resources, e13981.

Ottenburghs, J. (2021). The genic view of hybridization in the Anthropocene. Evolutionary Applications, 14(10), 2342-2360.

Ottenburghs, J. (2023). How common is hybridization in birds?. Journal of Ornithology, 164(4), 913-920.

Padhukasahasram, B. (2014). Inferring ancestry from population genomic data and its applications. Frontiers in Genetics, 5, 204.

Patterson, N., Price, A. L., & Reich, D. (2006). Population structure and eigenanalysis. PLoS Genetics, 2(12), e190.

Payseur, B. A., & Rieseberg, L. H. (2016). A genomic perspective on hybridization and speciation. Molecular Ecology, 25(11), 2337-2360.

Pearson, K. (1895). Note on regression and inheritance in the case of two parents. Proceedings of the Royal Society of London, 58, 240–242.

Pilot, M., Greco, C., vonHoldt, B. M., Randi, E., Jędrzejewski, W., Sidorovich, V. E., ... & 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(5), 662-680.

Pilot, M., Moura, A. E., Okhlopkov, I. M., Mamaev, N. V., Manaseryan, N. H., Hayrapetyan, V., ... & Bogdanowicz, W. (2021). Human-modified canids in human-modified landscapes: The evolutionary consequences of hybridization for grey wolves and free-ranging domestic dogs. Evolutionary Applications, 14(10), 2433-2456.

Porto-Foresti, F., Hashimoto, D. T., Prado, F. D., Senhorini, J. A., & Foresti, F. (2013). Genetic markers for the identification of hybrids among catfish species of the family Pimelodidae. Journal of Applied Ichthyology, 29(3), 643-647.

Price, A. L., Patterson, N. J., Plenge, R. M., Weinblatt, M. E., Shadick, N. A., & Reich, D. (2006). Principal components analysis corrects for stratification in genome-wide association studies. Nature Genetics, 38(8), 904-909.

Price, A. L., Tandon, A., Patterson, N., Barnes, K. C., Rafaels, N., Ruczinski, I., ... & Myers, S. (2009). Sensitive detection of chromosomal segments of distinct ancestry in admixed populations. PLoS Genetics, 5(6), e1000519.

Pritchard, J. K., Stephens, M., & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. Genetics, 155(2), 945-959.

Racimo, F., Sankararaman, S., Nielsen, R., & Huerta-Sánchez, E. (2015). Evidence for archaic adaptive introgression in humans. Nature Reviews Genetics, 16(6), 359-371.

Raj, A., Stephens, M., & Pritchard, J. K. (2014). fastSTRUCTURE: variational inference of population structure in large SNP data sets. Genetics, 197(2), 573-589.

Randi, E. (2008). Detecting hybridization between wild species and their domesticated relatives. Molecular Ecology, 17(1), 285-293.

Randi, E., Hulva, P., Fabbri, E., Galaverni, M., Galov, A., Kusak, J., ... & Caniglia, R. (2014). Multilocus detection of wolf x dog hybridization in Italy, and guidelines for marker selection. PLoS One, 9(1), e86409.

Rendón-Anaya, M., Wilson, J., Sveinsson, S., Fedorkov, A., Cottrell, J., Bailey, M. E., ... & Ingvarsson, P. K. (2021). Adaptive introgression facilitates adaptation to high latitudes in European aspen (*Populus tremula* L.). Molecular Biology and Evolution, 38(11), 5034-5050.

Rieseberg, L. H., & Carney, S. E. (1998). Plant hybridization. The New Phytologist, 140(4), 599-624.

Rodriguez, J. M., Bercovici, S., Elmore, M., & Batzoglou, S. (2013). Ancestry inference in complex admixtures via variable-length Markov chain linkage models. Journal of Computational Biology, 20(3), 199-211.

Rogers, S. M., Isabel, N., & Bernatchez, L. (2007). Linkage maps of the dwarf and normal lake whitefish (*Coregonus clupeaformis*) species complex and their hybrids reveal the genetic architecture of population divergence. Genetics, 175(1), 375-398.

Salter-Townshend, M., & Myers, S. (2019). Fine-scale inference of ancestry segments without prior knowledge of admixing groups. Genetics, 212(3), 869-889.

Sankararaman, S. (2020). Methods for detecting introgressed archaic sequences. Current Opinion in Genetics & Development, 62, 85-90.

Sankararaman, S., Mallick, S., Patterson, N., & Reich, D. (2016). The combined landscape of Denisovan and Neanderthal ancestry in present-day humans. Current Biology, 26(9), 1241-1247.

Schubert, R., Andaleon, A., & Wheeler, H. E. (2020). Comparing local ancestry inference models in populations of two-and three-way admixture. PeerJ, 8, e10090.

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: Biological Sciences, 363(1505), 2805-2811.

Secolin, R., Mas-Sandoval, A., Arauna, L. R., Torres, F. R., de Araujo, T. K., Santos, M. L., ... & Comas, D. (2019). Distribution of local ancestry and evidence of adaptation in admixed populations. Scientific Reports, 9(1), 13900.

Seldin, M. F., Pasaniuc, B., & Price, A. L. (2011). New approaches to disease mapping in admixed populations. Nature Reviews Genetics, 12(8), 523-528.

Stefanović, M., Bogdanowicz, W., Adavoudi, R., Martínez-Sosa, F., Doan, K., Flores-Manzanero, A., ... & Pilot, M. (2024). Range-wide phylogeography of the golden jackals

(*Canis aureus*) reveals multiple sources of recent spatial expansion and admixture with dogs at the expansion front. Biological Conservation, 290, 110448.

Stronen, A. V., Mattucci, F., Fabbri, E., Galaverni, M., Cocchiararo, B., Nowak, C., ... & Caniglia, R. (2022). A reduced SNP panel to trace gene flow across southern European wolf populations and detect hybridization with other Canis taxa. Scientific Reports, 12(1), 4195.

Suarez-Pajes, E., Díaz-de Usera, A., Marcelino-Rodríguez, I., Guillen-Guio, B., & Flores, C. (2021). Genetic ancestry inference and its application for the genetic mapping of human diseases. International Journal of Molecular Sciences, 22(13), 6962.

Sun, Q., Horimoto, A. R., Chen, B., Ockerman, F., Mohlke, K. L., Blue, E., ... & Li, Y. (2025). Opportunities and challenges of local ancestry in genetic association analyses. The American Journal of Human Genetics, 112(4), 727-740.

Tan, T., & Atkinson, E. G. (2023). Strategies for the Genomic Analysis of Admixed Populations. Annual Review of Biomedical Data Science, 6.

Tang, H., Choudhry, S., Mei, R., Morgan, M., Rodriguez-Cintron, W., Burchard, E. G., & Risch, N. J. (2007). Recent genetic selection in the ancestral admixture of Puerto Ricans. The American Journal of Human Genetics, 81(3), 626-633.

Taylor, S. A., & Larson, E. L. (2019). Insights from genomes into the evolutionary importance and prevalence of hybridization in nature. Nature Ecology & Evolution, 3(2), 170-177.

Thawornwattana, Y., Huang, J., Flouri, T., Mallet, J., & Yang, Z. (2023). Inferring the direction of introgression using genomic sequence data. BioRxiv, 2023-06.

Thornton, T. A., & Bermejo, J. L. (2014). Local and global ancestry inference and applications to genetic association analysis for admixed populations. Genetic Epidemiology, 38(S1), S5-S12.

Todesco, M., Pascual, M. A., Owens, G. L., Ostevik, K. L., Moyers, B. T., Hübner, S., ... & Rieseberg, L. H. (2016). Hybridization and extinction. Evolutionary Applications, 9(7), 892-908.

Tyagi, A., Godbole, M., Vanak, A. T., & Ramakrishnan, U. (2023). Citizen science facilitates first ever genetic detection of wolf-dog hybridization in Indian savannahs. Ecology and Evolution, 13(5), e10100.

Uren, C., Hoal, E. G., & Möller, M. (2020). Putting RFMix and ADMIXTURE to the test in a complex admixed population. BMC Genetics, 21, 1-8.

Utsunomiya, Y. T., Milanesi, M., Barbato, M., Utsunomiya, A. T. H., Sölkner, J., Ajmone-Marsan, P., & Garcia, J. F. (2020). Unsupervised detection of ancestry tracks with the GHap r package. Methods in Ecology and Evolution, 11(11), 1448-1454

van Waaij, J., Li, S., Garcia-Erill, G., Albrechtsen, A., & Wiuf, C. (2023). Evaluation of population structure inferred by principal component analysis or the admixture model. arXiv preprint arXiv:2302.04596.

Vi, T., Vigouroux, Y., Cubry, P., Marraccini, P., Phan, H. V., Khong, G. N., & Poncet, V. (2023). Genome-wide admixture mapping identifies wild ancestry-of-origin segments in cultivated Robusta coffee. Genome Biology and Evolution, 15(5), evad065.

Wang, J. (2022). Fast and accurate population admixture inference from genotype data from a few microsatellites to millions of SNPs. Heredity, 129(2), 79-92.

Wang, M. S., Wang, S., Li, Y., Jhala, Y., Thakur, M., Otecko, N. O., ... & Wu, D. D. (2020). Ancient hybridization with an unknown population facilitated high-altitude adaptation of canids. Molecular Biology and Evolution, 37(9), 2616-2629.

Wang, X., Zhu, X., Qin, H., Cooper, R. S., Ewens, W. J., Li, C., & Li, M. (2011). Adjustment for local ancestry in genetic association analysis of admixed populations. Bioinformatics, 27(5), 670-677.

Wangkumhang, P., & Hellenthal, G. (2018). Statistical methods for detecting admixture. Current Opinion in Genetics & Development, 53, 121-127.

Werhahn, G., Liu, Y., Meng, Y., Cheng, C., Lu, Z., Atzeni, L., ... & Senn, H. (2020). Himalayan wolf distribution and admixture based on multiple genetic markers. Journal of Biogeography, 47(6), 1272-1285.

Werhahn, G., Senn, H., Kaden, J., Joshi, J., Bhattarai, S., Kusi, N., ... & Macdonald, D. W. (2017). Phylogenetic evidence for the ancient Himalayan wolf: towards a clarification of its taxonomic status based on genetic sampling from western Nepal. Royal Society Open Science, 4(6), 170186.

Wu, J., Liu, Y., & Zhao, Y. (2021). Systematic review on local ancestor inference from a mathematical and algorithmic perspective. Frontiers in Genetics, 12, 639877.

Yuan, K., Zhou, Y., Ni, X., Wang, Y., Liu, C., & Xu, S. (2017). Models, methods and tools for ancestry inference and admixture analysis. Quantitative Biology, 5(3), 236-250.

Zhou, S., Ni, S., Dai, J., Zhou, Q., Zhou, R., & Liu, Y. (2020). Natural hybridization between Phyllagathis and Sporoxeia species produces a hybrid without reproductive organs. Plos One, 15(1), e02276.

3 Chapter 3

The Evolutionary Consequences of Hybridization in Grey Wolves, Golden Jackals, and Domestic Dogs

Abstract

Interspecific hybridization is a well-documented process, but its evolutionary consequences are poorly understood. Although hybridization can threaten species genetic integrity, it can also increase their adaptability by providing novel genetic variation. In this study, hybridization and adaptive introgression rates between two wild canids (Eurasian gray wolves and golden jackals) and free-ranging dogs have been evaluated using chromosome-level admixture analysis (ELAI). Our findings indicate that hybridization occurs across the distribution ranges of these canid species. The average dog ancestry proportions in admixed wolves and golden jackals (excluding F1 hybrids) were estimated at 6.4% and 1.2%, respectively. The average wolf ancestry proportion in admixed dogs was 1.4%, while no golden jackal ancestry was detected in dogs. The higher proportion of introgression in wolves and dogs compared to jackals may be explained by the closer evolutionary similarity between the first two canids. The average wolf ancestry proportion in admixed golden jackals was 1.7%, while no golden jackal ancestry was detected in wolves. In wolves and golden jackals, 114 and 94 genes were identified as candidate genes for adaptive introgression from dogs, respectively. In wolves, we observed adaptive introgression of MHC genes and the melanocortin receptor genes (MCRs). In golden jackals, the candidate adaptive genes from dogs were enriched for GO terms related to the ‘endoplasmic reticulum protein-containing complex (ER)’. Both MHC and ER play significant roles in the immune system. These findings support the hypothesis that gene flow from domesticated species, such as dogs, could enhance immune system diversity in wild canid populations. Free-ranging dogs appear to have acquired a large pool of beneficial genetic variants from wolves, which may have affected morphological, behavioural, and physiological traits. Among all genes that were under positive selection in canids, genes related to the immune and nervous systems were predominant in wild canids and free-ranging dogs, respectively. While gene flow from wolves may help mitigate some negative effects of domestication in free-ranging dogs, dogs may also introduce advantageous genetic variants into wild canid populations. Alongside of finding positive selection the signs of negative selection by identifying chromosomal blocks with underrepresented wolf ancestry in admixed dogs, and also chromosomal blocks with underrepresented wolf and dog ancestry in admixed jackals. These results suggest that some introgressed gene variants may also have a deleterious effect on these species, but they can be efficiently removed from their gene pools.

Keywords: Hybridization, Adaptive introgression, Canids

3.1. Introduction

Introgression results in the transfer of genetic variation between species, which can lead to various evolutionary consequences. Introgression from domesticated lineages into their wild relatives is one of the most commonly reported consequences of hybridization (Adavoudi and Pilot, 2021). While introgression has the potential to dilute a species' gene pool through genetic swamping, it is also recognized as an important source of adaptation (reviewed by Edelman and Mallet, 2021). As a result of natural selection pressures, domesticated species adapted to human-modified environments by developing new phenotypic variation (Fang et al., 2009; Wilkins et al., 2014; Milla et al., 2015; Solberg et al., 2020; Andersson and Purugganan, 2022). Progressing anthropogenic landscape transformation and habitat fragmentation have led to a rise in the population sizes of domestic species. This has resulted in more frequent contact between domestic and wild species, increasing the opportunities for hybridization.

Hybridization between domestic and wild species has been extensively documented across numerous animal and plant taxa (reviewed in: Gray, 2005; Todesco et al., 2016; McFarlane & Pemberton, 2019; Yadav et al., 2019; Purugganan, 2022; Westbrook and DiTommaso, 2023). In a rapidly changing environment, gaining advantageous variants through introgression may help species improve their fitness in a new habitat. This process might be more efficient than adapting through new mutations (Nelson et al., 2017), since the rate of acquisition of new alleles or haplotypes through introgression is faster than the mutation rate (Harrison et al., 2014).

Positive impacts of hybridization between wild and domestic species via adaptive introgression have been documented in several mammalian species. For instance, the adaptive introgression from domestic pigs (*Sus domesticus*) improved fitness and population growth in wild boars (*Sus scrofa*) (Mary et al., 2022). Moreover, adaptive introgression of the immune system gene variants from domestic goats (*Capra aegagrus hircus*) to Alpine ibex (*Capra ibex ibex*) has been observed (Grossen et al., 2014; Münger et al., 2024). Introgression as a source of adaptation has also been observed in cattle, where hybridization with yaks has played a key role in the adaptation of Tibetan cattle to high-altitude environments (Wu et al., 2018).

Hybridization within the family Canidae has been extensively documented in many studies (reviewed in Pendragon, 2011; Dziech, 2021; Adavoudi and Pilot, 2021). The domestic dog (*Canis lupus familiaris*) is the first species to have been domesticated, diverging from its primary ancestor, the gray wolf (*Canis lupus*), between 11,000 and 32,000 years ago (reviewed in Tancredi & Cardinali, 2023). The divergence between the golden jackal (*Canis aureus*) (hereafter called jackal) and the grey wolf occurred considerably earlier, between 1.5-2.4 million years ago (Koepfli et al., 2015). Due to underdeveloped reproductive barriers between these species, these species can interbreed and interbreed and produce fertile offspring (e.g. Randi, 2008; Harrison & Larson, 2014; Hindrikson et al., 2017). Growing anthropogenic landscape changes and habitat fragmentation have contributed to an increase in domestic dog population sizes. These large populations not only enhance dogs' adaptation to human-modified environments but also can lead to more frequent interactions with wild canids. Hybridization between

wild canids and domestic dogs may help admixed populations to quickly adapt to climate and landscape changes (Randi et al., 2014). For example, adaptation to high altitudes has been linked to adaptive introgression from ancient lineages in Tibetan dogs and wolves (Wang et al., 2020). On the other hand, high introgression rate from domestic dogs can also disrupt the local adaptation and unique genetic variation of wild canids and raise significant concerns for their conservation (e.g., Vilà and Wayne, 1999; Donfrancesco et al., 2019; Salvatori et al., 2019; Ninausz et al., 2023).

Hybridization in canids has been extensively studied, with particular attention given to its occurrence and rates. For example, the introgression rate of dog and wolf ancestries has been identified in hybrid samples between Eurasian wolf and free-ranging dog populations (Pilot et al., 2021). In this study, we analyzed SNP data from Eurasian wolf, golden jackal, and free-ranging dog populations to investigate the hybridization rate among these species. Since the evolutionary consequences of hybridization, particularly in cases involving wolf-jackal and jackal-domestic dog interactions, remain poorly understood, we assessed the hybridization pattern among these species. Since jackals are more distantly related to wolves and dogs, including jackal samples to this analysis will allow us to understand how the hybridization and introgression rates can be changed depending on the evolutionary distance between cross-breeding taxa. Additionally, it provides an opportunity to compare the effects of introgression from domestic dogs on two different wild canids, compare adaptive introgression rates in both species and assess whether it involves the same genes and/or the same biological processes. Furthermore, as jackals continue to expand their range, often into human-modified landscapes, it is important to investigate whether introgression from dogs or wolves plays a role in their adaptation to these novel environments.

The main aim of this study is to conduct a comparative analysis of hybridization and introgression rates between the domestic dog and two wild canids, and to explore how hybridization contributes to adaptation through the introgression of beneficial gene variants. By comparing patterns across species, we assess both shared and species-specific genomic responses to hybridization.

The main aim of this study is to compare the introgression rates from free-ranging dogs in wolves and golden jackals and explore how hybridization contributes to adaptation through the introgression of beneficial gene variants. This was achieved by (1) identifying genomic blocks showing overrepresentation of introgressed variants in wild canids and free-ranging dogs, (2) assessing how the degree of genetic differentiation between species can influence the proportion of introgressed gene variants under selection, (3) identifying genes within adaptive introgressed blocks that are under positive selection, and (4) determining the functions of these positively selected gene variants and their potential benefits for admixed individuals.

3.2. Methods

Sampling

We obtained tissue samples of gray wolves, jackals, and domestic dogs from across the Eurasian species distribution between 2018 and 2022. Saliva samples from domestic dogs were additionally collected using the PERFORMAgene animal DNA collection kit (DNA Genotek). Furthermore, tissue samples from a previous study (Rutkowski et al., 2015) were included. DNA of tissue samples was extracted using NucleoSpin Tissue Kit (Macherey Nagel, Duren, Germany) (For more details please see Chapter 2, method).

Dataset creation

We analysed a dataset consisting of wolf, jackal, and free-ranging dog (WJD) samples as well as three datasets subsampled from this first one: wolf and free-ranging dog (WD), jackal and free-ranging dog (JD), and wolf and jackal (WJ). These were the same datasets as those used as in the second chapter (see Chapter 2, Table S 2.1). A detailed description of the steps (e.g. sampling, genotyping, and data processing) involved in generating these datasets can be found in Chapter 2 (Methods section).

Admixture proportions in wolves, jackals and dogs

Based on the results of the previous chapter, we showed that local ancestry analysis is less influenced by confounding factors such as population substructure compared to global ancestry methods. ELAI and LAMP-LD produced more consistent results among local ancestry methods than Ghap. This could be because the Ghap analyses were run using imputed and phased data, which may have caused some admixed individuals to be misclassified as pure samples (see Chapter 2, Discussion). Therefore, we used ELAI v1.01 (Guan, 2014) to estimate individual ancestry proportions, with source populations defined based on LAMP-LD results (Table 3.1; for more details, please see Chapter 2, methods).

The datasets analysed using three-way and two-way ELAI analysis, respectively, were as follows: 568 admixed dogs, wolves, and jackals (the WJD dataset), 149 admixed dogs and 165 admixed wolves (the WD dataset), 124 admixed dogs and 191 admixed jackals (the JD dataset), and 252 admixed jackals and 5 admixed wolves (the WJ dataset). The admixed status of these samples was first identified in LAMP-LD and then reassessed in ELAI. To calculate the mean proportion of different ancestries in each species, first-generation hybrids were excluded. Moreover, since our results from the previous chapter showed that low genotype quality samples can cause bias in estimating ancestry proportions, we also calculated the proportions of hybrid ancestry only for samples with good quality data, and compared the results. The number of F1 hybrids and LQ samples removed from the WDJ dataset is presented in Table 3.2.

Table 3.1. Applied thresholds for selecting pure and admixed individuals for the ELAI analysis

Datasets	Threshold for pure wolves	Threshold for pure jackals	Threshold for pure dogs	Admixed samples (N)
WDJ dataset	0.01	0.003	0.003	568
WD dataset	0.01	-	0.003	314
JD dataset	-	0.005	0.001	315
WJ dataset	0.001	0.003	-	257

Table 3.2. Number of samples before and after removing first-generation hybrids.

Datasets	Total number of samples	First-generation hybrids	Number of LQ samples	Number of samples after removing F1 and LQ samples
WJD dataset	568	21	117	430
WD dataset	314	6	48	260
JD dataset	315	4	68	243
WJ dataset	257	10	71	176

Detection of chromosomal blocks with overrepresentation or underrepresentation of introgressed variants

For detecting chromosomal blocks with an overrepresentation or underrepresentation of introgressed variants, we focused on the results of ELAI in each dataset. The mean admixture proportions within each autosomal chromosome and across autosomal chromosomes in all datasets were calculated separately (i.e., dog ancestry in wolves and jackals, wolf ancestry in dogs and jackals, and jackal ancestry in wolves and dogs). Chromosomal blocks with hybrid ancestry greater than or lower than three standard deviations (SD) from the mean were identified for each chromosome, allowing us to detect blocks with overrepresented or underrepresentation hybrid ancestry, respectively. Since chromosomes are natural genetic units with independent recombination, overrepresented blocks were identified based on the SD at the level of individual chromosomes, rather than using the global SD across all autosomal loci (Pilot et al., 2021). These overrepresented chromosomal blocks in hybrid individuals containing excess introgressed ancestry are candidates for adaptive introgression, i.e. selection on introgressed variants. We only included ancestry blocks containing at least 10 consecutive SNPs to minimize the false-positive rate.

Identification of loci under positive selection

To identify loci under positive selection in each overrepresented introgressed chromosomal block, the integrated haplotype score (iHS) test (Voight et al., 2006), implemented in the rehh package version 3.2.2 (Gautier et al. 2017) in R was used. The iHS score is based on a ratio of extended haplotype homozygosity (EHH) associated with each allele which measures the decay of homozygosity in haplotypes as their length increases. An allele that increases in frequency rapidly due to selection will exhibit extended haplotype homozygosity over a greater distance than expected under a neutral model (Sabeti et al., 2002). Since the phased data is required before calculating iHS, the genotype data of wolf, dog, and jackal for each chromosome were phased using Beagle v. 5.4 (Browning et al., 2021), separately. After phasing, the iHS statistics and its two-sided p-value for each SNP were calculated using rehh package's functions scan_hh and ihh2ihs focusing on chromosomal blocks showing an overrepresentation of introgression. The phasing and the iHS test was conducted for wolf, dog, and jackal samples in each dataset separately. The iHS results showing $p < 0.05$ and $|iHS| > 2$, were considered as significant results. Therefore, SNPs selected based on this threshold were considered as SNP under positive selection. Correction for multiple testing was not applied since it was not recommended by Voight et al. (2006).

Gene ontology enrichment analysis

To identify protein-coding genes located within overrepresented chromosomal blocks, we used Ensembl (2024). The identified protein-coding genes in these blocks are candidates for adaptive introgression. For this purpose, the selected outlier SNPs in these regions (CAI SNPs) were lifted over to the 10K Boxer Tascha CanFam6 genome using the liftover tool in the UCSC Genome Browser (https://genome.ucsc.edu/cgi-bin/hgLiftOver) and annotated using the Ensembl Variant Effect Predictor tool (https://www.ensembl.org/info/docs/tools/vep/index.html). The Variant Effect Predictor is a powerful tool for analysing, annotating, and prioritizing genomic variants in both coding and non-coding regions (McLaren et al., 2016). Mapping was enabled only for SNPs that overlapped with at least one protein-coding Ensembl gene or were located within the flanking region of 100 kb downstream or upstream of a gene.

Functional characterization of the candidate genes for adaptive introgression

If more than one candidate gene overlapped with a CAI SNP, all genes were included for functional enrichment analyses. We created six gene sets, including dog ancestry in wolves and jackals, wolf ancestry in jackals and dogs, and also jackal ancestry in wolves and dogs. To test if the candidate genes are related to specific functions (e.g., Molecular Functions, Biological Processes, and Cellular Components), Gene Ontology (GO) enrichment analysis was performed for each gene set using the g:Profiler web server (Raudvere et al. 2019). Additionally, different gene sets were created by translating canine genes to orthologous human genes using the g:orth function in the g:Profiler web server, and enrichment analysis also were applied for these genes. A significance threshold of 0.05 was used and the p-value of the gene enrichment was corrected by Benjamini–Hochberg FDR (false discovery rate), as well as the more conservative g:SCS

(Set Counts and Sizes) false discovery rate correction method that accounts for multiple testing due to the overlap of functional terms (Reimand et al., 2007).

3.3. Results

The average proportion of different ancestries in admixed samples

Three-way ELAI

Among admixed wolves, five F1 hybrids between wolf and dog were identified (WSER466, WIRA616, WIndD2449, WBOS18, WSER483) (Fig 3.1; Table 3.4). These samples showed no jackal ancestry. In these samples, only three first individuals showed 50% dog ancestry in most of their chromosomes, and for the last two samples, the proportions of dog ancestry in each chromosome varied between 0 to 50%. Since the average proportions of dog ancestry based on the all 38 chromosomes were calculated around 0.55 and 0.45, these samples were considered as F1 hybrids, but their ancestry could have been more complex, e.g. they could have been F1 x F1 crosses. Among admixed jackals, three first-generation hybrids with dogs were identified (JROM10658_LQ, JHUN9531, JBEL598_LQ), and 50% dog ancestry has been shown in all chromosomes. These samples showed no wolf ancestry. We also found some samples that showed different proportions of the three ancestries in their all chromosomes. For instance, WIRA631_LQ showed 12% jackal, 44% dog, and 43% wolf ancestries at average, with consistent ancestry proportions across all autosomal chromosomes (Table 3.3, Fig 3.2). We also identified 12 individuals with jackal ancestry ranging from 41% to 58%, while their proportions of wolf and dog ancestries varied. For instance, sample JBUL78-19_LQ exhibited 52% jackal ancestry, along with 21% dog and 26% wolf ancestry, which was consistent across chromosomes (Table 3.3, Fig. 3.2). All dog samples, except those from India, had over 99% dog ancestry and were classified as pure dogs. After excluding first-generation hybrids (all samples with 40–55% ancestry from a single source), we estimated the mean proportions of dog, wolf, and jackal ancestry in admixed samples (Table 3.5). Additionally, to assess the impact of LQ samples, we recalculated the mean ancestry proportions after removing both LQ and F1 hybrid samples (Table 3.5). The inclusion or exclusion of LQ samples influenced the estimates of wolf ancestry in admixed jackals and dogs (except F1 hybrids). In only admixed jackal samples (except F1 hybrids), wolf ancestry was estimated at 1.7%, but this value dropped to nearly 0% when LQ samples were excluded. Similarly, in admixed dog samples, the estimated wolf ancestry decreased from 1.4% to 0.8% when LQ samples were removed.

It is important to note that non-admixed individuals were excluded from the ELAI analysis. To provide average ancestries for the whole population, including admixed (except F1 hybrids) and non-admixed individuals, we assumed that individuals from the reference panels had 100% ancestry from one species (Table 3.5).

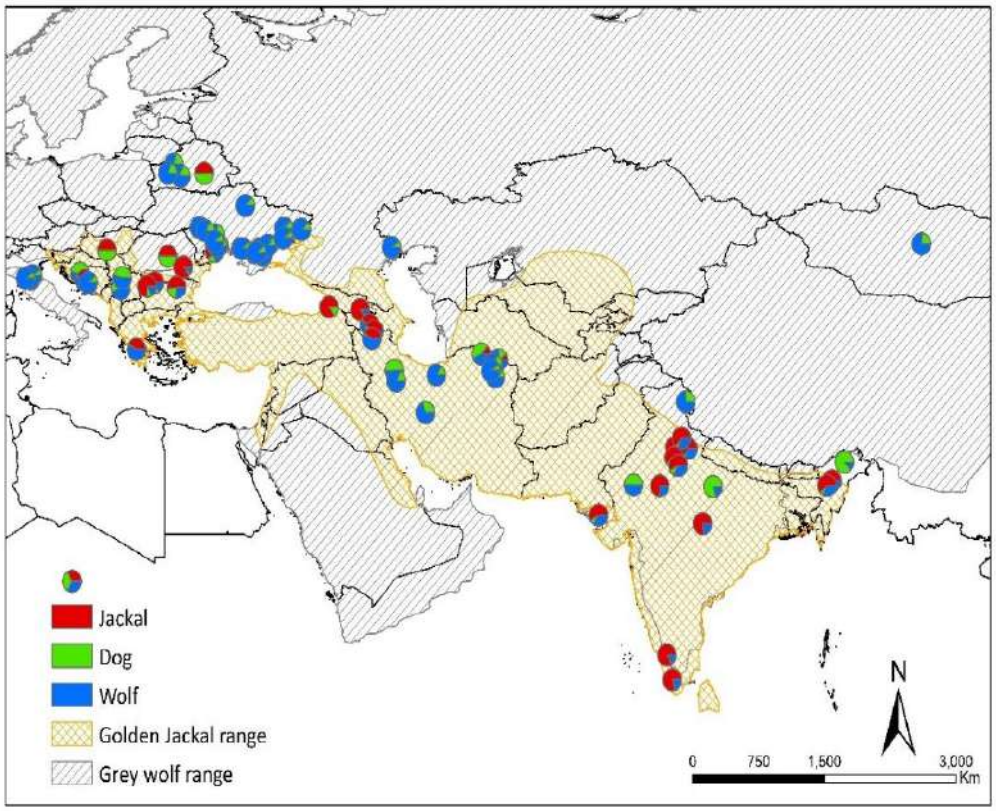


Fig. 3.1. The distribution range of admixed gray wolves and golden jackals in Eurasia.

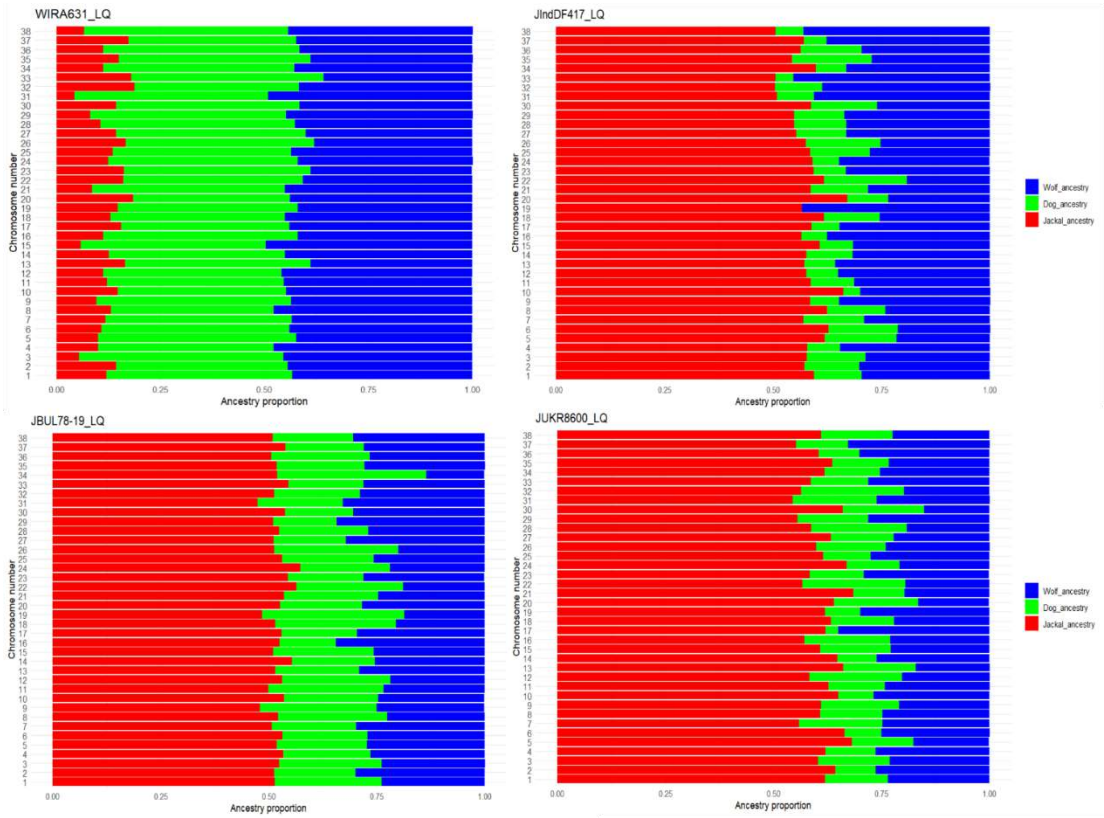


Fig. 3.2. The mean proportion of wolf, dog, and jackal ancestry in all autosomal chromosomes of individuals shown in Table 3.3.

Table 3.3. Samples showing more than 10% of different ancestries based on the three-way analysis.

Samples	Wolf ancestry	Jackal ancestry	Dog ancestry
WIRA631_LQ	0.432	0.125	0.442
JIndDF417_LQ	0.316	0.580	0.10
JBUL78-19_LQ	0.262	0.522	0.215
JUKR8600_LQ	0.239	0.614	0.146

Wolf-dog dataset

In the wolf-dog admixture analysis, the proportion of dog admixture in 165 potentially admixed wolves ranged from 0.005 to 0.53 (an average of 0.065, excluding F1 samples). The wolf admixture proportion in 149 potentially admixed dogs was lower, ranging from 0 to 0.20, and all but the Indian dogs showed less than 3% of wolf ancestry (an average of 0.017). 46 wolf samples had less than 90% wolf ancestry, including six individuals from India, Iran, and the Balkans, which were recognized as F1 hybrids (Table 3.4). Four of these individuals (WIndD2449, WIRA616, WIRA631_LQ, WSER466) showed 50% dog ancestry in most chromosomes. The proportions of dog ancestry in the two remaining individuals were between 0 to 50%. This result was aligned with the results of the three-way admixture, however, one sample (WIRA631_LQ), showed 12% jackal ancestry and 44% dog ancestry in the three-way analysis (Fig. 3.2). The jackal ancestry could not be inferred in the two-way wolf-dog admixture analysis. After excluding first-generation hybrids, the mean proportions of dog and wolf ancestry in the gene pools of admixed wolf and dog samples were estimated at 6.5% and 1.7%, respectively (Table 3.6). Removing LQ samples didn't change the proportion of dog ancestry in hybrid wolves, however, the proportion of wolf ancestry in hybrid dogs dropped slightly (1.7% vs. 1.1%).

Jackal-dog dataset

In the jackal-dog admixture analysis, the proportion of dog admixture in 191 potentially admixed jackals ranged from 0.002 to 0.5 (an average of 0.041, excluding F1 samples). Most dog samples showed limited signs of admixed ancestry and carried no more than 0.3% of chromosomal blocks originating from recent admixture with jackals. Three jackal samples from the Balkans and one sample from Belarus were identified as F1 hybrids since they displayed 50% dog ancestry in their chromosomes (JBUL78-19_LQ, JROM10658_LQ, JHUN9531, JBEL598_LQ) (Table 3.4). These individuals were also identified as F1 hybrids based on the three-way analysis, except for one sample (JBUL78-19_LQ), which showed 21% dog and 26% wolf ancestries. Most of the samples that showed 50% jackal ancestry with different proportions of dog and wolf ancestry in the WDJ dataset, showed 68-75% jackal ancestry proportions based on the JD dataset. After excluding first-generation hybrids, the mean proportions of jackal and dog ancestry in the gene pools of dogs and jackals were estimated at 0.1% and 4.1%, respectively. Removing LQ samples caused a significant drop in the proportion of dog ancestry in jackal samples (4.1% vs 1.3%) (Table 3 6).

Wolf-jackal dataset

In the jackal-wolf admixture analysis, among the five potentially admixed wolves that were included in the analysis, three samples from India showed more than 99.99% wolf ancestry and therefore were identified as pure wolf samples. Two wolf samples from Iran (WIRA631_LQ, WIRA1042_LQ) showed 22% and 10% jackal ancestry, respectively (Table 3.4). However, based on the results of the three-way analysis, WIRA631_LQ represents 12% jackal and 44% dog ancestry. Sample WIRA1042_LQ showed 9% jackal and 15% dog ancestry. Due to the small number of admixed individuals in the wolf–jackal dataset (only five admixed wolves), and the risk of producing biased results, we did not include these results here. For details, please refer to the Supplementary Material, Chapter 3. The proportion of wolf admixture in 252 potentially admixed jackals ranged from 0.0001 to 0.5 (an average of 0.031, excluding F1 hybrids). Out of 37 jackal samples that showed less than 90% jackal ancestry, four samples were identified as first-generation hybrids, showing 50-51% jackal ancestry. These samples were also identified as F1 hybrids using the JD and WDJ datasets. All samples showed around 50% jackal ancestry, and mixed proportions of dog/wolf ancestry in the WJD dataset showed 30-40% jackal ancestry proportions based on the WJ dataset. The high rate of hybridization between wolf and jackal was found in India since 19 of 28 samples had less than 90% jackal ancestry. These results were consistent with results from the three-way admixture analysis since jackals from India showed higher proportions of wolf ancestry compared to dog ancestry. After removing F1 hybrids, the mean proportions of wolf ancestry in the gene pools of admixed jackals were estimated at 3.1% (Table 3.6). After removing also LQ samples, the proportion significantly dropped (Table 3.6).

Table 3.4. Putative F1 hybrids identified based on ELAI results using all datasets.

Sample_ID	ELAI (WDJ)	ELAI (WD)	ELAI (JD)	ELAI (WJ)
WSER466	✓	✓		☒
WSER483	✓	✓		☒
WBOS18	✓	✓		☒
WIRA616	✓	✓		☒
WIndD2449	✓	✓		☒
WIRA631_LQ	☒	✓		☒
JBUL78_LQ	✓		✓	✓
JROM10658_LQ	✓		✓	✓
JHUN9531	✓		✓	✓
JBEL598_LQ	✓		✓	✓
JIndD3018_LQ	☒		☒	✓
JIndD45_LQ	✓		✓	✓
JIndD480_LQ	✓		✓	✓
JIndDF417_LQ	✓		✓	✓
JKAU8341_LQ	✓		☒	✓
JUKR8600_LQ	☒		☒	✓
JWKAU5740_LQ	✓		☒	✓
JURK8926_LQ	✓		☒	✓
JWKAU5741_LQ	✓		✓	✓
JKAU8321_LQ	✓		✓	✓
JWBOS38_LQ	✓		✓	✓
JGRE9066_LQ	✓		✓	✓
JKAU8086_LQ	✓		✓	✓

✓ Samples show hybrid ancestry in most of their chromosomes.

✓ Samples with pure ancestry in some chromosomes

✓ Samples show between 40-58% jackal ancestry and mixed proportions of dog/wolf ancestry

✓ Samples with proportions highly deviating from 50% (between 30%-70%)

☒ Samples not identified as F1 hybrids or showing hybrid ancestry less than 30%

Table 3.5. The mean percentage of different ancestries in admixed and all dogs, wolves, and jackals based on three-way analysis, after removing the first-generation hybrids and LQ samples. To provide average ancestries for the whole populations, including admixed and non-admixed individuals, we assumed that individuals from the reference panels had 100% ancestry from one species.

	Removing just the first-generation hybrids			Removing the first-generation hybrids and LQ samples			Considering all samples* (admixed and non-admixed)		
	Admixed wolves	Admixed jackals	Admixed dogs	Admixed wolves	Admixed jackals	Admixed dogs	All wolves	All jackals	All dogs
Wolf ancestry	-	0.017	0.014	-	0.000	0.008	-	0.009	0.003
Jackal ancestry	0.000	-	0.000	0.000	-	0.000	0.000	-	0.000
Dog ancestry	0.064	0.012	-	0.064	0.011	-	0.033	0.006	-

*All F1 samples were excluded from the calculation of average percentage of different ancestries.

Table 3.6. The mean percentage of different ancestries in admixed dogs, wolves, and jackals based on two-way ELAI analyses, after removing the first-generation hybrids and LQ samples. To provide average ancestries for the whole populations, including admixed and non-admixed individuals, we assumed that individuals from the reference panels had 100% ancestry from one species.

	Removing just the first-generation hybrids			Removing the first-generation hybrids and LQ samples			Considering all samples* (admixed and non-admixed)		
	Admixed wolves	Admixed jackals	Admixed dogs	Admixed wolves	Admixed jackals	Admixed dogs	All wolves	All jackals	All dogs
Wolf ancestry	-	0.031	0.017	-	0.012	0.011	-	0.016	0.004
Jackal ancestry	-	-	0.001	-	-	0.001	-	-	0.000
Dog ancestry	0.065	0.041	-	0.065	0.013	-	0.033	0.016	-

*All F1 samples were excluded from the calculation of average percentage of different ancestries.

Chromosomal blocks with overrepresentation or underrepresentation of introgressed variants

Wolf-dog dataset

The number and length of chromosomal blocks with significantly overrepresented or underrepresented hybridization-derived ancestry were identified based on the calculated standard deviation within each chromosome. In the wolf-dog dataset, eight blocks with significantly overrepresented dog ancestry were identified on seven chromosomes in wolves. The average proportion of dog ancestry within these blocks ranged from 0.104 to 0.157, with a global average of 0.127 across all blocks (Figure 3.3, Table S 3.1). In dogs, 31 blocks with overrepresented wolf ancestry were identified across 23 chromosomes, with an average proportion of wolf ancestry per block ranging from 0.041 to 0.164 and a global average of 0.066 (Figure 3.4, Table S 3.1). The average block size

was 982Kb in wolves and 1372 Kb in dogs. A total of 848 and 3,228 SNPs located in these regions were identified as CAI loci that are putatively under selection in wolves and dogs, respectively.

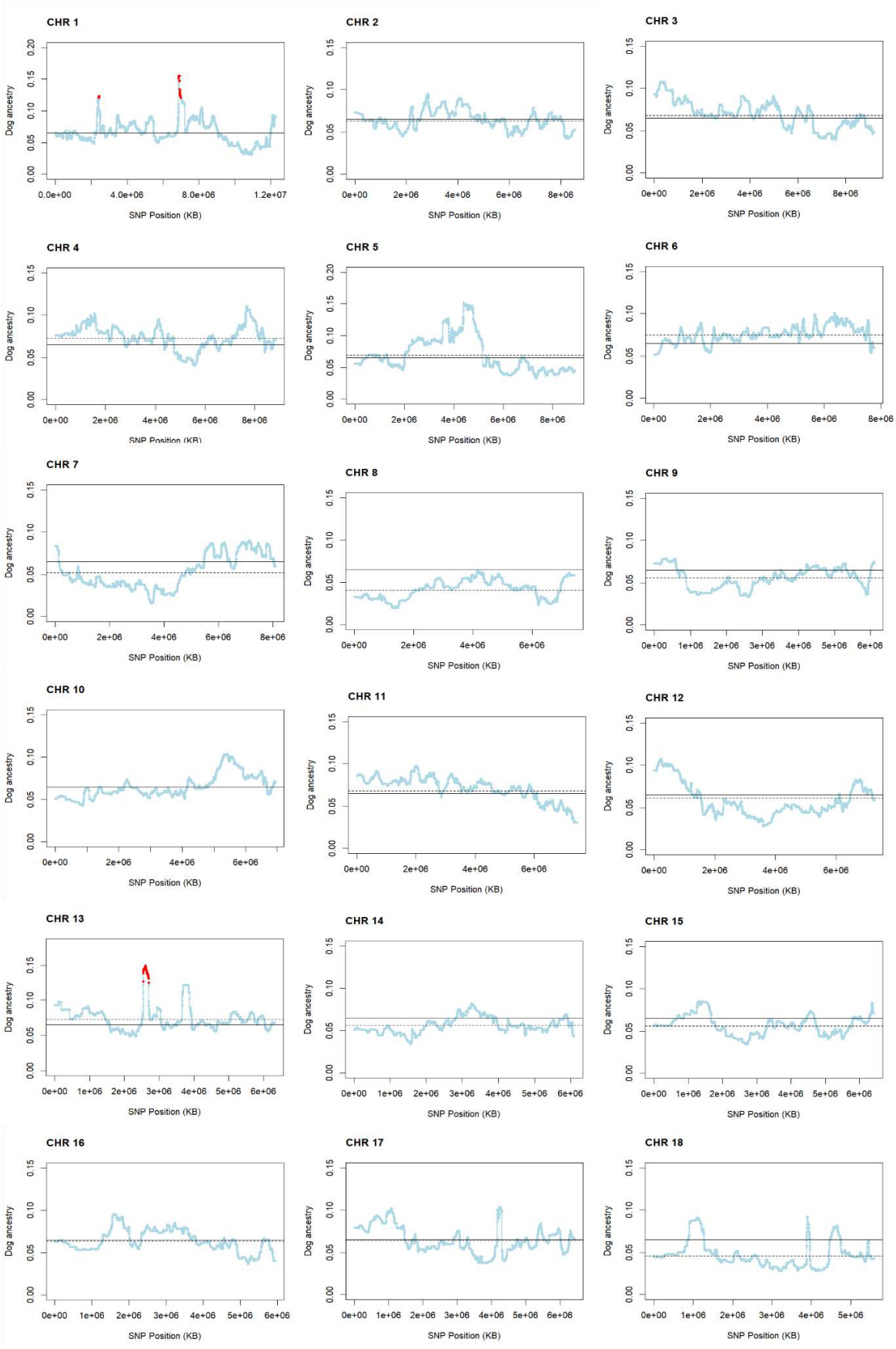
The threshold of less than three SD of the mean, was lower than zero in most of the chromosomes in dog and wolves samples. Therefore, we applied another threshold (less than 0.1% of hybrid ancestry) following Sankararaman et al. (2014). Based on this threshold, 71 underrepresented chromosomal blocks (hybrid ancestry deserts) were found in 28 chromosomes in admixed dogs (Fig 3.4, Table S 3.4), however, no hybrid ancestry desert was detected in admixed wolves.

Jackal-dog dataset

In the dog-jackal dataset, we identified seven blocks with significantly overrepresented dog ancestry on seven jackal chromosomes (Figure 3.5, Table S 3.2). The average proportion of dog ancestry per block ranged from 0.14 to 0.23, with a global average of 0.184 across all blocks. The average block size was 666 Kb in jackals and the total of 453 SNPs were identified as CAI loci. By applying a threshold of 0.1% hybrid ancestry, five hybrid ancestry deserts in three chromosomes were found in admixed jackals (Fig 3.5, Table S 3.5).Since all dog samples displayed more than 99% dog ancestry, chromosomal blocks with an overrepresentation of jackal variants were not identified for dog samples.

Wolf-jackal dataset

In the wolf-jackal dataset, only one chromosomal block with a wolf ancestry proportion of 0.086 was detected (Figure 3.6, Table S 3.3). The average block size was 861 Kb in jackals. In total, 14 SNPs were identified as CAI loci that are putatively under positive selection in jackals. Two chromosomal blacks in two chromosomes in admixed jackals showed signs of underrepresentation of wolf ancestry (Fig 3.6, Table S 3.6). Because we only identified five wolves with jackal admixture, we could not carry out the analysis of the jackal introgression in wolves, but the small number of admixed individuals suggests that such introgression is very limited.


CHR 1
CHR 2
CHR 3
CHR 4
CHR 5
CHR 6
CHR 7
CHR 8
CHR 9
CHR 10
CHR 11
CHR 12
CHR 13
CHR 14
CHR 15
CHR 16
CHR 17
CHR 18
Dog ancestry
SNP Position (KB)

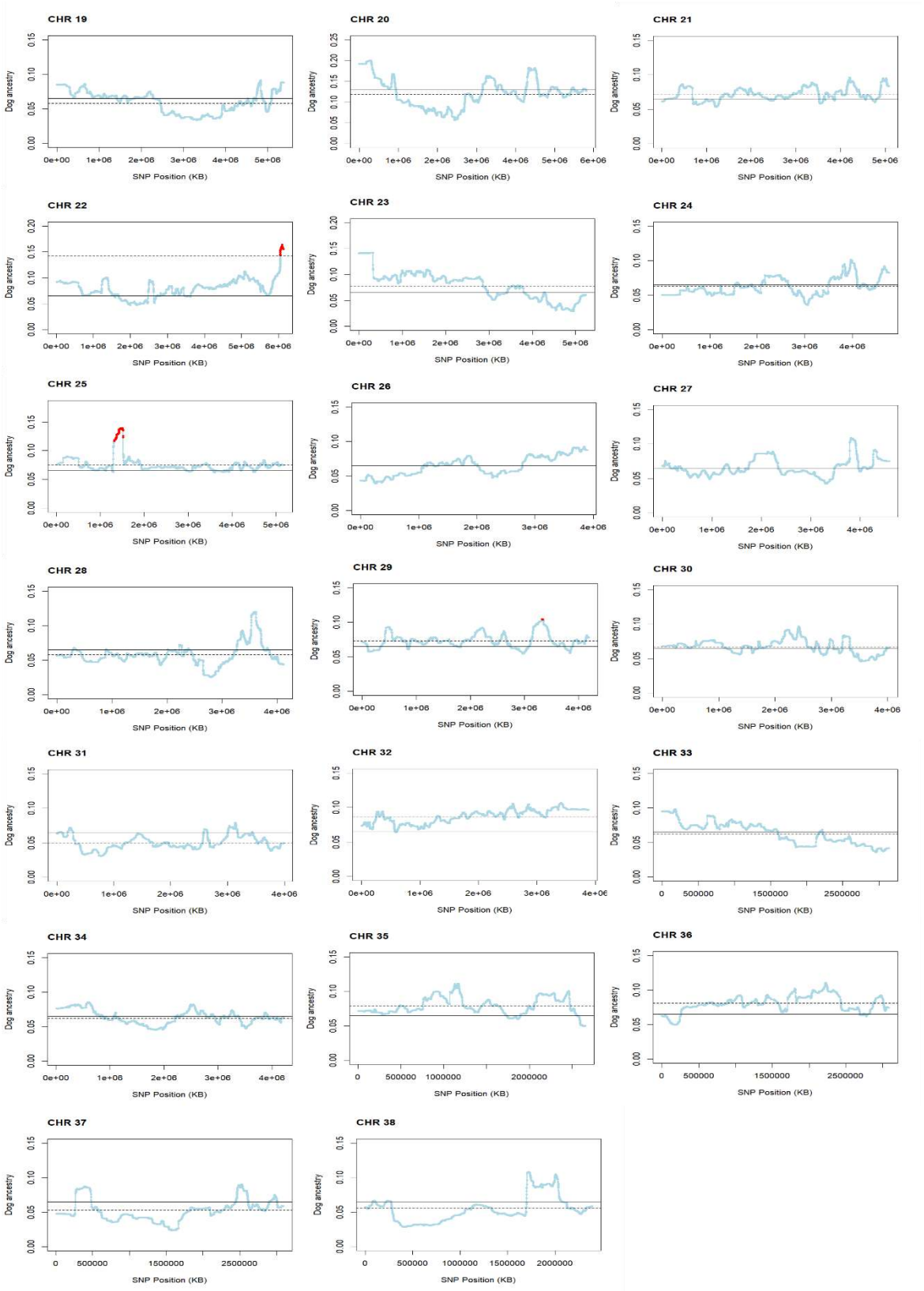
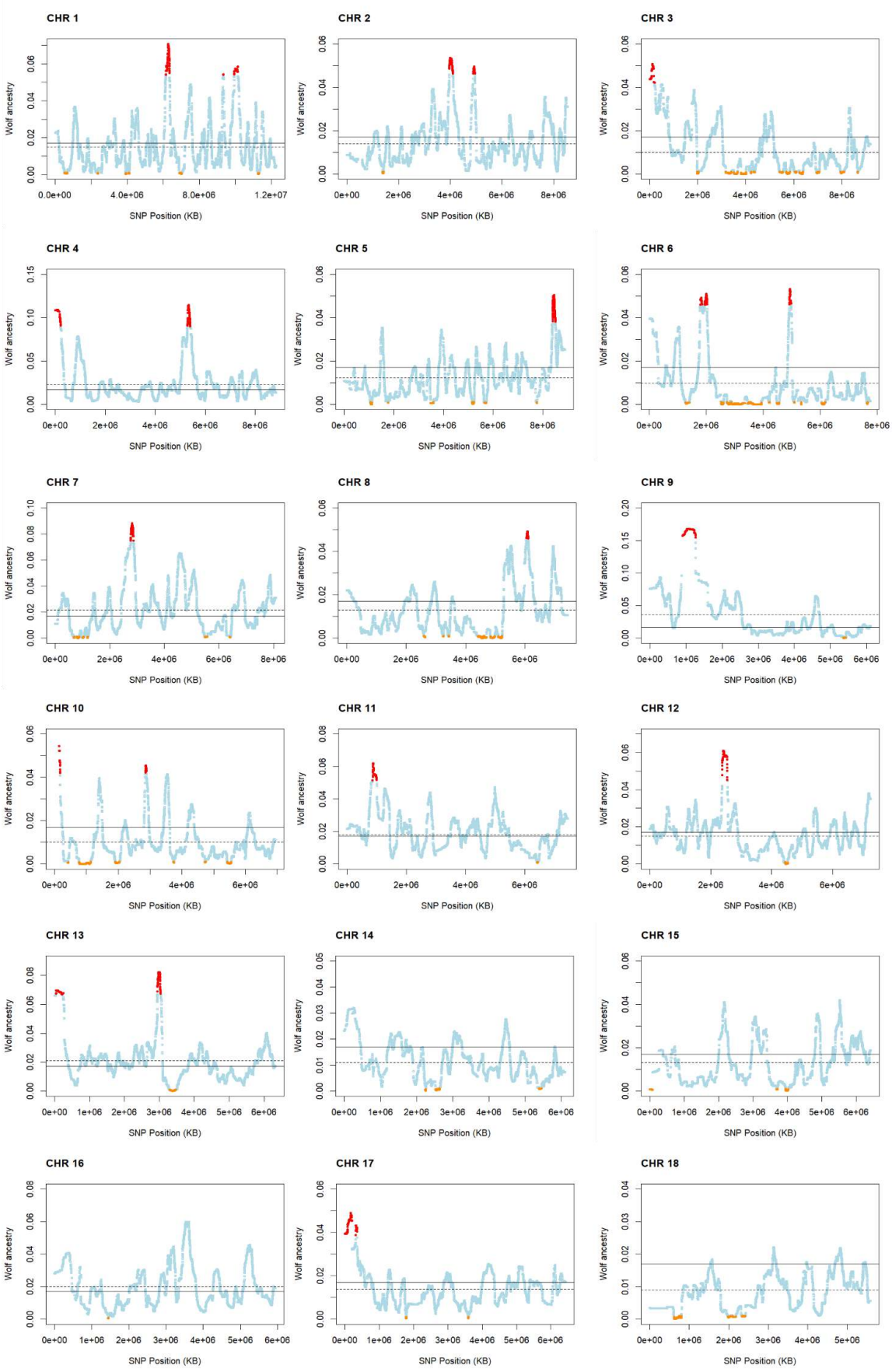


Fig. 3.3. Distribution of dog ancestry in admixed wolves. The x-axis represents SNP order along each autosomal chromosome, and the y-axis represents the proportion of dog admixture in admixed wolf. The solid horizontal line shows the mean dog admixture across autosomal chromosomes, and the dotted horizontal line shows the mean jackal admixture within each chromosome. Chromosomal blocks with overrepresented and underrepresented dog ancestry are marked in red and orange respectively.


CHR 1
CHR 2
CHR 3
CHR 4
CHR 5
CHR 6
CHR 7
CHR 8
CHR 9
CHR 10
CHR 11
CHR 12
CHR 13
CHR 14
CHR 15
CHR 16
CHR 17
CHR 18
Wolf ancestry
SNP Position (KB)

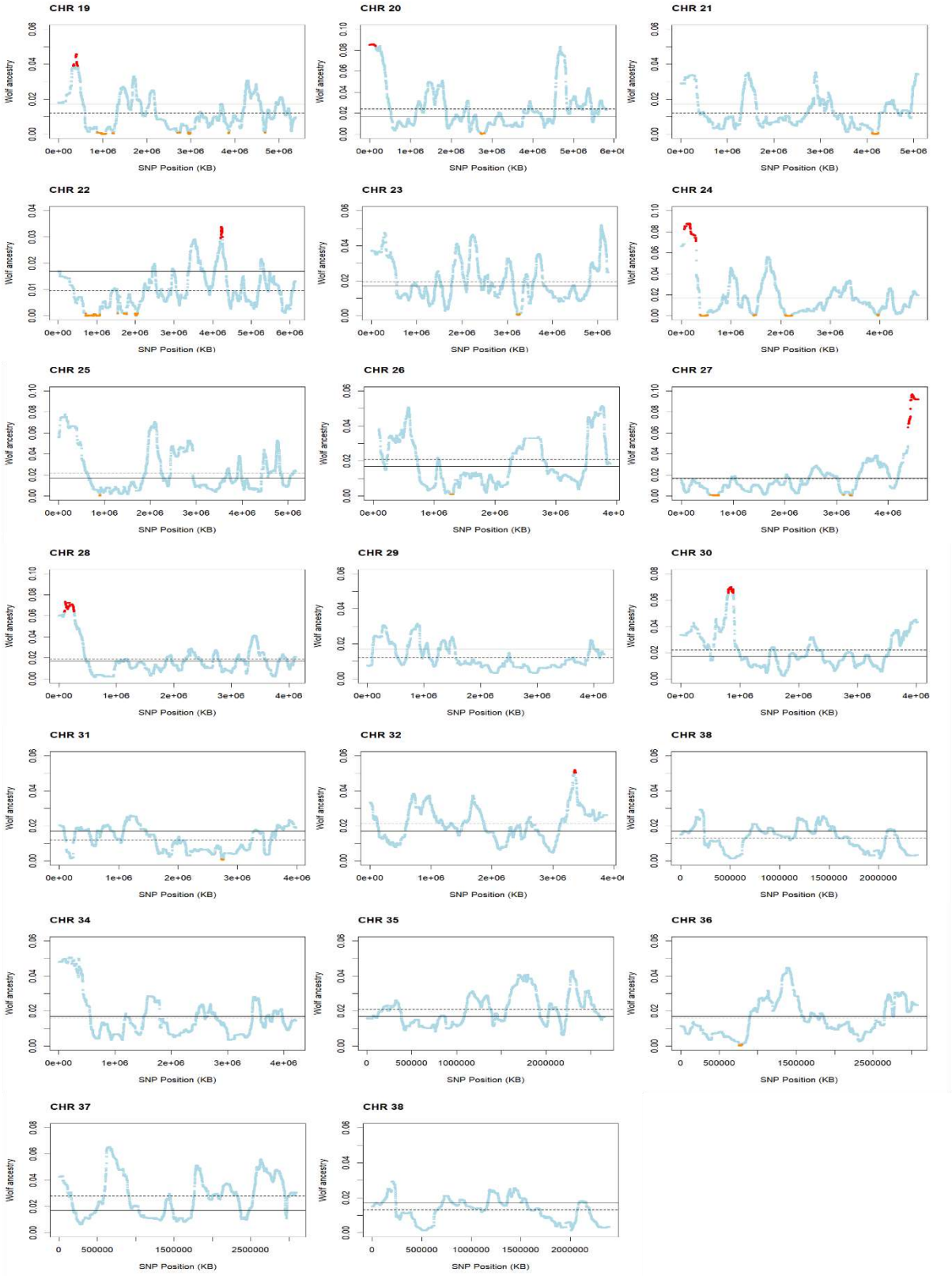
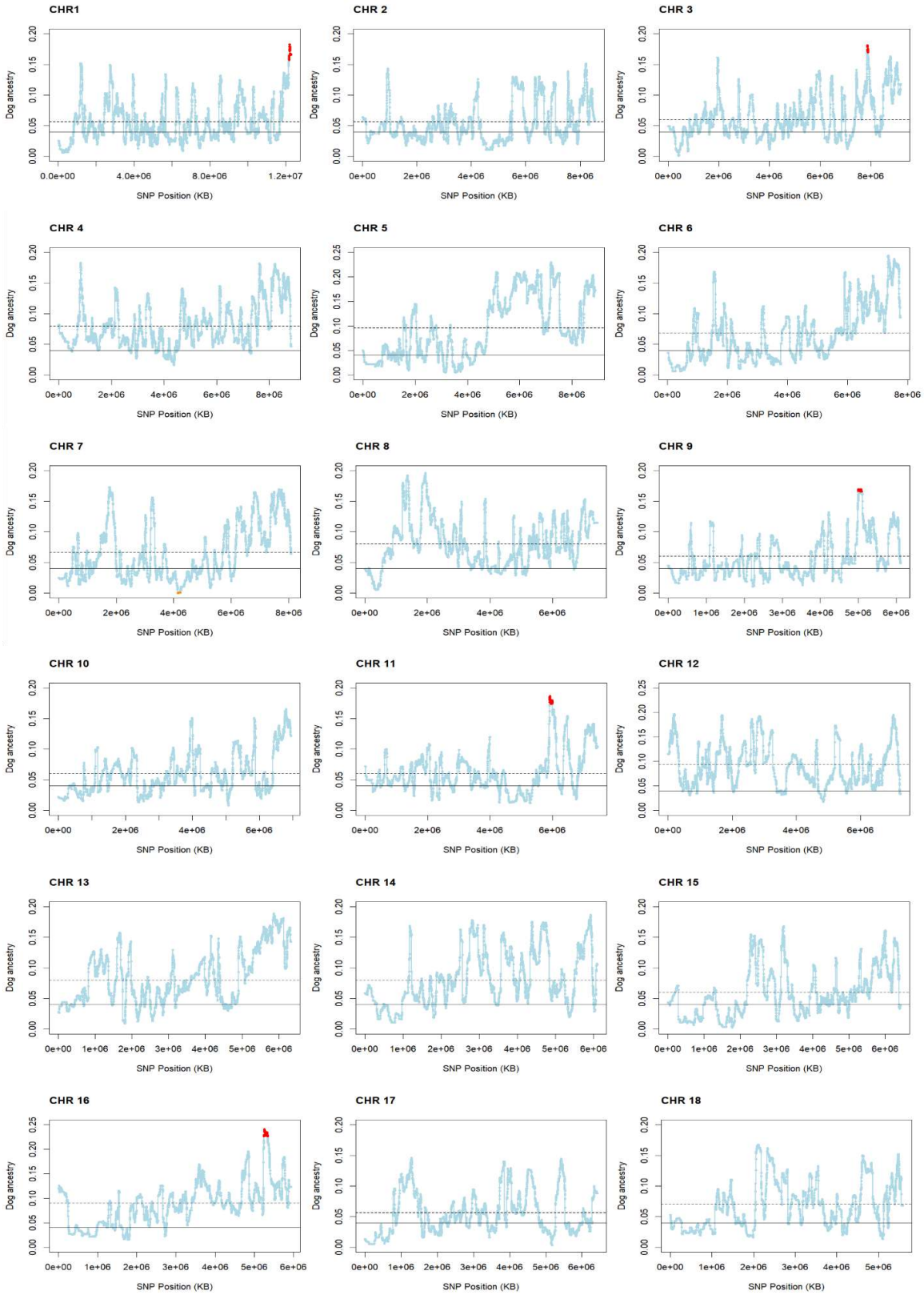


Fig. 3.4. Distribution of wolf ancestry in admixed dogs. The x-axis represents SNP order along each autosomal chromosome, and the y-axis represents the proportion of wolf admixture in admixed dogs. The solid horizontal line shows the mean wolf admixture across autosomal chromosomes, and the dotted horizontal line shows the mean wolf admixture within each chromosome. Chromosomal blocks with overrepresented and underrepresented wolf ancestry are marked in red and orange, respectively


CHR1
CHR 2
CHR 3
CHR 4
CHR 5
CHR 6
CHR 7
CHR 8
CHR 9
CHR 10
CHR 11
CHR 12
CHR 13
CHR 14
CHR 15
CHR 16
CHR 17
CHR 18
Dog ancestry
SNP Position (KB)

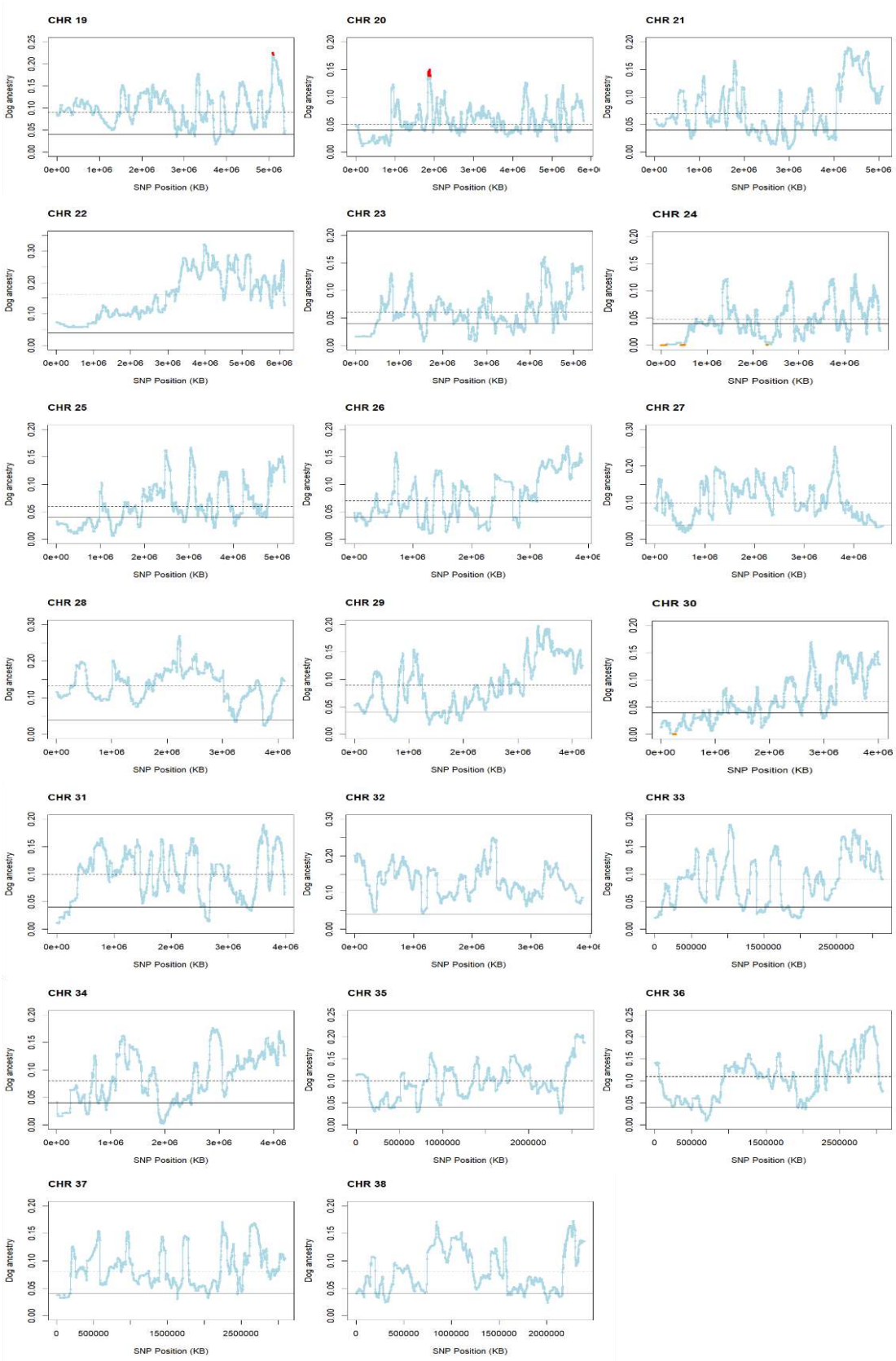
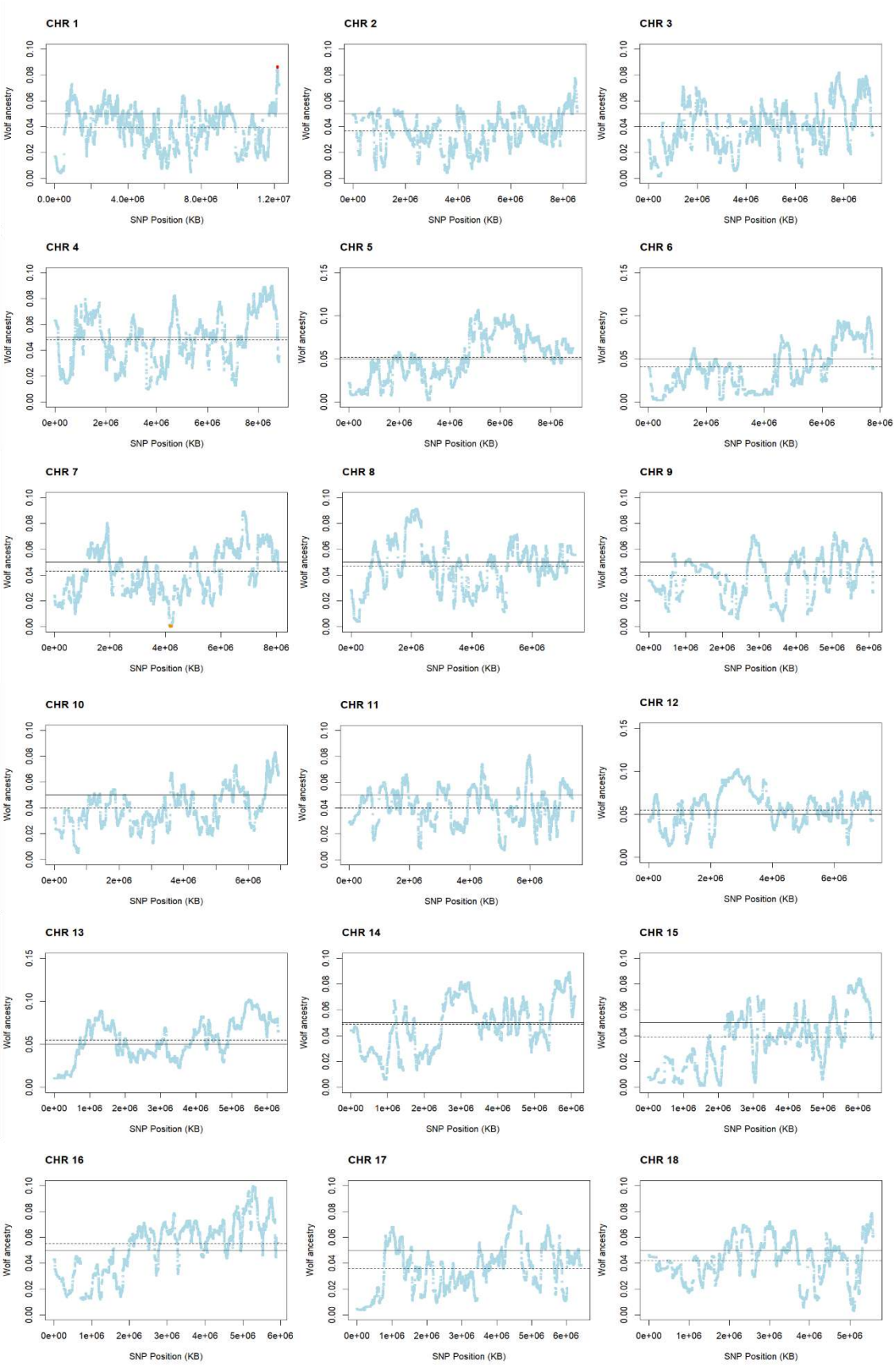


Fig. 3.5. Distribution of dog ancestry in admixed jackals. The x-axis represents SNP order along each autosomal chromosome, and the y-axis represents the proportion of dog admixture in admixed jackals. The solid horizontal line shows the mean dog admixture across autosomal chromosomes, and the dotted horizontal line shows the mean dog admixture within each chromosome. Chromosomal blocks with overrepresented and underrepresented dog ancestry are marked in red and orange respectively.


CHR 1
CHR 2
CHR 3
CHR 4
CHR 5
CHR 6
CHR 7
CHR 8
CHR 9
CHR 10
CHR 11
CHR 12
CHR 13
CHR 14
CHR 15
CHR 16
CHR 17
CHR 18
Wolf ancestry
SNP Position (KB)

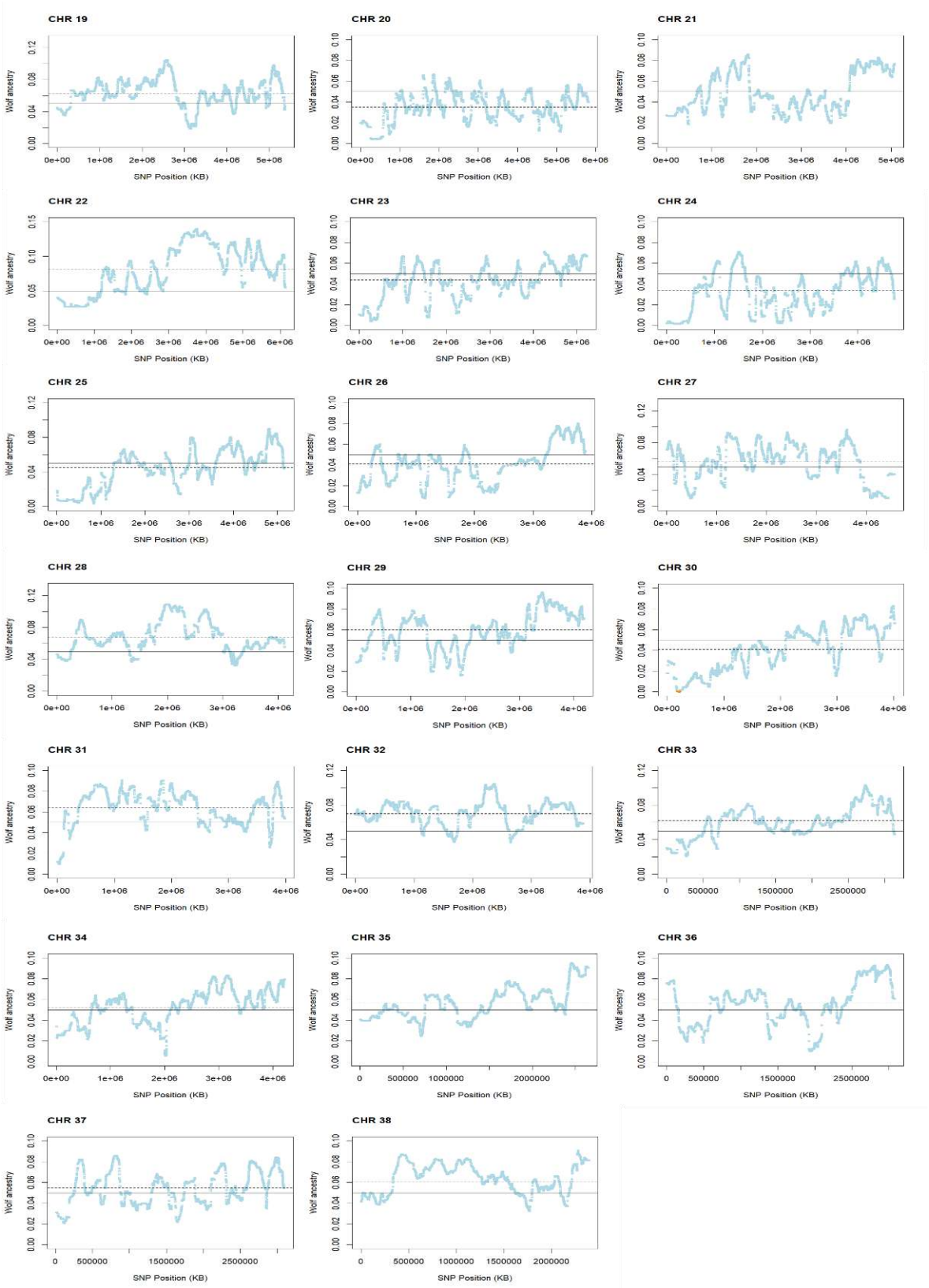


Fig. 3.6. Distribution of wolf ancestry in admixed jackals. The x-axis represents SNP order along each autosomal chromosome, and the y-axis represents the proportion of wolf admixture in admixed jackals. The solid horizontal line shows the mean wolf admixture across autosomal chromosomes, and the dotted horizontal line shows the mean wolf admixture within each chromosome. Chromosomal blocks with overrepresented and underrepresented wolf ancestry are marked in red and orange respectively.

Candidate genes located within introgressed chromosomal blocks

Before identifying SNPs mapped to protein-coding genes, all CAI SNPs were lifted-over to the CanFam6 genome assembly (10K Boxer Tascha). During this process, some SNPs failed to convert (see Table 3.7).

Table 3.7. The total number of CAI loci before and after the lift-over process.

Datasets	The total number of outlier SNPs	Failed SNPs to lift over	The number of SNPs used for the Ensembl analysis
Dog ancestry in wolves	848	2	846
Wolf ancestry in dogs	3,228	53	3,175
Dog ancestry in jackals	452	8	444
Wolf ancestry in jackals	14	0	14

Based on each dataset, the number of CAI loci that are located in intergenic regions, within genes, and within 100kb from protein-coding genes or long noncoding RNA were identified (Table 3.8). The genes associated with SNPs located within genes, and within 100kb of protein-coding genes, were included in the enrichment analysis. Because multiple SNPs can be assigned to the same gene (upstream or downstream), the total number of SNPs mapped to genes exceeds the number of unique genes.

Table 3.8. The number of CAI loci located in different parts of the genome

Dataset	Species	Number of CAI SNPs	Intergenic	Intron	Flanking regions (100kb)	Number of genes
WD	Wolves	846	89 (10.5%)	319 (37.7%)	438 (51.7%)	114
	Dogs	3175	449 (14.4%)	1113 (35.1%)	1613 (50.8%)	595
JD	Jackals	444	60 (13.5%)	152 (34.2%)	232 (52.2%)	94
WJ	Jackals	14	0	0	14 (100%)	4

Functional characterisation of the genes located in the overrepresented regions

We conducted GO enrichment analyses on the candidate genes for adaptive introgression for all datasets based on the canine genes and their human orthologues.

Wolf-dog dataset

The Gene Ontology (GO) analysis for the genes with overrepresented dog ancestry in wolves, based on the canine genome assembly, initially revealed an overrepresentation of molecular function and biological process categories with 32 significant GO terms (Table 3.9). However, when applying a more conservative threshold (g:SCS), only three GO terms (GO:0023023, GO:0042287, and GO:0042289) related to the molecular function "Major Histocompatibility Complex" (MHC) and one GO term (GO:0001909), associated with the biological process "leukocyte mediated cytotoxicity" remained

significant (Table 3.9). These four GO terms were consistently overrepresented, regardless of whether the Benjamini–Hochberg or the stricter g:SCS correction was used. When the analysis was conducted using human homologues, three significant GO terms associated with MHC molecular function were identified (Table 3.9). However, under the more stringent g:SCS threshold, only one GO term, "MHC protein complex binding" (GO:0023023) remained significant. This term was therefore overrepresented in the analyses based on both canine genes and human homologues and with both corrections.

The Gene Ontology (GO) analysis for the genes with overrepresented wolf ancestry in dogs, showed a larger set of enriched GO terms compared to wolves (Table 3.9). Based on the canine genome assembly, 36 GO terms were significantly overrepresented using the Benjamini–Hochberg correction. Applying the more restrictive threshold (g:SCS), 31 GO terms were significantly overrepresented. Four molecular functions and 23 biological processes GO terms were consistently overrepresented, regardless of whether the Benjamini–Hochberg or the stricter g: SCS correction was used. Based on the human homologues, 48 terms associated with molecular function, biological processes, and cellular component were significantly overrepresented (Table 3.9). The molecular function term " N,N-dimethylaniline monooxygenase activity" (GO:0004499) and 23 biological processes related to the “metabolic process” and “biological regulation” were identified as overrepresented in the analyses based on both canine genes and human homologues and with both corrections.

We compared these results with those from the previous study (Pilot et al., 2021). Since these overlapping chromosomal blocks between the two studies probably showed the strongest candidates genes for adaptive introgression, Gene Ontology analysis was carried out also for only genes that were located in the overlapped chromosomal blocks between the present study and the earlier study (Pilot et al., 2021) in wolves and dogs (Table S 3.7). In this study we identified chromosomal blocks of dog ancestry on seven chromosomes in wolves, whereas Pilot et al. (2021) found these blocks on 15 chromosomes, with only three (chromosomes 1, 13, and 28) overlapping between the two studies. In dogs, the current study identified chromosomal blocks from wolf ancestry on 23 chromosomes, while Pilot et al. (2021) identified them on 20 chromosomes with 15 chromosomes being common across both studies (chromosomes 1, 2, 3, 4, 5, 7, 9, 13, 17, 20, 22, 27, 28, 30, and 32).

In dogs based on the canine genome assembly, 24 GO terms were significantly overrepresented using the Benjamini–Hochberg correction. Applying the more restrictive threshold (g:SCS), 12 GO terms were significantly overrepresented. Six molecular functions related to the “monooxygenase activity” and six biological processes GO terms related to the “metabolic process” were consistently overrepresented, regardless of whether the Benjamini–Hochberg or the stricter g: SCS correction was used. In wolves, nine GO terms related to the “MHC” were significantly overrepresented using the Benjamini–Hochberg correction. However by applying the more restrictive threshold (g:SCS), no significant GO terms was found.

Dog-jackal dataset

The Gene Ontology (GO) analysis for the genes with overrepresented dog ancestry in jackals, based on the canine genome assembly, showed an overrepresentation of the cellular component “endoplasmic reticulum protein-containing complex” (GO:0140534; Table 3.10). The analysis based on human homologues revealed an overrepresentation of cellular component “Golgi trans cisterna” and “endoplasmic reticulum protein-containing complex”, however, “Golgi trans cisterna” was significant only when the Benjamini–Hochberg correction, but no g: SCS correction, was applied. Thus, the term “endoplasmic reticulum protein-containing complex” was identified as overrepresented in the analyses based on both canine genes and human homologues and with both corrections.

Wolf-jackal dataset

The Gene Ontology (GO) analysis for the genes with overrepresented wolf ancestry in jackals, based on the canine genome assembly showed significant enrichments of 19 GO terms related to the molecular function, biological process, and cellular component (Table 3.11). However, using the more restrictive correction (g: SCS), only one GO term “cyclin E1-CDK2 complex” (GO:0097134) was significant. The term “cyclin E1-CDK2 complex” was identified as overrepresented in cellular component in the analyses based on both canine genes and human homologues and with both corrections.

Table 3.9. Results of Gene Ontology analysis carried out for two sets of genes in wolves and dogs. The analysis was carried out for canine genes using the dog reference genome and for the human orthologues using the human reference genome. The significance was determined using the Benjamini-Hochberg correction (BH) and the more conservative g:SCS (Set Counts and Sizes) false-discovery rate correction method. Only P-values below 0.05 are shown.

Canids	Reference genome	Go source	Term name	Term ID	P (BH)	P (g:SCS)
Dog	Dog	GO:MF	N,N-dimethylaniline monooxygenase activity	GO:0004499	8.4E-06	4.5E-06
Dog	Dog	GO:MF	androstan-3-alpha,17-beta-diol dehydrogenase activity	GO:0047044	0.0025	0.0041
Dog	Dog	GO:MF	androsterone dehydrogenase activity	GO:0047023	0.0025	0.0041
Dog	Dog	GO:MF	NAD-retinol dehydrogenase activity	GO:0004745	0.0025	0.0069
Dog	Dog	GO:MF	oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen, NAD(P)H as one donor, and incorporation of one atom of oxygen	GO:0016709	0.0025	0.0060
Dog	Dog	GO:MF	steroid dehydrogenase activity, acting on the CH-OH group of donors, NAD or NADP as acceptor	GO:0033764	0.0198	
Dog	Dog	GO:MF	hypotaurine dehydrogenase activity	GO:0047822	0.0198	
Dog	Dog	GO:MF	steroid dehydrogenase activity	GO:0016229	0.0251	
Dog	Dog	GO:MF	NADP binding	GO:0050661	0.0382	
Dog	Dog	GO:MF	flavin adenine dinucleotide binding	GO:0050660	0.038	
Dog	Dog	GO:MF	ion binding	GO:0043167	0.048	
Dog	Dog	GO:BP	regulation of primary metabolic process	GO:0080090	1.9E-07	9.2E-08
Dog	Dog	GO:BP	regulation of nitrogen compound metabolic process	GO:0051171	9.2E-07	9.8E-07
Dog	Dog	GO:BP	regulation of nucleobase-containing compound metabolic process	GO:0019219	9.2E-07	1.3E-06
Dog	Dog	GO:BP	regulation of cellular metabolic process	GO:0031323	2.9E-06	5.6E-06
Dog	Dog	GO:BP	RNA biosynthetic process	GO:0032774	3.4E-06	8.1E-06
Dog	Dog	GO:BP	regulation of metabolic process	GO:0019222	5.2E-06	0.0000
Dog	Dog	GO:BP	regulation of DNA-templated transcription	GO:0006355	5.2E-06	0.0000
Dog	Dog	GO:BP	regulation of RNA biosynthetic process	GO:2001141	5.2E-06	0.0000
Dog	Dog	GO:BP	regulation of RNA metabolic process	GO:0051252	5.2E-06	0.0000
Dog	Dog	GO:BP	cellular nitrogen compound biosynthetic process	GO:0044271	5.2E-06	0.0000
Dog	Dog	GO:BP	nucleobase-containing compound biosynthetic process	GO:0034654	5.9E-06	0.0000
Dog	Dog	GO:BP	DNA-templated transcription	GO:0006351	5.9E-06	0.0000
Dog	Dog	GO:BP	aromatic compound biosynthetic process	GO:0019438	6.3E-06	0.0000
Dog	Dog	GO:BP	heterocycle biosynthetic process	GO:0018130	1.2E-05	0.0000
Dog	Dog	GO:BP	organic cyclic compound biosynthetic process	GO:1901362	1.8E-05	0.0001
Dog	Dog	GO:BP	regulation of macromolecule metabolic process	GO:0060255	6.3E-05	0.0004
Dog	Dog	GO:BP	regulation of gene expression	GO:0010468	0.0005	0.0045
Dog	Dog	GO:BP	regulation of macromolecule biosynthetic process	GO:0010556	0.0006	0.0057
Dog	Dog	GO:BP	regulation of cellular biosynthetic process	GO:0031326	0.0008	0.0074
Dog	Dog	GO:BP	regulation of biosynthetic process	GO:0009889	0.0010	0.0096
Dog	Dog	GO:BP	biological regulation	GO:0065007	0.0010	0.0104

Dog	Dog	GO:BP	regulation of biological process	GO:0050789	0.0021	0.0220
Dog	Dog	GO:BP	regulation of cellular process	GO:0050794	0.0022	0.0239
Dog	Dog	GO:BP	NADPH oxidation	GO:0070995	0.0308	
Dog	Dog	GO:BP	regulation of DNA metabolic process	GO:0051052	0.0420	
Dog	Dog	GO:CC	cytoplasm	GO:0005737		0.0422
Dog	Human	GO:MF	ion binding	GO:0043167	3.9E-06	2.4E-06
Dog	Human	GO:MF	hypotaurine dehydrogenase activity	GO:0047822	4.3E-05	0.0001
Dog	Human	GO:MF	growth hormone receptor binding	GO:0005131	4.3E-05	0.0000
Dog	Human	GO:MF	N,N-dimethylaniline monooxygenase activity	GO:0004499	4.3E-05	0.0001
Dog	Human	GO:MF	cation binding	GO:0043169	0.0001	0.0003
Dog	Human	GO:MF	metal ion binding	GO:0046872	0.0001	0.0004
Dog	Human	GO:MF	hormone activity	GO:0005179	0.0003	0.0015
Dog	Human	GO:MF	RNA polymerase II cis-regulatory region sequence-specific DNA binding	GO:0000978	0.0003	0.0018
Dog	Human	GO:MF	cis-regulatory region sequence-specific DNA binding	GO:0000987	0.0004	0.0027
Dog	Human	GO:MF	RNA polymerase II transcription regulatory region sequence-specific DNA binding	GO:0000977	0.0005	0.0038
Dog	Human	GO:MF	sequence-specific DNA binding	GO:0043565	0.0005	0.0038
Dog	Human	GO:MF	sequence-specific double-stranded DNA binding	GO:1990837	0.0008	0.0064
Dog	Human	GO:MF	transcription regulatory region nucleic acid binding	GO:0001067	0.0010	0.0096
Dog	Human	GO:MF	DNA binding	GO:0003677	0.0010	0.0102
Dog	Human	GO:MF	transcription cis-regulatory region binding	GO:0000976	0.0010	0.0093
Dog	Human	GO:MF	double-stranded DNA binding	GO:0003690	0.0028	0.0286
Dog	Human	GO:MF	hormone receptor binding	GO:0051427	0.0043	0.0463
Dog	Human	GO:MF	androsterone dehydrogenase activity	GO:0047023	0.0070	
Dog	Human	GO:MF	NAD-retinol dehydrogenase activity	GO:0004745	0.0074	
Dog	Human	GO:MF	11-cis-retinol dehydrogenase	GO:0106429	0.0077	
Dog	Human	GO:MF	trimethylamine monooxygenase activity	GO:0034899	0.0077	
Dog	Human	GO:MF	androstan-3-alpha,17-beta-diol dehydrogenase activity	GO:0047044	0.0085	
Dog	Human	GO:MF	receptor ligand activity	GO:0048018	0.0207	
Dog	Human	GO:MF	signaling receptor regulator activity	GO:0030545	0.0207	
Dog	Human	GO:MF	signaling receptor activator activity	GO:0030546	0.0237	
Dog	Human	GO:MF	cytokine receptor binding	GO:0005126	0.0250	
Dog	Human	GO:MF	protein binding	GO:0005515	0.0281	
Dog	Human	GO:MF	tRNA (cytosine-3-)-methyltransferase activity	GO:0052735	0.0339	
Dog	Human	GO:MF	steroid dehydrogenase activity, acting on the CH-OH group of donors, NAD or NADP as acceptor	GO:0033764	0.0352	
Dog	Human	GO:BP	regulation of primary metabolic process	GO:0080090	4.6E-06	2.2E-06
Dog	Human	GO:BP	regulation of nitrogen compound metabolic process	GO:0051171	1.0E-05	0.0000
Dog	Human	GO:BP	regulation of cellular metabolic process	GO:0031323	0.0005	0.0009
Dog	Human	GO:BP	positive regulation of cellular metabolic process	GO:0031325	0.0005	0.0012
Dog	Human	GO:BP	positive regulation of receptor signaling pathway via STAT	GO:1904894	0.0005	0.0010
Dog	Human	GO:BP	growth hormone receptor signaling pathway	GO:0060396	0.0005	0.0015

Dog	Human	GO:BP	cellular response to growth hormone stimulus	GO:0071378	0.0005	0.0020
Dog	Human	GO:BP	biological regulation	GO:0065007	0.0005	0.0021
Dog	Human	GO:BP	regulation of metabolic process	GO:0019222	0.0005	0.0024
Dog	Human	GO:BP	positive regulation of macromolecule metabolic process	GO:0010604	0.0005	0.0030
Dog	Human	GO:BP	cellular nitrogen compound biosynthetic process	GO:0044271	0.0005	0.0028
Dog	Human	GO:BP	regulation of nucleobase-containing compound metabolic process	GO:0019219	0.0006	0.0039
Dog	Human	GO:BP	positive regulation of metabolic process	GO:0009893	0.0006	0.0055
Dog	Human	GO:BP	regulation of RNA biosynthetic process	GO:2001141	0.0006	0.0045
Dog	Human	GO:BP	positive regulation of nitrogen compound metabolic process	GO:0051173	0.0007	0.0055
Dog	Human	GO:BP	RNA biosynthetic process	GO:0032774	0.0007	0.0059
Dog	Human	GO:BP	regulation of DNA-templated transcription	GO:0006355	0.0008	0.0068
Dog	Human	GO:BP	positive regulation of receptor signaling pathway via JAK-STAT	GO:0046427	0.0010	0.0090
Dog	Human	GO:BP	positive regulation of cellular process	GO:0048522	0.0013	0.0123
Dog	Human	GO:BP	DNA-templated transcription	GO:0006351	0.0014	0.0143
Dog	Human	GO:BP	aromatic compound biosynthetic process	GO:0019438	0.0014	0.0154
Dog	Human	GO:BP	regulation of RNA metabolic process	GO:0051252	0.0014	0.0161
Dog	Human	GO:BP	organic cyclic compound biosynthetic process	GO:1901362	0.0017	0.0193
Dog	Human	GO:BP	regulation of cellular process	GO:0050794	0.0018	0.0222
Dog	Human	GO:BP	nucleobase-containing compound biosynthetic process	GO:0034654	0.0018	0.0223
Dog	Human	GO:BP	response to growth hormone	GO:0060416	0.0018	0.0234
Dog	Human	GO:BP	regulation of macromolecule metabolic process	GO:0060255	0.0018	0.0247
Dog	Human	GO:BP	heterocycle biosynthetic process	GO:0018130	0.0018	0.0252
Dog	Human	GO:BP	regulation of biological process	GO:0050789	0.0021	0.0309
Dog	Human	GO:BP	positive regulation of chromosome organization	GO:2001252	0.0036	
Dog	Human	GO:BP	regulation of receptor signaling pathway via STAT	GO:1904892	0.0038	
Dog	Human	GO:BP	positive regulation of tyrosine phosphorylation of STAT protein	GO:0042531	0.0039	
Dog	Human	GO:BP	positive regulation of biological process	GO:0048518	0.0076	
Dog	Human	GO:BP	regulation of tyrosine phosphorylation of STAT protein	GO:0042509	0.0114	
Dog	Human	GO:BP	positive regulation of organelle organization	GO:0010638	0.0127	
Dog	Human	GO:BP	tyrosine phosphorylation of STAT protein	GO:0007260	0.0137	
Dog	Human	GO:BP	regulation of receptor signaling pathway via JAK-STAT	GO:0046425	0.0138	
Dog	Human	GO:BP	receptor signaling pathway via JAK-STAT	GO:0007259	0.0160	
Dog	Human	GO:BP	receptor signaling pathway via STAT	GO:0097696	0.0217	
Dog	Human	GO:BP	cellular homeostasis	GO:0019725	0.0226	
Dog	Human	GO:BP	vesicle fusion	GO:0006906	0.0246	
Dog	Human	GO:BP	positive regulation of RNA biosynthetic process	GO:1902680	0.0248	
Dog	Human	GO:BP	organelle membrane fusion	GO:0090174	0.0264	

Dog	Human	GO:BP	positive regulation of RNA metabolic process	GO:0051254	0.0332	
Dog	Human	GO:BP	positive regulation of DNA metabolic process	GO:0051054	0.0345	
Dog	Human	GO:BP	positive regulation of macromolecule biosynthetic process	GO:0010557	0.0345	
Dog	Human	GO:BP	positive regulation of telomere maintenance	GO:0032206	0.0387	
Dog	Human	GO:BP	positive regulation of DNA-templated transcription	GO:0045893	0.0389	
Dog	Human	GO:BP	positive regulation of nucleobase-containing compound metabolic process	GO:0045935	0.0393	
Dog	Human	GO:BP	regulation of DNA metabolic process	GO:0051052	0.0448	
Dog	Human	GO:BP	primary metabolic process	GO:0044238	0.0472	
Dog	Human	GO:CC	intracellular membrane-bounded organelle	GO:0043231	0.0142	0.0064
Dog	Human	GO:CC	membrane-bounded organelle	GO:0043227	0.0236	0.0216
Dog	Human	GO:CC	intracellular organelle	GO:0043229	0.0444	
Dog	Human	GO:CC	ATPase complex	GO:1904949	0.0444	
Wolf	Dog	GO:MF	MHC protein complex binding	GO:0023023	0.0001	0.0002
Wolf	Dog	GO:MF	MHC protein binding	GO:0042287	0.0008	0.0025
Wolf	Dog	GO:MF	MHC class II protein binding	GO:0042289	0.0045	0.0196
Wolf	Dog	GO:MF	rRNA (pseudouridine) methyltransferase activity	GO:0070037	0.0471	
Wolf	Dog	GO:MF	corticotropin receptor activity	GO:0004978	0.0471	
Wolf	Dog	GO:MF	interleukin-16 binding	GO:0042011	0.0471	
Wolf	Dog	GO:MF	interleukin-16 receptor activity	GO:0042012	0.0471	
Wolf	Dog	GO:MF	peroxisome matrix targeting signal-1 binding	GO:0005052	0.0471	
Wolf	Dog	GO:MF	signal sequence binding	GO:0005048	0.0471	
Wolf	Dog	GO:MF	1-acylglycerophosphoethanolamine O-acyltransferase activity	GO:0106262	0.0471	
Wolf	Dog	GO:MF	triose-phosphate isomerase activity	GO:0004807	0.0471	
Wolf	Dog	GO:MF	ATPase-coupled intramembrane lipid transporter activity	GO:0140326	0.0471	
Wolf	Dog	GO:BP	positive regulation of cell killing	GO:0031343	0.0204	
Wolf	Dog	GO:BP	regulation of lymphocyte mediated immunity	GO:0002706	0.0204	
Wolf	Dog	GO:BP	regulation of leukocyte mediated cytotoxicity	GO:0001910	0.0204	
Wolf	Dog	GO:BP	positive regulation of immune effector process	GO:0002699	0.0204	
Wolf	Dog	GO:BP	natural killer cell mediated immunity	GO:0002228	0.0204	
Wolf	Dog	GO:BP	regulation of lymphocyte activation	GO:0051249	0.0204	
Wolf	Dog	GO:BP	natural killer cell mediated cytotoxicity	GO:0042267	0.0204	
Wolf	Dog	GO:BP	leukocyte mediated cytotoxicity	GO:0001909	0.0204	0.0376
Wolf	Dog	GO:BP	cell killing	GO:0001906	0.0204	
Wolf	Dog	GO:BP	positive regulation of leukocyte mediated cytotoxicity	GO:0001912	0.0204	
Wolf	Dog	GO:BP	regulation of immune effector process	GO:0002697	0.0256	
Wolf	Dog	GO:BP	regulation of cell killing	GO:0031341	0.0252	
Wolf	Dog	GO:BP	lymphocyte activation	GO:0046649	0.0252	
Wolf	Dog	GO:BP	regulation of T cell activation	GO:0050863	0.0267	
Wolf	Dog	GO:BP	regulation of leukocyte activation	GO:0002694	0.0267	

Wolf	Dog	GO:BP	regulation of leukocyte mediated immunity	GO:0002703	0.0324	
Wolf	Dog	GO:BP	regulation of cell activation	GO:0050865	0.0337	
Wolf	Dog	GO:BP	regulation of immune system process	GO:0002682	0.0387	
Wolf	Dog	GO:BP	lymphocyte mediated immunity	GO:0002449	0.0450	
Wolf	Dog	GO:BP	leukocyte activation	GO:0045321	0.0466	
Wolf	Human	GO:MF	MHC protein complex binding	GO:0023023	0.0154	0.0389
Wolf	Human	GO:MF	MHC protein binding	GO:0042287	0.0154	
Wolf	Human	GO:MF	MHC class II protein binding	GO:0042289	0.0154	

Table 3.10. Results of Gene Ontology analysis carried out for a set of genes in jackals (jackal-dog dataset). The analysis was carried out for canine genes using the dog reference genome and for the human orthologues using the human reference genome. The significance was determined using the Benjamini-Hochberg correction (BH) and the more conservative g:SCS (Set Counts and Sizes) false-discovery rate correction method. Only P-values below 0.05 are shown.

Canids	Reference genome	Go source	Term name	Term ID	P (FDR)	P (g:SCS)
Jackal	Dog	GO:CC	endoplasmic reticulum protein-containing complex	GO:0140534	0.0176	0.0189
Jackal	Human	GO:CC	Golgi trans cisterna	GO:0000138	0.0298	
Jackal	Human	GO:CC	endoplasmic reticulum protein-containing complex	GO:0140534	0.0298	0.0483

Table 3.11. Results of Gene Ontology analysis carried out for two sets of genes in jackals. The analysis was carried out for canine genes using the dog reference genome and for the human orthologues using the human reference genome. The significance was determined using the Benjamini-Hochberg correction (BH) and the more conservative g:SCS (Set Counts and Sizes) false-discovery rate correction method. Only P-values below 0.05 are shown.

Canids	Reference genome	Go source	Term name	Term ID	P (FDR)	P (g:SCS)
Jackal	Dog	GO:MF	RNA polymerase II complex binding	GO:0000993	0.0219	
Jackal	Dog	GO:MF	basal transcription machinery binding	GO:0001098	0.0219	
Jackal	Dog	GO:MF	basal RNA polymerase II transcription machinery binding	GO:0001099	0.0219	
Jackal	Dog	GO:MF	RNA polymerase binding	GO:0070063	0.0219	
Jackal	Dog	GO:MF	cyclin-dependent protein serine/threonine kinase regulator activity	GO:0016538	0.0219	
Jackal	Dog	GO:MF	RNA polymerase core enzyme binding	GO:0043175	0.0219	
Jackal	Dog	GO:MF	enzyme binding	GO:0019899	0.0238	
Jackal	Dog	GO:MF	phosphoprotein binding	GO:0051219	0.0277	
Jackal	Dog	GO:MF	transcription corepressor activity	GO:0003714	0.0388	
Jackal	Dog	GO:MF	protein kinase regulator activity	GO:0019887	0.0455	
Jackal	Dog	GO:MF	kinase regulator activity	GO:0019207	0.0467	
Jackal	Dog	GO:BP	mesenchymal stem cell proliferation	GO:0097168	0.0460	
Jackal	Dog	GO:BP	positive regulation of mesenchymal stem cell proliferation	GO:1902462	0.0460	
Jackal	Dog	GO:BP	regulation of mesenchymal stem cell proliferation	GO:1902460	0.0460	
Jackal	Dog	GO:CC	cyclin E1-CDK2 complex	GO:0097134	0.0057	0.0499
Jackal	Dog	GO:CC	RPAP3/R2TP/prefoldin-like complex	GO:1990062	0.0144	
Jackal	Dog	GO:CC	cyclin-dependent protein kinase holoenzyme complex	GO:0000307	0.0279	
Jackal	Dog	GO:CC	protein kinase complex	GO:1902911	0.0467	
Jackal	Dog	GO:CC	serine/threonine protein kinase complex	GO:1902554	0.0467	
Jackal	Human	GO:MF	RNA polymerase II complex binding	GO:0000993	0.0211	
Jackal	Human	GO:MF	phosphatase inhibitor activity	GO:0019212	0.0211	
Jackal	Human	GO:MF	basal transcription machinery binding	GO:0001098	0.0211	
Jackal	Human	GO:MF	basal RNA polymerase II transcription machinery binding	GO:0001099	0.0211	
Jackal	Human	GO:MF	RNA polymerase binding	GO:0070063	0.0211	
Jackal	Human	GO:MF	protein phosphatase inhibitor activity	GO:0004864	0.0211	
Jackal	Human	GO:MF	RNA polymerase core enzyme binding	GO:0043175	0.0211	
Jackal	Human	GO:MF	enzyme regulator activity	GO:0030234	0.0211	
Jackal	Human	GO:MF	cyclin-dependent protein serine/threonine kinase regulator activity	GO:0016538	0.0211	
Jackal	Human	GO:MF	protein phosphatase regulator activity	GO:0019888	0.0254	
Jackal	Human	GO:MF	phosphatase regulator activity	GO:0019208	0.0254	
Jackal	Human	GO:MF	enzyme binding	GO:0019899	0.0254	
Jackal	Human	GO:MF	phosphoprotein binding	GO:0051219	0.0254	
Jackal	Human	GO:MF	molecular function regulator activity	GO:0098772	0.0254	
Jackal	Human	GO:MF	transcription corepressor activity	GO:0003714	0.0406	
Jackal	Human	GO:MF	protein kinase regulator activity	GO:0019887	0.0452	
Jackal	Human	GO:MF	kinase regulator activity	GO:0019207	0.0497	
Jackal	Human	GO:BP	positive regulation of mesenchymal stem cell proliferation	GO:1902462	0.0466	
Jackal	Human	GO:BP	regulation of mesenchymal stem cell proliferation	GO:1902460	0.0466	
Jackal	Human	GO:BP	mesenchymal stem cell proliferation	GO:0097168	0.0466	

Jackal	Human	GO:CC	cyclin E1-CDK2 complex	GO:0097134	0.0105	0.0499
Jackal	Human	GO:CC	RPAP3/R2TP/prefoldin-like complex	GO:1990062	0.0211	
Jackal	Human	GO:CC	protein folding chaperone complex	GO:0101031	0.0482	
Jackal	Human	GO:CC	cyclin-dependent protein kinase holoenzyme complex	GO:0000307	0.0484	

Candidate genes under positive selection within introgressed chromosomal blocks

If one or more SNPs showing signatures of positive selection in the iHS test were located within a protein-coding gene, the gene was considered a candidate for positive selection. In the wolf-dog dataset, we found 12 SNPs showing signature of selection within overrepresented introgressed chromosomal blocks of dog ancestry in wolves. Based on the genome annotation, four SNPs (33.4%), were placed within protein-coding genes, seven SNPs (58.3%) were located within 100 kb from protein-coding genes or long noncoding RNA, and only one SNP (8.3%) was placed outside the 100 kb from any annotated gene. In dogs, 45 SNPs showed signature of selection within overrepresented introgressed chromosomal blocks of wolf ancestry. 18 SNPs (40%) were located within protein-coding genes, 16 SNPs (35.5%) were placed within 100 kb of protein-coding genes and the remaining 11 SNPs (24.5%) were placed outside of 100 kb window. The false discovery rate was estimated at 8.3% in wolves and 24.5% in dogs, associated with detecting selection signatures. In a previous study (Pilot et al., 2021), the false discovery rate ranged from 10% to 12% in wolves and dogs. The relatively high error rate may be due to the less complete functional annotation of the dog genome compared to the human genome, as well as the unknown function of some constrained elements in mammalian genomes (Lindblad-Toh et al., 2011).

Among the 114 CAI genes that were placed in the overrepresented chromosomal blocks in admixed wolves, 14 were identified under positive selection based on the iHS results (Table 3.12). Four positively selected SNPs in wolves were placed within genes (TFDP1, TMEM255B, RASA3, and ATP8A2), and the others were located upstream or downstream of genes (Table 3.12). ATP8A2 gene is associated with diseases of the mammalian nervous system (Xu et al., 2012). RASA3 and TFDP1 are involved in fundamental cellular processes. RASA3 is a GTPase-activating protein (GAP) which is important for cell growth (Johansen et al., 2022), and TFDP1 plays important roles in cell proliferation (Nakajima et al., 2023). In dogs, of 595 candidate adaptive genes that were placed in the overrepresented chromosomal blocks, 58 genes were identified as genes under positive selection based on the iHS results. These genes were located in chromosomes 1, 13, 22, 25, 27, 28, and 29 (Table 3.12 and 3.13). 18 positively selected SNPs were located within x genes (NKAIN2, TG, SLA, SH3RF2, TCERG1, CACNA1C, HAVCR2, and HAVCR1). CACNA1C, TCERG1 and NKAIN2 genes (Tables 3.12 and 3.13) were associated with brain function, nervous system, immune activation, behaviour, and cognition (Gorokhova et al., 2007; Moon et al., 2018; Liu et al., 2022).

The Gene Ontology (GO) analysis for the genes under positive selection in dogs, based on the canine genome assembly, showed an overrepresentation of the biological processes “adaptive immune response” and “response to stress” (Table S 3.8). However, using the more restrictive correction (g: SCS), only two GO terms “immunological synapse” and “early endosome” related to the cellular component were significant. In

wolves, the Gene Ontology (GO) analysis for the genes under positive selection based on the canine genome assembly showed an overrepresentation of three molecular function terms related to “cellular homeostasis” and “signal regulation”, and nine cellular components related to “membrane-associated processes” (Table S 3.8). Using the more restrictive correction (g: SCS), only one GO terms, “phagolysosome membrane” was significant.

None of the outlier SNPs in the JD and WJ datasets were identified as significant in the iHS test for positive selection. These results suggests that introgression between these species is mostly neutral or disadvantageous.

Table 3.12. SNP loci located within chromosomal blocks with significant overrepresentation of introgressed ancestry in wolves and dogs show significant results in the iHS test for positive selection at P-value<0.05. Only SNPs located within protein-coding genes are listed. The table lists the position of each SNP on a chromosome, the gene within which the SNP is located, the value of the iHS statistic, the corresponding P-value, and the threshold.

Population	CHR	SNP	Gene (Intron)	Gene (Up/downstream)	iHS	P-value
Dog	1	63147747	NKAIN2		-3.010	0.0026
Dog	1	63155179	NKAIN2		-2.793	0.005
Dog	1	63157377	NKAIN2		-2.509	0.012
Dog	1	63273200	NKAIN2		2.111	0.034
Dog	1	63331062	NKAIN2		2.965	0.003
Dog	1	63395502	NKAIN2		2.390	0.01
Dog	2	40315410	-	SH3RF2, GRXCR2, PRELID2	-2.485	0.012
Dog	2	40402409	SH3RF2		-2.303	0.021
Dog	2	40889976	TCERG1		-2.609	0.009
Dog	2	40904689	TCERG1		-4.456	0.000
Dog	2	40921076	-	PPP2R2B, TCERG1, U6	-2.182	0.029
Dog	4	53051753	HAVCR1		-2.481	0.013
Dog	4	53056922	HAVCR2		2.257	0.0239
Dog	4	53064234	HAVCR1		-2.194	0.028
Dog	4	53067374	HAVCR1		-2.324	0.020
Dog	6	18253510		CORO1A, CLN3, SLX1A, SULT1A1, NUPR1, ATXN2L, APOBR, SGF29	2.296	0.021
Dog	6	18398476	-	ATXN2L, TUFM, SH2B1, ATP2A1, RABEP2, CD19, NFATC2IP, SPNS1, LAT, SBK1	2.063	0.039
Dog	6	49251785	-	EXTL2, VCAM1, SLC30A7	-4.807	0.000
Dog	6	49299085	-	EXTL2, VCAM1, SLC30A7	3.559	0.000
Dog	7	28221952	-	GORAB, PRRX1	-2.424	0.015
Dog	7	28441296	-	NTMT2	2.496	0.012
Dog	8	60876786	-	EFCAB11	2.746	0.006
Dog	9	10976881	-	TANC2, MARCHF10	3.440	0.000
Dog	10	28622800	-	APOL6, MP, RASD2, MCM5	2.779	0.005
Dog	12	24477577	-	RAB23	2.074	0.038
Dog	12	24813511	-	PRIM2	2.090	0.036
Dog	13	29546745	SLA		3.448	0.001
Dog	13	29581239	TG		-2.926	0.003
Dog	13	29679951	-	NDRG1, CCN4, TG	2.784	0.005
Dog	13	29682543	-	NDRG1, CCN4, TG	-4.578	0.000
Dog	27	44540416	CACNA1C		2.058	0.040
Dog	28	1792263	ENSCAFG00000029477		2.135	0.033
Dog	28	1795829	ENSCAFG00000029477		2.135	0.033
Dog	28	2265516	-	ALOX5, OR6D6, OR13A1, OR6D7	-2.048	0.040
Wolf	22	60725183	-	ADPRHL1, LAMP1, CUL4A, GRTP1, DCUN1D2, TMCO3	-2.329	0.020
Wolf	22	60895986	TFDP1		2.211	0.027
Wolf	22	60973456	TMEM255B		2.300	0.021
Wolf	22	61163629	RASA3		-2.159	0.031
Wolf	25	13633588	ATP8A2		-2.119	0.034
Wolf	25	14390299	-	ENSCAFG00000057729	-2.367	0.018
Wolf	25	14417692	-	ENSCAFG00000007106	-2.256	0.024

Wolf	25	14429605	-	ENSCAFG00000007106	-2.972	0.003
Wolf	25	14431300	-	ENSCAFG00000007106	-2.975	0.003
Wolf	28	35748262	-	C28H10orf90	-2.572	0.010

Table 3.13. The genes under positive selection, based on the iHS results in wolves and dogs, which are placed within protein-coding genes

Population	CHR	Gene	N SNPs	Functions
Wolf	22	TFDP1	1	DNA binding
Wolf	22	TMEM255B	1	membrane
Wolf	22	RASA3	1	calcium-release channel activity
Wolf	22	ATP8A2	1	nucleotide-binding, ATP binding
Dog	1	NKAIN2	6	regulation of sodium ion transport
Dog	2	SH3RF2	1	protein phosphatase 1, binding, transferase activity, metal ion binding
Dog	2	TCERG1	2	transcription coactivator activity, transcription corepressor activity, protein binding, mRNA processing
Dog	4	HAVCR1	3	phagocytosis, engulfment, positive regulation of mast cell activation
Dog	4	HAVCR2	2	cellular response to lipopolysaccharide, macrophage activation involved in immune response
Dog	13	SLA	1	regulation of MAPK cascade, signal transduction
Dog	13	TG	2	hormone biosynthetic process, regulation of myelination
Dog	27	CACNA1C	1	calcium channel activity, metal ion binding

3.4. Discussion

Hybridization between gray wolf, golden jackal, and domestic dog

Our results revealed that the occurrence of hybridization between wolves, jackals and dogs can be observed throughout their distribution ranges. We did not find any specific geographic region with a notably high frequency of hybridization, suggesting that hybridization is not confined to specific regions. Nonetheless, in some areas such as India, the Caucasus, Eastern Europe (Ukraine and Belarus), and the Balkans the occurrence of hybridization is more frequent.

In consistence with our findings, the presence of hybridization between canid species has been documented in these regions, e.g. Caucasus (Khosravi et al., 2013; Kopaliani et al., 2014; Asadi Aghbolaghi et al., 2014), India (Tyagi et al., 2023), Ukraine (Gursky, 1975; Dumenko, 2001; Stronen et al., 2013), Belarus (Molchan et al., 2023), and the Balkans (Stefanovic et al., 2024; Ninausz et al., 2023; Glove et al. 2015). The higher probability of hybridization in these regions can be attributed to four factors: (1) fewer restrictions on human activities affecting natural habitats (e.g. habitat modifications), (2) lower conservation status and illegal hunting of wild canids, (3) large population sizes of

free-ranging dogs, and (4) range expansion of wild canids. These factors have been elaborated in the sections below.

Hybridization due to anthropogenic disturbance has been documented in many species (Grabenstein et al., 2022; Nussberger et al., 2023; Fabbri et al., 2023). Human activities and habitat loss have increased the frequency of interspecific hybridization by disrupting mating patterns and creating habitat conditions that may favour hybrid individuals (Szynwelski et al., 2023). Although wolves generally avoid areas with high human densities and major roads (Kabir et al., 2017; Simpson et al., 2023), they demonstrate significant adaptability to landscape heterogeneity and can live in diverse environments, even in regions with larger human population densities (Reinhardt et al., 2019; Chapron et al., 2014; Fritts et al., 2003). Jackals as a generalist and opportunistic species can adapt to human-altered environments, where they exploit diverse anthropogenic resources (Fenton et al., 2021; Torretta et al., 2020). This adaptability to areas with higher human presence and tolerance to human disturbance in wolves and jackals may increase the likelihood of hybridization due to increasing range overlap of these species with domestic dogs.

The second reason that may increase the chance of hybridization is the lower conservation status and illegal hunting of wild canids. The conservation status of wild canids can be significantly affected by the absence or ambiguity of legal regulations, which may have varying implications for hybridization. For example, in Iran several studies indicate that gray wolves are frequently responsible for livestock depredation (Behdarvand et al., 2014; Ghoddousi et al., 2020; Akrim et al., 2021), leading to retaliatory killings by humans (Ghoddousi et al., 2016, 2020; Parchizadeh and Belant, 2021). Anthropogenic mortality, such as poaching, can disrupt the social cohesion within wolf packs (Borg et al., 2015; Brainerd et al., 2008; Cassidy et al., 2023). Losing even a single wolf, especially a breeding individual, can disrupt the pack's stability and lead to its dissolution, particularly when the pack is small (Zubiria Perez et al., 2024). Such disruptions can increase the likelihood of interbreeding with domestic dogs or admixed individuals, thereby facilitating hybridization (Santostasi et al., 2024). Similar dynamics have been observed in other canid systems, such as red wolf–coyote and eastern wolf–coyote interactions (Bohling & Waits, 2015; Rutledge et al., 2012).

The population size of domestic species is closely linked to the size of the human population (Gompper, 2014). Human population growth combined with the fragmentation of natural habitats increases both the number of domestic animals and the probability of encounters with their wild relatives (Adavoudi and Pilot, 2021). The global dog population is currently estimated at around one billion (Gompper, 2014). The presence of large, unmanaged free-ranging dog populations further exacerbates hybridization risks by increasing the frequency of encounters with wild canids. In regions where free-ranging dogs are abundant, the likelihood of interbreeding events rises, leading to gene flow between domestic and wild species. This process, coupled with human-induced landscape modifications, can accelerate the integration of domestic alleles into wild canid populations.

In addition to the effect of human-mediated factors on the hybridization frequency, hybridization appears to occur more frequently in regions of recent species expansion

compared to established core habitats (e.g., in jackals: Stefanovic et al., 2024; in coyotes: Kays et al., 2010). For instance, jackal densities in the areas of recent expansion are low, which may lead to an increased probability of interactions between jackals and dogs (Stefanovic et al., 2024). Finding the evidence of hybridization between jackals and dogs along the northern periphery of the jackal's range and in newly colonized areas aligned with previous research documenting similar patterns of hybridization in expanding populations (Stefanovic et al., 2024; Ninausz et al., 2023; Galov et al., 2015). Collectively, hybridization in canids may be driven by multiple factors such as habitat transformation, legal enforcement gaps, free-ranging dog populations, and species range shifts. These factors can lead to increased probability of the contact between these canids, therefore understanding them can improve wildlife management strategies aimed at reducing hybridization rates.

Introgression rates

Our analysis revealed that the gene pools of Eurasian wolves and jackals have been influenced by introgression from domestic dogs to different extents. The higher rate of dog introgression in wolves compared to jackals may result from the greater evolutionary distance between dog and jackal compared to dog's distance from wolf. The domestic dog diverged from its primary ancestor, the gray wolf, between 11,000 and 32,000 years ago (reviewed in Tancredi & Cardinali, 2023). The divergence between the golden jackal and the grey wolf occurred considerably earlier, between 1.5-2.4 million years ago (Koepfli et al., 2015). The evolutionary distance can be considered as one of the factors that could influence the introgression rate between species. In accordance with our finding, studies on other taxa also showed that the frequency of hybridization in sister species and also in recently diverged lineages is higher than in non-sister or in more distantly related species (Nesi et al., 2011; Gholamhosseini et al., 2013; Wei et al., 2023).

Our results also align with earlier studies that reported dog admixture in Eurasian wolves (e.g. Pilot et al. 2018; Harmoinen et al., 2021; Stronen et al., 2022; Sarabia et al., 2025). However, this study detected a higher level of introgression affecting the Eurasian wolf gene pool in comparison to an earlier study on wolf-dog hybridization across Eurasia (Pilot et al., 2021), which identified a 1.3% introgression of dog ancestry in wolves using ELAI analysis (including admixed and non-admixed wolves). This discrepancy can be attributed to the use of a larger number of SNPs in this study (229,120 vs. 106,549), which provided a higher resolution and enabled the detection of more subtle introgression signals. The larger wolf sample size compared to the previous study (315 vs 178 individuals), which includes wolves from regions with high admixture rates such as India and Iran, may have led to a higher observed introgression rate.

By excluding samples with low genotype quality (LQ samples), the estimated rate of dog ancestry in wolves remained unchanged, whereas the admixture proportion in jackals dropped significantly to 1.3%. This discrepancy results from the fact that a larger proportion of the jackal samples had low quality (57 of 191 admixed jackal samples, i.e., 30%, and 45 of 287 pure jackal samples, i.e., 16%). One possible explanation for having more LQ samples in jackals is that this study was based on the SNP arrays which were

developed based on the dog and wolf genomes. The probe design of these arrays is optimized for dogs, and when applied to species with greater evolutionary divergence, such as jackals, it can lead to an increased rate of missing data. The estimated wolf and jackal introgression rate in putatively admixed dogs, based on the pairwise comparison using ELAI analysis, was approximately 1.7% and 0%, respectively—substantially lower than the introgression rates observed in the opposite direction. The low admixture rates in dogs, compared to wild canids, may be explained by the large population size of domestic dogs (Pilot et al., 2021), which reduces the impact of hybridisation events on the gene pool. In large dog populations, a single back-crossing event—where a wolf-dog hybrid successfully reproduces with a dog—has only a minor effect on the overall gene pool, which explains the low levels of wolf admixture in the free-ranging dog population (Pilot et al., 2021). This study also observed a slightly lower level of admixture in the dog gene pool (1.7%) compared to the previous study's finding of 2.3% (Pilot et al., 2021). However, this difference likely falls within the statistical margin of error for admixture detection.

The estimated jackal ancestry in admixed wolves using a three-way analysis was nearly zero. The estimated rate of wolf ancestry introgression in admixed jackals was initially 3.1% based on the wolf-jackal dataset. However, removing the low-quality (LQ) samples reduced the estimated introgression rate to 1.2%. Based on the three-way admixture analysis, some of the low-quality samples, three with the prevailing jackal and one with the prevailing wolf ancestry, were identified with consistent proportions of dog, wolf, and jackal ancestry across all chromosomes (Table 3.3, Fig. 3.2). The consistent ancestry proportions in all chromosomes in these samples suggest that the inferred admixture is correct.

The presence of the samples with ancestries from the three canids suggest complex admixture patterns, e.g. jackals cross-breeding with wolf-dog hybrids or backcrosses. This is in line with the finding of putative F1 hybrids that do not have varying admixture proportions within individual chromosomes, which could be explained by F1 x F1 hybrid cross. F1 hybrids and recent backcrosses are rare and therefore cross-breeding between them is statistically unlikely. However, such admixed individuals may be locally abundant and potentially show behavioural patterns (e.g. habitat preferences or mate preferences) that may favour cross-breeding between them. This is an interesting topic to be explored in the future.

As shown in the previous chapter, low-quality data may bias the estimates of the introgression rates. On the other hand, it is also possible that the presence of admixture may result in reduced genotyping quality using the Axiom Array. However, high-quality genotypes of admixed samples were produced as well, including those of F1 individuals. Moreover, there were 45 LQ samples (16%) among 287 non-admixed jackals, implying that missing data does not always lead to the inference of admixture. Although admixture does not always reduce genotyping quality, in some samples, hybridization may increase the amount of missing data and contribute to low-quality genotyping. For this reason, all LQ samples were included in the remaining analyses. Using high-coverage whole-genome resequencing in future studies could help clarify this issue.

Chromosomal blocks with overrepresented or underrepresented introgressed variants

Although the rate of introgression from domestic dogs into the wolf gene pool was higher than in the opposite direction, the number and average length of chromosomal blocks with significantly overrepresented hybridization-derived dog ancestry in wolves were smaller compared to the opposite direction. Specifically, wolves had eight overrepresented introgressed blocks from dog ancestry across seven chromosomes, whereas dogs had 31 overrepresented introgressed chromosomal blocks from wolf ancestry across 23 chromosomes. This suggests that adaptive introgression from wolves to domestic dogs has likely been more prominent than the reverse. These findings are largely consistent with a previous study (Pilot et al., 2021). However, there are some differences between the two studies. For example, here we identified chromosomal blocks of dog ancestry on seven chromosomes in wolves, whereas Pilot et al. (2021) found these blocks on 15 chromosomes, with only three overlapping between the two studies. In dogs, the current study identified chromosomal blocks from wolf ancestry on 23 chromosomes, while Pilot et al. (2021) identified them on 20 chromosomes with 15 chromosomes being common across both studies (Please see results). The differences between the two studies could be attributed to several factors such as sample sizes and differences in geographical origins of samples. Moreover, the criteria used in both studies to identify blocks with overrepresented hybrid ancestry were strict (with a threshold of 3 SD from the mean), therefore, both studies can have a considerable number of false negatives.

Dogs appear to have acquired a larger pool of beneficial genetic variants from wolves, which supports the idea that higher genome-wide admixture proportions do not necessarily correlate with a larger number of blocks containing overrepresented hybrid ancestry, as noted by Pilot et al. (2021). One possible explanation for this result is the differentiation of population sizes between domestic dogs and wolves. Domestic dogs have much larger populations compared to wolves, which can substantially impact the effectiveness of natural selection and genetic drift. Due to the relatively weak effect of genetic drift in large populations, gene variants with a slight selective advantage may rise in frequency. In contrast, the much smaller population sizes of grey wolves make adaptive variants more likely to be eliminated by genetic drift, even those with large selective advantages (Manning et al., 2013). Therefore, this difference in population size likely accounts for the greater number of overrepresented introgressed chromosomal blocks in dogs, which may contain genetic variants that help mitigate some negative consequences of domestication, such as decreased genetic diversity and heightened tameness (Pilot et al., 2021).

In jackal samples, seven and one introgressed chromosomal blocks with an average of 0.184 and 0.086 of dog ancestry and wolf ancestry were found, respectively. This result showed that jackals are more influenced by adaptive introgression from dogs compared to wolves. Only five wolves showed small proportions of golden jackal admixture, suggesting that introgression from jackals has a very limited effect on the wolf gene pool. Accordingly, jackal introgression in dogs was negligible.

72 chromosomal blocks with underrepresented wolf ancestry in admixed dogs were found within 28 chromosomes, while no dog ancestry deserts in admixed wolves were detected. This finding largely aligns with the earlier study (Pilot et al., 2021), though the present study identified a greater number of underrepresented chromosomal blocks (71 vs. 10 blocks). Pilot et al. (2021) reported underrepresented chromosomal blocks in admixed dogs on seven chromosomes (Chr 1, 4, 6, 10, 20, 21, and 27), all of which were also identified in the present study as containing ancestry deserts in dogs. Furthermore, both studies consistently found that admixed dogs exhibited more underrepresented chromosomal blocks than admixed wolves. Higher number of admixed ancestry deserts in dogs compared to wolves may be explained by the fact that the majority of deleterious mutations in dogs are only mildly harmful (Marsden et al., 2016). When these mutations enter into the wolf gene pool, they can persist rather than be eliminated (Pilot et al., 2021). This is because, in species with small effective population sizes like wolves, genetic drift has a stronger influence than natural selection, making it less effective at removing weakly deleterious mutations (Kimura, 1964; Wright, 1931). In contrast, the large population size of dogs allows natural selection to act more efficiently against mildly harmful wolf ancestry in admixed dogs, leading to a higher number of ancestry deserts from wolves in dogs (Pilot et al., 2021).

In admixed golden jackals, dog ancestry deserts were identified on three chromosomes (Chr 7, 24, and 30), and wolf ancestry deserts were found on chromosomes 7 and 30. These results suggest that only a small proportion of genetic variation derived from dogs and wolves has a strongly deleterious effect on jackals.

The function of genes located in chromosomal regions with excess introgressed ancestry

Based on our results in wolf-dog dataset, dogs had 595 candidate genes under adaptive introgression, while Eurasian wolves had only 114 genes. In Eurasian wolves, candidate adaptive genes were enriched for GO terms related to the 'Major Histocompatibility Complex' (MHC) and 'leukocyte-mediated cytotoxicity.' Both MHC and leukocyte-mediated cytotoxicity play a significant role in the immune system. Moreover, when we ran gene ontology analysis just for 20 candidate adaptive genes that were located in the chromosomes (chr 1, 13, and 28) found as having overrepresented dog ancestry in both the present study and the earlier study (Pilot et al., 2021), these genes were enriched for GO terms related to the MHC and melanocortin receptor activity. These genes are therefore strong candidates for adaptive introgression, suggesting that gene flow from dogs may contribute to various aspects of wolf adaptation.

The MHC is a large genetic region, comprising multiple subregions, containing many genes that play a critical role in the adaptive immune response, particularly in antigen presentation pathways (Wong-Benito et al., 2023). High variation in immune system genes like MHC is often associated with increased parasite resistance (Šimková et al., 2021) and has been proposed to improve the survival of individuals (Sommer, 2005; Niskanen et al., 2013).

Adaptive introgression has been repeatedly observed in specific functional categories of genes, notably those related to the immune system. For instance, adaptive introgression from Neanderthal ancestry in modern humans is enriched for proteins that interact with viruses (Enard and Petrov, 2018). Similarly, adaptive introgression in immune-related genes has been suggested in several species, such as newts (Dudek et al., 2019; Gaczorek et al., 2023a), lizards (Gaczorek et al., 2023b), goats (Grossen et al., 2014; Munger et al., 2024), sheep (Barbato et al., 2017), wildcats (Howard-McCombe et al., 2023), wolves (Niskanen et al., 2014; Liu et al., 2017; Rocha et al., 2019; Kloch et al., 2021; Sarabia et al., 2025) and also in different vertebrates (Gaczorek et al., 2024).

MHC sequence variation in wolves is generally less polymorphic than in domestic dogs, likely due to factors such as fragmentation of wolf populations (Niskanen et al., 2012) and genetic bottlenecks (e.g., Fabbri et al., 2007). In contrast, free-ranging dogs typically maintain higher MHC variability (Runstadler et al., 2006). This greater polymorphism in dogs is likely driven by balancing selection, a larger effective population size, and denser dog populations compared to wolves, which may result in higher prevalence of infectious diseases (Niskanen et al., 2012). Additionally, the diverse range of human-dominated environments occupied by free-ranging dogs increases their exposure to pathogens, thereby promoting greater MHC polymorphism. MHC genes tend to be under balancing selection, given that high variability is beneficial, and novel variants from dogs may increase the overall MHC variability in wolves and thus individual fitness. Therefore, for wolves, adaptive introgression of MHC genes, along with other immune-related genes like those involved in leukocyte-mediated cytotoxicity, could enhance parasite resistance and immune responses. Factors like habitat modification and declining prey populations may bring wolves closer to human settlements, which can increase wolves exposure to novel pathogens and parasites (e.g. via contact with livestock and domestic dogs). In this situation adaptive introgression of immune-related genes from domestic dogs is beneficial for wolves. MHC alleles of domestic dog origin were detected in admixed Italian wolves (Galaverni et al., 2013; Sarabia et al., 2025).

We also found melanocortin 2 receptor (MC2R) gene related to the melanocortin receptors (MCRs) in wolves. MC2R is primarily involved in adrenal function, leading to cortisol production (Chida et al., 2007). In mammals, including canids, MC2R plays a key role in regulating stress responses and energy metabolism through the hypothalamic–pituitary–adrenal (HPA) axis (Smith and Vale, 2006). We suggest that adaptive introgression of MC2R from domestic dogs into wolf populations could potentially facilitate better adaptation of wolves to anthropogenically altered environments, especially in traits related to stress physiology, and metabolic regulation.

In admixed dogs, four molecular function GO terms were linked to multiple genes within the oxidoreductase activity group. By running GO analysis just for candidate adaptive genes that were located in the overlapped chromosomes between the present study and the earlier study (Pilot et al., 2021), we found enriched for GO terms related to the oxidoreductase activity group as well. Oxidoreductases are essential enzymes in plants, animals, and microorganisms (Das and Sen, 2024) that participate in a variety of physiological processes, for example in the synthesis of biomolecules, degradation, and

removal of molecules, and metabolism of exogenous molecules like drugs (Braune et al., 2019).

Additionally, 23 GO terms for biological processes were significantly enriched, including those related to metabolic processes, biosynthetic processes, and transcription regulation. Metabolism is fundamental to life-sustaining processes and represents a complex system of biochemical reactions that control the concentration and reaction rates of substrates and products (Brown et al., 2004). Candidate genes involved in metabolic processes, biosynthesis, and transcriptional regulation have been identified as key factors influencing the adaptation of both humans and animals to high-altitude environments (Qiu et al., 2012; Frede and Fandrey, 2013; Projecto-Garcia et al., 2013; Wang et al., 2014; Yang et al., 2017; Miao et al., 2017).

Golden jackals had 94 candidate genes under adaptive introgression from dogs. The candidate genes were enriched for GO terms related to the 'endoplasmic reticulum protein-containing complex' based on the canine and human genome assembly. The endoplasmic reticulum (ER) protein-containing complex is a crucial cellular component that plays a central role in cell homeostasis and survival. The ER is responsible for various essential cellular functions, including protein folding, the synthesis of a wide range of cellular lipids, and the regulation and maintenance of Ca^{2+} homeostasis (Chen et al., 2023; Schuldiner and Weissman, 2013; Sammels et al., 2010; Rutkowski and Kaufman, 2003; Sorger and Daum, 2003). Beyond its fundamental role in cellular maintenance, the ER's involvement in immune response highlights its broader significance in defending against infections and other environmental stressors. The ER serves as the primary site for the assembly of molecular complexes involved in antigen presentation via major histocompatibility complex (MHC) molecules (Roy et al., 2006). MHC I, located in the endoplasmic reticulum, is essential for presenting endogenous peptides on the cell surface (Thoma et al., 2023). This process is critical for the immune system's ability to recognize and respond to pathogens.

Given its importance in both cellular function and immune response, genes associated with the ER are likely to offer adaptive advantages, particularly when organisms encounter new environmental pressures, such as changes in diet, climate, or disease exposure. In the context of adaptive introgression, the transfer of genes from domestic dogs to jackals may have introduced beneficial genetic variants that enhance the function of the ER. Additionally, adaptive introgression may influence the jackals' ability to present antigens more efficiently through MHC molecules, improving their immune recognition and response. This would be particularly beneficial in environments where jackals encounter dogs frequently and where pathogens are constantly evolving, as it would allow the jackals to improve their immune defences. Thus, the introgression of these ER-associated genes from domestic dogs could enhance the jackals' overall fitness, providing them with the flexibility to adapt to rapidly changing conditions and improve their chances of survival in diverse environments.

In golden jackals with grey wolf admixture, genes located in regions with overrepresented wolf ancestry were enriched for the GO term "cyclin E1-CDK2 complex" based on the Benjamini–Hochberg FDR (false discovery rate), and the g:SCS (Set Counts and Sizes) thresholds. Cyclin E1-CDK2 plays a significant role in the cell

cycle, DNA replication, and cellular stress responses (Honda et al., 2005). Adaptive introgression of cyclin E1-CDK2 from wolves into the gene pool of jackals probably confers significant benefits related to stress response, immune function, and cellular repair, which can help jackals to better adapt and survive in challenging habitats.

The function of candidate genes showing signatures of adaptive introgression

Among genes located in chromosomal blocks with the excess of introgressed ancestry in each species, we identified genes under positive selection. These genes were strong candidates for adaptive introgression. However, it must be stressed that this approach does not identify genes that may be subject to balancing selection following the introgression of variants from another species.

In dogs, six SNPs found to be under positive selection were associated with NKAIN2 (Sodium/Potassium Transporting ATPase Interacting 2) on chromosome 1. Therefore, a variant of this gene introgressed from wolves likely undergoes positive selection in dogs. The function of NKAIN2 is not well investigated, however, the limited knowledge of this gene points to its abundant expression in the brain and nervous system development (Zhao et al., 2015). We also detected CACNA1C (Calcium Voltage-Gated Channel Subunit Alpha1 C) as a candidate gene under positive selection in dogs, consistent with previous findings (Pilot et al., 2021; Sarabia et al., 2025). CACNA1C plays a key role in dendritic development, synaptic plasticity, neuronal survival, memory, and learning (Bhat et al., 2012), making it a strong candidate for involvement in domestication-related behaviour changes. The importance of synaptic plasticity in domestication has been demonstrated in red foxes (*Vulpes vulpes*), where genes related to this process showed signatures of selection (Kukekova et al., 2018). Another gene that was identified as undergoing positive selection in dogs is SH3RF2 gene that plays a crucial role in neuronal signalling and synaptic plasticity, which are essential for learning and memory (Wang et al., 2018). Mice with SH3RF2 haploinsufficiency (having only one functional copy of the gene) exhibit significant deficits in social interaction and communication (Wang et al., 2018). We also found three olfactory receptor (OR) genes (OR6D6, OR6D7, and OR13A1) downstream of an SNP under positive selection, located in chromosome 28. This result is consistent with earlier studies, since 35 OR genes were found within overrepresented introgressed variants in free-ranging dogs (Pilot et al., 2021). Research on OR showed that domestic dogs might have lost functional OR genes commensurate with a documented reduction in nasal morphology as an outcome of the domestication process (Mouton et al., 2025). Therefore, adaptive introgression of OR genes from wolves may enhance olfaction in dogs to facilitate their detection ability of the sources of suitable and unsuitable food and detection of potential threats like larger predators (Pilot et al., 2021).

In wolves, genes with signatures of positive selection located in blocks with overrepresented dog ancestry are expressed in the brain and are related to nervous system function and cellular signalling. For example, ATP8A2 is expressed in the central nervous system and retina (Li et al., 2023) and is required for normal neural development and function through its role in cell proliferation, migration, and synaptic pruning

(Coleman and Molday, 2011; Xu et al., 2012). GRK1 is related to dim-light vision and photoresponse recovery (Weiss et al., 2001). The positive selection of genes associated with the dim-light vision, including GRK1, has been reported before in mammals (Wu et al., 2017). This gene was enriched for GO terms related to the "rhodopsin kinase activity" which plays a role in visual processes (Choi et al., 2001). We also found genes related to the immune system, such as LAMP1 gene (Xu et al., 2024). Which is related to immune system (Xu et al., 2024; Wang et al., 2024).

3.5. Conclusion

We found that hybridization occurs across the distribution ranges of canids, with the highest concentrations in the Balkans, India, the Caucasus, and northeastern Europe. Factors like high human disturbances, large population size of free-ranging dogs, and the range expansion of golden jackals can contribute as main factors for the high rates of hybridization in these regions. Based on the three-way ELAI analysis, the average dog ancestry proportions in all wolves and golden jackal samples (excluding F1 hybrids) were estimated at 3.3% and 0.9%, respectively. The average wolf ancestry proportion in admixed dogs was 0.3%, while no golden jackal ancestry was detected in dogs. The higher proportion of introgression in wolves and dogs compared to jackals may be explained by the closer evolutionary similarity between the first two canids. This is consistent with the expectation that the introgression rate decreases with increasing evolutionary distance between species.

Both wild canids and free-ranging dogs may gain benefits from hybridization. Through adaptive introgression, wild canids may acquire from dogs gene variants conferring adaptive advantage, including those that strengthen their immune systems. These beneficial genes may increase the resistance of wild canids to new pathogens, which would be particularly beneficial in environments where wild canids encounter dogs frequently and where pathogens are constantly evolving. Therefore, we can expect that introgression from dogs could provide an adaptive advantage for wild canids, especially those living in regions highly modified by humans. Free-ranging dogs appear to have acquired a larger pool of beneficial genetic variants from wolves, which may have contributed to some characteristics like morphological, behavioural, and physiological traits. Among genes that were under positive selection in canids, genes related to the nervous and immune systems were predominant in wild canids and free-ranging dogs. In addition to detecting signals of positive selection, we also found signs of negative selection in introgressed chromosomal blocks in dogs and golden jackals. Identifying chromosomal blocks with underrepresented wolf ancestry in admixed dogs and chromosomal blocks with underrepresented wolf and dog ancestry in admixed jackals suggested that introgression may have a deleterious effect on these species, but they can be efficiently removed from their gene pools. Overall, we highlight the complex nature of hybridization and introgression in the evolutionary process, showing that it can introduce both beneficial and maladaptive genetic variation.

3.6. Bibliography

Adavoudi, R., & Pilot, M. (2021). Consequences of hybridization in mammals: A systematic review. Genes, 13(1), 50.

Aghbolaghi, M. A., Rezaei, H. R., Scandura, M., & Kaboli, M. (2014). Low gene flow between Iranian Grey Wolves (*Canis lupus*) and dogs documented using uniparental genetic markers. Zoology in the Middle East, 60(2), 95-106.

Akrim, F., Mahmood, T., Belant, J. L., Nadeem, M. S., Qasim, S., Zangi, I. U. D., & Asadi, M. A. (2021). Livestock depredations by leopards in Pir Lasura National Park, Pakistan: characteristics, control and costs. Wildlife Biology, 2021(1), 1-7.

Anderson, E. (1949). Introgressive hybridization.

Anderson, E., & Hubricht, L. (1938). Hybridization in Tradescantia. III. The evidence for introgressive hybridization. American Journal of Botany, 396-402.

Andersson, L., & Purugganan, M. (2022). Molecular genetic variation of animals and plants under domestication. Proceedings of the National Academy of Sciences, 119(30), e2122150119.

Arendt, M., Cairns, K. M., Ballard, J. W. O., Savolainen, P., & Axelsson, E. (2016). Diet adaptation in dog reflects spread of prehistoric agriculture. Heredity, 117(5), 301-306.

Arnold, J., Humer, A., Heltai, M., Murariu, D., Spassov, N., & Hacklaender, K. (2012). Current status and distribution of golden jackals *Canis aureus* in Europe. Mammal Review, 42(1), 1-11.

Barbato, M., Hailer, F., Orozco-terWengel, P., Kijas, J., Mereu, P., Cabras, P., ... & Bruford, M. W. (2017). Genomic signatures of adaptive introgression from European mouflon into domestic sheep. Scientific Reports, 7(1), 7623.

Barton, N. H. (2001). The role of hybridization in evolution. Molecular ecology, 10(3), 551-568.

Behdarvand, A., Zamani, M. S., Sadeghi, F., Yahyapour, Y., Vaziri, F., Jamnani, F. R., ... & Siadat, S. D. (2017). Evaluation of Merkel cell polyomavirus in non-small cell lung cancer and adjacent normal cells. Microbial Pathogenesis, 108, 21-26.

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

Borg, B. L., Brainerd, S. M., Meier, T. J., & Prugh, L. R. (2015). Impacts of breeder loss on social structure, reproduction and population growth in a social canid. Journal of Animal Ecology, 84(1), 177-187.

Brainerd, S. M., Andrén, H., Bangs, E. E., Bradley, E. H., Fontaine, J. A., Hall, W., ... & Wydeven, A. P. (2008). The effects of breeder loss on wolves. The Journal of Wildlife Management, 72(1), 89-98.

Braune, A., Gütschow, M., & Blaut, M. (2019). An NADH-dependent reductase from Eubacterium ramulus catalyzes the stereospecific heteroring cleavage of flavanones and flavanonols. Applied and Environmental Microbiology, 85(19), e01233-19.

Brown, J. H., Gillooly, J. F., Allen, A. P., Savage, V. M., & West, G. B. (2004). Toward a metabolic theory of ecology. Ecology, 85(7), 1771-1789.

Cairns, K. M., Crowther, M. S., Parker, H. G., Ostrander, E. A., & Letnic, M. (2023). Genome-wide variant analyses reveal new patterns of admixture and population structure in Australian dingoes. Molecular Ecology, 32(15), 4133-4150.

Carneiro, M., Rubin, C. J., Di Palma, F., Albert, F. W., Alföldi, J., Barrio, A. M., ... & Andersson, L. (2014). Rabbit genome analysis reveals a polygenic basis for phenotypic change during domestication. Science, 345(6200), 1074-1079.

Cassidy, K. A., Borg, B. L., Klauder, K. J., Sorum, M. S., Thomas-Kuzilik, R., Dewey, S. R., ... & Smith, D. W. (2023). Human-caused mortality triggers pack instability in gray wolves. Frontiers in Ecology and the Environment, 21(8), 356-362.

Chapron, G., Kaczensky, P., Linnell, J. D., Von Arx, M., Huber, D., Andrén, H., ... & Boitani, L. (2014). Recovery of large carnivores in Europe's modern human-dominated landscapes. Science, 346(6216), 1517-1519.

Chen, H., Chen, E., Liu, M., Wang, J., Yin, J., Zhao, P., & Xu, Y. (2024). Identification of immune-related endoplasmic reticulum stress genes in proliferative diabetic retinopathy using bioinformatics analysis. Frontiers in Endocrinology, 15, 1341206.

Choi, S., Hao, W., Chen, C. K., & Simon, M. I. (2001). Gene expression profiles of light-induced apoptosis in arrestin/rhodopsin kinase-deficient mouse retinas. Proceedings of the National Academy of Sciences, 98(23), 13096-13101.

Cunze, S., & Klimpel, S. (2022). From the Balkan towards Western Europe: Range expansion of the golden jackal (*Canis aureus*)—A climatic niche modeling approach. Ecology and Evolution, 12(7), e9141.

Daïnou, K., Flot, J. F., Degen, B., Blanc-Jolivet, C., Doucet, J. L., Lassois, L., & Hardy, O. J. (2017). DNA taxonomy in the timber genus Milicia: evidence of unidirectional introgression in the West African contact zone. Tree Genetics & Genomes, 13, 1-12.

Das, P., & Sen, P. (2024). Relevance of Oxidoreductases in Cellular Metabolism and Defence. In Reactive Oxygen Species-Advances and Developments. IntechOpen.

Donfrancesco, V., Ciucci, P., Salvatori, V., Benson, D., Andersen, L. W., Bassi, E., ... & Mukherjee, N. (2019). Unravelling the scientific debate on how to address wolf-dog hybridization in Europe. Frontiers in Ecology and Evolution, 7, 175.

Dudek, K., Gaczorek, T. S., Zieliński, P., & Babik, W. (2019). Massive introgression of major histocompatibility complex (MHC) genes in newt hybrid zones. Molecular Ecology, 28(21), 4798-4810.

Dumenko, V. N., Kozlov, M. K., & Kulikov, M. A. (2001). Selective attention in dogs from energy characteristics of neocortical potentials in the 1-220 Hz band. Zhurnal Vysshei Nervnoi Deiatelnosti Imeni IP Pavlova, 51(6), 671-682.

Edelman, N. B., & Mallet, J. (2021). Prevalence and adaptive impact of introgression. Annual Review of Genetics, 55(1), 265-283.

Enard, D., & Petrov, D. A. (2018). Evidence that RNA viruses drove adaptive introgression between Neanderthals and modern humans. Cell, 175(2), 360-371.

Fabbri, E., Miquel, C., Lucchini, V., Santini, A., Caniglia, R., Duchamp, C., ... & Randi, E. (2007). From the Apennines to the Alps: colonization genetics of the naturally expanding Italian wolf (*Canis lupus*) population. Molecular Ecology, 16(8), 1661-1671.

Fan, Z., Silva, P., Gronau, I., Wang, S., Armero, A. S., Schweizer, R. M., ... & Wayne, R. K. (2016). Worldwide patterns of genomic variation and admixture in gray wolves. Genome Research, 26(2), 163-173.

Fang, M., Larson, G., Soares Ribeiro, H., Li, N., & Andersson, L. (2009). Contrasting mode of evolution at a coat color locus in wild and domestic pigs. PLoS Genetics, 5(1), e1000341.

Fenton, S., Moorcroft, P. R., Ćirović, D., Lanszki, J., Heltai, M., Cagnacci, F., ... & Ranc, N. (2021). Movement, space-use and resource preferences of European golden jackals in human-dominated landscapes: insights from a telemetry study. Mammalian Biology, 101, 619-630.

Frede, S., & Fandrey, J. (2013). Cellular and molecular defenses against hypoxia. In High altitude: human adaptation to hypoxia (pp. 23-35). New York, NY: Springer New York.

Fritts, S. H., Stephenson, R. O., Hayes, R. D., & Boitani, L. (2003). Wolves and humans. Wolves: Behavior, Ecology, and Conservation, 289-316.

Gabriele-Rivet, V., Brookes, V. J., Stephens, D., Arsenault, J., & Ward, M. P. (2021). Hybridisation between dingoes and domestic dogs in proximity to Indigenous communities in northern Australia. Australian Veterinary Journal, 99(9), 388-391.

Gaczorek, T. S., Chechetkin, M., Dudek, K., Caeiro-Dias, G., Crochet, P. A., Geniez, P., ... & Babik, W. (2023b). Widespread introgression of MHC genes in Iberian Podarcis lizards. Molecular Ecology, 32(14), 4003-4017.

Gaczorek, T. S., Marszałek, M., Dudek, K., Arntzen, J. W., Wielstra, B., & Babik, W. (2023a). Interspecific introgression of MHC genes in Triturus newts: Evidence from multiple contact zones. Molecular Ecology, 32(4), 867-880.

Gaczorek, T., Dudek, K., Fritz, U., Bahri-Sfar, L., Baird, S. J. E., Bonhomme, F., ... & Babik, W. (2024). Widespread adaptive introgression of major histocompatibility complex genes across vertebrate hybrid zones. Molecular Biology and Evolution, 41(10), msae201.

Galaverni, M., Caniglia, R., Fabbri, E., Lapalombella, S., & Randi, E. (2013). MHC variability in an isolated wolf population in Italy. Journal of Heredity, 104(5), 601-612.

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(9), 2324-2339.

Galov, A., Fabbri, E., Caniglia, R., Arbanasić, H., Lapalombella, S., Florijančić, T., ... & Randi, E. (2015). First evidence of hybridization between golden jackal (*Canis aureus*) and domestic dog (*Canis familiaris*) as revealed by genetic markers. Royal Society Open Science, 2(12), 150450.

Ghoddousi, A., Bleyhl, B., Sichau, C., Ashayeri, D., Moghadas, P., Sepahvand, P., ... & Kuemmerle, T. (2020). Mapping connectivity and conflict risk to identify safe corridors for the Persian leopard. Landscape Ecology, 35, 1809-1825.

Ghoddousi, A., Kh. Hamidi, A., Soofi, M., Khorozyan, I., Kiabi, B. H., & Waltert, M. (2016). Effects of ranger stations on predator and prey distribution and abundance in an I ranian steppe landscape. Animal Conservation, 19(3), 273-280.

Gorokhova, S., Bibert, S., Geering, K., & Heintz, N. (2007). A novel family of transmembrane proteins interacting with β subunits of the Na, K-ATPase. Human Molecular Genetics, 16(20), 2394-2410.

Gottelli, D., SILLERO-ZUBIRI, C., Applebaum, G. D., Roy, M. S., Girman, D. J., GARCIA-MORENO, J., ... & Wayne, R. K. (1994). Molecular genetics of the most endangered canid: the Ethiopian wolf *Canis simensis*. Molecular Ecology, 3(4), 301-312.

Grabenstein, K. C., Otter, K. A., Burg, T. M., & Taylor, S. A. (2023). Hybridization between closely related songbirds is related to human habitat disturbance. Global Change Biology, 29(4), 955-968.

Gray, A. J. (2005). Hybridization between crops and wild plants in the age of genetic engineering: new risks or new paradigms?.

Grossen, C., Keller, L., Biebach, I., International Goat Genome Consortium, & Croll, D. (2014). Introgression from domestic goat generated variation at the major histocompatibility complex of alpine ibex. PLoS Genetics, 10(6), e1004438.

Harmoinen, J., von Thaden, A., Aspi, J., Kvist, L., Cocchiararo, B., Jarausch, A., ... & Nowak, C. (2021). Reliable wolf-dog hybrid detection in Europe using a reduced SNP panel developed for non-invasively collected samples. BMC Genomics, 22(1), 473.

Harrison, P. W., Amode, M. R., Austine-Orimoloye, O., Azov, A. G., Barba, M., Barnes, I., ... & Yates, A. D. (2024). Ensembl 2024. Nucleic Acids Research, 52(D1), D891-D899.

Harrison, R. G., & Larson, E. L. (2014). Hybridization, introgression, and the nature of species boundaries. Journal of Heredity, 105(S1), 795-809.

Hindrikson, M., Remm, J., Pilot, M., Godinho, R., Stronen, A. V., Baltrūnaité, L., ... & Saarma, U. (2017). Wolf population genetics in E urope: a systematic review, meta-analysis and suggestions for conservation and management. Biological Reviews, 92(3), 1601-1629.

Honda, R., Lowe, E. D., Dubinina, E., Skamnaki, V., Cook, A., Brown, N. R., & Johnson, L. N. (2005). The structure of cyclin E1/CDK2: implications for CDK2 activation and CDK2-independent roles. The EMBO journal, 24(3), 452-463.

Howard-McCombe, J., Jamieson, A., Carmagnini, A., Russo, I. R. M., Ghazali, M., Campbell, R., ... & Beaumont, M. A. (2023). Genetic swamping of the critically endangered Scottish wildcat was recent and accelerated by disease. Current Biology, 33(21), 4761-4769.

Jarausch, A., von Thaden, A., Sin, T., Corradini, A., Pop, M. I., Chiriac, S., ... & Nowak, C. (2023). Assessment of genetic diversity, population structure and wolf-dog hybridisation in the Eastern Romanian Carpathian wolf population. Scientific Reports, 13(1), 22574.

Johansen, K. H., Golec, D. P., Huang, B., Park, C., Thomsen, J. H., Preite, S., ... & Schwartzberg, P. L. (2022). A CRISPR screen targeting PI3K effectors identifies RASA3 as a negative regulator of LFA-1–mediated adhesion in T cells. Science Signaling, 15(743), eabl9169.

Kamler, J. F., Minge, C., Rostro-García, S., Gharajehdaghipour, T., Crouthers, R., In, V., ... & Macdonald, D. W. (2021). Home range, habitat selection, density, and diet of golden jackals in the Eastern Plains Landscape, Cambodia. Journal of Mammalogy, 102(2), 636-650.

Kays, R., Curtis, A., & Kirchman, J. J. (2010). Rapid adaptive evolution of northeastern coyotes via hybridization with wolves. Biology Letters, 6(1), 89-93.

Khosravi, R., Rezaei, H. R., & Kaboli, M. (2013). Detecting hybridization between Iranian wild wolf (*Canis lupus pallipes*) and free-ranging domestic dog (*Canis familiaris*) by analysis of microsatellite markers. Zoological Science, 30(1), 27-34.

Kloch, A., Biedrzycka, A., Szewczyk, M., Nowak, S., Niedźwiedzka, N., Kłodawska, M., ... & W Mysłajek, R. (2021). High genetic diversity of immunity genes in an expanding population of a highly mobile carnivore, the grey wolf Canis lupus, in Central Europe. Diversity and Distributions, 27(9), 1680-1695.

Kopaliani, N., Shakarashvili, M., Gurielidze, Z., Qurkhuli, T., & Tarkhnishvili, D. (2014). Gene flow between wolf and shepherd dog populations in Georgia (Caucasus). Journal of Heredity, 105(3), 345-353.

Kudrenko, S., Fenchuk, V., Vollering, J., Zedrosser, A., Selva, N., Ostapowicz, K., ... & Heurich, M. (2023). Walking on the dark side: Anthropogenic factors limit suitable habitat for gray wolf (*Canis lupus*) in a large natural area covering Belarus and Ukraine. Global Ecology and Conservation, 46, e02586.

Kukekova, A. V., Johnson, J. L., Xiang, X., Feng, S., Liu, S., Rando, H. M., ... & Zhang, G. (2018). Red fox genome assembly identifies genomic regions associated with tame and aggressive behaviours. Nature Ecology & Evolution, 2(9), 1479-1491.

Li, H., Yang, D., Hao, M., & Liu, H. (2022). Differential expression of HAVCR2 gene in pan-cancer: A potential biomarker for survival and immunotherapy. Frontiers in Genetics, 13, 972664.

Lindblad-Toh, K., Garber, M., Zuk, O., Lin, M. F., Parker, B. J., Washietl, S., ... & Kellis, M. (2011). A high-resolution map of human evolutionary constraint using 29 mammals. Nature, 478(7370), 476-482.

Liu, S., Tang, W., Cao, J., Shang, M., Sun, H., Gong, J., & Hu, B. (2022). A comprehensive analysis of HAVCR1 as a prognostic and diagnostic marker for pan-cancer. Frontiers in Genetics, 13, 904114.

Liu, Y., Hardie, J., Zhang, X., & Rotello, V. M. (2017). Effects of engineered nanoparticles on the innate immune system. In Seminars in immunology (Vol. 34, pp. 25-32). Academic Press.

Lobo, D., Morales, H. E., Van Oosterhout, C., López-Bao, J. V., Silva, P., Llaneza, L., ... & Godinho, R. (2025). Ancient dog introgression into the Iberian wolf genome may have facilitated adaptation to human-dominated landscapes. Genome Research, 35(3), 432-445.

Marsden, C. D., Ortega-Del Vecchyo, D., O'Brien, D. P., Taylor, J. F., Ramirez, O., Vilà, C., ... & Lohmueller, K. E. (2016). Bottlenecks and selective sweeps during domestication have increased deleterious genetic variation in dogs. Proceedings of the National Academy of Sciences, 113(1), 152-157.

Mary, N., Iannuccelli, N., Petit, G., Bonnet, N., Pinton, A., Barasc, H., ... & Riquet, J. (2022). Genome-wide analysis of hybridization in wild boar populations reveals adaptive introgression from domestic pig. Evolutionary Applications, 15(7), 1115-1128.

McFarlane, S. E., & Pemberton, J. M. (2019). Detecting the true extent of introgression during anthropogenic hybridization. Trends in Ecology & Evolution, 34(4), 315-326.

McLaren, W., Gil, L., Hunt, S. E., Riat, H. S., Ritchie, G. R., Thormann, A., ... & Cunningham, F. (2016). The ensembl variant effect predictor. Genome Biology, 17, 1-14.

Miao, B., Wang, Z., & Li, Y. (2017). Genomic analysis reveals hypoxia adaptation in the Tibetan mastiff by introgression of the gray wolf from the Tibetan Plateau. Molecular Biology and Evolution, 34(3), 734-743.

Mihaylov, H., & Stoyanov, S. (2012, June). The wolf (*Canis lupus* l., 1758) in Bulgaria. In International symposium on hunting,"Modern aspects of sustainable management of game population", Zemun-Belgrade, Serbia (pp. 22-24).

Milla, R., & Morente-López, J. (2015). Limited evolutionary divergence of seedlings after the domestication of plant species. Plant Biology, 17(1), 169-176.

Miralles, A., Secondi, J., Pabijan, M., Babik, W., Lemaire, C., & Crochet, P. A. (2024). Inconsistent estimates of hybridization frequency in newts revealed by SNPs and microsatellites. Conservation Genetics, 25(1), 215-225.

Mohammadi, A., Alambeigi, A., López-Bao, J. V., & Kaboli, M. (2021). Fear of wolves in relation to attacks on people and livestock in Western Iran. Anthrozoös, 34(2), 303-319.

Molchan, V., Homel, K., Valnisty, A., Nikiforov, M., & Kheidorova, E. (2023). Genetic diversity of mtDNA in the grey wolf population of Belarus threatened by wolf-dog admixture. Theriol Ukr, 25, 87-99.

Moon, A. L., Haan, N., Wilkinson, L. S., Thomas, K. L., & Hall, J. (2018). CACNA1C: association with psychiatric disorders, behavior, and neurogenesis. Schizophrenia Bulletin, 44(5), 958-965.

Moura, A. E., Tsingarska, E., Dąbrowski, M. J., Czarnomska, S. D., Jędrzejewska, B., & Pilot, M. (2014). Unregulated hunting and genetic recovery from a severe population decline: the cautionary case of Bulgarian wolves. Conservation Genetics, 15, 405-417.

Mouton, A., Bird, D. J., Li, G., Craven, B. A., Levine, J. M., Morselli, M., ... & Murphy, W. J. (2025). Genetic and anatomical determinants of olfaction in dogs and wild canids. Molecular Biology and Evolution, 42(3).

Münger, X., Robin, M., Dalén, L., & Grossen, C. (2024). Facilitated introgression from domestic goat into Alpine ibex at immune loci. Molecular Ecology, 33(14), e17429.

Nakajima, R., Deguchi, R., Komori, H., Zhao, L., Zhou, Y., Shirasawa, M., ... & Ohtani, K. (2023). The TFDP1 gene coding for DP1, the heterodimeric partner of the transcription factor E2F, is a target of deregulated E2F. Biochemical and Biophysical Research Communications, 663, 154-162.

Nanda, S. (2011). HAVCR1 variants underlie susceptibility to liver failure in hepatitis A infection—an unlikely link with allergy. Nature Reviews Gastroenterology & Hepatology, 8(5), 244-244.

Nesi, N., Nakoune, E., Cruaud, C., & Hassanin, A. (2011). DNA barcoding of African fruit bats (Mammalia, Pteropodidae). The mitochondrial genome does not provide a reliable discrimination between Epomophorus gambianus and Micropteropus pusillus. Comptes Rendus. Biologies, 334(7), 544-554.

Ninausz, N., Fehér, P., Csányi, E., Heltai, M., Szabó, L., Barta, E., ... & Stéger, V. (2023). White and other fur colourations and hybridization in golden jackals (*Canis aureus*) in the Carpathian basin. Scientific Reports, 13(1), 21969.

Niskanen, A. K., Hagström, E., Lohi, H., Ruokonen, M., Esparza-Salas, R., Aspi, J., & Savolainen, P. (2013). MHC variability supports dog domestication from a large number of wolves: high diversity in Asia. Heredity, 110(1), 80-85.

Niskanen, A. K., Kennedy, L. J., Ruokonen, M., Kojola, I., Lohi, H., Isomursu, M., ... & Aspi, J. (2014). Balancing selection and heterozygote advantage in major histocompatibility complex loci of the bottlenecked Finnish wolf population. Molecular Ecology, 23(4), 875-889.

Parchizadeh, J., & Belant, J. L. (2021). Human-caused mortality of large carnivores in Iran during 1980–2021. Global Ecology and Conservation, 27, e01618.

Pendragon, B. (2011). A review of selected features of the family Canidae with reference to its fundamental taxonomic status. Journal of Creation, 25(3), 79-88.

Pilot, M., Moura, A. E., Okhlopkov, I. M., Mamaev, N. V., Manaseryan, N. H., Hayrapetyan, V., ... & Bogdanowicz, W. (2021). Human-modified canids in human-modified landscapes: The evolutionary consequences of hybridization for grey wolves and free-ranging domestic dogs. Evolutionary Applications, 14(10), 2433-2456.

Prerna, S., Edgaonkar, A., & Dubey, Y. (2015). Diet composition of Golden Jackals *Canis aureus* (Mammalia: Carnivora: Canidae) in Van Vihar National Park, India, a small enclosed area. Journal of Threatened Taxa, 7(8), 7422-7427.

Projecto-Garcia, J., Natarajan, C., Moriyama, H., Weber, R. E., Fago, A., Cheviron, Z. A., ... & Storz, J. F. (2013). Repeated elevational transitions in hemoglobin function during the evolution of Andean hummingbirds. Proceedings of the National Academy of Sciences, 110(51), 20669-20674.

Purugganan, M. D. (2022). What is domestication?. Trends in ecology & evolution, 37(8), 663-671.

Qiu, Q., Zhang, G., Ma, T., Qian, W., Wang, J., Ye, Z., ... & Liu, J. (2012). The yak genome and adaptation to life at high altitude. Nature Genetics, 44(8), 946-949.

Randi, E. (2008). Detecting hybridization between wild species and their domesticated relatives. Molecular Ecology, 17(1), 285-293.

Randi, E., Hulva, P., Fabbri, E., Galaverni, M., Galov, A., Kusak, J., ... & Caniglia, R. (2014). Multilocus detection of wolf x dog hybridization in Italy, and guidelines for marker selection. PLoS One, 9(1), e86409.

Raudvere, U., Kolberg, L., Kuzmin, I., Arak, T., Adler, P., Peterson, H., & Vilo, J. (2019). g: Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). Nucleic Acids Research, 47(W1), W191-W198.

Reimand, J., Kull, M., Peterson, H., Hansen, J., & Vilo, J. (2007). g: Profiler—a web-based toolset for functional profiling of gene lists from large-scale experiments. Nucleic Acids Research, 35(suppl_2), W193-W200.

Reinhardt, I., Kluth, G., Nowak, C., Szentiks, C. A., Krone, O., Ansorge, H., & Mueller, T. (2019). Military training areas facilitate the recolonization of wolves in Germany. Conservation Letters, 12(3), e12635.

Rocha, R. G., Magalhães, V., López-Bao, J. V., van der Loo, W., Llaneza, L., Alvares, F., ... & Godinho, R. (2019). Alternated selection mechanisms maintain adaptive diversity in different demographic scenarios of a large carnivore. BMC Evolutionary Biology, 19, 1-13.

Roy, C. R., Salcedo, S. P., & Gorvel, J. P. E. (2006). Pathogen–endoplasmic-reticulum interactions: in through the out door. Nature Reviews Immunology, 6(2), 136-147.

Runstadler, J. A., Angles, J. M., & Pedersen, N. C. (2006). Dog leucocyte antigen class II diversity and relationships among indigenous dogs of the island nations of Indonesia (Bali), Australia and New Guinea. Tissue Antigens, 68(5), 418-426.

Rutkowski, D. T., & Kaufman, R. J. (2004). A trip to the ER: coping with stress. Trends in Cell Biology, 14(1), 20-28.

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(1), 19-33.

Sammels, E., Parys, J. B., Missiaen, L., De Smedt, H., & Bultynck, G. (2010). Intracellular Ca2+ storage in health and disease: a dynamic equilibrium. Cell Calcium, 47(4), 297-314.

Sandee, D., Morrissey, K., Agrawal, V., Tam, H. K., Kramer, M. A., Tracy, T. S., ... & Miller, W. L. (2010). Effects of genetic variants of human P450 oxidoreductase on catalysis by CYP2D6 in vitro. Pharmacogenetics and Genomics, 20(11), 677-686.

Sarabia, C., Salado, I., Fernández-Gil, A., vonHoldt, B. M., Hofreiter, M., Vilà, C., & Leonard, J. A. (2025). Potential Adaptive Introgression From Dogs in Iberian Grey Wolves (*Canis lupus*). Molecular Ecology, e17639.

Schuldiner, M., & Weissman, J. S. (2013). The contribution of systematic approaches to characterizing the proteins and functions of the endoplasmic reticulum. Cold Spring Harbor Perspectives in Biology, 5(3), a013284.

Shakarashvili, M., Kopaliani, N., Gurielidze, Z., Dekanoidze, D., Ninua, L., & Tarkhnishvili, D. (2020). Population genetic structure and dispersal patterns of grey wolfs (*Canis lupus*) and golden jackals (*Canis aureus*) in Georgia, the Caucasus. Journal of Zoology, 312(4), 227-238.

Sidorovich, V., Schnitzler, A., Schnitzler, C., Rotenko, I., & Holikava, Y. (2017). Responses of wolf feeding habits after adverse climatic events in central-western Belarus. Mammalian Biology, 83, 44-50.

Šimková, A., Gettová, L., Civáňová, K., Seifertová, M., Janáč, M., & Vetešník, L. (2021). Diversity of MHC IIB genes and parasitism in hybrids of evolutionarily divergent cyprinoid species indicate heterosis advantage. Scientific Reports, 11(1), 16860.

Skoglund, P., Ersmark, E., Palkopoulou, E., & Dalén, L. (2015). Ancient wolf genome reveals an early divergence of domestic dog ancestors and admixture into high-latitude breeds. Current Biology, 25(11), 1515-1519.

Smith, W. J., Jezierski, M. T., Dunn, J. C., & Clegg, S. M. (2023). Parasite exchange and hybridisation at a wild-feral-domestic interface. International Journal for Parasitology, 53(14), 797-808.

Solberg, M. F., Robertsen, G., Sundt-Hansen, L. E., Hindar, K., & Glover, K. A. (2020). Domestication leads to increased predation susceptibility. Scientific Reports, 10(1), 1929.

Sommer, S. (2005). The importance of immune gene variability (MHC) in evolutionary ecology and conservation. Frontiers in Zoology, 2, 1-18.

Song, Y., Endepols, S., Klemann, N., Richter, D., Matuschka, F. R., Shih, C. H., ... & Kohn, M. H. (2011). Adaptive introgression of anticoagulant rodent poison resistance by hybridization between old world mice. Current Biology, 21(15), 1296-1301.

Sorger, D., & Daum, G. (2003). Triacylglycerol biosynthesis in yeast. Applied Microbiology and Biotechnology, 61, 289-299.

Stefanović, M., Bogdanowicz, W., Adavoudi, R., Martínez-Sosa, F., Doan, K., Flores-Manzanero, A., ... & Pilot, M. (2024). Range-wide phylogeography of the golden jackals (*Canis aureus*) reveals multiple sources of recent spatial expansion and admixture with dogs at the expansion front. Biological Conservation, 290, 110448.

Stefanović, M., Veličković, N., Bončina, A., Potušek, S., Matić, I., Djan, M., & Bužan, E. (2024). Duplication, recombination and weak selection shape evolution at the MHC class II SLA-DRB1 locus in wild boars from the western Balkans. Mammalian Biology, 1-10.

Stronen, A. V., Aspi, J., Caniglia, R., Fabbri, E., Galaverni, M., Godinho, R., ... & Harmoinen, J. (2022). Wolf-dog admixture highlights the need for methodological standards and multidisciplinary cooperation for effective governance of wild x domestic hybrids. Biological Conservation, 266, 109467.

Stronen, A. V., Jędrzejewska, B., Pertoldi, C., Demontis, D., Randi, E., Niedziałkowska, M., ... & Czarnomska, S. D. (2013). North-south differentiation and a region of high diversity in European wolves (*Canis lupus*). PLoS One, 8(10), e76454.

Szatmári, L., Cserkész, T., Laczkó, L., Lanszki, J., Pertoldi, C., Abramov, A. V., ... & Sramkó, G. (2021). A comparison of microsatellites and genome-wide SNPs for the detection of admixture brings the first molecular evidence for hybridization between *Mustela eversmanii* and *M. putorius* (Mustelidae, Carnivora). Evolutionary Applications, 14(9), 2286-2304.

Szynwelski, B. E., Kretschmer, R., Matzenbacher, C. A., Ferrari, F., Alievi, M. M., & de Freitas, T. R. O. (2023). Hybridization in canids—A case study of Pampas fox (*Lycalopex gymnocercus*) and domestic dog (*Canis lupus familiaris*) hybrid. Animals, 13(15), 2505.

Tancredi, D., & Cardinali, I. (2023). Being a dog: a review of the domestication process. Genes, 14(5), 992.

Thoma, V., Sakai, S., Nagata, K., Ishii, Y., Maruyama, S., Abe, A., ... & Tanimoto, H. (2023). On the origin of appetite: GLWamide in jellyfish represents an ancestral satiety neuropeptide. Proceedings of the National Academy of Sciences, 120(15), e2221493120.

Todesco, M., Pascual, M. A., Owens, G. L., Ostevik, K. L., Moyers, B. T., Hübner, S., ... & Rieseberg, L. H. (2016). Hybridization and extinction. Evolutionary Applications, 9(7), 892-908.

Torretta, E., Dondina, O., Delfoco, C., Riboldi, L., Orioli, V., Lapini, L., & Meriggi, A. (2020). First assessment of habitat suitability and connectivity for the golden jackal in north-eastern Italy. Mammalian Biology, 100, 631-643.

Tyagi, A., Godbole, M., Vanak, A. T., & Ramakrishnan, U. (2023). Citizen science facilitates first ever genetic detection of wolf-dog hybridization in Indian savannahs. Ecology and Evolution, 13(5).

Vestergaard, A. L., Coleman, J. A., Lemmin, T., Mikkelsen, S. A., Molday, L. L., Vilsen, B., ... & Andersen, J. P. (2014). Critical roles of isoleucine-364 and adjacent residues in a hydrophobic gate control of phospholipid transport by the mammalian P4-ATPase ATP8A2. Proceedings of the National Academy of Sciences, 111(14), E1334-E1343.

Vilà, C., Maldonado, J. E., & Wayne, R. K. (1999). Phylogenetic relationships, evolution, and genetic diversity of the domestic dog. Journal of Heredity, 90(1), 71-77.

Wang, G. D., Fan, R. X., Zhai, W., Liu, F., Wang, L., Zhong, L., ... & Zhang, Y. P. (2014). Genetic convergence in the adaptation of dogs and humans to the high-altitude environment of the Tibetan plateau. Genome Biology and Evolution, 6(8), 2122-2128.

Wang, J. Z., Wang, Y., & Jiang, P. (2015). The study and application of a novel hybrid forecasting model–A case study of wind speed forecasting in China. Applied Energy, 143, 472-488.

Wang, M. S., Wang, S., Li, Y., Jhala, Y., Thakur, M., Otecko, N. O., ... & Wu, D. D. (2020). Ancient hybridization with an unknown population facilitated high-altitude adaptation of canids. Molecular Biology and Evolution, 37(9), 2616-2629.

Wang, S., Tan, N., Zhu, X., Yao, M., Wang, Y., Zhang, X., & Xu, Z. (2018). Sh3rf2 haploinsufficiency leads to unilateral neuronal development deficits and autistic-like behaviors in mice. Cell Reports, 25(11), 2963-2971.

Wang, X., Fang, L., Xiao, L., Zhong, G., Han, M., Wang, B., ... & Zang, Y. (2024). Research on the effect of LAMP1 in the development and progression of ccRCC and its potential mechanism with LC3C-mediated autophagy. Frontiers in Immunology, 15, 1494005.

Westbrook, A. S., & DiTommaso, A. (2023). Hybridization in agricultural weeds: a review from ecological, evolutionary, and management perspectives. American Journal of Botany, 110(12), e16258.

Wilkins, A. S., Wrangham, R. W., & Fitch, W. T. (2014). The “domestication syndrome” in mammals: a unified explanation based on neural crest cell behavior and genetics. Genetics, 197(3), 795-808.

Wong-Benito, V., de Rijke, J., & Dixon, B. (2023). Antigen presentation in vertebrates: structural and functional aspects. Developmental & Comparative Immunology, 144, 104702.

Xu, S., Wang, Y., Wang, Y., Jiang, Y., Li, H., Han, C., ... & Wei, S. (2024). Development and immune evaluation of LAMP1 chimeric DNA vaccine against Singapore grouper iridovirus in orange-spotted grouper, Epinephelus coioides. Fish & Shellfish Immunology, 144, 109218.

Yadav, A., Jain, A., Sahu, J., Dubey, A., Gadpayle, R., Barwa, D. K., ... & Kashyap, K. (2019). An overview on species hybridization in animals. International Journal of Fauna and Biological Studies, 6(5), 36-42.

Yang, J., Jin, Z. B., Chen, J., Huang, X. F., Li, X. M., Liang, Y. B., ... & Qu, J. (2017). Genetic signatures of high-altitude adaptation in Tibetans. Proceedings of the National Academy of Sciences, 114(16), 4189-4194.

Zhao, S. C., Zhou, B. W., Luo, F., Mao, X., & Lu, Y. J. (2015). The structure and function of NKAIN2-a candidate tumor suppressor. International Journal of Clinical and Experimental Medicine, 8(10), 17072.

4 Chapter 4

The Role of Hybridization Between Wild Canids and Domestic Dogs in Adaptation to Environmental Change

Abstract

Human activities and climate changealter the species distribution ranges , which may bring closely related species with incomplete reproductive isolation into closer proximity, thereby facilitating hybridization. This study investigated how environmental variables contribute to the presence of dog ancestry in grey wolves and golden jackals across Eurasia. Random Forest (RF) models were used to investigate the influence of environmental variables on dog ancestry proportions in grey wolves and golden jackals. The environmental variables considered included climatic variables, topography, vegetation and land cover, and anthropogenic factors such as human footprint. Moreover, the association between environmental variables and adaptive introgression in admixed wolves, dogs, and golden jackals was determined using Redundancy Analysis (RDA). RF models showed that, milder winters were associated with increased dog ancestry in wolves, likely due to higher free-ranging dog densities in warmer regions. In contrast, lower annual temperatures correlated with increased dog ancestry in jackals, likely due to their ongoing northward range expansion of this species, with hybridization occurring more frequently at the expansion front. Additionally, during harsh winters, golden jackals tend to move closer to human settlements in search of food, which can increase the likelihood of contact with dogs. In both wolves and jackals, human footprint was positively associated with dog ancestry proportions, probably due to the large populations of free-ranging dogs and also land use change in these regions. The results of RDA showed a significant association between environmental variables and adaptive introgressed loci in admixed wolves with excess dog ancestry. No significant environmental associations were found for introgressed loci in admixed dogs and admixed jackals, likely due to their greater ecological flexibility. Significant associations between dog-derived loci in wolves and environmental variables suggest that these loci may play a role in local adaptation in human-modified habitats. These loci were linked in genes related to neural developmental processes (DOCK1), anti-inflammatory and neuroprotective activity (GAS6), and metabolism (LPCAT3). These genetic adaptations may help admixed wolves cope with new pathogens, environmental stressors, and dietary shifts associated with anthropogenic influences. The findings from this study emphasize the role of environmental variables, especially the human footprint, in increasing the dog introgression proportions in wild canids. Moreover, we showed how adaptive introgression from free-ranging dogs to wolves may facilitate wolf adaptation to changing environmental conditions.

Keywords: Hybridization, Genus *Canis*, environmental variables, human footprint, adaptative introgression

4.1. Introduction

The geographic distribution of species is influenced by diverse consequences of human activities such as land use change, introduction of invasive species, and climate change (Davison et al., 2021; Bellard et al., 2012). For example, human settlements, along with the expansion of croplands and industries, may lead to range expansion, contraction, or shift (Xia et al., 2023), which may bring ecologically segregated species into closer proximity, facilitating hybridization between them (Scheffers et al., 2016; Brennan et al., 2015; Chunco, 2014; Ottenburghs, 2021). The impacts of human activities on hybridization have long been acknowledged (e.g., Anderson, 1948). By creating new contact zones, climate change can further increase hybridization between previously isolated species (Brennan et al., 2015). The impact of climate change on hybridization in insect species showed that climate-induced range shifts can lead to the breakdown of isolation barriers, and thus, to an increase in hybridization frequency (Arce-Valdés and Sánchez-Guillén, 2022). In another example, climate change led to an increased hybridization rate between willow grouse (*Lagopus lagopus*) and rock ptarmigan (*L. muta*) in Scandinavia due to an increase in the range overlap between these species (Quintela et al., 2010). Moreover, climate change, by increasing habitat disturbance, can facilitate the establishment of invasive species, which in turn provides opportunities for hybridization with closely related native species (Chown et al., 2016). For instance, rapid climate warming has resulted in exacerbated interactions between native westslope cutthroat trout (*Oncorhynchus clarkii lewisi*) and non-native rainbow trout (*O. mykiss*) (Muhlfeld et al., 2014). Studies on plants showed that hybridization can affect and be affected by organism-environment interactions (Porretta and Canestrelli, 2023). For example, wildfires, as an environmental disturbance, play a significant role in increasing hybridization between two ecologically well-differentiated Californian oak species (Ortego et al., 2017). Although in some species the role of environmental variables on hybridization has been studied, there is a scarcity of knowledge on the role of environmental factors on hybridization. Understanding geographical patterns of hybridization is also important in order to identify where hybridization occurs more frequently and what environmental factors might drive it (Zbinden et al., 2023).

Most species from the genus *Canis* have wide geographic ranges (Kurtén and Anderson, 1980). The gray wolf (*Canis lupus*) and golden jackal (*Canis aureus*) are generalist carnivores with adaptable, opportunistic lifestyles, widely distributed across nearly entire Eurasia and southern Eurasia, respectively (Lanszki et al., 2016; Shakarashvili et al., 2020). Wolves exhibit high ecological resilience to habitat fragmentation and are capable of adapting to landscape modifications and persistent human presence (Chapron et al., 2014). While wolves generally avoid areas with high human density and landscapes dominated by croplands (Simpson et al., 2023; Kudrenko et al., 2023), in some areas, they display less avoidance of human settlements. As facultative scavengers, wolves can exploit a variety of anthropogenic food sources, such as livestock carcasses and garbage (Newsome et al., 2016; Mohammadi et al., 2019). This dietary flexibility increases their potential to persist in highly human-dominated landscapes (e.g., Kuijper et al., 2019; Mohammadi, et al., 2021). In these areas, due to

the high density of free-ranging dogs, hybridization between wolves and dogs has been reported frequently (Blanco et al., 1992).

The golden jackal's range was historically confined to southeastern Eurasia (Arnold et al., 2012). However, it has been rapidly expanding into northern and western Europe (Kowalczyk et al., 2020), driven by factors such as climate change (Serva et al., 2023), land-use changes (Šálek et al., 2014; Trouwborst et al., 2015), and adaptation to human-dominated environments (Fenton et al., 2021). This range expansion has led to potential overlap with other canid species and increased the rate of hybridization due to the low population density of expanding species (Nussberger et al., 2018). The occurrence of hybridization at the edge of the species' range has been reported for golden jackals (Stefanovic et al., 2024) and coyotes, which have experienced a rapid expansion in North America (Kays et al., 2010; Bohling et al., 2016). Low population densities of wolves and range expansion of jackals are identified as two factors that may increase the likelihood of hybridization with free-ranging dogs (Randi et al., 2000; Anderson et al., 2002; Lorenzini et al., 2014). Moreover, human disturbance is usually considered a main factor that may facilitate hybridization between wild canids and free-ranging dogs (e.g., Godinho et al., 2011; Lescureux and Linnell, 2014; Donfrancesco et al., 2019; Pilot et al., 2021). However, the effect of human disturbance on hybridisation rates has never been tested explicitly, and direct empirical evidence confirming this effect. In this chapter, we address this knowledge gap by testing the hypothesis that introgression rate in admixed individuals are associated with environmental variables. We aim to (1) identify the key environmental factors that may be associated with hybrid ancestry in wild canids, (2) detect adaptive introgressed loci associated with environmental variables, and (3) identify genes within adaptive introgressed regions that are linked to environmental factors.

4.2. Methods

Sampling

Gray wolves, golden jackals, and free-ranging dogs were sampled from their distribution range in Eurasia between 2018 and 2022. Moreover, frozen tissue samples from the previous study (Rutkowski et al., 2015) were included. Based on the type of samples, the DNA of tissue samples was extracted using NucleoSpin Tissue Kit (Macherey Nagel, Duren, Germany), and the PG-AC extraction kit was used for the saliva samples (PERFORMAgene).

Dataset creation

Two different datasets were prepared: (1) gray wolves and free-ranging dogs (WD dataset), and (2) golden jackals and free-ranging dogs (JD dataset). A detailed description of the steps (e.g. sampling, genotyping, and data processing) involved in generating these datasets can be found in Chapter 2 (see methods).

Admixed detection using ELAI

For estimating each sample's ancestries in the WD dataset and JD dataset, we used the results of ELAI v. 1.01 (Guan, 2014) from the previous chapters (For more details,

please see Chapter 3, methods). Based on the results of Chapter 3, in the WD dataset, 115 admixed wolves and 44 admixed dogs were identified. The rest of the samples (115 samples) were considered pure individuals. The mean proportion of wolf ancestry in dogs and dog ancestry in wolves was estimated at 0.017 and 0.065, respectively (Fig 4.1). In the JD dataset, we identified 48 golden jackal samples that showed less than 98% golden jackal ancestry and were considered admixed (Table S 4.2). All dog samples showed more than 98% dog ancestry (Fig 4.1). The mean proportion of golden jackal ancestry in dogs and dog ancestry in golden jackals was estimated at 0.001 and 0.041, respectively.

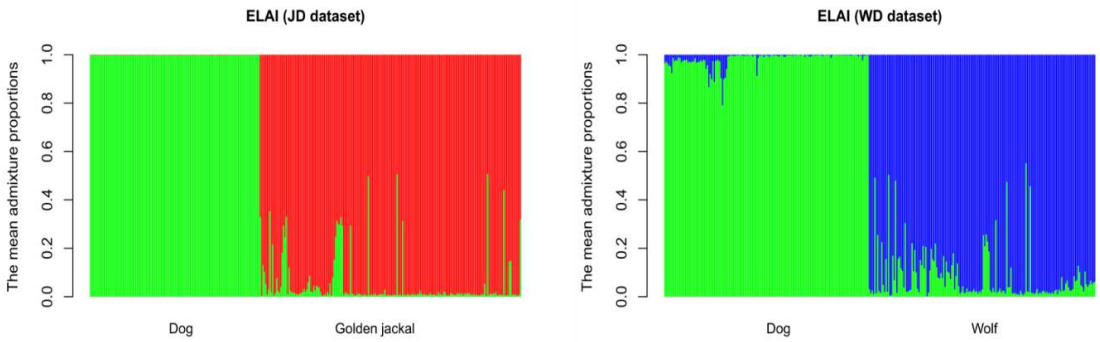


Fig. 4.1. The bar plot of the ELAI result using the WD and the JD datasets.

Environmental variables

Based on the literature review (Šálek et al., 2014; Wennink et al., 2019; Torretta et al., 2020; Ordiz et al., 2020; Silva et al., 2022; Dar et al., 2023; Bosch et al., 2023), 32 candidate variables related to the presence and habitat suitability of gray wolves, golden jackals, and free-ranging dogs were selected. All selected variables were categorized into four main categories: climate (CLM), topography (TOP), vegetation and land cover (VLC), and anthropogenic factors (ANT) (Table 2). CLM category consisted of 19 bioclimatic variables obtained from WorldClim v2 (Fick and Hijmans 2017) at ca. 1 km spatial resolution. Digital elevation model (DEM) of the study area was obtained from WorldClim v2 (Fick and Hijmans 2017) at ca. 1 km spatial resolution to extract three topographical predictors including the mean of elevation, standard deviation of slope, and topographic roughness using raster package v3.5-15 (Hijmans et al., 2022) in R version 4.1.2 (R Core Team, 2021). Vegetation and land cover (VLC) variables included open forest, closed forest, shrubland, herbaceous, and sparse vegetation, all of them at 100 m resolution (Buchhorn et al., 2020). They were downloaded from the Copernicus Global Land Service (https://lcviewer.vito.be/download). The diversity of land cover classes was calculated in FRAGSTATS (McGarigal et al., 2002) to examine how the degree of fragmentation across the study area may affect the hybridization and adaptive introgression among species. In addition to landcover, Normalized Difference Vegetation Index (NDVI) at 300 m (https://land.copernicus.eu/global/products/ndvi) was acquired. To assess the impact of anthropogenic factors on the spatial patterns of observed genetic variation, we downloaded and averaged human footprint data layers from 2015 and 2020 (https://wcshumanfootprint.org). Human Footprint index (HF) represents human influence by integrating the effects of population density, land use

change, road and railway density, and power infrastructure (Sanderson et al., 2002). The mean human footprint index was calculated for a 1 km grid to represent the level of human influence. We also calculated the proportion of croplands as another anthropogenic predictor.

All spatial layers were standardized to the same coordinate reference system (WGS84-EPSG:4326) and resampled to approximately 1 km resolution using the raster package (v3.5-15; Hijmans, 2022) in R version 4.1.2 (R Core Team, 2021). Environmental variable values for each individual were extracted by averaging cell values within a 2.5 km radius around the individual's coordinate location, utilizing the R package geobuffer (Ștefan, 2019). The home range of gray wolves, golden jackals, and free-ranging dogs was different, and also the home range of wolves and golden jackals can be influenced by the level of human disturbance in different regions. Therefore, we used a cell size of 20 km² as it represents an intermediate estimate of home range size in these species and is supported by previous studies (e.g. Vorel et al., 2024; Kamler et al., 2021; Wilson-Aggarwal et al., 2021; Meek, 1999).

Environmental factors affecting dog ancestry proportions

To investigate the influence of environmental variables on dog ancestry proportions in wolves and jackals, random forests (RF) models (Evans & Murphy, 2019; Franklin, 2009; Cutler et al., 2007; Breiman, 2001) were applied. The estimated dog ancestry proportions from ELAI were treated as dependent variables, while the environmental variables were considered independent variables. First, all environmental variables were tested for multicollinearity using the rfUtilities v 2.1-5 package (Evans & Murphy, 2019) as implemented in R (R Core Team, 2021). For each dataset (WD and JD), variables exhibiting multicollinearity above the threshold of 10^{-7} were removed and the association between ancestry values and environmental variables was assessed using RF models (regression mode). We used Random Forest (RF) models using the randomForest package (v. 4.7-1.1; Liaw, 2002) alongside rfUtilities (v. 2.1-5; Evans & Murphy, 2019) in R. Each model was configured with 1,000 bootstrap iterations (trees) and employed the permutation-based importance metric ("mir") to evaluate the contribution of each predictor variable. The final model was chosen based on its ability to explain the most variance, minimize the Root Mean Squared Error (RMSE), and balance simplicity with ecological relevance by limiting the number of predictors. To ensure the robustness of the selected models, we assessed fit statistics, including checks for overfitting, using the rf.regression.fit function. Once the best-performing models were identified for each transect, we conducted separate linear regression analyses in R (R Core Team, 2021) to determine the relationships' significance, direction, and strength.

To explore the relationship between hybrid ancestry and environmental variables across different regions, samples were categorized into six groups based on their distribution: North Asia, East Asia, the Caucasus, Central Europe, Northeast Europe, and Western Europe. RF model was applied for each region separately using the parameters described above.

Environmental association of adaptive introgressed regions

The association between environmental variables and adaptive introgressed regions was determined using Redundancy Analysis (RDA). RDA is a multivariate constrained ordination method that can combines multiple linear regression with PCA to identify loci significantly associated with environmental variables (Capblancq and Forester, 2021; Legendre and Legendre, 2012). Since RDA is a suitable method for identifying signatures of selection and adaptation (Forester et al., 2018; Rellstab et al., 2015), and also has a low rate of false positives (Capblancq et al., 2018), we used it to identify candidate adaptive loci that are associated with environmental variables. For this purpose, we used the chromosomal blocks with an overrepresentation of introgressed variants that were identified in the previous chapter (for more details, please see Chapter 3, Methods). SNPs located within these blocks are considered as strong candidates for adaptive loci (hereafter called CAI loci), that are located within or near genes putatively under selection (please see Chapter 3, Methods). The association between the CAI loci as the response variables and the environmental variables as the predictor variables (the same set of environmental variables with no detected collinearity as in the Random Forest analysis) was assessed using a partial RDA conditioned on the dbMEM spatial matrix using the vegan R package (Oksanen et al., 2019). The anova.cca function with a permutation test of 999 iterations and a p-value threshold of 0.05 was used to identify significant RDA axes. SNPs from these significant RDAs were extracted using a three-standard-deviation (two-tailed p-value = 0.0027) from the mean loading of these significant RDA axes. Each candidate adaptive loci was then associated with all environmental predictors in each dataset using Pearson's correlation coefficient (r) (Forester et al., 2018). SNPs with the strongest correlations were considered the best-supported SNP-environment associations. CAI SNPs that were associated with environmental variables were lifted over to the 10K Boxer Tascha genome (CanFam6) using the liftover tool in the UCSC Genome Browser (https://genome.ucsc.edu/cgi-bin/hgLiftOver) and annotated using the Ensembl Variant Effect Predictor tool (https://www.ensembl.org/info/docs/tools/vep/index.html). Finally, the function of each gene was evaluated to test whether the function may explain the environmental association.

4.3. Results

Environmental relationships with hybridization

The association between environmental variables and dog ancestry was estimated in admixed wolves and jackals using RF models. In wolves and jackals, 22 and 18 environmental variables were removed due to high correlation, respectively (Table 4.2, Fig 4.2). In wolf-dog dataset, five variables have been selected as a top-ranked RF model (Temperature Seasonality (Bio4), Min Temperature of Coldest Month (Bio6), Annual precipitation (Bio12), Precipitation of wettest month (Bio13) and human footprint index), that explained together 68.22% of the variation in dog ancestry proportions (Fig 4.2, Fig 4.3, Table 4.2). In jackal-dog dataset, four variables have been selected as a top-ranked RF model for the JD dataset, including Annual Mean Temperature (Bio1), Precipitation

Seasonality (Bio15), elevation, and human footprint index (Table 4.2, Fig 4.4), which explained together 77.48% of the variation in dog ancestry. For both datasets, a root-squared-mean and R^2 were calculated and testing the model fit showed that the model is not overfit (Table 4.1). After running a linear regression, dog ancestry in admixed wolves showed positive association with the minimum temperature of coldest month (Bio6) and human footprint, while in admixed jackals dog ancestry showed positive association with precipitation seasonality (Bio15), elevation and human footprint (Table 4.2). The RF results for each population are provided in the supplementary file (Table S 4.1 – S 4.13).

Table 4.1. The estimated root-mean-squared (RSME) and accuracy for each dataset.

Datasets	RSME*	R^2
Wolf-dog	0.276	0.682
Jackal-dog	0.242	0.774

*RSME reflects the average deviation of the predicted values from the observed values.

Table 4.2. The list of Environmental variables used in the analysis. Variables marked with (✓) were included in the Random Forest (RF) models, while variables marked with (×) were excluded due to high collinearity. Variables highlighted in bold (✓) represent those included in the top-ranked model. (+) indicate a positive association between hybrid ancestry and variable value, whereas (-) indicates a negative association

Predictors	**Group**	**Abbreviation**	**Final model (WD dataset)**	**Final model (JD dataset)**
Annual mean temperature	CLM	BIO1	×	✓ −
Mean diurnal range	CLM	BIO2	×	×
Isothermality	CLM	BIO3	×	×
Temperature seasonality	CLM	BIO4	✓ −*	✓
Max temperature of warmest month	CLM	BIO5	✓	×
Min temperature of coldest month	CLM	BIO6	✓ +	×
Temperature annual range	CLM	BIO7	×	×
Mean temperature of wettest	CLM	BIO8	×	×
Mean temperature of driest quarter	CLM	BIO9	×	×
Mean temperature of warmest quarter	CLM	BIO10	×	×
Mean temperature of coldest quarter	CLM	BIO11	×	×
Annual precipitation	CLM	BIO12	✓ −	✓
Precipitation of wettest month	CLM	BIO13	✓ −	✓
Precipitation of driest month	CLM	BIO14	×	×
Precipitation seasonality	CLM	BIO15	×	✓ +
Precipitation of wettest quarter	CLM	BIO16	×	×
Precipitation of driest quarter	CLM	BIO17	×	✓
Precipitation of warmest quarter	CLM	BIO18	✓	✓
Precipitation of coldest quarter	CLM	BIO19	✓	✓
Standard deviation of slope within a focal cell	TOP	SLO	×	×
Mean of elevation within a focal cell	TOP	ELV	✓	✓ +
Topographic roughness	TOP	ROU	✓	✓
Proportion of cell with > 10 within a focal cell	TOP	ESC	×	×
Proportion of open forests within a focal cell	VLC	OFR	×	×
Proportion of closed forests within a focal cell	VLC	CFR	×	×
Proportion of shrublands within a focal cell	VLC	SHR	×	×
Proportion of grasses within a focal cell	VLC	GRS	×	×
Proportion of sparse vegetation within a focal cell	VLC	SPR	×	×
Normalized difference vegetation index	VLC	NDV	×	×

Shannon' diversity of land cover types within a focal cell	VLC	SHD	✗	✓
Human footprint	ANT	HFI	✓ +	✓ +
Proportion of croplands within a focal cell	ANT	CRP	✗	✗

*The linear regression for Bio4 in the wolf-dog dataset was non-significant (P-value > 0.05)

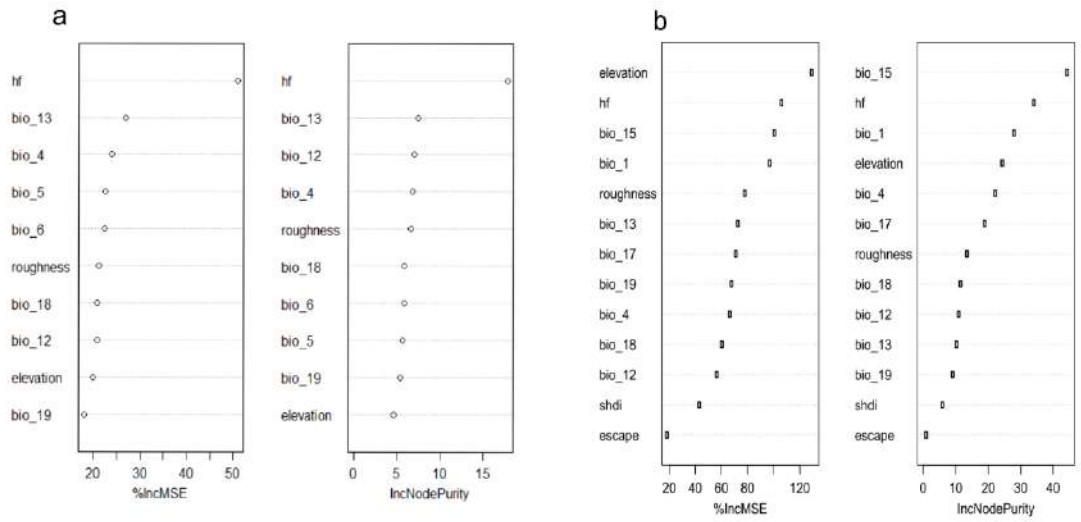


Fig 4.2. The importance measure for each variable of the WD dataset (a) and in the JD dataset (b) according to %IncMSE (Mean square-error) and IncNodePurity (node purity). %IncMSE reflects how much each feature contributes to reducing prediction error across the entire model, while IncNodePurity measures how much each feature improves node purity at each split (Liaw and Wiener, 2002).

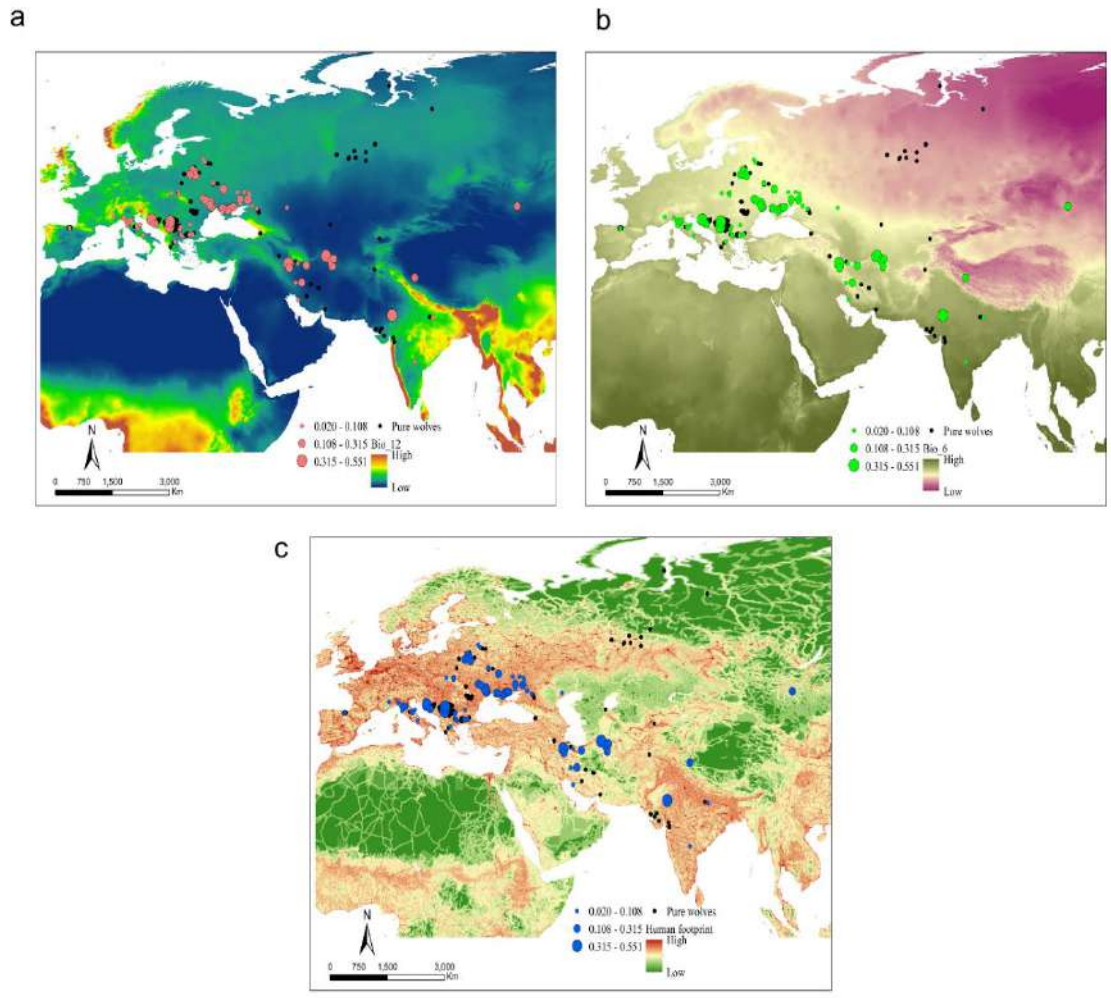


Fig. 4.3. The distribution of pure wolves (black dots) and admixed wolves (colored circles) that showed more than 2% dog ancestry on a map of (a) Annual precipitation (Bio12), (b) Minimum temperature of coldest month (Bio6), and (c) Human footprint. The size of each colored circle represents the proportion of dog ancestry.

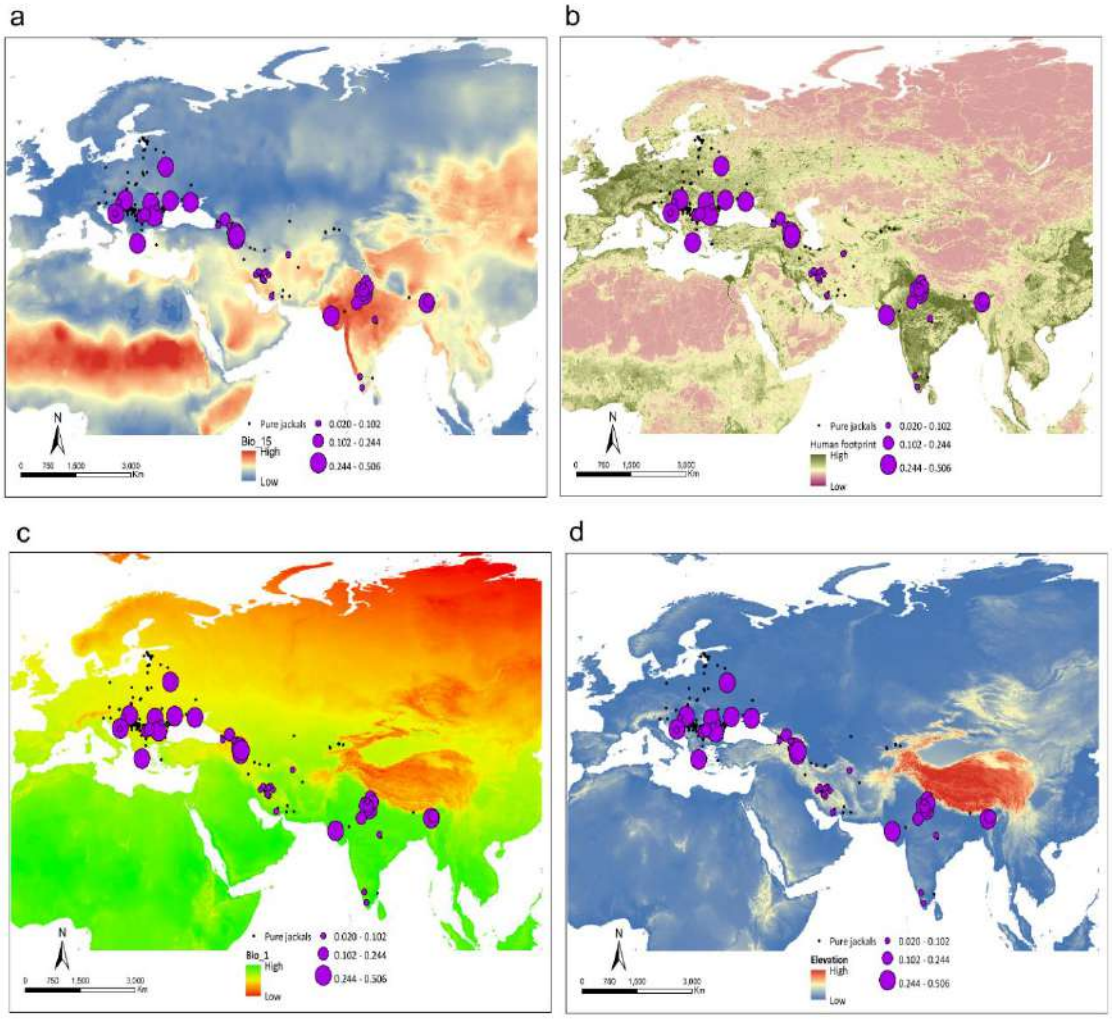


Fig. 4.4. The distribution of pure jackals (black dots), and admixed jackals (purple circles) showed more than 2% dog ancestry on a map of (a) Precipitation seasonality (bio_15), (b) Human footprint, (c) Annual mean temperature (Bio_1), and (d) Elevation. The size of each purple circle represented the proportion of dog ancestry.

Environmental association of adaptive introgressed regions

In the wolf-dog dataset, a total of 848 and 3,228 CAI loci were identified that had excess introgressed ancestry in wolves and dogs, respectively (Table S 3.1, Chapter 3). In the jackal-dog dataset, a total of 453 SNPs were identified as CAI loci that had excess dog ancestry (Table S 3.2, Chapter 3). Since all dog samples displayed less than 1% of jackal ancestry, chromosomal blocks with an overrepresentation of jackal variants weren't calculated for dog samples (for more details, see Chapter 3, Results).

In admixed wolves, two significant RDA axes ($p < 0.05$) explained 23.52% and 8.62% of the total variance. The first axis was primarily driven by bio12, bio19, and human footprint, whereas the second was associated with elevation, bio18, and roughness (Fig 4.5, Table 4.4). Seven candidate loci were associated with these environmental variables (Table 4.3).

Four of the seven candidate loci were mapped within 100kb of protein-coding genes. 14 genes located within 100 kb of these candidate loci were identified as candidate genes

for adaptive introgression in wolves, with potential associations with environmental variables. In cases where more than one gene was identified for one SNP, the nearest gene was considered (Table 4.5). The functions of these genes suggest potential roles in immune response, metabolic adaptation, and behavioral or neurological processes, which may explain their association with environmental variables and adaptive advantages in admixed wolves.

No significant RDA was identified in admixed dogs using CAI loci that showed an excess of wolf ancestry and admixed jackals using CAI loci that showed an excess of dog ancestry (Table 4.3). This result showed that adaptive introgression in hybrid dogs and golden jackals is probably not associated with environmental variability, or the sample sizes were too small to provide significant results.

Table 4.3. Regression coefficients from the overall RDAs for each dataset

	Loci set	Adjusted r^2	Overall RDA p-value	RDA1 p-value	The candidate's adaptive loci associate with environmental variables
Admixed wolves (N=114)	CAI loci with excess dog	0.109	0.001	0.001	chr28:35445148 chr28:36073423 chr13:25551106 chr13:25631078 chr13:25968426 chr22:61094736 chr27:38027433
Admixed dogs (N=44)	CAI loci with excess wolf ancestry of wolves	0.015	0.036	0.184	0 loci
Admixed jackals (N=48)	CAI loci with excess dog ancestry of dogs	0.044	0.162	0.166	0 loci

Table 4.4. Environmental variables values on the first and second RDA axes in admixed wolves.

	bio12	**bio13**	**bio18**	**bio19**	**bio4**	**bio5**	**bio6**	**elevation**	**roughness**	**hf**
RDA1	0.720	0.241	0.082	0.688	-0.63	0.290	0.613	-0.121	0.056	0.653
RDA2	0.241	-0.251	0.502	0.306	0.184	-0.346	-0.10	-0.641	-0.376	-0.169

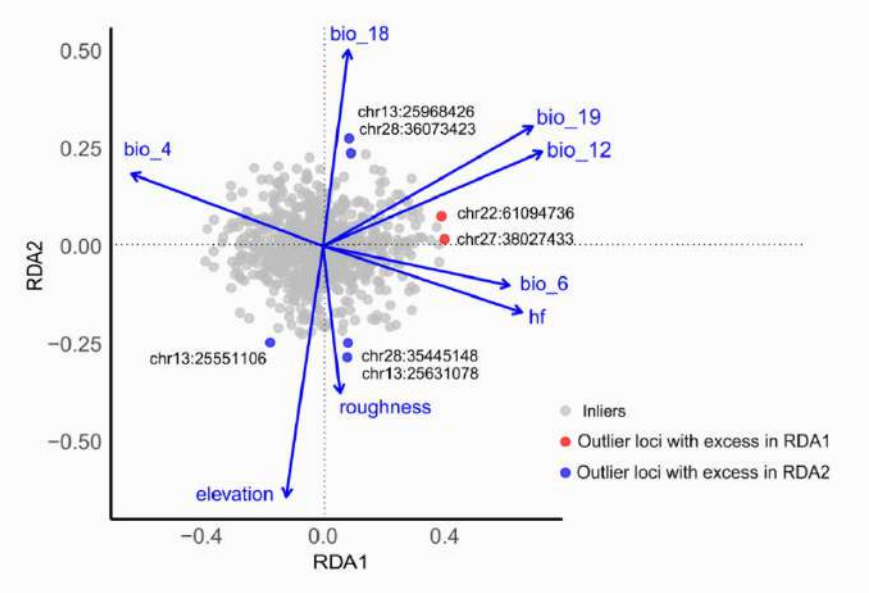


Fig. 4.5. Significant CAI loci that were associated with environmental variables based on the RDA in admixed wolves. Red circles indicate loci that were identified by the first RDA, and the blue circle indicates the loci identified by the second RDA. The length of the arrow indicates the strength of the correlation between loci and environmental variables.

Table 4.5. Candidate genes under adaptive introgression in wolves that are related to environmental variables. The nearest gene from each outlier SNP within a distance of 100 kb was considered.

CHR	Gene	Gene name	Distance from SNP	Environmental variables
22	GAS6	Growth Arrest Specific 6	88974	bio12, bio19
27	LPCAT3	lysophosphatidylcholine acyltransferase 3	687	bio 6, hf
28	DOCK1	Dedicator of cytokinesis protein 1	0	bio 12, bio19, bio 18
28	C28H10orf90	Chromosome 10 Open Reading Frame 90	69857	Roughness, bio6, hf

4.4. Discussion

The effect of environmental factors on dog admixture proportions in wild canids

Our results showed that environmental variables such as climate, anthropogenic factors, and topographic features can influence hybridization in canids. In both admixed wolves and admixed dogs, the dog ancestry proportion was associated with different climatic variables. For example, dog ancestry in wolves was associated with the minimum temperature of coldest month, annual precipitation, and temperature seasonality, while in jackals, it was associated with annual mean temperature, precipitation seasonality, and elevation. Gray wolves and jackals have different habitat suitability (e.g. Krofel et al., 2017; Garcia-Lozano et al., 2020; Bosch et al., 2022; Torretta et al., 2022; Rezaei et al., 2022; Serva et al., 2023), which may explain why

different environmental conditions affect the hybrid ancestry in these two species. Species may respond uniquely to environmental pressures, potentially due to differences in their habitat suitability, ecological preferences, dispersal capabilities, or behavioural adaptations

In wolves, we found positive associations between the minimum temperature of coldest month (Bio6) and dog ancestry. The Bio6 was derived from a multi-decade average of the minima of the coldest months (Hijmans et al., 2005). In southern regions of Eurasia (e.g., southern Europe, South Asia, the Middle East, and Southeast Asia), Bio6 values are higher, meaning milder winters. In these regions, free-ranging dogs are more abundant than in the north of Eurasia (Gomper, 2014). For example, it has been estimated that around 62 million dogs are living on the streets in India (Sensharma, et al., 2024). The positive association between Bio6 and dog ancestry rates may be explained by the fact that regions with higher Bio6 values likely support larger free-ranging dog populations, increasing the chances of contact and interbreeding between dogs and wolves. Conversely, in colder northern regions, where free-ranging dogs are scarce, hybridization opportunities are more limited.

However, the contrasting pattern is observed in golden jackals, since our result showed a negative association between hybridization and annual temperature. This indicates that in regions with lower temperatures, dog ancestry in golden jackals increases. Food availability is one of the main essential factors to shape the diet compositions of predators, and is highly related to environmental and climate conditions (Zhou et al., 2011; Soe et al., 2017; McCain et al., 2018). In golden jackals, the consumption of small mammals, livestock, and other food categories is predominantly related to annual temperature (Lanszki et al., 2022). Regions with lower annual temperatures, have lower abundance of some types of food like small mammals (e.g. rodents and voles) (Korslund and Steen, 2006), which may force the golden jackal as an opportunistic, omnivorous carnivore to approach human settlements to consume domestic animals and anthropogenic food (Lanszki et al., 2022). Golden jackals approaching human settlements to find food can increase the chance of encounters with domestic dogs and increase the hybridization rate between them. Moreover, as jackals have recently expanded northward (Stefanovic et al., 2024), their densities at the edges of the expanding range remain low, which may increase the likelihood of hybridization with dogs. Consequently, the observed correlation with temperature could be linked to the direction of this expansion (Stefanovic et al., 2024). Therefore, climatic variables can indirectly increase hybridization rates by imposing changes in food availability, dispersal ability, and creating barriers.

For both wolves and jackals, our findings reveal a positive correlation between the occurrence of hybridization and human footprint, suggesting that human activities can play a significant role in facilitating hybridization and increasing dog ancestry in wild canids. Regions with high human footprint often have larger densities of free-ranging dogs, increasing the chances of encounters with wild canids and potentially promoting hybridization (Pilot et al., 2021). Additionally, land use and land cover changes can occur in response to human activities, which may cause habitat destruction and fragmentation (Gounaridis et al., 2020). The natural ecosystem services, like availability of food and

water, can be disrupted due to anthropogenic habitat modification (Mmbaga et al., 2017). These kinds of disruptions may alter the distribution of many species and force them to move beyond their natural habitats and even approach human settlements in order to obtain food and/or avoid the intra-specific spatial competition (Scanes, 2018; Zanni et al., 2023). Large carnivores are most sensitive to land use changes due to large home range requirements and slow reproductive rate (Enserink and Vogel, 2006). Wolves tend to avoid regions with dense human settlements (Zanni et al., 2023; Simpson et al., 2023; Kudrenko et al., 2023), however, they show a plastic response to human footprint (Muhly et al., 2019) and can persist in these kinds of regions by adapting their behaviors to reduce the probability of direct encounters with humans (Llaneza et al., 2016; Kojola et al., 2016). Due to the low density of natural prey, wild canids as facultative scavengers may approach human settlements where garbage and livestock carcasses are available (Rotem et al. 2011; Newsome et al., 2016; Kapota et al. 2016; Mohammadi et al., 2019; Kamler et al. 2021). Easier access to anthropogenic food resources near human settlements, along with the availability of vacant territories, may attract wild canids to these regions. In areas where free-ranging dogs are abundant, this can increase the likelihood of encounters between wild canids and domestic dogs.

Besides habitat destruction, habitat fragmentation can also negatively impact the natural habitat of species (Fletcher et al., 2018). Fragmentation modifies the habitats and divides them into smaller habitat patches, creating habitat edges (Didham and Ewers, 2012; Lamb et al., 2016; Püttker et al., 2020). Increasing edge density can providing greater availability of resources, such as food, which in turn can impact the distribution and abundance of species, particularly opportunistic and generalist species. This, in turn, may influence interspecific interactions (Pereira et al., 2024), potentially increasing hybridisation rates.

Environmental associations of adaptive introgressed regions

Adaptive introgression can facilitate species' evolutionary responses to environmental changes by enabling rapid acquisition of new adaptive genetic variants (Adavoudi and Pilot, 2021). For example, hybridization between montane sedge species was associated with environmental shifts that caused hybrid individuals to occupy an environment that is largely unsuitable for the parental species (Hondel et al., 2022). Studies on hybridization in canids have suggested that introgression from dogs into the gene pool of wild canids might help them to better adapt to human-modified landscapes (Newsome et al. 2017; Pilot et al. 2021).

Moreover, investigating adaptive introgression in the Iberian wolf genome showed that MAST4 gene introgression from dogs may enhance cognitive traits in wolves and have facilitated adaptation to human-dominated landscapes (Lobo et al., 2025). The significant association between environmental variables, including the human footprint, and dog-derived adaptive variation introgressed into wolves identified in this study can validate the hypothesis that adaptive introgression plays a crucial role in wolf adaptation to human-modified habitats.

Based on the RDA results, the significant association between environmental variables and excessively introgressed loci was found in admixed wolves, while no

significant associations between environmental variables and CAI loci were observed in hybrid dogs and jackals. We found that the CAI loci in wolves were strongly correlated with topographic variables (elevation and roughness), climate variables (precipitation), and human disturbance (human footprint), which suggests that adaptive introgression from dogs may have facilitated the presence of admixed wolves in some specific regions where human footprint is high. In admixed wolves, adaptive introgression were found in genes were involved in neural developmental processes like DOCK1 (Shi, 2013), anti-inflammatory and neuroprotective like GAS6 (Kim et al., 2019), and metabolic functions and lipid processing such as LPCAT3 (Shi et al., 2023; Wang et al., 2024). These genes may play a role in the adaptation of admixed wolves in habitats where they may encounter new immunity challenges, pathogens, or environmental stressors. Studies showed that DOCK1 plays a conserved evolutionary role in Schwann cells (Cunningham et al., 2018; Doan and Monk, 2025), which highlights its fundamental role in the nervous system. Adaptive introgression of the DOCK gene family (DOCK3), has been reported in the genus *Panthera*, where introgressed variants in jaguars have been linked to axon development in the optic nerve (Figueiró et al., 2017). The adaptive introgression LPCAT3 gene from dogs may be linked to dietary shifts in admixed wolves, as they may rely more on anthropogenic food sources. Also adaptive introgression of genes relate to the olfactory receptor may increase

These findings suggest that adaptive introgression would be particularly beneficial in human-modified landscapes where access to anthropogenic food sources is common, encounters between wolves and dogs are frequent, and pathogens are constantly evolving.

In Chapter 3, we showed that adaptive introgression from wolves to domestic dogs has likely been more prominent than in the opposite direction, since more chromosomal blocks with overrepresented introgressed regions were found in dogs compared to wolves (31 *vs.* 8). Although admixed dogs acquired a larger pool of beneficial genetic variants of wolf ancestry compared to admixed wolves, none of these CAI loci (3,228 loci) was associated with environmental variables. This result suggests that adaptive introgression does not necessarily correlate with environmental variables and cannot always translate into detectable environmental adaptation. These results may be explained by the fact that wolves and dogs are different from the perspective of behavioral, ecology, and population dynamics. Free-ranging dogs as a generalist species quickly evolved in human-modified habitats (Udell and Brubaker, 2016), and by showing behavioral and ecological flexibility, can adapt and thrive in different environmental conditions like urban and rural (Acosta-Jamett et al., 2010; Boitani et al., 2022; Wheat and Wynne, 2025). Higher flexibility in dogs may allow them to successfully survive in diverse environments (Boitani and Ciucci, 2010). Additionally, while wolves have specific hunting strategies and territorial behaviors, dogs are dependent on human-derived food, which they can obtain independent of environmental conditions (Mech and Boitani, 2003; Marshall-Pescini et al., 2017; Tancredi and Cardinali, 2023). As a result, free-ranging dog distribution and survival are less influenced by environmental variables. However, we should note that the small sample sizes may have influenced our ability to detect significant results in dogs. Similarly, no significant associations between

environmental variables and CAI dog-derived variants (453 SNPs) were observed in admixed jackals. This lack of association may be due to the greater genetic divergence between dogs and golden jackals compared to dogs and wolves (Koepfli et al., 2015; Krofel et al., 2021), potentially limiting the impact of introgressed dog alleles on environmental adaptation in hybrid jackals. Additionally, previous studies showed that jackals are expanding their distribution range to central and northern Europe (Stefanovic et al., 2024), and hybridization between dogs and jackals is more frequent at the edge of the expanding range (see Chapter 3; Stefanovic et al., 2024). Since this expansion is quite recent, there wasn't enough time for natural selection to act on introgressed variants

4.5. Conclusion

Our findings highlight the complex relationship between environmental factors and hybrid ancestry among the representatives of the genus *Canis*. Our findings suggest that environmental factors that can support growing dog populations (e.g., warmer temperatures and increased human presence) are positively linked to dog ancestry in wild canids. This means such conditions may increase the chances of hybridization.

Previous studies have suggested that gene flow from dogs might help wolves adapt to human-related environmental pressures. Our findings provide direct support for this idea. We found a significant association between introgressed adaptive loci from dogs in wolves and environmental factors. Interestingly, the adaptive introgressed genes were associated with significant biological functions like the nervous system, immune system, and metabolism. These results emphasize the role of adaptive introgression, which can help wolves to better adapt to human-modified environments, where they may encounter new pathogens, environmental stressors, and dietary shifts. We highlighted the role of hybridization as an active evolutionary process that can play a more important role than previously thought. These results provide a valuable starting point for future research into how hybridization could help species deal with the future climate change challenges, and also ongoing land cover and land use changes.

4.6. Bibliography

Acosta-Jamett, G., Cleaveland, S., Cunningham, A. A., & Bronsvoort, B. D. (2010). Demography of domestic dogs in rural and urban areas of the Coquimbo region of Chile and implications for disease transmission. Preventive Veterinary Medicine, 94(3-4), 272-281.

Anderson, E. (1948). Hybridization of the habitat. Evolution, 1-9.

Anderson, E. C., & Thompson, E. (2002). A model-based method for identifying species hybrids using multilocus genetic data. Genetics, 160(3), 1217-1229.

Arce-Valdés, L. R., & Sánchez-Guillén, R. A. (2022). The evolutionary outcomes of climate-change-induced hybridization in insect populations. Current Opinion in Insect Science, 54, 100966.

Arnold, J., Humer, A., Heltai, M., Murariu, D., Spassov, N., & Hacklaender, K. (2012). Current status and distribution of golden jackals *Canis aureus* in Europe. Mammal Review, 42(1), 1-11.

Bellard, C., Bertelsmeier, C., Leadley, P., Thuiller, W., & Courchamp, F. (2012). Impacts of climate change on the future of biodiversity. Ecology Letters, 15(4), 365-377.

Blanco, J. C., Reig, S., & de la Cuesta, L. (1992). Distribution, status and conservation problems of the wolf *Canis lupus* in Spain. Biological Conservation, 60(2), 73-80.

Bohling, J. H., Dellinger, J., McVey, J. M., Cobb, D. T., Moorman, C. E., & Waits, L. P. (2016). Describing a developing hybrid zone between red wolves and coyotes in eastern North Carolina, USA. Evolutionary Applications, 9(6), 791-804.

Boitani, L., Ciucci, P., & Raganella-Pelliccioni, E. (2010). predation on livestock in Italy: a tool for conservation? Wildlife Research 37: 722-730. Wildlife Research, 37, 722-730.

Breiman, L. (2001). Random forests. Machine Learning, 45, 5-32.

Brennan, A. C., Woodward, G., Seehausen, O., Muñoz-Fuentes, V., Moritz, C., Guelmami, A., ... & Edelaar, P. (2015). Hybridization due to changing species distributions: adding problems or solutions to conservation of biodiversity during global change?. Evolutionary Ecology Research, 16(6), 475-491.

Brennan, A. C., Woodward, G., Seehausen, O., Muñoz-Fuentes, V., Moritz, C., Guelmami, A., ... & Edelaar, P. (2015). Hybridization due to changing species distributions: adding problems or solutions to conservation of biodiversity during global change?. Evolutionary Ecology Research, 16(6), 475-491.

Buchhorn, M., Lesiv, M., Tsendbazar, N. E., Herold, M., Bertels, L., & Smets, B. (2020). Copernicus global land cover layers—collection 2. Remote Sensing, 12(6), 1044.

Capblancq, T., & Forester, B. R. (2021). Redundancy analysis: A Swiss Army Knife for landscape genomics. Methods in Ecology and Evolution, 12(12), 2298-2309.

Capblancq, T., Luu, K., Blum, M. G., & Bazin, E. (2018). Evaluation of redundancy analysis to identify signatures of local adaptation. Molecular Ecology Resources, 18(6), 1223-1233.

Chapron, G., Kaczensky, P., Linnell, J. D., Von Arx, M., Huber, D., Andrén, H., ... & Boitani, L. (2014). Recovery of large carnivores in Europe's modern human-dominated landscapes. Science, 346(6216), 1517-1519.

Chown, S. L., Hodgins, K. A., & Griffin, P. C. (2016). Biological invasions, climate change, and genomics. Crop Breeding, 59-114.

Chunco, A. J. (2014). Hybridization in a warmer world. Ecology and Evolution, 4(10), 2019-2031.

Cunha Silva, L., Friker, B., Warembourg, C., Kanankege, K., Wera, E., Berger-González, M., ... & Dürr, S. (2022). Habitat selection by free-roaming domestic dogs in rabies endemic countries in rural and urban settings. Scientific Reports, 12(1), 20928.

Cunningham, R. L., Herbert, A. L., Harty, B. L., Ackerman, S. D., & Monk, K. R. (2018). Mutations in dock1 disrupt early Schwann cell development. Neural Development, 13, 1-15.

Cutler, D. R., Edwards Jr, T. C., Beard, K. H., Cutler, A., Hess, K. T., Gibson, J., & Lawler, J. J. (2007). Random forests for classification in ecology. Ecology, 88(11), 2783-2792.

Dar, S. A., Sharief, A., Kumar, V., Singh, H., Joshi, B. D., Bhattacharjee, S., ... & Sharma, L. K. (2023). Free-ranging dogs are seriously threatening Himalayan environment: delineating the high-risk areas for curbing free-ranging dog infestation in the Trans-Himalayan region. Environmental Monitoring and Assessment, 195(11), 1386.

Davison, C. W., Rahbek, C., & Morueta-Holme, N. (2021). Land-use change and biodiversity: Challenges for assembling evidence on the greatest threat to nature. Global Change Biology, 27(21), 5414-5429.

Didham, R. K., Kapos, V., & Ewers, R. M. (2012). Rethinking the conceptual foundations of habitat fragmentation research. Oikos, 121(2), 161-170.

Doan, R. A., & Monk, K. R. (2025). Dock1 functions in Schwann cells to regulate development, maintenance, and repair. Journal of Cell Biology, 224(5), e202311041.

Enserink, M., & Vogel, G. (2006). The carnivore comeback.

Evans, J. S., & Murphy, M. A. (2019). Package 'rfUtilities'. R Core Team: Vienna, Austria.

Fenton, S., Moorcroft, P. R., Ćirović, D., Lanszki, J., Heltai, M., Cagnacci, F., ... & Ranc, N. (2021). Movement, space-use and resource preferences of European golden jackals in human-dominated landscapes: insights from a telemetry study. Mammalian Biology, 101, 619-630.

Fick, S. E., & Hijmans, R. J. (2017). WorldClim 2: new 1-km spatial resolution climate surfaces for global land areas. International Journal of Climatology, 37(12), 4302-4315.

Fletcher Jr, R. J., Didham, R. K., Banks-Leite, C., Barlow, J., Ewers, R. M., Rosindell, J., ... & Haddad, N. M. (2018). Is habitat fragmentation good for biodiversity?. Biological Conservation, 226, 9-15.

Forester, B. R., Lasky, J. R., Wagner, H. H., & Urban, D. L. (2018). Comparing methods for detecting multilocus adaptation with multivariate genotype–environment associations. Molecular Ecology, 27(9), 2215-2233.

Franklin, E. B., Yee, L. D., Aumont, B., Weber, R. J., Grigas, P., & Goldstein, A. H. (2022). Ch3MS-RF: a random forest model for chemical characterization and improved quantification of unidentified atmospheric organics detected by chromatography–mass spectrometry techniques. Atmospheric Measurement Techniques, 15(12), 3779-3803.

Garcia-Lozano, C., Varga, D., Pintó, J., & Roig-Munar, F. X. (2020). Landscape connectivity and suitable habitat analysis for wolves (Canis lupus L.) in the Eastern Pyrenees. Sustainability, 12(14), 5762.

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(24), 5154-5166.

Gounaridis, D., Newell, J. P., & Goodspeed, R. (2020). The impact of urban sprawl on forest landscapes in Southeast Michigan, 1985–2015. Landscape Ecology, 35(9), 1975-1993.

Guan, Y. (2014). Detecting structure of haplotypes and local ancestry. Genetics, 196(3), 625-642.

Hijmans, R. J., Bivand, R., Forner, K., Ooms, J., Pebesma, E., & Sumner, M. D. (2022). Package 'terra'. Maintainer: Vienna, Austria.

Hijmans, R. J., Cameron, S. E., Parra, J. L., Jones, P. G., & Jarvis, A. (2005). Very high resolution interpolated climate surfaces for global land areas. International Journal of Climatology: A Journal of the Royal Meteorological Society, 25(15), 1965-1978.

Hodel, R. G., Massatti, R., & Knowles, L. L. (2022). Hybrid enrichment of adaptive variation revealed by genotype–environment associations in montane sedges. Molecular Ecology, 31(13), 3722-3737.

Kamler, J. F., Minge, C., Rostro-García, S., Gharajehdaghipour, T., Crouthers, R., In, V., ... & Macdonald, D. W. (2021). Home range, habitat selection, density, and diet of golden jackals in the Eastern Plains Landscape, Cambodia. Journal of Mammalogy, 102(2), 636-650.

Kamler, J. F., Minge, C., Rostro-García, S., Gharajehdaghipour, T., Crouthers, R., In, V., ... & Macdonald, D. W. (2021). Home range, habitat selection, density, and diet of golden jackals in the Eastern Plains Landscape, Cambodia. Journal of Mammalogy, 102(2), 636-650.

Kapota, D., Dolev, A., Bino, G., Yosha, D., Guter, A., King, R., & Saltz, D. (2016). Determinants of emigration and their impact on survival during dispersal in fox and jackal populations. Scientific Reports, 6(1), 24021.

Kays, R., Curtis, A., & Kirchman, J. J. (2010). Rapid adaptive evolution of northeastern coyotes via hybridization with wolves. Biology Letters, 6(1), 89-93.

Kim, K. H., Kim, E. Y., Ko, J. J., & Lee, K. A. (2019). Gas6 is a reciprocal regulator of mitophagy during mammalian oocyte maturation. Scientific Reports, 9(1), 10343.

Koepfli, K. P., Pollinger, J., Godinho, R., Robinson, J., Lea, A., Hendricks, S., ... & Wayne, R. K. (2015). Genome-wide evidence reveals that African and Eurasian golden jackals are distinct species. Current Biology, 25(16), 2158-2165.

Kojola, I., Hallikainen, V., Mikkola, K., Gurarie, E., Heikkinen, S., Kaartinen, S., ... & Nivala, V. (2016). Wolf visitations close to human residences in Finland: The role of age, residence density, and time of day. Biological Conservation, 198, 9-14.

Kollataj, W., Milczak, A., Kollataj, B., Sygit, M., & Sygit, K. (2012). The implementation of preventive vaccination of dogs and cats against rabies in rural areas. Environment and Pollution, 1(1), 20.

Korslund, L., & Steen, H. (2006). Small rodent winter survival: snow conditions limit access to food resources. Journal of Animal ecology, 156-166.

Kowalczyk, R., Wudarczyk, M., Wójcik, J. M., & Okarma, H. (2020). Northernmost record of reproduction of the expanding golden jackal population. Mammalian Biology, 100(1), 107-111.

Krofel, M., Giannatos, G., Cirovic, D., Stoyanov, S., & Newsome, T. (2017). Golden jackal expansion in Europe: a case of mesopredator release triggered by continent-wide wolf persecution?. Mammal, 28, 9–15.

Krofel, M., Južnič, D., & Allen, M. L. (2021). Scavenging and carcass caching behavior by European wildcat (*Felis silvestris*). Ecological Research, 36(3), 556-561.

Kudrenko, S., Fenchuk, V., Vollering, J., Zedrosser, A., Selva, N., Ostapowicz, K., ... & Heurich, M. (2023). Walking on the dark side: Anthropogenic factors limit suitable habitat for gray wolf (*Canis lupus*) in a large natural area covering Belarus and Ukraine. Global Ecology and Conservation, 46, e02586.

Kuijper, D. P. J., Churski, M., Trouwborst, A., Heurich, M., Smit, C., Kerley, G. I. H., & Cromsigt, J. P. G. M. (2019). Keep the wolf from the door: How to conserve wolves in Europe's human-dominated landscapes?. Biological Conservation, 235, 102-111.

Kurten, B., and E. Anderson. 1980. Pleistocene mammals of North America. Columbia Univ. Press, New York.

Lamb, A., Balmford, A., Green, R. E., & Phalan, B. (2016). To what extent could edge effects and habitat fragmentation diminish the potential benefits of land sparing?. Biological Conservation, 195, 264-271.

Lanszki, J., Hayward, M. W., Ranc, N., & Zalewski, A. (2022). Dietary flexibility promotes range expansion: The case of golden jackals in Eurasia. Journal of Biogeography, 49(6), 993-1005.

Lanszki, J., Kurys, A., Szabó, L., Nagyapáti, N., Porter, L. B., & Heltai, M. (2016). Diet composition of the golden jackal and the sympatric red fox in an agricultural area (Hungary). Folia Zoologica, 65(4), 310-322.

Legendre, P., & Legendre, L. (2012). Canonical analysis. In Developments in Environmental Modelling (Vol. 24, pp. 625-710). Elsevier.

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

Liaw, A. (2002). Classification and regression by randomForest. R news.

Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. *R news*, *2*(3), 18-22.

Llaneza, L., García, E. J., Palacios, V., Sazatornil, V., & López-Bao, J. V. (2016). Resting in risky environments: the importance of cover for wolves to cope with exposure risk in human-dominated landscapes. Biodiversity and Conservation, 25, 1515-1528.

Lobo, D., Morales, H. E., Van Oosterhout, C., López-Bao, J. V., Silva, P., Llaneza, L., ... & Godinho, R. (2025). Ancient dog introgression into the Iberian wolf genome may have facilitated adaptation to human-dominated landscapes. Genome Research, 35(3), 432-445.

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

Marshall-Pescini, S., Schwarz, J. F., Kostelnik, I., Virányi, Z., & Range, F. (2017). Importance of a species' socioecology: Wolves outperform dogs in a conspecific cooperation task. Proceedings of the National Academy of Sciences, 114(44), 11793-11798.

McCain, C. M., King, S. R., Szewczyk, T., & Beck, J. (2018). Small mammal species richness is directly linked to regional productivity, but decoupled from food resources, abundance, or habitat complexity. Journal of Biogeography, 45(11), 2533-2545.

McGarigal, K., Cushman, S. A., Neel, M. C., & Ene, E. (2002). FRAGSTATS: spatial pattern analysis program for categorical maps. Computer software program produced by the authors at the University of Massachusetts, Amherst. Available at the following web site: www. umass. edu/landeco/research/fragstats/fragstats. html, 6.

Mech, L. D., & Boitani, L. (2003). Wolf social ecology.

Meek, P. D. (1999). The movement, roaming behaviour and home range of free-roaming domestic dogs, Canis lupus familiaris, in coastal New South Wales. Wildlife Research, 26(6), 847-855.

Mmbaga, N. E., Munishi, L. K., & Treydte, A. C. (2017). How dynamics and drivers of land use/land cover change impact elephant conservation and agricultural livelihood development in Rombo, Tanzania. Journal of Land Use Science, 12(2-3), 168-181.

Mohammadi, A., Kaboli, M., Sazatornil, V., & López-Bao, J. V. (2019). Anthropogenic food resources sustain wolves in conflict scenarios of Western Iran. PloS One, 14(6), e0218345.

Mohammadi, A., Kaboli, M., Sazatornil, V., & López-Bao, J. V. (2019). Anthropogenic food resources sustain wolves in conflict scenarios of Western Iran. PloS one, 14(6), e0218345.

Mohammadi, A., Lunnon, C., Moll, R. J., Tan, C. K. W., Hobeali, K., Behnoud, P., ... & Farhadinia, M. S. (2021). Contrasting responses of large carnivores to land use management across an Asian montane landscape in Iran. Biodiversity and Conservation, 30, 4023-4037.

Muhlfeld, C. C., Kovach, R. P., Jones, L. A., Al-Chokhachy, R., Boyer, M. C., Leary, R. F., ... & Allendorf, F. W. (2014). Invasive hybridization in a threatened species is accelerated by climate change. Nature Climate Change, 4(7), 620-624.

Muhly, T. B., Johnson, C. A., Hebblewhite, M., Neilson, E. W., Fortin, D., Fryxell, J. M., ... & Musiani, M. (2019). Functional response of wolves to human development across boreal North America. Ecology and Evolution, 9(18), 10801-10815.

Newsome, T. M., Boitani, L., Chapron, G., Ciucci, P., Dickman, C. R., Dellinger, J. A., ... & Ripple, W. J. (2016). Food habits of the world's grey wolves. Mammal Review, 46(4), 255-269.

Nussberger, B., Currat, M., Quilodran, C. S., Ponta, N., & Keller, L. F. (2018). Range expansion as an explanation for introgression in European wildcats. Biological Conservation, 218, 49-56.

Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'hara, R. B., ... & Oksanen, M. J. (2013). Package 'vegan'. Community Ecology Package, Version, 2(9), 1-295.

Ordiz, A., Uzal, A., Milleret, C., Sanz-Perez, A., Zimmermann, B., Wikenros, C., ... & Sand, H. (2020). Wolf habitat selection when sympatric or allopatric with brown bears in Scandinavia. Scientific Reports, 10(1), 9941.

Ortego, J., Gugger, P. F., & Sork, V. L. (2017). Impacts of human-induced environmental disturbances on hybridization between two ecologically differentiated Californian oak species. New Phytologist, 213(2), 942-955.

Pereira, R., Matias, G., Santos-Reis, M., & Rosalino, L. M. (2024). Influence of habitat edges on spatial and spatio-temporal occurrence patterns of mesocarnivores in landscapes dominated by Eucalyptus plantations. Forest Ecology and Management, 572, 122257.

Porretta, D., & Canestrelli, D. (2023). The ecological importance of hybridization. Trends in Ecology & Evolution.

Püttker, T., Crouzeilles, R., Almeida-Gomes, M., Schmoeller, M., Maurenza, D., Alves-Pinto, H., ... & Prevedello, J. A. (2020). Indirect effects of habitat loss via habitat fragmentation: A cross-taxa analysis of forest-dependent species. Biological Conservation, 241, 108368.

Quintela, M., Thulin, C. G., & Höglund, J. (2010). Detecting hybridization between willow grouse (*Lagopus lagopus*) and rock ptarmigan *(L. muta*) in Central Sweden through Bayesian admixture analyses and mtDNA screening. Conservation Genetics, 11, 557-569.

R Core Team (2021) R: A language and environment for statistical computing. R Foundation for statistical computing. Vienna, Austria. Retrieved from http://www.R-project.org/

Randi, E., Lucchini, V., Christensen, M. F., Mucci, N., Funk, S. M., Dolf, G., & Loeschcke, V. (2000). Mitochondrial DNA variability in Italian and East European wolves: detecting the consequences of small population size and hybridization. Conservation Biology, 14(2), 464-473.

Range, F., & Marshall-Pescini, S. (2022). The socio-ecology of free-ranging dogs. Wolves and Dogs: between Myth and Science, 83-110.

Rellstab, C., Gugerli, F., Eckert, A. J., Hancock, A. M., & Holderegger, R. (2015). A practical guide to environmental association analysis in landscape genomics. Molecular Ecology, 24(17), 4348-4370.

Rezaei, S., Mohammadi, A., Shadloo, S., Ranaie, M., & Wan, H. Y. (2023). Climate change induces habitat shifts and overlaps among carnivores in an arid and semi-arid ecosystem. Ecological Informatics, 77, 102247.

Rotem, G., Berger, H., King, R., & Saltz, D. (2011). The effect of anthropogenic resources on the space-use patterns of golden jackals. The Journal of Wildlife Management, 75(1), 132-136.

Rutkowski, R., Krofel, M., Giannatos, G., Ćirović, D., Männil, P., Volokh, A. M., ... & Bogdanowicz, W. (2015). A European concern? Genetic structure and expansion of

golden jackals (*Canis aureus*) in Europe and the Caucasus. PLoS One, 10(11), e0141236.

Šálek, M., Červinka, J., Banea, O. C., Krofel, M., Ćirović, D., Selanec, I., ... & Riegert, J. (2014). Population densities and habitat use of the golden jackal (*Canis aureus*) in farmlands across the Balkan Peninsula. European Journal of Wildlife Research, 60, 193-200.

Sanderson, E. W., Jaiteh, M., Levy, M. A., Redford, K. H., Wannebo, A. V., & Woolmer, G. (2002). The human footprint and the last of the wild: the human footprint is a global map of human influence on the land surface, which suggests that human beings are stewards of nature, whether we like it or not. BioScience, 52(10), 891-904.

Scanes, C. G. (2018). Human activity and habitat loss: destruction, fragmentation, and degradation. In Animals and human society (pp. 451-482). Academic Press.

Scheffers, B. R., De Meester, L., Bridge, T. C., Hoffmann, A. A., Pandolfi, J. M., Corlett, R. T., ... & Watson, J. E. (2016). The broad footprint of climate change from genes to biomes to people. Science, 354(6313), aaf7671.

Sensharma, R., Reinhard, C. L., Powell, L., & Watson, B. (2024). Public perceptions of free-roaming dogs and cats in India and the United States. Journal of Applied Animal Welfare Science, 1-15.

Serva, D., Iannella, M., Cittadino, V., & Biondi, M. (2023). A shifting carnivore's community: habitat modeling suggests increased overlap between the golden jackal and the Eurasian lynx in Europe. Frontiers in Ecology and Evolution, 11, 1165968.

Shakarashvili, M., Kopaliani, N., Gurielidze, Z., Dekanoidze, D., Ninua, L., & Tarkhnishvili, D. (2020). Population genetic structure and dispersal patterns of grey wolfs (*Canis lupus*) and golden jackals (*Canis aureus*) in Georgia, the Caucasus. Journal of Zoology, 312(4), 227-238.

Shi, S., Zhang, H., Jiang, P., Zhou, Y., Zhu, Y., Feng, T., ... & Chen, J. (2024). Inhibition of LPCAT3 exacerbates endoplasmic reticulum stress and HBV replication. International Immunopharmacology, 143, 113337.

Simpson, T. L., Thiel, R. P., Sailer, D. T., Reineke, D. M., & Thomsen, M. (2023). Demographics of gray wolf (*Canis lupus*) packs recolonizing variable habitats in central Wisconsin. Northeastern Naturalist, 30(1), 75-98.

Soe, E., Davison, J., Süld, K., Valdmann, H., Laurimaa, L., & Saarma, U. (2017). Europe-wide biogeographical patterns in the diet of an ecologically and epidemiologically important mesopredator, the red fox Vulpes vulpes: a quantitative review. Mammal Review, 47(3), 198-211.

Ștefan, V. (2019). geobuffer: R package for constructing geodesic buffers using metric radiuses.

Stefanović, M., Bogdanowicz, W., Adavoudi, R., Martínez-Sosa, F., Doan, K., Flores-Manzanero, A., ... & Pilot, M. (2024). Range-wide phylogeography of the golden jackals (*Canis aureus*) reveals multiple sources of recent spatial expansion and admixture with dogs at the expansion front. Biological Conservation, 290, 110448.

Tancredi, D., & Cardinali, I. (2023). Being a dog: a review of the domestication process. Genes, 14(5), 992.

Torretta, E., Corradini, A., Pedrotti, L., Bani, L., Bisi, F., & Dondina, O. (2022). Hide-and-seek in a highly human-dominated landscape: Insights into movement patterns and selection of resting sites of rehabilitated wolves (*Canis lupus*) in northern Italy. Animals, 13(1), 46.

Torretta, E., Dondina, O., Delfoco, C., Riboldi, L., Orioli, V., Lapini, L., & Meriggi, A. (2020). First assessment of habitat suitability and connectivity for the golden jackal in north-eastern Italy. Mammalian Biology, 100, 631-643.

Trouwborst, A., Krofel, M., & Linnell, J. D. (2015). Legal implications of range expansions in a terrestrial carnivore: the case of the golden jackal (*Canis aureus*) in Europe. Biodiversity and Conservation, 24, 2593-2610.

Udell, M. A., & Brubaker, L. (2016). Are dogs social generalists? Canine social cognition, attachment, and the dog-human bond. Current Directions in Psychological Science, 25(5), 327-333.

van den Bosch, M., Kellner, K. F., Gantchoff, M. G., Patterson, B. R., Barber-Meyer, S. M., Beyer, D. E., ... & Belant, J. L. (2023). Habitat selection of resident and non-resident gray wolves: implications for habitat connectivity. Scientific Reports, 13(1), 20415.

Vorel, A., Kadlec, I., Toulec, T., Selimovic, A., Horníček, J., Vojtěch, O., ... & Barták, V. (2024). Home range and habitat selection of wolves recolonising central European human-dominated landscapes. Wildlife Biology, e01245.

Wang, B., Rong, X., Duerr, M. A., Hermanson, D. J., Hedde, P. N., Wong, J. S., ... & Tontonoz, P. (2016). Intestinal phospholipid remodeling is required for dietary-lipid uptake and survival on a high-fat diet. Cell Metabolism, 23(3), 492-504.

Wang, J., Li, J., Fu, Y., Zhu, Y., Lin, L., & Li, Y. (2024). Research progress, challenges and perspectives of phospholipids metabolism in the LXR-LPCAT3 signaling pathway and its relation to NAFLD. International Journal of Molecular Medicine, 53(4), 1-13.

Wennink, J., Lelieveld, G., de Knegt, H. J., & Klees, D. J. (2019). Habitat suitability analysis for the golden jackal (*Canis aureus*) in the Netherlands. Lutra, 62(1), 13-29.

Wheat, C. H., & Wynne, C. D. (2025). The unfulfilled potential of dogs in studying behavioural ecology and evolution during the Anthropocene. Animal Behaviour, 219, 123020.

Wilson-Aggarwal, J. K., Goodwin, C. E., Moundai, T., Sidouin, M. K., Swan, G. J., Léchenne, M., & McDonald, R. A. (2021). Spatial and temporal dynamics of space use by free-ranging domestic dogs *Canis familiaris* in rural Africa. Ecological Applications, 31(5), e02328.

Xia, W., Grueter, C. C., Zhang, C., Zhuang, H., Hu, J., Krzton, A., & Li, D. (2023). Distribution range contractions and identification of conservation priority areas for canids in Sichuan Province, China. Global Ecology and Conservation, 44, e02499.

Zanni, M., Brogi, R., Merli, E., & Apollonio, M. (2023). The wolf and the city: insights on wolves' conservation in the anthropocene. Animal Conservation, 26(6), 766-780.

Zbinden, Z. D., Douglas, M. R., Chafin, T. K., & Douglas, M. E. (2023). A community genomics approach to natural hybridization. Proceedings of the Royal Society B, 290(1999), 20230768.

Zhou, Y. B., Newman, C., Xu, W. T., Buesching, C. D., Zalewski, A., Kaneko, Y., ... & Xie, Z. Q. (2011). Biogeographical variation in the diet of Holarctic martens (genus *Martes*, Mammalia: Carnivora: Mustelidae): adaptive foraging in generalists. Journal of Biogeography, 38(1), 137-147.

General Discussion

This thesis offers a comprehensive investigation into the evolutionary consequences and ecological drivers of hybridization between wild canids (gray wolves and golden jackals) and free-ranging dogs. Using an integrative approach that combined genomic analyses, selection scans, and environmental associations, we achieved a deeper understanding of how hybridization operates not only as a source of interspecific gene flow but also as a mechanism for local adaptation and evolutionary change.

Our systematic review of studies investigating different evolutionary consequences of hybridization in mammalian orders and families showed that negative consequences of hybridization, like genetic swamping and introgression of variants from domestic animals, have been widely reported. Determining the consequences of hybridization is not always easy, and in some cases, long-term monitoring (for at least two consecutive generations) is needed to conclude about the advantages and disadvantages of introgressive hybridization (Hamilton and Miller, 2016; Chan et al., 2019). Additionally, the type of genetic markers used in a study can significantly influence the ability to detect different consequences of hybridization. For example, negative outcomes such as genetic swamping or loss of genetic integrity are often detectable using a limited number of neutral markers like microsatellites. In contrast, positive outcomes such as adaptive introgression typically require genome-wide data, especially markers located in or near coding regions, such as single-nucleotide polymorphisms (SNPs). The predominance of negative outcomes reported in our literature review can be explained by the fact that many studies rely primarily on neutral genetic markers. While these markers are useful for detecting changes in population structure and genetic diversity, they often miss the functional aspects of gene flow. SNP arrays, by providing a large number of loci (thousands to millions) spanning the entire genome, can improve the accuracy of ancestry estimates and enable the identification of older-generation backcrosses (Goli et al., 2024). Therefore, we concluded that integrating both neutral loci and markers located in coding regions can provide a more comprehensive and balanced understanding of hybridization, capturing not only its potential risks but also its adaptive benefits and the underlying factors that shape these outcomes.

Although many methods and approaches have been developed for inferring introgression rate and detecting hybridization, making decisions to select a method that accurately detects hybridization using genomic data remains challenging (Kong and Kubatko, 2021). In this study, we evaluated the methods of global and local ancestry reconstruction in the context of the analysis of introgressive hybridization. In some cases, the results obtained from these different approaches were inconsistent, which may result from differences in their methodological frameworks, the types of genetic information they utilize, and their strategies for handling missing data. We highlight two major factors, low-quality genotypes and the presence of subpopulation structure, as key contributors to these inconsistencies, both of which increase the uncertainty and variability in ancestry estimates. We found that global ancestry analyses such as ADMIXTURE are more likely to be affected by these confounding factors. Therefore, global ancestry methods may not be suitable as standalone approaches for the precise

inference of admixture proportions due to their susceptibility to confounding factors. These challenges highlight the importance of methodological precision and careful dataset curation to ensure accurate detection and interpretation of admixture. We recommend a joint use of local and global methods of ancestry analysis, with local ancestry results being prioritized for precise inference of introgression rates.

Using this methodological knowledge from the last chapter and the robust performance of ELAI in estimating ancestry proportions, in the next step, ELAI analysis was applied to estimate individual ancestry proportions, with source populations defined based on LAMP-LD results. We found that hybridization in the genus Canis is common in their distribution range. In some regions, including the Balkans, India, the Caucasus, and northeastern Europe, a higher frequency of hybridization was found, which may have resulted from high human disturbances, large population size of free-ranging dogs, and the range expansion of golden jackals. The role of these factors in increasing the chance of hybridization has been documented in other species as well (e.g., hybridization between wild and domestic cats, Nussberger et al., 2018; Matias et al., 2022). Human activities and habitat loss have increased hybridization between species by disrupting mating patterns and creating habitat conditions that may favour hybrid individuals (Szynwelski et al., 2023). Human population growth combined with the fragmentation of natural habitats increases both the numbers of domestic animals and the probability of encounters with their wild relatives (Adavoudi and Pilot, 2021). The global dog population is currently estimated at around one billion (Gompper, 2014). The presence of large, unmanaged free-ranging dog populations further exacerbates hybridization risks by increasing the frequency of encounters with wild canids. In addition to these factors, hybridization appears to occur more frequently in regions of recent species expansion compared to established core habitats (e.g., in jackals: Stefanovic et al., 2024; in coyotes: Kays et al., 2010). In this study, evidence of hybridization between jackals and dogs along the northern part of the jackal's range and in newly colonized areas was found. These findings align with previous research documenting similar patterns of hybridization in expanding jackal populations (Stefanovic et al., 2024; Ninausz et al., 2023; Galov et al., 2015). Therefore, identification of these factors and their direct and indirect effects on the hybridization frequency is important from the perspective of management of cross-breeding canids.

We clearly show the effect of evolutionary distances between the species on introgression rate between them, since a higher frequency of dog introgression was found in wolves compared to golden jackals. This finding supports the fact that the frequency of hybridization in sister species and also in recently diverged lineages is higher than in non-sister or more distantly related species (Gholamhosseini et al., 2013; Wei et al., 2023).

Both wild canids and free-ranging dogs may gain benefits from hybridization. Through adaptive introgression, wild canids may acquire from dogs gene variants conferring adaptive advantage, including those that strengthen their immune systems. These beneficial genes may increase the resistance of wild canids to new pathogens, which would be particularly beneficial in environments where wild canids encounter dogs frequently and where pathogens are constantly evolving. These results were aligned

with recent studies on hybridization between wolf and dog, which showed that adaptive introgression from domestic dogs into wolves may enhance canids' adaptation to human-dominated landscapes (Pilot et al., 2021; Lobo et al., 2025). Therefore, we can expect that introgression from dogs could provide an adaptive advantage for wild canids, especially those living in regions highly modified by humans. Free-ranging dogs appear to have acquired a larger pool of beneficial genetic variants from wolves, which may have contributed to some characteristics like morphological, behavioural, and physiological traits. Among genes in chromosomal blocks with excess introgression that were under positive selection in canids, genes related to the nervous and immune systems were predominant in wild canids and free-ranging dogs. In addition to detecting signals of positive selection, we also found signs of negative selection in introgressed chromosomal blocks in dogs and golden jackals. These results suggest that some introgressed gene variants may also have a deleterious effect on these species, but they can be efficiently removed from their gene pools. Overall, we highlight the complex nature of hybridization and introgression in the evolutionary process, showing that it can introduce both beneficial and maladaptive genetic variation.

Although numerous studies have suggested that hybridization between wolves and dogs is more frequent in human-modified landscapes, this pattern has remained largely unproven. In this study, we demonstrate that dog-derived ancestry in wild canids is positively associated with human footprint. This finding supports the hypothesis that increasing human disturbance or anthropogenic pressure may facilitate introgression between wild canids and free-ranging dogs. Regions with high human footprint may often have a larger number of free-ranging dogs, increasing the chances of encounters with wild canids and potentially promoting hybridization (Pilot et al., 2021). Additionally, land use and land cover change can occur in response to human activities, which may cause habitat destruction and fragmentation (Gounaridis et al., 2020). Such disruptions may alter the distribution of many species and force them to move beyond their natural habitats and even approach human settlements in order to obtain food and/or avoid the intra-specific spatial competition (Scanes, 2018; Zanni et al., 2023).

Additionally, studies on hybridization in canids have suggested that introgression from dogs into the gene pool of wolves might help wolves to better adapt to human human-modified landscapes (Newsome et al. 2017; Pilot et al. 2021; Lobo et al., 2025). Our findings provide direct support for this idea. We found a significant association between dog-derived introgressed adaptive loci in wolves and environmental factors. Interestingly, the adaptive introgressed genes were associated with the nervous system, immune system, and metabolism. These results emphasize the role of adaptive introgression, which can help wolves to better adapt to human-modified environments, where wolves may encounter new pathogens, environmental stressors, and dietary shifts. We highlighted the role of hybridization as an active evolutionary process that can play a more significant role than previously thought. Although finding a link between environmental variables and adaptive introgressed loci is an important step, it doesn't necessarily mean these genes directly improve fitness in wolves. To truly understand the benefits of these introgressed variants, we'll need experimental research and long-term monitoring of wild populations. These results provide a valuable starting point for future

research into how hybridization could help species deal with the future climate change challenges, and also ongoing land cover and land use changes. Exploring this concept further could offer insights into how hybridization might serve as an adaptive mechanism in rapidly changing environments.

Collectively, these results highlight the complex nature of hybridization and introgression in the evolutionary process. While adaptive introgressed variants may enhance local adaptation of species, deleterious alleles may also be introduced, potentially disrupting locally adapted gene complexes or increasing vulnerability to disease and other stressors. A well-known example from human evolution that illustrates this complexity is Neanderthal introgression, which contributed both beneficial traits, such as enhanced immune responses, and detrimental ones, including increased susceptibility to certain health conditions (Racimo et al., 2015; Dannemann, 2021). Therefore, to fully understand its evolutionary and conservation significance, it's crucial to identify candidate adaptive genes, detect their functional roles, and examine how they interact with the specific environmental conditions where they occur.

In Future work, we recommend applying whole-genome sequencing (WGS) to uncover broader introgressed regions in hybrid individuals. While SNP arrays provide valuable insights, WGS offers much higher resolution than SNP arrays and can detect small or rare introgressed chromosomal blocks (Theissinger et al., 2023). Also, WGS is recommended to explore the admixture patterns of individuals with different ancestries from wolves, golden jackals, and dogs (e.g., F1 × F1 crosses or backcrosses). Although F1 hybrids and recent backcrosses are rare and therefore cross-breeding between them is statistically unlikely, such hybrids may be locally common and show unique behaviours that promote further gene flow. Therefore, future studies using WGS could help clarify their genomic backgrounds and identify adaptive introgressed regions which linked to behaviour, morphology, or fitness.

The significant positive association between adaptively introgressed loci and environmental variables such as human footprint and climate variables presents a valuable opportunity for future research. Future studies can focus on clarifying the specific functions of these environmentally associated genes and evaluating whether they confer fitness advantages to wild canids across different habitats. Furthermore, given that climate change is a major threat to many species, we recommend assessing genomic vulnerability to rapid climate change in both pure and hybrid individuals and evaluating if admixed individuals with adaptive introgression exhibit greater resilience under future climate, compared to non-admixed individuals. Such assessments will help determine whether introgressive hybridization provides novel adaptive variation that could facilitate evolutionary rescue in species at risk from climate change.

Bibliography

Adavoudi, R., & Pilot, M. (2021). Consequences of hybridization in mammals: A systematic review. Genes, 13(1), 50.

Chan, W. Y., Hoffmann, A. A., & van Oppen, M. J. (2019). Hybridization as a conservation management tool. Conservation Letters, 12(5), e12652.

Dannemann, M. (2021). The population-specific impact of Neandertal introgression on human disease. Genome Biology and Evolution, 13(1), evaa250.

Galov, A., Fabbri, E., Caniglia, R., Arbanasić, H., Lapalombella, S., Florijančić, T., ... & Randi, E. (2015). First evidence of hybridization between golden jackal (*Canis aureus*) and domestic dog (*Canis familiaris*) as revealed by genetic markers. Royal Society Open Science, 2(12), 150450.

Gholamhosseini, A., Vardakis, M., Aliabadian, M., Nijman, V., & Vonk, R. (2013). Hybridization between sister taxa versus non-sister taxa: a case study in birds. Bird Study, 60(2), 195-201.

Goli, R. C., Chishi, K. G., Ganguly, I., Singh, S., Dixit, S. P., Rathi, P., ... & Kanaka, K. K. (2024). Global and local ancestry and its importance: a review. Current Genomics, 25(4), 237-260.

Gompper, M. E. (2014). The dog-human-wildlife interface: assessing the scope of the problem. Free-ranging Dogs and Wildlife Conservation, 1, 9-54.

Kays, R., Curtis, A., & Kirchman, J. J. (2010). Rapid adaptive evolution of northeastern coyotes via hybridization with wolves. Biology Letters, *6*(1), 89-93.

Kong, S., & Kubatko, L. S. (2021). Comparative performance of popular methods for hybrid detection using genomic data. Systematic Biology, 70(5), 891-907.

Lobo, D., Morales, H. E., Van Oosterhout, C., López-Bao, J. V., Silva, P., Llaneza, L., ... & Godinho, R. (2025). Ancient dog introgression into the Iberian wolf genome may have facilitated adaptation to human-dominated landscapes. Genome Research, 35(3), 432-445.

Matias, G., Rosalino, L. M., Alves, P. C., Tiesmeyer, A., Nowak, C., Ramos, L., ... & Monterroso, P. (2022). Genetic integrity of European wildcats: variation across biomes mandates geographically tailored conservation strategies. Biological Conservation, 268, 109518.

Miller, J. M., & Hamilton, J. A. (2016). Interspecies hybridization in the conservation toolbox: response to Kovach et al.(2016). Conservation Biology, 30(2), 431-433.

Newsome, T. M., Fleming, P. J., Dickman, C. R., Doherty, T. S., Ripple, W. J., Ritchie, E. G., & Wirsing, A. J. (2017). Making a new dog?. BioScience, 67(4), 374-381.

Ninausz, N., Fehér, P., Csányi, E., Heltai, M., Szabó, L., Barta, E., ... & Stéger, V. (2023). White and other fur colourations and hybridization in golden jackals (*Canis aureus*) in the Carpathian basin. Scientific Reports, 13(1), 21969.

Nussberger, B., Currat, M., Quilodran, C. S., Ponta, N., & Keller, L. F. (2018). Range expansion as an explanation for introgression in European wildcats. Biological Conservation, 218, 49-56.

Pilot, M., Moura, A. E., Okhlopkov, I. M., Mamaev, N. V., Manaseryan, N. H., Hayrapetyan, V., ... & Bogdanowicz, W. (2021). Human-modified canids in human-modified landscapes: The evolutionary consequences of hybridization for grey wolves and free-ranging domestic dogs. Evolutionary Applications, 14(10), 2433-2456.

Racimo, F., Sankararaman, S., Nielsen, R., & Huerta-Sánchez, E. (2015). Evidence for archaic adaptive introgression in humans. Nature Reviews Genetics, 16(6), 359-371.

Scanes, C. G. (2018). Human activity and habitat loss: destruction, fragmentation, and degradation. In Animals and Human Society (pp. 451-482). Academic Press.

Stefanović, M., Bogdanowicz, W., Adavoudi, R., Martínez-Sosa, F., Doan, K., Flores-Manzanero, A., ... & Pilot, M. (2024). Range-wide phylogeography of the golden jackals (*Canis aureus*) reveals multiple sources of recent spatial expansion and admixture with dogs at the expansion front. Biological Conservation, 290, 110448.

Szynwelski, B. E., Kretschmer, R., Matzenbacher, C. A., Ferrari, F., Alievi, M. M., & de Freitas, T. R. O. (2023). Hybridization in canids—A case study of Pampas fox (*Lycalopex gymnocercus*) and domestic dog (*Canis lupus familiaris*) hybrid. Animals, 13(15), 2505.

Theissinger, K., Fernandes, C., Formenti, G., Bista, I., Berg, P. R., Bleidorn, C., ... & Zammit, G. (2023). How genomics can help biodiversity conservation. Trends in Genetics, 39(7), 545-559.

Wei, S., Zhang, Q., Tang, S., & Liao, W. (2023). Genetic and ecophysiological evidence that hybridization facilitated lineage diversification in yellow Camellia (Theaceae) species: A case study of natural hybridization between C. micrantha and C. flavida. BMC Plant Biology, 23(1), 154.

Zanni, M., Brogi, R., Merli, E., & Apollonio, M. (2023). The wolf and the city: insights on wolves' conservation in the anthropocene. Animal Conservation, 26(6), 766-780.

Supplementary Information

Supporting Materials for the PhD Thesis:

The Role of Hybridization in the Evolutionary Response to Environmental Change in the Genus *Canis*

Author: Roya Adavoudi Jolfaei
Institution: Uniwersytet Gdański
Year: 2025

Chapter 2

Material and methods

DNA extraction

During the wash steps, all columns were centrifuged for 2 minutes, followed by removing residual ethanol through centrifugation of the columns for 3 minutes at 11,000xg. To increase the yield, DNA was eluted in two steps with 50µL double distilled water and incubated for 3 minutes at room temperature. After the incubation, all columns were centrifuged for 3 minutes at 11,000xg. Total DNA extract volume of 100 µL per sample. DNA from saliva samples was extracted by PG-AC extraction kit, following the manufacturer's instructions. Ice-cold ethanol was used to enhance DNA yield. Samples with low DNA concentration (below 15 ng/µL), were re-extracted with only 25 µL of double-distilled water used for DNA elution in each step. To monitor potential DNA contamination, negative controls were included in all protocols. The quality and quantity of DNA were measured using a NanoDrop ND-1000 (Thermo Fisher Scientific) spectrophotometer and fluorescent-based Qubit® quantitation assay (Thermo Fischer Scientific).

Datasets description

Table S2. 1. Summary of quality control analysis for all main datasets

Dataset	Total samples	Wolf (N)	Jackal (N)	Dog (N)	SNPs before QC	Missing data	MAF	LD pruning	Total SNP passed QC (LD pruning)	Total SNP passed QC (no LD pruning)
WJD	1386	315	478	593	229,120	8	21,795	172,357	34,960	207,317
WD	908	315	-	593	229,120	94	22,521	144,047	62,458	206,505
JD	1071	-	478	593	229,120	140	29,617	171,848	27,515	199,363
WJ	793	315	478	-	229,120	504	40,355	160,930	27,331	188,261

Regional datasets

Table S2. 2. Summary of quality control analysis for the Indian and Balkan datasets

Dataset	Total samples	Wolf (N)	Jackal (N)	Dog (N)
IWJD	100	21	33	46
IWD	67	21	-	46
IJD	79	-	33	46
IWJ	54	21	33	-
BWJD	327	110	167	50
BWD	160	110	-	50
BJD	217	-	167	50
BWJ	277	110	167	-

Population structure

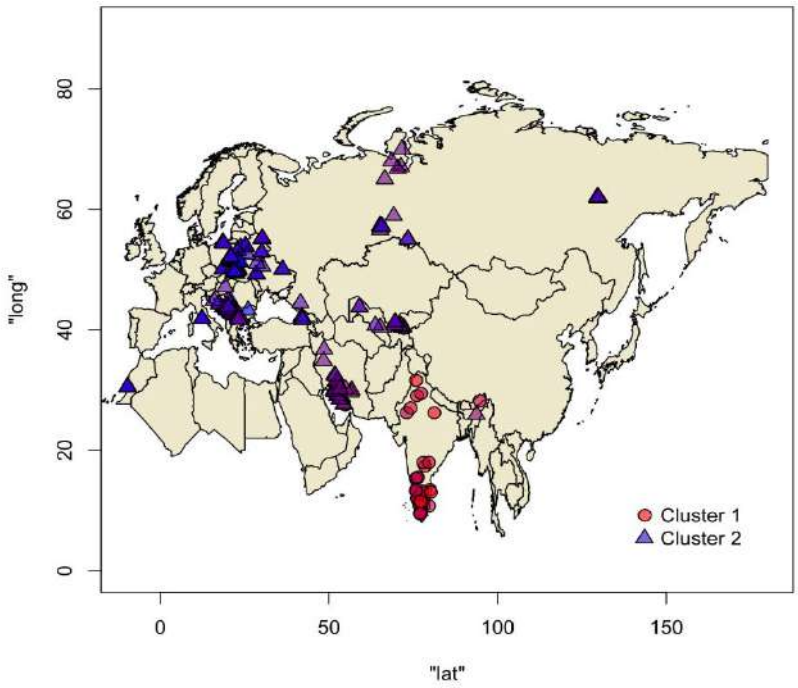


Fig. S 2. 1. Results of DAPC analysis for dog samples. Two different clusters are represented by different colors.

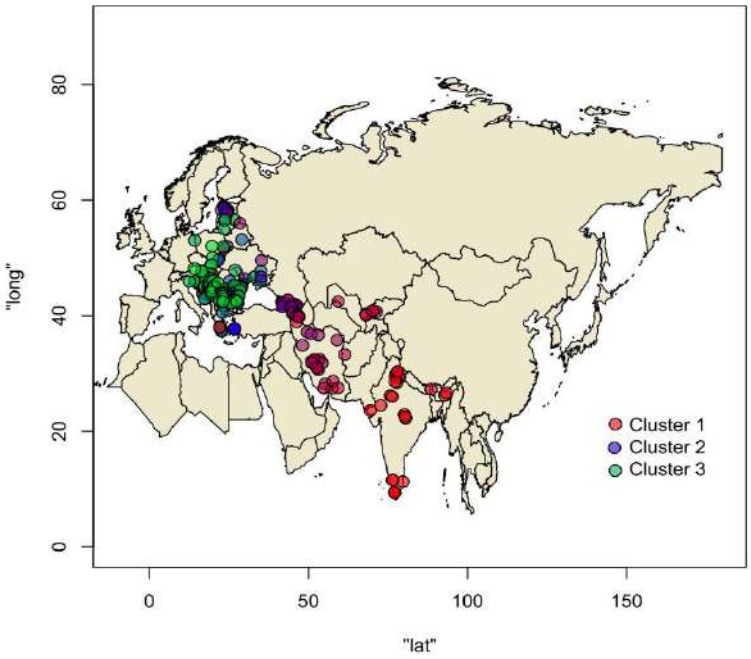


Fig. S 2. 2. Results of DAPC analysis for jackal samples. Two different clusters are represented by different colors.

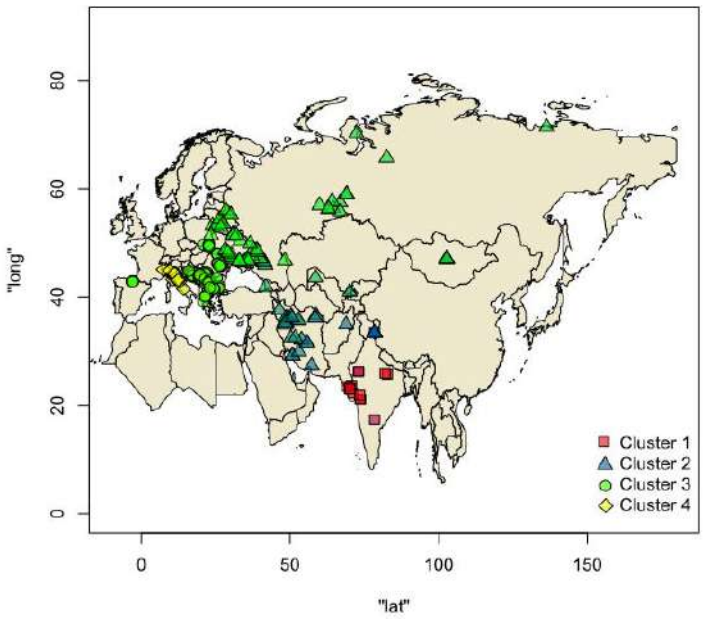


Fig. S 2. 3. Results of DAPC analysis for wolf samples. Four different clusters are represented by different colors.

Regional dataset

Global ancestry (PCA and ADMIXTURE)

The results of PCA for the India and the Balkan datasets are reported in Table S 2.3 (Fig S 2.4, Fig S 2.5). Based on the results of ADMIXTURE, At K=2, using the IWD and IWJ datasets, all Himalayan wolves were identified as F2 or F3 backcrosses (Fig S 2.6). The results of ADMIXTURE for the Balkan dataset (BWD and BJD) were mostly consistent with the findings from ADMIXTURE using the entire dataset and the results from PCA (Table S 2.4, Fig S 2.7). Three canids sampled as wolves (WBOS18, WSER483, WSER466) displayed a 50% assignment probability to both the wolf and dog clusters, indicating that they were first-generation wolf-dog hybrids. Regarding jackal samples, nine out of 167 jackal samples had less than 90% jackal ancestry and were LQ samples. Among these samples, four jackal samples (JBUL78-19_LQ, JBUL410-19_LQ, JBUL241-19_LQ, and JWBOS38_LQ) displayed coefficient assignments between 30 and 50-% to the jackal cluster (Table S 2.4, Fig S 2.7). The average ancestry values in wolves, jackals, and dogs are reported in Table S 2.5.

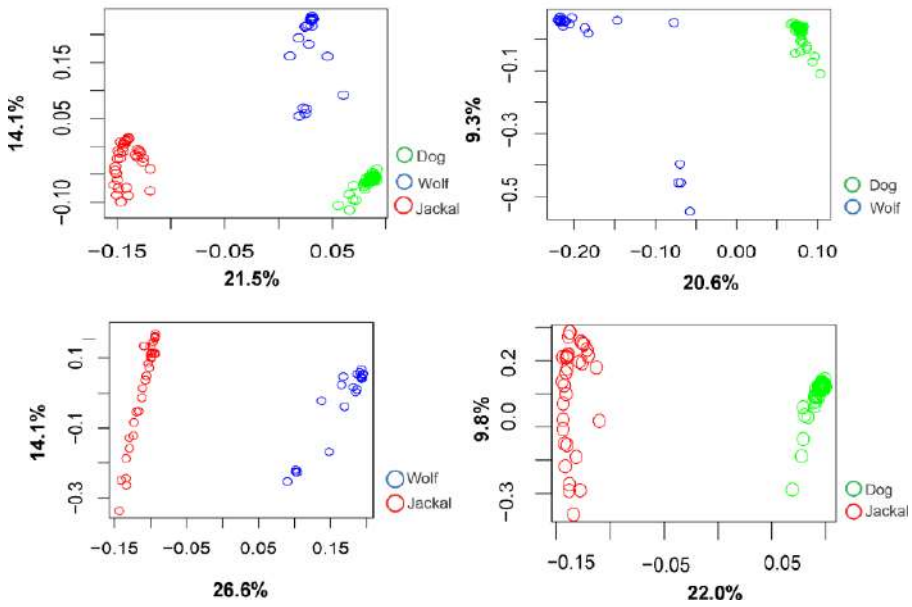


Fig. S 2. 4. Plots of two first principal components for all Indian datasets (IWJD, IWD, IJD, and IWJ).

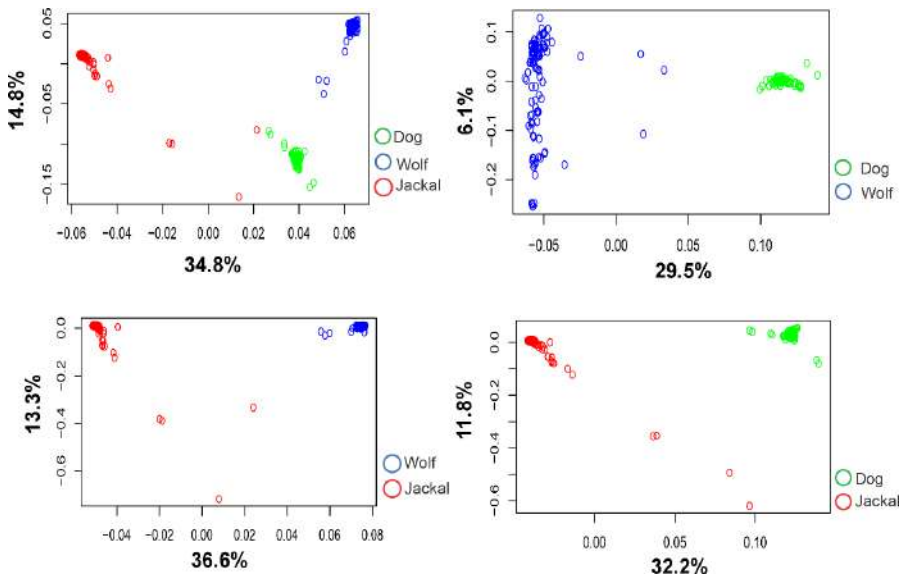


Fig. S 2. 5.Plots of two first principal components for all Balkan datasets (BWJD, BWD, BJD, and BWJ).

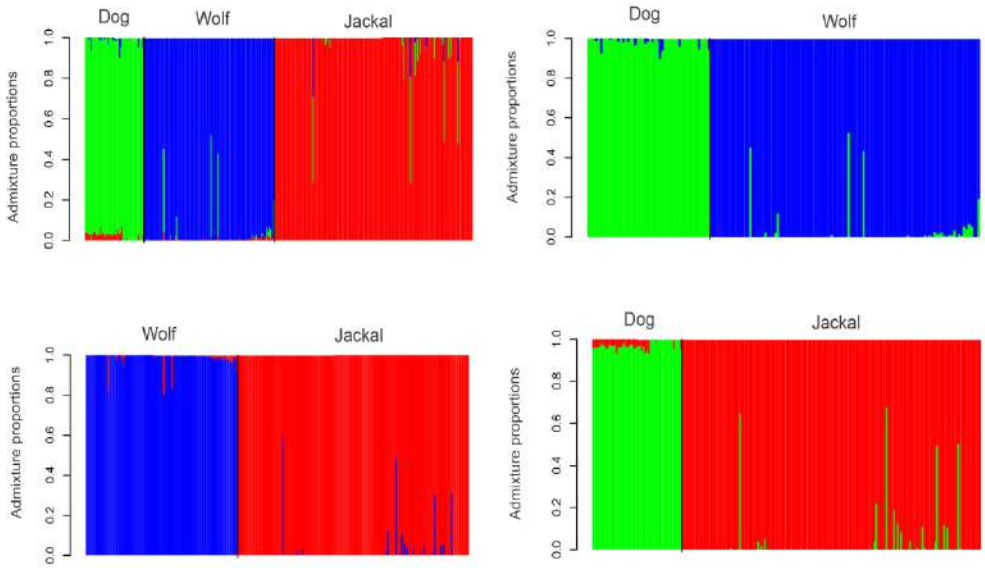


Fig. S 2. 6. Admixture plot for K=3 using the IWJD dataset, K =2 using the IWD, IJD, and IWJ

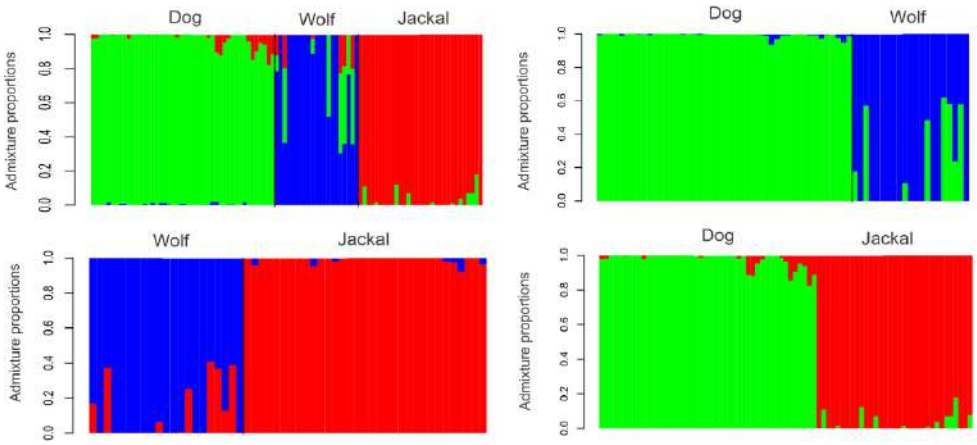


Fig. S 2. 7. Admixture plot for K=3 using the BWJD dataset, K =2 using the BWD, BJD, and BWJ

Local ancestry analysis

Based on the results of LAMP-LD, in the wolf and jackal samples from India (IWJ), all wolf samples displayed no more than 1% of jackal ancestry, and only five jackal samples displayed between 11-35% wolf ancestry. This result also was inconsistent with the results of LAMP-LD of the entire dataset. Tables S 2.6 and S 2.7, represent the putative hybrids and the mean proportion of wolf, dog, and jackal ancestry in Indian datasets based on the LAMP-LD results.

The results of LAMP-LD of the wolf and jackal dataset from the Balkans, showed that all wolf samples as pure samples (100% wolf ancestry). Three golden jackals had 50% wolf ancestry in their chromosomes (Table S 2.6). However, these samples were identified as backcrosses based on the results of ADMIXTURE.

Based on the results of ElAI, using the Indian wolf-jackal dataset (IWJ), all jackal samples showed more than 99% jackal ancestry. All Indian jackals identified as F1 or backcross based on ELAI using the entire dataset were confirmed as pure samples based on the Indian dataset (Table S 2.8). The mean proportion of wolf, dog, and jackal ancestry in Indian datasets is reported below (Table S 2.9).

In the wolf and jackal samples from the Balkans, all wolf samples were recognized as pure samples (100% wolf ancestry). Two jackal samples were recognized as f1 hybrids, however just only one of them shoewed wolf ancestry in most of the chromosomes (Table S 2.8). The mean proportion of wolf, dog, and jackal ancestry in the Balkan datasets is reported below (Table S 2.9). According to the results of GHap for the Indian wolf-jackal dataset (IWJ), all wolf and jackal samples had more than 99% assignment to their respective ancestries (Table S 2.10). In dog-jackal dataset (BJD) and wolf-jackal dataset (BWJ) from the Balkan, all dogs, jackals, and wolves showed more than 98% of their ancestries and were identified as pure samples (Table S 2.10). The mean proportion of wolf, dog, and jackal ancestry in Indian datasets is reported below (Table S 2.11).

Table S2. 3. Identified putative F1 hybrids based on PCA results using the entire and both regional datasets. Samples from the Balkans are indicated by an asterisk (*), while samples from India are marked with a double asterisk (**). Additionally, samples identified as F1 hybrids in both the regional and entire datasets are highlighted in bold

	Entire datasets				India				Balkan			
Putative hybrids	**WJD**	**WD**	**JD**	**WJ**	**IWJD**	**IWD**	**IJD**	**IWJ**	**BWJD**	**BWD**	**BJD**	**BWJ**
JROM10658_LQ	✓		✓	✓								
JHUN9531	✓		✓	✓								
JBEL598_LQ	✓		✓	✓								
WSER466*	✓	✓							✓	✓		✗
WSER483*	✓	✓							✓	✓		✗
WBOS18*	✓	✓							✓	✓		✗
WIRA616	✓	✓										
WIndD2449**	✓	✓			✓	✓		✗				
WIRA1042_LQ	✓											
WIRA631_LQ		✓		✓								
JBUL78-19_LQ*	✗		✗	✗					✓		✓	✓
JBUL410-19_LQ*	✗		✗	✗					✓		✓	✓
JBUL241-19_LQ*	✗		✗	✗					✓		✓	✓
JWBOS38_LQ*	✗		✗	✗					✓		✓	✓

✓ Samples show 50% of ancestry from each of the two canids in each chromosome

✓ Samples show an average of 50% ancestry from each of the two canids but not within each chromosome

✗ Samples not identified as F1 hybrid

Table S2. 4. Identified putative F1 hybrids based on ADMIXTURE results using the entire and both regional datasets. Samples from the Balkans are indicated by an asterisk (*), while samples from India are marked with a double asterisk (**). Additionally, samples identified as F1 hybrids in both the regional and entire datasets are highlighted in bold

	Entire datasets				India				Balkan			
Putative hybrids	**WJD**	**WD**	**JD**	**WJ**	**IWJD**	**IWD**	**IJD**	**IWJ**	**BWJD**	**BWD**	**BJD**	**BWJ**
WIndD2449**	✓	✓		✗	✓	✓		✗				
WIndD3078_LQ**	✗	✓		✗								
WIndD3011**					✓							
WIndD3012**					✓	✓		✗				
WIndDLDK_LQ**					✓	✓		✗				
WIndD48rep**					✓	✓		✗				
JIndD471_LQ**	✓		✓	✓	✗		✗	✗				
JIndD473_LQ**	✓			✓	✗			✗				
JIndD474_LQ**	✓		✗	✗								
JIndD2629_LQ**	✓		✗	✓	✗			✗				
JIndDF4398rep_LQ**	✓		✓	✓	✗		✗	✗				
JIndD2743_LQ**	✓		✓	✓	✗		✗	✗				
JIndD2639_LQ**	✓				✗							
JIndD3274_LQ**	✗		✓	✓			✗	✗				
JIndD2175_LQ**	✗		✓	✓			✗	✗				
JIndD3018_LQ**	✗		✓	✓			✗	✗				
JIndD45_LQ**	✗		✓	✗			✗					
JIndDF417_LQ**	✗		✓	✓			✗	✗				
JIndD480_LQ**	✗		✓				✗					
WSER466*	✓	✓		✗					✓	✓		✗
WSER483*	✓	✓		✗					✓	✓		✗
WBOS18*	✓	✓		✗					✓	✓		✗
JBUL241-19_LQ*	✓		✓	✓					✓		✓	✗
JBUL410-19_LQ*	✗		✓	✓					✓		✓	✗
JBUL78-19_LQ*	✗		✗	✗					✓		✓	✓
JWBOS38_LQ*	✓		✓	✗					✓		✓	✗

✓ Samples show 50% of ancestry from each of the two canids in each chromosome

✓ Samples show an average of 50% ancestry from each of the two canids but not within each chromosome

✗ Samples not identified as F1 hybrids

Table S2. 5. The average proportions of dog, wolf, and jackal ancestry in the entire and both regional datasets based on ADMIXTURE.

	Entire dataset			**India**			**Balkans**		
	Wolf	**Jackal**	**Dog**	**Wolf**	**Jackal**	**Dog**	**Wolf**	**Jackal**	**Dog**
Wolf ancestry ADMIXTURE	-	0.054	0.039	-	0.008	0.007		0.014	0.011
Jackal ancestry ADMIXTURE	0.031	-	0.012	0.102	-	0.023	0.010		0.027
Dog ancestry ADMIXTURE	0.064	0.048	-	0.159	0.023	-	0.019	0.021	

Table S2. 6. Identified putative F1 hybrids based on LAMP-LD results using the entire and both regional datasets. Samples from the Balkans are indicated by an asterisk (*), while samples from India are marked with a double asterisk (**). Additionally, samples identified as F1 hybrids in both the regional and entire datasets are highlighted in bold

	Entire datasets			India			Balkan		
Putative hybrids	**WD**	**JD**	**WJ**	**IWD**	**IJD**	**IWJ**	**BWD**	**BJD**	**BWJ**
WIndD2449**	✓			✓					
JIndD471_LQ**			✓			✗			
JIndD473_LQ**			✓			✗			
JIndDF4398rep_LQ**			✓			✗			
JIndD2743_LQ**		✓	✓		✗	✗			
JIndD3274_LQ**		✓	✓		✗	✗			
JIndD2175_LQ**		✓	✓		✗	✗			
JIndD3018_LQ**		✓	✓		✗	✗			
JIndD45_LQ**		✓	✓		✗	✗			
JIndDF417_LQ**		✓	✓		✗	✗			
JIndD480_LQ**		✓	✓		✗	✗			
WSER466*	✓						✓		
WSER483*	✓						✓		
WBOS18*	✓						✓		
JBUL241-19_LQ*		✓	✓					✓	✓
JBUL410-19_LQ*		✓	✓					✓	✓
JBUL78-19_LQ*		✓	✓					✓	✓
JWBOS38_LQ*		✓						✓	
JROM10658_LQ		✓	✓						
JHUN9531		✓	✓						
JBEL598_LQ		✓	✓						
WIRA631_LQ	✓		✓						
WIRA616	✓								
WIRA595_LQ	✓								
WMON8160_LQ	✓								
JWKAU5740_LQ		✓							
JWKAU5741_LQ		✓							
JKAU8086_LQ		✓							
JKAU8341_LQ		✓							
JKAU8321_LQ		✓							
JGEO42_LQ			✓						
JGRE9066_LQ		✓							
JURK8926_LQ		✓							
JUKR8600_LQ		✓	✓						

✓ Samples show 50% of ancestry from each of the two canids in each chromosome

✓ Samples show an average of 50% ancestry from each of the two canids but not within each chromosome

✗ Samples not identified as F1 hybrids

Table S2. 7. The mean percentage of SNP alleles of dog, wolf, and jackal ancestry based on LAMP-LD

	Entire dataset			**India**			**Balkans**		
	Wolf	Jackal	**Dog**	**Wolf**	**Jackal**	**Dog**	**Wolf**	**Jackal**	**Dog**
Wolf ancestry LAMP-LD	-	0.044	0.005	-	0.000	0.000	-	0.018	0.001
Jackal ancestry LAMP-LD	0.001	-	0.000	0.001	-	0.000	-	-	0.027
Dog ancestry LAMP-LD	0.044	0.031	-	0.082	0.000	-	0.023	0.000	-

Table S2. 8. Identified putative F1 hybrids based on ELAI results using the entire and both regional datasets. Samples from the Balkans are indicated by an asterisk (*), while samples from India are marked with a double asterisk (**). Additionally, samples identified as F1 hybrids in both the regional and entire datasets are highlighted in bold

	Entire datasets				**India**			**Balkan**		
Putative hybrids	**WJD**	**WD**	**JD**	**WJ**	**IWD**	**IJD**	**IWJ**	**BWD**	**BJD**	**BWJ**
WIndD2449**	✓	✓			✓					
JIndD3018_LQ**	✗			○			✗			
JIndD45_LQ**	✗			✓			✗			
JIndDF417_LQ**	✗			✓			✗			
JIndD480_LQ**	✗			✓			✗			
WSER466*	✓	✓						✓		
WSER483*	○	○						○		
JWBOS38_LQ*										○
WBOS18*	○	○						○		
JBUL78-19_LQ*	✗		✓	✓					○	✓

○ Samples show 50% of ancestry from each of the two canids in each chromosome

○ Samples Show an average of 50% ancestry from each of the two canids but not within each chromosome

✗ Samples not identified as F1 hybrids

Table S2. 9. The mean percentage of SNP alleles of dog, wolf, and jackal ancestry based on ELAI.

	Entire dataset			**India**			**Balkans**		
	Wolf	**Jackal**	**Dog**	**Wolf**	**Jackal**	**Dog**	**Wolf**	**Jackal**	**Dog**
Wolf ancestry ELAI	-	0.035	0.004	-	0.000	0.000	-	0.010	0.004
Jackal ancestry ELAI	0.000	-	0.000	-	-	0.000	-	-	0.000
Dog ancestry ELAI	0.042	0.020	-	0.097	0.000	-	0.022	0.010	-

Table S2. 10. Identified putative F1 hybrids based on GHap results using the entire and both regional datasets. Samples from the Balkans are indicated by an asterisk (*), while samples from India are marked with a double asterisk (**). Additionally, samples identified as F1 hybrids in both the regional and entire datasets are highlighted in bold

	Entire datasets			India			Balkan		
Putative hybrids	**WD**	**JD**	**WJ**	**IWD**	**IJD**	**IWJ**	**BWD**	**BJD**	**BWJ**
WIndD2449**	✗			✓					
WSER466*	✓						✓		
WSER483*	✓						✓		
WBOS18*	✓						✓		
JROM10658_LQ		✓	✓						
JHUN9531		✓	✓						
JBEL598_LQ		✓							
WIRA616	✓								

✓ Samples show 50% of ancestry from each of the two canids in each chromosome

✓ Samples Show an average of 50% ancestry from each of the two canids but not within each chromosome

✗ Samples not identified as F1 hybrids

Table S2. 11. The mean percentage of SNP alleles of dog, wolf, and jackal ancestry based on GHap results

	Entire dataset			India			Balkans		
	Wolf	**Jackal**	**Dog**	**Wolf**	**Jackal**	**Dog**	**Wolf**	**Jackal**	**Dog**
Wolf ancestry GHap	-	0.006	0.003	-	0.000	0.000	-	0.003	0.001
Jackal ancestry GHap	0.000	-	0.000	0.000	-	0.000	0.000	-	0.000
Dog ancestry GHap	0.034	0.006	-	0.034	0.000	-	0.021	0.003	-

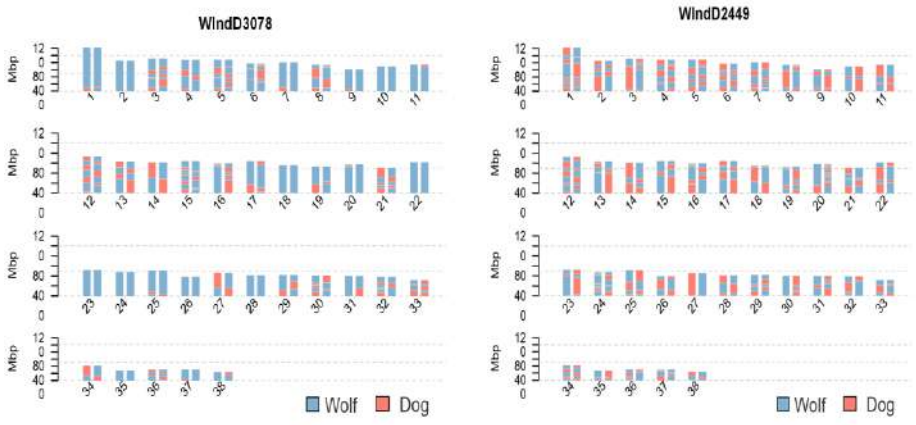


Fig. S 2. 8. Karyoplots of admixed wolves (Dataset IWD) in all 38 chromosomes based on the results of GHap

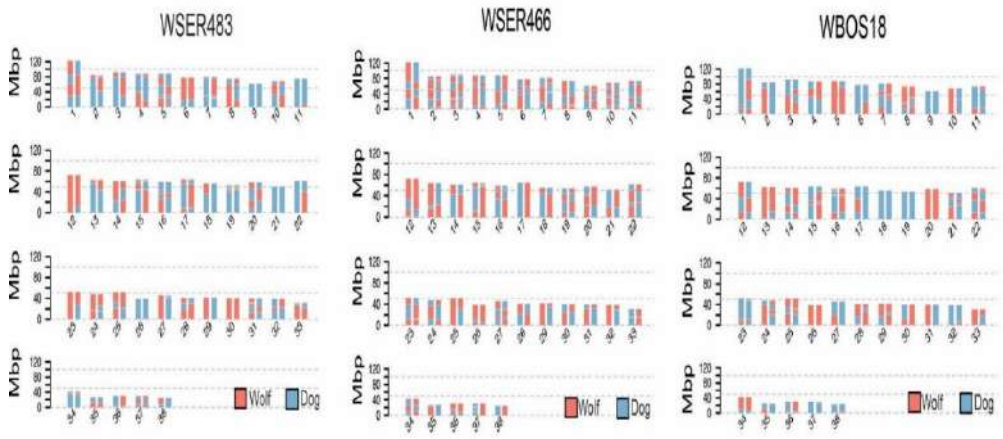


Fig. S 2. 9. Karyoplots of admixed wolves (Dataset BWD) in all 38 chromosomes based on the results of GHap

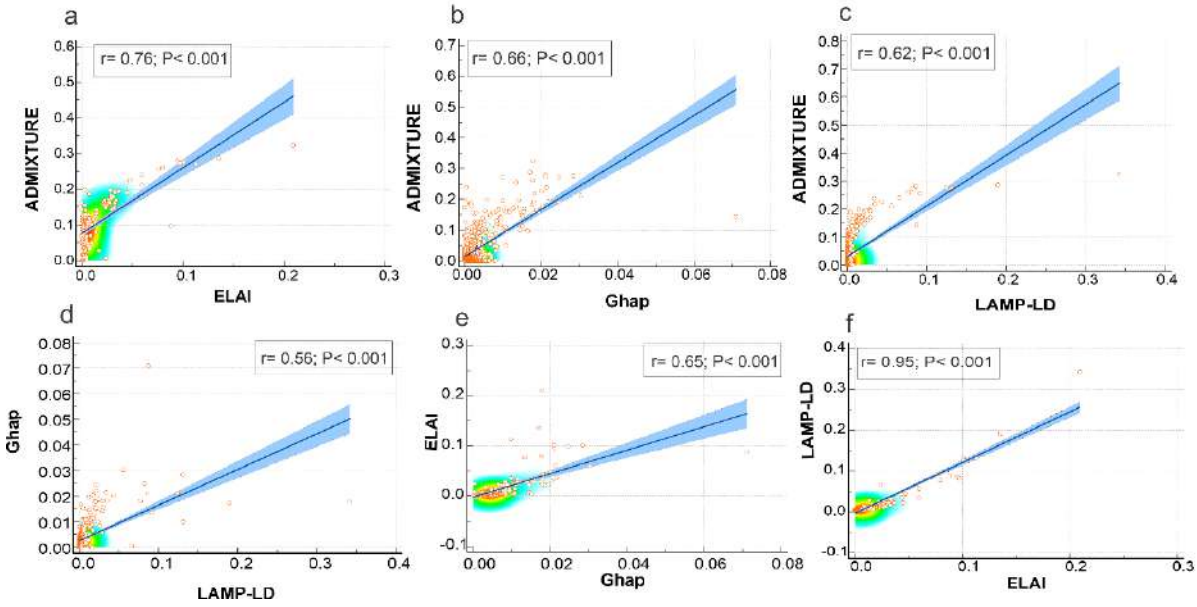


Fig. S 2. 10. The scatter plots between the estimated wolf ancestry in dogs between ADMIXTURE and ELAI (a), ADMIXTURE and GHap (b), ADMIXTURE and LAMP-LD (c), LAMP-LD and GHap (d), ELAI and GHap (e), and LAMP-LD and ELAI (f).

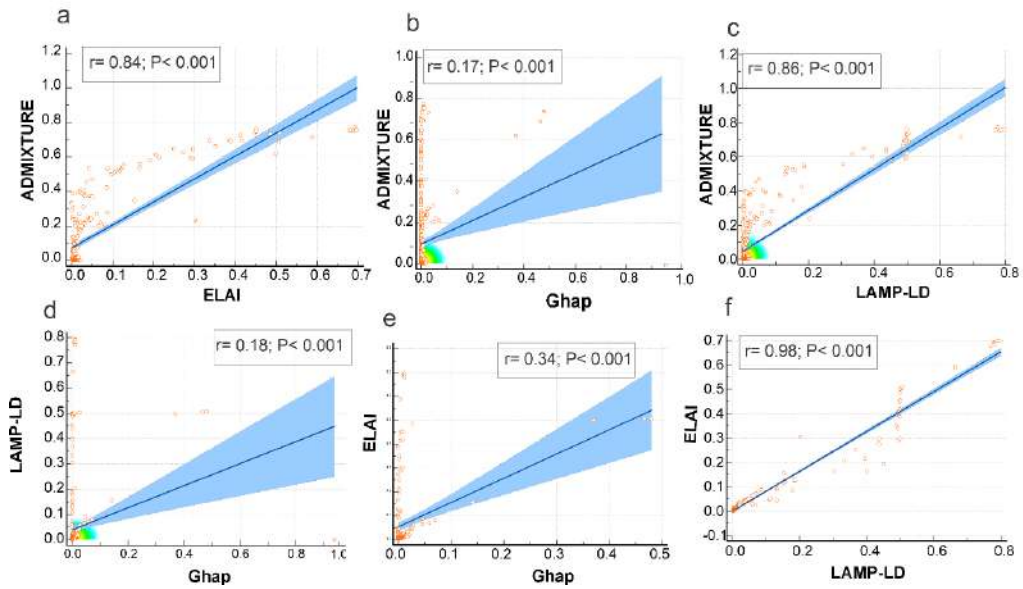


Fig. S 2. 11. The scatter plots between the estimated wolf ancestry in jackals between ADMIXTURE and ELAI (a), ADMIXTURE and GHap (b), ADMIXTURE and LAMP-LD (c), LAMP-LD and GHap (d), ELAI and GHap (e), and LAMP-LD and ELAI (f).

Table S2. 12. Number of detected outliers samples based on the Bland-Altman plot (dog ancestry in wolves from India)

Method	Number of outliers	Differences
ADMIXTURE	3	>= 0.38
LAMP-LD	2	>= 0.23
GHap	1	>= 0.045

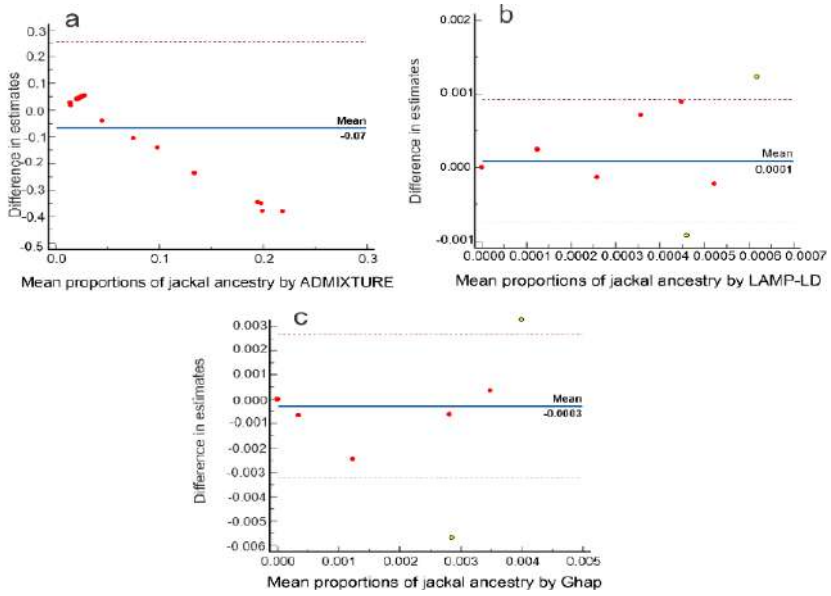


Fig. S 2. 12. Bland-Altman plots of average estimated jackal ancestry in wolf samples from India by ADMIXTURE (a), LAMP-LD (b), and GHap (c). The x-axis shows the average proportions of jackal ancestry using the entire and Balkan dataset and the y-axis shows the difference between estimated ancestry using the entire and Indian datasets.

Table S2. 13. Number of detected outliers samples based on the Bland-Altman plot (jackal ancestry in wolves from India)

Method	Number of outliers	Differences
ADMIXTURE	0	>= 0.35
LAMP-LD	2	>= 0.0008
GHap	2	>= 0.002

Table S2. 14. Number of detected outliers samples based on the Bland-Altman plot (dog ancestry in jackals from India)

Method	Number of outliers	Differences
GHap	2	>= 0.001

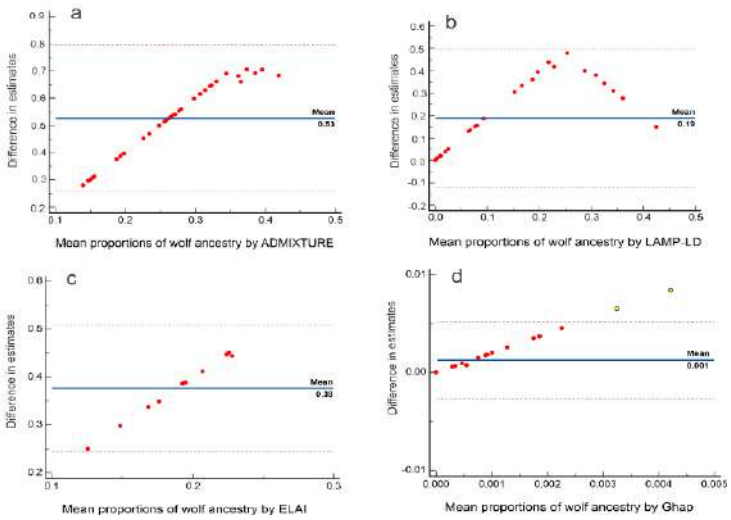


Fig. S 2. 13. Bland-Altman plots of average estimated wolf ancestry in jackals samples from India by ADMIXTURE (a), LAMP-LD (b), ELAI (c), and GHap (d). The x-axis shows the average proportions of wolf ancestry using the entire and Balkan datasets and the y-axis shows the difference between estimated ancestry using the entire and Indian datasets. The outlier samples are displayed in yellow color.

Table S2. 15. Number of detected outliers samples based on the Bland-Altman plot (wolf ancestry in jackals from India)

Method	Number of outliers	Differences
GHap	2	>= 0.006

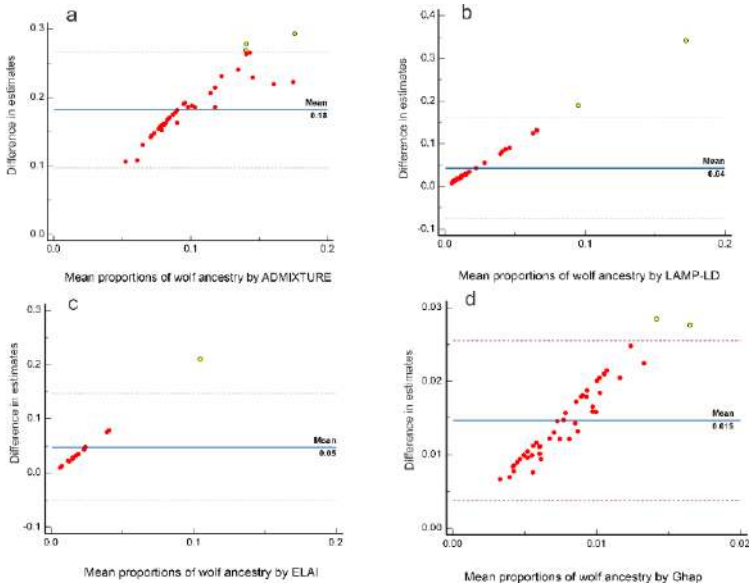


Fig. S 2. 14. Bland-Altman plots of average estimated wolf ancestry in dog samples from India by ADMIXTURE (a), LAMP-LD (b), ELAI (c), and GHap (d). The x-axis shows the average proportions of wolf ancestry using the entire and Indian datasets and the y-axis shows the difference between estimated ancestry using the entire and Indian datasets. The outlier samples are displayed in yellow color.

Table S2. 16. Number of detected outliers samples based on the Bland-Altman plot (wolf ancestry in dogs from India)

Method	Number of outliers	Differences
ADMIXTURE	3	>= 0.26
LAMP-LD	2	>= 0.18
GHap	2	>= 0.27
ELAI	1	>= 0.2

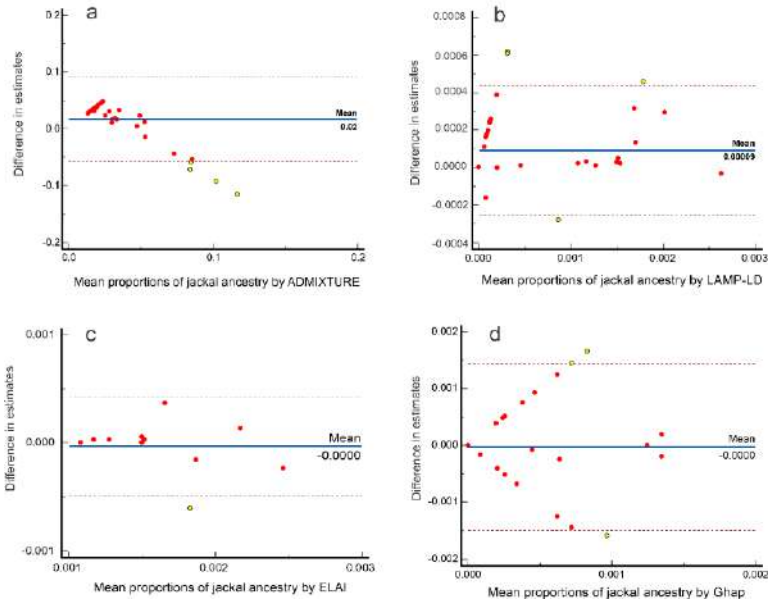


Fig. S 2. 15. Bland-Altman plots of average estimated jackal ancestry in dog samples from India by ADMIXTURE (a), LAMP-LD (b), ELAI (c), and GHap (d). The x-axis shows the average proportions of jackal ancestry using the entire and Indian datasets and the y-axis shows the difference between estimated ancestry using the entire and Indian datasets.

Table S2. 17. Number of detected outliers samples based on the Bland-Altman plot (jackal ancestry in dogs from India)

Method	Number of outliers	Differences
ADMIXTURE	4	>= 0.043
LAMP-LD	4	>= 0.0002
GHap	3	>= 0.0014
ELAI	1	>= 0.0006

Table S2. 18. The list of identified outlier samples based on the Bland-Altman plots using the entire and the India dataset.

	Outliers			
	ADMIXTURE	LAMP-LD	GHap	ELAI
Dog ancestry in Wolves	WIndD48rep WIndD3011 WIndD3012	WIndD48rep WIndDLDK_LQ	WIndD2449	-
Dog ancestry in jackals	-	-	JIndD2629_LQ JIndDF417_LQ	-
Wolf ancestry in jackals	-	-	JIndD480_LQ JIndDF417_LQ	-
Wolf ancestry in dogs	DIndDf2636rep_LQ DIndDf2647rep_LQ DIndDf2653rep_LQ	DIndD236_LQ DIndDf2636rep_LQ	DIndD2641rep_LQ DIndMirza2rep	DIndDf2636rep_LQ
Jackal ancestry in wolves	-	WIndD2449 WIndD3011	WIndD3011 WIndD3012	-
Jackal ancestry in dogs	DIndD236_LQ DIndD2487_LQ DIndDf2636rep_LQ DJIndD466_LQ	DIndB2 DIndB18rep DIndB25rep DIndB26rep	DIndB10 DIndB11 DIndD2641rep_LQ	DIndD2487_LQ

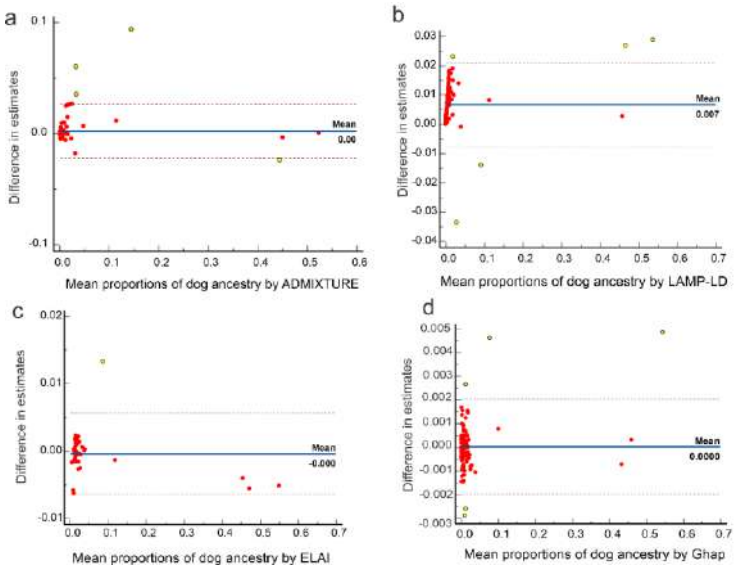


Fig. S 2. 16. Bland-Altman plots of average estimated dog ancestry in wolf samples from the Balkans by ADMIXTURE (a), LAMP-LD (b), ELAI (c), and GHap (d). The x-axis shows the average proportions of dog ancestry using the entire and Balkan dataset and the y-axis shows the difference between estimated ancestry using the entire and Balkan datasets. The outlier samples are displayed in yellow color.

Table S2. 19. Number of detected outliers samples based on the Bland-Altman plot (dog ancestry in wolves from the Balkans)

Method	Number of outliers	Differences
ADMIXTURE	4	>= 0.02
LAMP-LD	5	>= 0.02
GHap	5	>= 0.00031
ELAI	1	>= 0.01

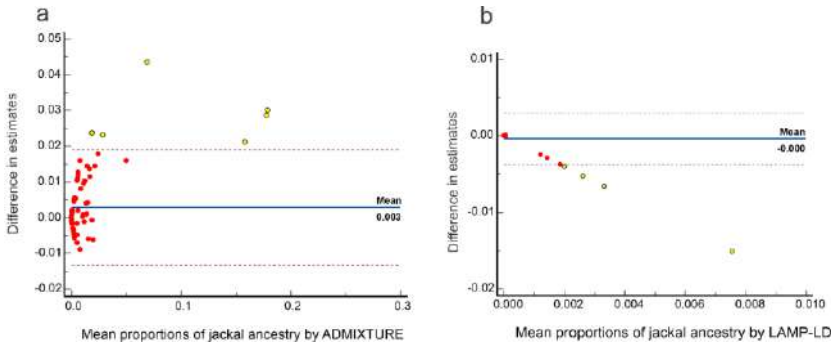


Fig. S 2. 17. Bland-Altman plots of average estimated jackal ancestry in wolf samples from the Balkans by ADMIXTURE (a), and LAMP-LD (b). The x-axis shows the average proportions of jackal ancestry using the entire and Balkan dataset and the y-axis shows the difference between estimated ancestry using the entire and Balkan datasets. The outlier samples are displayed in yellow color.

Table S2. 20. Number of detected outliers samples based on the Bland-Altman plot (jackal ancestry in wolves from the Balkans)

Method	Number of outliers	Differences
ADMIXTURE	6	>= 0.02
LAMP-LD	4	>= 0.004

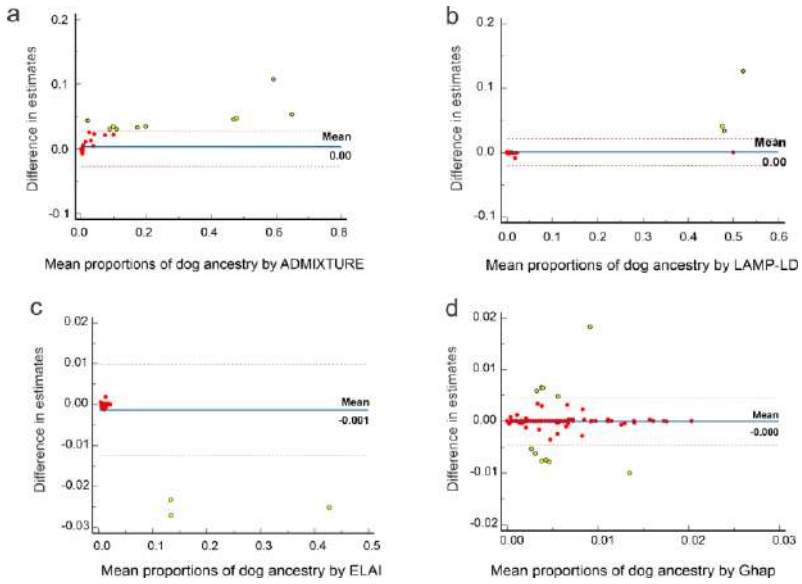


Fig. S 2. 18. Bland-Altman plots of average estimated dog ancestry in jackal samples from the Balkans by ADMIXTURE (a), LAMP-LD (b), ELAI (c), and GHap (d). The x-axis shows the average proportions of dog ancestry using the entire and Balkan dataset and the y-axis shows the difference between estimated ancestry using the entire and Balkan datasets.

Table S2. 21. Number of detected outliers samples based on the Bland-Altman plot (dog ancestry in jackals from the Balkans)

Method	Number of outliers	Differences
ADMIXTURE	10	>= 0.02
LAMP-LD	3	>= 0.02
GHap	11	>= 0.005
ELAI	3	>= 0.02

Table S2. 22. Number of detected outliers samples based on the Bland-Altman plot (wolf ancestry in jackals from the Balkans)

Method	Number of outliers	Differences
ADMIXTURE	7	>= 0.045
LAMP-LD	1	>= 0.082
GHap	2	>=0.007
ELAI	2	>= 0.13

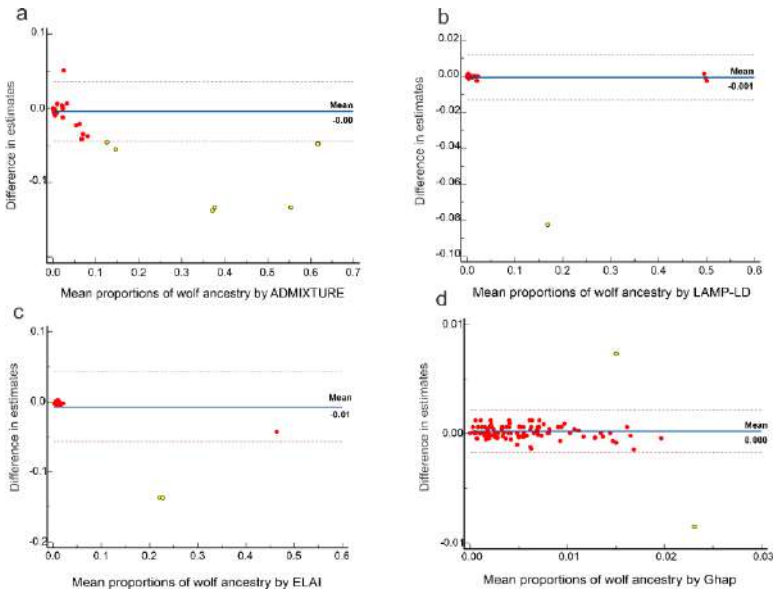


Fig. S 2. 19. Bland-Altman plots of average estimated wolf ancestry in jackal samples from the Balkans by ADMIXTURE (a), LAMP-LD (b), ELAI (c), and GHap (d). The x-axis shows the average proportions of wolf ancestry using the entire and Balkan dataset and the y-axis shows the difference between estimated ancestry using the entire and Balkan datasets. The outlier samples are displayed in yellow color.

Table S2. 23. Number of detected outliers samples based on the Bland-Altman plot (wolf ancestry in dogs from the Balkans)

Method	Number of outliers	Differences
ADMIXTURE	1	>= 0.075
LAMP-LD	1	>= 0.012
GHap	3	>=0.0035
ELAI	0	>= 0.045

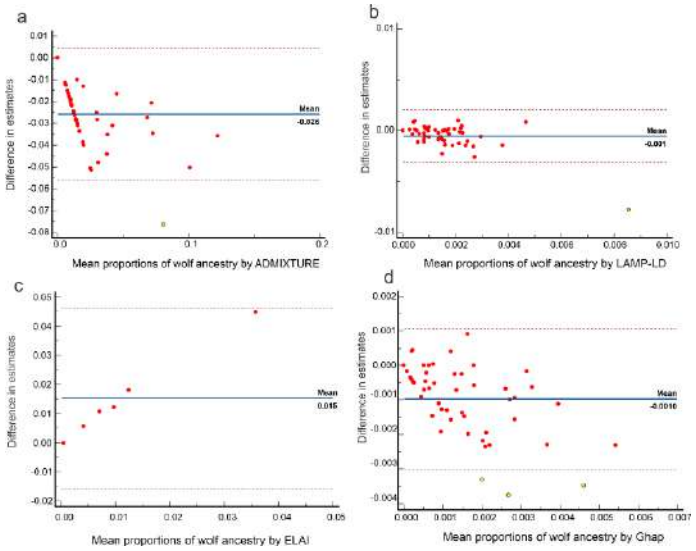


Fig. S 2. 20. Bland-Altman plots of average estimated wolf ancestry in dog samples from the Balkans by ADMIXTURE (a), LAMP-LD (b), ELAI (c), and GHap (d). The x-axis shows the average proportions of wolf ancestry using the entire and Balkan dataset and the y-axis shows the difference between estimated ancestry using the entire and Balkan datasets. The outlier samples are displayed in yellow color.

Table S2. 24. Number of detected outliers samples based on the Bland-Altman plot (jackal ancestry in dogs from the Balkans)

Method	Number of outliers	Differences
ADMIXTURE	3	>= 0.035
LAMP-LD	3	>= 0.001
GHap	4	>= 0.00031

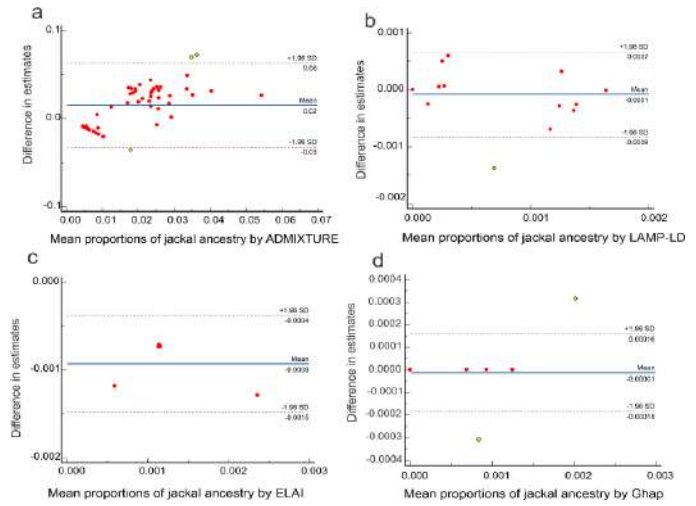


Fig. S 2. 21.Bland-Altman plots of average estimated jackal ancestry in dog samples from the Balkans by ADMIXTURE (a), LAMP-LD (b), ELAI (c), and GHap (d). The x-axis shows the average proportions of dog ancestry using the entire and Balkan dataset and the y-axis shows the difference between estimated ancestry using the entire and Balkan datasets. The outlier samples are displayed in yellow color.

Table S2. 25. The list of identified outlier samples based on the Bland-Altman plots using the entire and the Balkan dataset.

	Outliers			
	ADMIXTURE	**LAMP-LD**	**GHap**	**ELAI**
Dog ancestry in Wolves	WSER483 BW156 BW157 BWJ161	WBOS18 WSER466 WSER608 BW152_LQ BWJ161	WSER163 WSER442 WSER466 BW160 BWJ161	BWJ161
Dog ancestry in jackals	JWBOS38_LQ JBUL23-19_LQ JBUL78-19_LQ JBUL101-19_LQ JBUL110-19_LQ JBUL138-19 JBUL241-19_LQ JBUL283-19_LQ JBUL298-19_LQ JBUL410-19_LQ	JBUL241-19_LQ JBUL410-19_LQ JWBOS38_LQ	JBOS11 JBOS12 JBOS13 JBOS15 JBOS17 JBOS21 JBOS34 JBOS39 JBOS40 JBOS42 JWBOS38_LQ	JBUL78-19_LQ JBUL241-19_LQ JBUL410-19_LQ
Wolf ancestry in jackals	JWBOS38_LQ JBUL23-19_LQ JBUL78-19_LQ JBUL101-19_LQ JBUL138-19 JBUL241-19_LQ JBUL410-19_LQ	JWBOS38_LQ	JWBOS38_LQ JBUL78-19_LQ	JBUL241-19_LQ JBUL410-19_LQ
Wolf ancestry in dogs	DBUL7_LQ	DWBUL292	DSER2219 DWBUL301 DBUL12	-
Jackal ancestry in wolves	WBOS18 WSER466 WSER483 BW156 BW157 BWJ161	WBOS18 WSER466 WSER483 BW152_LQ	-	-
Jackal ancestry in dogs	DSER2226 DWBUL296 DBUL7_LQ	DBUL5 DBUL10 DBUL12	DSER2246 DBUL5 DBUL10 DBUL12	-

Chapter 3

Table S3. 1. Chromosomal blocks with overrepresentation of introgressed ancestry in the WD dataset

Species	chr	Chromosomal block position	Block size	Average inrogressed ancestry	N SNPs within chromosomal blocks	Genes candidate for Adaptive Introgression	Enriched genes
Wolf	1	chr1:24192886 - chr1:24427550	234,664	0.122	20	4	1
Wolf	1	chr1:68714786 - chr1:69928630	1,213,844	0.139	109	10	0
Wolf	13	chr13:25414838 - chr13:26999347	1,584,509	0.141	120	2	1
Wolf	22	chr22:60570318 - chr22:61380477	810,159	0.157	152	27	1
Wolf	25	chr25:13202008 - chr25:15189189	1,987,181	0.131	213	18	3
Wolf	27	chr27:37946382 - chr27:38973872	1,027,490	0.106	112	48	11
Wolf	28	chr28:35319394 - chr28:36185356	865,962	0.117	110	4	0
Wolf	29	chr29:33249669 - chr29:33382531	132,862	0.104	12	1	0
Wolf			**982,083.875**	**0.127**	**848**	**114**	**17**
Dogs	1	chr1:61529416 - chr1:63528127	1,998,711	0.062	140	8	5
Dogs	1	chr1:93328078 - chr1:93472154	144,076	0.054	7	6	1
Dog	1	chr1:99467483 - chr1:101511482	2,043,999	0.056	164	72	24
Dog	2	chr2:39694626 - chr2:41025654	1,331,028	0.051	135	11	4
Dog	2	chr2:48956689 - chr2:49512305	555,616	0.048	75	2	1
Dog	3	chr3:20920 - chr3:41955	21,035	0.043	2	6	1
Dog	3	chr3:229421 - chr3:1986252	1,756,831	0.046	112	8	2
Dog	4	chr4:21849 - chr4:2138082	2,116,233	0.105	121	15	3
Dog	4	chr4:52925637 - chr4:54002987	1,077,350	0.103	100	7	3
Dog	5	chr5:83971576 - chr5:85175224	1,203,648	0.045	115	2	0
Dog	6	chr6:17850095 - chr6:20228256	2,378,161	0.048	173	52	27
Dog	6	chr6:49183506 - chr6:49748296	564,790	0.049	56	11	1
Dog	7	chr7:27610479 - chr7:28586072	975,593	0.082	98	10	7
Dog	8	chr8:60541280 - chr8:61105727	564,447	0.047	58	6	1
Dog	9	chr9:9003519 - chr9:12597729	3,594,210	0.164	315	69	16
Dog	10	chr10:51538 - chr10:1702011	1,650,473	0.062	85	86	38

Dog	10	chr10:28538175 - chr10:28854653	316,478	0.044	40	11	5
Dog	11	chr11:8578208 - chr11:9976527	1,398,319	0.057	63	14	6
Dog	12	chr12:23806393 - chr12:25454840	1,648,447	0.057	125	9	3
Dog	13	chr13:327635 - chr13:2408078	2,080,443	0.068	136	20	6
Dog	13	chr13:29457133 - chr13:30387571	930,438	0.077	116	5	1
Dog	17	chr17:51048 - chr17:3571693	3,520,645	0.043	181	18	6
Dog	19	chr19:3308202 - chr19:4254441	946,239	0.041	41	8	4
Dog	20	chr20:19145 - chr20:1431900	1,412,755	0.085	72	16	5
Dog	22	chr22:42089667 - chr22:42518167	428,500	0.032	52	1	0
Dog	24	chr24:601689 chr24:2958054	2,356,365	0.082	143	12	4
Dog	26	chr26:8818 - chr26:915242	906,424	0.127	100	25	6
Dog	27	chr27:43905014 - chr27:45797335	1,892,321	0.087	140	27	4
Dog	28	chr28:965210 - chr28:2639267	1,674,057	0.069	141	30	12
Dog	30	chr30:8037955 - chr30:8915049	877,094	0.067	99	25	12
Dog	32	chr32:33510428 - chr32:33698953	188,525	0.051	22	3	1
Dog			1,372,685.516	0.066	3227	595	211

Table S3. 2. Chromosomal blocks with overrepresentation of introgressed in the JD datas

Species	chr	Chromosomal block position	Block size	Average inrogressed ancestry	N SNPs	Genes candidate for Adaptive Introgression	Enriched genes
Jackal	1	Chr1:121638246-Chr1:122670980	1,032,734	0.1692	105	8	0
Jackal	3	Chr3:78647199 Chr3:78960101	312,902	0.1760	28	10	0
Jackal	9	Chr9:50006743 Chr9:50839358	832,615	0.1681	64	19	1
Jackal	11	Chr11:58996495 Chr11:59947627	951,132	0.1782	84	11	0
Jackal	16	Chr16:52516621 Chr16:53436243	919,622	0.2321	104	10	1
Jackal	19	Chr19:50872952 Chr19:50970518	97,566	0.2242	12	0	0
Jackal	20	Chr20:18425282 Chr20:18942411	517,129	0.1451	56	36	2
Jackal			666242.85	0.1847	453 (in sum)	94 (in sum)	4 (in sum)

Table S3. 3. Chromosomal blocks with overrepresentation of introgressed ancestry in the WJ dataset

Species	chr	Chromosomal block position	Block size	Average inrogressed ancestry	N SNPs	Genes candidate for Adaptive Introgression	Enriched genes
Jackal	1	Chr1:121649389 - Chr1:121758416	109,027	0.0862	14	5	2
Jackal			109,027	0.0862	14	Sum 5	2
Wolf	13	Chr13:36250 - Chr13:2333650	2,297,400	0.410	150	22	0
Wolf	16	Chr16:12243526 - Chr16:12381294	137,768	0.177	11	1	0
Wolf	21	Chr21:34124601 Chr21:34273290	148,689	0.180	38	2	0
Wolf			861285.66	0.255	66.33	25	0

Table S3. 4. Chromosomal blocks with underrepresentation of introgressed ancestry in the WD dataset

Population	**chr**	**Chromosomal block position**	**Block size**	**N SNPs**
Dog	1	Chr1:5141525-chr1:5244422	102897	9
Dog	1	Chr1:5748787-Chr1:6684616	935829	81
Dog	1	Chr1:23404515-Chr1:23838138	433623	46
Dog	1	Chr1:39129731-Chr1:41072047	1942316	50
Dog	1	Chr1:68965472-Chr1:70174300	1208828	67
Dog	1	Chr1:112724434-Chr1:113156643	432209	46
Dog	2	Chr2:13857951-Chr2:14169959	312008	25
Dog	3	Chr3:19767961-Chr3:20252413	484452	45
Dog	3	Chr3:31601497-Chr3:43759397	12157900	442
Dog	3	Chr3:54123144-Chr3:63761534	9638390	456
Dog	3	Chr3: 69490781-Chr3:70570090	1079309	101
Dog	3	Chr3:79471589-Chr3:81026210	1554621	116
Dog	3	Chr3:86606033-Chr3:86716483	110450	11
Dog	5	Chr5:10580391-Chr5:11148074	567683	65
Dog	5	Chr5:17620318-Chr5:17668846	48528	7
Dog	5	Chr5:34883138-Chr5:35970901	1087763	105
Dog	5	Chr5:51578116-Chr5:52242709	664593	52
Dog	5	Chr5:56485412-Chr5:57299415	814003	61
Dog	5	Chr5:77590145-Chr5:77734981	144836	9
Dog	6	Chr6:12715599-Chr6:14029977	1314378	102
Dog	6	Chr6:25253982-Chr6:42190103	16936121	961
Dog	6	Chr6:44759043-Chr6:45586589	827546	68
Dog	6	Chr6:50704404-Chr6:50979757	275353	18
Dog	6	Chr6:53244206-	55241	4

		Chr:53299447		
Dog	6	Chr6:60537553 Chr6:61712345	1174792	129
Dog	6	Chr6:76394711- Chr6:76718149	323438	35
Dog	7	Chr7:6694892- Chr7:6971977	277085	40
Dog	7	Chrr7:7897328 Chr7:8775912	878584	98
Dog	7	Chr7:9931429- Chr7:11943965	2012536	63
Dog	7	Chr7:54804377- Chr7:55720493	916116	37
Dog	7	Chr7:64010039- Chr7:64229287	219248	20
Dog	8	Chr8:25920235- Chr8:26305548	385313	26
Dog	8	Chr8:32423625- Chr8:34374315	1950690	18
Dog	8	Chr8:44137702- Chr8:49832920	5695218	389
Dog	8	Chr8:51008743- Chr8:52296505	1287762	115
Dog	9	Chr9:53584297- Chr9:54174243	589946	66
Dog	10	Chr10:4067616- Chr10:4266476	198860	12
Dog	10	Chr10:7717109- Chr10:11351389	3634280	308
Dog	10	Chr10:19040316- Chr10:20390793	1350477	108
Dog	10	Chr10:37423111- Chr10:37570327	147216	13
Dog	10	Chr10:47236490- Chr10:47433802	197312	20
Dog	10	Chr10:54292371- Chr10:55476999	1184628	119
Dog	11	Chr11: 64036892- Chr11:64317903	281011	33
Dog	12	Chr12:44452181- Chr12:45371384	919203	126
Dog	13	Chr13:32983318- Chr13:34795708	1812390	192
Dog	14	Chr14:22356706- Chr14:26370694	4013988	103
Dog	14	Chr14:53874869- Chr14:54509772	634903	35
Dog	15	Chr15:23199- Chr15:782801	(not consecutive)	40
Dog	15	Chr15:36893368- Chr15:37109514	216146	26

Dog	15	Chr15:39487837-Chr15:40272258	784421	69
Dog	16	Chr16:14394444-Chr16:14587375	192931	28
Dog	17	Chr17:17803341-Chr17:17939635	136294	16
Dog	17	Chr17:35957253-CHr17:35997231	39978	17
Dog	18	Chr18: 6140478-Chr18:8065831	1925353	171
Dog	18	Chr18: 19814812-Chr18:21788776	1973964	100
Dog	18	Chr18: 23335052-Chr18: 24160679	825627	73
Dog	19	Chr19:8729374-Chr19: 10693457	1964083	105
Dog	19	Chr19:12257809-Chrr19:12384702	126893	11
Dog	19	Chr19:27012387-Chr19:27424176	411789	34
Dog	19	Chr19:29374978-Chr19:29830184	455206	48
Dog	19	Chr19:38524678-Chr19:38627392	102714	12
Dog	19	Chr19: 46784911-Chr19:46883923	99012	10
Dog	20	Chr20: 27266030-Chr20: 28160363	894333	75
Dog	21	Chr21: 41097834-Chr21: 42363927	1266093	159
Dog	22	Chr22: 7070084-Chr22: 10776365	3706281	307
Dog	22	Chr22: 15454292-Chr22:15467321	13029	2
Dog	22	Chrr22:17013031 Chr22:17679668	666637	66
Dog	22	Chr22:19971105-Chr22:20526613	555508	31
Dog	23	Chr23: 32196788-Chr23: 32688929	492141	42
Dog	24	Chr24: 3840685-Chr24: 5285896	1445211	122
Dog	24	Chr24: 14538712-Chr24: 14930769	392057	36
Dog	24	Chr24: 20949005-Chr24: 22299970	1350965	151
Dog	24	Chr24: 39471356-Chr24: 39713042	241686	32
Dog	25	Chr25: 8960137-Chr25:9051243	91106	8
Dog	26	Chr26: 12725431-	351977	16

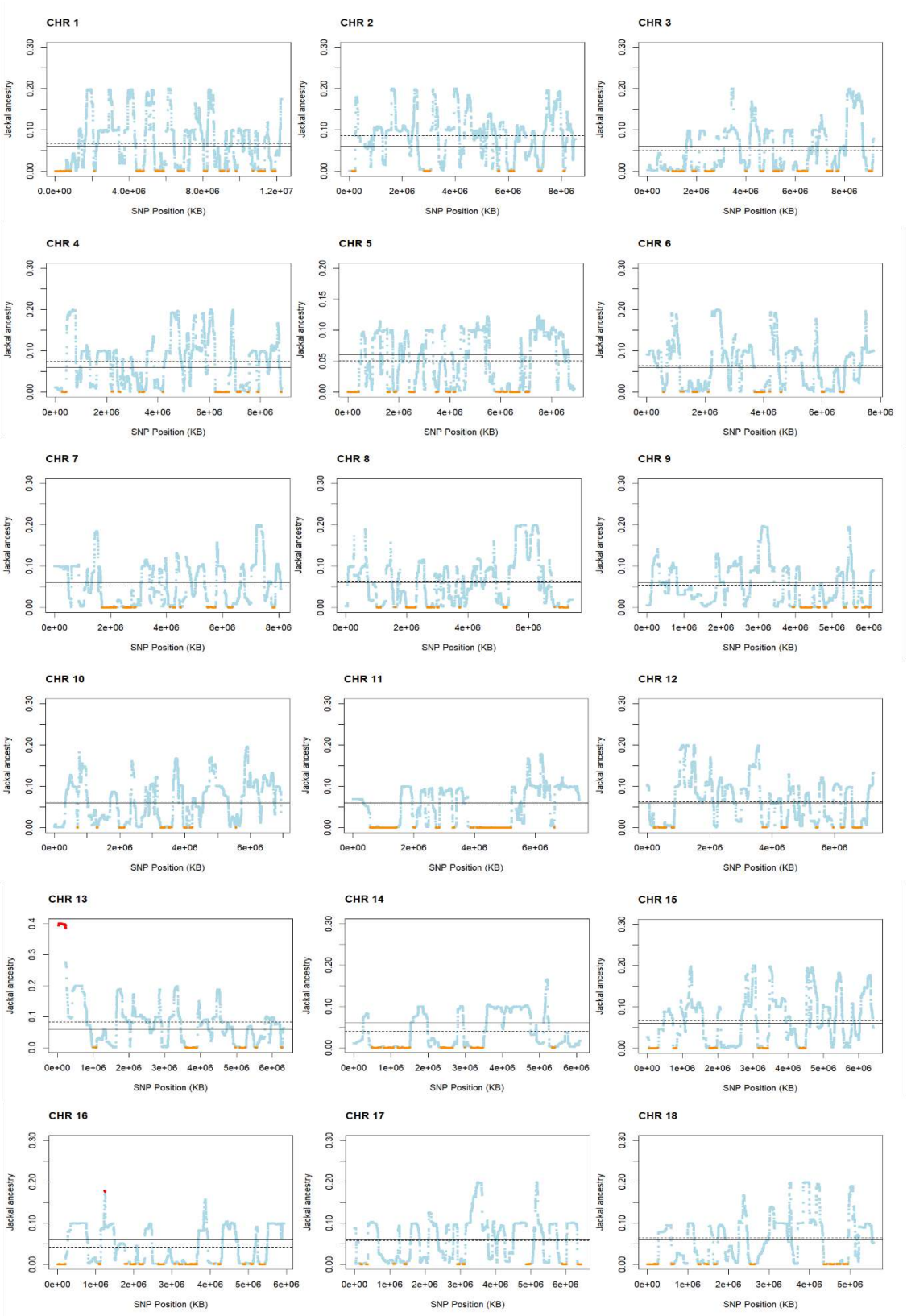
		Chr26:13077408		
Dog	27	Chr27: 5589715-Chr27:7198538	1608823	174
Dog	27	Chr27:31317852-Chr27:32961249	1643397	46
Dog	31	Chr31:27310539-Chr31:27576388	265849	30
Dog	36	Chr36: 7551774-Chr36:7922640	370866	48

Table S3. 5. Chromosomal blocks with underrepresentation of introgressed ancestry in the JD dataset

Population	**chr**	**Chromosomal block position**	**Block size**	**N SNPs**
Jackal	7	Chr7:41403776-Chr7:42217202	813426	59
Jackal	24	Chr24:25491-CHr24:1059902	(not consecutive)	73
Jackal	24	Chr24:4330346-Chr24:5092398	762052	46
Jackal	24	Chr24:22998333-Chr24:23112810	114477	11
Jackal	30	Chr30: 2351996-Chr30:2735801	383805	41

Table S3. 6. Chromosomal blocks with underrepresentation of introgressed ancestry in the WJ dataset

Population	**chr**	**Chromosomal block position**	**Block size**	**N SNPs**
Jackal	7	Chr7:41666086-Chr7:42304254	490268	18
Jackal	30	Chr30: 1755189-Chr30:2267036	511847	42


CHR 1
CHR 2
CHR 3
CHR 4
CHR 5
CHR 6
CHR 7
CHR 8
CHR 9
CHR 10
CHR 11
CHR 12
CHR 13
CHR 14
CHR 15
CHR 16
CHR 17
CHR 18
Jackal ancestry
SNP Position (KB)

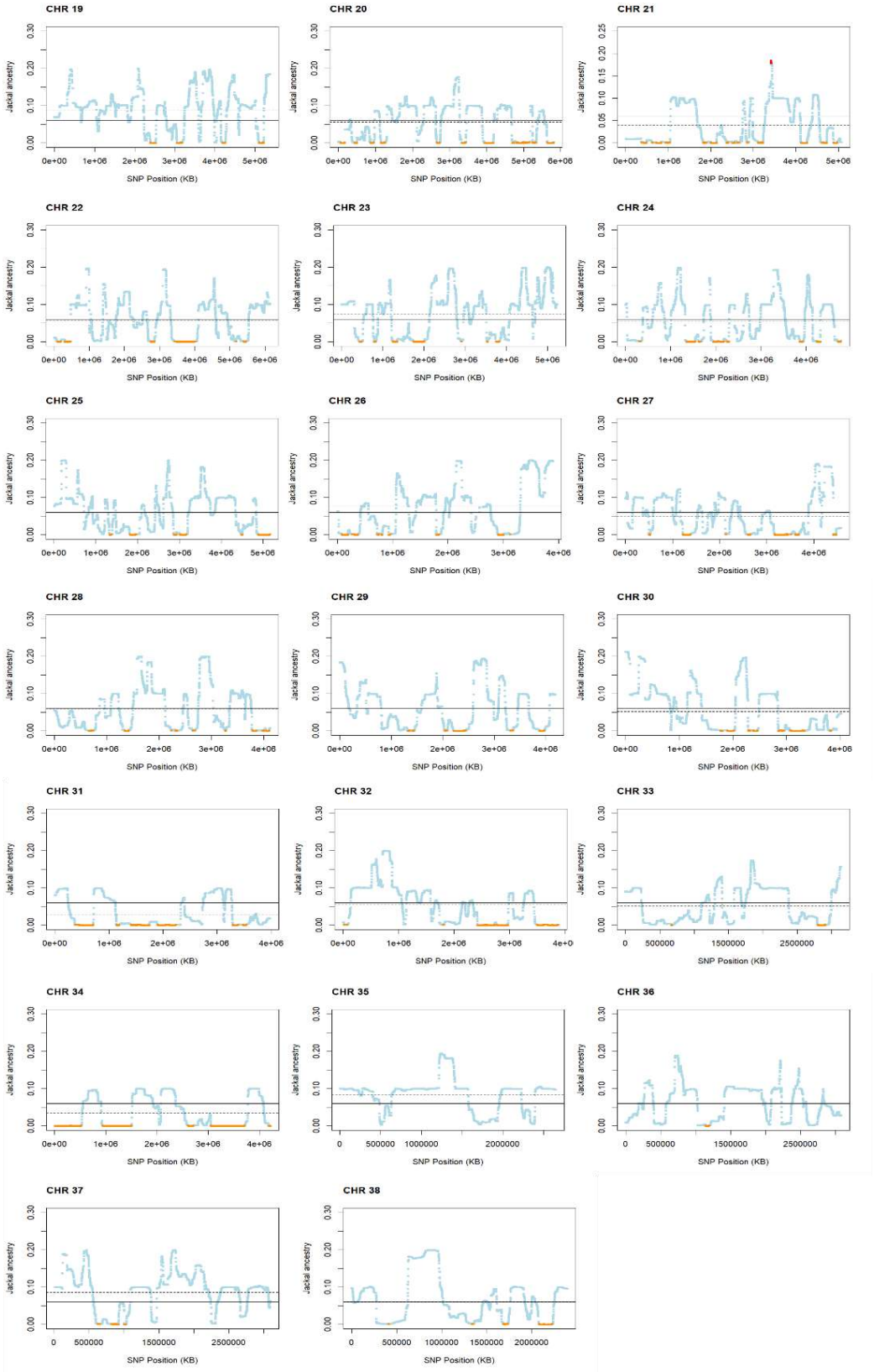


Fig. S.3. 1. Distribution of jackal ancestry in admixed wolves. The x-axis represents SNP order along each autosomal chromosome, and the y-axis represents the proportion of jackal admixture in admixed wolf. The solid horizontal line shows the mean jackal admixture across autosomal chromosomes, and the dotted horizontal line shows the mean jackal admixture within each chromosome. Chromosomal blocks with overrepresented and underrepresented jackal ancestry are marked in red and orange respectively.

Table S3. 7. Results of Gene Ontology analysis carried out for only genes that were located in the overlapped chromosomal blocks between the present study and the earlier study (Pilot et al., 2021) in wolves and dogs. The significance was determined using the Benjamini-Hochberg correction (BH) and the more conservative g:SCS (Set Counts and Sizes) false-discovery rate correction method. Only P-values below 0.05 are shown.

Canids	Reference genome	Go source	Term name	Term ID	P (FDR)	P (g:SCS)
Dog	Dog	GO:MF	N,N-dimethylaniline monooxygenase activity	GO:0004499	2.24E-07	2.78E-07
Dog	Dog	GO:MF	oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen, NAD(P)H as one donor, and incorporation of one atom of oxygen	GO:0016709	0.000	0.000
Dog	Dog	GO:MF	flavin adenine dinucleotide binding	GO:0050660	0.003	0.010
Dog	Dog	GO:MF	NADP binding	GO:0050661	0.004	0.018
Dog	Dog	GO:MF	monooxygenase activity	GO:0004497	0.005	0.034
Dog	Dog	GO:MF	hypotaurine monooxygenase activity	GO:0047822	0.005	0.031
Dog	Dog	GO:MF	oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen	GO:0016705	0.009	
Dog	Dog	GO:BP	regulation of primary metabolic process	GO:0080090	0.000	0.000
Dog	Dog	GO:BP	regulation of RNA biosynthetic process	GO:2001141	0.001	0.001
Dog	Dog	GO:BP	regulation of DNA-templated transcription	GO:0006355	0.001	
Dog	Dog	GO:BP	DNA-templated transcription	GO:0006351	0.001	0.002
Dog	Dog	GO:BP	regulation of RNA metabolic process	GO:0051252	0.001	0.003
Dog	Dog	GO:BP	regulation of metabolic process	GO:0019222	0.002	0.007
Dog	Dog	GO:BP	regulation of nucleobase-containing compound metabolic process	GO:0019219	0.002	0.008
Dog	Dog	GO:BP	regulation of macromolecule metabolic process	GO:0060255	0.010	
Dog	Dog	GO:BP	regulation of gene expression	GO:0010468	0.012	
Dog	Dog	GO:BP	NADPH oxidation	GO:0070995	0.014	
Dog	Dog	GO:BP	regulation of macromolecule biosynthetic process	GO:0010556	0.015	
Dog	Dog	GO:BP	regulation of biosynthetic process	GO:0009889	0.015	
Dog	Dog	GO:BP	negative regulation of fatty acid oxidation	GO:0046322	0.018	
Dog	Dog	GO:BP	embryonic organ development	GO:0048568	0.049	
Dog	Dog	GO:BP	alkanesulfonate biosynthetic process	GO:0046305	0.049	
Dog	Dog	GO:BP	taurine biosynthetic process	GO:0042412	0.049	
Dog	Dog	GO:BP	embryonic morphogenesis	GO:0048598	0.049	
Wolf	Dog	GO:MF	corticotropin receptor activity	GO:0004978	0.014	
Wolf	Dog	GO:MF	MHC class Ib protein complex binding	GO:0023025	0.014	
Wolf	Dog	GO:MF	MHC class Ib protein binding, via antigen binding groove	GO:0023030	0.014	
Wolf	Dog	GO:MF	mRNA 5'-cap (guanine-N7-)-methyltransferase activity	GO:0004482	0.016	
Wolf	Dog	GO:MF	MHC class Ib protein binding	GO:0023029	0.017	
Wolf	Dog	GO:MF	melanocortin receptor activity	GO:0004977	0.018	
Wolf	Dog	GO:MF	MHC protein complex binding	GO:0023023	0.025	
Wolf	Dog	GO:MF	mRNA methyltransferase activity	GO:0008174	0.038	
Wolf	Dog	GO:MF	MHC protein binding	GO:0042287	0.038	

Chapter 4

Table S.4. 1. The important measure for each variable of the WD dataset using East Asia samples according to %IncMSE and IncNodePurity

Variables	East_Asia..IncMSE	East_Asia.IncNodePurity
hf	35.122	36.743
bio_6	22.647	22.310
bio_4	22.529	21.580
bio_18	15.430	13.004
bio_12	18.402	11.654
bio_13	22.131	10.786
bio_19	23.617	10.323
bio_5	25.364	9.761
roughness	20.512	7.751
elevation	17.668	7.614

Table S.4. 2. The important measure for each variable of the WD dataset using North East Europe samples according to %IncMSE and IncNodePurity

Variables	North_East_Europe..IncMSE	North_East_Europe.IncNodePurity
hf	35.826	35.764
bio_6	22.893	21.472
bio_4	21.905	21.645
bio_18	15.845	12.758
bio_12	18.652	14.206
bio_13	22.660	10.720
bio_19	22.163	9.881
bio_5	25.015	9.179
roughness	20.912	8.323
elevation	19.957	7.682

Table S.4. 3. The important measure for each variable of the WD dataset using Central Europe samples according to %IncMSE and IncNodePurity

Variables	Central_Europe..IncMSE	Central_Europe.IncNodePurity
hf	35.137	37.826
bio_6	23.174	21.687
bio_4	22.189	20.896
bio_18	16.765	12.781
bio_12	18.225	12.155
bio_13	23.007	11.239
bio_19	24.573	9.818
bio_5	23.107	9.197
roughness	20.548	8.522
elevation	18.632	7.356

Table S.4. 4. The important measure for each variable of the WD dataset using West Europe samples according to %IncMSE and IncNodePurity

Variables	West_Europe..IncMSE	West_Europe.IncNodePurity
hf	36.714	36.698
bio_6	23.336	22.694
bio_4	22.728	21.464
bio_18	16.204	13.920
bio_12	18.703	11.770
bio_13	20.727	10.800
bio_19	23.901	9.481
bio_5	25.818	9.676
roughness	21.121	7.765
elevation	17.979	7.272

Table S.4. 5. The important measure for each variable of the WD dataset using Caucasus samples according to %IncMSE and IncNodePurity

Variables	Caucasus..IncMSE	Caucasus.IncNodePurity
hf	36.720	36.957
bio_6	23.404	23.682
bio_4	23.954	20.392
bio_18	14.978	12.147
bio_12	17.953	13.003
bio_13	22.298	11.505
bio_19	23.303	9.629
bio_5	24.045	9.051
roughness	21.035	7.999
elevation	17.110	7.252

Table S.4. 6. The important measure for each variable of the WD dataset using North Asia samples according to %IncMSE and IncNodePurity

Variables	North_Asia..IncMSE	North_Asia.IncNodePurity
hf	35.632	37.659
bio_6	22.374	20.937
bio_4	23.655	24.982
bio_18	16.028	11.588
bio_12	19.065	11.910
bio_13	21.997	10.402
bio_19	23.863	9.533
bio_5	24.414	9.168
roughness	18.661	7.902
elevation	19.191	7.436

Table S.4. 7. The important measure for each variable of the JD dataset using East Asia samples according to %IncMSE and IncNodePurity

Variables	East_Asia..IncMSE	East_Asia.IncNodePurity
bio_15	35.127	33.921
bio_1	34.395	28.380
hf	35.980	27.222
elevation	39.970	21.870
bio_4	24.093	15.464
bio_17	24.415	15.118
roughness	26.772	13.366
bio_18	24.643	9.722
bio_13	25.702	9.652
bio_19	20.895	8.460
bio_12	20.625	8.225
shdi	16.780	6.742
escape	6.298	1.134

Table S.4. 8. The important measure for each variable of the JD dataset using Caucasus samples according to %IncMSE and IncNodePurity

Variables	Caucasus..IncMSE	Caucasus.IncNodePurity
bio_15	34.676	33.027
bio_1	33.328	25.471
hf	39.169	27.446
elevation	40.987	22.082
bio_4	23.785	15.428
bio_17	25.054	18.173
roughness	26.366	13.294
bio_18	24.346	9.805
bio_13	23.182	8.888
bio_19	22.113	8.777
bio_12	20.964	8.946
shdi	17.072	6.672
escape	5.604	1.146

Table S.4. 9. The important measure for each variable of the JD dataset using North East Europe samples according to %IncMSE and IncNodePurity

Variables	North_East_Europe..IncMSE	North_East_Europe.IncNodePurity
bio_15	35.975	35.125
bio_1	33.747	26.091
hf	37.255	27.174
elevation	44.055	21.573
bio_4	24.393	14.093
bio_17	24.800	16.197
roughness	25.356	13.493
bio_18	23.202	10.293
bio_13	24.614	8.789
bio_19	21.283	8.995
bio_12	20.407	8.769
shdi	17.613	7.114
escape	6.396	1.207

Table S.4. 10. The important measure for each variable of the JD dataset using Central Europe samples according to %IncMSE and IncNodePurity

Variables	Central_Europe..IncMSE	Central_Europe.IncNodePurity
bio_15	36.572	33.703
bio_1	34.402	25.684
hf	40.967	28.410
elevation	40.010	21.256
bio_4	23.976	14.001
bio_17	26.022	17.443
roughness	27.451	13.614
bio_18	23.410	9.947
bio_13	23.243	8.946
bio_19	21.542	8.571
bio_12	20.338	9.011
shdi	16.947	7.212
escape	6.556	1.287

Table S.4. 11. The important measure for each variable of the JD dataset using North Asia samples according to %IncMSE and IncNodePurity

Variables	North_Asia..IncMSE	North_Asia.IncNodePurity
bio_15	33.947	34.695
bio_1	34.465	26.310
hf	39.453	27.570
elevation	39.376	21.919
bio_4	24.580	14.361
bio_17	23.626	16.450
roughness	27.141	12.938
bio_18	23.545	10.181
bio_13	25.264	9.612
bio_19	22.539	8.628
bio_12	20.814	8.582
shdi	16.499	6.979
escape	5.002	1.093

Table S.4. 12. The important measure for each variable of the JD dataset using West Europe samples according to %IncMSE and IncNodePurity

Variables	West_Europe..IncMSE	West_Europe.IncNodePurity
bio_15	36.097	33.275
bio_1	34.783	26.202
hf	38.586	27.989
elevation	42.766	22.280
bio_4	23.485	14.994
bio_17	24.610	16.504
roughness	26.589	13.428
bio_18	23.427	9.440
bio_13	23.015	9.357
bio_19	21.078	8.438
bio_12	20.716	9.178
shdi	16.398	6.845
escape	5.647	1.165

Table S.4. 13. R-squared and RMSE values for different populations using the WD and JD datasets

	WD dataset		**JD dataset**	
Populations	**RMSE**	**R^2**	**RMSE**	**R^2**
Caucasus	0.031	0.908	0.028	0.905
Central_Europe	0.011	0.927	0.009	0.936
East_Asia	0.015	0.957	0.031	0.901
North_Asia	0.017	0.954	0.016	0.932
North_East_Europe	0.018	0.920	0.009	0.955
West_Europe	0.002	0.991	0.0004	0.999

Dr hab. Małgorzata Pilot
Faculty of Biology
University of Gdańsk
ul. Wita Stwosza 59
80-308 Gdańsk, Poland
E-mail contact:
malgorzata.pilot@ug.edu.pl

26th of May 2025

Authorship statement

I confirm that I am a co-author of the paper:

Adavoudi, R., & Pilot, M. (2021). Consequences of hybridization in mammals: A systematic review. Genes, 13(1), 50.

I declare that I contributed to the study design, supervised its implementation, and revised the manuscript.

Małgorzata Pilot

dr hab. Małgorzata Pilot

Uniwersytet Gdański | 55 LAT

Gdańsk, 26.05. 2025.

I confirm that I am a co-author of the paper:

Adavoudi, R., & Pilot, M. (2021). Consequences of hybridization in mammals: A systematic review. *Genes*, *13*(1), 50.

I declare that I carried out the literature review and wrote the first draft of the manuscript.

Roya Adavoudi

Rada Dyscypliny Nauki biologiczne Wydział Biologii Uniwersytet Gdański	tel. +48 58 523 60 08 e-mail: iwona.mucha@ug.edu.pl	ul. Wita Stwosza 59 80-308 Gdańsk	

Appendix

Codes For Chapter 2

```
=====================================================

             Initial data processing

=====================================================

####### Plink software #########

Software: Plink  version 1.9 and 2.0

http://pngu.mgh.harvard.edu/purcell/plink/

Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D, Maller J,
Sklar P, de Bakker PIW, Daly MJ & Sham PC (2007)PLINK: a toolset for whole-
genome association and population-based linkage analysis. American Journal of
Human Genetics, 81.

---------------------------------------------------------------------

####### Filtering dataset #######


plink --bfile WDJ_merged --missing -- dog -- out WDJ_merged_missingness

plink --bfile WDJ_merged --dog --mind 0.2 --geno 0.1 -out WDJ_merged_QC1 --
make-bed --recode

plink --bfile WDJ_merged_QC1 --dog --out WDJ_merged_frequency --freq


plink --bfile WDJ_merged_QC1 --dog --maf 0.01 --out WDJ_merged_QC2 --recode --
make-bed

plink --bfile WDJ_merged_QC2 --dog --indep-pairwise 50 3 0.1 --out
WDJ_merged_linkage

plink --bfile WDJ_merged_QC2 --dog --extract WDJ_merged_linkage.prune.in --
recode --make-bed --out WDJ_merged_QC_LD

plink --bfile Wolf_QC --dog --make-king-table --out Wolf_QC_King_table

Plink --bfile Wolf_QC --dog --king-cutoff 0.48 --out Wolf_QC_king


=============================================================================
=
                    Global Ancestry Analysis
=============================================================================
=

######### PCA #########


plink --bfile WDJ_QC_LD --dog --pca --out WJD.pca


#######show result of PCA in R #############
```

```
pca <- read.table(file = "WJD.pca_.eigenvec",header = F,stringsAsFactors = F)
names(pca) <- c("FID","ID","PC1","PC2")

mypop <- pca$FID

mypop[mypop == "Wolf"] <- "blue"
mypop[mypop == "Jackal"] <- "red"
mypop[mypop == "Dog"] <- "green"

plot(pca$PC1,pca2$PC2,col=mypop)
legend("bottomleft",legend = unique(pca$FID),fill = c("blue","green", "red"))

=========================================================

######### ADMIXTURE #########


#ADMIXTURE software v. 1.3.0

#https://dalexander.github.io/admixture/download.html

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

-----------------------------------------------------------------------

#######RUN ADMIXTURE in LINUX #############

for i in {1..10}; do ./admixture --cv WDJ_QC_LD.bed $i >log${i}.out; done


#######show result of admixture in R/Barpllat #############

wjdK3 <- read.table("Wolf_Dog_Jackal_QC.3.Q")

#Make admixture barplot#

admixCol <- c( "blue","red", "green")
barplot(t(as.matrix(wjdK3)), col=admixCol,border=NA, ylab="Admixture
proportions", cex.main=1)
axis(1, at = c(400, 700, 1000), labels =c("Dog", "Wolf", "Jackal"), las=1,
tck=0, col="white", cex.axis=0.9)
abline(v=c(563,908))
axis(1, at = c(15, 45), labels =c("Dog", "Wolf", "Jackal"), las=1, tck=0,
col="white", cex.axis=0.9)

-------------------------------------------------------------

=====================================================================
                  Local Ancestry Analysis
=====================================================================

          ---------------------------------------------------------
          #########Local Ancestry in adMixed Populations (LAMP)########
          ---------------------------------------------------------

#Software Version: LAMP - Release 2.4
#http://lamp.icsi.berkeley.edu/lamp/
```

```
#S. Sankararaman, S. Sridhar, G. Kimmel, and E. Halperin. Estimating local
ancestry in admixed populations. American Journal of Human Genetics, 8(2):290-
303, 2008.


#=========================================
#RENAME DUPLICATED SNPs FOR LAMP ANALYSIS
#=========================================

bim <- read.table(file = "Wolf_dog_QC.bim",header = F)
dup.SNP <- which(duplicated(bim$V4))
bim$V7 <- bim$V4
bim$V7[dup.SNP] <- bim$V7[dup.SNP] + 1

write.table(x = bim[,c(2,7)],file = "update.map",quote = F,row.names =
F,col.names = F)
read.table("update.map")

# The positions need to be updated via PLINK
# I created the new binary files under name "Wolf_Dog_QC_updated"
system("./plink --dog --nonfounders --bfile  Wolf_Dog_QC --update-map
update.map --make-bed --out Wolf_Dog_QC_updated")

bim_updated <- read.table(file = "Wolf_Dog_QC_updated.bim",header = F)
which(duplicated(bim_updated$V4))

chromosome <- 38
# extract a chromosome (for each chromosome, the chromsome number and output
file name should be changed)
for (i in 1:chromosome){
  dir.create(path = paste("CHR",i,sep = ""))

  command <- "./plink --dog --bfile Wolf_dog_QC_updated"
  command <- paste(command,"--chr",i,"--make-bed --
out",paste("Wolf_dog_QC_updated_CHR",i,sep = ""))
  system(command)

}
  ##for i in {1..38}; do plink --bfile Wolf_dog_QC_updated --dog --chr ${i} --
out Wolf_dog_QC_updated_CHR${i} --make-bed --recode; done


  # create the recodeA recodel-allele input file of PLINK

  for(i in 1:38){
    bim <- read.table(file = paste("Wolf_dog_QC_updated_CHR",i,".bim",sep =
""),header = F,stringsAsFactors = F)
    write.table(x = cbind(bim$V2,"2"),file = paste("recodeAB_CHR",i,".txt",sep
= ""),quote = F,row.names = F,col.names = F)
  }

  # recode to genotypes (for each chromosome the input and output files names
should be changed) # this can also be written in a loop.
#plink --dog --bfile Wolf_dog_QC_updated_CHR1 --reference-allele
recodeAB_CHR1.txt --recodeA --out Wolf_dog_QC_updated_CHR1
#Or in loop
```

```
#for i in {1..38}; do plink --bfile Wolf_dog_QC_updated_CHR${i} --reference-
allele recodeAB_CHR${i}.txt --recodeA --out Wolf_dog_QC_updated_CHR${i} --dog;
done

 # ./plink --dog --bfile Wolf_dog_QC_updated-CHR1 --reference-allele
recodeAB_CHR1.txt --recodeA --out Wolf_dog_QC_updated_CHR1


#=====================================================================
  #PREPARE LAMP INPUT FILES USING THE SHELL SCRIPT prepfileforlamp.sh:
#=====================================================================

  #BASH SCRIPT (THIS IS A SEPARATE FILE and only copied here for information.
The bash script needs to be executed in a terminal/cmd window for each
CHROMOSOME:
  sh prefileforlamp.sh 1 Wolf_dog_QC
  sh prefileforlamp.sh 2 Wolf_dog_QC
  sh prefileforlamp.sh 3 Wolf_dog_QC

  #(...)

  #BASH SCRIPT:
  #!/bin/bash

  CHR=$1
  outfile=$2

  #mkdir CHR${CHR}
  #cd CHR${CHR}


  ## make LAMP input file
  cat  ${outfile}_CHR${CHR}.raw | awk 'NR > 1' | cut -d' ' -f 7- | sed 's/NA/-
1/g' > ${outfile}_CHR${CHR}_LAMPGENO.txt
  cat ${outfile}_CHR${CHR}.bim | awk '{print $4}' >
${outfile}_CHR${CHR}_LAMPMAP.txt

  echo "populations=2
genofile=GENOTYPES
posfile=SNPPOSITION
outputancestryfile=OUTFILE
offset=0.2
recombrate=1e-10
generations=10
alpha=0.53,0.47
ldcutoff=0.1" > config.txt

  #### MAKE LAMP configeration file
  cat config.txt | awk -v G=${outfile}_CHR${CHR}_LAMPGENO.txt -v
M=${outfile}_CHR${CHR}_LAMPMAP.txt -v O=${outfile}_CHR${CHR}_results.txt
'{gsub(/GENOTYPES/,G);gsub(/SNPPOSITION/,M);gsub(/OUTFILE/,O); print}' \
  >  ${outfile}_CHR${CHR}configfile.txt
----------------------------------------

  ### runing LAMP
  #lamp ${outfile}_CHR${CHR}configfile.txt

  #cd ..
```

```
#================================================================
#RUN LAMP on LINUX SERVER (or WINDOWS) FOR EACH CHROMOSOME
#================================================================

./lamp <config.txt> # outputfile = results.txt

### visualization the results of LAMP ###

Sudo ./ lamproot/bin/ generategraph.sh ancestry_lamp4.out

-----------------------------------------------------------------------

=======================================================
    ### Efficient Local Ancestry Inference(ELAI) ###
=======================================================


#Software Version: ELAI - version 11.3
#http://github.com/haplotype/ELAI/blob/main/elai-linux-multi-threading.tar.gz.

#Guan Y. Detecting structure of haplotypes and local ancestry. Genetics. 2014
Mar;196(3):625-42. doi: 10.1534/genetics.113.160697. Epub 2014 Jan 3. PMID:
24388880; PMCID: PMC3948796.
---------------------------------------------------------------


####### Keep reference set and admixed individuals ##########

plink --bfile Wolf_dog_QC --dog --keep Dog_pure.txt --out Pure_dog_QC --recode
--make-bed

plink --bfile Wolf_dog_QC --dog --keep Pure_wolf.txt --out Pure_wolf_QC --
recode --make-bed

plink --bfile Wolf_dog_QC --dog --keep Admixed.txt --out Admixed_QC --recode -
-make-bed




########### seprate chromosomes in each dataset ##############

 for i in {1..38}; do plink --bfile Admixed_QC --dog --make-bed --chr ${i} --
out Admixed_QC_CHR${i} --recode; done

for i in {1..38}; do plink --bfile Pure_wolf_QC --dog --make-bed --chr ${i} --
out Pure_wolf_QC_CHR${i} --recode; done

for i in {1..38}; do plink --bfile Pure_dog_QC --dog --make-bed --chr ${i} --
out Pure_dog_QC_CHR${i} --recode; done



########## converting to Bimbam format each chromosome ####################


for i in {1..38}; do plink --bfile Pure_jackal_QC_CHR${i} --dog --recode
bimbam  --out Pure_dog_QC_CHR${i}; done
```

```
for i in {1..38}; do plink --bfile Pure_wolf_QC_CHR${i} --dog --recode bimbam
--out Pure_wolf_QC_CHR${i}; done

for i in {1..38}; do plink --bfile Admixed_QC_CHR${i} --dog --recode bimbam  -
-out Admixed_wolf_dog_QC_CHR${i}; done


#### Move files "geno" and "pos" to diferent foldes for moving to Linux ###

for i in {1..38}; do mv "Admixed_wolf_jackal_QC_CHR${i}.recode.pos.txt"
"Run_elai_3/"; done

for i in {1..38}; do mv "Admixed_wolf_jackal_QC_CHR${i}.recode.geno.txt"
"Run_elai_3/"; done


for i in {1..38}; do mv "Pure_wolf_QC_CHR${i}.recode.pos.txt" "Run_elai_2/";
done
for i in {1..38}; do mv "Pure_wolf_QC_CHR${i}.recode.geno.txt" "Run_elai_2/";
done

for i in {1..38}; do mv "Pure_jackal_QC_CHR${i}.recode.pos.txt" "Run_elai/";
done
for i in {1..38}; do mv "Pure_jackal_QC_CHR${i}.recode.geno.txt" "Run_elai/";
done


######################## Run ELAI  ########################

cd Deskop
cd ELAI

./elai-lin -g Pure_jackal_QC_CHR1.recode.geno.txt -p 10 -g
Pure_wolf_QC_CHR1.recode.geno.txt -p 11 -g
Admixed_jackal_wolf_QC_CHR1.recode.geno.txt -p 1 -pos
Pure_jackal_QC_CHR1.recode.pos.txt -s 30 -mg 10 -C 2 -c 10 -R 45 -o
wolf_jackal_CHR1


########################## Run in Loop ########################


for i in {1..38}; do ./elai-lin -g Pure_jackal_QC_CHR${i}.recode.geno.txt -p
10 -g Pure_dog_QC_CHR${i}.recode.geno.txt -p 11 -g
Admixed_wolf_jackal_QC_CHR${i}.recode.geno.txt -p 1 -pos
Pure_jackal_QC_CHR${i}.recode.pos.txt -s 30 -mg 10 -C 2 -c 10 -R 30 -o
Wolf_jackal_CHR${i}; done


for i in {1..38}; do ./elai-lin -g Pure_dog_QC_CHR${i}.recode.geno.txt -p 10 -
g Pure_wolf_QC_CHR${i}.recode.geno.txt -p 11 -g
Admixed_wolf_dog_QC_CHR${i}.recode.geno.txt -p 1 -pos
Pure_dog_QC_CHR${i}.recode.pos.txt -s 30 -mg 10 -C 2 -c 10 -R 30 -o
Wolf_dog_CHR${i}; done

for i in {1..38}; do ./elai-lin -g Pure_dog_QC_ph_CHR${i}.recode.geno.txt -p
10 -g Pure_jackal_QC_ph_CHR${i}.recode.geno.txt -p 11 -g
Admixed_dog_jackal_QC_ph_CHR${i}.recode.geno.txt -p 1 -pos
```

```
Pure_dog_QC_ph_CHR${i}.recode.pos.txt -s 30 -mg 10 -C 2 -c 10 -R 30 -o
dog_jackal_ph_CHR${i}; done


=====================================================
    ### Beagle (phasing datasets) ###
=====================================================
#Software Version:  Beagle 5.4

#http://faculty.washington.edu/browning/beagle/beagle.html

#Browning, B. L., Tian, X., Zhou, Y., & Browning, S. R. (2021). Fast two-stage
phasing of large-scale sequence data. The American Journal of Human Genetics,
108(10), 1880-1890.
---------------------------------------------------------------------

######## preparing input files for each species separately ########


plink --bfile WJD_updated_QC --dog --out Wolf_QC --keep Wolf_samples.txt --
make-bed --recode


plink --bfile WJD_updated_QC --dog --out Jackal_QC --keep Jackal_samples.txt -
-make-bed --recode


plink --bfile WJD_updated_QC --dog --out Dog_QC --keep Dog_samples.txt --make-
bed --recode

################################### create vcf.bgz file ################

./plink2 --bfile Dog_QC --dog --export vcf bgz --ref-from-fa
Canis_lupus_familiaris.CanFam3.fa --out Dog_QC


./plink2 --bfile Jackal_QC --dog --export vcf bgz --ref-from-fa
Canis_lupus_familiaris.CanFam3.fa --out Jackal_QC


./plink2 --bfile Wolf_QC --dog --export vcf bgz --ref-from-fa
Canis_lupus_familiaris.CanFam3.fa --out Wolf_QC


##############################  based on each chromosomes ##############

for i in {1..38}; do plink --bfile Wolf_QC  --dog --chr ${i} --out
Wolf_QC_${i} --recode --make-bed; done


for i in {1..38}; do plink --bfile Dog_QC  --dog --chr ${i} --out Dog_QC_${i}
--recode --make-bed; done


for i in {1..38}; do plink --bfile Jackal_QC  --dog --chr ${i} --out
Jackal_QC_${i} --recode --make-bed; done


################################### create vcf.bgz file ################
```

```
for i in {1..38}; do ./plink2 --bfile Wolf_QC_${i} --dog --export vcf bgz --
ref-from-fa Canis_lupus_familiaris.CanFam3.fa --out Wolf_QC_${i}; done


for i in {1..38}; do ./plink2 --bfile Dog_QC_${i} --dog --export vcf bgz --
ref-from-fa Canis_lupus_familiaris.CanFam3.fa --out Dog_QC_${i}; done


for i in {1..38}; do ./plink2 --bfile Jackal_QC_${i} --dog --export vcf bgz --
ref-from-fa Canis_lupus_familiaris.CanFam3.fa --out Jackal_QC_${i}; done


########################## create Map file (plink format) ################

Based on the recombination map file we will create map file.
Chr      name     cm       position
1        .        0.294986106      4361401
1        .
1        .

For 38 chromosome we will create these files.


#################### Run Beagle ########################

java -jar beagle.22Jul22.46e.jar gt=Dog_QC_1.vcf.gz out=Dog_QC_1_imputed
map=CHR1_map_file.txt interarion=1000 burnin=10


######################  Run in a loop ######################

for i in {1..38};do java -jar beagle.22Jul22.46e.jar gt=Dog_QC_${i}.vcf.gz
out=Dog_QC_${i}_imputed map=CHR${i}_map_file.txt burnin=25 iterations=1000;
done

---------------------------------------------------------------------

=======================================================
    ### Genome-Wide Haplotyping (GHap) ###
=======================================================

#Software Version:  GHap r package version 2

#https://cran.r-project.org/package=GHap and
https://bitbucket.org/marcomilanesi/ghap/src/master/

#Utsunomiya, Y. T., Milanesi, M., Barbato, M., Utsunomiya, A. T. H., Sölkner,
J., Ajmone-Marsan, P., & Garcia, J. F. (2020). Unsupervised detection of
ancestry tracks with the GHap r package. Methods in Ecology and Evolution,
11(11), 1448-1454.
-------------------------------------------------

######### preparing input for GHap ###################

 ##### Combine all chromosomes after phasing #####

##First, give them an index
```

```
bcftools index Wolf_QC_1_imputed.vcf.gz
bcftools index Wolf_QC_2_imputed.vcf.gz
bcftools index Wolf_QC_3_imputed.vcf.gz
....

bcftools index Wolf_QC_38_imputed.vcf.gz


############## Combine them by bcftools ########

bcftools concat Wolf_QC_{1..38}_imputed.vcf.gz -Oz -o Wolf_QC_imputed.vcf.gz


#do all these steps for jackal and dog....


##### merging two vcf.gz files by bcftools #####

####  giving them an index:

bcftools index Wolf_QC.vcf.gz
bcftools index Dog_QC.vcf.gz


###### Merginng #########

bcftools merge Wolf_QC.vcf.gz Dog_QC.vcf.gz > Wolf_dog_QC_ph.vcf


###### keep samples from VCF files using bcftools ###(if we want to separate
some individuals)

bcftools view -S dog_wolf.txt  -o indiv_wolf_dog.vcf  Wolf_dog_QC_ph.vcf

note: -s is a file that contains one column of ID of individuals that we want
to separate
-----------------------------------------------------------------

######## RUN GHap in Rstudi############


library("GHap")

# #### DO NOT RUN IF NOT NECESSARY ###

# converting vcf file to ghap phased file

ghap.vcf2phase(vcf.files = "Wolf_dog_QC_ph.vcf", sample.files =
"Sample_wolf_dog.txt", out.file = "ghap_input_converted_Wolf_dog")


# compressing phased file

ghap.compress (input.file = "ghap_input_converted_wolf_dog", out.file =
"ghap_input_converted_wolf_dog" )

# loading compressed phased file

phase <- ghap.loadphase(input.file= "ghap_input_converted_Wolf_dog")
```

```
# ### RUN ###

# # Calculate marker density

mrkdist <- diff(phase$bp)
mrkdist <- mrkdist[which(mrkdist > 0)]
density <- mean(mrkdist)

# # Generate blocks for admixture events up to g = 10 generations in the past
# # Assuming mean block size in Morgans of 1/(2*g)
# # Approximating 1 Morgan ~ 100 Mbp

g <- 10
window <- (100e+6)/(2*g)
window <- ceiling(window/density)
step <- ceiling(window/4)
blocks <- ghap.blockgen(phase, windowsize = window,
                        slide = step, unit = "marker")

#
# # BestK analysis

bestK <- ghap.anctrain(object = phase, K = 10, tune = TRUE)
plot(bestK$ssq, type = "b", xlab = "K",ylab = "Within-cluster sum of squares")

#
# Unsupervised analysis with best K
# Construction of prototype alleles using best K,default 10 independent runs

prototypes <- ghap.anctrain(object = phase, K = 2, iter.max = 1, nstart = 1)

prototypes <- ghap.anctrain(object = phase, K=2, method = "unsupervised")

# Prediction of haplotype ancestry, smoothing, plotting
hapadmix <- ghap.anctest(object = phase,
                         blocks = blocks,
                         prototypes = prototypes,
                         test = unique(phase$id))

anctracks <- ghap.ancsmooth(object = phase, admix = hapadmix)
ghap.ancplot(ancsmooth = anctracks)


# Plot karyoplot, first make variables based on populations

Hybrid <- unique(phase$id[which(phase$pop == "Hybrid")])
Wolf_India <- unique(phase$id[which(phase$pop == "Wolf_India")])
Wolf_N_Asia <- unique(phase$id[which(phase$pop == "Wolf_N_Asia")])
Wolf_m_East <- unique(phase$id[which(phase$pop == "Wolf_m_East")])
Wolf_N_Europe <- unique(phase$id[which(phase$pop == "Wolf_N_Europe")])
Wolf_C_Europe <- unique(phase$id[which(phase$pop == "Wolf_C_Europe")])
Wolf_E_Eropue <- unique(phase$id[which(phase$pop == "Wolf_E_Europe")])
Dog_India <- unique(phase$id[which(phase$pop == "Dog_India")])
Dog_Bosnia <- unique(phase$id[which(phase$pop == "Dog_Bosnia")])
Dog_N_Asia <- unique(phase$id[which(phase$pop == "Dog_N_Asia")])
Dog_m_East <- unique(phase$id[which(phase$pop == "Dog_m_East")])
Dog_Russia <- unique(phase$id[which(phase$pop == "Dog_Russia")])
```

```
Dog_N_Europe <- unique(phase$id[which(phase$pop == "Dog_N_Europe")])
Dog_C_Europe <- unique(phase$id[which(phase$pop == "Dog_C_Europe")])
Dog_W_Europe <- unique(phase$id[which(phase$pop == "Dog_W_Europe")])
Dog_Morocco <- unique(phase$id[which(phase$pop == "Dog_morocco")])


# Plotting karyoplot separately for samples in each population, without ids
field plots should be done for all individuals, but with a lot of them Rstudio
is not showing all plots

ghap.karyoplot(ancsmooth = anctracks, ids=Hybrid, chr.line = 11, plot.line =
50, las=1, chr=NULL)
ghap.karyoplot(ancsmooth = anctracks, ids=Dog_India, chr.line = 11, plot.line
= 50, las=1, chr=NULL)
ghap.karyoplot(ancsmooth = anctracks, ids=Dog_m_East, chr.line = 11, plot.line
= 50, las=1, chr=NULL)
ghap.karyoplot(ancsmooth = anctracks, ids=Dog_N_Europe, chr.line = 11,
plot.line = 50, las=1, chr=NULL)
ghap.karyoplot(ancsmooth = anctracks, ids=Wolf_India, chr.line = 11, plot.line
= 50, las=1, chr=NULL)
ghap.karyoplot(ancsmooth = anctracks, ids=Wolf_C_Europe, chr.line = 11,
plot.line = 50, las=1, chr=NULL)
ghap.karyoplot(ancsmooth = anctracks, ids=Dog_C_Europe, chr.line = 11,
plot.line = 50, las=1, chr=NULL)
ghap.karyoplot(ancsmooth = anctracks, ids=Dog_W_Europe, chr.line = 11,
plot.line = 50, las=1, chr=NULL)
ghap.karyoplot(ancsmooth = anctracks, ids=Wolf_m_East, chr.line = 11,
plot.line = 50, las=1, chr=NULL)
ghap.karyoplot(ancsmooth = anctracks, ids=Wolf_N_Asia, chr.line = 11,
plot.line = 50, las=1, chr=NULL)
ghap.karyoplot(ancsmooth = anctracks, ids=Wolf_N_Europe, chr.line = 11,
plot.line = 50, las=1, chr=NULL)
.....

# Saving all karyplots (saving all visible plots from temporary Rstudio
folder)

plots.dir.path <- list.files(tempdir(), pattern="rs-graphics", full.names =
TRUE);
plots.png.paths <- list.files(plots.dir.path, pattern=".png", full.names =
TRUE)
file.copy(from=plots.png.paths, to="plots")

# Saving results

write.csv(anctracks[["haplotypes"]],file = "haplotypes.csv")
write.csv(anctracks[["proportions1"]],file = "proportions1.csv")
write.csv(anctracks[["proportions2"]],file = "proportions2.csv")

----------------------------------------------------------------

================================================================
                  Population Structure
================================================================

======================================================
 ### Discriminant Analysis of Principal Component(DAPC) ###
======================================================
#Software: adegenet 2.0.0
```

```
#https://adegenet.r-forge.r-project.org/files/tutorial-dapc.pdf#page=41.64
#Jombart T. and Ahmed I. (2011) adegenet 1.3-1: new tools for the analysis of
genome-wide SNP data. Bioinformatics. doi: 10.1093/bioinformatics/btr521
------------------------------------------------------------


# Working data set for DAPC
#system("plink --bfile Wolf_QC_king --dog --out Wolf_QC_King --export A")

# Load necessary libraries #


library(data.table)
jackal_qc_for_dapc<- fread(input = "Wolf_QC_king.raw", h=T)
write.table(Wolf_qc_for_dapc, "Wolf_QC_king.raw", col.names=T, row.names=F,
            quote=FALSE, sep=" ")

# Reading input file for DAPC
library(adegenet)

dapc_input <- read.PLINK("Wolf_QC_king.raw", parallel=F, sep="\t")

# K-means analysis on the principal components:
# Max K to test in K-means analysis
maxk <- 10
grp <- find.clusters(dapc_input, pca.select = "percVar",
                     perc.pca = 99, max.n.clust = maxk, choose.n.clust = TRUE)

  # 1st DAPC run to individuate the optimal number of principal components to
retain for not to incur into overfitting issues.

dapc <- dapc(dapc_input, grp$grp, pca.select = "percVar", perc.pca = 99, n.da
= 2)
print(dapc)

ascore <- optim.a.score(dapc)

# 2nd DAPC run with optimal number of PCs retained

dapc <- dapc(dapc_input, grp$grp, n.pca=ascore$best, n.da=length(grp$size) -
1)

print(dapc)

DAPC_ind.coord <- print (dapc$ind.coord)

write.csv(DAPC_ind.coord, 'D:/Roya/coor.csv')

windows()
# or: options( device = "windows" ); dev.new()
# quartz() # for MacOS
# x11() # for linux

scatter(dapc, bg="white", pch=21:23, cstar=0, col = c("lightsalmon2",
"darkseagreen3", "lightskyblue3", "lightpink2"))

scatter(dapc, 1, 1, bg = "white", col = c("lightsalmon2", "darkseagreen3",
"lightskyblue3", "lightpink2"),legend = F, cleg = 0.6, solid = 0.4)
```

```
############# Plotting DAPC results##########

windows()
# or:
# options(device = "windows"); dev.new()
# quartz() # for MacOS
# x11() # for linux

# background map

library(rgdal)
library(raster)
library(Sp)
library(ggplot2)

# loading elevation data for distribution map
E= raster("Elevation_masked_August_2023.tif")
plot(E)

# associating qc-ed individuals with DAPC cluster assignments and discriminant
function

DAPC_Wolf_QC_King.fam <- fread("Wolf_QC_King.fam")

colnames(DAPC_Wolf_QC_King.fam) <- c("FID", "ID", "ID_father", "ID_mother",
"Sex_code", "Phenotype")

head(DAPC_Wolf_QC_King.fam)

# are individuals in the 'coo_qced' and 'dapc' objects in the same order?

length(which(rownames(dapc$ind.coord)==coo_Wolf$ID))


# coordinates of the qc-ed individuals

library(plyr)

coo_qced <- match_df(coo_Wolf,DAPC_Wolf_QC_King.fam , on = "ID")

coo_qced$assign <- as.numeric(dapc$assign)

##as we have & clusters,ind.coord has 3 LD columns, then we should chose one
of them.
coo_qced$ind.coord <- as.vector(dapc$ind.coord [,c("LD3")])
coo_qced$pch[which(coo_qced$assign==1)] <- 21
coo_qced$pch[which(coo_qced$assign==2)] <- 22
coo_qced$pch[which(coo_qced$assign==3)] <- 23
#oo_qced$pch[which(coo_qced$assign==4)] <- 24

mycol <- colorRampPalette(colors = c("lightsalmon2", "darkseagreen3",
"lightskyblue3", "lightpink2"))

 coo_qced$col <- mycol(100)[as.numeric(cut(coo_qced$ind.coord, breaks = 100))]

head(coo_qced)

#each LD linear discriminat function
# genetic structure on the map
```

```
library(scales)
points(coo_qced$X, coo_qced$Y, pch=coo_qced$pch, bg=alpha(coo_qced$col, 0.6),
col="black", cex=1.4)

legend(x=130,y=10, legend=c("Cluster 1", "Cluster 2", "Cluster 3", "Cluster
4"), pch=c(21,24), bty="n", pt.bg="ghostwhite", pt.cex = 1.4)

title("DAPC - Wolf_LD3")

###these codes were used for jackals and dogs seperatly#######
----------------------------------------------------------------
```

Codes For Chapter 3

```
====================================================================
Finding Chromosomal blocks with overrepresentation  or underrepresentation of
introgressed variants
====================================================================

# Load necessary libraries #

library(data.table)
library(purrr)
library(dplyr)

#######CHR1
# we should use .ps21.txt file, the output of ELAI

dat <- fread(paste("Wolf_dog_CHR1.ps21.txt",sep=''), header=F,data.table=F)

##Selecting specific rows (removing First generation hybrids)##

data_modify <- dat[c(150:153, 155:163, 165:168, 170:249, 251:263, 265:266,
268:314),]

print(data_modify)

data3 <- data_modify/2

write.table(data3, "../Plotting_dog_in_wolf//wolves_CHR1.ps21.txt", row.names
=FALSE, col.names = FALSE)

dat <- fread(paste("wolves_CHR1.ps21.txt",sep=''), header=F,data.table=F)

##Load SNP positions

positions <- fread(paste("Wolf_dog_CHR1.snpinfo.txt",sep=''),
header=T,data.table=F) ## positions of SNPs

Wolf <- dat[seq(2,dim(dat)[2],by=2)]## corresponds to the ancestry of wolves,
read only one column for each individuals,we should check it before doing
further analysis

Dog <- dat[seq(1,dim(dat)[2],by=2)]## corresponds to the ancestry of dog, read
only one column for each individuals

#calculating average ancestry
```

```
avg_wolf <- apply(Wolf,2,mean,na.rm=T)
avg_dog <- apply(Dog,2,mean,na.rm=T)

write.table(avg_dog, "../Plotting_dog_in_wolf//avg_dog_chr1.txt", row.names
=FALSE, col.names = FALSE)

Mean <- mean(avg_dog, na.rm = TRUE)## average in all SNPs
#0.065
SD <- sd(avg_dog, na.rm = TRUE)*3
#0.0558
Threshould_positive <- Mean+SD
#0.121
Threshould_negative <- Mean-SD
#0.01


x= avg_dog
new_points<- keep(avg_dog, x> 0.121)##### 0.121 is 3SD+mean for the chromosome
number 1
#129 SNPs

positions2 <- fread(paste("SNP_position_chr1_modify.txt",sep=''),
header=F,data.table=F)##positions of SNPs with more than 3SD values

)##positions of SNPs with more than 3SD values
newpos2 <- round(positions2$V1/10

###Deserts_ancestry
x= avg_dog

)##### 0.001 is threshold for desert ancestry
desert_points<- keep(avg_dog, x< 0.01)
###0 SNPs


windows()
newpos <- round(positions$pos/10)
plot(newpos,avg_dog,type="l",lwd=2,col="lightblue",ylim=c(0,0.20),xlab="SNP
Position (KB)", ylab="Dog ancestry")
points(newpos,avg_dog,pch=16,col="lightblue", cex=0.5)
points(newpos2,new_points,pch=16,col="red", cex=0.5)


abline(h=0.065,lty=1,col="black")## 0.065 is the dig ancestry average in all
chromosomes (without F1 hybrids).
abline(h=0.065,lty=2,col="black")##average in all SNPs in the chromosome
title("CHR 1",adj=0 )#family = "A"

#######CHR2

# we should use .ps21.txt file, the output of ELAI
dat <- fread(paste("Wolf_dog_CHR2.ps21.txt",sep=''), header=F,data.table=F)

##Selecting specific rows (removing First generation hybrids)#####

data_modify <- dat[c(150:153, 155:163, 165:168, 170:249, 251:263, 265:266,
268:314),]
```

```
print(data_modify)

data3 <- data_modify/2

write.table(data3, "../Plotting_dog_in_wolf//wolves_CHR2.ps21.txt", row.names
=FALSE, col.names = FALSE)


dat <- fread(paste("wolves_CHR2.ps21.txt",sep=''), header=F,data.table=F)

#Load SNP positions

positions <- fread(paste("Wolf_dog_CHR2.snpinfo.txt",sep=''),
header=T,data.table=F)

####corresponds to the ancestry of wolf, read only one column for each
individuals,we should check it before doing further analysis
Wolf <- dat[seq(2,dim(dat)[2],by=2)]

Dog <- dat[seq(1,dim(dat)[2],by=2)]## corresponds to the ancestry of dog, read
only one column for each individuals,


avg_wolf <- apply(Wolf,2,mean,na.rm=T) #calculating average ancestry

avg_dog <- apply(Dog,2,mean,na.rm=T)  #calculating average ancestry

write.table(avg_dog, "../Plotting_dog_in_wolf//avg_dog_chr2.txt", row.names
=FALSE, col.names = FALSE)

Mean <- mean(avg_dog, na.rm = TRUE)## average in all SNPs
#0.063
SD <- sd(avg_dog, na.rm = TRUE)*3
#0.0338
Threshould_positive <- Mean+SD
#0.097
Threshould_negative <- Mean-SD
#0.03

x= avg_dog
new_points<- keep(avg_dog, x> 0.097)##### 0.03 is 3SD+mean for the chromosome
number 1
#0 SNPs

###Deserts_ancestry
x= avg_dog
desert_points<- keep(avg_dog, x< 0.03)##### 0.001 is threshold for desert
ancestry
###0 SNPs

positions2 <- fread(paste("SNP_position_SD_chr2.txt",sep=''),
header=F,data.table=F)##positions of SNPs with more than 3SD values
newpos2 <- round(positions2$V1/10)##positions of SNPs with more than 3SD
values


newpos <- round(positions$pos/10)
plot(newpos,avg_dog,type="l",lwd=2,col="lightblue",ylim=c(0,0.15),xlab="SNP
Position (KB)", ylab="Dog ancestry")
```

```
points(newpos,avg_dog,pch=16,col="lightblue", cex=0.5)
points(newpos2,new_points,pch=16,col="red", cex=0.5)


abline(h=0.065,lty=1,col="black")## 0.065 is the dig ancestry average in all
chromosomes (without F1 hybrids).
abline(h=0.063,lty=2,col="black")##average in all SNPs in the chromosome
title("CHR 2",adj=0 )#family = "A"
.
.
.
Do for all 38 chromosomes in all datasets


#############Identification of loci under positive selection#############

=========================================================
         integrated Haplotype Score (iHS)
```

```
Software: rehh package version 3.2.2 (Gautier et al. 2017)

https://cran.r-project.org/package=rehh

#Voight, B. F., Kudaravalli, S., Wen, X., & Pritchard, J. K. (2006). A map of
recent positive selection in the human genome. PLoS biology, 4(3), e72.

#Mathieu Gautier, Renaud Vitalis, rehh: an R package to detect footprints of
selection in genome-wide SNP data from haplotype structure, Bioinformatics,
Volume 28, Issue 8, April 2012, Pages 1176–1177,
-----------------------------------------------------------------------


# Load necessary libraries #

library(rehh)

hh <- data2haplohh(hap_file = "Wolf_hybrid_outlier.vcf.gz",
                   polarize_vcf = FALSE,
                   vcf_reader = "data.table", chr.name = "1")


scan <-scan_hh(
  hh,
  limhaplo = 2,
  limhomohaplo = 2,
  limehh = 0.05,
  limehhs = 0.05,
  phased = TRUE,
  polarized = FALSE,
  scalegap = NA,
  maxgap = NA,
  discard_integration_at_border = TRUE,
  interpolate = TRUE,
  threads = 1
)

ihs_results <- ihh2ihs(
  scan,
  freqbin = 1,
```

```
  min_maf = 0.05,
  min_nhaplo = NA,
  standardize = TRUE,
  include_freq = FALSE,
  right = FALSE,
  alpha = 0.05,
  p.side = NA,
  p.adjust.method = "none",
  verbose = TRUE
)


sum(is.na(ihs_results$ihs))

# Extract the iHS data frame from the list

ihs_df <- ihs_results$ihs

# Convert log p-value to p-value

ihs_df$PVALUE <- 10^(-ihs_df$LOGPVALUE)
write.csv(ihs_df, "ihs_df_chr13.csv")

# Select SNPs with significant selection signals

significant_snps <- subset(ihs_df, abs(IHS) > 2 & PVALUE < 0.05)

# Save results to CSV

write.csv(significant_snps, "significant_SNPs_iHS_chr1.csv", row.names =
FALSE)
.
.
.
######Do for all outlier loci that were found based on the previous
analysis######

-----------------------------------------------------------------
```

Codes For Chapter 4

```
================================================
                randomForest
================================================
#Software: randomForest version: 4.7-1.2
#https://www.stat.berkeley.edu/~breiman/RandomForests/
#Liaw A, Wiener M (2002). "Classification and Regression by randomForest." R
News, 2(3), 18-22. https://CRAN.R-project.org/doc/Rnews/.
```

```
# Load necessary libraries #

require(sp)
require(raster)
require(randomForest)
```

```
require(rfUtilities)
require(mapproj)
require(dismo)
require(maps)
require(proj4)
require(geoR)
require(spatialEco)
library(ggplot2)
library(dplyr)
library(caret)


################## Remove correlated Layers ##################

Oregon <- read.csv("All_Envi_layers_RF.csv", header = T)
xdata.Oregon <- Oregon[,c(1:32)]

cl.Oregon <- multi.collinear(xdata.Oregon, p=0.05)
for(l in cl.Oregon) {
  cl.test <- xdata.Oregon[,-which(names(xdata.Oregon)==l)]
  print(paste("REMOVE VARIABLE", l, sep=": "))
  multi.collinear(cl.test, p=0.05)
}

Oregon.trim <- Oregon[,-which(names(Oregon) %in% cl.Oregon)]

###### create The RandomForest model#######################################

df <- read.csv("wolf_dog_uncorrelated.csv")

# Split data into training and testing sets

set.seed(123)
trainIndex <- createDataPartition(df$Pop2_dog, p = 0.8, list = FALSE)
trainData <- df[trainIndex,]
testData <- df[-trainIndex,]

# Create the Random Forest model

set.seed(123)
rf_model <- randomForest(Pop2_dog ~ ., data = trainData, importance = TRUE)

# Print the variation explained (pseudo R-squared)

print(paste("Variation Explained:", rf_model$rsq[length(rf_model$rsq)] * 100,
"%"))


# View the default feature importance (Mean Decrease in Accuracy and Gini)

print(importance(rf_model))
windows()
varImpPlot(rf_model)

# Use the varImp function from caret to compute permutation-based importance

set.seed(123)
importance_mir <- varImp(rf_model, scale = FALSE)
```

```
# Display the permutation importance (MIR scale)

print(importance_mir)
write.csv(importance_mir, "imortant_layers_90_intensity.csv")

rmse <- function(observed, predicted) {
  sqrt(mean((observed - predicted)^2))
}

# Define the tuning grid for 'mtry' (number of features to sample)

tune_grid <- expand.grid(mtry = 1:21)  # Adjust mtry values based on your
dataset

# Set up trainControl for cross-validation

train_control <- trainControl(method = "cv", number = 5, search = "grid")

# Train the Random Forest model with hyperparameter tuning
set.seed(123)
rf_tuned <- train(
  Pop2_dog ~ .,
  data = trainData,
  method = "rf",
  trControl = train_control,
  tuneGrid = tune_grid,
  ntree = 500  # Number of trees (adjust based on your needs)
)

# Print the best tuned model based on accuracy

print(rf_tuned)

# Predict on test data and calculate RMSE
predictions <- predict(rf_tuned, newdata = testData)
rmse_value <- rmse(testData$Pop2_dog, predictions)

# Print the RMSE
print(paste("RMSE:", rmse_value))


# Set up RFE control
rfe_control <- rfeControl(functions = rfFuncs, method = "cv", number = 5)

# Perform RFE for feature selection

set.seed(123)
rfe_result <- rfe(
  trainData[, -ncol(trainData)],
  trainData$Pop2_dog,
  sizes = c(1:4),  # Specify subset sizes (e.g., number of features)
  rfeControl = rfe_control
)

# Print RFE result
print(rfe_result)

# Plot the RFE results to see how the RMSE changes with the number of features
plot(rfe_result, type = c("g", "o"))
```

```
# Get the optimal subset of features based on RFE
best_features <- predictors(rfe_result)
print(best_features)
write.csv(best_features, "best_feartures2.csv")

# Train final model using the selected features from RFE
set.seed(123)
final_rf_model <- randomForest(
  Pop2_dog ~ .,
  data = trainData[, c(best_features, "Pop2_dog")],
  ntree = 500,
  mtry = rf_tuned$bestTune$mtry  # Use the best 'mtry' from the previous
tuning
)

# Evaluate the final model on the test data
final_predictions <- predict(final_rf_model, newdata = testData[,
c(best_features, "Pop2_dog")])
final_rmse <- rmse(testData$Pop2_dog, final_predictions)

# Print final RMSE
print(paste("Final RMSE:", final_rmse))


# Print the variation explained (pseudo R-squared)
print(paste("Variation Explained:",
final_rf_model$rsq[length(final_rf_model$rsq)] * 100, "%"))

# Plot variable importance to see which factors contribute most to the model
windows()
importance(final_rf_model)
varImpPlot(final_rf_model)


# Predicted values
predicted <- predict(final_rf_model, df)

# Confusion matrix
conf_matrix <- table(Actual = df$Pop2_dog, Predicted = predicted)
print(conf_matrix)

#calculate permutation importance
set.seed(123)
importance_mir <- varImp(final_rf_model, scale = FALSE)
print(importance_mir)
write.csv(importance_mir, "important_permutation2.csv")
# Evaluate if the Random Forest model is overfitting
overfit_test <- rf.regression.fit(final_rf_model)

# Print the result of the overfit evaluation
print(overfit_test)


# Multiple linear regression (multiple predictors)
model <- lm(Pop2_dog ~ bio_13  + bio_12 + hf + bio_6 + bio_4, data = df)
summary(model)
```

```
------------------------------------------------------------------------

============================================================
            Redundancy Analysis (RDA)
============================================================
 Package: vegan, version: 2.6-10

# https://github.com/vegandevs/vegan/issues

#Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R.,
O'hara, R. B., ... & Oksanen, M. J. (2013). Package 'vegan'. Community ecology
package, version, 2(9), 1-295.
```

```
# Load necessary libraries #

library(vegan)
library(data.table)
library(ggplot2)
library(vcfR)
library(dplyr)

####### Conver VCF file (outlier SNPs) to CSV file######

# Input VCF file path
vcf_file <- "Pure_dogs_linear_LD.vcf"

# Output CSV file path
output_csv <- "Pure_dogs_linear_LD.csv"

vcf_to_csv(vcf_file, output_csv)

########################### Transposed data #################

# Load the CSV file without treating the first row as headers

input_file <- "Hybrid_wolf_dog_linear_LD.csv"  # Replace with your file path

data <- read.csv(input_file, header = FALSE, stringsAsFactors = FALSE)

# Transpose the entire dataset
transposed_data <- t(data)

# Convert the transposed matrix back to a data frame
transposed_df <- as.data.frame(transposed_data)

# Save the transposed data to a new CSV file

output_file <- "Transposed_Hybrid_wolf_dog_linear_LD.csv"
write.table(transposed_df, output_file, sep = ",", row.names = FALSE,
col.names = FALSE)

print(paste("Transposed file saved as:", output_file))

------------------------------------------------------------------------

#####Prepare input data#########
```

```
genotype_matrix <-
read.csv("Transpose_genotype_hybrid_dogs_oulier_wolves.csv", row.names = 1,
header = T)
  env_matrix <- read.csv("Env_hybrid_dog.csv", row.names = 1, header = T)


# Perform Redundancy Analysis (RDA)#####
rda_result <- rda(genotype_matrix ~ ., data = env_matrix)

# View the RDA result summary
summary(rda_result)

# Test the significance of the RDA model
anova_result <- anova(rda_result, permutations = 999)
print(anova_result)


# Calculate and print Adjusted R²
adjusted_r2 <- RsquareAdj(rda_result)$adj.r.squared
print(paste("Adjusted R²:", round(adjusted_r2, 4)))


# Perform an overall significance test for the RDA model
overall_pval <- anova.cca(rda_result, permutations = 999)

# Display the result
print(overall_pval)

#Determine Significant Axes
anova_result <- anova.cca(rda_result, by = "axis", permutations = 999)
print(anova_result)

# Extract significant axes
significant_axes <- which(anova_result[["Pr(>F)"]] < 0.05)
print(significant_axes)

# Extract species (loci) loadings for significant RDA axes
snp_loadings <- scores(rda_result, choices = significant_axes, display =
"species")
snp_loadings

###### Prepare data for ggplot
individual_scores <- scores(rda_result, display = "sites", scaling = 2)  #
Scores for individuals
species_scores <- scores(rda_result, display = "species", scaling = 2)  #
Scores for SNPs or variables
environmental_vectors <- scores(rda_result, display = "bp", scaling = 2)  #
Scores for environmental variables

# Convert individual scores to a data frame
individual_df <- data.frame(individual_scores)
individual_df$Cluster <- clusters_data$Cluster  # Add cluster information from
your data

# Convert species scores to a data frame (optional, for SNPs/variables)
species_df <- data.frame(species_scores)
species_df$Variable <- rownames(species_df)

# Convert environmental vectors to a data frame
```

```
env_df <- data.frame(environmental_vectors)
env_df$Variable <- rownames(env_df)

########## Enhanced RDA plot with ggplot2 ########

ggplot() +
  # Plot individual scores with clusters
  geom_point(data = individual_df, aes(x = RDA1, y = RDA2, color = Cluster),
size = 3, alpha = 0.7) +
  labs(title = "Enhanced RDA Triplot - Scaling 2", x = "RDA1", y = "RDA2",
color = "Cluster") +

  # Add environmental variable vectors
  geom_segment(data = env_df, aes(x = 0, y = 0, xend = RDA1, yend = RDA2),
               arrow = arrow(length = unit(0.2, "cm")), color = "blue") +
  geom_text(data = env_df, aes(x = RDA1, y = RDA2, label = Variable), color =
"blue", vjust = -0.5, size = 4) +

  # Add species (optional, if applicable)
  geom_text(data = species_df, aes(x = RDA1, y = RDA2, label = Variable),
color = "red", alpha = 0.6, size = 3) +

  # Improve theme and appearance
  theme_minimal() +
  theme(legend.position = "right") +
  scale_color_brewer(palette = "Set1")
```